Infection, Genetics and Evolution 46 (2016) 7–11

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

Research paper

Genomic diversity and phylogenetic relationships of human papillomavirus 16 (HPV16) in Nepal

Robert Makowsky [a,1], Pema Lhaki [b], Howard W. Wiener [c], Madhav P. Bhatta [d], Michael Cullen [e,f], Derek C. Johnson [c], Rodney T. Perry [c], Mingma Lama [b], Joseph F. Boland [e,f], Meredith Yeager [e,f], Sarita Ghimire [g], Thomas R. Broker [h], Sadeep Shrestha [c,*]

[a] Department of Biostatistics, School of Public Health, University of Alabama at Birmingham, AL, USA
[b] NFCC International, Kathmandu, Nepal
[c] Department of Epidemiology, School of Public Health, University of Alabama at Birmingham, AL, USA
[d] Department of Biostatistics, Environmental Health Sciences, and Epidemiology, College of Public Health, Kent State University, OH, USA
[e] Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Rockville, MD, USA
[f] Cancer Genomics Research Laboratory, Frederick National Laboratory for Cancer Research, Leidos Biomedical Research Inc., Gaithersburg, MD, USA
[g] Cancer Screening Center, Nepal Cancer Care Foundation, Lalitpur, Nepal
[h] Department of Biochemistry and Molecular Genetics, University of Alabama at Birmingham, Birmingham, AL, USA

ARTICLE INFO

Article history:
Received 2 September 2016
Received in revised form 5 October 2016
Accepted 6 October 2016
Available online 7 October 2016

Keywords:
HPV16 sequence
Nepal
Phylogenetic tree
HSIL
HPV16 lineage

ABSTRACT

Objectives/Background: Sequence variants in HPV16 confer differences in oncogenic potential; however, to date there have not been any HPV sequence studies performed in Nepal. The objective of this study was to characterize HPV16 viral genome sequences from Nepal compared to a reference sequence in order to determine their lineages. Additionally, we sought to determine if five High-grade Squamous Intraepithelial Lesion (HSIL) subjects were genetically distinct from the non-HSIL subjects.
Methods: DNA was isolated from exfoliated cervical cells from 17 individuals in Nepal who were previously identified to be HPV16-positive. A custom HPV16 Ion Ampliseq panel of multiplexed degenerate primers was designed that generated 47 overlapping amplicons and covered 99% of the viral genome for all known HPV16 variant lineages. All sequence data were processed through a custom quality control and analysis pipeline of sequence comparisons and phylogenetic analysis.
Results: There were high similarities across the genomes, with two major indels observed in the non-coding region between E5 and L2. Compared to the PAVE reference HPV16 genome, there were up to 9, 4, 38, 27, 8, 7, 52, and 32 nucleotide variants in the E6, E7, E1, E2, E4, E5, L2, and L1 genes in the Nepalese samples, respectively. Based on sequence variation, HPV16 from Nepal falls across the A, C, and D lineages in this study. We found no evidence of genetic distinctness between HSIL and non-HSIL subjects.
Conclusions: The evolutionary and pathological characteristics of the representative HPV16 genomes from Nepal seem similar to results from other parts of the world and provide the basis for further studies.
© 2016 Elsevier B.V. All rights reserved.

1. Introduction

HPV is a double-stranded, 8 kb DNA virus with a non-enveloped icosahedral virion of 55 nm diameter. It is a small tumor virus, encoding only 9–10 genes, and it interacts with the host-cell proteins to complete its productive life cycle (Iftner, 1990). The relative arrangement of 9–10 open reading frames (ORF) within the genome is conserved in all known papillomavirus types (de Villiers et al., 2004; Munger, 2002). In humans, high-risk HPV types are the main etiological agents for cervical cancer (zur Hausen, 2002). HPV type 16 (HPV16), belonging to the genus *Alphapapillomavirus* 9 and family Papillomaviridae (Chen et al., 2011), is the most potent carcinogen and found in the majority of cervical cancer cases (Munoz et al., 2003). Since the first sequencing of the HPV16 genome, accession K02718 (Seedorf et al., 1985), various naturally-occurring variants have been described. Phylogenetic analyses of primarily E6 and/or the upstream regulatory region (URR) have described various lineages and sub-lineages.

