

# The Biological Impact of Genomic Diversity in Cervical Cancer Development

Cristina Mendes de Oliveira José Eduardo Levi

Laboratório de Virologia, Instituto de Medicina Tropical, Universidade de São Paulo, São Paulo, Brazil

## Key Words

Human papillomavirus · Cervical cancer · Genetic diversity

## Abstract

Human papillomaviruses (HPVs) are the etiologic agents of cervical cancer, the unique human neoplasia that has one single necessary cause. The diversity of HPVs is well described, with 200 HPV types existing as distinct taxonomic units and each receiving an Arabic number. On a clinical basis, they are usually grouped by their site of occurrence and disease associations. Those types inhabiting the anogenital mucosa are more intensively studied and further divided into cancer-associated HPVs, which are termed 'high risk', while those linked to benign proliferative lesions are assigned as 'low risk'. HPV16 is responsible for approximately 50% of all ICC cases, and paradoxically is one of the most prevalent types among healthy women. Longitudinal studies have shown that when an incidental HPV16 infection becomes persistent it will result in an enhanced risk for the development of high-grade lesions. However, it is unknown why some persistent, HPV16 infections (or infections by other HR-HPV types) progress to CIN3+ while most clear spontaneously. Several epidemiological investigations have focused on cofactors, from the most obvious such as cigarette

and other carcinogenic exposures, to coinfections by other STDs such as chlamydia, with no significant findings. Thus, the current focus is on genomic variation from both virus and host. Such studies have been potentialized by the enormous technical advances in nucleic acid sequencing, allowing this relationship to be broadly interrogated. Corroborating subgenomic data from decades ago, an association between HPV16 lineages and carcinogenesis is being revealed. However, this effect does not seem to apply across female populations from different continents/ethnicities, again highlighting a role played by HPV16 adaptation and evasion from the host over time.

© 2016 S. Karger AG, Basel

## Introduction

Papillomaviruses (PVs) comprise a group of small, nonenveloped viruses with a double-stranded DNA genome of approximately 8 kbp that infect vertebrates. PVs have been isolated from diverse hosts, including mammals, birds, turtles, and snakes, suggesting they infect all amniotes [1]. It is believed that their origin dates to around 350 million years ago and is linked to changes in the epithelium of their ancestral host, the first reptiles.

Viruses that slowly evolve with their hosts, as PVs do, usually cause asymptomatic infections rather than severe disease [2]. However, certain human papillomavirus (HPV) infections can cause benign and malignant proliferative disorders, such as skin warts, epidermodysplasias, condylomata acuminata and malignant neoplasms of the anogenital region, mainly in the cervix. Molecular evidence suggests that HPV is also involved in a proportion of cancer cases in the vagina, vulva, anus, penis, and in the head and neck [3]. To date, more than 200 HPVs have been fully sequenced (<http://pave.niaid.nih.gov/>). HPVs are classified into 5 genera, with different types showing differential associations and disease spectra [4–6].

### Origin of PVs Diversity and Mutation Rates

In the infected epithelial cell, PV replication occurs during the S-phase using high-fidelity cellular DNA polymerases with error correction leading to mutation rates close to those of the hosts [1]. However, codon usage varies between species and between genes within the same genome [7], and HPVs do not have the same codon usage preferences of the host, since the viral genome is enriched in A+T [8, 9]. All PV genes show poor adaptation to human codon usage preference, and orthologous genes in PVs with a similar clinical presentation display similar codon usage preferences [10]. Regarding mutation, the APOBEC3 internal mutators have been shown to target PV DNA, introducing directional C>T changes [11, 12]. The APOBEC3 locus has undergone a large expansion in the human lineage [13], which may partially account for the broad diversity of human PVs. The substitution rate estimates for PV coding regions are between  $2 \times 10^8$  and  $5 \times 10^9$  substitutions per site per year [14, 15].

### HPV Diversity and Cancer

The phylogenetic relationship of HPVs is classified at the genera and species level, according to the International Committee on Taxonomy of Viruses (ICTV). Below the species level, the PV research community classifies viruses that are 10% dissimilar on the L1 gene as a different HPV type, and as a subtype when 2–10% dissimilar [4]. Recently, it was proposed that differences ranging from 1 to 2% of the full genome define the variant level and from 0.5 to 1% define sublineages [4, 16].

