MOLECULAR AND CELLULAR BIOLOGY, Feb. 1995, p. 704–710 Vol. 15, No. 2
0270-7306/95/$04.00+0

Expression Cloning of Oncogenes by Retroviral Transfer of cDNA Libraries

IAN WHITEHEAD, HEATHER KIRK, AND ROBERT KAY*

Department of Medical Genetics, University of British Columbia, and Terry Fox Laboratory, British Columbia Cancer Agency, Vancouver, British Columbia, Canada V5Z 4E6

Received 23 June 1994/Returned for modification 1 September 1994/Accepted 27 October 1994

A cDNA library transfer system based on retroviral vectors has been developed for expression cloning in mammalian cells. The use of retroviral vectors results in stable cDNA transfer efficiencies which are at least 100-fold higher than those achieved by transfection and therefore enables the transfer and functional screening of very large libraries. In our initial application of retroviral transfer of cDNA libraries, we have selected for cDNAs which induce oncogenic transformation of NIH 3T3 fibroblasts, as measured by loss of contact inhibition of proliferation. Nineteen different transforming cDNAs were isolated from a total of 300,000 transferred cDNA clones. Three of these cDNAs were derived from known oncogenes (*raf-1*, *lck*, and *ect2*), while nine others were derived from genes which had been cloned previously but were not known to have the ability to transform fibroblasts (β-catenin, thrombin receptor, phospholipase C-γ_2 and Spi-2 protease inhibitor genes). The Spi-2 cDNA was expressed in antisense orientation and therefore is likely to act as an inhibitor of an endogenous transformation suppressor. Seven novel cDNAs with transforming activities, including those for three new members of the CDC24 family of guanine nucleotide exchange factors, were also cloned from the retroviral cDNA libraries. Retroviral transfer of libraries should be generally useful for cloning cDNAs which confer selectable phenotypes on many different types of mammalian cells.

Cellular oncogenes or proto-oncogenes can be cloned by selecting for their ability to confer the phenotype of deregulated growth on cells in which they are expressed. This was first done by transfecting fragmented genomic DNA into cell lines such as NIH 3T3 fibroblasts which are susceptible to single-hit oncogenic transformation (13, 15, 31, 36, 40, 46). Subsequently, stable transfection of these cells with cDNA libraries in plasmid or phage expression vectors has been used as an alternative approach, to avoid the severe difficulties which have been encountered in recovering and analyzing oncogenes present in transfected genomic DNA (8, 9, 32–34). Despite the theoretical advantages of working with transfected cDNA expression libraries, only a small fraction of the many currently known oncogenes have been cloned in this way. The most significant limitation in the use of cDNA library transfer for cloning oncogenes has been the low efficiencies of cDNA transfer and expression which can be achieved by stable transfection methods. It is difficult to generate more than a few tens of thousands of transfectants in which cDNA clones are being expressed at adequate levels, but a comprehensive screening of a mammalian cDNA library demands the transfer and expression of several million clones. Therefore, the small number of oncogenes which have been cloned so far from transferred libraries is unlikely to be due to exhaustion of the pool of cDNAs with oncogenic potential but rather is a result of a failure to transfer, and thus detect, most such cDNAs.

In contrast to deliberately constructed cDNA libraries, natural populations of retroviruses have served as a prolific source of oncogenes, through their ability to incorporate and mutationally activate cDNA copies of host cell mRNAs and then to transfer them to recipient cells by infection (5). Artificial retroviral vectors with equivalent abilities to transfer and express cDNAs in mammalian cells (35) could be ideally suited to screening libraries for cDNAs conferring selectable phenotypes such as deregulated cell growth. Previous reports have indicated the potential for using retroviruses to clone cDNAs conferring selectable phenotypes, as demonstrated by the isolation of cDNAs encoding thymidine kinase (TK) via rescue of TK^- fibroblasts (39) and cDNAs encoding interleukin-3 and granulocyte-macrophage colony-stimulating factor via conversion of hemopoietic cell lines to cytokine-independent growth (42). This report describes the development of an efficient system for cDNA library transfer and expression by retroviral vectors and its application to the cloning of a large number of cDNAs which induce oncogenic transformation of murine fibroblasts.

MATERIALS AND METHODS

pCTV vectors. pCTV1 was constructed from the following elements: (i) an *Nhe*I-*Kpn*I fragment containing the U3 region and part of the R region of Moloney murine leukemia virus (MMLV), excised from the JZen retroviral vector (22); (ii) a *Kpn*I-*Xho*I fragment extending from the R region to base 1035 in the MMLV genomic sequence (GenBank accession number J02255), derived by PCR from the pBabeHygro retroviral vector (38); (iii) an *Xho*I-*Cla*I fragment containing a synthetic cloning site and an adjacent *supF* gene (43), derived by PCR from the pAX114 expression vector (24); (iv) a *Cla*I-*Bam*HI fragment from a subclone containing the myeloproliferative sarcoma virus (MPSV)-MMLV hybrid 3′ long terminal repeat (LTR) from JZen along with 110 bp of uncharacterized genomic DNA sequence downstream of the LTR; and (v) a *Bgl*II-*Spe*I fragment from pAX114, containing the β-globin and simian virus 40 (SV40) polyadenylation sites, the SV40 origin of replication, and the origin of replication from pUC8.

pCTV3 was constructed by inserting a *Cla*I-*Acc*I fragment containing a hygromycin resistance gene into the *Cla*I site of pCTV1. This gene consists of a synthetic polyomavirus enhancer and a herpes simplex virus TK promoter (49) attached to a hygromycin phosphohydrolase gene (3) with a modified translation initiation sequence.

pCTV1B and pCTV3B were derived by inserting the 380-bp *Bst*YI fragment containing the *Bst*XI stuffer fragment of pCDM8 (44) between the *Sal*I sites of pCTV1 or pCTV3.

pCTV3K was derived from pCTV3 by replacement of the *supF* gene with the P1 promoter of pBR322 and the neomycin phosphotransferase coding region of Tn*5* (1).

* Corresponding author. Mailing address: Terry Fox Laboratory, BC Cancer Agency, 601 W. 10th Ave., Vancouver, BC, Canada V5Z 4E6. Phone: (604) 877-6070. Fax: (604) 877-0712. Electronic mail address: robert@terryfox.ubc.ca.

