

CLINICAL LABORATORY METHOD FOR DETECTION OF *IGHV* MUTATION STATUS IN PATIENTS WITH CLL VALIDATED BY IGBLAST AND IMGTV-QUEST

Lasabova Z.¹, Plank L.², Flochova E.³, Burjanivova T.¹, Vanochova A.¹, Mihok L.⁴, Ilencikova D.⁴

¹Department of Molecular Biology, ²Department of Pathological Anatomy, ³Department of Hematology and Transfusiology, Jessenius Faculty of Medicine, Comenius University, University Hospital in Martin and

⁴National Cancer Institute in Bratislava, Slovak Republic

ABSTRACT

Chronic lymphocytic leukemia (CLL) is the most frequent type of adult leukemia in Western countries. Recently, new molecular prognostic markers like 17p deletion, 11q deletion, 13q deletion, trisomy 12, the mutational status of the immunoglobulin variable heavy chain genes (*IGHV*) genes, expression of ZAP-70 and CD38 were identified as prognostically significant. The CLL patients with mutated *IGHV* have a more favorable prognosis while non-mutated cases with the mutation's number less than 2% compared to the germline sequence suffer from more aggressive diseases. Here, we describe a clinical laboratory method for the detection of the mutation status of *IGHV* in patients with CLL using reverse transcription PCR and dideoxysequencing, and the evaluation using two immunoglobulin databases IMGTV-QUEST and IgbLAST. We analyzed 37 different clonal rearrangements in 35 patients. Using two different databases, we identified 13 mutated and 24 non-mutated clones. The most preferred subfamilies were VH1, VH3, and VH4. The CLLs using the subfamily 1-69 were all non-mutated. Unlike previous reports, there were no significant differences between the used databases observed. The clinical trials are already incorporating new prognostic molecular markers such *IGHV* mutational status, so it is important to use standardized clinical laboratory methods and databases for a reliable identification of the mutation status in CLL.

Key words: chronic lymphocytic leukemia, mutational status, dideoxysequencing, immunoglobulin databases

INTRODUCTION

Chronic lymphocytic leukemia (CLL) is the most frequent type of adult leukemia in Western countries which is characterized by accumulation and proliferation of functionally incompetent monoclonal B-lymphocytes with a typical phenotype CD5, CD19, CD23 and CD79a positive and with low surface expression of BCRs. There are two prognostic scoring systems currently used in the clinical praxis, the Rai and the Binet systems; however, the prognosis of patients with CLL is extremely variable (1, 2). The recognition that many patients with early-stage disease develop an aggressive clinical course led to the search for additional risk stratification tools. In recent years, new molecular prognostic markers like 17p deletion, 11q deletion, 13q deletion, trisomy 12, the mutational status of the immunoglobulin variable heavy chain genes (*IGHV*) genes, expression of ZAP-70 and CD38 were identified as prognostic significant (1, 2).

The rearrangement of *IGHV* gene segments occurs during the hematopoiesis and is based on allelic exclusion and random combination of V, D and J gene segments resulting in the surface membrane expression of unique functional receptor. Therefore, PCR can be used to identify lymphocyte populations derived from single cell by detecting

Address for correspondence:

Zora Lasabova, RNDr. PhD, Department of Molecular Biology, Jessenius Faculty of Medicine, Comenius University, Martin, Kalinciaka 2, 03861 Vrutyky, Slovakia. Phone: 00421-43-4286517; e-mail: lasabova@jfm.uniba.sk

the unique V-D-J rearrangement. The V segments are divided into 7 families based on sequence homology and primers specific for these families can be designed for PCR amplification (3, 4). The detection of the mutation status of IGHV is performed by the PCR analysis and direct DNA sequencing (4, 5). The CLL patients with mutated IGHV have a more favorable prognosis while non-mutated cases with the mutation's number less than 2% compared to the germline sequence suffer more aggressive diseases (1, 2, 5, 6).

