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\textsuperscript{a}, Vladimir Kashuba\textsuperscript{b}, Anna Baranova\textsuperscript{a,c}, Omid Rasool\textsuperscript{d}, Wim van Everdink\textsuperscript{d}, Yie Liu\textsuperscript{d}, Alexander Syomov\textsuperscript{a}, Martin Corcoran\textsuperscript{f}, Andrey Poltaraus\textsuperscript{b}, Vadim Brodyszky\textsuperscript{a}, Natalia Syomova\textsuperscript{a}, Alexey Kazakov\textsuperscript{a}, Rachel Ibbotson\textsuperscript{f}, Anke van den Berg\textsuperscript{a,e}, Martin Corcoran\textsuperscript{f}, Rinat Gizatullin\textsuperscript{b,g}, Ludmila Fedorova\textsuperscript{a}, Galina Sulimova\textsuperscript{a}, Alexander Zelenin\textsuperscript{a}, Larry Deaven\textsuperscript{b}, Hans Lehrach\textsuperscript{d}, Dan Grandner\textsuperscript{d}, Charles Buys\textsuperscript{e}, David Oscier\textsuperscript{d}, Eugene R. Zarabovsky\textsuperscript{b,g}, Stephan Einhorn\textsuperscript{d}, Nick Yankovsky\textsuperscript{a,c,*}

\textsuperscript{a}Yavlilov Institute of General Genetics, 117809 Moscow, Russia
\textsuperscript{b}MTC and GCR Karolinska Institute, S-104 01 Stockholm, Sweden
\textsuperscript{c}Bion Ltd., 117809 Moscow, Russia
\textsuperscript{d}Research Laboratory of Radiationhemmet, Department of Oncology/Pathology, Karolinska Hospital and Institute, S-171 76 Stockholm, Sweden
\textsuperscript{e}University of Groningen, 9713 AW Groningen, The Netherlands
\textsuperscript{f}Royal Bournemouth Hospital, Bournemouth BH7 7DW, UK
\textsuperscript{g}Engelhardt Institute of Molecular Biology, 117894 Moscow, Russia
\textsuperscript{h}LANL MS M 888, Los Alamos, NM 87545, USA
\textsuperscript{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 RBL 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 (W19598) 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 cosmids 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 and seems to be an additional interesting candidate for being a B-CLL TSG. LANL and ICRF chromosome 13 specific cosmid libraries have been shown to be good sources to construct such maps [12,13].

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 patients, 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.

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 loss of this gene is associated with the childhood neural retina tumor called retinoblastoma and the bi-allelic functional loss of this gene is associated with the first example of a TSG in humans was the RB1 gene [2]. The known genes involved in tumorigenesis (RET finger protein, BRCA1) and embryogenesis (MID1). We have termed the gene corresponding to YAC ICRF 61c1, as well as fine mapping of the genomic region lost in leukocytes of 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 cosmids 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...
the clone grid and used as probes but they are presented only in the probe bar on Fig. 1. These subcloned cosmids 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].

![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.](image-url)

---

**CEN**

**STS:** D13S1168

**YAC:** /-------------->ICRF 61c1------------>------------>-------->ICRF 126a9-------->

**CDNA:**

P cc1a5as61a6a1a1c191c1a1117s8a1172111a4791s3234412a434111143b653abbb8bb8bb6566p85
R 6127071b03176333364339513734132925402032745121e9993129339312223721c2545411172442112323
O 6132 a 1 k2b 2200093h3a 7706g3a197 906ec9g0df6021aega67a3a5a668a117fge66850 8265659197
LANL 150b 96f 9 14 4 a h h h 921 b b g2g12 2 6 1 1 18 8 1 i t 16 0 036342 6 1 1 1
S 1 2 2 . 0 0 1 . 1 0 0

**T3, T7, SP6 end:** 7 37 3 3 73373737 3 7 7 7 3 7 3 7773737 3 7 7 7373737 3 7 7 73373737 7 7 7 7

**Grid:**

LANL 157c4 . . . . . .
LANL 126h4 . . . . . .
LANL 105h5 . . . . . .
LANL 50a9 . . . . . .
LANL 67b4 . . . . . .
LANL 6a11 . . . . . .
LANL 162a2 . . . . . .
LANL 162a3 . . . . . .
LANL 130b8 . . . . . .
ICRF 133h12 . . . . . .
ICRF 95h2 . . . . . .
ICRF 60a1 . . . . . .
LANL 157b12 . . . . . .
LANL 29h9 . . . . . .
ICRF 19h2 . . . . . .
ICRF 43h3 . . . . . .
LANL 143g9 . . . . . .
LANL 170c7 . . . . . .
LANL 71a11 . . . . . .
LANL 150c9 . . . . . .
LANL 176c2 . . . . . .
LANL 109d4 . . . . . .
LANL 167a6 . . . . . .
LANL 178c3 . . . . . .
LANL 127f . . . . . .
LANL 74d12 . . . . . .
LANL 47g5 . . . . . .
LANL 168c . . . . . .
LANL 3e11 . . . . . .
LANL 116d10 . . . . . .
LANL 29a6 . . . . . .
LANL 3ea5 . . . . . .
LANL 127b6 . . . . . .
LANL 43a5 . . . . . .
LANL 49g7 . . . . . .
LANL 47f10 . . . . . .
LANL 126d9 . . . . . .
LANL 120a6 . . . . . .
LANL 138d4 . . . . . .
LANL 9ga4 . . . . . .
LANL 3e28 . . . . . .
LANL 35g9 . . . . . .
LANL 126a2 . . . . . .
LANL 49c6 . . . . . .
LANL 135b8 . . . . . .
LANL 109d2 . . . . . .
LANL 154b11 . . . . . .
LANL 157a5 . . . . . .
LANL 168c5 . . . . . .
LANL 5075 . . . . . .
LANL 115h2 . . . . . .
LANL 23a12 . . . . . .
LANL 55c11 . . . . . .

