

UNIVERSIDADE FEDERAL DO PARANÁ

LETÍCIA GONÇALVES POTAMIANOS

**ANÁLISE DE POLIMORFISMOS EM RNAS LONGOS NÃO-CODIFICANTES  
ASSOCIADOS À LEUCEMIA LINFOIDE AGUDA: ESTUDO CASO-CONTROLE**

CURITIBA

2024

LETÍCIA GONÇALVES POTAMIANOS

ANÁLISE DE POLIMORFISMOS EM RNAs LONGOS NÃO-CODIFICANTES  
ASSOCIADOS À LEUCEMIA LINFOIDE AGUDA: ESTUDO CASO-CONTROLE

Trabalho de conclusão de curso apresentado ao curso de Biomedicina, Setor de Ciências Biológicas, Universidade Federal do Paraná, como requisito parcial à obtenção do título de bacharel.

Orientadora: Prof<sup>a</sup>. Dr<sup>a</sup>. Jaqueline Carvalho de Oliveira

CURITIBA

2024

## **AGRADECIMENTOS**

À grandiosa Universidade Federal do Paraná, um antigo sonho que se tornou realidade.

À minha orientadora, pela oportunidade de fazer parte do Laboratório de Citogenética Humana e Oncogenética, pela gentileza, paciência e por todos os ensinamentos transmitidos ao longo dos anos.

A meus pais, que, sempre de braços abertos, me acolheram, apoiaram e instruíram, prestando todo o apoio necessário em dias difíceis e comemorando cada pequena vitória.

A Deus, por ter me permitido chegar até aqui.

À minha família, por todos os conselhos e suporte fornecido.

Ao meu namorado, pela parceria, carinho e amparo.

Aos meus amigos de Biomedicina, que tornaram a graduação mais leve, divertida e enriquecedora.

Aos meus colegas de laboratório, pela recepção e lições que levarei para toda trajetória acadêmica.

Aos mestres, por terem compartilhado seus conhecimentos e terem contribuído para minha formação.

A todas as crianças e jovens envolvidos neste estudo.

## RESUMO

A leucemia linfóide aguda (LLA) é uma neoplasia altamente heterogênea e pode ser classificada em muitos subtipos, importantes para a escolha do melhor protocolo de tratamento. É a leucemia mais comum na população pediátrica, correspondendo a até 80% dos casos nesse grupo versus 20% dos casos em adultos. Posto isto, é evidente a relevância de estudos que busquem uma melhor compreensão da base genética desta doença, visando um melhor manejo clínico dos pacientes. Um número crescente de RNAs longos não-codificantes (lncRNAs) tem sido descrito e estudos apontam que variantes podem funcionar como biomarcadores de predisposição, porém ainda são pouco estudadas na LLA. Sendo assim, o objetivo deste trabalho foi avaliar a frequência alélica de um polimorfismo (SNP) no lncRNA KCNMB2-AS1 em pacientes brasileiros com LLA, correlacionando com o quadro clínico. Os lncRNAs associados à LLA foram selecionados através da análise de dados de associação genética, obtidos na plataforma GWAS *Catalog*, que indicaram 12 variantes em 10 lncRNAs. O rs9290663 no KCNMB2-AS1 (A > T) foi selecionado para o desenho de *primers* e otimização da PCR alelo-específica para um estudo caso-controle. As 75 amostras de sangue periférico de pacientes com LLA pediátrica foram obtidas no Hospital Oncopediátrico Erastinho, enquanto as 90 amostras controle são provenientes de indivíduos saudáveis sem histórico familiar de câncer, armazenadas no Laboratório de Citogenética Humana e Oncogenética (LabCHO) da Universidade Federal do Paraná. As frequências genótípicas observadas em pacientes com LLA pediátrica foram 64% homocigotos AA; 32% heterocigotos AT e 4% homocigotos TT. As frequências genótípicas observadas nos indivíduos controle foram 67,8% homocigotos AA; 32,2 % heterocigotos AT e nenhum homocigoto para o alelo T. A frequência alélica do SNP rs9290663 calculada no grupo de pacientes foi de A: 80% e T: 20%, e no grupo de controles A: 84% e T: 16%, bem similares. As análises estatísticas de *Odds Ratio* (OR) não mostraram associação significativa para o SNP e a suscetibilidade à LLA pediátrica quando se analisaram todos os pacientes (OR: 0,84, intervalo: 0,44 a 1,63,  $p=0,61$ , AA versus AT/TT), ou apenas os pacientes com o fenótipo LLA-B (OR: 0,90, intervalo: 0,46 a 1,82,  $p=0,78$ ). Mesmo não obtendo o resultado esperado, esses estudos de associação e levantamento das frequências de SNPs em parte da população brasileira são de grande importância, visto que tal representatividade não está incluída nos estudos atuais. Pretendemos aumentar o número de amostras analisadas para confirmar os resultados preliminares obtidos no presente trabalho.

**Palavras-chave:** LLA; lncRNA; SNP; PCR alelo-específica; caso-controle.

## ABSTRACT

Acute lymphoblastic leukemia (ALL) is a highly heterogeneous cancer and can be classified into many subtypes, which are important for choosing the best treatment protocol. It is the most common leukemia in the pediatric population, accounting for up to 80% of cases in this group versus 20% of cases in adults. Having said this, the relevance of studies that seek a better understanding of the genetic basis of this disease, aiming at better clinical management of patients, is evident. An increasing number of long non-coding RNAs (lncRNAs) have been described and studies indicate that variants can function as predisposition biomarkers, but they aren't well studied in ALL. Therefore, the aim of this work was to evaluate the allelic frequency of a polymorphism (SNP) in the lncRNA KCNMB2-AS1 in Brazilian patients with ALL, correlating it with the clinical picture. The lncRNAs associated with ALL were selected through the analysis of genetic association data, obtained from the GWAS Catalog platform, which indicated 12 variants in 10 lncRNAs. rs9290663 in KCNMB2-AS1 (A > T) was selected for primer design and optimization of an allele-specific PCR for a case-control study. The 75 peripheral blood samples from patients with pediatric ALL were obtained at the Hospital Erastinho, while the 90 control samples came from healthy individuals with no family history of cancer, stored at the Human Cytogenetics and Oncogenetics Laboratory (LabCHO) at the Federal University of Paraná. Genotype frequencies observed in pediatric ALL patients were 64% AA homozygotes; 32% AT heterozygotes and 4% TT homozygotes. The genotype frequencies observed in control individuals were 67.8% homozygotes AA; 32.2% AT heterozygotes and no homozygotes for the T allele. The allelic frequency of SNP rs9290663 calculated in the patient group was A: 80% and T: 20%, and in the control group A: 84% and T: 16%, very similar. Statistical analyzes of Odds Ratio (OR) showed no significant association for the SNP and susceptibility to pediatric ALL when all patients were analyzed (OR: 0.84, range: 0.44 to 1.63,  $p=0.61$ , AA versus AT/TT), or only patients with the B-ALL phenotype (OR: 0.90, range: 0.46 to 1.82,  $p=0.78$ ). These weren't the expected results, but these association studies and survey of SNP frequencies in the Brazilian population are of great importance, because such representation is not included in current studies. We intend to increase the number of samples analyzed to confirm the preliminary results obtained in the present work.

Key-words: ALL; lncRNAs; SNP; allele-specific PCR; case-control.

## LISTA DE ILUSTRAÇÕES

FIGURA 1 – LOCALIZAÇÃO DO lncRNA KCNMB2-AS1.....	22
FIGURA 2 – SEQUÊNCIA AMPLIFICADA E <i>PRIMERS</i> UTILIZADOS.....	22
FIGURA 3 – PUREZA DO DNA EXTRAÍDO DE AMOSTRAS DE SANGUE DE PACIENTES COM LLA.....	23
FIGURA 4 – GEL DE ELETROFORESE DA PCR DA REGIÃO FLANQUEADORA DO SNP rs9290663.....	24
FIGURA 5 – CROMATOGRAMA DO <i>AMPLICON</i> DA REGIÃO FLANQUEADORA DO SNP rs9290663.....	24
FIGURA 6 – GEL DE AGAROSE PARA AVALIAÇÃO DA TEMPERATURA DE ANELAMENTO.....	25
FIGURA 7 – PADRÃO ELETROFORÉTICO DOS FRAGMENTOS AMPLIFICADOS.....	26
FIGURA 8 – GEL DE ELETROFORESE DA GENOTIPAGEM DO SNP rs9290663 DAS AMOSTRAS LLA24 – LLA32.....	27
FIGURA 9 – GEL DE ELETROFORESE DA GENOTIPAGEM DO SNP rs9290663 DAS AMOSTRAS LLA63 – LLA80.....	27

FIGURA 10 – GEL DE ELETROFORESE DA GENOTIPAGEM DO SNP rs9290663  
DAS AMOSTRAS LLA81 – LLA94.....28

FIGURA 11 – GEL DE ELETROFORESE DA GENOTIPAGEM DO SNP rs9290663  
DOS CONTROLES C49 – C64.....29

FIGURA 12 – GEL DE ELETROFORESE DA GENOTIPAGEM DO SNP rs9290663  
DOS CONTROLES C65 – C80.....30

## LISTA DE TABELAS

TABELA 1 - PADRÃO DE CICLAGEM DA PCR.....	19
TABELA 2 - lncRNAs E SNPs ASSOCIADOS À LLA.....	21
TABELA 3 - <i>PRIMERS</i> PARA PCR ALELO-ESPECÍFICA (FRAGMENTO DE 154 pb).....	23
TABELA 4 - CONDIÇÕES PARA PCR ALELO-ESPECÍFICA DO SNP rs9290663.....	25
TABELA 5 - FREQUÊNCIAS ALÉLICAS E GENOTÍPICAS EM PACIENTES E CONTROLES.....	31
TABELA 6 - FREQUÊNCIA DO SNP rs9290663 NO 1000 <i>GENOMES PROJECT PHASE 3</i> .....	32
TABELA 7 – ANÁLISES DE <i>ODDS RATIO</i> PARA GENÓTIPOS E ALELOS.....	33

## SUMÁRIO

<b>1.</b>	<b>INTRODUÇÃO.....</b>	<b>10</b>
1.1	CONTEXTO E PROBLEMÁTICA.....	10
1.2	JUSTIFICATIVA.....	11
1.3	OBJETIVOS.....	11
1.3.1	Objetivo geral.....	11
1.3.2	Objetivos específicos.....	11
<b>2.</b>	<b>REVISÃO BIBLIOGRÁFICA.....</b>	<b>12</b>
2.1	LEUCEMIAS.....	12
2.2	RNAs LONGOS NÃO-CODIFICANTES.....	13
2.2.1	RNAs longos não-codificantes no câncer.....	14
2.2.2	RNA longo não-codificante KCNMB2-AS1.....	15
<b>3.</b>	<b>METODOLOGIA.....</b>	<b>16</b>
3.1	ANÁLISES <i>IN SILICO</i> .....	16
3.1.1	Extração e filtragem dos dados.....	16
3.1.2	Seleção dos SNPs.....	16
3.1.3	Desenho dos <i>primers</i> .....	17
3.1.4	Casuística.....	17
3.2	ANÁLISES <i>IN VITRO</i> .....	18
3.2.1	Extração e quantificação de DNA.....	18
3.2.2	Otimização dos <i>primers</i> .....	19
3.2.3	Sequenciamento de Sanger.....	19
3.2.4	PCR para genotipagem.....	20
3.3	ANÁLISE ESTATÍSTICA.....	20
<b>4.</b>	<b>RESULTADOS.....</b>	<b>21</b>
4.1	SNPs CANDIDATOS PARA GENOTIPAGEM.....	21
4.2	<i>PRIMERS</i> PARA A REGIÃO DE rs9290663.....	22
4.3	QUANTIFICAÇÃO DAS AMOSTRAS DE SANGUE PERIFÉRICO.....	23
4.4	SEQUENCIAMENTO DE SANGER.....	23
4.5	PADRONIZAÇÃO DA PCR ALELO ESPECÍFICA.....	25
4.6	GENOTIPAGEM DAS AMOSTRAS DE PACIENTES.....	26
4.7	GENOTIPAGEM DAS AMOSTRAS CONTROLE.....	28

4.8	FREQUÊNCIAS ALÉLICAS E GENOTÍPICAS.....	30
4.9	ESTUDOS DE ASSOCIAÇÃO.....	32
5.	<b>DISCUSSÃO.....</b>	<b>34</b>
6.	<b>CONSIDERAÇÕES FINAIS.....</b>	<b>39</b>
	<b>REFERÊNCIAS.....</b>	<b>40</b>
	<b>APÊNDICE A - DADOS DOS PACIENTES COM LEUCEMIA.....</b>	<b>47</b>
	<b>APÊNDICE B – GENOTIPAGEM DAS AMOSTRAS DE SANGUE DOS PACIENTES.....</b>	<b>49</b>
	<b>APÊNDICE C - GENOTIPAGEM DAS AMOSTRAS CONTROLE.....</b>	<b>51</b>

# 1 INTRODUÇÃO

## 1.1 CONTEXTO E PROBLEMÁTICA

A leucemia linfóide aguda (LLA) é a forma mais comum de câncer infantil. Cerca de 3 mil crianças e adolescentes com menos de 20 anos são diagnosticados com LLA todo ano nos Estados Unidos. Já no Brasil, estima-se a ocorrência de aproximadamente 430 mil casos novos de câncer entre crianças e adolescentes (de 0 a 19 anos). Os principais tipos de câncer que ocorrem em crianças são leucemias, tumores do sistema nervoso central e linfomas (Instituto Nacional de Câncer, 2024).

Interações complexas entre fatores genéticos e ambientais estão na base das etiologias de leucemias pediátricas. Diante disso, é de extrema importância o aperfeiçoamento de diagnósticos diferenciais LLA, por possibilitar uma intervenção precoce no tratamento e uma estratégia terapêutica personalizada para cada paciente, levando em consideração suas alterações genéticas ao fazer uma estratificação de risco (Santos *et al.*, 2021). Os sistemas de estratificação de risco para LLA utilizam fatores clínicos, juntamente com a incorporação de aberrações citogenéticas de células leucêmicas no momento do diagnóstico. O notável progresso no tratamento da LLA tem se baseado justamente no ajuste da terapia para esses diferentes subgrupos da doença (Chadha *et al.*, 2023).

