INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road. Ann Arbor. MI 48106-1346 USA
313/761-4700 800/521-0600

Order Number 9510718

cDNA transformation markers from SV40 transformed human epidermal keratinocytes

Shen, Huan, Ph.D.

City University of New York, 1994

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

cDNA Transformation Markers From SV40 Transformed Human Epidermal Keratinocytes

by

Huan Shen

A dissertation submitted to the Graduate Faculty in Biochemistry in partial fulfillment of the requirement for the degree of Doctor of Philosophy, The City University of New York

1994

This manuscript has been read and accepted for the Graduate Faculty in Biochemistry in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

April 28, 1994
Date

Mark S. Steinberg
Chair of Examining Committee

April 28, 1994
Date

Executive Officer

Supervisory Committee

The City University of New York

Abstract

cDNA Transformation Markers From SV40 Transformed Human Epidermal Keratinocytes

by

Huan Shen

Adviser: Professor Mark L. Steinberg

Candidate cDNAs representing sequences specifically induced by SV40 were selected by eliminating cDNAs hybridizing to labeled heterogeneous cDNA probes derived from normal human epithelial cells (previous work). From these candidates two cDNAs, E12 and E13, have been found to be markers of epithelial transformation in that the mRNAs show highly induced expression not only in SV40 transformants but also in cell lines derived from various squamous cell carcinomas.

Highly induced expression of E12 and E13 emerges early during SV40 transformation process and which are shared in common by transformed keratinocytes which arise independently from one another. These stable changes of expression are likely to play some significant role in transformation even if the genes are not themselves oncogenes. This indicates that the expression of the E12 and E13 may be transactivated directly or indirectly by SV40 virus.

A full-length cDNA of the E12 gene has been obtained by using the protocol of RACE (rapid amplification of cDNA ends). The

complete nucleotide sequence of the E12 cDNA is 1251 bp in length and the deduced open reading frame sequence is 715 bp encoding a 235 amino acid protein. Computer analyses and sequence homology searches indicated that the E12 transcript contains a 5' upstream Alu sequence. The Alu sequence is about 300 bp in length at the 5' end of E12 and is apparently expressed in the open reading frame; the Alu sequences were found in many mRNA sequences catalogued in Genbank. The functions of the Alu sequences in mRNAs are not known.
The remainder of the E12 gene showed no significant homology to any sequence currently cataloged in the Genbank, except for a piece of unannotated sequence from randomly amplified cDNA, HM02E05.
Two genomic libraries were created in the lambda phage vector and screened with E12 and E13 probes, respectively. Two positive clones, GE12 for the E12 cDNA and GE13 for the E13 cDNA were isolated from the libraries. GE12 contains a 9.2 kb insert and GE13 contains a 9.5 kb insert. Both GE12 and GE13 were subcloned into the plasmid PGEM3Z designated PGE12 and PGE13, respectively. Restriction maps and partial sequences of PGE12 and PGE13 were obtained. PGE12 and PGE13 were used for transfection focus-forming assays as an *in vitro* indication that these genes may carry oncogenic potential.

Acknowledgment

First of all, I wish to thank my mentor, Dr. Mark L. Steinberg for his support, trust, and time through my graduate experience. Mark was always supportive when things went wrong, and enthusiastic when things went well. With incredible determination and patience, he guided me through a process of thinking and writing.

I wish to thank Dr. Cy Garte and Dr. Jack Valdovinos for their time and valuable advice. Dr. David Calhoun and Dr. Jerry Guyden were always willing to listen and provide constructive feedback and I thank them for their support and help.

I thank My colleagues in Mark's lab, who formed a friendly and cooperative team.

My thanks also goes to my wife Ming, my son Daniel, my parents and my parents-in-law. Their love and encouragement have always been, and will continue to be, an inspiration to me.

TABLE OF CONTENTS

	Page
Introduction	1
1. Multistage carcinogenesis	1
2. Oncogenic viruses	2
3. DNA viruses	3
4. SV40 large T antigen and its functions in transformation	9
5. Transformation of keratinocytes with SV40	13
Materials and methods	20
1. Cell culture	20
2. Isolation of total RNA and mRNA	20
3. Isolation of Genomic DNA and southern blot	22
4. Construction of genomic libraries	24
5. Deletion of the GC tails of cDNA clones	25
6. Elongation of the 5' end of E12 gene	26
7. DNA amplification in remelted agarose	27
8. Computer analyses	28
Results	30
1. Expression of novel sequences in SV40 transformed and human carcinoma cell lines	30
2. Sequencing of selected cDNA clones	44
3. Southern blot analyses	52
4. Elongation and cloning of the E12 gene	55

5. Sequencing analysis of E12 gene 63

6. Construction and screening of genomic libraries ... 70

Discussion 84

1. Expression of novel genes in SV40 transformed keratinocytes 84

2. Transformation related cDNA markers 87

3. Transfection with E12 and E13 91

4. Transcriptional activation of E12 and the role of SV40 large T antigen 92

Appendix 98

1. GENBANK entrte of the E12 sequence 98

Bibliography 100

List of tables

1. Properties and functions of SV40 T antigen 10

List of figures

1. A. Physical and functional map of SV40 7
B. The transcription map of T antigen and
T antigen binding sites 8
2. Diagram of the emergence of transformed properties
over time following infection by SV40 16
3. Schematic representation of the RACE protocol 29
4. Northern blot hybridization of various candidate
cDNAs to RNAs from different cells 31
5. Northern blot hybridization of various candidate
cDNAs to RNAs from different cells 32
6. Expression of cDNA E5 in normal, transformed,
and cancer cells 34
7. Expression of cDNA E11 in normal, transformed,
and cancer cells 35
8. Expression of cDNA E14 in normal, transformed,
and cancer cells 36
9. Expression of cDNA E15 in normal, transformed,
and cancer cells 36
10. Expression of cDNA H2 in normal, transformed,
and cancer cells 38
11. Expression of E13 cDNA in normal, transformed,
and carcinoma cells 39
12. Expression of E12 cDNA in normal, transformed,

and carcinoma cells 40
13. Cytoplasmic RNAs on an agarose gel 41
14. Rehybridization of an actin probe for E12 blot ... 42
15. Expression of E12 cDNA in normal and SV40 transformed keratinocytes at serial passages 43
16. Cytoplasmic RNAs on an agarose gel 45
17. Rehybridization of an actin probe for E12 blot ... 46
18. The partial sequence and homology analysis of E5 . 48
19. The partial sequence and homology analysis of E11 49
20. The partial sequence and homology analysis of H2 . 50
21. Homology alignment of E12 fragment and HUMHM02E05 51
22. E12 sequences in genomic DNAs 53
23. E13 sequences in genomic DNAs 54
24. Synthesized RACE primers 56
25 The first round amplification of the 5' end of the E12 DNA 58
26. Southern blot hybridization of the first round RACE product 59
27. The second round amplification of the 5' end of the E12 DNA 58
28. Reamplification of RACE products 61
29. Alignment of E12 gene from three fragment 62
30. The sequences of E12 64
31. The putative protein composition 65
32. The putative protein conformation 66
33. The protein hydrophobicity for E12 67

34. The Alu sequence at 5' end of E12 68

36. Digestion of genomic clone GE12 71

37. Southern blot of clones from genomic library 72

38. Southern blot of pGE12 with B1012 probe 73

39. Single and double digestion of pGE12 75

40. Restriction map of pGE12 76

41. Partial sequences of pGE12 77

42. Hybridization of GE13 to an E13 probe 78

44. The partial sequences of the pGE13 79

45. Restriction enzyme digestion of pGE13 81

46. Double restriction enzyme digestion of pGE13 82

47. Restriction map of pGE13 83

48. The putative protein antigenicity of E12 89

49. T antigen binding sites in E12 90

Introduction

1. Multistage carcinogenesis

Carcinogenesis is a multistep process and can result from any one or a combination of chemical, physical, biological and genetic events. The traditional view of carcinognesis is derived primarily from animal models. The process may be divided into at least 3 stages: initiation, promotion, and progression. The process of initiation is permanent and irreversible. The first stage of neoplastic development involves exposure of normal cells to chemical, physical or microbial carcinogens. Those carcinogens and cell division leads to fixation of those changes (1, 2). The efforts was made to evaluate the initiation of carcinogenesis by identifying the hepatic foci in hepatocarcinogenesis system. Those foci were considered as progeny of single initiated cells (3, 4, 5).

At the promotion stage, initiated cells proliferate and this enhances the probability that additional genetic damage may accumulate and may involve in defects in terminal differentiation and growth control, activation of proto-oncogenes and inactivation of tumor suppresser genes (6). Thus promotion involves, instead of changes in the structure of DNA, changes in the expression of genome.

The progression stage is characterized by the appearance of malignant neoplasm and intrinsic genomic instability that is manifest by the abnormal number and structure of chromosomes and gene amplification. It is believed that the basic molecular

alteration in the stage of progression is the development of karyotypic instability. Unlike normal cells which regulate the structure of genome and karyotype during multiple cell division, the malignant cells are constantly changing the structure of their genomes (7, 8). The critical targets affected by the changes induced by carcinogenic agents are not clear. Studies have identified a few groups of candidates which are likely to be the targets for carcinogenic agents during initiation, promotion and progression. They are proto-oncogenes (9), cellular oncogenes (10) and tumor suppressor genes (11).

2. Oncogenic viruses

Oncogenic viruses are those viruses that can cause cancers in animal models and transformation of cultured cells and are widely used to study the development of neoplasm. Studies employing oncogenic viruses provide a number of advantages to researchers. First, some viruses induce cancers in susceptible animal very efficiently and rapidly. Tumors develop within days or weeks. Second, the genome size of oncogenic virus is much smaller (~10kb) than the genome of the mammalian cell (~10^6kb) and this offers scientists an opportunity to study the virus genome piece by piece. Third, viruses can be easily handled in the laboratory. The molecular characteristics of viruses have allowed researchers to identify specific viral genes responsible for oncogenic transformation and finally to identify cellular oncogenes.

Retroviruses

Since the discovery of the first retroviral oncogene, the src of Rous Sarcoma virus (RSV) in 1970 (12, 13), there has been an extraordinary proliferation of studies about the action of oncogenic viruses. The discovery of the SRC oncogene was the first step towards understanding the function of gene specifically involved in induction of abnormal cell proliferation. The SRC oncogene of Rous sarcoma virus (RSV) was found as both a retroviral oncogene and a cellular proto-oncogene. Unlike DNA viruses, retroviruses have an RNA genome. The viron RNA is reverse-transcribed by a virus -coded RNA dependent DNA polymerase. The viral DNA is integrated into host DNA to form the DNA provirus. The proviral DNA is a template, both for the replication of viral DNA and for the synthesis of viral RNA. Synthesis of viral proteins occurs mainly in the final stages of virus replication. Progeny virus particles are released by budding from the plasma membrane without cell lysis. Retroviruses may induce cell transformation in two ways: First, retroviruses pick up normal cell genes by virus-cell recombination events. The altered versions of the transduced proto-oncogenes are different from their normal cell progenitor both in function and structure. Second, retrovirus without oncogene induce cancer following long period in infected animals. In many cases, activation of cellular oncogenes by adjacent integration of proviral DNA is involved (15, 16).

3. DNA viruses

DNA tumor viruses contribute to the development of 20% of world cancers (17). Generally, oncogenic DNA viruses fall into five groups: papovaviruses, adenoviruses, papillomaviruses, herpesviruses, and poxviruses. Oncogenic DNA viruses, as carcinogenic stimuli, also cause genomic alterations in the cell affected by these microorganisms. During viral transformation whole or part of the viral genomes of hepatitis B viruses, papovaviruses and adenoviruses is incorporated into host cell genome (18, 19, 20). This step is critical to the neoplastic transformation. The inserted parts may work as proviruses or cause activation of oncogenes mentioned above. In infected cells, the genomes of papillomaviruses, herpesviruses and proviruses work as episomes in the nuclei of infected cells, instead of being incorporated into host genomes. It appears that the episomes may alter the cell replication without directly changing the structure of the genome (21). A few families of DNA viruses carry oncogenes whose products appears to be responsible for transformation of normal cells.

Papillomaviruses

Two genes of papillomavirus, E5 and E6 are identified for inducing transformation of cultured cells (22). The E5 gene product is a small peptide which might stimulate the activity of cell surface receptors for growth factors. The E5 gene of the bovine papillomavirus, BPV-1, is responsible for the major transforming activity of the virus, producing foci in cultured rodent cells. The E6 gene product can bind the p53 tumor

suppressor protein in the cell cytoplasm, thus preventing p53 from exerting its tumor-suppresser function in nucleus (24). Recent reports have shown E7 gene product binds the retinoblastoma protein (RB) (25).

Adenoviruses

Transformation with adenoviruses is mediated by the combined action of two early region genes, E1A and E1B. They both encode multiple protein products as a result of alternative splicing. E1B encodes two major proteins. The 19 kd polypeptide appears to prevent E1A induced cytolysis and DNA degradation (26). The larger E1B product of approximately 55 kd may form a stable complex with p53 which is then sequestered in the cytoplasm(27). The two E1A products, 26 kd and 32 kd, function to activate transcription of both viral and cellular genes and interact with the RB gene product (28). In addition to its well known transforming activities, E1A also possesses several activities which suppress transformation, tumorigenesis and malignant progression (metastasis) of tumor cells (29). It appears to have opposing functions: a dominant oncogene and a tumor suppressor gene.

Simian Virus 40

SV40 is the most well characterized DNA virus known today. It has been an invaluable system both for the study of molecular details of eucaryotic cell processes and virus mediated cell transformation. The small size of its genome and the

availability in vitro culture systems make SV40 serve as a model for the study of basic mechanisms underlying DNA virus-mediated carcinogenesis. Simian Virus 40 (SV40) was originally isolated from cultures of primary rhesus monkey kidney cells.

The genome of SV40 is a double-stranded, circular DNA molecule, 5243 base-pairs in length (Fig. 1). The genome is divided into two functionally distinct genetic regions of about equal size. Each half of the genome encodes a different set of viral functions.

The " early " region codes for two regulatory proteins, the large tumor antigen (T-ag) and small tumor antigen (t-ag). The "late " region codes for three structural proteins of the virus coat, VP1, VP2, and VP3. A fourth late protein, LP-1, is encoded by an open reading frame in the 5' leader region of the mRNA for VP1 (30). LP1 is a nonstructural protein that appears to function in viron assembly (31). the " early " and " late " regions are expressed before and after the onset of viral DNA replication. SV40 DNA is packaged *in vivo* as chromatin, composed of about 25 nucleosomes, forming a nucleoprotein complex.

Depending upon cellular environment, SV40 can adopt two different life styles. African green monkey kidney cells are permissive for SV40 and support a lytic infection. Infection of permissive cells results in cell death and the release of progeny virus particles. All viral functions are expressed and the cells are lysed. During lytic infection, the early region

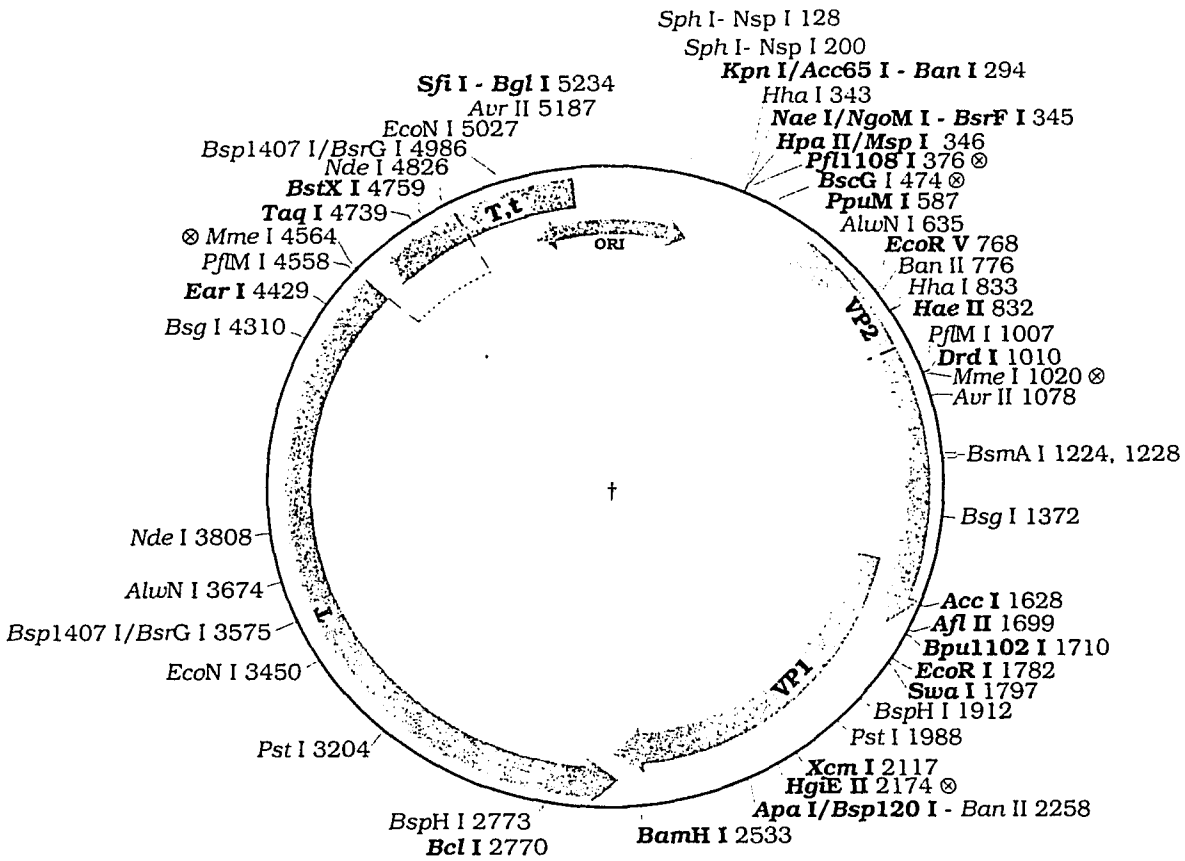
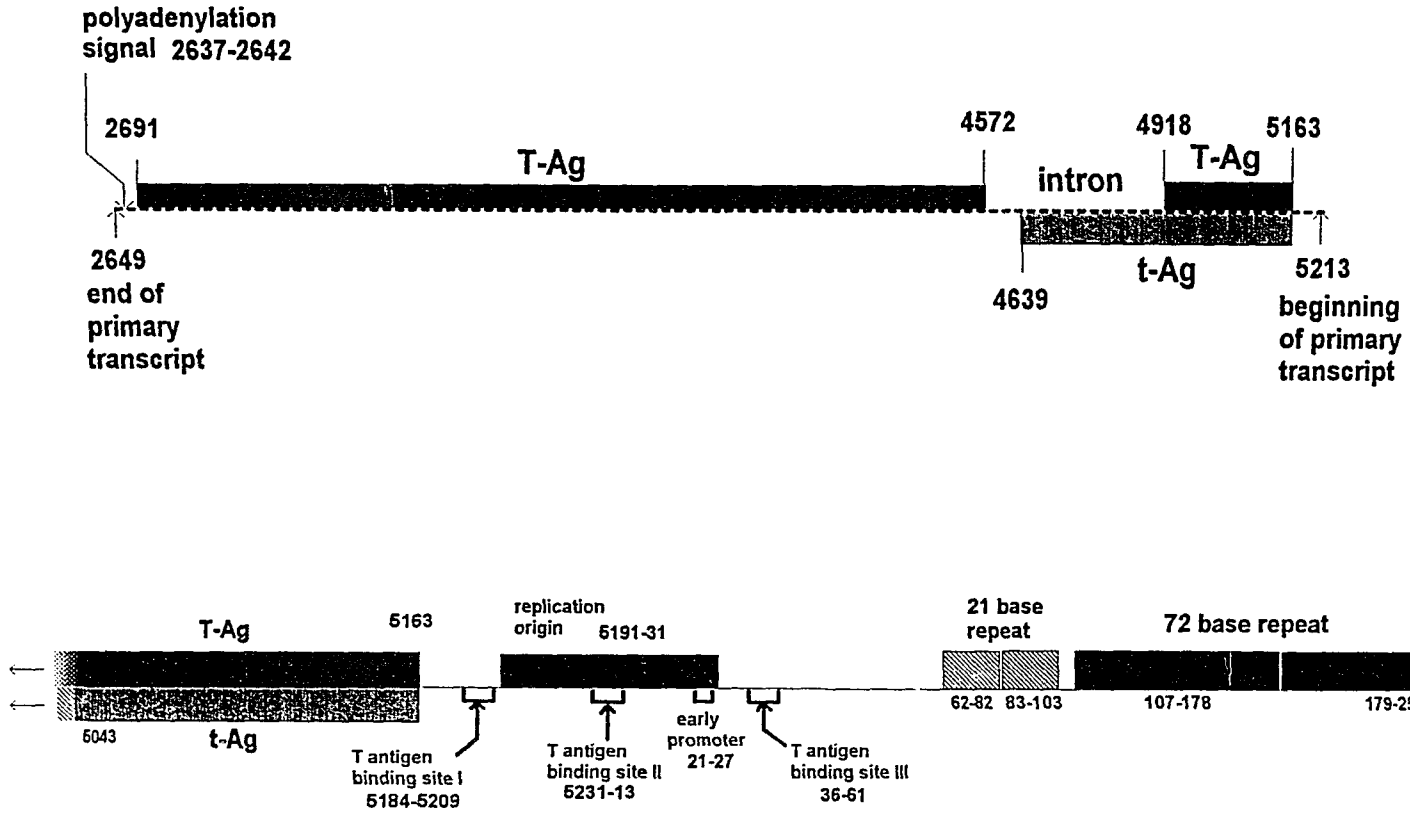


Figure 1

Physical and functional map of SV40 DNA. Nucleotide numbering (1-5243) starts clockwise from 5243/0 at the origin of replication around the molecule in the direction of late genes to early genes. The map shows the restriction sites of enzymes that cut the molecule. The map also shows the positions os the coding sequences and the origen of replication. The early genes (T and t) are transcribed clockwised. They are shown arranged in concentric rings to their reading frames. (from New England Biolab)

Figure 1. B
Shown below is a transcription map of T antigen and T antigen binding sites at and near the SV40 replication origion (from Dr. Steinberg).



proteins are expressed, which participate viral DNA synthesis and activation of transcription of viral late genes. T antigen plays the central role in initiation of viral DNA replication. T antigen binds to the origin, enters the duplex and catalyzes the ATP dependent unwinding of the DNA strands. The reaction forms two replication forks and the viral DNA is ready for elongation of nascent strands. The early transcription region encodes for mRNAs, large T antigen and small t antigen. When the level of large T antigen attains a critical level, large T antigen down regulates the early transcription and stimulates late transcription which includes synthesis of the structural proteins of the virus. Finally, a vacuolating process results in release of progeny virus particles and death of the host cell. Nonpermissive hosts are cell types in which SV40 fails to replicate and some cells, like human epithelial cells (33, 34) and rodent cells (35) , may be transformed by the virus. Only the early viral functions are expressed in transformed cells. Viral DNA is not replicated; progeny virons are not produced and cells are not lysed. Instead, the cells acquire an unlimited growth potential (i. e., immortalization), exhibit altered metabolic, morphologic, and growth properties (i. e., phenotypic transformation), and may be tumorigenic in vivo. In these cases, the viral genome is integrated into the genome of the infected cell. Integration occurs at many sites in the host cell genome, and the circular viral genome is cut at random sequences before or during integration.

TABLE 1

PROPERTIES AND FUNCTIONS OF SV40 T-ANTIGEN

A. Structural properties

Size:amino acid(81632Da); Mr 90,000-100,000

Modifications:

phosphorylation

N-terminal acetylation etc.

Supramolecular structure:

Monomers; Dimers; Larger homo-oligomers

B. Subcellular distribution

Nuclear:

Nucleoplasmic

Chromatin-bound

Nuclear matrix associated

Plasma membrane: NP40 soluble; NP40 insoluble (plasma membrane lamina)

C. Functions

Specific DNA binding (viral origin of replication)

Initiation of viral DNA replication

Autoregulation of viral early transcription

Induction of viral late transcription

ATPase activity

Complex formation with cellular protein P53 and pRb

Initiation and maintenance of cellular transformation

---

4. SV40 large T antigen and its functions in transformation

Large T antigen is one of the most intensely studied eukaryotic proteins because it performs an amazing variety of functions which are of interest to biologists studying replication, transcription and transformation.

T antigen is a phosphoprotein of 708 amino acids with several functional domains. Its properties and functions are briefly summarized in table 1.

