

Isolation of a Recombinant Murine Leukemia Virus Utilizing a New Primer tRNA

JOHN COLICELLI AND STEPHEN P. GOFF*

Department of Biochemistry and Molecular Biophysics, Columbia University, College of Physicians and Surgeons, New York, New York 10032

Received 3 June 1985/Accepted 11 September 1985

We have previously described the construction of a mutant of Moloney murine leukemia virus bearing a deletion at the normal site of integration of the viral DNA. We have now recovered a revertant of the virus after abortive infection of mouse cells and have determined the structure of the new virus. The revertant is a recombinant virus containing a 500-base-pair patch of new sequences derived from the mouse genome. The integration site was perfectly restored to the wild-type sequence, although the patch of DNA was overall only 80% homologous to Moloney murine leukemia virus. Surprisingly, the tRNA primer binding site was no longer homologous to the usual proline tRNAs, but was a perfect match for glutamine tRNA. This result suggests that the Moloney murine leukemia virus reverse transcriptase is not specific to one tRNA, but can utilize different tRNAs to prime the synthesis of viral DNA. Comparisons with published reports allowed the identification of sequences that are 94% homologous to the patch sequence, present in one of the endogenous retroviral sequences of the mouse. No replication-competent members of this family, utilizing the glutamine tRNA primer, have been previously isolated.

The long terminal repeat (LTR) sequences of the retroviral genome include several elements necessary for the completion of the viral life cycle (60, 61). Portions of the LTR sequences are required for reverse transcription of the viral DNA (11, 24); short sequences at the edges of the LTR are required for the integration of the proviral DNA (4, 38), and a large part of these sequences are required for transcription (2). Nucleotides internal to the LTRs are also essential for replication and are utilized in the initiation of synthesis of each DNA strand during reverse transcription (5, 6, 9, 13, 14, 31, 39, 40, 54). All of these elements are required in *cis* for replication; they constitute recognition sites for the binding of proteins and RNAs to the viral genome. A precise definition of these sequences and their interactions with viral proteins is important to our understanding of the retroviruses.

One of the sequences first identified as important for the replication of the retroviral genome is the binding site for the tRNA which primes synthesis of the first DNA strand during reverse transcription (5, 6, 14, 40). In nearly all mammalian retroviruses, the primer binding site is complementary to the 3' end of two proline tRNA isotypes (13, 40). In the case of the murine viruses, 18 bases in the viral RNA are complementary to this tRNA and are paired in the virion particle with the tRNA to form an RNA-RNA duplex. Although low levels of many tRNAs are found in the virion, the genomic RNA is tightly associated with this particular tRNA, and after the complex is purified the tRNA can be specifically released by heating (13, 67). The avian viruses make similar use of a tryptophan tRNA (5, 6, 14, 62) and show an additional selectivity for their primer tRNA; the avian reverse transcriptase has been shown to specifically bind to this tRNA in vitro and to be essential in vivo for the incorporation of the tRNA into particles (15, 36, 37). This has led to the suggestion that in the avian system reverse transcriptase is involved in the selection of the appropriate tRNA and perhaps in its pairing with the genomic RNA. No

such specific interaction between the murine reverse transcriptase and the proline tRNA has been detected by means of in vitro studies (36).

We have previously described the construction of a mutant of Moloney murine leukemia virus (M-MuLV) bearing an 8-base-pair (bp) deletion at the 3' terminus of the LTR sequence (4). Virions of this mutant cannot carry out a productive infection and are unable to establish the viral DNA in the infected cell. We now report the isolation of a revertant of this virus that arose by recombination with endogenous sequences in the host genome. The recovered sequences include a new tRNA primer binding site for the initiation of synthesis of the first strand of DNA during reverse transcription. We propose that this revertant virus replicates by means of a different primer tRNA.

MATERIALS AND METHODS

Cells and viruses. NIH/3T3 and NRK cells were grown in Dulbecco modified Eagle medium (GIBCO Laboratories) supplemented with 10% calf serum (MA Bioproducts). M-MuLV released by NIH/3T3 cells transfected with pZAP DNA (20) was the source of infectious wild-type virus. Infections were carried out in the presence of 8 μ g of polybrene (Aldrich Chemical Co.) per ml for 2 h at 37°C. Virus production was determined by the reverse transcriptase assay (12). Cells were cloned in 96-well cloning trays (48).

Protein analysis. Infected cells (2×10^6 per 10-cm dish) were starved for 30 min in 2 ml of Dulbecco modified Eagle medium lacking methionine, and [35 S]methionine (200 μ Ci; Amersham Corp.) was added for 15 min. Viral proteins were immunoprecipitated from lysates by using rabbit anti-*gag* serum 507 (gift of D. Baltimore) by standard procedures. Immunoprecipitated viral proteins were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis and fluorography as previously described (47). The proteins were treated with endo- β -*N*-acetylglucosaminidase (gift of P. Robbins) as described previously (59).

DNA manipulations. The methods used for bacterial transformation and plasmid DNA preparation were as described

* Corresponding author.

previously (48). Low-molecular-weight DNA containing the unintegrated viral DNA was prepared by the procedure of Hirt (17) at 30 h postinfection. Viral DNA was cloned into the *Hind*III restriction site of Charon 30A phage DNA. The phages were plated onto LE392 host cells and screened for viral DNA inserts by the procedure of Benton and Davis (1), and inserts of interest were subcloned into the *Hind*III site of the plasmid vector pBR322. The formation of a biologically active clone from the *dl587rev* genome was as follows. The permuted cloned *dl587rev* DNA was cleaved with *Cla*I plus *Xho*I, and the 2.1-kilobase (kb) fragment containing the new sequence was purified. This fragment was used to replace the left end of a wild-type, nonpermuted proviral DNA clone from the left LTR to the *Xho*I site. Preparation of cellular DNAs was as described previously (49). DNAs were analyzed by electrophoresis on agarose gels followed by blot transfer (56) and hybridization (64) to labeled viral probes. Preparation of DNA probes was by nick translation (42). A probe specific for the novel sequences of *dl587rev* was made as follows. The *dl587rev* DNA was digested with *Bal*I plus *Sac*I, the appropriate 220-bp fragment was purified and cloned into the plasmid vector pSP64 cleaved with *Sma*I plus *Sac*I. The resulting DNA was cleaved with *Eco*RI, and labeled runoff RNAs were synthesized with SP6 RNA polymerase (Boehringer Mannheim Biochemicals) in the presence of [α - 32 P]rCTP (Amersham Corp.) and three unlabeled triphosphates (71). DNA sequencing was performed by the method of Maxam and Gilbert (29).

RESULTS

Reversion of a deletion mutant of M-MuLV. We have previously described the construction of a mutant of M-MuLV, *dl587*, bearing a deletion of 8 bp at the 3' edge of the LTRs (4). This mutation disrupts the inverted repeat at the edge of the LTR and removes the highly conserved CA sequence at the site of joining of the viral DNA to host DNA. Producer cells were generated by cotransformation of NIH/3T3 cells with the altered DNA and a selectable marker (69), and four single-cell cloned producer cell lines were isolated. One such line, clone C11, was selected for further study. Cells expressing the mutant DNA were capable of releasing virion particles, but the virions were uninfecious when applied to new cells. The timing of the block was consistent with a block to proviral integration; virus successfully directed the synthesis of the unintegrated proviral DNA but could not establish the DNA in a productive state in the infected cells, and the cells did not become producers of new progeny virus (4).

