


RESEARCH ARTICLE

Open Access



Genetic variability in E6, E7 and L1 genes of Human Papillomavirus 62 and its prevalence in Mexico

Cristina Artaza-Irigaray^{1,2}, María Guadalupe Flores-Miramontes^{1,2}, Dominik Olszewski³, María Teresa Magaña-Torres¹, María Guadalupe López-Cardona⁴, Yelda Aurora Leal-Herrera⁵, Patricia Piña-Sánchez⁶, Luis Felipe Jave-Suárez^{1*†} and Adriana Aguilar-Lemarroy^{1*†} 

Abstract

Background: Human papillomavirus (HPV) is the main etiological agent of cervical cancer, the third most common cancer among women globally and the second most frequent in Mexico. Persistent infection with high-risk HPV genotypes is associated with premalignant lesions and cervical cancer development. HPVs considered as low risk or not yet classified, are often found in coinfection with different HPV genotypes. Indeed, HPV62 is one of the most prevalent HPV detected in some countries, but there is limited information about its prevalence in other regions and there are no HPV62 variants currently described. The aim of this study was to determine the prevalence of HPV62 in cervical samples from Mexican women and to identify mutations in the L1, E6 and E7 genes, which have never been reported in our population.

Methods: HPV screening was performed by Cobas HPV Test in women who attended prevention health programs and dysplasia clinics. All HPV positive samples ($n = 491$) and 87 additional cervical cancer samples were then genotyped with Linear Array HPV Genotyping test. Some samples were selected to corroborate genotyping by Next-Generation sequencing. On the other hand, nucleotide changes in L1, E6 and E7 genes were determined using PCR, Sanger sequencing and analysis with the CLC-MainWorkbench 7.6.1 software. L1 protein structure was predicted with the I-TASSER server.

Results: Using Linear Array, HPV62 prevalence was 7.6% in general population, 8% in Cervical Intraepithelial Neoplasia grade 1 (CIN1) samples and 4.6% in cervical samples. The presence of HPV62 was confirmed with Next-Generation sequencing. Regarding L1 gene, novel sequence variations were detected, but they did not alter the tertiary structure of the protein. Moreover, several nucleotide substitutions were found in E6 and E7 genes compared to reference HPV62 genomic sequence. Specifically, three non-synonymous sequence variations were detected, two in E6 and one in E7.

Conclusions: HPV62 is a frequent HPV genotype found mainly in general population and in women with CIN1, and in 90.5% of the cases it was found in coinfection with other HPVs. Novel nucleotide changes in its *L1*, *E6* and *E7* genes were detected, some of them lead to changes in the protein sequence.

Keywords: Cervical cancer, HPV62, E6, E7, L1

* Correspondence: lfjave@yahoo.com; adry.aguilar.lemarroy@gmail.com

†Equal contributors

¹División de Inmunología, Centro de Investigación Biomédica de Occidente (CIBO), Instituto Mexicano del Seguro Social (IMSS), Guadalajara, Jalisco, Mexico

Full list of author information is available at the end of the article



Background

Cervical Cancer (CC) is the fourth leading cause of cancer deaths in women worldwide with an estimate of more than 528,000 new diagnosed cases and 266,000 deaths in 2012. More than 85% of deaths occur in developing countries and in particular, in Mexico, CC is the second most frequent cancer leading to 4,769 deaths in 2012 [1]. This pathology is directly associated with HPV (Human Papillomavirus) infection and to date, around 200 HPVs have been described [2, 3]. Mucosal HPVs are grouped into low risk (LR-HPVs) and high risk HPVs (HR-HPVs), the latter being considered the etiologic agents of CC [3]. The *Alphapapillomavirus* genus harbors more than 60 types of HPVs, including the oncogenic HPVs associated to anogenital cancers according to the International Agency for Research on Cancer (IARC): types 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66 and 68 [4, 5]. Specifically, HPV16 and HPV18 are found in around 70% of the CC cases worldwide [6, 7].

HPVs have a circular double-stranded 8 kb DNA genome that typically contains eight genes [8]. The L1 gene, which encodes the principal virus capsid protein, is used for the classification and construction of phylogenetic trees as it is well conserved among different HPVs. In 2004, the HPV classification criteria were defined based on differences in the complete L1 Open Reading Frame (ORF): different genera share less than 60% nucleotide sequence identity, HPV species within a genus share between 60 and 70% identity, HPV types share between 71 and 89% nucleotide identity, HPV subtypes differ in 2-10% and HPV variants differ in 1-2% within the L1 ORF [9]. In 2013, the term variant was proposed to also include HPV subtypes. The use of full genome sequence information, instead of the L1 ORF, was recommended to classify a new variant genome. The alignment of complete viral genomes began to define variant lineages and sublineages using differences of 1-10% and 0.5-1%, respectively [10].

HR-HPV genomes encode three oncoproteins —E5, E6 and E7— that contribute to enhanced cell proliferation, initiation and progression of CC [11]. An interesting review contrasts the activities of the human alpha-HPV oncoproteins with their non-oncogenic counterparts based on cell culture studies [12]. The comparison of activities of LR- and HR-HPVs would lead to the identification of common activities probably needed for the viral life cycle, while additional functions of HR-HPVs could be crucial for the transformation/immortalization process.

HPV62 was characterized in 2004 by Fu et al. (accession number AY395706) from a cervical sample obtained from a 45-year-old woman with normal cytology [13] and it is considered as a LR-HPV. The E6 (447 bp), E7 (291 bp) and L1 (1512 bp) genes from this HPV62

reference genome (8092 bp) encode for 148, 96 and 503 amino acid proteins, respectively. In a first report including cervical samples from Mexican population, HPV62 was mainly detected in coinfection with other HPV genotypes and it was found in 5.1% of HPV positive patients with Cervical Intraepithelial Neoplasia Grade 1 (CIN1) and in 0.8% of HPV positive patients with CC using Linear Array HPV Genotyping test [14]. The aim of this study, was to determine the prevalence of this genotype in a greater number of samples among Mexican women and to detect possible mutations in L1, E6 and E7 genes of the HPV62 circulating in the Mexican population. Until now, only two complete genome sequences have been reported worldwide from 2 patients: the first one in 2004 (AY395706) [13] and the second one very recently uploaded (KU298924.1) [15].

