Murine leukemia virus sequences are encoded in the murine major histocompatibility complex
(viral integration/class I gene/TL region/cosmid library/gene expression)

DANIEL MERUELO*, RUTH KORNREICH*, ANTHONY ROSSOMANDO*, CHRISTINE PAMPENO*, ANDREW L. MELLOR†, ELIZABETH H. WEISS†, RICHARD A. FLAVELL†, AND ANGEL PELLICER*

*Irvington House Institute and Department of Pathology, New York University Medical Center, 550 First Avenue, New York, NY 10016; and †Biogen Research Corporation, 14 Cambridge Center, Cambridge, MA 02142

Communicated by Michael Heidelberger, November 30, 1983

ABSTRACT The studies reported here localize murine leukemia virus viral sequences to the TL region of the major histocompatibility complex, H-2. We examined a battery of 38 cosmids, isolated from two large genomic libraries constructed from C57BL/10 spleen DNA, that define 25 class I gene sequences. The viral probes used hybridized with only four cosmids, containing overlapping mouse sequences, that define four class I gene-related sequences in a region of 90 kilobases of DNA. The data show that two distinct viral envelope sequences are contained in the cluster. One of these sequences is situated with its 3' end next to the 3' end of a class I sequence. The other sequence, which does not contain the entire viral envelope, is proximal to the 3' end of a different class I sequence. Hybridization of the viral probes with the H-2 cosmid clones does not appear to be due to homology between viral and H-2 sequences. Rather, the viral sequences detected appear to be linked to or inserted amid class I genes. These findings may be significant in understanding molecular mechanisms involved in the generation of H-2 class I gene diversity.

The murine major histocompatibility complex (MHC), H-2, encodes the most polymorphic group of genes known in eukaryotes (10–50 alleles for each of its members) (1). The complex is comprised of the classic transplantation (class I) genes, immune response (class II) loci, and complement-related (class III) genes (2–4). Several features of the complex have hitherto drawn attention: extensive polymorphism (1), gene duplication (5, 6), gene dispersion (7), gene conversion (8), extraordinary mutation rates (9), and the association of class I and class II genes with susceptibility or resistance to a myriad of diseases (10, 11). A recurrent theme of many hypotheses to explain these characteristics has been the potential role of viruses (7, 12, 13). The studies reported here lend credence to this concept by localizing murine leukemia virus sequences within the H-2 complex.

The role of viruses in the above-mentioned properties of MHC genes became interesting to us when it was shown that transformation of thymus cells by an RNA type C leukemia virus, radiation-induced leukemia virus (RadLV), led to shutdown of synthesis and expression of antigens encoded by class I genes (14). Also, it was later shown that this effect of RadLV on MHC antigen expression probably plays a critical role in the escape of RadLV-transformed cells from the surveillance mechanisms of the immune system (15).

After several hypotheses had been tested (16, 17), it was concluded that viral integration may play a role in the observed alterations of antigen expression and that one of the critical integration sites of the virus may be adjacent to a MHC gene or a gene that regulates the expression of MHC antigens. The present studies locate viral sequences adjacent to class I genes in normal mice but do not address the potential location(s) of RadLV sequences within the MHC. However, they add further weight to the hypothesis that viruses such as RadLV may integrate within the MHC cluster of genes. They also set the stage for understanding the distinction between potential RadLV integration sites (if found) and endogenous viral sequences previously integrated next to class I genes in normal cells.

MATERIALS AND METHODS

Cloned Viral Probes. A p15E-specific probe was subcloned from a cloned 8.8-kilobase AKR-ecotropic (pAKR-Eco) murine leukemia virus DNA sequence (18), and pAKR-Eco was kindly provided by Malcolm Martin (National Institutes of Health).

