JOURNAL OF VIROLOGY, Dec. 1990, p. 5823–5832
0022-538X/90/125823-10$02.00/0

Vol. 64, No. 12

Polyomavirus Late Pre-mRNA Processing: DNA Replication-Associated Changes in Leader Exon Multiplicity Suggest a Role for Leader-to-Leader Splicing in the Early-Late Switch

ROBIN P. HYDE-DeRUYSCHER† AND GORDON G. CARMICHAEL*

Department of Microbiology, University of Connecticut Health Center, Farmington, Connecticut 06030

Received 6 June 1990/Accepted 28 August 1990

Polyomavirus late mRNAs contain at their 5′ ends multiple, tandem repeats of a 57-base noncoding sequence, the late leader, whose sequence appears only once in the viral genome. Pre-mRNA molecules are processed by a pathway that includes the splicing of late leader exons to each other in giant, multigenome-length precursors which are the result of inefficient transcription termination. We have devised a method involving reverse transcription and the polymerase chain reaction to determine the number of tandem late leader units on polyomavirus late RNA molecules. Using this technique, we have shown that each class of late viral mRNA (mVP1, mVP2, and mVP3) consists of molecules with between 1 and 12 tandem leader units at their 5′ ends. Importantly, single-leader RNAs are underrepresented in both the cytoplasm and the nucleus, suggesting that single-leader primary transcripts are preferentially degraded in the nucleus. In addition, the average number of leaders on late RNAs increases in the presence of DNA replication. Taken together with previous work from our laboratory, the results presented here are consistent with a model for the control of late gene expression at the level of RNA splicing and stability which is in turn controlled by the efficiency of transcription termination.

Polyomavirus is a small, double-stranded DNA virus whose genome consists of two transcription units and an intergenic regulatory region. Gene expression during lytic infection of permissive mouse cells proceeds in a well-defined temporally regulated manner (7, 24, 32, 36). Immediately after infection, RNA from the early transcription unit (E-RNA) begins to accumulate; however, RNA from the late transcription unit (L-RNA) accumulates more slowly. At 12 h after infection, the ratio of E-RNA to L-RNA is about 4:1 (7, 22, 24, 32, 36). At 12 to 15 h postinfection, viral DNA replication commences and L-RNA begins to accumulate rapidly while E-RNA accumulates more slowly. This switch results in a dramatic change in the relative abundances of E- and L-RNA; by 24 h postinfection, the ratio of E-RNA to L-RNA is about 1:50 (7, 22, 24, 32, 36). The switch is dependent on viral DNA replication; if replication is inhibited, E-RNA accumulates to abnormally high levels, with minimal accumulation of L-RNA (12, 14, 21–23).

How is the early-late switch controlled? Large T-antigen autoregulation of E-RNA accumulation certainly plays an important role (14, 35). However, L-RNA processing and stability are also important (22). To learn more about the role of L-RNA processing in the early-late switch, we have examined the events which are required for the production of polyomavirus late mRNAs, concentrating on the differences observed between infections under conditions where L-mRNA will (with DNA replication) or will not (without DNA replication) accumulate.

We have previously shown, by using a nuclear run-on assay, that the relative rates of early and late transcription do not change significantly throughout infection, even though the relative rates of RNA accumulation undergo dramatic changes (22). Therefore, the major control point for the accumulation of both E-RNA and L-RNA is likely to be at the posttranscriptional level. Although the relative strengths of the early and late promoters did not vary significantly throughout infection, one dramatic change in the late-strand transcription pattern was detected near the late poly(A) site: a higher level of transcription downstream of the late poly(A) site was found after the onset of DNA replication. One interpretation of this finding is that the change in run-on patterns is indicative of a decrease in transcription termination efficiency that is associated with 3′-end processing.

In addition to the transcription results, we found a parallel increase in the efficiency of one of the splices that is required for the production of mature mRNAs coding for the major structural protein, VP1 (22). This result was interpreted as suggestive of a change in the stability of the primary transcripts because extensive work on the splicing of polyomavirus late pre-mRNAs indicated that stability of L-RNAs is directly linked to their ability to undergo RNA splicing (4–6; N. L. Barrett, G. G. Carmichael, and Y. Luo, submitted for publication).

In this study, we wished to concentrate on a splicing event involved in the production of L-RNA, the leader-to-leader splice. Transcription from the late promoter of polyomavirus is unusual in that termination and polyadenylation are both inefficient at late times during infection (1, 2). This allows for the production of large multigenome-length transcripts which are resolved by the splicing of untranslated 57-nucleotide late leader exons to each other (3, 29, 37, 38). Previous work by us suggested that the termination efficiency for late transcription might change during the course of infection (22). This should result in the production of primary transcripts that are longer at late times than at early times. Since giant transcripts can be resolved only by leader-to-leader splicing, we set out to compare leader multiplicity and distribution in RNA isolated from infected cells grown in the absence and presence of a DNA synthesis

* Corresponding author.
† Present address: Lewis Thomas Laboratory, Department of Molecular Biology, Princeton University, Princeton, NJ 08544.

inhibitor, 1-β-D-arabinofuranosylcytosine (araC). The inhibition of replication essentially holds the virus in a early phase of infection, allowing RNA produced during this phase to accumulate to detectable levels. The results presented here show a dramatic change in leader multiplicity before and after DNA replication: in the absence of replication, most L-RNAs contain a single late leader unit at their 5′ ends; however, in the presence of replication, most L-RNAs contain two or more leaders. In addition, analysis of the distribution of leaders in both nuclear and cytoplasmic RNA from productively infected cells reveals that single-leader RNAs are present in smaller than expected amounts. This finding is consistent with the hypothesis that single-leader primary transcripts are less stable than multiple-leader transcripts, as has been suggested previously by investigators in our laboratory (4–6, 22). All of these results are consistent with a model which we present for the control of L-mRNA accumulation at the level of nuclear RNA stability.

MATERIALS AND METHODS

Cells and viruses. NIH 3T6 cells and polyomavirus 59RA were grown as described previously (22). Infections were carried out at a multiplicity of infection (MOI) of 1 or 20 and are indicated below and in the figure legends. In experiments in which DNA synthesis was inhibited, araC was added to the culture media to a final concentration of 20 μg/ml and the medium was replaced every 6 hours.

