E2 sequence variations of HPV 16 among patients with cervical neoplasia seen in the Indian subcontinent

Narayanan Sathish, Priya Abraham, Abraham Peedicayil, Gopalan Sridharan, R.V. Shaji, George Chandy

Abstract

Objectives. Specific nucleotide variations in the E2 DNA sequence were looked for in samples with an intact human papillomavirus (HPV) 16 episomal E2 DNA.

Methods. Ninety-two women, 76 with invasive cervical carcinoma and 16 with cervical intraepithelial neoplasia (CIN) were recruited. HPV DNA typing was performed by polymerase chain reaction (PCR) based restriction fragment length polymorphism (RFLP). Intact episomal E2 DNA of HPV 16 was detected by PCR. Important nucleotide variations in samples with amplifiable E2 DNA were detected by RFLP. Nucleotide sequencing was performed on representative samples to confirm RFLP findings.

Results. A total of 89 (96.7%) women were positive for HPV DNA. Of these, 56 (63%) were positive for HPV 16, and of these, 38 (68%) were positive for intact episomal HPV 16 E2 DNA while 18 (32%) were negative. Samples with intact episomal HPV 16 E2 DNA sequences were grouped into four different digestion profiles I to IV based on RFLP patterns. Digestion patterns revealed absence of any sequence variations in samples with digestion profile I and presence of a 2983 A-G variation in those with profile II. Samples with digestion profiles III and IV revealed three variations in the hinge region (3516 C-A, 3538 A-C, 3566 T-G) and two in the DNA binding domain (3684 C-A, 3694 T-A) of the E2 sequence. Sequencing performed on representative samples confirmed RFLP findings.

Conclusions. PCR-RFLP helped in the identification of important HPV 16 E2 sequence variations, circumventing the need for sequencing. The presence of the nucleotide variations in positions that could alter the biological and immunological functions of the E2 protein combined with its increased occurrence in this study bring out the importance of these variations.

© 2004 Elsevier Inc. All rights reserved.

Keywords: Human papillomavirus (HPV) 16; Cervical neoplasia; Episomal E2; E2 sequence variations

Introduction

Several epidemiological studies have consistently associated the presence of human papillomavirus (HPV) with the development of cervical neoplasia [1–3]. HPV 16 is the predominant type detected in high-grade cervical intraepithelial neoplasia (CIN) and in invasive cervical carcinoma [1,2].

The viral oncoproteins of HPV 16, E6 and E7, bring about the degradation of tumor suppressor proteins, p53 and pRb, respectively, resulting in the loss of cell-cycle control [4,5]. Expression of these oncoproteins is under the control of the viral E2 protein. It has been shown that integration of the HPV 16 viral genome into the cellular genome occurs during progression of cervical lesions from preinvasive to invasive
Disruption of the E2 gene is often considered a major risk factor for the progression from CIN to invasive carcinoma [10,11] and is also associated with poor disease survival from cervical cancer [11]. A majority of the studies, consistent with the above observation, have reported a relatively low frequency of episomal E2 forms of HPV 16 both in advanced dysplasia and in invasive carcinoma [12–15]. On the other hand, few studies have reported an increased frequency of intact episomal HPV 16 viral DNA in invasive cancer [16,17]. One important cause for such variations in the reported frequency of E2 detection in HPV 16-associated cervical lesions could be the differing sensitivity of the methodologies employed to detect the E2 gene.

Presence of an intact HPV 16 E2 gene with concomitant expression of the viral oncogenes in advanced cervical neoplastic lesions suggests that disruption of E2 alone is not a prerequisite for the development of cervical carcinomas and that alternative mechanisms for increasing the expression of viral oncogenes could also exist. One such mechanism suggested is the alteration of the E2 function as a result of mutation or variation [18,19]. As the E2 protein is an important viral protein predominantly involved in viral replication [20] and in regulation of expression of the viral oncogenes [21], nucleotide and amino acid variations in this protein could drastically alter the biological activity as well as the outcome of HPV 16-associated infections.

