Brazilian Journal of health Review

Use of bioinformatic tools for Etv6-Runx1 fusion detection associated with acute lymphocytic leukemia

Utilização de ferramentas de bioinformática para detecção da fusão *Etv6-Runx1* associada à leucemia linfocítica aguda

Inaldo Antônio dos Anjos Filho
Bacharel em Biomedicina pelo Centro Universitário UNIFAVIP | Wyden
Instituição: Centro Universitário UNIFAVIP | Wyden
Endereço: Avenida Adjar da Silva Casé, 800 - Indianópolis, Caruaru - PE, Brasil.
Orcid iD: https://orcid.org/0000-0001-6528-5316
E-mail: inaldoanjos_filho@hotmail.com

Ivson Warley Tôrres dos Anjos
Bacharel em Biomedicina pelo Centro Universitário UNIFAVIP | Wyden
Instituição: Centro Universitário UNIFAVIP | Wyden
Endereço: Avenida Adjar da Silva Casé, 800 - Indianópolis, Caruaru - PE, Brasil.
Orcid iD: https://orcid.org/0000-0002-7076-3121
E-mail: ivsonwarley@gmail.com

Manuela Eduarda de França
Biomédica pelo Centro Universitário UNIFAVIP | Wyden
Instituição: Centro Universitário UNIFAVIP | Wyden
Endereço: Avenida Adjar da Silva Casé, 800 - Indianópolis, Caruaru - PE, Brasil.
Orcid iD: https://orcid.org/0000-0002-9654-2247
E-mail: manuela.eduarda@gmail.com

Jonas José da Silva
Biomédico pelo Centro Universitário UNIFAVIP | Wyden
Instituição: Centro Universitário UNIFAVIP | Wyden
Endereço: Avenida Adjar da Silva Casé, 800 - Indianópolis, Caruaru - PE, Brasil.
Orcid iD: https://orcid.org/0000-0001-9348-2835
E-mail: jonnas-jose@hotmail.com

Juliana Laguzza de Oliveira Bustos Villabón
Biomédica pelo Centro Universitário UNIFAVIP | Wyden
Instituição: Centro Universitário UNIFAVIP | Wyden
Endereço: Avenida Adjar da Silva Casé, 800 - Indianópolis, Caruaru - PE, Brasil.
Orcid iD: https://orcid.org/0000-0003-1912-4698
E-mail: julianalaguzza@hotmail.com

Lidiane Gomes da Silva
Doutora em Genética pela Universidade Federal de Pernambuco
Instituição: Docente do Centro Universitário UNIFAVIP | Wyden

Brazilian Journal of health Review

Endereço: Avenida Adjar da Silva Casé, 800 - Indianópolis, Caruaru - PE, Brasil.
Orcid iD: https://orcid.org/0000-0003-4837-9016
E-mail: lidiane.silva@unifavip.edu.br

Nathalia Joanne Bispo Cezar
Doutora em Genética Humana pela Universidade de São Paulo
Instituição: Docente do Centro Universitário UNIFAVIP | Wyden
Endereço: Avenida Adjar da Silva Casé, 800 - Indianópolis, Caruaru - PE, Brasil.
Orcid iD: https://orcid.org/0000-0002-7383-0828
E-mail: njbcezar@gmail.com

ABSTRACT
Genetic mutations are the main causes that predispose onco-hematological pathologies. Acute lymphoblastic leukemia (ALL) is a neoplasia with high incidence in childhood and, in these patients are found mutations involving the *ETV6* and *RUNX1* genes. In the present study, we evaluated *in silico primers* previously described in literature to delimitate the translocation *ETV6-RUNX1* t (12; 21) (p13; q22). Also, the *in silico primers* were proposed to verify efficiency and quality in the detection of this fusion through bioinformatics tools. To detect the mutation, a *primer* pair was developed through bioinformatics mechanisms. Along with this, another 14 *primers* described in the available literature were rigorously evaluated. The National Center for Biotechnology Information (NCBI) and The European Bioinformatics Institute (EBI) databases were used to locate the target sequences and identify the genes melting points. The *Primer3Plus* tool was used to *primers* design, the *Oligo analyzer* software was used to evaluate the aforesaid *primers* using the following parameters: size 18-22 nucleotides; distribution of guanine and cytosine bases 40-60%; annealing temperature 52-60 °C; maximum of single base replicates 4 bp; *Delta G* value in *hairpins* formation above -9 kcal / mole^{-1} and *dimers* -9 kcal/mole. Finally, the gadget *MEGA7* and *NCBI BLAST* were used to evaluate the alignment and identity analysis of the *primers* and possible *amplicons*. Based on these analyses, it was possible to observe that from the 16 *primers*, only 7 qualified with all recommended values within the parameters. From this total, two pairs of *primers* presented better results through *in silico* analysis, giving rise for two possible *amplicons* that aligned adequately and, were useful at detecting the site of gene fusion. From this study, it was possible to observe the importance of *in silico* analysis for researches on complex diseases. Considering that from the *primers* design it is possible to perform a proper PCR, these analyzes are fundamental to reduce the experiments margin of error plus, the waste of time and inputs. The selected *primers* in this study presented parameters that indicate good results *in vitro* tests and can be used to delimitate the gene fusion.

Keywords: *primers*; bioinformatics; *in silico*; genes; ALL; *ETV6*; *RUNX1*; gene fusion.

RESUMO
Alterações genéticas são algumas das principais causas que predispõem patologias onco-hematológicas. A leucemia linfoide aguda (LLA) é uma neoplasia sanguínea com elevada incidência na infância, e apresenta com mais frequência em seus pacientes a mutação envolvendo os genes *ETV6* e *RUNX1*. No presente estudo foi avaliado *in silico primers* previamente descritos na literatura para delimitar a translocação *ETV6-RUNX1* t (12;21) (p13;q22). Adicionalmente, foram propostos *in silico primers* com intuito de constatar a eficácia e qualidade na detecção desta fusão por meio de ferramentas de bioinformática.