Methods

Sample collection

All samples were collected by gynecologists with a cytobrush inserted into the endocervical canal and placed into the transport medium PreservCyt (Hologic, Bedford, MA). Three large groups of samples were included: 1) cervical samples from women (general population), who attended cervical cancer prevention programs; 2) cervical samples with CIN1 from women who attended a Dysplasia Clinic, and 3) cervical samples with CC from women who attended the Oncology Hospital. The first group of samples include women from six different States of the Mexican Republic (Aguascalientes, Colima, Guanajuato, Jalisco, Michoacán, Nayarit and Yucatán), and they were obtained from the Regional Hospital Dr. Valentín Gómez Farías – ISSSTE (Guadalajara, Jalisco). The samples from the second group were recruited at the Dysplasia Clinic of the Regional General Hospital No. 12 Lic. Benito Juárez – IMSS (Mérida, Yucatán), and at the Dysplasia Clinic of the Western National Medical Center – IMSS (Guadalajara, Jalisco). Finally, the last group's samples were obtained exclusively from the Oncology Hospital of the Western National Medical Center – IMSS (Guadalajara, Jalisco). In all cases, the samples were taken as part of the routine diagnosis confirmation and an aliquot of the PreservCyt solution was given for our study, after informed consent was signed. Samples with excess of blood and mucus in the PreservCyt solution, low DNA quantity or quality samples, were eliminated. A total of 2835 samples were screened for HPV; 2399 from general population, 349 from Dysplasia Clinics and 87 from Oncology Hospital. Concerning only the first group of samples, the diagnosis was kept anonymous for this study, as authorized in the ethically approved protocol. However, the diagnosis of the second (CIN1) and third (CC) group of samples was obtained

by colposcopic observation and confirmed by histopathological analysis.

Samples were collected from July 2014 to July 2016. As detailed in the section of ethical considerations, collection of the samples from the different groups was authorized by the ISSSTE and the National Committee on Health Research and Ethics of the IMSS, for various research protocols.

HPV screening and genotyping

The samples from the first and the second group were first screened for HPV positivity with the Cobas HPV Test (Roche Molecular Systems, Inc). Afterwards, all HPV positive samples were genotyped with the Linear Array HPV Genotyping test (LA), Roche Molecular Diagnostics. Additionally, 48 CIN1 samples taken randomly, were genotyped with the 454 Next-Generation Sequencing (NGS) platform to confirm HPV genotyping. Regarding the third group, all CC samples were genotyped with LA and 48 of them were also selected randomly to confirm genotyping with NGS. The set of PGM11/09 primers (452 bp amplicon from nucleotide 949 to 1400 of L1 ORF) was used for NGS, as previously described [16]. In addition, NGS was also performed using the degenerated FAP primers set [17, 18]. A first conventional Polymerase Chain Reaction (PCR) was done with the primers FAP59 (forward 5'-TAACWG-TIGGICAYCCWTATT - 3') and FAP64 (reverse 5'-CCWATATCWVHCATITCICCATC - 3') to amplify a 478 bp fragment from the L1 region. PCR conditions were: 94 °C for 3 min, 35 cycles at 94 °C for 45 s, 50 °C for 30 s and 72 °C for 1 min, finally 72 °C for 10 min. A second PCR was done with the 478 bp amplicon using the following primers: Forward 5'-[Universal Multiplicom tail A: AAGACTCGGCAG-CATCTCCA] - [FAP6085: CCWGATCCHAATMRRT TTGC]-3'. and Reverse 5' [Universal Multiplicom tail B: GCGATCGTCACTGTTCTCCA] - [FAP64 primer]. The resulting amplicon is of 377 bp (from nucleotide 232 to 607 of L1 ORF). Universal Multiplicom tails A and B were the same used in Cat. no. MR-0020.024 (Multiplicom NV CFTR; Molecular Diagnostics, Niel, Belgium). This second PCR was run under the same conditions as the first PCR, except for the annealing temperature that was changed to 47 °C. Amplicons were visualized by gel electrophoresis in a 1.5% agarose gel. HPV positive samples were selected to undergo further screening by NGS using Multiplex identifiers (MID) barcodes for each sample as described in Flores-Miramontes et al. [16]. Quality control of the obtained sequences was carried out with the online platform Galaxy (version 16.04) and they were analyzed with Roche's GS Reference Mapper Software (version 2.9), using as references all

human papillomavirus sequences from the Papillomavirus Episteme (PaVE) database [3, 19].

Sanger sequencing for L1, E6 and E7 genes

E6, E7 and L1 genes were amplified from cervical samples positive to HPV62 by Linear Array using PCR with the following primer pairs specific for each gene of interest (all of them were designed with Oligo v6 software). E6: forward 5'-GGTCAGCACAGTAGCAATGACT-3' and reverse 5'-CGGGACGCTCTTGTAGGAC-3'; E7: forward 5'-CAGGAGTGTGGACAGGACGGTA-3' and reverse 5'-GCATCGGCCATGTCACTTATG-3'; L1: forward 5'-ACGCCTTCCTCCCTGCAACTA-3' and reverse 5'-CACTGACAAACGCGCACAACAC-3'. The reactions were performed in a final volume of 25 µL containing at least 200 ng of genomic DNA, 200 µM of each dNTP, 1X reaction buffer with 1.8 mM of MgCl₂, 12.5 pmol of each primer, and 1.25 units of Fast Start High Fidelity enzyme (Roche Applied Science, Cat. No. 04 738 284 001). PCR conditions were: initial denaturation at 95 °C for 2 min, 35 cycles of 95 °C for 30 s, annealing at 58 °C (*HPV62-E6* and *-E7*) or 62 °C (*HPV62-L1*) for 30 s, elongation at 72 °C for 45 s, and a final extension at 72 °C for 7 min. Afterwards, 5 µL of the PCR products were visualized on 1% agarose gel to corroborate the presence and size of the amplicon, and the other 20 µL were utilized for isopropanol purification. Purified amplicons were sequenced with BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Cat. No. 4337455) using the above mentioned forward primers for E6, E7, and L1 and additional primers designed for L1, which is too long to be sequenced with a single primer (5'-ACACGGAACGCATGGTATGGGC, 5'-GCAGAACCCTTATGGCGATTGTA-3' and 5'-TTGTGAAAATACAGTTAACCC-3'). Reactions were performed in a final volume of 20 µL with around 50 ng of DNA, 10 pmol of forward primer, 2 µL of 5X Sequencing Buffer, and 4 µL of Ready Reaction Premix. Cycling conditions were set as follows: initial denaturation at 96 °C for 1 min, and 25 cycles at 96 °C for 10 s, annealing at 50 °C for 5 s, elongation at 60 °C for 4 min. Finally, products were purified with Centri-Sep Spin Columns (ABI, Cat. No. 401762), and sequenced with the ABI PRISM 310 Genetic Analyzer (Applied Biosystems).

Sequence alignments and L1 protein structure prediction analysis

To detect genetic mutations or variants in E6, E7 and L1, the obtained gene sequences were aligned to the HPV62 reference sequence (reported as AY395706 in the NCBI database) using the *CLC-MainWorkbench 7.6.1* program (Qiagen). To look for amino acid changes, the DNA sequences were translated into protein with

the same software, and aligned to the reference protein. The phylogenetic tree showing alpha-3 group HPVs and all HPV62 sequences obtained in this work, was performed with MEGA v.7.014 software using Maximum Likelihood statistical method.

