

APPLICATION OF TARGETED NEXT GENERATION SEQUENCING FOR THE MUTATIONAL PROFILING OF PATIENTS WITH ACUTE LYMPHOBLASTIC LEUKEMIA

PRIMENA CILJANOG SEKVENCIRANJA NOVE GENERACIJE U ANALIZI MUTACIONOG PROFILA PACIJENATA SA AKUTNOM LIMFOBLASTNOM LEUKEMIJOM

Dragana Janic^{1,4}, Jelena Peric², Teodora Karan-Djurasevic², Tatjana Kostic², Irena Marjanovic², Bojana Stanic², Nadja Pejanovic², Lidija Dokmanovic^{1,4}, Jelena Lazic^{1,4}, Nada Krstovski^{1,4}, Marijana Virijevic^{3,4}, Dragica Tomin^{3,4}, Ana Vidovic^{3,4}, Nada Suvajdzic Vukovic^{3,4}, Sonja Pavlovic², Natasa Tomic²

¹Department of Hematology and Oncology, University Children's Hospital, University of Belgrade, Belgrade, Serbia

²Laboratory for Molecular Biomedicine, Institute of Molecular Genetics and Genetic Engineering, University of Belgrade, Belgrade, Serbia

³Clinic of Hematology, Clinical Center of Serbia, Belgrade, Serbia

⁴School of Medicine University of Belgrade, Belgrade, Serbia

Summary

Background: Acute lymphoblastic leukemia (ALL) is the most common cancer in children, whereas it is less common in adults. Identification of cytogenetic aberrations and a small number of molecular abnormalities are still the most important risk and therapy stratification methods in clinical practice today. Next generation sequencing (NGS) technology provides a large amount of data contributing to elucidation of mutational landscape of childhood (cALL) and adult ALL (aALL).

Methods: We analyzed DNA samples from 34 cALL and aALL patients, using NGS targeted sequencing TruSeq Amplicon – Cancer Panel (TSACP) which targets mutational hotspots in 48 cancer related genes.

Results: We identified a total of 330 variants in the coding regions, out of which only 95 were potentially protein-changing. Observed in individual patients, detected mutations predominantly disrupted Ras/RTK pathway (*STK11*, *KIT*, *MET*, *NRAS*, *KRAS*, *PTEN*). Additionally, we identified

Kratak sadržaj

Uvod: Akutna limfoblastna leukemija (ALL) je najčešće maligno oboljenje kod dece, dok je kod odraslih njena učestalost mnogo niža. U današnjoj kliničkoj praksi kao najvažnije metode stratifikacije pacijenata u određene grupe rizika koriste se metode identifikacije citogenetičkih aberacija i malog broja molekularnih markera. Tehnologija sekvenciranja nove generacije (SNG) obezbeđuje veliku količinu podataka koji doprinose razjašnjavanju mutacionog profila dečje (dALL) i adultne ALL (aALL).

Metode: Uzorci DNK iz 34 dALL i aALL pacijenata analizirani su primenom SNG ciljanog sekvenciranja (»TruSeq Amplicon Cancer Panel – TSACP«) kojim se sekvenciraju »hotspot« mutacije u 48 gena povezanih sa kancerom.

Rezultati: Identifikovano je ukupno 330 varijanti u kodirajućim regionima, od kojih je samo 95 njih za posledicu imalo potencijalnu promenu u proteinu. Posmatrano kod pojedinačnih pacijenata, detektovane mutacije su pretežno remetile Ras/RTK signalni put (*STK11*, *KIT*, *MET*, *NRAS*,

Address for correspondence:

Natasa Tomic
University of Belgrade
Institute of Molecular Genetics and Genetic Engineering
444a Vojvode Stepe Street, 11010 Belgrade, Serbia
Phone: +381 11 39 76 445
Fax: +381 11 39 75 808
natasa.tomic@imgge.bg.ac.rs

List of abbreviations: cALL, childhood acute lymphoblastic leukemia; aALL, adult acute lymphoblastic leukemia; NGS, next generation sequencing; NFM mutations, nonsense, frameshift and missense mutations.

5 patients with the same mutation in *HNF1A* gene, disrupting both Wnt and Notch signaling pathway. In two patients we detected variants in *NOTCH1* gene. *HNF1A* and *NOTCH1* variants were mutually exclusive, while genes involved in Ras/RTK pathway exhibit a tendency of mutation accumulation.

Conclusions: Our results showed that ALL contains low number of mutations, without significant differences between cALL and aALL (median per patient 2 and 3, respectively). Detected mutations affect few key signaling pathways, primarily Ras/RTK cascade. This study contributes to knowledge of ALL mutational landscape, leading to better understanding of molecular basis of this disease.

Keywords: acute lymphoblastic leukemia, next generation sequencing, somatic mutations

Introduction

Acute lymphoblastic leukemia (ALL) is a heterogeneous malignant disorder resulting from the accumulation of aberrantly transformed B or T lymphoid progenitors at different developmental stages (1). ALL is the most common cancer in children, representing about 80% of acute leukemia, whereas it is less common in adults (20%). The heterogeneity of this disease originates from various clinical, morphological and immunological phenotypes, but also from the fact that ALL is a genetically complex entity (2).

Contemporary approach to ALL treatment implies precise stratification into different risk groups that is primarily based on specific clinical and genetic characteristics. Significant advances have been made in the treatment of ALL with the cure rates for childhood ALL approaching 90%. Still, survival in the adult ALL population is only about 40% and decreases with age (3).

It has been noted that the frequency of recurring cytogenetic abnormalities present in both childhood and adult ALL, differ between these groups of patients. This difference may be the basis for the reported discrepancy in the survival rates in these age groups. Moreover, the frequency of high risk T-ALL, is much higher in adults than among children. Childhood ALL is in 85% of cases B-cell type, characterized by the presence of high hyperploid (>50 chromosomes) and t (12; 21) (p13; q22) i.e. *ETV6/RUNX1* rearrangement, both associated with favorable prognosis. Conversely, in adult leukemia, aberrations with poor prognostic significance, like the presence of hypodiploidy (30–39 chromosomes), translocations t (4; 11) (q21; q23) and t (9; 22) (q34; q11), i.e. *MLL/AFF1* and *BCR/ABL1* rearrangements, are much more frequent (4, 5).

