MOLECULAR AND CELLULAR BIOLOGY, Aug. 1991, p. 4088–4096
0270-7306/91/084088-09$02.00/0

Vol. 11, No. 8

Structure and Transforming Potential of the Human *cot* Oncogene Encoding a Putative Protein Kinase†

JUN MIYOSHI,* TAKATSUGU HIGASHI, HIROYUKI MUKAI, TOHRU OHUCHI, AND TAKEO KAKUNAGA

Department of Oncogene Research, Research Institute for Microbial Diseases, Osaka University, Osaka 565, Japan

Received 5 September 1989/Accepted 7 May 1991

A new transforming gene has been molecularly cloned from hamster SHOK cells transformed with DNA extracted from a human thyroid carcinoma cell line and named the *cot* (cancer Osaka thyroid) oncogene. cDNA sequencing disclosed that this oncogene codes for a protein with 415 amino acid residues, and computer matching showed 42 to 48% similarity matches with serine protein kinases. Its gene product was identified as a 52-kDa protein by transcription and translation in vitro. Expression of *cot* cDNA under transcriptional control by a retroviral long terminal repeat induced morphological transformation of NIH 3T3 cells as well as SHOK cells. Protein kinase activity associated with constructed $p60^{gag\text{-}cot}$ was detected by immune complex kinase assay with anti-*gag* antiserum. The *cot* oncogene was overexpressed in transformed SHOK cells and found to have a rearranged 3′ end in the last coding exon, which probably resulted in a deletion and an altered C′ terminus in the transforming protein. This DNA rearrangement appeared to have occurred during transfection of the tumor DNA into hamster SHOK cells and not in the original thyroid tumor.

A variety of oncogenes have been detected by transfection of NIH 3T3 mouse fibroblasts with genomic tumor DNAs (5, 6, 9, 28, 30, 31). About 80% of human tumor DNAs, however, are ineffective in inducing transformed foci of NIH 3T3 cells, and activated alleles of the c-*ras* gene family have been repeatedly detected by this standard approach. Thus, the focus-forming assay with NIH 3T3 cells may not be effective for detecting certain classes of oncogenes.

Several variations of the assay have been used in attempts to detect other oncogenes, including the use of other recipient cell lines (3, 12, 17, 33), assays for tumor formation in nude mice rather than transformation in cell culture (39), and even attempts to activate DNA from normal cells by adding strong promoter elements before DNA transfer. We have reported that a hamster embryonic cell line, SHOK (Syrian hamster Osaka Kanazawa), can be used as a good alternative to NIH 3T3 cells (14). SHOK cells appear to be more sensitive than NIH 3T3 cells to transformation by *mos*, as sensitive as NIH 3T3 cells to mutated *ras* genes, and less sensitive than NIH 3T3 cells to oncogenes of the tyrosine protein kinase family. SHOK cells are particularly useful for molecular cloning of mouse cellular oncogenes, because the mouse repetitive sequence serves as an effective genetic marker for identifying DNA sequences derived from mouse tumors under the background of hamster cellular DNA.

We have used the SHOK assay to survey a number of human and mouse tumor DNAs to detect potential oncogenes and have found some intriguing cases, such as alteration of N-*ras* gene mutation during the clinical course of acute lymphoblastic leukemia (35), an activated K-*ras* with a new mutation in codon 146 in the human HUT14 cell line, a rearranged *mos* from mouse osteosarcoma, and a novel oncogene from a human thyroid carcinoma cell line, TCO4. Here we report molecular cloning of this new oncogene and characterization of its structure and function. Our findings provide further evidence for the value of SHOK cells in identifying new classes of oncogenes.

MATERIALS AND METHODS

DNA transfection assay of SHOK and NIH 3T3 cells. SHOK cells were cultured in Dulbecco's modified Eagle's minimum essential medium (DMEM) supplemented with 10% fetal calf serum. For the focus-forming assay, SHOK cells were seeded at a density of 3×10^5 cells per 100-mm-diameter dish, transfected with 30 μg of tumor DNAs or 0.25 to 1 μg of recombinant plasmids per dish by the calcium phosphate precipitation method (10, 38), and maintained in DMEM supplemented with 3% fetal calf serum. Transformed foci were scored 14 to 21 days after DNA transfection. Neomycin-resistant colonies were selected with DMEM and G418 (400 μg/ml). NIH 3T3 cells were cultured in DMEM with 10% calf serum, maintained in DMEM with 5% calf serum for focus-forming assays, and selected with G418 (400 μg/ml).

Construction and screening of genomic and cDNA libraries. Genomic libraries were generated in λ2001 or cosmid pHSG274 vectors as described by Karn et al. (18) and Brady et al. (4). About 700,000 PFU of recombinant phage and 300,000 colonies of recombinant cosmid were screened with a human *Alu* sequence probe under standard conditions (21).

Two cDNA libraries were constructed with a λgt10 vector and poly(A)$^+$ RNA from transformed cells. The cDNA synthesis was primed in two ways, with oligo(dT) and with random oligonucleotides, and was performed with an Amersham kit as recommended by the manufacturer. After methylation of the cDNA with *Eco*RI methylase and the addition of *Eco*RI linkers, the DNA was digested with *Eco*RI and excess linkers were removed on a Bio-Gel A-50m column. The cDNA was then ligated to *Eco*RI-digested and dephosphorylated λgt10 arms. The ligated cDNA was packaged with Gigapak extracts (Stratagene Cloning System; Stratagene, San Diego, Calif.) and plated with C600 HFl$^+$ as a host

* Corresponding author.

† This work is dedicated to Takeo Kakunaga, who died on 21 September 1988.

bacterial strain. Phage DNA was transferred to duplicate nitrocellulose filters and hybridized to nick-translated ^{32}P-labeled probes. Approximately 1,000,000 phages from the oligo(dT)-primed library and 500,000 phages from the random-primed library were screened.

DNA and RNA blotting analyses. High-molecular-weight DNAs were prepared from human placenta and cultured cells, digested with restriction enzymes, size fractionated on a 0.4 or 0.7% agarose gel, and transferred to a nitrocellulose membrane. Northern blot analysis was performed by fractionating total RNA or poly(A)$^+$ RNA on a horizontal 1% agarose formaldehyde gel, and the samples were transferred to a nitrocellulose membrane. The filters were prehybridized for 2 h at 65°C in a solution containing 5× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate), 5× Denhardt's solution, 0.1% sodium dodecyl sulfate (SDS), and 100 μg of sonicated calf thymus DNA per ml. For hybridizations, filters were incubated overnight at 65°C with radio-labeled probes. After hybridization, the filters were washed twice for 5 min at room temperature and twice for 30 min at 65°C in 0.1% SDS–0.2× SSC.

DNA sequence and computer analysis. Appropriate restriction enzyme fragments from overlapping cDNA clones were subcloned into pUC18 or pUC19 and were sequenced by the dideoxy chain termination method (29). DNA sequences were analyzed with University of Wisconsin Genetics Computer Group programs.

