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

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

TRANSFORMATION-RELATED CHANGES IN THE SIMIAN VIRUS 40 EARLY GENES IN VIRAL-INFECTED HUMAN KERATINOCYTES IN VITRO

By

Raymond Vázquez

A Dissertation Submitted to the Graduate Faculty in Biochemistry in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy, The City University of New York.

1999

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.

10/28/98
Date

Mark L. Steinberg
Chair of Examining Committee

October 28, 1998
Date

Executive Officer

Supervisory Committee

The City University of New York

Abstract

TRANSFORMATION-RELATED CHANGES IN THE SIMIAN VIRUS 40 EARLY GENES IN VIRAL-INFECTED HUMAN KERATINOCYTES IN VITRO.

by

Raymond Vázquez

Adviser: Professor Mark Steinberg

The experimental procedures described below were designed to study the alterations in the viral early region of the oncogenic virus SV40 integrated into the transformed cells of human keratinocytes. To achieve this goal the endogenous viral sequences of different human keratinocyte cell lines that have become immortalized by the virus were studied. The transformation of these nonpermissive cells can be divided into early and late phases. The viral early genes encode for two

oncoproteins the large T-antigen and the small t-antigen. In the early stage of infection the cultured cell exhibit altered properties of growth which continue to change during the progression of long term culture. The late phase is characterized by the appearance of various transformed properties including accumulation of aberrant viral fragments and the expression of transformation related genes exhibiting altered patterns of gene expression. As transformed cells progress through several hundred cell generations the free viral DNA is gradually lost and only the viral sequences that are integrated into the host remain. Since the viral-infected cells have been carried in vitro for long periods of time, the viral sequences are found as either tandemly integrated full-length viral genomes or as subgenomic fragments. The expression of the viral early genes is considered to be essential for the maintenance of the transformed state, examination of the DNA sequence of the viral early region may identify transforming variants that are selected over long term culture. We sequenced the SV40 viral stock and compared it to the sequences of three different cell lines. Our sequence data has

revealed mutational clusters present in all three of the cell lines studied. The mutational clusters were found at different positions in each individual cell line ruling out the possibility of a mutational hot spot in the viral genome. Experimental analyses in conjunction with a detailed investigation of the mutation sites have demonstrated that each series of mutations confers enhanced growth properties to the cells, which is consistent with the theory that these mutated viral integrants are selectively retained by their host.

Acknowledgments

To my daughter Tamara. You are in my heart always. I love you very much.

To Teresa and Adrian, your love and support helped me get through this ordeal. There is not enough love in this world to thank you both.

To Veronica I am very proud of you. Keep up the good work.

To all my friends and family, you know who you are.

To my mentor Dr. Mark Steinberg, your patience and wisdom made this possible.

To Dr. Myer Fishman and the Minority Biomedical Research Support Program of the NIH (MBRS). Thanks for your help.

To Dr. Jerry Guyden and Dr. Mark Pezzano of the Research Center for Minority Institutes (RCMI). For your guidance.

To my thesis committee for your help.

To the members of the Biochemistry Department at the City University of New York. Thanks Sophie and Dr. Schultz.

Table of Contents:

Page

Introduction

1. Transformation Properties of Simian Virus 40 1

2. Structural Characteristics of the SV40 Viral Genome 1

3. Structural and Biological Characteristics of the Small t-Antigen 3

4. Structural and Biological Characteristics of the Large T-Antigen 4

5. Functional Domains of the Large T-Antigen 5

6. Large T-Antigen in the Viral DNA Replication Cycle 9

7. Large T-Antigen in the Regulation of Transcription 10

8. Regulation of Replication and Transcription by Phosphorylation of the Large T-Antigen 12

9. Role of the Large T-Antigen in Transformation 14

10. Transformation of Human Keratinocytes by the SV40 Virus 19

Materials and Methods

1. Cell culture of SV40 infected human keratinocytes 24

2. Extraction of the genomic DNA 24

3. Digestion of the genomic DNA 26

4. Bacterial Preparation 27

5. Bacteriophage Isolation 27

6. Ligation Procedure 31

7. Screening Procedure 32

8. Hybridization Procedure 34

9. Sequencing Vectors 36

10. Polymerase Chain Reaction 37

11. PCR Cloning Vector 38

12. Transformation of Competent Cells 40

13. cDNA Libraries 42

14. Sequencing Procedure 42

15. Southern Blot Analysis 44

16. Immunoprecipitation and Western Blot Analysis 45

17. Immunofluorescent Staining of SV40 Large T-antigen 48

18. Serum Dependent Analysis 48

19. Colony Formation of SV40-Infected Keratinocytes in Soft Agar 49

20. Computer Analysis of DNA Sequences 50

Results

A- Characterization of Genomic Libraries

1. Analysis of Genomic DNA 51

2. Genomic Libraries in Lambda Gem-11 53

3. cDNA Libraries 57

4. Polymerase Chain Reaction Clones 58

B- Enhanced Polymerase Chain Reaction Products Using Limited Digestion of Genomic DNA.

1. PCR Products Using Complete Digestion of Genomic DNA 61

2. PCR Products Obtained Using Time Dependent Digestion of Genomic DNA Using the SV40 No-cut Restriction Enzyme Xba I 63

3. PCR Products Obtained Using Time Dependent Digestion of Genomic DNA Using the SV40 No-cut Restriction Enzyme Sac I and the SV40 One-cut enzyme BamH I 66

4. PCR Products Obtained During Time Dependent Digestion of Genomic DNA from Cell line 98 68

C- Sequence Analyses

1. Characterization of Sequences Obtained from the SV40 Input Virus 69

2. Sequence Characterization of Clones from Cell Lines 98, 425 and 130 71

3. Sequence Analyses of SV40 Early Region Sequences From Cell Line 98 73

4. Sequence Analyses of SV40 Early Region Sequences From Cell Line 425 75

5. Sequence Analyses of SV40 Early Region Sequences From Cell Line 130 77

6. Compilation of Mutations Found in cell lines 98, 130 and 425 79

D- Immunofluorescent Staining of SV40 T-antigen in a Colony of Viral-infected Human Keratinocytes of Cell Line 98 81

E- Immunoprecipitation and Western Blot Analysis 82

F- Serum Dependent Analysis/Colony Formation Efficiency of Cell Lines 98, 130 and 425 84

Discussion

1- Analysis of the Sequences of the SV40 Input virus and SV40-Transformed Human Keratinocytes.

1. SV40 Input Virus 87

2. Mutations in the SV40 Integrants of Viral-transformed Cell Lines 425, 130, and 98 88

3. Cell line 98 91

4. Cell Line 425 93

5. Cell Line 130 94

6. Summary of the Total Mutations Accumulated in Cell Lines 98, 130 and 425 95

2- Errors in Replication That Lead to Mutational Clusters.

1. Proposed Mechanism by Which Mutations are Arranged into Clusters 96

3- Use of PCR to Study Genomic DNA Templates Infected With Viral Sequences 99

Bibliography 166

List of Tables

1. Position of the Functional Domains of the SV40 Large T-antigen 103

2. Structure and Position of all Primers used in the Polymerase Chain Reaction 104

3. Sequence Analysis of the SV40 Input Virus 105

4. Sequence Analysis of Cell Line 98, Passage 97 106

5. Sequence Analysis of Cell Line 425, Passage 12 107

6. Sequence Analysis of Cell Line 425 cDNA, Passage 46 108

7. Sequence Analysis of Cell Line 425, Passage 97 109

8. Sequence Analysis of Cell Line 130, Passage 131 110

9. Compilation of Mutations Examined 111

List of Figures

1. Structural Organization of the SV40 Genome 112

2. Functional Domains of the SV40 Large T-antigen 114

3. Transcriptional Control of the SV40 Genome Phosphorylation Cluster Sites of the Large T-antigen 115

4. Stages in the Transformation of Human Epidermal Keratinocytes by the SV40 Virus 117

5. Schematic Line Map of the Genomic Cloning Vector Lambda Gem-11 119

6. Schematic circular map of the sequencing vector M13mp18 DNA. Schematic line map of the sequencing vector pGem-3Z 121

7. Schematic Line Map Indicating the Orientation of the SV40 Genomic DNA Primers used in the Polymerase Chain Reaction 123

8. Schematic Line Map of the Polymerase Chain Reaction/ Sequencing Vector pT7Blue(R). Schematic Line Map of the Plasmid pSPORT 1..... 124

9. Southern Blot Analysis of the DNA From Transformed Human Keratinocyte Cell Lines 22 p.27 and 425 p. 32 126

10. Southern Blot Analysis of the DNA From Transformed Human Keratinocyte Cell Line 98 p. 76 128

11. Preparation and isolation of clones using SV40 transformed cell line 425, passage 66, into the lambda phage cloning vector Lambda Gem-11 130

12. Restriction digest of partially purified Lambda Gem-11 clone (3-1) from transformed human keratinocyte cell line 425 p. 66 132

13. Characterization of the Insert Cloned into Lambda Gem-11 Libraries From Cell Line 98 133

14. Genomic Profile of Different Cell Lines Transformed With SV40, and Characterization of the DNA From Cell Line 22 Cloned into Lambda Gem-11 135

15. Characterization of cDNA Libraries 136

16. Electrophoretic Analysis of PCR Products Used to Prepare Clones in the Vector T7Blue(R) 138

17. Characterization of Polymerase Chain Reaction Products From Cell Line 130 Cloned Into the Vector T7Blue(R) 139

18. Characterization of Polymerase Chain Reaction Products From Cell Line 130 Cloned Into the Vector T7Blue(R) 141

19. Polymerase Chain Reaction Products Using Complete Digestion of Genomic DNA 143

20. Time Dependent Restriction Digest of DNA From Transformed Human Keratinocyte Cell Line 130 145

21. PCR Products Obtained Using Time Dependent Digest of Genomic DNA 146

22. PCR Products Obtained Using Time Dependent Digest of Genomic DNA (Primers 19 and 222) 148

23. Time Dependent Restriction Digest of DNA From Transformed Human Keratinocyte Cell Line 130 Using the Restriction Enzymes Sac I and BamH I 149

24. PCR Products Obtained Using Xba I Digested Genomic DNA From Cell Line 98 151

25. Samples of DNA from Cell Lines 98 and 425 Used During Sequence Analysis 152

26. Immunoprecipitation and Western Blot Analysis of SV40 T-antigen, (p94), and Protein 53, (p53) 153

27. Western Blot of SV40 Transformed Cell Lines at Different Serial Passages 155

28. Compilation of all Mutations Found in the SV40 Stock and Cell Lines Transformed by the Virus 157

29. Immunofluorescent Staining of SV40 T-antigen in a Colony of Viral-infected Human Keratinocytes of Cell Line 98 158

30. Serum Dependent Growth Analysis of Cell Line 425 160

31. Serum Dependent Growth Analysis of Cell Line 130 162

32. Serum Dependent Growth Analysis of Cell Line 98 164

33. Mechanism by which the DNA binding affinity of the mutated large T-antigen product that has undergone a number of serial passages may return to a functional state that is similar to the unmutated protein 166

Introduction:

1- Transformation Properties of Simian Virus 40.

Simian Virus 40 (SV40), is a DNA virus member of the papovavirus group. It was first discovered by Sweet and Hilleman in cultures of rhesus monkey kidney cells which were used in the development of the poliovirus vaccine (1). The virus has a non cytopathic effect on the cells of the rhesus monkey but carries out its lytic cycle in the permissive African green monkey cells. Eddy and coworkers were the first to show that tumor induction in newborn hamsters was due to the oncogenic properties of the SV40 virus (2). Since then SV40 has been regarded as a genetic and biochemical system for the analysis of tumor progression in animals and in vitro studies of cultured cell lines (3,4).

2. Structural Characteristics of the SV40 Viral Genome.

The SV40 viral genome is a single molecule of double stranded, covalently closed circular DNA, 5243 base pairs in length, its

DNA is complexed with histones which gives SV40 DNA a superhelical structure. The chromosome is enveloped in an icosahedral shell which has a diameter of 45 nm. The viral particle has a sedimentation coefficient of 240s and a molecular weight estimated to be at 27 X 10^6 (5-7). The genomic DNA of the SV40 virus is divided into an early region and a late region (figure 1). The early region codes for two nonstructural proteins, the small t-antigen with its mRNA start site at base number 67 to 54 and the large T-antigen with its mRNA start site at base number 67 to 60 and continues after base number 54 up to 17. The late region codes for three structural proteins, VP1, VP2, and VP3, transcribed in the opposite direction to the early genes. These three proteins are all components of the icosahedral capsid. A fourth protein is encoded by the late region, LP-1 or agnoprotein which is responsible for viron assembly (8-11). The agnoprotein is a histone-like, 61 amino acid , highly basic DNA-binding protein, one of its roles is to enhance the efficiency of perinuclear-nuclear localization of VP1 (12-13). This protein may

also have a role in the temporary attenuation of late SV40 mRNA synthesis (14).

3- Structural and Biological Characteristics of the Small t-Antigen.

The small t-antigen is a nuclear and cytoplasmic protein that contains 174 amino acids (15). Differential splicing of the early region of SV40 DNA generates both the small t and large T-antigens. The mRNA that codes for the small t-antigen is the result of a splice between nucleotide position 4571 and 4638 (figure 1). The first 82 amino-terminal residues of these two antigens are the same, the difference lies within the carboxyl-terminal where the small t-antigen has 92 unique residues and the large T-antigen has 626 residues that are unique to its structure (figure 2, table 1) [13].

Although the role of the small t-antigen has not been completely ascertained, studies have shown that under specific conditions it is a vital component in the transformation of certain cell lines (16-17). It appears to be involved in the regulatory mechanism of phosphatase 2A. Small t-antigen seems to have a

growth-stimulating effect in transformed cell lines (18-19). SV40 small t-antigen activates the carboxyl-terminal transforming p53-binding domain of large T-antigen and therefore may play a role in the initiation of transformation (20).

4- Structural and Biological Characteristics of the Large T-Antigen.

The large T-antigen is a nuclear phosphoprotein that is 708 amino acids in length. Large T-antigen has a wide range of functions, including autoregulation, cell immortalization, activation of viral, host gene transcription, and in viral DNA replication. The role of the large T-antigen in both the lytic cycle of permissive cells and the abortive infection of nonpermissive cells has been well documented, (21-24). The lytic cycle takes place in the permissive monkey kidney cells of the African green monkey in a stage specific manner. In the first few hours after infection, mRNA is produced, that codes for the large T-antigen, this protein then regulates the rest of the lytic cycle. During the infection of nonpermissive cells viral DNA is replicated but late

genes are not transcribed. The viral DNA molecule will integrate nonspecifically into various sites of the host genomic DNA. Viral DNA which has now become part of the host DNA produces mRNA which encodes for the large T-antigen. The large T-antigen is then responsible for initiation and maintenance of the transformed state (25). In recent years, the transformation of non-permissive human keratinocyte cells has been studied in detail, and will be addressed later in this report (26-28).

5- Functional Domains of the Large T-Antigen.

Functional domains of the large T-antigen have been established by different experimental procedures including proteolysis of the protein, mutational and deletion analysis, phosphorylation and specific DNA binding experiments (figure 2, table 1) [29-34]. SV40 viral DNA has three binding sites to which the large T-antigen can bind. Binding site I has two recognition pentanucleotides to which large T-antigen binds. The binding of large T to binding site I is part of the system that autoregulates the transcription of early genes. Two penta-nucleotides located

within this region, have the sequence -GAGGC- and are separated from each other by a seven base pair -AT- rich spacer region (35-36). The large T-antigen binds to the pentanucleotide-spacer-pentanucleotide complex and to 5 through 7 base pairs on either side. The AT tract domain is a site for DNA polymerase-α binding which also stimulates the binding of T-antigen to this region (37). The binding of large T-antigen to this site regulates by repressing the transcription of the early genes (38-40). The TATA-box is located between binding sites I and II of the SV40 genome, and is essential for the initiation of transcription (41).

Site II is formed by four -GAGGC- elements which are made up of two pairs in opposite directions in a palindrome of 27 base pairs. The palindromic structure of Site II is part of the minimal viral origin of replication, which contains a 64 base pair regulatory unit and a palindromic element of seventeen -AT- rich base pairs. The three elements contained in Site II are important regulatory elements because they position the large T-antigen in

the proper orientation on the origin for its function in DNA replication (42-45).

The function of binding Site III has not yet been determined, although it has been implicated in the interaction of host proteins with the viral genome. ATP is directly involved in the induction of DNA binding/unwinding functions, as well as a catalytic role during hydrolysis by the ATPase domain of the large T-antigen. The conformational changes of the large T-antigen which are induced by the presence of the molecule of ATP, have a stabilizing effect during the protein-protein interactions between the antigen and the origin of replication (46). Experimental data obtained on the protein-DNA complex at the origin of replication has established a two-lobed structure in which each lobe is made up of a large T-antigen hexamer and the double strands of DNA that makeup the origin (47-49). The double stands of DNA are partially separated at the region of the origin by the action of a DNA topoisomerase, during hydrolysis of ATP by the ATPase activity of the large T-antigen further separation of the two strands is observed. This helicase activity leads to the unwinding

of the double strands. The helicase function of the large T-antigen is unique because unlike other helicases it can unwind double-stranded DNA from within internal sites (50-52). As the DNA becomes unwound, the origin of replication will no longer serve as a DNA binding site. The ATPase activity of the large T-antigen is located at the carboxy-terminal half of the protein between amino acid residues 418 and 627 (53).

Other important functional domain sites include the Zinc finger motif, which is found between the amino acid residues 302-320, and it is responsible for protein-protein interactions, and sequence-specific DNA binding activities (54). The cellular DNA protein DNA polymerase-α primase which is necessary for SV40 replication binds to two sites of the large T-antigen at residues 1-82 and residues 270-517. Binding of the retinoblastoma (Rb), and p53 tumor suppressor proteins occurs at sites 102-115 and 337-627 respectively (55-56).

6- Large T-Antigen in the Viral DNA Replication Cycle.

After infection of the host cell large T-antigen is found mostly in the nucleus where its function is to initiate viral DNA replication and to stimulate the synthesis of cellular DNA. Replication of viral DNA occurs in a sequence specific manner starting with the binding of the double hexamer complex of the large T-antigen to sequences of the viral origin, where the unwinding of the viral DNA begins. This is followed by the binding of the single-strand-specific DNA binding protein (SSB protein), which is necessary to stabilize the DNA in a single stranded state. This procedure is carried out in a bi-directional and semi-discontinuous manner in which the activity of a DNA topoisomerase and the presence of ATP are required (57-58). A tight bond forms between large T-antigen and DNA polymerase-α primase, the total structure is called the initiation complex and large T-antigen assumes its role as a helicase by unwinding DNA strands in a 3' to 5' direction which allows the formation of replication forks (59).

7- Large T-Antigen in the Regulation of Transcription.

One of the functions of the large T-antigen is to regulate transcription of the SV40 early and late genes but it also trans-activates a number of other viral and cellular genes. The SV40 early region is made up of three key regions, which include a TATA box that is responsible for the initiation of transcription, a 21 base pair repeat region rich in -GC- residues, and two identical 72 base pair repeat regions (60-64). The TATA box is located within base pairs 15 and 20, and is about 30 nucleotides from the initiation site of early transcription of viral RNA and acts as a stabilization factor by keeping the initiation of RNA transcription within a specific region (65-66). The action of large T-antigen on transcription is concentration dependent, at lower concentrations the T-antigen stimulates the transcription of early genes as the level of the concentration of T-antigen increases it becomes an inhibitor of early mRNA synthesis by interfering with the binding of the RNA polymerase. Autoregulation of early gene transcription seems to also depend on the interaction of the large T-antigen with the

enhancer-binding protein AP2 which must bind to the promoter/enhancer region so that the early genes can be completely expressed. The large T-antigen inhibits full expression of the early genes by forming a complex between the T-antigen protein and AP2 since it sequesters the AP2 protein thereby preventing binding to the promoter/enhancer region of the SV40 genome (67). There are three homologous 21 base pair repeats, which are located at nucleotide sites 40-60, 62-82 and 83-103, and all have -GC- rich regions, (GGCGGG), that are involved in the regulation of early RNA transcription by acting as binding sites for the general transcription factor SP1, which facilitates the binding of RNA polymerase to the promoter (68) [figure 3a]. The 72 base pair repeats, which are located between nucleotides 107-178 and 179-250, are enhancers essential for the early and late transcription of the SV40 viral genome. The enhancer regions on the 72 base pair repeats stimulate the formation of the transcriptional complex by serving as an entry site for the RNA polymerase and other transcriptional factors (69-71).

Other functions of the large T-antigen include the transcriptional activation of late genes and a role in the trans-activation of long the terminal repeat promoter of the Rous Sarcoma Virus and the promoter of the adenovirus (72-73). The large T-antigen may also regulate transcription by attenuation which is accomplished through the binding of the T-antigen to certain transcription factors.

8- Regulation of Replication and Transcription by Phosphorylation of the Large T-Antigen.

Large T-antigen is a phosphoprotein that carries two phosphorylation sites between residues 106-124 at the amino terminus and 639-701 at the carboxyl terminus (figure 3b). The differential phosphorylation of these two sites serve as a mechanism for the regulation of viral DNA replication and transcription of viral mRNAs. Removal of phosphate groups bound to Ser residues by alkaline phosphatase, which preferentially removes phosphate from Ser but not Thr, serves to increases DNA-binding activities and therefore to increase DNA replication

(33, 74-75). If the large T-antigen is treated with acid phosphatase the phosphate groups are removed from both the Ser and Thr residues resulting in the loss of the DNA-binding activity. If the Thr-124 residue is substituted with another amino acid partial or complete loss of DNA-binding activity results in non-initiation of SV40 replication. These experiments suggest that the phosphorylation of Thr, but not of Ser, has the ability to induce replication by increasing the binding affinity of the large T-antigen to site II in the SV40 origin of replication (76). Several models have been proposed to explain the role of phosphorylation in the regulation of transcription and replication (74, 77). In one such model, the infection of cells is followed by the phosphorylation of the Ser residues on the large T-antigen, which results in the autoregulation of early gene transcription, but does not allow viral DNA replication. The T-antigen can then bind to tumor-suppressor proteins such as p53 and pRb, which leads to progression of the cell cycle from the G1 phase to the S phase. At this point, the active forms of cell cycle-related protein kinase-2, cdc2, and protein phosphatase 2A (PP2A)

phosphorylate Thr-124 and dephosphorylates the Ser residues and the large T-antigen can now activate viral DNA replication (78).

9- Role of the Large T-Antigen in Transformation.

After infection of nonpermissive cells by the SV40 virus the viral DNA can integrate into the genomic DNA of the host, if the integrated DNA contains the early region, large T-antigen will be expressed and transformation of the host cell can occur. Recent evidence has demonstrated that interaction of the large T-antigen with cellular proteins involved in cell cycle control is essential for transformation, (79). In these experiments binding interactions between p53, Rb, p300 and large T-antigen have been studied to determine the role that these interactions in the process of transformation. The p300 protein binds to the amino terminus segment that contains amino acids 1-82. The role that p300 plays in transformation is the least known of the three cellular binding proteins. It shares homology with both Rb and p53. So far it is believed that p300 may play a role in the regulation of cellular DNA synthesis (79). The Rb gene product is

a nuclear phosphoprotein that is necessary for cell cycle regulation. Mutations in the Rb gene play an essential role in transformation and tumorigenesis not only of retinoblastomas, but in a variety of other human tumors. This is consistent with the role of the Rb gene being classified as a tumor suppressor (80). The pRb protein has been shown to bind to the SV40 viral large T-antigen, the human papilloma virus E7 protein, the adenovirus E1A protein and the polyoma T-antigen (81-84). When pRb is expressed the result is growth suppression and a decrease in tumorigenicity of transformed cells (85-86). The mode of action of the pRb protein seems to be due to the phosphorylation and dephosphorylation of the Rb gene product (87-89). The wild type pRb is phosphorylated in a cycle-dependent manner and does not bind to the large T-antigen and this allows the progression of the cell into the S phase of the cell cycle. Dephosphorylated pRb is the active form of the protein which will cause the arrest of the cells before they proceed from the G1 phase to the S phase. The Rb protein forms an inactive complex with the transcription factor called E2F.

This factor is essential for the transcription of cellular genes which drives cells into the S phase of the cell cycle (90-91). The SV40 large T-antigen can displace the transcriptional factor E2F and can therefore stimulate cellular and viral DNA replication by releasing cells from growth arrest. In 1979 the p53 protein was discovered as an oligomeric complex with the oncogene product of a DNA tumor virus (92). The protein was thought to be an oncogene product, because of its participation in the transformation of normal cells to transformed cells in culture (93-95). Later experiments showed that the original isolates were mutated forms of the p53 gene that produced proteins that had tumorigenic activities (96-97). The wild-type p53 acts as a tumor suppresser gene by causing cancer cells to arrest at the G1 phase before the cell is allowed to enter the S phase of the cell cycle. This would be the logical place to repair the cellular DNA since it has been shown that once the cell enters the S phase it is committed to cell division (98-102).

The mutant p53 protein has been found to cooperate with the *ras* oncogene in transforming rat cells in culture, and most recently,

it has been shown to have a cooperative effect with *ras* on the modulation of the activity of the human multidrug resistance (MDR1) gene. MDR1 promotes resistance to cytotoxic drugs in transformed cells (103-104). It binds to the oncogene products of several types of DNA viruses that produce tumors in animals, including the large T-antigen of SV40, E6 early proteins of the human papilloma virus type 16 and 18, Ad5 and E1B of the adenovirus, EBNA-5 of the Epstein-Barr virus and the 90 Kd phosphoprotein produced by the murine double minute 2 gene. Wild-type p53 has also been shown to bind to regions near the TATA box as a member of the TATA transcriptional complex where it functions to either activate or repress transcription (105-107). Another set of experiments have established that wild-type p53 protein is a possible transcriptional activator due to specific binding at promoter and enhancer sequences in vitro (108-112). Studies of human cancer cells derived from tumors of the colon, lung, ovary, brain, breast, bladder and other organs have shown that mutated forms of the p53 gene product are present in over 50% of the cases, and therefore the interaction

between the large T-antigen and that of p53 is of great interest (113-115). The large T-antigen and p53 are both phosphoproteins that must be phosphorylated so that a complex between these two proteins can occur (116). Wild-type p53 can bind to the large T-antigen at a site where DNA polymerase-a primase and ATP also bind (117-121). The site to which wild-type p53 binds the large T-antigen has consensus sequences (GGGCGG) which are also present in both the 21 and 72-base pair enhancer-promoter regions of the virus (122). These findings suggest that the binding of wild type p53 to the origin of replication of the SV40 virus inhibits the functions of the large T-antigen during DNA replication. When a complex is formed between the T-antigen and p53, the latter can no longer regulate cycle cell activity, which leads to the initiation of replication (123-125).

10- Transformation of Human Keratinocytes by the SV40 Virus.

In 1962 it was reported that SV40 could transform human fibroblast cells in vitro (126). The fibroblast system, although limited because of the number of transformation markers which are available, was instrumental in answering questions that related to growth regulation after viral infection. Due to the limitations of the fibroblast cell lines and the fact that more than 80% of all tumors occur in epithelial cells, a system that could study the interactions of viral genes with that of the cellular DNA of epithelial cells was sought. Our laboratory was the first to report the transformation and characterization of human epidermal keratinocytes by the SV40 virus (26). Terminally differentiated keratinocytes cells have the advantage over other systems in that the events occur according to a sequence of well defined periods of precrisis, crisis and postcrisis which demonstrate certain cytologic and histochemical features that have been addressed before (27,127). During the early stage of viral infection known as precrisis growth independence from

high serum concentrations can be detected within the first four passages, (about forty cell generations). Also the loss of cornified cell envelope and histochemical markers of differentiation are observed during this period. Clonal growth in the absence of feeder layers and the redistribution of fibronectin cables on the cell surface also occurs. Observation of infected cells by means of the microinjection of fluorescent dyes confirms that junctional communication between these cells has been altered (128-130). During this stage of transformation, the T-antigen acts as an effector molecule by providing certain growth advantages to the transformed cells, which are not seen in the nontransformed cells (3). As SV40 large T-antigen positive cells increase in number, there is also an increase in the growth potential of the keratinocyte cells. This growth increase continues until after 10-15 serial passages, after which the cells enter a period of crisis (figure 4) [26, 27, 127-130].

During <u>crisis</u> a marked decrease in growth potential is observed and the cells exhibit a cytopathic effect (CPE) that includes the appearance of large vacuolated cells, which become

detached from the rest of the cell culture. The period of crisis has been studied in a number of different cell lines and the severity of crisis varies from one cell line to another; in the worst case scenario, the cell culture will completely be destroyed within a ten day period after the first, signs of CPE. The survival rate of different strains of twelve independent cell lines that undergo crisis and survive past the 15th serial passage is roughly two thirds. The surviving cells exhibit an indefinite growth potential and therefore become "immortalized" cell lines (127-128). The period of crisis is an important stage between the early and late phases of the transformation process.

The cells now enter the stage of postcrisis, in which cells acquire the capacity to form colonies in agar and manifest anchorage independent growth. Markers observed during this period include altered actin and intermediate filament cytoskeletal elements and the formation of undifferentiated neoplastic tumors in nude mice (131). The SV40 viral infection of keratinocytes results in the alteration of the pattern of the keratin intermediate filament gene product which demonstrates

that the virus attacks certain phenotypic genes (132-134). The SV40 virus produces changes in the pattern of transcription of the keratin genes which include the loss of expression of stratified keratins and an induction of expression of genes coding for simple/fetal keratins. These molecular markers of transformation are also found in cells derived from carcinoma tumors, which is significant because the expression of epithelial fetal keratins in malignant cells is used in cancer diagnosis (135-139).

Our laboratory has observed the sequential transformation of epithelial cells by the SV40 virus over an extended period of time. During the early period post infection, most of the viral DNA is found to be present as either full length free supercoiled or truncated subgenomic fragments in high copy number. As the infected cells are carried through higher serial passage in culture, there is a change in the pattern of viral sequences such that only integrated fragments of viral DNA in low copy number are retained. Since the cells exhibit the ability to undergo constant changes while going through serial passage in culture, the mutations acquired after the advent of infection are the

result of a selection process in which mutations that confer growth advantages to the cells are selected. This lends support to the theory that only essential viral genes that have experienced selective mutagenesis are necessary for the transformation and maintenance of the "immortal" state of these cell lines. Based on this series of events it would be worthwhile to study the structural features of the integrated viral DNA and elucidate if these alterations may play a role in the transformed state of these cells. To study the integrated SV40 viral early genes we have chosen a series of established high passage human keratinocyte cell lines. Alterations in the viral sequences which selectively exhibit retention of mutations acquired during the transformation event are the subject of this dissertation.

Materials and Methods:

1. Cell Culture of SV40 Infected Human Keratinocytes:

Cells from newborn human foreskin specimens in suspension were disaggregated in trypsin and then infected by the virus at high multiplicity. The cells were then passaged at a split ratio of 1:3 in Dulbecco's minimal essential medium (DMEM), supplemented with 10% fetal calf serum, hydrocortisone (0.5μg/mL), cholera toxin (50μg/mL), and epidermal growth factor, (30 μg/mL). The established cell lines used in the following experiments were 98, 425, 130 and 22, (26, 27, 135).

