Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 5167–5171, June 1991
Cell Biology

Development of a highly efficient expression cDNA cloning system: Application to oncogene isolation

(mouse hepatoma/stable expression/plasmid rescue/B-*raf* gene/*in vivo* activation)

TORU MIKI*, TIMOTHY P. FLEMING*, MARCO CRESCENZI*†, CHRISTOPHER J. MOLLOY*‡, SHELLEY B. BLAM*§, STEVEN H. REYNOLDS¶, AND STUART A. AARONSON*‖

*Laboratory of Cellular and Molecular Biology, National Cancer Institute, Building 37, Room 1E24, Bethesda, MD 20892; and ¶National Institute of Environmental Health Sciences, Biological Risk Assessment Program, Research Triangle Park, NC 27709

Communicated by G. D. Aurbach, March 19, 1991 (received for review October 16, 1990)

ABSTRACT **We developed an expression cDNA cloning system capable of generating high-complexity libraries with unidirectionally inserted cDNA fragments and allowing efficient plasmid rescue. As an application of this system, a cDNA library was constructed from an NIH 3T3 transformant induced by mouse hepatocellular carcinoma DNA. Transfection of NIH 3T3 cells by the library DNA led to the detection of several transformed foci from which identical plasmids with transforming ability could be rescued. Structure and sequence analysis of the cDNA clones revealed that the oncogene was created by recombinational events involving an unknown gene and the mouse homologue of the B-*raf* protooncogene. Detection of the same genetic rearrangement in independent primary transformants implied that generation of the oncogene occurred within the tumor rather than during DNA transfection or cDNA library construction. The high frequency at which clones were identified and the large sizes of some of the transforming cDNA inserts isolated suggest wide applicability of this mammalian expression cloning system for isolating cDNAs of biologic interest.**

In bacterial and yeast systems, genes can be cloned by complementing mutant alleles with introduced genes. Similarly, in mammalian systems, introduction of cDNAs placed under a suitable promoter can complement mutations in recipient cells or allow for selection of any identifiable phenotype. We have recently reported the development of an efficient system to construct unidirectional cDNA libraries (1), which has made it possible to isolate full-length coding sequences of several genes (1–6). In this report, we refined the vector to allow for high-level cDNA expression in mammalian cells and the ability to perform efficient plasmid rescue. We tested the potential of this expression cloning system by cloning an oncogene cDNA from a mouse hepatocellular carcinoma (7).**

EXPERIMENTAL PROCEDURES

cDNA Library Construction. cDNA libraries were constructed as reported (1), except for use of a designed vector and adaptor more suitable for expression cloning (see *Results*). The revised *Sfi* I adaptor consisted of two oligonucleotides, 5′-CCAATCGCGACC-3′ and 5′-GGTCGCGATTGGTAA-3′. Amplification of the library and preparation of the DNA were done by standard procedures (8).

Cell Culture and DNA Transfection. All cells used were derivatives of NIH 3T3 cells (9). Calcium phosphate transfection (10) was used to stably introduce DNA into cells. Cells were maintained in Dulbecco's modified Eagle's medium (DMEM)/5% calf serum.

Plasmid Rescue. Genomic DNA (1.2 μg) was cleaved by *Xho* I and ligated at the concentration of 3 ng/μl. Four aliquots (100 μl each) of PLK-F′ competent cells (Stratagene) were transformed by the ligated DNA, as directed by the manufacturer. After heat shock, the cells were diluted 10-fold with S.O.C. medium (BRL) containing 1 mM isopropyl β-D-thiogalactopyranoside to induce expression of the *neo* gene driven by the *tac* promoter. The culture was incubated for 2 hr with shaking and plated on NZY hard agar (8) containing ampicillin (100 μg/ml), kanamycin (25 μg/ml), and isopropyl β-D-thiogalactopyranoside (100 μM).

Recombinant DNA Techniques. Preparation of λ and plasmid DNA was performed as described (1). Genomic DNA was extracted by the standard procedure (8). Total RNA was isolated and poly(A)-selected as described (1). DNA fragments were isolated by Geneclean (Bio 101 Inc., La Jolla, CA) and labeled by random priming using an Oligo labeling kit (Pharmacia). Hybridizations were performed at 42°C in 50% (wt/vol) formamide/5× standard saline citrate (8)/2.5× Denhardt's solution (8)/7 mM Tris hydrochloride, pH 7.5; denatured calf thymus DNA at 0.1 mg/ml/tRNA at 0.1 mg/ml. Nucleotide sequence was determined by the dideoxynucleotide chain-termination method (11).

RESULTS

Development of an Efficient Expression cDNA Cloning System. The λpCEV27 system was developed to clone cDNAs by means of stable phenotypic changes induced by a specific cDNA. Use of a λ–plasmid composite vector made it possible to generate high-complexity cDNA libraries and to efficiently excise the plasmid from the stably integrated phagemid DNA. This phagemid vector (Fig. 1) contained several features including two *Sfi* I sites for construction of cDNA libraries by using the automatic directional cloning (ADC) method (1), a Moloney murine leukemia virus long terminal repeat promoter suitable for cDNA expression in mammalian cells, the simian virus 40 (SV40) promoter-driven *neo* gene as a selectable marker, and multiple excision sites for plasmid rescue from genomic DNA. The λpCEV27 vector also incorporated the rat pre-pro-insulin poly(A) signal downstream from the cDNA cloning site (Fig. 1).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: SV40, simian virus 40.

