JOURNAL OF VIROLOGY, Sept. 1990, p. 4043–4050
0022-538X/90/094043-08$02.00/0

Vol. 64, No. 9

Altered Mouse Mammary Tumor Virus Transcript Synthesis in T-Cell Lymphoma Cells

JANIS RACEVSKIS

Department of Oncology, Montefiore Medical Center, and Department of Medicine, Albert Einstein Cancer Center, Bronx, New York 10467

Received 28 December 1989/Accepted 23 May 1990

Proviral copies of mouse mammary tumor virus (MMTV) are known to be amplified in certain T-cell lymphomas. Transcription of the amplified MMTV proviruses was studied in detail in two T-cell lymphoma lines and showed the production of deletions and premature termination of *env* mRNAs and the premature termination of *gag* transcripts. EL-4 cells produce three *env* mRNAs, and sequence analysis of cDNAs of the two smaller transcripts revealed large deletions encompassing the 3′ half of the *env* gene. The deletion in at least one of the altered transcripts appeared to be produced by a splicing mechanism. T-cell lymphoma line ML of DBA/2 mice also synthesizes two smaller *env* transcripts, both of which result from premature termination of transcription. Both lines transcribe high levels of *gag* mRNAs of about 0.8 kilobases in length, terminating at the end of the region encoding MMTV phosphoprotein pp21. Restriction enzyme *Bam*HI analysis of the amplified proviruses of EL-4 and ML cells as well as of additional non-mammary tumor cell types containing amplified MMTV proviruses suggested that the amplified proviruses were derived from exogenous viruses, or activated endogenous provirus MTV-1 in the case of DBA/2 strain tumor cells.

Mouse mammary tumor virus (MMTV), the causative agent of mammary tumors in susceptible mice, is also known to be expressed in tissues and tumors other than those of the mammary gland. The most extensively documented association of MMTV with nonmammary tumors is that with T-cell lymphomas (1, 8, 22, 27, 32). A characteristic common to all the MMTV-expressing T-cell lymphomas is the presence of numerous acquired (amplified) MMTV proviral genomes (8, 22), all containing long deletions in their long terminal repeat (LTR) regions (1, 14, 17, 18, 21, 23). In each case, the amplified proviral genomes appear to have originated from a single transcript bearing a deletion ranging in size between 300 and 500 base pairs located within the U_3 region of the LTR.

MMTV LTRs contain regulatory signals required for control of proviral gene expression, and these regulatory elements presumably interact with host cell transcription factors which determine viral expression in the particular cell type. A sequence exhibiting negative regulatory effects in transcription assays has been identified within the region of the LTRs that is always deleted in amplified T-cell lymphoma proviruses (14, 25). Transcription assay constructs containing MMTV LTRs bearing deletions appear to be more efficient in lymphoid cells than constructs with full-length LTRs (35). The T-cell lymphoma cells do not produce mature viral particles (27, 32), although they do transcribe high levels of MMTV RNA (9, 17, 23, 30). The mRNAs are translated into *gag* and *env* polyprotein precursors but are then not processed further into mature viral structural proteins because of an apparent block in the maturation pathway (27, 32). A C57BL/6 mouse-derived T-cell lymphoma line, EL-4, has also been reported to synthesize an unusual 1.0-kilobase (kb) LTR transcript (9, 30) which initiates within the envelope gene in response to phorbol ester treatment (9).

A detailed examination of the MMTV transcripts synthesized in two T-cell lymphoma lines revealed the presence of other unusual transcripts that contained deletions or were prematurely terminated. These observations are presented in this article and suggest the presence of mechanisms in T-cell lymphoma cells which create these MMTV transcript alterations.

MATERIALS AND METHODS

Cells and tissues. Transplantable T-cell lymphoma line ML cells were passaged in 6- to 8-week-old male DBA/2 mice as described previously (32). C57BL/6 mouse-derived T-cell lymphoma line EL-4 cells were grown in RPMI 1640 medium supplemented with 10% fetal bovine serum as reported before (17, 30). Lactating mammary glands (LMG) were obtained from C3H/Bi strain female mice. Macrophage cell line p388 and T-cell lymphoma S49 cells were both grown in RPMI 1640 medium supplemented with 10% fetal bovine serum (31).

DNA probes. EL-4 cell genomic DNA was digested with restriction enzyme *Bam*HI and fractionated, and the region of 1.2 to 1.3 kb was eluted for subcloning into plasmid vector pGem-4Z (Promega) by standard techniques (19). The MMTV *env* clone was selected, and a *Pst*I subfragment encompassing solely the *env* region was also subcloned. The 1.1-kb *Pst*I *gag* probe encompassing the 5′ half of the *gag* gene as well as 25 base pairs of leader sequence was a gift from E. Buetti (10).

RNA preparation. Total RNA was extracted from cells by the procedure of Okayama et al. (28). Pelleted cells were homogenized in 5.5 M guanidinium thiocyanate solution, followed by centrifugation through a cushion of cesium trifluoroacetate at 125,000 × *g* for 24 h at 15°C. The pelleted RNA was suspended in 10 mM Tris hydrochloride (pH 7.4)–1 mM EDTA and applied to a prepacked spin column containing oligo(dT)-cellulose (Pharmacia) according to the manufacturer's instructions to isolate polyadenylated [poly(A)$^+$] RNA. For preparing mRNA for cDNA library construction, the RNA samples were subjected to two rounds of oligo(dT)-cellulose selection. Analysis of poly(A)$^+$ RNAs was performed in vertical formaldehyde-containing 1.2% agarose gels (19). Transfer onto nitrocellulose membranes (Optibind; Schleicher & Schuell) and subsequent

hybridization with nick-translated probes were done according to the manufacturer's protocols. Standards for size estimation of separated mRNAs were an RNA ladder (Bethesda Research Laboratories) and RNA molecular weight marker I (Boehringer Mannheim).

cDNA library construction. Poly(A)$^+$ mRNA templates primed with oligo(dT)$_{12-18}$ were transcribed into cDNA copies by Moloney murine leukemia virus reverse transcriptase by using a cDNA synthesis kit (Pharmacia). Second-strand synthesis was performed by a modification of the procedure of Gubler and Hoffman (11), and blunt-ended cDNA was prepared for insertion into the *Eco*RI site of vector λgt11 (40) by ligation of *Eco*RI-*Not*I adaptors (Pharmacia) to the ends. The prepared cDNA was ligated into the phosphatase-treated arms of λgt11, followed by packaging into packaging extract Gigapack II (Stratagene). Recombinant bacteriophage libraries were screened on *Escherichia coli* Y1090$_{r-}$. Screening with antibodies was done by the method of Huynh et al. (15), with alkaline phosphatase-antibody conjugates (Promega) used for identification of positive clones. Screening of the cDNA libraries with nick-translated DNA probes was done by standard procedures (19).

