JCM Accepts, published online ahead of print on 16 January 2008
J. Clin. Microbiol. doi:10.1128/JCM.02202-07

HPV-16 Variant Analysis of E6, E7, L1 Genes and Long Control Region in Cervical Cancer Patients from North India

Shailja Pande[1], Neeraj Jain[1], Bhupesh K Prusty[1], Suresh Bhambhani[2], Sanjay Gupta[2], Rajyashri Sharma[3], Swaraj Batra[4], Bhudev C. Das[1]*

[1]Division of Molecular Oncology, [2] Division of Cytopathology, Institute of Cytology and Preventive Oncology (ICMR), I-7, Sector 39, Noida 201301, INDIA

[3]Department of Obsteritics and Gynecology, JNMCH, Aligarh

[4]Department of Obsteritics and Gynecology, LNJP, New Delhi

Running Title: HPV 16 variants in cervical cancer from India.

Key Words: HPV 16, Prototype, Variants, Sequence, Cervix cancer.

***Corresponding Author**: Bhudev C. Das, Ph.D., Division of Molecular Oncology, Institute of Cytology and Preventive Oncology, I-7, Sector 39, Noida 201301 INDIA Tel.: 91-120-2578837; Fax: 91-120-2579473.

E-mail: bcdas48@hotmail.com

ABSTRACT

High risk human papillomaviruses particularly, HPV types 16 and 18 play a cardinal role in the aetiology of cervical cancer. The most prevalent HPV type 16 shows intratypic sequence variants which are known to differ in oncogenic potential and geographic distribution. This study is designed to analyze sequence variations in E6, E7, L1 and LCR regions of HPV 16 in cervical cancer patients to identify most prevalent and novel HPV 16 variants and to correlate them with the severity of the disease. Cervical biopsies from sixty HPV16 positive cancer cases were analyzed by PCR and DNA sequencing. The most frequently observed variations were T350G (100%) in E6, T789C (87.5%) in E7, A6695C (54.5%) in L1 and G7521A (91.1%) in LCR. In addition, only one novel variant (T527A) in E6 and four new variants each in L1 (A6667C, A6691G, C6906T, A6924C) and in LCR (C13T, A7636C, C7678T, G7799A) were identified. While E7 was found to be highly conserved, the variant 350G of E6 was most prevalent in all the histopathological grades. Majority of LCR variants were found at the YY1 transcription factor binding sites. Interestingly, a complete absence of Asian Lineage and a high prevalence of European lineages in E6, E7, L1 and LCR (85%, 86.7%, 67.7% and 63.3% respectively) indicate possible epidemiological linkage between Europe and India with regard to the dissemination of HPV 16 infections in India.

HPV-16 Variant Analysis of E6, E7, L1 Genes and Long Control Region in Cervical Cancer Patients from North India

Cervical cancer is the major cancer in Indian women and a leading cause of cancer deaths. Every year more than 1, 30,000 new cases and about 70,000 deaths are recorded. Persistent infection of specific types of high risk human papillomaviruses (HR-HPVs) is essential for progression of cervical lesions(6, 11, 31, 60) and it is this subset of women that are likely to develop cancer subsequently (3, 4, 27, 40).Various studies have demonstrated that more than 70% of invasive cervical cancers harbor HPV16 and HPV18 (17, 31) and the products of viral transforming E6 and E7 genes are shown to contribute to tumorigenesis by functionally inactivating two important cellular tumor suppressor proteins, the p53 and the Rb, respectively (5, 14, 30, 55).

More than 100 types of HPV are known but only about 30 types are associated with anogenital cancer. According to Papillomavirus Nomenclature Committee, a new HPV type is defined by nucleotide sequence variation of more than 10% compared with other known HPV types in the E6, E7 and L1 open reading frames (ORFs). Those differing by 2-10% variations are referred to as subtypes whereas intratype variants may vary up to 2% in the coding region and 5% in the non-coding region when compared to that of prototype (2, 10).

On the basis of sequence variations in E6, L1, L2 and LCR HPV 16 variants have been identified and grouped into six distinct phylogenetic branches, such as E (European), AA (Asian-American), Af1 (African 1), Af2 (African 2), As (Asian), and NA (North American) (54, 56, 57). These variants have been found to show different geographic distribution with a varied oncogenic potential. A number of sequence variations have been reported in HPV 16 E6, E7, L1 gene as well as in LCR in cervical cancer (33, 39, 48, 55, 56). Studies have also shown that specific intratype variants may influence persistence of HPV infection and progression of precursor lesions to cancer (27, 58, 59).

HPV variants may also affect virus assembly, immunologic response, pathogenecity, p53 degradation, immortalization activity and regulation of transcription (15,18, 24, 25, 37, 51). An immediate impact of these variations will be on the sensitivity and specificity of different PCR based genotype diagnostic methods.

Of two HPV-16 oncogenes E6 and E7; the E6 has been found to show more variations compared to E7 which is relatively conserved (48, 50, 56, 58, 59). The analysis of the L1 gene that codes for the viral major capsid protein is of immense importance because of its high diagnostic value. The range of intratype variation observed in this region allows the distinction and assessment of known and novel HPV types (46). This is also an important target for the development of HPV vaccines. Very few reports are available on HPV variants from this region (9, 56). The long control region (LCR), which regulates viral transcription, has been found to be the most variable region of the HPV-16 genome (41, 53, 56). Several authors have evaluated association of specific HPV 16 variants with viral persistence and the development of CIN lesions (3, 27, 34, 50, 52).

In India, as high as 98% of the cervical carcinoma cases are found to harbor HPV infection and the most prevalent (~80%) type is HPV-16 (7, 8, 22). Although the prevalence of HPV and cervical cancer in India is highest in the world, there is not much data on HPV variants from different regions of the Indian subcontinent. In the present study we have examined the sequence variation in E6, E7, L1 and long control region (LCR) of the most prevalent high risk HPV type 16 and correlated them with the age of women, histopathologic grades of tumor and oncogenic potential.

