JOURNAL OF VIROLOGY, Nov. 1979, p. 483-496
0022-538X/79/11-0483/14$02.00/0

Vol. 32, No. 2

Organization of Mouse Mammary Tumor Virus-Specific DNA Endogenous to BALB/c Mice

J. CRAIG COHEN,[1]* JOHN E. MAJORS,[2] AND HAROLD E. VARMUS[2]

Department of Microbiology and Immunology, Tulane University School of Medicine, New Orleans, Louisiana 70112,[1] and Department of Microbiology, University of California, San Francisco, California 94143[2]

Received for publication 13 June 1979

We used restriction endonucleases to prepare physical maps of the mouse mammary tumor virus (MMTV)-specific DNA endogenous to the BALB/c mouse strain. The mapping was facilitated by the DNA transfer procedure, using complementary DNAs specific for the whole and for the 3′ terminus of MMTV RNA to detect fragments containing viral sequences. The strategies used for the arrangement of fragments into physical maps included sequential digestions with two or three enzymes; preparative isolation of *Eco*RI fragments containing viral sequences; and comparisons of virus-specific fragments derived from the DNA of several mouse strains. Most of the MMTV-related DNA in the BALB/c genome is organized into two units (II and III) which strongly resemble proviruses acquired upon horizontal infection with milk-borne strains of MMTV and other retroviruses. These units contain approximately 6.0×10^6 M_r of apparently uninterrupted viral sequences, they bear redundant sequences totaling at least 700 to 800 base pairs at their termini, and the terminal redundancies include sequences derived from the 3′ end of MMTV RNA. Units II and III are closely related in that they share 12 of 14 recognition sites for endonucleases, but cellular sequences flanking units II and III are dissimilar by this criterion. The remainder of the MMTV-related DNA endogenous to BALB/c mice is found in a single subgenomic unit (unit I) with a complexity of ca. 2×10^6 M_r; the structure of this unit has not been further defined. These results support the hypothesis that endogenous proviruses have been acquired by infection of germinal tissues with MMTV. The physical maps are also useful for identifying the MMTV genomes endogenous to BALB/c mice in studies of the natural history of mammary tumorigenesis.

DNA related to the genomes of retroviruses is endogenous to normal animals from several species (1, 34, 35), but the origin, structure, and function of such virus-specific DNA are uncertain. In particular, the organization of endogenous viral DNA has not been defined because only recently have techniques become available to permit the physical mapping of sequences occurring in low numbers of copies in the eucaryotic genome.

A few copies of DNA related to the genome of mouse mammary tumor virus (MMTV) are found endogenously in inbred strains of mice, in most wild mice, and in Asian mice (9, 18, 22, 24, 36). Genetically transmitted viruses with oncogenic activity can be recovered from some inbred mouse strains (3, 20, 23), indicating that at least some endogenous viral DNA must encode a complete viral genome and be arranged as a template suitable for transcription. In some strains from which infectious virus has not been recovered, most, if not all, of the genetic information in oncogenic milk-borne viruses has been identified in endogenous MMTV DNA by molecular hybridization (18, 22), and the endogenous viral DNA is often also transcribed into RNA (19, 37). In these cases and others, however, little is known about the organization of endogenous MMTV DNA.

We have recently presented evidence favoring the hypothesis that the endogenous proviruses of MMTV were acquired by multiple, independent, and relatively recent infections of germinal tissues (5). Feral mice vary markedly in the numbers of copies of MMTV DNA and the positions of viral DNA within the host genomes, and inbred strains of mice appear to have inherited various combinations of viral DNA in a manner suggesting that independent insertions of viral DNA have segregated like stable genetic markers during inbreeding.

If endogenous proviruses were, in fact, ac-

quired by independent infections of germ lines, they might be expected to resemble structurally the proviruses found in cells infected by horizontally transmitted retroviruses. Proviruses acquired by experimental infection can be accommodated at many sites in cellular DNA (4, 14, 16, 26, 28, 31), but they have similar features: they are coextensive with unintegrated viral DNA (and therefore approximately colinear with viral RNA); they are approximately 5×10^6 to 6×10^6 M_r in size; and they bear large terminal redundancies composed of sequences derived from both the 3′ and 5′ termini of viral RNA (4, 14, 16, 26, 28).

In this report, we used restriction endonucleases to analyze the MMTV DNA endogenous to one strain of inbred mice (BALB/c) from which no infectious MMTV has been isolated in order to determine whether such DNA is organized in a fashion similar to that described for proviruses known to be introduced by infection. We overcame the complexity of analyzing multiple units of viral DNA in a single cell by isolating *Eco*RI fragments containing viral DNA before secondary digestions, and we used hybridization reagents specific for the 3′ end of MMTV RNA to search for terminal redundancies in the endogenous viral DNA. The resulting physical maps reveal that the endogenous MMTV DNA in this strain is contained within three units located at different positions in the host genome. Two of these units closely resemble MMTV proviruses acquired by infection with respect to their complexity and terminal redundancies; the third unit contains only a portion of the MMTV genome. Thus, the structure of most of the endogenous MMTV DNA in BALB/c mice conforms to the hypothesis that this DNA was introduced by infection.

MATERIALS AND METHODS

DNA extraction. The liver from a male BALB/c mouse (Simonson Laboratories, Gilroy, Calif.) was homogenized in DNA extraction buffer (0.02 M Tris-chloride [pH 8.1]–0.01 M EDTA–0.1 M NaCl), using a motor-driven, Teflon-glass Dounce homogenizer. After deproteinization with self-digested pronase and sodium dodecyl sulfate (SDS) (1 mg/ml and 1.0%, respectively) at 37°C for 12 h, DNA was extracted twice with equal volumes of phenol-chloroform (2:1). The preparation was dialyzed extensively against 5 mM Tris-chloride (pH 7.4)–0.1 mM EDTA before analysis with restriction endonucleases.

Restriction endonuclease digestion of DNA. *Eco*RI restriction endonucleases was a generous gift from P. Greene; *Pst*I was purified from *Providencia stuartii* by the procedure of Greene et al. (11); and *Bgl*II and *Hpa*I were obtained from New England Biolabs (Beverly, Mass.). *Eco*RI digestions were performed in 0.1 M Tris-chloride (pH 7.4)–0.05 M NaCl–0.005 M $MgCl_2$–0.05% Nonidet P-40. The reaction buffer for *Pst*I included 0.006 M Tris-chloride (pH 7.2), 0.006 M $MgCl_2$, 0.05 M NaCl, 0.006 M 2-mercaptoethanol, and 100 μg of gelatin per ml. *Hpa*I digestion buffer consisted of 0.01 M Tris-chloride (pH 7.5), 0.01 M $MgCl_2$, 0.006 M KCl, and 0.001 M dithiothreitol. Reactions with *Bgl*II were performed in 0.006 M KCl, 0.01 M Tris-chloride (pH 7.4), 0.01 M $MgCl_2$, and 0.001 M dithiothreitol. Reactions were monitored for completeness by adding either lambda bacteriophage DNA (for *Eco*RI and *Bgl*II) or form I of plasmid pBR313 (for *Pst*I and *Hpa*I) to a portion of the digestion mixture. A 5- to 10-fold enzyme excess was used in each reaction volume of 0.1 to 0.5 ml, and in all cases reactions were shown to be complete by analysis of marker DNAs by gel electrophoresis.

Gel electrophoresis, DNA transfer, and hybridization. After restriction endonuclease digestion, DNA samples were electrophoresed in 0.8% agarose (Seakem) gels containing 0.04 M Tris-acetate (pH 8.15), 0.02 M sodium acetate, 0.018 M NaCl, and 0.002 M EDTA (12). Bromophenol blue was added as tracking dye, and electrophoresis was performed at 40 to 70 V for a time sufficient for the dye to migrate 12 to 18 cm. Ethidium bromide (2.5 μg/ml) was used to stain gels after electrophoresis for visualization of marker DNAs (*Hin*dIII-cleaved lambda bacteriophage DNA labeled with ^{32}P by nick translation with *Escherichia coli* DNA polymerase I; 15). Preparative fractionation of DNA was performed under similar conditions, but the buffer did not include NaCl. A special electrophoresis apparatus (17) was used to provide for discontinuous electroelution from a 0.8% agarose gel 2 cm thick and 6 cm long. Seventy fractions were collected and analyzed individually by gel electrophoresis and the DNA transfer-hybridization procedure of Southern (30).

For transfer of DNA to nitrocellulose sheets, we used the method described by Southern (30). Briefly, gels were treated with denaturing solution (0.5 N NaOH–1.5 M NaCl) for 1 h and then with neutralizing buffer (0.5 M Tris-chloride [pH 7.0]–1.5 M NaCl) for 3 to 5 h. DNA transfer was accomplished by capillary action from the gel to the filter with 6× SSC (SSC = 0.15 M NaCl–0.015 M sodium citrate) for 24 to 48 h, and the filter was dried at 80°C in vacuo for 2 h.