Although nucleotide variations in the E6 gene of HPV 16 have been extensively studied, [22–24] studies examining such variations in the E2 gene are only few in number [18,19,25,26,31], with no such studies reported from the Indian subcontinent. In our study, since a majority of cases of advanced cervical neoplasia showed the presence of an intact HPV 16 episomal E2 gene, we looked for the presence of specific nucleotide variations in the E2 DNA sequence.

Specific variations based on previous literature [19,26] present in the regions of the E2 gene of HPV 16 important for biological and immunological functions were looked for in this study. A polymerase chain reaction (PCR)-based restriction fragment length polymorphism (RFLP) was standardized and employed for this. Additionally, to validate the RFLP findings, nucleotide sequencing was carried out on a representative number of samples.

Materials and methods

Study subjects (n = 92)

Cervical biopsies were collected from 92 women, 76 with invasive cervical carcinoma and 16 with cervical intraepithelial neoplasia (CIN). These samples were obtained from women undergoing biopsy or surgery in the Obstetrics and Gynaecology Unit III of the Christian Medical College, Vellore, during the time period June 2001–September 2003. Ethical clearance from the institutional research committee and verbal consent from the patients were obtained for all the investigations.

Biopsied samples were transferred into sterile tubes containing viral transport medium (VTM) and transported to the virology laboratory immediately at +4°C. Upon receipt of tissues in the laboratory, they were cut into small pieces, weighed, and stored as multiple aliquots at −60°C.

Detection and typing of HPV DNA

DNA was extracted from the tissues using the QIAamp DNA Mini Kit (Qiagen, Germany). Following extraction, HPV DNA amplification was carried out using the consensus primers MY09/MY11 [27]. To ensure specimen adequacy and integrity of the DNA extraction process, concomitant amplification of the beta-globin gene was also performed using specific primers [28]. The PCR reaction conditions were similar to that described earlier [29].

Typing of samples positive for HPV DNA was performed by restriction fragment length polymorphism (RFLP) using a set of seven restriction endonucleases BamHI, DdeI, HaeIII, HinfI, PstI, RsaI, and Sau3AI (Roche Diagnostics, Mannheim, GmBh, Germany) as described earlier [29].

HPV 16 E2 DNA detection

DNA extracts positive for HPV 16 were processed with boiling at 100°C for 7 min, followed by immediate freezing for the denaturation of template DNA to start an effective amplification followed by PCR amplification as described earlier [29].

Detection of HPV 16 E2 variations by RFLP

Specific variations if present in the amplified HPV 16 E2 DNA product were detected by RFLP using a set of three restriction endonucleases DdeI, PstI, and RsaI (Roche Diagnostics). Digested products were detected by electrophoresis through an ethidium bromide containing 4% agarose gel. Digestion patterns obtained were analyzed with the aid of a gel documentation system (Gel Doc 2000, BioRad). Specific variations were detected, based on the interpretation of the different digestion patterns obtained with each of the enzymes.

The specific variations thus detected by RFLP were verified using the NEB cutter V2.0 (http://tools.neb.com/NEBcutter2/index.php). For this, the nucleotide variations detected by RFLP were incorporated into the prototype HPV 16 sequence (HPV-16R sequence, Los Alamos Laboratory), and such a putative sequence was analyzed by the NEB Cutter V2.0 with the corresponding enzyme.
used for RFLP to get a digestion pattern. The digestion pattern obtained by RFLP and that generated by NEB cutter V2.0 were compared.

**HPV 16 E2 DNA sequencing**

Specific variations in the HPV 16 E2 DNA detected by RFLP were further verified by sequencing of the E2 DNA in a representative number of samples. Products for sequencing were generated using the same primers as above used for the amplification of HPV 16 E2 DNA. Amplified products were subjected to cycle sequencing PCR with the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit (Perkin-Elmer). Sequencing reactions were then run on the ABI PRISM 310 Genetic Analyser (PE Applied Biosystems, Foster City, CA). Viral sequences were analyzed by sequencing analysis software and sequencing navigator. Specific variations in the sequences were detected by comparison of the obtained sequences with a standard sequence (HPV 16R) using the ClustalW software (European Bioinformatics Institute). E2 gene variations detected by RFLP and those detected by sequencing were correlated.

