JOURNAL OF VIROLOGY, July 1997, p. 5161–5172 Vol. 71, No. 7
0022-538X/97/$04.00+0

Characterization of Late Gene Transcripts Expressed during Vegetative Replication of Human Papillomavirus Type 31b

MICHELLE A. OZBUN AND CRAIG MEYERS*

Department of Microbiology and Immunology, The Pennsylvania State University College of Medicine, Hershey, Pennsylvania 17033

Received 21 January 1997/Accepted 17 March 1997

Human papillomaviruses (HPVs) are etiologic agents of anogenital cancers. The lack of an efficient in vitro system with which to study the differentiation-dependent viral life cycle has impeded most investigations of viral transcription and gene expression. The CIN-612 clone 9E cell line latently maintains episomal copies of HPV type 31b (HPV31b). The complete replicative life cycle of HPV31b can be studied by using the organotypic (raft) culture system. A number of spliced HPV31b early gene transcripts and two late gene transcripts have been described in studies using the raft system. An HPV31b early promoter, P_{97}, and a differentiation-induced promoter, P_{742}, have been characterized by using this system. In this study, we used the raft system to analyze the temporal expression patterns of HPV31b late gene transcripts during the viral life cycle. The expression of late RNAs peaked at day 12 after lifting to the air-liquid interface; the levels then declined dramatically by day 16. The peak of late RNA expression was coincident with the appearance of virus particles in the raft tissues. We characterized transcripts with the potential to encode late gene products, including 19 RNAs containing the L1 region and 4 RNAs containing the E5b and L2 open reading frames. We also found evidence for two novel promoters. Transcription of both L1- and L2-containing RNAs initiated at a region upstream of the early promoter. In addition, late gene RNAs were also transcribed by using a promoter in the E4 reading frame.

Human papillomaviruses (HPVs) are small DNA viruses similar in genomic structure and organization (reviewed in reference 13). To date, over 75 types of human papillomaviruses have been described (30). HPVs have a tropism for epithelial tissues and are capable of inducing benign and malignant lesions (3, 13, 51). Of those known to infect the anogenital mucosa, certain HPV types frequently are associated with lesions which progress to invasive cancers (24). These are known as high-risk HPVs and include HPV type 16 (HPV16), HPV18, HPV31, and HPV33. Low-risk types, such as HPV6 and HPV11, rarely lead to malignant progression.

The complete replicative cycle of HPVs is tightly linked to the differentiation state of the cells that they infect (28, 47). According to current models, infection of the basal cell layer occurs through a microabrasion; the circular viral genome replicates episomally and is maintained at ≈50 to 100 copies per cell (3). Early genes are expressed under the control of the enhancer and promoter elements contained in the upstream regulatory region (URR) of the viral genome (13). Viral DNA replication is mediated by E1 and E2 proteins and occurs along with cellular DNA replication in the mitotically active basal layer, ensuring that both the parent and daughter cells maintain a constant number of viral genomes (23, 48). As cells migrate up through the epithelium, they undergo a complex program of differentiation. Concomitant with cellular differentiation in virally infected cells is the amplification of viral genomes, the expression of the late proteins, and the assembly of virions (2, 9, 23, 28).

Because the viral life cycle is dependent on cellular differentiation, it has been particularly difficult to cultivate and analyze various types of HPVs in the laboratory. High-risk HPV-containing cell lines have been established from both genital neoplasias and primary cells transfected with the complete viral genome (2, 5, 11, 17, 26, 34, 38, 41, 46, 50). These cell lines generally harbor HPV genomes that are integrated into the host cell DNA and therefore are impaired in the ability to carry out the complete viral life cycle (5, 17, 26, 34, 38, 41, 50). The structures of transcripts encoding high-risk HPV early genes have been characterized in cells containing high-risk HPV types, mostly by using monolayer cultures (8, 11, 15, 30, 31, 35, 40, 42–45). However, a typical result of integration is the disruption of late gene transcription, leaving many unanswered questions regarding late gene expression. In recent years, substantial progress has been made in the in vitro cultivation of high-risk HPVs; Meyers et al. purified virions from a cell line (CIN-612 9E) latently infected with HPV31b and propagated in the organotypic (raft) tissue culture system (28). Subsequently, Hummel and coworkers described a variety of polycistronic early and late HPV31b transcripts which display differential splicing (15, 16). They also characterized an HPV31b early promoter, P_{97}, and a differentiation-specific promoter, P_{742} (15).

We have used the organotypic culture system to further characterize HPV31b late gene transcripts expressed in CIN-612 9E tissues during the complete viral life cycle. Our temporal studies using raft tissues show that the expression of both L1- and L2-containing transcripts peaks at day 12 after lifting to the air-liquid interface and decreases substantially by day 16. The peak in transcription coincides with the first observations of viral particles by electron microscopy. We report four different transcripts containing the E5b and L2 open reading frames (ORFs) and 19 different transcripts containing the L1 ORF. The transcripts differ in their splicing patterns and initiation sites. Along with the differentiation-specific P_{742} promoter, we have identified two additional promoters, P_L and P_{E4}, which mediate the transcription of RNAs containing the late gene ORFs.

* Corresponding author. Mailing address: Department of Microbiology and Immunology, The Pennsylvania State University College of Medicine, Hershey, PA 17033. Phone: (717) 531-6240. Fax: (717) 531-4600. E-mail: cmeyers@bcmic.hmc.psu.edu.

TABLE 1. PCR primers used in analysis of HPV31b gene expression in CIN-612 clone 9E rafts

Name	Sequence[a]	Sense or antisense	ORF[b]	HPV31 nt[a]
URR 5′	5′-TTA GGT GTC ACG CCA TAG-3′	Sense	URR	7381–7398
E6 5′	5′-CCT GCA GAA AGA CCT CGG-3′	Sense	E6	120–137
E7 5′	5′-CCC GAC AGC TCA GAT GAG-3′	Sense	E7	644–661
742 5′	5′-CTA CAA TGG CTG ATC CAG-3′	Sense	E7E1	852–874
E1-3 3′	5′-GTC CCA TGG ACA TTT GTC G-3′	Antisense	E1	2053–2071
L2 5′	5′-GCG GTC CAA ACG CTC TAC AAA ACG GCA-3′	Sense	L2	4173–4198
L2-3 3′	5′-GTA GAG CGT TTG GAC CGC-3′	Antisense	L2	4173–4190
L2 3′	5′-CGC TAG GCC GCC ACA GAG ACA TCT G-3′	Antisense	L2	5549–5573
L1 5′	5′-GTC TCT GTG GCG GCC TAG CGA GGC-3′	Sense	L1	5554–5577
L1-2 3′	5′-TAG CAC TGC CTG CGT G-3′	Antisense	L1	5657–5672
L1 3′	5′-GCT GGT GTA GTG GTA GAT GCT GAG GG-3′	Antisense	L1	7016–7041

[a] Corresponding to the sequence and numbering of HPV31 (10).
[b] ORF or region of HPV31.

MATERIALS AND METHODS

Cell and tissue culture, histochemical analyses, and electron microscopy. The CIN-612 cell line was established from a cervical intraepithelial neoplasia (CIN) grade I (CIN I) biopsy and contains HPV31b DNA (2). In the CIN-612 clonal derivative 9E, the HPV31b genome is maintained episomally at ≈50 copies per cell (2, 15). CIN-612 9E cells were maintained in E medium with mitomycin-treated murine J2 3T3 fibroblast feeders (26, 27). Epithelial organotypic (raft) tissue cultures for in vitro differentiation were maintained as previously described (14, 26–28, 34). Epithelial cells were seeded onto collagen matrices containing J2 3T3 fibroblast feeders. When the epithelial cells had grown to confluence, collagen matrices were lifted onto stainless steel grids, and the epithelial cells were fed by diffusion from under the matrix. Epithelial tissues were allowed to stratify and differentiate at the air-liquid interface over a 16-day period. Rafts were treated with 10 μM 1,2-dioctanoyl-*sn*-glycerol (C8:0; Sigma Chemical Co., St. Louis, Mo.) in E medium every other day. Raft tissues were harvested at various time points starting with day 2 after lifting to the air-liquid interface and extending to day 16 after lifting. Harvested rafts were fixed in 10% buffered formalin and embedded in paraffin, and 4-μm cross sections were prepared. Sections were stained with hematoxylin and eosin (28). Electron microscopy was performed on cross sections of raft tissues as previously described (28).

Nucleic acid extraction and RNA PCR analyses. Total cellular DNA was harvested as previously described (2). Total RNA was extracted from rafts and subconfluent monolayer cultures by using TRIzol reagent (Gibco BRL, Bethesda, Md.). To remove copurifying viral and cellular DNA, the RNA samples were treated with DNase I (1).

DNase I-treated total RNA was reverse transcribed by using oligo$(dT)_{16}$ or random hexamer primers, and PCR was performed by using a GeneAmp RNA PCR kit as instructed by the manufacturer (Perkin-Elmer, Branchburg, N.J.). All PCR primers (Table 1) were synthesized by Operon Technologies (San Diego, Calif.) and were used at 0.5 μM. The thermocycling profile was as follows: 4-min time delay at 94°C; 35 cycles of 94°C for 30 s, 58 to 60°C for 1 min, and 72°C for 2 min; concluding with a 15-min extension at 72°C.

Northern blotting and hybridization. Total RNA samples (20 μg) were separated on 1% agarose–0.66 M formaldehyde gels as described previously (1). The RNA was transferred to GeneScreen Plus membranes (NEN Research Products, Boston, Mass.), which were handled according to the manufacturer's instructions. For PCR-generated probes, primers L2 5′ and L2 3′ (Table 1) were used to amplify a 1,401-bp fragment of the L2 ORF from pBR-HPV31 (a gift from H. zur Hausen), including HPV31 nucleotides (nt) 4173 to 5573; primers L1 5′ and L1 3′ were used to amplify a 1,488-bp fragment from the L1 ORF of pBR-HPV31, including HPV31 nt 5554 to 7041. DNA sequences were labeled with [α-^{32}P]dCTP (3,000 Ci/mmol; DuPont NEN), using a Random Primed DNA labeling kit (Boehringer Mannheim Corp., Indianapolis, Ind.) according to the manufacturer's instructions. Labeled probe was separated from unincorporated nucleotides by centrifugation through a Sephadex G-50 column (Boehringer Mannheim Corp.). Hybridization was carried out with 5×10^5 to 1×10^6 cpm per ml of appropriate probe, and the membranes were washed to remove nonspecific hybridization and then exposed to Reflection film with intensifying screens (DuPont NEN).

