

RESEARCH ARTICLE

E6 and E7 Gene Polymorphisms in Human Papillomavirus Types-58 and 33 Identified in Southwest China

Zuyi Chen^{1,2,3}✉, Yaling Jing^{1,2,3}✉, Qiang Wen^{1,2,3}, Xianping Ding^{1,2,3*}, Tao Wang^{1,2,3}, Xuemei Mu^{1,2,3}, Yuwei Chenzhang^{1,2,3}, Man Cao^{1,2,3}

1 Key Laboratory of Bio-Resources and Eco-Environment, Ministry of Education, Sichuan University, Chengdu, China, **2** Bio-resource Research and Utilization Joint Key Laboratory of Sichuan and Chongqing, Sichuan and Chongqing, China, **3** Institute of Medical Genetics, College of Life Science, Sichuan University, Chengdu, China

✉ These authors contributed equally to this work.

* brainding@scu.edu.cn



OPEN ACCESS

Citation: Chen Z, Jing Y, Wen Q, Ding X, Wang T, Mu X, et al. (2017) E6 and E7 Gene Polymorphisms in Human Papillomavirus Types-58 and 33 Identified in Southwest China. PLoS ONE 12(1): e0171140. doi:10.1371/journal.pone.0171140

Editor: Maria Lina Tornesello, Fondazione IRCCS Istituto Nazionale dei Tumori, ITALY

Received: October 13, 2016

Accepted: January 16, 2017

Published: January 31, 2017

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

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

Funding: The authors received no specific funding for this work.

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

Abstract

Cancer of the cervix is associated with infection by certain types of human papillomavirus (HPV). The gene variants differ in immune responses and oncogenic potential. The E6 and E7 proteins encoded by high-risk HPV play a key role in cellular transformation. HPV-33 and HPV-58 types are highly prevalent among Chinese women. To study the gene intratypic variations, polymorphisms and positive selections of HPV-33 and HPV-58 E6/E7 in southwest China, HPV-33 (E6, E7: n = 216) and HPV-58 (E6, E7: n = 405) E6 and E7 genes were sequenced and compared to others submitted to GenBank. Phylogenetic trees were constructed by Maximum-likelihood and the Kimura 2-parameters methods by MEGA 6 (Molecular Evolutionary Genetics Analysis version 6.0). The diversity of secondary structure was analyzed by PSIPred software. The selection pressures acting on the E6/E7 genes were estimated by PAML 4.8 (Phylogenetic Analyses by Maximum Likelihood version 4.8) software. The positive sites of HPV-33 and HPV-58 E6/E7 were contrasted by ClustalX 2.1. Among 216 HPV-33 E6 sequences, 8 single nucleotide mutations were observed with 6/8 non-synonymous and 2/8 synonymous mutations. The 216 HPV-33 E7 sequences showed 3 single nucleotide mutations that were non-synonymous. The 405 HPV-58 E6 sequences revealed 8 single nucleotide mutations with 4/8 non-synonymous and 4/8 synonymous mutations. Among 405 HPV-58 E7 sequences, 13 single nucleotide mutations were observed with 10/13 non-synonymous mutations and 3/13 synonymous mutations. The selective pressure analysis showed that all HPV-33 and 4/6 HPV-58 E6/E7 major non-synonymous mutations were sites of positive selection. All variations were observed in sites belonging to major histocompatibility complex and/or B-cell predicted epitopes. K93N and R145 (I/N) were observed in both HPV-33 and HPV-58 E6.

Introduction

Cervical cancer is the third most common cancer in women worldwide, and a persistent infection of the high-risk human papillomavirus (HPV) is a major risk factor for cervical cancer

[1,2]. Approximately 500000 new cases of cervical cancer are diagnosed every year, and the disorder causes 250000 deaths; more than 85% of all patients are from low-income countries [3,4]. The risk of developing cervical cancer in HPV-infected patients is 50-fold higher than uninfected women [5,6]. Genital HPV types are typically divided into two groups according to their presumed oncogenic potential. HPV-16, 18, 31, 33, 35, 39, 45, 51, 52, 53, 56, 58, and 59 are the common high oncogenic risk types [7].

The early expressing proteins *E6* and *E7* of high-risk HPV are the primary oncoproteins involved in human epithelial cell immortalization and transformation, and act through their interactions with numerous host proteins [8]. A primary target of *E6* is the p53 tumor suppressor protein. Additionally, *E6* activates telomerase expression and modulates the activities of PDZ domain-containing proteins and tumor necrosis factor receptors [9,10]. The *E7* proteins have evolved as a primary target of the retinoblastoma (Rb) family of proteins that control the activity of E2F transcription factors for degradation, resulting in the increase in the expression of E2F-responsive genes [11]. Moreover, knockdown of *E6* and *E7* expression in cervical cancer cells successfully suppresses cell growth and induces apoptosis [12]. Thus, HPV *E6/E7* can form an ideal target for diagnostic detection and therapeutic vaccine design.

Worldwide, HPV-33 and HPV-58 account for approximately 5% and 2% of cervical cancer cases, respectively; nevertheless, these HPV types show remarkably high infection rates in East Asia [13–16]. The areas of China, where cases of carcinoma *in situ* and cervical cancer caused by HPV-58 were more than the cases caused by HPV-18, were rated second; while HPV-33 was ranked third or fourth [17–19]. Furthermore, in our previous study, we reported that the detection rates of HPV-16 and 18 had decreased and that of HPV-58 and HPV-33 had increased over a 6-year period [20]. Therefore, researchers in China should devote greater attention to the more prevalent high-risk types of HPV-58 and HPV-33 for vaccine design than areas HPV-58 and HPV-33 are not so common.

The distribution and variation of HPV types show some degree of ethnic and geographical differences; consequently, the diagnosis, treatment, and ideal vaccine constructs necessarily need to be local [21,22]. Amino acid changes of high-risk HPV *E6/E7* might affect carcinogenic potential, host immune responses, and therapeutic effects of the vaccine. Among different types of HPV, there are types and variants that can acquire biological advantages through fixed mutations in their genomes, and even small variations could result in small adaptive improvements, possibly altering the composition of a HPV population [23]. This study may contribute more information in understanding HPV distribution by contrasting the HPV-33 and HPV-58 positive selections. Unfortunately, the data available on HPV-33 and HPV-58 *E6/E7* are still limited, especially in China. Therefore, the present study aimed to investigate the HPV-33 and HPV-58 *E6/E7* gene polymorphisms, intratypic variations and positive selections in southwest China by examining a large series of samples covering the Chinese population. This study can provide essential data for future research on viral prevention and therapeutics. Above all, our findings may offer critical information for developing diagnostic probes and vaccines, specifically designed based on HPV-33 and HPV-58 *E6/E7*.

