Enhanced oncogenicity of Asian-American human papillomavirus 16 is associated with impaired E2 repression of E6/E7 oncogene transcription

Rosa M. Ordoñez,1 Ana María Espinosa,1 Dolores Javier Sánchez-González,2 Juan Armendáriz-Borunda3 and Jaime Berumen1

1Laboratorio de Medicina Genómica, Hospital General de México/UNISER-Facultad de Medicina, UNAM, Dr Balmis # 148, Colonia Doctores, Delegación Cuauhtémoc, CP 06726 Ciudad de México, Mexico
2Departamento de Biología Celular, Escuela Médico Militar, Universidad del Ejército y Fuerza Aerea, Cerrada de Palomas s/n, Colonia Lomas de San Isidro, Ciudad de México, Mexico
3Instituto de Biología Molecular y Terapia Genética, Centro Universitario de Ciencias de la Salud, Universidad de Guadalajara, Apartado Postal 2-123, 44281 Guadalajara, Jalisco, Mexico

Asian-American (AA) variants of human papillomavirus type 16 (HPV-16) are linked to a high incidence of cervical cancer in Mexico, with some evidence strongly suggesting that they are more oncogenic than European (E) variants, including their association with younger women and their higher associated risk of cervical cancer. Differences in the regulation of viral E6/E7 oncogene expression by the E2 protein may be involved in the higher oncogenicity of AA variants. In E variants, E6/E7 oncogene transcription is repressed by the E2 protein and is frequently up-regulated by the destruction of the E2 gene during viral integration. In contrast, the E2 gene is retained in full in most AA-positive carcinomas, raising the possibility of alternative mechanisms for increasing viral oncogene transcription. The authors investigated whether the higher oncogenicity of AA variants is linked to differences in E6/E7 oncogene transcription and the mechanism of E2 deactivation. E6/E7 and E1/E2 transcripts were explored by RT-PCR in 53 HPV-16-positive cervical carcinomas, 39 retaining (20 European and 19 AA) and 14 having lost (12 European and 2 AA) the E1/E2 genes, and transcription repression activity of the AA E2 genes was tested in four cell lines that constitutively express the β-galactosidase reporter gene driven by the viral long control region. E6/E7 oncogene transcripts were found in all carcinomas, but only those positive for AA variants with E1/E2 genes had complete E2 transcripts. E2 transcripts were down-regulated by splicing in E-positive carcinomas retaining E1/E2. AA E2 genes were impaired for repression of E6/E7 oncogene transcription in vivo. These results suggest that E6/E7 oncogene expression starts earlier in AA than E variant infections, since E variants need E2 to be destroyed or down-regulated.

INTRODUCTION

Almost a quarter of all cervical carcinomas in Mexico are attributable to Asian-American (AA) variants (subclasses AA-a and AA-c) of human papillomavirus type 16 (HPV-16) (Berumen et al., 2001). The high incidence of cervical cancer in Mexico may be explained in part by the poor coverage of the PAP screening programme, which reaches only 30% of adult women (Lazcano et al., 1999). However, the high frequency of HPV-16 AA variants may also contribute to the high incidence of cervical cancer in Mexican women, as there is some clinical and epidemiological evidence suggesting that AA variants of HPV-16 are more oncogenic than E variants. AA variants confer a nine times higher risk than E variants for cervical cancer (Berumen et al., 2001) and, together with other non-E variants, a 4–5 times higher risk than E variants for high-grade cervical intraepithelial neoplasia (CIN) lesions (Xi et al., 1997). Furthermore, AA variants are associated with women on average 8 years younger and with more aggressive tumours compared with E variants (Berumen et al., 2001).

The higher oncogenicity of an HPV-16 variant could result from increased neoplastic activity of the E6/E7 oncoproteins, more efficient virus replication or better stimulation of viral oncogene expression. Some experimental evidence supports a higher oncogenicity of AA variants. The E6 protein from an AA-a variant (isolate 512) consistently produced more serum and calcium differentiation-resistant colonies in primary human foreskin keratinocytes and...
stimulated p53 degradation better than the E variant E6 reference protein (Stoppler et al., 1996). Activity of the p97 promoter in AA variants was 1.7- to 3.3-fold higher than that of the E reference virus, suggesting increased expression of viral oncogenes (Veress et al., 1999; Kammer et al., 2000). The viral DNA copy number of AA variants is also much higher than that of E variants in cervical carcinomas, suggesting that AA variants replicate better (Casas et al., 1999).

Differences in the regulation of viral E6/E7 oncogene transcription by the E2 protein may also be involved in the higher oncogenicity of AA variants. For E variants, E6/E7 oncogene transcription is repressed by the E2 protein and is frequently up-regulated by the destruction of the E2 gene during viral integration (Schwarz et al., 1985; Schneider-Maunoury et al., 1987; Wilczynski et al., 1988; Cullen et al., 1991). In contrast, the E2 gene is retained in full in most AA-positive carcinomas (Casas et al., 1999), raising the possibility of alternative mechanisms for increasing viral oncogene transcription. The E2 protein of AA variants has 16 non-synonymous mutations distributed throughout the protein. This E2 variant protein could regulate the expression of E6/E7 genes differently. However, other factors not related to E2 variation may also be associated with the deregulation of E2, since half of the E1/E2-positive carcinoma samples in a study by Casas et al. (1999) had E variants.

The aim of this work was to analyse the E6/E7 and E1/E2 gene transcripts in cervical carcinomas positive for AA or E variants retaining the E1/E2 genes and to analyse the E6/E7 oncogene transcription. The E2 protein of AA variants has higher oncogenicity of AA variants. For E variants, E6/E7 transcription by the E2 protein may be also involved in the deregulation of E2, since half of the E1/E2-positive carcinoma samples in a study by Casas et al. (1999) had E variants.