Para detecção da mutação foi desenvolvido um par de *primer* através de ferramentas de bioinformática. Adicionalmente, outros 14 *primers,* descritos na literatura foram rigorosamente avaliados. Os bancos de dados NCBI e EBI foram utilizados para localização das sequências alvo e identificação dos pontos de fusão dos genes. A ferramenta *Primer3Plus* foi empregada para desenho dos *primers*, o *software Oligo analyzer* foi usado na avaliação dos *primers* utilizando os seguintes parâmetros: tamanho 18-22 nucleotídeos; distribuição de bases guanina e citosina 40-60%; temperatura de anelamento 52-60 °C; máximo de repetições de base única 4 pb; valor do *Delta G* na formação de *harpins* acima de -9 kcal/mole[-1] e de *dímeros* -9 kcal/mole. Por fim, utilizou-se as ferramentas *MEGA7* e *NCBI BLAST* para avaliação do alinhamento e análise da identidade dos *primers* e dos possíveis *amplicons*. Com base nestas análises, foi possível observar que dos 16 *primers*, apenas 7 estavam com todos os valores dentro dos parâmetros recomendáveis. Deste total, dois pares de *primers* apresentaram melhores resultados através da análise *in silico*, originando dois possíveis *amplicons* que alinharam adequadamente e, adicionalmente, foram úteis na detecção do local da fusão gênica. A partir desse estudo foi possível observar a importância da analise *in silico* para pesquisas de doenças complexas. Considerando-se que a partir do desenho de *primers* é possível realizar uma PCR de qualidade, estas análises são fundamentais para reduzir a margem de erro dos experimentos e o desperdício de tempo e de insumos. Os *primers* selecionados neste estudo apresentaram parâmetros que apontam bons resultados em testes *in vitro*, podendo ser utilizados na delimitação da fusão gênica.

Palavras-Chave: *primers*; bioinformática; *in silico*; genes; LLA; *ETV6*; *RUNX1*; fusão gênica.

1 BACKGROUND

Leukemia is a malignant pathology that affects white blood cells (leukocytes), causing the decrease and/or loss of the natural differentiation ability and the unregulated multiplication of progenitors. Causing a clonal proliferation of young cells (blasts), and without functionality in the bone marrow, having distinct laboratory and prognostic characteristics [1,2]. They are classified especially according to the most prevalent hematopoietic lineage. When these alterations are predominant in B or T lymphocytes it is called lymphocytic and when it changes the myeloid line, myelocytic. Regarding the maturation level of leukemic cells, they are classified as acute due to the presence of a larger number of blastic and, chronic has most of the cells maturated [3–5]. It is estimated for the 2018-2019 biennium in Brazil, an incidence of 10,800 new cases per 100,000 inhabitants, being 5,940 men and 4,860 women [6].

Acute lymphocytic leukemia (ALL) is a malignant alteration resulting from an unregulated preproduction of B and T lymphocytes called lymphoblast [7]. Being the most common childhood neoplasm, occurring more frequently in boys and, has its apex between 2 and 5 years old [8,9]. In the United States approximately 6000 cases of ALL are diagnosed, where half of these are found in infant patients [10].

Brazilian Journal of health Review

Genetic alterations are among the main factors that trigger ALL, associated with more than 75% of cases [7]. Among the changes, chromosomal translocations are directly related to ALL activating the mainly transcription factor genes leading to uncontrolled cell proliferation. The most common genetic alteration in children is the translocation *ETV6-RUNX1* t (12; 21) observed in approximately 25% of cases being rare in adults diagnosed with ALL (Figure 1) [11].

Figure 1. Genetic changes in ALL. Half of ALL cases are related to hyperploidy (more than 50 chromosomes per blast) and t (12; 21) forming the *ETV6-RUNX1* fusion. Source: *Journal of Clinical Oncology 29*, no. 5 (February 2011) 551-565 (PUI et al., 2011).
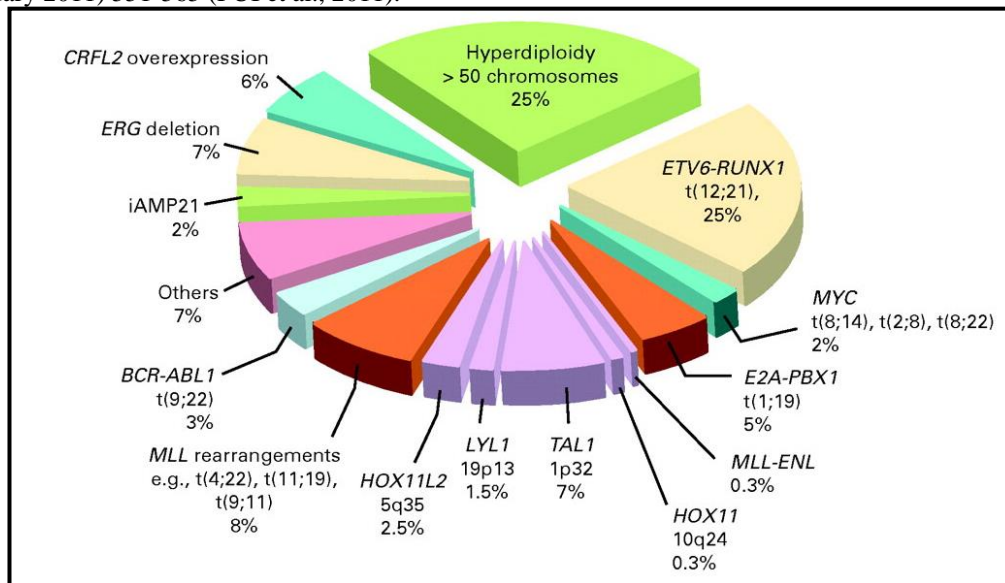