To detect virus-specific DNA by hybridization, filters were incubated with ^{32}P-labeled MMTV complementary DNA (cDNA) synthesized by using calf thymus oligomers as primers ($cDNA_{rep}$; 16) or oligodeoxythymidylic acid [oligo(dT)] as primer ($cDNA_{3'}$) with either sucrose gradient-purified virion RNA (10 to 20 S) or oligo(dT)-cellulose-selected MMTV RNA (10 to 20 S; 33), respectively. Specificity of viral cDNA's was evaluated by hybridization to filters with restriction endonuclease-digested unintegrated MMTV DNA (29). The annealing mixture included 3× SSC, 50% formamide, 0.05 M HEPES buffer (*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid; pH 7.0), 200 μg of yeast RNA per ml, 50 μg of sheared and denatured salmon sperm DNA per ml, 1×10^6 to 2×10^6 cpm of [^{32}P]cDNA, and Denhardt buffer (0.02% each of bovine serum albumin, polyvinylpyrrolidone, and Ficoll; 6). Filters were incubated for 12 h at 41°C with annealing mix lacking cDNA and then with 1 to 2 ml of annealing

mix with cDNA for 48 to 72 h. After incubation, filters were washed for 1 h in 2× SSC at room temperature; in 0.1× SSC–0.1% SDS at 50°C for 90 min; and with multiple rinses at room temperature with 0.1× SSC–0.1% SDS and then 0.1× SSC only. After air drying, the filters were exposed to Kodak RP-Royal X-Omat film at −70°C in the presence of Dupont Cronex Lightning Plus screens (32). Reannealing filters with different cDNA's after initial analysis required that filters be placed in Seal-a-meal food container bags with annealing mix containing 1×10^6 to 2×10^6 cpm/ml and incubated for 30 min in a 68°C water bath before incubation at 41°C for 48 to 72 h.

Nomenclature. To simplify identification of BALB/c restriction endonuclease fragments containing MMTV DNA, we have named them by letter and enzyme in order of molecular weight, indicating the $M_r \times 10^{-6}$ in parentheses [e.g., *Eco*RI-A(10.0) is the largest *Eco*RI fragment containing MMTV-specific sequences and has an M_r of 10.0×10^6].

RESULTS

Restriction fragments of BALB/c DNA containing MMTV-specific sequences. By annealing virus-specific cDNA to cellular DNA, we estimated that BALB/c mice contain approximately five to six copies of viral DNA per diploid genome (4, 22). The extent of annealing of virus-specific cDNA and RNA to nucleic acids from uninfected cells indicated that most, if not all, of the sequences in the genomes of horizontally transmitted MMTV were present in the germ lines of inbred strains. Analysis of the kinetics and extent of annealing, however, does not indicate whether the mass of virus-specific DNA is organized into a few provirus-like units or widely dispersed throughout the mouse genome in many disordered, subgenomic units.

An initial appraisal of this problem can be obtained by digestion of cellular DNA with various restriction endonucleases, followed by identification of fragments containing MMTV-specific sequences with molecular hybridization. Enzymes chosen for this survey were among those previously tested for their ability to cleave viral DNA synthesized after infection by MMTV strains transmitted in the milk of GR or C3H mice (see Fig. 7A): *Eco*RI, which recognizes a single site in the DNA of MMTV(C3H) or MMTV(GR); *Pst*I, which recognizes five sites; *Bgl*II, which cleaves at two sites; and *Hpa*I, which does not cleave such DNA (29). We made the provisional assumption, later shown to be correct, that such enzymes would also be useful for generating a simple physical map of endogenous MMTV DNA, although at the outset we did not know whether the maps of endogenous DNA would bear any relationship to the maps of viral DNA synthesized after horizontal infection. We also assumed, based upon annealing experiments noted above, that cDNA's prepared from the RNA of milk-borne MMTV(C3H) would detect most or all of the sequences present within the units of endogenous MMTV DNA. Again, the results of mapping studies shown below support the validity of this assumption, though minor differences would not have been detected. The cDNA's used represented either the entire genome of MMTV(C3H) ($cDNA_{rep}$, generated by random priming with oligomers of calf thymus DNA) or the 200 to 300 nucleotides at the 3′ terminus of viral RNA [$cDNA_{3'}$, prepared by priming templates of polyadenylated viral RNA with oligo(dT)]. $cDNA_{3'}$ detects sequences that are redundant at the ends of unintegrated linear and integrated MMTV(C3H) and MMTV(GR) DNA (4, 29). Occasionally, minor bands were found on single autoradiograms but were not reproducible in multiple analyses (e.g., band at 3.4 in Fig. 4, lane C); these were not considered in the generation of the maps.

Digestion of BALB/c DNA with *Eco*RI produced five virus-specific fragments, all of which annealed with both of the MMTV-specific cDNA's (Fig. 1A and 1E; Table 1; 4, 5). This result suggested that the endogenous MMTV DNA is not widely scattered in small units throughout the mouse genome. It would accord well with the estimates of copy number from association kinetics if, for example, two proviral elements had a single internal *Eco*RI site and the third lacked an internal site. In this case, four fragments would be generated from the first two proviruses, and all four fragments would anneal with $cDNA_{3'}$, assuming the proviruses bore terminal redundancies that included sequences from the 3′ end of viral RNA; the third proviral element would be contained within the fifth fragment.

Digestion of BALB/c liver with *Pst*I (Fig. 1B and 1F; Table 1) yielded a larger number of virus-specific fragments, as would have been predicted from the restriction map of horizontally transmitted MMTV DNA. Moreover, some of these fragments have the same molecular weights (e.g., 1.1×10^6 and 0.9×10^6 M_r) as fragments derived from internal regions of MMTV(C3H) and MMTV(GR), suggesting that homologous regions might be present in endogenous proviruses. *Hpa*I, on the other hand, generated fragments smaller than would have been expected if it had failed to cleave proviruses of standard size (ca. 5.9×10^6 M_r; Fig. 1C and 1G), implying either that the endogenous MMTV DNA contained recognition sites for *Hpa*I or that some of the DNA was organized in subgenomic units. *Bgl*II also generated a modest num-

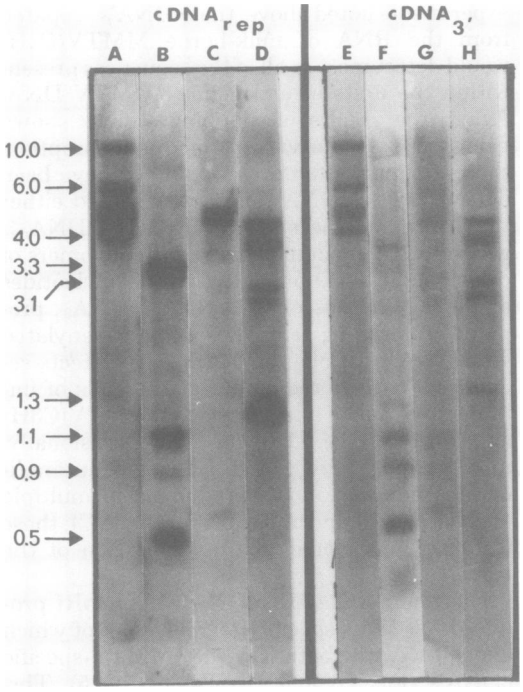

FIG. 1. *MMTV-specific fragments obtained from BALB/c liver DNA by digestion with various restriction endonucleases. Samples (10 μg) of high-molecular-weight DNA extracted from the liver of a male BALB/c mouse were digested with the designated restriction endonuclease as described in Materials and Methods. The samples were subjected to electrophoresis in an 0.8% agarose gel, and DNA fragments were transferred to nitrocellulose sheets (30), hybridized with ^{32}P-labeled MMTV-specific cDNA$_{rep}$ (lanes A–D) or cDNA$_{3'}$ (lanes E–H), and detected by autoradiography. A HindIII digest of lambda bacteriophage ^{32}P-labeled DNA (not shown) was included to obtain the $M_r \times 10^{-6}$ values indicated on the left. A and E, EcoRI digests; B and F, PstI digests; C and G, HpaI digests; D and H, BglII digests.*

ber of virus-specific fragments (Fig. 1D and 1H) without any obvious relationship to the *Bgl*II fragments expected from proviruses acquired by infection.

Although definitive deductions would have been premature at this stage of the analysis, the variations in intensity of bands and the discordances between annealing with cDNA$_{3'}$ and cDNA$_{rep}$ could clearly be helpful in ordering fragments into a restriction map. For example, a large fragment detected readily with cDNA$_{3'}$ but poorly with cDNA$_{rep}$ [e.g., *Pst*I-A(3.4) and *Pst*I-D(1.2)] would probably contain mostly host DNA linked to the end of a provirus; on the other hand, a small fragment that annealed extremely efficiently with either cDNA [e.g., *Pst*I-H(0.5)] might be derived from more than one unit of endogenous viral DNA.