O estudo com RNAs para almejar uma terapia assertiva para diferentes tipos de câncer já é realidade. Dentre eles, alguns RNAs não-codificantes (ncRNAs) exercem efeitos anti-tumorais, por inibirem a transcrição de RNAs mensageiros (mRNAs) ou se ligarem a proteínas para bloquear suas funções, por exemplo (Jiang *et al.*, 2019). Apesar da compreensão abrangente do papel dos lncRNAs no câncer ainda estar em estágios iniciais, evidências sugerem que a maioria dos lncRNAs funcionam através de interações específicas com bio-macromoléculas.

Os estudos de associação genômica ampla (GWAS) também revolucionaram o estudo de doenças complexas, como o câncer, permitindo a avaliação de polimorfismos de nucleotídeo único (SNPs) em lncRNAs e a determinação de sua associação com um fenótipo de doença. Dessa maneira, a avaliação dos lncRNAs e seus polimorfismos também é promissora para a descoberta de novas associações relacionadas ao desenvolvimento da LLA (Inaba; Mullighan, 2020).

## 1.2 JUSTIFICATIVA

O diagnóstico precoce é um determinante crítico para o sucesso do tratamento da doença e pode aumentar a taxa de sobrevivência dos pacientes. Os fatores que limitam as atuais abordagens de rastreamento da leucemia incluem baixa sensibilidade e especificidade e custos elevados. Sendo assim, uma abordagem baseada em biomarcadores novos e inovadores do sangue periférico, e não medula óssea, oferece uma alternativa confortável e atraente aos pacientes, levando potencialmente a uma maior taxa de participação.

Os SNPs mapeados em lncRNAs, nesse sentido, podem ser bons candidatos como biomarcadores de impacto clínico. Segundo Inaba e Mullighan (2020), apesar do risco relativo ser tipicamente baixo para uma variante, esses SNPs, cumulativamente, podem resultar em um aumento de até 10 vezes no risco para LLA. O presente trabalho busca uma melhor compreensão da base genética desta neoplasia na população sul-brasileira, para contornar a falta de representatividade de diferentes populações em estudos genômicos humanos, que resulta em informações incompletas ou equivocadas (Sirugo; Williams; Tishkoff, 2019). Para esse fim, a genotipagem de amostras por PCR alelo-específica é uma ferramenta útil e de baixo custo.

## 1.3 OBJETIVOS

### 1.3.1 OBJETIVO GERAL

Avaliar a frequência alélica, em uma coorte brasileira, de um SNP mapeado em RNA longo não-codificante e associado à LLA.

### 1.3.2 OBJETIVOS ESPECÍFICOS

- Identificar na literatura e nos dados depositados no catálogo GWAS os principais SNPs mapeados em lncRNAs associados à LLA pediátrica;
- Realizar a genotipagem das amostras de pacientes e controles por PCR alelo-específica;
- Verificar se há associação entre o polimorfismo rs9290663, mapeado no lncRNA KCNMB2-AS1, e a suscetibilidade à LLA pediátrica em pacientes brasileiros;
- Comparar as frequências alélicas e genotípicas com outros bancos de dados.

## 2 REVISÃO BIBLIOGRÁFICA

### 2.1 LEUCEMIAS

Na leucemia, uma célula sanguínea que ainda não atingiu a maturidade sofre uma mutação genética que a transforma em uma célula cancerosa, a qual não funciona de forma adequada, multiplica-se mais rápido e possui mecanismos anti-apoptóticos (Instituto Nacional de Câncer (INCA), 2024). As leucemias são divididas em mais de doze tipos, sendo as quatro primárias a leucemia mieloide aguda (LMA), a leucemia mieloide crônica (LMC), a leucemia linfóide aguda (LLA) e a leucemia linfóide crônica (LLC). A classificação se baseia no crescimento, que pode ser rápido (leucemia aguda) ou lento (leucemia crônica), e a partir do tipo de célula acometida: células linfóides e células mielóides.

A leucemia linfóide aguda (LLA) é uma doença maligna dos linfoblastos B ou T caracterizada pela proliferação descontrolada de linfócitos imaturos anormais e seus progenitores, sendo uma doença heterogênea com variações moleculares significativas (Bhojwani; Yang; Pui, 2015). Esta neoplasia é desencadeada por uma variedade de alterações genéticas, incluindo translocações cromossômicas, mutações e aneuploidias que envolvem genes responsáveis pela regulação do ciclo celular e pelo desenvolvimento de células linfóides, sendo que os subtipos B e T de LLA mostram padrões distintos de alterações genéticas (Coccaro *et al.*, 2019).

Historicamente, acreditava-se que a maioria dos cânceres infantis, incluindo a LLA, não tinham base hereditária. No entanto, estudos genômicos recentes revelaram que pelo menos 4% das crianças com este tipo de leucemia desenvolvem a doença devido a uma predisposição genética (Bloom *et al.*, 2020). Algumas síndromes genéticas (Síndrome de Down, Anemia de Fanconi, Li-Fraumeni), infecções por EBV e HTLV-1 também são fatores de risco para a LLA. Entre os agentes cancerígenos com evidência suficiente de aumento de risco às leucemias, destaca-se a exposição ao benzeno, ao formaldeído e às radiações X e gama (Associação Americana do Câncer (ACS), 2024).

Nas últimas três décadas, vários estudos citogenéticos convencionais forneceram informações sobre a patogênese da LLA. As translocações comuns em crianças com LLA-B incluem t(12;21) [ETV6–RUNX1], t(1;19) [TCF3–PBX1], e

translocações envolvendo o gene KMT2A com vários genes de fusão parceiros (5%), interferindo em genes que regulam o desenvolvimento linfóide normal (Bhojwani; Yang; Pui, 2015). Além disso, os diferentes subtipos de LLA são caracterizados por assinaturas distintas de metilação do DNA, que por sua vez se correlacionam com perfis de expressão gênica, mostrando a importância dos mecanismos reguladores epigenéticos na hematopoiese normal e maligna (Figuerola *et al.*, 2013).

Conseqüentemente, incorporar informações genéticas somáticas nos cuidados clínicos moveram o tratamento de LLA na direção da medicina de precisão, com o uso de agentes direcionados, como inibidores de tirosina quinase, para crianças com LLA cujos blastos expressam a oncoproteína BCR-ABL, que é responsável pelo crescimento e reprodução descontrolada das células leucêmicas.

## 2.2 RNAs LONGOS NÃO-CODIFICANTES

Entre o genoma transcrito, menos de 2% do genoma são genes codificadores de proteínas, enquanto mais de 90% do restante são transcritos em RNAs não-codificantes (ncRNAs), que podem ser subdivididos em pequenos ncRNAs, como MicroRNAs (miRNAs), RNAs que interagem com PIWI (piRNAs) e pequenos RNAs derivados de RNA de transferência (tsRNAs), bem como a classe lncRNA recentemente caracterizada (Lin *et al.*, 2020).

lncRNAs são transcritos de RNA com mais de 500 nucleotídeos, sem potencial de codificação de proteínas ou com potencial limitado e conservação evolutiva limitada, e podem ser intergênicos, intrônicos ou antissenso (Mattick *et al.*, 2023). Entre os genes expressos no transcriptoma humano, mais de 68% são transcritos em lncRNAs (Lin *et al.*, 2020).

Sua biogênese se assemelha a dos RNAs mensageiros (mRNAs). Eles são transcritos pela polimerase II (Pol II) e podem ser poliadenilados 3' e capeados em 5'. A maioria dos lncRNAs possui diferentes padrões de expressão e preferencialmente localização nuclear, em contraste com os mRNAs, porque estão envolvidos na regulação epigenética da expressão gênica (Aliperti; Skonieczna; Cerase, 2021).

Apesar da maioria dos lncRNAs estar localizada no núcleo, alguns deles também desempenham funções no citoplasma (Chi *et al.*, 2019). Os lncRNAs apresentam papel importante vários processos biológicos, como remodelamento da

cromatina, ativação da transcrição, interferência na transcrição, processamento de RNA e tradução de mRNAs (Zhang *et al.*, 2019).

### 2.2.1 RNAs LONGOS NÃO-CODIFICANTES NO CÂNCER

Estudos de associação genômica amplas (GWAS) identificaram variantes de DNA mapeadas em sequências relacionadas à transcrição de um grande número de lncRNAs associados com os mais variados tipos de câncer. Os lncRNAs exibem um papel importante no câncer, podendo desempenhar papel supressor tumoral ou promover a oncogênese (Bhan; Soleimani; Mandal, 2017).

Devido aos seus padrões de expressão e suas características tecido-específicas, diversos estudos apontam que os lncRNAs podem, de fato, funcionar como biomarcadores prognósticos para diferentes tipos de patologias, a exemplo do câncer. Os lncRNAs também regulam importantes processos celulares, como sinalização proliferativa, vigilância imunológica, imortalidade replicativa, angiogênese, entre outros, funcionando como reguladores-chave das vias envolvendo vários “*Hallmarks*” do câncer (Bach; Lee, 2018).

Como exemplo, o lncRNA HOTTIP, derivado do gene *HOXA*, demonstrou ser altamente expresso em muitos tipos de câncer, atuando como oncogene na leucemia mieloide aguda (LMA) (Luo *et al.*, 2019). No estudo, descobriram que o HOTTIP estava aberrantemente elevado na LMA e funcionava como um regulador epigenético para modular a transcrição de genes associados ao processo hematopoiético.

Os primeiros ncRNAs que demonstraram estar superexpressos no câncer foram o antígeno 3 do câncer de próstata (PCA3), e o transcrito específico da próstata 1 (PCGEM1), sendo que ambos são atualmente usados como biomarcadores para câncer de próstata, altamente específicos (Taniue; Akimitsu, 2021).

Portanto, o desenvolvimento de tecnologias que permitam fácil quantificação de moléculas e monitoramento de sua expressão, por abordagens não invasivas, aumenta muito o potencial de utilização de lncRNAs na prática clínica.

### 2.2.2 RNA LONGO NÃO-CODIFICANTE KCNMB2-AS1

A subunidade  $\beta 2$  codificada pelo gene *KCNMB2* medeia a inativação rápida dos canais de potássio cálcio-dependentes, uma classe de canais de potássio que são ativados por voltagem e pelo aumento de cálcio intracelular (Zarei *et al.*, 2007). Estes canais são reguladores importantes do cálcio e da sinalização elétrica em vários tecidos, desempenhando funções em processos fisiológicos diversos, como a excitabilidade neuronal, a contração suave dos músculos e a secreção hormonal (Vergara *et al.*, 1998). Existem quatro proteínas KCNMB2, e todas elas compartilham uma topologia prototípica de dois domínios transmembrana, com o domínio N-terminal sendo necessário para a oclusão do canal. Funcionalmente, cada uma dessas proteínas altera distintamente as taxas de abertura e fechamento do canal, bem como a sensibilidade do canal ao cálcio (Bentrop *et al.*, 2001).

O lncRNA KCNMB2-AS1, localizado no cromossomo 3 na fita *anti-sense* ao gene *KCNMB2*, vem sendo estudado em diferentes tipos de câncer. No carcinoma pulmonar de células não pequenas, este lncRNA aumenta a expressão de ROCK1, facilitando a progressão tumoral (Yang; Wang, 2022), e no câncer de ovário atua como oncogene (Cao *et al.*, 2023). Zhu *et al.* (2021) também avaliaram o papel do KCNMB2-AS1 na progressão do câncer de bexiga e verificaram que, funcionalmente, seu *knockdown* inibiu dramaticamente a proliferação, migração e invasão das células cancerosas (linhagem T24) *in vitro* e suprimiu o crescimento tumoral *in vivo*.

Além disso, o lncRNA KCNMB2-AS1 desempenha papel considerável no desenvolvimento do câncer de esôfago, através do direcionamento do eixo miR-3194-3p/PYGL (Xu *et al.*, 2021). Sabe-se, também, que MiR-3194-3p suprime a proliferação, migração e invasão celular, bem como promove a apoptose celular no câncer de bexiga, visando a Aquaporina 1 (AQP1), um tipo de proteína integral de membrana com distribuição ampla pelos sistemas fisiológicos, que transporta especificamente água, sem permitir a passagem de solutos para o interior das células (Wei *et al.*, 2020).

Polimorfismos de nucleotídeo único (SNPs) na região do lncRNA KCNMB2-AS1 podem ser evidenciados em bancos de dados públicos e em GWAS, porém não se tem fortes evidências de SNPs específicos nesse lncRNA associados a doenças.

### 3 METODOLOGIA

#### 3.1 ANÁLISES *IN SILICO*

##### 3.1.1 EXTRAÇÃO E FILTRAGEM DOS DADOS

Todas as análises *in silico* deste estudo foram realizadas a partir de dados de outros trabalhos do laboratório (Beal, 2022; Sampaio, 2023). Os dados de associação genética foram extraídos da plataforma GWAS *Catalog* (Buniello *et al.*, 2019). Dado que cada *dataset* (conjunto de dados) da plataforma é identificado com um EFO ID (*Experimental Factor Ontology Identifier*), os termos “ALL”, “Acute Lymphoid Leukemia” e “Acute Lymphoblastic Leukemia” foram pesquisados e somente os resultados dentro da categoria *traits* (fenótipos) foram considerados. Para a continuidade das análises, o conjunto de dados EFO\_0000220 foi analisado (Beal, 2022).

Na etapa de filtragem dos dados, foram excluídas as variantes associadas a LLC e estudos que não apresentaram associação, permanecendo apenas os GWAS relacionados à LLA. Ainda, dentro os dados de LLA, foram excluídas as variantes associadas a fenótipos secundários, como resistência ou hipersensibilidade a quimioterápicos. O *software BioMart* da plataforma *Ensembl* (Cunningham *et al.*, 2022) foi usado para identificar e selecionar apenas os SNPs localizados em lncRNAs, a partir do identificador gênico.

##### 3.1.2 SELEÇÃO DOS SNPs

Os estudos que identificaram associações genéticas de variantes de lncRNAs com LLA foram lidos e analisados quanto à idade da coorte, subtipo de LLA, MAF (*minor frequency allele*) do SNP e valor de OR (*odds ratio*) da associação estatística. Os SNPs com MAF > 0,05, OR > 1,2 e encontrados em estudos de coorte com idade menor que 19 anos (crianças, adolescentes e jovens adultos) foram selecionados para o desenho de *primers*.

