JOURNAL OF VIROLOGY, Oct. 1990, p. 5188–5191
0022-538X/90/105188-04$02.00/0

Vol. 64, No. 10

Human Papillomavirus Type 6 Long Control Region and Human Cellular DNA Contain Related Sequences

TOM F. HRISOMALOS,[1] DENISE L. BOGGS,[1] AND KENNETH H. FIFE[1,2]*

Departments of Medicine[1] and Microbiology and Immunology,[2] Indiana University School of Medicine, Indianapolis, Indiana 46202-5124

Received 29 January 1990/Accepted 21 June 1990

We have identified a region of human papillomavirus type 6 (HPV-6) DNA that hybridizes with human cellular DNA containing no detectable HPV DNA sequences. The region of hybridization has been localized to a segment of the viral long control region between the end of the L1 open reading frame and the late polyadenylation signal and is likely contained within a 94-base-pair insertion at nucleotide 7350 which is present in the cloned HPV-6b DNA used for these studies. Restriction fragments of HPV-6 DNA from seven patients suggested that this insert was present in these naturally occurring viral genomes as well. The presence of this insert was confirmed by direct sequence analysis of polymerase chain reaction-amplified segments from four naturally occurring HPV-6 genomes. By analogy with other systems, this insert and surrounding sequences may function to destabilize the HPV-6 late mRNA.

The 7,900-base-pair (bp) human papillomavirus (HPV) genome can be divided into an early region, a late region, and a segment which contains no large open reading frames which has been called the long control region (LCR) or upstream regulatory region. The LCR contains regulatory sequences such as promoters, enhancers, and target sequences for *trans*-acting proteins which modulate HPV gene expression. The 5′ end of the LCR is a purine- and thymidine-rich region of variable length which contains the polyadenylation signal for the late region transcript. This region has also been shown to contain enhancer activity and to bind cellular proteins (1). A variant of HPV type 6 (HPV-6vc) has also been shown to have enhancer activity in the 5′ end of the LCR (12). Recently, it has been shown that cloning HPV-6b DNA in *Escherichia coli* may result in rearrangement or deletion of sequences at the 5′ end of the LCR (8). Thus, structural variations in the LCR from cloned HPV-6b DNA from a variety of sources which have been reported (3, 7, 9, 11) must be interpreted cautiously. Because of preliminary studies which suggested that a portion of the HPV-6b LCR may be related to cellular DNA sequences (18), we investigated these structural similarities in cloned and uncloned HPV-6b DNA.

Viral and cellular DNA hybridizations. Using the reverse blot technique in which radiolabeled cellular DNA is used as a hybridization probe for cloned HPV DNA immobilized on nitrocellulose (5, 18), we had previously observed the presence of low levels of cross-hybridization between the largest *Pst*I fragment of cloned HPV-6b DNA and human cellular DNA probes even under relatively high-stringency conditions. For the following experiments, the HPV-6 cloned DNA (referred to as pHPV-6b-RI in this report) was provided by E.-M. de Villiers and H. zur Hausen (Institut für Virusforschung, Heidelberg, Federal Republic of Germany) and consists of the HPV-6b genome cloned at the *Eco*RI site in pBR322 (4). Sequence coordinates for HPV-6b used in this paper are those from the original sequence determination (16). To better localize the region of cross-hybridization, pHPV-6b-RI DNA was digested with several relevant restriction enzymes, electrophoresed, transferred to nitrocellulose, and hybridized under moderately stringent conditions (T_m − 20°C) with radiolabeled human DNA from a cervical scrape specimen which was known to be negative for HPV