†Present address: Dipartimento di Biologia Cellulare e Dello Sviluppo, Universita di Roma "La Sapienza", 00185 Rome, Italy.
‡Present address: Department of Cardiovascular Biochemistry, Bristol-Myers Squibb Pharmaceutical Research Institute, Princeton, NJ 08543-4000.
§Present address: Berlex Biosciences, Alameda, CA 94501.
‖To whom reprint requests should be addressed.

**The sequence reported in this paper has been deposited in the GenBank data base (accession no. M64429).

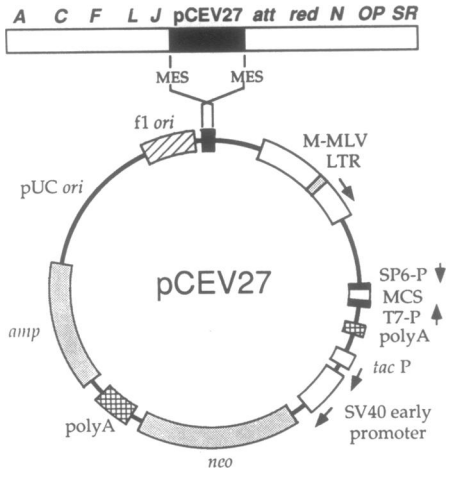


FIG. 1. Structure of the cDNA cloning-expression vector λpCEV27. Structure of the λpCEV27 vector is shown above with location of λ genes. Plasmid portion is enlarged and shown below as circular map. The multiple excision site (MES) contains restriction sites for infrequent cutters: *Not* I, *Xho* I, *Pvu* I, and *Mlu* I. The multiple cloning site (MCS) contains restriction sites for *Bam*HI, *Sal* I, *Sfi* I(A), *Eco*RI, *Bgl* II, *Hin*dIII, *Sfi* I(B), *Sal* I, and *Bst*EII in clockwise order. The two *Sfi* I sites are used to insert cDNA molecules by the automatic directional cloning method (1), and the two *Sal* I sites are used to release the inserts. SP6-P and T7-P represent phage promoters for SP6 and T7 RNA polymerases, respectively. The *trp–lac* fused promoter *tac* and SV40 early promoter are used to express the *neo* structural gene in *Escherichia coli* (kanamycin resistance) and eukaryotic cells (G-418 resistance), respectively. In contrast to λpCEV15 (1), the *bona fide* promoter of the *neo* gene was removed so as to fuse the SV40 early promoter directly to the *neo* structural gene. Thus, in λpCEV27, expression of the *neo* gene in *E. coli* was achieved by transcription from the *tac* promoter, inserted upstream from the SV40 early promoter. Directions of transcription from the promoters are shown by arrows. Polyadenylylation signals are labeled as poly(A). In earlier vectors (1), replication of rescued plasmids in *E. coli* was unstable. In λpCEV27, the replication origin of pUC19 was used to increase copy number. Locations of replication origins (*ori*) and ampicillin resistant gene (*amp*) are shown. M-MLV LTR, Moloney murine leukemia virus long terminal repeat.

The strategy for expression cloning of oncogene cDNAs is summarized in Fig. 2. When library cDNA is used to transfect mammalian cells, cDNA clones are integrated with recombination between λ DNA and host genomic DNA. For plasmid rescue, genomic DNA extracted from the transformant is subjected to digestion with an enzyme that can cleave the λ–plasmid junctions. The resulting DNA can then be circularized and used for bacterial transformation. For this purpose, the sites for two additional infrequent cutters, *Xho* I and *Mlu* I, were included along with the *Not* I site.

Expression cDNA Cloning of an Oncogene in a Furfural-Induced Hepatoma. We had analyzed (7) hepatocellular tumors of the mouse strain B6C3F1 for presence of activated oncogenes. Although the majority were activated *ras* or c-*raf* oncogenes, four could not be identified. The sources of these oncogenes were tumors designated T4, T18, T23, and T28. One (T23) was spontaneously generated, whereas the others were associated with chronic furfural exposure (7). Genomic DNAs of NIH 3T3 transformants containing each of the unidentified oncogenes were examined under low-stringency hybridization conditions by using a number of known and potential oncogene probes. None showed DNA fragments with either increased intensity or abnormal sizes relative to those detected in NIH 3T3 control DNA (data not shown). Thus, none of these transforming genes appeared closely related to any of the genes used as probes.

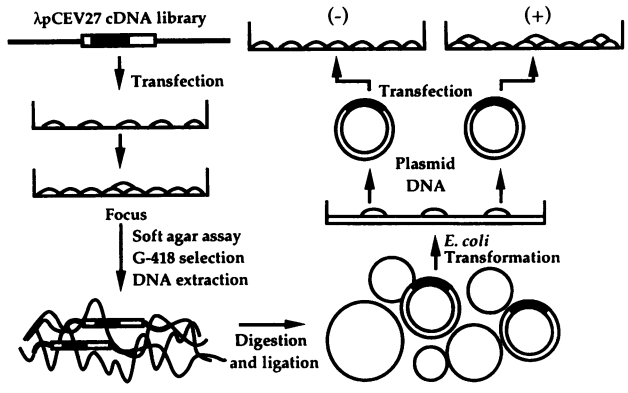


FIG. 2. Strategy for expression cloning of transforming gene cDNAs. NIH 3T3 cells are transfected by λpCEV27 cDNA library DNA and scored at 14–17 days for transformed foci. Transformed cells can then be assayed for G-418 resistance and colony formation in soft agar. After expansion to mass culture, genomic DNA is isolated and subjected to plasmid rescue by digestion with either *Not* I, *Xho* I, or *Mlu* I, followed by ligation at low DNA concentration and transformation to a suitable bacterial strain. Bacterial colonies resistant to both ampicillin and kanamycin are isolated. Plasmid DNA extracted from each colony is tested by transfection analysis on NIH 3T3 cells to identify the transforming cDNA clone.