Cloning and sequencing. pGOZ31-L1 was constructed by inserting the 216-bp *Pst*I fragment of the pBR-HPV31 L1 region (HPV31 nt 6705 to 6921) in pGEM-3Zf(−) (Promega Corp., Madison, Wis.). PCR products were cloned by using a TA cloning kit (Invitrogen, San Diego, Calif.). pCR31-L2 was produced by TA cloning of the 1,401-bp PCR product generated with primers L2 5′ and L2 3′ on pBR-HPV31, as described above. The partial cDNA constructs pCR31b-E6L1, pCR31b-E7L1, pCR31b-2c5, pCR31b-742L1, and pCR31b-E6L2 were generated by TA cloning from RNA PCR products (see Fig. 5). Subsequent cloning was performed by using standard techniques (37). p31URRE was made by ligating the 795-bp *Acc*I-*Pst*I fragment from pBR-HPV31 (HPV31 nt 7238 to 121 [see Fig. 5]) into pGEM-3Zf(−). The URR was added upstream of each of the partial cDNA constructs pCR31b-E6L1 and pCR31b-E6L2 by ligating the *Acc*I-*Pst*I URR fragment from p31URRE to the *Pst*I site (HPV31 nt 121) located at the 5′ end of each of the cDNA plasmids. These constructs were designated p31U*E6L1 (Fig. 7, construct A) and p31U*E6L2 (Fig. 7, construct E), respectively. The "U*" denotes that the constructs contained the URR and E6* ORF, in addition to the E7,E1^ region, at their 5′ ends. The URR,E6*,E7 region (*Acc*I-*Sau*3AI fragment, nt 7238 to 858; [see Fig. 5]) from p31U*E6L1 (Fig. 7, construct A) was inserted upstream of the partial cDNA constructs pCR31b-E7L1, pCR31b-2c5, and pCR31b-742L1 (Fig. 5, constructs B, C, and D, respectively). Each cDNA plasmid was digested with an enzyme cleaving in the multiple cloning site (MCS) and with *Sau*3AI to remove the sequences 5′ to the *Sau*3AI site and to facilitate the ligation of the *Acc*I-*Sau*3AI URR,E6*,E7 region from p31U*E6L1 in their place. The resulting constructs were designated p31U*E7L1, p31U*2c5, and p31U*742L1 (Fig. 7, constructs D, C, and B), respectively. The constructs shown in Fig. 7 collectively were named the URRE6*-late constructs.

URR-late gene constructs which contained the full E6 ORF, in place of the E6* region, were also created. p31b-URRE1 was made by TA cloning the 2,604-bp PCR product generated with primers URR 5′ and E1-3 3′ (see Fig. 5) and total DNA from C8-treated CIN-612 9E rafts harvested at 8 days. p31b-URRE1 was digested with an enzyme which cut in the 5′ MCS and with *Ban*II to release the URR,E6,E7 region (HPV31 nt 7381 to 811). Each of the URRE6*-late constructs also was digested with an enzyme specifically cleaving within the 5′ MCS and with *Ban*II to remove the URR,E6*,E7 sequences. The URR,E6,E7 region from p31b-URRE1 was ligated to each late cDNA sequence at the *Ban*II sites, and the resulting constructs were named p31b-UE6L1, p31b-UE7L1, p31b-U2c5, p31b-U742L1, and p31b-UE6L2, respectively (not shown); the plasmids were designated the URRE6-late constructs.

Double-stranded DNA sequencing was performed by the dideoxy method according to the protocol for Sequenase version 2.0 (United States Biochemical, Cleveland, Ohio). The reaction products were separated on 7% polyacrylamide–8 M urea sequencing gels.

RNase protection assays (RPAs). Antisense RNA probes were synthesized by using [α-^{32}P]CTP (800 Ci/mmol; DuPont NEN) and a MAXI script/RPA II kit as instructed by the manufacturer (Ambion Inc., Austin, Tex.). pCR31-L2 was digested with *Dde*I, which yielded a 300-nt antisense probe predicted to protect 221 nt of the L2 ORF. pGOZ31-L1 was linearized with *Hin*dIII, which gave a 271-nt riboprobe expected to protect 216 nt of the L1 ORF. pCR31b-E7L1 was digested with *Hin*fI, giving rise to a 336-nt riboprobe predicted to protect 257 nt of the E7,E1^E4*,L1 transcript. *Pvu*II digestion of pCR31b-E6L1 yielded a 383-nt antisense probe. This riboprobe was predicted to protect 315 nt from the E7,E1*II,L1 transcript and 194 nt of transcripts containing the region E7,E1^. Full-length probes were gel purified on 7 M urea–5% polyacrylamide gels. The probes were eluted from gels in 0.5 M ammonium acetate–1 mM EDTA–0.1% sodium dodecyl sulfate for 3 to 12 h at 37 to 50°C. By using the RPA II kit, 15 to 20 μg of total RNA or yeast RNA was hybridized with 4×10^4 cpm of the appropriate probe at 43°C for 16 to 20 h. Unhybridized RNA was digested with 500 U of RNase T_1 per ml. Samples were analyzed by electrophoresis through a 7 M urea–5% polyacrylamide gel followed by autoradiography. RNA Century standards were prepared as recommended by the manufacturer (Ambion). The intensity of protected fragments was measured by scanning laser densitometry. RNA concentrations were based on optical densities; concentrations were verified by electrophoresis in agarose gels and staining with ethidium bromide.

S1 and exoVII protection assays. Probes for nuclease S1 (S1) and exonuclease VII (exoVII) nuclease protection analyses were prepared by PCR amplification using a 5′-end-labeled primer complementary to the sense strand of DNA (either L1-2 3′ or L2-3 3′) and an unlabeled primer complementary to the antisense

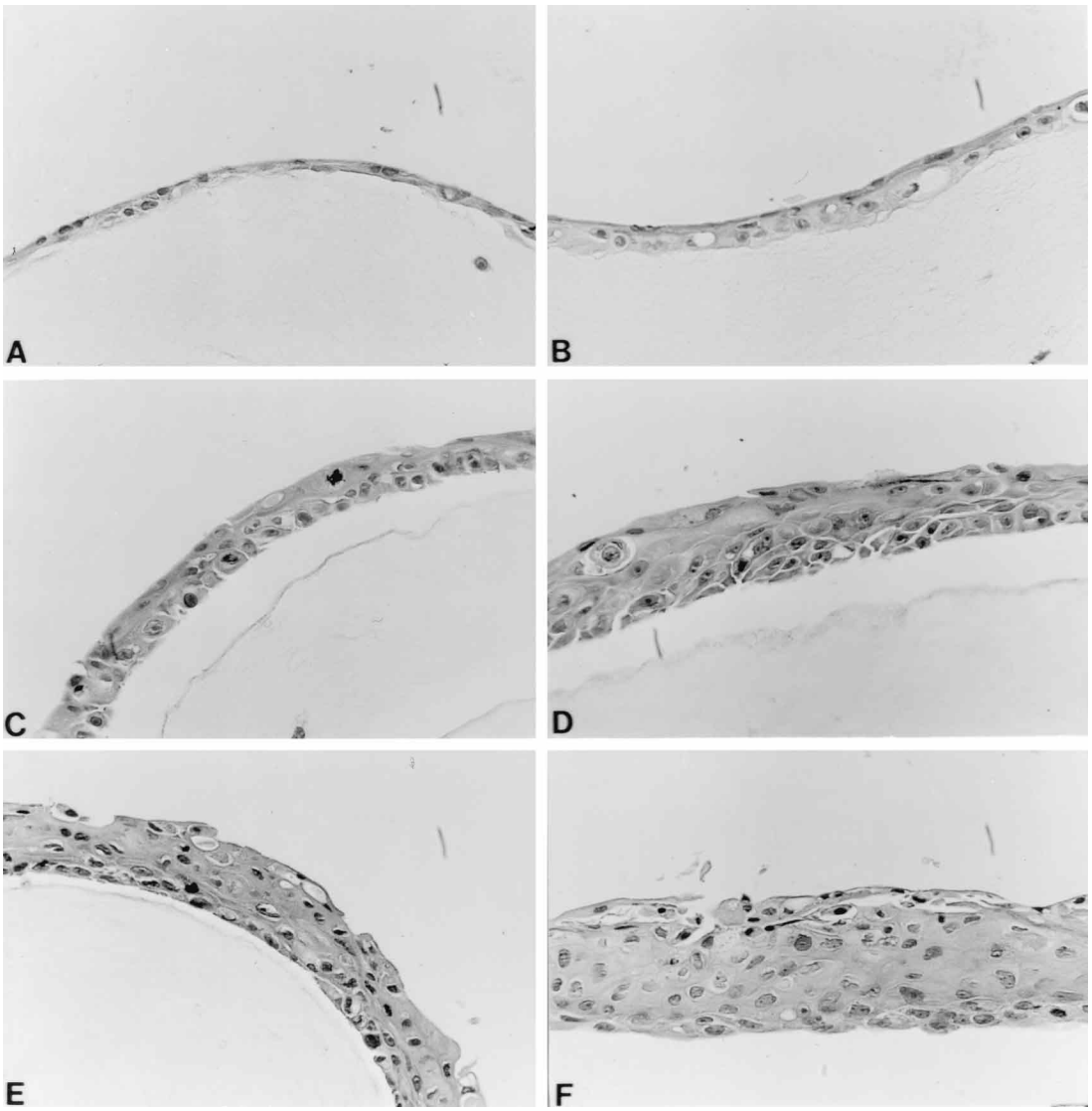

FIG. 1. Hematoxylin and eosin staining of CIN-612 9E tissues grown in the organotypic culture system. CIN-612 clone 9E cells were cultured on rafts which were treated with 10 μM C8:0 every second day. Rafts were harvested following lifting to the air-liquid interface after 2 days (A), 4 days (B), 6 days (C), 8 days (D), 11 days (E), and 13 days (F).

