Journal of General Virology (1996), **77**, 1987–1991. Printed in Great Britain

REVIEW ARTICLE

Reverse transcriptase jumps and gaps

Jean-Jacques Kupiec and Pierre Sonigo

ICGM-CNRS UPR 415, Génétique des Virus, Institut Cochin de Génétique Moléculaire, 22 rue Méchain, F-75014 Paris, France

Introduction

Retroviruses are single-stranded RNA viruses of eukaryotes. Different subfamilies have been described. Avian or murine oncoviruses induce neoplasms, whereas lentiviruses, typified by human immunodeficiency virus and spumaviruses, produce persistent infections. Lentiviral infections may cause chronic disease whereas spumaviruses are apparently non-pathogenic. The retroviral life cycle is characterized by reverse transcription of their single-stranded plus (i.e. coding) RNA genome into a double-stranded DNA intermediate that integrates into the host genome. Over the last 25 years the mechanism of reverse transcription has been studied in great detail and these studies have led to the model shown in Fig. 1. The result of reverse transcription is a linear double-stranded DNA molecule with a long terminal repeat (LTR) at each extremity. Synthesis of each DNA strand by the virus-encoded reverse transcriptase requires one template switch, also called a jump: the first jump is needed for synthesis of a minus DNA strand complementary to the viral RNA, the second jump for plus DNA strand synthesis. A specific RNA primer is required for synthesis of each strand: a tRNA for the minus DNA strand and a polypurine RNA for the plus DNA strand (see Coffin, 1990; Whitcomb & Hughes, 1992, for reviews). In spite of a large body of evidence in support of this model a series of reports has indicated that there could be some differences among retroviruses concerning the mechanisms of the second jump and plus strand DNA synthesis. Here, we review these data and their possible biological significance.

Plus strand DNA synthesis

All the retroviruses analysed so far possess a purine-rich sequence (polypurine tract or PPT) that lies immediately 5′ of the 3′ LTR. The RNA sequence of the PPT resists the retroviral RNase H activity, which degrades the RNA genome after minus cDNA synthesis. This sequence is used as the primer for initiating plus strand DNA synthesis. Production of the correct primer by the endonucleolytic activity of RNase H is dependent upon the specific recognition of the PPT with its cognate RNase H (Huber & Richardson, 1990; Luo *et al.*, 1990; Pullen *et al.*, 1993; Rattray & Champoux, 1989; Wöhrl & Moelling, 1990). For some murine oncoviruses, it seems that one PPT is involved in a single priming event (Kung *et al.*, 1981).

It has been shown that purine-rich sequences bearing a resemblance to the PPT can provide additional initiation sites (Finston & Champoux, 1984; Rattray & Champoux, 1987) and, in avian oncoviruses, the plus strand of linear unintegrated DNA actually contains multiple discontinuities, suggesting a multiple priming mechanism (Boone & Skalka, 1981*a*, *b*; Hsu & Taylor, 1982; Kung *et al.*, 1981; Taylor *et al.*, 1983; Varmus *et al.*, 1978). However, internally initiated DNAs are not utilized for strand transfer (Bowman *et al.*, 1996). These discontinuities could be repaired either before (Lee & Coffin, 1991) or after integration into the host DNA (Miller & Bushman, 1995), possibly by ligase activity or strand displacement (Boone & Skalka, 1981*b*).

The 'gap'

Lentiviruses and spumaviruses seem to use a slightly different mode of multiple priming. In the linear unintegrated DNA of these viruses, a unique discontinuity has been found in an invariant position at the 3′ end of the *pol* open reading frame, adjacent to a copy of the PPT, suggesting a dual mode of initiation of plus strand DNA synthesis.

For lentiviruses, the gapped DNA was reported by Harris *et al.* (1981) to be present in culture cells infected with maedi–visna virus (MVV). In this case, more than 99% of the viral DNA was found to be a linear unintegrated duplex, with a single-stranded gap in the plus strand. Subsequently, Blum *et al.* (1985) analysed the time-course of viral DNA synthesis and showed that in MVV, a major origin of plus DNA synthesis is close to the centre of the genome. In fact, sequence analysis of the MVV genome revealed that the PPT located at the boundary of the 3′ LTR is duplicated at the end of the *pol* open reading frame, near the single-stranded gap, strongly supporting this hypothesis (Sonigo *et al.*, 1985).