Finally, the L1 protein structures were predicted with the I-TASSER server (Iterative Threading ASSEMBly Refinement), which identifies structural templates from the Protein Data Bank by multiple threading approach and constructs full-length atomic models by iterative template fragment assembly simulations [20–23]. Protein structure alignment was performed with *CLC MainWorkbench 7.6.1* program.

Results

HPV62 prevalence genotyped by Linear Array HPV Genotyping Test

Regarding the first group of cervical samples described in “Methods section” (general population of women who attended cervical cancer prevention health programs), from the 2399 samples screened for HPV positivity by Cobas HPV Test, 291 were HPV positive (12.1%) and 22/291 (7.6%) were HPV62 positive. Concerning the second group (CIN1 samples), from the 349 screened with Cobas, 200 (57.3%) were HPV positive and 16/200 (8%) were HPV62 positive. Finally, the third group of samples (CC samples) showed 100% of HPV positivity with LA, and from those, only 4 samples (4.6%) were HPV62 positive. It is important to highlight that in only 4 samples from the total of 42 HPV62 positive samples, this genotype was found as single infection (exclusively in CIN1 samples), and in the other 38 (90.5%) in coinfection with 1 to 5 additional HPV genotypes (detailed in Table 1).

Interestingly, as depicted in Fig. 1, HPV62 was more frequently found in coinfection with 16, 39, 59, 51 and 83 HPV genotypes.

HPV genotyping by NGS

To confirm HPV genotyping, 48 CIN1 samples were sequenced with NGS, 8 of them HPV62 positive with Linear Array. In those 8 samples, HPV62 presence was corroborated using NGS; however, 3 additional samples were positive to HPV62 only with the last methodology.

Concerning the presence of HPV62 exclusively in cervical cancer, 48 samples diagnosed with squamous cervical carcinoma were selected for HPV genotyping by NGS (choosing preferentially those that showed more than two HPV genotypes detected by Linear Array). To detect a broader HPV genotype spectrum in those samples, NGS was performed utilizing PGMY11/09 and FAP primer sets. As depicted in Table 2, four of the samples mapped with high identity to HPV62, and they were found in coinfection with additional HPV genotypes,

Table 1 HPV genotypes detected in coinfection with HPV62 by Linear Array in each of the 38 HPV62 positive samples

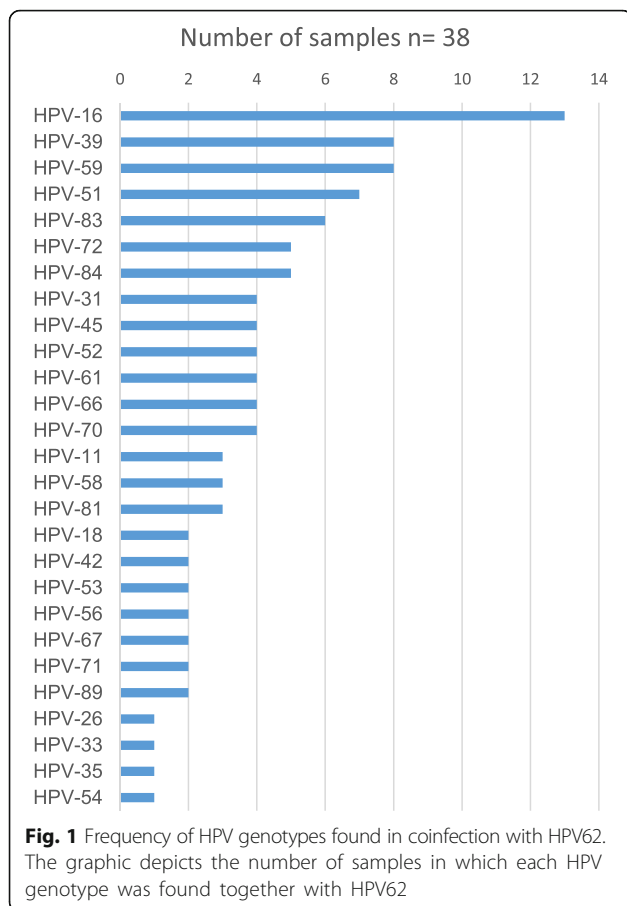
GROUP 1 SAMPLES (general population)			
Sample Code	HPV genotypes detected by Linear Array	Sample Code	HPV genotypes detected by Linear Array
GP-1	16, 42, 58 , 81, 83	GP-12	39, 59 , 61
GP-2	31, 52 , 61, 72, 84	GP-13	42, 45 , 84
GP-3	39, 51, 56, 58 , 84	GP-14	16, 18
GP-4	11, 52, 59 , 83	GP-15	16, 51
GP-5	31, 39 , 66, 67	GP-16	16, 52
GP-6	39 , 61, 84, 89	GP-17	31, 59
GP-7	52, 56 , 70, 72	GP-18	35 , 67
GP-8	11, 16, 39	GP-19	45 , 72
GP-9	16, 51, 59	GP-20	51, 59
GP-10	26, 45 , 66	GP-21	58 , 66
GP-11	31 , 72, 83	GP-22	51
GROUP 2 SAMPLES (CIN1)			
Sample Code	HPV genotypes detected by Linear Array	Sample Code	HPV genotypes detected by Linear Array
CIN1-1	33, 45, 59 , 71, 83	CIN1-7	39 , 81
CIN1-2	16, 51, 59 , 81	CIN1-8	51 , 53
CIN1-3	11, 70, 83	CIN1-9	70, 83
CIN1-4	16, 39 , 66	CIN1-10	16
CIN1-5	61, 72, 84	CIN1-11	59
CIN1-6	16 , 53	CIN1-12	89
GROUP 3 SAMPLES (CC)			
Sample Code	HPV genotypes detected by Linear Array	Sample Code	HPV genotypes detected by Linear Array
CC-1	16 , 54, 70	CC-3	39 , 71
CC-2	16, 18	CC-4	16

In bold: HPV genotypes classified as carcinogenic to humans by the IARC (group 1); in italics: possibly carcinogenic to humans (group 2B)

including beta-1 papillomavirus (12, 21 and 118), which were only detected with FAP primers.

