

# Selection of cervical keratinocytes containing integrated HPV16 associates with episome loss and an endogenous antiviral response

Mark R. Pett\*, M. Trent Herdman\*, Roger D. Palmer\*, Giles S. H. Yeo†, Mahmud K. Shivji\*, Margaret A. Stanley‡, and Nicholas Coleman\*\*§

\*Medical Research Council Cancer Cell Unit, Cambridge CB2 2XZ, United Kingdom; †Cambridge Institute for Medical Research, Cambridge CB2 2XY, United Kingdom; and ‡Department of Pathology, University of Cambridge, Cambridge CB2 1QP, United Kingdom

Communicated by Sheldon Penman, Massachusetts Institute of Technology, Cambridge, MA, January 4, 2006 (received for review December 15, 2005)

Integration of high-risk human papillomavirus (HRHPV) into the host genome is a key event in cervical neoplastic progression. Integration is associated with deregulated expression of the viral oncogenes E6 and E7 and acquisition of a selective growth advantage for cells containing integrants. Overexpression of the viral transcriptional regulator E2 from heterologous promoters has an inhibitory effect on transcription from integrated HRHPV. Therefore, we hypothesized that loss of E2-expressing episomes from cells in which integration had previously occurred would be required for such cells to gain a growth advantage. Using the unique W12 model of cervical squamous carcinogenesis, we show that cells containing integrated HPV16 reproducibly emerged during long-term culture when there had been a rapid fall in episome numbers. During the period of emergence, it is possible to isolate single-cell clones containing an intracellular mixture of the integrant being selected and episomes at reduced load. The lower level of E2 expression seen in such cells is associated with partial inhibition of transcription from the HPV16 integrant. Full deregulation is not observed until complete loss of E2-expressing episomes occurs. Microarray analysis showed that episome loss was closely associated with endogenous activation of antiviral response genes that are also inducible by the type I IFN pathway. Taken together, our results indicate that episome loss, associated with induction of antiviral response genes, is a key event in the spontaneous selection of cervical keratinocytes containing integrated HPV16. We conclude that cervical carcinogenesis requires not only HRHPV integration, but also loss of inhibitory episomes.

human papillomavirus | cervix | integration | interferon | progression

Integration of high-risk human papillomavirus (HRHPV) into the host genome is an important step in cervical neoplastic progression (1, 2). Integrated viral genomes from which HRHPV early genes are transcribed have been detected in  $\approx 87.5\%$  of cervical malignancies (3). Integration usually causes deletion or disruption of the viral regulatory E2 gene, while retaining a variable segment including the E6 and E7 oncogenes and the upstream regulatory region (4, 5). Overexpression of E2 from heterologous promoters in cells harboring integrated HRHPV can repress the early promoter of the integrated virus, causing a sharp reduction in E6 and E7 expression (6). Thus, HRHPV integration and disruption/deletion of E2 leads to increased expression of the viral oncogenes (7, 8). Cells containing integrated HRHPV acquire a growth advantage over cells harboring episomal HRHPV (the natural viral state in productive infections) and show increased genomic instability (9–11).

Cervical keratinocyte cell lines established from precursor low-grade squamous intraepithelial lesions have indicated that episomal HRHPV genomes are maintained at  $\approx 100$  copies per cell in the basal region of an infected epithelium (12, 13). Viral integration is therefore most likely to occur in cells containing this number of episomes. It has recently been suggested that, whereas overexpression of E2 can inhibit the early promoter of

integrated HRHPV, it has little or no effect on the transcription from episomal HRHPV (14). It follows that physiological levels of E2 expressed in episomally infected keratinocytes could inhibit transcription from coexistent integrants, preventing deregulated viral oncogene expression, cell selection, and clonal outgrowth. We therefore hypothesize that loss of E2-expressing episomes plays a key role in the emergence of cervical keratinocytes containing “selectable” HRHPV integrants (i.e., those that retain the upstream regulatory region, E6, and E7 and have disrupted/deleted E2 genes). This model represents a departure from prevailing views of HRHPV-related oncogenesis, because progression is generally assumed to be due simply to cells containing exclusively integrated HRHPV outgrowing cells with only episomal genomes (7).

