Cell, Vol. 45, 711–719, June 6, 1986, Copyright © 1986 by Cell Press

Isolation and Characterization of a New Cellular Oncogene Encoding a Protein with Multiple Potential Transmembrane Domains

Dallan Young, Gayle Waitches, Carmen Birchmeier, Ottavio Fasano,* and Michael Wigler
Cold Spring Harbor Laboratory
P.O. Box 100
Cold Spring Harbor, New York 11724

Summary

We have cloned and sequenced a new human oncogene and have named it *mas*. This oncogene was detected by its tumorigenicity in nude mice using the cotransfection and tumorigenicity assay previously described. The *mas* oncogene has a weak focus-inducing activity in transfected NIH 3T3 cells. A DNA rearrangement in the 5′ noncoding sequence, which occurred during transfection, is probably responsible for activation of the *mas* gene. The cDNA sequence of the *mas* oncogene reveals a long open reading frame that codes for a 325 amino acid protein. This protein is very hydrophobic and has seven potential transmembrane domains. In this respect, the structure of the *mas* protein is novel among cellular oncogene products and may reflect a new functional class of oncogenes.

Introduction

The first oncogenes discovered were the transforming genes of the oncogenic viruses (reviewed by Bishop, 1985). The subsequent discovery that the oncogenes of retroviruses were derived from normal host cellular genes provided the first direct evidence that cellular genomes contain genes with transforming potential. More recently, the development of techniques for DNA transfer in eukaryotic cells led to the discovery of cellular transforming genes in tumor cells by their ability to induce foci of transformed NIH 3T3 cells (reviewed by Land et al., 1983). Several new oncogenes have been discovered this way, including N-*ras* (Shimizu et al., 1983), *met* (Cooper et al., 1984), *neu* (Bargmann et al., 1986), and possibly others (Goubin et al., 1983; Lane et al., 1984; Takahashi et al., 1985).

To search for oncogenes that may escape detection by the focus induction assay, our lab has explored a sensitive bioassay for transforming genes based on the tumorigenicity in nude mice of NIH 3T3 cells following cotransfection with a selectable marker and DNA from tumor cells. Using this assay, three transforming genes were derived from the DNA of the MCF-7 cell line (Fasano et al., 1984). Here, we describe a new human oncogene we call "*mas*" that was detected by the tumorigenicity assay following cotransfection with DNA isolated from a human epidermoid carcinoma. This gene efficiently induces tumorigenicity and has a weak focus-inducing activity in NIH 3T3 cells. We have cloned cDNAs containing the entire coding sequence of *mas*. The *mas* gene encodes a protein with seven hydrophobic regions that are potential transmembrane domains, suggesting that *mas* is an integral membrane protein. The structure of *mas* protein is unique among cellular oncoproteins and may represent a new functional class.

* Present address: EMBL Meyerhof Strasse 1, D 6900 Heidelberg, Federal Republic of Germany.

Results

Isolation of the *mas* Oncogene

The *mas* oncogene was detected using the cotransfection and tumorigenicity assay we have previously described (Fasano et al., 1984). This method is a modification of the tumorigenicity assay described by Blair et al. (1982). DNA from a human epidermoid carcinoma was used to cotransfect NIH 3T3 cells with the plasmid pKOneo, and transfected cells were selected with the neomycin analogue G418. In one experiment, one of six nude mice injected with these cells developed a "primary" tumor within 4 weeks. Nude mice injected with cells cotransfected with DNA isolated from this primary tumor developed "secondary" tumors within 2 weeks. After a third round, DNA was purified from a "tertiary" tumor, and a genomic library was constructed in the cosmid vector pHC79. Four overlapping cosmid clones that contained human DNA were isolated by filter hybridization to total human genomic DNA (Gusella et al., 1980; Shih and Weinberg, 1982). Characterization of these cosmids by restriction mapping and Southern hybridization to total human DNA revealed that one of the cosmids, pMAS1, contains a 22 kb stretch of human DNA flanked by mouse DNA on both sides (Figure 1).

The pMAS1 cosmid was tested by the cotransfection and tumorigenicity assay to determine whether it contained transforming potential. Nude mice developed tumors within 2 weeks after injection of cells that had been cotransformed with pMAS1 and pKOneo and selected for resistance to G418 (see Table 1). Furthermore, in our standard focus assay, NIH 3T3 cells transfected with pMAS1 formed foci within 16 days. The foci of cells transformed with pMAS1 are unlike foci of cells transformed with the activated human H-*ras*^{val12} oncogene, isolated from the T24 bladder carcinoma cell line. They appear at a lower frequency about 4–6 days later (Table 2). The foci induced by pMAS1 are characterized not by an abnormal morphology of the constituent cells, but by an exceedingly high cell density (see Figure 2).

To define the regions of pMAS1 essential for transforming activity, various restriction endonuclease digests of pMAS1 were tested by the NIH 3T3 focus assay. When pMAS1 was cut with EcoRI or SalI, the transforming potential was destroyed, indicating that regions including one or more of each of these sites are essential for transforming activity (Table 2). In contrast, pMAS1 DNA cut with SmaI or XhoI retained its transforming ability. A 7.3 kb re-

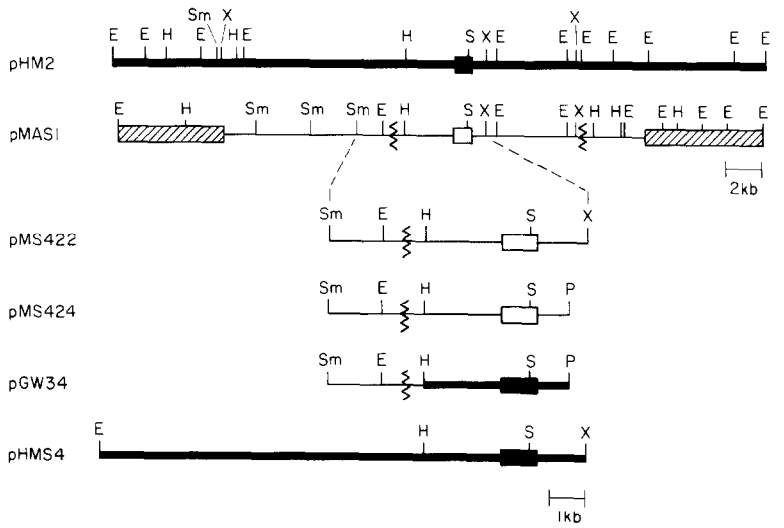

Figure 1. Maps of Restriction Endonuclease Sites in *mas* Clones

Filled boxes and thick lines indicate the coding region and flanking sequences derived from the placental *mas* allele. Open boxes and thin lines indicate the coding region and flanking sequences derived from the rearranged *mas* allele. The hatched areas indicate regions containing mouse repetitive DNA. Squiggles indicate the sites of DNA rearrangements. Restriction sites are indicated as follows: E, EcoRI; H, HpaI; S, SalI; Sm, SmaI; P, PstI; and X, XhoI.

gion of pMAS1, which is defined by SmaI and XhoI sites and contains single EcoRI and SalI sites, is capable of transforming NIH 3T3 cells. This region was subcloned into pUC8 to generate pMS422 (Figure 1).

