The integration of HPV-18 DNA in cervical carcinoma

S A Corden, L J Sant-Cassia, A J Easton, A G Morris

Abstract

Aims—Little information is available on the patterns of integration into the host chromosomal DNA of cervical carcinomas of human papillomavirus type 18 (HPV-18) DNA, which is associated with up to 20% of these carcinomas. Because integration of the viral genome may be extremely important in the pathogenesis of cervical carcinoma, the aim of this study was to investigate which regions of HPV-18 DNA are integrated into the cellular DNA of cervical carcinomas.

Methods—Southern analysis using four subgenomic probes covering the entire HPV-18 genome was used to map viral DNA integrated within cellular DNA. The polymerase chain reaction (PCR) was used to confirm the presence of specific regions of the viral genome.

Results—In all 11 carcinomas there was a single major HPV-18 DNA integrant, retaining ~ 4000 bp of HPV-18 DNA, indicating that approximately half of the virus genome had been lost upon integration. Southern analysis suggested strongly that the viral breakpoint was within the E1/E2 gene boundary, with concomitant loss of part or all of the E2 ORF (open reading frame), all of the E4, E5, and L2 ORFs and part of the L1 ORF. These data were supported by the PCR results, which confirmed that the region of integrated HPV-18 DNA from nucleotides 6558 to 162 was present in all the carcinoma samples studied. Assuming that no genomic rearrangements, deletions, or insertions had occurred, 4131 bp of integrated HPV-18 DNA could be accounted for in eight cervical carcinoma samples. The results of Southern analysis also suggested that integration of HPV-18 DNA may have occurred at a specific host chromosomal site.

Conclusions—Broadly, the viral sequences retained upon HPV-18 integration resemble those found when HPV-16 is integrated. However, it appears that the HPV-18 E2 region is more consistently deleted.

(*J Clin Pathol: Mol Pathol* 1999;**52**:275–282)

Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, UK
L J Sant-Cassia
A J Easton
A G Morris

PHLS Wales, University Hospital of Wales, Heath Park, Cardiff CF4 4XW, UK
S A Corden

Correspondence to:
Dr Morris.
email:
am@dna.bio.warwick.ac.uk

Cervical cancer is the second most common cancer in women worldwide.[1][2] Human papillomavirus (HPV) DNA is present in ~ 90% of cervical carcinomas, with HPV-16 and HPV-18 being the two most prevalent types.[3–6]

Although HPV-16 and HPV-18 are regarded as essentially the same, it should be noted that these two viruses only share ~ 50% homology at the nucleotide level. This is reflected in the different array of cellular transcriptional regulators that bind to the HPV long control region (LCR).[7] These differences might have a bearing on the pathogenesis of the two viruses.

It is generally accepted that HPV is the primary causal agent of cervical carcinoma and the World Health Organisation (WHO) has declared that HPV-16 and HPV-18 are, indeed, carcinogens.[8] Therefore, the role of HPV-16 and HPV-18 in the pathogenesis of cervical carcinoma is of great importance. It is well established that viral proteins derived from the E6 and E7 genes are essential in the development of cervical carcinoma. However, the importance and necessity of viral DNA integration in the pathogenesis of cervical carcinoma is not well established.

After the identification of HPV DNA in cervical carcinomas there was great interest in determining its physical state in the progenitor cervical lesions and carcinomas. A variety of techniques, including those based on the polymerase chain reaction (PCR), non-isotopic in situ hybridisation (NISH), and Southern analysis, have shown that in HPV DNA positive carcinomas the viral DNA is often integrated into the host genome.[9–17] Therefore, it has been suggested that the integration of HPV DNA into host chromosomes plays an important role in the progression to malignancy. However, there have been several reports describing HPV-16 positive cervical carcinomas that have clearly shown the presence of either integrated HPV DNA or episomal HPV DNA, and also carcinomas with both episomal and integrated HPV DNA.[18–23] Up to 30% of cervical carcinomas contain episomal forms only of HPV-16.[22][24] This led Das *et al* to suggest that HPV DNA integration may not be necessary for malignant progression.[24] However, one very important finding is that 100% of HPV-18 cervical carcinomas contain integrated forms only of HPV-18 DNA.[21][22][24][25] Although the importance of this is not yet clear, it may be that the pathogenesis of HPV-16 and HPV-18 are different.

Although integration studies on HPV have been numerous, most of the data collected to date have come from studies of HPV-16, probably because of its greater prevalence over HPV-18. These studies have revealed no precise pattern regarding the regions of virus DNA that are lost upon HPV DNA integration into the host genomic DNA. However, in general, integration involves E1/E2 gene dis-

Sample	Carcinoma type	Differentiation	Pelvic nodal status	Age (years)	HPV type
1	Squamous	Poor	Involved	37	18
2	Squamous	Poor	Involved	41	18
3	Squamous	Moderate	No involvement	69	18
4	Squamous	NA	NA	69	18
5	Squamous	NA	No involvement	45	16 and 18
6	Squamous	Moderate/Poor	Involved	41	16 and 18
7	Adenocarcinoma	NA	No involvement	41	18
8	Squamous	Moderate	No involvement	46	18
9	Squamous	Moderate/Well	Involved	58	18
10	Adenocarcinoma	Moderate/Well	Involved	51	18
11	Squamous	Moderate/Poor	Involved	36	18
12	Squamous	Moderate/Poor	Involved	35	18
13	Squamous	Moderate/Well	Involved	30	16 and 18

Most patients were found to have a squamous carcinoma of the cervix. The exceptions were patient samples 7 and 10 who were found to have an adenocarcinoma of the cervix.
NA, not available.

ruption, together with retention of the LCR, E6, and E7 genes.[11–13][26–29] The disruption or loss of the E2 gene is believed to have a pivotal role in relieving the repression of HPV E6 and E7 gene expression because of the involvement of E2 in the regulation of viral transcription.[30–34] Integration is also believed to stabilise E6/E7 mRNA transcripts by use of a cellular polyadenylation signal.[10] However, because of the paucity of information regarding HPV-18 integration, we cannot conclude that similar mechanisms function with this virus, which as indicated differs considerably from HPV-16.