Functions of T antigen in transformation

During the early phase of infection of SV40, the most important event is stimulation of host cell DNA synthesis, as viral DNA replication is dependent on the host cell DNA synthesis machinery. T antigen accumulates during the early phase of infection in the nucleus. It alters the cellular transcription pattern and stimulates quiescent cells to enter S phase, even an amino-terminal fragment of SV40 T antigen induces cellular DNA synthesis in quiescent cells (130). The mechanism of transformation by T antigen remains to be discovered. An amino-terminal domain (amino acids 1-82) is considered as a transforming domain by studies of T antigen mutations(129), but as yet no known fuction for this region of T antigen has been directly identified as been involved in transformation. A lot of work has been directed at trying to solve the puzzle. First, it is unlikely that the mechanism involves cellular genes activated by a promoter insertion event, since the integration sites for SV40 DNA into the cellular genome appear to be random (36). Several laboratories have used different deleted T antigens to

immortalize cell strains and found that different segments of T antigen appear to be dispensable for the immortalization of different cell strains, even within a single species (45, 52, 53, 54). A possible hypothesis is that T antigen may cause the immortalization through more than one pathway. There may be multiple routes open to T antigen. T antigen may participate in a key transformation - related reaction in one cell species and it could not do so in the other (55).

Tumor suppressor genes, also called antioncogenes, are wild type alleles of genes that play negative regulatory roles in cell proliferation, differentiation and other cellular processes. The inactivation, loss and disturbance of antioncogenes is oncogenic. The RB gene is a tumor suppresser gene. Molecular analysis has fully confirmed that retinoblastomas arise by the loss of both alleles at the same locus, the gene localized at region 13 q 14 on the human chromosome map. Tumor formation or transformation occurs when these genes are inactivated. In 1988, DeCaprio et al (56) observed that SV40 T antigen forms a specific complex with p110-114Rb, the product of the retinoblastoma gene. Complex formation was found to depend upon the integrity of the 105-114 amino acid sequence(T antigen area) of SV40, as does full transformation activity. Several reports(56, 57, 58, 131) suggest that T antigen may perform its transforming functions by modulating the Rb gene product. T antigen also binds to another antioncogene, p53(59, 60, 61). T antigen appears to interfere with the functions of antioncogenes, probably by competing with the normal cellular

target for antioncogenes. Studies also showed transfection with T-mutants defective in either the Rb or p53 binding domain yielded no morphological alterations characteristic of transformation(62, 63, 64). Those results are consistent with the idea mentioned previously, that T antigen as a multi functional oncogenic protein may use different pathways to transform cells from different sources.

5. Transformation of keratinocytes with SV40

In vitro cell culture systems have provided researchers with the ability to analyze and compare the properties of normal and cancer cells under controlled conditions. The basic distinction in vivo is that cancer cells proliferate under conditions where normal cells do not. In cell culture system, most cancer cells also show their growth which is lack of control. A primary distinction, between normal and neoplastic cells in culture is that normal cells display density-dependent control of cell growth. In contrast, neoplastic cell continue growing to much higher cell density. Second distinction between normal and neoplastic cells in culture is their requirement for serum growth factors. Neoplastic cells frequently display a reduced dependence upon extracellular growth factors. A third distinction between normal and neoplastic cells in culture is anchorage-dependent growth. Normal fibroblasts and epithelial cells must attack to a solid matrix in order to grow. Many transformed fibroblasts and epithelial cells can grow without attachment to a solid surface, which is anchorage independence.

A fourth distinction of normal and neoplastic cells in culture is immortality. Most normal cells have limited proliferative capacity in vitro. Most neoplastic cells can grow indefinitely in culture. Normal cells move across the surface of a culture dish, but stop such a movement upon contact with neighboring cells. Such contact inhibition is not displayed by tumor cells which continue moving after cell contact. Stratified squamous epithelium consists of many layers of cells, making this tissue relatively thick. This tissue forms the antermost layer of the skin (epidermis). The deepest layer of the epidermis, which are next to the dermis, are nourished by dermal blood vessels and capable of reproducing as the cells of this layer divide, older epidermal cells are slowly pushed away from the dermis toward the surface of the skin. At the same time the cells begin to undergo a hardening process called keratinization, during which strands of tough, fibrous, waterproof keratin develop within the cell membranes. Many layers of these tightly packed cells accumulate in the outer portion of the epidermis. This produces a covering of dry tough, protective material that prevents the escape of water from underlying tissues and the entrance of various microorganisms.

Carcinomas which arise from epithelial cells comprise about 90% of human cancers. That makes the study of transformation of epidermal keratinocytes very valuable model for understanding the development of epithelial cell neoplasm. The terminal differentiation program of epidermal keratinocytes takes place in cell culture (65, 66, 67) and the behavior of keratinocytes

in vitro closely resembles to their behavior *in vivo*. Thirty years ago, Koprowski et al (68) first reported the transformation of fibroblasts *in vitro* by SV40 virus. Fifteen years later, Steinberg and Defendi used SV40 to transform human keratinocytes *in vitro* after improved culture conditions for this cell type became available(66, 69). This virus-mediated transformation system has been well characterized with stage specific altered phenotypic properties (Fig. 2). Those properties include: the disruption of the normal pattern of growth and differentiation (69, 71, 132), loss of dependence on high serum concentration, the ability to form colonies in soft agar (70, 71), loss of polarity with respect to cell surface fibronectin distribution(72), changes in patterns of junctional communication (34, 46)and altered expression of various dermatologic antigens (73). The stage specific emergence of transformed properties over time is illustrated in Figure 2. The diagram shows that transformation of human keratinocytes by SV40 is a progressive process. It suggests that SV40 transformation of epithelial cells follow the paradigm of multistage of carcinogenesis. Different phenotypic properties emerge over many cell generations following infection by SV40. In the initial days following infection only a few cells in each of the keratinocyte colonies (comprising only about 1% of the total cell number exposed to virus) are T antigen positive. As the cultures are passaged, the percentage of T antigen positive cells progressively increases around passage 4 (30 cell generations) and most of the cells are T antigen positive.

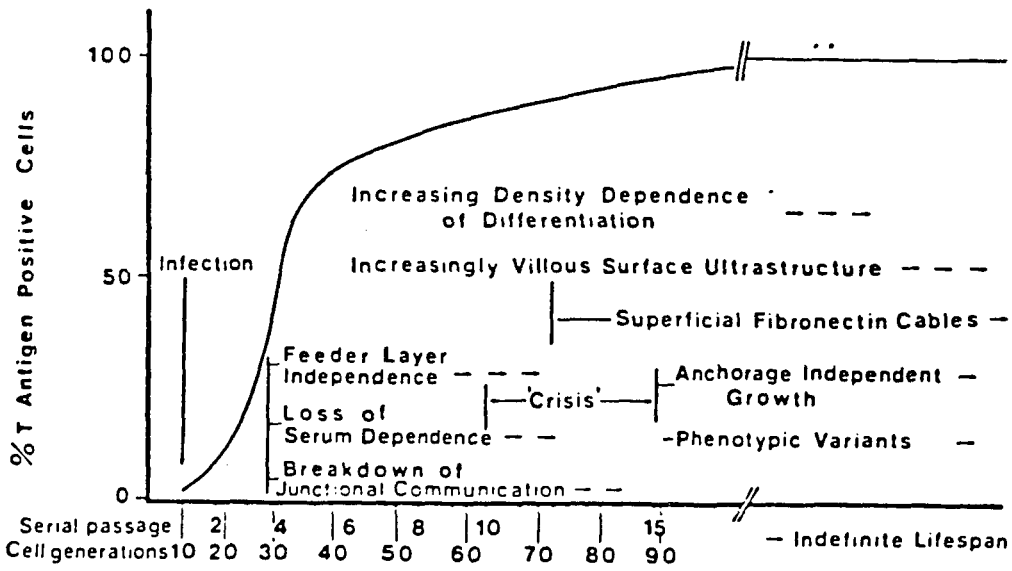


Figure 2

Diagram of the emergence of transformed properties over time following infection of human epidermal keratinocytes by SV40 (Defendi and Steinberg 1982)

Between passage 4 and 10, cells start to lose growth dependence upon high serum concentrations and become independent of the requirement for fibroblast feeder layers during this period colonies begin to lose the differetiated multilayers characteristic of normal keratinocytes. A period from passage 10 to 15 after SV40 infection is known as the crisis period. When keratinocytes enter this period, cells show lower growth potential, exhibit a greatly enlarged, bizarre appearance and are vacuolated. Many cells die at this point, probably due to the end of their natural lifespan (73). By the 15th passage, about 85% of the cells are T antigen positive. Cells that survive crisis are immortalized and have infinite growth potential. This survival rate is greatly enhanced if they are maintained at high cell densities (33, 69). The surviving cells can be established as transformed keratinocyte cell lines (23). Since cells were first transformed with SV40 thirty years ago, the critical targets of SV40 in transformed host cells still remain unclear.

In our laboratory recent studies have focused on the molecular changes in host cell gene expression which accompany transformation, e.g. changes in the pattern of expression keratins in SV40 transformed cells(74, 75, 76). Keratins are a family of intermediate filament proteins consisting of more than 30 polypeptides. According to their biochemical and immunochemical properties, they are divided into two subfamilies,

one acidic and ranging in size from about 40-65kd (type 1). SV40

transformed keratinocytes ultimately lose expression of keratins characteristic of stratified epithelia and express simple epithelial keratins instead. The cellular function of keratins other than structural elements remains unknown.

The mechanism of SV40 transformation of human epidermal keratinocytes appears to be complex. A single protein, T antigen, is implicated in a variety of alterations in host cells. The study of the pattern of biochemical changes during SV40 transformation may shed the light on the transformation process. One striking action of oncogenes is based upon regulation, or deregulation, of transcription. Modulation of gene expression is frequently an important ramification of intracellular signaling and plays a vital role in the control of cellular proliferation. T antigen may directly or indirectly influence transcription of cellular genes. In this laboratory we have previously demonstrated dramatically altered patterns of expression of keratins. To isolate cDNAs coding for keratin-like genes, poly A^+ RNAs derived from SV40-transformed keratinocytes were enriched for keratin sequences by size fractionation on a sucrose gradient. RNAs from gradient fractions were translated in vitro in a rabbit reticulocyte lysate system. The ^{35}S-methionine labeled products were immunoprecipitated using a polyclonal antiserum to human stratum corneum keratins and displayed by SDS/PAGE and autoradiography. C-tailed cDNAs made from poly A^+ RNAs sedimenting in fractions exhibiting the largest amount of keratin synthesis were ligated to G-tailed pBR322 plasmid vector and transformed into competent E. Coli

recipients. From the original library of 13,000 bacterial colonies 300 were picked for further study on the basis of hybridization to cDNA probes made from poly A^+ RNAs from the enriched sucrose gradient fractions.

Since we reasoned that novel mRNAs should be present in high abundance in the SV40 transformed keratinocytes and absent or reduced in normal epithelial cells and fibroblasts, cDNA probes made from total poly A^+ RNA from each of these cell types was used to screen nitrocellulose filters containing immobilized plasmid DNA from candidate clones. On this basis 25 candidates clones were isolated and believed to be induced by SV40. Because this system appears to represent a clear example of transformation related changes in gene expression, it is of interest to characterize these induced genes. The exciting observation is that those genes are not only induced to high levels of expression in viral transformants but also in several cell lines derived from squamous cell carcinomas. Those findings indicate that the induced genes are common features both in transformed and carcinomas cells. Two cDNA clones E12 and E13 from SV40 transformed keratinocytes were isolated and characterized to understand their role in transformation process in this thesis.

Materials and methods

1. Cell culture

Fresh human skin samples were received in DMEM culture medium containing ampicillin and streptomycin (from NYU nursery). Samples were washed once in 70% ethanol and then three times in DMEM medium. Samples were trimmed to get rid of fat and the dermal layer. Minced sample pieces were placed on the bottom of tissue culture flask and flasks were left upside down for 2 hour, while the minced pieces attached to the flask. The skin samples were cultured in DMEM supplemented with 10% fetal calf serum, 0.1 mg of cholera toxin (Sigma Chemicals) per liter, 0.01 mg of epidermal growth factor (Sigma Chemicals) per liter, 0.5 mg of hydrocortisone (Sigma Chemicals) per liter. SV40 transformed keratinocyte cell lines, squamous cell lines, HTB cell lines and Hela cells were cultured in DMEM medium supplemented with 10% fetal calf serum, 0.5 mg/liter hydrocortisone, 100 mg/liter penicillin , 50 mg/liter streptomycin and 0.0085% amphotericin.

2. Isolation of total RNA and mRNA

A modification of the NP40 lysis method (Morris et al. 1989) for isolation of cytoplasmic RNA was employed. Confluent cell cultures were harvested by trypsinization or scraping. Cell

pellets were washed twice with 1 x PBS and stored in liquid nitrogen until use. Before the cell pellet thawed, the pellet was suspended in ice cold lysis buffer (0.4M NaCL, 1.5mM $MgCL_2$, 10mM Tris pH 8.0, 0.5% NP40, 500 units RNasin/ml) and kept on ice for 5 min. Nuclei were spun down and the supernatant was extracted twice with phenol:chloroform (1:1) and once with chloroform. Then 1/10 volume of 3M NaAc and two volumes of ethanol were added to the mixture. Cytoplasmic RNA was kept at -70°C and precipitate was kept at 4°C until use. mRNAs were prepared by the method of Davis(80). Briefly oligo(dT) cellulose was suspended in loading B (20 mM Tris, pH 7.4, 1M NaCL, 1mM EDTA, 0.1%) and loaded to a plastic syringe up to 0.5 ml. The final packed column was washed with 3 ml NaCL, 5mM EDTA and then washed with water until the pH of the effluent was less than pH 8. Finally, the column was washed with 5 ml loading buffer A (40mM Tris, pH7.4, 1M NaCl, 1 mM EDTA and 1% SDS). The RNA sample (about 5 mg) in 0.5 ml H_2O was heated to 65°C for 5 min., 0.5 ml of prewarmed (65°C) buffer A was added and mixture was cooled to room temperature. the 1 ml sample was loaded onto a prepared column. The eluate was collected, heated to 65°C and then cooled to room temperature. The eluate was loaded again on the column and the column was washed with 5 ml of loading buffer B. The RNA was eluted with 1.5 elution buffer(10 mM Tris, pH 7.4, 1 mM EDTA, 0.05% of SDS) and 10 drop fractions were collected.

RNA-containing fractions were identified by fluorescence in ethidium bromide agarose plates. Fractions containing RNA were

pooled and 1/10 of volume of NaAc and 2.5 volume of ethanol was added to precipitate mRNA.

Formaldehyde gels and Northern blot hybridization

g of agarose were suspended in 42 ml H_2O and 5 ml MOPS buffer (0.2 M MOPS, 0.05 NaAc, 0.01 M EDTA) and suspended by heating in microwave. The agarose suspense was cooled to 50^{0}C and 2.7 ml of 37% formaldehyde and 4 µl ethidium bromide(10 mg/ml) to gel mixture. 10 µg of total RNA in loading buffer were heated at 90^{0}C for 2 min and loaded into each gel well after which electrophoresis was carried out.

Gel separated RNAs were transferred to nitrocellulose paper by blotting in 10x SSC(3M NaCL, 0.15M Na_3 Citrate). The filter was air dried and baked for 2 h in a oven at 80^{0}C. Prehybridization was performed in 2x PIPES buffer, 50% deionized formamide, 0.5% SDS, and 100 µg/ml sheared salmon sperm DNA for 2 hr at 42^{0}C. Hybridization was done in the same solution for overnight at 42^{0}C with a nick translated or end labeled probe. Filters were washed twice at 55^{0}C in 0.1x SSC solution, 0.1% SDS (Sodium Dodecyl Sulfate) for 20 min, then once in 0.1x SSC for another 20 min. To confirm the quantity and quality of the RNAs the blot which had been hybridized was washed in boiling 0.1x SSC for 2 min and kept at 90^{0}C for 20 min. The filter was rehybridized to a chicken actin probe(ATTC).

3. Isolation of Genomic DNA and southern blot

T75 cell culture flasks containing 10^8 cells were rinsed twice with 1 x PBS. 5 ml of extraction buffer (100 mM Tris, pH 7.9; 10

mM NaCl; 10 mM EDTA; 0.5% SDS [Sodium Dodecyl Sulfate]) and 500 µg of proteinase K were added to each flask. Flasks were incubated at 37^{0}C overnight and scraped gently to collect the lysate mixture. The DNA suspension was placed in a 50 ml centrifuge tube and extracted with an equal volume of a one to one phenol/chloroform:isoamyl alcohol (24:1). The 50 ml tubes were placed on a rotator and mixed gently for 10 min and then centrifuged at 3,000 rpm for 10 min. The supernatant was recovered and extracted with phenol/chloroform-isoamyl alcohol as DNA was precipitated with two volumes of ethanol, washed with 70% of ethanol, and dried for 15 minutes under a vacuum. DNA was dissolved in 2 ml of TE buffer (10 mM Tris, pH7.4 and 0.5 mM EDTA) and RNase A was added at a final concentration of 10µg/ml and incubated at 37^{0}C overnight. The phenol-chloroform extraction and ethanol precipitation was repeated. DNA pellets were resuspended in 0.5 ml of TE (pH 7.5) and dialyzed against TE (pH 7.5) for two days.

DNA samples were digested with proper restriction enzymes, electrophoresesed in 1% agarose gel. DNA gel was soaked in denaturant (0.5 M NaOH, 1 M NaCl) for 30 min. Replaced with fresh

denaturant for another 30 min. Then the denatured gel was soaked in neutriant (2M Tris pH 7.4, 1M NaCl) for 20 min. Changed the neutriant for another 20 min and blotted the gel to a nitrocellulose filter. The filter was air dried and baked for 2 h in a oven at 80^{0}C. Prehybridization was performed in 2x PIPES buffer, 50% deionized formamide, 0.5% SDS and 100 µg/ml sheared

salmon sperm DNA for 2 hr at 420C. Hybridization was performed in the same solution with addition of a nick translated or end labeled probe at 42^{0}C overnight. Filters were washed twice at 55^{0}C in 0.1x SSC solution (Sodium Chloride, Sodium Citrate), 0.1% SDS (Sodium Dodecyl Sulfate) for 20 min, then once in 0.1x SSC for another 20 min.

4. Construction of genomic libraries

To clone E12 and E13 containing fragments, genomic DNA libraries were constructed using BamH I and EcoR I digested DNAs from normal keratinocytes. 10 µg of digested DNA fragments were extracted with phenol-chloroform and then ligated into EMBL4 and EMBL3 bacteriophage vector DNA in a molar ratio of approximately 3:1 (insert to vector). The ligation reaction was carried out using T4 ligase in 15^{0}C water bath for 24 hours. The ligation results were checked by gel electrophoresis to observe the formation of high molecular weight complexes. Ligation mixtures were packaged *in vitro* using Gigapac packaging extract system (Stratagene), followed by a drop of chloroform and 500 µl of sterile SM buffer (100 mM NaCl; 17 mM $MgSO_4 \cdot 7H_2O$; 1 M Tris-Cl, PH7.5 and 2% Gelatin). The E.Coli host NM539 (Promega Biotec) was used to grow and title the phage recombinants in a 200 µl of suspension of host bacterial in 10 mM $MgSO_4$. The amount of packaged phage used for library screening was determined from the plaque forming unit (PFU) title. About 10^4 plaques were plated in a 150 mm plate. The mixture of phage and

bacteria was incubated for 15 min for absorption at 37^{0}C. 3 ml of top agar was added and spread evenly on bottom agar plate. The plates were incubated at 37^{0}C overnight. Recombinant phages, propagated in the E.Coli host NM539 (Promega Biotec), were screened by plaque hybridization using nick-translated E12 and E13 probes. Selected positive clones were isolated. Bacterial phage from a selected plaque was suspended in 1 ml of SM buffer. 20, 30, 40 and 50 µl of the recombinant phage were mixed with 200 µl NM539 and plated to get a complete lysate. 5 ml of SM buffer was added to a completely lysed plate and gently agitated for 2 hours at room temperature.

The phage stock solution was collected and stored at 4^{0}C. The ratio of stock solution:NM539 suspension was 1:25 for amplification of phage. Phage DNA was isolated from CsCl gradient banded phage in a ultracentrifuge rotor (55,000 rpm for 16 hours at 22^{0}C).

5. Deletion of the GC tails of cDNA clones

The GC tails added during cDNA construction form very strong secondary structures and abort the sequencing process. The Cyclone I biosystem (IBI) was used to delete the GC tails of selected cDNA clones.

2µg of single stranded DNAs of mp18 or mp19, containing inserts of selected clones, were annealed with 2µl of primers (Hind III 18-mer for mp18; RD22-mer for mp19) and digested with Hind III for mp18 and EcoR I for mp19 at 500C for 2 hour. The digestion mixtures were displayed on agarose gel to determine that the cleavages were complete. 2µl of T4 polymerase was added to the

reaction mixtures containing linearized DNA and incubated at 37^{0}C. 4µl aliquots of each mixture was removed at 20 second intervals and terminated by heating at 65^{0}C for 10 min. All the aliquots were combined into a single tube. The deleted single stranded DNAs were tailed with dATP or dGTP. The specific primers were annealed to the deleted products joining the two ends of the molecules. Finally, T4 ligase was added to seal the remaining nick. 10^5 transformants per µg of DNA was obtained and isolated plaques were picked. The sequence data were obtained se deleted cDNA clones.

6. Elongation of the 5' end of E12 gene

The modified RACE (rapid amplification of cDNA ends. Frohman, M. A..) method was used to get the full length cDNA of E12. The protocol is illustrated in Figure 3. Amplification of the 5' end of the E12 gene was performed by PCR reaction. 1µg of poly(A)+ RNA in 15.5µl H_2O and 2µg of total RNA in 15.5µl H_2O were heated for 3 min at 65^{0}C and quenched on ice, respectively. 2µl 10xRTC buffer, 0.5 µl of E12 specific primer 53 (5'-ACGAGACTCTAGAGAGGCAT-3'), 0.5µl of RNasin(Promega Biotec) and 40 units of Mo-MuLV reverse transcriptase were added to the RNA samples. The mixtures were incubated at 42^{0}C for 1 hour, and then at 52^{0}C for 30 min. The reaction mixtures were diluted to 500µl and stored at -20^{0}C. Excess primer was removed by passing through a sephacryl S-400 spun column (Promega). The volumes of the cDNA solutions were reduced to 10µl by vacuum evaporation in a Speed Vac. 4µl of 5x tailing buffer(IBI), 4µl of 1 mM dATP,

and 10 units of terminal transferase were added to the cDNA mixture and incubated for 5 min at 37°C, and then for 5 min at 65°C. The tailed cDNA mixtures were diluted to 500µl with TE. PCR reactions were carried out in reaction mixtures containing: 5µl of 10xPCR buffer, 5µl of DMSO, 5µl of dNTPs (15 mM each), 27µl of H_2O, 3µl of primers(1µl of each primer,

5'-GACTCGAGTCGACATCGAT_{17}-3'

5'GACTCGAGTCGACATCGAT-3'

5'-TCAGATCTGGTGACAACACT-3')

and 5µl of the appropriate cDNA pool. The mixtures were denatured at 950C for 5 min and then cooled to 72°C. 0.5 µl of Taq polymerase was added and the mixtures were overlayed with 30µl of mineral oil. The primers were annealed at 55°C for 5 min and extended at 72°C for 40 min. Each PCR cycle was: 95°C for 40 seconds/55°C for 1 min/72°C for 3 min and amplification was carried out for 40 cycles. The reactions were extended at 72°C for 15 min.

