

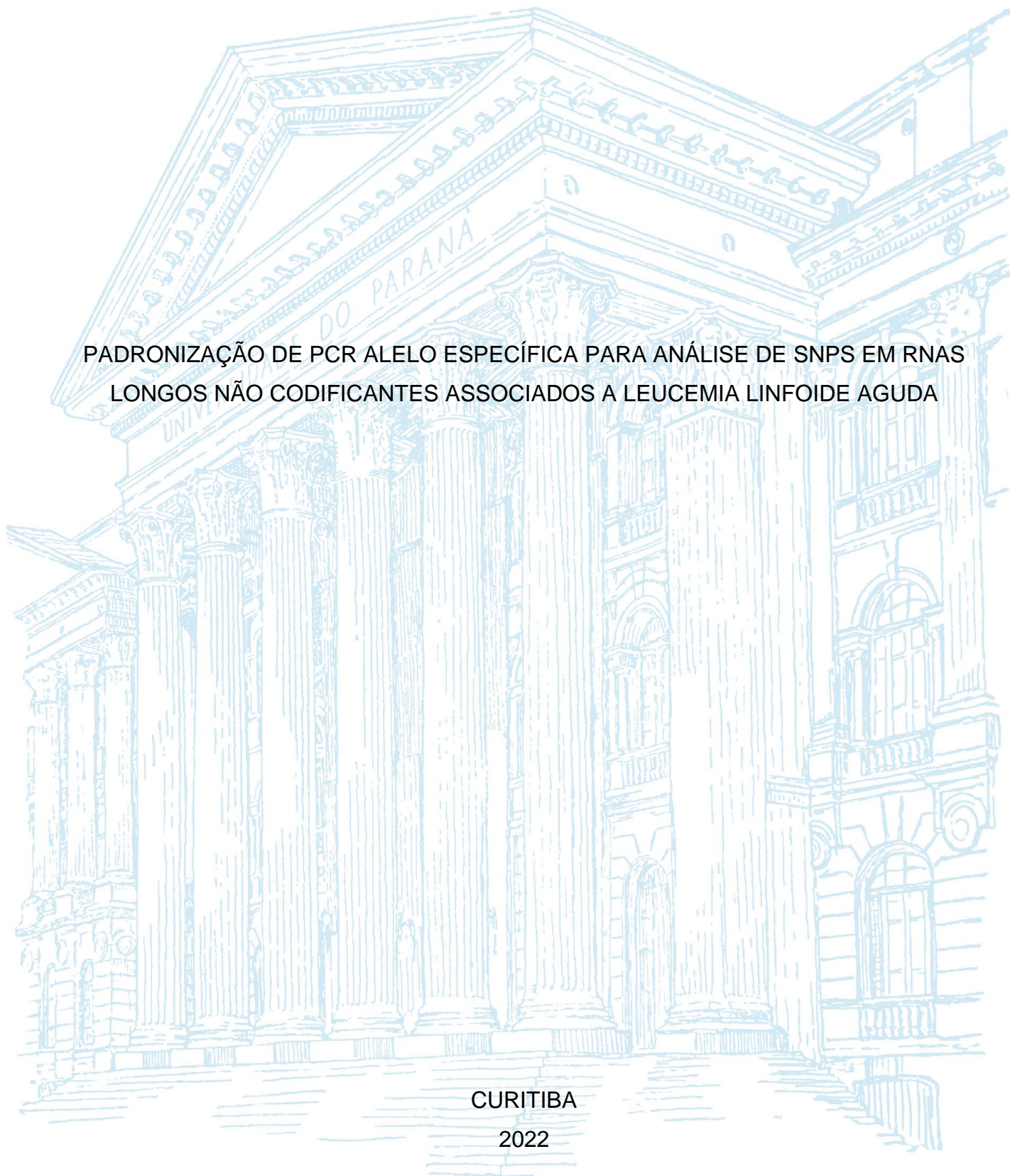
UNIVERSIDADE FEDERAL DO PARANÁ

ANA LAURA BERNHARD BEAL

PADRONIZAÇÃO DE PCR ALELO ESPECÍFICA PARA ANÁLISE DE SNPS EM RNAs
LONGOS NÃO CODIFICANTES ASSOCIADOS A LEUCEMIA LINFOIDE AGUDA

CURITIBA

2022



ANA LAURA BERNHARD BEAL

PADRONIZAÇÃO DE PCR ALELO ESPECÍFICA PARA ANÁLISE DE SNPS EM RNAS
LONGOS NÃO CODIFICANTES ASSOCIADOS A LEUCEMIA LINFOIDE AGUDA

Trabalho de Conclusão do Curso apresentado ao curso de Engenharia de Bioprocessos e Biotecnologia, Setor Palotina, Universidade Federal do Paraná, como requisito parcial à obtenção do título de bacharel.

Orientadora: Profa. Dra. Jaqueline Carvalho de Oliveira

Coorientadora: Dra. Carolina Mathias

CURITIBA

2022



UNIVERSIDADE FEDERAL DO PARANÁ

ATA DE REUNIÃO

Aos 5 dias do mês de maio do ano de dois mil e vinte e dois, às 13:30 horas, na Sala virtual https://teams.microsoft.com/l/meetup-join/19%3apdn4T3tTevujBkLgFrLQ8bc67sziKT_oC-vy3yW18qs1%40thread.tacv2/1650551216568?context=%7b%22Tid%22%3a%22c37b37a3-e9e2-42f9-bc67-4b9b738e1df0%22%2c%22Oid%22%3a%22fa0c854f-a9f7-41fd-b270-3b5ec6d767f6%22%7d da Plataforma Teams, Universidade Federal do Paraná, Setor Palotina, realizou-se a Defesa Pública e Oral do Trabalho de Conclusão de Curso intitulado "**PADRONIZAÇÃO DE PCR ALELO ESPECÍFICA PARA ANÁLISE DE SNPS EM RNAS LONGOS NÃO CODIFICANTES ASSOCIADOS A LEUCEMIA LINFOIDE AGUDA**" apresentado pela discente Ana Laura Bernhard Beal, orientada pela Profa. Dra Jaqueline Carvalho de Oliveira e co-orientada pela Dra Carolina Mathias, como um dos requisitos obrigatórios para conclusão do curso de graduação em Engenharia de Bioprocessos e Biotecnologia. Iniciados os trabalhos, a orientadora e Presidente da Banca concedeu a palavra à discente, para exposição do seu trabalho. A seguir, foi concedida a palavra em ordem sucessiva aos membros da Banca de Exame, os quais passaram a arguir a discente. Ultimada a defesa, que se desenvolveu nos termos normativos, a Banca de Exame, em sessão secreta, passou aos trabalhos de julgamento, tendo atribuído à discente as seguintes notas: Profa. Dra. Ana Cláudia Bonatto, nota: 95,0 (noventa e cinco), Profa. Dra. Daniela Malheiros, nota: 93,0 (noventa e três), e Profa. Dra. Jaqueline Carvalho de Oliveira, nota: 95,0 (noventa e cinco). A nota final da discente, após a média aritmética dos três membros da banca de exame, foi **94,0 (noventa e quatro)**. As considerações e sugestões feitas pela Banca de Exame deverão ser atendidas pela discente sob acompanhamento de sua orientadora. Nada mais havendo a tratar foi lavrada a presente ata, que, lida e aprovada, vai por todos assinada eletronicamente.



Documento assinado eletronicamente por **JAQUELINE CARVALHO DE OLIVEIRA, PROFESSOR DO MAGISTERIO SUPERIOR**, em 09/05/2022, às 11:29, conforme art. 1º, III, "b", da Lei 11.419/2006.



Documento assinado eletronicamente por **DANIELLE MALHEIROS FERREIRA, PROFESSOR DO MAGISTERIO SUPERIOR**, em 09/05/2022, às 11:30, conforme art. 1º, III, "b", da Lei 11.419/2006.



Documento assinado eletronicamente por **ANA CLAUDIA BONATTO, PROFESSOR DO MAGISTERIO SUPERIOR**, em 09/05/2022, às 11:35, conforme art. 1º, III, "b", da Lei 11.419/2006.



A autenticidade do documento pode ser conferida [aqui](#) informando o código verificador **4491105** e o código CRC **7FBD9FFE**.

AGRADECIMENTOS

Primeiramente, a minha orientadora por ter me concedido a oportunidade de fazer parte do Laboratório de Citogenética Humana e Oncogenética, me proporcionando uma das experiências mais incríveis da minha vida, e contribuindo significativamente para o meu crescimento pessoal e profissional.

À minha família, por todo o seu amor e suporte dado, que me ajudaram a trilhar novos caminhos e conquistar grandes oportunidades.

À minha irmã e meus amigos, por suas palavras de incentivo e todos os momentos de alegria e apoio proporcionados.

Aos meus amigos do período acadêmico, por compartilharem de diversos momentos de alegrias, e de dificuldades, que foram importantes para a nossa evolução.

Aos meus colegas do Laboratório de Citogenética Humana e Oncogenética, que me acolheram, me auxiliaram e tornaram a rotina do laboratório algo leve e divertido.

A todos meus professores, pelos conhecimentos e lições de vida compartilhados, que certamente contribuíram para o meu 'eu' de hoje, e me inspiram a ir mais longe.

E o mais importante, à Deus, por ter me permitido chegar onde estou hoje, me abençoando com sua infinita bondade e me possibilitado sonhar cada vez mais alto.

RESUMO

As leucemias são neoplasias de células hematopoiéticas caracterizadas pelo acúmulo de células transformadas na medula óssea, que substituem as células normais sanguíneas. São caracterizadas de acordo com a sua linhagem precursora e o grau de maturação das suas células, sendo divididas em mieloide ou linfoide, e aguda ou crônica, respectivamente. Uma vez que a leucemia linfoide aguda, ou LLA, é o tipo mais comum de câncer infantil, e tem por característica variações moleculares significativas, estudos que busquem uma melhor compreensão da base genética dessa doença são de extrema relevância. Assim, os RNAs longos não codificantes (lncRNAs), transcritos de mais de 200 nucleotídeos de tamanho, conhecidos por seu padrão de expressão tecido-específico, são um importante objeto de estudo, visto que a desregulação da sua expressão ou mutação já vem sendo associada a inúmeras doenças humanas, incluindo diferentes tipos de tumores malignos. Dessa maneira, a avaliação dos lncRNAs e seus polimorfismos são promissores para a descoberta de novas associações relacionadas ao desenvolvimento da LLA. Uma ferramenta que vem auxiliando, nesse sentido, para a identificação de variantes genéticas associadas a fenótipos em amostras populacionais são os estudos de associação genômica ampla (do inglês, *Genome-wide Association Study*- GWAS). Nesses estudos, milhares de polimorfismos de nucleotídeo único (SNPs) associados a risco são identificados, principalmente em regiões não codificantes do genoma. No entanto, dado que as regiões codificantes são mais conhecidas e apresentam funções mais bem caracterizadas, a maioria dos trabalhos focam sua atenção nessas moléculas, ficando os SNPs mapeados em lncRNAs pouco explorados. Assim, o presente trabalho objetiva identificar SNPs em lncRNAs associados com LLA, em GWAS, e padronizar uma reação em cadeia da polimerase alelo específica para duas regiões selecionadas. Para tal, através da análise de dados de associação genética relacionados a LLA, obtidos na plataforma *GWAS Catalog*, obtivemos 12 variantes em 10 lncRNAs, no qual os SNPs rs75777619, no CCDC26, e rs9290663 do lncRNA KCNMB2-AS1 foram selecionados para o desenho de *primers* e otimização da PCR alelo-específica. Após a testagem de diferentes temperaturas de pareamento de *primers*, o SNP da região lncRNA CCDC26 foi padronizado com 67°C e para o lncRNA KCNMB2-AS1, a temperatura ideal de pareamento foi de 63°C, sendo possível padronizar a reação dos dois lncRNAs selecionados em PCR convencional, e análise por eletroforese em gel de agarose. Desse modo, com os resultados do presente trabalho, concluímos que será possível realizar uma análise caso-controle com pacientes que já desenvolveram LLA e foram atendidos pelo hospital Erasto Gaertner, Curitiba-PR, visando-se uma melhor compreensão da base genética dessa doença na população brasileira.