Therefore, we have analysed the physical state of HPV-18 DNA in cervical carcinomas. Using the combined techniques of Southern hybridisation and PCR analysis we have demonstrated that HPV-18 DNA in cervical carcinomas shows a specific integration pattern with respect to the viral DNA, with a preference to retain a specific region of viral DNA. We find that, in general, the regions of HPV-18 that are integrated are similar to HPV-16 with respect to the LCR, E6, and E7 genes only. The total region of integrated viral DNA appears to be much more conserved for HPV-18 than for HPV-16. This is particularly noticeable within the L1 and E1 regions.
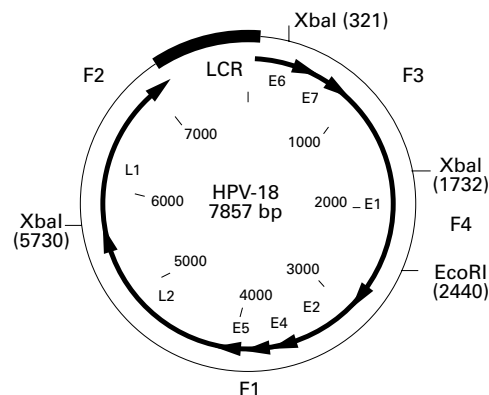

Figure 1 A schematic representation of the subgenomic fragments of HPV-18 created by restriction endonuclease digestion using XbaI and EcoRI. The nucleotide positions of the endonuclease restriction sites within the HPV-18 genome are shown. The regions of the HPV-18 genome covered by the subgenomic probes are also shown. LCR, long control region.

Materials and methods

SAMPLE COLLECTION AND DNA EXTRACTION

All samples studied were resected cervical carcinomas at varying histological stages (table 1), and were collected at the time of surgery. Fresh samples were minced, and digested overnight with DNase I (Sigma-Aldrich, Poole, Dorset, UK), hyaluronidase (Boehringer-Mannheim, Mannheim, Germany), and collagenase (Sigma-Aldrich). The cell suspension was filtered to remove any large debris, centrifuged at $258\times g$ and lysed using Hirt buffer (10 mM Tris/HCl, pH 7.4, 20 mM EDTA, 0.5% (wt/vol) sodium dodecyl sulphate (SDS)). Proteinase K (Life Technologies, Paisley, UK) was added to a final concentration of 100 μg/ml and the cell lysate left overnight at 37°C. DNA was phenol/chloroform extracted, precipitated, and resuspended in TE (10 mM Tris/HCl, pH 7.4, 1 mM EDTA). Amplification of HPV-18 DNA using HPV type specific primers HPV-18 P1 (tacccggat[7459]CCGATTTCGGTTGCC[7473]) and P2 (tacccgggta[162]CCGTGCACAGAT CAG[148]) was used to show that HPV-18 DNA was present in all carcinomas studied (data not shown).

PREPARATION OF SUBGENOMIC HPV-18 DNA PROBES FOR SOUTHERN ANALYSIS

Full length HPV-18 in the plasmid vector pBR322 (clone 18) was a kind gift from Dr E-M DeVilliers, Deutschkrebsforschungszentrum, Heidelburg, Germany. Full length HPV-18 DNA was excised from pBR322 using the EcoRI restriction endonuclease and purified from agarose gels using the Qiaquick™ (Qiagen, Crawley, Sussex, UK) agarose gel extraction kit, according to the manufacturer's instructions. The HPV-18 DNA was then digested with XbaI to provide four subgenomic fragments (F1–F4, fig 1). These sub-genomic fragments were ligated into the plasmid vector pBluescript SK+ (Stratagene, Amsterdam, The Netherlands) using restriction endonucleases XbaI and EcoRI, and plasmid stocks prepared. Before use as DNA probes, the cloned HPV-18 DNA subgenomic fragments were excised from pBluescript SK+, using the appropriate restriction endonucleases, and purified from agarose gels.

SOUTHERN ANALYSIS

DNA probes were prepared with the random hexamer oligonucleotide labelling kit (Pharmacia Biotech, Little Chalfont, Buckinghamshire, UK) according to the manufacturer's instructions, using α^{32}P labelled dCTP (3000 Ci/ mmol; Amersham International, Little Chalfont, Buckinghamshire, UK). Genomic DNA (5–10 μg) extracted from the tissues was digested with one of the restriction enzymes HindIII, PstI, PvuII, or Sau3AI (New England Biolabs, Hitchin, Hertfordshire, UK), and the resultant fragments separated by electrophoresis on a 0.8% agarose gel. After Southern transfer on to nylon membranes (Hybond-N™; Amersham) the DNA was UV crosslinked to the membrane.

The membranes were prehybridised with Quickhyb™ (Stratagene) containing dena-

5418		7133		87	578	904			2884

Figure 2 showing schematic representation with:

| L1 | LCR | E6 | E7 | E1 |

MY11
(6558)

LPCR3
(7473)

N0253
(151)

E1L1
(2041)

E1L2
(2485)

E1R1
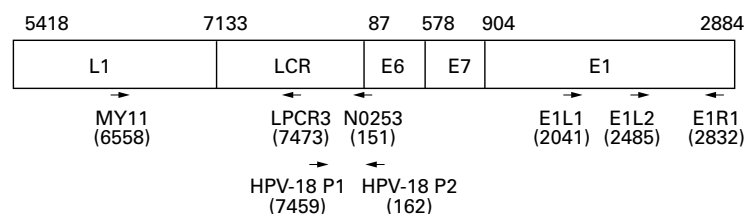(2832)

HPV-18 P1
(7459)

HPV-18 P2
(162)

Figure 2   A schematic representation of the position of the oligonucleotides used in the amplification of the regions of integrated HPV-18 DNA in cervical carcinomas. The relevant HPV-18 open reading frames (ORFs), 5' nucleotide position of each oligonucleotide and size (bp) of the HPV-18 DNA amplicons are shown. LCR, long control region.
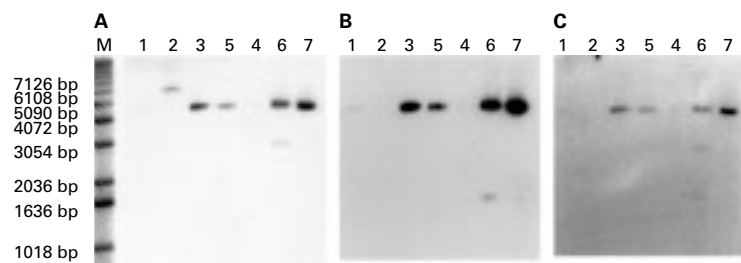

| A | M | 1 | 2 | 3 | 5 | 4 | 6 | 7 | B | 1 | 2 | 3 | 5 | 4 | 6 | 7 | C | 1 | 2 | 3 | 5 | 4 | 6 | 7 |

7126 bp
6108 bp
5090 bp
4072 bp
3054 bp
2036 bp
1636 bp
1018 bp

Figure 3   (A–C) HindIII restriction enzyme digestion of clinical samples 1–7. HindIII generated HPV-18 DNA fragments were detected using ^{32}P-labelled subgenomic fragments (F2–F4) and sized using a ^{35}S labelled 1 kb DNA ladder (M). The numbers at the top of the lanes refer to the patient sample numbers. Figures A–C have been probed with subgenomic HPV-18 probes F2, F3, and F4, respectively.

tured, sheared salmon sperm DNA at a final concentration of 100 µg/ml. The denatured ^{32}P labelled probe was added and left to hybridise at 68°C for two hours. The membranes were washed with 2× saline sodium citrate (SSC) containing 0.1% SDS at room temperature and 0.1× SSC containing 0.1% SDS at 60°C. The membranes were then autoradiographed for 12–48 hours.