In an attempt to revert the mutation, we infected NIH/3T3 cells with *dl587* virus and passaged the cells for many days. Supernatant media from the cells were collected and tested for the presence of progeny virus by a reverse transcriptase assay (Fig. 1). Initially no virus was released, but after 16 days high levels of reverse transcriptase appeared in the medium. The recovered virus population, termed *dl587rev*, was phenotypically indistinguishable from wild-type virus when tested for infectivity in the same assay (Fig. 1).

To determine whether retroviral sequences endogenous to the mouse genome were required for the reversion event, we infected normal rat kidney (NRK) cells with *dl587* virus and tested for release of virus as before. No virus were ever recovered from these infections, even after 8 weeks (Fig. 1). The wild-type virus and the revertant population recovered from mouse cell infections were both able to replicate well on the rat cells. Thus, if revertants had arisen they should have been able to spread throughout the rat cell culture and

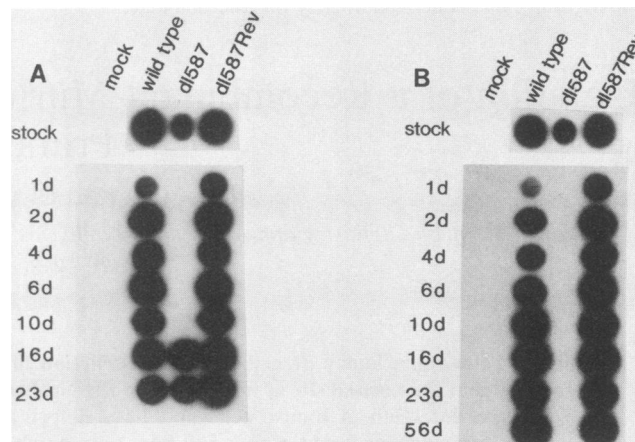


FIG. 1. Kinetics of release of progeny virus after infection with various virus preparations. Stock: Virus was collected from the supernatant medium above cells expressing various proviral DNAs, and reverse transcriptase assays were performed to determine the virus concentration (12). These preparations were used to infect either NIH/3T3 cells (A) or NRK cells (B). The medium was collected at the indicated times after infection and assayed for the presence of progeny virus by reverse transcriptase assay. Infection of NIH/3T3 cells with *dl587* virus resulted in the appearance of revertant virus after about 2 weeks, whereas infection of rat cells with the mutant virus never yielded revertants.

would have been readily detected. These results suggest that mouse cells are uniquely able to give rise to revertants, consistent with the notion that homologous endogenous sequences are required.

Structure of the revertant virus. To determine the basis for the phenotypic reversion, the genome structure of the recovered virus was analyzed. NIH/3T3 cells were infected with the *dl587rev* virus, and after 30 h the DNA fraction enriched for low-molecular-weight material was isolated by the Hirt procedure (17). The DNA was cleaved with a battery of restriction enzymes, the resulting fragments were separated by gel electrophoresis, and the viral DNA fragments were detected by blot hybridization with a viral DNA probe. Normal amounts of all of the viral DNA forms were synthesized by the revertant virus (Fig. 2). The restriction patterns of the new virus were similar to those of the wild type, but there were new sites for cleavage by *Sac*I, *Bgl*II, and *Bgl*III (Fig. 2). Thus, there had been substantial changes in the virus.

The further characterization of the virus was achieved by molecular cloning. The low-molecular-weight DNA was cleaved with *Hind*III and cloned in the phage vector Charon 30A; phage bearing M-MuLV inserts were identified by hybridization. A detailed restriction map of two such clones showed that there were several changes, all clustered in the portion of the genome upstream from the *gag* gene (Fig. 3). No changes elsewhere in the genome were detected. The map of the cloned viruses predicted the appearance of all of the fragments detected in the revertant population by blot hybridization, showing that the clones were representative of the bulk of the replicating virus. The complete DNA sequence of this region from one of the clones was determined and aligned with that of wild-type M-MuLV to provide maximal pairing (Fig. 4). The sequence comparison shows that the revertant virus, as expected, is a recombinant; a patch of sequence approximately 500 bp long had been replaced with homologous, but distinct, sequences.

The sequence immediately revealed why the virus was no

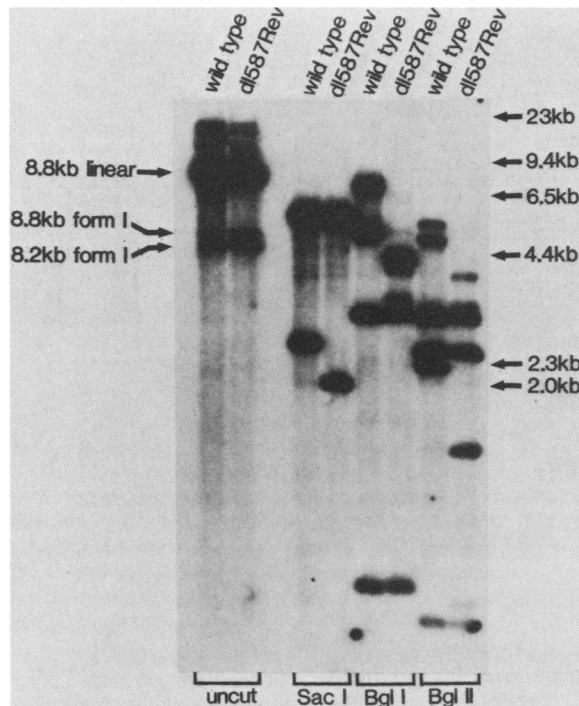


FIG. 2. Southern blot hybridization of preintegrative viral DNAs. NIH/3T3 cells were infected with either wild-type or *dl587rev* virus, and after 24 h DNA was extracted. The DNAs were cleaved with the indicated restriction enzymes and fractionated by agarose gel electrophoresis, and the viral DNAs were detected by blot hybridization. The revertant genome contained new sites for cleavage by *SacI*, *BglII*, and *BglIII*.

longer defective for replication. The new patch of DNA restored the deleted 8 bp perfectly and regenerated a completely wild-type sequence at the LTR edge. The patch was asymmetrically placed over the original deletion, extending from a point in the R region of the LTR, about 120 bp upstream of the deletion, to a point near the *gag* gene, about 400 bp downstream. Several sequence features in this region known to be important for proper expression of the genome were retained, such as the AATAAA signal (41) for polyadenylation (bases 496 through 501) and the splice donor (51) for the *env* mRNA (bases 654 through 660). The new sequence has an average of 80% homology to the parental M-MuLV sequence in the affected area. There are two regions of relatively high nonhomology: one region (bases 670 through 724) immediately downstream from the *env* mRNA splice donor sequence, and another (bases 859 through 952) further downstream, spanning the *PvuI* site of wild-type M-MuLV. The homology in these two regions is only 62%, and the homology excluding these regions is 87%.

