JOURNAL OF VIROLOGY, Apr. 1983, p. 42–49
0022-538X/83/040042-08$02.00/0

Vol. 46, No. 1

Transcription of Mouse Mammary Tumor Virus: Identification of a Candidate mRNA for the Long Terminal Repeat Gene Product

DAVID A. WHEELER,[1]* JANET S. BUTEL,[2] DANIEL MEDINA,[3] ROBERT D. CARDIFF,[4] AND GORDON L. HAGER[1]

Laboratory of Tumor Virus Genetics, National Cancer Institute, Bethesda, Maryland 20205[1]; Department of Virology and Epidemiology[2] and Department of Cell Biology,[3] Baylor College of Medicine, Houston, Texas 77030; and Department of Pathology, School of Medicine, University of California, Davis, California 95616[4]

Received 13 September 1982/Accepted 14 December 1982

We have examined an assortment of preneoplastic and neoplastic mouse mammary tissues for the presence of an mRNA which could encode the putative long terminal repeat gene product of mouse mammary tumor virus. We report here the detection of a novel mouse mammary tumor virus-specific, polyadenylic acid-containing transcript in certain preneoplastic and neoplastic mammary tissue of BALB/c mice. The molecule is 1.6 kilobases in length and contains sequences from the transcriptional leader and the U3 region of the proviral DNA. The upstream terminus of the 3′ information lies 75 to 80 nucleotides from the beginning of the long terminal repeat open reading frame, in close proximity to a consensus splice acceptor in the DNA. The transcript was detected in hormonally or chemically induced neoplastic, preneoplastic, and lactating mammary tissue of BALB/c mice, but not in preneoplastic or tumor tissue induced by exogenous viruses in any strain of mice examined. This implies that the RNA we observed is transcribed from an endogenous provirus template.

The long terminal repeat (LTR) of mouse mammary tumor virus (MMTV) contains an open reading frame of sufficient coding capacity to produce a 37,000-dalton protein. This novel gene was discovered by DNA sequencing of the provirus LTR (13) and by in vitro translation of 3′ polynucleotide fragments from the RNA genome (11, 34). The presence of the open reading frame has now been confirmed in four strains of MMTV: the exogenous viruses of C3H (11, 13; J. Majors, unpublished data) and GR (11, 17, 34) and the endogenous proviruses called unit II (12a, 21) and unit V (L. Johnson, unpublished data). Conservation of the open reading frame in these several virus strains strongly suggests that this putative gene is utilized at some time in the life cycle of the virus.

There are two strategies by which gene products are known to be expressed from retrovirus genomes. One is via translation of a multigenic mRNA and subsequent cleavage of a large polyprotein into individual gene products as in the expression of the *gag* and *pol* genes (1, 20). The other involves splicing of the transcriptional leader sequence onto genes that lie at the 3′ end of the viral genome. This strategy is exemplified by the expression of the *env* (32, 39) gene in replication-competent retroviruses and the *src* (32) gene in Rous sarcoma virus. If the putative LTR gene product (pLTR) were produced by synthesis and subsequent cleavage of a polyprotein, an env-pLTR fusion protein of about 103,000 daltons should be observed. Analysis of MMTV translation products by in vitro translation of 24S envelope mRNA (11, 15, 34, 35) or by pulse-labeling of virus-producing cells in vivo (1, 10) has never shown such an *env* precursor protein. Therefore, it is possible that the LTR gene may be expressed from a unique spliced mRNA.

Our approach to this problem was to analyze the RNA from mouse breast tissue in which MMTV is known to be transcriptionally active. These samples included normal lactating mammary gland, preneoplastic and tumor tissue induced by the milk-transmitted virus, and preneoplastic tissue and tumors induced by treatment with chemical carcinogens or mammotropic hormones. We found a novel MMTV-specific transcript 1.6 kilobases (kb) in length with properties expected for the mRNA that encodes pLTR.

MATERIALS AND METHODS

Mice and tissues. BALB/cCrl mice are free from the milk-borne virus and have a negligible incidence of breast cancer (26, 29). BALB/cV mice are a substrain of BALB/cCrlMed harboring a milk-borne virus which gives rise to a 47% incidence of breast tumors (14). BALB/c mice foster nursed on C3H mice (BALB/cfC3H) have a 90% incidence of breast tumor at 9 months. Tumors were induced in BALB/c mice by intragastric administration of 7,12-dimethylbenz[*a*]anthracene (DMBA) as previously described (6). Multiparous C3H/HeN mice were obtained from the small animal section at the National Institutes of Health, GRD/A mice are maintained in a colony at the Pathology Department at University of California, Davis, and the RIII tumor was the gift of Len Benade.

D1 preneoplastic tissue was derived from a hyperplastic alveolar nodule (HAN) that originated in a BALB/c mouse after hormonal stimulation with prolactin and estradiol (27) and is propagated in BALB/c mice by serial transplantation (26). The D1 tissue line originally had a low rate of spontaneous tumor production, but it has recently undergone a radical change in tumor potential and now has a 50% spontaneous tumor incidence at 6 months after transplantation (2). However, the D1 line still does not produce virus or viral antigens (2). The Z5d and $Z5c_1$ nodule outgrowth lines were developed from HAN occurring in BALB/cfC3H mice (3). These lines have a 15 and 44% spontaneous tumor rate, respectively, and produce viral RNA and antigens.

Lactating mammary glands were obtained from 6- to 8-month-old pregnant BALB/c mice.

