Volume 14 Number 2 1986 Nucleic Acids Research

Construction of a retroviral cDNA version of the *int*-1 mammary oncogene and its expression *in vitro*

Frans A.M.Rijsewijk, Maarten van Lohuizen, Albert van Ooyen* and Roel Nusse

Department of Molecular Biology, The Netherlands Cancer Institute (Antoni van Leeuwenhoekhuis) Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands

Received 12 November 1985; Accepted 2 December 1985

ABSTRACT

The int-1 mammary oncogene is activated by proviruses of the Mouse Mammary Tumor Virus in many different mammary tumors. We have inserted a genomic fragment containing the protein-encoding domain of the gene into the retroviral shuttle vector pZIPneoSV(X)1. After one round of virus replication we recovered recombinant proviral DNA containing a correctly spliced copy of int-1. In vitro transcription of this cDNA version of int-1 using SP6 polymerase and translation in a reticulocyte lysate yielded a protein of approximately 37,000 daltons. High expression of int-1 in NIH-3T3 cells infected with recombinant virus did not lead to morphological transformation.

INTRODUCTION

Mammary tumors in mice can be induced by a milk-transmitted retrovirus, the Mouse Mammary Tumor Virus (MMTV) (1). During infection of the mammary gland cells by the virus, proviral DNA of MMTV becomes integrated in the host cell genome, causing mutations of various sorts (2). In the genome of the tumor cells, proviral DNA is often found inserted near cellular genes, called int-1 and int-2, that are thought to be host cell oncogenes responsible for viral oncogenesis (3-5). These genes are activated as a consequence of nearby proviral insertion, leading to transcripts exclusively found in virus-induced mammary tumors. We have recently established the complete nucleotide sequence of the int-1 gene and the structure of its transcript by nuclease S1 mapping (6). The structure predicted a protein of 370 amino acids, that has no conspicuous homology with other known gene products. The proviral insertions in mammary tumors are frequently found within exons of int-1, but the domain that encodes the protein is never disrupted. The protein has a hydrophobic leader and its sequence is extremely highly conserved between mouse and man (7).

Further characterization of the structure and the properties of the int-1 protein would greatly benefit from the isolation of cDNA clones. In

this paper, we report on the construction of a cDNA version of int-1 by using the retroviral vector pZIPneoSV(X)1 (ref.8). A fully spliced copy of the gene was obtained after one round of virus replication. RNA transcribed from the cDNA template directed synthesis of a protein of the expected length in an in vitro translation assay.

MATERIALS AND METHODS

Cells and culture conditions. Ψ-2 cells (9), obtained through Dr. E. Zwarthoff from Dr. R. Mulligan, and NIH-3T3 cells were cultured in Dulbecco's modified Eagle's medium (DMEM, Gibco) supplemented with 10% newborn calf serum. The COS-5 cell line, obtained from Dr. J. Tasseron, is a subclone of COS-1 cells from Y. Gluzman (10) and was grown in DMEM with 10% newborn and 10% fetal calf serum. All experiments were performed under biosafety conditions recommended by the Netherlands Committee on recombinant DNA work.

Origin of recombinant clones. The genomic fragment containing the int-1 gene is a 12-kb BglII fragment subcloned from a bacteriophage clone (4) isolated from a library from normal BALB/c DNA. The retroviral shuttle vector pZIPneoSV(X)1 was kindly provided by Dr. C. Cepko (8).

Transfection and selection procedures. The calcium phosphate transfection procedure of Graham and Van der Eb (11) was used for introduction of DNA into the Ψ-2 cells. In the transient virus experiments, virus was harvested 18 hours after transfection. In case of preselection of the Ψ-2 cells for neo expression, cells were grown in normal complete medium for 2 days, and then split at a 1 to 5 ratio and grown in medium containing 400 µg/ml G418 (Gibco). Colonies resistant to the drug were collected after 14 days. Virus from the pooled colonies was harvested 18 hours after medium change.
Virus-containing medium was filtered through a 45 µm nitrocellulose filter. For infections, NIH-3T3 cells at $2x10^5$ cells per 100 cm^2 dish were incubated with 1 ml of virus containing medium plus 8 µg/ml polybrene (Sigma) for two hours. Medium was then changed to DMEM with 10% newborn calf serum, and, two days later, selection with G418 was started.