7. DNA amplification with samples in remelted agarose

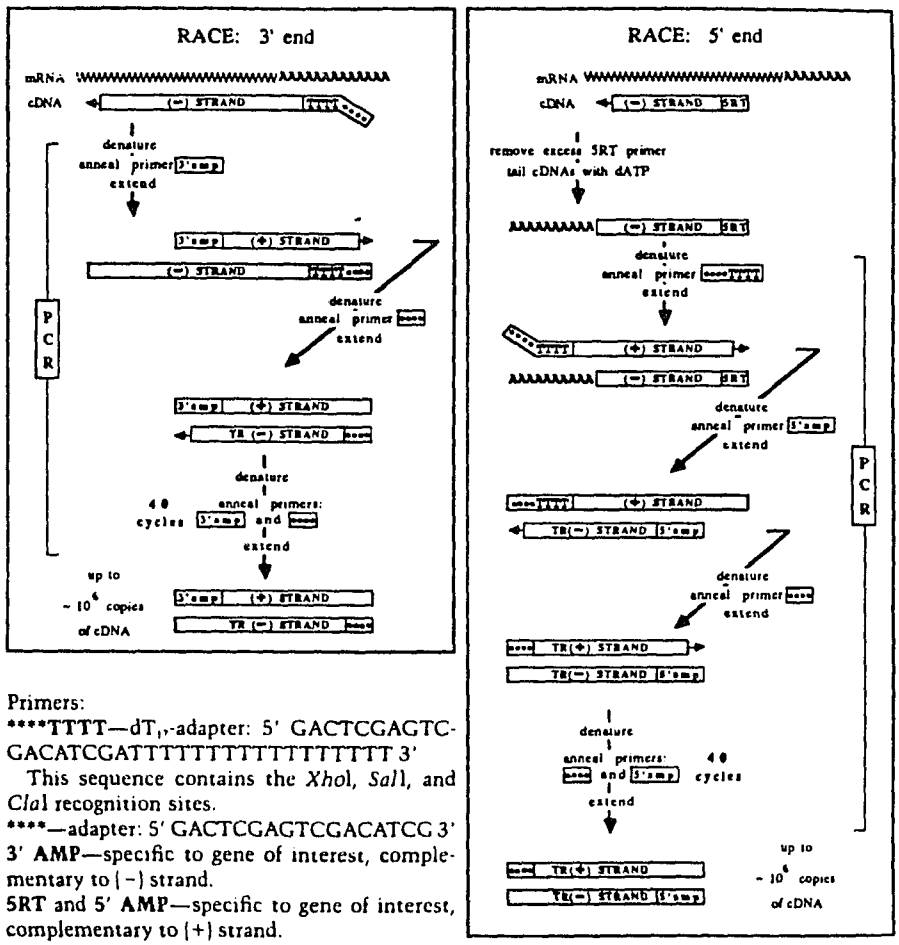
White's protocol was used to amplify the PCR products. PCR products were separated on a low melting point agarose gel and the band of interest was excised and melted at 65°C for 10 min. The liquified gel was diluted to a final DNA concentration of approximately 0.1ng/µl and stored at -20°C to be used as a DNA template. 5ng of the DNA template was used in each PCR reaction. The amplification reactions were performed for 30 cycles.

8. Computer Analyses

The LFASTA program of Pearson and Lipman (94) was used for sequence homology searches against the Genbank database. The PCGENE program (Intelligenetics) was also employed for homology searches of putative polypeptides from open reading frame against the Swiss Protein Database. GBWARE program is employed to analyze the properties of the putative E12 protein.

Figure 3

Schematic representation of the RACE protocol.



Primers:
****TTTT—dT_{17}-adapter: 5' GACTCGAGTCGACATCGATTTTTTTTTTTTTTTT 3'
This sequence contains the *Xho*I, *Sal*I, and *Cla*I recognition sites.
****—adapter: 5' GACTCGAGTCGACATCG 3'
3' AMP—specific to gene of interest, complementary to (−) strand.
5RT and **5' AMP**—specific to gene of interest, complementary to (+) strand.

Schematic representation of the RACE protocol. Open rectangles represent DNA strands actively being synthesized; shaded rectangles represent DNA previously synthesized. At each step the diagram is simplified to illustrate only how the new product formed during the previous step is utilized. A (−) or (+) strand is designated as "truncated" (TR) when it is shorter than the original (−) or (+) strand, respectively.

Results:

1. Expression of novel sequences in SV40 transformed and human carcinoma cell lines.

Rescreening of novel sequences induced in SV40 transformed keratinocytes.

Differential screening of cDNA libraries in previous work in our laboratory led to the isolation of cDNAs whose expression might be induced in SV40 transformed cells and not detected in normal cultured keratinocytes (78). To confirm the identity of those clones representing mRNAs specifically induced in SV40 transformed cells, additional northern blot hybridization studies were conducted. The inserts of cDNA clones were separated on low melting agarose gels and isolated by phenol extraction. Each of the cloned inserts was labeled with ^{32}P by nick translation and used as a probe. The first set of northern blots was performed to determine if those inserts were specifically induced in SV40 transformed cell line. Cytoplasmic RNAs from normal cultured keratinocytes (EP) cells and SV40 transformed cell lines were hybridized to cDNA probes (Figures 4, 5, and 15). The results of these northern blots showed that cDNAs, E4, E11, E13, E5, E14, E15 and H2 appeared to show expression in SV40 transformed keratinocyte cells but not in normal cultured keratinocytes. The expression of E12 was induced

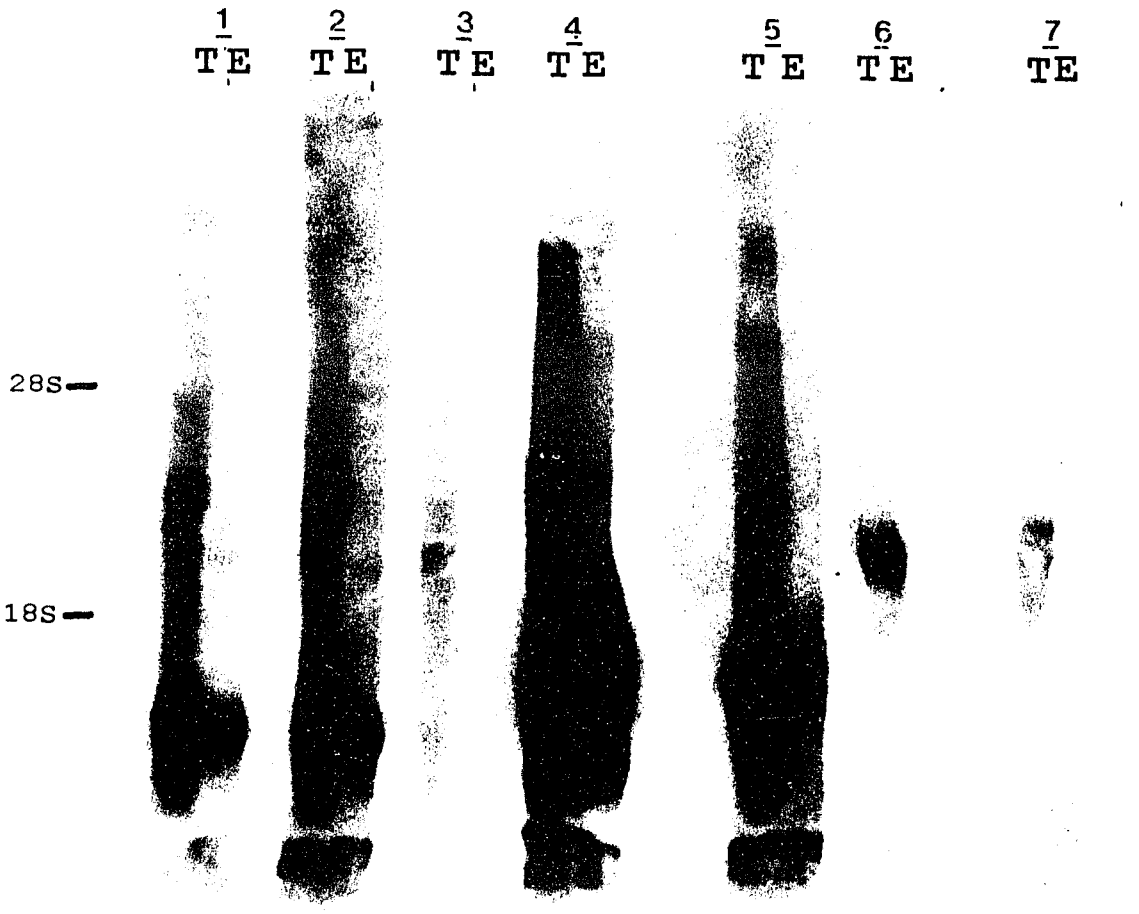


Figure 4:

Northern blot hybridization of various candidate cDNAs to total RNAs from normal and transformed cells. Cytoplasmic RNA samples were run on denaturing agarose gels, transferred to nitrocellulose and hybridized to ^{32}P-labeled nick translated cDNA probes. cDNA probes, E1(1), E2(2), E4(3), E9(4), E10(5), E11(6), E13(7), were hybridized to total RNAs derived from normal cultured epidermal keratinocytes (E) and SV40 transformed keratinocyte cell line 425 (T).

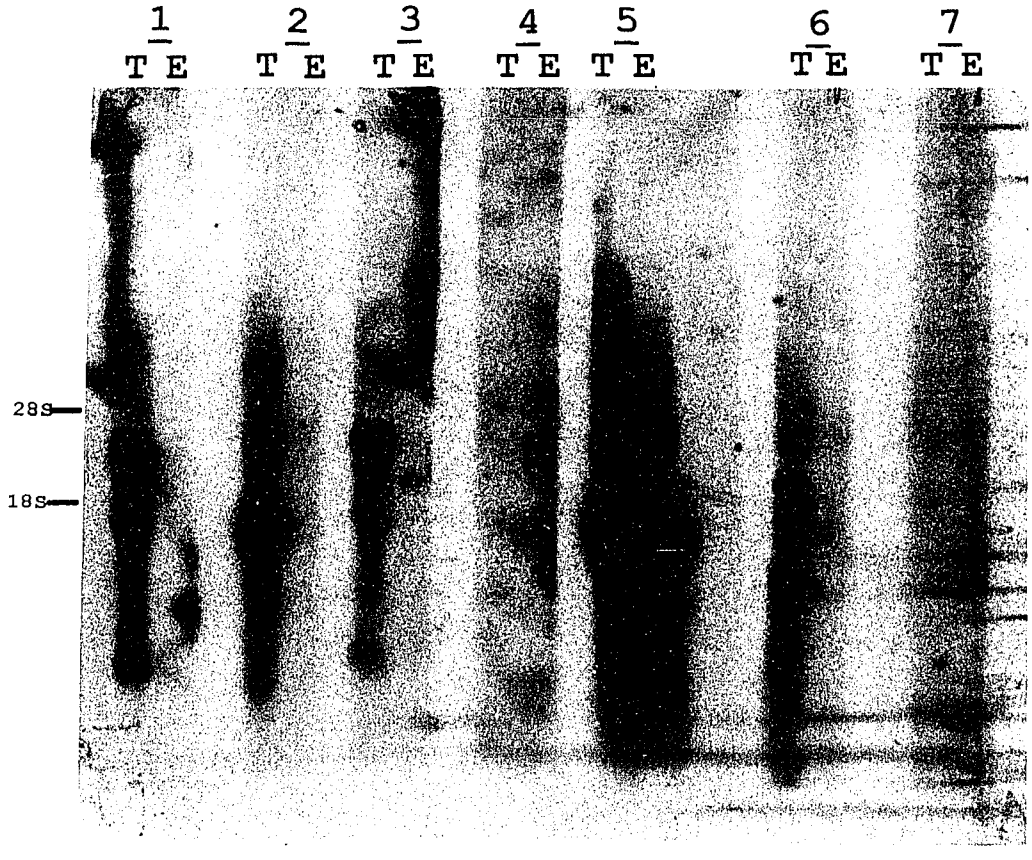


Figure 5:

Northern blot hybridization of various candidate cDNAs to total RNAs from normal and transformed cells as in Figure 4. cDNA probes, E5(1), E14(2), E15(3), E17(4), E18(5), H2(6), G15(7), were hybridized against total RNAs derived from normal cultured epidermal keratinocytes (E) and SV40 transformed keratinocyte cell line 425 (T).

to much higher levels in SV40 transformed cells than it is in normal cultured keratinocytes. These clones were selected for further study to determine if they were not only induced in transformed keratinocyte cells but also induced in cell lines derived from carcinoma tumors. Cytoplasmic RNAs from EP cells, SV40 transformed cell lines, squamous cell carcinoma (SCC) cell lines, Hela cells and breast carcinoma (HTB) cell lines were used to hybridize against these selected cDNA probes.cDNA clones E5 (Fig. 6), E11 (Fig.7), E14 (Fig.8), E15 (Fig.9), H2 (Fig. 10) and E13 (Fig. 11) were not detected in normal keratinocytes but were induced both in SV40 transformed cell lines and in human carcinoma cell lines. Figure 12 shows that a major band of about 1.2 k was detected by E12. The expression of E12 was detected in very low level in normal cells (EP) but present in dramatically higher levels in SV40 transformed cell lines and human cancer cell lines(SCC25, SCC324, Hela, Mcf7, and HTB3, 37, 182). Figure 13 shows RNA samples on the formaldehyde gel which was used for the blot used in Fig. 12; this blot hybridized to an actin gene probe to demonstrate equal sample loading and the quality of the RNA samples in Figure 14. Figure 15 shows that the expression of E12 was induced slightly even at the first passage following infection but was greatly induced at the second serial passage. The expression of E12 slightly increased in passage 1 and highly increased in passage 2.The high expression of E12 was maintained in long term transformants (passage 86). Figure 16 shows the RNA samples on the formaldehyde gel used for the blot in Fig. 15 was

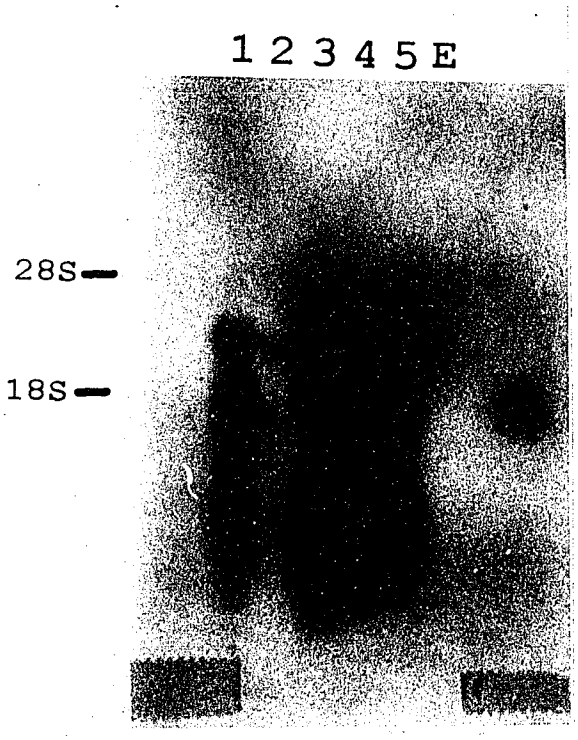


Figure 6:

Expression of cDNA E5 in normal, transformed, and tumor cell lines. The nick translated probe, cDNA E5, was hybridized against total RNAs derived from normal epidermal keratinocytes (E); SV40 transformed cell lines: 98(4), 425 (1); carcinoma cell lines: Hela (2), Mcf7(3), SCC 324 (5)

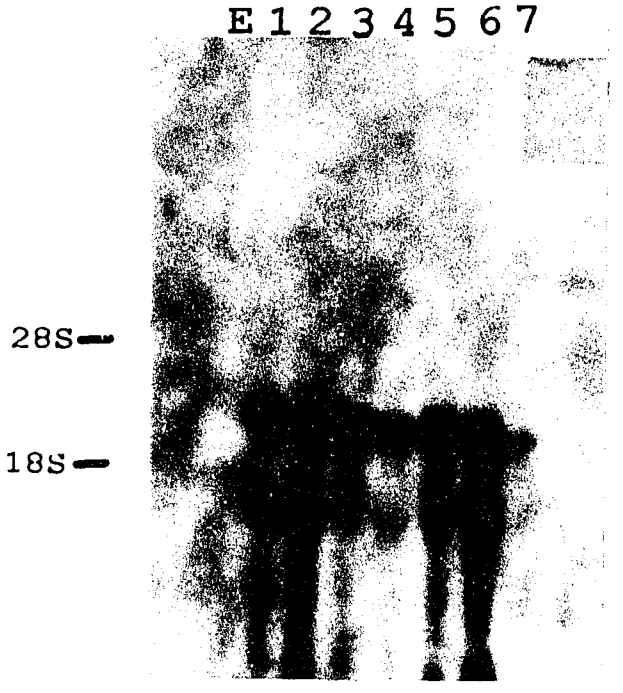


Figure 7:

Expression of cDNA E11 in normal, transformed, and cancer cells. The nick translated probe, cDNA E11, was hybridized against total RNAs from normal epidermal keratinocytes (E); SV40 transformed keratinocytes: 98(5), 130(6), 425(7); carcinoma cell lines: Hela (1), Mcf7(2), SCC 25(3), SCC 324(4).

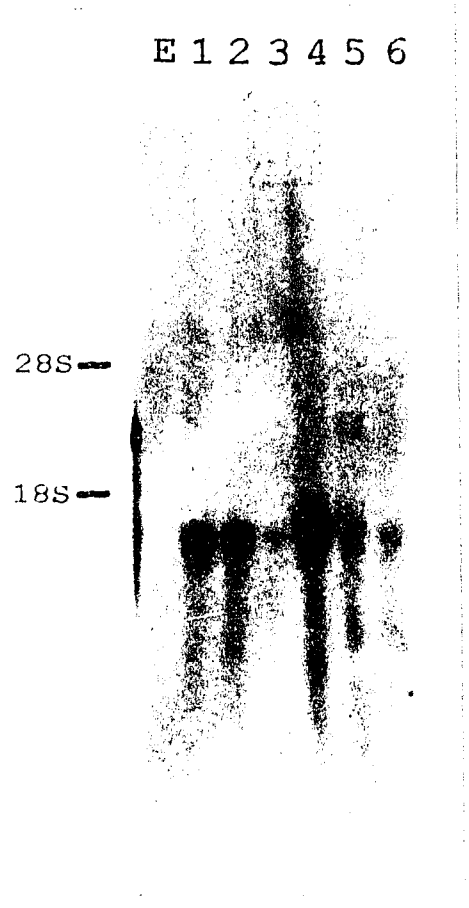


Figure 8:

Expression of cDNA E14 in normal, transformed, and cancer cells. The nick translated probe, cDNA E14, was hybridized against total RNAs from normal cells: epidermal keratinocytes (E); SV40 transformed keratinocytes: 98(1), 425(2); carcinoma cell lines: Hela (3), Mcf7(4), SCC 25(5), SCC 324 (6).

Figure 9:

Expression of cDNA E15 in normal, transformed, and cancer cells. The nick translated probe, cDNA E15, was hybridized against total RNAs from normal epidermal keratinocytes (E); SV40 transformed keratinocytes: 98(1), 425(2); carcinoma cell lines: Hela (3), Mcf7 (4), SCC 25(5), SCC 324 (6).

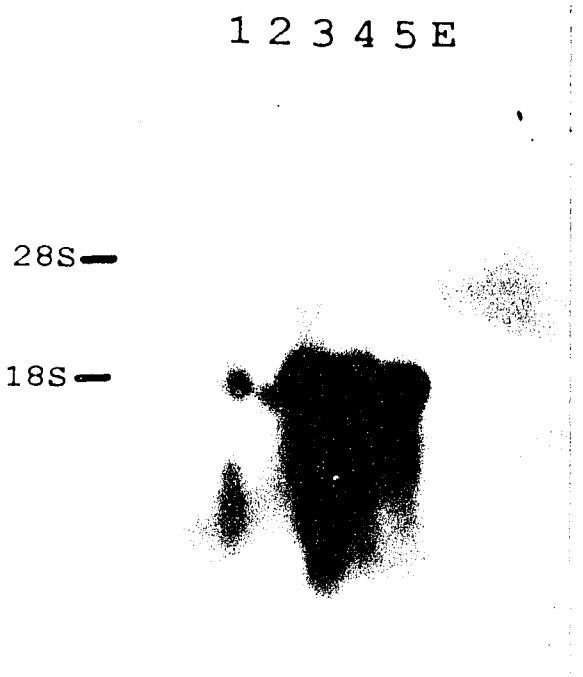


Figure 10:

Expression of cDNA H2 in normal, transformed, and cancer cells. The nick translated probe, cDNA H2, was hybridized against total RNAs from normal epidermal keratinocytes (E); SV40 transformed keratinocytes: 98(5), 130(6), 425(7); carcinoma cell lines: Hela (1), Mcf7(2), SCC 25(3), SCC 324(4).

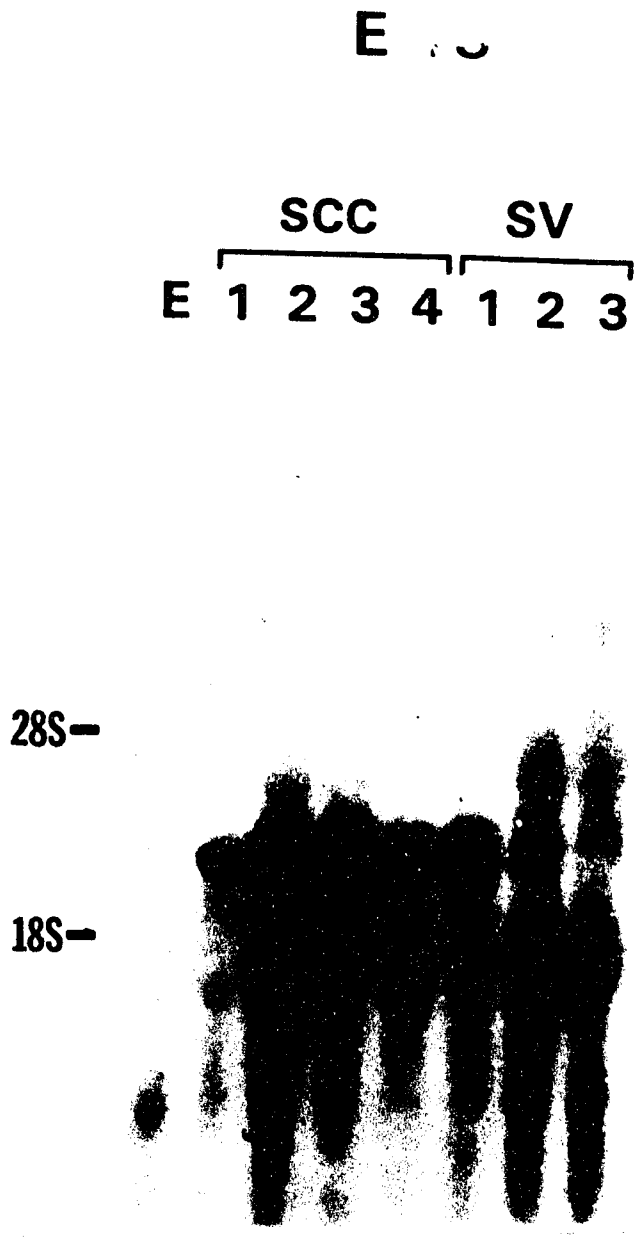


Figure 11

Expression of E13 cDNA in normal, transformed and carcinoma cells. The E13 probe was hybridized against total RNAs derived from normal cultured epidermal keratinocytes (E); carcinoma cell lines: SCC 1 , SCC 2 , SCC 3 , SCC 4 ; and SV40 transformed cell lines: SV 1 , SV 2 , SV 3.

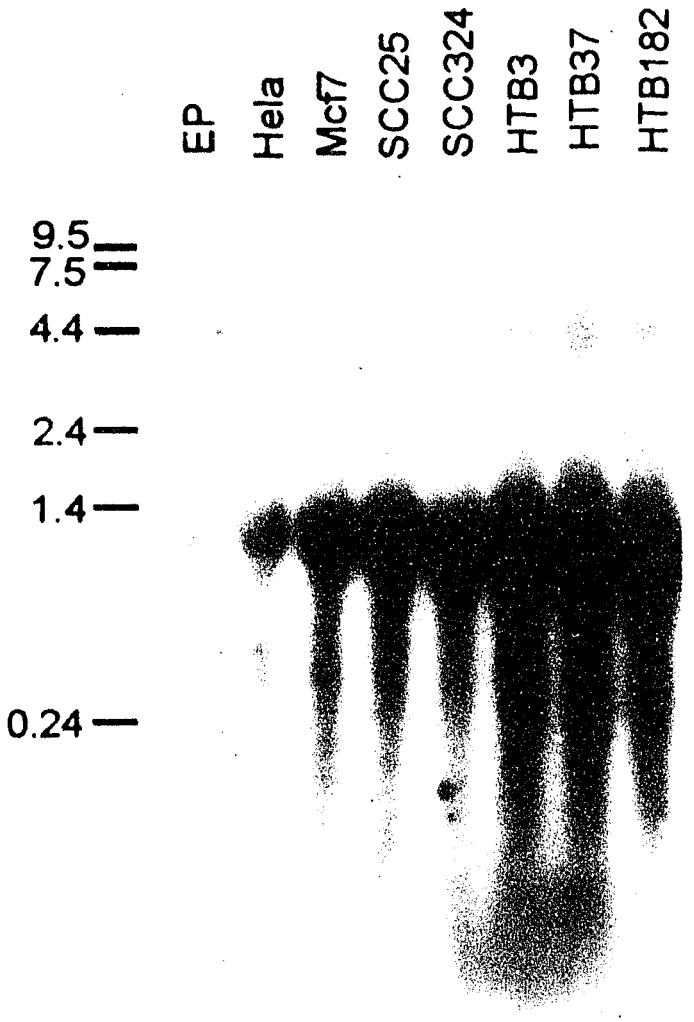


Figure 12:

Expression of E12 cDNA in normal, transformed and carcinoma cells. The E12 probe was hybridized against total RNAs derived from normal cultured epidermal keratinocytes (EP); and carcinoma cell lines: Hela (cervical), mcf7 (breast), SCC25 (oral), SCC324 (facial), HTB 3(bladder), HTB 37 (colon), HTB 182 (lung).