Preparation of fractions of *Eco*RI-cleaved BALB/c DNA containing separated virus-specific fragments. Cleavage of BALB/c liver DNA with various enzymes indicated that relatively simple patterns of virus-specific fragments could be obtained, but the patterns seemed too complicated to permit construction of physical maps for individual units of endogenous MMTV DNA by direct sequential digestions. We therefore elected to separate on a preparative scale some of the *Eco*RI fragments of virus-specific DNA before attempting to chart more detailed maps by secondary cleavages. The *Eco*RI fragments were chosen because each fragment seemed likely to provide us with a large portion or all of an endogenous element, based upon the tentative interpretation offered in the preceding section. One obvious limitation of the choice of the *Eco*RI fragments was the inherent difficulty of separating *Eco*RI-C(5.0) and *Eco*RI-D(4.7).

Five milligrams of DNA from BALB/c livers was digested with *Eco*RI and subjected to preparative electrophoresis in a thick, short 0.8% agarose gel, using an apparatus equipped for timed collection of DNA fragments exiting from

TABLE 1. *Virus-specific fragments from restriction endonuclease digests of BALB/c DNA*

Enzyme	Fragment	$M_r \times 10^{-6}$[a]	Detected with:	
			cDNA$_{rep}$	cDNA$_{3'}$
*Eco*RI	A	10.0	+	+
	B	6.0	+	+
	C	5.0	+	+
	D	4.7	+	+
	E	4.0	+	+
*Pst*I	A	3.4	−	+
	B	3.3	+	−
	C	3.1	+	−
	D	1.2	−	+
	E	1.1	+	−
	F	1.0	+	+
	G	0.9	+	+
	H	0.5	+	+
	I	0.2	−	+
*Hpa*I	A	4.4	+	+
	B	4.2	+	+
	C	1.1	−	+
	D	0.9	−	+
	E	0.6	+	+
*Bgl*II	A	4.0	+	+
	B	3.5	+	+
	C	2.8	+	+
	D	2.5	+	+
	E	1.3	+	+
	F	1.2	+	−

[a] These values were obtained by comparison of the data shown in Fig. 1 with *Hin*dIII digests of ^{32}P-labeled lambda bacteriophage DNA electrophoresed in parallel lanes.

the end of the gel (see Materials and Methods). Seventy fractions were then analyzed by repeated electrophoresis, DNA transfer, and molecular hybridization for their content of virus-specific fragments. Four pools (subsequently called I, II, III, and IV) were made from fractions found to contain the *Eco*RI fragments previously identified in digestions of total BALB/c liver DNA (Fig. 2). Pool I contained the *Eco*RI-E(4.0) fragment, apparently uncontaminated with other virus-specific fragments; pool II contained the *Eco*RI-C(5.0) and *Eco*RI-D(4.7) fragments, contaminated with a small amount of the *Eco*RI-E(4.0) fragment; pool III contained the *Eco*RI-B(6.0) fragment, minimally contaminated with smaller fragments (not visible in the test shown in Fig. 2C); and pool IV contained

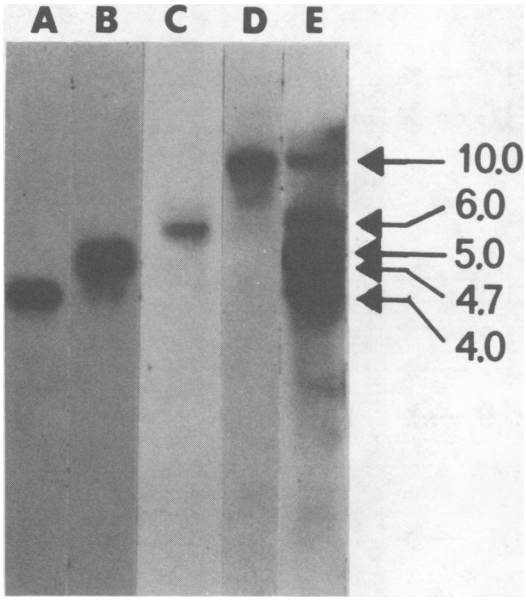

FIG. 2. *Fractionated EcoRI fragments of MMTV-specific DNA from BALB/c mouse liver DNA. Five milligrams of BALB/c liver DNA was digested with EcoRI and subjected to preparative electrophoresis through a 0.8% agarose gel on a fractionating gel electrophoresis apparatus (see Materials and Methods). Seventy fractions were obtained by collecting electroeluted material at frequent intervals, and each fraction was analyzed by gel electrophoresis, transfer to nitocellulose sheets, and hybridization with MMTV cDNA$_{rep}$. Four pools were made from multiple fractions which contained virus-specific EcoRI fragments. A sample of each pool was reanalyzed to evaluate homogeneity. $M_r \times 10^{-6}$ for each viral fragment is indicated on the right and was determined by including ^{32}P-labeled HindIII lambda DNA in a parallel lane (not shown). A, pool I; B, pool II; C, pool III; D, pool IV; E, unfractionated EcoRI-digested BALB/c liver DNA.*

the *Eco*RI-A(10.0) fragment. The amount and purity of these fragments were sufficient to proceed with secondary digestions. In some cases, however, fragments apparently derived from secondary digestion of contaminating (usually smaller) *Eco*RI fragments complicated the analysis, but these could generally be identified as contaminants without major difficulty (see, for example, fragments of 2.8×10^6 and 2.6×10^6 M_r in Fig. 3C or the fragment of 1.4×10^6 M_r in Fig. 4B).

Secondary cleavages of isolated *Eco*RI fragments of MMTV DNA. Secondary digestions of the DNA in the four pools of *Eco*RI fragments with *Pst*I, *Hpa*I, and *Bgl*II (Fig. 3–5; Table 2) permitted us to make several deductions important to the construction of restriction endonuclease maps for the MMTV DNA endogenous to BALB/c mice. The following guidelines were helpful in sorting out the complex data obtained in these experiments.

First, fragments that are novel to the sequential digestion products (i.e., not present among products of digestion with the second enzyme alone; Fig. 1 and Table 1) must be positioned at an end of the *Eco*RI fragment used for the secondary digestion. For example, *Pst*I digestion of *Eco*RI-C(5.0) and *Eco*RI-D(4.7) generated major, novel fragments of 2.8×10^6 and 2.6×10^6 M_r (Fig. 3B); the appearance of these two fragments indicated that *Eco*RI sites must be present in *Pst*I-B(3.3) and *Pst*I-C(3.1), as confirmed by the absence of these two *Pst*I fragments in all sequential digestions with *Eco*RI and *Pst*I (Fig. 3; also see below). Similar arguments apply to the fragments of 1.4×10^6, 3.0×10^6, and 1.2×10^6 M_r derived by *Hpa*I cleavage of DNA from pools I, II, and III, respectively, and to the fragments of 0.8×10^6 and approximately 0.3×10^6 M_r obtained by *Bgl*II digestion of DNA from pools I, II, and III.

There is a significant possible exception to this first rule, in that prior fractionation of the *Eco*RI fragments may permit detection of fragments that have insufficient amounts of virus-specific sequence to be observed in digests of unfractionated DNA. There is, for example, no other way to explain the relatively faint bands representing fragments of 1.5×10^6 and 1.3×10^6 M_r observed after *Hpa*I digestion of *Eco*RI-C(5.0) and *Eco*RI-D(4.7) (Fig. 4B). In fact, these fragments appear to be composed largely of cellular sequences (see below), making detection of the minor virus-specific sequences difficult in unfractionated DNA.

Second, when *Pst*I, *Hpa*I, and *Bgl*II fragments of MMTV DNA (Table 1) are also found among products of secondary digestions of *Eco*RI frag-

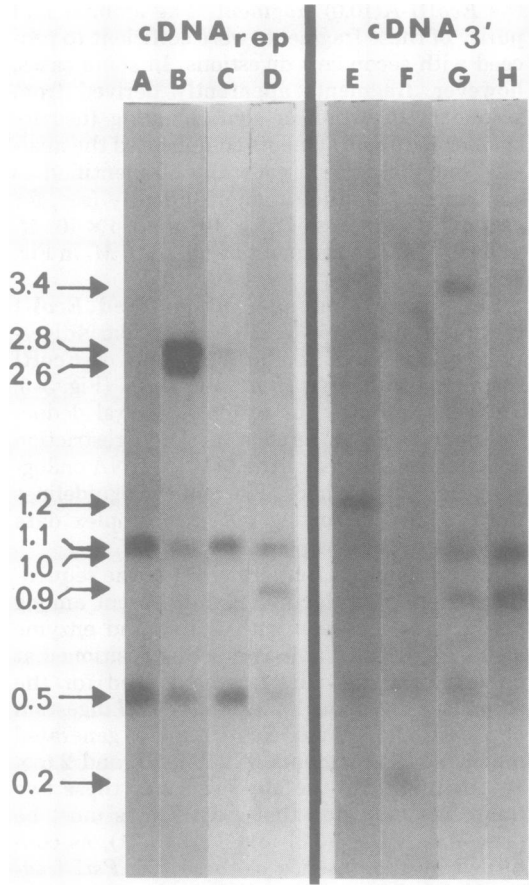

FIG. 3. *Products of PstI digestion of separated EcoRI fragments of BALB/c mouse liver DNA. Samples of each pool of EcoRI fragment (see Fig. 2) were digested with PstI, electrophoresed on 0.8% agarose gels, and transferred to nitrocellulose sheets for hybridization with either [^{32}P]cDNA$_{rep}$ (lanes A–D) or [^{32}P]cDNA$_{3'}$ (lanes E–H). HindIII-digested ^{32}P-labeled lambda bacteriophage DNA (not shown) was included as marker to obtain the $M_r \times 10^{-6}$ shown on the left. A and E, pool I; B and F, pool II; C and G, pool III; D and H, pool IV.*

ments (Table 2), these fragments are likely to be positioned internally in the relevant *Eco*RI fragment. However, it is possible to generate by these double digestions new fragments, bounded by an *Eco*RI site, which fortuitously comigrate with products of digestion with the second enzyme alone. For example, the *Pst*I-*Eco*RI fragments of 0.5×10^6 are likely to be a mixture of *Pst*I-H(0.5) and the products of *Eco*RI digestion of *Pst*I-B(3.3) and *Pst*I-C(3.1) (Fig. 3B and 3F). This inference is substantiated in a later section (see Fig. 8).

