

RESEARCH ARTICLE

Study of association and molecular analysis of human papillomavirus in breast cancer of Indian patients: Clinical and prognostic implication

Saimul Islam¹, Hemantika Dasgupta¹, Anirban Roychowdhury¹, Rittwika Bhattacharya¹, Nupur Mukherjee¹, Anup Roy², Gautam Kumar Mandal³, Neyaz Alam⁴, Jaydip Biswas⁴, Shyamsundar Mandal⁵, Susanta Roychoudhury⁶, Chinmay Kumar Panda^{1*}

1 Department of Oncogene Regulation, Chittaranjan National Cancer Institute, Kolkata, West Bengal, India, **2** Department of Pathology, North Bengal Medical College and hospital, Sushruta Nagar, Darjeeling, West Bengal, India, **3** Department of pathology, Chittaranjan National Cancer Institute, Kolkata, West Bengal, India, **4** Department of Surgical Oncology, Chittaranjan National Cancer Institute, Kolkata, West Bengal, India, **5** Department of Epidemiology & Biostatistics, Chittaranjan National Cancer Institute, Kolkata, West Bengal, India, **6** Saroj Gupta Cancer Centre and Research Institute, Thakurpukur, Kolkata, India

* ckpanda.cnci@gmail.com



OPEN ACCESS

Citation: Islam S, Dasgupta H, Roychowdhury A, Bhattacharya R, Mukherjee N, Roy A, et al. (2017) Study of association and molecular analysis of human papillomavirus in breast cancer of Indian patients: Clinical and prognostic implication. PLoS ONE 12(2): e0172760. doi:10.1371/journal.pone.0172760

Editor: Javier S. Castresana, University of Navarra, SPAIN

Received: January 1, 2017

Accepted: February 9, 2017

Published: February 28, 2017

Copyright: © 2017 Islam et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by University Grant Commission Fellowship grant F.2-3/2000 (SA-I) (Sr. No. 2061030813, Ref. No.: 20-06 / 2010) to Mrs. H. Dasgupta and grants from Council of Scientific and Industrial Research, Government of India [no. 60(0111)/14/EMR-II of dt 03/11/2014] to Dr. C. K. Panda. The funders had

Abstract

Objectives

Human papillomavirus (HPV) causes tumors primarily Cervical cancer. Recently, inconsistent reports came up in Breast cancer (BC) too. In India, despite treatment 70,218 BC patients die each year. So, we explored the association of HPV, if any, with BC prognosis in Indian pre-therapeutic (PT) and Neo-adjuvant chemotherapy (NACT) patients with subsequent analysis of HPV profile.

Methods

HPV prevalence was checked and analysis of physical status, copy number, genome variation, promoter methylation and expression (mRNA and protein) of the prevalent subtype was done.

Results

High prevalence of HPV was observed in both PT (64.0%) and NACT (71.0%) cases with significant association with younger (20–45 yrs) PT patients. Interestingly, HPV infection was significantly increased from adjacent normal breast (9.5%, 2/21), fibro adenomas (30%, 3/10) to tumors (64.8%, 203/313) samples. In both PT and NACT cases, HPV16 was the most prevalent subtype (69.0%) followed by HPV18 and HPV33. Survival analysis illustrated hrHPV infected PT patients had worst prognosis. So, detailed analysis of HPV16 profile was done which showed European-G350 as the most frequent HPV16 variant along with high rate of integration. Moreover, low copy number and hyper-methylation of P97 early promoter were concordant with low HPV16 E6 and E7 mRNA and protein expression. Notably,

no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abbreviations: hrHPV, High risk HPV; BC, Breast Cancer; PT, Pre-therapeutic; NACT, Neo-adjuvant chemotherapy treated; P97, Early promoter.

four novel variations (KT020838, KT020840, KT020841 and KT020839) in the LCR region and two (KT020836 and KT020837) in the E6 region were identified for the first time along with two novel E6^{E7*1} (KU199314) and E6^{E7*II} (KU199315) fusion transcript variants.

Conclusion

Thus, significant association of hrHPV with prognosis of Indian BC patients led to additional investigation of HPV16 profile. Outcomes indicated a plausible role of HPV in Indian BC patients.

Introduction

Human papillomavirus (HPV) is a DNA virus having a 9.2Kb genome. The high-risk (hr) subtypes were frequently associated with different cancers, primarily cervical cancer (CACX) and Head and neck squamous cell carcinoma (HNSCC) [1]. Among the subtypes of hrHPV, HPV16 was the most prevalent subtype in these tumors [2]. It is well known that HPV16 mainly infects the basal epithelial stem cell through break of stratified epithelial layer [3]. The transformation HPV depends on expression of E6/E7 oncoproteins which contributes to the process of carcinogenesis by increasing cellular proliferation leading to more genomic instability and inhibition of apoptosis [4]. The expression of both the oncoproteins is controlled by viral protein E2, which often gets abrogated due to viral integration in the host genome through the E2 region [5]. In addition, expression of E6 and E7 has been regulated by the long control region (LCR) harboring different transcription factors binding sites along with the viral copy number variation though the activity of P97 promoter and enhancer [6, 7]. The activity of this promoter and enhancer were further controlled by methylation though host DNA Methyltransferase enzymes [8]. Moreover, it was also reported that RNA splicing may control the expression of these oncoproteins as well [9].

In recent times, several worldwide reports have associated HPV with Breast cancer (BC) while other investigators have strongly negated it. A current meta-analysis encompassing 21 different studies had also revealed this inconsistency in prevalence of high risk HPV16 (ranging from 3% to 92%), HPV18 (2% to upto 100%) and HPV33 (4% to 70%) [10]. HPV16 infection (40%, 4/10) was also reported in neo-adjuvant chemotherapy treated (NACT) samples [11]. In addition, HPV was also detected in normal (1–29%) and benign (13.7–55%) breast tissues [12–14]. It was further reported that the E6 and E7 oncoproteins of HPV16 could immortalize human mammary epithelial cells indicating their importance in cellular transformation [15]. Previous studies reported frequent integration of HPV16 in BC patients, whereas, contradictory reports of low viral copy number and reduced expression of HPV16 oncoproteins E7 (protein) has also been found [14, 16]. Interestingly, BC patients with HPV16 infection have been reported to show better prognosis [17]. So, association of HPV with BC still remains controversial.

On the other hand, in India, where 144,937 BC cases were newly diagnosed and 70,218 BC patients die per year in spite of therapy, variable frequencies (from negative to up to 26.5%) of HPV infection in BC have been found [18, 19]. But till now no in depth study to analyze the association of HPV with BC has been carried out.

Thus, in the present study, attempts have been made to analyze the association of HPV, if any, in BC of Indian patients. At first, we analyzed the prevalence of HPV in 272 pre-therapeutic and 41 NACT BC of Indian patients followed by prevalence of high risk HPV16, 18 and 33.

Our study showed high frequency (63.9%-71%) of HPV infection in the both pre-therapeutic and NACT samples with highest prevalence of HPV16. PT patients with hrHPV infection showed worst prognosis as well. So, further analyses were done to decipher the genetic and epigenetic status of HPV16. Finally, integrated genome, P97 promoter methylation, low viral copy number along with reduced expression (mRNA/protein) of its oncoproteins E6/E7 were the attributes of HPV16, frequently encountered in Indian BC patients.

Materials and methods

Patient population, tumor tissues and cell lines

A total of 272 freshly operated pre-therapeutic and 41 neo-adjuvant chemotherapy treated breast tumor specimens along with 21 adjacent normal tissues, 10 fibro adenomas and 7 benign phyllode were collected from Chittaranjan National Cancer Institute(CNCI), Kolkata, India, after appropriate approval of Institutional ethical committee and written consent from individuals. After surgery, some portion of the operated specimen was collected by the surgeons in sterile pots in the operating room. These were then cut into pieces by sterile surgical blades. One part of the tissues were stored in 10% formalin for immunohistochemistry (IHC) analysis and the remaining part stored at -80°C in isolated rack until further use. In addition, standardized precautions were taken to avoid contamination during tissue preparation. For RNA analysis some of the pre-therapeutic samples ($N = 10$) were kept in RNAlater[®] (Ambion, USA) solution which is separately stored. Among these samples 56.8% (178/313) patients were below the average age of 45.6 years and 43.1% (135/313) were above the average age (S1 Table). As samples were collected before and after therapy, we have designated as pre-therapeutic and chemotherapy treated samples (NACT). Breast cancer cell line MCF7 and cervical cancer cell line SiHa were obtained from National Centre for Cell Science, Pune, India (S1 Fig).

Isolation of DNA from tumor samples

The contaminant normal cells in the breast lesions were removed by micro-dissection procedure [20] to enrich (>80%) tumor cells for isolation of high-molecular weight DNA by proteinase-K digestion followed by phenol-chloroform extraction [21]. Separate sterile work place was used and every feasible precaution was additionally taken so that no contaminations from surroundings occur.