2. Extraction of the Genomic DNA:

The extraction of 10 x 100 mm plates containing the transformed cell lines was initiated by pouring off the growth medium and rinsing the cells with 1 X PBS twice, followed by the addition of 0.5 mg protinase K and 5 mL of extraction buffer (100 mM Tris-HCl, pH 7.9; 10 mM Ethylene Diamine Tetra

Acetate { EDTA }, 10 mM NaCl, and 0.5% Sodium Dodecyl Sulfate { SDS }). This mixture was allowed to incubate at 37°C overnight. The plates were scraped gently and the DNA mixture was placed in 50 mL centrifuge tubes. Equal volumes of phenol/chloroform-isoamly alcohol, (24:1), in a ratio of 1:1 was then added to each tube. The 50 mL tubes were then placed on a gentle mixer for 10 minutes. This was then centrifuged at 3 K rpm for 10 minutes. The supernatant was saved and care was taken not to suction the white interface. The last steps starting with the phenol/chloroform-alcohol extraction were repeated. To the supernatant two volumes of 95% ethanol were added and after gently agitation, it was placed in the -70°C freezer for 30 minutes. This was then spun down for 10 minutes at 3 K rpm. The ethanol was poured off. The same amount of 70% ethanol was added to this, and the procedure to recover the DNA was repeated. The DNA was then dried for 15 minutes under vacuum. To each tube were added 2 mL of 10 mM tris, pH 7.4 and 0.5 mM in EDTA and the tubes allowed to incubate at 37°C overnight. RNase A was

added to a final concentration of 10 μg/mL using a stock solution of 10 mg/mL. The phenol/chloroform extract and the ethanol precipitation was then repeated twice. Addition of 0.5 mL of 10 mM tris, pH 7.4 and 0.5 mM EDTA followed. This was allowed to incubate at 37°C overnight. The mix was then dialyzed for three days against 10 mM tris, pH 7.4 and 0.5 mM EDTA. The purity of the DNA was checked at 260 and 280 Lambda. The 280 reading should be 0.5 that of the 260 reading.

3. Digestion of the Genomic DNA:

The genomic DNA was cut with different restriction endonucleases including Xba I and others, which are no-cut enzymes for the virus SV40. A typical restriction digest set up took anywhere from 2 to 8 hours and included 100 μL of DNA, 3 μL of the restriction endonuclease, 1μL of BSA, 12 μL of 10 X buffer and 5 μL of H_2O. Once finished the digestion was extracted with phenol/chloroform followed extraction with chloroform, precipitation with 100% ethanol, washed with 70%

ethanol and resuspended in TE, (10 mM Tris-Cl, pH 8.0 and 1mM EDTA, pH 8.0). This procedure was checked by electrophoresis. The agarose gel was prepared by using 0.5 g Agarose, 10 μL Ethidium Bromide and 1 mL 50 X TAE buffer (242 g of Tris base, 57.1 mL of glacial acetic acid and 100 mL of 0.5 EDTA, pH 8.0).

4. Bacterial Preparation:

The first step was the preparation of the two host bacteria NM538 and NM539. This was done by incubating the host bacteria in sterile NZYM medium for 24 hours (Per liter: 10 g NZ amine, 5 g NaCl, 5 g yeast extract and 2 g $MgSO_4 \cdot 7H_2O$ pH to 7.5). The bacteria was recovered by centrifugation and then resuspended in sterile 10 mM $MgSO_4$, (0.4 X the volume of the original culture) and stored at 4 °C for up to 14 days.

5. Bacteriophage Isolation:

The phage Lambda GEM-11 (figure 5), was incubated in 5 mL of NM538 bacterial host for 10 minutes at 37°C. The lysate was then added to fresh NZYM medium and placed in a shaker overnight.

This mix was then poured into 500 mL of sterile NZYM medium and allowed to incubate at 37 °C overnight. To the lysed bacterial suspension were added 10 mL of chloroform and allowed to stand for 10 minutes at room temperature. This mix was then spun down at 8000 X g and 4 °C for 10 minutes. The decanted supernatant was filtered by using a flask topped with a cheese cloth filter. The pellet and chloroform were then discarded. This was followed by the addition of both RNAse and DNAse to a final concentration of 10 mg/mL and this mix was placed at 37 °C for 1 hour. An addition of 29.2 g of NaCl and 50g of polyethylene glycol to the supernatant followed. This was allowed to dissolve at room temperature and was then placed in ice for 1 hour. The solution was then centrifuged for 10 minutes at 8000 X g and 4 °C. After which the pellet was recovered and resuspended in 8 mL of SM (Per liter: 5.8 g NaCl, 2 g $MgSO_4 \cdot 7H_2O$, 50 mL 1M Tris-Cl [pH 7.5] and 5 mL 2% Gelatin), using 6 mL of SM to resuspend the pellet and 2 mL to rinse. An equal volume of chloroform is then added to the resuspended pellet, vortexed for 30 seconds and spun down by centrifugation at 80 RPM

for 15 minutes. The aqueous phase was then measured accurately and a volume of CsCl corresponding to the following formula was used:

$$\frac{\text{Final Volume X 67}}{82} = \text{grams of CsCl to be added.}$$

The CsCl added was then dissolved completely and the resulting mix was centrifuged at 40,000 RPM for 24 hours at 22 to 28 °C by using the TST60.4 rotor. The following day, the small plastic tubes which contained the centrifuged mix were removed from the rotor with a set of special forceps. When placed under a scattering light, they exhibited small light blue circular bands. Tubes were punctured with a 20 gauge needle and the bands were collected. These collected bands were pooled together and placed in dialysis bags to be dialyzed against 1000 mL of TE, (Per liter: 0.5844 g 10 mM NaCl, 5 mL of 2 M Tris-Cl pH 8.0, 2.0331 g 10 mM $MgCl_2$). The phage preparation was dialyzed for 1 hour and then transferred into 50 mL plastic tubes into which EDTA added

was to a final concentration of 20 mM, (For example to a 20 μl sample you would add 8 μl from a stock solution of 0.5 M, pH 8.0 of EDTA). Proteinase K (20 mg/mL), was then added to a final concentration of 50 μg/mL. To this mixture was then added sodium dodecyl sulfate, SDS, to a final concentration of 5% and each tube was mixed several times by inversion. The sample was incubated for one hour at 68 °C followed by extraction with an equal volume of phenol, centrifuged and the aqueous phase recovered. It was then extracted with an equal volume of chloroform twice. The corresponding aqueous phase was pooled and dialyzed against 1 liter of TE buffer at 4 °C overnight. Dialysis was then repeated in the morning with a fresh change of TE buffer for 1 hour. A 5 μL sample of the dialyzed solution was checked for concentration and purity of the phage by the use of gel electrophoresis. The isolated phage was then digested with the appropriate enzyme.

6. Ligation Procedure:

The ligation procedure was carried out by using T4 ligase. Amounts of each reactant used were as follows:

Lambda Gem vector	3 μL
Insert	1 μL
T4 ligase	2 μL
10 X ligation buffer	2 μL
H_2O	12 μL

Total volume was 20 μL and placed in an eppindorf tube. To test that the vector has not reannealed, a blank was prepared that did not include the insert but did have a total of 13 μL of H_2O. This was placed in a 15 °C water bath for 24 hours. Once the ligation mix was removed it was tested by gel electrophoresis. After the ligation procedure was termed positive the DNA was recovered by using phenol/chloroform extraction. 4 μL of the ligation mix was then used for packaging, followed by the addition of 30 μL of chloroform and 400 μL of sterile SM.

7. Screening Procedure:

First the vector carrying the insert was used to infect the bacterial host NM539, (spi^+), bacterial cell line by placing 200 μL of bacteria in a 15 mL test tube in the presence of 1 μL, 5 μL, and 0.5 μL of packaged ligation this was allowed to incubate for 15 minutes at 37°C. The incubation mix was then plated by using 3 mL of top agar which was kept at 42 °C so that it would remain in its liquid form. It was then poured into the test tube that contained the infected NM 539. This mix was spread evenly on the bottom agar plate and was allowed to harden at room temperature. The plates were then inverted and placed in a 37 °C incubator overnight. As a control 200 μL of the bacteria which was not infected was used. To prepare the top and bottom agar 0.75 and 1.5% of bacto-agar In NZYM medium was used respectively. The plates were then checked for plaque frequency. Depending on this frequency large plates were prepared with an amount of packaged material not to exceed 12,000 plaques. To transfer the clear plaques which were produced overnight the following procedure was used. The plates

were placed on a flat surface. A sterile nitrocellulose filter marked with a pencil was placed on top of the agar plate. The filter's position was marked by making asymmetric holes using a 18 gauge needle, marked with a permanent pen top and bottom plates. After the filter was completely wet it was peeled off by using a pair of forceps. A piece of Whatman 3 MM paper was cut so that the nitrocellulose filters would fit and have about one inch between each filter. The peeled filter was then transferred to the 3 MM paper which was saturated with denaturing solution, (0.5 M NaOH and 1.5 M NaCl), so that the filter is completely wet with the solution. The filter is then left for 5 minutes in the denaturalization solution. The filter was then transferred to a second 3 MM sheet of paper that had been saturated with neutralizing solution, (1.5 M NaCl and 0.5 M Tris.Cl pH 8.0). This was allowed to stand for 5 minutes transferred to 3 MM paper and dry at room temperature for 30 t0 60 minutes. After the drying period another two pieces of 3 MM Whatman paper was used to sandwich the filter. This was then baked for 2 hours at 80 °C in a vacuum oven.

8. Hybridization Procedure:

The SV40 DNA probe was prepared with 3 μL of ^{32}P-labeled ATP, GTP, TTP and CTP which were dried under vacuum for 30 minutes. Followed by the addition of 1 μg of SV40 DNA, 2 μg of nick translation buffer and 13 μL of H_2O. This mixture was then incubated for 10 minutes at 37°C followed by the addition of 2 μL of DNA polymerase and 1 μL of a 10^{-4} dilution of DNAse I. The probe was then allowed to incubate in a 15°C water bath for a period of one hour after which 10 μL of EDTA where added and the mix was placed in a 68°C water bath for 15 minutes followed by the addition of 400 μL of TE buffer were added. The nitrocellulose filters were then placed in a prehybridization solution which includes the following reagents:

Formamide	5 mL
2% polyvinypyrrolidone, (PVP).	0.5 mL
2% ficol	0.5 mL
5 X SSPE from 20 x solution	2.5 mL
2% Bovine Serum Albumin, (BSA).	0.5 mL
100 μg/mL non-specific DNA	0.2 mL
H_2O	0.8 mL

This was placed at 42 °C for 1 hour. The filters were then placed into a hybridization solution overnight, this solution is prepared by using;

Formamide	5 mL
2% PVP	0.1 mL
2% Ficol	0.1 mL
2% BSA	0.1 mL
non-specific DNA	0.2 mL
5 X SSPE	2.5 mL
SV40 probe	0.1 mL
H_2O	1.9 mL

After this procedure the filters were washed using 0.1% sodium dodecyl sulfate and 0.1 X SSC. This was done for 20 minutes at 50 °C and repeated a second time. Then the filters were washed in 0.1 X SSC for another 20 minutes at 50°C. SSC (0.1) was prepared from a stock solution of 20 X SSC which contains 178.3 g of NaCl and 88.2 g of sodium citrate by dissolving this in 800 mL of H_2O. This is adjusted to a pH of 7.0 with a 10 N solution of NaOH and taken to a volume of 1000 mL. The filters are then ready for autoradiography. The filters are placed on 3 mm paper and covered with cellophane paper and placed in a film case for 24 hrs. at -70°C. The films are then

exposed and the positive signals are marked on the original plates. The positive plaques were recovered in a sterile environment. The plaques which are still on a peace of agar are placed in 1.5 mL tubes and to this is added 500 μL of SM. These were then screened at least 3 times to purify the clone.

9. Sequencing Vectors:

After the insert was cut out of the amplification vector it was transformed into one of the following the sequencing vectors, Mp18, Mp19 or pGem 3Z, (figure 6a and b). The preparation of the single stranded template DNA was carried out as follows. JM101 was grown overnight in LB medium. The overnight culture was then diluted in a 1:50 ratio. One plaque of the DNA to be sequenced was used and placed in 2 mL of culture. This mixture was then incubated for 6 hrs. in the 37°C shaker, placed in a microcentrifuge tube and a pellet was recovered by centrifugation. To 1 mL of the supernatant 0.25 mL of a 20% PEG-6000/3.5 M ammonium acetate solution was added and allowed to stand on ice for one hour. The mix was centrifuged and a grayish-

white pellet was recovered. The phage pellet was then resuspended in 100 μL of TE buffer. The DNA was then extracted with phenol/chloroform and the aqueous layer was transferred to a clean tube, the phenol/chloroform layer was then back extracted with 100 μL of TE buffer. The extraction with phenol/chloroform was repeated until no interface was visible. Ethanol precipitate the DNA using sodium acetate, (3 M before you start the precipitation), at -70°C for 30 minutes. Dried in vacuum for 30 minutes and resuspend in 10-20μL of de-ionized H_2O.

10- Polymerase Chain Reaction:

The polymerase chain reaction, (PCR), was carried out on different types of DNA samples, including cut and uncut genomic DNA infected with the SV40 virus, established cell lines cloned into phage Lambda GEM-11, and the DNA introduced into the sequencing vector pGEM-3Z. A typical polymerase chain reaction included, 5 μL of template DNA, (approximately 1 pmol), 1 μL each of the two primers selected, (figure 7, table 2), 5 μL of the reaction buffer, 2 μL of the

dNTP's, and 35 μL of sterile H_2O. The reaction setup included a 95 °C denaturation period followed by the addition of 1 μL of Tag DNA polymerase, (3 units/μL), the reaction was then allowed to proceed through 38 cycles, the conditions of which included a period of denaturation at 94°C, an annealing period at 50°C and an extension/termination period at 70 °C. This PCR protocol was used for the amplification of selected DNA sites, both in the integrated viral DNA and that of the flanking human sequences.

11- PCR Cloning Vector:

The vector used for the amplification and sequencing of the products of the polymerase chain reaction was the T-vector pT7Blue(r) (figure 8a). This vector simplifies the cloning of PCR products which typically leave single 3' A-nucleotide overhangs on their reaction products, since the vector provides a compatible single T-nucleotide overhang. The PCR products can be directly cloned into the vector without purification by extracting a determined amount of the total PCR product with one volume of chloroform/ isoamyl alcohol, (24:1), followed by vortexing for

one minute and centrifugation at 12,000 X g for another minute. An alternative method of cloning purified PCR products is by cutting a desired band out of an electrophoretic low melting point gel, followed by either extraction by phenol/chloroform or by using the GENECLEAN kit, (Bio 101 Inc). The phenol/chloroform purification is done by cutting out the desired band from the electrophoretic gel, placed in an eppindorf tube at 68°C for 5 minutes, followed by extraction of the DNA with an equal volume of phenol, then an equal volume of phenol/chloroform, (1:1), extraction with one volume of chloroform. The aqueous phase is then precipitated by the use of ethanol. GENECLEAN uses a silica matrix to selectively bind the DNA which is then eluted using a polar solvent such as sterile water. The PCR product is ligated to the vector using established procedure previously explained in this document. Competent cells are then transformed with the ligated vector.

12- Transformation of Competent Cells:

Two types of competent cells were used in these experiments. The first were produced using the E. *coli* strain JM101, which was grown on minimal media agar plates, to ensure that the F episome is not lost. The F episome is required for the cells to grow on minimal media, since the proline region is deleted on the chromosome and is only present on the F episome. One colony from the minimal medium agar plates is placed into 3 mL of LB media and incubated on a 37°C shaker overnight. The overnight culture is diluted to a factor of 1/200 into 50 mL of LB, and grown on a shaker at 37°C for a period of 2-3 hours so that an OD_{590} reading is 0.4 t0 0.6. The cells are then placed on ice and all of the materials used should be pre-chilled. The cells are centrifuged at 7000 X g for 5 minutes, resuspended in 1/5 volume of ice cold 0.1M $MgCl_2$ and place on ice for 20 minutes, followed by centrifugation as described above, and resuspended in ice cold 0.1 M $CaCl_2$ at 1/50 of the original culture volume. The plating culture is prepared by using a single colony

of the JM101 from the minimal agar plates by incubating in LB for a period of about 4-6 hours until a OD_{550} reading of 1.0 is obtained. 100 μL of competent cells are inoculated with 3 μL of the vector, (10 ng of vector), incubated on ice for a period of 40 minutes with periodic mixing of the solution, placed at 42°C for two minutes, then plated using 200 μL of plating culture and 3 mL of molten top agar. The plates are incubated overnight at 37°C and screened for positive plaques by hybridization techniques described before. The second procedure makes use of competent cells provided with the pT7Blue(r) vector kit. In this procedure 20 μL of the competent cells are inoculated with 1 μL of the ligation mix, placed on ice for 30 minutes, heat treated at 42 °C for a period of 40 seconds, placed on ice for 2 minutes, followed by the addition of 80 μL of SOC medium at room temperature and shacked at 37 °C for 1 hour. 50 μL of the transformed cells are then spread on LB plates that contain 15 μg/mL of tetracycline which insures the selectable F' containing lacZΔM15 is maintained, with the elimination of undesired

background of non-recombinants. This is followed by the selection of positive colonies using hybridization techniques.

13- cDNA Libraries:

cDNA libraries containing SV40 transformed genomic DNA from the immortalized cell line 425 passage 46 were provided by Dr. Gruber, (laboratories of Life Technologies). These cDNA libraries were cloned directionally into the vector pSPORT at the Not I, Sau I site, and grown on LB plates, the resulting colonies were screened using the hybridization techniques described above, (figure 8b).

14- Sequencing Procedure:

Three different sequencing procedures were used during these experiments. The first was done manually using the Sequenase DNA sequencing system. For each set of four sequencing lanes, a single annealing reaction is used that contains 1 μL of the desired primer, 2 μL of the reaction buffer and 7 μL of template DNA, (1 μg for M13 or 3-5 μg of plasmid DNA). The mix is heated at 65 °C for 2 minutes and allowed to cool at room temperature for a

period of 30 minutes so that the proper annealing is achieved. The labeling reaction is prepared by the addition of 1 μL of 0.1 M DTT, 2 μL of labeling mix, 0.5 μL of [α-^{35}S]dATP and 2 μL of diluted Sequenase enzyme, (dilution is 1:8).

The solution is mixed and incubated for 5 minutes at room temperature followed by the termination reaction in which 2.5 μL of ddATP, ddCTP, ddGTP and ddTTP are placed into four different tubes. The tubes are then heated for 1 minute at 37 °C to each a 3.5 μL aliquot of labeling incubation is added and returned to the 37 °C for a period of 5 minutes, followed by the addition of 4 μL of stop solution. A 6% acrylamide gel is prepared by using 5.7 g of electrophoresis grade acrylamide, 0.3 g of electrophoresis grade bis-acrylamide, 48 g of urea, 10 mL of 10 X TBE buffer, 40 mL of de-ionized H_2O, 500 μL of 10% ammonium persulfate and 50 μL of TEMED. The gel is poured and allowed to polymerize at room temperature for at least 1 hour at which time the gel is run at constant power, between 50-90 watts, and a voltage of 1500-2000 watts, using a 1X TBE buffer solution, until the temperature

of the gel is constant, (temperature of the glass should not exceed 60°C). The samples are then heated between 75-80°C at which time 2.5 μL were loaded onto the gel. The gel is then run for a period of 4-6 hours, fixed in a 10% acetic acid, 10% methanol solution for ten minutes, placed on a sheet of Whatman 3 MM paper and placed on a gel dryer at 80 °C for 30-60 minutes. The gel is exposed on film for a period of 48 hours. The second method makes use of Tag DNA cycle sequencing. In this procedure both DNA cloned into sequencing vectors and direct PCR products were sequenced manually. The last method of sequencing has made use of the automated system at the RCMI facility at CCNY.

15- Southern Blot Analysis:

For southern blot analysis a sample of 5-10 μg of DNA were loaded into a 1% agarose electrophoretic gel and transferred to a nitrocellulose filter using passive blotting techniques. The filter was then hybridized to full length SV40 DNA using the procedure already described above. The filter was then exposed to film for a

period of 24 hours or placed on a phosphoimager for a shorter exposure time, (0.5-1 hour).

16- Immunoprecipitation and Western Blot Analysis:

Western blot analysis of T-antigen proteins produced from transformed keratinocyte cell lines numbers 22 passage 113, 425 passage 68, 130 passage 116 , and AG34 an agar-selected clone from cell line 98. Cells were passaged in 6 cm plates until confluent, removed using a solution of trypsin-EDTA, neutralized using 2 mL of DMEM, and centrifuged for 10 minutes at 3,000 rpm's. The resulting pellet was resuspended in 1.0 mL of 1 X PBS, centrifuged, frozen and thawed at 37°C. The pellet was dissolved in 50 μL SDS gel-loading buffer, boiled for 5 min, and placed on ice for 15 minutes. The sample was then centrifuged at 10,000 g for 2 minutes and run on a SDS-stacking/running gel combination. The sample was centrifuged for 2 minutes and run on a SDS-stacking/running gel combination. The stacking gel was prepared by using:

Acrylamide Stock Solution	500 μL
1 M Tris-HCl, pH = 6.8	515 μL
20% SDS	25 μL
10% $(NH_4)_2SO_4$	50 μL
TEMED	10 μL
H_2O	3.85 mL

The running gel was prepared using the following:

Acrylamide Stock Solution	5.0 mL
1 M Tris-HCl, pH = 8.8	5.6 mL
20% SDS	75 μL
10% $(NH_4)_2SO_4$	100 μL
TEMED	20 μL
H_2O	4.35 mL

The proteins were then transferred from the polyacrylamide gel onto a nitrocellulose filter, which was then stained with Ponceau S, at this time the molecular marker was cut out. The nitrocellulose was immunoblotted using the anti-phosphotyrosine system. The primary antibodies used were the large T-antigen SV40 antibody, (Ab-2), [CALBIOCHEM], that is specific to protein products of this antigen and small t-antigen SV40 antibody (Ab-3). The secondary antibody used was an anti-mouse-alkaline phosphatase, (AP), conjugate, this secondary antibody catalyzes a colorimetric reaction when the appropriate substrate is used,

(BCIP-colorimetric substrate, Promega Corp.), the sites of antigen localization turn dark purple. Immunoprecipitation of the large T-antigen was performed on cell lines 98 passage 103, 130 passage 140 and 425 passage 93. Cells were treated as described above, after the freeze/thawing procedure the cells were dissolved in ice cold 1 mL of PBSTDS solution, (100 mL 10 X PBS, 10 mL of 100% TRITON X-100, 5 g of deoxycholate, 0.5 μg/mL of leupeptin, 1 mM EDTA, 1μg/mL pepstatin and 0.2 mM PMSF. Dilute to 1 L), and incubated at 4°C for 10 minutes. The cellular debris was removed by centrifugation at 4,000 rpm for 2 minutes at room temperature. The lysate was cleared by incubating for one hour at 4°C with 1 μg of normal mouse IgG and 20 μL of agarose conjugant, followed by centrifugation at 1,500 rpm to pellet out the agarose. The large T-antigen SV40 primary antibody, (Ab-1) [CALBIOCHEM], is then added to the cell lysate, followed by the addition of 15 μL of protein A-agarose, at which time the mixture is incubated on a shaker, at 4° C for a period of 2-24 hrs. The immunoprecipitate is collected by centrifugation at 2500 rpm for 15 minutes, washed in PBSTDS, and analyzed by the western

blot procedure described above. Two immunoprecipitations were carried out one using large T-antigen antibody (Ab-2) and the other using p53 monoclonal antibody.

17- Immunofluorescent Staining of SV40 Large T-antigen:

Cover slip cultures of human keratinocytes were rinsed in phosphate buffered saline (PBS), fixed in cold acetone-methanol (70:30). Air dried and incubated at 37°C with hamster antiserum to SV40 T-antigen which was provided by Dr. Robert Carroll (NYU Medical Center). The cover slips were then washed three times with PBS and incubated for 1 hour with goat anti-hamster IgG fluorescent conjugate. The cells were then rinsed three times with PBS, mounted in glycerol-PBS and view under fluorescent microscopy.

18- Serum Dependent Analysis:

Cells from lines 98 passage 98, 130 passage 131 and 425 passage 93 at an initial concentration of 1.25×10^5 cells/plate were

plated into 6 cm plates. Each cell line was analyzed in duplicates at fetal calf serum concentrations of 0.1, 0.5, 2, 5, and 10% in DMEM supplemented with hydrocortisone, (0.5 μg/mL), and antibiotics, (100 X). Cells were counted for a period of seven days. The data was then graphed as (cells/plate) versus days.

19- Colony Formation of SV40-Infected Keratinocytes in Soft Agar:

Cells from lines 98 passage 105, 130 passage 145, and 425 passage 99 were harvested from monolayer cultures and seeded into soft agar plates, 6 cm, in duplicate at the cell density of 10^6. The base layer stock solution was prepared by mixing 40 mL of 1.25% Difco Bacto agar, 40 mL of 2 X supplemented DMEM, 10 mL of Difco Bacto tryptose phosphate broth and 10 mL of fetal calf serum. Five milliliters of the base layer was pipetted into each plate and allowed to set at room temperature. The top layer contained one volume of the cell suspension medium mixed with two volumes of 0.5% agar (140). Colonies were counted after 4

days and colony formation efficiency, (CFE), were calculated as described by Defendi and Steinberg [141].

20- Computer Analysis of DNA Sequences:

The computer program LFASTA and PCGENE were used to align the SV40 input sequences to that of established clones to search for mutations that may arise during the progression of the transformed state of the infected keratinocyte cell lines. PCGENE has also been applied to nucleotide analysis. The program GENEBANK was used to compare catalogued SV40 sequences to that of the SV40 input virus.

Results:

A- CHARACTERIZATION OF GENOMIC LIBRARIES.

1. Analysis of Genomic DNA.

Figure 9 shows a Southern blot analysis of the viral sequences in SV40- transformed lines 22 and 425 at the 27th and 32nd passages respectively. At the 27th serial passage line 22 appears to display mostly unintegrated ('free') viral DNA as indicated by the presence of two bands electrophoretically separated from undigested cellular DNA and migrating with the apparent sizes corresponding to the supercoiled (form I) and nicked circular form (form II) of the free viral DNA. However, while digestion with enzymes having single restriction sites on SV40 did produce bands migrating at ~5.2 kb as expected for unintegrated viral DNA, minor subgenomic bands of about 4 and 4.3 kb for BamHI and EcoRI respectively were also seen. In contrast, no free viral DNA forms were seen in undigested line 425 DNA run on the same blot. EcoRI

digestion of line 425 DNA produced two viral containing bands of about 2.4 and 3.0 kb instead of the single 5.2 kb band which would be expected from native viral DNA. In addition Bam HI digests of the same DNA produced at least four bands of aberrant sizes including three which were supragenomic in size, consistent with the presence of integrated viral sequences. The results shown on the left were produced using a 96 hour exposure, the results shown on the right were obtained from the same blot using a 24 hour exposure. All of the experiments that involved DNA analysis of infected human keratinoctes had first been compared to human keratinocyte cells that were not infected. This procedure served as a control against any background that may occur due to cellular DNA. The southern blot analysis clearly shows that unintegrated SV40 virus in these low passage cells is maintained in episomal form at high copy number, followed by random integration shortly thereafter (142). As the infected cells are carried through higher serial passage in culture, there is a change in the pattern of viral sequences such that only integrated fragments of viral DNA in low copy number are retained, as shown in experimental data obtained

from the high passage cell line 98, passage number 76 (figure 10). Densitometric analysis done on this cell line and partial sequencing of human DNA flanking the integrated virus demonstrate that the viral sequences are present in low copy number, about 1-3 genome equivalents/cell (143). Southern blot analysis of this transformed cell line substantiate these results.

2. Genomic Libraries in Lambda Gem-11.

The Lambda Gem-11 phage cloning vector, was selected because of its high cloning capacity (0 to 14 Kb), [144-145]. The vector was digested either with Xba I or Sac I and ligated to the genomic DNA of different transformed cell lines using T4 DNA ligase. Figure 11 (A) shows the ligation profile of transformed cell line 425, passage 66. The Lambda Gem-11 vector was cut either with the restriction enzyme Xba I or Sac I and religated to itself to make sure that the isolated vector was pure. Genomic DNA was digested with the same enzymes and also religated. The ligation of vector and insert was then packaged using packaging extracts, (Stratagene), plated onto the *E. coli* host NM539 and

screened for positive plaques using ^{32}P-labeled full length SV40 genomic DNA as a probe, (figure 11 B). The primary screening yielded five positive recombinants out of a total of 2×10^5 screened. Further rounds of clone purification produced two independent isolates, 3-1 and 5-1,(figure11C). DNA from the partially pure isolates as digested using the restriction enzymes, Avr II, (two sites on SV40), BamH I, (one site on SV40), EcoR I, (one site on SV40), and Xba I, (no sites on SV40), to establish the size of the genomic insert, (figure 12A and B). Southern blot analysis of the samples in (A) were conducted using ^{32}P-labeled full length SV40 genomic DNA as a probe. The uncut 3-1 clone shows hybridization of DNA that is still trapped at the loading point on the agarose gel. This is most likely due to protein contaminants. Protein-DNA interactions of this type often lead to the incomplete digestion of the clone and may account for the banding pattern observed during this analyses. Another possibility is that a portion of the polylinker has been lost and viral sequences may still be attached to one of the arms of the vector. This would explain why the Avr II digest produces a band that is

~1 kb in size which has viral sequences. The analysis of the BamH I digest is more complex. There is hybridization to a fragment of about ~5 kb and to a band that migrates slightly lower than the band that hybridized in all of the samples, this band may include human sequences. The BamH I digest was used to establish clones prepared by using the sequencing vector Mp18 and is currently under study. Agarose gel electrophoresis restriction digests on the repurified Lambda Gem-11 clone, 3-1 indicate that there is 5 kb fragment that contains viral sequences, (figure 12 C). Other clones were established using the genomic DNA from the cell lines 98, 22 and most recently from cell line 130, (figures 13 A-D). Figure 13 (A) shows different clones from cell line 98. The gel was analyzed using the southern blot techniques. The results demonstrate that the size of the insert is approximately 5.2 Kb, figure 13 (b). Figure 13 (c-d), are southern blot analyses of clones prepared in cell lines 98 and 22. These clones vary in size and in intensity, the size difference can be attributed to human sequences that flank the viral genome since the restriction enzyme used to obtain these clones was Xba I, a no-cut enzyme for

SV40. The difference in intensity is due to the purification technique which results in different concentrations of product. Figure 14 (A-B) is a genomic profile of different cell lines transformed with SV40, and further characterization of the DNA of cell Line 22 cloned into Lambda Gem-11. Restriction digest of DNA from transformed human keratinocyte cell lines 425 passage 12 and 96, 130 passage 131, and 98 passage 76 were conducted using the enzyme Xba I. Two clones from the SV40 transformed keratinocyte cell line 22 were also examined in this experiment; 22(B3), and 22(B4). Southern blot analysis was conducted using ^{32}P-labeled full length SV40 genomic DNA as a probe. Figure 14 (A) is the result of a 24 hour exposure, (B) is the same blot exposed for a period of 5 days. Lane 1, which is genomic DNA from cell line 425 using low passage 12, results in a series of bands that produce intense signals during blot analysis. Lanes 2 through 4, that consist of genomic DNA from cell lines 425, p. 96, and two different DNA preparations using the high passage cell line 130, p. 131, did not result in a signal after the 24 hour period of exposure. Lane 5 is genomic DNA from cell line 98, p. 76,

which produces bands of lower intensity than that of the DNA from cell line 425 p. 12. After a five day exposure figure 14 (B), bands start to appear lane 7 is the clone 22(B3), lane 8 is clone 22(B4) and lane 9 is SV40 control DNA. and clearly shows that these clones contained the viral sequences.