Third, fragments larger than 0.3×10^6 M_r that are detected principally or solely with cDNA$_{3'}$ [e.g., *Pst*I-A(3.4) or *Pst*I-D(1.2)] must be composed mostly of cellular DNA, since cDNA$_{rep}$ is capable of detecting fragments as small as 0.3×10^6 M_r, which are likely to be composed mostly or entirely of viral sequences (see Fig. 5A and 5C).

Fourth, the total molecular weight of virus-

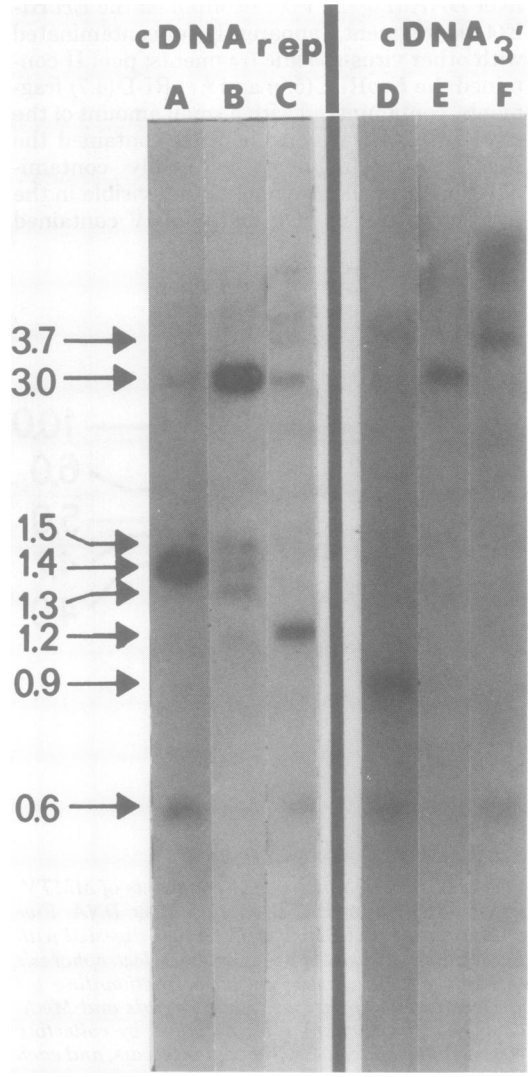

FIG. 4. *Products of HpaI digestion of separated EcoRI fragments from BALB/c liver DNA. HpaI was used to digest samples of pools of EcoRI fragments derived from BALB/c liver (Fig. 2); the products were analyzed with cDNA$_{rep}$ (lanes A–C) and cDNA$_{3'}$ (lanes D–F) as described in the legend to Fig. 1. Numbers and arrows on the left indicate the $M_r \times 10^{-6}$ of viral fragments as determined by inclusion of radiolabeled HindIII fragments of lambda DNA (not shown) in a parallel lane. A and D, pool I; B and E, pool II; C and F, pool III.*

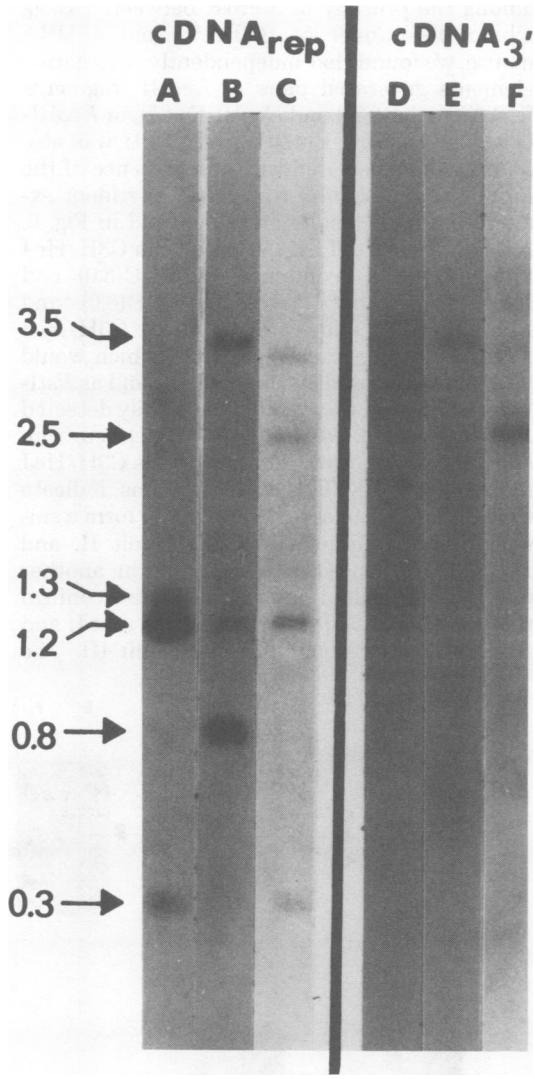

FIG. 5. *Products of BglII digestion of separated EcoRI fragments from BALB/c liver DNA. Samples of the various pools of EcoRI fragments shown in Fig. 2 were treated with BglII, subjected to electrophoresis in 0.8% agarose gels, transferred to nitrocellulose sheets, and analyzed by hybridization with either cDNA$_{rep}$ (lanes A–C) or cDNA$_{3'}$ (lanes D–F). $M_r \times 10^{-6}$ indicated on the left are based on bacteriophage lambda DNA markers included in adjacent lanes (not shown). A and D, pool I; B and E, pool II; C and F, pool III. The faint band in lane C at the position of ca. 3.2 × 10^6 M_r was not observed in repeated tests and has therefore not been considered further in the construction of maps.*

specific fragments derived from the secondary digestions provides an upper limit for the amount of viral DNA in each *Eco*RI fragment. For instance, there can be no more than 2.8 × 10^6 M_r of viral DNA in the *Eco*RI-E(4.0), based upon the sum of the fragments generated by secondary digestion with *Pst*I (Fig. 3A).

Using these principles and their corollaries as guides, it was possible to draw several conclusions about the viral DNA in the five *Eco*RI fragments and, in some cases, to establish maps for *Pst*I, *Hpa*I, and *Bgl*II sites within the viral and cellular DNA in those fragments.

First, *Eco*RI-A(10.0) contains less than 2.5 × 10^6 M_r of viral sequences, and the *Eco*RI sites appear to be located in cellular DNA outside the viral DNA (Fig. 3D and 3H). Thus, this *Eco*RI fragment harbors a subgenomic unit of viral DNA which we have termed unit I. Limited recovery of *Eco*RI-A(10.0) during fractionation procedures and the unusual structure of unit I have compromised further characterization. However, previous data concerning the segre-

TABLE 2. *Virus-specific fragments from secondary restriction endonuclease digests of pools of EcoRI fragments of BALB/c DNA*

Enzyme	Pool of *Eco*RI fragments	*Eco*RI fragments	$M_r \times 10^{-6}$[a]	Detected with: cDNA$_{rep}$	Detected with: cDNA$_{3'}$
*Pst*I	I	E(4.0)[b]	1.2	−	+
			1.1	+	−
			0.5	+	+
	II	D(4.7)	2.8	+	−
		C(5.0)	2.6	+	−
			0.5	+	+
			0.2	−	+
	III	B(6.0)	3.4	−	+
			1.1	+	−
			0.5	+	+
	IV	A(10.0)	1.0	+	+
			0.9	+	+
*Hpa*I	I	E(4.0)	1.4	+	−
			0.9	−	+
			0.6	+	+
	II	D(4.7)	3.0	+	+
		C(5.0)	1.5	+	−
			1.3	+	−
	III	B(6.0)	3.7	−	+
			1.2	+	−
			0.6	+	+
*Bgl*II	I	E(4.0)	1.3	+	+
			1.2	+	−
			0.3	+	−
	II	D(4.7)	3.5	+	+
		C(5.0)	0.8	+	−
	III	B(6.0)	2.5	+	+
			1.2	+	−
			0.3	+	−

[a] These values were obtained from *Hin*dIII-digested ^{32}P-labeled lambda bacteriophage DNA electrophoreses in a parallel lane.