HPV detection and sub-typing

The presence of HPV in BC was detected by PCR using primers (MY09 and MY11) from the consensus L1 region [22]. Typing of HPV-16/18/33 in the L1-positive samples was done by PCR using specific primers (S2 Table) The HPV16, HPV18 and HPV33 plasmids were used as positive control for the respective cases [22, 23]. Further validation was done by Southern hybridization using ^{32}P -labeled HPV16 and HPV18 type-specific probes [24]. All PCR reactions were done in clean restricted work surface where other laboratory works were not performed.

Determination of HPV16 physical status

The physical status of HPV16 genome was determined by multiplex PCR using three different primer sets (S2 Table) and the ratio of E2/E6 copy numbers was calculated as described previously [25]. For episomal and integration control, HPV16 plasmid and SiHa (HPV16 positive) were used respectively. Further validation was done in 23 randomly selected samples by real time PCR using Power SYBR Green (Applied Biosystems, USA) with the same primer sets and controls [26].

Sequencing analysis of HPV16 LCR, E6 and E7 regions

Using two set of primers (S2 Table), LCR and E6-E7 regions were sequenced in both sense and antisense directions with a BigDye Terminator Cycle Sequencing 3.1 Kit (Applied Biosystems, USA) according to the manufacturer's instructions in a 3130xl Genetic Analyzer (Applied Biosystems, USA) [25]. Designation of phylogenetic clusters of HPV16 E6 region was done according to Yamada et al [27].

Methylation analysis of LCR

The methylation status of the p97 promoter and enhancer region was analyzed by methylation sensitive restriction enzyme analysis (MSRA) using enzymes HpaII and HhaI for digestion of P97 promoter and enhancer respectively [28]. The 445 bp fragment of β -3A adaptin gene (K1) and 229 bp fragment of RAR β 2 (K2) were used as digestion and integrity controls [29].

Estimation of HPV16 copy number

The copy numbers of HPV16 in the samples were determined by TaqMan absolute Real-Time PCR method using specific probe (S2 Table) as described previously [30].

Quantification of HPV16 E6 and E7 transcripts

Total RNA was extracted from HPV16 positive breast tumors and the two cell lines MCF7 and SiHa [31]. Relative quantification of E6/E7 expression was performed using a power SYBR-green assay (Applied Biosystems, USA) with β 2-microglobulin as endogenous control [25]. The RT-PCR products were electrophoresed in 2% agarose gel for detection of different splice products of E6/E7 mRNA and were further characterized by Sanger sequencing using specific primers (S2 Table).

Detection of HPV16 E6 and E7 proteins expression

The goat polyclonal antibody for E6 (sc-1584) and mouse monoclonal antibody for E7 (sc-6981) were used at a dilution of 1:200. The HRP conjugated rabbit anti-goat IgG (sc-2768) and goat anti-mouse IgG (sc-2005) were used at a dilution of 1:500 as secondary antibody in respective cases, followed by scoring as mentioned previously [30]. The antibodies were purchased from M/s Santa Cruz Biotechnology, CA, USA.

Bio-informatics analysis for E6 and E7 mRNA and protein stability

Transcription factor prediction was done using on line server *Alibaba 2.1* [32]. The protein sequence and half-life of different splice variants of E6/E7 transcript were predicted by transcription-translation tool and *ProtLifePred* wave server respectively [33]. The stability of E6 variants in both mRNA and protein were predicted by *mFOLD* and *I-Mutant 2.0* followed by validation in *Proven* wave server [34–36].

Statistical analysis

Chi-square analysis was used to determine association between HPV profiles and different clinicopathological parameter of tumors using Epi Info 6.04. Survival analysis (upto 5 years) was performed by Kaplan–Meier method using SPSS 10.0. P-value ≤ 0.05 was considered statistically significant in all analysis.

Results

Prevalence of HPV in breast tissues

Using primers MY09/MY11, total HPV prevalence was detected in 63.9% (174 / 272) in pre-therapeutic and 71.0% (29/41) in NACT BC samples (Fig 1A and 1B). Among the HPV positive pre-therapeutic BC samples, the prevalence of HPV16, HPV18 and HPV33 were in the following order: 69.0% (120/174) > 35.0% (61/174) > 2.9% (5/174) respectively (Fig 1C, 1D, 1E & 1F and S2A & S2B Fig). The co-infection of HPV16 and HPV18 was seen in 25.2% (44/174) samples. About 18.3% (32/174) samples showed other than HPV16, HPV18 and HPV33 infection. Similar trend of the HPV subtypes infection was seen in NACT BC samples with 69.0% (20/29) HPV16, 65.0% (19/29) HPV18, 3.4% (1/29) HPV33 and 6.9% (2/19) other HPV subtypes (Fig 1F). About 48.2% (14/29) samples showed co-infection of HPV16 and HPV18, and 5% (1/19) samples showed co-infection of HPV16, HPV18 and HPV33. The breast cancer cell line MCF7 showed HPV16 infection (Fig 1A, 1C & 1D and S2A & S2B Fig). However, only 9.5% (2/21) adjacent normal breast tissues showed HPV infection although their tumor tissues were HPV positive. In addition, fibro adenomas and phyllode samples showed 30% (3/10) and 71.5% (5/7) HPV infection. Surprisingly, the HPV infection was significantly ($p = 0.001$) increased during progression from normal to tumor (Table 1). Interestingly, significant high HPV infection was seen within average value of age of onset (45 yrs) of pre-therapeutic BC samples ($p = 0.02$) (S3 Table). However, no such correlation was seen in NACT samples. On the other hand, no significant correlation of total HPV/different HPV subtypes infection were seen in different clinicopathological parameters in univariate and multivariate analysis of pre-therapeutic BC samples (S3 Table). Similar trend was also seen in NACT samples.

The Kaplan–Meier survival analysis showed significantly ($p = 0.04$) poor survival of hrHPV positive than overall HPV negative pre-therapeutic patients (Fig 2A). However, no such significance has been observed in NACT cases (Fig 2B). Similar trend in survival was seen when overall HPV positive pre-therapeutic and NACT samples were compared with their respective

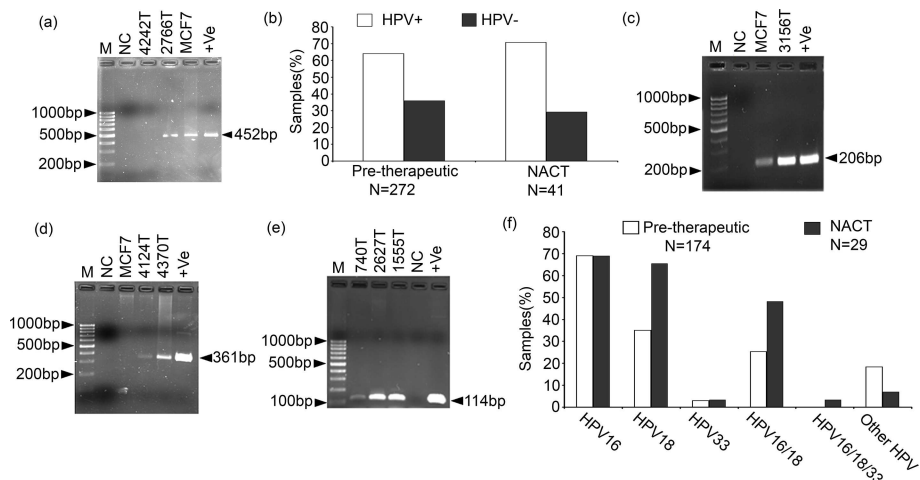


Fig 1. Determination of HPV prevalence in primary breast tumors and breast cancer cell line MCF7. (a) Representative agarose gel showing positive HPV infection in pre-therapeutic sample 2766T and MCF7 cell line while, neo-adjuvant chemotherapy treated (NACT) sample 4242T is HPV negative. +Ve means Positive control having HPV16 plasmid. (b) Frequency of HPV infection in BC samples. Samples were identified as positive when PCR bands were seen with respective subtype specific primers. Representative agarose gel of (c) HPV16 (d) HPV18 (e) HPV33 detection PCR. [M: 100bp marker, NC represent Negative control with no DNA, +Ve represent Positive control having HPV16, HPV18 and HPV33 plasmid in their respective subtype detection gels] (f) Distribution of different genotypes of HPV in different type of BC samples.

doi:10.1371/journal.pone.0172760.g001

Table 1. Comparative HPV infection in adjacent normal breast, fibro adenomas, benign phyllode and breast tumors.

Sample Type	HPV positive (%)	HPV negative (%)	P value
Normal n = 21	2(9.5)	19(90.5)	0.001*
Fibro adenomas n = 10	3(30.0)	7(70.0)	
Benign phyllode n = 7	5(71.5)	2(28.5)	
Breast cancer n = 213	203(64.8)	110(35.2)	

* indicate significant correlation ($P \leq 0.05$).

doi:10.1371/journal.pone.0172760.t001

HPV negative group, though it was not statistically significant ($p = 0.05$) (S3A and S3B Fig). As hrHPV infection showed significant correlation with worst BC prognosis, so detailed genetic (physical status, copy number variation and sequence variation) and epigenetic (P97 promoter and enhancer methylation) profiles of HPV16 were performed in BC samples.

