JOURNAL OF VIROLOGY, Feb. 1994, p. 1191–1195
0022-538X/94/$04.00+0

Vol. 68, No. 2

Characterization of a Marek's Disease Virus *Bam*HI-L-Specific cDNA Clone Obtained from a Marek's Disease Lymphoblastoid Cell Line

KAZUHIKO OHASHI, WENPING ZHOU, PRISCILLA H. O'CONNELL, AND KAREL A. SCHAT*

Department of Avian and Aquatic Animal Medicine, College of Veterinary Medicine, Cornell University, Ithaca, New York 14853

Received 30 July 1993/Accepted 29 October 1993

Two Marek's disease (MD) virus *Bam*HI-L-specific cDNA clones were isolated from a cDNA library constructed from poly(A)$^+$ RNA fractions of an MD lymphoblastoid cell line, MDCC-CU41 (CU41). These clones were mapped to the region corresponding to the *Bam*HI-Q_2 and -L regions. These clones hybridized with 2.5-, 0.8-, and 0.6-kb transcripts prepared from CU41. The transcriptional unit of the 0.6-kb transcript was determined by RNase protection assays. An open reading frame encoding a 107-amino-acid polypeptide was identified in the 0.6-kb transcript. Reverse transcriptase-PCR demonstrated the presence of this transcript in both CU41 and a reticuloendotheliosis virus-transformed cell line latently infected with MD virus.

Marek's disease (MD) is a lymphoproliferative disease of chickens caused by MD virus (MDV) (3). On the basis of its lymphotropism, MDV is classified as a gammaherpesvirus (14), yet the genome structure of MDV is similar to that of herpes simplex virus, an alphaherpesvirus (6). The MDV genome consists of two unique regions (US and UL) flanked by inverted repeats (IRS, IRL, TRS, and TRL). Several MDV homologs of herpes simplex virus genes have been identified and mapped to the US and IRS regions (1, 2, 15, 16).

Presently, the process of transformation of T cells by MDV is poorly understood. In lymphoblastoid cell lines established from MD tumors, viral gene expression is very limited, whereas many viral transcripts are found in lytically infected cells (21). This observation suggests that MDV oncogenicity and latency are closely related and that these few transcripts and their gene products may be important for the initiation and/or maintenance of latency and/or transformation by MDV. Recent studies (18, 22) demonstrated that the regions actively transcribed in tumor cell lines are the US and IRS regions, which correspond to the *Bam*HI H, I_2, L, and A fragments according to the restriction enzyme map of MDV (6).

Although the *Bam*HI-L region is highly transcribed in MDV-induced lymphoblastoid tumors and tumor cell lines, detailed information characterizing these transcripts is lacking. A cDNA library was prepared from poly(A)$^+$ RNA isolated from the MDV-transformed lymphoblastoid cell line MDCC-CU41 (CU41), using RNAzol B (Biotecx Laboratories Inc., Houston, Tex.) followed by purification using the PolyA Tract mRNA isolation system (Promega Corp., Madison, Wis.). CU41 was selected because we had previously demonstrated the presence of only six MDV-specific transcripts, including a 0.6-kb *Bam*HI-L transcript, in this cell line (18). The cDNA library was constructed by using the ZAP-cDNA synthesis kit (Stratagene, La Jolla, Calif.).

