



Ministério da Saúde

Instituto Nacional de Câncer José Alencar Gomes da Silva

Coordenação de Pesquisa

Pós-Graduação em Oncologia - INCA

CAROLINE DE AGUIAR PIRES POUBEL

Impacto dos tipos de deleção do gene supressor tumoral *IKZF1* sobre a expressão de seus genes-alvo nas leucemias linfoblásticas de células precursoras B

Orientadora: Dra. Mariana Emerenciano Cavalcanti de Sá

Rio de Janeiro

2018



Ministério da Saúde
Instituto Nacional de Câncer José Alencar Gomes da Silva
Coordenação de Pesquisa

Pós-Graduação em Oncologia - INCA

CAROLINE DE AGUIAR PIRES POUBEL

Impacto dos tipos de deleção do gene supressor tumoral *IKZF1* sobre a expressão de seus genes-alvo nas leucemias linfoblásticas de células precursoras B

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Oncologia, Instituto Nacional de Câncer, como parte dos requisitos necessários à obtenção do título de Mestre em Oncologia.

Orientadora: Dra. Mariana Emerenciano Cavalcanti de Sá

Rio de Janeiro
2018

P872i Poubel, Caroline de Aguiar Pires.

Impacto dos tipos de deleção do gene supressor tumoral IKZF1 sobre a expressão de seus genes: alvo nas leucemias linfoblásticas de células precursoras B. / Caroline de Aguiar Pires Poubel. _ Rio de Janeiro: INCA, 2018.

xviii; 90 fl.; il.

Dissertação (Mestrado em Oncologia) – Instituto Nacional de Câncer José Alencar Gomes da Silva, 2018.

Orientador(a): Mariana Emerenciano Cavalcanti de Sá.

1. Leucemia Linfoblástica Aguda de Células B. 2. Deleção de Genes. 3. IKZF1. 4. Pediatria. I. Sá, Mariana Emerenciano Cavalcanti de (Orient.). II. Instituto Nacional de Câncer José Alencar Gomes da Silva. III. Título.

CDD 616.994



Ministério da Saúde
Instituto Nacional de Câncer José Alencar Gomes da Silva
Coordenação de Pesquisa

Pós-Graduação em Oncologia - INCA

CAROLINE DE AGUIAR PIRES POUBEL

Impacto dos tipos de deleção do gene supressor tumoral *IKZF1* sobre a expressão de seus genes-alvo nas leucemias linfoblásticas de células precursoras B

Orientadora: Dra. Mariana Emerenciano Cavalcanti de Sá

EXAMINADORES:

Dra. Raquel Ciivalschi Maia (Titular interna)

Dra. Bárbara da Costa Reis Monte-Mór (Titular interna)

Dr. Marcelo Ribeiro Alves (Titular externo)

Dra. Anke Bergmann (Suplente interno)

Dr. José Andrés Yunes (Suplente externo)

Rio de Janeiro

2018

Dedico este trabalho a todas as crianças e seus familiares, aos quais sou totalmente grata pela contribuição e confiança.

AGRADECIMENTOS

À **Deus** por ser meu guia e minha força. Agradeço pelo amor incondicional todos os dias e por cada oportunidade me dada. Seu minucioso cuidado me faz ser extremamente abençoada e cercada de pessoas maravilhosas que só me fizeram crescer nessa trajetória chamada vida.

A **Dra. Mariana Emerenciano** por ter me dado a oportunidade de iniciar e avançar na pesquisa. Me faltam palavras para agradecer todo o aprendizado, incentivo e paciência que recebi nesses anos. Deixo explícito aqui minha admiração por sua paixão em ensinar, não apenas com a finalidade de formar bons profissionais, mas também incríveis seres humanos.

Ao **Dr. Bruno Lopes** por contribuir e muito com a idealização deste trabalho. Agradeço toda a ajuda na bancada e fora dela, principalmente nos momentos de dúvidas que surgiram nesse caminho. Sendo assim, também agradeço à **Dra. Thayana Barbosa**, que vai muito além de uma colaboradora, é uma amiga que ganhei para vida. Não canso de dizer o quanto te admiro! Agradeço pela paciência em caminhar comigo desde a iniciação científica, por me ensinar tanto, por compreender meus erros, por me ajudar a consertá-los, pelos puxões de orelha, pelas risadas. Agradeço também a minha amiga **Letícia Hummel** pelo suporte e companhia na bancada. Muito obrigada.

À família **MECS**, isso mesmo, família! Passamos por muitas coisas juntos e juntos permanecemos. Agradeço pelos conselhos, por dividir experiências, aprendizado... crescer com vocês é uma honra para mim. Em especial, agradeço à **Ana Luíza** por ter se tornado uma amiga dentro e fora do laboratório. Muito obrigada por todas as conversas, desabafos, milhares de risadas, horas de almoço, paciência, ansiedade compartilhada, incentivo (ao extremo) e sinceridade. É muito bom saber que posso contar como você.

À **Dra. Maria do Socorro** e todo o **Programa de Hematologia-Oncologia Pediátrico** pela disponibilidade das amostras e por ceder o espaço físico para a realização dos experimentos. À **Dra. Mariana Boroni** por ter aberto as portas da bioinformática para essa leiga. Agradeço muito o aprendizado que tenho tido dia após dia. Saiba que foi muito importante para mim o seu incentivo desde o início

dessa aventura. Agradeço também todo o **grupo do Laboratório de Bioinformática e Biologia Computacional** por ter me recebido de braços abertos. A disponibilidade e paciência de vocês em tirar minhas dúvidas e ensinar tanta coisa nova, até me constrange. Vou levá-los para sempre comigo.

Aos meus pais, **Edilênio** e **Vania**, por comprarem meus sonhos e não medirem esforços para fazê-los acontecer. Até me emociono ao escrever, pois sei que não há palavras suficientes que demonstrarão minha gratidão e amor por vocês. Obrigada por cada palavra de ânimo e renúncia, por entenderem a distância e por me receberem com abraços apertados nas chegadas e “vai com Deus, que Ele te abençoe” a cada partida. O apoio de vocês me torna capaz de alçar voos como este. E dessa forma, também agradeço aos meus irmãos, **Nayane** e **Davi**, todos **familiares** e **amigos** que festejaram minha aprovação na seleção do Mestrado, e agora torcem para mais uma vitória. Obrigada pelas orações de vocês.

Ao meu melhor amigo e namorado, **Caíque**, por sempre ouvir meus desabafos e retribuir com conselhos, que, mais do que nunca, foram fundamentais. Você foi e é meu porto seguro. Obrigada por me apoiar incondicionalmente; por compartilhar meus sonhos e planos. Desculpe por minha ausência algumas vezes, mas agradeço muito a compreensão que você teve em todas, principalmente nesse momento pré-defesa. Amo você.

A todos da **Banca Examinadora** pela disponibilidade e aceite em contribuir com este trabalho. É uma honra ser avaliada por vocês. Muito obrigada!

ÍNDICE

LISTA DE TABELAS.....	x
LISTA DE FIGURAS.....	xi
LISTA DE ABREVIATURAS.....	xiii
RESUMO.....	xvii
ABSTRACT.....	xviii
1 INTRODUÇÃO	1
1.1 LEUCEMIA LINFOBLÁSTICA AGUDA DE CÉLULAS PRECURSORAS B (LLA-CPB)	2
1.2 PERFIL CITOGENÉTICO-MOLECULAR DA LLA-CPB.....	3
1.2.1 Biomarcadores de alto risco/prognóstico adverso.....	4
1.2.2 Biomarcadores de risco/prognóstico intermediário	6
1.2.3 Biomarcadores de baixo risco/bom prognóstico	7
1.2.4 Alterações adicionais recorrentes.....	8
1.3 O GENE IKZF1	9
1.4 TIPOS DE DELEÇÕES EM IKZF1 E CLASSES FUNCIONAIS	11
1.5 O FATOR DE TRANSCRIÇÃO IKAROS.....	13
1.6 GENES DIFERENCIALMENTE EXPRESSOS	14
1.6.1 RAG1	15
1.6.2 EBF1.....	15
1.6.3 FIGNL1	16
1.6.4 GRB10.....	17
1.7 DELEÇÃO EM IKZF1 E BTG1.....	17
2 JUSTIFICATIVA DO ESTUDO	19
3 OBJETIVOS.....	20
3.1 GERAL.....	20
3.2 ESPECÍFICOS	20
4 MATERIAL E MÉTODOS	21
4.1 CASUÍSTICA	21
4.2 DETERMINAÇÃO DO STATUS DE IKZF1.....	22
4.2.1 Extração de ácido desoxirribonucleico (DNA)	22
4.2.1.1 Ensaio comercial QIAmp DNA MiniBlood.....	22
4.2.1.2 Reagente TRIzol.....	23
4.2.2 Quantificação do DNA	24
4.2.3 Identificação de alterações no número de cópias por MLPA.....	24

4.2.3.1	<i>Ensaio comercial com sondas para IKZF1</i>	25
4.2.3.2	<i>Reação de MLPA</i>	25
4.2.4	Confirmação de deleções recorrentes por MP-PCR	26
4.3	QUANTIFICAÇÃO RELATIVA DO NÍVEL TRANSCRICIONAL DOS GENES DE INTERESSE	28
4.3.1	<i>Extração de ácido ribonucleico (RNA)</i>	28
4.3.2	<i>Quantificação e qualidade do RNA</i>	28
4.3.3	<i>Purificação de RNA</i>	29
4.3.4	<i>Síntese de DNA complementar (cDNA)</i>	30
4.3.4.1	<i>Reação de síntese de cDNA</i>	30
4.3.4.2	<i>RT-PCR para amplificação do gene constitutivo GAPDH</i>	30
4.3.5	Quantificação gênica por RT-qPCR	31
4.3.5.1	<i>Ensaio com sondas hidrolisáveis</i>	31
4.3.5.2	<i>Ensaio com agentes intercalantes de fita-dupla</i>	33
4.3.5.3	<i>Determinação do Delta Ct (ΔCt)</i>	33
4.4	VALIDAÇÃO DOS RESULTADOS	34
4.4.1	<i>Determinação do status de IKZF1</i>	35
4.4.2	<i>Expressão relativa de genes (RNA-seq)</i>	36
4.5	ANÁLISES ESTATÍSTICAS	36
5	RESULTADOS	38
5.1	CARACTERIZAÇÃO DOS PACIENTES	38
5.1.1	<i>Coorte de Descoberta</i>	38
5.1.2	<i>Coorte de Validação</i>	39
5.2	NÍVEIS DE EXPRESSÃO DOS GENES	40
5.3	ANÁLISES DE EXPRESSÃO VERSUS DADOS CLÍNICO-LABORATORIAIS	42
5.3.1	<i>Idade</i>	42
5.3.2	<i>Sexo</i>	43
5.3.3	<i>Leucometria</i>	44
5.3.4	<i>Classificação de risco prognóstico</i>	45
5.4	ANÁLISES DE EXPRESSÃO VERSUS STATUS DE IKZF1	45
5.5	ANÁLISES DE SOBREVIVÊNCIA	46
6	DISCUSSÃO	50
7	CONCLUSÃO	56
8	REFERÊNCIAS BIBLIOGRÁFICAS	57
9	ANEXOS	68

LISTA DE TABELAS

Tabela 4.1. Sequência dos oligonucleotídeos de <i>IKZF1</i> para as reações do MP-PCR.....	27
Tabela 4.2. Sequência dos oligonucleotídeos do gene <i>GAPDH</i> para reação de RT-PCR.....	30
Tabela 4.3. Sequência dos oligonucleotídeos usados na reação de RT-qPCR..	33
Tabela 5.1. Características dos pacientes com LLA-CPB.....	39

LISTA DE FIGURAS

Figura 1.1. Conjunto de causas da leucemia linfoblástica aguda pediátrica.....	2
Figura 1.2. Alterações genéticas da LLA-CPB.....	3
Figura 1.3. Frequência das alterações primárias da LLA-CPB pediátrica.....	4
Figura 1.4. Taxas de sobrevida livre de doença e incidência acumulativa de recaída.....	8
Figura 1.5. Estrutura de <i>IKZF1</i> e isoformas de Ikaros.....	9
Figura 1.6. Frequência das alterações em <i>IKZF1</i>	11
Figura 1.7. Taxas de incidência acumulativa de recaída e sobrevida livre de doença em pacientes com e sem alterações em <i>IKZF1</i>	13
Figura 1.8. Expressão gênica em pacientes LLA-CPB <i>IKZF1</i> selvagem e deletado.....	14
Figura 1.9. Taxas de incidência acumulativa de recaída e sobrevida livre de doença em pacientes com alterações em <i>IKZF1</i> e <i>BTG1</i>	18
Figura 4.1. Desenho esquemático do estudo.....	21
Figura 4.2. Desenho esquemático da técnica de MLPA.....	25
Figura 4.3. Eletroforese em gel de agarose para a reação de RT-PCR.....	31
Figura 4.4. Ciclagem das reações para RT-qPCR.....	32
Figura 4.5. Exemplo de reação processada pelo programa QuantStudio Real-Time PCR.....	34
Figura 5.1. Níveis de expressão dos genes nas coortes de descoberta e validação.....	40
Figura 5.2. Correlação entre a expressão do gene <i>IKZF1</i> e os genes-alvo nas coortes de descoberta e validação.....	41
Figura 5.3. Perfil de expressão gênica entre os pacientes, de acordo com suas características.....	42
Figura 5.4. Análises de expressão entre os pacientes <i>IKZF1</i> selvagem nas coortes de descoberta e validação, de acordo com a idade.....	43
Figura 5.5. Análises de expressão entre os pacientes <i>IKZF1</i> selvagem nas coortes de descoberta e validação, de acordo com o sexo.....	43
Figura 5.6. Análises de expressão entre os pacientes <i>IKZF1</i> selvagem nas coortes de descoberta e validação, de acordo com a leucometria.....	44

Figura 5.7. Análises de expressão entre os pacientes <i>IKZF1</i> selvagem nas coortes de descoberta e validação, de acordo com a classificação de risco prognóstico.....	44
Figura 5.8. Análises de expressão entre os pacientes <i>IKZF1</i> selvagem e deletado nas coortes de descoberta e validação.....	45
Figura 5.9. Análises de expressão de acordo com o status de <i>IKZF1</i> nas coortes de descoberta e validação.....	46
Figura 5.10. Análise de sobrevida global e livre de eventos dos pacientes de acordo com o risco definido pelos critérios do NCI.....	47
Figura 5.11. Análise de sobrevida global e livre de eventos dos pacientes de acordo com o status de <i>IKZF1</i>.....	47
Figura 5.12. Análise de sobrevida global e livre de eventos dos pacientes de acordo com a expressão gênica.....	49

LISTA DE ABREVIATURAS

<i>.gtf</i>	<i>General Transfer Format</i>
<i>.tsv</i>	<i>Tab Separated Values</i>
<i>.txt</i>	<i>Text Extension</i>
<i>.xlsx</i>	<i>Excel Extension</i>
Δ dn	Deleção do tipo dominante-negativo
Δ haplo	Deleção do tipo haploinsuficiente
$^{\circ}$ C	Escala em grau Celsius
μ L	Microlitro
μ M	Micromolar
ABL1	<i>Abelson Tyrosine-Protein Kinase 1</i>
AFF1 (AF4)	<i>AF4/FMR2 Family Member 1</i>
ATP	<i>Adenosine triphosphate</i>
BCR	<i>Breakpoint Cluster Region</i>
BTG1	<i>B-Cell Translocation Gene 1</i>
CDKN2A	<i>Cyclin Dependent Kinase Inhibitor 2A</i>
CDKN2B	<i>Cyclin Dependent Kinase Inhibitor 2B</i>
cDNA	DNA complementar
céls	células
CEP	Comitê de ética e pesquisa
CRLF2	<i>Cytokine Receptor Like Factor 2</i>
CSF2RA	<i>Colony Stimulating Factor 2 Receptor Alpha</i>
Ct	<i>Cycle Threshold</i>
DD	Domínio de dimerização
DDC	<i>Dopa Decarboxylase</i>
DEPC	<i>Diethylpyrocarbonate</i>
DLD	Domínio de ligação ao DNA
DNA	<i>Deoxyribonucleic Acid</i>
dNTP	<i>Deoxynucleotide Triphosphates</i>
Dr.	Doutor
Dra.	Doutora
EBF1	<i>Early B-Cell Factor 1</i>

EDTA	<i>Ethylenediaminetetraacetic Acid</i>
EGIL	<i>European Group for the Immunological Characterization of Leukemias</i>
ETV6	<i>ETS Variant 6</i>
FIGNL1	<i>Fidgetin Like 1</i>
FISH	<i>Fluorescence in situ hybridization</i>
GAPDH	<i>Glyceraldehyde-3-Phosphate Dehydrogenase</i>
GRB10	<i>Growth Factor Receptor Bound Protein 10</i>
Hi-Di	<i>Highly deionized</i>
HLF	<i>Hepatic Leukemia Factor</i>
iAMP21	<i>Intrachromosomal amplification of chromosome 21</i>
IGH	<i>Immunoglobulin Heavy Locus</i>
IGHD	<i>Immunoglobulin Heavy Constant Delta</i>
IKZF1	<i>IKAROS Family Zinc Finger 1</i>
IKZF2	<i>IKAROS Family Zinc Finger 2</i>
IKZF3	<i>IKAROS Family Zinc Finger 3</i>
IL3RA	<i>Interleukin 3 Receptor Subunit Alpha</i>
INCA	Instituto Nacional de Câncer
KIAA0284	<i>Centrosomal Protein 170B</i>
KMT2A (MLL)	<i>Lysine Methyltransferase 2A (Mixed-Lineage Leukemia)</i>
L	Litro
LBBC	Laboratório de Bioinformática e Ciência Computacional
LD-PCR	<i>Long-distance Polymerase Chain Reaction</i>
LLA	Leucemia linfoblástica aguda
LLA-CPB	Leucemia linfoblástica aguda de células precursoras B
LMA	Leucemia mieloide aguda
MAF	<i>Mutation Annotation Format</i>
MgCl ₂	Cloreto de magnésio
MgSO ₄	Sulfato de magnésio
min	Minutos
MIQE	<i>Minimum Information for Publication of Quantitative Real-Time PCR Experiments</i>
MIR31	<i>MicroRNA 31</i>

mL	Mililitro
<i>MLLT1 (ENL)</i>	<i>Mixed-Lineage Leukemia; Translocated To, 1</i>
<i>MLLT10 (AF10)</i>	<i>Mixed-Lineage Leukemia; Translocated To, 10</i>
<i>MLLT3 (AF9)</i>	<i>Mixed-Lineage Leukemia; Translocated To, 3</i>
<i>MLLT4 (AF6)</i>	<i>Mixed-Lineage Leukemia; Translocated To, 4</i>
MLPA	<i>Multiplex Ligation-dependent Probe Amplification</i>
mM	Milimolar
MO	Medula óssea
MP-F	Mistura de <i>primer forward</i>
MP-PCR	<i>Multiplex Polymerase Chain Reaction</i>
MP-R1	Mistura de <i>primer reverse</i> (ímpar)
MP-R2	Mistura de <i>primer reverse</i> (par)
<i>MTA1</i>	<i>Metastasis Associated 1</i>
NCI	<i>National Cancer Institute</i>
nm	Nanômetro
ng	Nanograma
NuRD	<i>Nucleosome Remodeling Histone Deacetylase</i>
<i>P2RY8</i>	<i>Purinergic Receptor P2Y8</i>
PAR1	Pseudoautosomal Region 1
<i>PAX5</i>	<i>Paired Box 5</i>
pb	pares de base
<i>PBX1</i>	<i>PBX Homeobox 1</i>
PCR	<i>Polymerase Chain Reaction</i>
PHOP	Programa de Hematologia-Oncologia Pediátrico
qPCR	<i>Quantitative Polymerase Chain Reaction</i>
<i>RAG1</i>	<i>Recombination Activating 1</i>
<i>RB1</i>	<i>Retinoblastoma 1</i>
RNA	<i>Ribonucleic Acid</i>
RNA-seq	<i>RNA sequencing</i>
rpm	rotações por minuto
RT	<i>Reverse Transcriptase</i>
RT-PCR	<i>Reverse Transcription Polymerase Chain Reaction</i>
RT-qPCR	<i>Reverse Transcription Quantitative Polymerase Chain Reaction</i>

<i>RUNX1</i>	<i>Runt Related Transcription Factor 1</i>
seg	Segundos
SG	Sobrevida global
<i>SHOX</i>	<i>Short Stature Homeobox</i>
SLE	Sobrevida livre de eventos
SNC	Sistema Nervoso Central
SP	Sangue periférico
t	translocação
Taq	<i>Thermus aquaticus</i>
TARGET	<i>Therapeutically Applicable Research to Generate Effective Treatments</i>
<i>TCF3</i>	<i>Transcription Factor 3</i>
U	Unidade
UV	Ultravioleta
WGS	<i>Whole Genome Sequencing</i>
<i>ZBPB</i>	<i>Zona Pellucida Binding Protein</i>



RESUMO

As deleções em *IKZF1* ($\Delta IKZF1$) ocorrem em aproximadamente 15% dos casos pediátricos com leucemia linfoblástica aguda de células precursoras B (LLA-CPB) e estão associadas com aumento nas taxas de recaída em diversos protocolos. A consequência funcional das $\Delta IKZF1$ varia de acordo com os éxons envolvidos, podendo resultar em efeito dominante-negativo (deleção dos éxons 4 a 6) ou haploinsuficiência (deleções envolvendo os éxons 1, 2 e/ou 8). Visto que Ikaros (codificada por *IKZF1*) regula a expressão gênica através do remodelamento de cromatina, nós hipotetizamos que $\Delta IKZF1$ podem modificar a expressão de genes envolvidos na linfopoiese B (*RAG1*, *EBF1* e *BTG1*), no reparo de DNA (*FIGNL1*) e na interação com receptores de tirosina quinase (*GRB10*). Para avaliar o impacto dos tipos de $\Delta IKZF1$ na expressão destes genes, analisamos 72 pacientes diagnosticados com LLA-CPB antes dos 16 anos. Estes foram agrupados de acordo com o status de *IKZF1* obtido através das técnicas de MLPA e PCR multiplex. Os níveis de expressão gênica de *IKZF1* e os demais genes foram obtidos por PCR quantitativo em tempo real. Em seguida, as análises foram realizadas de forma comparativa com os dados clínico-laboratoriais e status de *IKZF1* usando o teste estatístico Mann-Whitney. A fim de validar nossos achados, também realizamos análises *in silico* com dados de pacientes pediátricos disponibilizados pelo projeto TARGET. Os dados de *whole genome sequencing* foram usados para determinar o status de *IKZF1* e RNA-seq para expressão dos genes. As análises de sobrevida livre de evento (SLE) e global (SG) foram realizadas pelo método de Kaplan-Meier. Em relação às características clínico-laboratoriais dos pacientes, observamos que pacientes com idade <120 meses apresentam uma tendência de expressão aumentada de *IKZF1*, *RAG1* e *BTG1*, e baixa para *GRB10*, quando comparados àqueles com idade ≥ 120 meses. Já o gene *RAG1* apresentou uma mediana de expressão maior no grupo de baixo risco quando comparado tanto com o grupo de risco intermediário quanto com o de alto risco. As análises de expressão demonstraram que o gene *EBF1* apresenta uma tendência de níveis reduzidos diante da Δdn quando comparado com os outros grupos. Já o gene *GRB10* tem níveis aumentado nos pacientes com $\Delta haplo$. Além disso, foi possível observar que pacientes com baixa expressão de *EBF1* tiveram maior tempo de SG quando comparado com os de alta expressão, enquanto que indivíduos considerados com baixa expressão de *GRB10* apresentaram pior SG e SLE em relação aqueles com alta expressão. Portanto, nossos resultados demonstram que os diferentes status de *IKZF1* estão associados a diferentes níveis de expressão de genes-alvo.

Palavras-chave: LLA-CPB; *IKZF1*; deleção.



Ministério da Saúde

Instituto Nacional de Câncer José Alencar Gomes da Silva

Coordenação de Pesquisa

ABSTRACT

IKZF1 deletions ($\Delta IKZF1$) occur in 15% of pediatric B cell precursor acute lymphoblastic leukemia (BCP-ALL) and are associated with an increased relapse rate in several protocols. The functional consequence of $\Delta IKZF1$ depends on the exons involved, and may result in dominant-negative effect (deletions of exons 4, 5 and/or 6) or haploinsufficiency (deletions involving the exons 1, 2 and/or 8). Because Ikaros (encoded by *IKZF1*) regulates gene expression via chromatin remodeling, we hypothesize that $\Delta IKZF1$ may change the expression of genes that are involved in B lymphopoiesis (*RAG1*, *EBF1* and *BTG1*), DNA repair (*FIGL1*) and interaction with receptor tyrosine kinases (*GRB10*). To assess the impact of $\Delta IKZF1$ types on the gene expression, we analysed 72 children diagnosed with BCP-ALL before 16 years-old. They were grouped according to the status of *IKZF1* as determined by MLPA assay and multiplex PCR. The expression level analysis of *IKZF1* and other genes was assessed by quantitative PCR. Then the analyzes were performed in a comparative manner with the clinical-laboratory data and *IKZF1* status using the Mann-Whitney test. In order to validate our findings, we also performed *in silico* analysis with pediatric patient data provided by the TARGET project. For this validation cohort, whole genome sequencing were used to determine the *IKZF1* status and RNA-seq for gene expression. The overall (OS) and event-free survival (EFS) were performed by the Kaplan-Meier method. We observed that patients younger than 120 months showed an increased expression of *IKZF1*, *RAG1* and *BTG1*, and decreased for *GRB10*, when compared to those older than 120 months. The *RAG1* gene was also found with higher expression within the good risk group when compared to both intermediate and high risk groups. Expression analysis demonstrated that *EBF1* have a tendency of decreased levels in Δdn when compared to the other groups. Increased levels of *GRB10* gene were found in patients with *IKZF1* $\Delta haplo$. In addition, it was observed that patients with low expression of *EBF1* had a better OS when compared to those with high expression, whereas individuals considered with low expression of *GRB10* presented worse OS and EFS in relation to those with high expression. Therefore, our results demonstrate that the different *IKZF1* status are associated with different expression levels of target genes.

Key words: BCP-ALL; *IKZF1*; deletion.

1 INTRODUÇÃO

O câncer corresponde a um conjunto de neoplasias malignas caracterizadas pela proliferação descontrolada de células anormais em diferentes tecidos ou órgãos. De forma geral, essa doença é diagnosticada com mais frequência entre pacientes idosos, entretanto, alguns subtipos podem ser encontrados em maiores números entre os pacientes pediátricos.

Em 2017, os Estados Unidos estimaram a ocorrência de 10.270 novos casos de câncer entre crianças com idade até 14 anos. Já no Brasil, este número foi de 12.600 entre crianças e adolescentes com idade entre 1 e 19 anos. Embora o tratamento para estes indivíduos tenha alcançado grandes progressos com taxa de cura alcançando 70-80%, o câncer continua sendo a principal causa de morte por doença entre os casos pediátricos (INCA; NIH).