Infections by most HPVs are asymptomatic [17], but 12 types are classified by the International Agency for Re-

search on Cancer as definitely oncogenic: HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58 and 59 [18]. Carcinogenic potential is not evenly distributed among HPVs, with types 16 and 18 responsible for 70% of invasive cervical cancer (ICC) cases worldwide [19, 20].

Different HPV16 and 18 variants and sublineages have been identified based on partial regions of the genome [21–25] and the complete genome [26–29], and variants of both types are associated with different risks of ICC development [30–33]. Differences in the oncogenic potential displayed by variants of the same HPV type have been described; HPV16 B, C, D2 and D3 variants promote a 3-fold higher risk of cervical cancer development compared to the A1 sublineage. The non-European HPV18 variants are identified more frequently in cancer tissues and high-grade squamous lesions [31, 34–36]. HPV33 (C7732G) and HPV58 (C632T and G760A) variants have been associated with a higher risk of cervical cancer development [37, 38]. In addition, the HPV16 D2 and D3 sublineages are associated with adenocarcinoma [36, 39]. Some studies suggest that a higher risk for cervical cancer development is not only related to the HPV variant/sublineage, but also dependent on ethnicity. A study that analyzed the risk of CIN3+ development in women infected by HPV16 according to ethnicity shows that white women infected by the HPV16 A1/A2 sublineages have a higher risk of CIN3+ compared to all women from other ethnicities. In contrast, this risk is higher for Asian women with HPV16 A4 infection and Hispanic women with the D2/D3 infection sublineages [36]. Moreover, rare cases of cervical cancer in which the only HPV type isolated was a low-risk HPV have been described [19, 40], suggesting a particular susceptibility in the host.

Several research groups have sought to identify mutations or a mutation profile that could be associated with the transformation activity of an individual isolate. Researchers analyzed the PVs upstream regulatory region (URR), which contains the transcriptional factor binding sites observed that are partly common to all PVs and are partly type specific [41–43]. All these studies led to the conclusion that changes in the replicative behavior and in the ability to generate carcinogenesis by the different variants could be due to point mutations in the URR [44, 45]. However, in-depth analysis of the E2–L2 region, which is responsible for encoding the E5 protein of mucosal HPVs, led to 4 phylogenetically distinct families that could be related to the ability to produce benign or malignant proliferative disorders [46]. These data suggest that a single region of the HPV genome is not sufficient to explain the different infection outcomes.

A study conducted by Chen et al. [16] underscores the importance of sequencing complete HPV genomes in such studies, since mutations are not uniformly distributed through the genome. More recently, with the advent of next-generation sequencing, some groups have started to work with technology that is able to generate 3 to 4 orders of magnitude more information than Sanger sequencing technology [47]. This information density, provided by the deep sequencing approach, is fundamental for the study of viral diversity as it allows the simultaneous detection of minor variants. The resolution power makes these techniques the current standard tool for inter- and intrahost viral diversity analyses.

Studies that analyzed the near full HPV16 genome, using deep sequencing, did not observe any deletion event in the invasive cervical cancer samples analyzed [39, 48]. However, a significant increase in read depth close to L1 and L2, respectively, compatible with a duplication event of a genome fragment spanning these regions was observed [39]. Similar results of partial duplications of the viral genome have been independently reported by different laboratories in lung cancer cases evolving from respiratory papillomatosis caused by HPV low-risk types [49, 50].

Intrahost HPV genetic diversity is poorly studied, but deserves more attention since it could be the main explanation for different outcomes after infections by the same HPV type. Intrahost next-generation sequencing data have shown that HPV16 episomes present in the W12 cell line did not accumulate sequence variation above 0.5% [48]. However, the same authors reported that more than 40% of the clinical samples contained polymorphic sites reaching frequencies up to 5% [48]. In another study, it was detected in polymorphic sites in all ICC samples submitted to deep sequencing [39]. Viral genomic diversity accumulated during the course of a natural infection contrasts sharply to the absence of variation generated during *in vitro* subculture of the W12 episome-containing cells.