We have tested our hypothesis by using the unique HPV16-containing cervical keratinocyte cell line W12 in monolayer culture (12); this represents a useful system to investigate the effects of HPV16 infection in basal cervical squamous cells, the key site of deregulation of HRHPV viral oncogenes in cervical neoplasia (10). W12 was derived from a low-grade squamous intraepithelial lesion (LG-SIL), which resulted from “natural” infection *in vivo* with HPV16, the HRHPV type most commonly detected in cervical carcinomas (15). At early passages, W12 retains  $\approx 100$  HPV16 episomes per cell and recapitulates a LG-SIL in organotypic culture (8, 12). We have previously demonstrated that W12 accurately models cervical neoplastic progression during long-term culture, with spontaneous transition from cells containing only episomal HPV16 to a population containing only integrated HPV16 (8, 10). Therefore, W12 represents a valuable system for studying events associated with spontaneous selection of cervical keratinocytes containing only integrated HRHPV.

By growing W12 in long-term culture and undertaking single-cell cloning we demonstrate here that selection of cells containing integrated HPV16 is indeed associated with episome loss from cells also containing the integrant. Moreover, episome loss is itself associated with endogenous activation of antiviral genes. Our data suggest that models of HRHPV-related carcinogenesis must include not only viral integration but also the steps leading to loss of episome-mediated inhibition of selectable integrants.

## Results

**Physical State of HPV16 in Multiple Long-Term Passage Series of W12.** We previously showed that in a long-term passage series of W12, referred to here as W12.Series1, there was loss of episomal

Conflict of interest statement: No conflicts declared.

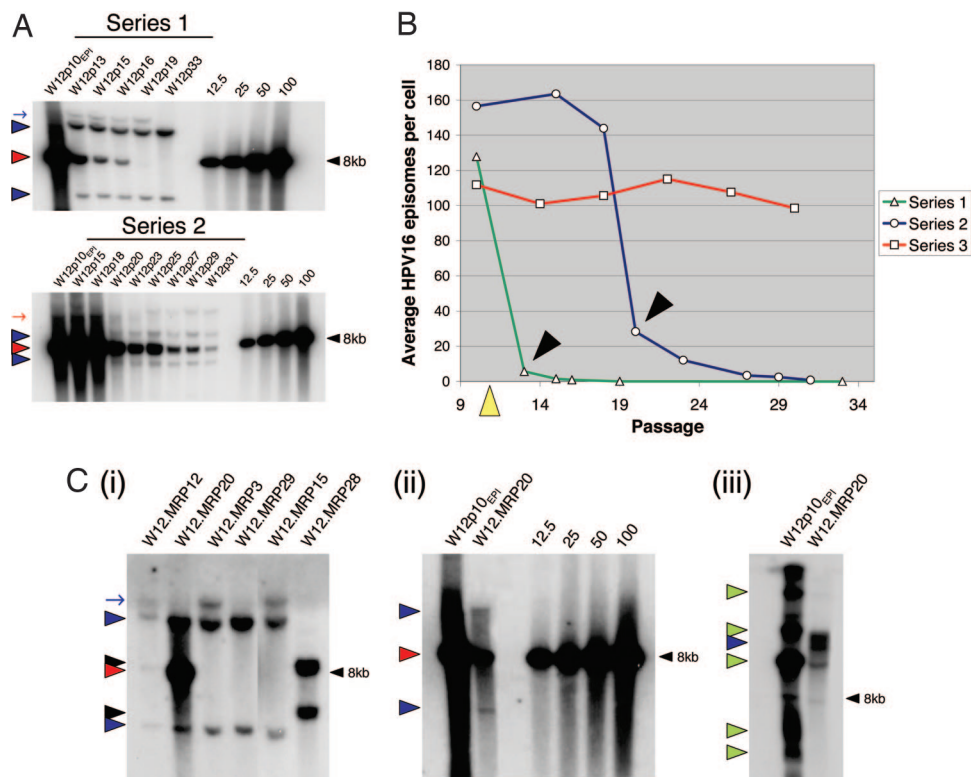
Freely available online through the PNAS open access option.

