Volume 16 Number 23 1988 Nucleic Acids Research

Isolation of human cDNA clones of *myb*-related genes, A-*myb* and B-*myb*

Nobuo Nomura*, Masayoshi Takahashi[1], Minami Matsui, Shunsuke Ishii[2], Takayasu Date[3], Shigemi Sasamoto and Ryotaro Ishizaki

Molecular Oncology Laboratory, Nippon Veterinary and Zootechnical College, Sakuragi, 1-10-19 Uenosakuragi, Taito-ku, Tokyo 110, [1]Institute of Applied Microbiology, University of Tokyo, Bunkyo-ku, Tokyo 113, [2]Laboratory of Physical and Chemical Research, RIKEN, 3-1-1 Koya-dai, Tsukuba, Ibaraki 305 and [3]Department of Biochemistry, Kanazawa Medical University, Uchinada, Ishikawa 920-02, Japan

Received September 15, 1988; Revised and Accepted November 8, 1988 Accession nos X13293, X13294

ABSTRACT

cDNA clones of the myb-related genes A-myb and B-myb were obtained by screening human cDNA libraries. The predicted open reading frame of B-myb could encode a protein of 700 amino acid residues. Although the C-terminal end has not been cloned yet, an almost entire coding region of A-myb, which is 745 amino acid long, was determined. The A-myb and B-myb proteins are highly homologous with the myb protein in three regions. Domain I, which is 161 amino acid long, is well conserved in the myb gene family. The homology between human-myb and A-myb in domain I is 90% at the amino acid level. Domain II, which is about 85 amino acid long, is less well conserved. Although it is a short stretch, domain III is found in the C-terminal region. The mRNAs of A-myb and B-myb were 5.0 and 2.6kb, respectively. The mRNA expression pattern of the myb gene family in various tumors is presented.

INTRODUCTION

The myb protooncogene is an evolutionarily conserved locus identified by its homology with the transforming gene v-myb of avian myeloblastosis virus (AMV) and avian erythroblastosis virus E26 (1-4). The products of the c-myb and v-myb genes appear to be nuclear DNA binding proteins (2,5,6,7). Recently the v-myb protein was shown to specifically recognize the nucleotide sequence pyAAC$^{G}/_{T}$G (8). It has been reported that c-myb mRNA is expressed predominantly in normal and tumor cells of hematopoietic origin, and that its level of expression is much higher in immature cells than in mature cells of each lineage examined (9). In vitro induction of terminal differentiation is associated with early disappearance of myb transcripts in several myeloid cell lines (9). Thus the c-myb gene product may be involved in the control of growth and/or differentiation of hematopoietic cells. AMV causes myeloblastic

or monocytic leukemia in chickens (10), whereas E26 causes erythroblastic as well as myeloblastic or monocytic leukemia in chickens (11). Structural abberations of the c-myb gene in human, murine, and chicken tumors have also been reported: (i) several murine tumors are associated with insertion of a defective Molony murine leukemia virus into a c-myb gene and with expression of an abnormal c-myb transcript (12-17); (ii) rearrangement of the c-myb was noted in chicken B-cell lymphomas (18) and human melanoma cells (19); (iii) amplification of c-myb has been reported in human myeloid leukemia (20) and colon carcinoma (21). Therefore, c-myb may play a key role in oncogenesis.

Families of nuclear oncogenes (22,23: N.Nomura et al., in preparation), of protein-tyrosine kinases (24) and of thyroid hormone receptor genes (25) have been reported. Each gene family harbors a conserved region that should encode proteins with common function. To obtain an insight into the functions of c-myb in transformation and cell growth, we looked for a gene(s) related to c-myb. Here we report the isolation and characterization of cDNA clones of the human myb-related genes A-myb and B-myb.

MATERIALS AND METHODS

Cells:

Cell lines were derived from a neuroblastoma (NB-1, TGW-III-nu, NB39-nu), Burkitt lymphoma (JBL-1, JBL-3, JBL-5), myeloid cells (KG-1), a T cell lymphoma (Molt-4), a mesothelioma (TC8), an arrhenoblastoma (TC25), a malignant fibrous histiocytoma (NMS10), and carcinomas of the stomach (MKN-1, MKN-28, MKN-45, MKN-74, KATO-III), prostate (PC3, 1013L), thyroid (TC78, TC80), lung (NMS83), colon (CL, Colo320DM), vulva (A431), breast (MCF-7), kidney (253J), and uterus (T24). MKN-1, MKN-28, MKN-45, MKN-74, KATO-III, NB-1, TGW-III-nu and NB39-nu were from T.Suzuki (Niigata University), JBL-1, JBL-3 and JBL-5 were from I.Miyoshi (Kohchi Medical College), T24, PC3, 1013L and 253J were from Y.Nakagami (Nippon Medical School), TC8, TC25, TC78 and TC80 were from S.Maeda (Nippon Medical School), and KG-1, A431 and Molt4 were

from the Japanese Cancer Research Resources Bank. Colo320DM was purchased from Dainihonseiyaku, Osaka, Japan.

cDNA library:

The human cDNA libraries used in this work were generously provided by D.P.Dialynas [a λgt10 cDNA library from mRNAs of the T cell line HPB-MLT (26)], J.E.Sadler [λgt11, endothelial cells from umbilical vein (27)], P.Chambon [λgt11, a breast cancer cell line MCF-7 (28)], J.R.de Wet [λgt11, a hepatoma cell line Li-7 (29)], S.L.C.Woo [λgt11, liver (30)], W.L.Miller [λgt10, adrenal (31)], J.M.Puck [λgt11, peripheral blood lymphocytes (32)], J.L.Millan [λgt11, testis (33)], C.Betsholtz [λgt10, a glioma cell line U-343MGa clone 2:6 (34)] and G.J.Roth [λgt11, an erythroleukemia cell line HEL (35)]. Human placenta and IMR32 (a neuroblastoma cell line) cDNA libraries which were constructed in λgt11 and λgt10 phage vector respectively, were purchased from Clontech Lab., Inc. (Palo Alto, CA, U.S.A.)

Screening of the cDNA library:

A 2.6kb EcoRI fragment of pE2.6 (36), the 1.3kb HpaII-NcoI fragment of λ-Amyb1 (this work) and 0.85kb and 1.4kb EcoRI fragments of λ-Bmyb1 (this work) were random-primed (37) with [α-^{32}P]dCTP (3000Ci/mmol) to a specific activity of $2x10^6$cpm/ng. Hybridization was performed in solution containing either 30% (relaxed condition) or 50% (stringent condition) formamide, 5xSSC, 0.5%SDS, 5xDenhardt's solution, 100μg/ml of sonicated salmon testis DNA and ^{32}P-labeled probe ($2x10^6$cpm/ml) at 37°C for 16 hours. After several washings in 1xSSC, 0.1%SDS at room temperature, filters were finally washed with 0.1xSSC, 0.5%SDS either at 35°C (relaxed condition) or 50°C (stringent condition) for 1 hour.