A leucemia linfoblástica aguda (LLA) é a neoplasia pediátrica diagnosticada com mais frequência (cerca de 25% dos casos), seguida por tumor no sistema nervoso central (SNC) e linfomas (INCA; NIH; de Camargo et al, 2010). A LLA é uma neoplasia hematológica maligna heterogênea, caracterizada pela rápida expansão e acúmulo de células progenitoras linfoides, contendo múltiplas lesões genéticas, na medula óssea (MO). Este quadro afeta a produção das células sanguíneas normais, levando os pacientes a apresentarem sintomas como: palidez, cansaço e sonolência, devido ao quadro de anemia; infecções constantes decorrentes de imunossupressão; hematomas, petéquias e sangramentos prolongados por causa da trombocitopenia; além de outras manifestações clínicas causadas pela infiltração das células leucêmicas nos tecidos (Pui, 2012).

O processo de desenvolvimento da LLA pode ser iniciado por causas multifatoriais desconhecidas ou ainda pouco compreendidas como exposição a fatores exógenos ou endógenos e susceptibilidade genética (Inaba et al, 2013) (Figura 1.1). Dependendo da linhagem celular comprometida, a LLA ainda pode ser classificada em células precursoras B ou T. Neste trabalho, especificamente, serão discutidos aspectos da doença derivada de células precursoras B e seu comportamento em pacientes pediátricos.

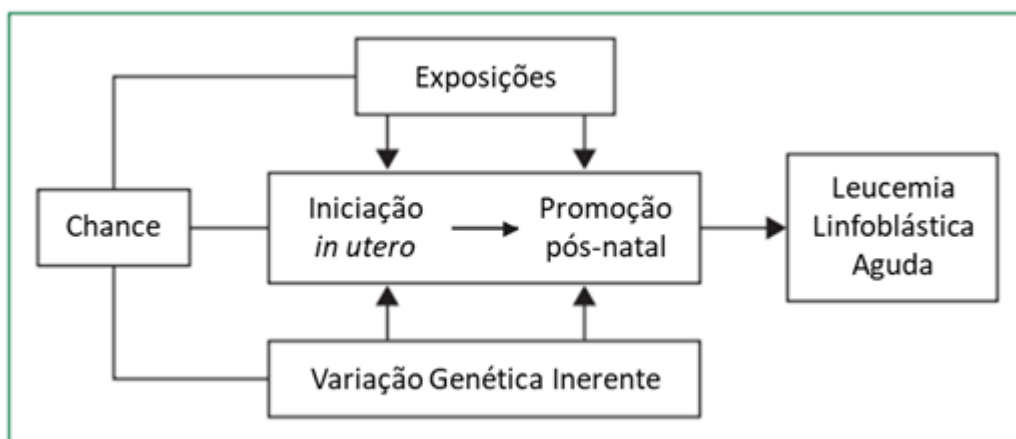


Figura 1.1. Conjunto de causas da leucemia linfoblástica aguda pediátrica. Exposições exógenas e endógenas, como infecção e inflamação, variação genética inerente e chance exercem um papel na história natural da leucemia linfoblástica aguda pediátrica (Modificado de Inaba et al, 2013).

1.1 LEUCEMIA LINFOBLÁSTICA AGUDA DE CÉLULAS PRECURSORAS B (LLA-CPB)

O diagnóstico da LLA-CPB pode ser estabelecido principalmente pela análise imunofenotípica dos blastos infiltrados na MO ou sangue periférico (SP), através da presença de marcadores de células B como CD79a e CD22 citoplasmático, e CD19. Cerca de 85% dos casos de LLA pediátrica são LLA-CPB, sendo considerada a neoplasia mais comum em crianças. O pico de incidência ocorre entre as idades de 2 e 5 anos, indo ao encontro de teorias baseadas no envolvimento de infecções no desenvolvimento da LLA (Kinlen, 2004; Greaves, 2006). Além disso, observa-se uma maior frequência de casos entre os indivíduos do sexo masculino do que feminino (1,4:1), ainda que não exista alguma justificativa plausível (Moorman, 2016).

A fim de que um protocolo adequado de tratamento seja designado aos pacientes, em um primeiro momento, estes são estratificados de acordo com o risco, seguindo os critérios do Instituto Nacional de Câncer (NCI: do inglês *National Cancer Institute*). Indivíduos entre 1 e 10 anos de idade, com leucometria inferior a 50×10^9 células/L são classificados como de risco padrão; enquanto pacientes com idade inferior a 1 ano ou superior a 10 anos, com leucometria superior a 50×10^9 células/L são classificados como de alto risco (Smith et al, 1996).

Com o passar das décadas, a LLA-CPB tem sido caracterizada por alterações cromossômicas estruturais e numéricas frequentemente associadas a estratificação

de risco prognóstico. Apesar deste perfil citogenético-molecular ser extremamente útil para o refinamento do prognóstico, na prática clínica ainda ocorrem muitas recaídas e falhas de tratamento mesmo nos grupos definidos como de bom risco prognóstico (Schultz et al, 2007; Woo et al, 2014). Entretanto, estudos demonstram que com abordagens mais modernas, as taxas de sobrevida livre de evento (SLE) podem atingir cerca de 90% (Vora et al, 2013; Vora et al, 2014). Isto é, estratificar de forma precisa cada paciente de acordo com o risco de recaída e submetê-los a tratamentos ditos personalizados.

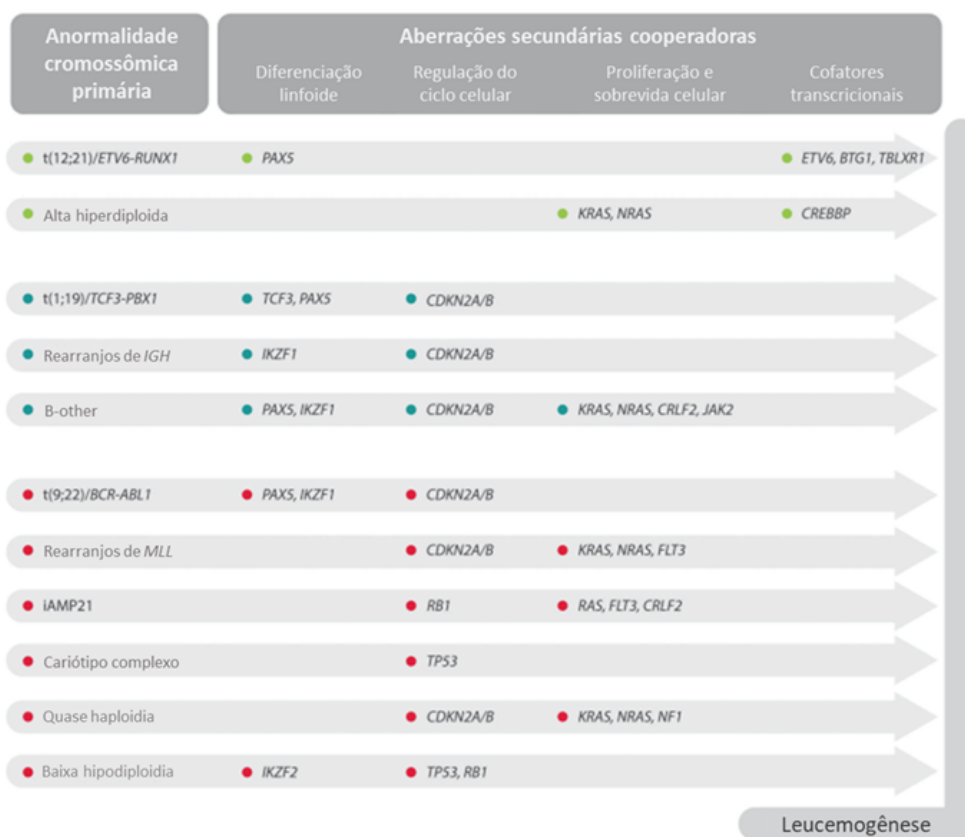


Figura 1.2. Alterações genéticas da LLA-CPB. Para que ocorra o processo de leucemogênese, são necessários eventos primários e secundários. Nesta figura, as cores indicam o risco prognóstico associado a cada grupo de alterações. Verde: bom prognóstico. Azul: intermediário. Vermelho: prognóstico adverso (Modificado de Moorman, 2016).

1.2 PERFIL CITOGÉNÉTICO-MOLECULAR DA LLA-CPB

Aproximadamente 75% dos casos de LLA-CPB apresentam alterações cromossômicas, denominadas primárias, juntamente com aberrações genéticas adicionais, sendo responsáveis pela ativação e manutenção dos clones leucêmicos

(Figura 1.2). Atualmente, a Organização Mundial de Saúde (OMS) utiliza as principais alterações primárias para classificar as LLAs-CPB (Arber et al, 2016). Assim como estas, outras alterações são comumente usadas como marcadores de prognóstico, estratificando os pacientes em três grandes grupos de acordo com o risco: i) alto, ii) intermediário, e iii) baixo (Figura 1.3) (Moorman, 2016). Em sua maioria, são rastreadas através da combinação das técnicas de hibridização *in situ* por fluorescência (FISH) e reação em cadeia da polimerase via transcriptase reversa (RT-PCR).

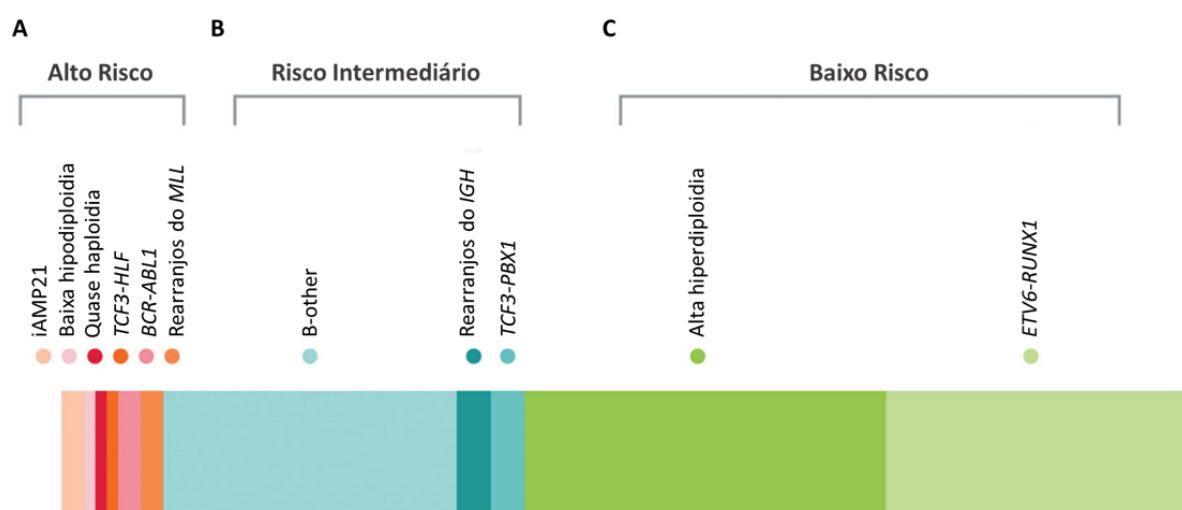


Figura 1.3. Frequência das alterações primárias da LLA-CPB pediátrica. (A) Para alto risco, seis alterações foram associadas e estas estão presentes na minoria dos casos pediátricos. (B) Enquanto isso, cerca de 30% dos pacientes são classificados como B-other, apresentam rearranjos do *IGH* ou fusão gênica *TCF3-PBX1*. Para estes, a associação de risco é determinada como intermediária. (C) As duas anomalias associadas ao baixo risco representam cerca de 60% das LLAs-CPB (Modificado de Moorman, 2016).

1.2.1 Biomarcadores de alto risco/prognóstico adverso

São consideradas anormalidades de alto risco rearranjos envolvendo o gene *KMT2A* (*MLL*), as translocações $t(9;22)/BCR-ABL1$, $t(17;19)/TCF3-HLF$, além da amplificação intracromossomal no cromossomo 21 (*iAMP21*) e de alterações numéricas como ‘quase haploidia’ (menos que 30 cromossomos) e ‘baixa hipodiploidia’ (entre 30 e 39 cromossomos) (Figura 1.3A).

O gene *KMT2A* pode ser encontrado rearranjado com diversos parceiros, sendo os principais: *AFF1* (*AF4*), *MLLT1* (*ENL*), *MLLT4* (*AF6*), *MLLT3* (*AF9*) e

MLLT10 (AF10) (Meyer et al, 2018). As principais consequências dessas alterações interferem tanto na capacidade de auto-proliferação quanto no crescimento descontrolado das células tronco. De forma geral, essas alterações acometem cerca de 9% dos casos de LLA-CPB pediátrica, entretanto, ao analisar a leucemia de lactentes (pacientes com idade inferior a 12 meses) essa frequência aumenta para mais de 80%. As baixas taxas de sobrevida e altas probabilidades de recaída fazem com que esses indivíduos sejam tratados de forma mais agressiva (Pieters et al, 2007).

A fusão gênica *BCR-ABL1* decorre da translocação $t(9;22)(q34;q11)$ e acomete cerca de 3-5% dos casos. Essa anomalia é considerada como alvo terapêutico para inibidor de tirosina quinase, isso porque seu produto proteico é responsável por aumentar de forma constitutiva a atividade da enzima tirosina quinase, resultando na ativação de diversas vias de sinalização, assim como no aumento da proliferação e desregulação da diferenciação celular (Lugo et al, 1990; Bose et al, 1998). Além disso, está associada a pacientes de idade avançada, com alta contagem leucocitária e infiltração do SNC (Crist et al, 1990).

Através dos estudos de expressão gênica de larga escala foi possível observar que alguns casos de LLA-CPB, apesar da ausência de qualquer alteração primária conhecida, apresentavam perfil de expressão gênica semelhante ao observado em pacientes com *BCR-ABL1* (den Boer et al, 2009). Estes casos geralmente recebem tratamento inadequado, visto que muitas vezes são classificados como de risco padrão pelos critérios do NCI (idade inferior a 10 anos e contagem leucocitária superior a 25×10^9 céls/L). Entretanto, assim como os pacientes *BCR-ABL1* positivos, estes casos estão associados a um prognóstico adverso. Denominados como *BCR-ABL1-like*, devido à dificuldade de identificação, muitos pacientes ainda são classificados no subgrupo *B-other*, que será discutido com mais detalhes no tópico seguinte.

A $t(17;19)(q22;p13)$ é responsável pela formação da fusão gênica *TCF3-HLF*, que acomete menos que 1% dos casos. Esta anomalia está associada a um pior prognóstico, assim como a pacientes mais velhos. O produto proteico codificado por essa fusão regula de forma aberrante genes responsáveis pela programação de morte celular em progenitores linfoides (de Boer et al, 2011).

A alteração iAMP21 caracteriza-se pela amplificação de no mínimo três cópias de uma grande e variável região do cromossomo 21. Ocorre em aproximadamente 2% dos casos de LLA-CPB, sendo estes pacientes mais velhos e com baixa contagem leucocitária. Sua identificação é importante, pois quando os indivíduos são submetidos a protocolos de risco padrão, há alta taxa de recaída, o que não é observado naqueles tratados pelo protocolo de alto risco (Harewood et al, 2003; Moorman et al, 2007; Harrison et al, 2014).

Por fim a hipodiploidia pode acometer 5-8% dos casos, compondo um pequeno, porém heterogêneo grupo. Pacientes classificados com 'quase haploidia' frequentemente tem os cromossomos 21, X, Y, 14 e/ou 18 afetados e estão associados à leucometria baixa. Da mesma forma são os indivíduos com 'baixa hipodiploidia', entretanto, é possível observar alteração também dos cromossomos 1, 11, 19, entre outros e idade mais avançada. Ambos os grupos demonstraram associação a uma pior resposta terapêutica, com SLE em 3 anos de 29% (com intervalo de confiança entre 10-53%) (Harrison et al, 2004).

1.2.2 Biomarcadores de risco/prognóstico intermediário

Entre as alterações associadas ao risco intermediário estão os rearranjos envolvendo o gene *IGH*, a $t(1;19)/TCF3-PBX1$ e o grupo denominado *B-other* (Figura 1.3B).

Os rearranjos envolvendo o gene *IGH* são de frequência baixa, ocorrendo entre 2-5% de todos os casos de LLA-CPB pediátrica. Muitos parceiros envolvidos nos rearranjos já foram identificados, entretanto, o gene *CRLF2* é o mais comum deles. A justaposição do gene parceiro na região de *enhancer* de *IGH*, devido a translocação, resulta na sua superexpressão (Russell et al, 2009; Dyer et al, 2010).

A fusão gênica *TCF3-PBX1* é resultante da translocação $t(1;19)(q23;p13)$ e está presente em até 6% dos casos. Tanto o gene *TCF3* quanto o *PBX1* codificam proteínas importantes para o desenvolvimento linfóide. Dessa forma, observa-se que o produto proteico gerado pela justaposição destes genes atua reprimindo a diferenciação celular. Estes pacientes são classificados como sendo de risco intermediário, pois o desfecho clínico depende da modalidade terapêutica utilizada (Harrison, 2001).

Pacientes sem nenhuma alteração primária conhecida são subcategorizados em um grupo chamado *B-other*. Por se tratar de um grupo altamente heterogêneo, inúmeros esforços foram realizados na última década para estratificar de forma adequada e precisa estes pacientes. Atualmente, estes indivíduos são separados de acordo com duas estratégias: i) presença de alterações secundárias preditoras de prognóstico (vide tópico 1.2.4); e ii) análise do perfil de expressão gênica para definir *BCR-ABL1-like* (Moorman, 2016).

1.2.3 Biomarcadores de baixo risco/bom prognóstico

Tanto a alta hiperdiploidia (entre 51-65 cromossomos) quanto a $t(12;21)/ETV6-RUNX1$ estão associadas ao baixo risco (Figura 1.3C). Juntas, estas alterações podem acometer cerca de 60% dos casos com LLA-CPB (Moorman, 2016).

A alteração numérica conhecida por alta hiperdiploidia é decorrente do desenvolvimento de clones celulares com múltiplas cópias de um ou mais cromossomos, como o X, 4, 6, 10, 14, 17, 18 e 21 (Moorman et al, 1996; Paulsson et al, 2010). Além disso, é considerada a anomalia mais frequente na LLA-CPB pediátrica acometendo cerca de 28% dos casos, com taxas de SLE e sobrevida global (SG) de até 90% (Moorman et al, 2003; Dastugue et al, 2013; Ghazavi et al, 2015). Entretanto, ainda há um grupo de pacientes associados a eventos de recaída (Pui et al, 1988; Heerema et al, 2000; Chen et al, 2015).

A translocação mais comum na LLA-CPB, acometendo cerca de 25% dos casos, é $t(12;21)(p13;q22)$ que resulta na fusão gênica *ETV6-RUNX1*. A proteína resultante dessa fusão é capaz de alterar a expressão normal de genes regulados tanto por *ETV6* quanto por *RUNX1*, contribuindo para a iniciação da leucemogênese visto que ambos são fatores de transcrição necessários para a hematopoese normal (Fuka et al, 2011; Linka et al, 2013). Contudo, assim como em outras alterações, são necessários eventos genéticos adicionais para que ocorra a leucemia (Wiemels et al, 1999; Mullighan et al, 2007).

1.2.4 Alterações adicionais recorrentes

As anormalidades adicionais, também conhecidas como eventos secundários, são frequentemente mutações pontuais e/ou alterações no número de cópias, como deleções e ampliações. Os genes frequentemente alterados estão envolvidos em vias de diferenciação linfóide (*PAX5*, *IKZF1*, *TCF3*), regulação do ciclo celular (*CDKN2A/B*, *RB1*, *TP53*), proliferação e sobrevivência celular (*CRLF2*, *JAK2*, *RAS*, *FLT3*), ou são cofatores de transcrição (*ETV6*, *BTG1*, *CREBBP*) (Figura 1.2). Técnicas como PCR seguido de sequenciamento e *multiplex ligation-dependent probe amplification* (MLPA) são utilizadas para o rastreamento destas alterações. Algumas destas alterações têm funções importantes na leucemogênese, na estratificação de risco e/ou na definição terapêutica (Ghazavi et al, 2015).

Neste estudo, focaremos nas deleções encontradas no gene *IKZF1*. Diversos estudos demonstraram uma forte correlação entre a presença dessa alteração com o aumento da taxa de recaída e diminuição da SLE dos pacientes (Figura 1.4) (Martinelli et al, 2009; Mullighan et al, 2009; Kuiper et al, 2010; Chen et al, 2012; Dorge et al, 2013; Boer et al, 2016). As deleções em *IKZF1* podem atingir 15% ou mais dos casos de LLA-CPB pediátrica, sendo muito prevalentes em dois subgrupos: *BCR-ABL1*-positivo, cerca de 70%; e *BCR-ABL1*-like, aproximadamente 40% dos indivíduos (Mullighan, 2012). Estes resultados indicam que a detecção dessas alterações ao diagnóstico pode ser importante na identificação de pacientes com elevado risco de falha no tratamento.

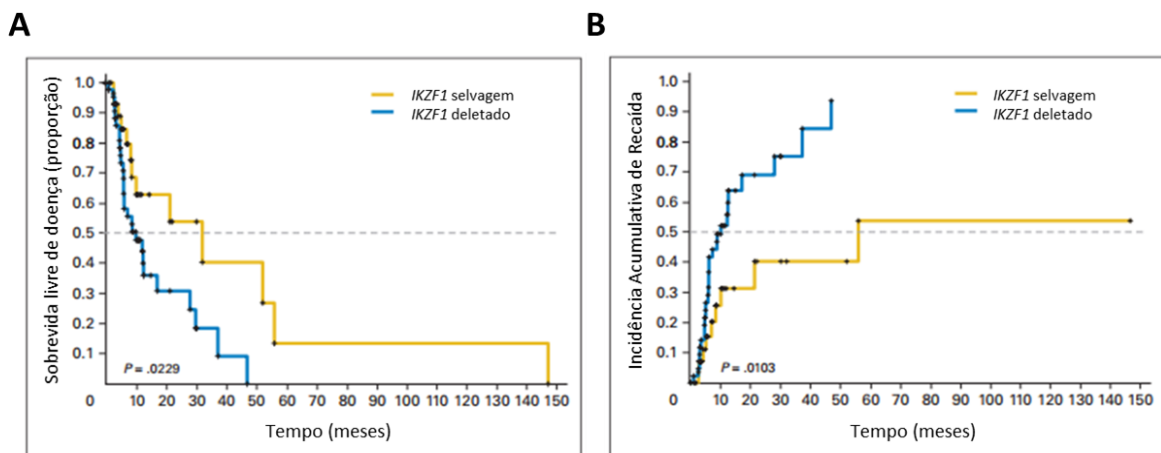


Figura 1.4. Taxas de sobrevida livre de doença e incidência cumulativa de recaída. Análise representativa das taxas de (A) sobrevida e (B) recaída, de acordo com a presença (linha azul) ou não (linha amarela) de deleções no gene *IKZF1* (Modificado de Martinelli et al, 2009).

1.3 O GENE *IKZF1*

O gene *IKZF1* está localizado no braço curto do cromossomo 7, na região 1, banda 2, sub-banda 2 (7p12.2) e é composto por 8 éxons, sendo apenas as sequências entre o 2 e o 8 responsáveis pela codificação do fator de transcrição Ikaros (Figura 1.5A). Contudo, acredita-se que o éxon 1 seja necessário para regulação da transcrição (Iacobucci et al, 2012; Yoshida et al, 2013; Olsson & Johansson, 2015). Ikaros é responsável pela regulação do desenvolvimento das linhagens linfoides B e T, assim como rearranjo da imunoglobulina.

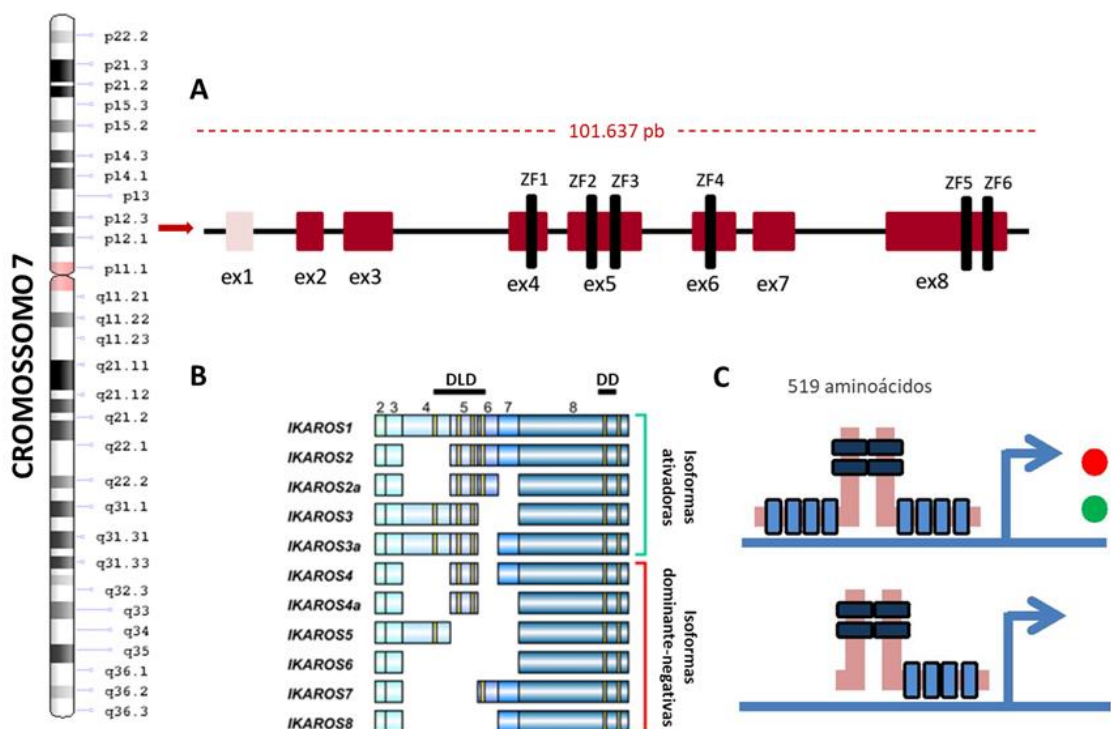


Figura 1.5. Estrutura de *IKZF1* e isoformas de Ikaros. (A) Representação esquemática do gene *IKZF1*, composto por 8 éxons (retângulos) e 6 dedos de zinco (ZF) representados pelas barras pretas. O fator de transcrição Ikaros é codificado pelos éxons de 2 a 8 (retângulos cor vinho). (B) Ikaros pode ser representado em 11 isoformas, ativadoras ou dominante-negativas, e é formado por um domínio de ligação ao DNA (DLD) e um domínio de dimerização (DD). (C) Com a formação de homodímeros pela ligação ao DD (retângulos cor azul-escuro), o complexo se liga ao DNA pelo DLD (retângulos cor azul-claro) sendo capaz de exercer sua função regulatória sobre outros genes. Seja ativando (círculo verde) ou reprimindo (círculo vermelho) seu alvo. Diante de uma proteína dominante-negativa, a resposta é uma formação de heterodímeros, que perde essa capacidade funcional (Modificado de Olsson & Johansson, 2015; Meyer et al, 2013; Raetz & Carroll, 2014).

