HPV 16 E6 sequence variations in Indian patients with cervical neoplasia

Narayanan Sathish\textsuperscript{a}, Priya Abraham\textsuperscript{a,}\textsuperscript{*}, Abraham Peedicayil\textsuperscript{b}, Gopalan Sridharan\textsuperscript{a}, George Chandy\textsuperscript{c}

\textsuperscript{a}Department of Clinical Virology, Christian Medical College, Vellore 632004, India
\textsuperscript{b}Department of Obstetrics and Gynaecology, Christian Medical College, Vellore, India
\textsuperscript{c}Department of Clinical Gastroenterology and Hepatology, Christian Medical College, Vellore, India

Received 24 March 2005; received in revised form 15 April 2005; accepted 20 April 2005

Abstract

In a cross-sectional study performed between June 2001 and November 2003, the HPV 16 E6 gene of 50 women with cervical neoplasia and 20 cytologically normal women (‘controls’) was sequenced following amplification by PCR. The 350T to G variant was seen in 35 (70%) of 50 patients’ isolates while it was seen in only 3 (15%) of the 20 isolates from the ‘controls’. The higher occurrence of the 350G variant among the patients was statistically significant ($P < 0.01$). Isolates from patients were grouped into the European (E), Asian-American (AA) and North-American (NA-1) phylogenetic clusters with 46 (92%) belonging to the E cluster. All the 20 isolates from ‘controls’ belonged to the E cluster. This study suggests an association of the HPV 16 350G variant with a higher risk of cervical neoplasia and a predominance of E lineage strains among Indian HPV 16 isolates.

© 2005 Elsevier Ireland Ltd. All rights reserved.

Keywords: Human papillomavirus (HPV)-16; Cervical neoplasia; E6 variants

1. Introduction

The HPV 16 E6 oncoprotein specifically interacts and brings about the degradation of the tumor suppressor protein, p53 \cite{1}. The E7 oncoprotein of HPV 16 on the other hand binds to the retinoblastoma gene product, pRb \cite{2}. This brings about the dissociation of complexes containing the cellular transcription factor (E2F) bound to pRb resulting in the release of E2F that promotes cellular proliferation \cite{3}. These mechanisms mediated by the E6 and E7 oncoproteins ultimately result in loss of cell-cycle control promoting cervical carcinogenesis. Transcription of these viral oncogenes is controlled by the p97 promoter located in the proximal region of the long control region (LCR) of the HPV 16 genome \cite{4,5}. These observations highlight the importance of the HPV 16 E6, E7 genes and the LCR in the pathogenesis of cervical neoplasia. Thus sequence variations in all of the above could drastically
influence the oncogenic potential of the virus and thereby the outcome of HPV 16 associated infections.

Naturally occurring variants of HPV 16 have been reported to alter certain biological and biochemical properties [6]. This has been evidenced by their reduced capacity to suppress keratinocyte differentiation responses and their increased ability to induce p53 degradation. Sequence variations in the HPV 16 genome were also seen to affect the virus assembly [7], immunological recognition by the host [8] and immortalization activity [6].

The HPV 16 genome is polymorphic with respect to the different geographical regions over the globe and has been classified into different phylogenetic clusters (variants). This classification is based on the sequence variations of the E6, L1, L2 and the LCR regions [9,10]. The E6 region of the HPV 16 genome seems to be less conserved and shows a lot of variations throughout its length [9,10]. The significance attributed to the E6 gene variations is due to its occurrence in regions critical for p53 degradation [11] and in regions involved in host immune recognition [12]. The most studied among the HPV 16 E6 gene variations includes a change from T to G at nucleotide position 350 which results in the change of the encoded amino acid from leucine (L) to valine (V). Several earlier studies have brought out the significance of this variation in the context of its association with progression of cervical neoplasia [13–17].

The present study was undertaken to investigate the occurrence of the HPV 16 350G variant among Indian women with cervical neoplasia. Additionally, the distribution of different phylogenetic clusters of HPV 16 was also looked at.