MATERIALS AND METHODS

Study population and specimen collection. Out of a total of 90 tumor samples screened by standard procedure (20, 38), sixty HPV-16 positive cervical cancer cases were recruited for the variant analysis in the present study. The tumor samples were part of biopsies collected for routine diagnosis of cervical cancer from the Cancer Clinic of the Department of Gynecology and Obstetrics, Lok Nayak Hospital, New Delhi and Jawaharlal Nehru Medical College, Aligarh. Fresh tumor biopsies were collected in chilled phosphate buffered saline (PBS) and stored at -70°C deep freezer for further processing later. Informed consent was obtained from all the patients and the study was ethically approved by the Institutional Ethical Committee. Diagnosis of histopathological grades was done by two Pathologists independently. In the case of discrepancies, it was reconfirmed by the third Pathologist to arrive at a final decision.

DNA extraction and typing of HPV. High molecular weight genomic DNA from cervical biopsies was isolated by the standard method of proteinase K digestion and phenol chloroform extraction routinely followed in our laboratory (8, 20).

Detection of HPV and typing of high-risk HPV types 16 and 18 was carried out using consensus and type-specific primers described earlier (20, 38). Amplification of *β-globin* gene served as the internal control. PCR was performed using the in-house PCR-protocol routinely followed in our laboratory (7, 8, 17). Briefly, the method involved a 25µl reaction mix containing 100-200ng DNA, 10mM Tris-Cl (pH 8.4), 50mM KCl, 1.5mM $MgCl_2$, 12.5µM of each dNTPs (dATP, dCTP, dGTP and dTTP), 5 pmoles of each oligonucleotide primer and 0.5U Taq DNA Polymerase (Applied Biosystems, US). The temperature profile used for amplification constituted an initial denaturation at 95°C for 5min followed by 30 cycles with denaturation at 95°C for 30sec, annealing at 55°C for 30sec and extension at 72°C for 30sec which was extended for 4min in the final cycle. The oligonucleotide primers were

synthesized in an automated Applied Biosystems DNA Synthesizer (Model 381A, Applied Biosystems Inc., Foster City, CA, USA) and HPLC purified.

Variant analysis of E6, E7, L1 gene and LCR of HPV type 16 by PCR and direct sequencing. HPV 16 E6, E7 and LCR specific PCR was performed with specific primers for HPV 16 E6 flanking the region (nt 83-559) with amplicon size 476 bp: 5'-GAA ACC GGT TAG TAT AAA AGC AGA C-3' and 5'-AGC TGG GTT TCT CTA CGT GTT CT-3'; for E7 (nt 562-858) with amplicon size 296 bp: 5'-CCA TAA TAT AAG GGG TCG GTG GA-3' and 5'-TTT TTC CAC TAA CAG CCT CTA CAT-3'; for LCR (nt 7437-7456, 119-147) amplicon size 617 bp: 5'-CCA TTT TGT AGC TTC AAC CCG-3' and 5'-AAG TGT GGT AAC TTT CTG GGT CGC TCC TG-3'(54). L1 gene was amplified partially (450 bp) using MY 11/MY 09 consensus primer sequences: 5'-GCM CAG GGW CAT AAY AAT GG-3' and 5'-CGT CCM ARR GGA WAC TGA-3' where M=A+C, W=A+T, Y=C+T and R=A+G (38).

For variant analysis, the PCR products were directly sequenced on an automated DNA sequencer (310 ABI Prism Genetic Analyzer, Applied Biosystems, USA) according to the manufacturer's protocol. PCR products were first purified using ammonium acetate/ ethanol precipitation method to remove unused dNTPs and primers, and then cycle sequenced using the BIG Dye Terminator Sequencing Ready Reaction Mix (Applied Biosystems, USA) on Gene Amp PCR 9700 (Applied Biosystems, USA). The raw data was collected using ABI Prism 310 collection software, analyzed using sequencing analysis software V3.4.1 on a MAC operating system V9.1 and compared with the sequence Gen Bank/EMBL/DDBJ U89348. All the samples were reverse sequenced to corroborate the findings.

Identification of variants. Variants were identified using the prototype sequence (HPV 16R), which belongs to the European lineages, as standard for comparisons and nucleotide position numbering (32) and variants were classified in six major branches as European (E), Asian-American

(AA), Asian (As), African1 (Af1), African 2 (Af2) and North American (NA) as described by Yamada et al. (56).

Statistical analysis. Statistical analysis was performed using the χ^2 test to determine the association between various HPV16 variants, age and histopathologic status of the cancer patients. Fisher's exact test was employed where small numbers were compared. P-value less than 0.05 has been considered as stastically significant.

RESULTS

Out of 90 cervical tumor screened by L1 consensus primers, 65(72.2%) were positive for HPV, and further genotyping with type-specific primers revealed 92.3% (n = 60) positivity for HPV 16 and 7.7% (n = 5) for HPV 18. All sixty patients recruited were exclusively infected with HPV-16 and none of them showed presence of HPV type 18. The patients belonged to age group 30-80 year with a mean age of 51.3 year (SD = 12.8). The patients were divided into two age groups; 30 to 50 years and above 50 years. All the cases were of squamous cell carcinoma; 23 (38.3%) well-differentiated (WDSCC), 35 (58.3%) moderately differentiated (MDSCC) and 2 (3.4%) were poorly differentiated squamous cell carcinomas (PDSCC). The E6, E7, L1 and the LCR were analyzed for nucleotide variations which were compared with the prototype sequence 16 R of the European lineage (E).