**METHODS**

**Samples.** Fifty-three samples of invasive cervical carcinoma positive for HPV-16 DNA, 39 retaining (20 E and 19 AA) and 14 having lost (12 E and 2 AA) the E1/E2 genes, were included in this study. Samples were selected from a previous study where HPV-16 variants were detected (Berumen et al., 2001), based essentially on the availability of a tissue biopsy fragment stored at −70 °C when analysis of RNA was performed. The status of the E1/E2 region was investigated by PCR as described previously (Berumen et al., 1994). Tissue samples were obtained by direct open biopsy procedures; one fragment was used for routine pathological diagnosis and the other was snap-frozen in liquid nitrogen and stored at −70 °C for molecular analysis.

**Analysis of the HPV-16 genome by Southern blotting.** Southern blot analysis was performed in 11 of the 39 carcinomas positive for E1/E2 (five positive for E variants and six positive for AA variants). For Southern blots, 10 μg sample DNA was digested with BamHI and PstI (Gibco-BRL Life Technologies), following the manufacturer’s recommendations. In each experiment, the first lane contained a molecular mass standard (HindIII-digested phage lambda) and the second lane contained 50 copies per cell of HPV-16 linear DNA (725 pg) and 1450 pg HPV-16 DNA plus 10 μg human placental genome (Sigma) as carrier, digested with BamHI/PstI. HPV-16 DNA was used as a probe under stringent conditions and hybridization was conducted as described previously (Casas et al., 1999).

**Analysis of transcripts from E6/E7, E1/E2 and E6–E2 early regions by RT-PCR.** Total RNA was extracted with Trizol (Gibco-BRL Life Technologies) following the manufacturer’s recommendations. Purified RNA was treated with DNase I before being reverse-transcribed. One microgram total RNA was reverse-transcribed using the GeneAmp RNA PCR kit (Perkin-Elmer) in a 20 μl reaction containing 5 mM MgCl2, 1X PCR buffer (50 mM KCl, 10 mM Tris/HCl, pH 8.3), 1 mM each dATP, dCTP, dGTP and dTTP, 1 mM RNase inhibitor, 2.5 μM random hexamer primers and 1 U reverse transcriptase. The reaction was allowed to proceed for 15 min at 42 °C and was inactivated at 95 °C for 5 min. Five microlitres of this reaction was used for the amplification of specific cDNA by the addition of 20 μl of reaction mix containing 2 mM MgCl2, 1X PCR buffer, 1 μM each primer and 2.5 U AmpliTaq DNA polymerase (Perkin-Elmer). Reactions were carried out in a DNA GeneAmp PCR System 9600 (Perkin-Elmer). After denaturing for 2 min at 94 °C, 40 cycles were performed of 94 °C for 30 s, 60 °C for 1 min and 72 °C for 1 min. The annealing temperature was varied according to the Tm of the primers used for each region. A final elongation step was carried out at 72 °C for 8 min. The location of the primers used to detect the different transcripts is shown in Fig. 1. The sequence and position of forward (F) and reverse (R) primers were as follows.

F104 (5’-ATGTTTCAGGACCCACAGGA-3’, nt 104–123)

F856 (5’-TTAATCTACCATGGCTGATCC-3’, nt 856–876)

F1244 (5’-TTGAAAGCGAAGACAGGGG-3’, nt 1244–1263)

F2975 (5’-GAACCTCAACTAACGTTAGA-3’, nt 2975–2994)

F3192 (5’-CTGTTGAGGGTGTTAGTT-3’, nt 3192–3211)

R858 (5’-TTATGTTTCTGTGAAGCAGA-3’, nt 858–838)

R2865 (5’-GGGATGTGTTTCAAATAATG-3’, nt 2865–2845)

R3124 (5’-TCCATCAACTGCACTTCCA-3’, nt 3124–3105)

R3467 (5’-AGTCGTTGTGTTTACTCTCGG-3’, nt 3467–3447)

R3624 (5’-TGCATCATGGGTAGTTAGTT-3’, nt 3624–3605)

Transcripts from the E6/E7 region were detected with primers F104/R858; from the E1/E2 region, including E4, with primers F856/R3467; from the E1/E2 region, excluding E4, with primers F856/R2865; from the E2 gene alone, excluding E4, with primers F2975/R3124 and

**Fig. 1.** Genome map and schematic representation of mRNA transcripts from the early region of the reference HPV-16 strain. The upper panel shows the genome map with the ORFs in their proper reading frames shown as coloured rectangles. The location of the primers used for RT-PCR and sequencing are indicated by arrows on the genome. In the lower panel, coloured boxes represent the translated regions, bold black lines represent unspliced non-translated regions and dotted lines show the spliced regions. The donor and acceptor splice sites are shown by nucleotide numbers above each box. The proteins encoded by each transcript are shown on the right. The species of transcripts were named according to the 1996 sequence database (Los Alamos National Laboratory), and a new polycistronic transcript, described in this report (named Q and including the E6^E7 transcript), has been added. LCR, long control region.
F3192/R3624; from the E2C region with primers F1244/R3467; and from the E6–E2 region with primers F104/R3624. As a control for the stability and concentration of RNA, two primers to detect actin transcripts were used: ACT1 (5′-ATGATGGTGAGAATCC-3′, nt 1854–1874) and ACTInv (5′-GTTTCCTGTCAATGAGTC-3′, nt 2172–2152) (Dodd et al., 1992). To confirm the absence of DNA, an aliquot of each RNA sample was analysed by PCR for the E6/E7 region before the reverse transcription reaction was carried out. Biopsies of placenta were used to control the sensitivity of HPV DNA and were explored in parallel with the tumour samples. To control the sensitivity of PCR, 500 ng placental DNA with 100 fg viral DNA were used in each series of experiments. The analysis of PCRs was performed by electrophoresis in agarose gels, and the presence of transcripts from each viral region (E6/E7 and E1/E2) was confirmed by Southern blotting (data not shown), performed as described previously (Berumen et al., 1994). Precise identification of transcripts from the E6/E7, E1/E2 and E6–E2 regions was carried out by sequencing the bands obtained by RT-PCR in both directions. For sequencing, the fluorescent cycle sequencing method (BigDye Terminator Ready Reaction Kit; Perkin-Elmer) was used as described previously (Casas et al., 1999). Sequence analysis was performed using an ABI PRISM 310 Genetic Analyser System (Perkin-Elmer). Data analysis was carried out using the DNAsis software (Hitachi Software Engineering Co.). HPV sequences and base positions are numbered according to the 1997 sequence database (Los Alamos National Laboratory).