Rescue of integrated virus by fusion with COS cells. The virus-infected and G418 selected NIH-3T3 cells were co-cultivated with COS-5 cells at a 1:3 ratio in DMEM with 10 % newborn and 10 % fetal calf serum. Confluent cultures were washed in serum-free DMEM before addition of 2 ml 50% polyethylene glycol 1500 (DBH) in DMEM (12). Cells were allowed to fuse for 60-180 seconds, after which they were washed twice with serum-free DMEM and once with DMEM containing 10% newborn calf serum. The cells were then cultured for two days in DMEM with 10% fetal and 10% newborn calf serum, lysed and extracted according to Hirt (13) to isolate extrachromosomal DNA. The Hirt supernatant was used to transform E. coli C600. Recombinant plasmids from kanamycin selected colonies were isolated according to Birnboim and Doly (14).

Isolation and analysis of cellular DNA and RNA. Extraction and Southern blot (15) analysis of DNA was according to standard procedures (16). Cellular RNA was prepared by the lithium chloride method (17). Polyadenylated RNA was enriched over oligo-dT cellulose, size-fractionated over formaldehyde-agarose gels and blot-hybridized. Probes for blots were labeled by nick-translation (18).

Plasmid construction and DNA sequence analysis. The various constructs were assembled using standard procedures (16). DNA sequence analysis was done according to the Smith-Calvo modified Maxam and Gilbert technique (19,20).

RNA synthesis and translation in vitro. We followed the procedure of Melton (21), using reagents from Promega Biotec, to transcribe the int-1 cDNA version cloned into the SP65 cloning vector. This RNA was directly translated in a reticulocyte extract (16) with [^{35}S]-methionine as labeled amino acid. The reaction products were separated on a 11% polyacrylamide gel in the presence of SDS.

RESULTS

From the previously established nucleotide sequence and structure of the genomic clone of int-1, it is known that the entire protein-encoding domain is contained within a 3.4-kb StuI fragment (Fig. 1A, ref.6). The postulated most proximal start of transcription of the gene is also located within this fragment, 25 bp from the end, but the functional promoter of the gene is more upstream. At the 3' end of the gene, the StuI site is in front of the polyadenylation signal. This fragment was therefore very suitable to generate a cDNA clone in a retroviral vector: it lacks a promoter and a polyadenylation signal, each of which can interfere with replication of retroviral vectors (22-24), and it should be able to direct synthesis of the complete int-1 protein.

This StuI fragment was isolated from a larger subclone (pMT34) that contains the whole int-1 gene (4), and ligated with the aid of BglII linkers, into the BamHI site of the retroviral vector pZIPneoSV(X)1, hereafter called SV(X) (ref.8, Fig. 1B). The BamHI site in the vector is preceded by a packaging signal, a splice donor site, and the proviral LTR. Transcription from the LTR generates both full-length viral RNA (containing the inserted gene and, by differential splicing ,subgenomic RNA (for expression of neomycin resistance). We selected plasmids with the inserted int-1 fragment in the same transcriptional orientation as the viral vector. These plasmids were transfected into Ψ-2 cells, which supply all viral proteins in trans, to generate virus particles containing processed int-1 inserts, free of helper virus.

To establish integrated proviral DNA, the virus produced by the transfected Ψ-2 cells was subsequently used to infect NIH-3T3 cells, either transiently or after selection of the Ψ-2 cells for resistance to G418. In the transient infections, medium from the Ψ-2 cells was harvested 18 hours after transfection. The transient virus stock had a titer of 10^2 CFU/ml of G418 resistant NIH-3T3 colonies. When the transfected Ψ-2 cells were first selected for expression of the acquired DNA and grown into mass-culture, a much higher virus titer was achieved, $7x10^4$ CFU/ml. Transient and stable virus titers of the SV(X) vector without insert were $2x10^2$ and $3x10^5$ CFU/ml, respectively.

