JOURNAL OF VIROLOGY, Dec. 1983, p. 654–659
0022-538X/83/120654-06$02.00/0

Vol. 48, No. 3

Discontinuities in the DNA Synthesized by an Avian Retrovirus

JOHN M. TAYLOR,* ANITA CYWINSKI, AND JULIE K. SMITH

The Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111

Received 20 June 1983/Accepted 9 September 1983

The unintegrated linear DNA synthesized in cells infected by Rous sarcoma virus is a predominantly double-stranded structure in which most of the minus-strand DNA, complementary to the viral RNA genome, is genome sized, whereas the plus-strand DNA is present as subgenomic fragments. We previously reported the application of benzoylated naphthoylated DEAE-cellulose chromatography to demonstrate that of the linear viral DNA species synthesized in quail embryo fibroblasts infected with Rous sarcoma virus greater than 99.5% contain single-stranded regions and these regions are predominantly composed of plus-strand DNA sequences (T. W. Hsu and J. M. Taylor, J. Virol. **44**:47–53, 1982). We now present the following additional findings. (i) There were on the average 3.5 single-stranded regions per linear viral DNA, and these single-stranded regions could occur at many locations. (ii) With a probe to the long terminal repeat, we detected, in addition to a heterogeneous size distribution of subgenomic plus-strand DNA species, at least three prominent discrete size classes. Each of these discrete species had its own specific initiation site, but all had the same termination site. Such species were analogous to those reported by Kung et al. (J. Virol. **37**:127–138, 1981). (iii) These discrete size classes of plus-strand DNA were present not only on the major size class of linear DNA but also on a heterogeneous population of slower-sedimenting species, which we have called immature linears. Our interpretation is that we have thus detected several additional sites for the initiation of plus-strand DNA. (iv) The 340-base plus-strand strong-stop DNA was only found associated with the immature linears. (v) From a size and hybridization comparison of these discrete size classes of plus-strand DNA with minus-strand DNA species, as synthesized in the endogenous reaction of melittin-disrupted virions, it was found that the putative additional initiation sites for plus-strand DNA synthesis corresponded to many of the pause sites in the synthesis of minus-strand DNA.

The general model for the synthesis and structure of linear retroviral DNA (as recently reviewed in reference 14) is a composite of studies of viral DNA from infected cells in vivo and DNA as synthesized outside the cell, either in the endogenous reaction of disrupted virions or with purified viral RNA and reverse transcriptase. It is also a composite of studies with murine or avian retroviruses. However, a major difference between the linear viral DNA of certain avian retroviruses (Rous sarcoma viruses and related avian leukosis viruses) and that of the murine leukemia viruses is in the mechanism of synthesis of plus-strand DNA, the strand homologous to the viral RNA genome. With the avian viruses the plus-strand DNA in the linear species is predominantly, or even exclusively, composed of subgenomic fragments with lengths of 1 to 3 kilobases. In contrast, with the murine leukemia viruses there is a significant amount of plus-strand DNA that is of genome size. This latter fact also appears to be true for mouse mammary tumor virus (14). Viral DNA of the ovine retrovirus, visna virus, may be intermediate, with exactly two fragments of plus-strand DNA (5). Our studies have been concerned with the discontinuities in the viral plus-strand DNA of linear viral DNAs present in quail cells infected with a Rous sarcoma virus.

In a previous study single-stranded regions (SSRs) on the viral DNA were detected by chromatography on benzoylated naphthoylated DEAE-cellulose (6). It was found in agreement with studies of Boone and Skalka (1), that the SSRs were predominantly, but by no means exclusively, composed of plus-strand DNA. Studies described in this manuscript were undertaken to further characterize the frequency and location of the discontinuities.

