Proc. Natl. Acad. Sci. USA
Vol. 75, No. 12, pp. 5812–5816, December 1978
Biochemistry

Recombination between viral and cellular sequences generates transforming sarcoma virus

(transformation-defective mutants/recovered avian sarcoma virus/*src*-specific sequences/oligonucleotide fingerprinting)

L.-H. WANG, C. C. HALPERN, M. NADEL, AND H. HANAFUSA

The Rockefeller University, New York, New York 10021

Communicated by Purnell W. Choppin, August 30, 1978

ABSTRACT **A series of sarcoma viruses has been obtained from tumors induced by transformation-defective (td) mutants of the Schmidt–Ruppin strain of Rous sarcoma virus, subgroup A (SR-A). The RNA sequences of these "recovered avian sarcoma viruses" (rASVs) were compared with those of td mutants and of SR-A by oligonucleotide fingerprinting. Of six sarcoma-specific oligonucleotides present in SR-A RNA, three to six were missing in the RNAs of the four td mutants examined. All six isolates of rASV examined have regained these six oligonucleotides. In addition, most rASV RNAs have three new oligonucleotides not present in the RNA either of td mutants or of SR-A. The newly obtained oligonucleotides are located between 800 and 2600 nucleotides from the 3′ end of rASV RNA, which corresponds to the *src* region of SR-A RNA mapped previously. Furthermore, viral RNAs of two td mutants isolated from a clone of rASV lack most *src*-specific oligonucleotides, including the three new ones. No differences were found among RNAs of td, SR-A, and rASV in the regions outside of *src*. Our results indicate that RNA sequences that rASVs have acquired from cells in the process of conversion from td virus to transforming virus are mapped within the *src* region and segregate with the transforming function. Some of the sequences are new and some are identical with those in SR-A RNA.**

Transformation-defective (td) mutants of avian sarcoma viruses carry deletions of the *src* gene (1). The deletion was found to be incomplete in some td mutants derived from nondefective Schmidt–Ruppin Rous sarcoma virus (RSV) of subgroup A (SR-A). RNAs of these mutants and other similar isolates are larger than those of avian leukosis viruses (2–4) and hybridize with as much as 20–30% of DNA complementary to SR-A sarcoma (*src*)-specific sequences (5). We have reported the isolation of new sarcoma viruses from chicken tumors produced after long latent periods after injection of some of these td mutants (6). The sarcoma viruses thus isolated were called "recovered avian sarcoma viruses" (rASVs) (6). Biological and biochemical properties of rASVs, such as focus morphology, subgroup specificity, size of viral structural proteins, and size of genome RNA, indicate that rASVs were probably generated by recombination between td viral genomes and cellular sequences (5, 6). The present study was undertaken to characterize and compare the genomic RNA sequences of rASVs, parental td viruses, and SR-A, with particular emphasis on their *src*-specific sequences.

MATERIALS AND METHODS

Cells and Viruses. Chicken embryo fibroblasts were cultured as described (7). SR-A 85-1 and SR-A 85-7 were isolated by Sadaaki Kawai; the two clones differed slightly in their attachment to cells (5). td mutants of SR-A (td101, td105, td107A, and td108) were isolated as spontaneous mutants by Kawai *et al.* (3) and further purified by one or two terminal dilutions (6). All rASVs have been described (5, 6) except nos. 362 and 367, which were isolated from chicken tumors induced by injection of td105 into 1-day-old chickens in the same manner as described (6). td111, td112, and td113 were isolated from a clone of rASV no. 1441 by the published procedure (8, 9). The derivation of the viruses is summarized in Fig. 1.

RNA Analyses. Preparation of isotope-labeled viral RNAs, RNase T1 fingerprinting, and RNase A analysis have been described (10, 11). Polyacrylamide gel electrophoresis of viral RNAs followed the published procedure (1). In all cases, viral 60–70S RNAs were used for analyses. For most rASVs, genomic 39S RNAs (12), purified from heat-denatured, viral 60–70S RNAs, were also analyzed, and the same results were obtained as with 60–70S RNAs.