Conventional approaches for molecular cloning of transforming genes identified by transfection analysis have generally taken advantage of differences in repetitive sequences between the transfecting DNA and those of the recipient cell. Such sequences provide molecular probes for use in identifying and isolating genomic fragments of the transforming genes. Rodent oncogenes have been particularly difficult to molecularly clone because relatedness of their repetitive sequences precludes discrimination of such genes from genomic DNA of mouse NIH 3T3, commonly used to isolate oncogenes by DNA transfection. To overcome such inherent

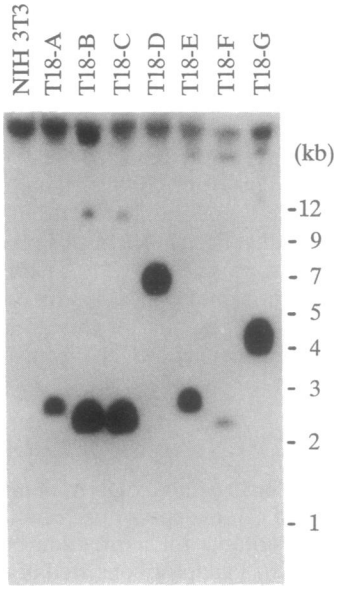


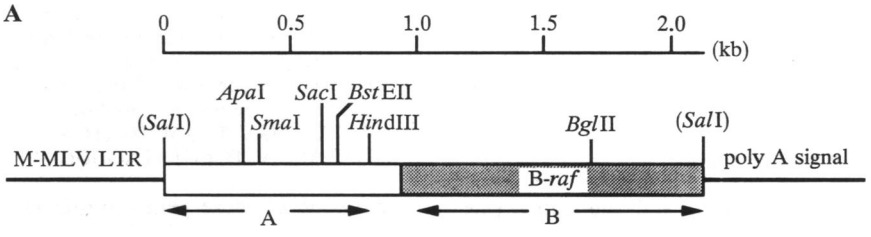
FIG. 3. Detection of T18 oncogene cDNA inserts in the library-induced transformants. Genomic DNAs from individual transformants (T18-A to T18-G) were digested with *Sal* I to release cDNA inserts (see Fig. 1). Digested DNA (5 μg) was separated on a 0.5% agarose gel by electrophoresis and transferred to nitrocellulose membrane (Nitrocellulose GTG; FMC). The Southern blot was probed with the cDNA insert rescued from transformant T18-B. NIH 3T3 genomic DNA was used as control. Location of each fragment of the molecular-size-marker (1-kb ladder; BRL) is shown at right in kb.

difficulties, we attempted to use the λpCEV27 expression cloning system to isolate transforming cDNA from one furfural-induced mouse tumor, T18. A cDNA library (3×10^6 independent clones) was constructed from poly(A)$^+$ RNA extracted from a second-cycle DNA transfectant of the tumor. Transfection of 80 plates of NIH 3T3 cells with the expression cDNA library at 5 μg per plate led to the detection of seven foci that demonstrated G-418 resistance. These results indicated that each transformant had taken up and stably integrated vector DNA, making it likely that the transformed foci were induced by exogenous cDNA, rather than arising as a result of spontaneous transformation.

Two of these transformants, designated T18-B and T18-C, were selected for plasmid rescue. By restriction mapping of several distinct plasmids obtained from each transformant, it was possible to establish that one plasmid rescued from each had the identical insert. These results suggested that this cDNA might encode the oncogene product. Transfection analysis of each rescued plasmid DNA demonstrated that these same two cDNA clones possessed high-titered transforming activity of $\approx 10^3$ focus-forming units per nmol of DNA, whereas none of the other plasmids rescued from the same transformants showed detectable activity.

To determine whether other transformants induced by the cDNA library contained the T18 oncogene, genomic DNA extracted from each primary transformant was digested with *Sal* I to release cDNA inserts from the vector and subjected to Southern blotting analysis with the rescued T18-B cDNA insert as probe. Because *Sal* I is an infrequent cutter for mammalian DNA, genomic DNA was cleaved to very large fragments, which remained near the origin of the gel. Thus, the relatively small cDNA fragments released from cellular DNAs by *Sal* I cleavage could be separated from the bulk of genomic fragments. Fig. 3 shows that each of the cDNA library transformants contained the T18 oncogene cDNA insert. The sizes ranged from ≈2.1 kilobases (kb) to 7 kb, suggesting that several of the inserts represented independent cDNA clones of the oncogene.