Production and in vitro translation of SP6 *cot* RNA. The pSP65 recombinant plasmids were linearized with *Sac*I, *Bam*HI, and *Hin*dIII and used as templates for in vitro transcription of *cot* RNA (24). SP6 *cot* RNA was translated in vitro in a rabbit reticulocyte lysate (Amersham). The reaction mixture for translation (50 μl) contained 40 μl of lysate, 0.75 MBq of [^{35}S]methionine (>37 TBq/mmol; Amersham), and 1 μg of SP6 *cot* RNA. Translation products were separated on an SDS–10% polyacrylamide gel and located by fluorography.

Construction of expression plasmids carrying the *cot* oncogene. The plasmids expressing p52^{cot} (pJJ26 and pJF1) and p60$^{gag\text{-}cot}$ (pJA25) were constructed by substituting the *cot* cDNA for the v-*fgr* oncogene in Gardner-Rasheed feline sarcoma virus (GR-FeSV) (26). Vector sequences were derived from pSV2*neo* and *Eco*RI-linker insertion mutants of GR-FeSV as previously described (25). We manipulated two independent insertion mutants in the 5′ and 3′ untranslated regions of v-*fgr* by *Eco*RI digestion and religation to remove the whole v-*fgr* sequence, and the construct was used for pJJ26 and pJF1. In a similar way, the insertion mutant *in354*, containing a 12-bp *Eco*RI linker at the *Hae*III site located in the *gag*-actin junction, and the mutant in the 3′ untranslated region as described were combined to restore only the *gag* part of v-*fgr* and an in-frame *Eco*RI site to be linked to the *cot* cDNA sequence, and the combined construct was used for constructing pJA25. These modified parts of GR-FeSV containing a single *Eco*RI site were excised at the *Sac*I site located in two long terminal repeats (LTRs) and subcloned into a modified version of pUC13 that lacked an *Eco*RI site. Then we constructed the expression plasmids in two steps, ligating the *Eco*RI fragments bearing the *cot* cDNA to these intermediates and putting the modified *Sac*I-*Sac*I fragments back in the right place between LTRs of GR-FeSV in pSV2*neo*GR-FeSV. The *cot* cDNA sequence of pJJ26 was derived from a chimeric construct of the cDNA clones *a* and *e* (Fig. 1B) at the *Kpn*I site, and the sequences of pJF1 and pJA25 were derived from clones *g* and *h*, respectively. p52^{cot} is encoded by the entire open reading frame of the *cot* cDNA, whereas p60$^{gag\text{-}cot}$ consists of three parts. The first 118 amino acids of p60$^{gag\text{-}cot}$ correspond to p15^{gag}; the second group, amino acids 119 to 122 between the *gag* and *cot* domains, is encoded by the linker sequence GCTGGAATTCCT; and the third group, amino acids 123 to 527, is derived from a *cot*-coding region beginning with Lys-11 and extending to Cys-415.

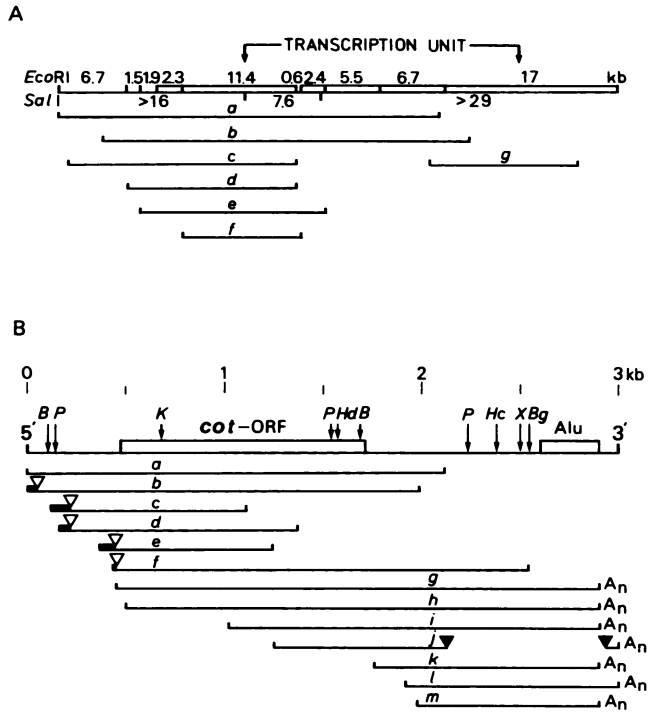

FIG. 1. Restriction map of genomic DNA and cDNA clones of the *cot* oncogene in transformed SHOK cells. (A) Structure of *cot* genomic DNA clones. The 56 kb of DNA surrounding the *cot* oncogene locus is indicated. The open boxes represent human *Alu*-positive *Eco*RI fragments. Clones *a* and *b* were recombinant cosmids, and clones *c* through *g* were recombinant phages. (B) Structure of *cot* cDNA clones. The upstream open box indicates the 1,245-bp open reading frame of the *cot* oncogene, and the downstream open box indicates one copy of the *Alu* sequence present in the 3′ untranslated region. Clones *a* through *f* were isolated from a random oligonucleotide-primed cDNA library. Clones *g* through *m* were isolated from an oligo(dT)-primed cDNA library. Three splicing sites, indicated by open triangles, are present in the 5′ upstream sequence, and alternative exons are linked as shown by closed boxes. Closed triangles indicate the splicing sites which were detected in the 3′ untranslated region of clone *j*. Restriction sites are as follows: *K*, *Kpn*I; *P*, *Pst*I; *B*, *Bam*HI; *Bg*, *Bgl*II; *Hd*, *Hin*dIII; *Hc*, *Hin*cII; *X*, *Xba*I.

RESULTS

Molecular cloning of a new transforming gene. On the basis of a survey of 27 tumor DNAs for transforming activities on SHOK cells, DNA derived from a human thyroid carcinoma cell line, TCO4, gave one characteristic focus with bizarre morphology. The presence of a new transforming gene, named *cot* (cancer Osaka thyroid), was repeatedly detected by Southern blot analysis with a human repetitive *Alu* sequence probe and by lack of hybridization to other known oncogenes during the serial process of DNA transfection.