An MDV-specific cDNA clone, L1, was obtained through screening of the CU41 cDNA with gel-purified, ^{32}P-labeled MDV (strain GA) *Bam*HI L fragment (6), kindly provided by M. Nonoyama, Tampa Bay Research Institute, Tampa, Fla. A second clone, L2, was obtained by PCR amplification (7) of the cDNA library, using the L1 cDNA-specific primer (5′-CGT AATGGATCCCGTCC-3′; primer 2) and the T7 promoter primer (Promega Corp.) (Fig. 1). Southern blot analysis showed that the L1 cDNA clone hybridized with 1.0-kb *Bgl*II and 1.2-kb *Hpa*II subfragments of the *Bam*HI L fragment, whereas the L2 clone did not hybridize with any of these fragments (data not shown). Alignment of these cDNA sequences with the sequence of *Bam*HI-L or *Eco*RI-Q of the MDV genome (8) revealed the exact locations of the cDNA clones. The 3′ end of the L1 clone was localized at 221 bases to the right of the left *Bam*HI site of the *Bam*HI L fragment. The L2 clone was mapped to the region left of the *Bam*HI L fragment, which corresponds to the *Bam*HI-Q_2 region.

To characterize the transcripts coded for by the cDNA clones, radiolabeled probes were prepared from the L1 cDNA clone and used for Northern (RNA) blot analysis. Two micrograms of poly(A)$^+$ RNA fractions from CU41 cells and MDV strain RB1B-infected chicken kidney cells (CKC) (19) were used; the results are shown in Fig. 2. Three transcripts of 2.5, 0.8, and 0.6 kb were detected, which were more abundant in the RNA sample of CU41 than in that of RB1B-infected CKC. The L1 probe did not hybridize with the RNA sample from uninfected CKC. The probe prepared from the L2 cDNA clone also hybridized with these three transcripts (data not shown). Previous studies (18, 22) also demonstrated that the *Bam*HI-L-specific transcripts detected in tumor cell lines (CU41 and MKT-1) were 0.6, 0.8, and 2.4 kb in sizes. In addition, since the L1 and L2 probes hybridized with all three of these transcripts, they are transcribed, at least in part, from the same region in which these cDNA clones were mapped. These data suggest the existence of complex transcriptional patterns from this region.

The sequences of the cDNA clones L1 and L2 were determined by the dideoxy sequencing method of Sanger et al. (17), using either the Sequenase version 2.0 sequencing kit (United States Biochemical Corp., Cleveland, Ohio) or the Cyclist Exo$^-$ Pfu DNA sequencing kit (Stratagene). In addition, a genomic fragment, named Q2L1, of *Bam*HI-Q_2 of MDV strain

* Corresponding author. Department of Avian and Aquatic Animal Medicine, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853. Phone: (607) 253-3364. Fax (607) 253-3369.

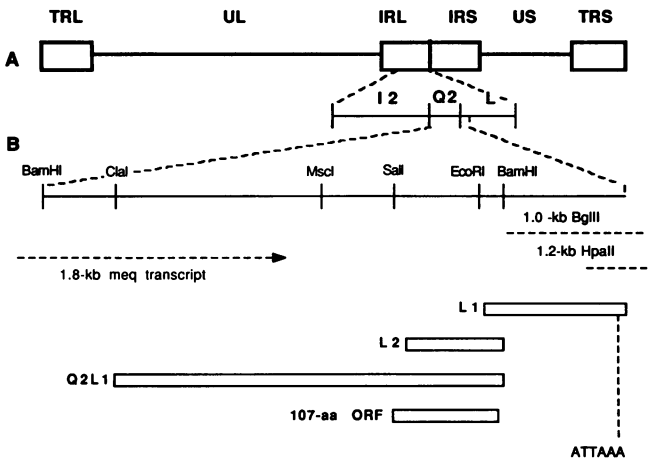

FIG. 1. Schematic representation of the genomic structure of MDV and locations of the *Bam*HI-L-specific cDNA clones. (A) Genomic structure of MDV with map positions of *Bam*HI-I_2, -Q_2, and -L regions (6). UL, unique long region; US, unique short region; IRL, long inverted repeat region; IRS, short inverted repeat region; TRL, long terminal repeat region; TRS, short terminal repeat region. (B) Restriction enzyme map of the *Bam*HI-Q_2 region and part of the *Bam*HI-L region and locations of *Bam*HI-L-specific cDNA clones (L1 and L2). Locations of the L1 and L2 cDNA clones, the Q2L1 PCR product, and an ORF with 107 amino acids (aa) are shown as open boxes. Locations of the 1.8-kb *meq* gene (8) and of the 1.0-kb *Bgl*II and 1.2-kb *Hpa*II subfragments of the *Bam*HI L fragment are shown as broken lines. The location of the polyadenylation signal (ATTAAA) in the L1 clone is also shown.