RNA purification. Whole cell RNA was purified from mammary tissues by either the guanidine hydrochloride extraction procedure of Deely et al. (9), with modifications, or by a phenol extraction procedure outlined below. Tissues weighing from 0.4 to 2 g were frozen in liquid nitrogen and pulverized to a granular powder. The granules were dissolved in 10 volumes of 8 M guanidine hydrochloride–20 mM sodium acetate (pH 5.0), and the mixture was homogenized by a Brinkmann polytron for 20 s at full speed. After three successive ethanol precipitations from guanidine, the RNA was dissolved in a minimum volume of 0.1 M sodium acetate–20 mM EDTA (pH 7.0) and precipitated by the addition of 1.45 volumes of 5 M sodium acetate (pH 5.0). This procedure was repeated twice, after which the pellet was dissolved in sterile distilled water and heated at 65°C for 2 min. The RNA was ethanol precipitated twice from distilled water and then stored under ethanol at −20°C until used.

RNA from hyperplastic outgrowths and tumors weighing less than 0.5 g were isolated and purified by phenol extraction. Water-saturated phenol was mixed with an equal volume of buffer composed of 0.1 M Tris-hydrochloride (pH 9.0), 1 mM EDTA, and 0.5% sodium dodecyl sulfate (SDS). This mixture was added to pulverized tissue in a ratio of 10 ml/0.1 g of tissue. The mixture was immediately homogenized as above. The phases were separated by centrifugation, and the phenol was extracted with an equal volume of the Tris-EDTA-SDS buffer. The aqueous phases, including the white interface, were pooled and extracted with phenol until the white interface disappeared. This usually required two to four extractions, after which the clear aqueous phase was extracted with phenol-chloroform-isoamyl alcohol (24:24:1) and finally with chloroform-isoamyl alcohol (24:1). The aqueous phase was then made 0.1 M in sodium acetate (pH 5.0), and the nucleic acids were precipitated by addition of 2 volumes of ethanol. The precipitate was dissolved in 1.7 ml of sterile distilled water, and 2 volumes of 5 M sodium acetate (pH 5.0) was added. The RNA was precipitated for 5 to 6 h at 4°C and was collected by centrifugation at 30,000 rpm for 30 min in a type 50 rotor. The pellet was dissolved in water, ethanol precipitated, and stored at −20°C until used.

Oligodeoxythymidylic acid-cellulose chromatography. The RNA precipitate was collected and dissolved in 0.1 M LiCl–10 mM Tris-hydrochloride–2 mM EDTA–0.1% SDS (pH 7.2) and then heated to 55°C for 2 min. The solution was quick chilled and then passed over an oligodeoxythymidylic acid-cellulose column (Bethesda Research Laboratories), preequilibrated in the same buffer, at a ratio of 1 mg of RNA to 50 mg of oligodeoxythymidylic acid. The flow-through was collected and passed back over the column two more times before the column was washed with 20 volumes of buffer. The bound RNA was eluted with 10 mM Tris-hydrochloride–2 mM EDTA–0.1% SDS, and the RNA concentration was determined. The sodium acetate concentration was made 0.25 M at pH 5.0, and 20 μg of tRNA was added before ethanol precipitation. The precipitated RNA was collected and dissolved in the column elution buffer at a concentration of 1 mg/ml with respect to cellular RNA and then stored at −196°C until used.

Blot hybridizations. RNA was denatured in formaldehyde as described elsewhere (16a) or by glyoxal as described by McMaster and Carmichael (25), and electrophoresed through 1% agarose and then transferred to nitrocellulose paper and hybridized by the procedure of Thomas (38). Hybridizations were carried out in 50% formamide, 5× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate), 25 mM sodium phosphate (pH 6.8), 0.5% SDS, 250 μg of sonicated and denatured salmon sperm DNA per ml, and 2% each of bovine serum albumin, polyvinylpyrrolidone, and Ficoll. Probes with an average specific activity of 10^8 cpm/μg were prepared by nick translation and added to the hybridization mixture at 10^6 cpm/ml. The reactions were sealed in a plastic bag and hybridized for 40 h at 42°C. Filters were then washed by the method of Thomas (38) and autoradiographed on Kodak XR-1 film with DuPont Cronex intensifying screens at −70°C.

Subgenomic MMTV DNA fragments. Gene fragments of MMTV have been subcloned in pBR322 from the original *Eco*RI partial digest library in the phage vector Charon 4A (12). MMTV-specific fragments were purified from these plasmids by cutting with the appropriate restriction enzymes followed by agarose gel electrophoresis and electroelution of the viral fragments or by *N,N'*-bisacrylylcystamine-acrylamide (Bio-Rad Laboratories) electrophoresis as described previously (18). Eluted fragments were purified by DEAE-cellulose chromatography and concentrated by ethanol precipitation. The DNA fragments used as probes in this report are represented schematically in Fig. 1. The 220-base-pair (bp) fragment, used as a leader probe, extends from 30 bp inside U5 to approximately 15 bp inside the *gag* gene (L. Johnson, unpub-

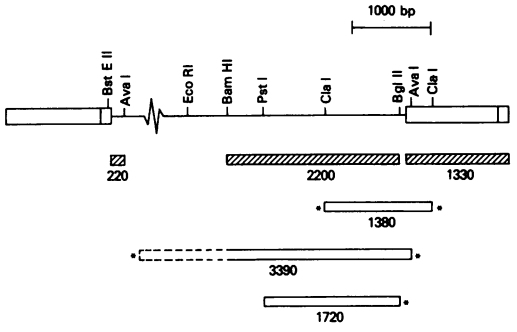

FIG. 1. Schematic diagram of the linear MMTV provirus and subgenomic fragments of MMTV used as probes. The *Eco*RI site is near the center of the linear MMTV genome. Restriction sites to the left of *Eco*RI are from the unit V provirus (L. Johnson, unpublished data), and sites to the right are from the C3H exogenous virus. The restriction sites define the termini for the DNA fragments below them. The cross-hatched fragments were nick translated and used in the blot hybridization analysis. The open bar fragments were end labeled and used in the S1 nuclease experiments. The asterisk indicates the location of the ^{32}P label. The dashed region of the 3,390-bp probe indicates a portion of pBR322 sequences present in the probe. The numbers below the bars are the sizes of the fragments in base pairs. The location of the restriction sites in the LTR, numbering base pairs from the right virus-host junction, are as follows: *Cla*I, 993; *Ava*I, 1,292; *Bgl*II, 1,367. Other information concerning these DNA fragments is given in the text.

lished data). The 2,200-bp fragment, used as an *env* gene probe, encompasses *env* and a small portion of *pol*. The 1,330-bp LTR probe was subcloned as described previously (13).