Studies have shown that some HPV16 variants are more oncogenic than others and could also result in differential viral persistence and progression to cervical cancer (Schiffman et al., 2010; Zehbe et al.,

* Corresponding author at: University of Alabama at Birmingham, School of Public Health, Department of Epidemiology, Ryals Rm217L, 1720 2nd Ave S., Birmingham, AL 35294-0022, USA.
E-mail address: sshrestha@uab.edu (S. Shrestha).
[1] Current address: HTG Molecular Diagnostics, Arizona, USA.

http://dx.doi.org/10.1016/j.meegid.2016.10.004
1567-1348/© 2016 Elsevier B.V. All rights reserved.

Table 1
Total number of nucleotide variants by gene per individual compared to the reference genome (NC_001526) with overall summaries.

	E6	E7	E1	E2	E4	E5	L2	L1
Achham1	0	1	1	0	0	0	5	2
Achham2	6	2	24	19	7	4	28	14
Achham3	6	3	25	19	7	4	29	16
Achham4	0	1	1	1	1	0	5	2
Achham5	0	0	3	3	1	1	2	2
Jhapa1	1	0	2	2	1	2	2	1
Jhapa2	1	0	1	1	1	2	2	1
Jhapa3	1	0	4	2	1	1	2	2
Jhapa4	1	0	2	1	1	2	4	3
Jhapa5	1	0	1	1	1	2	3	1
Jhapa6	1	0	1	1	1	2	3	3
Jhapa7	0	0	0	1	1	2	3	4
Jhapa8	1	0	3	3	1	1	2	3
Jhapa9	0	0	4	2	1	1	2	2
Jhapa10	1	0	1	1	1	3	3	2
Jhapa11	6	2	23	18	5	4	25	19
Jhapa12	1	0	1	1	1	3	5	1
# Total sites	477	297	1950	1103	288	252	1422	1518
# Variable sites	9	4	38	27	8	7	52	32
% Variable Sites	1.89%	1.35%	1.95%	2.45%	2.78%	2.78%	3.66%	2.11%

1998a; Londesborough et al., 1996). For example, the incidence of cervical cancer varies between 4.4 (Western Asia) and 75.9 (Malawi) per 100,000 women (World Health Organization, 2012). In the US, the rates are 9.4, 8.9 and 7.5 per 100,000 women in Hispanics, African American and European Americans, respectively (Howlader et al., 2016). However, African-American women have the highest mortality from invasive cervical cancer, at 3.9 per 100,000 US women, compared to 2.1 for European Americans and 2.6 for Hispanics (Howlader et al., 2016). Although racial differences in incidence and mortality for cervical cancer can be viewed in relation to accessibility and use of screening, diagnosis, and treatment, different genetic makeups of both the host and virus and cellular mechanisms apparently regulate the biology and pathogenesis of the disease. In particular, the non-European HPV16 lineage contributes more to the infection persistence and cervical cancer progression than the European-lineage (Freitas et al., 2014; Sichero et al., 2007; Mirabello et al., 2016). Likewise, studies have shown that a change from T to G at nucleotide position 350 in HPV 16 E6 gene, resulting in change of amino acid leucine (L) to valine (V), is associated progression of cervical neoplasia (Zehbe et al., 1998b).

The countries of South Asia contribute approximately one-third of the world's cervical cancer cases. With 24.2 cases per 100,000 individuals, Nepal has one of the highest rates of cervical cancer in South Asia (Bruni et al., 2016). Approximately 2150 cases of cervical cancer occur in Nepal each year, resulting in over 1100 deaths for a mortality rate of over 50% (Bruni et al., 2016). Despite the unknown yet potential burden of cervical cancer in Nepal, there is very little information on the prevalence of the prevailing HPV genotypes and lineages. To address the missing gap of HPV16 lineages in Nepal, we have sequenced the entire HPV16 genome from 17 individuals in two major regions of Nepal and examined the classification of HPV16 variant sub-lineages by using the known lineage and sub-lineage (geographical) references.

