Physical state and copy numbers of HPV16 in oral asymptomatic infections that persisted or cleared during the 6-year follow-up

Adriana Lorenzi,1,2,3 Jaana Rautava,1,4 Katja Kero,5 Kari Syrjänen,6 Adhemar Longatto-Filho,2,7 Seija Grenman5 and Stina Syrjänen1,4,*

Abstract
Persistent human papillomavirus (HPV) infection is a key event in HPV-induced carcinogenesis. As part of the prospective Finnish Family HPV Study, we analysed the physical state and viral copy numbers of HPV16 in asymptomatic oral infections that either persisted or cleared during the 6-year follow-up. The persister group comprised 14 women and 7 men with 51 and 21 HPV16-positive brush samples. The clearance group included 41 women and 13 men, with 64 and 24 samples, respectively. Physical state and viral DNA load were assessed by using quantitative PCR for HPV16 E2 and E6 genes. E2/E6 ratio was calculated and HPV16 was classified as episomal, mixed or integrated with values of 0.93–1.08, <0.93 and 0, respectively. In both genders, the physical state of HPV16 was significantly different between the cases and controls (P<0.001). HPV16 was episomal in all men and 66% (27/41) of women who cleared their infection. HPV16 was mixed and/or integrated in 71% and 57% of the women and men persisters, respectively. The mean HPV16 copy number per 50 ng genomic DNA was nearly 5.5-fold higher in the women than in the men clearance group (P=0.011). Only in men, HPV16 copy numbers were higher in persisters than in the clearance group (P=0.039). To conclude, in both genders, persistent oral HPV16 infections were associated with the mixed or integrated form of HPV16, while in the clearance groups, episomal HPV16 predominated. This indicates that HPV16 integration is a common event even in asymptomatic oral infections, which might predispose the infected subjects to progressive disease.

INTRODUCTION
The incidence of head and neck squamous cell carcinoma (HNSCC) has increased in the Western world since the late 1960s. Smoking and alcohol consumption remain the key risk factors, but increased incidence of HNSCCs has been linked in part to human papillomavirus (HPV) infections, particularly those of HPV16 [1–4]. HPV involvement in HNSCC was first presented in 1983, based on HPV-suggesting morphological changes and expression of HPV structural antigens [5]. Soon, the presence of HPV DNA in HNSCC samples was confirmed by Southern blot hybridization [6] and localized in cancer cell nuclei by in situ hybridization [7]. Since then, thousands of HNSCC cases have been analysed for the presence of HPV DNA or mRNA using different molecular techniques. The results have been summarized in several meta-analyses; one of the most frequently cited showing pooled estimates for HPV DNA prevalence of 46% (95% CI 38.9–52.9) in cancers of the oropharynx, 22% (16.4–28.3) in those of the larynx and 24% (18.7–30.2) in oral cancer [2]. HPV prevalence of HNSCC is globally variable, but HPV16 is the most prevalent genotype worldwide.

Only a few natural history studies are available on HPV infection in head and neck region, with the main focus in oral cavity and/or oropharynx [8–11]. From these studies, it is clear that asymptomatic HPV infections exist also in oral mucosa. Most of these asymptomatic infections clear spontaneously, but a minor subgroup remain persistent similarly
to in the genital tract [11, 12]. However, a firm documentation of the progressive potential of such persistent oral/oropharyngeal HPV infections is still lacking.

Integration of viral episomes into human chromosomes is a common phenomenon in cervical carcinogenesis [13–15]. During HPV integration, ORFs E2 and E1 are frequently disrupted or deleted. Given that E2 controls the expression of E6 and E7 oncogenes during a normal viral life cycle, deletion of E2 will result in uncontrolled expression of E6 and E7, leading to genomic instability and eventually malignant transition [14, 16, 17]. Although well established in cervical carcinogenesis, opinions are still divided as to whether HPV integration is an early or a late event in this process [15, 16, 18]. Recent wide genome analyses have shown that HPV integrations can occur throughout the human genome in both HNSCC and cervical cancer, and the viral genome can be disrupted also elsewhere than in the E1–E2 region [19–22].