PCR. The polymerase chain reaction (PCR) technique (33) was performed with reagents provided in the Gene Amp DNA amplification reagent kit from Perkin-Elmer Cetus. Oligonucleotide primers described in the Results section were purchased from Research Genetics, Huntsville, Ala. Reaction conditions were as follows: 2 min of denaturation at 94°C, annealing for 1 min at 65°C, and extension for 15 min at 70°C. The magnesium concentration was 1.5 mM, and 2.5 U of *Thermus aquaticus* DNA polymerase was added per 100-μl sample. Reactions were carried out in a Perkin-Elmer Cetus DNA thermal cycler for a total of 30 cycles. Samples of the λgt11 cDNA libraries were extracted with phenol-chloroform, and approximately 10 ng of phage DNA was used as a template.

DNA sequence analysis. For DNA sequence analysis, inserts were subcloned into vectors pGem-3z and -4z (Promega Biotech), and sequences were determined by the dideoxy technique (34) with T7 DNA polymerase (Pharmacia) in conjunction with the exonuclease III deletion subcloning technique (12). The sequences of both ends of inserts were also determined in the vector λgt11 by using λgt11 forward and reverse primers and *T. aquaticus* DNA polymerase (Promega).

RESULTS

MMTV RNA analysis. cDNA libraries of ML and EL-4 T-cell lymphomas were constructed in protein expression vector λgt11 and screened with antiserum raised against ML cells (32) in an effort to isolate cDNAs encoding tumor-specific antigens. When tested in an immunoprecipitation assay against [^{35}S]methionine-labeled ML cell extract, the antiserum precipitated the MMTV *env* and *gag* precursor polyproteins (32), as well as some additional unknown proteins (data not shown). The majority of antiserum-reactive clones isolated were cDNAs of MMTV *env* and *gag* transcripts. Further characterization of the MMTV transcript cDNA clones revealed that many *env* clones contained deletions, and all *gag* transcript clones were truncated. Specifically, two small *env* gene transcripts, approximately 1.0 and 1.9 kb in length, both containing long deletions, were isolated from the EL-4 cell cDNA library, and truncated *gag* transcripts consisting of only a 5′ segment of the *gag* gene were isolated from both the EL-4 and ML cell cDNA libraries. These initial observations prompted the present study of a detailed analysis of MMTV transcript synthesis in the T-cell lymphomas.

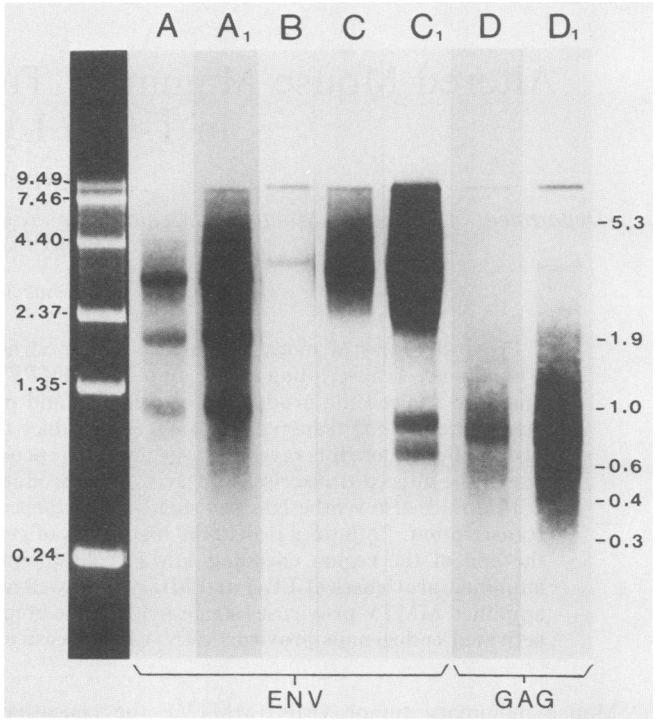

FIG. 1. Northern blot analysis of poly(A)$^+$ RNAs extracted from EL-4 cells (lane A), C3H mouse LMG (lane B), and ML cells (lanes C and D). Lanes A_1, C_1, and D_1 are longer exposures of lanes A, C, and D, respectively. The leftmost lane shows ethidium bromide-stained RNA ladder size markers and their lengths (in kilobases). The sizes shown on the right were obtained from RNA Molecular Weight Marker I. Lanes A, B, and C were hybridized with the 1.28-kb *Bam*HI probe. Lane D was hybridized with the 1.1-kb *Pst*I *gag* probe. Approximately 10 μg of poly(A)$^+$ RNA was loaded per lane.

To verify the presence of these transcripts in cellular RNA, the poly(A)$^+$ RNA used to construct the cDNA libraries was fractionated on agarose-formaldehyde gels, transferred to nitrocellulose membranes, and probed with *env* and *gag* gene probes. Cellular RNA was extracted by the procedure of Okayama et al. (28) to ensure minimum degradation, and poly(A)$^+$ RNA was selected on oligo(dT) spin columns, a rapid procedure which also contributes to the recovery of high-quality undegraded mRNA. Northern (RNA) blots were hybridized with a 1.28-kb *Bam*HI fragment probe, spanning the polymerase and *env* genes, a 1.1-kb *Pst*I *gag* probe (10), and a 1.4-kb *Pst*I LTR fragment. The *Bam*HI fragment was cloned from the amplified provirus of EL-4 cells in order to analyze the region encoding the 5′ junction sequences of the deletions in the EL-4 cell *env* transcripts.

