JOURNAL OF VIROLOGY, June 1986, p. 955–959
0022-538X/86/060955-05$02.00/0

Vol. 58, No. 3

Chromosomal Locations of Members of a Family of Novel Endogenous Human Retroviral Genomes

TOBY M. HORN,[1] KAY HUEBNER,[2] CARLO CROCE,[2] AND ROBERT CALLAHAN[1]*

Laboratory of Tumor Immunology and Biology, National Cancer Institute, Bethesda, Maryland 20892,[1] and Wistar Institute, Philadelphia, Pennsylvania 19104[2]

Received 13 November 1985/Accepted 15 January 1986

Human cellular DNA contains two distinguishable families of retroviral related sequences. One family shares extensive nucleotide sequence homology with infectious mammalian type C retroviral genomes (T. I. Bonner, C. O'Connell, and M. Cohen, Proc. Natl. Acad. Sci. USA 79:4709–4713, 1982; M. A. Martin, T. Bryan, S. Rasheed, and A. S. Khan, Proc. Natl. Acad. Sci. USA 78:4892–4896, 1981). The other family contains major regions of homology with the *pol* genes of infectious type A and B and avian type C and D retroviral genomes (R. Callahan, W. Drohan, S. Tronick, and J. Schlom, Proc. Natl. Acad. Sci. USA 79:5503–5507, 1982; I. M. Chiu, R. Callahan, S. R. Tronick, J. Schlom, and S. A. Aaronson, Science 223:364–370, 1984). Analysis of the human recombinant clone HLM-2 has shown that the *pol* gene in the latter family is located within an endogenous proviral genome (R. Callahan, I. M. Chiu, J. F. H. Wong, S. R. Tronick, B. A. Roe, S. A. Aaronson, and J. Schlom, Science 228:1208–1211, 1985). We show that the proviral genome in HLM-2 and the related recombinant clone HLM-25 are located, respectively, on human chromosomes 1 and 5. Other related proviral genomes are located on chromosomes 7, 8, 11, 14, and 17.

Retroviruses appear to have had an intimate association with avian and mammalian species throughout much of their evolution. The occasional infection of germ cells or early-developing embryos in utero has resulted in the accumulation of endogenous proviral genomes which are transmitted as stable Mendelian genes (8). Most mammalian species contain from several hundred to thousands of copies of endogenous proviral genomes (8, 18, 21, 27). Certain inbred strains of mice with a high incidence of leukemia (AKR strain) or mammary tumors (GR strain) contain endogenous proviral genomes which are responsible for the production of highly infectious retroviruses (23, 28, 30). Depending on the site of integration, endogenous proviral genomes can influence the expression of cellular genes either by altering their normal transcriptional control or by physically disrupting the genes. The dilute coat color and recessive, lethal, collagen I gene mutations in certain inbred mouse strains represent examples of this phenomenon (3, 17).

The use of low-stringency blot hybridization has recently permitted the detection of sequences related to infectious murine and primate retroviral genomes in human cellular DNA. These endogenous proviral genomes can be separated into two groups (7): (i) proviral genomes which share extensive nucleotide sequence homology with mammalian type C retroviral genomes (2, 19, 29) and (ii) a novel class of proviral genomes which represents a mosaic of sequences characteristic of different infectious retroviral genera (5). Analysis of a recombinant clone of human cellular DNA (designated HLM-2) containing one of these mosaic proviral genomes demonstrated that it possesses *env* gene sequences related to a murine type A retrovirus, long terminal repeat (LTR) sequences related to the type D squirrel monkey retrovirus, and *pol* gene sequences related to each of these as well as to type B mouse mammary tumor virus and avian type C Rous sarcoma virus (4, 5, 20). There are 50 and 1,000 copies, respectively, of the HLM-2 *pol*- and the LTR-related sequences in human cellular DNA (T. M. Horn, M. Gonda, and R. Callahan, manuscript in preparation). Although expression of endogenous human type C proviral genomes has been detected in normal and tumor tissues (26), the role which they or the HLM-2 proviral genomes play in the etiology of human neoplasia is unknown.