We constructed genomic DNA libraries of transformed

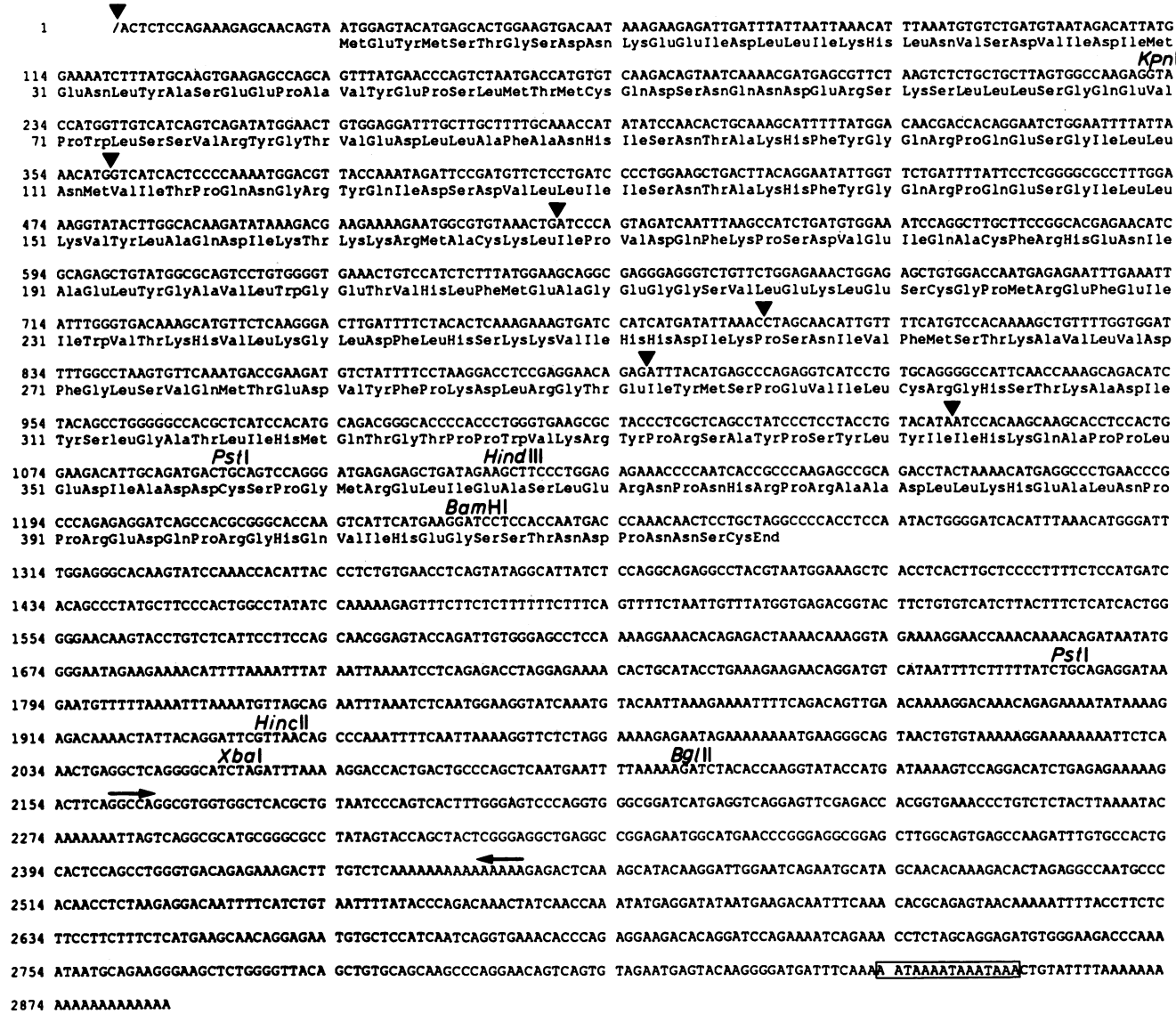

FIG. 2. Nucleotide sequence and predicted amino acid sequence of the *cot* cDNA. Nucleotides start at the cDNA region corresponding to the beginning of exon 3 of the *cot* oncogene. The nucleotide sequences of two noncoding exons, exons 1 and 2, are not shown. Amino acids are numbered in the lower line from the first methionine of the open reading frame. The DNA sequence between the two arrows is that of the *Alu* sequence. Splicing sites in the coding region are shown by closed triangles. The boxed sequence is that of three overlapping recognition sites for polyadenylation of *cot* mRNA (8).

SHOK cells and isolated several overlapping recombinant cosmid and phage clones that strongly hybridized to the *Alu* sequence and appeared to contain the *cot* oncogene. The restriction map of genomic clones is shown in Fig. 1A. None of these clones transformed SHOK cells because they lacked parts of the transforming sequence.

In parallel experiments, two λgt10 cDNA libraries constructed from mRNA of transformed SHOK cells were screened with *Alu*-negative genomic DNA probes and clone *g* was obtained. By using the cDNA insert of this 2.4-kb clone as a probe, additional clones were obtained as shown in Fig. 2B. The entire overlapping sequence in the cDNA clones spanned about 3 kb and hybridized with a cluster of all the *Alu*-positive genomic DNA fragments that were cosegregated with the transformed phenotype after three cycles of transfection with genomic DNA, suggesting that the whole sequence of *cot* mRNA was represented in the cDNA clones recovered. Clones *b* through *f* have sequences at the 5′ end that are unrelated to the sequence at the 5′ end of clone *a* (Fig. 1, closed boxes). These different 5′ ends may arise by alternate splicing, since they are joined to the body of the mRNA at the splice acceptors for exons 1 through 3. However, these sequences have not yet been mapped to the *cot* locus. A less likely possibility is that these 5′ ends resulted from rearrangements during cloning. If the 5′ end of the *cot* gene is alternately spliced, its significance is unclear, since with the exception of clone *e*, which has an extra 34 amino acids, the other clones represent mRNAs that should all encode identical proteins.

Nucleotide sequence of the *cot* cDNA and predicted protein

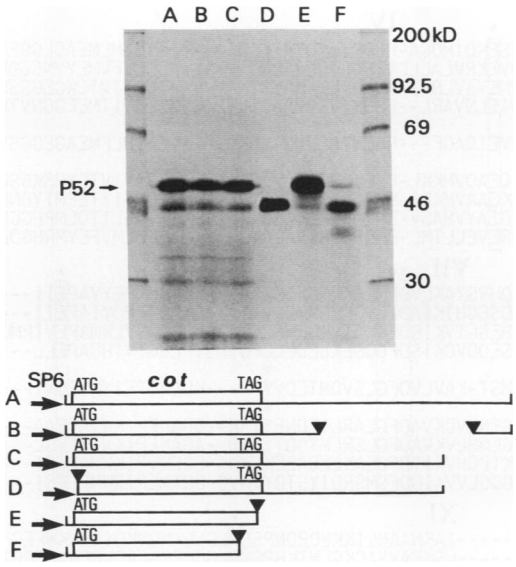

FIG. 3. Proteins synthesized in vitro from SP6-*cot* RNAs. SP6-*cot* RNAs were translated in vitro in rabbit reticulocyte lysates, and translation products were analyzed by separation on an SDS–10% polyacrylamide gel and fluorography. The arrow represents p52, the major form of Cot protein translated from the first initiation AUG codon. The structures of the templates for RNA synthesis are indicated (A through F): A contains the *Alu* sequence in the 3′ untranslated region; B lacks the Alu sequence by RNA splicing; C has a deletion of 900 bp in the distal part of the 3′ untranslated region; D has a deletion of 30 bp (10 amino acid residues) including the first AUG codon in the 5′ end of the *cot*-coding sequence and would be translated from the second AUG; E and F were digested with *Bam*HI and *Hin*dIII at the 3′ end of the open reading frame before in vitro transcription.

structure. The entire cDNA sequence has an open reading frame capable of encoding a polypeptide backbone of 415 amino acids, with a predicted molecular mass of 48 kDa (nucleotides 24 to 1271 [Fig. 2]). The nucleotide sequence surrounding the first methionine codon at position 24 corresponds well to the consensus sequence ccA/GccATGG for a translation initiation site (19). The assignment was confirmed by computer analysis for nonrandom nucleotide composition (7) and for codon preference (11).