Palavras-chaves: LLA, lncRNAs, GWAS, PCR alelo-específica.

ABSTRACT

Leukemias are neoplasms of hematopoietic cells characterized by the accumulation of transformed cells in the bone marrow, which replace normal blood cells. They are characterized according to their precursor lineage and the degree of maturation of their cells, being divided into myeloid or lymphoid, and acute or chronic, respectively. Since acute lymphoblastic leukemia, or ALL, is the most common type of childhood cancer, and is characterized by significant molecular variations, studies that seek a better understanding of the genetic basis of this disease are extremely relevant. Thus, long non-coding RNAs (lncRNAs), transcripts over 200 nucleotides in size, known for their tissue-specific expression pattern, are an important object of study, since the deregulation of their expression or mutation has already been associated to numerous human diseases, including different types of malignant tumors. Therefore, the evaluation of lncRNAs and their polymorphisms are promising for the discovery of new associations related to the development of ALL. A tool that has been helping, in this sense, for the identification of genetic variants associated with phenotypes in population samples are the Genome-wide Association Study (GWAS). In these studies, thousands of risk-associated single nucleotide polymorphisms (SNPs) are identified, mainly in non-coding regions of the genome. However, given that the coding regions are better known and have better characterized functions, most works focus their attention on these molecules, remaining the mapped SNPs in lncRNAs less explored. Consequently, the present work aims to identify SNPs in lncRNAs associated with ALL, in GWAS, and optimize an allele-specific polymerase chain reaction for two selected regions. To this end, through the analysis of genetic association data related to LLA, obtained from the GWAS Catalog platform, we obtained 12 variants in 10 lncRNAs, in which the SNPs rs75777619, in CCDC26, and rs9290663 of lncRNA KCNMB2-AS1 were selected for primer design and optimization of allele-specific PCR. After testing different primer annealing temperatures, the SNP of the lncRNA CCDC26 region was standardized at 67°C and for the lncRNA KCNMB2-AS1, the ideal annealing temperature was 63°C, making it possible to standardize the reaction of the two lncRNAs selected in conventional PCR, and analysis by agarose gel electrophoresis. So, with the results of the present work, we conclude that it will be possible to carry out a case-control analysis with patients who have already developed ALL and were treated at the Erasto Gaertner hospital, Curitiba-PR, aiming at a better understanding of the genetic basis of this disease in Brazilian population.

Keywords: ALL, lncRNAs, GWAS, allele-specific PCR.

SUMÁRIO

1 INTRODUÇÃO	8
2 REVISÃO BIBLIOGRÁFICA	10
2.1 LEUCEMIA LINFOIDE AGUDA	10
2.2 RNAS NÃO CODIFICANTES	11
2.3 RNAS LONGOS NÃO CODIFICANTES	11
2.3.1 lncRNAs no câncer e leucemias	13
3 JUSTIFICATIVA	14
4 OBJETIVOS	14
4.1 OBJETIVO GERAL	14
4.2 OBJETIVOS ESPECÍFICOS.....	15
5 MATERIAL E MÉTODOS	15
5.1 ANÁLISES <i>IN SILICO</i>	15
5.1.1 Extração de dados GWAS	15
5.1.2 Filtragem dos dados GWAS	15
5.1.3 Seleção dos SNPs	16
5.1.4 Desenho e Avaliação dos <i>primers</i>	16
5.2 CASUÍSTICA	18
5.3 ANÁLISES <i>IN VITRO</i>	19
5.3.1 Otimização dos <i>Primers</i>	19
5.3.2 Extração e quantificação de DNA	20
6 RESULTADOS	20
6.1 LNCRNAS SELECIONADOS	20
6.2 DESENHO DOS PRIMERS	22
6.3 OTIMIZAÇÃO DA REAÇÃO DE PCR PARA GENOTIPAGEM.....	23
6.4 QUANTIFICAÇÃO DE DNA.....	25
7 DISCUSSÃO	27
8 CONSIDERAÇÕES FINAIS	30
REFERÊNCIAS	31

1 INTRODUÇÃO

Nos últimos anos, com os avanços em tecnologias de sequenciamento de DNA e RNA, foi possível identificar que apenas uma fração do genoma de aproximadamente 22.300 representam genes codificadores de proteínas. Apesar dessa baixa taxa de tradução, cerca de 80% do genoma é realmente transcrito, o que levanta questões sobre a natureza dos demais RNAs, também conhecidos como RNAs não codificantes (ncRNAs) (YAO, WANG, CHEN, 2019; POLICARPO et al., 2021).

Uma classe importante dentro dos ncRNAs, e que vem ganhando destaque em decorrência ao seu papel na regulação da expressão gênica, são os RNAs longos não codificantes (lncRNA), transcritos de mais de 200 nucleotídeos de tamanho (WU et al., 2014). Conhecidos por seu padrão de expressão tecido-específico, a superexpressão, deficiência ou mutação dos lncRNAs já vem sendo associadas a inúmeras doenças humanas, incluindo diferentes tipos de cânceres (ESTELLER, 2011).

Uma vez que os principais resultados de pesquisas básicas são transformados em métodos de diagnóstico e estratégias terapêuticas, a identificação e estudo acerca os lncRNAs vem proporcionando sua caracterização como a próxima geração de biomarcadores para doenças humanas (WU, DU, 2017). Desse modo, a avaliação desse transcrito pode auxiliar na elaboração de um diagnóstico e novas classificações moleculares mais precisas (WU, DU, 2017).

Um tipo de câncer que vem sendo bem descrito quanto a importância dos lncRNAs em seu processo carcinogênico são as leucemias, visto que os lncRNAs apresentam um papel funcional no desenvolvimento das células sanguíneas (ALVAREZ-DOMINGUEZ et al., 2014; ALVAREZ-DOMINGUEZ, LODISH, 2017). Por conseguinte, o estudo e caracterização de novos lncRNAs são essenciais para uma melhor compreensão da fisiopatologia dessa doença, bem como para a descoberta de potenciais novos biomarcadores (CHEN et al., 2017).

As leucemias são neoplasias de células hematopoiéticas classificadas de acordo com a origem celular e o grau de maturação das células neoplásicas, podendo ser divididas em mieloide ou linfóide, e aguda ou crônica, respectivamente (SZCZEPAŃSKI, VAN DER VELDEN, VAN DONGEN, 2003). São divididas em mais de 12 tipos, no qual as quatro primárias são a leucemia mieloide aguda (LMA), leucemia mieloide crônica (LMC), leucemia linfóide aguda (LLA) e leucemia linfóide crônica (CLL) (INCA, 2021).

A LLA é o tipo mais comum de câncer infantil e tem como perfil a proliferação descontrolada de células B e T anormais e imaturas, que comprometem a hematopoiese normal e órgãos como o fígado, baço, linfonodos, timo, meninges e as gônadas (PUI, JEHA, 2007).

Embora o tratamento de LLA venha melhorando nos últimos anos, principalmente em crianças, com taxas de sucesso de 90% em grandes centros, em comparação aos 40% para adultos, alguns pacientes ainda desenvolvem complicações devido ao tratamento intensivo (HUNGER, MULLIGHAN, 2015; PUI, JEHA, 2007). Ademais, pesquisas indicam que alguns pacientes que atingem a remissão clínica após a fase inicial do tratamento podem sofrer recaída da doença, sendo que para pacientes com recaída precoce (<18 meses do diagnóstico), apenas 11% têm chance de cura (CARROLL, RAETZ, 2011; HUNGER ET AL., 2011).

Dessa maneira, a avaliação dos lncRNAs e seus polimorfismos são promissores para a descoberta de novas associações relacionadas ao desenvolvimento da LLA (INABA, MULLIGHAN, 2020). Uma ferramenta que vem auxiliando na identificação de variantes genéticas associadas a fenótipos em amostras populacionais são os estudos de associação genômica ampla (do inglês, *Genome-wide Association Study- GWAS*). Nesses estudos milhares de polimorfismos de nucleotídeo único (no inglês, *Single Nucleotide Polymorphisms – SNPs*) associados a risco são identificados, principalmente em regiões não codificantes do genoma (ZHANG, LUPSKI, 2015).

Apesar disso, a relação entre os loci relacionados a fenótipos e os lncRNAs ainda são pouco compreendidas (JIN et al., 2011; MIRZA et al., 2014; ZHANG, LUPSKI, 2015). Além do mais, a falta de representatividade de diferentes ancestralidades em estudos genômicos humanos, uma vez que são predominantemente baseados em populações europeias, pode resultar em informações incompletas ou equivocadas (SIRUGO, WILLIAMS, TISHKOFF, 2019).

Assim sendo, o presente trabalho objetiva padronizar uma reação em cadeia da polimerase alelo específica para a análise de lncRNA-SNPs associados a leucemia linfóide aguda, identificados por GWAS. Através dos resultados gerados, estudos de avaliação de frequência e associação com LLA serão realizados, em uma pesquisa caso-controle com pacientes que já desenvolveram LLA e foram atendidos pelo hospital Erasto Gaertner, Curitiba-PR, visando-se uma melhor compreensão da base genética dessa doença na população brasileira.

2 REVISÃO BIBLIOGRÁFICA

2.1 LEUCEMIA LINFOIDE AGUDA

A leucemia linfóide aguda é uma neoplasia maligna caracterizada pela proliferação clonal e acúmulo de linfócitos B e T imaturos na medula óssea. Embora ocorra em qualquer idade, é uma doença que apresenta seu pico de incidência em crianças de 2 a 5 anos de idade, representando o tipo mais comum de câncer infantil (SILVA, 2009).

Ainda que a etiologia da LLA permaneça desconhecida, essa neoplasia já vem sendo associada a diferentes fatores de risco, como genéticos e/ou ambientais (FUJITA et al., 2021). Além disso, sua investigação clínica vem sendo realizada através de estudos morfológicos, imunofenotípicos e, em alguns casos, genéticos (ONCIU, 2009).

