FEBS Letters 426 (1998) 266–270 FEBS 20124

A cosmid and cDNA fine physical map of a human chromosome 13q14 region frequently lost in B-cell chronic lymphocytic leukemia and identification of a new putative tumor suppressor gene, Leu5

Bagrat Kapanadze[a], Vladimir Kashuba[b], Anna Baranova[a,c], Omid Rasool[d], Wim van Everdink[e], Yie Liu[d], Alexander Syomov[a], Martin Corcoran[f], Andrey Poltaraus[g], Vadim Brodyansky[a], Natalia Syomova[a], Alexey Kazakov[a], Rachel Ibbotson[f], Anke van den Berg[e], Rinat Gizatullin[b,g], Ludmila Fedorova[g], Galina Sulimova[a], Alexander Zelenin[g], Larry Deaven[h], Hans Lehrach[i], Dan Grander[d], Charles Buys[e], David Oscier[f], Eugene R. Zabarovsky[b,g], Stephan Einhorn[d], Nick Yankovsky[a,c,*]

[a] *Vavilov Institute of General Genetics, 117809 Moscow, Russia*
[b] *MTC and GCR Karolinska Institute, S-104 01 Stockholm, Sweden*
[c] *Bion Ltd., 117809 Moscow, Russia*
[d] *Research Laboratory of Radiumhemmet, Department of Oncology/Pathology, Karolinska Hospital and Institute, S-171 76 Stockholm, Sweden*
[e] *University of Groningen, 9713 AW Groningen, The Netherlands*
[f] *Royal Bournemouth Hospital, Bournemouth BH7 7DW, UK*
[g] *Engelhardt Institute of Molecular Biology, 117894 Moscow, Russia*
[h] *LANL MS M 888, Los Alamos, NM 87545, USA*
[i] *MPI für Molekulare Genetik, Ihnestrasse 73, D-14195 Berlin, Germany*

Received 11 March 1998

Abstract B-cell chronic lymphocytic leukemia (B-CLL) is a human hematological neoplastic disease often associated with the loss of a chromosome 13 region between RB1 gene and locus D13S25. A new tumor suppressor gene (TSG) may be located in the region. A cosmid contig has been constructed between the loci D13S1168 (WI9598) and D13S25 (H2-42), which corresponds to the minimal region shared by B-CLL associated deletions. The contig includes more than 200 LANL and ICRF cosmid clones covering 620 kb. Three cDNAs likely corresponding to three different genes have been found in the minimally deleted region, sequenced and mapped against the contigged cosmids. cDNA clone 10k4 as well as a chimeric clone 13g3, codes for a zinc-finger domain of the RING type and shares homology to some known genes involved in tumorigenesis (RET finger protein, BRCA1) and embryogenesis (MID1). We have termed the gene corresponding to 10k4/13g3 clones LEU5. This is the first gene with homology to known TSGs which has been found in the region of B-CLL rearrangements.

Key words: Tumor suppressor gene; Location; B-cell chronic lymphocytic leukemia; Zinc binding; Human

1. Introduction

Genes that normally function to prevent or suppress malignancy are known as 'tumor suppressor genes' or TSG [1]. The first example of a TSG in humans was the RB1 gene [2]. The bi-allelic functional loss of this gene is associated with the childhood neural retina tumor called retinoblastoma and the gene itself is located on human chromosome 13q14.3. This genomic region is also frequently lost in B-cell chronic lymphocytic leukemia (B-CLL), the most frequent hematopoetic neoplastic disease in elderly [3,4]. B-CLL patients with 13q14 deletions were, however, found to retain a functional RB1 gene in the malignant cells. These observations indicated that a defect of a gene telomeric to RB1 may be involved in the malignant transformation of B cells [3–5]. The minimal region shared by B-CLL associated deletions includes loci D13S272 to D13S25 at the telomeric side [6–8]. According to other data [9] the deletions expand from D13S272 in the opposite direction to the D13S273 locus. Our recent data have minimized the region of interest to a 130-kb area around the locus D13S319 [10] and further down to an approximately 10-kb region [11].