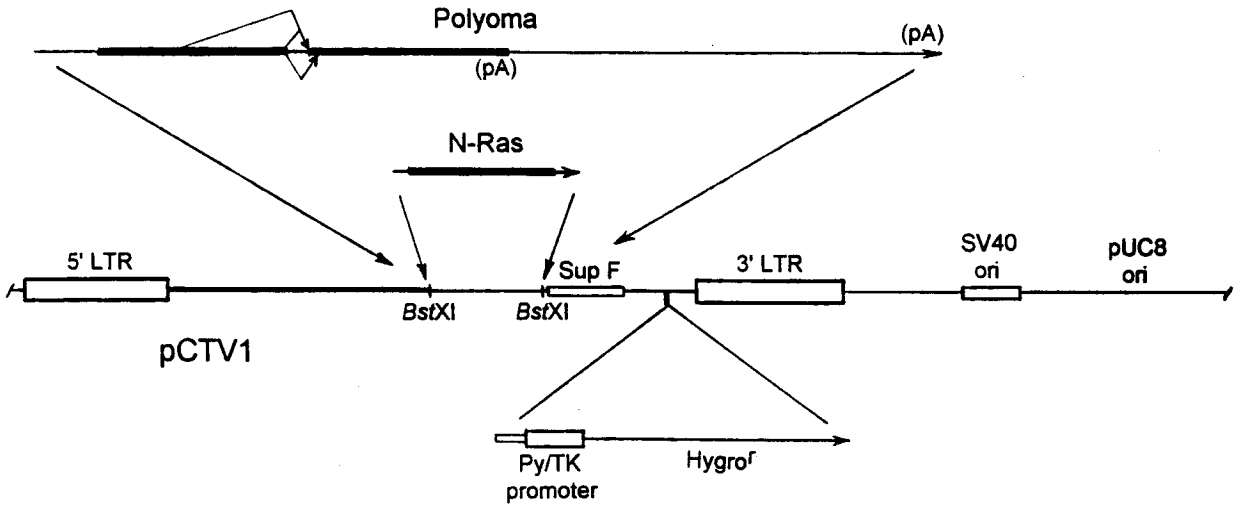

FIG. 1. Structures of retroviral vectors for cDNA library transfer. The pCTV1 vector consists of a 5′ MMLV LTR and extended *gag* region lacking the normal splice donor and initiation codon, a polylinker with a *Bst*XI stuffer which is displaced by cDNA insertion, a *supF* gene which provides a compact marker for bacterial transformation as well as a marker for cDNA recovery, a composite MPSV-MMLV 3′ LTR, and replication origins from pUC8 and SV40. The total size of pCTV1 is 4,350 bp. The pCTV3 vector is identical except for the presence of a 1,210-bp hygromycin phosphohydrolase gene driven by a polyomavirus (Py) enhancer and herpes simplex virus TK promoter. An exact description of the structure and construction of these vectors is available on request. After viral transmission, the region including and between the two LTRs becomes integrated into the recipient cell's genome as a provirus, with both 5′ and 3′ LTRs now being derived from the composite MPSV-MMLV LTR. Transcription of the provirus is initiated within the 5′ LTR, and polyadenylation occurs within the 3′ LTR. The N-Ras cDNA and polyomavirus early genes used to test for cDNA transfer and expression efficiencies are shown above the vector, with the alternative splicing and potential polyadenylation signals (pA) marked on the polyomavirus construct.

cDNA constructs. The CD24 cDNA used for quantifying cDNA transfer and expression was the 950-bp clone described previously (25). The N-Ras cDNA was amplified by PCR from a cDNA library prepared from the SK-N-SH neuroblastoma cell line (13). This cDNA has a Gln-61-to-Lys mutation which results in oncogenic activation. The polyomavirus early region construct extended from 153 to 2926 in the polyomavirus genomic sequence (GenBank accession number J02288).

Construction of cDNA libraries in pCTV vectors. Poly(A) RNA was prepared from the murine hemopoietic cell lines GM979 (7), 32D (17), and B6SUtA$_1$ (17) by guanidinium isothiocyanate lysis and oligo(dT) chromatography as described previously (12, 21). cDNA was synthesized from the poly(A) RNA by using MMLV reverse transcriptase (Life Technologies, Gaithersburg, Md.) and random sequence primers, with the procedures and reagents recommended by the supplier. 5′ phosphorylated *Bst*XI adapters (TCAGTTACTCAGG and CCTGA GTAACTGACACA) were ligated to the double-stranded cDNA, which was then size fractionated by agarose gel electrophoresis. Twenty nanograms of *Bst*XI-digested and dephosphorylated pCTV3B or pCTV1B plasmid DNA was ligated with an equimolar amount of the size-fractionated cDNA in 10 μl of 25 mM Tris-Cl (pH 7.8)–5 mM $MgCl_2$–1 mM dithiothreitol–1 mM ATP–0.5 U of T4 DNA ligase (Life Technologies); the mixture was incubated at 16°C for 3 h and then for 20 min at 72°C. The ligation reaction was used to transform *Escherichia coli* MC1061/p3 (45) by electroporation, using the procedures described in the operating manual for the GenePulser electroporation apparatus (BioRad, Richmond, Calif.). After a 90-min incubation at 37°C, the transformed bacteria were plated in soft agar, incubated at 37°C for 20 h, and then recovered from the agar as described previously (50). Plasmid DNA was prepared from the pooled bacterial colonies by the alkaline lysis procedure (4), digested with RNase A and RNase T_1, and precipitated with ammonium acetate and ethanol.

Transfection of packaging cells and infection of NIH 3T3 cells. Six-centimeter-diameter culture dishes containing 1.5×10^6 GP+E-86 cells (29) were washed twice with Dulbecco's modified Eagle's medium (DME; Stem Cell Technologies, Vancouver, British Columbia, Canada) containing 20 mM *N*-2-hydroxyethylpiperazine-*N*-2-ethanesulfonic acid (HEPES; pH 7.2); 2 ml of DME-HEPES containing 2 μg of plasmid DNA and 200 μg of DEAE-dextran (molecular weight, 500,000; Pharmacia) per ml was then added, and the cells were incubated for 60 min at 37°C. The cells were washed twice with DME-HEPES containing 10% calf serum and incubated for 3 h in DME-HEPES containing 10% calf serum and 200 μM chloroquine, after which the medium was replaced with DME–10% calf serum; 45 h later, the culture medium was replaced with 4 ml of DME–10% calf serum containing 5 mM sodium butyrate to enhance transcription from the transfected plasmids (16). The medium containing secreted viruses was collected 18 h later and filtered, and Polybrene was added to a concentration of 10 μg/ml. This viral supernatant was then added to a 10-cm-diameter dish containing 5×10^5 NIH 3T3 cells. After a 16-h incubation at 37°C, the viral supernatant was replaced with DME–10% calf serum; the NIH 3T3 culture was maintained with twice-weekly feedings of this medium for 3 to 4 weeks to allow the emergence of transformed foci.