Analysis of EL-4 cell mRNA with the *Bam*HI probe (Fig. 1, lane A) revealed the presence of three major *env* transcripts, a 3.0-kb mRNA, which probably represents the major spliced envelope transcript, and two smaller transcripts of about 1.9 and 1.0 kb, as predicted by the isolated cDNA *env* clones. The major *env* mRNA was only approximately 3.0 kb in size instead of the normal 3.5 kb because of the 491-base-pair deletion in the LTR of the amplified EL-4

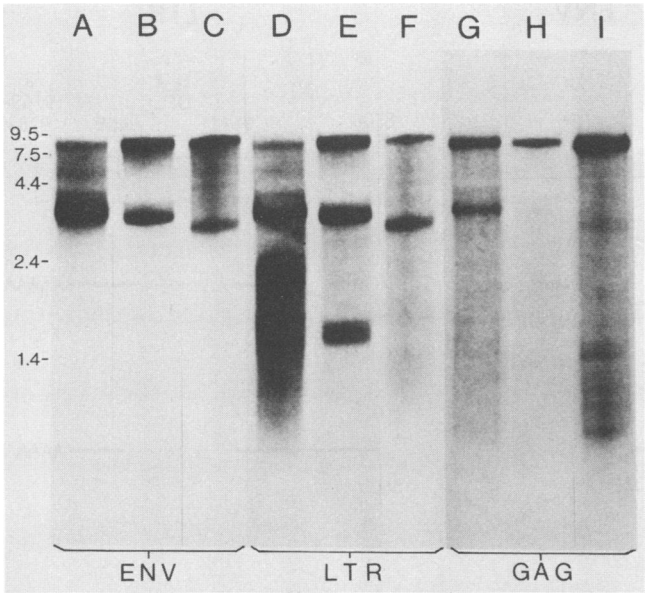

FIG. 2. Northern blot analysis of poly(A)$^+$ RNAs extracted from p388 monocyte-macrophage cells (lanes A, D, and G), C3H mouse LMG (lanes B, E, and H), and S49 T-cell lymphoma cells (lanes C, F, and I). Lanes A, B, and C were hybridized with the 1.28-kb *Bam*HI probe; lanes D, E, and F were hybridized with the 1.4-kb *Pst*I LTR probe; and lanes G, H, and I were hybridized with the 1.1-kb *Pst*I *gag* probe. Numbers on the left represent RNA ladder size markers (in kilobases). Approximately 10 μg of poly(A)$^+$ RNA was loaded per lane.

cell provirus (17). A longer exposure of the same blot showed the presence of the 8.1-kb full-length genomic RNA transcript (Fig. 1, lane A_1). Analysis of MMTV RNA isolated from LMG of C3H/Bi mice, the usual site for MMTV infection and replication, showed only the presence of 8.5-kb genomic and 3.5-kb envelope mRNAs (Fig. 1, lane B). No smaller *env* transcripts were detectable in the C3H LMG RNA, even after extended exposure of film (Fig. 2, lane B).

C3H LMG RNA analysis also demonstrated the integrity of the isolated RNA by the complete absence of degradation products. The C3H LMG RNA revealed the presence of the 1.7-kb LTR transcript (36, 39) when probed with the LTR probe (Fig. 2, lane E). ML cells were found to express a high level of *env* message (Fig. 1, lanes C and C_1) in addition to genome-length RNA, as well as two discrete smaller transcripts approximately 0.7 and 0.9 kb in length. The small *env* transcripts in ML cells were relatively much less abundant than the full-size *env* transcripts, which probably accounts for the failure to isolate cDNAs of these transcripts during the initial antiserum screening. Hybridization of ML cell RNA with the *gag* probe detected genome-length RNA (Fig. 1, lanes D and D_1) but revealed that the majority of transcripts were smaller than 1.0 kb. Although there seemed to be some heterogeneity, the predominant RNA forms appeared to be two closely spaced bands approximately 0.8 kb in length. After a longer film exposure, a faint band comigrating with *env* mRNA could be detected (Fig. 1, lane D_1); it probably reflects the presence of 25 base pairs of leader sequence in the 1.1-kb *Pst*I *gag* probe. The short *gag* transcripts were also detected in EL-4 cell RNA, although in lower amounts than in ML cells (data not shown).

Additional analysis was also performed on mRNA isolated from p388 cells, a monocyte-macrophage cell line containing amplified MMTV proviral copies with unaltered LTRs (31), and an additional T-cell lymphoma, S49 of BALB/c mice, which also harbors amplified MMTV proviruses with altered LTRs (8). Hybridization with the *env Bam*HI probe showed that both cell types synthesized genome-length and *env* transcripts (Fig. 2, lanes A and C). There was also a slight hybridization with a 2.3-kb band in the p388 cell RNA (Fig. 2, lane A). No smaller *env* transcripts were evident in the S49 cells (Fig. 2, lane C). Analysis with the LTR probe revealed the presence of at least two small transcripts in p388 cells, the 1.7-kb LTR mRNA in C3H LMG cells, and no additional mRNAs in S49 cells (Fig. 2, lanes D, E, and F). *gag* probe analysis indicated the probable presence of a heterogeneous population of short *gag* transcripts in the S49 cells (Fig. 2, lane I). The hybridization of the *env* message of p388 cells with the *gag* probe was probably caused by the 25 bases of leader sequence present in the probe and the very high relative level of *env* message in p388 cells (Fig. 2, lane G).

Sequence analysis of envelope transcript cDNAs. DNA sequence analysis of the two small *env* transcripts detected in EL-4 cells revealed that both species contained the correct 5′ leader sequence and splice site (10) but had long deletions in the *env* gene region (Fig. 3). Four cDNA clones of each were isolated from an unamplified cDNA library; two of each were sequenced and shown to be identical. Both RNA species also contained the identical deletion in the LTR region, previously shown to be present in the amplified provirus of EL-4 (14, 17). The shorter (1.05 kb) mRNA contained a deletion of 1,963 base pairs which extended into the LTR region (Fig. 3), while the 1.85-kb transcript had a 1,161-base-pair deletion encompassing the 3′ half of the *env* gene.

