is alive and reasonably healthy—but it could be much more so. To flourish vigorously, to utilize to the full our vast human resources for greater progress in health research, additional sums are needed for many purposes, but equally we need a renewal of the long-range commitment to excellence and accelerated progress, and to preservation of managerial flexibility within the enterprise itself.

The AAMC has consistently and effectively supported this position. In that connection, I urge you to read carefully the analysis and exposition of principles for the support of biomedical research just issued by the AAMC executive council. This little blue book, titled *Preserving America's Preeminence in Medical Research*, is an exceptionally clear, balanced, and persuasive statement deserving of widespread attention.

Most of you will remember another influential report issued by the AAMC in 1965. It was written by Lowell T. Coggeshall and titled *Planning for Progress Through Medical Education*. In the report was an observation that fits exactly the context of my remarks and, in a sentence, captures the ideas I have endeavored to present. In Dr. Coggeshall's words—as valid now as when they were written—"The important question for the future is whether the present system is sufficiently flexible and imaginative to keep pace with the contemporary revolution in medical sciences and the changing expectations of the American people" (*4*).

References and Notes

1. L. Thomas, from the foreword of *Laboratories and Clinics: An Account of Research at the National Institutes of Health* (Academic Press, New York, in press).
2. Report on 1976 Appropriations for Departments of Labor, Health, Education, and Welfare and Related Agencies. U.S. Senate 1975.
3. A. Kornberg, from "Biology and technology," unpublished manuscript, 1982.
4. L. T. Coggeshall, "Planning for medical progress through education," a report submitted to the Executive Council of the Association of American Medical Colleges (Association of American Medical Colleges, Washington, D.C., 1965).

RESEARCH ARTICLE

Major *pol* Gene Progenitors in the Evolution of Oncoviruses

Ing-Ming Chiu, Robert Callahan, Steven R. Tronick
Jeffrey Schlom, Stuart A. Aaronson

Oncoviruses, a subfamily of Retroviridae (*1*), are the causative agents of naturally occurring tumors in diverse vertebrate species. Unlike most viruses, which are spread only as infectious agents, oncoviruses can also be transmitted within the germ line of the host. Under such conditions, these viruses are passed from one generation to the next and often in an unexpressed form. The widespread distribution of oncoviruses among vertebrates implies that this intimate association has persisted through a considerable period of evolution.

The oncovirus genera have been classified by morphologic criteria (*2*). Defective intracisternal type A viral particles were initially observed in early embryos of mice and in certain murine tumors (*3*). Later studies revealed the existence of infectious retroviruses containing extensive homology to the A particle genome as well (*4*). Type B viruses have been found only in murine species. Such viruses have been established as etiologically responsible for mammary tumors of the mouse (*5*). Type C viruses, which are widely distributed among birds and mammals, cause leukemia and other tumors [see (*6*) for reviews]. The most recently described type D oncoviruses are so far limited to primate species, and their oncogenicity remains to be established (*7*).

In recent years, efforts have been made to ascertain the evolutionary relationships among different oncovirus genera. One of the most useful approaches has been the demonstration of shared antigenic determinants in their translational products. Interspecies cross-reactivity was initially observed for several early isolates of type C viruses (*8*). The advent of radioimmunological techniques made it possible to demonstrate the presence of interspecies determinants common to the respective *gag*, *pol*, and *env* gene products of all known mammalian type C viruses (*9*). Such studies have led to the conclusion that mammalian type C viruses arose from a common progenitor. The detection of immunological relatedness between the major structural proteins of type B and D viruses, as well as between mammalian type C and D viruses, has suggested that evolutionary links may also exist among these three major oncovirus genera (*10*).

Efforts to analyze the structural and evolutionary relationships between different oncoviruses have been facilitated by the ability to isolate and amplify these viral genomes by molecular cloning techniques. In the present studies, we used molecular hybridization and nucleotide sequence analysis to detect and localize related genes of viruses representing different oncovirus genera. We have now established the existence of major *pol* gene families in the evolution of oncoviruses, as well as other previously undetected evolutionary linkages.

Oncoviruses have been classified on the basis of their morphological properties (*2*). Even though four different genera have been recognized (Table 1), vi-

Abstract. *The genetic relationships among molecularly cloned prototype viruses representing all of the major oncovirus genera were investigated by molecular hybridization and nucleotide sequence analysis. One of the major progenitors of the* pol *genes of such viruses gives rise to mammalian type C viruses and another gives rise to type A, B, D, and avian type C oncoviruses. Evidence of unusual patterns of homology among the* env *genes of mammalian type C and D oncoviruses illustrates that genetic interactions between their progenitors contributed to the evolution of oncoviruses.*

I.-M. Chiu, S. R. Tronick, and S. A. Aaronson are in the Laboratory of Cellular and Molecular Biology and R. Callahan and J. Schlom are in the Laboratory of Tumor Immunology and Biology, National Cancer Institute, Bethesda, Maryland 20205.

ruses belonging to the type A, B, and D genera share certain morphogenic and biochemical properties. We used cloned squirrel monkey retrovirus (SMRV) (*11*), a prototype D virus, as a molecular hybridization probe to detect genetic relatedness among oncoviruses. The cloned viral DNA's studied included the recombinant type A virus M432 (*4*); a mammalian type B virus, mouse mammary tumor virus (MMTV) (*12*); the mammalian type C viruses, Rauscher murine leukemia virus (R-MuLV) (*13*), Moloney murine leukemia virus (M-MuLV) (*14*), simian sarcoma–associated virus (SSAV) (*15*), and baboon endogenous virus (BaEV) (*16*); and avian type C viruses, Rous sarcoma virus (RSV) (*17*) and avian myeloblastosis–associated virus (MAV) (*18*). The oncoviral inserts in each clone were subjected to electrophoresis on an agarose gel (0.8 percent), transferred to nitrocellulose, and then hybridized under relaxed conditions with labeled SMRV DNA. The SMRV probe demonstrated homology with all of the viral genomes tested (Fig. 1).