In humans, certain cancers occur in family clusters, which implies a genetic predisposition to the development of disease in these individuals (1, 14). Similarly, there are several hereditary illnesses which appear to be linked to defective cellular genes (22). To assess the potential involvement of endogenous retroviral genomes in these genetic disorders, we have begun to determine the distribution of HLM-2-related proviral genomes on the different human chromosomes. We have analyzed cellular DNA from somatic cell hybrids, formed between human lymphocytes or cell lines and mouse or Chinese hamster cell culture cells, that contain different complements of human chromosomes (9–11, 13, 15, 16; K. Huebner, M. Isobe, C. M. Croce, D. W. Golde, S. E. Kaufman, and J. Gasson, Science, in press). The origins of the recombinant DNA probes used in this analysis are depicted in Fig. 1. *Eco*RI-restricted human spleen DNA was electrophoretically separated on an 0.8% agarose gel. The gel was processed for Southern blotting onto Genatrans (Plasco) using 20× SSC as the transfer medium (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate). Strips were prehybridized at 55°C in Genatrans hybridization solution (5× Denhardt solution [12], 1% sodium dodecyl sulfate, 3× SSPE [20× SSPE is 3 M NaCl–0.2 M NaH_2PO_4, pH 6.5], 2.5% dextran sulfate, 0.001 M sodium phosphate, [pH 6.7]) containing 50% formamide (deionized). Filters were hybridized with 2.5×10^6 cpm of denatured ^{32}P-labeled probe per ml under relaxed conditions (20% formamide in Genatrans hybridization solution) at 37°C for 36 h and then washed to stringent conditions (0.1× SSC–0.1% sodium dodecyl sulfate, 65°C). The filters were exposed to Kodak XAR film at −70°C, and Cronex intensifier screens were used for the following lengths of time: probe A (HLM-2 *pol-env*), 4 h;

* Corresponding author.