3. cDNA Libraries.

Two cDNA libraries, (G1 and G2), using cell line 425 passage 46 prepared as described in "Materials and Methods" were analyzed with a view towards obtaining transcripts that produced aberrant viral proteins, we have observed T-immuno reactive bands of unusual size during western blot analysis, which will be addressed later. The cDNA libraries were grown on LB medium plates and the resulting colonies were screened by hybridization to ^{32}P-labeled full length SV40 genomic DNA as a probe. The first three independent isolates obtained from G1, were G1(1), G1(2), G1(3), followed by the isolates G2(4), G2(5) from the G2 library, (figures 15 A-D). The result of southern blot analyses confirms that these libraries contained SV40. Size

differentials were observed amongst the different clones examined, G1(1) resulted in two bands that were of lower molecular weight then that of clones G1(2) and G1(3), which may be due to aberrant transcripts. Similar differences in size have been observed between clones prepared using the G1 library and that of G2. G1 clones 2 and 3 are of lower molecular weight then G2(4) and G2(5). After many attempts were made to sequence all of these clones only data from G(2) and G(3) have been analyzed due to the fact that the other clones have not produced discernible results, this may be due to the purity of the PCR templates used to obtain the sequencing data.

4. Polymerase Chain Reaction Clones.

We have used the polymerase chain reaction in the sequence analyses of various clones. The PCR products were prepared by using primers that are specific to the sequences of the SV40 DNA, and others that do not map in SV40 and were expected to anneal to sites in the human genome. The latter should produce clones that have both SV40 and human sequences, these clones could be used

to find the junctional sites between the two genomes. All of the PCR products have been cloned into the vector T7Blue(R), [Novagen]. The first two of these clones were prepared using the genomic cell line 425, one using 425 early and the other using 425 late. The primers used in the preparation of these PCR products were primer 17, which maps on the SV40 genome at the nucleotide positions 3756 <-> 3752, and primer 222 which maps on the SV40 genome at nucleotide positions 291 <-> 257, (figure 16). The PCR products were then run on an agarose gel and the results show that this primer combination produced a DNA fragment which had the expected size of 1728 kb. The cloning procedure was carried out by cutting the desired band from a low melting point agarose gel, purification, ligation to the vector, followed by transformation into JM101 competent cells. Two other sets of PCR clones have also been obtained using similar techniques. The DNA used in these experiments is from cell line 130, passage 131; the first were produced using combinations of the SV40 primers, 17, 19 at nucleotide positions 3756 <-> 3772, and primer 21, (nucleotide positions 4472 <-> 4489), paired with

primer 222, and have been given the designation H-1 through H-1 + n, (where n so far equals 1-9). Southern blot analyses clearly demonstrated that these clones carry a portion of the SV40 genome, (figure 17). The second set of clones is of interest because they have been prepared using one SV40 primer 97D-2, which maps on the SV40 genome at the nucleotide positions 5147 <-> 5127, and the arbitrary primers 9-1, and 10-1, which should map in a region of the human DNA that flanks the virus. These set of clones have been given the nomenclature of S-1 through S-1 + n, (where n so far equals 1-12). This set of clones has been analyzed using the southern blot technique. The results show that the clone identified as S-2 produced an intense signal on the autoradiograph obtained from the blotting procedure, while clones S-1 and S-3 through S-13 result in bands of much lower intensity (figure 18).

B- ENHANCED POLYMERASE CHAIN REACTION PRODUCTS USING LIMITED DIGESTION OF GENOMIC DNA.

1. PCR Products Using Complete Digestion of Genomic DNA.

The rationale for these set of experiments came from the variability in the success of the application of PCR techniques on genomic DNA, which had been isolated from the SV40 transformed cell lines used in our laboratory. The products obtained during the PCR of genomic DNA did not always conform to the results expected. Since purified genomic DNA is a molecule of high molecular weight and the keratinocyte cells used in our experiments have low copy number of viral sequences with respect to that of total genome, the normal DNA-DNA interactions may restrict the availability of the viral sequences to the enzyme used in the chain reaction by exhibiting a stearic effect. The use of restriction endonuclease would cause the DNA to relax, that is open up its nucleotides to the Taq-polymerase protein. Figure 19 shows the results of one such experiment, all of the samples were cut with the restriction enzyme Xba I, and it is

important to note that the digestion time used exceeded two hours. The pattern of digestion differs between the cell lines used in this experiment. The DNA of the infected cell line 130 was digested for a period of four hours, whereas the DNA from the infected cell line 425 was cut for a period of three hours, this coincides with the results obtained from the PCR reactions. The PCR products obtained with the cut genomic DNA of cell line 130 produced diffuse bands, on the other hand the bands recovered from cell line 425 are well defined. The results obtained from southern blot analysis indicate that the bands that positively hybridized to the ^{32}P-labeled full length SV40 genomic DNA probe contained viral sequences and were of the size expected for the primers used in the reaction. The desired bands were then cut out of a low melting point gel and cloned into the T7Blue(R) vector as described above.

2. PCR Products Obtained Using Time Dependent Digestion of Genomic DNA Using the SV40 No-cut Restriction Enzyme Xba I.

Using the data obtained in the PCR reactions from the complete digest of the transformed genomic DNA as a model, experiments were conducted using time dependent digest of the cellular DNA obtained from the infected cell line 130, (high passage number 134). Figure 20 is a Xba I digest of the genomic DNA, aliquots were removed at different time intervals, time zero, 15 minutes, 30 minutes, 60 minutes, 90 minutes, 120 minutes, 240 minutes, 8 μL of each sample was then placed on an 0.8% agarose gel and the results are shown here. The pattern of digestion is clearly defined in the range of least cut DNA at 15 minutes to a complete digest which is achieved at 240 minutes. Southern blot analysis was conducted using ^{32}P-labeled full length SV40 genomic DNA as a probe. The results shown on the right were produced using a 4 day exposure. The polymerase chain reaction, (38 cycles), was performed on 5 μL aliquots from each sample, the primer sets used were 13 <-> 222 and 17 <->

222, these primers had produced products in previous experiments. A sample of 10 μL of the resulting reactions were run on an electrophoretic gel, followed by southern blot analysis, the results were documented on figure 21a and b. The upper panel was obtained using primers 17 <-> 222, which should produce a PCR product that is 1728 kb in length. The reaction done on the genomic DNA which is not cut produced a smear that does not hybridize to the full length SV40 probe. The 15 minute digest results in an intense smear of DNA that demonstrates hybridization to the probe, producing a product that results in a defuse signal on the exposed autoradiograph. At 30 minutes banding starts to appear, but the resultant hybridization gives a product which is less intense and found at a lower molecular weight then that expected. The result at 60 minutes were puzzling; at this time the resultant band is even at a lower molecular weight then the one observed at 30 minutes and hybridization does not yield a discernible product. At 90 minutes the desired band appears and hybridization confirms these results, although bands other then the ones expected develop the

desired band can be obtained by cutting the sample from a low melting point gel. Of interest are the results obtained from the digestion at 240 minutes, here we see the bands start to disappear and the results of the southern blot analysis show a smear that is of a higher intensity of previous digest, but less specific then the product obtained at 90 minutes. The observations that result from the use of primer pair 13 <-> 222, which should produce a product that is 2418 bp in length, differed from the experimental data just presented, although the results clearly demonstrate that the signal produced from the sample that has undergone restriction digestion for a period of 90 minutes results in the best banding pattern and autoradiograph signal when subjected to southern blot analysis, the bands produced during this reaction were not of the expected size for native SV40 DNA. There are many scenarios that may lead to these results, including mutations in the viral DNA at the primer site or mis-priming at a site other then the primers binding site. Other experiments were conducted to prove that these results obtained by the limited digest of genomic DNA were

reproducible. The same type of experiments using a different set of primers 19 <-> 222, which should give PCR products that have a size of 1491 bp, was conducted and the results are shown in figures 22a and b. At time zero when the genomic DNA is not cut no PCR product is observed, but a band that has the desired molecular weight emerged at 15 minutes of initiation of the Xba I restriction digest; bands are also observed at digestion times of 30 minutes, 90 minutes, 120 minutes and 240 minutes, but no PCR product is seen at 60 minutes. Southern blot analysis confirms the results that stem from the gel electrophoresis of these samples and the best PCR results are achieved after 90 minutes.

3- PCR Products Obtained Using Time Dependent Digestion of Genomic DNA Using the SV40 No-cut Restriction Enzyme Sac I and the SV40 One-cut enzyme BamH I.

To verify that the results obtained with the SV40 no-cut enzyme Xba I are reproducible we used another SV40 no-cut enzyme Sac I. These set of experiments were conducted under the same

conditions as described above, using the cellular DNA isolated from the cell line 130. Figure 23a, b and c are the result of the Sac I digest, followed by southern blot analysis of the PCR products. When the SV40 sequences in the genomic DNA is not cut, a band is produced that has a size of approximately 450 bp, when compared to DNA molecular weight marker VI, this band does not hybridize to the SV40 probe. After 15 minutes have passed from the addition of enzyme, a band that has a size of 1491 bp, which is the desired size of the PCR product using the primer combination 19 <-> 222, produces a clear band as shown in the blot analysis. At 30 minutes of digestion this band disappears and then reappears at 60 minutes. From the blot analysis the PCR products obtained at 90 and 120 minutes post addition of enzymes produced the best results, these bands displaying strongest signal during the hybridization. Finally at 240 minutes of digestion the band disappears as seen both on the agarose gel and the southern blot analysis. The results obtained using the SV40 one-cut enzyme BamH I were also significant, (figure 23 d-f). Southern blot analysis shows that there are no clear bands in any of the samples

obtained from the BamH I digest, although binding of the SV40 probe to some of the PCR samples is observed. This may be due to PCR products that are produced in low concentrations and therefore are not observed during gel analyses.

4- PCR Products Obtained During Time Dependent Digestion of Genomic DNA from Cell line 98.

To rule out the possibility that the observed enhancement might reflect idiosyncrasies associated with viral sequences in line 130, we tested another cell line using Xba I. Figure 24 shows that treatment of line 98 genomic DNA with Xba I produced two distinct bands at 90 minutes of digestion while no product was observed at any of the other time points. The larger band was of the size expected for intact viral sequences (1492 kb), while the smaller band migrated at about 670 kb. Interestingly, this smaller band had never been seen in our previous PCR reactions using this same cell line without restriction enzyme treatment.

C- Sequence Analyses:

1. Characterization of Sequences Obtained from the SV40 Input Virus.

The SV40 input virus was sequenced by three methods: 1) Mp18 clones were prepared from Xba I lambda Gem-11 libraries and sequenced using manual techniques, 2) isolates from lambda Gem-11 libraries were used as templates for PCR reaction conducted with primers that are unique to the viral genome, and 3) viral DNA isolated from permissive African green monkey cells was purified and sequenced using PCR mediated direct sequencing techniques, (Table 3). The sequence of the SV40 input viral stock was compared to that of sequences catalogued in GENBANK. Three of the alterations were found in the T-antigen domain of which two changed the amino acid composition of the large T-antigen protein. The substitution at base 3755 (A-> G) is found in the domain where p53 and DNA polymerase α-primase compete for binding to the T-antigen, (150). The binding of p53 to this site could prevent the formation of a replication initiation complex, since the input virus remains viable this change does not effect

the replication capabilities of the viral stock. The amino acid changed was number 355, from leucine to arginine which has a guanidine function that is one of the strongest organic bases and creates a close interaction with the phosphate groups of DNA (151). The substitution at base number 4071 (T-> A), changes the nonpolar amino acid proline for the equally nonpolar residue leucine. A silent substitution occurred at base number 4299 (C-> T). Three substitutions were observed in the T-antigen intron which is region that undergoes change in all of the cell lines analyzed during this study. An alteration was found at base number 5209 (C-> T), this is located within the T-antigen binding site 1, an area responsible for the regulation of the early gene transcripts, mutations in this region result in the overexpression of the early genes, (40). A substitution/deletion was found in the 72-bp palindromic enhancer between SV40 map positions 179 and 185, which is a region of known strain variation, (CTGGTTGC-> C--CTAA). Two substitutions were found in the agno protein site and two more in the late genes of structural protein VP2.

2. Sequence Characterization of Clones from Cell Lines 98, 425 and 130.

Sequencing data was obtained from transformed cell line 98 passage 96 using three different procedures. Genomic DNA from these cells was isolated, digested with Xba I, and ligated to the vector lambda Gem-11. The DNA from the resulting isolates was digested and cloned into the sequencing vector Mp 18 and analyzed using manual sequencing techniques as described in Materials and Methods. This cell line has also produced sequencing data by direct analysis of PCR products using recombinant phage DNA as a template. The primers used in these reactions were combinations of the SV40 primers, 17 and 19 with 222. Figure 25a, shows the results of the PCR reaction. Multiple bands were observed when the samples were screened on an agarose gel, this can be accounted for by a mis-priming event sometimes detected while using PCR with genomic templates. Southern blot analysis was performed on these PCR products to determine which samples could be used for sequencing. The bands which resulted in positive signals were cut out of a low melting point gel,

reamplified using PCR and sequenced. Cell line 98 was also sequenced by using genomic DNA as a template for the PCR reaction, (table 4). Sequence analysis of cell line 425, (passages 12, 46, and 98), was also completed. Data from line 425 at passages 12 and 98 was obtained using both genomic DNA as a template for the PCR reaction and T7blue(R) PCR clones prepared using primers 17 and 222. cDNAs isolated from cell line 425 at the 46^{th} passage were cloned into the plasmid pSPORT1 was sequenced using the M13 forward sequencing primer. DNA from 425 passage 66 was cloned into the sequencing vector pGEM-3Z, but sequence analysis of these clones designated 3-1 and 5-1 have yet to yield any significant data, (figure 25b), [tables 5,6 and 7]. Two other sets of T7blue(R) PCR clones have also been sequenced, these clones were obtained using DNA from cell line 130 passage 131, these clones were given the nomenclature of H-1 through H-1 + n, (where n so far equals 1-9), and S-1 through S-1 + n, (where n so far equals 1-12) as described in material and methods. Cell line 130 has also been sequenced from PCR products acquired from the genomic DNA template, (table 8).

3. Sequence Analyses of SV40 Early Region Sequences From Cell Line 98.

Cell line 98 passage 97 was sequenced as described above, all of the mutations found were substitutions. Unlike cell lines 130 and 425 previous analyses of this cell has proven that it is integrated into the host genome (143). The integrated viral sequences were found to be contained on two BamH I fragments of about 6.9 and 5.2 kb in an anchorage-independent clonal subline (Ag34) prepared from cell line 98. This was verified by sequence analysis done directly on the parent cell line using Mp18 clones obtained from lambda GEM-11 genomic libraries (figure 13). Sequence analyses of cell line 98 p. 97 revealed a cluster of mutations between amino acids 3111 and 3206, seven mutation that code for five amino acids were examined, (table 4). Changes in bases 3111 (A-> C) and 3112 (A-> C) substituted the nonpolar amino acid isoleucine for the strong DNA binding residue arginine. The substitution in base 3179 (A-> T) results in changing the nonpolar amino acid phenylalanine for the nonpolar residue isoleucine. A lysine is changed for an asparagine at base number

3189, which should result in somewhat of a loss in DNA binding affinities. At base number 3199 an arginine is replaced with a lysine, both of which bind DNA strongly. Substitutions at bases 3204 and 3206 change the polar amino acid glutamine for the also polar tyrosine. A survey of these mutations shows that amino acids which may be involved in DNA binding of the T-antigen protein product are changed in such a way that the DNA binding affinity of the protein is equivalent to that of the unmutated product.

Of interest is the fact that interactions of the T-antigen protein product with the p53 protein stimulate transformation, (152), since the mutations described above were located in the functional domain of the T-antigen that binds to the p53 protein, immunoprecipitation, followed by western blot analysis was examined to compare the binding of p53 to T-antigen in all of the cell lines studied (fig 26b). The results of these experiments did not clearly demonstrate any significant alterations in the binding of the T-antigen protein to p53 the loss of a band was observed in immunoprecipitation analyses of cell line 98. Other mutations

were also recorded in the T-antigen intron, a total of eleven of these were substitutions. One mutation was found in the 72-bp palindromic enhancer, at base 205 where a C was replaced by a G. Four mutations were found in the late gene of the structural protein VP1.

4. Sequence Analyses of SV40 Early Region Sequences From Cell Line 425.

SV40 sequences from an early passage of line 425 (passage 12) was sequenced and compared to the input virus to elucidate changes present at this time in culture. At base number 4436 a (T-> C) substitution was recorded, this resulted in the change of the amino acid lysine (residue 128) for glutamine, a basic polar amino acid for a polar amino acid. This substitution was found in the T-antigen domain that contains the nuclear localization signal, a basic region from amino acid 126-132, that directs proteins from the cytoplasm to the nucleus, (153-155). Two substitutions were found in the early-mRNA start site, at bases 5180 (G-> C), 5181 (G-> A), and one addition at the late-mRNA

start site, at base 283 (C). One substitution was recorded at the T-antigen binding site 1, [base number 5206 (T-> C)]. As discussed above this may result in the overexpression of early genes. Another substitution was found in the 72-bp palindromic enhancer at base 209 (A-> G). All of these mutations were also carried by 425 cells studied at higher passages.

cDNA prepared from cells at passage 46 showed two substitutions at bases 3557 (A-> T) and 3559 (C-> T), which where found to map in the ATPase activity, helicase activity, and ATP, p53, polymerase α binding sites. The mutation at amino acid site 420 was silent, the change that occurred at site 419 was an arginine for a lysine both basic polar amino acids, DNA binding at this site should not be effected.

The SV40 early region from a higher passage (p. 97) of this cell line was also sequenced and showed a significant number of mutations clustered within a region between bases 3930 and 4041, affecting amino acids between residues 258-296 within the overlap between the DNA binding and polymerase-α binding sites. Fourteen mutations were recorded in this site, all were

substitutions, of which seven were silent. At base positions 3930 and 4017 the polar amino acid glutamine was changed for the basic amino acid histidine, this substitution would enhance DNA binding. Three amino acid substitutions at bases 3935, 4016 and 4025, are for residues with the same properties, polar for polar or nonpolar for nonpolar. The change at base 3937 results in a valine, nonpolar, instead of a glutamine, polar, which is significant and may effect the binding capabilities of DNA polymerase α. A stop signal results from the substitution at base 3942 (G-> T), this may develop into the formation of aberrant transcripts which have been observed during western blot analysis (figure 27b). Four mutations were found in the T-antigen intron, all of which were substitutions.

5. Sequence Analyses of SV40 Early Region Sequences From Cell Line 130.

Sequence analysis was done on passage number 131 of this cell line (table 8). Five mutations were found at base numbers 3117, 4110, 4113, 4288 and 4299, all of which were silent

substitutions. A mutational cluster was found between bases 4391 and 4486, all of which produced an amino acid change. Out of eleven substitutions found in this area four were changes of residues that exhibit the same functional properties and one mutation at base 4477 (T-> C) exchanged the polar amino acid asparagine for the nonpolar residue glycine. The mutations found between bases 4447 and 4486 are located in the amino-terminus phosphorylation sites, residues 106-124. All of the amino acids that show the potential to be phosphorylated have been changed to amino acids that can not be phosphorylated, with the exception of serine 106, which was not mutated in this cell line. Amino acid 111, serine, has been changed to phenylalanine, threonine 117 has been changed to isoleucine, serine 120 has been exchanged for phenylalanine, serine 123 was substituted by leucine and threonine 124 has been changed to isoleucine. Three mutations were also observed in the T-antigen intron, all of which are substitutions. Two changes were recorded in the small t-antigen, large T-antigen overlap site at bases 4919 (T-> C) and 4987 (C-> T), the first results in an amino acid substitution of the polar

glutamine to the nonpolar glycine, the latter change results into a silent mutation. One substitution was found in the 72-bp palindromic enhancer region at base number 145.

6. Compilation of Mutations Found in cell lines 98, 130 and 425.

A total of 83 mutations were catalogued during this study, of which 12 are located in the input virus, 6 were found in the early passage of cell line 425 and the other 65 or 89% were selected by cells that have been carried in vitro for long periods of time, (figure 28, table 9). A total of 43 mutations or 51.8% were located in the T-antigen functional domain, 15 of these were silent mutations. A total of 13 mutations translated into complete changes in the properties of the amino acids; that is a nonpolar amino acid was changed for a polar residue and vice versa. Out of the total of 43 changes found in this region 15 were pseudo-silent mutations, where a polar amino acid was exchanged for a different polar amino acid that has similar or equal properties. The T-antigen subdomains affected in cell line 130

were the Rb binding site (amino acids 111 and 114), the phosphorylation sites (amino acids 111-124), DNA binding site, helicase (amino acids 143, 172, 177, 233-234) and the ATPase site (amino acid 566). The accumulation of mutations in cell line 425 were found in the following T-antigen subdomains; nuclear location signal (amino acid 128), DNA binding, poly α, helicase regions (amino acids located within numbers 259-295) and ATPase function (419-420). The T-antigen subdomains affected by the mutations in cell line 98 were poly α, helicase region, p53 and the ATPase function (amino acids located within numbers 538-569). A total of 21 or 25.4% were located in the T-antigen intron. Four mutations were found in the 72-bp palindromic enhancer an area known for its different strain variants. Mutations that substitute a C for an A seem to accumulate in this area as observed from the sequence analysis presented. Another 4 mutations were located in the structural protein VP1 and two were mapped to VP2. Two mutations were found in each of the following regions, (t,T)-antigen site, T-antigen binding site I, the early-mRNA start site, and one mutation was found in the late-

mRNA site. The distribution of these mutations is significant because the great majority are either silent, pseudo-silent, or located in regions that would not affect the immortalized state of these SV40 transformed cell lines.

D- Immunofluorescent staining of SV40 T-antigen in a colony of viral-infected human keratinocytes of cell line 98.

Immunofluorescent staining of the large T-antigen was used to study changes in the normal pattern of growth and differentiation which results from the infection of human epithelial cells by the oncogenic virus SV40. Only a few cells in the epithelial colony become positive for T-antigen during the initial days post-infection., as seen in figure 29 (A) which is an epithelial colony at the third passage. As the cells progress through different passages there is an increase in the number of cells that become T-antigen positive. After a period of 2-3 weeks post-infection about 30% of the colonies have T-antigen positive cells randomly distributed throughout the cell culture. At the tenth serial

passage as cells enter a period of "crisis" about 90% of the cells are T-antigen positive [figure 29b]. Once the cells have survived the 15th passage 100% of the cells are T-antigen positive and will eventually become "immortalized".

E- Immunoprecipitation and Western Blot Analysis.

Western blot analysis was conducted on SV40 transformed keratinocyte cell lines 22 p. 113, 425 p. 68, 130 p. 116, and AG34 an agar-selected clone derived from cell line 98 using the large T-antigen specific monoclonal antibody (Ab-2) an antibody raised the amino-terminal portion of the 94 kDa large T-antigen. Figure 27 (A) shows an increased level of antibody binding to cell lines 130 and 425, with a substantial decrease in the binding to cell lines 22 and clonal subline AG34. Cell line 22 is at passage 48 and therefore should have a higher viral copy number then cell lines 425 and 130, however for reasons unknown the results do not conform to this fact. These results point to a possible series of alterations in the expression of the protein product of this cell line. The observation that AG34 shows decreased binding

affinities can be explained due to the fact that it is an agar-selected clone from cell line 98 and during clonal selection mutational changes might have occurred which alter the protein product of the virus. Figure 26 (A) is an immunoprecipitation of large T-antigen using the primary antibody (Ab-1) which is specific to p94 and p21 of SV40 large T and small t antigens respectively, followed by western blotting using (Ab-2) as the secondary antibody. The banding pattern observed in these three cell lines was similar, with a strong interaction of the primary antibody to the p94 protein product which is expected since this protein is required for the maintenance of the transformed state in these cells. Figure 26 (B) is an immunoprecipitation of large T-antigen using the primary antibody (Ab-1), followed by western blot analysis conducted with a p53 monoclonal antibody as the secondary (Ab P53-12). The p53 antibody exhibits a high binding affinity for the T-antigen protein 94 on all three lines, with a loss of a band in line 98. These findings may prove to be significant since a substantial amount of the mutations found in this cell line were localized to the p53 binding domain of the

large T-antigen. These mutations may play a role in the interaction between the large T-antigen product and p53.

F- Serum Dependent Analysis/ Colony Formation Efficiency of Cell Lines 98, 130 and 425:

Cells from lines 98 passage 98, 130 passage 131 and 425 passage 93 at an initial concentration of 1.25 X 10^5 cells/plate were plated into 6 cm plates. Each cell line was analyzed in duplicate at fetal calf serum concentrations of 0.1, 0.5, 2.0, 5.0, and 10.0% in DMEM for a period of seven days. All three cell lines examined exhibited a progressive loss of growth rate when grown in 0.1, 0.5 and 2.0% serum, with little or no viable cells observed after the seven day period, (figures 30-32). Cell line 130 grew more rapidly three days after plating at serum concentration of 10.0% then did the other two cell lines. It grew 33 percent more rapidly then cell line 425 and 66 percent more rapidly then cell line 98. Four days after plating cell line 98 had grown 75 percent more rapidly then cell line 130 and 86 percent more rapidly then cell

line 425 at this concentration. At day six and a concentration of 10.0% serum the maximum amount of cells observed for cell line 425 was 3.3 X 10^5 cells/plate after which time the number of cells started to decrease. Cell numbers reached a maximum of cells per plate for cell line 130 at day seven, 7.2 X 10^5 cells/plate were observed at a serum concentration of 10%. Cell line 98 showed up to a fivefold increase in cell number over a seven day period when compared to 425 cells and almost a twofold increase over cell line 130. The maximum number of cells per plate, 1.35 X 10^6, was observed on day seven using a serum concentration of 10%. These findings suggest that the mutational clusters unique to each cell line may effect growth rate of the mutated cell.

Cells harvested from monolayer cultures were also seeded into 6 cm soft agar plates, in duplicate at a cell density of 10^6/plate and colonies were counted after a 4 day period and colony forming efficiencies were calculated. Cell numbers were obtained from averaged microscope fields (20 per plate) and colony formation

efficiencies normalized to the total number of cells per plate as described previously (141). The results showed that cell lines 98 (3.2%) and 130 (3.1%) exhibited about the same colony forming efficiency on soft agar, while cell line 425 (CFE = 5.0%) formed colonies at an increased ratio. The colonies formed by cell line 425 were also substantially larger, which also reflects better growth capabilities under anchorage independent conditions. These two studies may reflect that differences in growth rate parallel the selective pressure exerted by long term culture. Indeed some of the growth advantages observed in each cell line may arise from the mutations located at the different sites unique to each of the transformed cell lines.

Discussion:

I- Analysis of the Sequences of the SV40 Input virus and SV40-Transformed Human Keratinocytes.

1. SV40 Input Virus:

The SV40 input virus was sequenced as described above, although the viral stock exhibits significant changes from those reported in the GENBANK libraries this viral genome was able to infect and grow in permissive cells and to transform human keratinocyte cells with the subsequent passage to the immortalized state (figure 28, table 3). The deletion/substitution found in between bases 179 and 185 is in a region known to be present in strain variants. It is also important to note that the changes found in the input virus were also carried by all of the cell lines sequenced in this study.

2. Mutations in the SV40 integrants of viral-transformed cell lines 425, 130, and 98.

In this study, integrated viral sequences from SV40 transformed keratinocytes were sequenced with a view towards examining stable alterations in the viral genome acquired during long term serial passage of cells. Analysis of RNA viruses undergoing increased number of serial passages has been examined by others (156-158). This type of virus exhibits a high percent of mutations on passage in the absence of mutagenic agents. Spontaneous mutations occurring in the RNA genome with rates between 10^{-3} to 10^{-4} per incorporated nucleotide have been observed (159). These high rates of change in the genome of these viral molecules has been attributed to errors during replication (RNA to DNA to RNA), and to the lack of proof reading enzymes during the replicative process. On the other hand during the evolutionary changes that take place in DNA viruses have been observed to accumulate at much lower rates in the order of 10^{-8} to 10^{-11} per incorporated nucleotide in the DNA genome, and some

of these mutations can revert to the original phenotype (160-161). Analysis of certain mutational events observed during the replication of SV40 in infected cells have been studied, high multiplicity infection, induces the replication and duplication of defective viruses due to genetic complementation. Viral molecules that contain defective genomes start to accumulate when taken through serial passage in this environment. These SV40 variants exhibit multiple deletions, duplications and insertions in their viral DNA (162). Some of the SV40 genomes are covalently linked to the cellular DNA of the host and eventually start to lose most of the viral DNA, and this leads to viral particles that are incapable of transformation. Mutants such as these have multiple sites in which a substantial number of nucleotides have been affected. We have studied three independent cell lines all of which showed great fidelity in their DNA sequences with the exception of the T-antigen intron and a mutational cluster unique to each one of the cell lines. Sequence analyses of viral sequences from cell line 425 at p.12, p.46 and p.97 have shown that the accumulation of the mutations, most of

which are point mutations, found in this and in cell lines 98, and 130 is a progressive process over several hundred cell generations. The low passage number 12 of cell line 425 had a total of six mutations that were also carried by subsequent passages. The intermediate passage 46 had only two more mutations than p.12, for a total of 14. Late passage 97 of cell line 425 has amassed a total of 32 changes, or 56% more mutations than the intermediate passage. The appearance of mutations and the changes observed in the other two cell lines studied result in an increase in certain growth capabilities of the cell. These observations lead to the hypothesis that such alterations confer growth advantage(s) as a result of cell selection occurring over a period of long term culture. Our laboratory is the first to study the accumulation of point mutations acquired in DNA viruses that have been carried in dilute serial passage over long term culture.

3. Cell line 98.

A cluster of mutations were found between amino acids 3111 and 3206, seven mutations that affect five amino acids were examined (table 4). The changes in amino acids are summarized as follows: one nonpolar amino acid was substituted by a polar amino acid, while a nonpolar residue was replaced by the same, three polar amino acids were changed to three polar amino acids. The T-antigen domain affected was the p53 binding site, binding of p53 to this site results in a decrease in the growth potential of the virus. This decrease is due to a competition between protein-53 and DNA polymerase α-primase which also binds to this site, this would prevent the formation of the replication initiation complex formed by T-antigen, DNA polymerase α-primase and other cellular factors, therefore increasing the growth suppressing activity of p53 with a concomitant decrease in the transformation potential of the virus (19, 152). On the other hand, if T-antigen protein binds to p53 in a protein-protein complex this would lead to the neutralization of the growth-suppressing properties of p53. Changes observed in the amino

acids in this region may lead to structural differences in the T-antigen protein product, which could enhance the affinity of T-antigen for p53, thereby stimulating transformation (152). Immunoprecipitation, followed by western blotting techniques were examined to compare the binding of p53 to T-antigen in all of the independent cell lines. Cell line 98 showed a loss of a band during this analyses which may reflect changes in the binding of p53 to the large T-antigen. If the mutations found in the p53 binding site contribute to a higher binding affinity between p53 and T-antigen, this could alter the growth capabilities of this cell line. Other mutations were also recorded in the T-antigen intron which has a tendency for mutational drifts of which a total of eleven of these were substitutions. A higher incidence of mutations was observed in the intron region of the virus in cell line 98, this observation also leads us to believe that the mutations described above were due to the selective nature of the viral-cellular interaction and not due to random events.