2. Materials and methods

2.1. Origin of clinical specimens

We previously conducted health camps for HPV screening in two regions of Nepal, first in the Accham district in far-west Nepal (Johnson et al., 2014) and second in the Jhapa district in far-east Nepal. The protocols have been previously described in detail for Accham (Johnson et al., 2014) and the same procedures were followed for Jhapa. Briefly, cervical specimens were a) self-collected using the APTIMA Cervical Specimen Collection and Transport (CSCT) kit (Hologic/Gen-Probe, San Diego, CA) collected by Auxiliary Nurse Midwives (ANMs) using ThinPrep® PreservCyt® medium (Hologic/Gen-Probe, San Diego, CA). Cervical specimens were tested for HPV genotype first using generic APTIMA® HR-HPV mRNA (APTIMA® HPV) (Hologic/Gen-Probe, San Diego, CA), which detects the presence of E6/E7 mRNA for 14 types of HR-HPV (HPV 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, 68) and all positives were retested using APTIMA® HPV16 18/45 Genotype (Hologic/Gen-Probe, San Diego, CA) Assays. In Accham, 5/261 individuals tested positive for HPV 16 in both self-collected and clinician-collected samples (Johnson et al., 2014). In Jhapa, 12 individuals tested positive for HPV16 in both samples. The remaining aliquots of all samples that tested positive for HPV16 were used for sequencing in this study. Cervical cytology was assessed for research purposes using clinician-collected ThinPrep® PreservCyt® medium (Hologic/Gen-Probe, San Diego, CA), with results classified according to the Bethesda criteria. In Accham, three women had high-grade squamous intraepithelial lesion (HSIL) and one had squamous cell cancer (SCC) (Johnson et al., 2014; Johnson et al., 2016). In Jhapa, only one individual had HSIL.

Among married women, the geographic location of the husband's migration was categorized as “Migrated within Nepal for work”, “Migrated outside of Nepal for work”, or “Has never migrated for work”. Further, in Jhapa since there is a Bhutanese refugee camp, several questions about immigration and place of birth were asked in the survey questionnaire to assess the refugee status of each individual. If the individual was born in Bhutan and answered that they lived in the Bhutanese refugee camp, they were considered “refugee”.

This study was conducted following the principles of the Helsinki Declaration of 1975. Institutional review boards from the University of Alabama at Birmingham and the Nepal Health Research Council approved this study.

2.2. DNA isolation

DNA was extracted using QIAamp MiniElute Media Kit, following the manufacturer's manual (Qiagen, Valencia, CA).

2.3. Ion AmpliSeq library preparation and sequencing

A custom Ion AmpliSeq HPV16 panel consisting 47 overlapping amplicons that ranged from 181 bp to 375 bp in size was employed to amplify 99% of 7906 bp HPV16 genome, as previously described (Cullen et al., 2015). Because the genome is circular, the first amplicon and the last amplicon overlapped and therefore contained a short

Table 2
Total number of amino acid variants by gene per individual compared to the reference genome (NC_001526) with overall summaries (AA = Amino Acid).

	E6	E7	E1	E2	E4	E5	L2	L1
Achham1	0	1	1	0	0	0	3	1
Achham2	3	0	10	13	2	1	10	7
Achham3	3	0	10	13	2	1	11	7
Achham4	0	1	1	1	1	0	2	0
Achham5	0	0	1	1	0	1	1	1
Jhapa1	1	0	2	1	0	2	1	1
Jhapa2	1	0	1	1	0	2	1	1
Jhapa3	1	0	2	1	0	1	1	1
Jhapa4	1	0	1	1	0	1	3	2
Jhapa5	1	0	1	1	0	2	2	1
Jhapa6	1	0	1	1	0	1	2	1
Jhapa7	0	0	0	1	0	2	2	2
Jhapa8	1	0	1	1	0	1	1	1
Jhapa9	0	0	2	1	0	1	1	1
Jhapa10	1	0	1	1	0	2	2	2
Jhapa11	3	0	6	12	2	1	10	7
Jhapa12	1	0	1	1	0	2	2	1
# Total AA	158	98	649	365	95	83	473	507
# Variable AA	5	1	14	15	3	2	22	11
% Variable AA	3.16%	1.02%	2.16%	4.11%	3.16%	2.41%	4.65%	2.17%

stretch of identical sequence. Custom HPV16 degenerate primers were designed using a consensus sequence with ambiguity codes (IUPAC) derived from seven sequences representing major HPV16 lineages, 2 European (lineage A), 1 European/Asian (sub-lineage A-4), 1 African-1 (lineage B), 1 African-2 (lineage C) and 2 Asian (lineage D) (Burk et al., 2013).

Libraries were generated following the manufacturer's Ion AmpliSeq Library Preparation kit 2.0-96LV protocol (Life Technologies, Part #4480441) with modifications, as previously described (Cullen et al., 2015). PCR based viral sequencing was performed with 2 μl of samples utilizing the Ion Torrent Personal Genome Machine (PGM; Life Technologies, Carlsbad CA).