The correct determination of the mutational status of IGHV gene segment obtained by PCR strictly depends on the comparison of sequenced nucleotides to the closet germline sequence and is possible only when using bioinformatics tools and databases. The most comprehensive and most updated are databases IMGT (<http://imgt.cines.fr>) and IgBLAST (<http://www.ncbi.nlm.nih.gov/igblast/>) that can be used for determination of the mutational status of IGHV genes. The IMGT uses unique numbering for anchor positions of framework (FWR) and complementarity determining region (CDR). For *VH*, the anchor amino acids are cystein C23, conserved tryptophan W41, conserved hydrophobic amino acid L89 and second W118. 104 and 118 are anchor position for the CDR3. Anchor position for CDR1 are amino acids at positions 26 and 39, and CDR2 55 and 66, respectively (7). IgBLAST is also suitable for analysis of immunoglobulin V region sequences using BLAST search algorithm and reports germline V,D and J segments with framework and CDR annotation according the Kabat numbering when CDR1 is anchored by amino acids 31 and 35b, CDR2 by 50 and 65 and CDR3 by 95 and 102 (8). There were reported about 4% differences in the interpretation of the mutational status when different databases were used (9,10).

The aim of this study was the development of clinical laboratory method for the detection of the mutation status of IGHV in patients with CLL using peripheral blood and RNA reverse transcription. For validation of this method, we compared the alignment results released by the IMGT/V-QUEST and IgBLAST and analyzed the differences in sequence identity.

PATIENTS AND METHODS

Patients

We obtained peripheral blood clinical samples from patients with CLL diagnosed in the Department of Hematology and Transfusiology in University Hospital in Martin after informed consent. The study was approved by the institutional ethical board.

Nucleic acid preparation

Peripheral blood was collected in EDTA tubes and 3 ml were separated by density gradient centrifugation using Ficoll-Hipaque (Sigma, USA) and counted. The RNA was isolated using TRIzol reagent (MRC, USA) according to manufacturer's protocol. Briefly, 20 - 30 million cells were lysed by adding 300 μ l TRIZol reagent and proteins were removed by adding of 300 μ l of chlorophorm solution with subsequent vortexing and centrifugation, RNA was precipitated using 1 volume of isopropanol, dried and resolved in DEPC water (Gibco, USA). The RNA concentration was measured at 260 nm and the quality of RNA was controlled by ethidium-bromide stained agarose gel electrophoresis.

cDNA synthesis

RNA was reverse transcribed into cDNA using Verso cDNA kit (Thermo Scietific, United Kingdom) according to the manufacturer's instructions. Briefly, 1 μ l hexanucleotide random primer was added to 1 μ g RNA and the mixture was denatured at 65°C for 10 minutes. Then after, the premix with dNTPs, reverse transcriptase and RT enhancer was added and the mixture was incubated for 50 minutes at 47°C. The reverse transcriptase was inactivated for 10 min at 75°C.

The PCR reaction

We used degenerated PCR primers from the seven IGVHFR1 gene segment families and a mixture of J gene segment families (11, 12). The PCR reaction was performed with

1.0 µl cDNA or 0.2 µg of template DNA in 25 µl of reaction mixture containing 2 mmol/L MgCl₂, 10 pmol/L of each forward and reverse primer, 0.5 mmol/L of each of the four dNTPs, 2.5 µl of 10x PCR Buffer (ABgene®, United Kingdom) and 1 unit of Thermostart Taq polymerase (ABgene®, United Kingdom). The PCR mix was subjected to hot start at 95 °C for 8 min followed by 35 cycles of denaturation at 95 °C for 30 seconds, annealing at 58 °C for 30 seconds and extension at 72 °C for 30 seconds, and the final step of 4 minutes at 72 °C. The PCR products from single PCR reactions were analyzed in 2% ethidium-bromide stained agarose gel and the appropriate DNA fragments were excised from the agarose gel and purified using NucleoSpin Extract II kit (Macherey-Nagel, United Kingdom) prior to dideoxysequencing reaction.

Dideoxysequencing and sequence alignment

The purified PCR product from predominant bands in PCR analysis were directly sequenced using the Big Dye Terminator kit v.1.1 (Applied Biosystems, USA) and ABI PRISM 3100 Genetic Analyzer (Applied Biosystems, USA). The sequences were processed in the program Chromas 1.5 (Technelysium, Australia) and aligned to IMGT/V-QUEST database (<http://imgt.cines.fr>). The nucleotides involved were counted from the codon 1 to codon 104 according to IMGT unique numbering and the percentage was calculated based on the ratio between the numbers of nucleotide differences according the ERIC recommendations (8). As mutated were identified sequences which differ more than 2% from the germline sequence. We aligned the sequences to the BLAST Ig database according to the KABAT numbering and compared the results obtained by both databases.

