more reactive than duplex DNA. Possibly the ~12-bp hybrid does exist but has a structure sufficiently different from extended hybrids that its DNA component is more reactive than a strand of DNA in a duplex. Alternatively, the ~12-bp hybrid might be in equilibrium with free RNA and DNA and thus somewhat open to both nuclease and $KMnO_4$ attack.

REFERENCES AND NOTES

1. A. J. Carpousis and J. D. Gralla, *J. Mol. Biol.* **183**, 165 (1985).
2. W. Metzger, P. Shickor, H. Heumann, *EMBO J.* **8**, 2745 (1989); B. Krummel and M. J. Chamberlin, *Biochemistry* **28**, 7829 (1989).
3. H. Gamper and J. Hearst, *Cell* **29**, 81 (1982).
4. M. Hanna and C. F. Meares, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 4238 (1983).
5. T. Yager and P. H. von Hippel, in Escherichia coli *and* Salmonella typhimurium *Cellular and Molecular Biology*, F. C. Neidhardt, Ed. (American Society for Microbiology, Washington, DC, 1987), pp. 1241–1275.
6. L. A. Wagner, R. B. Weiss, R. Driscoll, D. S. Dunn, R. F. Gesteland, *Nucleic Acids Res.* **18**, 3529 (1990).
7. G. A. Rice, C. M. Kane, M. J. Chamberlin, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 4245 (1991).
8. E. J. Grayhack, X. Yang, L. F. Lau, J. W. Roberts, *Cell* **42**, 259 (1985).
9. X. Yang, J. A. Goliger, J. W. Roberts, *J. Mol. Biol.* **210**, 453 (1989).
10. H. C. Guo, M. Kainz, J. W. Roberts, *J. Bacteriol.* **173**, 1554 (1991).
11. S. Sasse-Dwight and J. D. Gralla, *J. Biol. Chem.* **264**, 8074 (1989).
12. X. Yang, C. M. Hart, E. J. Grayhack, J. W. Roberts, *Genes Dev.* **1**, 217 (1987).
13. X. Yang, thesis, Cornell University (1988).
14. Point mutations are named according to position relative to the RNA start site and to the sequence change in the top (RNA-like) strand of DNA.
15. The paused transcription complex of λ DNA characterized in vitro contains RNAs at both +16 and +17 (W. Yarnell and J. Roberts, unpublished results), although the RNA was originally described as "+16" (*8*).
16. Although adenine adjacent to reactive thymine may be identified by primer extension (*11*), we have noted only pyrimidines in Fig. 3.
17. M. Kainz and J. W. Roberts, unpublished data.
18. In vitro transcription reactions were similar to those described (*8*). Open complexes were formed by incubating reactions containing 5 nM supercoiled plasmid, 25 nM *E. coli* RNAP (purified in this laboratory by G. DiCenzo), and 150 nM NusA protein (purified in this laboratory by J. Goliger) in transcription buffer (20 mM tris-HCl, pH 7.5, 0.1 mM EDTA, 1 mM dithiothreitol (DTT), 10 mM KCl, and 4 mM $MgCl_2$), with or without rifampicin at 10 μg/ml, at 37° for 10 min. Transcription was begun by addition of 200 μM adenosine triphosphate (ATP), cytidine triphosphate (CTP), and guanosine triphosphate (GTP) and 50 μM uridine triphosphate (UTP), and incubation was continued at 37°C for 5 min.
19. J. A. Boroweic, L. Zhang, S. Sasse-Dwight, J. D. Gralla, *J. Mol. Biol.* **196**, 101 (1987).
20. Note that there is more signal in lanes 1 and 2 of Fig. 1B than in other lanes; lanes 4 through 7 better display the strong overall similarity of reactivity in vivo and in vitro.
21. Y. B. Shi, H. Gamper, B. Van Houten, J. E. Hearst, *J. Mol. Biol.* **199**, 277 (1988).
22. K. Kirkegaard, H. Buc, A. Spassky, J. C. Wang, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 2544 (1983); U. Siebenlist, R. B. Simpson, W. Gilbert, *Cell* **20**, 269 (1980).
23. S. Barak, B. Ghosh, W. Whalen, D. Lazinski, A. Das, *Cell* **50**, 885 (1987); R. J. Horwitz, J. Li, J. Greenblatt, *ibid.* **51**, 631 (1987).
24. $KMnO_4$ reactivity in vivo was analyzed essentially as described (*11*). *Escherichia coli* HB101 transformed with plasmid was grown to mid-log phase in M9 glucose supplemented with 0.5% casamino acids and 100 μg of ampicillin per milliliter. Culture aliquots were treated at 37°C with 10 mM $KMnO_4$ for 2 min or with 200 μg of rifampicin per milliliter for 5 min followed by 10 mM $KMnO_4$ for 2 min. Plasmid DNA was extracted by a boiling lysis method as described (*25*), except that 0.06 mg of lysozyme per milliliter was added, and then purified [as described (*11*)]. Plasmid yields were quantified with a slot blot hybridization assay in which aliquots of purified $KMnO_4$-modified plasmid were transferred to nylon hybridization membrane (GeneScreen Plus, New England Nuclear, Boston, MA) and hybridized with ^{32}P-labeled oligonucleotides used as primers in the primer extension assay. The hybridization signal was quantified by densitometry of the autoradiogram of the membrane or by direct counting of the membrane using a Betascope blot analyzer (Betagen). Equal amounts of plasmid were analyzed by primer extension as described (*19*).
25. G. Del Sol, G. Manfioletti, C. Schneider, *Biotechniques* **7**, 514 (1989).
26. Because only pyrimidines react significantly with $KMnO_4$, some boundaries of melted segments are not determined exactly. For open complexes, uncertainties in upstream and downstream boundaries are, respectively: λ, 1 and 2 bp; 82, 1 and 0 bp; and 21, 1 and 3 bp. For paused elongation complexes: the λ +16–+17 complex, 2 and 1 bp; the 82 +15 complex, 1 and 1 bp; the 82 +25 complex, 0 and 0 bp; and the 21 +18 complex, 1 and 0 bp.

