Increase of Integration Events and Infection Loads of Human Papillomavirus Type 52 with Lesion Severity from Low-Grade Cervical Lesion to Invasive Cancer


Departments of Microbiology, Obstetrics, and Gynaecology and School of Public Health, The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, New Territories, Hong Kong Special Administrative Region, People’s Republic of China

Infection load and the integration of human papillomaviruses (HPV) have been implicated as determinants for oncogenesis, but whether variation among different HPV types exists remains unclear. We investigated 91 women infected with HPV type 52 (HPV-52), a type that is rare worldwide but common in East Asia. The median viral load increased with the severity of the lesion (248 copies/cell equivalent for normal/cervical intraepithelial neoplasia [CIN] grade 1, 402 copies/cell equivalent for CIN 2, 523 copies/cell equivalent for CIN 3, and 1,435 copies/cell equivalent for invasive cancer). The proportion of specimens with integration increased significantly with the severity of the lesion (P < 0.001). The viral load was associated with the physical status of the viral genome, with higher levels for the pure episomal form (P = 0.001). Infection status should be considered when interpreting viral load data for HPV-52, as single infections with this HPV type were found to have marginally higher viral loads than coinfections (P = 0.051). All except one sample had E2 disruption restricted to only a part of the gene. Integration is a critical step in HPV-52-induced carcinogenesis. The profile of E2 disruption was different from that described for HPV-16, with the amino-terminal region being most frequently involved. Selecting the appropriate E2 region for amplification is critical in studying the integration of HPV-52. In summary, the HPV-52 viral load and the integrated proportion increased with the severity of the cervical lesions but had a different pattern than that of HPV-16.

Infection with human papillomavirus (HPV) is known to play a pivotal role in the development of cervical cancer (4, 37, 41). Some previous studies have suggested that not the presence of viral DNA per se but the viral load of HPV is an important determinant for the development of high-grade lesions (18, 25, 27, 34, 35, 36, 40). However, a substantial number of studies did not observe a correlation between viral load and the severity of cervical disease (5, 8, 26, 29, 38). The techniques that have been used to measure HPV infection load vary across studies. Some studies measured the overall viral load across a panel of high-risk HPV types without considering the specific type of HPV actually harbored in the specimen. Although coinfections with multiple types are commonly found throughout different grades of cervical lesion, previous studies often did not consider single infection or coinfection status. It is now recognized that the measurement and interpretation of the viral load in HPV infection are far more complex than previously perceived (39).

The HPV genome can exist in two physical states: it can either be a circular episomal form or integrated into the human genome (10). The integration of viral DNA is considered to result in transformation of the infected cells, as integration usually disrupts the E2 gene whose normal function is to down-regulate the expression of the E6 and E7 viral oncoproteins (11, 16, 17, 33). It has therefore been proposed that viral integration could be an adjunctive diagnostic or prognostic marker (19, 21). Studies based on HPV type 18 (HPV-18)-positive clinical specimens showed that integration is virtually completed in high-grade lesions and invasive cancers (3, 30). However, the results obtained from HPV-16-positive specimens are more diverse (5, 8, 20). Furthermore, it has recently been demonstrated that the integration of the HPV-16 genome does not invariably result in high levels of oncogene E6/E7 expression (12).

Infection load and viral integration are interrelated, as the replication of HPV can only occur in the presence of the episomal form of the genome. At present, only a few studies have examined these two events in the same cohort of subjects (1, 13, 20, 31). Furthermore, most studies have targeted the type that is the most common worldwide, HPV-16, whereas data on other HPV types are very limited (6, 13, 32). Whether the association of infection load and viral integration with cervical cancer development varies with the type of HPV infection remains unknown. HPV-52, though generally regarded as an uncommon type, has been found to exist with an unusually high prevalence in a few Chinese populations, including the population in Hong Kong (7, 9, 15, 22, 23, 24). In some of these areas, HPV-52 was the second- or third-most-common HPV type found in cervical cancer (7, 13, 15, 24). We took advantage of this special epidemiological feature in our local population to examine the two key viral parameters, infection...
load and viral genome integration, of HPV-52 infection in a cohort of southern Chinese women with various degrees of cervical lesions.

