Virology 402 (2010) 197–202

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

Association of viral load with HPV16 positive cervical cancer pathogenesis: Causal relevance in isolates harboring intact viral E2 gene

Damayanti Das [a], Bornali Bhattacharjee [b], Shrinka Sen [a], Indranil Mukhopadhyay [a], Sharmila Sengupta [a,*]

[a] Human Genetics Unit, Indian Statistical Institute, 203, B.T. Road, Kolkata 700108, India
[b] University of Massachusetts Medical School, Department of Molecular Genetics and Microbiology, Worcester, MA, USA

ARTICLE INFO

Article history:
Received 27 November 2009
Returned to author for revision
3 February 2010
Accepted 17 March 2010
Available online 14 April 2010

Keywords:
HPV16 infection
Cervical cancer
Intact E2 gene
High viral load
E2 disruption

ABSTRACT

We tested the hypothesis that cervical cancers (CaCx) harbor high HPV16 viral load compared to controls and this is influenced by E2 status and age of subjects. Viral load (natural log transformed values) per 100 ng genomic DNA was estimated (152 cases and 87 controls) by Taqman assay. Median viral load was significantly higher (Mann–Whitney *U* test) among cases (17.21) compared to controls (9.86), irrespective of E2 status or upon considering E2 status as a covariate in logistic regression model ($p<0.001$). Viral load of E2 intact cases (17.80) was significantly higher ($p<0.001$) compared to those with disrupted E2 (9.78). At equivalent probability of being a case, viral load was higher among individuals (i) of lower age, irrespective of E2 status, and (ii) with intact E2 but of similar age as those with disrupted E2. Thus viral load in association with E2 status and/or age might be of causal relevance in CaCx pathogenesis.

Introduction

Cervical cancer (CaCx) is one of the leading causes of death among Indian women, contributing to one-fourth of the global total of deaths incurred due to this. Recent reports (http://www-dep.iarc.fr/) depict that 132,082 new cases and 74,118 deaths are recorded every year (age-standardized rates: incidence = 30.7; mortality = 17.8) in India. While HPV infections appear to be the major etiological factors for CaCx development (Schlecht et al., 2001), HPV16 is the most common high-risk type identified among such cases both worldwide (Munoz et al., 2003) as well as in India (De Andrea et al., 2007; Franceschi et al., 2005; Laikangbam et al., 2007; Sowjanya et al., 2005). Majority of HPV infections regress spontaneously. Only a small percentage of high-risk HPV-infected women harbor persistent infections that lead to cervical carcinogenesis, after a long latent period. This points to the need of HPV testing and the identification of HPV-related cofactors that might be relevant for detecting those at risk of developing CaCx in the long run.

It is known that viral genome in episomal form replicates along with the differentiating epithelial cells from basal membrane to the superficial zone and is shed off along with the sloughed-off epithelial cells resulting in transient infection (zur Hausen, 2002). Persistent infection and viral genome integration into the host genome is known to mediate oncogenicity (Woodman et al., 2007). E1 and E2 are the early viral proteins needed for viral replication and translation, while E6 and E7 are the oncoproteins responsible for cellular transformation by inactivation of p53 and pRb proteins, respectively. E2 protein represses E6 and E7 expression. Integration of viral genome into the host genome, chiefly at fragile sites (Kalantari et al., 2001; Wentzensen et al., 2004), not only affects various cellular pathways of the host cell-cycle machinery, but also disrupts the viral E2 gene most commonly in the hinge region of the HPV16 E2 protein. In absence of E2-driven repression, E6 and E7 are over-expressed driving infected cells toward transformation.

On the contrary, our study (Bhattacharjee and Sengupta, 2006a) as well as a few others (Narayanan et al., 2004) has identified that a substantial proportion of individuals with CaCx harbor intact E2, which could be either purely intact or concomitant, i.e., a mixture of intact and disrupted forms. Such observations, point toward the biological plausibility of cervical carcinogenesis under the impact of HPV16 intact E2 gene or intact viral genomes, in addition to E2 disruption due to viral genome integration into the host genome. Many study groups have proposed viral load estimates per cell or per unit amount of genomic DNA, as a potential HPV-related biomarker, which could be used for predicting those at risk of CaCx development (Franco and Coutlée, 2009). However, there are several reports, which have failed to relate high HPV16 DNA copy number with CaCx development (Swan et al., 1999; Josefsson et al., 2000; van Duin et al., 2002; Hernandez-Hernandez et al., 2003; Abba et al., 2003; de Boer et al., 2007).

We undertook the present study to investigate the association of HPV16 viral load, if any, with CaCx by comparing HPV16 positive cytologically normal women with those diagnosed with CaCx (squamous cell carcinomas). Apart from considering the potential of viral load

* Corresponding author. Fax: +1 91 33 2577 6680/3049.
E-mail addresses: sharmilasg2003@yahoo.com, sharmila@isical.ac.in (S. Sengupta).

0042-6822/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2010.03.030

as an independent factor of disease risk, we further determined its association with the risk of CaCx in consideration with other factors related to the viral genome such as (i) disruption or intactness of the viral E2 gene and (ii) ratios of viral load based on E2 and E6 gene copies, i.e., E2/E6 ratios. A number of studies (Peitsaro et al., 2002; Cricca et al., 2007; Saunier et al., 2008) have proposed the latter factor as a reliable indicator of the physical status of viral genomes within the host cells.

Results

Distribution of viral load among cases and controls

Number of copies of HPV16 viral genome (viral load) per 100 ng of HPV16 positive genomic DNA (natural log transformed) was evaluated in 152 cases and 87 controls. The viral load per 100 ng HPV16 positive genomic DNA was calculated from the standard plot constructed with various concentrations of HPV16 plasmid DNA per 100 ng genomic DNA, after amplification of a portion of HPV16 E6 by qRT-PCR. Viral load among cases ranged between 3.977 and 22.764, as shown in Fig. 1, with a median viral load of 17.21. Viral load among controls ranged between 3.736 and 20.527, with a median viral load of 9.86. Using Mann–Whitney *U* test, we identified that irrespective of viral E2 gene status (i.e., E2 intact or disrupted), median viral load in cases (17.21) was significantly (p-value<0.001) higher compared to that in controls (9.86).