27. We thank W. Yarnell and H.-C. Guo for advice and discussion, G. DiCenzo and J. Goliger for materials, J. Gralla for assistance with the $KMnO_4$ footprinting procedure, and J. Lis, J. Helmann, and members of the laboratory for their critique of the manuscript. Oligonucleotides were synthesized by the Cornell University oligonucleotide synthesis facility. This research was supported by NIH grant GM 21941 and a grant from the Cornell Biotechnology Program.

27 August 1991; accepted 3 December 1991

Mechanism of Transduction by Retroviruses

AMANDA SWAIN AND JOHN M. COFFIN*

Retroviruses can capture cellular sequences and express them as oncogenes. Capture has been proposed to be a consequence of the inefficiency of polyadenylation of the viral genome that allows the packaging of cellular sequences flanking the integrated provirus in virions; after transfer into virions, these sequences could be incorporated into the viral genome by illegitimate recombination during reverse transcription. As a test for this hypothesis, a tissue culture system was developed that mimics the transduction process and allows the analysis and quantitation of capture events in a single step. In this model, transduction of sequences adjacent to a provirus depends on the formation of readthrough transcripts and their transmission in virions and leads to various recombinant structures whose formation is independent of sequence similarity at the crossover site. Thus, all events in the transduction process can be attributed to the action of reverse transcriptase on readthrough transcripts without involving deletions of cellular DNA.

RETROVIRUSES HAVE BEEN STUDIED for their ability to acquire and express cellular sequences as oncogenes (*1, 2*). However, the mechanism by which these viruses incorporate cellular genes into their genome is still undefined. Transduction of oncogenes is thought to proceed from a provirus integrated within or near a proto-oncogene (*3–5*). Joining of cell and viral sequences can proceed by deletions that remove portions of viral and cellular sequences (*4*) or by formation of joint "readthrough" transcripts (*5*). Illegitimate recombination of the packaged hybrid transcripts during reverse transcription will incorporate proto-oncogene sequences into the provirus (Fig. 1).

That readthrough transcripts could be intermediates in transduction is implied by the observation that cleavage and polyadenylation of viral transcripts in the 3′ long terminal repeat (LTR) of avian leukosis virus is inefficient (*6*). A significant portion (about 15%) of viral transcripts from infected cells contain neighboring cell sequences, are packaged efficiently, and can serve as templates for reverse transcription. Indeed, a mutation inactivating the polyadenylation signal of Rous-associated virus (RAV-1) does not greatly affect viral replication. Most such mutant viral genomes have cellular sequences appended to their 3′ ends, but these sequences are usually lost during reverse transcription (*7*).

To test the role of readthrough transcripts in the transduction of cellular sequences, we constructed a model resembling a provirus integrated upstream of a proto-oncogene and introduced it into cells in culture. The SV-*neo* cassette (*8*), containing the selectable *neo* marker, flanked by DNA containing promoter and polyadenylation signals from SV40, was placed downstream of an RAV-1 provirus with a mutation in the polyadenylation signal (*7*) (Fig. 2A). QT35 cells [a transformed quail cell line (*9*)] were transfected with this DNA, and *neo*-expressing

Tufts University School of Medicine, Department of Molecular Biology and Microbiology, 136 Harrison Avenue, Boston, MA 02111.

*To whom correspondence should be addressed.

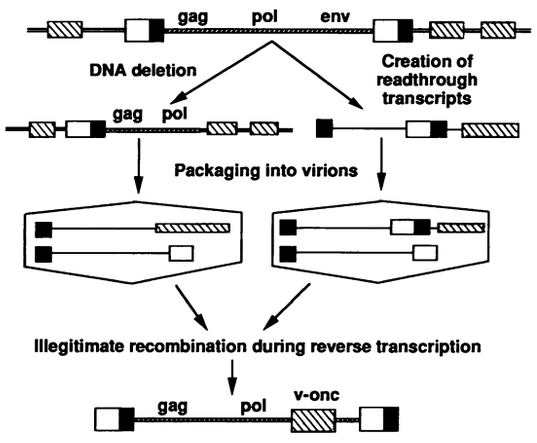

Fig. 1. Models for transduction of oncogenes, showing postulated events in the creation of viral oncogenes (*3–5*). The necessary steps are: (i) Integration of a provirus within a proto-oncogene, shown as a series of exons (hatched boxes) and introns (lines). (ii) The models differ at this step. A DNA deletion that would eliminate the 3′ LTR (*4*) would give rise to transcripts containing both viral (black line) and cellular (hatched box to indicate that the introns have been spliced out) sequences. Inefficient polyadenylation would give rise to readthrough transcripts containing both viral and cellular sequences (*5*). (iii) The hybrid transcripts would then be packaged into virions along with wild-type RNA's. (iv) Illegitimate recombination during reverse transcription would give rise to proviruses that contained cellular sequences. (v) Additional rearrangements and point mutations during several rounds of replication would enhance the transforming phenotype of the transduced oncogene.