For sequence alignment purposes, the first and last amplicons were bioinformatically split to account for the sequence overlap thus creating 48 overlapping sequence contigs. Raw sequencing reads generated by the Ion Torrent sequencer were quality and adaptor trimmed by Ion Torrent Suite prior to alignment to the HPV16R reference sequence (7906 bp) (Myers et al., 1997) using TMAP. A human hg19-HPV16 hybrid reference genome was used to filter out human reads prior to generating QC metrics and HPV16 variant calls.

2.4. Phylogenetic tree construction and statistical analyses

Eighteen reference sequences were included (Smith et al., 2011) and detailed in Appendix Table 1. These reference sequences were downloaded from Genbank and included to represent currently recognized lineages and sub-lineages (Burk et al., 2013). Alignment of the seventeen study sequences (KX947269, KX947270, KX947271, KX947272, KX947273, KX947274, KX947275, KX947276, KX947277, KX947278, KX947279, KX947280, KX947281, KX947282, KX947283,

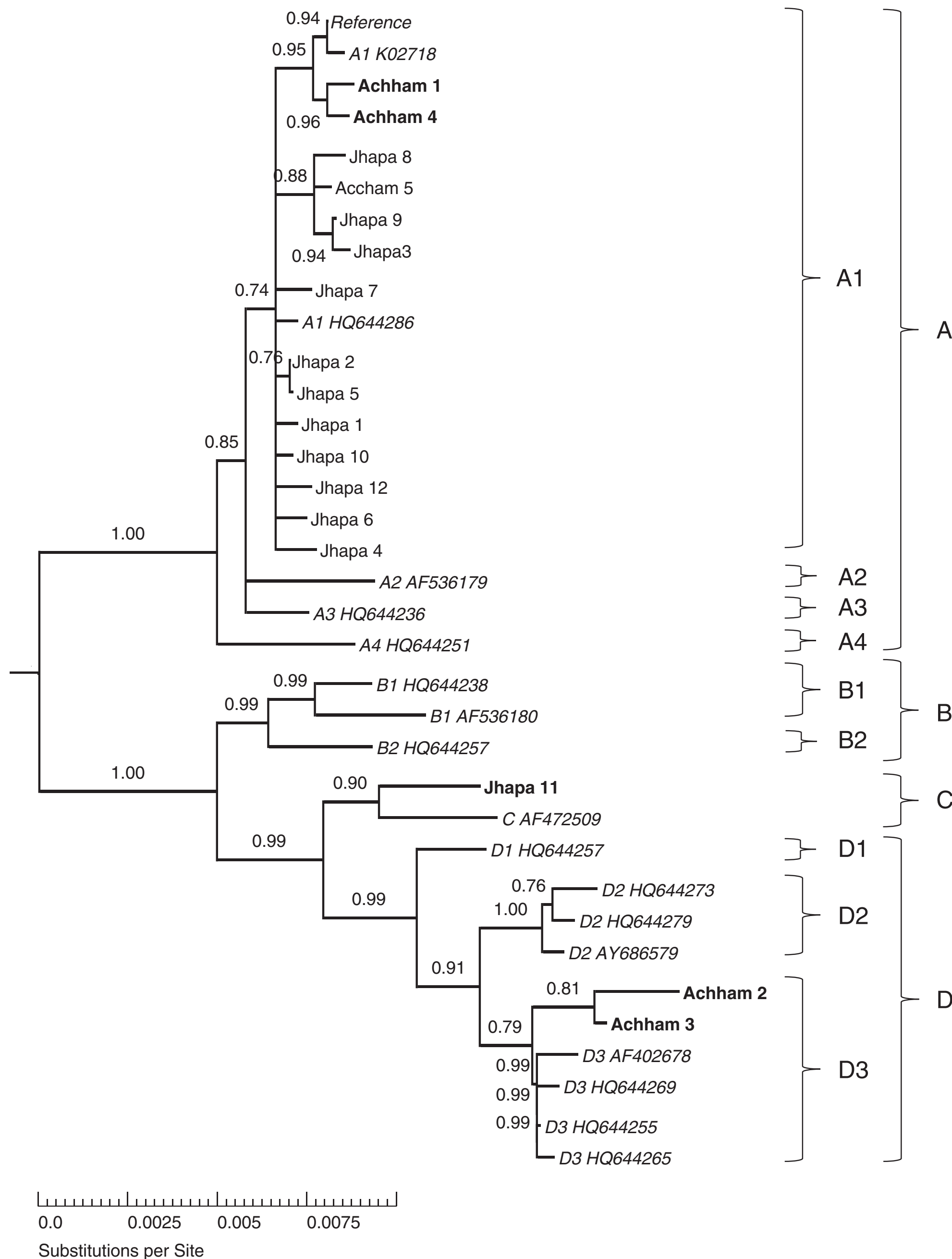

Fig. 1. Phylogenetic tree based on 35 complete HPV genome sequences, 17 from two distinct regions in Nepal and 18 reference sequences comprising all known lineages and sub-lineages. Nodes with bootstrap support <0.70 were collapsed. Reference sequences and their Genbank ID were italicized (further details in Appendix Table 1). HSIL subjects are bolded.

KX947284, KX947285) and the eighteen reference sequences was initially done using the Muscle option in MEGA version 6.0 and then edited by hand. Maximum-likelihood was performed using MEGA software version 6.0 assuming the General Time Reversible model with Gamma rate distribution and invariant sites (GTR + I + G) using a maximum parsimony starting tree. Node support was assessed by bootstrapping with 1000 replications, and nodes with bootstrap proportions <0.70 were collapsed. A discrete gamma distribution was used to model evolutionary rate differences among sites (5 categories (+ G, parameter = 0.1000)). The rate variation model allowed for some sites to be evolutionarily invariable (+ I, 7.4% sites). The tree was drawn to scale, with branch lengths measured in the number of substitutions per site.