Measurement of the transcription repression activity of HPV-16 E2 variants

Plasmid constructions. The E2 genes from the E reference and AA-a and AA-c variants were amplified by PCR using the Pfu DNA polymerase (Stratagene) using the following primers containing restriction sites in the 5′ ends (underlined in the following primers): 5′-ACGGAGAATTCAGGAGAATTCGTTGCAACATGTAATAG-3′ (nt 2756–2784) and 5′-ACGGAGAATTCAGGAGAATTCGTTGCAACATGTAATAG-3′ (nt 3853–3825). E2 variant genes were amplified from genomic DNA derived from AA-a-positive (isolate MX68) and AA-c-positive (isolate MX70) cervical carcinomas, and the reference E2 gene was amplified from the cloned HPV-16 reference genome. The PCR mixture was as follows: 500 ng tumour DNA (or 1 ng cloned HPV-16) was added to achieve a final volume of 50 μl containing 30 mM Tris/HCl, pH 8.8, 2 mM MgCl2, 50 mM KCl, 1 mM DTT, 200 μM each dATP, dTTP, dGTP and dCTP, 1 μM each primer, 100 ng BSA μl−1 and 1 l-O U Pfu DNA polymerase (Stratagene). The reaction was carried out in a DNA GeneAmp PCR System 9600 (Perkin-Elmer). After denaturing for 3 min at 94 °C, 15 cycles were performed of 94 °C for 30 s, 54 °C for 1.5 min and 72 °C for 1.5 min. A further 25 cycles were then carried out under the same conditions but with an annealing temperature of 58 °C. A final elongation step was carried out at 72 °C for 7 min. PCR products were purified by two phenol/chloroform extractions, digested with the appropriate restriction enzymes (BamHI/EcoRI, purified using the Qiaex II Kit (Qiagen) and then inserted into the BamHI/EcoRI sites of the Bluescript SK− vector using 6 U T4 DNA ligase (Gibco-BRL, Life Technologies) in 20 μl 50 mM Tris/HCl, 10 mM MgCl2, 1 mM ATP, 1 mM DTT and 5 % PEG-8000 at 16 °C for 1 h. Recombinant vectors were cloned in Escherichia coli DH5α. E2 genes from the E reference and AA-a and AA-c variants were subcloned into the BamHI/EcoRI restriction sites of pcDNA3 expression vector (Invitrogen) generating plasmids p16E2/E, p16E2/AA-a and p16E2/ AA-c, respectively (see Fig. 5a). The complete long control region (LCR) of the E HPV-16 reference (nt 7036–103), cloned upstream from the β-galactosidase reporter gene into the pc4-AUG vector (Thummel et al., 1988) (construction plCR16-β-gal; see Fig. 5a), was kindly provided by Alejandro Garcia-Carranca. All molecular constructs were verified by DNA sequencing on an Applied Biosystems automated sequencer.

Cell cultures and transfections. The human cervical carcinoma cell lines C33a, HaCat, SiHa and HeLa were cultured in Dulbecco’s modified Eagle’s medium (Gibco-BRL, Life Technologies) supplemented with 10 % fetal bovine serum (Gibco-BRL, Life Technologies) and 120 mg penicillin and streptomycin l−1 in a 5 % CO2 atmosphere. Cells were transfected by the standard calcium phosphate co-precipitation method in 60 mm Petri dishes, as described previously (Guido et al., 1992). Briefly, 24 h before transfection, the cells were grown in 60 mm tissue culture dishes to a density of 7.5 × 105 cells per dish. C33A and HaCat cell lines were transfected with 10 μg plasmid plCR16-β-gal and stable transfectants were selected with 800 μg genetin ml−1 (Gibco-BRL Life Technologies) over a 2-month period. Single-cell clones were generated by limiting dilution in 96-well plates. The cell clones were frozen in liquid nitrogen until required. Selected clones of C33a and HaCat, which constitutively express the β-galactosidase protein, as well as SiHa and HeLa cell lines, which constitutively express the E6/E7 oncoproteins of HPV-16 and HPV-18, respectively, were transiently co-transfected with 10 μg of the empty vector (pcDNA3) or the E2 expression vectors plus 10 μg green fluorescent protein (GFP)-expressing plasmid pQBi25-f (Q-Biogene). Sixteen hours after transfection, the cells were washed, supplied with fresh medium and cultured for another 24 h. For C33a and HaCat experiments, the cells were harvested and cellular extracts were obtained by four freeze–thaw cycles in 250 mM Tris/HCl, pH 8.0, 15 % glycerol and 5 mM DTT buffer. Proteins were quantified (Bradford, 1976) and cellular extracts were normalized for protein concentration. For experiments with all four cell lines, total RNA was extracted with guanidine isothiocyanate using the GlassMax RNA microisolation kit (Gibco-BRL, Life Technologies) following the procedure recommended by the manufacturer. Transfection efficiency, measured by immunofluorescence, ranged between 50 and 70 %, being highest in C33a cells followed successively by HeLa, SiHa and HaCat cells. Experiments were performed in triplicate and repeated independently at least three times.