MATERIALS AND METHODS

Study subjects. This study was conducted in Hong Kong, where cervical cancer ranks ninth for new cancers in women, with an age-standardized incidence of 7.8 per 100,000 (14). The cervical samples for this study were obtained from women attending the colposcopy clinic at the Prince of Wales Hospital. Women with normal colposcopic findings were classified into the “normal” group provided that their subsequent 6-month follow-up cervical cytology results did not reveal abnormalities. The cervical disease status of the specimens was defined histologically and classified as cervical intraepithelial neoplasia (CIN) grade 1, 2, or 3 or invasive cervical cancer (ICC). Pregnant or immunocompromised women or those with a previous history of cervical disease or surgery were excluded. Consent was obtained from all study subjects, and the study was approved by the institutional ethics committee.

HPV-52 detection and genotyping. Total DNA was extracted from cervical scrape samples by using a QIAamp DNA mini kit (Qiagen). HPV was genotyped by using a linear array HPV genotyping test (Roche Molecular Systems, Inc., CA) which can detect 37 HPV types. Samples found to contain HPV-52 either as a single infection or as a coinfection with other HPV types were subjected to further examination for viral load and E2 gene disruption status.

Viral load, host gene quantitation, and integration. The total (episomal plus integrated) crude viral load was determined by real-time PCR targeting the E7 gene of HPV-52. The E7 gene was used as it is retained both in the episomal and integrated forms. To account for variations in the amount of cells collected in each sample, the housekeeping gene beta-actin was also quantified. The normalized viral load was obtained with the equation (E7_all/beta-actin_all) × 2 and expressed as viral copies/cell equivalent.

The disruption in the E2 gene was inferred by comparing the levels of the E2 and E7 gene copies in the viral load. E2 gene disruption was taken as a maker of HPV integration. To account for the possibility that disruption might involve only part of the E2 gene, four sets of nonoverlapping real-time PCR primers were designed to encompass the whole E2 gene, with the first two sets covering the amino (N)-terminal region and the other two sets targeting, respectively, the hinge (H) region and the carboxy (C)-terminal region. When any of the E2 gene real-time PCR assays revealed an undetectable gene copy number, the specimen was regarded as harboring the pure episomal form.

The six real-time PCRs targeting HPV-52 E7: the E2 C-terminal, E2 H, E2 N-terminal 1, and E2 N-terminal 2 regions; and beta-actin were performed by using power SYBR green PCR master mix (Applied Biosystems, Foster City, CA) which can detect 37 HPV types. Samples found to contain HPV-52 either as a single infection or as a coinfection with other HPV types were subjected to further examination for viral load and E2 gene disruption status.

RESULTS

Ninety-one women with HPV-52 infections were included in this study. Their ages ranged from 22 to 88 (mean, 50.0; standard deviation, 15.2) years. Twelve women had normal cervixes, 12 had CIN 1, 20 had CIN 2, 27 had CIN 3, and 20 had ICC. All invasive cancers were of the squamous-cell type, except for one case of adenocarcinoma.

Viral load. The crude viral load of the 91 specimens spread over a wide range of from 644 to 9 × 10^8 copies/μl (median, 1,669,812; interquartile range, 267,937 to 8,384,753) (Table 2). When normalized to the beta-actin DNA level, the viral load ranged from 0.03 to 89,344 copies/cell equivalent (median, 462; interquartile range, 41 to 1,564). The distribution of viral loads and beta-actin levels with respect to the degree of cervical lesion is shown in Fig. 1. The normalized median viral load appeared to increase with the severity of the cervical lesion (248 copies/cell equivalent for normal/CIN 1, 402 copies/cell equivalent for CIN 2, 523 copies/cell equivalent for CIN 3, and 1,435 copies/cell equivalent for ICC) (Table 2). However, the interquartile ranges overlapped, and the differences were not statistically significant (P = 0.308 by Kruskal-Wallis test).