Distribution of different physical forms (intact or disrupted) of E2 gene among cases and controls

Having confirmed the E2 disruption status as disrupted (pure disruption) or intact, as per the protocol adopted earlier in our laboratory (Bhattacharjee and Sengupta, 2006a,b), we undertook the method of Peitsaro et al. (2002) to determine the proportion of E2 intact samples that were of the mixed type, i.e., harbored disruption of E2 as well as intact E2. As per this protocol (Peitsaro et al., 2002), on the basis of the ratio of copy numbers of E2 and E6 (E2/E6), the E2 gene is interpreted as intact when $E2/E6 \geq 1$, mixed or concomitant forms of intact and disrupted when $0<E2/E6<1$, and disrupted when $E2/E6 = 0$, i.e., E2 fails to amplify.

We observed that, out of 152 cases, 139 (91.45%) had intact E2 (22 purely intact and 117 concomitant) and 13 (8.55%) had purely disrupted E2. Out of 87 controls, 77 (88.5%) had intact E2 (26 purely intact and 51 concomitant) and 10 (11.49%) had purely disrupted E2. Case control analysis of this data revealed that while pure E2 disruption or viral integration failed to show any significant difference between the sample groups, concomitant or mixed forms of E2 or viral genomes were significantly higher among cases compared to controls (OR = 2.71; CI = 1.41–5.22; p-value = 0.002) among those identified with intact E2 by PCRs covering the entire E2. Similar observation was also recorded, when disrupted and concomitant forms were grouped together in comparison to the purely intact E2 (OR = 2.52; CI = 1.32–4.79; p-value = 0.004) and case control analysis was performed.

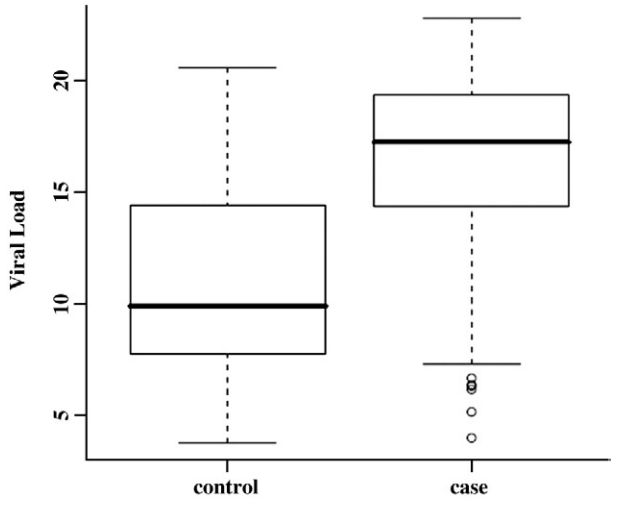

Fig. 1. Box plots representing distribution of HPV16 viral load (natural log transformed) per 100 ng genomic DNA among case and control samples.

Viral load among CaCx cases in consideration of E2 disruption status or various physical forms of the samples

Based on Mann–Whitney *U* test, we observed that the cases with intact E2 ($n = 139$) harbored a significantly ($p<0.001$) higher viral load (median value = 17.80) than those with disrupted E2 ($n = 13$; median value = 9.78). The range of viral load within E2 intact cases was between 6.376 and 22.764, while that within E2 disrupted cases was between 9.778 and 19.336. Further analysis of the data revealed that viral load was significantly higher ($p = 0.001$ and $p<0.001$, respectively) among CaCx cases having intact E2 or $E2/E6 \geq 1$ ($n = 22$; median value = 17.053) and those having mixed forms or $0<E2/E6<1$ ($n = 117$; median value = 17.930) compared to CaCx cases having disrupted E2. Among the CaCx cases having mixed forms of E2, we determined the HPV16 viral load of the isolates harboring disrupted E2, by subtracting the copy number of intact E2 from the total viral load represented by E6. It was observed that the median viral load of such isolates within the pool of mixed samples was 1.380 (range 0.024–1.907) and this was significantly lower ($p<0.001$) compared to that of CaCx cases having disrupted E2 only. The observations indicate a possible association between viral load and CaCx, which is likely to be influenced by E2 status (intact or disrupted).

Logistic regression analysis of the association of viral load with CaCx

It was observed that there was a significant difference ($p<0.001$) in the mean age of CaCx cases (49.60 ± 11.06 years) and that of controls (33.52 ± 11.014 years). However, the mean age of those CaCx cases harboring intact E2 gene (49.64 ± 11.10 years) failed to differ ($p = 0.899$) from those harboring disrupted E2 (49.23 ± 10.12 years). We therefore determined the potential of viral copy number or load as a risk factor of HPV16-related CaCx, either independently or by considering it as a cofactor of disease risk along with E2 disruption status and determined the $OR_{age\text{-}adjusted}$ and p-values, based on logistic regression models. Such analysis is depicted in Table 1, which revealed that viral load was significantly associated with CaCx [$OR_{age\text{-}adjusted} = 1.261$; CI = 1.161–1.369; $p<0.001$] irrespective of other factors. Logistic regression analysis considering the two viral factors as covariates showed that both viral load [$OR_{age\text{-}adjusted} = 1.333$; CI = 1.207–1.472; $p<0.001$] and E2 disruption status [$OR_{age\text{-}adjusted} = 5.225$; CI = 1.426–19.149; $p = 0.013$] were significantly associated with CaCx.

Table 1
Age-based logistic regression analysis of the association of viral load with CaCx either as an independent risk factor or in consideration of other viral risk factors.