RB1B was amplified by PCR using primers 1 (5′-CAACATA AAGGAAAGGG-3′) and 2 (5′-CGTAATGGATCCCGTCC-3′), cloned into pBluescript (KS+), and sequenced. The primers for Q2L1 were designed from sequence data of the *Eco*RI Q fragment (8) and the L1 clone, respectively (Fig. 3). The sequence data were analyzed with the Genetics Computer Group software package (4). Sequence homology searches were carried out by using FASTA (11). The nucleotide sequences of the three clones were aligned and are shown in Fig. 3. Analysis of the sequences resulted in identification of one open reading frame (L1 ORF) which can encode a predicted polypeptide of 107 amino acids with a calculated molecular weight of 12,385. The potential translation initiation site (ATG) of the ORF was located at positions 725 to 727. A 237-base 3′ untranslated sequence was present in the cDNA, and a potential poly(A) signal, ATTAAA, was present at positions 1262 to 1267. This sequence has been shown to be effective as a poly(A) signal for the US8 and US9 transcripts in herpes simplex virus (13).

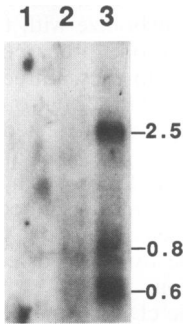

FIG. 2. Northern blot analysis to confirm the origins of the L1 cDNA clone. Poly(A)$^+$ RNA isolated from either uninfected CKC (lane 1), RB1B-infected CKC (lane 2), or CU41 (lane 3) was hybridized with a probe prepared from the L1 cDNA clone. Sizes of transcripts are shown in kilobases.

DNA sequence analysis of the upstream region of the L1 ORF resulted in identification of two potential TATA box-like sequences: TATAATT at positions 444 to 450 and TA-ATATATA at positions 665 to 673, from which the lengths of the predicted transcripts, without poly(A) tails, are approximately 0.8 and 0.6 kb, respectively. Both of these TATA elements are located downstream of the polyadenylation signal (AATAAG; positions 402 to 407) for the 1.8-kb *meq* transcript (8). Furthermore, some potential transcriptional regulatory sequences were identified in the region adjacent to the L1 ORF. An Oct-1 binding motif (ATGCAAAT) (5) was found at positions 648 to 655. In addition, three copies of similar sequences, ATGCACAGA (positions 656 to 664 and 679 to 687) and ATGCACCGA (positions 679 to 687), were found near the Oct-1 motif. However, further studies will be necessary to determine whether these sequences are indeed involved in the regulation of L1 transcription.

A comparison of the predicted amino acid sequence of the L1 ORF with known protein sequences deposited in GenBank has not shown any significant homology. Some potential sites for phosphorylation by protein kinase C (threonine residues at positions 11 and 51), casein kinase II (a serine residue at position 33), and tyrosine kinase (a tyrosine residue at position 32) (10) were found in this ORF.