Viral genome status. Of the 91 specimens examined, 40 (44.0%) did not show any disruption of the E2 gene. These specimens were regarded as harboring the pure episomal form.

### Table 1. Primer sets for real-time PCRs

<table>
<thead>
<tr>
<th>Amplification target</th>
<th>Primer sequences (nt*)</th>
<th>Annealing temp (°C)</th>
<th>Concn (μM)</th>
<th>Product size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E2 N-terminal region 1</td>
<td>5'-GATACCGGCACTTAAATG-3' (2751–2772) and 5'-CCTTCGCTTAAAGACACCTGCA-3' (2937–2964)</td>
<td>58</td>
<td>0.75</td>
<td>196</td>
</tr>
<tr>
<td>E2 N-terminal region 2</td>
<td>5'-AGTCTAGAAATGCGCGTGC-3' (3034–3053) and 5'-CAATTGCATCTACACTCC-3' (3163–3184)</td>
<td>58</td>
<td>1.0</td>
<td>151</td>
</tr>
<tr>
<td>E2 H region</td>
<td>5'-GTCACACTGCGACGGAAC-3' (3373–3392) and 5'-TCAGTCGACGTACGACCTGC-3' (3544–3563)</td>
<td>58</td>
<td>0.75</td>
<td>191</td>
</tr>
<tr>
<td>E2 C-terminal region</td>
<td>5'-GTTCAATTCTCCTACCTGGC-3' (3685–3706) and 5'-GACATGACACCTTGATATAACCTGC-3' (3819–3843)</td>
<td>58</td>
<td>1.0</td>
<td>158</td>
</tr>
<tr>
<td>E7</td>
<td>5'-GCAACCTGAACACATGACCT-3' (597–617) and 5'-TCCTCGCCATGCTATGAAT-3' (763–782)</td>
<td>58</td>
<td>0.75</td>
<td>186</td>
</tr>
<tr>
<td>Beta-actin</td>
<td>5'-GACCGGATCTGCTACCAACT-3' (1283–1303) and 5'-CACTTTCTGGGCCGCT-3' (1404–1424)</td>
<td>0.125</td>
<td>0.125</td>
<td>142</td>
</tr>
</tbody>
</table>

* nt, nucleotide position numbering according to the HPV52 reference strain (GenBank accession no. X74481).
The proportion of specimens harboring integrated genomes in either the pure or concomitant form increased significantly with the increase in severity of the cervical lesion (chi-squared value for linear trend of proportions, 4.2; *P* = 0.001) (Fig. 2). The pure integrated form was found in 25.0% of ICC cases, as well as in a similar proportion (20.8%) of normal/CIN 1 cases. The proportion of the pure episomal form in ICC cases, as well as in a similar proportion (20.8%) of normal/CIN 1 cases. The proportion of the pure episomal form in ICC cases (10.0%) was significantly lower than its proportion in the other groups (44.4% versus 16.7 to 22.2%) (Table 2). All 12 of these specimens suggested the presence of multiple E2 gene disruption patterns. These multiple disruption patterns were more commonly found in normal/CIN 1 cases than in the other groups (44.4% versus 16.7 to 22.2%) (Table 2). All 12 of these specimens contained a disruption that occurred across the N-terminal 1 and N-terminal 2 regions and which coexisted with a disruption that involved only the N-terminal 1, N-terminal 2, or C-terminal region. Overall, the N-terminal 1 and N-terminal 2 regions were the two most frequently disrupted regions, with each found to be disrupted in 76.5% (39/51) of specimens harboring integrated (pure or concomitant forms) viral genomes, whereas the less-involved C-terminal and H regions were found to be disrupted in 17.6% and 5.9% of specimens (Table 2).