4. Cell line 425.

Sequence analysis of this cell line produced fourteen mutations all clustered between bases 3930 and 4041 (tables 5-8) affecting amino acids between residues 258-296 found within the overlap between the DNA and DNA polymerase-α binding sites. Of the fourteen mutations recorded in this cluster seven were silent. The accumulation of the mutations recorded in the three passages examined shows a progression consistent with the fact that certain mutations are selectively retained by cells over long term culture. This is evident if we follow the mutational pattern exhibited by these three passages, early, intermediate and late as described above. This series of alterations confer enhanced features as observed during colony formation in soft agar, where this cell line was by far the more efficient at forming colonies in soft agar; these colonies were larger and grew more rapidly then those of cell lines 98 and 130. This advantage in colony growth formation is just another example of cell selection. Four mutations were also observed in the T-antigen intron.

5. Cell line 130.

Mutations that were observed in a cluster between amino acid sites 4447 and 4486 are involved in the regulation of replication due to phosphorylation or dephosphorylation of certain serine and threonine residues. Dephosphorylation of serine residues 120 and 123 have been shown to enhance T-antigen binding to site II, with an increase in replication activity (74,77). Another study has focused on the phosphorylation of threonine 124, mutants that carry an alanine at this site led to loss of DNA replication activity both in vivo and in vitro and also exhibit defective site II DNA binding activity (76). It is worthy to note that serum dependent analysis combined with colony formation efficiency on soft agar show this cell line exhibits intermediate growth potential when compared to the other two cell lines. Since the mutation and dephosphorylation of threonine 124 destroys replication in other systems studied, this makes the point that the combination of mutations in this area may overcome the debilitating effects produced by this single mutation since this

has not been observed in our studies. Three mutations were observed in the T-antigen intron of the virus in this cell line.

6. Summary of the Total Mutations Accumulated in cell lines 98, 130 and 425.

The cumulative effect of the mutations studied in all of the cell lines further points to a selective mechanism(s) employed by the cell to enhance its growth potential. For the three cell lines studied, 43 mutations were located in the T-antigen domain, of which 30 or 70% in this region, were either silent or pseudo-silent, and only 13 or 30% expressed a complete change in the functional characteristics of the amino acids transcribed. The T-antigen intron was also found to have a high incidence of mutations, 21 in total, this is considerable since this region is located within bases 4571 and 4918, which only comprises 345 bases of a total of 2469 that makeup the T-antigen gene. The mutations located in the T-antigen intron will not translate into a defective or mutant protein while on the other hand, debilitating mutations found in the rest of T-antigen may indeed produce a

non-functional protein. Since the continued expression of the viral oncogene is essential for the maintenance of the transformed state, the mutations in the T-antigen where such that only a few would result in significant changes, and only those changes that enhance the growth potential of the cell without destroying the T-antigen protein are retained. This points to the fact that although the viral DNA is subject to selective pressure due to progressive evolution it maintains similar functional properties in T-antigen.

2- Errors in Replication That Lead to Mutational Clusters.

1. Proposed Mechanism by Which Mutations are Arranged into Clusters.

How viral DNA replication results in the accumulation of mutations in cells over long term culture has not been determined. Replication of the SV40 genome requires the presence of T-antigen and the replication system of the host. Replication begins at the viral origin and like its mammalian

host, proceeds in a bi-directional semi-discontinuous manner (19). Studies using cell-free replication extracts have helped to identify some of the viral and cellular proteins necessary for replication event to occur (153,163-164). Two distinct complexes are formed during the replicative process. The first is known as the pre-initiation complex, it is formed by the binding of T-antigen, assuming the structure of a double hexamer, with the origin of replication. This pre-initiation complex requires other factors such as the human single-strand-specific DNA binding protein (human SSB protein), DNA topoisomerase activity, and the presence of ATP, which results in unwinding of the viral origin. The second, termed initiation complex is formed when T-antigen and associated co-factors bind to DNA polymerase-α primase, the protein-protein interactions between T-antigen and the primase are highly specific (150,165-167). The initiation complex is thought to play a role in the proper alignment of DNA polymerase-α, so that this enzyme may initiate DNA chain elongation (19). Sequence analysis of SV40 DNA from cell lines 98, 130 and 425 does note answer the question of why there is an accumulation of

mutations within a relatively small area of the viral genome, the mutational clusters found in the three cell lines studied range in size between 39 bases in cell line 130, to 111 in cell line 425 (figure 28). If the first mutations observed in the mutational clusters lead to errors in the initiation complex of replication, subsequent rounds of replication may induce more changes in the affected area leading to the accumulation of mutations in a specific region as shown by the sequence data presented herein. Studies done on RNA viruses demonstrated that the rate of replication of some retroviruses is more than four orders of magnitude higher than the replication system of the host (167-169). This is important because an increase in spontaneous mutations have been observed in viruses that exhibit higher turnover rates. Since SV40 combines many of the replication activities on one protein, the replication event may take place at a higher rate, which may induce an accumulation of mutations observed at the cluster sites. Once all of the mutations have been amassed in the T-antigen domain, the protein expressed by the

mutated genome may have the same or similar features as the wild-type, in a way acting as a structural revertant as is the case in cell line 98. Figure 33 shows a mechanism by which the DNA binding affinity of the mutated large T-antigen product that has undergone a number of serial passages may return to a functional state that is similar to the unmutated protein. What is certain is that the mutational clusters observed in the three different cell lines result in cells that exhibit increased growth potentials. The mechanism described above can also account for the loss of viral copies since mutations that persist through the genome of certain viruses initially integrated may lead to the loss of the SV40 variants that no longer contribute to the immortalized state of the cell.

3- Use of PCR to Study Genomic DNA Templates Infected With Viral Sequences.

Genomic DNA preparations derived from all three cell lines studied exhibited poor template activity when PCR was carried out on target sequences that included the sites in which the

cluster of mutations were located. In particular cell line 130 remained refractory to amplification yielding no reproducible amplification products within this region until pre-treatment was conducted with the restriction enzyme Xba I (figure 22). Techniques for preparing genomic DNAs from higher eukaryotes are tailored to minimize nicking caused by physical manipulation of samples during extraction. However, these procedures often result in viscous suspensions in which some of the DNA may be incompletely solubilized or partially aggregated. Conditions for PCR might be enhanced by small amounts of nicking by making the target template sequences more available for reaction or optimizing primer interaction. The presence of aggregates might explain the unexpected finding that template activity can sometimes exhibit multiple peaks of enhancement. Thus the purity of the DNA sample could dictate how effective the restriction treatment will be and therefore define the actual time course for PCR products obtained. Whatever the actual mechanism, substantial enhancement of amplification of the viral templates was always observed within 2 hours of restriction enzyme

treatment and often after only 15 minutes. Because restriction sites for enzymes with hexameric recognition sequences are expected to occur on the average only once in every 4^6 base pairs, there is a good likelihood that any given restriction enzyme will not cut within a typical amplimer-defined DNA segment. Therefore, only a small panel of different candidate enzymes need be tested even in cases where no sequence information for the PCR target segment is available. Selection of restriction enzymes with low cutting efficiencies, such as rare cutters would also minimize the possibility of cleavage within the target segment which must be avoided. Selection of low efficiency enzymes might also provide an additional advantage inasmuch as too rapid digestion of template would be prevented so that an optimal digestion time range may be selected.

Table 1. Position of the functional domains of the SV40 large T-antigen.

Position of Domain	Function
1-82, N-Terminus	Binding to polymerase α-primase, (Polα), Binding to cellular protein p300.
102-115	Binding to proteins p105 and pRb.
106-124	Phosphorylation cluster sites.
126-132	Nuclear location signal.
131-259	SV40 DNA binding domain.
131-627	Helicase activity.
270-517	Binding to p53 and polymerase α-primase.
302-320	Zinc finger potencial metal binding region.
418-627	ATP binding and ATPase activity.
639-701	Phosphorylation cluster sites.

Table 2. Structure and position of all primers used in the polymerase chain reaction.

Primer	Orientation	Position	SV40/Arbitary	Sequence
10	⇦	2331-2315	SV40	CCAGAAGTGTTGGTAAA
14	⇦	1658-1641	SV40	AAAGCACTCCACCTCAG
16	⇦	1334-1317	SV40	ACTTGCCTGAATACTGTC
18	⇦	1002-986	SV40	AAGATACTGAACACTGT
20	⇦	617-601	SV40	GCAGCTTCAGACACAGT
222	⇦	291-257	SV40	CTGAGGCGGAAAGAACCA-GCTGTGGAATGTGT
97D-2	⇦	5147-5127	SV40	ACGAGAGGAATCTTTGCAGC
SV-E1	⇦	5131-5115	SV40	GCAGCTAATGGACCTTC
13	⇨	3066-3082	SV40	AGGTCTGTACCAAATTA
15	⇨	3415-3431	SV40	AACTGGTCAATAGCTAC
17	⇨	3756-3772	SV40	TTGTAGGCTATCAACCC
19	⇨	4043-4059	SV40	GATTAAAATCATGCTCC
21	⇨	4472-4489	SV40	CCTCATCATCACTAGAT
23	⇨	4879-4896	SV40	ATTGCATCAACACCAGG
87G-1	⇨	361-379	SV40	GGCCTCCGTTAAGGTTCGT
9-1	NA	?	Arbitary	CTGCTCTCA
10-1	NA	?	Arbitary	CTGATCCATG

Table 3. Sequence analysis of the SV40 input virus.

SV40 GENBANK	Amino Acid	SV40 Input	Amino Acid Change	Type of Change	Functional Domain
3755 (A) 4071 (T) 4299 (C)	Leu Pro Lys	G A T	Arg Leu Lys	Substitution Substitution Substitution	T-antigen
4642 (G) 4849 (C) 4879 (C)		T T T		Substitution Substitution Substitution	T-antigen intron
5209 (C)		T		Substitution	T-antigen binding site I
179 (CTGGTTGC)		C - - - CTAA		Substitution/ Deletion	72-bp palindromic enhancer
399 (A) 445 (G)		G C		Substitution Substitution	Agno protein
574 (T) 583 (G)		G C		Substitution Substitution	VP2 VP2

Table 4. Sequence analysis of cell line 98, passage 97.

SV40 Input	Amino Acid	Change	New Amino Acid	Type of Change	Functional Domain
3111 (A)	Ile	C	Arg	Substitution	T-antigen
3112 (A)	Ile	C	Arg	Substitution	
3179 (A)	Phe	T	Ile	Substitution	
3189 (T)	Lys	G	Asn	Substitution	
3199 (C)	Arg	T	Lys	Substitution	
3204 (C)	Gln	A	Tyr	Substitution	
3206 (G)	Gln	A	Tyr	Substitution	
4709 (A)		G		Substitution	T-antigen intron
4721 (A)		C		Substitution	
4743 (A)		T		Substitution	
4771 (A)		C		Substitution	
4787 (T)		C		Substitution	
4788 (T)		C		Substitution	
4789 (T)		C		Substitution	
4813 (T)		C		Substitution	
4820 (C)		A		Substitution	
4839 (C)		T		Substitution	
4852 (T)		C		Substitution	
205 (C)		G		Substitution	72-bp palindromic enhancer
1523 (A)		C		Substitution	VP1
1597 (A)		T		Substitution	
1627 (A)		T		Substitution	
1629 (T)		G		Substitution	

Table 5. Sequence analysis of cell line 425, passage 12.

SV40 Input	Amino Acid	Change	New Amino Acid	Type of Change	Functional Domain
4436 (T)	Lys	C	Gln	Substitution	T-antigen
5180 (G) 5181 (G)		C A		Substitution Substitution	E-mRNA start site
5206 (T)		C		Substitution	T-antigen binding site I
209 (A)		G		Substitution	72-bp palindromic enhancer
283		C		Insertion	L-mRNA start site

Table 6. Sequence analysis of cell line 425 cDNA, passage 46.

SV40 Input	Amino Acid	Change	New Amino Acid	Type of Change	Functional Domain
3557 (A)	Tyr	T	Tyr	Substitution	T-antigen
3559 (C)	Arg	T	Lys	Substitution	

Table 7. Sequence analysis of cell line 425, passage 97.

SV40 Input	Amino Acid	Change	New Amino Acid	Type of Change	Functional Domain
3930 (C)	Gln	G	His	Substitution	T-antigen
3935 (A)	Phe	G	Leu	Substitution	
3937 (T)	Glu	A	Val	Substitution	
3940 (A)	Leu	T	Leu	Substitution	
3942 (G)	Tyr	T	Stop	Substitution	
3987 (T)	Ala	C	Ala	Substitution	
3999 (T)	Val	C	Val	Substitution	
4002 (A)	Leu	T	Leu	Substitution	
4011 (G)	Ser	A	Ser	Substitution	
4016 (C)	Val	G	Leu	Substitution	
4017 (T)	Gln	A	His	Substitution	
4024 (G)	Thr	T	Asn	Substitution	
4035 (T)	Glu	C	Glu	Substitution	
4041 (T)	Pro	C	Pro	Substitution	
4690 (A)		G		Substitution	T-antigen intron
4720 (C)		C		Substitution	
4730 (C)		T		Substitution	
4839 (G)		T		Substitution	

Table 8. Sequence analysis of cell line 130, passage 131.

SV40 Input	Amino Acid	Change	New Amino Acid	Type of Change	Functional Domain
3117 (T)	Arg	C	Arg	Substitution	T-antigen
4110 (C)	Leu	T	Leu	Substitution	
4113 (G)	Ala	T	Ala	Substitution	
4288 (T)	Glu	G	Glu	Substitution	
4299 (C)	Lys	T	Lys	Substitution	
4391 (G)	Leu	C	Val	Substitution	
4447 (G)	Thr	A	Ile	Substitution	
4449 (A)	Ser	T	Leu	Substitution	
4450 (G)	Ser	A	Leu	Substitution	
4459 (G)	Ser	A	Phe	Substitution	
4461 (G)	Asp	T	Glu	Substitution	
4468 (G)	Thr	A	Ile	Substitution	
4470 (A)	Ala	T	Val	Substitution	
4471 (G)	Ala	A	Val	Substitution	
4477 (T)	Asp	C	Gly	Substitution	
4486 (G)	Ser	A	Phe	Substitution	
4606 (A)		G		Substitution	T-antigen intron
4839 (C)		T		Substitution	
4841 (G)		A		Substitution	
4919 (T)	Glu	C	Gly	Substitution	t, T-antigen
4987 (C)	Lys	T	Lys	Substitution	t, T-antigen
145 (C)		A		Substitution	72-bp palindromic enhancer

Table 9. Compilation of Mutations Examined.

Location	Type of Amino Acid Change			Location Other Then T-antigen	Total Number of Mutations	Percent
	Silent	Pseudo-silent [a]	Complete			
T-antigen	15	15	13	N.A.[b]	43	51.8%
T-antigen intron	N.A.	N.A.	N.A.	N.A.	21	25.4%
t, T-antigen	0	1	1	N.A.	2	2.4%
T-antigen binding site I	N.A.	N.A.	N.A.	2	2	2.4%
Early-mRNA start site	N.A.	N.A.	N.A.	2	2	2.4%
72-bp palindromic enhancer	N.A.	N.A.	N.A.	4	4	4.8%
Late-mRNA start site	N.A.	N.A.	N.A.	1	1	1.2%
Agno protein	N.D.[c]	N.D.	N.D.	2	2	2.4%
VP2	N.D.	N.D.	N.D.	2	2	2.4%
VP1	N.D.	N.D.	N.D.	4	4	4.8%

a = The term pseudo-silent was used to identify amino acids that have the same or similar properties as explained in the text. b = nonapplicable. c = not determined

Figure 1. Structural organization of the SV40 genome. The early region encodes for the small t-antigen and the large T-antigen. The late region encodes for three capsid proteins, VP1, VP2, and VP3. The late region also encodes for the viron assembly protein LP-1 (agno). Included are the origin of replication, (Ori), the transcription map and the large T-antigen binding sites at and near the SV40 origin of replication.

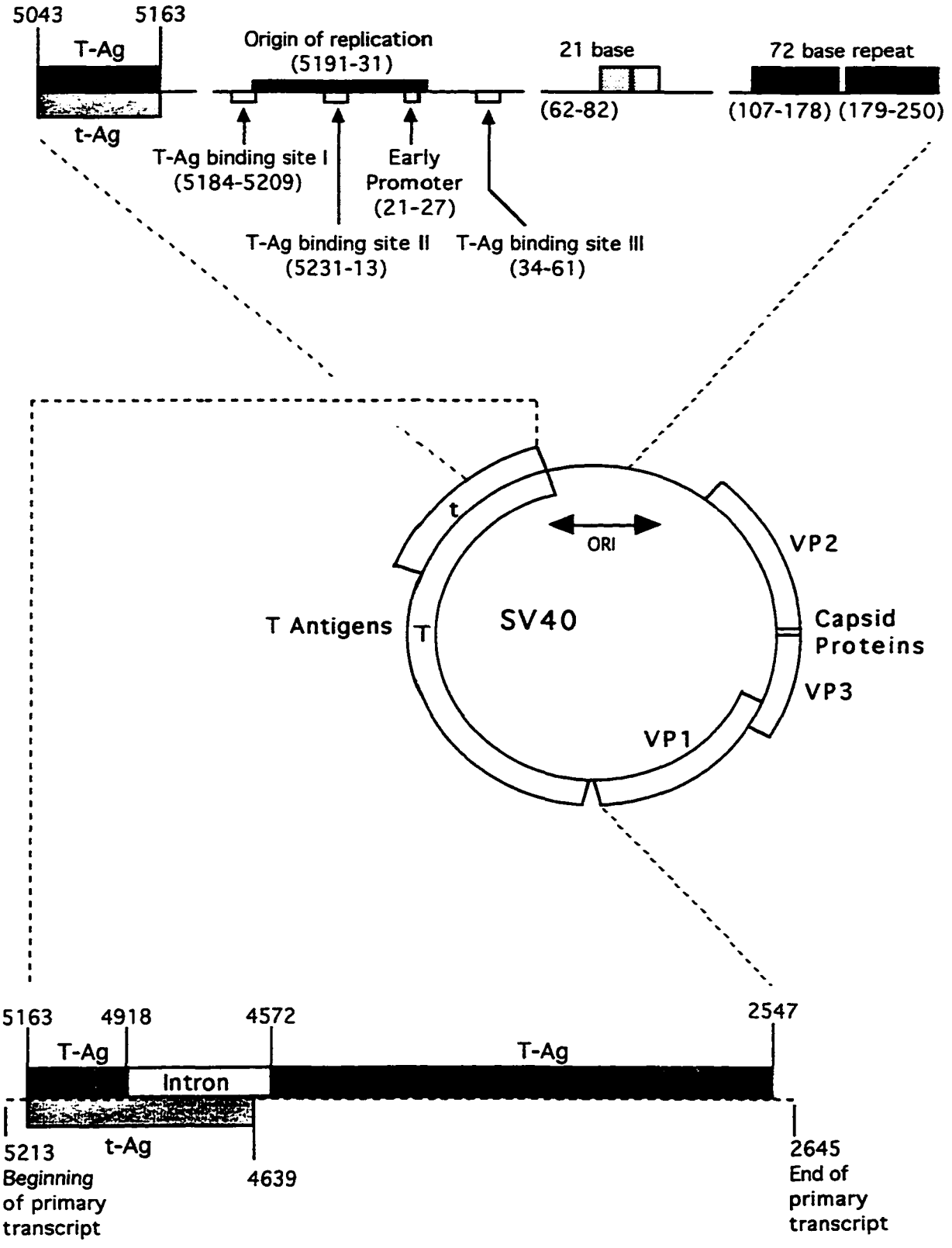


Figure 2. Functional domains of the SV40 large T-antigen. There are two regions of the protein that bind to DNA polymerase α-primase, (pol α). The product of the retinoblastoma gene, (Rb), and the cellular protein p107, which binds to the same region. Included are the binding sites of the nuclear location signal, (NLS), the zinc finger metal binding region, the control region of SV40 DNA, (DNA binding), the regions of ATP binding, ATPase and helicase activity. The host range, which is required in certain cells for virion production.
Adapted from Fanning and Knippers et al.1992

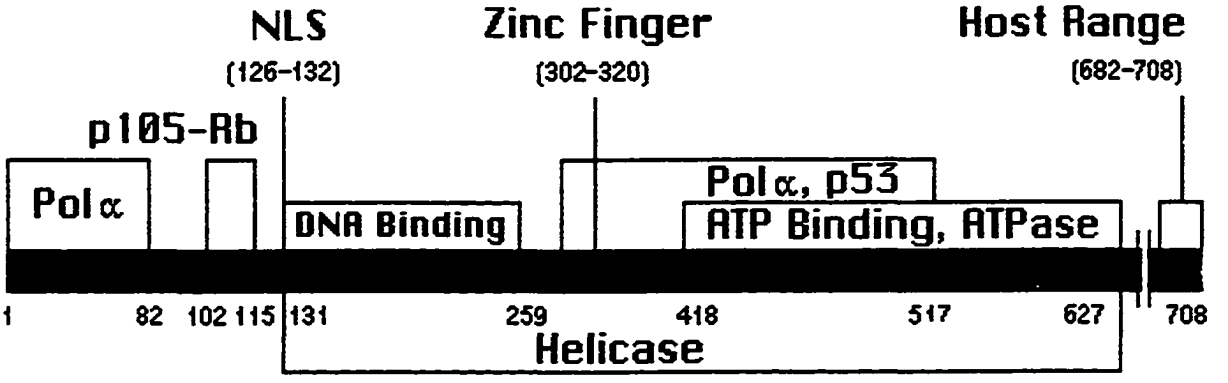


Figure 3. (A) Transcriptional control of the SV40 genome. Early transcription is made up of three key regions, a TATA box, that is responsible for the initiation of transcription, a 21 base pair repeat region rich in -GC- residues, and two identical 72 base pair repeat regions. The two identical 21 base pair repeats have a 21 base pair homologous repeat and are located at nucleotide sites 40-60, 62-82 and 83-103 and all have -GC- rich regions, (GGCGGG), which are involved in the regulation of early RNA transcription by acting as binding sites for the general transcription factor, (SP1). (B) Phosphorylation cluster sites of the large T-antigen. The two phosphorylation cluster sites are found between residues 106-124 at the aminoterminus and 639-701 at the carboxyl terminus. All residues except for three are found to be overexpressed in mammalian cells, the exceptions are residues, 111, 665, and 667, which have only been found in the cells of insects.

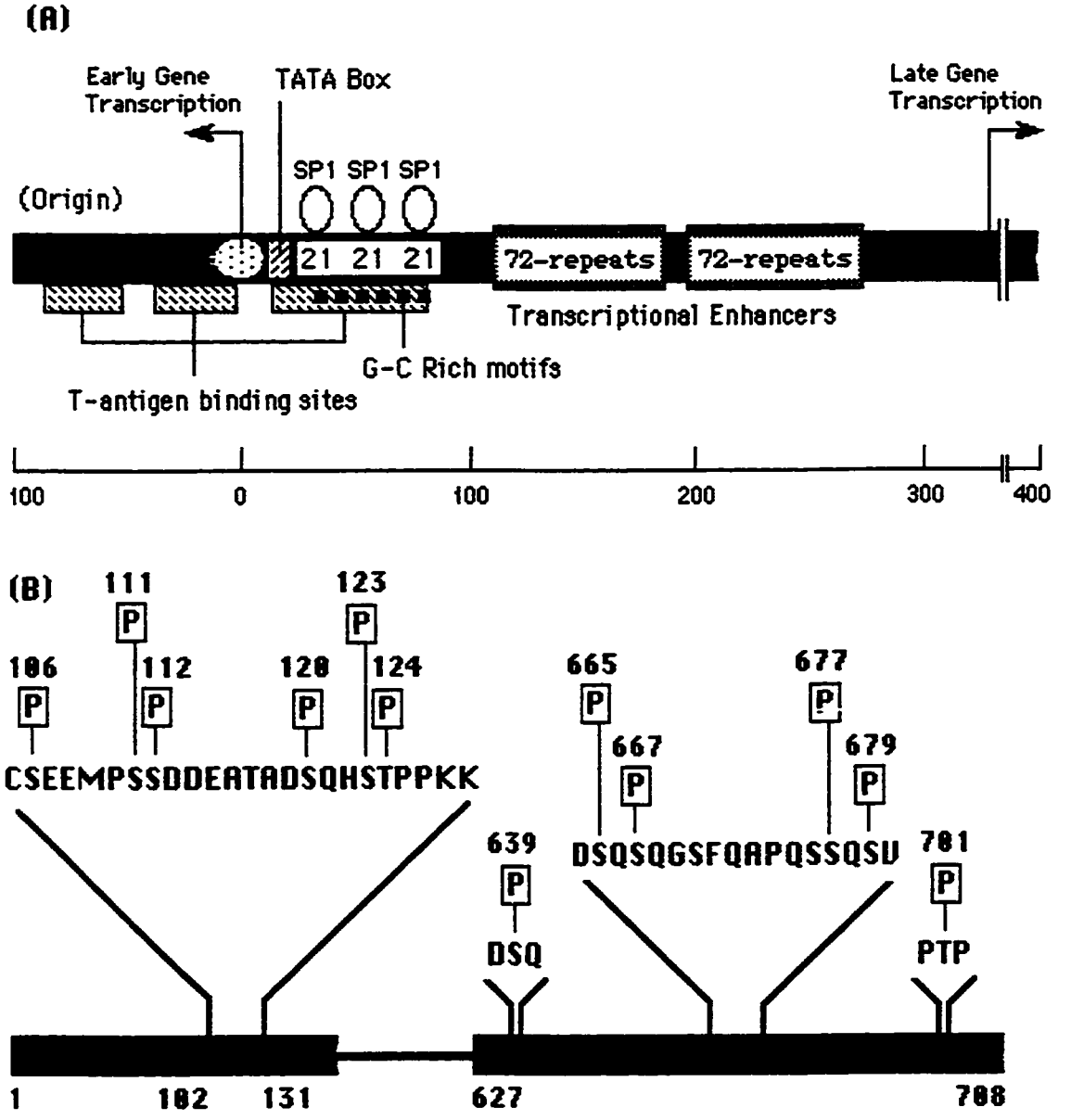
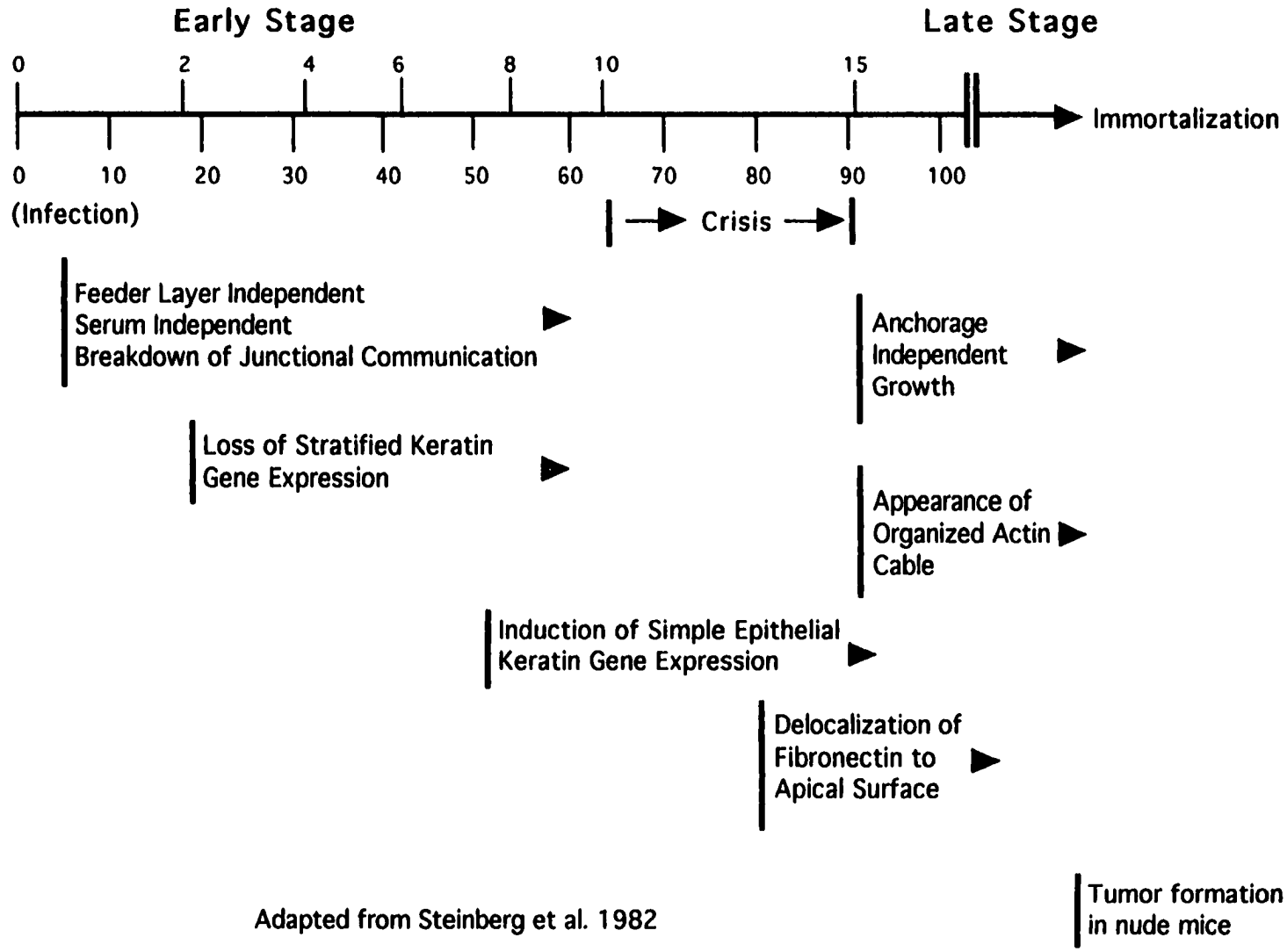


Figure 4. Stages in the transformation of human epidermal keratinocytes by the SV40 virus. There are three defined periods in the generation of "immortalized" cell lines, precrisis, crisis and postcrisis which demonstrate certain cytologic and histochemical features. During the early stage of viral infection known as precrisis feeder layer and serum independence, breakdown of junction communication, loss of stratified keratin, induction of simple epithelial, and expression of certain genes are observed. During crisis a marked decrease in growth potential is observed and the cells exhibit a cytopathic effect, (CPE), that includes the delocalization of fibronectin to apical surface. The cells now enter the stage of postcrisis in which cells acquire the capacity to form colonies in agar and experience anchorage independent growth and the appearance of organized actin cables.



Adapted from Steinberg et al. 1982

Figure 5 Schematic line map of the genomic cloning vector Lambda Gem-11. T7 RNA polymerase promoter, (19406-19422), T7 RNA polymerase initiation site, (19423), SP6 RNA polymerase promoter, (32231-32247), SP6 RNA polymerase initiation site, (32230). Spi selection for nonrecombinant background. Genomic DNA fractionation is required, (144-145).