The gapped DNA structure has also been demonstrated in spumaviruses. In cultured cells infected with human and simian foamy virus type 1 (HFV and SFV-1), only linear unintegrated DNA molecules could be detected by Southern blotting. In this DNA, a single-stranded region was localized near the middle of the molecule (Kupiec *et al.*, 1988), and nucleotide sequence analysis showed a PPT duplication in the same region (Kupiec *et al.*, 1988, 1991). The gap was then shown to be located on the plus DNA strand and to be 120 nucleotides long,

Author for correspondence: P. Sonigo.
Fax +33 1 40 51 72 10. e-mail sonigo@cochin.inserm.fr

0001-3925

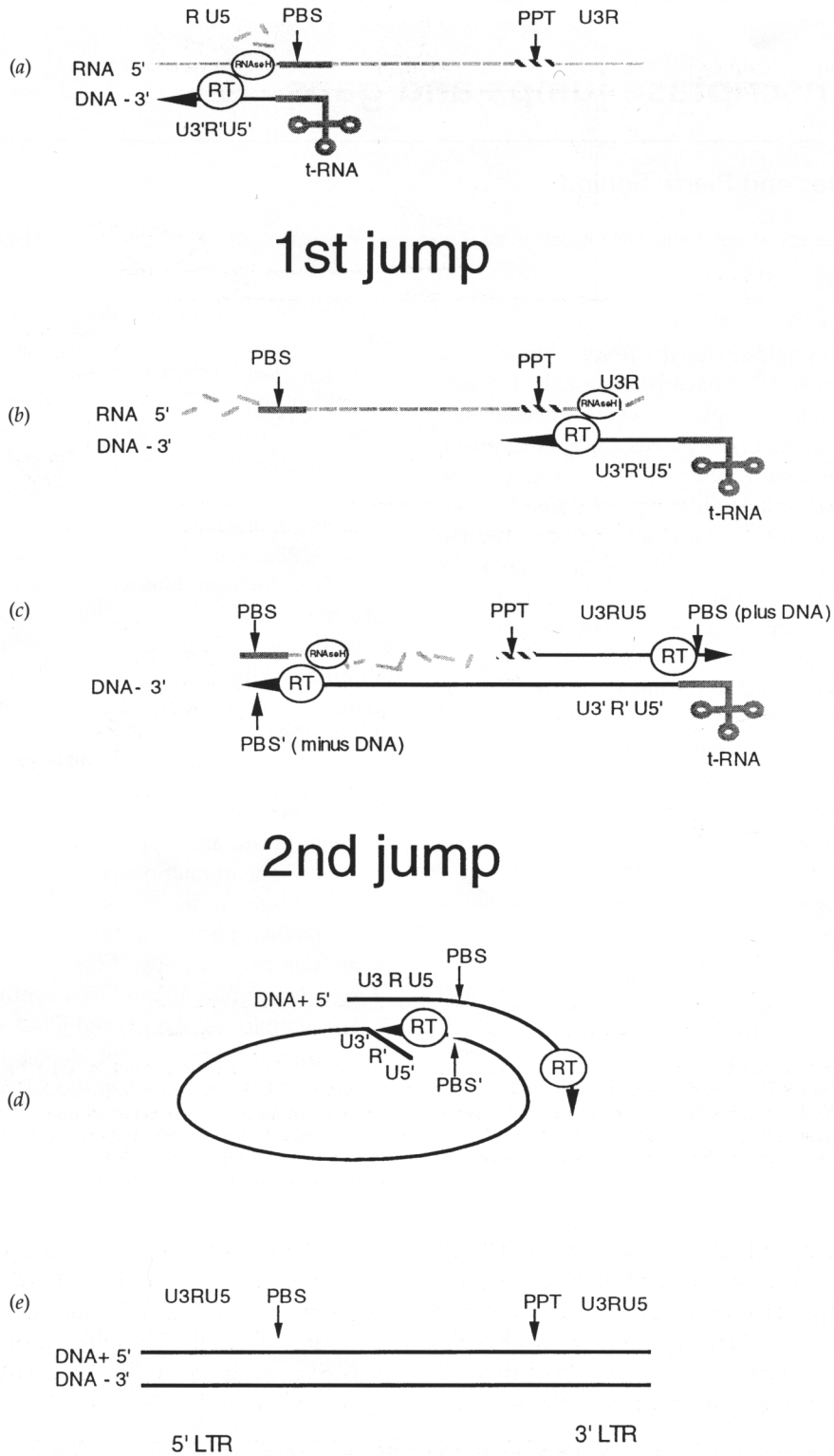

Fig. 1. Classical model of reverse-transcription. PBS, primer binding site; PPT, polypurine tract; U3, R and U5, components of the LTRs. For clarity, actual proportions of these sequences have been altered in the diagrams. Shaded areas correspond to primers or to hybridized sequences that are essential during the jumps. Minus DNA synthesis is primed by a tRNA paired to a complementary sequence, called the primer binding site (*a*). From this site, which defines the 3′ limit of the 5′ LTR, U5 (a sequence unique to the 5′ extremity of the RNA genome) and R (a sequence repeated at both extremities of the genome) are first copied by the reverse transcriptase (RT; *a*). The RNase H activity associated with the RT hydrolyses the RNA template and the newly synthesized DNA sequences are transferred to the 3′ end of the genome (first jump). After pairing of the complementary R sequences, minus DNA synthesis resumes (*b*). As minus DNA synthesis proceeds, plus DNA synthesis is