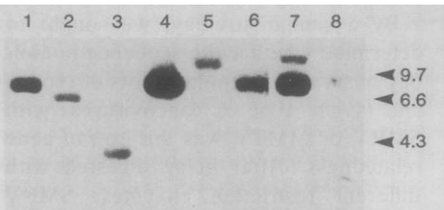


Fig. 1. Related nucleotide sequences detected by SMRV in other oncoviral genomes. Viral DNA's cloned in pBR322 were digested with the appropriate restriction enzymes and separated on an agarose gel (0.8 percent). The DNA's were transferred to nitrocellulose paper by the method of Southern (*34*) and hybridized under nonstringent conditions in a buffer [1*M* NaCl, 10× Denhardt's solution (*35*), salmon sperm DNA (100 μg/ml), and 20 percent formamide] with nick-translated (*36*) SMRV probe (1×10^7 count/min per milliliter) at 37°C for 36 hours. The filters were then washed in triple-strength standard saline citrate, 0.1 percent sodium dodecyl sulfate, and 0.1 percent sodium pyrophosphate at 50°C. One microgram of DNA was loaded in each lane except lane 4 (0.02 μg). The viruses tested were M432 (lane 1); MMTV, clone GR7-1a (lane 2); MAV, 4-kbp Eco RI subgenomic fragment (lane 3); SMRV (lane 4); R-MuLV (lane 5); SSAV (lane 6); and RSV (lane 7). Each clone represents the entire viral genome with the exception of MMTV and MAV. An equimolar amount of pBR322 DNA was present in each lane, including lane 8. Autoradiograms of lanes 5, 6, and 8 were exposed four times longer than those of lanes 1, 2, 3, 4, and 7.

To confirm and extend these findings, we performed reciprocal hybridizations using each of the other cloned viruses as a probe. The extent of homology was quantified by determining the amount of radioactivity in each band. As shown in Table 2, members of each oncovirus genus hybridized most strongly among themselves. Furthermore, reciprocal relationships were observed not only between SMRV and representative type A, B, and avian C viruses, but among each of these other groups as well.

Homologous regions of oncoviral pol *genes.* In order to localize regions of MMTV homologous with SMRV, we molecularly dissected clone GR7-1a, which contains 6.8 kilobase pairs (kbp) of the MMTV proviral genome permuted at its single Eco RI site (*12*). Digestion with Eco RI and Pst I generated six fragments, only two of which (3.1 and 0.9 kbp) contained *pol* sequences. Only these two fragments were specifically hybridized by SMRV DNA (data not shown). Further digestion of the 3.1-kbp fragment with Bgl II yielded two fragments, both of which were hybridized by the SMRV DNA probe. In a reciprocal experiment, three fragments (1.8-kbp Pst I–Bgl II, 1.3-kbp Bgl II–Eco RI, and 0.9-kbp Eco RI–Pst I) representing the 5′, middle, and 3′ end of the MMTV *pol* gene, respectively, were isolated and used as hybridization probes. Each probe detected specific SMRV DNA fragments after cleavage of SMRV DNA with either Bam HI or Bgl II plus Pst I, or with Sph I (Fig. 2). The strongest

Table 1. Comparative morphological and biochemical properties of infectious type A, B, C, and D oncoviruses.

Characteristic	Type A	Type B	Type C		Type D
			Avian	Mammalian	
Prototypes	M432*	MMTV	RSV MAV	MuLV SSAV BaEV	SMRV
Presence of intracytoplasmic A particles	−	+	−	−	+
Complete nucleoid at budding	+	+	−	−	+
Preferred cation for DNA polymerase activity	Mg^{2+}	Mg^{2+}	Mg^{2+}	Mn^{2+}	Mg^{2+}
Mature particle					
Nucleoid morphology	Centric	Eccentric	Centric	Centric	Eccentric
Envelope spikes	N.A.†	Long, with knobs	Short	Short	Short
Hormone responsiveness of virus	No	Yes	No	No	No

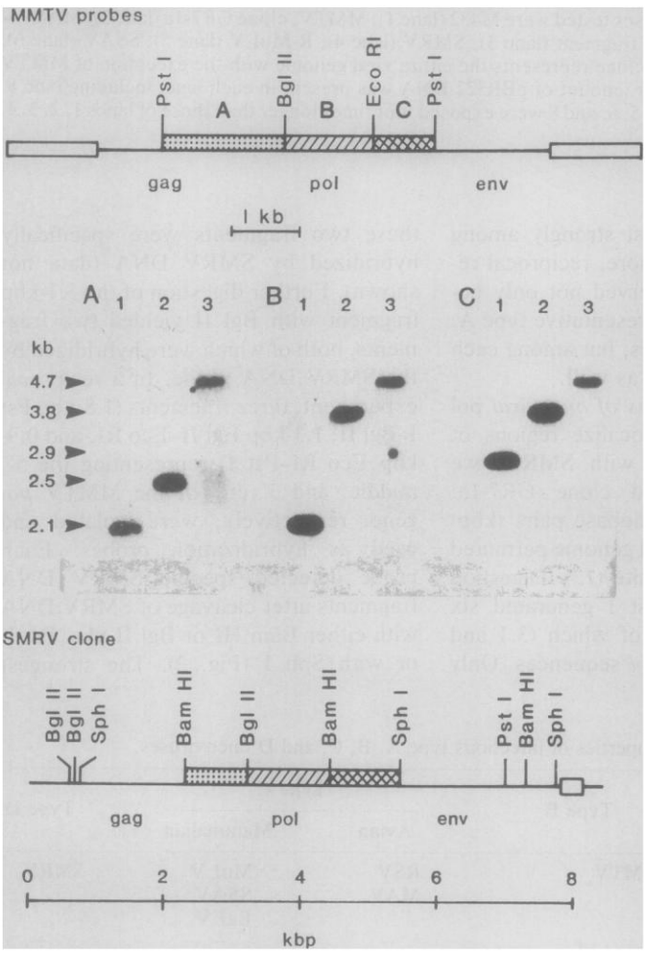
*M432 is an endogenous virus from *Mus cervicolor* containing extensive sequence homology with type A intracisternal A particle (IAP) genes (*4*). There is no extracellular form of the IAP, budding taking place into endoplasmic reticulum. †N.A., not applicable.