Labeled probes and DNA sequencing. DNA fragments were uniformly ^{32}P labeled by nick translation (30), modified as described previously (19). ^{32}P end-labeled probes and DNA sequencing were accomplished by the procedure of Maxam and Gilbert (22).

S1 nuclease mapping of transcripts. The procedure of Berk and Sharp (4), with modifications as described previously (19), was used to map RNA splice acceptor sites.

RESULTS

MMTV transcripts detected in mammary tumor tissue. Initial examination of virus-induced tumors showed the presence of four viral RNA species. An 8.9-kb mRNA representing the full-length genome and a 3.6-kb mRNA representing the spliced envelope message were readily apparent in the majority of virus-induced tumors in all high-tumor-incidence strains of mice (Fig. 2). In addition, there were two minor species of MMTV-specific RNA that occurred infrequently in tumors and at low concentrations relative to the genomic and *env* mRNA species. These species migrated at 2.5 and 1.0 kb and hybridized most efficiently with the *env* probe (Fig. 2; lanes 1, 2, 3, 4, and 7 for the 2.5-kb transcript; lanes 1, 2, and 4 for the 1.0-kb transcript). LTR and leader sequences were also detectable in the 2.5-kb RNA, but not the 1.0-kb species (data not shown). The 1.0-kb transcript was synthesized at levels so low that it was not possible to determine unambiguously that the LTR and leader sequences were absent from the molecule. The 2.5- and 1.0-kb species may possibly be partial *env* gene transcripts, but their complete structure and function are presently undetermined.

MMTV RNA was also produced in mammary tumors induced by hormones and chemicals (6, 16, 23, 24, 28, 29, 33) in the absence of infectious virus. In BALB/c tumors induced by carcinogens, we (2, 6, 16) and others (28) have observed low levels of RNA expression with an absence of viral antigen synthesis. These observations raised the possibility that some previously uncharacterized gene product was being expressed. Therefore, we examined a series of such tumors to characterize the viral transcripts being produced. The results showed that a novel transcript of 1.6 kb was synthesized in several DMBA-induced BALB/c mammary tumors (Fig. 3A and Table 1). This transcript displayed strong

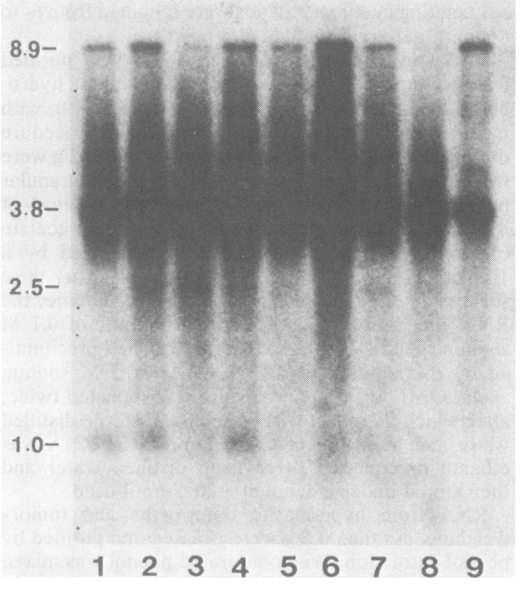

FIG. 2. Northern blot analysis of RNA from mouse mammary tumors. RNA was purified by guanidine hydrochloride extraction (see text). Approximately 5 g of polyadenylic acid-selected tumor RNA was denatured by glyoxal, electrophoresed through 1% agarose, blotted, and hybridized with the 2,200-bp *env* fragment (Fig. 1). Lanes: 1, transplanted tumor originating in virus-infected D2 tissue; 2, 3, and 4, C3H/HeN tumors; 5, 7, 8, and 9, transplanted BALB/cV tumors; 6, BALB/cfcV tumor.

TABLE 1. Distribution of the 1.6-kb RNA in mammary tissues from various strains of mice

Mouse strain	Exogenous virus[a]	Type of mammary tissue tested[b]	Presence of 1.6-kb RNA[c]
BALB/cCrl	−	LMG[d]	4/4
		HOG (D1)	5/5[e]
		Tumor (D1)	3/8[f]
		Tumor (DMBA)	7/10
BALB/cV	+	Tumor	0/2
BALB/cfC3H	+	HOG (Z)	0/3
		Tumor	0/5
		Tumor (DMBA)	0/1
C3H/HeN	+	Tumor	0/10
GR/A	+	Tumor	0/1
RIII	+	Tumor	0/1

[a] −, Absence of exogenous virus; +, presence of exogenous virus.

[b] "Tumor" designates mammary tumors which appeared spontaneously in breeding females. "Tumor (DMBA)" indicates that the mouse was administered the chemical carcinogen DMBA. The origins of the D1 and Z HOG lines are detailed in Materials and Methods.

[c] The fraction represents the number of specimens positive for the 1.6-kb RNA species over the number of specimens tested.