Above-mentioned chromosomal rearrangements are common in ALL and are critical events in leukemogenesis. These, so called primary genetic events, usually affect lymphoid differentiation and

KRAS, *PTEN*). Pored toga, identifikovano je 5 pacijenata sa istom mutacijom u *HNF1A* genu, koja je uzrokovala poremećaje u Wnt i Notch signalnom putu. Kod dva pacijenta otkrivene su varijante u *NOTCH1* genu. Nije detektovano istovremeno prisustvo varijanti u *HNF1A* i *NOTCH1* genu, dok su geni uključeni u Ras/RTK signalni put pokazali tendenciju ka akumuliranju mutacija.

Zaključak: Naši rezultati pokazuju da ALL sadrži mali broj mutacija, bez značajnih razlika između dALL i aALL (medijana po pacijentu 2 odnosno 3). Detektovane mutacije izazivaju poremećaje u nekoliko ključnih signalnih puteva, prvenstveno Ras/RTK kaskade. Ova studija doprinosi ukupnom znanju o mutacionom profilu ALL, što vodi ka boljem razumijevanju molekularne osnove ovog oboljenja.

Cljučne reči: akutna limfoblastna leukemija, sekvenciranje nove generacije, somatske mutacije

proliferation processes, but for the induction of full-blown leukemia, multiple mutations are required. New technologies like next generation sequencing (NGS) offers great potential for variants identification and genomic profiling of ALL. Utilization of NGS has enabled detection of additional submicroscopic alterations in the genes involved in tumor suppression, apoptosis, and cell-cycle regulation, contributing to more comprehensive insight into leukemogenesis. And not only that, these new markers have been used in diagnosis, risk-stratification and targeted therapy application, leading to improvement of current protocols and patient management (6, 7).

In this study, we applied targeted next generation sequencing on MiSeq System for analyzing somatic mutations in groups of adult (aALL) and childhood (cALL) ALL patients, in order to facilitate recognition and better understanding of the genetic profile of the disease.

Materials and Methods

Subjects

Bone marrow samples from the 17 adult and 17 childhood ALL patients at diagnosis were collected. Adult ALL patients came from the Clinic of Hematology, Clinical Center of Serbia, and childhood patients came from the Department of Hematology, University Children Hospital in Belgrade. The study was approved by the Ethics Committee of the Clinical Center of Serbia. Research was conducted in accordance with the ethical standards of the World Medical Association's Declaration of Helsinki. Informed consent was obtained from each patient or patient's parent or guardian.

Mononuclear cells were separated by Ficoll density gradient centrifugation and cryopreserved until mutational analyses. Some clinical characteristics of the patients are listed in *Tables I and II*.

Table I Clinical characteristic of cALL patients.

Patient No.	Sex	Age	immunophenotype	karyotype	RT-PCR analysis
1	F	34	B-ALL (common)	NA	TEL/AML1
2	F	30	B-ALL (common)	46,XX [20]	neg
3	M	46	B-ALL (common)	NA	neg
4	F	18	B-ALL (common)	hyperdiploidy (51–55, XX)	neg
5	F	76	B-ALL (common)	hyperdiploidy (49–52, XX)	neg
6	M	97	B-ALL (common)	NA	TEL/AML1
7	F	29	B-ALL (common)	NA	neg
8	M	23	pre-B-ALL	46,XY [20]	neg
9	F	30	pre-B-ALL	46,XX [20]	BCR/ABL
10	M	61	B-ALL (common)	hyperdiploidy (55–60, XY)	neg
11	F	30	B-ALL (common)	46,XX [20]	TEL/AML1
12	F	125	B-ALL (common)	46,XX [20]	neg
13	F	77	B-ALL (common)	hyperdiploidy (47, XX)	TEL/AML1
14	M	63	B-ALL (common)	NA	neg
15	M	213	pre-B-ALL	NA	neg
16	M	145	pre-B-ALL	NA	TEL/AML1
17	M	126	B-ALL (common)	NA	NA

M – male, F – female; NA – not available

Table II Clinical characteristic of aALL patients.

Patient	Sex	Age (years)	immunophenotype	Karyotype	RT-PCR analysis
1	M	29	pro B-ALL	46,XY [20]	ND
2	F	26	B-ALL (common)	46,XX [6]	MLL/AFF1
3	M	37	B-ALL (common)	46,XY, t(4,11)(q21;q23)[2]/62–82,XY,t(4,11)(q21;q23) [18]	ND
4	M	44	B-ALL (common)	46,XY [20]	neg
5	M	19	B-ALL (common)	46XY [20]	neg
6	F	64	B-ALL (common)	46,XX [20]	neg
7	M	40	T-ALL	46,XY [13]	BCR/ABL
8	F	24	B-ALL (common)	46,XX,del(cq)[5]/46,XX,t(9;22)(q34;q11)[2]/46,XX [7]	BCR/ABL
9	F	45	B-ALL (common)	NA	BCR/ABL
10	M	33	B-ALL (common)	46,XY, t(8,22)(q24;q11) [20]	ND
11	F	19	B-ALL (common)	46 XX/46XX, -B, -C, +M1, +M2	ND
12	M	41	B-ALL (common)	46XX [20]	ND
13	M	28	B-ALL (common)	46XY[20]	ND
14	F	61	B-ALL (common)	46,XX [20]	BCR/ABL
15	M	43	B-ALL (common)	NA	ND
16	M	28	T-ALL	46, XY [20]	ND
17	F	35	B-ALL (common)	46, XX [20]	ND

M – male, F – female; NA – not available; ND – not done

TruSeq Amplicon – Cancer Panel library preparation and sequencing

TruSeq Amplicon – Cancer Panel, TSACP (Illumina Inc., San Diego, CA, USA) targets mutational hotspots in 48 cancer-related genes. TSACP consists of 212 amplicons captured by pairs of oligonucleotides designed to hybridize flanking targeted regions of interest. Genomic DNA from mononuclear cells was extracted using Qiagen Blood Mini kit (Heidelberg, Germany). The library preparation was performed using 250 ng of genomic DNA, according to the manufacturer's protocol. The equal volumes of normalized libraries were pooled and prepared for subsequent cluster generation and sequencing on the MiSeq system (Illumina Inc., San Diego, CA, USA). Paired-end sequencing was performed using the MiSeq Reagent Kit v3 (600-cycle) and the sequencing quality was demonstrated by the percentage of bases having the Q30 score (1 error in 1000 bases) of 97.2%.