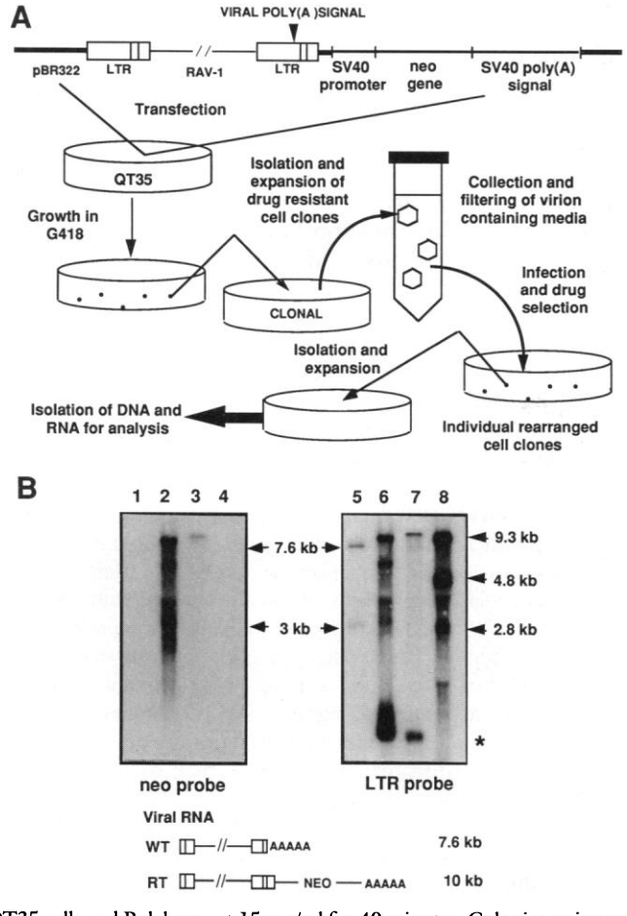

Fig. 2. Experimental approach. (**A**) The DNA was derived from a molecular clone of RAV-1 (*19*) with either the wild-type (AAUAAA) or mutant (AAGGAA) polyadenylation signal in the LTRs (7) and from pSV2neo (*8*). The Acc I to Eco RI fragment of pSV2neo was introduced at the Bst EII site downstream of the LTR (position 103 in Pr-RSV-C sequence) (*20*). The fragment containing the provirus and pSV2neo sequences was introduced into pBR322 at the Eco RV and Pvu II sites. This DNA was linearized at the only Cla I site in the pBR322 sequence before transfection. The QT35 cells (*9*) were transfected by incubation for 6 hours with 1 to 3 mg of DNA and Polybrene at 30 mg/ml, and then by treatment with DMSO (*21*). After 2 days, G418 was added, at 300 mg/ml, to the medium and colonies were picked 2 weeks later. Cell clones were expanded, supernatants were collected as a source of virion particles, and cell DNA was prepared (7). To remove remaining cells, the supernatants were either filtered or centrifuged and then incubated with fresh QT35 cells and Polybrene at 15 mg/ml for 40 minutes. Colonies resistant to G418 were again selected and used as a source of DNA for analyses of rearranged viral structures. (**B**) RNA was isolated, subjected to electrophoresis in formaldehyde agarose gels, and analyzed by hybridization (Northern) (7). The RNA's analyzed were poly(A)-containing RNA from the M7 cell clone (lanes 2 and 6), virion RNA from virions produced by M7 (lanes 3 and 7), and total cellular RNA from a nonclonal population of turkey cells infected with wild-type RAV-1 virus (lanes 1 and 5) and with the Prague C strain of Rous sarcoma virus (lanes 4 and 8). The RNA-containing filter was first hybridized to an antisense ^{32}P-labeled RNA probe containing the *neo* gene (Hind III to Sma I fragment of pSV2neo) (*neo* probe). The probe was removed at 80°C (1 hour), and the filter was then hybridized to an RNA probe containing the LTR and 5′ viral sequences up to the Sac I site (nt 255) (LTR probe). A representation of the structures of viral wild-type (WT) and readthrough (RT) transcripts is shown. The small transcript indicated by a star was determined from the sequence of a PCR product obtained by "rapid amplification of cDNA ends" (*22*). With the use of a hybrid dT-adapter primer, cDNA was made from virions and amplified by an LTR and an adapter primer.

clones were selected for their resistance to the antibiotic G418. Analysis of DNA from the cell clone (M7) chosen for study showed that one or two unrearranged copies of the structure introduced were present. RNA analysis showed that readthrough RNA was made as a 10-kb transcript and packaged into virions (Fig. 2B, lanes 3 and 7). As was expected for this mutant virus (7), transcripts that were polyadenylated in the LTR (a 7.6-kb RNA) could not be detected in the infected cells or in the virions. Several other *neo*-containing RNA's (probably splicing variants) were observed in the cells but were not packaged (lane 2). In addition to these, we also observed a small transcript that did not contain *neo* sequences and was packaged efficiently (lane 7) but seemed unimportant to the transduction process (*10*).