RESULTS

td SR-A Viruses That Carry Partial Deletions of *src* Gene. Fig. 1 shows the lineage of the viral strains analyzed. td107A has consistently failed to induce chicken tumors and to generate rASVs. ^{32}P-Labeled 60–70S RNAs of various viruses were subjected to complete RNase T1 digestion followed by two-dimensional separation of the digestion products (10). The resulting fingerprinting patterns of RNase T1-resistant oligonucleotides are shown in Fig. 2 *A–D*. SR-A 85-1 and 85-7 gave identical fingerprints; six oligonucleotides, 8a, 10a, 13, 27, 32b, and 37, were identified as *src* specific based on previous studies (13) and present mapping data with various poly(A)-containing RNA fragments (see Figs. 3 and 5). Both td101 and td105 lack five *src*-specific oligonucleotides of SR-A but retain oligonucleotide 27. td108 lacks three oligonucleotides, 8a, 10a, and 37, whereas td107A lacks all six *src*-specific oligonucleotides present in parental SR-A. From the relative locations of the six *src*-specific oligonucleotides, the extents and locations of the deletions in td viruses can be roughly estimated (see Fig. 5).

Genetic Identity of rASVs. Direct analyses of RNA sequences are necessary to clarify the origin of rASVs. Six independent isolates of rASVs derived from three td viruses were analyzed. The size of heat-denatured 60–70S RNAs of rASVs, estimated by sodium dodecyl sulfate/polyacrylamide gel electrophoresis, was indistinguishable from that of Prague strain of RSV-B (PR-B) (data with rASV no. 1441 are shown in Fig. 4; others not shown). This confirms our earlier conclusions from analyses of rASV RNAs by velocity sedimentation in sucrose gradients (5).

The fingerprints of 60–70S [^{32}P]RNAs of rASV are similar, but not identical, to those of parental td and SR-A viruses (see Figs. 2 and 5). rASV RNAs contain additional oligonucleotides not present in td and SR-A viruses. All rASVs tested, except

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: td, transformation-defective; RSV, Rous sarcoma virus; SR-A, Schmidt–Ruppin strain of RSV, subgroup A; PR-B, Prague strain of RSV, subgroup B; rASV, recovered avian sarcoma virus; *src*, gene responsible for sarcoma formation in animals and for cell transformation.

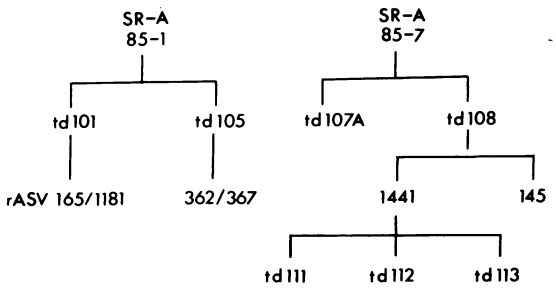

FIG. 1. Relationship among SR-A, td SR-As, rASVs, and td viruses derived from rASV.

rASV no. 1441, contain six *src*-specific oligonucleotides present in SR-A RNA. All rASVs except no. 362 contain three oligonucleotides not present in SR-A RNA; rASV no. 362 contains two of the three new oligonucleotides. rASV no. 1441 lacks the *src*-specific oligonucleotide 8a present in SR-A but has one new oligonucleotide 8c (Fig. 2*G*). RNase A-resistant fragments of oligonucleotides 8a and 8c reveal that they are homologous oligonucleotides (Table 1). The appearance of oligonucleotide 8c in RNA from rASV no. 1441 is most likely due to a single mutation that converts AU, present in oligonucleotide 8a, into AC. As shown in Figs. 2, 3, and 5, td108 contains two oligonucleotides, 32d and 39, not present in parental SR-A 85-7. These two td-108-specific oligonucleotides appear in both rASV nos. 1441 and 145; this is compelling evidence that both rASVs have indeed been derived from td108 used in the injection into chickens.

Each numbered oligonucleotide shown in Fig. 2 was analyzed for its RNase A-resistant fragments; the results confirm and extend data published previously (13). Therefore, only oligonucleotides to be discussed here and those not analyzed before are listed in Table 1.