Bioinformatics Analysis

After library sequencing, FASTQ files were processed in multiple stages starting with quality control and trimming, alignment and pre-processing, followed by additional quality control, variant calling and filtration. The bioinformatics pipeline was designed by Seven Bridges Genomics (SBG), containing open source bioinformatics tools as well as in-house developed tools. The basic quality control was performed with FastQC (8) and after that, low quality bases were trimmed from read ends using FastqMcf (9). Then the production of BAM file(s) was done with BWA-MEM (10–12). This step involved the alignment of sequences with the reference genome GRCh37. The insertions/deletions (indel) realignment over the reads overlapping target regions was performed using RealignerTargetCreator and IndelRealigner (13, 14). SBG developed custom scripts that was used for additional quality control which consisted of counting reads of each amplicon and identifying amplicons presenting low read-coverage across all samples. The variant calling and filtration was done by UnifiedGenotyper and VariantFiltration tools (13, 14). Comparison between resulting sequence and the reference genome GRCh37 sequence was done using UnifiedGenotyper which applies a Bayesian approach and produces a VCF file containing single nucleotide variants (SNVs) and insertions/deletions (indel) variants. To filter out low quality variants from VCF file, Variant Filtration tool was used, and for variant annotation we used Ensembl Variant Effect Predictor (VEP) (15). At the final step, a report has been generated for each sample, containing all amplicones with the reading depth for each one, with present mutation (if called) both on DNA and protein level and dbSNP identifier. The Integrated Genomics Viewer was used for visual evaluation of the data (16).

Results

We have analyzed approximately 11.9×10^8 bp sequence from 34 ALL patients (17 cALL and 17 aALL) by targeted NGS using TSACP. The average coverage of high-quality sequences was $2609 \times$ per amplicon. Ten genes were discarded due to insufficient coverage, therefore a total of 183 amplicons from 38 genes was used for subsequent analysis. Variants were identified in relation to the GRCh37 reference genome by applying a Bayesian approach and compared to public genetic variation databases.

We identified 72 different variants across the samples in both coding and non-coding targeted regions, out of which 37 (21 in cALL, 22 in aALL) variants were in the coding regions and 35 (28 in cALL, 30 in aALL) outside of the targeted amplicons. Among the 37 different variants in the coding region, we found 3 different types of insertions/deletions (indels) in 3 cALL patients and in 2 aALL patients, while in the non-coding regions we found 9 different types of indels (in 9 cALL and in 8 aALL patients). We also identified 34 different single nucleotide variants-SNVs (in 18 cALL and in 20 aALL samples) in the coding and 26 different SNVs (in 19 cALL and in 22 aALL patients) in the non-coding regions.

Additionally, in all patients' samples, we identified a total of 330 (157 in cALL, 173 in aALL) variants (including synonymous variants) in the coding regions, (median per patient: 9, range: 6–12; median per cALL: 9, range: 6–12; median per aALL: 10, range: 7–12) (Figure 1). In the non-coding regions, we found 429 (211 cALL, 218 aALL) (median per patient: 13 range: 10–15; median per cALL: 13, range: 10–14; median per aALL: 13, range 10–15) (Figure 1).

Only mutations located within the coding regions were considered for further analysis, and from those only protein-changing mutations (nonsense, frameshift and missense (NFM) mutations). The total number NFM mutations was 95 (45 in cALL, 50 in aALL), median per patient 2, range: 1–7 (median per cALL: 2, range: 1–6; median per aALL: 3, range: 1–5). The majority of patients had no more than 3 NFM mutations, whereas only one cALL patient (#4) had 6 (Figure 2).

Our analysis revealed that 21 different genes had at least one NFM mutation in the coding regions (17 in cALL, 15 in aALL). Out of these, we identified variants in 6 cALL-specific genes (*CDKN2A*, *GNAQ*, *HRAS*, *PTPN11*, *AKT1*, and *ERBB2*) and 4 genes containing NFM mutations only in aALL patients (*NRAS*, *CSF1R*, *RET*, and *FLT3*). Mutations identified in the coding regions of following targeted genes: *KIT* (5 cases), *HNF1A*, *STK11* and *KRAS* were present in at least 4 cases (more than 10%), whereas substitution variants *KDR* Q472H and *TP53* P72R were detected in at least 18 cases (more than 50%). The