RNase protection assays were carried out to obtain a transcriptional map of the origin of the cDNA clones, using the method of Melton et al. (9). Several subclones were obtained by cloning restriction enzyme-digested cDNA and PCR products into pBluescript (KS+) (Fig. 4). With these subclones used as templates, ^{32}P-labeled antisense RNA probes were synthesized with either T3 or T7 RNA polymerase and the Riboprobe system (Promega Corp.). Each probe (3×10^5 cpm) was annealed to a total RNA sample (50 μg) prepared from either CU41, the reticuloendotheliosis virus-transformed cell line RECC-CU197 (CU197) (20), RB1B-infected CKC, or uninfected CKC for 18 h at 42°C and then digested with RNase ONE (Promega). Protected fragments were resolved on 6% polyacrylamide–8 M urea sequencing gels and detected by autoradiography. Because of incomplete digestion by RNase after hybridization, some undigested probes were present in this experiment. However, it is clear that these four probes specifically hybridized with RNAs from CU41 and RB1B-infected CKC but not with those from uninfected CKC and CU197. The 514-base *Sal*I-*Hin*fI region (probes 2, 3, and 4), which corresponds to the L1 ORF and 3′ noncoding region of the L1 cDNA, was fully protected when hybridized with the RNA sample from either CU41 or RB1B-infected CKC. A probe covering the full 0.6-kb fragment was not used, but we do not believe that the observed protected fragments came from different transcripts, because there have been only three transcripts reported from this region (18, 22). Moreover, the L1 and L2 cDNA probes hybridized with these transcripts. In addition to a fully protected 193-base fragment, smaller fragments of two different sizes (115 and 42 bases) were detected when probe 1, corresponding to the *Msc*I-*Sal*I fragment, was hybridized with either of these RNAs. According to DNA sequence data (Fig. 3), a TATA box-like sequence was localized 55 bases upstream of the *Sal*I site, and the predicted length of an unspliced transcript from this TATA box is approximately 600 bases; this agrees with the result of North-