Association between HPV variants and histopathological grades. The E6 variants were found to be significantly higher as compared to prototype in both WDSCC (68.2%) and MDSCC (69.4%) cases which include only 2 PDSCC cases while prototypes of E7 (86.4%, 86.2% respectively) and L1 (60.0%, 73.6% respectively) (**Table 1**) were most frequent in them. The frequency of LCR variants was also found to be more in WDSCC as well as MDSCC (81.8% and 75% respectively) (see table 1).

E6 sequence variations. Nucleotide sequences of the complete E6 ORF in all 60 cancer patients screened were compared with the HPV16 reference sequence. Forty-two (70%) patients showed seven different types of E6 gene mutations of which four; G145T, C335T, T350G and T527A were missense mutations; Glutamine → Histidine (Q14H), Histidine → Tyrosine (H78Y), Leucine → Valine (L83V) and Serine → Threonine (S142T) respectively. The remaining three T286A, A289G and A532 G led to silent mutations at codon 61, 62 and 143 for amino acids Alanine, Valine and Serine respectively .The most frequently observed mutation was the European variant T350G **(Fig. 2a)** which was detected in all 42 (100%) patients either alone (34; 81%) or along with other variants(8; 19%) **(Fig. 1)**. Additional E6 gene variations were observed only in those cases which showed 350G

variation and no case with 350T sequence had any nucleotide change. Majority of cases (85%) were found to be of European origin (E), while only 10% to branch NA and 3.3% AA, but none to As, Af1 or Af2 **(Table 3).** Interestingly, a novel mutation T527A was recorded along with 350G **(Fig.1)** for the first time.

E7 sequence variations. The E7 gene revealed only five nucleotide variations in 8 (13.3%) patients whereas rest 52 (86.7%) cases showed prototype sequence. All the variants were silent nucleotide alterations of which, one was transversion and four were transition substitution .The silent mutation detected were G666A, T732C, T789C, T795G and T828C for amino acids Glutamic Acid (E), Phenylalanine (F), Isoleucine (Il), Threonine (T) and Isoleucine (Il) at codons 35, 57, 76, 78 and 89 respectively. E7 sequencing data indicated 86.7% to branch E (P), 11.6% to Af2, 1.7% to As and none to NA, AA, Af1 (**Table 3**).

L1 sequence variations. A partial sequence of L1 gene was screened using 450 bp in 31 cervical cancer patients. A total of 13 nucleotide variations were detected in 11 (35.5%) patients while the majority (20; 64.5%) showed prototype sequence. Of 13 variants, 5 (38.5%) led to missense and 8 (61.5%) to silent mutations.

Of four variants found independently **(Fig. 1)**; two were silent mutations, A6667C and A6691G coding for Serine. at codon 343 and 351, whereas A6803T and A6964C led to missense mutations, Threonine → Serine (T389S) and Lysine → Asparagine (K442N). Most frequently detected L1 variation A6695C **(Fig. 2c)** was found in 6 (54.5%) cases. Interestingly, four novel L1 variants consisting of 2 silent mutations, at A6667C and A6691G, and 2 missense mutations, Serine → Phenylalanine (S423F) at C6906T and Glutamine → Proline (Q429P) at A6924C were observed **(Table 2)**. The case with novel L1 variants which exhibited silent mutations had only prototype sequences at E6, E7 and LCR **(Fig. 1)**.The L1 sequence data showed 67.7% to branch E, 3.2% to branch E (G131), 12.9% to AA, 6.5% to NA and none to As, Af1, Af2 while 9.7% were new variants **(Table 3).**

LCR sequence variations. The LCR of HPV-16 showed highest nucleotide variations in various known and unknown regulatory binding sites of number of cellular and/or viral transcription/regulatory factors. A total of 45 (75%) variant cases that were analyzed showed 17 different point mutations (9 transitions and 8 transversions)**.** Of these 7(G7521A, A7636C, C7689A, T7714G, C7792T, G7826A, A7839C) were detected in the known binding sites for various transcription factors such as YY1, AP1, TEF1, NF1 and Oct-1. The most commonly observed LCR variation was a transition substitution G7521A **(Fig. 2d)** detected in 41 (91.1%) patients at one of the many YY1 binding sites. This mutation occurred alone in as many as 22 (53.7%) cases **(Fig. 1).**

Another frequently observed LCR variation was a transversion mutation T7714G, detected in 9 (20%) cases mostly along with G7521A located in the NF1 transcription factor binding site. Two other LCR variants T7743G and C7764T were detected upstream to NF1 binding site in 11.1% and 22.2% cases respectively. We report here four novel LCR variants, C13T, A7636C, C7678T and G7799A **(Table 2)** of which 3 were transition and 1 (A7636C) was transversion mutations located at the AP1 binding site.

Phylogenetic classification of LCR variants was found to be in the same pattern as observed in E6, E7 and L1 showing majority (63.3%) to branch E (P), and only 13.3% to AA, 11.6% to Af1, 3.3% to Af2, and none to NA, As. Four (8.3%) novel LCR variants were observed for the first time from India **(Table 3).**

DISCUSSION

HPV variants data is important in developing HPV diagnostics, vaccines and other therapeutic approaches to control the virus-induced diseases. HPV 16 variants have been shown to have different biological as well as biochemical effects resulting in altered oncogenic potential (1, 47, 59). Oncogenicity of distinct HPV variants may also differ between geographical regions because of difference in population related distribution of HLA alleles (28).

Nucleotide sequence variations observed in E6, E7, L1 genes and LCR of HPV type 16 in 60 cervical carcinoma cases showed a high prevalence of 350G E6 variants as compared to prototype 350T in all different histopatholgic grades (WDSCC, MDSCC/ PDSCC). This is in good agreement with the previous reports (1, 59) which demonstrated an increased frequency of 350G E6 variants in cervical cancer leading to an amino acid change of Leucine to Valine at position 83 (L83V) and this has been associated with progression of cervical lesions (59).In contrast , several authors failed to find significant role of E6 variants in cervical cancer (19, 33, 50). The prevalence of E6 gene mutation studied in a subset of south Indian cervical cancer patients, showed only 20% of cancer patients with 350G variants which is in sharp contrast to the present study showing 78% of the patients with 350G variants.Our results however, are in good agreement with another study by Sathish et al. (36) from south-eastern part of India.