To test whether readthrough transcripts could be rearranged to incorporate flanking sequences during reverse transcription, filtered medium containing virions produced by the M7 cell clone was used to infect QT35 cells, and *neo*-expressing colonies were selected. Two weeks later, 10 to 80 resistant clones per plate were observed. Some of these clones were isolated, and cells were grown to serve as a source of DNA. Analysis of the DNA after it was digested with appropriate restriction enzymes showed that each clone yielded a distinct fragment with both *neo* and viral sequences and thus contained a single and different rearranged provirus. Thus, we were able to select for the rare cases where the *neo* sequences were incorporated into the provirus in an organization appropriate for their expression.

Since initial analyses of the proviral DNA implied that a rearrangement of viral and *neo* sequences had occurred in each clone, we used a strategy involving PCR (polymerase chain reaction) amplification to isolate and sequence junctions between viral and *neo* sequences from the DNA of several of these clones (Fig. 3). The data revealed several properties of the rearranged sequences.

1) The structures observed contained *neo*- and SV40-related sequences flanked by different portions of viral sequences, resembling oncogene-containing viruses (Fig. 4A). In two cases, clones 18 and 42, a long stretch of A's was found at the junction between parental sequences (Fig. 3B). As is the case for some *fps* (*11*) and *erb B* (*12*) oncogene transductions, recombination occurred in the polyadenylate [poly(A)] tail of the transduced *neo* gene and in different regions of the virus for both clones. As with oncogenes, an intron present in the SV40 sequences was spliced out from the two rearranged clones that were analyzed (clones 32 and 42).