A significant difference between the parental M-MuLV sequence and the revertant sequence is a change in the tRNA primer binding site immediately adjacent to the edge of the LTR. In the wild-type M-MuLV, there are 18 bases which are perfectly complementary to the 3' end of the two proline tRNA isotypes present in mammalian cells. In the recombinant, 5 of the 18 bases have been changed (Fig. 4), such that the binding site sequence would no longer be expected to pair with proline tRNA. A comparison of the new binding site with the 3' ends of known mammalian tRNAs (10) revealed that only one of the sequenced tRNAs is homologous: the sequence is a perfect match (18 of 18) for

pairing with mammalian glutamine tRNA (70) (Fig. 5). These results suggest that the recombinant virus replicates with a glutamine tRNA primer rather than the usual proline tRNA.

New sequences from the cloned *dl587rev* genome are biologically active. The cloned genome of the revertant virus contains many sequence differences from the parental M-MuLV and might be representative of a minor, defective subpopulation of the total replicating virus. To determine whether the cloned genome was functional, the permuted viral DNA was excised from the vector by cleavage with *HindIII*, ligated to form oligomers, and introduced into NIH/3T3 cells by the DEAE-dextran procedure (30). No replicating virus could be recovered as judged by XC plaque assay (44) or by reverse transcriptase assay (12). Some portion of the cloned DNA was therefore defective. In our experience, about half of all cloned viral genomes contain lethal defects, probably due to errors occurring during reverse transcription (12), and it was possible that the defect lay outside the new sequences. To test this, a 2.1-kb fragment containing all of the new sequences was produced by cleavage with *ClaI* plus *XhoI* and used to replace the 5' end of a full-length cloned proviral DNA. The resulting cloned DNA was applied to NIH/3T3 cells as before and was found to be as infectious as the wild-type DNA by both XC plaque assay (44) and reverse transcriptase assay (12). Thus, the new sequences were fully functional and were able to interact correctly with the gene products encoded by the wild-type M-MuLV genome.

Analysis of *gag* proteins encoded by the revertant virus. Wild type M-MuLV encodes two distinct *gag* polypeptides. A cytoplasmic protein termed Pr65^{gag} is the major species, and a glycosylated membrane protein termed gPr80^{gag} is made in smaller amounts. Translation of the major protein is initiated at an AUG codon at nucleotide position 1070 of the proviral DNA, whereas the glycosylated protein contains extra sequences at the N terminus which are encoded in the region between the LTR and the initiator codon for Pr65^{gag} (8, 46, 47). The glycosylated protein is known to be dispensable for replication in culture (47). To determine whether the *dl587rev* virus could direct the synthesis of both proteins, cells were pulse-labeled with [³⁵S]methionine, total proteins

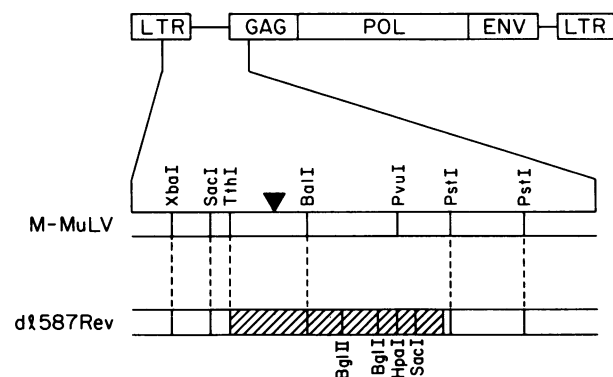


FIG. 3. Restriction map of the cloned *dl587rev* genome. The top line shows a genetic map of the wild-type M-MuLV genome as a linear proviral DNA. The second line shows a restriction map of wild-type M-MuLV in the region from the 5' LTR to the 5' part of the *gag* gene. The position of the 8-bp deletion in *dl587* is indicated by the solid triangle. The third line shows a restriction map of *dl587rev* in the region containing new substituted sequences. The region containing nucleotide differences from Mu-MuLV is indicated by the hatched area.