Activation of the *mas* Oncogene by Rearrangement

Comparison of Southern blots of normal human DNA with DNA derived from the tertiary nude mouse tumor revealed a difference in the size and intensity of EcoRI restriction fragments homologous to *mas* (Figure 3), suggesting that the *mas* gene was rearranged and amplified in transformants. To determine the nature of the DNA rearrangement, clones containing homology to a region of pMAS1 were isolated from a human placenta cosmid library. The probe used was the 2.2 kb SalI–EcoRI restriction fragment, which does not contain Alu-repetitive sequences and includes the 3′ portion of the *mas* coding sequence (see next section). Three independent overlapping clones designated pHM1, pHM2, and pHM3 were obtained. Comparison of maps of restriction endonuclease sites in pHM2 and pMAS1 confirm that there is a break in homology localized between the EcoRI and HpaI sites in the 5′ noncoding region of the *mas* gene (Figure 1). Although this break point is not in the *mas* coding region (see next section), it does occur in a region essential for transformation of NIH 3T3 cells by the *mas* gene (see above and Table 2). This suggests that the rearrangement found in the transformant is of functional significance.

The cosmid pHM2, which contains the normal human homolog of *mas*, was tested to determine whether it has transforming activity. By our standard transfection assay, pHM2 did not induce foci of transformed NIH 3T3 cells even after 4 weeks. However, pHM2 did induce tumors in a cotransfection and tumorigenicity assay, although with a longer lag time than pMAS1 (see Table 1). Therefore, the normal *mas* clone, pHM2, has a weak biological transforming activity detected by our tumorigenicity assays.

To test whether the DNA rearrangement was responsible for activation of *mas* transforming potential, we constructed the hybrid clone pGW34, which contains the 5′ noncoding region of the rearranged gene and the entire coding region from the normal human homolog (see Figure 1 and next section). This hybrid clone had focus-inducing activity similar to the *mas* clone, pMS422 (Table 2). In contrast, pHMS4, a subclone of the normal placenta

Table 1. Tumorigenicity Assays

	Mean Tumor Diameter (mm) after Week:					
Test Plasmid	1	2	3	4	5	6
pMAS-1	0.0	11.1	19.6			
	0.0	9.8	16.0			
	0.0	10.9	17.0			
	0.0	10.6	18.0			
pHM-2	0.0	0.0	4.5	10.6	11.2	
	0.0	0.0	5.4	10.6	14.0	
	0.0	0.0	0.0	5.8	13.6	
	0.0	0.0	0.0	8.7	13.7	
	0.0	0.0	0.0	0.0	5.0	9.8
	0.0	0.0	0.0	0.0	0.0	0.0
None	0.0	0.0	0.0	0.0	0.0	0.0
	0.0	0.0	0.0	0.0	0.0	0.0
	0.0	0.0	0.0	0.0	0.0	0.0
	0.0	0.0	0.0	0.0	0.0	0.0
pT24	10.6	22.8				
	9.0	22.8				

Tumorigenicity assays were performed as previously described (Fasano et al., 1984). NIH 3T3 cells (8×10^5 cells per plate) were cotransformed with 200 ng pKOneo, 30 μg of high molecular weight NIH 3T3 DNA, and test plasmid (300 ng of pMAS1, 300 ng of pHM2, or 50 ng of pT24) per plate. Following transfection, cells were split 1:5 and were selected for resistance to the antibiotic G418. After 2 weeks the cells were confluent and were injected into nude mice (10^7 cells per mouse). Tumor formation at the site of injection was followed for 6 weeks, and the mean tumor diameter was measured. Each line in the above table represents a single animal injected with independently cotransformed pools of cells.

Table 2. Focus Assays

Experiment	DNA	ng/Plate	Foci/ng
Experiment A	pMAS1	200	0.105[a]
	pHM2	200	< 0.0025[a]
	pHM2	1000	< 0.0005[a]
	pT24	50	2.0[a]
Experiment B	pMAS1	100	0.36[b]
	pMAS1/SmaI	100	0.41[b]
	pMAS1/XhoI	100	0.50[b]
	pMAS1/EcoRI	100	< 0.01[b]
	pMAS1/SalI	100	< 0.01[b]
Experiment C	pMS422	400	0.61[b]
	pMS424	400	0.50[b]
	pGW34	400	0.60[b]
	pHMS4	400	< 0.001[b]

Focus assays in NIH 3T3 cells were performed as previously described (Perucho et al., 1981). NIH 3T3 cells were transformed with the indicated amount of test DNAs, either cleaved with the indicated restriction enzymes (Experiment B), or uncleaved (Experiment A and Experiment C). High molecular weight NIH 3T3 DNA (30 μg/plate) was used as carrier DNA. After the indicated times, the number of foci were scored, and the number of foci per nanogram of test DNA was calculated. See Figure 1 for a description of the plasmid clones.
[a] Foci scored after 2 weeks.
[b] Foci scored after 3 weeks.

clone (Figure 1), did not induce foci, suggesting that the DNA rearrangement is responsible for activation of *mas* transforming potential. Since this rearrangement lies outside of coding sequences, we suspect that activation of the *mas* gene results from overproduction or inappropriate production of its normal product. However, in order to make a definitive conclusion we need to know the nature of the normal gene transcript.

Finally, we sought to determine whether the *mas* gene was activated in the donor tumor DNA or, rather, became activated during cotransfection into NIH 3T3 cells. Southern blotting experiments demonstrated that the *mas* genes resident in the original human epidermoid carcinoma did not have the rearranged structure found in the NIH 3T3 cotransformant, nor were they amplified (Figure 3). Moreover, we have not observed the transfer of the *mas* gene from the original tumor DNA in other cotransformants. Thus, we have no evidence that the *mas* gene was active in the original human tumor. Rather, it is likely that the *mas* gene became rearranged and activated during gene transfer from the original human tumor DNA.

Organization and Sequence of the *mas* Gene Transcript

To define the transcription unit and coding potential of the *mas* gene, we cloned cDNAs complementary to *mas* mRNA. Poly(A)$^+$ mRNA was purified from a cell line, MAS-133, derived from the nude mouse tumor from which pMAS1 was isolated. Blot hybridization of this RNA with the *mas* gene insert from pMS422 revealed a homologous mRNA approximately 2.5 kb in length (data not shown). A cDNA library was constructed from this poly(A)$^+$ RNA in the λgt10 vector and was screened for homology to pMS422. Sixteen overlapping cDNA clones were isolated and characterized by restriction endonuclease site mapping and DNA sequencing (see Figure 4). From sequence data we assembled a composite nucleotide sequence that contains a complete open reading frame of 975 bp (Figure 5). The first ATG in this reading frame is preceded by an in-frame stop codon at position −12. The entire open reading frame was contained on a single cDNA clone, pMC130. From these data we have assumed the orientation of transcription as shown in Figure 4. To demonstrate that the *mas* transforming product is coded by pMC130, we constructed a plasmid, pKC30, containing the coding region of pMC130 linked to the SV40 promoter. NIH 3T3 cells cotransfected with pKC30 and pKOneo and selected with G418 are tumorigenic in nude mice and give rise to foci (data not shown).