Materials and Methods

Ethics statement

The study was approved by the Education and Research Committee and the Ethics Committee of Sichuan University, Sichuan, China. Before the samples were collected, a written informed consent was obtained from all the patients or their guardians, and patient/study subject privacy was carefully protected.

Study population and specimen collection

Between January 1, 2009 and December 31, 2015, 16793 (age range 16–87, average age 33.02, median age 29) cervical specimens were collected from women who underwent cervical screenings, histology and cytology evaluations for cervical diseases at Sichuan Reproductive Health Research Center Affiliated Hospital, The Angel Women's and Children's Hospital, The Chengdu Western Hospital Maternity Unit, The People's Hospital of Pengzhou, and Chongqing The Fourth Hospital. Women over 14 years of age and with visible cervical lesions and/or HPV-related diseases (e.g., cervicitis, cervical intraepithelial neoplasia, and cervical cancer) were eligible for inclusion [20]. Specimens were collected from the participants and stored in a preservative buffer (NaCl 9g, C₆H₅CO₂Na 10g, H₂O 1L) at -20°C.

Genomic DNA extraction and HPV typing

HPV-DNA was extracted and examined using the Human Papillomavirus Genotyping Kit For 23 Types (Yaneng Bio, Shenzhen, China) according to the manufacturer's instructions. This kit enabled the classification of the 23 HPV types (HPV 16, 18, 31, 33, 35, 39, 45, 51, 52, 53, 56, 58, 59, 66, 68, 73, 83, MM4, 6, 11, 42, 43, and 44). 326 HPV-33 and 592 HPV-58 positive samples were collected.

PCR amplification and identification of variants

The complete *E6/E7* genes of HPV-33 were first amplified by polymerase chain reaction (PCR) in the thermal cycler (Longgene, Hangzhou, China) using the primers: 5' -AAAAAAGTAGG GTGTAACCGA-3' and 5' -TGCCACTGTTCATCTGCTGT-3'; this step was followed by a second round of amplification using the inner primers: 5' -ACGGTGCATATATAAAGCAAACA TT-3' and 5' -CTTCTACCTCAAACCAACC-AGTACA-3', when needed. The HPV-58 *E6/E7* fragment was amplified using specific primers described previously [24]. Each 50 µL PCR reaction mixture contained 5 µL of extracted DNA (10–100 ng), 200 µmol MgCl₂ and dNTPs, 2 U of Pfu DNA polymerase (Sangon Biotech, Shanghai, China), and 0.25 µmol of each primer. The cycling conditions employed were as follows: 95°C for 10 min; 35 cycles of 94°C for 50 sec, 54°C (different for each gene) for 60 sec, 72°C for 60 sec; and a final step of 72°C for 7 min. The PCR amplification products were visualized on 2% agarose gels stained with GeneGreen nucleic acid dye under the ultraviolet light WFH-202. Target products were sequenced at Sangon Biotech (Shanghai, China), and all the data were confirmed by repeating the PCR amplification and sequence analysis at least twice.

The sequences were subsequently analyzed by NCBI Blast and DNAMAN version 5.2.2 (Lynnon Biosoft, Quebec, Canada). HPV-33 and HPV-58 DNA nucleotide positions were numbered according to their reference sequences M12732.1 and D90400, respectively. The *E6/E7* sequences of 33HE01-33HE15 in HPV-33 and 58HE01-58HE17 in HPV-58 were published with the GenBank accession codes KX354744-KX354775.

Sequence analysis

The secondary structures of the reference proteins were predicted by PSIPred server (<http://bioinf.cs.ucl.ac.uk/psipred/>) with the default parameters, which provided a simple and accurate secondary structure prediction method. Using a very stringent cross validation technique to evaluate the method's performance, PSIPRED 3.2 achieves an average Q3 score of 81.6% [25]. The data were analyzed using SPSS version 19 (IBM, Armonk, NY, USA). The Pearson χ^2 test was used to confirm the results. $P < 0.05$ was considered statistically significant. A mutation of which the frequency $\geq 10\%$ was considered as a major mutation.

Phylogenetic analysis of HPV-33 and HPV-58

Phylogenetic trees of respective HPV-33 *E6/E7* and HPV-58 *E6/E7* variation patterns were constructed through Maximum-likelihood trees by MEGA 6 using Kimura's two-parameter model. The tree topology was evaluated by employing bootstrap resampled 1000 times [26]. The reference sequences used to construct the phylogenetic branches were collected from the GenBank sequence database and were listed in S1 and S2 Tables. HPV-33 and HPV-58 lineage and sublineage classification based on other unique *E6/E7* have been described in previous studies [24,27–36]. Numbers above the branches indicated the bootstrap values that are greater than 75%.

Epitope prediction

ProPred-I server (<http://www.imtech.res.in/raghava/propred1/>) [37] was used to predict the human leukocyte antigen (HLA) class I binding promiscuous epitope(s) at the default settings (allele selection: all, threshold: 4%, tabular results: 4, proteasome filter: off, immunoproteasome: off). To predict epitope(s) for HLA class II alleles, ProPred software (<http://www.imtech.res.in/raghava/propred/>) [38] was used at the default settings (threshold %: 3.0, allele: all, display top scorers: blank). ABCpred server (http://www.imtech.res.in/raghava/abcpred/ABC_submission.html), which uses an artificial neural network, was used for the prediction of B-cell epitope(s) in the *E6* and *E7* sequences at the default settings (threshold value: 0.51, window length: 16; overlapping filter: no) [39].

Selective pressure analysis and homology comparison

To estimate for positive selection at particular sites of the HPV-33 and HPV-58 *E6/E7* gene sequences, the *codeml* program in the PAML 4.8 was used. This program performed the likelihood ratio tests (LRTs) to infer nonsynonymous and synonymous nucleotide divergence for coding regions employing the method described by Nei and Gojobor [40–42]. HPV-33 and HPV-58 *E6/E7* protein sequences were aligned by ClustalX 2.1 (<ftp://ftp.ebi.ac.uk/pub/software/clustalw2/>) [43].