Sequences bordering the 5′ and 3′ deletion junction sites of the two transcripts are shown at the bottom of Fig. 3. The 1.28-kb *Bam*HI fragment extending from bases 6250 through 7536 of the amplified provirus of EL-4 cells was subcloned in order to determine the sequences surrounding the 5′ portion of the deletions. Inspection of the sequences bordering the 5′ end of the deletion in the 1.85-kb mRNA showed that they constituted a consensus splice donor site (26). The sequences shown near the 3′ end of the deletions are those of Kwon and Weissman (17), and they indicate that the 1.85-kb mRNA makes use of the splice acceptor site that is normally used to generate the 1.7-kb LTR transcript (36, 39). Thus, it appears that the deletion in the 1.85-kb mRNA is generated as a result of a splicing mechanism. The identical deletion has been observed to be present in a 1.0-kb phorbol ester-induced transcript that is initiated from a promoter within the *env* gene of the amplified provirus of EL-4 cells (9). The altered *env* transcripts were not found to be phorbol ester dependent for transcription.

The sequences bordering the deletion of the 1.05-kb transcript do not appear to fit into consensus splice donor and acceptor models. These *env* mRNAs do not appear to be transcribed from integrated proviruses bearing the same deletions, since analysis of EL-4 cell proviral DNA with restriction enzymes did not indicate the presence of proviral fragments consistent with such a possibility (9) (data not shown). Furthermore, if proviruses bearing *env* gene deletions were present, then one would expect to see multiple size species of genome-length RNA, which was not the case (Fig. 1).

PCR analysis. The two small *env* transcripts present in ML cells were characterized by making use of the PCR technique (33) and the λgt11 cDNA libraries. The analysis was carried out by using a primer consisting of the first 20 bases

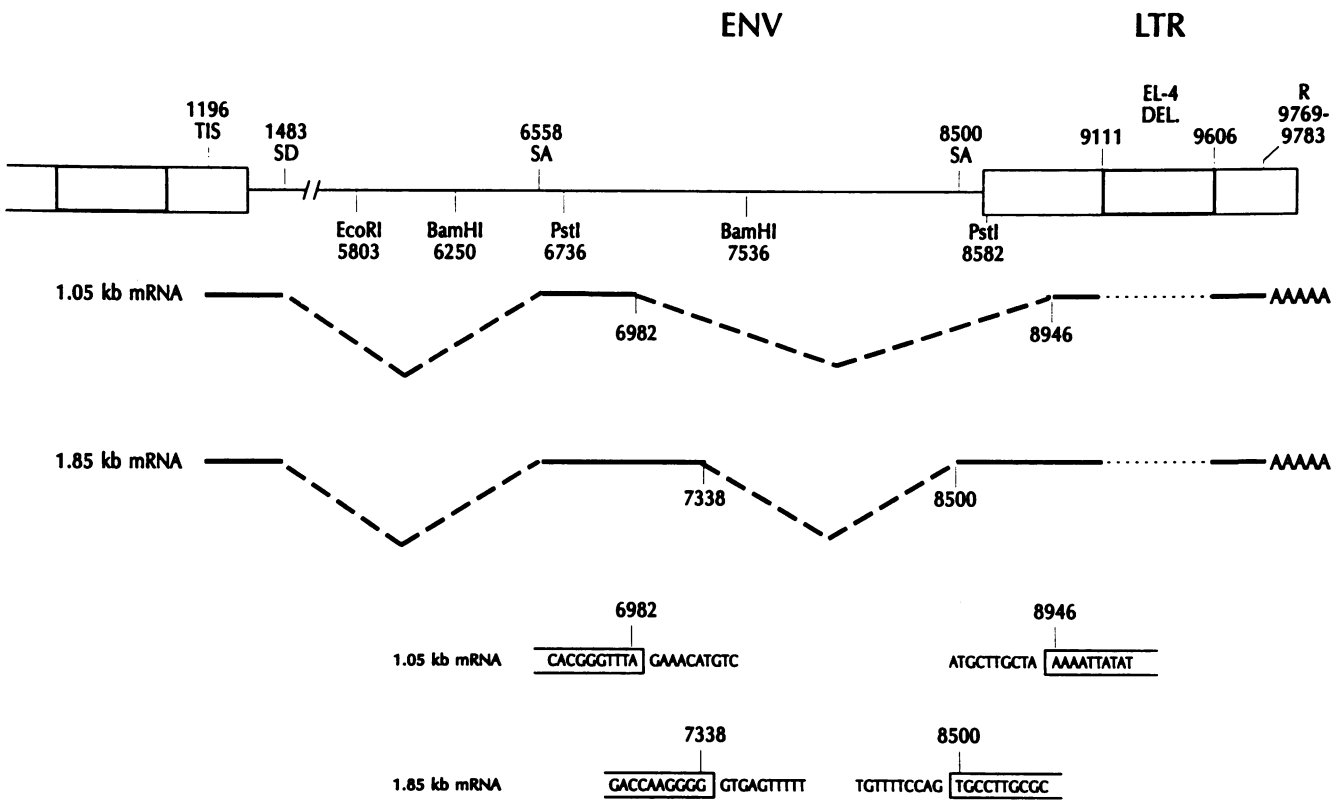

FIG. 3. Schematic representation of the amplified EL-4 cell provirus and of the two shortened *env* mRNAs. Base numbering corresponds to the analogous bases of the milk-transmitted MMTV of BR6 mice, as reported by Moore et al. (23). Loci identified above the proviral depiction are transcription initiation site (TIS), splice donor site (SD), splice acceptor site (SA), EL-4 LTR-deleted region (DEL), and R region of the 3′ LTR. Dashed lines indicate introns and deleted portions. Sequences bordering both ends of the deletions are shown at the bottom.

immediately to the right of the *env* gene splice acceptor site (GATGCCGAATCACCAATCTG) in combination with the reverse λgt11 sequencing primer (GTAATGGTCAACC AAACCACAGTT). In addition, an oligonucleotide containing an *Eco*RI restriction site (GGAGAATTC) was attached to the 5′ end of the *env* gene primer in order to facilitate subcloning of PCR products. Reaction with the ML cDNA library as the template generated two prominent bands, one approximately 450 and the other 650 base pairs long (Fig. 4, lane A). No product reflecting the full-length *env* transcript was observed because of the limitation of the PCR technique in generating molecules larger than approximately 2 kb in length. No small products were observed with the C3H LMG library template (Fig. 4, lane B). Reaction with the EL-4 cDNA library synthesized two prominent closely spaced bands of about 750 to 800 base pairs (Fig. 4, lane C).