Physical status of HPV16 in BC

In pre-therapeutic BC samples, significantly high HPV16 integration (87.5%, 105/120) was seen compared to the episomal form (4.2%, 5/120) ($p = 0.01$) (Fig 3A, 3B, 3C and 3D). Both episomal and integrated forms i.e mixed form, was evident in 8.3% (9/120) samples (Fig 3D). Similar trend was seen in NACT samples with 81.0% (17/21) integrated form, 4.8% (1/21) episomal form and 14.3% (3/21) mixed form (Fig 3D). Significant concordance ($p = 0.75$) was evident between the multiplex PCR and the Real Time PCR method of analysis of HPV16 physical status (data not shown). Interestingly, HPV16 integration was seen mainly in the hinge region of E2 in 97.0% (102/105) pre-therapeutic and 100.0% (17/17) NACT samples (Fig 3E and 3F). In MCF7, HPV16 was present in an episomal form with insertion/ deletion in the N-terminal region due to the presence of two altered PCR products of E2A primer (Fig 3A, 3B and 3C). The HPV16 physical status showed no significant correlation with different clinicopathological parameters due to high prevalence of the integrated form (data not shown).

Sequence variation analysis of LCR and E6-E7 region of HPV16 in BC

The total sequence variation was seen in 70.8% (34/48) pre-therapeutic BC samples with 21 variant in the LCR region, 11 variant at E6 and 3 variant at E7 regions (Fig 4A & 4B). In LCR region, scattered sequence variation was evident with high 62.5% (30/48) 7521G>A transition

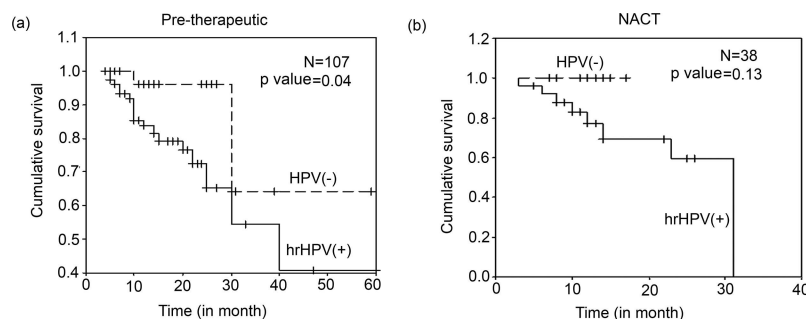


Fig 2. Kaplan-Meier 5-year survival probability curves with cumulative survival of breast cancer patients (BC) based on hrHPV status. Survival probability of BC patients in (a) pre therapy group ($p = 0.04$) and (b) neo-adjuvant chemotherapy treated (NACT) group ($p = 0.13$). [N: total number of samples included in this study. hrHPV(+) indicates infection by either HPV16, HPV18 or HPV33].

doi:10.1371/journal.pone.0172760.g002

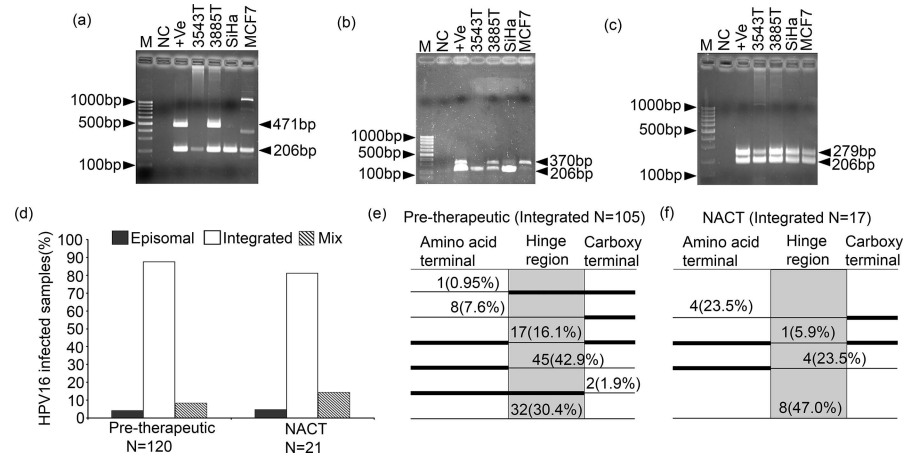


Fig 3. Analysis of HPV16 physical status in breast tumor samples and MCF7 breast cancer cell line. Representative agarose gel of physical status of HPV16 genome at (a) E2A region. (b) E2B region and (c) E2C region (d) Histogram represent significant high frequency of integrated viral genome in both pre-therapeutic and NACT samples ($p \leq 0.01$). Frequency of integration at three regions of the E2 gene in (e) pre-therapeutic cases and (f) NACT samples. [Here M: 100bp marker, NC represent Negative control with no DNA, +Ve represent Episomal control where HPV16 plasmid was used. SiHa is the HPV16 positive Cervical cancer cell line used as Integration control for E2A and E2B region while Episomal control for E2C region]

doi:10.1371/journal.pone.0172760.g003

in the samples (Fig 4A). The majority of sequence variations were overlapped with different transcription factors binding sites like YY1, GRE1, Oct-1 etc (Fig 4A). Five samples showed 788 6C>G transversion in the replication origin containing YY1 repressor binding site. Four novel variants at 7628A>T (KT020838), 7800A>G(KT020840), 7837A>G(KT020841) and 7839A>C (KT020839) were identified in 7 samples (Fig 4A). The variant at 7839A>C (KT020839) was seen in four samples and other variants were seen in one sample each (Fig 4A). These novel variants overlapped with Oct-1 and Pit-1a transcription factors binding sites (Fig 4A).

In E6 region, the sequence variations were evident in flanking of the two zinc finger domains with high frequency (70.0%, 34/48) of 350T>G transversion followed by 18.7% (9/48) variation at 286T>A and 289A>G sites (Fig 4B). Two novel variations were identified at 395G>T (KT020836)(2.0%, 1/48) and 519G>A (KT020837) (2.0%, 1/48) sites (Fig 4B). Majority 54.5% (6/11) of the variations affected biological function of E6 protein. In the E7 region, the sequence variation was seen mainly in the zinc finger 3 domain with 19.0% (9/48) at 795T>G and 14.5%(7/48) at 789T>C sites (Fig 4B). Seven samples showed both 795T>G and 789T>C variations. In addition, one sample showed variation in flanking of the zinc finger domain at 732T>C. After compilation of the sequence variation at E6 and E7 regions, it was evident that the European lineage with T350G variant (E-G350) (1 of which was E-G350 G519A) was frequent 39.6% (19/48) followed by 29.0% (14/48) of European prototype (Ep), 19.0% (9/48) of North American 1 (NA1), 8.0% (4/48) of North American 1 variant (NA1A336G), 2.0% (1/48) of Asian American variant (AA G395T) and 2.0% (1/48) of African1 (AF1).

Analysis of LCR region methylation of HPV16

Significant high methylation (96.7%, 30/31) was seen in P97 promoter region than the enhancer region (51.6%, 16/31) ($p = 0.004$) (Fig 4C, 4D & 4E). The methylation in both enhancer and promoter regions were seen in 48.0% (15/31) of the samples (Fig 4E). The overall LCR region methylation showed no association with different clinicopathological parameters (data not shown). The genetic variations seen in the LCR region (Fig 4A) were not overlapped with the P97 promoter and enhancer methylation sites (nt58 and nt7672).