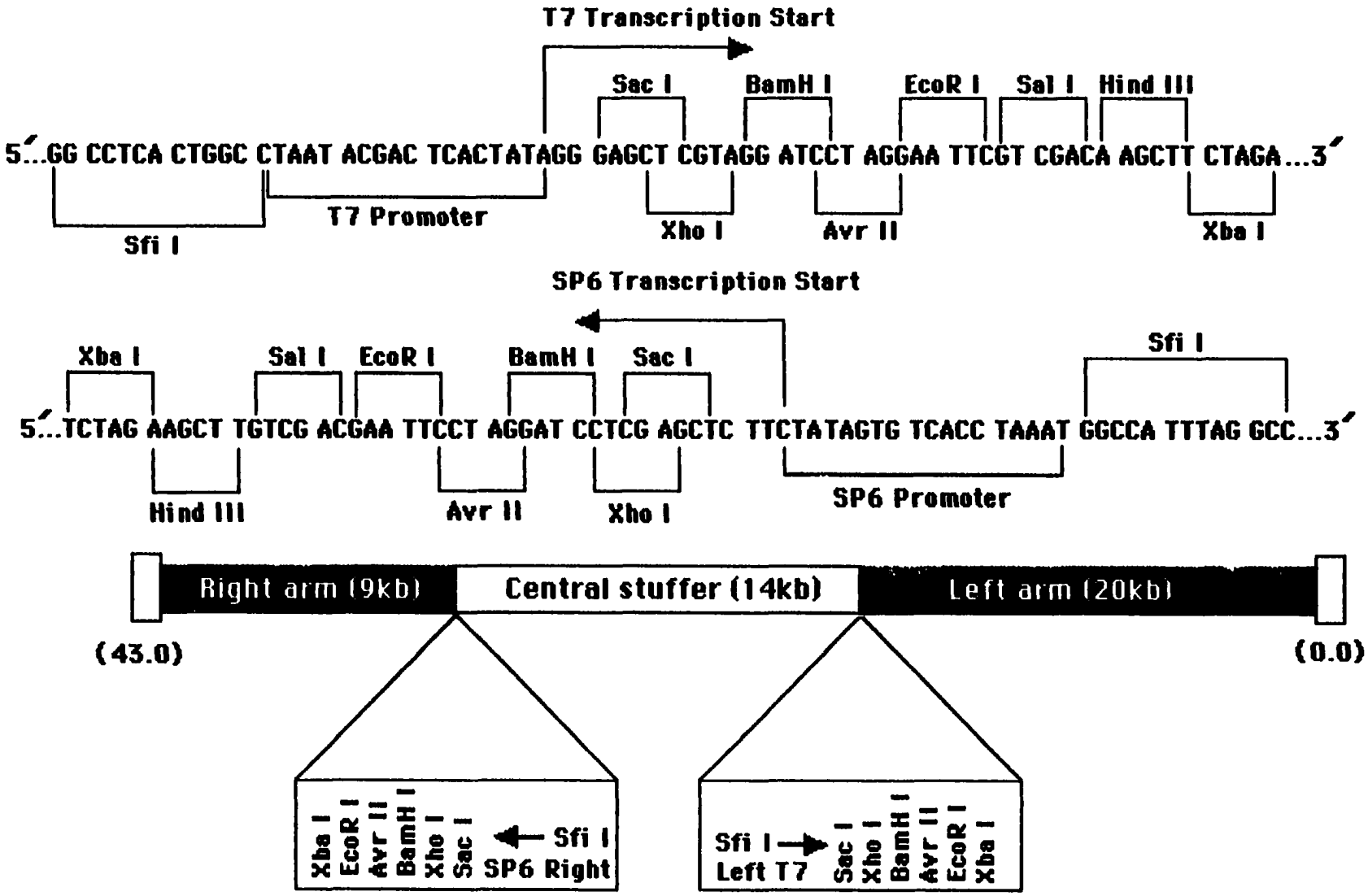
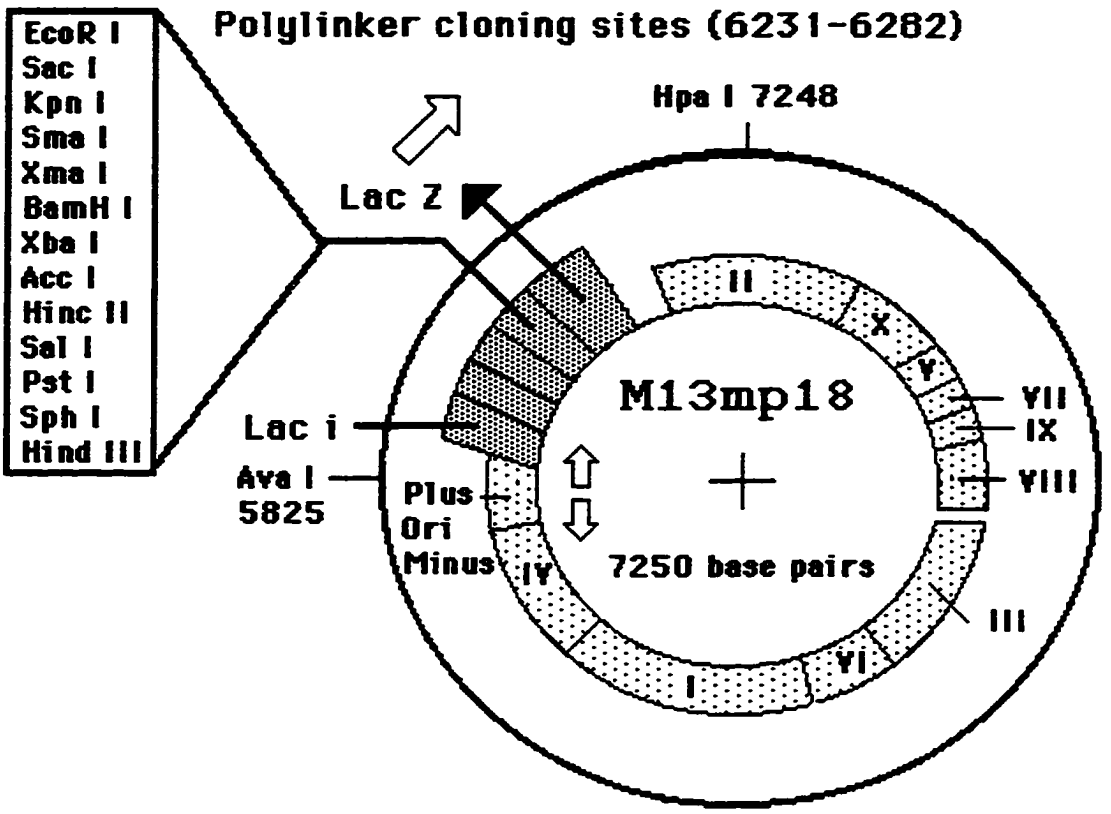


Figure 6. (A) Schematic circular map of the sequencing vector M13 mp18 DNA. The nucleotide numbering is clockwise around the viral plus strand in the 5' to 3' direction. The polylinker includes 10 hexanucleotide recognition sequences for 13 different restriction enzymes. (B) Schematic line map of the sequencing vector pGem-3Z. T7 RNA polymerase initiation site, (1), SP6 RNA polymerase initiation site, (69). Lac Z start codon at 108, lac operon sequences, (2561-2724; 94-223), β-lactamase (Amp-r) coding region, (1265-2125). The polylinker includes 10 hexanucleotide recognition sequences for 13 different restriction enzymes.

(A)



(B)

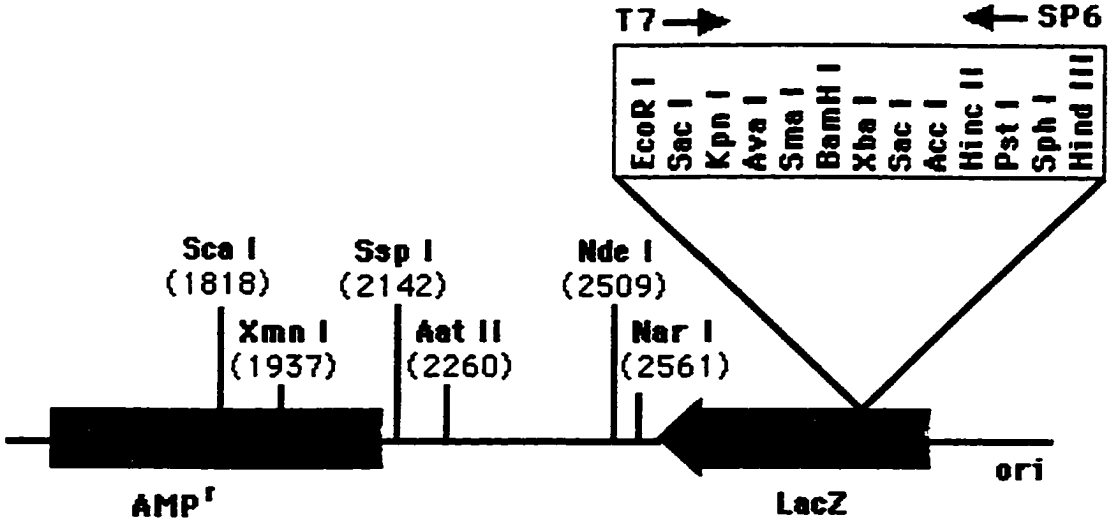


Figure 7. Schematic line map indicating the orientation of the SV40 genomic DNA primers used in the polymerase chain reaction.

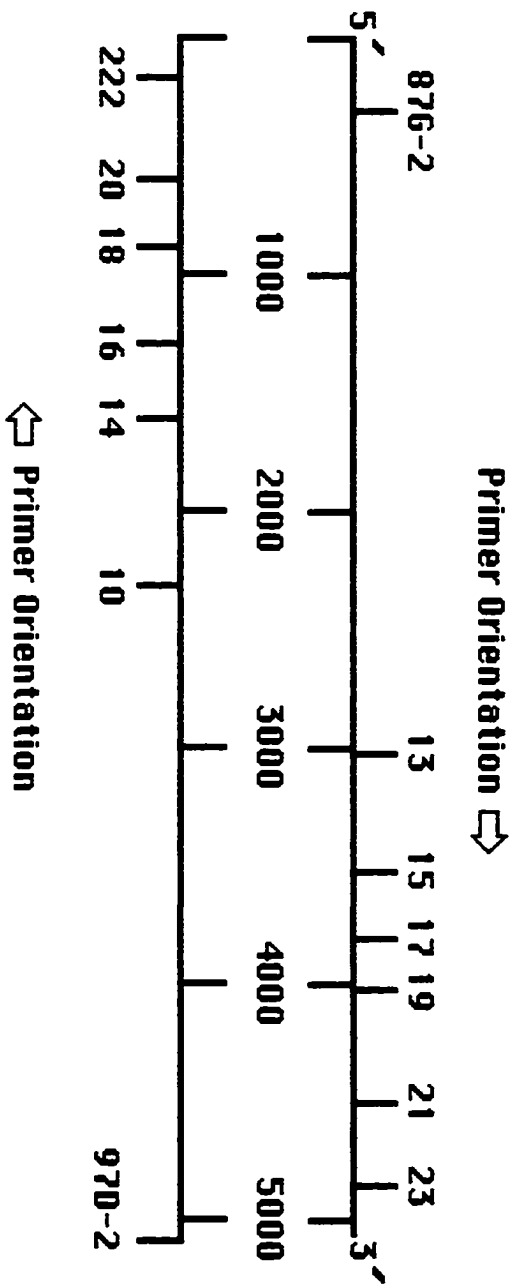


Figure 8. (A) Schematic line map of the polymerase chain reaction/ sequencing vector pT7Blue(R). Lac Z start codon at 1, lac Z α-peptide, (1-282), T7 promoter, (24-40), f1 origin of replication, (284-739), *bla* coding sequence, (871-1728), pUC origin, (2489). The polylinker includes15 recognition sequences for 18 different restriction enzymes, [146-149]. (B) Schematic line map of the plasmid pSPORT 1. The map includes SP6 promoter, (174-183), T7 promoter, (298-318), and the sites for lac OPZ', fl intergenic region, AMPr, lac I and the origin. The polylinker includes 19 recognition sequences.

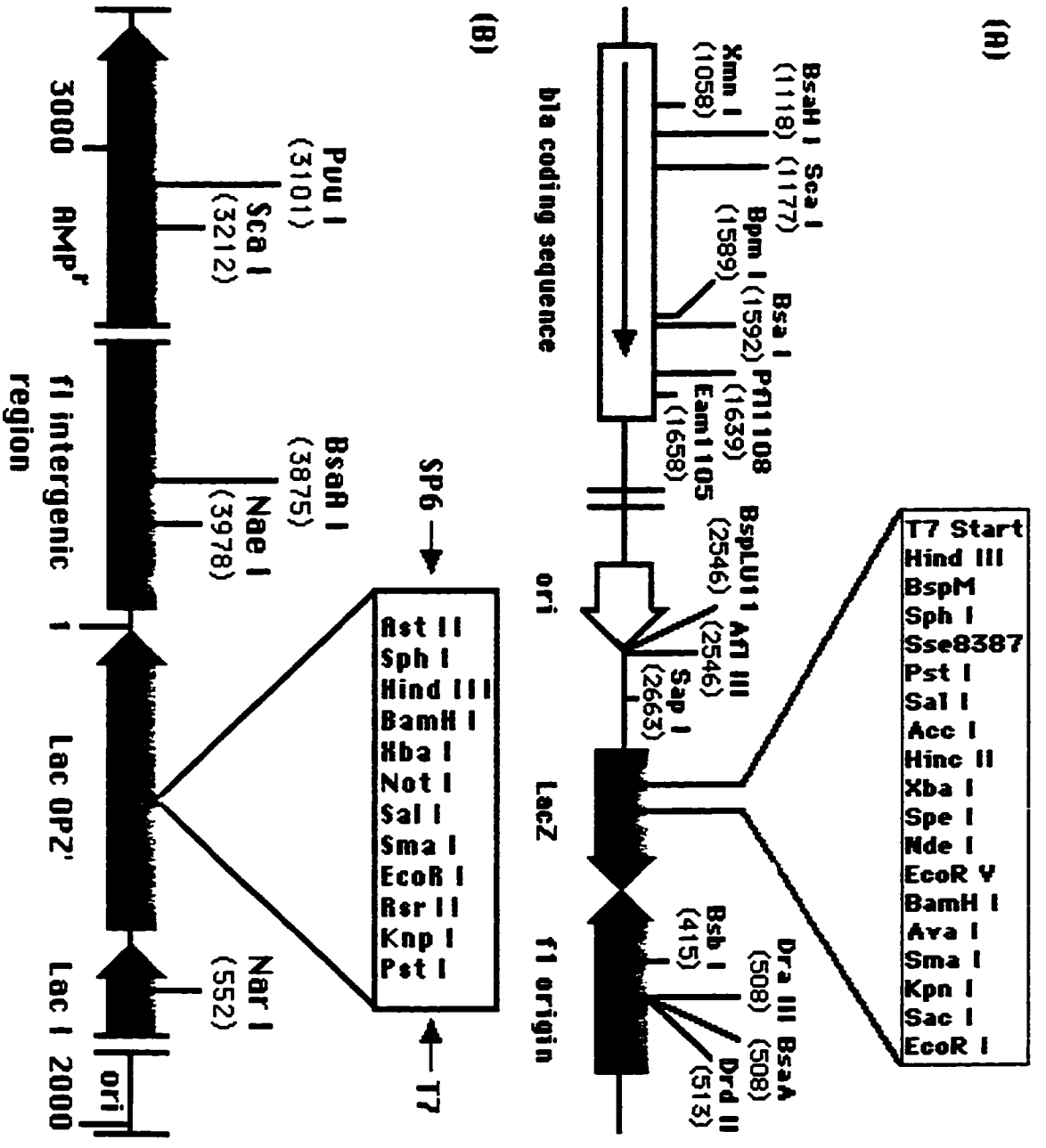


Figure 9 (A) Southern blot analysis of the DNA from transformed human keratinocyte cell lines #22, (passage number 27), and #425, (passage number 32). The DNA was digested with the restriction enzymes BamH I and EcoR I, products were run on an 0.8% agarose gel, stained with ethidium bromide. The southern blot analysis was conducted using ^{32}P-labeled full length SV40 genomic DNA as a probe. The results shown on the left were produced using a 96 hour exposure. The marker used was lambda DNA digested with Hind III. (B) The results shown on the right were obtained from the same blot using a 24 hour exposure.

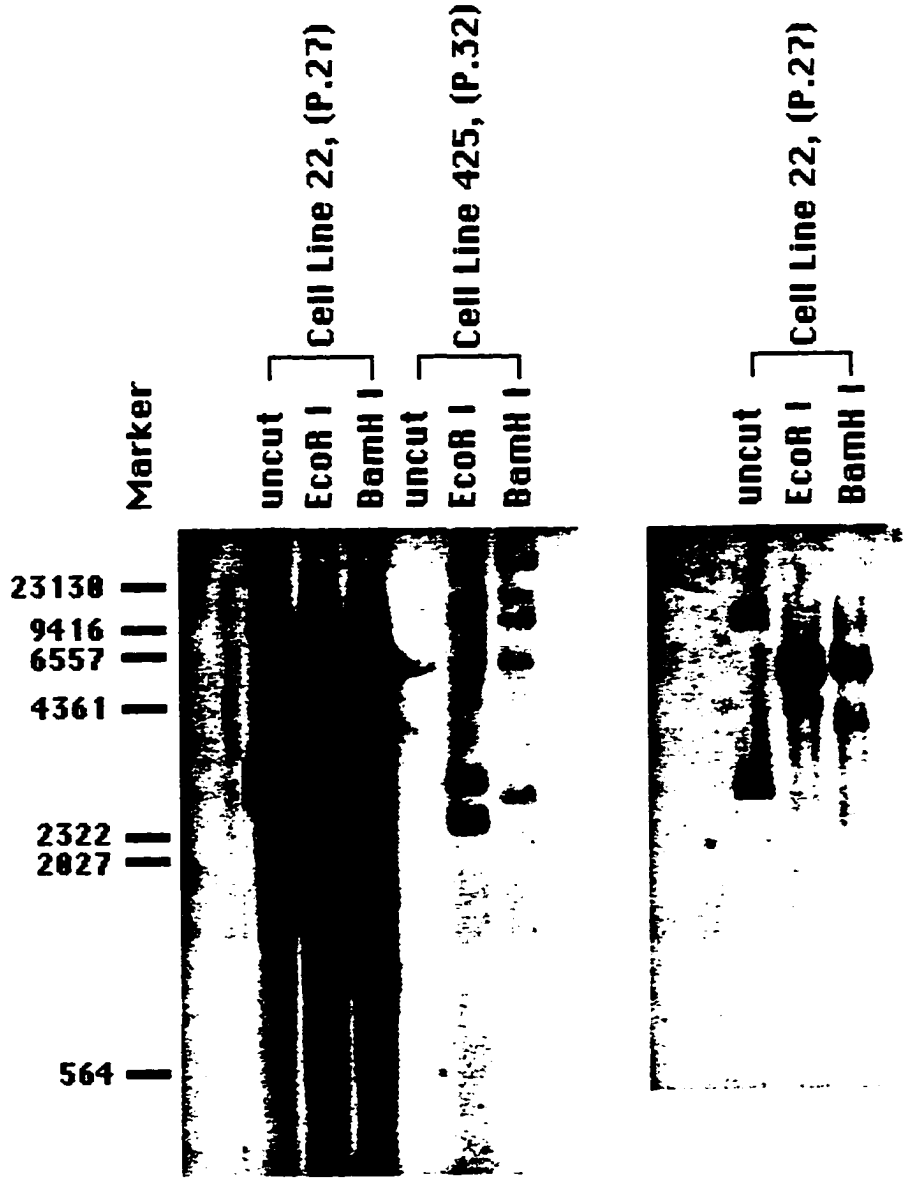


Figure 10. Southern blot analysis of the DNA from transformed human keratinocyte cell line #98, (passage number 76). The digestion products were run on an 0.8% agarose gel, stained with ethidium bromide. The marker used was lambda DNA digested with Hind III. The restriction enzymes used were Ava I, EcoR I, Sac I, Sal I, Xba I and Xho I. All enzymes are no-cut for SV40 except for EcoR I, which has a unique site at position 1782. Southern blot analysis was conducted using ^{32}P-labeled full length SV40 genomic DNA as a probe. The results shown were produced using a 120 hour exposure.

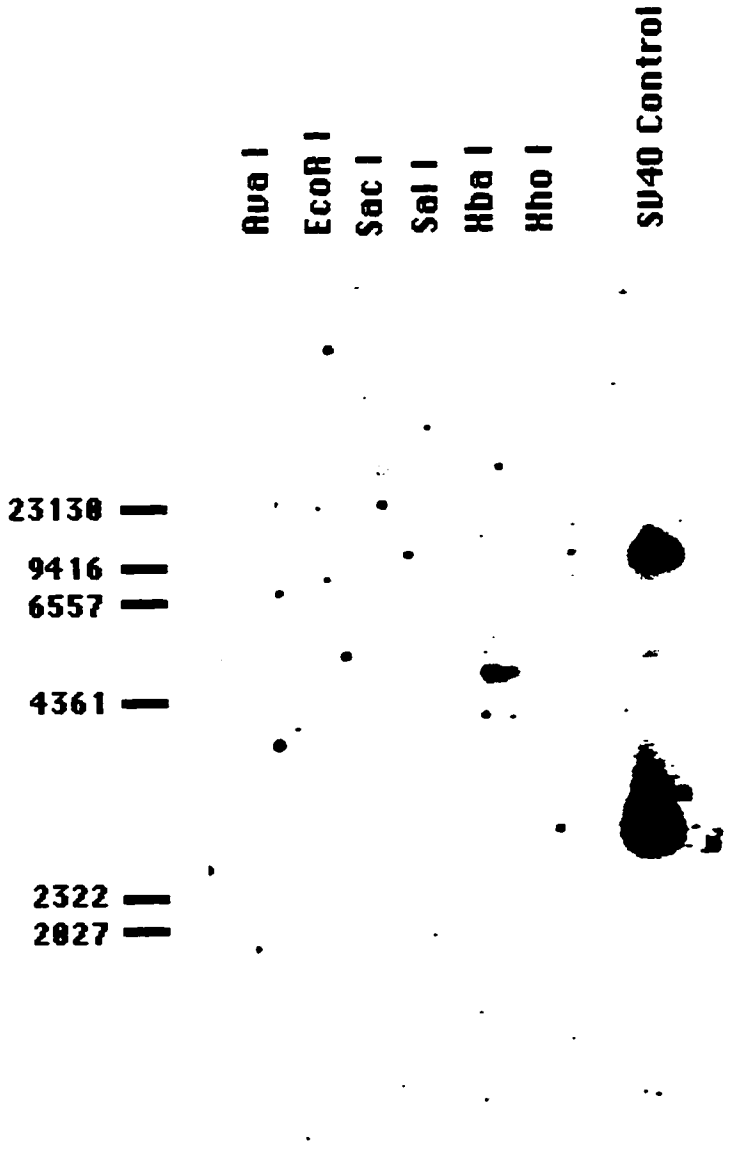
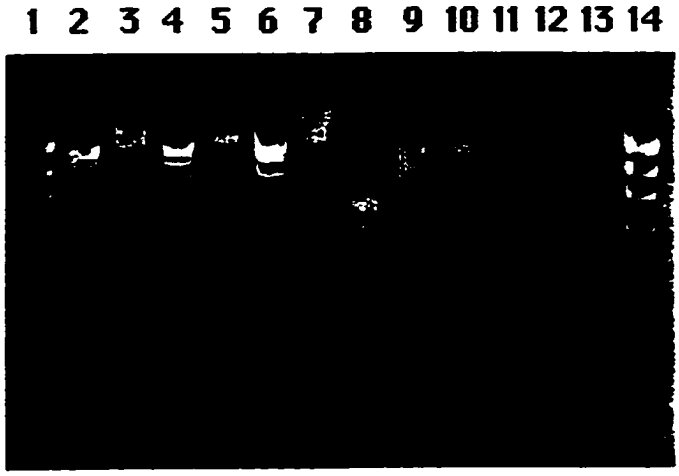
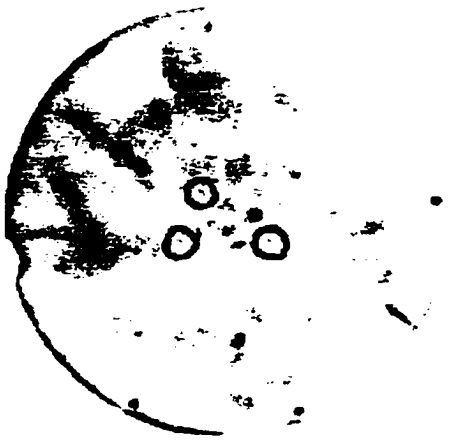


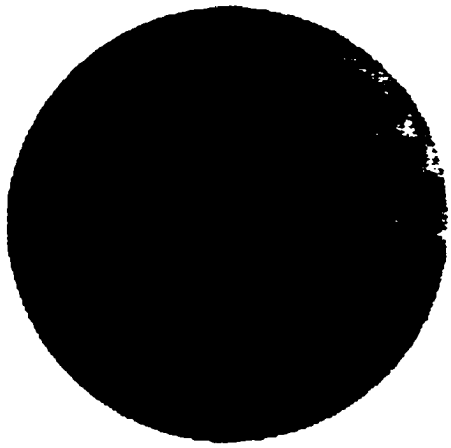
Figure 11. (A) Preparation and isolation of clones using SV40 transformed cell line 425, passage 66, into the lambda phage cloning vector Lambda Gem-11. Lanes 1 and 14 are the marker lambda DNA digested with Hind III. Lanes 2 and 4 are Lambda Gem-11 vector cut with the restriction enzyme Xba I. Lanes 3 and 5 are ligations of the Xba I cut vectors. Lane 6 is the vector cut with Sac I, lane 7 is the Sac I digested vector religated. Lanes 8 and 9 are the genomic DNA cut with Xba I and then religated. Lanes 10 through 13 are a different extraction of the same genomic DNA, first cut with Xba I and religated, (lanes 10-11), then cut with Sac I and religated, (lanes 12-13). All samples were run on an 0.8% agarose gel stained with ethidium bromide. (B) Plaque purification of an independent isolate from the lambda phage library using transformed cell line 425, the first round plate was blotted to a nitrocellulose membrane and hybridized to ^{32} P-labeled full length SV40 genomic DNA as a probe, the circular marks enclose SV40 hybridizing plaques that were further purified. (C) Shows the autoradiogram of a nitrocellulose membrane blotted to a plate of purified clone, (3-1).



(A)



(B)



(C)

Figure 12. (A) Restriction digest of partially purified Lambda Gem-11 clone, (3-1), from transformed human keratinocyte cell line #425, (passage 66). The digestion products were run on an 0.8% agarose gel, stained with ethidium bromide. The size marker used was lambda DNA digested with Hind III. Lane 1 is the marker, lane 2 is the uncut clone, lane 3 through 6 were cut with the restriction enzymes shown. The arrow points to the 5 kb fragment that was subsequently recovered from the purified clone. (B) Southern blot analysis of samples in (A) were conducted using ^{32}P-labeled full length SV40 genomic DNA as a probe. The arrow indicates hybridization to the 5 kb fragment. (C) Agarose gel electrophoresis restriction digests on the purified Lambda Gem-11 clone, 3-1. Lane 1 is the marker, lane 2 is the uncut clone.

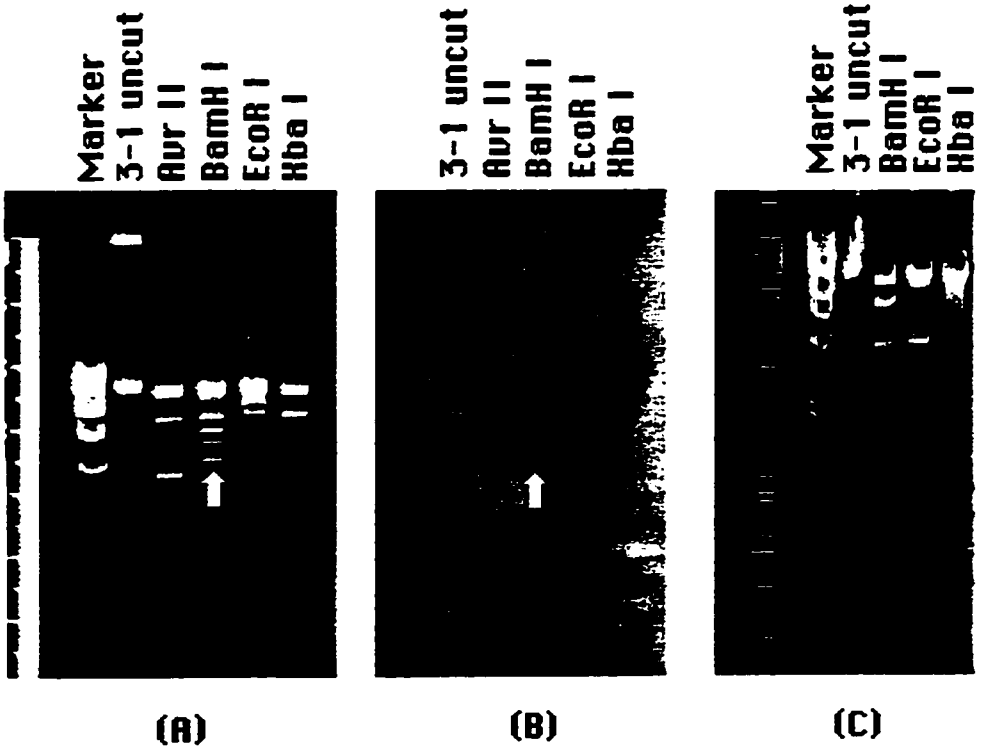
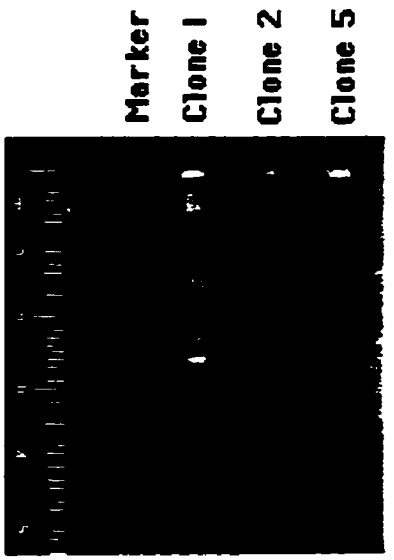
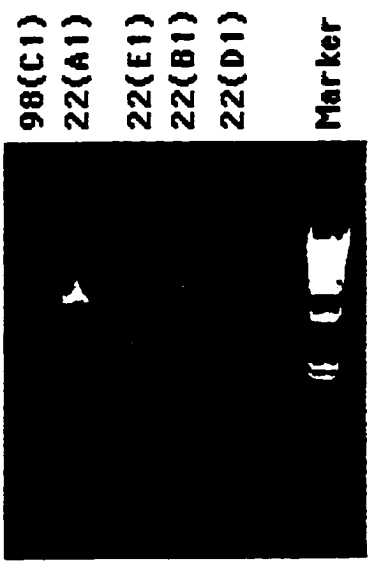


Figure 13 (A) Characterization of the insert cloned into Lambda Gem-11 libraries from line 98. Three phage clones were analyzed in this experiment. The clones designated C-1, C-2 and C-5 were digested with the restriction enzyme Xba I, the products obtained were run on an 0.8% agarose gel, stained with ethidium bromide. The marker used was lambda DNA digested with Hind III. (B) Lane 1 is the marker, lane 2 is clone 1, lane 3 is clone 2, lane 4 is clone 5. The size of the insert has been estimated to be above 5.2 kb. Southern blot analysis produces one band at the site of the insert, when the clonal DNA is screened with ^{32} P-labeled full length SV40 genomic DNA. (C) DNA from the cell line #22 cloned into Lambda Gem-11. Four clones from the SV40 transformed keratinocyte cell line #22 were obtained in this experiment, lane 1 is the clone C-1 from the cell line #98, (used as a control), lane 2 is 22(A1), lane 3 is 22(E1), lane 4 is 22(B1), lane 5 is 22(D1), lane 6 is the marker. (D) Southern blot analysis produced positive signals, for all clones described, when the clonal DNA was screened with ^{32} P-labeled full length SV40 DNA.