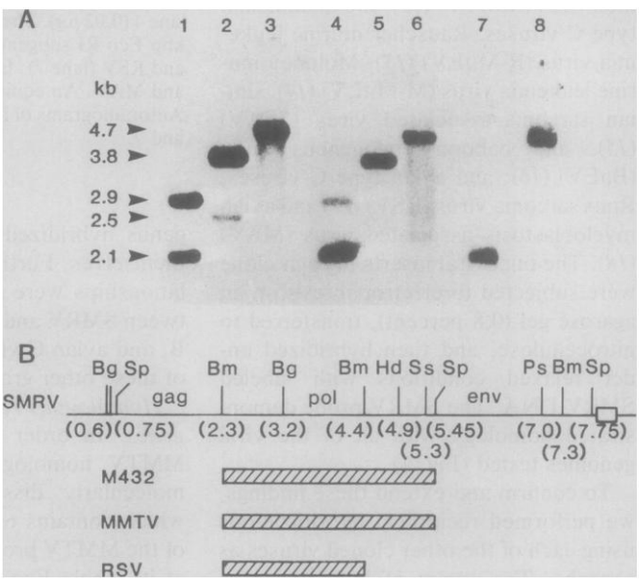
Table 2. Molecular hybridization analysis of cloned oncoviral DNA's. Radioactive bands were cut out of filters and their activity was measured in a liquid scintillation spectrometer. Symbols: −, < 300 count/min; +, 2,000 to 10,000 count/min; ++, 10,000 to 100,000 count/min; +++, 100,000 to 1,000,000 count/min; ++++, > 1,000,000 count/min; N.T., not tested.

Hybridization probes	Reaction with cloned viral DNA							
	M432	MMTV	MAV	RSV	SMRV	M-MuLV	SSAV	BaEV
M432	++++	++	+	N.T.	++	N.T.	−	−
MMTV	++	++++	N.T.	+	++	−	N.T.	N.T.
MAV	+	N.T.	++++	+++	+	−	N.T.	N.T.
RSV	+	+	+++	++++	+	N.T.	−	−
SMRV	++	++	+	N.T.	++++	N.T.	+	++
M-MuLV	−	−	−	N.T.	+	++++	+++	+++
SSAV	−	N.T.	N.T.	−	+	+++	++++	N.T.
BaEV	−	N.T.	N.T.	−	++	+++	N.T.	++++

signals were obtained with the probe representing the 3′ end of the MMTV *pol* gene. None of these probes detected sequences to the right of the Sph I site at 5.45 kbp on the SMRV map (Fig. 2). Thus, these findings established genetic relatedness between *pol* gene sequences of prototype B and D oncoviruses.

By a similar strategy, we sought to determine whether the sequence homology observed in genomic blots of type A and avian type C oncoviruses with SMRV or MMTV was due to *pol* gene relatedness. After being digested with different restriction enzymes, SMRV was subjected to Southern blotting analysis with M432, MMTV, and RSV each used as probes. As shown in Fig. 3A, each of these probes specifically hybridized to SMRV DNA fragments containing its *pol* gene sequences. In reciprocal studies, genomic SMRV probes hybridized to the *pol* genes in digests of M432, MAV, and RSV (Fig. 4A). These find-





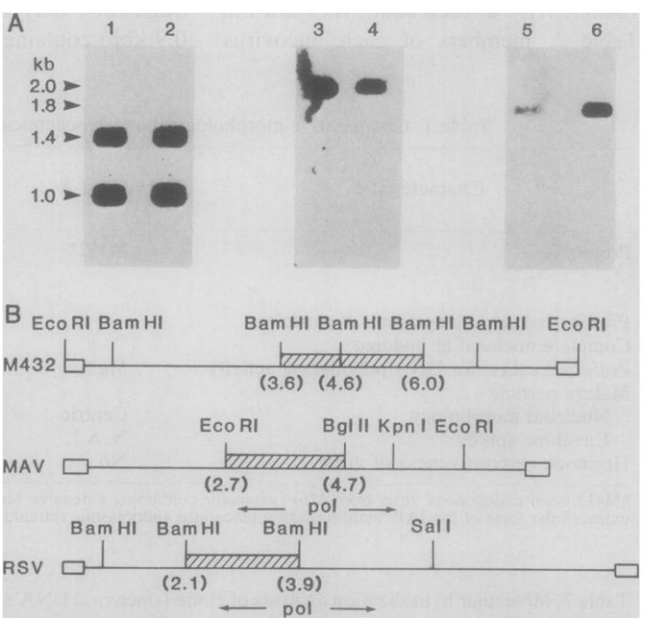


Fig. 2 (left). Homology between *pol* gene sequences of SMRV and MMTV. MMTV clone pMMTV-(C3H)2.5, which contained 4.0 kbp of sequences coding for MMTV *gag-pol*, was digested with Pst I, Bgl II, and Eco RI. Three fragments, Pst I–Bgl II (1.8 kbp), Bgl II–Eco RI (1.3 kbp), and Eco RI–Pst I (0.9 kbp), designated A, B, and C on the MMTV map shown on the top line, were used as probes in (A), (B), and (C), respectively. In (A), (B), and (C), λSMRV DNA was digested with Bam HI (lane 1), Bgl II and Pst I (lane 2), and Sph I (lane 3) and subjected to electrophoresis on an agarose gel (1.2 percent). The regions in SMRV DNA homologous to each of the three probes are shown in the lower map as stippled (probe A), hatched (probe B), and cross-hatched (probe C) boxes. Open boxes represent long terminal repeats (LTR's). Nonstringent hybridization conditions were used as described in Fig. 1. The exposure time for (A) was four times longer than for (B) and (C). The gene order of SMRV shown in the map above was determined on the basis of these and previously reported results (*11*). Fig. 3 (top right). Localization of the regions in SMRV homologous to type A, B, and avian type C oncoviruses. (A) SMRV DNA was digested with Bam HI (lanes 1, 4, and 7), Bgl II plus Pst I (lanes 2 and 5), and Sph I (lanes 3, 6, and 8) and separated on an agarose gel. The DNA was then transferred to nitrocellulose paper and hybridized with nick-translated viral DNA probes M432 (lanes 1 to 3), MMTV clone GR7-1a (lanes 4 to 6), and RSV (lanes 7 and 8). From the results, the maximum regions of homology in SMRV could be deduced and are shown in (B) as hatched boxes. The open box represents the LTR sequence and the map coordinates are given in parentheses. Abbreviations: *Bg*, Bgl II; *Sp*, Sph I; *Bm*, Bam HI; *Hd*, Hind III; *Ss*, Sst I; and *Ps*, Pst I. Only the pertinent restriction enzyme sites are shown. Fig. 4 (bottom right). Demonstration of reciprocal homologies between M432, MMTV, MAV, RSV, and SMRV *pol* genes. (A) DNA's were digested with appropriate enzymes, separated on a 1.2 percent agarose gel, transferred to nitrocellulose paper, and hybridized with the probes, as described in Fig. 1. M432 DNA was digested with Bam HI and hybridized with SMRV (lane 1) or MMTV (lane 2) DNA probes. MAV DNA (4-kbp Eco RI subgenomic fragments) was digested with Bgl II and Kpn I simultaneously and hybridized with SMRV (lane 3) or M432 DNA probes (lane 4). Sal I–permuted 9.6-kbp RSV DNA was digested with Bam HI and hybridized with SMRV (lane 5) or MMTV (lane 6) DNA probes. The regions of homology are shown in (B) as hatched boxes. The *pol* regions of MAV and RSV are demarcated by arrows. The open boxes represent LTR's.