The *ETV6* gene (ETS variant gene 6) has its cytogenetic location at 12p13.2, the short arm of chromosome 12 at position 13.2 (Figure 2) and its molecular location of the base pairs 11,649,601 to 11,895,402 [12]. This gene has an N-terminal stranded helix-helix domain requested to protein-protein interaction and homodimerization plus, a C-terminal portion which is an ETS DNA binding domain formed by a GGAA/T central series sequence filled with nearby purines. There are also 8 *exons* encoding a 452 amino acid protein that is involved in 28 different chromosomal translocations, all of them linked to leukemia, one with the ALL-associated *RUNX1* gene [13,14]. The *ETV6* gene encodes an ETS family transcription factor. This family produces transcription factors that are crucial modulators of cellular homeostasis in various tissues and, when altered affect the normal functionality of these factors present in different cancers. Specifically, the *ETV6* gene has supreme

importance in hematopoiesis natural function within the bone marrow, that is why translocations involving *ETV6* are widely observed in different hematologic disorders [14].

The most common aberration involving *ETV6* is the t (12; 21) (p13; q22) translocation that fuses *ETV6* to the *RUNX1* gene resulting in an *in-frame ETV6-RUNX1* chimeric protein [15].

Figure 2. Cytogenetic localization of the *ETV6* gene. Source: Genome Decoration Page / NCBI.
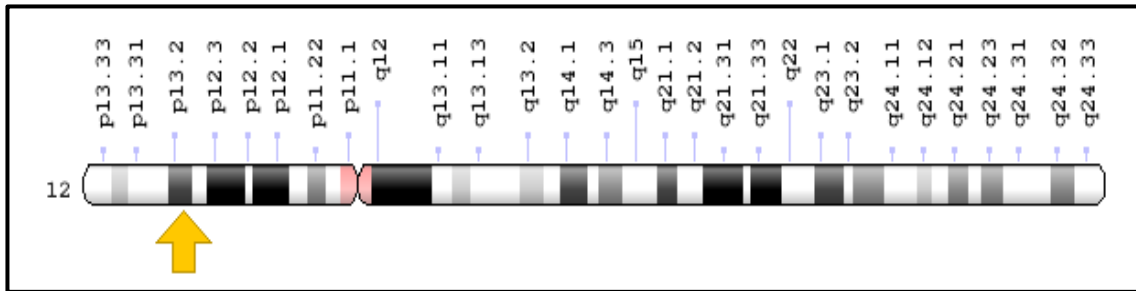
The *RUNX1* (dwarfism-related transcription factor 1) gene is cytogenetically located at 21q22.12 (Figure 3). That is, on the long arm of chromosome 21 at position 22.12 and its base pair molecular location is 34,787,801 to 35,049 .334 of the same chromosome [12]. The nucleus transcription factor (CBF) is heterodimeric (so-called because it is made of chains with different quantity of amino acids) that joins the main element of various promoters and enhancers. The *RUNX1* gene encodes a protein that represents the CBF alpha subunit and, is related to the normal development of hematopoiesis in the embryonic period. This gene expression occurs widely in the appendix, bone marrow, and 24 other tissues [16].

Figure 3. Cytogenetic localization of the *RUNX1* gene. Source: Genome Decoration Page / NCBI.
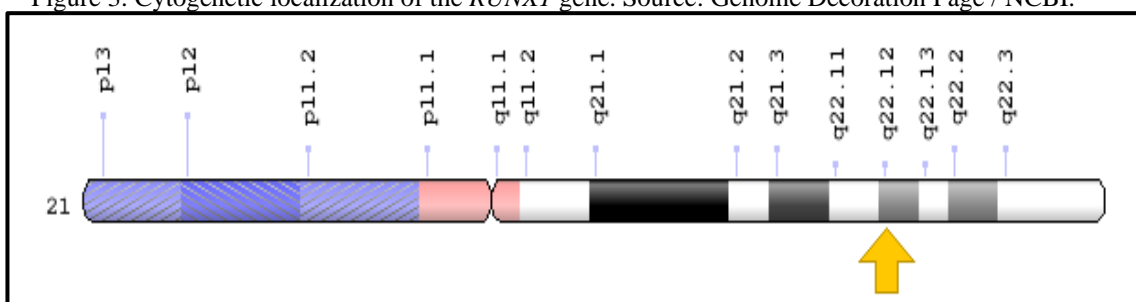
Continuously expressed, *ETV6-RUNX1* rearrangement influences leukemia subtypes boosting its proliferation and maintenance. *ETV6-RUNX1*-positive cells may appear in the womb and undergo genetic adjustment leading to leukemia [17]. Many studies in the last

decades revealed important advances on the exact function of *ETV6-RUNX1* rearrangement and the impact on ALL, both in the disease's onset and in relapse situations [18]. The fusion of *ETV6-RUNX1* unite its repressive central domain and the helix-loop-helix, to *ETV6* N-terminal tip with the transactivation and DNA binding domains. This modification transforms the *RUNX1* gene, which is a transcriptional activator, into a repressive suppressor [14].

To assist studies and research, bioinformatics tools are used to study and interpret biological data [19]. The most relevant use of bioinformatics involves genomic data due not only to the large amount of this type of information but, especially regarding technological advances in sequencing. These sequences are transformed into biological evidence through different studies identifying nucleotide sequence characteristics [20].

Considering *ETV6-RUNX1* the most common gene fusion in ALL children cases, is important to enable its identification using conventional molecular diagnosis. As follows, the present study aimed to evaluate from *in silico* studies *primers* previously described in the literature to delimit the *ETV6-RUNX1* translocation and, to produce it in the same purpose to verify the detecting effectiveness and accuracy for this *primer* fusion through bioinformatics gadgets.