Nucleotide variations in HPV62-L1

To study whether the HPV62 circulating in the Mexican population exhibits nucleotide changes in L1, all HPV62-L1 sequences obtained by NGS with PGMY or FAP primers (independently of the diagnosis of the samples) were aligned to the HPV62 complete genome (AY395706). The contigs were obtained from 11 samples amplified with PGMY primers and 9 samples amplified with FAP primers. Alignment of the eleven HPV62-L1 contigs revealed the presence of 9 mutations distributed among all the samples in the region amplified with



PGMY primers. As depicted in Table 3, three out of the nine mutations were non-synonymous. Moreover, nine HPV62-L1 contigs from the region amplified with FAP primers revealed 10 mutations, five of them non-synonymous (Table 3).

To study whether the HPV62 that infects the Mexican population is a variant of the virus, the whole conserved HPV62-L1 ORF (1690 bp amplicon) from genomic DNA of a cervical sample was amplified and sequenced (HPV62-L1 P9). A total of 17 nucleotide changes were detected in HPV62-L1: four non-synonymous c.250A > G (p.T84A), c.263C > G (p.A88G), c.1139A > G (p.K380R) and c.1489G > A (p.A497T); and 13 synonymous c.111G > A, c.117C > T, c.126 T > C, c.324C > T, c.627C > T, c.663G > A, c.1014 T > C, c.1017G > A, c.1161C > T, c.1416A > G, c.1425 T > A, c.1443 T > C, c.1497A > C (Fig. 2a). The first 3 amino acid substitutions are in the same spatial region in L1, while the last substitution is in the C-terminal domain, as shown by protein structure prediction (Fig. 2b, c, d). None of the four amino acid changes affect the tertiary structure of the L1 protein according to the predicted models.

Nucleotide variations in HPV62-E6 and HPV62-E7

The E6 and E7 genes amplified from genomic DNA of 13 cervical samples infected with HPV62 were purified and sequenced to detect nucleotide variations in the HPV62 that circulates in Mexico. Altogether, seven nucleotide changes were identified in HPV62-E6, two of which were non-synonymous and led to an alteration in

Table 2 HPV genotypes detected in samples from cervical cancer that were positive to HPV62 by NGS

Sample code	HPV types found by Linear Array	HPV types found by NGS	Reads using FAP primers	Reads using PGMY primers	Total Reads
CC-1	16, 54, 62, 70	16	-	174	174
		33	-	1	1
		54	776	3	779
		62	105	-	105
		70	163	348	511
		118	2	-	2
CC-2	16, 18, 62	16	9	481	490
		62	895	3	898
CC-3	39, 62, 71	16	-	50	50
		39	-	2	2
		62	1086	218	1304
		71	-	98	98
CC-5	33	81	2	-	2
		12	21	-	21
		21	1069	-	1069
		33	1	507	508
		62	3	-	3

Reads obtained from each HPV type by using FAP or PGMY11/09 primers are detailed for each sample. Linear Array results are also included

Table 3 Genetic mutations found in the 5' and 3'-ends of HPV62-L1

Nucleotide changes (PGMY primers)	Amino acid change	Number of samples with the mutation <i>n</i> = 11	Nucleotide changes (FAP primers)	Amino acid change	Number of samples with the mutation <i>n</i> = 9
c.969 T > C		2	c.250A > G	p.T84A	9
c.987A > C *	p.E329D	3	c.263C > G	p.A88G	7
c.1017G > A		1	c.265A > T	p.T89S	1
c.1071 T > C *		3	c.413C > T	p.A138V	1
c.1104G > A		1	c.436A > G	p.I146V	1
c.1236 T > C		3	c.438C > T		1
c.1256A > G	p.H419R	4	c.468A > G		1
c.1279A > G *	p.T427A	1	c.495A > T		5
c.1287A > G		3	c.495A > G		1
			c.516C > A		1

Nucleotide changes are shown in samples amplified with PGMY or FAP primers. The amino acid changes are described for those with non-synonymous mutations. The number of samples that carry each nucleotide change are included. (*) Already reported mutations

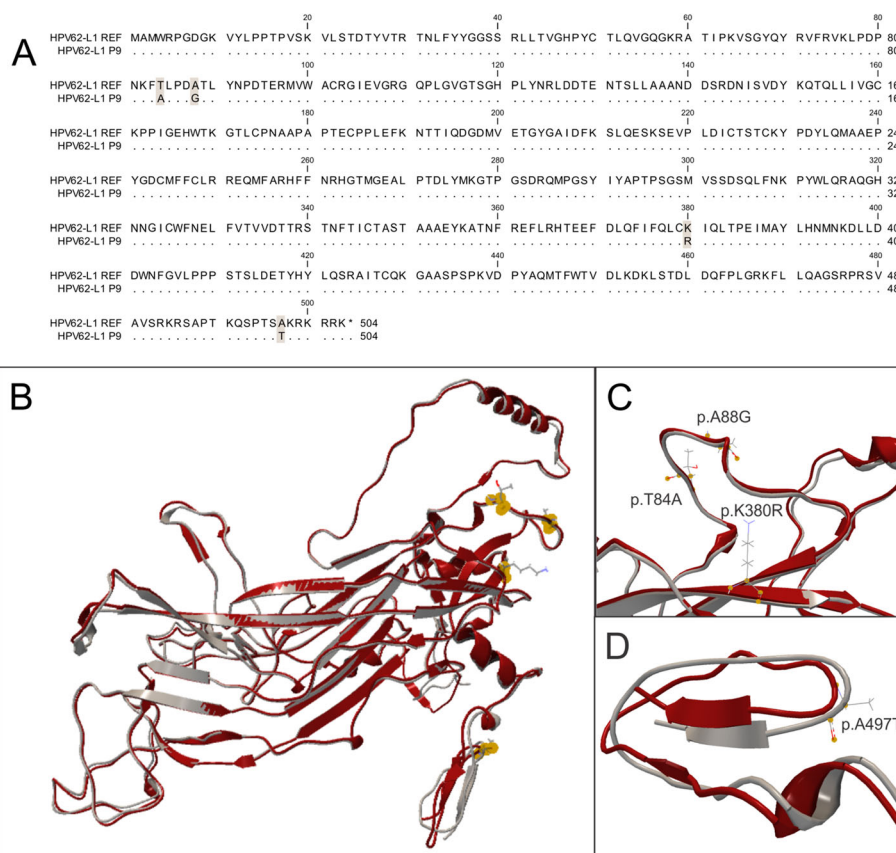


Fig. 2 Sequence and structure alignment of HPV62 L1 protein. **a** Protein sequence alignment of the AY395706 NCBI sequence (HPV62-L1 REF) with that obtained from cervical sample P9 (HPV62-L1 P9); dots indicate matching residues, amino acid changes are darkened. **b** Structure alignment of both 503 amino acid complete proteins (HPV62-L1 REF in red and HPV62-L1 P9 in grey) and location of the four detected mutations shown as yellow dots. **c** Magnification of the structure alignment region containing the mutated amino acids threonine (p.T84A), alanine (p.A88G), lysine (p.K380R); and **d**) alanine (p. A497T)

the protein sequence (c.157C > T, p.R53W and c.404A > G, p.Y135C). Regarding HPV62-E7, three nucleotide changes were detected and one of them affected the protein sequence (c.125C > G, p.A42G) (Table 4). The non-synonymous changes were present in all the samples, while the other nucleotide variations that do not change the amino acid were found only in some of them. The genetic sequences of E6 and E7 from HPV62 amplified from the 13 samples were translated into protein and aligned to the reference sequence for an easier visualization of the nucleotide changes and their location (Fig. 3). Finally, to visualize the relationship between all HPVs from Alpha-3 species and the HPV62 sequences obtained from the 13 cervical samples, a phylogenetic tree was built based on E6/E7 gene sequences (Fig. 4).