DNA sequencing:

Sequence analysis was carried out by the dideoxynucleotide chain terminator method with a modification (7-deaza dGTP instead of dGTP) (38,39). Relevant DNA fragments were isolated from λ phage clones by digestion with restriction endonucleases and were cloned into M13mp11, M13mp18 and pUC18 (40). Some sequencing was performed by subcloning appropriate restriction fragments into M13mp11 and M13mp18.

Computer analysis:

Homology studies and other computer analyses were carried out with the University of Wisconsin Computing Group package (41) and IDEAS (42) programs in a VAX/VMS computer (Institute of Medical Science, Tokyo University).

Northern and Southern blot analyses:

Cytoplasmic RNA from cells (43) was passed over oligo (dT)-cellulose. The glyoxylated poly(A)$^+$ RNA (3μg) was fractionated on 0.7% agarose gel and transferred to a Biodyne A filter (Pall, New York, U.S.A.) (44). Genomic DNAs were digested with BamHI (Takara Shuzo, Kyoto, Japan), electrophoresed in 0.7% agarose gel, treated, and blotted onto a Biodyne A filter essentially as described by Soutern (45).

RESULTS

Isolation of cDNA clones of myb and myb-related genes

To obtain cDNA clones of the human-myb gene, we screened six kinds of cDNA libraries with a 2.6kb EcoRI fragment of the h-myb genomic clone pE2.6 (36) as a probe. On screening $3x10^5$ phages of each library, 17 and 7 positive clones, respectively, were isolated from the T cell (HPB-MLT) (26) and the breast cancer cell (MCF-7) cDNA libraries (28). No positive clones were obtained from the other four libraries, namely placenta, liver, umbilical vein and Li-7 cDNA libraries. Several clones from the T cell library were characterized further by physical mapping and partial sequence analysis, and results were consistent with reported findings (46,47). As h-myb, D-myb [the myb gene of Drosophila melanogaster (2,48)], and maize myb (49) have large homologous domains in the N-terminal portion (Fig.4), we thought that this region might be conserved in a gene(s) related to myb. Therefore, we excised the 570bp EcoRI fragment of h-myb from a cDNA clone and used it as a probe [Fig.1 (a)]. Two positive clones were isolated by screening six kinds of cDNA libraries under conditions of reduced stringency. Hybridization with a h-myb probe and partial sequence analysis revealed that these clones had myb-related genes that differed from each other. These newly identified genes related to myb were named A-myb and B-myb, respectively. Both clones were

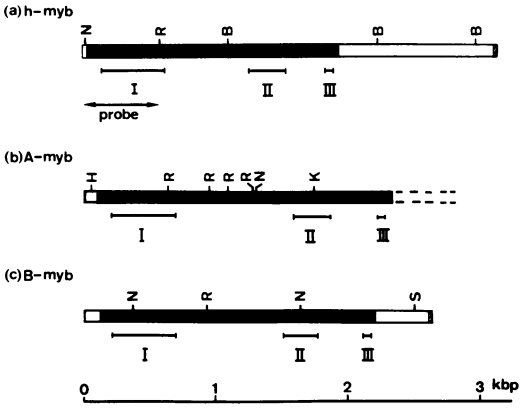
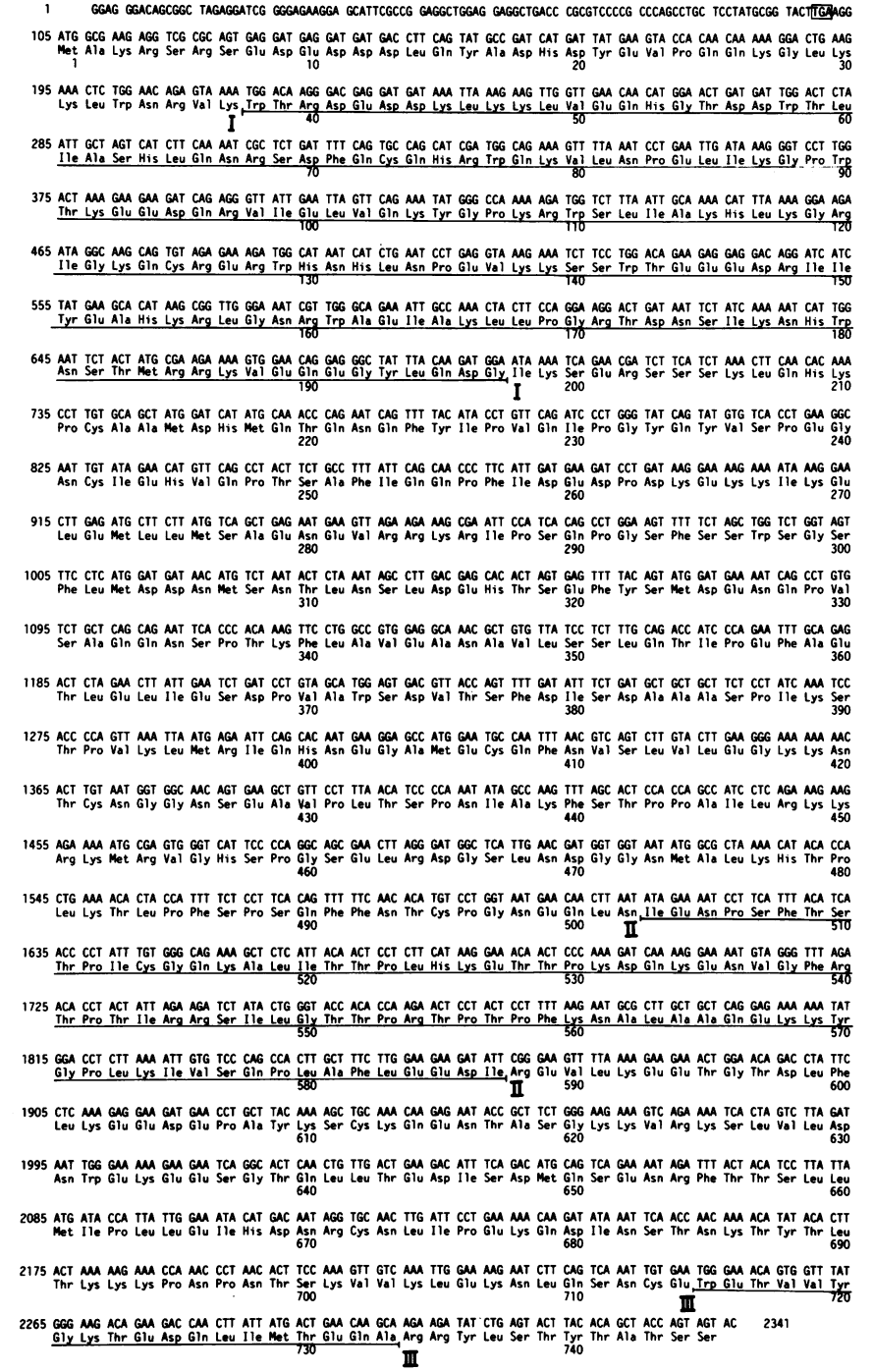