Model no.	Model variables	Regression coefficient	S.E.	p-value	OR (95% CI)
1	Age	0.124	0.020	0.000	1.132 (1.089–1.177)
	Viral load	0.232	0.042	0.000	1.261 (1.161–1.369)
	Constant	−7.747	1.078	0.000	0.000
2	Age	0.128	0.021	0.000	1.137 (1.091–1.184)
	Viral load	0.287	0.051	0.000	1.333 (1.207–1.472)
	E2 disruption status	1.653	0.663	0.013	5.225 (1.426–19.149)
	Constant	−8.855	1.271	0.000	0.000

Prediction of CaCx risk under the impact of viral load, E2 disruption status, and age of subjects

Viral load and E2 disruption status were significantly associated with CaCx risk as revealed by the age-adjusted OR (95% CI) and *p*-values in Table 1. In addition, there was significant difference in mean age between the CaCx cases and controls, irrespective of E2 disruption status. Therefore, taking into consideration the effects of all these explanatory variables, two probability-prediction graphs for the disease risk were constructed, using the fitted logistic regression models of case control status on viral load, E2 disruption status, and age. Fig. 2A and B represents the probability of disease risk for different viral loads at different ages ranging from 20 to 70 years, in the presence of intact E2 (Fig. 2A) or disrupted E2 (Fig. 2B), respectively. Based on comparative analysis of Fig. 2A and B, it was evident that the disease risk increased with increase in viral load and also with increase in age of the individual, irrespective of their E2 disruption status. Moreover, at a particular age, for any viral load, probability of being classified as a CaCx case was higher for individuals harboring disrupted E2 than for those harboring intact E2. Irrespective of E2 status, at similar risk levels, individuals of lower age were found to harbor higher viral load, compared to those of higher age. In addition, at equal risk level and being of similar age, those harboring intact E2 portrayed a higher viral load compared to those with disrupted E2, as evident from the plots.

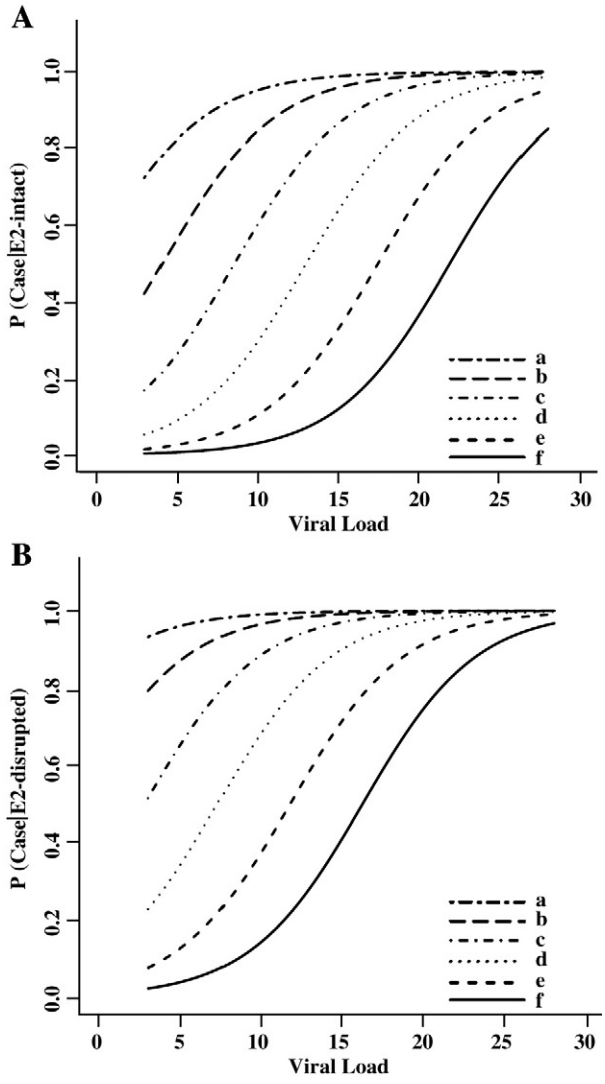

Fig. 2. (A) Probability scores (*Y*-axis) for the disease risk versus viral load (natural log transformed) per 100 ng genomic DNA (*X*-axis) among individuals of different ages harboring intact HPV16 E2 gene: a (age 20 years); b (age 30 years); c (age 40 years); d (age 50 years); e (age 60 years); f (age 70 years). (B) Probability scores (*Y*-axis) for the disease risk versus viral load (natural log transformed) per 100 ng genomic DNA (*X*-axis) across different ages, harboring disrupted HPV16 E2 gene: a (age 20 years); b (age 30 years); c (age 40 years); d (age 50 years); e (age 60 years); f (age 70 years).

Discussion

HPV16 E2 gene disruption, as a consequence of viral integration into the host genome, is considered as a critical event in the pathogenesis of cervical neoplasia due to the loss of negative feedback control of viral oncogene expression (Tan et al., 1994; Jeon and Lambert, 1995). Recent reports (Narayanan et al., 2004; Anderssen et al., 2005; Cheung et al., 2006), including one from our laboratory (Bhattacharjee and Sengupta, 2006a) portraying the presence of intact E2 genes in almost 60% of HPV16 positive CaCx cases, have subsequently paved the way for new paradigms of cervical carcinogenesis (Bhattacharjee and Sengupta, 2006b; Bhattacharjee et al., 2008). These are directed toward exploring alternative mechanisms of loss of E2 repression that could lead to sustained E6/E7 expression even with intact E2 gene, i.e., in the presence of concomitant viral genomes (both episomal and integrated) or purely episomal viral genomes.

Apart from viral integration status or the status of viral E2 genes (intact or disrupted), other HPV-related factors are also being considered for predicting the risk of CaCx and viral load appears to be one such factor. We, therefore, estimated the viral load of HPV16 in our samples (cases and controls) on the basis of E6 copy numbers irrespective of other viral factors or taking into consideration the status of E2 (intact or disrupted).

Our study reflected a significant association of viral load with CaCx, which was supported both by non-parametric Mann–Whitney *U* test and logistic regression analysis. This was also obvious, when viral load was considered along with E2 disruption status. The logistic regression model (Table 1) justified the significant role of both viral load and E2 disruption status in the pathogenesis of CaCx under the impact of HPV16 infection. Earlier, we reported that E2 disruption was significantly higher among CaCx cases compared to controls (Bhattacharjee and Sengupta, 2006a,b). The present analysis, considering viral load as a risk factor of CaCx development, also confirmed the association of viral E2 disruption with CaCx.