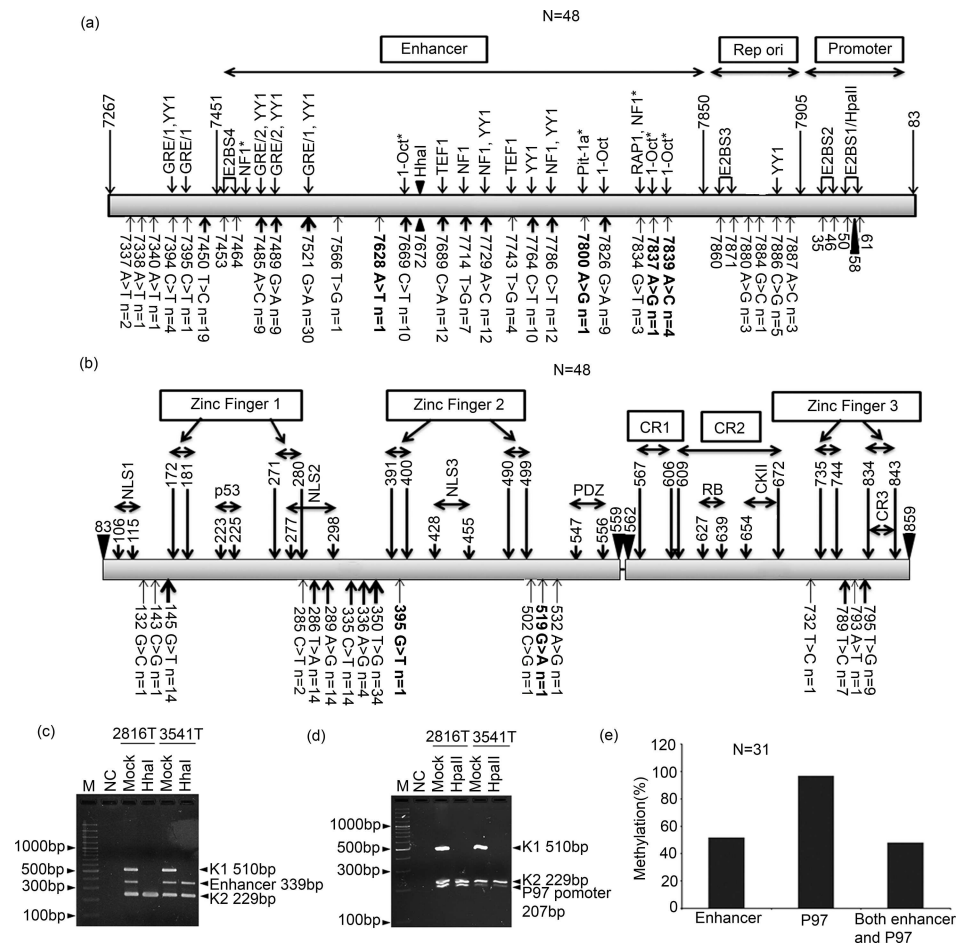


Fig 4. Sequence variation of LCR, E6 and E7 region and methylation status of LCR of HPV16 in pre-therapeutic breast tumor cases. Schematic representation of identified sequence variations in (a) the LCR region and (b) the E6 and E7 gene. Here, bold lettered variants were novel ones. Upper head arrow thickness indicate the frequency of the given variation where “N” represent number of samples. In LCR, lower head arrows indicates binding of the transcription factors while in E6 and E7 gene it highlighted different protein domains. Methylation sites in the LCR were also marked (Bold triangle). “*” indicates transcription factors predicted by Alibaba 2.1 TF Binding Prediction software. Representative agarose gel showing methylation status at (c) the enhancer region (d) the P97 promoter region. (e) Histogram showed significant high frequency methylation in P97 promoter than enhancer region ($p = 0.004$). [M: 1000bp marker, NC represent Negative control with no DNA, K1 used as digestion control, Mock: mock digestion without enzyme, HhaI: DNA digested with HhaI enzyme, HpaII: DNA digested with HpaII enzyme].

doi:10.1371/journal.pone.0172760.g004

Analysis of copy number of HPV16 in pre-therapeutic BC

The wide variation of HPV16 copy number (0.58–1044.6 copies/ 50ng gDNA) with the median value of 9.3 copies/ 50ng gDNA was observed in BC (S4A Table). There was gradual increase in HPV16 copy numbers with progressive cytological severity, stage and nodes at pathology (S4B Table). Among the different HPV16 lineages, the E-G350 lineage showed high copy number (S4B Table).

Expression analysis E6 and E7 of HPV16

mRNA. There were wide variations in expressions of E6, E7 and E6/E7 with mean value of 9.03, 9.6 and 8.7 respectively in different BC samples and the cell lines MCF7 and SiHa (Fig 5A–5E). Different splice variants of E6 and E6/E7 such as E6**I*, E6**II*, E6**I*/E7, E6**II*/E7 and

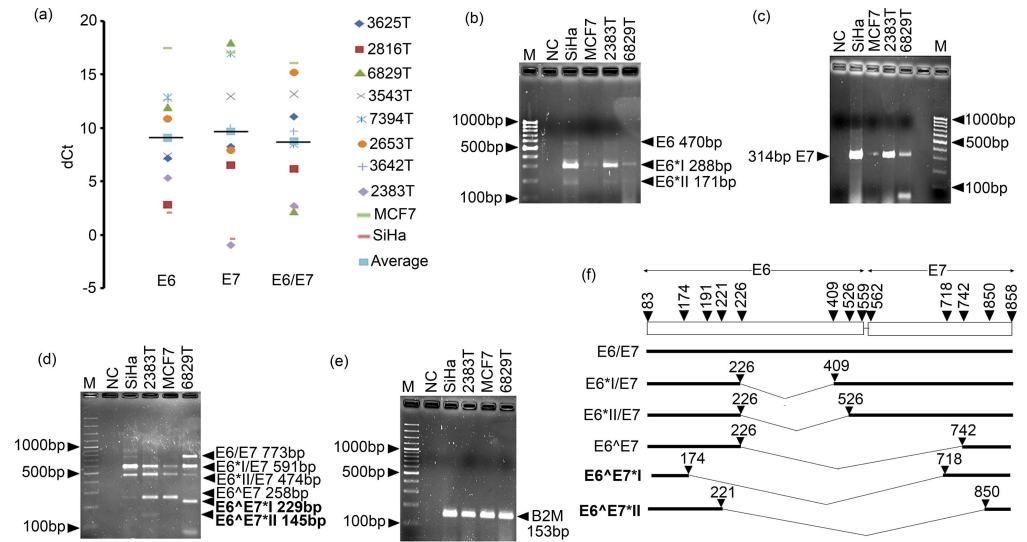


Fig 5. Detection and quantification of E6 and E7 mRNA of HPV16 in pre-therapeutic breast tumor samples and MCF7 breast cancer cell line. (a) The dot plot showed wide variation of dCt value of E6, E7 and full length E6/E7 mRNA expression. High dCt value indicates low expression and vice versa. Here, SiHa cell line was used as positive control for E6 and E7 expression. Bold horizontal line represent average dCt value. (b) Representative agarose gel of post real time PCR of E6 showed different splice forms of E6. (c) Agarose gel represents post real time PCR of E7. (d) Representative agarose gel of post real time PCR showed full length E6/E7 mRNA with its different splice forms. Bold lettered represent novel splice form. (e) Representative gel showed post real time PCR of β 2-microglobulin used as endogenous control gene. (f) Schematic diagram of different splice forms of E6 and E6-E7 transcripts. Bold line represent exon and V-shaped thin line represents intron. Nucleotide position in HPV16 genome was also indicated. Number of left and right arm represent splicing donor and acceptor sites respectively. [Here M: 1000bp marker; NC represent Negative control with no cDNA].

doi:10.1371/journal.pone.0172760.g005

E6[^]E7, were seen in different BC samples (Fig 5A & 5F and S4A & S4B Fig). Among the splice variants, E6^{*}I/E7 transcript was the most prevalent (8/8). Two novel fusion transcripts E6[^]E7^{*}I and E6[^]E7^{*}II in two samples were sequenced and registered in Gene Bank with accession numbers KU199314 and KU199315 respectively (Fig 5F and S4C & S4D Fig). The bio-informatics analysis predicted that the novel E6 variant 519G>A (KT020836) seen in E-G350 lineage decreased the stability of both E6/E7 and E6^{*}I/E7 mRNAs, whereas another novel E6 variant 395G>T (KT020837) seen in AA lineage decreased stability of E6/E7 mRNA only. In addition, low stability of two novel transcripts E6[^]E7^{*}I and E6[^]E7^{*}II were predicted than full length E6/E7 transcript using mFOLD wave server (data not shown). Moreover, transcription—translation tool and ProtLifePread server predicted that only E6[^]E7^{*}II transcript could produce E6/E7 fusion protein with comparable half-life of E6[^]E7 transcript (data not shown).

Protein. In immunohistochemical analysis, nuclear expression of E6 and E7 was seen in the HPV16 positive samples (Fig 6A–6F). About 53.3% (16/30) samples showed low E6 and E7 protein expression. The protein expression of both E6 and E7 showed concordance with respective mRNA expression (S5 Table). The bio-informatics analysis (I-Mutant 2.0 server) predicted that the novel E6 variant of 519 G>A of E-G350 lineage might have low protein stability than another novel E6 variant 395G>T in AA lineage with increased stability (data not shown).

Discussion

The primary aim of the study was to understand the association of HPV infection with BC. As hrHPV infection showed statistically significant correlation with worst prognosis of BC

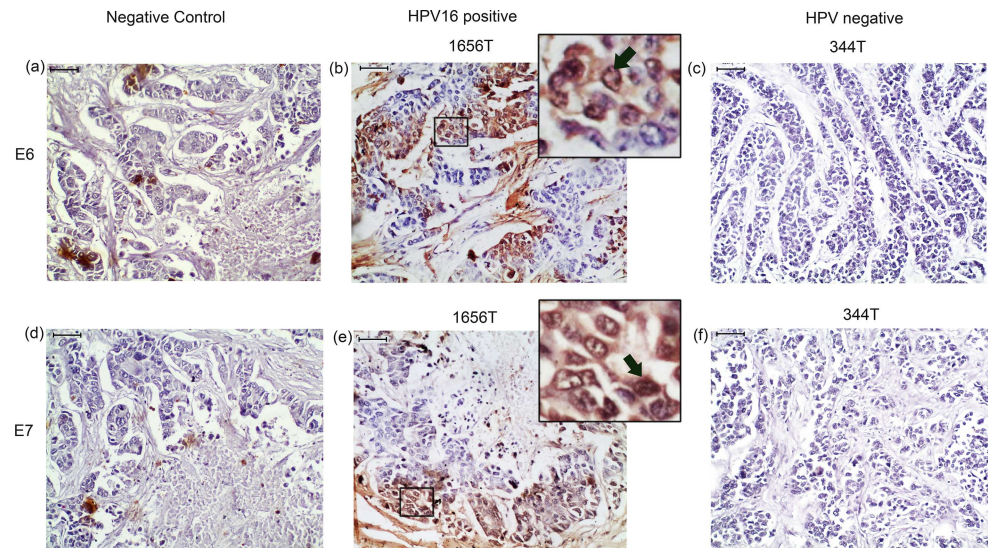


Fig 6. Immunohistochemical detection of E6 and E7 protein of HPV16 in pre-therapeutic breast tumor tissues. (b) & (e) Representative immunohistochemical staining of E6 and E7 in HPV16 positive samples. (c) & (f) Representative immunohistochemical staining in HPV negative sample (a) & (d) Immunohistochemical staining with out primary antibody represented Negative control (NC). [Magnification of tissue samples is 20X and for inset, magnification is 40X, Scale bars = 50 μ m].

doi:10.1371/journal.pone.0172760.g006

patients so, further its genetic and epigenetic landscape of the most prevalent subtype (HPV16) was analyzed. At first, prevalence of HPV was analyzed in (Eastern) Indian BC patients ($n = 313$) and subsequently comprehensive analysis of HPV16 profile was performed. To the best of our knowledge such comprehensive analysis was done for the first time among Indian BC patients.