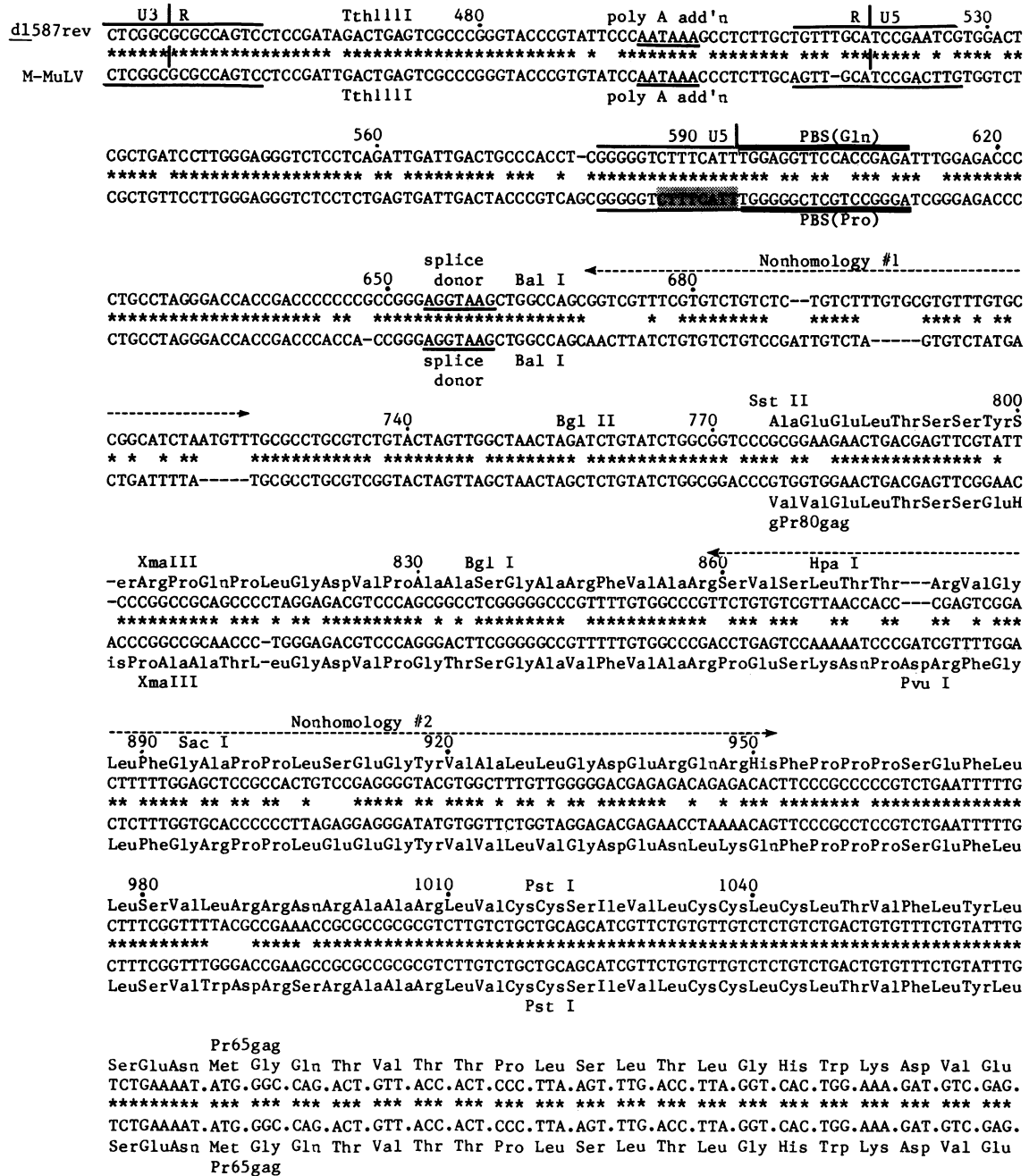


FIG. 4. Sequence of the substitution region of *d1587rev* and comparison with sequence of homologous region of M-MuLV. The upper sequence is *d1587rev*, and the lower sequence is M-MuLV. Asterisks indicate the conserved bases. Dashes in the sequences are gaps introduced to achieve maximal pairing of the two sequences. The nucleotides of the new sequence are numbered relative to the left edge of the 5' LTR. The boundaries of the U3, R, and U5 regions are indicated by vertical lines; restriction sites and other important features of the sequences, such as the primer binding site, are indicated. The 8 bases deleted in *d1587* at the edge of the LTR are shaded. The two regions of greatest nonhomology are delimited by the broken lines. There are three differences between the sequence of our wild-type M-MuLV (lower sequence) and the published sequence (53): bases 449, 628, and 1066.

were extracted, and the *gag* proteins were immunoprecipitated and analyzed by sodium dodecyl sulfate-gel electrophoresis followed by fluorography. Normal levels of Pr65^{gag} were synthesized by the mutant, but no glycosylated *gag* protein could be detected (Fig. 6). Treatment of the extracts with endoglycosidase H resulted in the normal shift in mobility of the wild-type gPr80^{gag} and did not reveal any changes in the mutant *gag* proteins. One or more of the

changes between the parental M-MuLV and *d1587rev* apparently prevented the synthesis of the glycosylated *gag* protein.

Origin of the new sequences. The source of the novel sequences in the revertant virus was likely to be sequences present in the mouse genome and lacking in the rat genome. Hybridization experiments were performed to test this notion. A 220-bp DNA fragment within the novel sequences



FIG. 5. Primer binding site sequences of M-MuLV and *dl587rev* and pairing with primer tRNAs. The upper lines show the sequence of M-MuLV near the edge of the 5' LTR. The last bases of the LTR are boxed. The next 18 bases are homologous to the 3' end of proline tRNA. The lower lines show the sequence of *dl587rev* in the corresponding region. The corresponding 18 bases are homologous to the 3' end of glutamine tRNA (70). In both cases the next base is 1-methyl adenine, which cannot participate in normal base pairing.

released by cleavage of the revertant genome with *BalI* plus *SacI* was isolated and subcloned into the plasmid vector pSP64 downstream from a phage SP6 transcriptional promoter. The plasmid was used as template for the formation of radioactively labeled runoff transcripts, and the labeled RNA was used to probe a Southern blot of mouse DNAs. Many fragments with homology to the probe were produced by cleavage of NIH/3T3, BALB/c, or C57BL6 DNAs with any of three restriction enzymes (Fig. 7). The labeling was retained even at stringent washing. Rehybridization of the blot with a labeled DNA probe representative of a similar region of the wild type M-MuLV genome (the 500-bp fragment released by cleavage with *BalI*) revealed a similar

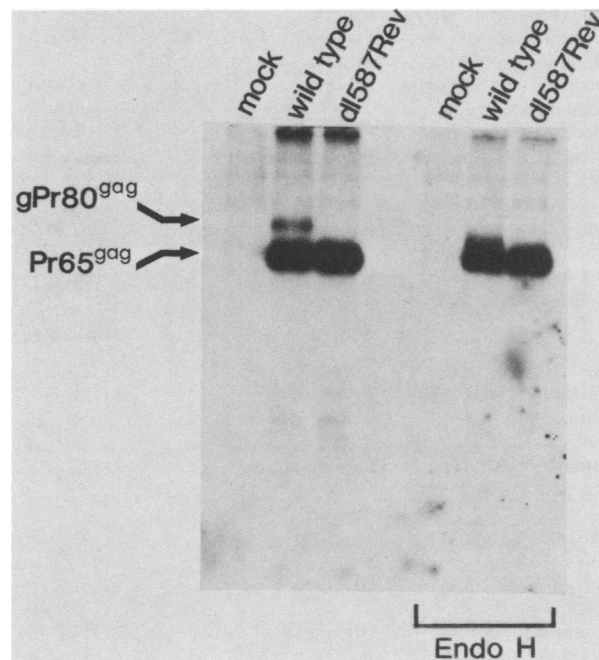


FIG. 6. *gag* gene products encoded by wild-type M-MuLV and *dl587rev* genomes. Cells infected with the indicated virus were metabolically labeled with [³⁵S]methionine, and *gag* gene products were examined by immunoprecipitation, sodium dodecyl sulfate-polyacrylamide gel electrophoresis, and fluorography. Samples on the right were treated with endoglycosidase H (Endo H) before electrophoresis. Whereas M-MuLV directed the synthesis of both glycosylated and nonglycosylated *gag* proteins, *dl587* made only the nonglycosylated form.

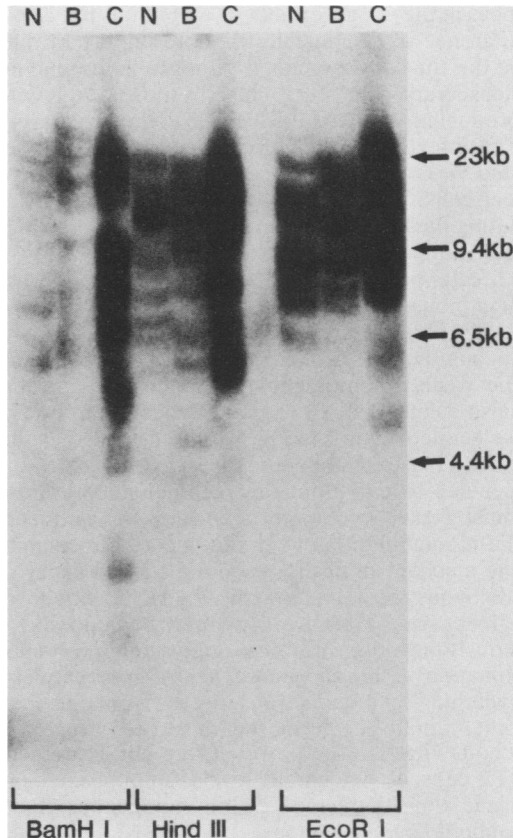


FIG. 7. Detection of endogenous sequences in mouse genomic DNA showing homology to *dl587rev* by Southern blot hybridization. Genomic DNA from various mouse strains was cleaved with the indicated restriction enzymes, fractionated by electrophoresis, blotted, and hybridized to a radioactive RNA probe corresponding to the new sequences in *dl587rev*. Lanes: N, NIH/3T3 cell line DNA; B, BALB/c liver DNA; C, C57BL6 liver DNA. Twice as much DNA was loaded in lanes C. The positions of migration of DNA standards are indicated.

pattern of homologous bands (data not shown). Thus, the mouse genome does contain sequences homologous to the novel DNA in *dl587rev*. These sequences are probably the endogenous retroviral DNAs which are also detected by M-MuLV probes. The results suggest that one of these endogenous DNAs served as donor for the repair event.