2. Materials and methods

2.1. Patients (n = 50)

The patient group consisted of the first 50 (Mean age ± SD: 42 ± 6.1) of a total of 70 women with cervical neoplasia positive for HPV 16. The clinical staging of the 50 women comprised of one each diagnosed with CIN I and II, respectively, seven with CIN III and 41 with invasive carcinoma. A total of 39 women were from south India while the remaining 11 were from east India. These were women attending the departments of Obstetrics and Gynaecology (over a time period of June 2001–November 2003) at the Christian Medical College and Hospital, Vellore, India. Clinical staging of invasive cervical carcinoma was done by the examining clinician according to the International Federation of Gynaecology and Obstetrics (FIGO) criteria.

Ethical clearance to perform this study was obtained from the institutional research committee. Verbal consent was also obtained from all patients after explaining the nature and the expected outcome of the study.

2.2. Controls (n = 20)

This consisted of samples obtained from 20 (Mean age ± SD: 35.8 ± 4.2) HPV 16 positive but cytologically normal women all of whom were from south India. This was drawn from a total of 250 women from the general community who were screened for the presence of HPV DNA.

2.3. HPV 16 DNA detection

HPV DNA was extracted from the cervical biopsies with the QIAGEN DNA Mini Kit (Qiagen, Germany). Presence of HPV 16 in the cervical biopsies obtained from women with cervical neoplasia and in cervical swabs collected from the cytologically normal control women was determined by the line blot assay (following amplification of HPV DNA with the PGMY primers) as described earlier [18].

2.4. HPV 16 E6 PCR

Amplification of the HPV 16 E6 gene was performed by a non-nested PCR in a reaction mixture consisting of 10 μL of the extracted DNA, 1.5 mM MgCl2, 200 μM dNTPs, 20 pmol of specific primers and 1 U of Taq DNA polymerase enzyme (Roche Diagnostics, Gmbh, Mannheim, Germany) in a final reaction volume of 50 μL. The primers used span a region of both the E6/E7 open reading frames (ORFs) of HPV 16 from nucleotides 142–666 [19] thus allowing for the amplification of both the HPV 16 E6 and the E7 ORFs.

The thermal cycling profile consisted of 30 cycles of amplification each consisting of denaturation at
94 °C for 1 min, annealing at 55 °C for 1 min, and extension at 72 °C for 1 min with a post-PCR extension step at 72 °C for 10 min. Amplified HPV 16 E6/E7 DNA product (525 bp) was detected by electrophoresis in ethidium bromide containing agarose gel.

2.5. HPV 16 E6 sequencing

HPV 16 E6 DNA was amplified for sequencing using the same primers as above. Amplified products were subjected to cycle sequencing with the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit (Perkin-Elmer). Sequencing reactions were performed in the automated DNA sequencer, the ABI PRISM 310 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA). The obtained electropherograms were analysed using the sequencing analysis software (version 1.01, Applied Biosystems).

The sequences obtained were confirmed as HPV 16 by a basic local alignment search tool (BLAST) search. Identification of specific variations in the HPV 16 E6 DNA sequences was done by aligning the study sequences with a HPV 16 prototype sequence (GenBank K02718/HPV 16R) using the Clustal W multiple alignment software programme (European Bio Informatics Institute). Designation of phylogenetic clusters of the HPV 16 E6 DNA sequences was done as described earlier [10].

2.6. Statistical analysis

Data were analysed by the Chi² square test using the statistical package EPI INFO (version 6.04b). Differences between groups were considered statistically significant at \( P < 0.05 \). Logistic regression analysis was performed using the NCCS/PASS 2000 Dawson edition.

3. Results

3.1. Patients

Of the 50 HPV 16 E6/E7 DNA positive samples subjected to nucleotide sequencing, 38 (76%, 95% CI 62.8–86.3) samples revealed nucleotide variations in the E6 sequence while 12 (24%, 95% CI 13.7–37.2) samples revealed the prototype sequence.

A total of eight different E6 gene nucleotide variations were seen among the samples including the 350T to G variation. This also included three novel variations not reported thus far (the 276A to C, 421T to G and the 522G to C variation). The 276A to C variation resulted in a change of the encoded amino acid from asparagine (N) to serine (S) while the other two variations did not. The different HPV 16 E6 gene variations seen among these samples is shown in Table 1. The distribution of the different E6 variations in the HPV 16 positive samples with respect to staging

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Different HPV 16 E6 gene variations and predicted amino acids among patients with cervical neoplasia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nucleotide position</td>
<td>276</td>
</tr>
<tr>
<td>Reference nucleotide</td>
<td>A</td>
</tr>
<tr>
<td>C</td>
<td>–</td>
</tr>
<tr>
<td>–</td>
<td>A</td>
</tr>
<tr>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Variations other than 276A to C, 335C to T and 350T to G did not result in an amino acid change.