Our results showed frequent infection (63.9–71.0%) of HPV in pre-therapeutic and NACT BC samples, indicating plausible importance of HPV infection in development of this disease irrespective of therapy. The high HPV infection may be due to poor hygienic condition and malnutrition of patients together with ethnicity [19, 37]. Among the HPV positive samples, the prevalence of HPV16 was high (69.0%) followed by HPV18 (35.0%) and HPV33 (2.9%). For HPV detection, the gold-standard L1 consensus primers MY09/11 primers was used which can detect up to 40 mucosal HPV genotypes, including the most prevalent high risk (HPV 16 & 18) as well as the low risk (HPV 6 & 11) types [38]. Every precaution was taken to prevent any possible risk of contamination of PCR reactions [39]. Furthermore, the result was validated using Southern blot.

Similar to our data, high frequency (74.0%) of HPV infection was previously reported though there HPV33 was found to be the most prevalent subtype [40]. In contrast, investigators also informed moderate (21.0–48.0%), low (1.6–13.3%) or absence of HPV infection in BC with varied prevalence of HPV16 or HPV18 subtypes [10]. This ambiguity in frequency of HPV infection in BC might be due to differences in etiological factors, ethnicity and analysis methods [37]. Moreover, all necessary precautions were taken strictly while carrying out all the PCR reactions used for prevalence analysis (see [Materials and Methods](#)).

The presence of HPV in normal and benign tumor and sharp increase in malignant breast tumors indicates its pathological importance in breast cancer. Similar trend of HPV infection from normal to benign and tumor was also reported by Lawson et al [14]. Consistent with our data, recent meta-analysis showed that HPV infection was four fold increased in BC compared with normal breast tissues [41]. Similar to Lawson and Bae et al reports, in our study also high

HPV infection ($p = 0.02$) was found in lower age group (≤ 45 yrs) patients, which suggests that HPV infection, may aid in reducing the age of tumor onset [12]. Like earlier reports, no significant correlation of HPV infection was found with clinico-pathological parameters including, stage, grade, lymph node, age of onset and parity [17, 42, 43]. In addition, poor prognosis was found in hrHPV infected pre-therapeutic patients when compared with other groups ($p = 0.04$). Interestingly, no significant association with prognosis was seen when total HPV positive samples were considered similar to previous studies where no association of HPV-positive BC patients with clinical outcome was reported [11, 17]. It appears that BC patients with low-risk HPV infection may be at lesser risk unless they are co-infected or solely infected with hrHPV types though it is known that differences may occur due to ethnicity and sample size [37].

In HPV16 positive samples, HPV16 was integrated frequently (87.5%) in the host genome by disrupting the hinge region (E2B) of the E2 gene with gradual increase with stage. Such observation clearly indicates rise in virulence of HPV during progression of the tumor. Though frequent integration was seen previously in BC [11, 16] but only such disruption of E2B was reported in cervical carcinoma [44]. Presence of two altered PCR products of E2A region in MCF7 cell line, similar to cervical carcinoma, indicates plausible site of rearrangement of viral genome [45].

Final validation of HPV16 occurrence in BC samples was done using Sanger sequencing of the LCR and E6/E7 region of the HPV16 genome. Additionally, sequence variation analysis revealed 7521G>A as the major (62.5%) sequence variant in the LCR region. This site overlapped with the binding of transcriptional repressor, YY1, which already been reported to regulate E6/E7 expression [46]. Other than the common ones, four novel variants in the LCR regions (7628A>T, 7800A>G, 7837A>G and 7839A>C) was also observed, which coincided with the predicted Oct-1/Pit-1a transcription factors binding sites [47]. In E6 locus, the major (70.0%) variant was 350T>G along with the familiar variants [48]. This 350T>G was also significantly associated with pathogenesis of HPV [49]. Moreover, *in silico* analysis revealed that the two novel variants 519G>A and 395G>T appears to destabilize the E6 mRNA and protein stability. In E7 gene, two observed synonymous sequence variation of 795T>G and 789T>C sites were already been reported in cervical carcinoma [25].

It is known that in cervical carcinoma AA lineage show higher virulence than other lineages [50]. But among Indian BC patients, E-G350 showed highest frequency (39.6%) followed by Ep (29.0%), NA1 (19.0%), NA1 A336G (8.0%), AA G395T (2.0%) and Af1 (2.0%). It is known that genome integration is associated with promoter hypomethylation [28]. Although majority of BC samples showed integration but comparatively higher methylation (96.7%) of P97 promoter and enhancer region (51.6%) was seen. On the contrary, predominant hypomethylation of P97 and enhancer region were seen in the integrated form of HPV16 in cervical carcinoma and head and neck squamous cell carcinoma (HNSCC) [28, 51]. Overall these results display significant differences in the pathogenesis process of HPV in BC, arising may be due to dissimilar tissue differentiation and microenvironment status. In addition, the absence of sequence variation in P97 promoter and enhancer sites indicates that the nucleotide variation may have occurred by other mechanism rather than DNA deamination after methylation. It appears that overall the mode of pathogenesis of HPV is different from that of cervical cancer.

Similar to other investigators, [16, 43] low copy number of HPV genome was seen in the present study among Indian BC patients. Likewise, such reports were present in non-genital cancers (like esophageal cancer and HNSCC) along with HPV16 positive cervical cancer cell line SiHa [52, 53]. This suggests that low viral load may be sufficient to promote carcinogenesis, especially if the viral DNA is integrated into the host genome [26]. Like cervical cancer,

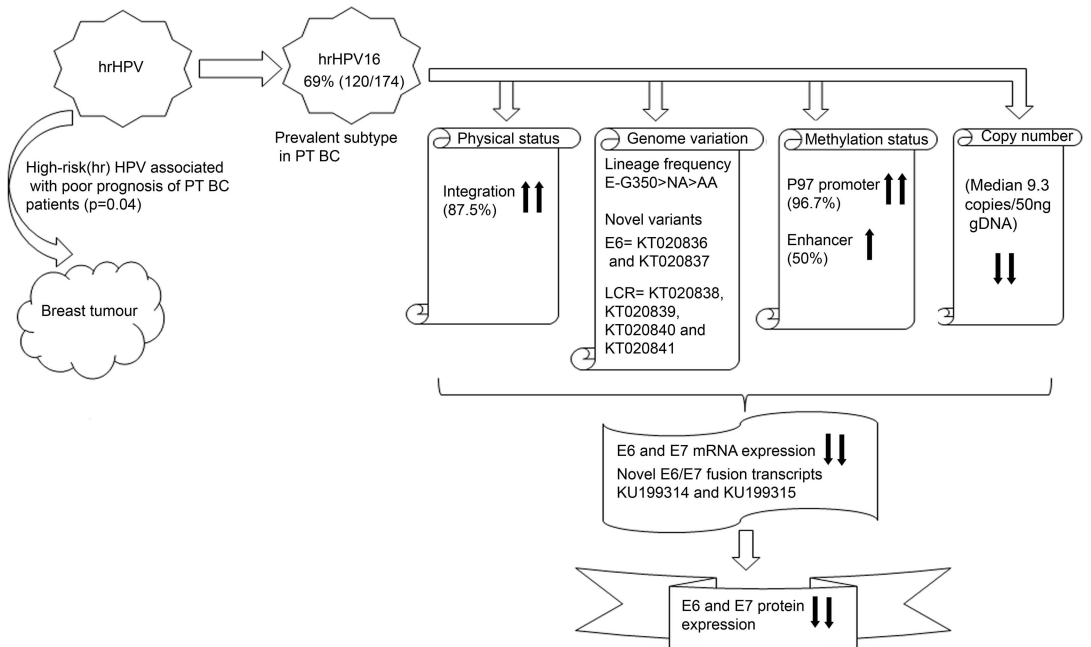


Fig 7. Schematic diagram represent summarised results of the current study. Here, upper head arrow indicates high, lower head arrow indicates low. PT represent pre-therapeutic BC. gDNA; Genomic DNA of patients.

doi:10.1371/journal.pone.0172760.g007

gradual increase in viral load with progression and severity of the disease was evident in the BC samples [22].