It could be claimed that intralesion viral diversity could rather reflect an original infection by several very closely related HPV16 genotypes. However, our current understanding of HPV-associated disease suggests that individual lesions are associated with individual infection events [51, 52]. Instead, we propose here that at least a fraction of such changes accumulated during infection may arise from selection pressures associated with the immune response. Regarding the adaptive immune response, polymorphisms observed in the E6 gene could be a result of an immune-selective pressure. Many linear epitopes have already been described in this region [53]

and could also be related to the affinity of the interactions of the E6 protein with E6-AP [54] and p53 [55]. Similarly, polymorphisms in the E1 gene, present in different HPV16 subvariants, could be related to the DNA binding domain capacity, but do not affect the replication activity [48]. Finally, humoral epitopes are known in the E2 region [53] and could drive immune selection.

Genetic variability of the host may play an important role in HPV infection outcome. Several host genetic studies show that genes in the major histocompatibility complex region are involved in the susceptibility to HPV-associated diseases. HLA-DRB1 and HLA-DQB1 are associated with the susceptibility to cervical carcinoma [56, 57]. A positive association between SCC (squamous cell carcinoma) and human leukocyte antigen has also been described for DRBQ\*15, DRB1\*11, DRB1\*04, and DRB1\*07, and protection against SCC is described especially for DRB1\*13 [58].

## Conclusion

The association of HPV type with cancer risk is well established. However, an important scientific gap remains in our understanding of why the same HPV type leads to such variable clinical outcomes. Variants of HPV16 are related to an increased risk of CIN3+, a finding restricted to the ethnical group where these variants were firstly described, suggesting an evolving ability of the prevailing variants to evade the host's immune control. Besides this, intralesion diversity, a common feature in ICC, may be generated during the course of a chronic infection, possibly fueled by immune selection pressures, either innate, for example in response to an APOBEC3 directional mutation, or adaptive, such as in response to humoral or cellular directed responses.

Mutational patterns that can be responsible for the differential transforming activity within and between HPV types, subtypes, variants and lineages may emerge in the near future, as more whole-genome sequences linked to epidemiological and clinical data becomes available.

## Disclosure Statement

The authors have no conflicts of interest to declare.