Abbreviations: HRHPV, high-risk human papillomavirus; pn, passage *n*; GO, Gene Ontology.

Data deposition: Data from this study were deposited in the National Institutes of Health Gene Expression Omnibus database (accession no. GSE4289).

§To whom correspondence should be addressed. E-mail: nc109@cam.ac.uk.

© 2006 by The National Academy of Sciences of the USA



**Fig. 1.** Southern blot analysis of HPV16 physical state in multiple long-term passage series of polyclonal W12. (A) Genomic DNA from independent long-term passage series of W12, each starting from episome-only cells (W12p10<sub>EPI</sub>), were digested with the HPV16 single-cutter BamHI and subjected to Southern blot analysis using <sup>32</sup>P-labeled full-length HPV16 as probe. The images depict autoradiographs of the hybridization pattern from W12 long-term passage series 1 and 2 (for series 3, see Fig. 5), together with copy number controls (right lanes). Red and blue arrowheads indicate bands generated by unit length HPV16 and virus–host junctions, respectively. Red and blue arrows indicate partially digested HPV16 episomes and HPV16-containing fragments, respectively. (B) Line graphs depicting PhosphorImager quantification of the average HPV16 episome copy number per cell in each of the long-term passage series. Black arrowheads indicate the passage at which integrated HPV16 was first detected in W12.Series1 and Series2. Limiting dilution cloning of cells in W12.Series1 was performed at p11 (yellow arrowhead). (C) Southern blot analysis of selected single-cell clones generated from W12.Series1. (i) Selected clones digested with BamHI. (ii and iii) Comparison of W12p10<sub>EPI</sub> and clone W12.MRP20, digested with BamHI (ii) or the HPV16 noncutter HindIII (iii). Copy number controls are also shown in ii. The red and blue arrowheads and blue arrow are as in A. Black arrowheads in i indicate bands generated by nonselected virus–host junctions in W12.MRP28. Green arrowheads in iii indicate various forms of uncut HPV16 episomes.

HPV16 and selection of cells containing only integrated virus (8, 10). We aimed to analyze the period of selective outgrowth more closely and to determine the characteristics of episome loss in additional long-term series (W12.Series2 and W12.Series3) originating from the same starting polyclonal population at passage 10 (W12p10<sub>EPI</sub>, which contains  $\approx$ 100 episomes per cell; refs. 8 and 12).

We performed Southern blot analysis using the HPV16 single-cutter BamHI (Fig. 1A, for W12.Series3, see Fig. 5, which is published as supporting information on the PNAS web site) and quantified episomes based on the intensity of the 7.9-kb episomal band (Fig. 1B). W12.Series1 demonstrated an early rapid reduction in episome numbers. By passage 13 (p13), the average episome content was reduced by more than 95%, and this was followed by a more gradual reduction in residual episomes until none was detectable by p19. The selected HPV16 integrant (indicated by the presence of two additional BamHI bands) emerged when episome numbers had rapidly fallen and was subsequently retained.

A comparable pattern of events was observed in W12.Series2. Rapid episome loss was observed between p18 and p20, where average episome loss content was reduced by  $>$ 80%. A different selected integrant (producing two virus–host junction bands of different sizes to those in Series1) emerged when episome numbers had fallen rapidly. There was a more gradual reduction of residual episomes during subsequent passages. In contrast, in