The presence of the predicted open reading frame was confirmed by analysis of in vitro protein products directed by *cot* cDNA clones. Two forms of p52 and p46 were detected as shown in Fig. 3. Since the cDNA clone lacking 30 bp at the 5′ end of the coding sequence (10 residues at the N terminus) generated only p46 (Fig. 3, lane D), we concluded that p52 and p46 are products of alternative initiation of translation at two in-frame AUG codons on the same mRNA. The ratio of p52 to p46 synthesized in rabbit reticulocyte lysates indicated that in vitro translation was initiated predominantly at the first AUG (nucleotide 24) rather than the second AUG (nucleotide 111 or 168). The 3′ end of the open reading frame was confirmed by linearization of the *cot* cDNA in the SP6 expression vector. Digestion of templates for RNA synthesis at a *Bam*HI site (nucleotide 1236) 33 bp upstream from a TAG stop codon and a *Hin*dIII site (nucleotide 1120) directed the syntheses of the 52-kDa protein (Fig. 3, lane E) and the truncated 48-kDa protein (Fig. 3, lane F), respectively.

The predicted amino acid sequence of the Cot protein was used to search the National Biomedical Research Foundation protein sequence data bank for similar sequences by using the program of Lipman and Pearson (20). The initial homology search showed a significant match with bovine cGMP-dependent protein kinase (34). This match (>45% similarity) covers the 271 carboxy-terminal amino acids of the two polypeptides but not their amino-terminal sequences. All protein kinases are known to have striking sequence similarity in their catalytic domains (16), and this is also the case with Cot. The amino acid sequence of Cot has a series of short sequence motifs that are highly conserved, including the ATP binding site, Gly(Pro)-X-Gly-X-X-Gly (residues 145 to 150) followed 17 residues downstream by a Lys-167, and a 60-amino-acid region containing the conserved sequences Arg(His)-Asp-Leu(Ile) (residues 252 to 254), Asp-Phe-Gly (residues 270 to 272), and Ala(Ser)-Pro-Glu (residues 295 to 297) (amino acids in parentheses are those of Cot).

The Cot protein was compared with other members of the protein kinase family (Fig. 4). The similarities of Cot to serine-threonine kinases such as C kinase and to the c-*mos* and c-*raf* proteins are 43, 42, and 48%, respectively, starting at part of the ATP-binding site (residues 145 to 415), whereas the similarity of Cot to the tyrosine kinase family is about 40 to 45%. There was no significant difference in overall similarities of Cot to these two kinase families, but segments in subdomains VI and IX showed strong similarity to those of serine protein kinases. In particular, the amino acid sequence between the invariant residues Asp-253 and Asn-258 in subdomain VI, which is reported to be the most striking indicator of amino acid specificity (13), was closely similar to those of serine protein kinases, suggesting that the Cot protein is a serine-threonine protein kinase.

Other noteworthy features of Cot were the Pro-145 and Tyr-282 residues. The Pro-145 residue instead of Gly as the first residue in the ATP-binding motif appeared to be a novel feature, because this has not been reported for other mammalian protein kinases although it is found in the herpes simplex virus US3 gene product HSVK (23). The Tyr-282 residue could correspond to the tyrosine frequently conserved as a major phosphorylation site in the tyrosine kinase family. Evaluation of these structural features must await the construction of specific antibodies and further characterization of Cot.

Transforming potential and protein kinase activity associated with chimeric constructs of the *cot* cDNA. Since no genomic DNA clones with transforming activity have been isolated, we tried to identify the transformed phenotypes associated with the expression of the *cot* cDNA. We constructed plasmids of *cot* cDNA under the control of GR-FeSV long terminal repeat with the *neo* gene as a selection marker as shown in Fig. 5A. The plasmids pJJ26 and pJF1 were directed to express the Cot protein from its own AUG codon, whereas pJA25 was constructed to express the Gag-Cot fusion protein. These constructs were transfected into SHOK and NIH 3T3 cells. The number of G418-resistant colonies, as an internal control, was used to verify that the efficiency of transfection was the same in all cases. As shown in Table 1, chimeric plasmids were capable of inducing morphological transformation of SHOK and NIH 3T3 cells at efficiencies comparable to those of known oncogenes. Morphological effects were evident after transfection of SHOK and NIH 3T3 cells with these constructs, as shown in Fig. 5C.

So far we had not succeeded in raising an antibody specific to Cot. We therefore tested for protein kinase activity

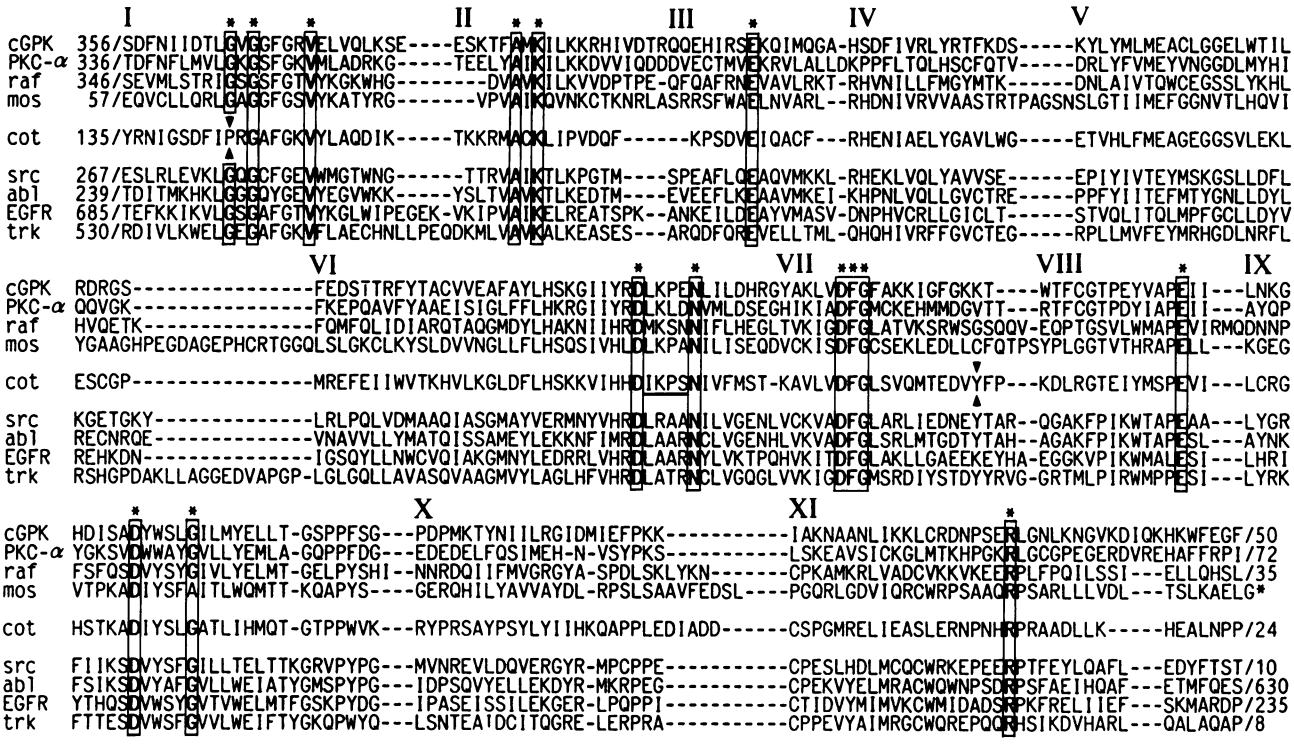

FIG. 4. Comparison of the Cot protein with selected members of the protein kinase gene family. The amino acid sequences of the protein kinase genes and subdomains I to XI were aligned as previously described (13). Dashes indicate alignment gaps. Fifteen strongly conserved amino acid residues in the protein kinases are boxed. Two amino acid residues of interest in Cot, Pro-145 and Tyr-282, are indicated by closed triangles. The sequence predicting the amino acid specificity of protein kinase is underlined. The amino acid sequences of the following proteins are presented: bovine cGMP-dependent protein kinase (cGPK) (34); bovine protein kinase C (PKC-α) (27); human c-*raf* (raf) (2); human c-*mos* (mos) (37); human c-*src* (scr) (1); human c-*abl* (abl) (32); EGF receptor (EGFR) (36); and *trk* (trk) (22).