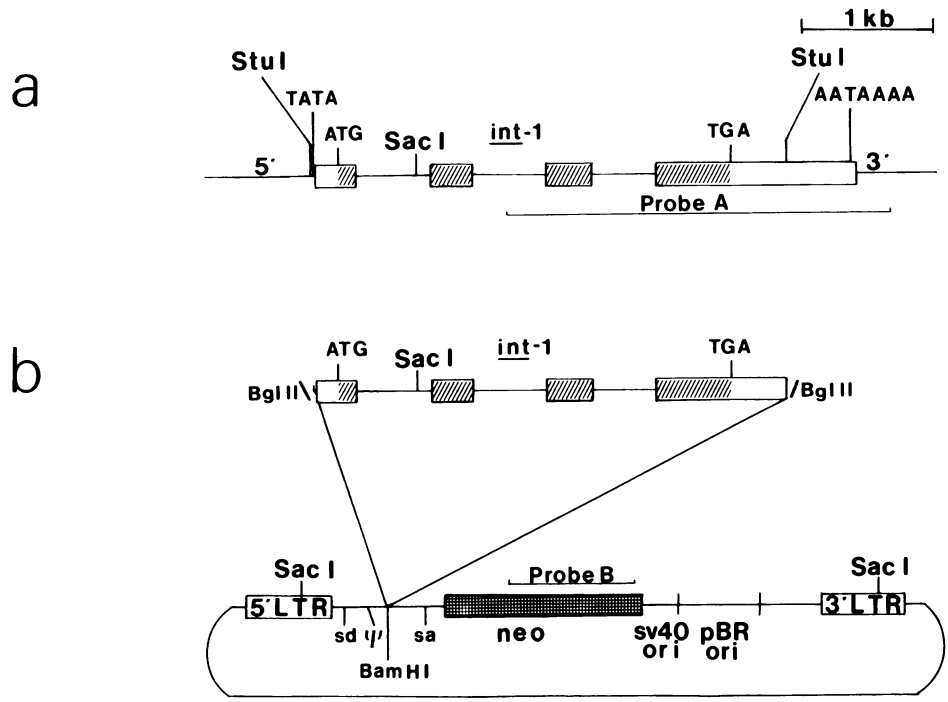

Fig. 1: DIAGRAMS OF THE MOUSE INT-1 GENE AND THE INT-1-pZIPneoSV(X)1 CONSTRUCT

a. The structure of the mouse int-1 gene. The boxed areas indicate the exons. The striped boxes represent the protein-encoding domain from the start codon (ATG) to the stop codon (TGA). TATA= TATA box region; AATAAAA=polyadenylation signal. StuI and SacI= restriction enzyme recognition sites, used for molecular cloning or hybridization analysis.

b. Structure of the 3.4-kb StuI genomic int-1 fragment, inserted into the BamHI site of pZIPneoSV(X)1 shuttle vector by the use of BglII linkers. The vector is described by Cepko et al. (8). The proviral sequences originate from a M-MuLV provirus, are bordered by mouse genomical sequences and ligated in pBR322. LTR= long terminal repeat; sd= splice donor site; ψ= packaging signal; sa= splice acceptor site; Neo= Neomycin resistance gene (on the 1.4 kb-HindIII-SalI fragment from Tn5); SV40 ori= origin of replication of SV40. pBR ori= origin of replication of pBR322. SacI and BamHI= restriction enzyme sites.

Analysis of proviral DNA in infected NIH-3T3 cells.

The structure of the proviruses in the infected cells was examined by restriction digests and blot hybridization. Digestion with SacI proved to be informative. This enzyme has one recognition site in the inserted int-1 fragment, within an intron, and it cuts in the LTR of the vector (Fig.1). If the int-1 insert is spliced correctly, a fragment of 5.8-kb, equivalent to

