Human Papillomavirus 16 Lineage D is Associated with High Risk of Cervical Cancer in the Brazilian Northeast Region

Papillomavirus humano 16 da linhagem D associado a alto risco de câncer de colo do útero em região do nordeste brasileiro

Luís Felipe Leite Martins¹  Miguel Ângelo Martins Moreira²  Rodrigo Alves Pinto³  Neilane Bertoni dos Reis⁴  Shayany Pinto Felix²  João Paulo Castello Branco Vidal⁵  Leuridan Cavalcante Torres³  Ariani Impieri Souza³  Liz Maria de Almeida⁶

¹ Cancer Surveillance and Data Analysis Division, Instituto Nacional de Câncer, Rio de Janeiro, RJ, Brazil
² Tumor Genetics and Virology Program, Instituto Nacional de Câncer, Rio de Janeiro, RJ, Brazil
³ Instituto de Medicina Integral Prof. Fernando Figueira, Recife, PE, Brazil
⁴ Division of Population Research, Instituto Nacional de Câncer, Rio de Janeiro, RJ, Brazil
⁵ Universidade Federal do Maranhão, São Luis, MA, Brazil
⁶ Coordination of Prevention and Surveillance, Instituto Nacional do Câncer, Rio de Janeiro, RJ, Brazil

Address for correspondence Ariani Impieri Souza, MD, PhD, Rua dos Coelhos, 300, 50070-902, Boa Vista, Recife, PE, Brazil (e-mail: ariani.impieri@gmail.com; ariani@imip.org.br).

Keywords
► HPV16
► lineage A
► lineage D
► CIN2/CIN3
► cervical cancer

Abstract
Objective  Similar to Human Papillomavirus (HPV) genotypes, different lineages of a genotype also have different carcinogenic capabilities. Studies have shown that specific genotype lineages of oncogenic HPV are associated with variable risks for the development of cervical intraepithelial neoplasia (CIN2/CIN3) and cervical cancer. The present study aimed to analyze the genetic diversity of the HPV16 genotype in women with CIN2/CIN3 and cervical cancer, from the northeast region of Brazil.

Methods  A cross-sectional multicenter study was conducted in the northeast region of Brazil, from 2014 to 2016. This study included 196 cases of HPV16 variants (59 and 137 cases of CIN2/CIN3 and cervical cancer, respectively). The difference of proportion test was used to compare patients with CIN2/CIN3 and cervical cancer, based on the prevalent HPV16 lineage (p < 0.05).

Results  According to the histopathological diagnosis, the percentage of lineage frequencies revealed a marginal difference in the prevalence of lineage A in CIN2/CIN3, compared with that in cervical cancer (p = 0.053). For lineage D, the proportion was higher in cancer cases (32.8%), than in CIN2/CIN3 cases (16.9%), with p = 0.023.
Introduction

Cervical cancer is the fourth most common cancer, in terms of incidence and mortality, among women worldwide. In Brazil, it is the third and fourth most common cancer among women, in terms of incidence and mortality, respectively. Several studies have shown that the Human Papillomavirus (HPV) is a predominant, but not the only, factor for cervical cancer development. The viral genotype HPV16 is the most prevalent in cervical cancer and is considered a Group 1 carcinogenic agent for humans by the International Agency for Cancer Research. The study of the genetic diversity of HPV16 has enabled the characterization of specific viral lineages associated with a higher risk of cervical intraepithelial neoplasia (CIN2/CIN3) and cervical cancer development. For HPV16, the lineages were initially named according to their geographical prevalence and separated into five groups: European (EUR), Asian (As), Asian-American (AA), African 1 (Af1), and African 2 (Af2). In 2013, Burk et al. proposed a new numeric nomenclature for all HPV lineages based on the differences in the complete viral genome sequence. The lineages for HPV16 were renamed as follows: A (corresponding to EUR and As), B (corresponding to Af1), C (corresponding to Af2), and D (corresponding to AA).