Para que sua função normal seja exercida, duas estruturas proteicas devem estar presentes. A primeira trata-se do domínio de ligação ao DNA (DLD) na região N-terminal, formado por quatro dedos de zinco codificados pelos éxons 4, 5 e 6. A presença de, pelo menos, três destes já é suficiente para que ocorra a ligação de Ikaros a regiões específicas do DNA. A segunda estrutura compõe o domínio de dimerização (DD) na região C-terminal, formado por dois dedos de zinco codificados pelo éxon 8, essenciais para formação de homo ou heterodímeros com diferentes isoformas ou com outros membros de sua família, como Helios, Aiolos, Eos e Pegasus (codificados por *IKZF2*, *IKZF3*, *IKZF4* e *IKZF5*, respectivamente) (Morgan et al, 1997; Kelley et al, 1998; Kaufmann et al, 2003; Iacobucci et al, 2012).

Atualmente, já foram descritas onze isoformas de Ikaros devido ao mecanismo de *splicing* alternativo que afeta *IKZF1* (Figura 1.5B) (Molnár et al, 1996; Klug et al, 1998; Rebollo & Schmitt, 2003). Seis destas proteínas são conhecidas como isoformas dominante-negativas, pois apresentam perda do DLD, mas o DD permanece inalterado. Isso possibilita a formação de dímeros sem que haja a função de Ikaros. Normalmente, essas isoformas são expressas em baixos níveis capazes de regular a atividade de outras cinco proteínas geradas. Estas são denominadas de ativadoras, pois permanecem com capacidade transcricional visto que ambos domínios foram preservados (Figura 1.5C) (Rebollo & Schmitt, 2003).

A perda da atividade de Ikaros é geralmente uma consequência da presença de deleções genômicas em *IKZF1* (descritas no tópico a seguir), mas também pode ser decorrente de alterações no mecanismo de *splicing*, elevando os níveis das isoformas dominante-negativas, e modificações pós-transducionais (Popescu et al, 2009). A indicação de que este evento poderia estar envolvido no processo de leucemogênese veio de estudos utilizando camundongos. Observou-se que, na presença de mutações heterozigotas em *Ikzf1*, os animais desenvolviam leucemia ou linfoma de células T, resultante da diminuição da expressão de isoformas ativadoras ou aumento das dominante-negativas. Entretanto, esta condição não conduzia à leucemia de células B (Winandy et al, 1995; Papathanasiou et al, 2003, Dumortier, 2006). O que seria um resultado intrigante, pois em humanos, as alterações em *IKZF1* acometem cerca de apenas 5% dos casos de LLA de células T e 15% da LLA-CPB (Marcais et al, 2010; Zhang et al, 2012).

A associação ganhou força por meio de outros estudos em modelos animais, onde se notou que a haploinsuficiência de Ikaros, na presença da fusão gênica *BCR-ABL1*, levava ao desenvolvimento acelerado da LLA-CPB (Virely et al, 2010). Atualmente, acredita-se que a perda de Ikaros, ocasionada por deleções em *IKZF1* levando a haploinsuficiência ou aumento da expressão de isoformas dominante-negativas, seja um evento secundário na LLA-CPB. Evento este essencial para evolução do clone leucêmico. (Joshi et al, 2014; Schwickert et al, 2014).

1.4 TIPOS DE DELEÇÕES EM *IKZF1* E CLASSES FUNCIONAIS

As deleções em *IKZF1* observadas nos blastos dos pacientes variam entre alterações pontuais, intragênicas ou completas (Figura 1.6) (Mullighan et al, 2009; Caye et al, 2013; Dupuis et al, 2013; Meyer et al, 2013; Boer et al, 2016). A maioria são deleções monoalélicas, com apenas 12% dos casos sendo bialélicas. O gene *IKZF1* também pode ser acometido por mutações, entretanto, estas representam 10% dos casos de LLA-CPB e não estão associadas ao prognóstico destes pacientes.

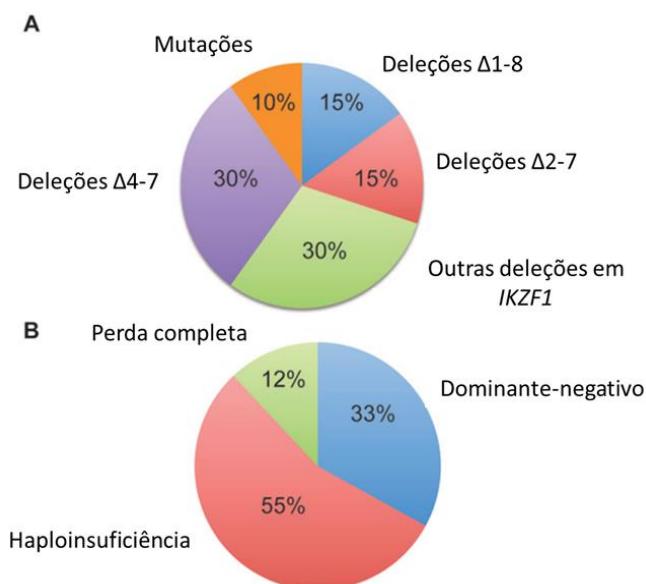


Figura 1.6. Frequência das alterações em *IKZF1*. (A) Frequência de acordo com os tipos de alterações no gene *IKZF1*: mutações, deleções intragênicas comuns (4-7 e 2-7), deleções intragênicas menos comuns e pontuais (outras deleções), e deleções completas. (B) Frequência das deleções de acordo com as classes funcionais: perda completa (devido as deleções bialélicas), haploinsuficiência (acometimento dos éxons 1, 2 e/ou 8) e dominante-negativo (deleções do 4-7) (Modificado de Kastner et al, 2013).

As deleções pontuais são aquelas que acometem um único éxon ou mais, separadamente, como por exemplo: deleção nos éxons 2 e 5. Geralmente, estas alterações são agrupadas com outras deleções, chamadas de intragênicas, que ocorrem com uma frequência mais baixa, como as deleções dos éxons 2-3, 2-8 e 4-8. Juntas, elas representam cerca de 30% dos casos com alteração em *IKZF1*. Outro tipo de deleção intragênica é a que envolve os éxons 2 a 7, presente em aproximadamente 15% dos pacientes LLA-CPB com deleção em *IKZF1*, sendo a segunda mais frequente. A maioria dessas deleções levam a haploinsuficiência do gene, devido ao acometimento do éxon 1 ou 2, importantes para a tradução da proteína. Esta classe funcional também é associada as deleções envolvendo o éxon 8, responsável pelo DD, pois sem este domínio Ikaros não é capaz de formar dímeros e exercer sua função (Dupuis et al, 2012; Kastner et al, 2013).

Dessa forma, deleções completas, ou seja, de todo o gene também causam um efeito de haploinsuficiência. Estas, por sua vez, podem acontecer por diferentes motivos: deleção completa apenas do gene *IKZF1*; deleção do gene com pontos de quebra em regiões adjacentes; grandes deleções ou perda do braço pequeno do cromossomo 7; ou monossomia do 7 (Lopes et al, 2016).

A deleção dos éxons de 4 a 7 é a mais comum, acometendo 30% dos indivíduos com *IKZF1* alterado. Ela é responsável pela síntese aberrante de uma das isoformas dominante-negativas, a Ikaros 6 (ou Ik6), pois perde a capacidade de ligar ao DNA (DLD), embora ainda consiga formar dímeros com as isoformas ativadoras produzidas pelo alelo normal (Dupuis et al, 2012). Dessa forma, ela sequestra as suas proteínas normais, assim como as de sua família, reduzindo sua atividade.

Recentemente, nosso grupo participou de um estudo colaborativo, que teve como objetivo avaliar o valor prognóstico das deleções raras de *IKZF1* na LLA-CPB (Boer et al, 2016). Isso porque a associação do prognóstico adverso às deleções em *IKZF1* são bem estabelecidas para 4-7 e 1-8. Os resultados demonstraram que pacientes com deleções raras em *IKZF1* também apresentam diminuição das taxas de SLE e aumento de recaída (Figura 1.7A e B). Entretanto, é possível observar uma variabilidade na SLE de acordo com cada tipo de deleção (Figura 1.7C), sendo, desta forma, importante avaliar a consequência funcional destas alterações para

melhor compreensão do papel de *IKZF1*. Ainda não é totalmente claro se estas diferentes deleções estão associadas a diferentes mecanismos moleculares.

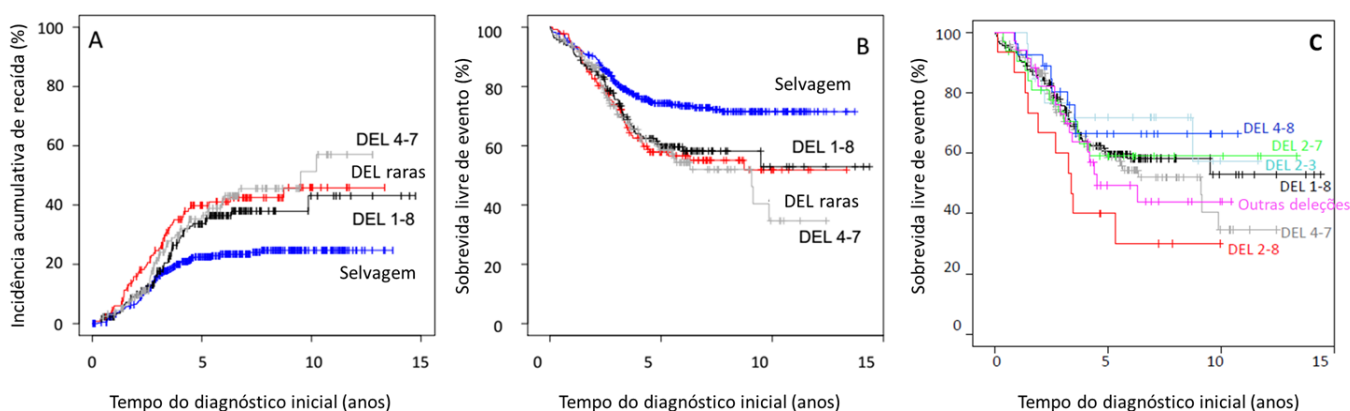


Figura 1.7. Taxas de incidência acumulativa de recaída e sobrevida livre de doença em pacientes com e sem alterações em *IKZF1*. (A) Incidência acumulativa de recaída e (B) sobrevida livre de eventos dos pacientes com LLA-CPB e deleções raras de *IKZF1* (DEL 2-3, DEL 2-7, DEL 2-8, DEL 4-8 e outras deleções) em relação àqueles com *IKZF1* selvagem, DEL 4-7 e DEL 1-8. (C) Sobrevida livre de evento de acordo com cada tipo de deleção (Modificado de Boer et al, 2016).

1.5 O FATOR DE TRANSCRIÇÃO IKAROS

A linfopoiese depende de uma fina regulação da expressão de inúmeros genes específicos a cada estágio e linhagem celular através dos fatores de transcrição. Ikaros é considerado um dos principais reguladores deste processo, com função complexa e ainda com vários aspectos a serem esclarecidos (Georgopoulos et al, 1994; Kirstetter et al, 2002; Papathanasiou et al, 2003).

Estudos demonstraram que altos níveis de Ikaros se encontram em regiões de heterocromatina pericentromérica, correlacionando com a localização de genes reprimidos transcricionalmente (Brown et al, 1997; Cobb et al, 2000) e embasando a hipótese de que esse fator de transcrição também exerce um papel no silenciamento gênico. Atualmente, acredita-se que Ikaros recruta fatores remodeladores de cromatina, regulando, de forma ativa ou repressora, a expressão gênica (Arenzana et al, 2015). Isso ocorre principalmente devido a sua associação com o complexo Mi-2/NuRD, no qual combina remodelamento de nucleossomo dependente de ATP (Mi-2) e atividades histona desacetilase, sendo capaz de ativar ou suprimir a expressão

de genes-alvo (Zhang et al, 1998; Shimono et al, 2003; Yoshida et al, 2008; Zhang et al, 2011).

1.6 GENES DIFERENCIALMENTE EXPRESSOS

Com o objetivo de elucidar os mecanismos pelos quais as deleções em *IKZF1* levam a leucemogênese e contribuem para um prognóstico adverso, alguns estudos realizaram análises em larga escala para expressão gênica em amostras com *IKZF1* selvagem e deletado (Iacobucci et al, 2012; Vitanza et al, 2014). Em outros trabalhos, por ensaios *in vitro*, também foi possível analisar estas duas condições (Ferreiros-Vidal et al, 2013; Schwickert et al, 2014).

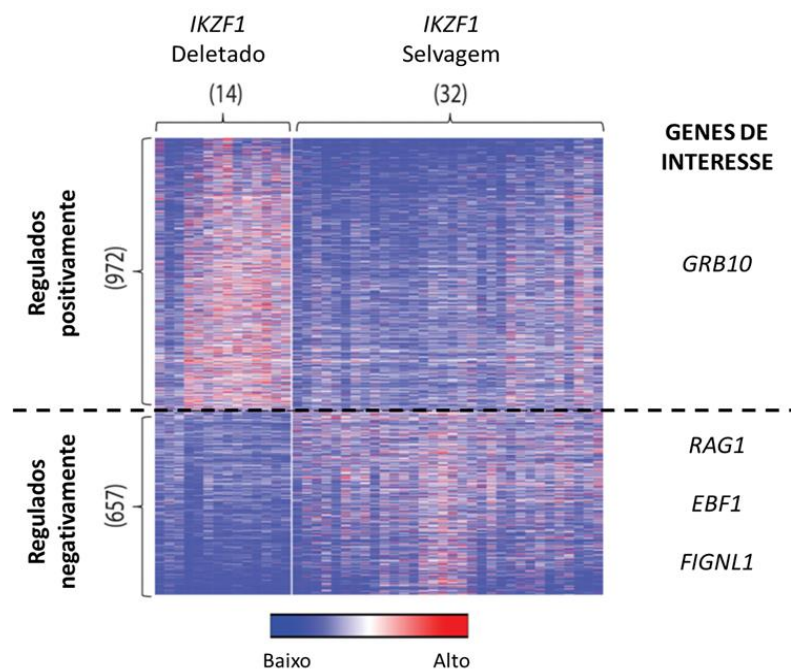


Figura 1.8. Expressão gênica em pacientes LLA-CPB *IKZF1* selvagem e deletado. Heatmap esquemático de microarray demonstrando um perfil diferencial de expressão gênica entre 46 pacientes com LLA-CPB separados em dois grupos: *IKZF1* selvagem (n=32) e deletado (n=14). Observa-se que na presença da deleção, 972 genes são regulados positivamente, com destaque para *GRB10*. Enquanto isso, outros 657 são regulados negativamente, entre eles estão *RAG1*, *EBF1* e *FIGNL1*, alvos do estudo (Modificado de Vitanza et al, 2014).

O resultado observado foi um padrão de expressão gênica diferencial entre os dois grupos (Figura 1.8), sendo *IKZF1* deletado caracterizado por regulação negativa de genes envolvidos na linhagem celular B e na via de reparo do DNA, assim como

regulação positiva de genes envolvidos na progressão do ciclo celular, vias de sinalização e auto-renovação de células progenitoras. Todas as análises demonstraram uma enorme quantidade de genes diferencialmente expressos, dentre os quais destacamos quatro.

1.6.1 RAG1

RAG1 (*Recombination Activating Gene 1*) é um gene localizado na região cromossômica 11p12 responsável por codificar uma proteína essencial para formação e função adequada de células B, através do processo de recombinação V(D)J (Jung et al, 2006). Acredita-se que Ikaros, através de ligação a sítios em regiões promotoras ou *enhancers*, possa regular este gene.

Ferreirós-Vidal e colaboradores, tentando elucidar o papel de Ikaros na especificação e diferenciação da linhagem celular B, analisaram a expressão diferencial de seus genes-alvo em linhagem celular (pré-B) codificando Ikaros selvagem ou Ikaros com deficiência no DLD, ou seja, presença de isoforma dominante-negativa. Foram observados 1982 genes regulados positivamente e 2486 regulados negativamente na presença de Ikaros, envolvidos principalmente na sinalização do receptor de células B, progressão do ciclo celular e recombinação V(D)J. Dentre eles estava o gene *Rag1*, apresentando expressão aumentada dos níveis de expressão no grupo Ikaros selvagem quando comparado com o grupo alterado, sugerindo uma regulação positiva (Ferreiros-Vidal et al, 2013).

1.6.2 EBF1

O fator transcricional *EBF1* (*Early B-Cell Factor 1*) se encontra na região cromossômica 5q33.3 e faz parte de uma família que desempenha papel importante no desenvolvimento de linfócitos, induzindo a expressão de genes específicos de células B e recombinação da imunoglobulina *IGH*, além de reprimir o desenvolvimento mieloide (Pongubala et al, 2008).

Iacobucci e colaboradores investigaram os mecanismos associados com as deleções em *IKZF1* em uma coorte de 51 pacientes adultos com LLA-B. Neste estudo, eles também demonstraram um perfil de expressão diferencial entre os

pacientes com *IKZF1* selvagem (n=20) e deletado (n=31). Os resultados apontaram 294 genes regulados positivamente, enquanto 128 outros genes estavam regulados negativamente, diante da deleção em *IKZF1*. Dentre estes, *EBF1* foi selecionado para análises *in vitro* onde os autores confirmaram seu envolvimento com Ikaros, podendo ser observados sítios de ligação desta proteína na região promotora do gene. Além disso, eles demonstraram fortes linhas de evidência de que a deleção de Ikaros reduz diretamente a ativação de *EBF1* em células leucêmicas, contribuindo para um bloqueio extensivo da diferenciação B (Iacobucci et al, 2012).

1.6.3 FIGNL1

Pouco se sabe sobre o gene *FIGNL1* (*Fidgetin Like 1*) em toda literatura científica. Ele está localizado adjacente a *IKZF1* na região cromossômica 7p12.1 e codifica um membro da família de proteínas AAA+ (ATPases associadas a várias atividades celulares), que desempenham funções importantes em diversos eventos celulares como desdobramento e degradação de proteínas, fusão de membrana e remodelamento de nucleossomos (Hanson & Whiteheart, 2005; Snider & Houry, 2008). Dessa maneira, quando ocorre dano ao DNA, a proteína codificada por *FIGNL1* é recrutada para os sítios específicos e, através da formação de um complexo com outras proteínas, realiza o reparo do DNA com eficiência via recombinação homóloga (Yuan & Chen, 2013).

Schwickert e colaboradores, ao estudar sobre o controle de Ikaros nas células pró-B expressando transcritos de *Ikzf1* alterado (deleção do éxon 8) e com transcritos completos, observaram 215 genes regulados ativamente contra 295 regulados negativamente na presença do modelo selvagem. Os autores relataram que a expressão de *Figl1* estava entre 5-10x aumentada (Schwickert et al, 2014). Além disso, dados recentes do nosso grupo, demonstraram que a maioria das deleções completas em *IKZF1* abrangem *FIGNL1* (Lopes et al, 2016).

1.6.4 GRB10

O gene *GRB10* (*Growth Factor Receptor Bound Protein 10*) também se encontra na região cromossômica 7p12.1, adjacente a *IKZF1*, e codifica uma proteína pertencente a uma família de proteínas adaptadoras, que interagem com alguns receptores de tirosina quinase e moléculas de sinalização. No que diz respeito a leucemia, pouco se conhece sobre esse gene. Entretanto, já foi observado que em células de leucemia mieloide aguda (LMA), a proteína Grb10 se liga ao receptor FLT3, levando a ativação das vias de sinalização *PI3K-Akt* e *STAT5*, tendo como resultado um aumento na proliferação, sobrevivência e formação de colônias de células hematopoiéticas. Além de uma parcial correlação da elevada expressão dessa proteína com recaída e prognóstico adverso (Kazi & Ronnstrand, 2013).

Na LLA-CPB, um recente estudo demonstrou que Grb10 pode estar envolvido na leucemogênese mediada pela fusão gênica *BCR/ABL1*. Através de modelo *in vivo* com *knockdown* dessa proteína, observou-se diminuição da proliferação, sobrevivência prolongada e reduzida capacidade de formação de colônia (Illert et al, 2015). Vitanza e colaboradores, em 2014, através do painel de expressão gênica diferencial, demonstraram que o gene *GRB10* estaria regulado positivamente em pacientes pediátricos com LLA-CPB e *IKZF1* deletado (Vitanza et al, 2014).

1.7 DELEÇÃO EM IKZF1 E BTG1

O gene *BTG1* (*B-Cell Translocation Gene 1*) está localizado na região cromossômica 12q21.33 e é membro de uma família anti-proliferativa. A expressão deste gene é elevada nas fases G0/G1 do ciclo celular, sendo posteriormente reduzida. *BTG1* é responsável por codificar uma proteína capaz de co-regular o crescimento e co-ativar a diferenciação celular, interagindo com diversos receptores nucleares. Deleções neste gene são considerados eventos secundários na LLA-CPB, sendo frequentemente encontrado nos subgrupos com as fusões *ETV6-RUNX1* e *BCR-ABL1* (Tsuzuki et al, 2007; Xie et al, 2014).

Recentemente, tem sido observado uma alta frequência de deleções no gene *BTG1* dentro dos casos de LLA-CPB *IKZF1* deletado. Scheijen e colaboradores, neste ano, associaram a presença concomitante de deleções em *IKZF1* e *BTG1* com

a redução de SLE e aumento de recaída em uma coorte de 533 crianças (Figura 1.9). Além disso, ao tentar estabelecer que este genótipo combinado cooperava para a leucemogênese através de ensaios *in vivo*, observaram um aumento na incidência do tumor e aumento da resistência aos glicocorticoides (Scheijen et al, 2017).

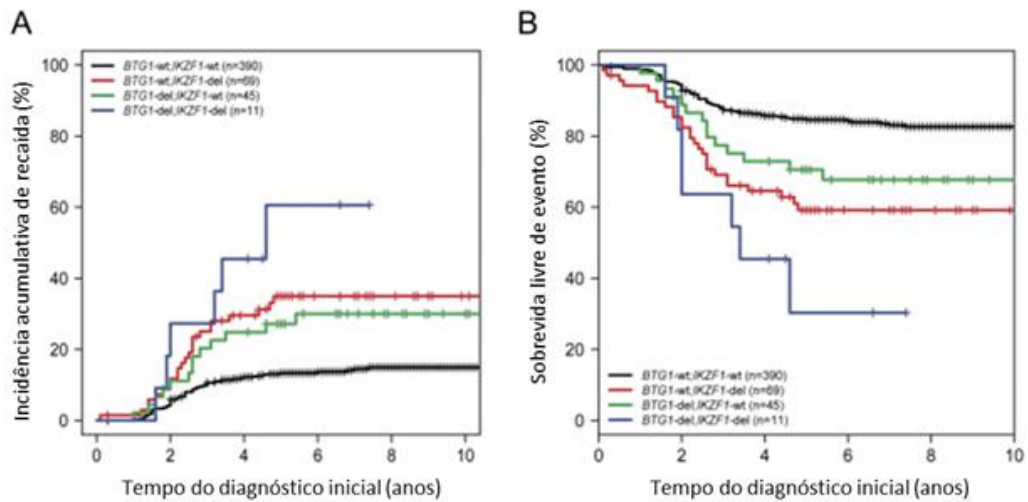


Figura 1.9. Taxas de incidência acumulativa de recaída e sobrevida livre de doença e em pacientes com alterações em *IKZF1* e *BTG1*. (A) Incidência acumulativa de recaída e (B) Sobrevida livre de eventos de pacientes com LLA-CPB do grupo de oncologia pediátrica da Nova Zelândia e Austrália. wt: selvagem. del: deletado (Modificado de Scheijen et al, 2017).

2 JUSTIFICATIVA DO ESTUDO

Como apresentado anteriormente, a LLA-CPB é a neoplasia mais comum entre as crianças, com cerca de 15% dos casos sendo acometidos por deleções em *IKZF1*. No decorrer dos anos, foi observado sua associação com o aumento das taxas de recaída em diversos protocolos, despertando o interesse do nosso grupo em contribuir, através de análises genômicas, para uma melhor compreensão destas alterações na leucemogênese.

Recentemente, em um estudo colaborativo, foi demonstrado que a presença de deleções em *IKZF1*, independente do tipo, também confere uma pior SLE a estes pacientes (Boer et al, 2016). Apesar disso, hipotetizamos que a consequência funcional das deleções em *IKZF1* pode apresentar variações, pois algumas destas alterações, descritas anteriormente, resultam em isoformas dominantes-negativas, enquanto outras na haploinsuficiência do gene.

Estudos recentes já identificaram uma diferença de expressão gênica entre casos com ou sem deleção em *IKZF1* (Iacobucci et al, 2012; Vitanza et al, 2014). Entretanto, não há relato se o perfil observado é influenciado pelos diferentes tipos de deleção encontrados em *IKZF1*. Portanto, com este estudo, pretendemos avaliar a hipótese de que os diferentes status de *IKZF1* são capazes de mudar a expressão de genes-alvo. Entre estes, destacamos os genes *EBF1*, *RAG1* e *BTG1* por estarem envolvidos na linfopoiese B, além dos genes *FIGNL1*, envolvido na via de reparo do DNA, e *GRB10*, que interage com receptores de tirosina quinase.

3 OBJETIVOS

3.1 GERAL

- Avaliar o impacto das deleções recorrentes no gene *IKZF1* na regulação da expressão de seus genes-alvo em crianças com leucemia linfoblástica aguda de células precursoras B.

3.2 ESPECÍFICOS

- Avaliar os níveis de expressão relativa de *IKZF1* e dos seus genes-alvo (*RAG1*, *EBF1*, *FIGNL1*, *GRB10* e *BTG1*);
- Associar a expressão de genes-alvo de *IKZF1* com dados clínico-laboratoriais de pacientes com LLA-CPB;
- Correlacionar a expressão dos genes-alvo com os tipos de deleção encontrados no *IKZF1*;
- Validar os resultados através de análises *in silico*;
- Correlacionar os níveis de expressão com o prognóstico dos pacientes.

4 MATERIAL E MÉTODOS

4.1 CASUÍSTICA

Foram incluídas amostras de pacientes com idade igual ou inferior a 16 anos, diagnosticados com LLA-CPB entre o período de janeiro de 2004 a julho 2017. Pacientes com tratamento prévio com quimioterápicos e portadores de síndromes genéticas congênitas foram excluídos do estudo, assim como amostras com percentual de blastos inferior a 30% ou inviáveis para as análises laboratoriais. O presente estudo foi aprovado pelo comitê de ética e pesquisa (CEP) do INCA sob o número de protocolo #33709814.7.0000.5274 (Anexo 1).

As amostras foram caracterizadas morfológicamente e imunofenotipicamente como LLA-CPB conforme os critérios de diagnóstico de LLAs previamente descritos (Catovsky et al, 1991; Béné, 2005), seguindo os procedimentos laboratoriais do Programa de Hematologia-Oncologia Pediátrico (PHOP) do INCA (Pombo de Oliveira et al, 2009). O perfil citogenético-molecular também foi determinado e com a caracterização completa dos casos (Arber et al, 2016), as amostras seguiram o fluxograma do estudo (Figura 4.1).