Um MAF > 0,05 garante que o SNP seja comum o suficiente para ser relevante na população estudada. 1,2 para OR é um limiar razoável que indica um aumento significativo no risco, sem ser tão alto a ponto de levantar suspeitas de viés ou problemas metodológicos. Para a análise dos SNPs, foram usadas as plataformas

*Ensembl* (Cunningham *et al.*, 2022) para dados populacionais e dbSNP-NCBI (*Database of Single Nucleotide Polymorphisms*) (Sherry; Ward; Sirotkin, 1999) para dados genômicos.

### 3.1.3 DESENHO DOS PRIMERS

Para avaliar o polimorfismo no lncRNA, *primers* alelo-específicos, que amplificam apenas o alelo de interesse por complementariedade em sua extremidade 3', foram desenhados para a realização da genotipagem através da Reação em Cadeia da Polimerase (PCR). Sendo assim, foi feito o desenho de um *primer forward* comum para dois *primers reverse* alelo-específicos, considerando reações separadas.

A sequência flanqueadora do SNP (aproximadamente 1000 pb) foi baixada em formato FASTA na plataforma dbSNP do NCBI (Sherry; Ward; Sirotkin, 1999). Para o desenho dos *primers*, a sequência flanqueadora de cada SNP foi inserida na ferramenta *Primer-Blast* do NCBI (Ye *et al.*, 2012) como PCR *template* e uma sequência de 19-25 pb contendo o alelo específico terminal foi inserida como “*own primer*”. Ainda, o tamanho do produto da PCR foi alterado para 100-300 pb, e para os *primers reverse* alelo-específicos, a sequência foi transformada em sua reversa complementar para inserção na plataforma.

Os *primers* gerados foram analisados quanto à temperatura de *melting* ( $T_m$ ) ( $\Delta T_m < 5,0$ ), autocomplementaridade (*self-complementarity*  $< 2,0$ ), autocomplementaridade 3' (*self-3'-complementarity*  $< 2,0$ ) e conteúdo GC (%GC  $\sim 55\%$ ). Os *primers* que atingiram os valores de cada parâmetro foram analisados na ferramenta OligoAnalyzer (IDT™ *Integrated DNA Technologies*) sob os parâmetros de qPCR. Nessa ferramenta, foi analisada a estrutura tridimensional da molécula para verificar a possibilidade de formação de *hairpins* ( $\Delta G < -6$ ;  $\Delta G > 6$ ), *self-dimers* ( $\sim 0$ ) e *hetero-dimers* ( $\sim 0$ ) entre as moléculas de cada PCR. Os *primers* que atingiram os valores estabelecidos foram sintetizados.

### 3.1.4 CASUÍSTICA

O trabalho foi realizado com amostras de sangue periférico coletadas em tubo de EDTA, provenientes do Hospital Erasto Gaertner (PR), de pacientes na faixa etária de um a 19 anos (*vide* apêndice A), mediante a assinatura do Termo de Consentimento Livre e Esclarecido (TCLE) para maiores de idade e pais de menores de idade e Termo

de Assentimento Livre e Esclarecido (TALE) para pacientes menores de idade. As amostras de leucemia foram coletadas sob aprovação do Comitê de Ética em Pesquisa da Liga Paranaense de Combate ao Câncer - Hospital Erasto Gaertner, sob o número CAAE: 08809419.0.0000.0098.

Quanto às amostras de saliva utilizadas como controle (*vide* apêndice C), estas foram obtidas do banco de amostras do Laboratório de Citogenética Humana e Oncogenética (LabCHO), da Universidade Federal do Paraná, também mediante assinatura de TCLE e resposta a um formulário, sendo selecionadas somente amostras de indivíduos o qual não apresentavam histórico de dois ou mais cânceres na família (CAAE: 43945121.4.0000.0102). As amostras ficam armazenadas a  $-20^{\circ}\text{C}$  no próprio laboratório.

## 3.2 ANÁLISES *IN VITRO*

### 3.2.1 EXTRAÇÃO E QUANTIFICAÇÃO DE DNA

As amostras de sangue periférico foram, primeiramente, centrifugadas a 2500 rpm por cinco minutos, para a separação do plasma e recuperação de 200  $\mu\text{L}$  da camada leucocitária, utilizada para extração de material genético. A extração do DNA das 93 amostras de pacientes (LLA = 75, LMA = 18) foi realizada com o Biopur™ Kit Mini Spin Plus, protocolo 1, segundo recomendação do fabricante.

Na etapa final, um volume de 50  $\mu\text{L}$  do tampão de eluição foi adicionado diretamente sobre a membrana de sílica, formando o volume final da solução. O DNA das amostras controle foi extraído pelo método de beads magnéticas, com a utilização do Extracta Kit da Locus (MVXA P016) no sistema automatizado de extração Extracta® 32.

A medida da concentração de DNA e a identificação da pureza das amostras foram feitas pelo espectrofotômetro NanoDrop® 2000 (Thermo Scientific™). A pureza foi verificada através da razão entre as absorbâncias medidas a 260 nm e 280 nm, cujo valor deve estar entre 2,0 e 1,8 (Viljoen; Booyesen; Tantuan, 2022).

### 3.2.2 OTIMIZAÇÃO DOS *PRIMERS*

A técnica de PCR convencional, com reações separadas em paralelo, foi utilizada para avaliar a especificidade dos *primers* desenhados para o lncRNA e seu polimorfismo selecionado para estudo. Os testes foram realizados seguindo o padrão de ciclagem exposto na tabela 1. A temperatura de anelamento ideal foi aquela que permitiu uma reação alelo-específica.

TABELA 1 - PADRÃO DE CICLAGEM DA PCR

	ETAPA DO CICLO	TEMPERATURA	TEMPO
	Desnaturação inicial	95°C	10 minutos
x40	Desnaturação	96°C	30 segundos
	Anelamento	A definir	30 segundos
	Extensão	72°C	30 segundos
	Extensão final	72°C	10 minutos

FONTE: A autora (2024).

### 3.2.3 SEQUENCIAMENTO DE SANGER

Sampaio (2023), previamente em trabalho do grupo, amplificou a região flanqueadora do SNP candidato (226 pb) a partir dos *primers* desenhados. A PCR foi montada a partir de 2 µL do DNA *template* (20 ng/µL), 0,5 µL da solução contendo os *primers forward e reverse* (10 mM), 17,5 µL de água ultrapura e 10µL de Master Mix para PCR convencional (1x) desenvolvido pelo IBMP – Fiocruz.

Para o Sequenciamento de Sanger, as amostras foram tratadas com exonuclease (exo1) e fosfatase alcalina (SAP). Para cada 9 µL de produto de PCR, foram adicionados 0,33 µL de exo1, 0,33 µL de SAP e 0,34 µL de água ultrapura. Essa solução foi colocada em termociclador a 37°C por 1 hora, e depois a 85°C por 15 minutos. Após, foi feito o protocolo padrão de marcação fluorescente pelo *BigDye™ Terminator*: para cada 4 µL de produto de PCR tratado com exo/SAP, foram adicionados 2,0 µL de tampão *BigDye™*, 2,6 µL de água ultrapura e 0,7 µL de solução de *primer (forward e reverse)* a 10 pmol/µL. Essa solução foi colocada em um termociclador seguindo 40 ciclos de 15 segundos a 96°C para desnaturação, 15

segundos a 64°C para anelamento e 4 minutos a 60°C para extensão (Sampaio, 2023).

A última fase consistiu na purificação do DNA marcado por precipitação, com utilização de EDTA (125 mM) e etanol (absoluto e 70%), a fim de precipitar o produto de PCR marcado. Após a evaporação completa do etanol, as amostras foram analisadas no equipamento 3500 Genetic Analyzer (Thermo Fisher).

### 3.2.4 PCR PARA GENOTIPAGEM

A PCR, tanto para as amostras de pacientes quanto para os controles, foi preparada com 0,3 µL da solução contendo os *primers forward e reverse* (10 mM), 8,2 µL de água ultrapura, 1,5 µL do DNA *template* a 20 ng/µL e 10 µL de Master Mix para PCR convencional (2x), desenvolvido pelo IBMP – Fiocruz.

Os produtos da PCR foram genotipados por meio do padrão eletroforético dos fragmentos amplificados em gel de agarose 1,5% suplementado com GelRed™ Biotium 10.000x, diluído 1:500. O gel de eletroforese foi interpretado a partir da presença ou ausência dos produtos de PCR no gel, verificado no fotodocumentador GD-1000 (Axygen®). Para estimar o tamanho dos fragmentos, foi usado marcador de peso molecular (CellCo) de 1000 – 100 pb.

### 3.3 ANÁLISE ESTATÍSTICA

As frequências alélicas e genotípicas foram calculadas diretamente no grupo de pacientes e no grupo controle, utilizando-se o teste de desvios das proporções do teorema de equilíbrio de Hardy-Weinberg por Qui-quadrado.

Adicionalmente, por ser um teste particularmente indicado para estudos de caso-controle (Wagner; Callegari-Jacques, 1998), foi utilizado o cálculo de *Odds ratio* (OR), e o teste de Qui-Quadrado para avaliar se as variáveis alélicas/genotípicas estão associadas com o fenótipo estudado (LLA pediátrica). Os testes estatísticos foram realizados com o *software GraphPad Prism*. Para todos os testes acima descritos, foram considerados como significativos valores de  $p < 0,05$ .

## 4 RESULTADOS

### 4.1 SNPs CANDIDATOS PARA GENOTIPAGEM

Ao todo, dados de 58 GWAS foram filtrados em busca de SNPs mapeados em lncRNAs e associados à LLA pediátrica. Desses, 445 associações foram filtradas, sendo 257 delas associadas à LLA, das quais apenas 137 eram associadas à LLA pediátrica. 33 SNPs foram encontrados mapeados em 32 lncRNAs diferentes e, após os filtros de *odds-ratio* ( $> 1,2$ ) e MAF ( $> 0,05$ ), restaram 12 SNPs mapeados em 10 lncRNAs diferentes (Tabela 2).

TABELA 2 - lncRNAs E SNPs ASSOCIADOS À LLA

SNP	lncRNA	OR	CI (95%)	MAF	Alleles	GWAS
rs2069426	<i>CDKN2B-AS1</i>	1.35	[1.26 - 1.70]	0.06 (G)	G   T	Wiemels <i>et al.</i> , 2018
rs75777619	<i>CCDC26</i>	1.26	[1.17 - 1.36]	0.1 (A)	A   G	Vijayakrishnan <i>et al.</i> , 2019
rs28665337	<i>CCDC26</i>	1.34	[1.21 - 1.47]	0.1 (C)	C   A   T	Vijayakrishnan <i>et al.</i> , 2019
rs4617118	<i>CCDC26</i>	1.28	[1.19 - 1.37]	0.19 (A)	A   G	Wiemels <i>et al.</i> , 2018
		1.69	[1.40 - 1.90]	0.23 (G)	A   G	Trevino <i>et al.</i> , 2009
rs9290663	<i>KCNMB2-AS1</i>	1.58	[1.20 - 1.90]	<b>0,25 (T)</b>	<b>A   T   G</b>	Trevino <i>et al.</i> , 2009
rs11978267	<i>AC124014.1</i>	1.43	[1.28 - 1.59]	0.23 (G)	A   G	Ellinghaus <i>et al.</i> , 2011
		1.43	[NR]	0.23 (G)	A   G	Wiemels <i>et al.</i> , 2018
rs7156960	<i>AC016526.2</i>	1.2	[1.12 - 1.33]	0.44 (G)	G   C	Ellinghaus <i>et al.</i> , 2011
rs630662	<i>AC025508.1</i>	[NR]	[NR]	0.2 (G)	G   A   T	Ellinghaus <i>et al.</i> , 2011
rs10018622	<i>AC079921.1</i>	1.9	[1.45 - 2.63]	0.49 (G)	G   A   C	Archer <i>et al.</i> , 2017
rs11155133	<i>AL035446.1</i>	3.62	[2.10 - 6.00]	0.13 (A)	A   G	Trevino <i>et al.</i> , 2009
rs7738636	<i>AL355612.1</i>	1.2	[1.15 - 1.41]	0.28 (A)	A   C	Ellinghaus <i>et al.</i> , 2011
rs10170236	<i>MMADHC-DT</i>	1.45	[1.24 - 1.69]	0.33 (A)	A   G	Yang <i>et al.</i> , 2012

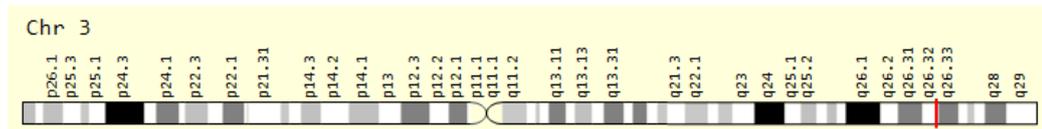
FONTE: Adaptado de Sampaio (2023).

LEGENDA: OR: Odds Ratio (no português, razão de possibilidade); CI (95%): Intervalo de confiança de 95%; MAF: frequência do alelo menor.

Assim, o SNP selecionado para o estudo caso-controle foi o rs9290663 do *KCNMB2-AS1*, que está localizado em 3q26.32 (figura 1). Apesar desse transcrito já ter sido estudado em cânceres como o de esôfago (Xu *et al.*, 2021) e carcinoma pulmonar de células não pequenas (Yang; Wang, 2022), bem como o rs9290663 ter sido associado a um maior risco de LMA em uma coorte chinesa (Cao *et al.*, 2016), e o alelo T como alelo de risco para LLA (Treviño *et al.*, 2009), ainda não foram conduzidos estudos com esse SNP na população brasileira para verificar a frequência

alélica. Desse modo, a avaliação desse lncRNA e seu SNP pode ser importante para uma melhor compreensão da base genética deste tipo de leucemia.