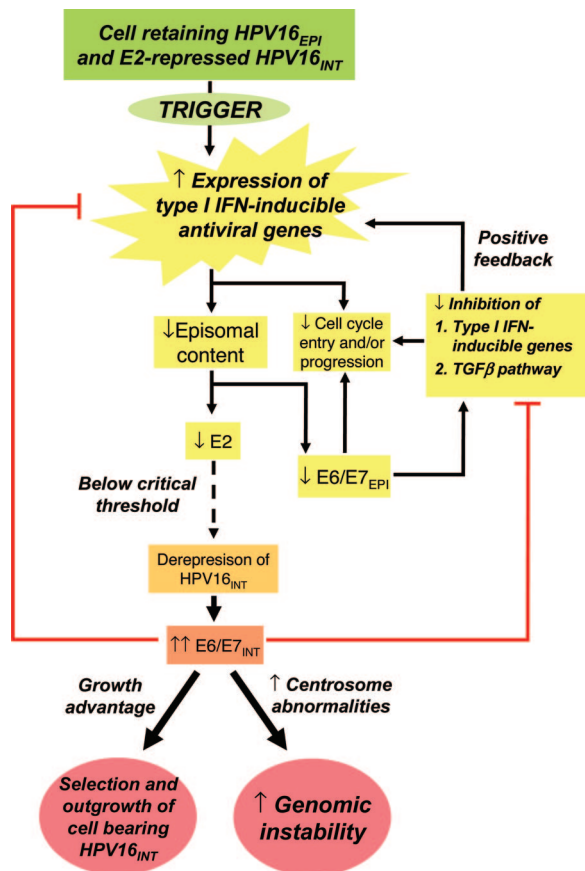
W12.Series3 neither loss of episomes nor selection of integrated HPV16 were observed after 20 passages and  $\approx$ 100 population doublings (Fig. 5). These observations indicate that spontaneous selection of integrated HPV16 is associated reproducibly with rapid loss of episomes, although these events are not inevitable and their timing varies.

**Clonal Analysis of W12 Long-Term Passage Series1 During the Period of Rapid Episome Loss.** Although Southern blot analysis was consistent with our hypothesis, we considered the alternative explanation that the kinetics of episome loss may simply have reflected integrant-only cells outgrowing variably less “fit” episome-only populations. Therefore, we generated single-cell clones by limiting dilution of W12.Series1 at p11; i.e., during the period of rapid episome loss and selection of integrated HPV16 (Fig. 1B). Thirty clones were generated, designated W12.MRP1–30.

We have shown that the upstream virus–host junction of the selected HPV16 integrant in W12.Series1 comprises a disrupted L2 gene and sequences at chromosome 5p15.1 (10). Therefore, we were able to design a PCR protocol for screening the physical state of HPV16 in the single-cell clones (see Fig. 6 and *Supporting Text*, which are published as supporting information on the PNAS web site). Based on the PCR results, Southern blot analysis was performed on six selected clones; W12.MRP12 and -20 (suggested by PCR to contain episomes and the integrant being selected), W12.MRP3 and -29 (containing the integrant







**Fig. 4.** Proposed scheme of events leading to the selection of keratinocytes containing only integrated HPV16 in W12. E6/E7<sub>EPI</sub> and E6/E7<sub>INT</sub> refer to episome- and integrant-derived viral oncogene expression, respectively. HPV16<sub>EPI</sub> and HPV16<sub>INT</sub> refer respectively to episomal and integrated HPV16.

Taken together, our results are consistent with the scheme of events proposed in Fig. 4. Endogenous activation of antiviral genes in a cell containing a “selectable” HPV16 integrant induces a decrease in episome content, with proportionate reduction in episomal expression of E2, E6, and E7. Reduced E6 and E7 will lead to reduced inhibition of type I IFN-inducible genes (19–22) and the TGF- $\beta$  pathway (23–25). These events would account for decreased expression of cell cycle genes (17, 26), and would also lead to rapid further reduction in episome levels in a positive feedback loop. Episome loss will eventually cause E2 expression to reach a critical low level, at which expression from the coexisting integrated HRHPV can no longer be inhibited. Deregulated expression of E6 and E7 from the integrant will confer a selective advantage and increase genomic instability, and at the same time inhibit type I IFN-inducible genes and activation of the TGF- $\beta$  pathway. Thus, the expression levels of these host genes in emergent cells containing integrated HPV16 only will revert to baseline levels.

We have not yet identified the initial trigger for episome loss. Previous studies have suggested that the HRHPV E2 gene is required for a full antiviral response to be elicited by exogenous IFN (17). On the other hand, inhibition of type I IFN-inducible antiviral genes by the viral oncogenes that are also expressed by HRHPV episomes (19–22) suggests that there is a delicate balance between elicitation and inhibition of an antiviral response in cells containing HPV16 episomes. Perturbation of this balance may be the key to induction of the events shown in Fig. 4, and this in turn may be related to deregulation of virus or host gene expression due to genetic and/or epigenetic mechanisms.