DNA Procedures. DNA was digested with the appropriate restriction endonuclease (New England BioLabs or Bethesda Research Laboratories) as recommended by the supplier. Gel electrophoresis and hybridization were done by the method of Southern (19). Only stringent conditions were used. After transfer of DNA to the filters was complete, they were washed with 6× standard saline citrate (NaCl/Cit; 1× NaCl/Cit is 0.15 M NaCl/0.015 M Na citrate, pH 7), air dried, baked for 5 hr at 80°C under vacuum, and then each was submerged in 100 ml of prehybridization fluid (6× NaCl/Cit/0.02% bovine serum albumin/0.02% Ficoll and 0.02% polyvinylpyrrolidone, prewarmed to 65°C). After incubation for 5 hr in a water bath at 65°C with shaking, the filters were removed from the blocking solutions, air dried, and placed in heat-sealable plastic bags. Then, the hybridization solution [2× NaCl/Cit/0.01 M EDTA, 50-labeled denatured probe DNA (labeled to at least 2 × 10⁸ cpm/μg), and denatured calf thymus DNA at 5 μg/ml] was added and hybridization was continued overnight at 65°C. The filters were then washed three times for 20 min each in a tray containing 2× NaCl/Cit/0.05% NaDodSO₄/20 mM NaH₂PO₄/0.06% Na₂H₂PO₄ at 65°C, once with 1× NaCl/Cit, once with 0.4× NaCl/Cit, and finally numerous times with 0.1× NaCl/Cit until the Cerenkov count was <100 cpm/5 μl.

RESULTS

Detection of Viral Sequence in H-2 Cosmids. Because H-2-linked resistance to RadLV-induced thymomas involves viral interaction with H-2 at the DNA level (14–17) we tested the possibility that viral DNA sequences may be located within the MHC. Preliminary experiments suggested that viral sequences are integrated in the region of chromosome 17 containing the H-2 genes in RadLV-induced tumors as well as in tissues from normal mice. Digestion of DNA from thymocytes of several inbred mouse strains with BamHI, Xba I,

Abbreviations: MHC, major histocompatibility complex; RadLV, radiation-induced leukemia virus; NaCl/Cit, standard saline citrate; kb, kilobase(s).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.
or Kpn I produced fragments that appeared to hybridize with H-2 and viral probes. However, because many genomic fragments hybridize with sequences that are detected by the broadly crossreactive H-2 probes we used (20), the significance of hybridization of one of these fragments with a viral probe was unclear and these results might be merely coincidental. Greater certainty would result if the fragment containing the H-2 sequences and cross-hybridizing with the viral probe could be isolated and cloned. For this reason, we examined 38 cosmids isolated from two large genomic libraries constructed from C57BL/10 (H-2b haplotype) spleen DNA (unpublished data; L. Golden, K. Fahner, A. L. Mellor, J. J. Devlin, H. Bullman, H. Tiddens, and H. Bud, personal communication) for viral sequences. (Taken together, these cosmids defined all class I genes in the C57BL/10 H-2 complex.) In the experiments reported below, all inserts used as probes were removed from their vectors to eliminate hybridization to nongenomic sequences in the cosmids.

The cosmids were digested with BamH I, which does not cut within the cosmid vectors used [pOPF I and pTM (21, 22)], and were hybridized under stringent conditions with an ecotropic env viral probe, encompassing sequences encoding gp70. The virus probe only hybridized with four cosmids, H6, B1.19, H43, and B1.15 (Fig. 1).

Localization of the Viral Sequences Within the MHC Complex and With Respect to the Specific Class I Genes. The four cosmids studied contain overlapping mouse sequences and belong to a group of 10 overlapping cosmid clones that define five class I gene-related sequences in a region of 90 kilobases (kb) of DNA (Fig. 2). This cluster has been mapped to the TL region of the MHC by restriction enzyme polymorphism (unpublished data; K. Fahner, A. L. Mellor, J. J. Devlin, H. Bullman, H. Tiddens, and H. Bud, personal communication).

Since hybridization with cosmids H43 and H6 shows two bands compared with only a single band in B1.19 (Figs. 1 and 3), one of the two regions hybridizing with our viral probe must be in the left segment of cosmids H6 and H43, which is missing in cosmid B1.19 (Fig. 2). By the same reasoning, the second viral hybridizing fragment, which is the only one present in cosmid B1.19, must be located in the region of D7 and common to H6, H43, and B1.19, and been positioned at the rightmost side of H6 and H43.

