MOLECULAR AND CELLULAR BIOLOGY, June 1988, p. 2328–2334
0270-7306/88/062328-07$02.00/0

Vol. 8, No. 6

Retroviral Vector System for the Study of cDNA Gene Formation

RALPH DORNBURG AND HOWARD M. TEMIN*

McArdle Laboratory for Cancer Research, University of Wisconsin, 450 North Randall Ave., Madison, Wisconsin 53706

Received 23 December 1987/Accepted 6 March 1988

A retroviral vector system was developed to study the retrotransposition of RNAs lacking all *cis*-acting sequences required for normal retroviral replication. Our experiments indicate that such RNAs can be encapsidated in retroviral proteins, reverse transcribed, and integrated to form functional cDNA genes in infected cells. The frequency of this process, however, was approximately 8 orders of magnitude less than that of normal retroviral replication. The efficiency was limited at each step in this process. Investigation of seven cDNA genes by Southern blot analysis revealed that all of them were truncated at either the 3′ or the 5′ end or both. These truncations are not seen with natural cDNA genes and raise the question of retroviral involvement in their formation.

The discovery of reverse transcription as a key step in retroviral replication (1, 28) led to the proposal that the flow of genetic material into the genome via RNA intermediates plays a role in various biological processes and evolution (25, 26). Data supporting this hypothesis were obtained over the past decade by the molecular cloning and DNA sequencing of many eucaryotic genes. It appears that more than 10% of the mammalian genome consists of sequences that originated by retrotransposition of cellular RNAs (2, 4, 21, 27, 34). These types of sequences comprise retroviral and retroviruslike sequences (including retrotransposons), highly and middle-repetitive DNA elements (for example, long and short interspersed nucleus repeat sequences), and cDNA pseudogenes of protein-coding genes (4, 21, 29, 31, 34).

Reverse transcription of retroviral RNAs and probably of retrotransposons is mediated by proteins and *cis*-acting sequences encoded by these elements. Retroviral cDNA synthesis is primed by a tRNA which hybridizes to a specific tRNA primer binding site. A polypurine tract primes second-strand DNA synthesis, and a viral integrase and a viral attachment site guarantee high efficiency of integration (19, 30). As a consequence of integration, short direct repeats are generated in the chromosomal DNA at the site of insertion (17).

No specific *trans*- or *cis*-acting sequences for retrotransposition are known to be present in pseudogenes of normal protein-coding genes or many repetitive DNA elements. Thus, the mechanism of formation of these sequences is unknown, and the evidence for retrotransposition is only circumstantial. The hypothesis of retrotransposition for some of these sequences is based on the findings that the structure of some cDNA pseudogenes reflects the structure of mature (spliced) RNAs and contains poly(A) sequences. Moreover, many of these sequences are flanked by direct repeats apparently formed in the course of DNA integration (2, 21, 27, 29, 34). Usually, one of these direct repeats abuts the 5′ end of the mature RNA transcript sequence.

To investigate whether retroviral proteins could have participated in the retrotransposition of mRNAs, we constructed a retroviral vector system to study this process. We find that RNA without retroviral *cis*-acting sequences can be encapsidated in retroviral proteins, reverse transcribed, and integrated to form a functional cDNA gene. The frequency of this gene transfer, however, is about 8 orders of magnitude less than that of an RNA with all retroviral *cis*-acting sequences.

MATERIALS AND METHODS

Nomenclature. All vectors used in this study were derived from spleen necrosis virus (SNV), an avian reticuloendotheliosis virus. Plasmid constructions are indicated by the letter p (e.g., pRD17) to distinguish them from virus (e.g., RD17) derived from the plasmids. Hygromycin (HY) is the hygromycin B phosphotransferase gene isolated from pLg89 (7). Hygro^r (hygromycin resistant) refers to the phenotype.

Plasmid constructions. All plasmids (Fig. 1) were constructed by standard cloning procedures (14) and were derived from pJD220HY, pJD220SVHY (5), pTK1 (15, 32), pJE189, and pME142 (10, 23) in several sequential cloning steps. Cloning strategies and detailed restriction enzyme maps of these vectors are available upon request. JD214HY has all the *cis*-acting sequences required for normal retroviral replication. The viral protein-coding sequences are replaced by the hygromycin resistance gene (7), the expression of which is driven by the SNV long terminal repeat (LTR) promoter. In pJD220SVHY the U3 region of the right LTR (except for 10 base pairs at the 5′ end) was substituted by an *Xho*I linker (U3$^-$). Earlier studies have shown that U3 sequences are required for polyadenylation of the retroviral RNA (5). Thus, the polyadenylation sequence of the simian virus 40 (SV40) late genes (35) was introduced at the 3′ LTR of this vector.

pRD17, pRD18, and pRD19 are similar to pJD220SVHY. The hygromycin resistance gene, however, is placed in the opposite orientation to the retroviral sequences. Gene expression is controlled by the internal promoters as well as by an internal polyadenylation signal (derived from the herpes simplex virus type 1 thymidine kinase [*tk*] gene [15, 32]), the orientation of which is the same as that of the hygromycin resistance gene. pRD17 and pRD18 differ by the combination of the internal promoters. A bacterial suppressor tRNA gene (*supF*) (22) was inserted between the hygromycin resistance gene and the polyadenylation signal to facilitate cloning of integrated proviruses (20).