The levels of normalized viral load were significantly different among specimens harboring viral genomes with different physical statuses (*P* value of 0.001 by Kruskal-Wallis test) (Fig. 3a). Specimens harboring the pure episomal form had higher viral loads with a median (interquartile range) of 877 (281 to 3,208) copies/cell equivalent, whereas those harboring the pure integrated form had a lower viral load of 33 (3 to 260) copies/cell equivalent.

**Single infection and coinfection.** Of the 91 subjects, 53 had single infections, with HPV-52 alone. The other 38 had coinfections with HPV-52 and other HPV type(s) (Table 2). The most common coinfection was with HPV-16 (12/38; 31.6%), followed by HPV-18 (18.4%) and HPV-58 and HPV-66 (5.3% each). Four cases had triple infections, three with HPV-16 and -18 and one with HPV-18 and -33. Other coinfecting types, with each found in one case, included HPV-33, -51, -53, -54, -55, -56, -61, -69, -81, -CP6108, and -MM8. The viral loads in specimens with single infections were marginally higher than in those with coinfections (median [interquartile range] of 798 [93 to 3,191] versus 255 [36 to 1,383] copies/cell equivalent; *P* value of 0.051 by Mann-Whitney U test) (Fig. 3B). When the physical status of the viral genome was correlated with the infection status of single infection or coinfection, no significant association was observed.

**E2 disruption patterns.** The E2 gene copy number profiles of 12 specimens suggested the presence of multiple E2 gene disruption patterns. These multiple disruption patterns were more commonly found in normal/CIN 1 cases than in the other groups (44.4% versus 16.7 to 22.2%) (Table 2). All 12 of these specimens contained a disruption that occurred across the N-terminal 1 and N-terminal 2 regions and which coexisted with a disruption that involved only the N-terminal 1, N-terminal 2, or C-terminal region. Overall, the N-terminal 1 and N-terminal 2 regions were the two most frequently disrupted regions, with each found to be disrupted in 76.5% (39/51) of specimens harboring integrated (pure or concomitant forms) viral genomes, whereas the less-involved C-terminal and H regions were found to be disrupted in 17.6% and 5.9% of specimens (Table 2).

The proportion of specimens with disruption involving the N-terminal 1 region increased significantly from CIN 2 to ICC (chi-squared value for linear trend of proportions, 4.2; *P* =

---

**TABLE 2. Viral load and physical status of the viral genome in HPV-52 infection according to the degree of cervical lesion**