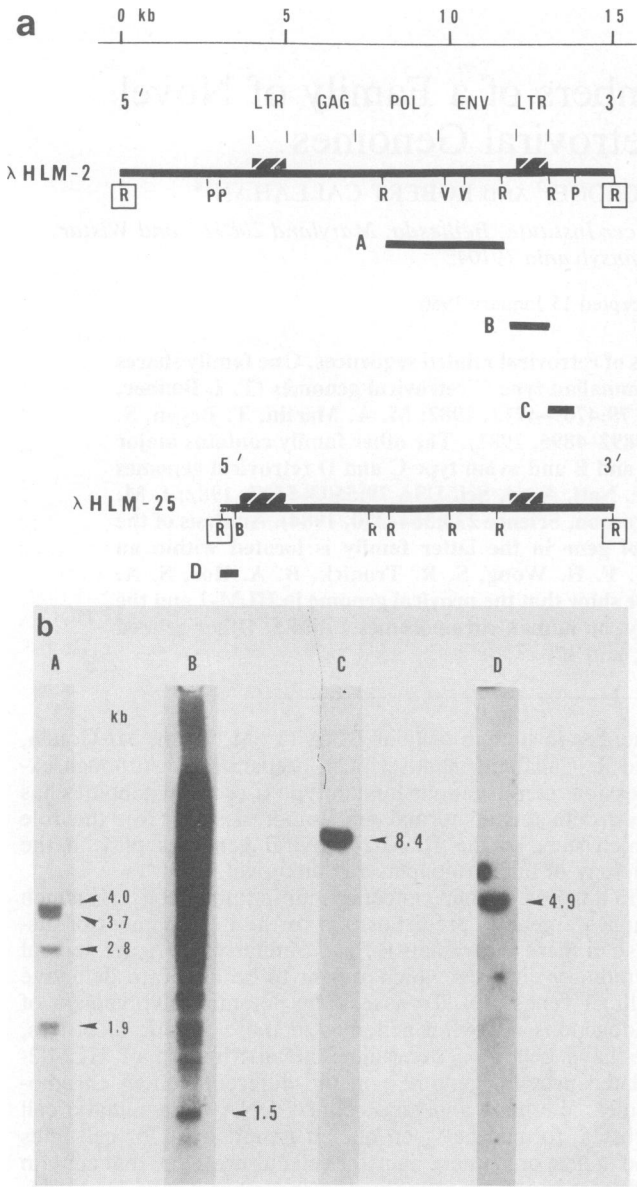

FIG. 1. Probes used to localize HLM sequences on human chromosomes. (a) Partial restriction maps of HLM-2 and HLM-25. Abbreviations: B, *Bam*HI; P, *Pst*I; R, *Eco*RI; V, *Eco*RV. The probes were 3.7-kbp *Eco*RI fragment (A), 1.5-kbp *Eco*RI fragment (B), 0.8-kbp *Eco*RI-*Pst*I fragment (C), and 0.55-kbp *Eco*RI-*Bam*HI fragment (D). (b) Hybridization patterns to human spleen cellular DNA. kb, Kilobases.

probe B (HLM-2 *env*-LTR), overnight; probe C (HLM-2, 3′ flank), overnight; probe D (HLM-25, 5′ flank), overnight.

Probes A and B correspond respectively to a 3.7-kilobase-pair (kbp) *Eco*RI fragment containing *pol-env* gene sequences and a 1.5-kbp *Eco*RI fragment containing *env*-LTR sequences. Each probe hybridizes to a complex array of fragments in *Eco*RI-restricted normal human lymphocyte DNA. Probes C and D are restriction fragments flanking, respectively, the proviral genomes in recombinant clone HLM-2 and the related proviral genome in recombinant clone HLM-25 (Horn et al., in preparation). Each of these probes detects a single *Eco*RI restriction fragment in human cellular DNA (Fig. 1).

Southern blot analysis was performed on *Eco*RI- and *Xba*I-restricted DNAs from the somatic cell hybrids and parental cells. Mouse and Chinese hamster DNA failed to react with probe A (Fig. 2A, lanes 2 and 3). Probe A detected 10 restriction fragments ranging in sizes from 2 to 25 kbp in *Eco*RI-restricted human DNA (Fig. 2A, lanes 1, 9, and 11). DNAs digested with *Xba*I yielded a pattern of 11 restriction fragments, with the major fragment of 24 kbp (Fig. 2B, lane 1). Analysis of DNAs from several hybrids suggests that most of the restriction fragments occur on a number of different chromosomes (Fig. 2). Several of the restriction fragments were found on chromosomes 7, 8, 11, and 17. Only the 4-kbp *Eco*RI and the 24-kbp *Xba*I fragments appear to have a limited chromosomal distribution. These fragments were present in hybrids containing chromosome 14. When DNA from the hybrid retaining only chromosome $14q^+$ (most of chromosome 14 and a small portion of chromosome 8) was digested with *Eco*RI or *Xba*I, the 4- and 24-kbp fragments, respectively, were present (Fig. 2A, lane 6, and 2B, lane 2). This hybrid contains the human $14q^+$ chromosome (14pter → 14q32::8q24 → 8qter) of a Burkitt lymphoma cell line carrying an 8;14 chromosome translocation (11). The 4-kbp *Eco*RI fragment in this hybrid is not derived from a piece of chromosome 8, since a hybrid containing only an intact human chromosome 8 did not contain the 4-kbp fragment (Fig. 2A, lane 7). A similar finding was observed with the 24-kbp *Xba*I fragment (Fig. 2B, lane 3). We conclude that the 4-kbp *Eco*RI and 24-kbp *Xba*I fragments are located on chromosome 14. The intensity of the hybridization signal with these two fragments in human DNA suggests that they are present in the genome in multiple copies. Whether this signal intensity reflects gene duplication on chromosome 14 or the presence of the fragments on other chromosomes not present in these somatic cell hybrids remains to be determined.