DNA strand [either M13(−40), 5′-GTTTTCCCAGTCACGAC-3′ (complementary to plasmid sequences 92 to 76 nt upstream of the HPV31 URR sequences), or URR 5′ (Table 1)]. Twenty picomoles of primer was 5′ end labeled by using 120 μCi of [γ-^{32}P]ATP (6,000 Ci/mmol; DuPont NEN) in 70 mM Tris-Cl (pH 7.6)–10 mM $MgCl_2$–5 mM dithiothreitol–50 μg of bovine serum albumin per ml–15 U of T4 polynucleotide kinase (New England Biolabs, Beverly, Mass.). Using 10 ng of each of the URRE6*-late or URRE6-late constructs as the template, 30 cycles of PCR were performed as described above. Single-stranded, 5′-end-labeled probes were purified by electrophoresis through 7 M urea–3.5% polyacrylamide gels. Hybridizations and S1 digestions were performed with an S1-Assay kit (Ambion) as instructed by the manufacturer. *Exo*VII digestions were performed as described previously (4). Forty micrograms of total RNA or yeast RNA was hybridized with 10^4 cpm of each probe at 43°C for 16 to 20 h. Unhybridized nucleic acids were digested with 250 U of S1 (Ambion) per ml or with 80 U of exoVII (Gibco BRL) per ml. Samples were analyzed by electrophoresis through a 7 M urea–4% polyacrylamide gel followed by autoradiography.

RESULTS

Temporal stratification of CIN I epithelium in the organotypic culture system. In the organotypic (raft) culture system, the differentiation of cells isolated from biopsies of CIN lesions or cervical carcinomas morphologically mimics that of their in vivo counterparts (2, 5, 27, 28, 31, 34). Addition of protein kinase C (PKC) pathway activators, such as C8:0, to the raft culture medium induces a more complete differentiation program (28, 31). To demonstrate the temporal stratification of the CIN I-derived CIN-612 9E cell line in the organotypic culture system, C8:0-treated CIN-612 9E rafts were harvested for paraffin embedding on various days following lifting to the air-liquid interface. Thin sections of the raft cultures were histochemically stained with hematoxylin and eosin. On day 2, a single layer of epithelial cells was observed (Fig. 1A). By day 6, the spinous layer was evident and a thin corneum was beginning to develop (Fig. 1C), and on day 8, the granular layer became visible (Fig. 1D). Over the 2-week period, the tissue thickened, forming a stratified epithelium (Fig. 1) closely resembling tissue from a CIN I lesion (2, 5, 20, 34). The lower third of the epithelium contained basal-like cells, whereas differentiation occurred in the upper epithelium. In addition, koilocytic cells were observed in the upper layers beginning at

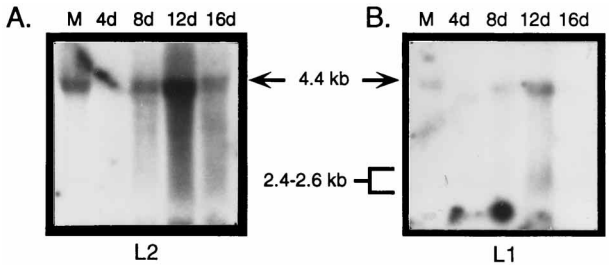

FIG. 2. Northern analyses of HPV31b L1 and L2 transcripts. CIN-612 clone 9E cells were cultured as monolayers (M) or as raft tissues; the rafts were treated with 10 μM C8:0 every second day. Rafts were harvested following lifting to the air-liquid interface after 4 days (4d), 8 days (8d), 12 days (12d), and 16 days (16d). Twenty micrograms of each sample of DNase I-treated total RNA was analyzed. Probes were specific to either the L2 ORF (HPV31 nt 4173 to 5573) (A) or the L1 ORF (HPV31 nt 5554 to 7041) (B). Sizes of transcripts were calculated by comparison to RNA markers. The blots were exposed for >6 days at −70°C.

day 8 (Fig. 1D). These characteristics are typical of CIN I morphology associated with HPV infection (20).

Temporal expression of L1- and L2-containing transcripts in untreated rafts and rafts induced to differentiate and support the vegetative life cycle of HPV31b. The more complete differentiation of CIN-612 9E rafts upon PKC activation is accompanied by a strong induction of HPV31b late gene expression and the assembly of virions (9, 16, 28). We analyzed the temporal expression of L1 and L2 transcripts by performing Northern blot hybridization on total RNA harvested from untreated CIN-612 9E monolayer cells and from C8:0-treated CIN-612 9E rafts harvested at 4, 8, 12, and 16 days after lifting to the air-liquid interface (Fig. 2). Consistent with previous reports (2, 16), we detected two size classes of HPV31b late gene transcripts. Both L1 and L2 probes identified transcripts of 4.4 kb. A diffuse band corresponding to a 2.4- to 2.6-kb size class was detected with the L1 probe in the RNA from 12 day rafts (Fig. 2B, lane 12d). In the treated rafts, after little expression on day 4, expression of both L1 and L2 transcripts peaked at day 12. It was striking that the levels of L1 and L2 RNAs decreased dramatically by day 16 in the raft system. As reported by Bedell and coworkers (2), we detected the larger late gene transcripts in the lanes containing RNA from untreated monolayers. We did not detect the smaller L1 transcripts in the monolayer cultures; however, it was unclear whether this inability reflected the sensitivity of the assay. Coincident with the peak of late gene expression between 8 and 12 days, we observed viral particles as early as day 10 by electron microscopy of raft tissue cross sections (Fig. 3).

To quantitate the relative changes in temporal L1 and L2 mRNA expression among CIN-612 9E untreated monolayer cultures, untreated rafts, and C8:0-treated rafts, samples were analyzed with antisense RNA probes specific to 3′ internal regions of the L1 and L2 ORFs (Fig. 4C). The L2-specific riboprobe protected the expected 221-nt fragment of RNA (Fig. 4A), whereas the L1-specific riboprobe protected a 216-nt fragment (Fig. 4B). For the two probes, the RPA patterns were similar and exhibited temporal expression patterns consistent

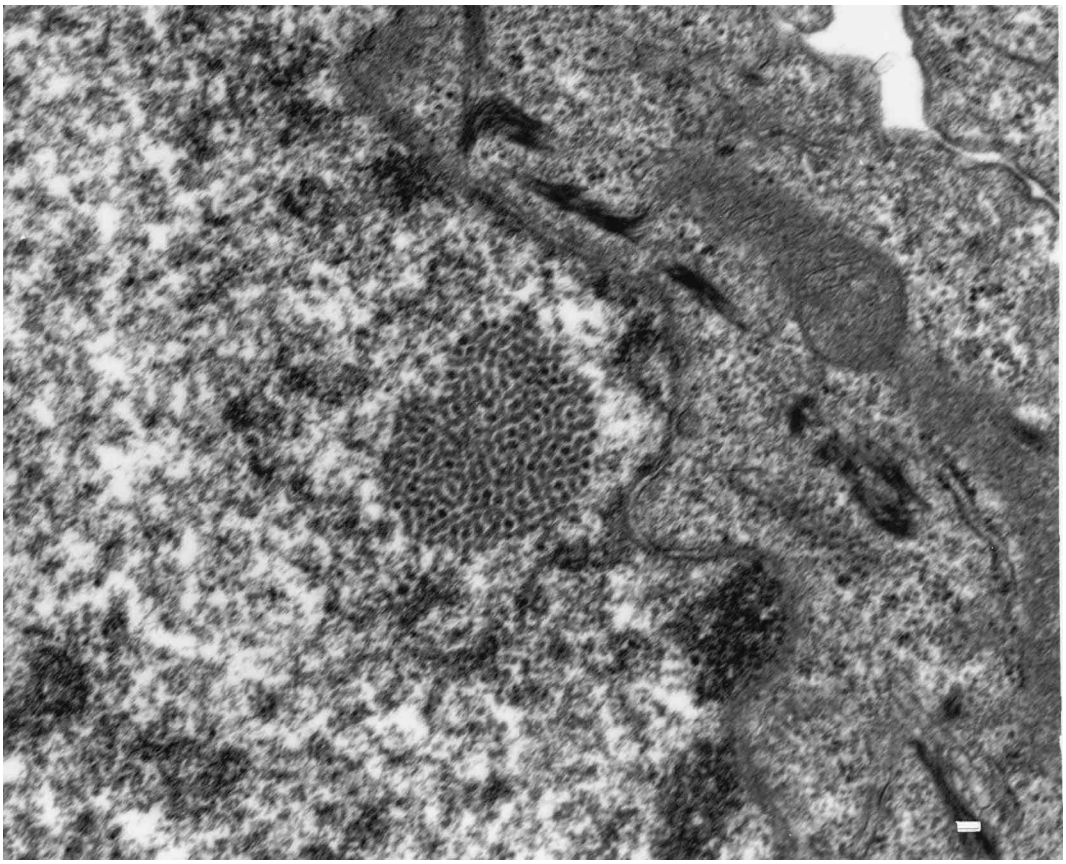

FIG. 3. Examination of C8:0-treated CIN-612 9E raft tissue by electron microscopy. Raft tissue was harvested at day 10 after lifting to the air-liquid interface, fixed, and stained. An intranuclear patch of viral particles can be observed. The bar is equal to 100 nm.