Cells. D17 C3A2 dog helper cells and D17 dog cells were grown as previously described (5). The helper cells (C3A2) were derived from D17 cells and supply all retroviral proteins essential for viral replication without production of replication-competent helper virus (33). Selection for Hygro^r

* Corresponding author.

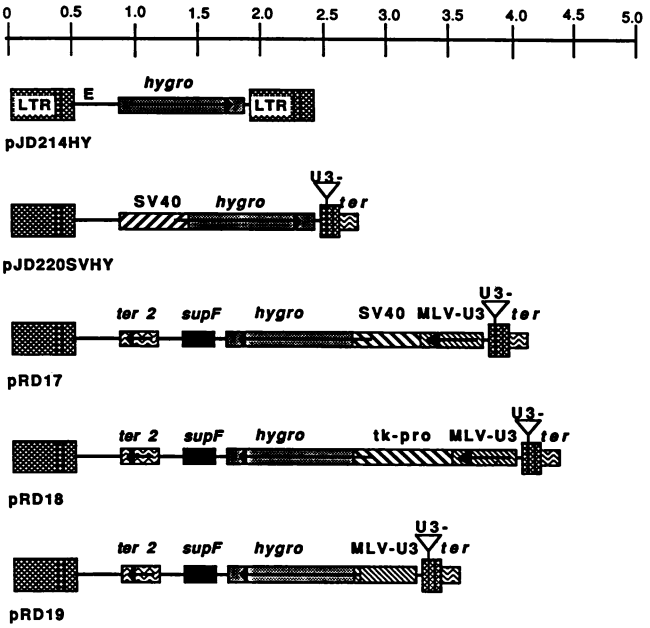

FIG. 1. Structure of retroviral vectors. All vectors are based on SNV and contain SNV LTR and encapsidation (E) sequences. Plasmid sequences are not shown. Arrows indicate the orientation of the genes. Abbreviations: hygro, hygromycin resistance gene; SV40, promoter of SV40 early genes; ter, polyadenylation signal of the SV40 late genes; U3$^-$, U3-deleted LTR of SNV; ter 2 and tk-pro, polyadenylation sequence and promoter of the thymidine kinase gene of herpes simplex virus type 1, respectively; MLV-U3, U3 promoter of Moloney murine leukemia virus LTR; *supF*, suppressor tRNA gene of *Escherichia coli*. The scale (in kilobase pairs) is given at the top.

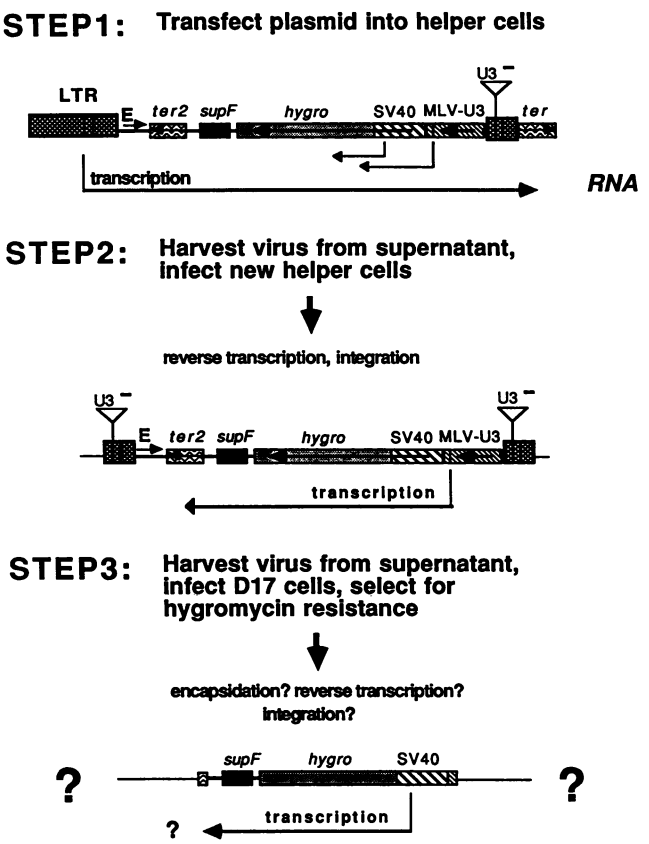

FIG. 2. Experimental protocol. A retrovirus vector (pRD17) is shown at the top (step 1). After transfection DNA of this structure is in step 1 cells. The structure of the DNA provirus formed from the transcript (starting at the left LTR) of this vector is shown in the middle (step 2). This provirus would not have LTR U3 sequences (5) and would only be transcribed by the internal promoters. If this transcript were encapsidated, reverse transcribed, and integrated (step 3), it could be expressed from the remaining internal promoter. Polyadenylation of this transcript will depend on the presence of a poly(A) signal in the chromosomal DNA. Abbreviations are as in the legend of Fig. 1.

cell colonies was done in the presence of hygromycin B at 80 μg/ml.

Transfections and infections. Cells were transfected by the polybrene-dimethyl sulfoxide method (11). Virus titers were determined as described previously (5) without freezing and thawing the virus solutions. To remove cells, the virus solutions were centrifuged for 10 min at 3,000 × *g*, and the upper half of each solution was carefully removed and used for infections. Step 3 infections were done with 0.5 ml of undiluted supernatant medium obtained from step 2 helper cells.