Results

Among all the HPV-58 and HPV-33 samples, 405 sequences of HPV-58 *E6/E7* gene and 216 sequences of HPV-33 *E6/E7* gene were obtained owing to the small number of copies of infected HPV in some women and limited amplicons obtained for sequencing.

Gene polymorphism of HPV-33 *E6/E7*

Compared with the HPV-33 reference sequence (GenBank: M12732.1), 104 of the 216 HPV-33 isolates (48.15%) showed complete *E6/E7* sequence homology with the reference and the remaining 112 (51.85%) isolates showed nucleotide variation.

Altogether, 11 nucleotide positions of *E6/E7* fragment showing sequence polymorphisms were identified among the 216 isolates. 8 biallelic mutations occurred in the 450-bp *E6* Open Reading Frame (ORF); more specifically, 6/8 were non-synonymous substitutions and 2/8 were synonymous mutations. In addition, two biallelic mutations and one triallelic (C706A/T) mutation were found over the 294-bp *E7* ORF, resulting in 4 amino acid changes of S29T, A45V, A45E, and Q97L (Table 1). The sequence variability of *E7* was lower than that of *E6*. The average probability of a nucleotide sequence deviation from the prototype was 20.37 substitutions per 10000 bp for *E6* and 15.28 substitutions per 10000 bp for *E7*. Mutations generating a frame shift or a premature stop codon were not observed.

Table 1. Nucleotide sequence mutation at E6/E7 of 15 HPV 33 isolates.

sequence pattern	HPV33 E6								HPV33 E7			n (216)	sub-lineages
	2	2	3	3	3	4	5	5	6	7	8		
	1	7	2	6	8	4	4	4	5	0	6		
	3	3	9	4	7	6	2	9	8	6	2		
M12732	A	A	G	A	A	A	G	T	G	C	A		
33HE01	-	-	-	-	-	-	-	-	-	-	-	104	A1
33HE02	-	-	-	-	-	-	-	-	C	-	-	3	A1
33HE03	-	-	-	-	-	-	-	-	C	T	-	1	A1
33HE04	-	-	-	-	-	-	-	-	-	T	-	31	A1
33HE05	-	-	-	-	-	-	T	-	-	T	-	4	A1
33HE06	-	-	-	-	-	-	T	-	-	-	-	16	A1
33HE07	-	-	-	-	-	-	T	-	-	-	T	1	A1
33HE08	-	-	-	-	-	-	T	-	C	-	T	1	A1
33HE09	-	-	-	-	-	-	T	-	C	-	-	4	A1
33HE10	-	-	C	-	-	-	T	-	C	-	-	7	A1
33HE11	-	-	-	-	-	-	-	-	-	A	-	1	A1
33HE12	-	-	-	-	C	-	-	-	-	A	-	1	A1
33HE13	C	-	-	-	C	-	-	C	-	A	-	25	A1
33HE14	C	G	-	-	-	-	-	-	-	-	T	1	A3
33HE15	C	G	-	C	C	G	-	-	-	-	T	16	A3
aa mutation						Q	R						
	K		S	N	K	1	1		S	A	Q		
	3	-	7	8	9	1	4	-	2	4	9		
	5		4	6	3	3	5		9	5	7		
	N		T	H	N	R	I		T	V/E	L		
second structure	-	S	H	-	-	H	-	-	-	-	-		

The nucleotides conserved with respect to the reference sequence were marked with a dash (-), whereas a variation position was indicated by a letter. The “S” in the last row of the table means Sheet, the “H” means Helix. “n” represents the number of each sequence pattern among the 216 samples examined.

doi:10.1371/journal.pone.0171140.t001

Two non-synonymous mutations (S74T and Q113R) occurred in the HPV-33 E6 sequences encoding the alpha helix. No non-synonymous mutation was detected in the HPV-33 E7 sequences encoding the alpha helix or the beta sheet.

Among the 216 determinable samples of HPV-33, the A1 and A3 variants were found in 199 (92.13%) and 17 (7.87%) samples, respectively (Table 1 and Fig 1). All of the variants belonged to prototype-like groups (lineage A); whereas no nonprototype-like variant (lineage B) of HPV-33 was observed in our study (Fig 1).

Gene polymorphism of HPV-58 E6/E7

When compared with the HPV-58 reference sequences (GenBank: D90400), 11.85% (48/405) of the isolates showed complete E6/E7 sequence homology with the reference, and the nucleotide variation rate of HPV-58 E6/E7 was 88.15% (357/405) among the 405 HPV-58 isolates. Overall, 8 nucleotide sequence variations were detected in the 450-bp E6 ORF; specifically, 4 mutations were non-synonymous and 4 mutations were synonymous. 13 single nucleotide changes were identified in the 297-bp E7 ORF, out of which 10 substitutions were non-synonymous and 3 substitutions were synonymous mutations (Table 2). Sequence variability was higher for E7 than for E6. The average probability of a nucleotide sequence deviation from