One strategy to discover the putative TSG gene involves a fine mapping of the genomic region lost in leukocytes of B-CLL patients, followed by screening of cDNA clones belonging to this region and subsequent characterization of a probable candidate TSG. LANL and ICRF chromosome 13 specific cosmid libraries have been shown to be good sources to construct such maps [12,13].

Previously a YAC-STS map of the region, frequently lost in B-CLL [13,14], as well as a smaller contig with YAC derived cosmids covering the region [10] has been constructed. A recent map of the 13q14.3 region included 13 PAC clones, 42 cosmids and several EST markers without any homology to known genes [15]. Here we describe a cosmid contig (based on genomic libraries LANL and ICRF) covering the minimal region shared by B-CLL-associated deletions, cDNA screening corresponding to YAC ICRF 61c1, as well as fine mapping and sequencing of the cDNA clones. One of the cDNA clones, 10k4, contains a zinc-finger domain and seems to be an additional interesting candidate for being a B-CLL TSG.

2. Materials and methods

2.1. Chromosome 13 cosmid libraries

Two genomic chromosome 13-specific cosmid libraries LA13NC01 and ICRF C108 (average size of insert 38 kb) have been characterized in our previous work [12]. The cosmids subcloned from YACs have been described in [10]. The subcloned cosmids were both included into

*Corresponding author. Fax: (7) (095) 1351289.
E-mail: yankovsky@vigg.ru

PII S0014-5793(98)00357-3

the clone grid and used as probes but they are presented only in the probe bar on Fig. 1. These subcloned cosmid names on Fig. 1 start from a letter ('a', 'b' or 'c' for YAC ICRF 61c1, 126a9 and CEPH 922a8, respectively).

2.2. YAC clones

The CEPH and ICRF YAC clones and contig used to make the cosmid contig and screen cDNA libraries have been described before [14,16].