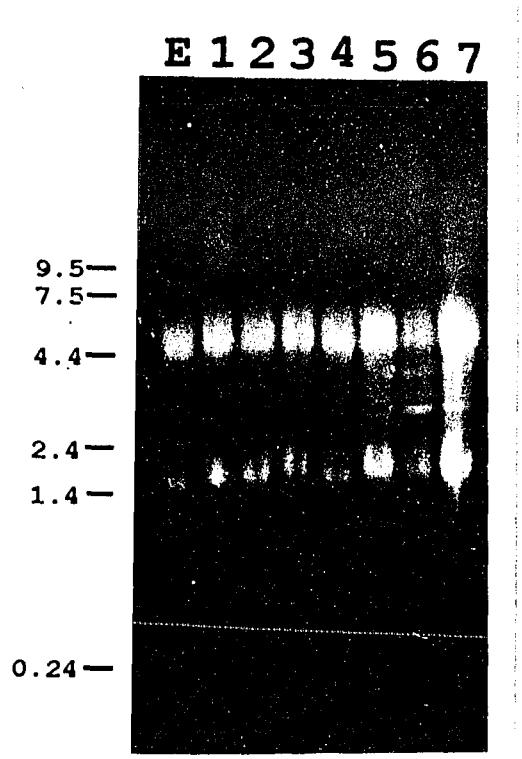


Figure 13:

Cytoplasmic RNAs on an agarose gel containing methymercury hydroxide. Cytoplasmic RNAs were isolated from EP (E), Hela (lane 1), mcf7 7(lane 2), SCC 25(lane 3), SCC 324(lane 4), HTB 3(lane 5), HTB 37(lane 6) and HTB182(lane 7), and run on formaldehyde agarose gel with methymercury hydroxide. Sizes of markers are indicated in kb.

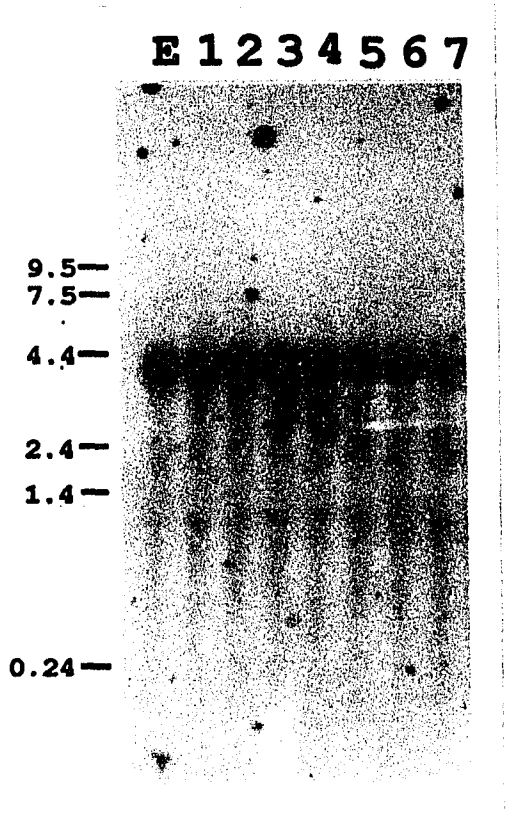


Figure 14:

Rehybridization of an actin probe to the northern blot shown in Fig. 12. The hybridization filter used in Fig. 12 was washed with boiling 0.2x SSC containing 0.2% SDS. The actin probe was hybridized against total RNAs derived from normal cultured epidermal keratinocytes (EP); and carcinoma cell lines as shown in Fig. 12.

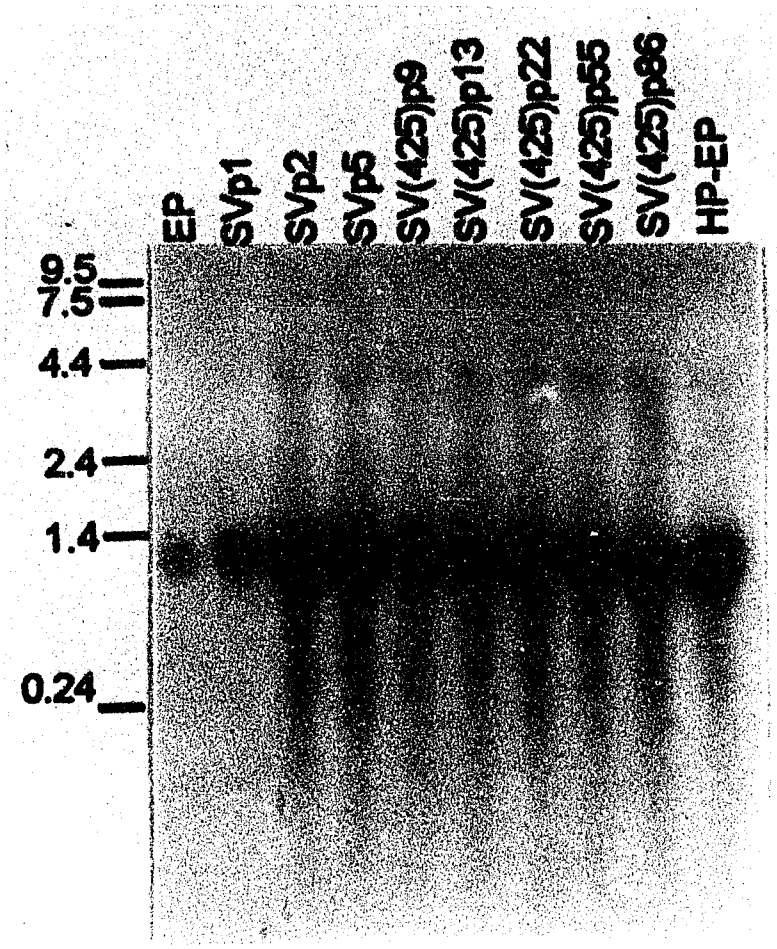


Figure 15:

Expression of E12 cDNA in normal and SV40 transformed keratinocytes at various serial passage levels. The E12 probe was hybridized against total RNAs derived from normal cultured epidermal keratinocytes, SV40 transformed keratinocytes at passages 1, 2, 5, 9, 13, 22, 55 and 86; and Hp-EP. A continuos line of cervical epithelial cells immortalized by HPV-16 supplied by Dr. V. Defendi (Dept. of Pathology, NYU medical Center).

hybridized to an actin gene probe to demonstrate equal loading and the quality of RNA samples used in Figure 17.

2. Sequencing of selected cDNA clones

To characterize the cDNA clones induced during SV40 transformation, sequencing of these cDNA clones were carried out to search for meaningful homologies between sequences of the selected cDNA clones and sequences catalogued in the GENBANK data base. Those inserts of cDNA clones were flanked with extensive G-C tails on both ends during the construction of the cDNA libraries and the G-C tails at both ends of the inserts form very strong secondary structures which block DNA polymerase and cause aborted sequencing reactions. Even at 90°C Taq polymerase failed to produce readable sequence data within the secondary structure region. Therefore, the Cyclone I Biosystem was used to delete the GC tails of selected cDNA clones. The inserts of cDNA clones were subcloned into the M13 phages, mp18 and mp19 and single stranded DNAs were isolated. An oligomer containing the sequence of a polylinker enzyme cutting site was annealed to the single stranded of mp18 or mp19 into which the insert had been cloned. The annealed hybrid was then digested with an appropriate restriction enzyme, EcoR 1 for mp19 and Hind III for mp18. The presence of recognition sites elsewhere in the insert dose not lead to digestion since only the double stranded region is digestible. The linearized molecules were

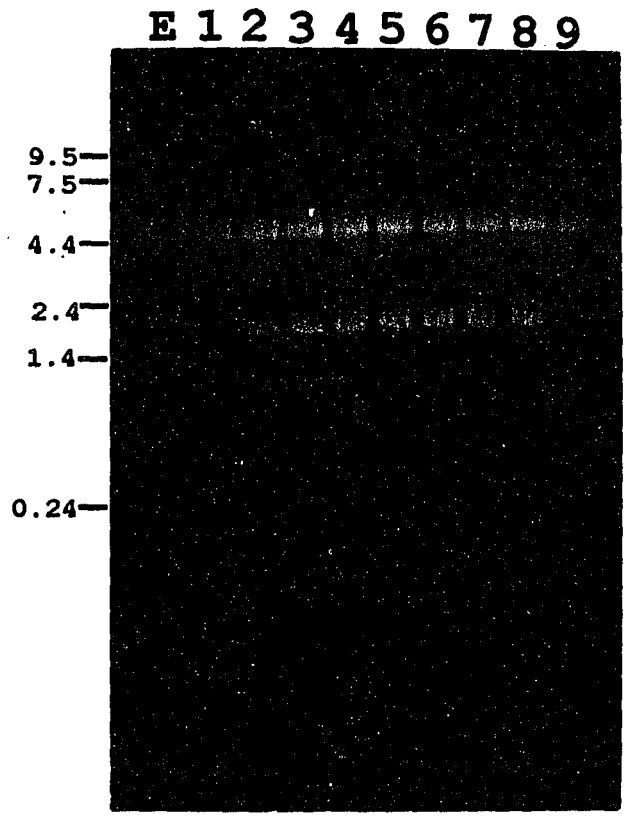


Figure 16:

Cytoplasmic RNAs run on a denaturing agarose gel containing methymercury hydroxide. RNAs were isolated from normal cultured keratinocytes (EP cell); SV40 transformed keratinocytes were tested at various serial passage levels: lane 1. passage 1; lane 2. passage 2; lane 3. passage 5; lane 4. passage 9, lane 5. passage 13; lane 6. passage 22; lane 7. passage 55; lane 8. passage 86; and lane 9. Hp-EP (HPv-16 transformed keratinocytes).

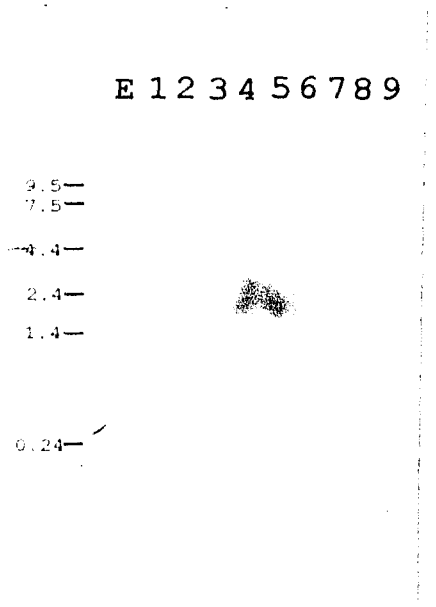


Figure 17:

Hybridization of actin to a northern blot shown in Fig. 15. The northern blot hybridization filter used in Fig. 15 was washed with boiling 0.2 x SSC solution containing 0.2% SDS. The actin probe was hybridized against total RNAs derived from normal cultured epidermal keratinocytes, SV40 transformed keratinocytes as shown in Fig. 16.

then selectivelydigested by T4 polymerase to generate an overlapping set of subclones without GC tails. The deleted clones were transformed into JM101 and sequenced by using standard dideoxy method. Sequences of E5, E11, E12, E13, E14, E15 and H2 were all analyzed for sequence homology against GENBANK catalogued sequences using the FASTA algovithw of Pearson and Lipman (94). These homology searches showed that E5 and E11 exhibited strong homology to some 23S rRNAs (Fig. 18, 19) while H2 had high homology with some 16S rRNAs (Fig. 20). Clone E14 was determined to be identical to PG-1 which had been studied and published previously(78).Neither E12 nor E13 fragments were found to exhibit significant homology to any other sequences catalogued in the GENBANK database at that time and further studies were performed. Recent searches of GENBANK has revealed identity to a randomly amplified fragment of a human mRNA designated as clone HM02E05 from a large scale cDNA sequencing for analysis of quantitative aspects gene expression (128). HM02E05 is a 223 bp fragment of the 3' end of a cDNA which overlaps the 3' end of E12 (Fig. 21). HM02E05, the cDNA fragment cloned from human liver, therefore appears to represent the same gene with E12 expressed in human keratinocytes.

Figure 18

Figure 18 A shows Partial sequence of E5. B shows homology between the E5 sequence and the TOBCPRG23 sequence which codes for a tobacco 23S ribosome gene. The LFASTA program is used for the homology search.

A. Partial sequence of E5

CGAGACCACGTTGCGAATTGAAACATCTTAGTAGCAACAGGAAAAGAAAATAAATAATGATTTCGTCAGTAGTGGCGA
GCGAAACAGAAGAGCCCAAACCTGTAAAAAGGGGTTGTAGGACATCTTACATTGAGTTACAAAATTTTATGATAGTAG
AAGAAGTTGGAAAGCTTCAACATAGAAGGTGATATTCCTGTATACGAAATCATAAAATCTCATAGATGTATCCTGAGT
AGGGCGGGGCACGTGAAACCCTGTCTGAATCTGCCGGGACCACCCGTAAGGCTAAATACTAAT

B. Homology between E5 and TOBCPRG23

75% identity in 298 nt overlap

```
                10        20        30        40        50
e5.    CGAGACCACGTTGCGAATTGAAACATCTTAGT-AGCAACAGGAAAAGAAAATAAATAATG
        ::::: :: : ::::: X:::::::::::::: ::: : ::::::::::::  ::: :  :
TOBCPR AGAGACAACCTGGCGAACTGAAACATCTTAGTGAGCCAGAGGAAAAGAAAGCAAA-AGCG
      260         270       280       290       300       310
      60        70        80        90        100       110
e5.    ATTTCGTCAGTAGTGGCGAGCGAAACAGAAG-AGCCCAAACCTGTAAAAAGGGGTTGTAG
        ::: :   ::::: ::::::::::::  : :: :::: ::::: :: :::: ::::::X :
TOBCPR ATTCCCGTAGTAGCGGCGAGCGAAATGGGAGCAGCCTAAACC-GTGAAAACGGGTTGTGG
      320       330       340       350       360       370
      120       130       140       150       160       170
e5.    GACATC-TTACATTGAGTTACAAAATTTTATGATAGTAGAAGAAGTTGGAAAGCTTCAAC
        :: : :  ::::  : ::          :  :: :::  ::::: :   ::: ::: :: :
TOBCPR GAGAGCAATACA-AGCGT-------CGTGCTGCTAGGCGAAGACGCCCGAATGCTGCACC
       380       390              400       410       420
       180       190       200       210       220       230
e5.    ATAGAAGGTGATATTCCTGTATACGAAATCAT-AAAATCTCATAGATGTATCCTGAGTAG
        :::: :: :: : ::: :::  ::::: ::: :  : :: :: : : :  :: ::::::
TOBCPR CTAGATGGCGAAAGTCCAGTAGCCGAAAGCATCACTAGCTTAT-GCTCTGACCCGAGTAG
     430       440       450       460       470       480
       240       250       260       270       280       290
e5.    GGCGGGGCACGTGAAACCCTGTCTGAATCTGCCGGGACCACC-CGTAAGGCTAAATACTA
         :::::::::: :: :: :: ::::::  :  ::::::::  : ::::::::::::
TOBCPR CATGGGGCACGTGGAATCCCGTGTGAATCGACAAGGACCACCTTGCAAGGCTAAATACTC
      490       500       510       520       530       540
```

Figure 19

Partial sequence of E11
CAGGATACTCTATGAGATTTTATGATTTCGTATACGGAATATCACCTTCTATGTTGAAGC
TTTCCAACTTCTTCTACTATATAAAATTTTGTAACTCAATGTAAGATGTCCTACACCCCT
TTTTACAGGTTTGGGCTCTTTCGCTTCGCTCGCCACTACTGACGAAATCATTATTTATTT
TCTTTTCCTGTTGCTACTAAGATGTTTCAATT

Results of homology search for E11
LEPRRN23S 83.7% identity in 86 nt overlap

```
            10        20        30        40        50        60
e11.in AATTGAAACATCTTAGTAGCAACAGGAAAAGAAAATAAATAATGATTTCGTCAGTAGTGG
       X:::::::::::::::::: :   :  :::::::::: ^v ::::::: ::::::::: ::
LEPRRN AATTGAAACATCTTAGTACCCGGAATAAAAGAAAAT-AAGAATGATTCCGTCAGTAGCGG
           180       190       200       210       220       230

                70        80        90       100       110
e11.in CGAGCGAA-GCGAAAGAGCCCAAACCTGTAAAAAGGGGTGTAGGACATCTTACATTGAGT
        ::::::: ::: ::::::: :::::
LEPRRN TGAGCGAACGCGGAAGAGCCTAAACCCCTGCCTACGTTACAGATCTAATTCGCTGTAGCA
           240       250       260       270       280       290
```

RUARRN23S
75.5% identity in 94 nt overlap

```
             10        20        30        40        50        60
e11.in AATTGAAACATCTTAGTAGCAACAGGAAAAGAAAATAAATAATGATTTCGTCAGTAGTGG
       :: X::::::::: :::: :   :::::::: :::: ^v    :::::: : ::::::::
RUARRN AAGTGAAACATCTCAGTACCCCGAGGAAAAG-AAATCAACCGAGATTTCCTAAGTAGTGG
        180       190       200       210       220       230

                70        80         90       100       110
e11.in CGAGCGAA-GCGAAAGAGCCC-AAACCTGTAAAAAGGGGTGTAGGACATCTTACATTGAG
       ::: :::: : :::  ::::: :::  :::: ::
RUARRN CGAACGAACGGGAAGCAGCCCAAAAGTTGTATAAGTCATAGTTGAATTATCTGGAAAGGT
         240       250       260       270       280       290
```

Figure 20

Partial sequence of H2

```
GGGGAGCTCAGGATGAACGCTCGCTGTGTGCCTAATACATGCATGTTGAACGGGATGTAGCA
ATACATTCAGTAGCGAATGGGTGAGTAACACGTACCTAACCTACCTTTAAGACTGGGATAAC
TATTGGAAACAATAGCTAATACCGGATATAGTTATTTATCGCATGATGAGTAATAGAAAGGA
GCTTCACAGCTTCACTTAAAAATGGGGTGCGGAACATTAGTTAGTTGGTAGGGTAATGGCCT
ACCAAGACGATGATGTTTAGCCGGGCCGAGACGTGTACGGCCACACTCCACTGAGATACGGC
CAGACTCTACGGAGGCAGCAGTAAGGAATTTGCG
```

Results of homology search for H2

CHLCPCE16S
62.7% identity in 357 nt overlap

```
               10        20        30        40        50        60
h2.    GGGGAGCTCAGGATGAACGCTCGCTGTGTGCCTAATACATGCATGTTGAACGGGATGTAG
           X:::::::::::::::: :: :  ::: ::: ::::::: :: :::: ^v: ::
CHLCPC TCCTGGCTCAGGATGAACGCTGGCGGCATGCTTAACACATGCAAGTCGAAC-GGAGGTTT
       20        30        40        50        60          70

               70          80        90       100       110
h2.    CAATACATTCAG--TAGCGAATGGGTGAGTAACACGTACCTAACCTACCTTTAAGACTGG
         :   :: ::  :::   :      :::::    :   :  :  ::   : ::    :
CHLCPC GCTTTTTTTAAGCTTAGGTTAAAAAAAAGTAAATTCTTAGTGGCGGACGGGTGAGTAACG
        80        90       100       110       120       130

      120            130       140       150       160
h2.    GATAA----CTA-TTGGAAACAATAGCTAA-TACCGGATATAGTTATT--TATCGCAT--
         :::    :::  :  :  :    :  ::  :: ::: :  :::  :  :: : :::
CHLCPC CGTAAGAACCTACCTTTAGGCGGGGGACAACAACTGGAAACGGTTGCTAATACCCCATAA
       140       150       160       170       180       190

      170       180       190       200       210       220
h2.    GATGAGTAATAGAAAGGAGCTTCACAGCTTCACT-TAAAAATGGG-GTGCGGAACATTAG
       : :::: : :: ::::::     :    :::  : :::: :::::  ::::    :::::
CHLCPC GCTGAGGAGTA-AAAGGAATACTA----TTCTGTCTAAAGATGGGCTTGCGTCTGATTAG
       200        210       220           230       240       250

        230       240       250       260       270       280
h2.    TTAGTTGGTAGGGTAATGGCCTACCAAGACGATGATGTTTAGCCGGGCCGAGACG-TGTA
        : ::::::  :::::: :: :::::::: : : :::   :::: :: : :::: : ::
CHLCPC CTTGTTGGTGAGGTAATTGCTTACCAAGGCAACGATCAGTAGCTGGTCTGAGAGGATGAT
            260       270       280       290       300       310

         290       300       310       320       330       340
h2.    CGGCCACACT-CCACTGAGATACGG-CCAGAC-TCTACGG-AGGCAGCAGTAAGGAATTT
       : ::::::::   ::::::: :::: :::::: ::::::: :::::::::: ::::::::
CHLCPC CAGCCACACTGGGACTGAGACACGGCCCAGACTTCTACGGAAGGCAGCAGTGAGGAATTT
            320       330       340       350       360       370


h2.    GCG
        :
CHLCPC TCC
```

Figure 21:

Homology alignment of E12 fragment and HUMHM02E05.

>HUMHM02E05 Human HepG2 3'-directed MboI cDNA clone
92.8% identity in 223 nt overlap

```
          1030       1040       1050       1060       1070       1080
e12w1  GATCTGATTG-TAACCCACTGAGAGGACAGAGAGAAATAAGTGCCCTCTCCCACTCTCTT
       :::: :.:.: .::::: ::.X:::::::::::::::::::::::::::::::: :::::
HUMHM0 GATCCGNTNGANAACCCCCTNAGAGGACAGAGAGAAATAAGTGCCCTCTCCCACCCTCTT
                10         20         30         40         50         60

           1090       1100       1110       1120       1130       1140
e12w1  CCTACTGGTCTCTCTATGCCTCTCTAGAGTCTCGTCTCTTTTACCCTGGCCCCTCTCCCT
       :::::::::::::::::::::::::: :::::::::::.::::::::::::::::::.:
HUMHM0 CCTACTGGTCTCTCTATGCCTCTCTACAGTCTCGTCTCTNTTACCCTGGCCCCTCTCCNT
                70         80         90        100        110        120

           1150       1160       1170       1180       1190       1200
e12w1  TGGGCTCTGATGAAAAATTGCTGACTGTAGCTTTGGAAGTTTAGCTCTGAGAACCGTAGA
       :::::::::::::::::..:::::::::::::::::::::::::::::::::::::::::
HUMHM0 TGGGCTCTGATGAAAAANNGCTGACTGTAGCTTTGGAAGTTTAGCTCTGAGAACCGTAGA
                130        140        150        160        170        180

           1210       1220       1230       1240
e12w1  TGATTTCAGTTCTAGGAAAATAAAACCCGTTGATTACTAAAAA
       ::::::::::.::::::::::::::::::.:::::::::.::X
HUMHM0 TGATTTCAGTNCTAGGAAAATAAAACCCGTNGATTACTANAAA

                190        200        210        220
```

3. Southern blot analyses

Southern blot analyses of E12 and E13 were done to address the possibility that the two genes may be amplified by rearrangement or translocation.In order to compare the restriction patternsof genomic DNAs from normal and transformed keratinocytes, genomic DNAs from normal epithelial cells, transformed cell lines, and SCC cell lines were isolated, digested with restriction enzymes and transferred to nitrocellulose filters. The filters were hybridized to ^{32}P labeled E12 and E13 probes. Figure 22 shows the results of the Southern blot analysis using E12. In EcoR 1 digest, the E12 probe hybridized to a major band of about 9.2Kb and a minor band of about 8.5 kb. In BamH 1 digested pattern for E12, two bands were shown on the Southern blot: a major band was about 25 kb and one minor band of about 7.5 kb. Southern blot analysis (Figure 23) showed that the E13 probe hybridized to a band of about 9.5kb in BamH I digests and a band of about 6.5kb in EcoR I digests. There were also no detectable changes in restriction digestion pattern found for E12 and E13 in SV40 transformed keratinocytes. these results indicate that E12 and E13 are probably not altered by amplification and rearrangement in either viral transformants and SCC cells.

Based on northern blots E12 and E13 are believed to represent two genes whose expression are specifically and dramatically induced after SV40 infection. We further suggest that they may play a role in the transformation process.