RNA PCR. Total nuclear and total cytoplasmic RNA were isolated by using guanidine isothiocyanate as described previously (22). Oligonucleotides specific for each late splice junction were synthesized by using a Milligen Biosearch Cyclone DNA synthesizer. Oligonucleotide 275 (5′-TATCA CCGTACAGCCTTG) is specific for mVP1. Oligonucleotide 276 (5′-TGAAAATTCACTTACTTG) is specific for mVP2. Oligonucleotide 277 (5′-GGTACCGCTGTATTCTTG) is specific for mVP3. Oligonucleotide 278 (5′-CTTGACATTT TCTATTTTAAG) will bind just upstream of the late leader and is complementary to cDNA synthesized for mVP1, mVP2, and mVP3. Reverse transcription with oligonucleotides 275, 276, 277, or an equimolar mixture of all three was carried out as follows. The appropriate amount of RNA was added to an extension mix to create final concentrations of 1× *Taq* polymerase buffer (10 mM Tris [pH 8.3], 50 mM KCl, 1.5 mM $MgCl_2$, 0.001% gelatin), 1 mM deoxynucleoside triphosphates, and 10 ng of each oligonucleotide per μl in 10 μl. Just prior to the addition of RNA, the extension mix was UV irradiated in a tissue culture hood for 5 min as described previously (34) to eliminate background signals due to any contaminating DNA or RNA in the mix. The RNA was added, and the tubes were heated at 90°C for 2 min, held at room temperature for 5 min, and finally cooled on ice for at least 2 min. Then 11 U of avian myeloblastosis virus reverse transcriptase (11 U/μl; Dupont, NEN Research Products) was added, and the mixture was incubated at room temperature for 5 min and then at 50°C for 30 min. An additional 11 U of reverse transcriptase was added, and the mixture was incubated for 30 min more at 50°C. A 40-μl volume of polymerase chain reaction (PCR) mix (1× Taq polymerase buffer containing 2.5 U of Amplitaq DNA polymerase [Cetus] and 1 ng of ^{32}P-end-labeled oligonucleotide 278, irradiated as above just before use) was then added, and each tube was topped with mineral oil. Temperature cycling was carried out in a Coy temperature cycler for the following times: 90°C, 2 min; 15 cycles of 90°C for 1 min, 30-s ramp to 45°C for 2 min, 72°C for 2 min, 30-s ramp back to 90°C, 15 cycles of 90°C for 1 min, 30-s ramp to 45°C for 2 min, 72°C for 5 min, 30-s ramp back to 90°C, 72°C for 10 min. The mixture was then held at 4°C. One-tenth of the amplified mixture was then analyzed by denaturing electrophoresis on 6% acrylamide–7 M urea gels.

Sucrose gradients. Nuclear RNA was fractionated essentially as described by Treisman and Kamen (38). Linear 5 to 20% sucrose gradients containing 50% deionized formamide were prepared for a Beckman SW41 rotor. A 100-μg portion of total nuclear RNA was resuspended in 50 μl of RNase-free H_2O, 50 μl of deionized formamide was added, and the sample was heated to 90°C for 2 min and then quick-chilled on ice. The sample was then carefully layered over the gradient and spun at 28,000 rpm at 20°C for 17 h. Fractions of 0.4 ml were collected from the top of the gradient by using an Isco model 640 density gradient fraction collector. The refractive index of each fraction was monitored, and a parallel gradient of cytoplasmic RNA was analyzed by agarose gel electrophoresis for the positions of 18S and 28S rRNA. The fractions were then analyzed by using PCR as described above.

RESULTS

Experimental design. The objective of this study was to examine the multiplicity (the number of leaders on individual L-mRNAs) and distribution (the percentage of the L-mRNA population with one, two, etc., leaders) of the 57-nucleotide late leader exon on polyomavirus L-RNA in the absence or presence of DNA replication. To accomplish this, we designed the strategy outlined in Fig. 1. This stepwise analysis was necessary because of numerous unusual features of polyomavirus L-mRNAs. The late promoter does not possess a TATA box, which results in the use of at least 15 start sites which span >150 bases (10, 13, 15, 16, 37–39). Primer extension analysis cannot be used to quantitate leader multiplicity because this technique gives a complex repeating pattern with heterogeneous start sites appearing with a spacing of 57 nucleotides (G. Adami and G. Carmichael, unpublished data). Our attempts to use S1 nuclease mapping or RNase protection with probes containing three tandem late leader exons have yielded unreliable and variable results owing to artifactual full-length probe protection by single-leader transcripts (C. Marlor and G. Carmichael, unpublished data). To circumvent these difficulties, we adopted a modified strategy containing an additional step outlined in Fig. 1. First, a 15-base oligonucleotide specific for a leader-to-body splice junction was hybridized and extended by reverse transcription. Next, a 21-base ^{32}P-5′-end-labeled oligonucleotide was hybridized to this cDNA and extended by using a DNA-dependent DNA polymerase. The second oligonucleotide hybridizes downstream of most of the start sites for the late promoter and removes the 5′-end heterogeneity. The resulting products were separated on denaturing polyacrylamide gels. RNAs with a single leader resulted in a band at 93 nucleotides; RNAs with multiple leaders generated a ladder with successive bands at intervals of 57 nucleotides.

The method outlined above proved ideal for further amplification of the DNA products by using the PCR. Such amplification would allow the examination of leader multiplicity in the absence of DNA replication. Also, the high temperature of extension (72°C) increased the specificity of the hybridization, lowered the background, and increased sensitivity. However, to verify that short DNA fragments (i.e., single-leader RNAs) are not preferentially amplified

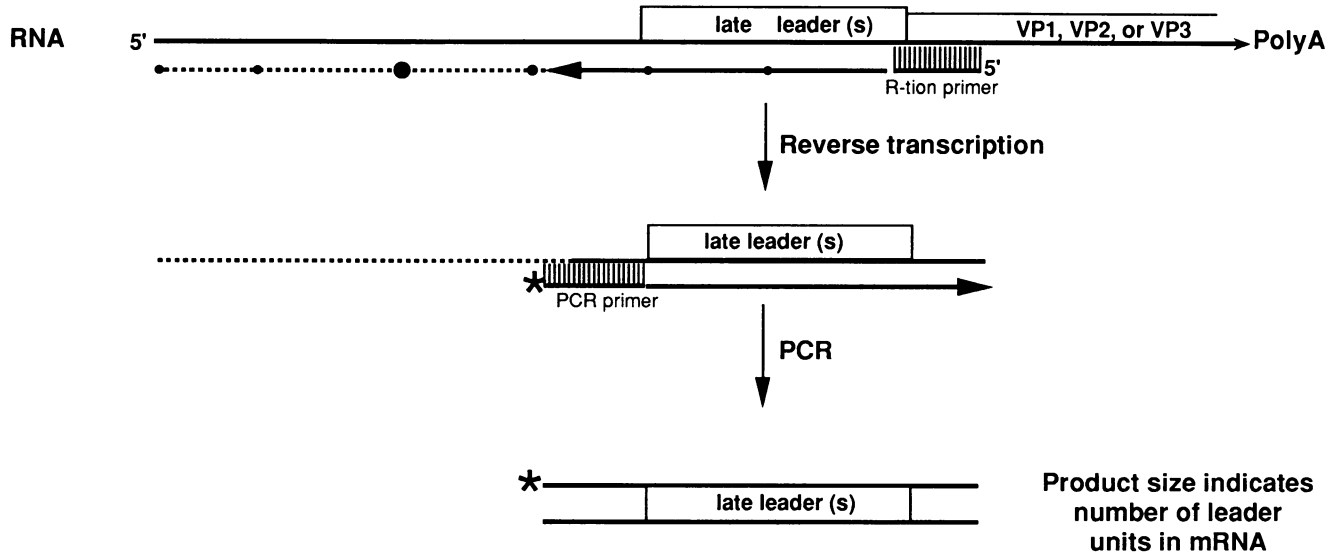
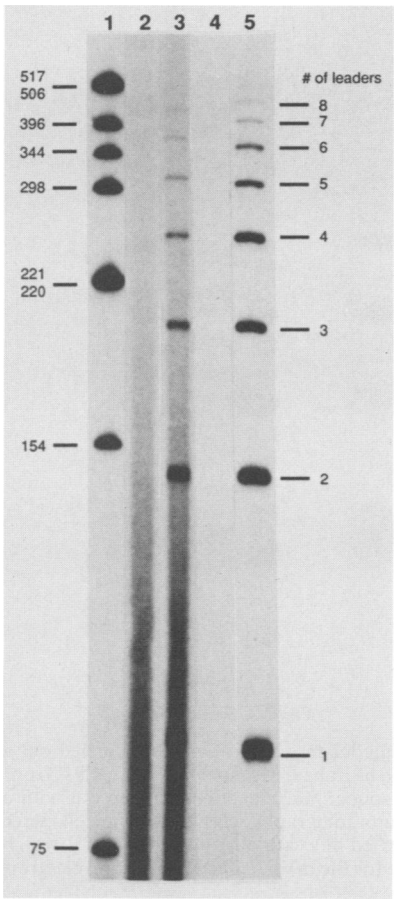