2 METHODS

2.1 OBTAINING REFERENCE SEQUENCES AND GENE FUSION POINTS

Initially, the genome sequence reference was searched in the National Center for Biotechnology Information (NCBI) in the "Gene" tab [12], using the *ETV6* and *RUNX1* genes as keywords. By obtaining the complete sequences, the identification of the breakpoints of the translocation *ETV6-RUNX1* was initiated. To assist in the detection of the main *exons*, coding regions of the genes, associated with the targeted translocation, the European Bioinformatics Institute database was used. First, the genes were searched in the "Search" tab and then the "View in Ensemble" option was used to identify the genes. Subsequently, the sequence was downloaded in the "*exons*" option. All sequences obtained were archived in FASTA format. After target sequence selection, *primer* analysis was initiated.

2.2 *PRIMER* DESIGN, INCLUSION AND EXCLUSION PARAMETERS

Primer3Plus software was used to locate all possible *primer* sequences that could be originated from the target sequences. The best parameters selected admitted: *primer* size, GC content (guanine and cytosine) and *melting* temperature, preset in the program "General Settings" tab.

As inclusion and exclusion criteria for *primer* selection of the following parameters, were considered: I) length between 18 and 22 nucleotides, II) GC content amid 40% and 60%, III) *melting* temperature (Mt) between 52 ° C and 60 ° C, IV) maximum acceptable number of single base repeats 4 bp, V) ending at 3 ' with Guanine or Cytosine and VI) Mt difference among *primers* less than 5 ° C [21,22].

2.3 OBTAINING PRIMERS FROM BIBLIOGRAPHIC DATABASES

The articles were found in the following scientific databases: Scientific Electronic Library Online (SciELO), Medical Literature Analysis and Retrieval System Online (Medline) and National Center for Biotechnology Information (NCBI). It is noteworthy that the selection of articles was made without geographical and temporal restriction and, only information regarding the *primers* sequences were collected.

2.4 PRIMERS ALIGNMENT AND EVALUATION

The *MEGA7* software was used to perform the 16 *primers* alignment. Two generated in the *Primer3* software and fourteen obtained from the published articles, with the *exons* involved in the *ETV6* and *RUNX1* fusion. The *ClustalW* algorithm that performs the "Multiple and Global" alignment was selected. After the alignment, was possible to verify which *primers* were complete and without their respective *exons gaps*. From there on, an evaluation was made using the *Oligo Analyzer* software, that searched for possible arrangement of *hairpins*, *self-dimers* and *hetero dimers*. Likewise, the *primers* GC content, size and Mt were confirmed. Looking at *hairpins* standards, Delta G values could not be below -9 kcal/mole^{-1} [21]. Regarding the *dimer* analysis besides the Delta G evaluation, the possible occurrence of *primers* annealing was also observed. Conclusively, the alignment of possible amplified segments by the *primers* was made using the NCBI BLAST tool.

Figure 4. Step-by-step flowchart for *primer* design, acquisition, and evaluation.
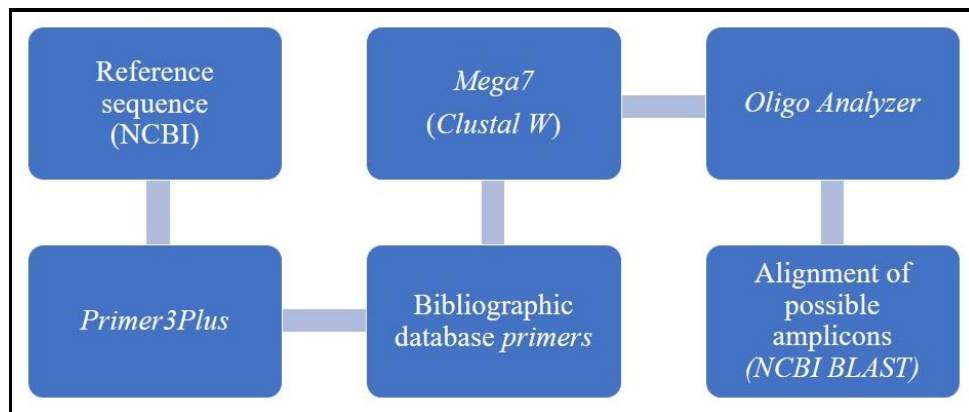
Source: From the author.

3 RESULTS

3.1 DEVELOPED *PRIMERS*

By obtaining and analyzing the genome of the genes involved in *ETV6-RUNX1* translocation, it was possible to detect breakpoints. Target regions were located in the *EVT6* gene including: 546 bp *exon* 5. And the *RUNX1* gene including: 254 bp *exon* 3, consistent with literature data [16]. *Primer3* software designed *primers* for the mentioned breakpoints with the parameters previously described in methods paragraph, where 21bp and 20bp size, 50% and 57.1% GC content and the melting temperature selected was 54.8 °C and 59.6 °C, not exceeding 5 °C the variation between the forward and reverse *primers*, in that order (Table 1).

Table 1. *Primers* designed by the *Primer3Plus* software.

Primers					
	Sequence 5'- 3'	**Direction**	**Size (pb)**	**% CG**	**Melt Temp (°C)**
ETV6-F	CCACCATTGAACTGTTGCAC	*Forward*	20	50	54,8
RUNX1-R	AGCACGGAGCAGAGGAAGTTG	*Reverse*	21	57,1	59,6

Primers obtained through Primer3 software using predetermined parameters.

3.2 *PRIMER* EVALUATION

All *primers* already described in the literature, were included in the evaluation. The parameters used are expressed in this study methods component. Fourteen *primer* sequences were selected (Table 2). In alignment, using the *MEGA7* software all *primers* yearned with

their respective target sequences. The *Oligo Analyzer* tool was used to perform the final *primer* evaluation. Afterwards, the *hairpins* formation was analyzed, using Delta G values greater than -9 kcal/mole^{-1}, the resulting structures and the annealing temperature (Figure 5). In this context, the highest temperature in *hairpin* formation was referred to *primer* 7 (47.3 ° C). Regarding the *self-dimer* analysis, there was a variation from 2 to 4 consecutive bases between the complementary bases, as well as the respective Delta G values above -9 kcal/mole. All *primers* showed 100% identity with their respective target sequences.