TABLE 1. Frequency and distribution of SSRs on fragments of linear viral DNA

Fragment length (b)[a]	Fragment location[b]	Fraction that lacks SSR[c]	Avg no. of SSRs per:	
			Fragment[d]	Linear[e]
9,312	Intact genome	<0.005	>5.3	>5.3
5,231	Left-side *Kpn*I	<0.005	>5.3	>9.6
4,397	Right-side *Kpn*I	<0.005	>5.3	>11.4
3,825	Middle *Eco*RI	0.21	1.6	3.9
3,094	Right internal *Eco*RI	0.37	1.0	3.0
2,369	Left internal *Eco*RI	0.32	1.1	4.6
1,384	Internal *Bam*HI	0.72	0.33	2.3

[a] Indicated sizes and locations of restriction fragments are from the published nucleotide sequence of Schwartz et al. (12).

[b] Linear viral DNA containing SSRs was subjected to restriction enzyme digestion followed by BND-cellulose chromatograhy to separate fragments with and without SSRs. Fragments were subjected to agarose gel electrophoresis and Southern analysis (13) to detect viral sequences (6).

[c] For each fragment, the fraction that lacks SSRs was determined by densitometer analysis of the autoradiogram. A relatively overexposed autoradiogram was used to set the indicated lower level of sensitivity.

[d] With the fraction P_0, we deduced the average number of SSRs per fragment (z) with the Poisson formula $z = -\ln P_0$.

[e] With the number of SSRs per fragment, we deduced the average number of SSRs per total linear of 9,312 nucleotides in length.

MATERIALS AND METHODS

Cells and viruses. The source of the Prague strain of Rous sarcoma virus was the quail clone Q-PrA-4, transformed by the integration of a single provirus (11). Growth of this virus and its use for the infection of quail embryo fibroblasts were as previously described (6).

Isolation and analysis of unintegrated viral DNA. At 22 h after infection, unintegrated viral DNA was isolated by a modified Hirt extraction (6). Linear DNA with SSRs was isolated by two passages over BND-cellulose and used in the studies described below. Agarose gel analysis with or without prior glyoxalation was followed by a Southern transfer to nitrocellulose (13). All of these procedures have been previously described (6).

M13 probes for the plus and the minus strands of the LTR. Plasmid pSRA2, which contains the entire avian sarcoma virus genome inserted in pBR322, was generously provided by William DeLorbe and co-workers (3). DNA of pSRA2 was digested with *Eco*RI, and the 340-base (b) long terminal repeat (LTR) fragment released was subcloned into M13mp7, using materials from Bethesda Research Laboratories. We thus selected M13 clones that contained the plus or the minus strand of the LTR. Radioactive probes were made as previously described (10). We copied the single-stranded genome of unmodified M13mp7, using the Klenow fragment of DNA polymerase I from *Escherichia coli*, random oligonucleotide primers, and [^{32}P]TTP. This DNA was extracted, denatured, and used to make each of the LTR-specific probes. To make 10 ml of probe, 16 ng of product was hybridized with 32 ng of one of the M13 recombinant clones. Hybridization was in 16 μl of 0.6 M NaCl for 2 h at 68°C.

RESULTS

Frequency and distribution of SSRs. It previously has been shown that linear viral DNA both before and after digestion with certain restriction enzymes behaves during chromatography on BND-cellulose in a manner indicative of the presence of SSRs (6). Such studies have been extended and quantitated by densitometry (Table 1). It can be seen that as we produced smaller fragments of the viral DNA there was an increased probability, P_0, that some fragments would lack SSRs. From this value we deduced the average number of SSRs per fragment, z, using the Poisson formula $P_0 = e^{-z}$. Then, utilizing the length of the fragment relative to the length of the uncut linear, we deduced the average number of SSRs per linear. The various results of these calculations fell into two groups (Table 1). One group consisted of the four small fragments considered in the table. For these, the deductions were quite uniform, with an average value of 3.5 SSRs per linear viral DNA. This uniformity was consistent with a random distribution of SSRs throughout the DNA. None of these fragments, however, contained a right or a left end of the linear. The second group, composed of the intact linear or the left or the right side, yielded higher deduced values. Our interpretation was that there was an additional relatively higher frequency of SSR at each end of the linears.