ings strongly imply genetic relatedness between the *pol* genes of representatives of the four major oncovirus genera. In contrast, when analogous studies were performed with mammalian type C viral DNA's, no *pol* gene homologies were observed. The relatedness of SMRV to other sequences in these viral DNA's are described below.

A common pol *gene progenitor for oncoviruses with Mg^{2+}-dependent reverse transcriptases.* To determine precisely the extent of homology between the *pol* genes of these oncoviruses, we undertook comparative nucleotide sequence analysis. The complete nucleotide sequences of RSV (*19*) and M-MuLV (*20*) *pol* genes as well as the 3′ terminal region of the MMTV *pol* gene (*21*) have been reported. We therefore chose to sequence the 3′ terminal region

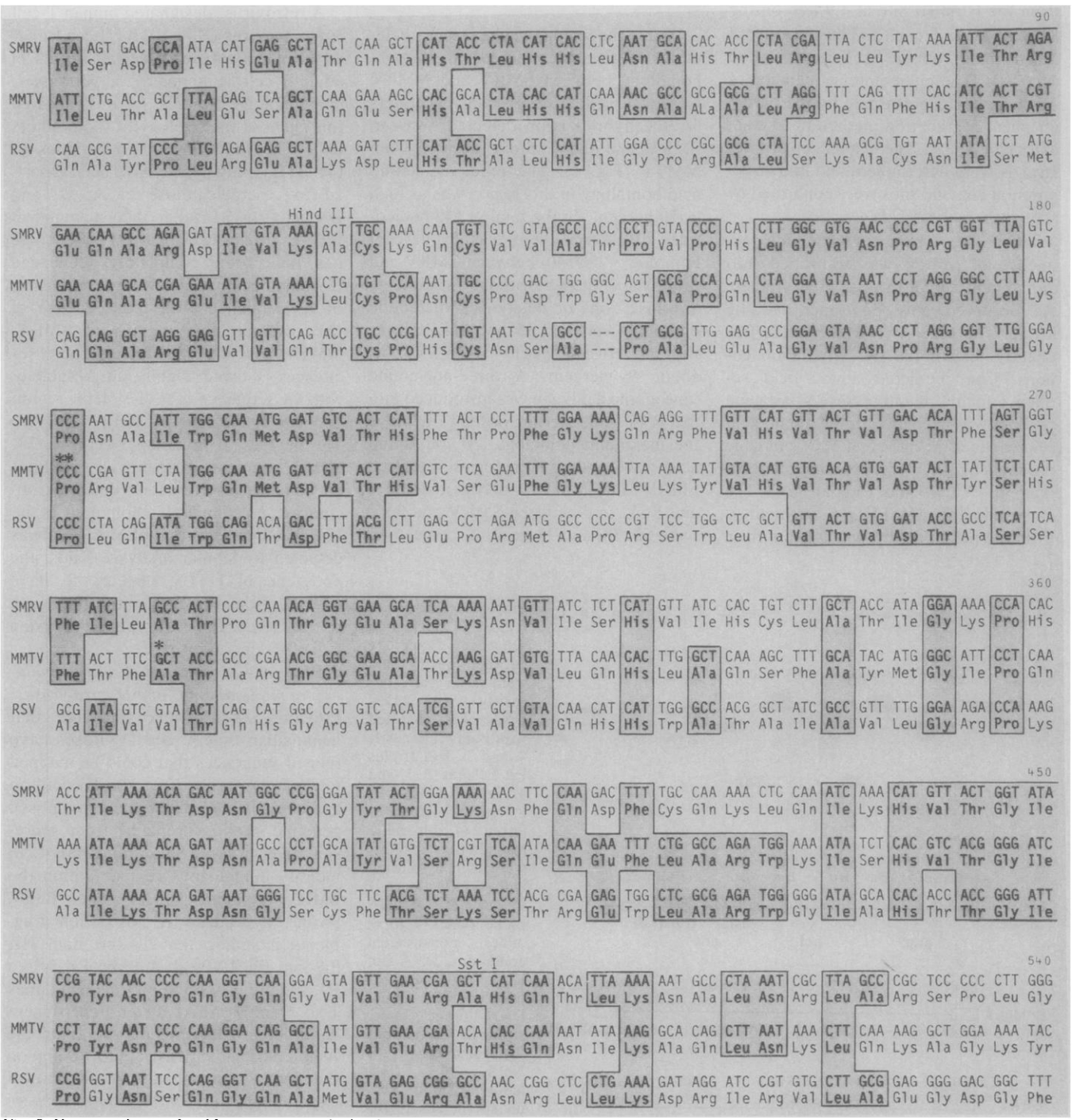


Fig. 5. Comparative nucleotide sequence analysis of the *pol* gene segments of SMRV, MMTV, and RSV. DNA sequences were determined by the method of Maxam and Gilbert (*22*) and translated into amino acids. The predicted amino acid sequences were aligned by using Wilbur and Lipman's PRTALN computer program (*24*). A K-tuple size of 1, window size of 20, and gap penalty of 4 were used in this analysis. The MMTV sequence determined by us differs from the published sequence of Redmond and Dickson (*21*) by the insertion of CC at positions 181 and 182 and the insertion of G at position 280. The differences, indicated in the MMTV sequences by asterisks, result in a frame shift of 32 amino acids as compared with the published MMTV sequence. However, the reading frames before nucleotide 181 and after nucleotide 280 are not altered. The symbols for the nucleotides are A, adenine; C, cytosine; G, guanine; and T, thymine. The symbols for the amino acids are Ala, alanine; Arg, arginine; Asn, asparagine; Asp, aspartic acid; Cys, cysteine; Glu, glutamic acid; Gln, glutamine; Gly, glycine; His, histidine; Ile, isoleucine; Leu, leucine; Lys, lysine; Met, methionine; Phe, phenylalanine; Pro, proline; Ser, serine; Thr, threonine; Trp, tryptophan; Tyr, tyrosine; and Val, valine. Autoradiograms of the DNA sequences are available upon request.