The PCR products were digested with restriction enzyme *Eco*RI and separated on acrylamide gels, and the prominent bands were excised, eluted, and subcloned for sequencing. The two species from the ML library were sequenced in their entirety and turned out to be prematurely terminated *env* transcripts. Both had polyadenylated 3′ ends; the smaller one terminated at base 6940, and the larger one terminated at base 7161. Taking into account the 289-base-pair leader sequence, and excluding poly(A) tails, the predicted sizes of the two transcripts would be 673 and 893 base pairs, respectively. These lengths are in good agreement with the initial size estimates of 0.7 and 0.9 kb, which were based on size marker comparisons on Northern blots (Fig. 1, lane C_1). Just upstream of the termination sites of the two ML cell *env* transcripts were polyadenylation signal consensus sequences (2) (Fig. 5A and B).

Sequence analysis of the two prominent EL-4 cDNA *env* products revealed that the shorter one (750 base pairs) was identical to the region 3′ of the *env* splice acceptor site of the

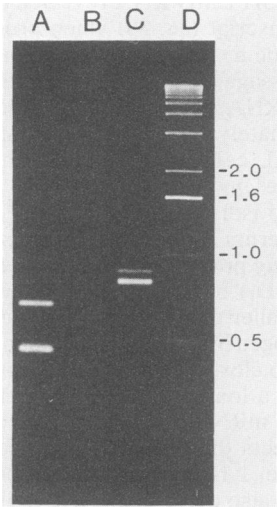

FIG. 4. Ethidium bromide-stained agarose gel of the PCR products of cDNA λgt11 libraries of ML cells (lane A), C3H LMG (lane B), and EL-4 cells (lane C). Lane D shows 1-kb ladder size markers identified (in kilobases) on the right.

```
A)                                                        6941
                                                          |
      CCGTGGGATG GGAAAGTACA GACCCCATTA AAGTTCTGAC CAATCTAAAA

B)                                                  7161
                                                    |
      TAATAAACCT GGGGGAAAGG GTGATAAAAG GCGTATGTGG AAAAAAAAAA

C)    1810
      |
      CGCTGAGGCC  AAGTCTGTGA  CTGAAGAAGA  GTTAGAAGAA  GGTTTAACCG
      GGCTACTATC  AACAAGTTCA  CAAGAAAAAA  CTTATGGGAC  CAGGGGAACA
      GCATATGCAG  AAATAGACAC  AGAGGTAGAC  AAATTGTCTG  AACACATTTA
      TGATGAACCA  TAAGAAGAAG  AAGAAGAAGA  AGAAGAAGAA  GAAGAAAT
                        |
                        1975

D)    1940
      |
      AAATTGTCTG  AACACATTTA  TGATGAACCA  TATGAAGAAA  AGGAGAAGGC
      AGATAAAAAT  GAGGAAAAGG  ACCATGTTAG  AAAGGTAAAG  AAAATAGTAC
      CAAAAAAAA
      |
      2040
```

FIG. 5. (A) 3′ end of the 344-base-pair PCR product of ML cell cDNA library. (B) 3′ end of the 544-base-pair PCR product of ML cell cDNA library. Potential polyadenylation signals are underlined. (C) Sequence of the *gag* cDNA isolated with antiserum. (D) 3′ end of the 843-base-pair *gag* cDNA isolated by hybridization. Base numbering corresponds to that of Moore et al. (24).

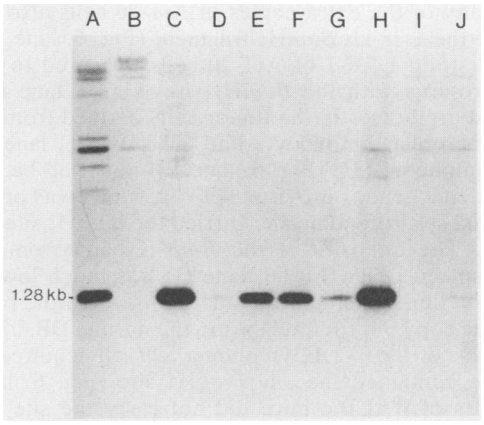

FIG. 6. Southern blot analysis of *Bam*HI-digested genomic DNA hybridized with the 1.28-kb *Bam*HI probe. Genomic DNA was isolated from (A) T-cell lymphoma EL-4, (B) C57BL/6 mouse liver, (C) AtT-20 pituitary tumor cells, (D) LAF1 mouse liver, (E) p388 monocyte-macrophage cell line, (F) T-cell lymphoma ML, (G) DBA/2 mouse spleen, (H) T-cell lymphoma S49, (I) BALB/c mouse liver, and (J) Leydig testicular tumor cell line I-10.

previously described 1.05-kb transcript, and the longer one (850 base pairs) contained the deletion of the 1.85-kb transcript but terminated at base 8553, just 5′ of the polypurine tract which precedes the 3′ MMTV LTR. The 850-base-pair EL-4 product probably represents an artifact of cDNA library construction, whereupon reverse transcription is primed by oligo(dT) on the polypurine tract, which contains 12 A residues interrupted by one G. A band of about 1.6 kb was visible in the EL-4 PCR product (Fig. 4, lane C) and probably arose from the 1.85-kb transcript cDNA. A faint band of approximately 2.0 kb was detectable in all three cDNA library PCR samples and probably represents the full-length *env* transcript cDNA primed on the polypurine tract.

***gag* gene transcripts.** Screening of cDNA libraries of EL-4 and ML cells with the anti-ML serum containing anti-*env* and anti-*gag* polyprotein activity identified a number of short cDNA clones (about 200 base pairs in length) from both libraries corresponding to sequences within the 5′ end of the *gag* gene region. Five cDNA clones were isolated (three from the DBA/2 and two from the EL-4 libraries) consisting of sequences encoding most of the phosphoprotein pp21 of the *gag* gene (13). The 3′ ends of all five isolates were identical, ending at base 1975, followed by the noncoding sequence GAA repeated 10 times (Fig. 5C). When the cDNA libraries were screened with the 1.1-kb *Pst*I *gag* fragment, the isolated clones extended into the leader sequence, and an additional group of four clones ending at base 2039 followed by a poly(A) tail were also isolated (Fig. 5D). Because of the difficulties encountered when subcloning sequences from this region of the MMTV genome into plasmid vectors (3), the isolated cDNA clones were only sequenced from both ends in the λgt11 vector. Assuming that both groups of transcripts initiate at the cap site, one can calculate a length of 779 and 843 base pairs for each type of transcript, not including the poly(A) or GAA tail. These sizes are in agreement with the estimates obtained by Northern blot analysis (Fig. 1, lane D). cDNAs of genome-length transcripts were, of course, not isolated since the entire transcript is too long to fit into the λgt11 vector (15, 40). Whether the truncated *gag* gene transcripts are translated in cells was analyzed by immunoprecipitation of labeled cell extracts with the antiserum used for library screening. The analysis was inconclusive for, although the 75,000-molecular-weight *gag* polyprotein precursor (6) was clearly identifiable, possible smaller translation products were not very prominent (data not shown).