The leftmost viral hybridizing fragment shown in Fig. 2 was positioned on the basis of two observations. First, H6 and H43 show two bands on hybridization with the env probe, while B1.19 yields only one band (Figs. 1 and 3). Therefore, the hybridizing region must be to the left of B1.19. Second, one of the two BamH I fragments in cosmids H43 and H6 hybridizing with the viral pAKR-Eco probe does not hybridize with H-2 probes, pH-2II or pH-2III (20) (Fig. 3). This fragment cannot be common to B1.19 because the viral hybridizing fragment of the latter cosmid hybridizes with the H-2 probes (Fig. 3). The H43 and H6 fragments in question must therefore be outside the H-2 domains and, accordingly (see Fig. 2), can only correspond to the leftmost BamH I fragment. The location of the rightmost viral fragment is derived from the following observations. The fragment is not present in cosmid B1.15. In B1.19 only one pAKR-Eco-hybridizing fragment is detected and this fragment appears identical in molecular weight to one of the BamH I fragments of cosmid H6 (Figs. 1 and 3). In addition, the common viral fragment in cosmids B1.19 and H6 hybridizes with both pH-2II and pH-2III. The equivalent fragment in cosmid H43 is smaller by about 1.5 kb because this is the right terminal BamH I fragment of H43, which is truncated. This fragment of H43 hybridizes with pH-2II but not with pH-2III, which is consistent with the cosmid map because H43 does not contain the 5′ region of the third rightmost H-2 gene. The location of the p3SE-hybridizing fragment was determined as follows, using a p3SE probe derived by digestion with Pst I and Xba I from λAKR (23). This viral fragment is not present in B1.19 and must therefore be in the leftmost portion of cosmids H6 and H43. The p3SE fragment does not hybridize with either H-2 probe (pH-2II or pH-2III; Fig. 3). The 4.2-kb BamH I p3SE-hybridizing fragments in both cosmids are of identical molecular weight (Fig. 3) and too small to contain cosmid vector DNA (because BamH I does not cut the 8-kb vector) and as expected does not hybridize with pBR322 (Fig. 3). Thus p3SE sequences must lie between those hybridizing with pAKR-Eco and the boxes containing H-2-hybridizing sequences. B1.15 is also expected to contain p3SE sequences, but it was not hybridized with the probe and therefore a p3SE location is not shown in Fig. 2. The different nomenclature used for the various cosmid clones (H vs. B) reflects the fact they were derived from independent cosmid libraries.

Cosmids H6, H43, and B1.19 were also hybridized under stringent conditions (Fig. 1) with a probe derived from the

![Fig. 1. Viral envelope sequences are found within H-2 cosmids clones. Southern blots (19) of BamH I-digested DNA of 19 H-2 cosmids are shown. (In addition to the cosmids shown, two other filters containing an additional 19 H-2 cosmids were hybridized but no other viral hybridizing sequences were found.) Each cosmid DNA (1 μg) was digested overnight with a 10-fold excess of BamH I and electrophoresed on 0.8% agarose gels (Bethesda Research Laboratories) containing ethidium bromide. After electrophoresis, the gels were visualized under UV light and photographed (B). The DNA was then denatured and transferred to nitrocellulose as described (19). The probe used was a pAKR-Eco 2.7-kb Sal I/BamH I fragment within the envelope (env) gene (4.3-7.0 kb) of complete infectious AKR provirus cloned probe, λAKR 623 (23, 24). Nick-translation was done using New England Nuclear kit NEK 004B. (A) Results after autoradiography. Molecular weight markers run on the same gel were HindIII/EcoR I-digested phage λ DNA.
p15E region of 1AKR (23, 24), which is encoded on the 3' side of the viral envelope region coding for gp70. The p15E probe hybridizes with sequences in cosmids H6 and H43 but not B1.19 (Fig. 3). The absence of p15E hybridizing sequences in B1.19 suggests that a defective genome is encoded in this class I cosmid.