Probe B (HLM-2 proviral *env*-LTR) detects at least 30 distinct *Eco*RI fragments in human cellular DNA (Fig. 3, lane H) but does not hybridize to mouse cellular DNA (Fig. 3, lane M). Many of the LTR-related sequences in human cellular DNA are not associated with retroviral related structural genes (J. Fetherston and R. Callahan, manuscript in preparation). DNAs from a panel of hybrid cells (9–11, 13, 15, 16; Huebner et al., in press), characterized for the presence of specific human chromosomes by isozyme analysis and, in some cases, karyotypic analysis and DNA-DNA hybridization with DNA probes for genes assigned to specific chromosomes (Table 1), were analyzed for the presence of sequences homologous to those of probes B, C, and D as described above. Hybridization of probe B under stringent conditions (30% formamide, 37°C) to blots of *Eco*RI-restricted cellular DNA from hybrids containing chromosomes 7, $14q^+$, or 17 reveals a complex pattern of related fragments (Fig. 3, lanes 1, 5, and 8). Many of the fragments appear to be unique to each chromosome. Similar analysis of hybrids containing multiple human chromosomes (Fig. 3 and Table 1) suggests that the LTR-related sequences are found throughout the cellular genome. These results demonstrate that the disparity in copy number between the HLM-2 *pol* gene-related sequences (50 copies) and the LTR-related sequences (1,000 copies) in human cellular DNA is not due to the tandem duplication of the latter sequences on a few human chromosomes.

The chromosomal locations of the proviral genomes in recombinant clones HLM-2 and HLM-25 were determined