RESULTS

Detection of clonality using seven different primers from the FR1 of VH gene segment families

As shown in the Fig.1, we used seven different VH family primers in RT-PCR and isolated tumor clones from 35 patients with CLL. We identified a clone by PCR when we have seen one or two discrete bands in the ethidium-bromide stained agarose gel (Fig.1B, C). In a tested control sample we observed bands in each lane (Fig. 1A) corresponding to the polyclonal background.

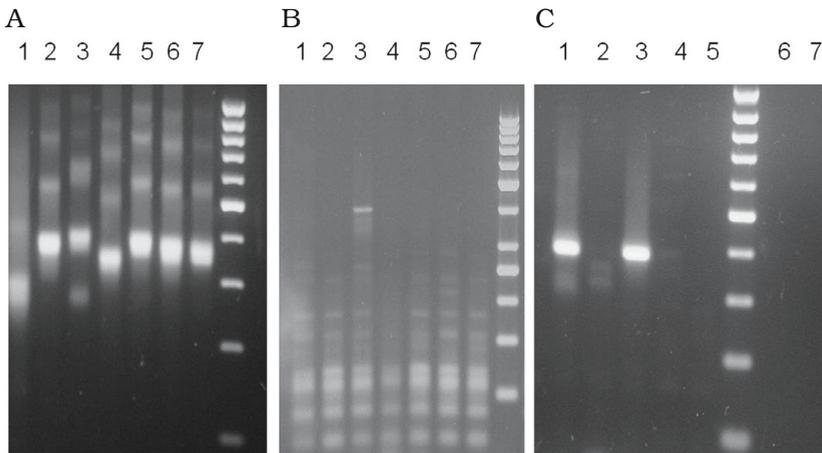


Fig. 1 Ethidium-bromide stained gels after PCR of cDNA from patients with CLL when the IGHV gene family 1,2,3,4,5,6 and 7 specific primers were used. **A.** polyclonal **B.** monoclonal; in the lane 3, a positive clone of VH3 gene family is shown, which is indicative for clonality; **C.** double rearrangement specific for families 1 and 3

Analysis of the IGHV sequences by IgBLAST and IMGT/V-QUEST databases

For sequence evaluation, we used IgBLAST and IMGT/V-QUEST. Only the IMGT/V-Quest interprets the result as productive or unproductive rearrangement. A rearranged sequence is productive if no stop codon has been detected in the V-D-J-region of VH as it is shown in an example in figure 2 (Fig. 2A). All analyzed PCR products with the exception of two were productive rearrangements. In one case, we found unproductive rearrangement of 4-34*01 allele with 96.34% identity to the germline sequence with a stop codon in place of the W118 (Fig. 2B). In another case, we were not able to detect complete junction between V-D-J, which was interpreted as “no rearrangement found” by IMGT (Fig. 3). We were not able to find other clonal segments in these cases so we decided to include them in the subsequent evaluations.

Using above mentioned methods, we identified 37 different clones in 35 patients; 13 clones showed somatic hypermutation according to the both bioinformatics protocols and 24 clones showed less than 2% difference to the closet germline sequence. These were interpreted as non-mutated (Table 1). In two CLL cases, we detect two different clones in one clinical sample - 1-8*08 and 3-11*03 (Figure 1.C) with 100% and 99.58 % identity to the germline sequence, respectively. The other case with two clones showed 4-39*01 and 4-2*01 with 100% and 99.6% identity, respectively. The most preferred subfamilies were VH1 (12 cases), VH3 (12 cases) and VH4 (8 cases). The clones from the VH1 subfamily were non-mutated with the exception of one case (subfamily 1.46). The CLLs using the subfamily 1-69 were all non-mutated (Table 1).