associated with Gag-Cot by using anti-feline leukemia virus $p15^{gag}$ antibody. As shown in Fig. 5B, $p60^{gag\text{-}cot}$ was immunoprecipitated from lysates of transformed SHOK cells metabolically labeled with [^{35}S]methionine, and autophosphorylation of $p60^{gag\text{-}cot}$ was detected with anti-feline leukemia virus p15 serum in a standard retrovirus protein kinase assay. The protein kinase activity of $p60^{gag\text{-}cot}$ appeared to be more potent than that of $p70^{gag\text{-}actin\text{-}fgr}$, as judged from the ratio of the ^{35}S and ^{32}P radioactivities incorporated into the oncoproteins.

Elevated expression and DNA rearrangement of the *cot* oncogene in transformed SHOK cells. Information about transcripts and coding sequences of the *cot* oncogene and its normal counterpart was required to evaluate the contributions of overexpression, altered regulation, and structural mutation to the transforming activity. Transcripts of the *cot* oncogene in TCO4 and transformed SHOK cells were analyzed by RNA blotting. As shown in Fig. 6A, an approximately 3.7-kb transcript was readily detected in transformed SHOK cells but not in TCO4 cells. Similar results were obtained with poly(A)$^+$ RNA purified from TCO4 and SHOK transformants at each stage of serial transfection (data not shown). A high level of expression of the *cot* oncogene thus appeared to be involved, at least in part, in transformation of SHOK cells.

There was a possibility that the mutation in coding sequences of the *cot* oncogene might be involved in oncogenic activation. We therefore examined the genome organization of the *cot* locus in transformed cells, in the original tumor TCO4 cells, and in human placenta cells by Southern hybridization. A probe prepared from a 0.7-kb *Sau*3A restriction fragment in the 5′ upstream region hybridized to a 21-kb DNA fragment in *Eco*RI-digested TCO4 and human placental DNAs, whereas it hybridized to an 11.4-kb fragment in *Eco*RI-digested transformed SHOK DNAs (Fig. 6B). As the same probe hybridized to a 7.4-kb fragment in *Bgl*II-digested DNAs from all three sources (Fig. 6B), the *cot* oncogene in transformed SHOK cells was rearranged at the 5′ end within the 3-kb region between the *Eco*RI and *Bgl*II sites. By hybridizing *cot* cDNA probes to *Eco*RI-digested genomic DNAs, three fragments of 2.4, 6.7, and 7.5 kb were detected in human placenta and TCO4 DNAs, whereas the 7.5-kb fragment at the 3′ end was rearranged to a 17-kb fragment in transformed SHOK cells (Fig. 6C). These data indicate that the *cot* oncogene was rearranged at both ends in transformed SHOK cells but not in TCO4 cells or human placenta cells. We next confirmed these results by cloning the normal *cot* locus from human placenta cells and DNA sequencing of all known exons. Results showed that the *cot* oncogene contained eight exons consisting of two noncoding and six coding exons; that the 5′ end of the *cot* oncogene was rearranged upstream of the noncoding exon 1, whereas the 3′ rearrangement point was within the last exon, exon 8 (Fig. 7A); and that no coding mutation existed in exons other than exon 8. As shown in Fig. 7B, the *cot* DNA sequence downstream from the 3′ rearranged point was totally different from the *cot* proto-oncogene sequence and the following coding sequence of the *cot* oncogene was abruptly