***Bam*HI restriction sites in amplified proviruses.** As described above, the amplified provirus of EL-4 cells was found to contain a *Bam*HI site in the *env* gene, and digestion with *Bam*HI gave rise to an internal fragment of 1.28 kb encompassing base pairs 6250 through 7536. The presence of a *Bam*HI site in the *env* gene which yields the characteristic 1.28-kb internal *Bam*HI fragment is usually a characteristic of exogenous, horizontally transmitted MMTVs (4). In addition, the *env* gene *Bam*HI site is present in MTV-2 (22), the provirus of GR strain mice which is responsible for mammary tumor induction in those mice (20), and is also found to be amplified in T-cell lymphomas of GR strain mice (21–23). This *Bam*HI site is also present in MTV-1 (5), the endogenous provirus which causes late-onset mammary tumors in strains C3Hf and DBA/2f, in which the exogenous virus has been removed by foster nursing of offspring (20, 37). MTV-1 and MTV-2 are the only endogenous proviral loci known to induce mammary tumors (20). Of the three proviral genomes of C57BL/6 mice (the parental strain of EL-4 cells), none carry the *env Bam*HI site (29), suggesting that the amplified provirus in EL-4 cells might have originated from a horizontally transmitted exogenous infectious MMTV.

Southern blot analysis of additional non-mammary tumor cell types known to contain amplified MMTV genomes (31) and hybridization with the 1.28-kb *Bam*HI probe demonstrated that in every case, most of the acquired proviral genomes contained the *env* BamHI site (Fig. 6). Although most of the acquired proviruses in EL-4 cells contained the *env* BamHI site, four of them did not (Fig. 6, lane A). This observation confirms the reported high level of polymorphism present in the amplified proviral genomes (7). Simi-

larly, many of the extra copies in AtT-20 cells also did not generate the 1.28-kb *Bam*HI fragment (Fig. 6, lane C). The parental strain LAF-1 of AtT-20 cells appeared to carry a single provirus with the *Bam*HI site (Fig. 6, lane D). The amplified proviruses in the tumor cells derived from DBA/2 mice (macrophage-monocyte line P388 [Fig. 6, lane E] and T-cell lymphoma ML [Fig. 6, lane F]) might all be derived from the endogenous provirus MTV-1, which was present in the DBA/2 strain genome and carried the *Bam*HI site (Fig. 6, lane G). The intensity of the *Bam*HI band from DBA/2 normal spleen DNA (Fig. 6, lane G) was much lower than that of the tumor cell DNAs, since it reflected the presence of a single copy of proviral DNA in the normal DBA/2 tissue. In the case of the T-cell lymphoma S49, all acquired copies save one contained the *env Bam*HI site (Fig. 6, lane H). Proviruses of BALB/c mice did not carry the site (Fig. 6, lane I), and the one extra proviral copy present in Leydig testicular tumor line I-10 (31) did contain the *Bam*HI site (Fig. 6, lane J).

DISCUSSION

The synthesis of *env* transcripts containing deletions in EL-4 cells suggests the presence of splicing or other mechanisms that are not operational in mammary epithelial cells, the usual site of MMTV replication. Restriction enzyme analysis did not indicate the presence of proviruses bearing the analogous deletions and suggested a posttranscriptional modification mechanism. In the case of the 1.85-kb *env* mRNA of EL-4 cells, the 5′ junction site of the deletion contained a splice donor consensus sequence, while the 3′ junction site consisted of the splice acceptor site normally used in the synthesis of the spliced 1.7-kb LTR transcript (36, 39). The identical deletion has been reported to be present in the 1.0-kb phorbol ester-induced LTR transcript shown to be initiated from a promoter within the *env* gene (9). The small *env* transcripts were not, however, dependent on phorbol ester treatment for transcription. The borders of the deletion in the 1.05-kb mRNA did not conform to a splicing mechanism model, and the deletion might have been produced by some other process. The presence of small polypeptides in EL-4 cells immunoprecipitable with anti-gp52^{env} antiserum has been observed (30) and suggests that the shortened *env* mRNAs are translated. These transcripts obviously contain functional open reading frames, since their cDNAs were originally selected with antiserum containing anti-*env* activity. These small *env* proteins have never been observed in MMTV-producing epithelial cells (6) and are detectable after very short labeling periods, precluding their derivation from the processing of precursors. Similarly, small *env* translation products, not corresponding to normal MMTV structural proteins, have been detected in ML cells (32). The presence of a deletion-producing mechanism in these cells might offer a convenient explanation for the genesis of deletions observed in the LTRs of the amplified proviruses. One might further speculate that only the appropriate deletion in the LTR region would confer a selective advantage to a newly integrated MMTV provirus in a lymphoma.

The two smaller *env* transcripts observed in ML cells were found to be produced by premature termination of transcription. In both instances, potential polyadenylation signals (2) were located just upstream of the termination site. These sequences are, however, also present in normally transcribed proviruses, which suggests the influence of additional factors unique to the ML cells. Transcription termination requires additional signals (2), and it is also possible that base changes in the amplified proviruses (7) have altered these motifs or created new ones.