FIGURA 1 – LOCALIZAÇÃO DO lncRNA KCNMB2-AS1



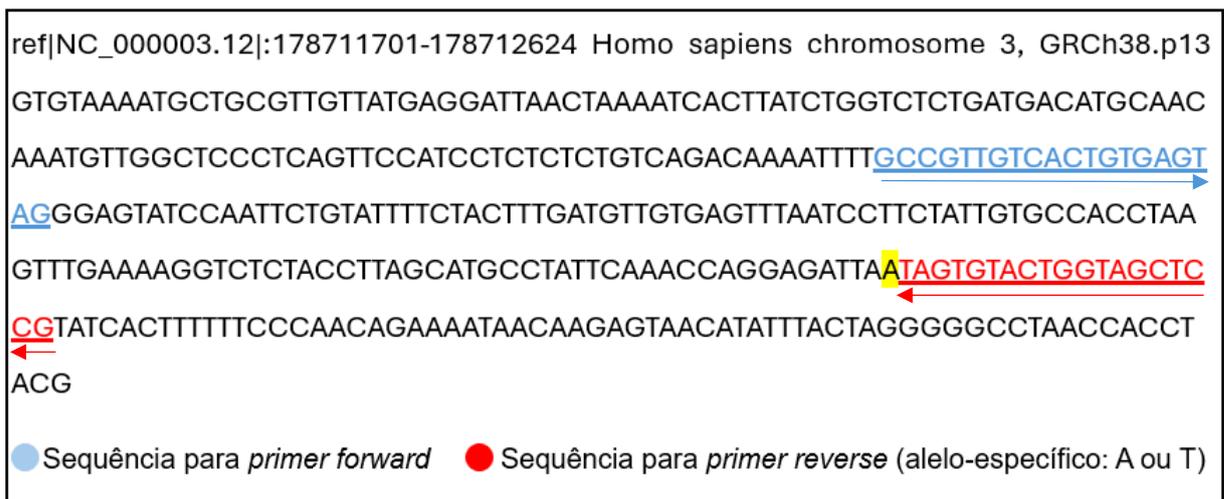
FONTE: GeneCaRNA (2024).

LEGENDA: linha vermelha representa a localização cromossômica de KCNMB2-AS1

#### 4.2 PRIMERS PARA A REGIÃO DE rs9290663

Respeitando os critérios estabelecidos na análise *in silico*, foram encontrados *primers reverse* alelo-específicos adequados para o SNP rs9290663 (Figura 2, Tabela 3). Existem 3 alelos do SNP (A/G/T, Ensembl) e a maioria dos trabalhos mostra um MAF próximo a 0,25 (T).

FIGURA 2 - SEQUÊNCIA AMPLIFICADA E PRIMERS UTILIZADOS



FONTE: A autora (2024)

LEGENDA: em amarelo - SNP rs9290663 (alelos A ou T). A seta representa onde os *primers* se anelam.

TABELA 3 - PRIMERS PARA PCR ALELO-ESPECÍFICA (FRAGMENTO DE 154 pb)

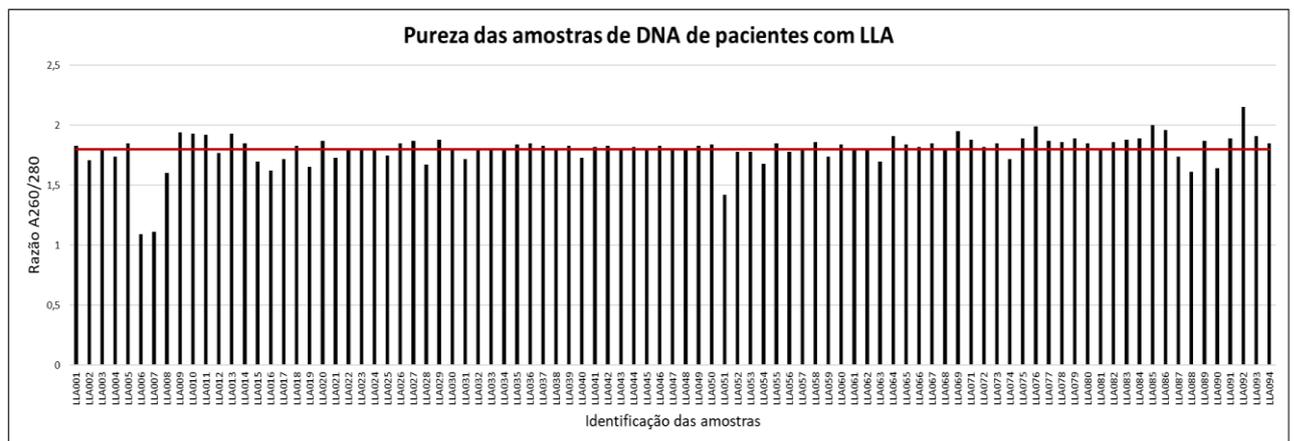
<i>Forward 3' – 5'</i>	<i>Reverse (A) 5' – 3'</i>	<i>Reverse (T) 5' – 3'</i>
GCCGTTGTCAGTGTGAGTAG	CGGAGCTACCACTACTA <sup>A</sup>	CGGAGCTACCACTACTA <sup>T</sup>

FONTE: A autora (2024)

#### 4.3 QUANTIFICAÇÃO DAS AMOSTRAS DE SANGUE PERIFÉRICO

Ao todo foram quantificadas 93 amostras de DNA extraído do sangue periférico de pacientes com leucemia (*vide* apêndice B), sendo 75 delas de LLA. Em relação à razão 260/280, que indica pureza do material genético, e espera-se valores próximos a 1,80 para DNA, verificou-se uma média de 1,7986, com desvio-padrão de 0,14 (Figura 3). Para o experimento, as amostras com concentração maior que 20 ng/μL foram diluídas até esse valor.

FIGURA 3 - PUREZA DO DNA EXTRAÍDO DE AMOSTRAS DE SANGUE DE PACIENTES COM LLA



FONTE: A autora (2024).

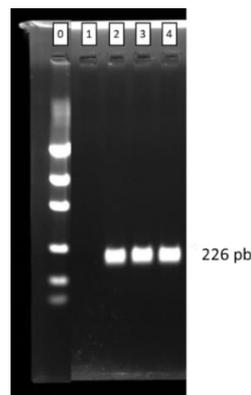
LEGENDA: A linha vermelha indica a média da razão 260/280.

#### 4.4 SEQUENCIAMENTO DE SANGER

Para a confirmação de genotipagem visualizada na PCR alelo-específica, foi realizada a genotipagem de alguns indivíduos por Sequenciamento de Sanger. O gel de eletroforese do *amplicon* da região flanqueadora (226 pb) do SNP rs9290663 mostrou bandas do tamanho esperado em relação ao que foi desenhado *in silico* (Figura 4).

Os cromatogramas do sequenciamento desses *amplicons* revelaram dois picos de fluorescência na posição do SNP para as amostras heterozigotas testadas, confirmando genótipo heterozigoto (AT) (Figura 5). Ao todo, 3 amostras foram sequenciadas para esse teste. As amostras A1 (LLA29) e A3 (LLA31) se mostraram heterozigotas (AT), enquanto a amostra A2 (LLA30) é homozigota para o alelo A.

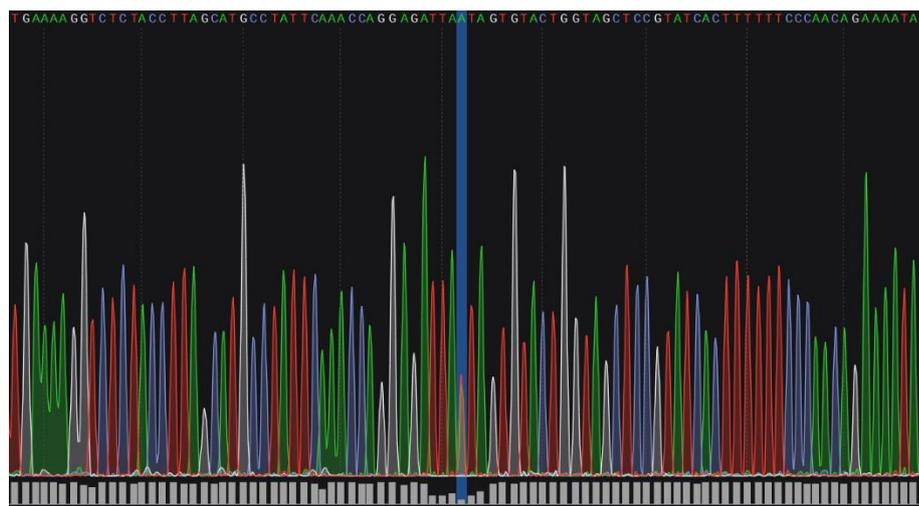
FIGURA 4 - GEL DE ELETROFORESE DA PCR DA REGIÃO FLANQUEADORA DO SNP rs9290663



FONTE: Sampaio (2023).

LEGENDA: 0 – ladder; 1 – branco; 2 – A1 (LLA29); 3 – A2 (LLA30); 4 – A3 (LLA31).

FIGURA 5 - CROMATOGRAMA DO AMPLICON DA REGIÃO FLANQUEADORA DO SNP rs9290663



FONTE: Sampaio (2023).

LEGENDA: faixa azul – alelo A/T do rs9290663 (fita -), pico duplo, heterozigoto.

#### 4.5 PADRONIZAÇÃO DA PCR ALELO-ESPECÍFICA

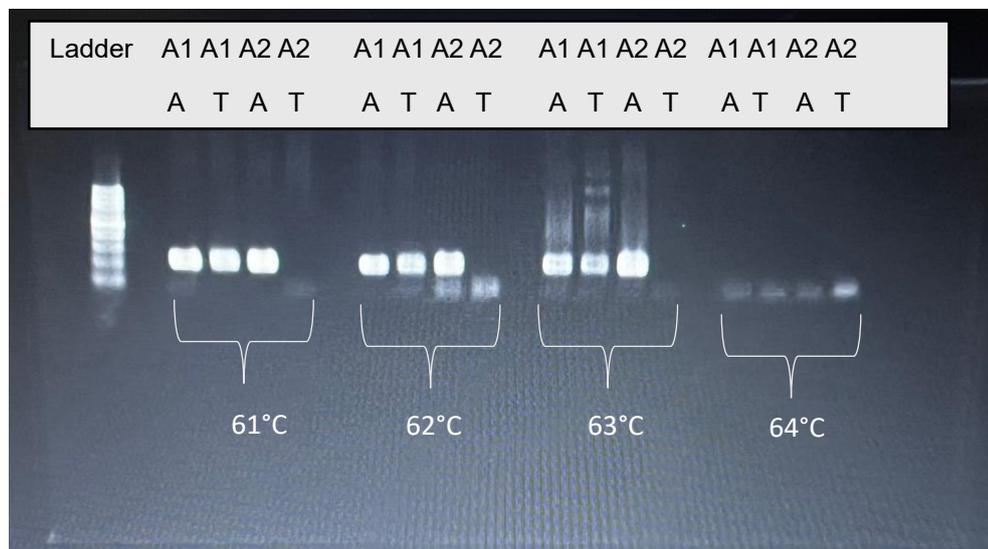
A tabela 4 apresenta as condições encontradas para a padronização da PCR alelo-específica. A temperatura ideal de anelamento foi baseada na última temperatura a qual permitiu a amplificação e diferenciação dos genótipos, como mostrado no gel de agarose da Figura 6. Assim, na temperatura de 63°C foi alcançado a diferenciação dos genótipos para o lncRNA KCNMB2-AS1 e seu SNP (alelo A/T).

TABELA 4 – CONDIÇÕES PARA PCR ALELO-ESPECÍFICA DO SNP rs9290663

	ETAPA DO CICLO	TEMPERATURA	TEMPO
	Desnaturação inicial	95°C	10 minutos
x40	Desnaturação	96°C	30 segundos
	Anelamento	63°C	30 segundos
	Extensão	72°C	30 segundos
	Extensão final	72°C	10 minutos

FONTE: A autora (2024).

FIGURA 6 – GEL DE AGAROSE PARA AVALIAÇÃO DA TEMPERATURA DE ANELAMENTO



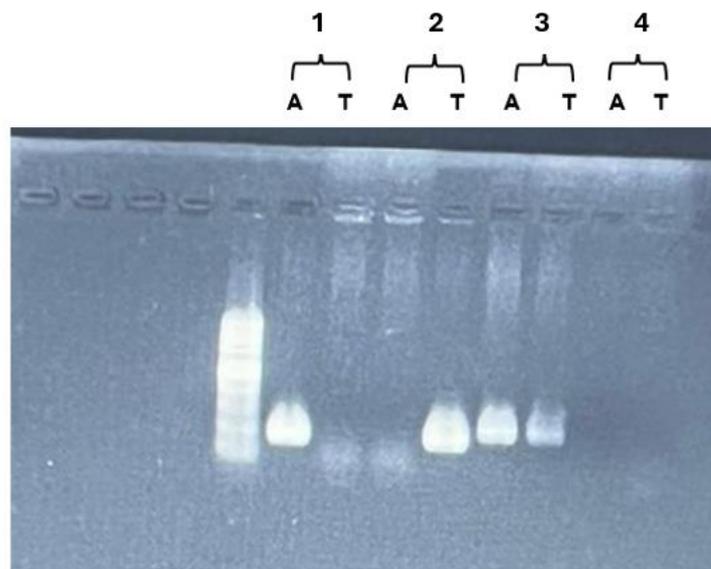
FONTE: A autora (2024).

LEGENDA: Ladder: Marcador de peso molecular; A1 - Amostra heterozigota e A2 - Amostra homozigota AA nas temperaturas de anelamento de 61°C, 62°C, 63°C e 64°C.

#### 4.6 GENOTIPAGEM DAS AMOSTRAS DE PACIENTES

A figura 7 mostra o padrão eletroforético esperado para as amostras de pacientes com leucemia e amostras controle. As figuras 8, 9 e 10 mostram os géis de agarose referentes à genotipagem do SNP rs9290663 nas amostras de DNA extraídas de sangue de pacientes com leucemia. É possível perceber as bandas do tamanho esperado (154 pb), indicando os genótipos de cada indivíduo: considerando um indivíduo a cada duas bandas, uma banda à esquerda indica homozigotos AA, uma banda à direita indica homozigotos TT e duas bandas indicam heterozigoto AT. Também foi observado um excesso de *primers* (banda abaixo de 100 pb nas figuras).