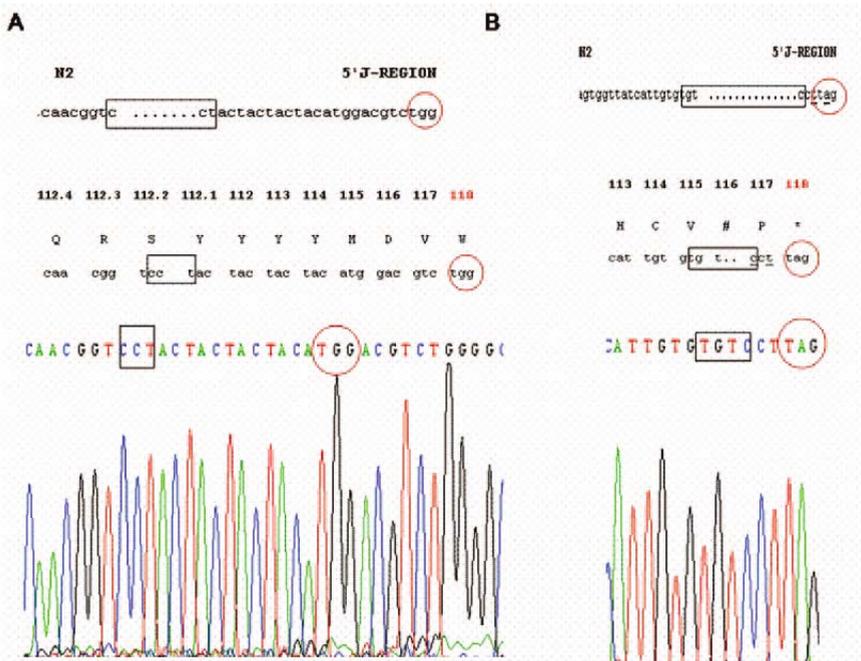


Fig. 2 A. Example of a normal productive rearrangement with an in-frame junction (black rectangle in A) and W118 as anchor of the CDR3 (red circle in A) **B.** Non-productive rearrangement with an out-of-frame junction (black rectangle in B) and stop codon resulting from the frameshift (red circle in B); example from the patient Nr.1

Table 1 Clonality results and mutation status identified by RT-PCR from peripheral blood in 35 CLL cases

CLL patient	V _H gene family segment	Sequence identity (%) BLAST	Sequence identity (%) IMGT	M mutated N non-mutated
1.	4-34*01	96	96.34	M
2.	3-33*01	89.5	89.1	M
3.	6-1*01	100	100	N
4.	1-3*01	100	100	N
5.	1-3*01	100	100	N
6.	3-48*03	100	100	N
7.	1-46*01	97.1	96.94	M
8.	1-69*01	100	100	N
9.	6-1*01	95.7	95.59	M
10.	1-2*04	99.6	99.57	N
11.	4-61*02	91	89.45	M
12.	1-69*01	100	100	N
13.	4-31*02	96	95.81	M
14.	4-34*01	92.9	93.23	M
15.	3-53*01	90.2	91.02	M
16a.	1-8*08	100	100	N
16b.	3-11*03	99.6	99.58	N
17.	3-21*01	97.9	97.56	M, but 3-21
18.	3-7*01	91.1	91.25	M
19.	1-3*01	100	100	N
20.	3-33*01	100	100	N
21.	1-18*01	100	100	N
22.	1-69*01	100	100	N
23.	4-34*01	89.8	89.92	M
24.	1-2*04	99.2	99.16	N
25.	4-39*01	100	100	N
26.	3-15*01	98	98.03	N
27.	3-43*01	99.6	99.6	N
28.	1-69*12	100	100	N
29.	1-69*01	99.2	100	N
30.	4-34*01	100	100	N
31.	6-1*01	95.1	94.92	M
32.	3-11*01	100	100	N
33.	3-33*01	100	100	N
34.	7-4*02	96.3	96.15	M
35a	4-39*01	100	100	N
35b	4-2*01	99.6	99.60	N