To cast further light on HPV integration in asymptomatic oral infections, we analysed the frequency of HPV16 integration in oral HPV infections that either (1) persisted or (2) cleared during a prospective follow-up (FU) of 6 years as part of the Finnish Family HPV Study. We also assessed the potential differences between the genders because HPV-associated HNSCCs are more prevalent among males.

RESULTS

Clinical characteristics of the subjects

In total, 55 women and 38 men with 115 and 45 oral samples collected during the FU were included as summarized in Table 1. We analysed the demographic data including age, sexual habits, smoking, use of alcohol, sexually transmitted diseases and contraception to identify predictors for oral HPV16 persistence and clearance in women and men, separately. The women persisters were significantly younger (mean 24.5±3.4 years) than the women who cleared oral HPV16 infection (27.3±3.9 years; P=0.021) while in men no such difference was found (26.6±3.05 years and controls 27.8±5.0 years, P=0.574). Interestingly, the use of hormonal intrauterine device protected against oral HPV16 persistence (OR 1.341, 95% CI 1.040–1.730). No other predictors were found as the number of subjects were so limited.

Physical state of HPV16 in oral infections

Table 2 summarizes the physical state and the copy numbers of HPV16 E6 in asymptomatic oral HPV infections. None of the female persisters had only episomal form of HPV16, as the physical state detected was either in mixed (7/14) or purely integrated (3/14) form. In male persisters, HPV16 integration was the most prevalent state (4/7). All 13 men who cleared their oral HPV16 infection had only episomal HPV16, while 66% (27/41) of women with episomal form cleared their HPV. During the FU, the physical state of HPV16 changed in three female persisters, from episomal to mixed (one case) or from mixed to integrated (two women), and in one male persister from episomal to integrated. Interestingly, no HPV16 E6 but only E2 was detected in two women and in one man.

Copy number of HPV16 in oral infections

In women, the mean copy number of HPV16 in asymptomatic oral infections that cleared was 639 copies per 50 ng human DNA (approximately 76.6 copies per 1000 cells), which is nearly 5.5-fold higher than in their male counterparts (P=0.011) (Fig. 1). With persistent oral HPV16 infection, both genders had similar copy numbers. Male persisters had higher HPV16 E6 copy numbers than males who cleared their oral infection (P=0.039), whereas no such difference was found among women. Fig. 2 shows the physical state and copy numbers of oral HPV16 according to the gender and the outcome of the infection. The copy numbers of purely integrated HPV16 were significantly lower in women than in male persisters (P=0.022), the contrary being true for episomal HPV among those who cleared (P=0.011).

Fig. 3 summarizes the dynamics of HPV16 copy numbers over time. Interestingly, women who eventually cleared (a) their oral HPV infection had a higher peak in HPV16 copy numbers at the baseline visit when pregnant. The copy numbers declined dramatically after delivery, as evident already at the second visit. The cleared oral HPV16 infection in men (c) was characterized by highest copy numbers detected at the second (2-month) visit and subsequent decline after the third visit. Most of the women with persistent HPV16 infection had highest viral loads also at the second FU visit (b), followed by progressive decline during the subsequent visits. In the male persisters (d), HPV16 copy numbers declined after the third visit, followed by progressive decline during the subsequent visits. In the male persisters (d), HPV16 copy numbers declined after the third visit, followed by progressive decline during the subsequent visits.

Table 1. Number of subjects and their samples during the 6-year follow-up according to visit, gender and outcome of oral HPV16 infection

<table>
<thead>
<tr>
<th>Gender</th>
<th>Outcome of oral HPV16 infection</th>
<th>No. of samples</th>
<th>Follow-up (months)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Baseline</td>
</tr>
<tr>
<td>Women, n=55</td>
<td>Persisters (n=14)</td>
<td>51</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Clearance group (n=41)</td>
<td>64</td>
<td>15</td>
</tr>
<tr>
<td>Men, n=38</td>
<td>Persisters (n=7)</td>
<td>21</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Clearance group (n=13)</td>
<td>24</td>
<td>3</td>
</tr>
</tbody>
</table>

–, non applicable.
numbers fluctuated although an increase was seen between 12- and 36-month visits.