**Statistical analysis**

Data were analyzed by the $\chi^2$ test using the statistical package EPI INFO (Version 6.04b). Results were considered statistically significant at $P < 0.05$.

**Results**

**HPV DNA detection results**

Of the total 92 women recruited for the study, 89 (96.7%) were positive for HPV DNA. HPV 16 was detected in a total of 56 (63%) women. These 56 women constituted the study population.

**Histopathological diagnosis**

Forty-six of the HPV 16-positive women were diagnosed to have invasive cervical carcinoma, and 10 were diagnosed to have CIN (1 diagnosed as CIN I, 2 as CIN II, and seven as CIN III).

**HPV 16 episomal E2 DNA results**

Thirty-eight (68%) samples were positive for the presence of intact episomal E2 DNA, while 18 (32%) were negative for episomal E2 DNA suggesting the presence of integrated forms. The increased prevalence of intact episomal HPV 16 E2 DNA as compared to the prevalence of nonepisomal (Integrated) forms was statistically significant ($P < 0.001$).

**HPV 16 E2 RFLP digestion profile**

Based on the digestion patterns obtained with the three different restriction endonucleases (in the samples positive for episomal HPV 16 E2 DNA), isolates were grouped into four different digestion profiles (designated types I to IV). Shown in Table 1 are the different digestion profiles obtained, their interpretation, and total number of samples showing the different profiles along with the clinical staging of these samples.

**HPV 16 E2 DNA sequencing results**

All the four samples with type I digestion profile and four samples each from those showing types II–IV digestion profiles were sequenced. Samples with type I digestion profile upon sequencing did not reveal the presence of any

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Different digestion profiles obtained upon RFLP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of digestion profile</td>
<td>Restriction enzyme</td>
</tr>
<tr>
<td>I</td>
<td>DdeI</td>
</tr>
<tr>
<td></td>
<td>PstI</td>
</tr>
<tr>
<td></td>
<td>RsaI</td>
</tr>
<tr>
<td>II</td>
<td>DdeI</td>
</tr>
<tr>
<td></td>
<td>PstI</td>
</tr>
<tr>
<td></td>
<td>RsaI</td>
</tr>
<tr>
<td>III</td>
<td>DdeI</td>
</tr>
<tr>
<td></td>
<td>PstI</td>
</tr>
<tr>
<td></td>
<td>RsaI</td>
</tr>
<tr>
<td>IV</td>
<td>DdeI</td>
</tr>
<tr>
<td></td>
<td>PstI</td>
</tr>
<tr>
<td></td>
<td>RsaI</td>
</tr>
</tbody>
</table>

*Note: Sequence variations are indicated in bold.

a Cervical intraepithelial neoplasia.

b Squamous cell carcinoma.

c Adenocarcinoma.
sequence variations (absence of variations was also detected by RFLP). Samples with type II digestion profile did not reveal the presence of any additional variation other than that detected by RFLP. Samples with types III and IV digestion profiles, however, revealed variations in addition to that detected by RFLP. These nucleotide variations seen in the E2 DNA sequences in the representative samples belonging to each of the profiles is shown in Fig. 1.

Discussion

In this study, a significantly higher proportion of samples (68%) revealed the presence of an intact HPV 16 episomal E2 DNA sequence as compared to those harboring purely integrated forms (P < 0.001). These results differ from those of other studies that documented an increased prevalence of integrated forms of the viral sequence in advanced cervical neoplastic lesions [12–15], wherein the E2 sequence tends to get disrupted. Thus, our study seems to contradict the fact that integration is one of the most important prerequisites for the development of invasive cervical cancer as, of the 38 samples positive for episomal HPV 16 E2 DNA, 30 (79%) were samples of invasive cervical cancer. A plausible explanation for the absence of disrupted E2 sequences in a high proportion of samples could be sampling error. Thus, there is a probability that the nonbiopsied sites could have harbored disrupted HPV 16 E2 sequences.