Southern blots. Experiments were performed by standard techniques (14, 24). All blots shown were hybridized with a purified hygromycin probe isolated from pJD220SVHY.

RNA isolation, Northern analysis, and RNA dot blots. Cellular RNAs were purified by standard techniques (14). For viral RNA isolation, tissue culture supernatant medium (obtained from confluent cultures) was centrifuged at 10,000 × *g* for 10 min to remove cells, and RNA was isolated from a pellet obtained by centrifugation at 100,000 × *g* (90 min) by phenol extraction and ethanol precipitation (14). Northern (RNA) and dot blots were performed by standard techniques (14).

RESULTS

Our experimental design is outlined in Fig. 2. Helper cells were transfected with vector DNAs, selected for Hygro^r, and passaged. These cells contain DNA like that in the plasmid. Virus was collected from confluent cultures. Fresh helper cells and D17 cells were infected, and virus titers were determined for each construct (Table 1, step 2), JD214HY gave the highest titer. The titer of JD220SVHY was reduced by a factor of approximately 20 compared with that of JD214HY, probably as a result of the difference of the promoter strengths and/or interference between the SNV and SV40 promoters (9). The titers of RD17 and RD19 were reduced by a factor of approximately 50 compared to that of JD220SVHY. This finding might reflect the presence of anti-sense RNAs (6) and/or disturbance of RNA transcription starting from the left LTR caused by the internal promoters working in the opposite direction. The titers of all viruses on D17 cells were 20- to 50-fold higher than those on helper cells as a result of lack of superinfection interference (12).

To determine the rate at which the hygromycin resistance gene could be transferred by an RNA lacking all retroviral *cis*-acting sequences, supernatant medium was harvested from confluent step 2 cultures containing about 5×10^2 step 2 helper cell (C3A2) clones (in the following referred to as mass infection), and fresh D17 cells were infected. Approximately 1 Hygro^r colony per ml of supernatant medium (average of 30 plates tested for each) was obtained with

TABLE 1. Virus titers at steps 2 and 3[a]

Virus	Virus titers		
	Step 2		Step 3, D17
	C3A2 cells	D17 cells	
JD214HY	5×10^5	2×10^7	5×10^7
JD220SVHY	3×10^4	1×10^6	ND
RD19	7×10^2	2×10^4	0[b]
RD17	7×10^2	2×10^4	1
RD18	2×10^2	8×10^3	1

[a] Helper cells (C3A2) and D17 cells were infected with viruses harvested from transfected C3A2 helper cells. Titers at step 2 (see text) (expressed as hygromycin resistance transforming units per milliliter of tissue culture supernatant medium) were determined as described previously (5). Virus titers of all step 3 RD viruses were determined by infecting 30 dishes of D17 cells with 0.5 ml of supernatant medium harvested from step 2 mass-infected confluent C3A2 helper cells. Colonies were counted 3 weeks after infection. ND, Not determined.

[b] <0.06.

RD17 and RD18. However, no colonies could be detected from RD19-infected cells, probably reflecting the lack of a step 3 internal promoter. The titers derived from JD214-infected helper cells were in the same range as those obtained from JD214HY-transfected helper cells (Table 1, step 3).

To study this gene transfer in detail, experiments were performed with single step 2 helper cell colonies. Helper cells were infected with virus from step 1 cells as described above and were selected for hygromycin resistance. Single step 2 clones were isolated and analyzed for the structure of the integrated provirus. During the construction of pJD-220SVHY, the deleted U3 sequences were replaced by an *Xho*I linker, which is duplicated in the left LTR during formation of the step 2 provirus (Fig. 2) (5). Thus, the gene structure of the promoterless LTR provirus formed from all U3$^-$ pRD vectors could be monitored by Southern blot analysis of *Xho*I-digested chromosomal DNAs. Most of the cell clones investigated contained the expected proviral structure (Fig. 3A and Table 2). Northern blot analysis of RNAs isolated from step 2 cells showed bands of the predicted sizes (Fig. 3B), confirming that RNA transcription from step 2 proviruses was initiated and terminated as expected.

Infectivity studies performed with supernatant media harvested from confluent cultures of such step 2 clones revealed that 16 of 22 RD17-infected clones and 2 of 3 RD18-infected clones transferred hygromycin resistance (Table 2). The total number of step 3 colonies derived from each clone ranged from one to six per 3 ml of supernatant medium (data not shown). The average frequency of the Hygro^r transfer was identical to that obtained from mass-infected cell cultures (compare Tables 1 and 2). One colony was obtained with supernatant medium from 1 of 10 RD19 step 2 clones. No Hygro^r transfer was observed when fresh D17 cells were infected with supernatant media taken from step 2 D17 cells (RD17, RD18, and RD19 mass infections and single-cell colonies) superinfected with a wild-type retrovirus (reticuloendotheliosis virus A; data not shown).