Fig.1 cDNA clones of (a) human-myb, (b) human A-myb, and (c) human B-myb
The solid and open boxes represent the coding and non-coding regions, respectively. The hatched box shows a polyA tail. Three conserved domains are indicated by Roman numerals (I,II and III). (a) The 570bp myb probe used for initial screening of myb-related genes is indicated by an underline. (b) Dashed lines indicate a 3' non coding region and a C-terminal portion which have not been cloned yet. Abbreviations: R,EcoRI; B,BamHI; H, HpaII; N,NcoI; K,KpnI; S,StuI. HpaII sites of (a) h-myb and (c) B-myb are not shown.

isolated from the T cell (HPB-MLT) library (26) and were named λ-Amyb1 and λ-Bmyb1, respectively.

DNA sequence of the human-A-myb cDNA clone

About 1450bp nucleotide sequence of the insert of the λ-Amyb1 phage was determined (Fig.2). An open reading frame starting with the first ATG codon at position 105 was identified, but the C-terminal region was deleted in this cDNA clone. To obtain other A-myb clones, we screened twelve cDNA libraries with the 1273bp HpaII-NcoI fragment of λ-Amyb1 as a probe under stringent conditions. Six clones were obtained from the MCF-7 cDNA library, and four clones were isolated from the testis cDNA library. In addition, single clones were isolated from the hepatoma cell Li-7, the liver and the peripheral blood cDNA libraries, respectively. Southern hybridization and DNA sequencing analysis showed that a clone isolated from the testis cDNA library, named λ-Amyb2, carried the C-terminal region of A-myb. The 2341bp nucleotide sequence of A-myb cDNA,