We conclude that rASVs have regained the *src*-specific oligonucleotides missing in the corresponding td viruses; in

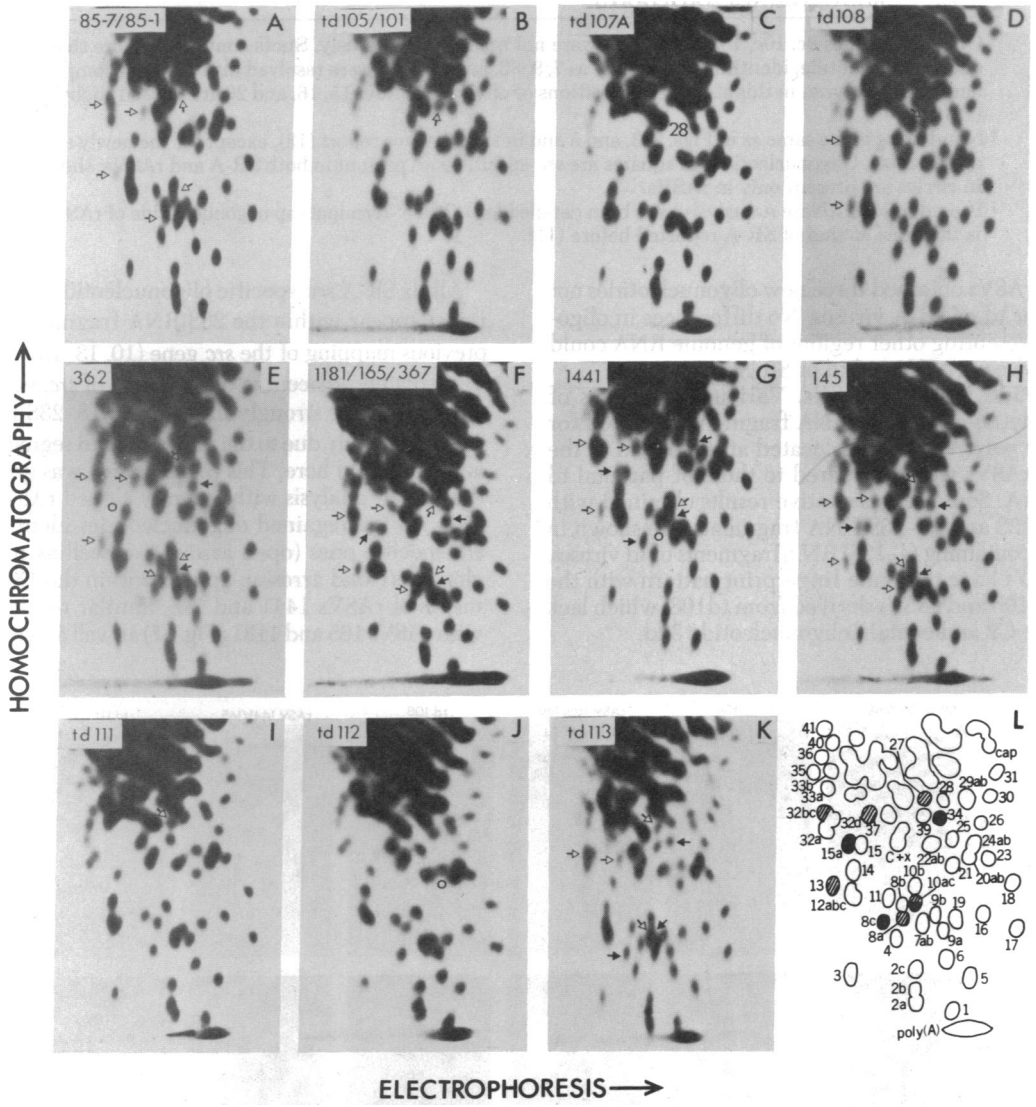

FIG. 2. Fingerprint patterns of RNase T1-resistant oligonucleotides of 60–70S viral [^{32}P]RNAs. Only one fingerprint is shown for different viruses that give identical patterns: 1181/165/367 in *F* means that the actual fingerprint of 60–70S RNA of rASV no. 1181 is shown and the pattern is the same for nos. 165 and 367. Similar designations apply in *A* and *B*. Numbering of spots in each fingerprint follows the tracing (*L*) of a typical pattern of rASV no. 145 (*H*). Spot 8c was added to the tracing to indicate the relative chromatographic position of that spot. Spots darkened in *L* and pointed to by solid arrows in other panels are present only in rASVs. Spots shaded in *L* and pointed to by open arrows in other panels represent *src*-specific oligonucleotides present in both SR-A and rASVs. Open circles in *E*, *G*, and *J* indicate the positions of missing oligonucleotides that would have been located in the corresponding fingerprints. *A* is the fingerprint of SR-A 85-7 60–70S RNA after alkaline fragmentation and removal of poly(A)-containing fragments. Oligonucleotide 28 of td 107A viral RNA chromatographed at a position similar to that of oligonucleotide 34 in that particular fingerprint (*C*); therefore, spot 28 is numbered to avoid confusion.