Table 2. Sequence and results of evaluated *primers*

ID	Sequence	Direction	Size (pb)	GC %	MT (°C)	BLAST	Self dimer (ΔG: kcal/mole)	Hairpin (ΔG: kcal/mole^{-1})	Source
1	TGCACCCTCTGATCCTGAAC	Forward	20	55	56,9	100	-7.05	0.04	Dongen *et al* 1999
2	AACGCCTCGCTCATCTTGC	Reverse	19	57,9	58,4	100	- 3,61	0,13	Dongen *et al* 1999
3	AAGCCCATCAACCTCTCTCATC	Forward	22	50	56,7	100	- 3,14	0,13	Dongen *et al* 1999
4	TGGAAGGCGGCGTGAAGC	Reverse	18	66,7	61,5	100	- 3,61	- 0,72	Dongen *et al* 1999
5	CGCACCAGGAGAACAACCAC	Forward	20	60	58,8	100	- 3,61	- 0,2	Dongen *et al* 1999
6	GAACCACATCATGGTCTCTG	Forward	20	50	53	100	- 6,37	- 2,2	Sudhakar *et al* 2011
7	TGGCTTACATGAACCACATCATGG	Forward	24	45,8	57,5	100	- 6,96	- 3,19	Sudhakar *et al* 2011
8	AACCTCTCTCATCGGGAAGA	Forward	20	50	54,9	100	-3,61	- 0,51	Tiensiwakul *et al* 2004
9	CAGAGTGCCATCTGGAACAT	Reverse	20	50	54,8	100	- 5,13	- 1,12	Tiensiwakul *et al* 2004
10	AACGCCTCGCTCATCTTGCCTG	Reverse	22	59,1	62,2	100	- 3,61	0,13	Tiensiwakul *et al* 2004
11	TGCACCCTCGATCCTGAAC	Forward	19	57,9	57,1	100	- 7,05	1,0	Pakakasama *et al* 2008
12	TGCATAGGGAAGGGAAG	Forward	17	52,9	50,9	100	- 7,05	0,00	Scurto *et al* 1998
13	TTCTTTTTTCACCATTCTTCC	Forward	21	33,3	49	100	- 1,47	0,00	Scurto *et al* 1998
14	GCAGAGGAAGTTGGGGCT	Reverse	18	61,1	57,7	100	- 3,14	0,7	Scurto *et al* 1998
15	CCACCATTGAACTGTTGCAC	Forward	20	50	54,8	100	- 7,05	0,33	From the author
16	AGCACGGAGCAGAGGAAGTTG	Reverse	21	57,1	59,6	100	- 3,61	- 1,49	From the author

Parameters obtained after analysis in *Mega7* and *Oligo Analyzer* software.

3.3 PRIMERS CHOICE AND POSSIBLE AMPLICONS ANALYSIS

Among the *primers* described in the studies by Dongen et al 1999, Sudhakar et al 2011, Tiensiwakul et al 2004, Pakakasama et al 2008 and Scurto et al 1998 [24–28], numbered respectively from 1 to 14. The selected *primers* were used to later stages of analysis. Due to the best parameters, *primer* pairs 1-2 and 15-16 were analyzed for *hetero dimer* formation, with Delta G above -9 kcal/mole and 3 complementary bases in sequence (Figure 7). Lastly, the alignment of possible amplicons formed by the selected *primer* pairs through NCBI BLAST was performed. The region that will possibly be amplified by pair 1-2 is 259 bp, while pair 15-16 is 697 bp. Thus, it was found that both regions delimited by the *primers* showed 100% identity when aligned to the Homo sapiens genome at two sites that correspond to the *ETV6* and *RUNX1* genes (Figures 8 and 9).

Figure 5. *Hairpin* formation analysis.

structure	Image	ΔG (kcal.mole⁻¹)	T_m (°C)	structure	Image	ΔG (kcal.mole⁻¹)	T_m (°C)
1		0.04	24.3	1		0.33	19.1
				2		1.21	8.1

| structure	Image	ΔG (kcal.mole⁻¹)	T_m (°C)	structure	Image	ΔG (kcal.mole⁻¹)	T_m (°C)
1		0.13	22.8	1		-1.49	46.5
2		0.31	20.3				

Source: from the author

Figure 6. *Self-dimer* formation analysis.
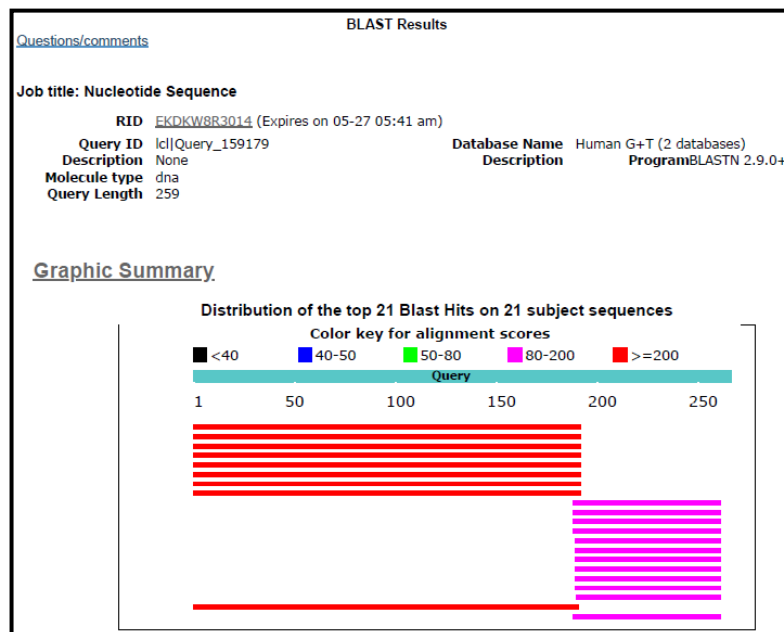
Primer 1