Discussion

A fragment of the L1 ORF —MY09/11 region— of HPV62 was first sequenced in 1994 and the complete genome was isolated and sequenced to characterize the novel HPV type 62 in 2004 [13, 24]. In 2006, the Linear Array HPV Genotyping Test was launched; an assay that identifies up to 37 HPV genotypes in cervical samples, including HPV62. HPV62 belongs to the *Alphapapillomavirus* genus, species 3 group (α-3 group) together with HPV 61, 72, 81, 83, 84, 86, 87, 89, 102 and 114.

In the present study, from 578 HPV positive samples obtained from general population, CIN1 and CC women, the HPV62 frequency found with LA was 7.6, 8, and 4.6%, respectively, showing that this genotype has an important prevalence in Mexico; and it is mainly found in coinfection with HR-HPV genotypes (Table 1 and Fig. 1). Importantly, by using NGS, 11 out of 48 individually sequenced CIN1 samples were also positive for HPV62, despite only 8 of them were positive for this HPV

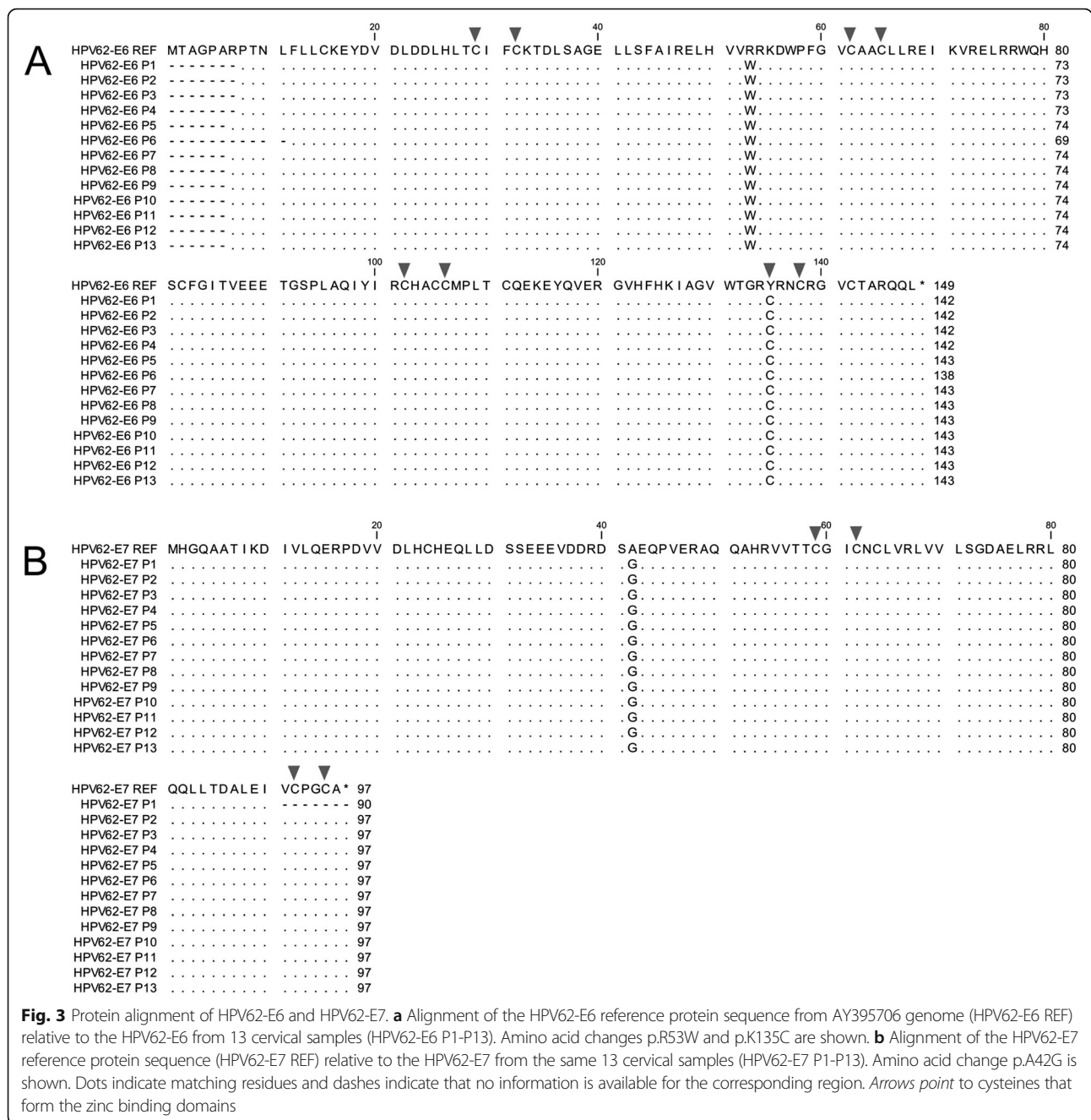
genotype by Linear Array; moreover, concerning CC samples, only 1 of the 4 HPV62 positive samples detected by NGS, was positive with the last methodology. Therefore, the frequency of HPV62 could be underestimated. It is worth mentioning that, because cervical cancer cells were collected from cervical swabs and not from biopsies, the samples may contain normal cells together with cancer cells, so this could be a limitation of this study, since it is not possible to know which cells each HPV is infecting.