Evidence from our parallel study (M.T.H., M.R.P., I. Roberts, W. O. F. Alazawi, A. E. Tescherndorff, X.-Y. Zhang, M.A.S., and N.C., unpublished data) that selectable integrated HPV16 can exist in a minority of cells in a polyclonal population for long periods without exerting a selective growth advantage argues against integration itself serving as the stimulus for episome loss. In an earlier study using W12.Series1 (8), we demonstrated increased expression (relative to an episome only population) of IFN-inducible genes in late passage integrant-only cells that had acquired high-level genomic instability after viral integration. These cells had undergone  $\approx 40$  population doublings after the selected integrant first became detectable, and had acquired many more chromosomal abnormalities than seen in any of the samples in the present study (Fig. 9, which is published as supporting information on the PNAS web site). This finding argues that high-level genomic instability in a HRHPV-infected cell may also activate genes inducible by the type I IFN pathway, despite the deregulated expression of the viral oncogenes.

It will now be important to assess whether the events that we have observed in W12 are also seen during spontaneous selection of other integrated HRHPV types. It may be possible to use other naturally infected keratinocytes, such as the HPV31b-containing cell line CIN612, for which episome-only and integrant-only forms are described (13). Given our evidence that episome loss and selection of integrated HPV16 occurs very rapidly *in vitro*, it will be difficult to detect similar events occurring *in vivo*. Nevertheless, our data are supported by the observation that most cervical carcinomas containing integrated HRHPV have little or no detectable episomal DNA (3, 27–31). Although some carcinomas do contain both episomal and integrated virus, *in situ* analysis has shown that regions containing only integrated HRHPV exist adjacent to regions containing apparently only episomes (32). Thus, in these cases, episome loss in the context of integrant selection is also applicable. On the other hand,  $\approx 12.5\%$  of cervical carcinomas appear to contain transcripts derived only from episomal HRHPV (3), suggesting an alternative pathway of episome-driven carcinogenesis that warrants further investigation.

We conclude from the W12 model system that induction of episome loss, associated with activation of antiviral response genes, is a key event in spontaneous selection of cells containing integrated HPV16. We propose that a revision of the current model of HPV16-induced cervical neoplasia is required. Progression of lesions in which HPV16 integration plays a role requires not only integration *per se*, but also loss of regulatory episomes.

## Materials and Methods

**Cell Culture.** Cell culture was as described (12, 33). Single-cell clones were generated by limiting dilution (34) from W12.Series1 at passage 11, when rapid episome loss was occurring (Fig. 1B). The single colonies generated were expanded in six-well plates for analysis of HPV16 physical state and levels of expression of host and viral genes. At  $\approx 80\%$  confluence, fibroblast feeder cells were removed, followed by extractions of genomic DNA as described (8) and of RNA using TRIzol (Invitrogen).

**Southern Blot Analysis of the Physical State of HPV16.** Five micrograms of genomic DNA was restriction enzyme digested, electrophoresed through a 0.8% agarose gel, with appropriate copy number controls, and transferred to Hybond-N+ nylon membrane (Amersham Pharmacia) (8, 10). Probe was prepared by excision of full-length HPV16 DNA from the pspHPV16 plasmid (12), followed by labeling with [ $\alpha$ - $^{32}$ P]dCTP by random priming. HPV16 was detected and quantified by using a PhosphorImager (Fuji).

### Preparation and Hybridization of Probes for Microarray Analysis.

Total RNA from W12 samples was used to generate biotin-labeled cRNA for microarray analysis. Two technical replicates were performed for each sample to control for variation in labeling and hybridization efficiency. Double-stranded cDNA was synthesized by using SuperScript (Invitrogen), employing the (dT)<sub>24</sub>-T7 promoter primer. Biotin-labeled cRNA was then generated by Bioarray *in vitro* transcription (Enzo), and fragmented by metal-induced hydrolysis. Probe from each replicate was hybridized, washed, stained, and scanned by the Medical Research Council Geneservice (Cambridge, U.K.) using standard Affymetrix procedures. We used GeneChip HG-U133 Plus 2.0 Arrays (Affymetrix).