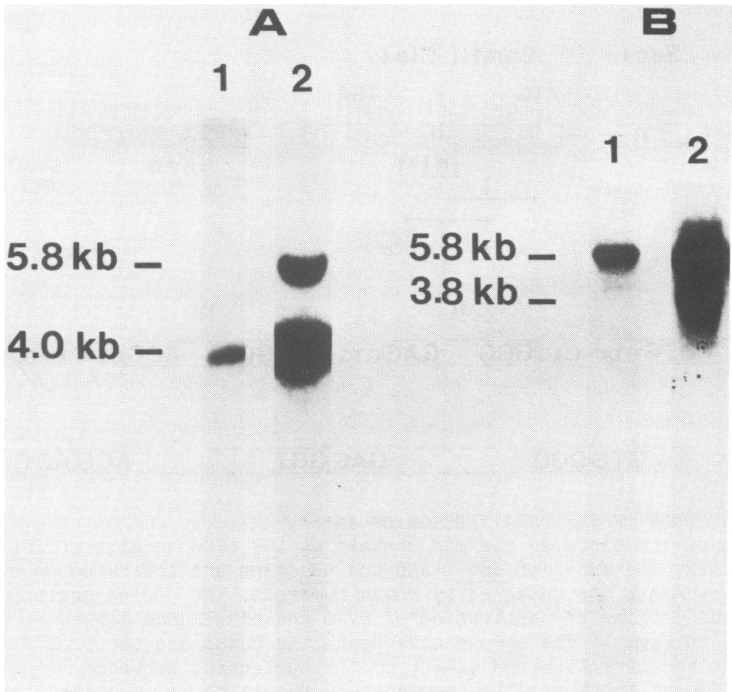

Fig. 2: BLOT ANALYSIS OF THE NIH-3T3 CELLS INFECTED BY THE INT-1-pZIPneoSV(X)1 VIRIONS.
A. Autoradiogram of a Southern blot analysis of DNA digested with SacI, separated on a 0,7% agarose gel and hybridized with int-1 probe A (see Fig. 1A). Lane 1: DNA of NIH-3T3 cells. Lane 2: DNA of NIH-3T3 cells infected with int-1-pZIPneoSV(X)1 virions.
B. Northern blot analysis of RNA of NIH-3T3 cells infected with int-1-ZIPneoSV(X)1 virions and separated on a 1% agarose gel. Lane 1: hybridized with int-1 probe A (see Fig. 1A). Lane 2: hybridized with neo probe B (see Fig. 1 B).

the unit length of the recombinant provirus, is expected to hybridize to int-1 probes. A fragment of this size was observed in all three NIH-3T3 clones which were infected with transient virus, next to the endogenous int-1 SacI fragment of 4.0-kb (Fig.2A). In contrast, all eight NIH-3T3 clones that had arisen from infection by the stable high-titer virus, contained aberrantly sized fragments, or no acquired int-1 copies at all, whereas they did hybridize to neo probes. Apparently, preselection of the transfected Ψ-2 cells for expression of neo results in emergence of clones in which the provirus had sustained deletions. These deletions may allow more efficient expression of neomycin resistance. We also analyzed the