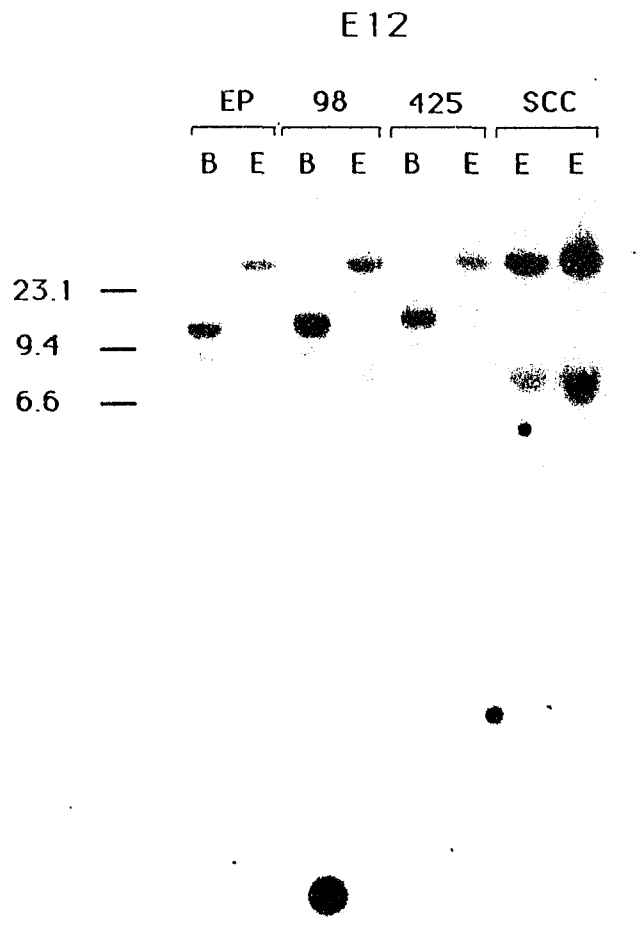


Figure 22:

Southern blot hybridization analysis of E12 sequences in genomic DNAs normal epithelial cells (EP), transformed epithelial (98, 425) and SCC (SCC324) cell lines after BamH 1 and EcoR 1 digestion. The migration of Hind 111 size markers is indicated at the left.

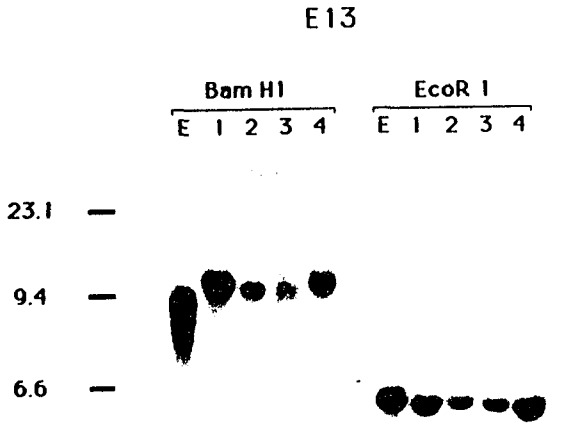


Figure 23

E13 sequences in DNAs from normal epithelial cells (E and 3) and transformed epithelial (1: 425; 2 and 4: 98) cell lines after BamH 1 and EcoR 1 digestion. The Hind III-digested markers are indicated in kilobase.

4. Elongation and cloning of the E12 gene

Due to the degradation of mRNA during construction of the cDNA library, both E12 and E13 are only about 400 bp long and not full length cDNA sequences.

From sequence analysis of the fragment of E12, poly A tail and poly A addition site were identified in the 3' end of E12. In order to obtain the missing 5' end of the cDNA, the RACE (rapid application of cDNA ends) method was used to elongate the E12 gene from the 3' end of the E12 fragment. The primers used for the RACE reaction for the E12 gene were:

5'-ACGAGACTCTCGAGAGGCAT-3' (primer 53)

5'-GACTCGAGTCGACATCGAT17-5' (primer 54)

5'-GACTCGAGTCGACATCGAT-3' (primer 55)

5'-TCAGATCTGGTGACAACACT-3' (primer 56)

The primers were synthesized using a Cyclone Plus DNA Synthesizer(Millipore) and the lenghths checked on 6% acrylamide gel (Figure 24).

The mRNA from transformed cell line 425 was isolated and reverse transcribed using primer 53. The 3' end of the first strand of the cDNA was tailed with oligo (A). Amplification of the 5' end of E12 was accomplished by PCR reaction with primers 54, 55 and 56. Figure 25 shows the first round of the products of RACE for E12. Both total RNA and mRNA were used as templates; in lanes 2 and 3, cytoplasmic RNA was used as a template. mRNA was used in lane 4 and 5. A clear band of about 450 bp is seen in lane 5.The

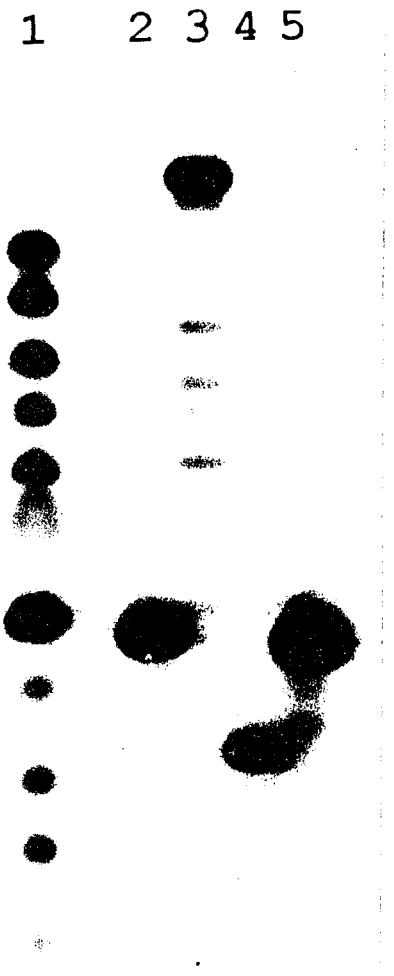


Figure 24:

Synthesized RACE primers. 4 primers were synthesized by using Cyclone Plus DNA Synthesizer (Millipore).

The primers and the size markers were end labeled and run on a 6% acrylamide gel. Lane 1 contains the size markers from 32 mer (top one) to 8 mer (bottom one). Lane 2 is primer 53 (20 mer); lane 3 is primer 54 (35 mer); lane 4 is primer 55 (17 mer); lane 5 is primer 56 (20 mer).

elongated 5' end and E12 should share about 200 bp overlapping the original cDNA. The E12 probe was used to hybridize to the products of RACE. Hybridization of E12 to the RACE products is shown in Figure 26. It is clear that mRNA templates (lane 4 and 5) give longer and more specific products than total RNA templates do. The products of RACE then were blunt ended and cut with a restriction enzyme at one end of the blunt ended products. Those products were cloned into mp18 and mp19 for sequencing. The E12 fragment was extended about 400 bp at 5' end (including 200 bp overlap region), but it was not enough to contain the whole E12 mRNA based on the length of E12 seen in the northern blot(Figure 12). To get the full length gene, a second round of RACE amplification of the 5' end of E12 was performed. Three specific primers (primer 57: 5'-TATGACTATAAGCTGGC-3', primer 58: 5'-CTTTACGTCTGGGTGATATA-3', and primer 131: 3'-TGGAGAGCACACAAAGGG-5') derived from successive sequencing runs of E12 were synthesized. The second round amplification produced an extended fragment of about 700 bp (Figure 27: lane 3 and lane 5). of RACE was performed by using rePCR method. The fragment ofabout 700 bp was excised from a low melting agarose gel and theagarose slice was melted at 65^{0}C. RePCR was performed by using DNA samples in melted agarose as templates. The band in lane 4 of Figure 28 is the reamplified band shown in lane 5 of Figure 27. The reamplified PCR fragment called B1012 as cloned into mp18 and mp19 and sequenced.

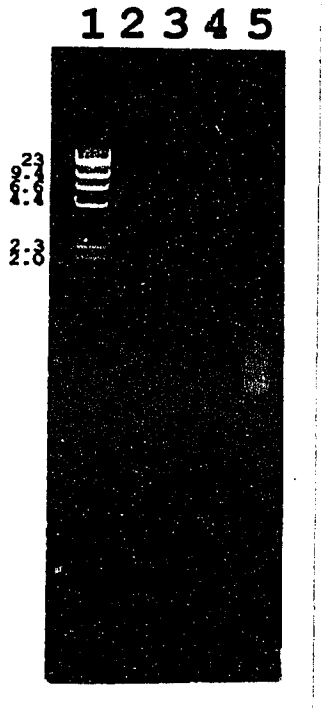


Figure 25:

The first round amplification of the 5' end of the E12 cDNA. Total RNAs and mRNAs of SV40 transformed cell line 425 were reverse transcribed with primer 53 into cDNAs. The cDNAs were used as templates for the PCR reaction to amplify the 5' end of the E12 cDNA. Lane 1 contains molecular weight marker and sizes are indicated in kb. Lane 2 is the PCR product of RACE, using a single stranded cDNA template reverse transcribed from total RNA and primers: 54, 55 and 53. Lane 3 is the same as lane 2, except for the addition of primer 56 instead of primer 53. Lane 4 is the PCR product of RACE, using a single stranded cDNA template reverse transcribed from mRNA and primers 53, 54, 55. Lane 5 is the same as lane 4, except for the primer 56 instead of primer 53.

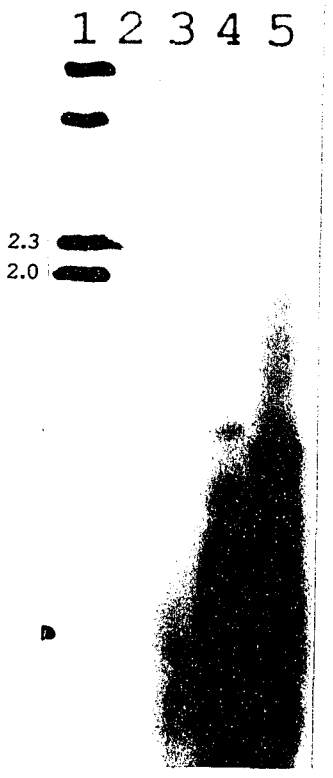


Figure 26:

Southern blot hybridization of the first round PCR products to the E12 probe. The cDNAs from the PCR reactions in Fig. 25 were transferred onto nitrocellulose filter and hybridized to E12 probe. Lane 1 contains molecular weight marker and sizes are indicated in kb. Lane 2 is the PCR product of RACE, using a single stranded cDNA template reverse transcribed from total RNA and primers: 54, 55 and 53. Lane 3 is the same as lane 2, except for the addition of primer 56 instead of primer 53. Lane 4 is the PCR product of RACE, using a single stranded cDNA template reverse transcribed from mRNA and primers 53, 54, 55. Lane 5 is the same as lane 4, except for the primer 56 instead of primer 53.

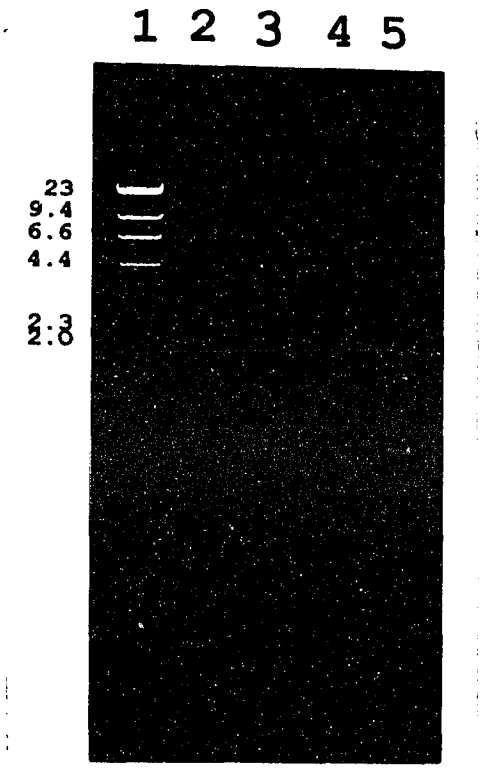


Figure 27:

Second round amplification of the 5' end of E12 cDNA. mRNAs from SV40 transformed keratinocyte cell line 425 were reverse transcribed into cDNAs with E12 specific primer 53, 56, and 131, respectively. The cDNAs were used as templates for PCR reaction to amplify the 5' end of E12 DNA. Lane 1 contains molecular size markers whose migration is indicated in kb on the left. Lanes 2, 3, 4 and 5 are PCR products of RACE for 5' end of E12 cDNA. In lane 2 primer 53 was used for reverse transcription and primers 54, 55 and 56 were used for RACE. Lane 3 is the same as lane 2 except that primer 57 was used instead of primer 56. In lane 4 primer 56 was used for reverse transcription and primers 54, 55 and 57 were used for RACE. In lane 5 primer 131 was used for reverse transcription and primers 54, 55 and 57 were used for RACE.

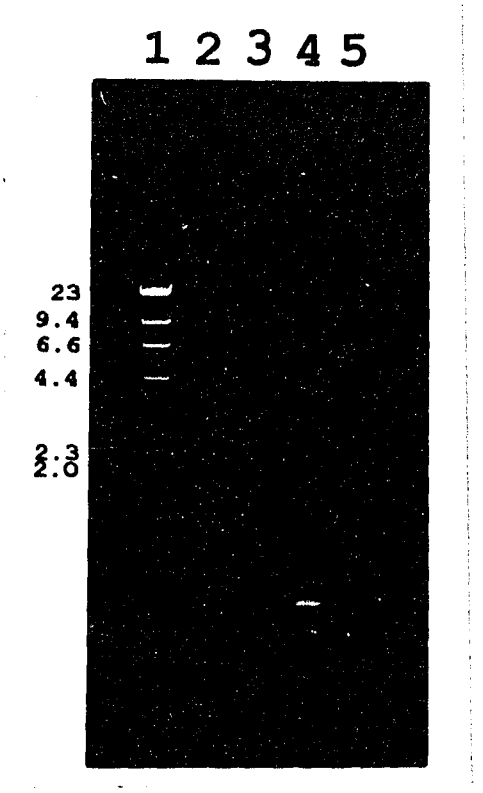


Figure 28:

A band of about 700 bp (Fig. 27 lane 5) was excised and melted at 65°C. The amplification as performed by using the cDNA in the melted agarose as a template. The primers used in PCR reactions were: primers 55 and 56 (lane 2); primers 55 and 57 (lane 3); primers 55 and 57 (lane 4); primers 55 and 131 (lane 5).

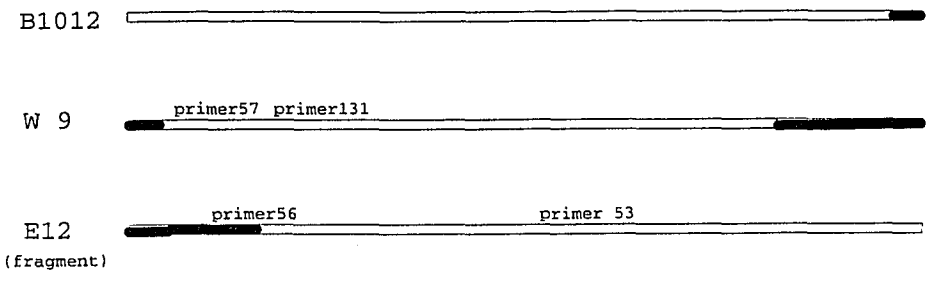


Figure 29:
Alignment of E12 gene from three fragments.
The E12 fragment is a cDNA clone. The W9 is a clone from first round PCR. The B1012 is a clone from second round PCR. Positions of the various primers are indicated. The solid lines are overlapping areas.

5. Sequencing analysis of E12 gene

Sequences of the three RACE products were aligned according to overlapping areas (Figure 29). The nucleotide sequence of E12 and the deduced amino acid sequence from the longest open reading frame are presented in Figure 30. The complete E12 cDNA is 1251 bp in length. A putative translation initiation site ATG is at position 401 and an in-frame stop codon is at position 1107. The open reading frame is 715 bp encoding a 235 amino acid protein. The polyadenylation site (AATAAA) is at position 1220. The PCGENE program (intelligenetics) was used to analyze the putative protein composition, conformation and hydrophobicity of the E12 gene. The LFASTA program and the GENBANK database were emploied to search the sequence homologies between the E12 and genes catalogued in the database. The putative protein composition and molecular weight for E12 is shown in Figure 31. The prediction of putative protein conformation for E12 is shown in Figure 32. The protein hydrophobicity dada is shown in Figure 33. Computer analyses and sequence searches indicated that the E12 transcript contains a 5' upstream Alu sequence. The Alu sequence is about 300 bp in length at the 5' end of E12 and is apparently expressed in the open reading frame; the Alu sequences were found in many mRNA sequences catalogued in Genbank (Figure 34). The functions of the Alu sequences in mRNAs are not known. The remainder of the E12 gene showed no significant homology to any sequence currently cataloged in the Genbank, except for the cDNA fragment HM02E05 (128), a randomly amplified and unannotated sequence (Fig. 21).

```
 GAGACAGACTTTCGCTCTTGTCACCCAGGCTGGAGTGCAATGGAATGATCTTGGCTCACT     60
GCAAACTTCGCCTTCCAGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCCAGTAGCTGGG    120
ATTACAGGCATGTGCCACCACGCCAGGTTCATTTTGTATTTTCAGTAGAGACTGGGTTTC    180
ATCGTGCTGGCCAGGCTGGTCTCGAACTCCTGACCTCAAGTGTTCCACCCACCTCAGCCT    240
TCCCAAAGTGCTGGGATTACAGGCGTGACGACCAAGCCTGGCCAACCCTCTCGTCTTTCA    300
GGGGCTTAAATAAATTTGCTTTTCTCCCCAAATTCTTTGCCCACCCCATCAGCCAGTTTC    360
AAAGGGTGGAGTGTAACGTGGGTTGCCCAATCCTGCTTGCCATGAAGTACTTAGCTTACA    420
                                      METLysTyrLeuAlaTyrS

GCAGTTTGCCAGGCGCAGACACCATGCTTTATTTCTATTTCTACGAACAAGAAGCGTCTC    480
erSerLeuProGlyAlaAspThrMETLeuTyrPheTyrPheTyrGluGlnGluAlaSerL

TTGCAGTCTGCAATATTTGTCGTCAAAAATTTCACTGGGTGTTGTATCAAATCAGCCACC    540
euAlaValCysAsnIleCysArgGlnLysPheHisTrpValLeuTyrGlnIleSerHisL

TGTATAGAGGTGTGATCGTTGACAATTTTTTGTTACATCCCGATGGCCGTTTTACCTGGA    600
euTyrArgGlyValIleValAspAsnPheLeuLeuHisProAspGlyArgPheThrTrpT

CCATTTTCTTCCTTTCATGGGTGAAGCAAAACTCCTTGGTAGACTTTTTTTTTGGGACAG    660
hrIlePhePheLeuSerTrpValLysGlnAsnSerLeuValAspPhePhePheGlyThrG

AGTCTCGCTCTGTCGCTCTGTTGCCCAGGCTGGAGTGCAGTGGTGCGATGTCGACTCTCC    720
luSerArgSerValAlaLeuLeuProArgLeuGluCysSerGlyAlaMETSerThrLeuH

ACACAGTGCTCAGGCCAGCTTATAGTCATATATATCACCCAGACGTAAAGGAAAAGACAC
isThrValLeuArgProAlaTyrSerHisIleTyrHisProAspValLysGluLysThrH    780

ATTTTTTAGGAAATGTTTTTAATAAAAGAAAATTACAAAAAAAAAATTTTAAAGACCCCTA    840
isPheLeuGlyAsnValPheAsnLysArgLysLeuGlnLysLysIleLeuLysThrProA

ACCCTTTGTGTGCTCTCCATTCTGCTCCTTCCCCATCGTTGCCCCCATTTTTGAGGTGCA    900
snProLeuCysAlaLeuHisSerAlaProSerProSerLeuProProPheLeuArgCysT

CTGGGAGGCTCCCCTTCTATTTGGGGCTTGATGACTTTCTTTTTGTAGCTGGGGCTTTGA    960
hrGlyArgLeuProPheTyrLeuGlyLeuAspAspPheLeuPheValAlaGlyAlaLeuM

TGTTCCTTCCAGTGTCATTTCTCAATCCACATACCCTAACCTGGCCCCCTCAGTGTTGCA   1020
ETPheLeuProValSerPheLeuAsnProHisThrLeuThrTrpProProGlnCysCysT

CCAGATCTGATTGTAACCCACTGAGAGGACAGAGAGAAATAAGTGCCCTCTCCCACTCTC   1080
hrArgSerAspCysAsnProLeuArgGlyGlnArgGluIleSerAlaLeuSerHisSerL

TTCCTACTGGTCTCTCTATGCCTCTCTAGAGTCTCGTCTCTTTTACCCTGGCCCCTCTCC   1140
euProThrGlyLeuSerMETProLeu

CTTGGGCTCTGATGAAAAATTGCTGACTGTAGCTTTGGAAGTTTAGCTCTGAGAACCGTA   1200
GATGATTTCAGTTCTAGGAAAATAAAACCCGTTGATTACTAAAAAAAAAAA            1260
```

Figure 30.
The nucleotide sequence and deduced amino acid sequece of E12. The putative translation initiation site ATG is at 401. The terminate site TAG is at 1170. The polyadenylation signal AATAAA is at 1220.

SEQAID II PROTEIN COMPOSITION ANALYSIS OF E12 (from 1 to 235)
Putative Amino Acid Composition: total residues=235

Amino acid	Count	Mole%
Ala	12	5.11
Cys	8	3.40
Asp	8	3.40
Glu	6	2.55
Phe	18	7.66
Gly	11	4.68
His	10	4.26.09
Gln	7	2.98
Arg	12	5.11
Ser	20	8.51
Ile	7	2.98
Lys	10	4.26
Leu	35	14.89
Met	5	2.13
Asn	8	3.40
Pro	19	8
Thr	13	5.53
Val	12	5.11
Trp	4	1.70
Tyr	10	4.26
Ter	0	0.00

Protein Molecular Weight = 26885.40

Figure 31
The table above is the putative protein composition of E12 as determined using PCGENE program.

```
MKYLAYSSLPGADTMLYFYFYEQEASLAVCNICRQKFHWVLYQISHLYRGVIVDNFLLHP 60

HHEEEETTCTTCCHEEEEHHHHHHHHHHHHHHHHHHHHHHHEEEHHHEEEEEEEEEEEECC


DGRFTWTIFFLSWVKQNSLVDFFFGTESRSVALLPRLECSGAMSTLHTVLRPAYSHIYHP 120

TTTEEEEEEEHHHHHTTTHEEEEEHHHHHHHHEECHHHHHHTHHEEEEEEEETEEEEECH


DVKEKTHFLGNVFNKRKLQKKILKTPNPLCALHSAPSPSLPPFLRCTGRLPFYLGLDDFL 180

HHHHHHHEHHHHHHHHHHHHHEECCCCTEEEEECCCCCCCCCEEEETTTCCEEETCHHHH


FVAGALMFLPVSFLNPHTLTWPPQCCTRSDCNPLRGQREISALSHSLPTGLSMPL      235

HHHHHHHEEHEEEECTEEEECCTTTEEETTTCCTTTHHHEEEEEETCCETEEEEE
```

Figure 32
Shown above is the prediction of the putative protein conformation of E12 as determined using the PCGENE program (SEQUID II-G.O.R.). H=alpha Helix, E=beta sheet (Extended), T=Turn, C=random
Coil Decision constants: alpha:-100; beta:-87; turn:0; coil:0.

Figure 33

The protein hydrophobicity for E12. The data were analyzed with PCGENE.