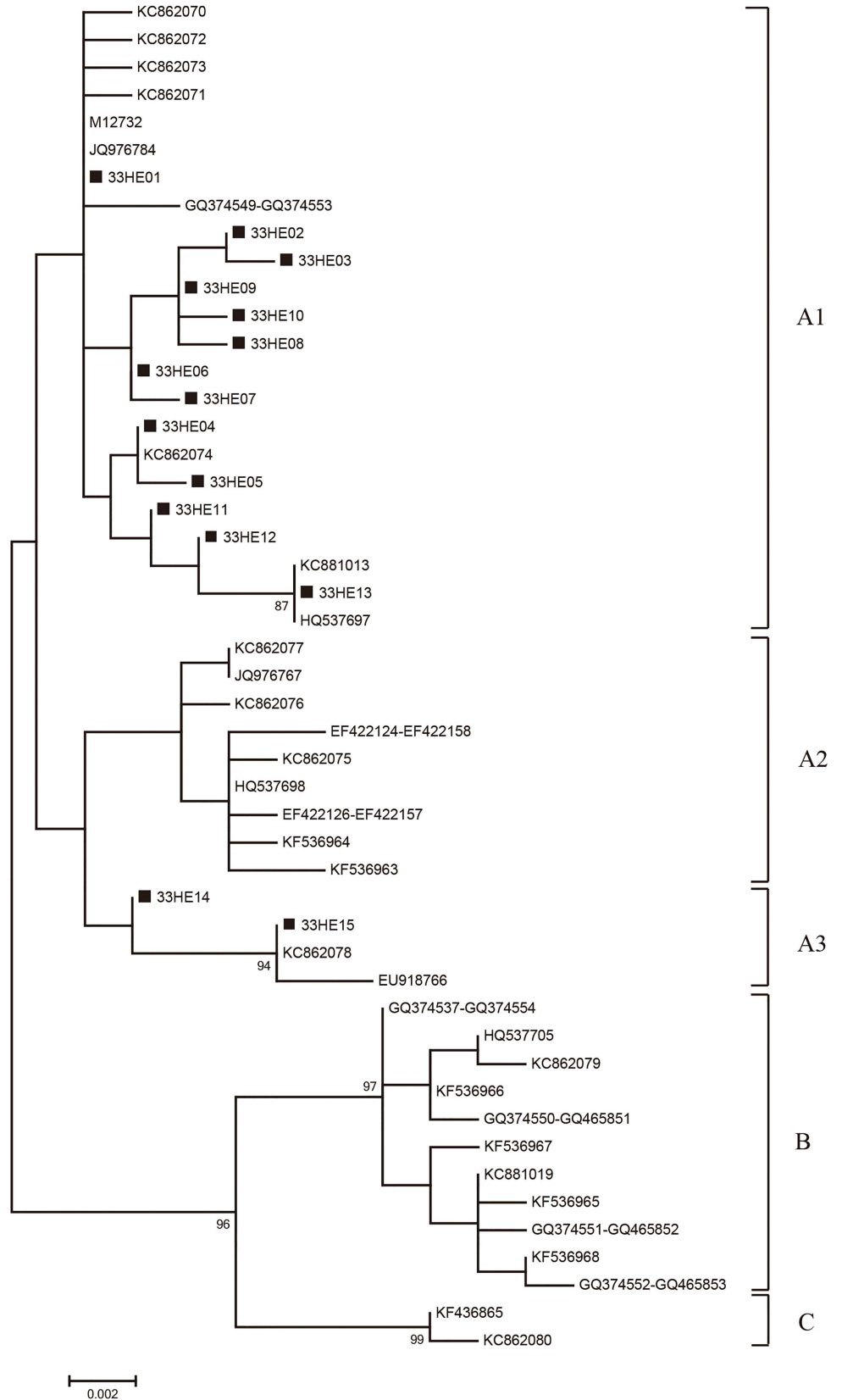


Fig 1. The Maximum-likelihood tree of HPV-33. The Maximum-likelihood tree of HPV-33 variants based on E6-E7 combined sequences. Numbers above the branches indicate the bootstrap values that are greater than 75%.

doi:10.1371/journal.pone.0171140.g001

prototype was 20.63 substitutions/10000 bp for E6 and 78.90 substitutions/10000 bp for E7. There were no mutations generating a frame shift or a premature stop codon.

In addition, there were no non-synonymous mutations in the HPV-58 E6 sequences encoding the alpha helix or the beta sheet; while three non-synonymous mutations (T74A, D76E, and V77A) occurred in HPV-58 E7 sequences encoding the alpha helix.

Among 405 isolates of HPV-58, the A and B variants were found in 404 (99.75%) and 1 (0.25%) isolates, respectively. The nonprototype-like variant (lineage B) of HPV-58 was rare in our study. The sub-lineage A1, A2, and A3 variants were found in 168 (41.48%), 163 (40.25%), and 73 (18.02%) HPV-58 isolates, respectively (Table 2 and Fig 2). Lineages C and D were not observed in the present study (Fig 2).

Table 2. Nucleotide sequence mutation at E6/E7 of 17 HPV 58 isolates.

sequence pattern	HPV58 E6										HPV58 E7										n(405)	sub-lineages	
	1	2	2	3	3	3	3	5	5	6	6	7	7	7	7	7	7	7	7	8			8
	8	0	5	0	6	8	9	4	9	3	9	2	4	5	6	6	6	9	9	0	0		
	7	3	9	7	7	8	5	3	9	2	4	6	4	5	0	1	3	3	8	1	3		
D90400	C	G	A	C	C	A	T	G	G	C	G	T	T	C	G	G	A	A	C	C	T		
58HE01	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	48	A1
58HE02	T	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	7	A1
58HE03	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	30	A1
58HE04	-	-	-	-	-	C	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	A1
58HE05	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	C	1	A1
58HE06	-	-	G	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	C	2	A1
58HE07	-	-	-	-	-	C	-	A	-	-	-	-	G	-	-	-	-	-	-	-	C	5	A1
58HE08	-	-	-	-	-	C	-	-	A	-	-	-	G	-	-	-	-	-	-	-	C	13	A1
58HE09	-	-	-	-	-	C	-	-	-	-	-	-	G	-	-	-	-	-	-	-	C	61	A1
58HE10	-	-	-	T	-	-	-	-	-	-	-	-	G	-	-	A	-	-	-	-	-	1	A2
58HE11	-	-	-	T	-	-	-	A	-	-	A	-	G	-	-	A	-	-	-	-	-	7	A2
58HE12	-	-	-	T	-	-	-	-	-	T	A	-	G	-	-	A	-	-	-	-	-	1	A2
58HE13	-	-	-	T	-	-	-	-	-	-	A	-	G	A	-	A	G	-	-	-	-	26	A2
58HE14	-	-	-	T	-	-	-	-	-	-	A	-	G	-	-	A	-	-	-	-	-	128	A2
58HE15	-	-	-	T	-	-	-	-	-	T	-	-	G	-	A	-	-	-	-	-	-	69	A3
58HE16	-	-	-	T	-	-	-	A	-	T	-	-	G	-	A	-	-	-	-	-	-	4	A3
58HE17	-	C	-	T	A	C	-	-	-	-	-	-	G	-	A	-	-	G	T	A	-	1	B1
aa mutation								R															
		E			D	K		1	R	T	G			T	G	G	T	T		D	V		
	-	3	-	-	8	9	-	4	9	2	4	-	-	6	6	6	6	7	-	7	7		
		2			6	3		5		0	1			1	3	3	4	4		6	7		
		Q			E	N		K	K	I	R			N	S	D	A	A		E	A		
second structure	S	-	S	H	-	-	-	-	-	-	-	-	S	-	-	-	-	H	H	H	H		