We have observed nucleotide positions 83-144, 146-285 and 351-526 as the most conserved regions of HPV 16 E6 which seems to be important for silencing E6 expression using siRNA and/or ribozyme treatment. Also, the amino acids located in these regions may play an important immunogenic role for new vaccine strategies (21). The amino acid substitution analyzed in this study are Q14H, H78Y, L83V which are also observed by Stoppler and his group (47). Most interestingly, a hundred percent distribution (100%, 42/42) of HPV 16 E6 European variant 350G was detected either alone (81%, 34/42) or in combination with additional E6 gene mutations. The E6 variations, G145T and C335T detected with 350G have also been observed in AA, NA, Af1 and Af2 lineages. A novel E6

gene mutation T527A (S142T) that replaced serine by threonine was observed in only one (2.4%) case. In sharp contrast to reports from China which showed high prevalence of HPV 16 Asian lineage (55) ,our results as well as the reports of other authors from India (36, 39) suggest complete absence of Asian lineage in India.

HPV 16 E7 gene mutation has been reported to be rare in various geographic and ethnic populations world over (16 36). Yet several studies particularly from Japan, Korea and Indonesia have reported a high frequency of E7 gene mutation (65-75%) in cervical cancer patients (9, 45). Interestingly, a recent report from China showed, 100% of cervical cancer cases having nucleotide variation in the HPV16 E7 region (55). In contrast, we observed a highly conserved E7 and as high as 86.7% patients showed no sequence variation and only 13.3% showed variation in the form of silent mutation thus leaving no effect on proteins. Because of conserved E7, it can be far easier target to silence its effects at RNA or protein level. Most frequently reported E7 variation A647G has also been found to be completely absent in India. Any change in the two transforming genes E6 and E7 may lead to altered biological function and oncogenicity of the proteins encoded, which can affect the natural history of HPV infection (13).

The L1 gene also did not reveal any significant variation and 64.5% (20/31) of the samples showed prototype sequence. Very few reports are available on the HPV 16 L1 variants (9, 53, 54, 56) and none from south Asia. Interestingly, the common L1 variant T6862C observed in all previous studies was not detected instead,4 novel nucleotide variations, two silent and two missense mutations at codon 343, 351, 423 and 429 were detected **(Table 2)**.This is indicative of hypervariability of L1 gene in this region. Importance of most frequently observed L1 gene variation A6695 (T353P) lies in the fact that a polar uncharged amino acid threonine is replaced by non-polar aliphatic amino acid that may have effect on structure or function of L1 protein, which may play an important role in immune recognition and vaccine development strategies. Amino acid changes among molecular variants have also been shown to affect the efficiency of HPV 16 L1 proteins to self assemble into VLPs (25, 49).

This variability could lead to conformational changes within epitopes relevant for viral neutralization (44).

The LCR which contains the upstream regulatory region is the binding sites of various cellular and viral transcription factors that either activate or suppress the p97 promoter activity which regulates the transcription of various HPV 16 genes, especially the E6 and E7 oncogenes. It has been reported to be most variable segment of the HPV 16 genome in different populations (23, 26, 53, 56). We also observed similar pattern as 75% (45/60) patients revealed presence of 17 different LCR variants of which 4 (C13T, A7636C, C7678T and G7799A) were found to be novel types, including, a A7636C substitution was detected in AP1 binding site (4.4%). Since LCR is the major site for transcriptional control of the virus; studies are underway to correlate the observed alterations with the biological/biochemical properties with different clinicopathological features of the disease. The number of nucleotide alterations was highest in LCR, followed by L1, E6 and E7 gene, and this is in good agreement with the previous report (56).The most frequently observed mutation in the LCR was G7521A located at the YY1-binding site of as high as 91% cases that causes repression of HPV transcription. This mutation has been found to be uniformly distributed in majority of cervical cancer patients throughout the world (26, 41, 53, 56) but not in asymptomatic carriers (41). Mutations in the YY1 binding sites have been shown to promote the P97 promoter activity by 3 to 6 folds (12, 29). It has also been shown to quench AP1 activity thereby repressing HPV 16 transcription (35). The role of mutation at YY1 binding site in cancer cases assumes a greater importance as this mutation is absent in non-cancerous lesions and asymptomatic carriers (42). Presence of mutation in the binding sites of various transcription factors or to their close vicinity is intriguing since these factors enhance the transcription activity of HPV 16 oncogenes. A low frequency (4.4%) of mutation found in Oct-1 binding site appears to be important since oct-1 is an important part of basal transcriptional machinery and is known to down-regulate HPV expression (43).

Most interestingly, when the distribution of HPV16 variant with respect to E6, E7, L1 and LCR was compared with that of different lineages, an extremely high prevalence of European lineage (85%, 86.7%, 67.7% and 63.3% respectively) and complete absence of Asian lineage (36, 39) were observed. The occurrence of AA and NA classes in this study is in accordance with other studies reported from Southern part of India (56). It is important that HPV variants are correlated with the severity of lesions through follow-up data to elucidate the biological significance of HPV 16 variants during cervical carcinogenesis thus delineating its utility in developing reliable diagnostics and effective therapeutics including vaccines against HPV.

ACKNOWLEDGEMENTS

Shailja Pande is thankful to Department Of Science and Technology (DST), Government of India, New Delhi, for providing financial support under Women Scientists' scheme to her. Authors are also thankful to Dr.L.Satyanarayana, Dr.Suresh Hedau and Dr. Alok C Bharti of ICPO for their help with this manuscript.