FIGURA 7 – PADRÃO ELETROFORÉTICO DOS FRAGMENTOS AMPLIFICADOS

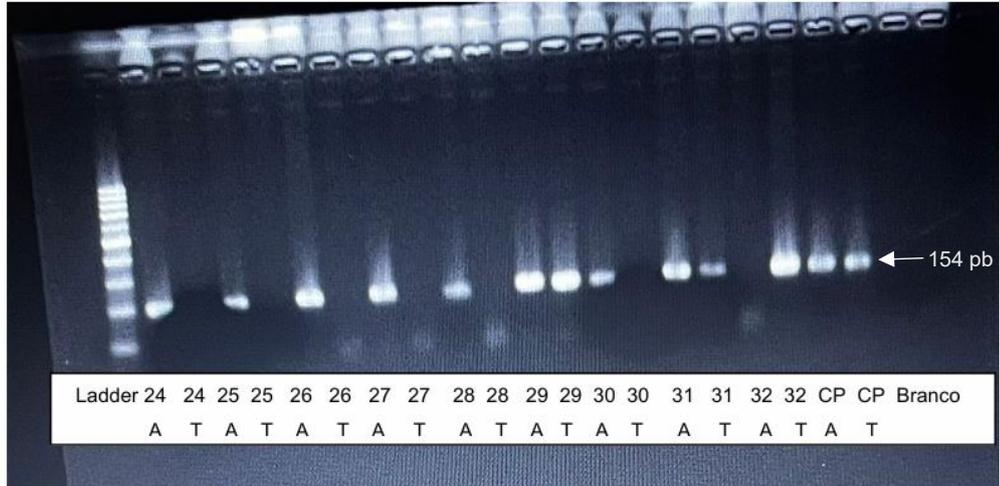


FONTE: A autora (2024).

LEGENDA: À esquerda: Marcador de peso molecular; 1 - Amostra homozigota AA; 2 - Amostra homozigota TT; 3 - Amostra heterozigota AT; 4 – branco.

Ao todo, 93 amostras de pacientes com leucemia (LLA = 75, LMA = 18) foram genotipadas por PCR alelo-específica para essa região (*vide* Apêndice B). O controle positivo refere-se às amostras sabidamente heterozigotas, devido ao Sequenciamento de Sanger realizado na etapa anterior. A amostra LLA70 foi removida do estudo por não apresentar material suficiente para as análises.

FIGURA 8 - GEL DE ELETROFORESE DA GENOTIPAGEM DO SNP rs9290663 DAS AMOSTRAS LLA24 – LLA32



FONTE: A autora (2024).

LEGENDA: As amostras estão representadas pelo código interno, referente aos alelos do SNP rs9290663: T - alelo T e A - alelo A.

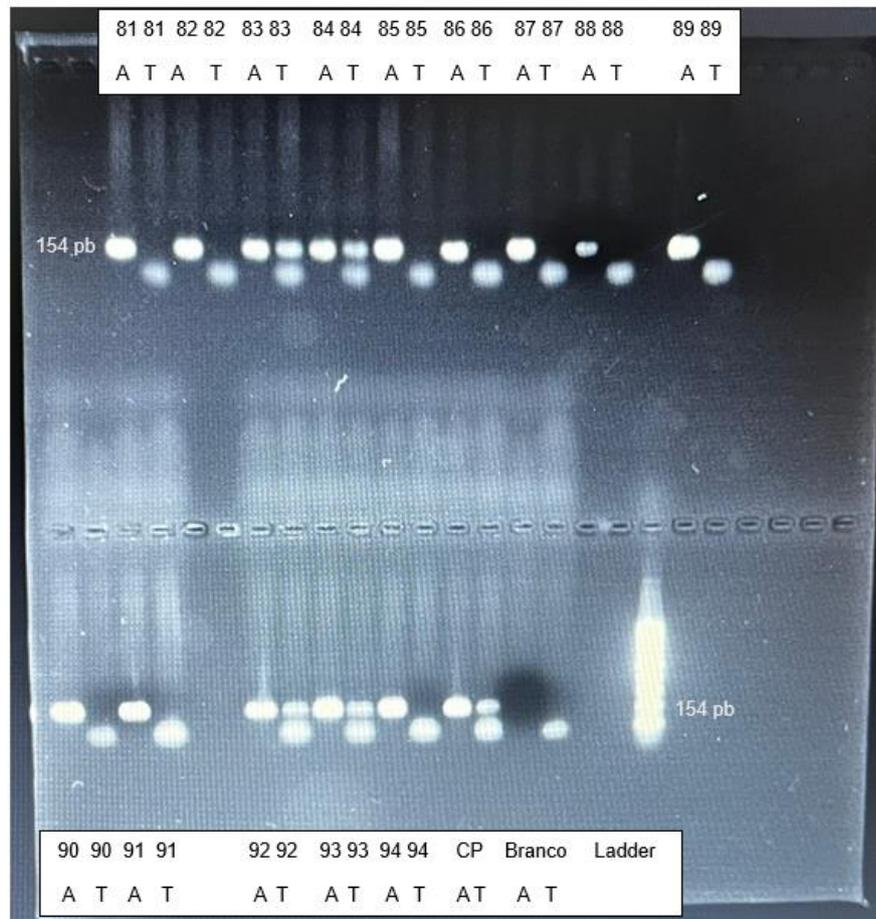
FIGURA 9 - GEL DE ELETROFORESE DA GENOTIPAGEM DO SNP rs9290663 DAS AMOSTRAS LLA63 – LLA80



FONTE: A autora (2024).

LEGENDA: As amostras estão representadas pelo código interno, referente aos alelos do SNP rs9290663: T - alelo T e A - alelo A.

FIGURA 10 - GEL DE ELETROFORESE DA GENOTIPAGEM DO SNP rs9290663 DAS AMOSTRAS LLA81 – LLA94



FONTE: A autora (2024).

LEGENDA: As amostras estão representadas pelo código interno, referente aos alelos do SNP rs9290663: T - alelo T e A - alelo A.

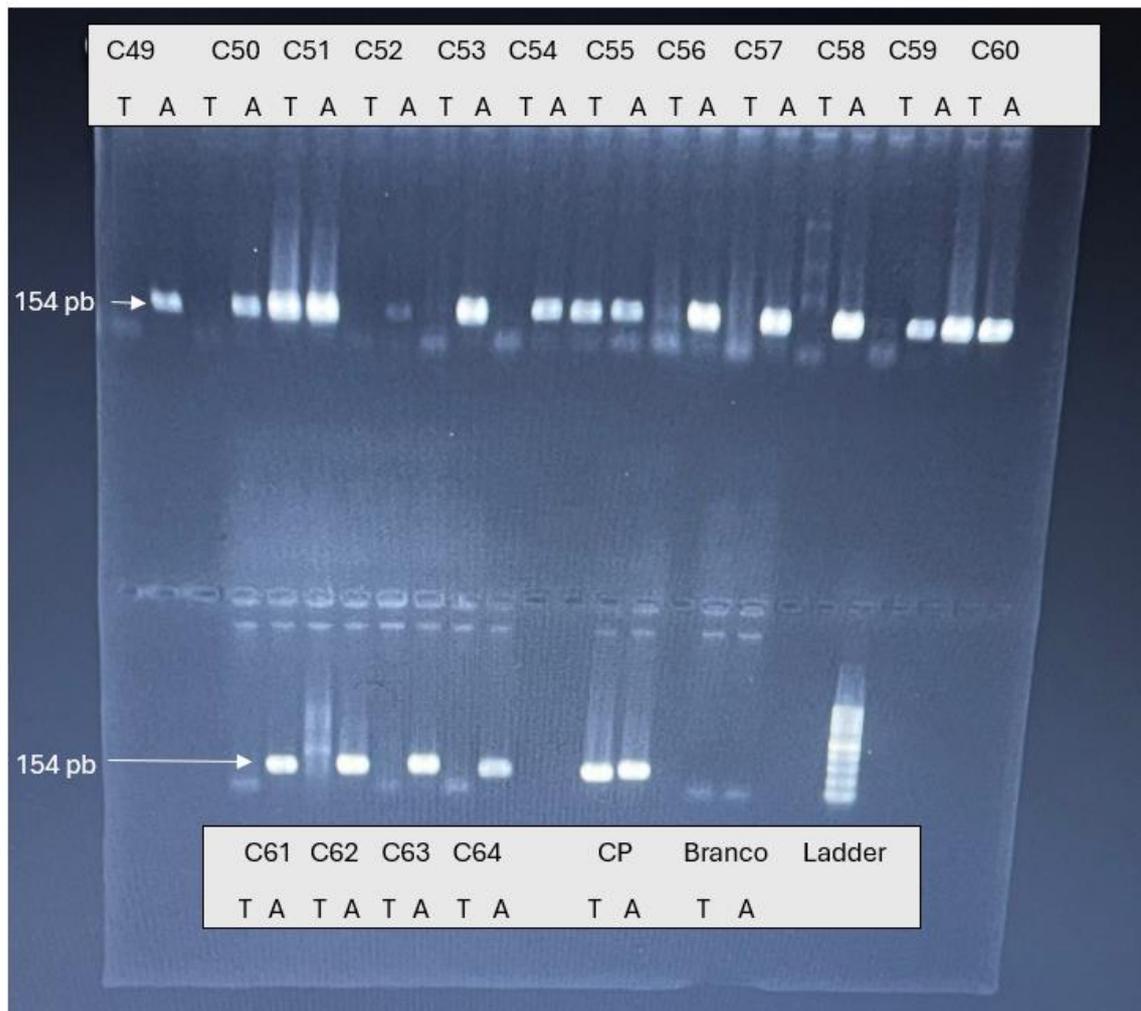
#### 4.7 GENOTIPAGEM DAS AMOSTRAS CONTROLE

As figuras 11 e 12 mostram os géis de agarose referentes à genotipagem do SNP rs9290663 nas amostras controle. É possível perceber as bandas do tamanho esperado (154 pb), indicando os genótipos de cada indivíduo: considerando um indivíduo (identificado pelo código interno) a cada duas bandas, uma banda à esquerda indica homozigotos TT, uma banda à direita indica homozigotos AA e duas bandas indicam heterozigoto AT.

Ao todo, 90 amostras controles foram genotipadas por PCR alelo-específica para essa região (*vide* Apêndice C). A amostra C9 foi removida das análises

estatísticas por não ter apresentado nenhuma banda no gel de eletroforese, que foi repetido duas vezes.

FIGURA 11 - GEL DE ELETROFORESE DA GENOTIPAGEM DO SNP rs9290663 DOS CONTROLES C49 – C64



FONTE: A autora (2024).

LEGENDA: As amostras estão representadas pelo código interno, referente aos alelos do SNP rs9290663: T - alelo T e A - alelo A.

FIGURA 12 - GEL DE ELETROFORESE DA GENOTIPAGEM DO SNP rs9290663 DOS CONTROLES C65 – C80



FONTE: A autora (2024).

LEGENDA: As amostras estão representadas pelo código interno, referente aos alelos do SNP rs9290663: T - alelo T e A - alelo A.

#### 4.8 FREQUÊNCIAS ALÉLICAS E GENOTÍPICAS

As genotipagens do lncRNA KCNMB2-AS1 foram realizadas em amostras de 93 pacientes e 90 controles. No entanto, dentre os pacientes, apenas 75 amostras são de indivíduos com leucemia linfóide aguda pediátrica, alvo principal deste trabalho. Primeiramente, foram realizadas análises estatísticas com todas as amostras de leucemia. Em seguida, as análises estatísticas foram feitas separando as 18 amostras de pacientes que não são de LLA pediátrica. Também foram feitas análises apenas com os 60 pacientes de LLA-B, o subtipo mais comum da doença.

Dessa forma, as frequências genotípicas observadas em pacientes com algum tipo de leucemia foram 66,7% (n = 62) homozigotos AA; 29% (n = 27) heterozigotos AT e 4,3% (n = 4) homozigotos TT. As frequências genotípicas observadas em pacientes com LLA pediátrica foram 64% (n = 48) homozigotos AA; 32% (n = 24) heterozigotos AT e 4% (n = 3) homozigotos TT (Tabela 5).

Já as frequências genotípicas observadas nos indivíduos controle foram 67,8% (n = 61) homozigotos AA; 32,2 % (n = 29) heterozigotos AT e zero homozigotos TT (Tabela 5). A frequência alélica do SNP rs9290663 calculada no grupo de pacientes com LLA ou somente LLA-B foi de A: 80% e T: 20%, e no grupo de controles A: 84% e T: 16%, bem similares.

TABELA 5 – FREQUÊNCIAS ALÉLICAS E GENOTÍPICAS EM PACIENTES E CONTROLES

GENÓTIPO OU ALELO	PACIENTES COM LEUCEMIA (n=93)		PACIENTES COM LLA (n=75)		CONTROLES (n=90)	
	%	n	%	n	%	n
GENÓTIPO AA	66,7	62	64	48	67,8	61
GENÓTIPO AT	29	27	32	24	32,2	29
GENÓTIPO TT	4,3	4	4	3	0	0
ALELO A	81		80		84	
ALELO T	19		20		16	

FONTE: A autora (2024).

LEGENDA: n indica o número amostral em cada grupo.

A tabela 6 mostra a frequência alélica e genotípica encontrada para o rs9290663 em diferentes populações, do estudo 1000 *Genomes Project Phase 3*. Há predominância do alelo A em todas as populações estudadas. No presente estudo a maior frequência genotípica observada em indivíduos sem leucemia também foi AA, seguido de AT, em consonância com diferentes bancos de dados.

Interessantemente, no *Gambian Genome Variation Project*, a frequência alélica do SNP rs9290663 calculada no grupo foi de A: 0.585 e T: 0.415, um valor alto para o alelo T quando comparado com outras populações externas ao continente africano.