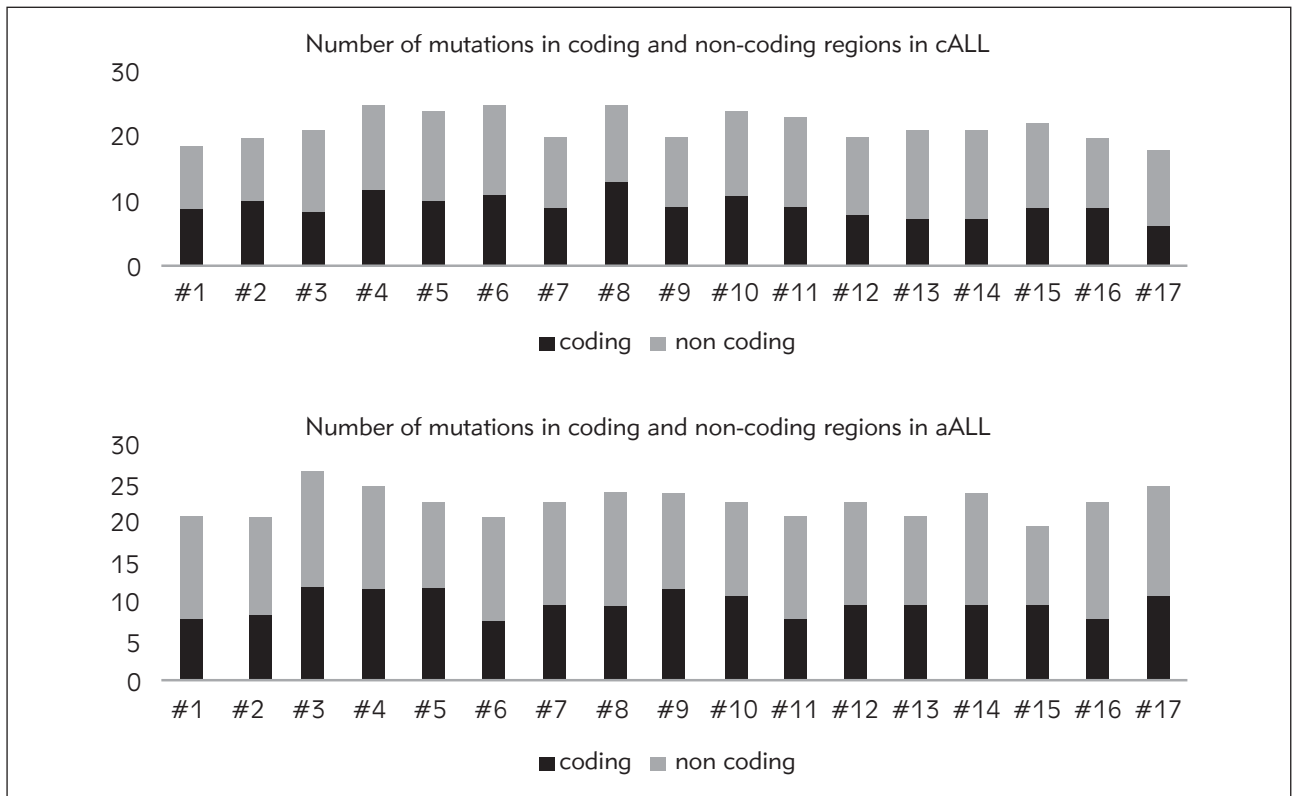


Figure 1 Total number of mutations in coding and non-coding regions identified by targeted NGS in cALL and aALL patients.

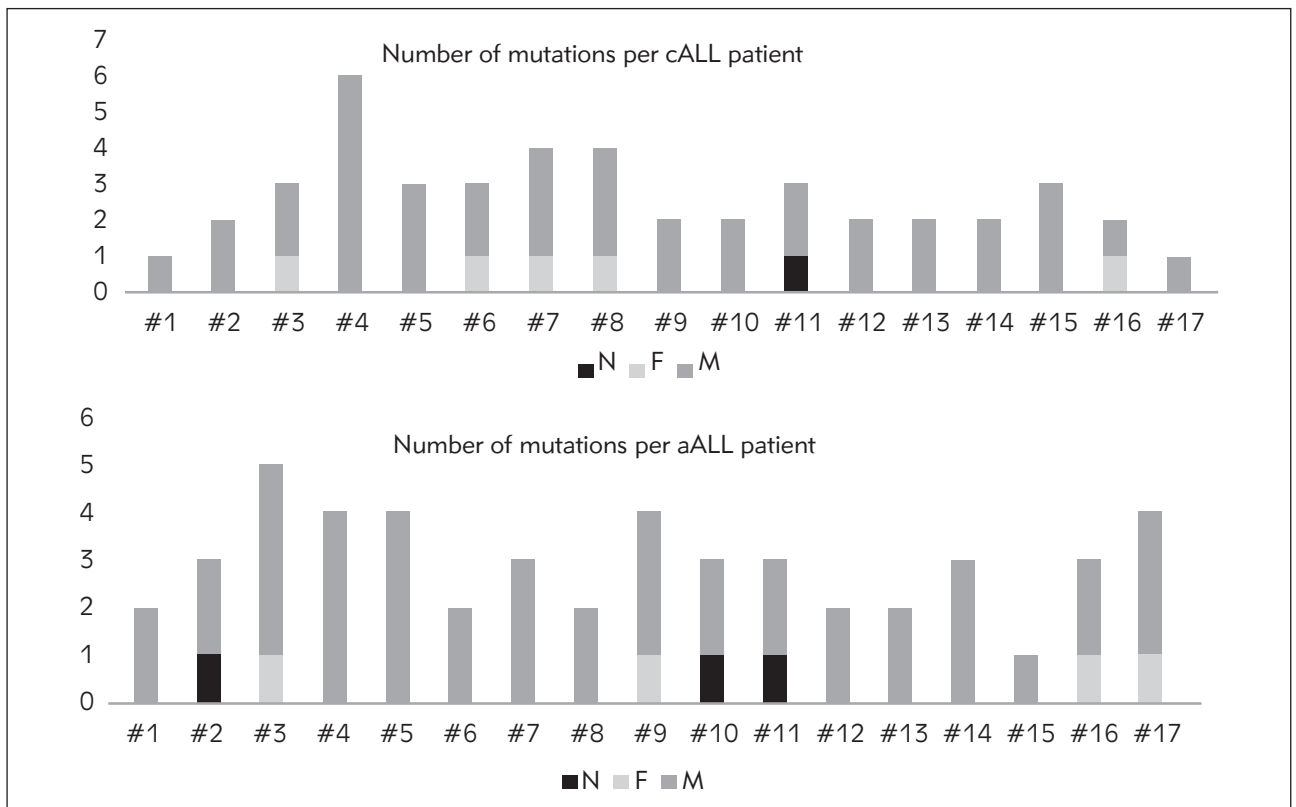


Figure 2 Distribution of nonsense (N), frameshift (F), and missense (M), mutations in the coding regions of targeted genes per cALL and aALL patient.

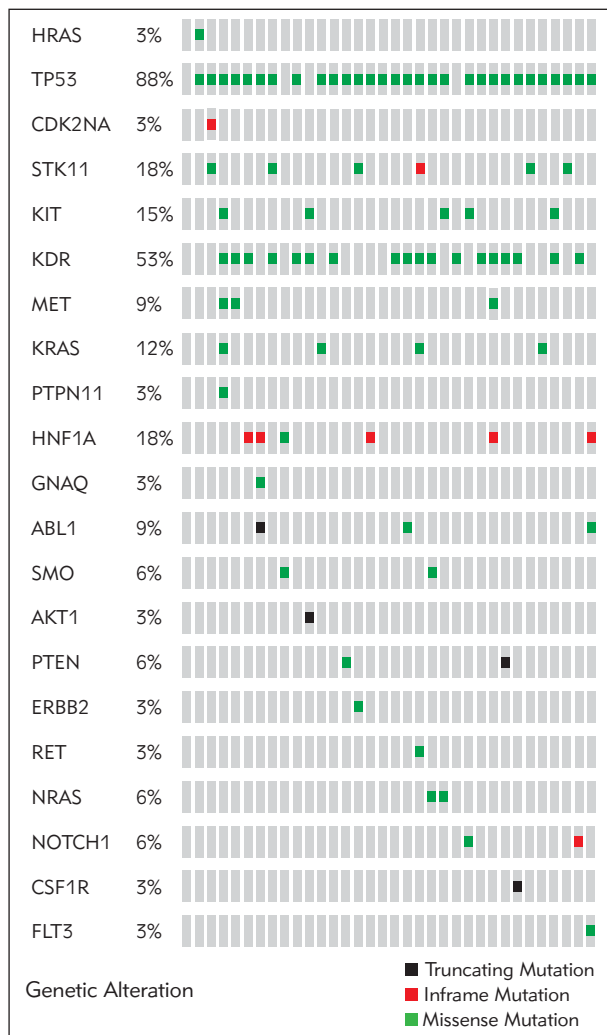


Figure 3 OncoPrint showing the distribution of genetic alterations in 38 targeted tumor suppressor and oncogenes in 34 ALL patients. The type of mutations are labeled in the color legend, particular genes in rows and tumor samples in columns.

list of the genes and their mutational types per patients are represented in OncoPrint (Figure 3).