FIG. 1. Strategy used to determine the number of tandem late leader exons on the 5′ ends of polyomavirus mRNAs. The top line represents L-mRNA molecules. The dots represent the major start sites found for RNA produced from the late promoter, with the relative abundances (10, 13, 15, 16, 37–39) indicated by dot size. The large open rectangle represents the late leader structure, which can contain one or more tandem leader exons. The narrow open-ended rectangle represents one of the coding exons attached to the leader exon. The short line labeled R-tion primer represents an oligonucleotide specific for one of the leader-to-body splices (see Materials and Methods for sequences). The box labeled PCR primer represents an end-labeled oligonucleotide complementary to the cDNA. See text for details of the method.



under these conditions and that the PCR conditions yield reliable quantitative comparisons of mRNAs with different leader structures, we performed a test of validity by using Sequenase, a modified form of T7 DNA polymerase known to be extremely processive (30). In this experiment, temperature cycling (and hence amplification) was not used. Results comparing the amplification procedure with Taq DNA polymerase and one round of synthesis with T7 DNA polymerase are shown in Fig. 2, and quantitation of the gel is shown in Table 1. The patterns are extremely similar. Thus, under the conditions used here, the leader multiplicity of polyomavirus L-mRNAs can be accurately determined by using the PCR.

Multiple leader structures on mVP1, mVP2, and mVP3. Cytoplasmic RNA isolated at 36 h postinfection was analyzed for the presence of multiple leaders by using oligonucleotides specific for mVP1, mVP2, or mVP3. Each class of polyomavirus L-mRNAs contains members with multiple leaders, and the distribution of leaders on each class of mRNAs is virtually identical (Fig. 3). This is an important point because although mVP1 and mVP3 contain a leader-to-body splice, mVP2 does not. Therefore, all three polyomavirus L-mRNAs appear to be processed in a similar manner and leader-to-leader splicing appears to be independent of mRNA body splicing. mRNAs with up to 12 leaders have been detected in several experiments (data not shown).

FIG. 2. Comparison of late leader multiplicity and distribution detected by using Sequenase or Amplitaq. RNA was isolated from infected or mock-infected cells and subjected to reverse transcription, as described in Materials and Methods, by using oligonucleotide 275, which is specific for the mVP1 splice. Lanes: 1, DNA molecular size markers (in nucleotides); 2, Sequenase reaction with RNA from mock-infected cells; 3, Sequenase reaction with RNA from cells infected for 24 h at an MOI of 20; 4, PCR with RNA from mock-infected cells; 5, PCR with RNA from cells infected for 24 h at an MOI of 20. See Table 1 for quantitation.

TABLE 1. Leader distribution determined by using Sequenase or PCR

No. of leaders	Distribution (% of total)[a] from:	
	Sequenase[b]	Amplitaq
1	39[c]	26
2	30	35
3	17	17
4	9	14
5	5	7

[a] To calculate the percentage of the total, the signal for each band (after subtracting out the corresponding signal in the mock lane) was divided by the total signal detected.

[b] The leader distribution obtained by using Sequenase was quantitated by densitometry in a Bio-Rad model 620 densitometer.

[c] This number may be an overestimate as a result of the high background in this region of the gel.

Calculation of termination efficiency by using leader distribution. To quantitate the distribution of leaders on mVP1, mVP2, and mVP3, the gel shown in Fig. 3 was analyzed by using a Betagen Betascope 603 Blot Analyzer. The percentage of each species is shown in Table 2. Assuming that the relative level of each species is a direct result of transcription termination, and using the percentage of each species of RNA present (with one, two, three, etc., leaders), the termination efficiency for each traverse of the genome can be calculated. For the first round of transcription, the termination efficiency should be equal to the percentage of RNAs containing one leader. For the second and succeeding rounds, the termination efficiency should be equal to the percentage of RNAs in that round divided by the percentage of molecules containing a greater number of leaders. The termination efficiency for each round of transcription was calculated by using the leader distribution from mVP1, mVP2, and mVP3 and is presented in Table 2. The use of leader distribution to calculate the termination efficiency for mRNAs containing two or more leaders generates numbers which are very similar for each traverse of the genome. In addition, this efficiency is the same in any given infection regardless which of the mRNAs (mVP1, mVP2, or mVP3) is examined. This number is characteristic for RNA isolated from cells infected at a particular MOI: infection at a low MOI results in a higher termination efficiency, whereas infection at a high MOI yields a lower calculated efficiency. In addition, termination efficiency calculated from nuclear