**DISCUSSION**

Physical state of HPV16 in HPV-associated precancer lesions and cancers, especially in the uterine cervix, has been widely studied. To the best of our knowledge, our study is the first to describe viral load and physical state of HPV16 in oral asymptomatic infections that either persisted or cleared during the 6-year FU. HPV persistence is a key event in HPV-induced carcinogenesis [23–26]. In 1985, it was shown that cervical HPV infection with genotypes 16 and 18 progressed from asymptomatic infection to cervical intraepithelial neoplasia (CIN) and finally to invasive cancer in less than 3 years. The patient did not accept any treatment due to religious reasons [23].

Accordingly, we expected here to find integrated or mixed state in persistent oral HPV16 infections. As hypothesized, HPV16 was found to be integrated in all malepersisters, except one with episomal HPV16. None of the female persisters had purely episomal form of HPV16, mixed form being the most predominant state. Men who cleared their asymptomatic oral HPV16 infection had only episomal HPV16. Although the episomal state was prevalent (66 %, 27/41) also in women who cleared, there were cases of HPV16 with either a mixed or an integrated state. Our results indicate that HPV integration is an early event in HPV16 persistence. Thus, the subjects with persistent oral HPV16 might be at higher risk for HPV-associated diseases, too. As there are no previous data on physical state of HPV16 in oral infections, we extrapolate our results in the context of similar data available on genital HPV infections. In addition, the results are explicated by earlier in vitro results, which have shown that the cells with integrated virus will expand and those with episomal virus are gradually depleted, which change is detectable also in the morphology and growth kinetics of the cells.
Most studies agree that the HPV genome is episomal in the majority of benign lesions and even in low-risk squamous intraepithelial lesions [27, 28], while HPV integration is frequent in cervical cancers and cell lines established from them. However, only HPV episomes have also been reported in some of the cervical cancers [29]. Integration of HPV16 has been regarded earlier as a late event in cervical carcinogenesis as the methods used were based on fusion transcripts [24]. This view was revisited in the mid-2000s by our in vitro studies, which showed convincing evidence of an early integration and loss of episomes in an HPV33-positive cell line (UT-DEC-1), providing the cells with a growth advantage [16, 30]. Our finding was confirmed by Pett et al. [17], who re-analysed those W12 cell lines, which were previously characterized expressing only episomal HPV16. They also showed that cells containing integrated HPV16 reproducibly emerged during long-term culture when there had been a rapid fall in episome numbers [17]. Thus, their conclusion was similar to ours: cervical carcinogenesis requires not only high-risk HPV integration but also loss of inhibitory episomes [16, 17, 30]. In 2006, we were able to confirm our hypothesis also in a clinical setting [31]. In a long-term FU study on cervical HPV infections, a mixed pattern was the most prevalent physical state of HPV16, sometimes seen in cases with normal Pap smear. Purely integrated state was the most prevalent one in high-grade CIN and cancers, but also detectable in some low-grade abnormalities [31]. Also, Dutta and coworkers identified integrated HPV in 27% of asymptomatic cervical infections, albeit the episomal form predominated [32]. Their conclusion was that integration of HPV16 along with changes in methylation pattern of early and late promoters is essential for neoplastic transformation of asymptomatic cervical infections.