## References

- 1 Bravo IG, Felez-Sanchez M: Papillomaviruses: viral evolution, cancer and evolutionary medicine. *Evol Med Public Health* 2015;2015: 32–51.
- 2 Doorbar J, et al: Human papillomavirus molecular biology and disease association. *Rev Med Virol* 2015;(suppl 1):2–23.
- 3 Cubie HA: Diseases associated with human papillomavirus infection. *Virology* 2013;445: 21–34.
- 4 Bernard HU, et al: Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology* 2010;401:70–79.
- 5 Bravo IG, de Sanjose S, Gottschling M: The clinical importance of understanding the evolution of papillomaviruses. *Trends Microbiol* 2010;18:432–438.
- 6 de Villiers EM: Cross-roads in the classification of papillomaviruses. *Virology* 2013;445: 2–10.
- 7 Marin A, et al: Variation in G + C-content and codon choice: differences among synonymous codon groups in vertebrate genes. *Nucleic Acids Res* 1989;17:6181–6189.
- 8 Zhao KN, Liu WJ, Frazer IH: Codon usage bias and A+T content variation in human papillomavirus genomes. *Virus Res* 2003;98: 95–104.
- 9 Cladel NM, Bertotto A, Christensen ND: Human alpha and beta papillomaviruses use different synonymous codon profiles. *Virus Genes* 2010;40:329–340.
- 10 Felez-Sanchez M, et al: Cancer, warts, or asymptomatic infections: clinical presentation matches codon usage preferences in human papillomaviruses. *Genome Biol Evol* 2015;7:2117–2135.
- 11 Vartanian JP, et al: Evidence for editing of human papillomavirus DNA by APOBEC3 in benign and precancerous lesions. *Science* 2008;320:230–233.
- 12 Wang Z, et al: APOBEC3 deaminases induce hypermutation in human papillomavirus 16 DNA upon beta interferon stimulation. *J Virol* 2014;88:1308–1317.
- 13 Munk C, Willemsen A, Bravo IG: An ancient history of gene duplications, fusions and losses in the evolution of APOBEC3 mutators in mammals. *BMC Evol Biol* 2012;12:71.
- 14 Rector A, et al: Ancient papillomavirus-host co-speciation in Felidae. *Genome Biol* 2007; 8:R57.
- 15 Shah SD, Doorbar J, Goldstein RA: Analysis of host-parasite incongruence in papillomavirus evolution using importance sampling. *Mol Biol Evol* 2010;27:1301–1314.
- 16 Chen Z, et al: Evolution and taxonomic classification of human papillomavirus 16 (HPV16)-related variant genomes: HPV31, HPV33, HPV35, HPV52, HPV58 and HPV67. *PLoS One* 2011;6:e20183.
- 17 Antonsson A, et al: The ubiquity and impressive genomic diversity of human skin papillomaviruses suggest a commensalic nature of these viruses. *J Virol* 2000;74:11636–11641.
- 18 Bouvard V, et al: A review of human carcinogens – part B: biological agents. *Lancet Oncol* 2009;10:321–322.
- 19 de Sanjose S, et al: Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study. *Lancet Oncol* 2010;11:1048–1056.
- 20 Li N, et al: Human papillomavirus type distribution in 30,848 invasive cervical cancers worldwide: variation by geographical region, histological type and year of publication. *Int J Cancer* 2011;128:927–935.
- 21 Ho L, et al: Sequence variants of human papillomavirus type 16 in clinical samples permit verification and extension of epidemiological studies and construction of a phylogenetic tree. *J Clin Microbiol* 1991;29:1765–1772.
- 22 Yamada T, et al: 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* 1995;69:7743–7753.
- 23 Arias-Pulido H, et al: Human papillomavirus type 18 variant lineages in United States populations characterized by sequence analysis of LCR-E6, E2, and L1 regions. *Virology* 2005; 338:22–34.
- 24 Calleja-Macias IE, et al: Genomic diversity of human papillomavirus-16, 18, 31, and 35 isolates in a Mexican population and relationship to European, African, and Native American variants. *Virology* 2004;319:315–323.
- 25 Cornet I, et al: Human papillomavirus type 16 genetic variants: phylogeny and classification based on E6 and LCR. *J Virol* 2012;86:6855–6861.
- 26 Chen Z, et al: Diversifying selection in human papillomavirus type 16 lineages based on complete genome analyses. *J Virol* 2005;79: 7014–7023.
- 27 Smith B, et al: Sequence imputation of HPV16 genomes for genetic association studies. *PLoS One* 2011;6:e21375.
- 28 Chen Z, et al: Evolutionary dynamics of variant genomes of human papillomavirus types 18, 45, and 97. *J Virol* 2009;83:1443–1455.
- 29 Chen Z, et al: Evolution and taxonomic classification of *Alphapapillomavirus 7* complete genomes: HPV18, HPV39, HPV45, HPV59, HPV68 and HPV70. *PLoS One* 2013;8:e72565.
- 30 Berumen J, et al: Asian-American variants of human papillomavirus 16 and risk for cervical cancer: a case-control study. *J Natl Cancer Inst* 2001;93:1325–1330.
- 31 Hildesheim A, et al: Human papillomavirus type 16 variants and risk of cervical cancer. *J Natl Cancer Inst* 2001;93:315–318.
- 32 Villa LL, et al: Molecular variants of human papillomavirus types 16 and 18 preferentially associated with cervical neoplasia. *J Gen Virol* 2000;81(pt 12):2959–2968.
- 33 Cornet I, et al: HPV16 genetic variation and the development of cervical cancer worldwide. *Br J Cancer* 2013;108:240–244.
- 34 Xi LF, et al: Human papillomavirus type 16 and 18 variants: race-related distribution and persistence. *J Natl Cancer Inst* 2006;98:1045–1052.
- 35 Sichero L, et al: High grade cervical lesions are caused preferentially by non-European variants of HPVs 16 and 18. *Int J Cancer* 2007; 120:1763–1768.
- 36 Mirabello L, et al: HPV16 sublineage associations with histology-specific cancer risk using HPV whole-genome sequences in 3,200 women. *J Natl Cancer Inst* 2016;108.
- 37 Chan PK, et al: Association of human papillomavirus type 58 variant with the risk of cervical cancer. *J Natl Cancer Inst* 2002;94:1249–1253.
- 38 Khouadri S, et al: Human papillomavirus type 33 polymorphisms and high-grade squamous intraepithelial lesions of the uterine cervix. *J Infect Dis* 2006;194:886–894.
- 39 de Oliveira CM, et al: High-level of viral genomic diversity in cervical cancers: a Brazilian study on human papillomavirus type 16. *Infect Genet Evol* 2015;34:44–51.
- 40 de Oliveira CM, Aquiar LS, Genta ML, Alves VA, Levi JE: HPV-11 associated metastatic cervical cancer. *Gynecol Oncol Case Rep* 2012;2:18–19.
- 41 Villa LL, Schlegel R: Differences in transformation activity between HPV-18 and HPV-16 map to the viral LCR-E6-E7 region. *Virology* 1991;181:374–377.
- 42 O'Connor M, Chan S-Y, Bernard HU: Transcription factor binding sites in the long control region of genital HPVs; in Meyers G, Bernard HU, Delius, et al. (eds): *Human Papillomaviruses, 1995 Compendium*. Los Alamos, Los Alamos National Laboratory, 1995.
- 43 Garcia-Vallve S, et al: Different papillomaviruses have different repertoires of transcription factor binding sites: convergence and divergence in the upstream regulatory region. *BMC Evol Biol* 2006;6:20.
- 44 Rose B, et al: Point mutations in SP1 motifs in the upstream regulatory region of human papillomavirus type 18 isolates from cervical cancers increase promoter activity. *J Gen Virol* 1998;79(pt 7):1659–1663.
- 45 Hubert WG: Variant upstream regulatory region sequences differentially regulate human papillomavirus type 16 DNA replication throughout the viral life cycle. *J Virol* 2005;79: 5914–5922.
- 46 Bravo IG, Alonso A: Mucosal human papillomaviruses encode four different E5 proteins whose chemistry and phylogeny correlate with malignant or benign growth. *J Virol* 2004;78:13613–13626.
- 47 Liu L, et al: Comparison of next-generation sequencing systems. *J Biomed Biotechnol* 2012;2012:251364.