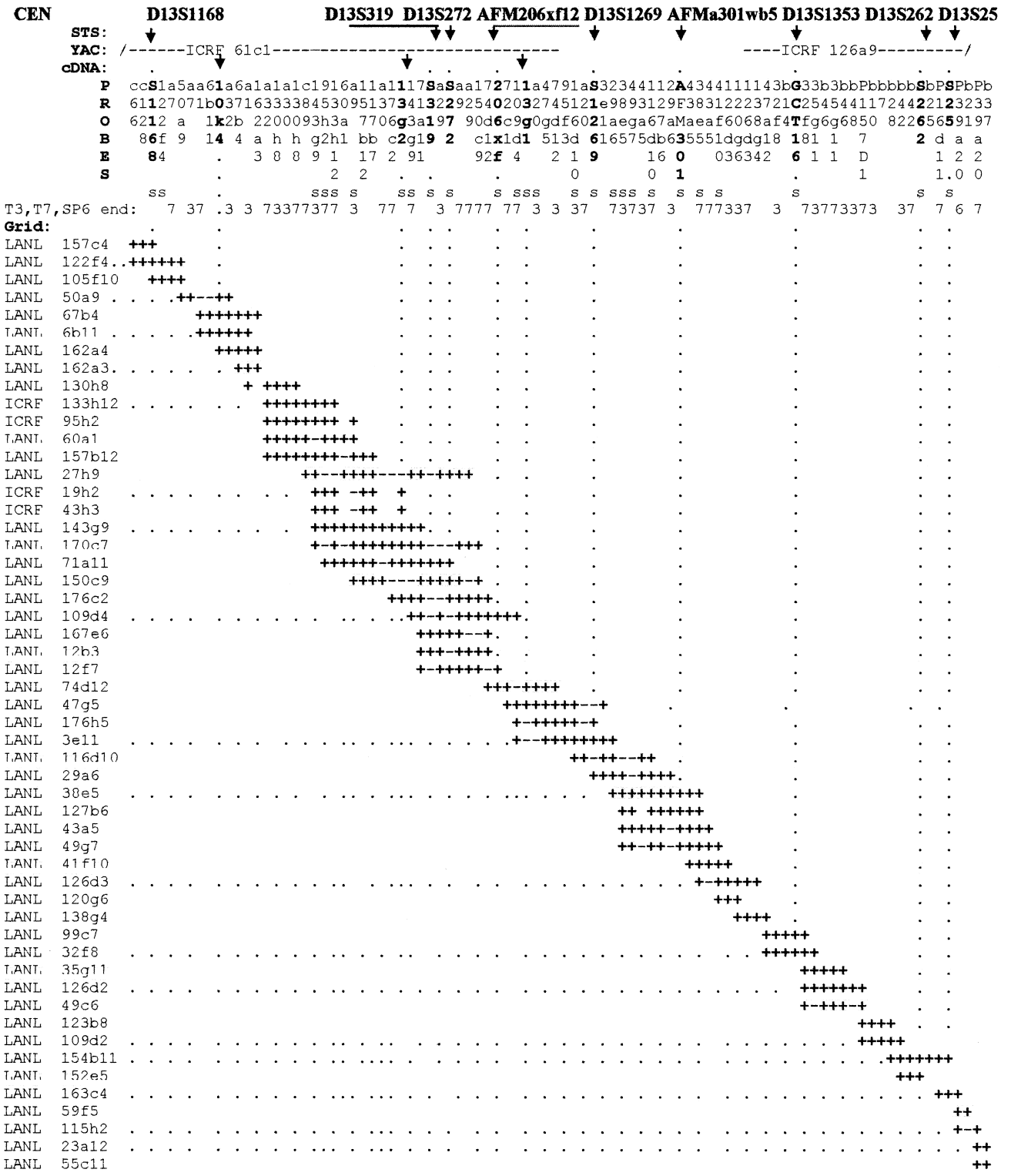

Fig. 1. Cosmid contig between loci D13S1168 and D13S25. STS markers assigned to the cosmids are shown in the upper line. The next line represents YAC structure. The YAC structure is shown as superposition of cosmids subcloned from the YAC or screened with the YAC as probe. Vertical arrow marks cDNA probe name; 's' character below the probe name indicates that the region corresponding to the probe is sequenced. If only one of the vector promoters was used to generate riboprobe, the name of the promoter is shown under the probe name. (+) or (−) indicates positive or negative results of hybridization. Horizontal dot lines mark cosmids forming minimal tilling path. Shaded rectangle marks the cosmids containing 13g3/CLL2-5 gene.

2.3. FISH

Biotin labelled YAC or cosmid DNA to human chromosome metaphase spreads of the lymphoblastoid cell line ROM-2 in combination with R–G banding was done as described by others [17].

2.4. Clone grid

High-density cosmid clone grids were used with each clone double spotted. The grids were from ICRF (UK) and RZPD (DB) (36864 spots on 22×22-cm nylon membrane) or prepared in the study (1536 spots on 8×12-cm membrane) and processed as described earlier [18].

2.5. YAC probes

YAC probes for cosmid grid screening were prepared according to [19] and labelled by random priming. End labelling and hybridization of oligonucleotide probes to the cosmid grids were carried out as described elsewhere [14], and positive cosmids were checked by Southern or PCR analysis.

2.6. Computer program

Computer program 'Contig' [20] was used for cosmid contig assembly. The program is available via ftp://ftp.icnet.uk/icrf-public/GenomeAnalysis/icrf_contigv2.-tar.Z

3. Results

3.1. A cosmid sublibrary in the RB1-D13S25 region of chromosome 13

Three YAC clones (CEPH745e3, ICRF 61c1 and ICRF101d6) were used as probes for screening of the cosmid library grids to find more than 400 clones belonging to the region. We checked 95 cosmids randomly picked up from the subset by FISH to human metaphase spread. Seventy-one of the cosmids belonged to 13q14. Forty-eight out of the 71 cosmids had a unique location in the human genome. For 23 out of the 71 an additional location(s) in the human genome was found. Some double locations were found on chromosome 13 (+13q12, +13q21, +13q34) including a double signal in 13q14. The revealed high level of imperfect duplications involving the region indicates the need for caution in the interpretation of further cDNA screening results and assignment of DNA probes from the region.