In our study, differential expression pattern of HPV16 E6, E7 and full length E6/E7 transcripts were seen. Though it was concordant with Lawson et al findings but, others author were unable to detect any E6 and E7 transcripts [14, 54]. This might be due to low copy number of viral genome and use of fresh tissue rather than paraffin embedded specimen used in previous analyses. Moreover, two spliced transcripts (E6^{*I} and E6^{*II}) of E6 and three spliced transcripts (E6^{*I}/E7, E6^{*I}/E7, E6[^]E7) of E6/E7 were detected in the BC samples as previously seen in cervical carcinoma [25]. Additionally, two novel fusion transcripts E6[^]E7^{*I} and E6[^]E7^{*II} of E6/E7 were observed. Among these, donor and acceptor splicing sites of E6[^]E7^{*II} transcript was previously predicted bioinformatically [55]. Lower stability of E6[^]E7^{*II} than E6[^]E7^{*I} transcripts but comparable half-life of its corresponding fusion proteins was further predicted *in silico*. Therefore, the identified HPV genome showed functional transcriptional activity in BC tissues. This was further concordant with the differential nuclear expression of E6 and E7 proteins in the BC samples similar to cervical carcinoma and HNSCC [14, 56].

A sum up of results was provided to give an overall picture of genetic (physical status, genome variations and copy number) and epigenetic (methylation status) spectrum of HPV16 in breast tumors (Fig 7). Frequent infection of hrHPV in BC along with significant association of hrHPV with poor prognosis suggests that HPV may be associated with BC in Indian (Eastern) population. When compared with CACX, high integration of HPV16 was found similar. But low copy number, hypermethylated p97 promoter and enhancer along with low E6 and E7 expression and novel fusion transcripts of E6/E7 may provide a clue for different mode of pathogenesis of HPV16 in BC. Thus, these outcomes indicated plausible involvement of HPV in BC. However, more study is needed to completely understand the role of HPV infection in development of BC.

Supporting information

S1 Table. Distribution of samples and different clinico-pathological parameters of pre-therapeutic and neo-adjuvant chemotherapy treated (NACT) breast cancer (BC) patients.
(DOC)

S2 Table. Detail information of primers used.
(DOCX)

S3 Table. Status of HPV infection associated with different clinico-pathological parameters of the pre-therapeutic and neo-adjuvant chemotherapy treated (NACT) BC patients.
(DOC)

S4 Table. (a) Median value of HPV16 copy number in BC samples. (b) Distribution of HPV16 copy number according to grade, stage, age, median age of onset and HPV16 lineage.
(DOC)

S5 Table. Concordance between mRNA and protein expression of E6 and E7 gene of HPV16.
(DOC)

S1 Fig. Utilization of pre-therapeutic and neo-adjuvant chemotherapy treated primary breast tumor samples. Schematic diagram represent work flow and distribution of samples in different experimental procedure. [‘N’ represent number of samples; IHC: immunohistochemistry]
(TIF)

S2 Fig. Validation of presence of HPV16 and HPV18 subtypes in breast tumor and MCF7 cell line by southern blot. Corresponding autoradiograph of (a) HPV16 (b) HPV18 detection agarose gel in NACT and pre-therapeutic sample. [Here M: 100bp marker, NC (Negative control) represent with out DNA, +Ve (positive control) represent HPV16, HPV18 and HPV33 plasmid in their respective subtypes]
(TIF)

S3 Fig. Kaplan–Meir 5-year survival probability curves with cumulative survival of breast cancer patients (BC) based on total HPV status. No statistically significant association was observed in survival probability of HPV infected (a) pre-therapeutic and (b) NACT patients.
(TIF)

S4 Fig. Characterization of different splice form of E6 and E6/E7 transcripts of HPV16 in pre-therapeutic breast tumor. Representative sequence chromatogram of (a) E6^{*}II/E7 transcript showing junction of the splicing site (nt225/526) (b) E6[^]E7 transcript showing junction of the splicing site (nt226/742) (c) E6[^]E7^{*}I transcript showing junction of the splicing site (nt174/718) (d) E6[^]E7^{*}II transcript showing junction of the splicing site (nt221/850). Novel splice form was shown in bold.
(TIF)

Acknowledgments

The authors thank the Director, Chittaranjan National Cancer Institute, Kolkata, India. They thank Dr. Partha Sarathi Dasgupta, Emeritus Scientist, Chittaranjan National Cancer Institute, for his valuable suggestions during the study. We would also like to thank Mr. Sudip Samadder for his technical help. We would also like to extend our gratitude to Dr. Michel Favre, Genetics, Papillomavirus and Human Cancer Unit, Institut Pasteur, France for providing HPV33

plasmid without which the present study would have been incomplete. We would like to thank Dr. H. zur Hausen and Dr. E. M. de Villiers for their generous gifts of HPV16/18 plasmids. Financial support for this work was provided by UGC-NET Fellowship grant F.2-3/2000 (SA-I) (Sr. No. 2061030813, Ref. No.: 20–06 / 2010(i) EU-IV dated 22.10.2010) to Mrs. H. Dasgupta and grants from Council of Scientific and Industrial Research, Government of India [No. 60 (0111)/14/EMR-II of dt 03/11/2014] to Dr. C. K. Panda.

Author Contributions

Conceptualization: SI CKP.

Data curation: SI A. Roy GKM.

Formal analysis: SI SM.

Funding acquisition: HD CKP.

Investigation: SI HD RB NM.

Methodology: SI CKP.

Project administration: CKP SR.

Resources: NA GKM CKP JB.

Supervision: CKP.

Validation: A. Roychowdhury SI HD.

Visualization: SI A. Roychowdhury.

Writing – original draft: SI.

Writing – review & editing: A. Roychowdhury CKP SR.

References

1. Psyrrri A, DiMaio D. Human papillomavirus in cervical and head-and-neck cancer. *Nat Clin Pract Oncol*. 2008; 5(1):24–31. doi: [10.1038/ncponc0984](https://doi.org/10.1038/ncponc0984) PMID: [18097454](https://pubmed.ncbi.nlm.nih.gov/18097454/)
2. Bruni L, Diaz M, Castellsague X, Ferrer E, Bosch FX, de Sanjose S. Cervical human papillomavirus prevalence in 5 continents: meta-analysis of 1 million women with normal cytological findings. *J Infect Dis*. 2010; 202(12):1789–99. doi: [10.1086/657321](https://doi.org/10.1086/657321) PMID: [21067372](https://pubmed.ncbi.nlm.nih.gov/21067372/)
3. Egawa K. Do human papillomaviruses target epidermal stem cells? *Dermatology*. 2003; 207(3):251–4. PMID: [14571065](https://pubmed.ncbi.nlm.nih.gov/14571065/)
4. Moody CA, Laimins LA. Human papillomavirus oncoproteins: pathways to transformation. *Nat Rev Cancer*. 2010; 10(8):550–60. doi: [10.1038/nrc2886](https://doi.org/10.1038/nrc2886) PMID: [20592731](https://pubmed.ncbi.nlm.nih.gov/20592731/)
5. Kalantari M, Karlsen F, Kristensen G, Holm R, Hagmar B, Johansson B. Disruption of the E1 and E2 reading frames of HPV 16 in cervical carcinoma is associated with poor prognosis. *Int J Gynecol Pathol*. 1998; 17(2):146–53. PMID: [9553812](https://pubmed.ncbi.nlm.nih.gov/9553812/)
6. Dell G, Gaston K. Human papillomaviruses and their role in cervical cancer. *Cell Mol Life Sci*. 2001; 58(12–13):1923–42. PMID: [11766888](https://pubmed.ncbi.nlm.nih.gov/11766888/)
7. Kozuka T, Aoki Y, Nakagawa K, Ohtomo K, Yoshikawa H, Matsumoto K, et al. 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*. 2000; 91(3):271–9. PMID: [10760685](https://pubmed.ncbi.nlm.nih.gov/10760685/)
8. Kim K, Garner-Hamrick PA, Fisher C, Lee D, Lambert PF. Methylation patterns of papillomavirus DNA, its influence on E2 function, and implications in viral infection. *J Virol*. 2003; 77(23):12450–9. doi: [10.1128/JVI.77.23.12450-12459.2003](https://doi.org/10.1128/JVI.77.23.12450-12459.2003) PMID: [14610169](https://pubmed.ncbi.nlm.nih.gov/14610169/)
9. Johansson C, Schwartz S. Regulation of human papillomavirus gene expression by splicing and polyadenylation. *Nat Rev Microbiol*. 2013; 11(4):239–51. doi: [10.1038/nrmicro2984](https://doi.org/10.1038/nrmicro2984) PMID: [23474685](https://pubmed.ncbi.nlm.nih.gov/23474685/)