Structural Analysis of T18 Oncogene cDNA. A detailed restriction map of the 2.1-kb insert of one of the transforming




```
B   1  AC CAA TCG CGA CCG AGC GCG CTC CCG CTT CCC CGG CGG CGG CGC AGG ACA ATG GAG GGT GGC TGT GGA GAA GGC GGC GGC GGC ACG GGC TCC GGC CGG AGC GCC GCA
    1     Gln Ser Arg Pro Ser Ala Leu Pro Leu Pro Arg Arg Arg Arg Arg Thr MET Glu Gly Gly Cys Gly Glu Gly Gly Gly Gly Thr Gly Ser Gly Arg Ser Ala Ala

  108  GCG GCG CGC CGA GCG GGG AGA ATG AGG CCG AGA GCG CAG GGC CCG GAC TCG GAG AGC GGC GGC GAG GCG TCC CGG CTC AAC CTG TTG GAC ACT TGC GCC GTG TGC CAC
   36  Ala Ala Arg Arg Ala Gly Arg MET Arg Pro Arg Ala Gln Gly Pro Asp Ser Glu Ser Gly Gly Glu Ala Ser Arg Leu Asn Leu Leu Asp Thr Cys Ala Val Cys His

  216  CAG AAC ATC CAG AGC CGG GTG CCC AAG CTG CCC TGC CTG CAC TCG TTC TGC CAG CGC TGT TTG CCC GCG CCG CAG CGC TAT CTC ATG CTG ACG GCG CCC GCG CTG
   72  Gln Asn Ile Gln Ser Arg Val Pro Lys Leu Leu Pro Cys Leu His Ser Phe Cys Gln Arg Cys Leu Pro Ala Pro Gln Arg Tyr Leu MET Leu Thr Ala Pro Ala Leu

  324  GGC TCG GCA GAG ACC CCT CCA CCC GCT CCC GCC CCC GCC CCC GCC CCG GGC TCC CCG GCC GGT GGT CCT TCG CCA TTC GCC ACC CAA GTT GGA GTC ATT CGA TGC CCA
  108  Gly Ser Ala Glu Thr Pro Pro Pro Ala Pro Ala Pro Ala Pro Ala Pro Gly Ser Pro Ala Gly Gly Pro Ser Pro Phe Ala Thr Gln Val Gly Val Ile Arg Cys Pro

  432  GTT TGC AGT CAA GAG TGT GCT GAG AGA CAC ATC ATA GAC AAC TTT TTT GTG AAG GAC ACC ACT GAA GTT CCT AGT AGT ACA GTA GAA AAG TCT AAT CAG GTA TGT ACA
  144  Val Cys Ser Gln Glu Cys Ala Glu Arg His Ile Ile Asp Asn Phe Phe Val Lys Asp Thr Thr Glu Val Pro Ser Ser Thr Val Glu Lys Ser Asn Gln Val Cys Thr

  540  AGC TGT GAA GAC AAT GCA GAA GCT AAT GGG TTT TGT GTA GAG TGT GTT GAA TGG CTC TGC AAG ACA TGT ATT AGA GCT CAC CAG AGG GTG AAG TTC ACA AAA GAC CAC
  180  Ser Cys Glu Asp Asn Ala Glu Ala Asn Gly Phe Cys Val Glu Cys Val Glu Trp Leu Cys Lys Thr Cys Ile Arg Ala His Gln Arg Val Lys Phe Thr Lys Asp His

  648  ACA GTC AGG CAG AAA GAA GAA GTA TCT CCA GAG GCA GTT GGG GTG ACC AGT CAG CGA CCA GTG TTT TGT CCC TTC CAT AAA AAG GAG CAG TTG AAA CTT TAC TGT GAA
  216  Thr Val Arg Gln Lys Glu Glu Val Ser Pro Glu Ala Val Gly Val Thr Ser Gln Arg Pro Val Phe Cys Pro Phe His Lys Lys Glu Gln Leu Lys Leu Tyr Cys Glu

  756  ACA TGT GAT AAA CTG ACC TGT CGA GAC TGC CAG CTG CTA GAA CAC AAA GAA CAC AGG TAT CAA TTT ATA GAA GAA GCT TTT CAG AAT CAA AAA GTG ATC ATA GAT ACT
  252  Thr Cys Asp Lys Leu Thr Cys Arg Asp Cys Gln Leu Leu Glu His Lys Glu His Arg Tyr Gln Phe Ile Glu Glu Ala Phe Gln Asn Gln Lys Val Ile Ile Asp Thr

  864  CTA ATC ACC AAA CTG ATG GAA AAA ACA AAA TAT ATA AAG TAT ACA GGA AAT CAG ATC CAA AAT AGG ATA ATT GAA ATA AAT CAA AAC CAA AAG CAG GTG GAA CAG GAT
  288  Leu Ile Thr Lys Leu MET Glu Lys Thr Lys Tyr Ile Lys Tyr Thr Gly Asn Gln Ile Gln Asn Arg Ile Ile Glu Ile Asn Gln Asn Gln Lys Gln Val Glu Gln Asp
                                                                                                                  ⌐ B-raf ⟶
  972  ATT AAA GTT GCC ATC TTC ACA TTG ATG GTG GAG ATA AAC AAA AAA GGG AAA GCT CTG CTG CAC CAG CTT GAG AAA ACA CTT GGT AGA AGA GAT TCA AGT GAT GAC TGG
  324  Ile Lys Val Ala Ile Phe Thr Leu MET Val Glu Ile Asn Lys Lys Gly Lys Ala Leu Leu His Gln Leu Glu Lys Thr Leu Gly Arg Arg Asp Ser Ser Asp Asp Trp

 1080  GAG ATT CCT GAT GGA CAG ATT ACA GTG GGA CAG AGA ATT GGA TCT GGG TCA TTT GGA ACT GTC TAC AAG GGA AAG TGG CAT GGT GAT GTG GCA GTG AAA ATG TTG AAT
  360  Glu Ile Pro Asp Gly Gln Ile Thr Val Gly Gln Arg Ile Gly Ser Gly Ser Phe Gly Thr Val Tyr Lys Gly Lys Trp His Gly Asp Val Ala Val Lys MET Leu Asn

 1188  GTG ACA GCA CCC ACA CCT CAA CAG CTA CAG GCC TTC AAA AAT GAA GTA GGA GTG CTC AGG AAA ACT CGA CAT GTG AAT ATC CTC CTT TTC ATG GGC TAT TCT ACA AAG
  396  Val Thr Ala Pro Thr Pro Gln Gln Leu Gln Ala Phe Lys Asn Glu Val Gly Val Leu Arg Lys Thr Arg His Val Asn Ile Leu Leu Phe MET Gly Tyr Ser Thr Lys

 1296  CCA CAA CTG GCA ATT GTT ACA CAG TGG TGT GAG GGC TCC AGC TTA TAT CAC CAT CTC CAC ATC ATT GAG ACC AAA TTT GAG ATG ATC AAA CTT ATA GAT ATT GCT CGG
  432  Pro Gln Leu Ala Ile Val Thr Gln Trp Cys Glu Gly Ser Ser Leu Tyr His His Leu His Ile Ile Glu Thr Lys Phe Glu MET Ile Lys Leu Ile Asp Ile Ala Arg