TABELA 6 – FREQUÊNCIA DO SNP rs9290663 NO 1000 GENOMES PROJECT PHASE 3

POPULAÇÃO	FREQUÊNCIA ALÉLICA		FREQUÊNCIA GENOTÍPICA		
	ALELO A	ALELO T	A/A	A/T	T/T
AFRICANA	0.630 (833)	0.370 (489)	0.396 (262)	0.467 (309)	0.136 (90)
AMERICANA	0.879 (610)	0.121 (84)	0.778 (270)	0.202 (70)	0.020 (7)
LESTE ASIÁTICO	0.752 (758)	0.248 (250)	0.569 (287)	0.365 (184)	0.065 (33)
EUROPEIA	0.825 (830)	0.175 (176)	0.690 (347)	0.270 (136)	0.040 (20)

FONTE: *National Center for Biotechnology Information - NCBI databases (2024).*

LEGENDA: número entre parênteses indica o total de indivíduos.

#### 4.9 ESTUDOS DE ASSOCIAÇÃO

Semelhante ao risco relativo (RR), a *odds ratio* retorna um valor maior ou igual a zero e sua interpretação é feita de acordo com valores menores, iguais ou maiores do que 1. Se  $OR=1$ , ter ou não ter o fator de risco não influencia na ocorrência do desfecho (LLA pediátrica). Por outro lado, se  $OR>1$ , a chance de desenvolver o desfecho é mais provável no grupo dos indivíduos que têm o fator de risco do que entre aqueles que não o têm.

A comparação das frequências genotípicas entre pacientes e controles não apresentou associação com a suscetibilidade à LLA quando comparados AxA/T/TT (Tabela 7), nas análises com todos os pacientes com leucemia ( $OR = 0,96$ ;  $p = 0,87$ ;  $IC = 95\% 0,52 - 1,73$ ), só os 75 pacientes LLA ( $OR = 0,84$ ;  $p = 0,61$ ;  $IC 95\% = 0,44 - 1,63$ ), ou ainda analisando apenas os 60 pacientes com LLA-B ( $OR = 0,90$ ;  $p = 0,78$ ;  $IC 95\% 0,46 - 1,82$ ).

Para verificar a existência de relação alélica, foi feita a comparação entre os alelos A e T, que também não apresentou relação estatística significativa. As análises feitas considerando os pacientes com outros tipos de leucemia ou somente pacientes LLA-B também não demonstraram nenhuma associação estatística significativa (Tabela 7).

TABELA 7 - ANÁLISES DE ODDS RATIO PARA GENÓTIPOS E ALELOS

<i>Odds Ratio (IC 95%)</i>		
	<b>A vs T</b>	<b>AA vs AT/TT</b>
<b>Leucemias x controles</b>	0,83 (0,48 – 1,41)	0,96 (0,52 - 1,73)
	p = 0,49	p = 0,87
<b>LLA x controles</b>	0,77 (0,44 - 1,32)	0,84 (0,44 - 1,63)
	p = 0,36	p = 0,61
<b>LLA-B x controles</b>	0,75 (0,41 - 1,36)	0,90 (0,46 - 1,82)
	p = 0,36	p = 0,78

FONTE: A autora (2024).

## 5 DISCUSSÃO

Os estudos de associação genômica ampla (GWAS) possibilitaram uma infinidade de associações robustas para uma série de fenótipos e doenças, com o número de variantes associadas crescendo constantemente à medida que o tamanho das amostras do GWAS aumenta. Os resultados obtidos através de plataformas de acesso público têm uma gama de aplicações, tais como obter informações sobre a biologia subjacente a um fenótipo, estimar a sua herdabilidade, calcular correlações genéticas e inferir potenciais relações causais entre fatores de risco e desfechos na saúde (Uffelmann *et al.*, 2021).

Evidências cumulativas indicam a ligação entre SNPs, ncRNAs e várias doenças humanas (Ning; Zhang, 2018). Muitos lncRNAs têm sido associados à causa e progressão de cânceres, através da expressão alterada e/ou mutações em lncRNAs que atuam como oncogenes ou supressores de tumor (Mattick *et al.*, 2023).

Embora os lncRNAs representem uma parte relevante do transcriptoma, e desempenhem papel vital na complexidade humana, mantendo a função celular normal, esses transcritos ainda carecem de estudos quanto à sua caracterização funcional e a identificação das vias moleculares nas quais estão envolvidos (Janaththani; Srinivasan; Batra, 2021).

Neste estudo foi utilizado o banco de dados GWAS para identificar de forma abrangente variantes genéticas em lncRNAs associadas à suscetibilidade à LLA pediátrica. Os resultados das análises computacionais mostraram *loci* interessantes para validação da associação estatística registrada na literatura. A partir de dados públicos, o presente trabalho identificou de forma criteriosa 12 SNPs de lncRNAs associados à LLA pediátrica, sendo o SNP rs9290663 localizado no KCNMB2-AS1 o escolhido para estudo em uma população brasileira, predominantemente do Paraná.

Embora o locus de *CDKN2A/B* contenha vários genes codificadores de proteínas importantes, estudos sugerem que polimorfismos em ANRIL - lncRNA antisenso codificado neste locus - estão associados ao risco para muitos tipos de câncer diferentes, bem como ao risco para doença cardiovascular aterosclerótica, obesidade e diabetes tipo II (Kong; Hsieh; Alonso, 2018). Em células tumorais, o

ganho de função do ANRIL parece aumentar a proliferação, metástase e sobrevivência celular, enquanto a perda de função do ANRIL reduz o crescimento do tumor, a invasão e a metástase, e aumenta a apoptose e a senescência celular (Kong; Hsieh; Alonso, 2018).

No câncer de mama, vários estudos indicam que CDKN2B-AS1 pode ser um biomarcador para diagnóstico devido ao seu padrão de expressão aberrante. Qin *et al.* (2021) mostraram que o *knockdown* de CDKN2B-AS1 promoveu apoptose e reprimiu a proliferação maligna de células MCF7 e T4D7, ao interagir com miR-122-5p, um microRNA que desempenha papel importante em diversos processos, como inflamação, fibrose e regulação do metabolismo.

Wang *et al.* (2020) verificaram um aumento na expressão de CDKN2B-AS1 na leucemia mieloide aguda (LMA). O lncRNA CDKN2B-AS1 atua no desenvolvimento da LMA através da supressão epigenética de uma proteína estimuladora de TP53 (ASPP2), através da regulação negativa de miR-34a, alvo *downstream* do ANRIL. O *knockdown* do lncRNA, portanto, inibiu a proliferação, migração e invasão celular, e promoveu a apoptose de células leucêmicas.

Na LLA-T, o silenciamento de CDKN2B-AS1 atenuou a carcinogênese e aumentou a sensibilidade à adriamicina *in vivo*. Isso porque o lncRNA regula a expressão de TRAF5 (membro da família de fatores associados ao receptor de necrose tumoral) também através do sequestro de um microRNA, o miR-335-3p (Chen *et al.*, 2020).

O gene CCDC26 codifica um lncRNA de mesmo nome, o qual está associado à leucemia mieloide aguda e outras neoplasias. Foi reportado que o lncRNA CCDC26 regula a proliferação e morte de células leucêmicas através de um receptor de tirosina-quinase denominado KIT (Hirano *et al.*, 2015). Este lncRNA também é um modulador na mudança de transcrição dos genes da globina na diferenciação das células eritróides (Hirano *et al.*, 2020).

Chen *et al.* (2019) examinaram a expressão de CCDC26 em pacientes com LMA e observaram níveis de expressão elevados nos pacientes recentemente diagnosticados e recidivantes quando comparado com o grupo controle. Logo, este

lncRNA pode servir como um novo biomarcador para monitorar a progressão e prever resultados clínicos na LMA.

Em suma, embora o lncRNA CCDC26 já esteja bem descrito quanto ao seu papel na LMA, a sua associação com a LLA apenas foi reportada por polimorfismos identificados por GWAS (Vijayakrishnan *et al.*, 2018; Wiemels *et al.*, 2018; Vijayakrishnan *et al.*, 2019). Por conta disso, também será interessante avaliar a frequência do SNP rs75777619 em uma coorte brasileira em futuros trabalhos do LabCHO, uma vez que no GWAS de Vijayakrishnan e colaboradores (2019) esse polimorfismo foi apenas reportado em populações com ascendência europeia.

O presente estudo também evidenciou regiões de lncRNAs com função biológica ainda não conhecida como potenciais marcadores de suscetibilidade à LLA pediátrica (*AC124014.1*, *AC016526.1*, *AC025508.1*, *AC079921.1*, *AL035446.1*, *AL355612.1*, *MMADHC-DT*). Sendo assim, esse levantamento de dados aqui exposto fornece uma nova visão sobre a etiologia molecular desta neoplasia em crianças e adolescentes e concede informações pertinentes para trabalhos futuros.

Quanto ao *KCNMB2-AS1*, o lncRNA de escolha para este trabalho, ele já vem sendo estudado em diferentes tipos de câncer, como já mencionado. No GWAS de Treviño e colaboradores (2009), o SNP rs9290663 e sua associação com LLA foram analisados apenas em uma coorte com ascendência europeia. Além disso, quando o SNP foi investigado em uma coorte chinesa, verificou-se que o alelo rs9290663-T pode aumentar o risco para LMA, havendo relação com idade (Cao *et al.*, 2016).

Fica claro, portanto, a lacuna existente na literatura sobre a associação desse SNP em *KCNMB2-AS1* à leucemia linfóide aguda e quais as funções desempenhadas por este lncRNA no processo tumorigênico da LLA. Ainda, tal polimorfismo não foi avaliado na população brasileira, marcada por sua natureza multiétnica.

Apesar da população, tamanho geográfico e importância econômica emergente do Brasil, existem poucas pesquisas em escala genômica sobre os fatores genéticos que predispõem a população brasileira a doenças. Os estudos deste nível com a população europeia e norte-americana são inegavelmente mais robustos (Mychaleckyj, 2017).

O conhecimento dos padrões de variação genética dentro e entre as populações humanas contribuiu para uma melhor compreensão da relação entre genética e etnia, e também aprimorou a análise de estudos de associação caso-controle. E a população brasileira, devido ao seu alto grau de miscigenação, pode fornecer um recurso valioso para estratégias que visem mapear características complexas em humanos (Giolo *et al.*, 2012).

A partir dos resultados das análises deste trabalho, foi observada uma variação do MAF (frequência do alelo menor) tanto nas amostras de pacientes (0,19) como em controles (0,16), se comparadas aos dados divulgados de outras populações. No *1000 Genomes Project phase 3* (Tabela 6) também há variação nas frequências alélicas, com a população europeia (A=0,825; T=0,175) tendo valores mais similares aos encontrados no presente estudo.

No presente estudo, os valores das frequências genotípicas AT foram bem próximos no grupo LLA (0,320) e grupo controle (0,322). Apesar de 29 indivíduos no grupo controle serem heterozigotos para rs9290663, não foi encontrado nenhum indivíduo com genótipo TT neste grupo, corroborando com os dados que mostram o alelo T como o menos frequente.

Os resultados dos estudos *gnomAD – Genomes* (0,20) e *TOPMED* (0,21) possuem o MAF muito semelhantes e seus dados são compostos por mais de 50% de amostras europeias. Os valores próximos já eram esperados, visto que a população sul-brasileira possui, em sua maioria, ancestralidade europeia (Probst *et al.*, 2000).

Nas análises de OR, entretanto, não foram encontradas associações estatisticamente relevantes quando comparados os genótipos (AA x AT/TT), uma vez que todos os valores-p foram superiores a 0,05, e o OR não retornou valores maiores que um. Logo, nenhum genótipo sugeriu associação ao aumento de risco para leucemia linfóide aguda pediátrica.

As pesquisas “básicas” como essa são cruciais, pois incrementam nossa compreensão da base molecular de doenças malignas – explicando como os tumores crescem, invadem e evoluem. Tais conhecimentos são críticos para a implementação

de estratégias inovadoras e eficazes para diagnóstico, prevenção e tratamento, aprimorando a prática clínica e promovendo terapias mais direcionadas e assertivas.

Para alguns tipos de leucemia as terapias direcionadas também se tornaram parte do tratamento padrão. Esses tratamentos têm como alvo proteínas que controlam como as células cancerosas crescem, se dividem e se espalham. Diferentes tipos de leucemia requerem diferentes combinações de terapias (National Cancer Institute (NCI), 2024).

Nos últimos anos, foi estabelecido que os lncRNAs desempenham um papel importante em muitos processos biológicos, incluindo regulação genética (Batista; Chang, 2013), pontos de verificação do ciclo celular e migração celular. Seu papel na ocorrência e desenvolvimento do câncer já foi comprovado. Diante disso, o uso de lncRNAs no diagnóstico e tratamento do câncer está atraindo cada vez mais atenção de pesquisadores e médicos (Beylerli *et al.*, 2022).

Huang e colaboradores (2021) exploraram o perfil de expressão de diversos lncRNAs em amostras de LLA-B e LLA-T pediátrica por microarranjo, e observaram que os lncRNAs desregulados podem funcionar como potenciais biomarcadores que distinguem entre os dois tipos de LLA e o prognóstico (favorável ou desfavorável). Não só esse, mas vários estudos apontam os lncRNAs como bons biomarcadores para leucemia.

Existem algumas limitações do presente estudo, principalmente pelo reduzido número de amostras em cada grupo. Em vista disso, é interessante continuar com a genotipagem de rs9290663 (KCNMB2-AS1) em um número amostral maior para verificar se há mudança relevante nas frequências genotípicas e alélicas, bem como para buscar uma associação estatística significativa com a suscetibilidade à LLA pediátrica.

## 6 CONSIDERAÇÕES FINAIS

A lista de 12 SNPs em 10 lncRNAs associados à LLA pediátrica compilada nesse trabalho contribui para o avanço das pesquisas oncogenéticas com os RNAs longos não-codificantes.

Toda a metodologia foi realizada de maneira minuciosa para obtenção de resultados confiáveis, desde a seleção dos SNPs até as análises estatísticas. Nesse sentido, a genotipagem das amostras de LLA e controles por PCR alelo-específica se mostrou como uma opção vantajosa e acessível.

A média de pureza das amostras (~1,80) indicou uma qualidade aceitável, enquanto a diluição para 20 ng/μL e o ajuste de quantidade de DNA na aplicação do gel visou uniformidade nos experimentos. O Sequenciamento de Sanger possibilitou a confirmação dos genótipos encontrados na população, e todas as análises estatísticas foram realizadas utilizando um  $p < 0,05$ .