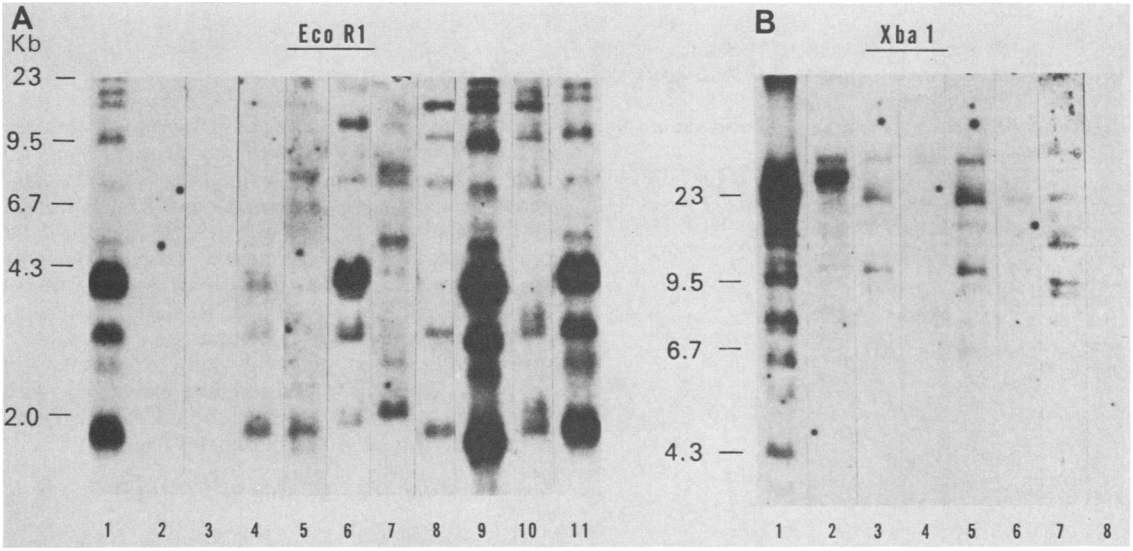

FIG. 2. Hybridization of HLM-2 *pol-env* to somatic cell hybrids containing only one or a few human chromosomes. Genomic DNAs were extracted from clones of mouse × human or Chinese hamster × human somatic cell hybrids. DNAs were digested with *Eco*RI (A) or *Xba*I (B) and then subjected to agarose gel electrophoresis and transferred to nitrocellulose. The blots were hybridized with pBR322 containing probe region A of Fig. 1 at low stringency. (A) *Eco*RI-cleaved DNA (10 μg per lane) from normal human lymphocytes (lane 1), mouse cell line (lane 2), Chinese hamster cell line (lane 3), mouse-human hybrid retaining human chromosomes 6, 7, 17q, and 21 (lane 4), mouse-human hybrid retaining chromosome 7 (lane 5), mouse-human hybrid retaining translocation chromosome $14q^+$ (14qter → 14q32::8q24 → 8qter) (lane 6), Chinese hamster-human hybrid retaining chromosome 8 (lane 7); mouse-human hybrid retaining chromosome 11 (lane 8), human cell line (lane 9), mouse-human hybrid retaining chromosome 17 (lane 10), and simian virus 40-transformed human cell line (lane 11). (B) *Xba*I digests of DNA (10 μg per lane) from human cell line (lane 1) and hybrids retaining the following human chromosomes: $14q^+$ (lane 2), chromosome 8 (lane 3), chromosome 17 (lane 4), chromosomes 6, 7, 17q, and 21 (lane 5), chromosome 7 (lane 6), chromosome 11 (lane 7), mouse cell line (lane 8).

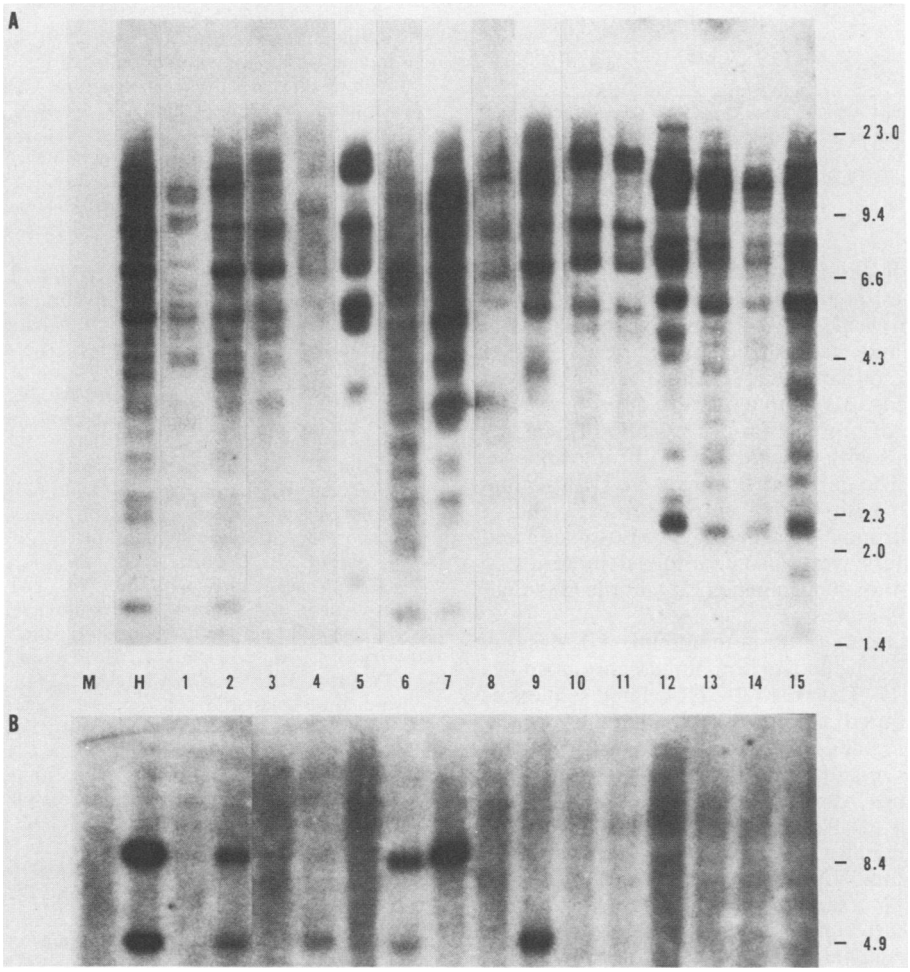

FIG. 3. Hybridization of probe B (panel A) and a mixture of probes C and D (panel B) to DNA from BW5417 mouse cells (lane M), human peripheral lymphocytes (lane H), and human × mouse somatic cell hybrids containing mixtures of human chromosomes (lanes 1–15 described in Table 1).