DISCUSSION

An important question raised by the above findings is whether hybridization of the viral probes with the H-2 cosmid clones occurs because homology exists between the viral and H-2 segments or because both types of sequences are linked in the same cosmid DNA. This question is even more relevant in view of the recent paper by Clarke et al. (25), published while this manuscript was being reviewed. Those authors have suggested weak homology between HLA class I and human T-leukemia virus sequences, based on hybridization experiments. However, these data are difficult to explain in view of the lack of homology at the nucleotide sequence level between human T-leukemia virus (26) and HLA-B7 (ref. 27; S. Weissman, personal communication) as determined by computer analysis. Our data do not support a simple homology interpretation for several reasons. First, as pointed out above, for at least one of the viral sequences, some cosmid BamHI restriction fragments hybridize to the H-2 probe but not to the viral probe (Fig. 3) and vice versa. Similar results are observed after digestion by EcoRI, HindIII, and Pst I (data not shown). Further analysis of the second viral sequence by restriction analysis also shows clear separation between the viral and H-2 sequences (data not shown). Second, the probes themselves do not cross-hybridize; i.e., none of the viral probes hybridize to the H-2 probes or vice versa under the conditions used. Furthermore, the data in Fig. 1 indicate that the majority of H-2 genes of the b haplotype do not hybridize to the viral probes. Finally, a computer comparison of H-2 nucleotide sequences for K^b, D^t, and L^d genes (28-31) with the ectopic envelope sequence (32) indicates no obvious homology and certainly insufficient homology to obtain the degree of hybridization observed in our experiments. For these reasons, we believe that the sequences detected are linked to or inserted amidst class I genes.

A potential caveat to the findings reported here is that the particular mouse used to derive the cosmid libraries may coincidentally have had an inserted viral genome at H-2. However, as stated above, the cosmid clones used were derived from independently created cosmid libraries and results with overlapping cosmid clones from the two libraries are internally consistent.

The findings reported herein may have significant implications for the generation of H-2 class I gene diversity. For example, it has been suggested that the high mutation rate associated with class I genes may result from the incorporation of viral genomes into germinal cells (7, 12) since the spontaneous mutation rate is significantly influenced by environmental conditions during breeding (7, 9). Similarly, the mechanism by which H-2 mutations arise appears to result from multiple and simultaneous nucleotide substitutions and has been likened to gene conversion (8, 33) as it occurs in yeast in which segments of one gene are replaced by segments of another gene. Such a mechanism appears to be likely since potential donors of the mutant sequence have been detected [at least a centimorgan away from the acceptor sequences (unpublished results)]. Several mechanisms might be responsible for this type of gene conversion. One would involve transposon-like structures such as Alu-like repeat elements or murine leukemia viruses. These sequences could play a role in insertion or deletion of nonhomologous sequences. A similar mechanism may account for the extensive gene duplication associated with H-2 genes.

The potential role for viruses in these phenomena is more credible in view of the findings reported herein. Further support for such hypotheses comes from our recent studies on the effects of viruses on H-2 gene expression and organization. These studies demonstrate associations between RadLV-induced transformation, changes in H-2 antigen expression, and hypermethylation and rearrangements of H-2 DNA (35). Hence, the results of this study suggest that the presence of a certain virus or viruses in H-2 can influence the expression of MHC antigens. This has important implications for the study of viruses and their role in the expression of MHC antigens.
of endogenous viral sequences is in the vicinity of lymphocyte differentiation and histocompatibility loci (37).

The mapping of viral sequences to the TL region of the MHC complex is significant in view of the postulate that the TL region may encode an integrated viral genome (38). Although this postulate was later withdrawn (39), the data reported here support the original postulate. It remains to be seen whether such viral sequences play a role in the conversion of TL" strains to TL" during leukemogenesis, as originally proposed (38).

Although viral sequences have been found so far in only one H-2 region, represented by cosmid clones H6, H43, B1.19, and B1.15, detailed analysis of other H-2 cosmid clusters with additional viral probes should indicate whether viral sequences are associated with more than one group of H-2 genes. Our findings open the way for a molecular approach to studying the potential role of viruses in the generation of histocompatibility gene diversity.

This work was supported by Grants CA22247 and CA31346 from the National Institutes of Health and Grant CD-177 from the American Cancer Society. Work by A.L.M., E.H.W., and R.A.F. was supported by Biogen N.V. D.M. is a Leukemia Society of America Scholar. C.P. is a Fellow of the Cancer Research Institute/F. M. Kirby Foundation and was also supported by NIH Training Grant...
5T32CA091-61-0T.