MOLECULAR AND CELLULAR BIOLOGY, Aug. 1982, p. 1014–1019
0270-7306/82/081014-06$02.00/0

Vol. 2, No. 8

NOTES

Human Gene (c-*fes*) Related to the *onc* Sequences of Snyder-Theilen Feline Sarcoma Virus

GENOVEFFA FRANCHINI,* EDWARD P. GELMANN, RICCARDO DALLA FAVERA, ROBERT C. GALLO, AND FLOSSIE WONG-STAAL

Laboratory of Tumor Cell Biology, Division of Cancer Treatment, National Cancer Institute, Bethesda, Maryland 20205

Received 9 February 1982/Accepted 31 March 1982

The *onc* gene (v-*fes*) of the acutely transforming feline sarcoma virus (Snyder-Theilen strain) has homologous cellular sequences (c-*fes*) in all vertebrate species, including humans. We isolated from a human DNA library recombinant phages containing overlapping c-*fes* sequences. The human c-*fes* locus spans a region of 3.4 kilobases and contains 1.4 kilobases of DNA homologous to the viral *onc* sequence interspersed with three intervening sequences.

Feline sarcoma viruses (FeSVs) are type C retroviruses that cause fibrosarcomas in cats and transform cultured fibroblasts of different mammalian species. Three isolates of FeSV have been identified: the Snyder-Theilen (ST) strain (25), the Gardner-Arnstein strain (10), and the Sarma-McDonough strain (13). These viruses are replication defective; their gene order is 5′-Δ*gag-onc*-Δ*env-c* region-3′ (22, 23). The *onc* genes of FeSV are thought to be derived from normal vertebrate DNA. The ST and Gardner-Arnstein strains have, in fact, acquired the same set of cat cellular sequences (c-*fes*) (1, 8, 28). Three avian sarcoma viruses, Fujinami strain, PRC II, and UR-I (14, 24) have recovered the chicken cellular gene homologous to the cat c-*fes*.

The transformation-specific proteins of the FeSVs are fused *gag-onc* polypeptides which exhibit associated tyrosine protein kinase activity (2, 20). Although the nature and the role of the c-*fes* protein product are not known, they are likely to be similar to those of the v-*fes* gene in a manner analogous to other characterized *onc* gene products (15, 17, 21, 31).

Total human DNA digested with *Eco*RI and hybridized with ^{32}P-labeled ST-FeSV-pBR322 recombinant plasmid (pFeSV) discloses one hybridizing band of 14 kilobases (kb), suggesting that the human c-*fes* gene is confined to a single locus (Fig. 1, lane a). To study the structure and arrangement of c-*fes* in more detail, we screened a human *Alu*I-*Hae*III DNA library (11) by using pFeSV as a probe. Plating 6×10^5 recombinant phages (corresponding to the equivalent of about two genomes), we isolated five clones (3), three of which, λ-N-15, λ-N-17, and λ-N-26, were analyzed further. Since the DNA library had been constructed by ligating DNA fragments to Charon 4A vector arms with artificial *Eco*RI linkers, digestion of the recombinant phages with *Eco*RI regenerates the vector arms and the insert DNA in one or more pieces, depending on the natural *Eco*RI sites present in the insert. In the N-15, N-17, and N-26 clones, *Eco*RI generated fragments of 14 and 2 kb, 10 and 5 kb, and 8.4, 2.5, and 2 kb, respectively. In each case only the larger fragment hybridized to pFeSV (Fig. 1, lanes c, f, and h). N-15 contained the entire 14-kb *Eco*RI fragment detected in total genomic DNA. When digested with *Sac*I, human DNA yielded three fragments of 4, 2.5, and 0.8 kb that hybridized to pFeSV (Fig. 1, lane b). The same three fragments were generated by *Sac*I digestion of the N-15 and N-26 clones, suggesting that all of the c-*fes* coding sequences are present in these two recombinants (Fig. 1, lanes g and i). *Sac*I digestion of N-17 yielded the 2.5- and 0.8-kb fragments, whereas the 4-kb fragment was replaced by a higher-molecular-weight band of 20.8 kb (Fig. 1, lane d). *Sac*I and *Eco*RI codigestion converted the latter band to a 1-kb fragment detected by pFeSV, suggesting that N-17 contains only this fraction of the 4-kb *Sac*I fragment in human DNA (Fig. 1, lane e). However, N-17 may still contain most of the coding c-*fes* sequences. N-17, radiolabeled by nick translation, detected all of the restriction fragments of ST-FeSV that contain v-*fes* sequences (data not shown).