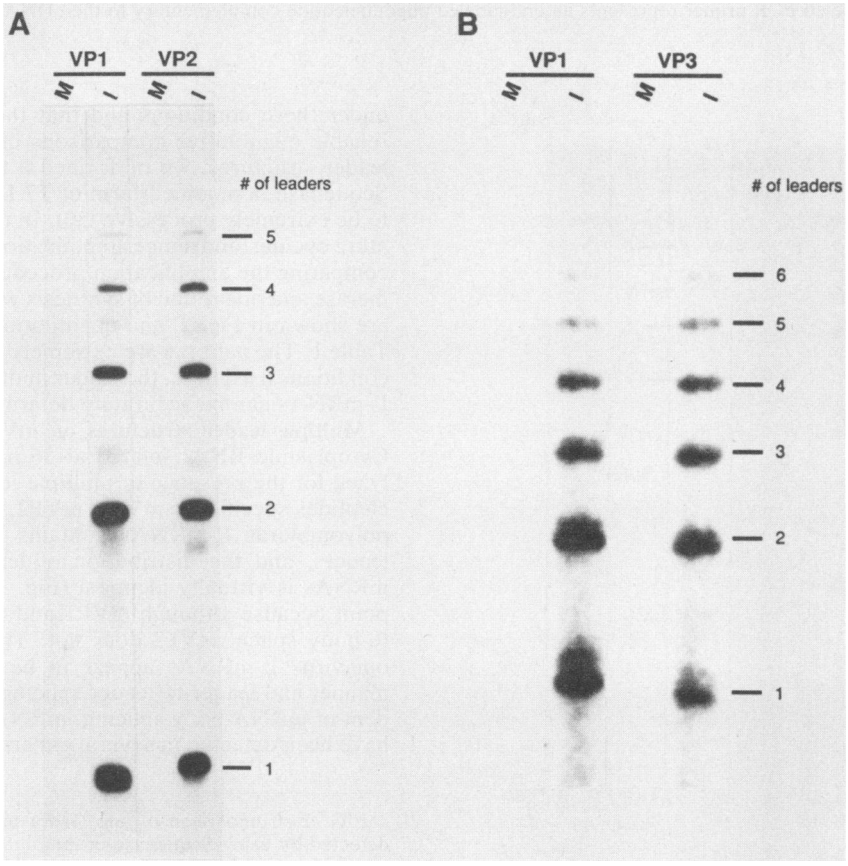


FIG. 3. All three classes of polyomavirus L-mRNAs contain members with multiple leaders. Analyses were carried out as described in Materials and Methods. (A) Cytoplasmic RNA from cells infected at an MOI of 1 for 36 h. Lanes: M, mock RNA; I, RNA from infected cells; VP1, Reactions carried out with oligonucleotide 275, which is specific for the mVP1 splice; VP2, reactions carried out with oligonucleotide 276, which is specific for mVP2. Following electrophoresis, bands were revealed by autoradiography. (B) RNA from cells infected at an MOI of 20 for 24 h. Lanes: M, mock RNA; I, RNA from infected cells; VP1, reactions carried out with oligonucleotide 275, which is specific for the mVP1 splice; VP3, reactions carried out with oligonucleotide 277, which is specific for the mVP3 splice. Following electrophoresis, bands were visualized and quantitated by using a Betagen Betascope 603 Blot Analyzer.

TABLE 2. Leader distribution of polyomavirus late cytoplasmic RNAs

No. of leaders	% Total distribution[a]							
	Infection 1[b]		Infection 2[b]		Predicted[c]	Mix −araC[d]	Mix +araC[d]	Nuclear (fraction 7)[e]
	mVP1	mVP2	mVP1	mVP3				
1	34 (34)	31 (31)	30 (30)	19 (19)	74	35	72	53 (53)
2	49 (74)	45 (66)	29 (42)	28 (35)	19	37	16	33 (70)
3	12 (69)	15 (64)	19 (46)	23 (43)	5	17	7	9 (64)
4	4 (79)	7 (78)	15 (66)	19 (62)	1.3	7	3	4 (80)
5	1	2	8	11	0.3	3	1	1

[a] The percentage of each species was calculated as in Table 1. Values in parentheses represent the calculated termination efficiency of that round of transcription. The termination efficiency for each sample was calculated by using the formula $\%T_n$ = percentage of total RNA with n leaders/[100 − Σ(percentage of total in single-leader RNA to percentage of total in RNAs containing $n-1$ leaders)], where n is a whole number between 1 and 5.

[b] mVP1 and mVP2 infection 1 is from Fig. 3A; mVP1 and mVP3 infection 2 is from Fig. 3B. Infection 1 was carried out at an MOI of 1, and infection 2 was carried out at an MOI of 20.

[c] Calculated by assuming a 74% termination efficiency, which was the average determined from mVP1, infection 1. At each round, the total remaining percentage was multiplied by 0.74 to determine the fraction predicted which should be present in each species.

[d] Fig. 5 (Mix ± araC).

[e] Fig. 6 (nuclear; lane 7 only). Quantitation was carried out by using a Betagen Betascope Blot Analyzer.

run-on analysis gives comparable numbers to those obtained here (compare infection 2 with results in reference 22) and is in the same range calculated by using in vivo pulse-labeling of nascent late transcripts (27, 28, 35, 39). Thus, this appears to be a valid method of determining the efficiency of termination in the polyomavirus system.

Underrepresentation of single-leader mRNAs. When using the leader distribution from mVP1 in infection 1, an average of 74% of the polymerases terminate each traverse of the genome two to five times. By using this average and assuming a simple model for transcription termination in which a constant percentage of polymerase molecules stop each time a terminator element is encountered, a predicted distribution of leaders was calculated and is presented along with the observed distributions in Table 2. To better illustrate the close correlation between the predicted and observed distributions in RNAs containing two to five leaders, we artificially matched the relative contribution of two-leader RNAs by multiplying all of the predicted distribution percentages by 2.57. The results are shown graphically in Fig. 4. It is apparent that for RNAs containing two or more leaders, the predicted and observed distributions were nearly identical. However, the relative contribution of one-leader mRNAs was depressed approximately 5.5-fold from the predicted level. This dramatic difference between the predicted and observed amounts is consistent with the hypothesis that RNAs containing one leader are less stable than multiple-leader RNAs or are derived from less-stable precursors (6).

Changes in leader multiplicity in the absence of DNA replication. The main objective of this study was to look for any changes in the processing of late transcripts during the course of infection, with the hope that any differences might

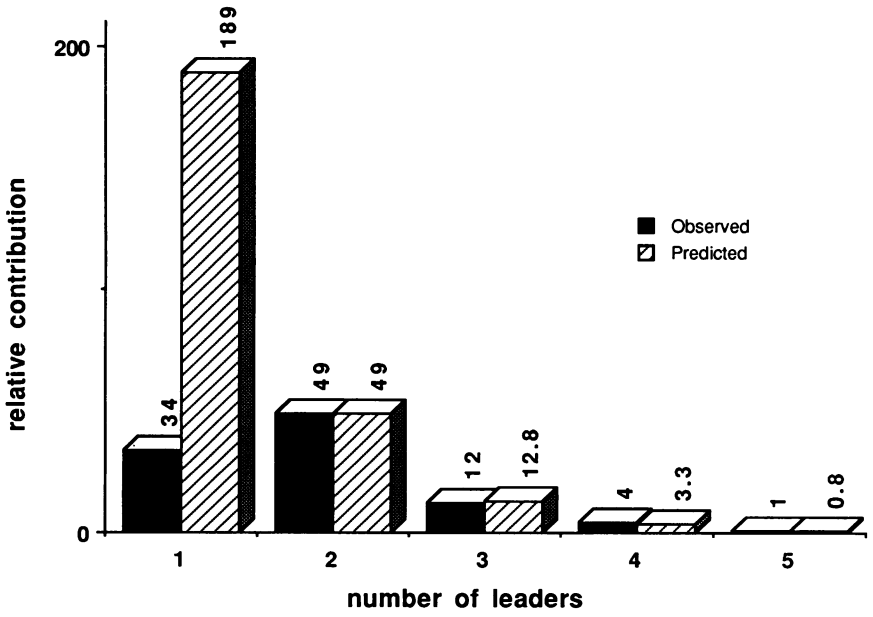


FIG. 4. Observed and predicted leader distributions for mVP1 molecules. The observed distribution of mVP1 molecules with different numbers of late leaders analyzed in Fig. 3A is plotted along with the predicted distribution, assuming no bias against single-leader mRNAs and a 74% termination efficiency. To best illustrate the congruence of the observed and predicted distributions, the percentages of the predicted distributions were multiplied by 2.57 to match exactly the relative contributions of two-leader RNAs.