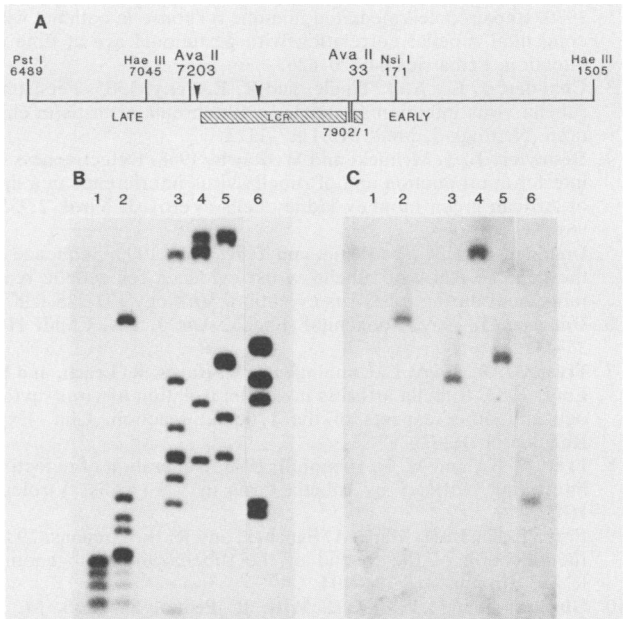

FIG. 1. (A) Locations of relevant restriction enzyme cleavage sites in and near the HPV-6 LCR. The arrowheads show the location of the polymerase chain reaction primers which flank the amplified sequence. The lower part of the figure demonstrates hybridization of human cellular DNA to subgenomic fragments of cloned HPV-6b DNA. (B) Transfer hybridized with pHPV-6b-RI DNA. (C) Transfer hybridized with human cervical cellular DNA (known to be negative for HPV DNA on Southern blot). Plasmid DNA was digested with restriction endonucleases as follows: lanes 1, pBR322 digested with *Hae*III; lanes 2, pHPV-6b-RI digested with *Hae*III; lanes 3, pHPV-6b-RI digested with *Nsi*I plus *Pst*I; lanes 4, pHPV-6b-RI digested with *Pst*I; lanes 5, pHPV-6b-RI digested with *Nsi*I; and lanes 6, pHPV-6b-RI digested with *Ava*II.

* Corresponding author.

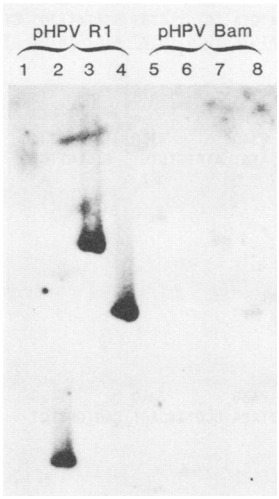

FIG. 2. Hybridization of human cellular DNA with subgenomic fragments of two different HPV-6b clones. DNA from pHPV-6b-RI and pHPV-6b-Bam was digested with the restriction endonucleases *Fnu*DII (lanes 1 and 5), *Ava*II (lanes 2 and 6), *Hae*III (lanes 3 and 7), or *Nsi*I plus *Pst*I (lanes 4 and 8), electrophoresed, transferred to nitrocellulose, and hybridized with radiolabeled HPV-negative DNA from cervical cells.

DNA by low-stringency Southern blotting (Fig. 1). By comparison of overlapping fragments, the region of cross-hybridization was found to be limited to within the borders of the *Ava*II E restriction fragment spanning nucleotides (nt) 7203 to 33. This *Ava*II fragment roughly corresponds to the LCR of HPV-6b. The lack of cross-hybridization between an *Fnu*DII restriction fragment spanning nt 7471 to 297 and the cellular DNA probe limited the potential region of cross-hybridization to the 5′ end of the LCR between nt 7203 and 7471 (data not shown). Similar experiments were done by using human leukocyte DNA as a further HPV-negative control, with the same results (data not shown).

To further localize the cross-hybridizing region, a second HPV-6b clone (pHPV-6b-Bam) was also examined for cross-hybridization with cellular DNA. This plasmid was provided by A. Roman and consists of the HPV segment from pHPV-6b-RI ligated and recloned at the *Bam*HI site of pBR322 (8). Restriction endonuclease digests and hybridizations with cellular DNA were conducted on both HPV-6b clones in parallel. Although cross-hybridization between normal human cellular DNA and fragments representing the 5′ end of the LCR was detected for both pHPV-6b-RI and pHPV-6b-Bam DNA, the degree of cross-hybridization at T_m − 20°C was markedly greater with the pHPV-6b-RI clone than with the pHPV-6b-Bam clone (Fig. 2). These two cloned DNAs have been shown to differ by the size of an insert at nt 7350 (54 bp for pHPV-6b-Bam versus 94 bp for pHPV-6b-RI) as well as a deletion of nt 7351 to 7399 in pHPV-6b-Bam (8). These data suggest that the region of cross-hybridization involves primarily the sequences between the 5′ 54 nt of the 94-bp insert of pHPV-6b-RI (which are identical to the 54-nt insert of pHPV-6b-Bam) and nt 7400 of the prototype HPV-6b.