Size distribution of plus-strand DNA in linears. The size distribution of plus-strand species present in linear avian retrovirus DNA has been reported as heterogeneous, with a size range of 1 to 3 kilobases (4). To simplify the picture and possibly detect specific initiation or termination sites or both for plus-strand DNA, we used probes to limited regions of the genome. As previously reported, the use of such a plus-

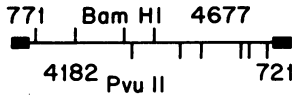

FIG. 1. Restriction enzyme sites on linear viral DNA. This diagram, based on the nucleotide sequence data of Schwartz et al. (12), indicates the relevant sites of *Bam*HI and *Pvu*II and the lengths of restriction fragments that would be detected using probes specific for LTRs (■).

strand-specific probe to a 1,384-b *Bam*HI fragment still gave a predominantly heterogeneous pattern of size classes (6). We therefore used a plus-strand probe specific for the 340-b LTR. Our expectation was that we would thus detect only those plus-strand species that began or ended with an LTR sequence. To distinguish between the plus-strand species that began with an LTR from those that ended with an LTR, we chose appropriate prior restriction enzyme digestions. As summarized in Fig. 1, *Bam*HI and *Pvu*II make such cuts. Results of our analysis are shown in Fig. 2. In this experiment, SSR-containing linear viral DNA from infected cells was isolated by BND-cellulose chromatography, glyoxalated under denaturing conditions, and then subjected to electrophoresis on a gel of 1% agarose. After the SSR-containing linear viral DNA was transferred to nitrocellulose by the procedure of Southern (13), we used a specific M13 clone to detect only those species that contain sequences from the plus strand of the LTR. As shown in Fig. 2, lane 1, the major species, indicated by SS+, had the approximate size expected for the viral LTR, 340 b (12). In fact, Fig. 2, lane 5, shows the migration of an LTR, as released by *Eco*RI, from a clone of the whole viral genome cloned in pBR322 (3). However, in Fig. 2, lane 1, additional discrete size classes of plus-strand DNA were detected. Relative to markers of phage lambda DNA cut with *Hin*dIII, we deduced the species indicated as a, b, and c to be 3,710, 1,660, and 1,070 b, respectively. In Fig. 2, lane 1, the intensities of a, b, and c relative to SS+ were low and required an increased exposure for the autoradiogram. However, in some experiments the relative intensities were comparable (see Fig. 4, lane 1).

To determine whether the three larger species began or ended with an LTR, we subjected the native linear DNA to a prior digestion with *Bam*HI, *Pvu*II, or both (Fig. 2, lanes 2, 3, and 4, respectively). As summarized in Fig. 1, prior digestion with *Bam*HI would create a 771-b fragment for species that initiated at the left-side LTR and leave uncut those species shorter than 4,667 b that terminated at the right-side LTR. The unique plus-strand species a, b, and c were not cut by *Bam*HI, consistent with them having