DISCUSSION

It has been known for many years that replicating retroviruses can recombine with homologous sequences of the host (66, 68). The structure of the *dl587rev* genome indicates that reversion of a lethal deletion mutation can also occur by recombination of a mutant retrovirus with related, but distinct, sequences present in the mouse genome. The reversion event was similar in many ways to the reversion of a larger deletion in the *pol* gene, *dl5401*, which we reported previously (49). The size of the patches of new DNA in these two cases were very similar (530 and 387 bp), and the mechanism by which they arose may well be the same. In the present case the revertant virus appeared much sooner after infection; only 2 weeks rather than 6 weeks was required to generate a replicating virus. The reason for the difference in timing is not clear. The novel sequences in *dl587rev* are slightly more homologous to M-MuLV (80%) than are the

sequences in the *pol* revertant (70%), but it is not certain that this difference is responsible. A more important difference may be the titer of the applied virus. Cells producing *dl587* can release very high titers, close to those of the wild type; cells producing *dl5401* make 1/10 to 1/20 of wild-type levels (48). The infections in the two cases were performed with undiluted virus preparations, directly as harvested from the producer cells. Thus, the number of virus particles capable of initiating the recombination event may be much higher for the *dl587* mutant than for the *dl5401* mutant.

The results presented here show that reversion occurred rapidly in mouse cells and did not occur at all after infection of rat cells. Thus, the reversion depends on the existence of homologous sequences in the murine host and cannot occur with the weakly homologous sequences in rat cells. This result also suggests that the event depends on the presence of these sequences in the recipient cell, and that the virus cannot make use of the sequences in the genome of the producer cell to carry out the recombination. The mutant virus must persist in some cryptic form in the infected cell until the recombination event can occur. We cannot determine the mechanism of the reaction; it could occur directly by homologous recombination of viral DNA with host DNA. This direct recombination would be unusual, since exogenously introduced DNAs rarely undergo homologous recombination with endogenous genes. Alternatively, template switching by reverse transcriptase from one viral RNA to another could result in the observed structures. The endogenous viral DNAs, however, are not expressed in the NIH/3T3 cells at detectable levels (unpublished observations), and thus there may be no endogenous RNAs that could undergo such reactions.

The source of the new sequences is uncertain, but must be one of the many endogenous retroviral DNAs. Examination of the structures of several cloned endogenous sequences shows that many do contain the *SacI* and *BglII* sites of *dl587rev* (22, 25, 43). The recovery of recombinant viruses containing these 5' sequences, together with previous results demonstrating the acquisitions of other regions of the genome, shows that multiple regions of the endogenous viral genomes are functional. It seems likely that the mouse genome may contain sufficient endogenous proviral DNAs to provide at least one functional copy of virtually any region of the retroviral genome. The formation of these chimeric viruses may play a significant role in the evolution of the murine retroviruses.

The most striking difference between the revertant virus and the parental M-MuLV is the change in the tRNA primer binding site; the new virus contains a sequence which is homologous to glutamine tRNA rather than the usual proline tRNA. We feel that the new site, with 28% nonhomology, is unlikely to be compatible with the use of the original proline tRNA for replication. The virus almost certainly uses the glutamine tRNA as primer, because the primer binding site is in fact regenerated from the tRNA itself during reverse transcription (11). The new sequence is clearly functional, since this region of the cloned viral genome can be used to form a biologically active cloned DNA. Furthermore, it is clear that this position on the genome is indeed used to initiate DNA synthesis and form the 3' end of the minus strand, since analysis of the unintegrated DNA of the virus (Fig. 2) revealed that the termini of the full-length linear DNA were unaltered from the wild type relative to the conserved restriction sites. The result conclusively shows that there is no specific interaction between the murine reverse transcriptase and the proline primer tRNA and

suggests that different tRNAs may serve equally well as primers. This is in agreement with previous binding studies which failed to detect any specific interaction between the murine enzyme and proline tRNA (36). The situation appears to be different in those avian viruses for which specific binding of the enzyme to the primer tRNA has been demonstrated (15, 36, 37).

The *dl587rev* virus is the first replication-competent virus utilizing glutamine tRNA as primer. A summary of the primer tRNAs used by various retroviruses is given in Table 1. The vast majority of the mammalian viruses, including murine, feline, and primate viruses, utilize proline tRNA. Several viruslike elements from the mouse genome have been cloned and sequenced. The intracisternal A-type particle genomes (26) include a primer binding site homologous to phenylalanine, and examples of the replication-defective viruslike 30S elements have been found with primer binding sites which match with proline, glycine, and glutamine (19, 21). The literature includes one extensive sequence of a defective element, AL10, endogenous to BALB/c mice, which has been noted as containing a near-perfect match (17 of 18) of the primer binding site with glutamine tRNA (35). If this endogenous element were related to the element that served as donor in the formation of *dl587rev*, then the homology should extend over much of the novel sequences

TABLE 1. Primer tRNAs utilized by various retroviruses

Retrovirus	Primer tRNA	Infectious	References
Murine leukemia viruses ^a	Proline	+	(3, 16, 53) (13, 40, 63)
Feline leukemia virus	Proline	+	(23)
Bovine leukemia virus	Proline	+	(45)
Simian sarcoma virus	Proline	+	(7)
Human T-cell leukemia viruses I and II	Proline	+	(50, 55)
Spleen necrosis virus	Proline	+	(52)
Viruslike 30S genes VL3, VM1	Proline	-	(21)
Mouse mammary tumor virus	Lysine	+	(27, 39)
Human t-cell lymphotropic virus III	Lysine	+	(65)
Rous sarcoma virus	Tryptophan	+	(5, 6, 14, 58)
Endogenous leukosis virus ev1	Tryptophan	-	(18)
Intracisternal A-type particle Gene	Phenylalanine	-	(34)
Viruslike 30S gene BVL-1	Glycine	-	(19)
Viruslike 30S gene NVL-3	Glycine	-	(33)
Viruslike 30S genes VL5, VL6	Glycine	-	(21)
Endogenous sequences AL10, RFM1, 9, 16, 17	Glutamine	-	(32, 35)
Viruslike 30S gene VL8	Glutamine	-	(21)
<i>dl587rev</i>	Glutamine	+	This work

^a M-MuLV, Gross and Friend murine leukemia viruses, AKV, and spleen focus-forming virus.