 1404  CAG ACT GCA CAG GGC ATG GAT TAC TTA CAC GCC AAG TCA ATC ATC CAC AGA GAC CTC AAG AGT AAT AAT ATA TTT CTT CAT GAA GAC CTC ACG GTA AAA ATA GGT GAC
  468  Gln Thr Ala Gln Gly MET Asp Tyr Leu His Ala Lys Ser Ile Ile His Arg Asp Leu Lys Ser Asn Asn Ile Phe Leu His Glu Asp Leu Thr Val Lys Ile Gly Asp

 1512  TTT GGT CTA GCC ACA GTG AAA TCT CGG TGG AGT GGG TCC CAT CAG TTT GAA CAG TTG TCT GGA TCT ATT TTG TGG ATG GCA CCA GAA GTA ATC AGA ATG CAA GAT AAA
  504  Phe Gly Leu Ala Thr Val Lys Ser Arg Trp Ser Gly Ser His Gln Phe Glu Gln Leu Ser Gly Ser Ile Leu Trp MET Ala Pro Glu Val Ile Arg MET Gln Asp Lys

 1620  AAC CCG TAT AGC TTT CAG TCA GAT GTG TAT GCG TTT GGG ATT GTT CTG TAC GAA CTG ATG ACC GGC CAG CTA CCT TAT TCA AAC ATC AAC AAC AGG GAT CAG ATA ATT
  540  Asn Pro Tyr Ser Phe Gln Ser Asp Val Tyr Ala Phe Gly Ile Val Leu Tyr Glu Leu MET Thr Gly Gln Leu Pro Tyr Ser Asn Ile Asn Asn Arg Asp Gln Ile Ile

 1728  TTT ATG GTG GGA CGA GGA TAC CTA TCT CCA GAT CTC AGT AAG GTA CGG AGT AAC TGT CCA AAA GCC ATG AAG AGA TTA ATG GCA GAG TGC CTC AAA AAG AAA AGA GAC
  576  Phe MET Val Gly Arg Gly Tyr Leu Ser Pro Asp Leu Ser Lys Val Arg Ser Asn Cys Pro Lys Ala MET Lys Arg Leu MET Ala Glu Cys Leu Lys Lys Lys Arg Asp

 1836  GAG AGA CCA CTC TTT CCC CAA ATT CTC GCC TCC ATT GAG CTG CTG GCC CGC TCA TTG CCA AAA ATT CAC CGC AGT GCA TCA GAA CCT TCC TTG AAT CGG GCT GGT TTC
  612  Glu Arg Pro Leu Phe Pro Gln Ile Leu Ala Ser Ile Glu Leu Leu Ala Arg Ser Leu Pro Lys Ile His Arg Ser Ala Ser Glu Pro Ser Leu Asn Arg Ala Gly Phe

 1944  CAA ACA GAA GAT TTT AGT CTG TAT GCT TGT GCT TCT CCG AAA ACA CCC ATC CAA GCA GGG GGA TAT GGT GGG TTT CCA GTC CAC TGA AACACCCTGAGTGAGACATTAAAGTAA
  648  Gln Thr Glu Asp Phe Ser Leu Tyr Ala Cys Ala Ser Pro Lys Thr Pro Ile Gln Ala Gly Gly Tyr Gly Gly Phe Pro Val His
                                                                                                                  Ala

 2059  GGTAGCAACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

FIG. 4. Nucleotide and deduced amino acid sequences of the T18 oncogene cDNA. (*A*) Restriction map of the transforming cDNA insert rescued from the T18-B transformant. Locations of the DNA probes used are shown below map. M-MLV LTR, Moloney murine leukemia virus long terminal repeat. (*B*) Nucleotide sequence of the entire T18-B cDNA insert. Nucleotides 1–13 are from the *Sfi* I adaptor sequence. The presumable initiation codon is underlined. The function between the unknown and B-*raf* genes is indicated. Amino acid residue differing from that of human B-*RAF* product is shown below residue. Residues that are substituted by the *env* sequence in R-*mil* are underlined.

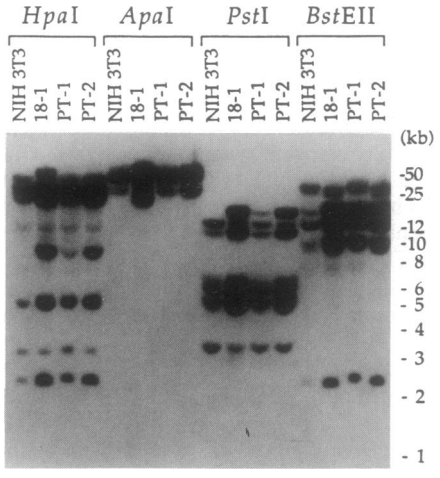


FIG. 5. Rearrangement and amplification of the T18 oncogene in the primary and secondary transformants. Sources of DNA and restriction enzymes used are shown at top. PT-1 and PT-2 are primary transformants induced by original T18 tumor DNA. 18-1 is a second-cycle transformant induced by PT-2 DNA and was the source of the cDNA library. Genomic DNA (5 μg) was digested by *Sal* I to release the cDNA inserts, separated on a 0.5% agarose gel by electrophoresis, and transferred to nitrocellulose membrane. The Southern blot was probed by the cDNA insert rescued from T18-B. NIH 3T3 genomic DNA was used as control.

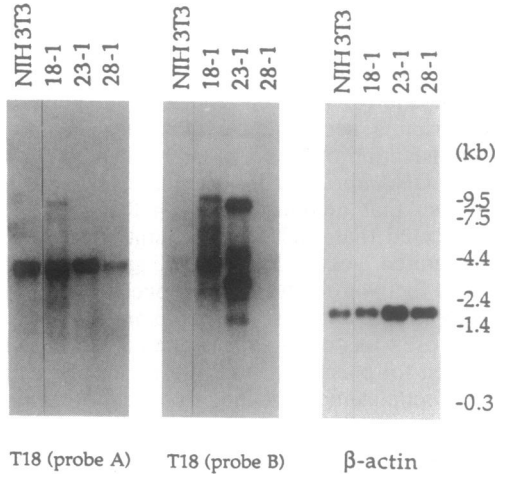


FIG. 6. Detection of mRNAs of the T18 oncogene. Poly(A)$^+$ RNAs extracted from the cell lines indicated were denatured and separated on a formaldehyde gel. 18-1 is a second-cycle transformant derived from T18 tumor. Primary transformants 23-1 and 28-1 were induced by T23 and T28 tumor DNAs, respectively. RNA was transferred to a nitrocellulose membrane for analysis with each probe. The 5′ probe was isolated as the *Sal* I–*Hin*dIII fragment (Fig. 4*A*). The 3′ probe was prepared by PCR (19) with the B-*raf* primer (5′-CCTCGAGATTCAAGTGATGAC-3′) and the T7 primer (5′-CTAATACGACTCACTATAGGGG-3′).