Comparison of the restriction endonuclease cleavage sites of genomic and cDNA sequences indicated that the cDNA is entirely colinear with genomic sequences. This observation was subsequently confirmed for the coding region by direct nucleotide sequencing of the homologous

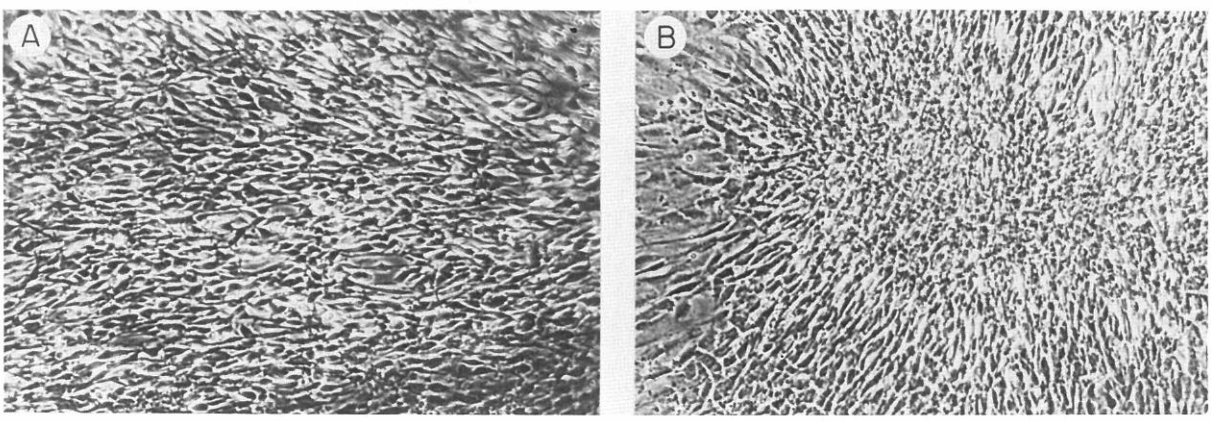

Figure 2. Foci from *mas* and *ras* Transformed Cells

Foci of NIH 3T3 cells transformed with pT24 (A) and pMS422 (B) after 14 days. pT24 carries an activated human H-*ras*^{val12} gene (Fasano et al., 1983).

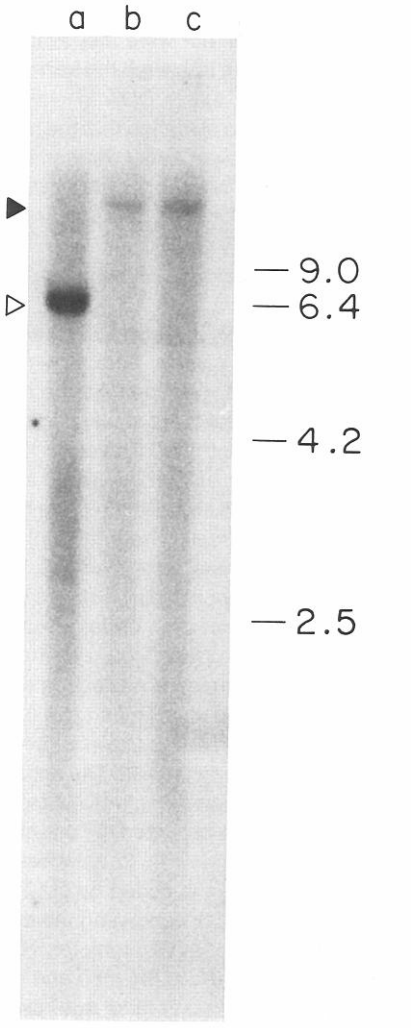

Figure 3. Southern Blot of Human and *mas* Transformant DNA

EcoRI-cut DNAs derived from: lane a, MAS-133 cell line; lane b, the original human epidermoid carcinoma, and lane c, human placenta. Blots were hybridized to the 2.2 kb SalI–EcoRI alu-repeat free fragment of pMAS1. Open and closed arrows indicate positions of *mas*-related sequences in MAS-133 and human genomic DNAs, respectively. The molecular weights of size markers are indicated in kilobases on the right.

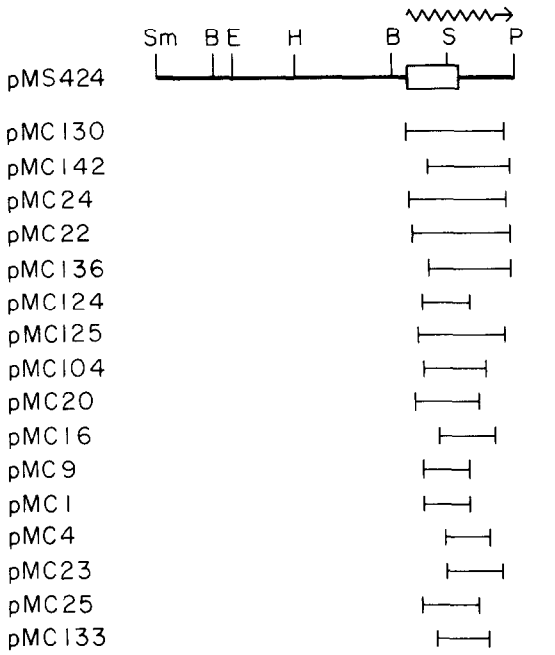

Figure 4. Relation of *mas* cDNA Clones to Genomic Clone pMS424

Inserts from the indicated cDNA clones are aligned under the genomic sequence. The open box indicates the coding region. The arrow shows the direction of transcription. Abbreviations for restriction endonuclease sites are as described in Figure 1.

region of the genomic DNA clone pMS422 (see Figure 5). By S1 analysis we determined that the colinearity of the mRNA with the genomic DNA extends in the 5′ direction approximately 45 bases beyond the first ATG in the coding region (Figure 6). The site marking the beginning of colinearity between RNA and genomic sequences may reflect a transcription initiation site or a splice site. Since no cDNA clones containing poly(A) were found, it is likely that *mas* transcripts extend beyond the 3′ end of the cDNA clone pMC142 (see Figure 4).

To learn more about the organization of the normal *mas* gene and the events that can lead to its activation, we examined RNA in cells that were cotransformed with the cosmid pHM2, selected for tumorigenicity in nude mice, and subsequently placed back into culture. Total RNA was prepared from these cells and was analyzed in the manner described for MAS-133 cells (Figure 6). This analysis indicates that transcripts of *mas* in these cells are also colinear with genomic DNA 5′ to the open reading frame already identified. Hence, the protein encoded in these cells is probably the same as in MAS-133.