A common feature of all step 3 colonies was very slow colony growth. To quantify this behavior, we measured the growth rate of 6 step 2 clones and 16 step 3 clones starting at day 13 after infection (Fig. 4 and data not shown). At this time all cells had died in uninfected control cultures. Ten of the step 3 clones died within the third week after infection. The other six colonies grew very poorly compared with step

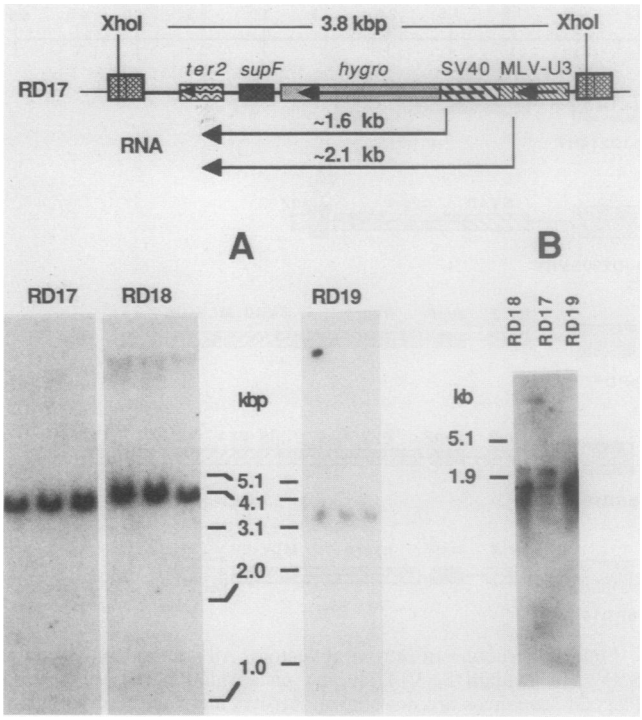

FIG. 3. Analysis of DNA and RNA from step 2 helper cell clones. The hypothetical structure of a step 2 provirus (RD17) and the RNAs transcribed from it are shown at the top. The sizes of the bands expected by *Xho*I digests of chromosomal DNAs from RD18 and RD19 proviruses are 4.1 and 3.3 kilobase pairs, respectively. The sizes of RNAs transcribed from RD18 and RD19 are 2.3 plus 1.6 kilobases and 1.6 kilobases, respectively. (A) Southern blot analysis of *Xho*I-digested chromosomal DNAs isolated from single step 2 colonies infected with RD17, RD18, and RD19. (B) Northern blot analysis of RNAs isolated from RD18-, RD17-, and RD19-infected step 2 cells. The filters were hybridized with a hygromycin resistance gene-specific probe.

2 clones. Reduction of the hygromycin concentration resulted in a large increase in cell growth in two of these six clones after another week. When the hygromycin concentration was increased, the four slow-growing colonies died, whereas growth of the other two colonies was not affected. The poor growth of the step 3 clones indicates poor expression. The low final survival indicates that the lack of retroviral *cis*-acting sequences hampers formation and/or integration of a functional gene.

To study the structure of the cDNAs formed in step 3 clones, 31 step 3 RD17 and RD18 clones were grown up after

TABLE 2. Transfer of the hygromycin gene from individual step 2 helper cell clones[a]

Virus	No. of clones with intact provirus[b]/total	Fraction of clones transferring Hygro^r	Avg titer[c] at step 3
RD19	5/5	1/10	0.04
RD17	16/23	16/22	1
RD18	8/9	2/3	1

[a] Step 2 helper cell clones were investigated for the presence of the predicted proviral structure. Supernatant medium was harvested from each individual clone, and five dishes of D17 cells were infected with 0.5 ml of supernatant medium.

[b] *Xho*I digestion gave the predicted size band.

[c] Titers are expressed as in Table 1.

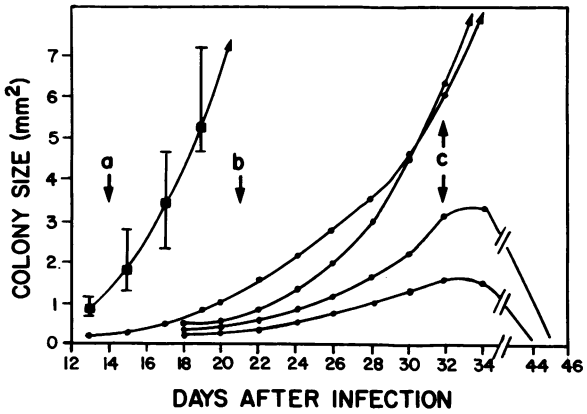

FIG. 4. Growth of step 2 and step 3 colonies. D17 cells were infected with virus harvested from transfected helper cells (■, average size of 6 colonies) and with virus harvested from infected step 2 helper cells (●). Arrows: a, all cells dead in control cultures; b, reduction of the hygromycin concentration to 20 μg/ml; c, increase of the hygromycin concentration to 60 μg/ml. The sizes of the colonies were measured with an ocular grid.

hygromycin selection was removed 3 weeks after infection. Chromosomal DNAs were isolated and analyzed by Southern blots. If full-length cDNA copies were generated, digestion of the chromosomal DNAs with *Bam*HI and *Bgl*II would generate defined bands (Fig. 5A). *Sac*I digestion should result in DNA fragments of various lengths, because no sequences recognized by this enzyme are located within the postulated RNA transcripts.