Não foram encontradas as associações estatísticas esperadas, porém este trabalho pode servir como suporte para uma melhor compreensão da base genética da LLA na população brasileira de parte da região sul, uma vez que poucos estudos desse porte são conduzidos aqui.

A genotipagem da região estudada no trabalho continuará e espera-se que no futuro a função desse SNP em KCNMB2-AS1 seja elucidada, para compreendermos o papel dessa variante na suscetibilidade à LLA pediátrica.

## REFERÊNCIAS

ALIPERTI, V., SKONIECZNA, J., CERASE, A. 2021. **Long Non-Coding RNA (lncRNA) Roles in Cell Biology, Neurodevelopment and Neurological Disorders.** *Non-coding RNA*, 7(2), 36.

AMERICAN CANCER SOCIETY (ACS). 2024. **Risk Factors for Childhood Leukemia.** Disponível em: [Risk Factors for Childhood Leukemia | American Cancer Society](#)

BACH, D. H., LEE, S. K. 2018. **Long noncoding RNAs in cancer cells.** *Cancer letters*, 419, 152–166.

BATISTA P.J., CHANG H.Y. 2013. **Long noncoding RNAs: cellular address codes in development and disease.** *Cell*, 152(6):1298–1307.

BEAL, A. L. B. 2022. **Padronização de PCR alelo-específica para análise de SNPs em RNAs longos não codificantes associados à leucemia linfóide aguda.** Trabalho de conclusão de curso (Engenharia de Bioprocessos e Biotecnologia) - Universidade Federal do Paraná, Brasil.

BENTROP, D., BEYERMANN, M., WISSMANN, R., FAKLER B. 2001. **NMR Structure of the “Ball-and-chain” Domain of KCNMB2, the  $\beta$ 2-Subunit of Large Conductance  $\text{Ca}^{2+}$ - and Voltage-activated Potassium Channels.** *Journal of Biological Chemistry*, Volume 276, Issue 45.

BEYLERLI, O., GAREEV, I., SUFIANOV, A., ILYASOVA, T., & GUANG, Y. 2022. **Long noncoding RNAs as promising biomarkers in cancer.** *Non-coding RNA research*, 7(2), 66–70.

BHAN, A., SOLEIMANI, M., MANDAL, S. S. 2017. **Long Noncoding RNA and Cancer: A New Paradigm.** *Cancer research*, 77(15), 3965–3981.

BHOJWANI, D., YANG, J. J., PUI, C. H. 2015. **Biology of childhood acute lymphoblastic leukemia.** *Pediatric clinics of North America*, 62(1), 47–60.

BLOOM, M., MACIASZEK, J. L., CLARK, M. E., PUI, C. H., NICHOLS, K. E. 2020. **Recent advances in genetic predisposition to pediatric acute lymphoblastic leukemia.** *Expert review of hematology*, 13(1), 55–70.

BUNIELLO, A. et al. 2019. **The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics.** *Nucleic Acids.*

CAO, S. et al. 2016. **Replication analysis confirms the association of several variants with acute myeloid leukemia in Chinese population.** *Journal of cancer research and clinical oncology*, v. 142, n. 1, p. 149-155.

CAO, S., LIU, M., GUO, Z., LI, Y. 2023. **Long noncoding RNA KCNMB2-AS1 acts as an oncogene in ovarian cancer.** *Acta biochimica et biophysica Sinica*, 55(11), 1844-1846.

CHADHA, R., UDAYAKUMAR, D. S., SANGWAN, S., GORE, A., JHA, B., GOEL, S., MATHUR, N., RASTOGI, N., DIXIT, R., SOOD, N., YADAV, S. P., SAXENA, R. 2023. **Cytogenetic Risk Stratification of B-Acute Lymphoblastic Leukemia and Its Correlation with Other Prognostic Factors.** *Indian journal of hematology & blood transfusion*, 39(1), 141–145.

CHEN, C., WANG, P., MO, W., ZHANG, Y., ZHOU, W., DENG, T., ZHOU, M., CHEN, X., WANG, S., WANG, C. 2019. **lncRNA-CCDC26, as a novel biomarker, predicts prognosis in acute myeloid leukemia.** *Oncology letters*, 18(3), 2203–2211.

CHEN, L., SHI, Y., LI, J., YANG, X., LI, R., ZHOU, X., ZHU, L. 2020. **lncRNA CDKN2B-AS1 contributes to tumorigenesis and chemoresistance in pediatric T-cell acute lymphoblastic leukemia through miR-335-3p/TRAF5 axis.** *Anti-cancer drugs*.

CHI, Y., WANG, D., WANG, J., YU, W., YANG, J. 2019. **Long Non-Coding RNA in the Pathogenesis of Cancers.** *Cells*, 8(9), 1015.

COCCARO, N., ANELLI, L., ZAGARIA, A., SPECCHIA, G., ALBANO, F. 2019. **Next-Generation Sequencing in Acute Lymphoblastic Leukemia.** *International journal of molecular sciences*, 20(12), 2929.

CUNNINGHAM, F. et al. Ensembl 2022. *Nucleic acids research*, v. 50, n. D1, p. D988-D995, 2022.

FIGUEROA, M. E., CHEN, S. C., ANDERSSON, A. K., PHILLIPS, L. A., LI, Y., SOTZEN, J., KUNDU, M., DOWNING, J. R., MELNICK, A., MULLIGHAN, C. G. 2013. **Integrated genetic and epigenetic analysis of childhood acute lymphoblastic leukemia.** *The Journal of clinical investigation*, 123(7), 3099–3111.

GIOLO, S. R., SOLER, J. M., GREENWAY, S. C., ALMEIDA, M. A., DE ANDRADE, M., SEIDMAN, J. G., SEIDMAN, C. E., KRIEGER, J. E., PEREIRA, A. C. 2012. **Brazilian urban population genetic structure reveals a high degree of admixture.** *European journal of human genetics : EJHG*, 20(1), 111–116.

HIRANO, T., TSURUDA, T., TANAKA, Y., HARADA, H., YAMAZAKI, T., ISHIDA, A. 2021. **Long noncoding RNA CCDC26 as a modulator of transcriptional switching between fetal and embryonic globins.** *Biochimica et biophysica acta. Molecular cell research*, 1868(3), 118931.

HIRANO, T., YOSHIKAWA, R., HARADA, H., HARADA, Y., ISHIDA, A., YAMAZAKI, T. 2015. **Long noncoding RNA, CCDC26, controls myeloid leukemia cell growth through regulation of KIT expression.** *Molecular cancer*, 14, 90.

HUANG, X., HUANG, L., XIE, Q., ZHANG, L., HUANG, S., HONG, M., LI, J., HUANG, Z., & ZHANG, H. 2021. **LncRNAs serve as novel biomarkers for diagnosis and prognosis of childhood ALL.** *Biomarker research*, 9(1), 45.

INABA, H., & MULLIGHAN, C. G. 2020. Pediatric acute lymphoblastic leukemia. *Haematologica*, 105(11), 2524–2539.

INSTITUTO NACIONAL DO CÂNCER (INCA). 2023. Leucemia. Disponível em: <https://www.gov.br/inca/pt-br/assuntos/cancer/tipos/leucemia>.

JANATHHANI, P., SRINIVASAN, S. L., BATRA, J. 2021. **Long non-coding RNAs at the Chromosomal Risk Loci Identified by Prostate and Breast Cancer GWAS.** *Genes*, 12(12), 2028.

JIANG, M. C., NI, J. J., CUI, W. Y., WANG, B. Y., ZHUO, W. 2019. **Emerging roles of lncRNA in cancer and therapeutic opportunities**. *American journal of cancer research*, 9(7), 1354–1366.

KONG, Y., HSIEH, C. H., ALONSO, L. C. 2018. **ANRIL: A lncRNA at the CDKN2A/B Locus with Roles in Cancer and Metabolic Disease**. *Frontiers in endocrinology*, 9, 405.

LIN, W., ZHOU, Q., WANG, C. Q., ZHU, L., BI, C., ZHANG, S., WANG, X., JIN, H. 2020. **LncRNAs regulate metabolism in cancer**. *International journal of biological sciences*, 16(7), 1194–1206.

LUO H., ZHU G., XU J., LAI Q., YAN B., GUO Y. ET AL. 2019. **HOTTIP lncRNA promotes hematopoietic stem cell self-renewal leading to AML-like disease in mice**. *Cancer Cell* 36, 645–659.

MATTICK, J.S., AMARAL, P.P., CARNINCI, P. et al. 2023. **Long non-coding RNAs: definitions, functions, challenges and recommendations**. *Nat Rev Mol Cell Biol* 24, 430–447.

MYCHALECKYJ, J. C., HAVT, A., NAYAK, U., PINKERTON, R., FARBER, E., CONCANNON, P., LIMA, A. A., GUERRANT, R. L. 2017. **Genome-Wide Analysis in Brazilians Reveals Highly Differentiated Native American Genome Regions**. *Molecular biology and evolution*, 34(3), 559–574.

National Cancer Institute (NCI). 2024. **Advances in Leukemia Research**. Disponible em: <https://www.cancer.gov/types/leukemia/research>

NING, S; ZHANG, Y. 2018. **Genome-Wide Mapping of SNPs in non-coding RNAs**. *Advances in Experimental Medicine and Biology*, v. 1094, p. 39-48.

PROBST, C. M. et al. 2000. **HLA polymorphism and evaluation of European, African, and Amerindian contribution to the white and mulatto populations from Paraná, Brazil**. *Human biology*, v. 72, n. 4, p. 597–617.

QIN, S., NING, M., LIU, Q., DING, X., WANG, Y., LIU, Q. 2021. **Knockdown of long non-coding RNA CDKN2B-AS1 suppresses the progression of breast cancer by miR-122-5p/STK39 axis**. *Bioengineered*, 12(1), 5125–5137.

SAMPAIO, M. A. C. 2023. **Investigação de polimorfismos de nucleotídeo único em RNAs longos não-codificantes associados à leucemia linfóide aguda infantil.** Monografia (Ciências Biológicas) - Universidade Federal do Paraná, Brasil.

SANTOS, N. B. G., NOGUEIRA, B. M. D., SOUSA, I. C., MORAES, M. E. A., MOREIRA-NUNES, C. A. 2021. **Identification of genetic alterations and their prognostic influence in Pediatric Acute Lymphoblastic Leukemia in Brazil: a systematic review.** Brazilian Journal of Development, v.7, n.3, p. 26220-26238.

SHERRY, S. T., WARD, M., SIROTKIN, K. 1999. **dbSNP - Database for single nucleotide polymorphisms and other classes of minor genetic variation.** Genome Research, v. 9, n. 8, p. 677-679.

SIRUGO, G., WILLIAMS, S. M., TISHKOFF, S. A. 2019. **The Missing Diversity in Human Genetic Studies.** Cell, 177(1), 26–31.

TANIUE, K., AKIMITSU, N. 2021. **The Functions and Unique Features of LncRNAs in Cancer Development and Tumorigenesis.** *International journal of molecular sciences*, 22(2), 632.

TREVIÑO, L. R., YANG W., FRENCH D., et al. 2009. **Germline genomic variants associated with childhood acute lymphoblastic leukemia.** Nature Genetics. 41(9):1001-1005.

UFFELMANN, E., HUANG, Q.Q., MUNUNG, N.S. et al. 2021. **Genome-wide association studies.** *Nat Rev Methods Primers* 1, 59.

VERGARA, C., LATORRE, R., MARRION, N. V., ADELMAN, J. P. 1998. **Calcium-activated potassium channels.** Current opinion in neurobiology, 8(3), 321–329.

VIJAYAKRISHNAN J. et al. 2019. **Identification of four novel associations for B-cell acute lymphoblastic leukaemia risk.** Nature communications, v. 10, n. 1, p. 1-9.

VIJAYAKRISHNAN, J. et al. 2018. **Genome-wide association study identifies susceptibility loci for B-cell childhood acute lymphoblastic leukemia.** Nature communications, v. 9, n. 1, p.1-9.

VILJOEN, C. D., BOOYSEN, C., TANTUAN, S.S. 2022. **The suitability of using spectrophotometry to determine the concentration and purity of DNA extracted from processed food matrices.** Journal of Food Composition and Analysis, Volume 112.

WAGNER, M. B., CALLEGARI-JACQUES, S. M. 1998. **Medidas de associação em estudos epidemiológicos.** Jornal de Pediatria, 74: 247-251.

WANG, C.H., LI, Q.Y., NIE, L., MA, J., YAO, C.J., CHEN, F.P. 2020. **LncRNA ANRIL promotes cell proliferation, migration and invasion during acute myeloid leukemia pathogenesis via negatively regulating miR-34a.** The International Journal of Biochemistry & Cell Biology, Volume 119.

WEI M., YU, H., CAI, C. 2020. **MiR-3194-3p inhibits breast cancer progression by targeting aquaporin1.** Front Oncol. 10:1513

WIEMELS, J. L. et al. 2018. **GWAS in childhood acute lymphoblastic leukemia reveals novel genetic associations at chromosomes 17q12 and 8q24.21.** Nature communications, v. 9, n. 1, p. 1-8.

XU, J., LIU, X., LIU, X., ZHI, Y. 2021. **Long noncoding RNA KCNMB2-AS1 promotes the development of esophageal cancer by modulating the miR-3194-3p/PYGL axis.** Bioengineered, 12(1), 6687–6702.

YANG, H., WANG, Z. 2022. **Long Noncoding RNA KCNMB2-AS1 Increases ROCK1 Expression by Sponging microRNA-374a-3p to Facilitate the Progression of Non-Small-Cell Lung Cancer.** Cancer management and research, 14, 2457–2458.

YE, J. et al. 2012. **Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction.** BMC Bioinformatics, v. 13, n. 1, p. 1-11.

ZAREI, M.M., SONG, M., WILSON, R.J., COX, N., COLOM, L.V., KNAUS, H.G., STEFANI, E., TORO, L. 2007. **Endocytic trafficking signals in KCNMB2 regulate**

**surface expression of a large conductance voltage and Ca<sup>2+</sup>-activated K<sup>+</sup> channel.** *Neuroscience*, Volume 147, Issue 1, 80-89.

ZHANG, X., WANG, W., ZHU, W., DONG, J., CHENG, Y., YIN, Z., SHEN, F. 2019. **Mechanisms and Functions of Long Non-Coding RNAs at Multiple Regulatory Levels.** *International journal of molecular sciences*, 20(22), 5573.