REFERENCES

1. **Andersson, S., M. Alemi, E. Rylander, A. Strand, B. Larsson, J. Sallstrom, and E. Wilander.** 2000. Uneven distribution of HPV 16 E6 prototype and variant (L83V) oncoprotein in cervical neoplastic lesions. Br J Cancer 83:307-310.

2. **Bernard, H. U., S. Y. Chan, M. M. Manos, C. K. Ong, L. L. Villa, H. Delius, C. L. Peyton, H. M. Bauer, and C. M. Wheeler.** 1994. Identification and assessment of known and novel human papillomaviruses by polymerase chain reaction amplification, restriction fragment length polymorphisms, nucleotide sequence, and phylogenetic algorithms. J Infect Dis **170:**1077-1085.

3. **Bontkes, H. J., M. van Duin, T. D. de Gruijl, M. F. Duggan-Keen, J. M. Walboomers, M. J. Stukart, R. H. Verheijen, T. J. Helmerhorst, C. J. Meijer, R. J. Scheper, F. R. Stevens, P. A. Dyer, P. Sinnott, and P. L. Stern.** 1998. HPV 16 infection and progression of cervical intra-epithelial neoplasia: analysis of HLA polymorphism and HPV 16 E6 sequence variants. Int J Cancer **78**:166-171.

4. **Castle, P. E., M. Schiffman, R. Herrero, A. Hildesheim, A. C. Rodriguez, M. C. Bratti, M. E. Sherman, S. Wacholder, R. Tarone, and R. D. Burk.** 2005. A prospective study of age trends in cervical human papillomavirus acquisition and persistence in Guanacaste, Costa Rica. J Infect Dis **191:**1808-1816.

5. **Crook, T., J. A. Tidy, and K. H. Vousden.** 1991. Degradation of p53 can be targeted by HPV E6 sequences distinct from those required for p53 binding and trans-activation. Cell **67:**547-556.

6. **Cuzick J, G. Terry , L. Ho , T. Hollingworth ,and M. Anderson .**1992.Human papillomavirus type 16 in cervical smears as predictor of high-grade cervical intraepithelial neoplasia [corrected] Lancet **18:** 959-960.

7. **Das, B. C., J. K. Sharma, V. Gopalakrishna, and U. K. Luthra.** 1992. Analysis by polymerase chain reaction of the physical state of human papillomavirus type 16 DNA in cervical preneoplastic and neoplastic lesions. J Gen Virol **73** (Pt 9):2327-2336.

8. **Das, B. C., J. K. Sharma, V. Gopalkrishna, D. K. Das, V. Singh, L. Gissmann, H. zur Hausen, and U. K. Luthra.** 1992. A high frequency of human papillomavirus DNA sequences in cervical carcinomas of Indian women as revealed by Southern blot hybridization and polymerase chain reaction. J Med Virol **36:**239-245.

9. **de Boer, M. A., L. A. Peters, M. F. Aziz, B. Siregar, S. Cornain, M. A. Vrede, E. S. Jordanova, S. Kolkman-Uljee, and G. J. Fleuren.** 2004. Human papillomavirus type 16 E6, E7, and L1 variants in cervical cancer in Indonesia, Suriname, and The Netherlands. Gynecol Oncol **94:**488-494.

10. **de Villiers, E. M.** 1994. Human pathogenic papillomavirus types: an update. Curr Top Microbiol Immunol **186:**1-12.

11. **de Villiers, E. M., C. Fauquet, T. R. Broker, H. U. Bernard, and H. zur Hausen.** 2004. Classification of papillomaviruses. Virology **324:**17-27**.**

12. **Dong, X. P., and H. Pfister.** 1999. Overlapping YY1- and aberrant SP1-binding sites proximal to the early promoter of human papillomavirus type 16. J Gen Virol **80 (Pt 8):**2097-2101.

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

14. **Dyson, N., P. M. Howley, K. Münger and E. Harlow.** 1989. The human papilloma virus-16 E7 oncoprotein is able to bind to the retinoblastoma gene product. Science **243:** 934-937.

15. **Ellis, J. R., P. J. Keating, J. Baird, E. F. Hounsell, D. V. Renouf, M. Rowe, D. Hopkins, M. F. Duggan-Keen, J. S. Bartholomew, L. S. Young, and et al.** 1995. The association of an HPV16 oncogene variant with HLA-B7 has implications for vaccine design in cervical cancer. Nat Med **1:**464-470.

16. **Eschle, D., M. Durst, J. ter Meulen, J. Luande, H. C. Eberhardt, M. Pawlita, and L. Gissmann.** 1992. Geographical dependence of sequence variation in the E7 gene of human papillomavirus type 16. J Gen Virol **73 (Pt 7):**1829-1832.

17. **Gupta, A., R. Arora, S. Gupta, B. K. Prusty, U. Kailash, S. Batra, and B. C. Das.** 2006. Human papillomavirus DNA in urine samples of women with or without cervical cancer and their male partners compared with simultaneously collected cervical/penile smear or biopsy specimens. J Clin Virol **37:**190-194.

18. **Hildesheim, A., M. Schiffman, C. Bromley, S. Wacholder, R. Herrero, A. Rodriguez, M. C. Bratti, M. E. Sherman, U. Scarpidis, Q. Q. Lin, M. Terai, R. L. Bromley, K. Buetow, R. J. Apple, and R. D. Burk.** 2001. Human papillomavirus type 16 variants and risk of cervical cancer. J Natl Cancer Inst **93:**315-318.

19. **Hu, X., T. Pang, Z. Guo, N. Mazurenko, F. Kisseljov, J. Ponten, and M. Nister.** 2001. HPV16 E6 gene variations in invasive cervical squamous cell carcinoma and cancer in situ from Russian patients. Br J Cancer **84:**791-795.