PCR recovery of proviral cDNA inserts. Transformed cells were plucked from infected NIH 3T3 cell cultures and expanded to 10^6 cells, and genomic DNA was prepared by proteinase K digestion, phenol extraction, and ethanol precipitation. Fifty-microliter PCR reaction mixes contained 300 μg of genomic DNA, 100 ng of primers annealing to retroviral vector sequences (5′ primer, CTCACTCCT TCTCTAGCTC; 3′ primer, CACACTGCTCGAATCAAGC), 200 μM each deoxynucleoside triphosphate, 1.25 U of *Pfu* DNA polymerase (Stratagene), and the buffer provided with the polymerase. Thermal cycles were as follows: 1 at 95°C for 60 s; 5 at 95°C for 60 s, 50°C for 30 s, and 72°C for 300 s; and 30 at 95°C for 30 s, 50°C for 30 s, and 72°C for 300 s. After cycling, 150 μl of 1.5 M ammonium acetate was added, and the reaction mix was extracted with chloroform and precipitated with 2 volumes of ethanol. Specific PCR products in a portion of the reaction were identified by Southern blotting, using a probe detecting the coamplified *supF* gene. The amplified DNA was then digested with *Mlu*I and *Bsi*WI, and PCR products were purified by agarose gel electrophoresis, electroelution, and ethanol precipitation. The purified PCR products were ligated with *Mlu*I-*Bsi*WI-digested pCTV3K and used for transformation of *E. coli* MC1061/p3. Plasmids were then isolated from single bacterial clones. Plasmid clones which tested positive for transformation were recloned by bacterial transformation and retested for transformation to ensure that transformation was reproducible and was not due to a contaminating plasmid.

RESULTS AND DISCUSSION

Design of retroviral vectors. A set of retroviral vectors was constructed specifically for cDNA library transfer and expression cloning (Fig. 1). The pCTV vectors were made as compact as possible to maximize stability of cDNA inserts during propagation both as plasmids in *E. coli* and as retroviruses. They contain a pair of *Bst*XI sites separated by a stuffer fragment to allow the use of noncomplementary *Bst*XI adapters for cDNA insertion (45). Inserted cDNAs are included in the genomic viral transcript which initiates in the 5′ LTR. This transcript also contains an extended packaging signal to maximize viral titers (2). The *gag* initiation codon within this region has been removed by mutagenesis to allow efficient translation from initiator codons within inserted cDNAs (38). All splicing signals have been removed from the pCTV vectors to avoid the activation of cryptic splice sites within inserted cDNAs, which

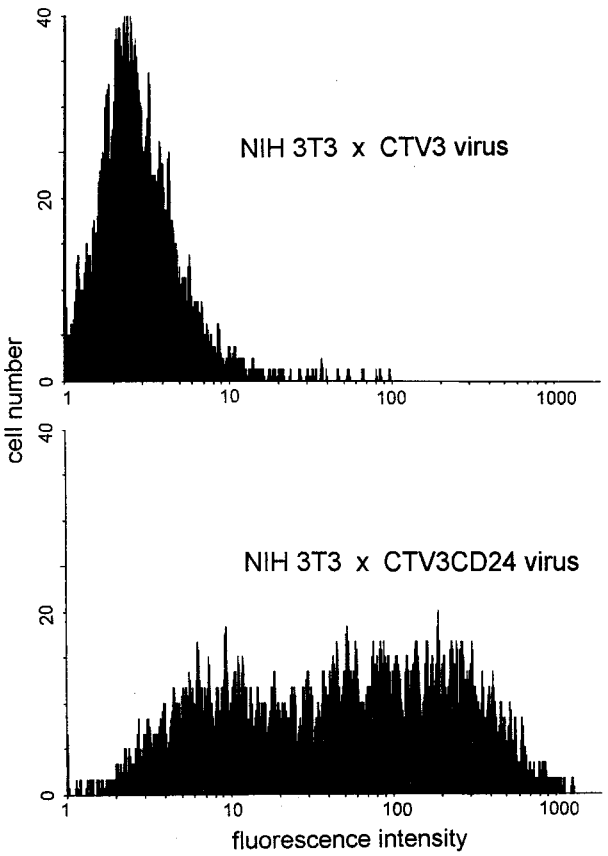

FIG. 2. Determination of cDNA transfer and expression efficiencies. GP+E-86 packaging cells were transfected with pure pCTV3 or pCTV3CD24 plasmid clones, the latter carrying a cDNA encoding the cell surface protein CD24 (25). NIH 3T3 cells were infected with the resulting viral supernatants and 2 days later were stained with an anti-CD24 monoclonal antibody complexed with the fluorochrome phycoerythrin. Expression of CD24, as measured by fluorescence intensity, was analyzed by flow cytometry as described previously (25).

would prevent cDNA transmission. The pCTV vectors also contain SV40 origins of replication so that the libraries can be used for expression cloning strategies dependent on episomal replication in COS cells (24, 45).

Production of retroviruses. Transient transfection of retroviral packaging cell lines was used to achieve the simultaneous, rapid, and proportionate conversion of very large numbers of cDNA clones into cDNA-bearing retroviruses. With a DEAE-dextran-mediated transfection protocol combined with chloroquine and butyrate treatments, titers of about 10^5 infectious viruses per ml, as determined by hygromycin resistance, were obtained with the GP+E-86 packaging line (29). Infection of NIH 3T3 cell cultures with the viral supernatants produced by transient transfection resulted in the transfer of functional viruses to about 70% of the cells, as measured by expression of a cell surface protein encoded by a virus-borne cDNA (Fig. 2). Transient transfection of GP+E-86 cells with pCTV3-based cDNA libraries typically resulted in somewhat lower titers, in the range of 1×10^4 to 4×10^4/ml, possibly reflecting the lower purity of the plasmid DNA obtained directly from bacterial colonies of the plated library. Such titers were still sufficiently high to allow the transfer of 100,000 cDNA clones by using 5 ml of supernatant for the infection of about 300,000 recipient cells. This is in striking contrast to our previous experience with direct transfection of NIH 3T3 and other cell lines, in which case the use of fully optimized transfection procedures (10, 11) and expression vectors (24, 37) resulted in cDNA transfer to only 0.1 to 0.5% of the recipient cells.