1 GGAG GGACAGCGGC TAGAGGATCG GGGAGAAGGA GCATTCGCCG GAGGCTGGAG GAGGCTGACC CGCGTCCCCG CCCAGCCTGC TCCTATGCGG TACTTGAAGG
105 ATG GCG AAG AGG TCG CGC AGT GAG GAT GAG GAT GAT GAC CTT CAG TAT GCC GAT CAT GAT TAT GAA GTA CCA CAA CAA AAA GGA CTG AAG
Met Ala Lys Arg Ser Arg Ser Glu Asp Glu Asp Asp Asp Leu Gln Tyr Ala Asp His Asp Tyr Glu Val Pro Gln Gln Lys Gly Leu Lys
1 10 20 30
195 AAA CTC TGG AAC AGA GTA AAA TGG ACA AGG GAC GAG GAT GAT AAA TTA AAG AAG TTG GTT GAA CAA CAT GGA ACT GAT GAT TGG ACT CTA
Lys Leu Trp Asn Arg Val Lys Trp Thr Arg Asp Glu Asp Asp Lys Leu Lys Lys Leu Val Glu Gln His Gly Thr Asp Asp Trp Thr Leu
I 40 50 60
285 ATT GCT AGT CAT CTT CAA AAT CGC TCT GAT TTT CAG TGC CAG CAT CGA TGG CAG AAA GTT TTA AAT CCT GAA TTG ATA AAG GGT CCT TGG
Ile Ala Ser His Leu Gln Asn Arg Ser Asp Phe Gln Cys Gln His Arg Trp Gln Lys Val Leu Asn Pro Glu Leu Ile Lys Gly Pro Trp
70 80 90
375 ACT AAA GAA GAA GAT CAG AGG GTT ATT GAA TTA GTT CAG AAA TAT GGG CCA AAA AGA TGG TCT TTA ATT GCA AAA CAT TTA AAA GGA AGA
Thr Lys Glu Glu Asp Gln Arg Val Ile Glu Leu Val Gln Lys Tyr Gly Pro Lys Arg Trp Ser Leu Ile Ala Lys His Leu Lys Gly Arg
100 110 120
465 ATA GGC AAG CAG TGT AGA GAA AGA TGG CAT AAT CAT CTG AAT CCT GAG GTA AAG AAA TCT TCC TGG ACA GAA GAG GAG GAC AGG ATC ATC
Ile Gly Lys Gln Cys Arg Glu Arg Trp His Asn His Leu Asn Pro Glu Val Lys Lys Ser Ser Trp Thr Glu Glu Glu Asp Arg Ile Ile
130 140 150
555 TAT GAA GCA CAT AAG CGG TTG GGA AAT CGT TGG GCA GAA ATT GCC AAA CTA CTT CCA GGA AGG ACT GAT AAT TCT ATC AAA AAT CAT TGG
Tyr Glu Ala His Lys Arg Leu Gly Asn Arg Trp Ala Glu Ile Ala Lys Leu Leu Pro Gly Arg Thr Asp Asn Ser Ile Lys Asn His Trp
160 170 180
645 AAT TCT ACT ATG CGA AGA AAA GTG GAA CAG GAG GGC TAT TTA CAA GAT GGA ATA AAA TCA GAA CGA TCT TCA TCT AAA CTT CAA CAC AAA
Asn Ser Thr Met Arg Arg Lys Val Glu Gln Glu Gly Tyr Leu Gln Asp Gly Ile Lys Ser Glu Arg Ser Ser Ser Lys Leu Gln His Lys
190 I 200 210
735 CCT TGT GCA GCT ATG GAT CAT ATG CAA ACC CAG AAT CAG TTT TAC ATA CCT GTT CAG ATC CCT GGG TAT CAG TAT GTG TCA CCT GAA GGC
Pro Cys Ala Ala Met Asp His Met Gln Thr Gln Asn Gln Phe Tyr Ile Pro Val Gln Ile Pro Gly Tyr Gln Tyr Val Ser Pro Glu Gly
220 230 240
825 AAT TGT ATA GAA CAT GTT CAG CCT ACT TCT GCC TTT ATT CAG CAA CCC TTC ATT GAT GAA GAT CCT GAT AAG GAA AAG AAA ATA AAG GAA
Asn Cys Ile Glu His Val Gln Pro Thr Ser Ala Phe Ile Gln Gln Pro Phe Ile Asp Glu Asp Pro Asp Lys Glu Lys Lys Ile Lys Glu
250 260 270
915 CTT GAG ATG CTT CTT ATG TCA GCT GAG AAT GAA GTT AGA AGA AAG CGA ATT CCA TCA CAG CCT GGA AGT TTT TCT AGC TGG TCT GGT AGT
Leu Glu Met Leu Leu Met Ser Ala Glu Asn Glu Val Arg Arg Lys Arg Ile Pro Ser Gln Pro Gly Ser Phe Ser Ser Trp Ser Gly Ser
280 290 300
1005 TTC CTC ATG GAT GAT AAC ATG TCT AAT ACT CTA AAT AGC CTT GAC GAG CAC ACT AGT GAG TTT TAC AGT ATG GAT GAA AAT CAG CCT GTG
Phe Leu Met Asp Asp Asn Met Ser Asn Thr Leu Asn Ser Leu Asp Glu His Thr Ser Glu Phe Tyr Ser Met Asp Glu Asn Gln Pro Val
310 320 330
1095 TCT GCT CAG CAG AAT TCA CCC ACA AAG TTC CTG GCC GTG GAG GCA AAC GCT GTG TTA TCC TCT TTG CAG ACC ATC CCA GAA TTT GCA GAG
Ser Ala Gln Gln Asn Ser Pro Thr Lys Phe Leu Ala Val Glu Ala Asn Ala Val Leu Ser Ser Leu Gln Thr Ile Pro Glu Phe Ala Glu
340 350 360
1185 ACT CTA GAA CTT ATT GAA TCT GAT CCT GTA GCA TGG AGT GAC GTT ACC AGT TTT GAT ATT TCT GAT GCT GCT TCT CCT ATC AAA TCC
Thr Leu Glu Leu Ile Glu Ser Asp Pro Val Ala Trp Ser Asp Val Thr Ser Phe Asp Ile Ser Asp Ala Ala Ala Ser Pro Ile Lys Ser
370 380 390
1275 ACC CCA GTT AAA TTA ATG AGA ATT CAG CAC AAT GAA GGA GCC ATG GAA TGC CAA TTT AAC GTC AGT CTT GTA CTT GAA GGG AAA AAA AAC
Thr Pro Val Lys Leu Met Arg Ile Gln His Asn Glu Gly Ala Met Glu Cys Gln Phe Asn Val Ser Leu Val Leu Glu Gly Lys Lys Asn
400 410 420
1365 ACT TGT AAT GGT GGC AAC AGT GAA GCT GTT CCT TTA ACA TCC CCA AAT ATA GCC AAG TTT AGC ACT CCA CCA GCC ATC CTC AGA AAG AAG
Thr Cys Asn Gly Gly Asn Ser Glu Ala Val Pro Leu Thr Ser Pro Asn Ile Ala Lys Phe Ser Thr Pro Pro Ala Ile Leu Arg Lys Lys
430 440 450
1455 AGA AAA ATG CGA GTG GGT CAT TCC CCA GGC AGC GAA CTT AGG GAT GGC TCA TTG AAC GAT GGT GGT AAT ATG GCG CTA AAA CAT ACA CCA
Arg Lys Met Arg Val Gly His Ser Pro Gly Ser Glu Leu Arg Asp Gly Ser Leu Asn Asp Gly Gly Asn Met Ala Leu Lys His Thr Pro
460 470 480
1545 CTG AAA ACA CTA CCA TTT TCT CCT TCA CAG TTT TTC AAC ACA TGT CCT GGT AAT GAA CAA CTT AAT ATA GAA AAT CCT TCA TTT ACA TCA
Leu Lys Thr Leu Pro Phe Ser Pro Ser Gln Phe Phe Asn Thr Cys Pro Gly Asn Glu Gln Leu Asn Ile Glu Asn Pro Ser Phe Thr Ser
490 II 500 510
1635 ACC CCT ATT TGT GGG CAG AAA GCT CTC ATT ACA ACT CCT CTT CAT AAG GAA ACA ACT CCC AAA GAT CAA AAG GAA AAT GTA GGG TTT AGA
Thr Pro Ile Cys Gly Gln Lys Ala Leu Ile Thr Thr Pro Leu His Lys Glu Thr Thr Pro Lys Asp Gln Lys Glu Asn Val Gly Phe Arg
520 530 540
1725 ACA CCT ACT ATT AGA AGA TCT ATA CTG GGT ACC ACA CCA AGA ACT CCT ACT CCT TTT AAG AAT GCG CTT GCT GCT CAG GAG AAA AAA TAT
Thr Pro Thr Ile Arg Arg Ser Ile Leu Gly Thr Thr Pro Arg Thr Pro Thr Pro Phe Lys Asn Ala Leu Ala Ala Gln Glu Lys Lys Tyr
550 560 570
1815 GGA CCT CTT AAA ATT GTG TCC CAG CCA CTT GCT TTC TTG GAA GAA GAT ATT CGG GAA GTT TTA AAA GAA GAA ACT GGA ACA GAC CTA TTC
Gly Pro Leu Lys Ile Val Ser Gln Pro Leu Ala Phe Leu Glu Glu Asp Ile Arg Glu Val Leu Lys Glu Glu Thr Gly Thr Asp Leu Phe
580 II 590 600
1905 CTC AAA GAG GAA GAT GAA CCT GCT TAC AAA AGC TGC AAA CAA GAG AAT ACC GCT TCT GGG AAG AAA GTC AGA AAA TCA CTA GTC TTA GAT
Leu Lys Glu Glu Asp Glu Pro Ala Tyr Lys Ser Cys Lys Gln Glu Asn Thr Ala Ser Gly Lys Lys Val Arg Lys Ser Leu Val Leu Asp
610 620 630
1995 AAT TGG GAA AAA GAA GAA TCA GGC ACT CAA CTG TTG ACT GAA GAC ATT TCA GAC ATG CAG TCA GAA AAT AGA TTT ACT ACA TCC TTA TTA
Asn Trp Glu Lys Glu Glu Ser Gly Thr Gln Leu Leu Thr Glu Asp Ile Ser Asp Met Gln Ser Glu Asn Arg Phe Thr Thr Ser Leu Leu
640 650 660
2085 ATG ATA CCA TTA TTG GAA ATA CAT GAC AAT AGG TGC AAC TTG ATT CCT GAA AAA CAA GAT ATA AAT TCA ACC AAC AAA ACA TAT ACA CTT
Met Ile Pro Leu Leu Glu Ile His Asp Asn Arg Cys Asn Leu Ile Pro Glu Lys Gln Asp Ile Asn Ser Thr Asn Lys Thr Tyr Thr Leu
670 680 690
2175 ACT AAA AAG AAA CCA AAC CCT AAC ACT TCC AAA GTT GTC AAA TTG GAA AAG AAT CTT CAG TCA AAT TGT GAA TGG GAA ACA GTG GTT TAT
Thr Lys Lys Lys Pro Asn Pro Asn Thr Ser Lys Val Val Lys Leu Glu Lys Asn Leu Gln Ser Asn Cys Glu Trp Glu Thr Val Val Tyr
700 710 III 720
2265 GGG AAG ACA GAA GAC CAA CTT ATT ATG ACT GAA CAA GCA AGA AGA TAT CTG AGT ACT TAC ACA GCT ACC AGT AGT AC 2341
Gly Lys Thr Glu Asp Gln Leu Ile Met Thr Glu Gln Ala Arg Arg Tyr Leu Ser Thr Tyr Thr Ala Thr Ser Ser
730 III 740

which encodes 745 amino acid residues, was shown in Fig.2. Although no termination codon was observed, we speculate that the open reading frame of A-myb might end within 20 additional codons (see DISCUSSION).