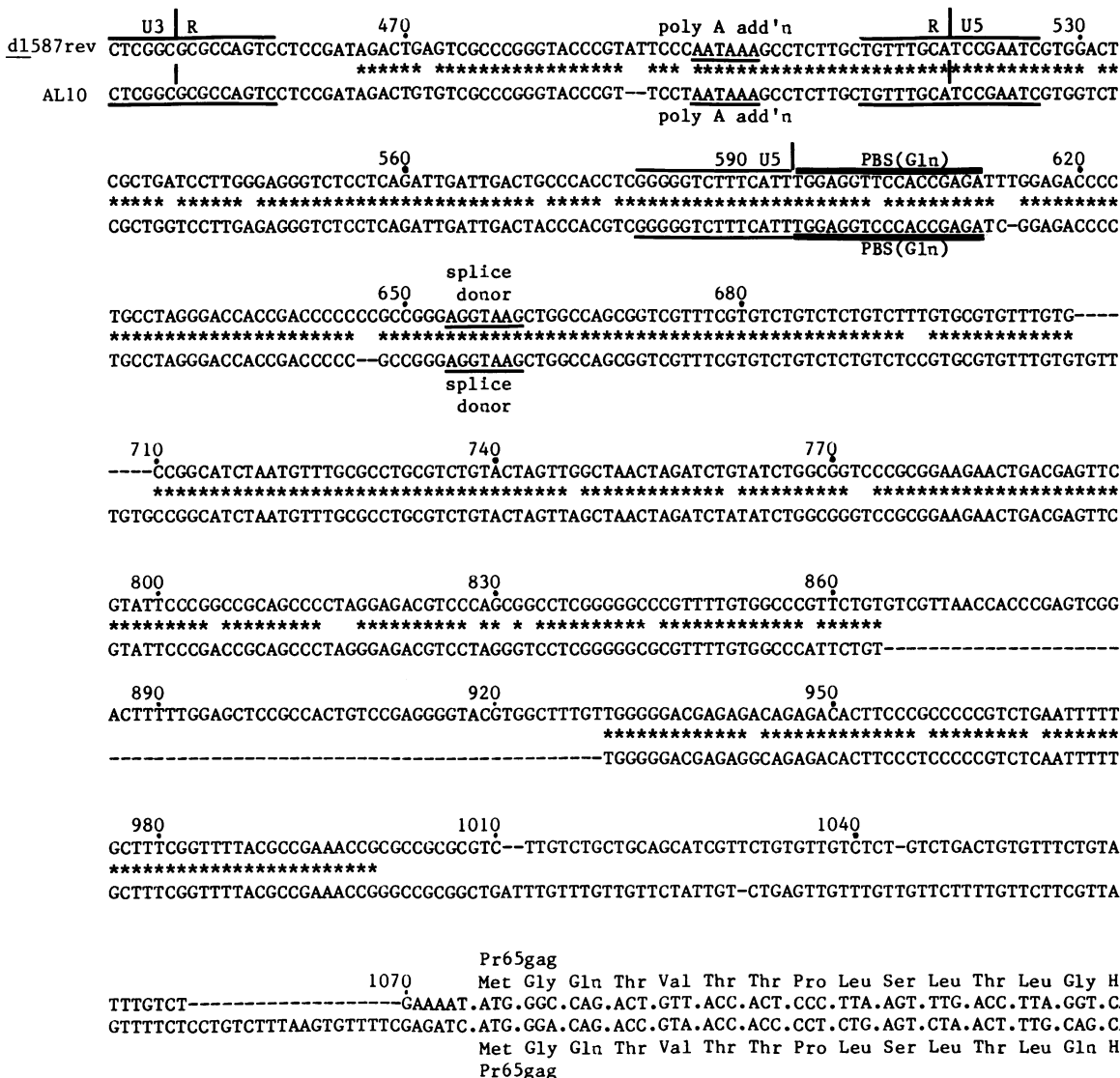


FIG. 8. Comparison of DNA sequence of *d1587rev* and AL10, an endogenous retroviral genome (35). The upper sequence is *d1587rev*, and the lower sequence is AL10. Asterisks indicate matches between the two sequences and are used only in the region in which *d1587rev* sequences differ from M-MuLV. Dashes are gaps introduced to achieve maximal pairing of the two sequences.

in the revertant. Comparison of AL10 with the revertant indeed reveals extensive homology (Fig. 8). The AL10 sequence contains a large deletion relative to *d1587rev*, but in the regions which can be aligned the sequences are 94% homologous. This finding suggests that an endogenous virus quite similar to AL10 was the donor for the recombination event. The sequences may belong to one of the two major families of endogenous sequences in mouse cells (57), since many cloned copies exhibit restriction maps similar to that of *d1587rev* (22, 43). Recently, four more proviral clones endogenous to the RFM/un mouse have been isolated and characterized (32). The primer binding sites for these four genomes have been determined: three of the sites are perfect (18 of 18) matches for glutamine tRNA, and the fourth is a 16 of 18 match. These results suggest that the majority of the endogenous proviral DNAs may use this tRNA as primer.

There are many other major differences between M-MuLV and the new sequences in *d1587*. One region (nonhomology no. 1 in Fig. 4) contains numerous changes in the region thought to be essential for packaging of virion

RNA into particles (28, 47); this subregion may turn out to be dispensable for this process. It is also possible that the secondary and tertiary structures of the wild-type and revertant sequences are similar. Such features as the site for polyadenylation of the viral RNA, the repeat sequences at the edge of the LTR, and the splice donor for the *env* mRNA were all retained in the recombinant.

The failure of the new virus to produce the glycosylated *gag* protein gPr80^{gag} sheds some light on the means by which this protein is formed in the wild type. The origin of the glycosylated *gag* protein has been unclear because there are no AUG codons in the appropriate reading frame in the region, and no spliced mRNAs have been detected which would remove this problem. It has been suggested that a GUG codon might be utilized to initiate translation (47; I. M. Verma, personal communication), since a GUG codon is found at the appropriate place and reading frame in M-MuLV and in the related AKV (16). Examination at the sequence of the recombinant virus showed that both of two GUG codons found at positions 767 through 772 have been

destroyed by mutation, and the glycosylated *gag* protein encoded by the wild-type M-MuLV is apparently not made by the revertant. Thus, this result is consistent with the notion that the wild-type virus uses the GUG codons to initiate translation of the protein and confirms the observation that the gPr80^{gag} protein is dispensable.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant CA 30488 from the National Cancer Institute and by an award from the Chicago Community Trust/Searle Scholars Program. Salary support for S.P.G. was provided by the Irma T. Hirsch Trust.