[d] LMG, Lactating mammary gland.

[e] D1 tissue maintained separately by R.D.C. and D.M. gave the same results.

[f] One tumor that contained the 1.6-kb transcript was from a mouse that had been given a pituitary isograft at the time of transplantation of the D1 HOG.

hybridization to the leader (Fig. 3A, lane 1) and LTR (lane 3) probes and relatively weak hybridization to the *env* probe (lane 2). The molecule therefore contained 5′ leader sequences juxtaposed to the 3′ end of the genome. The weak hybridization to the *env* probe indicated that only a small portion of the *env* sequences were present in the RNA, consistent with the juncture between the 5′ and 3′ information residing close to the *Bgl*II site that defines the right terminus of the *env* probe (see below).

MMTV transcripts detected in mammary preneoplastic tissue. Normal mouse breast tissue proceeds to a tumor through a well-defined preneoplastic state called an HAN (26). HAN are morphologically distinct foci of proliferating tissue in the normal mammary gland that can be excised and transplanted into the breast fat pad of a syngeneic recipient in which the resident mammary ducts have been removed. Serial transplantation of the HAN results in the establishment of a quasi-stable tissue line called a hyperplastic outgrowth (HOG). BALB/c mice have a negligible incidence of spontaneous breast tumors and a correspondingly negligible incidence of spontaneous HAN. However, HAN may be induced by chemical carcinogens or prolonged hormonal stimulation (26). A well-characterized HOG, D1, was developed from an HAN induced in BALB/c by a combination of prolactin and estradiol (27). Other HOG lines have been established from HAN arising in BALB/cfC3H mice (3) and are referred to as Z lines. We examined D1 and two Z lines, $Z5c_1$ and Z5d, for viral RNA transcription. The results obtained (Fig. 3B) show that D1 tissue contains a 1.6-kb transcript similar to that described above. This RNA was found to hybridize with the LTR and leader probes in a pattern similar to that of the DMBA-induced tumor, except that in this case hybridization to the *env* probe was not detectable (Fig. 3B, lane 2). The inability to detect hybridization with the *env* probe was probably due to the relatively low amount of transcript expressed in the D1 tissue (data not shown).

In contrast to the results obtained with the D1 tissue, HOG lines developed from virus-producing BALB/cfC3H mice contained the usual viral messages of 8.9 and 3.8 kb, plus a smaller 1.2-kb transcript (Fig. 4A and B, lanes 1 and 2). The 1.2-kb RNA species was detected by the LTR probe (Fig. 4A), but did not hybridize with either the leader probe (Fig. 4B) or the *env* probe (data not shown).

Mapping of potential splice acceptor. The information present in the 1.6-kb transcript suggested that it resulted from splicing the MMTV leader to a position upstream from the 3′ LTR.

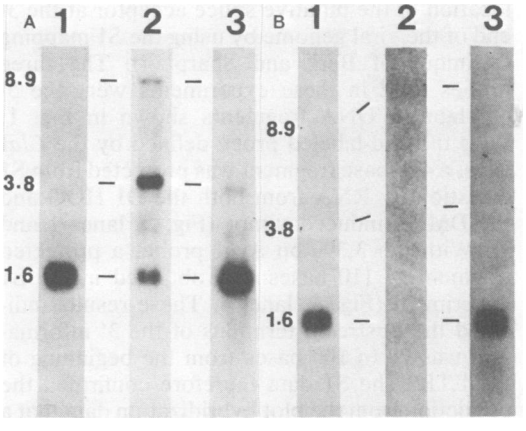

FIG. 3. Blot analysis of RNA from BALB/c mammary tissue. A, RNA from a DMBA-induced tumor. B, RNA from D1 HOG. RNA was purified from the DMBA-induced tumor by the guanidine hydrochloride technique and from the D1 tissue by phenol extraction (see text). Approximately 3 g of polyadenylated RNA per lane was denatured in formaldehyde, electrophoresed, blotted, and hybridized to MMTV DNA from the 220-bp transcriptional leader (lane 1), the 2,200-bp *env* fragment (lane 2), or the 1,330-bp LTR fragment (lane 3).

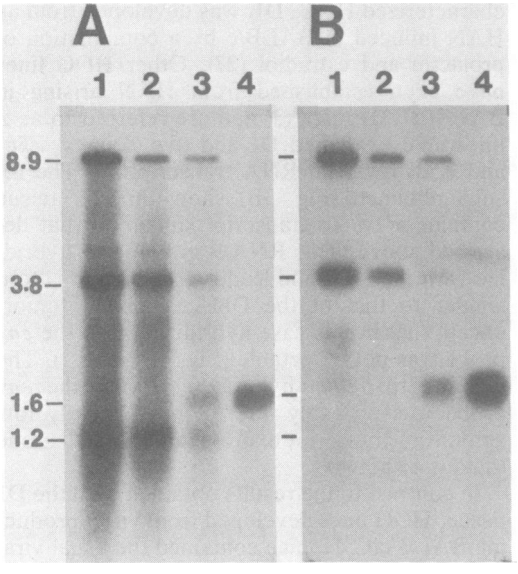

FIG. 4. Northern blot analysis of RNA from BALB/c HOGs. RNA from HOGs was treated as described in the legend to Fig. 3. A, Hybridization to the 1,328-bp LTR probe. B, Hybridization to the 220-bp leader probe. The HOG lines tested were as follows: lane 1, Z5d; lane 2, $Z5c_1$; lane 3, D1 transplanted into a BALB/cfC3H mouse; lane 4, D1 transplanted into a BALB/c mouse.