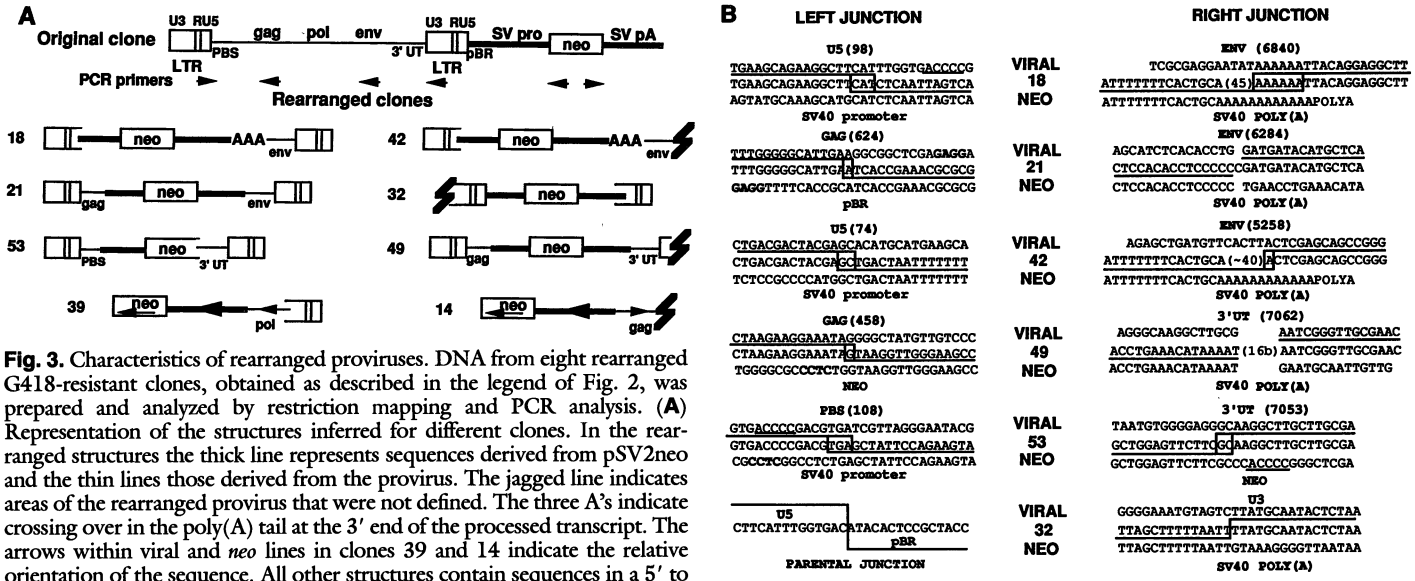

Fig. 3. Characteristics of rearranged proviruses. DNA from eight rearranged G418-resistant clones, obtained as described in the legend of Fig. 2, was prepared and analyzed by restriction mapping and PCR analysis. (**A**) Representation of the structures inferred for different clones. In the rearranged structures the thick line represents sequences derived from pSV2neo and the thin lines those derived from the provirus. The jagged line indicates areas of the rearranged provirus that were not defined. The three A's indicate crossing over in the poly(A) tail at the 3′ end of the processed transcript. The arrows within viral and *neo* lines in clones 39 and 14 indicate the relative orientation of the sequence. All other structures contain sequences in a 5′ to 3′ orientation for both virus and *neo*. (**B**) Sequences at the junction between the virus and *neo* at the 5′ end (left) and the 3′ end (right) of the *neo* sequences. These sequences were obtained by dideoxy sequencing from M13 clones of the fragments amplified by PCR with both *neo*- and viral-specific primers shown by the small arrows in (A). For all these sequences the appropriate viral sequences are shown above and the *neo*-related sequences are below the rearranged sequences. At each recombination site is shown the position in the genome or the pSV2neo sequences where this event has occurred. The numbers associated with the viral sequences were derived from the equivalent position in the Pr-RSV-C sequence (*20*). Abbreviations: pBR, pBR322; SVpro, SV40 promoter; SV pA, SV40 polyadenylation region; U3, R, and U5, components of the LTR; PBS, primer binding site; 3′ UT, 3′ untranslated region.

2) Analysis of the junctions between virus and *neo*-derived sequences (Fig. 3B) showed that little sequence identity (six or fewer common nucleotides) was needed for the recombination event to occur. In fact, in two cases (clones 21 and 32) no common bases at the site of the junction were observed, despite the presence of a four-base stretch of identical sequences adjacent to the crossover point in clone 21. Recombinatory junctions for known viral oncogenes are usually devoid of long stretches of homologous sequence (*2*), although junctions of some oncogenes, such as *ski* (*13*), are derived from short regions of shared sequences between virus and proto-oncogene.

3) Neighboring consensus sequences were not immediately apparent in these clones. Stretches that had been identified in the transduction of oncogenes by feline leukemia virus were seen but without any striking frequency (*14*). In 3 cases out of 11, the consensus ACCCC was observed (Fig. 3B) close to the junction between *neo* and viral sequences. The sequence GAGG or CCTC (Fig. 3B) was found four times near the left junction and CTCCTC was not found in any instance at the right junction.

4) Insertion of sequences at the junction between virus and *neo* genes was found in two cases, at the right junctions. Clone 21 had an additional C in a succession of C's, and clone 49 showed an insert of 16 bases identical in sequence to a region in the *pol* gene. These structures are consistent with a broad spectrum of mutations and rearrangements reported to occur during replication of other retroviruses (*15, 16*).

5) Several structures had properties that were characteristic of this system. Analysis of two clones (clones 39 and 14) suggested that the *neo* gene was in the opposite orientation relative to the viral genes. Sequencing of the amplified products for the right junction in each case showed that recombination took place in the same area of the vector sequences but different regions of the virus in the two clones (Fig. 3A). Clone 39 showed two recombination events at one junction; the U3 region of the LTR was connected to a region in the *pol* gene in the opposite orientation, which was then connected to the *neo* gene in the same orientation as the *pol*-related sequences. Some clones, of which 32 is an example, retained the left junction between *neo* and viral sequences present in the original clone. This suggested that only one recombination event, at the right junction, was necessary to move the *neo* gene between LTR sequences. These results suggest that this system, while supplying information on the mechanism of oncogene transduction, will also be useful for characterizing the range and nature of aberrations created and tolerated by the virus replication machinery.

As in the case of oncogene transfer by transforming retroviruses, two clones, 49 and 53, were able to transfer the *neo* gene by infection of fresh cells with viral particles. Infection of fresh QT35 with virions from clone 42, however, did not give rise to drug-resistant clones. This deficiency was consistent with the rearrangements in clone 42 DNA, which led to the loss of cis-acting sequences (including the primer binding site) necessary for reverse transcription.

In addition to providing a model for illegitimate recombination during reverse transcription, our model provided a means of testing directly the role of readthrough transcripts as intermediates in transduction. The experiment described above was repeated with virions from a series of cell clones that produced virions with different ratios of *neo*-containing to wild-type viral transcripts. The different ratios were obtained by changing the numbers of polyadenylation signals upstream of the *neo* gene. The DNA's introduced into QT35 cells contained the *neo* gene and the provirus with the following polyadenylation signals: the inactivating mutant signal [up to 99% of viral transcripts would be readthrough RNA's (*7*)]; wild-type signal [up to 15% of viral transcripts would be readthrough RNA's (*6*)]; and both the LTR and a second polyadenylation signal, from the L3 gene of adenovirus-2 (*17*), introduced between the 3′ LTR and the *neo* gene (a very low level of viral transcripts would contain *neo* sequences) (Fig. 4A). Cell clones resistant to G418 that contained the intact DNA sequences were isolated and used as sources of virions for the subsequent infections. Analysis of RNA from virions from the mutant (M7), two wild-type clones (WT1 and WT11), and three double polyadenylation site clones