3.2. A cosmid contig between loci D13S1168 and D13S25

As starting points for chromosome walking in the region we used STS assigned cosmids in addition to FISH characterized cosmids corresponding to both YAC CEPH 745e3 and ICRF 61c1. Riboprobes were generated from 172 individual cosmid insert ends and hybridized to the cosmid subset representing the 13q14 region or to the total chromosome 13 library grid. The hybridization data gained with the riboprobes as well as with 21 oligonucleotide probes for the STS and EST in the region have been entered and interpreted by the 'Contig' computer program. Though the region was mainly covered by the LANL contigged cosmids after the subset analysis, we found some gaps in the cosmid map and branching points. To bridge the gaps we used a set of cosmids [10] prepared by subcloning of YACs covering the region.

The resulting cosmid contig between loci D13S1168 and D13S25 is presented in Fig. 1. Twelve EST and STS markers have been assigned to the contigged cosmids. A minimal tilling path for the region can be formed of 19 LANL and ICRF cosmids (marked with a horizontal dot line in Fig. 1). The estimated total length of the contigged region is approximately 620 kb. Nucleotide sequences have been established for some cosmid insert ends marked by the 's' letter line in Fig. 1. The complete information on the contig hybridization and sequence data is available from N.Y.

3.3. cDNA library screening, characterization and mapping

In screening for cDNA clones we concentrated on the region between D13S273 and AFMA301wb5 which is the focus of deletions associated with B-CLL [6,9,10]. Screening a pre-B lymphocyte cDNA library was done with the labelled ICRF YAC 61c1 (370 kb) bridging without gaps the interval between loci D13S1150 and D13S272 and nearly reaching AFMA301wb5 on the telomeric end (see Fig. 1). The YAC corresponds to the centromeric half of the region associated with B-CLL deletions. Screening of placenta and fetal brain libraries (Stratagene) was done with labelled fragments of cosmids 1a, 32a, 30a.

Five clones were isolated and one of them, 10k4 (1.4 kb in size), was analyzed here. It appeared that this clone and CLL2-5 gene, described in [11], contains the same sequences. cDNA clone 10k4 contains an ORF coding for a protein of 407 amino acids. A middle part of the amino acid sequence contains a zinc-finger domain of the RING type and shares homology with some human transcription factors taking part in early embryogenesis and tumor progression (Fig. 2). The maximal homology (P-value = $1.8e^{-20}$) was shown for XPRF transcription factor described in [22]. This protein is a product of MID1 gene, which acts in early human embryogenesis and is deleted in Opitz G/BBB syndrome (midline formation defect). The gene is evolutionarily conserved as can be concluded from a protein homology to A33 protein from ribbed newt and to F54G8.4 protein from *C. elegans* (see Fig. 2). Some of the proteins sharing homology to 10k4 are present in human B-lymphocytes, which are the target cells for B-CLL. These are transcription factors Rpt-1, Staf-50, and a Sjögren syndrome antigen RO-SS/A. 10k4 also shares a homology with BRCA-1, which is a well established TSG. Beside 10k4 is the minimally deleted genome region shown for B-CLL in our recent study [11]. That makes cDNA 10k4 a probable candidate for B-CLL TSG.