In the present study, we identified changes in the physical state of HPV16 during the FU in some of the women. This could implicate that persistent oral infections are already progressing from asymptomatic stage towards intraepithelial lesion. However, we do not have any evidence to demonstrate that these infections are progressing to clinical lesions or towards malignancy. In a careful clinical examination performed after 3 years of FU for the entire cohort, there were no signs of potentially malignant lesions in the oral mucosa. Currently, the reservoir of oral HPV infection is unknown. Unfortunately, FU studies on the progression of asymptomatic oral HPV infection to oral potentially malignant lesions such as leukoplakia are still missing. There are two old retrospective studies available; one showing the progression of HPV-positive leukoplakias to oral cancer within the 10-year FU [33]. In the other study by Nielsen and coworkers, 39 leukoplakias were followed with in situ hybridization and PCR with HPV16 E6 primers [34]. Altogether, 40.8% of leukoplakias were HPV DNA positive. They concluded that it is likely that HPV is a cofactor of cancer progression, as 100% of patients who developed oral cancers within 4–12 years were all positive for HPV. Here, the number of individuals in persistent and clearance groups were too limited to identify any predictors for oral HPV16 persistence and clearance except the age in women and the use of intrauterine device. However, we have earlier published the risk factors of persistence and clearance of oral HPV infection in women and men of the entire Finnish Family HPV Study cohort [9–12]. We showed that in men, the history of genital warts decreased the risk of persistent oral high-risk (clade alpha, species 7/9) HPV infections while smoking increased it. No predictors of high-risk HPV clearance were disclosed in men. In women, being seropositive for low-risk HPV at baseline increased the risk of
persistence whereas a second pregnancy and use of oral contraceptives protected from it. Oral HPV clearance was explained by older age (more common with increasing age, as found in this subcohort), history of atopic reactions (increased clearance), history of sexually transmitted diseases (decreased clearance) and new pregnancy (protective) during FU. In those earlier studies, the definition of oral HPV persistence was at least two subsequent HPV-positive tests with the same genotype while here the definition was at least 24 months’ HPV16 persistence.

Recent studies on HPV16 integration in oropharyngeal cancers have shown that pure integrated form is less common than in genital cancers. Recently, both Faust et al. [35] and Deng et al. [36] reported that only integrated HPV16 was present in 6% of the 78 and 18 fresh biopsies, respectively, while mixed status was the most prevalent one, found in 42 and 72% of the samples. Koskinen and coworkers found integrated HPV in 48% of paraffin-embedded HNSCC samples. One would expect to find lower integration rates in paraffin-embedded samples than in fresh biopsies due to methodologic reasons [37]. New next-generation sequencing technologies have identified much higher integration rates of HPV and distribution of viral genome not only in the E1–E2 region but in the entire HPV genome [20]. Thus, it is to be expected that the mixed pattern with low copies of integrated HPV is more common than found before. Also, the excess of episomal form in a mixed HPV16 status might decrease the detection of integrated form with the E2/E6 ratio PCR method. We showed earlier that a 10-fold excess of episomal form over integrated form interferes with the test regardless of the amount of viral DNA. The same was true with the background DNA exceeding 1500 ng in the reaction [38]. In line with our results, Arias-Pulido and coworkers showed that the detection of low-frequency HPV16 integration events was markedly limited by the common presence and abundance of HPV episomal forms [29].

Interestingly, we identified two women and one HPV16-positive man where only E2 was detectable without any E6. Recent studies on genome-wide profiling of HPV integration in cervical cancer have profiled not only the integration site of HPV but also the integration pattern of the virus as such. Hu et al. in 2015 annotated breakpoints on the
HPV16 genome [39]. Unexpectedly, in contrast to the reports that integrated HPV16 should retain intact oncogenes E6 and E7 with the long control region, they found that breakpoints could occur in any part of the viral genome, perhaps enabling the virus to adapt to the changing environment during carcinogenesis. Thus, methods like ours to identify HPV16 integration status on the basis of the E2/E6 ratio may be inaccurate because the E6 gene may be disrupted in some events. However, they found that breakpoints were most prone to occur in E1 instead of E2, followed by L1, L2 and E6 [39]. Studies on the genetic profile of HPV16 in HNSCC have also reported a few cases with undetectable E6 [37]. These few cases of oral HPV16 without detectable E6 gene need further studies using viral sequencing or wide genome analysis to understand the significance of this finding.