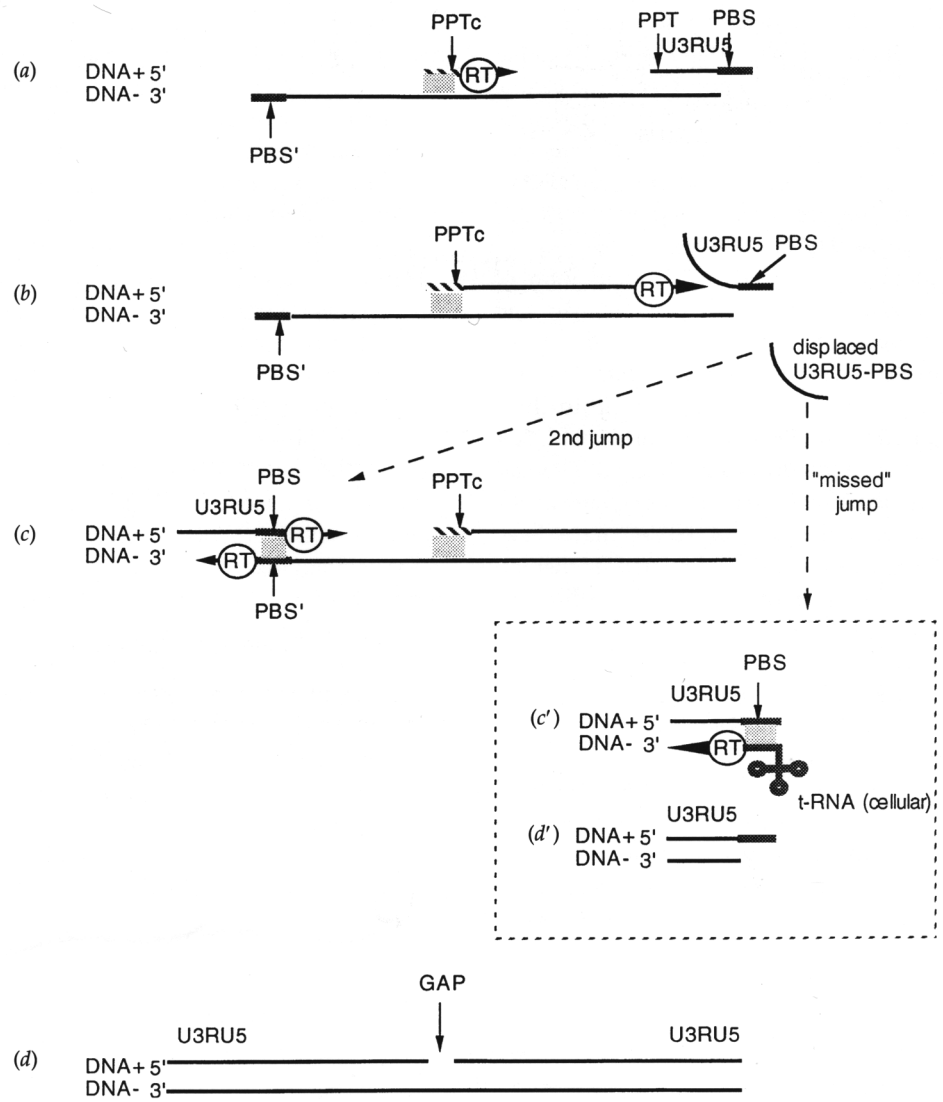

Fig. 2. Model proposed by Li *et al.* (1993) for the second jump. Step (*a*) of this figure corresponds to (*d*) in Fig. 1. PPTc, central copy of the PPT found in lenti- and spumaviruses. The strong stop plus DNA (U3RU5) is displaced by DNA elongating from PPTc (*b*). It can either be transferred to the 3′ extremity of the minus strand DNA (c) or serve as a template for synthesis of a double-stranded LTR (*c′* and *d′*). This 'orphan' LTR has actually been detected during HIV and spumavirus replication. It could either be an unused by-product or serve an unknown function as a source of strong stop DNA.

immediately 5′ to the central PPT (Tobaly *et al.*, 1991). It is worthy of note that in these studies a small double-stranded DNA fragment corresponding to the LTR was also detected. These results have been confirmed in every spumavirus that has been looked at (Schweizer *et al.*, 1989; Renshaw *et al.*, 1991), showing that it is a general property of this group.

A discontinuity had been predicted to be present in human immunodeficiency virus type 1 (HIV-1) DNA since a PPT duplication was found in its nucleotide sequence (Wain-Hobson *et al.*, 1985). The presence of a single-stranded region was then confirmed in the linear HIV-1 DNA from infected lymphoid cells (Charneau & Clavel, 1991; Hungnes *et al.*,

initiated from an RNA primer corresponding to a sequence (the PPT) rich in purine that defines the 5′ limit of the 3′ LTR (*c*). Plus DNA sequences corresponding to U3 (a sequence unique to the 3′ extremity of the RNA genome), R and U5 are copied from their complementary minus DNA, as well as a plus DNA sequence corresponding to the PBS from the tRNA primer (*c*). Once the minus PBS is made and the tRNA primer is removed, it is suggested that circularization occurs by pairing