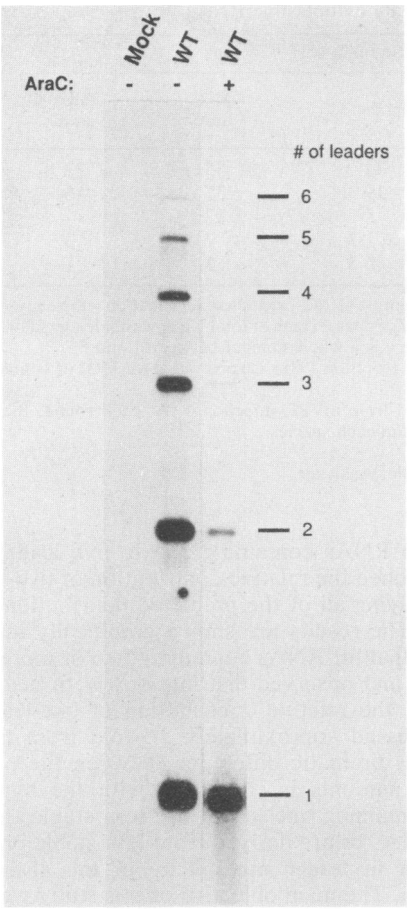


FIG. 5. Leader multiplicity in the presence or absence of DNA replication. Analyses were carried out as described in Materials and Methods, with an equimolar mixture of oligonucleotides 275, 275, and 277, which will detect all L-mRNAs. Lanes: Mock, RNA from mock-infected cells; WT, RNA isolated from cells infected at an MOI of 1 for 36 h. + and − indicate whether cells were maintained after infection in the presence or absence of 20 μg of araC per ml.

shed light on the early-late switch. A useful way to amplify the difference between the early stage of infection (before DNA replication) and late stages (after DNA replication) is to infect cells and to isolate RNA 36 h after infection from cultures maintained in either the absence or presence of araC, a DNA synthesis inhibitor. Without DNA replication, the early phase of infection is extended, allowing for a higher level of accumulation of early-phase RNA.

By using the RNA PCR technique described above, leader multiplicity was determined for polyomavirus cytoplasmic RNA isolated from infected cells grown for 36 h in the absence or presence of araC. An equimolar mixture of oligonucleotides 275, 276, and 277 was used to assess leader distribution on all three polyomavirus L-mRNAs. The results shown in Fig. 5 and Table 2 clearly demonstrate that in the absence of DNA replication, most of the L-RNAs that accumulate have a single late leader at their 5′ ends. In contrast, if DNA replication is not blocked, a majority of the polyomavirus mRNAs contain two or more leaders.

Underrepresentation of single-leader L-RNAs in the nucleus. It was possible that the deficit of single-leader RNAs in the cytoplasm was due to less efficient transport from the nucleus. Alternatively, single-leader RNAs might be preferentially degraded. To distinguish between these models, we examined the distribution of leaders on polyomavirus nuclear RNAs. However, we wished to look at molecules that had been processed to mature size: giant primary transcripts and any splicing intermediates would cloud analysis because unspliced late leader exons would be detected as RNAs containing a single leader. To assess leader multiplicity and distribution on mature-sized nuclear polyomavirus mRNAs, nuclear RNA isolated 36 h after infection (from infection 1) was separated on linear 5 to 20% sucrose gradients containing 50% formamide. Thirty 0.4-ml fractions were collected and analyzed for multiple leaders by using an oligonucleotide specific for mVP1. The results are shown in Fig. 6. Leader-to-leader splicing and leader-to-VP1 splicing can be detected throughout the gradient. Whether these represent true RNA processing intermediates or are only RNAs smeared down the gradient in the process of collection awaits further experimentation. Importantly, however, quantitation of RNAs that are the size of mature mVP1 (fraction 7) shows that the level of single-leader RNAs is still lower than that predicted by termination efficiency (Table 2). Thus, both cytoplasmic and nuclear RNAs display relatively low levels of single-leader RNAs. This is consistent with the hypothesis that primary transcripts that cannot be spliced leader-to-leader are preferentially degraded in the nucleus (6).

DISCUSSION

In this work we have examined in detail the multiplicity and distribution of late leader exons on polyomavirus L-RNAs. We have developed a method to determine the number of late leaders on mRNAs by using reverse transcription and the PCR. A comparison of this method involving 30 cycles of the PCR and a single round of second-strand synthesis with Sequenase shows that under the conditions used here, PCR does not preferentially amplify short molecules (100 nucleotides versus 500 nucleotides). Therefore, this is a convenient, rapid, and sensitive method of examining leader structures on polyomavirus L-RNAs.

An examination of cytoplasmic RNA from polyomavirus-infected cells reveals that all three L-mRNA types (mVP1, mVP2, and mVP3) can contain between 1 and 12 tandem leader units at their 5′ ends. This is the first direct demonstration that mVP2 contains multiple leaders. Treisman (37) showed by sequencing cDNAs that mVP1 and mVP3 can both contain multiple leaders; however, a cDNA for mVP2 was not found. Legon et al. (29) determined that the average number of leaders on each of the three L-mRNAs was three to five by using RNase T_1 fingerprinting of labeled species isolated after a high-multiplicity infection. Owing to the difficulty in separating mVP3 and mVP2, it was not clear whether the multiple leaders detected in these experiments were from mVP2 or from contaminating mVP3. The use of oligonucleotides specific for the splice junctions of each mRNA in the original reverse transcription step and in the PCR circumvents this problem. In this study we found that under the conditions used in our laboratory, the average leader multiplicity on each of the three L-mRNAs during the late phase of infection was two to three. Although the use of *Taq* DNA polymerase and Sequenase led to similar results, it is still possible that reverse transcriptase converts short mRNAs (single-leader RNAs) to DNA more efficiently than it converts longer mRNAs (multiple-leader mRNAs), biasing the results toward single leaders. If this were true, the underrepresentation of single-leader RNAs late after infection would be even more dramatic than shown here.