20. **Jain, N., V. Singh, S. Hedau, S. Kumar, M. K. Daga, R. Dewan, N. S. Murthy, S. A. Husain, and B. C. Das.** 2005. Infection of human papillomavirus type 18 and p53 codon 72 polymorphism in lung cancer patients from India. Chest **128:**3999-4007.

21. **Jiang, M., and J. Milner.** 2002. Selective silencing of viral gene expression in HPV-positive human cervical carcinoma cells treated with siRNA, a primer of RNA interference. Oncogene **21:**6041-6048.

22. **Kailash, U., C. C. Soundararajan, R. Lakshmy, R. Arora, S. Vivekanandhan, and B. C. Das.** 2006. Telomerase activity as an adjunct to high-risk human papillomavirus types 16 and 18 and cytology screening in cervical cancer. Br J Cancer **95:**1250-1257.

23. **Kammer, C., U. Warthorst, N. Torrez-Martinez, C. M. Wheeler, and H. Pfister.** 2000. Sequence analysis of the long control region of human papillomavirus type 16 variants and functional consequences for P97 promoter activity. J Gen Virol **81:**1975-1981.

24. **Kast, W. M., R. M. Brandt, J. Sidney, J. W. Drijfhout, R. T. Kubo, H. M. Grey, C. J. Melief, and A. Sette.** 1994. Role of HLA-A motifs in identification of potential CTL epitopes in human papillomavirus type 16 E6 and E7 proteins. J Immunol **152:**3904-3912.

25. **Kirnbauer, R., J. Taub, H. Greenstone, R. Roden, M. Durst, L. Gissmann, D. R. Lowy, and J. T. Schiller.** 1993. Efficient self-assembly of human papillomavirus type 16 L1 and L1-L2 into virus-like particles. J Virol **67:**6929-6936.

26. **Kurvinen, K., M. Yliskoski, S. Saarikoski, K. Syrjanen, and S. Syrjanen.** 2000. Variants of the long control region of human papillomavirus type 16. Eur J Cancer **36:**1402-1410.

27. **Londesborough, P., L. Ho, G. Terry, J. Cuzick, C. Wheeler, and A. Singer.** 1996. Human papillomavirus genotype as a predictor of persistence and development of high-grade lesions in women with minor cervical abnormalities. Int J Cancer **69:**364-368.

28. **Matsumoto, K., H. Yoshikawa, S. Nakagawa, X. Tang, T. Yasugi, K. Kawana, S. Sekiya, Y. Hirai, I. Kukimoto, T. Kanda, and Y. Taketani.** 2000. Enhanced oncogenicity of human papillomavirus type 16 (HPV16) variants in Japanese population. Cancer Lett **156:**159-165.

29. **May, M., X. P. Dong, E. Beyer-Finkler, F. Stubenrauch, P. G. Fuchs, and H. Pfister.** 1994. The E6/E7 promoter of extrachromosomal HPV16 DNA in cervical cancers escapes from cellular repression by mutation of target sequences for YY1. Embo J **13:**1460-1466.

30. **Munger, K., and P. M. Howley.** 2002. Human papillomavirus immortalization and transformation functions. Virus Res **89:**213-228.

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

32. **Myers, G., H. U. Bernard, H. Delius, C. Baker,J. Icenogle, A. Halpern, and C. Wheeler.** 1995**.** Human papillomavirus.A compilation of analysis of nucleic acid and amino acid sequences. Publication LA-UR 95-3675.Los Alamos National Laboratory,Los Alamos, N.M.

33. **Nindl, I., K. Rindfleisch, B. Lotz, A. Schneider, and M. Durst.** 1999. Uniform distribution of HPV 16 E6 and E7 variants in patients with normal histology, cervical intra-epithelial neoplasia and cervical cancer. Int J Cancer **82:**203-207.

34. **Nobbenhuis, M. A., J. M. Walboomers, T. J. Helmerhorst, L. Rozendaal, A. J. Remmink, E. K. Risse, H. C. van der Linden, F. J. Voorhorst, P. Kenemans, and C. J. Meijer.** 1999. Relation of human papillomavirus status to cervical lesions and consequences for cervical-cancer screening: a prospective study. Lancet **354:**20-25.

35. **O'Connor, M., and H. U. Bernard.** 1995. Oct-1 activates the epithelial-specific enhancer of human papillomavirus type 16 via a synergistic interaction with NFI at a conserved composite regulatory element. Virology **207:**77-88.

36. **Radhakrishna Pillai, M., S. Sreevidya, B. H. Pollock, P. G. Jayaprakash, and B. Herman.** 2002. Human papillomavirus type 16 E6 and E7 gene variations in Indian cervical cancer. Gynecol Oncol **87:**268-273.

37. **Rastogi, T., A. Hildesheim, and R. Sinha.** 2004. Opportunities for cancer epidemiology in developing countries. Nat Rev Cancer **4:**909-917.

38. **Resnick, R. M., M. T. Cornelissen, D. K. Wright, G. H. Eichinger, H. S. Fox, J. ter Schegget, and M. M. Manos.** 1990. Detection and typing of human papillomavirus in archival cervical cancer specimens by DNA amplification with consensus primers. J Natl Cancer Inst **82:**1477-1484.

39. **Sathish, N., P. Abraham, A. Peedicayil, G. Sridharan, and G. Chandy.** 2005. HPV 16 E6 sequence variations in Indian patients with cervical neoplasia. Cancer Lett **229:**93-99.

40. **Schlecht, N. F., R. W. Platt, E. Duarte-Franco, M. C. Costa, J. P. Sobrinho, J. C. Prado, A. Ferenczy, T. E. Rohan, L. L. Villa, and E. L. Franco.** 2003. Human

papillomavirus infection and time to progression and regression of cervical intraepithelial neoplasia. J Natl Cancer Inst **95:**1336-1343.