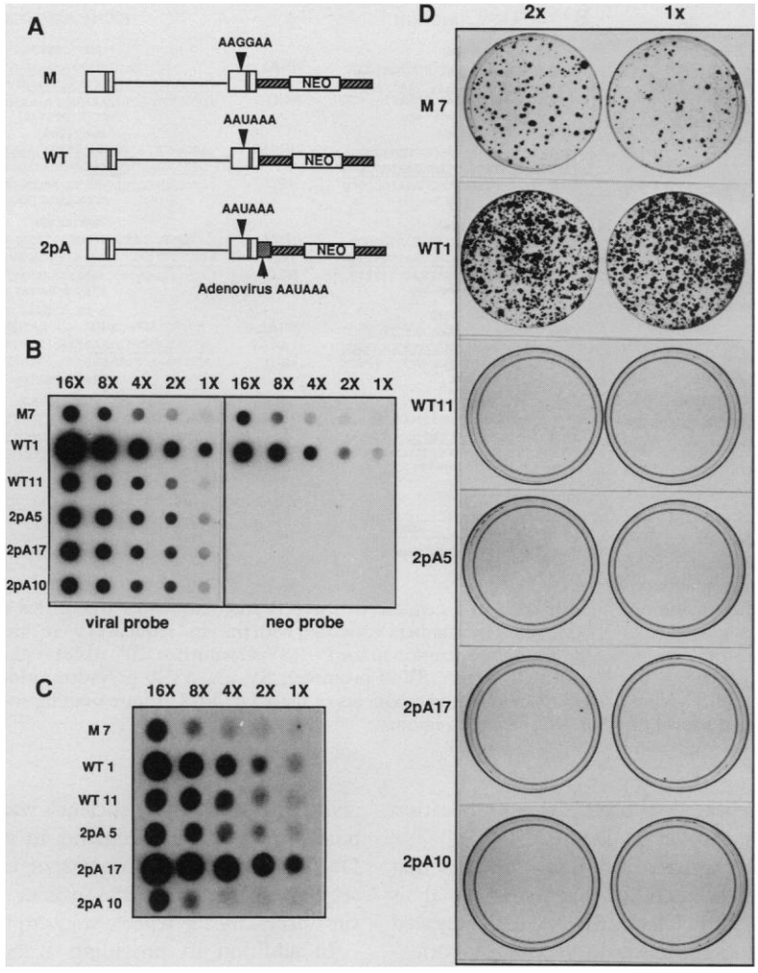

Fig. 4. Correlation between frequency of transduction and level of readthrough transcripts. (**A**) Representation of the DNA constructs introduced into cells. All constructs contained the *neo* cassette downstream of an RAV-1 provirus. Construct M and WT contain mutant (AAGGAA) and wild-type (AAUAAA) polyadenylation signal in the LTR's, respectively. Construct 2pA contained wild-type polyadenylation signals in the LTR's plus a fragment (insert in pT3L3) containing the L3 polyadenylation signal of adenovirus-2 (*17*) (shaded box) introduced at the *T*th111 I site 27 bp downstream of the pSV2neo Acc I site. (**B**) Virion RNA was isolated (*7*) and serial twofold dilutions were spotted onto nitrocellulose. Duplicate filters were hybridized to antisense ^{32}P-labeled riboprobes described in Fig. 2. (**C**) A standard reverse transcriptase assay was performed with sedimented virions from media of the cell clones described in (B) (*23*) and twofold dilutions were spotted onto DEAE cellulose paper. (**D**) Transduction ability of virus produced by the various clones. Undiluted (2×) (6 ml) or a twofold dilution of clarified medium from each clone was used to infect a culture of QT35 cells, and G418-resistant clones were selected as described in Fig. 2 and stained with methylene blue (*24, 25*).

(2pA5, 2pA10, and 2pA17) confirmed that the viral RNA's were of the expected composition and sizes (*18*).

To determine the relative ability of these clones to yield rearranged proviruses, virions were collected and used to infect QT35 cells. The amount of *neo*-containing RNA relative to the amount of total viral RNA in the virions used to infect the cells, the reverse transcriptase activity (a measure of virion number), and the number of drug-resistant clones obtained for two dilutions of virus from each cell clone were determined (Fig. 4, B to D). In all cases, the relative numbers of drug-resistant clones correlated with the relative amounts of readthrough RNA containing *neo* sequences in the infectious virions. These results suggest that the presence of readthrough transcripts is a key factor in the transduction of downstream *neo* sequences in this system.

Several aspects of our studies show that the mechanism of transduction involving readthrough transcripts is the main way that the neo sequences are captured in our system and suggest that this pathway is a probable mechanism for oncogene transduction by retroviruses. The inability of DNA structures containing two different polyadenylation signals to lead to the transduction of *neo* sequences implies that rearrangements at the DNA level (of the sort shown in Fig. 1), which should not be affected by the extent of readthrough transcription, did not contribute to this event. The heterogeneity of the *neo*-containing structures obtained after infection with virions from the M7 clone implied that they were created during infection and were not the result of a preexisting rearrangement in the original virion-producing clone. Recombination must have occurred after transcription and RNA processing to give rise to rearranged structures lacking introns as well as recombination junctions that contained sequences derived from the poly(A) tail. The presence of *neo*-containing proviruses lacking important cis-acting signals (such as primer binding sites) provides evidence that these rearrangements did not occur before the initial infection because genomes with such defects would not have been able to be mobilized by packaging into virions or reverse transcription. Our results imply that transfer of host genetic information is commonplace during retroviral infection. The system that we have developed will allow further characterization of the mechanism and consequences of this process.