We also found interesting gender differences in the viral copy number of HPV16 E6. In the clearance group, HPV16 copy numbers were significantly higher in women than in men, whereas no gender difference was seen among persisters. Among males, persisters had higher viral load of HPV16 than those who cleared their HPV16 infection. As far as the authors know, there are no previous studies on the gender differences in HPV16 copy numbers. However, there was an interesting study from the early 1990s where E4 and L1 gene products and copy numbers of HPV6 and HPV11 were analysed in benign genital condylomas [40]. In that study, 17 out of 50 women were pregnant, and also 9 men were included. Both the E4 and L1 proteins were found more frequently in lesions from pregnant women than in non-pregnant women or men. Higher E4 and L1 protein levels correlated with the higher copy numbers [40]. These observations are in agreement with our present results and stress the importance of pregnancy as one of the determinants of HPV viral load. Here, the copy numbers of HPV16 were fivefold higher in samples taken during the third trimester of pregnancy in women who eventually cleared their oral infection. The opposite was found in women with persistent infections, in whom the viral load was sevenfold higher 2 months after the delivery than during the pregnancy. In our entire Finnish Family HPV Study cohort, there were almost 100 women who became pregnant for the second time during the 6-year FU. We have previously shown that a new pregnancy protected the women against incident oral HPV infections [10, 11]. However, if a woman had oral HPV infection already at the onset of the second pregnancy, the virus was not prone to clear [10, 11]. The present study confirms that also the index pregnancy shows a similar behaviour. Thus, women who had oral HPV already during the index pregnancy cleared their infection afterwards, as shown by a dramatic decline of viral load after the delivery. In persisters, oral HPV was acquired at high viral load soon after pregnancy, but the copy numbers declined in 6 months. This suggests that HPV infection at high viral load might increase the risk for viral integration and HPV persistence. This study provides further confirmative evidence on the temporal multistep events in progression of HPV infection, as hypothesized earlier [16].

As previously stated, there are no FU studies available on the role of HPV viral load and physical state on the progression of asymptomatic oral or oropharyngeal HPV infections to cancer precursor lesions or cancer. However, several studies have analysed the viral load of HPV16 in HNSCC. Most of these studies confirm that there is a wide variation in HPV16 viral load among the samples [36, 37, 41–43]. Recently, Faust and coworkers reported a variation up to 100 000-fold for HPV16 viral load in different samples [35]. The reported copy numbers have varied from 10 to 4 901 400 copies per 10 000 cells or from less than 1 copy cell−1 to over 1000 copies cell−1 [35, 37]. Especially in oropharyngeal cancers, the copy numbers have been highest in all studies and high viral load and episomal form of HPV16 have been associated with more favourable prognosis [35, 43–45]. As expected, the copy numbers of HPV16 E6 in asymptomatic oral HPV infection are here much lower than found in head and neck cancers varying from 7.6 to 96.6 copies per 1000 cells. As we used oral brush samples, the sample always contains bacterial DNA in addition to human DNA. Thus, the actual copy numbers per cell are most probably higher. We rather preferred here to express the copies per 50 ng DNA to allow the comparison between those studies using E2/E6 ratio PCR method.

Taken together, persistent oral HPV16 infections in both genders are clearly associated with mixed or integrated state of HPV16. Pregnancy might impact on the viral load and outcome of oral HPV16 infection. However, further confirmative studies are required to confirm the significance of viral integration and loss of episomes as a key event in oral HPV16 persistence and progression. If this is shown to be true, detection of mixed or integrated HPV16 infection in oral mucosa might predict an increased risk for HPV-associated cancer.

**METHODS**

**Study design**

This study is part of the prospective Finnish Family HPV Study ongoing at the Department of Oral Pathology, Institute of Dentistry, Faculty of Medicine, University of Turku, and the Department of Obstetrics and Gynecology, Turku University Central Hospital, since 1998. The study protocol and its amendment (no. 2/1998 and no. 2/2006) were approved by the Research Ethics Committee of Turku University Hospital. Members of 329 families (329 mothers, 131 fathers, 331 offspring) were enrolled into the study [46]. Written informed consent was granted from all subjects. The present subcohort included women (mothers) and men (fathers) who had asymptomatic oral HPV16 infection which either persisted or cleared during the 6-year FU. The cases included 14 women with 51 samples and 7 men with 21 samples, and the clearance group consisted of 41 women with 64 samples and 13 men with 24 samples (Table 1).
collected during the FU. In the present substudy, HPV persistence was defined as ≥24 months’ 16 type-specific duration. In women, the mean time of persistence was 45.0 ±17 months (median 44.4 months, range 25.6–81.2 months). In men, the mean duration of HPV16 persistence in oral mucosa was 45.1±7.0 months (median 43.9 months, range 37.2–55.1 months).