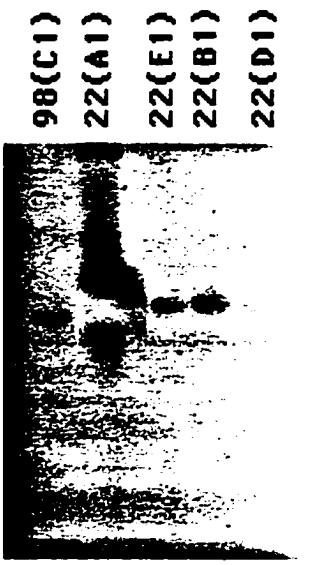


(A)

(B)



(C)



(D)

Figure 14 Genomic Profile of Different Cell Lines Transformed with SV40, and Characterization of the DNA from Cell Line #22 Cloned into Lambda Gem-11. (A) Southern Blot analysis of the restriction digest of genomic DNA from transformed human keratinocyte cell lines #425 passage 12 and 96, #130 passage 131, #98 passage 76 and from the cloned cell line #22 passage 66, using the enzyme Xba I are shown. The results shown in (A) were produced using a 24 hour exposure. (B) Is the same southern blot exposed for a period of five days. The blot was conducted using ^{32}P-labeled full length SV40 genomic DNA as a probe.

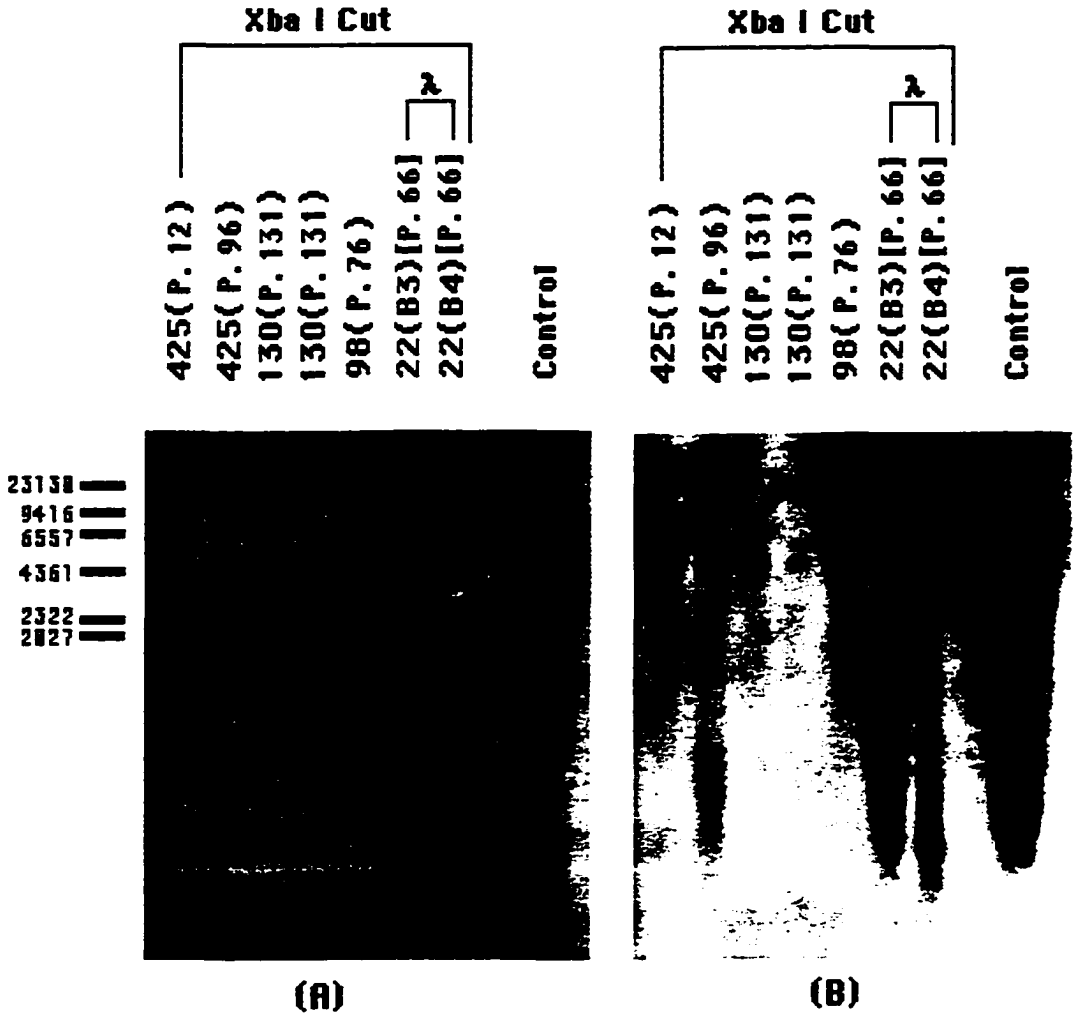
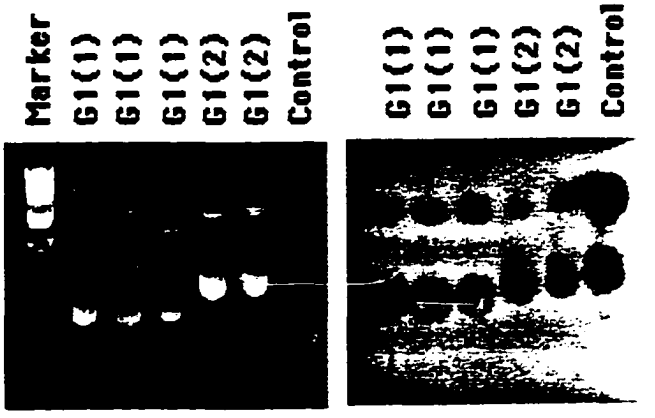
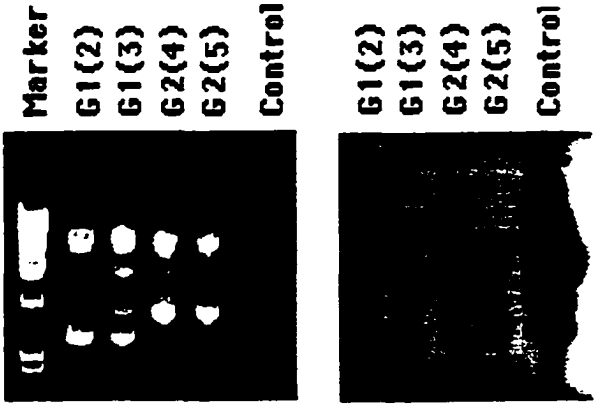


Figure 15. Characterization of cDNA Libraries. Three clones were obtained from the stock G-1, the nomenclature used was G1(1), G1(2) and G1(3), while two clones were obtained from the stock G-2, named G2(4) and G2(5). All of the samples were run on an 0.8% agarose gel, stained with ethidium bromide. The marker used was lambda DNA digested with Hind III. Southern blot analysis was conducted using ^{32}P-labeled full length SV40 genomic DNA as a probe. (A) Agarose gel of clones G1(1) and G1(2). Lane 1 is the marker and lane 7 is SV40 DNA used as a control. (B) Southern blot analysis of samples G1(1) and G1(2) which confirms that these libraries contained SV40. (C) Agarose gel of clones G1(3), G2(4) and G2(5). Lane 1 is the marker and lane 6 is SV40 DNA used as a control. (D) Southern blot analysis of samples G2(4) and G2(5). The result of southern blot analysis confirms that these libraries contained SV40. The results are significant because the isolates from the G-2 stock, are distinct from the clones previously screened.



(A) (B)



(C) (D)

Figure16 Electrophoretic analysis of PCR products used to prepare clones in the vector T7Blue(R). The genomic DNA were #22 and #425. The PCR products were run on an 0.8% agarose gel and stained with ethidium bromide. The marker used was a mixture of pBR322 DNA cleaved with Bgl I and Hinf I. The samples were cut out of a low melting point gel and cloned into the T7Blue(R) vector.

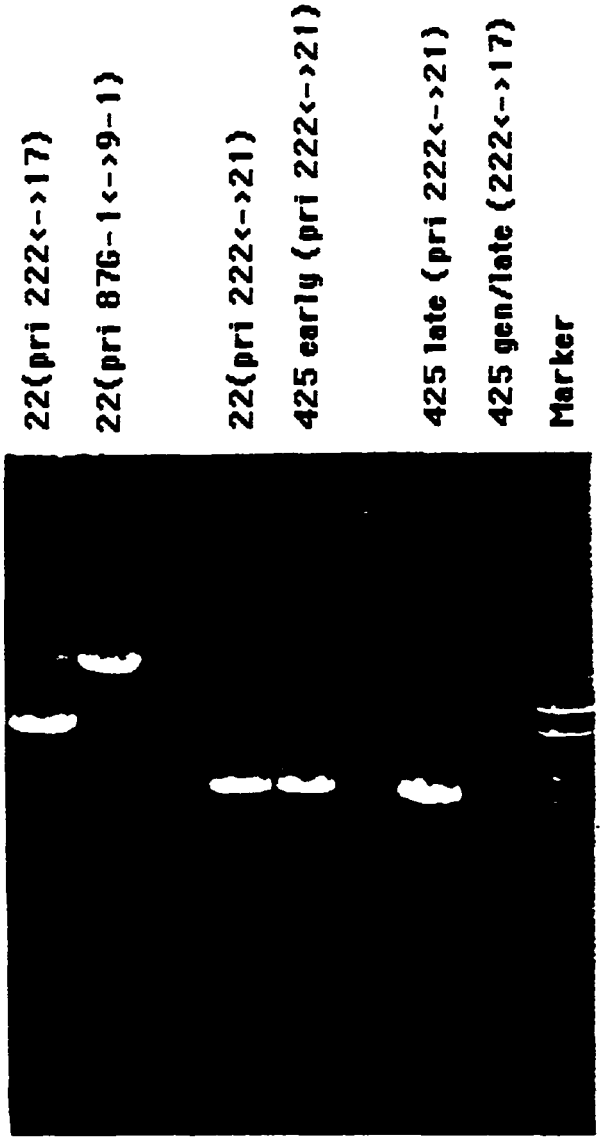
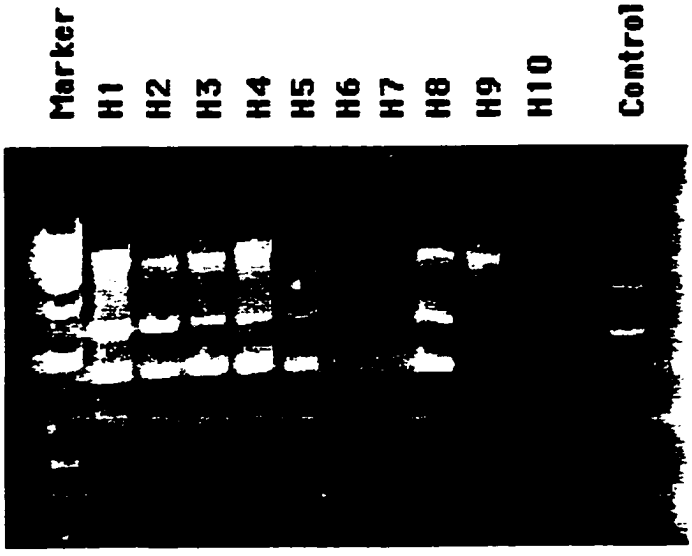
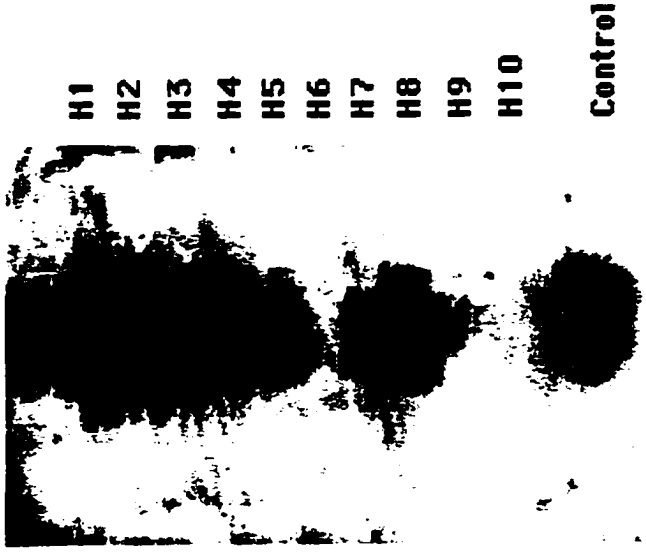


Figure 17. (A) Characterization of Polymerase Chain Reaction Products from Cell Line #130 Cloned into the Vector T7Blue(R). The DNA used in these experiments is from the high passage cell line #130, (passage number 131), the combination of primers used was the SV40 primer 22 with the primers 17, 19, and 21. The nomenclature given these sets of clones was H-1 through H-1 + n, (where n so far equals 1-9). The clones were run on an 0.8% agarose gel, stained with ethidium bromide. The marker used was lambda DNA digested with Hind III. Lane 1 is the marker and lane 12 is SV40 DNA used as a control,[left]. (B) Southern blot analysis was conducted using ^{32} P-labeled full length SV40 genomic DNA as a probe, confirms that all clones, with the exception of H-6, H-7 and H-10 contained SV40, [right].

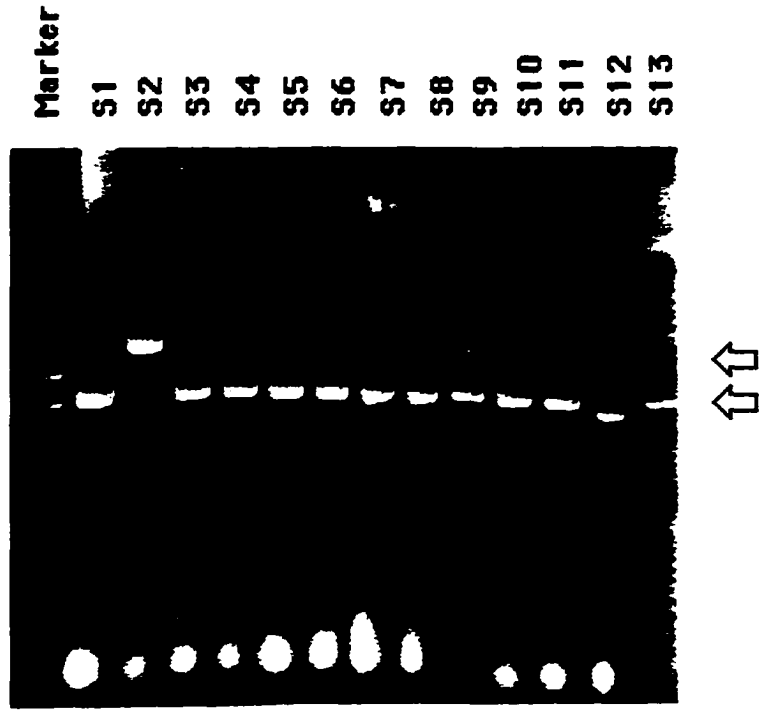


(A)

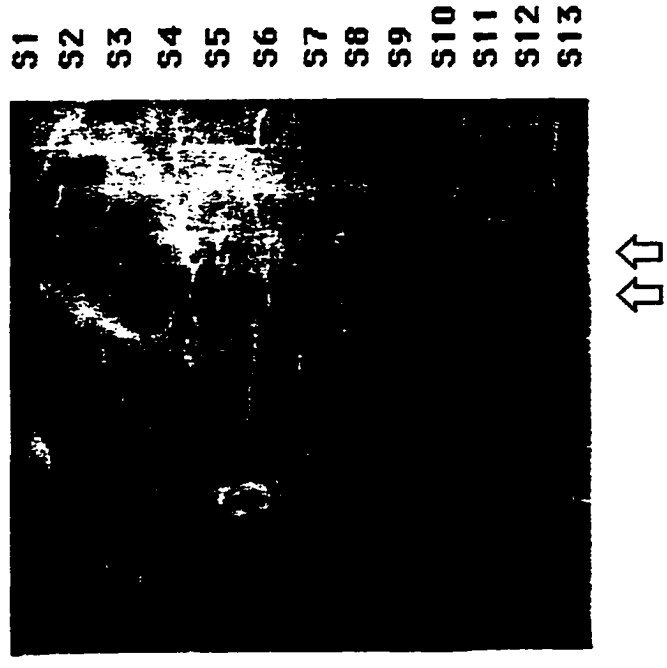


(B)

Figure 18. (A) Characterization of Polymerase Chain Reaction Products from Cell Line #130 Cloned into the Vector T7Blue(R). The DNA used in these experiments is from the high passage cell line #130, (passage number 131), the combination of primers used one SV40 primer 97D-2, which maps on the SV40 genome and the arbitrary primers 9-1, and 10-1, which should map in a region of the human DNA that flanks the virus. These set of clones have been given the nomenclature of S-1 through S-1 + n, (where n so far equals 1-12). The clones were run on an 0.8% agarose gel, stained with ethidium bromide. The marker used was lambda DNA digested with Hind III. Lane 1 is the marker, [left]. (B) Southern blot analysis was conducted using ^{32}P-labeled full length SV40 genomic DNA as a probe, the results show that all clones contained SV40, but clone S-2 produced an intense signal when compared to the other PCR products. The arrows indicate the location of the positive signals [right].



(A)



(B)

Figure 19. (A) Polymerase Chain Reaction Products using Complete Digestion of Genomic DNA. All of the samples were cut with the restriction enzyme Xba I, for a period of time that exceeded two hours. The samples were run on an 0.8% agarose gel, stained with ethidium bromide. The marker used was lambda DNA digested with Hind III. Lane 1 is the marker. (B) The results obtained from southern blot analysis indicate that the bands that positively hybridized to the ^{32}P-labeled full length SV40 genomic DNA probe contained viral sequences and were the size expected for the primers used in the reaction.

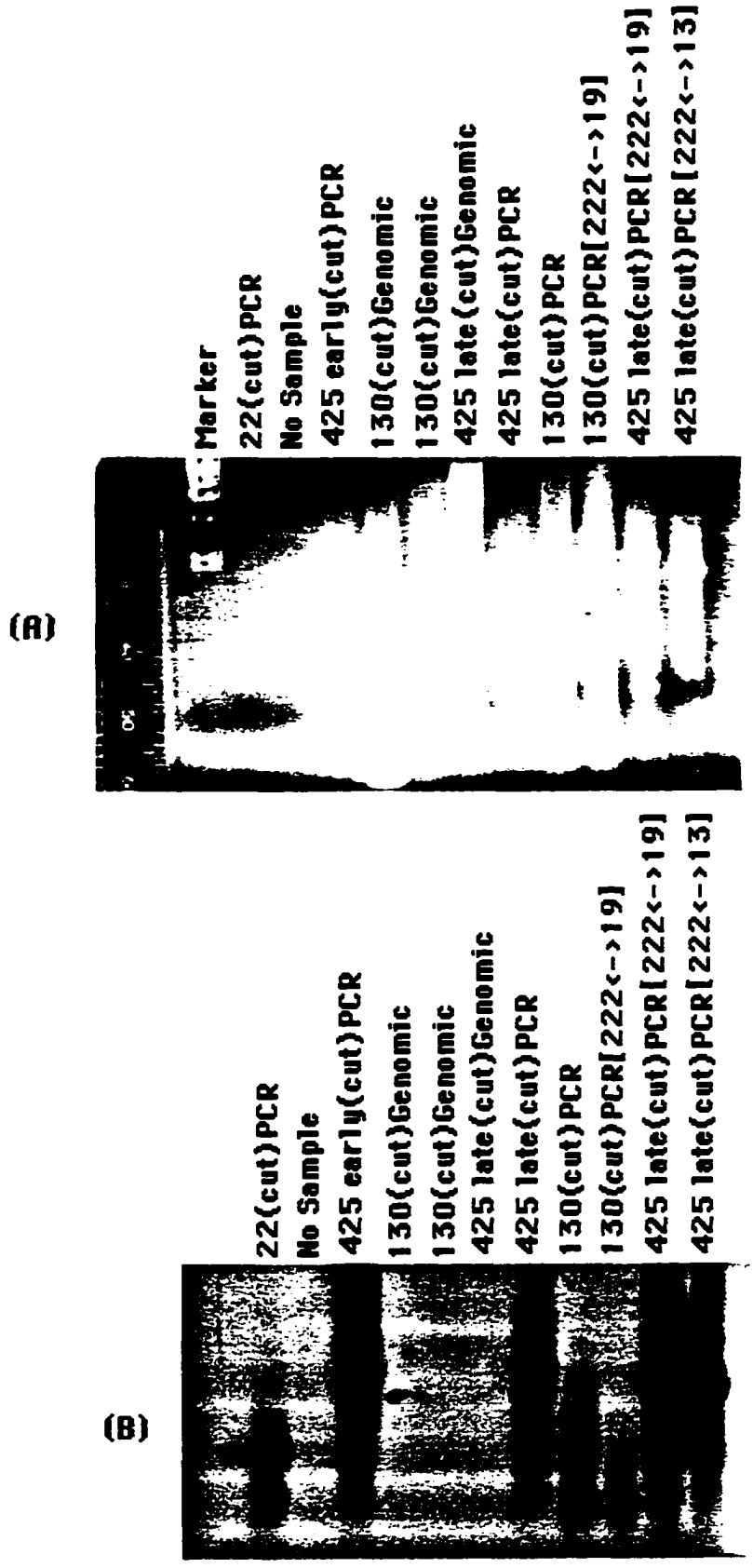


Figure 20 (A) Time Dependent Restriction Digest of DNA from Transformed Human Keratinocyte Cell Line #130. The digest was performed using the enzyme Xba I, aliquots were removed at different time intervals as indicated, 8 μL of each sample was then placed on an 0.8% agarose gel, stained with ethidium bromide. The marker used was lambda DNA digested with Hind III. The pattern of digestion is clearly defined in the range of least cut DNA at 15 minutes to a complete digest which is achieved at 240 minutes. (B) Southern blot analysis was conducted using ^{32}P-labeled full length SV40 genomic DNA as a probe. The results shown on the right were produced using a 24 hour exposure.

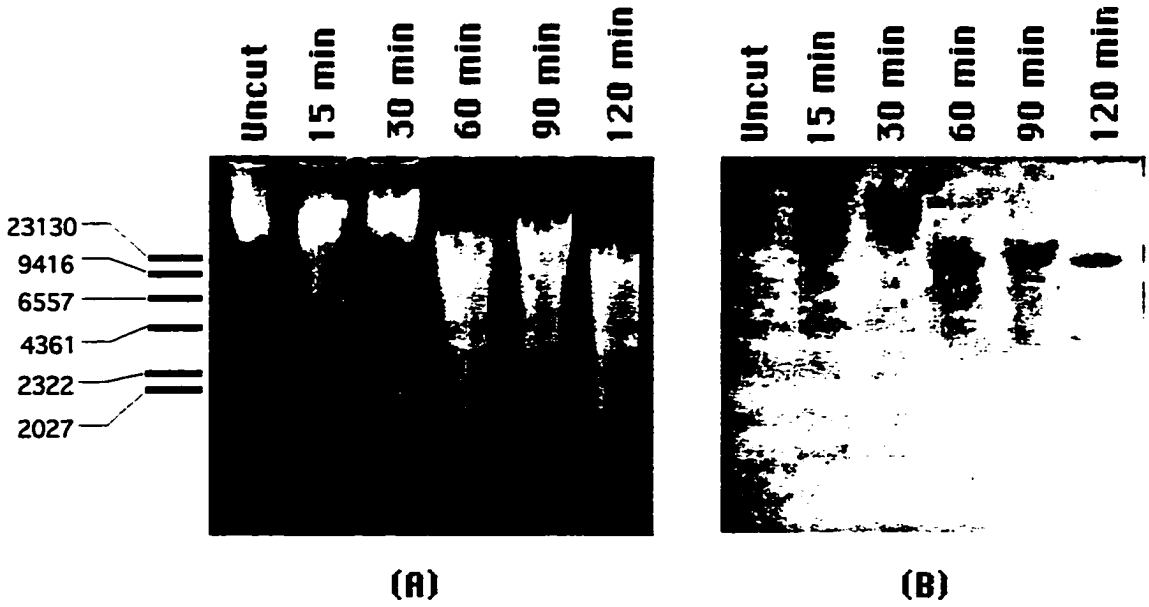


Figure 21. (A) PCR Products Obtained Using Time Dependent Digest of Genomic DNA. All samples shown are PCR products that were prepared by using different digestion times, starting at time zero to 240 minutes, sample were then placed on an 0.8% agarose gel and stained with ethidium bromide. The marker used was pBR322 DNA cleaved with Bgl I and Hinf I. The upper panel of the gel was obtained using primers 17 <-> 222, which should produce a PCR product that is 1728 kb in length. The lower panel of the gel was obtained using primers 13 <-> 222, which should produce a product that is 2418 bp in length. In both cases the results demonstrate that at time zero, (DNA not cut), no PCR products are obtained and that the best results are achieved at 90 minutes after the digestion began, [left]. (B) The results obtained from southern blot analysis indicate that the bands that positively hybridized to the ^{32}P-labeled full length SV40 genomic DNA probe contained viral sequences. The exposure time for the blot was four hours, [center]. (C) Is the same blot exposed for a period of 24 hours, [right].

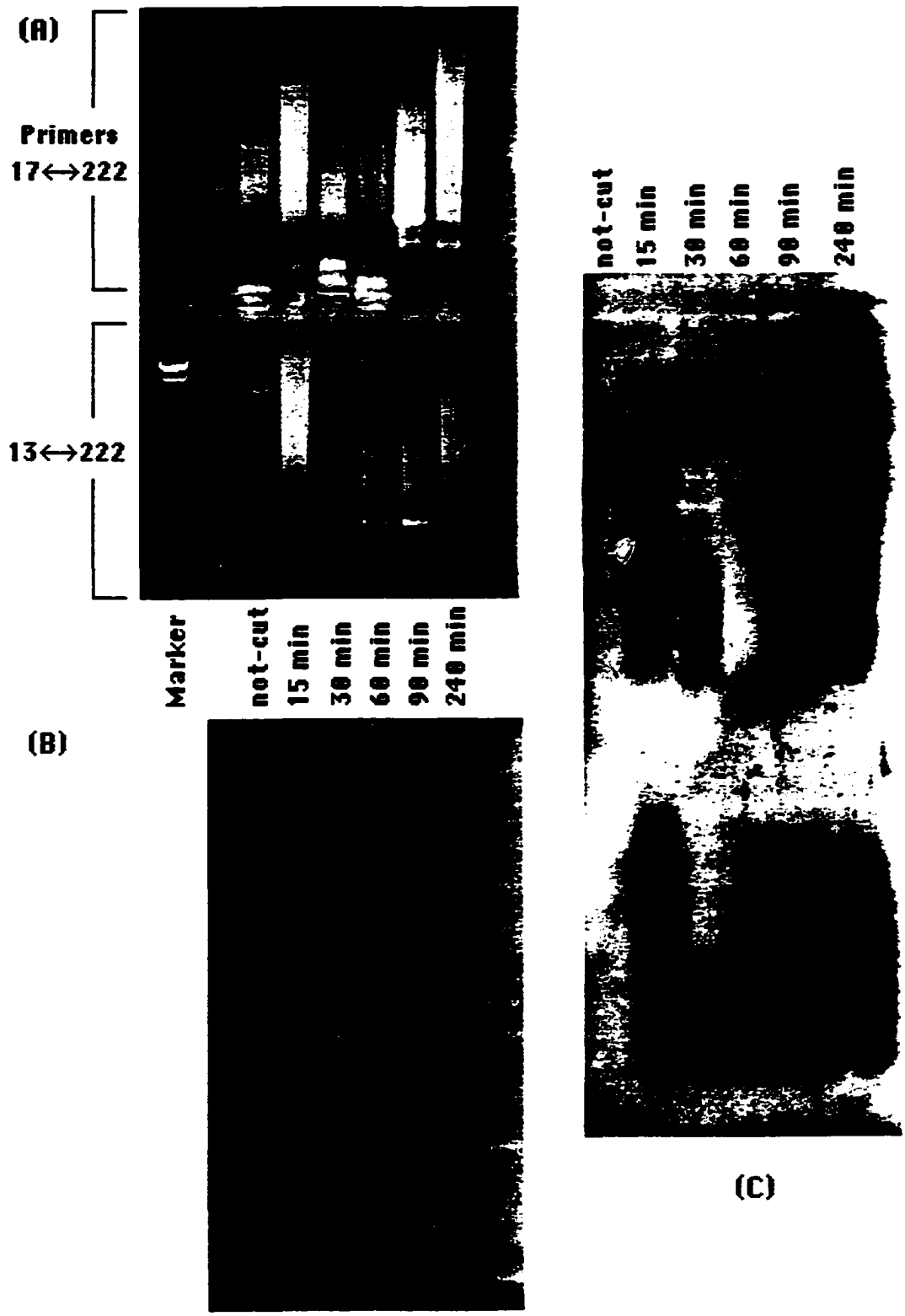


Figure 22. (A) PCR Products Obtained Using Time Dependent Digest of Genomic DNA. All samples shown are PCR products that were prepared by using different digestion times, starting at time zero to 240 minutes, sample were then placed on an 0.8% agarose gel and stained with ethidium bromide. The marker used was pBR322 DNA cleaved with Bgl I and Hinf I. The samples were obtained using primers 19 <-> 222, which should produce a PCR product that is 1492 kb in length. The results coincide with the data obtained in figure 21, at time zero, (uncut DNA), no PCR products are obtained and that the best results are achieved at 90 minutes after the digestion began, [left]. (B) Southern blot analysis indicate that the bands that positively hybridized to the ^{32} P-labeled full length SV40 genomic DNA probe contained viral sequences. The exposure time for the blot was three hours, [right].