\(^a\) N58S denotes a change in the encoded amino acid from asparagine (N) to serine (S).

\(^b\) H78Y denotes a change in the encoded amino acid from histidine (H) to tyrosine (Y).

\(^c\) L83V denotes a change in the encoded amino acid from leucine (L) to valine (V).
of cervical neoplasia among the patients is shown in Table 2.

Based on the HPV 16 E6 sequences, the 50 samples could be grouped into the three different variant classes. These included the European (E), Asian-American (AA) and the North-American (NA-1) variant classes as shown in Fig. 1. Samples belonging to the E variant class comprised both the prototype E6 350T and the variant 350G. A total of 46 (92%) samples were identified as belonging to the E cluster while two (4%) samples each were identified as belonging to the AA and the NA-1 phylogenetic classes, respectively.

Of the total 50 samples, 15 (30%, 95% CI 18.55–43.7) samples were positive for the presence of 350T whilst the remaining 35 (70%, 95% CI 56.3–81.4) were positive for the 350G variant (substitution of T by G).

### 3.2. Controls

Based on the E6 sequences obtained, all the isolates were classified as belonging to the E cluster. HPV 16 E6 sequence analysis revealed that of the 20 HPV 16 isolates only three (15%) isolates showed the presence of the 350G variant whilst the remaining 17 (85%) isolates revealed the presence of the 350T prototype.

The mean age ± SD of the patient group was significantly higher as compared to the control group (P<0.01). There was, however, no statistically significant difference between both groups with respect to geographical origin. Multivariate analysis revealed that the increased prevalence of the HPV 16 350G variant in the patient group (70%) versus the control group (15%) was statistically significant (P<0.01) despite differences in mean age ± SD between the groups.

### 4. Discussion

Based on the E6 sequences obtained in this study, a majority (92%) of the HPV 16 isolates (from the patients) sequenced belonged to the E variant class whilst two (4%) isolates each were identified as belonging to the AA and the NA-1 classes.

![Nucleotide sequence variations (written vertically) in the HPV 16 E6 open reading frames (ORFs) among the study samples (from patients) that formed the basis for classification into different phylogenetic classes. Phylogenetic classes are noted by letters: E-P, European prototype; E-350G, European 350G variant; NA-1, North-American-1; AA, Asian-American. Nucleotides varying from the prototype are indicated in bold letters. Dashes, nucleotide positions where no variations were found.](image)
respectively. On the other hand, all isolates from the cytologically normal women belonged to the E cluster.

Though the number of samples analysed in this study were not high, to the best of our knowledge, similar studies on the distribution of HPV 16 variant classes have not been reported from the Indian subcontinent. The occurrence of AA and the NA-1 classes in this study is to be noted. The absence of HPV 16 strains belonging to the AA and the NA-1 classes in a previous Indian study published from south India [15] as compared to ours, suggests that HPV 16 strains in India could vary regionally. This however requires further substantiation by large-scale studies.

Another observation seen in this study was the absence of HPV 16 strains belonging to the Asian lineage (characterized by the presence of 178T to G variation). The previously reported Indian study [15] shows the same findings. This information contrasts with other studies from Asia [20,21] wherein a high prevalence of Asian lineage strains of HPV 16 has been reported. The absence of strains belonging to the Asian lineage in this study could be explained by the fact that majority of the samples studied were from south Indian women. A clearer picture with regard to HPV 16 strains of Asian lineage HPV 16 strains within India will emanate from studies from other regions of the country. Additionally, the part of the HPV 16 E7 gene sequenced in this study (from nucleotide positions 562–666) did not reveal any nucleotide variations. However, to obtain gainful conclusions on the conserved nature of the E7 region, sequencing of the entire region needs to be performed which was not pursued in this study.