β-Galactosidase assays. Extracts of C33a and HaCat cells were prepared and assayed for β-galactosidase activity essentially as described previously (Norton & Coffin, 1985). Briefly, 50 μl lysate was made up to a final volume of 800 μl with PBS, mixed with 200 μl freshly prepared 4 mg o-nitrophenyl galactoside (ONPG) ml−1 and then incubated at 37 °C until a yellow colour was apparent. The samples were placed on ice and enzyme reactions were terminated by the addition of 500 μl 1 M Na2CO3. The amount of ONPG hydrolysed was determined spectrophotometrically and β-galactosidase activity was estimated as follows: units = (A405 × 380)/min.

Transcript detection by hybrid capture assay. Transcripts from E6/E7, E2 and β-actin genes were determined using a modified hybrid capture method using the HPV-Test Hybrid Capture II Kit (Digene), adapted to analyse total RNA with DNA probes. Probes for the E6/E7 genes of HPV-16 and HPV-18, the E2 gene of HPV-16 and β-actin were generated by PCR using primers and a PCR protocol described previously (Berumen et al., 1994; Dodd et al., 1992). For hybrid capture, 1 μg total RNA extracted from cell lines and 10 μg probes were used. Probe DNA was dissolved in 75 μl denaturation reagent and incubated for 45 min at 65 °C in a water bath. Cellular RNA was dissolved in 25 μl probe diluent solution. Both solutions were then mixed for 3 min at 1100 r.p.m. and 25 °C in an orbital shaker and incubated for 60 min at 65 °C in a water bath for hybridization. The hybridization reactions were transferred to the capture microplate wells and the procedure to detect the signal was performed as recommended by the manufacturer. The capture microplate was read in a DML-2000 Luminometer (Digene). Expression of E6/E7 was normalized against expression of the β-actin gene.

Determination of viral DNA and RNA copy number by hybrid capture assay. Viral DNA copy number and the level of E6/E7 transcripts from each viral region (E6/E7 and E1/E2) was confirmed by quantitative PCR using 500 ng placental DNA with 100 fg viral DNA. To standardize the sensitivity of PCR, 500 ng placental DNA with 100 fg viral DNA were used in each series of experiments. The analysis of PCRs was performed by electrophoresis in agarose gels, and the presence of transcripts from each viral region (E6/E7 and E1/E2) was confirmed by Southern blotting (data not shown), performed as described previously (Berumen et al., 1994). Precise identification of transcripts from the E6/E7, E1/E2 and E6–E2 regions was carried out by sequencing the bands obtained by RT-PCR in both directions. For sequencing, the fluorescent cycle sequencing method (BigDye Terminator Ready Reaction Kit; Perkin-Elmer) was used as described previously (Casas et al., 1999). Sequence analysis was performed using an ABI PRISM 310 Genetic Analyser System (Perkin-Elmer). Data analysis was carried out using the DNAsis software (Hitachi Software Engineering Co.). HPV sequences and base positions are numbered according to the 1997 sequence database (Los Alamos National Laboratory).
transcripts were measured in 26 tumours [20 E1/E2-positive (10 AA and 10 E variants) and 6 E1/E2-negative (all E variants)] by the hybrid capture method using the HPV-Test Hybrid Capture II Kit. For viral DNA copy number, the method was used with the RNA probes provided with the kit under the conditions recommended by the manufacturer. To measure the level of E6/E7 transcripts, the method was adapted to analyse total RNA with DNA probes, as has been described above. For hybrid capture, 500 ng total RNA or DNA from each sample was used. To calculate the viral DNA or RNA copy number per cell, a standard curve was constructed as described previously (Berumen et al., 1994), but instead of measuring the curve points (1, 10, 50, 100, 500, 1000, 3000 and 10 000 viral DNA copies) by the dot blot technique, they were measured by the hybrid capture method. A non-linear regression analysis of this data gave the best fit into a hyperbolic equation ($r^2=0.98$, $P<0.001$) and the copy number per cell of a given tumour was extrapolated from this curve. The amount of RNA was normalized against the levels of β-actin transcripts.

Apoptosis assays. The induction of apoptosis by E2 was measured in HeLa and SiHa cells by flow cytometry and confocal laser-scanning microscopy using the terminal deoxynucleotidyl transferase-mediated dUTP-rhodamine nick end labelling (TUNEL) technique (In situ Cell Death Detection kit, TMR red; Roche Applied Science). For flow cytometry, culture cells and transfections were performed under the same conditions as E2 repression assays (see above). Forty hours after transfection, cells were washed in PBS and fixed with 4% paraformaldehyde/PBS at 22 °C for 1 h. Cells were permeabilized with PBS containing 0.1% Triton X-100 and 0.1% sodium citrate and incubated with 50 μl TUNEL enzyme mixture following the manufacturer’s recommendations. Cells were analysed using a FACScan flow cytometer (Becton Dickinson). Cells positive for GFP were gated for analysis of apoptosis. For confocal microscopy, SiHa and HeLa cells were seeded at 2 × 10⁴ cells per well on to a tissue culture plate or slide (Lab-Tek) 24 h before transfection. The co-transfections were carried out using 250 ng of each plasmid under the same conditions described above. Forty hours after transfection, the chamber/slides were washed in PBS, fixed, permeabilized and analysed by the TUNEL technique under the conditions recommended by the manufacturer. Labelled specimens were scanned with an LSM-5 Pascal confocal laser-scanning microscope (Carl Zeiss) linked to a Zeiss inverted microscope Axiovert 200 M equipped with a Zeiss 40 × plan-neofluar oil immersion lens with a numerical aperture of 1.3 (Carl Zeiss). The 488 nm line from an argon laser was used for the GFP signal and the 543 nm line from a helium–neon laser was used for the TMR red TUNEL signal. Fluorescence signals from GFP were observed using a dichroic beam splitter (HFT 488/543/633; Carl Zeiss) and emission filter (BP505-530; Carl Zeiss). Optical fluorescence signals of apoptotic cells labelled with TMR red were observed using a dichroic beam splitter (HFT 488/543/633; Carl Zeiss) and emission filter (LP560; Carl Zeiss). A computer equipped with KS-300 software version 3.0 (Carl Zeiss) was used for operating the system. Detector gain and pinhole aperture were automatically adjusted. The image resolution was 1024 × 1024 pixels (8 bits, 256 colour levels). Experiments were performed in triplicate and repeated independently three times.