Table 1. RNase T1-resistant oligonucleotides of SR-A, tdSR-As, rASVs, and td rASVs*

Spot no.†	RNase A digestion products‡	Spot no.†	RNase A digestion products‡
7a	5U,5C,(AAG),(A_5N)	27	2U,2C,(AAG),(A_{4-5}N)
7b	5U,4C,(AC),0.5(AU),(AAAN),(A_5G)	32a	4C,G,4(AC)
8a	3U,5C,G,2(AC),(AU),(AAAC)	32b	4C,G,(AC),(A_4N)
8b	5U,9C,G,2(AC)	32c	4C,3(AC),(AG)
8c	3U,4C,G,3(AC),(AAAC)	32d	3C,G,2(AC),(AAAC)
9a	5U,7C,G,3(AU)	33a	4C,2(AC),(AG),(AAC)
9b	3U,4C,2(AU),(AG),(AAC),(AAU)	33bc	6C,3(AC),(AG),(AAC),(AAG)
10a	4U,5C,G,2(AC),(AU)	34	4U,3C,G,(AC),(AAC)
10b	3U,4C,G,3(AC),2(AU)	35	U,5C,(AC),(AG)
10c	3U,6C,2(AC),(AU),(AG)	36	4C,2(AC),(AG)
13	4C,(AC),(AAC),(AAG)	37	U,4C,(AC),(AU),(AG)
15	4C,(AU),(AG),(A_{4-5}N)	39	2U,3C,G,(AC),(AU),(AAU)
15a	6C,G,2(AC),(AU)	40	2C,2(AC),(A_4G)
16	5U,4C,G,2(AU),(A_4N)	41	2C,G,3(AC)
20	3U,C,2G,4(AU),2(AAU),2(AAAC)	C	G,(AC),(AU),(AAU),(AAAC)
21	3U,2C,G,(AC),(AU),(AAC)	Cap	$m^7G^{5'}ppp^{5'}$GmCp(2U,C,AU)Gp
22	6U,7C,2G,2(AC),2(AU),(AAU)		

* Oligonucleotides 8c, 10c, 15a, and 34–41 were not analyzed previously. Spots containing more than one oligonucleotide, identified previously as 7, 9, 32, and 33 (13), were resolved into their subcomponents and analyzed in this study. Compositions of oligonucleotides 15, 16, and 20 are revised slightly (13).

† Numbering is the same as in Figs. 2, 3, and 5 and in the previous report (13), except for the newly analyzed spots. Oligonucleotides in squares are *src*-specific ones present in both SR-A and rASVs; those in circles are present only in rASVs.

‡ Procedures for RNase A analysis have been detailed (11). The 5′-terminal cap oligonucleotide of rASVs is the same as that of SR-A reported before (11).

addition, most rASVs obtained three new oligonucleotides not present in either td or SR-A viruses. No differences in oligonucleotides representing other regions of genome RNA could be detected between rASVs and td or SR-A viruses.

Oligonucleotide Maps of rASVs. Various size classes of ^{32}P-labeled, poly(A)-containing RNA fragments [selected for the presence of poly(A), which is located at the 3′ end of the viral RNA] of rASVs were compared to those of parental td viruses and SR-A. Some representative results obtained with two classes, 4–12S and 18–23S RNA fragments, are shown in Fig. 3. Poly(A)-containing (4–12S) RNA fragments of td viruses, SR-A, and rASVs have the same fingerprint pattern with the exception of td108 and rASVs derived from td108, which lack oligonucleotide C2 and contain oligonucleotide 32d.