In seven cALL patients and in eight aALL patients we detected 16 unreported NFM mutations in 10 genes (Table III and Table IV). The largest number of new mutations were detected in the *STK11* gene with four, followed by *ABL1* gene with three and *NOTCH1* gene with two mutations. These mutations were prevalently substitutions – missense type, but we also detected three nonsense and three frameshift truncating mutations.

Discussion

Acute lymphoblastic leukemia represents hematopoietic malignancy whose main feature is its clinical heterogeneity reflecting the heterogeneity

that exists on the genetic level. As the development of full-blown leukemia implies a multistep process of gradual accumulation of genetic and epigenetic alterations, ALL represents a mixture of the sub clones, characterized by a special combination of the mutations (17). Each mutation, characterized as »driver« or »passenger« mutation, in its own way contributes to complete leukemic phenotype and clinical characteristics. In order to study such a complex nature of the disease next generation sequencing (NGS) methodology was used, enabling the detection of new somatic mutations that are contributing to the pathogenesis of ALL.

In this study, the application of TSACP cancer panel to analyze the mutational pattern of childhood and adult ALL samples we have analyzed the role of genes previously described primarily in solid tumors. Moreover, by applying targeted re-sequencing method we have achieved a high accuracy in variant detection, with an average coverage of 2609× per amplicon. High coverage is required for detection of somatic mutations in the samples with large number of sub clones, characteristic for hematological malignancies.

We have detected 95 potentially protein-changing variants, (45 in cALL and 50 in aALL patients). Our finding of low number of mutations in both cALL (median per patient 2, range 1–6) and aALL (median per patient 3, range 1–5) is in accordance with previous studies with reported frequency of 0–7 mutations per patient (18, 19). Moreover, in comparison to other types of both adult and childhood cancers, acute leukemias were described as low mutation rate cancers (20, 21). In particular, in many of our patients we were not able to detect any of the mutations, excluding common germline polymorphism in *TP53*, and in *KDR* gene (Table III and Table IV). Polymorphism P72R in *TP53* gene, characteristic for 70% of European population, was found in 32 patients, while eighteen out of 34 contained Q472H variant in *KDR* gene (22).

In this study, we have noticed that commonly mutated genes belong to Ras/RTK signaling pathway, which is in accordance with previously published data (18, 23). Deregulation of Ras signaling pathway is very common feature among all cancers, because activated RAS proteins affect multiple downstream pathways (Raf/MEK/ERK and PI3K/Akt), and thus deregulate many important cellular processes (24).

One of the main mechanism of Ras deregulation is through acquisition of oncogenic mutations in 3 RAS genes: *NRAS*, *KRAS*, and *HRAS* (25). In our study, we identified hotspot mutations in *HRAS* gene affecting Glycine at G48, in *KRAS* gene affecting Glycine at positions G12 and G13, and in *NRAS* gene the same amino-acid at position G12 and Alanine at A59. Identified mutations were mutually exclusive. This observation is in accordance with the traditional concept according to which only one

Table III Mutations identified in childhood ALL patients using NGS.

Sample No.	Mutation detected by MySeq	Mutation Status	dbSNP	COSMIC
1	TP53, c.215C>G, p.P72R	Homozygous	rs1042522	COSM250061
2	TP53, c.215C>G, p.P72R	Homozygous	rs1042522	COSM250061
	HRAS, c.142G>A, p.G48R	Heterozygous		COSM55555612
3	TP53, c.215C>G, p.P72R	Heterozygous	rs1042522	COSM250061
	STK11, c.1023G>T, p.L341F	Heterozygous	unreported	unreported
	CDKN2A, c.175_176insG, p.V59fs*61	Heterozygous		COSM13715
4	KDR, c.1416A>T, p.Q472H	Heterozygous	rs1870377	COSM149673
	TP53, c.215C>G, p.P72R	Heterozygous	rs1042522	COSM250061
	KIT, c.1621A>C, p.M541L	Heterozygous	rs3822214	COSM28026
	MET, c.3029C>T, p.T1010I	Heterozygous	rs56391007	COSM707
	KRAS, c.35G>A, p.G12D	Heterozygous	rs121913529	COSM521
	PTPN11, c.205G>A, p.E69K	Heterozygous	rs397507511	COSM13013
5	MET, c.3029C>T, p.T1010I	Heterozygous	rs56391007	COSM707
	KDR, c.1416A>T, p.Q472H	Heterozygous	rs1870377	COSM149673
	TP53, c.215C>G, p.P72R	Homozygous	rs1042522	COSM250061
6	KDR, c.1416A>T, p.Q472H	Heterozygous	rs1870377	COSM149673
	TP53, c.215C>G, p.P72R	Heterozygous	rs1042522	COSM250061
	HNF1A, c.864_865insC p.P292fs [*] 25	Heterozygous		COSM4611384
7	TP53, c.215C>G, p.P72R	Homozygous	rs1042522	COSM250061
	HNF1A, c.864_865insC p.P292fs [*] 25	Heterozygous		COSM4611384
	GNAQ, c.842A>G, p.E281G	Heterozygous	unreported	unreported
	ABL1, c.754C>A, p.Q252 [*]	Heterozygous	unreported	unreported
8	KDR, c.1416A>T, p.Q472H	Heterozygous	rs1870377	COSM149673
	TP53, c.215C>G, p.P72R	Heterozygous	rs1042522	COSM250061
	STK11, c.769delG, p.G257fs*28	Heterozygous	unreported	unreported
	STK11, c.802G>A, p.G268R	Heterozygous		COSM4559384
9	HNF1A, c.862G>T, p.G288W	Heterozygous	rs539507291	
	SMO, c.1916T>C, p.V639A	Heterozygous	unreported	unreported
10	KDR, c.1416A>T, p.Q472H	Heterozygous	rs1870377	COSM149673
	TP53, c.215C>G, p.P72R	Heterozygous	rs1042522	COSM250061
11	KIT, c.1621A>C, p.M541L	Heterozygous	rs3822214	COSM28026
	KDR, c.1416A>T, p.Q472H	Heterozygous	rs1870377	COSM149673
	AKT1, c.66G>A, p.22*stop	Heterozygous	unreported	unreported
12	KRAS, c.38G>A, p.G13D	Heterozygous	rs112445441	COSM532
	TP53, c.215C>G, p.P72R	Heterozygous	rs1042522	COSM250061
13	KDR, c.1416A>T, p.Q472H	Homozygous	rs1870377	COSM149673
	TP53, c.215C>G, p.P72R	Heterozygous	rs1042522	COSM250061
14	PTEN, c.64G>A, p.D22N	Heterozygous	unreported	unreported
	TP53, c.215C>G, p.P72R	Homozygous	rs1042522	COSM250061
15	ERBB2, c.2341C>T, p.R811W	Heterozygous	unreported	unreported
	STK11, c.1087A>G, p.T363A	Heterozygous	unreported	unreported
	TP53, c.215C>G, p.P72R	Homozygous	rs1042522	COSM250061
16	HNF1A, c.864_865insC p.P292fs [*] 25	Heterozygous		COSM4611384
	TP53, c.215C>G, p.P72R	Homozygous	rs1042522	COSM250061
17	TP53, c.215C>G, p.P72R	Homozygous	rs1042522	COSM250061