- 48 Kukimoto I, et al: Genetic variation of human papillomavirus type 16 in individual clinical specimens revealed by deep sequencing. *PLoS One* 2013;8:e80583.
- 49 DiLorenzo TP, et al: Human papillomavirus type 6a DNA in the lung carcinoma of a patient with recurrent laryngeal papillomatosis is characterized by a partial duplication. *J Gen Virol* 1992;73(pt 2):423–428.
- 50 Yuan H, et al: Use of reprogrammed cells to identify therapy for respiratory papillomatosis. *N Engl J Med* 2012;367:1220–1227.
- 51 Murray RF, Hobbs J, Payne B: Possible clonal origin of common warts (*Verruca vulgaris*). *Nature* 1971;232:51–52.
- 52 Quint W, et al: One virus, one lesion – individual components of CIN lesions contain a specific HPV type. *J Pathol* 2012;227:62–71.
- 53 Dillner J: Mapping of linear epitopes of human papillomavirus type 16: the E1, E2, E4, E5, E6 and E7 open reading frames. *Int J Cancer* 1990;46:703–711.
- 54 Watts KJ, et al: Sequence variation and physical state of human papillomavirus type 16 cervical cancer isolates from Australia and New Caledonia. *Int J Cancer* 2002;97:868–874.
- 55 Beckman G, et al: Is p53 polymorphism maintained by natural selection? *Hum Hered* 1994;44:266–270.
- 56 Zoodsma M, et al: Analysis of the entire HLA region in susceptibility for cervical cancer: a comprehensive study. *J Med Genet* 2005;42:e49.
- 57 Engemark M, et al: Affected sib-pair analysis of the contribution of HLA class I and class II loci to development of cervical cancer. *Hum Mol Genet* 2004;13:1951–1958.
- 58 de Araujo Souza PS, Sichero L, Maciag PC: HPV variants and HLA polymorphisms: the role of variability on the risk of cervical cancer. *Future Oncol* 2009;5:359–370.