Structure of the LCR in naturally occurring HPV-6 genomes. Because nearly all HPV-6 genomes characterized to date have been from cloned DNA and the region of HPV-6 DNA which was of interest on the basis of our hybridization experiments was the same region which has been shown to be subject to cloning artifacts (8), we examined uncloned HPV-6 DNA from clinical specimens to determine whether either cloned HPV-6b DNA was representative of the DNA associated with clinical lesions. DNA was purified from seven cervical scrape specimens, as previously described (18). These specimens had been shown by *Pst*I digestion and Southern blot hybridization to contain HPV-6b DNA (data not shown). DNA from these seven cervical scrape specimens was digested with *Ava*II, Southern blotted, and probed with ^{32}P-labeled pHPV-6b-RI (Fig. 3). All specimens had *Ava*II E fragments with sizes of approximately 850 bp, which is 120 bp longer than the length predicted from the published prototype sequence (nt 7203 to 33). The *Ava*II E fragment from these clinical specimens, like that of the HPV-6b-RI cloned DNA, is actually larger than the *Ava*II D fragment (nt 257 to 1075), and the E fragments from the clinical specimens comigrated with that of the HPV-6b-RI clone. The identity of the *Ava*II E fragment from the clinical specimens was confirmed by hybridization of a replicate blot with a gel-purified *Fnu*DII LCR fragment (nt 7471 to 297) from the HPV-6b-RI cloned DNA (data not shown).

DNA sequencing. While the previous results were consistent with the hypothesis that the pHPV-6b-RI cloned DNA was more representative of the "natural" HPV-6 sequence with respect to the 5′ end of the LCR, nucleotide sequence analysis of an authentic HPV-6 was needed to directly prove the hypothesis. Therefore, the polymerase chain reaction was employed, using *Thermus aquaticus* DNA polymerase (14) to amplify a predicted 260-bp segment of the 5′ end of the HPV-6b LCR (nt 7230 to 7490; Fig. 1). Because of concerns about sequence errors which may have been

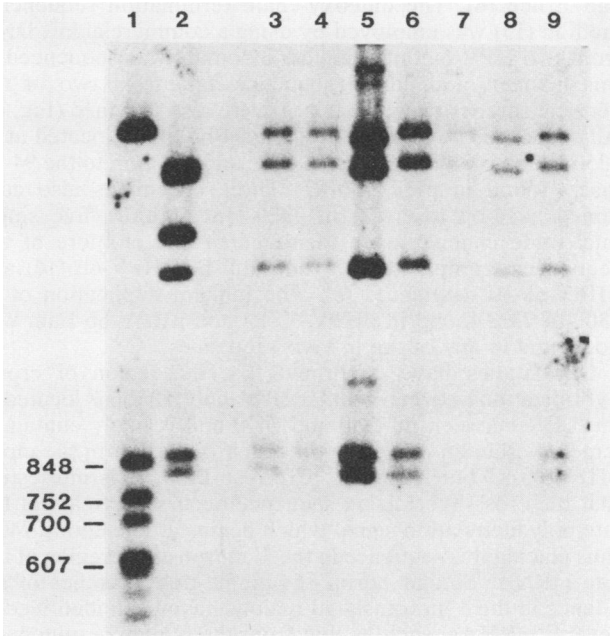

FIG. 3. Southern blot analysis of the size of the LCR fragment from HPV-6-containing clinical specimens. Lanes 1 and 2 contain pHPV-6b-RI DNA digested with *Hae*III and *Ava*II, respectively, and the sizes of fragments in the relevant range are shown to the left (in base pairs). Lanes 3 to 9 are *Ava*II digests of cervical cellular DNA from seven patients. These seven specimens were previously shown to contain HPV-6 DNA by Southern blot. The gel containing these digests was transferred and hybridized with labeled pHPV-6b-RI DNA.