When HPV62 was characterized, it was one of the ten most prevalent HPV types detected in women with normal Pap smears from Costa Rica and New Mexico [13]. Later, more studies supported this finding. In the Italian population, HPV62 was detected in 1.5% of high-grade squamous intraepithelial lesions and also in 1.5% of CC, being the only LR-HPV found in CC samples [25]. There was no HPV62 detection in samples without cervical lesions in both populations. Among LR-HPV types detected in the Northern Indian population, HPV62 was the most common (10.5% of HPV positive samples) and in Egypt HPV62 was also the most prevalent LR-HPV among HPV positive women (17.4 and 9.7% in two different research works) [26–28]. In Thai women, HPV62 is also among the most frequent LR-HPV (11.3% of HPV positive samples) [29]. Another interesting result was found in females from the United States where the most common HPV type was HPV62 (found in 6.5% of all the subjects, where 42.5% are HPV positive) [30]. In Croatian women, HPV62 was among the most prevalent LR-HPVs in women with abnormal cervical cytology (23.3%) [31]. In Northeast Brazil, HPV62 prevalence among the overall population was of 3.6% and in Korea, 2% of atypical squamous cell and low-grade squamous intraepithelial lesions were HPV62 positive [32, 33]. In Mexico, a IMSS Research Network

Table 4 Nucleotide changes found in *HPV62-E6* and *HPV62-E7* gene sequences from 13 cervical samples infected with HPV62 (P1-P13) compared to the reported HPV62 genome (AY395706, NCBI)

Mutations	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13
<i>HPV62-E6</i>													
c.27G > A	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	
c.37 T > C	✓	✓		✓	✓	✓		✓	✓	✓			
c.157C > T*	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
c.177 T > C							✓				✓		
c.199 T > C	✓			✓		✓			✓	✓			
c.201G > C	✓			✓		✓			✓	✓			
c.404A > G*	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
<i>HPV62-E7</i>													
c.125C > G*	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
c.183 T > C							✓						
c.199A > C	✓			✓		✓			✓	✓			

(*) Nucleotide substitutions that alter protein sequence

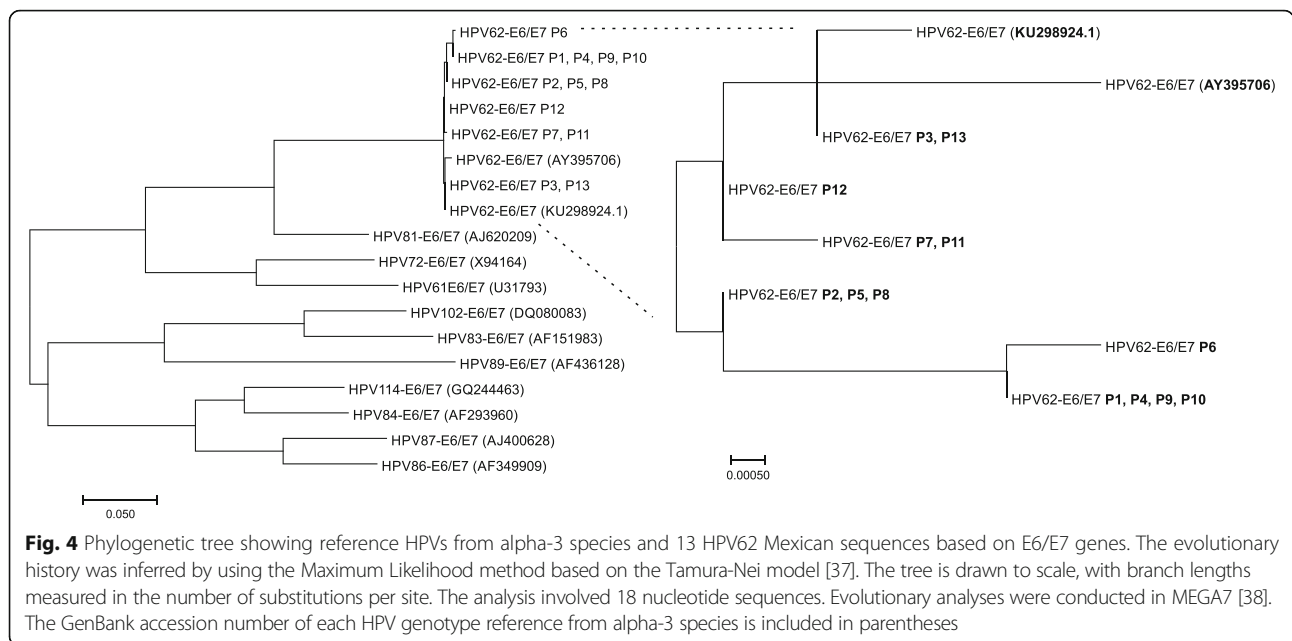


report on HPV, including 822 women, found HPV62 infection in 3.1% of women without cervical lesions, in 5.1% with CIN1, in 6.7% with CIN3 and in 0.8% in CC samples [14].

All these studies agree in the high frequency of HPV62 in cervical samples and its omnipresence in all kind of diagnosed samples; it is therefore of great interest to further study this genotype.

This work describes novel nucleotide changes in the HPV62-L1 complete gene (1512 bp). The 17 variations identified in the L1 ORF amplified from one sample

determine a difference of 1.12% (17/1512 bp) compared to the reference sequence. A nucleotide sequence difference of 1% or more would define a new variant lineage, but according to the recommended new HPV variant classification criteria, the complete genome (and not only the L1 ORF) has to be sequenced [10]. The alignment of 20 HPV62-L1 sequences obtained by NGS using PGM Y or FAP primer pairs from 16 cervical samples revealed the presence of a total of 19 mutations, 8 of them being non-synonymous. Interestingly, to our knowledge, only 3 out of the 19 mutations have been



previously described [34]. It is worth mentioning that concerning those nucleotide changes, there was not a different distribution between CIN1 and CC samples, the mutations were distributed randomly among them.

Concerning HPV62-E6 and -E7 ORFs, the presence of genetic variations is described for the first time in the present study. Specifically, concerning HPV62-E6, non-synonymous nucleotide changes c.157C > T (p.R53W) and c.404A > G (p.Y135C) were found; additionally, c.125C > G (p.A42G) substitution in HPV62-E7 was found in all the 13 samples under study. However, the c.404A > G (p.Y135C) substitution that leads to a change of a tyrosine by a cysteine in all the 13 analyzed samples is located in a key position essential for the formation of a zinc binding domain which needs two CxxC motifs. Indeed, eight cysteines involved in the formation of two zinc binding sites in E6 protein are conserved among the different HPV types and the seventh cysteine in HPV62 is located in position 135, where a tyrosine was reported instead [12, 35]. HPV62 was originally characterized with the overlapping PCR method and according to the described methodological process, PCR products were visualized with UV illumination in agarose gels before product purification, ligation into pGEM-Teasy vector and sequencing. UV radiation is a strong mutagen that can induce conversion from one base to another [36]. Therefore, the adenine reported in position 404 of the reference genome could probably be due to a spontaneous change caused by UV radiation. The reference AY395706 HPV62 genomic sequence could have been reported with some mistakes and the three non-synonymous changes in E6 and E7 described in

this work might be found of the HPV62 sequence worldwide. Earlier this year, a new complete HPV62 genome sequence has been uploaded in the GenBank database (KU298924.1) [15]; in which some of the nucleotide changes described in the present research are confirmed.