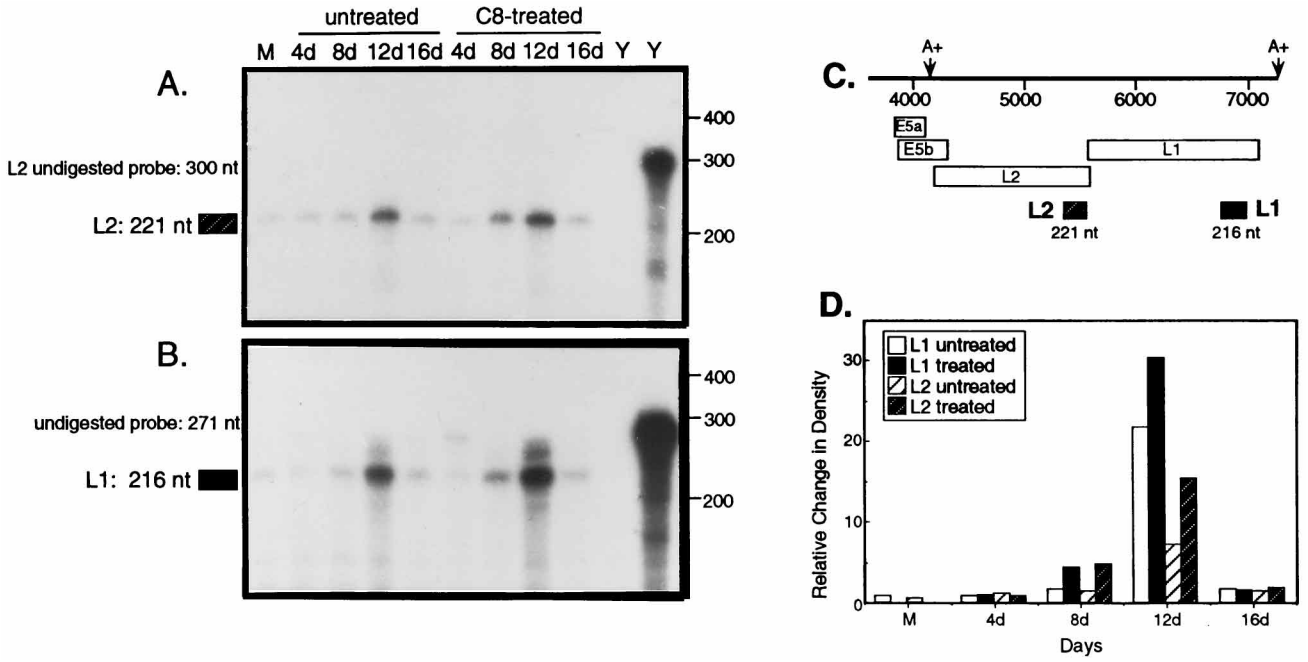

FIG. 4. RPAs quantitating the levels of HPV31b L1 and L2 transcripts. CIN-612 9E cells were cultured as monolayers (M) or on rafts; the rafts were either untreated or treated with 10 μM C8:0 every second day. Rafts were harvested following lifting to the air-liquid interface after 4 days (4d), 8 days (8d), 12 days (12d), and 16 days (16d) as indicated. Twenty micrograms of DNase I-treated total RNA or yeast RNA (Y) was analyzed. Two yeast RNA samples were included as controls; the leftmost was RNase digested to show probe specificity, and the rightmost was not RNase digested to indicate the size of the input probe. The sizes of the probes and predicted sizes of the protected fragments are indicated to the left of panels A and B, and RNA Century markers (Ambion) are indicated to the right. (A) Probe is specific to the 3′ internal L2 ORF (HPV31 nt 5353 to 5573). (B) Probe is specific to the 3′ internal L1 ORF (HPV31 nt 6705 to 6921). (C) Schematic of HPV31b late region with riboprobes illustrated. (D) Densitometric data from panels A and B; autoradiograms were scanned in the regions containing the full-length protected fragments. Absolute readings were plotted and represent the relative changes during stratification and differentiation. Open boxes, L1 expression in untreated rafts; closed boxes, L1 expression in C8:0-treated rafts; lightly striped boxes, L2 RNA expression in untreated rafts; heavily striped boxes, L2 expression in C8:0-treated rafts.

with the results of the Northern analyses. Monolayer cultures of CIN-612 9E cells were found to express both L1 and L2 RNAs (Fig. 4A and B, lanes M). The L1 and L2 transcript levels peaked at day 12 in untreated and C8:0-treated CIN-612 9E rafts. However, PKC-induced rafts had greater levels of late gene mRNAs than did their untreated counterparts (compare lanes 12d in Fig. 4A and B). The relative changes in the levels of late gene RNAs were determined by densitometric scanning (Fig. 4D). The results indicate that PKC induction led to an increase in the levels of late gene transcripts. This representation clearly illustrates the striking peak in the expression of late gene RNAs in the raft system at day 12 and the significant drop in expression by day 16. The levels of L1-containing transcripts were two times those of the L2-containing RNAs at day 12, whether the raft tissues were treated or untreated. These data were representative of several analyses using different preparations of RNA.

Cloning and sequencing of late gene transcripts expressed during the HPV31b life cycle. To define the structures of late gene mRNAs expressed in CIN-612 9E raft tissues, rafts were treated with C8:0 for 12 to 16 days and then harvested. Total RNA was extracted from rafts, DNase I treated, and subjected to reverse transcription. PCR was performed with various primer pairs (Table 1). Primers E6 5′, E7 5′, and 742 5′ were paired with primer L1-2 3′ or L2-3 3′ to assay for cDNA molecules containing L1 or L2 sequences, respectively (Fig. 5). Three novel HPV31b cDNA structures contained the L1 region (Fig. 5, transcripts A to C). Another L1-containing cDNA (Fig. 5, transcript D) and an L2 cDNA (Fig. 5, transcript E) were similar to transcripts previously reported (16). Structures of the transcripts were determined from the cloned cDNAs; however, we will subsequently refer to them as RNA transcripts. In each transcript, the E1 splice donor at HPV31b nt 877 was used. In one transcript, following the E6* and E7 ORFs, the E1 donor was spliced directly to the L1 splice acceptor at nt 5552, the first base of the first AUG translational start for the L1 ORF (Fig. 5, transcript A). This splice was predicted to result in the termination of translation 17 nt downstream of the L1 AUG (fused ORF designated E1*II). All other late gene transcripts contained the E1^E4 splice (Fig. 5, transcripts B to E). The E7,E1^E4*,L1 transcript included E7 and the E1^E4 region to nt 3322 spliced to a nonconsensus site at nt 4369 in the L2 region, out of frame (Fig. 5, transcript B). The transcript went on to splice from nt 4422 to the AUG in the L1 ORF at nt 5552. The predicted E1^E4* protein encoded by this transcript should contain 15 amino acids of E1^E4, 18 amino acids in the L2 region (out of frame with L2), and 21 amino acids out of frame in the L1 ORF (Fig. 5, transcript B). The E1^E4,E5*,L1s transcript (Fig. 5, transcript C), which potentially encodes a normal E1^E4 fusion protein, continued past the E4 splice donor at nt 3590. This tricistronic transcript used a nonconsensus splice donor at nt 3832 and a nonconsensus acceptor site in L1 at nt 5555 (Fig. 5, transcript C). Beginning with the E5a start codon at HPV31 nt 3816, the transcript could potentially encode a 26-amino-acid E5a fusion (E5*) that terminates out of frame in the 5′ L1 region, similar to E1^E4*. Finally, in this transcript, the first AUG of the L1 ORF was spliced out, leaving the second in-frame L1 AUG at 5999 as the first available start site. Translation initiated at this site is predicted to give rise to a smaller L1 protein containing the C-terminal 355 amino acids of the full-length 504-amino-acid L1 protein. The structure of

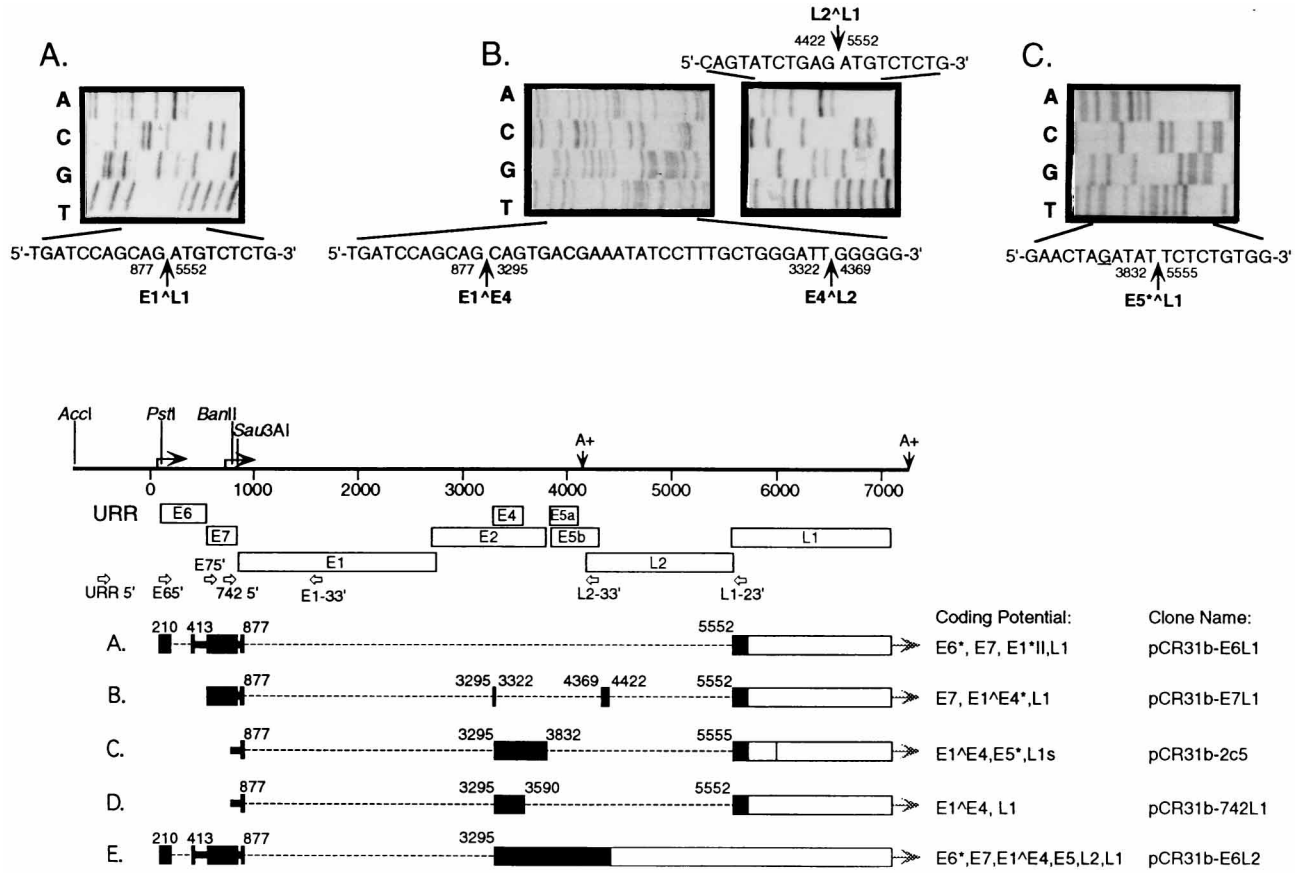

FIG. 5. cDNA-derived HPV31b sequences from C8:0-treated CIN-612 9E tissues. The top panels show the results of dideoxy sequencing. The lower schematic illustrates the HPV31b genome arbitrarily linearized following the late polyadenylation signal (A+). The following restriction enzyme sites were important for cloning: *Acc*I at HPV31 nt 7238, *Pst*I at nt 121, *Ban*II at nt 811, and *Sau*3AI at nt 858. Bent arrows indicate the reported early promoter P_{97} and the differentiation-specific promoter P_{742} (15). Polyadenylation sites at nt 4138 to 4143 and 7227 to 7231 are designated by A+ (10). The ORFs are indicated by open boxes. PCR primer placement and orientation are shown by open arrows. Transcripts characterized in this study are shown (A to E), with the filled regions representing areas included in the clones and sequenced. The dotted lines mark areas removed by splicing. The numerals above the transcript structures indicate the nucleotides at the splice junctions. The amino acid coding potential of each transcript is illustrated to the right, followed by the respective clone name of the cDNA-containing construct. Panels A, B, and C at the top correspond to the transcripts designated A, B, and C, respectively, at the bottom. A, RNA-PCR using primers E6 5′ and L1-2 3′ gave a 677-bp partial cDNA containing the E6* splice and the E1^L1 splice. B, RNA-PCR using primers E7 5′ and L1-2 3′ gave a 436-bp partial cDNA containing the E1^E4 splice, a nonconsensus E4^L2 splice, and a nonconsensus L2^L1 splice. C, RNA-PCR using primers 742 5′ and L1-2 3′ gave a 676-bp partial cDNA containing the E1^E4 splice and an E5^L1 splice which deletes the first AUG of the L1 ORF at nt 5552. The underlined base deviates from the published sequence of HPV31 (10). The vertical bar in the open box corresponding to L1 indicates the second AUG at nt 5999 of the ORF. D, RNA-PCR using primers 742 5′ and L1-2 3′ gave a 438-bp partial cDNA containing the E1^E4 splice and a E4^L1 splice. E, RNA-PCR using primers E6 5′ and L2-3 3′ gave a 1,450-bp partial cDNA containing the E6* splice and the E1^E4 splice. Transcripts similar to transcripts D and E were reported previously (16). HPV31b nt numbering is based on the sequence of HPV31 (10).