of the SMRV *pol* gene for comparative purposes, even though relaxed hybridization analyses of this region did not show detectable homology with RSV (Fig. 3B); this region of the RSV *pol* gene encodes the pp32 peptide that has endonuclease activity (*22*). The SMRV segment extended from 115 bp upstream from the Hind III site at 4.9 kbp on the SMRV physical map to 53 base pairs (bp) downstream from the Sst I site at 5.3 kbp (Fig. 3B). The nucleotide sequencing of this region was performed by the partial chemical degradation method of Maxam and Gilbert (*23*). The sequences of both strands were determined, and known restriction enzyme sites were confirmed by sequence analysis (*11*).

When the primary nucleotide sequence of SMRV was translated into amino acids in all three reading frames, we observed only one open reading frame. The other two frames were frequently interrupted by termination codons. The predicted amino acid sequence for the putative SMRV *pol* gene product was aligned with the amino acid sequences of MMTV and RSV *pol* gene products by the PRTALN computer program of Wilbur and Lipman (*24*). The corresponding MMTV *pol* segment is from the position 3 bp downstream from the single Eco RI site to the position 261 bp upstream from the putative termination codon (*21*). The analogous sequences of the RSV genome extended from nucleotide 4210 to 4746 (*19*). The first codon after the single Eco RI site within the MMTV genome was arbitrarily assigned as the first nucleotide (Fig. 5).

Within the region encompassing nucleotide 193 to nucleotide 303, 69 percent homology between SMRV and MMTV was observed. This probably accounts for the hybridization detected between SMRV and the C probe derived from MMTV (Fig. 2). The predicted amino acid homology in this region was as high as 70 percent. In the total of 540 nucleotides displayed, there was 55 percent homology (296/540) between SMRV and MMTV *pol* genes without introducing any insertions or deletions. Nucleotide homology between SMRV and RSV, as well as between MMTV and RSV, was about 46 percent. A three-nucleotide (one amino acid) gap was introduced into the RSV sequence in order to achieve the maximum homology. At the amino acid level, the homologies between SMRV and MMTV, between MMTV and RSV, and between SMRV and RSV were 52, 38, and 38 percent, respectively. These findings firmly establish the related but divergent nature of the *pol* genes of these viruses. In contrast, this analysis detected no comparable homology at either nucleotide or amino acid sequence levels between any of these *pol* genes and that of M-MuLV (data not shown).

A retrovirus, designated human T-cell leukemia virus (HTLV), has been isolated from patients with certain forms of adult T-cell leukemia (*25*). Moreover, epidemiologic studies have implicated this agent in the etiology of such tumors (*26*). The functional characteristics of the HTLV polymerase have been reported to resemble those of A, B, and D oncoviruses, rather than mammalian type C viruses (*27*). The HTLV genome has now been completely sequenced (*28*), making possible direct comparisons of its sequence with the sequence of the corresponding regions of SMRV, MMTV, and RSV *pol* genes. These sequences showed significant homology with the sequence of HTLV in the region between amino acids 57 and 171 (Fig. 5). The degrees of homology were 46, 38, and 39 percent, respectively, without introducing insertions or deletions. In contrast, no comparable homology of the HTLV *pol* gene to that of M-MuLV was detected by similar analysis. Thus, the *pol* gene of HTLV appears to have evolved from the progenitor that gave rise to the *pol* genes of types A, B, avian C, and D oncoviruses.

Env *gene recombination in the evolution of mammalian C and D retroviruses.* Previous studies have indicated that mammalian type C and D viruses have related sequences that could be mapped to their *env* genes (*29, 30*). To confirm and extend these observations, we used genomic probes prepared from representative viruses, SSAV and BaEV, to identify homologous regions within the SMRV genome. SSAV and BaEV probes hybridized to restriction fragments including the 2.9-kbp Bam HI, 3.8-kbp Bgl II–Pst I, 1.4-kbp Pst I–Eco RI, and the 2.3-kbp Sph I fragment (lanes 1 to 6 in Fig. 6A) under nonstringent conditions. These results identified the region of homology as 5.45 to 7.3 kbp on the SMRV map (Fig. 6B). Reciprocal experiments with SMRV as a probe were consistent with these results (lanes 10 to 13 in Fig. 6A). As summarized in Fig. 6B, the region of homology between SMRV and R-MuLV and between SMRV and SSAV genomes corresponded to their *env* gene sequences.

Under stringent hybridization conditions (lanes 7 to 9 in Fig. 6A), the region of homology between BaEV and SMRV was reduced to a 0.3-kbp stretch from