plasmids was constructed (Fig. 4*A*), and the clone was subjected to sequence analysis. The cDNA clone contained a large open reading frame starting from the 5′-end (Fig. 4*B*). The first methionine codon at positions 51–53 is likely to be the translation start site because it matched well the consensus sequence (12). Larger cDNAs (Fig. 3) may contain other initiation sites. No poly(A) signal was found near the 3′ end, suggesting that the cDNA was synthesized by internal priming. A database search indicated that the 5′ portion of the cDNA contained an unknown sequence, whereas the 3′ region was closely related to human B-*RAF* (13) and avian R-*mil* (14) genes (Fig. 4). To determine the breakpoint, we compared the T18 nucleotide sequence with that of proto-B-*raf* (T. Yamamoto, personal communication) and v-R-*mil* (14). There was no homology with either sequence upstream from position 1044 in the T18 oncogene. Thus, position 1044 represents the junction between an unknown sequence and the B-*raf* gene. R-*mil* is a viral oncogene and encodes a *gag*–R-*mil*–*env* fusion protein. The junction of *gag* and R-*mil* has been mapped 144 nucleotides upstream from the T18 break point (14), whereas there was a junction of a different sequence and the human B-*RAF* oncogene (T. Yamamoto, personal communication) 174 nucleotides upstream from the junction in the T18 oncogene. In each of the B-*raf* oncogenes, including T18, the breakpoints did not disrupt the predicted kinase domain of the protein. Comparison of the predicted amino acid sequences with that of B-*raf* indicated identity, except for a single amino acid difference at position 671, in which glycine was substituted for alanine in human B-*RAF* (Fig. 4*B*). There was also complete identity with avian R-*mil*, except for a small stretch of nine amino acids at the R-*mil* COOH terminus (Fig. 4*B*), where recombination with an avian retroviral *env* gene caused this substitution.

Evidence for *in Vivo* Oncogene Activation. The human B-*RAF* gene product is ≈84% related to the amino acid sequence of the c-*raf* oncogene (13); another member of the *raf* family, A-*raf*, is also structurally similar (15). Most *raf* oncogenes have been activated by mechanisms involving structural rearrangements due to recombination and loss of NH_2-terminal sequences of the *raf*-coding sequence (16). Moreover, most reported instances have involved *in vitro* activation of these genes during DNA transfection, rather than by mechanisms leading to oncogene activation within the tumor itself (13, 17, 18). Thus, the rearrangement activating the T18 oncogene might have occurred during the course of cDNA library construction or as an artifact of DNA transfection with the original-tumor DNA. Alternatively, the oncogene might have been activated within the original tumor itself.

Although original-tumor DNA was not available, two primary transformants that had been independently induced by this tumor DNA could be analyzed. Fig. 5 shows that rearrangement as well as amplification of both non-B-*raf* and related B-*raf* portions of the gene were found not only in the secondary transfectant, which was the source of the cDNA library, but in both primary transformants, PT-1 and PT-2. Because such *in vitro* rearrangements are very rare, these findings strongly argue that the oncogene was activated *in vivo* in the hepatoma as part of the neoplastic process.

Evidence for Activation of Other B-*raf* Oncogenes. In an effort to characterize its transcript and search for evidence of additional B-*raf* oncogenes among the three other hepatoma transforming genes, we subjected poly(A)-selected RNAs from primary or secondary NIH 3T3 transfectants containing each oncogene to analysis with DNA probes from 5′ (B-*raf*-unrelated) and 3′ (B-*raf*-related) portions of the T18 oncogene. Fig. 6 shows that control NIH 3T3 cells contained a 4.2-kb RNA that hybridized with the 5′ non-B-*raf*-related portion of the oncogene, but there was no detectable B-*raf* transcript. In contrast, the second-cycle T18 transfectant, which was the source of the expression cDNA library, showed a major 4.2-kb as well as minor 10- and 3-kb transcripts, which appeared to hybridize with both probes. Fig. 6 further shows that a primary T23 oncogene transfectant contained multiple B-*raf* hybridizing transcripts, indicating that it was induced by B-*raf* oncogene as well. Of note, the multiple transcripts detected differed in sizes from those of the T18 oncogene. Moreover, no abnormal transcript was detected by the B-*raf*-unrelated probe derived from the T18 oncogene (Fig. 6). Thus, if the T23 oncogene arose by a mechanism involving B-*raf* gene rearrangement, this rearrangement was different from that associated with activation