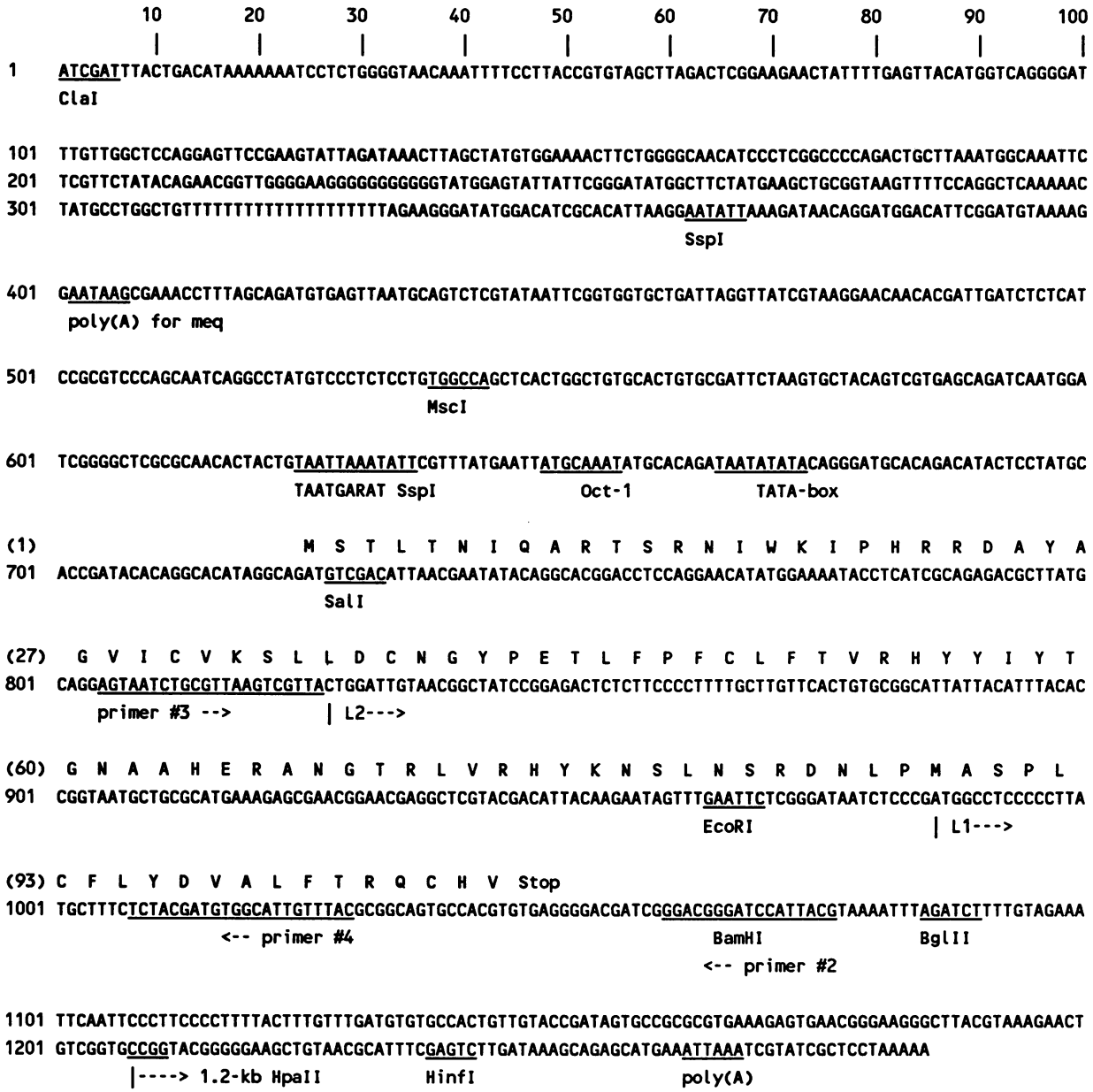

FIG. 3. Nucleotide sequence of the *Bam*HI-L-specific cDNA clone and its upstream genomic sequence. The upstream genomic sequence was obtained from analysis of the Q2L1 clone (PCR product). Alignment of these sequences is shown. The predicted amino acid sequence of the ORF is shown by the single-letter code above the nucleotide sequence, with amino acid numbers in parentheses. The poly(A) signals for the 1.8-kb *meq* transcript (8) and for the L1 transcript(s) are underlined and indicated as "poly(A) for meq" and "poly(A)," respectively. Restriction enzyme sites for *Cla*I, *Msc*I, *Ssp*I, *Sal*I, *Eco*RI, *Bam*HI, and *Bgl*II are underlined and labeled as such. A *Hin*fI site used for subcloning of the L1 cDNA clone is also underlined. The start of the 1.2-kb *Hpa*II subfragment is underlined and indicated. The vertical bars followed by "L1" and "L2" represent the starts of these cDNA clones. The locations of potential promoter sequences are also underlined and indicated (TAATGARAT, Oct-1, and TATA-box). The locations of primers (2, 3, and 4) used in RT-PCR are also shown.

ern blot analysis. The 42-base fragment was also detected when the 100-base *Ssp*I-*Sal*I fragment (Fig. 3) was used as a probe (data not shown). Therefore, we concluded that the 42-base protected fragment should correspond to the region immediately upstream of the *Sal*I site. We were not able to confirm which region corresponds to another protected fragment of 115 bases. Neither a splice acceptor site nor a potential TATA box sequence was found around 115 bases upstream of the *Sal*I site.

The initiation sites for the other two transcripts (2.5 and 0.8 kb) remain to be established. On the basis of sequence data for the *Eco*RI Q fragment and the cDNA clones (L1 and L2), the poly(A) signal for the L1 transcript was localized approximately 2.5 kb downstream of the potential transcriptional initiation site for the *meq* transcripts (8). Therefore, taken together with the result of RNase protection assays, this finding suggests that the 2.5-kb transcript detected by the L1 and L2 probes could be a readthrough transcript from the *meq*