All six SR-A *src*-specific oligonucleotides (Fig. 3*G*, open arrows) appear within the 23S RNA fragment, confirming the previous mapping of the *src* gene (10, 13–15). Oligonucleotides immediately adjacent to the 5′ region of *src*, such as 9b, 21, and 22, also appear strongly in the SR-A 18–23S RNA fragments. This is probably due to the presence of td segregants in the SR-A virus analyzed here. This complication was greatly reduced in the similar analysis with recently cloned rASV no. 1441 (Fig. 3*K*). All the regained oligonucleotides identical to the SR-A *src*-specific ones (open arrows), as well as the new oligonucleotides (solid arrows), appear within the 18–23S RNA fragments of rASVs 1441 and 145. Similar results were obtained with rASVs 165 and 1181 (Fig. 3*I*) as well as with rASVs 362 and

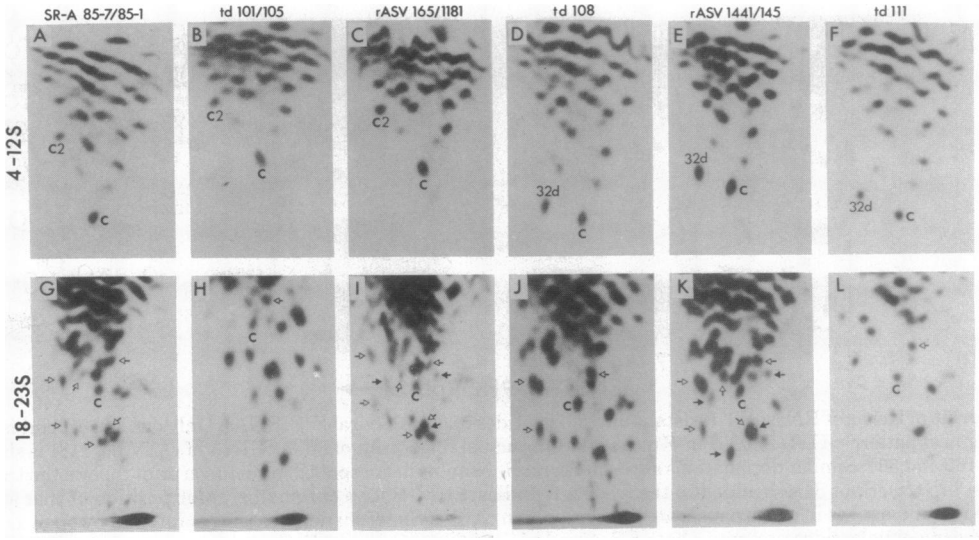

FIG. 3. Fingerprints of viral poly(A)-containing RNA fragments. Preparation of different sizes of poly(A)-tagged RNA fragments has been detailed (10). Only the upper two-thirds of the autoradiograms are shown for fingerprints of 4–12S RNA fragments since no other oligonucleotides except poly(A) appear in the lower one-third. Large RNase T1-resistant oligonucleotides present in each RNA fragment were recovered from DEAE-cellulose plates and characterized by RNase A analysis (11). Open and solid arrows, see legend for Fig. 2.