Since the major species of MMTV-specific RNA in the D1 and DMBA-induced tumor tissues was the 1.6-kb transcript, it was possible to map the location of the putative splice acceptor at the 3′ end of the viral genome by using the S1 mapping technique of Berk and Sharp (4). The three probes used in these experiments were the 5′ end-labeled DNA fragments shown in Fig. 1. With the end-labeled probe defined by the *Cla*I sites, a 420-base fragment was protected from S1 digestion by RNA from both the D1 HOG and the DMBA-induced tumor (Fig. 5, lanes 1 and 2). With the 3,390-bp *Ava*I probe, a protected fragment of 110 bases was obtained in the S1 experiment (Fig. 5, lane 3). These results indicated the upstream terminus of the 3′ information was 70 to 100 bases from the beginning of the LTR. The S1 data therefore confirmed the prediction from the blot hybridization data that a relatively small amount of the *env* probe sequence was represented in the transcript.

To map more precisely the location of the potential splice boundary, a sequencing ladder, constructed from the 1,720-bp *Bgl*II probe, was used as a sizing marker for the S1-resistant *Bgl*II fragment-RNA hybrid. The S1 nuclease digestion product of the tumor RNA-*Bgl*II probe hybrid gave a constellation of bands ranging from 75 to 80 bases from the beginning of the LTR (Fig. 6, lanes 2 and 3), in close agreement with the results obtained with the *Cla*I and *Ava*I probes. The sequence in Fig. 6 was transposed to its plus strand complement in Fig. 7 to clarify the genetic relationships in this region of the genome. At a position 75 bases upstream from the LTR there was an octanucleotide that agreed closely with the consensus sequence for a splice acceptor, YNYYYNCAG, where Y is a pyrimidine and N is any nucleotide (36). The smallest S1 fragment migrated at a position 1.5 bases smaller than the predicted acceptor nucleotide, a thymidine residue. The 1.5-base lag of the S1 product was presumably due to the difference in the terminal nucleotide structure between the chemically hydrolyzed DNA of the sequence ladder and the S1-treated probe (37).

Distribution of the 1.6-kb RNA in mammary tissue. The presence of information from both the 5′ end and 3′ end of the viral genome in the 1.6-kb transcript, the location of the left terminus of the 3′ information upstream from the initiation codon beginning the pLTR reading frame, and the coincidence of a consensus splice

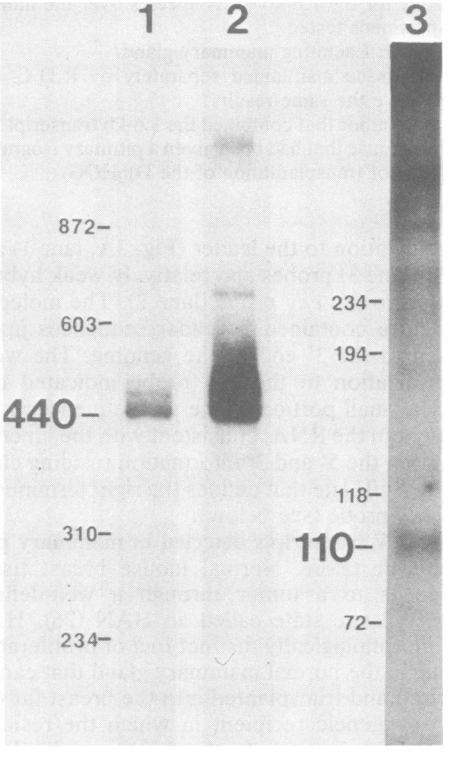

FIG. 5. S1 nuclease analysis of RNA from BALB/c mammary tissue. Approximately 100 μg of unfractionated RNA was hybridized to the end-labeled probes (Fig. 1). The hybrids were processed as described in the text and then electrophoresed through 5% polyacrylamide. RNA was from the D1 hyperplastic outgrowth (lane 1) and the DMBA-induced BALB/c tumor (lanes 2 and 3). The probes used were the 1,380-bp *Cla*I fragment (lanes 1 and 2) and the 3,390-bp *Ava*I fragment (lane 3).

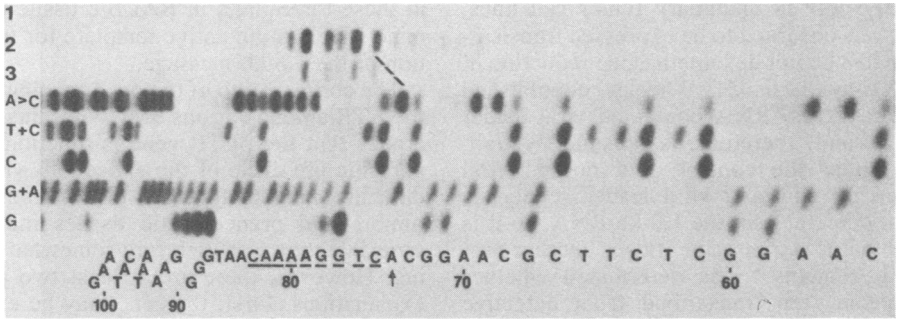

FIG. 6. Precise localization of the 1.6-kb RNA splice acceptor. Approximately 50 g of polyadenylic acid-selected RNA from the DMBA-induced BALB/c tumor was hybridized to the 1,720-bp *Bgl*II fragment, digested by S1 nuclease as described in the text, and electrophoresed on a 12% sequencing gel. Before electrophoresis a sample of the hybridization reaction was trichloroacetic acid precipitated and counted. Lanes: 1, 2,000 cpm of a hybridization to RNA from the cell line HTC; 2 and 3, 1,000 and 250 cpm, respectively, of the hybridization to the DMBA-induced tumor RNA. The S1 digestion products are run in parallel with the sequencing ladder generated by hydrolytic cleavage of the probe. The minus strand sequence read from the ladder is indicated below the ladder. The numbering represents the distance in base pairs from the 3′ end of the LTR. The underlined bases correspond to the complement of the consensus splice acceptor, and the arrow points to the nucleotide which accepts the spliced leader. The dashed line shows the proximity of the smallest major S1-generated fragment to the splice acceptor.