Studies have shown that HPV16 lineages B, C, and D with higher pathogenicity are associated with a higher viral persistence, compared with the lineage A. Additionally, the HPV16 lineages B, C, and D are also associated with a higher risk of CIN2/CIN3 and cervical cancer. The D lineage is also associated with adenocarcinoma. A longitudinal study has shown that the HPV16 sub-lineages D2 and D3 are more significantly associated with CIN2/CIN3 and cervical cancer, compared with other lineages/sub-lineages. Similar characteristics were reported in lineages of other HPV genotypes, such as HPV 18 and HPV 45.

In this study, we analyzed the genetic diversity of the HPV16 genotype in CIN2/CIN3 and cervical cancer cases in women from the northeast region of Brazil. In this region, cervical cancer was ranked second in terms of cancer incidence, excluding nonmelanoma skin cancer (20.48/100,000), and first in terms of mortality due to cancer (9.52/100,000) among women, in 2020. The present work was part of a multicenter study on the HPV genotypes prior to the introduction of the HPV vaccine in the National Immunization Program, which was performed in two other cities in Brazil (Rio de Janeiro and Belém).

Methods

A cross-sectional multicenter study was performed at two hospitals for cancer treatment in Cidade do Recife, Pernambuco, Brazil, between July 2014 and December 2016.
Women ≥18 years old, with pap smear tests indicative of high-grade intraepithelial lesion (HGL; CIN2/CIN3) or cervical cancer, and who were referred for a colposcopy exam, were invited to participate in the present study. Women who underwent a biopsy and were diagnosed for CIN2/CIN3 or cervical cancer, based on histopathological examination, were included in this analysis. The exclusion criteria included women previously treated for cervical cancer (cancer surgery, radiotherapy, or chemotherapy) or those with cognitive or physical disabilities that could prevent them from answering a questionnaire.

After receiving signed consent, the women were interviewed by trained nurses using an epidemiological questionnaire. We collected data about their socioeconomic status, knowledge about cervical cancer prevention, access to healthcare services for diagnosis and treatment, hormonal and reproductive histories, and tobacco usage. Additional clinical information was obtained from their medical records. The biopsy samples were stored in RNALater until nucleic acid isolation, and were then sent to the research laboratory of the hospital.

DNA was isolated from the biopsy samples using the QIAamp DNA Mini Kit (Qiagen; Cat. Number 51306, North Rhine-Westphalia, Germany). HPV was detected through PCR amplification using the primer sets PGMY07 and PGMY09, whereas reactions without PCR products underwent nested PCR, using the primers GP5+/GP6+. Samples negative for HPV DNA amplification after nested PCR were subjected to PCR for β-globin, and samples positive for β-globin and negative for HPV through nested PCR were considered negative for HPV. For HPV identification, the PCR products were purified using the Illustra GFX PCR DNA and Gel Band Purification Kit (GE Healthcare; Cat. Number 28403471, Buckinghamshire, UK), and were further subjected to DNA sequencing in both directions, using the Big Dye Terminator Cycle Sequencing Ready Reaction V3.1 Kit (Applied Biosystems; Cat. Number 4336919, Texas, USA), as per manufacturer’s instructions, and sequenced in an ABI 3730xL DNA Analyzer (Applied Biosystems, Osaka, Japan). The electropherograms of each sample were checked manually and a consensus sequence of the bidirectional sequencing was subjected to HPV genotype identification using the Blast software (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

Among the 415 samples evaluated for HPV genotyping through DNA sequencing, the 8 most common types were: HPV16 (58.6%), HPV45 (7.2%), HPV18 (7.0%), HPV35 (4.6%), HPV58 (3.6%), HPV31 (3.1%), HPV33 (2.7%), and HPV52 (2.4%). HPV16 was the most common HPV genotype in CIN2/CIN3 (53.9%) and cervical cancer (60.7%). A total of 243 samples contained the HPV16 genotype, and 196 samples (80.7%) were further assessed to identify their lineages.

Samples exhibiting the HPV16 genotypes were subjected to PCR amplification of the viral genomic regions LCR and E6, in two overlapping fragments, as previously described. The resultant sequence, obtained by sequencing the overlapping fragments, was aligned to HPV16 lineage reference sequences suggested by Burk et al. HPV16 lineage identification was performed by detecting a sequence signature (the presence of specific nucleotides at specific sequence positions), as previously described. HPV16 lineage nomenclature used in this study followed those provided by Burk et al.