Predicted Primary and Secondary Structure of the *mas* Protein

The complete nucleotide sequence of the coding region determined from *mas* cDNAs and the corresponding predicted amino acid sequence of 325 amino acids are shown in Figure 5. The deduced amino acid sequence does not share significant homology with any published sequence. Analysis of the hydrophobicity plot of the sequence by the method of Kyte and Doolittle (1982) reveals that the *mas* protein has seven distinct hydrophobic regions (Figure 7). Computational methods based on hydrophobicity profiles have been developed to distinguish transmembrane regions from hydrophobic internal regions of globular proteins (for review see Eisenberg, 1984). A recently developed algorithm correctly predicts membrane-spanning segments in many proteins, including the seven segments in bacteriorhodopsin, and correctly disqualifies hydrophobic regions from many soluble proteins (Eisenberg et al., 1984). This algorithm predicts seven transmembrane domains in the *mas* protein corresponding to the seven hydrophobic regions. Each of the hydrophobic regions is separated by hydrophilic regions that contain a predicted β-turn secondary structure (Chou

```
      GGATCCAGAAGGGTCATTCAATCAGTTCTCAGTCTTATCAGGTCTAAGTTCCTTTCTTATCAGGTCCTAAAGG
CCTAATCTTATCATTGTGACAAAGATAACTGTAGAGTCTGTTAAACTTTTTTTTAATAACATGAAGATTATGATTTATA
GCTGAATTTCTCCCTTTTATTCCAATTCAACAATTTTCATGGCTTTTTGTGTTTGTTTTGTTCTGGACATATTTACAGA

                          •••                 MET ASP GLY SER ASN VAL THR SER PHE VAL    10
AAATTACCTGAAGAGTTCCAACC TGA GGC CTC CTC ATG GAT GGG TCA AAC GTG ACA TCA TTT GTT

VAL GLU GLU PRO THR ASN ILE SER THR GLY ARG ASN ALA SER VAL GLY ASN ALA HIS ARG    30
GTT GAG GAA CCC ACG AAC ATC TCA ACT GGC AGG AAC GCC TCA GTC GGG AAT GCA CAT CGG

GLN ILE PRO ILE VAL HIS TRP VAL ILE MET SER ILE SER PRO VAL GLY PHE VAL GLU ASN    50
CAA ATC CCC ATC GTG CAC TGG GTC ATT ATG AGC ATC TCC CCA GTG GGG TTT GTT GAG AAT

GLY ILE LEU LEU TRP PHE LEU CYS PHE ARG MET ARG ARG ASN PRO PHE THR VAL TYR ILE    70
GGG ATT CTC CTC TGG TTC CTG TGC TTC CGG ATG AGA AGA AAT CCC TTC ACT GTC TAC ATC

THR HIS LEU SER ILE ALA ASP ILE SER LEU LEU PHE CYS ILE PHE ILE LEU SER ILE ASP    90
ACC CAC CTG TCT ATC GCA GAC ATC TCA CTG CTC TTC TGT ATT TTC ATC TTG TCT ATC GAC

TYR ALA LEU ASP TYR GLU LEU SER SER GLY HIS TYR TYR THR ILE VAL THR LEU SER VAL   110
TAT GCT TTA GAT TAT GAG CTT TCT TCT GGC CAT TAC TAC ACA ATT GTC ACA TTA TCA GTG

THR PHE LEU PHE GLY TYR ASN THR GLY LEU TYR LEU LEU THR ALA ILE SER VAL GLU ARG   130
ACT TTT CTG TTT GGC TAC AAC ACG GGC CTC TAT CTG CTG ACG GCC ATT AGT GTG GAG AGG

CYS LEU SER VAL LEU TYR PRO ILE TRP TYR ARG CYS HIS ARG PRO LYS TYR GLN SER ALA   150
TGC CTG TCA GTC CTT TAC CCC ATC TGG TAC CGA TGC CAT CGC CCC AAG TAC CAG TCG GCA

LEU VAL CYS ALA LEU LEU TRP ALA LEU SER CYS LEU VAL THR THR MET GLU TYR VAL MET   170
TTG GTC TGT GCC CTT CTG TGG GCT CTT TCT TGC TTG GTG ACC ACC ATG GAG TAT GTC ATG

CYS ILE ASP ARG GLU GLU GLU SER HIS SER ARG ASN ASP CYS ARG ALA VAL ILE ILE PHE   190
TGC ATC GAC AGA GAA GAA GAG AGT CAC TCT CGG AAT GAC TGC CGA GCA GTC ATC ATC TTT

ILE ALA ILE LEU SER PHE LEU VAL PHE THR PRO LEU MET LEU VAL SER SER THR ILE LEU   210
ATA GCC ATC CTG AGC TTC CTG GTC TTC ACG CCC CTC ATG CTG GTG TCC AGC ACC ATC TTG

VAL VAL LYS ILE ARG LYS ASN THR TRP ALA SER HIS SER SER LYS LEU TYR ILE VAL ILE   230
GTC GTG AAG ATC CGG AAG AAC ACG TGG GCT TCC CAT TCC TCC AAG CTT TAC ATA GTC ATC

MET VAL THR ILE ILE ILE PHE LEU ILE PHE ALA MET PRO MET ARG LEU LEU TYR LEU LEU   250
ATG GTC ACC ATC ATT ATA TTC CTC ATC TTC GCT ATG CCC ATG AGA CTC CTT TAC CTG CTG

TYR TYR GLU TYR TRP SER THR PHE GLY ASN LEU HIS HIS ILE SER LEU LEU PHE SER THR   270
TAC TAT GAG TAT TGG TCG ACC TTT GGG AAC CTA CAC CAC ATT TCC CTG CTC TTC TCC ACA

ILE ASN SER SER ALA ASN PRO PHE ILE TYR PHE PHE VAL GLY SER SER LYS LYS LYS ARG   290
ATC AAC AGT AGC GCC AAC CCT TTC ATT TAC TTC TTT GTG GGA AGC AGT AAG AAG AAG AGA

PHE LYS GLU SER LEU LYS VAL VAL LEU THR ARG ALA PHE LYS ASP GLU MET GLN PRO ARG   310
TTC AAG GAG TCC TTA AAA GTT GTT CTG ACC AGG GCT TTC AAA GAT GAA ATG CAA CCT CGG

ARG GLN LYS ASP ASN CYS ASN THR VAL THR VAL GLU THR VAL VAL •••                   325
CGC CAG AAA GAC AAT TGT AAT ACG GTC ACA GTT GAG ACT GTC GTC TAA GAACTGTGAGGGAAG

TTGTGGATAAAAATGGTGGAACACAGGTCATTTTTAGTTTGTGCTTGGAATATGACTTAAGTATCTCCTAAATGTGATA
CAGAAGAACATCTCATCCCATATGCATGAGATACTAATTAATGATGAAA
```

Figure 5. Nucleotide Sequence of Coding and Flanking Regions of the *mas* Gene

The DNA sequence was derived from cDNA clones and the genomic subclone pMS422. The amino acid sequence deduced from the coding region is shown above the DNA sequence. The in-frame stop codons of the open reading frame are indicated by asterisks. The DNA sequence from the fourteenth nucleotide position 5′ from the start ATG to the 3′ end was derived from the cDNA clones. The sequence of the coding and 5′ regions were determined from the genomic clone pMS422. The numbers on the right are amino acid coordinates.

and Fasman, 1978). Both the amino- and carboxy-terminal ends of the *mas* protein are hydrophilic. This analysis strongly suggests that the *mas* protein is an integral membrane protein with many transmembrane domains.

Several proteins that span the membrane multiple times have been identified and studied. These include bacteriorhodopsin and the eukaryotic visual rhodopsins (Ovchinnikov, 1982), lactose permease (Foster et al., 1983), the acetylcholine receptor (Ross et al., 1977; Popot and Changeux, 1984), the sodium ion channel (Noda et al., 1984), the (Na^+ K^+)ATPase (Kawakami et al., 1985; Shull et al., 1985), the Ca^{2+}-ATPase (MacLennan et al., 1985), and the erythrocyte anion exchange protein (Jay and Cantley, 1986). In Figure 7 we show the hydrophobicity plots of bovine rhodopsin and the α-subunit of acetylcholine receptor for comparison with *mas* protein. There is a striking similarity in the hydrophobicity patterns of *mas* and rhodopsin, which may reflect structural and functional similarities in these proteins.