Mesmo que o diagnóstico a nível morfológico- classificação L1, L2, L3, e imunofenotípico- LLA precursora de células B (B-LLA) e LLA precursora de células T (T-LLA) - seja relativamente homogêneo, a LLA demonstra uma heterogeneidade significativa a nível genético, visto que alterações genéticas e anomalias cromossômicas estão comumente envolvidas na diferenciação e proliferação das células precursoras linfóides (FUJITA et al., 2021). Dessa maneira, a avaliação da base genética dessa neoplasia vem se mostrando uma ferramenta importante para uma estratificação de risco mais precisa, contribuindo para os protocolos de tratamentos atuais (ONCIU, 2009).

Segundo Inaba e Mullighan (2020), importantes moléculas associadas ao risco de predisposição genética a LLA são os ncRNAs e seus polimorfismos, no qual GWAS já indicaram pelo menos 13 loci associados à LLA. Apesar do risco relativo ser tipicamente baixo para cada variante (correspondendo a um aumento de até 1,5 ou 2 vezes), esses SNPs, cumulativamente, podem resultar em um aumento de até 10 vezes no risco de LLA (INABA, MULLIGHAN, 2020).

Desse modo, os estudos dos ncRNAs e seus polimorfismos são uma ferramenta importante para uma melhor compreensão dessa neoplasia, bem como para o desenvolvimento de novas abordagens direcionadas, em que transcritos pouco explorados são os lncRNAs.

2.2 RNAs NÃO CODIFICANTES

O dogma central da biologia afirma que a informação genética está essencialmente contida no DNA que é transcrito em RNA e posteriormente traduzido em proteínas (CRICK, 1970). Não obstante, através dos conhecimentos acumulados nas últimas décadas, sabe-se que esse processo é muito mais complexo que isso, uma vez que a grande maioria da sequência do DNA é transcrita em RNAs não codificantes de proteína, ou ncRNAs (BERRETTA, MORILLON, 2009).

Em geral, os ncRNAs são conhecidos por seu papel em múltiplos processos celulares e podem ser divididos em (1) ncRNAs *housekeeping* e (2) ncRNAs regulatórios (ZHANG et al., 2019). Os primeiros são RNAs constitutivamente transcritos e apresentam funções essenciais para o funcionamento celular, tais como o RNA ribossomal (rRNA) e RNA transportador (tRNA) com papéis estruturais e regulatórios na maquinaria de tradução (MATTICK, MAKUNIN, 2006; PALAZZO, LEE, 2015). Os segundos, por sua vez, são RNAs associados a regulação da expressão gênica, em que os microRNAs (miRNAs), RNAs que interagem com Piwi (piRNAs), RNAs longos não codificantes (lncRNAs) são alguns exemplos (HOMBACH, KRETZ, 2016).

Classicamente, para efeito de sistematização, os ncRNAs são classificados acerca do seu tamanho, em que pequenos ncRNAs, incluindo os miRNAs e piRNAs, estão em uma faixa de 17 nucleotídeos (nt) à 200 nt, enquanto os lncRNAs são maiores que 200 nt (LI et al., 2021). Geralmente, os pequenos RNAs são conhecidos por serem reguladores negativos da expressão de RNAs alvos, enquanto os lncRNAs possuem funções regulatórias mais complexas e flexíveis nos processos biológicos, que incluem a compensação de dosagem, *splicing* alternativo, *imprinting* genômico, inativação do cromossomo X, e assim por diante (LI et al., 2021). Dessa maneira, os ncRNAs são relevantes por seu papel tanto na modulação de processos moleculares quanto celulares, e cada vez mais estudos estão sendo realizados com enfoque nessas moléculas (ZHANG et al., 2019).

2.3 RNAs LONGOS NÃO CODIFICANTES

Os RNAs longos não codificantes representam uma família heterogênea de RNAs, que podem diferir tanto em sua biogênese como origem genômica (STATELLO et al., 2021). Apesar disso, a maioria dos lncRNAs são transcritos pela RNA polimerase II,

capeados pela 7-metil guanosina em sua extremidade 5', e poliadenilados em sua extremidade 3', sendo similares aos RNAs que codificam proteínas, também conhecidos como RNA mensageiros (mRNA) (LI et al., 2021).

Apesar das semelhanças com os mRNAs, os lncRNAs apresentam características distintas no que diz respeito ao seu nível de expressão, função e localização celular (ULITSKY, BARTEL, 2013). Dessa maneira, estão associados a diferentes papéis regulatórios no contexto celular, como na remodelação da cromatina, modificação da histona, metilação do DNA, além de poder atuar como fatores de transcrição, acentuadores, etc. (WU, DU, 2017).

Um grande desafio nos últimos anos envolvendo os lncRNAs é uma classificação global desses transcritos, visto que diferentes nomenclaturas vêm sendo utilizadas de acordo com os atributos empíricos originalmente usados para detectá-los (LAURENT, WAHLESTEDT, KAPRANOV, 2015).

Dessa forma, uma classificação comumente realizada se baseia na localização genômica dos lncRNAs, sendo esses classificados em relação aos genes codificadores adjacentes em sense, antisense, bidirecional, intrônico ou intergênico (LI et al., 2021). Ainda, podem ser classificados quanto a sua função, sendo divididos de acordo com o seu mecanismo regulatório, no qual alguns exemplos são seu modo de ação como andaimes, guia, chamariz, esponja, precursor de pequenos ncRNAs, ou outros (JARROUX, MORILLON, PINSKAYA, 2017).

Assim sendo, os lncRNAs apresentam ação em diversos processos biológicos, em que seus padrões específicos de expressão coordenam o estado da célula, diferenciação e desenvolvimento de doenças. Desse modo, a superexpressão, deficiência ou mutação de lncRNAs vem sendo associado a inúmeras doenças (QUINN, CHANG, 2016).

Nesse sentido, cada vez mais estudos voltados ao reconhecimento do papel dos lncRNAs no sistema nervoso, muscular, cardiovascular, adiposo, hematopoiético, imune, e suas patologias associadas, vem sendo realizados (STATELLO et al., 2021). Com isso, busca-se explorar a possibilidade dos lncRNAs como potenciais alvos terapêuticos, bem como biomarcadores para diagnóstico e prognósticos das diferentes patologias (LÓPEZ-JIMÉNEZ, ANDRÉS-LEÓN, 2021).

2.3.1 lncRNAs no câncer e leucemias

Estudos sugerem que os lncRNAs exercem um papel insubstituível na iniciação e progressão do câncer, visto sua interação com diferentes proteínas, DNA e RNAs, bem como sua atuação na regulação da sobrevivência celular, crescimento, invasão, metástase e angiogênese tumoral, quando diferencialmente expressos (GAO et al., 2020).

Nesse sentido, a desregulação dos lncRNAs pode resultar em um modo de ação tanto como oncogene quanto como supressor de tumor, em que alguns lncRNAs já descritos por seu papel promotor de tumor são os HOTAIR, MALAT1, PCAT-1, PCGEM1, TUC338, e como supressores de tumor são os GAS5, linc-p21, MEG3 e PTENP1 (ZHANG et al., 2013).

O RNA antisenso de transcrição HOX (HOTAIR), RNA longo não codificante intergênico (lincRNA), é um dos lncRNAs mais estudados e conhecidos dado sua superexpressão na maioria dos tumores sólidos e correlação com invasão, progressão, metástase e mau prognóstico (WU et al., 2014). Esse lincRNA atua como um andaime molecular, ligando e direcionando modificações de histonas e reprogramando o estado da cromatina, favorecendo a progressão dessa malignidade (WU et al., 2014).

Outro lncRNA muito bem estudado é o gene 3 expresso maternamente, ou MEG3, expresso nos tecidos normais. O MEG3 induz o acúmulo da proteína p53, estimulando sua transcrição e regulando sua expressão (ZHOU, 2007). Por conseguinte, em tumores, é observado que a inativação do MEG3 leva a um aumento significativo na expressão de genes promotores de angiogênese, demonstrando seu papel como supressor de tumor (ZHOU, ZHANG, KLIBANSKI, 2012).

No contexto das leucemias, esses dois lncRNAs, HOTAIR e MEG3, também já foram descritos quanto ao seu papéis como supressor e promotor de tumor (HAO, SHAO, 2015; ZHANG et al., 2016; LI et al., 2018) Não obstante, em comparação com a hematopoiese normal e outros tipos de cânceres, os lncRNAs em malignidade hematopoiéticas ainda não foram extensivamente caracterizados (HAN, CHEN, 2013).

Dessa forma, estudos que visem a avaliação de lncRNAs quanto a sua superexpressão, deficiência ou mutação no contexto das leucemias são importantes para uma melhor compreensão do desenvolvimento dessa neoplasia. Tendo isso em vista, uma ferramenta importante que vem sendo amplamente utilizada para a identificação de variantes em lncRNAs associados ao câncer são os GWAS (ZHANG et al., 2013). Apesar de já serem descritos alguns lncRNAs e seus SNPs no contexto da LLA, estudos

adicionais ainda são necessários para uma compreensão mais precisa da função desses transcritos e suas variantes nesse tipo de câncer (CHEN et al., 2017).

3 JUSTIFICATIVA

A leucemia linfóide aguda é uma doença heterogênea com variações moleculares significativas, sendo essenciais trabalhos que identifiquem variações genéticas de predisposição ou proteção associados a essa neoplasia. Recentemente, um crescente volume de trabalhos nessa área permitiram que um conjunto de informações fossem integradas em banco de dados, auxiliando na busca por variações que possam favorecer uma melhor compreensão da base genética dessa doença, como, por exemplo, a publicação de GWAS. Apesar de a grande maioria dos polimorfismos identificados por esses estudos serem mapeados em ncRNAs, ainda pouco se é descrito sobre os seus loci associados, dado que o foco permanece nas regiões codificantes do genoma. Dessa maneira, a identificação de SNPs em lncRNAs por meio de GWAS e a avaliação desses por estudos caso-controle podem contribuir para novos conhecimentos, seja no âmbito fisiopatológico como clínico, dado o papel regulatório dos lncRNAs em leucemias. Para esse fim, a genotipagem de amostras por PCR alelo-específicas pode se tornar uma ferramenta útil e de baixo custo. Porém, deve ser realizada com uma otimização refinada da metodologia, anterior as análises, para a obtenção de resultados confiáveis e diferenciação dos genótipos.

4 OBJETIVOS

4.1 OBJETIVO GERAL

O objetivo deste trabalho é a padronização de reações em cadeia da polimerase (PCRs) alelo-específicas para análises de SNPs em lncRNAs, previamente associados a LLA por GWAS, visando-se uma futura avaliação da frequência alélica desses polimorfismos em uma coorte brasileira.