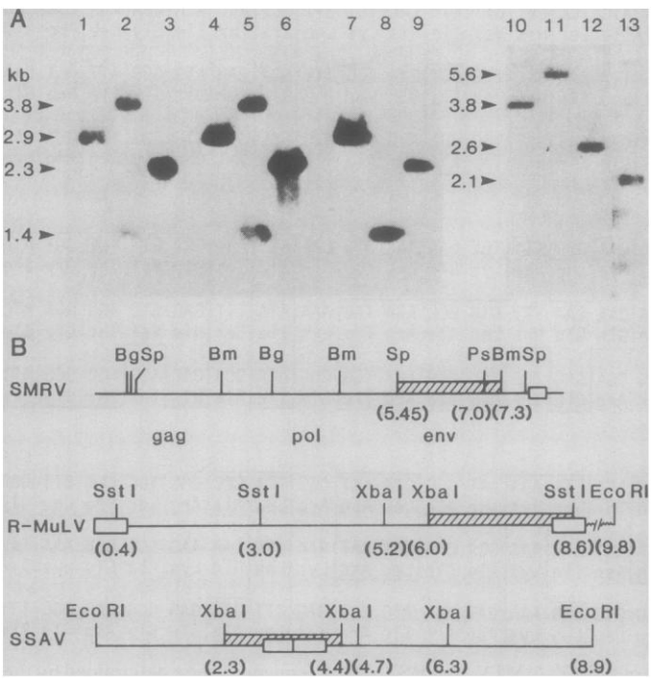


Fig. 6. Localization of regions of nucleotide sequence homology between SMRV and mammalian type C retroviruses. (A) SMRV DNA was digested with Bam HI (lanes 1, 4, and 7), Bgl II plus Pst I (lanes 2, 5, and 8), or Sph I (lanes 3, 6, and 9) and transferred to nitrocellulose filters. The filters were then hybridized with SSAV (lanes 1 to 3) or BaEV DNA probe (lanes 4 to 6) under nonstringent conditions. DNA's of lanes 7 to 9 were hybridized with BaEV DNA probe under stringent conditions [hybridization in 1*M* NaCl, 10× Denhardt's solution, salmon sperm DNA (100 μg/ml), 40 percent formamide, and probe (1×10^7 cpm/ml) at 37°C overnight, then washing in triple-strength standard saline citrate, 0.1 percent sodium dodecyl sulfate, and 0.1 percent sodium pyrophosphate at 64°C]. R-MuLV DNA was digested with Xba I (lane 10), Sst I (lane 11), or Xba I plus Sst I (lane 12), and SSAV DNA was digested with Xba I (lane 13). The filters were then hybridized with SMRV DNA probe under relaxed conditions. (B) The resulting hybridized regions are shown as hatched boxes. Open boxes represent LTR's. The direction of transcription for each virus is from left to right. The region of homology detected between SMRV and SSAV or BaEV under nonstringent conditions spans the region between the Sph I and the Bam HI sites (5.45 to 7.3 kb) on the SMRV physical map. Under stringent conditions with the BaEV probe, the homologous region is present between the Pst I and Bam HI sites (7.0 to 7.3 kb) on the SMRV map.

Pst I to Bam HI, corresponding to 7.0 kbp to 7.3 kbp on the SMRV map (Fig. 6B); in contrast, there was no detectable homology between SSAV and SMRV under the same hybridization conditions (data not shown). These findings indicate that the type D and mammalian type C *env* genes diverged from a common progenitor but that SMRV and BaEV share a more highly conserved sequence.

In order to better define this unusual relationship, we performed heteroduplex analysis of the SMRV and BaEV genomes. A representative heteroduplex is shown in Fig. 7A. We observed a continuous 300-bp stretch of homology (segment 4 in Fig. 7B). Since the orientation of the two clones within their plasmid vectors was known (data not shown), the homologous region could be assigned to the p15E coding segment of BaEV (Fig. 7C). These results strongly suggest that BaEV and SMRV became evolutionarily linked by a more recent recombinational event involving their p15E coding regions.

Discussion. Oncoviruses exhibit a unique association with vertebrates. These agents are transmitted as endogenous viral sequences within the germ lines of a wide variety of species. Moreover, in some cases oncoviruses have been established as infectious agents responsible for the induction of naturally occurring tumors of animals, probably including man. Earlier studies have shown that mammalian type C viruses share antigenic determinants among several of their respective *gag*, *pol*, and *env* gene products; this implies that these viruses have a common progenitor (*9*). The major structural proteins encoded within the *gag* gene of type C and D oncoviruses share antigenic determinants, as do the major proteins of type B and D oncoviruses (*10*). Moreover, certain mammalian type C and type D oncoviruses share common antigenic determinants in their *env*-coded gp70 protein (*29*), as well as sequence homology in their p15E coding region (*30*). These patterns of homology argue that the evolution of present oncovirus groups has occurred through genetic interactions among their progenitors.

The promiscuous nature of the oncovirus genome is strikingly illustrated by our present studies. Under appropriate hybridization conditions, distant but significant homology could be demonstrated between type C virus *env* genes and that of SMRV, findings consistent with their having a common *env* gene progenitor. By use of stringent hybridization conditions, we also detected a much greater degree of homology between the p15E coding regions of BaEV and SMRV than is observed between either of these viruses and the *env* genes of other type C viruses (Fig. 6). By heteroduplex analysis it was possible to demonstrate a region of complete colinearity between the corresponding p15E coding regions of BaEV and SMRV. These results strongly imply that BaEV and SMRV became evolutionarily more closely linked by a relatively recent recombinational event involving their p15E coding regions.

In contrast to the relationships among *gag* or *env* genes that have been extensively studied by immunological methods, relatively little is known about the genetic relationships among *pol* genes of different oncovirus genera. Our present studies establish by both molecular hybridization analysis and nucleotide sequence comparison that the *pol* genes of prototype type A, type B, avian type C, and type D oncovirus genera are all genetically related. These findings correlate well with the known functional similarities in the reverse transcriptases coded for by their *pol* genes. These enzymes exhibit similar template and cation preferences that differ markedly from those of mammalian type C viruses (*31*). Thus, our studies provide strong support for the concept of a major *pol* gene progenitor for mammalian type C viruses and another for types A, B, avian type C, and D oncoviruses.