10. Li N, Bi X, Zhang Y, Zhao P, Zheng T, Dai M. Human papillomavirus infection and sporadic breast carcinoma risk: a meta-analysis. *Breast Cancer Res Treat.* 2011; 126(2):515–20. doi: [10.1007/s10549-010-1128-0](https://doi.org/10.1007/s10549-010-1128-0) PMID: [20740311](https://pubmed.ncbi.nlm.nih.gov/20740311/)
11. Herrera-Goeptfert R, Khan NA, Koriyama C, Akiba S, Perez-Sanchez VM. High-risk human papillomavirus in mammary gland carcinomas and non-neoplastic tissues of Mexican women: no evidence supporting a cause and effect relationship. *Breast.* 2011; 20(2):184–9. doi: [10.1016/j.breast.2010.11.006](https://doi.org/10.1016/j.breast.2010.11.006) PMID: [21146410](https://pubmed.ncbi.nlm.nih.gov/21146410/)
12. Bae JM, Kim EH. Human papillomavirus infection and risk of breast cancer: a meta-analysis of case-control studies. *Infect Agent Cancer.* 2016; 11:14. doi: [10.1186/s13027-016-0058-9](https://doi.org/10.1186/s13027-016-0058-9) PMID: [26981149](https://pubmed.ncbi.nlm.nih.gov/26981149/)
13. Manzouri L, Salehi R, Shariatpanahi S, Rezaie P. Prevalence of human papilloma virus among women with breast cancer since 2005–2009 in Isfahan. *Adv Biomed Res.* 2014; 3:75. doi: [10.4103/2277-9175.125873](https://doi.org/10.4103/2277-9175.125873) PMID: [24627883](https://pubmed.ncbi.nlm.nih.gov/24627883/)
14. Lawson JS, Glenn WK, Salyakina D, Delprado W, Clay R, Antonsson A, et al. Human Papilloma Viruses and Breast Cancer. *Front Oncol.* 2015; 5:277. doi: [10.3389/foonc.2015.00277](https://doi.org/10.3389/foonc.2015.00277) PMID: [26734565](https://pubmed.ncbi.nlm.nih.gov/26734565/)
15. Woods Ignatoski KM, Dziubinski ML, Ammerman C, Ethier SP. Cooperative interactions of HER-2 and HPV-16 oncoproteins in the malignant transformation of human mammary epithelial cells. *Neoplasia.* 2005; 7(8):788–98. PMID: [16207481](https://pubmed.ncbi.nlm.nih.gov/16207481/)
16. Khan NA, Castillo A, Koriyama C, Kijima Y, Umekita Y, Ohi Y, et al. Human papillomavirus detected in female breast carcinomas in Japan. *Br J Cancer.* 2008; 99(3):408–14. doi: [10.1038/sj.bjc.6604502](https://doi.org/10.1038/sj.bjc.6604502) PMID: [18648364](https://pubmed.ncbi.nlm.nih.gov/18648364/)
17. Ohba K, Ichiyama K, Yajima M, Gemma N, Nikaido M, Wu Q, et al. In vivo and in vitro studies suggest a possible involvement of HPV infection in the early stage of breast carcinogenesis via APOBEC3B induction. *PLoS One.* 2014; 9(5):e97787. doi: [10.1371/journal.pone.0097787](https://doi.org/10.1371/journal.pone.0097787) PMID: [24858917](https://pubmed.ncbi.nlm.nih.gov/24858917/)
18. Statistics Of Breast Cancer In India 2012. http://www.breastcancerindia.net/statistics/stat_global.html
19. Hedau S, Kumar U, Hussain S, Shukla S, Pande S, Jain N, et al. Breast cancer and human papillomavirus infection: no evidence of HPV etiology of breast cancer in Indian women. *BMC Cancer.* 2011; 11:27. doi: [10.1186/1471-2407-11-27](https://doi.org/10.1186/1471-2407-11-27) PMID: [21247504](https://pubmed.ncbi.nlm.nih.gov/21247504/)
20. Dasgupta S, Mukherjee N, Roy S, Roy A, Sengupta A, Roychowdhury S, et al. Mapping of the candidate tumor suppressor genes' loci on human chromosome 3 in head and neck squamous cell carcinoma of an Indian patient population. *Oral Oncol.* 2002; 38(1):6–15. PMID: [11755815](https://pubmed.ncbi.nlm.nih.gov/11755815/)
21. Maiti GP, Mondal P, Mukherjee N, Ghosh A, Ghosh S, Dey S, et al. Overexpression of EGFR in head and neck squamous cell carcinoma is associated with inactivation of SH3GL2 and CDC25A genes. *PLoS One.* 2013; 8(5):e63440. doi: [10.1371/journal.pone.0063440](https://doi.org/10.1371/journal.pone.0063440) PMID: [23675485](https://pubmed.ncbi.nlm.nih.gov/23675485/)
22. Dutta S, Begum R, Mazumder Indra D, Mandal SS, Mondal R, Biswas J, et al. Prevalence of human papillomavirus in women without cervical cancer: a population-based study in Eastern India. *Int J Gynecol Pathol.* 2012; 31(2):178–83. doi: [10.1097/PGP.0b013e3182399391](https://doi.org/10.1097/PGP.0b013e3182399391) PMID: [22317877](https://pubmed.ncbi.nlm.nih.gov/22317877/)
23. Baay MF, Quint WG, Koudstaal J, Hollema H, Duk JM, Burger MP, et al. Comprehensive study of several general and type-specific primer pairs for detection of human papillomavirus DNA by PCR in paraffin-embedded cervical carcinomas. *J Clin Microbiol.* 1996; 34(3):745–7. PMID: [8904451](https://pubmed.ncbi.nlm.nih.gov/8904451/)
24. Bhattacharya N, Singh RK, Mondal S, Roy A, Mondal R, Roychowdhury S, et al. Analysis of molecular alterations in chromosome 8 associated with the development of uterine cervical carcinoma of Indian patients. *Gynecol Oncol.* 2004; 95(2):352–62. doi: [10.1016/j.ygyno.2004.07.036](https://doi.org/10.1016/j.ygyno.2004.07.036) PMID: [15491757](https://pubmed.ncbi.nlm.nih.gov/15491757/)
25. Mazumder Indra D, Singh RK, Mitra S, Dutta S, Chakraborty C, Basu PS, et al. Genetic and epigenetic changes of HPV16 in cervical cancer differentially regulate E6/E7 expression and associate with disease progression. *Gynecol Oncol.* 2011; 123(3):597–604. doi: [10.1016/j.ygyno.2011.08.004](https://doi.org/10.1016/j.ygyno.2011.08.004) PMID: [21911249](https://pubmed.ncbi.nlm.nih.gov/21911249/)
26. Si HX, Tsao SW, Poon CS, Wong YC, Cheung AL. Physical status of HPV-16 in esophageal squamous cell carcinoma. *J Clin Virol.* 2005; 32(1):19–23. doi: [10.1016/j.jcv.2004.04.004](https://doi.org/10.1016/j.jcv.2004.04.004) PMID: [15572001](https://pubmed.ncbi.nlm.nih.gov/15572001/)
27. Yamada T, Manos MM, Peto J, Greer CE, Munoz N, Bosch FX, et al. Human papillomavirus type 16 sequence variation in cervical cancers: a worldwide perspective. *J Virol.* 1997; 71(3):2463–72. PMID: [9032384](https://pubmed.ncbi.nlm.nih.gov/9032384/)
28. Dutta S, Chakraborty C, Dutta AK, Mandal RK, Roychowdhury S, Basu P, et al. Physical and methylation status of human papillomavirus 16 in asymptomatic cervical infections changes with malignant transformation. *J Clin Pathol.* 2015; 68(3):206–11. doi: [10.1136/jclinpath-2014-202611](https://doi.org/10.1136/jclinpath-2014-202611) PMID: [25563334](https://pubmed.ncbi.nlm.nih.gov/25563334/)
29. Loginov VI, Maliukova AV, Seregin Iu A, Khodyrev DS, Kazubskaja TP, Ermilova VD, et al. [Methylation of the promoter region of the RASSF1A gene, a candidate tumor suppressor, in primary epithelial tumors]. *Mol Biol (Mosk).* 2004; 38(4):654–67.