4.2 OBJETIVOS ESPECÍFICOS

- Avaliar a presença de SNPs associados com LLA e mapeados em lncRNAs através de dados depositados no catálogo GWAS;
- Selecionar SNPs em lncRNAs de interesse, através da análise do GWAS e dados da literatura;
- Desenhar os *primers* para os SNPs selecionados, atendendo aos critérios das análises *in silico*;
- Otimização da PCR alelo-específica *in vitro*.

5 MATERIAL E MÉTODOS

5.1 ANÁLISES *IN SILICO*

5.1.1 Extração de dados GWAS

Os GWAS vem sendo amplamente utilizados para a identificação de loci de interesse responsáveis pela variação fenotípica (SILVA, 2018). Dessa maneira, considerando-se a categoria *traits* (fenótipos) da plataforma GWAS *Catalog* (BUNIELLO et al., 2019) foram pesquisados os seguintes termos em busca de associações genéticas relacionados a LLA: “*ALL*”, “*Acute Lymphoid Leukemia*”, “*Acute Lymphoblastic Leukemia*”.

Dado que cada conjunto de dados na plataforma é identificado com um EFO ID (*Experimental Factor Ontology Identifier*), 2 fenótipos associados a LLA em si foram encontrados (EFO_0000220 e EFO_0000094). A partir da leitura dos dados, observou-se que o EFO_0000094 é um subconjunto de EFO_0000220, e assim, para a continuidade das análises, apenas o conjunto de dados EFO_0000220 foi analisado.

5.1.2 Filtragem dos dados GWAS

A etapa de filtragem do conjunto de dados selecionado consistiu na exclusão das variantes associadas a LLC e de estudos que não apresentaram associação, permanecendo apenas os GWAS relacionados a LLA. Ainda, dentro os dados de LLA, foram excluídas as variantes associadas a fenótipos secundários (e.g.: hipersensibilidade ao quimioterápico PEG-asparaginase- HØJFELDT et al., 2019) e selecionado apenas os

SNPs mapeados em regiões que transcrevem lncRNAs. Para isso, o *software BioMart* da plataforma *Ensembl* (CUNNINGHAM et al., 2022) foi utilizado, em que através do identificador gênico de cada polimorfismo foi possível identificar e selecionar apenas os SNPs localizados em lncRNAs.

5.1.3 Seleção dos SNPs

Os estudos que identificaram associações genéticas de variantes de lncRNAs com LLA foram lidos e analisados quanto a 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 < 19 anos (crianças, adolescentes e jovens adultos) foram selecionados para o desenho de *primers*.

Para a análise dos SNPs, foram usadas as plataformas *Ensembl* para dados populacionais e a dbSNP-NCBI (do inglês, *Database of Single Nucleotide Polymorphisms*) (SHERRY et al., 2001) para dados genômicos.

5.1.4 Desenho e Avaliação dos *primers*

Buscando-se avaliar os SNPs selecionados nos lncRNAs, *primers* alelo-específicos (ou seja, 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). Para isso, duas estratégias foram adotadas baseadas: (1) na realização de duas PCRs distintas, a serem realizadas em tubos diferentes, em paralelo, ou (2) na realização da PCR *multiplex*, no qual ambas as reações podem ser executadas tanto por PCR convencional como por PCR em tempo real (qPCR). Para o caso da qPCR, o objetivo será a análise por curva de *melting*.

Para as PCRs em paralelo, o desenho dos *primers* foi realizado através de um *primer reverse* comum para dois *primers forward* alelo-específicos, considerando-se reações separadas (Figura 1). Ou dois *primers reverse* alelo-específicos e o mesmo *primer forward*.

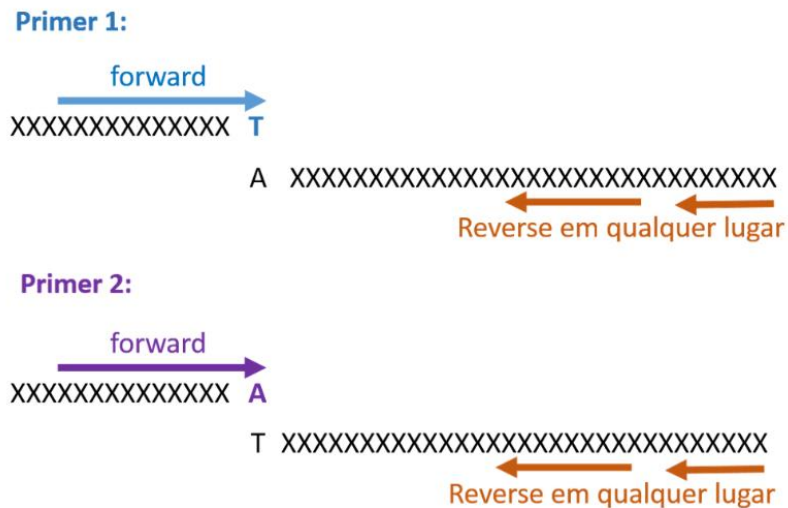


FIGURA 1. Desenho dos *primers* para genotipagem por reações separadas, em paralelo.

Já para a PCR *multiplex*, foram desenhados um *primer forward* alelo-específico e um *reverse* alelo-específico, com seus respectivos pares, para uma mesma reação de PCR (Figura 2), de forma que o par de *primer forward* alelo-específico/reverso comum amplifique um fragmento de tamanho X, e o reverso alelo-específico/*forward* comum amplifique um outro fragmento de tamanho Y. Com isso, um dos fragmentos amplificados apresentará uma sequência de tamanho e de composição de nucleotídeos diferentes, o que permitirá que ambos os alelos possam ser analisados simultaneamente por suas respectivas temperaturas de *melting* (T_m), em que uma diferença de T_m ideal entre os pares de *primers* é igual ou maior que 5°C (PAPP et al., 2003).

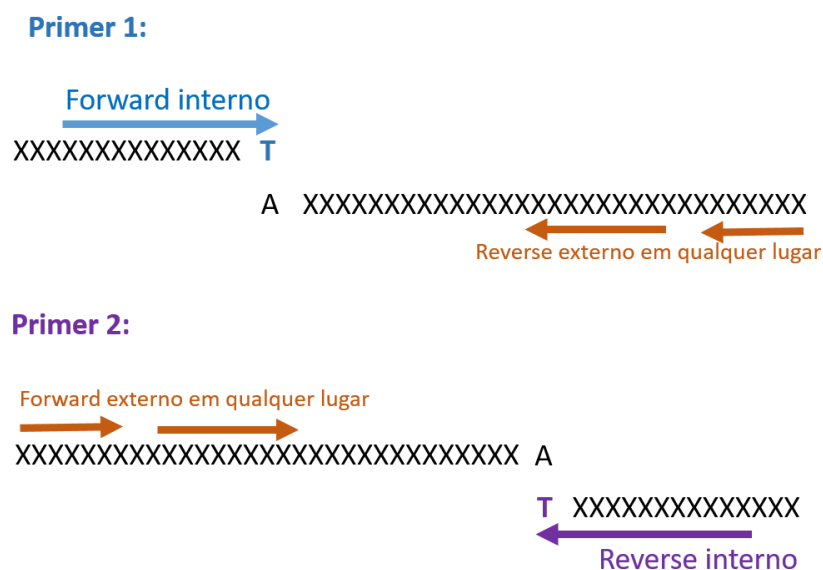


FIGURA 2. Desenho dos *primers* para genotipagem *multiplex*.

Assim, a primeira etapa para o desenho dos *primers* em ambos os casos consistiu no *download* da sequência flanqueadora de cada SNP- aproximadamente 1000 pb- em *fasta*, através da dbSNP do NCBI. Essa sequência foi inserida na ferramenta *Primer-Blast* do NCBI como PCR *template* e uma sequência de 19-25 pb contendo o alelo específico na extremidade 3' foi inserido 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 anterior a inserção na plataforma. Dessa forma, 10 *primers* para cada *forward* e *reverse* alelo-específicos foram gerados.

Tendo em vista que alguns problemas comuns no desenho dos *primers* englobam à chance de formação de dímeros (auto e hetero dímeros) e grampos (dobramentos da cadeia de um *primer*, ocorrendo o pareamento consigo mesma), esses parâmetros foram analisados (VIEIRA, 2011). Além disso, foram evitados *primers* com grandes diferenças de temperaturas de *melting* - visto que dificulta uma seleção eficiente da temperatura de anelamento - assim como de sequências com repetições de dinucleotídeos (e.g., ATATATAT) ou de base única (e.g., CCCCCC) (LORENZ, 2012).

Por conseguinte, os 10 *primers* gerados foram analisados através da *Primer-Blast* do NCBI quanto à sua temperatura de *melting* (T_m) ($\Delta T_m < 5,0$), autocomplementaridade (*self-complementarity* $\leq 8,0$), autocomplementaridade 3' (*self-3'-complementarity* $\leq 8,0$) e conteúdo GC (~ 40-60%). Já pela ferramenta OligoAnalyzer (IDT™ *Integrated DNA Technologies*), conforme a estrutura tridimensional da molécula, foi avaliado a possibilidade de formação de *grampos* ($\Delta G < -6$; $\Delta G > 6$), auto-dímeros (~ 0) e hetero-dímeros (~ 0). Ao final, os *primers* que atingiram os valores estabelecidos de cada parâmetro foram pedidos para síntese com a empresa *Invitrogen* (*Thermo Scientific*).

5.2 CASUÍSTICA

Tendo em vista uma futura avaliação da frequência dos SNPs selecionados em pacientes brasileiros com LLA e indivíduos controle sem câncer, o presente trabalho está sendo realizado com amostras de sangue periférico de pacientes de LLA obtidas no Hospital Erasto Gaertner, Curitiba, Paraná, e de pacientes controles provenientes do banco de amostras do Laboratório de Citogenética Humana e Oncogenética (LabCHO).

As amostras de câncer foram coletadas sob aprovação do Comitê de Ética em Pesquisa em Seres Humanos do Setor de Ciências da Saúde, UFPR, sob o número de CAAE: 08809419.0.0000.0098, parecer: 3.256.165. E com a assinatura do termo de

Consentimento Livre e Esclarecido (TCLE) pelos pacientes e o Termo de Assentimento Livre e Esclarecido (TALE) para pacientes menores de idade.

Quanto as amostras controle, essas foram obtidas a partir da assinatura de TCLE (CAAE: 31687620.2.0000.0096), coleta de saliva e resposta à um formulário, sendo selecionadas somente amostras de pacientes o qual não apresentavam histórico de 2 ou mais cânceres na família.

5.3 ANÁLISES *IN VITRO*

5.3.1 Otimização dos *Primers*

A fim de verificar a especificidade dos *primers* desenhados para os lncRNAs selecionados e seus polimorfismos, os *primers* sintetizados foram testados *in vitro* pela técnica de PCR convencional, por reações separadas em paralelo.

Os testes foram realizados seguindo o padrão de ciclagem exposto na tabela 1, em que a temperatura de anelamento ideal será aquela que permitirá uma reação alelo específica.

TABELA 1. Padrão de ciclagem da PCR.