Statistical analyses. All experiments in culture cells were performed in triplicate and repeated independently at least three times. Representative experiments are shown. For DNA and RNA copy number and E2 activity, the results were expressed as the mean ± SD, and the Student’s t-test or Mann–Whitney U-test were used, as appropriate, to assess the statistical significance of differences between the groups. The χ² test or Fisher’s exact test was used, as appropriate, to assess the statistical significance of differences in the frequency of transcript detection. The association of DNA copy number and RNA copy number was measured by linear regression. Differences were considered statistically significant if $P<0.05$. The Sigma Stat (SPSS Inc.) software program was used for statistical analyses. All statistical tests were two-sided.

RESULTS AND DISCUSSION

E6/E7 transcripts were explored by RT-PCR using primers at the start of E6 (F104) and at the end of E7 (R858; Fig. 1). Four different E6/E7 transcripts were detected in the total population analysed, which were 754, 571, 454 and 238 bp in length (Fig. 2a). The first three transcripts have been reported previously (Doorbar et al., 1990; Sherman et al., 1992) and correspond to unspliced E6, E6*, E7 (spliced from nt 226 to 409) and E6**E7 (spliced from nt 226 to 526) products, respectively (Figs 1 and 2e). The transcript of 238 bp was spliced from nt 226 to 742 and included the 5’ region of E6 (nt 104–226) and the 3’ region of E7 (nt 742–855) (Figs 1 and 2a and e). This novel transcript, named E6^E7, encoded a truncated E6/E7 fused protein of 79 aa, 41 derived from the N terminus of the E6 protein and 38 from the C terminus of the E7 protein. All 53 tumours explored, irrespective of the status of the E1/E2 genes, showed one or more of the four detected E6/E7 transcripts. The E6^*E7 transcript was detected in almost all the samples explored (48 out of 53; 90.6%), whereas the full E6, E7, E6**E7 and the E6^*E7 transcripts were detected only in 66%–0% of 32%–1% of tumour samples, respectively. All but one (E6*, E7) transcripts were distributed evenly between the E1/E2-positive and E1/E2-negative tumours (Table 1). The detection of the E6^*E7 transcript was higher in E1/E2-positive samples (97.4%) than in E1/E2-negative ones (71.4%; $P=0.014$, Fisher’s exact test). On the other hand, the frequency of the known E6/E7 transcripts (full E6, E7, E6**E7 and E6^*E7) was not different between AA- and E-positive samples (Table 1). However, the novel E6^*E7 transcript was detected mostly in the AA-positive samples. This transcript was detected in 66%–7% of AA-positive samples, but only in 9.4% of E-positive samples ($P<0.001$, χ² test; Table 1). These data indicated that, for both AA and E variants, the expression of E6/E7 genes was not impaired by the presence of the E2 gene and suggested that the E2 gene was not expressed or was functionally inactive in the E1/E2-positive tumours.

As the retention of E1/E2 is directly associated with copy number (Berumen et al., 1994), quantitative differences in E6/E7 expression between these groups were further explored. E6/E7 transcripts and DNA viral copy number were analysed by hybrid capture in 20 E1/E2-positive and six E1/E2-negative samples. All E1/E2-positive samples had a high number of E6/E7 transcripts and the mean copy number of these transcripts (1352 ± 363·5) was 13·8-fold higher than that of E1/E2-negative (98±3·4·5) samples (Table 2; $P=0.001$, Mann-Whitney U-test). The mean copy number of E6/E7 transcripts was 2·1-fold higher in the AA-positive (1563 ± 682) than in the E-positive (750·6 ± 669·3) tumours ($P=0.006$; t-test). However, this difference disappeared when the tumours were classified according to the presence ($P=0.14$; t-test) or absence of...
Fig. 2. Expression of the E6/E7, E1/E2 and E6–E2 regions in HPV-16-positive cervical carcinomas retaining the E1/E2 region detected by RT-PCR. (a) Transcripts from E6/E7 genes of representative cervical carcinomas. Experiments with E carcinomas are shown in lanes 1–4 and AA carcinomas in lanes 5–16. Transcripts were amplified with primers F104/R858. Bands of 754, 571, 454 and 238 bp correspond to transcripts E6,E7, E6*I,E7, E6*II,E7 and E6^E7, respectively. (b) Transcripts from β-actin of the samples shown in (a). (c) RT-PCR of the E1/E2 region from cervical carcinomas samples positive for AA (T1–T3) and E (T4–T6) variants. A band of 133 bp was amplified with primers F856/R3467 and corresponded to the E1^E4 transcript; a band of 149 bp, obtained with primers F2975/R3124, corresponded to the 5' region of the E2 transcript; and bands of 180 and 307 bp were amplified with primers F856/R2865 and corresponded to the E2 and E1C,E2 transcripts, respectively (Fig. 1). (d) RT-PCR of the E6–E2 region from seven cervical carcinoma samples positive for AA (T9) and E (T7, T8, T10–T13) variants. Bands are labelled with the same letter as the polycistronic species in Fig. 1. The bands were amplified using a primer at the beginning of E6 (F104) and one at the end of E2 (R3624). Lane M contains a 1 kb marker plus DNA ladder. (e) Electropherograms of the partial DNA sequence of some of the transcripts in (a), (c) and (d). Sequences show the nucleotides around the splice junction, which are indicated above the electropherograms.
transcripts was detected in E1/E2-negative samples (E1–E4, E2, E1C and E2C; Fig. 1). While none of these transcripts reported from the E1/E2 region were analysed to determine whether the E2 gene is expressed, all transcripts explored in the E- or AA-positive samples showed the expected band for this transcript (149 bp; Fig. 2c, lane b), but it was not found in any E-positive carcinomas (Fig. 3). The fact that the expression of E6/E7 is much higher in E1/E2-positive than in E1/E2-negative tumours also supports the hypothesis that the E2 gene could be down-regulated in these carcinomas and that E6/E7 expression levels depend on the amount, rather than the status, of the viral genomes.