Southern blots revealed that six step 3 clones did not have the expected *Bam*HI-*Bgl*II band and therefore contained truncated cDNA genes missing some sequences from the 3′ and/or 5′ end (Fig. 5). One clone had two separate insertions of the hygromycin resistance gene (Fig. 5B, lane b). *Sac*I digestion of DNAs from these clones resulted in bands of different lengths as expected. Digests of DNAs from RD17-derived clones with *Eco*RI plus *Sph*I and from RD18-derived clones with *Eco*RI showed defined 500-base-pair bands in all digests (data not shown). Thus, 5′-end deletions did not include the SV40 or *tk* promoters which are located on these fragments (downstream of an *Sph*I or *Eco*RI site in RD17 or RD18, respectively; Fig. 6). Further restriction enzyme mapping by double digestions with *Hin*dIII-*Bam*HI, *Bgl*II-*Xba*I, and *Hin*dIII-*Xba*I revealed that six of the seven integrates contained a complete hygromycin resistance gene (defined *Hin*dIII-*Xba*I fragment; Fig. 6). Sequences 3′ of the hygromycin resistance gene were missing in the seventh integrate (Fig. 6, map b). Five integrates were truncated at the 3′ end (fragments resulting from a *Bgl*II-*Xba*I digest were bigger than expected; data not shown). Three integrates contained the *Bam*HI site located 35 bases downstream of the initiation of RNA transcription from the Moloney murine leukemia virus promoter (23). Thus, they could contain all sequences up to the 5′ end. Rehybridization of the filters used in these analyses with SNV LTR or *gag-pol*-specific probes confirmed that the Hygro^r transfer was mediated without retroviral sequences (data not shown).

(Four other clones showed the *Bam*HI-*Bgl*II pattern as predicted for a full-length integrated retrotranscript [data not shown]. Moreover, the *Sac*I digest of DNA from these clones resulted in bands corresponding to the length of a *Sac*I-digested step 2 provirus. Further investigations of these clones indicated that two of them contained SNV

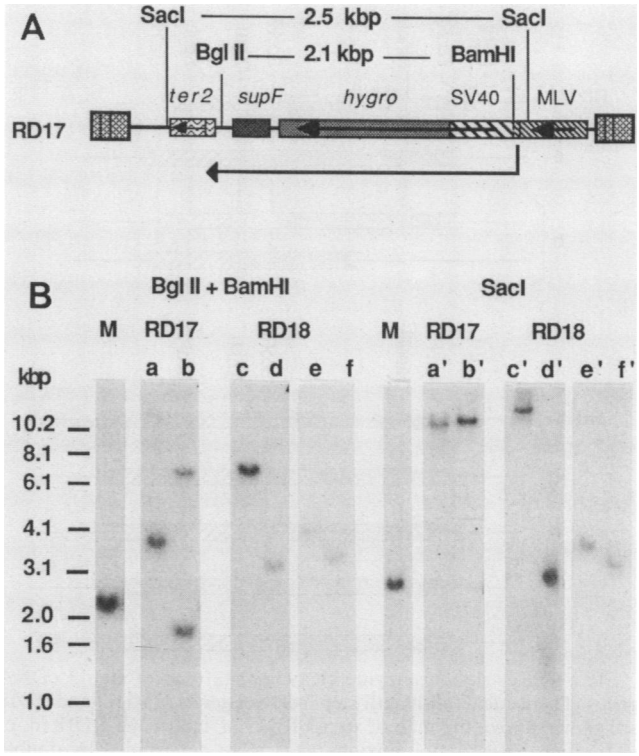

FIG. 5. Structure of step 3 cDNA genes. (A) Restriction enzyme cleavage site map of a step 2 provirus (RD17). The arrow below indicates the RNA containing the internal promoter transcribed from this integrate. A full-length step 3 retrotranscript generated from this RNA would contain the *Bgl*II and *Bam*HI cleavage sites. The sizes of the bands obtained from a step 2 RD18 provirus were 2.4 (*Bgl*II-*Bam*HI) and 2.8 (*Sac*I) kilobase pairs. (B) Southern blot of *Bam*HI-*Bgl*II- and *Sac*I-digested DNAs extracted from different step 3 colonies. Lanes: M, control DNA of a parental step 2 RD18 provirus digested with the same enzymes; a through f and a′ through f′, DNAs from different clones digested with *Bgl*II-*Bam*HI and *Sac*I, respectively. The *Sac*I digest of RD17 (lane b) contained two large bands with approximately the same mobility. The filters were hybridized with a hygromycin resistance gene-specific probe.

sequences and, thus, were derived from recombinant retroviruses. The other two clones [including the only clone derived from RD19] probably originated from unexpected transcripts driven by cellular promoters outside of the step 2 provirus [data not shown]. Such transcripts might contain some retroviral *cis*-acting sequences [encapsidation, primer-binding site].)

No hygromycin resistance gene was detected in the other 21 step 3 clones. When these clones were put back under hygromycin (60 μg/ml) selection, all cells died within 12 days (data not shown). This finding further indicates that integration of the hygromycin resistance gene retrotranscript was inefficient, in agreement with the growth rate of step 3 clones (Fig. 4).