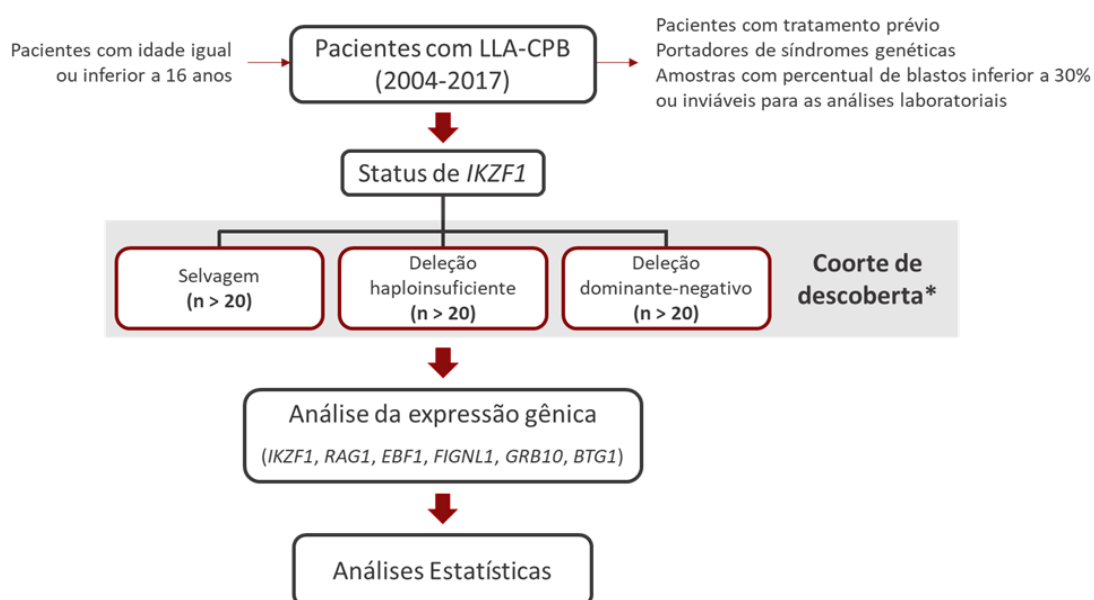


Figura 4.1. Desenho esquemático do estudo. *Após a determinação do status de *IKZF1*, foram selecionados cerca de 20 casos em cada grupo com a finalidade de compor a coorte de descoberta, tornando possível as análises estatísticas comparativas entre eles.

4.2 DETERMINAÇÃO DO STATUS DE IKZF1

O status de *IKZF1* foi determinado como: a) selvagem; b) deletado do tipo haploinsuficiente (Δ haplo), quando as deleções envolviam os éxons 1, 2 e/ou 8; e c) deletado do tipo dominante-negativo (Δ dn), quando as deleções eram nos éxons 4, 5 e 6. Para isso, foi utilizada a combinação das técnicas de MLPA e PCR *Multiplex* (MP-PCR), conforme descrito a seguir.

4.2.1 Extração de ácido desoxirribonucleico (DNA)

Os DNAs foram obtidos das amostras de aspirados de MO e/ou SP coletadas em EDTA, através de dois métodos diferentes: a) ensaio comercial *QIAmp DNA Blood Mini Kit* (Qiagen); e b) fenol-clorofórmio com uso do reagente TRIzol (Invitrogen).

4.2.1.1 Ensaio comercial QIAmp DNA MiniBlood

As amostras frescas, com quantidade suficiente de células, seguiram para o método de extração através do ensaio comercial *QIAmp DNA Blood Mini Kit* (Qiagen), seguindo o protocolo fornecido pelo fabricante.

Em tubos de 1,5 mL previamente identificados, foram adicionados 20 μ L da proteinase K, 200 μ L de amostra e 200 μ L de tampão de lise (*Buffer AL*). Após homogeneização em vórtex, as amostras foram incubadas a 56°C por 10 minutos no banho seco. Centrifugou-se brevemente os tubos para remover gotas da tampa e então, foram adicionados 200 μ L de etanol 100%. Após nova homogeneização em vórtex, cada mistura foi transferida para uma coluna *QIAamp Mini Spin*, acoplada a um tubo coletor de 2 mL. As colunas foram centrifugadas a 8000 rpm por 1 minuto e o líquido filtrado foi descartado. Em seguida, 500 μ L do tampão de lavagem 1 (*Buffer AW1*) foi adicionado a cada coluna, que foram centrifugadas a 8000 rpm por 1 minuto e novamente, o líquido filtrado foi descartado. Adicionou-se 500 μ L do tampão de lavagem 2 (*Buffer AW2*), as colunas foram centrifugadas na velocidade máxima de 14000 rpm por 3 minutos e o líquido filtrado foi novamente descartado. Esta última etapa de centrifugação foi repetida na mesma velocidade por 1 minuto e tanto o líquido filtrado quanto o tubo coletor foram descartados. Cada coluna foi

transferida para um tubo de 1,5 mL, precisamente identificado, sendo adicionados 15 µL de tampão de eluição (*Buffer AE*). As amostras foram incubadas à temperatura ambiente por 1 minuto e, então, centrifugadas a 8000 rpm por 1 minuto. A etapa de eluição foi repetida, resultando em DNA no volume final de 30 µL.

4.2.1.2 Reagente TRIzol

As amostras que, após o procedimento para lise de hemácias, foram armazenadas a -80°C em 1 mL de TRIzol (Invitrogen), posteriormente, foram submetidas ao método de fenol-clorofórmio com base no protocolo fornecido pelo fabricante do reagente.

Primeiramente, foram adicionados 200 µL de clorofórmio em cada amostra, homogeneizadas vigorosamente por alguns segundos. Após incubação à temperatura ambiente por 3 minutos, estas foram centrifugadas a 4°C por 15 minutos a 14000 rpm. Dessa forma, foi obtida a separação das camadas orgânica, intermediária e aquosa (transparente). A fase aquosa foi transferida para um novo tubo precisamente identificado, que seguiu para extração do ácido ribonucleico (RNA) (*vide* tópico 4.3.1). As fases de baixo foram submetidas ao protocolo de extração do DNA descrito a seguir.

Foram adicionados 300 µL de etanol 100% em cada tubo de amostra, e então, homogeneizadas por inversão. Após incubação à temperatura ambiente por 3 minutos, estas foram centrifugadas a 10000 rpm por 5 minutos a 4°C e o sobrenadante descartado. O precipitado foi ressuspenso com 1 mL de solução de citrato de sódio 0,1M em etanol 10%. Após nova incubação à temperatura ambiente por 30 minutos, sempre misturando por inversão, foi realizada uma centrifugação a 10000 rpm por 5 minutos a 4°C e o sobrenadante descartado. Então, a etapa de lavagem com a solução de citrato de sódio foi repetida. Em seguida, o precipitado foi eluído em 1,5-2 mL de etanol 75% e incubado por 15 minutos à temperatura ambiente, misturando por inversão periodicamente. Então, após nova centrifugação a 10000 rpm por 5 minutos a 4°C, os tubos contendo os precipitados foram deixados abertos à temperatura ambiente para que fosse retirado o máximo de etanol presente. Dessa forma, cada amostra foi dissolvida em 50 µL de tampão de eluição e incubadas por 6 minutos no banho seco a 68°C. Após a incubação, centrifugou-se a 12000 rpm por 10 minutos para que o material insolúvel fosse removido e em

seguida, o sobrenadante contendo o DNA foi transferido para um novo tubo identificado.

O DNA obtido por este método foi submetido a uma etapa de precipitação com a finalidade de reduzir a concentração de contaminantes. Tal passo foi necessário devido as condições recomendadas para o uso com as sondas de MLPA.

4.2.2 Quantificação do DNA

A quantificação do DNA foi realizada após os processos de extração através da leitura no espectrofotômetro *NanoDrop 1000* (Thermo Fisher Scientific). A mensuração da quantidade de DNA obtida das amostras foi realizada utilizando 1 μ L de cada. A unidade de concentração foi dada em ng/ μ L, sob absorvância medida à 260 nm. Contudo, foi utilizado como valores de referência, da razão dos comprimentos de onda 260 nm e 280 nm, números próximos a 1,8, para verificar a pureza das amostras.

4.2.3 Identificação de alterações no número de cópias por MLPA

Após a obtenção do DNA, as amostras foram submetidas a técnica de MLPA para detecção de deleções em *IKZF1* (Schwab et al, 2010; Barbosa et al, 2015). Considerada como técnica padrão para tal finalidade, a MLPA se baseia em uma reação em cadeia da polimerase onde as sondas hibridizadas em uma sequência de DNA alvo específica são amplificadas utilizando apenas um par de oligonucleotídeos, permitindo a quantificação comparativa de múltiplos alvos (Schouten et al, 2002). Dessa forma, a reação pode ser dividida em cinco etapas: i) desnaturação do DNA; ii) hibridização das sondas; iii) reação de ligação; iv) reação de PCR; e v) análise dos fragmentos (Figura 4.2).

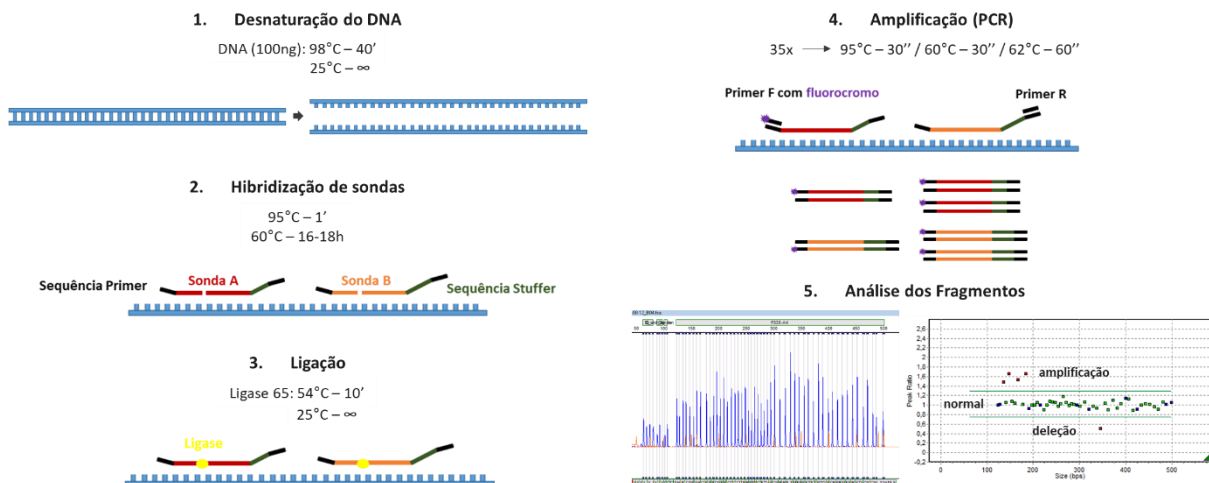


Figura 4.2. Desenho esquemático da técnica de MLPA. Essa técnica consiste em 5 etapas: desnaturação do DNA; hibridização das sondas (quer sejam comerciais ou não); ligaçao pela enzima ligase; reaçao de PCR propriamente dita; e por fim, análise dos fragmentos.

4.2.3.1 Ensaios comerciais com sondas para IKZF1

Primeiramente, foi utilizado o ensaio comercial *SALSA MLPA P335 ALL-IKZF1* (MRC-Holland) que contém 1 sonda localizada em cada éxon de *IKZF1*, além de sondas em outros genes como *PAX5*, *ETV6*, *RB1*, *BTG1*, *EBF1*, *CDKN2A-CDKN2B* e região PAR1 (SHOX area, *CRLF2*, *CSF2RA*, *IL3RA* e *P2RY8*).

Com o objetivo de confirmar as alterações observadas em *IKZF1*, uma outra reaçao foi realizada, agora com o ensaio comercial *SALSA MLPA P202 IKZF1* (MRC-Holland). Este ensaio contém 2 sondas em cada éxon de *IKZF1*, além de sondas em outros genes como *FIGNL1*, *DDC*, *ZBPB*, *IKZF2*, *IKZF3*, *CDKN2A-CDKN2B*, *MIR31*, *IGHD*, *MTA1* e *KIAA0284*.

4.2.3.2 Reaçao de MLPA

Para ambos os ensaios, foram utilizadas pelo menos 3 amostras controles por reaçao. Estas são derivadas de uma mistura de amostras de crianças cujo diagnóstico de leucemia foi descartado, submetidas às mesmas condições de extraçao daquelas a serem comparadas.

Dessa forma, 100 ng de DNA de todas as amostras foram desnaturados a 98°C por 40 minutos e em seguida, uma mistura contendo as sondas do respectivo ensaio comercial foi adicionada. A soluçao resultante foi aquecida a 98°C por 1

minuto, com posterior incubação a 60°C por 16-18 horas, para hibridização das sondas. Na etapa seguinte, a enzima *Ligase-65* foi adicionada e incubou-se por mais 15 minutos a 54°C. Após inativação da mesma a 98°C por 5 minutos, uma nova mistura contendo os oligonucleotídeos direto e reverso, com um deles marcado com fluorocromo, do respectivo ensaio comercial foi adicionada para etapa de PCR. A solução final foi submetida às seguintes condições: 95°C por 30 segundos, 60°C por 30 segundos e 72°C por 1 minuto em 35 ciclos, e 72°C por 20 minutos.

Para cada amostra, 0,8 µL do produto e 0,2 µL do marcador padrão *GeneScan LIZ 500* (Applied Biosystems) foram adicionados em 9 µL de Formamida Hi-Di (Life Technologies) e transferidas para uma placa de 96 poços. As amostras foram desnaturadas a 95°C por 5 minutos e, posteriormente, resfriadas por 2 minutos no gelo. No sequenciador ABI 3130xl (Applied Biosystems), os fragmentos de PCR foram separados por eletroforese capilar e os dados coletados através de fluorescência. Como etapa final, as análises dos fragmentos foram realizadas através do programa *GeneMarker* versão 2.2.0 (SoftGenetics), onde os parâmetros utilizados para determinação do valor de pico relativo foram: valores entre 0,75 e 1,3 considerados padrão normal; valores abaixo de 0,75 considerados padrão de perda (deleção); e valores acima de 1,3 considerados padrão de ganho (amplificação) (Schwab et al, 2010; Barbosa et al, 2015).

4.2.4 Confirmação de deleções recorrentes por MP-PCR

A técnica de MP-PCR consiste em uma variante de PCR que permite a amplificação de várias regiões de interesse em uma única reação devido ao uso de mais de um par de oligonucleotídeos. As reações, denominadas A e B, empregadas para a identificação de deleções em *IKZF1* foram previamente descritas por Meyer et al. em 2013. A reação A permite a detecção de deleções que acometem os éxons 2 e 3 ($\Delta 2-3$), enquanto que a reação B detecta deleções que afetam os éxons 2 a 7 ($\Delta 2-7$), 2 a 8 ($\Delta 2-8$), 4 a 7 ($\Delta 4-7$) ou 4 a 8 ($\Delta 4-8$) (Meyer et al, 2013). No presente estudo, esta metodologia foi utilizada como teste confirmatório das deleções encontradas por MLPA (ensaio comercial *SALSA MLPA P335 ALL-IKZF1*).

Os oligonucleotídeos utilizados em ambas as reações estão listados na tabela 4.1. Para a reação A os oligonucleotídeos foram separados em 3 diferentes

misturas: i) MP-F: oligonucleotídeos de sentido direto; ii) MP-R1: oligonucleotídeos de sentido reverso numerados ímpares; e iii) MP-R2: oligonucleotídeos de sentido reverso numerados pares. Cada oligonucleotídeo, ou mistura, foi usado na concentração de 0,4 μ M, além de dNTP 0,2 μ M (GE Healthcare). Os demais reagentes disponibilizados pelo ensaio comercial *Platinum Taq DNA polymerase High Fidelity* (Invitrogen), foram usados nas seguintes concentrações: solução de tampão 1x, MgSO₄ 2,5 mM e 1U de *Taq Polymerase High Fidelity*. A concentração de DNA utilizada foi de 200 ng por reação, possuindo cada uma um volume final de 50 μ L.

Tabela 4.1. Sequência dos oligonucleotídeos de *IKZF1* para as reações do MP-PCR.

Reação A	Oligonucleotídeos	Sequência	Reação A	Oligonucleotídeos	Sequência
MP-F	IKZF1.I1.F1B	5'-AGTTCACCTTCTGTCAAGCGTCTGTTGCTCT-3'	MP-R2	IKZF1.I3.R2	5'-GATGGCACTGGCAGTCATTCTCTATGTCT-3'
	IKZF1.I1.F2	5'-TGGATGTGTGTGTTTCATGCGTGGTTAATA-3'		IKZF1.I3.R4	5'-TCTAGGAAGGACTTGGGCACATTGAAGAAT-3'
	IKZF1.I1.F3	5'-TCATGTGGACCATGGCTTTCTGTATTCT-3'		IKZF1.I3.R6	5'-TCTCGGCACTACACACACTCTCTTTAGGC-3'
	IKZF1.I1.F4	5'-TGGCTGAAAATGGGCTCTAATTAGTGGAAA-3'		IKZF1.I3.R8	5'-CTGGCACTTCTGTCAAACCTCACATCTCT-3'
MP-R1	IKZF1.I3.R1	5'-CCCAATTGATGTTAAAATTGGACTGGACCT-3'	IKZF1.I3.R10	5'-TTTCATATAAAAATGCTGCCAACACCTTGA-3'	
	IKZF1.I3.R5	5'-CTGTTACTGCCTGCAGGATAGACTTCTGGA-3'	IKZF1.I3.R10B	5'-GCAAGCAGCCAGTTTTACAATAACCTTTGC-3'	
	IKZF1.I3.R7	5'-GGTACCCCAACCCATCCTTATACATGACAC-3'	IKZF1.I3.R12	5'-AATGTACACTGTTAGTCCCACTGACCAA-3'	
	IKZF1.I3.R9	5'-CTTCCGGTCCAGGATCTCCATATAACAAT-3'	IKZF1.I3.R14	5'-CAAGGACTCTATGACTCGGTACCACCTTGA-3'	
	IKZF1.I3.R11	5'-TATTCTTTACAGGACAGTTTCCAGCA-3'			
	IKZF1.I3.R13	5'-TGACTGAGACATAATGGACAAGGCCAAT-3'			
Reação B	Oligonucleotídeos	Sequência	Reação B	Oligonucleotídeos	Sequência
	IKZF1.I1.F1B	5'-AGTTCACCTTCTGTCAAGCGTCTGTTGCTCT-3'		IKZF1.I3.F2	5'-GAATTGCTTCACGATCACATAACACAATGC-3'
	IKZF1.I1.F2	5'-TGGATGTGTGTGTTTCATGCGTGGTTAATA-3'		IKZF1.I7.R4	5'-ACAGACCATAGAGTCCCTCTAGGGGAAAA-3'
	IKZF1.I1.F3	5'-TCATGTGGACCATGGCTTTCTGTATTCT-3'		IKZF1.E8.R6	5'-TTCATCATTTGTTCTCCTCTCGTAGCTG-3'
	IKZF1.I1.F4	5'-TGGCTGAAAATGGGCTCTAATTAGTGGAAA-3'		IKZF1.3U.R1	5'-CTTCTCCCTCTCATATTGGGATCTCATTGG-3'
				IKZF1.HS.R2	5'-TGTTCTCTTGTCTCAACAGACACAATG-3'

Legenda do nome de cada oligonucleotídeos: nome do gene, seguido do local de anelamento (I – íntron; E – éxon; 3U – 3-UTR; HS – Hotspot) e da direção da fita (F – direto/ R – reverso)

A estratégia de ciclagem adotada foi de PCR de longa distância (LD-PCR) com duração média de 13 horas. Em uma fase inicial, as amostras são submetidas a temperatura de 93°C por 3 minutos, seguida de 10 ciclos com desnaturação em 15 segundos a 93°C, anelamento em 30 segundos a 62°C e extensão em 18 minutos a 68°C. A partir daqui, iniciam-se 32 ciclos com desnaturação a 93°C por 15 segundos, anelamento a 62°C por 30 segundos e extensão a 68°C por 18 minutos, sendo que nesta última são adicionados 20 segundos por ciclo; e então a ciclagem é finalizada com uma fase de extensão de 68°C por 10 minutos.

Os produtos foram separados por eletroforese em gel de agarose 1,5%, pós corado em solução de brometo de etídio (Invitrogen) e visualizados com o auxílio de um transiluminador UV ChemiDoc™ (Bio-Rad).

4.3 QUANTIFICAÇÃO RELATIVA DO NÍVEL TRANSCRICIONAL DOS GENES DE INTERESSE

Após a determinação do status de *IKZF1* nos blastos dos pacientes, os níveis de RNA foram quantificados, de acordo com os genes selecionados para este estudo, através da técnica de PCR quantitativo (qPCR), seguindo as diretrizes propostas no *Minimum Information for Publication of Quantitative Real-Time PCR Experiments* (MIQE). Para isso, foram observados padrões de preparo e de qualidade das amostras até a execução da metodologia propriamente dita (Nolan et al, 2006; Bustin et al, 2009).

4.3.1 Extração de ácido ribonucleico (RNA)

O RNA foi obtido pelo método de fenol-clorofórmio com o uso do reagente TRIzol (Invitrogen) antes da extração de DNA (*vide* tópico 4.2.1.2). Após a separação das camadas, a fase aquosa de cada amostra foi transferida para um novo tubo precisamente identificado e seguiu para isolamento de RNA conforme protocolo disponibilizado pelo fabricante do reagente.

Em cada tubo foram adicionados 500 µL de isopropanol 100%. Após incubação à temperatura ambiente por 10 minutos, o material foi centrifugado a 4°C por 10 minutos a 14000 rpm. O sobrenadante foi descartado e então, o precipitado foi ressuspenso com 1 mL de etanol 75%. Centrifugou-se novamente a 14000 rpm, entretanto, por apenas 5 minutos a 4°C e o sobrenadante foi descartado. Os tubos foram deixados abertos por 5-10 minutos à temperatura ambiente para secagem do precipitado. Ao final, foram adicionados 20 µL de água DEPC e após homogeneização, as amostras foram incubadas por 10 minutos a 55-60°C em banho seco.

4.3.2 Quantificação e qualidade do RNA

O RNA extraído foi quantificado através da leitura no espectrofotômetro *NanoDrop 1000* (Thermo Fisher Scientific), utilizando apenas 1 µL de cada amostra. A unidade de concentração foi dada em ng/µl, sob absorvância medida à 260 nm.

Para verificar a pureza, foram considerados como valores de referência números próximos a 2,0, resultantes da razão dos comprimentos de onda 260 nm e 280 nm.

A qualidade dessas amostras foi avaliada através de eletroforese em gel de agarose 1,5% pré-corado com 3 μ L *GelRed* (Uniscience) ou pós-corado em solução de brometo de etídio (Invitrogen). Pelo menos 200 ng de RNA foram aplicados no gel, que logo após a corrida eletroforética, puderam ser visualizados com o auxílio do transiluminador UV E-Gel Imager (Life Technologies) ou UV ChemiDoc™ (Bio-Rad), dependendo do corante utilizado. As amostras consideradas íntegras ainda foram submetidas ao processo de purificação.

4.3.3 Purificação de RNA

Após estas etapas, o RNA foi purificado utilizando o ensaio comercial *PureLink RNA Mini Kit* (Invitrogen), seguindo o protocolo padrão disponibilizado pelo fabricante.

Cerca de 10-20 μ L das amostras foram separadas para o procedimento de purificação. O mesmo volume foi utilizado para adicionar a solução de tampão de lise com 1% de 2-mercaptoetanol, assim como para adicionar etanol 100% em cada tubo. Após homogeneizar por vórtex, o volume total da amostra foi transferido para uma coluna *Spin Cartridge*, com um tubo coletor. Centrifugou-se por 15 segundos a 14000 rpm e descartou-se o líquido filtrado. Em seguida, foram adicionados 500 μ L de tampão de lavagem 2 com etanol. Centrifugou-se novamente por 15 segundos a 14000 rpm e descartou-se o líquido filtrado. A etapa de lavagem foi repetida mais uma vez. Por fim, a coluna foi novamente centrifugada a 14000 rpm por 1 minuto, para secar a membrana contendo o RNA. O tubo coletor foi descartado e a coluna foi colocada em um novo tubo de 1,5 mL identificado. O volume de água DEPC utilizada para eluir a amostra foi dependente da concentração e do volume inicial do RNA. Uma parte do volume foi adicionada à coluna, que ficou incubada à temperatura ambiente por 1 minuto. Em seguida, foi realizada uma centrifugação de 14000 rpm por 2 minutos. O restante do volume foi adicionado posteriormente, repetindo as etapas de incubação e centrifugação.

Com o RNA purificado, as análises de quantificação e qualidade foram novamente realizadas de modo semelhante ao descrito no tópico 4.3.2. Apenas as

amostras com boa quantificação e qualidade seguiram para a síntese do DNA complementar (cDNA).

4.3.4 Síntese de DNA complementar (cDNA)

O cDNA foi obtido através do ensaio comercial *SuperScript VILO Master Mix* (Invitrogen). Esta mistura combina, em um único tubo, *SuperScript III Reverse Transcriptase* e outros componentes de reação da transcriptase reversa (RT).

4.3.4.1 Reação de síntese de cDNA

Para síntese de 500 ng de RNA em no máximo 18 µL (volumes abaixo foram completados com água DEPC), foram utilizados 2 µL de *SuperScript VILO Master Mix* (Invitrogen). A mistura foi homogeneizada e levada ao termociclador sob as condições: 25°C por 10 minutos, 42°C por 60 minutos, e 85°C por 5 minutos.

4.3.4.2 RT-PCR para amplificação do gene constitutivo GAPDH

A técnica de RT-PCR foi empregada neste momento do estudo para analisar a integridade dos cDNAs através da amplificação do gene constitutivo *GAPDH*. A sequência dos oligonucleotídeos usados estão na tabela 4.2.

Tabela 2. Sequência dos oligonucleotídeos do gene *GAPDH* para reação de RT-PCR.

Oligonucleotídeos	Sequência
<i>GAPDH</i> sentido direto	5' TGACCCCTTCATTGACCTCA 3'
<i>GAPDH</i> sentido reverso	5' AGTCCTTCCACGATACCAAA 3'

As condições para a reação com volume final de 20 µL foram: 0,8 µl de cDNA, oligonucleotídeos 0,2 µM, dNTPs 0,15 µM (GE Healthcare); e solução tampão 1X, MgCl₂ 2,5 mM e 1U de *Taq* DNA polimerase, reagentes disponibilizados pelo ensaio comercial *Taq DNA Polymerase* (Invitrogen). Em um termociclador, foi programada a seguinte ciclagem: desnaturação inicial a 95°C por 30 segundos; 35 ciclos de desnaturação a 94°C por 30 segundos, anelamento a 60°C por 30 segundos e extensão a 72°C por 45 segundos; e extensão final a 72°C por 10 minutos.

Após corrida eletroforética, os produtos de PCR foram visualizados em gel de agarose 2% pré-corado com 3 μ L *GelRed* (Uniscience) ou pós-corado em solução de brometo de etídio (Invitrogen), com auxílio do transiluminador UV E-Gel Imager (Life Technologies) ou UV ChemiDoc™ (Bio-Rad), respectivamente. O tamanho do produto de PCR esperado foi de 419 pb (Figura 4.3).