The intactness or disruption status of viral E2 gene concomitant with viral load thus appears to be relevant for HPV16 positive CaCx pathogenesis. The biological plausibility of this observation is likely to be supported by the fact that E2 protein enhances viral DNA replication by interacting with the viral replication factor E1 and recruiting it to the origin of replication (Gillitzer et al., 2000; Berg and Stenlund, 1997; Yasugi et al., 1997) and also plays a more direct role in replication facilitating viral genome segregation by tethering the viral genomes to host mitotic chromosomes (Jaquelline et al., 2006).

An attempt to predict the probability of disease risk on the basis of the interplay of the significantly associated viral factors like viral load and E2 disruption status along with the individual's age revealed that probability of CaCx increased with viral load irrespective of age and E2 status (Fig. 2A and B). Among individuals of similar age and viral load, the probability of disease risk was higher among those having disrupted E2 than those with intact E2. It was also clear from such analysis that among individuals having similar probability of disease risk, viral load was higher among (i) those of lower age, irrespective of E2 status, and (ii) among those of similar age but having intact E2. These observations clearly justify the role of high HPV16 viral load in enhancing the risk of CaCx and portray that attainment of high viral load could probably complement the lack of E2 disruption in mediating cervical carcinogenesis.

Our analysis also portrayed that viral load was significantly higher among CaCx cases harboring intact E2 (either purely intact or concomitant) than among cases with purely disrupted E2. We observed for the first time in this study that copy number of viral isolates with

disrupted E2 within CaCx cases harboring concomitant forms (mixtures of E2 intact and disrupted) was significantly lower compared to that within CaCx cases harboring purely disrupted E2 only. We therefore speculate that intact E2 in CaCx cases may be relevant for disease pathogenesis in the absence of E2 disruption, or in the presence of reduced disruption of E2. Hence there is a need to validate the hypothesis that another complementary pathway of cellular transformation might exist, which could be mediated through a high viral load in the absence of or decreased E2 disruption events.

Measurement of HPV16 load has long been suggested as a potential tool to indicate severity of cervical neoplasia (Peitsaro et al., 2002), showing increasing viral load from mild dysplasia to CaCx (Abba et al., 2003; Hernandez-Hernandez et al., 2003; Josefsson et al., 2000) and our study is in agreement with this, reflecting significantly higher viral load in CaCx cases compared to controls. However, some studies have failed to identify higher viral loads in cervical dysplasia or cancer (Franco and Coutlée, 2009), which could be attributable to a number of factors, mainly the lack of standardized protocols in viral load estimation (Saunier et al., 2008), different methods used for the calculation of viral load (Roberts et al., 2008), study design, or samples tested (Wang and Hildesheim, 2003) etc.

In this study, for estimation of viral load, we used Taqman-based real-time PCR technique, which appears to be both sensitive and specific (Häfner et al., 2008; Gallo et al., 2003) and interpreted viral copy numbers present per 100 ng of input genomic DNA. In a randomly selected subset of DNA samples (from both exfoliated cervical cells and biopsy tissues), we estimated the copy number of the host gene G6PD by a Taqman assay using 100 ng genomic DNA per assay. Such assay failed to differ between the two sample types (data not shown), substantiating the estimation of viral copy numbers present per 100 ng of genomic DNA. We also generated titration curves of HPV16 viral load (natural log transformed) covering a concentration range of 1–200 ng genomic DNA (natural log transformed) for a randomly selected subset of DNA samples from control scrapes as well as case biopsies. The curves were linear in nature and the slopes were similar in all cases analyzed (data not shown), suggesting that the nature of samples, whether cervical scrapes or biopsy material, is unlikely to influence the viral load values recorded as per the assay protocol followed by us.

It is worth noting that intact HPV16 E2 in the episomal form is often found to coexist with disrupted forms in CaCx (Matsukura et al., 1986), often together with high viral load in such samples (Peitsaro et al., 2002; Saunier et al., 2008; Kulmala et al., 2006), which is similar to our observations. A few studies have also reported an increased frequency of intact episomal HPV16 DNA in CaCx (Cullen et al., 1991). Furthermore, existence of intact E2 in cervical cancer cell lines such as CaSki or CC7T/IVGH (Choo et al., 1989) has also been observed. In fact, some earlier studies have recorded the presence of episomal and integrated HPV within the same cells in clinical samples confirmed by in situ hybridization (Cooper et al., 1991; Kristiansen et al., 1994). Together, such observations justify our approach of analyzing viral load in DNA samples harboring intact or disrupted E2, given the possibility of the existence of both forms in CaCx tissues.

The prevailing concept regarding cervical carcinogenesis is in favor of enhanced cellular proliferation following E2 disruption as a consequence of viral integration, due to withdrawal of E2-induced repression of E6 and E7 oncogenic transcription (Saunier et al., 2008; Kulmala et al., 2006; Häfner et al., 2008). However, contradictory observations have also been recorded, which revealed no significant increase in E6 or E7 expression following E2 disruption (Häfner et al., 2008; Cheung et al., 2008). Furthermore, because of viral integration into the host genome leading to chromosomal instability, the effectors of cellular transformation could probably follow different pathways for different subgroups of CaCx samples (Arias-Pulido et al., 2006; Kulmala et al., 2006; Gallo et al., 2003; Duensing and Munger, 2002; Scheurer et al., 2007). A study (Kozuka et al., 2000) based on SiHa and CaSki cell lines also proposed that high copy number of HPV DNA would directly contribute to increased expression of E6 and E7. Some studies have identified a decrease in viral load with E2 disruption attributable to interruption of viral replication machinery (Saunier et al., 2008; Kulmala et al., 2006) that depends on E1 and E2. Taken together, these studies render strong biological plausibility to our observation of high viral load in CaCx cases harboring intact E2, in comparison to those harboring disrupted E2.