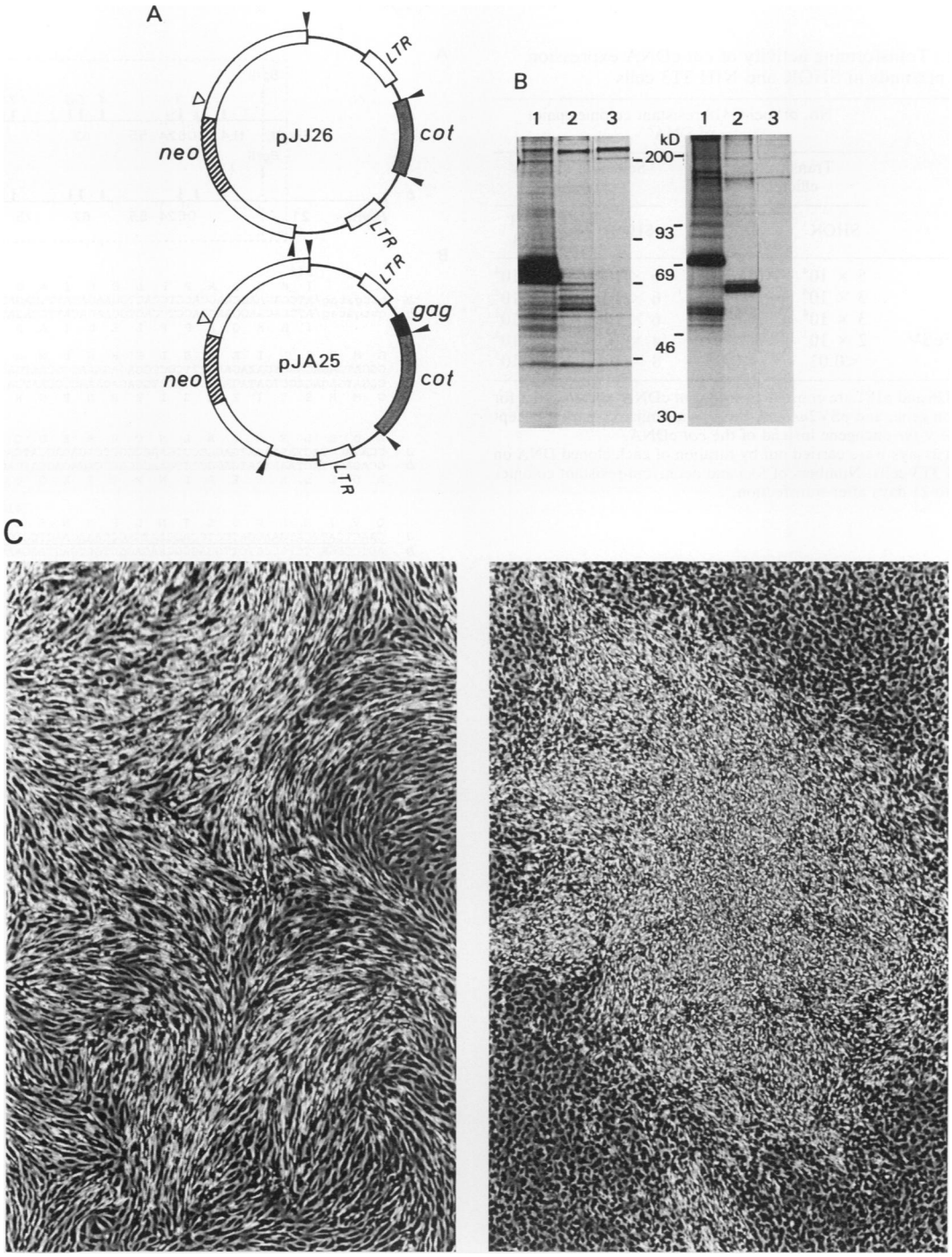

FIG. 5. Biological activity associated with expression of *cot* cDNA. (A) Structure of *cot* cDNA expression plasmids. The expression vectors consist of pSV2*neo* and retroviral sequences derived from GR-FeSV (26) and linker insertion mutants (25). The *cot* cDNA was expressed under control of the 5′ long terminal repeat, whereas the selection marker *neo* gene was transcribed independently from the simian virus 40 early promoter. The arrowheads indicate *Eco*RI sites, and the open triangles indicate simian virus 40 *ori*. (B) Protein kinase activity associated with $p70^{gag\text{-}actin\text{-}fgr}$ and $p60^{gag\text{-}cot}$. Cell extracts containing 200 μg of soluble protein were prepared from transformed SHOK cells labeled with [^{35}S]methionine. Immunoprecipitates with anti-*gag* antiserum were analyzed by separation on an SDS–10% polyacrylamide gel and fluorography (left panel). To detect protein kinase activity, cell extracts were prepared separately from parallel cultures that were not metabolically labeled. Immunoprecipitation was carried out in the same condition as metabolic labeling. Immunoprecipitates were subjected to in vitro kinase assays for analyzing autophosphorylation of transforming proteins (right panel). Radioactivities of ^{35}S and ^{32}P incorporated into transforming proteins were quantified independently by scintillation counting of bands excised from the SDS-polyacrylamide gel. The relative specific activity was calculated by dividing the ^{32}P radioactivity by the ^{35}S radioactivity. Transfected plasmids: lane 1, GR-FeSV; lane 2, pJA25; lane 3, pJJ26. (C) Transformation of SHOK and NIH 3T3 cells by the plasmid pJJ26. Phase-contrast photographs (40× magnification) of transformed foci of SHOK and NIH 3T3 cells induced by pJJ26 are presented. Transformed SHOK cells are refractile and spindle-shaped and tend to be arranged in a whorled fashion (left panel). These features were reproducible with other expression plasmids of *cot* cDNA. In contrast, the plasmid appeared to induce ordinary forms of transformed foci of NIH 3T3 cells (right panel), which were not readily distinguishable from those induced by activated *ras* and other oncogenes.

TABLE 1. Transforming activity of *cot* cDNA expression plasmids in SHOK and NIH 3T3 cells

Plasmid[a]	No. of foci G418-resistant colonies/pmol of DNA[b]			
	Transforming efficiency		Selection for Neo^r expression	
	SHOK	NIH 3T3	SHOK	NIH 3T3
pJJ26	5×10^4	6×10^4	6×10^4	7×10^4
pJF1	3×10^4	3×10^4	6×10^4	6×10^4
pJA25	3×10^4	2×10^4	6×10^4	5×10^4
pSV2*neo*GR-FeSV	2×10^3	3×10^4	4×10^4	5×10^4
pSV2*neo*	<0.01	<0.01	8×10^4	1×10^5

[a] Plasmids pJJ26 and pJF1 are constructs of the *cot* cDNA, pJA25 codes for the *gag-cot* fusion gene, and pSV2*neo*GR-FeSV is a similar construct except that it carries the v-*fgr* oncogene instead of the *cot* cDNA.

[b] Transfection assays were carried out by titration of each cloned DNA on SHOK and NIH 3T3 cells. Numbers of foci and neomycin-resistant colonies were scored 14 to 21 days after transfection.

terminated by a stop codon TAG. These results suggest that the *cot* oncogene might be activated by a C′-terminal truncation during the process of DNA transfer, although we have no idea about the C′ terminus of the normal Cot kinase because we have no information about the normal transcript of the *cot* proto-oncogene.

We conclude from these findings that both overexpression and structural mutation could be involved in oncogenic activation of the *cot* oncogene.

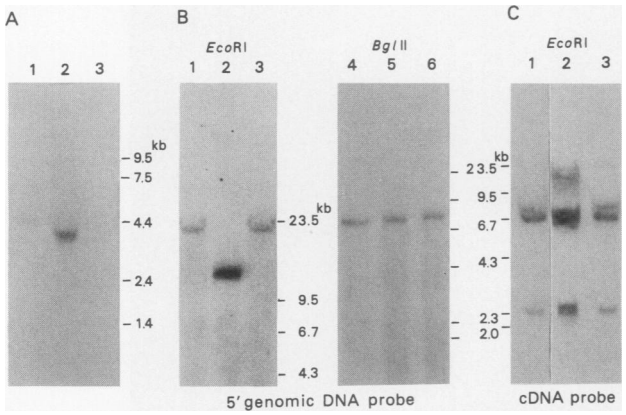

FIG. 6. Northern and Southern blot hybridization of the *cot* oncogene. (A) Steady-state *cot* mRNA expression in TCO4 and SHOK cells. Total RNAs (10 μg per lane) were analyzed on a 1% agarose formaldehyde gel. RNAs were blotted and hybridized to ^{32}P-labeled *cot* cDNA probe. The autoradiogram was exposed for 2 days. RNA standards are shown. Lane 1, TCO4; lane 2, transformed SHOK; lane 3, normal SHOK. (B) Southern blot analysis of 5′ untranslated regions in the *cot* locus. Samples of 20 μg of DNAs from human placenta, SHOK transformants, and TCO4 cells were digested with *Eco*RI (lanes 1 through 3) and *Bgl*II (lanes 4 through 6), size fractionated on 0.4 and 0.7% agarose gels, and transferred to nitrocellulose membranes. The blots were hybridized with nick-translated probes encoding an *Alu*-negative genomic sequence in the 5′ upstream region (0.7 kb). Lane 1, human placenta; lane 2, transformed SHOK; lane 3, TCO4; lane 4, human placenta; lane 5, TCO4; lane 6, transformed SHOK. (C) Southern blot analysis of coding regions in the *cot* locus. The blot was hybridized with a nick-translated *cot* cDNA probe containing the entire open reading frame. Lane 1, human placenta; lane 2, transformed SHOK; lane 3, TCO4.