For orientation and mapping of the three c-*fes* clones, three probes were used in Southern blot

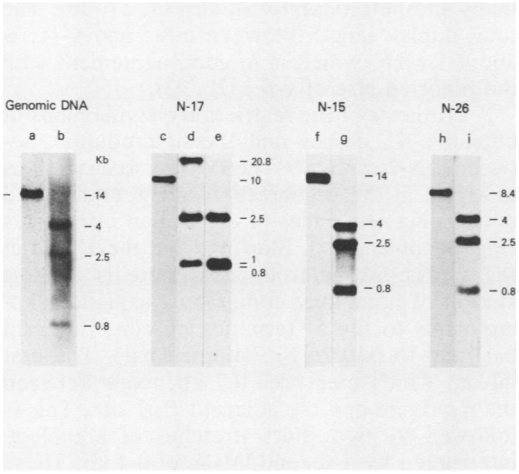

FIG. 1. Autoradiograms of DNA samples. High-molecular-weight human placenta DNA and purified phage DNAs were digested under standard conditions with the restriction enzymes *Eco*RI and *Sac*I (2 U of enzyme per μg of DNA) and were electrophoresed on 0.8% agarose gels. The DNA was transferred to nitrocellulose filters and hybridized to ^{32}P-nick-translated pFeSV (18, 26, 29). The length of the DNA fragments is expressed in kilobases. Lanes: genomic DNA digested with *Eco*RI (a) and *Sac*I (b); N-17 clone DNA digested with *Eco*RI (c), *Sac*I (d), and *Eco*RI-*Sac*I (e); N-15 clone DNA digested with *Eco*RI (f) and *Sac*I (g); and N-26 clone DNA digested with *Eco*RI (h) and *Sac*I (i). The construction of pFeSV by using the FeSV insert purified from λ-ST-FeSV was according to standard protocols (4, 12).

analyses. These probes were pFeSV and two plasmid subclones of v-*fes*, S_L and S_R, which represent two 500-base pair *Pst*I fragments encompassing 80% of v-*fes*. The derivation of these fragments from FeSV (Fig. 2A) has been described previously (9). S_L is proximal to the 5′ end of v-*fes* and is identified by the presence of the restriction sites for *Kpn*I, *Sal*I, and *Sac*I. S_R is adjacent to S_L and covers most of the 3′ end of v-*fes*. The DNA of N-15, N-17, and N-26 phages was digested with various enzymes, and triplicate filters were prepared for hybridization to labeled pFeSV, S_L, and S_R. Figure 2B shows the results of a *Pst*I digestion of clone N-15 DNA. The pFeSV detected bands of 1.8, 1.2, and 0.4 kb and a doublet at 0.6 kb (lane 1). Codigestion with *Eco*RI did not change the pattern (data not shown). The S_L probe only detected the 1.8- and 0.6-kb bands (lane 2), and S_R detected the 1.2- and 0.6-kb bands (lane 3). The extra DNA fragments detected exclusively by pFeSV are derived from the extreme ends of v-*fes* beyond the regions of S_L and S_R. Codigestion experiments with other enzymes enabled us to place the 0.4-kb *Pst*I band detected by pFeSV adjacent to the 0.6-kb band hybridizing to S_L. Therefore, the 0.4-kb fragment contains the terminal 5′ sequences of v-*fes*. A more detailed restriction map of the three clones is shown in Fig. 3.

The internal *Eco*RI site present in the N-15 and N-17 clones corresponds to the 5′ end of the 14-kb genomic *Eco*RI fragment. N-15 overlaps completely with N-17 and contains 4 kb of additional sequences at the 3′ end. One of the internal *Eco*RI sites of the N-26 clone corresponds to the 3′ site of the genomic *Eco*RI DNA fragment. N-26 overlaps the N-15 clone completely and overlaps N-17 only partially. Whereas the complexity of v-*fes* is only 1.4 kb, the complexity estimated by restriction mapping of the human c-*fes* gene is 4.6 kb.