SEMI-NESTED PCR TO DETECT L1/LCR AND E1/E2 REGIONS

The oligonucleotide primers (fig 2) used to detect the L1/LCR region were MY11,[35] N0253 (^{151}TCAGGTAGCTTGTAGGGTCG CCGT^{128}), and LPCR3 (^{7473}GGCAACC GAAATCGGTTGCAC^{7453}). The PCRs were divided into first round (MY11 and N0253) to amplify a 1450 bp DNA fragment and second round (MY11 and LPCR3) to amplify a 915 bp DNA fragment. The first round PCR mixture consisted of 100–250 ng of target DNA; 0.4 µM of each primer; 10× PCR reaction buffer, supplied and used according to the manufacturer's instructions (Life Technologies); 250 µM of each of dGTP, dATP, dCTP, and dTTP (Promega, Southampton, UK); and 2.5 U of Taq DNA polymerase (Life Technologies) in a final reaction volume of 50 µl. The first round cycles were 95°C for one minute, followed by 35 cycles of 95°C for 30 seconds, 56°C for 20 seconds, 74°C for 1.5 minutes, and 74°C for five minutes. The second round PCR mixture was essentially the same as the first round, with the exceptions that only 0.2 µM of each primer was used and 1 µl of first round PCR mixture was used as the target DNA. The PCR cycles for the second round were 95°C for one minute, followed by 35 cycles of 95°C for 30 seconds, 56°C for 20

seconds, 74°C for one minute, and a final cycle of 74°C for five minutes.

The oligonucleotide primers that were used to detect the E1/E2 region (fig 2) were E1-R1 (^{2832}CCTTCGGTGTCTGCATCTTC^{2813}), E1-L1 (^{2041}GACAGATGAAAGCGATAT GGC^{2061}), and E1-L2 (^{2485}GGTGGCCATGT-TAGATGATGC^{2505}). First round PCR (E1-R1 and E1-L1) amplified a 791 bp DNA fragment and the second round PCR (E1-R1 and E1-L2) amplified a 347 bp DNA fragment. The first round PCR mixture consisted of 100–250 ng of target DNA; 0.4 µM of each primer; 10× Thermopol reaction buffer, supplied and used according to the manufacturer's instructions (New England Biolabs); 250 µM of each of dGTP, dATP, dCTP, and dTTP (Promega); 1.0 U of Vent® DNA polymerase (New England Biolabs); and sterile distilled water to a final reaction volume of 50 µl. The first round cycles were 95°C for one minute, followed by 35 cycles of 95°C for 30 seconds, 60°C for 20 seconds, 74°C for one minute, and a final cycle of 74°C for five minutes. The second round PCR mixture was essentially the same as the first round, with the exceptions that only 0.3 µM of each primer was used, 1 µl of the first round PCR mixture was used as the target DNA, and only 0.5 U of Vent DNA polymerase was used. The PCR cycles for the second round were 95°C for one minute, followed by 35 cycles of 95°C for 30 seconds, 60°C for 20 seconds, 74°C 1.5 minutes, and a final cycle of 74°C for five minutes.

Results

HindIII DIGESTION OF GENOMIC DNA FROM CLINICAL SAMPLES AND SOUTHERN ANALYSIS SHOW A SINGLE MAJOR INTEGRANT

The four subgenomic probes used in the Southern hybridisation of cervical carcinoma samples were shown to behave in a predictable manner, hybridising only to DNA fragments of HPV-18 generated by digestion of the cloned full length HPV-18 DNA with HindIII, PstI, PvuII, and Sau3AI (data not shown). Figure 3 shows the results of the Southern blots of HindIII restriction enzyme digests of DNA from seven of 13 cervical carcinomas probed with subgenomic fragments. Not all subgenomic fragments were used because the physical removal of bound probe eventually made the nylon membranes unusable.