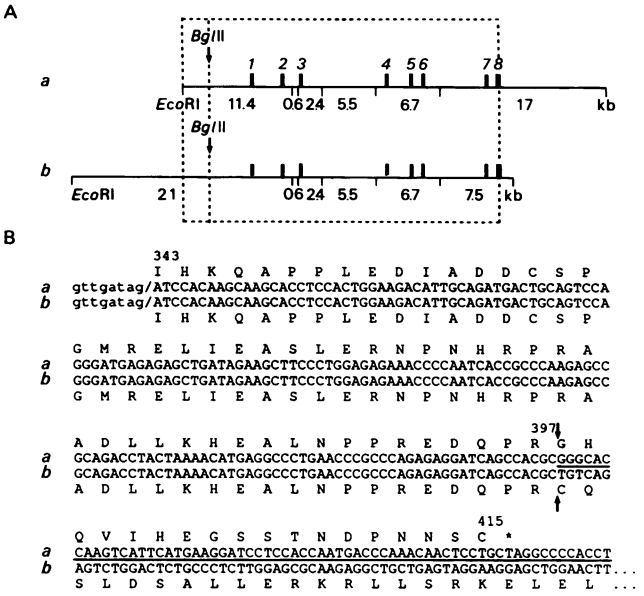

FIG. 7. DNA rearrangement in exon 8 of the *cot* oncogene resulting in C′-terminal truncation of the Cot protein. (A) Rearrangements in the 5′ and 3′ regions of the *cot* oncogene are schematically shown. Rearrangements were deduced by restriction mapping of the *cot* oncogene (*a*) and the cot proto-oncogene (*b*). The region enclosed by broken lines was conserved in both genes. The 5′ rearranged site was mapped between the *Eco*RI and *Bgl*II sites, and the 3′ rearranged site was within exon 8. (B) DNA sequencing of exon 8 in the *cot* oncogene (*a*) and its normal counterpart (*b*) indicated that the 3′ rearranged point was as indicated by arrows. Exon 8 of the *cot* oncogene was rearranged at the junction of codons 397 and 398 and followed by the short sequence encoding 18 amino acid residues and a stop codon. The DNA sequence downstream of the 3′ rearranged site is underlined. The 3′ boundary of exon 8 of the *cot* proto-oncogene was not identified, and the presence of following exons was not confirmed because no information about normal *cot* transcripts is available.

DISCUSSION

This paper describes the molecular cloning of the novel *cot* oncogene and analyses of biological activities associated with the *cot* cDNA expressed in a proviral vector. By sequence analysis of the *cot* cDNA and a computer search, we found significant similarity between the amino acid sequence of the predicted Cot protein and those of the protein kinase family. In addition, we constructed a chimeric plasmid encoding the *gag-cot* protein and detected the protein kinase activity that resides in p60$^{gag\text{-}cot}$ by immune complex kinase assay with an antiserum raised against retroviral p15^{gag} protein. With regard to the phosphoamino acid specificity of the *gag-cot* protein kinase, we have detected phosphoserine but not phosphotyrosine in the *gag-cot* protein after in vitro phosphorylation (data not shown). However, further investigation is needed to show that this activity is truly intrinsic to the Cot protein. Cot was found to have a unique ATP-binding site starting with proline instead of glycine and a tyrosine residue as a possible autophosphorylation site, but it is noteworthy that the amino acid sequences in subdomains VI and VIII predicting the amino acid specificity of protein kinases were more closely

related to those of serine-threonine protein kinases than to those of tyrosine protein kinases. A novel kinase, STY, has recently been cloned with an antibody directed against phosphotyrosine from a murine embryonal carcinoma cell line (15). It appears to have Ser-, Thr-, and Tyr-phosphorylating activities, although its cDNA shares sequence homology with serine-threonine kinases. Thus, it would be of some interest to know whether Cot is of the STY type. The structure of the amino-terminal region of Cot, however, was considerably different from that of STY, which contains a putative nuclear localization signal. In addition, we have obtained preliminary data on the subcellular localization of Cot suggesting that Cot is localized mainly in the cytoplasm. There seems to be no evidence that Cot is of the STY type or a nuclear protein. Our previous observations that SHOK cells tended to be transformed preferentially by *mos* and related oncogenes and our isolation of a rearranged *mos* oncogene from mouse osteosarcoma by SHOK assay (unpublished results) also appeared to favor the idea that the Cot protein is a serine-threonine protein kinase. We are now trying to generate antisera against Cot, which are crucial for assessing this issue in detail.

In principle, oncogenic activation could occur through two mechanisms, namely, deregulation of expression and alteration of function. In the case of the *cot* oncogene, both mechanisms were likely to be involved. Comparison of the structure of the *cot* oncogene with that of its normal allele showed that the sequence downstream from the *Bgl*II site in the 5′ untranslated region was conserved during development of the original thyroid carcinoma and serial DNA transfection. Recently, we observed promoter activity of this conserved sequence (unpublished results). In contrast, a transcript of the *cot* oncogene was readily detected in transformed SHOK cells but not in thyroid tumor TCO4 cells. It is thus probable that some mechanism(s) regulating the expression of the *cot* proto-oncogene in TCO4 cells was less effective or absent in SHOK cells. On the other hand, it is conceivable that the rearrangement of the protein coding region of the *cot* oncogene that resulted in a truncated and substituted carboxyl terminus could lead to removal of regulatory constraints so that the Cot protein remained constitutively and inappropriately active. Thus, the oncogenic activation in SHOK cells may have involved unregulated expression as well as structural change in the Cot protein as a consequence of DNA transfection. These inferences suggest that the *cot* oncogene did not participate in the carcinogenesis of the original thyroid cancer.

The present results show that the *cot* oncogene is a new member of the protein kinase family. Further investigation of the *cot* protein function may throw light on other protein kinase pathways involved in the control of cell growth and proliferation. Our results also indicate that SHOK cell assay is useful for detecting additional transforming genes in DNAs that cannot be detected by NIH 3T3 assay, although the basic differences between these two assays remain to be determined.

ACKNOWLEDGMENTS

We are grateful to Yoko Sugiyama for generous support and encouragement.

This work was supported in part by a Grant-in-Aid for Special Project Research on Cancer-Bioscience from the Ministry of Education, Science, and Culture of Japan.