The human clone λ-N-26 was used in heteroduplex formation (6) with λ-FeSV since they both have cloned inserts oriented 5′-3′ with respect to the left and right phage arms. The other two human clones, λ-N-15 and λ-N-17, are oriented 3′-5′. Charon 4A, the λ-N-26 vector, and λ-WES, the FeSV vector, are homologous along the 19.8 kb of the Charon 4A left arm (λ_L). The right phage arms (λ_R) contain regions of nonhomology and, due to their unequal length, form a deletion loop. These heteroduplex features of the λ vectors were seen in all heteroduplex molecules. An example is shown in Fig. 4A. Heteroduplex comparisons of the two insert sequences consistently showed certain features. (i) The 3′-flanking sequences of the c-*fes* opposite a segment containing 3′-viral helper-derived

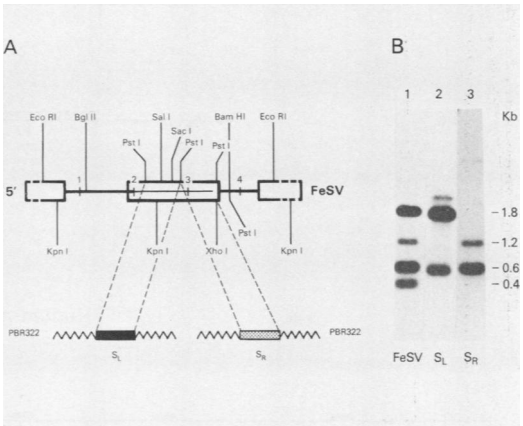

FIG. 2. (A) Restriction map of ST-FeSV (23). The central box represents the v-*fes* gene. The two *Pst*I fragments indicated were subcloned at the *Pst*I site of pBR322 as previously described (9). The stippled bar indicates the 5′ *Pst*I fragment (S_L) of the v-*fes* gene, whereas the solid bar indicates the 3′ *Pst*I fragment (S_R). (B) Autoradiogram of *Pst*I-digested N-15 DNA hybridized with labeled plasmids pFeSV (lane 1), S_L (lane 2), and S_R (lane 3).

sequences and adjacent λ-WES sequences form a large substitution loop near λ_R. (ii) A stem-and-loop structure, suggesting the presence of interspersed inverted repeat sequences, is formed within the c-*fes* sequences of this substitution loop. (iii) At the 5′ limit of this large substitution loop there is a duplex region representing the 3′ homologous sequences of c-*fes* and v-*fes*. Due to twisting of the hybrid molecule, the 5′ region of the c-*fes*-v-*fes* heteroduplex was, in most cases, uninterpretable. One molecule, however, contains a break in the 3′-flanking sequences of N-26, which allowed chain relaxation and clear display of the entire c-*fes*-v-*fes* duplex region (Fig. 4B and C). This region contains three deletion loops which, most likely, represent intervening sequences in c-*fes* toward the 5′ end of the gene. A substitution loop extending from the 3′ end of one c-*fes*-v-*fes* duplex and λ_L is formed by 1.5 kb of the 5′-flanking sequences of N-26 opposite 1.7 kb of 5′-viral helper-derived sequences of FeSV and 1.5 kb of nonhybridized λ-WES left arm. The measurements of all elements of the heteroduplex molecule are summarized in Fig. 4D. Regions containing the λ arms agree well with published maps and heteroduplex studies (7, 27, 30). The total duplex length between c-*fes* and v-*fes* is about 1.4 kb, which is in good agreement with the reported size of v-*fes* (22, 23).

A synthesis of the restriction enzyme maps of human c-*fes* clones and the heteroduplex between λ-N-26 and λ-ST-FeSV (Fig. 4) enabled us to arrive at the organization of the coding and intervening sequences of the human c-*fes* locus (Fig. 3, lower part). Starting from the 5′ end of the v-*fes*-c-*fes* heteroduplex, there is a short stretch of homology comprising sequences homologous to the 5′ terminus of v-*fes* (hatched bar) and to part of S_L (stippled bar). The first intron, which measures 0.7 kb, maps between the *Kpn*I site and the adjacent *Pst*I site. This is followed by two short stretches of homology interrupted by a second intron of 0.4 kb. These map between a *Pst*I site and a *Sac*I site. The third intron, of 0.8 kb, maps between the same *Sac*I site and a *Bam*HI site. The long stretch of homology after the third intron accounts for sequences homologous to the remaining S_L region, all of S_R (solid bar), and the 3′-terminal sequences of v-*fes* (white bar). We searched also for the presence of repetitive sequences in our