41. **Schmidt, M., W. Kedzia, and A. Gozdzicka-Jozefiak.** 2001. Intratype HPV16 sequence variation within LCR of isolates from asymptomatic carriers and cervical cancers. J Clin Virol **23:**65-77.

42. **Seedorf, K., G. Krammer, M. Durst, S. Suhai, and W. G. Rowekamp.** 1985. Human papillomavirus type 16 DNA sequence. Virology **145:**181-185.

43. **Sibbet, G. J., S. Cuthill, and M. S. Campo.** 1995. The enhancer in the long control region of human papillomavirus type 16 is up-regulated by PEF-1 and down-regulated by Oct-1. J Virol **69:**4006-4011.

44. **Sichero, L., S. Ferreira, H. Trottier, E. Duarte-Franco, A. Ferenczy, E. L. Franco, and L. L. Villa.** 2007. High grade cervical lesions are caused preferentially by non-European variants of HPVs 16 and 18. Int J Cancer **120:**1763-1768.

45. **Song, Y. S., S. H. Kee, J. W. Kim, N. H. Park, S. B. Kang, W. H. Chang, and H. P. Lee.** 1997. Major sequence variants in E7 gene of human papillomavirus type 16 from cervical cancerous and noncancerous lesions of Korean women. Gynecol Oncol **66:**275-281.

46. **Stewart, A. C., A. M. Eriksson, M. M. Manos, N. Munoz, F. X. Bosch, J. Peto, and C. M. Wheeler.** 1996. Intratype variation in 12 human papillomavirus types: a worldwide perspective. J Virol **70:**3127-3136.

47. **Stoppler, M. C., K. Ching, H. Stoppler, K. Clancy, R. Schlegel, and J. Icenogle.** 1996. Natural variants of the human papillomavirus type 16 E6 protein differ in their abilities to alter keratinocyte differentiation and to induce p53 degradation. J Virol **70:**6987-6993.

48. **Tornesello, M. L., M. L. Duraturo, I. Salatiello, L. Buonaguro, S. Losito, G. Botti, G. Stellato, S. Greggi, R. Piccoli, S. Pilotti, B. Stefanon, G. De Palo, S. Franceschi, and F. M. Buonaguro.** 2004. Analysis of human papillomavirus type-16 variants in Italian women with cervical intraepithelial neoplasia and cervical cancer. J Med Virol **74:**117-126.

49. **Touze, A., S. El Mehdaoui, P. Y. Sizaret, C. Mougin, N. Munoz, and P. Coursaget.** 1998. The L1 major capsid protein of human papillomavirus type 16 variants affects yield of virus-like particles produced in an insect cell expression system. J Clin Microbiol **36:**2046-2051.

50. **van Duin, M., P. J. Snijders, M. T. Vossen, E. Klaassen, F. Voorhorst, R. H. Verheijen, T. J. Helmerhorst, C. J. Meijer, and J. M. Walboomers.** 2000. Analysis of human papillomavirus type 16 E6 variants in relation to p53 codon 72 polymorphism genotypes in cervical carcinogenesis. J Gen Virol **81:**317-325.

51. **Veress, G., M. Murvai, K. Szarka, A. Juhasz, J. Konya, and L. Gergely.** 2001. Transcriptional activity of human papillomavirus type 16 variants having deletions in the long control region. Eur J Cancer **37:**1946-1952.

52. **Villa, L. L., P. Rahal, and E. L. Franco.** 1997. Molecular variant analysis as a tool in natural history studies of human papillomavirus infection and cervical neoplasia, New Developments in Cervical Cancer Screening and Prevention. Blackwell Science, Oxford p. 379-385.

53. **Villa, L. L., L. Sichero, P. Rahal, O. Caballero, A. Ferenczy, T. Rohan, and E. L. Franco.** 2000. Molecular variants of human papillomavirus types 16 and 18 preferentially associated with cervical neoplasia. J Gen Virol **81:**2959-2968.

54. **Wheeler, C. M., T. Yamada, A. Hildesheim, and S. A. Jenison.** 1997. Human papillomavirus type 16 sequence variants: identification by E6 and L1 lineage-specific hybridization. J Clin Microbiol **35:**11-19.

55. **Wu, Y., Y. Chen, L. Li, G. Yu, Y. He, and Y. Zhang.** 2006. Analysis of mutations in the E6/E7 oncogenes and L1 gene of human papillomavirus 16 cervical cancer isolates from China. J Gen Virol **87:**1181-1188.

56. **Yamada, T., M. M. Manos, J. Peto, C. E. Greer, N. Munoz, F. X. Bosch, and C. M. Wheeler.** 1997. Human papillomavirus type 16 sequence variation in cervical cancers: a worldwide perspective. J Virol **71:**2463-2472.

57. **Yamada, T., C. M. Wheeler, A. L. Halpern, A. C. Stewart, A. Hildesheim, and S. A. Jenison.** 1995. Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, L2, and L1 coding segments. J Virol **69:**7743-7753.

58. **Zehbe, I., G. Voglino, H. Delius, E. Wilander, and M. Tommasino.** 1998. Risk of cervical cancer and geographical variations of human papillomavirus 16 E6 polymorphisms. Lancet **352:**1441-1442.

59. **Zehbe, I., E. Wilander, H. Delius, and M. Tommasino.** 1998. Human papillomavirus 16 E6 variants are more prevalent in invasive cervical carcinoma than the prototype. Cancer Res **58:**829-833.

60. **zur Hausen, H.** 2002. Papillomaviruses and cancer: from basic studies to clinical application. Nat Rev Cancer **2:**342-350.

FIG. 1. Nucleotide sequence variations among the HPV16 isolates. E6, E7, LCR and L1 nucleotide positions at which variations were observed are written vertically across the top. The identification codes of the samples are indicated along the left. The HPV 16 reference DNA sequence is indicated as REF. For each variant sequence, positions that do not vary relative to the HPV reference sequences are marked with dashes.