of the PBS complementary sequences (*d*) (second jump). In this model, strand displacement is not needed for the second jump itself but for completion of the 5′ LTR, which occurs at the subsequent step while plus DNA synthesis proceeds on the remainder of the genome, leading to the linear DNA with its two LTRs (*e*). This molecule can then integrate into the host genome. The first jump can be either an intra- or an intermolecular transfer, leading to minor variations in the model. The second jump is always an intramolecular event.

1991). However, this single-stranded region seems to correspond to an overlap of the two plus strand segments rather than a gap (Charneau *et al.*, 1994). Mutagenesis of the central PPT led to a marked decrease in infectivity (Hungnes *et al.*, 1992; Charneau *et al.*, 1992), indicating that the dual priming is important for HIV-1 replication, at least *in vitro*. Additional discontinuities have also been reported to be present in the 3′ half of HIV DNA (Miller & Bushman, 1995).

Reverse transcription is also used in the replication of hepadnaviruses, small DNA viruses of man and animals, and caulimoviruses, a group of DNA viruses of plants. Interestingly, discontinuous plus DNA synthesis has also been described for caulimoviruses, which can contain up to three gaps. In cauliflower mosaic virus, a small proportion of DNA molecules can even contain additional gaps (Maule & Thomas, 1985). In these viruses, actual gaps as well as overlaps of the plus strand segment, as in the case of HIV-1, have been described (reviews: Bao & Hull, 1992, and references therein; Mason *et al.*, 1987). Finally, discontinuous plus DNA has also been described in the Ty1 retrotransposon (Pochart *et al.*, 1993). In all these cases, the PPT sequence is repeated within the genome and provides additional priming sites.