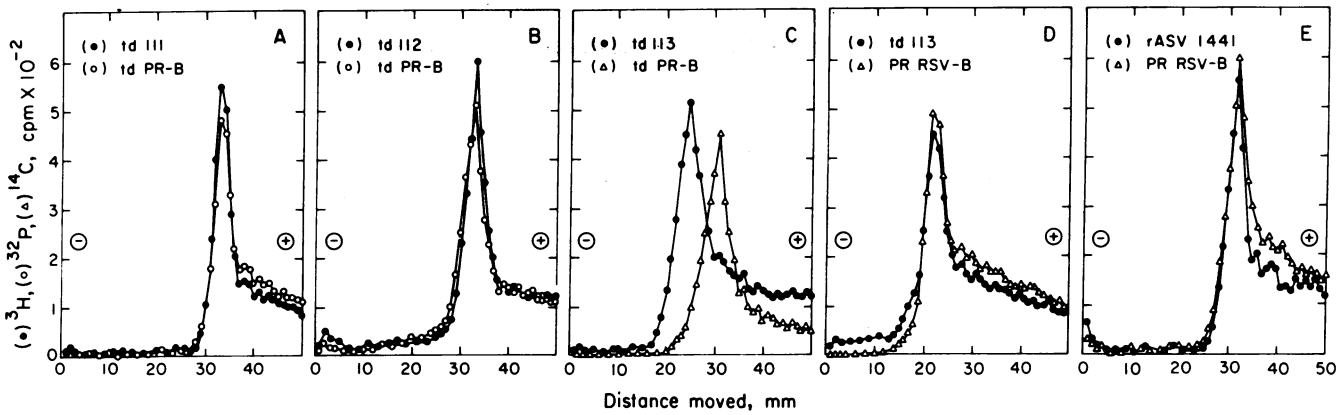

FIG. 4. Polyacrylamide gel electrophoresis of viral RNAs. The viral 60–70S RNAs were heat-denatured and electrophoresed in 2.2% polyacrylamide gels at 6 V/cm of gel for 7–8 hr by the published procedures (1).

367 (data not shown). The relative locations of oligonucleotides present in 18–23S poly(A)-containing RNA fragments of rASVs were deduced from the data shown in Fig. 3, as well as from other data (not shown) obtained with fragments between 12 and 18S and from data quantitating the yield of each oligonucleotide present in 18–23S RNA fragments (13). The resulting order of oligonucleotides is shown in Fig. 5.

td Viruses Derived from rASV. To further define the *src*-specific sequences in rASV, we isolated three td viruses, td111, td112, and td113, from a clone of rASV no. 1441. The size of RNA of td111 and td112 was similar to that of a td virus, td PR-B (Fig. 4). The RNA of td113 is much larger than that of td PR-B and is indistinguishable in size from that of PR-B (Fig. 4). The RNA of parental rASV no. 1441 is also indistinguishable in size from that of PR-B (Fig. 4*E*).

td111 lacks eight oligonucleotides present in rASV no. 1441 (Fig. 2 *G* and *I*): five *src*-specific ones, present in both rASV and SR-A, as well as the three rASV-specific oligonucleotides. td112 lacks the same eight oligonucleotides deleted in td111, as well as oligonucleotide 21, which maps immediately adjacent to the 5′ end of the *src* region in rASV no. 1441 (Fig. 2 *G* and *J*; Fig. 5). td113 lacks only two oligonucleotides, including the rASV-specific oligonucleotide 15a. All three td viruses retain oligonucleotides 32d and 39, characteristic of rASV no. 1441 originally derived from td108 (Figs. 2 and 3). Comparison of poly(A)-containing RNA fragments of td111 with those of td108 and of rASV no. 1441 is shown in Fig. 3. The extents and locations of the deletions in the three td viruses are estimated (Fig. 5). These results provide direct evidence that the three newly obtained oligonucleotides are part of the *src* gene of rASVs since these RNA sequences segregate with the transforming function. Thus, a total of nine oligonucleotides can be identified as *src* specific in rASVs.