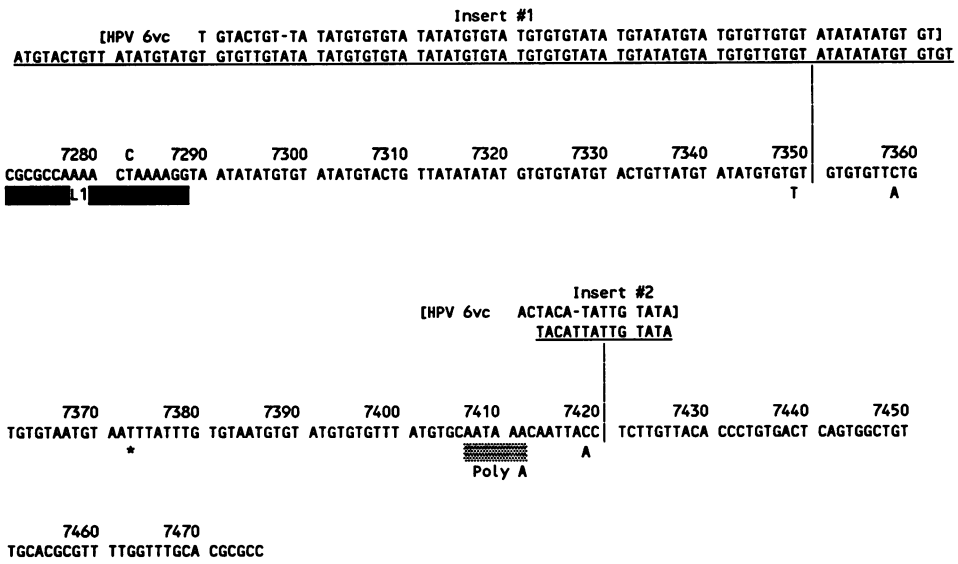

FIG. 4. Nucleotide sequences of the amplified segments of HPV-6 DNA from two condylomata acuminata and two cervical scrapings. The numbers used refer to the published HPV-6b nucleotide sequence (16). Four single nucleotide changes were seen in three of the four specimens and are shown above or below the sequence at the appropriate locations. The fourth specimen was the same as the published sequence at those locations. The nucleotide marked with an asterisk differs from the published sequence and was the same in all four specimens. Insert 2 was present in only three of the four specimens. For comparison, the sequences of analogous inserts reported in HPV-6vc DNA (11) are aligned above those we identified. Only the coding strand is shown; the end of the L1 open reading frame and the late polyadenylation signal are indicated.

introduced during the amplification reaction, sequencing was performed on the products of several independent amplification reactions. In addition, the sequencing templates for each strand were made in independent reactions by using the single-strand amplification reaction described by Gyllensten and Erlich (6). The dideoxy-chain termination sequencing method (15) was employed by using a commercial kit. DNA from two HPV-6-containing condylomata was sequenced in this manner, and partial sequence data from two of the cervical scrapes shown in Fig. 3 were also obtained (Fig. 4). All sequences contained the same 94-bp insert located at nt 7350 which is identical in sequence and location to the 94-bp insert found in pHPV-6b-RI. Three specimens also contained a 14-bp insert at nt 7420. In addition, five single nucleotide changes were identified in one or more of the sequences, compared with the published HPV-6b (16) and pHPV-6b-RI sequences (8). The tandem duplication of nt 7300 to 7323 found in pHPV-6b-RI and pHPV-6b-Bam was not found in any of the in vivo sequences.

Our studies have confirmed that the region of cross-hybridization between cellular DNA and HPV-6 is located in the LCR between nt 7230 and 7490 and is likely contained largely within an 89-bp region which is present in the clone pHPV-6b-RI but absent in pHPV-6b-Bam. It is interesting that the cross-hybridizing sequence occurs upstream of the late polyadenylation signal which begins at nucleotide 7407, thus placing this sequence in the 3′ untranslated region of the late mRNA. Similar purine-thymidine-rich stretches of sequence in the 3′ untranslated regions have been identified in several mRNAs, such as that from the human granulocyte-macrophage colony-stimulating factor gene (17). Insertion of such a sequence into the 3′ untranslated region of the rabbit β-globin gene has been shown to markedly destabilize the transcript and shorten its half-life (17). It is possible that this sequence in the HPV-6b genome serves to destabilize the late transcript under most situations, making production of virions unlikely. Such a mechanism has been proposed for regulation of late gene expression in bovine papillomavirus type 1 (10), although in that system, evidence of transcription attenuation in the late region has also been reported (2).

We have also shown that the HPV-6 genomes present in seven independent clinical specimens have *Ava*II E fragments larger than those predicted by the published HPV-6b sequence but similar to that present in pHPV-6b-RI DNA and that at least two of these specimens and two additional HPV-6 specimens contain a common 94-bp insertion. These observations emphasize the point made by Kasher and Roman (8) that cloned HPV-6 genomes may not accurately represent the sequence of the HPV-6 LCR. This discrepancy has been previously noted (3), but the exact nature and location of the difference have not been reported. This is especially critical for studies that attempt to correlate variations in the nucleotide sequence of the 5′ end of the LCR of cloned DNA with the biological behavior of the original tissue from which the DNA was cloned. For example, Rando and associates (11, 13) cloned an HPV-6 genome from an invasive vulvar carcinoma and identified several differences in the LCR of the DNA of this virus which they called HPV-6vc. They also identified enhancer activity in the region of two of the inserts in the 5′ end of the LCR (12). However, the 74-bp insert reported at position 7348 for HPV-6vc (11, 12) is nearly identical to part of the 94-bp insert we identified in four naturally occurring HPV-6 genomes (Fig. 4), and the 15-bp insert at position 7418 in HPV-6vc is nearly identical to the 14-bp insert we identified at position 7420. Thus, the HPV-6vc genome may vary much less from that of wild-type HPV-6 than was initially thought. The reference sequence of HPV-6b should probably be amended to include the 94-bp insert; the smaller insert does not appear to be uniformly present.