the E1^E4,L1 transcript was reported previously and would encode E1^E4 fusion and L1 proteins (Fig. 5, transcript D) (16). In addition, we characterized a transcript structurally similar to the E1^E4,E5,L2,L1 mRNA described in the same study (16). However, our transcript included the sequences for E6* and E7, well upstream of the P_{742} promoter assigned to this RNA (Fig. 5, transcript E). Thus, our transcript contained ORFs E6*, E7, E1^E4, E5a, E5b, L2, and L1.

Verification and temporal expression of specific late transcripts by RPAs. RPAs using cDNA-derived probes were performed for two reasons. First, we wished to verify that the cDNAs cloned from RNA-PCR products represented authentic mRNA molecules. Second, we wanted to determine the temporal expression pattern of specific viral mRNAs during PKC-induced differentiation of the raft tissues. Total RNAs from C8:0-treated rafts were harvested at intervals following lifting to the air-liquid interface; representative results of these experiments are shown in Fig. 6. The cDNA clones from which sequence data in Fig. 5 (transcripts A to D) were derived served as templates for antisense RNA for RPAs. This strategy was not feasible for investigation of the L2 transcript (Fig. 5, construct E), as technical limitations precluded the synthesis of an antisense riboprobe of the length required for specificity. Our results verified that each of the L1-containing cDNA species represented viral RNAs (Fig. 6 and data not shown). Furthermore, the temporal expression levels of the L1-specific mRNAs were similar to each other, to the results obtained by Northern analysis (Fig. 2), and to the results obtained by RPA of the total L1 and L2 RNAs (Fig. 4). L1 expression was easily detected in RNA samples derived from untreated monolayer cultures of CIN-612 9E cells. Low levels of the L1 RNAs were observed in 4-day rafts, L1-specific RNA levels peaked in C8:0-treated rafts harvested at 12 days, and levels were significantly lower by 16 days (Fig. 6). The E7,E1*II,L1 probe also strongly protected 194 nt of RNA corresponding to the E7,E1^ region (Fig. 6B). The E7,E1^ region is contained in the majority of the early and late HPV31b transcripts thus far identified (15, 16, 31).

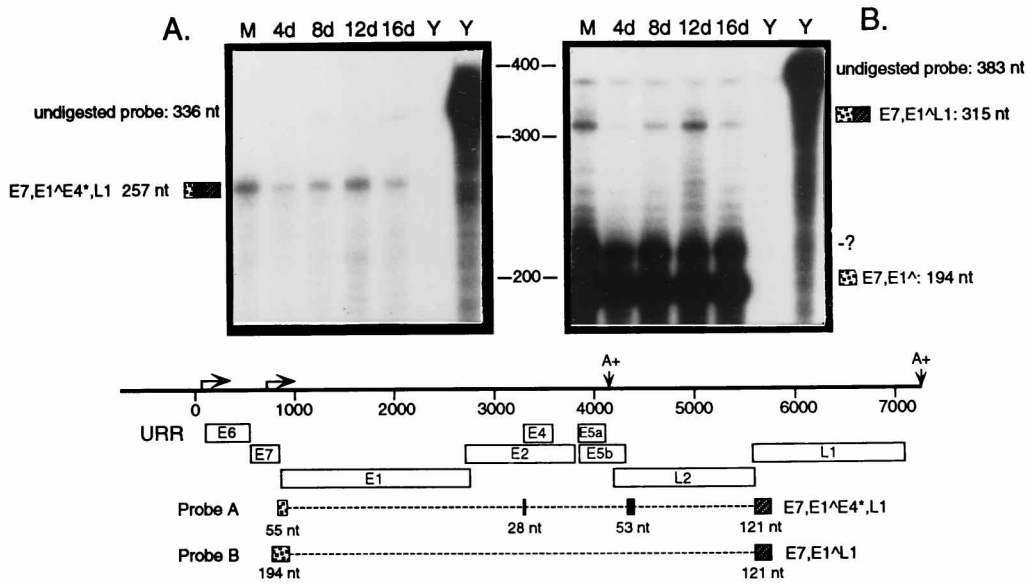

FIG. 6. RPAs of HPV31b L1 transcripts. CIN-612 9E cells were cultured as monolayers (M) or on rafts treated with 10 μM C8:0 every second day. Rafts were harvested following lifting to the air-liquid interface after 4 days (4d), 8 days (8d), 12 days (12d), and 16 days (16d) as indicated. Fifteen micrograms of total RNA or yeast RNA (Y) was analyzed. Two yeast RNA samples were included as controls; the leftmost was RNase digested to show probe specificity, and the rightmost was not RNase digested to indicate the size of the input probe. The sizes of the probes and sizes of the protected fragments are shown to the left and right of panels A and B, respectively. RNA Century markers (Ambion) are indicated between panels A and B. (A) The probe is derived from HPV31b cDNA clone pCR31b-E7L1 and is specific to 257 nt of E7,E1^E4*,L1 (HPV31 nt 823 to 877^3295 to 3322^4369 to 4422^5552 to 5672). (B) The probe is derived from HPV31b cDNA clone pCR31b-E6L1 and is specific to 315 nt of E7,E1*II,L1 (E7,E1^L1; HPV31 nt 683 to 877^5552 to 5672). Predicted protection for E7,E1^ is 194 nt. Below panels A and B is a schematic of HPV31b illustrating the riboprobe sequences.

5′ end-mapping of HPV31b late gene transcripts by S1 and exoVII protection assays. To map the 5′ regions of the characterized L1 and L2 transcripts, HPV31b sequences were cloned upstream of the 5′ ends of the partial cDNAs obtained via RNA PCR and cloning (Materials and Methods, Fig. 5, and Fig. 7). Two sets of constructs were prepared for each late gene cDNA; the URRE6*-late constructs contained the URR, E6*, E7, E1^, and plus late sequences, whereas the URRE6-late constructs contained the URR, full-length E6, E7, E1^, and late sequences. Single-stranded, 5′-end-labeled probes were made from plasmid constructs by amplifying the region shown in Fig. 7 with a primer labeled at the 5′ end. To ensure that the probes would be larger than the transcripts, the second, unlabeled primer was M13(−40) (see Materials and Methods). With each probe, we analyzed total RNA from CIN-612 9E monolayer cultures and C8:0-treated rafts harvested at day 12. Only results of assays using the URRE6*-late probes are shown (Fig. 7). Each URRE6*-late probe gave identical results in S1 and exoVII digestion assays. The exoVII-digested samples migrated slightly slower than the S1-digested samples, a common phenomenon when large amounts of RNA are analyzed (4). Each URRE6*-late probe protected a product with a 5′ end in the vicinity of P_{97} (Fig. 7, upward-pointing arrows). More detailed analyses have demonstrated that this subset of late gene transcripts actually initiates ≈20 nt upstream of P_{97}, at a novel HPV31b promoter which we designated P_L (32). Nevertheless, we are not able to rule out the possibility that late gene-containing transcripts are also initiated at P_{97}. In each sample of 12-day C8:0-treated RNA, each probe also protected a doublet mapping 5′ ends to the region of P_{742} (Fig. 7, arrowheads). There was no evidence of P_{742} protection in RNA samples derived from 9E monolayers (Fig. 7, lanes M). In addition, probes containing the E4 splice acceptor (Fig. 7, constructs B, C, D, and E) protected a fragment corresponding to a 5′ end near the E4 splice acceptor sequences; we designated this promoter P_{E4} (32) (Fig. 7, downward-pointing arrows; data from probe E not shown). The P_{E4} protection product was observed in the monolayer-derived RNA samples, albeit at lower levels (Fig. 7B to D, lanes M).

None of the techniques used thus far provided any evidence for late transcripts containing the full-length E6 ORF. As the use of cDNAs to perform nuclease protection experiments can mask unspliced regions, we created probe templates which were similar to the L1-containing constructs described above (Fig. 7, constructs A to E), but they contained the full-length E6 ORF in place of the E6* ORF. In addition to P_{E4} and P_{742}, the S1 assay revealed protected fragments with a 5′ end at the E6* splice acceptor and a P_L 5′ end. ExoVII, which lacks endonuclease activity, protected fragments corresponding to 5′ ends at P_{E4}, P_{742}, and P_L (data not shown). Taken together, the results of S1 and exoVII protection assays using 5′-end-labeled URRE6-late probes indicated that a fraction of the RNAs initiating at P_L contained the full E6 ORF. Based on the results of the nuclease protection assays (Fig. 6 and 7 and data not shown), we estimate approximately equal relative abundance of transcripts A, B, C, and E (Fig. 7). The transcript shown in Fig. 7D appeared to be present at lower levels relative to the others. Furthermore, the transcripts initiated from the P_L and P_{742} promoters were present at greater levels than those transcripts initiated from the P_{E4} region promoter. Figure 8 summarizes the structures of the late gene-containing transcripts, their putative promoters, and their coding potentials. There appeared to be three to four versions of each viral RNA structure characterized by RNA PCR. One version started at P_L and contained E6, E7, and E1^; a second initiated at P_L and contained E6*, E7, and E1^; and a third began at P_{742}. In addition, each transcript containing the E4 region had a start site in the area of the E4 splice acceptor at nt 3295 (P_{E4}).