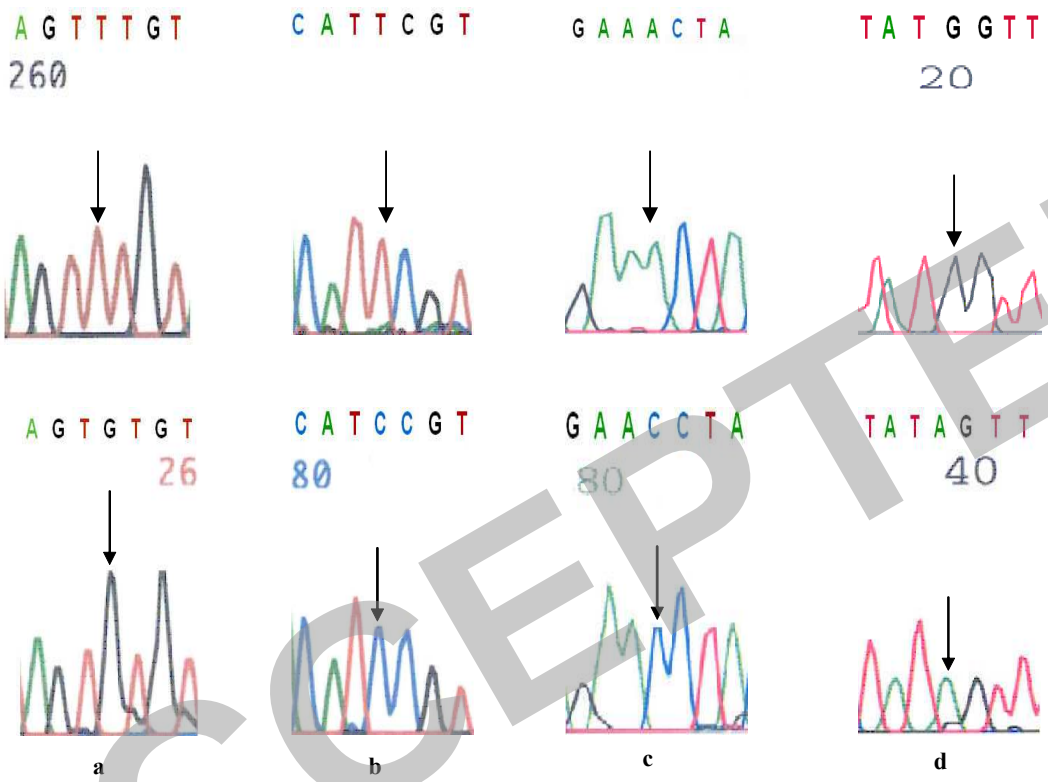

FIG.2 (a-d) Sequencing Electropherogram showing frequently detected nucleotide variations in different genomic segments of HPV-16. (a) E6 T350G, (b) E7 T789C (c) L1 A6695C and (d) LCR G7521A Upper panel shows prototype sequence while lower panel shows variant sequence indicated by arrow.

TABLE 1. Association of histo-pathological diagnosis with variants of E6, E7, LCR AND L1

Gene	WDSCC N = 22		MDSCC[b] N = 38		p-Value
	Prototype	Variant	Prototype	Variant	
E6	7 (31.8 %)	15 (68.2 %)	13 (34.2 %)	25 (69.4%)	0.84
E7	19 (86.4 %)	3 (13.6 %)	33 (86.6 %)	5 (13.8 %)	0.95
LCR	4 (18.2 %)	18 (81.8 %)	11(28.9 %)	27 (75.0%)	0.35
L1[a]	6 (60.0 %)	4 (40.0 %)	14 (73.6 %)	7 (36.8%)	0.71

[a] Variant analysis for L1 was done only in 31cases (N=10 in WDSCC, N=19 in MDSCC and N=2 in PDSCC)

[b] Two cases of PDSCC with only prototype were clubbed with MDSCC

TABLE 2: Novel variants of HPV16 E6, L1 and LCR according to nucleotide position, variant, altered amino acid and mutation

	Nucleotide Position	Prototype	Variant	Original amino acid	Altered amino acid	Codon No.	Mutation [a]	No&% of Mutation
E6	527	T	A	Serine	Threonine	142	MTs	1 (2.4%)
L1	6667	A	C	Serine	Serine	343	STv	2 (18.2%)
	6691	A.	G	Serine	Serine	351	STs	1 (9.1%)
	6906	C	T	Serine	Phenylalanine	423	MTs	2 (18.2%)
	6924	A	C	Glutamine	Proline	429	MTv	1 (9.1%)
LCR	13	C	T	-	-	-	Ts	2 (4.4%)
	7636	A	C	-	-	-	Tv	2 (4.4%)
	7678	C	T	-	-	-	Ts	1 (2%)
	7799	G	A	-	-	-	Ts	2 (4.4%)

[a] Type of mutation (M=missense, S=silent, Ts=transition, Tv=transversion)

TABLE 3: Frequency distribution of HPV TYPE 16 variants

in different phylogenetic lineages

LINEAGE	E6 N= 60	E7 N=60	L1 N= 31	LCR N= 60
European E-(P)	-	52 (86.7 %)	21 (67.7 %)	38 (63.3%)
E - P (350T)	18 (30%)	-	-	-
E - P (350G)	33 (55%)	-	-	-
E - G (131)	-	-	1 (3.2%)	-
Asian-American (AA)	2 (3.3%)	-	4 (12.9%)	8 (13.3 %)
North-American (NA)	6 (10%)	-	2 (6.5%)	-
African-1 (Af-1)	-	-	-	7 (11.6%)
African-2 (Af-2)	-	7 (11.6%)	-	2 (3.3%)
Asian (As)	-	1 (1.7%)	-	-
New variants	1 (1.7%)	-	3 (9.7%)	5 (8.3%)