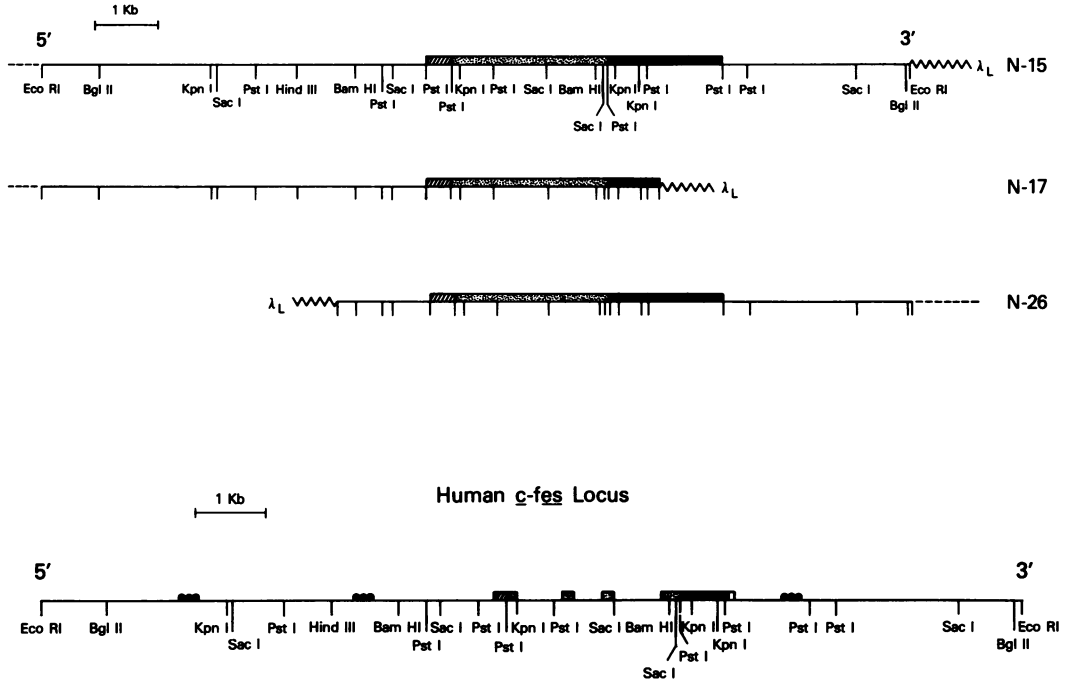

FIG. 3. Restriction map of human c-*fes* clones. The hybridizing regions are highlighted by bars. The hatched bar represents the sequences detected by the 5′-terminal sequences of v-*fes*; the stippled bar represents the region detected by S_L; and the solid bar represents the region detected by S_R. The lower part shows a schematic representation of the human c-*fes* locus constructed from restriction enzyme mapping and heteroduplex analyses. The designations of areas corresponding to different regions of v-*fes* are as described above. The black dots represent the approximate localization of sequences homologous to a member of the *Alu*I family.