Delta G: -7.05 kcal/mole Base Pairs: 4

```
5'              TGCACCCTCTGATCCTGAAC
                ||||
3' CAAGTCCTAGTCTCCCACGT
```

Primer 2

Delta G: -3.61 kcal/mole Base Pairs: 2

```
5'              AACGCCTCGCTCATCTTGC
                ||
3' CGTTCTACTCGCTCCGCAA
```

Primer 15

Delta G: -7.05 kcal/mole Base Pairs: 4

```
5' CCACCATTGAACTGTTGCAC
                ||||
3'              CACGTTGTCAAGTTACCACC
```

Primer 16

Delta G: -7.05 kcal/mole Base Pairs: 4

```
5' GTTAGGACCCTGCAAACAGC
        :       ||||       :
3'      CGACAAACGTCCCAGGATTG
```

Source: from the author
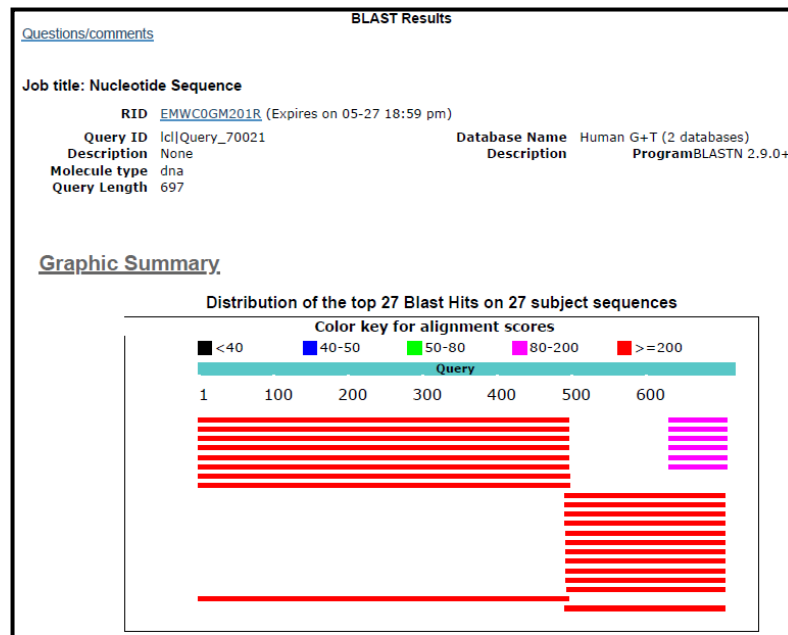
Figure 7. *Hetero-dimer* formation analysis.
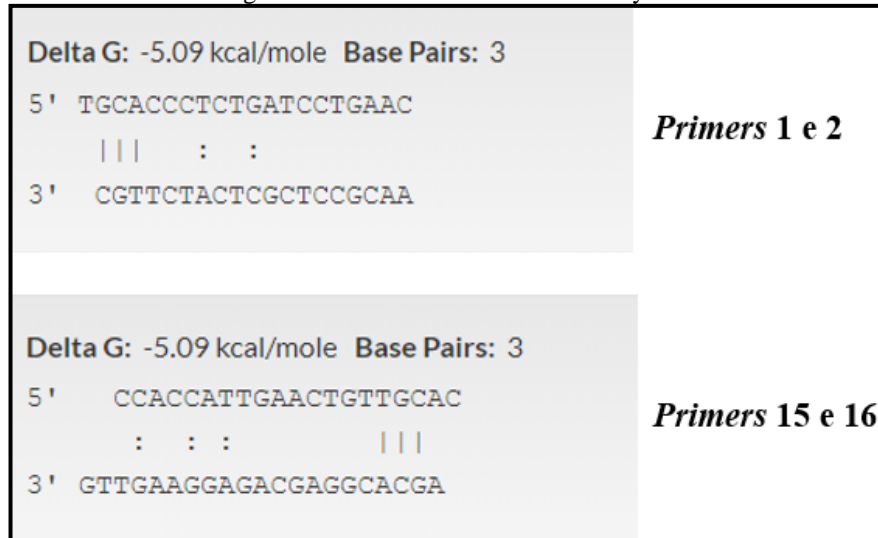
Source: from the author

Figure 8. Analysis of the region amplified by *primers* 1 and 2 after alignment with BLAST. Accessed on 25/05/2019.

Source: from the author

Brazilian Journal of health Review

Figure 9. Analysis of the region amplified by *primers* 15 and 16 after BLAST alignment. Accessed on 25/05/2019.

Source: from the author

4 DISCUSSION

Currently most studies have used modern techniques to detect translocations that trigger hematological neoplasms such as ALL, but this is not the reality of most diagnostic centers. In order to confirm the presence of the *ETV6-RUNX1* gene fusion by conventional PCR, is essential to use *primers* that delimit and amplify the genomic fusion point [29]. In the present work, primers analysis *in silico* were performed in order to optimize the detection mechanism of this change. For this, the selected sequences were evaluated by different software, regarding the detection quality of *ETV6-RUNX1* translocation.

According to the performed analyzes was possible to verify that most of the *primers* are within the established parameters. Being more specific, the *primers* size remained between 18 of 22 bp, standing alone *primer* 7 with 24 bp. This parameter importance comes from the immutable annealing step from primer to target sequence. It is important to emphasize that the short *primer* sequence may nonspecifically amplify; on the other hand, long *primers* increase the chances of *hairpins* occurrence (Table 2) [21,22].