In fig 3, samples 3, 6, 5, and 7 show the presence of only one major DNA fragment of ∼ 4700–4900 bp when probed with fragments F2, F3, and F4. Because HindIII has no restriction recognition site within the published HPV-18 DNA sequence, the results indicate that in the samples there is a single major HPV-18 DNA integrant. This HPV-18 integrant cannot be full length because the observed DNA fragment, which will contain some cellular sequences, is smaller than the entire length of the HPV-18 genome. Sample 4, although shown to be HPV-18 DNA positive by PCR (data not shown), did not hybridise with any of the subgenomic fragments used. This was also seen for sample 12 (data not shown). This may be because the
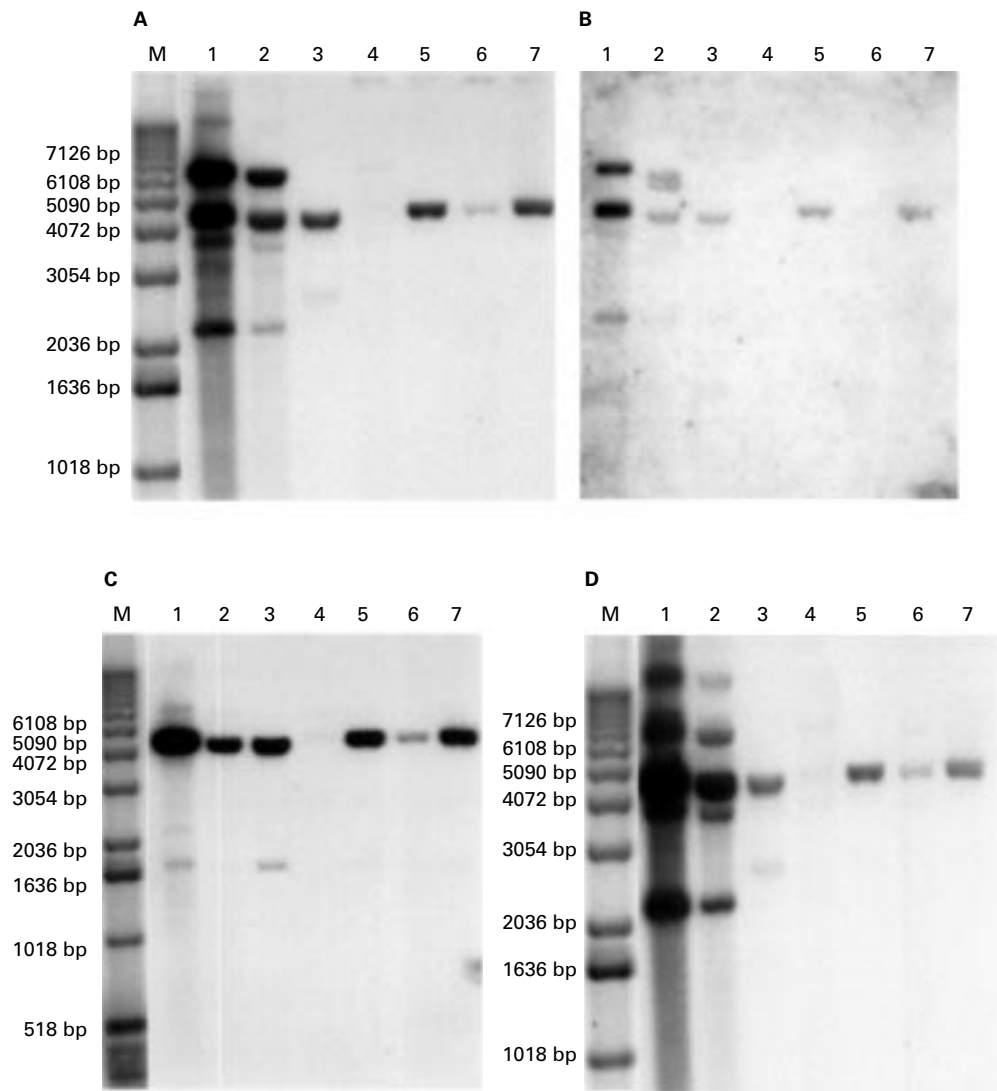
A

B

C

D

Figure 4   (A–D) Autoradiographs of Southern blots showing the physical state of HPV-18 DNA in clinical samples 1–7. PstI generated HPV-18 DNA fragments were detected using ^{32}P labelled subgenomic fragments (F1–F4) and sized using a ^{35}S labelled 1 kb DNA ladder (M). The numbers at the top of the lanes refer to the patient sample numbers. Figures A–D have been probed with the subgenomic HPV-18 probes F1, F2, F3, and F4, respectively.

amount of HPV-18 DNA present was below the level of detection. The integrated HPV-18 DNA fragments detected in 11 of the 13 cervical carcinomas studied were approximately the same size. This suggests that HPV-18 DNA integration into the host chromosomal DNA occurs within a similar region of both the virus genome and the host chromosomal DNA.

PstI AND PvuII DIGESTS OF GENOMIC DNA FROM CLINICAL SAMPLES AND SOUTHERN ANALYSIS INDICATE THE LOSS OF HALF THE HPV-18 GENOME UPON INTEGRATION

Restriction enzyme digests of integrated HPV-18 DNA using enzymes cutting within the HPV genome should generate characteristic HPV-18 subgenomic fragments, together with cellular junction fragments. If regions of the HPV-18 genome have been deleted upon integration, some of the expected subgenomic fragments will be lost and new ones created. To identify the sites of integration in the virus DNA, the genomic DNA was digested indi-

vidually with several restriction enzymes known to digest within HPV-18 DNA. Figures 4 and 5 show the results of the PstI and PvuII restriction enzyme digests, respectively, and table 2 gives the sizes of the predicted fragments determined from the published sequence of HPV-18 DNA.

Figure 4 shows the presence of a common PstI generated DNA fragment of ~ 4800–4600 bp. This is smaller than the 6337 bp DNA fragment expected if this region of the genome were intact. The result indicates that part of this DNA fragment has been lost upon viral DNA integration into the host chromosomal DNA. As seen in fig 4, all four subgenomic probes hybridised to the predominant 4800–4600 bp DNA fragment, indicating that the junctions of the subgenomic fragments at nucleotide positions 321 (F2/F3), 1732 (F3/F4), and 2440 (F4/F1) must be present (fig 1). This represents a minimum of 2119 bp of integrated HPV-18 DNA and includes all of the E7 open reading frame (ORF) and at least large parts of the E6 and E1 ORFs. Fragments of
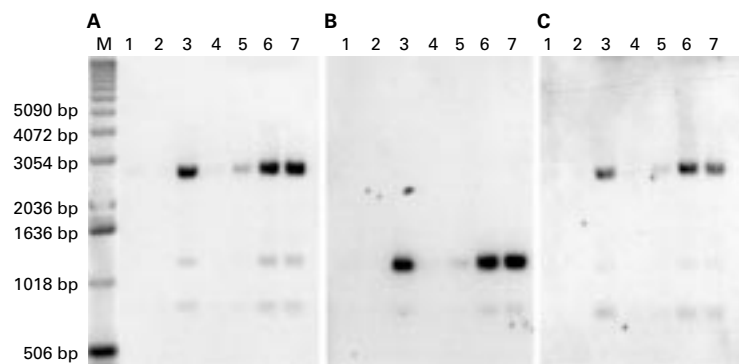
Figure 5   (A–C) Autoradiographs of Southern blots showing the physical state of HPV-18 DNA in clinical samples 1–7. PvuII generated HPV-18 DNA fragments were detected using ^{32}P labelled subgenomic fragments (F1–F4) and sized using a ^{35}S labelled 1 kb DNA ladder (M). The numbers at the top of the lanes refer to the patient sample numbers. Figures A–C have been probed with subgenomic probes F1, F3, and F4 respectively.

Table 2   Expected HPV-18 DNA fragments from PstI and PvuII enzyme cleavage digests

Enzyme	Nucleotide position	Fragment size (bp)
PstI	6766–5246	6337
	5246–5313	67
	5313–5770	457
	5770–6325	555
	6325–6766	441
PvuII	851–2100	1249
	2100–6473	4373
	6473–851	2235

441 bp, 457 bp, and 555 bp generated by cleavage at PstI recognition sites at nucleotide positions 5313, 5770, 6325, and 6766 were not detected, and were therefore probably deleted (data not shown). Therefore, the hybridisation seen to the major fragment with subgenomic probe F2 in fig 4B must be caused by the LCR component of the PstI generated DNA fragment. The region within the HPV-18 genome that appears to be deleted upon integration corresponds with that encoding the E1/E2 and L2/L1 proteins (fig 1). Similar hybridisation patterns were seen with 11 of the 13 cervical carcinomas studied.