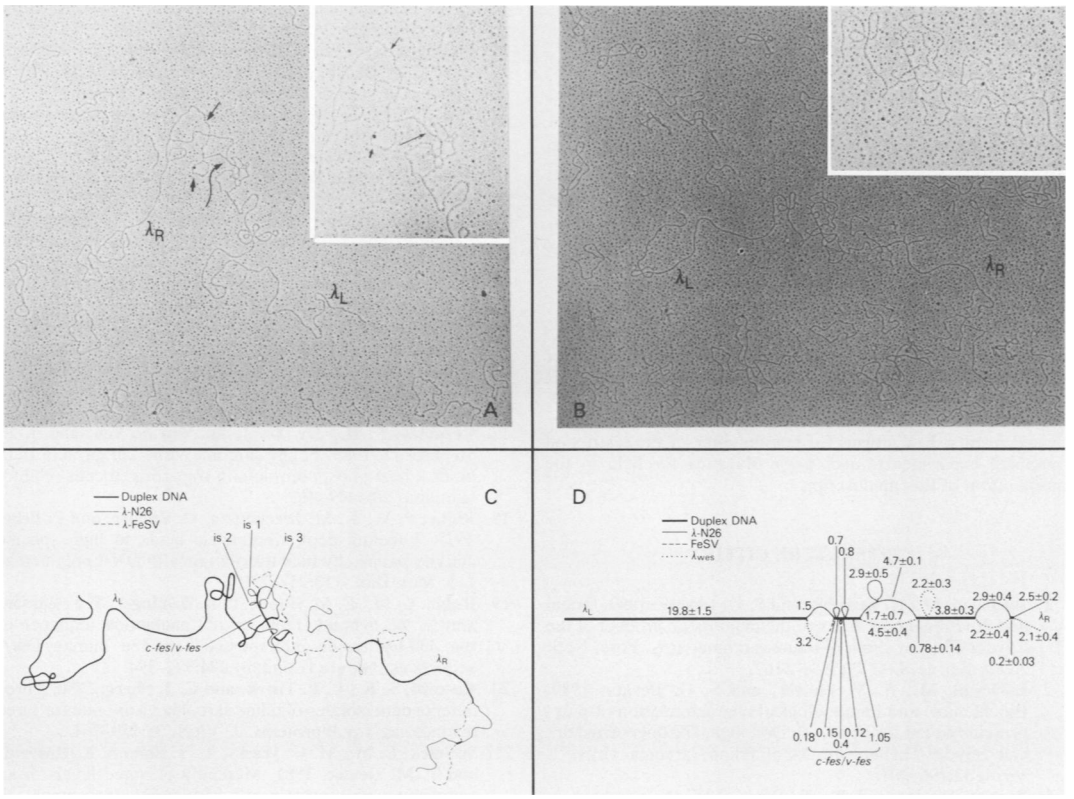

FIG. 4. Heteroduplex of λ-N-26 and λ-ST-FeSV. Heteroduplex molecules were formed in 50% formamide at 32°C after total phage particle disruption and DNA denaturation in 0.1 N NaOH. DNA was spread in 50% formamide onto a hypophase of 20% formamide, recovered on parlodian-coated copper grids, and rotary shadowed with platinum-palladium (6). (A) Typical heteroduplex molecule. The left (λ_L) and right (λ_R) arms of the vector are labeled. The substitution loop includes the 3′ end of both inserts. The 3′-cellular flanking region (short, thin arrow) contains a self-annealed sequence and a loop. The opposing DNA strand includes both FeSV and λ-WES sequences. The 3′ region of the c-*fes*/v-*fes* duplex is clearly seen (long, thin arrow). Magnification: main figure, ×26,500; insert, ×123,000. (B) Heteroduplex showing entire c-*fes*/v-*fes* region of homology and three intervening sequences. Magnification: main figure, ×26,500; insert, ×123,000. (C) Tracing of the molecule shown in (B); intervening sequences are designated is1, is2, and is3. (D) Summary of length measurements (in kilobases) from 12 heteroduplex molecules (standard deviations are given) and the molecule shown in (B) (no standard deviations are given). The region of *onc* gene homology is 1.4 kb long.

human clones by using, as a labeled DNA probe, a plasmid containing a member of the *Alu*I family, designated BLUR 8 (19). The 5′ side of the c-*fes* clones contains two regions of homology to the *Alu*I repeat. These map at 1,600 and 4,000 nucleotides upstream from the 5′ end of the c-*fes* gene (Fig. 3).

Another region of these sequences was found at 1,000 nucleotides downstream from the 3′ end of the c-*fes* gene. The N-26 clone may contain even more regions of homology in its 3′-flanking sequence. In summary, restriction enzyme mapping and heteroduplex studies confirm that we have isolated human DNA fragments homologous to the entire 1.4-kb sequence of the *onc* gene of ST-FeSV. The human c-*fes* locus, which includes both coding and intervening sequences, spans about 3.4 kb, with all three introns located in the 5′ half of the c-*fes* gene. The human c-*fes* clones are similar in complexity to the cat c-*fes* locus which we have mapped previously (9). We compared the human c-*fes* gene to another human *onc* gene, c-*sis*, recently cloned in our laboratory (5). The c-*sis* gene is homologous to the *onc* gene of the simian sarcoma virus. We found no sequence homology between the two genes in any coding, intervening, or flanking sequences, except for sequences of the *Alu*I repeat family found in both the human c-*sis* and the human c-*fes* genes.