Comparison of the alignment results from IgBLAST and IMGT/V-QUEST

The valuable lengths of sequenced PCR products were different. We were able to analyze in 29 cases at least some nucleotides in the FWR1 and complete sequences of CDR1, FWR2, CDR2 and FWR3 for sequence homology search with the closest germline sequences. In 10 cases, we started the alignment from the CDR1 sequence. When the IMGT was used, the length of the evaluated sequences was from 227 to 249 bp with FWR1; the shortest sequence without FWR1 was 190 bp in length. Using the IgBLAST, the sequences with FWR1 were from 243 to 255 bp in length, the shortest sequence without FWR1 was 197 bp in length (data not shown). The comparison by IMGT is performed from the position 1 to the position 104 of the second conserved cystein. The comparison by IgBLAST was performed to the closest germline V sequence what included up to 6 nucleotides following the anchor second cystein (Fig. 4). To obtain greater certainty in interpretation of results, we collected and aligned the sequences with their closets germline counterpart using the IMGT database and the “new” version of the GeneBank/IgBlast. Because of these different alignment protocols, we evaluated differences in the mutational status when the mutations occurred at the end of the V sequence followed the second cystein at position 104. Only functional sequences marked as productive rearrangements, one with no rearrangement and one with the stop codon (as shown in the figures 2 and 3) were taken into consideration. We found only 19 subtle discrepancies between results obtained by IgBLAST and IMGT/V-QUEST databases, 16 of them were bellow 0.5% without any effect on the mutational status. In three cases, the discrepancies were between 0.8 and 1.5 % without any effect on the mutational status too (Table 1).The discrepancies were based on different bp lengths evaluated by IgBLAST and IMGT/V-QUEST based on the differences in the numbering. In all cases with 100% homology with the exception of one, an equal degree of identity was found (Table 1). In this one case, there was different number of mutations reported because the mutation was placed behind the second cystein (Fig.4). These mutations were without any effect on the final mutational status which was interpreted as non-mutated (patient number 29 in the table 1, Fig. 4).

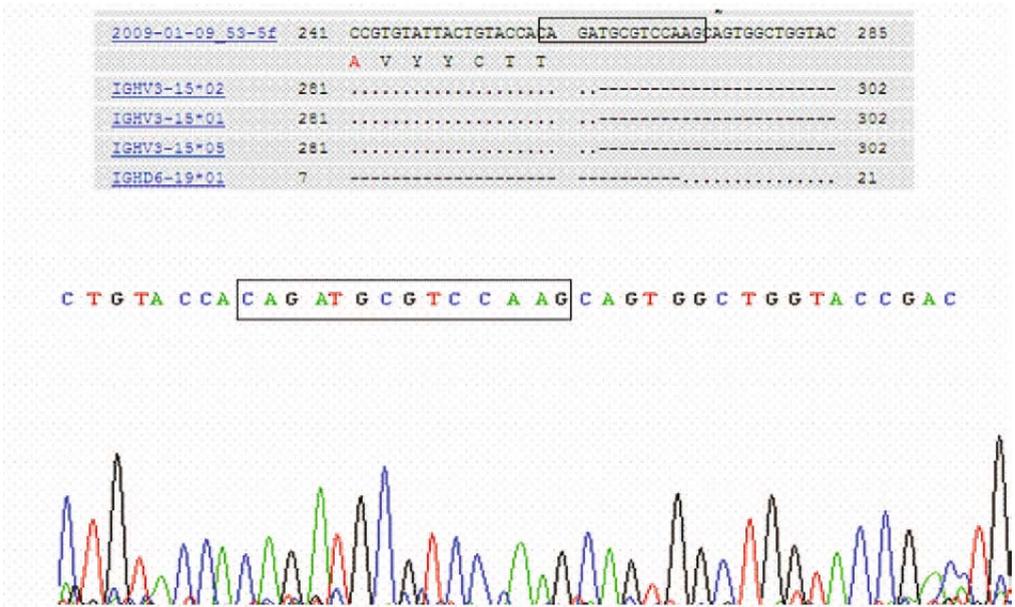


Fig. 3 Example of “no rearrangement found” sequence with the V-D junction only; example form the patient Nr.26