A very sensitive methodology as described earlier [14] was employed for the amplification of episomal HPV 16 E2 DNA in this study. Here, before amplification, the DNA extracts were boiled for 7 min at 100°C followed by immediate freezing. This ensures that the DNA strands get separated and are held apart as single strands. As the HPV 16 E2 open reading frame (ORF) amplified in this study spanned a long region of 1027 bp, the abovementioned additional step becomes essential to start of an efficient amplification. This could ensure that even minimal amounts of episomal E2 DNA get amplified, which will go undetected if an insensitive PCR methodology is adopted.

The HPV 16 E2 ORF in this study was amplified by the use of a single primer pair that thereby avoids the amplification of different regions of the E2 ORF by separate primer pairs as described earlier [18,19]. Amplification of the HPV 16 E2 ORF with a single primer set followed by RFLP using a set of three different restriction endonucleases to detect important nucleotide variations is, to the best of our knowledge, described for the first time in this study. The different digestion profiles (I to IV) presented in this study and their interpretation could definitely help in the identification of important sequence variations in the HPV 16 E2 DNA sequence (including the 3516 C-A, 3538 A-C, 3566 T-G variations present in the DNA binding domain of the E2 ORF). This could thus circumvent the need to identify these important variations by nucleotide sequencing which could be far too expensive to be carried out on all samples.

Of the two variations detected in the DNA binding domain of the E2 ORF by PCR-RFLP, the importance of the 3684 C-A variation (that results in an amino acid change from threonine to lysine) has already been brought out in two earlier studies [18,19]. This variation lies adjacent to the conformational structure of this helix and the conformation of the E2 protein. It has been suggested in an earlier study [19] that the interaction between the cellular transcription factors the E2 protein and the viral promoter/enhancer region may be altered by this variant. Although the 3694 T-A variation does not result in a change in the coded amino acid, from its proximity to the 3684 C-A variation, it could be speculated that this variation along with the 3684 C-A variation could contribute to altering the conformational structure of the DNA binding helix.

Of the three variations detected in the hinge region of the E2 ORF by PCR-RFLP, the 3516 C-A variation (resulting in amino acid change from threonine to asparagine) and the 3566 T-G variation (resulting in amino acid change from phenylalanine to valine) are of significance. These variations are located in two of the most reactive B-cell epitopes [30]; as a result of which, there could be potential alteration of humoral response to the virus. Although the 3538 A-C variation did not result in an amino acid change, presence of this in conjunction with the above two variations in all of the samples should be attributed importance.

Twenty-four (63%) of the 38 samples positive for HPV 16 episomal E2 DNA revealed the presence of the above-mentioned 5 variations (3 in the hinge region and 2 in the DNA binding domain) in comparison to 14 (37%) that did
not harbor these variations. This difference was statistically significant \( (P < 0.05) \). Although the examined cases of CIN I and II were very few, none of these variations were detected in these early preinvasive lesions.

Of significance is the fact that, in contrast to an earlier study [19] where there was a total absence of the 3516 C-A and 3538 A-C variations in cervical neoplastic lesions, our study reports a high prevalence of these variations thereby bringing out the importance of these variations in our study group. Interestingly in all of the 11 samples with type III digestion profile, an additional variation, the 3371 T-C variation in the hinge region of E2, was detected.

A new variation, the 2983 A-G, in the transactivation domain of the HPV-16 E2 ORF is reported for the first time in this study, detected in all samples belonging to the type II digestion profile. Although the functional significance of these variations could not be assessed by us, the presence of this variation in 10 (26.3%) of the 38 samples positive for these variations could not be assessed by us, the presence of this variation in 10 (26.3%) of the 38 samples positive for episomal E2 could have implication that warrants further studies.

On sequencing four samples each from the group showing digestion profiles III and IV, numerous variations (in addition to those detected by RFLP) were detected as represented in Fig. 1. Among these variations detected, the 3159 C-A, 3161 C-T, 3517 T-C, and the 3566 T-G variations are of functional significance, as these positions represent some of the important humoral epitopes in the E2 protein [30]. These variations hence could result in alteration of immune epitopes, resulting in altered humoral responses to the virus as alluded to earlier in the discussion. All four samples sequenced from groups showing types III and IV digestion profiles revealed an identical set of variations. Thus, there arises a possibility, wherein the other samples belonging to these digestion profiles could also be harboring the same variations as seen in the representative samples sequenced. This could be of significance as a majority (63%) of samples in our study showed the types III and IV digestion profiles which harbored important sequence variations. Presence of E2 variations in regions which could alter the immune epitopes/function of the protein seems to be in concordance with findings of earlier studies [18,19,25,31].