	ETAPA DO CICLO	TEMPERATURA	TEMPO
x40	Desnaturação inicial	95°C	10 min
	Desnaturação	96°C	30 seg
	Anelamento	A definir	30 seg
	Alongamento	72°C	30 seg
	Alongamento final	72°C	10 min

As PCRs foram realizadas utilizando-se 2 µL de DNA, 0,5 µL da solução contendo os *primers* forward e reverse (100 nmol/100 µL), 17,5 µL de água ultrapura e 10 µL do Master Mix para PCR convencional (1x) desenvolvido pelo IBMP, ICC/FioCruz, totalizando um volume de 30 µL, seguindo sugestão do fabricante.

Os resultados gerados foram interpretados por meio de análise de eletroforese em gel de agarose 2% com o uso do marcador de peso molecular (*DNA Ladder Low Range- Thermo Scientific*) de 80-1031 pb e de 1,5 uL de Gel Red 10.000 x (*Biotium*), diluído 1:500, por amostra.

5.3.2 Extração e quantificação de DNA

Seguindo a logística da extração de material genético de amostras que são testadas para COVID do LabCHO, o DNA das amostras controle e inicialmente dos pacientes com leucemia foi extraído com a utilização do Extracta Kit da *Loccus* (MVXA-P016) - Kit de extração de ácidos nucleicos virais por *beads* magnéticas- no sistema automatizado de extração Extracta 32, sendo diluído em 80 µL.

A fim de se comparar o rendimento dessa metodologia com um protocolo específico para amostras de sangue total, a partir da amostra LLA032 a extração do DNA foi realizada por meio do Biopur™ Kit Extração Mini Spin Plus, protocolo 01. Na etapa final, um volume de 50 µL do tampão de eluição foi adicionado diretamente sobre a membrana de sílica, formando o volume final da solução.

Para ambos os protocolos, anterior a etapa de extração, as amostras de sangue periférico coletadas foram primeiramente centrifugadas a 2500 rpm por 5 minutos, em que 200 µL da porção leucocitária foram recuperados.

A quantificação do DNA de cada amostra foi realizada no equipamento NanoDrop® 2000 (*Thermo Scientific*) através da absorbância medida em 260nm e 280nm (LUCENA-AGUILAR et al., 2016). Posteriormente, as amostras foram armazenadas a -20 °C no LabCHO.

6 RESULTADOS

6.1 LNCRNAs SELECIONADOS

Através dos critérios aplicados para a filtragem e seleção dos SNPs foram obtidos ao final 12 variantes em 10 lncRNAs diferentes, conforme disposto na tabela 2. Desse modo, dois SNPs foram selecionados com base nos dados literários correspondentes aos seus lncRNAs e por suas respectivas sequências flanqueadoras, que permitiram um desenho seguro dos *primers*.

TABELA 2. lncRNAs e SNPs associados a LLA.

lncRNAs	SNP	OR	95% CI	Alelo	MAF	Referência
ACO16526.2	rs7156960	1.2	[1.09-1.33]	G	0.44	Ellinghaus et al. 2012
AC025508.1	rs630662	-	-	A	0.20	Ellinghaus et al. 2012
AC079921.1	rs10018622	1.9	[1.45-2.63]	A	0.49	Archer et al., 2017
AC079921.2						

AC124014.1	rs11978267	1.69	[1.40-1.90]	G	0.23	Treviño et al., 2009
		1.43	[1.28-1.59]			Ellinghaus et al., 2012
		1.43	-			Wielmels et al., 2018
AL035446.1	rs11155133	3.62	[2.10-6.00]	G	0.13	Treviño et al., 2009
AL355612.1	rs7738636	1.2	[1.15-1.41]	C	0.28	Ellinghaus et al., 2012
CCDC26	rs75777619*	1.26	[1.17-1.36]	G	0.10	Vijayakrishnan et al., 2019
	rs28665337	1.34	[1.21-1.47]	A	0.10	Vijayakrishnan et al., 2018
	rs4617118	1.28	[1.19-1.37]	G	0.19	Wielmels et al., 2018
CDKN2B-AS1	rs2069426	1.35	-	T	0.06	Wielmels et al., 2018
KCNMB2-AS1	rs9290663**	1.58	[1.20-1.90]	T	0.25	Treviño et al., 2009
MMADHC-DT	rs10170236	1.45	[1.24-1.69]	G	0.33	Yang et al., 2012

Fonte: Os autores (2021). Legenda: OR: *Odds Ratio* (no português, razão de possibilidade); 95% CI: Intervalo de confiança de 95%; MAF: frequência do alelo menor.

Assim, o primeiro lncRNA selecionado foi o CCDC26, visto que esse transcrito já vem sendo associado a diferentes tipos de cânceres, como os gliomas (GONZALEZ-CASTRO et al., 2019), câncer de pâncreas (PENG e JIANG, 2016), tireoide (MA et al., 2021), e etc. Apesar do estudo de Hirano e colaboradores (2015) sugerir que a expressão do CCDC26 ocorre apenas em linhagens de células hematopoiéticas de origem mieloide, os resultados GWAS enfatizam a importância do estudo da associação desse lncRNA e seus SNPs no contexto das leucemias linfoides. Para esse fim, o SNP selecionado foi o rs75777619, uma vez que sua sequência flanqueadora atendeu ao critério de conteúdo GC adequado e ausência de repetições de dinucleotídeos e de base única, em comparação aos outros polimorfismos.

Adicionalmente, o segundo SNP e lncRNA selecionado foi o rs9290663 do KCNMB2-AS1. Apesar desse transcrito já ter sido estudado em cânceres como o de bexiga (ZHU et al., 2021; CHEN et al., 2021) e de pulmão de células não pequenas (YANG, WANG, WANG, 2020), bem como o seu rs9290663 ter sido associado a um maior risco de LMA em uma coorte chinesa (CAO et al., 2016), ainda pouco se é relatado sobre a sua função regulatória na LLA.

Desse modo, a avaliação desses lncRNAs e seus SNPs é importante para uma melhor compreensão da base genética da LLA. Através da plataforma *Ensembl* foi verificado que ambos os SNPs são variantes bialélicas - A/G para o rs75777619 e A/T para o rs9290663 - em concordância ao reportado pelo GWAS de Vijayakrishnan et al. (2019) e Treviño et al. (2009), respectivamente. Ainda, foi visto que o SNP rs75777619 do CCDC26 apresenta um alto desequilíbrio de ligação com o rs28665337 em todas as populações da plataforma, exceto a Tâmil do Sri Lanka no Reino Unido (r^2 : 0.78).

6.2 DESENHO DOS PRIMERS

A sequência dos *primers* dos lncRNAs e seus SNPs, respeitando aos critérios estabelecidos na análise *in silico*, estão dispostos na tabela 3 e 4. Ambos foram sintetizados objetivando-se a reação de PCR alelo específica.

Para o SNP do lncRNA KCNMB2-AS1 optou-se por uma metodologia de PCR em paralelo, no qual os *primers reverse* são alelos específicos. Já para o SNP do lncRNA CCDC26 optou-se por uma metodologia de PCR multiplex, em que a curva de *melting* prevista para os diferentes pares de *primers* sintetizados pode ser visualizado na figura 3.

TABELA 3. Sequência dos *primers* para o rs9290663 (tamanho do fragmento: 154 pb)

FORWARD 3' – 5'	REVERSE (A) 5' – 3'	REVERSE (T) 5' – 3'
GCCGTTGTCACTGTGAGTAG	CGGAGCTACCAGTACTAA ^A	CGGAGCTACCAGTACTAT ^T

Fonte: Os autores (2022).

TABELA 4. Sequência dos *primers* para o rs75777619.

PAR DE PRIMER 1 (118 pb)		PAR DE PRIMER 2 (291 pb)	
FORWARD (A) 3' – 5'	REVERSE 5' – 3'	FORWARD 3' – 5'	REVERSE (G) 5' – 3'
GTATGGTATGCACCTGG GTCAT ^A	TGCTTTGGTTGTTCTTG ACTGAC	GACGTAATGCCCTTA CCTG	GCCTGAACCCTTAAAC CCAC ^C

Fonte: Os autores (2022)

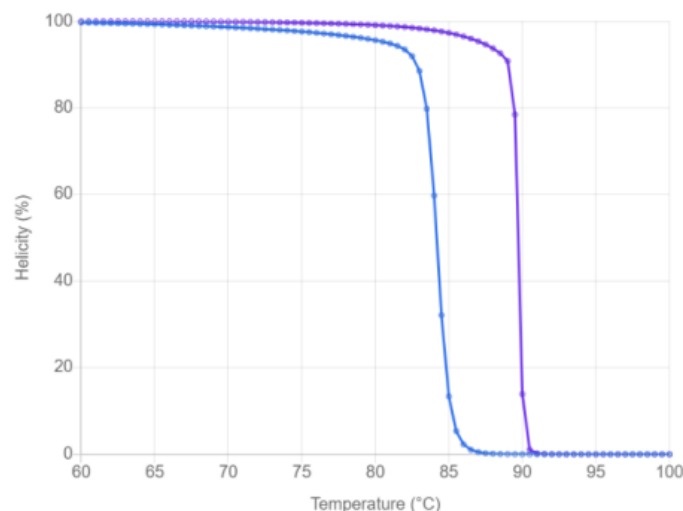


FIGURA 3. Curva de *melting* prevista para os pares de *primers* do SNP rs75777619 do lncRNA CCDC26 através do *Software DNA-UTAH*. Legenda: curva em azul- *primer* 118 pb (alelo A); roxo- *primer* 291 pb (alelo G).

6.3 OTIMIZAÇÃO DA REAÇÃO DE PCR PARA GENOTIPAGEM

Visando-se a realização da genotipagem por PCR alelo-específica, testes iniciais foram realizados a fim de encontrar a temperatura de anelamento ideal para uma reação alelo-específica, assim como para otimizar as condições da PCR (Figura 4). Desse modo, a escolha da temperatura ideal foi baseada na última temperatura a qual permitiu à amplificação e diferenciação dos genótipos (Figura 5).

Sob componentes e condições da PCR otimizados, os padrões da banda de PCR que se formaram após a eletroforese em gel de agarose podem ser visualizados na figura 6 e 7. Assim, na temperatura de 63°C foi alcançado a diferenciação dos genótipos para o lncRNA KCNMB2-AS1 e seu SNP (alelo A/T), e à 67°C para o lncRNA CCDC26.

A fins de exemplificação, o genótipo de duas amostras controles podem ser verificados, no qual para o rs9290663 do lncRNAs KCNMB2-AS1 (Figura 6) tanto indivíduos homozigotos (alelo A- comum), como heterozigotos (alelo A e T), foram observados, em um n=30. E para o rs75777619 do lncRNAs CCDC26 (Figura 7), até o presente momento, apenas indivíduos homozigotos (alelo A- comum) foram identificados, em um n=10. Entretanto, para ambos os casos, ainda não foi detectado apenas o genótipo de homozigose no alelo de menor frequência dos SNPs.