PROTEIN HYDROPHOBICITY DATA FOR E12
Averages and values plotted are for that base plus the next 6

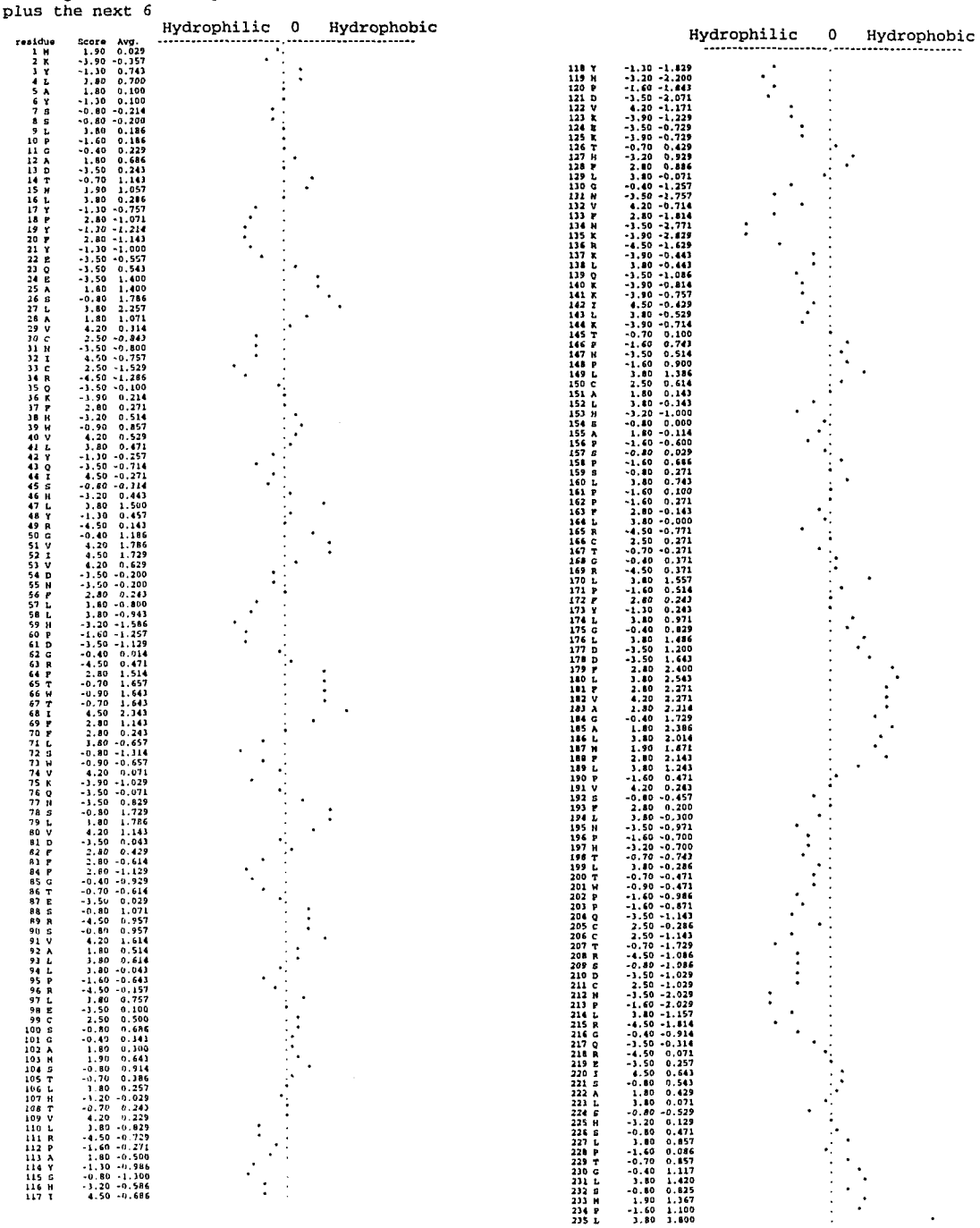


Figure 33

The protein hydrophobicity for E12. The data were analyzed with PCGENE.

Figure 34

Figure 34

The Alu sequence at 5'end of E12 shows high homolgies with other cDNAs and exons previously catologued in Genbank. A are best search scores. B are homology alignments of E12 with two other mRNAs catologued in Genbank.

A
The best scores are:

HUMALBGC	Human serum albumin gene, complete cds.
HUMALBGC	Human serum albumin gene, complete cds.
S60052	oestrogen receptor (3' untranslated region).
HUMHPARS2	Human haptoglobin-related gene, exons 2, 3.
HUMTTK	Human kinase (TTK) mRNA, complete cds.
HUMPRCA	Human protein C gene, complete cds.
HUMKER18	Human keratin 18 (K18) gene, complete cds.
HSCERBA8	Human c-erbA-1 gene, exon 8 for thyroid hormon.
HUMLYTOXBB	Homo sapiens lymphotoxin-beta gene, compl
HSCGMP4	H.sapiens DNA for cGMP phosphodiesterase (exon).
HUMPP14B	Human placental protein 14 (PP14) gene, comlete cds.
HUMORF08	Human mRNA for ORF, complete cds.
HUMDES	Human desmin gene, complete cds.
HUMWT1	Human Wilms tumor gene encoding WT1 protein.
HUMLIMGP	Human lens membrane protein (mp19) gene, exon.

B
HUMTTK Human kinase (TTK) mRNA, complete cds.
77.5% identity in 285 nt overlap

```
              10        20        30        40        50        60
e12w1  GAGACAGACTTTCGCTCTTGTCACCCAGGCTGGAGTGCAATGGAATGATCTTGGCTCACT
       :: ::::  : :: :: : ::: X:::::::::::::: :::   ::::::::::::::
HUMTTK GAAACAGGGTCTCCCT-TGGTCGCCCAGGCTGGAGTGCAGTGGTGCGATCTTGGCTCACT
         340       350       360       370       380       390

              70        80        90       100       110       120
e12w1  GCAAACTTCGCCTTCCAGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCCAGTAGCTGGG
       :: :: : :: ::: :: :::::: ::: ::::  : : :::::::  :::::::: :
HUMTTK ATAACCTCCACC-TCCTGGTTTCAAGTGATCCTCCCACTTTAGCCTCCTGAGTAGCTGTG
         400       410       420       430       440       450

             130       140       150       160       170
e12w1  ATTACAGGCATGTGCCACCACGCCAGGTTCA-TTTTGTATTTTCAGTAGAGACTGGGTTT
       ::::::::: ::  ::::::: :: :: : : :::::::::::: : ::::::: ::::::
HUMTTK ATTACAGGCGTGCACCACCACACCCGGCTAATTTTTGTATTTTTATTAGAGACAGGGTTT
           460       470       480       490       500       510

     180       190       200       210       220       230
e12w1  CATCGTGCTGGCCAGGCTGGTCTCGAACTCCTGACCTCAAGTGTTCCACCCACCTCAGCC
       :: : :: :::::::::::: :::: :::::::::  :::::: : ::: ::  ::::   :
HUMTTK CACCATGTTGGCCAGGCTGTTCTCAAACTCCTGGACTCAAGGGATCCGCCTGCCTC-CAC
           520       530       540       550       560       570

     240       250       260       270       280       290
e12w1  TTCCCAAAGTGCTGGGATTACAGGCGTGA-CGACCAAGCCTGGCCAACCCTCTCGTCTTT
       ::::::^v:: : : ::::::::: ::::   :::: ::::: ::
HUMTTK TTCCCAAAGTCCCGAGATTACAGGTGTGAGTCACCATGCCTGACCTTATAATTCTTAAGT
           580       590       600       610       620       630
```

HUMALBGC Human serum albumin gene, complete cds.
84.2% identity in 285 nt overlap

```
              10        20        30        40        50        60
e12w1  GAGACAGACTTTCGCTCTTGTCACCCAGGCTGGAGTGCAATGGAATGATCTTGGCTCACT
       ::::::: :::::::::::  ::::::::: ::::::::  ::::::::::::
HUMALB AAGACAGAGTTTCGCTCTTGTTTCCCAGGCTGGGGTGCAATGGTGTGATCTTGGCTCAGC
    9910      9920      9930      9940      9950      9960

              70        80        90       100       110       120
e12w1  GCAAACTTCGCCTTCCAGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCCAGTAGCTGGG
       :::: ::  ::: ::: X:::::::: :::: :: :::::::::::::: :::::::::
HUMALB GCAACCTCTGCC-TCCTGGGTTCAAGTGATTTTCATGCCTCAGCCTCCCAAGTAGCTGGG
    9970      9980       9990     10000     10010     10020

             130       140       150       160       170       180
e12w1  ATTACAGGCATGTGCCACCACGCCAGGTTCATTTTGTATTTTCAGTAGAGACTGGGTTTC
       :::::::::::: :::::::: :: :: : :::::::::::: ::::::: : :::::::
HUMALB ATTACAGGCATGCGCCACCACACCTGGCTAATTTTGTATTTTTAGTAGAGGCGGGGTTTC
    10030     10040     10050     10060     10070     10080

             190       200       210       220       230       240
e12w1  ATCGTGCTGGCCAGGCTGGTCTCGAACTCCTGACCTCAAGTGTTCCACCCACCTCAGCCT
       : : :  :: :::: ::::::::::::::::::::::: ::: ::::::X :::  :::
HUMALB ACCATATTGTCCAGACTGGTCTCGAACTCCTGACCTCAGGTGATCCACCCGCCTTGGCC-
    10090     10100     10110     10120     10130     10140

             250       260       270       280       290
e12w1  TCCCAAAGTGCTGGGATTACAGGCGTGA-CGACCAAGCCTGGCCAACCCTCTCGTCTTTC
       ::::::::::::::::::::::::: ::: : :::  :::  ::: :
HUMALB TCCCAAAGTGCTGGGATTACAGGCATGAGCCACCTTGCCCAGCCTAAGAAGATTTTTTGA
```

Figure 36 shows the nucleotide sequence of E13 fragment. No significant homology to any sequence currently catalogued in Genbank was found and both E12 and E13 appear to represent novel genes.

6. Construction and screening of genomic libraries

In order to examine the regulatory sequences of E12 and E13, two genomic libraries were created in the lambda phage vector,EMBLE, and screened with E12 and E13 probes, respectively. Genomic libraries were created from DNA which was isolated from epidermal keratinocytes and cut with restriction enzymes, BamH 1 for E12 cloning and EcoR 1 for E13 cloning. Digested genomic DNA fragments were cloned into Lambda phage EMBL 4 DNA digested with the appropriate restriction enzyme. $1x10^6$ clones were plated and screened by using the original E12 cDNA fragment as a probe. One positive clones containing E12 was selected and designated GE12. The 9.2 kb insert released by digestion of GE12 with BamH I was shown in Figure 36. The identity of the insert was confirmed by Southern blot of the same gel (Figure 37). To determine if the GE12 clone contained the whole E12 coding sequence, the B1012 probe which represents the 5' end of E12 was used to hybridize to GE12. GE12 DNA was digested with BamH I in lane 5 (Fig.38 A) and hybridized to B1012 probe in lane 5(Fig. 38: B). These results suggest that GE12 probably contains the all the E12 exons. The insert of GE12 was then subcloned into a

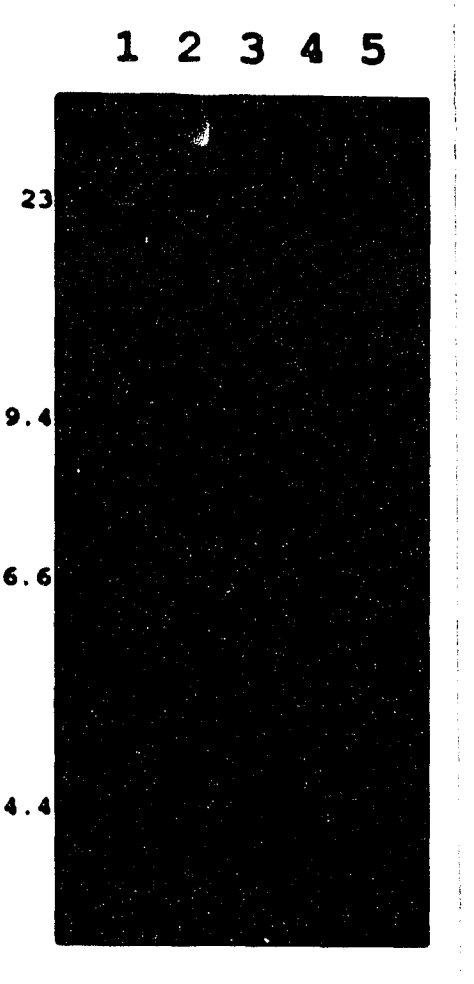


Figure 36:

Digestion of genomic clone GE12. Lane 1 shows molecular weight markers with sizes indicated in kb at the left. Lane 2 contains GE12 cut with Sal 1. Lane 3 is GE12 cut with EcoR 1. Lane 4 is GE12 cut with BamH 1. Lane 5 is EMBL 4 cut with BamH 1.

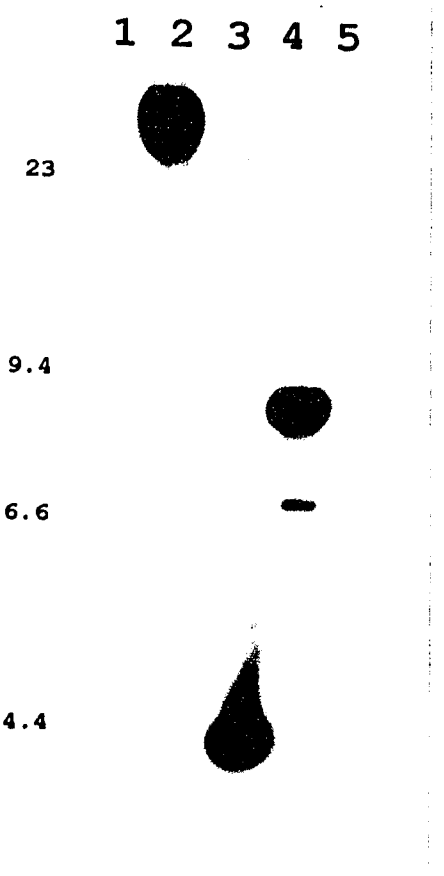


Figure 37:

Southern blot hybridization of clones from genomic library. The DNAs in Fig. 36 were transferred onto a nitrocellulose filter and hybridized to a ^{32}P labeled E12 cDNA probe. Lane 1 contains molecular weight markers and sizes are indicated in kb on the left. Lane 2

contains GE12 cut with Sal 1. Lane 3 is GE12 cut with EcoR 1. Lane 4 is GE12 cut with BamH 1. Lane 5 is EMBL 4 cut with BamH 1.

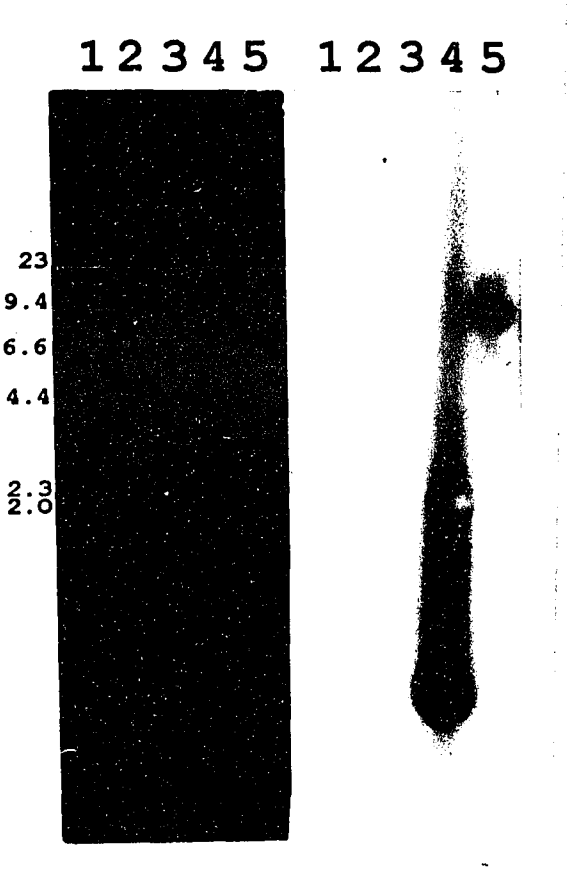


Figure 38:

Southern blot hybridization of PGE12 DNA, using the 5' end cDNA clone, B1012, as a probe. A: DNAs were run on an agarose gel and B: the DNAs were transferred onto a nitrocellulose filter and a ^{32}P-labeled B1012 probe was hybridized to the DNAs. lanes 1 are molecular weight markers and sizes are indicated in kilobases at the left; lanes 2 is uncut plasmid PGEM3Z; lanes 3 is plasmid PGEM3Z cut with BamH I; lanes 4 is cDNA B1012; lanes 5 is GE12 cut with BamH I.

multifunctional plasmid PGEM 3Z and named PGE12 for further study. Restriction enzyme digestion were performed (Figure 39) and the restriction map of PGE12 was drawn(Figure 40) based on the results from Figure 39. The partial sequence of PGE12 is in Figure 41.

The genomic library for E13 was constructed essentially in the same way as the library for E12 except that genomic DNA was digested with the EcoR I and the fragments were cloned into the lambda phage vector EMBL 3.The positive clones were isolated and their inserts DNAs were digested by EcoR 1 into fragments of 9.4 kb, 3.7 kb, 2,6 kb, 2.5 kb and 1.0 kb(Fig. 42: A). When the blot was hybridized with the E13 probe, the 9.4 kb fragment gave a positive signal. This suggested that the 9.4 kb fragment, which contains E13 sequence, may overlap with the short arm of the phage(Fig. 42: B). In order to confirm this possibility, the EcoR

I digested product as electrophoresed for two days to separate the presumed overlapping fragments. A band of about 9.5 kb was observed just above the short arm of the vector (Fig. 43); the total length of the insert was 19.2 kb. The 9.5 kb insert fragment was subcloned into plasmid PGEM3Z and partially sequenced (Fig. 44).

The plasmid PGE13 was used for restriction enzyme mapping. The PGE13 was digested with single enzyme(Fig. 45) and with double enzymes(Fig. 46). A restriction map was constructed as in Figure 47.

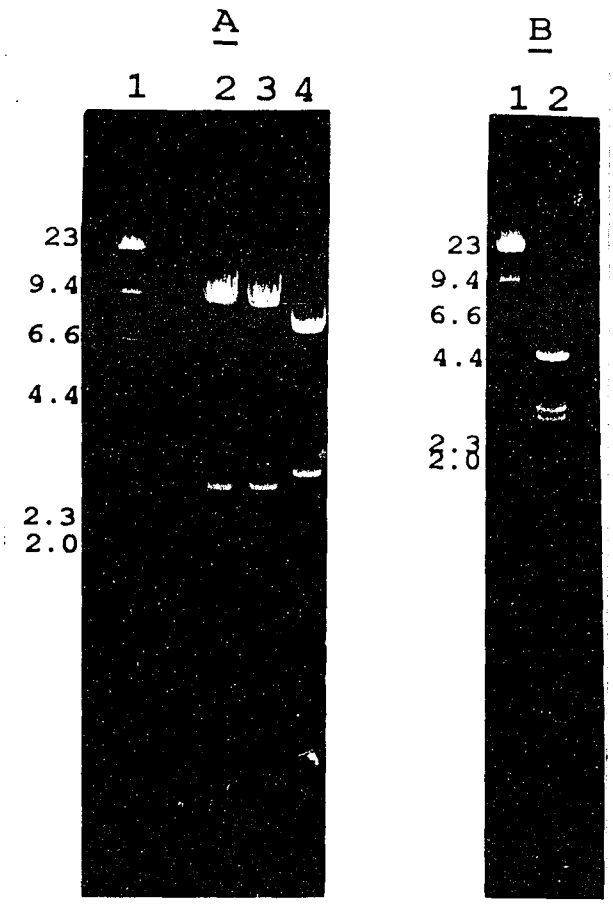
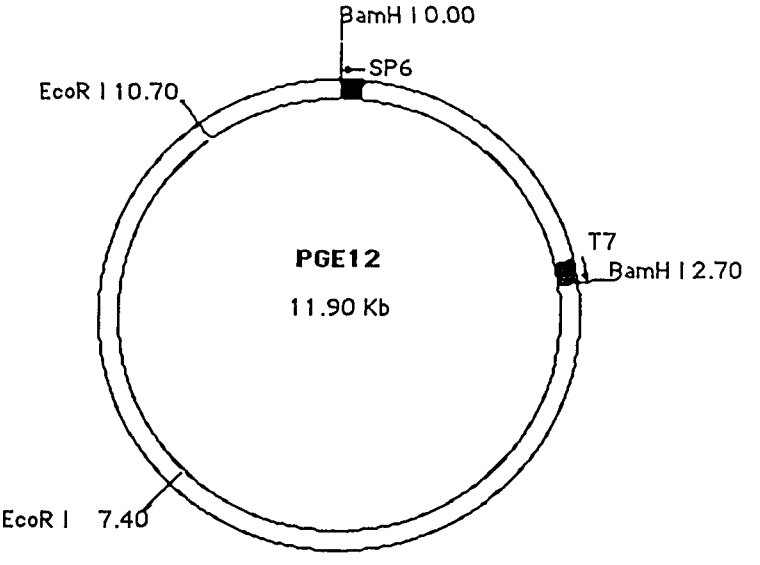


Figure 39:

Single and double restriction enzyme digestions of PGE12. PGE12 is the insert of GE12 subcloned into a plasmid, PGEM3Z. A: Single digestion of DNAs with BamH I or EcoR I. Lane 1 are molecular weight markers and sizes are indicated by kilobase. Lane 2 and 3 are PGE12 cut with BamH I. Lane 4 shows PGE12 cut with EcoR I. B: Double digestion of PGE12 with BamH I and EcoR I. Lane 1 contains molecular weight markers. Lane 2 shows PGE12 digested with BamH I and EcoR I.

A



B

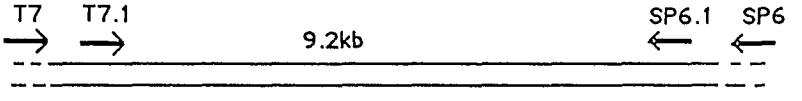


Figure 40
A. Restriction map of PGE12.
B. Positions of sequece primer in PGE12.

Figure 41

The partial sequences of the PGE12 were obtained using primers SP6 and SP6.1 in A; primers T7 and T7.1 in B.

A. The nucleotide sequence of the PGE12 starts from the end of the

Sp6 promoter.

CCTGCAGGTCGACTCTAGACCAATCCAATGAGAGATTTTACTCAGGGCAGCCTGAAGGAGAAAAGTGAAG

AAGAGTTTAAACTATGTGTCAATGAGAGAAGCTGGAAATCAGAAAAGCCAGAAAAAACCAGCACCCATAA

TTGAACTGAGGAGTAAGGCAAGGTCACAAACATGAAAACAAAGGAAAAGTTTCTTGCAGTAACACCAGTT

GGAACTTCTCATCATTAACTGTATTTTTCCTAGAAATTTCTACAGAAAATAAAATAAGAATGCTGTGGTC

AAAAGGGCAGAGTATTCAAACAGAAAGCAGAAGAAAACTGCTTTCTCTGTCCAGAGAAACTAAGCTCCGT

CACTTACTAATTATATCCATGGCA

B. The nucleotide sequence of the PGE12 starts from the end of the T7 promoter area.

CTGGCTAATAAAATAACACACAAAGGCAAAGCGTAGTACTCTTTACCAGTGTATGTGCCATGAAAAAGTG

GAACAATGCCCATGACTGACTGAATATAAATAGAGAGTCAGTGAAACGATAATGCAGGAACCTTGGGATC

ATAATGAGAGATCTAATGGTCTTTGGTTCTCAATAGACTTTCTTTACCCTCCTGATGGATAATAAATTTC

TACATACTTGACAATTTGCACATGGCGACCAGAATTTCTGTGCATAAAAGTCATATACCACAGAGTCACG

GTTAGTAACACATGGAATGGGGAAAAACTGAAAGCCTTTTCTGTAAGATCTGGAACATGACAAGGATGCC

CACTTTCACCACTGTTATTCAACATAGTACTGGAATTCCTAGCTAGAGCAACCAGACAAGAGAAAGAAAT

AAACGGATCAAATTGGAATGAAAGAAGTTAAATTATCTGTTTGCAGATGATATAGCTATATTGGAAACTG

AGACTCACA

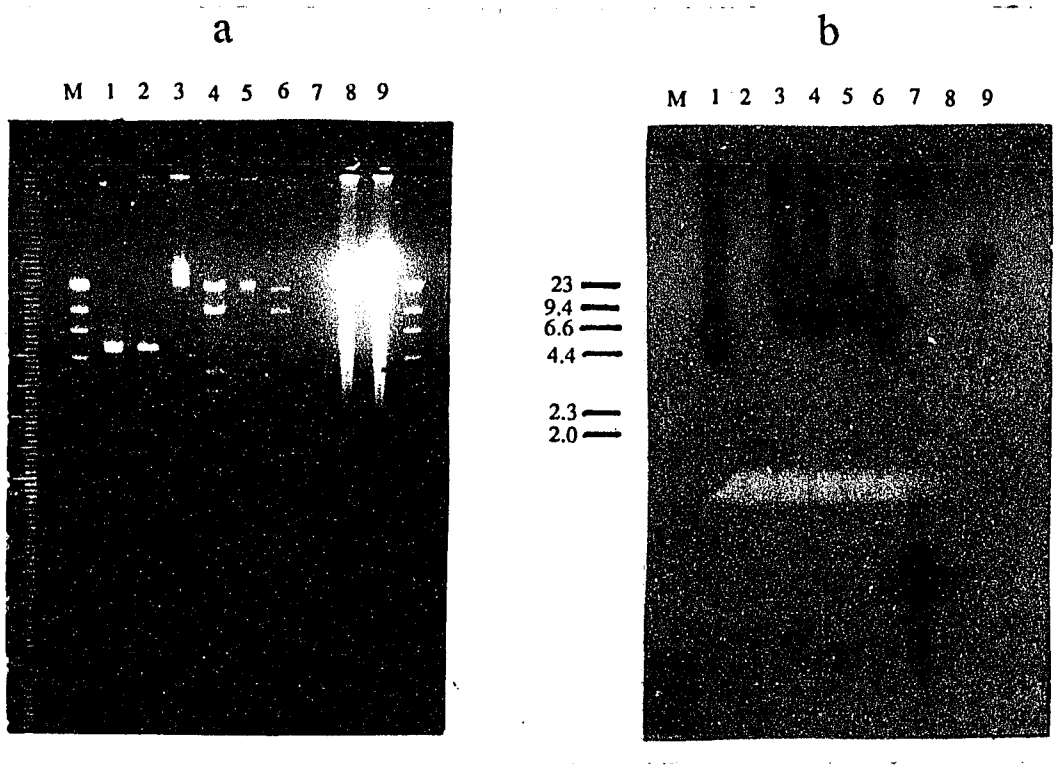


Figure 42:

Hybridization of the GE13 insert and genomic DNA to an E13 probe.