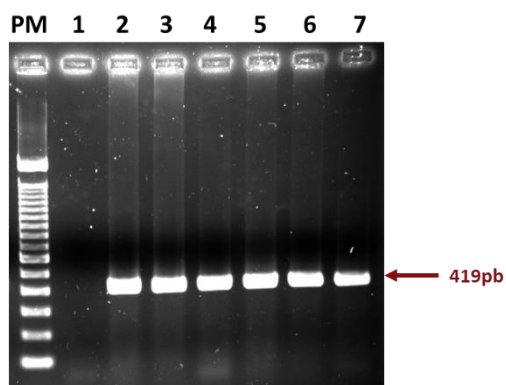


Figura 4.3. Eletroforese em gel de agarose para a reação de RT-PCR. Visualização dos cDNAs íntegros de pacientes (2 ao 7); junto com controle negativo (1). PM: peso molecular de 100pb.

4.3.5 Quantificação gênica por RT-qPCR

Com os cDNAs funcionantes, foi iniciado o procedimento para quantificação dos níveis de expressão gênica nas amostras selecionadas. A técnica empregada, nesta etapa, foi a RT-qPCR, que permite a detecção e quantificação, em tempo real, dos produtos amplificados a cada ciclo, devido a sua proporcionalidade com sinal fluorescente mensurado.

Dois sistemas de detecção foram usados na metodologia, por questões de custos. Enquanto os genes *RAG1*, *EBF1*, *FIGNL1*, *GRB10* e *BTG1* foram submetidos a técnica de RT-qPCR com sondas hidrolisáveis, *IKZF1* foi submetido a mesma, contudo pelo método de agentes intercalantes de fita-dupla, ambos discutidos separadamente nos tópicos abaixo. O gene de referência utilizado nos dois sistemas foi o *18s rRNA*, identificado através de dados obtidos previamente no nosso grupo (Barbosa, 2017).

4.3.5.1 Ensaio com sondas hidrolisáveis

Esta metodologia consiste numa reação de PCR utilizando oligonucleotídeos específicos para a região de interesse, além de sondas também específicas

contendo em suas extremidades um fluoróforo e um *quencher* (absorve a fluorescência). Ocorrendo a hibridização destas estruturas ao cDNA, a síntese começa a ser realizada por uma enzima também capaz de degradar a sonda anelada. Uma vez que a ligação entre os compostos é rompida, o fluoróforo passa a emitir fluorescência que é capturada pelo aparelho adequado.

Para este ensaio, utilizamos as placas customizadas *TaqMan array 96-well plate* (Thermo Fisher Scientific) com os oligonucleotídeos dos genes *18s rRNA*, *RAG1*, *EBF1*, *FIGNL1*, *GRB10* e *BTG1* liofilizados por linha. Cada placa foi preparada com 12 amostras. Para o preparo da reação, foram usados *TaqMan Master Mix* (Thermo Fisher Scientific): *AmpliTaq Gold® DNA Polymerase, UP (Ultra Pure)*, dNTPs (com dUTP), *ROX™ Passive Reference, Uracil-N glycosylase (UNG)* e componentes de tampão otimizados; água DEPC e cDNA na concentração de 9,3 ng. De acordo com o protocolo fornecido pelo fabricante, para cada amostra foi preparada uma mistura de 8 reações, contendo 90 µL de *TaqMan Master Mix* 90 µL de cDNA mais água DEPC. Após homogeneização, foram aplicados 20 µL em cada poço e a reação seguiu para o aparelho sendo submetida a ciclagem esquematizada na figura 4.4A.

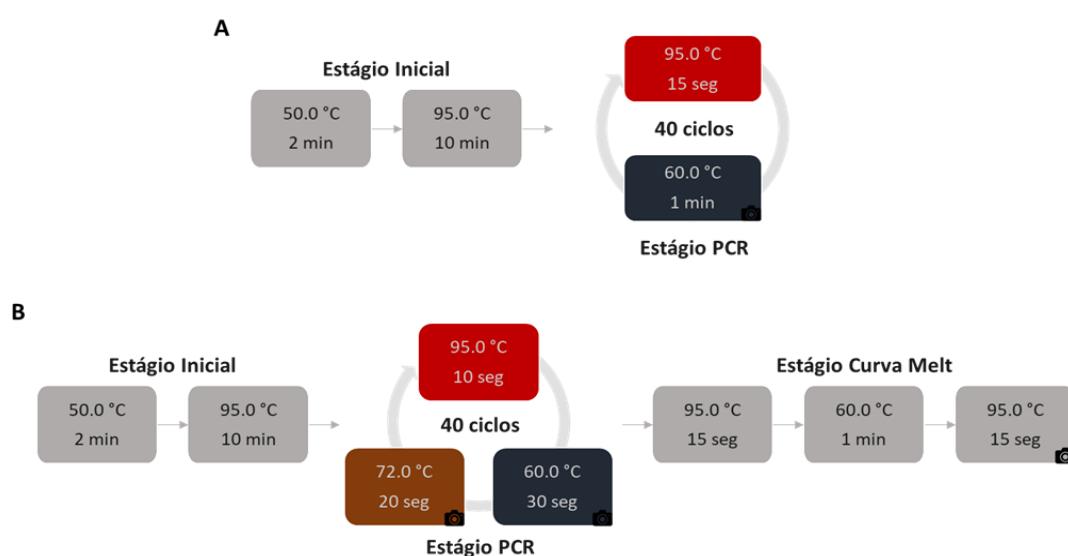


Figura 4.4. Ciclagem das reações para RT-qPCR. (A) Esquema de ciclagem utilizada para as reações com as sondas *TaqMan*. (B) Etapas realizadas na ciclagem para reações com o corante *BRYT Green*.

4.3.5.2 Ensaio com agentes intercalantes de fita-dupla

Para esta metodologia foi utilizado o ensaio comercial *GoTaq qPCR master mix*, no qual o *BRYT Green* é o agente intercalante de fita-dupla. Dessa forma, o cDNA é desnaturado e os oligonucleotídeos específicos para a região de interesse se anelam a fita molde. Conforme a nova fita vai sendo sintetizada, moléculas deste agente se intercalam nos sulcos formados, emitindo fluorescência que será capturada pelo equipamento.

Ambos os genes *IKZF1* e *18s rRNA* foram quantificados por este método (Tabela 4.3). As amostras foram realizadas em duplicatas e o Ct (*cycle threshold*), detalhado no tópico abaixo, utilizado foi a média daqueles obtidos em cada amostra. A concentração utilizada de cDNA foi de 12,5 ng, com *GoTaq qPCR master mix*, água livre de nucleases e oligonucleotídeos de sentido direto e reverso, de *IKZF1* para reação 1 e *18s rRNA* para reação 2. O volume final em cada poço foi de 30 µL e a reação foi submetida a ciclagem esquematizada na figura 4.4B.

Tabela 4.3. Sequência dos oligonucleotídeos usados na reação de RT-qPCR.

Oligonucleotídeos	Sequência
<i>18s rRNA</i> E1.F1	5' AAACGGCTACCACATCCAAG 3'
<i>18s rRNA</i> E1.R1	5' CCTCCAATGGATCCTCGTTA 3'
<i>IKZF1</i> E4.S ^a	5' GATGAAGAGAATGGGCGTGC 3'
<i>IKZF1</i> E4.AS ^a	5' CATTCAATTTCTCTCCCGAGGC 3'

^aObtidos de Iacobucci et al, 2009. E: éxon; F e S: sentido direto; R e AS: sentido reverso.

4.3.5.3 Determinação do Delta Ct (ΔCt)

Todas as reações de RT-qPCR foram conduzidas no aparelho *ViiA 7 Real-Time PCR System* (Applied Biosystems) e processadas pelo programa *QuantStudio Real-Time PCR* versão 1.2 (Applied Biosystems) (Figura 4.5). A partir do gráfico gerado, um *threshold*, ou seja, um limiar de detecção do início da fase exponencial de amplificação foi determinado para cada reação. As reações derivadas do ensaio por sondas hidrolisáveis foram submetidas a um *threshold* de 0,03, enquanto as derivadas do ensaio por agente ligante ao DNA a um *threshold* de 0,14. Desta forma, o valor do Ct, ou seja, o número de ciclos necessários para o início da amplificação, pôde ser determinado para cada amostra, permitindo a quantificação

relativa do gene de interesse. Após a correção pelo Ct do gene de referência ($\Delta Ct = Ct_{\text{gene de interesse}} - Ct_{\text{gene de referência}}$), o valor utilizado para análise de expressão foi o de $1/\Delta Ct$ (Schmittgen & Livak, 2008).



Figura 4.5. Exemplo de reação processada pelo programa *QuantStudio Real-Time PCR*. A figura demonstra um exemplo de reação, usando *TaqMan*, analisada pelo programa. À esquerda: gráfico gerado, em função logarítmica, pelo sinal de fluorescência obtido em cada ciclo com *threshold* determinado (linha azul). À direita: representação do posicionamento de cada alvo e amostra nesta placa.

4.4 VALIDAÇÃO DOS RESULTADOS

A validação dos resultados obtidos neste estudo foi realizada em uma coorte independente de pacientes através de análises *in silico*, utilizando dados disponibilizados *online* pela iniciativa *Therapeutically Applicable Research to Generate Effective Treatments* (TARGET) (<https://ocg.cancer.gov/programs/target/data-matrix>). Esta etapa foi realizada em conjunto com a Dra. Mariana Boroni, no Laboratório de Bioinformática e Ciência Computacional (LBBC) do INCA.

Os dados foram obtidos do projeto *TARGET Acute Lymphoblastic Leukemia*, que tem como objetivo elucidar a caracterização molecular para determinar as mudanças genéticas responsáveis pela iniciação e progressão das LLAs pediátricas difíceis de serem tratadas. Este projeto foi gerado em duas fases distintas: a) fase

piloto (*ALL Phase I*); e b) fase de expansão, subdividida em descoberta e validação (*ALL Phase II*). Apenas as fases piloto e de descoberta foram utilizadas para a etapa de validação do presente projeto, devido a disponibilidade dos dados.

Através do arquivo *.xlsx* contendo as características clínicas dos pacientes, foi possível selecionar apenas os casos diagnosticados com LLA-CPB e excluir aqueles com síndrome de Down. Todas as amostras utilizadas nas metodologias abaixo foram derivadas de MO e/ou SP colhidas ao diagnóstico. Os detalhes sobre a manipulação dos materiais e dos dados gerados estão depositados no endereço eletrônico do projeto, com o título *TARGET Project Experimental Methods*.

4.4.1 Determinação do status de *IKZF1*

O passo seguinte foi determinar o status de *IKZF1* nas amostras dos pacientes. Para isso, foram utilizados dados abertos de *whole genome sequencing* (WGS), possibilitando a identificação de alterações no número de cópias. Cada arquivo MAF foi gerado pela análise da amostra de diagnóstico pareada com a amostra referência (de remissão). Selecionamos apenas os arquivos para número de cópias cromossômicas, disponíveis por paciente no formato *.tsv*. O material era composto por colunas anotadas com: cromossomo analisado; início e final da região; classificação da ploidia; genes presentes na região; entre outras informações.

Estes arquivos foram trabalhados utilizando o ambiente estatístico R. Através de uma busca com o comando *grep* pelo parâmetro '*IKZF1*', selecionamos todos os casos que apresentavam alguma alteração nesse gene. Até este momento, os pacientes foram identificados, em uma nova coluna, como *IKZF1* selvagem, deletado ou amplificado. Entretanto, os casos com amplificação de *IKZF1* foram excluídos para a etapa de análises de expressão.

Para determinar qual o tipo de deleção presente nos indivíduos com *IKZF1* deletado, foi utilizado o genoma de referência GRCh37 em arquivo *.gtf*. Extraímos do mesmo as regiões correspondentes ao início e final de cada éxon do transcrito *IKZF1-001*. A tabela contendo essas informações, assim como aquela com as regiões afetadas de *IKZF1* de todos os pacientes, foram transformadas para classe *GRanges* usando o pacote *GenomicRanges* versão 1.26.4 (Bioconductor) a fim de serem submetidas ao comando *findOverlaps*. Dessa forma, cada região pode ser

identificada e os pacientes classificados pelo tipo de deleção em *IKZF1*: Δ haplo ou Δ dn (vide tópico 4.2).

4.4.2 Expressão relativa de genes (RNA-seq)

Para determinar a expressão dos genes de interesse, os dados selecionados foram de RNA-seq obtidos pelo sequenciamento do cDNA na plataforma HiSeq 2000 (Illumina), no modo '*paired-end*'. As *reads* sequenciadas foram alinhadas ao genoma de referência GRCh37. Os dados pré-processados da quantificação gênica foram disponibilizados em arquivos *.txt* para cada paciente.

Utilizando novamente o ambiente estatístico R, foi possível extrair apenas a contagem bruta das *reads*, assim como os genes quantificados por cada uma delas, de todos os pacientes e depositar em nova tabela. Em seguida, através do pacote *DESeq2* versão 1.14.1 (Bioconductor), foi realizada a análise de expressão diferencial pelo comando *DESeq*, após filtro de contagens reduzidas de *reads* (no mínimo 5). Os dados de entrada foram decorrentes da tabela com a contagem de *reads* e daquela com dados clínico-laboratoriais e status de *IKZF1* de cada paciente, mensurando o efeito desta última condição. Os valores normalizados foram obtidos pelo comando *counts* e usados para as análises de expressão.

4.5 ANÁLISES ESTATÍSTICAS

Todos os gráficos e análises apresentados neste estudo também foram realizados utilizando o ambiente estatístico R.

Dessa forma, os *heatmaps* foram gerados com os dados de expressão utilizando o pacote *Pheatmap* versão 1.0.8 (CRAN-R). Já os demais gráficos foram gerados por comandos a partir do pacote *Ggplot2* versão 2.2.1 (CRAN-R).

Através do pacote *Ggpubr* versão 0.1.6 (CRAN-R), o teste de Mann-Whitney foi aplicado para análise comparativa dos dados de expressão de todos os genes estudados, seja com os dados clínico-laboratoriais dos pacientes ou com o status de *IKZF1*. Os resultados com P-valor menor do que 0,05 foram considerados estatisticamente significantes.

A SG e SLE foram estimadas através do método de Kaplan-Meier nos casos da coorte de validação, de acordo com os critérios de estratificação de risco do NCI,

status de *IKZF1* e expressão dos genes estudados. Para SG, consideramos o tempo, em meses, desde o diagnóstico até o óbito, enquanto que para SLE, o tempo foi determinado até o primeiro evento, sendo este o óbito, recaída ou neoplasia maligna secundária. Todas as análises foram realizadas pelo pacote *Survival* versão 2.41-3 (CRAN-R), com resultados considerados estatisticamente significantes diante de P-valor menor que 0,05.

5 RESULTADOS

5.1 CARACTERIZAÇÃO DOS PACIENTES

Os pacientes das coortes de descoberta e validação foram caracterizados de acordo com os dados clínico-laboratoriais, como: idade (meses), sexo, leucometria ($\times 10^9$ células/L), subtipo citogenético-molecular e classificação de risco prognóstico (Moorman, 2016), além do status de *IKZF1*.

5.1.1 Coorte de Descoberta

No período deste estudo (2004-2017), 540 pacientes foram diagnosticados com LLA-CPB. A maioria destes tinham idade inferior à 120 meses (80,7%) e eram do sexo masculino (56,5%). Além disso, 70% apresentavam leucometria inferior ou igual a 50 bilhões de células por litro e 47,0% puderam ser classificados como *B-other* (Tabela 5.1).

Os casos diagnosticados entre os anos de 2004 a 2011 foram previamente caracterizados quanto ao status de *IKZF1* (Barbosa et al, 2015). A outra parte dos casos também foi submetida a técnica de MLPA (*SALSA MLPA P335 ALL-*IKZF1**) e em caso de deleção, confirmada por MLPA (*SALSA MLPA P202 *IKZF1**) ou MP-PCR. Esta etapa foi realizada juntamente com a Dra. Thayana da Conceição Barbosa, no PHOP, durante meu projeto de iniciação científica e no presente estudo. Ao final das análises, 22,6% (122/540) dos pacientes apresentaram deleção em *IKZF1*, sendo 79,5% (97/122) deleção do tipo haploinsuficiente (Δ haplo) e 20,5% (25/122) deleção do tipo dominante-negativa (Δ dn). Os demais casos (418/540) foram considerados do tipo selvagem.

Após a determinação do status de *IKZF1* e conforme os critérios definidos na metodologia, 72 amostras foram selecionadas com a finalidade de compor a coorte de descoberta e prosseguir com o fluxograma do estudo. Sendo assim, 37,5% (27/72) dos casos eram do tipo selvagem, enquanto 34,7% (25/72) eram Δ haplo e 27,8% (20/72) eram Δ dn. A maioria dos pacientes na coorte de descoberta tinha idade inferior a 120 meses (70,8%), era do sexo masculino (55,6%), com contagem leucocitária inferior ou igual a 50 bilhões de células por litro (59,7%) e foram classificados pelo subtipo citogenético-molecular *B-other* (64,5%) (Tabela 5.1).

Tabela 5.1. Características dos pacientes com LLA-CPB.

Variáveis	Inicial (N=540) ^a		Coorte de Descoberta (N=72) ^b		Coorte de Validação (N=129) ^c	
	n	%	n	%	n	%
Idade (meses)						
<120	436	80,7	51	70,8	78	60,5
≥120	104	19,3	21	29,2	51	39,5
Sexo						
Masculino	305	56,5	40	55,6	67	51,9
Feminino	235	43,5	32	44,4	62	48,1
Leucometria (x10⁹/L)						
≤50	378	70,0	43	59,7	78	60,5
>50	162	30,0	29	40,3	51	39,5
Subtipo citogenético-molecular^d						
Alta hiperdiploidia	61	20,2	4	6,5	15	11,6
<i>ETV6-RUNX1</i>	52	17,2	7	11,3	11	8,5
<i>TCF3-PBX1</i>	38	12,6	2	3,2	14	10,9
<i>B-other^e</i>	142	47,0	40	64,5	75	58,1
<i>BCR-ABL1</i>	11	3,6	3	4,8	5	3,9
<i>KMT2A-r</i>	31	10,3	6	9,7	4	3,1
<i>TCF3-HLF</i>	NA ^f	NA ^f	NA ^f	NA ^f	1	0,8
<i>iAMP21</i>	NA ^f	NA ^f	NA ^f	NA ^f	4	3,1
hipodiploidia	3	1,0	0	0,0	0	0,0
ND ^g	202	-	10	-	-	-
Classificação de risco prognóstico^h						
Baixo	113	20,9	11	15,3	26	20,2
Intermediário	180	33,3	42	58,3	89	69,0
Alto	45	8,3	9	12,5	14	10,9
ND ^g	202	37,4	10	13,9	-	-
Status de <i>IKZF1</i>						
Selvagem	418	77,4	27	37,5	100	77,5
Deleção Haploinsuficiente	97	18,0	25	34,7	20	15,5
Deleção Dominante-Negativo	25	4,6	20	27,8	9	7,0

^aTodos os casos no período definido do estudo; ^bCasos selecionados para o estudo; ^cSérie independente de casos obtidos pelo projeto TARGET; ^dOs percentuais foram calculados pelo número total de pacientes classificados (N=302, N=62 e N=129, respectivamente); ^eNegativo para as alterações citogenética-moleculares avaliadas; ^fNão avaliado; ^gNão definido; ^hClassificados de acordo com o subtipo citogenético-molecular (Moorman, 2016).

5.1.2 Coorte de Validação

Foram disponibilizados, pelo projeto *TARGET Acute Lymphoblastic Leukemia*, dados de 1035 pacientes nas fases selecionadas (*vide* tópico 4.4), mas apenas aqueles com doença decorrente de células precursoras B e sem Síndrome de Down

(n=957) prosseguiram para as próximas etapas. Destes, somente 149 indivíduos apresentaram dados disponibilizados em ambas as metodologias de WGS e RNA-seq, sendo 132 realizados com amostras colhidas no diagnóstico.

Após extração dos dados de WGS, foi possível classificar estes pacientes de acordo com as alterações no número de cópias no gene *IKZF1*. Aqueles com deleção bialélica (n=1) ou amplificação (n=2) foram excluídos neste estudo. Dessa forma, a coorte de validação foi composta por 129 crianças, com 22,5% (29/129) dos casos apresentando deleção em *IKZF1*, sendo 69,0% (20/29) Δ haplo e 31,0% (9/29) Δ dn, enquanto 77,5% (100/129) foram considerados do tipo selvagem.

Assim como na coorte de descoberta, a maioria dos pacientes tinha idade inferior a 120 meses (60,5%), era do sexo masculino (51,9%), com contagem leucocitária inferior ou igual a 50 bilhões de células por litro (60,5%) e foram classificados pelo subtipo citogenético-molecular *B-other* (58,1%) (Tabela 5.1).

5.2 NÍVEIS DE EXPRESSÃO DOS GENES

A quantificação dos níveis de expressão relativa dos genes obtida por RT-qPCR e RNA-seq das coortes de descoberta e validação, respectivamente, resultou no perfil demonstrado na figura 5.1. É possível observar que os genes envolvidos na linfopoiese B (*IKZF1*, *RAG1*, *EBF1*, *BTG1*) estão mais expressos quando comparados aos genes *GRB10* (interação com receptores tirosina quinase) e *FIGNL1* (reparo de DNA).

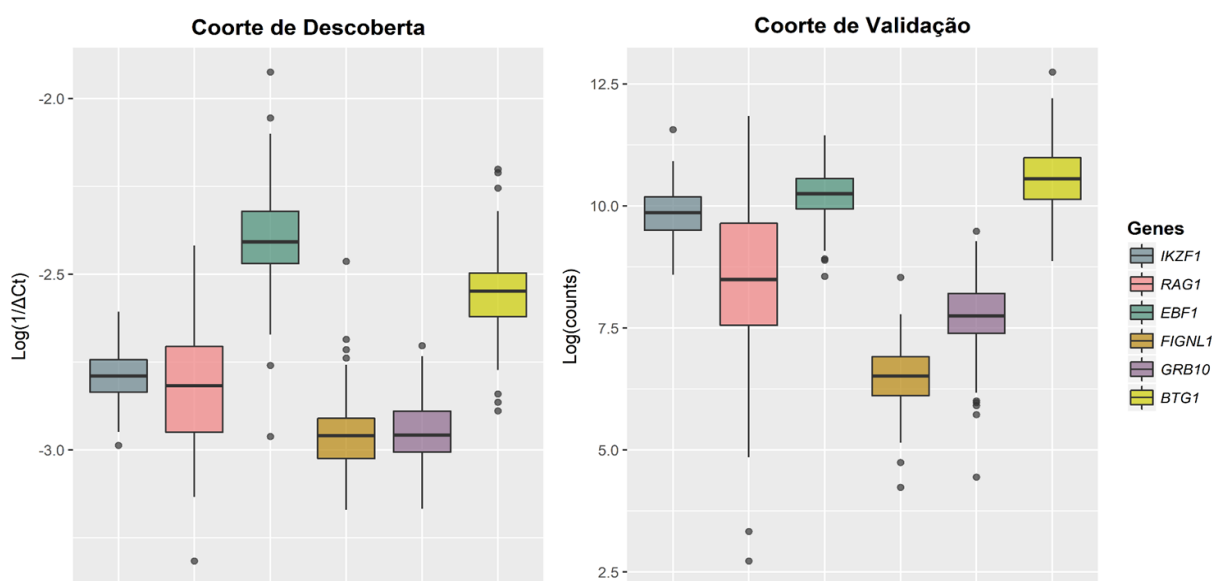


Figura 5.1. Níveis de expressão dos genes nas coortes de descoberta e validação.

Ao associar os níveis de expressão relativa do gene *IKZF1* com os genes-alvo, observou-se uma correlação positiva, porém fraca, entre eles, com exceção de *EBF1* e *BTG1* (Figura 5.2).

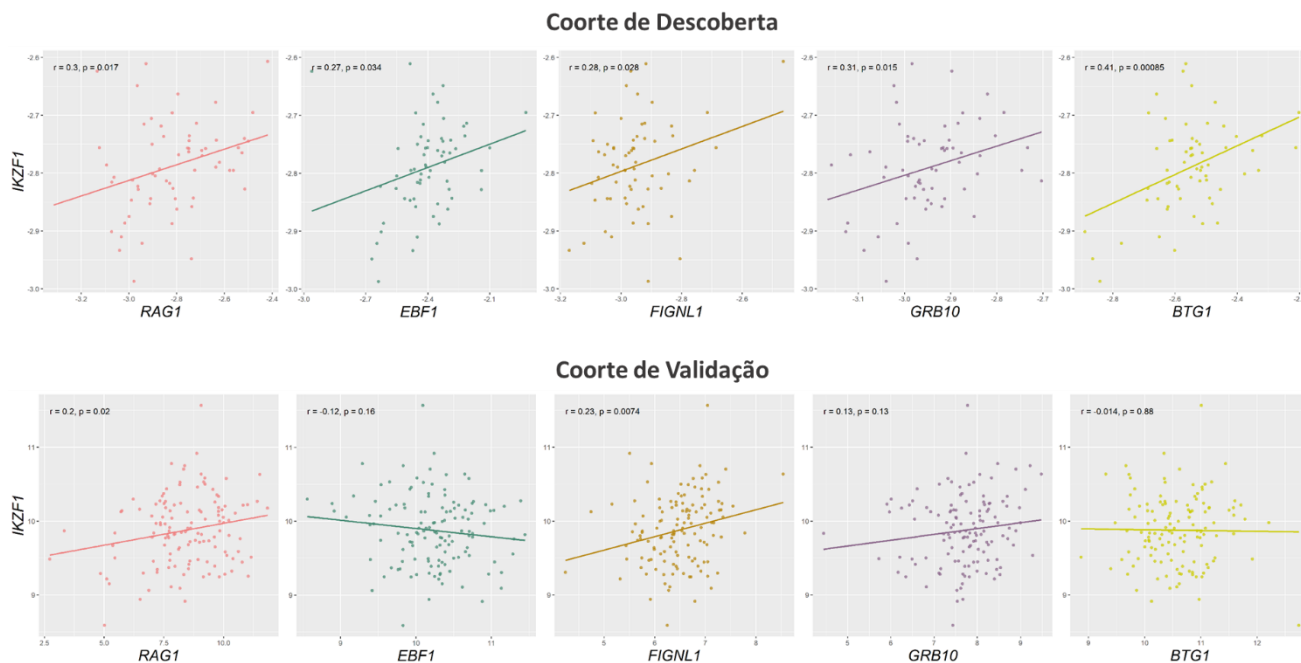


Figura 5.2. Correlação entre a expressão do gene *IKZF1* e os genes-alvo nas coortes de descoberta e validação.

O perfil de expressão gênica foi analisado de acordo com as características clínico-laboratoriais, como: status de *IKZF1*, risco prognóstico de acordo com os subtipos citogenético-moleculares, leucometria, gênero e idade. Os resultados foram demonstrados graficamente através de *heatmap* (Figura 5.3). Os casos foram agrupados por correlação e foi observado um painel bastante heterogêneo, não sendo possível associar os perfis encontrados com as características avaliadas.