Logistic regressions were utilized to determine if the five HSIL subjects (Achham 1, Achham 2, Achham 3, Achham 4, and Jhapa 11) with HSIL or SCC had any differences that separated them from the remaining subjects. To accomplish this, we tried two approaches. First, we analyzed each variable nucleotide site singly. Second, each gene was translated and we looked to see if any combination of variable amino acids could perfectly classify the two groups. Finally, we calculated the odds ratio for the 350G/T variant with HSIL given its previously reported importance.

3. Results

The aligned dataset consisted of 35 nucleotide sequences and a total of 7918 bp. The sequences matched well with the reference genome with the exception of Lineages B, C, and D, which had several insertions and deletions in the non-coding region between E5 and L2. The maximum overall pairwise Jukes-Cantor distance observed among Nepalese and reference subjects was 1.892% and Table 1 depicts the number of pairwise nucleotide differences between each Nepalese subject and the reference genome per gene. Across the alignment of 7918 nucleotides, there were 406 variables sites (5.1%). Table 2 depicts the observed amino acid differences between each Nepalese subject and the reference genome per gene.

The best ML tree had a log likelihood of −14,900.3 (Fig. 1). After collapsing nodes containing bootstrap support proportions <0.70, the Nepalese samples fell across 3 lineages (A, C, and D). Within the A lineage, all Nepalese samples clustered within the A1 sub-lineage reference samples. Within the D lineage, all Nepalese subjects clustered within the D3 sub-lineage. The five HSIL subjects did not comprise a single monophyletic lineage but were instead spread out across the phylogenetic tree (Table 3).

Table 3
Migration, refugee status and cervical cytology of individuals with HPV16 infection at two screening sites.

Tree name	Lineage	Sub-lineage	Female migration	Husband migration	Cytology	Refuge
Achham1	A	A1	No	No	HSIL*	NA
Achham2	D	D3	No	India	HSIL*	NA
Achham3	D	D3	No	India	SCC§	NA
Achham4	A	A1	No	India	HSIL*	NA
Achham5	A	A1	No	No	normal	NA
Jhapa1	A	A1	NA	No	normal	Yes
Jhapa2	A	A1	No	No	normal	Yes
Jhapa3	A	A1	No	No	normal	No
Jhapa4	A	A1	NA	Qatar	normal	No
Jhapa5	A	A1	No	No	normal	No
Jhapa6	A	A1	No	No	normal	Yes
Jhapa7	A	A1	No	Kathmandu	normal	Yes
Jhapa8	A	A1	No	No	normal	No
Jhapa9	A	A1	No	NA	normal	No
Jhapa10	A	A1	No	Qatar	normal	No
Jhapa11	C	NA	No	NA	HSIL*	No
Jhapa12	A	A1	NA	Kathmandu	normal	No

*HSIL — high-grade squamous intraepithelial lesions; §SCC — squamous cell carcinoma.