TABLE 1. HLM-2 and HLM-25 specific sequences in mouse × human somatic cell hybrids

Human chromosome or probe	Hybrid no.[a]														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1		+				+	+								
2				+			+								
3		+		+	+	+		+							
4		+	+			+	+		+			+	+	+	
5		+		+		+			+						
6		+				+	+		+				+	+	+
7	+	+	+				+								
8		+	+		+										
9		+					+		+	+	+				
10		+				+			+						
11						+	+								
12				+			+			+	+		+		
13		+	+						+	+	+				
14		+			+	+	+		+	+	+				+
15							+								
16															
17		+	+	+		+		+	+	+	+				
18			+			+	+					+			
19				+											
20		+	+	+			+		+				+		+
21			+								+				
22		+					+		+	+	+				
X		+	+			+	+		+			+	+	+	+
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
HLM-2 probe C	−	+	−	−	−	+	+	−	−	−	−	−	−	−	−
HLM-25 probe D	−	+	−	+	−	+	−	−	+	−	−	−	−	−	−

[a] Hybrid 5 contained a 14q$^+$ (14pter → 14q32::8q24 → 8qter) chromosome from a Burkitt lymphoma and thus contained most of chromosome 14 and a small part of the long arm of chromosome 8. Hybrid 15 retained normal chromosomes 6 and X and a t(14;20) (14q11 → 14qter::20pter → 20q1.1).

by using unique host flanking restriction fragments (probes C and D, Fig. 1). The same blot used with probe B was dehybridized by two washes in 0.4 M NaOH for 20 min at room temperature and neutralized in 0.2 M Tris (pH 7.5)–0.1× SSC–0.5% sodium dodecyl sulfate before the next hybridization cycle (Fig. 3B) with a mixture of probes C and D. The concordance of segregation of restriction fragments related to these probes with specific human chromosomes is shown in Table 1. The proviral genomes in HLM-2 and HLM-25 cosegregate with human chromosomes 1 and 5, respectively. To confirm the presence of chromosomes 1 and 5 in the hybrids, filters were also hybridized in series to probes for the human α-spectrin gene (15) and the GM-CSF gene (Huebner et al., in press).

Our results indicate that, at a minimum, HLM-2 *pol* gene-related sequences are located on human chromosomes 1, 5, 7, 8, 11, 14, and 17. The HLM-2 LTR-related sequences appear to be distributed throughout the human cellular genome. In similar studies, endogenous proviral genomes related to infectious mammalian type C retrovirus were found on chromosomes 7 and 18 (24, 25). Human chromosomes 1, 8, 11, and 14 frequently either are altered by deletion or are participants in translocations (32). Many of these genetic alterations appear to be specific for different neoplasias. The LTR element of endogenous proviral genomes contains transcription start signals as well as sequences capable of enhancing the transcription of adjacent cellular genes (31). It seems possible that endogenous proviral genomes could contribute to tumor development by altering the expression of cellular proto-oncogenes brought into close proximity by chromosomal deletion or translocation events. Alternatively, some of the human proviral genomes may act as transposable insertional mutagens. Such an event seems to have occurred in the activation of the c-*mos* proto-oncogene in mouse plasmacytomas by a retroviral-like intracisternal A-particle genome (6). Moreover, the high frequency of solitary LTR elements throughout the cellular genome makes it tempting to speculate that they, too, may be associated with some hereditary illnesses or neoplasia. The study reported here provides a starting point for the further exploration of the potential role of endogenous retroviral genomes in human disease.