[b] $M_r \times 10^{-6}$ of *Eco*RI virus-specific fragments in designated pool.

gation of endogenous MMTV proviral DNA indicated that the *Eco*RI-A(10.0) fragment segregated as an independent unit (5).

Second, each of the remaining four *Eco*RI fragments contains up to 3×10^6 to 4×10^6 M_r of viral sequences, including sequences specific to the 3′ end of MMTV RNA.

Third, the products of secondary digestions indicate that *Eco*RI-B(6.0) and *Eco*RI-E(4.0) are each derived from one of two units of viral DNA (units II and III) that also yield *Eco*RI-C(5.0) and *Eco*RI-D(4.7). As one example of evidence for this conclusion, the fragments generated by *Eco*RI cleavage of *Hpa*I-A(4.4) and *Hpa*I-B(4.2) must be the 3.0×10^6 M_r fragments in the *Hpa*I digest of *Eco*RI-C(5.0) and *Eco*RI-D(4.7) (Fig. 4B) and the 1.4×10^6 and 1.2×10^6 M_r fragments in the *Hpa*I digests of *Eco*RI-E(4.0) and *Eco*RI-B(6.0), respectively (Fig. 4A and 4C). In other words, these data locate *Eco*RI sites within *Hpa*I-A(4.4) and *Hpa*I-B(4.2). Similar arguments can be made about the location of *Eco*RI sites within *Pst*I-B(3.3) and *Pst*I-C(3.1) and within *Bgl*II-F(1.2).

Fourth, within each of the four *Eco*RI fragments from these two units, various *Hpa*I, *Pst*I, and *Bgl*II fragments unaffected by *Eco*RI digestion can be arranged (see Table 2 and discussion of Fig. 7, below). The ordering was based initially upon the relationship of fragment size and relative efficiencies of annealing with $cDNA_{rep}$ and $cDNA_{3'}$, as discussed earlier. Since the complexity of viral sequences was considerably less than the M_r of each *Eco*RI fragment (see second conclusion, above) and since each *Eco*RI cleavage appeared to affect only one *Pst*I, *Bgl*II, or *Hpa*I fragment (see third conclusion, above), we deduced that the viral sequences were approximately contiguous at one end of each *Eco*RI fragment. The proposals for overlapping sets of *Pst*I, *Bgl*II, and *Hpa*I fragments (see Fig. 7) were further validated by additional sequential digestions presented in a later section.

Fifth, *Eco*RI-B(6.0) and *Eco*RI-E(4.0) appear to be derived from closely related regions of MMTV proviruses as indicated by common *Pst*I fragments (1.1×10^6 and 0.5×10^6 M_r; Fig. 3A and 3C) and *Bgl*II fragments (0.6×10^6 M_r; Fig. 5A and 5C).

Using genetic segregation of endogenous viral DNA to facilitate mapping. To determine which of the four smaller *Eco*RI fragments came from the same proviral elements, we exploited our recent finding that endogenous MMTV DNA has been segregating like stable genetic elements during the inbreeding of contemporary mouse strains (5). By analysis of endonuclease digests of DNA from several strains among the progeny of a cross between a Bagg albino (forerunner of BALB/c) and a DBA mouse, we found that independently segregating elements generated pairs of *Eco*RI fragments [e.g., *Eco*RI-B(6.0) and *Eco*RI-D(4.7), or *Eco*RI-C(5.0) and *Eco*RI-E(4.0)]; *Pst*I-C(3.1) was also found to be associated with the presence of the former pair of *Eco*RI fragments. Pertinent examples of such results are illustrated in Fig. 6. An *Eco*RI digest of liver DNA from a C3H/HeJ mouse (lane F) contained *Eco*RI-C(5.0) and *Eco*RI-E(4.0), but lacked *Eco*RI-B(6.0) and *Eco*RI-D(4.7). The *Pst*I digest of C3H/HeJ DNA did not contain *Pst*I-C(3.1), which would be expected to produce as strong a band as *Pst*I-B(3.3); however, there was a less readily detected *Pst*I fragment of ca. 3.1×10^6 M_r derived from an endogenous provirus unique to the C3H/HeJ mouse (lane B). These observations indicate that *Eco*RI-B(6.0) and *Eco*RI-D(4.7) form a single proviral unit, which we call unit II, and *Eco*RI-C(5.0) and *Eco*RI-E(4.0) form another proviral unit, called unit III. They also confirm that the *Pst*I-B(3.3) fragment is from unit II and the *Pst*I-C(3.1) fragment is from unit III. The

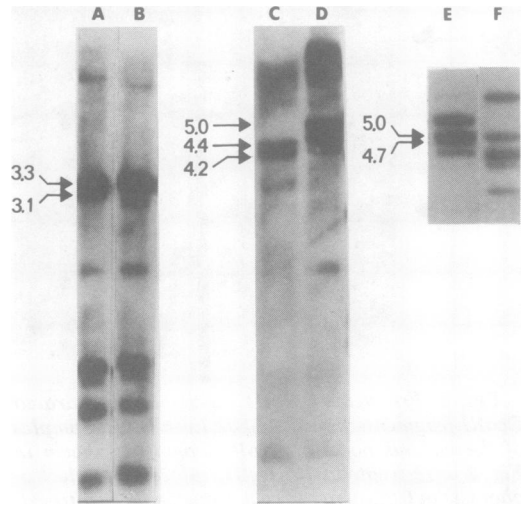

FIG. 6. *Comparison of restriction endonuclease patterns of MMTV-specific DNA in BALB/c and C3H/HeJ mouse liver DNAs. A 10-μg sample of DNA extracted from the liver of either a male BALB/c or a male C3H/HeJ mouse was digested with PstI, HpaI, or EcoRI and analyzed for MMTV-specific fragments with $cDNA_{rep}$ as described in the legend to Fig. 1. $M_r \times 10^{-6}$ of significant viral fragments are indicated to the left of each panel and were determined from HindIII-digested lambda [^{32}P]DNA included in parallel lanes (not shown). PstI digests of BALB/c (A) and C3H/HeJ (B) DNAs; HpaI digests of BALB/c (C) and C3H/HeJ (D) DNAs; EcoRI digests of BALB/c (E) and C3H/HeJ (F) DNAs.*

additional *Hpa*I site in one of these two proviruses can be mapped by comparison of *Hpa*I digestions of DNAs from these two strains (lanes C and D). The *Hpa*I-A(4.4) fragment is common to both BALB/c and C3H/HeJ DNAs, indicating that the *Hpa*I-B(4.2) fragment is derived from unit III. [The minor fragment of approximately 4.2×10^6 M_r in C3H/HeJ could represent either a sequence unique to this strain or a comigrating fragment from *Hpa*I-B(4.2) in digests of BALB/c DNA and derived from unit I.]

Testing proposed physical maps of endogenous MMTV DNA with additional sequential digestions. With the aid of the genetic arguments presented in the previous section, it is possible to organize all of the fragments obtained by secondary *Pst*I, *Bgl*II, and *Hpa*I digestions of *Eco*RI fragments B through E (Fig. 3–5) into physical maps of units II and III (Fig. 7). The only arrangement fully consonant with the rules governing interpretation of the double-digestion products indicates that these units contain approximately 6.0×10^6 M_r of apparently continuous viral sequences, with sequences from the 3′ end of viral RNA positioned near both ends of both units. In the proposed terminal redundancies, there are three sites (two *Pst*I and one *Hpa*I) that have the same relationship to each other at both ends of both units; this finding supports the evidence for redundancies from annealings with $cDNA_{3'}$, suggests that the terminal redundancies in units II and III are very similar, and shows that the size of the redundant sequence is at least 700 to 800 bases. The orientation of the maps with respect to viral RNA was based upon the evidence (discussed below in relation to Fig. 9) that $cDNA_{3'}$ detected only the 100 to 200 nucleotides adjacent to polyadenylic acid at the 3′ end of viral RNA. Thus, we placed to the right the ends of units II and III yielding fragments that were composed largely of cellular DNA and annealed efficiently with $cDNA_{3'}$. In these constructions, the two units appear closely related in that they share many restriction sites within the proposed viral sequences, but the sites in flanking regions of cellular DNA are different. Thus, units II and III differ only with respect to one additional site for *Pst*I and one for *Hpa*I in unit III, with 12 sites held in common; in flanking cellular DNA, however, all sites except a *Bgl*II site to the left of viral DNA are independently positioned in relation to the viral DNA of units II and III. (The similar positions of the *Bgl*II site may be coincidental, but further study of the homology of the cellular sequences flanking these proviruses is warranted.)

To challenge the proposed maps, we performed additional sequential digestions with various combinations of the four enzymes used in this study (Fig. 8 and 9). Several features of the maps were confirmed by these tests.