There were no variable nucleotides or combination of amino acids within genes that would perfectly separate the HSIL from non-HSIL subjects. Three of the 11 subjects with the 350 T variant exhibited HSIL while 2 of the 12 subjects with the G variant exhibited HSIL (odds ratio of developing HSIL with a T compared to a G of 1.875; 95% CI of 0.25 to 14.08).

4. Discussion

This study sought to explore the genetic diversity and phylogenetic signal across HPV samples from people living in Nepal. Whole HPV16 genomes were sequenced for 17 individuals located throughout the country with varying travel histories. We observed that HPV samples fell across the A, C, and D lineages, with the majority of subjects within the A lineage. We found that genetic variation was not constant across genes. At the nucleotide level, variability per gene ranged from 1.35% for E7 to 3.66% for E4. At the amino acid level, variability per gene ranged from 1.02% to 4.65% for the E7 and L2 genes, respectively. Some of the variants were in the overlapping E2/E4 regions and it is not clear if these variants would change one or both genes as we did not have expression data. Finally, we were unable to find any genetic variants that could explain HSIL status.

Our phylogenetic analyses produced a tree where the Nepalese samples were clustered within the A, C, and D HPV lineages. Multiple sub-lineages are recognized within the A (n = 4) and D (n = 3) sub-lineages (B also has two recognized sub-lineages, but no representative Nepalese subjects in this study), but for both lineages, Nepalese subjects clustered within one sub-lineage (A1 and D3). All but one of the samples from Jhapa grouped with the A lineage, while the remaining sample (Jhapa 11) grouped with lineage C. The Achham samples were more diverse, with 3 samples falling within the A lineage (A1 sub-lineage) and 2 grouping with lineage D (D3 sub-lineage). The subject's reported travels, spouse's travels, or refugee status did not consistently offer an explanation for these patterns, meaning that multiple lineages must have been actively spreading within Jhapa and Achham. Most HPV16 lineage studies in neighboring country India has also been based on E6, E7, L1 and URR regions and not the entire HPV genome; however, in most case lineages A occurs in high prevalence (Pande et al., 2008; Sathish et al., 2005; Kabekkodu et al., 2015).

Finding one or more genetic variants that are strongly associated with HSIL status could have important clinical ramifications. For example, the 350G/T variant has been found to be associated with HSIL, but the variant that is associated with higher HSIL risk is population dependent (Cornet et al., 2013). In Nepal, Cornet et al. (2013) only observed HSIL in subjects with the T variant while in neighboring India, a study found the G variant to increase the odds of HSIL (Sathish et al., 2005). In this study, the odds ratio of developing HSIL with a T compared to a G was >1 (1.875), but the 95% CI included 1 and the p-value was far from statistically significant (0.5412). Therefore, our study finds no evidence that either variant is associated with greater odds of developing HSIL.

Given our small sample size (5 subjects with HSIL), we opted to not conduct significance tests but to instead look for perfect separation between HSIL and non-HSIL subjects. This was performed at both the univariate nucleotide level as well as the multivariate amino acid level within genes. At both levels, we did not find any variants (combinations thereof within genes) that perfectly grouped HSIL and non-HSIL subjects. Additionally, we found that HSIL subjects did not contain HPV lineages that clustered together phylogenetically, but instead were spread out across the phylogenetic tree and together accounted for 167 of the 406 total variable sites. Specifically, two of the five HSIL subjects (Achham 1 and 4) were infected with strains from the A lineage, one subject (Jhapa 11) was infected with a strain most closely related to the C lineage, and the remaining two (Achham 2 and 3) with strains most closely related to the D lineage.

The results from this study contribute to the field of molecular epidemiology of HPV16 in two regions in Nepal. Interestingly, even in a small sample-size, we observed samples closely related to three of the four (A, C, D) currently recognized HPV lineages. While we observed several variants, the phylogenetic tree suggests no major differences in lineages of HPV16 from other parts of the world. Studies with a larger sample size as well as in other high-risk HPV variants are warranted to fully understand the circulating lineages in the region.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.meegid.2016.10.004.

Disclosure of conflicts of interest

None.