TABLE 1. Efficiencies of transfer and expression of transforming cDNAs

Transfected plasmid(s)	Amt of DNA	No. of transformed foci
pCTV1	1 μg	0
pCTV1 + pCTV1/N-Ras	1 μg + 1 ng	~300[a]
pCTV1 + pCTV1/N-Ras	1 μg + 0.1 ng	39
pCTV1 + pCTV1/Polyoma	1 μg + 1 ng	~500[a]
pCTV1 + pCTV1/Polyoma	1 μg + 0.1 ng	62

[a] Estimated focus number, due to high density of foci on dishes.

Detection of rare oncogenic cDNAs within transferred viral populations. One of our initial concerns about using retroviral vectors for cDNA library transfer was the possibility that many cDNA clones, especially longer ones, would be inefficiently transmitted or partially deleted during transmission. This might occur either through interference of the cDNA inserts with the transcription or packaging of the viral genomic RNA or through recombination between different cDNA clones sharing a common packaging cell or virus particle. To competitively test the fidelity of cDNA transmission, we constructed pCTV1 vectors carrying two distinctly different oncogenic cDNAs, i.e., an activated N-Ras cDNA of only 650 bp and a 2,850-bp region encompassing the entire early gene region of polyomavirus. In addition to its much larger size, the polyomavirus clone included alternatively utilized splice sites and two potential polyadenylation signals, features which would be expected to be found in some cDNAs and which have the potential to block formation of the intact genomic RNA needed for transmission. The N-Ras and polyomavirus vectors were highly diluted with pCTV1 vector to mimic a library containing rare oncogenic cDNAs. These DNA mixtures were then used to transfect GP+E-86 cells, and the resulting supernatants containing released virus were used to infect NIH 3T3 cells. The small N-Ras clone and the large and complex polyomavirus clone were transmitted to the NIH 3T3 cells with similar efficiencies, which were roughly equivalent to their abundances within the population of plasmids used for the initial transfection of the packaging cells (Table 1). Large numbers of transformed foci were induced in the infected NIH 3T3 cell cultures even when the abundances of the vectors carrying the N-Ras or polyomavirus clones were only 1:10,000 relative to inert vector. This result suggested that retrovirally transferred cDNA libraries could be used to detect the presence of transforming cDNAs with abundances of less than 1:100,000.

Recovery of retrovirally transferred cDNAs. To facilitate the analysis of the large number of transformed cell clones which were expected to be generated from cDNA library transfers, a PCR strategy was developed for rapidly recovering cDNAs from selected recipient cell clones and immediately reinserting them into the pCTV3 vector (Fig. 3). Several thousand bacterial transformants were obtained from cDNA recovery procedures performed on transformed NIH 3T3 cell clones derived from infection with the 10,000:1 mixture of pCTV1 with pCTV1/N-Ras. The yields of bacterial transformants from recovery procedures performed on the polyomavirus-transformed cell clones were about 20-fold lower, in proportion to the lower yield of PCR amplification of the longer polyomavirus cDNA. All recovered cDNA clones which were examined had retained the original cDNA structure, with the exception in all polyomavirus clones of the loss of the 61-bp intron spe-

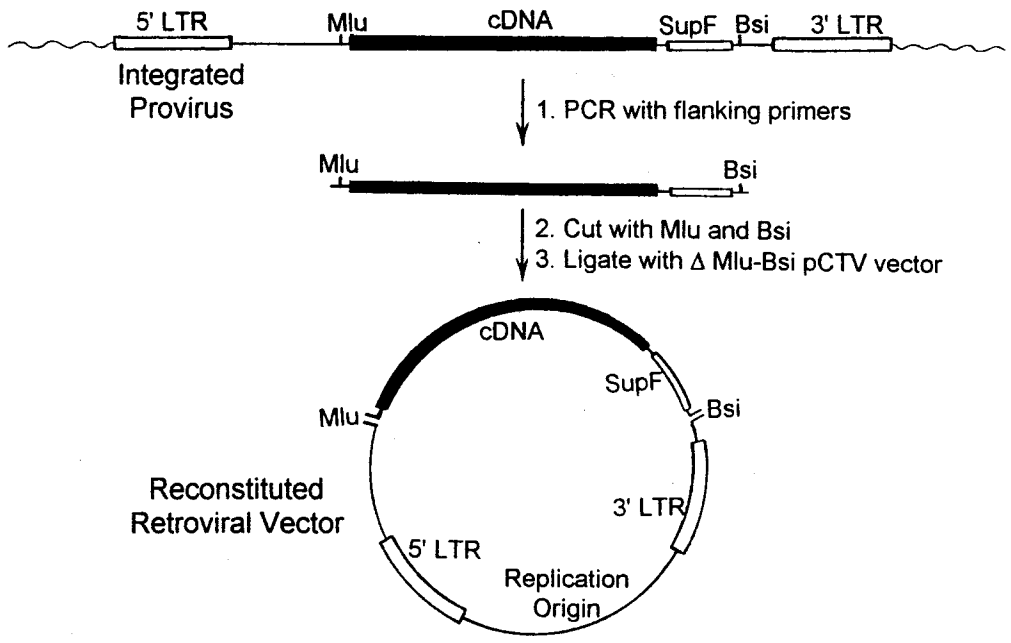

FIG. 3. Procedure for recovering cDNAs from integrated proviruses. Genomic DNA is prepared from infected cell clones and used as template for PCR amplification with a pair of primers complementary to vector sequences upstream of the site of cDNA insertion and downstream of the *supF* gene. After digestion with restriction enzymes *Mlu*I and *Bsi*WI, PCR products are gel purified and ligated between the same sites in pCTV3K, a variant of pCTV3 which has a kanamycin resistance gene rather than a *supF* gene between the *Mlu*I and *Bsi*WI sites. The formation of a complete plasmid which can be selected on the basis of suppressor tRNA activity is therefore dependent on the incorporation of the amplified *supF* gene, along with the accompanying cDNA insert. The plasmids isolated from transformed *E. coli* are fully functional retroviral vectors which can be immediately used for testing the activity of the recovered cDNA by transfection into packaging cells and infection of secondary recipients.

cific for the mRNA encoding the oncogenic middle T protein. This intron would be expected to be spliced out during viral transmission. Three recovered N-Ras cDNA clones and seven recovered polyomavirus cDNA clones were reconverted to retroviral form by transient transfection of GP+E-86 cells and then used to infect NIH 3T3 cells. All 10 of these clones were highly transforming, demonstrating that their oncogenic potential had been faithfully maintained through the process of transmission, PCR recovery, and retransmission.