To determine whether the E2 gene is expressed, all transcripts reported from the E1/E2 region were analysed (E1–E4, E2, E1C and E2C; Fig. 1). While none of these transcripts was detected in E1/E2-negative samples (n = 14), one or more were detected in all E1/E2-positive samples explored (n = 39). With primers located at the start and end (F856/R3467; Fig. 1) of the E1/E2 region, only the E1/E4 transcript (133 bp band; Fig. 2c, lane a) was detected in all E- or AA-positive samples. This transcript excluded a spliced region from nt 881 to 3357, including E2 (Figs 1 and 2e) and encodes the E4 protein. Since E1/E4 is the most abundant transcript detected in tumours (Stoler et al., 1992; Bohm et al., 1993; Sherman et al., 1992) and cell lines (Sherman & Alloul, 1992; Rohlf et al., 1991), the lack of other E1/E2-derived transcripts could be the result of preferential PCR amplification. Therefore, a short 5′ region of the full E2 transcript, not including E1/E4, was investigated with primers F2975/R3124 (Fig. 1). Surprisingly, all AA-positive samples showed the expected band for this E2 transcript (149 bp; Fig. 2c, lane b), but it was not found in any E-positive carcinomas (P < 0.001; χ² test). This E2 transcript can be part of eight different polycistronic transcripts (E–K and M in Fig. 1) and includes a 3′ region from the E1 gene, of either 106 bp (transcript E2; Fig. 1) or 233 bp (transcript E1C,E2; Fig. 1). Both transcripts were detected with primers F856/R2865 (Fig. 1) in all AA-positive samples, but not in E-positive samples [E2: 180 bp, Fig. 2c (lane c) and e; and E1C,E2: 307 bp; Fig. 2c (lane c) and e]. To determine whether the mRNAs contained the entire E2 transcript and encoded the full-length E2 protein, the 3′ end of this RNA species was investigated with primers located at the 3′ end of E2 (F3192/R3624; Fig. 1). Only the AA-positive samples were positive for this transcript and a PCR product of 432 bp was detected (data not shown). The E2C transcript was analysed in only six AA-positive samples and was present in them all. The abundant expression of E1/E4 and the lack of E2 and E1 transcripts strongly suggest that E2 is down-regulated.

---

**Table 1.** Frequency of E6/E7 gene transcripts in cervical carcinomas positive for HPV-16 classified by virus class and the status of E1/E2 genes

<table>
<thead>
<tr>
<th>Groups</th>
<th>n</th>
<th>Frequency of E6/E7 gene transcripts (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>E6,E7</td>
<td>E6*/E1,E7</td>
</tr>
<tr>
<td>E1/E2 status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>39</td>
<td>26 (66-7)</td>
</tr>
<tr>
<td>Negative</td>
<td>14</td>
<td>9 (64-3)</td>
</tr>
<tr>
<td>Virus class</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>21</td>
<td>14 (66-7)</td>
</tr>
<tr>
<td>E</td>
<td>32</td>
<td>21 (65-6)</td>
</tr>
<tr>
<td>Total†</td>
<td>53</td>
<td>33 (65-0)</td>
</tr>
</tbody>
</table>

†All 53 tumours explored showed one or more of the four detected E6/E7 transcripts.

a vs b, P = 0.014, Fisher’s exact test; c vs d, P < 0.001, χ² test.

**Table 2.** Viral copy number (DNA and RNA) in HPV-16-positive carcinomas classified by the status of E1/E2 genes and viral classes

<table>
<thead>
<tr>
<th>Groups</th>
<th>Viral copy number per cell (mean±SD) (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>E1/E2-positive</td>
</tr>
<tr>
<td>E6/E7 DNA</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>142±3±6±8 (10)</td>
</tr>
<tr>
<td>AA</td>
<td>215±95±4 (10)</td>
</tr>
<tr>
<td>Both</td>
<td>180±7±87±8 (20)</td>
</tr>
<tr>
<td>E6/E7 RNA</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>114±2±54±0 (10)</td>
</tr>
<tr>
<td>AA</td>
<td>1563±682 (10)</td>
</tr>
<tr>
<td>Both</td>
<td>1352±636±5 (20)</td>
</tr>
</tbody>
</table>

a vs b, P = 0.079, t-test; c vs d, P < 0.001, Mann–Whitney U-test; c vs f, P = 0.003, t-test; g vs h, P = 0.14, t-test; i vs j, P = 0.001, Mann–Whitney U-test; k vs l, P = 0.006, t-test.
AA, Asian-American variants; E, European variants.
NI, Not included.
by alternative splicing in E-positive carcinomas retaining complete E1 and E2 genes. These results are in agreement with the detection of high levels of E1^E4 in the cytosol and E1 and E2 transcripts in the nuclei of cancer cells from cervical carcinomas positive for HPV-16 (Stoler et al., 1992). Similar E1/E2 transcript patterns have been found in cervical carcinomas derived from Caucasian patients (Stoler et al., 1992), in whom it is assumed that E variants predominate (Yamada et al., 1997). E2 transcripts have been previously found mainly in CIN lesions and in the cell lines HPKII and CaSki (Sherman et al., 1992; Bohm et al., 1993; Rohlfis et al., 1991; Sherman & Alloul, 1992). The nucleotide changes in the sequence of the E2 gene from AA variants (Casas et al., 1999) could be related to differences in the transcription pattern. However, there are no changes in the sequence of known consensus splicing sites located in the E2 gene (Doorbar et al., 1990; Sherman et al., 1992), nor in primer annealing sites. The translation of E2 from such transcripts in in vitro and in vivo experiments (Alloul & Sherman, 1999) suggests that E2 could also be translated in tumour samples.