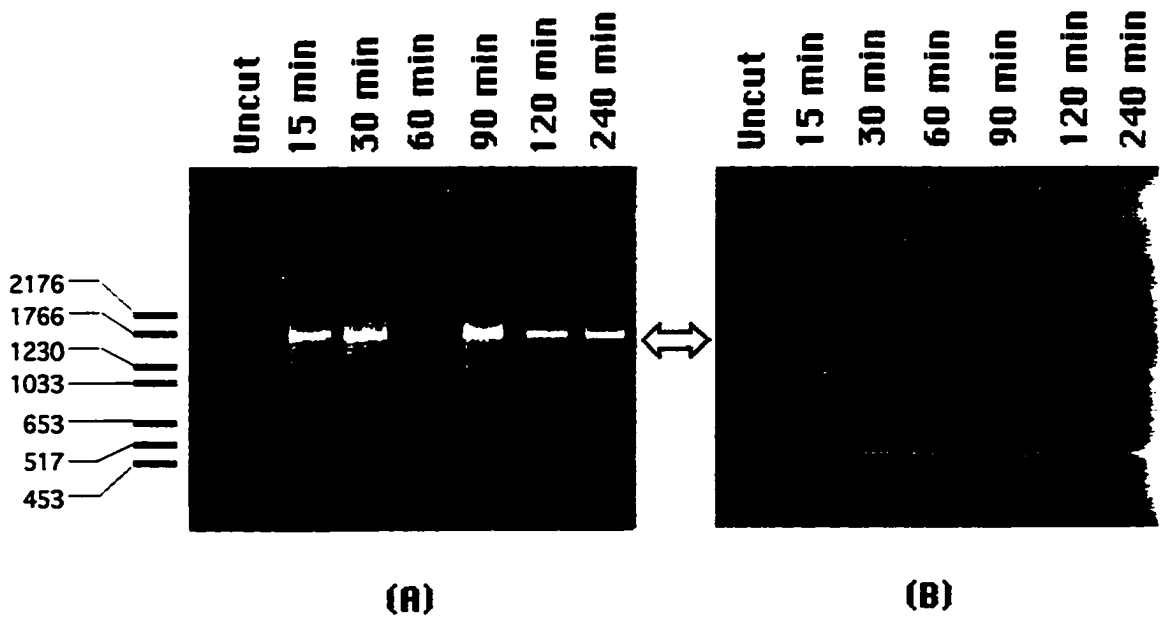


Figure 23. Time Dependent Restriction Digest of DNA from Transformed Human Keratinocyte Cell Line #130, Using the Restriction Enzymes Sac I and BamH I. All samples shown are PCR products that were prepared by using different digestion times, starting at time zero to 240 minutes, sample were then placed on an 0.8% agarose gel and stained with ethidium bromide. The marker used was pBR322 DNA cleaved with Bgl I and Hinf I. All of the samples were obtained using primers 19 <-> 222, which produces a PCR product that is 1491 kb in length. (A) Digest using the restriction enzyme Sac I, which is a no-cut enzyme for the SV40 genome, Lane 2 through 8 are samples run using the time intervals described above, [upper left]. (B) Southern blot analysis was conducted using ^{32} P-labeled full length SV40 genomic DNA as a probe, with an exposure time of 24 hours, [upper center]. (C) The results were produced using a 48 hour exposure, [upper right]. (D) Digest using BamH I, which is a one-cut enzyme for the SV40 genome, [lower left]. (E) Southern blot analysis of the BamH I digest using a 24 hour exposure time, [lower center]. (F) Southern blot analysis of the BamH I digest using a 48 hour exposure time, [lower right].

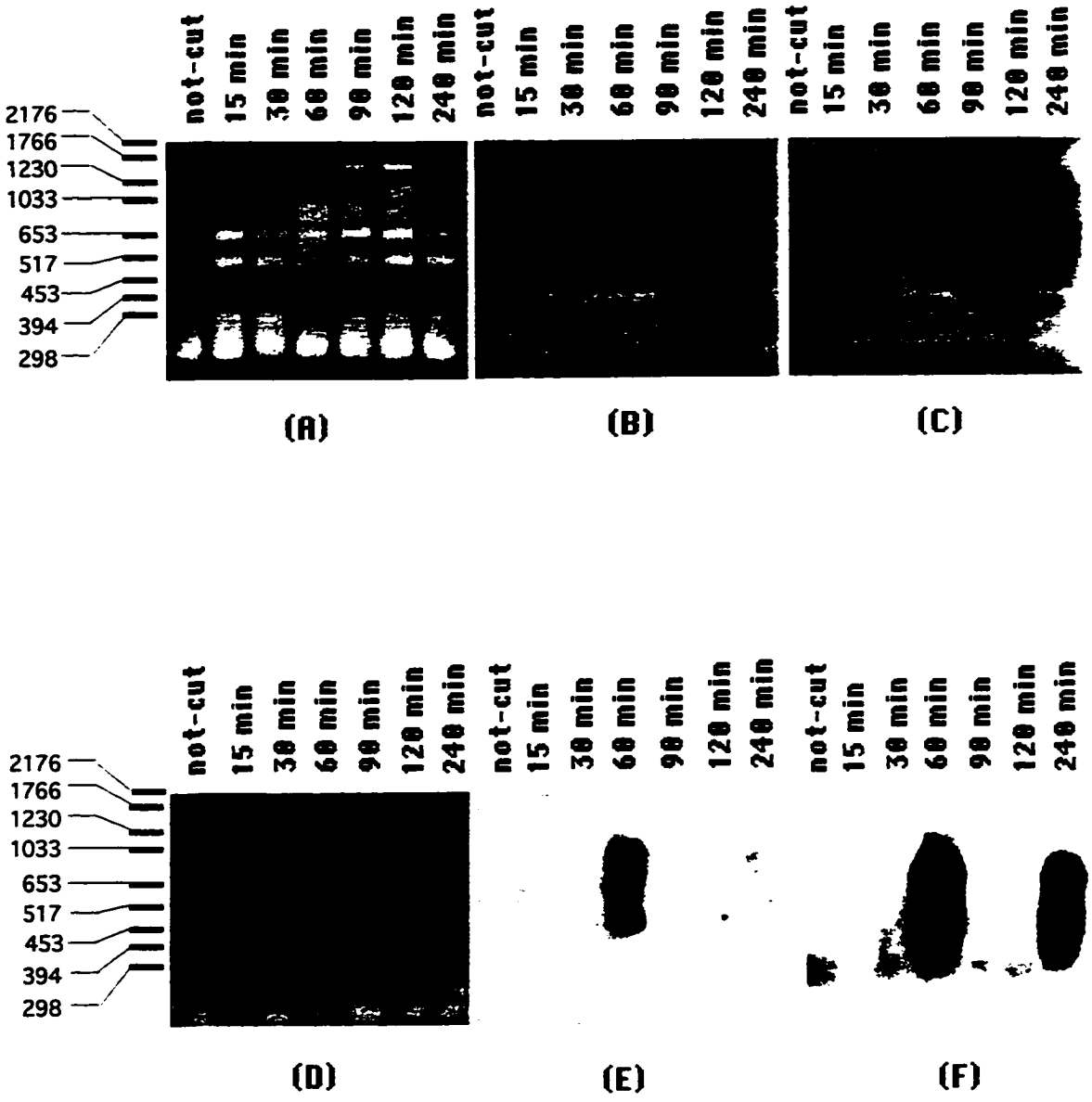


Figure 24 (A) PCR products obtained using Xba I digested genomic DNA from cell line 98. All samples shown are PCR products using different digestion times up to 240 minutes. Samples were run on an 0.8% ethidium bromide agarose gel. The marker used was pBR322 DNA cleaved with Bgl I and Hinf I. Two PCR products are observed at 90 minutes. (B) Southern blot hybridization using a ^{32}P-labeled full length SV40 genomic DNA probe. The exposure time for the blot was 24 hours.

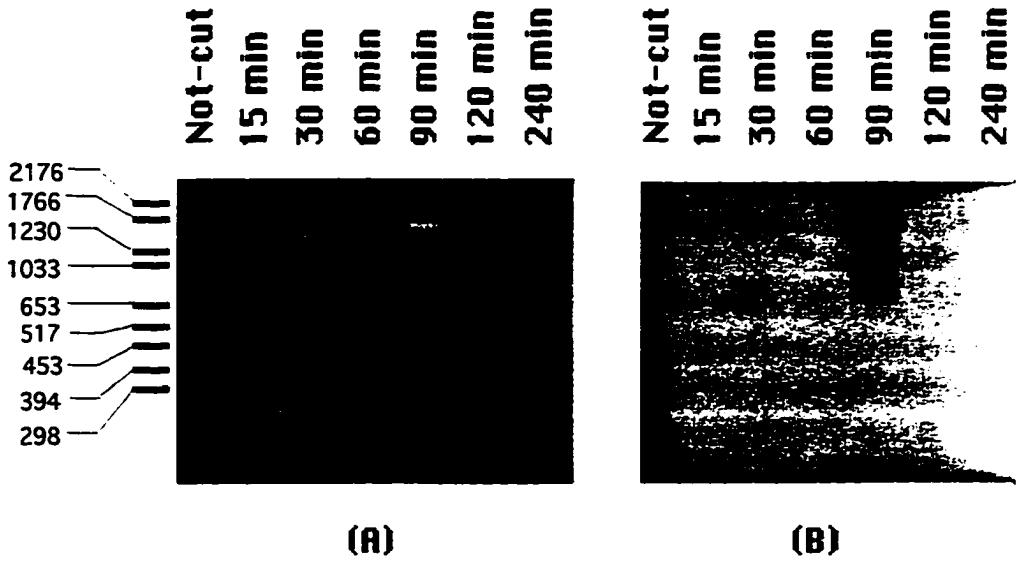


Figure 25 (A) Samples of DNA from Cell Lines 98 and 425 used During Sequence Analysis. PCR products from cell line 98 using lambda Gem-11 clones as template. Samples were loaded in to an 0.8% agarose gel and stained with ethidium bromide. The samples were obtained using primer combinations 19 <-> 222, upper tier, and 17 <-> 222, second and third rows. (B) Southern blot analysis indicate that the bands that positively hybridized to the P-labeled full length SV40 genomic DNA probe contained the desired viral sequences. The exposure time for the blot was 24 hours. (C) Agarose gel analysis of cell line 425 cloned into the vector sequencing vector pGem-3Z, the marker used was lambda DNA digested with Hind III. 3-1 and 5-1 represent individual isolates.

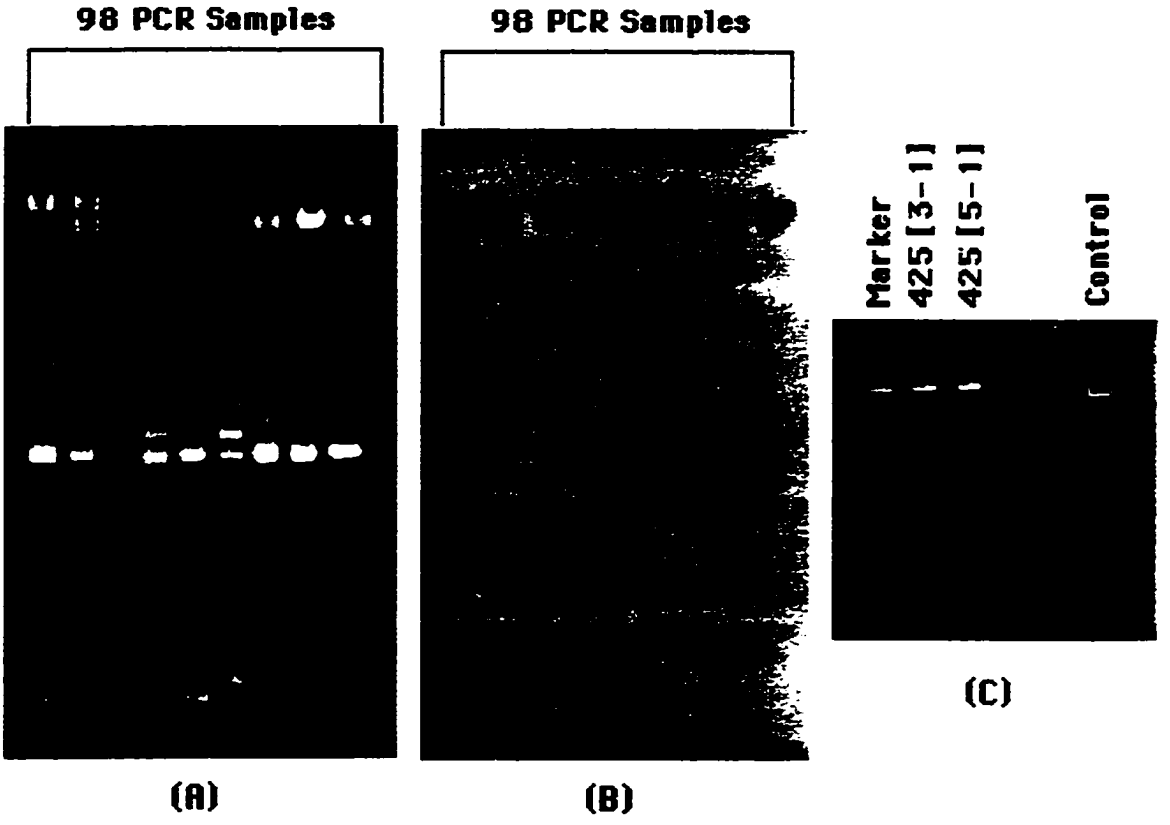
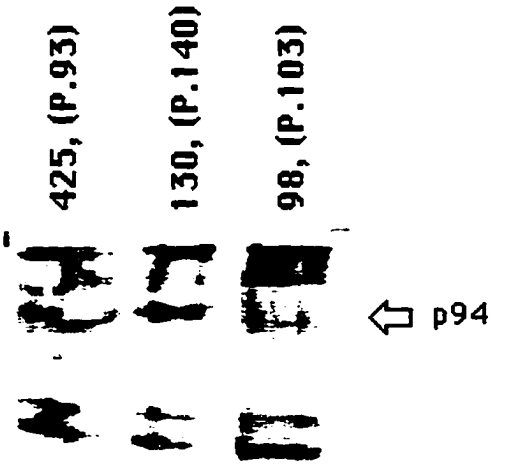
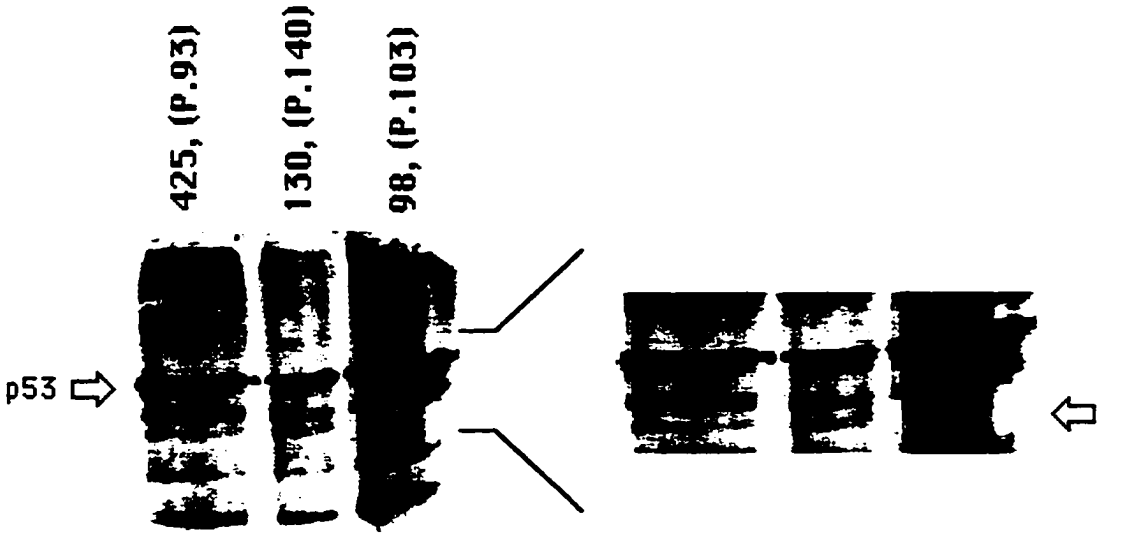


Figure 26. Immunoprecipitation and Western Blot Analysis of SV40 T-antigen, (p94), and Protein 53, (p53). [A] Immunoprecipitation of large T-antigen followed by western blot analysis using monoclonal antibody (Ab-2). Arrow indicates the location of p94. (B) Immunoprecipitation of large T-antigen followed by western blot analysis using monoclonal antibody (p53). Arrow on left indicates the location of p53. Arrow on right marks missing band.

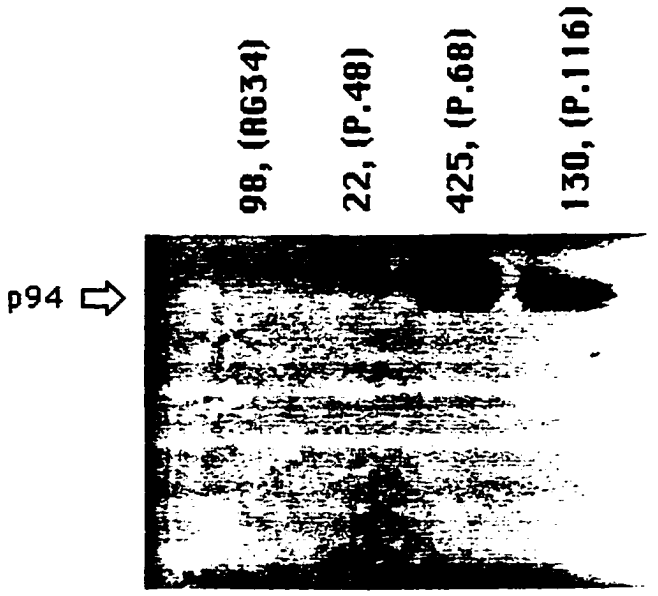


(A)

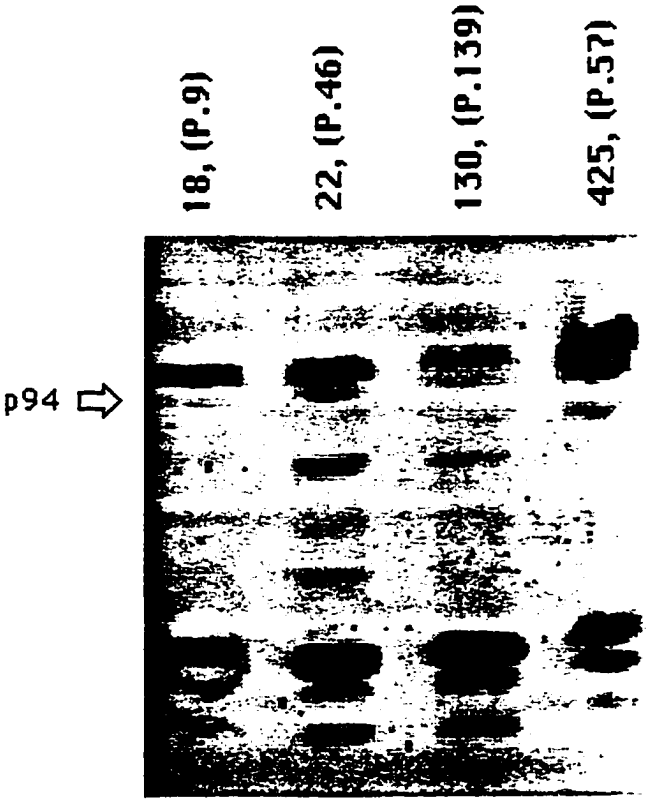


(B)

Figure 27. Western blot of SV40 transformed cell lines at different serial passages. Samples on both gels (A) and (B) were run on a SDS-stacking/running gel combination, and transferred on to nitrocellulose. Arrows indicates the location of p94.



(A)



(B)

Figure 28. Compilation of all Mutations Found in the SV40 Stock and Cell Lines Transformed by the Virus. The viral genome is divided into functional domains represented by the letters, A through F, as follows. A= T-antigen, B= T-antigen intron, C= t,T-antigen, D= T-antigen binding site, E= 72 bp palindromic enhancer, F= Late genes. SVI = the SV40 input virus. Cell lines are represented by the number, 98, 130, and 425. Cell line 425 is subdivided by its passage, 425E is passage 12, 425 CDNA is passage 46 and 425L is passage 97. The Symbol * represents a change in original amino acid.

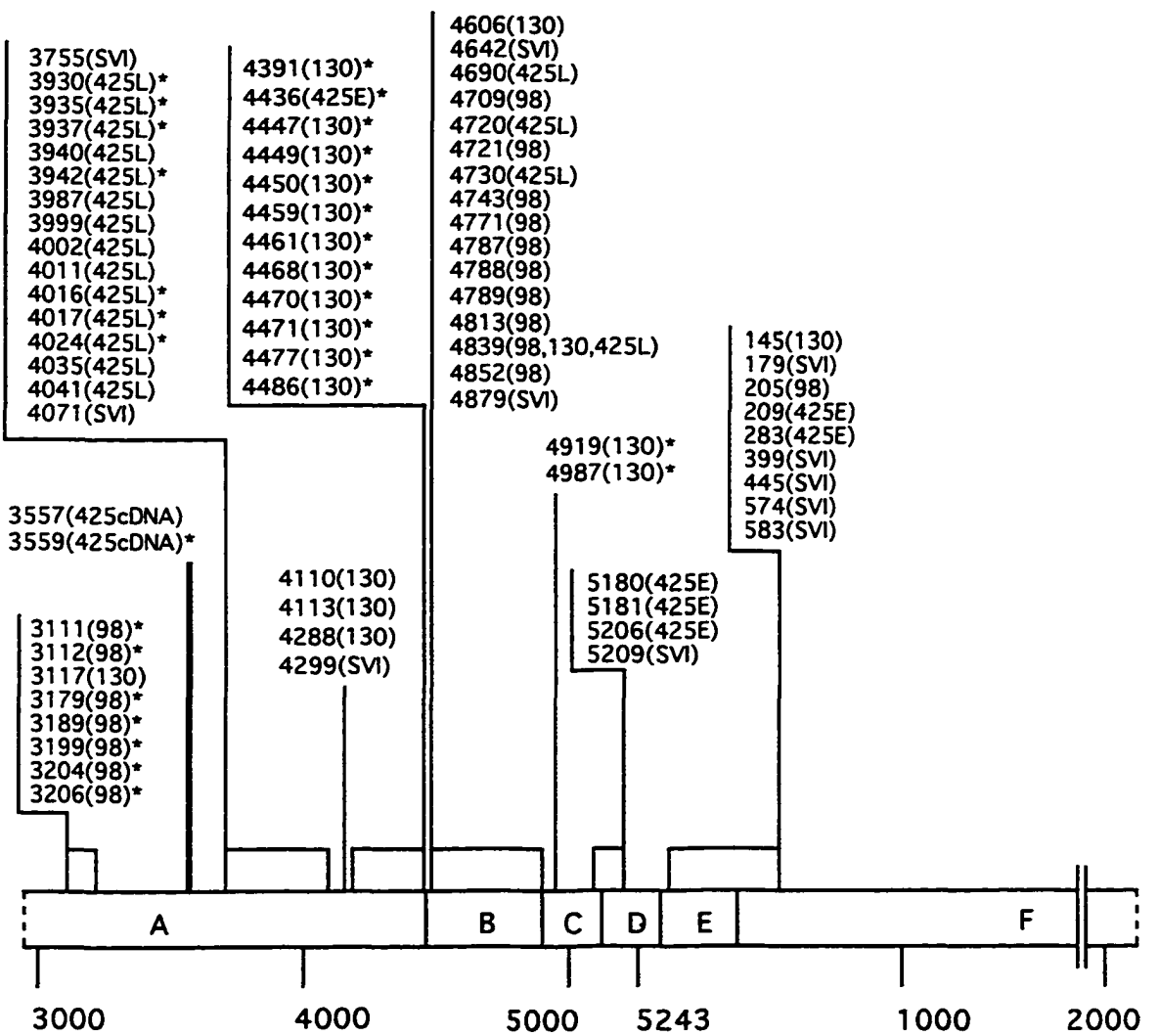
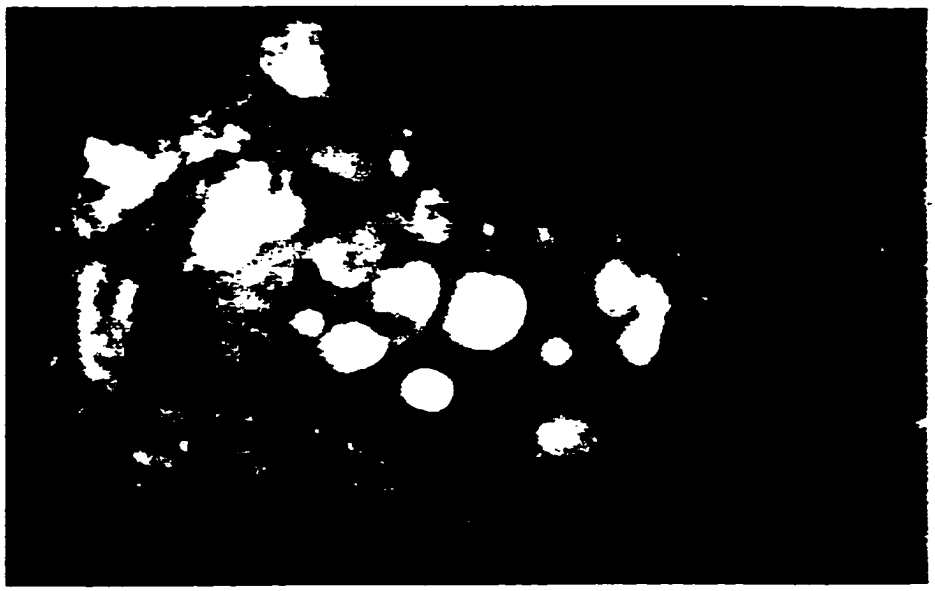
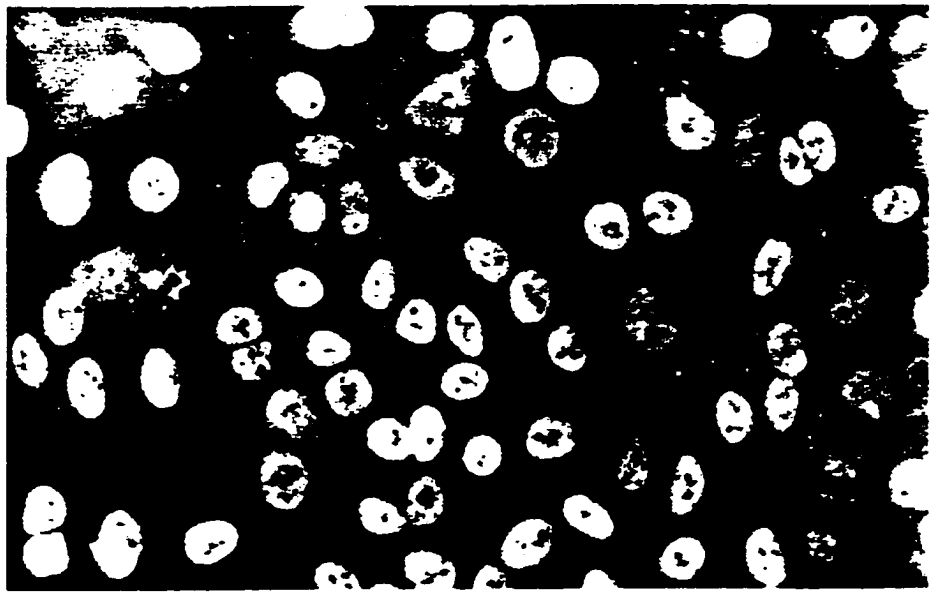


Figure 29. Immunofluorescent staining of SV40 T-antigen in a colony of viral-infected human keratinocytes of cell line 98. The immunofluorescent staining was performed as described in materials and methods. (A) At the third passage post infection. (B) At the tenth serial passage as cells enter a period of "crisis".



(A)



(B)

Figure 30. Serum Dependent Growth Analysis of Cell Line 425. Cells were plated at an initial concentration of 1.25 X 10^5 and passaged for a period of seven days. Symbols representing the different concentrations of serum are indicated below.

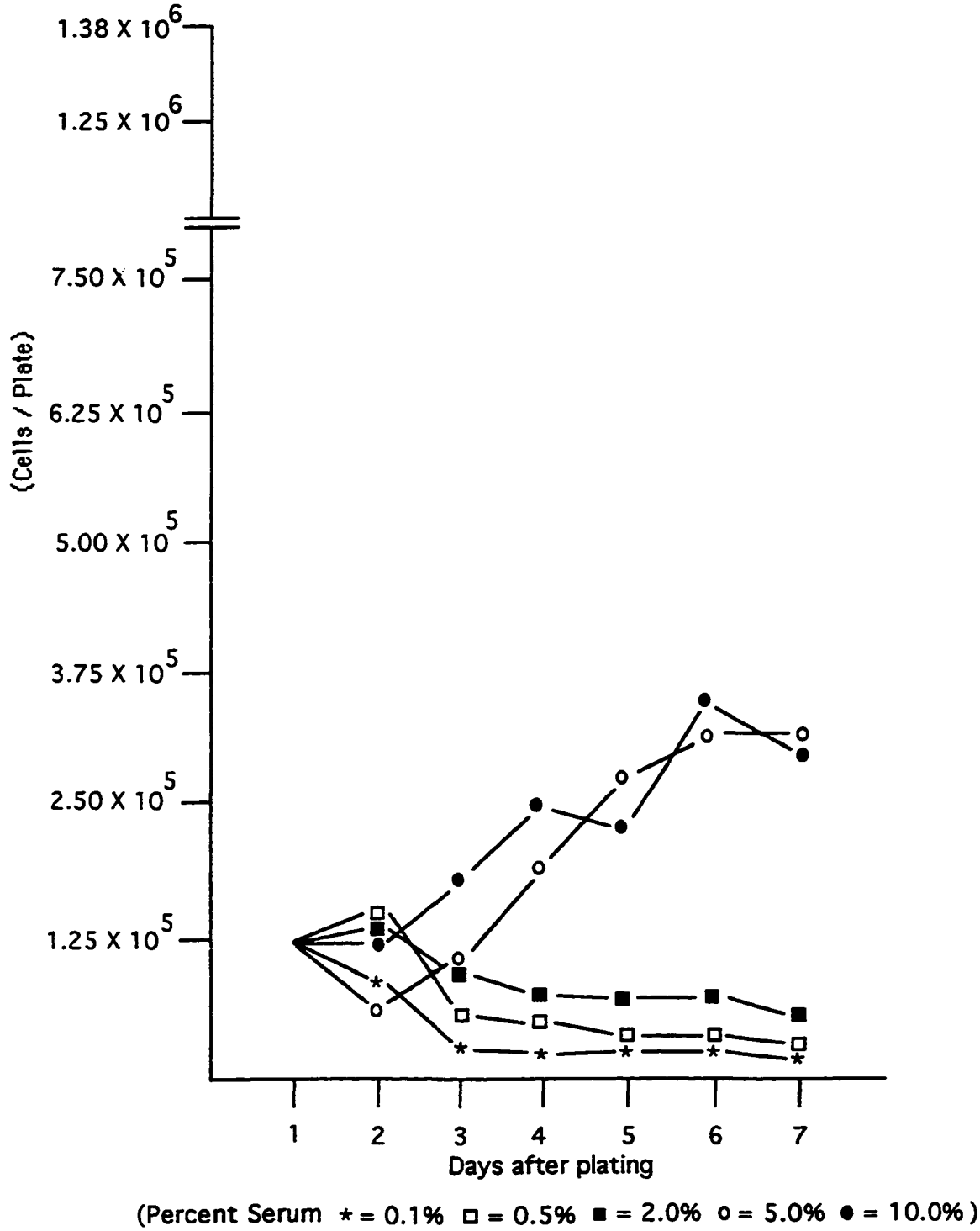


Figure 31. Serum Dependent Growth Analysis of Cell Line 130. Cells were plated at an initial concentration of 1.25 X 10^5 and passaged for a period of seven days. Symbols representing the different concentrations of serum are indicated below.