<table>
<thead>
<tr>
<th>Viral characteristics</th>
<th>Cervical pathology</th>
<th>Normal/CIN 1 (n = 24)</th>
<th>CIN 2 (n = 20)</th>
<th>CIN 3 (n = 27)</th>
<th>ICC (n = 20)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median viral load</td>
<td>(interquartile range)</td>
<td>777,734 (129,962–3,744,271)</td>
<td>712,453 (166,251–3,995,568)</td>
<td>2,817,698 (525,172–9,348,353)</td>
<td>3,325,361 (87,655–100,000,000)</td>
</tr>
<tr>
<td>Crude total viral load (copies/µl)</td>
<td>11,909 (1,827–51,890)</td>
<td>6,564 (1,461–18,151)</td>
<td>18,008 (7,898–62,012)</td>
<td>523 (93–1,410)</td>
<td>29,201 (6,341–53,428)</td>
</tr>
<tr>
<td>Normalized viral load (copies/cell equivalent)</td>
<td>248 (38–1,051)</td>
<td>402 (45–3,318)</td>
<td>523 (93–1,410)</td>
<td>1,435 (86–4,980)</td>
<td></td>
</tr>
<tr>
<td>Viral genome physical status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pure episomal form</td>
<td>15 (62.5)</td>
<td>11 (55.0)</td>
<td>12 (44.4)</td>
<td>2 (10.0)</td>
<td></td>
</tr>
<tr>
<td>Pure integrated form</td>
<td>5 (20.8)</td>
<td>2 (10.0)</td>
<td>2 (7.4)</td>
<td>5 (25.0)</td>
<td></td>
</tr>
<tr>
<td>Concomitant forms</td>
<td>4 (16.8)</td>
<td>7 (35.0)</td>
<td>13 (48.1)</td>
<td>13 (65.0)</td>
<td></td>
</tr>
<tr>
<td>HPV-52 infection status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single infection</td>
<td>15 (62.5)</td>
<td>13 (65.0)</td>
<td>16 (59.3)</td>
<td>9 (45.0)</td>
<td></td>
</tr>
<tr>
<td>Coinfection with other HPV type(s)</td>
<td>9 (37.5)</td>
<td>7 (35.0)</td>
<td>11 (40.7)</td>
<td>11 (55.0)</td>
<td></td>
</tr>
<tr>
<td>E2 gene region involved in disruption (no. [%])</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N-terminal region 1</td>
<td>7 (77.8)</td>
<td>4 (44.4)</td>
<td>10 (66.7)</td>
<td>18 (100.0)</td>
<td></td>
</tr>
<tr>
<td>N-terminal region 2</td>
<td>9 (100.0)</td>
<td>7 (77.8)</td>
<td>13 (86.7)</td>
<td>10 (55.6)</td>
<td></td>
</tr>
<tr>
<td>H region</td>
<td>2 (22.2)</td>
<td>0 (0.0)</td>
<td>2 (13.3)</td>
<td>2 (10.0)</td>
<td></td>
</tr>
<tr>
<td>C-terminal region</td>
<td>4 (44.4)</td>
<td>2 (22.2)</td>
<td>2 (13.3)</td>
<td>2 (10.0)</td>
<td></td>
</tr>
<tr>
<td>N-terminal region 1 or 2</td>
<td>9 (100.0)</td>
<td>9 (100.0)</td>
<td>15 (100.0)</td>
<td>18 (100.0)</td>
<td></td>
</tr>
<tr>
<td>Whole E2 gene</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>1 (5.0)</td>
<td></td>
</tr>
<tr>
<td>Pattern of E2 disruption (no. [%])</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>5 (55.6)</td>
<td>7 (77.8)</td>
<td>12 (80.0)</td>
<td>15 (83.3)</td>
<td></td>
</tr>
<tr>
<td>Multiple</td>
<td>4 (44.4)</td>
<td>2 (22.2)</td>
<td>3 (20.0)</td>
<td>3 (16.7)</td>
<td></td>
</tr>
</tbody>
</table>

* The number of specimens with integration was the denominator. Twelve specimens showed mixed patterns of disruption.
However, the proportion was paradoxically high for the normal/CIN 1 group (77.8%) (Fig. 4).

The proportion of specimens with disruption involving the C-terminal region appeared to decrease with an increase in cervical disease severity. However, the trend was not significant when subjected to statistical analysis (chi-squared value for linear trend, 1.8; \( P = 0.18 \)).

**DISCUSSION**

While at least 15 HPV types are known to cause cervical cancer (28), the model of oncogenesis for HPV is mainly established from studies of HPV-16 and HPV-18 (41). Whether these observations and hypotheses can be generalized to other high-risk HPV types remains to be verified. Studies of the viral
characteristics of non-16/-18 HPV types and their correlation with the progression of infected cervical lesions are very limited. The current study provides data on an HPV type, HPV-52, that is associated with a substantial proportion of cervical cancers in East Asia.