Using computer analysis to search for local, dense homology [LSRCHP program (*24*)] beween the most conserved regions of these four oncoviral *pol* genes and other published amino acid sequences, we were able to detect homology with the M-MuLV *pol* gene. This region, corresponding to amino acids 117 to 132 (Fig. 5) is as follows (*32*):

SMRV:	G	K	P	H	T	I	K	T	D	N	G	P	G	Y	T	G
MMTV:	G	I	P	Q	K	I	K	T	D	N	A	P	A	Y	V	S
RSV:	G	R	P	K	A	I	K	T	D	N	G	S	C	F	T	S
HTLV:	G	K	P	S	Y	I	N	T	D	N	G	P	A	Y	I	S
M-MuLV:	G	M	P	Q	V	L	G	T	D	N	G	P	A	F	V	S

Amino acids in each sequence matching the consensus sequence, GKPQXIKTDNGPAYTS, are underlined. The sequence for each virus in this region corresponds to the consensus sequence in 10 to 12 of the 16 positions. To evaluate the statistical significance of similarities to the consensus sequence, we used the SEQDP program (*33*). The program com-

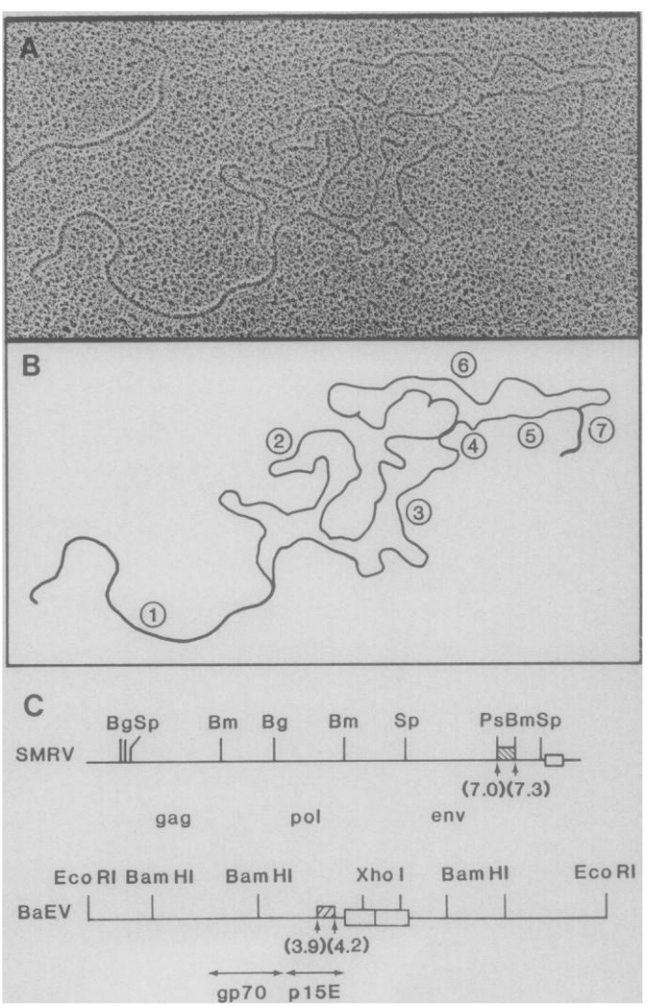


Fig. 7. Heteroduplex analyses of SMRV and BaEV. Supercoiled plasmid DNA's containing Eco RI–permuted SMRV (8.4 kbp) and BaEV (8.8 kbp) DNA were linearized by digestion with Pvu I. The heteroduplexing reaction was performed as described by Davis *et al.* (*37*). A representative heteroduplex is displayed in (A). Contour lengths shown in the schematic representation (B) were based on measurements of 20 molecules, and are as follows: region 1, 3.8 ± 0.2; region 2, 7.0 ± 0.4; region 3, 3.9 ± 0.3; region 4, 0.3 ± 0.08; region 5, 1.1 ± 0.1; region 6, 4.6 ± 0.4; and region 7, 0.6 ± 0.04 kb. Regions 1 and 7 are pBR322 DNA segments. The region of homology (region 4) between SMRV and BaEV is shown in (C) as a hatched box. The boundaries of gp70 and p15E genes of BaEV are demarcated by arrows (*30*). The open boxes represent LTR's.

pares a similarity score between two given sequences and then generates a set of reference scores by randomly shuffling the sequences and computing the similarity. The number of standard deviations from the mean score of the reference set may be used as an estimate of significance. The pairwise comparisons with the consensus sequence yielded values ranging from 8.4 to 13.7, which were judged to be highly significant. Whether this highly conserved region reflects the convergent evolution toward some important shared enzymatic function or an evolutionary remnant of a linkage between the highly diverged mammalian type C virus enzyme and other oncoviral reverse transcriptases is not known. In either case, this conserved region suggests an important function common to all oncoviral polymerases.

There has been conjecture that viral reverse transcriptases may have evolved from cellular progenitors. Our detection of a highly conserved sequence shared by all known oncovirus *pol* genes may be of interest in this regard. Antibodies directed against peptides deduced from this conserved region or specific DNA probes based on the sequence itself might be used in efforts to identify homology between the viral reverse transcriptase and DNA polymerases of cellular origin.