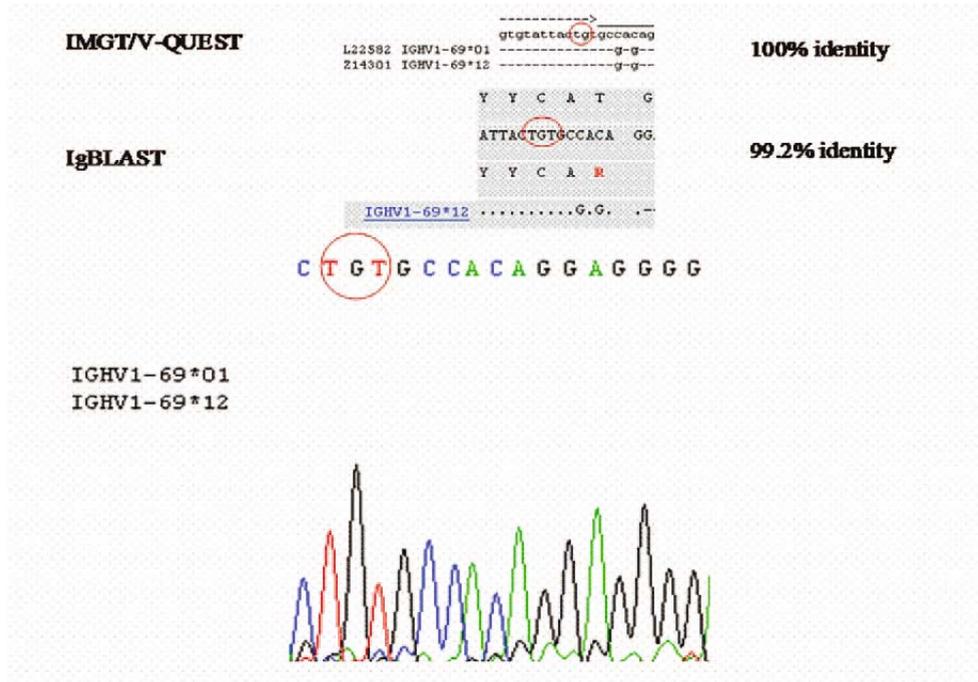


Fig. 4 Differences in the identity calculation between the IMGT/V-QUEST algorithm and IgBLAST in one 1-69 positive case. The calculation up to second cystein codon (red circle) results into 100% identity between the germline and patient sequence. IgBLAST calculates the complete VH region which include the two mutations behind the second cystein resulting into 99.2% identity; example from patient Nr. 29

DISCUSSION

In present study, we established the method for detection of mutational status of IGHV genes in patients with CLL starting with RNA extraction from peripheral blood. We used primers from different VH gene subfamilies and single RT-PCR reaction for analysis of each subfamily as described by Pekova et al. (12). cDNA has an advantage over gDNA, of preferentially identifying the functional productive rearrangements and moreover it is overexpressed in B cells leading to higher detection sensitivity (8, 12). When we used cDNA, the repeated sequencing can be avoided when two rearrangements are present. Although when we used the cDNA, we found in two cases double productive rearrangements. The interpretation in our cases was not complicated, because both patients had productive non-mutated rearrangements. The analysis of DNA is more frequently used, can be more easily performed and we consider this application for the analysis of archived formalin-fixed paraffin-embedded (FFPE) section. The disadvantage is the possible presence of more bands during clonality testing a finding of nonproductive rearrangements. The reported recommendations for laboratory protocols used to detect clonality in FFPE sections can also applied in determining IGHV mutational status of DNA isolated from FFPE section of patients with CLL (3).

Using both databases and above mentioned criteria, we found 16 very subtle differences up to 0.5 % when used different databases. These differences were with no effect on the mutational status. In three cases were the differences 0.8, 0.82 and 1.5% without any effect on the mutational status too. This result is different from previous reports, when there were observed discordant results leading to the changes of the

prognostic category based on the mutational status and/or different CLL alleles identified by the particular database (9,10). The first reason that we did not observe these differences is the very small number of our cases. The other reason may be also the updating of the IgBLAST that is now able to detect mutational status including the deletions and insertion more efficiently than previously (<http://www.ncbi.nlm.nih.gov/igblast/>). According to our experience, the combination of both databases can be useful in the analysis of the mutation status of the IGHV genes. We need also more samples to be analyzed in order to say whether the updating of IgBLAST reduces differences between databases.

In our study, we were able to analyze the complete CDR1, FWR2, and CDR2 up to anchor second cystein in CDR3 which represents a limited length of VH sequence. Previously, Marasca et al. (15) reported analysis of limited length of VH sequence when a degenerate FR2 primer was used and which includes the 3' portion of FR2, the entirety of CDR2, FR3, CDR3, and a small portion of FR4 region. This VH part corresponds to 56% of the complete VH sequence and it was showed that patients with a percentage of somatic mutation less than 2% in the FR2-CDR3 segment analyzed have a worse prognosis (13). We can also take into consideration a part of the VH sequence starting with FWR2 what is important when we consider analysis of the FFPE tissue.

The clinical trials are already incorporating new prognostic molecular markers such IGHV mutational status. In some circumstances, a flow cytometry is a very suitable method for clonality testing in lymphoproliferative diseases (14), but in case of CLL, the non-mutated IGHV determined by PCR and DNA sequencing is considered to predict the shorter progression free survival and overall survival after treatment with purine nucleoside analogs and alkylating agents with or without rituximab (15, 16), so it is important to use standardized clinical laboratory methods and databases for a reliable identification of the mutation status in CLL.