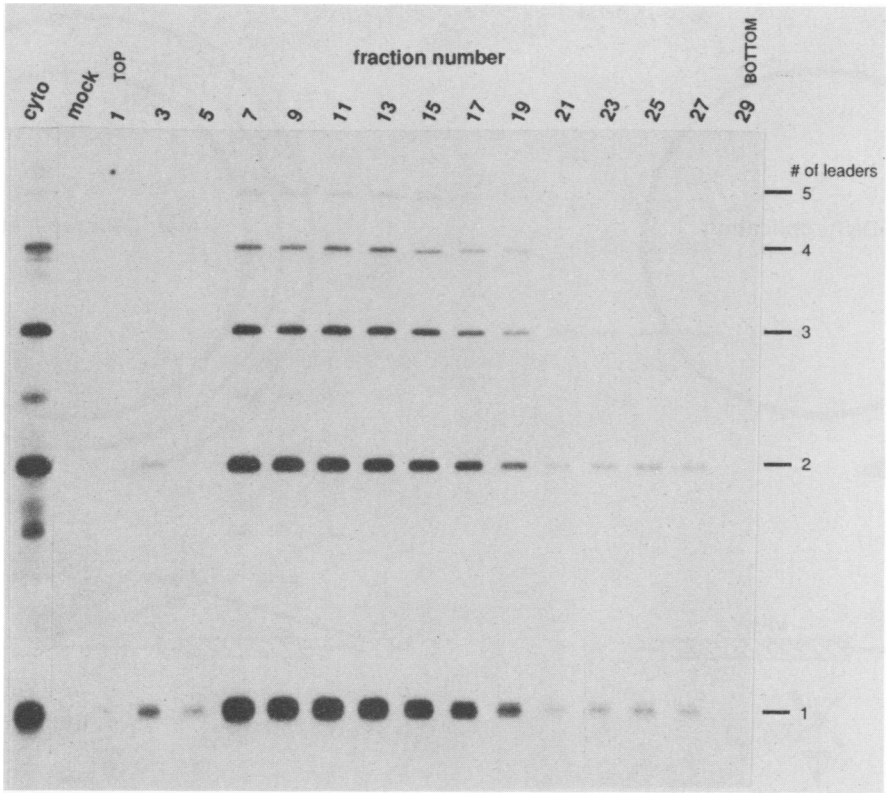


FIG. 6. Leader multiplicity in size-fractionated nuclear RNA. Nuclear RNA from infected NIH 3T6 cells was separated on linear 5 to 20% sucrose gradients, and fractions were analyzed for multiple leaders, as described in Materials and Methods, with oligonucleotide 275, which is specific for the leader-to-VP1 splice junction. Cyto indicates cytoplasmic RNA from the same infection; mock indicates the analysis carried out on nuclear RNA from mock-infected cells. Fraction 1 is the top of the gradient, and fraction 29 is the bottom. Mature mVP1 molecules sedimented to fractions 7 to 9.

The distribution of late leaders can be used to calculate a termination efficiency for late transcription. Assuming that the stability of the RNAs is the same, the percentage of single-leader RNAs represents the fraction of transcription complexes terminating the first time around the genome, the percentage of RNAs containing two leaders represent those that terminated the second time around, and so on. For polyomavirus, the termination efficiency calculated in this manner yields efficiencies which are similar regardless of which mRNA (mVP1, mVP2, or mVP3) is used to calculate this efficiency. A comparison of the termination efficiency calculated by using the RNA PCR method gives results in agreement with the efficiency calculated by using nuclear run-on analysis (compare infection 2 in Table 2 with results in reference 22) and with results of the sensitive method of Tseng and Acheson (39). A polyomavirus mutant with a more efficient late poly(A) site (about 90% efficient) has been constructed by Lanoix et al. (28). We have tested this mutant in our PCR assay (R. Hyde-DeRuyscher and G. Carmichael, unpublished results), and our results show a 90% drop between successive leader bands, again arguing for the validity of the method. Therefore, this method could be a rapid, sensitive assay to investigate the sequence requirements of termination or polyadenylation.

Interestingly, there appears to be a correlation between leader multiplicity and MOI. In RNA isolated from cells infected at an MOI of approximately 1, the greatest number of leaders that could be detected was six to seven (Fig. 3A). However, infection at an MOI of 10 to 20 allowed for the production of up to 12 leaders on the 5′ end of polyomavirus L-RNAs (Fig. 2, lane 5). This has been a reproducible finding. Thus, infecting cells at a lower MOI appears to result in more efficient termination, a finding that is consistent with the proposal that a limiting factor is titrated out at late times after infection (26, 28).

Regardless of the efficiency of reverse transcription, the distribution of leaders on polyomavirus RNAs show that cytoplasmic mRNAs containing only one leader are less abundant than predicted by a simple model of transcription termination. However, the distribution of mRNAs containing two to five leaders fits this predicted distribution extremely well. The distribution of leaders on mRNA-sized transcripts in the nucleus also reveals a paucity of single-leader mRNAs. One way to account for the observed distributions is to assume that the splicing of polyomavirus L-mRNAs occurs in a random fashion where leader exons can splice to other leaders (or coding body exons) far downstream. Other studies from this laboratory, to be presented elsewhere (Y. Luo and G. Carmichael, submitted for publication), have indicated that the vast majority of splicing for polyomavirus L-mRNAs is sequential; leaders splice to the nearest downstream leader, and coding exons are spliced to the nearest upstream leader. In addition, the distribution of leaders on mVP2 cannot be explained by a random-splicing model because these RNAs do not contain a leader-to-body splice. The only way a single leader mVP2 molecule can be generated is by the use of first the poly(A) site downstream of the 5′ end of the mRNA. Therefore, the data