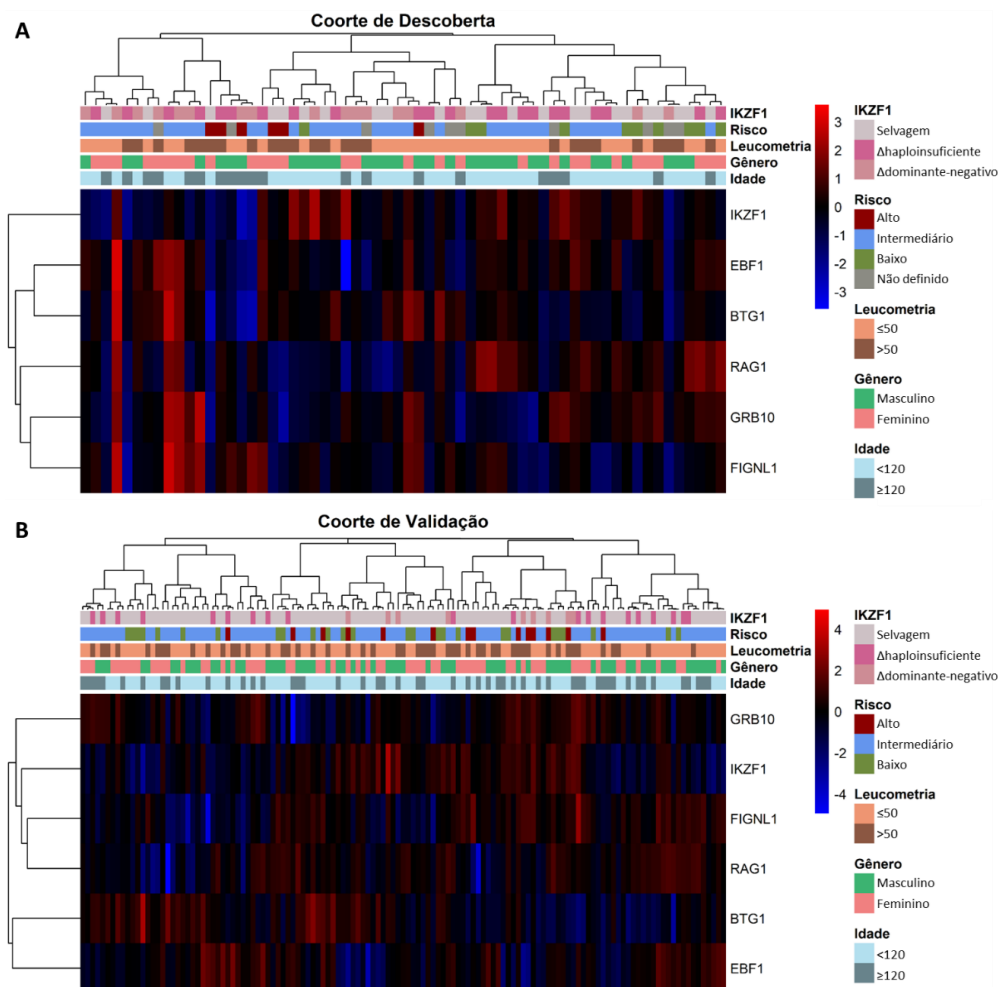


Figura 5.3. Perfil de expressão gênica entre os pacientes, de acordo com suas características. *Heatmap* representativo da expressão gênica em cada paciente (A) na coorte de descoberta e (B) na coorte de validação.

5.3 ANÁLISES DE EXPRESSÃO VERSUS DADOS CLÍNICO-LABORATORIAIS

As análises de expressão de cada gene-alvo, de acordo com as características clínico-laboratoriais, foram realizadas utilizando apenas os pacientes considerados selvagem para o gene *IKZF1*.

5.3.1 Idade

De acordo com a idade (Figura 5.4), observamos que pacientes com idade inferior a 120 meses apresentaram mediana de expressão de *IKZF1* aumentada quando comparados àqueles com idade igual ou maior a 120 meses ($p=0,048$). Entretanto, este resultado não foi validado pela nossa segunda coorte, onde ambos os grupos demonstraram níveis de expressão semelhantes.

Os genes *RAG1*, *EBF1*, *FIGNL1*, *GRB10* e *BTG1* não apresentaram diferenças estatisticamente significantes entre os dois grupos de idade.

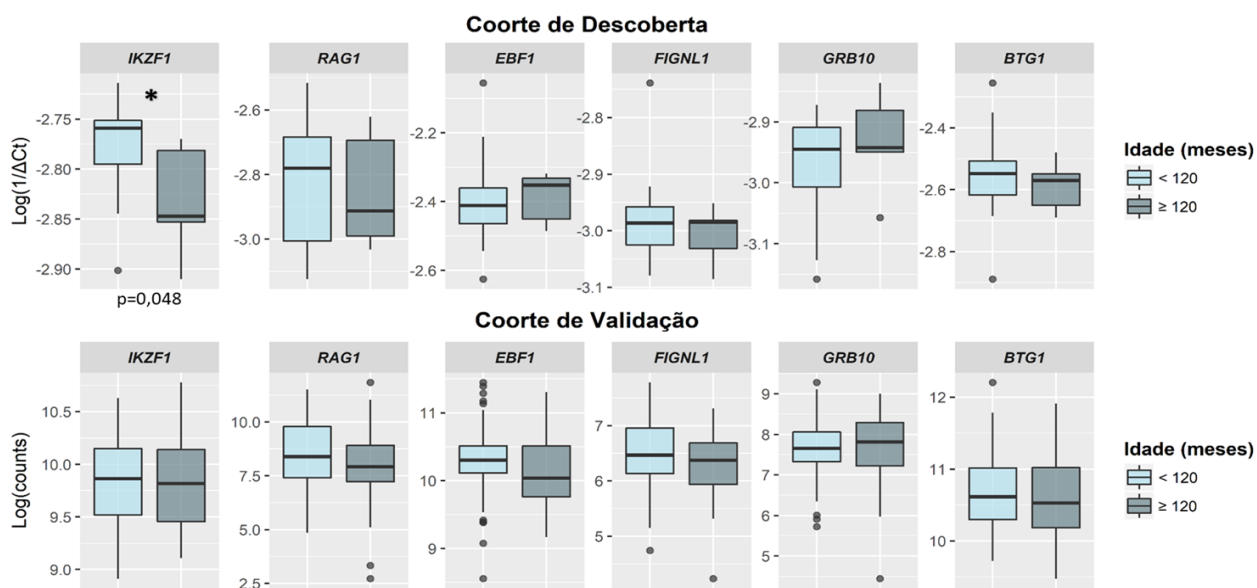


Figura 5.4. Análises de expressão entre os pacientes *IKZF1* selvagem nas coortes de descoberta e validação, de acordo com a idade.

5.3.2 Sexo

Foi observado que pacientes do sexo feminino apresentam maior mediana de expressão de *EBF1* quando comparados aos pacientes do sexo masculino ($p=0,02$). Contudo, este resultado não foi corroborado na coorte de validação. Para a expressão dos demais genes, não observamos diferenças estatisticamente significantes (Figura 5.5).

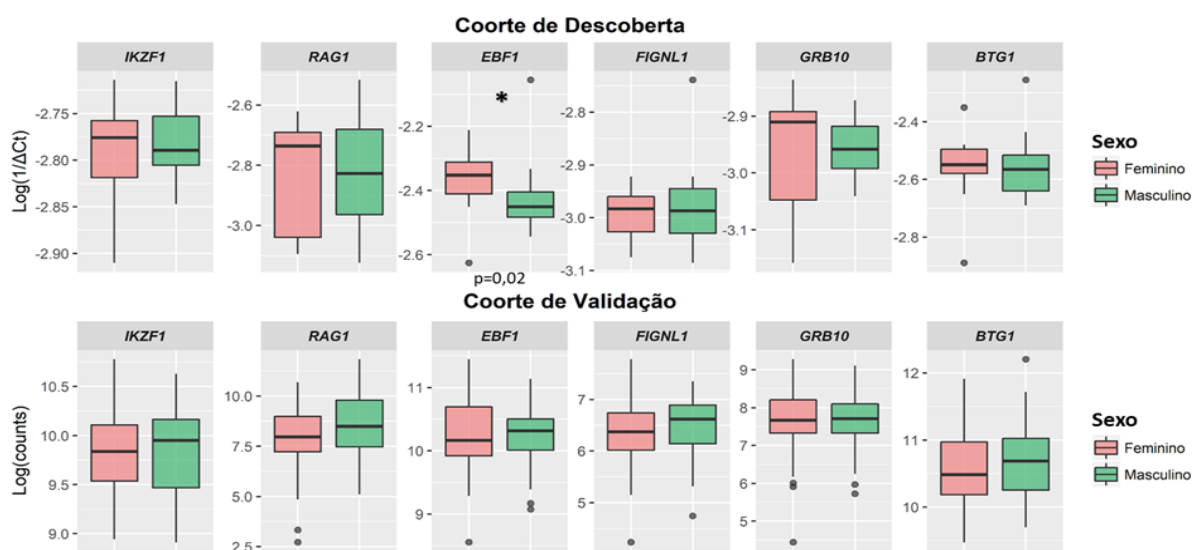


Figura 5.5. Análises de expressão entre os pacientes *IKZF1* selvagem nas coortes de descoberta e validação, de acordo com o sexo.

5.3.3 Leucometria

De acordo com a leucometria dos pacientes ao diagnóstico (Figura 5.6), não foi observada diferença estatisticamente significante.

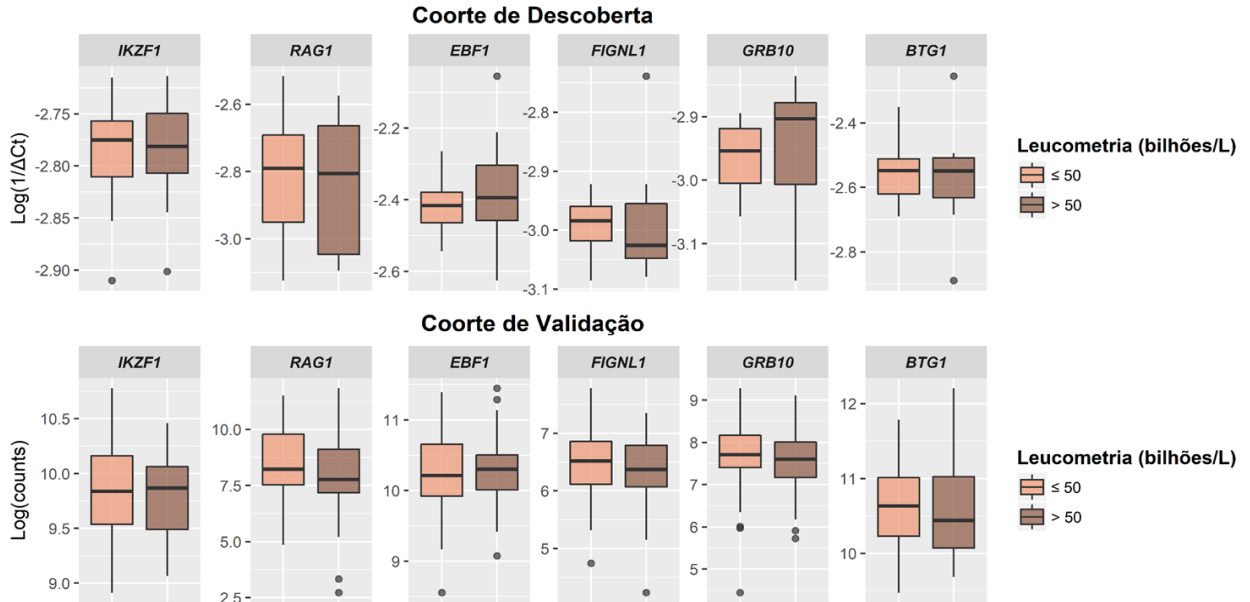


Figura 5.6. Análises de expressão entre os pacientes *IKZF1* selvagem nas coortes de descoberta e validação, de acordo com a leucometria.

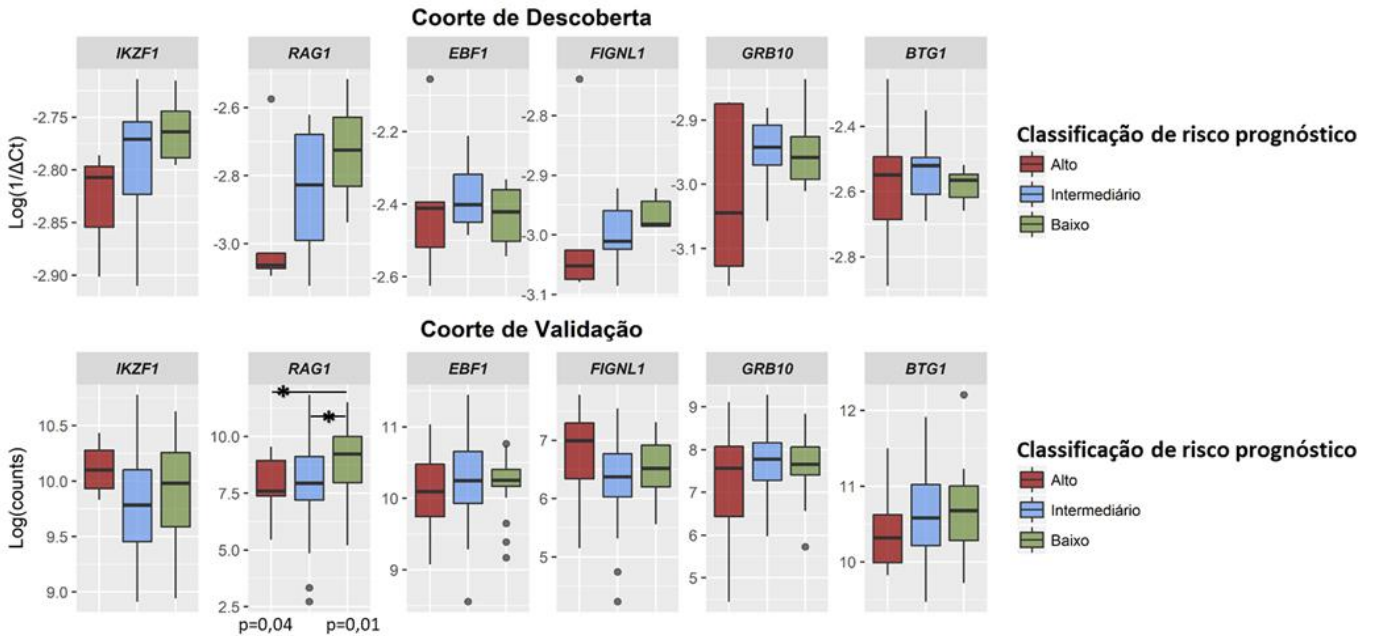


Figura 5.7. Análises de expressão entre os pacientes *IKZF1* selvagem nas coortes de descoberta e validação, de acordo com a classificação de risco prognóstico. São considerados como de alto risco pacientes com as alterações *KMT2A-r*, *BCR-ABL1*, *TCF3-HLF*, *iAMP21* e hipodiploidia; risco intermediário, pacientes com *TCF3-PBX1* e *B-other*; e baixo risco, pacientes com alta hiperdiploidia e *ETV6-RUNX1*.

5.3.4 Classificação de risco prognóstico

As análises foram realizadas também de acordo com a classificação de risco prognóstico (Figura 5.7). Para o gene *RAG1*, o grupo de baixo risco apresentou uma tendência de mediana de expressão maior quando comparado tanto com o grupo de risco intermediário quanto com o de alto risco. Este resultado foi validado pela segunda coorte (p=0,01 e 0,04, respectivamente).

Não observamos diferença estatisticamente significativa entre os grupos para os genes *IKZF1*, *EBF1*, *FIGNL1*, *GRB10* e *BTG1*.

5.4 ANÁLISES DE EXPRESSÃO VERSUS STATUS DE IKZF1

Em um primeiro momento, as análises de expressão foram comparadas entre os grupos com *IKZF1* selvagem e deletado (Figura 5.8), entretanto, não observamos diferença estatisticamente significativa para os genes *IKZF1*, *EBF1* e *BTG1*.

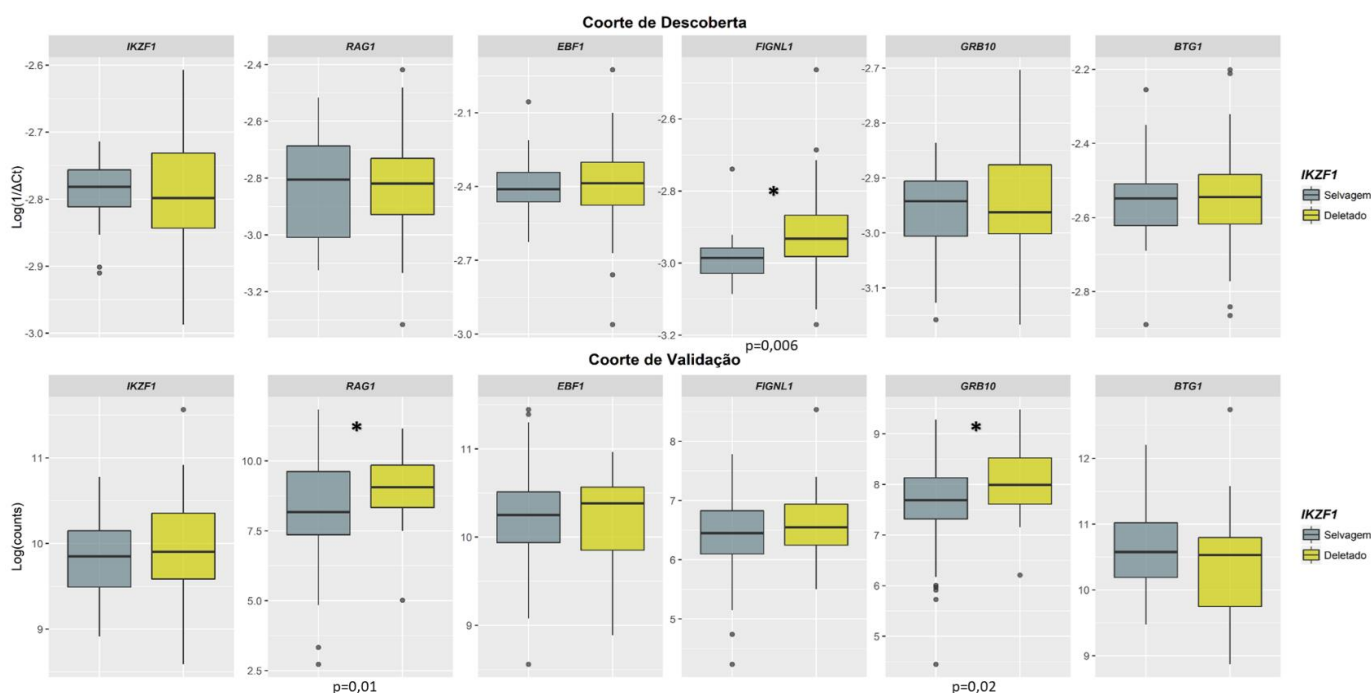


Figura 5.8. Análises de expressão entre os pacientes com *IKZF1* selvagem e deletado nas coortes de descoberta e validação.

Embora a expressão dos genes *RAG1* e *GRB10* apresente este mesmo cenário na coorte de descoberta, foi observado uma mediana menor no grupo selvagem quando comparado ao grupo deletado nos indivíduos da coorte de validação (p=0,01 e 0,02, respectivamente). Para o gene *FIGNL1*, a associação foi

observada na coorte de descoberta, onde este gene teve menor expressão no grupo selvagem do que no grupo deletado ($p=0,006$). Já na coorte de validação, foi possível observar apenas uma tendência de diferença entre os grupos.

Em um segundo momento, estes pacientes com *IKZF1* deletado foram estratificados de acordo com o tipo de deleção: Δ haplo e Δ dn, e as análises de expressão estão demonstradas na figura 5.9. Assim como na primeira análise, os genes *EBF1* e *BTG1* não apresentaram diferenças estatisticamente significantes entre os grupos. Já para o gene *IKZF1*, o grupo com Δ dn teve maior expressão que o grupo com Δ haplo ($p<0,001$) e selvagem ($p<0,0001$) na coorte de validação.

Para a expressão dos genes *RAG1* e *GRB10*, foi observado mediana menor no grupo selvagem quando comparado ao grupo com Δ haplo na coorte de validação ($p<0,0001$ e $p=0,04$, respectivamente). Também observamos que este grupo apresentou menor expressão de *FIGNL1*, quando comparado com ambos os grupos Δ haplo ($p=0,002$) e Δ dn ($p<0,001$).

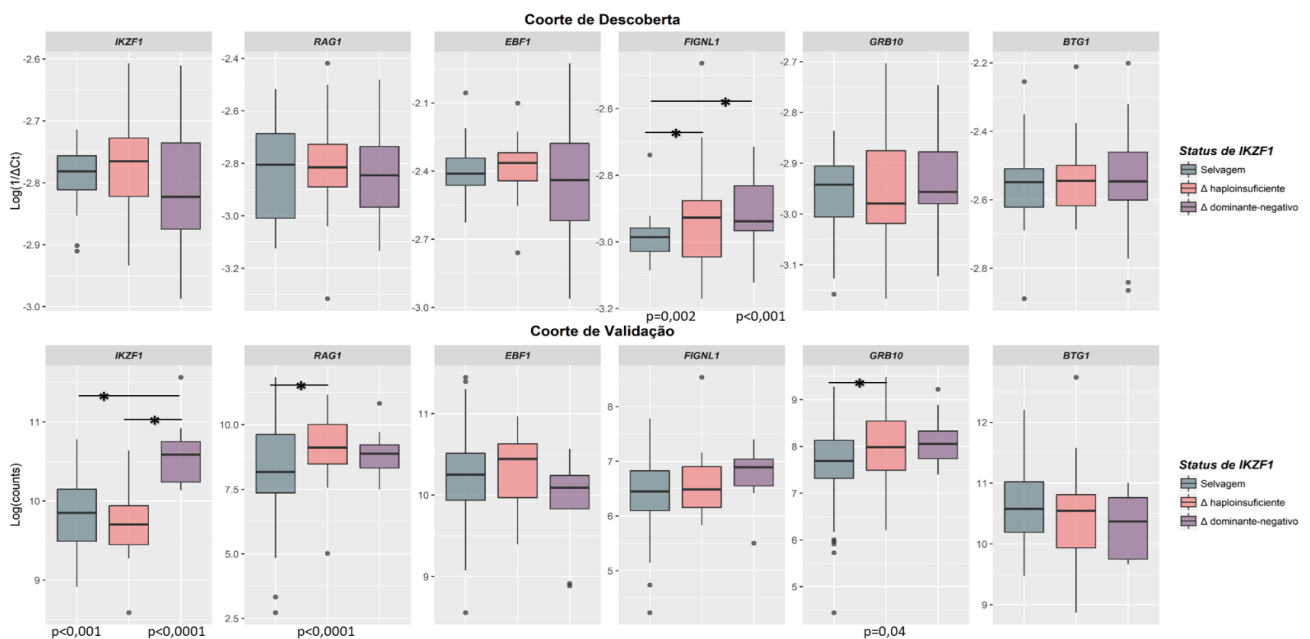


Figura 5.9. Análises de expressão de acordo com o status de *IKZF1* nas coortes de descoberta e validação.

5.5 ANÁLISES DE SOBREVIDA

A SG e a SLE foram avaliadas com os dados de seguimento disponibilizados de 118 pacientes incluídos na coorte de validação. Sabendo que estes indivíduos são estratificados de acordo com os critérios do NCI para definir o protocolo de tratamento, analisamos estes parâmetros e não foi observado diferença significativa

entre os grupos para SG ($p=0,55$, mediana de 38,5 meses e 48,5 meses para risco alto e padrão, respectivamente). Entretanto, os pacientes classificados como de risco padrão apresentaram pior SLE quando comparados àqueles de risco alto ($p=0,017$) (Figura 5.10).

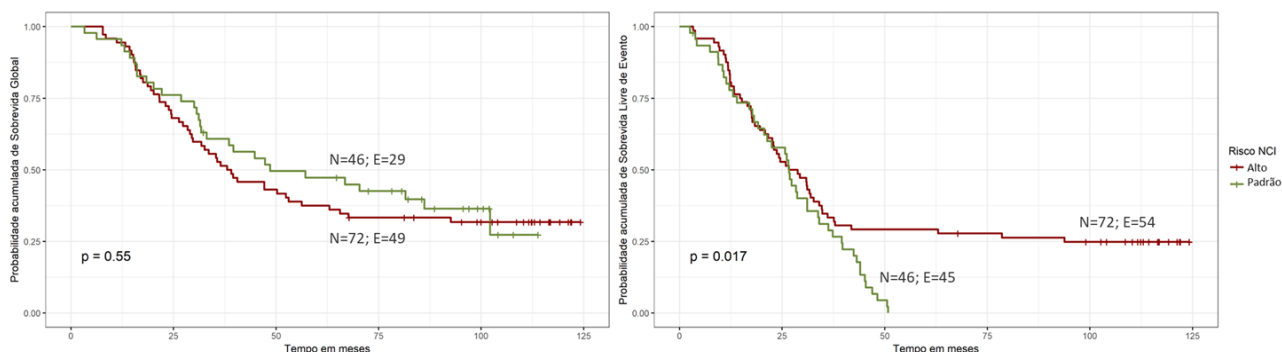


Figura 5.10. Análise de sobrevida global e livre de eventos dos pacientes de acordo com o risco definido pelos critérios do NCI. N = número de pacientes analisados; E= número de pacientes acometidos pelo evento.

Na figura 5.11, estão os gráficos representativos das análises de sobrevida de acordo com o status de *IKZF1*. Observamos maior SG entre os pacientes selvagem para o gene *IKZF1* (48,5 meses) em comparação com o grupo deletado (33,6 meses) ($p=0,014$), seja diante da Δ haplo (35,4 meses) ou Δ dn (29,0 meses) ($p=0,047$). Entretanto, não houve diferença significativa na SLE entre estes grupos.