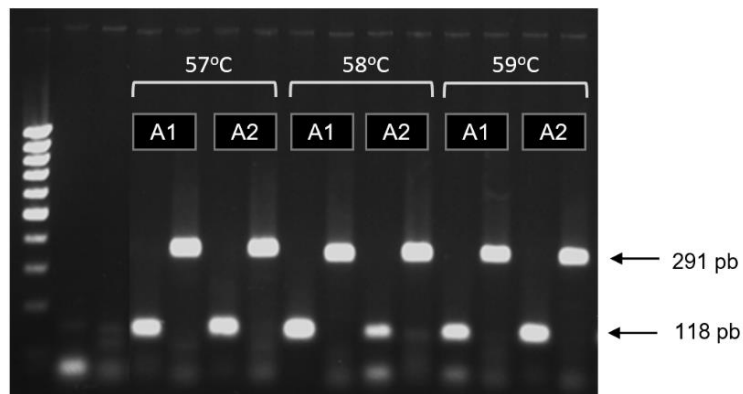


FIGURA 4. Teste inicial do lncRNA CCDC26 e seu rs75777619 nas temperaturas de 57, 58, 59°C. Legenda: A1- amostra 1; A2- amostra 2. Resultado: as temperaturas não permitiram uma reação alelo-específica. Ademais, um excesso de *primers* foi gerado, sendo ajustado a concentração para 0,2 µL da solução contendo os *primers forward* e *reverse* (100 nmol/100 µL) para 17,8 µL de água ultrapura.

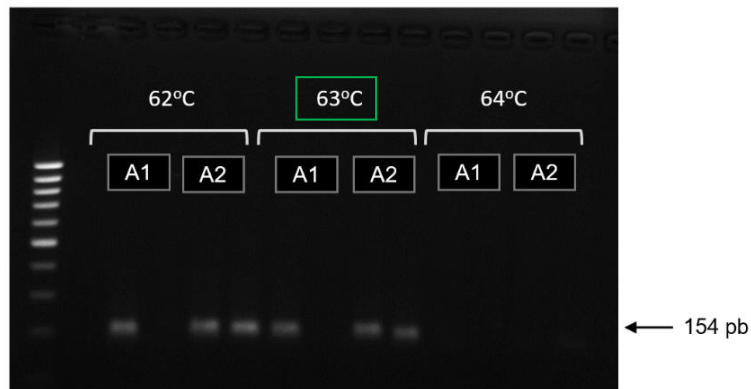


FIGURA 5. Teste final do lncRNA KCNMB2-AS1 quanto a temperatura de anelamento ideal. Legenda: A1- amostra 1; A2- amostra 2. Resultado: 63 °C é a temperatura ideal.

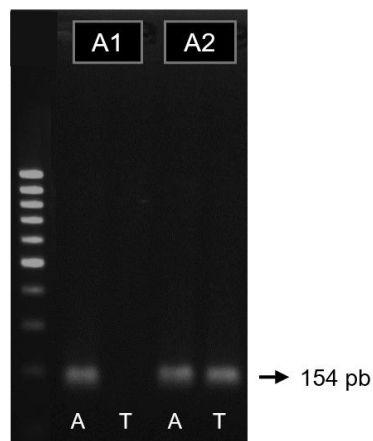


FIGURA 6. Gel de eletroforese do produto da PCR à 63°C do rs9290663 do lncRNAs KCNMB2-AS1. Legenda: A1- amostra 1; A2- amostra 2. Resultado: A1 (homozigoto- alelo A); A2 (heterozigoto- alelo A e T).

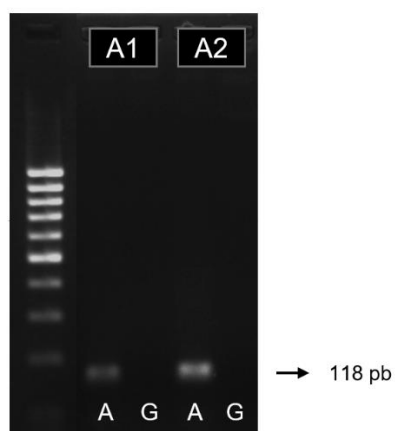


FIGURA 7. Gel de eletroforese do produto da PCR à 67°C do rs75777619 do lncRNAs CCDC26. Legenda: A1- amostra 1; A2- amostra 2. Resultado: A1 (homozigoto- alelo A); A2 (homozigoto- alelo A).

Dessa forma, a partir dos resultados gerados foi possível padronizar a reação dos dois lncRNAs selecionados. Assim, as tabelas abaixo resumem as condições finais a serem utilizadas para a realização da genotipagem por PCR alelo-específica no estudo de caso-controle dos pacientes com LLA.

TABELA 5. Ciclagem da reação do lncRNA KCNMB2-AS1.

	ETAPA DO CICLO	TEMPERATURA	TEMPO
	Desnaturação inicial	95°C	10 min
x40	Desnaturação	96°C	30 seg
	Anelamento	63°C	30 seg
	Alongamento	72°C	30 seg
	Alongamento final	72°C	10 min

TABELA 6. Ciclagem da reação do lncRNA CCDC26

	ETAPA DO CICLO	TEMPERATURA	TEMPO
	Desnaturação inicial	95°C	10 min
x40	Desnaturação	96°C	30 seg
	Anelamento	67°C	30 seg
	Alongamento	72°C	30 seg
	Alongamento final	72°C	10 min

Por fim, cabe destacar que ainda não foram testadas e otimizadas reações *multiplex*, bem como PCRs em tempo real. Não obstante, a partir dos resultados obtidos nesse trabalho é possível inferir que a realização da genotipagem das amostras de LLA por PCR convencional é uma opção vantajosa e acessível.

6.4 QUANTIFICAÇÃO DE DNA

Os dados referentes à quantificação das amostras de DNA dos pacientes de leucemia que formam a coorte do presente estudo estão apresentados na tabela 7. Até o presente, 57 amostras de leucemias foram coletadas, sendo 48 amostras de pacientes de LLA.

Os resultados alcançados pelas diferentes metodologias de extração do DNA podem ser visualizados na tabela abaixo, no qual a utilização do *Biopur*TM Kit Extração Mini Spin Plus permitiu um maior rendimento- quantidade média de DNA de 5633,61 ng em comparação a 2002,14 ng do *Extracta* Kit.

Em relação a razão 260/280, que indica pureza do material genético, e espera-se valores próximos a 1,8 para DNA, verificou-se uma média de 1,64 para o DNA extraído pelo *Extracta* Kit e 1,76 para o *Biopur* Kit. Quanto a razão 260/230, indicativa de contaminação, obteve-se uma média de 1,60 e 1,78, respectivamente, em que valores entre 1,8 e 2,2 são adequados. Desse modo, a partir desses resultados observou-se também melhores valores para o DNA extraído pelo *Biopur* Kit. Entretanto, futuras PCRs devem ser realizadas a fim de verificar a integridade das amostras, visando-se os testes de genotipagem.

TABELA 7. Quantificação das amostras de DNA de pacientes com leucemia (0-19 anos) tratados no Hospital Erasto Gaertner.

ID AMOSTRA	CONCENTRAÇÃO (ng/μl)	QUANTIDADE (ng)	A260 (ABS)	A280 (ABS)	260/280	260/230
EXTRACTA KIT (MVXA-P016)						
LLA001	69,5	5560	1,390	0,758	1,83	1,27
LLA002	3,9	312	0,079	0,061	1,28	1,02
LLA004	33,6	2688	0,672	0,372	1,81	1,49
LLA005	62,7	5016	1,254	0,677	1,85	2,03
LLA006	1,6	128	0,031	0,029	1,09	2,49
LLA007	3,4	272	0,067	0,061	1,11	1,00
LLA008	14,7	1176	0,294	0,184	1,60	1,51
LLA009	21,9	1752	0,438	0,233	1,88	1,82
LLA010	69,5	5560	1,389	0,720	1,93	2,07
LLA011	47,8	3824	0,955	0,511	1,87	1,84
LLA012	21,1	1688	0,421	0,237	1,77	2,68
LLA013	78,0	6240	1,560	0,806	1,93	2,53
LLA014	53,7	4296	1,073	0,563	1,91	2,63
LLA015	4,6	368	0,092	0,061	1,50	-1,84
LLA016	2,2	176	0,044	0,037	1,18	0,5
LLA017	6,6	528	0,133	0,076	1,75	0,78
LLA018	12,7	1016	0,255	0,122	2,08	1,45
LLA019	4,1	328	0,082	0,053	1,55	0,71
LLA020	3,3	264	0,067	0,063	1,06	-2,74
LLA021	10,1	808	0,203	0,117	1,73	1,31
LLA022	41,6	3328	0,833	0,452	1,84	1,73
LLA023	8,4	672	0,169	0,096	1,76	1,57
LLA024	6,5	520	0,130	0,078	1,67	1,02
LLA025	9,8	784	0,197	0,099	1,98	1,46
LLA026	33,0	2640	0,661	0,358	1,85	1,65

LLA027	34,0	2720	0,681	0,364	1,87	1,38
LLA028	15,9	1272	0,318	0,191	1,67	0,91
LLA029	31,0	2480	0,619	0,330	1,88	1,73
LLA030	35,6	2848	0,712	0,395	1,80	1,70
LLA031	10,0	800	0,200	0,101	1,98	1,26
BIOPUR™ KIT EXTRAÇÃO MINI SPIN PLUS						
LLA032	24,6	1230	0,492	0,217	2,27	0,03
LLA033	353,3	17665	7,065	3,805	1,86	2,39
LLA034	99,4	4970	1,987	1,101	1,80	2,18
LLA037	240,5	12025	4,810	2,613	1,84	2,34
LLA039	243,3	12165	4,867	2,636	1,85	2,38
LLA040	50,5	2525	1,009	0,566	1,78	2,11
LLA041	50,9	2545	1,018	0,563	1,81	2,17
LLA043	92,9	4645	1,858	1,010	1,84	2,14
LLA044	25,0	1250	0,050	0,129	0,39	-0,02
LLA046	87,1	4355	1,742	0,944	1,85	2,38
LLA048	66,3	3315	1,325	0,741	1,79	1,64
LLA050	118,7	5935	2,374	1,307	1,82	2,02
LLA051	58,5	2925	1,170	0,650	1,80	1,66
LLA052	101,4	5070	2,028	1,119	1,81	1,99
LLA053	44,6	2230	0,891	0,499	1,79	1,65
LLA054	9,5	475	0,190	0,105	1,81	0,81
LLA055	280,4	14020	5,609	3,054	1,84	2,19
LLA057	81,2	4060	1,623	0,906	1,79	1,96

7 DISCUSSÃO

Os cânceres são causados por mutações que podem ser herdadas, induzidas por fatores ambientais ou resultantes de erros de replicação de DNA (TOMASETTI, LI, VOGELSTEIN, 2017). Uma vez que um interesse comum é a compreensão da base genética dessa doença, uma ferramenta importante que vem sendo utilizada nos últimos anos com essa finalidade é o GWAS, visto que sugere a relação causal entre variantes genéticas dentro de uma espécie e as diferenças fenotípicas observadas entre os indivíduos (KORTE, FARLOW, 2013).