To determine whether differences in E2 gene expression between AA- and E-positive carcinomas are related to differences in the status of the viral genome, 11 of the 39 carcinomas positive for E1/E2 genes by PCR were further investigated by Southern blot analysis, including five samples positive for E variants and six positive for AA variants (nine are shown in Fig. 4). All tumours had intact E1/E2 genes (2–8 kb band; Fig. 4) and showed the typical seven-band PstI/BamHI cleavage pattern of HPV-16; however, seven (three E and four AA) also showed larger-sized bands above the E1/E2 band (Fig. 4; lanes 2, 3, 4, 6 and 7), a pattern consistent with tandem integration or episomal multimeric forms. The remaining four (two E and two AA) showed episomal patterns (Fig. 4; lanes 1, 5, 8 and 9). All 11 E1/E2-positive tumours contained more than 50 viral copies per cell and the bands observed were reasonably strong, including the E1/E2-containing band (2–8 Kb). These patterns are similar to those reported previously in a larger sample of carcinomas positive for E1/ E2 by PCR (Casas et al., 1999). Since Southern blotting is much less sensitive than PCR, the Southern blotting pattern found in a specific tumour may represent the predominant pattern in most of its tumour cells. These data indicate that there were no important differences in the viral genome status between AA- and E-positive samples, and thus differences in the expression of the E2 gene cannot be explained by this factor. The data also suggest that expression of E6/E7 genes seems to be independent of the status of the viral genome, since those tumours showing either episomal or tandem integration patterns both had high levels of E6/E7 expression.

E6/E7 transcripts in tumours positive for E1/E2 by either PCR or Southern blot do not necessarily come from tumours cells that contain intact E1/E2 genes. The ability to amplify E1/E2 by PCR does not exclude the presence of integrated viral forms with E1/E2 genes destroyed in some cells. In fact, tumours positive for E1/E2 may contain both cells with complete E1/E2 genes and cells with the E1/E2 genes disrupted (Berumen et al., 1995). Therefore, in these tumours, the expression of viral oncogenes could come rather from those cells containing the interrupted the E2 gene. To investigate whether the E6/E7 and E1/E2 transcripts came from the same viral genomes containing a complete early region, we explored seven E1/E2-positive samples (one AA and six E) with primers located at the start of the E6 gene and at the end of the E2 gene (F104/ R3624, Fig. 1). All seven samples showed transcripts composed from the E6/E7 and E1/E2 regions (Fig. 2d). Six different composed transcripts were detected in the samples analysed, including five reported previously (A, H, B, C and D) and one non-reported transcript containing the hybrid E6^E7 transcript and the spliced product E1^E4 (Q in Figs 1 and 2d). These and previously reported data (Sherman et al., 1992; Sherman & Alloul, 1992) clearly indicate that transcripts from E6/E7 genes can arise from genomes with an intact E2 gene.
Since the E2 gene of AA-a and AA-c variants has 16 non-conserved mutations distributed throughout the entire open reading frame (Casas et al., 1999), we determined whether these proteins repressed the viral oncogene transcription in four cell lines, two stably transfected with the β-galactosidase reporter gene driven by the LCR of the reference HPV-16 strain (C33a and HaCat) and two cell lines derived from cervical carcinomas with integrated HPV-16 (SiHa) or HPV-18 (HeLa) genomes, which constitutively express the E6/E7 oncogenes and do not express the E2 gene (Schwarz et al., 1985; Sherman et al., 1992).

While the construct with the E E2 variant gene (p16E2/E) strongly repressed the basal β-galactosidase activity (59.5% and 18.5% activity compared with the vector only) or E6/E7 oncogene expression (51.4% and 40.1% activity compared with the vector only), the construct with the AA-c E2 variant gene (p16E2/AA-c) repressed the LCR to a much lesser degree in all four cell lines tested, although the difference was statistically significant (Fig. 5b, c) (P < 0.05; t-test). The E2 protein of the AA-a variant repressed the LCR to a lesser degree than the European E2 protein in the four cell lines (Fig. 5b, c), but was statistically significant for C33a, SiHa and HeLa cells (P < 0.05; t-test). The difference in E2 repression activity between the AA-c and AA-a variants was statistically significant in all cell lines (P < 0.05; t-test). No difference in the transcription level of E2-transfected genes was detected in experiments with the three E2 variant genes (data not shown). These results could suggest that E2 repression is impaired by sequence variation in AA variants and may explain the expression of E6/E7 oncogenes in all AA-positive carcinomas, even the presence of transcripts with coding potential for the complete E2 protein. Although AA-a and AA-c E2 proteins were examined in the context of the reference E LCR, it can be reasonably assumed that they could behave similarly in the context of their own LCR, since the three LCRs do not differ in the nucleotide sequence of E2-binding sites (Veress et al., 1999; Kamber et al., 2000). Furthermore, E2 proteins from AA-a and E variants had similar transactivation effects on the LCR of both variant classes in transient transfection assays (Veress et al., 1999), and in the present work the repressor activity of E2 from the three HPV-16 variants tested was not significantly different for the LCR of HPV-16 or HPV-18 (Fig. 5). We cannot exclude the possibility that the lower transcription repression activity of AA E2 variant proteins could be also associated with a decreased stability or impaired nuclear localization of these proteins (Abrio et al., 1996).