REFERENCES AND NOTES

1. R. Weiss, N. Teich, H. Varmus, J. Coffin, Eds., *RNA Tumor Viruses* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1985).
2. J. M. Bishop, *Annu. Rev. Biochem.* **52**, 301 (1983).
3. M. P. Goldfarb and R. A. Weinberg, *J. Virol.* **38**, 136 (1981).
4. R. Swanstrom, R. C. Parker, H. E. Varmus, J. M. Bishop, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 2519 (1983).
5. S. A. Herman and J. M. Coffin, *Science* **236**, 845 (1987).
6. ———, *J. Virol.* **60**, 497 (1986).
7. A. Swain and J. M. Coffin, *ibid.* **63**, 3301 (1989).
8. P. J. Southern and P. Berg, *J. Mol. Appl. Genet.* **1**, 327 (1982).
9. C. Moscovici *et al.*, *Cell* **11**, 95 (1977).
10. Sequence analysis revealed that this small RNA was a spliced variant that joined 5′ LTR to SV40 sequences with the use of the splice donor in *gag* and the splice acceptor flanking the intron in SV40 sequences 3′ of the *neo* gene (Fig. 2B). The presence of this transcript in the virion was unexpected because it contained the same viral 5′ sequences as the spliced *env* mRNA that was not packaged. Subsequent experiments imply that this species did not participate in recombination, and we believe that this transcript was not necessary for the transduction of *neo* sequences.
11. C. C. Huang, N. Hay, J. M. Bishop, *Cell* **44**, 935 (1986).
12. M. A. Raines, N. J. Maihle, C. Moscovici, L. Crittenden, H.-J. Kung, *J. Virol.* **62**, 2437 (1988).
13. E. Stavnezer, D. Brodeur, L. Brennan, *Mol. Cell. Biol.* **9**, 4038 (1989).
14. D. L. Doggett *et al.*, *J. Virol.* **63**, 2108 (1989).
15. V. K. Pathak and H. M. Temin, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 6024 (1990).
16. ———, *ibid.*, p. 6019.
17. H. Skolnik-David, C. Moore, P. Sharp, *Genes Dev.* **1**, 672 (1987).
18. We do not know why WT11 produced so little readthrough RNA (Fig. 4B). In that analysis of the DNA from the clone showed no rearrangements, the effect might be due to the site of integration of the construct in the genome. The amounts of *neo*-containing readthrough transcripts seemed to be higher for clone WT1 than M7 because of WT1 producing more total viral RNA even though most of this consisted of transcripts ending in the 3′ viral LTR. The total amounts of viral RNA for all the

clones reflected the character of the particular clone and the number of cells on the plate from which the virions were obtained.
19. L. Sealy, M. L. Privalsky, G. Moscovici, C. Moscovici, J. M. Bishop, *Virology* **130**, 155 (1983).
20. D. E. Schwartz, R. Tizard, W. Gilbert, *Cell* **32**, 853 (1983).
21. S. Kawai and M. Nishizawa, *Mol. Cell. Biol.* **4**, 1172 (1984).
22. M. Frohman, in *PCR Protocols. A Guide to Methods and Applications*, M. Innis, D. Gelfand, J. Sninsky, T. White, Eds. (Academic Press, New York, 1990), p. 28.
23. S. Goff, P. Traktman, D. Baltimore, *J. Virol.* **38**, 239 (1981).
24. T. Maniatis, E. F. Fritsch, J. Sambrook, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982).
25. We thank N. Rosenberg and T. Stacy for discussion, S. Dobson for technical assistance, and C. Moore for the pT 3L3 plasmid. Supported by research grant R35 CA 44385 from the National Cancer Institute.

12 June 1991; accepted 4 November 1991

Early Evolution of Avian Flight and Perching: New Evidence from the Lower Cretaceous of China

PAUL C. SERENO* AND RAO CHENGGANG

Fossil bird skeletons discovered in Lower Cretaceous lake deposits in China shed new light on the early evolution of avian flight and perching. The 135 million-year-old sparrow-sized skeletons represent a new avian, *Sinornis santensis*, n. gen. n. sp., that preserves striking primitive features such as a flexible manus with unguals, a footed pubis, and stomach ribs (gastralia). In contrast to *Archaeopteryx*, however, *Sinornis* exhibits advanced features such as a broad sternum, wing-folding mechanism, pygostyle, and large fully reversed hallux. Modern avian flight function and perching capability, therefore, must have evolved in small-bodied birds in inland habitats not long after *Archaeopteryx*.

THE SKELETON OF THE OLDEST BIRD, *Archaeopteryx*, is characterized by elongate grasping forelimbs and a long balancing tail, and its skeleton resembles in many regards that of its nearest theropod relatives (*1*). *Archaeopteryx*, nevertheless, was clearly capable of gliding or primitive powered flight, as evidenced by flight feathers of modern avian aerodynamic design and arrangement along the forearm and manus (*2*). Late Cretaceous carinate birds such as *Ichthyornis*, in contrast, exhibit an advanced avian flight apparatus, including an expansive keeled sternum for bulky flight musculature and wing and tail modifications for aerial maneuverability (*3*). The basic components of the modern avian flight apparatus, thus, must have arisen in the Early Cretaceous, during the first third of avian history.