**Analysis of Microarray Data.** Data were analyzed by using GENE-SPRING (Agilent Technologies). Initially, a “per chip” normalization to the 50th percentile was performed, followed by a “per gene” normalization using the mean expression values for each gene across the replicates of the reference W12p10<sub>EPI</sub>. Statistical analysis was performed in the “log of ratio” mode to give equal weighting to increases and decreases in expression level. Given that geometric mean values were used, this centered the expression values of all genes in the reference sample at or around 1, and gave expression values for the clones that equated to the expression ratio relative to the W12p10<sub>EPI</sub> reference. Stringent criteria were then applied to highlight significant gene expression changes. First, we filtered out genes with expression levels

not altered by at least 1.5 fold in either of the clones relative to the reference. Second, to include OFF → ON and ON → OFF genes, we further considered genes flagged as present or marginal in both replicates of at least one of the three populations analyzed. Because low-range expression values may be unreliable, only genes with at least one raw value of >100 when flagged as present or marginal were analyzed. After filtering, a one-way ANOVA was performed to detect genes with significantly altered expression ( $P < 0.05$ ) in at least one of the three populations analyzed. *K* means cluster analysis was then applied to the set of significantly altered genes. GO biological process analysis was performed by using the GOMINER tool (35). Significant enrichment of specific GO biological processes in each *k* means cluster was determined by using a one-sided Fisher’s exact test (35). *P* values were adjusted by using a false discovery rate correction (36).

**Real-Time RT-PCR.** Quantitative PCR of cDNA was performed to quantify viral gene expression and validate changes in expression of selected host genes. We adapted SYBR green protocols and used four housekeeping genes to normalize expression levels (see *Supporting Text*).

We thank Ian Roberts and Emma Gooding (Medical Research Council Cancer Cell Unit) and Barry Zeeberg (National Cancer Institute, Bethesda) for help and advice. This work was supported by the Medical Research Council and Cancer Research UK.