Three cDNA clones have been found in the pre-B lymphocyte cDNA library. These are 13g1 (1.6 kb), 13g2 (2.7 kb), and 13g3 (1.2 kb). Clones 13g1 and 13g2 contain a poly-A tail. Northern analysis shows that 13g1 hybridizes to four different bands corresponding to 1.3-, 1.6-, 2.9- and 3.5-kb mRNA; 13g2 did not produce a clear Northern signal; 13g3 corresponds to 1.7-, 2.4-, and 7.5-kb mRNA.

All the cDNA clones have been sequenced. The sequence analysis revealed that 13g3 cDNA contains sequences partly

Protein	P-value	
Leu5		**ME**L**LEEDLTCPICCSLFDDPRVLPCSHNFCKKCL**
XPRF	$1.8e^{-20}$	**MET**L**ESELTCPIC**L**ELFEDP**LL**LPCAH**SL**CFNCA**
A33	$7.9e^{-17}$	**ED**DFT**EDLTCPLCRSLF**K**EPVILECGHNFC**KH**CI**
Rpt-1	$9.5e^{-17}$	**LE**MIK**EEVTCPIC**L**ELLK**EP**VSADC**N**H**SFCRAC**I
XNF7	$1.1e^{-16}$	AGDFA**EELTCPLCVELFKDP**VMVAC**GHNFCRSC**I
RFP	$3.7e^{-15}$	A**EC**L**QQE**T**TCPVC**LQY**FVEP**MM**L**DC**GHNIC**CA**CL**
Staf-50	$1.7e^{-13}$	KVDIEK**EVTCPIC**L**ELLTEP**LS**LDCGHSFCQAC**I
F54G8.4	$3.7e^{-13}$	**LE**KI**EQLLTCPIC**LDR**Y**KQ**PKLLPCQHTFCYPCL**
Ro SS/A	$4.5e^{-11}$	**LTMMWEEVTCPIC**LDP**F**VEPVSIEC**GHSFCQEC**I**
BRCA1	$9.6e^{-6}$	**L**HA**MQ**KI**LECPIC**L**ELIK**EP**VSTKC**DHIFCKFCM**

Fig. 2. Comparison of the predicted amino acid sequences of a Leu5 zinc-finger domain fragment with database protein sequences. Amino acid identities and conservative substitutions are shown in bold. All protein sequences except Leu5 were retrieved from the NCBI/GenBank database.

overlapping with 10k4 (EMBL Acc. AJ224819) and CLL2-5 clones [11]. 13g3 cDNA had a homology to a number of ESTs. These ESTs were also found homologous to 10k4 [11]. The cDNA sequence analysis shows that clone 13G1 is virtually identical to p48/Hip (GenBank Acc. U28918) [21]. Clone 13g2 shows no clear homology in the GeneBank sequences, except for an Alu repeat presented in 13g2.

ESTs corresponding to 13g1, 13g2 and 10k4/13g3 cDNA clones have been established. The primer structure and PCR conditions are the following:

EST13g1: 13g1-1 5′-CTCAGAGCAGTAATCTTCC-3′,
13g1-2 5′-GAAGATAATCCCTACCCC-3′, $t = 55–57$, 256 bp.
EST 13g2: 13g2-1 5′-AGTAAGGCAGTGAGCTAGG-3′,
13g2-2 5′-GGCATAAACATCTGCCTGC-3′, $t = 55–57$, 241 bp.
EST 13g3: 13g3-1 5′-TCAAGTGTCCTACATGCCG-3′,
13g3-2 5′-AGGCATCCCTTTCCTGAGC, $t = 55–57$, 268 bp.

EST 13g1 is located between 206xf12 and D13S1269, and EST 13g2 and 10k4/13g3 are located between D13S1168 and D13S319 (see Fig. 1). The cDNA clones also have been mapped against the cosmid panel by hybridization. cDNA 13g1 corresponds to cosmids 47g5 and 74d12; 13g2 corresponds to cosmids 143g9, 170c7, 71a11 and 176c2. cDNA 10k4/13g3 corresponds to cosmids 6b11 and 67b4 (see Fig. 1).