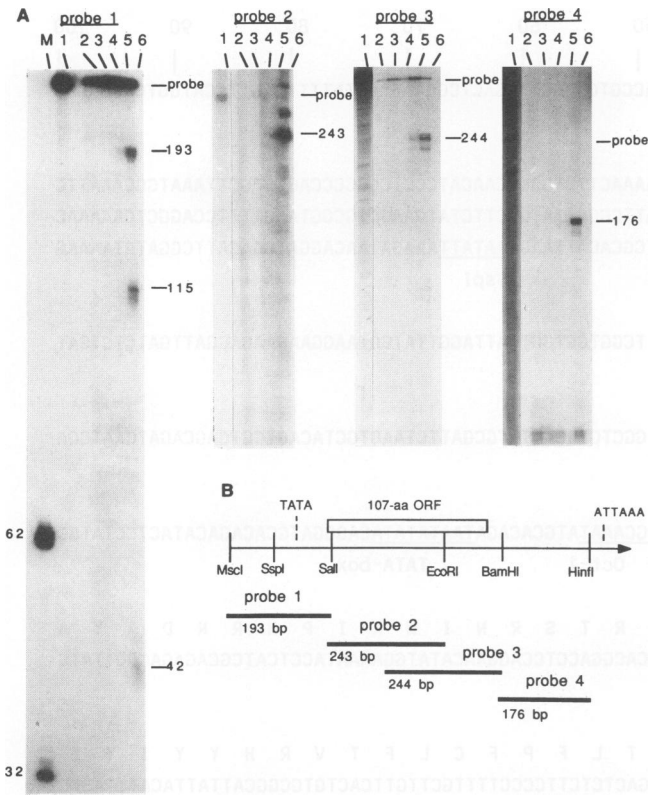

FIG. 4. RNase protection assays of RNAs from RB1B-infected CKC and CU41. (A) Each probe was hybridized with either tRNA (lane 2) or RNA from uninfected CKC (lane 3), from CU197 (lane 4), from RB1B-infected CKC (lane 5), or from CU41 (lane 6). Lane 1 represents each undigested probe. Sizes of the protected fragments are indicated in bases. (B) Locations of four probes used in this experiment. aa, amino acid.

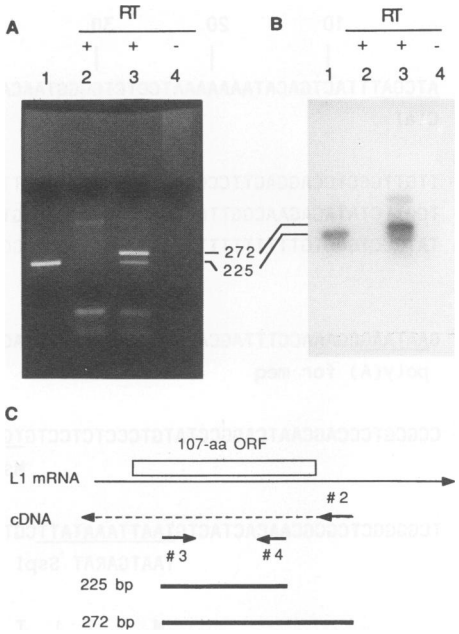

FIG. 5. RT-PCR of RNA prepared from CU210. Resultant PCR products were resolved on the agarose gel and analyzed by ethidium bromide staining (A) and Southern blot hybridization using the L2 cDNA clone as a probe (B). Lane 1 represents the PCR product derived from total DNA of CU41, using the primers 3 and 4. Lanes 2 and 3 represent the RT-PCR products derived from CU197 and CU210, respectively, using the same combination of primers. Lane 4 represents the result of PCR amplification of RNA prepared from CU210 (no reverse transcription). The presence of RT in the experimental reactions is indicated at the top. The sizes of PCR products are shown in base pairs. (C) Locations of primers (2, 3, and 4) used in this experiment and the predicted amplified fragments. aa, amino acid.