The epidemiological data were stored using Epinfo 3.5.1 and then linked to both the clinical and HPV DNA sequence data. The final database was analyzed using Stata v.15.0. The chi-squared test (or the Fisher exact test) was used to compare the distribution of the patients based on selected characteristics, according to their histopathological diagnosis. The ratio difference test was used to compare the prevalence of HPV16 viral strains, considering two groups: CIN2/CIN3 and cervical cancer (p < 0.05). The B and C lineages were grouped due to low frequency.

All study procedures were approved by the Institutional Ethics Committees at both hospitals (CAAE 24687713.8.0000.5201 and CAAE 40349014.0.0000.5205).

Results

Selected characteristics of the 196 female patients included in the present study, according to disease status (CIN2/CIN3 versus cervical cancer), are presented in Table 1. We observed that the women diagnosed with CIN2/CIN3 were younger than those diagnosed with cervical cancer. No statistical differences were found between the women with CIN2/CIN3 and cervical cancer for oral contraceptive use or tobacco exposure. The number of childbirths was greater in women with cervical cancer than in those with CIN2/CIN3.

The LCR and E6 regions of each sample with an HPV16 genotype were aligned with HPV16 reference sequences representing the lineages A, B, C, and D, and the presence of specific SNPs was used to identify HPV16 lineages present in each sample. The distribution of the HPV16 lineages is as follows: lineage A, 130 women; lineage B, 1 woman; lineage C, 10 women; and lineage D, 55 women. We observed distinct frequencies of HPV16 lineages between CIN2/CIN3 and cervical cancer, with lineage A being more frequent in CIN2/CIN3 (p = 0.053). Moreover, based on the histopathological diagnosis, the comparison of HPV16 lineage frequencies showed a higher proportion of the D lineage in cervical cancer than that in CIN2/CIN3 (p = 0.023) (Table 2).

Discussion

The lineage A of HPV16 was the most frequent in the samples examined in the present study, for both CIN2/CIN3 and cervical cancer, while lineage D was predominant in cervical cancer samples, suggesting an association between the lineage D of HPV16 and cervical cancer.

The genetic variations in HPV16 can influence the risk of cervical cancer, which vary based on the different lineages or sublineages present in different regions of the world. Previous studies attempted to demonstrate an association between non-European lineages (B/C/D) and a higher risk for developing CIN2/CIN3 and cervical cancer compared with the European lineage (A). A major limitation of
these studies was the grouping of lineages B, C, and D as non-European lineages, which did not allow the differentiation of the carcinogenic potential among these lineages. A recent study on >3,200 women, comparing women without lesions and with CIN1 and women with CIN2 or CIN3 or invasive cancer, have reported an association between the HPV16 sub-lineages D2 and D3 and CIN3 and/or cervical cancer.\(^\text{17}\) Clifford et al. analyzed samples from >7,100 women, comparing those with normal cells, or atypical squamous cells of undetermined significance, or low-grade squamous intraepithelial lesions or CIN1, with those with invasive cervical cancer, and found an association between the HPV16 lineage D and sublineage A4 and cancer diagnosis.\(^\text{8}\)

A recent study on Croatian women confirmed that HPV16, mainly belonging to the European branch, was most frequently associated with histologically confirmed high-grade intraepithelial lesions (CIN2 or CIN3) and cervical cancer.\(^\text{28}\)

In the present study, the lineage A of HPV16 was the most frequently detected lineage, and although it was associated with high-grade lesions (CIN2/CIN3), we could not demonstrate a difference between these high-grade lesions and cervical cancer. Another study conducted in a different region in Brazil also found a high prevalence of the lineage A of HPV16 in samples of high- and low-grade lesions, including cervical cancer.\(^\text{29}\)

In contrast, the D lineage of HPV16 showed a significant association with cervical cancer, compared with high-grade lesions, suggesting that the lineage D of HPV16 is involved in the progression of HPV16 infection to cervical cancer. Although this is a cross-sectional study, the present study provides novel data on the association between specific lineages of HPV16 and cervical cancer. Our data suggest that the lineage D could be more aggressive with respect to the progression of high-grade lesions into cancer, without