In Figure 8 we illustrate the relative positions of amino acids in the transmembrane domains of the *mas* protein assuming an α-helical structure. This figure shows the hydrophobic amino acid residues and their relation to charged and polar residues. The pattern of hydrophobic and hydrophilic domains can be clearly seen. This protein has a total of 28 positively charged and 21 negatively charged residues distributed primarily in the hydrophilic regions; however, there are some charged residues within the hydrophobic domains. Charged residues have been observed in many of the transmembrane domains of membrane proteins.

The *mas* protein does not contain an amino-terminal hydrophobic signal sequence characteristic of many membrane proteins (see Wickner and Lodish, 1985, for review on protein translocation). However, some proteins with multiple transmembrane domains, such as bovine rhodopsin, the erythrocyte anion exchange protein, and the sodium channel protein, lack amino-terminal signal sequences. In the first two cases, insertion into the membrane is cotranslational and requires internal signal sequences (Braell and Lodish, 1982; Friedlander and Blobel, 1985). Since the *mas* protein does not contain an amino-terminal hydrophobic signal sequence, its integration into the membrane may also depend on internal signal sequences, or alternatively, it may spontaneously become inserted into the membrane because of its hydrophobic nature (Blobel, 1980; Engelman and Steitz, 1981).

It has been established that the tripeptide sequence Asn-X-Thr/Ser is a site for N-glycosylation in secreted and membrane proteins (reviewed by Kornfeld and Kornfeld, 1985), although not all of these sequences are glycosylated (Struck and Lennarz, 1980). There are four potential sites for N-glycosylation in the predicted protein sequence of *mas* at positions 5, 16, 22, and 272. The first three of these sites are clustered in the first hydrophilic amino-terminal region, while the fourth site is in the seventh hydrophobic domain.

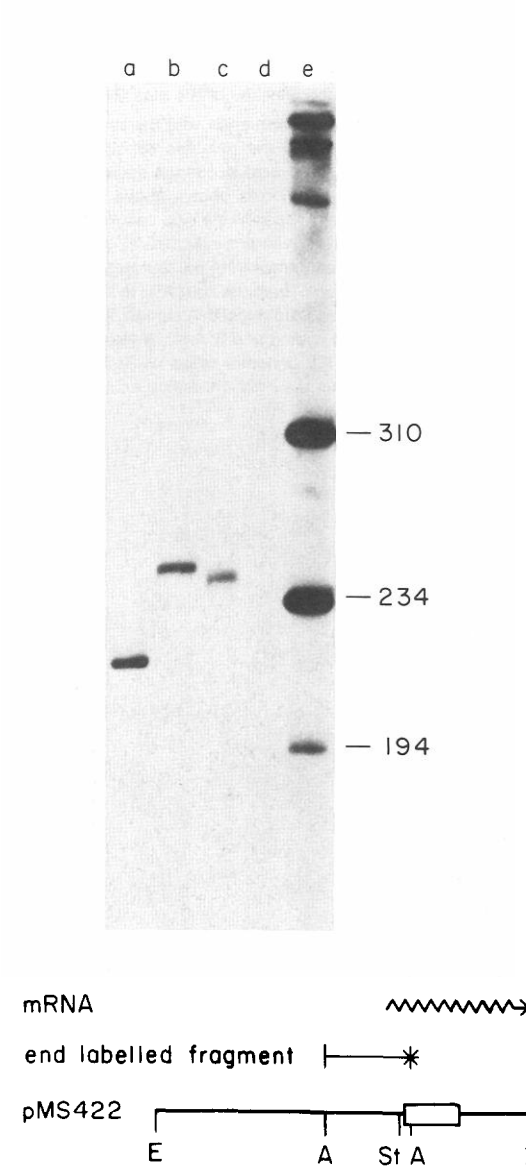

Figure 6. S1 Analysis of RNA Transcripts

The AccI fragment of the genomic clone pMS422 was end-labeled, hybridized to RNA, and digested with S1 nuclease. RNA used was isolated from: lane d, NIH 3T3 cells; lane c, MAS-133; or lane b, NIH 3T3 cells transformed with the normal *mas* clone pHM2. This figure also shows: lane a, the 211 bp StuI–AccI fragment of pMS422; and lane e, molecular weight markers. Restriction endonuclease sites are indicated as follows: A, AccI; St, StuI; X, XhoII; and E, EcoRI. The StuI site is 10 bp 5′ to the ATG initiation codon and includes the last base pair of the first in-frame stop codon. These data indicate that colinearity between RNA transcripts and genomic sequences from the *mas* gene extend 5′ beyond the termination codon at the StuI site.

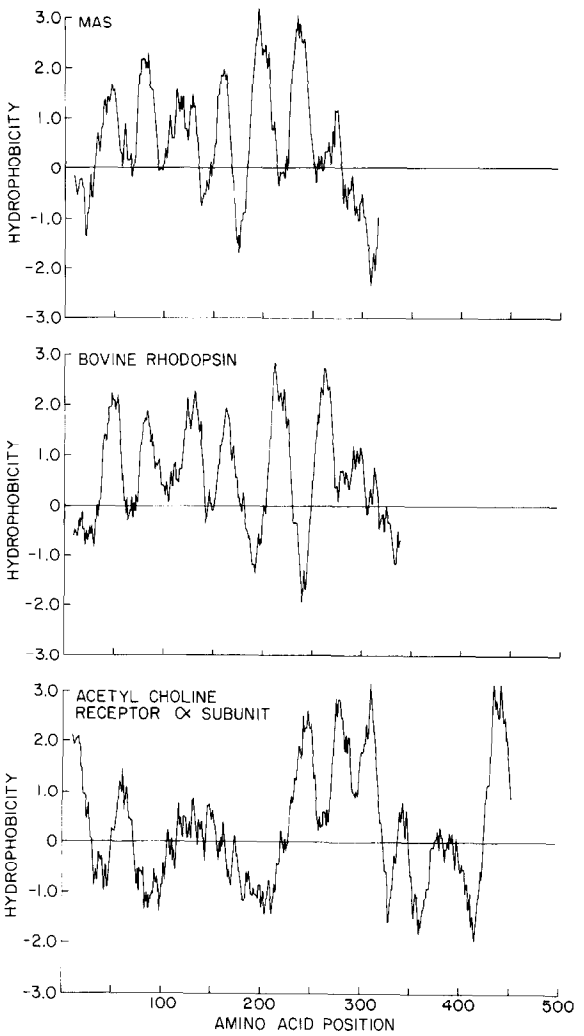

Figure 7. Hydrophobicity Plots of the Predicted *mas* Protein

Average hydrophobic values were determined for spans of 19 residues using the method and hydrophobic values of Kyte and Doolittle (1982). The plots of the predicted *mas* protein, bovine rhodopsin, and the α-subunit of acetylcholine receptor are shown for comparison.