Dado que quase 90% de todos os SNPs associados a fenótipos são identificados em regiões não codificantes do genoma, nos GWAS, uma classe importante entre os ncRNAs são os lncRNAs (MIRZA et al., 2014). Embora os lncRNAs representem uma

parte relevante do transcriptoma, e sejam importantes para diversos processos celulares, é notável que esses transcritos ainda carecem de estudos quanto a sua caracterização funcional e a identificação das vias moleculares nos quais estão envolvidos (JANATHHANI, SRINIVASAN, BATRA, 2021).

O lncRNA CCDC26 e seus SNPs vem sendo associados a diversos tipos de cânceres, incluindo a leucemia mieloide (WILSON, KANHERE, 2021). Localizado no locus 8q24.21, o CCDC26 encontra-se na região de importantes proteínas associadas a tumorigênese- FAM84B, GSDMC, FAM49B e ASAP1, assim como do oncogene MYC e de lncRNAs com funções independentes ao MYC (WILSON, KANHERE, 2021).

Hirano e colaboradores (2015) reportaram que o nocaute do CCDC26 em células K562 derivadas de LMC resulta em uma expressão transcricional alterada de vários genes, incluindo a ativação do KIT, CD24 e PASD1. No estudo de Li et al. (2021), o CCDC26 foi expresso de forma aberrante em células mieloides crônicas e agudas, sendo regulado positivamente e promovendo a proliferação e invasão através da via CELF2/circRNA_ANKIB1/miR-195- 5p/Eixo PRR11/PI3K/AKT. Ainda, Chen et al. (2019) observou o papel desse lncRNA na regulação de uma série de vias, incluindo 'diferenciação de osteoclastos', 'via de sinalização de receptor semelhante a NOD', 'vias metabólicas', 'biogênese de ribossomo em eucariotos' e 'droga metabolismo-citocromo P450', que apresentam um papel importante na patogênese e progressão da LMA.

Além da sua função na leucemia mieloide, o CCDC26 está associado ao risco de gliomas (GONZALEZ-CASTRO et al., 2019), câncer de pâncreas (PENG e JIANG, 2016), tireoide (MA et al., 2021), bem como a fenótipos secundários, como resistência a imatinibe em tumores estromais gastrointestinais (CAO et al, 2018; YAN et al., 2019)

Também, esse lncRNA tem sido caracterizado em relação ao seu papel na diferenciação das células eritróides (HIRANO et al., 2021) e quanto a sua influência na localização subcelular da DNA-metiltransferase 1, em que sua inibição resulta na hipometilação do DNA (JONES et al., 2021).

Em suma, embora o lncRNA CCDC26 já venha sendo bem descrito quanto a suas funções, a sua associação com a LLA apenas foi reportada através de polimorfismos identificados por GWAS (VIJAYAKRISHNAN et al., 2018; WIEMELS et al., 2018; VIJAYAKRISHNAN et al., 2019). Desse modo, o estudo desse lncRNA em LLA e a avaliação da frequência do SNP rs7577761 na coorte brasileira é algo interessante e inovador, uma vez que no GWAS de Vijayakrishnan e colaboradores (2019) esse polimorfismo foi apenas reportado em populações com ascendência europeia.

Abordando-se sobre o lncRNA KCNMB2-AS1, esse está localizado no cromossomo 3, e vem sendo descrito quanto ao seu papel no câncer de endométrio (TANG et al., 2019) câncer cervical (ZHANG et al., 2020), câncer de pulmão de células não pequenas (NSCLC) (YANG, WANG, WANG, 2020) e câncer de bexiga (ZHU et al., 2021; CHEN et al., 2021), em que para todos os estudos, com exceção do primeiro, foi verificado um efeito negativo desse lncRNA no processo da carcinogênese.

No GWAS de Treviño e colaboradores (2009), o SNP rs9290663 e sua associação com LLA também foi estudada apenas em uma coorte com ascendência europeia. Além disso, esse SNP foi investigado em uma coorte chinesa, em que os autores discutiram sua localização no íntron do KCNMB2, que codifica a subunidade $\beta 2$ do canal de potássio de grande condutância ativado por cálcio, e verificaram que o alelo rs9290663-T pode aumentar o risco de LMA, havendo relação com idade (CAO et al., 2016).

Dessa forma, o estudo e avaliação de lncRNAs e seus SNPs na população brasileira são importantes para uma melhor compreensão genética da LLA, visto que mais de 80% de todos os dados gerados por GWAS são de ancestralidade europeia, indicando uma assimetria e hegemonia na identificação de novas associações (MILLS, RAHAL, 2019). Assim, por meio dos resultados obtidos nesse trabalho, será possível realizar uma análise de frequência e associação dos SNPs selecionados, em regiões que transcrevem lncRNA, em um estudo de caso-controle no Paraná- Brasil.

Tendo isso em vista, as amostras de LLA obtidas do Hospital Erasto Gaertner e as amostras controle do LabCHO serão analisadas por PCR alelo-específico, em que uma informação importante a ser verificada é a ancestralidade dos pacientes selecionados.

Para os genótipos obtidos pela PCR em um teste inicial será realizado a validação pela técnica de sequenciamento de Sanger e, ao final, os resultados obtidos serão analisados estatisticamente através do cálculo de *Odds ratio* (OR) e teste do Qui-Quadrado. Ainda, análises funcionais podem ser realizadas.

Assim sendo, a otimização de reações alelo-específicas é um passo importante, demorado e que necessita ser realizado de maneira minuciosa para obtenção de resultados confiáveis, o que demonstra a importância do presente trabalho.

8 CONSIDERAÇÕES FINAIS

A partir dos resultados gerados no presente trabalho foi possível identificar e selecionar 12 SNPs em 10 lncRNAs associados a LLA, por GWAS. O rs75777619, no CCDC26, e o rs9290663 do lncRNA KCNMB2-AS1 foram selecionados, e seguiram para a etapa de desenho de *primers* e padronização da PCR alelo-específica. Após testagem de diferentes temperaturas de pareamento de *primers*, o SNP da região lncRNA CCDC26 foi padronizado com 67°C e o SNP do lncRNA KCNMB2-AS1 à 63°C, permitindo a otimização da reação dos dois lncRNAs selecionados em PCR convencional, e análise por eletroforese em gel de agarose. Ainda, a genotipagem das amostras de LLA por PCR convencional se mostrou como uma opção vantajosa e acessível.

Dessa maneira, com os resultados desse trabalho, concluímos que será possível realizar uma análise caso-controle com pacientes que já desenvolveram LLA e foram atendidos pelo hospital Erasto Gaertner, Curitiba-PR, visando-se uma melhor compreensão da base genética dessa doença na população brasileira.

REFERÊNCIAS

- ALVAREZ-DOMINGUEZ, J. R.; LODISH, H. F. Emerging mechanisms of long noncoding RNA function during normal and malignant hematopoiesis. **Blood, The Journal of the American Society of Hematology**, v. 130, n. 18, p. 1965-1975, 2017.
- ALVAREZ-DOMINGUEZ, J. R. et al. Long noncoding RNAs during normal and malignant hematopoiesis. **International journal of hematology**, v. 99, n. 5, p. 531-541, 2014.
- ARCHER N. P. et al. Family-based exome-wide association study of childhood acute lymphoblastic leukemia among Hispanics confirms role of ARID5B in susceptibility. **PLoS One**, v. 12, n. 8, p. e0180488, 2017.
- BERRETTA, J.; MORILLON, A. Pervasive transcription constitutes a new level of eukaryotic genome regulation. **EMBO reports**, v. 10, n. 9, p. 973-982, 2009.
- BUNIELLO, A. et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. **Nucleic acids research**, v. 47, n. D1, p. D1005-D1012, 2019.
- CAO, K. et al. CCDC26 knockdown enhances resistance of gastrointestinal stromal tumor cells to imatinib by interacting with c-KIT. **American Journal of Translational Research**, v. 10, n. 1, p. 274, 2018.
- CAO, S. et al. 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, 2016.
- CARROLL, W. L.; RAETZ, E. A. Clinical and laboratory biology of childhood acute lymphoblastic leukemia. **The Journal of pediatrics**, v. 160, n. 1, p. 10-18, 2012.
- CHEN, C. et al. lncRNA-CCDC26, as a novel biomarker, predicts prognosis in acute myeloid leukemia. **Oncology letters**, v. 18, n. 3, p. 2203-2211, 2019.
- CHEN, S. et al. Long non-coding RNAs: the novel diagnostic biomarkers for leukemia. **Environmental toxicology and pharmacology**, v. 55, p. 81-86, 2017.
- CHEN, Y. et al. Long Noncoding RNA KCNMB2-AS1 Promotes SMAD5 by Targeting miR-3194-3p to Induce Bladder Cancer Progression. **Frontiers in Oncology**, v. 11, 2021.
- CRICK, F. Central dogma of molecular biology. **Nature**, v. 227, n. 5258, p. 561-563, 1970.
- CUNNINGHAM, F. et al. Ensembl 2022. **Nucleic acids research**, v. 50, n. D1, p. D988-D995, 2022.
- ELLINGHAUS, E. et al. Identification of germline susceptibility loci in ETV6-RUNX1-rearranged childhood acute lymphoblastic leukemia. **Leukemia**, v. 26, n. 5, p. 902-909, 2012.
- ESTELLER, M. Non-coding RNAs in human disease. **Nature reviews genetics**, v. 12, n. 12, p. 861-874, 2011.

FUJITA, T. C. et al. Acute lymphoid leukemia etiopathogenesis. **Molecular Biology Reports**, v. 48, n. 1, p. 817-822, 2021.

GAO, N. et al. Long non-coding RNAs: the regulatory mechanisms, research strategies, and future directions in cancers. **Frontiers in Oncology**, v. 10, p. 2903, 2020.

GONZÁLEZ-CASTRO, T. B. et al. Genetic Polymorphisms of CCDC26 rs891835, rs6470745, and rs55705857 in Glioma Risk: A Systematic Review and Meta-analysis. **Biochemical genetics**, v. 57, n. 4, p. 583-605, 2019.

HAN, B.; CHEN, Y. Potential pathological and functional links between long noncoding RNAs and hematopoiesis. **Science signaling**, v. 6, n. 289, p. re5-re5, 2013.

HAO, S.; SHAO, Z. HOTAIR is upregulated in acute myeloid leukemia and that indicates a poor prognosis. **International journal of clinical and experimental pathology**, v. 8, n. 6, p. 7223, 2015.

HIRANO, T. et al. Long noncoding RNA CCDC26 as a modulator of transcriptional switching between fetal and embryonic globins. **Biochimica et Biophysica Acta (BBA)-Molecular Cell Research**, v. 1868, n. 3, p. 118931, 2021.

HIRANO, T. et al. Long noncoding RNA, CCDC26, controls myeloid leukemia cell growth through regulation of KIT expression. **Molecular cancer**, v. 14, n. 1, p. 1-16, 2015.