Table IV Mutations identified in adult ALL patients using NGS.

Sample No.	Mutation detected by MySeq	Mutation Status	dbSNP	COSMIC
1	KDR, c.1416A>T, p.Q472H	Heterozygous	rs1870377	COSM149673
	TP53, c.215C>G, p.P72R	Heterozygous	rs1042522	COSM250061
2	KDR, c.1416A>T, p.Q472H	Heterozygous	rs1870377	COSM149673
	TP53, c.215C>G, p.P72R	Homozygous	rs1042522	COSM250061
	ABL1, c.782G>A, p.W261*stop	Heterozygous	unreported	unreported
3	KRAS, c.38G>A, p.G13D	Heterozygous	rs121913488	COSM532
	RET, c.2383A>T, p.S795C	Heterozygous	unreported	unreported
	STK11, c.769delG, p.G257fs*28	Heterozygous	unreported	unreported
	KDR, c.1416A>T, p.Q472H	Heterozygous	rs1870377	COSM149673
	TP53, c.215C>G, p.P72R	Heterozygous	rs1042522	COSM250061
4	KDR, c.1416A>T, p.Q472H	Heterozygous	rs1870377	COSM149673
	TP53, c.215C>G, p.P72R	Heterozygous	rs1042522	COSM250061
	SMO, c.1916T>C, p.V639A	Heterozygous	unreported	unreported
	NRAS, c.176C>A, p.A59D	Heterozygous		COSM253327
5	TP53, c.215C>G, p.P72R	Homozygous	rs1042522	COSM250061
	NRAS, c.35G>A, p.G12D	Heterozygous	rs121913237	COSM564
	KIT, c.1621A>C, p.M541L	Heterozygous	rs3822214	COSM28026
	TP53, c.614A>T, p.Y205F	Heterozygous		COSM11351
6	KDR, c.1416A>T, p.Q472H	Heterozygous	rs1870377	COSM149673
	TP53, c.215C>G, p.P72R	Homozygous	rs1042522	COSM250061
7	KIT, c.1621A>C, p.M541L	Heterozygous	rs3822214	COSM28026
	NOTCH1, c.4690C>T, p.H1564Y	Heterozygous	unreported	unreported
	TP53, c.215C>G, p.P72R	Homozygous	rs1042522	COSM250061
8	KDR, c.1416A>T, p.Q472H	Heterozygous	rs1870377	COSM149673
	TP53, c.215C>G, p.P72R	Homozygous	rs1042522	COSM250061
9	MET, c.3029C>T, p.T1010I	Heterozygous	rs56391007	COSM707
	HNF1A, c.864_865insC	Heterozygous		COSM4611384
	KDR, c.1416A>T, p.Q472H	Heterozygous	rs1870377	COSM149673
	TP53, c.215C>G, p.P72R	Homozygous	rs1042522	COSM250061
10	PTEN, c.19G>T, p.E7*stop	Heterozygous		COSM5298
	KDR, c.1416A>T, p.Q472H	Heterozygous	rs1870377	COSM149673
	TP53, c.215C>G, p.P72R	Homozygous	rs1042522	COSM250061
11	CSF1R, c.2862C>A, p.C954*stop	Heterozygous	unreported	unreported
	KDR, c.1416A>T, p.Q472H	Heterozygous	rs1870377	COSM149673
	TP53, c.215C>G, p.P72R	Homozygous	rs1042522	COSM250061
12	TP53, c.215C>G, p.P72R	Homozygous	rs1042522	COSM250061
	STK11, c.1046A>G, p.E349G	Heterozygous	unreported	unreported
13	KRAS, c.35G>T, p.G12V	Heterozygous	rs121913529	COSM520
	TP53, c.215C>G, p.P72R	Heterozygous	rs1042522	COSM250061
14	KIT, c.1621A>C, p.M541L	Heterozygous	rs3822214	COSM28026
	KDR, c.1416A>T, p.Q472H	Heterozygous	rs1870377	COSM149673
	TP53, c.215C>G, p.P72R	Homozygous	rs1042522	COSM250061
15	TP53, c.215C>G, p.P72R	Heterozygous	rs1042522	COSM250061
16	KDR, c.1416A>T, p.Q472H	Heterozygous	rs1870377	COSM149673
	NOTCH1, c.4729_4734delGTGGTG	Heterozygous	unreported	unreported
	TP53, c.215C>G, p.P72R	Heterozygous	rs1042522	COSM250061
17	FLT3, c.2503G>T, p.D835Y	Heterozygous	rs121913488	COSM783
	HNF1A, c.864_865insC	Heterozygous		COSM4611384
	TP53, c.215C>G, p.P72R	Heterozygous	rs1042522	COSM250061
	ABL1, c.880A>G, p.K294E	Heterozygous	unreported	unreported

mutation in each pathway is sufficient for disease development. In one model of ALL genesis, it was suggested that G12D variant in *KRAS* is a first genetic event responsible for malignant transformation of hematopoietic stem cells (26). Still, the majority of studies have focused on *KRAS* and *NRAS* mutant forms and suggest that oncogenic RAS alone is insufficient to drive leukemogenesis and cooperating genetic events are necessary for full-blown leukemia. It was found that mutations affecting RAS gene family and entire Ras pathway as well, were associated with other aberrations (27). In our cohort of patients, it was the case in MLL1/AF4-driven leukemogenesis.