Discussion

The cotransfection and tumorigenicity assay was developed to search for transforming genes from tumor cells that may not be readily detected by the standard focus assay. Using this assay, three transforming genes were previously isolated from the MCF-7 cell line (Fasano et al., 1984). One of these genes is a normal N-*ras* gene, which is amplified in MCF-7 DNA. Another, designated *mcf2*, is currently being investigated, and the third gene, *mcf3*, is the human homolog of v-*ros* (Birchmeier et al., unpublished). In the latter case, the human *ros* gene was rearranged during or after gene transfer, probably resulting in its activation. A similar event appears to have occurred in the activation of the *mas* gene. The *mas* gene was found to be rearranged and amplified in the tertiary nude mouse tumor DNA. In contrast, the *mas* genes of the original human tumor DNA, used in the first round of cotransfection experiments, have the normal configuration and are not amplified. Although the normal human *mas* gene has weak tumor-inducing activity, it has no detectable focus-inducing activity. By contrast, the rearranged *mas* gene has strong tumor-inducing activity and can induce foci of NIH 3T3 cells. Analysis by chimeric gene construction indicates that a similar rearrangement can activate the normal placental allele. Therefore, we conclude that the cotransfection and tumorigenicity assay has a propensity

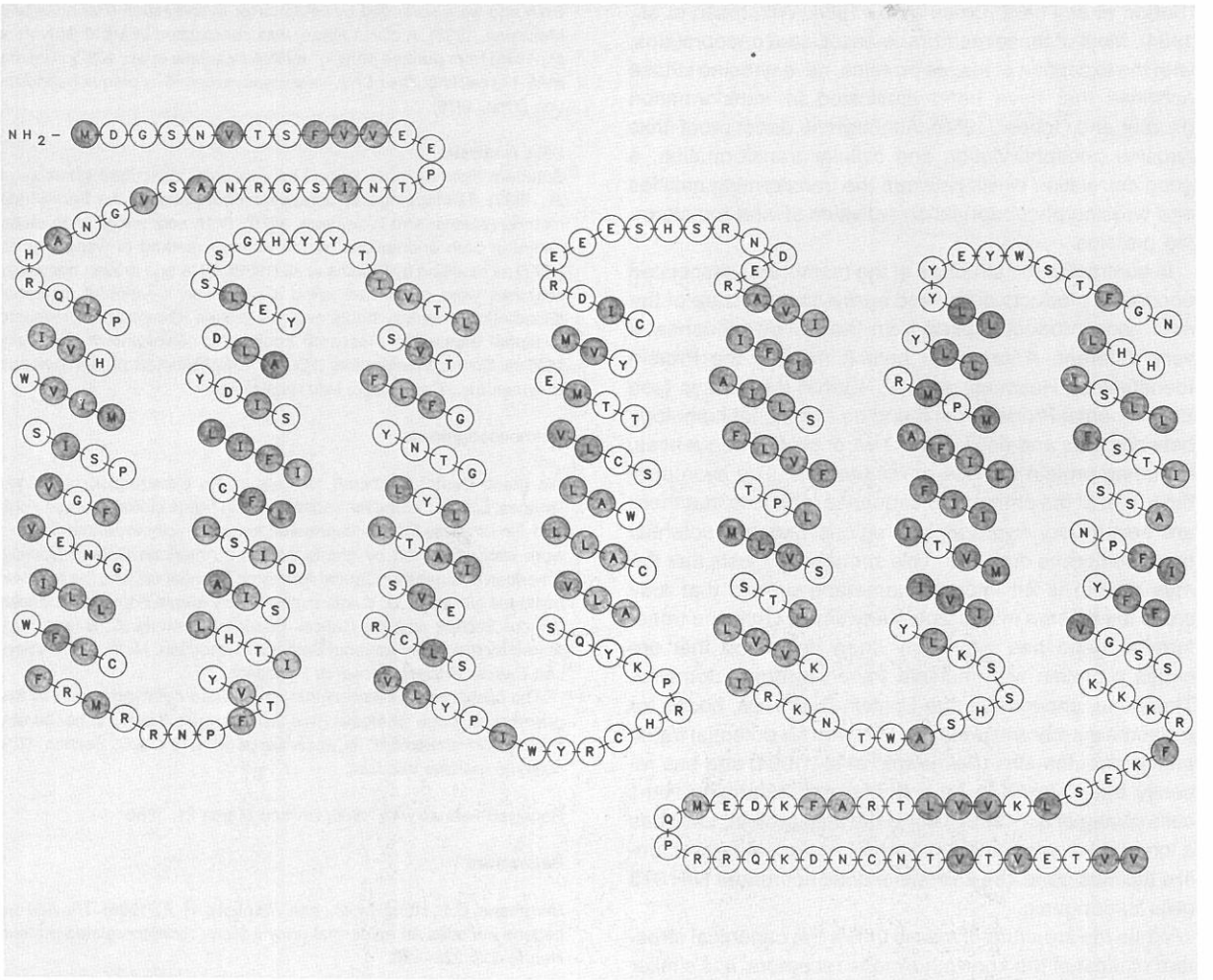

Figure 8. *mas* Protein Sequence

The hydrophobic domains are drawn assuming an α-helical structure with 3.6 residues per turn. The hydrophobic amino acids Leu, Ile, Val, Phe, Ala, and Met are shaded.

for detecting those proto-oncogenes that can induce tumorigenicity in NIH 3T3 cells by rearrangement and/or amplification following cotransfection. In this respect, the NIH 3T3 cotransfection and tumorigenicity assay differs considerably from the NIH 3T3 focus assay, which very rarely scores positive with DNA that does not already contain an activated oncogene. There is one reported case where the NIH 3T3 focus assay detected a transforming gene that may have been activated by a DNA rearrangement (Takahashi et al., 1985).

We do not understand precisely how the *mas* gene became activated during gene transfer. In transformants, the *mas* gene has been rearranged, but this rearrangement does not appear to involve the coding domain. It is clear that we can activate the oncogenic potential of a normal placental allele by reconstructing a chimeric gene that replaces the normal 5′ sequence with a sequence 5′ to the coding region of the rearranged gene. This chimeric gene leaves intact the long open reading frame of the placental allele. Thus it seems likely that transformation by *mas* results from inappropriate expression of a normal gene product. However, we cannot rigorously exclude the possibility that the rearranged alleles have an altered splicing pattern, which results in an altered protein product. To resolve this problem completely, we must identify *mas* transcripts in a normal cell, which we have not yet been able to do. However, we have examined transcripts of the *mas* gene in cells cotransformed with the normal allele. S1 analysis indicates that in such cells we can exclude the existence of an additional amino-terminal coding domain.

There are currently about 30 known viral or cellular oncogenes; many code for proteins associated with the plasma membrane (Land et al., 1983). Three oncogenes, v-*erb*B, v-*fms*, and v-*ros* code for proteins with single transmembrane domains and are probably related to growth factor receptors (Downward et al., 1984; Hampe et al., 1984; Neckameyer and Wang, 1985). While *ras* and *src* do not encode transmembrane domains, they are associated with the membrane and have fatty acid residues that are added to the proteins by a posttranslational modification

(Sefton et al., 1982; Cross et al., 1984; Willumsen et al., 1984). Most of these membrane-associated oncoproteins, with the exception of the *ras* proteins, have tyrosine kinase activities that have been implicated in transformation (Hunter and Cooper, 1986). Although no direct proof links tyrosine phosphorylation and cellular transformation, a good correlation exists between the transforming abilities and tyrosine phosphorylation activities of viral transforming proteins.