REFERENCES

1. Shanafelt TD. Predicting clinical outcome in CLL: how and why. *Hematology. Am Soc Hematol Educ Program.* 2009;421-9.
2. Hallek M, German CLL Study Group, 2008. Prognostic factors in chronic lymphocytic leukemia. *Ann Oncol.* 2008; Suppl.4: iv51-iv53.
3. Van Dongen, JJM, Langerak AW, Bruggemenn M, Evans PAS, Hummel M et al. Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: Report of the BIOMED-2 concerted action BMH4-CT98-3936. *Leukemia* 2003; 17: 2257-2317.
4. Szankasi P, Bahler DW. Clinical laboratory analysis of immunoglobulin heavy chain variable region genes for chronic lymphocytic leukemia prognosis. *J Molec Diagn.* 2010; 12: 244-249.
5. Tobin G, Thunberg U, Karlsson K, Murray F, Laurell A, et al. Subsets with restricted immunoglobulin gene rearrangements features indicate a role of antigen selection in the development of chronic lymphocytic leukemia. *Blood* 2004; 104: 2879-2885.
6. Hamblin TJ, Davis Z, Gardiner A, Oscier DG, Stevenson FK. Non-mutated Ig VH genes are associated with a more aggressive form of chronic lymphocytic leukemia. *Blood* 1999; 94: 1848-1854.
7. Giudicelli V, Chaume D, Lefranc MP. IMGIT/V-QUEST, an integrated software program for immunoglobulin and T cell receptor V-J and V-D-J rearrangement analysis. *Nucleic Acid Res.* 2004; 32: W435-W440.
8. Ghia P, Stamatopoulos K, Belessi C, Moreno C, Stilgenbauer S, et al.: ERIC recommendations on IGHV gene mutational status analysis in chronic lymphocytic leukemia. *Leukemia* 2007; 21: 1-3
9. Pekova S, Baran-Marszak F, Schwarz J, Matoska V. Mutated or non-mutated? Which database to choose when determining the IgVH mutational status in chronic lymphocytic leukemia? *Haematologica* 2006; 91: e11-e12.
10. Davi F, Rosenquist R, Ghia P, Belessi C, Stamatopoulos K. Determination of IGHV gene mutational status in chronic lymphocytic leukemia: bioinformatics advances meet clinical needs. *Leukemia* 2008; 22, 212-214; doi:10.1038/sj.leu.2404969;

11. Souto-Carneiro MM, Krenn V, Hermann R, Konig A, Muller- Hermelink HK. IgVH genes from different anatomical regions, with different histopathological patterns of rheumatoid arthritis patient suggest cyclic-reentry of mature synovial B-cells in the hypermutation process. *Arthritis Res.* 2000; 2: 303-314.
12. Pekova S, Markova J, Pajer P, Dvorak M, Cetkovsky, P, Schwarz J. Touch-down reverse transcriptase-PCR detection of IgVH rearrangement and Sybr-green-based real-time RT-PCR quantitation of minimal residual disease in patients with chronic lymphocytic leukemia. *Mol Diagn.* 2005; 9: 23-34.
13. Marasca R, Maffei R, Morselli M, Zucchini P, Castelli I, et al. Immunoglobulin mutational status detected through single-round amplification of partial VH region represents a good prognostic marker for clinical outcome in chronic lymphocytic leukemia. *J Molec Diagnostics.* 2005; 7: 566-574.
14. Valekova L, Fedorova J, Rumanova S, Stasko J, Flochova E, et al. Gastric MALT lymphoma: the role of flow cytometry. *Acta Med Martiniana* 2007; 3: 24-28.
15. Kjarfan-Dabaja MA, Chavez JC, Khorfan KA, Pinilla-Ibarz J. Clinical and therapeutic implications of the mutational status of IGHV in patients with chronic lymphocytic leukemia. *Cancer* 2008; 13: 898-906.
16. Shanafelt TD. Predicting clinical outcome in CLL: how and why? *Hematology Am Soc Hematol Educ Program.* 2009; 421-429.

Acknowledgement: This work was supported by the project "Center of Excellency for Perinatology Research II" (ITMS code 26220120036) co-financed from EU sources

Received: April, 7, 2011

Accepted: May, 31, 2011