Many attempts have been made to ascertain the role of cellular *onc* genes in cellular transformation. The oncogenic potential of a DNA sequence can be tested by transformation of

cultured cells at high efficiency. Thus far, molecularly cloned cellular analogs of viral *onc* genes, by themselves, do not seem to have transforming activity. However, the mouse cellular gene, c-*mos*, which is homologous to the *onc* gene of Moloney murine sarcoma virus, can transform fibroblasts in vitro when linked to an active viral promoter (16). We are currently testing the transforming capability of the cloned human c-*fes* sequences by linking them to a variety of viral promoters.

We thank T. Maniatis for the human library, C. J. Sherr for the cloned λ-WES-ST-FeSV, and C. W. Schmid for the BLUR 8 plasmid. We also thank William A. Haseltine, Michael D. Trus, and Joseph G. Sodroski for communication of unpublished results, J. V. Ingari for help in some of the restriction mapping experiments, and Anna Mazzuca for help in the preparation of this manuscript.

LITERATURE CITED

1. **Barbacid, M., K. Beemon, and S. G. Devare.** 1980. Origin and functional properties of the major gene product of the Snyder-Theilen strain of feline sarcoma virus. Proc. Natl. Acad. Sci. U.S.A. **77**:5158–5162.
2. **Barbacid, M., A. V. Lauver, and S. G. Devare.** 1980. Biochemical and immunological characterization of polyproteins coded for by the McDonough, Gardner-Arnstein, and Snyder-Theilen strains of feline sarcoma virus. J. Virol. **33**:196–207.
3. **Benton, W. D., and R. W. Davis.** 1977. Screening λ gt recombinant clones by hybridization to single plaques *in situ*. Science **196**:180–182.
4. **Bolivar, F., R. Rodriguez, P. J. Greene, M. C. Bethach, H. C. Hegneker, and M. W. Boyer.** 1977. Construction and characterization of new cloning vehicles. Gene **2**:95–113.
5. **Dalla Favera, R., E. P. Gelmann, R. C. Gallo, and F. Wong-Staal.** 1981. A human *onc* gene homologous to the transforming gene (v-*sis*) of simian sarcoma virus. Nature (London) **295**:31–35.
6. **Davis, R. W., M. Simon, and N. Davidson.** 1971. Electron microscope heteroduplex methods for mapping regions of base sequence homology in nucleic acids. Methods Enzymol. **21**:413–428.
7. **de Wet, J. R., D. L. Daniels, J. L. Schroeder, B. G. Williams, K. Denniston-Thompson, D. D. Moore, and F. R. Blattner.** 1980. Restriction maps for twenty-one Charon vector phages. J. Virol. **33**:401–410.
8. **Fedele, A. F., J. Even, C. F. Garon, L. Donner, and C. J. Sherr.** 1981. Recombinant bacteriophages containing the integrated transforming provirus of Gardner-Arnstein feline sarcoma virus. Proc. Natl. Acad. Sci. U.S.A. **78**:4036–4040.
9. **Franchini, G., J. Even, C. J. Sherr, and F. Wong-Staal.** 1981. *onc* sequences (v-*fes*) of Snyder-Theilen feline sarcoma virus are derived from noncontiguous regions of a cat celluler gene. Nature (London) **290**:154–157.
10. **Gardner, M. B., R. W. Rongey, P. Arnstein, J. D. Estes, P. Sarma, R. J. Huebner, and C. G. Rickard.** 1970. Experimental transmission of feline fibrosarcoma to cats and dogs. Nature (London) **226**:807–809.
11. **Maniatis, T., R. C. Hardison, E. Lacy, J. Lauer, C. O'Connell, D. Quon, G. K. Sim, and A. Efstratiadis.** 1978. The isolation of structural genes from libraries of eucaryotic DNA. Cell **15**:687–701.
12. **McDonell, M. W., M. N. Simon, and F. W. Studier.** 1977. Analysis of restriction fragments of T7 DNA and determination of molecular weights by electrophoresis in neutral and alkaline gels. J. Mol. Biol. **110**:119–146.
13. **McDonough, S. K., S. Larsen, R. S. Brodey, N. D. Stock, and W. D. Hardy, Jr.** 1971. A transmissible feline fibrosarcoma of viral origin. Cancer Res. **31**:953–956.