HØJFELDT, S. G. et al. Genetic predisposition to PEG-asparaginase hypersensitivity in children treated according to NOPHO ALL 2008. **British journal of haematology**, v. 184, n. 3, p. 405-417, 2019.

HOMBACH, S.; KRETZ, M. Non-coding RNAs: classification, biology and functioning. **Non-coding RNAs in colorectal cancer**, p. 3-17, 2016.

HUNGER, S. P.; MULLIGHAN, C. G. Acute lymphoblastic leukemia in children. **New England Journal of Medicine**, v. 373, n. 16, p. 1541-1552, 2015.

HUNGER, S. P. et al. Improving outcomes for high-risk ALL: Translating new discoveries into clinical care. **Pediatric blood & cancer**, v. 56, n. 6, p. 984-993, 2011.

INABA, H.; MULLIGHAN, C. G. Pediatric acute lymphoblastic leukemia. **Haematologica**, v. 105, n. 11, p. 2524, 2020.

INCA. **Leucemia**. 2021, Disponível em: <https://www.inca.gov.br/tipos-de-cancer/leucemia#:~:text=A%20leucemia%20%C3%A9%20uma%20doen%C3%A7a,substituem%20as%20c%C3%A9lulas%20sangu%C3%ADneas%20normais>. Acesso em: 27 fev. 2022.

JANATHHANI, P.; SRINIVASAN, S. L.; BATRA, J. Long Non-Coding RNAs at the Chromosomal Risk Loci Identified by Prostate and Breast Cancer GWAS. **Genes**, v. 12, n. 12, p. 2028, 2021.

JARROUX, J.; MORILLON, A.; PINSKAYA, M. History, discovery, and classification of lncRNAs. **Long Non-Coding RNA Biology**, p. 1-46, 2017.

JIN, G. et al. Human polymorphisms at long non-coding RNAs (lncRNAs) and association with prostate cancer risk. **Carcinogenesis**, v. 32, n. 11, p. 1655-1659, 2011.

JONES, R. et al. A long intergenic non-coding RNA regulates nuclear localization of DNA methyl transferase-1. **Science**, v. 24, n. 4, p. 102273, 2021.

KORTE, A.; FARLOW, A. The advantages and limitations of trait analysis with GWAS: a review. **Plant methods**, v. 9, n. 1, p. 1-9, 2013.

LAURENT, G. St; WAHLESTEDT, C.; KAPRANOV, P. The Landscape of long noncoding RNA classification. **Trends in genetics**, v. 31, n. 5, p. 239-251, 2015.

LI, C. et al. LncRNA CCDC26 Interacts with CELF2 protein to enhance myeloid leukemia cell proliferation and invasion via the circRNA_ANKIB1/miR-195-5p/PRR11 Axis. **Cell Transplantation**, v. 30, p. 0963689720986080, 2021.

LI, J. et al. An overview on identification and regulatory mechanisms of long non-coding RNAs in fungi. **Frontiers in Microbiology**, v. 12, 2021.

LI, Z. et al. Long noncoding RNA MEG3 inhibits proliferation of chronic myeloid leukemia cells by sponging microRNA21. **Biomedicine & Pharmacotherapy**, v. 104, p. 181-192, 2018.

LÓPEZ-JIMÉNEZ, E.; ANDRÉS-LEÓN, E. The implications of ncRNAs in the development of human diseases. **Non-coding RNA**, v. 7, n. 1, p. 17, 2021.

LORENZ, T. C. Polymerase chain reaction: basic protocol plus troubleshooting and optimization strategies. **JoVE (Journal of Visualized Experiments)**, n. 63, p. e3998, 2012.

LUCENA-AGUILAR, G. et al. DNA source selection for downstream applications based on DNA quality indicators analysis. **Biopreservation and Biobanking**, v. 14, n. 4, p. 264-270, 2016.

MA, X. et al. Long Noncoding RNA CCDC26 Promotes Thyroid Cancer Malignant Progression via miR-422a/EZH2/Sirt6 Axis. **OncoTargets and therapy**, v. 14, p. 3083, 2021.

MATTICK, J. S.; MAKUNIN, I. V. Non-coding RNA. **Human molecular genetics**, v. 15, n. suppl_1, p. R17-R29, 2006.

MIRZA, A. H. et al. Effects of GWAS-associated genetic variants on lncRNAs within IBD and T1D candidate loci. **PLoS One**, v. 9, n. 8, p. e105723, 2014.

ONCIU, M. Acute lymphoblastic leukemia. **Hematology/oncology clinics of North America**, v. 23, n. 4, p. 655-674, 2009.

- PALAZZO, A. F.; LEE, E. S. Non-coding RNA: what is functional and what is junk? **Frontiers in genetics**, v. 6, p. 2, 2015.
- PAPP, A. C. et al. Single nucleotide polymorphism genotyping using allele-specific PCR and fluorescence melting curves. **Biotechniques**, v. 34, n. 5, p. 1068-1072, 2003.
- PENG, W.; JIANG, A. Long noncoding RNA CCDC26 as a potential predictor biomarker contributes to tumorigenesis in pancreatic cancer. **Biomedicine & Pharmacotherapy**, v. 83, p. 712-717, 2016.
- POLICARPO, R. et al. From junk to function: LncRNAs in CNS health and disease. **Frontiers in molecular neuroscience**, p. 151, 2021.
- PUI, C.; JEHA, S. New therapeutic strategies for the treatment of acute lymphoblastic leukaemia. **Nature Reviews Drug Discovery**, v. 6, n. 2, p. 149-165, 2007.
- QUINN, J. J.; CHANG, H. Y. Unique features of long non-coding RNA biogenesis and function. **Nature Reviews Genetics**, v. 17, n. 1, p. 47-62, 2016.
- SHERRY, S. T. et al. dbSNP: the NCBI database of genetic variation. **Nucleic acids research**, v. 29, n. 1, p. 308-311, 2001.
- SILVA, F. A. Avaliação epidemiológica das leucemias linfoblásticas em crianças brasileiras e implicação de infecções na sua patogênese. **Rio de Janeiro (RJ): Instituto Nacional de Câncer**, 2009.
- SILVA, L. F. **Estudo de associação genômica ampla (GWAS) em *Coffea canephora***. 42 f. Dissertação, Universidade Federal de Viçosa, Minas Gerais, 2018.
- SIRUGO, G.; WILLIAMS, S. M.; TISHKOFF, S. A. The missing diversity in human genetic studies. **Cell**, v. 177, n. 1, p. 26-31, 2019.
- STATELLO, L. et al. Gene regulation by long non-coding RNAs and its biological functions. **Nature Reviews Molecular Cell Biology**, v. 22, n. 2, p. 96-118, 2021.
- SZCZEPAŃSKI, T.; VAN DER VELDEN, V. H.; VAN DONGEN, J. J. Classification systems for acute and chronic leukaemias. **Best Practice & Research Clinical Haematology**, v. 16, n. 4, p. 561-582, 2003.
- TANG, H. et al. Identification and function analysis of a five-long noncoding RNA prognostic signature for endometrial cancer patients. **DNA and cell biology**, v. 38, n. 12, p. 1480-1498, 2019.
- TOMASETTI, C.; LI, L.; VOGELSTEIN, B. Stem cell divisions, somatic mutations, cancer etiology, and cancer prevention. **Science**, v. 355, n. 6331, p. 1330-1334, 2017.
- TREVIÑO L. R. et al. Germline genomic variants associated with childhood acute lymphoblastic leukemia. **Nature Genetics**, v. 41, n. 9, p. 1001-1005, 2009.
- ULITSKY, I.; BARTEL, D. P. lincRNAs: genomics, evolution, and mechanisms. **Cell**, v. 154, n. 1, p. 26-46, 2013.

VIEIRA, D. P. **Técnicas de PCR: aplicações e padronização de reações**. 2011. Disponível em: <https://www.imt.usp.br/wp-content/uploads/proto/protocolos/aula2.pdf>. Acesso em: 23 abr. 2022.

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

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

WIEMELS, J. L. et al. 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, 2018.

WILSON, C.; KANHERE, A. 8q24. 21 locus: a paradigm to link non-coding RNAs, genome polymorphisms and cancer. **International Journal of Molecular Sciences**, v. 22, n. 3, p. 1094, 2021.

WU, T.; DU, Y. LncRNAs: From basic research to medical application. **International Journal of Biological Sciences**, v. 13, n. 3, p. 295–307, 2017.

WU, Y. et al. Long noncoding RNA HOTAIR involvement in cancer. **Tumor Biology**, v. 35, n. 10, p. 9531-9538, 2014.

WU, Z. et al. Regulation of lncRNA expression. **Cellular & molecular biology letters**, v. 19, n. 4, p. 561-575, 2014.

YAN, J. et al. Downregulation of lncRNA CCDC26 contributes to imatinib resistance in human gastrointestinal stromal tumors through IGF-1R upregulation. **Brazilian Journal of Medical and Biological Research**, v. 52, 2019.

YANG, H.; WANG, Z.; WANG, Z. 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**, v. 12, p. 12679, 2020.

YANG J. J. et al. Genome-wide association study identifies germline polymorphisms associated with relapse of childhood acute lymphoblastic leukemia. **Blood, The Journal of the American Society of Hematology**, v. 120, n. 20, p. 4197-4204, 2012

YAO, R.; WANG, Y.; CHEN, L. Cellular functions of long noncoding RNAs. **Nature cell biology**, v. 21, n. 5, p. 542-551, 2019.

ZHANG, F.; LUPSKI, J. R. Non-coding genetic variants in human disease. **Human molecular genetics**, v. 24, n. R1, p. R102-R110, 2015.

ZHANG, H. et al. Long non-coding RNA: a new player in cancer. **Journal of hematology & oncology**, v. 6, n. 1, p. 1-7, 2013.

ZHANG, P. et al. Non-coding RNAs and their integrated networks. **Journal of integrative bioinformatics**, v. 16, n. 3, 2019.

ZHANG, Y. et al. Long Noncoding RNA KCNMB2-AS1 Stabilized by N6- Methyladenosine Modification Promotes Cervical Cancer Growth Through Acting as a Competing Endogenous RNA. **Cell transplantation**, v. 29, p. 0963689720964382, 2020.

ZHANG, Y. et al. Role of HOTAIR in the diagnosis and prognosis of acute leukemia. **Oncology reports**, v. 36, n. 6, p. 3113-3122, 2016.

ZHOU, Y. et al. Activation of p53 by MEG3 non-coding RNA. **Journal of Biological Chemistry**, v. 282, n. 34, p. 24731-24742, 2007.

ZHOU, Y.; ZHANG, X.; KLIBANSKI, A. MEG3 noncoding RNA: a tumor suppressor. **Journal of molecular endocrinology**, v. 48, n. 3, p. R45-R53, 2012.

ZHU, J et al. KCNMB2-AS1 Promotes Bladder Cancer Progression Through Sponging miR-374a-3p to Upregulate S100A10. **Frontiers in Genetics**, p. 1023, 2021.