Deregulation of Ras/RTK signaling pathway can also occur due to constitutive activation of protein tyrosine kinase located upstream of RAS (25). Many studies described receptor tyrosine kinases as a key regulator of the process of hematopoiesis, as well as leukemogenesis (28). In our cohort of ALL patients, we have found mutations of missense type in receptor tyrosine kinase genes *FLT3* and *ERBB2* (at positions D835Y and R781W, respectively). Additionally, we identified variants in non-receptor tyrosine kinase *ABL1* gene including two stop gain mutations at position Q252* and W261*, as well as one substitution mutation K294E in one patient. All of these mutations are unreported, although mutations in *ABL1* gene have been previously associated with ALL occurrence or therapy resistance (7). We have also detected one mutation in *PTPN11* gene coding for non-receptor tyrosine phosphatase SHP2. SHP2 is a putative positive regulator of the Ras signaling pathway and mutations in *PTPN11* gene have been described as a «driver» mutations in B-ALL development (19).

Mutations in *HNF1A* gene, encoding transcriptional factor affect both Ras/RTK and Notch1 pathways. We have detected six patients with mutations in this gene; five frameshift mutations (P292fs*25), and 1 missense variant (G288W). In our study, these frameshift mutations were associated with the presence of *TEL/AML1* rearrangement in cALL, probably as a «second hit» mutation, as an additional genetic event required for development of full-blown leukemia (30, 31). Changes in the *HNF1A* gene are associated with liver/pancreatic tumors, not with hematological malignancies (32, 33).

Another gene whose mutations are not usually associated with hematological malignancies is *STK11* (*LKB1*) gene. *STK11* gene is encoding serine/threonine kinase protein, which has been involved in the cell cycle and apoptotic processes. It is assumed that this gene has the role of a classical tumor suppressor gene, because its loss-of-function somatic mutations lead to deactivation of the PI3K/Akt signaling pathway (34). Significant frequency of somatic mutations in *STK11* gene were reported only in lung and cervical tumors, while in other types of human cancers, the occurrence of these mutations is a sporadic

event (35, 36). In our study, it was one of the most mutated one with one frameshift (G257fs*28) and four missense mutations (G268R, L341F, E349G, and T363A) found in 6 patients. All of the detected *STK11* mutations were previously unreported, and mutations in this gene are not specific for leukemias.

In *NOTCH1* gene, we identified one missense and one frameshift mutation in 2 T-ALL patients. Activating mutations in Notch1 signaling pathway have been described as a crucial factor in T-ALL development and their identification could lead to prognostic marker discovery and therapy improvement (37, 38). In five patients, 2 cALL and 3 aALL, we identified M541L mutation in *KIT* gene that encodes proto-oncogene receptor tyrosine kinase, while in another receptor tyrosine kinase coding proto-oncogene *MET*, we detected T1010I mutation that was present exclusively among cALL patients. *KIT* gene belongs to receptor tyrosine kinase gene family involved in hematopoietic stem cells (HSCs) self-renewal and differentiation, suggesting that any activating mutation in the receptor could alter hematopoietic development. Mammalian cells transformed with *KIT* gene that contained activating mutations exhibited increased growth both *in vitro* and *in vivo*, suggesting important role of this mutation in leukemogenesis (39). Pathogenic mutations in *KIT* tyrosine kinase gene have been reported in various diseases. One of the recently published results emphasizes the role of M541L variant in the therapy response in chronic eosinophilic leukemia patients (40).

In conclusion, by using targeted NGS method in studying the mutational landscape of our cohort of ALL patients, we have found that low number of mutations are implicated in the pathogenesis of this disease. The impact of the detected mutations is focused on few key signaling pathways, primarily on Ras/RTK cascade. Our findings provide additional information for mutational portrait of ALL and the results could be used as a supplement to classical therapy stratification methods. In that way, we would be able to apply therapeutics that target specific signaling pathways in each individual patient. It is possible that better outcome among ALL patients of all ages could finally be accomplished through such personalized treatment approach.

Acknowledgments. This work was funded by the Ministry of Education, Science and Technological Development, Republic of Serbia (grant no. III 41004) and by the European Commission, EU-FP7-REGPOT-316088, 2013–2016. We thank colleagues from Seven Bridges Genomics for assistance with bioinformatics analysis.

Conflict of interest statement

The authors stated that they have no conflicts of interest regarding the publication of this article.