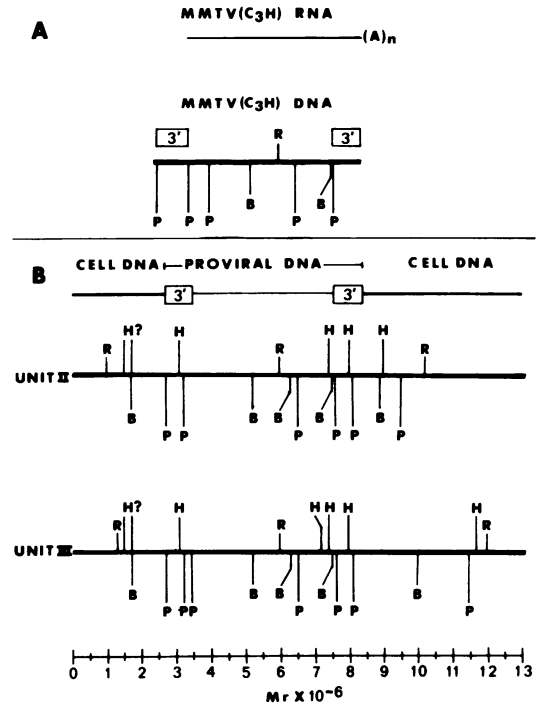

FIG. 7. *Physical maps of restriction endonuclease sites on unintegrated MMTV(C3H) DNA (panel A; 29) and on two units of endogenous MMTV DNA in the BALB/c mouse strain (panel B). Data presented in Fig. 1 through 6 and Tables 1 and 2 are summarized as linear maps of both viral and flanking cell sequences in designated units II and III of endogenous MMTV DNA. The relative orientation of unintegrated MMTV DNA to viral RNA is indicated by the RNA molecule at the top of panel A and the relative positions of viral and cellular sequences are indicated at the top of panel B; the boxes indicate terminally redundant sequences which include sequences from the 3′ end of viral RNA. The scale at the bottom indicates $M_r \times 10^{-6}$. Specific cleavage sites for EcoRI (R), PstI (P), BglII (B), and HpaI (H) are indicated for unintegrated and endogenous (units II and III) MMTV DNAs. Question marks indicate sites not mapped on a specific unit of MMTV DNA.*

First, the data in Fig. 3 suggested, but did not prove, that *Eco*RI cleaved *Pst*I-B(3.3) and *Pst*I-C(3.1) from units II and III to generate fragments of 0.5×10^6 M_r from the right-hand side of the *Eco*RI site; these fragments presumably comigrated with *Pst*I-H(0.5), which is derived from the terminally repeated structure and hence from all *Eco*RI fragments from units II and III. Since the maps predict that *Pst*I-H(0.5),

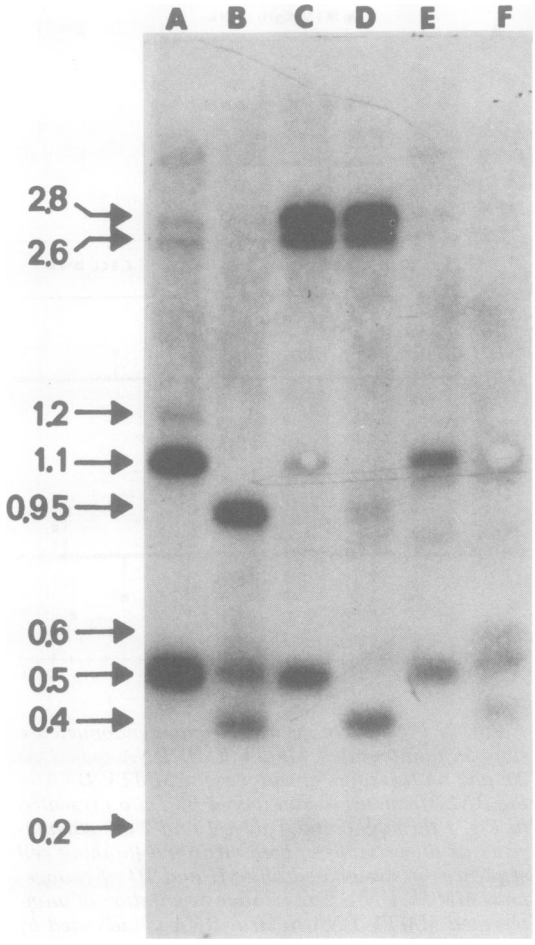

FIG. 8. *Virus-specific products of digestion of separated EcoRI fragments from BALB/c liver DNA with two restriction enzymes. Pools containing EcoRI fragments (see Fig. 2) were digested with either PstI (A, C, and E) or PstI and HpaI (B, D, and F) and analyzed as described in Fig. 1 with [^{32}P]cDNA$_{rep}$. $M_r \times 10^{-6}$ of viral fragments are indicated on the left and determined by inclusion of HindIII-digested lambda [^{32}P]DNA. DNA from pool I digested with PstI (A) or PstI and HpaI (B); DNA from pool II digested with PstI (C) or PstI and HpaI (D); DNA from pool III digested with PstI (E) or PstI and HpaI (F).*

but not the 0.5×10^6 M_r fragments from the double digestion, should be cleaved with *Hpa*I, the products of *Pst*I digestion of *Eco*RI-E(4.0) and *Eco*RI-B(6.0) were subjected to redigestion with *Hpa*I (Fig. 8, lanes A, B, E, and F). As predicted, the relative intensity of the bands in the region of 0.5×10^6 M_r was diminished after the *Hpa*I digestion, and new fragments of 0.4×10^6 M_r were generated (compare lanes A and E with lanes B and F).

Second, neither the 0.5×10^6 nor 0.4×10^6 fragments in Fig. 8, lanes B and F, annealed with cDNA$_{3'}$ (data not shown), indicating that *Hpa*I digestion removed the portion of *Pst*I-H(0.5) capable of reacting with this reagent. This result supports the proposed orientation of the maps with respect to viral RNA by demonstrating that cDNA$_{3'}$ is specific for only 100 to 200 nucleotides at the 3′ end of viral RNA.

Third, the experiments in Fig. 8 also show, as predicted by the maps, that *Pst*I-E(1.1) from unit II cleaved by *Hpa*I to produce a fragment of 0.95×10^6 M_r (lane B); *Pst*I-E(1.1) from unit III was cleaved at an additional *Hpa*I site to yield a fragment of 0.6×10^6 M_r (lane F). (The new fragments of 0.1×10^6 to 0.15×10^6 M_r to be expected after digestion with *Eco*RI, *Pst*I, and *Hpa*I were presumably too small to be detected with our methods.)

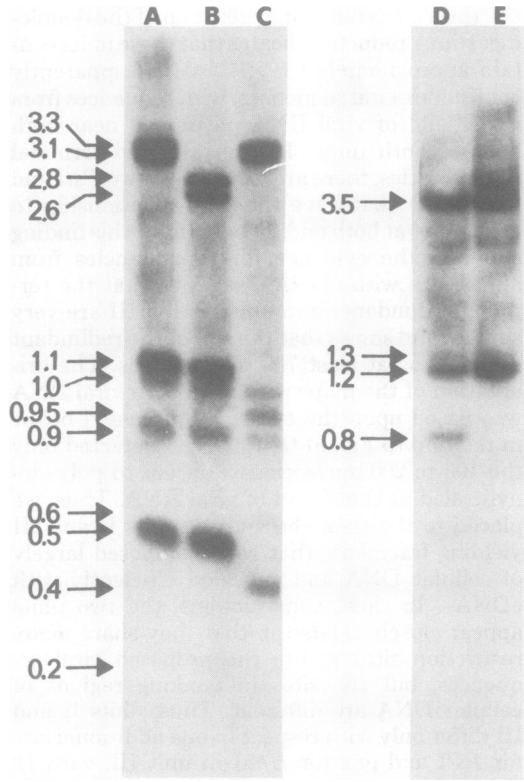

FIG. 9. *Virus-specific products of BALB/c liver DNA digested with two restriction endonucleases. Ten-microgram samples of liver DNA were digested sequentially with two restriction enzymes and analyzed with cDNA$_{rep}$ as described in the legend to Fig. 1. $M_r \times 10^{-6}$ of viral fragments were determined from HindIII-digested ^{32}P-labeled lambda bacteriophage DNA and are indicated to the left of each panel. The enzymes used were: (A) PstI; (B) PstI and EcoRI; (C) PstI and HpaI; (D) EcoRI and BglII; (E) BglII.*

Fourth, the maps predict that the products of double digestion with *Pst*I and *Eco*RI should include two fragments from the left side of each provirus, one of 2.8×10^6 or 2.6×10^6 M_r which is not susceptible to recleavage with *Hpa*I and one of 0.5×10^6 M_r [*Pst*I-H(0.5)] which bears a site for *Hpa*I. Since we propose that the ends of units II and III are terminally redundant, the 0.5×10^6 M_r fragments and their *Hpa*I cleavage products should be identical to those derived from the right side of units II and III. All of these predictions are clearly fulfilled, using *Eco*RI fragments B and E from the right sides of units II and III (Fig. 8, lanes A, B, E, and F) and *Eco*RI fragments C and D from the left sides (Fig. 8, lanes C and D).