It has been suggested earlier [21] that high concentrations of episomal E2 could repress the expression of the viral oncogenes by preventing the binding of cellular transcription factors (that enhance the expression of the oncogenes) to their respective sites within the P97 promoter. Thus, it has been suggested that the loss of E2 function as a consequence of disruption or mutation (as in this study) could result in the upregulation of the viral promoter and hence an increased expression of the oncogene transcripts [32,33]. In accordance with this, in our study, samples showing the presence of HPV 16 episomal E2 DNA also revealed the concomitant expression of the viral oncogenes (data not shown). The extensive sequence variations seen (in samples belonging to digestion profiles III and IV) could have altered drastically the conformation of the E2 protein. As a result of this, either an absence of E2 binding or an inefficient E2 binding to the P97 promoter sites could occur. Both of these possibilities could result in the binding of the respective cellular transcription factors to the P97 promoter sites, resulting in enhanced expression of these oncogenes. Alternatively, it could be that, due to the increased sensitivity of the PCR methodology employed, even minimal amounts of E2 present in the sample could have been amplified. Such minimal E2 concentration has been shown to stimulate P97 transcription, thereby resulting in increased oncogene expression [34].

E2 protein is a potent transcriptional activator and greatly increases viral DNA replication by colocalizing the viral E1 protein to the origin of replication [20]. The region of the E2 protein required for this association with the E1 protein is located near the N-terminal transactivation domain [18,20]. The total absence of variations in this region (amino acid positions 18–41) in our study isolates suggests that the binding capacity of E2 to E1 could remain unaltered, also suggesting that the replication efficiency of E2 would remain unaffected.

The difference in the percentage prevalence of squamous cell carcinoma isolates with types I and II profile as compared to those with types III and IV profiles was not of any statistical significance. The same was also true for cases of adenocarcinoma. Thus, no discernible difference was observed in the association of the E2 variants with histological type of cancer.

As part of a separate study, HPV 16 E6/E7 nucleotide sequencing was performed on all samples of this study (data not shown). The region sequenced encompassed nucleotide positions 142–666 encompassing the HPV 16 E6/E7 ORFs. A total of 21 isolates revealed presence of both episomal HPV 16 E2 DNA and the E6 350G variant. Among these, 5 (23.8%) samples did not harbor the important E2 variations (types III and IV digestion profiles), while 16 (76.2%) samples harbored these variations, this difference being statistically significant. Thus, a significant proportion of HPV 16 350G variants harbored the important E2 sequence variations. It could thus be speculated from this that both the HPV 16 E2 and E6 variations could act in conjunction, influencing the pathogenic potential of HPV 16.

In conclusion, this is the first study of its kind reporting on specific E2 sequence variations in strains of HPV 16 from the Indian subcontinent. The sensitive PCR methodology adopted here will definitely help in amplification of even low amounts of the E2 ORF. PCR using a single pair of primers to amplify the E2 ORF followed by RFLP as described in this study should be of value in the identification of important sequence variations. This hence could serve as an easier alternative to sequencing as well as to other techniques that involve amplification of several regions of the E2 ORF in overlapping fragments. The increased prevalence of episomal HPV 16 E2 DNA in
advanced cervical neoplastic lesions seems to substantiate the finding that integration does not necessarily occur during the progression of cervical neoplasia. The presence of E2 sequence variations in the study isolates in positions that could potentially alter the immune epitopes and function of the protein suggests that the molecular activity of the E2 protein and thereby the outcome of HPV infections could be altered as a result of this. The E2 sequence variations could also serve as an alternative mechanism deregulating the expression of the HPV E6/E7 oncoproteins. To conclude, future studies are required to provide comprehensive information on the significance and the potential role of these E2 variations.

References