**Oral brush samples**

Oral brush samples (Cytobrush; MedScan) were collected from both spouses at baseline, and at month-2, -6, -12, -24, -36 and -72 visits [9, 11, 46]. HPV DNA was extracted from oral samples with the high-salt method. For HPV detection, nested PCR with primer pairs MY09/MY11 and GP05+/GP06+ was performed, followed by HPV genotyping with Multimix kit (Progen Biotechnik), which detects 24 HPV genotypes [9]. Table 1 summarizes the numbers of subjects and the outcome of their oral HPV16 infection as related to time points of oral sampling.

**Analyses of HPV16 integration and copy number of E6**

HPV16 integration was analysed by using our method described in 2002 [16]. Briefly, a region of the E2 ORF being most often deleted during HPV16 integration is targeted by one set of PCR primers and a probe, while another set targets the E6 ORF. In episomal state, both targets should be equivalent, while in the integrated state, the copy numbers of E2 would be less than those of E6. E2 and E6 ORFs are amplified simultaneously with quantitative PCR (qPCR).

The primers and probes used in the qPCR run with the LightCycler96 System (Roche Diagnostics) are given in Table S1 (available in the online Supplementary Material). The HPV16 E2 and E6 probes were labelled with VIC and 6-FAM dyes, respectively.

qPCR amplification was performed in a 20 µl volume containing 20 pmol µl⁻¹ of each primer and 100 pmol µl⁻¹ of each probe, 10 µl volume of ready-to-use hot-start PCR mix containing Taq DNA Polymerase, reaction buffer, dNTP mix and MgCl₂ (FastStart Essential DNA Probes Master; Roche Diagnostics). Totally, 100 ng of target DNA from oral brush samples was added into the reaction mixture. The qPCR amplification conditions were used as described previously: 10 min at 95°C, followed by a two-step cycle of 15 s at 95°C and 60 s at 60°C, repeated 40 times [31]. However, specifically for LightCycler96, we increased the cycle number to 50 as suggested by the manufacturer’s technical support.

In each run, standard curves were obtained by amplification of a dilution series of 500 000, 50 000, 5000 and 50 copies of a clone of HPV16 in pBR322 (kindly provided by H. zur Hausen, German Cancer Research Centre, Heidelberg, Germany). At least three no-template control reaction mixtures were included in each run. All experiments were performed in triplicate. The raw data gathered by the instrument was transferred to the LightCycler96 Application Software version 1.1 for analyses.

The absolute copy numbers of HPV16 E6 in 50 ng cellular DNA were recorded. The copy numbers were expressed also as copies per 1000 cells by estimating the DNA content of one diploid human cell to be 6 ng. The integrated E6 was calculated by subtracting the copy numbers of E2 (episomal) from the total copy numbers of E6 (episomal and integrated). Ratios of E2/E6 of less than 0.93 indicate the presence of both integrated and episomal forms. Values of 0.93–1.08 indicate predominance of the episomal form, whereas in integrated state, E2 is not detected.

**Statistical analyses**

The results were analysed using the IBM SPSS Statistics software version 20.0 for Mac (IBM). Chi-squared ($\chi^2$) test was used to compare the frequencies among groups, accompanied by Fisher’s exact test when indicated. Differences in the means of continuous variables (copy numbers) were analysed using the $t$-test for two independent samples. In all tests, $P<0.05$ was considered statistically significant.

**References**


---

**Five reasons to publish your next article with a Microbiology Society journal**

1. The Microbiology Society is a not-for-profit organization.
2. We offer fast and rigorous peer review – average time to first decision is 4–6 weeks.
3. Our journals have a global readership with subscriptions held in research institutions around the world.
4. 80% of our authors rate our submission process as ‘excellent’ or ‘very good’.
5. Your article will be published on an interactive journal platform with advanced metrics.

Find out more and submit your article at microbiologyresearch.org.