Several groups have reported that full-length E2 expression induces apoptosis in many types of cell (Sanchez-Perez et al., 1997; Webster et al., 2000, 2001). Although it is possible that the ‘repression’ effect demonstrated above could be due to cell apoptosis rather than E2-induced repression of the target gene, this seems to be unlikely, since the expression level of the β-actin gene was not different among the experimental groups (data not shown). Furthermore, in previous reports some authors have shown that E2 failed to induce apoptosis in C33 and HaCat cell lines (Desaintes et al., 1997), two of the cell lines where we demonstrated the lower E2 repression.

![Fig. 5. Gene expression regulation of the viral E6/E7 oncogenes by E2 variant genes in vivo.](http://vir.sgmjournals.org)
Fig. 6. Confocal laser-scanning microscopy of SiHa (a) and HeLa (b) cells co-transfected with E2- and GFP-expressing plasmids and stained using the TUNEL apoptosis assay. For both panels, the first row (a–e) shows the green channel, which detects GFP; the second row (f–j) shows the red channel, which detects the TUNEL signal; and the third row (k–o) shows the co-localization channel. The first column (a, f and k) shows non-transfected cells, the second column (b, g and l) shows cells transfected only with pBQI25 plasmid, the third column (c, h and m) shows cells co-transfected with pBQI25-f and p16E2/E, the fourth column (d, i and n) shows cells co-transfected with pBQI25-f and p16E2/AA-a, and the fifth column (e, j and o) shows cells co-transfected with pBQI25-f and p16E2/AA-c. Examples of apoptotic signals are indicated with arrows. The apoptotic signals in transfected cells stain yellow in the third channel. Magnification × 400.
activity of AA variants. To confirm that differences in the E2 repression activity between the AA and E variants was not associated with differences in the capacity of E2 to induce apoptosis, we measured apoptosis using the TUNEL technique in SiHa and HeLa cells after co-transfection with E2 plasmids and a GFP-expressing plasmid (pQBI25-f) under the same conditions used for the E2 repression assays. Similar levels of apoptosis were observed with the three E2 variant genes by flow cytometry analysis. The proportion of apoptosis induced by E, AA-a and AA-c E2 genes in the transfected cells was $26.1 \pm 6.7$, $17.5 \pm 13.6$ and $25.9 \pm 9.8\%$, respectively, in HeLa cells, and $52.2 \pm 19.6$, $51.7 \pm 43.4$ and $41.1 \pm 23.2\%$, respectively, in SiHa cells. Similar results were obtained by confocal microscopy analysis. A representative set of experiments explored by confocal microscopy is shown in Fig. 6. These data indicated that differences in the expression of E6/E7 oncogenes between cells transfected with E and AA E2 variant genes are directly related to differences in the repression activity of E2 proteins, rather than differences in the apoptosis activity of E2 proteins.

E2 inactivation by alternative splicing or sequence variation suggests that a selective pressure would not be exerted that would lead to the obliteration of the E2 gene. Disruption of E1/E2 genes could instead be associated with cellular differentiation and sequence variation of the E1/E2 genes. For instance, whereas all well-differentiated tumours retained the complete E1/E2 genes, only one-third of less-differentiated carcinomas did so (Berumen et al., 1994). In agreement with these data, the splicing of E2 through the formation of E1/E4 transcripts occurs essentially in well-differentiated cells in vitro (Pray & Laimins, 1995; Frattini et al., 1996) or in the upper well-differentiated cells of tumour samples (Stoler et al., 1992). The disruption of E1/E2 genes occurs much more frequently in E than in AA variants (Casas et al., 1999). Therefore, as the E viral genome remains in the basal undifferentiated cells, the integration and disruption of the E2 gene could be favoured. In contrast, in well-differentiated cells, the E viral genome retaining the complete E1/E2 region could deactivate E2 by splicing. On the other hand, for AA variants there would not be a selective pressure to induce the disruption or splicing of the E2 gene.

In addition to differences in the biological activities of the E6 protein (Stoppler et al., 1996), viral copy number (Casas et al., 1999) and up-regulation of the p97 promoter (Veress et al., 1999; Kammer et al., 2000), the enhanced oncogenicity of AA variants could be further associated with the novel E6/E7 transcript and with the lower p97 promoter repression of E2 variant proteins. The association of the E6/E7 transcript with AA-positive tumours is clear; however, not all AA tumours had this transcript, while some of the E variants did. Furthermore, the protein encoded by this transcript does not include regions involved with cell transformation and immortalization (Myers & Androphy, 1995; Halpern & Munger, 1995). This protein contained 41 of the 43 amino acids of the E6*I truncated protein and, like this protein, could function as a transactivator of the p97 promoter (Shirasawa et al., 1994). The low repression activity of E2 suggests that increased expression of E6/E7 oncogenes may occur much earlier in AA tumours. Therefore, the time period between viral infection and the emergence of a frankly invasive cancer may be decreased. This hypothesis could explain the association of the AA-c variant with patients 11 years younger than those with E variants (Berumen et al., 2001). In the case of E HPV-16 variants, the increased expression of the E6/E7 oncogenes is linked with disruption of the E2 gene during viral integration (Schwarz et al., 1985; Schneider-Maunoury et al., 1987; Romanczuk et al., 1990) or down-regulation of E2 by splicing (this paper), which may depend on cellular factors. Consequently, the lapse from infection time to development of an invasive tumour could be longer. These hypotheses need to be confirmed since, in the W12 cell line, E2 has no effect on transcription of E6/E7 from episomal viral DNA (Bechtold et al., 2003).

ACKNOWLEDGEMENTS

This work was supported by the National Council of Science and Technology (CONACYT), grant number 8135/A1 (to J. B.) and was undertaken mainly in the Postgraduate School of Medicine of the Mexican Army. We are grateful to Dr Susana Kofman and Alejandro Garcia for help with laboratory resources.

REFERENCES


papillomavirus type 16 have extensive mutations in the E2 gene and are highly amplified in cervical carcinomas. Int J Cancer 83, 449–455.