Several studies have reported a positive association between the viral load of HPV-16 and current cervical disease status or prognosis (18, 25, 27, 34, 35, 36, 40). At present, there has been only one study reporting the viral load of HPV-52, which showed a higher viral load for invasive cancers than for CIN 2/3 (13). Results from the current study also suggested that, though not reaching statistical significance, the viral load of HPV-52 increased with the degree of cervical lesion. However, it should be noted that the wide range of variation and the overlaps in the viral load levels between normal and high-grade lesions do not favor a clinical role for viral load determination for HPV-52. We also observed that the interpretation of viral load can be complicated by its association with the viral genome status (episomal or integrated) and infection status (single or coinfection). Our finding that specimens harboring the

![FIG. 3. Distribution of viral load according to infection characteristics. (a) Physical status of viral genome. (b) Infection status. The middle line indicates the median. The box represents the interquartile range. The lines extending from each box represent the upper and lower limit. The normalized viral load was calculated by the equation \((E7_{\text{copy}}/\beta\text{-actin}_{\text{copy}}) \times 2\).](#)

![FIG. 4. E2 gene disruption involvement according to the degree of cervical lesion.](#)
episomal form of the HPV-52 genome were associated with a higher viral load can be explained, at least partly, by the fact that only the episomal form of the viral genome is capable of vegetative replication. On the other hand, the reason for the observed tendency for a higher viral load among cases of single infection than among cases of coinfection is not as obvious. One hypothesis could be that in the presence of other HPV types during a coinfection, HPV-52 becomes less successful in maintaining the infection and, thus, might be less transmissible. This raises a question about the possible type replacement effect following the widespread use of vaccines targeting HPV-16, but not HPV-52. Nonetheless, it should be noted that the viral load levels in cases of single infection with HPV-52 were only marginally higher than those in cases of coinfection. More studies are needed to confirm the current observation.

Studies of HPV-16 have repeatedly shown that clinical samples, regardless of lesion grade, often contain a mixture of the episomal and integrated forms of the viral genome (1, 8, 20). Our study also revealed this phenomenon for HPV-52. Given this feature, techniques employed for determining integration status should be capable of detecting integration over the episomal genome background. Real-time PCR comparing the levels of the E2 and E6 or E7 genes as applied in this study is a commonly used method to achieve this purpose. However, real-time PCR has a technical limitation in that the target amplification product should be around 100 to 150 bp in length to provide an efficient amplification. Our results showed that only one specimen had disruption involving the whole E2 gene, implying that studies using amplification targets not covering the whole E2 gene could have underestimated the integration of HPV-52 (13). Our results showed that the N-terminal region of E2 was most frequently involved in disruption and, thus, the most sensitive target for identifying integration of HPV-52. In contrast to the results of a report on HPV-16 (2), the H region is rarely disrupted in HPV-52. These results suggest that the E2 gene disruption pattern of HPV could be type specific.

The concept that viral integration is a prerequisite for the oncogenic progression of an HPV-infected cell has been challenged by a few recent studies using HPV-16 as the study model (5, 8, 26, 29). It has been shown that the integration of HPV-16 does not invariably result in high levels of E6/E7 expression (12). The only such study of HPV-52, which was reported by Ho et al. (13), used the H region of E2 as the amplification target to detect integration in specimens from 13 invasive cancers and 18 CIN 2/3 cases. It was found that 92% of CIN 2/3 cases but only 25% of cancers harbored integrated genomes (either as the pure integrated or concomitant form). In contrast, our data show a significant correlation between integrated viral genomes and lesion severity. Although we observed two invasive cancers harboring the pure episomal form of the viral genome, the overall picture indicates that viral integration is still a key event leading to the oncogenic progression of an HPV-52-infected cervical lesion. Furthermore, our results show that the integration event of HPV-52 occurred early in the course of infection and might be one of the key initiators of oncogenetic progression.

The two invasive cancers that harbored the pure episomal form of the viral genome were both instances of HPV-52 single infection, without the involvement of other HPV types. Although we cannot exclude the possibility that viral genome disruption had occurred outside the E2 region, it is worthwhile to investigate whether events such as methylation had bypassed viral integration in the development of invasive cancers in these patients.

The correlation of infection load and integration of HPV with oncogenic progression and its clinical usefulness may vary with the type of HPV in question. More studies on HPV types other than HPV-16 are warranted.

ACKNOWLEDGMENTS

The work described in this paper was fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, People’s Republic of China (project no. CUHK4429/03 M). The authors declare no conflict of interest.

REFERENCES