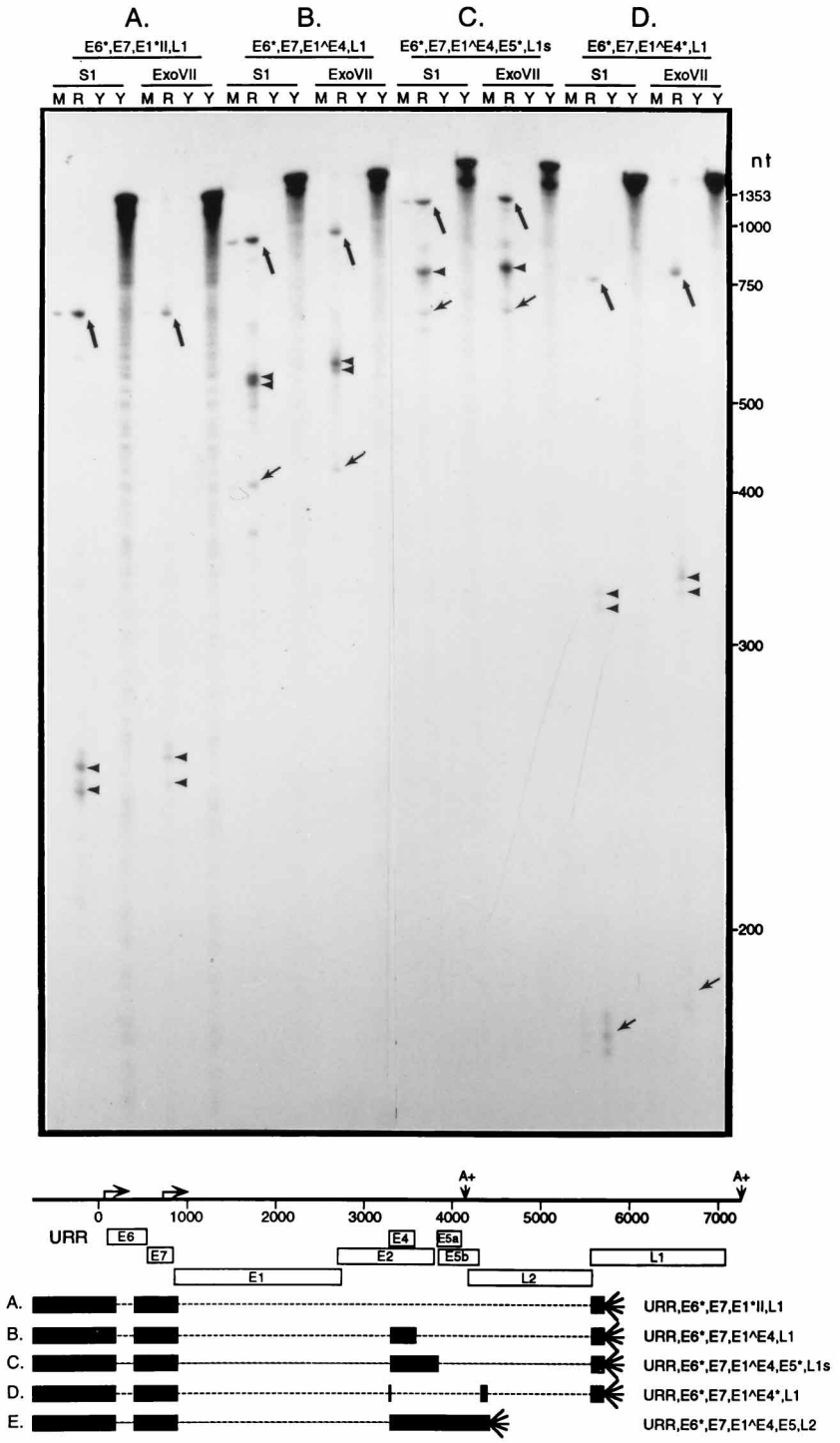

FIG. 7. S1 and exoVII nuclease protection analyses of HPV31b L1 transcripts. CIN-612 9E cells were cultured as monolayers (M) or as C8:0-treated rafts harvested at day 12 (R). Forty micrograms of total RNA or yeast RNA (Y) was hybridized analyzed by digestion with S1 nuclease or exoVII nuclease as indicated. Two yeast RNA samples were included as controls; for each probe, the leftmost yeast RNA sample was nuclease digested to show probe specificity and the rightmost was not nuclease digested to indicate the size of the input probe. RNA Century markers (Ambion) and 5′-end-labeled ∅X174 DNA digested with *Hae*III were used as standards, indicated at the right. Probe specificity is indicated at the bottom. The plasmids used as templates and the lengths of the probe species are as follows: (A) p31U*E6L1, 1,572 nt; (B) p31U*742L1, 1,900 nt; (C) p31U*2c5, 2,098 nt; (D) p31U*E7L1, 1,653 nt; (E) p31U*E6L2, 2,120 nt (data not shown). Upward-pointing arrows show a protection to the P_L promoter. Arrowheads indicate protection corresponding to the P_{742} promoter. Downward-pointing arrows note protection to the P_{E4} promoter (32).

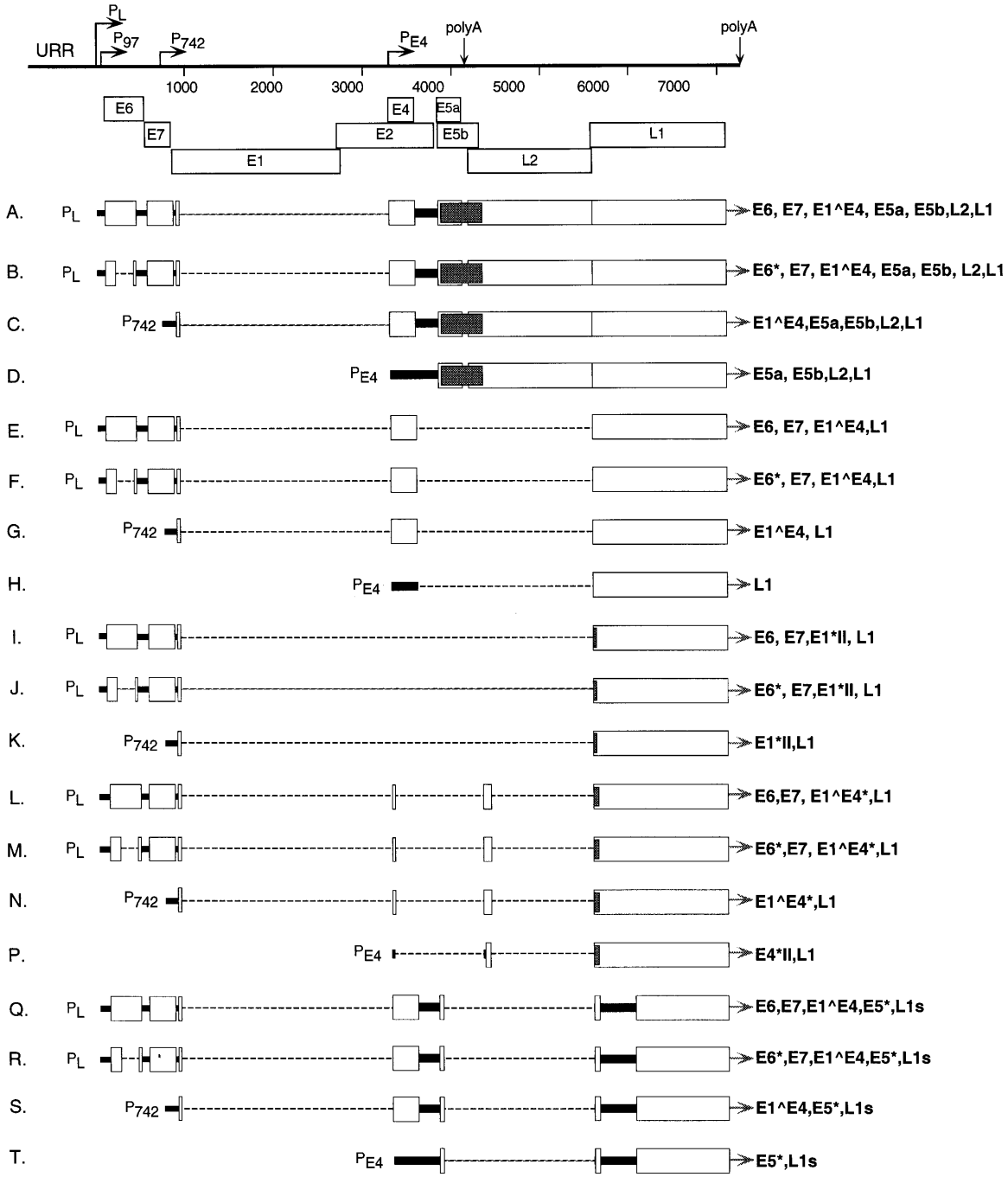

FIG. 8. Structures of the late transcripts expressed during the life cycle of HPV31b. Open boxes indicate ORFs, with smaller stippled boxes illustrating overlapping ORFs. Thick black lines represent noncoding sequences. Dotted lines show regions spliced out of transcripts. The arrow at the 3′ ends indicates that the transcripts likely proceed to the late polyadenylation site. The promoters are shown by bent arrows; polyadenylation sites are given as downward arrows. The basic RNA structures were characterized by RNA PCR and sequencing. Nuclease protection assays provided evidence that each basic structure has three to four different versions. The first set of these species have a transcriptional start at the P_L promoter, positioned ≈20 nt upstream from P_{97} (32), and include a full-length E6 ORF (A, E, I, L, and Q). The second set uses the P_L promoter but contains the E6* ORF (B, F, J, M, and R). The third set of late transcripts start at the differentiation-specific P_{742} promoter (C, G, K, N, and S). Lastly, the transcripts containing E4 sequences can use a translational start, P_{E4}, in the region of the E4 splice acceptor (D, H, P, and T) (32).