The nucleotides conserved with respect to the reference sequence were marked with a dash (-), whereas a variation position was indicated by a letter. The "S" in the last row of the table means Sheet, the "H" means Helix. "n" represents the number of each sequence pattern among the 405 samples examined.

doi:10.1371/journal.pone.0171140.t002

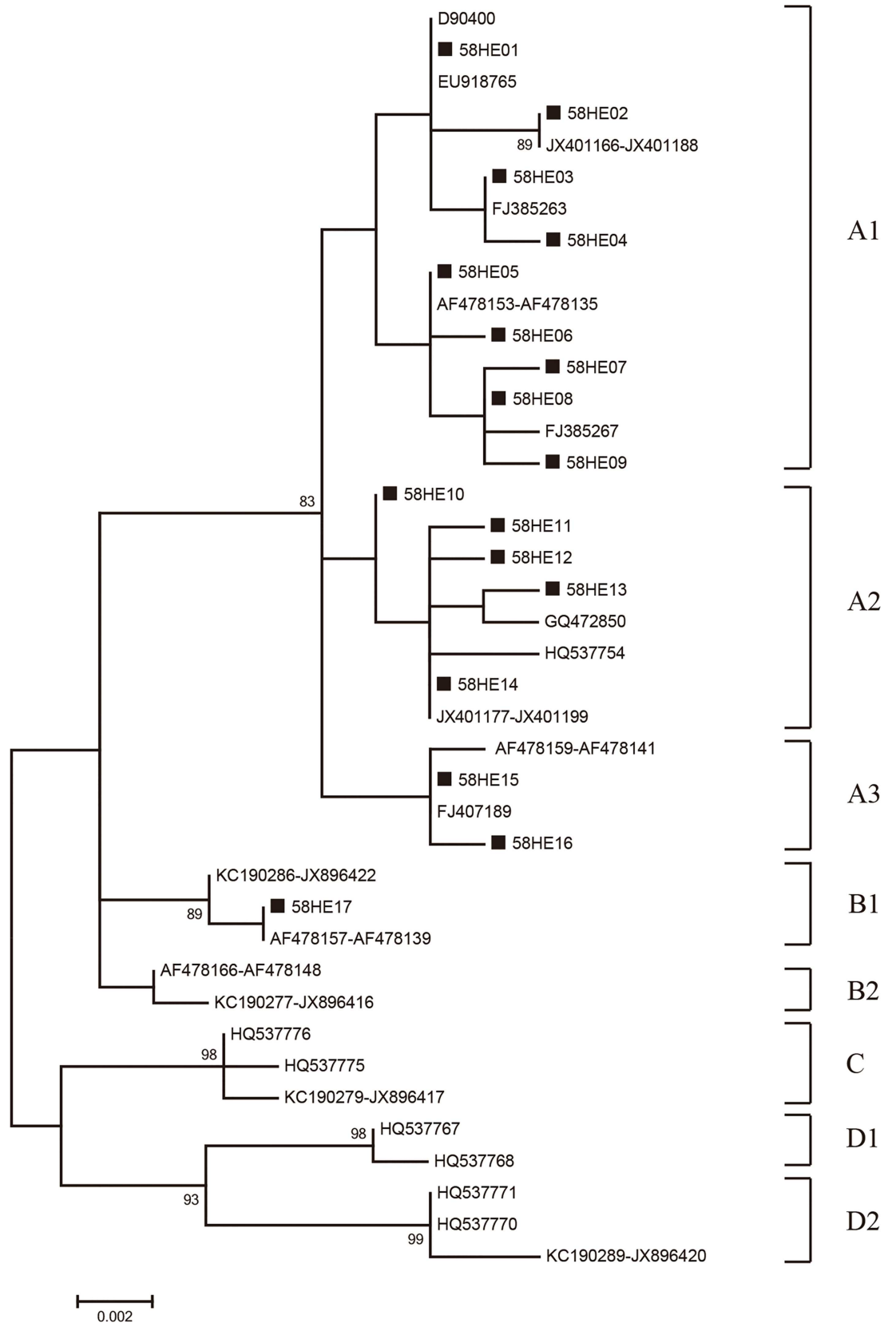


Fig 2. The Maximum-likelihood tree of HPV-58. The Maximum-likelihood tree of HPV-58 variants based on *E6-E7* combined sequences. Numbers above the branches indicate the bootstrap values that are greater than 75%.

doi:10.1371/journal.pone.0171140.g002

Table 3. Site-specific tests for positive selection on HPV-33 E6/E7.

Models	InL	Estimates of parameters	2ΔI	E6 Positively selected sites	E7 Positively selected sites
M7	-1193.188	p = 0.005, q = 0.047		NA	
M8	-1166.06	p0 = 0.970, p = 0.005, q = 0.012, p1 = 0.030, ω = 86.838	54.256 p<0.01	35K**, 93K**, 145R**	29S**, 45A**, 97Q*

doi:10.1371/journal.pone.0171140.t003

Selective pressure analysis and homology comparison

The variable dN/dS ratios were tested among various lineages using the PAML 4.8 software. The results of the selective pressure analysis of HPV-58 and HPV-33 E6 and E7 genes (P-value, 0.1) have been summarized in Tables 3 and 4. The positively selected sites for HPV-33 E6 were K35N, K93N, and R145I; for HPV-33 E7 were S29T, A45V, A45E, and Q97L; for HPV-58 E6 were K93N, R145K; and for HPV-58 E7 were T20I, G41R, G63S, and G63D. The homology comparison results were shown in S1 Fig. K93N and R145 (I/N) of HPV-33 and HPV-58 E6 were observed at the same two positions.