Several hypotheses can be envisaged to explain the significance of the dual initiation mechanism. The plus strand discontinuity might be involved in a stage of the viral infectious cycle subsequent to reverse transcription, as suggested by Hungnes *et al.* (1992). The plus strand discontinuity might constitute, for example, the binding site of a protein involved in the nuclear transport or integration of the provirus. Another hypothesis, this time at the reverse transcription stage itself, was proposed by Charneau & Clavel (1991). They suggested that the dual initiation of plus strand DNA synthesis ensures faster syntheses and faster virus replication. However, it is not clear that polymerization of plus strand DNA is the limiting parameter or that this determines the overall speed of reverse transcription. The two jumps, which imply complementary strand dissociation, or RNase H digestion, followed by reannealing at a different position, could be more important in this respect. However, if dual initiation of plus DNA synthesis is accommodated into the existing model (Fig. 1), the mechanics of the jumps remain unchanged.

A new model has been proposed for the second jump: it requires the occurrence of at least two plus strand DNA priming events and could provide an additional explanation for the PPT primer duplication (Fig. 2). The second jump is classically explained by circularization and pairing of complementary primer binding site sequences, but not strand displacement, which is only needed at the subsequent step to complete the 5′ LTR (Fig. 1*d*). In contrast, Li *et al.* (1993) suggested that in HIV, the second jump occurs via strand displacement of the DNA synthesized from the 3′ PPT, requiring a second priming event (Fig. 2*b*). The displaced fragment (usually termed 'strong stop plus DNA') could subsequently reanneal with its complementary sequence at the 3′ end of the minus strand DNA (Fig. 2*c*). In the context of this new model, initiation of plus strand synthesis upstream from the 3′ PPT is required to ensure the second jump by displacement of strong stop plus DNA. This model predicts that if this fragment 'misses the jump', it could serve as a template for synthesis of a complementary minus sequence, primed by a host cellular tRNA (Fig. 2*c*). The resulting double-stranded DNA fragment, corresponding to unintegrated 'orphan' LTR sequences, has actually been detected in HIV-1 infected cells (Li *et al.*, 1993). A similar DNA fragment was also detected in cells infected by spumaviruses (Kupiec *et al.*, 1988; Renshaw *et al.*, 1991; Tobaly *et al.*, 1991).

It is interesting to note that the retrovirus replication model was established for oncoviruses, which transform cells and survive through single provirus integration into the host genome (lysogenic-like cycle). In contrast, lenti- and spumaviruses establish productive infections with a high quantity of linear unintegrated DNA (lytic-like cycle). However, for MVV and the spumaviruses, the integrated proviral state is, if it ever exists, very difficult to detect (Harris *et al.*, 1981; Kupiec *et al.*, 1988). Recently, it has been shown that spumaviruses encapsidate DNA molecules and replicate through an unusual pathway reminiscent of hepadnaviruses (Yu *et al.*, 1996). It is thus tempting to establish a link between the accumulation of unintegrated DNA and the productive mode of replication. In this context, reconsidering the importance of 'gaps', strand displacements, template transfers and 'orphan LTRs' opens new possibilities for replicative mechanics that justify further investigation.

We should like to thank B. Shacklett, J. Richardson and M. Sitbon for suggestions and critical review of the manuscript, and Bob Naviaux for helpful discussions.

References

Bao, Y. & Hull, R. (1992). Characterization of the discontinuities in rice bacilliform virus DNA. *Journal of General Virology* **73**, 1297–1301.

Blum, H. E., Harris, J. D., Ventura, P., Walker, D., Staskus, K., Retzel, E. & Haase, A. T. (1985). Synthesis in cell culture of the gapped linear duplex DNA of the slow virus Visna. *Virology* **142**, 270–277.