Fifth, several of the results illustrated in Fig. 8 were confirmed and extended by redigestion of *Pst*I fragments in unfractionated BALB/c DNA with either *Eco*RI or *Hpa*I (Fig. 9, lanes A–C). Thus, *Eco*RI cleaved only *Pst*I-B(3.3) and *Pst*I-C(3.1) of those *Pst*I fragments detectable with $cDNA_{rep}$; the products of the double digestion were 2.8×10^6, 2.6×10^6, and 0.5×10^6 M_r. *Hpa*I did not cleave within *Pst*I-B(3.3) or *Pst*I-C(3.1), but generated fragments of 0.95×10^6, 0.6×10^6, and 0.4×10^6 M_r by digestion of *Pst*I-E(1.1) and *Pst*I-H(0.5).

Sixth, the maps for units II and III predict that *Bgl*II-F(1.2) represents not one but two fragments which are adjacent in the internal region of both proviruses. One of the *Bgl*II-F(1.2) fragments from each unit should be cleaved by *Eco*RI to produce fragments of 0.8×10^6 and 0.3×10^6 M_r; the other *Bgl*II-F(1.2) fragments should remain intact. These predicted results were observed after redigestion of *Bgl*II fragments of unfractionated BALB/c DNA with *Eco*RI (Fig. 9, lanes D and E). The relative intensity of the band at 1.2×10^6 M_r was decreased, and the appropriate new bands appeared. *Eco*RI also cleaved *Bgl*II-A(4.0); this fragment must be derived from unit I, but the products of the secondary digestion with *Eco*RI have not been identified.

DISCUSSION

In this report, we used restriction endonucleases to prepare physical maps of the MMTV-related DNA endogenous to the genome of BALB/c mice. The major conclusion we drew is that most of the MMTV DNA in the BALB/c genome is organized into two units (II and III) which resemble the proviruses acquired after infection with retroviruses, i.e., structures with ca. 6.0×10^6 M_r of viral DNA containing terminally redundant sequences derived from the end(s) of viral RNA. A small amount of viral DNA in BALB/c DNA is also present in a unit of subgenomic size (unit I).

This study represents the first effort to determine the structure of endogenous viral DNA by direct physical mapping procedures. The technical approach required the use of agarose gel electrophoresis, the DNA transfer procedure (30), and molecular hybridization with virus-specific, radiolabeled cDNA's. These methods allowed detection of the small number of virus-specific restriction fragments without purification of the viral DNA from the ca. 10^6-fold excess of cellular DNA. The mapping procedure included conventional sequential digestions with various combinations of two or three endonucleases (Fig. 8 and 9). To minimize the complexity of analyzing multiple units within a single cellular genome, we also used preparative fractionation (17) of the large *Eco*RI fragments bearing MMTV sequences before some of the secondary digestions with other enzymes (Fig. 2–5). The use of hybridization reagent specific for a few hundred bases at the 3′ end of MMTV RNA ($cDNA_{3'}$) was particularly helpful, since we were able to show that this sequence was repeated at both ends of units II and III, as it is in proviruses acquired by infection (4, 14, 26). Furthermore, the low complexity of $cDNA_{3'}$ allowed us to detect certain fragments that contained a small amount of viral DNA linked to cellular DNA. The most unconventional aspect of the mapping procedure involved the use of mouse genetics. We have shown elsewhere that endogenous MMTV proviruses have segregated like stable genetic elements during inbreeding (5), and we deduced that some restriction fragments must have come from the same endogenous unit of viral DNA by showing that the appearance of those fragments was coordinate in digests of DNA from various strains (Fig. 6; 5).

These strategies for physical mapping were adequate for an unambiguous ordering of restriction fragments in units II and III and their immediately flanking cell sequences. Nevertheless, more detailed analysis will be necessary to exclude the presence of small regions of DNA unrelated to MMTV (e.g., cellular intervening sequences) within the structures defined and to determine the precise length and origins of sequences repeated at the ends of the endogenous proviruses. In MMTV DNA introduced by infection, sequences from both the 3′ and 5′ ends of viral RNA are terminally repeated as a sequence of ca. 1,200 bases (29; J. E. Majors, unpublished data). We showed here that sequences from the 3′ end of MMTV RNA are present at both ends of units II and III (as based upon annealings with $cDNA_{3'}$) and that the ter-

minal redundancies are at least 700 to 800 base pairs in length (as based upon *PstI* and *HpaI* recognition sites repeated at both ends). However, we have not as yet used cDNA transcribed from the 5′ terminus of viral RNA to analyze these redundancies, nor have we determined the precise length of the redundant sequence. Presumably these and other refinements can now be approached after amplification of endogenous proviruses in procaryotic systems.

Since no endogenous MMTV strains have been recovered from BALB/c mice, we were obliged to use cDNA's prepared from the RNA of milk-borne MMTV. We then assumed that these cDNA's would detect all the restriction fragments generated from each unit that contained endogenous DNA related to the genomes of milk-borne MMTV. Within the sensitivity of the analysis, this assumption appears to have been valid, since we did not find anomalous reductions in the combined molecular weights of virus-specific fragments, using multiple enzymes to cleave endogenous DNA. It is possible, however, that the unit we have denoted as subgenomic (unit I) is a conventionally sized provirus, only part of which is homologous to milk-borne strains of MMTV. The size of the other two units (ca. 6.0×10^6 M_r) is not significantly different from the size of proviruses acquired by infection with milk-borne virus (ca. 5.9×10^6 M_r; 29), and we were able to detect all the fragments derived from units II and III with our hybridization reagent. These findings support our previous claims that most, if not all, of the genomic sequences of milk-borne viruses have homologs in the DNA of inbred mice (4, 22, 25). However, the locations of restriction sites for *HpaI*, *PstI*, and *BglII* within the endogenous proviruses differ appreciably from the location of sites for those enzymes within the DNA of milk-borne strains of MMTV (4, 29), suggesting that significant differences in base sequences are distributed along the genomes. It is possible that regions of greatest difference account for the evidence that endogenous MMTV DNA may lack sequences present in milk-borne viruses (7, 18); the results presented here, however, argue against the notion that the genomes of milk-borne viruses carry genetic elements of significant complexity totally unrepresented in endogenous proviruses.

Our finding that most endogenous MMTV DNA is organized into structures resembling conventional proviruses acquired by infection offers further support for our contention that endogenous proviruses of MMTV were acquired in the relatively recent past, long after speciation by multiple, independent infections of germinal tissues (5). Our argument was previously based upon a demonstration of the heterogeneity of wild mice with respect to the number of copies and sites of integration of endogenous MMTV DNA. A complementary analysis of endogenous MMTV DNA in inbred mice suggested that the heterogeneity of wild mice was not due to a high frequency of deletions, transpositions, or amplification of endogenous proviral DNA, since such DNA has segregated like stable elements during inbreeding. Support for this hypothesis has come from similar studies of other endogenous viruses (2, 8a, 13) and from the demonstration that the four endogenous MMTV proviruses in the A/HeJ strain mouse are located on at least three chromosomes (21). The present report shows that most of the endogenous MMTV DNA in BALB/c mice has the structure to be predicted for a provirus acquired by infection, rather than for sequences that have evolved directly from cellular genes. Of course, the present results, of themselves, do not indicate how or when the proviruses might have been introduced into the germ line.

The results presented here have an intermediate practical implication in that they provide several markers (in the form of restriction sites) with which to identify the MMTV genomes endogenous to BALB/c and to differentiate them from the genomes of other strains of MMTV, both horizontally and genetically transmitted. We have already exploited the differences in patterns of restriction fragments from endogenous MMTV DNA and milk-borne MMTV DNA in order to assay organs of BALB/c mice for infection by virus ingested via foster nursing with C3H mice (4). V. L. Morris et al. (personal communication) have used a similar approach to show that MMTV occasionally recovered from tumors of BALB/c mice is not endogenous to those mice and was presumably introduced by inadvertent infection with virus strains transmitted in the milk of females with high tumor incidence.

We chose to perform these initial mapping studies of endogenous MMTV DNA with BALB/c mice for at least three reasons: (i) the units of viral DNA appeared relatively low (4, 22); (ii) the strain is extremely susceptible to milk-borne viruses and hence widely used to study virus-induced mammary cancer (3, 20, 23); and (iii) there is very limited expression of endogenous viral DNA, even in the lactating mammary glands of multiparous animals or in mammary tumors arising in uninfected animals (8, 19, 37). For the last reason, we supposed that this strain might be the most likely to lack MMTV DNA organized into conventional pro-

viral structures. In other mouse strains that lack MMTV in the milk, there is evidence for production of infectious virus late in life (3, 20, 23) or for synthesis of viral RNA similar in size to the MMTV RNA identified in infected cells (27), thus suggesting the presence of endogenous proviruses that resemble viral DNA acquired by infection with MMTV. It would be of obvious interest to define the factors that impair the expression of the two apparently complete proviruses endogenous to BALB/c mice (units II and III).

ACKNOWLEDGMENTS

We thank J. M. Bishop, S. Hughes, and D. Robertson for useful discussions of these data and the manuscript and Y. W. Kan for use of instrumentation essential to the completion of this work.

This study was supported by Public Health Service grants CA 12705, CA 19287, and CA 21454 and by American Cancer Society grant VC 70. J.C.C. was supported by a fellowship from the National Institutes of Health (grant CA 04471) and by grant IN-133 from the American Cancer Society.