To study the efficiency of the other steps involved in this transfer of nonretroviral mRNAs, RNA transcription and encapsidation were measured. RNAs were extracted from step 2 cells and from particles harvested from supernatant media of these cells and were subjected to RNA dot-blot analysis (Fig. 7). The results of this experiment showed that the SNV LTR promoter is about 1 order of magnitude stronger in helper cells than the SV40, *tk*, and Moloney murine leukemia virus U3 promoters as estimated by the

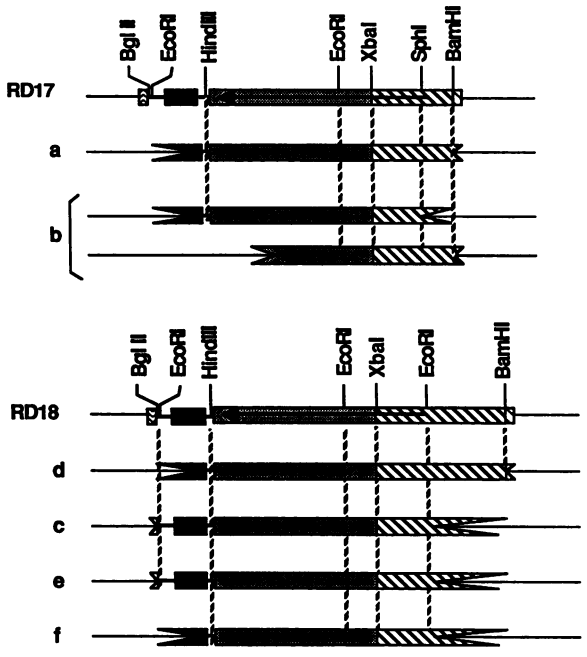

FIG. 6. Restriction enzyme cleavage site map of step 3 cDNA genes. Theoretical full-length step 3 cDNA genes of RD17 and RD18 are shown above the map of step 3 RD17 (a and b) and RD18 (d, c, e, f) clones. Restriction enzyme sites were determined by double digests of chromosomal DNAs and Southern analysis (see text). Dashed lines represent restriction enzyme sites present in the cDNA genes as shown at the top. Indented rectangles indicate that cDNA sequences end in these regions.

intensity of the RNA dots (8). Encapsidation of the internal hygromycin resistance gene transcripts of RD17, RD18, and RD19 was reduced by a factor of at least 10^4. Encapsidation of RD19 RNA was more efficient than that of RD17 and RD18. This finding agrees with earlier investigations that showed that efficiency of encapsidation is increased with decreasing length of RNA (J. E. Embretson, Ph.D. thesis, University of Wisconsin, Madison, 1986).

DISCUSSION

A retroviral vector system was developed to study the formation of cDNA genes (Fig. 1 and 2). A gene unit (promoter[s], hygromycin B phosphotransferase gene, 3′ RNA-processing site) was introduced into helper cell genomes with a retroviral vector that loses its U3 LTR promoter after one round of replication (Fig. 2, step 2) (5). Thus, RNA transcription is driven only by the internal promoters, and RNA transcripts stop at the internal 3′ RNA processing site. Thus, these RNAs do not contain the *cis*-acting sequences required for normal retroviral replication. Following the experimental design outlined in Fig. 2, infectivity studies with media harvested from step 2 mass infections and from individual step 2 clones showed that Hygro^r could be transferred to D17 cells by the medium from infected helper cells (Tables 1 and 2). No Hygro^r transfer could be obtained with supernatant media harvested from step 2 D17 cells which lacked retroviral helper functions or which expressed a homologous replication-competent retrovirus (reticuloendotheliosis virus A) (data not shown). This data establishes that the Hygro^r transfer is dependent on retroviral proteins and the absence of homologous retroviral RNA competing for encapsidation.

The efficiency of the Hygro^r transfer by RD17 and RD18 was compared with that of a retroviral vector (JD214HY) containing all retroviral *cis*-acting sequences required for retroviral replication. Studies with step 2 mass infections and individual cell clones showed that this nonretroviral RNA transfer was more than 7 orders of magnitude less efficient than that of normal retroviral replication (Tables 1 and 2). Moreover, formation of stable integrated functional cDNA genes occurred only in about 30% of Hygro^r clones (see below and Fig. 4). Decreased levels of RNA transcription and encapsidation reduced the efficiency of retrotransposition by 1 and 4 orders of magnitude, respectively (Fig. 7). With these data it may be estimated that successful reverse transcription of a nonviral RNA into a functional cDNA gene was reduced by a factor of 10^{-2}. This estimation is drawn from the finding that the overall formation of a stable integrated functional cDNA gene was reduced by a factor of 10^{-8}, and that RNA transcription, encapsidation, and integration limited this process by factors of about 10^{-1}, 10^{-4}, and 10^{-1}, respectively. However, since we selected for functional cDNA genes, the frequency determined is underestimated. For example, we have previously shown that the absence of a functional poly(A) sequence reduces the efficiency of transfer by a factor of 5 (5).

Step 2 and 3 clones were analyzed for the structure of integrated hygromycin resistance genes. Southern blot analysis of chromosomal DNAs purified from isolated step 2 clones established that a promoterless LTR proviral structure was formed as predicted, and RNAs were transcribed from the internal promoters as expected (Fig. 3).

Investigation of 31 RD17 and RD18 step 3 clones indicated that 21 separate clones did not contain an integrated copy of the hygromycin resistance gene. Since they had survived more than 3 weeks of hygromycin selection (Fig. 4), we hypothesize that in these clones step 2 transcripts were reverse transcribed and transiently expressed without integration. We hypothesize that the product from a single

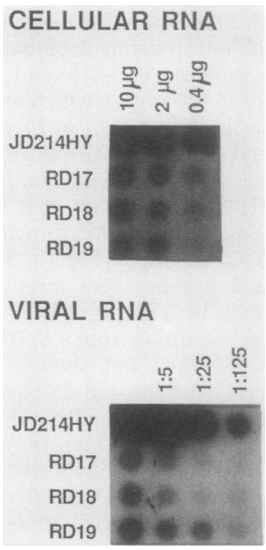

FIG. 7. Dot-blot analysis of RNAs from step 2 cells. RNAs were isolated from step 2 helper cells and from viruses harvested from these cells. The viral RNAs applied to the filter correspond to the total amount isolated from 4 ml (JD214HY) and 400 ml (RD viruses) of tissue culture supernatant medium. The blots were hybridized with a hygromycin B phosphotransferase gene-specific probe. For more explanations, see the text.