Boone, L. R. & Skalka, A. M. (1981 *a*). Viral DNA synthesized in vitro by avian retrovirus particles permeabilized with mellitin. I. Kinetics of synthesis and size of minus- and plus-strand transcripts. *Journal of Virology* **37**, 109–116.

Boone, L. R. & Skalka, A. M. (1981 *b*). Viral DNA synthesized in vitro by avian retrovirus particles permeabilized with mellitin. II. Evidence for a strand displacement mechanism in plus-strand synthesis. *Journal of Virology* **37**, 117–126.

Bowman, E. H., Pathak, V. K. & Hu, W. S. (1996). Efficient initiation and strand transfer of polypurine tract-primed plus-strand DNA prevent strand transfer of internally initiated plus strand transfer DNAs. *Journal of Virology* **70**, 1687–1694.

Charneau, P. & Clavel, F. (1991). A single-stranded gap in human immunodeficiency virus unintegrated linear DNA defined by a central copy of the polypurine tract. *Journal of Virology* **65**, 2415–2421.

Charneau, P., Alizon, M. & Clavel, F. (1992). A second origin of DNA plus-strand synthesis is required for optimal human immunodeficiency virus replication. *Journal of Virology* **66**, 2814–2820.

Charneau, P., Mirambeau, G., Roux, P., Paulous, S., Buc, H. & Clavel, F. (1994). HIV-1 reverse transcription: a termination step at the center of the genome. *Journal of Molecular Biology* **241**, 651–662.

Coffin, J. M. (1990). Retroviridae and their replication. In *Virology*, pp. 1437–1500. Edited by B. N. Fields & D. M. Knipe. New York: Raven Press.

Finston, W. I. & Champoux, J. J. (1984). RNA-primed initiation of Moloney murine leukemia virus plus strands by reverse transcriptase in vitro. *Journal of Virology* **51**, 26–33.

Harris, J. D., Scott, J. V., Traynor, B., Brahic, M., Stowring, L., Ventura, P., Haase, A. T. & Peluso, R. (1981). Visna virus DNA: discovery of a novel gapped structure. *Virology* **113**, 573–583.

Hsu, T. W. & Taylor, J. M. (1982). Single-stranded regions on unintegrated avian retrovirus DNA. *Journal of Virology* **44**, 47–53.

Huber, H. E. & Richardson, C. C. (1990). Processing of the primer for plus strand DNA synthesis by human immunodeficiency virus type-1 reverse transcriptase. *Journal of Biological Chemistry* **265**, 10565–10573.

Hungnes, O., Tjotta, E. & Grinde, B. (1991). The plus strand is discontinuous in a subpopulation of unintegrated HIV-1 DNA. *Archives of Virology* **116**, 133–141.

Hungnes, O., Tjotta, E. & Grinde, B. (1992). Mutations in the central polypurine tract of HIV-1 result in delayed replication. *Virology* **190**, 440–442.

Kung, H. J., Fung, Y. K., Majors, J. E., Bishop, J. M. & Varmus, H. E. (1981). Synthesis of plus strands of retroviral DNA in cells infected with avaian sarcoma virus and mouse mammary tumor virus. *Journal of Virology* **37**, 127–138.

Kupiec, J. J., Tobaly, J., Canivet, M., Santillana, H. M., Flugel, R. M., Peries, J. & Emanoil, R. R. (1988). Evidence for a gapped linear duplex DNA intermediate in the replicative cycle of human and simian spumaviruses. *Nucleic Acids Research* **16**, 9557–9565.

Kupiec, J. J., Kay, A., Hayat, M., Ravier, R., Peries, J. & Galibert, F. (1991). Sequence analysis of the simian foamy virus type 1 genome. *Gene* **101**, 185–194.

Lee, Y. M. & Coffin, J. M. (1991). Relationship of avian retrovirus DNA synthesis to integration in vitro. *Molecular and Cellular Biology* **11**, 1419–1430.

Li, P., Stephenson, A. J., Kuiper, L. J. & Burrel, C. J. (1993). Double-stranded strong-stop DNA and the second template switch in human immunodeficiency virus (HIV) DNA synthesis. *Virology* **194**, 82–88.