30. Dutta S, Chakraborty C, Mandal RK, Basu P, Biswas J, Roychoudhury S, et al. Persistent HPV16/18 infection in Indian women with the A-allele (rs6457617) of HLA-DQB1 and T-allele (rs16944) of IL-1beta -511 is associated with development of cervical carcinoma. *Cancer Immunol Immunother.* 2015; 64(7):843–51. doi: [10.1007/s00262-015-1693-5](https://doi.org/10.1007/s00262-015-1693-5) PMID: [25893807](https://pubmed.ncbi.nlm.nih.gov/25893807/)
31. Sambrook J, Russell DW, Sambrook J. The condensed protocols from Molecular cloning: a laboratory manual. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press; 2006. v, 800 p. p.
32. Grabe N. AliBaba2: context specific identification of transcription factor binding sites. *In Silico Biol.* 2002; 2(1):S1–15. PMID: [11808873](https://pubmed.ncbi.nlm.nih.gov/11808873/)
33. Bachmair A, Finley D, Varshavsky A. In vivo half-life of a protein is a function of its amino-terminal residue. *Science.* 1986; 234(4773):179–86. PMID: [3018930](https://pubmed.ncbi.nlm.nih.gov/3018930/)
34. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 2003; 31(13):3406–15. PMID: [12824337](https://pubmed.ncbi.nlm.nih.gov/12824337/)
35. Capriotti E, Fariselli P, Casadio R. I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Res.* 2005; 33(Web Server issue):W306–10. doi: [10.1093/nar/gki375](https://doi.org/10.1093/nar/gki375) PMID: [15980478](https://pubmed.ncbi.nlm.nih.gov/15980478/)
36. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the functional effect of amino acid substitutions and indels. *PLoS One.* 2012; 7(10):e46688. doi: [10.1371/journal.pone.0046688](https://doi.org/10.1371/journal.pone.0046688) PMID: [23056405](https://pubmed.ncbi.nlm.nih.gov/23056405/)
37. Wang T, Chang P, Wang L, Yao Q, Guo W, Chen J, et al. The role of human papillomavirus infection in breast cancer. *Med Oncol.* 2012; 29(1):48–55. doi: [10.1007/s12032-010-9812-9](https://doi.org/10.1007/s12032-010-9812-9) PMID: [21318737](https://pubmed.ncbi.nlm.nih.gov/21318737/)
38. da Fonseca AJ, Galvao RS, Miranda AE, Ferreira LC, Chen Z. Comparison of three human papillomavirus DNA detection methods: Next generation sequencing, multiplex-PCR and nested-PCR followed by Sanger based sequencing. *J Med Virol.* 2016; 88(5):888–94. doi: [10.1002/jmv.24413](https://doi.org/10.1002/jmv.24413) PMID: [26496186](https://pubmed.ncbi.nlm.nih.gov/26496186/)
39. Di Lonardo A, Venuti A, Marcante ML. Human papillomavirus in breast cancer. *Breast Cancer Res Treat.* 1992; 21(2):95–100. PMID: [1320958](https://pubmed.ncbi.nlm.nih.gov/1320958/)
40. Gumus M, Yumuk PF, Salepci T, Aliustaoglu M, Dane F, Ekenel M, et al. HPV DNA frequency and subset analysis in human breast cancer patients' normal and tumoral tissue samples. *J Exp Clin Cancer Res.* 2006; 25(4):515–21. PMID: [17310842](https://pubmed.ncbi.nlm.nih.gov/17310842/)
41. Lawson JS, Glenn WK, Whitaker NJ. Human Papilloma Viruses and Breast Cancer—Assessment of Causality. *Front Oncol.* 2016; 6:207. doi: [10.3389/fonc.2016.00207](https://doi.org/10.3389/fonc.2016.00207) PMID: [27747193](https://pubmed.ncbi.nlm.nih.gov/27747193/)
42. Damin AP, Karam R, Zettler CG, Caleffi M, Alexandre CO. Evidence for an association of human papillomavirus and breast carcinomas. *Breast Cancer Res Treat.* 2004; 84(2):131–7. doi: [10.1023/B:BREA.0000018411.89667.0d](https://doi.org/10.1023/B:BREA.0000018411.89667.0d) PMID: [14999143](https://pubmed.ncbi.nlm.nih.gov/14999143/)
43. Herrera-Goepfert R, Vela-Chavez T, Carrillo-Garcia A, Lizano-Soberon M, Amador-Molina A, Onate-Ocana LF, et al. High-risk human papillomavirus (HPV) DNA sequences in metaplastic breast carcinomas of Mexican women. *BMC Cancer.* 2013; 13:445. doi: [10.1186/1471-2407-13-445](https://doi.org/10.1186/1471-2407-13-445) PMID: [24083491](https://pubmed.ncbi.nlm.nih.gov/24083491/)
44. Kulmala SM, Syrjanen SM, Gyllensten UB, Shabalova IP, Petrovichev N, Tosi P, et al. Early integration of high copy HPV16 detectable in women with normal and low grade cervical cytology and histology. *J Clin Pathol.* 2006; 59(5):513–7. doi: [10.1136/jcp.2004.024570](https://doi.org/10.1136/jcp.2004.024570) PMID: [16484445](https://pubmed.ncbi.nlm.nih.gov/16484445/)
45. Arias-Pulido H, Peyton CL, Joste NE, Vargas H, Wheeler CM. Human papillomavirus type 16 integration in cervical carcinoma in situ and in invasive cervical cancer. *J Clin Microbiol.* 2006; 44(5):1755–62. doi: [10.1128/JCM.44.5.1755-1762.2006](https://doi.org/10.1128/JCM.44.5.1755-1762.2006) PMID: [16672403](https://pubmed.ncbi.nlm.nih.gov/16672403/)
46. Pientong C, Wongwarissara P, Ekalaksananan T, Swangphon P, Kleeboon P, Kongyingyoes B, et al. Association of human papillomavirus type 16 long control region mutation and cervical cancer. *Virology.* 2013; 45(1):30. doi: [10.1016/j.virol.2012.11.030](https://doi.org/10.1016/j.virol.2012.11.030) PMID: [23343096](https://pubmed.ncbi.nlm.nih.gov/23343096/)
47. Pande S, Jain N, Prusty BK, Bhambhani S, Gupta S, Sharma R, et al. Human papillomavirus type 16 variant analysis of E6, E7, and L1 genes and long control region in biopsy samples from cervical cancer patients in north India. *J Clin Microbiol.* 2008; 46(3):1060–6. doi: [10.1128/JCM.02202-07](https://doi.org/10.1128/JCM.02202-07) PMID: [18199779](https://pubmed.ncbi.nlm.nih.gov/18199779/)
48. Pillai MR, Hariharan R, Babu JM, Lakshmi S, Chiplunkar SV, Patkar M, et al. Molecular variants of HPV-16 associated with cervical cancer in Indian population. *Int J Cancer.* 2009; 125(1):91–103. doi: [10.1002/ijc.24322](https://doi.org/10.1002/ijc.24322) PMID: [19358280](https://pubmed.ncbi.nlm.nih.gov/19358280/)
49. van Duin M, Snijders PJ, Vossen MT, Klaassen E, Voorhorst F, Verheijen RH, et al. Analysis of human papillomavirus type 16 E6 variants in relation to p53 codon 72 polymorphism genotypes in cervical carcinogenesis. *J Gen Virol.* 2000; 81(Pt 2):317–25. doi: [10.1099/0022-1317-81-2-317](https://doi.org/10.1099/0022-1317-81-2-317) PMID: [10644829](https://pubmed.ncbi.nlm.nih.gov/10644829/)
50. Chang YJ, Chen HC, Pan MH, Lee BH, You SL, Lin CY, et al. Intratypic variants of human papillomavirus type 16 and risk of cervical neoplasia in Taiwan. *J Med Virol.* 2013; 85(9):1567–76. doi: [10.1002/jmv.23651](https://doi.org/10.1002/jmv.23651) PMID: [23852681](https://pubmed.ncbi.nlm.nih.gov/23852681/)

51. Park IS, Chang X, Loyo M, Wu G, Chuang A, Kim MS, et al. Characterization of the methylation patterns in human papillomavirus type 16 viral DNA in head and neck cancers. *Cancer Prev Res (Phila)*. 2011; 4(2):207–17.
52. Shuyama K, Castillo A, Aguayo F, Sun Q, Khan N, Koriyama C, et al. Human papillomavirus in high- and low-risk areas of oesophageal squamous cell carcinoma in China. *Br J Cancer*. 2007; 96(10):1554–9. doi: [10.1038/sj.bjc.6603765](https://doi.org/10.1038/sj.bjc.6603765) PMID: [17453003](https://pubmed.ncbi.nlm.nih.gov/17453003/)
53. Koskinen WJ, Chen RW, Leivo I, Makitie A, Back L, Kontio R, et al. Prevalence and physical status of human papillomavirus in squamous cell carcinomas of the head and neck. *Int J Cancer*. 2003; 107(3):401–6. doi: [10.1002/ijc.11381](https://doi.org/10.1002/ijc.11381) PMID: [14506740](https://pubmed.ncbi.nlm.nih.gov/14506740/)
54. Frega A, Lorenzon L, Bononi M, De Cesare A, Ciardi A, Lombardi D, et al. Evaluation of E6 and E7 mRNA expression in HPV DNA positive breast cancer. *Eur J Gynaecol Oncol*. 2012; 33(2):164–7. PMID: [22611956](https://pubmed.ncbi.nlm.nih.gov/22611956/)
55. Ajiro M, Jia R, Zhang L, Liu X, Zheng ZM. Intron definition and a branch site adenosine at nt 385 control RNA splicing of HPV16 E6*1 and E7 expression. *PLoS One*. 2012; 7(10):e46412. doi: [10.1371/journal.pone.0046412](https://doi.org/10.1371/journal.pone.0046412) PMID: [23056301](https://pubmed.ncbi.nlm.nih.gov/23056301/)
56. Smeets SJ, Hesselink AT, Speel EJ, Haesevoets A, Snijders PJ, Pawlita M, et al. A novel algorithm for reliable detection of human papillomavirus in paraffin embedded head and neck cancer specimen. *Int J Cancer*. 2007; 121(11):2465–72. doi: [10.1002/ijc.22980](https://doi.org/10.1002/ijc.22980) PMID: [17680565](https://pubmed.ncbi.nlm.nih.gov/17680565/)