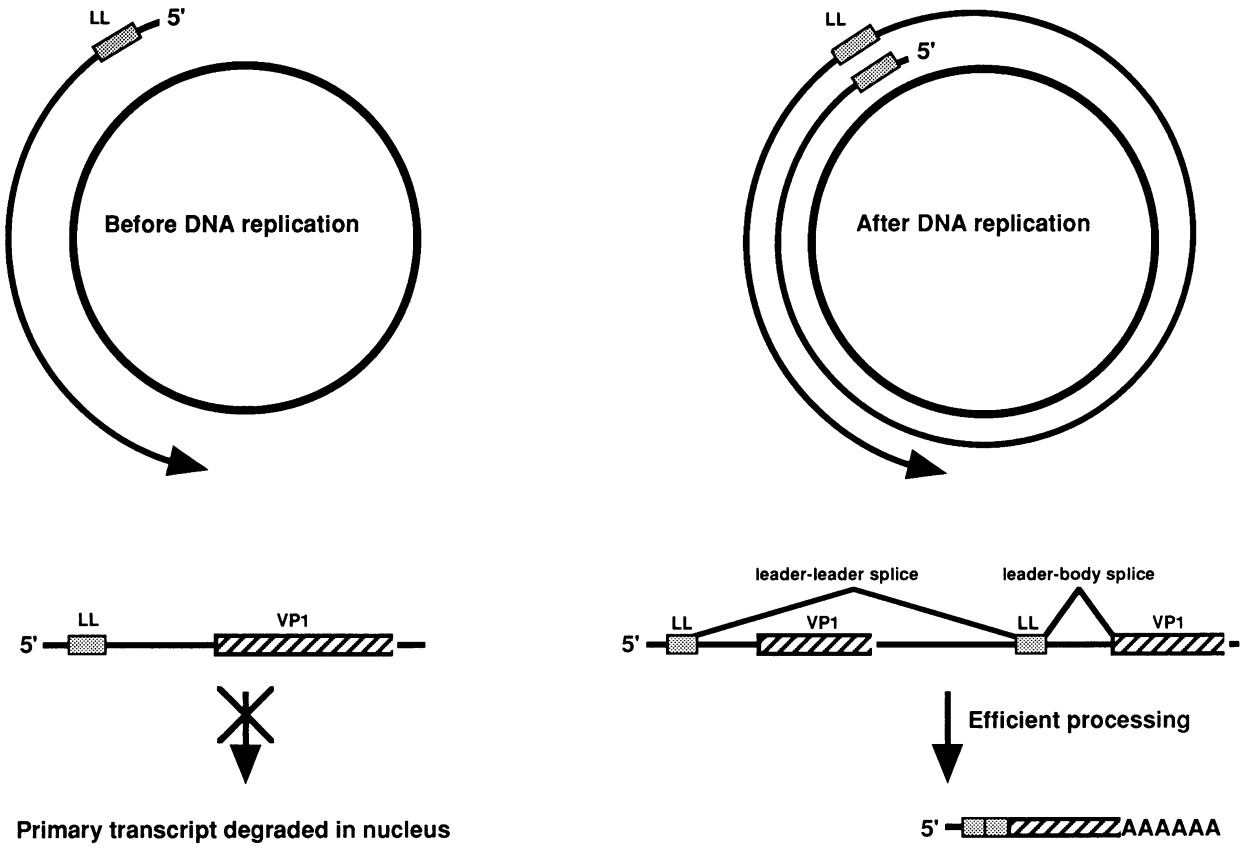


FIG. 7. Model for the control of polyomavirus L-mRNA accumulation. At the top is shown a schematic representation of the polyomavirus late transcription unit. The shaded box (LL) near the 5′ end of the late transcription unit represents the 57-nucleotide noncoding late leader exon (not to scale). The arrow at the end of the transcription unit represents the poly(A) site. In this model, late transcription is efficiently terminated before DNA replication, producing primary transcripts which contain only one late leader exon. These molecules are inefficiently spliced, and most are degraded in the nucleus. After the onset of DNA replication, termination efficiency decreases, allowing for the production of multigenome-length primary transcripts. This allows leader-to-leader splicing, which stabilizes the RNAs. In the example shown here, the 3′-terminal leader exon splices to a VP1 coding exon to produce a mature mRNA with two tandem leader units at its 5′ end. Following splicing and polyadenylation, mature mRNAs are transported to the cytoplasm.

presented here cannot be explained by a random-splicing model for the processing of polyomavirus late pre-mRNAs.

A more complicated model for transcription termination, in which a polymerase must encounter a terminator twice before efficiently terminating, may also explain the observed distribution. This possibility seems unlikely in light of the study on duplicated poly(A) regions, which do not appear to increase the efficiency of termination for polyomavirus (28); however, it is still a formal possibility that an element not present in the duplicated constructs is responsible for this effect.

The simplest model which can account for the observed distribution of late leaders is that primary transcripts containing one leader are preferentially degraded in the nucleus. Studies in several systems have indicated that the splicing or at least the competence to splice can be important for the efficient accumulation of RNA (8, 9, 11, 17–20, 31, 33). Our laboratory has demonstrated that any mutation which blocks leader-to-leader splicing interferes with the accumulation of L-RNA in both the cytoplasm and the nucleus (4–6). In addition, Kern et al. (25, 26) have shown that cells containing a single integrated copy of polyomavirus do not produce a detectable amount of L-RNA; however, a line which contains multiple genomes inserted in tandem, allowing for the production of primary transcripts which can be spliced leader-to-leader, produces a high level of polyomavirus L-RNA containing multiple leaders. Thus, evidence from a number of different approaches has indicated that the production of large, multigenome-length transcripts which can be spliced leader-to-leader is a prerequisite for the efficient accumulation of polyomavirus L-mRNA.

An examination of leader multiplicity in the absence or presence of DNA replication reveals that in the absence of DNA replication there are fewer tandem leader units on L-mRNAs. This result can be interpreted in one of two ways: either the splicing of leaders is a regulated event and is very inefficient in the absence of DNA replication, or the termination efficiency of late transcription is changing and the leader distribution reflects this change. Our previous work, using nuclear run-on transcription, suggested that termination for late transcription becomes less efficient as infection proceeds (22). This would predict that longer pre-mRNA molecules would be produced at late times during infection. These molecules would then be resolved by the splicing of late leader exons to each other. The finding, using two different methods (run-on transcription and RNA PCR), that the efficiency of transcription termination for polyomavirus late transcription decreases in the presence of DNA replication, coupled with the data summarized above on the requirement for leader-to-leader splicing for the

efficient accumulation of L-RNA, is strong supporting evidence that these events are involved in the accumulation of L-RNA after the onset of viral DNA replication.

Model for the control of polyomavirus late-gene expression. Figure 7 is a schematic representation of our working model for the regulation of polyomavirus late-gene expression. In this model, before DNA replication, transcription termination is efficient, allowing for the production of mostly half-genome-length primary transcripts. This is based on nuclear run-on transcription (22) and the leader multiplicity of mRNAs in the absence of DNA replication presented here. These primary transcripts are inefficiently spliced, and most are not transported to the cytoplasm and are degraded in the nucleus. This is supported by the change in splicing efficiencies observed before and after DNA replication (22) and by the leader multiplicity analysis of nuclear RNA and cytoplasmic RNA shown here (Fig. 3, 5, and 6). In addition, other studies from our laboratory have shown that splicing is required for stabilization of L-RNA (4–6; Barrett et al., submitted). Therefore, before replication, the small amount of L-RNA that accumulates is predominantly unspliced RNA (single-leader mVP2) that has presumably escaped the degradation pathway.

Once replication begins, transcription termination becomes less and less efficient. This might be due to a titration of a limiting polyadenylation or termination factor, as has been suggested by others (26, 27). This model is consistent with the correlation between MOI and leader multiplicity shown here. A decrease in termination efficiency would allow the production of large, multigenome-length transcripts. Leader-to-leader splicing would then stabilize these pre-mRNAs, allowing 3′-end processing and mRNA body splicing to occur. This model is supported by transcription studies (both run-on transcription [22] and pulse-labeling [35]), studies of leader multiplicity in the presence and absence of DNA replication (shown above), work on late-leader-splicing mutants (4–6; Barrett et al., submitted), and analyses of viral RNAs found in cells containing integrated polyomavirus genomes (25, 26). In addition to the stabilization of mRNAs after DNA replication, the increase in the number of templates being transcribed would increase the rate at which L-RNA accumulates.

Our model is consistent with all of the data available on the accumulation and regulation of polyomavirus L-RNA and provides the virus with a mechanism by which the production of capsid proteins is coupled to the extent of DNA replication. This model, however, does not address another important aspect of the early-late switch: the down regulation of E-RNA accumulation. Although levels of E-RNA may be influenced by large T-antigen autoregulation of the early promoter (36), some studies have suggested that E-RNA levels may be controlled after transcription initiation (14, 35). We are currently investigating the possibility that giant late primary transcripts themselves contribute to the down regulation of E-RNA at late times, perhaps by forming RNA-RNA duplexes in the nucleus.

ACKNOWLEDGMENTS

We thank N. Barrett, Y. Luo, and A. Roome for helpful comments on the manuscript.

This work was supported by Public Health Service grant CA 45382 from the National Cancer Institute.