Predicted MHC and B-cell epitopes

The results of major histocompatibility complex (MHC) epitopes (only epitopes binding not less than 10 HLA class alleles were shown) and B-cell epitopes (only score ≥ 0.85 were shown), as predicted, were summarized in Fig 3, the details were shown in S3–S5 Tables. For MHC I, 8 good epitopes were obtained for HPV-33 E6, 5 for HPV-33 E7, 11 for HPV-58 E6 and 5 for HPV-58 E7, respectively; For MHC II, 6 good epitopes were obtained for HPV-33 E6, 6 for HPV-33 E7, 9 for HPV-58 E6 and 6 for HPV-58 E7, respectively; For B-cell, 1 high score epitope was obtained for HPV-33 E6, 3 for HPV-33 E7, 2 for HPV-58 E6 and 3 for HPV-58 E7, respectively. Amino acids change may influence the epitope’s binding ability, decrease or increase the binding ability, or even lead to disappearance of epitopes or appearance of new epitopes (S3–S5 Tables).

Discussion

HPV types are divided into different genus according to their biological characteristics. HPV-33 and 58 are known to be closely related to each other, and belong to the α-9 species, which is the principal species consisting of almost all carcinogenic types [44]. The intratypic variations observed in E6 and E7 can offer useful information for the distinction and identification of known or new HPV types [45,46]. The present study revealed higher frequencies of HPV-58 E6/E7 variations than HPV-33 E6/E7. Furthermore, HPV-33 E7 was observed to be steadier than E6; therefore, E7 was selected as a more suitable target for diagnostic detection of HPV-

Table 4. Site-specific tests for positive selection on HPV-58 E6/E7.

Models	InL	Estimates of parameters	2ΔI	E6 Positively selected sites	E7 Positively selected sites
M7	-1273.301	p = 0.005, q = 0.012		NA	
M8	-1256.611	p0 = 0.972, p = 35.347, q = 99.000, p1 = 0.028, ω = 24.138	33.38 p<0.01	93K**, 145R**	20T**, 63G**

Tables 3 and 4: In L, the log-likelihood difference between the two models; 2ΔI, twice the log-likelihood difference between the two models; the positively selected sites were identified with posterior probability ≥ 0.9 using Bayes empirical Bayes (BEB) approach, an asterisk indicates posterior probability ≥ 0.95, and two asterisks indicate posterior probability ≥ 0.99. NA means not allowed. NS means the sites under positive selection, but not reaching the significance level of 0.9.

doi:10.1371/journal.pone.0171140.t004

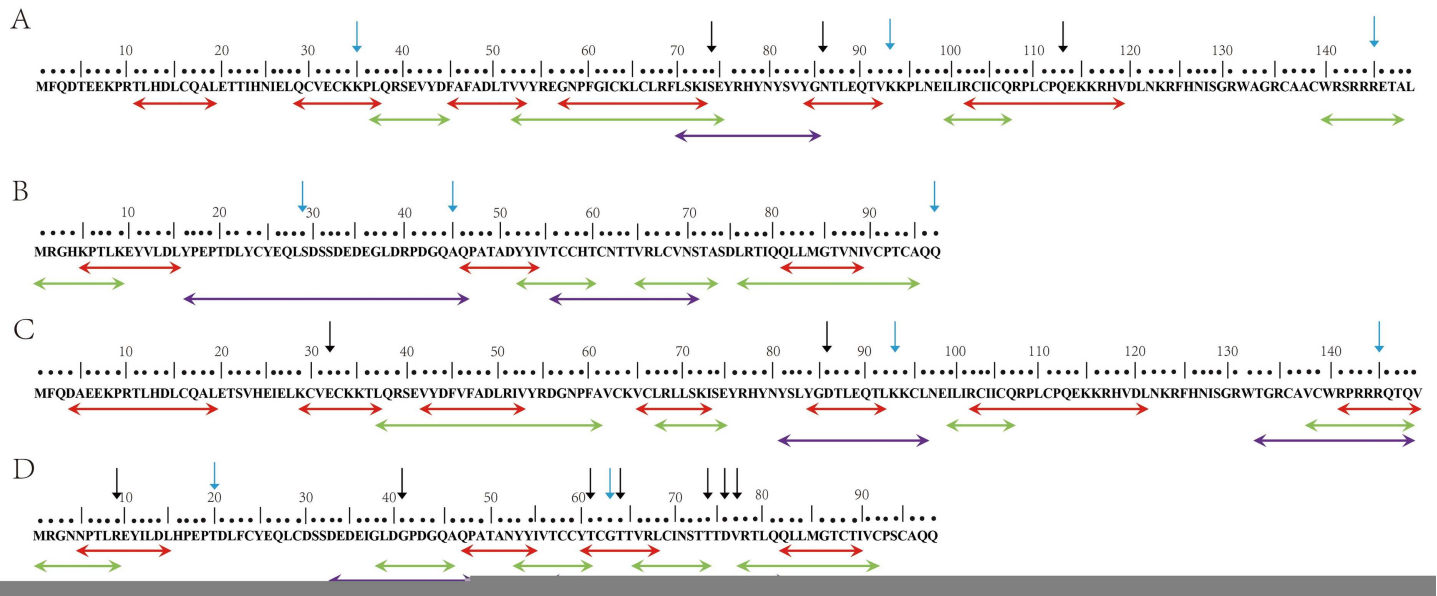


Fig 3. Predicted MHC and B-cell epitopes. Epitopes within the reference sequence of (A) HPV-33 *E6* protein; (B) HPV-33 *E7* protein; (C) HPV-58 *E6* protein; and (D) HPV-58 *E7* protein. Red indicates a consensus of the predicted MHC I epitopes binding not less than 10 HLA I alleles. Green indicates a consensus of the predicted MHC II epitopes binding not less than 10 HLA II alleles. Purple indicates a consensus of the predicted B-cell epitopes of which the score ≥ 0.85 for the combined physical-chemical properties. Black arrow indicates corresponding mutation is only a non-synonymous mutation, and blue arrow indicates corresponding mutation is a non-synonymous mutation and positive selection, simultaneously.

doi:10.1371/journal.pone.0171140.g003

33 than *E6*. On the contrary, HPV-58 *E6* was steadier than *E7* ($\chi^2 = 16.015$, $P < 0.01$); and was chosen as a more appropriate target for diagnostic detection of HPV-58 than *E7*.