of the T18 oncogene. A transfectant induced by the T28 oncogene did not show abnormal B-*raf*-hybridizing RNAs (Fig. 6), arguing that this oncogene must be distinct from B-*raf*.

DISCUSSION

The present study describes the development of a stable mammalian expression cloning system and its successful application to the molecular cloning of oncogenes. Although mammalian cell expression has been used to isolate cDNAs for cell-surface antigens (20), growth factors (21), and their receptors (22), these systems have often been based on transient expression from the SV40 promoter in COS cells. In the COS cell system (20, 23), expression is very efficient due to replication of introduced vectors from SV40 origin of replication. However, selection of cells showing a specific phenotype is generally practical only by means of stable expression. Efficient synthesis of full-length cDNAs, unidirectional insertion of cDNA fragments into vectors, and high-level expression of cDNA inserts are all required to establish such a system. Each of these features was designed into the λpCEV27 vector.

The capacity for efficient rescue of cDNA clones from mammalian cells is another important function of a stable expression cloning system. When plasmid cDNA libraries are used to transfect mammalian cells, single plasmids integrated in genomic DNA are difficult to release. Plasmid rescue is readily achieved only when multiple copies are clustered at a single integration site (24, 25). Excision of the plasmid by induction of replication from the SV40 origin using COS cell fusion often results in rearrangement or truncation of cDNA inserts (26). To rescue cDNA clones efficiently from stable integration sites within mammalian host cells, we used a strategy involving λ–plasmid composite vectors. The presence of sites for three infrequent cutters at the plasmid–λ junctions facilitated the plasmid-rescue step and minimized possibility of cleavage within the cDNA insert.

The largest oncogene cDNA insert identified was 7 kb, indicating that very large cDNAs can be isolated by this technique. We estimate that the frequency at which the oncogene was represented in the library was ≈1 per 10^4–10^5 cDNAs. The oncogene was identified as the mouse homologue of human B-*RAF*, a recently described third member of the *raf* protooncogene family. A human B-*RAF* oncogene arose as an artifact of DNA transfection (13), and the other occurred as a result of retroviral transduction of the avian oncogene, R-*mil* (14). We showed that at least two independent primary transfectants of the original tumor DNA contained the identical rearrangement involving B-*raf* and the same unknown sequence. Thus, this B-*raf* oncogene likely originated as part of the neoplastic process.

Sequence analysis revealed that the mechanism of B-*raf* activation was due to genetic rearrangement involving its 3′ kinase domain with an unknown upstream sequence. This rearrangement resulted in NH_2-terminal truncation of the B-*raf* product, consistent with analogous rearrangements truncating the human B-*RAF* and avian R-*mil* products (13, 14). With c-*raf*, NH_2-terminal truncation at a similar position has been demonstrated to activate its inherent serine/threonine kinase activity (27). Thus, their NH_2-terminal domains probably play important roles in regulating the catalytic domains of *raf* family members. Removal of the inhibitory influence of the upstream domain by genetic rearrangements, such as the one described in our study, represents the likely mechanism responsible for activation of transforming function. However, further studies will be required to directly establish the validity of this hypothesis.