References and Notes

1. F. Fenner, *Virology* **71**, 371 (1976).
2. G. Schidlovsky, in *Cancer Research: Cell Biology, Molecular Biology and Tumor Virology*, R. Gallo, Ed. (CRC Press, Cleveland, 1977), pp. 189–245; A. Dalton, *J. Natl. Cancer Inst.* **49**, 323 (1972).
3. W. Biczysko, M. Pienkowski, D. Solter, H. Koprowsky, *J. Natl. Cancer Inst.* **51**, 1041 (1973); P. G. Callarco and D. Szöllösi, *Nature (London)* **243**, 91 (1973); P. G. Chase and L. Piko, *J. Natl. Cancer Inst.* **51**, 1971 (1973); A. J. Dalton, M. Potter, R. M. Merwin, *ibid.* **26**, 1221 (1961); K. Perk and J. E. Dahlberg, *J. Virol.* **14**, 1304 (1974).
4. R. Callahan, E. L. Kuff, K. K. Lueders, E. Birkenmeier, *J. Virol.* **40**, 901 (1981).
5. D. H. Moore in *Cancer: A Comprehensive Treatise*, F. F. Becker, Ed. (Plenum, New York, 1975), vol. 2, pp. 131–167.
6. N. Teich, J. Wyke, T. Mak, A. Bernstein, W. Hardy, in *Molecular Biology of Tumor Viruses: RNA Tumor Viruses*, R. Weiss, N. Teich, H. Varmus, J. Coffin, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., ed. 2, 1982), pp. 785–998; L. Gross, *Oncogenic Viruses* (Pergamon, Oxford, ed. 2, 1970).
7. R. L. Heberling, S. T. Barker, S. S. Kalter, G. C. Smith, R. J. Helmke, *Science* **195**, 289 (1977).
8. G. Geering, T. Aoki, L. J. Old, *Nature (London)* **226**, 265 (1970); R. V. Gilden, S. Oroszlan, R. J. Huebner, *Nature (London) New Biol.* **231**, 107 (1971).
9. J. R. Stephenson and S. A. Aaronson, *J. Virol.* **12**, 564 (1973); M. Strand and J. T. August, *ibid.* **13**, 171 (1974); C. J. Scherr, L. A. Fedele, R. E. Benveniste, G. J. Todaro, *ibid.* **15**, 1440 (1975); M. Barbacid, J. R. Stephenson, S. A. Aaronson, *Cell* **10**, 641 (1977); J. M. Krakower, M. Barbacid, S. A. Aaronson, *J. Virol.* **22**, 331 (1977); S. A. Aaronson, M. Barbacid, S. Hino, S. R. Tronick, J. M. Krakower, in *Advances in Comparative Leukemia Research*, R. Bentvelzen, J. Hilgers, D. S. Yohn, Eds. (Elsevier/North-Holland, Amsterdam, 1978), pp. 127–134.
10. M. Barbacid, L. K. Long, S. A. Aaronson, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 72 (1980).
11. I. M. Chiu, P. R. Andersen, S. A. Aaronson, S. R. Tronick, *J. Virol.* **47**, 434 (1983).
12. J. E. Majors and H. E. Varmus, *Nature (London)* **289**, 253 (1981).
13. A. Habara, E. P. Reddy, S. A. Aaronson, *J. Virol.* **44**, 731 (1982).
14. E. P. Reddy, M. J. Smith, E. Canaani, K. C. Robbins, S. R. Tronick, S. Zain, S. A. Aaronson, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 5234 (1980); A. Srinivasan, E. Reddy, S. A. Aaronson, *ibid.* **78**, 2077 (1981).
15. S. G. Devare, E. P. Reddy, K. C. Robbins, P. R. Andersen, S. R. Tronick, S. A. Aaronson, *ibid.* **79**, 3179 (1982); E. P. Gelman, F. Wong-Staal, R. A. Kramer, R. C. Gallo, *ibid.* **78**, 3373 (1981).
16. N. Battula, G. Hager, G. J. Todaro, *J. Virol.* **41**, 583 (1982).
17. D. Shalloway, A. D. Zelentz, G. M. Cooper, *Cell* **24**, 531 (1981).
18. L. M. Souza, J. N. Strommer, R. L. Hillyard, M. C. Komaromy, M. A. Baluda, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 5177 (1980).
19. D. E. Schwartz, R. Tizard, W. Gilbert, *Cell* **32**, 853 (1983).
20. T. M. Shinnick, R. A. Lerner, J. G. Sutcliff, *Nature (London)* **253**, 543 (1983).
21. S. M. S. Redmond and C. Dickson, *EMBO J.* **2**, 125 (1983).
22. G. F. Gerard and D. P. Grandgenett, in *Molecular Biology of RNA Tumor Viruses*, J. R. Stephenson, Ed. (Academic Press, New York, 1980), pp. 345–395; P. K. Misra, D. P. Grandgenett, J. M. Parsons, *J. Virol.* **44**, 330 (1982).
23. A. M. Maxam and W. Gilbert, *Methods Enzymol.* **65**, 499 (1980).
24. W. J. Wilbur and D. J. Lipman, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 726 (1983).
25. B. J. Poiesz, F. W. Ruscetti, A. F. Gazdar, P. A. Bunn, J. D. Minna, R. C. Gallo, *ibid.* **77**, 7415 (1980); M. Yoshida, I. Miyoshi, Y. Hinuma, *ibid.* **79**, 2031 (1982).
26. W. A. Blattner *et al.*, *J. Infect. Dis.* **147**, 406 (1983).
27. H. M. Rho, B. Poiesz, F. W. Ruscetti, R. C. Gallo, *Virology* **112**, 355 (1981).
28. M. Seiki, S. Hattori, Y. Hirayama, M. Yoshida, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 3618 (1983).
29. J. R. Stephenson, S. Hino, E. W. Garrett, S. A. Aaronson, *Nature (London)* **261**, 609 (1976); S. G. Devare, R. E. Hanson, J. R. Stephenson, *J. Virol.* **26**, 316 (1978).
30. M. Cohen, N. Rice, R. Stephens, C. O'Connell, *J. Virol.* **41**, 801 (1982).
31. J. Schlom, in *Molecular Biology of RNA Tumor Viruses*, J. Stephenson, Ed. (Academic Press, New York, 1980), pp. 447–484; S. H. Wilson and E. L. Kuff, *Proc. Natl. Acad. Sci. U.S.A.* **69**, 1531 (1972); E. Scolnick, E. Rands, S. A. Aaronson, G. J. Todaro, *ibid.* **67**, 1789 (1970).
32. The one letter symbols for the amino acids are A, alanine; R, arginine; N, asparagine; D, aspartic acid; C, cysteine; Q, glutamine; E, glutamic acid; G, glycine; H, histidine; I, isoleucine; L, leucine; K, lysine; M, methionine; F, phenylalanine; P, proline; S, serine; T, threonine; W, tryptophan; Y, tyrosine; V, valine; and X, undecided.
33. W. B. Goad and M. I. Kanehisa, *Nucleic Acids Res.* **10**, 247 (1982).
34. E. M. Southern, *J. Mol. Biol.* **98**, 503 (1975).
35. D. T. Denhardt, *Biochem. Biophys. Res. Commun.* **23**, 641 (1966).
36. P. W. J. Rigby, M. Dieckmann, C. Rhodes, P. Berg, *J. Mol. Biol.* **113**, 237 (1977).
37. R. W. Davis, M. Simon, N. Davidson, *Methods Enzymol.* **21**, 413 (1971).

38. We thank S. Devare and P. Reddy for advice on nucleotide sequence analysis, D. Lipman for aid in computer analysis and helpful discussions, S. Skuntz for technical assistance, J. Dalhberg for critical reading of the manuscript, and O. Gritz for preparing the manuscript. We also thank J. Majors and H. Varmus for the MMTV clones GR7-1a and pMMTV (C3H)2.5, M. Cohen for the BaEV clone, G. Cooper for the RSV clone, and M. Baluda for the MAV subclone.

12 October 1983; accepted 6 December 1983