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.

Keywords: human papillomavirus type 18 (HPV-18); integration; Southern analysis, cervical carcinoma

Cervical cancer is the second most common cancer in women worldwide.1–3 Human papillomavirus (HPV) DNA is present in ∼90% of cervical carcinomas, with HPV-16 and HPV-18 being the two most prevalent types.4–5 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). 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.6 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.7–22 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.10–23 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.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-
Most patients were found to have a squamous carcinoma of the cervix. The exceptions were samples 1–13.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Carcinoma type</th>
<th>Differentiation</th>
<th>Pelvic nodal status</th>
<th>Age (years)</th>
<th>HPV type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Squamous</td>
<td>Poor</td>
<td>Involved</td>
<td>37</td>
<td>18</td>
</tr>
<tr>
<td>2</td>
<td>Squamous</td>
<td>Poor</td>
<td>Involved</td>
<td>41</td>
<td>18</td>
</tr>
<tr>
<td>3</td>
<td>Squamous</td>
<td>Moderate</td>
<td>No involvement</td>
<td>69</td>
<td>18</td>
</tr>
<tr>
<td>4</td>
<td>Squamous</td>
<td>NA</td>
<td>NA</td>
<td>69</td>
<td>18</td>
</tr>
<tr>
<td>5</td>
<td>Squamous</td>
<td>NA</td>
<td>No involvement</td>
<td>45</td>
<td>16 and 18</td>
</tr>
<tr>
<td>6</td>
<td>Squamous</td>
<td>Moderate/Poor</td>
<td>Involved</td>
<td>41</td>
<td>16 and 18</td>
</tr>
<tr>
<td>7</td>
<td>Adenocarcinoma</td>
<td>NA</td>
<td>No involvement</td>
<td>41</td>
<td>16 and 18</td>
</tr>
<tr>
<td>8</td>
<td>Squamous</td>
<td>NA</td>
<td>NA</td>
<td>46</td>
<td>18</td>
</tr>
<tr>
<td>9</td>
<td>Squamous</td>
<td>NA</td>
<td>NA</td>
<td>69</td>
<td>18</td>
</tr>
<tr>
<td>10</td>
<td>Squamous</td>
<td>Poor</td>
<td>Involved</td>
<td>41</td>
<td>18</td>
</tr>
<tr>
<td>11</td>
<td>Squamous</td>
<td>Poor</td>
<td>Involved</td>
<td>37</td>
<td>18</td>
</tr>
</tbody>
</table>

Table 1 A summary of the histological analysis, age, and HPV type detected for patient samples 1–13.

Of note, all samples showed a disruption of the E6 and E7 genes. However, the loss of E2 was not observed.

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×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 (tacccggatccggatttcggttgctcc162) and P2 (tacccggatccggatttcggttgctcc162) 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 DNA was cloned into plasmid vector pBR322 (clone 18) and used as a DNA probe.

**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-32P 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-

**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), S' nucleotide position of each oligonucleotide and size (bp) of the HPV-18 DNA amplicons are shown. LCR, long control region.

**Figure 3** (A–C) HindIII restriction enzyme digestion of clinical samples 1–7. HindIII generated HPV-18 DNA fragments were detected using 32P-labelled subgenomic fragments (F2–F4) and sized using a 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.

**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
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 individually 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
Integration of HPV-18 DNA in cervical carcinoma

PvuII generated HPV-18 DNA fragments were detected using "S" labelled subgenomic fragments (F1–F4) and sized using a 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

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Nucleotide position</th>
<th>Fragment size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PstI</td>
<td>6766–5246</td>
<td>6337</td>
</tr>
<tr>
<td></td>
<td>5246–5313</td>
<td>518</td>
</tr>
<tr>
<td></td>
<td>5313–5770</td>
<td>557</td>
</tr>
<tr>
<td></td>
<td>5770–6325</td>
<td>555</td>
</tr>
<tr>
<td></td>
<td>6325–6766</td>
<td>441</td>
</tr>
<tr>
<td>PvuII</td>
<td>851–2100</td>
<td>1249</td>
</tr>
<tr>
<td></td>
<td>2100–6473</td>
<td>4373</td>
</tr>
<tr>
<td></td>
<td>6473–851</td>
<td>2235</td>
</tr>
</tbody>
</table>

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. 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.

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 position 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, in conjunction with oligonucleotides N0253 and LPCR3. Figure 2 shows the relative locations and orientations of...
DNA positive control; −, negative control. The numbers at the top of each lane refer to the patient sample numbers. M, 1 kb DNA ladder; + HeLa fragments of 347 bp were electrophoresed on a 1.5% (wt/vol) agarose gel.

**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 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. 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 cervical material, is that the LCR and the viral oncogenes E6 and E7 are always present. 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. 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. However, in view of the numerous HPV-16 positive cervical 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. 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. However, this latter study did not demonstrate which specific HPV-18 se-
The integration patterns of HPV-18 DNA in cervical carcinoma samples were investigated using semi-nested PCR and Southern analysis. The results showed that a considerable part of the viral genome is retained and E2 was lost, indicating a preferential integration site within the host chromosomal flanking regions of different sequences. 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 episonomal 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.