These findings open a new door to the study of LR-HPVs commonly found in HPV positive women. To date, oncogenic HPVs have been studied primarily, thus information on rare or LR-HPVs is limited, although they are often more prevalent than HR-HPVs. When found in coinfection with other HPV types, these non-oncogenic viruses may play an important role in the progression or regression of a cervical lesion, but this remains unstudied. Undoubtedly, more research is needed in this field, particularly with respect to the E5, E6 and E7 proteins and their different domains shared between HR-HPVs and LR-HPVs to understand their way of action. Furthermore, the L1 major capsid protein and its different variants may influence crucial steps of the viral infection cycle due to altered affinity to other proteins.

HPV infection in a single patient often comes along with more than one HPV genotype. Hence, clustering of the different genotypes present in a single sample might be crucial for understanding the roles of each genotype in carcinogenesis. This will lead to a better understanding of possible interactions between HPVs found in coinfections, both low and high risk.

Conclusions

HPV62 was found in Mexican women who attended their preventive routine check-up and in women with

cervical cancer. In the general population and in the CIN1 samples, HPV62 was often present both in single and multiple infection; however, in cervical cancer samples it was only found in coinfections with at least one HR-HPV type. To our knowledge, this is the first study that describes the presence of mutations in HPV62-E6 and -E7 genes. Moreover, newly observed nucleotide changes in the L1 gene were found to alter the L1 protein sequence. Upcoming discoveries in this field will complement the current information on variants of human papillomavirus and on still unclassified genotypes in their carcinogenicity risk to humans.

Abbreviations

CC: Cervical Cancer; CIN1: Cervical Intraepithelial Neoplasia Grade 1; HR-HPV: High Risk Human Papillomavirus; LR-HPV: Low-Risk Human Papillomavirus; NGS: Next-Generation Sequencing; ORF: Open Reading Frame

Acknowledgements

CA-I and MGF-M are grateful for a scholarship from Consejo Nacional de Ciencia y Tecnología (CONACyT)- Mexico. PP-S and MTM-T are scholarship holders of the IMSS Foundation and are grateful for their support.

Funding

This work was supported by Fondo de Investigación en Salud – IMSS, grants numbers FIS/IMSS/PROT/PRI0/14/033 to AA-L and FIS/IMSS/PROT/PRI0/15/046 to LFJ-S.

Availability of data and materials

The datasets generated and/or analyzed during the current study are not publicly available, but are available from the corresponding author on reasonable request.

Authors' contributions

MGL-C, YAL-H and PP-S were involved in patient interviews and sample recruitment. MGL-C and PP-S performed the Cobas HPV Test, MGF-M and DO conducted the Linear Array HPV test and NGS. CA-I carried out PCR amplifications, Sanger sequencing, the bioinformatics analyses and draft the manuscript; MTM-T gave support with the Sanger sequencing, AA-L and LFJ-S conceived of and designed the study, supervised all experiments and analyses, and wrote the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Consent for publication

Not applicable.

Ethics approval and consent to participate

Before the samples were taken, participants authorized and signed informed consent to participate in the study. Ethical approval of this study was obtained from the ISSSTE Research Committee (registration number ISSSTE/CEI/2013/080) and the National Committee on Health Research and Ethics of the IMSS (with the registration numbers R-2012-785-090, R-2014-785-036 and R-2016-785-028).

Author details

¹División de Inmunología, Centro de Investigación Biomédica de Occidente (CIBO), Instituto Mexicano del Seguro Social (IMSS), Guadalajara, Jalisco, Mexico. ²Programa de Doctorado en Ciencias Biomédicas, Centro Universitario de Ciencias de la Salud (CUCS), Universidad de Guadalajara, Jalisco, Mexico. ³Institute of Pharmacy and Molecular Biotechnology, University of Heidelberg, Heidelberg, Germany. ⁴Unidad de Medicina Genómica y Genética, Hospital Regional Dr. Valentín Gómez Farías - ISSSTE, Guadalajara, Jalisco, Mexico. ⁵Unidad de Investigación Médica Yucatán (UIMY) - IMSS, Mérida, Yucatán, Mexico. ⁶Laboratorio de Oncología Molecular, Unidad de Investigación Médica en Enfermedades Oncológicas (UIMEO) - IMSS, Ciudad de México, Mexico.