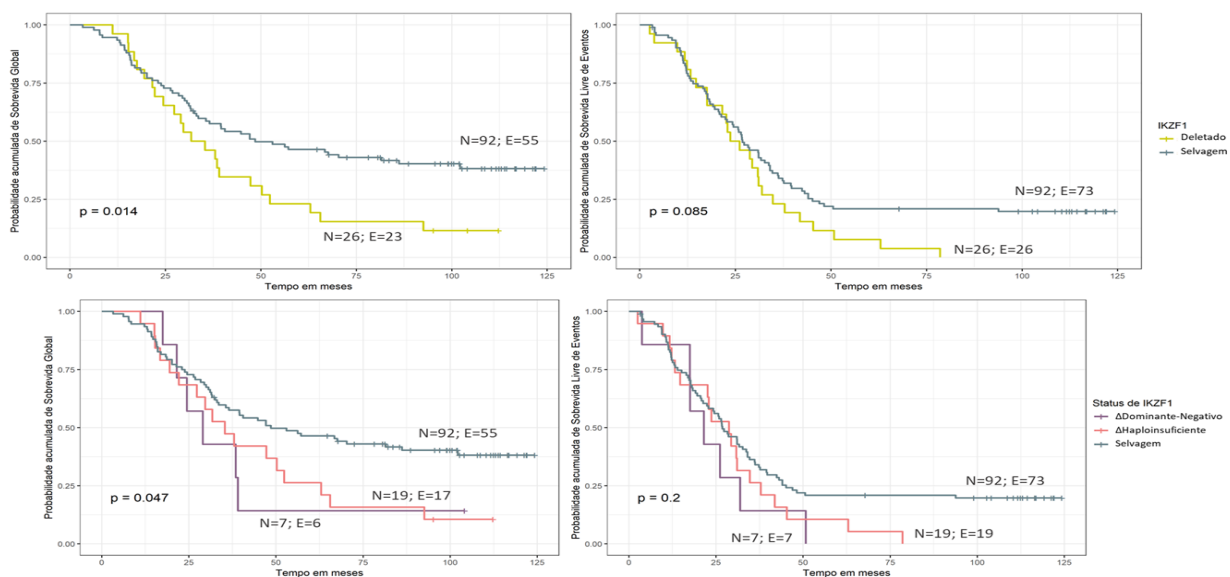


Figura 5.11. Análise de sobrevida global e livre de eventos dos pacientes de acordo com o status de *IKZF1*. N = número de pacientes analisados; E= número de pacientes acometidos pelo evento.

Ao analisar a alta (acima da mediana) e baixa (abaixo da mediana) expressão dos genes, foi possível observar que pacientes com baixa expressão de *EBF1* tem maior SG (63,0 meses) quando comparado com alta expressão (33,1 meses) ($p=0,042$). Para indivíduos considerados com baixa expressão de *GRB10*, observou-se o contrário. Estes apresentaram menor SG (38,5 meses) em relação aqueles com alta expressão (66,8 meses) ($p=0,035$). Este cenário também foi demonstrado pela análise de SLE (20,8 meses vs 31,9 meses, $p=0,001$). Para os demais genes, não houve diferença no prognóstico entre os grupos (Figura 5.12).

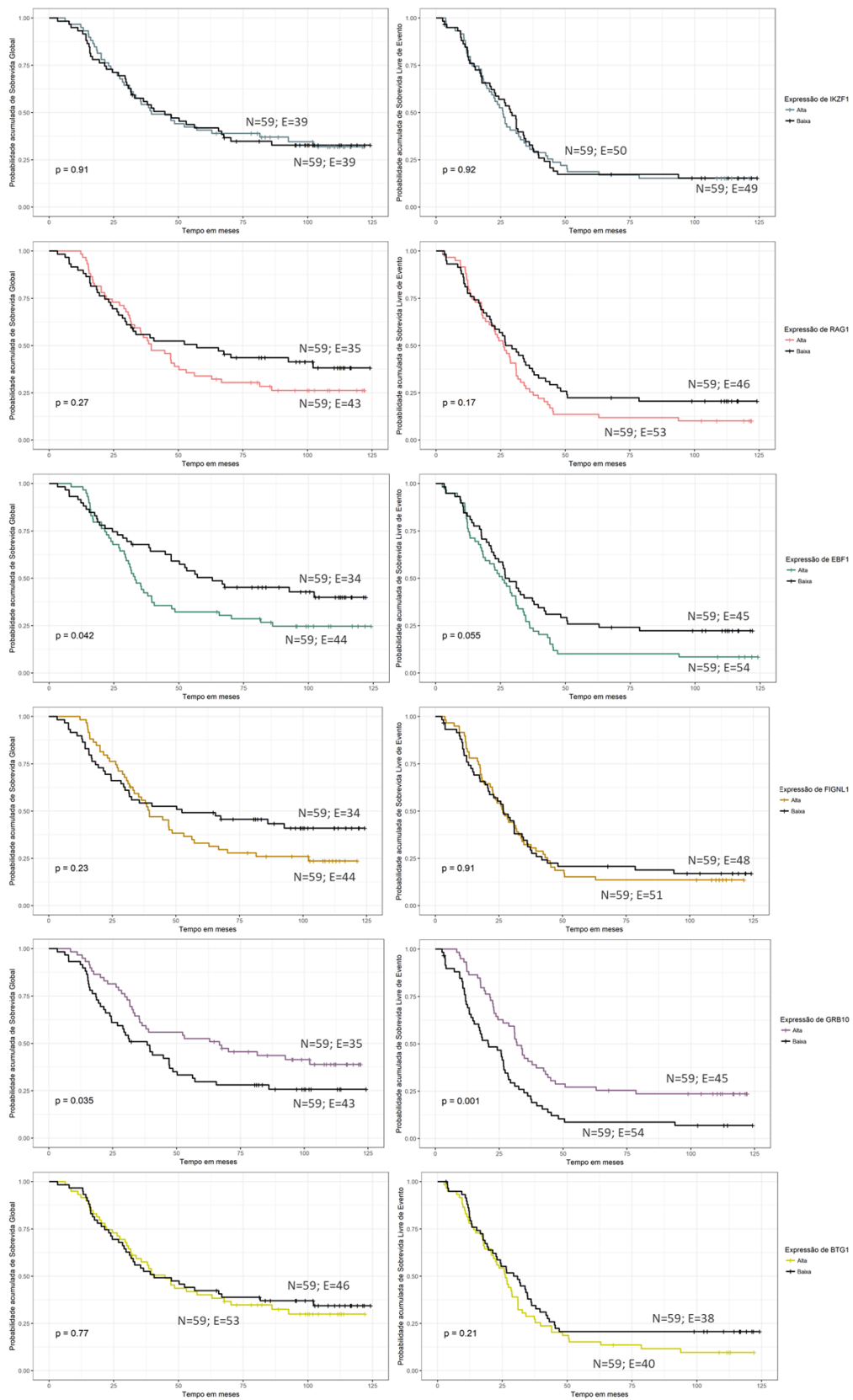


Figura 5.12. Análise de sobrevida global e livre de eventos dos pacientes de acordo com a expressão gênica. N = número de pacientes analisados; E= número de pacientes acometidos pelo evento.

6 DISCUSSÃO

A LLA-CPB pode ser caracterizada como uma doença altamente heterogênea, manifestando diferenças em nível clínico, genético e até mesmo demográfico. Diante disso, estudos que buscam maior entendimento sobre esta neoplasia são importantes para melhorias no diagnóstico, estratificação de risco e prognóstico, direcionando os pacientes para o tratamento mais adequado. É possível observar que a maioria das crianças com LLA-CPB apresentam idade inferior a 120 meses, leucometria baixa e são do sexo masculino. Neste estudo, observamos as mesmas características entre as séries de casos investigadas, que demonstraram ser representativas deste subgrupo. Em relação aos subtipos citogenético-moleculares, o conjunto conhecido como *B-other* foi o mais frequente tanto na coorte de descoberta como na coorte de validação, seguido pelas alterações de bom prognóstico (*ETV6-RUNX1* e alta hiperdiploidia). Este resultado deve-se, provavelmente, pela inclusão de casos com deleção em *IKZF1*, frequentes no grupo *B-other*, além disso, podemos considerar as formas variadas de identificação destes indivíduos entre diferentes estudos (Moorman, 2016).

A identificação de alterações secundárias também corrobora com o perfil heterogêneo da LLA-CPB. Decorrentes do comprometimento de genes envolvidos no desenvolvimento linfóide, ciclo celular, regulação transcricional, sobrevivência e proliferação celular, estas alterações contribuem para o processo leucemogênico por diversas vias e podem ser associadas a diferentes subtipos citogenético-moleculares (Mullighan et al, 2007; Moorman, 2016). Como exemplo, observamos as deleções no gene *IKZF1*, presentes recorrentemente nos subtipos *BCR-ABL1* e *BCR-ABL1-like*. Estudos demonstram que a frequência dessas deleções sofre pequenas variações em diferentes coortes, podendo representar cerca de 12-28% dos casos de LLA-CPB (Mullighan et al, 2007; Mullighan et al, 2009; Dorge et al, 2013; Barbosa et al, 2015). Em nossa coorte de validação, obtivemos semelhante frequência desta alteração (aproximadamente 22%). No que diz respeito aos tipos de deleção em *IKZF1*, observamos predomínio de deleção haploinsuficiente (69,0%) e cerca de um terço dos pacientes com deleção dominante-negativa (31,0%), proporção de casos semelhante ao descrito na literatura (Kastner et al, 2013). O mesmo não pode ser analisado na coorte de descoberta, pois foi necessária uma seleção destes casos para as análises subsequentes por tipos de deleção.

IKZF1 é responsável por codificar a proteína Ikaros que atua como fator transcricional regulando o desenvolvimento das células B através da ativação ou repressão de genes-alvo (Arenzana et al, 2015). Os resultados decorrentes das nossas análises de expressão, demonstraram uma heterogeneidade dos níveis deste transcrito entre as coortes. No entanto, devemos considerar as diferenças metodológicas. A técnica de RT-qPCR é frequentemente usada para determinar expressão gênica devido sua sensibilidade, acurácia e rápida execução, no entanto, a qualidade do RNA, a eficiência de amplificação e a escolha do gene de referência podem afetar estes parâmetros. Também é necessário o conhecimento prévio da sequência de interesse para o pareamento dos oligonucleotídeos, que, neste caso, anelam no éxon 4 da isoforma principal de *IKZF1*. Além disso, o cDNA foi obtido a partir do RNA total do material colhido de pacientes com percentual de blastos superior a 30% (coorte de descoberta). O fato de ter sido analisado RNA proveniente de células normais e neoplásicas pode interferir nessa quantificação, visto que, para a técnica de RNA-seq, as moléculas de RNA foram obtidas de pacientes com percentual de blastos superior a 70% (coorte de validação). Esta metodologia, por sua vez, tem sido uma poderosa tecnologia usada para análises de transcriptoma. Os resultados gerados pela RNA-seq oferecem uma aproximação quantitativa dos níveis do transcrito na forma de contagem das *reads*. Estas são pré processadas, mapeadas em um genoma de referência e, por fim, normalizadas para a remoção de vieses técnicos como o comprimento das diferentes moléculas de RNAs e a profundidade do sequenciamento.

As análises de expressão de *IKZF1* demonstraram que não houve diferença entre os pacientes com e sem deleção neste gene. Entretanto, pacientes com Δ dn apresentam maior expressão de *IKZF1* quando comparados com pacientes selvagem ou com Δ haplo na coorte de validação. Este resultado também foi observado por um estudo recente com pacientes pediátricos LLA-CPB *BCR-ABL1-like*. Neste trabalho, eles também demonstraram que casos com Δ haplo apresentaram uma redução da expressão quando comparados aos casos *IKZF1* selvagem (Vitanza et al, 2014).

Com o avanço dos estudos de larga escala, foram observados possíveis efeitos na expressão gênica entre as condições com e sem deleção (Iacobucci et al, 2012; Ferreiros-Vidal et al, 2013; Schwickert et al, 2014; Vitanza et al, 2015). Diante

da alteração em *IKZF1*, o perfil obtido é caracterizado, principalmente, pela regulação negativa de genes importantes para o desenvolvimento das células B (como *RAG1* e *EBF1*), assim como para aqueles envolvidos em vias de reparo do DNA (como *FIGNL1*). Enquanto que genes requeridos no ciclo celular, apoptose e em vias de sinalização (como *GRB10*), são regulados positivamente. Nossos resultados demonstram que os diferentes tipos de deleção encontrados em *IKZF1* podem influenciar este perfil, colaborando para o melhor entendimento das variações na consequência funcional dessas alterações.

Ferreirós-Vidal e colaboradores, através de linhagem celular não leucêmica pré-B, demonstraram que *Rag1* apresenta expressão reduzida diante da isoforma dominante-negativa de Ikaros quando comparada com a condição normal (Ferreiros-Vidal et al, 2013). Sugerimos que diante dessa alteração (Δ dn), *RAG1* deixa de ser regulado devido ao sequestro das isoformas normais de Ikaros pelas deficientes alterando, assim, sua função normal. Entretanto, acreditamos que a redução da expressão não pode ser observada em nossos pacientes, com LLA-CPB e Δ dn em *IKZF1*, devido a mecanismos inerentes a doença que alteram a expressão de *RAG1* por outras vias. Observamos também que os casos da coorte de validação com Δ haplo em *IKZF1* apresentaram maiores níveis de expressão de *RAG1* quando comparados com pacientes com *IKZF1* selvagem. Este resultado pode ser devido a uma diferente regulação de *RAG1* diante da redução de função de Ikaros.

Foi observada uma tendência de expressão diminuída de *EBF1* diante da Δ dn quando comparado com os outros grupos. Apesar de não terem alcançado significância estatística, estes resultados estão de acordo com os dados demonstrados por um estudo realizado em adultos com LLA-B *BCR-ABL1*. Após observar que o gene *EBF1* está regulado negativamente diante da deleção de *IKZF1* nos pacientes, os autores realizaram um ensaio com luciferase onde foi possível demonstrar uma redução dos níveis proteicos de *EBF1* diante da isoforma dominante-negativa de Ikaros (Iacobucci et al, 2012).

Schwickert e colaboradores, ao analisar células pró-B expressando transcritos de *Ikzf1* com deleção do éxon 8 e com transcritos completos, relataram que a expressão de *Figl1* está aumentada diante da função normal de Ikaros (Schwickert et al, 2014). No contexto leucêmico, também observamos expressão diferencial deste gene em pacientes com *IKZF1* selvagem e deletado, entretanto, os níveis

aumentados de *FIGNL1* foram observados diante da deleção, independentemente do tipo, em nossa coorte de descoberta. Não há estudos esclarecendo a função deste gene na leucemia, mas, recentemente, um trabalho envolvendo pacientes com câncer de pulmão de células pequenas relatou que a ativação anormal da via de recombinação homóloga pela superexpressão de *FIGNL1* contribui para a rápida progressão e recaída, além da resistência à quimioterapia, nestes indivíduos (Ma et al, 2017). Dessa forma, é possível que este gene seja modulado por outras vias independentes de *IKZF1* e que o aumento de sua expressão pode contribuir para o cenário agressivo apresentado pela LLA-CPB com *IKZF1* deletado.

Os casos com deleção em *IKZF1* também apresentaram maiores níveis de expressão do gene *GRB10*, especificamente do tipo haploinsuficiente, quando comparados com pacientes com *IKZF1* selvagem. Esse resultado observado na coorte de validação corrobora com os dados apresentados por Vitanza e colaboradores, que demonstraram níveis aumentados deste gene em pacientes pediátricos com LLA-CPB e *IKZF1* deletado (Vitanza et al, 2014).

Embora *BTG1* também seja considerado um fator essencial na linfopoiese, os estudos mencionados acima não listaram a expressão deste gene como alterada diante da deleção em *IKZF1*. Nossos resultados corroboraram esta informação demonstrando que pacientes com deleção em *IKZF1* não apresentam diferença na expressão de *BTG1* quando comparado aos pacientes com *IKZF1* selvagem.

Neste estudo avaliamos a alteração/perda da função Ikaros pela presença de deleções genômicas em *IKZF1*, entretanto, outros fatores podem ser levados em consideração para o melhor entendimento dos achados aqui encontrados. Algumas alterações no mecanismo de *splicing* podem resultar no aumento dos níveis de diferentes isoformas de Ikaros formando heterodímeros que alteram sua função, assim como modificações pós-transducionais em resíduos específicos de Ikaros também podem regular sua atividade (Gomez-del Arco et al, 2005; Popescu et al, 2009). Outro ponto são os fatores remodeladores de cromatina recrutados por essa proteína para regulação da expressão gênica de diversos alvos. Uma vez alterados, estes fatores podem modificar a função do complexo formado resultando em um efeito contrário àquele observado diante das deleções em *IKZF1*. Juntos, estes mecanismos poderiam explicar os achados discordantes entre este trabalho e outros descritos pela literatura.

Com o objetivo de compreender a associação entre a expressão de *IKZF1* e dos genes de interesse neste estudo, nós realizamos uma correlação destes dados através do gráfico de dispersão, sendo possível observar correlação positiva fraca de *IKZF1* com os genes *RAG1*, *FIGNL1* e *GRB10*, enquanto que para *EBF1* e *BTG1* os resultados foram contraditórios entre as coortes.

Em relação às características clínico-laboratoriais dos pacientes, não há muitos estudos correlacionando estes dados com a expressão dos genes estudados no estudo presente. Dessa forma, procuramos estabelecer um perfil de expressão gênica de acordo com estas variáveis e observamos que não há correlação entre os indivíduos, revelando um quadro de heterogeneidade mesmo que de forma ínfima. Através das nossas análises com os casos *IKZF1* selvagem, foi possível ampliar um pouco este conhecimento. Observamos que pacientes com idade inferior a 120 meses apresentam uma tendência a possuir expressão aumentada de *IKZF1*, *RAG1* e *BTG1*, e baixa para *GRB10*, quando comparados àqueles com idade igual ou maior a 120 meses. Já o gene *RAG1* revelou-se com mediana de expressão maior no grupo de baixo risco quando comparado tanto com o grupo de risco intermediário quanto com o de alto risco.

Os dados disponibilizados em endereço eletrônico pelo projeto TARGET nos possibilitaram analisar a SG e SLE dos pacientes estudados de acordo com a expressão destes genes. Primeiramente, classificando estes indivíduos como de alto risco (≥ 120 meses de idade e leucometria alta) e risco padrão (< 120 meses de idade e leucometria baixa), observamos diferença da SLE. Os pacientes de risco padrão apresentaram pior SLE quando comparados àqueles de risco alto, provavelmente indivíduos que deveriam ser tratados por protocolos mais agressivos estão sendo submetidos a terapias brandas levando a falhas na indução, recaídas e até mesmo óbito. Este resultado enfatiza a importância da estratificação de risco de forma adequada antes de iniciar o tratamento. Assim, outros fatores devem ser levados em consideração, como os grupos citogenético-moleculares e alterações secundárias que estão frequentemente correlacionadas com um prognóstico desfavorável. Alguns estudos demonstraram que a deleção em *IKZF1* pode ser considerada um fator de prognóstico adverso independente das características clínico-laboratoriais dos pacientes (Mullighan et al, 2009; Dorge et al, 2013; Schwab et al, 2013). Em nosso estudo foi possível demonstrar a correlação dessas alterações com o

prognóstico, onde a SG dos pacientes com *IKZF1* selvagem foi maior em comparação com o grupo deletado (48,5 vs 33,6 meses), independente do tipo de deleção (Δ haplo: 35,4 meses; Δ dn: 29,0 meses), corroborando com os resultados apresentados previamente (Boer et al, 2016).

O gene *EBF1* é um importante regulador transcricional do desenvolvimento linfóide, sendo observado o aumento de sua expressão de acordo com o estágio maturativo das células B. Alterações neste gene estão associadas a baixa taxa de sobrevida livre de recaída em pacientes com LLA-CPB. Observa-se que a perda de sua função pode contribuir para a progressão da doença e resistência à quimioterapia (Mullighan et al, 2007). Entretanto, ao analisar a expressão dos genes, pacientes com baixa expressão de *EBF1* apresentaram maior SG do que aqueles com alta expressão (63,0 vs 33,1 meses).

Nossos achados também revelaram que indivíduos considerados com baixa expressão de *GRB10* apresentam menor SG e SLE em relação aqueles com alta expressão (38,5 vs 66,8 meses; 20,8 vs 31,9 meses). Na LMA, já foi observada uma correlação parcial da expressão elevada da proteína codificada por *GRB10* com recaída e prognóstico adverso em pacientes com mutação no gene *FLT3* (Kazi & Ronnstrand, 2013). Na LLA-CPB, através de modelo *in vivo* com *knockdown* para a mesma proteína e com a presença da alteração *BCR-ABL1*, observou-se diminuição da proliferação, sobrevivência prolongada e reduzida capacidade de formação de colônia (Illert et al, 2015). Estes dados podem reforçar a ideia do envolvimento do gene *GRB10* na progressão da leucemia, entretanto, acreditamos que a associação com o prognóstico desses pacientes seja devido a outras alterações.

7 CONCLUSÃO

- Deleções em *IKZF1* foram encontradas em aproximadamente 22% dos casos de LLA-CPB, sendo a maioria deleção do tipo haploinsuficiente e cerca de um terço dominante-negativa;
- Em relação às características dos pacientes, o gene *RAG1* revelou-se com mediana de expressão maior no grupo de baixo risco quando comparado com os demais grupos decorrentes da classificação de risco por alterações citogenética-moleculares;
- Este estudo avaliou que os diferentes status de *IKZF1* estão associados a diferentes níveis de expressão de genes-alvo: o gene *EBF1* apresentou uma tendência de expressão reduzida diante da Δ dn quando comparado com os outros grupos, enquanto que o gene *GRB10* tem níveis aumentados em pacientes com LLA-CPB e *IKZF1* deletado, especificamente do tipo haploinsuficiente;
- Pacientes com baixa expressão de *EBF1* tem melhor SG, enquanto que indivíduos considerados com baixa expressão de *GRB10* apresentam pior SG e SLE, em relação aqueles com alta expressão.

8 REFERÊNCIAS BIBLIOGRÁFICAS

Arber DA, Orazi A, Hasserjian R, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*. 2016; 127(20):2391-2405.

Arenzana TL, Schjerven H, Smale ST. Regulation of gene expression dynamics during development transitions by the Ikaros transcription factor. *Genes & Development*. 2015; 29:1801-1816.

Barbosa TC. Caracterização da via de RB em leucemias linfoblásticas pediátricas com a fusão *TCF3-PBX1*. Tese (Doutorado em Oncologia) – Instituto Nacional de Câncer José Alencar Gomes da Silva. Rio de Janeiro. 2017.

Barbosa TC, Terra-Granado E, Quezado Magalhães IM, et al. Frequency of copy number abnormalities in common genes associated with B-cell precursor acute lymphoblastic leukemia cytogenetic subtypes in Brazilian children. *Cancer Genet*. 2015; 208(10):492-501.

Béné MC. Immunophenotyping of acute leukaemias. *Immunol Lett*. 2005; 98(1):9-21.

Boer JM, van der Veer A, Rizopoulos D, et al. Prognostic value of rare IKZF1 deletion in childhood B-cell precursor acute lymphoblastic leukemia: an international collaborative study. *Leukemia*. 2016; 30(1):32-38.

Bose S, Deininger M, Gora-Tybor J et al. The presence of typical and atypical BCR-ABL fusion genes in leukocytes of normal individuals: biologic significance and implications for the assessment of minimal residual disease. *Blood*. 1998; 92(9):3362-3367.

Brown KE, Guest SS, Smale ST, et al. Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. *Cel*. 1997; 91:845–854.

Bustin SA, Benes V, Garson JA, et al. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem*. 2009; 55(4):611-622.

Catovsky D, Matutes E, Buccheri V, et al. A classification of acute leukaemia for the 1990s. *Ann Hematol.* 1991; 62(1):16-21.

Caye A, Beldjord K, Mass-Malo K, et al. Breakpoint-specific multiplex polymerase chain reaction allows the detection of IKZF1 intragenic deletions and minimal residual disease monitoring in B-cell precursor acute lymphoblastic leukemia. *Haematologica.* 2013; 98(4):597-601.

Chen C, Bartenhagen C, Gombert M, et al. Next-generation-sequencing of recurrent childhood high hyperdiploid acute lymphoblastic leukemia reveals mutations typically associated with high risk patients. *Leuk Res.* 2015; 39(9):990-1001.

Chen M, Harvey RC, Mullighan CG, et al. Outcome modeling with CRLF2, IKZF1, JAK, and minimal residual disease in pediatric acute lymphoblastic leukemia: a Children's Oncology Group Study. *Blood.* 2012; 119:3512-3522.

Cobb BS, Morales-Alcelay S, Kleiger G, et al. Targeting of Ikaros to pericentromeric heterochromatin by direct DNA binding. *Genes Dev.* 2000; 14: 2146–2160.

Crist W, Carroll A, Shuster J, et al. Philadelphia chromosome positive childhood acute lymphoblastic leukemia: clinical and cytogenetic characteristics and treatment outcome. A Pediatric Oncology Group study. *Blood.* 1990; 76(3):489–494.

Dastugue N, Suci S, Plat G, et al. Hyperdiploidy with 58-66 chromosomes in childhood B-acute lymphoblastic leukemia is highly curable: 58951 CLG-EORTC results. *Blood.* 2013; 121:2415–2423.

de Boer J, Yeung J, Ellu J, et al. The E2A-HLF oncogenic fusion protein acts through Lmo2 and Bcl-2 to immortalize hematopoietic progenitors. *Leukemia.* 2011; 25:321–330.

de Camargo B, de Oliveira Santos M, Rebelo MS, et al. Cancer incidence among children and adolescents in Brazil: first report of 14 population-based cancer registries. *Int J Cancer.* 2010; 126(3):715-720.

den Boer ML, van Slegtenhorst M, De Menezes RX, et al. A subtype of childhood acute lymphoblastic leukaemia with poor treatment outcome: a genome-wide

classification study. *Lancet Oncology*. 2009; 10:125–134.

Dörge P, Meissner B, Zimmermann M, et al. IKZF1 deletion is an independent predictor of outcome in pediatric acute lymphoblastic leukemia treated according to the ALL-BFM 2000 protocol. *Haematologica*. 2013; 98(3):428-32.

Dumortier A, Jeannet R, Kirstetter P, et al. Notch activation is an early and critical event during T-Cell leukemogenesis in Ikaros deficient mice. *Mol Cell Biol*. 2006; 26:209-220.

Dupuis A, Gaub MP, Legrain M, et al. Biclonal and biallelic deletions occur in 20% of B-ALL cases with IKZF1 mutations. *Leukemia*. 2013; 27(2):503-507.

Dyer MJ, Akasaka T, Capasso M, et al. Immunoglobulin heavy chain locus chromosomal translocations in B-cell precursor acute lymphoblastic leukemia: rare clinical curios or potent genetic drivers? *Blood*. 2010; 115(8):1490-1499.

Ferreirós-Vidal I, Carroll T, Taylor B, et al. Genome-wide identification of Ikaros targets elucidates its contribution to mouse B-cell lineage specification and pre-B-cell differentiation. *Blood*. 2013; 121:1769-1782.

Fuka G, Kauer M, Kofler R, et al. The leukemia-specific fusion gene ETV6/RUNX1 perturbs distinct key biological functions primarily by gene repression. *PLoS ONE*. 2011; 6:e26348.

Georgopoulos K, Bigby M, Wang JH, et al. The Ikaros gene is required for the development of all lymphoid lineages. *Cell*. 1994; 79:143–156.

Ghazavi F, Lammens T, Van Roy N, et al. Molecular basis and clinical significance of genetic aberrations in B-cell precursor acute lymphoblastic leukemia. *Exp Hematol*. 2015; 43(8):640-653.

Gomez-del Arco P, Koipally J, Georgopoulos K. Ikaros SUMOylation: switching out of repression. *Mol Cell Biol*. 2005; 25:2688–2697.

Greaves M. Infection, immune responses and the aetiology of childhood leukaemia. *Nat Rev Cancer*. 2006; 6(3):193-203.

Hanson PI, Whiteheart SW. AAA+ proteins: have engine, will work. *Nat Rev Mol Cell Biol.* 2005; 6:519-529.