transcriptional unit. In addition, the 115-base protected fragment detected by probe 1 (Fig. 4) might correspond to the region spanning 115 bases downstream of the *Msc*I site, since another potential poly(A) signal (ATTAAA) was found 85 bases downstream of the *Msc*I site. In fact, a number of different sizes of the *meq* transcripts were detected in the previous study (8). Thus, there might be a relationship or coordination in gene expression between the *meq* gene and the origin of these *Bam*HI-L-specific cDNA clones. If this is the case, the 0.8-kb transcript might be generated by splicing so that the 5′ part of the 0.8-kb transcript would be derived from the *meq* gene, although the possibility that a TATA box-like sequence at positions 444 to 450 initiates transcription of this 0.8-kb messenger cannot be excluded. Both Northern blot analysis and RNase protection assays showed that the amount of the L1 transcript was higher in CU41 than in lytically infected CKC.

Reverse transcriptase (RT)-PCR was done to determine whether the L1 transcript is expressed in CU41 and in CU210, a reticuloendotheliosis virus-transformed cell line latently infected with MDV (12). RT-PCR was carried out by using the GeneAmp RNA PCR kit (Perkin-Elmer Cetus, Norwalk, Conn.) as instructed by the manufacturer. After treatment with DNase (Promega), 1-μg samples of total RNAs prepared from CU41, CU197, and CU210 were reverse transcribed with primer 2. Primers used for subsequent PCR amplification of the cDNA were primers 3 (5′-AGTAATCTGCGTTAAG TCGTTA-3′) and 4 (5′-GTAAACAATGCCACATCGTA GA-3′), which amplified a 225-bp fragment located within the L1 ORF from CU210 and CU41 but not from the control cell line CU197 (Fig. 5). An additional 272-bp fragment was also amplified in the same reaction. Since both the 225- and 272-bp bands hybridized with the L1 probe, it was concluded that the 272-bp fragment was generated as a result of the amplification between primers 3 and 2. The latter was used for reverse transcription and coexisted during the PCR (Fig. 5A). No DNA was amplified from CU210 or CU41 RNA in the absence of reverse transcriptase.

The functional importance of the L1 transcripts and their gene products, if any, remains unknown. The L1 ORF has been identified in the predicted L1 transcripts. A comparison of the predicted amino acid sequence with other known sequences in GenBank did not yield any significant homology. Pratt et al. (12) reported that a 0.6-kb *Bam*HI-L-specific transcript is expressed in a relatively large amount in CU210, a reticuloendotheliosis virus-transformed cell line latently infected with MDV. The RT-PCR data clearly showed that CU210 expresses the origin of the L1 and L2 cDNA clones. From the sizes of these transcripts, it is logical that the 0.6-kb *Bam*HI-L-specific transcript expressed in CU210 should be identical to the 0.6-kb L1 transcript. This also suggests that the L1 transcript might be associated with MDV latency in tumor cells. This hypothesis is further supported by the finding that the L1 transcripts are more abundant in CU41 than in lytically infected cells.

Nucleotide sequence accession number. The nucleotide sequence data reported here have been assigned GenBank accession number L19763.