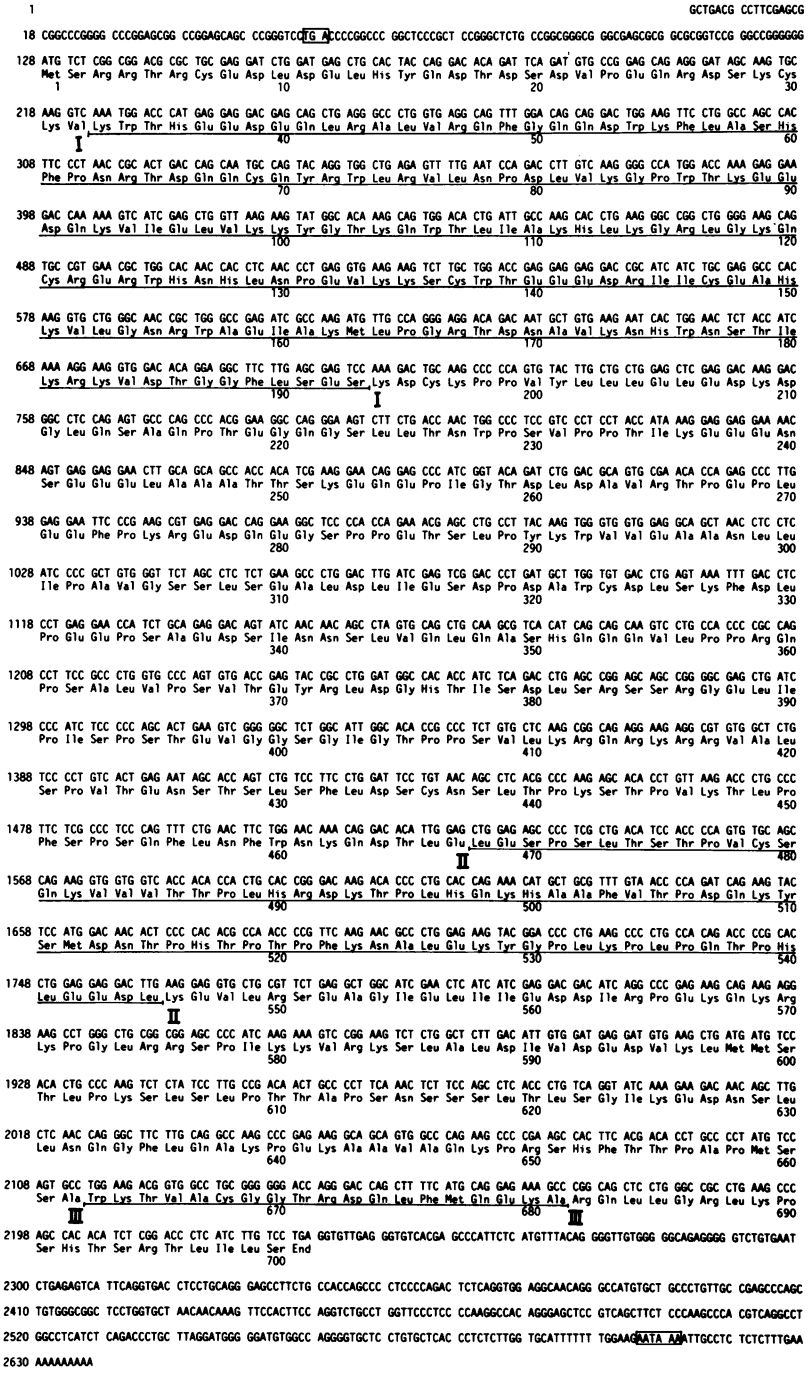
DNA sequence of the human-B-myb cDNA clones

To obtain other B-myb clones, we screened eight cDNA libraries further with the 0.85kb and 1.4kb EcoRI fragments of λ-Bmyb1 as probes under stringent conditions. Another ten clones were obtained from the T cell cDNA library, and 19 clones were isolated from the peripheral blood lymphocyte cDNA library. In addition, two clones were obtained from the liver cDNA library, four from the umbilical vein library and one each from the placenta and Li-7 cell cDNA libraries. The sequence of human-B-myb was determined from the DNA sequences of several representative clones (Fig.3). An open reading frame of 2100bp starting with the first ATG codon at position 128 was identified. The flanking nucleotides do not match the consensus sequence of Kozak (50), but an in-frame termination codon was found 70bp upstream of this ATG codon. The predicted open reading frame could encode a protein of 700 amino acid residues, with a calculated molecular weight of 78,791. Two polyadenylation sites were found, although these were two base pairs, apart. The insert of one clone ended at position 2627, followed by a polyA tail. The inserts of the other two clones ended at position 2625.

Homology among h-myb, A-myb, B-myb and D-myb

Fig.4 shows the alignment of homologous domains of the h-myb, A-myb, B-myb and D-myb proteins. Domain I, which is 161 amino acid long, is well conserved in the myb gene family. For example, the homology of this domain in h-myb and A-myb is 90% at the amino acid level. Moreover 8 of 17 amino acid changes

Fig.2 Human A-myb cDNA nucleotide sequence and the deduced amino acid sequence
Although the C-terminal end is not known, the sequence of 745 amino acids is shown. Three conserved domains (I,II and III) are underlined. The in-frame stop codon, which is 4 bp upstream of the putative initiation codon, is boxed. The insert of the λ-Amyb1 phage was terminated and polyadenylated at nucleotide position 1448.



are conservative ones. Domain II, which is about 85 amino acid long, is less well conserved. Domain III is seen in the C-terminal region as a short strech. Domains I and II are located in the region that is well conserved in human, mouse and chicken c-myb.