Possibilities of spread of the molecular variants of HPV 16 from one continent to another have been shown to exist [22]. The predominance of E variants seen in this study comprising of south Indian patients mainly is difficult to explain. It is possible that these strains were acquired over a long time period as a result of Aryan migration into the Indian subcontinent. The Aryans share common linguistic, cultural and racial aspects with Europeans which may be traceable to a common ancestral home in central Asia. This could partially explain the predominance of E variants in this study.

Analysis of sequence variations in the HPV 16 positive tumour samples collected globally has shown the presence of the AA variants in the South East Asian countries (Indonesia, Thailand and Philippines) [10]. This possibly suggests that AA variants could be prevalent in other Asian countries and hence could explain the presence of the AA variants in our study also.

A limitation of this study was that the primers employed did not span the entire HPV 16 E6 ORF but instead spanned a region starting from nucleotide position 142 of the HPV 16 E6 gene. Hence, the possible presence of other variations in the E6 gene such as the 109T to C and the 131A to G could not be detected. However, these primers served to detect the variation that has been attributed maximum significance, i.e. the 350T to G variation. Additionally, the designation of the variant classes was still possible with the E6 region amplified by primers used in this study.

Another limitation of our study was the very few isolates of CIN I and II examined. Increased number of isolates in these categories could not be examined as most of the patients who presented to the outpatient clinics were only those with advanced cervical lesions. Among the study patients with high-grade cervical lesions (seven with CIN III and 41 with invasive cancer), a high proportion of them (70.8%) revealed the 350G variant as compared to those with 350T. To assess the true significance of the high prevalence of the 350G variant in the patient group, the presence of this variant was also looked for in 20 HPV 16 positive cytologically normal women. Among the 20 women, only three (15%) revealed the presence of the 350G variant. There was a statistically significant ($P$ < 0.01) difference in the prevalence of 350G in this group as compared to the patient group.

Few characteristics like the mean age and the geographical origin of the patient and the control groups were examined to find out if the two groups were comparable. Though the mean age ± SD of the patient group was significantly higher as compared to the control group ($P$ < 0.01) there was no statistically significant difference between the two groups with respect to geographical origin. The higher detection rates of the 350G variant among patients (70%) as compared to controls (15%) was highly significant ($P$ < 0.01) as per multivariate analysis. Among the other factors, which could differ between the two
groups is the socioeconomic status. As information on this was not available for a proportion of patients this parameter was not taken into account. However, available data revealed that all women in the control group and a high proportion (80%) of them among patients were either themselves or spouses of labourers thus belonging to the lower socioeconomic strata.

Earlier studies [13–15,23,24] from different parts of the globe have looked at the occurrence of the HPV 16 350G variant among women with cervical cancer. In two of these studies [13,23], the prevalence of HPV 16 350G variant among cervical cancer isolates were high with percentage detection rates of 88% [13] and 55.5% [23], respectively. In one other study [14] performed on Swedish and Finnish women with cervical cancer, the 350G variant was seen to be enriched in patients with high-grade lesions and cancer as compared to variants in the long control region. In yet another recent study [24] carried out on Swedish women, the 350G variant in association with certain HLA haplotypes seemed to increase the risk of cervical cancer approximately 4–5 fold. In the only published Indian study [15], a preponderance of 350G was seen among younger women with cervical neoplasia and the authors had suggested that this variant could be associated with the more aggressive nature of cervical cancer.

Against the above background of higher pathogenicity associated with the 350G variant, the findings of our study reiterate the same. The significance of the 350T to G variant (in the context of cervical carcinogenesis) would have been more startling in this study, if a higher number of cytologically normal women were examined.

The true significance of such variations can be generated only if various other factors that could also influence the outcome of the HPV 16 variant associated infections are taken into account. These include presence of different HLA genotypes, polymorphisms of host cellular proteins like p53, environmental factors and duration of infection.

In conclusion, this study reports the prevalence of different variations in the HPV 16 E6 gene among patients with different stages of cervical neoplasia. We have also reported on the distribution of different phylogenetic clusters of HPV 16. An increased prevalence of the E6 350T to G variation was seen in patients as compared to controls. The fact that HPV 16 is the predominant type in India, necessitates more of such comprehensive studies from different parts of the country on these variants. Such studies would significantly contribute to the existing knowledge on the role of HPV 16 variants in cervical neoplasia.

References