In contrast to the structure of the membrane-associated oncogene products discussed above, the structure of the *mas* gene product implied from the cDNA sequence is very different. A computer search through the Protein Identification Resource and GENBANK data bases (see Experimental Procedures) found no significant homology between *mas* and any known DNA or protein sequences, including protein kinases or *ras* proteins. The hydrophobicity plot of the amino acid sequence indicates that there are seven very hydrophobic regions that are potential transmembrane domains. This strongly suggests that the *mas* protein is an integral membrane protein that may cross the plasma membrane many times. Only one transforming gene has previously been described that encodes a protein with multiple transmembrane domains. This gene, encoded by the Epstein-Barr virus, codes for a plasma membrane protein (LMP) with six potential transmembrane domains (Fennewald et al., 1984) and has recently been shown to be capable of transforming Rat-1 cells (Wang et al., 1985). Unlike the *mas* protein, LMP has a long carboxy-terminal hydrophilic region. Moreover, unlike the *mas* gene, the LMP gene does not render NIH 3T3 cells tumorigenic.

While the structure of *mas* is unlike the canonical structure of most of the known hormone receptors, it is similar to a class of proteins that includes the acetylcholine receptor and the visual rhodopsins. The acetylcholine receptor functions as a hormonally regulated ion channel. Visual rhodopsin is a light receptor that functions to activate transducin, an intracellular guanine nucleotide binding protein. We speculate that *mas* may be a receptor that activates a critical component in a growth regulatory pathway, perhaps by serving in signal transduction or as a membrane channel. The unique nature of *mas* leads us to suspect that it may provide a new link in understanding growth control.

Experimental Procedures

Focus and Tumorigenicity Assays

High molecular weight DNA was purified from cell lines (Perucho et al., 1981) and solid tumors (Fasano et al., 1984) as described. DNA transfer into NIH 3T3 cells were performed by a modified calcium phosphate precipitation method (Wigler et al., 1979). Focus assays (Perucho et al., 1981) and tumorigenicity experiments (Fasano et al., 1984) were performed as previously described. The plasmids pT24 (containing the activated human H-*ras*^{val12} gene) and pKOneo (containing a neomycin/G418 antibiotic resistance gene) were previously described (Fasano et al., 1983, 1984).

Construction of Libraries

Genomic libraries were constructed in the cosmid vector pHC79 (Hohn and Collins, 1980; Maniatis et al., 1982) from EcoRI partially cleaved DNA and were screened by colony filter hybridization (Hanahan and Meselson, 1983). A cDNA library was constructed in λgt10 (Huynh et al., 1985) from purified poly(A)$^+$ mRNA (Maniatis et al., 1982) from the MAS-133 cell line. The cDNA library was screened by plaque hybridization (Woo, 1979).

DNA Analysis

Southern blots were performed as previously described (Shimizu et al., 1983). S1 mapping was done by a modification of the Berk-Sharp method (Weaver and Weissman, 1979). DNA sequences were determined in both orientations by the dideoxy method of Sanger et al. (1977) as modified by Biggins et al. (1983). DNA and protein homology searches were performed using a previously developed algorithm (Goad and Kanehisa, 1982), and the Protein Identification Resource (National Biomedical Research Foundation, Georgetown University Medical Center, Washington, D.C.) and GENBANK (Bolt, Beranek and Newman Inc., Cambridge, MA) data banks.

Acknowledgments

We thank Daniel Birnbaum for help in the construction of cosmid libraries, Louise Edlund for assistance with tissue culture, Jorgen Fogh and Jim Loveless for assistance with tumorigenicity experiments. This work was supported by grants from the American Cancer Society, American Business for Cancer Research Foundation, and the National Institutes of Health. D. Y. was supported by grants from the American Cancer Society and the Cancer Research Institute. C. B. was supported by the Swiss National Science Foundation. M. W. is an American Cancer Society Research Professor.

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

Received February 13, 1986; revised March 21, 1986.

References

Bargmann, C. I., Hung, M.-C., and Weinberg, R. A. (1986). The *neu* oncogene encodes an epidermal growth factor receptor-related protein. Nature *319*, 226–230.

Biggins, M. D., Gibson, T. J., and Hong, G. F. (1983). Buffer gradient gels and ^{35}S label as an aid to rapid DNA sequence determination. Proc. Natl. Acad. Sci. USA *80*, 3963–3965.

Bishop, J. M. (1985). Viral oncogenes. Cell *42*, 23–38.

Blair, D. G., Cooper, C. S., Oskarsson, M. K., Eader, L. A., and Vande Woude, G. F. (1982). New method for detecting cellular transforming genes. Science *281*, 1122–1125.

Blobel, G. (1980). Intracellular protein topogenesis. Proc. Natl. Acad. Sci. USA *77*, 1496–1500.

Braell, W. A., and Lodish, H. F. (1982). The erythrocyte anion transport protein is cotranslationally inserted into microsomes. Cell *28*, 23–31.

Chou, P. Y., and Fasman, G. D. (1978). Prediction of the secondary structure of proteins from their amino acid sequence. Adv. Enzymol. *47*, 45–148.

Cooper, C. S., Park, M., Blair, D. G., Tainsky, M. A., Huebner, K., Croce, C. M., and Vande Woude, G. F. (1984). Molecular cloning of a new transforming gene from a chemically transformed human cell line. Nature *311*, 29–33.

Cross, F. R., Garber, E. A., Pellman, D., and Hanafusa, H. (1984). A short sequence in the p60-src N-terminus is required for p60-src myristylation and membrane association for cell transformation. Mol. Cell. Biol. *4*, 1834–1842.

Downward, J., Yarden, Y., Mayes, E., Scrace, G., Totty, N., Stockwell, P., Ullrich, A., Schlessinger, J., and Waterfield, M. D. (1984). Close similarity of epidermal growth factor receptor and v-erb-B oncogene protein sequences. Nature *307*, 521–527.

Eisenberg, D. (1984). Three-dimensional structure of membrane and surface proteins. Ann. Rev. Biochem. *53*, 595–623.

Eisenberg, D., Schwarz, E., Komaromy, M., and Wall, R. (1984). Analy-

sis of membrane and surface protein sequences with the hydrophobic moment plot. J. Mol. Biol. *179*, 125–142.

Engelman, D. M., and Steitz, T. A. (1981). The spontaneous insertion of proteins into and across membranes: the helical hairpin hypothesis. Cell *23*, 411–422.

Fasano, O., Taparowsky, E., Fiddes, J., Wigler, M., and Goldfarb, M. (1983). Sequence and structure of the coding region of the human H-ras-1 gene from T24 bladder carcinoma cells. J. Mol. Appl. Genet. *2*, 173–180.

Fasano, O., Birnbaum, D., Edlund, L., Fogh, J., and Wigler, M. (1984). New human transforming genes detected by a tumorigenicity assay. Mol. Cell. Biol. *4*, 1695–1705.

Fennewald, S., van Santen, V., and Kieff, E. (1984). Nucleotide sequence of an mRNA transcribed in latent growth-transforming virus infection indicates that it may encode a membrane protein. J. Virol. *51*, 411–419.

Foster, D. L., Boublik, M., and Kaback, H. R. (1983). Structure of the lac carrier protein of E. coli. J. Biol. Chem. *258*, 31–34.

Friedlander, M., and Blobel, G. (1985). Bovine opsin has more than one signal sequence. Nature *318*, 338–343.

Goad, W. B., and Kanehisa, M. I. (1982). Pattern recognition in nucleic acid sequences. I. A general method for finding local homologies and symmetries. Nucl. Acids Res. *10*, 247–263.