4. Discussion

We have constructed a cosmid contig in the 13q14 region between the RB1 gene and D13S25 locus, which is expected to contain a putative B-CLL TSG [4–9]. We faced branching in the left part of the contig covering 200 kb and presented by five-cosmid tilling path (122F4, 6b11, 162a3, 60a1, 71a11). Indeed each of the cosmids has at least one more location in the human genome according to FISH. Contig branching was also shown in a study of chromosome 16 abnormalities in acute non-lymphocytic leukemia [23]. This branching was due to a chromosome 16 specific repeat. Some sites of translocations and inversions in persons affected by leukemia coincided with the sites of the repeat location. It could also be a possible explanation for B-CLL associated rearrangements found in the 13q14 region [4–9].

In a previous study we found two genes (Leu1 and Leu2) that span the minimally deleted region. Detailed mutational analysis of these genes in a large B-CLL material, however, did not reveal any cases with point mutations or small mutations in any of these genes. It can therefore not be ruled out that other adjacent or overlapping genes may be of importance. In the present study we describe three additional cDNA clones that belong to three different loci on human chromosome 13 within the region of interest.

The distances between the loci are approximately 40 kb (see Fig. 2). The cDNA clones likely correspond to three different genes from their different location and according to the sequence and Northern hybridization data.

BLASTN and FASTA analysis of the sequence derived from the 10k4 clone revealed that it is homologous to 15 clones from different libraries made from spleen, colon, placenta as well as T and B lymphocytes. Each of these clones, however, corresponds to short fragments of the 10k4 sequence. Interestingly, the deduced 10k4 protein sequence shares significant homology to the RET finger protein and the well-known TSG BRCA1. Furthermore, 10k4 is bordering the minimally deleted region at 13q described in a recent study that we performed on 209 B-CLL clones [11]. Taking the data together we believe that the gene corresponding to the 10k4 clone could be an additional interesting candidate for B-CLL TSG. We have termed this gene as Leu5.

Acknowledgements: The work was supported by the Russian National Human Genome Program, by the US DOE Grant OR00033-93CIS016, The Royal Swedish Academy of Science Grant #1450, by the Dutch Cancer Society, the Kay Kendall Leukaemia Research Fund UK, the Cancer Society of Sweden, and by the King Gustav V. Jubilee Foundation. R.G. and V.K. were recipients of fellowships from the Concern Foundation in Los Angeles and the Cancer Research Institute in New York. The authors are grateful to L.G. Savina for excellent technical assistance and to S.A. Borinskaya for help in preparation of the manuscript.