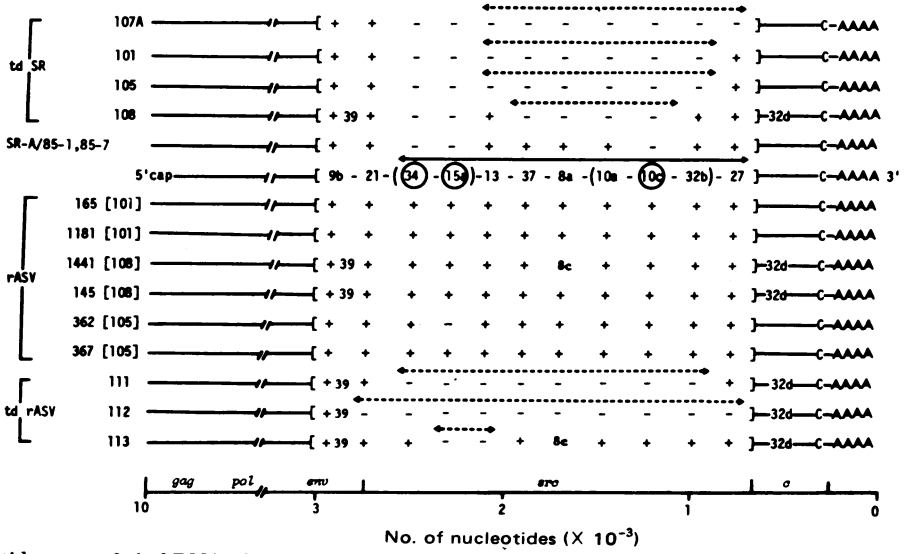

FIG. 5. Oligonucleotide maps of viral RNAs. Numbering of oligonucleotides is the same as in Figs. 2 and 3 and Table 1. The four regions of sarcoma viral RNA (16) (i.e., *src* for sarcoma formation, *env* for envelope glycoprotein, *pol* for DNA polymerase, and *gag* for group-specific antigens) are indicated in the established map order (17) on the bottom scale, representing the 39S viral RNA. *c*, A highly conserved region, characterized by the oligonucleotide C which is present in all exogenous avian RNA tumor viruses analyzed to date (10). Each RNA is represented by a horizontal bar with bracketed *src* region. The 3′-terminal 800 nucleotides (except oligonucleotide 32d) and the 5′ two-thirds of the genome are indistinguishable for all viral RNAs analyzed here and are represented by solid bars. The oligonucleotide map of rASV 39S RNA is shown by line no. 6. The relative location of *src*-specific oligonucleotides within parentheses is uncertain. The + and − symbols above and under each oligonucleotide designate the presence or absence of that oligonucleotide in viral RNA. Solid line within the bracket defines the *src*-specific sequences of rASV; broken lines define the deletions in the td viruses. The portion of RNA in SR-A and its td derivatives corresponding to the region characterized by oligonucleotides 15a and 34 in rASV should be interpreted most likely as lack of large RNase T1-resistant oligonucleotide markers rather than as a deletion. Therefore it is not certain whether RNA sequences within that region are missing in td SR viruses; the lengths of deletion of td101, td105, and td107A may not be shorter than that of td111. Each td viral RNA should actually be shorter than the RNA of SR-A or rASV by the length of its deletion.

DISCUSSION

The length of the deletions in td viruses derived from SR-A, estimated by oligonucleotide mapping, agrees with studies of td viruses on their RNA sizes (5), on their contents of remaining *src* sequences measured by hybridization (5), and on their ability to induce tumors and to give rise to rASVs after injection into chickens (6). td101 and td105 differ from td107A only by the presence of one additional oligonucleotide at the 3′ portion of the *src* region. Since td107A has been unable to generate rASVs, the 3′ *src* sequences remaining in td101 and td105 may be necessary for recombination with the cellular sequences to generate rASVs.

Our results provide evidence that strongly suggests that rASVs have been generated by recombination between td viral and cellular *src* sequences which have been identified in normal chicken and other avian cells (18, 19, 20). All rASVs except rASV no. 362 obtained the same sequences not present in either td or SR-A viruses. The newly obtained sequences of rASVs map within the *src* region of viral RNA and segregate with the transforming function. rASV no. 362 contains only two of the three new *src*-specific oligonucleotides. This may be due to a mutation after recombination or to a double crossover in the region containing oligonucleotide 15a. No other sequence differences could be detected between rASVs and td or SR-A viruses.

Our results rule out the possibilities that rASVs represent an endogenous sarcoma virus induced by injection of td viruses into chickens or that rASVs are revertants of td viruses carrying point mutations in *src* and deletions in other part(s) of the genome. One can also rule out the argument that, despite purification by endpoint dilution, the td viruses still contained a very low level of parental SR-A virus that was enhanced through *in vivo* passage, because the sarcoma virus recovered from tumors consistently carries the three oligonucleotide sequences that are not present in SR-A RNA. The presence of oligonucleotides 32d and 39 in rASV derived from td108 further demonstrates the direct genealogical relationship between td virus and rASV. Further, there is a good correlation between the size of td virus RNA and the ability of the td virus to generate rASV.

Six of the nine *src*-specific oligonucleotides of rASVs are also present in SR-A. This raises a possibility that the SR-A *src* gene is of cellular origin. At least one of the three new *src* oligonucleotides of rASVs, namely, oligonucleotide 10c, is also present in PR-B and in Bratislava 77 strain of avian sarcoma virus, subgroup C (10, 13–15). Furthermore, five of nine rASV *src*-specific oligonucleotides are present in PR-B (unpublished observations). Thus, different avian sarcoma viruses may have a common origin for their *src* genes. This notion is consistent with the findings by hybridization that different avian sarcoma viruses share 80–90% of their *src* sequences with one another and at least 75% with cellular *src* sequences (12, 19, 21). The differences in *src*-specific oligonucleotides for different strains of sarcoma viruses could be due to simple mutations. Alternatively, there may be more than one copy of cellular *src* genes which are highly related but not identical to each other. Different sarcoma viruses may have inherited different *src* genes from cells. It would be very interesting to see whether td viruses of different strains of RSV could recover different cellular *src* sequences.