unintegrated cDNA gene was distributed to progeny cells, giving rise to small colonies. The extrachromosomal copies of these retrotranscripts were lost during further cell divisions in the absence of selection. This hypothesis is supported by the finding that *att*$^-$ SNV proviruses produce progeny virus in the absence of integration (18). SNV sequences were detected in two other step 3 clones, indicating that a recombinant retrovirus had been formed, although the parental step 2 cells contained a provirus as predicted (Fig. 3A; RD18, right lane). Two step 3 clones (including the clone derived from RD19) contained unexpected integrates, probably a result of RNA transcription in step 2 cells from a promoter outside the provirus.

In six clones cDNA genes with the predicted structure were found. The structure of these genes strongly suggests that they were derived from transcripts starting at the MLV promoters of RD17 and RD18 (Fig. 5 and 6). This conclusion is drawn from the finding that no step 2 proviral sequences outside of these transcription units were detectable in these step 3 clones by restriction enzyme digestions. Hybridizations with SNV LTR- or *gag-pol*-specific probes established that the Hygro^r transfer in these clones was not mediated by a recombinant retrovirus.

While this work was in progress, it was shown that a cell line containing an encapsidation-defective avian sarcoma provirus retrotransposed an introduced neomycin mRNA. The formation of retrotransposed neomycin-resistant cell clones was the result of abnormal initiations of RNA transcription in the parental helper cell line, resulting in transcripts containing an internal (the original) promoter (13). In our studies, the formation of Hygro^r cell clones was shown to be dependent on the presence of an internal promoter (SV40 or *tk*) driving hygromycin resistance gene expression. This conclusion is based on Southern blot analysis (Fig. 6) and the finding that RD19 (which has no second internal promoter)-infected step 2 cells could not transfer Hygro^r, although RNAs transcribed from this provirus were encapsidated with higher efficiency than RD17 and RD18 RNAs (Fig. 7).

In addition, our data show that the predicted cDNA genes investigated were truncated at either the 3′ or 5′ end or both. Five integrates did not contain the *Bgl*II site, which is located about 140 bases upstream of the start of the poly(A) tail of the step 2 RNA. 3′-end truncation even included hygromycin-resistance-coding sequences in one integrate (see below). Thus, these cDNA genes do not contain poly(A) sequences like many natural cDNA pseudogenes. Four of seven integrates did not contain the *Bam*HI site located 35 base pairs downstream of the 5′ end of step 2 RNAs. The truncations could reflect the lack of specific primers for the synthesis of the first and second cDNA strands. Alternatively, the truncations could reflect the lack of terminal *att* sequences, resulting in unspecific integrations. In our experiments, we selected for the formation of functional cDNA genes containing an active promoter, a full-length hygromycin resistance gene, and a polyadenylation signal on the chromosomal DNA adjacent to the integrate. Thus, formation of nonfunctional (even more truncated) cDNA genes almost certainly occurred much more frequently. However, such integrates were not detected, except in one clone which contained an additional functional cDNA gene.

In summary, our data show that retroviral proteins can contribute to the genesis of cDNA genes. However, the finding that none of the cDNA genes represented a full-length copy of the original RNA raises the question whether the formation of cDNA pseudogenes [with poly(A) sequences] and analogous repetitive DNA elements was mediated by retroviruses. Cellular enzymes (derived and modified from ancient retroviruses or progenitors of retroviruses) might be involved (3, 16). Molecular cloning and DNA sequencing of the integration sites of the cDNA genes described here will give more insight into this process.

ACKNOWLEDGMENTS

We thank S. Coe, S. Hinz, J. Couch, and J. Rein for technical assistance; J. Coffin for a gift of pAS7 and *supF*-selectable bacteria and phage; J. Dougherty, T. Gilmore, and C. Gelinas for many useful discussions and suggestions during the course of this work; and M. Hannink, W.-S. Hu, A. Panganiban, and B. Sudgen for helpful comments on the manuscript.

This research was supported by Public Health Service grants CA-22443 and CA-07175 from the National Cancer Institute. R.D. was supported by a fellowship from the Deutsche Forschungsgemeinschaft. H.M.T. is an American Cancer Society Research Professor.