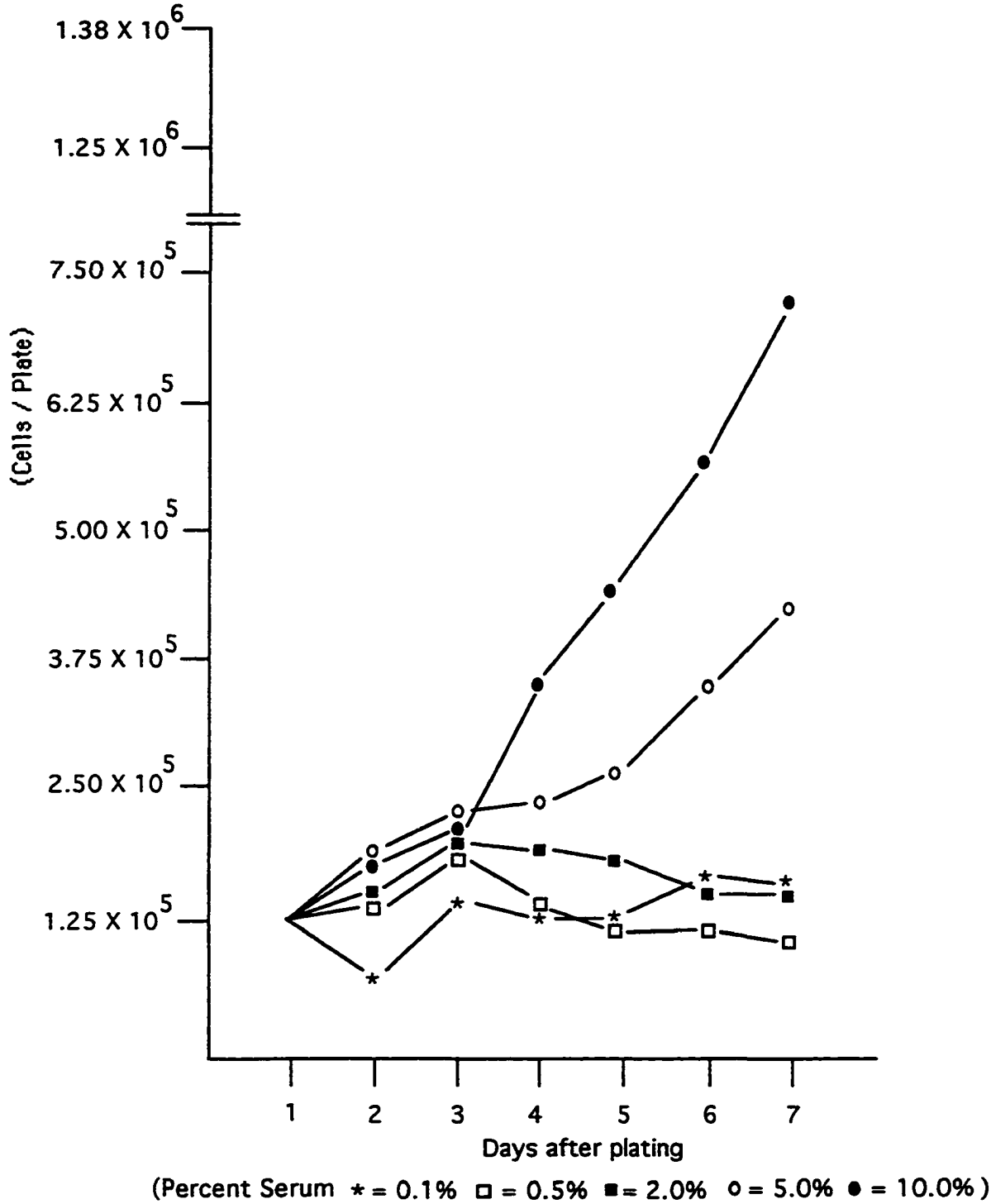


Figure 32. Serum Dependent Growth Analysis of Cell Line 98. Cells were plated at an initial concentration of 1.25×10^{5} and passaged for a period of seven days. Symbols representing the different concentrations of serum are indicated below.

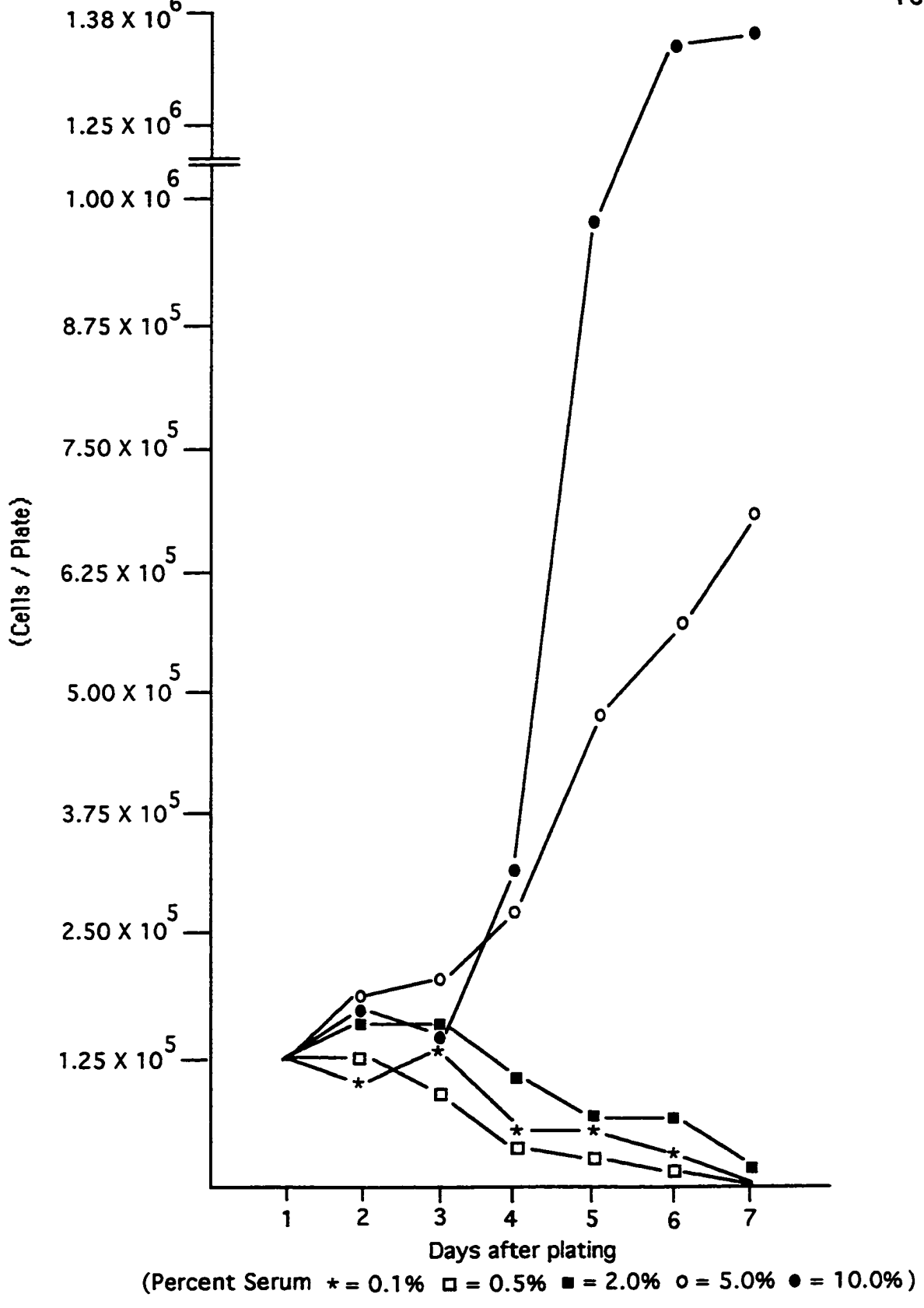
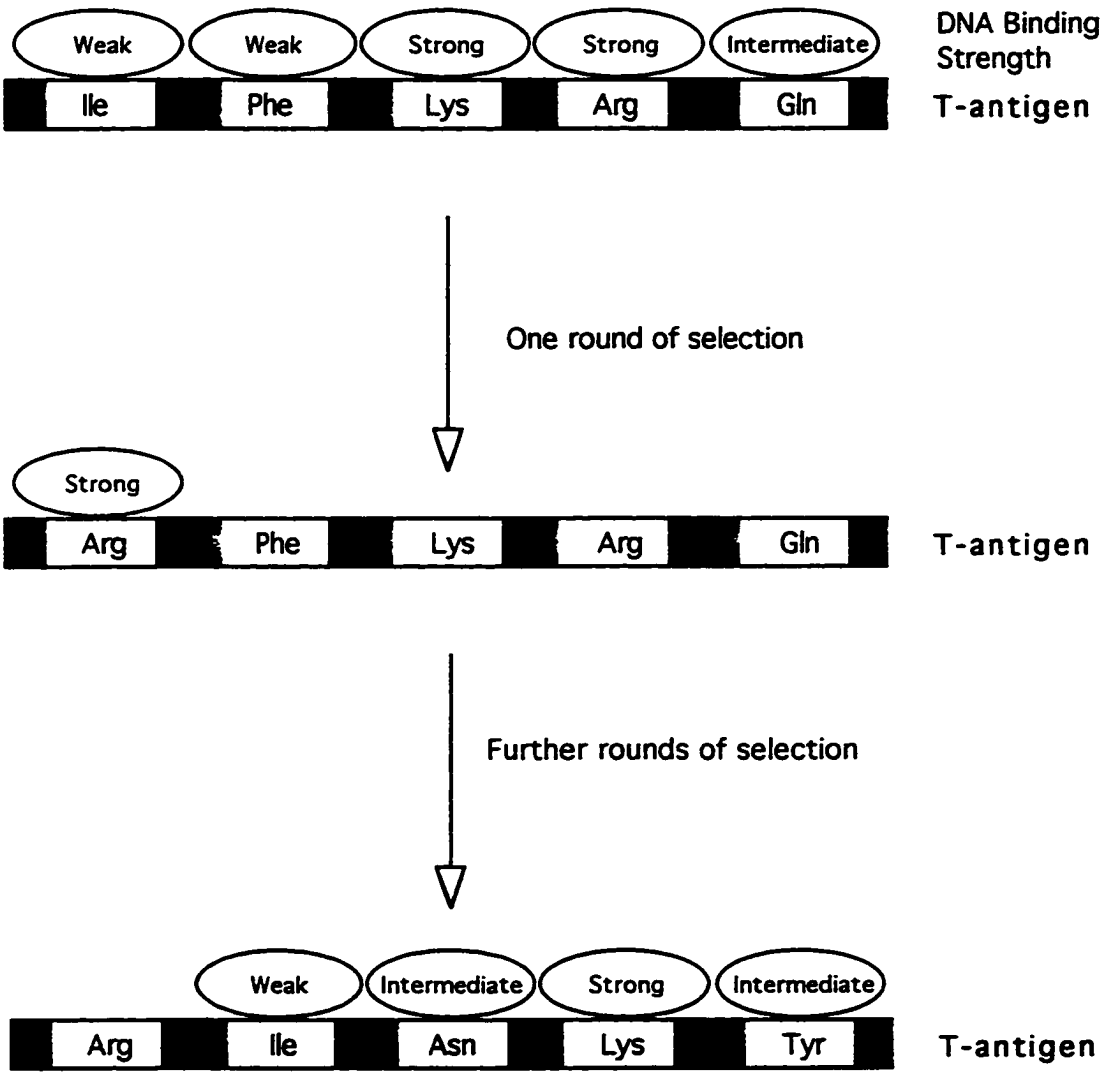


Figure 33. Mechanism by which the DNA binding affinity of the mutated large T-antigen product that has undergone a number of serial passages may return to a functional state that is similar to the unmutated protein.



Bibliography:

1- Sweet, B.H. and Hilleman, M.R. 1960. Proc. Soc. Exp. Biol. Med. 105: 420.

2- Eddy, B.E., et al. 1962. Virology 17: 65.

3- Tooze, J., ed. 1980. Molecular Biology of Tumor Viruses, 2nd edition. DNA Tumor Viruses, part 2. Cold Spring Harbor Laboratory, New York.

4- Weinberg, R.A., ed 1989. Oncogenes and the Molecular Origins of Cancer,
Cold Spring Harbor Laboratory Press, New York.

5- Klug, A.J., 1965. J. Mol. Biol. 11: 424.

6- Finch, J.T., et al. 1975. Structure of Small DNA Containing Animal Viruese. Comprehensive Virology 5: 119-131. Plenum Press, New York.

7- Botchan, M., et al. 1986. Cancer Cells. DNA Tumor Viruses 4: 1-16. Cold Spring Harbor Laboratory Press, New York.

8- DePamphilis, M.L. and Wasserman, P.M., 1983. Organization and Replication of Papovavirus DNA, in: Organization and Replication of Viral DNA (Kaplan, A.S. ed.), 37-114, CRC Press, Boca Raton, Fla.

9- Ng, S.C., et al. 1985. J. Biol. Chem. 260: 1127.

10- Hay, N., et al. 1982. Cell 29: 183.

11- Privies, C., et al. 1979. J. Virol. 29:1204.

12- Resnick, J. and Shenk, T., 1986. J. Virol. 60(3):1098

13- Carswell, S., et al. 1986. J. Virol. 60(3):1055

14- Hay, N., et al. 1985. Mol. Cell Biol. 5(6):1327

15- Ellman, M., et al. 1984. J. Virol. 50: 623

16- Bikel, H., et al. 1986. Cancer Cells. DNA Tumor Viruses 4: 405. Cold Spring Harbor Laboratory Press, New York.

17- Rubin, H.J., et al. 1982. Cell 30: 469

18- Sompayrac, L. 1983. J. Virol. 46: 620

19- Fanning, E., et al. 1992. Annu. Rev. Biochem. 61: 55

20- Zerrahn J., et al. 1996. J. Virol. 70: 6781

21- Sambrook, J., et al. 1968. Proc. Natl. Acad. Sci. USA., 60: 1288

22- Gelb, L.D., et al. 1971. J. Mol. Biol. 57: 129.

23- Hirai, K., et al. 1972. J. Virol. 9: 705.

24- Holzel, F., et al. 1974. J. Mol. Biol. 84: 423

25- Brugge, J.S., et al. 1975. J. Virol. 15: 619.

26- Steinberg, M.L., et al. 1979. Proc. Natl. Acad. Sci. USA., 76: 801-805

27- Steinberg, M.L., et al. 1993. J. Invest. Dermatol. 81: 131s

28- Steinberg, M. Immortalization of Human Epidermal Keratinocytes by SV40. In Culture of Immortalized Cells. Chapter 5, pp. 95-120, Ian Freshney (Ed.), Wiley-Liss, New York, 1996

29- Schwyzer, M., et al. 1980. J. Biol. Chem. 255: 5627

30- Simmons, D.T., et al. 1988. Proc. Natl. Acad. Sci. USA., 85: 2086

31- Arthur, A.K., et al. 1988. J. Virol. 62: 1999

32- Strauss, M., et al. 1987. J. Virol. 61: 3326

33- Grässer, F.A., et al. 1987. J. Virol. 61: 3373

34- Höss, A., et al. 1990. J. Virol. 64: 6291

35- Tijan, R., 1978. Cell 13: 165-179.

36- Tegtmeyer, P., et al. 1981. J. Virol. 46: 151-161.

37- Murakami Y., et al. 1993. J. Biol. Chem. 268: 11018-11027

38- Rio, D., et al. 1980. Proc. Nat. Acad. Sci. USA 77: 5706-5710.

39- Hanson, U., et al. 1981. Cell 27: 603-612.

40- DiMaio, D., et al. 1982. J. Mol. Biol. 156: 531-548

41- Gilinger, G., et al. 1993. J. Virol. 67(11): 6682-6688

42- Deb. S., et al. 1986. Mol. Cell Biol. 6: 1663-1680.

43- Deb.S., et al. 1987. J. Virol. 61: 2143-2149.

44- Li, J., et al. 1985. Mol. Cell. Biol. 5: 1238.

45- Li, J., et al. 1986. Mol. Cell. Biol. 6: 1117-1128.

46- Boroweic, J.A., et al. 1988. Proc. Nat. Acad. Sci. USA. 85: 64-68.

47- Parsons, R., et al. 1991. J. Virol. 65: 2798-2806.

48- Loeber, G., et al. 1991. J. Virol. 65: 3167-3174.

49- Mastrangelo, I.A., 1989. Nature 338: 658-682.

50- Wold, M.S., et al. 1987. Proc. Nat. Acad. Sci. USA. 84: 3643-3647.

51- Parsons, R., et al. 1990. J. Virol. 64: 509-518.

52- Weikowski, M., et al. 1988. J. Biol. Chem. 263: 436-442.

53- Clertant, P., et al. 1984. J. Biol. Chem. 259: 15196-15203.

54- Höss, A., et al. 1990. J. Virol. 64: 6291-6296.

55- DeCaprio, J.A., et al. 1988. Cell 54: 275-283.

56- Schmieg, F.I., 1988. Virology 164: 132-140.

57- Clark, R., et al. 1983. Mol. Cell Biol. 3: 220-228.

58- Bullock, P., et al. 1990. J. Virol. 64: 6234-6245.

59- Taljanidisz, J., et al. 1987. Nucleic Acids Res. 19: 7877-7888.

60- Schutzbank, T., et al. 1982. Cell 30: 481-490.

61- Soprano, K., et al. 1983. Mol. Cell Biol. 3: 214-219.

62- Taylor, I., et al. 1989. J. Biol. Chem. 264: 16160-16164.

63- Wagner, S., et al. 1990. Oncogene 5: 353-359.

64- Wagner, S., et al. 1991. J. Virol. 65: 3259-3267.

65- Breathnach, R., et al. 1981. Annu. Rev. Biochem. 50: 349-383.

66- Myers, R., et al. 1981 Cell 25: 373-384.

67- Michael, P.J., et al. 1987. Cell 50: 847-861.

68- Das, G., et al. 1985. J. Mol. Biol. 182: 229-239.

69- Fromm, M., et al. 1983. Mol. Cell Biol. 3: 991.

70- Mattaj, I., et al. 1985. Nature 316: 163.

71- Gruda, M.C., 1991. J. Virol. 65: 3553-3558.

72- Alwine, J.C. 1985. Mol. Cell Biol. 5: 1034-1042.

73- Spence, S.L., et al. 1994. J. Virol. 68: 4227-4240.

74- Scheidtman, K.H., et al. 1991. J. Virol. 65: 2089-2101.

75- Cann, A.J. 1997. Principles of Molecular Virology. 2nd Edition. Academic Press Inc. California, USA.

76- Schnieider, J. et al. 1988. J. Virol. 62: 1598-1605.

77- Virshup, D., et al. 1989. EMBO J. 8: 3891-3898.

78- McVay, D., et al. 1989. Nature 341: 503-507.

79- Howe, J.A., et al. 1990. Proc. Nat. Acad. Sci. USA. 87: 5883-5887.

80- Sager, R. 1989. Science 246: 1406-1412.

81- DeCaprio, J., et al. 1989. Cell 58: 1085-1095.

82- Dyson, N., et al. 1989. Science 243: 934-937.

83- Dyson, N., et al. 1989. J. Virol. 64: 1353-1356.

84- Whyte, P., et al. 1988. Nature 334: 124-129.

85- Takahashi, R., et al. 1991. Proc. Nat. Acad. Sci. USA. 88: 5257-5261.

86- Banerjee, A., et al. 1992. Cancer Res. 52: 6297-6304.

87- Ludlow, J., et al. 1989. Cell 56: 57-65.

88- DeCaprio, J., et al. 1992. Proc. Nat. Acad. Sci. USA. 89: 1795-1798.

89- Furukawa, Y., et al. 1990. Proc. Nat. Acad. Sci. USA. 87: 2770-2774.

90- Helin, K., et al. 1992. Cell 70: 337-350.

91- Kaelin, W., et al. 1992. Cell 70: 351-364.

92- Lane, D.P., et al. 1979. Nature 278: 261-263.

93- Eliyahu, D., et al. 1984. Nature 312: 646-649.

94- Parada, L.F., et al 1984. Nature 312: 649-651.

95- Jenkins, J.R., et al. 1984. Nature 312: 651-654.

96- Hinds, P.W., et al. 1989 J. Virol. 63: 739-746.

97- Wang, E.H., et al. 1989. Cell 57: 379-392.

98- Baker, S.J., et al. 1989. Science 244: 217.

99- Paulovich, A.G., et al. 1997. Cell. 88: 315-321.

100- Watson et al. 1988. Molecular Biology of the Cell. Garland Publishing Inc.

101- Parde, A.B., et al. 1978. Annu. Rev. Biochem. 47: 715-750.

102- Parde, A.B. 1974. Prod. Natl. Acad. Sci. U.S.A. 71: 1286-1290.

103- Soussi, T., et al. 1990. Oncogene 5, 945-952.

104- Oren, M. et al. 1990. Science 250, 113-116.

105- Truant, R., et al. 1993. J. Biol. Chem. 268: 2284

106- Martin, D.W., et al. 1993. J. Biol. Chem. 268: 13062

107- Levine, A.J., 1997. Cell. 88: 323

108- Prives, C., et al. 1992. Nature 358: 83

109- Vogelstein, B., 1992. Science 256: 827

110- Levine, A.J., et al. 1992. Genes & Development 6: 1143

111- Michieli, P., et al. 1994. Cancer Research 54: 3391

112- El-Deiry, W.S., et al. 1994. Cancer Research 54: 1169

113- Takahashi, T., et al. 1990. Prod. Natl. Acad. Sci. U.S.A. 87: 5863

114- Vogelstein, B. 1990. Nature 348: 681

115- Levine, A. J., et al. 1994. J. Lab. Clin. Med. 123-6: 817

116- Ludlow, J., et al. 1993. FASEB J. 7: 866

117- Soussi, T., et al. 1990. Oncogene 5: 945

118- Oren, M., et al. 1990. Science 250: 113

119- Chin, K., et al. 1992. Science 255: 459

120- Lane, D.P., et al. 1987. Nature 329: 456

121- Howley, P.M. et al. 1990. Cell 63: 1129.

122- Bargonetti, J. et al. 1991. Cell 65: 1083

123- Ishii, S., et al. 1992. Proc. Natl. Acad. Sci. USA 89: 5403

124- Weintraub, H., et al. 1991. Proc. Natl. Acad. Sci. USA 88: 4570

125- Kern, S.E. et al. 1991. Science 252: 1708

126- Koprowski, H., et al. 1962. J. Cell Comp. Physiol. 59: 281

127- Defendi, V., Naimski, P., and Steinberg, M.L., 1982. J. Cell Physiol. [Suppl] 2:131.

128- Steinberg, M.L., et al. 1981. J. Cell Physiol. 109: 153-159.

129- Steinberg, M.L., et al. 1982. Exp. Cell Res. 139: 369-375.

130- Edelman, B., Steinberg, M.L., and Defendi, V. 1985. Int. J. Cancer
35(2): 219-225.

131- Steinberg, M.L., et al. 1986. Processes in Cutaneous Epidermal Differentiation, Bernstein, I.A. and Hirone, T. ed. Praeger Scientific, New York, N.Y. 287-299.

132- Okada, N., Steinberg, M.L. and Defendi, V. 1984. Exp. Cell Res., 153:198-207.

133- Hronis, T.S., Steinberg, M.L., Defendi, V. and Sun T.-T. 1984. Cancer Res., 44:5797-5804.

134- Morris, A., Steinberg, M.L. and Defendi, V. 1985. Proc. Natl. Acad. Sci., 82: 8498-8502.

135- Cooper, D., et al. 1985. Lab. Inv., 52: 243-256.

136- Nelson, W.G., et al. 1984. Cancer Res., 44: 1600-1603.

137- Morris, A.E., Steinberg, M.L., Li, J.-Y. and Defendi, V. 1989. Exp. Cell Res. 182: 461-472.

138- Stieber, P., et al. 1993. Bodenmueller, H., et al. 1993. Clin. Biochem. 26(4): 301-304.

139- Takada,M., et al. 1995. Br. J. Cancer 71(1): 160-165.

140- Macpherson, I., et al. 1964. Virology. 24: 291

141- Steinberg, M.L., et al. 1982. Exp. Cell. Res. 139: 369

142- Steinberg, M.L., et al. 1985. J. Cellular Biochem. 29: 95

143- Chen, G., et al. 1996. Gene 183: 41

144- Frischauf, A.M., et al. 1983. J. Mol. Biol. 170, 827.

145- Zabarovsky, E.R. and Allikmets, R.L., 1986. Gene 42, 119.

146- Chen, E.Y. and Seeburg, P.H., 1985. DNA 4, 165-170.

147- Hattori, M. and Sakaki, Y., 1986. Anal. Biochem. 152, 232.

148- Mierendorf, R. and Pfeffer, D., 1987. Methods in Enzymology 152, 556-562, Academic Press, New York.

149- Ausubel, F.M., et al. (eds) 1989. Current Protocols in Molecular Biology, John Wiley & Sons, New York.

150- Gannon, J.V., et al. 1987. Nature 329: 456

151- Dugas, H., et al. 1981. Bioorganic Chemistry, Springer-Verlag, New York.

152- Tiemann, F., et al. 1994. Oncogene 9: 1907

153- Hurwitz, J., et al. 1990. J. Biol. Chem. 265: 18043

154- Fairman, M.P., et al. 1988. EMBO J. 7: 1211

155- Wold, M.S., et al. 1989. J. Biol. Chem. 264: 2801

156- Parvin, J.D., et al. 1986. J. Virol. 59: 377

157- Smith, D.B., et al. 1987. J. Gen. Virol. 1987. 68: 2729

158- Ward, C.D., et al. 1988. J. Virol. 62: 558

159-Holland, J., et al. 1982. Science 215: 1577

160- Murphy, B.R., et al. 1980. Ann NY Acad. Sci. 354: 172

161- Tolpin, M.D., et al. 1982. Infect. Immun. 36: 645

162- Symposia on Quantitative Biology 1974. Vol. XXXIX-Part 1, Cold Spring Harbor Laboratory Press, New York.

163- Ghallberg, M.D., et al. 1989. Annu. Rev. Biochem. 58: 671

164- Stillman, B. 1989. Annu. Rev. Cell Biol. 5: 197

165- Gough, G., et al. 1988 Cancer Cells 6: 153

166- Gannon, J.V., et al. 1990. New Biol. 2: 84

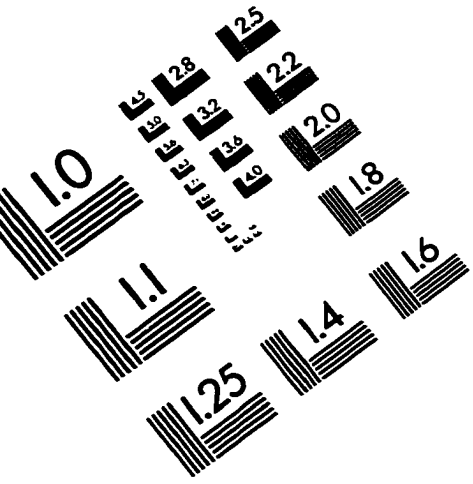
167- Collins, K.L., et al. 1991. Mol. Cell. Biol. 11: 2108

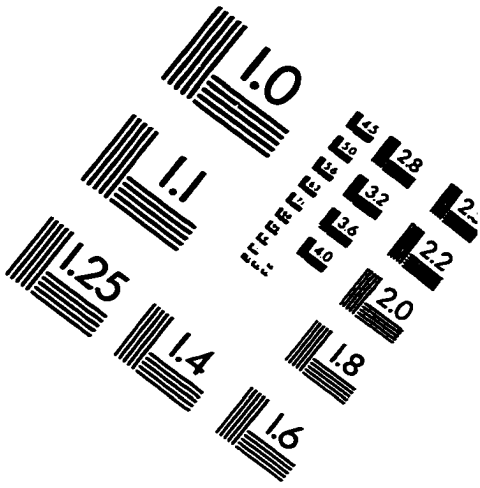
168- Dougherty, J., et al. 1988. J. Virol. 62: 2817

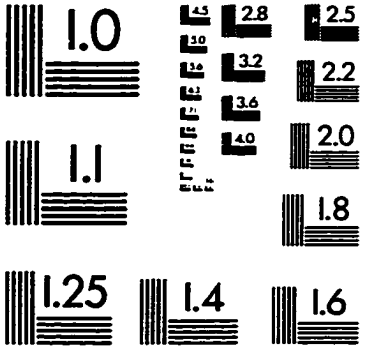
168- Leider, J.M., et al. 1988. J. Virol. 62: 3084

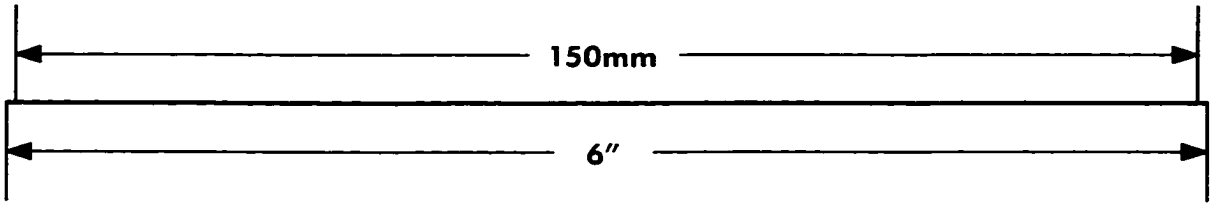
169- Britten, R.J., et al. 1986. Science 231: 1393

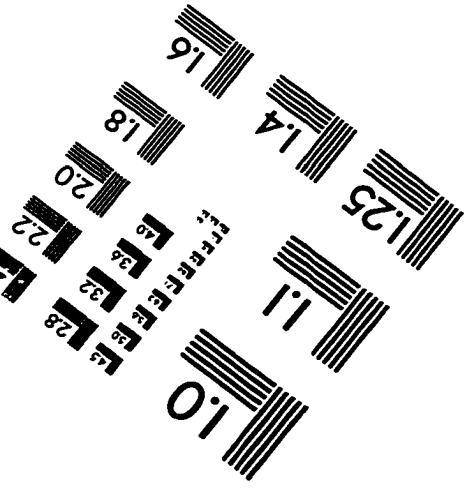
IMAGE EVALUATION TEST TARGET (QA-3)

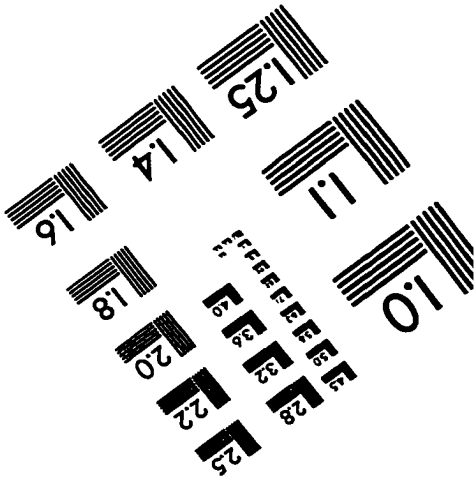












APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989