In fig 5A and C, the clinical samples were digested with PvuII and they showed the presence of a major DNA fragment of ~2800–2500 bp, which hybridised strongly with subgenomic probes F1 and F4, respectively. The presence of this DNA fragment suggests that all, or part, of the 4373 bp restriction fragment generated by digestion at PvuII recognition sites at nucleotide positions 2100–6473 (table 2) in the HPV-18 genome is lost upon the integration of HPV-18 DNA into the host chromosomal DNA. The data confirm the presence of the PvuII restriction sites at nucleotide positions 851 and 2100, which are contained within the large PstI generated DNA fragment (table 2), because the resulting 1249 bp DNA fragment hybridises strongly with subgenomic probe F3 (fig 5B). In addition, the expected 2235 bp PvuII generated fragment was not seen, suggesting that the PvuII recognition site at nucleotide position 6473 was not present. This would place one region of the viral integration site in the L1 region of the virus genome, which suggests that the L2 and probably E5 ORFs are lost during or after the inte-
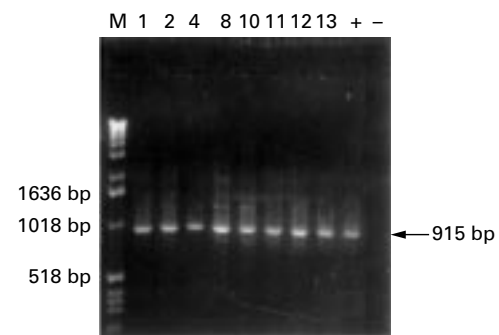

Figure 6   HPV-18 DNA was amplified from cervical carcinoma samples using semi-nested PCR. The second round of the semi-nested PCR used the MY11 consensus oligonucleotide and a second HPV-18 type specific oligonucleotide, LPCR3, to generated a 915 bp fragment. Amplified DNA fragments were electrophoresed on a 1.5% (wt/vol) agarose gel. The numbers at the top of each lane refer to the patient sample numbers. M, 1 kb DNA ladder; +, HeLa DNA positive control; −, negative control.

gration of HPV-18 DNA into the host chromosomal DNA.

A further restriction enzyme digestion using Sau3AI (data not shown) confirmed that the maximum extent of the HPV genome (reading clockwise) that is retained is up to nucleotide positon 2900, which is in close proximity to the E1/E2 boundary, and results in loss of the downstream genes E2, E4, E5, and possibly L2.

L1/LCR AND E1/E2 SEMI-NESTED PCR USED TO DETERMINE THE BOUNDARIES OF INTEGRATED HPV-18 DNA IN CLINICAL SAMPLES

The results of Southern analysis suggested that the HPV-18 DNA integration point in the clinical samples analysed in our study lies, on the one side, close to nucleotide 2900, within the E1/E2 ORFs and, on the other side, within the L1 region around nucleotide 6325. To analyse this further, PCR was used to determine the most terminal regions of the integrated HPV-18 genome, basing the primer locations on the Southern blot data.

In the 13 original carcinomas studied (table 1), HPV-18 DNA was shown by the presence of a 560 bp PCR product amplified from nucleotides 7459 to 162. The oligonucleotides used (HPV-18 P1/P2) were designed with mismatches at their 3' ends when aligned with the HPV-16 sequence. Preliminary data showed that the oligonucleotides used would only amplify either HPV-16 or HPV-18 DNA. Similarly, the HPV-18 oligonucleotides also carried mismatches at their 3' ends when aligned with the HPV-45 sequence. Although no confirmatory PCR showed that HPV-45 was not amplified, a similar result was expected to that of HPV-16. These oligonucleotides were adequately able to amplify HPV-18 DNA from the carcinomas studied without the need for semi-nested PCR. Semi-nested PCR was only required when amplifying the much larger L1/LCR fragment of 1450 bp.

The oligonucleotide set chosen to determine the left boundary included one of the L1 consensus primers, MY11,[35] in conjunction with oligonucleotides N0253 and LPCR3. Figure 2 shows the relative locations and orientations of
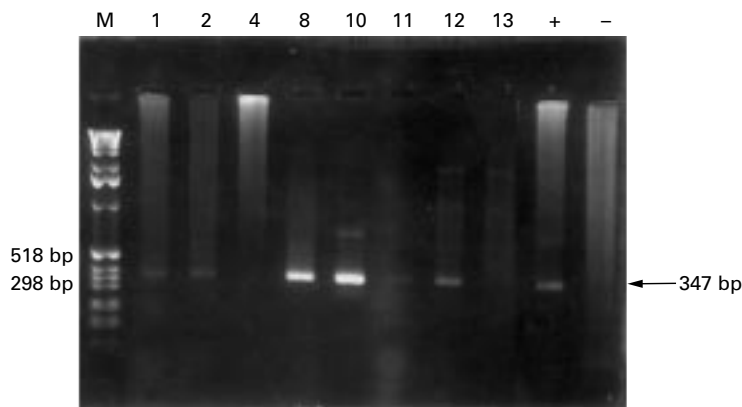
Figure 7   HPV-18 DNA was amplified from cervical carcinoma samples using HPV-18 type specific oligonucleotides E1L2 and E1R1 in a semi-nested PCR. Amplified DNA fragments of 347 bp were electrophoresed on a 1.5% (wt/vol) agarose gel. The numbers at the top of each lane refer to the patient sample numbers. M, 1 kb DNA ladder; + HeLa DNA positive control; −, negative control.

the oligonucleotides. These oligonucleotides are leftwards of the original HPV-18 type specific primers used to show the presence of HPV-18 DNA in the carcinoma samples. Although the MY11 oligonucleotide has been used to amplify HPV DNA from several HPV types, when used with an HPV-18 type specific primer, only HPV-18 DNA is amplified. MY11 was also chosen because it would allow DNA amplification across the PstI site at nucleotide position 6766, a region that was hypothesised to be present in the integrated HPV-18 DNA of the clinical samples analysed by Southern analysis. The oligonucleotide set chosen to determine the right boundary were from within the E1 region, which was also believed to be present in the integrated HPV-18 DNA of the clinical samples studied. Figures 6 and 7 show the results of the semi-nested PCRs.