LITERATURE CITED

- Benton, W. D., and R. W. Davis. 1977. Screening Lambda gt recombinant clones by hybridization to single plaques in situ. *Science* **196**:180-182.
- Benz, E. W., Jr., R. M. Wydro, B. Nadal-Ginard, and D. Dina. 1980. Moloney murine sarcoma proviral DNA is a transcriptional unit. *Nature (London)* **288**:665-669.
- Clark, S. P., and T. W. Mak. 1982. Nucleotide sequences of the murine retrovirus Friend SFFVp long terminal repeats: identification of a structure with extensive dyad symmetry 5' to the TATA box. *Nucleic Acids Res.* **10**:3315-3330.
- Colicelli, J., and S. P. Goff. 1985. Mutants and pseudorevertants of Moloney murine leukemia virus with alterations at the integration site. *Cell* **42**:573-580.
- Cordell, B., E. Stavnezer, R. Friedrich, J. M. Bishop, and H. M. Goodman. 1976. Nucleotide sequence that binds primer for DNA synthesis to the avian sarcoma virus genome. *J. Virol.* **19**:548-558.
- Dahlberg, J. E., R. C. Sawyer, J. M. Taylor, A. J. Faras, W. E. Levinson, H. M. Goodman, and J. M. Bishop. 1974. Transcription of DNA from the 70S RNA of Rous sarcoma virus. I. Identification of a specific 4S RNA which serves as primer. *J. Virol.* **13**:1126-1133.
- Devare, S. G., E. P. Reddy, J. D. Law, K. C. Robbins, and S. A. Aaronson. 1983. Nucleotide sequence of the simian sarcoma virus genome: demonstration that its acquired cellular sequences encode the transforming gene product p28^{src}. *Proc. Natl. Acad. Sci. USA* **80**:731-735.
- Edwards, S. A., and H. Fan. 1980. Sequence relationship of glycosylated and unglycosylated *gag* polyproteins of Moloney murine leukemia virus. *J. Virol.* **35**:41-51.
- Finston, W. I., and J. J. Champoux. 1984. RNA-primed initiation of Moloney murine leukemia virus plus strands by reverse transcriptase in vitro. *J. Virol.* **51**:26-33.
- Gauss, D. H., and M. Sprinzl. 1983. Compilation of tRNA sequences. *Nucleic Acids Res.* **11**:1-53.
- Gilboa, E., S. W. Mitra, S. Goff, and D. Baltimore. 1979. A detailed model of reverse transcription and tests of crucial aspects. *Cell* **18**:93-100.
- Goff, S. P., P. Traktman, and D. Baltimore. 1981. Isolation and properties of Moloney murine leukemia virus mutants: use of a rapid assay for release of virion reverse transcriptase. *J. Virol.* **38**:239-248.
- Harada, F., G. G. Peters, and J. E. Dahlberg. 1979. The primer tRNA for Moloney murine leukemia virus DNA synthesis. *J. Biol. Chem.* **254**:10979-10985.
- Harada, F., R. C. Sawyer, and J. E. Dahlberg. 1975. A primer ribonucleic acid for initiation of *in vitro* Rous sarcoma virus deoxyribonucleic acid synthesis. *J. Biol. Chem.* **250**:3487-3497.
- Haseltine, W. A., A. Panet, D. Smoler, D. Baltimore, G. Peters, F. Harada, and J. E. Dahlberg. 1977. Interaction of tryptophan tRNA and avian myeloblastosis virus reverse transcriptase: further characterization of the binding reaction. *Biochemistry* **16**:3625-3632.
- Herr, W. 1984. Nucleotide sequence of AKV murine leukemia virus. *J. Virol.* **49**:471-478.
- Hirt, B. 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. *J. Mol. Biol.* **26**:365-371.
- Hishinuma, F., P. J. DeBona, S. Astrin, and A. M. Skalka. 1981. Nucleotide sequence of acceptor site and termini of integrated avian endogenous provirus evl: integration creates a 6 bp repeat of host DNA. *Cell* **23**:155-164.
- Hodgson, C. P., P. K. Elder, T. Ono, D. N. Foster, and M. J. Getz. 1983. Structure and expression of mouse VL30 genes. *Mol. Cell. Biol.* **3**:2221-2231.
- Hoffmann, J. W., D. Steffen, J. Gusella, C. Tabin, S. Bird, D. Cowing, and R. A. Weinberg. 1982. DNA methylation affecting the expression of murine leukemia proviruses. *J. Virol.* **44**:144-157.
- Itin, A., and E. Keshet. 1985. Primer binding sites corresponding to several tRNA species are present in DNAs of different members of the same retrovirus-like gene family (VL30). *J. Virol.* **54**:236-239.
- Khan, A. S., W. P. Rowe, and M. A. Martin. 1982. Cloning of endogenous murine leukemia virus-related sequences from chromosomal DNA of BALB/c and AKR/J mice; identification of an *env* progenitor of AKR-247 mink cell focus-forming proviral DNA. *J. Virol.* **44**:625-636.
- Laprevotte, I., A. Hampe, C. J. Sherr, and F. Galibert. 1984. Nucleotide sequence of the *gag* gene and *gag-pol* junction of feline leukemia virus. *J. Virol.* **50**:884-894.
- Lobel, L. I., and S. P. Goff. 1984. Reverse transcription of retroviral genomes: mutations in the terminal repeat sequences. *J. Virol.* **53**:447-455.
- Lowy, D. R., E. Rands, S. K. Chattopadhyay, C. F. Garon, and G. L. Hager. 1980. Molecular cloning of infectious integrated murine leukemia virus DNA from infected mouse cells. *Proc. Natl. Acad. Sci. USA* **77**:614-618.
- Lueders, K. K., and E. L. Kuff. 1977. Sequences associated with intracisternal A particles are reiterated in the mouse genome. *Cell* **12**:963-972.
- Majors, J. E., and H. E. Varmus. 1983. Nucleotide sequencing of an apparent proviral copy of *env* mRNA defines determinants of expression of the mouse mammary tumor virus *env* gene. *J. Virol.* **47**:495-504.
- Mann, R., R. Mulligan, and D. Baltimore. 1983. Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. *Cell* **33**:153-159.
- Maxam, A., and W. Gilbert. 1980. Sequencing end-labelled DNA with base-specific chemical cleavages. *Methods Enzymol.* **65**:499-599.
- McCutchan, J. H., and J. S. Pagano. 1968. Enhancement of the infectivity of SV40 deoxyribonucleic acid with diethylaminoethyl-dextran. *J. Natl. Cancer Inst.* **41**:351-357.
- Mitra, S. W., M. Chow, J. Champoux, and D. Baltimore. 1982. Synthesis of murine leukemia virus plus strong stop DNA initiates at a unique site. *J. Biol. Chem.* **257**:5983-5986.
- Nikbakht, K. N., C.-Y. Ou, L. R. Boone, P. L. Glover, and W. K. Yang. 1985. Nucleotide sequence analysis of endogenous murine leukemia virus-related proviral clones reveals primer-binding sites for glutamine tRNA. *J. Virol.* **54**:889-893.
- Norton, J. D., J. Connor, and R. J. Avery. 1984. Unusual long terminal repeat sequence of a retrovirus transmissible mouse (VL30) genetic element: identification of functional domains. *Nucleic Acids Res.* **12**:3445-3459.
- Ono, M., and H. Ohishi. 1983. Long terminal repeat sequences of intracisternal A particle genes in the Syrian hamster genome: identification of tRNA Phe as a putative primer tRNA. *Nucleic Acids Res.* **11**:7169-7179.
- Ou, C., L. R. Boone, and W. K. Yang. 1983. A novel sequence segment and other nucleotide structural features in the long terminal repeat of a BALB/c mouse genomic leukemia virus-related DNA clone. *Nucleic Acids Res.* **11**:5603-5620.
- Panet, A., and H. Berliner. 1978. Binding of tRNA to reverse transcriptase of RNA tumor viruses. *J. Virol.* **26**:214-220.
- Panet, A., W. A. Haseltine, D. Baltimore, G. Peters, F. Harada, and J. E. Dahlberg. 1975. Specific binding of tryptophan transfer RNA to avian myeloblastosis virus RNA-dependent DNA polymerase (reverse transcriptase). *Proc. Natl. Acad. Sci. USA* **72**:2535-2539.
- Panganiban, A. T., and H. M. Temin. 1983. The terminal