acceptor sequence in the DNA strongly suggested that the transcript was an mRNA for the pLTR protein. We examined a variety of normal, preneoplastic, and malignant breast tissues, but detected this transcript only in the absence of exogenous MMTV infection (Table 1). It was detected consistently in lactating mammary gland (data not shown) and the D1 HOG line, in 7 of 10 DMBA-induced tumors in BALB/c mice, and in 3 of 8 tumors arising spontaneously from D1 tissue. The 1.6-kb transcript was not detected in mammary tumors obtained from virus-infected BALB/cfC3H, BALB/cV, C3H, GR, or RIII mice. We have not investigated MMTV transcription in the lactating mammary glands or preneoplastic tissue from any strain other than BALB/c.

DISCUSSION

In an attempt to detect the expression of the pLTR gene we analyzed MMTV-specific RNA in chemically, hormonally, or virus-induced preneoplastic and neoplastic mouse breast tissue. Expression of MMTV antigens or RNA (or both) is known to accompany transformation of murine mammary cells by these agents (6, 16, 23, 24, 28, 33). In virus-induced tumors a series of subgenomic mRNA species migrating at 3.8, 2.5, and 1.0 kb arose that hybridized strongly with the *env* gene probe (Fig. 2). The 3.8-kb transcript is known to be translated into the gp52 and gp36 envelope proteins (11, 15, 34, 35), but the function of the smaller *env* transcripts is unknown. The 2.5-kb transcript has been identified

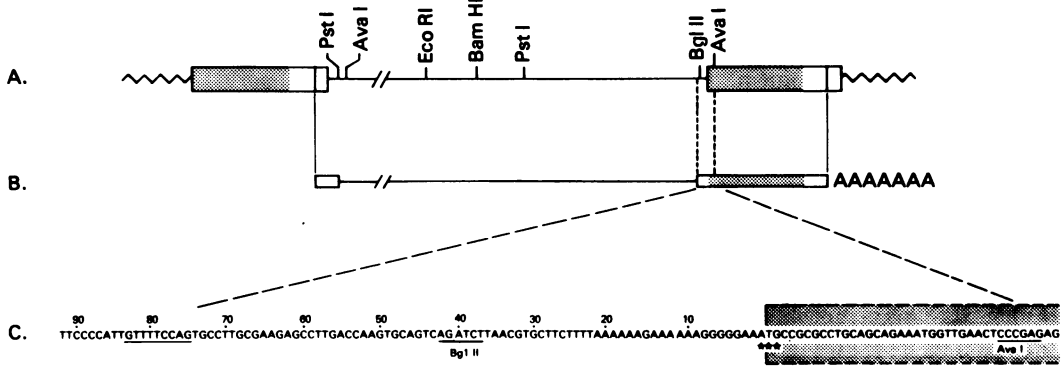

FIG. 7. Schematic representation of the LTR transcript. A, Integrated MMTV provirus. The wavy lines indicate host DNA, and the stippling indicates the location of the open reading frame. B, Structure of the 1.6-kb RNA derived by the data in this paper. The open boxes represent sequences in the final spliced RNA, and the straight line represents sequences removed by splicing. C, DNA sequence of the splice acceptor region. The numbering of the bases is the same as in Fig. 6. The underlined sequence is the splice acceptor consensus. The stippling delineates the LTR. The stars highlight the translation initiation codon for pLTR.

previously only in mammary tumor cell lines, where it was observed to be expressed transiently for up to 8 h after dexamethasone induction of viral RNA synthesis (31; Wheeler, unpublished data). The 2.5-kb RNA contained viral leader sequences and, therefore, is presumably transcribed under the control of a normal viral promoter. In contrast, viral leader sequences were not detectable in the 1.0-kb RNA, so it is possible that this transcript utilizes another promoter. It remains to be determined whether these species are transcribed from defective proviruses containing deletions in the *env* gene or are products of alternative processing pathways of *env* mRNA.

In addition, a molecule of 1.2 kb was observed in three preneoplastic tissues with exogenous virus involvement. This molecule contained LTR sequences, but no viral leader sequences were detectable (Fig. 4A), and it has not been analyzed further.

A novel MMTV-specific transcript of 1.6 kb appeared regularly in BALB/c lactating mammary glands and mammary neoplasms of nonviral etiology (Table 1). This transcript had characteristics expected of an mRNA for the LTR gene product: it was a polyadenylic acid-containing molecule, sequences from the 5′ transcriptional leader were adjoined to the U3 region of the genome (Fig. 3, 5, and 6), and the boundary of the 3′ information occurred at a consensus splice acceptor sequence (Fig. 6 and 7), arguing that a splicing mechanism accounted for the ligation of these sequences. These data constitute the first evidence that the pLTR gene is expressed and support our earlier hypothesis (6, 16) that partial MMTV genome transcription in BALB/c mammary tumors could reflect the expression of a previously uncharacterized MMTV gene product.