---

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 meta-
phase 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) (3686
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 South-
ern or PCR analysis.

2.6. Computer program
Computer program ‘Contig’ [20] was used for cosmID contig assem-
bly. The program is available via ftp://ftp.icnet.uk/icrf-public/Ge-
nomeAnalysis/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 lib-
ary 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 ge-
nome was found. Some double locations were found on chro-
mosome 13 (+13q12, +13q21, +13q34) including a double
signal in 13q14. The revealed high level of imperfect dupli-
cations involving the region indicates the need for caution in the
interpretation of further cDNA screening results and assign-
ment 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. Ribopropes 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 ribopropes as well as
with 21 oligonucleotide probes for the STS and EST in the
region have been entered and interpreted by the ‘Contig’ com-
puter 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 till-
ping 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 approxi-
mately 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 AFMA301w5 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 be-
tween loci D13S1150 and D13S272 and nearly reaching AF-
MA301w5 on the telomeric end (see Fig. 1). The YAC cor-
responds to the centromeric half of the region associated with
B-CLL deletions. Screening of placenta and fetal brain libra-
ries (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
CCL2-5 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 de-
fect). The gene is evolutionarily conserved as can be con-
scluded 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
Sjoºgren syndrome antigen RO-SS/A. 10k4 also shares a ho-
mology 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 lympho-
cyte 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
mRNA; 13g2 did not produce a clear Northern signal; 13g3
hybridizes to different bands corresponding to 1.3-, 1.6-, 2.9-
and 3.5-kb mRNA. 13g1 corresponds to the centromeric half of the region
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 be-
tween loci D13S1150 and D13S272 and nearly reaching AF-
MA301w5 on the telomeric end (see Fig. 1). The YAC cor-
responds to the centromeric half of the region associated with
B-CLL deletions. Screening of placenta and fetal brain libra-
ries (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
CCL2-5 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 de-
fect). The gene is evolutionarily conserved as can be con-
scluded 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
Sjoºgren syndrome antigen RO-SS/A. 10k4 also shares a ho-
mology 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 lympho-
cyte 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
mRNA; 13g2 did not produce a clear Northern signal; 13g3
hybridizes to 1.7-, 2.4-, and 7.5-kb mRNA.

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.

Protein  P-value  M.R.LEEDELTCPICSLSLDFDRPVLPCSHPFCKKCL
Leu5  1.8e-20  M.TELESLTCPLCSLDFDRPVLPCSHPFCKKCL
XPRF  7.9e-17  A33  9.5e-17  LEHMKEETTCPICLSLDFDRPVLPCSHPFCKKCL
Rpt-1  1.1e-16  XPXPFCLCPICLDFDRPVLPCSHPFCKKCL
RFP  3.7e-15  A.RCLOQETCTPCICLSLDFDRPVLPCSHPFCKKCL
Staf-50  1.7e-13  KVEDEKETTCPICLSLDFDRPVLPCSHPFCKKCL
F54G8.4  3.7e-13  LEKIKIQCLCPICLDFDRPVLPCSHPFCKKCL
Ro SS/A  4.5e-11  L.RHEIESLTCPICLDFDRPVLPCSHPFCKKCL
BRCA1  9.6e-6  CLSRLHIEETTCPICLDFDRPVLPCSHPFCKKCL

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 present 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′-GCATACGCTTCTGACCC-3′, 13g1-2 5′-GAGATATCCCCCTACCC-3′, t = 55–57, 256 bp.

EST 13g2: 13g2-1 5′-AGTAAAGCTGAGCTTCG-3′, 13g2-2 5′-GGCTAAATACCTGCCCTCCG-3′, t = 55–57, 241 bp.

EST 13g3: 13g3-1 5′-TCAAGTGTCCTACATGGG-3′, 13g3-2 5′-AGGCCATCTCTTCTCTGAGC, t = 55–57, 268 bp.

EST 13g1 is located at 206fx12 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 RI 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

[20] Mott, R.F., Grigoriev, A.V., Maier, E., Hoheisel, J.D. and Leh- 
[21] Prapapanich, V., Chen, S., Nair, S.C., Rimerman, R.A. and 
[22] Quaderi, N.A., Schweiger, S., Gaudenz, K., Franco, B., Rugarli, 
E.I., Berger, W., Feldman, G.I., Volta, M., Angolfi, G. and Gi- 