The truncated *gag* gene cDNAs were initially thought to be a phenomenon explainable by the so-called "poison" sequences (3) shown to be present in that region of the MMTV genome. The presence of transcript fragments in cDNA libraries is expected; however, they are almost always 3′-end fragments when oligo(dT) priming is used for library construction. The 5′-end region of the MMTV *gag* gene is known to contain sequences which inhibit the propagation of plasmid or phage vectors carrying these poison sequences (3). Recombinant λgt11 clones containing the short *gag* transcripts did indeed grow slowly and produced very small plaques. The short *gag* cDNA inserts were sequenced in the phage vector to avoid the need for subcloning into plasmids, a procedure shown to induce rearrangements and insertion of bacterial elements in the poison sequence region. These sequences have been reported to be especially troublesome when derived from exogenous, milk-transmitted MMTVs, to which the amplified EL-4 cell provirus appears to be related. Northern blot analysis of EL-4 and ML cell RNA, however, demonstrated the presence of high levels of *gag* transcripts of about 0.8 kb in size hybridizing with the *Pst*I *gag* probe, which is derived from the 5′ half of the *gag* gene. A similar analysis of S49 cells, a T-cell lymphoma of BALB/c mice, also indicated the presence of short *gag* transcripts.

The procedures used to purify and fractionate mRNA ensured the recovery of intact transcripts, as could be verified by the complete absence of RNA degradation products in the C3H LMG RNA. The much shorter 200-base-pair *gag* cDNAs isolated by antiserum screening probably reflect the selection of cDNAs with a continuous open reading frame needed for the formation of the β-galactosidase fusion products which are produced by the recombinant λgt11 vectors (15, 40). Therefore, cDNAs extending 5′ of the *gag* gene initiation codon cannot be used to form fusion products.

A striking feature of MMTV transcription in the EL-4 and ML cell lymphomas is the very high ratio of *env* to genome-length RNA synthesized. In the normal site of MMTV replication, the LMG, equivalent amounts of *env* and genome-length mRNAs were synthesized (Fig. 1, lane B). However, when the short *gag* transcripts were taken into account, these ratios in the lymphomas appeared to be more even. Since there do not appear to be proviral fragments in the genomes of the lymphoma cells, it is probable that the short *gag* transcripts are generated as a result of premature termination.

The amplified proviruses in each tumor line appear to have originated from a single transcript (9, 14). According to the analysis with restriction enzyme *Bam*HI, the original source of this transcript might have been infection with an exogenous virus or reinfection with activated provirus MTV-1 for DBA/2 strain-derived tumor cells. Although a great deal of polymorphism has been shown to occur in the amplified proviruses (7), the majority of the acquired proviral copies do contain the *env Bam*HI site, suggesting its presence in the original transcript. The LTRs of MTV-1 have been shown to contain a unique sequence region (5); unfortunately, this region is located within the deleted portion of the amplified provirus of ML cells (18) and cannot be used to identify the source of the amplified genome. In the case of T-cell lymphomas of GR mice, it has been shown that the amplified provirus is always derived from proviral locus MTV-2 (21,

23), the highly tumorigenic, endogenous virus which causes a 100% incidence of mammary tumors in female mice of the GR strain (20). Similarly, an amplified MMTV provirus with LTR alterations, found in a kidney adenocarcinoma line, bears the characteristic *Bam*HI site (38).

Whether MMTV is the actual etiological agent in the induction of all T-cell lymphomas with amplified proviruses is still open to question. The multiple proviral copies in these cells might simply reflect the end result of many reintegration events during the course of cell passage, whether in vivo or in vitro. In the case of the infectious T-cell-tropic B-type virus (1), which has a typical deletion in the LTRs, there is no question that it is the leukemogenic agent. There is also strong circumstantial evidence for the involvement of MTV-2, the virulent mammary tumor-inducing provirus of GR mice, in the induction of T-cell lymphomas in male mice of that strain (21–23). For carcinogen-induced lymphomas of RF/J mice, however (16), the amplification of MMTV proviral copies has been reported to be a tissue culture phenomenon.

ACKNOWLEDGMENTS

I am grateful to Harry Beyer for expert technical assistance and for the contributions made by Gary Zimmer.

This work was supported by Public Health Service grants CA-43864 and P30 CA-13330 from the National Cancer Institute.