In conclusion, our exploratory analysis adds to the body of literature that confirms the significant association of high viral loads with CaCx, independent of other viral factors. Additionally, our approach further highlights the significance of viral load in CaCx pathogenesis in association with age (high viral load at low age) and E2 status (high viral load among E2 intact individuals). We have further identified for the first time that CaCx cases harboring intact HPV16 E2 gene (purely intact or concomitant) are capable of mounting a high viral load overall, wherein viral load of isolates with disrupted E2 among cases with concomitant forms of E2 was significantly lower, when compared to those harboring disrupted E2 only. Such observations are likely to have significant implications in disease pathogenesis, probably by contributing to sustained E6 and E7 expression. Based on the findings recorded in this study, as well as our previous observations (Bhattacharjee and Sengupta, 2006a,b; Bhattacharjee et al., 2008), we reconfirm that multiple pathways other than E2 disruption are likely to be instrumental in CaCx causation.

Materials and methods

Samples and subjects

The samples used for this study were nested to an ongoing natural cohort study (Bhattacharjee and Sengupta, 2006a,b; Laikangbam et al., 2007; Bhattacharya and Sengupta, 2007; Bhattacharjee et al., 2008). The case samples (biopsy tissues) were collected from married women who were attending a cancer referral hospital (Cancer Centre Welfare Home and Research Institute, South 24 Parganas, West Bengal, India) during the period 2000–2007. Out of a total of 300 malignant cervical samples (tissue biopsies) collected during this period, 160 were identified to be HPV16 positive, of which DNA samples from 152 were available for the present analysis. These malignant samples were histopathologically confirmed as invasive squamous cell carcinomas and clinically diagnosed as tumor stage III and above as per FIGO classification, and the age of these women varied between 27 and 80 years of age (mean age: 49.60 years; S.D.: 11.06 years).

The control samples were normal cervical scrapes confirmed by Pap smear test and derived from married and non-pregnant (or, 6 months post-partum) women with no previous history of cervical dysplasia/malignancy. These women were attending a Reproductive Health (RH) Clinic (Child in Need Institute or CINI, South 24 Parganas, West Bengal, India), for routine reproductive healthcare counseling. Such samples were also collected through cervical screening camps organized jointly by CINI and us in the neighborhood villages from where women came to attend the RH clinic. All the control samples were collected within the period of 1998–2007. Out of a total of 1112 cervical scrapes, 94 were identified to be HPV16 positive, of which 87 DNA samples were available for the current analysis. The age of these 87 women varied between 18 and 77 years (mean age: 33.52 years; S.D.: 11.014 years).

All samples, cases, and controls were collected from the subjects with informed consent approved by the institutional ethical committee for human experimentation. Details regarding subjects, samples, DNA isolation, HPV screening and determination of HPV16, and E2 disruption status are described elsewhere (Laikangbam et al., 2007; Bhattacharjee and Sengupta, 2006a).

Viral load estimation by real-time PCR of E6 gene

The quality and concentration of the stock DNA isolated from HPV16 positive individuals (both cases and controls) was determined

spectrophotometrically. Working DNA solutions of concentration 100 ng/μl were prepared by diluting the stock with TE buffer and stored in aliquots at 4 °C to be used for subsequent PCRs. A primer pair for β-globin gene, which amplifies a 268-bp fragment, was used as internal control, to ensure the integrity of the DNA samples (Tachezy et al., 1994).

The DNA samples used for this study all portrayed the presence of β-globin gene amplicon and viral load was expressed as viral copy numbers present per 100 ng of genomic DNA. Quantitative Real-Time PCR (qRT-PCR) was performed for viral load assessment following absolute quantification method using TAQMAN hydrolysis probe on ABI 7900 HT Fast Real-Time PCR platform according to the method of Peitsaro et al. (2002). The E6 oncogene is established as a relatively stable viral gene. Therefore, a small region of E6 (81 bp: nucleotides 94 to 174) was targeted for this assay, keeping the qRT-PCR efficiency around 100%. The PCR program was set with initial UNG activation at 50 °C for 2 min and denaturation at 95 °C for 10 min, followed by an amplification phase of 40 cycles each including two steps of denaturation and annealing at 95 °C for 15 s and 60 °C for 1 min, respectively. Hundred nanograms of HPV16 positive genomic DNA was targeted for qRT-PCR in a reaction mixture of 20 μl containing 25 ng of primers (forward: 5'-GAGAACTGCAATGTTTCAGGACC-3'; reverse: 5'-TGTATAGTTGTTTGCAGCTCTGTGC-3') and 0.1 μM of hydrolysis probe [(6-FAM)-CAGGAGCGACCCAGAAAGTTACCACAGTT-(DQ)] (Peitsaro et al., 2002) in Taqman Universal Master Mix (P.E. Applied Biosystems, Perkin-Elmer). A standard curve, plot ($R^2 = 0.99$) of threshold cycles (C_T) against log copy number, was generated using serial dilutions of 1.75×10^9, 1.75×10^7, 1.75×10^5, and 1.75×10^3 copies of the HPV16 plasmid insert (pUC 19 plasmid vector with HPV16 ref. sequence insert = 10.582 kb; concentration = 20×10^3 ng/μl) along with 100 ng genomic DNA (placental DNA). Viral copy number per 100 ng of genomic DNA was calculated by the following formula and was then log transformed (natural log):

Number of copies per μl = (amount of plasmid in grams) ÷ (weight of one molecule of plasmid in grams) = 1.75×10^{12} copies = 1.75×10^{12} copies/μl
[Amount of plasmid in grams = $20\times10^3\times10^{-9}$]
[1 kb of double stranded DNA = 6.5×10^5 Daltons; Avogadro's number = 6.023×10^{23}; 1 Dalton = $(1/6.023\times10^{23})$ g]
[Weight of one molecule of plasmid in grams = $10.582\times6.5\times10^5/6.023\times10^{23}$]

Three non-template controls (NTCs) were used in each assay to exclude the chances of non-specific amplification. The viral copy number values (per 100 ng of genomic DNA) of the unknown samples were interpreted from the standard curve, which showed a linear relationship between the threshold cycle values plotted against the log of the copy number over the entire range of dilutions. The slope of the standard curve was used to determine the efficiency [Efficiency = $10^{(-1/\text{slope})} - 1$] of the PCR reaction. An assay in which 90–100% efficiency was obtained was considered for analysis. Each such assay was done at least twice, with three replicates per sample in each assay.