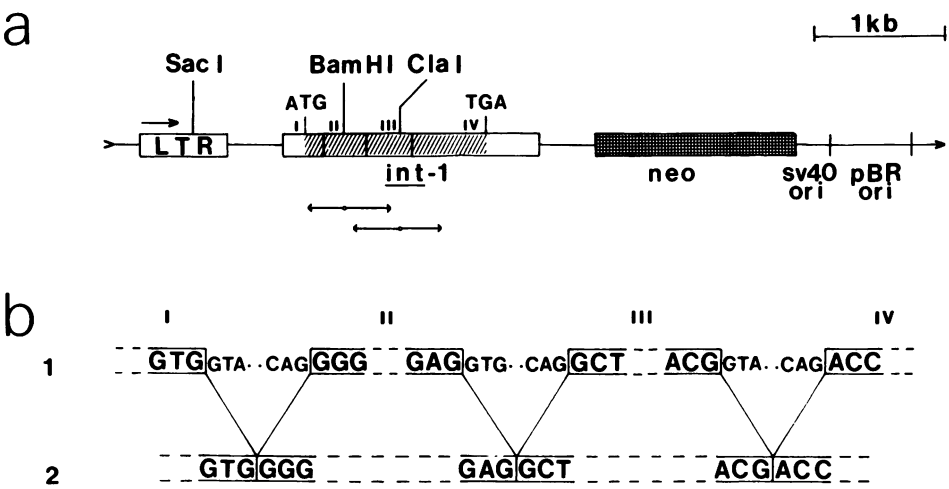

Fig. 3 STRUCTURE OF THE cDNA VERSION OF INT-1.
a. Diagram of the circular plasmid containing the cDNA version of int-1. The M-MuLV-LTR, the exons of int-1 and the neo gene are indicated by boxes. The int-1 exons are numbered by roman numerals. The coding region of int-1 and the origins of replication of SV40 and pBR322 are marked as in the legends to Fig. 1. The arrows underneath the BamHI and the ClaI site indicate the direction and extent of the nucleotide sequence.
b. Comparison of the nucleotide sequence of the intron-exon borders as postulated by Van Ooyen and Nusse (6), with the exon-exon borders found in the int-1 cDNA version. The roman numerals indicate the exons of int-1.

transcripts in the NIH-3T3 cells infected with transient virus. A probe for the int-1 insert hybridized to an RNA species of 5.8-kb; and a probe for the neo gene recognized, in addition to the 5.8-kb species, a transcript of 3.8-kb (Fig.2B).

We conclude that proviruses established from the transient virus stocks were reverse transcripts from correctly spliced viral RNA, and that they generated the expected full-length and subgenomic transcripts.

The NIH-3T3 cells expressing the viral transcripts containing int-1 sequences did not significantly differ from the parental cell line with respect to morphology. In addition, the cells were not tumorigenic upon injection in a nude mouse. Expression of int-1 from the inserted provirus is thus not sufficient to transform a fibroblastic cell line.

Recovery of the integrated proviruses.

The molecular cloning of the integrated provirus containing the spliced int-1 version was facilitated by the presence of a functional origin of replication of both SV40 and pBR322 in the SV(X) vector (Fig. 1B,ref.8).

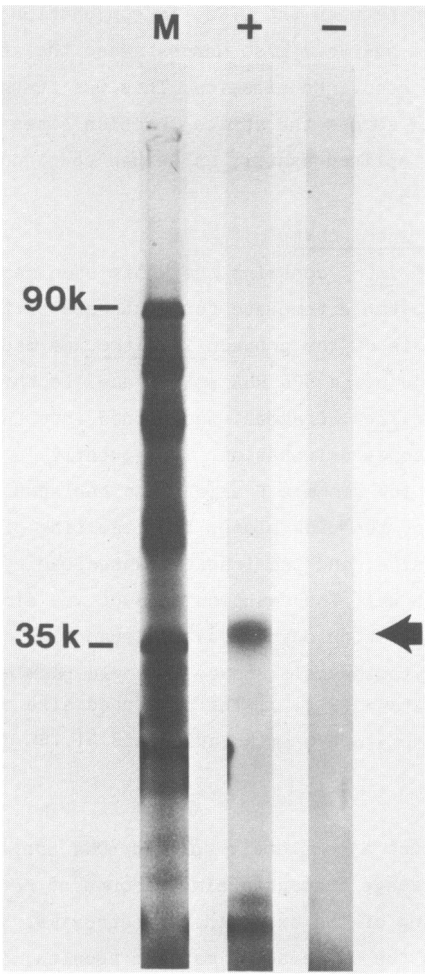

Fig. 4 Analysis of the translation products of SP6 RNA transcribed from int-1 cDNA in a reticulocyte lysate. The [^{35}S] methionine-labeled proteins were separated by SDS-PAGE on an 11% gel and visualized by fluorography. 150-200 ng RNA was added (+) or no RNA was added (-). M= Markers in kilodaltons.

The infected NIH-3T3 cells were fused to COS cells, which express high levels of the SV40 T antigen (10). This leads to "onion skin" replication from the SV40 origin, and rapid formation of extrachromosomal proviral DNA. This DNA was isolated by Hirt extraction, and used to transform E. coli, which were selected for kanamycin resistance. Subsequent restriction analysis of the plasmids showed that circular DNA had been recovered containing one copy of the proviral LTR (Fig. 3A), apparently arisen from

homologous recombination between the LTRs. In conjunction with heteroduplex mapping (not shown) this analysis also demonstrated the absence of all three introns of int-1 in the recovered plasmids. This was further substantiated by nucleotide sequencing across the splice junction sites (Fig 3B). The four exons of int-1 had been spliced exactly as we had concluded from the S1 analysis of tumor mRNA (6).