mRNAs of A-myb and B-myb

poly(A)$^+$ RNAs were prepared from various cell lines and subjected to Northern blotting as described by Thomas (44). After hybridization with the A-myb probe under stringent conditions, a band of 5.0kb was detected [Fig.5 (a)]. Hematopoietic cell lines, including Burkitt lymphoma (JBL-1, -3, -5) and T cell lymphoma (Molt-4) expressed A-myb mRNA at high levels. Carcinoma cell lines from the kidney (253J), uterus (T24) and colon (Colo320DM) and a sarcoma cell line (NMS10) also expressed the A-myb message at high levels, but the other cell lines gave either a weak or no visible band of A-myb mRNA. As the transcript of A-myb is 5.0kb, the A-myb cDNA sequence determined, which is 2341bp in length, should have a deletion of a large portion of either 5' or 3' non translated region. On hybridization with the 0.85kb and 1.4kb EcoRI fragments of λ-Bmyb1, B-myb mRNA of 2.6kb was detected [Fig.5 (b)]. Augmented expression of B-myb was observed in JBL-1, -3, -5, Molt-4, Colo320DM, A431, NB39-nu, KG-1 and MKN-1, but B-myb mRNAs were detected in all the cells examined. As the B-myb mRNA is 2.6kb, the B-myb cDNA sequence determined, which is 2638bp in length, should represent the entire B-myb mRNA. Enhanced c-myb expression was detected in JBL-1, -3, -5, KG-1 and Molt-4 [Fig.5 (c)]. Messages of c-myb were weakly expressed in Colo320DM, CL, TGW-III-nu, A431, NB39-nu, MKN-74, KATO III, PC3 and 1013L. Although strong c-myb expression was detected in hematopoietic cells, the A-myb and B-myb gene seem to be more broadly expressed than c-myb.

Fig.3 Nucleotide sequence of human B-myb cDNA and deduced amino acid sequence
Three conserved domains (I,II and III) are underlined. The putative polyadenylation signal (AATAAA) is boxed. The in-frame stop codon, which is 70bp upstream of the putative initiation codon, is also boxed.

```
(I)   h-myb   42 RWTREETEKLKKLVEQNGTDDWKVIANYLPNRTDVQCQHRWQKVLNPELIKGPWTKEEDQRVIELVQKYGPKRWSVIAKHL 122
      A-myb   37 K...D..D........H.....TL..SH.Q..S.F.....................................L..... 117
      B-myb   33 K..H....Q.RA..R.F.QQ...FL.SHF.....Q...Y..LR....D.V..........K.....K...T.Q.TL..... 113
      D-myb   87 ..SKS..VL..Q...TH.-EN.EI.GPHFKD.LEQ.V.Q..A.............RD..DM..K..RNF...K.TL..RY. 166

      h-myb 123 KGRIGKQCRERWHNHLNPEVKKTSWTEEEDRIIYQAHKRLGNRWAEIAKLLPGRTDNAIKNHWNSTMRRKVEQEGYLQES  202
      A-myb 118 .....................S..........E.....................S.....................DG  197
      B-myb 114 ...L.R..............SC.........CE...V..........M........V.......IK...DTG.F.S..  193
      D-myb 167 N.................NI...A...K..E......LE...Q..K...R...................YDV.RRSVNA  246


(II)  h-myb 414 LEMPSLTS-TPL-IGHKLTVTTPFHR-DQTVKTQKENTVFRTPAIKRS 458
      A-myb 503 I.N..F..-..I-C.Q.ALI...L.K-ET.P.D....VG....T.R.. 547
      B-myb 468 ..S.....-..V-CSQ.VV....L..-.K.PLH..-HAA.V..DQ.Y. 511
      D-myb 394 .KS.C..TFEDMDLRASTP..KVYN.VGMEI.KEM.TSSIE..H-.-. 539

      h-myb 459 ILESSPRTPTPFKHALAAQEIKYGPLKMLPQTPSHLVEDL 498
      A-myb 548 ..GTT.........N......K......IVS.PLAF.E..I 587
      B-myb 512 -MDNT.H.......N..---.-.......P.....-..E... 545
      D-myb 540 Q.-G-.........K....IGK.RDGRRYE.SS..S..... 477


(III) h-myb 608 WEPASCGKMEEQMTSSSQA 626
      A-myb 715 ..TVVY..T.D.LIMTE.. 733
      B-myb 663 .KTVA..GTRD.LFMQEK. 681
      D-myb 607 .ARVA...SRD..FMEE.. 625
```

Fig.4 Comparison of the amino acid sequences of human-myb, human-A-myb, human-B-myb, and Drosophila-myb
Three conserved domains are shown. (I)-(III) Amino acids identical to those of human-myb are indicated by dots. Dashes indicate gaps introduced for alignment.

Evolutionary conservation of A-myb and B-myb

Southern blots of BamHI-cleaved mammalian and avian genomic DNAs were hybridized with the human A-myb or the human B-myb probe. As shown in Fig.6, A-myb and B-myb are evolutionarily conserved. Possible amplifications and/or rearrangements of A-myb and B-myb were examined in 140 human tumors, inluding 60 cell lines and 80 fresh tumors, but no structural aberrations were observed (data not shown).

DISCUSSION

We screened twelve kinds of cDNA libraries and isolated cDNAs of the human A-myb and B-myb genes. Three domains conserved in h-myb, A-myb, B-myb and D-myb were noted. Domain I, which is 161 amino acid long, is well conserved in the myb gene family. This region also shows potentially important features: (i) three tandemly organized direct repeats of 51-52 amino acids; (ii) 98%-99% homology among human, mouse and chicken c-myb at the amino acid level, and (iii) relation of this gene to myb in maize (49). As the c-myb protein is localized in the nucleus, the gene products of A-myb and B-myb can also function in the nucleus. The highly conserved domain I