Isolation of oncogenic cDNAs from retrovirally transferred libraries. Three separate cDNA libraries derived from hemopoietic cell lines were constructed in the pCTV vectors. An estimated 300,000 viral clones were generated from these libraries by transient transfection of GP+E-86 packaging cells and transferred to NIH 3T3 cells by infection. Cell cultures infected with these libraries developed a total of 83 distinctly transformed foci, with a variety of morphologies and growth rates, while uninfected cell cultures or those infected with viruses derived from the pCTV3 vector alone had 3- to 10-fold-lower frequencies of occurrence of transformed foci, equivalent to the spontaneous rate of transformation of the NIH 3T3 cells used in these experiments.

The foci of transformed cells which developed in the library-infected cultures were isolated and expanded, and transferred cDNAs within them were amplified by PCR and then cloned by ligation into the pCTV3K vector. Plasmids containing cDNA inserts of the expected sizes were tested for transforming activity by infection of NIH 3T3 cells, and the positive clones were sequenced to determine their identities. The mean sizes of recovered cDNAs with transforming activity were very similar to the mean sizes of the cDNAs in the originating libraries (Table 2). Therefore, there did not appear to be significant discrimination against the transmission of larger cDNAs under the competitive circumstances of library transfer.

Table 3 lists the 19 cDNAs with transforming activity which were recovered from PCRs performed on 47 of the transformed recipient cell clones. Twelve of the cDNAs encode proteins which have exact sequence identities to entries in the sequence databases. Three of these (Raf-1, Ect2, and Lck) are known to be able to oncogenically transform NIH 3T3 cells. Transforming activity of Raf-1 and Ect2 requires N-terminal truncation (28, 34, 47), as had occurred in the cDNAs cloned from our retroviral libraries. Lck transforms NIH 3T3 cells when its kinase activity is stimulated by mutations in negative regulatory sites (30), but the Lck cDNA which we cloned is predicted to encode a protein lacking kinase activity. Therefore, its weak transforming activity may occur through a different mechanism. The truncated Lck cDNA included three out-of-frame ATG codons upstream of the first ATG in the Lck-encoding reading frame. When these upstream ATGs were removed, which would be predicted to improve the translation efficiency of the cDNA (26), both the frequency and the potency of transformation were increased (Table 3). This finding implies that transformation was indeed mediated by the predicted translation product.

TABLE 2. Comparison of cDNA sizes in libraries versus sizes of recovered cDNAs

Library		Recovered cDNAs	
Name	Mean size (bp) of cDNAs	No.	Mean size (bp)
L16, L17	2,200	6	2,300
L18, L20	1,600	9	1,700
L19	2,200	2	1,600
L25	1,800	1	1,200
L27	2,500	1	2,400

TABLE 3. Transforming cDNAs isolated from retrovirally transferred cDNA libraries

cDNA[a]	Size (bp)	Identity	Transforming potency[b]	Transforming frequency[c]
TL16-8cC9	3,200	β-Catenin, N truncated	Moderate	0.05
TL16-8/ATG[d]		β-Catenin, ATG enhanced	Moderate	0.10
TL16-11cB1	1,600	No similarity to database sequences	Weak	0.01
TL16-13cB2	2,800	No similarity to database sequences	Moderate	0.01
TL16-25c2	1,900	Raf-1, N truncated	Strong	0.4
TL17-5cA3	2,500	Ect2, N truncated	Strong	0.5
TL17-11cB5	1,800	Lck, N truncated	Weak	0.005
TL17-11/ATG[d]		Lck, ATG enhanced	Moderate	0.03
TL18c2	800	Phospholipase C-γ_2, N + C truncated	Moderate	0.01
TL18c2/ATG[d]		Pholipase C-γ_2, ATG enhanced	Moderate	0.05
TL18-2c1	2,000	Thrombin receptor, full length	Strong	ND
TL18-3c1	1,600	Thrombin receptor, full length	Strong	0.1
TL18-5c1	1,700	Thrombin receptor, full length	Strong	ND
TL18-6cB4	1,200	Thrombin receptor, C truncated	Strong	ND
TL18-8c6	2,100	Thrombin receptor, full length	Strong	ND
TL18-9c1	1,800	CDC24 GNEF family, full length	Strong	0.2
TL18-10cA1	2,700	Low similarity to G_βs, antisense	Weak	0.02
TL19c39	1,300	Thrombin receptor, C truncated	Strong	ND
TL19-10c2	1,900	CDC24 GNEF homolog, N truncated	Strong	0.5
TL20-6cB8	1,500	Spi-2 protease inhibitor, antisense	Moderate	0.05
TL25-12-4-4	1,200	No similarity to database sequences	Strong	0.05
TL27-19-1-1	2,400	CDC24 GNEF homolog, C truncated	Strong	0.3

[a] TL16 and TL17 cDNAs were cloned from the GM979 erythroleukemic cell line cDNA library, TL18, TL19, and TL20 cDNAs were from the 32D myeloid cell line library, and the TL25 and TL27 cDNAs were from the B6SUtA$_1$ hemopoietic progenitor cell line library.

[b] Maximal density of cell growth and degree of morphological transformation observed in foci arising in NIH 3T3 cultures after infection with retroviral vectors carrying the indicated cDNA.

[c] Ratio of transformed colonies to total colonies arising in low-density NIH 3T3 cultures infected with pCTV3 retroviruses carrying the indicated cDNA. Colonies were counted and scored as transformed or nontransformed after 5 to 14 days under hygromycin selection. The spontaneous transformation frequencies of NIH 3T3 cultures infected with the pCTV3 vector alone ranged from 0.0001 to 0.001. The pCTV3/N-Ras virus had a transformation frequency of 0.7. ND, not determined.

[d] The sequences flanking presumptive start codons in these cDNAs were altered, as listed below, to optimize translation efficiency as follows: TL16-8cC9, CTGTTATGG; TL16-8/ATG, CCACCATGG, and one upstream ATG removed; TL17-11cB5, TGTGGATGG; TL17-11/ATG, CCACCATGG, and three upstream ATGs removed; TL18c2, TGGACATGC; TL18c2/ATG, CCACCATGG, and one upstream ATG removed. The cDNAs were also truncated immediately upstream of the modified start codons, to remove out-of-frame ATG codons which would be expected to reduce translation efficiency from the presumptive start codons.