Estimation of E2 copy number and determination of E2/E6 ratios of the copy numbers by real-time PCR

In general, disruption of E2 gene releases the repressive effect of E2 protein on viral oncogene (E6 and E7) expression, leading to host cellular transformation. The most commonly reported (Kalantari et al., 2001; Saunier et al., 2008; Arias-Pulido et al., 2006; Kulmala et al., 2006) region that harbors the majority of disruption events within HPV16 E2 gene is the one that codes for the hinge of the E2 protein. A small portion (nucleotides 3361 to 3442; 82 bp) of this region was targeted to determine the presence and copy number of the E2 gene by qRT-PCR (Peitsaro et al., 2002) keeping efficiency of the assay around 100%. One hundred nanograms of HPV16 positive genomic DNA was used as template DNA in a reaction mixture of 20 μl containing 25 ng of primers (forward: 5'-AACGAAGTATCCTCTCCTGAAATTATTAG-3'; reverse: 5'-CCAAGGCGACGGCTTTG-3') and 0.1 μM of hydrolysis probe [(6-VIC)-CACCCCGCCGCGACCCATA-(DQ)] in Taqman Universal Master Mix (P.E. Applied Biosystems, Perkin-Elmer). The PCR program used and the method of estimation of E2 copy number was the same as that used for viral E6 copy number estimation. Copy numbers of E2 were determined from the standard plot ($R^2 = 0.98$), which was obtained in the same way as in case of qRT-PCR of E6, and log transformed (natural log). E2/E6 ratios were subsequently determined for each of the samples analyzed.

Confirmation of the disruption status of E2 gene

Samples were considered as having intact E2 only when amplification was recorded with real-time PCR for E2 as described above, and (i) PCR for the larger fragment of E2 (nucleotides 2811–3837; amplicon size 1027 bp) and (ii) PCR with three sets of overlapping primer pairs spanning nucleotides 2810–3471, 3448–3649, and 3596–3872, yielding amplicon sizes of 661, 203, and 277 base pairs, respectively, in three separate reactions. The latter PCR protocols have been published earlier from our laboratory (Bhattacharjee and Sengupta, 2006a,b). The samples that failed to amplify the larger E2 fragment as well as one or more of the overlapping fragments, irrespective of real-time PCR-based amplification of E2, were considered to harbor purely disrupted E2.

Statistical analyses

Box plots were constructed to observe the difference in distribution of viral load among cases and controls. Kolmogorov–Smirnov test identified viral load (copy number of viral genome per 100 ng of HPV16 positive genomic DNA) as a variable not following normal distribution. Therefore, non-parametric test (Mann–Whitney *U* test) was performed to study association of viral load with the disease phenotype. Logistic regression analysis was carried out to determine the association of viral copy numbers with CaCx. To investigate the individual risk factor effects, we have included the viral factors, viz. copy numbers and E2 disruption status in the logistic regression model, adjusting the possible effect of age. The significance of the effects of the individual risk factors or covariates is determined in terms of OR (95% CI) and *p*-values.

Based on such logistic regression models, probability-prediction graphs for disease risk were constructed against viral load at different ages of individuals considering E2 intact and E2 disrupted viral genome statuses as covariates. All statistical analyses were done using software packages SPSS (version 16.0 for windows) and R (www.r-project.org).

Acknowledgments

We thank the Cancer Centre Welfare Home and Research Institute (Thakurpukur, South 24 Parganas, West Bengal, India) and the Child in Need Institute (Pailan, South 24 Parganas, West Bengal, India) for their support in sample collection; Professor Partha P. Majumder and Dr. Saurabh Ghosh of Human Genetics Unit, Indian Statistical Institute, Kolkata, India, for providing guidelines for statistical analysis of the data; and all members of the Human Genetics Unit, Indian Statistical Institute, Kolkata, India, for technical support during the work. Special thanks are also given to the Indian Statistical Institute for providing Ms. Damayanti Das with a Fellowship (JRF and SRF) and the Council of Scientific and

Industrial Research, Government of India for providing Dr. Bornali Bhattacharjee with a Fellowship (JRF and SRF) to work on this project. This work was funded partially by the Indian Statistical Institute, Kolkata and partially by the Ministry of Environment and Forests; Grant No. F. No.19-50/2005-RE, Government of India.

References

Abba, M.C.S., Mouron, A., Gomez, M.A., Dulout, F.N., Golijow, C.D., 2003. Association of human papillomavirus viral load with HPV16 and high-grade intraepithelial lesion. Int. J. Gynecol. Cancer 13, 154–158.

Anderssen, S., Safari, H., Mints, M., Lewensohn-Fuchs, L., Gyllensten, U.B., Johansson, B., 2005. Type distribution, viral load and integration status of high-risk human papillomaviruses in pre-stages of cervical cancer (CIN). Br. J. Cancer 92, 2195–2200.

Arias-Pulido, H., Peyton, C.L., Joste, N.E., Vargas, H., Wheeler, C.M., 2006. Human papillomavirus type 16 integration in cervical carcinoma in situ and in invasive cervical cancer. J. Clin. Microbiol. 44, 1755–1756.

Berg, M., Stenlund, A., 1997. Functional interactions between papillomavirus E1 and E2 proteins. J. Virol. 71, 3853–3863.

Bhattacharjee, B., Sengupta, S., 2006a. HPV16 E2 gene disruption and polymorphisms of E2 and LCR: some significant associations with cervical cancer in Indian women. Gynecol. Oncol. 100, 372–378.

Bhattacharjee, B., Sengupta, S., 2006b. CpG methylation of HPV 16 LCR at E2 binding site proximal to P97 is associated with cervical cancer in presence of intact E2. Virology 354, 280–285.

Bhattacharjee, B., Mandal, N.R., Roy, S., Sengupta, S., 2008. Characterisation of sequence variations within HPV16 isolates among Indian women: prediction of causal role of rare non-synonymous variations within intact isolates in cervical cancer pathogenesis. Virology 377, 143–150.