Since the 1.6-kb transcript was present in BALB/c tissues which do not contain a milk-borne (exogenous) virus, the molecule must have been transcribed from an endogenous template. BALB/c mice contain three copies of endogenous provirus, called units I, II and III (8). Units II and III are complete copies of the viral genome at the level of resolution afforded by Southern blot analysis, whereas unit I is a defective provirus containing a maximum of 3.2 kb of 3′ viral information (7). There is one report that in tumors arising from DI HOGs the unit I provirus is selectively hypomethylated (5), suggesting that this template is available for transcription. When subgenomic MMTV probes are employed, the unit I provirus hybridizes efficiently only with leader and LTR-containing sequences (D. Wheeler, unpublished data), thus provoking the speculation that unit I is a reverse transcript of the 1.6-kb RNA. We are currently investigating the structure of the LTRs present in these three units in BALB/c tissue to determine which is the active template for transcription of the 1.6-kb message.

The conservation of the open reading frame in the LTR among various MMTV strains strongly argues that the pLTR gene is functional in the parasitic life cycle of the exogenous virus. The difficulty in detecting the pLTR transcript in tumors and preneoplastic tissues induced by virus infection was therefore somewhat perplexing. However, there are at least two plausible explanations. First, the gene may be expressed only at levels too low to be readily detectable by the blot hybridization assay. In preliminary experiments, S1 mapping of virus-specific RNA from 34i cells, a line derived from a virus-induced tumor in a C3H mouse, revealed the same splice junction as seen in the DMBA tumor, but the level of the RNA containing this splice was approximately 1% of that detected with the DMBA-induced tumor RNA reported here. The pLTR transcript could not be detected by blot hybridization in RNA from this line.

Second, the pLTR gene might be expressed at a specific stage in the life cycle of the exogenous virus. The possibility that the LTR message is transiently expressed during hormone induction of viral RNA transcription has been investigated by assaying the relative concentrations of viral messages at different times after the addition of hormone to MMTV-infected cell cultures. The LTR transcript was not discernible before or at any time up to 48 h after induction of transcription by dexamethasone in 34i cells, MMTV-infected HTC cells (D. Wheeler, unpublished data), cultured GR mouse mammary tumor cells, or MMTV-infected XC cells (31). Nevertheless, temporal expression of pLTR could be coupled to other variables in the hormonal milieu of the differentiating breast tissue during the life of the mouse.

The results reported here clearly identify an RNA species uniquely encoding the putative MMTV LTR gene in lactating, preneoplastic, and neoplastic mammary tissues of BALB/c mice. Further investigation is necessary to determine the factors regulating the expression of this new MMTV gene and its biological function.

ACKNOWLEDGMENTS

We thank Ronald G. Wolford for his expert technical assistance in the preparation of the sequencing ladder.

This work was supported in part by Public Health Service contract NOI-CP9-1020 awarded by the National Cancer Institute.

LITERATURE CITED

1. **Anderson, S. J., R. B. Naso, J. Davis, and J. M. Bowen.** 1979. Polyprotein precursors to mouse mammary tumor virus proteins. J. Virol. 32:507–516.
2. **Ashley, R. L., R. D. Cardiff, and T. G. Fanning.** 1980.