(a): Digestion of GE13 and genomic DNA with EcoR I. Lane M contains molecular markers; lanes 1 and 2 are inserts from different genomic clones; lanes 3 and 5 are undigested GE13; lanes 4 and 6 are GE13 digested with EcoR I; lanes 8 and 9 are GE13 are digested with BamH I; (b): The DNAs on the gel in (a) was blotted and hybridized with an E13 probe.

Figure 44

The partial sequences of the PGE13 were obtained using primers SP6 and PSP6.2 in A; primers T7 and PT7.2 in B; specific E13 primers PE13 and PE13F in C.

A. The partial sequence of GE13 from the direction of Sp6 promoter primer

ATTTGCAGTTGGTTCTACTTCATAAACAGAGTCTAACAAGCGAGATAAACTAGGATTTAAGAACTGGCTT
TCTGGGTTATTATTTTGAGTTTTTTCTTCATTTTTAAAGGAAACTAGATTAGGTGTGGCACAAGCTACAA
ATAAATGAGAAATTGCAAAAAAAGTCGTAAAAATTAACGGGAGTTTTTTGTTCATATTATTCCTTTATTT
GTTCTTCTATGACTTGATTGTAGTAAGATCATTTAGATTTTTGGTTTTTTGCTCCAGTGTAGTAAATTTT
TCCTTTATTTATTAAAGTAACTGAATCTACTATGTGATCAATTTCAGAAAGATTATGACTGCTTAAAAAA
ATAGTTTTTCCTTCTTTTTTAAGATTTTGAATGATTTGGATAAAATCAAATCTACTTGTGGGATCTAGAT
TAGCTAAGGGTTCATCTAAAATAATTAATTCTGGATCGTGAATTAGAGCTTGAATTAATGCAATTTTTCT
TTTTTGACCTGAAGAAAAATCATATGGTTTTGCATTTGCAAGACTGAATATTAATGAAGATAAAAATAAA
TCAATATCTTTGCTTGCTTTTTGTTTATCAAATTGAAGTGTTGCTAATAAATTAAATATCTTCG

B. The partial sequence of GE13 from the direction of the T7 promoter primer

AGTGAAGCGGTGGAATGCGTAGATATTAGGAAGAACACCAATAGGCGAAGGCAGCTAACTGGTTATATAT
TGACACTAAGGGACGAAAGCTTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGAT
GATCATTAGTTGGTGGAATAATTTCACTAACGCAGCTAACGCGTTAAATGACTCGCCTGAGTAGTATGCC
TGCAAGAGTGAAACTTAAAGGAATTGACGGGAACCCTGCACAAGCTGTTGGAGCATGTGGTTAATTTGAT
TCTACGCTGATGAACCTTACCCACTCTTGACACTTCTGCAAAGCATATAGAGATATGTGGAGGTTAACAG
AATGACAGATGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTAGGTTAAGTCCTGCAACGAGCGC
AACCCTTTTCTTTAGTTACTAATATTAAGTTAAGGACTCTAGAGATACTGCCTGGGTAACCAGGAGGAAG
GTGGGGACGACGTCAAATCATCATGCCTCTTACGAGTGGGGCAACACACGTGCTACAATGGTCGGTACAA
AGAGAAGCAATATGGTGACATGGAGCAAATCTCAAAAAACCGATCTCAGTTCGGATTGAAGTCTGCAACT
CGACTTCATGAACTCGGAATCGCTAGTAATCGTAGATCAGCTACGCTACGGTGAATACGTTCTCGGGTTT
GTACACACTCGC

C. The partial sequence of GE13 from the E13 specific area

ATTGCAGGTCTACATGATTCTGCTGCTAGATACCTATCTGGAAGGGTCGTTCTCGGTACAGGTTGATAAA
TAACACTAGAGCTTCTGAACATGAAATCATTCAATTCGTTACTCGACCGAAGTCTTCACTATGCATCACA
ACTCAAGTATGCTATACGGATTTGCCTATATAACTCTCTCTTGTTGCTTACCCCACAATCCAGTAAGTGG
TAGAACTATCCTTTTTCGTCACTCCATCATTCTTTTACCAAGTACAGGAATATTAACCTGTTGTCCATCG
ACTATGCATTTCTGCCTCATCTTAGGCCCTGACTAACCCTGGTGGACGAACCTTGCCCAGGAAACCTTTC
CCAATAGGCGTCGAAGATTCTCACTTCGAATCGTTACTCATGCCGGCATTCTCACTTTTAATCTCTCCAC
CAGTCCTCACGGTCTAACTTAACTAAATTAAAACGCTCTCCTAACGCACATAATGTGCCCGTAGCTTCGG
TATTGTGTTTGAGTCCCATTACATTATTGGCGCAAGATCTCTTGACTAGTGAGCAATTACGCACTCTTTA
AAAGGTGGCTGCTTCTAAGCCAACTTCCTAGTTGTTTATGAAATCTCACAACCTTTCTCACTTAACACAA
TTTTGGGACCTTAGCTGACGATCTGGGTTGTTTCCCTCGCGAGCGTGGACGTTATCACCCACGTTCCGAC

TGCATAGCTTAATTTTAATGGTATTCGGAGTTTGATTATAGTCAGTACAGCTAGATGCCGCCATTCCACT
TTCAGTTCTCTACTCCCATTAAACAATACTATTCTCTAGCCCTAAAGCTATTTCGAAGAGAACCAGCTAT
CTCCAGGTTCGATTGGAATTTCACCGCTATCCACAAGTCATCCGGGCACTTTTTAGCGTACTACGGTTCG
GTCCTCCACTTAGTTTTACCTAAGCTTCAACCTGCTCATGGATAGATCACCTGGTTTCGGGTATATGCCA
ATATACTAAAGTCGCCCTATTCAGACTCGGTTTCCCTACGGCTCCGCTTTTTTCTGCTTAACCTTGCATA
TAACATAACTCGCCGGCCCATACTGCAAGATGTACGCCATCACACTTAACGTTGCTCTGACTACTTGTAA
GTAAATGGTTTCAGAATCTATTCACTCCCCTCTGGTTCTTGCACTCCTCACGTACTAGTCACTATCGGCT
GATTAGATTAGCTACGGTGTCGC

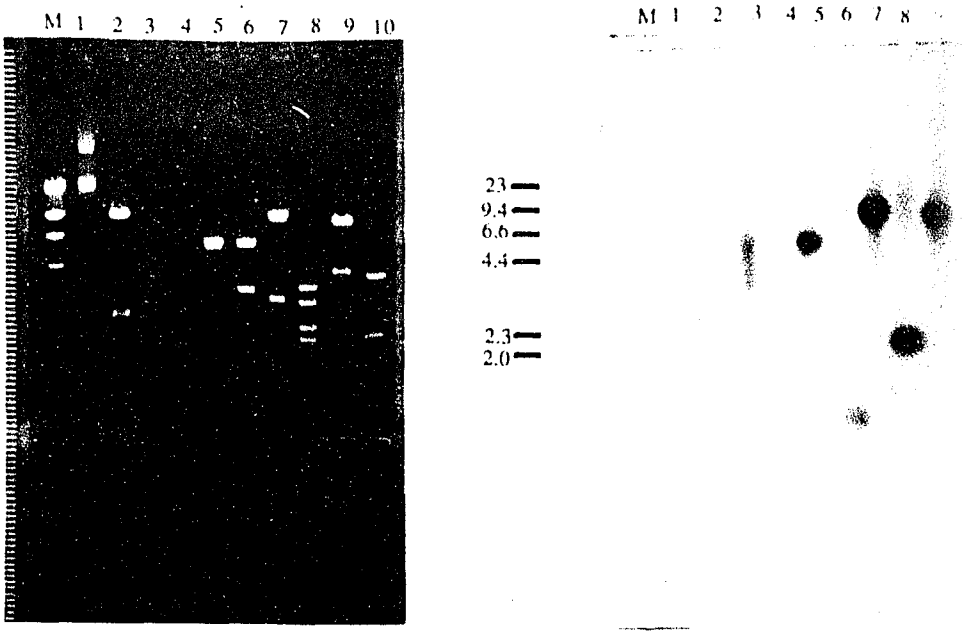


Figure 45:

Restriction enzyme digestion of PGE13. A: lane M contains molecular weight markers; lane 1 contains uncut PGE13; PGE13 was digested with EcoR I(lane 2), Sac I(lane 3), Kpn I(lane 4), Sma I(lane 5), Ava I(lane 6), Xba I(lane 7), Hinc II(lane 8), Pst I(lane 9) and Hind III(lane 10). B: Hybridization of the blot in A to an E13 probe.

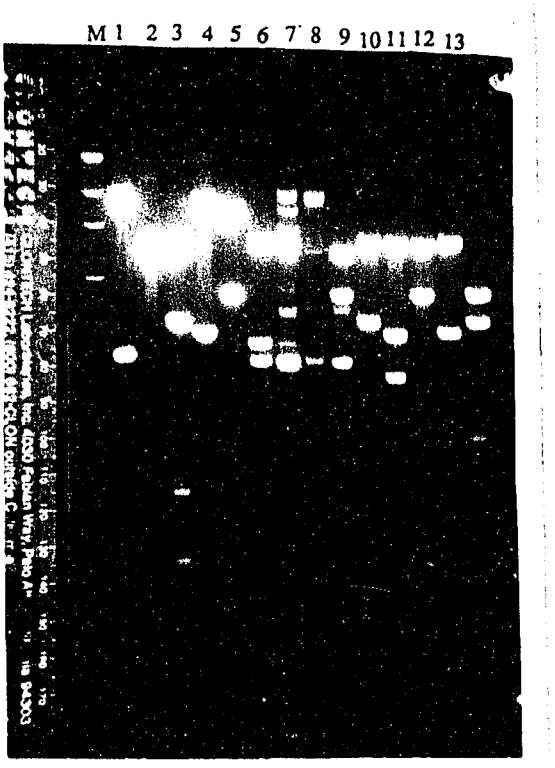
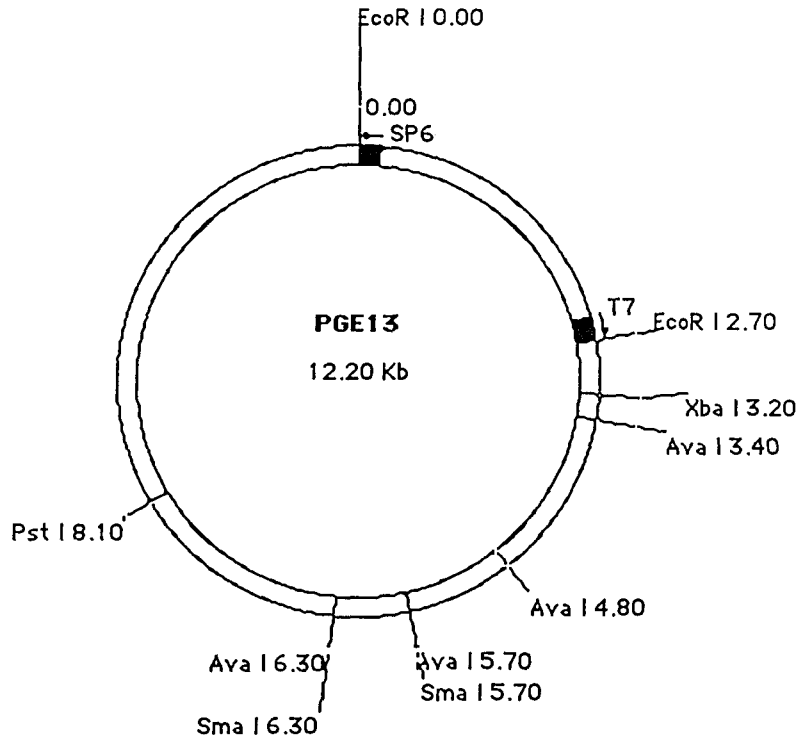


Figure 46:

Restriction enzyme digestion of PGE13. Lane M contains molecular weight markers. PGE13 was digested with EcoR I (lane 1), Sma I(lane 2), Ava I(lane 3), Xba I(lane 4), Pst I(lane 5), and double digested with EcoR I and Sma I(lane 6), EcoR I and Ava I(lane 7), EcoR I and Xba I (lane 8), EcoR I and Pst I(lane 9), Sma I and Ava I (lane 10), Sma I and Xba I(lane 11), Sma I and Pst I (lane 12), Ava I and Xba I(lane 13), Ava I and Pst I (lane 14).



B

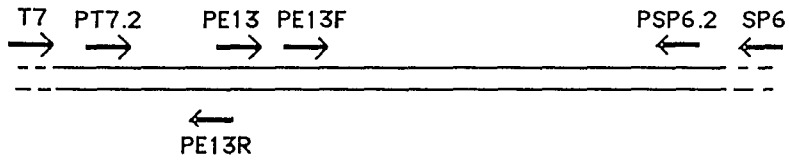


Figure 47
A. Restriction map of PGE13
B. Positions of sequence primers in PGE13

DISCUSSION

1. Expression of novel genes in SV40 transformed keratinocytes

Candidate cDNAs representing sequences specifically induced by SV40 were selected by eliminating cDNAs hybridizing to labeled heterogenous cDNA probes derived from normal human epithelial cells (subtraction hybridization). Further northern blot studies showed that from these candidates two cDNAs, E12 and E13, were found to be markers of epithelial transformation in that the mRNAs show highly induced expression not only in SV40 transformants but also in cell lines derived from various squamous cell carcinomas (Fig. 12). Malignant transformation is known to require altered gene expression in the transformed cells which ultimately give rise to tumors. Altered expression of oncogene in transformed cells is perhaps the most obvious and widely appreciated example of this process (84, 85). Highly induced expression of E12 and E13 emerges early during SV40 transformation process and which are shared in common by transformed keratinocytes which arise independently from one another. These stable changes of expression are likely to play some significant role in transformation even if the target genes of SV40 are not themselves oncogenes. This indicates that the expression of the E12 and E13 may be transactivated directly or indirectly by SV40 virus. The significance of E12 and E13 is based on the belief that cDNA whose expression is induced in the long term transformants may represent genes with oncogenic or growth promoting properties. The process of transformation of

human keratinocytes by SV40 takes place through distinct stages over a period of many months following infection (33, 51, 63, 80). Fully malignant tumors are believed to arise from a single cell or a small number of initiated cells whose progeny are exposed to strong selective pressures which shape the malignant phenotype. Immortalization occurs in a few SV40 infected cells which emerge from a crisis period to become capable of indefinite life span and form permanent cell lines in culture. This could be interpreted as the result of the loss of a dominantly acting growth repressor gene activity (82, 83) and activation of proliferation related gene activity. Recently the supressor genes, RB, P53 and WT1, have been studied, but the targets of these genes are still not well known. The fact that expressions of E12 and E13 which are highly induced in both transformed and carcinoma cell lines indicate the E12 gene product may be involved in normal regulation of cellular growth and proliferation as well as in neoplastic transformation. We also found that the E12 gene was induced in a very early stage of SV40 infection, as early as the first passage following infection (Fig. K). This indicates that E12 expression is enhanced immediately after infection and kept preferentially expressed at high levels throughout the transformation process. It is possible that the T antigen may directly activate the expression of E12 gene and the E12 gene may be the one of genes involved in the early loss of control of proliferation. The E12 gene product may participate in transformation as a cellular oncogene or may simply be a relay in the elaborate biochemical

circuit that governs the proliferation of cells. Another interesting finding from the analysis of Northern blot is that E12 expression in Hela cell is always lower than in SCC lines and other transformed cell lines. It may be that E12 expression may not be as critical for Hela cell transformation or tumorigenesis as for the viral transformants. Pereizasmith (92) applied cell fusion to assign immortal human cells to distinct groups and provide the genes important in normal growth regulation that have been modified in immortal cells. Hybrids from the fusion of normal human diploid fibroblasts to different immortal human cell lines exhibited limited division potential. This lead to the conclusion that the phenotype of cellular senescence is dominant and that immortal cells arise as a result of recessive changes in the growth control mechanism of the normal cell. Exploiting the fact that immortality is recessive, they fused immortal human cell lines with each other, assigned 21 cell lines to four complementation groups for indefinite division (92, 93). SV40 transformed cells and Hela cells were assigned into two distinct groups, group A and group B, respectively. If SV40 transformed cells are fused with Hela cell, the hybrid will exhibit limited division potential. This assignment indicates that SV40 transformed cells and Hela cells exhibit complimentarity and may be immortalized through somewhat different pathways. If the E12 product is involved in the immortalization of SV40 transformed cells but not in Hela cells, this would be consistent with this hypothesis. It would be interesting to test the fusion hybrid of SV40 transformed

keratinocytes and Hela cells for complimentarity in the same way. According to the assignment of immortal human cells developed by Pereizasmith(92), the hybrids will have limited division potential because of complementation of different recessive changes that had led to immortalization in the parent cell lines. We can test the expression of E12 and E13 in cells which are fused from SV40 transformants and Hela cells and the expression of E12 and E13 genes in the hybrid might shed some light on the roles of E12 and E13 genes during immortalization, i. e., transformation. If the products of E12 and E13 are the targets of repressor genes, they may be suppressed in the hybrid. This could be studied directly by immunocytochemistry and/or *in situ* hybridization.

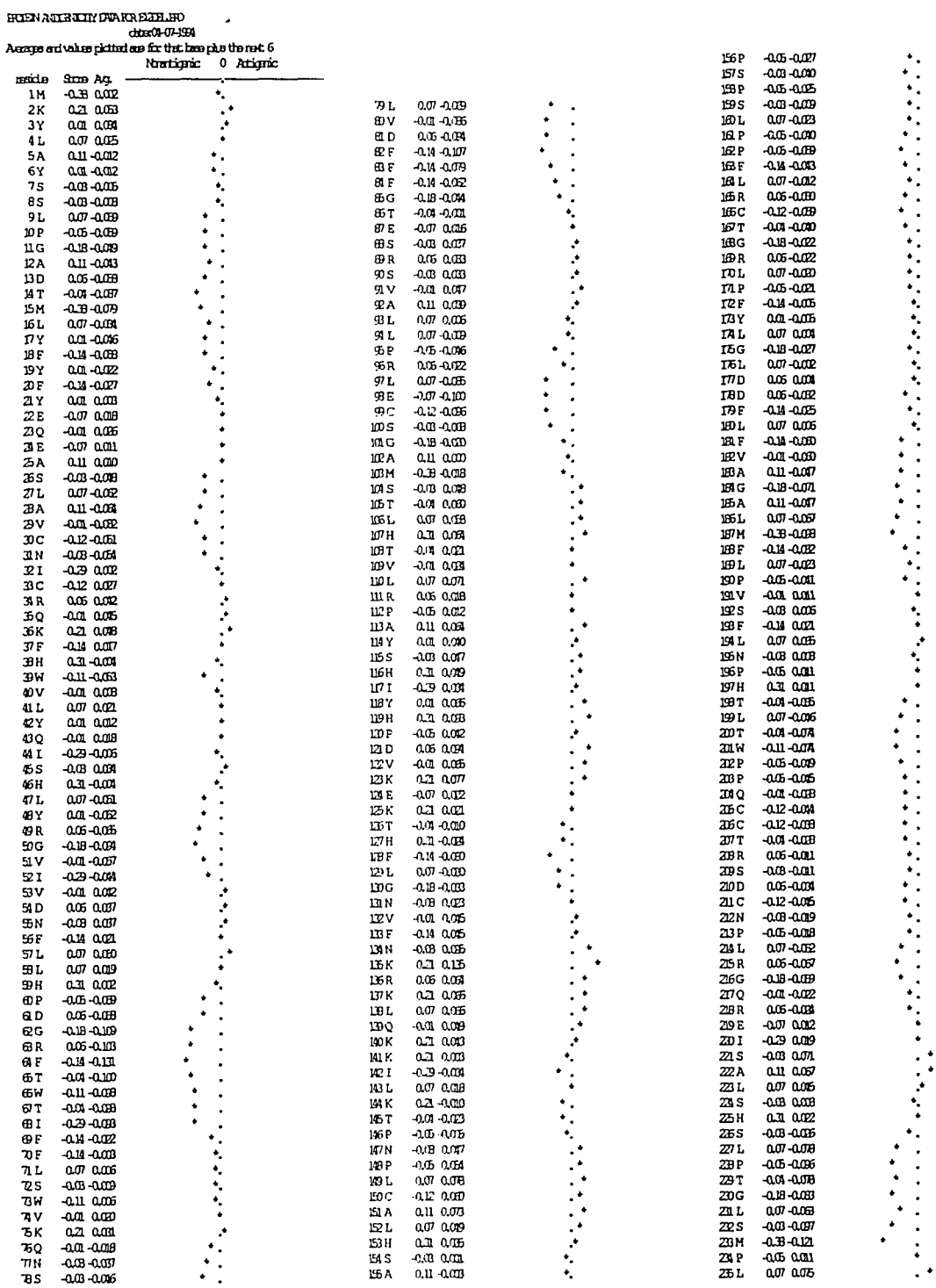
2. Transformation related cDNA markers

Altered gene expression also can be used as a marker of transformation. Our laboratory has demonstrated viral dependent alteration of transcription of the keratin intermediate filament genes as molecular markers of transformation (51, 73, 80,87). The expression of certain keratins is related to the proliferative state of cells(74, 75, 76). In keratinocytes transformed by the oncogenic virus, SV40, viral-dependent alteration of transcription of the keratins including ectopic expression of genes coding for simple epithelial/fetal keratins. The significance of these findings stems, in part, from the clinical observation that the expression of fetal keratins accompanies malignant transformation of epithelial cells and is

now widely used as a tool in tumor diagnosis (90, 91, 129, 130). The keratin 1, 2 and 10 are expressed mainly in cornified stratified epithelial. The smaller keratins (7, 8, 18 and 19) are formed mainly in simple epithelia. A SV40 transformed keratinocyte line, SV-HE(line 115), isolated by Steinberg and Defendi(33), was found that SV-HE had lost most of its keratin 5, 6, 14/15 16 and 17, a set highly characteristic of simple epithelial cell, instead expressed simple epithelial keratins 8, 18 and 19 (74). Thus the transformed cells evidently have a simple epithelial nature and are unable to stratify or accumulate the cell envelope precursor, involncrin (74). E12 gene is induced in the very early stage after infection and maintained during long term transformation (Fig. K). This finding shows great potential for use of the E12 gene as an marker for transformation and tumor diagnosis.

The putative E12 protein sequence is known from the open reading frame. Antibodies to the oligopeptide representing subfragments containing antigenic epitopes in E12 can be determined by computer analysis of the polypeptide (91, Figure 48) and can be commercially made. Such antibodies might be used to compare E12 protein expression in both normal and transformed keratinocytes and tissue biopsy samples, to assess the potential for clinical cancer diagnosis. Immunocytochemic staining of cells will directly show the cytological localization *in situ* and relative amounts of E12 products. Further characterization of E12 product will be performed. More biochemical data of the product of E12

Figure 49. Computer analysis of the open reading frame of E1 shows the putative protein antigenicity.