Bhattacharya, P., Sengupta, S., 2007. Predisposition to HPV16/18-related cervical cancer because of proline homozygosity at codon 72 of p53 among Indian women is influenced by HLA-B*07 and homozygosity of HLA-DQB1*03. Tissue Antigens 70, 283–293.

Cheung, J.L., Lo, K.W., Cheung, T.H., Tang, J.W., Chan, P.K., 2006. Viral load, E2 gene disruption status and lineage of human papillomavirus type 16 infections in cervical neoplasia. J. Infect. Dis. 194, 1706–1712.

Cheung, J.L.K., Cheung, T.H., Tang, J.W.T., Chan, P.K.S., 2008. Increase of integration events and infection loads of human papillomavirus type 52 with lesion severity from low grade cervical lesion to invasive cancer. J. Clin. Microbiol. 46, 1356–1362.

Choo, K.B., Cheung, W.F., Liew, L.N., Lee, H.H., Han, S.H., 1989. Presence of catenated human papillomavirus type 16 episomes in a cervical carcinoma cell line. J. Virol. 63, 782–789.

Cooper, K., Herrington, C.S., Stickland, J.E., Evans, M.F., McGee, J.O., 1991. Episomal and integrated human papillomavirus in cervical neoplasia shown by non-isotopic in situ hybridisation. J. Clin. Pathol. 44, 990–996.

Cricca, M., Morselli-Labate, A.M., Venturoli, S., Ambretti, S., Gentilomi, G.A., Gallinella, G., Costa, S., Musiani, M., Zerbini, M., 2007. Viral DNA load, physical status and E2/E6 ratio as markers to grade HPV16 positive women for high-grade cervical lesions. Gynecol. Oncol. 106, 549–557.

Cullen, A.P., Reid, R., Lorincz, A.T., 1991. Analysis of the physical state of different human papillomavirus DNAs in intraepithelial and invasive neoplasm. J. Virol. 65, 606–612.

De Andrea, M., Mondini, M., Azzimonti, B., Dell'Oste, V., Germano, S., Gaudino, G., Musso, T., Landolfo, S., Gariglio, M., 2007. Alpha- and beta-papillomavirus E6/E7 genes differentially modulate pro-inflammatory gene expression. Virus Res. 124, 220–225.

de Boer, M.A., Jordanova, E.S., Kenter, G.G., Peters, A.A., Corver, W.E., Trimbos, J.B., Fleuren, G.J., 2007. High human papillomavirus oncogene mRNA expression and not viral DNA load is associated with poor prognosis in cervical cancer patients. Clin. Cancer Res. 13, 132–138.

Duensing, S., Munger, K., 2002. The human papillomavirus type 16 E6 and E7 oncoproteins independently induce numerical and structural chromosome instability. Cancer Res. 62, 7075–7082.

Franceschi, S., Rajkumar, R., Snijders, P.J., Arslan, A., Mahé, C., Plummer, M., Sankaranarayanan, R., Cherian, J., Meijer, C.J., Weiderpass, E., 2005. Papillomavirus infection in rural women in southern India. Br. J. Cancer 92, 601–606.

Franco, E.L., Coutlée, F., 2009. Prognostic value of measuring load of human papillomavirus DNA in cervical samples: an elusive target. J. Natl. Cancer Inst. 101, 153–161.

Gallo, G., Bibbo, M., Bagella, L., Zamparelli, A., Sanseverino, F., Giovagnoli, M.R., Vecchione, A., Giordano, A., 2003. Study of viral integration of HPV-16 in young patients with LSIL. J. Clin. Pathol. 56, 532–536.

Gillitzer, E., Chen, G., Stenlund, A., 2000. Separate domains in E1 and E2 proteins serve architectural and productive roles for cooperative DNA binding. EMBO J. 19, 3069–3079.

Häfner, N., Driesch, C., Gajda, M., Jansen, L., Kirchmayr, R., Runnebaum, I.B., Dürst, M., 2008. Integration of the HPV16 genome does not invariably result in high levels of viral oncogene transcripts. Oncogene 27, 1610–1617.

Hernandez-Hernandez, D.M., Ornelas-Bernal, L., Guido-Jiménez, M., Apresa-Garcia, T., Alvarado-Cabrero, I., Salcedo-Vargas, M., Mohar-Betancourt, A., Garcia-Carranca, A., 2003. Association between high-risk human papillomavirus DNA load and precursor lesions of cervical cancer in Mexican women. Gynecol. Oncol. 90, 310–317.

Jaquelline, G., Oliveira, Colf, Leremy A., McBride, Alison A., 2006. Variations in the association of papillomavirus E2 proteins with mitotic chromosomes. Proc. Natl. Acad. Sci. U. S. A. 103, 1047–1052.

Jeon, S., Lambert, P.F., 1995. Integration of human papillomavirus type 16 DNA into the human genome leads to increased stability of E6 and E7 mRNAs: implications for cervical carcinogenesis. Proc. Natl. Acad. Sci. U. S. A. 92, 1654–1658.

Josefsson, A.M., Magnusson, P.K., Ylitalo, N., Sørensen, P., Qwarforth-Tubbin, P., Andersen, P.K., Melbye, M., Adami, H.O., Gyllensten, U.B., 2000. Viral load of human papillomavirus 16 as a determinant for development of cervical carcinoma in situ: a nested case-control study. Lancet 355, 2189–2193.

Kalantari, M., Blennow, E., Hagmar, B., Johansson, B., 2001. Physical state of HPV16 and chromosomal mapping of the integrated form in cervical carcinomas. Diagn. Mol. Pathol. 10, 46–54.

Kozuka, T., Aoki, Y., Nakagawa, K., Ohtomo, K., Yoshikawa, H., Matsumoto, K., Yoshiike, K., Kanda, T., 2000. Enhancer–promoter activity of human papillomavirus type 16 long control regions isolated from cell lines SiHa and CaSki and cervical cancer biopsies. Jpn. J. Cancer Res. 91, 271–279.