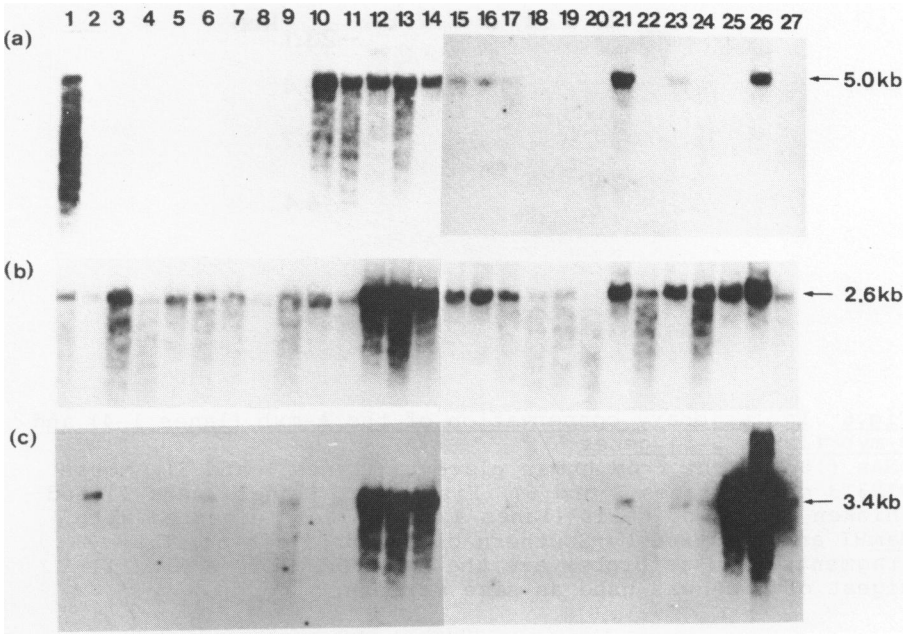


Fig.5 mRNAs of (a) A-myb, (b) B-myb and (c) c-myb
poly(A)$^+$ RNA was analyzed as described in Materials and Methods. The fragment used as A-myb probe is the 1.2kb HpaII-NcoI fragment of λ-Amyb1 (nucleotides 43-1317 in Fig.2). The B-myb probe used is the insert of λ-Bmyb1 (nucleotides 98-2366 in Fig.3). The fragments which include entire coding sequences of c-myb were used as a probe of c-myb. Lanes: 1,T24; 2,CL; 3,MKN-1; 4, MKN-28; 5,MKN-45; 6,MKN-74; 7,KATO-III; 8,PC3; 9,1013L; 10,253J; 11,NMS10; 12,JBL-1; 13,JBL-3; 14,JBL-5; 15,NMS83; 16,TC78; 17, NB-1; 18,TC25; 19,TC8; 20,TC80; 21,Colo320DM; 22,TGW-III-nu; 23,A431; 24,NB39-nu; 25,KG-1; 26,Molt-4; 27,MCF-7.

may be responsible for common functions of the myb gene family, such as binding to a specific region(s) of DNA (8) and/or interaction with a specific protein(s), as in the case of fos (51). The v-myb genes of AMV and E26 virus do not have part of the first stretch of three tandem repeats. Therefore, two of the three repeated sequences may be enough for the function of domain I. As the C-terminal region including domains II and III is deleted in the v-myb gene, these domains may play a regulatory role in the function of the myb gene family. Retroviral integration in leukemia cell lines WEHI-265,

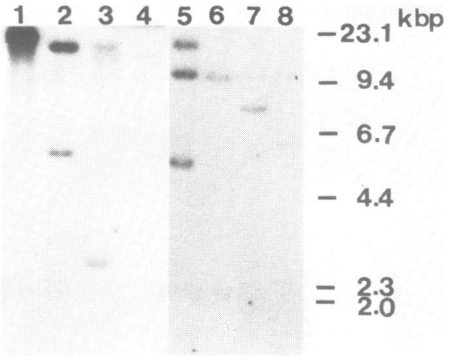


Fig.6 Evolutionary conservation of the A-myb (lanes 1-4) and B-myb (lanes 5-8) genes
DNAs (10μg each) from human placenta (lanes 1 and 5), mouse NIH3T3 cells (lanes 2 and 6), Fisher rat (lanes 3 and 7) and chicken red blood cells (lanes 4 and 8) were digested with BamHI and analyzed by Southern blot hybridization. The fragments used as probes are the same as Fig.5. A HindIII digest of λDNA was used as size markers.

WEHI-274 and NSF-60 was observed at the positions corresponding to amino acid residues 48, 72 and 396, respectively, within the mouse c-myb gene (12,14,15,16,17). Therefore, either truncation of the N-terminal region or deletion of the C-terminal region including domains II and III might activate or deregulate the myb gene, playing a key role in neoplastic transformation. The numbers of codons downstream of domain III of h-myb, B-myb and D-myb are 14, 19 and 32, respectively, wheares 12 codons were observed beyond the boundary of A-myb domain III (Fig.2). Therefore we speculate that the missing codons of the A-myb C-terminal end might be no more than 20 codons. In preliminary studies, A-myb was mapped in chromosome 8 of the human genome (M.Yoshida and N.Nomura, unpublished).

The mRNA expression pattern of the myb gene family in 27 cell lines was examined. Lymphoid cell lines, including Burkitt lymphoma (JBL-1, -3, -5) and T cell lymphoma (Molt-4) express high levels of A-myb, B-myb and c-myb mRNAs. This suggests that concomitant expression of the myb gene family might be relevant to genesis and/or progression of T cell and B cell lymphoma. Analyses of possible changes in A-myb and B-myb expression

associated with changes in differentiation states of hematopoietic cells are in progress.

ACKNOWLEDGMENTS

We thank T.Kunieda for DNAs of mouse NIH3T3 cells and Fisher rats, M.Ide for excellent technical assistance and T.Yamamoto and H.Mitsui for helpful discussions. We also thank D.P.Dialynas, P.Chambon, J.E.Sadler, D.R.Helinski, S.L.C.Woo, T.Chandra, W.L.Miller, J.M.Puck, J.L.Millan, C.Betsholtz and G.J.Roth for providing cDNA libraries. We are grateful to N.Miyajima, A.Ito, M.Yamamoto for help in computer analysis, and T.Suzuki, I.Miyoshi, Y.Nakagami and S.Maeda and the Japanese Cancer Research Resources Bank for cell lines. This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan.

*To whom correspondence should be addressed

REFERENCES

1. Bergmann, D.G., Souza, L.M., and Baluda, M.A. (1981) J.Virology 40, 450-455.
2. Peters, C.W.B., Sippel, A.E., Vingron, M. and Klempnauer, K.-H. (1987) EMBO J 6, 3085-3090.
3. Roussel, M., Saule, S., Lagrou, C., Rommens, C., Beug, H., Graf, T. and Stehelin, D. (1979) Nature (London) 281, 452-455.
4. Souza, L.M., Strommer, J.N., Hillyard, R.L., Komaromy, M.C. and Baluda. M.A. (1980) Proc.Natl.Acad.Sci. USA 77, 5177-5181.
5. Boyle, W.S., Lampert, M.A., Lipsick, J.S. and Baluda, M.A. (1984) Proc.Natl.Acad.Sci. USA 81, 4265-4269.
6. Klempnauer, K.-H., Ramsay, G., Bishop, J.M., Moscovici, M.G. , Moscovici, C., McGrath, J.P. and Levinson, A.D. (1983) Cell 33, 345-355.
7. Klempnauer, K.-H., Symonds, G., Evan, G.I. and Bishop, J.M. (1984) Cell 37, 537-547.
8. Biedenkapp, H., Borgmeyer, U., Sippel, A.E. and Klempnauer, K.-H. (1988) Nature (London) 335, 835-837.
9. Westin,E.H., Gallo, R.C., Arya, S.K., Eva, A., Souza, L.M., Baluda, M.A., Aaronson, S.A. and Wong-Staal, F. (1982) Proc. Natl.Acad.Sci. USA 79, 2194-2198.
10. Moscovici, C. (1975) Curr. Top. Microbiol. Immunol. 171, 79 -101.
11. Radke, K., Beug, H., Kornfield, S., and Graf, T. (1982) Cell 31, 643-653.
12. Gonda, T.J., Cory, S., Sobieszczuk, P., Holtzman, D., and Adams, J.M. (1987) J.Virology 61, 2754-2763.
13. Mushinski, J.F., Potter, M., Bauer, S.R. and Reddy, E.P. (1983) Science 220, 795-798.
14. Rosson, D., Dugan, D. and Reddy, E.P. (1987) Proc.Natl.Acad. Sci. USA 84, 3171-3175.