- Longworth, M. S. & Laimins, L. A. (2004) *Microbiol. Mol. Biol. Rev.* **68**, 362–372.
- Munger, K., Baldwin, A., Edwards, K. M., Hayakawa, H., Nguyen, C. L., Owens, M., Grace, M. & Huh, K. (2004) *J. Virol.* **78**, 11451–11460.
- Klaes, R., Woerner, S. M., Ridder, R., Wentzensen, N., Duerst, M., Schneider, A., Lotz, B., Melsheimer, P. & von Knebel Doeberitz, M. (1999) *Cancer Res.* **59**, 6132–6136.
- Schwarz, E., Freese, U. K., Gissmann, L., Mayer, W., Roggenbuck, B., Stremlau, A. & zur Hausen, H. (1985) *Nature* **314**, 111–114.
- Choo, K. B., Pan, C. C. & Han, S. H. (1987) *Virology* **161**, 259–261.
- Dowhanick, J. J., McBride, A. A. & Howley, P. M. (1995) *J. Virol.* **69**, 7791–7799.
- Jeon, S., Allen-Hoffmann, B. L. & Lambert, P. F. (1995) *J. Virol.* **69**, 2989–2997.
- Alazawi, W., Pett, M., Arch, B., Scott, L., Freeman, T., Stanley, M. A. & Coleman, N. (2002) *Cancer Res.* **62**, 6959–6965.
- Duensing, S. & Munger, K. (2002) *Cancer Res.* **62**, 7075–7082.
- Pett, M. R., Alazawi, W. O., Roberts, I., Downen, S., Smith, D. I., Stanley, M. A. & Coleman, N. (2004) *Cancer Res.* **64**, 1359–1368.
- White, A. E., Livanos, E. M. & Tlsty, T. D. (1994) *Genes Dev.* **8**, 666–677.
- Stanley, M. A., Browne, H. M., Appleby, M. & Minson, A. C. (1989) *Int. J. Cancer* **43**, 672–676.
- Bedell, M. A., Hudson, J. B., Golub, T. R., Turyk, M. E., Hosken, M., Wilbanks, G. D. & Laimins, L. A. (1991) *J. Virol.* **65**, 2254–2260.
- Bechtold, V., Beard, P. & Raj, K. (2003) *J. Virol.* **77**, 2021–2028.
- Bosch, F. X., Manos, M. M., Munoz, N., Sherman, M., Jansen, A. M., Peto, J., Schiffman, M. H., Moreno, V., Kurman, R. & Shah, K. V. (1995) *J. Natl. Cancer Inst.* **87**, 796–802.
- Der, S. D., Zhou, A., Williams, B. R. & Silverman, R. H. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 15623–15628.
- Chang, Y. E., Pena, L., Sen, G. C., Park, J. K. & Laimins, L. A. (2002) *J. Virol.* **76**, 8864–8874.
- Turek, L. P., Byrne, J. C., Lowy, D. R., Dvoretzky, I., Friedman, R. M. & Howley, P. M. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 7914–7918.
- Chang, Y. E. & Laimins, L. A. (2000) *J. Virol.* **74**, 4174–4182.
- Nees, M., Geoghegan, J. M., Hyman, T., Frank, S., Miller, L. & Woodworth, C. D. (2001) *J. Virol.* **75**, 4283–4296.
- Ronco, L. V., Karpova, A. Y., Vidal, M. & Howley, P. M. (1998) *Genes Dev.* **12**, 2061–2072.
- Barnard, P. & McMillan, N. A. (1999) *Virology* **259**, 305–313.
- Nees, M., Geoghegan, J. M., Munson, P., Prabhu, V., Liu, Y., Androphy, E. & Woodworth, C. D. (2000) *Cancer Res.* **60**, 4289–4298.
- Lee, D. K., Kim, B. C., Kim, I. Y., Cho, E. A., Satterwhite, D. J. & Kim, S. J. (2002) *J. Biol. Chem.* **277**, 38557–38564.
- Favre-Bonvin, A., Reynaud, C., Kretz-Remy, C. & Jalinot, P. (2005) *J. Virol.* **79**, 4229–4237.
- Herrera, R. E. (1998) *Prog. Mol. Subcell. Biol.* **20**, 11–24.
- Cullen, A. P., Reid, R., Campion, M. & Lorincz, A. T. (1991) *J. Virol.* **65**, 606–612.
- Das, B. C., Sharma, J. K., Gopalakrishna, V. & Luthra, U. K. (1992) *J. Gen. Virol.* **73**, 2327–2336.
- Durst, M., Kleinheinz, A., Hotz, M. & Gissman, L. (1985) *J. Gen. Virol.* **66**, 1515–1522.
- Kalantari, M., Blennow, E., Hagmar, B. & Johansson, B. (2001) *Diagn. Mol. Pathol.* **10**, 46–54.
- Tonon, S. A., Picconi, M. A., Bos, P. D., Zinovich, J. B., Galuppo, J., Alonio, L. V. & Teyssie, A. R. (2001) *J. Clin. Virol.* **21**, 129–134.
- Hopman, A. H., Smedts, F., Dignef, W., Ummelen, M., Sonke, G., Mravunac, M., Vooijs, G. P., Speel, E. J. & Ramaekers, F. C. (2004) *J. Pathol.* **202**, 23–33.
- Coleman, N., Greenfield, I. M., Hare, J., Kruger-Gray, H., Chain, B. M. & Stanley, M. A. (1993) *Am. J. Pathol.* **143**, 355–367.
- Freshney, R. I. (1994) in *Culture of Animal Cells: A Manual of Basic Technique* (Wiley-Liss, New York), pp. 161–178.
- Zeeberg, B. R., Feng, W., Wang, G., Wang, M. D., Fojo, A. T., Sunshine, M., Narasimhan, S., Kane, D. W., Reinhold, W. C., Lababidi, S., *et al.* (2003) *Genome Biol.* **4**, R28.
- Benjamini, Y., Drai, D., Elmer, G., Kafkafi, N. & Golani, I. (2001) *Behav. Brain Res.* **125**, 279–284.