The ability to use this stable expression cloning approach to rapidly isolate and characterize oncogenes associated with rodent models should aid in studying mechanisms of carcinogenic action. More importantly, our demonstration of the feasibility of this stable-expression cDNA cloning approach suggests its wide applicability to the isolation of other molecules of biologic interest. This applicability could include growth factors or receptors where introduction into a suitable cell might create an autocrine transforming loop and allow for selection of cells with a growth advantage (28). The same strategy might also lead to isolation of other limiting molecules in mitogenic signaling pathways or any cDNA the stable expression of which confers a phenotype that allows for its selection.

We thank Tadashi Yamamoto (Tokyo) for unpublished observations; Cheryl Smith, Jennifer Artrip, and Marthanna Moore for excellent technical support; and Steven R. Tronick for computer analysis.

1. Miki, T., Matsui, T., Heidaran, M. A. & Aaronson, S. A. (1989) *Gene* **83**, 137–146.
2. Matsui, T., Heidaran, M. A., Miki, T., Popescu, N., La Rochelle, W., Kraus, M., Pierce, J. & Aaronson, S. A. (1989) *Science* **243**, 800–804.
3. Finch, P. W., Rubin, J. S., Miki, T., Ron, D. & Aaronson, S. A. (1989) *Science* **245**, 752–755.
4. Kraus, M. H., Issing, W., Miki, T., Popescu, N. C. & Aaronson, S. A. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 9193–9197.
5. Kruh, G. D., Perego, R., Miki, T. & Aaronson, S. A. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 5802–5806.
6. Rubin, J. S., Chan, A. M.-L., Bottaro, D. P., Burgess, W. H., Taylor, W. G., Cech, A. C., Hirschfield, D. W., Wong, J., Miki, T., Finch, P. W. & Aaronson, S. A. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 415–419.
7. Reynolds, S. H., Stowers, S. J., Patterson, R. M., Maronpot, R. R., Aaronson, S. A. & Anderson, M. W. (1987) *Science* **237**, 1309–1316.
8. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) in *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
9. Jainchill, J. L., Aaronson, S. A. & Todaro, G. J. (1969) *J. Virol.* **4**, 549–553.
10. Wigler, M., Silverstein, S., Lee, L.-S., Pellicer, A., Cheng, Y. C. & Axel, R. (1977) *Cell* **11**, 223–232.
11. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
12. Kozak, M. (1987) *Nucleic Acids Res.* **15**, 8125–8148.
13. Ikawa, S., Fukui, M., Ueyama, Y., Tamaoki, N., Yamamoto, T. & Toyoshima, K. (1988) *Mol. Cell. Biol.* **8**, 2651–2654.
14. Marx, M., Eychene, A., Laugier, D., Bechade, C., Crisanti, P., Dezelee, P., Pessac, B. & Calothy, G. (1988) *EMBO J.* **7**, 3369–3373.
15. Huleihel, M., Goldsborough, M., Cleveland, J. L., Gunnell, M., Bonner, T. & Rapp, U. R. (1986) *Mol. Cell. Biol.* **6**, 2655–2662.
16. Rapp, U. R., Cleveland, J. L., Bonner, T. I. & Storm, S. M. (1988) in *The Oncogene Handbook*, ed. Reddy, E. P. (Elsevier, Amsterdam), pp. 211–251.
17. Ishikawa, K., Takaku, F., Hayashi, K., Nagao, M. & Sugimura, T. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 3209–3212.
18. Stanton, V. P., Jr., & Cooper, G. M. (1987) *Mol. Cell. Biol.* **7**, 1171–1179.
19. Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A. & Arnheim, N. (1985) *Science* **230**, 1350–1354.
20. Aruffo, A. & Seed, B. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 8573–8577.
21. Lee, F., Yokota, T., Otsuka, T., Gemmell, L., Larson, N., Luh, J., Arai, K.-I. & Rennick, D. (1985) *Proc. Natl. Acad. Sci. USA* **85**, 4360–4364.
22. D'Andrea, A. P., Lodish, H. F. & Wong, G. G. (1989) *Cell* **57**, 277–285.
23. Okayama, H. & Berg, P. (1983) *Mol. Cell. Biol.* **3**, 280–289.
24. Noda, M., Kitayama, H., Matsuzaki, T., Sugimoto, Y., Okayama, H., Bassin, R. & Ikawa, Y. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 162–166.
25. Grant, S. G. N., Jessee, J., Bloom, F. R. & Hanahan, D. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 4645–4649.
26. Munro, S. & Maniatis, T. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 9248–9252.
27. Moelling, K., Heimann, B., Beimling, P., Rapp, U. R. & Sander, T. (1984) *Nature (London)* **312**, 558–561.
28. Miki, T., Fleming, T. P., Bottaro, D. P., Rubin, J. S., Ron, D. & Aaronson, S. A. (1991) *Science* **251**, 72–75.