Acknowledgements

We thank the study participants and the staff at Nepal Fertility Care Center (NFCC) who helped with the sample collection, particularly Pankaj Bhattarai who was instrumental in the health camps. The Family Health Division, within the Nepal ministry of health also assisted with the logistics of the health camp and data collection. The study was supported in part by the pilot fund from the Department of Epidemiology, School of Public Health, University of Alabama at Birmingham and College of Public Health, Kent State University. Derek Johnson was supported by a National Institutes of Health Cancer Prevention and Control Training Grant (R25CA47888).

References

Bruni, L., Barrionuevo-Rosas, L., Albero, G., Aldea, M., Serrano, B., Valencia, S., Brotons, M., Mena, M., Cosano, R., Muñoz, J., Bosch, F.X., de Sanjosé, S., Castellsagué, X., 2016. ICO Information Centre on HPV and Cancer (HPV Information Centre). Human Papillomavirus and Related Diseases in Nepal. Summary Report 2016-02-26 (Date Accessed).

Burk, R.D., Harari, A., Chen, Z., 2013. Human papillomavirus genome variants. Virology 445 (1–2), 232–243.

Chen, Z., Schiffman, M., Herrero, R., Desalle, R., Anastos, K., Segondy, M., Sahasrabuddhe, V.V., Gravitt, P.E., Hsing, A.W., Burk, R.D., 2011. Evolution and taxonomic classification of human papillomavirus 16 (HPV16)-related variant genomes: HPV31, HPV33, HPV35, HPV52, HPV58 and HPV67. PLoS One 6 (5), e20183.

Cornet, I., Gheit, T., Iannacone, M.R., Vignat, J., Sylla, B.S., Del Mistro, A., Franceschi, S., Tommasino, M., Clifford, G.M., 2013. HPV16 genetic variation and the development of cervical cancer worldwide. Br. J. Cancer 108 (1), 240–244.

Cullen, M., Boland, J.F., Schiffman, M., Zhang, X., Wentzensen, N., Yang, Q., Chen, Z., Yu, K., Mitchell, J., Roberson, D., Bass, S., Burdette, L., Machado, M., Ravichandran, S., Luke, B., Machiela, M.J., Andersen, M., Osentoski, M., Laptewicz, M., Wacholder, S., Feldman, A., Raine-Bennett, T., Lorey, T., Castle, P.E., Yeager, M., Burk, R.D., Mirabello, L., 2015. Deep sequencing of HPV16 genomes: a new high-throughput tool for exploring the carcinogenicity and natural history of HPV16 infection. Papillomavirus Res. 1, 3–11.

de Villiers, E.M., Fauquet, C., Broker, T.R., Bernard, H.U., zur Hausen, H., 2004. Classification of papillomaviruses. Virology 324 (1), 17–27.

Freitas, L.B., Chen, Z., Muqui, E.F., Boldrini, N.A., Miranda, A.E., Spano, L.C., Burk, R.D., 2014. Human papillomavirus 16 non-European variants are preferentially associated with high-grade cervical lesions. PLoS One 9 (7), e100746.

Howlader, N., Noone, A.M., Krapcho, M., Miller, D., Bishop, K., Altekruse, S.F., Kosary, C.L., Yu, M., Ruhl, J., Tatalovich, Z., Mariotto, A., Lewis, D.R., Chen, H.S., Feuer, E.J., Cronin, K.A. (Eds.), April 2016. SEER Cancer Statistics Review, 1975–2013. National Cancer Institute, Bethesda, MD http://seer.cancer.gov/csr/1975_2013/, based on November 2015 SEER data submission, posted to the SEER web site.

Iftner, T., 1990. Papillomavirus genomes sequence analysis related to functional aspects. In: H., P. (Ed.), Papillomaviruses and Human Cancer. CRC Press, Boca Raton Florida, pp. 181–203.

Johnson, D.C., Bhatta, M.P., Smith, J.S., Kempf, M.C., Broker, T.R., Vermund, S.H., Chamot, E., Aryal, S., Lhaki, P., Shrestha, S., 2014. Assessment of high-risk human papillomavirus infections using clinician- and self-collected cervical sampling methods in rural women from far western Nepal. PLoS One 9 (6), e101255.

Johnson, D.C., Lhaki, P., Bhatta, M.P., Kempf, M.C., Smith, J.S., Bhattarai, P., Aryal, S., Chamot, E., Regmi, K., Vermund, S.H., Shrestha, S., 2016. Spousal migration and human papillomavirus infection among women in rural western Nepal. Inter. Health 8 (4), 261–268.