Kristiansen, E., Jenkins, A., Holm, R., 1994. Coexistence of episomal and integrated HPV16 DNA in squamous cell carcinoma of the cervix. J. Clin. Pathol. 47, 253–256.

Kulmala, S.M., Syrjänen, S.M., Gyllensten, U.B., Shabalova, I.P., Petrovichev, N., Tosi, P., Syrjänen, K.J., Johansson, B.C., 2006. Early integration of high copy HPV16 detectable in women with normal and low grade cervical cytology and histology. J. Clin. Pathol. 59, 513–517.

Laikangbam, P., Sengupta, S., Bhattacharya, P., Duttagupta, C., Singh, T.D., Verma, Y., Roy, S., Das, R., Mukhopadhyay, S., 2007. A comparative profile of the prevalence and age distribution of human papillomavirus type 16/18 infections among three states of India with focus on northeast India. Int. J. Gynecol. Cancer 17, 107–117.

Matsukura, T., Kanda, T., Furuno, A., Yoshikawa, H., Kawana, T., Howley, P.M., 1986. Cloning of monomeric human papillomavirus type 16 DNA integrated within cell DNA from cervical carcinoma. J. Virol. 58, 979–982.

Munoz, N., Bosch, F.X., de Sanjose, S., Herrero, R., Castellsagué, X., Shah, K.V., Snijders, P.J., Meijer, C.J., 2003. Epidemiologic classification of human papillomavirus types associated with cervical cancer. N. Engl. J. Med. 348, 518–527.

Narayanan, S., Abraham, P., Peedicayil, A., Sridharan, G., Shaji, R.V., Chandy, G., 2004. E2 sequence variations of HPV 16 among patients with cervical neoplasia seen in the Indian subcontinent. Gynecol. Oncol. 95, 363–369.

Peitsaro, P., Johansson, B., Syrjänen, S., 2002. Integrated human papillomavirus type 16 is frequently found in cervical cancer precursors as demonstrated by a novel quantitative real-time PCR technique. J. Clin. Microbiol. 40, 886–891.

Roberts, I., Ng, G., Foster, N., Stanley, M., Herdman, M.T., Pett, M.R., Teschendorff, A., Coleman, N., 2008. Critical evaluation of HPV16 gene copy number quantification by SYBR green PCR. BMC Biotechnol. 8, 57.

Saunier, M., Monnier-Benoit, S., Mauny, F., Dalstein, V., Briolat, J., Riethmuller, D., Kantelip, B., Schwarz, E., Mougin, C., Prétet, J.L., 2008. Analysis of human papillomavirus type 16 (HPV16) DNA load and physical state for identification of HPV16 infected women with high grade lesions or cervical carcinoma. J. Clin. Microbiol. 46, 3678–3685.

Scheurer, M.E., Dillon, L.M., Chen, Z., Follen, M., Adler-Storthz, K., 2007. Absolute quantitative real-time polymerase chain reaction for the measurement of human papillomavirus E7 mRNA in cervical cytobrush specimens. Infect. Agent Cancer 2, 8.

Schlecht, N.F., Kulaga, S., Robitaille, J., Schlecht, N.F., Kulaga, S., Robitaille, J., Ferreira, S., Santos, M., Miyamura, R.A., Duarte-Franco, E., Rohan, T.E., Ferenczy, A., Villa, L.L., Franco, E.L., 2001. Persistent human papillomavirus infection as a predictor of cervical intraepithelial neoplasia. J. Am. Med. Assoc. 286, 3106–3114.

Sowjanya, A.P., Jain, M., Poli, U.R., Padma, S., Das, M., Shah, K.V., Rao, B.N., Devi, R.R., Gravitt, P.E., Ramakrishna, G., 2005. Prevalence and distribution of high-risk human papilloma virus (HPV) types in invasive squamous cell carcinoma of the cervix and in normal women in Andhra Pradesh. India. BMC Infect. Dis. 5, 116.

Swan, D.C., Tucker, R.A., Tortolero-Luna, G., Mitchell, M.F., Wideroff, L., Unger, E.R., Nisenbaum, R.A., Reeves, W.C., Icenogle, J.P., 1999. Human papillomavirus (HPV) DNA copy number is dependent on grade of cervical disease and HPV type. J. Clin. Microbiol. 37, 1030–1034.

Tachezy, R., Van Ranst, M.A., Cruz, Y., Burk, R.D., 1994. Analysis of short novel human papillomavirus sequences. Biochem. Biophys. Res. Commun. 204, 820–827.

Tan, S.H., Leong, L.E.C., Walker, P.A., Bernard, H.U., 1994. The human papillomavirus type 16 E2 transcription factor binds with low cooperativity to two flanking sites and represses the E6 promoter through displacement of Sp1 and TFIID. J. Virol. 68, 6411–6420.

van Duin, M., Snijders, P.J., Schrijnemakers, H.F., Voorhorst, F.J., Rozendaal, L., Nobbenhuis, M.A., van den Brule, A.J., Verheijen, R.H., Helmhorst, T.J., Meijer, C.J., 2002. Human papillomavirus 16 load in normal and abnormal cervical scrapes: an indicator of CIN II/III and viral clearance. Int. J. Cancer 98, 590–595.

Wang, S.S., Hildesheim, A., 2003. Viral and host factors in human papillomavirus persistence and progression. J. Natl. Cancer Inst. Monogr. 31, 35–40.

Wentzensen, N., Vinokurova, S., von Knebel Doeberitz, M., 2004. Systematic review of genomic integration sites of human papillomavirus genomes in epithelial dysplasia and invasive cancer of the female lower genital tract. Cancer Res. 64, 3878–3884.

Woodman, B.J.C., Collins, S.I., Young, L.S., 2007. The natural history of cervical HPV infection: unresolved issues. Nat. Rev. Cancer 7, 11–22.

Yasugi, T., Benson, J.D., Sakai, H., Vidal, M., Howley, P.M., 1997. Mapping and characterization of the interaction domains of human papillomavirus type 16 E1 and E2 proteins. J. Virol. 71, 891–899.

zur Hausen, H., 2002. Papillomaviruses and cancer: from basic studies to clinical application. Nature 2, 342–350.