REFERENCES

1. **Anderson, A. S., A. Francesconi, and R. W. Morgan.** 1992. Complete nucleotide sequence of the Marek's disease virus ICP4 gene. Virology **189:**657–667.
2. **Buckmaster, A. E., S. D. Scott, M. J. Sanderson, M. E. G. Boursnell, N. L. J. Ross, and M. M. Binns.** 1988. Gene sequence and mapping data for Marek's disease virus and herpesvirus of turkey: implications for herpesvirus classification. J. Gen. Virol. **69:**2033–2042.
3. **Calnek, B. W., and R. L. Witter.** 1991. Marek's disease, p. 342–385. *In* B. W. Calnek, J. L. Barnes, C. W. Beard, W. M. Reid, and H. W. Yoder, Jr. (ed.), Diseases of poultry. Iowa State University Press, Ames, Iowa.
4. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:**387–395.
5. **Falkner, F. G., R. Moickat, and H. G. Zachau.** 1986. Sequences closely related to an immunoglobulin gene promoter/enhancer element occur also upstream of other eukaryotic and prokaryotic genes. Nucleic Acids Res. **14:**8819–8827.
6. **Fukuchi, K., M. Sudo, Y. S. Lee, A. Tanaka, and M. Nonoyama.** 1984. Structure of Marek's disease virus DNA: detailed restriction enzyme map. J. Virol. **51:**102–109.
7. **Innis, M. A., D. H. Gelfand, J. J. Sninsky, and T. J. White.** 1990. PCR protocols: a guide to methods and applications. Academic Press, San Diego, Calif.
8. **Jones, D., L. Lee, J.-L. Liu, H.-J. Kung, and J. K. Tillotson.** 1992. Marek's disease virus encodes a basic-leucine zipper gene resembling the fos/jun oncogenes that is highly expressed in lymphoblastoid tumors. Proc. Natl. Acad. Sci. USA **89:**4042–4046.
9. **Melton, D. A., P. A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green.** 1984. Efficient in vitro synthesis of biologically active RNA and DNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res. **13:** 7035–7056.
10. **Pearson, R. B., and B. E. Kemp.** 1991. Protein kinase phosphorylation site sequences and consensus specificity motifs: tabulations. Methods Enzymol. **200:**63–77.
11. **Pearson, W. R., and D. J. Lipman.** 1988. Improved tools for biological sequence analysis. Proc. Natl. Acad. Sci. USA **85:**2444–2448.
12. **Pratt, W. D., R. W. Morgan, and K. A. Schat.** 1992. Characterization of reticuloendotheliosis virus-transformed avian T lymphoblastoid cell lines infected with Marek's disease virus. J. Virol. **66:**7239–7244.
13. **Rixon, F. J., and D. J. McGeoch.** 1985. Detailed analysis of the mRNAs mapping in the short unique region of herpes simplex virus type 1. Nucleic Acids Res. **13:**953.
14. **Roizman, B.** 1982. The family herpesviridae: general description, taxonomy, and classification, p. 1–23. *In* B. Roizman (ed.), The herpesviruses, vol. 1. Plenum Press, New York.
15. **Ross, L. J. N., and M. M. Binns.** 1991. Properties and evolutionary relationships of the Marek's disease virus homologues of protein kinase, glycoprotein D and glycoprotein I of herpes simplex virus. J. Gen. Virol. **72:**939–947.
16. **Sakaguchi, M., T. Urakawa, Y. Hirayama, N. Miki, M. Yamamoto, and K. Hirai.** 1992. Sequence determination and genetic content of an 8.9-kb restriction fragment in the short unique region and the internal repeat of Marek's disease virus type 1 DNA. Virus Genes **6:**365–378.
17. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
18. **Schat, K. A., A. Buckmaster, and L. J. N. Ross.** 1989. Partial transcription map of Marek's disease herpesvirus in lytically infected cells and lymphoblastoid cell lines. Int. J. Cancer **44:**101–109.
19. **Schat, K. A., B. W. Calnek, and J. Fabricant.** 1982. Characterization of two highly oncogenic strains of Marek's disease virus. Avian Pathol. **11:**593–605.
20. **Schat, K. A., W. D. Pratt, R. Morgan, D. Weinstock, and B. W. Calnek.** 1992. Stable transfection of reticuloendotheliosis virus-transformed lymphoblastoid cell lines. Avian Dis. **36:**432–439.
21. **Silver, S., A. Tanaka, and M. Nonoyama.** 1979. Transcription of the Marek's disease virus genome in a nonproductive chicken lymphoblastoid cell line. Virology **93:**127–133.
22. **Sugaya, K., G. Bradley, M. Nonoyama, and A. Tanaka.** 1990. Latent transcripts of Marek's disease virus are clustered in the short and long repeat regions. J. Virol. **64:**5773–5782.