Received: 25 November 2016 Accepted: 20 February 2017

Published online: 04 March 2017

References

- GLOBOCAN. Estimated Cancer Incidence, Mortality and Prevalence Worldwide in 2012. http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx. Accessed 6 June 2016.
- Zur Hausen H. Papillomaviruses and cancer: from basic studies to clinical application. *Nat Rev Cancer*. 2002;2(5):342–50.
- Pedraza-Brindis EJ, Sanchez-Reyes K, Hernandez-Flores G, Bravo-Cuellar A, Jave-Suarez LF, Aguilar-Lemarroy A, Gomez-Lomeli P, Lopez-Lopez BA, Ortiz-Lazareno PC. Culture supernatants of cervical cancer cells induce an M2 phenotypic profile in THP-1 macrophages. *Cell Immunol*. 2016;310:42–52.
- Ma Y, Madupu R, Karaoz U, Nossa CW, Yang L, Yooseph S, Yachinski PS, Brodie EL, Nelson KE, Pei Z. Human papillomavirus community in healthy persons, defined by metagenomics analysis of human microbiome project shotgun sequencing data sets. *J Virol*. 2014;88(9):4786–97.
- Bravo IG, Féliz-Sánchez M. Papillomaviruses viral evolution, cancer and evolutionary medicine. *Evol med public health*. 2015;2015(1):32–51.
- Comprehensive Cervical Cancer Prevention and Control. WHO. 2013.
- de Sanjose S, Quint WG, Alemany L, Geraets DT, Klaustermeier JE, Lloveras B, Tous S, Felix A, Bravo LE, Shin H-R. Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study. *Lancet Oncol*. 2010;11(11):1048–56.
- Munoz N, Castellsagué X, de González AB, Gissmann L. HPV in the etiology of human cancer. *Vaccine*. 2006;24:S1–S10.
- De Villiers E-M, Fauquet C, Broker TR, Bernard H-U, zur Hausen H. Classification of papillomaviruses. *Virology*. 2004;324(1):17–27.
- Burk RD, Harari A, Chen Z. Human papillomavirus genome variants. *Virology*. 2013;445(1):232–43.
- Moody CA, Laimins LA. Human papillomavirus oncoproteins: pathways to transformation. *Nat Rev Cancer*. 2010;10(8):550–60.
- Klingelutz AJ, Roman A. Cellular transformation by human papillomaviruses: lessons learned by comparing high-and low-risk viruses. *Virology*. 2012;424(2):77–98.
- Fu L, Terai M, Matsukura T, Herrero R, Burk RD. Codetection of a mixed population of candHPV62 containing wild-type and disrupted E1 open-reading frame in a 45-year-old woman with normal cytology. *J Infect Dis*. 2004;190(7):1303–9.
- Aguilar-Lemarroy A, Vallejo-Ruiz V, Cortes-Gutierrez EI, Salgado-Bernabe ME, Ramos-Gonzalez NP, Ortega-Cervantes L, Arias-Flores R, Medina-Diaz IM, Hernandez-Garza F, Santos-Lopez G, et al. Human papillomavirus infections in Mexican women with normal cytology, precancerous lesions, and cervical cancer: type-specific prevalence and HPV coinfections. *J Med Virol*. 2015;87(5):871–84.
- Siqueira JD, Alves BM, Prellwitz IM, Furtado C, Meyrelles AR, Machado ES, Seanez HN, Soares MA, Soares EA. Identification of novel human papillomavirus lineages and sublineages in HIV/HPV-coinfected pregnant women by next-generation sequencing. *Virology*. 2016;493:202–8.
- Flores-Miramontes MG, Torres-Reyes LA, Alvarado-Ruiz L, Romero-Martínez SA, Ramírez-Rodríguez V, Balderas-Peña LM, Vallejo-Ruiz V, Piña-Sánchez P, Cortés-Gutiérrez EI, Jave-Suárez LF. Human papillomavirus genotyping by linear array and next-generation sequencing in cervical samples from Western Mexico. *Virol J*. 2015;12(1):161.
- Forslund O, Antonsson A, Nordin P, Stenquist B, Hansson BG. A broad range of human papillomavirus types detected with a general PCR method suitable for analysis of cutaneous tumours and normal skin. *J Gen Virol*. 1999;80(Pt 9):2437–43.
- Li J, Pan Y, Xu Z, Wang Q, Hang D, Shen N, Liu M, Zhang C, Abliz A, Deng Q, et al. Improved detection of human papillomavirus harbored in healthy skin with FAP6085/64 primers. *J Virol Methods*. 2013;193(2):633–8.
- Papillomavirus Episteme (PaVE). <https://pave.niaid.nih.gov>. Accessed 12 July 2016.
- Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc*. 2010;5(4):725–38.
- Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER suite: protein structure and function prediction. *Nat Methods*. 2015;12(1):7–8.
- I-Tasser Protein Structure and Function Predictions. <http://zhanglab.ccmb.med.umich.edu/I-TASSER/>. Accessed 20 July 2016.
- Zhang Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*. 2008;9(1):40.

24. Bernard H-U, Chan S-Y, Manos MM, Ong C-K, Villa LL, Delius H, Peyton CL, Bauer HM, Wheeler CM. 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.* 1994;170(5):1077–85.
25. Tornesello ML, Duraturo ML, Botti G, Greggi S, Piccoli R, De Palo G, Montella M, Buonaguro L, Buonaguro FM. Prevalence of alpha-papillomavirus genotypes in cervical squamous intraepithelial lesions and invasive cervical carcinoma in the Italian population. *J Med Virol.* 2006;78(12):1663–72.
26. Datta P, Bhatla N, Dar L, Patro AR, Gulati A, Kriplani A, Singh N. Prevalence of human papillomavirus infection among young women in North India. *Cancer Epidemiol.* 2010;34(2):157–61.
27. Shaltout MF, Sallam HN, AbouSeeda M, Moiety F, Hemedat H, Ibrahim A, Sherbini ME, Rady H, Gopala K, DeAntonio R. Prevalence and type distribution of human papillomavirus among women older than 18 years in Egypt: a multicenter, observational study. *Int J Infect Dis.* 2014;29:226–31.
28. Youssef MA, Abdelsalam L, Harfoush RA, Talaat IM, Elkattan E, Mohey A, Abdella RM, Farhan MS, Foad HA, Elsayed AM. Prevalence of human papilloma virus (HPV) and its genotypes in cervical specimens of Egyptian women by linear array HPV genotyping test. *Infect agents cancer.* 2016;11(1):1.
29. Kantathavorn N, Mahidol C, Sritana N, Sricharunrat T, Phoolcharoen N, Auewarakul C, Teerayathanakul N, Taepisitpong C, Saeloo S, Sornsamdang G. Genotypic distribution of human papillomavirus (HPV) and cervical cytology findings in 5906 Thai women undergoing cervical cancer screening programs. *Infect agents cancer.* 2015;10(1):7.
30. Hariri S, Unger ER, Sternberg M, Dunne EF, Swan D, Patel S, Markowitz LE. Prevalence of genital human papillomavirus among females in the United States, the national health and nutrition examination survey, 2003–2006. *J Infect Dis.* 2011;204(4):566–73.
31. Roksandić-Križan I, Bošnjak Z, Perić M, Đurkin I, Zujčić Atalić V, Vuković D. Distribution of Genital Human Papillomavirus (HPV) Genotypes in Croatian Women with Cervical Intraepithelial Neoplasia (CIN)—A Pilot Study. *Coll Antropol.* 2013;37(4):1179–83.
32. Santos FM, Gurgel A, Lobo C, Freitas A, Silva-Neto J, Silva L. Prevalence of human papillomavirus (HPV), distribution of HPV types, and risk factors for infection in HPV-positive women. *Genetics and molecular research. GMR.* 2016;15(2):1–9.
33. So KA, Kim MJ, Lee K-H, Lee I-H, Kim MK, Lee YK, Hwang C-S, Jeong MS, Kee M-K, Kang C. The impact of high-risk HPV genotypes other than HPV 16/18 on the natural course of abnormal cervical cytology: a Korean HPV cohort study. *Cancer Res Treat.* 2016;48(4):1313.
34. Gurgel APAD, Chagas BS, Amaral CMM, Albuquerque EMB, Serra IGSS, Silva Neto JC, Muniz MTC, Freitas AC. Prevalence and genetic variability in capsid L1 gene of rare human papillomaviruses (HPV) found in cervical lesions of women from North-East Brazil. *Biomed Res Int.* 2013;2013:546354.
35. Thomas M, Narayan N, Pim D, Tomaić V, Massimi P, Nagasaka K, Kranjec C, Gammoh N, Banks L. Human papillomaviruses, cervical cancer and cell polarity. *Oncogene.* 2008;27(55):7018–30.
36. Rastogi RP, Kumar A, Tyagi MB, Sinha RP. Molecular mechanisms of ultraviolet radiation-induced DNA damage and repair. *J Nucleic Acids.* 2010;2010:592980.
37. Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol.* 1993;10(3):512–26.
38. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol.* 2016;33(7):1870–4.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