The results of the PCR analysis of the left boundary (fig 6) showed that in all eight of the selected clinical samples the predicted amplified DNA product of 915 bp was present; therefore, integrated HPV-18 DNA was present between nucleotides 6558 and 7473, the region bounded by oligonucleotides MY11 and LPCR3. Single rounds of PCR with both sets of primers was not sufficient to visualise any amplified DNA product on ethidium bromide stained agarose gels (data not shown). The detection of the second round 915 bp amplified HPV-18 DNA fragment in all clinical samples merely indicates that semi-nested PCR is more sensitive.[36] Furthermore, because the second round of the semi-nested PCR relies on the success of the first round, the region of HPV-18 DNA from nucleotides 151 to 6558, amplified in the first round of the PCR using oligonucleotides MY11 and N0253, must be present. The presence of contaminating DNA can be excluded because the negative control was shown to be free of product in both the first and second rounds of the semi-nested PCR. However, a further, and more important, conclusion that can be drawn from these data is that because the oligonucleotides N0253 and LPCR3 used in the semi-nested PCR overlap with the initial type specific oligonucleotides HPV-18 P1/P2 at nucleotide positions 7459

and 162, respectively (fig 2), the region from nucleotide positions 6558 to 162 must be present.

The results of the PCR analysis of the right boundary (fig 7) showed that the expected amplified DNA product of 347 bp was not seen in two of the eight selected clinical samples. Because single rounds of PCR were insufficient to visualise any amplified products, the semi-nested PCR must have involved amplification of the region from nucleotide positions 2041 to 2832 of the integrated HPV-18 DNA in clinical samples 1, 2, 8, and 10–12. The absence of amplified HPV-18 DNA fragments in the remaining two clinical samples, samples 4 and 13, could have been the result of either the target DNA being below the level of detection or the fact that the target DNA was absent in these samples. If we assume the latter possibility to be correct, it would suggest that in clinical samples 4 and 13 the viral integration point is around nucleotide position 2832 (fig 2).

Discussion
The integration of HPV DNA into the host genome is incidental to the virus life cycle, and does not confer an advantage to the virus, because viral integration often results in deletion of part of the virus genome and the generation of infectious progeny ceases. However, integration is probably an important step in cervical carcinogenesis because it is so often associated with the progression of cervical carcinoma. Hence, it follows that the nature of the integration, that is, the regions of the genome that are preserved or lost, may be crucial to the carcinogenic process.

The general conclusion of substantial integration studies on HPV-16, both in cell lines and, more importantly, in carcinoma material, is that the LCR and the viral oncogenes E6 and E7 are always present.[11 13 26–28 37] However, as discussed earlier, several studies have shown variability in the loss of part or all of the E2 gene. Indeed, Berumen *et al* show that the E1/E2 region is retained in HPV-16 positive cervical carcinomas.[25] This latter observation is important because the E2 protein is a transcriptional repressor of E6 and E7 gene expression and so the loss of E2 is thought to result in overexpression of the E6 and E7 genes, thereby aiding in the progression of cervical carcinomas.[31 32 38] However, in view of the numerous HPV-16 positive cervial carcinomas that contain only episomal forms of the virus, this general view requires modification.

Studies of HPV-18 DNA integration have been less frequent. Kitagawa *et al* showed that out of nine HPV-18 positive cervical carcinomas, only three showed deletions in either the E1 gene, E2 gene, or both, and one carcinoma showed a deletion in the E5/L2 genes only.[29] In the remaining five carcinomas the HPV-DNA appeared to be full length. This is in contrast to the work of Berumen *et al*, who reported the absence of the E1/E2 region in 20 HPV-18 positive cervical carcinomas, as revealed by PCR analysis.[25] However, this latter study did not demonstrate which specific HPV-18 se-
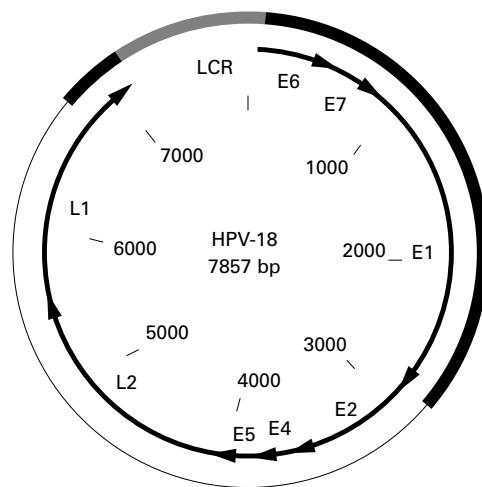
Figure 8   Diagrammatic representation of the region of HPV-18 DNA retained upon viral DNA integration into the host chromosomal DNA. The HPV-18 genes and long control region (LCR) are shown. The shaded region of the HPV-18 genome is that believed to be present after viral DNA integration.

quences were present in the carcinomas, and neither did it establish the extent of the deletions. In our studies, we sought to extend the knowledge of the integration patterns of HPV-18 to determine whether it resembles those of HPV-16 in the constant retention of the LCR, E6, and E7 genes, and the preferential loss of E2. We found that, in the 11 cervical carcinoma samples for which we obtained Southern analysis data, E6 and E7 were retained and E2 was lost.

The restriction digestion fragment generated with HindIII, which does not cut within the HPV-18 genome, was in every case in the region of 4800 bp; because this fragment must contain some cellular material, this indicates that a considerable part of the viral genome is deleted. The use in Southern analysis of HPV-18 subgenomic probes from distinct regions of the virus genome suggested that HPV-18 DNA integration had occurred close to the E1/E2 boundary, with the concomitant loss of part or all of the E2, all of the E4, E5, and L2, and part of the L1 genes. This conclusion was supported by the data presented in figs 6 and 7, which confirmed (assuming no unusual genomic deletions, insertions, or rearrangements) that a total of 4131 bp of HPV-18 DNA is present in an integrated form in samples 1, 2, 8, 10, 11, and 12, from nucleotide positions 6558 to 2832 on the HPV-18 prototype genome. Therefore, the data obtained by semi-nested PCR analysis add support to the data obtained by Southern analysis, where ~ 4000 bp of integrated HPV-18 DNA was shown to be present in the samples. Although 1491 bp of integrated HPV-18 DNA was shown to be present in samples 4 and 13, from nucleotides 6558 to 151, the expected DNA fragment from the E1 region was absent. This was thought to be caused by small amounts of target DNA or the loss of target DNA in this region of the integrated virus. The results of the Southern analyses on these samples showed that the junction between subgenomic fragments F1 and F4, at nucleotide position 2440,

must be present (fig 1). Therefore, unless any unusual genomic rearrangement had occurred in the integrated HPV-18 DNA, both the E1L1 and E1L2 oligonucleotides should have been able to hybridise with target DNA. If it is assumed that the lack of amplified HPV-18 DNA in samples 4 and 13 was caused by the limits of detection; the only other likely explanation is that the E1R1 oligonucleotide at position 2832 was unable to hybridise to the target DNA. In turn, this suggests that in these samples the viral integration point is near to nucleotide position 2832, which is close to the E1/E2 boundary. This would support the Southern analysis data, which indicate that HPV-18 DNA integration occurs at the E1/E2 boundary, with the concomitant loss of the downstream ORFs E4, E5, L2, part of L1, and part or all of E2. Figure 8 shows the region of HPV-18 that we conclude to be present after viral DNA integration.