LITERATURE CITED

1. **Baltimore, D.** 1970. RNA-dependent DNA polymerase in virions of RNA tumor viruses. Nature (London) **226**:1209–1211.
2. **Baltimore, D.** 1985. Retroviruses and retrotransposons: the role of reverse transcriptase in shaping the eukaryotic genome. Cell **40**:481–482.
3. **Bauer, H. G., and P. H. Hofschneider.** 1976. An RNA-dependent DNA polymerase, different from the known viral reverse transcriptases, in the chicken system. Proc. Natl. Acad. Sci. USA **73**:3025–3029.
4. **Deininger, P. L., and G. R. Daniels.** 1986. The recent evolution of mammalian repetitive DNA elements. Trends Genet. **3**: 76–80.
5. **Dougherty, J. P., and H. M. Temin.** 1987. A promoterless retroviral vector indicates that there are sequences in U3 required for 3′ RNA processing. Proc. Natl. Acad. Sci. USA **84**:1197–1201.
6. **Green, P. J., O. Pines, and M. Inouye.** 1986. The role of antisense RNA in gene regulation. Annu. Rev. Biochem. **55**: 569–597.
7. **Gritz, L., and J. Davies.** 1983. Plasmid encoded hygromycin B resistance: the sequence of hygromycin B phosphotransferase gene and its expression in *Escherichia coli* and *Saccharomyces cerevisiae*. Gene **25**:179–188.
8. **Embretson, J. E., and H. M. Temin.** 1986. Pseudotyped retroviral vectors reveal restrictions to reticuloendotheliosis virus replication in rat cells. J. Virol. **60**:662–668.
9. **Emerman, M., and H. M. Temin.** 1984. Genes with promoters in retrovirus vectors can be independently suppressed by an epigenetic mechanism. Cell **39**:459–467.
10. **Emerman, M., and H. M. Temin.** 1986. Comparison of promoter suppression in avian and murine retrovirus vectors. Nucleic Acids Res. **14**:9381–9396.
11. **Kawai, S., and M. Nishizawa.** 1984. New procedure for DNA transfection with polycation and dimethyl sulfoxide. Mol. Cell. Biol. **4**:1172–1174.
12. **Kazushige, K., S. Hidetoshi, and T. Odaka.** 1986. Relationship between the cellular resistance to Friend murine leukemia virus infection and the expression of murine leukemia virus-gp70-related glycoprotein on cell surface of BALB/c-*Fv-4w*^{r} mice. Virology **150**:509–512.
13. **Linial, M.** 1987. Creation of a processed pseudogene by retroviral infection. Cell **49**:93–102.
14. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
15. **McKnight, S. L.** 1980. The nucleotide sequence and transcript map of the herpes simplex virus thymidine kinase gene. Nucleic Acids Res. **8**:5949–5964.
16. **Mondal, H., and P. H. Hofschneider.** 1983. Demonstration of free reverse transcriptase in the nuclei of embryonic tissues of

the Japanese quail. Biochem. Biophys. Res. Commun. **116:** 303–311.
17. **Panganiban, A. T.** 1985. Retroviral DNA integration. Cell **42:** 5–6.
18. **Panganiban, A. T., and H. M. Temin.** 1983. The terminal nucleotides of retrovirus DNA are required for integration but not for virus production. Nature (London) **306:**155–160.
19. **Panganiban, A. T., and H. M. Temin.** 1984. The retrovirus *pol* gene encodes a product required for DNA integration: identification of a retrovirus *int* locus. Proc. Natl. Acad. Sci. USA **81:**7885–7889.
20. **Reik, W., H. Weiher, and R. Jaenisch.** 1985. Replication-competent Moloney murine leukemia virus carrying a bacterial suppressor tRNA gene: selective cloning of proviral and flanking host sequences. Proc. Natl. Acad. Sci. USA **82:**1141–1145.
21. **Rogers, J. H.** 1985. The origin and evolution of retroposons. Int. Rev. Cytol. **93:**187–279.
22. **Seed, B.** 1983. Purification of genomic sequences from bacteriophage libraries by recombination and selection *in vivo*. Nucleic Acids Res. **11:**2427–2445.
23. **Shinnick, T. M., R. A. Lerner, and J. G. Sutcliffe.** 1981. Nucleotide sequence of Moloney murine leukemia virus. Nature (London) **293:**543–548.
24. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98:**503–517.
25. **Temin, H. M.** 1971. The protovirus hypothesis. J. Natl. Cancer Inst. **46:**III–VIII.
26. **Temin, H. M.** 1980. Origin of retroviruses from cellular movable genetic elements. Cell **21:**599–600.
27. **Temin, H. M.** 1985. Reverse transcription in the eukaryotic genome: retroviruses, pararetroviruses, retrotransposons, and retrotranscripts. Mol. Biol. Evol. **2:**455–468.
28. **Temin, H. M., and S. Mitzutani.** 1970. RNA-dependent DNA polymerase in virions of Rous sarcoma virus. Nature (London) **226:**1211–1213.
29. **Vanin, E. F.** 1985. Processed pseudogenes. Characteristics and evolution. Annu. Rev. Genet. **19:**253–272.
30. **Varmus, H. E.** 1982. Form and function of retroviral proviruses. Science **216:**812–820.
31. **Wagner, M.** 1986. A consideration of the origin of processed pseudogenes. Trends Genet. **5:**134–137.
32. **Wagner, M. J., J. A. Sharp, and W. C. Summers.** 1981. Nucleotide sequence of the thymidine kinase gene of herpes simplex virus type 1. Proc. Natl. Acad. Sci. USA **78:**1441–1445.
33. **Watanabe, S., and H. M. Temin.** 1983. Construction of a helper cell line for avian reticuloendotheliosis virus cloning vectors. Mol. Cell. Biol. **3:**2241–2249.
34. **Weiner, A. M., P. L. Deininger, and A. Efstratiadis.** 1986. Nonviral retroposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. Annu. Rev. Biochem. **55:**631–661.
35. **Wickens, M., and P. Stephenson.** 1984. Role of the conserved AAUAAA sequence: four AAUAAA point mutants prevent messenger RNA 3′ end formation. Science **226:**1045–1051.