The other cDNAs represented in the databases were derived from genes which were not previously identified as having transforming activity, and therefore their mechanisms of transformation are speculative. Thrombin receptor expression and its cleavage by thrombin in the serum used for cell culture presumably activates a G-protein-coupled mitogenic signalling pathway (27, 48) in NIH 3T3 cells, as has been demonstrated for ligand-stimulated serotonin and muscarinic receptors (18, 23). The transforming effects of the truncated β-catenin may be due to its disruption of the function of β-catenin-associated cell adhesion molecules such as E-cadherin (19), leading to transformation via loss of contact-induced growth repression. Alternatively, expression of truncated β-catenin may itself trigger a mitogenic signal transduction pathway in NIH 3T3 cells. The mechanism of action of the truncated phospholipase C-γ_2 is even more obscure, as it does not encode the catalytic domains required for mitogenic signal transduction (14). It could conceivably act by interfering with the function of an unknown inhibitor of the endogenous phospholipase C-γ_1 of NIH 3T3 cells. As seen for the Lck cDNA, the transforming activities of the truncated β-catenin and phospholipase C-γ_2 cDNAs were increased (Table 3) when upstream ATGs were removed and the sequence contexts of the presumptive initiation codons were optimized for translation (26).

The transforming activity of the antisense Spi-2 cDNA was presumably due to an inhibitory effect on the stability or translation of mRNAs encoding the Spi-2 protease inhibitor (20), leading to increased secreted protease activity and disruption of cell-cell or cell-matrix contacts which normally repress proliferation.

The remaining seven of the transforming cDNAs do not have high sequence identity with any entries in the DNA sequence databases. Three of these cDNAs encode proteins which include domains with homology to the CDC24 family of guanine nucleotide exchange factors (6) and therefore are presumed to transform NIH 3T3 cells by constitutively activating signal transduction through Ras-like proteins. One of the other novel clones has an antisense-oriented reading frame with weak but significant sequence similarity to beta subunits of G proteins and thus potentially acts as a suppressor of an endogenous G-protein-coupled signal transduction pathway which inhibits transformation of NIH 3T3 cells.

Of the 14 different types of cDNAs which were isolated in our screening of the retroviral cDNA libraries, 8 encoded proteins which are known or strongly implicated as positive inducers of mitogenic signals, and one (TL18-10cA1) may be an antisense suppressor of an inhibitor of mitogenic signals. This high yield of known or presumptive growth regulators confirms the validity of using the NIH 3T3 focus formation assay for detecting cDNAs encoding proteins with these functions. The fact that four of these cDNAs had not previously been cloned confirms our expectation that many novel growth regulators remain to be cloned via screening for transformation of NIH 3T3 cells or other transformation-susceptible cell lines. Given the proportion of growth regulators isolated in our screen, there is a high probability that the three novel cDNAs with no obvious coding sequence similarity to known proteins will also prove to be bona fide growth regulators. Alternatively, one or more of these cDNAs may encode or repress the expression of proteins which morphologically transform NIH 3T3 cells by

altering cell surface interactions (possible examples being the Spi-2 and β-catenin clones).

Some of the cDNAs described above appear to have acquired transforming activity via mutations arising in the process of cDNA library construction, i.e., truncation or inversion. The expression of such altered cDNAs can therefore be critical to the induction of a desired phenotype, either through the removal of negative regulatory domains from a phenotype inducer (e.g., truncated Raf-1 and Ect2) or by the generation of suppressors of phenotype repressors (e.g., antisense Spi-2). The possible dependence on specific mutation for the revelation of a phenotype such as oncogenic transformation means that only a small subset of the cDNAs derived from transcripts of a given gene may be functionally detectable. Furthermore, many of the cDNAs cloned in this study had transformation frequencies considerably less than 1. While this may in part be due to the transmission of a partially deleted virus or to promoter interference resulting in the inactivation of the LTR, in most cases it is likely due to critical thresholds of expression not being reached by most of the proviruses carrying the cDNA. Thus, the ability to transfer and screen very large cDNA libraries can be an essential requirement for the cloning of such cDNAs.

The NIH 3T3 cell line is obviously a very effective recipient cell for detecting and cloning oncogenes. This is largely because of its high susceptibility to transformation by the expression of proteins which activate the Ras-Raf signal transduction pathway. When expressed in the C3H10T1/2 fibroblast line, only the truncated Raf cDNA and the TL19-10c2 cDNA caused transformation, while all other cDNAs listed in Table 3 were negative. However, we have recently used the C127 epithelial cell line as a recipient for cDNA library transfer and have isolated one cDNA encoding a putative transcription factor which transforms this line but does not transform NIH 3T3 cells. This implies that the use of alternative recipient cells can either restrict or expand the spectrum of transforming cDNAs which can be detected and cloned.

The 19 transforming cDNAs listed in Table 3 were isolated from a relatively small number of transferred clones, representing just 7% of the 4 million clones in the available libraries. The scale of our library screens has been restricted not by the number of cDNA clones which can be transferred but rather by the number of transforming cDNAs which can be processed and analyzed. There are presumably many more transforming cDNAs remaining to be detected in these libraries. Of the 14 different types of transforming cDNAs isolated, multiple independent clones were obtained only in the case of the thrombin receptor, reflecting the relatively high abundance of its mRNA in 32D cells (25a) and the apparent ability of this receptor to transform NIH 3T3 cells without a requirement for specific truncating mutations.

Two major advantages were obtained through the use of retroviral library transfer for the isolation of cDNAs capable of inducing proliferation of NIH 3T3 cells. The first was the ability to screen very large numbers of cDNA clones on an equivalent number of recipient cells. This resulted in a high rate of cDNA-induced versus spontaneous transformation events and thus a high rate of recovery of transforming versus inert cDNAs. The second advantage was the relatively high levels of expression obtained with retrovirally transferred cDNAs, which in our hands are 10-fold greater on average than those obtained with the best available stable transfection vectors. Efficient expression can be critical for obtaining transformation and other targeted phenotypes in recipient cells, particularly when an antisense mechanism is involved.