Goubin, G., Goldman, D. S., Luce, J., Neiman, P. E., and Cooper, G. (1983). Molecular cloning and nucleotide sequence of a transforming gene detected by transfection of chicken B-cell lymphoma DNA. Nature *302*, 114–119.

Gusella, J. F., Keys, C., Varsanyi-Breiner, A., Kao, F., Jones, C., Puck, T. T., and Housman, D. (1980). Isolation and localization of DNA segments from specific human chromosomes. Proc. Natl. Acad. Sci. USA *77*, 2829–2833.

Hampe, A., Gobet, M., Sherr, C., and Galibert, F. (1984). Nucleotide sequence of the feline retroviral oncogene v-fms shows unexpected homology with oncogenes encoding tyrosine-specific protein kinases. Proc. Natl. Acad. Sci. USA *81*, 85–89.

Hanahan, D., and Meselson, M. (1983). Plasmid screening at high colony density. Meth. Enzymol. *100*, 333–342.

Hohn, B., and Collins, J. (1980). A small cosmid for efficient cloning of large DNA fragments. Gene *11*, 291–298.

Hunter, T., and Cooper, J. A. (1986). Viral oncogenes and tyrosine phosphorylation. In The Enzymes "Enzyme Control by Phosphorylation," P. D. Boyer and E. G. Krebs, eds. (New York: Academic Press), in press.

Huynh, T. V., Young, R. A., and Davis, R. W. (1985). Constructing and screening cDNA libraries in lambda gt10 and lambda gt11. In DNA Cloning Techniques: A Practical Approach, D. Glover, ed. (Oxford: IRL Press), pp. 49–78.

Jay, D., and Cantley, L. (1986). Structural aspects of the red cell anion exchange protein. Ann. Rev. Biochem., in press.

Kawakami, K., Noguchi, S., Noda, M., Takahashi, H., Ohta, T., Kawamura, M., Nojima, H., Nagano, K., Hirose, T., Inayama, S., Hayashida, H., Miyata, T., and Numa, S. (1985). Primary structure of the alpha-subunit of *Torpedo californica* ($Na^+ + K^+$)ATPase deduced from cDNA sequence. Nature *316*, 733–736.

Kornfeld, K., and Kornfeld, S. (1985). Assembly of asparagine-linked oligosaccharides. Ann. Rev. Biochem. *54*, 631–664.

Kyte, J., and Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. *157*, 105–132.

Land, H., Parada, L. F., and Weinberg, R. A. (1983). Cellular oncogenes and multistep carcinogenesis. Science *222*, 771–778.

Lane, M. A., Sainten, A., Doherty, K. M., and Cooper, G. M. (1984). Isolation and characterization of a stage-specific transforming gene, Tlym-1, from T-cell lymphomas. Proc. Natl. Acad. Sci. USA *81*, 2227–2231.

MacLennan, D. H., Brandl, C. J., Korczak, B., and Green, N. M. (1985). Amino-acid sequence of a $Ca^{2+} + Mg^{2+}$-dependent ATPase from rabbit muscle sarcoplasmic reticulum, deduced from its complementary DNA sequence. Nature *316*, 696–700.

Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982). Molecular Cloning: A Laboratory Manual. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory).

Neckameyer, W. S., and Wang, L. H. (1985). Nucleotide sequence of avian sarcoma virus UR2 and comparison of its transforming gene with other members of the tyrosine protein kinase oncogene family. J. Virol. *53*, 879.

Noda, M., Shimizu, S., Tanabe, T., Takai, T., Kayano, T., Ikeda, T., Takahashi, H., Nakayama, H., Kanaoka, Y., Minamino, N., Kangawa, K., Matsuo, H., Raftery, M. A., Hirose, T., Inayama, S., Hayashida, H., Miyata, T., and Numa, S. (1984). Primary structure of Electrophorus electricus sodium channel deduced from cDNA sequence. Nature *312*, 121–127.

Ovchinnikov, Y. A. (1982). Rhodopsin and bacteriorhodopsin: structure-function relationships. FEBS Lett. *148*, 179–191.

Perucho, M., Goldfarb, M., Shimizu, K., Lama, C., Fogh, J., and Wigler, M. (1981). Human-tumor-derived cell lines contain common and different transforming genes. Cell *27*, 467–476.

Popot, J., and Changeux, J. (1984). Nicotinic receptor of acetylcholine: structure of an oligomeric integral membrane protein. Physiol. Rev. *64*, 1162–1239.

Ross, M. J., Klymkowsky, M. W., Agard, D., and Stroud, R. M. (1977). Structural studies of a membrane-bound acetylcholine receptor from Torpedo californica. J. Mol. Biol. *116*, 635–659.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA *74*, 5463–5467.

Sefton, B. M., Trowbridge, I. S., Cooper, J. A., and Scolnick, E. M. (1982). The transforming proteins of Rous sarcoma virus, Harvey sarcoma virus and Abelson virus contain tightly bound lipid. Cell *31*, 465–474.

Shih, C., and Weinberg, R. A. (1982). Isolation of a transforming sequence form a human bladder carcinoma cell line. Cell *29*, 161–169.

Shimizu, K., Goldfarb, M., Suard, Y., Perucho, M., Li, Y., Kamata, T., Ferimisco, J., Stavnezer, E., Fogh, J., and Wigler, M. (1983). Three human transforming genes are related to the viral *ras* oncogenes. Proc. Natl. Acad. Sci. USA *80*, 2112–2116.

Shull, G. E., Schwartz, A., and Lingrel, J. B. (1985). Amino-acid sequence of the catalytic subunit of the ($Na^+ + K^+$)ATPase deduced from a complementary DNA. Nature *316*, 691–695.

Struck, D. K., and Lennarz, W. J. (1980). The function of saccharide-lipids in synthesis of glycoproteins. In The Biochemistry of Glycoproteins and Proteoglycans, W. Lennarz, ed. (New York: Plenum), pp. 35–83.

Takahashi, M., Ritz, J., and Cooper, G. M. (1985). Activation of a novel human transforming gene, *ret*, by DNA rearrangement. Cell *42*, 581–588.

Wang, D., Liebowitz, D., and Kieff, E. (1985). An EBV membrane protein expressed in immortalized lymphocytes transforms established rodent cells. Cell *43*, 831–840.

Weaver, R. F., and Weissman, C. (1979). Mapping of RNA by a modification of the Berk-Sharp procedure: the 5′ termini of 15S beta-globin mRNA precursor and mature 10S beta-globin mRNA have identical map coordinates. Nucl. Acids Res. *7*, 1175–1193.

Wickner, W. T., and Lodish, H. F. (1985). Multiple mechanisms of protein insertion into and across membranes. Science *230*, 400–407.

Wigler, M., Sweet, R., Sim, G. K., Wold, B., Pellicer, A., Lacy, E., Maniatis, T., Silverstein, S., and Axel, R. (1979). Transformation of mammalian cells with genes from procaryotes and eucaryotes. Cell *16*, 777–785.

Willumsen, B. M., Christensen, A., Hubbert, N. L., Papageorge, A. G., and Lowy, D. R. (1984). The p21 ras C-terminus is required for transformation and membrane association. Nature *310*, 583–586.

Woo, S. L. C. (1979). A sensitive and rapid method for recombinant phage screening. Meth. Enzymol. *68*, 389.