In vitro translation of int-1 cDNA

The cDNA version of int-1 contains the whole open reading frame,and should therefore be a suitable template for synthesis of the protein. In order to attempt synthesis of the protein in vitro, we made int-1 RNA from the retroviral cDNA clone using SP6 RNA polymerase. To this end, a BglII/StuI-BglII genomic int-1 fragment was cloned into the BamHI site of the SP65 vector in the sense orientation. Subsequently, a NcoI-AccI fragment covering the introns of the genomic fragment was replaced by the corresponding fragment of the cDNA clone. The resulting plasmid was linearized at the StuI site, and the insert transcribed by SP6 RNA polymerase into a 1.6-kb RNA. The reaction product was directly, without capping, translated in a reticulocyte lysate. Separation of the translation products on SDS polyacrylamide gels showed a single prominent product of approximately 37,000 daltons (Fig.4). The predicted size of the int-1 protein, which should contain 370 amino acids, is 41,185 daltons.

DISCUSSION.

Our results show that a completely spliced cDNA copy of the int-1 gene has been obtained by passage through a single round of retroviral replication. The splicing of the exons in the retroviral vector has occurred exactly as deduced from the nuclease S1 mapping results, illustrating the general usefulness of retroviral vectors for obtaining cDNA clones (8). Conversely, the result of this cloning procedure confirms that our proposal for the structure of the int-1 gene (6), based on S1 mapping of exons and nucleotide sequencing is correct. Very recently, Fung et al (25). have isolated a cDNA clone of int-1 by reverse transcribing mRNA from mammary tumors induced by MMTV. The nucleotide sequence of this clone is also identical to the sequence of the genomic int-1 exons.

The clone that we have constructed encodes the complete protein of int-1 and will hence be used to synthesize the int-1 gene product. We have shown in this paper that RNA made from the cDNA template directs translation of a 37,000 dalton protein. We are currently inserting the cDNA clone into a variety of bacterial expression vectors, aiming at synthesis of large

quantities of the protein to be used as immunogen.

An additional spin-off from the retroviral vector now available may be the definition of the biological properties of int-1. The assignment of int-1 as an oncogene for mammary tumors is still largely based on the high frequency of insertional activation by MMTV in mammary tumors. MMTV has no viral oncogene by itself, and depends on activation of cellular genes for tumor formation. Unlike other host genes that are activated by slow retroviruses in the course of oncogenesis, such as c-myc (26), the int genes have not been encountered as host-derived oncogenes in rapidly transforming retroviruses. We now have a retroviral variant that contains int-1 as its host cell derived oncogene, and we hope to study the biological properties of the gene by infection of normal cells. The NIH-3T3 cells infected by the int-1 recombinant virus showed no conspicuous sign of morphological transformation although high levels of RNA were detected. Transfection of established fibroblast cell lines with int-1 constructs does not lead to manifest transformation either (unpublished results). It is possible that expression of int-1 in fibroblast lines results in more subtle changes in growth properties, but alternatively, the biological effects of activated int-1 may be limited to certain cell types, perhaps mammary gland cells only.

Retroviral vectors could be suitable to introduce activated oncogenes into primary mammary gland cells, but this will require some additional manipulation of the recombinant clones, in order to obtain virus stocks with high titers. As we and others have observed, high-titer production of recombinant virus from transfected Ψ-2 cells is only obtained when the producer cells are selected for G418 resistance. This often leads to rearrangements of the provirus, notably in cases where the insert is large and contains splice signals, apparently because expression of neo from the subgenomic mRNA is less efficient. In contrast, simultaneous expression of both the inserted sequence and neo is observed in the NIH-3T3 cells which contain a processed provirus (Fig.2), originally established from a transient round of virus replication. The titers of transient virus stocks from transfected ψ-2 cells are, however, too low for biological purposes. To circumvent these problems, we are introducing vector constructs having cDNA inserts rather than a genomic copy into the ψ-2 packaging cells. Selection of the transfected cells for expression of neo will then lead to production of high-titer virus with an intact genome. Virus variants with int-1 cDNA clones as inserts are now being produced for biological assays.