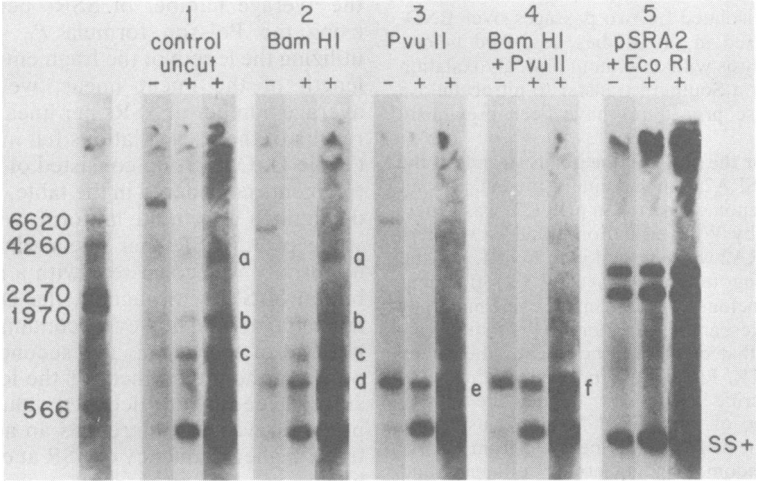

FIG. 2. Agarose gel electrophoresis of glyoxal-treated linear viral DNA. For lanes 1 through 4, linear DNA was isolated from infected cells by a Hirt extraction and BND-cellulose chromatography. Samples of this DNA were subjected to the indicated restriction digests, after which they were glyoxalated and subjected to electrophoresis on a slab gel of 1% agarose. Lane 5 is an *Eco*RI digest of recombinant plasmid pSRA2. After linear DNA was transferred to nitrocellulose, an M13 recombinant probe was used to detect sequences related to the plus strand of the LTR. As shown, autoradiograms corresponding to two different exposures were obtained. The filter was then stripped of probe and hybridized with another M13 probe, this time to detect sequences related to the minus strand of the LTR, followed again by autoradiography. The left-side lane is phage lambda DNA cut with *Hin*dIII, with lengths as indicated. Those bands with adjacent lettering are discussed in the text, and the migration of the plus-strand strong-stop DNA is indicated by SS+.

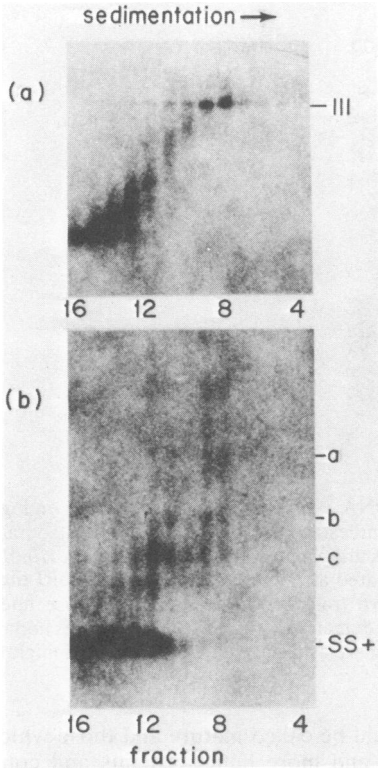

FIG. 3. Rate zonal sedimentation of linear viral DNA followed by agarose gel electrophoresis. Samples of unintegrated linear viral DNA were subjected to sedimentation on a 5 to 20% sucrose gradient (in 10 mM Tris-hydrochloride [pH 7.4]–10 mM EDTA) in a Beckman SW60 rotor at 50,000 rpm for 3 h at 20°C. Gradient fractions were then subjected to agarose gel electrophoresis without (a) or with (b) prior glyoxalation. After the DNA was transferred to nitrocellulose, sequences related to the plus-strand LTR were detected as described in the legend to Fig. 2. In (a), the migration of mature linear viral DNA is indicated by III; in (b), the lettering at the right side is discussed in the text.

initiated at discrete locations near the right side and terminated after the LTR at the right side (Fig. 2, lane 2). These species were analogous to species seen by Kung et al. (9) using a related avian retrovirus. As a consequence of the *Bam*HI digestion, there was in addition a new band, d, which migrated relative to markers at about 774 b. This band was interpreted as being derived by a *Bam*HI cut on species of plus-strand DNA longer than 774 b that initiated at the left side of the left LTR. To look more closely at such species, we submitted linear DNA to a prior digestion with *Pvu*II. This would cut the three discrete species ending in the right-side LTR to 721 b and might allow us to detect longer species that initiated in the left-side LTR (Fig. 1). We observed a species, e, of about 740 b, presumably from the right-side LTR, but we still could not detect any discrete species that initiated in the left-side LTR (Fig. 2, lane 3). Thus, those species that initiated in the left-side LTR and were detected by the *Bam*HI digest were heterogeneous in their termination sites. As a control we found that when the linear DNA was digested with both *Bam*HI and *Pvu*II (Fig. 2, lane 4), only a single band, f, was obtained and was interpreted as being the unresolved mixture of fragments d and e. Also shown in Fig. 2 are the autoradiograms obtained when the nitrocellulose filter was stripped of probe and rehybridized with an M13 probe specific not for the plus-strand DNA but for the minus-strand DNA of the LTR. These data confirmed that the minus-strand DNA in the linear was predominantly of genome size and that the indicated restriction enzyme digestions were complete.