An unexpected observation arising from the HindIII Southern analysis data was that the HPV-18 DNA fragments generated from all the cervical carcinoma samples studied were remarkably similar. These data suggest that there was a preferential integration site within both the host and virus genomes, otherwise fragment size variation would be expected because of variation in the host chromosomal flanking regions of different sequences. The results described here are in agreement with others, which suggest that HPV-18 DNA integration occurs at a preferential site within the viral and host genomes.[13 15 39–42] The site of HPV-18 DNA integration in a cell line derived from an HPV-18 DNA positive cervical carcinoma has been localised to chromosomes 8q24 and 2p24.[43] Interestingly, chromosome positions 8q24–qter and 2p23–24 coincide with the location of c-myc and N-myc, respectively.[41 43–45] However, more experimental data are required to determine the importance of these observations.

Our observation that the E2 region is always lost in integrated HPV-18 implies that the transcriptional regulator encoded by this gene is always absent. In this, HPV-18 differs from HPV-16, where in at least a few carcinomas the E2 product is present, either because of episomal forms of the virus or because, in the integrated form, the gene is not deleted. However, the regions of the HPV-18 genome that are integrated are similar to those of HPV-16 with respect to E6, E7, and the LCR only.

Our findings, namely the retention of the "carcinogenic" region of the HPV-18 genome and the predominantly single major integrant at specific viral and host genomic sites, suggest several mechanisms for the deregulation of cell cycle control.

First, a combination of the loss of viral transcriptional regulation and active transcription of the integrated viral DNA could aid in the overexpression of the transforming genes E6 and E7 that is seen in cervical carcinomas. Alternatively, integration of HPV DNA could be adjacent to either a protooncogene or an uncharacterised tumour suppressor gene. The effect of this could be to place the adjacent

gene under the regulation of viral elements and, consequently, to enhance the deregulation of the tight control exerted over cellular replication and to aid the progression to a malignant phenotype. Because of the need to acquire more information about the host chromosomal integration site, the isolation of the viral–cellular junction fragments is of great importance.