Acknowledgements.

We thank many colleagues for supplying reagents, Karin Langereis for initial help with the in vitro translation experiments and Dr. E. Matthews and Dr. D. Greaves for comments on the manuscript. We also thank Dr. H. Varmus for communication of unpublished results. FAMR and AvO were supported by grants by the Netherlands Cancer Society.

*Present address: Gist Brocades, Postbus 1, 2600 MA Delft, The Netherlands

REFERENCES

1. Teich, N., Wyke, J.A., Mak, T., Bernstein, A. and Hardy, W. (1982) Pathogenesis of retrovirus-induced disease. In: The Molecular Biology of Tumor Viruses. Part III RNA Tumor Viruses. Weiss, R.A., Teich, N., Varmus, H.E. and Coffin, J.M. (eds.). Cold Spring Harbor Laboratory, p. 999-1109.
2. Varmus, H.E. (1983) In: Mobile Genetic Elements. Shapiro, J.A. (ed.). p. 411-503. Academic Press.
3. Peters, G., Brookes, S., Smith, R. and Dickson, C. (1983) Cell 33, 369-377.
4. Nusse, R. and Varmus, H.E. (1982) Cell 31, 99-109.
5. Nusse, R., van Ooyen, A., Cox, A., Fung, Y.-K.T. and Varmus, H.E. (1984) Nature 307, 131-136.
6. Van Ooyen, A. and Nusse, R. (1984) Cell 39, 233-240.
7. Van Ooyen, A., Kwee, V., and Nusse, R. (1985) EMBO. J.(in press)
8. Cepko, C.L., Roberts, B.E. and Mulligan, R.C. (1984) Cell 37, 1053-1062.
9. Mann, R., Mulligan, R.C. and Baltimore, D. (1983) Cell 33, 153-159.
10. Gluzman, Y (1981) Cell 23, 175-182.
11. Graham, F.L. and van der Eb, A.J. (1973) Virology 52, 456-467.
12. Davidson, R.L., O'Malley, K.A. and Wheeler, T.B. (1976) Somatic Cell Genet. 2, 271-280.
13. Hirt, D. (1967) J. Mol. Biol. 26, 365-369.
14. Birnboim, H.C. and Doly, J. (1979) Nucl. Acids Res. 7, 1513-1523.
15. Southern, E.M. (1975). J. Mol. Biol. 98, 503-517.
16. Maniatis, T., Fritsch, P. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory).
17. Auffray, C., Rougeon, F. (1980) Eur. J. Biochem. 107, 303-314.
18. Rigby, P.W.J., Dieckmann, M., Rhodes, C. and Berg, P. (1977) J. Mol. Biol. 113, 237-251.
19. Maxam, A.M. and Gilbert, W. (1980) Methods in Enzymology, vol. 65, p. 499-560, Academic Press Inc., New York.
20. Smith, D.R., and Calvo, J.M. (1980) Nucl. Acids Res. 8, 2255-2274.
21. Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K. and Green, M.R. (1984) Nucleic Acids Res. 12, 7035-7056.
22. Emerman, M. and Temin, H.M. (1984) J. Virol. 50, 42-49.
23. Miller, A.d., Verma, I.M., Friedmann, T. and Jolly, A.J. (1983) Gene expression, p. 555-563 Alan R. Liss Inc., New York.
24. Joyner, A.L. and Bernstein, A. (1983) Mol. Cell. Biol. 3, 2180-2190.
25. Fung, Y.-K.T., Schackleford, G.M., Brown, A.M.C., Sanders, G.S. and Varmus, H.E. (1985) Mol. Cell. Biol. (in press).
26. Hayward, W.S., Neel, B.G. and Astrin, S.M. (1981) Nature 290, 475-480.