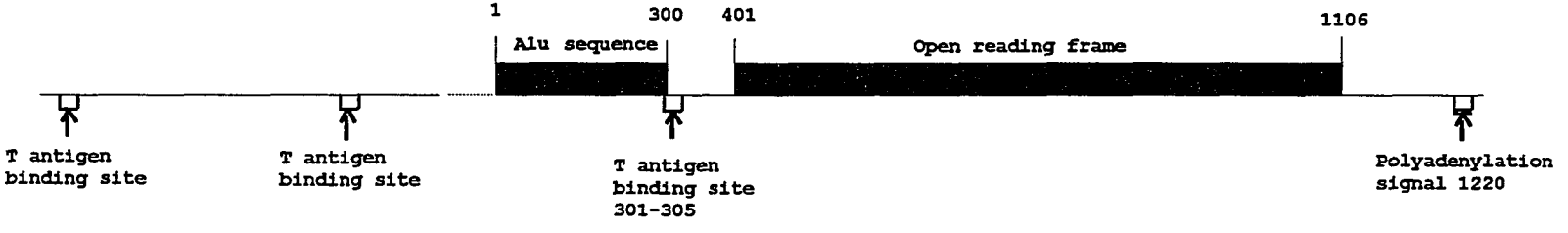


Figure 49
Above shown is the transcription unit of E12 and the T antigen binding site near the Alu sequence.

will be obtained by Western blot experiments. A two dimensional Western blot will identify the E12 product and proteins which share the same epitopes with E12. Assuming that proteins have same structure, i, e, epitopes, may have similar functions and belong to a same group, to analyze those proteins may lead to more information about E12 product. The product of E12 can be quantitatively determined by immunocytochemical staining and protein isolation. Comparing the amount of E12 product in normal and transformed cells, more information about the control of E12 expression will be obtained. Even if more mRNAs of the E12 are transcribed after the transformation with SV40 virus, the amount of E12 product in normal and in transformed cells are in the same level, it may mean that expression of E12 is also under posttranscriptional control.

3. Transfection with E12 and E13

The ability to detect viral oncogenes by gene transfer set the stage for identifying biologically active cellular oncogenes. Robert Weinberg and colleagues found that DNAs from chemically transformed rodent cell lines induced transformation of recipient NIH 3T3 cells (88). The transforming activities of E12 and E13 were assessed by transfer of E12 and E13 into recipient NIH 3T3 cells. A standard focus forming assay (89) were used. For further study, normal human keratinocytes may be used as recipient cells, because those genes are possibly critical for transformation only in epithelial cells. Also, the cDNA clones of E12 and E13 from transformed keratinocytes

instead of genomic DNA clones from normal keratinocytes may be used for transfection. The efficiency of transformation induced by normal cell DNA is about 100-fold lower than that induced by an active viral oncogene (90). The clones of the E12 and E13 from normal cells should be obtained first and transfection with cDNAs of E12 from normal and transformed cells may give different results If the E12 gene is altered during the transformation.

4. Transcriptional activation of E12 and the role of SV40 large T antigen

The potential importance of DNA rearrangements, including proviral insertion and chromosome translocation, has been recognized. One model for tumorigenesis by retroviruses lacking oncogenes is known as insertional mutagenesis (91). According to this model, integration of proviral DNA near a specific host cell target gene is a mutagenic event which initiates neoplasia. In principle, insertional mutagenesis could occur by either of two mechanisms. First, insertion of proviral DNA within a gene could result in its inactivation. Alternatively, integration of proviral DNA could activate gene expression by disrupting normal regulatory sequences and bring a target host gene under control of the viral LTRs. In both of these cases, the virus induce neoplasia indirectly. To test whether changes in the E12 and E13 genomic sequences induced by viral infection, the restriction patterns of genomic sequences were examined in Southern blots.

Based on the southern blot analyses (Fig. 22, 23), no obvious change in restriction pattern was observed. These results indicate that SV40 infection per se does not lead to rearrangement of E12 and E13 sequences. The enhanced expression of E12 and E13 seems not to be the results of insertion mutagenesis.

T antigen auto regulates the level of SV40 early mRNA transcription through direct binding to sequences within the SV40 origin of replication(124). Transcription activation of the SV40 late promoter by T antigen has been well studied (112, 113, 114, 115, 116). T antigen is able to transactivated transcription from a variety of other viral and cellular promoters including E2 and E3 promoters (117, 118), the chicken _ 2 collagen (118), the human heat shock protein 70 (119), and _ and ß globin promoters(120), and intracisternal- A particle endogenous retroviral LTRs (121). The RTVL-HLTRs, long terminal repeats of a large family of human endogenous retrovirus like sequences, are present in human genome. In human, these elements are dispersed on all chromosomes and are expressed in several transformed cell lines (122). Interestingly, recent studies (Antila et al, 122) have shown that T antigen can induce expression in vitro from an RTVL-HLTR which is normally quiescent in addition to significantly increasing the promoter activities of several, but not all, RTVL-HLTRs. These findings suggest that promoters and their associated transcripts may be transcriptionally stimulated by SV40 T antigen. Is it possible that SV40 T antigen may transactivate the expression of E12 gene

by stimulating the E12 promoter. Primer extension experiments could be utilized to obtain the sequence of promoter region and more information about the regulation of E12 could be obtained. Homology searches of the E12 sequence using the Fast A algorithm of Pearson and Lipman (94) against the Genbank sequence database showed that the first 300 bases of 5' end of E12 have a high degree of homology to a number of other genes previously catalogued in the database (Fig. Z9). However this area contains human Alu repeated sequences. Alu sequences constitute 3 to 6% of the total mass of the human DNA and are of the most abundant family of moderately repetitive DNA sequence in the human genome. A typical human Alu family member is a sequence of about 300 base pairs (9) and It consists of two directly repeating monomer units. The first monomer is usually about 120 bp, and the second monomer is about 150 bp. Two A-rich regions were present at the end of the first monomer and at the end of the second monomer (96, 97). The Alu family represents one of the most prevalent classes of mobile elements, having amplified to a copy number in excess of 500,000 within the human genome (97). They are usually located in intergenic DNA or in introns. Alu members have been reported to be transcribed by RNA polymerase III from a polymerase III promoter that is located inside the Alu sequence (98, 99, 100). The 5' region of the Alu sequence of the E12 gene is interesting since normally the appearance of Alu sequences in mature mRNA is not observed. Alu sequences were first observed in the early 1980s() in mature mRNA. Since then many other mRNAs containing Alu sequences have

been reported (101, 102, 103, 104, Fig. 34). What type of role that Alu families might play *in vivo* is not clear. Various speculations have been made, e.g., that they are the start sites of normal cellular DNA replication (105, 106), modulators of chromatin structure (107), hot spots in recombination (108), inhibitors of gene conversion (109) or mRNA stabilizes in the cytoplasm (110). However, no convincing experimental evidence supporting any one of these speculations has yet been reported. Oliviero (111) reported that a DNA fragment containing a copy of the Alu repeated element enhances transcription of the linear specific heptoglobin related (Hpr) promoter in hepatoma cell lines. In transfection experiments in the human hepatoma cell line Hep G2 chimeric constructs of the Hpr promoter, which contain an Alu repeated element upstream from the transcriptional start site, are much more transcriptionally active than the constructs without Alu repeated elements. It was surprising that an Alu repeated element could enhance the accurate transcription of a downstream RNA polymerase II promoter. However, it dose appear that ALu sequences may play a role in transcription activation. There is also a possibility that an Alu sequence such us that found in the E12 gene may be involved in the regulation of transcription. During the infection by SV40, a high level of E12 expression may be induced by T antigen directly or indirectly. T antigen may directly transactivate the E12 gene by binding to nucleotide sequence motifs important for transcriptional control. These motifs in E12 gene could be inside of the promoter area or the inside of

the Alu sequence area. T antigen autoregulates the level of SV40 early mRNA through direct binding to sequences on the early side of the origin of replication (124). To determine the presence of such motifs experimentally, gel shift and DNase I footprint experiments could be performed to demonstrate the T antigen binding sites on E12 gene. If there are T antigen binding sites within the E12 gene, these sites can be introduced into a CAT assay system. The transactivation function can be checked by cotransfecting a CAT construct system with a T antigen expression plasmid. On the other hand, T antigen more likely transactivate genes indirectly though cellular regulatory proteins. In fact, a mutant heat shock 70 promoter with all known T antigen binding sites removed still retains the ability to be transactivated by T antigen(119). T antigen can stimulate transcription in the absence of sequence specific binding to the activated promoter (125). Experiments using a mutant T antigen which dose not enter into nuclei show that the transactivation by T antigen does not require its nuclear localization (126). The amount of transcription factor Sp1 is increased in SV40-infected cells (127). This finding indicates that the direct binding to promoter or control region as a transcriptional stimulator is unlikely take place in most cases. Gorbunova (131) employed a temperature sensitive mutant of SV40, tsA239, which synthesizes a defective T antigen at $40^{0}C$, to investigate the mutagenic activity of T antigen in Chinese hamster cells under permissive and nonpermissive temperature. These tsA mutants transform cells under permissive conditions and do not transform

them under nonpermissive conditions (132,133). Thus at permissive temperature tsA239 SV40 induced both gene mutations and chromosome aberrations in hamster cells. The results suggest that T antigen transform cells possibly by inducing gene mutations and chromosome aberrations. It is possible that T antigen can activate the E12 gene through indirect interaction with the cellular protein normally involved in regulating E12 gene transcription. To test the transactivation of E12 by T antigen, a CAT plasmid construct with the E12 promoter area could be cotransfected with a T antigen expression construct. If T antigen dose activate the promoter of E12 *in vivo*, identification of those factors and the sequences with which they interact will aid our understanding of the mechanisms governing the regulation of E12 and the role of E12 during the transformation of SV40 virus.

Appendix: GENBANK entrte of the E12 sequence

```
LOCUS       HUMTRRP      1240 bp ss-mRNA      PRI    24-SEP-1993
DEFINITION  Human transformation-related protein mRNA, 3' end.
ACCESSION   L24521
KEYWORDS    Alu-like repeat; transformation-related protein.
SOURCE      Homo sapiens epithelium cDNA to mRNA.
  ORGANISM  Homo sapiens
            Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
            Eutheria; Primates; Haplorhini; Catarrhini; Hominidae.
REFERENCE   1  (bases 1 to 1240)
AUTHORS   Shen,H. and Steinberg,M.L.
JOURNAL   Unpublished (1993)
STANDARD  full staff_review
FEATURES             Location/Qualifiers
CDS             <1..1109
                /product="transformation-related protein"
                /codon_start=3

/translation="DRLSLLSPRLECNGMILAHCKLRLPGFKRFSCLSLPSSWDYRHV

PPRQVHFVFSVETGFHRAGQAGLELLTSSVPPTSAFPKCWDYRRDDQAWPTLSSFRGL

NKFAFLPKFFAHPISQFQRVECNVGCPILLAMKYLAYSSLPGADTMLYFYFYEQEASL

AVCNICRQKFHWVLYQISHLYRGVIVDNFLLHPDGRFTWTIFFLSWVKQNSLVDFFFG

TESRSVALLPRLECSGAMSTLHTVLRPAYSHIYHPDVKEKTHFLGNVFNKRKLQKKIL

KTPNPLCALHSAPSPSLPPFLRCTGRLPFYLGLDDFLFVAGALMFLPVSFLNPHTLTW

PPQCCTRSDCNPLRGQREISALSHSLPTGLSMPL"
repeat_region   1..283
                /rpt_family="Alu"
                /note="Alu-like repeat"
polyA_signal    1221..1226
polyA_site      1240
source          1..1240
                /organism="Homo sapiens"
                /cell_type="SV40-transformed keratinocyte"
                /sequenced_mol="cDNA to mRNA"
                /tissue_type="epithelium"
BASE COUNT      276 a    327 c    260 g    377 t
ORIGIN
gagacagact ttcgctcttg tcacccaggc tggagtgcaa tggaatgatc ttggctcact
gcaaacttcg ccttccaggg ttcaagcgat tctcctgcct cagcctcccc agtagctggg
attacaggca tgtgccacca cgccaggttc attttgtatt ttcagtagag actgggtttc
atcgtgctgg ccaggctggt ctcgaactcc tgacctcaag tgttccaccc acctcagcct
tcccaaagtg ctgggattac aggcgtgacg accaagcctg gccaaccctc tcgtctttca
ggggcttaaa taaatttgct tttctcccca aattctttgc ccaccccatc agccagtttc
aaagggtgga gtgtaacgtg ggttgcccaa tcctgcttgc catgaagtac ttagcttaca
gcagtttgcc aggcgcagac accatgcttt atttctattt ctacgaacaa gaagcgtctc
ttgcagtctg caatatttgt cgtcaaaaat ttcactgggt gttgtatcaa atcagccacc
tgtatagagg tgtgatcgtt gacaattttt tgttacatcc cgatggccgt tttacctgga
```

ccattttctt cctttcatgg gtgaagcaaa actccttggt agacttttt tttgggacag
agtctcgctc tgtcgctctg ttgcccaggc tggagtgcag tggtgcgatg tcgactctcc
acacagtgct caggccagct tatagtcata tatatcaccc agacgtaaag gaaaagacac
atttttagg aaatgttttt aataaaagaa aattacaaaa aaaaatttta aagaccccta
acccttttgtg tgctctccat tctgctcctt ccccatcgtt gcccccattt ttgaggtgca
ctgggaggct ccccttctat ttggggcttg atgactttct ttttgtagct ggggctttga
tgttccttcc agtgtcattt ctcaatccac ataccctaac ctggcccccct cagtgttgca
ccagatctga ttgtaaccca ctgagaggac agagagaaat aagtgccctc tcccactctc
ttcctactgg tctctctatg cctctctaga gtctcgtctc ttttaccctg gcccctctcc
cttgggctct gatgaaaaat tgctgactgt agctttggaa gtttagctct gagaaccgta
gatgatttca gttctaggaa aataaaaccc gttgattact

mia T10 9/27/93
Mia McLeod mzscc@cunyvm.cuny.e 9/27/93 Annotation of your
submission

BIBLIOGRAPHY

1. Pitot, H. C., et al. (1988). Multistage carcinogenesis. In "Theories of Carcinogenesis" (O. Sversen, ed.) pp. 159-177 Hemispgere, New York.
2. McCormick, P. J., et al. (1982). Proc. Natl. Acad. Sci. U. S. A. 79, 4342-46.
3. Pitot, H. C., et al. (1990). Annu. Rev. Pharmacol Toxicol 30, 465-500.
4. Moore, A., et al. (1987). Carcinogenesis 8, 483-6.
5. Pitot, H. C., et al. (1987). Carcinogenesis 8, 1491-99.
6. Harris, C. (1991). Cancer Res. 51, 5023s-5044s.
7. Fearon, E. R. (1990). Cell 61, 759-67.
8. Lasko, D., et al. (1991). Annu. Rev. Genet. 25, 281-314.
9. Parker, C., et al. (1981). Proc. Natl. Acad. Sci. U. S. 78, 5842-46.
10. Cooper, G. (1990). Cellular Oncogenes. In "Oncogenes", pp. 65-95. Jones and Barllett Publishers.
11. Marshall, C. J., et al., (1991). Cee 64, 313-26.
12. Martin, G. S. (1970). Nature 227, 1021-1023.
13. Vogt, P. K. (1971). Virology 46, 939-946.
14. Stehelin, D. (1976). Nature 260, 170-173.
15. Corcoran, L. M., et al. (1984). Cell 37, 113-122.
16. Nilsen, T. W., et al. (1985). Cell 41, 719-26.
17. Marx, J. (1989). Science 243, 1012-13.
18. Slagle, B. L., et al. (1992). Prog. Med. Virol. 39, 167-203.

19. Naimski, P., Steinberg, L. M. (1985). J. Cellul. Biochem. 29, 95-103.

20. Branston, E., et al. (1985). Biochim. Biophys. Acta 780, 67-94.

21. Pitot, H., et al. (1993). Cancer 72, 3: 962-970.

22. Martin, P., et al. (1989). Cell 59, 21-23.

23. Ostrow, R. S., et al. (1993). Virology 196(2), 861-7.

24. Liang, X. H., et al. (1993). Oncogene 8(10), 2645-52.

25. Yamashita, T., et al. (1993). Oncogene 8(9), 2433-41.

26. Zhang, S., et al. (1992) Virology 191(2), 793-802.

27. Van den Henvel, S. T. (1993). J. Virology 67(9), 5226-34.

28. Whyte, P., et al. (1988). Nature 334, 124-129.

29. Yu, D. et al. (1991). Mol. Cell Biol. 6, 887-89.

30. Nomura, J. G., et al. (1981). Nature 291, 346-349.

31. Ng, S-C., et al. (1985). J. Biological Chem. 1127-1132.

32. Kelly, T. (1988). J. B. C. 263(34), 17889-92.

33. Steinberg, M., et al. (1983). J. Invest. Dermatology 81, 131s-136s.

34. Steinberg, M., et al. (1982). Exp. Cell Res. 139, 369-75.

35. Tooz, et al. (1981). DNA Tumor Virus. Cold Spring Harbor Lab.

36. Stahl, H., et al. (1987). Biochem. Biophys. Acta 910, 1-10.

37. Kimura, G., et al. (1975). Proc. Natl. Acad. Sci. 72, 673-77.

38. Martin, R. G., et al. (1975). J. Virol. 15, 599-612.

39. Brockman, W. W. (1978). J. Virol. 25, 860-870.

40. Robinson, J. K., et al. (1987). J. Dermatol. Surg. Ocol. 13:

283-294.

41. Garin-Chesa, P., et al. (1986). Am. J. Surg. Pathol. 10: 829-835.

42. Gorbunova, L. V. et al. (1982). Mol. Gen. Genet. 187: 473-476.

43. Martin, R. G., et al. (1975). J. Virol. 15: 599-612

44. Rigby, P., (1979). Nature 282: 781-784.

45. Colby, W., et al. (1982). Proc. Nat. Acad. Sci. U. S. A. 79, 5189-93.

46. Steinberg, M., et al. (1981). J. Cellul. Physiol. 109,153-59.

47. Tegtmeyer, P. (1975). J. Virol. 15, 613-68.

48. Aaronson, S. A., et al. J. Cell. Physiol. 72, 141-148.

49. Winesbourne, D. J. (1978). J. Biol. Chem. 253, 5109-5120.

50. Lewis, A. M., et al. (1979). Pro. Nat. Acad. Sci. U. S. A. 76, 4299-4302.

51. Steinberg, M., et al. (1986). pp. 285-299. In "Process in Culaneous Epidermal Differentiation" (Praeger Scientific: New York).

52. Tevethia, M. J., et al. (1984). Virology 137, 414-421.

53. Petit, C., et al. (1983). Virology 127, 74-82.

54. Asselin, C., et al. (1985). J. Virol. 56, 958-968.

55. Livingston, D., et al. (1987). Mol. Biol. Med. 4, 63-80.

56. DeCaprio, J. A., et al. (1988). Cell. 54: 275-283

57. Trifillis, P., et al. (1990). J. Virol. 64(3): 1345-1347

58. Cherinton, V., et al.(1988). Mol. Cell. Biol. 8: 1380-1384

59. Tan, T., et al. (1986). J. Virol. 59(3): 574-583

60. Deppert, W., et al. (1987). Mol. Cell. Biol. 7: 4453-4463

61. Yewdell, J., et al. (1985). J. Virol. 59: 444-452

62. Peden, K., et al. (1989). Virol. 46, 151-161

63. Defendi, V., et al. (1982). J. Cell. Physiol. Suppl. 2: 131-140

64. Lin. J., et al. (1990). Virol. 176(1): 302-305

65. Green, H. (1977). Cell 11, 405-16.

66. Pheinwald, J., et al. (1975). Cell 3, 331-344.

67. Gross, M., et al. (1987). Exp. Cell. Res. 171, 460-474.

68. Koprowski, H., et al. (1962). J. Cell Comp. Physiol. 59, 281-292.

69. Steinberg, M., et al. (1979). Pro. Natl. Acad. Sci. U. S. A. 76(2), 801-805.

70. Naimski, P., et al. (1985). J. Cell. Biochem. 29, 95-103.

71. Defendi, V., et al. (1982). J. Cell Physiol. Suppl. 2, 131-140.

72. Edelman, B., et al. (1985). Int. J. Cancer 35(2), 219-225.

73. Chang, S. (1986). Biochem. Biophy. Acta 823, 161-194.

74. Hronis, T., et al. (1984). Cancer Res. 44, 5797-5804.

75. Steinberg, M., et al. (1985). J. Cellular. Physiol. 123, 117-125.

76. Morris, A., et al. (1985). Proc. Natl. Acad. Sci. U. S. A. 82, 8498-8502.

77. Wang, E., et al. (1985). pp. 282-306. In "Intermediate filaments". (The New York Acadmy of Sciences).

78. Morris, A., et al. (1989). Exp. Cell Res. 82, 461-72.

79. Bishop, M. (1991). Cell 64, 235-248.

80. Steinberg, M. (1985). J. Cell Physiol. 123, 117-125.

81. Wright, W. E., et al. (1989). Mol. Cell Biol. 9, 3088-3092.

82. Pereira-Smith, O. M., et al. (1987). Mol. Cell Biol. 7, 1541-1544.

83. Pereira-Smith, O. M., et al. (1988). Proc. Natl. Acad. Sci. U. S. A. 85, 6042-6046.

84. Armelin, H. A., et al. (1984). Nature 228, 1313-1315.

85. Kruyer, W., et al. (1984). Nature 312, 711-716.

86. Gozette, M., et al. (1984). Mol. Cell Biol. 4, 1493-1498.

87. Hronis, T. S., et al. (1984). Cancer Res. 44, 5797-5804.

88. Shih, C., et al. (1979). Proc. Natl. Acad. Sci. U. S. A. 76, 5714-5718.

89. Land, H., et al. (1983). Nature 304, 596-602.

90. Cooper, G. M., et al. (1980). Nature 284,418-421.

91. Neol, B. G., et al. (1981). Cell 23, 323-334.

92. Pereira-Smith, M., et al. (1988). Proc. Natl. Acad. Sci. U. S. A. 85, 6042-6046.

93. Pereira-Smith, O. M., et al. (1983). Science 221, 964-966.

94. Pearson, W. R., et al. (1988). Proc. Natl. Acad. Sci. U. S. A. 85, 2444-2448.

95. Kariya, Y., et al. (1987). Gene 53, 1-10.

96. Jelinek, W. R., et al. (1982). Annu. Rev. Biochem. 51, 813-844.

97. Deininger, P. L., et al. (1981). J. Mol. Biol. 151, 17-33.

98. Elder, J. T., et al. (1981). Nucl. Acids Res. 9, 1171-1189.

99. Allan, M., et al. (1983). Cell 35, 187-197.

100. Hess, J., et al. (1985). J. Mol. Biol. 184, 7-21.

101.Yamamoto, T., et al. (1984). Cell 39, 27-38.

102. Quatrochi, L., et al. (1986). Proc. Natl. Acad. Sci. U. S. A. 83, 6731-6735.

103. Hood, L., et al. (1983). Annu. Rev. Immunol. 1, 529-568.

104. Rothkopf, G. S., et al. (1986). Biochemistry 25, 993-1002.

105. Jelinek, W. R., et al.(1982). Annu. Rev. Biochem. 51, 813-844.

106. Ariga, H. (1984). Mol. Cell Biol. 4, 1476-1482.

107. Duncan, C. H. (1981). Gene 13, 185-196.

108. Rogers, J. (1985). Nature 317, 559.

109. Hess, J., et al. (1985). J. Mol. Biol. 184, 7-21.

110. Robertson, H. D., et al. (1984). Mol. Cell Biol. 4, 310-316.

111. Olivieo, S., et al. (1988). Nucl. Acids Res. 16(4), 1285-1293.

112. Brady, J., et al. (1984). Proc. Natl. Acad. Sci. U. S. A. 81, 2040-44.

113. Keller, J. M., et al. (1985). Mol. Cell Biol. 5, 1859-1869.

114. Brady, J., et al. (1985). Mol. Cell Biol. 5, 1391-1399.

115. Gallo, G. J., et al. (1988). Mol. Cell Biol. 8, 1648-1656.

116. Wildeman, A. (1989). Proc. Natl. Acad. Sci. U. S. A. 86, 2123-2127.

117. Loeken, M. R., et al. (1986). Mol. Cell Biol. 6, 2020-2026.

118. Alwine, J. C. (1985). Mol. Cell Biol. 5, 1034-42.

119. Taylor, I., et al. (1989). J. Biol. Chem. 264, 16160-16164.

120. Cao, S. X., et al. (1989). Biochem. J. 258, 769-776.

121. Luria, S., et al. (1986). J. Virol. 998-1003.

122. Wilkinson, D. A., et al. (1990). J. Virol. 64m 2157-2167.

123. Fenchterm, A., et al. (1991). V. 187, 242-250.

124. Rio, D. C., et al. (1983). Cell 32, 1227-1240.

125. Tack, L., et al. (1985). J. Virol. 54, 207-218.

126. Wildeman, A. (1989). Proc. Natl. Acad. Sci. U. S. A. 86, 2123-27.

127. Saffer, J., et al. (1990). Genes Dev. 4, 659-666.

128. Okubam, K., et al. (1992). Nature Genetics 2, 173-179.

129. Manfredi, J. J., et al. (1994). Biochi. Biophys. Acta 1198, 65-83

130. Sompayrac, L., et al. (1994). Virol. 200(2), 849-53

131. Saenz, R. T., et al. (1994). Mol. Cell. Biol. 14(4), 2686-98

132. Yilmaz, A., et al. (1993). British J. Cancer 68(5), 867-73