1 Boring CC, Squires TS, Tong T. Cancer Statistics, 1993. *Cancer* 1993;**42**:19–38.
2 Parkin DM, Pisani P, Ferlay J. Estimates of the world-wide incidence of eighteen major cancers in 1985. *Int J Cancer* 1993;**54**:594–606.
3 Lorincz AT, Reid R, Bennett-Jenson A, *et al*. Human papillomavirus infection of the cervix: relative risk associations of 15 common anogenital types. *Obstet Gynecol* 1992;**79**:328–37.
4 Xavier-Bosch F, Manos MM, Muñoz N, *et al*. Prevalence of human papillomavirus in cervical cancer: a worldwide perspective. *J Natl Cancer Inst* 1995;**87**:796–803.
5 Iwasawa A, Nieminen P, Lehtinen M, *et al*. Human papillomavirus DNA in uterine cervix squamous cell carcinoma and adenocarcinoma detected by polymerase chain reaction. *Cancer* 1996;**77**:2275–9.
6 Nakagawa S, Yoshikawa H, Onda T, *et al*. Type of human papillomavirus is related to clinical features of cervical carcinoma. *Cancer* 1996;**78**:1935–41.
7 O'Connor MJ, Chan S-H, Bernard H-U. Transcription factor binding sites in the long control region of genital HPVs. In: Meyers G, Bernard H-U, Delius H, eds. *Human papillomavirus*. Los Alamos, Mexico: Los Alamos National Laboratory, 1995:21–40.
8 Mayor S. Human papillomavirus classified as carcinogenic. *BMJ* 1996;**313**:70.
9 Boshart M, Gissman L, Ikenberg H, *et al*. A new type of papillomavirus DNA, its presence in genital cancer biopsies and in cell lines derived from cervical cancer. *EMBO J* 1984;**3**:1151–7.
10 Schwarz E, Freese UK, Gissmann L, *et al*. Structure and transcription of human papillomavirus sequences in cervical carcinoma cells. *Nature* 1985;**314**:111–14.
11 Pater MM, Pater A. Human papillomavirus types 16 and 18 sequences in carcinoma cell lines of the cervix. *Virology* 1985;**145**:313–18.
12 Shirasawa H, Tomita Y, Kubota K, *et al*. Transcriptional differences of the human papillomavirus type 16 genome between precancerous lesions and invasive carcinomas. *J Virol* 1988;**62**:1022–7.
13 Shirasawa H, Tomita Y, Fuse A, *et al*. Structure and expression of an integrated human papillomavirus type 16 genome amplified in a cervical carcinoma cell line. *J Gen Virol* 1989;**70**:1913–19.
14 Wilczynski SP, Pearlman L, Walker J. Identification of HPV-16 early genes retained in cervical carcinomas. *Virology* 1988;**166**:624–7.
15 Wagatsuma M, Hashimoto K, Matsukura T. Analysis of integrated human papillomavirus type 16 DNA in cervical cancers: amplification of viral sequences together with cellular flanking sequences. *J Virol* 1990;**64**:813–21.
16 Daniel B, Mukherjee G, Seshadri L, *et al*. Changes in the physical state and expression of human papillomavirus type 16 in the progression of cervical intraepithelial neoplasia lesions analysed by PCR. *J Gen Virol* 1995;**76**:2589–93.
17 Daniel B, Rangarajan A, Mukherjee G, *et al*. The link between integration and expression of human papillomavirus type 16 genomes and cellular changes in the evolution of cervical intraepithelial neoplastic lesions. *J Gen Virol* 1997;**78**:1095–101.
18 Lehn H, Villa LL, Marziona F, *et al*. Physical state and biological activity of human papillomavirus genomes in precancerous lesions of the female genital tract. *J Gen Virol* 1988;**69**:187–96.
19 Fuchs PG, Girardi F, Pfister H. Human papillomavirus 16 DNA in cervical cancers and in lymph nodes of cervical cancer patients: a diagnostic marker for early metastases? *Int J Cancer* 1989;**43**:41–4.
20 Matsukura T, Koi S, Sugase M. Both episomal and integrated forms of human papillomavirus type 16 are involved in cervical cancers. *Virology* 1989;**172**:63–72.
21 Cullen AP, Reid R, Campion M, *et al*. Analysis of the physical state of different human papillomavirus DNAs in intraepithelial and invasive cervical neoplasm. *J Virol* 1991;**65**:606–12.
22 Das BC, Sharma JK, Gopalakrishna V, *et al*. A high frequency of human papillomavirus DNA sequences in cervical carcinomas of Indian women as revealed by Southern blot hybridisation and polymerase chain reaction. *J Med Virol* 1992;**36**:239–45.
23 May M, Dong X-P, Beyer-Finkler E, *et al*. The E6/E7 promoter of extrachromosomal HPV16 DNA in cervical cancers escapes from cellular repression by mutation of target sequences for YY-1. *EMBO J* 1994;**13**:1460–6.
24 Das BC, Sharma JK, Gopalakrishna V, *et al*. Analysis by polymerase chain reaction of the physical state of human papillomavirus type 16 DNA in cervical preneoplastic and neoplastic lesions. *J Gen Virol* 1992;**73**:2327–36.
25 Berumen J, Casas L, Segura E, *et al*. Genome amplification of human papillomavirus types 16 and 18 in cervical carcinomas is related to the retention of the E1/E2 genes. *Int J Cancer* 1994;**56**:640–5.
26 Choo K-B, Pan C-C, Han S-H. Integration of human papillomavirus type 16 into cellular DNA of cervical carcinoma: preferential deletion of the E2 gene and invariable retention of the long control region and the E6/E7 open reading frames. *Virology* 1987;**161**:259–61.
27 Choo K-B, Lee H-H, Pan C-C, *et al*. Sequence duplication and internal deletion in the integrated human papillomavirus type 16 genome cloned from a cervical carcinoma. *J Virol* 1988;**62**:1659–66.
28 Chen C-M, Shyu M-P, Chun H-W, *et al*. Analysis of deletion of the integrated human papillomavirus 16 sequence in cervical cancer: a rapid multiplex polymerase chain reaction approach. *J Med Virol* 1994;**44**:206–11.
29 Kitagawa K, Yoshikawa H, Onda T, *et al*. Genomic organisation of human papillomavirus type 18 in cervical cancer specimens. *Jpn J Cancer Res* 1996;**87**:263–8.
30 Chong T, Chan W-K, Bernard H-U. Transcriptional activation of human papillomavirus 16 by nuclear factor 1, AP1, steroid receptors and a possibly novel transcription factor, Pvf: a model for the composition of genital papillomavirus enhancers. *Nucleic Acids Res* 1990;**18**:465–70.
31 Tan S-H, Gloss B, Bernard H-U. During negative regulation of the human papillomavirus-16 E6 promoter, the viral E2 protein can displace Sp1 from a proximal promoter element. *Nucleic Acids Res* 1992;**20**:251–6.
32 Tan S-H, Leong L, Walker E-C, *et al*. The human papillomavirus type 16 E2 transcription factor binds with low cooperativity to two flanking sites and represses the E6 promoter through displacement of Sp1 and TFIID. *J Virol* 1994;**68**:6411–20.
33 Demeret C, LeMoal M, Yaniv M, *et al*. Control of HPV-18 replication by cellular and viral transcription factors. *Nucleic Acids Res* 1995;**23**:4777–84.
34 Steger G, Corbach S. Dose-dependent regulation of the early promoter of human papillomavirus type 18 by the viral E2 protein. *J Virol* 1997;**71**:50–8.
35 Manos MM, Ting Y, Wright DK, *et al*. Use of polymerase chain reaction amplification for the detection of genital human papillomaviruses. *Cancer Cells* 1989;**7**:209–14.
36 Cartwright NH, Sant-Cassia LJ, Easton AJ, *et al*. Detection of human papillomavirus in vulval carcinoma using a semi-nested PCR and restriction enzyme typing: a rapid and sensitive technique. *J Clin Pathol: Mol Pathol* 1996;**49**:M236–9.
37 Shirasawa H, Tomita Y, Sekiya S, *et al*. Integration and transcription of human papillomavirus type 16 and 18 sequences in cell lines derived from cervical carcinomas. *J Gen Virol* 1987;**68**:583–91.
38 Gloss B, Bernard H-U. The E6/E7 promoter of human papillomavirus type 16 is activated in the absence of E2 proteins by a sequence-aberrant Sp1 distal element. *J Virol* 1990;**64**:5577–84.
39 Lazo P. Structure, DNase I hypersensitivity and expression of integrated papillomavirus in the genome of HeLa cells. *Eur J Biochem* 1987;**165**:393–401.
40 Choo K-B, Lee H-H, Liew L-N, *et al*. Analysis of the unoccupied site of an integrated human papillomavirus 16 sequence in a cervical carcinoma. *Virology* 1990;**178**:621–5.
41 Choo K-B, Chen C-M, Han C-P, *et al*. Molecular analysis of cellular loci disrupted by papillomavirus 16 integration in cervical cancer: frequent viral integration in topologically destabilized and transcriptionally active chromosomal regions. *J Med Virol* 1996;**49**:15–22.
42 Gallego MI, Lazo P. Deletion in human chromosome 12q13–15 by integration of human papillomavirus DNA in a cervical carcinoma cell line. *J Biol Chem* 1995;**270**:24321–6.
43 Couturier J, Sastre-garau X, Schneider-Maunoury S, *et al*. Integration of papillomavirus DNA near c-myc genes in genital carcinomas and its consequences for proto-oncogene expression. *J Virol* 1991;**65**:4534–8.
44 Dallafavera R, Bregni M, Erikson J, *et al*. Human c-myc oncogene is located on the region of chromosome-8 that is translocated in Burkitt-lymphoma cells. *Proc Natl Acad Sci USA* 1982;**79**:7824–7.
45 Schwab M, Varmus HE, Bishop JM, *et al*. Chromosome localisation in normal human cells and neuroblastomas of a gene related to c-myc. *Nature* 1984;**308**:288–91.