References

[1] Marshall, C.J. (1991) Cell 64, 313–326.
[2] Weinberg, R.A. (1992) Cancer Surveys 12, 43–57.
[3] Liu, Y., Szekely, L., Grander, D., Soderhall, S., Julisson, G., Gahrton, G., Linder, S. and Einhorn, S. (1993) Proc. Natl. Acad. Sci. USA 90, 8697–8701.
[4] Hawthorn, L.A., Chapman, R., Oscier, D. and Cowell, J.K. (1993) Oncogene 8, 1415–1419.
[5] Brown, A.G., Ross, F.M., Dunne, E.M., Steel, C.M. and Weir-Thompson, E.M. (1993) Nat. Genet. 3, 67–72.
[6] Liu, Y., Hermanson, M., Grander, D., Merup, M., Wu, X., Heyman, M., Rasool, O., Juiliusson, G., Gahrton, G., Detlofsson, R., Nikiforova, N., Buys, C., Soderhall, S., Yankovsky, N., Zabarovsky, E. and Einhorn, S. (1995) Blood 86, 1911–1915.
[7] Devilder, M.C., Francois, S., Bosic, A., Moreau, A., Mellerin, M.P., Le Paslier, D., Bataille, R. and Moisin, J.P. (1995) Cancer Res. 55, 1355–1357.
[8] Bullrich, F., Veronese, M.L., Kitada, S., Jurlander, J., Caligiuri, M.A., Reed, J. and Croce, C.M. (1996) Blood 88, 3109–3115.
[9] Stilgenbauer, S., Nickolenko, E., Leupolt, G., Weib, G., Schroder, M., Fischer, K., Bentz, M., Lichter, P. and Dohner, H. (1996) Ann. Oncol. 7 (abstr.), 41.
[10] Corcoran, M.M., Rasool, O., Liu, Y., Iyengar, A., Grander, D., Ibbotson, R.E., Merup, M., Wu, X., Brodyansky, V., Gardiner, A.C., Juliusson, G., Chapman, R.M., Ivanova, G., Tiller, M., Gahrton, G., Yankovsky, N., Zabarovsky, E., Oscier, D.G. and Einhorn, S. (1998) Blood 91, 1382–1390.
[11] Liu, Y., Corcoran, M., Rasool, O., Ivanova, G., Ibbotson, R., Grander, D., Iyengar, A., Baranova, A., Kashuba, V., Merup, M. and Wu, X. et al. (1997) Oncogene 15, 2463–2473.
[12] Kapanadze, B.I., Brodyanskii, V.M., Baranova, A.V., Sevast'yanov, S.Y., Fedorova, N.D., Kurskov, M.M., Kostina, M.A., Mironov, A.A., Sineokii, S.P., Zakhar'ev, V.M., Grafodatskii, A.S., Modyanov, N.N. and Yankovskii, N.K. (1996) Genetics (in Russ.) 32, 331–340.
[13] Fisher, S.G., Cayanis, E., Russo, J.J., Sunjevaric, I., Boukhgalter, B., Zhang, P., Yu, T., Rothstein, R., Warburton, D., Edelman, I.S. and Efstradiatis, A. (1994) Genomics 21, 525–537.
[14] Brodyanskii, V.M., Sulimova, G.E., Udina, I.G., Aitova, S.S., Shaikhaev, G.O., Sharikova, O.A., Zakharev, V.M., Fedorova, L.I., Zelenin, A.V., Einhorn, S., Buys, C., Laland, M., Ross, M. and Yankovsky, N.K. (1995) Mol. Biol. (in Russ.) 29, 1126–1136.
[15] Kalachikov, S., Migliazza, A., Cayanis, E., Fracciolla, N.S., Bonaldo, M.F., Lawton, L., Jelenc, P., Ye, X., Qu, X., Chien, M. and Hauptshtein, R. et al. (1997) Genomics 42, 369–377.
[16] Hawthorn, L., Roberts, T., Verlind, E., Kooy, R.F. and Cowell, J.K. (1995) Genomics 30, 425–429.
[17] Senger, G., Ragousis, J., Trewsdale, J. and Sheer, D. (1993) Cytogenet. Cell Genet. 61, 49–53.
[18] Bentley, D.R., Todd, C., Collins, J., Holland, J., Dunham, I., Hassock, S., Bankier, A. and Gianelly, F. (1992) Genomics 12, 534–541.
[19] Baxendale, S., Bates, G.P., MacDonald, M.E., Gusella, J.F. and Lehrach, H. (1991) Nucleic Acids Res. 19, 6651.

[20] Mott, R.F., Grigoriev, A.V., Maier, E., Hoheisel, J.D. and Lehrach, H. (1993) Nucleic Acids Res. 21, 1965–1974.
[21] Prapapanich, V., Chen, S., Nair, S.C., Rimerman, R.A. and Smith, D.F. (1996) Mol. Endocrinol. 10, 420–431.
[22] Quaderi, N.A., Schweiger, S., Gaudenz, K., Franco, B., Rugarli, E.I., Berger, W., Feldman, G.J., Volta, M., Angolfi, G. and Gidenkranz, S. et al. (1997) Nat. Genet. 17, 285–291.
[23] Stallings, R.L., Doggett, N.A., Okumura, K. and Ward, D.C. (1992) Genomics 13, 332–338.