LITERATURE CITED

1. **Aaronson, S. A., and J. R. Stephenson.** 1976. Endogenous type C RNA viruses of mammalian cells. Biochim. Biophys. Acta **458**:323-354.
2. **Astrin, S. M.** 1978. Endogenous viral genes of the White Leghorn chicken: common site of residence and sites associated with specific phenotypes of viral gene expression. Proc. Natl. Acad. Sci. U.S.A. **75**:5941-5945.
3. **Bentvelzen, P.** 1974. Host-virus interactions in murine mammary carcinogenesis. Biochim. Biophys. Acta **355**: 236-259.
4. **Cohen, J. C., P. R. Shank, V. L. Morris, R. Cardiff, and H. E. Varmus.** 1979. Integration of the DNA of mouse mammary tumor virus in virus-infected normal and neoplastic tissues of the mouse. Cell **16**:333-345.
5. **Cohen, J. C., and H. E. Varmus.** 1979. Endogenous mammary tumour virus DNA varies among wild mice and segregates during inbreeding. Nature (London) **278**:418-423.
6. **Denhardt, D. T.** 1966. A membrane-filter technique for the detection of complementary DNA. Biochem. Biophys. Res. Commun. **23**:641-648.
7. **Drohan, W., R. Kettman, D. Colcher, and J. Schlom.** 1977. Isolation of the mouse mammary tumor virus sequences not transmitted as germinal provirus in the C3H and RIII mouse strains. J. Virol. **21**:986-995.
8. **Dudley, J. P., J. M. Rosen, and J. S. Butel.** 1978. Differential expression of poly(A)-adjacent sequences of mammary tumor virus RNA in murine mammary cells. Proc. Natl. Acad. Sci. U.S.A. **75**:5797-5801.

8a. **Frisby, D. P., R. A. Weiss, M. Roussel, and D. Stehelin.** 1979. The distribution of endogenous chicken retrovirus sequences in the DNA of Galliform birds does not coincide with avian phylogenetic relationships. Cell **17**: 623-634.

9. **Gillespie, D., S. Gillespie, R. C. Gallo, J. L. East, and L. Dmochowski.** 1973. Genetic origin of RD 144 and other RNA tumour viruses assayed by molecular hybridization. Nature (London) New Biol. **244**:51-54.
10. **Goulian, M.** 1968. Incorporation of oligonucleotides into DNA. Proc. Natl. Acad. Sci. U.S.A. **61**:284-291.
11. **Greene, P. J., H. L. Heyneker, F. Bolivar, R. L. Rodriguez, M. C. Betlach, A. A. Couarrubias, K. Backman, D. J. Russel, R. Tait, and H. W. Boyer.** 1978. A general method for the purification of restriction enzymes. Nucleic Acids Res. **5**:2372-2380.
12. **Helling, R. B., H. M. Goodman, and H. W. Boyer.** 1974. Analysis of endonuclease *Eco*RI fragments of DNA from lambdoid bacteriophages and other viruses by agarose gel electrophoresis. J. Virol. **14**:1235-1244.
13. **Hughes, S. H., F. Payvar, D. Spector, R. T. Schimke, H. L. Robinson, J. M. Bishop, and H. E. Varmus.** 1979. Heterogeneity of genetic loci in chickens: analysis of endogenous viral and nonviral genes by cleavage of DNA with restriction endonucleases. Cell **18**:347-359.
14. **Hughes, S. H., P. R. Shank, D. H. Spector, H.-J. Kung, J. M. Bishop, and H. E. Varmus.** 1978. Proviruses of avian sarcoma virus are terminally redundant, co-extensive with unintegrated linear DNA and integrated at many sites. Cell **15**:1397-1410.
15. **Kelly, R. B., N. R. Cozzarelli, M. P. Deutscher, I. R. Lehman, and A. Kornberg.** 1970. Enzymatic synthesis of deoxyribonucleic acid. XXXII. Replication of duplex deoxyribonucleic acid by polymerase at a single strand breach. J. Biol. Chem. **245**:39-45.
16. **Keshet, E., and H. M. Temin.** 1978. Sites of integration of reticuloendotheliosis virus DNA in chicken DNA. Proc. Natl. Acad. Sci. U.S.A. **75**:3372-3376.
17. **Leder, P., S. M. Tilghman, D. C. Tiemeier, K. I. Polsky, J. G. Seidman, M. H. Edgell, L. W. Enquist, A. Leder, and B. Norman.** 1978. The cloning of mouse globin and surrounding gene sequences in bacteriophage lambda. Cold Spring Harbor Symp. Quant. Biol. **42**:915-920.
18. **Michalides, R., and J. Schlom.** 1975. Relationship in nucleic acid sequences between mouse mammary tumor virus variants. Proc. Natl. Acad. Sci. U.S.A. **72**:4635-4639.
19. **Michalides, R., L. Van Deemter, R. Nusse, G. Röpcke, and L. Boot.** 1978. Involvement of mouse mammary tumor virus in spontaneous and hormone-induced mammary tumors in low-mammary-tumor mouse strains. J. Virol. **27**:551-559.
20. **Moore, D. H.** 1975. Mammary tumor virus, p. 131-167. *In* F. B. Becker (ed.), Cancer: a comprehensive treatise, vol. 2. Plenum Press, New York.
21. **Morris, V. L., C. Kozak, J. C. Cohen, P. R. Shank, P. Jolicoeur, F. Ruddle, and H. E. Varmus.** 1979. Endogenous mouse mammary tumor virus DNA is distributed among multiple mouse chromosomes. Virology **92**: 46-55.
22. **Morris, V. L., E. Medeiros, G. M. Ringold, J. M. Bishop, and H. E. Varmus.** 1977. Comparison of mouse mammary tumor virus-specific DNA in inbred, wild, and Asian mice, and in tumors and normal organs from inbred mice. J. Mol. Biol. **114**:73-91.
23. **Nandi, S., and C. M. McGrath.** 1973. Mammary neoplasia in mice. Adv. Cancer Res. **17**:353-414.
24. **Parks, W. P., and E. M. Scolnick.** 1973. Murine mammary tumor cell clones with varying degrees of virus expression. Virology **55**:163-173.
25. **Ringold, G. M., P. B. Blair, J. M. Bishop, and H. E. Varmus.** 1976. Nucleotide sequence homologies among mouse mammary tumor viruses. Virology **70**:550-553.
26. **Ringold, G. M., P. R. Shank, H. E. Varmus, J. Ring, and K. R. Yamamoto.** 1979. Integration and transcription of mouse mammary tumor virus DNA in rat hepatoma cells. Proc. Natl. Acad. Sci. U.S.A. **76**:665-669.
27. **Robertson, D. L., and H. E. Varmus.** 1979. Structural analysis of the intracellular RNAs of murine mammary tumor virus. J. Virol. **30**:576-589.
28. **Sabran, J. L., T. W. Hsu, C. Yeater, A. Kaji, W. S. Mason, and J. M. Taylor.** 1979. Analysis of integrated avian RNA tumor virus DNA in transformed chicken, duck, and quail fibroblasts. J. Virol. **29**:170-178.
29. **Shank, P. R., J. C. Cohen, H. E. Varmus, K. R.**

Yamamoto, and G. M. Ringold. 1978. Mapping of linear and circular forms of mouse mammary tumor virus DNA with restriction endonucleases: evidence for a large specific deletion occurring at high frequency during circularization. Proc. Natl. Acad. Sci. U.S.A. **75:** 2112-2116.

30. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **38:**503-517.
31. **Steffen, D., and R. A. Weinberg.** 1978. The integrated genome of murine leukemia virus. Cell **15:**1003-1010.
32. **Swanstrom, R., and P. R. Shank.** 1978. X-ray intensifying screens greatly enhance the detection by autoradiography of the radioisotope ^{32}P and ^{125}I. Anal. Biochem. **86:**184-192.
33. **Tal, J., H.-J. Kung, H. E. Varmus, and J. M. Bishop.** 1977. Characterization of DNA complementary to nucleotide sequences adjacent to poly (A) at the 3' terminus of the avian sarcoma virus genome. Virology **79:** 183-197.
34. **Todaro, G. J., R. E. Benveniste, R. Callahan, M. M. Leiber, and C. J. Sheer.** 1975. Endogenous primate and feline type C viruses. Cold Spring Harbor Symp. Quant. Biol. **39:**1159-1168.
35. **Todaro, G. J., and R. J. Huebner.** 1972. The viral oncogene hypothesis: new evidence. Proc. Natl. Acad. Sci. U.S.A. **96:**1009-1015.
36. **Varmus, H. E., J. M. Bishop, R. C. Nowinski, and N. H. Sarkar.** 1972. Mammary tumor virus specific nucleotide sequences in mouse DNA. Nature (London) New Biol. **238:**189-191.
37. **Varmus, H. E., N. Quintrell, E. Medeiros, J. M. Bishop, R. C. Nowinski, and N. H. Sarkar.** 1973. Transcription of mouse mammary tumor virus genes in tissues from high and low tumor incidence mouse strains. J. Mol. Biol. **79:**663-679.