The GC content of *primers* 4, 13 and 14 presented different values than recommended respectively: 66.7%, 33.3% and 61.1%. The other initiators were within the suggested range: 40 to 60%. Values above or below this range may influence melting temperature, as the case with *primer* 4 with Mt of 61.5ºC and *primer* 13 with Mt of 49ºC. This measurement is directly proportional to the variation of GC content concentration [30].

Brazilian Journal of health Review

Regarding the consecutive single base repeat, *primer* 13 presented six consecutive thymines. All other observed *primers* had less than 4 repetitions. Too many repetitions increase the primers potential of nonspecific annealing. Considering that melting temperature influences the specificity and efficiency of PCR [30], *primers* 4, 10, 12 and 13 respectively, were considered inadequate, according to the suggested temperature (52ºC - 60ºC), which might reflect in low amplification rate. Additionally, it is recommended that *primers* have cytosine or guanine at the 3 'end in order to increase binding efficiency, as adenine or thymine do not provide pairing stability [22,30]. In this sense, *primers* 8, 9 and 14 do not meet such requirements by having at their 3 'end the respective bases sequence: adenine, thymine and thymine.

Regarding the formation of secondary structures analysis known as *self-dimer* and *hairpin*, all *primers* obtained the following results: - 9 kcal/mole^{-1} and the base pairing was between 2 bp and 4 bp, presenting within the pre- established requirements [21]. The observed variation in *primers* was between - 7.05 kcal/mole and - 1.47 kcal/mole for *self-dimers* and between - 3.19 kcal/mole^{-1} and 1.0 kcal/mole^{-1} in *hairpins*, decreasing the chances of formation of these secondary structures occur [21], which corroborates for greater reaction efficiency.

According to our results, obtained from *in silico* analyzes, the *primer* pairs presenting best results were 1 and 2 described in Dogen et al 1999, and the 15 and 16 developed in this study. Both pairs have: i) difference in Mt less than 5ºC; ii) Delta G value for *hetero dimer* formation equal to - 5.09 kcal/mole. Likewise, *primer* pair-based amplicons have been successful in *in silico* analysis to delimit the major region described in studies where *ETV6* and *RUNX1* fusion occurs [23]. Thus and so, the detection of this gene rearrangement through conventional PCR has the possibility of presenting greater specificity and assertiveness, emphasizing the crucial contribution of bioinformatics tools.

5 CONCLUSIONS

Based on this study, was concluded that through free bioinformatics tools it is possible to provide satisfactory results due to *primers* efficiency in the DNA sequences amplification during PCR to detect ALL related mutations, helping to guide more effective therapies and treatment behaviors.

The *primers* selected in this study for amplification of the gene melting point involving *ETV6-RUNX1* translocation showed values in the *in silico* evaluation that lead to good

results to *in vitro* tests and can be used to delimitate the fusion and consequently the ALL diagnosis.

Beyond that, this research showed that is essential to thoroughly evaluate the parameters during the *primer* design process, as this approach directly reflects in the PCR quality. At last, the *in silico* analysis will grant cost savings in the laboratory testing phases.

Brazilian Journal of health Review

Braz. J. Hea. Rev., Curitiba, v. 3, n. 5, p. 12754-12770 set./out. 2020. ISSN 2595-6825

Brazilian Journal of health Review

REFERÊNCIAS

1.   Dantas GKS, Silva LTA, Passos XS, Carneiro CC. Diagnóstico diferencial da leucemia linfóide aguda em pacientes infanto-juvenis [Internet]. Revista da Universidade Vale do Rio Verde. 2015 [citado 17 de novembro de 2018]. Disponível em: https://dialnet.unirioja.es/servlet/articulo?codigo=5298238

2.   Santos MHS. Mortalidade e fatores associados ao óbito em crianças e adolescentes com leucemia linfoide aguda tratadas em um serviço de referência no nordeste do Brasil. 2017;55.

3.   Rizzati F, Buldini B, Basso G. Leukemia blast cell identification. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. março de 2015;5(2):74–85.

4.   Dourado KMC. Papel da Endoglina/CD105 em leucemias agudas: um potencial alvo para intervenção terapêutica. 2016;118.

5.   Moreira LA, Batista SC. Diagnóstico de leucemias linfóides agudas: uma revisão. Revista Saúde em Foco. 2018;9.

6.   INCA - CÂNCER - Tipo - Leucemia [Internet]. [citado 22 de outubro de 2018]. Disponível   em: http://www2.inca.gov.br/wps/wcm/connect/tiposdecancer/site/home/leucemia

7.   SILVA EW da. Mecanismos de citotoxicidade de chalconas isoladas e encapsuladas em nanopartículas lipídicas sobre uma linhagem celular de leucemia linfoblástica aguda e identificação de novos inibidores da proteína de resistência ABCG2 [Internet] [text]. Universidade Federal de Santa Catarina; 2013 [citado 29 de outubro de 2018]. Disponível em: https://repositorio.ufsc.br/xmlui/bitstream/handle/123456789/123135/325542.pdf?sequenc e=1&isAllowed=y]

8.   Ganazza MA. Estudo de doença residual minima em leucemia linfoide aguda da criança e do adolescente. Minimal residual disease study in acute lymphoblastic leukemia of child and adolescent [Internet]. 2014 [citado 15 de outubro de 2018]; Disponível em: http://repositorio.unicamp.br/jspui/handle/REPOSIP/309912

9.   Vieira AF, Neves B, Tonelli SR. Perfil epidemiológico da leucemia linfoide nas regiões do Brasil. UNILUS Ensino e Pesquisa. 31 de março de 2018;14(37):130–43.

10.   Hunger SP, Mullighan CG. Acute Lymphoblastic Leukemia in Children. Longo DL, organizador. New England Journal of Medicine. 15 de outubro de 2015;373(16):1541–52.

11.   Pui C-H, Carroll WL, Meshinchi S, Arceci RJ. Biology, Risk Stratification, and Therapy of Pediatric Acute Leukemias: An Update. JCO. 10 de fevereiro de 2011;29(5):551–65.