DISCUSSION

We used the organotypic (raft) tissue culture system to characterize the structures and temporal expression patterns of HPV31b late gene transcripts during the complete viral life cycle. Our data show that PKC-induced CIN-612 9E tissues had undergone a program of differentiation by day 12 in the raft system, morphologically resembling the CIN I lesion from which the cells were originally derived (2, 34). Both untreated and PKC-induced CIN-612 9E tissues maximally expressed late gene transcripts around day 12 in the raft system. However, PKC induction resulted in higher levels of late gene transcript expression. In addition, clumps of viral particles could readily

be observed in the PKC-activated CIN-612 9E tissues as early as day 10, indicating the late gene transcripts were functional for translation. Quantitative analysis of the levels of L1- and L2-containing transcripts showed that L1 transcripts outnumbered L2 RNAs by a ratio of 2 to 1. This is not surprising considering that L1 protein is expressed in excess of L2 protein in productively infected HPV1a cells (7). We found it striking that the levels of late RNAs decreased some 8- to 16-fold from days 12 to 16 in the rafts tissues. These data suggest that rafts harvested after day 12 are likely to be suboptimal for analyses of late genes.

Differentiation alone is not responsible for inducing transcription of late genes, as these transcripts were also detected in untreated monolayers, 4-day rafts, and 8-day rafts. Neither was induction of the PKC pathway required for late gene transcription. It long has been an assumption that proliferating monolayer cultures functionally resemble the mitotically active basal layer of the epithelium. However, we found that monolayer cultures consistently expressed higher total levels of late gene transcripts than did rafts harvested at 4 days, which consist predominantly of basal-like cells. This result suggests that proliferating monolayer cultures are not analogous to basal cells. It is important to note that although late genes are transcribed at relatively high levels in monolayer cultures, the synthesis of capsid proteins in monolayers has not been reported. Likewise, late proteins have not been detected in untreated CIN-612 9E raft tissues even though a certain level of differentiation is achieved and late transcripts are readily detected in untreated rafts (2, 28, 33). These circumstances imply that some type of differentiation-specific translational control may be exerted over late gene expression. To our knowledge, investigations of translational control have been limited to undifferentiated monolayer cultures, and presently used techniques would be difficult to perform in epithelial tissues. It is also possible that there is induction of more translationally prone late mRNAs during the differentiation process than in monolayers. For example, there may be a difference in the late gene RNA stability or processing at the 3′ ends that our assays have not addressed. Late gene transcripts initiating at P_{742} were undetectable in monolayer cultures but greatly upregulated in PKC-induced raft tissues (Fig. 7 and references 15 and 16). Thus, we favor the theory that these higher levels of late gene transcripts in the differentiating tissues provide a threshold needed to produce a critical mass of capsid proteins, permitting detection by techniques such as Western blotting and immunochemical staining. For reasons detailed below, we also believe that these P_{742}-initiated transcripts are most likely to be translated into late proteins. Also in support of our theory is the finding that HPV31b E1^E4 transcripts are detected in monolayers and rafts; however, E1^E4 transcripts from P_{742} are greatly upregulated in rafts (15). E1^E4 proteins can be detected in CIN-612 9E monolayers, untreated rafts, and PKC-induced rafts but levels are highest in the latter (33). Thus, we believe that late proteins may be synthesized in monolayers and untreated rafts but are present at levels below the sensitivity of currently used techniques.

The HPV31 late region-containing transcripts can be grouped either by promoter usage or by general splicing patterns (Fig. 8). We detected two to three possible 5′ ends (promoters) for the late gene-containing RNAs (e.g., Fig. 8, transcripts A, C, and D) and five general patterns of splicing (e.g., Fig. 8, transcripts A, E, I, L, and Q). Further, our analyses showed that individual RNA structures had well-defined and reproducible temporal expression patterns during the viral life cycle. For these reasons, we believe the cDNAs were derived from bona fide viral RNAs. However, because we analyzed total RNA and not cytoplasmic RNA, it remains a possibility that some of the RNA structures identified represent precursor RNAs and not cytoplasmic mRNA species. A subset of the late transcripts used a promoter that we mapped ≈20 nt upstream of the P_{97} early promoter (32). We designated this promoter P_L, for late promoter. For transcripts using P_L, a subset contained the full-length E6 and E7 ORFs, whereas others contained the E6* and E7 ORFs. Consistent with an earlier report, we found strong use of the differentiation-specific P_{742} promoter in C8:0-treated rafts (16). In addition, we found a subset of the late transcripts which contained E4 sequences used as promoter, P_{E4}, near the E4 splice acceptor (32).

Different promoter usage combined with various splicing patterns gave rise to a complex array of HPV31b late gene transcripts. Two transcripts initiating at P_{742}, shown as C and G in Fig. 8, were described previously (16). However, our data showed that these transcripts exist in other forms, initiating at P_L and at P_{E4}. These first two sets of transcripts contain the E1^E4,E5a,E5b,L2,L1 ORFs (transcripts A to C) and E1^E4,L1 ORFs (transcripts E to G). The third set of transcripts (I to K) potentially encode an 11-amino-acid fusion protein, E1*II, and L1. The fourth set of late gene RNA structures (L to N) potentially encode a novel 54-amino-acid fusion protein, E1^E4*, and L1. The fifth set of transcripts (Q to S) has the potential to encode E1^E4 plus two previously unidentified ORFs, E5* and L1s. The E5* ORF should encode a 26-amino-acid fusion protein. The L1s ORF should encode a truncated L1 protein, beginning at the second in-frame L1 AUG, predicted to be ~40 kDa, compared to a full-sized L1 of ~55 kDa. Smaller L1 proteins have been detected in purified HPV1a virions and in cells expressing L1 ORFs of HPV1, -6, -11, -16, and -33 (7, 12, 18, 19, 36, 39, 49). These were presumed to be breakdown products; however, our data suggest that the synthesis of smaller L1 proteins can arise from transcripts initiated from an internal AUG(s). Transcripts initiating at P_{E4} (D, H, P, and T) would not encode the E1^E4 fusion due to the absence of the E1 AUG. The first ORFs in transcript D are E5a and E5b. The first translational start in transcript H is the L1 AUG. A second in-frame AUG in the E1^E4* ORF from transcript P should initiate the 34-amino-acid E4*II fusion. Transcript T maintained the E5* and L1s ORFs.

The prominent question raised by these data is, why are there so many structurally polycistronic transcripts containing the late ORFs? One possibility is that the transcription of any ORF past the early polyadenylation signal includes the L1 and L2 ORFs merely as transcriptional readthrough to the late polyadenylation signal. If so, the presence of the late ORFs in the majority of these transcripts may be insignificant with respect to the translation of the late ORFs. Elegant and thorough analyses of the variables governing translational efficiencies of ORFs in polycistronic mRNAs have been performed by Kozak (21, 22). Numerous viruses express mRNAs that are structurally polycistronic. However, most evidence suggests that these mRNAs are functionally monocistronic and that AUG context determines whether translation of a downstream AUG is favored (21). The majority of the transcripts that we characterized contain the E6 or E6*, E7, and/or E1^E4 ORFs, which have AUG contexts that remain constant upstream of L2 or L1. Thus, based on Kozak's work, these transcripts may not be responsible for efficient expression of late gene products. Certain transcripts initiating at P_{742} (Fig. 8, transcripts C, G, and S) probably are required primarily to make E1^E4, the major viral gene product (6). However, P_{742}-initiated transcripts with a single ORF upstream of L1 (e.g., G, K, and N) likely contribute to L1 expression, especially

in PKC-induced rafts where P_{742} is greatly upregulated (reference 16 and our data). In addition, the P_{E4}-initiating transcripts seem likely candidates for encoding the late proteins, as they contain no or one upstream ORF. Transcript D in Fig. 8 may be responsible for the expression of E5a, E5b, and L2, proteins expressed in relatively low amounts in virus-infected cells (7, 25). The early polyadenylation site occurs within the E5b ORF, making E5b the 5′-most ORF in the HPV31b genome which requires the late polyadenylation site. The P_{742}-initiated, E1*II-containing transcript, K, may also produce L1 protein. These conclusions must be tempered by the fact that the rules for translational efficiency are not based on studies in differentiated tissues, as indicated above. Finally, although we have no solid evidence for translational control, it is tempting to speculate that the putative E1*II, E1^E4*, E4*II, and/or E5* late proteins may be differentiation-specific accessory proteins which help determine which late RNAs are translated. The latter possibility might help to explain the role of the P_L-initiating transcripts which, based on Kozak's work, appear not to be functional for late gene translation.

In conclusion, we have described the structures and temporal expression patterns of HPV31b late gene transcripts. This study was possible only because we used the raft culture system, which permits duplication of the complete HPV31b life cycle in vitro. Late gene transcription peaks at day 12 in the raft system, concomitant with the observation of viral particles in the tissues. Seventeen novel transcripts and two new late promoters were identified. The transcripts have the potential to encode a number of characterized and putative late gene products, including E1^E4, E1^E4*, E4*II, E5a, E5b, E5*, L2, L1, and L1s. We are presently investigating the ability of the RNAs to encode specific late proteins. We recently reported the biosynthesis of infectious HPV18 in primary keratinocytes transfected with HPV18 DNA and allowed to differentiate in the raft system (29). We expect that this procedure will allow the analysis reported here to be conducted on any naturally occurring HPV type or mutant for which cloned DNA is available. In addition, genetic analyses of viral late gene expression should reveal much about the regulation of late gene expression during cellular differentiation and viral morphogenesis.

ACKNOWLEDGMENTS

We are grateful to David J. Spector and Timothy J. Mayer for critical reading of the manuscript. We thank Felix O. Wettstein and Carl Baker for advice and helpful discussions.

This work was supported by Public Health Service grant CA-66316 from the National Cancer Institute (M.A.O.) and National Cancer Institute grant CA-64624 (C.M.).