In summary, we detected not only the 340-b plus-strand DNA, which is the classical strong-stop plus DNA (15), but also at least three other plus-strand species which initiated at different locations than did the 340-b species but which appeared to have the same termination site. We also detected plus-strand species that initiated at the left-side LTR, but these species were heterogeneous with respect to their 3′ termini.

At one point we wished to quantitate data such as those in Fig. 2 and determine what fraction of the plus strand DNA species initiated and terminated in a given way. However, a complication became apparent when the linear DNA was subjected to prior rate zonal sedimentation. Linear viral DNA with SSRs was first isolated by BND-cellulose chromatography and then subjected to rate zonal sedimentation followed by agarose gel electrophoresis without or with prior denaturation in the presence of glyoxal. After Southern transfer (13) we assayed for those plus-strand sequences complementary to the LTR (Fig. 3). Some of the linears sedimented to about eight fractions from the bottom of the gradient, although there was also a heterogeneous distribution of species with slower sedimentation (Fig. 3a).

When the gradient fractions were denatured before electrophoresis (Fig. 3b), we found that the previously mentioned discrete species of plus-strand DNA (a, b, and c) were associated with both the slower-sedimenting species and the major band in fraction 8. In contrast, the major 340-b, plus-strand strong-stop DNA was only associated with the slower-sedimenting species.

Pauses in endogenous minus-strand DNA synthesis. We then made what proved to be an interesting comparison between the specific plus-strand species that we described on the

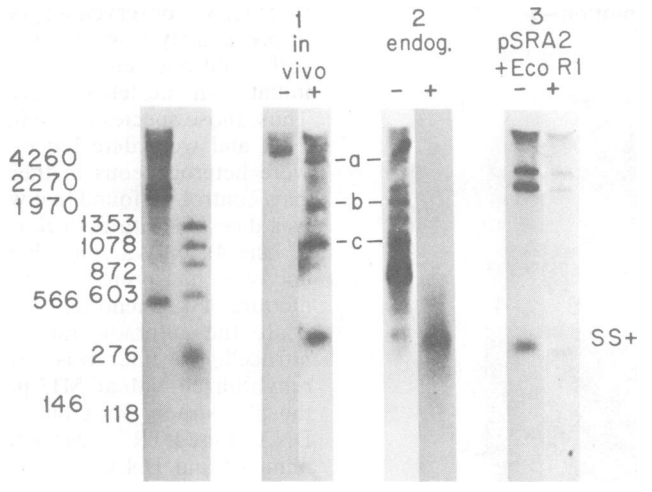

FIG. 4. Agarose gel electrophoresis of glyoxal-treated viral DNA synthesized in vivo or in an endogenous reaction of melittin-disrupted virions. Lane 1 is the linear DNA synthesized in vivo; lane 2 represents viral DNA synthesized with melittin-disrupted virions. The two left-side lanes are phage lambda DNA cut with *Hin*dIII and phage φX replicative form cut with *Hae*III, with lengths as indicated at the left. Lane 3 is an *Eco*RI digest of recombinant plasmid pSRA2. After electrophoresis and Southern transfer (13), the nitrocellulose filter was sequentially hybridized as described in the legend to Fig. 2, first to detect plus-strand LTR sequences and then to detect minus-strand LTR sequences. Bands with adjacent lettering are as discussed in the text, and migration of plus-strand strong stop DNA is indicated by SS+.