12.   Home - Gene - NCBI [Internet]. [citado 26 de novembro de 2018]. Disponível em: https://www.ncbi.nlm.nih.gov/gene/

13. Kowarz E, Dingermann T, Marschalek R, Kowarz E, Dingermann T, Marschalek R. Do Non-Genomically Encoded Fusion Transcripts Cause Recurrent Chromosomal Translocations? Cancers. 18 de outubro de 2012;4(4):1036–49.

14. Neveu B, Caron M, Lagacé K, Richer C, Sinnett D. Genome wide mapping of ETV6 binding sites in pre-B leukemic cells. Sci Rep [Internet]. 19 de outubro de 2018 [citado 18 de novembro de 2018];8. Disponível em: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6195514/

15. Schäfer D, Olsen M, Lähnemann D, Stanulla M, Slany R, Schmiegelow K, et al. Five percent of healthy newborns have an ETV6-RUNX1 fusion as revealed by DNA-based GIPFEL screening. Blood. 15 de fevereiro de 2018;131(7):821–6.

16. Silva-Cruz R, Bobadilla-Morales L, Corona-Rivera JR, Corona-Rivera A. Impacto del gen de fusión etv6/runx1 en leucemia linfoblástica aguda pediátrica. Revista Venezolana de Oncología. 2016;28(4):261–71.

17. Böiers C, Richardson SE, Laycock E, Zriwil A, Turati VA, Brown J, et al. A Human IPS Model Implicates Embryonic B-Myeloid Fate Restriction as Developmental Susceptibility to B Acute Lymphoblastic Leukemia-Associated ETV6-RUNX1. Dev Cell. 05 de 2018;44(3):362-377.e7.

18. Sun C, Chang L, Zhu X. Pathogenesis of ETV6/RUNX1 -positive childhood acute lymphoblastic leukemia and mechanisms underlying its relapse. Oncotarget. 18 de março de 2017;8(21):35445–59.

19. Pitassi C, Gonçalves AA, Moreno Júnior V de A. Fatores que influenciam a adoção de ferramentas de TIC nos experimentos de bioinformática de organizações biofarmacêuticas. Ciênc saúde coletiva. janeiro de 2014;19:257–68.

20. Pagani I, Liolios K, Jansson J, Chen I-MA, Smirnova T, Nosrat B, et al. The Genomes OnLine Database (GOLD) v.4: status of genomic and metagenomic projects and their associated metadata. Nucleic Acids Res. janeiro de 2012;40(Database issue):D571–9.

21. Queiroz JA da S, Alves LS, Dall'acqua DSV, Souza LFB. Desenho e Validação de Primers In Silico para Detecção do Vírus Sincicial Respiratório Humano. Design and Validation of Primers In Silico for Detection of Human Respiratory Syncytial Virus [Internet]. 2017 [citado 24 de maio de 2019]; Disponível em: https://www.arca.fiocruz.br/handle/icict/32950

22. Moreira DMB, Reis AS dos, Sales ML. Desenho de iniciadores para amplificação do vírus linfotrópico de células t humano tipo 1 (HTLV-1). Revista Brasileira de Ciências da Vida [Internet]. 25 de abril de 2018 [citado 28 de maio de 2019];6(Especial). Disponível em: http://jornal.faculdadecienciasdavida.com.br/index.php/RBCV/article/view/753

23. Jin Y, Wang X, Hu S, Tang J, Li B, Chai Y. Determination of ETV6-RUNX1 genomic breakpoint by next-generation sequencing. Cancer medicine. 29 de dezembro de 2015;5.

24. Dongen J van, Macintyre EA, Gabert JA, Delabesse E, Rossi V, Saglio G, et al. Standardized RT-PCR analysis of fusion gene transcripts from chromosome aberrations in

Brazilian Journal of health Review

acute leukemia for detection of minimal residual disease. Leukemia. dezembro de 1999;13(12):1901.

25.   Sudhakar N, Rajalekshmy KR, Rajkumar T, Nancy KN. RT-PCR and real-time PCR analysis of E2A-PBX1, TEL-AML1, mBCR-ABL and MLL-AF4 fusion gene transcripts in de novo B-lineage acute lymphoblastic leukaemia patients in south India. J Genet. agosto de 2011;90(2):349–53.

26.   Tiensiwakul P. Cloning and sequencing of ETV6/RUNX1 (TEL/AML1) variant in acute lymphoblastic leukemia. Cancer Genet Cytogenet. fevereiro de 2004;149(1):85–8.

27.   Pakakasama S, Kajanachumpol S, Kanjanapongkul S, Sirachainan N, Meekaewkunchorn A, Ningsanond V, et al. Simple multiplex RT-PCR for identifying common fusion transcripts in childhood acute leukemia. Int J Lab Hematol. agosto de 2008;30(4):286–91.

28.   Scurto P, Hsu Rocha M, Kane JR, Williams WK, Haney DM, Conn WP, et al. A multiplex RT-PCR assay for the detection of chimeric transcripts encoded by the risk-stratifying translocations of pediatric acute lymphoblastic leukemia. Leukemia. dezembro de 1998;12(12):1994–2005.

29.   Fueller E, Schaefer D, Fischer U, Krell PFI, Stanulla M, Borkhardt A, et al. Genomic Inverse PCR for Exploration of Ligated Breakpoints (GIPFEL), a New Method to Detect Translocations in Leukemia. PLoS One [Internet]. 19 de agosto de 2014 [citado 25 de maio de 2019];9(8). Disponível em: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4138100/

30.   Lorenz TC. Polymerase Chain Reaction: Basic Protocol Plus Troubleshooting and Optimization Strategies. J Vis Exp [Internet]. 22 de maio de 2012 [citado 28 de maio de 2019];(63). Disponível em: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4846334/