References

1. Pui CH, Robison LL, Look AT. Acute lymphoblastic leukaemia. *Lancet* 2008; 371: 1030–43.
2. Mrozek K, Harper DP, Aplan PD. Cytogenetics and molecular genetics of acute lymphoblastic leukemia. *Hematol Oncol Clin North Am* 2009; 23: 991–1010.
3. Milosevic G, Kotur N, Krstovski N, Lazic J, Zukic B, Stankovic B, Janic D, Katsila T, Patrinos GP, Pavlovic S, Dokmanovic L. Variants in TPMT, ITPA, ABCC4 and ABCB1 genes as predictors of 6-mercaptopurine induced toxicity in children with acute lymphoblastic leukemia. *J Med Biochem* 2018; 37: 320–7.
4. Moorman AV, Chilton L, Wilkinson J, Ensor HM, Bown N, Proctor SJ. A population-based cytogenetic study of adults with acute lymphoblastic leukemia. *Blood* 2010; 115: 206–14.
5. Toft N, Birgens H, Abrahamsson J, Bernell P, Griškevičius L, Hallböök H, et al. Risk group assignment differs for children and adults 1-45 yr with acute lymphoblastic leukemia treated by the NOPHO ALL-2008 protocol. *Eur J Haematol* 2013; 90: 404–12.
6. Mullighan CG. The molecular genetic makeup of acute lymphoblastic leukemia. *Hematology Am Soc Hematol Educ Program* 2012: 389–96.
7. Gasic V, Stankovic B, Zukic B, Janic D, Dokmanovic L, Krstovski N, Lazic J, Milosevic G, Lucafò M, Stocco G, Decorti G, Pavlovic S, Kotur N. Expression pattern of long Non-coding RNA growth arrest-specific 5 in the remission induction therapy in childhood acute lymphoblastic leukemia. *J Med Biochem* 2018; 38: 292–8.
8. www.bioinformatics.babraham.ac.uk/projects/fastqc. 2010; Available from: www.bioinformatics.babraham.ac.uk/projects/fastqc.
9. <http://code.google.com/p/ea-utils/>. 2011.
10. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009; 25: 1754–60.
11. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 2010; 26: 589–95.
12. <https://arxiv.org/abs/1303.5075>. 2013.
13. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 2011; 43: 491–8.
14. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010; 20: 1297–303.
15. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, et al. The Ensembl Variant Effect Predictor. *Genome Biol* 2016; 17: 122.
16. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nat Biotechnol* 2011; 29: 24–6.
17. Warner JK, Wang JC, Hope KJ, Dick JE. Concepts of human leukemic development. *Oncogene* 2004; 23: 7164–77.
18. Ebrahimi-Rad M, Khatami S, Ansari S, Jalylfar S, Valadbeigi S, Saghiri R. Adenosine deaminase 1 as a biomarker for diagnosis and monitoring of patients with acute lymphoblastic leukemia. *J Med Biochem* 2018; 37: 128–33.
19. Mullighan CG. The genomic landscape of acute lymphoblastic leukemia in children and young adults. *Hematology Am Soc Hematol Educ Program* 2014: 174–80.
20. Liu YF, Wang BY, Zhang WN, Huang JY, Li BS, Zhang M, et al. Genomic Profiling of Adult and Pediatric B-cell Acute Lymphoblastic Leukemia. *EBioMedicine* 2016; 8: 173–83.
21. Marjanovic I, Kostic J, Stanic B, Pejanovic N, Lucic B, Karan-Djurasevic T, et al. Parallel targeted next generation sequencing of childhood and adult acute myeloid leukemia patients reveals uniform genomic profile of the disease. *Tumour Biol* 2016; 37: 13391–401.
22. Bodian DL, McCutcheon JN, Kothiyal P, Huddleston KC, Iyer RK, Vockley JG, et al. Germline variation in cancer-susceptibility genes in a healthy, ancestrally diverse cohort: implications for individual genome sequencing. *PLoS One* 2014; 9: e94554.
23. Ding LW, Sun QY, Tan KT, Chien W, Mayakonda A, Yeoh AEJ, et al. Mutational Landscape of Pediatric Acute Lymphoblastic Leukemia. *Cancer Res* 2017; 77: 390–400.
24. Bos JL. Ras oncogenes in human cancer: a review. *Cancer Res* 1989; 49: 4682–9.
25. Shannon K. The Ras signaling pathway and the molecular basis of myeloid leukemogenesis. *Curr Opin Hematol* 1995; 2: 305–8.
26. Zhang J, Wang J, Liu Y, Sidik H, Young KH, Lodish HF, et al. Oncogenic Kras-induced leukemogenesis: hematopoietic stem cells as the initial target and lineage-specific progenitors as the potential targets for final leukemic transformation. *Blood* 2009; 113: 1304–14.
27. Stam RW. The ongoing conundrum of MLL-AF4 driven leukemogenesis. *Blood* 2013; 121: 3780–1.
28. Reilly JT. Receptor tyrosine kinases in normal and malignant haematopoiesis. *Blood Rev* 2003; 17: 241–8.
29. Branford S, Rudzki Z, Walsh S, Grigg A, Arthur C, Taylor K, et al. High frequency of point mutations clustered within the adenosine triphosphate-binding region of BCR/ABL in patients with chronic myeloid leukemia or Ph-positive acute lymphoblastic leukemia who develop imatinib (STI571) resistance. *Blood* 2002; 99: 3472–5.
30. Yamamoto T, Isomura M, Xu Y, Liang J, Yagasaki H, Kamachi Y, et al. PTPN11, RAS and FLT3 mutations in childhood acute lymphoblastic leukemia. *Leukemia Res* 2006; 30: 1085–9.
31. Zuna J, Madzo J, Krejci O, Zemanova Z, Kalinova M, Muzikova K, et al. ETV6/RUNX1 (TEL/AML1) is a fre-

- quent prenatal first hit in childhood leukemia. *Blood* 2011; 117: 368–9.
32. Luo Z, Li Y, Wang H, Fleming J, Li M, Kang Y, et al. Hepatocyte nuclear factor 1A (HNF1A) as a possible tumor suppressor in pancreatic cancer. *PLoS ONE* 2015; 10: e0121082.
33. Pilati C, Letouze E, Nault JC, Imbeaud S, Boulai A, Calderaro J, et al. Genomic profiling of hepatocellular adenomas reveals recurrent FRK-activating mutations and the mechanisms of malignant transformation. *Cancer Cell* 2014; 25: 428–41.
34. Gao Y, Ge G, Ji H. LKB1 in lung cancerigenesis: a serine/threonine kinase as tumor suppressor. *Protein Cell* 2011; 2: 99–107.
35. Sanchez-Cespedes M, Parrella P, Esteller M, Nomoto S, Trink B, Engles JM, et al. Inactivation of LKB1/STK11 is a common event in adenocarcinomas of the lung. *Cancer Res* 2002; 62: 3659–62.
36. Wingo SN, Gallardo TD, Akbay EA, Liang MC, Contreras CM, Boren T, et al. Somatic LKB1 mutations promote cervical cancer progression. *PLoS ONE* 2009; 4: e5137.
37. Ferrando AA. The role of NOTCH1 signaling in T-ALL. *Hematology Am Soc Hematol Educ Program* 2009; 2009: 353–61.
38. Zhu YM, Zhao WL, Fu JF, Shi JY, Pan Q, Hu J, et al. NOTCH1 mutations in T-cell acute lymphoblastic leukemia: prognostic significance and implication in multifactorial leukemogenesis. *Clin Cancer Res* 2006; 12: 3043–9.
39. Hussain SR, Raza ST, Babu SG, Singh P, Naqvi H, Mahdi F, et al. Screening of C-kit gene mutation in acute myeloid leukaemia in Northern India. *Iran J Cancer Prev* 2012; 5: 27–32.
40. Iurlo A, Gianelli U, Beghini A, Spinelli O, Orofino N, Lazzaroni F, et al. Identification of kit (M541L) somatic mutation in chronic eosinophilic leukemia, not otherwise specified and its implication in low-dose imatinib response. *Oncotarget* 2014; 5: 4665–70.

Received: December 16, 2018

Accepted: March 22, 2019