The sequences in Fig. 4 have been submitted to GenBank under accession numbers M35091 and M35092.

We thank E.-M. de Villiers, H. zur Hausen, and A. Roman for providing the cloned DNA used in these studies. We also thank C. H. Lee for helpful technical suggestions and A. Roman and D. Brown for a critical review of the manuscript.

LITERATURE CITED

1. **Auborn, K. J., R. L. Galli, T. P. Dilorenzo, and B. M. Steinberg.** 1989. Identification of DNA-protein interactions and enhancer activity at the 5′ end of the upstream regulatory region in human papillomavirus type 11. Virology **170**:123–130.
2. **Baker, C. C., and J. S. Noe.** 1989. Transcriptional termination between bovine papillomavirus type 1 (BPV-1) early and late polyadenylation sites blocks late transcription in BPV-1-transformed cells. J. Virol. **63**:3529–3534.
3. **Boshart, M., and H. zur Hausen.** 1986. Human papillomaviruses in Buschke-Löwenstein tumors: physical state of the DNA and identification of a tandem duplication in the noncoding region of a human papillomavirus 6 subtype. J. Virol. **58**:963–966.
4. **de Villiers, E.-M., L. Gissmann, and H. zur Hausen.** 1981. Molecular cloning of viral DNA from human genital warts. J. Virol. **40**:932–935.
5. **de Villiers, E.-M., A. Schneider, G. Gross, and H. zur Hausen.** 1986. Analysis of benign and malignant urogenital tumors for human papillomavirus infection by labelling cellular DNA. Med. Microbiol. Immunol. **174**:281–286.
6. **Gyllensten, U. B., and H. A. Erlich.** 1988. Generation of single-stranded DNA by the polymerase chain reaction and its application to direct sequencing of the HLA-DQA locus. Proc. Natl. Acad. Sci. USA **85**:7652–7656.
7. **Kasher, M. S., and A. Roman.** 1988. Characterization of human papillomavirus type 6b isolated from an invasive squamous carcinoma of the vulva. Virology **165**:225–233.
8. **Kasher, M. S., and A. Roman.** 1988. Alterations in the regulatory region of the human papillomavirus type 6 genome are generated during propagation in *Escherichia coli*. J. Virol. **62**:3295–3300.
9. **Kulke, R., G. E. Gross, and H. Pfister.** 1989. Duplication of enhancer sequences in human papillomavirus 6 from condylomas of the mamilla. Virology **173**:284–290.
10. **Lambert, P. F., C. C. Baker, and P. M. Howley.** 1988. The genetics of bovine papillomavirus type 1. Annu. Rev. Genet. **22**:235–258.
11. **Rando, R. F., D. E. Groff, J. G. Chirikjian, and W. D. Lancaster.** 1986. Isolation and characterization of a novel human papillomavirus type 6 DNA from an invasive vulvar carcinoma. J. Virol. **57**:353–356.
12. **Rando, R. F., W. D. Lancaster, P. Han, and C. Lopez.** 1986. The noncoding region of HPV-6VC contains two distinct transcriptional enhancing elements. Virology **155**:545–556.
13. **Rando, R. F., T. V. Sedlacek, J. Hunt, A. B. Jenson, R. J. Kurman, and W. D. Lancaster.** 1986. Verrucous carcinoma of the vulva associated with an unusual type 6 human papillomavirus. Obstet. Gynecol. **67**:70s–75s.
14. **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.
15. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
16. **Schwarz, E., M. Durst, C. Demankowski, O. Lattermann, R. Zech, E. Wolfsperger, S. Suhai, and H. zur Hausen.** 1983. DNA sequence and genome organization of genital human papillomavirus type 6b. EMBO J. **2**:2341–2348.
17. **Shaw, G., and R. Kamen.** 1986. A conserved AU sequence from the 3′ untranslated region of GM-CSF mRNA mediates selective mRNA degradation. Cell **46**:659–667.
18. **Webb, D. H., R. E. Rogers, and K. H. Fife.** 1987. A one-step method for detecting and typing human papillomavirus DNA in cervical scrape specimens from women with cervical dysplasia. J. Infect. Dis. **156**:912–919.