The isolation of murine sarcoma viruses by passage of murine leukemia viruses *in vivo* has been reported (22–24), although the mechanism for the generation of the sarcoma viruses (25, 26) is not entirely clear. Our findings differ from those reported with the murine system in that rASVs described here have gained sarcoma-genic information and have at the same time retained all the replicative functions of parental td virus.

We thank Rosemary DiGiorgio for assistance with these experiments. This work was supported by Grant CA14935 from the National Cancer Institute and Grant VC-225 from the American Cancer Society. L.H.W. and M.N. are recipients of fellowships from the Damon Runyon–Walter Winchell Cancer Fund. C.C.H. was supported by Institutional Training Grant T32CA09256 from the National Cancer Institute.

1. Duesberg, P. H. & Vogt, P. K. (1973) *Virology* **54**, 207–219.
2. Lai, M. M.-C., Hu, S. S. F. & Vogt, P. K. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 4781–4785.
3. Kawai, S., Duesberg, P. H. & Hanafusa, H. (1977) *J. Virol.* **24**, 910–914.
4. Yoshida, M. & Ikawa, Y. (1977) *Virology* **83**, 444–448.
5. Halpern, C. C., Hayward, W. S. & Hanafusa, H. (1979) *J. Virol.*, in press.
6. Hanafusa, H., Halpern, C. C. Buchhagen, D. L. & Kawai, S. (1977) *J. Exp. Med.* **146**, 1735–1747.
7. Hanafusa, H. (1969) *Proc. Natl. Acad. Sci. USA* **63**, 318–325.
8. Vogt, P. K. (1971) *Virology* **46**, 939–946.
9. Kawai, S. & Hanafusa, H. (1972) *Virology* **49**, 37–44.
10. Wang, L.-H., Duesberg, P. H., Beemon, K. & Vogt, P. K. (1975) *J. Virol.* **16**, 1051–1070.
11. Wang, L.-H., Duesberg, P. H., Robins, T., Yokota, H. & Vogt, P. K. (1977) *Virology* **82**, 472–492.
12. Hayward, W. S. (1977) *J. Virol.* **24**, 47–63.
13. Wang, L.-H., Duesberg, P. H., Kawai, S. & Hanafusa, H. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 447–451.
14. Joho, R. H., Billeter, M. A. & Weissmann, C. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 4772–4776.
15. Coffin, J. M. & Billeter, M. A. (1976) *J. Mol. Biol.* **100**, 293–318.
16. Baltimore, D. (1975) *Cold Spring Harbor Symp. Quant. Biol.* **39**, 1187–1200.
17. Wang, L.-H., Galehouse, D., Mellon, P., Duesberg, P., Mason, W. S. & Vogt, P. K. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 3952–3956.
18. Stehelin, D., Varmus, H. E., Bishop, J. M. & Vogt, P. K. (1976) *Nature (London)* **260**, 170–173.
19. Wang, S. Y., Hayward, W. S. & Hanafusa, H. (1977) *J. Virol.* **24**, 64–73.
20. Spector, D. H., Baker, B., Varmus, H. E. & Bishop, J. M. (1978) *Cell* **13**, 381–386.
21. Stehelin, D., Guntaka, R. V., Varmus, H. E. & Bishop, J. M. (1976) *J. Mol. Biol.* **101**, 349–365.
22. Harvey, J. J. (1964) *Nature (London)* **204**, 1104–1105.
23. Moloney, J. B. (1966) *Natl. Cancer Inst. Monogr.* **22**, 139–142.
24. Roy-Burman, P. & Klement, V. (1975) *J. Gen. Virol.* **28**, 193–198.
25. Scolnick, E. M., Rands, E., Williams, D. & Parks, W. P. (1973) *J. Virol.* **12**, 458–463.
26. Frankel, A. E. & Fischinger, P. J. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 3705–3709.