A third major advantage of retroviral cDNA library transfer, not exploited in this study, is the potential to use cell lines which have previously been completely inaccessible to expression cloning because of low transfection efficiencies. There are only a limited number of fibroblast and epithelial cell lines which can be stably transfected with efficiencies of greater than 0.1% (10, 11). In contrast, almost all cell lines and some primary cell types can be infected with retroviruses. We have recently begun generating retroviral cDNA libraries with the BOSC23 packaging cell line, which was established for the specific purpose of yielding high viral titers via transient transfection (41). The viral titers of up to 10^6/ml obtained with these packaging cells are sufficiently high to enable the transfer of million-member cDNA libraries to hemopoietic cell lines and the detection of rare cDNA clones conferring selectable phenotypes such as conversion to growth in the absence of cytokines (50). Use of the BOSC23 cell line should permit the transfer of large cDNA libraries to cell lines representing many different lineages as well as to some primary cell types. This would make retroviral transfer of cDNA libraries uniquely suitable for the cloning of cDNAs which can impose new phenotypes or complement genetic defects in specialized cell types.

ACKNOWLEDGMENTS

We thank H. Land and S. Cory for providing the pBabeNeo and JZen retroviral vectors used in the assembly of the pCTV vectors, B. Seed for providing the *supF* gene and *Bst*XI stuffer from his pCDM8 vector, M. Capecchi for the polyomavirus/TK promoter from pMC1, A. Bank and W. Pear for providing the GP+E-86 and BOSC23 packaging cell lines, J. Hassell for providing the polyomavirus early region clone, S. Nagata for the EF-1α promoter, and P. Lansdorp for CD24 detection reagents.

This work was funded by grants from the Medical Research Council of Canada, the National Cancer Institute of Canada, and the British Columbia Health Research Foundation and by a fellowship to I.W. from the Leukemia Research Foundation.

REFERENCES

1. **Beck, E., G. Ludwig, E. A. Auerswald, B. Reiss, and H. Schaller.** 1982. Nucleotide sequence and exact localization of the neomycin phosphotransferase gene from transposon Tn5. Gene **19:**327–336.
2. **Bender, M. A., T. D. Palmer, R. E. Gelinas, and A. D. Miller.** 1987. Evidence that the packaging signal of Moloney murine leukemia virus extends into the *gag* region. J. Virol. **61:**1639–1646.
3. **Bernard, H.-U., G. Krammer, and W. G. Rowekamp.** 1985. Construction of a fusion gene that confers resistance against hygromycin B to mammalian cells in culture. Exp. Cell Res. **158:**237–243.
4. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7:**1513–1523.
5. **Bishop, J. M.** 1991. Molecular themes in oncogenesis. Cell **64:**235–248.
6. **Boguski, M. S., and F. McCormick.** 1993. Proteins regulating Ras and its relatives. Nature (London) **366:**643–654.
7. **Brown, B. A., R. W. Padgett, S. C. Hardies, C. A. Hutchinson, and M. H. Edgell.** 1982. β-Globin transcript found in induced murine erythroleukemia cells is homologous to βh0 and βh1 genes. Proc. Natl. Acad. Sci. USA **79:**2753–2757.
8. **Chan, A. M.-L., M. Chedid, E. S. McGovern, E. S. Popescu, T. Miki, and S. A. Aaronson.** Expression cDNA cloning of a serine kinase transforming gene. Oncogene **8:**1329–1333.
9. **Chan, A. M.-L., T. P. Fleming, E. S. McGovern, M. Chedid, T. Miki, and S. A. Aaronson.** 1993. Expression cDNA cloning of a transforming gene encoding the wild-type Gα12 gene product. Mol. Cell. Biol. **13:**762–768.
10. **Chen, C., and H. Okayama.** 1987. High-efficiency transformation of mammalian cells by plasmid DNA. Mol. Cell. Biol. **7:**2745–2752.
11. **Chen, C. A., and H. Okayama.** 1988. Calcium phosphate-mediated gene transfer: A highly efficient transfection system for stably transforming cells with plasmid DNA. BioTechniques **6:**632–638.
12. **Chomczynski, P., and N. Sacchi.** 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. **162:**156–159.
13. **Cooper, C. S., M. Park, D. G. Blair, M. A. Tainsky, K. Huebner, C. M. Croce, and G. F. Vande Woude.** 1984. Molecular cloning of a new transforming gene from a chemically transformed human cell line. Nature (London) **311:**29–33.
14. **Emori, Y., Y. Homma, H. Sorimachi, H. Kawasaki, O. Nakanishi, K. Suzuki,**