linear DNA synthesized in infected cells and the pauses in the minus-strand DNA synthesis as carried out with an endogenous reaction. For some time it has been known that the synthesis of minus-stand DNA pauses at a number of discrete sites. They are not permanent stops because with increased time of synthesis the average size of the DNA products increases (1). Our results were similar. In an endogenous reaction there were synthesized multiple species of minus-strand DNA that contained the LTR sequence (Fig. 4, lane 2). The plus-strand DNA synthesized in vivo contained fewer discrete species (Fig. 4, lane 1). However, for each of these plus-strand species (a, b, and c), and possibly additional minor species, there was a species of similar size in the minus-strand DNA synthesized in the endogenous reaction. This correspondence extended to the relative intensities of the sets of plus- and minus-strand species. Also, because each of the minus strands started with an LTR and, as we have shown earlier, each of the plus strands ended with an LTR, we inferred that a pair of species of equal length must be complementary, that is, encoded from the same region of the linear.

DISCUSSION

Because of the data obtained, we divided the so-called unintegrated linear viral DNA species into two populations: those which appear homogeneous by electrophoresis or sedimentation and could be called mature and those which are smaller and more heterogeneous and could be called immature. Quantitation of linears that contained plus-strand species, as detected by the LTR probe, indicated that about 60% of the linears were immature.

As described here and previously (6) virtually all of the mature linears contained SSRs that were composed predominantly of plus-strand DNA. Boone and Skalka (1) originally made this observation and concluded that this occurs via displacement synthesis. Such synthesis also occurs with cauliflower mosaic virus, which is retrovirus-like in many respects (7). The displaced regions on cauliflower mosaic virus begin at three or four specific locations that have been well characterized (7). This is in contrast to the avian retrovirus situation. Although we have shown 3.5 single-stranded plus-strand tails per average mature linear, there were actually many more potential sites. The discrete size classes that we did find were on both the mature and the immature linears. We would guess, but have not shown, that the SSRs on the immature linears were composed of minus-strand DNA, rather than plus-strand DNA. We have shown that the discrete species corresponded to unique initiation and shared termination sites. Distribution of these species differed from that of the strong-stop plus-strand species which was only detected on the immature linears. Our interpretation was that the slower species were the less mature forms of linear viral DNA and that maturation

included elongation of the minus strand. Thus, when the minus strand reached genome size, the 340-b, plus-strand strong-stop species could then continue synthesis by using the 3′ end of the minus strand, and so the discrete size class of strong-stop DNA would no longer be found, as such, on mature linears. (Our interpretation was that the discrete size class of strong-stop DNA becomes those plus-strand species that we detected to have a common 5′ terminus at the left side of the linears but are heterogeneous with respect to their 3′ termini.) Since the additional discrete size classes of plus-strand DNA were present not only on immature linears but also on mature linears, we interpreted that they may be initiated from sites that were used not as alternatives but as sites in addition to the site for the 340-b, plus-strand strong-stop DNA.

We detected a correspondence between the additional sites for the initiation of plus-strand synthesis and some, but not all, of the sites at which there were pauses during endogenous synthesis of minus-strand DNA. Our data could not distinguish whether a minus-strand pause site and a plus-strand initiation site may be shared during an isolated reverse transcription event for a viral genome. However, in some experiments with in vivo linears, we also detected discrete species of minus-strand DNA that contain the minus-strand sequence of the LTR and have the same electrophoretic mobility as the described species of plus-strand DNA (data not shown). The possibility that such a correspondence could occur on a single molecule is consistent with the model of retrovirus DNA synthesis and recombination proposed by Coffin (2). That model was attractive because it suggested a reason for the viral genome existing as a dimer of two identical RNAs. Coffin suggested that nicks generated in genomic RNA could simultaneously create pauses in minus-strand DNA synthesis and primers for plus-strand DNA synthesis. Further elongation of the minus strand could be via a switch in template presumably to the other viral genome in the dimer. Junghans et al. (8) recently argued in favor of a displacement-assimilation model of viral DNA synthesis and recombination, but neither of these two models offers an explanation for the preferred sites of plus-strand initiation that we have observed. With this in mind we recently reconstructed the specific initiation of plus-strand DNA synthesis, using viral RNA, minus-strand DNA, and reverse transcriptase (J. K. Smith, A. Cywinski, and J. M. Taylor, submitted for publication).