Kabekkodu, S.P., Bhat, S., Pandey, D., Varghese, V.K., Shukla, V., Ghosh, S., Kushtagi, P., Bhat, P., Gopinath, P.M., Satyamoorthy, K., 2015. Prevalence of human papillomavirus types and phylogenetic analysis of HPV-16 L1 variants from Southern India. Asian Pac. J. Cancer Prev. 16 (5), 2073–2080.

Londesborough, P., Ho, L., Terry, G., Cuzick, J., Wheeler, C., Singer, A., 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 (5), 364–368.

Mirabello, L., Yeager, M., Cullen, M., Boland, J.F., Chen, Z., Wentzensen, N., Zhang, X., Yu, K., Yang, Q., Mitchell, J., Roberson, D., Bass, S., Xiao, Y., Burdett, L., Raine-Bennett, T., Lorey, T., Castle, P.E., Burk, R.D., Schiffman, M., 2016. HPV16 sublineage associations with histology-specific cancer risk using HPV whole-genome sequences in 3200 women. J. Natl. Cancer Inst. 108 (9).

Munger, K., 2002. The role of human papillomaviruses in human cancers. Front. Biosci. 7, d641–d649.

Munoz, N., Bosch, F.X., de Sanjose, S., Herrero, R., Castellsague, X., Shah, K.V., Snijders, P.J., Meijer, C.J., International Agency for Research on Cancer Multicenter Cervical Cancer Study G, 2003. Epidemiologic classification of human papillomavirus types associated with cervical cancer. N. Engl. J. Med. 348 (6), 518–527.

Myers, G., Baker, C., Münger, K., Sverdrup, F., McBride, A., Bernard, H.-U., Meissner, J., 1997. Human Papillomaviruses: A Compilation and Analysis of Nucleic Acid and Amino Acid Sequences. Theoretical Biology and Biophysics Division, Los Alamos National Laboratory, Los Alamos, N. Mex.

Pande, S., Jain, N., Prusty, B.K., Bhambhani, S., Gupta, S., Sharma, R., Batra, S., Das, B.C., 2008. 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. 46 (3), 1060–1066.

Sathish, N., Abraham, P., Peedicayil, A., Sridharan, G., Chandy, G., 2005. HPV 16 E6 sequence variations in Indian patients with cervical neoplasia. Cancer Lett. 229 (1), 93–99.

Schiffman, M., Rodriguez, A.C., Chen, Z., Wacholder, S., Herrero, R., Hildesheim, A., Desalle, R., Befano, B., Yu, K., Safaeian, M., Sherman, M.E., Morales, J., Guillen, D., Alfaro, M., Hutchinson, M., Solomon, D., Castle, P.E., Burk, R.D., 2010. A population-based prospective study of carcinogenic human papillomavirus variant lineages, viral persistence, and cervical neoplasia. Cancer Res. 70 (8), 3159–3169.

Seedorf, K., Krammer, G., Durst, M., Suhai, S., Rowekamp, W.G., 1985. Human papillomavirus type 16 DNA sequence. Virology 145 (1), 181–185.

Sichero, L., Ferreira, S., Trottier, H., Duarte-Franco, E., Ferenczy, A., Franco, E.L., Villa, L.L., 2007. High grade cervical lesions are caused preferentially by non-European variants of HPVs 16 and 18. Int. J. Cancer 120 (8), 1763–1768.

Smith, B., Chen, Z., Reimers, L., van Doorslaer, K., Schiffman, M., Desalle, R., Herrero, R., Yu, K., Wacholder, S., Wang, T., Burk, R.D., 2011. Sequence imputation of HPV16 genomes for genetic association studies. PLoS One 6 (6), e21375.

World Health Organization, 2012. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase no.11 [internet]. Available from: http://globocan.iarc.fr. Archived at: http://www.webcitation.org/6b3ANK8Ne. In.; 2012.

Zehbe, I., Voglino, G., Delius, H., Wilander, E., Tommasino, M., 1998a. Risk of cervical cancer and geographical variations of human papillomavirus 16 E6 polymorphisms. Lancet 352 (9138), 1441–1442.

Zehbe, I., Wilander, E., Delius, H., Tommasino, M., 1998b. Human papillomavirus 16 E6 variants are more prevalent in invasive cervical carcinoma than the prototype. Cancer Res. 58 (4), 829–833.

zur Hausen, H., 2002. Papillomaviruses and cancer: from basic studies to clinical application. Nat. Rev. Cancer 2 (5), 342–350.