and T. Takanawa. 1989. A second type of rat phosphoinositide-specific phospholipase C containing a src-related sequence not essential for phosphoinositide hydrolyzing activity. J. Biol. Chem. **264:**21885–21890.
15. **Eva, A., and S. A. Aaronson.** 1985. Isolation of a new human oncogene from a diffuse B-cell lymphoma. Nature (London) **316:**273–275.
16. **Gorman, C. M., B. H. Howard, and R. Reeves.** 1983. Expression of recombinant plasmids in mammalian cells is enhanced by sodium butyrate. Nucleic Acids Res. **11:**7631–7648.
17. **Greenberger, J. S., M. A. Sakakeeny, R. K. Humphries, C. J. Eaves, and R. J. Eckner.** 1983. Demonstration of permanent factor-dependent multipotential (erythroid/neutrophil/basophil) hematopoietic progenitor cell lines. Proc. Natl. Acad. Sci. USA **80:**2931–2935.
18. **Gutkind, J. S., E. A. Novotny, M. R. Brann, and K. C. Robbins.** 1991. Muscarinic acetylcholine receptor subtypes as agonist-dependent oncogenes. Proc. Natl. Acad. Sci. USA **88:**4703–4707.
19. **Hamaguchi, M., N. Matsuyoshi, Y. Ohnishi, B. Gotoh, M. Takeichi, and Y. Nagai.** 1993. $p60^{v\text{-}src}$ causes tyrosine phosphorylation and inactivation on the N-cadherin-catenin cell adhesion system. EMBO J. **12:**307–314.
20. **Inglis, J. D., M. Lee, D. R. Davidson, R. E. Hill.** 1991. Isolation of two cDNAs encoding novel alpha 1-antichymotrypsin-like proteins in a murine chondrocytic cell line. Gene **106:**213–220.
21. **Jacobson, A.** 1987. Purification and fractionation of poly(A)$^+$ RNA. Methods Enzymol. **152:**254–261.
22. **Johnson, G. R., T. J. Gonda, D. Metcalf, I. K. Hariharan, and S. Cory.** 1989. A lethal myeloproliferative syndrome in mice transplanted with bone marrow cells infected with a retrovirus expressing granulocyte-macrophage colony stimulating factor. EMBO J. **8:**441–448.
23. **Julius, D., T. J. Livelli, T. M. Jessell, and R. Axel.** 1989. Ectopic expression of the serotonin 1c receptor and the triggering of malignant transformation. Science **244:**1057–1062.
24. **Kay, R., and R. K. Humphries.** 1991. New vectors and procedures for isolating cDNAs encoding cell surface proteins by expression cloning in COS cells. Methods Mol. Cell. Biol. **2:**254–265.
25. **Kay, R., P. M. Rosten, and R. K. Humphries.** 1991. CD24, a signal transducer modulating B cell activation responses, is a very short peptide with a glycosyl phosphatidylinositol membrane anchor. J. Immunol. **147:**1412–1416.
25a. **Kirk, H., and R. Kay.** Unpublished data.
26. **Kozak, M.** 1991. Structural features in eukaryotic mRNAs that modulate the initiation of translation. J. Biol. Chem. **266:**19867–19870.
27. **LaMorte, V. J., E. D. Kennedy, L. R. Collins, D. Goldstein, A. T. Harootunian, J. H. Brown, and J. R. Feramisco.** 1993. A requirement for Ras protein function in thrombin-stimulated mitogenesis in astrocytoma cells. J. Biol. Chem. **268:**19411–19415.
28. **Li, P., K. Wood, H. Mamon, W. Haser, and T. Roberts.** 1991. Raf-1: a kinase currently without a cause but not lacking in effects. Cell **64:**479–482.
29. **Markowitz, D., S. Goff, and A. Bank.** 1988. A safe packaging line for gene transfer: separating viral genes on two different plasmids. J. Virol. **62:**1120–1124.
30. **Marth, J. D., J. A. Cooper, C. S. King, S. F. Ziegler, D. Tinker, R. W. Overell, E. G. Krebs, and R. M. Perlmutter.** 1988. Neoplastic transformation induced by an activated lymphocyte-specific protein tyrosine kinase ($pp56^{lck}$). Mol. Cell Biol. **8:**540–550.
31. **Martin-Zanca, D., S. H. Hughes, and M. Barbacid.** 1986. A human oncogene formed by the fusion of truncated tropomyosin and protein tyrosine kinase sequences. Nature (London) **319:**743–748.
32. **Miki, T., T. P. Fleming, D. P. Bottaro, J. S. Rubin, D. Ron, and S. A. Aaronson.** 1991. Expression cDNA cloning of the KGF receptor by creation of a transforming autocrine loop. Science **251:**72–75.
33. **Miki, T., T. P. Fleming, M. Crescenzi, C. J. Molloy, S. B. Blam, S. H. Reynolds, and S. A. Aaronson.** 1991. Development of a highly efficient expression cDNA cloning system: application to oncogene isolation. Proc. Natl. Acad. Sci. USA **88:**5167–5171.
34. **Miki, T., C. L. Smith, J. E. Long, A. Eva, and T. P. Fleming.** 1993. Oncogene *ect2* is related to regulators of small GTP-binding proteins. Nature (London) **362:**462–465.
35. **Miller, A. D., and G. J. Rosman.** 1989. Improved retroviral vectors for gene transfer and expression. BioTechniques **7:**980–990.
36. **Miyoshi, J., T. Higashi, H. Mukai, T. Ohuchi, and T. Kakunaga.** 1991. Structure and transforming potential of the human *cot* oncogene encoding a putative protein kinase. Mol. Cell. Biol. **11:**4088–4096.
37. **Mizushima, S., and S. Nagata.** 1990. pEF-BOS, a powerful mammalian expression vector. Nucleic Acids Res. **18:**5322.
38. **Morgenstern, J. P., and H. Land.** 1990. Advanced mammalian gene transfer: high titre retroviral vectors with multiple drug selection markers and a complementary helper-free packaging cell line. Nucleic Acids Res. **18:**3587–3596.
39. **Murphy, A. J. M., and A. Efstratiadis.** 1987. Cloning vectors for expression of cDNA libraries in mammalian cells. Proc. Natl. Acad. Sci. USA **84:**8277–8281.
40. **O'Bryan, J. P., R. A. Frye, P. C. Cogswell, A. Neubauer, B. Kitch, C. Prokop, R. Espinosa III, M. M. LeBeau, H. S. Earp, and E. T. Liu.** 1991. *axl*, a transforming gene isolated from primary human myeloid leukemia cells, encodes a novel receptor tyrosine kinase. Mol. Cell. Biol. **11:**5016–5031.
41. **Pear, W. S., G. P. Nolan, M. L. Scott, and D. Baltimore.** 1993. Production of high-titer helper-free retroviruses by transient transfection. Proc. Natl. Acad. Sci. USA **90:**8392–8396.
42. **Rayner, J. R., and T. J. Gonda.** 1994. A simple and efficient procedure for generating stable expression libraries by cDNA cloning in a retroviral vector. Mol. Cell. Biol. **14:**880–887.
43. **Seed, B.** 1983. Purification of genomic sequences from bacteriophage libraries by recombination and selection in vivo. Nucleic Acids Res. **11:**2427–2445.
44. **Seed, B.** 1987. An LFA-3 cDNA encodes a phospholipid-linked membrane protein homologous to its receptor CD2. Nature (London) **329:**840–842.
45. **Seed, B., and A. Aruffo.** 1987. Molecular cloning of the CD2 antigen, the T-cell erythrocyte receptor, by a rapid immunoselection procedure. Proc. Natl. Acad. Sci. USA **84:**3365–3369.
46. **Shimizu, K., M. Goldfarb, M. Perucho, and M. Wigler.** 1983. Isolation and preliminary characterization of the transforming gene of a human neuroblastoma cell line. Proc. Natl. Acad. Sci. USA **80:**383–387.
47. **Stanton, V. P., Jr., D. W. Nichols, A. P. Laudano, and G. M. Cooper.** 1989. Definition of the human *raf* amino-terminal regulatory region by deletion mutagenesis. Mol. Cell. Biol. **9:**639–647.
48. **Thien-Khai, H., D. T. Hung, V. I. Wheaton, and S. R. Coughlin.** 1991. Molecular cloning of a functional thrombin receptor reveals a novel proteolytic mechanism of receptor activation. Cell **64:**1057–1068.
49. **Thomas, K. R., and M. R. Capecchi.** 1987. Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. Cell **51:**503–512.
50. **Tognon, C., and R. Kay.** Unpublished data.
51. **Vogeli, G., E. Horn, M. Laurent, and P. Nath.** 1985. Recombinant DNA techniques: storage and screening of cDNA libraries with large numbers of individual colonies from initial transformations. Anal. Biochem. **151:**442–444.