LITERATURE CITED

1. **Ball, J. K., H. Diggelmann, G. A. Dekaban, G. F. Grossi, R. Semmler, P. A. Waight, and R. F. Fletcher.** 1988. Alterations in the U_3 region of the long terminal repeat of an infectious thymotropic type B retrovirus. J. Virol. **62**:2985–2993.
2. **Birnstiel, M. L., M. Busslinger, and K. Strub.** 1985. Transcription termination and 3′ processing: the end is in site! Cell **41**:349–359.
3. **Brookes, S., M. Placzek, R. Moore, M. Dixon, C. Dickson, and G. Peters.** 1986. Insertion elements and transitions in cloned mouse mammary tumour virus DNA: further delineation of the poison sequence. Nucleic Acids Res. **14**:8231–8245.
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 tissue of the mouse. Cell **16**:333–345.
5. **Crouse, C. A., and R. J. Pauley.** 1989. Molecular cloning and sequencing of the MTV-1 LTR: evidence for a LTR sequence alteration. Virus Res. **12**:123–138.
6. **Dickson, C., and G. Peters.** 1983. Proteins encoded by mouse mammary tumor virus. Curr. Top. Microbiol. Immunol. **106**: 1–34.
7. **Dudley, J., A. Arfsten, C. L. L. Hsu, C. Kozak, and R. Risser.** 1986. Molecular cloning and characterization of mouse mammary tumor proviruses from a T-cell lymphoma. J. Virol. **57**:385–388.
8. **Dudley, J., and R. Risser.** 1984. Amplification and novel locations of endogenous mouse mammary tumor virus genomes in mouse T-cell lymphomas. J. Virol. **49**:92–101.
9. **Elliott, J. F., B. Pohajdak, D. J. Talbot, J. Shaw, and V. Paetkau.** 1988. Phorbol diester-inducible, cyclosporine-suppressible transcription from a novel promoter within the mouse mammary tumor virus *env* gene. J. Virol. **8**:1373–1380.
10. **Fasel, N., E. Buetti, J. Firzlaff, K. Pearson, and H. Diggelmann.** 1983. Nucleotide sequence of the 5′ noncoding region and part of the *gag* gene of mouse mammary tumor virus: identification of the 5′ splicing site for subgenomic mRNAs. Nucleic Acids Res. **11**:6943–6955.
11. **Gubler, V., and B. J. Hoffman.** 1983. A simple and very effective method for generating cDNA libraries. Gene **25**:263–269.
12. **Henikoff, S.** 1984. Unidirectional digestion with exonuclease III creates targeting breakpoints for DNA sequencing. Gene **28**: 351–359.
13. **Hizi, A., L. E. Henderson, T. D. Copeland, R. C. Sowder, H. C. Krutzsch, and S. Oroszlan.** 1989. Analysis of Gag protein from mouse mammary tumor virus. J. Virol. **63**:2543–2549.
14. **Hsu, C.-L. L., C. Fabritius, and J. Dudley.** 1988. Mouse mammary tumor virus proviruses in T-cell lymphomas lack a negative regulatory element in the long terminal repeat. J. Virol. **62**:4644–4652.
15. **Huynh, T. V., R. A. Young, and R. W. Davis.** 1985. Constructing and screening cDNA libraries in lambda gt10 and lambda gt11, p. 49–78. *In* D. M. Glover (ed.), DNA cloning: a practical approach, vol. 1. IRL Press, Oxford.
16. **Kraft, R., S. T. Ishizaka, S. A. Okenquist, G. Childs, F. Lilly, and J. Lenz.** 1989. Absence of mouse mammary tumor virus proviral amplification in chemically induced lymphomas of RF/J mice. J. Virol. **63**:3200–3204.
17. **Kwon, B. S., and S. M. Weissman.** 1984. Mouse mammary tumor virus-related sequences in mouse lymphocytes are inducible by 12-*O*-tetradecanoylphorbol-13-acetate. J. Virol. **52**: 1000–1004.
18. **Lee, W. T., O. Prakash, D. Klein, and N. H. Sarkar.** 1987. Structural alterations in the long terminal repeat of an acquired mouse mammary tumor virus provirus in a T-cell leukemia of DBA/2 mice. Virology **159**:39–48.
19. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
20. **Michalides, R., R. van Nie, and R. Nusse.** 1981. Mammary tumor induction loci in GR and DBAf mice contain one provirus of the mouse mammary tumor virus. Cell **23**:165–173.
21. **Michalides, R., and E. Wagenaar.** 1986. Site-specific rearrangements in the long terminal repeat of extra mouse mammary tumor proviruses in murine T cell leukemias. Virology **154**: 76–84.
22. **Michalides, R., E. Wagenaar, J. Hilkens, J. Hilgers, B. Groner, and N. E. Hynes.** 1982. Acquisition of proviral DNA of mouse mammary tumor virus in thymic leukemia cells from GR mice. J. Virol. **43**:819–829.
23. **Michalides, R., E. Wagenaar, and P. Weijers.** 1985. Rearrangements in the long terminal repeat of extra mouse mammary tumor proviruses in T-cell leukemias of mouse strain GR result in a novel enhancerlike structure. Mol. Cell. Biol. **5**:823–830.
24. **Moore, R., M. Dixon, R. Smith, G. Peters, and C. Dickson.** 1987. Complete nucleotide sequence of a milk-transmitted mouse mammary tumor virus: two frameshift suppression events are required for translation of *gag* and *pol*. J. Virol. **61**:480–490.
25. **Morley, K. L., M. G. Toohey, and D. O. Peterson.** 1987. Transcriptional repression of a hormone-responsive promoter. Nucleic Acids Res. **15**:6973–6989.
26. **Mount, S. M.** 1982. A catalogue of splice junction sequences. Nucleic Acids Res. **10**:459–472.
27. **Nusse, R., L. van der Ploeg, L. van Duijn, R. Michalides, and J. Hilgers.** 1979. Impaired maturation of mouse mammary tumor virus precursor polypeptides in lymphoid leukemia cells, producing intracytoplasmic A particles and no extracellular B-type virions. J. Virol. **32**:251–258.
28. **Okayama, H., M. Kawaichi, M. Brownstein, F. Lee, T. Yokata, and K. Arai.** 1987. High-efficiency cloning of full-length cDNA: construction and screening of cDNA expression libraries for mammalian cells. Methods Enzymol. **154**:3–27.
29. **Peterson, D. O., K. G. Kriz, J. E. Marich, and M. G. Toohey.** 1985. Sequence organization and molecular cloning of mouse mammary tumor virus DNA endogenous to C57BL/6 mice. J. Virol. **54**:525–531.
30. **Racevskis, J.** 1986. Expression of the protein product of the mouse mammary tumor virus long terminal repeat gene in phorbol ester-treated mouse T-cell leukemia cells. J. Virol. **58**:441–449.
31. **Racevskis, J., and H. Beyer.** 1989. Amplification of mouse mammary tumor virus genomes in non-mammary tumor cells. J. Virol. **63**:456–459.
32. **Racevskis, J., and N. H. Sarkar.** 1982. ML antigen of DBA/2 mouse leukemias: expression of an endogenous murine mammary tumor virus. J. Virol. **42**:804–813.
33. **Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi,**

G. T. Horn, K. B. Mullis, and H. A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239**:487–491.

34. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
35. **Theunissen, H. J. M., M. Paardekooper, L. J. Maduro, R. J. A. M. Michalides, and R. Nusse.** 1989. Phorbol ester-inducible T-cell-specific expression of variant mouse mammary tumor virus long terminal repeats. J. Virol. **63**:3466–3471.
36. **van Ooyen, A. J. J., R. J. A. M. Michalides, and R. Nusse.** 1983. Structural analysis of a 1.7-kilobase mouse mammary tumor virus-specific RNA. J. Virol. **46**:362–370.
37. **Verstraeten, A. A., and R. van Nie.** 1978. Genetic transmission of mammary tumor virus in the DBAf mouse strain. Int. J. Cancer **21**:473–475.
38. **Wellinger, R. J., M. Garcia, A. Vessaz, and H. Diggelmann.** 1986. Exogenous mouse mammary tumor virus proviral DNA isolated from a kidney adenocarcinoma cell line contains alterations in the U3 region of the long terminal repeat. J. Virol. **60**:1–11.
39. **Wheeler, D. A., J. S. Butel, D. Medina, R. D. Cardiff, and G. L. Hager.** 1983. Transcription of mouse mammary tumor virus: identification of a candidate mRNA for the long terminal repeat gene product. J. Virol. **46**:42–49.
40. **Young, R. A., and R. W. Davis.** 1983. Efficient isolation of genes by using antibody probes. Proc. Natl. Acad. Sci. USA **80**: 1194–1198.