REFERENCES

1. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.).** 1995. Current protocols in molecular biology. John Wiley & Sons, Inc., Brooklyn, NY.
2. **Bedell, M. A., J. B. Hudson, T. R. Golub, M. E. Turyk, M. Hosken, G. D. Wilbanks, and L. A. Laimins.** 1991. Amplification of human papillomavirus genomes in vitro is dependent on epithelial differentiation. J. Virol. **65:**2254–2260.
3. **Broker, T. R., and M. Botchan.** 1986. Papillomaviruses: retrospectives and prospectives. Cancer Cells **4:**17–36.
4. **Calzone, F. J., R. J. Britten, and E. H. Davidson.** 1987. Mapping of gene transcripts by nuclease protection assays and cDNA primer extension. Methods Enzymol. **152:**611–632.
5. **De Geest, K., M. E. Turyk, M. I. Hosken, J. B. Hudson, L. A. Laimins, and G. D. Wilbanks.** 1993. Growth and differentiation of human papillomavirus type 31b positive human cervical cell lines. Gynecol. Oncol. **49:**303–310.
6. **Doorbar, J., D. Campbell, R. J. A. Grand, and P. H. Gallimore.** 1986. Identification of the human papilloma virus-1a E4 gene products. EMBO J. **5:**355–362.
7. **Doorbar, J., and P. H. Gallimore.** 1987. Identification of proteins encoded by the L1 and L2 open reading frames of human papillomavirus 1a. J. Virol. **61:**2793–2799.
8. **Doorbar, J., A. Parton, K. Harley, L. Banks, T. Crook, M. Stanley, and L. Crawford.** 1990. Detection of novel splicing patterns in a HPV16-containing keratinocyte cell line. Virology **178:**254–262.
9. **Frattini, M. G., H. B. Lim, and L. A. Laimins.** 1996. *In vitro* synthesis of oncogenic human papillomaviruses requires episomal genomes for differentiation-dependent late expression. Proc. Natl. Acad. Sci. USA **93:**3062–3067.
10. **Goldsborough, M. D., D. DiSilvestre, G. F. Temple, and A. T. Lorincz.** 1989. Nucleotide sequence of human papillomavirus type 31: a cervical neoplasia-associated virus. Virology **171:**306–311.
11. **Grassmann, K., B. Rapp, H. Maschek, K. U. Petry, and T. Iftner.** 1996. Identification of a differentiation-inducible promoter in the E7 open reading frame of human papillomavirus type 16 (HPV-16) in raft cultures of a new cell line containing high copy numbers of episomal HPV-16 DNA. J. Virol. **70:**2339–2349.
12. **Hagensee, M. E., N. Yaegashi, and D. A. Galloway.** 1993. Self-assembly of human papillomavirus type 1 capsids by expression of the L1 protein alone or by coexpression of the L1 and L2 capsid proteins. J. Virol. **67:**315–322.
13. **Howley, P. M.** 1996. *Papillomavirinae*: the viruses and their replication, p. 2045–2076. *In* B. N. Fields and D. M. Knipe (ed.), Fields virology, 3rd ed. Raven Press, New York, N.Y.
14. **Hudson, J. B., M. A. Bedell, D. J. McCance, and L. A. Laimins.** 1990. Immortalization and altered differentiation of human keratinocytes in vitro by the E6 and E7 open reading frames of human papillomavirus type 18. J. Virol. **64:**519–526.
15. **Hummel, M., J. B. Hudson, and L. A. Laimins.** 1992. Differentiation-induced and constitutive transcription of human papillomavirus type 31b in cell lines containing viral episomes. J. Virol. **66:**6070–6080.
16. **Hummel, M., H. B. Lim, and L. A. Laimins.** 1995. Human papillomavirus type 31b late gene expression is regulated through protein kinase c-mediated changes in RNA processing. J. Virol. **69:**3381–3388.
17. **Kaur, P., and J. K. McDougall.** 1988. Characterization of primary human keratinocytes transformed by human papillomavirus type 18. J. Virol. **62:**1917–1924.
18. **Kirnbauer, R., F. Booy, N. Cheng, D. R. Lowy, and J. T. Schiller.** 1992. Papillomavirus L1 major capsid protein self-assembles into virus-like particles that are highly immunogenic. Proc. Natl. Acad. Sci. USA **89:**12180–12184.
19. **Kirnbauer, R., J. Taub, H. Greenstone, R. Roden, M. Dürst, L. Gissmann, D. R. Lowy, and J. T. Schiller.** 1993. Efficient self-assembly of human papillomavirus type 16 L1 and L1-L2 into virus-like particles. J. Virol. **67:**6929–6936.
20. **Koss, L. G.** 1987. Cytologic and histologic manifestations of human papillomavirus infection of the female genital tract and their clinical significance. Cancer **60:**1942–1950.
21. **Kozak, M.** 1986. Regulation of protein synthesis in virus-infected animal cells. Adv. Virus Res. **31:**229–292.
22. **Kozak, M.** 1991. An analysis of vertebrate mRNA sequences: intimations of translational control. J. Cell Biol. **115:**887–903.
23. **Lambert, P. F.** 1991. Papillomavirus DNA replication. J. Virol. **65:**3417–3420.
24. **Lorincz, A. T., R. Reid, A. B. Jenson, M. D. Greenburg, W. Lancaster, and R. J. Kurman.** 1992. Human papillomavirus infection of the cervix: relative risk associations of 15 common anogenital types. Obstet. Gynecol. **79:**328–337.
25. **Mayer, T. J., and C. Meyers.** Unpublished data.
26. **McCance, D. J., R. Kopan, E. Fuchs, and L. A. Laimins.** 1988. Human papillomavirus type 16 alters human epithelial cell differentiation *in vitro*. Proc. Natl. Acad. Sci. USA **85:**7169–7173.
27. **Meyers, C.** 1996. Organotypic (raft) epithelial tissues culture system for the differentiation-dependent replication of papillomavirus. Methods Cell Sci. **18:**201–210.
28. **Meyers, C., M. G. Frattini, J. B. Hudson, and L. A. Laimins.** 1992. Biosynthesis of human papillomavirus from a continuous cell line upon epithelial differentiation. Science **257:**971–973.
29. **Meyers, C., T. J. Mayer, and M. A. Ozbun.** Synthesis of infectious human papillomavirus type 18 in differentiating epithelium transfected with viral DNA. Submitted for publication.
30. **Myers, G., C. Baker, C. Wheeler, A. Halpern, A. McBride, and J. Doorbar.** 1996. Human papillomaviruses 1996. A compilation and analysis of nucleic acid and amino acid sequences. Los Alamos National Laboratory, Los Alamos, N.Mex.
31. **Ozbun, M. A., and C. Meyers.** 1996. Transforming growth factor β1 induces differentiation in human papillomavirus-positive keratinocytes. J. Virol. **70:**5437–5446.
32. **Ozbun, M. A., and C. Meyers.** Unpublished data.
33. **Pray, T. R., and L. A. Laimins.** 1995. Differentiation-dependent expression of E1^E4 proteins in cell lines maintaining episomes of human papillomavirus type 31b. Virology **206:**679–685.
34. **Rader, J. S., T. R. Golub, J. B. Hudson, D. Patel, M. A. Bedell, and L. A. Laimins.** 1990. *In vitro* differentiation of epithelial cells from cervical neo-

plasias resembles *in vivo* lesions. Oncogene **5:**571–576.
35. **Rohlfs, M., S. Winkenbach, S. Meyer, T. Rupp, and M. Dürst.** 1991. Viral transcription in human keratinocyte cell lines immortalized by human papillomavirus type-16. Virology **183:**331–342.
36. **Rose, R. C., W. Bonnez, R. C. Reichman, and R. L. Garcea.** 1993. Expression of human papillomavirus type 11 L1 protein in insect cells: in vivo and in vitro assembly of viruslike particles. J. Virol. **67:**1936–1944.
37. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, New York, N.Y.
38. **Sang, B.-C., and M. S. Barbosa.** 1992. Increased E6/E7 transcription in HPV 18-immortalized human keratinocytes results from inactivation of E2 and additional cellular events. Virology **189:**448–455.
39. **Sasagawa, T., P. Pushko, G. Steers, S. E. Gschmeissner, M. A. N. Hajibagheri, J. Finch, L. Crawford, and M. Tommasino.** 1995. Synthesis and assembly of virus-like particles of human papillomavirus type 6 and type 16 in fission yeast *Schizosaccharomyces pombe*. Virology **206:**126–135.
40. **Schneider-Gädicke, A., and E. Schwarz.** 1986. Different human cervical carcinoma cell lines show similar transcription patterns of human papillomavirus type 18 early genes. EMBO J. **5:**2285–2292.
41. **Schwarz, E., U. K. Freese, L. Gissmann, W. Mayer, B. Roggenbuck, A. Stremlau, and H. zur Hausen.** 1985. Structure and transcription of human papillomavirus sequences in cervical carcinoma cells. Nature **314:**111–114.
42. **Schwarz, E., A. Schneider-Gädicke, and H. zur Hausen.** 1987. Human papillomavirus type-18 transcription in cervical carcinoma cell lines and in human cell hybrids. Cancer Cells **5:**47–53.
43. **Sherman, L., and N. Alloul.** 1992. Human papillomavirus type 16 expresses a variety of alternatively spliced mRNAs putatively encoding the E2 protein. Virology **191:**953–959.
44. **Sherman, L., N. Alloul, I. Golan, M. Durst, and A. Baram.** 1992. Expression and splicing patterns of human papillomavirus type-16 mRNAs in pre-cancerous lesions and carcinomas of the cervix, in human keratinocytes immortalized by HPV 16, and in cells lines established from cervical cancers. Int. J. Cancer **50:**356–364.
45. **Smits, H. L., M. T. E. Cornelissen, M. F. Jebbink, J. G. van den Tweel, A. P. H. B. Struyk, M. Briët, and J. ter Schegget.** 1991. Human papillomavirus type 16 transcripts expressed from viral-cellular junctions and full-length viral copies in CaSki cells and in a cervical carcinoma. Virology **182:**870–873.
46. **Stanley, M. A., H. M. Browne, M. Appleby, and A. C. Minson.** 1989. Properties of a non-tumorigenic human cervical keratinocyte cell line. Int. J. Cancer **43:**672–676.
47. **Taichman, L. B., and R. F. LaPorta.** 1986. The expression of papillomaviruses in epithelial cells, p. 109–139. *In* N. P. Salzman and P. M. Howley (ed.), The Papovaviridae, vol. 2. The papillomaviruses. Plenum Press, New York, N.Y.
48. **Ustav, M., and A. Stenlund.** 1991. Transient replication of BPV-1 requires two viral polypeptides encoded by the E1 and E2 open reading frames. EMBO J. **10:**449–457.
49. **Volpers, C., P. Schirmacher, R. E. Streeck, and M. Sapp.** 1994. Assembly of the major and minor capsid protein of human papillomavirus type 33 into virus-like particles and tubular structures in insect cells. Virology **200:**504–512.
50. **Woodworth, C. D., J. Doniger, and J. A. DiPaolo.** 1989. Immortalization of human foreskin keratinocytes by various human papillomavirus DNAs corresponds to their association with cervical carcinoma. J. Virol. **63:**159–164.
51. **zur Hausen, H., and A. Schneider.** 1987. The role of papillomaviruses in human anogenital cancer, p. 245–263. *In* N. P. Salzman and P. M. Howley (ed.), The Papovaviridae, vol. 2. The papillomaviruses. Plenum Press, New York, N.Y.