Previous studies revealed two main groups—prototype-like group (lineage A) and nonprototype-like group (lineage B) among HPV-33 variants [27, 29]. However, all of the HPV-33 variants belonged to the prototype-like group, and the nonprototype-like group was not observed in the determinable samples of our study; these findings were similar to the results obtained by Chen et al and Godinez et al [29,30]. The A1 sub-lineage (92.13%) was determined as the major type in southwest China. In a worldwide study on *E6/E7* segment of 213 HPV-33 samples, the A2 sub-lineage (59.72%) was the main type in the Asia and Oceania region [29]. However, no A2 sub-lineage was detected in this study. 7.87% of the sub-lineage belonged to A3.

In the current research, the prototype-like group of HPV-58 was the dominating variant; this classification of lineages was in agreement with the results of previous studies [36,47]. Nevertheless, the distribution of HPV-58 variant lineages around China was observed to be different in our study; B1 variant was a nonprototype variant, while the C variant was a nonprototype variant lineage in Hong Kong and Tai Wan [36,47]. When compared with a previous report on HPV-58 in southwest China [33], B lineages were newly found.

The most common mutations observed in HPV-33 *E6* were A231C (K35N, 42/216, a positive mutation) and A387C (K93N, 42/216, a positive mutation); in HPV-33 *E7* were G658C (S29T, 36/216, a positive mutation) and C706A (A45E, 27/216); in HPV-58 *E6* were C307T (237/405) and A388C (K93N 111/405); while the most common non-synonymous mutations in HPV-58 *E7* were T744G (319/405) and G761A (G63D, 163/405, a positive mutation). These mutations may be considered important when *E6/E7* are chosen as targets for primer design or diagnostic detection.

To the best of our knowledge, the present study is the first to contrast the HPV-33 and HPV-58 *E6/E7* sites under positive selection. The key characteristic of positive selection is that

it causes an unusually rapid rise in allele frequency, under positive selection, the positive mutation(s) may rise to high frequency rapidly, help species to adapt to the environments [48]. The selective pressure analysis showed that all the sites that evolved under positive selection were common non-synonymous mutations, indicating that the positively selected variations beneficial for HPV-33 and HPV-58 to accommodate their environments are wide-spread. HPV-33 falls next to HPV-58 in the phylogenetic tree [44]; and remarkably, the positive sites K93N and R145 (I/N) were observed in both HPV-33 and HPV-58 *E6*, they may have evolutionary significance in making HPV-33 and HPV-58 adaptive to their environments. The HPV-58 *E6* mutations T20I and G63S have been reported to increase the risk of developing cervical cancer [36]; interestingly, in the present study, we found that these two mutations were positively selected. Specific intratypic HPV genome variations may be related to virus infectivity, pathogenicity, progression to cervical cancer, viral particle assembly, and host immune response. Of all these variations in the present study, the five newly-reported mutations have been only found in southwest China until now: G329C (S74T) and G542T (R145I, a positive variation) for HPV-33 *E6*, D658C (S29T, a positive variation) for HPV-33 *E7* as well as A259G and T395C for HPV-58 *E6* [24,29,34–36,47,49–51].

Amino acid positions 145–149 form the PDZ binding domain in the *E6* protein; amino acid positions 21–29 form short linear motif responsible for Rb binding in *E7* protein; whereas the positions 58, 61, 91, and 94 act as Zn binding sites in the *E7* protein [29,52]. In the present study, G542T (R145I, a positive mutation) in HPV-33 *E6* and G543A (R145K, a positive mutation) were found at residues 145–149; G658 (S29, a positive mutation) in HPV-33 *E7* was found at residues 21–29, G632T (T20I, a positive mutation) in HPV-58 *E7* was found beside residues 21–29; C755A (T61N) in HPV-58 *E7* was found at residues involved in Zn binding for *E7* protein. Until now, there is no published data to prove that immunity to one variant can prevent another variant in HPV infections [53].

Modern immunoinformatics provide new strategies for the design and identification of antigen-specific epitopic sites that could be used as vaccine targets. The prediction of MHC and B-cell epitope in this study can be potentially used for the vaccine development against specific HPV variants in the Chinese population. In the present study, variations other than HPV-33 *E6* K93N, HPV-33 *E7* Q97L and HPV-58 *E7* T20I were observed in sites belonging to ideal B-cell and/or MHC predicted epitopes. Some variations, like HPV-58 *E6* R145K, occurred at the sites belonging to both B-cell and MHC epitopes; Amino acids change may influence the epitopes, and then the immune recognition of HPV-infected cells. For example, the score of HPV-58 *E6* B-cell epitope 81-96YSLYGDTLEQTLKKCL was 0.90, because of the mutation D86E, the score of epitope 81-96YSLYGETLEQTLNKCL decreased to 0.86; the HPV-58 *E7* predicted epitope 77-85VRTLQQLLM disappeared because of the mutation V77A. However, further experimentation is required to validate the prediction through immunoinformatics.

This is the first study to examine the changes in *E6/E7* epitopes of HPV-58 variants in southwest China and to report the variants of HPV-33 *E6/E7* in China. The data presented in this study may have significant implications in understanding the intrinsic geographical relatedness and biological differences between HPV-33 and HPV-58 *E6/E7*, and may also contribute to the design of clinical diagnostic probes and second-generation therapeutic vaccines based on HPV-33 and HPV-58 *E6/E7*.

Supporting Information

S1 Table. HPV-33 reference sequences used in phylogenetic analysis.
(DOCX)

S2 Table. HPV-58 reference sequences used in phylogenetic analysis.
(DOCX)

S3 Table. ProPred I analysis for binding of E6/E7 sequences to HLA class I.
(DOCX)

S4 Table. ProPred analysis for binding of E6/E7 sequences to HLA class II.
(DOCX)

S5 Table. Predicted B-cell epitopes of the E6/E7 gene.
(DOCX)

S1 Fig. The homology comparison results of HPV-33 and HPV-58 E6/E7. Note: The residues conserved across HPV types were shown in capital letters, whereas the nonconserved residues are given in lowercase letters.
(TIF)

Acknowledgments

We thank for the following hospitals for the sample collection and technical assistance with pathology: Sichuan Reproductive Health Research Center Affiliated Hospital, The Angel Women's and Children's Hospital, The Chengdu Western Hospital Maternity Unit, and The Peoples' Hospital of Pengzhou, and Chongqing the fourth hospital.

Author Contributions

Conceptualization: ZC XD.

Data curation: XD YJ ZC.