But the fossil record for the Early Cretaceous, thus far, has provided little documentation of this critical transformation (*4*). Recently, partial bird skeletons have been discovered in the Lower Cretaceous of Asia (*5*) and Europe (*6*) (Fig. 1) that provide important clues to the early evolution of flight, but none of these specimens is complete. We report on the discovery of sparrow-sized bird skeletons in Lower Cretaceous beds in northeastern China (*7*). The holotype skeleton (Fig. 2B) of the new bird, *Sinornis santensis*, n. gen. n. sp. (*8*), is preserved on part and counterpart slabs of fine-grained freshwater lake sediment and is associated with abundant fish (*Lycoptera*), insect, and plant remains (*9*). The associated pollen and spore assemblage suggests a Valanginian age for the bird (*10*), which, if correct, would make it second only to *Archaeopteryx* in age among birds (Fig. 1C).

The skull and skeleton exhibit a number of striking primitive features that have not been reported thus far in any bird except *Archaeopteryx*. The skull has a proportionately short, toothed snout (Fig. 2A) as in *Archaeopteryx*. The carpus and manus in the forelimb are separate, rather than fused into a unit carpometacarpus, and the manus is composed of freely articulating metacarpals, with well-formed phalanges and unguals on the first and second digits (Fig. 3C). The manual unguals are relatively small and only moderately recurved, in contrast to the slender, highly recurved unguals in *Archaeopteryx* (*11, 12*). The pelvis is remarkably primitive and closely resembles that of *Archaeopteryx* (Fig. 4B). The elements of the pelvic girdle are free rather than coossified (*6, 12*), the iliac blades are erect rather than converging toward the midline, and the ischium is blade-shaped rather than strap-shaped. The rodlike shaft of the pubis appears to be directed more ventrally than posteriorly and terminates distally in a hook-shaped foot similar to that in *Archaeopteryx* and theropod dinosaurs. As in *Archaeopteryx*, the metatarsals are separate, rather than coossified, along all but their proximal ends. Rows of slender stomach ribs (gastralia) are preserved on the ventral aspect of the trunk, passing between the hind limbs (Fig. 4B). Gastralia have been lost in all other birds except *Archaeopteryx* (*12*). These retained archaic features are not specifically involved in flight or perching and add to current evidence that favors theropod dinosaurs as the nearest avian relatives. Lack of fusion in the manus, pelvis, and hind limbs seems to document the primitive avian condition, rather than signify immaturity in the holotype skeleton, because bone surfaces are finished throughout the skeleton and the components in the dorsal vertebrae and pygostyle are completely fused (*13*).

Advanced avian characters in the skeleton of *Sinornis* are almost all directly related to flight or perching. There appear to be no more than 11 dorsal vertebrae in the trunk as in the Spanish bird (*6*), rather than 14 as occur in *Archaeopteryx* and most theropods. The tail is short with only eight free vertebrae and a large pygostyle for attachment of the rectricial fan. The short trunk and tail in *Sinornis* shift the center of mass toward the forelimbs, as in modern powered fliers, as opposed to a center of mass near the hind limbs, as in the terrestrial cursor *Archaeopteryx*. In living birds the pygostyle is closely correlated with the size of the rectricial fan

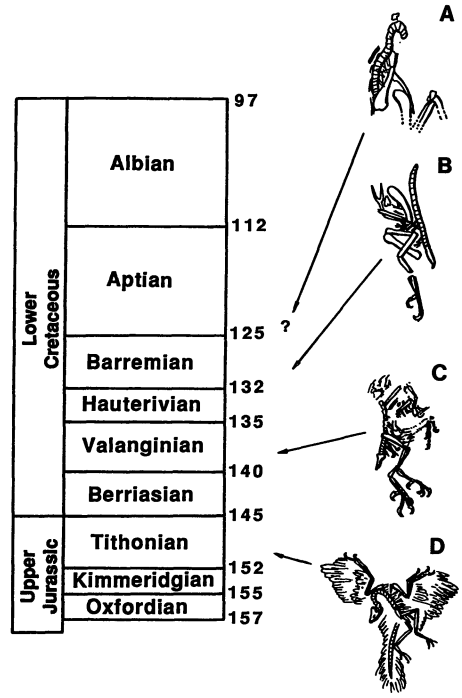

Fig. 1. Temporal position of Upper Jurassic and Lower Cretaceous fossil birds. (**A**) *Ambiortus dementjevi*, (**B**) Las Hoyas bird, (**C**) *Sinornis santensis*, and (**D**) *Archaeopteryx lithographica*.

P. C. Sereno, Department of Organismal Biology and Anatomy, and Committee on Evolutionary Biology, University of Chicago, Chicago, IL 60637.
Rao C., Beijing Natural History Museum, Beijing, People's Republic of China.

*To whom correspondence should be addressed.