14. **Neil, J. C., J. F. Delamarter, and P. Vogt.** 1981. Evidence for three clones of avian sarcoma viruses: comparison of the transformation-specific proteins of PRC II, Y73, and Fujinami viruses. Proc. Natl. Acad. Sci. U.S.A. **78**:1906–1910.
15. **Opperman, H. O., A. D. Levinson, H. E. Varmus, L. Levintow, and J. M. Bishop.** 1979. Uninfected vertebrate cells contain a protein that is closely related to the product of the avian sarcoma virus transforming gene (*src*). Proc. Natl. Acad. Sci. U.S.A. **76**:1804–1808.
16. **Oskarsson, M., W. L. McClements, D. G. Blair, J. V. Maizel, and G. F. Vande Woude.** 1980. Properties of a normal mouse cell DNA sequences (*sarc*) homologous to the *src* sequence of Moloney sarcoma virus. Science **207**:1222–1224.
17. **Reynolds, F. H., Jr., W. J. M. Van de Ven, and J. R. Stephenson.** 1980. Feline sarcoma virus polyprotein P115 binds a host phosphoprotein in transformed cells. Nature (London) **286**:409–411.
18. **Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg.** 1977. Labeling deoxyribonucleic acids to high specific activity *in vitro* by nick translation with DNA polymerase I. J. Mol. Biol. **113**:237–251.
19. **Rubin, C. M., C. M. Houck, P. L. Deininger, T. Friedman, and G. W. Schmid.** 1980. Partial nucleotide sequence of the 300-nucleotide interspersed repeated human DNA sequences. Nature (London) **284**:372–374.
20. **Ruscetti, S. K., L. P. Turek, and C. J. Sherr.** 1980. Three independent isolates of feline sarcoma virus code for three distinct gag-x polyproteins. J. Virol. **35**:259–264.
21. **Scolnick, E. M., M. O. Weeks, T. Y. Shih, S. K. Ruscetti, and T. M. Dexter.** 1981. Markedly elevated levels of an endogenous *sarc* protein in a hemopoietic precursor cell line. Mol. Cell. Biol. **1**:66–74.
22. **Sherr, C. J., L. A. Fedele, L. Donner, and L. P. Turek.** 1979. Restriction endonuclease mapping of unintegrated proviral DNA of Snyder-Theilen feline sarcoma virus: localization of sarcoma-specific sequences. J. Virol. **32**:860–875.
23. **Sherr, C. J., L. A. Fedele, M. Oskarsson, J. Maizel, and G. Vande Woude.** 1980. Molecular cloning of Snyder-Theilen feline leukemia and sarcoma viruses: comparative studies of feline sarcoma virus with its natural helper virus and with Moloney murine sarcoma virus. J. Virol. **34**:200–212.
24. **Shibuya, M., T. Hanafusa, H. Hanafusa, and J. R. Stephenson.** 1980. Homology exists among the transforming sequences of avian and feline sarcoma viruses. Proc. Natl. Acad. Sci. U.S.A. **77**:6536–6540.
25. **Snyder, S. P., and G. H. Theilen.** 1969. Transmissible feline fibrosarcoma. Nature (London) **221**:1074–1075.
26. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
27. **Thomas, M., J. R. Cameron, and R. W. Davis.** 1974. Viable molecular hybrids of bacteriophage lambda and eukaryotic DNA. Proc. Natl. Acad. Sci. U.S.A. **71**:4579–4583.
28. **Van de Ven, W. J. M., F. W. Reynolds, and J. R. Stephenson.** 1980. The nonstructural components of polyproteins encoded by replication-defective mammalian transforming retroviruses are phosphorylated and have associated protein kinase activity. Virology **101**:185–197.
29. **Wahl, G., M. Stern, and G. Stark.** 1979. Efficient transfer of large DNA fragments from agarose gels to diazyobenzyloxymethyl paper and rapid hybridization by using dextran sulfate. Proc. Natl. Acad. Sci. U.S.A. **76**:3683–3687.
30. **Williams, B. G., and F. R. Blattner.** 1979. Construction and characterization of the hybrid bacteriophage lambda Charon vectors for DNA cloning. J. Virol. **29**:555-575.

31. **Witte, O. N., N. Rosenberg, and D. Baltimore.** 1979. A normal cell protein cross-reactive to the major Abelson murine leukemia virus gene product. Nature (London) **281**:396–398.

32. **Wong-Staal, F., R. Dalla Favera, G. Franchini, E. P. Gelmann, and R. C. Gallo.** 1981. Three distinct genes in human DNA related to the transforming genes of mammalian sarcoma retroviruses. Science **213**:226–228.