### Table 1 Distribution of female patients based on selected characteristics, according to their histopathological diagnosis

<table>
<thead>
<tr>
<th>Variable</th>
<th>CIN2/CIN3</th>
<th>Cervical cancer</th>
<th>p-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n = 59</td>
<td>%</td>
<td>n = 137</td>
</tr>
<tr>
<td>Age (years old)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20–39</td>
<td>37</td>
<td>62.7</td>
<td>43</td>
</tr>
<tr>
<td>40–49</td>
<td>16</td>
<td>27.1</td>
<td>40</td>
</tr>
<tr>
<td>50–59</td>
<td>6</td>
<td>10.2</td>
<td>23</td>
</tr>
<tr>
<td>≥ 60</td>
<td>0</td>
<td>0.0</td>
<td>31</td>
</tr>
<tr>
<td>Number of childbirths</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>3</td>
<td>5.1</td>
<td>3</td>
</tr>
<tr>
<td>1–2</td>
<td>28</td>
<td>47.5</td>
<td>43</td>
</tr>
<tr>
<td>3–4</td>
<td>17</td>
<td>28.8</td>
<td>44</td>
</tr>
<tr>
<td>5–6</td>
<td>5</td>
<td>8.5</td>
<td>14</td>
</tr>
<tr>
<td>≥ 7</td>
<td>6</td>
<td>10.1</td>
<td>33</td>
</tr>
<tr>
<td>Oral contraceptive use</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>42</td>
<td>71.2</td>
<td>88</td>
</tr>
<tr>
<td>No</td>
<td>17</td>
<td>28.8</td>
<td>49</td>
</tr>
<tr>
<td>Tobacco exposure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>26</td>
<td>44.1</td>
<td>76</td>
</tr>
<tr>
<td>No</td>
<td>33</td>
<td>55.9</td>
<td>61</td>
</tr>
</tbody>
</table>

*Test for proportion difference among patients with CIN2/CIN3 and cervical cancer.

### Table 2 Distribution of female patients, according to HPV16 lineage frequencies and their histopathological diagnosis

<table>
<thead>
<tr>
<th>HPV16 lineages</th>
<th>Histopathological diagnosis</th>
<th>p-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>CIN2/CIN3</td>
</tr>
<tr>
<td></td>
<td>n (%)</td>
<td>n (%)</td>
</tr>
<tr>
<td>A</td>
<td>130 (66.3)</td>
<td>45</td>
</tr>
<tr>
<td>B/C</td>
<td>11 (5.6)</td>
<td>4</td>
</tr>
<tr>
<td>D</td>
<td>55 (28.1)</td>
<td>10</td>
</tr>
</tbody>
</table>

Abbreviation: CIN, cervical intraepithelial neoplasia.

* Chi-squared test or Fisher exact test.
restricting the association from linking normal/low grade lesions and invasive cancer.

**Conclusion**

The A lineage of HPV16 was the most frequent in both the CIN2/CIN3 and cervical cancer samples, whereas the lineage D of HPV16 was predominant in cervical cancer, suggesting an association between HPV16 lineage D and cervical cancer.

**Contributions**


**Conflict of Interests**

The authors have no conflict of interest to declare.

**Acknowledgments**

We would like to thank Antônio Maria S. Negrão, Neile A. de Carvalho, Evaneide A. de Morais, and all nurses for their contribution to the execution of the study. Moreover, we would like to thank the people of Pernambuco Cancer Hospital and Instituto de Medicina Integral Prof. Fernando Figueira (IMIP) for their support during the study. We also would like to thank Editage (www.editage.com) for English language editing. The present study was financially supported by the following grants: OPAS/INCA, CEPESC: no. BR/LOA/1200085–001; Instituto Nacional de Ciência e Tecnologia para Controle do Câncer (INCT-Cancer Control); Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq: grants: 573806/2008–0 and 304339/2018–0); Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ, grants: E26/170.026/2008 and 200.928/2021).

**References**