- nucleotides of retrovirus DNA are required for integration but not virus production. *Nature* (London) **306**:155-160.
39. Peters, G., and C. Glover. 1980. tRNA's and priming of RNA-directed DNA synthesis in mouse mammary tumor virus. *J. Virol.* **35**:31-40.
 40. Peters, G., F. Harada, J. E. Dahlberg, A. Panet, W. A. Haseltine, and D. Baltimore. 1977. Low-molecular-weight RNAs of Moloney murine leukemia virus: identification of the primer for RNA-directed DNA synthesis. *J. Virol.* **21**:1031-1041.
 41. Proudfoot, N. J., and G. G. Brownlee. 1974. Sequence at the 3' end of globin mRNA shows homology with immunoglobulin light chain mRNA. *Nature* (London) **252**:359-362.
 42. Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg. 1977. Labelling of DNA to high specific activity by nick translation. *J. Mol. Biol.* **113**:237-258.
 43. Roblin, R., J. M. Young, R. J. Mural, T. E. Bell, and J. N. Ihle. 1982. Molecular cloning and characterization of murine leukemia virus-related DNA sequences from C3H/HeN mouse DNA. *J. Virol.* **43**:113-126.
 44. Rowe, W. P., W. E. Pugh, and J. W. Hartley. 1970. Plaque assay techniques for murine leukemia viruses. *Virology* **42**:1136-1139.
 45. Sagata, N., T. Yasunaga, Y. Ogawa, J. Tsuzuku-Kawamura, and Y. Ikawa. 1984. Bovine leukemia virus: unique structural features of its long terminal repeats and its evolutionary relationship to human T-cell leukemia virus. *Proc. Natl. Acad. Sci. USA* **81**:4741-4745.
 46. Saris, C. J. M., J. van Eenbergen, R. M. J. Liskamp, and H. P. J. Bloemers. 1983. Structure of glycosylated and unglycosylated *gag* and *gag-pol* precursor proteins of Moloney murine leukemia virus. *J. Virol.* **46**:841-859.
 47. Schwartzberg, P., J. Colicelli, and S. P. Goff. 1983. Deletion mutants of Moloney murine leukemia virus which lack glycosylated *gag* protein are replication competent. *J. Virol.* **46**:538-546.
 48. Schwartzberg, P., J. Colicelli, and S. P. Goff. 1984. Construction and analysis of deletion mutations in the *pol* gene of Moloney murine leukemia virus: a new viral function required for productive infection. *Cell* **37**:1043-1052.
 49. Schwartzberg, P., J. Colicelli, and S. P. Goff. 1985. Recombination between a defective retrovirus and homologous sequences in host DNA: reversion by patch repair. *J. Virol.* **53**:719-726.
 50. Seiki, M., S. Hattori, Y. Hirayama, and M. Yoshida. 1983. Human adult T-cell leukemia virus: complete nucleotide sequence of the provirus genome integrated in leukemia cell DNA. *Proc. Natl. Acad. Sci. USA* **80**:3618-3622.
 51. Sharp, P. A. 1981. Speculations on RNA splicing. *Cell* **23**:643-646.
 52. Shimotohno, K., S. Mizutani, and H. M. Temin. 1980. Sequence of retrovirus provirus resembles that of bacterial transposable elements. *Nature* (London) **285**:550-554.
 53. Shinnick, T. M., R. A. Lerner, and J. G. Sutcliffe. 1981. Nucleotide sequence of Moloney murine leukemia virus. *Nature* (London) **293**:543-548.
 54. Smith, J. K., A. Cywinski, and J. M. Taylor. 1984. Specificity of initiation of plus-strand DNA by Rous sarcoma virus. *J. Virol.* **52**:314-319.
 55. Sodroski, J., M. Trus, D. Perkins, R. Patarca, F. Wong-Staal, E. Gelmann, R. Gallo, and W. A. Haseltine. 1984. Repetitive structure in the long-terminal-repeat element of a type II human T-cell leukemia virus. *Proc. Natl. Acad. Sci. USA* **81**:4617-4621.
 56. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**:503-517.
 57. Steffen, D. L., R. Mural, D. Cowing, J. Mielcarz, J. Young, and R. Roblin. 1982. Most of the murine leukemia virus sequences in the DNA of NIH/Swiss mice consist of two closely related proviruses, each repeated several times. *J. Virol.* **43**:127-135.
 58. Swanstrom, R., H. E. Varmus, and J. M. Bishop. 1982. Nucleotide sequence of the 5' noncoding region and part of the *gag* gene of Rous sarcoma virus. *J. Virol.* **41**:535-541.
 59. Tarentino, A. L., and F. Maley. 1974. Purification and properties of an endo B-N-acetyl-glucosaminidase from *Streptomyces griseus*. *J. Biol. Chem.* **249**:811-816.
 60. Temin, H. M. 1981. Structure, variation and synthesis of retrovirus long terminal repeat. *Cell* **27**:1-3.
 61. Varmus, H. E. 1982. Form and function of retroviral proviruses. *Science* **216**:812-820.
 62. Verma, I. M., N. L. Meuth, and D. Baltimore. 1972. Covalent linkage between ribonucleic acid primer and deoxyribonucleic acid product of the avian myeloblastosis virus deoxyribonucleic acid polymerase. *J. Virol.* **45**:622-627.
 63. Villemur, R., E. Rassart, L. DesGrosseillers, and P. Jolicœur. 1983. Molecular cloning of viral DNA from leukemogenic Gross passage A murine leukemia virus and nucleotide sequence of its long terminal repeat. *J. Virol.* **45**:539-546.
 64. Wahl, G. M., M. Stern, and G. R. Stark. 1979. Efficient transfer of large DNA fragments from agarose gels to diazobenzoyloxymethyl paper and rapid hybridization using dextran sulphate. *Proc. Natl. Acad. Sci. USA* **76**:3683-3687.
 65. Wain-Hobson, S., P. Sonigo, O. Danos, S. Cole, and M. Alizon. 1985. Nucleotide sequence of the AIDS virus, LAV. *Cell* **40**:9-17.
 66. Wang, L.-H., C. C. Halpern, M. Nadel, and H. Hanafusa. 1978. Recombination between viral and cellular sequences generates transforming sarcoma virus. *Proc. Natl. Acad. Sci. USA* **75**:5812-5816.
 67. Waters, L. C., B. C. Mullin, E. G. Bailiff, and R. A. Popp. 1980. Differential association of transfer RNAs with the genome of murine, feline, and primate retroviruses. *Biochim. Biophys. Acta* **608**:112-126.
 68. Weiss, R. A., W. S. Mason, and P. K. Vogt. 1973. Genetic recombinants and heterozygotes derived from endogenous and exogenous avian RNA tumor viruses. *Virology* **52**:535-552.
 69. Wigler, M., R. Sweet, G. K. Sim, B. Wold, A. Pellicer, E. Lacy, T. Maniatis, S. Silverstein, and R. Axel. 1979. Transformation of mammalian cells with genes from procaryotes and eukaryotes. *Cell* **16**:777-785.
 70. Yang, J. A., L. W. Tai, P. F. Agris, C. W. Gehrke, and T. Wong. 1983. The nucleotide sequence of a major glutamine tRNA from rat liver. *Nucleic Acids Res.* **11**:1991-1996.
 71. Zinn, K., D. DiMaio, and T. Maniatis. 1983. Identification of two distinct regulatory regions adjacent to the human beta-interferon gene. *Cell* **34**:865-879.