We thank Josephine Romano for the preparation of some hybrid DNAs.

This research was supported in part by Public Health Service grants CA-21124, CA-10805, CA-25875, and CA-39860 from the National Institutes of Health.

LITERATURE CITED

1. **Bishop, D. T., and F. J. Gardner.** 1980. Analysis of genetic predisposition to cancer in individual pedigrees, p. 389–406. *In* J. Cairns, J. Lyon, and M. Skolnick (ed.), Cancer incidence in defined populations. Banbury report 4. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
2. **Bonner, T. I., C. O'Connell, and M. Cohen.** 1982. Cloned endogenous retroviral sequences from human DNA. Proc. Natl. Acad. Sci. USA **79**:4709–4713.
3. **Breindl, M., K. Harbers, and R. Jaenisch.** 1984. Retrovirus-induced lethal mutation in collagen I gene of mice is associated with an altered chromatin structure. Cell **38**:9–16.
4. **Callahan, R., I. M. Chiu, J. F. H. Wong, S. R. Tronick, B. A. Roe, S. A. Aaronson, and J. Schlom.** 1985. A new class of endogenous human retroviral genomes. Science **228**:1208–1211.
5. **Callahan, R., W. Drohan, S. Tronick, and J. Schlom.** 1982. Detection and cloning of human DNA sequences related to the mouse mammary tumor virus genome. Proc. Natl. Acad. Sci. USA **79**:5503–5507.
6. **Canaani, E., O. Dreagen, A. Khar, G. Rechavi, D. Rain, J. B. Cohen, and D. Givol.** 1983. Activation of the c-*mos* oncogene in a mouse plasmacytoma by insertion of an endogenous intracisternal A-particle genome. Proc. Natl. Acad. Sci. USA **80**:7118–7122.
7. **Chiu, I. M., R. Callahan, S. R. Tronick, J. Schlom, and S. A. Aaronson.** 1984. Major *pol* gene progenitors in the evolution of oncoviruses. Science **223**:364–370.
8. **Coffin, J.** 1982. Endogenous viruses, p. 1109–1204. *In* R. Weiss, N. Teich, H. Varmus, and J. Coffin (ed.), RNA tumor viruses. Cold Spring Harbor monograph series. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
9. **Cooper, C. S., M. Park, D. G. Blair, M. A. Tainsky, K. Huebner, C. M. Croce, and G. F. Vande Woude.** 1984. Molecular cloning of a new transforming gene from a chemically transformed human cell line. Nature (London) **311**:29–33.
10. **Croce, C. M., M. Isobe, A. P. Palumbo, J. Puck, J. Ming, D. Tweardy, J. Erikson, M. Davis, and G. Rovera.** 1985. Gene for α-chain of human T-cell receptor: location on chromosome 14 region involved in T-cell neoplasms. Science **227**:1044–1047.
11. **Dalla-Favera, R., M. Bregni, J. Erikson, D. Patterson, R. C. Gallo, and C. M. Croce.** 1982. Human c-myc oncogene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. Proc. Natl. Acad. Sci. USA **79**: 7824–7828.
12. **Denhardt, D. T.** 1966. A membrane filter technique for the detection of complimentary DNA. Biochem. Biophys. Res. Commun. **26**:641–646.
13. **Erikson, J., D. L. Williams, J. Finan, P. C. Nowen, and C. M. Croce.** 1985. Locus of the α-chain of the T-receptor is split by chromosome translocation in T-cell leukemias. Science **229**:784–787.