Harewood L, Robinson H, Harris R, et al. Amplification of AML1 on a duplicated chromosome 21 in acute lymphoblastic leukemia: a study of 20 cases. *Leukemia.* 2003; 17:547–553.

Harrison CJ, Moorman AV, Broadfield ZJ, et al. Three distinct subgroups of hypodiploidy in acute lymphoblastic leukaemia. *British Journal of Haematology.* 2004; 125:552–559.

Harrison CJ, Moorman AV, Schwab C, et al. An international study of intrachromosomal amplification of chromosome 21 (iAMP21): cytogenetic characterization and outcome. *Leukemia.* 2014; 28:1015–1021.

Harrison CJ. The detection and significance of chromosomal abnormalities in childhood acute lymphoblastic leukaemia. *Blood Rev.* 2001; 15(1):49-59.

Heerema NA, Sather HN, Sensel MG, et al. Prognostic impact of trisomies of chromosomes 10, 17, and 5 among children with acute lymphoblastic leukemia and high hyperdiploidy (> 50 chromosomes). *Journal of Clinical Oncology.* 2000; 18:1876–1887.

Iacobucci I, Storlazzi CT, Cilloni D, et al. Identification and molecular characterization of recurrent genomic deletions on 7p12 in the IKZF1 gene in a large cohort of BCR-ABL1–positive acute lymphoblastic leukemia patients: on behalf of Gruppo Italiano Malattie Ematologiche dell'Adulto Acute Leukemia Working Party (GIMEMA AL WP). *Blood.* 2009; 114:2159-2167.

Iacobucci I, Iraci N, Messina M, et al. IKAROS deletions dictate a unique gene expression signature in patients with adult B-cell acute lymphoblastic leukemia. *PLoS One.* 2012; 7(7):e40934.

Illert AL, Albers C, Kreutmair S, et al. Grb10 is involved in BCR-ABL-positive leukemia in mice. *Leukemia.* 2015; 29(4):858-868.

Inaba H, Greaves M, Mullighan CG. Acute lymphoblastic leukaemia. *Lancet.* 2013;

381(9881):1943–1955.

INCA – Instituto Nacional de Câncer. INCA - CÂNCER - Tipo - Infantil. Disponível em: <<http://www2.inca.gov.br/wps/wcm/connect/tiposdecancer/site/home/infantil>>. Acesso em: 24 de janeiro de 2018.

Joshi I, Yoshida T, Jena N, et al. Loss of Ikaros DNA binding function confers integrin-dependent survival on pre-B cells and progression to acute lymphoblastic leukemia. *Nat Immunol.* 2014; 15:294–304.

Jung D, Giallourakis C, Mostoslavsky R, et al. Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. *Annu Rev Immunol.* 2006; 24:541–570.

Kastner P, Dupuis A, Gaub MP, et al. Function of Ikaros as a tumor suppressor in B cell acute lymphoblastic leukemia. *Am J Blood Res.* 2013; 3(1):1-13.

Kaufmann C, Yoshida T, Perotti EA, et al. A complex network of regulatory elements in Ikaros and their activity during hemo-lymphopoiesis. *The EMBO Journal.* 2003; 22:2211–2223.

Kazi JU, Rönstrand L. FLT3 signals via the adapter protein Grb10 and overexpression of Grb10 leads to aberrant cell proliferation in acute myeloid leukemia. *Mol Oncol.* 2013; 7(3):402-418.

Kelley CM, Ikeda T, Koipally J, et al. Helios, a novel dimerization partner of Ikaros expressed in the earliest hematopoietic progenitors. *Current Biology.* 1998; 8:508–515.

Kinlen L. Infections and immune factors in cancer: the role of epidemiology. *Oncogene.* 2004; 23(38):6341-6348.

Kirstetter P, Thomas M, Dierich A, et al. Ikaros is critical for B cell differentiation and function. *Eur J Immunol.* 2002; 32:720–730.

Klug CA, Morrison SJ, Masek M, et al. Hematopoietic stem cells and lymphoid progenitors express different Ikaros isoforms, and Ikaros is localized to heterochromatin in immature lymphocytes. *Proc Natl Acad Sci USA.* 1998. 95:657-

662.

Kuiper RP, Waanders E, van der Velden VH, et al. IKZF1 deletions predict relapse in uniformly treated pediatric precursor B-ALL. *Leukemia*. 2010; 24(7):1258-1264.

Linka Y, Ginzl S, Krüger M, et al. The impact of TEL-AML1 (ETV6-RUNX1) expression in precursor B cells and implications for leukaemia using three different genome-wide screening methods. *Blood Cancer*. 2013; 3:e151.

Lopes BA, Meyer C, Barbosa TC, et al. COBL is a novel hotspot for IKZF1 deletions in childhood acute lymphoblastic leukemia. *Oncotarget*. 2016; 7(3):53064-53073.

Lugo TG, Pendergast AM, Muller AJ et al. Tyrosine Kinase-Activity and Transformation Potency of Bcr-Abl Oncogene Products. *Science*. 1990; 247:1079–1082.

Ma J, Li J, Yao X, et al. FIGNL1 is overexpressed in small cell lung cancer patients and enhances NCI-H446 cell resistance to cisplatin and etoposide. *Oncol Rep*. 2017; 37(4):1935-1942.

Marcais A, Jeannet R, Hernandez L, et al. Genetic inactivation of Ikaros is a rare event in human T-ALL. *Leuk Res* 2010; 34:426-429.

Martinelli G, Iacobucci I, Storlazzi CT, et al. IKZF1 (Ikaros) deletions in BCR-ABL1-positive acute lymphoblastic leukemia are associated with short disease-free survival and high rate of cumulative incidence of relapse: a GIMEMA AL WP report. *J Clin Oncol*. 2009; 27(31):5202-5207.

Meyer C, Burmeister T, Gröger D, et al. The MLL recombinome of acute leukemias in 2017. *Leukemia*. 2018; 32(2):273-284.

Meyer C, Zur Stadt U, Escherich G, et al. Refinement of IKZF1 recombination hotspots in pediatric BCP-ALL patients. *Am J Blood Res*. 2013; 3(2):165-173.

Molnár A, Wu P, Largespada DA, et al. The Ikaros gene encodes a family of lymphocyte-restricted zinc finger DNA binding proteins, highly conserved in human and mouse. *J Immunol*. 1996; 156:585-592.

Moorman AV, Clark R, Farrell, et al. Probes for hidden hyperdiploidy in acute lymphoblastic leukaemia. *Genes Chromosomes Cancer*. 1996; 16(1):40-45.

Moorman AV, Richards SM, Martineau M, et al. Outcome heterogeneity in childhood high-hyperdiploid acute lymphoblastic leukemia. *Blood*. 2003; 102:2756–2762.

Moorman AV, Richards SM, Robinson HM, et al. Prognosis of children with acute lymphoblastic leukemia (ALL) and intrachromosomal amplification of chromosome 21 (iAMP21). *Blood*. 2007; 109:2327–2330.

Moorman AV. New and emerging prognostic and predictive genetic biomarkers in B-cell precursor acute lymphoblastic leukemia. *Haematologica*. 2016; 101(4):407-416.

Morgan B, Sun L, Avitahl N, et al. Aiolos, a lymphoid restricted transcription factor that interacts with Ikaros to regulate lymphocyte differentiation. *The EMBO Journal*. 1997; 16:2004–2013.

Mullighan CG, Goorha S, Radtke I, et al. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature*. 2007; 446:758–764.

Mullighan CG, Su X, Zhang J, et al. Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia. *N Engl J Med*. 2009; 360(5):470-480.

Mullighan CG. The molecular genetic makeup of acute lymphoblastic leukemia. *Hematology Am Soc Hematol Educ Program*. 2012; 2012:389-396.

NIH – National Cancer Institute. Childhood Cancers. Disponível em: <<https://www.cancer.gov/types/childhood-cancers>>. Acesso em: 24 de janeiro de 2018.

Nolan T, Hands RE, Bustin SA. Quantification of mRNA using real-time RT-PCR. *Nat Protoc*. 2006; 1(3):1559-1582.

Olsson L, Johansson B. Ikaros and leukaemia. *Br J Haematol*. 2015; 169(4):479-491.

Papathanasiou P, Perkins AC, Cobb BS, et al. Widespread failure of hematolymphoid differentiation caused by a recessive niche-filling allele of the Ikaros transcription factor. *Immunity*. 2003; 19:131-144.

Paulsson K, Forestier E, Lilljebjörn H, Heldrup J, Behrendtz M, Young BD, Johansson B: Genetic landscape of high hyperdiploid childhood acute lymphoblastic leukemia. *Proc Natl Acad Sci USA*. 2010; 107:21719–21724.

Pieters R, Schrappe M, De Lorenzo P, et al. A treatment protocol for infants younger than 1 year with acute lymphoblastic leukaemia (Interfant-99): an observational study and a multicentre randomised trial. *Lancet*. 2007; 370(9583):240-250.

Pombo de Oliveira MS, Koifman S, Vasconcelos GM, et al. Development and perspective of current Brazilian studies on the epidemiology of childhood leukemia. *Blood Cells Mol Dis*. 2009; 42(2):121-125.

Pongubala JM, Northrup DL, Lancki DW, et al. Transcription factor EBF restricts alternative lineage options and promotes B cell fate commitment independently of Pax5. *Nat Immunol*. 2008; 9:203–215.

Popescu M, Gurel Z, Ronni T, et al. Ikaros stability and pericentromeric localization are regulated by protein phosphatase 1. *The Journal of Biological Chemistry*. 2009; 284:13869–13880.

Pui C. Acute lymphoblastic leukemia. In C. Pui (Ed.), *Childhood Leukemias*. Cambridge: Cambridge University Press. 2012. (pp. 332-366).

Pui CH, Raimondi SC, Williams DL. Isochromosome 17q in Childhood Acute Lymphoblastic-Leukemia - an Adverse Cytogenetic Feature in Association with Hyperdiploidy. *Leukemia*. 1988; 2:222–225.

Raetz EA, Carroll WL. Refining prognosis in BCR-ABL1–positive ALL. *Blood*. 2014; 123:1626-1627.

Rebollo A, Schmitt C. Ikaros, Aiolos and Helios: Transcription regulators and lymphoid malignancies. *Immunology and Cell Biology*. 2003; 81:171-175.

Russell LJ, Capasso M, Vater I, et al. Deregulated expression of cytokine receptor gene, CRLF2, is involved in lymphoid transformation in B-cell precursor acute lymphoblastic leukemia. *Blood*. 2009; 114:2688–2698.

Scheijen B, Boer JM, Marke R, et al. Tumor suppressors BTG1 and IKZF1 cooperate

during mouse leukemia development and increase relapse risk in B-cell precursor acute lymphoblastic leukemia patients. *Haematologica*. 2017; 102(3):541-551.

Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc*. 2008; 3(6):1101-1108.

Schouten JP, McElgunn CJ, Waaijer R, et al. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res*. 2002; 30(12):e57.

Schultz KR, Pullen DJ, Sather HN, et al. Risk- and response-based classification of childhood B-precursor acute lymphoblastic leukemia: a combined analysis of prognostic markers from the Pediatric Oncology Group (POG) and Children's Cancer Group (CCG). *Blood*. 2007; 109(3):926-935.

Schwab CJ, Jones LR, Morrison H, et al. Evaluation of multiplex ligation-dependent probe amplification as a method for the detection of copy number abnormalities in B-cell precursor acute lymphoblastic leukemia. *Genes Chromosomes Cancer*. 2010; 49(12):1104-1113.

Schwab CJ, Chilton J, Morrison H, et al. Genes commonly deleted in childhood B-cell precursor acute lymphoblastic leukemia: association with cytogenetics and clinical features. *Haematologica*. 2013; 98(7):1081-1088.

Schwickert TA, Tagoh H, Gültekin, et al. Stage-specific control of early B cell development by the transcription factor Ikaros. *Nat Immunol*. 2014; 15(3):283-293.

Shimono Y, Murakami H, Kawai K, et al. Mi-2 beta associates with BRG1 and RET finger protein at the distinct regions with transcriptional activating and repressing abilities. *J Biol Chem*. 2003; 278:51638–51645.

Smith M, Arthur D, Camitta B, et al. Uniform approach to risk classification and treatment assignment for children with acute lymphoblastic leukemia. *J Clin Oncol*. 1996; 14(1):18-24.

Snider J, Houry WA. AAA+ proteins: diversity in function, similarity in structure. *Biochem Soc Trans*. 2008; 36:72-77.

TARGET (NIH). Acute Lymphoblastic Leukemia. Disponível em: <<https://ocg.cancer.gov/programs/target/acute-lymphoblastic-leukemia>>. Acesso em: 24 de janeiro de 2018.

Tsuzuki S, Karnan S, Horibe K, et al. Genetic abnormalities involved in t(12;21) TEL-AML1 acute lymphoblastic leukemia: analysis by means of array-based comparative genomic hybridization. *Cancer Sci.* 2007; 98(5):698-706.

Virely C, Moulin S, Cobaleda C, et al. Haploinsufficiency of the IKZF1 (IKAROS) tumor suppressor gene cooperates with BCR-ABL in a transgenic model of acute lymphoblastic leukemia. *Leukemia.* 2010; 24:1200–1204.

Vitanza NA, Zaky W, Blum R, et al. Ikaros deletions in BCR-ABL-negative childhood acute lymphoblastic leukemia are associated with a distinct gene expression signature but do not result in intrinsic chemoresistance. *Pediatr Blood Cancer.* 2014; 61(10):1779-1785.

Vora A, Goulden N, Mitchell C, et al. Augmented post-remission therapy for a minimal residual disease-defined high-risk subgroup of children and young people with clinical standard-risk and intermediate-risk acute lymphoblastic leukaemia (UKALL 2003): a randomised controlled trial. *Lancet Oncology.* 2014; 15(8):809–818.

Vora A, Goulden N, Wade R, et al. Treatment reduction for children and young adults with low-risk acute lymphoblastic leukaemia defined by minimal residual disease (UKALL 2003): a randomised controlled trial. *Lancet Oncology.* 2013; 14(3):199–209.

Wiemels JL, Ford AM, Van Wering ER, et al. Protracted and variable latency of acute lymphoblastic leukemia after TEL-AML1 gene fusion in utero. *Blood.* 1999; 94(3):1057-1062.

Winandy S, Wu P, Georgopoulos K. A dominant mutation in the Ikaros gene leads to rapid development of leukemia and lymphoma. *Cell* 1995; 83:289-299.

Woo JS, Alberti MO, Tirado CA. Childhood B-acute lymphoblastic leukemia: a genetic update. *Exp Hematol Oncol.* 2014; 3:16.

Xie J, Wang Q, Wang Q, et al. High frequency of BTG1 deletions in patients with BCR-ABL1-positive acute leukemia. *Cancer Genet.* 2014; 207(5):226-230.

Yoshida T, Hazan I, Zhang J, et al. The role of the chromatin remodeler Mi-2beta in hematopoietic stem cell self-renewal and multilineage differentiation. *Genes Dev.* 2008; 22:1174–1189.

Yoshida T, Georgopoulos K. Ikaros fingers on lymphocyte differentiation. *Int J Hematol.* 2014; 100(3):220-229.

Yoshida T, Landhuis E, Dose M, et al. Transcriptional regulation of the *Ikzf1* locus. *Blood.* 2013; 122(18):3149-3159.

Yuan J, Chen J. FIGNL1-containing protein complex is required for efficient homologous recombination repair. *Proc Natl Acad Sci USA.* 2013; 110(26):10640-10645.

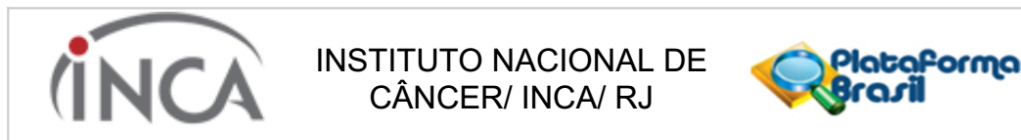
Zhang J, Jackson AF, Naito T, et al. Harnessing of the nucleosome-remodeling-deacetylase complex controls lymphocyte development and prevents leukemogenesis. *Nat Immunol.* 2011; 13:86–94.

Zhang J, Ding L, Holmfeldt L, et al. The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. *Nature* 2012; 481:157-163.

Zhang Y, LeRoy G, Seelig HP, et al. The dermatomyositis-specific autoantigen Mi2 is a component of a complex containing histone deacetylase and nucleosome remodeling activities. *Cell.* 1998; 95:279–289.

9 ANEXOS

Anexo 1 - Parecer do CEP



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: Análise Genômica de Deleções no gene IKZF1 em Leucemias Linfoblásticas Agudas Infantis

Pesquisador: MARIANA EMERENCIANO CAVALCANTI DE SÁ

Área Temática:

Versão: 2

CAAE: 33709814.7.0000.5274

Instituição Proponente: Coordenação de Pesquisa

Patrocinador Principal: FUN CARLOS CHAGAS F. DE AMPARO A PESQUISA DO ESTADO DO RIO DE JANEIRO - FAPERJ
MINISTERIO DA CIENCIA, TECNOLOGIA E INOVACAO

DADOS DO PARECER

Número do Parecer: 944.566

Data da Relatoria: 28/02/2015

Apresentação do Projeto:

Conforme descrito no Parecer Consubstanciado CEP/INCA Nº 748.804 de 13 de Agosto de 2014.

Objetivo da Pesquisa:

Conforme descrito no Parecer Consubstanciado CEP/INCA Nº 748.804 de 13 de Agosto de 2014.

Avaliação dos Riscos e Benefícios:

Conforme descrito no Parecer Consubstanciado CEP/INCA Nº 748.804 de 13 de Agosto de 2014.

Comentários e Considerações sobre a Pesquisa:

Conforme descrito no Parecer Consubstanciado CEP/INCA Nº 748.804 de 13 de Agosto de 2014.

Considerações sobre os Termos de apresentação obrigatória:

Conforme descrito no Parecer Consubstanciado CEP/INCA Nº 748.804 de 13 de Agosto de 2014.

Endereço: RUA DO RESENDE, 128 - SALA 203
Bairro: CENTRO **CEP:** 20.231-092
UF: RJ **Município:** RIO DE JANEIRO
Telefone: (21)3207-4550 **Fax:** (21)3207-4556 **E-mail:** cep@inca.gov.br



INSTITUTO NACIONAL DE
CÂNCER/ INCA/ RJ



Continuação do Parecer: 944.566

Em resposta às pendências apontadas pelo CEP/INCA foram submetidos em 28/01/2015 os seguintes arquivos:

- PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_341318.pdf
- Termo_de_Acordo_28Jan2015.pdf
- Resposta Parecer 748.804.pdf
- TCLE_IKZF1.pdf
- Projeto_IKZF1_CEP_v2_28.01.15.doc
- documentos 441.11.pdf
- Declaracao_Ausencia_Termo_Assentimento_28Jan2015.pdf

Recomendações:

Não se aplica.

Conclusões ou Pendências e Lista de Inadequações:

Trata-se da análise das respostas às pendências apontadas no Parecer Consubstanciado CEP/INCA Nº 748.804 de 13 de Agosto de 2014:

PENDÊNCIA 1.1

RESPOSTA: Na seção "População de Estudo" foram atendidos os esclarecimentos solicitados. "As amostras biológicas utilizadas neste estudo são de medula óssea (MO) obtidas antes de qualquer procedimento terapêutico e provenientes das instituições colaboradoras no Brasil, de acordo com os critérios de inclusão e exclusão abaixo descritos.

1. Seleção dos casos – coorte retrospectiva: Serão avaliadas as amostras de MO que foram recebidas das coletas já realizadas em pacientes com até 16 anos de idade, com diagnóstico de LLA-cpB, e atendidos nas instituições no Brasil. Somente serão aceitas para análises deste estudo, as amostras biológicas referentes aos pacientes cujos responsáveis tenham assinado o Termo de Consentimento Livre e Esclarecido (TCLE) específico do Programa de Hematologia-Oncologia Pediátricos (PHOP) localizado no Centro de Pesquisas (CPQ), INCA-RJ, e também aceitem assinar o TCLE específico para este estudo.

É previsto o recebimento de amostras que foram coletadas durante o período de Janeiro de 2009 a dezembro de 2014. Estas serão consideradas, para efeitos de análises deste estudo, como "amostras retrospectivas" ou "AR".

2. Seleção dos casos – coorte prospectiva: Serão avaliadas as amostras de MO que serão recebidas das coletas a serem realizadas em pacientes com até 16 anos de idade, para diagnóstico de LLA-

Endereço: RUA DO RESENDE, 128 - SALA 203
Bairro: CENTRO **CEP:** 20.231-092
UF: RJ **Município:** RIO DE JANEIRO
Telefone: (21)3207-4550 **Fax:** (21)3207-4556 **E-mail:** cep@inca.gov.br



INSTITUTO NACIONAL DE
CÂNCER/ INCA/ RJ



Continuação do Parecer: 944.566

cpB, e que serão atendidos nos centros de pesquisa colaboradores no Brasil. Somente serão aceitas para as análises deste estudo, as amostras biológicas referentes aos pacientes que assinarem o TCLE específico do PHOP, e também o TCLE específico para este estudo. Estas amostras biológicas serão consideradas, para efeitos de análises deste estudo, como "amostras prospectivas" ou "AP".

O recrutamento de pacientes elegíveis para este estudo, será realizado em cada instituição, pelo pesquisador participante e/ou profissional/is da área clínica encarregado/s do contato com o paciente. Este profissional/is deverão ser designado(s) pelo pesquisador participante responsável pela instituição, e ter os respectivos registros nos arquivos deste estudo (Currículo Vitae, licença profissional, treinamentos, Identificações no formulário de profissionais designados para o estudo, por exemplo), como forma de garantirmos que as amostras biológicas serão selecionadas de acordo com este projeto de pesquisa. Somente serão selecionados os pacientes que estejam de acordo com os critérios de inclusão e exclusão acima descritos. Para esta população de pacientes será apresentado o TCLE para o responsável do paciente elegível. A aplicação destes documentos também será realizada pelos profissionais qualificados e treinados para esta tarefa, nas instituições participantes.

i. Critérios de inclusão: a) Paciente com idade até 16 anos (incompletos); b) Paciente diagnosticado com LLA-cpB, de acordo com critérios da Organização Mundial da Saúde (OMS); c) Disponibilidade de amostras de MO; d) Paciente e responsável que aceitem participação neste estudo, através de assinaturas nos TCLEs.

ii. Critérios de exclusão: a) Paciente com leucemias associadas a síndromes congênitas e mielodisplasias; b) Paciente com material insuficiente ou células de baixa viabilidade; c) Paciente com 16 anos de idade ou mais."

ANÁLISE: PENDÊNCIA ATENDIDA

PENDÊNCIA 1.2

RESPOSTA: Estes esclarecimentos foram dados nas seções "Tipo do estudo" e "População de Estudo": "Estudo retrospectivo e prospectivo de análise molecular qualitativa."

ANÁLISE: PENDÊNCIA ATENDIDA

PENDÊNCIA 1.3

RESPOSTA: O esclarecimento quanto ao critério adotado para definir o número amostral encontra-se na seção "Desenho do Estudo". "Estima-se analisar aproximadamente 30 amostras de medula

Endereço: RUA DO RESENDE, 128 - SALA 203
Bairro: CENTRO **CEP:** 20.231-092
UF: RJ **Município:** RIO DE JANEIRO
Telefone: (21)3207-4550 **Fax:** (21)3207-4556 **E-mail:** cep@inca.gov.br



INSTITUTO NACIONAL DE
CÂNCER/ INCA/ RJ



Continuação do Parecer: 944.566

óssea (MO) de pacientes com diagnósticos de LLA-cpB. Este número foi estimado através de cálculo amostral a partir da prevalência da lesão molecular em questão (8)."

ANÁLISE: PENDÊNCIA ATENDIDA

PENDÊNCIA 1.4

RESPOSTA: O texto foi re-escrito, conforme no item "Análise dos dados". "Os resultados brutos do 100k SNP Array serão adquiridos pelo programa Affymetrix GCOS (versão 1.4) e os dados analisados serão gerados pelo programa GTYPE (versão 4.0). Os valores de cópia numérica e LOH serão inferidos pelo algoritmo CNAT (versão 3.0). O algoritmo CNAG v.2.0 (Copy Number Analyzer for GeneChip Mapping 100K arrays) também será usado para a análise dos dados de cópia numérica e LOH. Arquivos de referência serão gerados a partir da amostra de remissão (quando disponível) para a análise pareada com a amostra tumoral. Caso contrário, os dados serão comparados com os dados de um agrupamento de amostras em remissão."

ANÁLISE: PENDÊNCIA ATENDIDA

PENDÊNCIA 1.5

RESPOSTA: O item referente a "BENEFICIOS" foi removido do projeto e inserido no TCLE.

ANÁLISE: PENDÊNCIA ATENDIDA

PENDÊNCIA 1.6

RESPOSTA: O projeto foi adequado segundo as observações explicitadas nesta pendência.

ANÁLISE: PENDÊNCIA ATENDIDA

PENDÊNCIA 2

RESPOSTA: Seguem abaixo as respostas para cada subitem.

2.1, 2.2, 2.3 - As declarações pertinentes ao projeto de acordo com os requisitos estabelecidos na Resolução Nº 441 de 12 de maio de 2011 estão apresentadas.

2.4 - Foi elaborado um modelo de Termo de Acordo para aprovação do CEP.

2.5 - A informação referente ao armazenamento de amostras foi revisada.

ANÁLISE: PENDÊNCIA ATENDIDA

PENDÊNCIA 3

RESPOSTA: O cronograma foi adequado conforme solicitado.

Endereço: RUA DO RESENDE, 128 - SALA 203

Bairro: CENTRO

CEP: 20.231-092

UF: RJ

Município: RIO DE JANEIRO

Telefone: (21)3207-4550

Fax: (21)3207-4556

E-mail: cep@inca.gov.br



INSTITUTO NACIONAL DE
CÂNCER/ INCA/ RJ



Continuação do Parecer: 944.566

ANÁLISE: PENDÊNCIA ATENDIDA

PENDÊNCIA 4

RESPOSTA: Diante dos esclarecimentos explicitados nesta pendência, foi elaborado o TCLE apropriado para o projeto de pesquisa.

ANÁLISE: PENDÊNCIA ATENDIDA

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

Considerações Finais a critério do CEP:

Diante do exposto, o Comitê de Ética em Pesquisa do Instituto Nacional de Câncer (CEP-INCA), de acordo com as atribuições definidas na Resolução CNS 466/12 e na Norma Operacional N° 001/2013 do CNS, manifesta-se pela aprovação do projeto de pesquisa proposto.

Ressalto o(a) pesquisador(a) responsável deverá apresentar relatórios semestrais a respeito do seu estudo.

RIO DE JANEIRO, 04 de Fevereiro de 2015

Assinado por:
Carlos Henrique Debenedito Silva
(Coordenador)

Endereço: RUA DO RESENDE, 128 - SALA 203
Bairro: CENTRO **CEP:** 20.231-092
UF: RJ **Município:** RIO DE JANEIRO
Telefone: (21)3207-4550 **Fax:** (21)3207-4556 **E-mail:** cep@inca.gov.br

Página 05 de 05