Luo, G., Sharmeen, L. & Taylor, J. (1990). Specificities involved in the initiation of retroviral plus-strand DNA. *Journal of Virology* **64**, 592–297.

Mason, W. S., Taylor, J. M. & Hull, R. (1987). Retroid virus genome replication. *Advances in Virus Research* **32**, 35–96.

Miller, M. D., Wang, B. & Bushman, F. D. (1995). Human immunodeficiency virus type 1 preintegration complexes containing discontinuous plus strands are competent to integrate in vitro. *Journal of Virology* **69**, 3938–3944.

Maule, A. J. & Thomas, C. M. (1985). Evidence from cauliflower mosaic virus virion DNA for additional discontinuities in the plus strand. *Nucleic Acids Research* **13**, 7359–7373.

Pochart, P., Agoutin, B., Rousset, S., Chanet, R., Doroszkiewicz, V. & Heyman, T. (1993). Biochemical and electron microscope analyses of the DNA reverse transcripts present in the virus-like particles of the yeast transposon Ty1. Identification of a second origin of Ty1 DNA plus strand synthesis. *Nucleic Acids Research* **21**, 3513–3520.

Pullen, A. K., Rattray, A. J. & Champoux, J. J. (1993). The sequence features important for plus strand priming by human immunodeficiency virus type 1 reverse transcriptase. *Journal of Biological Chemistry* **268**, 6221–6227.

Rattray, A. J. & Champoux, J. J. (1987). The role of Moloney murine leukemia virus RNAse H activity in the formation of plus-strand primers. *Journal of Virology* **61**, 2843–2851.

Rattray, A. J. & Champoux, J. J. (1989). Plus strand priming by Moloney murine virus. *Journal of Molecular Biology* **208**, 445–456.

Renshaw, R. W., Gonda, M. A. & Casey, J. W. (1991). Structure and transcriptional status of bovine syncytial virus in cytopathic infections. *Gene* **105**, 179–184.

Schweizer, M., Renne, R. & Neumann-Haefelin, D. (1989). Structural analysis of proviral DNA in simian foamy virus (LK-3)-infected cells. *Archives of Virology* **109**, 103–114.

Sonigo, P., Alizon, M., Staskus, K., Klatzmann, D., Cole, S., Danos, O., Retzel, E., Tiollais, P., Haase, A. & Wain-Hobson, S. (1985). Nucleotide sequence of the Visna lentivirus: relationship to the AIDS virus. *Cell* **42**, 369–382.

Taylor, J. M., Cywinski, A. & Smith, J. K. (1983). Discontinuities in the DNA synthesized by an avian retrovirus. *Journal of Virology* **48**, 654–659.

Tobaly, J., Kupiec, J. J., Santillana, H. M., Canivet, M., Peries, J. & Emanoil, R. R. (1991). Further characterization of the gapped DNA intermediates of human spumavirus: evidence for a dual initiation of plus-strand DNA synthesis. *Journal of General Virology* **72**, 605–608.

Varmus, H. E., Heasley, S., Kung, H. J., Oppermann, H., Smith, V. C., Bishop, J. M. & Shank, P. R. (1978). Kinetics of synthesis, structure and purification of avian sarcoma virus-specific DNA made in the cytoplasm of acutely infected cells. *Journal of Molecular Biology* **120**, 55–82.

Wain-Hobson, S., Sonigo, P., Danos, O., Cole, S. & Alizon, M. (1985). Nucleotide sequence of the AIDS virus, LAV. *Cell* **40**, 9–17.

Whitcomb, J. M. & Hughes, S. H. (1992). Retroviral reverse transcription and integration: progress and problems. *Annual Review of Cell Biology* **8**, 275–206.

Wöhrl, B. M. & Moelling, K. (1990). Interaction of HIV-1 ribonuclease H with polypurine tract containing RNA–DNA hybrids. *Biochemistry* **29**, 10141–10147.

Yu, S. F., Baldwin, S. R., Gwynn, S. R., Yendapalli, S. & Linial, M. L. (1996). Human foamy virus replication: a pathway distinct from that of retroviruses and hepadnaviruses. *Science* **271**, 1579–1582.