ZHU, J., HUANG, Y., ZHANG, Y., HUANG, R., HUANG, C. 2021. **KCNMB2-AS1 Promotes Bladder Cancer Progression Through Sponging miR-374a-3p to Upregulate S100A10.** *Frontiers in genetics*, 12, 655569.

## APÊNDICE A - DADOS DOS PACIENTES COM LEUCEMIA

(Continua)

ID Amostra	Diagnóstico	Data da coleta	Idade no diagnóstico
LLA001	LLA-B	17/11/2021	4 anos
LLA002	LLA-B	17/11/2021	4 anos
LLA003	LMA	17/11/2021	2 anos
LLA004	LLA	17/11/2021	10 anos
LLA005	LLA Pré-B	19/11/2021	19 anos
LLA006	LLA Pré-B	19/11/2021	11 anos
LLA007	LLA-B	19/11/2021	3 anos
LLA008	LLA	19/11/2021	2 anos
LLA009	LLA-B	19/11/2021	12 anos
LLA010	LLA-B	22/11/2021	2 anos
LLA011	LLA Pré-B	22/11/2021	12 anos
LLA012	LLA-B	24/11/2021	5 anos
LLA013	LLA-B	24/11/2021	6 anos
LLA014	LLA-B	24/11/2021	4 anos
LLA015	LLA-B	24/11/2021	11 anos
LLA016	LLA-B	26/11/2021	4 anos
LLA017	LLA-T	26/11/2021	8 anos
LLA018	LLA Pré-B	26/11/2021	3 anos
LLA019	LLA-B	26/11/2021	13 anos
LLA020	LLA-B	26/11/2021	13 anos
LLA021	LLA-B	29/11/2021	8 anos
LLA022	LLA-B	29/11/2021	4 anos
LLA023	LLA	01/12/2021	14 anos
LLA024	LLA-B	01/12/2021	5 anos
LLA025	LLA-B	01/12/2021	6 anos
LLA026	LLA-B	01/12/2021	3 anos
LLA027	LLA-B	01/12/2021	2 anos
LLA028	LLA-T	03/12/2021	17 anos
LLA029	LLA-B	13/12/2021	3 anos
LLA030	LLA-B	15/12/2021	9 anos
LLA031	LLA-B	15/12/2021	2 anos
LLA032	LLA-B	15/12/2021	3 anos
LLA033	LLA-B	03/01/2022	5 meses
LLA034	LLA-B	07/01/2022	5 anos
LLA035	LMA-M3	12/01/2022	18 anos
LLA036	LMA-M7	12/01/2022	1 ano
LLA037	LLA-B	12/01/2022	4 anos
LLA038	LMA-M3	12/01/2022	14 anos
LLA039	LLA	25/01/2022	2 anos
LLA040	LLA-B	25/01/2022	17 anos
LLA041	LLA-B	25/01/2022	13 anos
LLA042	LMA	28/01/2022	4 anos
LLA043	LLA-B	28/01/2022	3 anos
LLA044	LLA	28/01/2022	
LLA045	LMA-M3	31/01/2022	16 anos
LLA046	LLA-T	02/02/2022	
LLA047	LMMJ	11/02/2022	
LLA048	LLA-B	14/02/2022	3 anos

(Conclusão)

ID Amostra	Diagnóstico	Data da coleta	Idade no diagnóstico
LLA049	TL-DS	14/02/2022	
LLA050	LLA Pré-B	14/02/2022	2 anos
LLA051	LLA-B	18/02/2022	5 anos
LLA052	LLA-B	18/02/2022	2 anos
LLA053	LLA-B	18/02/2022	5 anos
LLA054	LLA-B	04/03/2022	7 anos
LLA055	LLA-B	04/03/2022	1 ano
LLA056	LMA-M3	15/03/2022	
LLA057	LLA-B	17/03/2022	2 anos
LLA058	LMA-M5	29/04/2022	
LLA059	LLA	29/04/2022	
LLA060	LMA-M7	29/04/2022	
LLA061	LMA-M3	29/04/2022	9 anos
LLA062	LLA-B	03/05/2022	2 anos
LLA063	LLA-B	04/05/2022	2 anos
LLA064	LMA-M3	04/05/2022	
LLA065	LLA	10/05/2022	12 anos
LLA066	LLA-B	10/05/2022	16 anos
LLA067	LLA-B	19/05/2022	9 anos
LLA068	LLA	19/05/2022	4 anos
LLA069	LMA-M5	20/05/2022	
LLA070	LLA-B	23/05/2022	4 anos
LLA071	LMA	23/05/2022	
LLA072	LLA	26/05/2022	5 anos
LLA073	LMA-M3	08/06/2022	
LLA074	LLA pré-B	15/06/2022	4 anos
LLA075	LLA-B	15/06/2022	4 anos
LLA076	LLA-B	17/06/2022	3 anos
LLA077	LLA pré-B	30/06/2022	4 anos
LLA078	LLA-B	04/07/2022	3 anos
LLA079	LLA-B	04/07/2022	11 anos
LLA080	LLA pré-B	07/07/2022	11 anos
LLA081	LLA B	08/07/2022	8 anos
LLA082	LLA pré-B	11/07/2022	17 anos
LLA083	LLA pré-B	11/07/2022	5 anos
LLA084	LLA pré-B	28/07/2022	18 anos
LLA085	LMA-M5	29/07/2022	13 anos
LLA086	LLA-B	29/07/2022	6 anos
LLA087	LLA-B	29/07/2022	5 anos
LLA088	LMA-M5	01/08/2022	
LLA089	LLA pré-B	18/10/2022	7 anos
LLA090	LLA-T	26/10/2022	9 anos
LLA091	LMC	09/11/2022	
LLA092	LLA	11/11/2022	10 anos
LLA093	LLA pré-B comum	11/11/2022	6 anos
LLA094	LLA pré-B	25/11/2022	7 anos

## APÊNDICE B – GENOTIPAGEM DAS AMOSTRAS DE SANGUE DOS PACIENTES

(Continua)

ID Amostra	Concentração (ng/μL)	260/280	Genótipo rs9290663
LLA001	72,6	1,83	AT
LLA002	28,8	1,71	AA
LLA003	20,5	1,8	AA
LLA004	48,7	1,74	AA
LLA005	61,4	1,85	AA
LLA006	2,5	1,09	AA
LLA007	3,4	1,11	AT
LLA008	14,7	1,6	AA
LLA009	25,1	1,94	AT
LLA010	69,5	1,93	AA
LLA011	74,5	1,92	AA
LLA012	21,1	1,77	AT
LLA013	78	1,93	AA
LLA014	49,8	1,85	AT
LLA015	17,6	1,7	AA
LLA016	14,8	1,62	AA
LLA017	57,5	1,72	AT
LLA018	107,2	1,83	AA
LLA019	27,3	1,65	AA
LLA020	13,7	1,87	AA
LLA021	5	1,73	AA
LLA022	40,7	1,8	AA
LLA023	51,3	1,81	AA
LLA024	44,5	1,79	AA
LLA025	27,6	1,75	AA
LLA026	15,7	1,85	AA
LLA027	40,8	1,87	AA
LLA028	8,9	1,67	AA
LLA029	48,1	1,88	AT
LLA030	20,8	1,8	AA
LLA031	44,1	1,72	AT
LLA032	91,7	1,81	TT
LLA033	277,2	1,79	AT
LLA034	80,6	1,79	TT
LLA035	295,4	1,84	AA
LLA036	457,8	1,85	AA
LLA037	522	1,83	AA
LLA038	23,4	1,8	AT
LLA039	210,7	1,83	AA
LLA040	44,3	1,73	AA
LLA041	35,9	1,82	TT
LLA042	94,4	1,83	AA
LLA043	76	1,79	AA
LLA044	104,4	1,82	AA
LLA045	88	1,81	AA
LLA046	157,3	1,83	AA
LLA047	88	1,81	TT
LLA048	54,1	1,8	AA
LLA049	172,9	1,83	AT
LLA050	178,2	1,84	AA

(Conclusão)

ID Amostra	Concentração (ng/μL)	260/280	Genótipo rs9290663
LLA051	36,5	1,42	AA
LLA052	87,5	1,78	AT
LLA053	88,6	1,78	AA
LLA054	25,1	1,68	AA
LLA055	453,3	1,85	AA
LLA056	93,4	1,78	AA
LLA057	91,3	1,8	AA
LLA058	195,2	1,86	AT
LLA059	30,2	1,74	AT
LLA060	195,1	1,84	AA
LLA061	85,8	1,8	AA
LLA062	112	1,81	AT
LLA063	18,5	1,7	AA
LLA064	83,7	1,91	AA
LLA065	132	1,84	AT
LLA066	44,8	1,82	AT
LLA067	79,4	1,85	AT
LLA068	61,5	1,81	AT
LLA069	44,1	1,95	AA
LLA070	-	-	-
LLA071	211,7	1,88	AT
LLA072	264,9	1,82	AA
LLA073	221,7	1,85	AA
LLA074	117,5	1,72	AT
LLA075	20,3	1,89	AA
LLA076	42,4	1,99	AA
LLA077	376,6	1,87	AA
LLA078	419,3	1,86	AT
LLA079	129,9	1,89	AA
LLA080	100,1	1,85	AT
LLA081	25,5	1,8	AA
LLA082	299,6	1,86	AA
LLA083	74	1,88	AT
LLA084	59,3	1,89	AT
LLA085	25,1	2	AA
LLA086	20,5	1,96	AA
LLA087	26,7	1,74	AA
LLA088	55	1,61	AA
LLA089	83,7	1,87	AA
LLA090	26,7	1,64	AA
LLA091	95,8	1,89	AA
LLA092	25	2,15	AT
LLA093	68,1	1,91	AT
LLA094	68,5	1,85	AA

## APÊNDICE C – GENOTIPAGEM DAS AMOSTRAS CONTROLE

(Continua)

<b>Código interno</b>	<b>Data de nascimento</b>	<b>Sexo</b>	<b>Genótipo rs9290663</b>
C1	11/05/1968	Masculino	AA
C2	16/04/1962	Masculino	AA
C3	25/05/1970	Masculino	AT
C4	29/10/1961	Masculino	AA
C5	04/12/1964	Masculino	AA
C6	29/03/1961	Masculino	AA
C7	10/04/1957	Masculino	AT
C8	12/04/1964	Feminino	AA
C9	29/01/1969	Masculino	-
C10	04/07/1969	Masculino	AA
C11	04/07/1965	Masculino	AA
C12	03/11/1963	Masculino	AT
C13	30/03/1963	Feminino	AA
C14	23/05/1967	Feminino	AA
C15	18/04/1967	Masculino	AA
C16	29/03/1968	Feminino	AT
C17	28/06/1963	Feminino	AA
C18	14/10/1969	Feminino	AA
C19	06/05/1963	Feminino	AA
C20	24/10/1964	Feminino	AA
C21	06/03/1957	Masculino	AA
C22	11/06/1966	Feminino	AA
C23	18/12/1964	Feminino	AA
C24	04/11/1965	Masculino	AA
C25	04/10/1963	Masculino	AA
C26	27/09/1969	Masculino	AT
C27	05/10/1969	Feminino	AT
C28	05/06/1970	Masculino	AT
C29	18/12/1954	Masculino	AA
C30	25/11/1957	Feminino	AT
C31	12/08/1970	Feminino	AA
C32	18/09/1964	Masculino	AA
C33	03/04/1964	Feminino	AA
C34	06/10/1968	Masculino	AT
C35	28/08/1966	Feminino	AT
C36	14/12/1969	Masculino	AA
C37	26/10/1968	Masculino	AA
C38	09/09/1961	Feminino	AA
C39	16/05/1961	Feminino	AA
C40	04/07/1958	Feminino	AA
C41	22/12/1953	Feminino	AA
C42	12/06/1966	Feminino	AA
C43	10/10/1969	Feminino	AA
C44	28/04/1969	Feminino	AA
C45	16/10/1958	Masculino	AA
C46	21/10/1954	Feminino	AA
C47	17/05/1965	Feminino	AA
C48	27/02/1963	Masculino	AA
C49	10/08/1965	Masculino	AA
C50	03/02/1958	Feminino	AA

(Conclusão)

<b>Código interno</b>	<b>Data de nascimento</b>	<b>Sexo</b>	<b>Genótipo rs9290663</b>
C51	18/05/1970	Masculino	AT
C52	10/01/1968	Feminino	AA
C53	17/05/1958	Feminino	AA
C54	29/05/1961	Feminino	AA
C55	21/09/1960	Masculino	AT
C56	03/12/1958	Masculino	AA
C57	13/04/1970	Masculino	AA
C58	14/03/1964	Feminino	AA
C59	20/04/1995	Feminino	AA
C60	21/07/1969	Feminino	AT
C61	06/11/1970	Masculino	AA
C62	25/03/1970	Feminino	AA
C63	21/01/1958	Feminino	AA
C64	15/10/1966	Feminino	AA
C65	20/04/1970	Masculino	AT
C66	11/04/1960	Feminino	AA
C67	13/12/1964	Feminino	AT
C68	10/04/1967	Masculino	AT
C69	04/11/1963	Feminino	AT
C70	26/02/1968	Feminino	AA
C71	05/05/1969	Masculino	AT
C72	10/04/1969	Feminino	AT
C73	04/09/1960	Feminino	AA
C74	26/10/1967	Feminino	AA
C75	27/03/1958	Feminino	AA
C76	25/06/1969	Masculino	AT
C77	27/10/1968	Masculino	AA
C78	09/10/1966	Masculino	AA
C79	27/01/1970	Feminino	AT
C80	10/07/1967	Feminino	AT
C81	11/09/1966	Feminino	AA
C82	13/07/1965	Masculino	AT
C83	05/08/1967	Feminino	AA
C84	15/03/1967	Masculino	AT
C85	13/01/1964	Feminino	AA
C86	24/07/1970	Feminino	AT
C87	20/12/1965	Feminino	AT
C88	23/07/1967	Feminino	AT
C89	10/04/1961	Masculino	AT
C90	05/01/1958	Masculino	AT
C91	26/11/1966	Masculino	AA