LITERATURE CITED

1. **Acheson, N.** 1976. Transcription during productive infection with polyoma virus and SV40. Cell **8**:1–12.
2. **Acheson, N., E. Buetti, K. Scherrer, and R. Weil.** 1971. Transcription of the polyoma virus genome: synthesis and cleavage of giant late polyoma specific RNA. Proc. Natl. Acad. Sci. USA **68**:2231–2235.
3. **Acheson, N. H.** 1978. Polyoma giant RNAs contain tandem repeats of the nucleotide sequence of the entire viral genome. Proc. Natl. Acad. Sci. USA **75**:4754–4758.
4. **Adami, G. R., and G. G. Carmichael.** 1986. Polyomavirus late leader region serves an essential spacer function necessary for viability and late gene expression. J. Virol. **58**:417–425.
5. **Adami, G., and G. G. Carmichael.** 1987. The length but not the sequence of the polyoma virus late leader exon is important for both late RNA splicing and stability. Nucleic Acids Res. **15**: 2593–2610.
6. **Adami, G. R., C. W. Marlor, N. L. Barrett, and G. G. Carmichael.** 1989. Leader-to-leader splicing is required for the efficient production and accumulation of polyomavirus late mRNAs. J. Virol. **63**:85–93.
7. **Beard, P., N. H. Acheson, and I. H. Maxwell.** 1976. Strand-specific transcription of polyoma virus DNA early in productive infection and in transformed cells. J. Virol. **17**:20–26.
8. **Brinster, R. L., J. M. Allen, R. R. Behringer, R. E. Gelinas, and R. D. Palmiter.** 1988. Introns increase transcriptional efficiency in transgenic mice. Proc. Natl. Acad. Sci. USA **85**:836–840.
9. **Buchman, A. R., and P. Berg.** 1988. Comparison of intron-dependent and intron-independent gene expression. Mol. Cell. Biol. **8**:4395–4405.
10. **Cahill, K. B., A. J. Roome, and G. G. Carmichael.** 1990. Replication-dependent transactivation of the polyomavirus late promoter. J. Virol. **64**:992–1001.
11. **Callis, J., M. Fromm, and V. Walbot.** 1987. Introns increase gene expression in cultured maize cells. Genes Dev. **1**:1183–1200.
12. **Cogen, B.** 1978. Virus-specific early RNA in 3T6 cells infected by a tsA mutant of polyoma virus. Virology **85**:222–230.
13. **Cowie, A., C. Tyndall, and R. Kamen.** 1981. Sequences at the capped 5′-ends of polyoma virus late region mRNAs: an example of extreme terminal heterogeneity. Nucleic Acids Res. **9**:6305–6322.
14. **Farmerie, W. G., and W. R. Folk.** 1984. Regulation of polyomavirus transcription by large tumor antigen. Proc. Natl. Acad. Sci. USA **81**:6919–6923.
15. **Flavell, A., A. Cowie, J. R. Arrand, and R. Kamen.** 1980. Localization of three major capped 5′ ends of polyoma virus late mRNAs within a single tetranucleotide sequence in the viral genome. J. Virol. **33**:902–908.
16. **Falvell, A. J., A. Cowie, S. Legon, and R. Kamen.** 1979. Multiple 5′-terminal cap structures in late polyoma virus RNA. Cell **16**:357–372.
17. **Gasser, C. S., C. C. Simonsen, J. W. Schilling, and R. T. Schimke.** 1982. Expression of abbreviated mouse dihydrohydrofolate reductase genes in cultured hamster cells. Proc. Natl. Acad. Sci. USA **79**:6522–6526.
18. **Gruss, P., and G. Khoury.** 1980. Rescue of a splicing defective mutant by insertion of an heterologous intron. Nature (London) **286**:634–637.
19. **Gruss, P., C. Lai, R. Dhar, and G. Khoury.** 1979. Splicing as a requirement for biogenesis of functional 16S mRNA of simian virus 40. Proc. Natl. Acad. Sci. USA **76**:4317–4321.
20. **Hamer, D., K. Smith, S. Boyer, and P. Leder.** 1979. SV40 recombinants carrying rabbit beta-globin coding sequences. Cell **17**:725–735.
21. **Heiser, W. C., and W. Eckhart.** 1982. Polyoma virus early and late mRNAs in productively infected mouse 3T6 cells. J. Virol. **44**:175–188.
22. **Hyde-DeRuyscher, R., and G. G. Carmichael.** 1988. Polyomavirus early-late switch is not regulated at the level of transcription initiation and is associated with changes in RNA processing. Proc. Natl. Acad. Sci. USA **85**:8993–8997.
23. **Kamen, R., P. Jat, R. Treisman, J. Favaloro, and W. R. Folk.** 1982. 5′ Termini of polyoma virus early region transcripts synthesized in vivo by wild-type virus and viable deletion mutants. J. Mol. Biol. **159**:189–224.

24. **Kamen, R., D. M. Lindstrom, H. Shure, and R. W. Old.** 1974. Virus-specific RNA in cells productively infected or transformed by polyoma virus. Cold Spring Harbor Symp. Quant. Biol. **39**:187–198.
25. **Kern, F. G., P. D. Bovi, and C. Basilico.** 1987. A reiterated leader sequence is present in polyomavirus late transcripts produced by a transformed rat cell line. J. Virol. **61**:4055–4059.
26. **Kern, F. G., S. Pellegrini, and C. Basilico.** 1986. Cis and trans-acting factors regulating gene expression from the polyomavirus late promoter. Cancer Cells **4**:115–124.
27. **Lanoix, J., and N. Acheson.** 1988. A rabbit β-globin polyadenylation signal directs efficient termination of transcription of polyomavirus DNA. EMBO J. **7**:2515–2522.
28. **Lanoix, J., R. W. Tseng, and N. H. Acheson.** 1986. Duplication of functional polyadenylation signals in polyomavirus DNA does not alter efficiency of polyadenylation or transcription termination. J. Virol. **58**:733–742.
29. **Legon, S., A. J. Flavell, A. Cowie, and R. Kamen.** 1979. Amplification in the leader sequence of late polyoma virus mRNAs. Cell **16**:373–388.
30. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
31. **Neuberger, M. S., and G. T. Williams.** 1988. The intron requirement for immunoglobulin gene expression is dependent upon the promoter. Nucleic Acids Res. **16**:6713–6724.
32. **Piper, P.** 1979. Polyoma virus transcription early during productive infection of mouse 3T6 cells. J. Mol. Biol. **131**:399–407.
33. **Ryu, W.-S., and J. E. Mertz.** 1989. Simian virus 40 late transcripts lacking excisable intervening sequences are defective in both stability in the nucleus and transport to the cytoplasm. J. Virol. **63**:4386–4394.
34. **Sarkar, G., and S. S. Sommer.** 1990. Shedding light on PCR contamination. Nature (London) **343**:27.
35. **Skarnes, W. C., D. C. Tessier, and N. A. Acheson.** 1988. RNA polymerases stall and/or prematurely terminate nearby both early and late promoters on polyomavirus DNA. J. Mol. Biol. **203**:153–171.
36. **Tooze, J. (ed.).** 1980. DNA tumor viruses: molecular biology of tumor viruses, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
37. **Treisman, R.** 1980. Characterization of polyoma late mRNA leader sequences by molecular cloning and DNA sequence analysis. Nucleic Acids Res. **8**:4867–4888.
38. **Treisman, R., and R. Kamen.** 1981. Structure of polyoma virus late nuclear RNA. J. Mol. Biol. **148**:273–301.
39. **Tseng, R. W., and N. H. Acheson.** 1986. Use of a novel S1 nuclease RNA-mapping technique to measure efficiency of transcription termination on polyomavirus DNA. Mol. Cell. Biol. **6**:1624–1632.