Reevaluation of the effect of mouse mammary tumor virus infection on the BALB/c mouse hyperplastic outgrowth. J. Natl. Cancer Inst. 65:977–983.
3. **Ashley, R. L., R. D. Cardiff, D. J. Mitchell, L. J. Faulkin, and J. K. Lund.** 1980. Development and characterization of mouse hyperplastic mammary outgrowth lines from BALB/cfC3H hyperplastic alveolar nodules. Cancer Res. 40:4232–4242.
4. **Berk, A. J., and P. A. Sharp.** 1978. Spliced early mRNAs of simian virus 40. Proc. Natl. Acad. Sci. U.S.A. 75:1274–1278.
5. **Breznik, T., and J. C. Cohen.** 1982. Altered methylation of endogenous viral promoter sequences during mammary carcinogenesis. Nature (London) 295:255–257.
6. **Butel, J. S., S. Dusing-Swartz, S. H. Socher, and D. Medina.** 1981. Partial expression of endogenous mouse mammary tumor virus in mammary tumors induced in BALB/c mice by chemical, hormonal and physical agents. J. Virol. 38:571–580.
7. **Cohen, J. C., J. E. Majors, and H. E. Varmus.** 1979. Organization of mouse mammary tumor virus-specific DNA endogenous to BALB/c mice. J. Virol. 32:483–496.
8. **Cohen, J. C., and H. E. Varmus.** 1979. Endogenous mammary tumor virus DNA varies among wild mice and segregates during inbreeding. Nature (London) 278:418–423.
9. **Deeley, R. G., J. I. Gordon, A. T. H. Burns, K. P. Mullinix, M. Bina-Stein, and R. F. Goldberger.** 1977. Primary activation of the vitellogenin gene in the rooster. J. Biol. Chem. 252:8310–8319.
10. **Dickson, C., and M. Atterwill.** 1980. Structure and processing of the mouse mammary tumor virus glycoprotein precursor Pr73^{env}. J. Virol. 35:349–361.
11. **Dickson, C., and G. Peters.** 1981. Protein-coding potential of mouse mammary tumor virus genome RNA as examined by *in vitro* translation. J. Virol. 37:36–47.
12. **Donehower, L. A., J. Andre, D. S. Berard, R. G. Wolford, and G. L. Hager.** 1980. Construction and characterization of molecular clones containing integrated mouse mammary tumor virus sequences. Cold Spring Harbor Symp. Quant. Biol. 44:1153–1159.
12a. **Donehower, L. A., B. Fleurdelys, and G. L. Hager.** 1983. Further evidence for the protein coding potential of the mouse mammary tumor virus long terminal repeat: nucleotide sequence of an endogenous proviral long terminal repeat. J. Virol. 45:941–949.
13. **Donehower, L. A., A. L. Huang, and G. L. Hager.** 1981. Regulatory and coding potential of the mouse mammary tumor virus long terminal redundancy. J. Virol. 37:226–238.
14. **Drohan, W., Y. A. Teramoto, D. Medina, and J. Schlom.** 1981. Isolation and characterization of a new mouse mammary tumor virus from BALB/c mice. Virology 114:175–186.
15. **Dudley, J. P., and H. E. Varmus.** 1981. Purification and translation of murine mammary tumor virus mRNAs. J. Virol. 39:207–218.
16. **Dusing-Swartz, S., D. Medina, J. S. Butel, and S. H. Socher.** 1979. Mouse mammary tumor virus genome expression in chemical carcinogen induced mammary tumors in low- and high- tumor-incidence mouse strains. Proc. Natl. Acad. Sci. U.S.A. 76:5360–5364.
16a. **Ellis, R. W., D. DeFeo, M. E. Furth, and E. M. Scolnick.** 1982. Mouse Cells contain two distinct *ras* gene mRNA species that can be translated into a p21 *onc* protein. Mol Cell. Biol. 2:1339–1345.
17. **Fasel, N., K. Pearson, E. Buetti, and H. Digglemann.** 1982. The region of mouse mammary tumor virus DNA containing the long terminal repeat includes a long coding sequence and signals for hormonally regulated transcription. EMBO J. 1:3–7.
18. **Hansen, J. N., B. H. Pheiffer, and J. A. Boehnert.** 1980. Chemical and electrophoretic properties of solubilized disulfide gels. Anal. Biochem. 105:192–201.
19. **Haung, A. L., M. C. Ostrowski, D. Berard, and G. L. Hager.** 1981. Glucocorticoid regulation of the Ha-MuSV p21 gene conferred by sequences from mouse mammary tumor virus. Cell 27:245–255.
20. **Jamjoom, G. A., R. B. Naso, and R. B. Arlinghaus.** 1977. Further characterization of intracellular precursor polyproteins of Rauscher leukemia virus. Virology 78:11–34.
21. **Kennedy, N., G. Knedlitschek, B. Groner, N. E. Hynes, P. Herrlich, R. Michalides, and A. J. J. van Ooyen.** 1982. Long terminal repeats of endogenous mouse mammary tumor virus contain a long open reading frame which extends into adjacent sequences. Nature (London) 295:622–624.
22. **Maxam, A., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base specific chemical cleavage. Methods Enzymol. 65:499–560.
23. **McGrath, C. M., and R. F. Jones.** 1978. Hormonal induction of mammary tumor viruses and its implications for carcinogenesis. Cancer Res. 38:4112–4125.
24. **McGrath, C. M., E. J. Marineau, and B. A. Voyles.** 1978. Changes in MuMTV DNA and RNA levels in BALB/c mammary epithelial cells during malignant transformation by exogenous MuMTV and by hormones. Virology 87:339–353.
25. **McMaster, G. K., and G. C. Carmichael.** 1977. Analysis of single- and double-stranded nucleic acids in polyacrylamide and agarose gels by using glyoxal and acridine orange. Proc. Natl. Acad. Sci. U.S.A. 74:4835–4838.
26. **Medina, D.** 1973. Preneoplastic lesions in mouse mammary tumorigenesis. Methods Cancer Res. 7:3–53.
27. **Medina, D., and K. B. DeOme.** 1968. Influence of mammary tumor virus on the tumor-producing capabilities of nodule outgrowth free of mammary tumor virus. J. Natl. Cancer Inst. 40:1303–1308.
28. **Michalides, R., L. van Deemter, R. Nusse, and P. Hageman.** 1979. Induction of mouse mammary tumor virus RNA in mammary tumors of BALB/c mice treated with urethane, X irradiation, and hormones. J. Virol. 31:63–72.
29. **Pauley, R. J., D. Medina, and S. H. Socher.** 1979. Murine mammary tumor virus expression during mammary tumorigenesis in BALB/c mice. J. Virol. 29:483–493.
30. **Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg.** 1977. Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. J. Mol. Biol. 113:237–251.
31. **Robertson D. J., and H. E. Varmus.** 1980. Dexamethasone induction of the intracellular RNA of mouse mammary tumor virus. J. Virol. 40:673–683.
32. **Rothenberg, E., D. J. Donoghue, and D. Baltimore.** 1978. Analysis of a 5′ leader sequence on murine leukemia virus 21S RNA: heteroduplex mapping with long reverse transcriptase products. Cell 13:435–451.
33. **Schlom, J., R. Michalides, D. Kufe, R. Hehlmann, S. Spiegelman, P. Bentvelzen, and P. Hageman.** 1973. A comparative study of the biologic and molecular basis of murine mammary carcinoma: a model for human breast cancer. J. Natl. Cancer Inst. 51:541–551.
34. **Sen, G. C., J. Racevsis, and N. H. Sarkar.** 1981. Synthesis of murine mammary tumor viral proteins in vitro. J. Virol. 37:963–975.
35. **Sen, G. C., S. W. Smith, S. L. Marcus, and N. H. Sarkar.** 1979. Identification of the messenger RNAs coding for the *gag* and *env* gene products of the murine mammary tumor virus. Proc. Natl. Acad. Sci. U.S.A. 76:1736–1740.
36. **Sharp, P. A.** 1981. Speculations on RNA splicing. Cell 23:643–646.
37. **Sollner-Webb, B., and R. H. Reeder.** 1979. The nucleotide sequence of the initiation and termination sites for ribosomal RNA transcription in X. laevis. Cell 18:485–499.
38. **Thomas, P.** 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. U.S.A. 77:5201–5205.
39. **Weiss, S. R., H. E. Varmus, and J. M. Bishop.** 1977. The size and genetic composition of virus-specific RNAs in the cytoplasm of cells producing avian sarcoma-leukosis viruses. Cell 12:983–992.