REFERENCES

1. **Anderson, S. K., C. P. Gibbs, A. Tanaka, H.-J. Kung, and D. J. Fujita.** 1985. Human cellular *src* gene: nucleotide sequence and derived amino acid sequence of the region coding for the carboxy-terminal two-thirds of $pp60^{c\text{-}src}$. Mol. Cell. Biol. **5:** 1122–1129.
2. **Bonner, T. I., H. Oppermann, P. Seeburg, S. B. Kerby, M. A. Gunnell, A. C. Young, and U. R. Rapp.** 1986. The complete coding sequence of the human *raf* oncogene and the corresponding structure of the c-*raf*-1 gene. Nucleic Acids Res. **14:**1009–1015.
3. **Borek, C., A. Ong, and H. Mason.** 1987. Distinctive transforming genes in X-ray-transformed mammalian cells. Proc. Natl. Acad. Sci. USA **84:**794–798.
4. **Brady, G., H. M. Jantzen, H. U. Bernard, R. Brown, G. Schutz, and T. Hashimoto-Gotoh.** 1984. New cosmid vectors developed for eukaryotic DNA cloning. Gene **27:**223–232.
5. **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.
6. **Eva, A., and S. A. Aaronson.** 1985. Isolation of a new human oncogene from a diffuse B-cell lymphoma. Nature (London) **316:**273–275.
7. **Fickett, J. W.** 1982. Recognition of protein coding regions in DNA sequences. Nucleic Acids Res. **10:**5303–5318.
8. **Fitzgerald, M., and T. Shenk.** 1981. The sequence 5′-AAUA AA-3′ forms part of the recognition site for polyadenylation of late SV40 mRNAs. Cell **24:**251–260.
9. **Goldfarb, M., K. Shimizu, M. Perucho, and M. Wigler.** 1982. Isolation and preliminary characterization of a human transforming gene from T24 bladder carcinoma cells. Nature (London) **296:**404–409.
10. **Graham, F. L., and A. J. van der Eb.** 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology **52:**456–467.
11. **Gribskov, M., J. Devereux, and R. R. Burgess.** 1984. The codon preference plot: graphic analysis of protein coding sequences and prediction of gene expression. Nucleic Acids Res. **12:**539–549.
12. **Guerrero, I., P. Calzada, A. Mayer, and A. Pellicer.** 1984. A molecular approach to leukemogenesis: mouse lymphomas contain an activated c-*ras* oncogene. Proc. Natl. Acad. Sci. USA **81:**202–205.
13. **Hanks, S. K., A. M. Quinn, and T. Hunter.** 1988. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. Science **241:**42–52.
14. **Higashi, T., H. Sasai, F. Suzuki, J. Miyoshi, T. Ohuchi, S. Takai, T. Mori, and T. Kakunaga.** 1990. Hamster cell line suitable for transfection assay of transforming genes. Proc. Natl. Acad. Sci. USA **87:**2409–2413.
15. **Howell, B. W., D. E. H. Afar, J. Lew, E. M. J. Douville, P. L. E. Icely, D. A. Gray, and J. C. Bell.** 1991. STY, a tyrosine-phosphorylating enzyme with sequence homology to serine/threonine kinases. Mol. Cell. Biol. **11:**568–572.
16. **Hunter, T.** 1987. A thousand and one protein kinases. Cell **50:**823–829.
17. **Hynes, N. E., R. Jaggi, S. C. Kozma, R. Ball, D. Muellener, T. Wetherall, B. W. Davis, and B. Groner.** 1985. New acceptor cell for transfected genomic DNA: oncogene transfer into a mouse mammary epithelial cell line. Mol. Cell. Biol. **5:**268–272.
18. **Karn, J., H. W. D. Matthes, M. J. Gait, and S. Brenner.** 1984. A new selective phage cloning vector, λ2001, with sites for XbaI, BamHI, HindIII, EcoRI, SstI and XhoI. Gene **32:**217–224.
19. **Kozak, M.** 1987. An analysis of 5′-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res. **15:**8125–8132.
20. **Lipman, D. J., and W. R. Pearson.** 1985. Rapid and sensitive protein similarity searches. Science **227:**1435–1441.
21. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
22. **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.

23. **McGeoch, D. J., and A. J. Davison.** 1986. Alpha herpesviruses possess a gene homologous to the protein kinase gene family of eukaryotes and retroviruses. Nucleic Acids Res. **14**:1765–1777.
24. **Melton, D. A., P. A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green.** 1984. Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res. **12**:7035–7056.
25. **Miyoshi, J., Y. Miyoshi, H. Sasai, N. Sakai, T. Katsumata, and T. Kakunaga.** 1989. Differential requirements of *gag* and γ-actin domains for transforming potential of Gardner-Rasheed feline sarcoma virus. J. Virol. **63**:1174–1180.
26. **Naharro, G., S. R. Tronick, S. Rasheed, M. B. Gardner, S. A. Aaronson, and K. C. Robbins.** 1983. Molecular cloning of integrated Gardner-Rasheed feline sarcoma virus: genetic structure of its cell-derived sequence differs from that of other tyrosine kinase-coding *onc* genes. J. Virol. **47**:611–619.
27. **Parker, P. J., L. Coussens, N. Totty, L. Rhee, S. Young, E. Chen, S. Stabel, M. D. Waterfield, and A. Ullrich.** 1986. The complete primary structure of protein kinase C—the major phorbol ester receptor. Science **233**:853–859.
28. **Pulciani, S., E. Santos, A. V. Lauver, L. K. Long, K. C. Robbins, and M. Barbacid.** 1982. Oncogenes in human tumor cell lines: molecular cloning of a transforming gene from human bladder carcinoma cells. Proc. Natl. Acad. Sci. USA **79**:2845–2849.
29. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
30. **Shih, C., and R. A. Weinberg.** 1982. Isolation of transforming sequence from a human bladder carcinoma cell line. Cell **29**:161–169.
31. **Shimizu, K., Y. Nakatsu, M. Sekiguchi, K. Hokamura, K. Tanaka, M. Terada, and T. Sugimura.** 1985. Molecular cloning of an activated human oncogene, homologous to v-*raf*, from primary stomach cancer. Proc. Natl. Acad. Sci. USA **82**:5641–5645.
32. **Shtivelman, E., B. Lifshitz, R. P. Gale, B. A. Roe, and E. Canaani.** 1986. Alternative splicing of RNAs transcribed from the human *abl* gene and from the *bcr-abl* fused gene. Cell **47**:277–284.
33. **Sutherland, B. M., P. V. Bennett, A. G. Freeman, S. P. Moore, and P. T. Strickland.** 1985. Transformation of human cells by DNAs ineffective in transformation of NIH3T3 cells. Proc. Natl. Acad. Sci. USA **82**:2399–2403.
34. **Takio, K., R. D. Wade, S. B. Smith, E. G. Krebs, K. A. Walsh, and K. Titani.** 1984. Guanosine cyclic 3′,5′-phosphate dependent protein kinase, a chimeric protein homologous with two separate protein families. Biochemistry **23**:4207–4218.
35. **Terada, N., J. Miyoshi, K. Kawaha, H. Sasai, S. Orita, K. Yumura-Yagi, J. Hara, A. Fujinami, and T. Kakunaga.** 1990. Alteration of N-*ras* gene mutation after relapse in acute lymphoblastic leukemia. Blood **75**:453–457.
36. **Ullrich, A., L. Coussens, J. S. Hayflick, T. J. Dull, A. Gray, A. W. Tam, J. Lee, Y. Yarden, T. A. Libermann, J. Schlessinger, J. Downward, E. L. V. Mayes, N. Whittle, M. D. Waterfield, and P. H. Seeburg.** 1984. Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells. Nature (London) **309**:418–425.
37. **Watson, R., M. Oskarsson, and G. F. Vande Woude.** 1982. Human DNA sequence homologous to the transforming gene (*mos*) of Moloney murine sarcoma virus. Proc. Natl. Acad. Sci. USA **79**:4078–4082.
38. **Wigler, M., R. Sweet, G. K. Sim, B. Wold, A. Pellicer, E. Lacy, T. Maniatis, S. Silverstein, and R. Axel.** 1979. Transformation of mammalian cells with genes from prokaryotes and eukaryotes. Cell **16**:777–785.
39. **Young, D., G. Waitches, C. Birchmeier, O. Fasano, and M. Wigler.** 1986. Isolation and characterization of a new cellular oncogene encoding a protein with multiple potential transmembrane domains. Cell **45**:711–719.