ACKNOWLEDGMENTS

This work was supported by Public Health Service grants CC-22651 and RR-05539 from the National Institutes of Health, by grant MV-7 from the American Cancer Society, and by an appropriation from the Commonwealth of Pennsylvania. J.K.S. was supported by Public Health Service predoctoral grant 5 T 32 GM-07071-08 awarded to the University of Pennsylvania.

We thank W. Mason and J. Summers for critical reading of the manuscript and L. Coates for assistance with the tissue culture.

LITERATURE CITED

1. **Boone, L. R., and A. M. Skalka.** 1981. Viral DNA synthesized in vitro by avian retrovirus particles permeabilized with melittin. II. Evidence for a strand displacement mechanism in plus-strand synthesis. J. Virol. **37**:117–126.
2. **Coffin, J.** 1979. Replication and recombination of retrovirus genomes: some underlying hypotheses. J. Gen. Virol. **42**:1–26.
3. **DeLorbe, W. J., P. A. Luciw, H. M. Goodman, H. E. Varmus, and J. M. Bishop.** 1980. Molecular cloning and characterization of avian sarcoma virus circular DNA molecules. J. Virol. **36**:50–61.
4. **Guntaka, R. V., O. C. Richards, P. R. Shank, H.-J. Kung, N. Davidson, E. Fritsch, J. M. Bishop, and H. E. Varmus.** 1976. Covalently closed circular DNA of avian sarcoma virus: purification from nuclei of infected quail tumor cells and measurement by electron microscopy and gel electrophoresis. J. Mol. Biol. **106**:337–357.
5. **Harris, J. D., J. V. Scott, B. Traynor, M. Brahic, L. Stowring, P. Ventura, A. T. Haase, and R. Peluso.** 1981. Visna virus DNA: discovery of a novel gapped structure. Virology **113**:573–583.
6. **Hsu, T. W., and J. M. Taylor.** 1982. Single-stranded regions on unintegrated avian retrovirus DNA. J. Virol. **44**:47–53.
7. **Hull, R., and S. N. Covey.** 1983. Does cauliflower mosaic virus replicate by reverse transcription? Trends Biochem. Sci. **8**:119–121.
8. **Junghans, R. P., L. R. Boone, and A. M. Skalka.** 1982. Retroviral DNA H structures: displacement-assimilation model of recombination. Cell **30**:53–62.
9. **Kung, H. J., Y. K. Fung, J. E. Majors, J. M. Bishop, and H. E. Varmus.** 1981. Synthesis of plus strands of retroviral DNA in cells infected with avian sarcoma virus and mouse mammary tumor virus. J. Virol. **37**:127–138.
10. **Mason, W. S., C. Aldrich, J. Summers, and J. M. Taylor.** 1982. Asymmetric replication of duck hepatitis B virus DNA in liver cells: free minus-strand DNA. Proc. Natl. Acad. Sci. U.S.A. **79**:3997–4001.
11. **Sabran, J. L., T. W. Hsu, C. Yeater, A. Kaji, W. S. Mason, and J. M. Taylor.** 1979. Analysis of integrated avian RNA tumor virus DNA in transformed chicken, duck, and quail fibroblasts. J. Virol. **29**:170–178.
12. **Schwartz, D. E., R. Tizard, and W. Gilbert.** 1983. Nucleotide sequence of Rous sarcoma virus. Cell **32**:853–869.
13. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
14. **Varmus, H., and R. Swanstrom.** 1982. Replication of retroviruses, p. 369–512. *In* R. Weiss, N. Teich, H. Varmus, and J. Coffin (ed.), RNA tumor viruses, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
15. **Varmus, H. E., S. Heasly, H.-J. Kung, H. Opperman, V. C. Smith, J. M. Bishop, and P. R. Shank.** 1978. Kinetics of synthesis, structure and purification of avian sarcoma virus-specific DNA made in the cytoplasm of acutely infected cells. J. Mol. Biol. **120**:55–82.