14. **Go, R. C. P., M. C. King, J. Bailey-Wilson, R. C. Elston, and H. T. Lynch.** 1983. Genetic epidemiology of breast cancer and associated cancers in high risk families. I. Segregation analysis. J. Natl. Cancer Inst. **71**:455–461.
15. **Huebner, K., A. P. Palumbo, M. Isobe, C. A. Kozak, S. Monaco, G. Rovera, C. M. Croce, and P. J. Curtis.** 1985. The α-spectrin gene is on chromosome 1 in mouse and man. Proc. Natl. Acad. Sci. USA **82**:3970–3974.
16. **Isobe, M., K. Huebner, J. Erikson, R. C. Peterson, F. J. Bollum, L. M. S. Chang, and C. M. Croce.** 1985. Chromosome localization of the gene for human deoxynucleotidyltransferase to region 10q23→10q25. Proc. Natl. Acad. Sci. USA **82**: 5836–5840.
17. **Jenkins, N. A., N. G. Copeland, B. A. Taylor, and B. K. Lee.** 1981. Dilute (d) coat color mutation of DBA/2J mice is associated with the site of integration of an ecotropic MuLV genome. Nature (London) **293**:370–374.
18. **Lueders, K. K., and E. Kuff.** 1977. Sequences associated with intracisternal A particles are reiterated in the mouse genome. Cell **12**:963–972.
19. **Martin, M. A., T. Bryan, S. Rasheed, and A. S. Khan.** 1981. Identification and cloning of endogenous retroviral sequences present in human DNA. Proc. Natl. Acad. Sci. USA **78**: 4892–4896.
20. **May, F. E. B., B. R. Westley, H. Rochefort, E. Buetti, and H. Digglemann.** 1983. Mouse mammary tumor virus related sequences are present in human DNA. Nucleic Acids Res. **11**:4127–4139.
21. **McAllister, R. M., M. Nicolson, R. Heberling, H. Charmann, N. Rice, and R. Gilden.** 1978. Infectivity of endogenous baboon type C virus related genes, p. 135–138. *In* P. Bentvelzen, J. Hilgers, and D. S. Yolan (ed.), Advances in comparative leukemia virus research. Elsevier Biomedical Press, Amsterdam.
22. **McKusick, V. A.** 1984. The human gene map. Genet. Maps **3**:417–446.
23. **Michalides, R., R. Van Nie, R. Nusse, N. E. Hynes, and B. Grover.** 1981. Mammary tumor induction loci in GR and DBAF mice contain one provirus of the mouse mammary tumor virus. Cell **23**:165–173.
24. **O'Brien, S. J., T. I. Bonner, M. Cohen, C. O'Connell, and W. G. Nash.** 1983. Mapping of an endogenous retroviral sequence to human chromosome 18. Nature (London) **303**:74–77.
25. **O'Connell, C., S. O'Brien, W. G. Nash, and M. Cohen.** 1984. ERV3, a full-length human endogenous provirus: chromosomal localization and evolutionary relationship. Virology **138**: 225–235.
26. **Rabson, R. A., P. E. Steele, C. F. Garon, and M. A. Martin.** 1983. mRNA transcripts related to full-length endogenous retroviral DNA in human cells. Nature (London) **306**:604–607.
27. **Rotman, G., A. Itin, and E. Keshet.** 1984. Solo large terminal repeats (LTR) of an endogenous retrovirus-like family (VL30) in the mouse genome. Nucleic Acids Res. **12**:2273–2282.
28. **Rowe, W. P.** 1972. Studies on genetic transmission of murine leukemia virus by AKR mice. I crosses with Fv-1^n strains of mice. J. Exp. Med. **136**:1272–1285.
29. **Steele, P. E., A. B. Rabson, T. Bryan, and M. A. Martin.** 1984. Distinctive termini characterize two families of human endogenous retroviral sequences. Science **225**:943–947.
30. **Van Nie, R., A. A. Verstracten, and J. DeMoses.** 1977. Genetic transmission of mammary tumor virus by GR mice. Int. J. Cancer **16**:922–931.
31. **Varmus, H. E.** 1982. Form and function of retroviral proviruses. Science **216**:812–820.
32. **Yunis, J. J., and A. L. Soreng.** 1984. Constitutive fragile sites and cancer. Science **226**:1199–1204.