15. Shen-Ong, G.L.C., Morse III, H.C., Potter, M. and Mushinski, S. (1986) Mol.Cell.Biol. 6, 380-392.
16. Weinstein, Y., Cleveland, J.L., Askew, D.S., Rapp, U.R. and Ihle,J.N. (1987) J.Virology 61, 2339-2343.
17. Weinstein, Y., Ihle, J.N., Lavu, S. and Reddy, E.P. (1986) Proc.Natl.Acad.Sci. USA 83, 5010-5014.
18. Kanter, M.R., Smith, R.E., and Hayward, W.S., (1988) J.Virology 62, 1423-1432.
19. Linnenbach, A.J., Huebner, K., Reddy, E.P., Herlyn, M. Parmiter, A.H., Nowell, P.C. and Koprowski, H. (1988) Proc. Natl.Acad.Sci. USA 85, 74-78.
20. Pelicci, P.G., Lanfrancone, L., Brathwaite, M.D., Wolman, S.R., and Dalla-Favera, R. (1984) Science 224, 1117-1121.
21. Alitalo, K., Winqvist, R., Lin, C.C., De La Chapelle, A., Schwab, M., and Bishop, J.M. (1984) Proc.Natl.Acad.Sci. USA 81, 4534-4538.
22. Ryder, K., Lau, L.F. and Nathans, D. (1988) Proc.Natl.Acad. Sci. USA 85, 1487-1491.
23. Kaye, F., Battey, J., Nau, M., Brooks, B., Seifter, E., De Greve, J., Birrer, M., Sausville, E. and Minna. J. (1988) Mol.Cell.Biol. 8, 186-195.
24. Hunter, T. (1987) Cell 50, 823-829.
25. Giguere, V., Yang, N., Segui, P. and Evans, R.M. (1988) Nature (London) 331, 91-94.
26. Dialynas, D.P., Murre, C., Quertermous, T., Boss, J.M., Leiden, J.M., Seidman, J.G. and Strominger,J.L. (1986) Proc. Natl.Acad.Sci.USA. 83, 2619-2623.
27. Sadler, J.E., Shelton-Inloes, B.B., Sorace, J.M., Harlan, J.M., Titani, K. and Davie, E.W. (1985) Proc. Natl. Acad. Sci. USA. 82, 6394-6398.
28. Walter, P, Green, S., Greene, G., Krust, A., Bornert, J.-M. Jeltsch, J.-M., Staub, A., Jensen, E., Scrace, G., Waterfield, M. and Chambon, P. (1985) Proc.Natl.Acad.Sci. USA. 82, 7889-7893.
29. de Wet, J.R., Fukushima, H., Dewji, N.N., Wilcox, E., O'Brien, J.S. and Helinski, D.R. (1984) DNA 3, 437-447.
30. Kwok, S.C.M., Ledley, F.D., DiLella, A.G., Robson, K.J.H. and Woo, S.L.C. (1985) Biochemistry 24, 556-561.
31. Chung, B.-C., Matteson, K.J., Voutilainen, R., Mohandas, T.K. and Miller, W.L. (1986) Proc.Natl.Acad. Sci.USA. 83, 8962-8966.
32. Gold, D.P., Puck, J.M., Pettey, C.L., Cho, M., Coligan, J., Woody, J.N. and Terhorst,C. (1986) Nature (London) 321, 431 -434.
33. Millan, J.L., Driscoll, C.E., LeVan, K.,M. and Goldberg, E. (1987) Proc.Natl.Acad.Sci.USA. 84, 5311-5315.
34. Betsholtz, C., Johnsson, A., Heldin, C.-H., Westermark, B., Lind, P., Urdea, M.S., Eddy, R., Shows, T.B., Philpott, K., Mellor, A.L., Knott, T.J. and Scott, J. (1986) Nature (London) 320, 695-699.
35. Lopez, T.A., Chung, D.W., Fujikawa, K., Hagen, F.S., Papayannopoulou, T. and Roth, G.J. (1987) Proc.Natl.Acad. Sci.USA 84, 5615-5619.
36. Franchini, G., Wong-Staal, F., Baluda, M.A., Lengel, C. and Tronick, S.R. (1983) Proc.Natl.Acad.Sci. USA 80, 7385-7389.
37. Feinberg, A.P. and Vogelstein, B. (1983) Analytical Biochemistry 132, 6-13.

38. Mizusawa, S., Nishimura, S. and Seela, F. (1986) Nucleic Acids Res. 14, 1319-1324.
39. Sanger, F., Nicklen, S. and Coulson, A.T. (1977) Proc.Natl. Acad.Sci.USA 74, 5463-5467.
40. Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Gene 33, 103-119.
41. Devereux, J., Haeberli, P. and Smithies, O. (1984) Nucleic Acids Res. 12, 387-395.
42. Kanehisa, M. (1984) Nucleic Acids Res. 12, 203-213.
43. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning (Cold Spring Harbor Laboratory),191-193.
44. Thomas, P.S. (1983) In Wu,R., Grossman, L. and Moldave, K. (eds), Methods in Enzymology, Academic Press, New York, Vol 100, pp.255-266.
45. Southern, E. (1975) J.Mol.Biol. 98, 503-517.
46. Majello, B., Kenyon, L.C. and Dalla-Favera, R. (1986) Proc. Natl.Acad.Sci. USA. 83, 9636-9640.
47. Slamon, D.J., Boone, T.C., Murdock, D.C., Keith, D.E., Press, M.F., Larson, R.A. and Souza, L.N. (1986) Science 233, 347-351.
48. Katzen, A.L., Kornberg, T.B. and Bishop, J.M. (1985) Cell 41, 449-456.
49. Javier, P.-A., Ghosal, D., Wienand, U., Peterson, P.A. and Saedler, H. (1987) EMBO J. 6, 3553-3558.
50. Kozak, M. (1984) Nucleic Acids Res 12, 857-872.
51. Franza, B.R.Jr., Rauscher III, F.J., Josephs, S.F. and Curran, T. (1988) Science 239, 1150-1153.
52. Angel, P., Allehretto, E.A., Okino, S.T., Hattori, K., Boyle, W.J., Hunter, T. and Karin, M. (1988) Nature (London) 332, 166-171.
53. Bohmann, D., Bos, T.J., Admon, A., Nishimura, T., Vogt, P.K. and Tjian, R. (1987) Science 238, 1386-1392.