Formal analysis: TW XM YC.

Funding acquisition: XD ZC.

Investigation: ZC YJ QW TW XM YC MC XD.

Methodology: ZC YJ QW.

Project administration: XD ZC.

Resources: XD.

Software: ZC YJ MC.

Supervision: XD ZC.

Validation: XD MC.

Visualization: ZC YJ XD.

Writing – original draft: ZC YJ.

Writing – review & editing: ZC XD.

References

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA Cancer J Clin.* 2011; 61: 69–90. doi: [10.3322/caac.20107](https://doi.org/10.3322/caac.20107) PMID: [21296855](https://pubmed.ncbi.nlm.nih.gov/21296855/)

2. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer*. 2010; 127: 2893–2917. doi: [10.1002/ijc.25516](https://doi.org/10.1002/ijc.25516) PMID: [21351269](https://pubmed.ncbi.nlm.nih.gov/21351269/)
3. Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol*. 1999; 189: 12–19. doi: [10.1002/\(SICI\)1096-9896\(199909\)189:1<12::AID-PATH431>3.0.CO;2-F](https://doi.org/10.1002/(SICI)1096-9896(199909)189:1<12::AID-PATH431>3.0.CO;2-F) PMID: [10451482](https://pubmed.ncbi.nlm.nih.gov/10451482/)
4. Shen M, Ding X, Li T, Chen G, Zhou X. Sequence Variation Analysis of HPV-18 Isolates in Southwest China. *PLoS One*. 2013; 8: e56614. doi: [10.1371/journal.pone.0056614](https://doi.org/10.1371/journal.pone.0056614) PMID: [23451059](https://pubmed.ncbi.nlm.nih.gov/23451059/)
5. Bosch FX, Manos MM, Muñoz N, Sherman M, Jansen AM, Peto J, et al. Prevalence of human papillomavirus in cervical cancer: a worldwide perspective. *J Natl Cancer Inst*. 1995; 87: 796–802. PMID: [7791229](https://pubmed.ncbi.nlm.nih.gov/7791229/)
6. Bosch FX, Lorincz A, Muñoz N, Meijer CJ, Shah KV. The causal relation between human papillomavirus and cervical cancer. *J Clin Pathol*. 2002; 55: 244–265. PMID: [11919208](https://pubmed.ncbi.nlm.nih.gov/11919208/)
7. Muñoz N, Bosch FX, de Sanjosé S, Herrero R, Castellsagué X, Shah KV, et al. Epidemiologic classification of human papillomavirus types associated with cervical cancer. *N Engl J Med*. 2003; 348: 518–527. doi: [10.1056/NEJMoa021641](https://doi.org/10.1056/NEJMoa021641) PMID: [12571259](https://pubmed.ncbi.nlm.nih.gov/12571259/)
8. Rampias T, Sasaki C, Psyrii A, Psyrii A. Molecular mechanisms of HPV induced carcinogenesis in head and neck. *Oral Oncol*. 2014; 50: 356–363. doi: [10.1016/j.oraloncology.2013.07.011](https://doi.org/10.1016/j.oraloncology.2013.07.011) PMID: [23953776](https://pubmed.ncbi.nlm.nih.gov/23953776/)
9. Scheffner M, Huijbregtse JM, Vierstra RD, Howley PM. The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. *Cell*. 1993; 75: 495–505. PMID: [8221889](https://pubmed.ncbi.nlm.nih.gov/8221889/)
10. Thomas M, Narayan N, Pim D, Tomaić V, Massimi P, Nagasaka K, et al. Human papillomaviruses, cervical cancer and cell polarity. *Oncogene*. 2008; 27: 7018–7030. doi: [10.1038/onc.2008.351](https://doi.org/10.1038/onc.2008.351) PMID: [19029942](https://pubmed.ncbi.nlm.nih.gov/19029942/)
11. Moody CA, Laimins LA. Human papillomavirus oncoproteins: pathways to transformation. *Nat Rev Cancer*. 2010; 10: 550–560. doi: [10.1038/nrc2886](https://doi.org/10.1038/nrc2886) PMID: [20592731](https://pubmed.ncbi.nlm.nih.gov/20592731/)
12. Chang JT, Kuo TF, Chen YJ, Chiu CC, Lu YC, Li HF, et al. Highly potent and specific siRNAs against E6 or E7 genes of HPV16- or HPV18-infected cervical cancers. *Cancer Gene Ther*. 2010; 17: 827–836. doi: [10.1038/cgt.2010.38](https://doi.org/10.1038/cgt.2010.38) PMID: [20885450](https://pubmed.ncbi.nlm.nih.gov/20885450/)
13. Chan PK, Ho WC, Yu MY, Pong WM, Chan AC, Chan AK, et al. Distribution of human papillomavirus types in cervical cancers in Hong Kong: current situation and changes over the last decades. *Int J Cancer*. 2009; 125: 1671–1677. doi: [10.1002/ijc.24495](https://doi.org/10.1002/ijc.24495) PMID: [19507252](https://pubmed.ncbi.nlm.nih.gov/19507252/)
14. Chen CA, Liu CY, Chou HH, Chou CY, Ho CM, Twu NF, et al. The distribution and differential risks of human papillomavirus genotypes in cervical preinvasive lesions: A Taiwan Cooperative Oncologic Group Study. *Int J Gynecol Cancer*. 2006; 16: 1801–1808. doi: [10.1111/j.1525-1438.2006.00655.x](https://doi.org/10.1111/j.1525-1438.2006.00655.x) PMID: [17009975](https://pubmed.ncbi.nlm.nih.gov/17009975/)
15. Asato T, Maehama T, Nagai Y, Kanazawa K, Uezato H, Kariya K. A Large Case-Control Study of Cervical Cancer Risk Associated with Human Papillomavirus Infection in Japan, by Nucleotide Sequencing-Based Genotyping. *J Infect Dis*. 2004; 189: 1829–1832. doi: [10.1086/382896](https://doi.org/10.1086/382896) PMID: [15122519](https://pubmed.ncbi.nlm.nih.gov/15122519/)
16. Hwang T. Detection and typing of human papillomavirus DNA by PCR using consensus primers in various cervical lesions of Korean women. *J Korean Med Sci*. 1999; 14: 593–599. doi: [10.3346/jkms.1999.14.6.593](https://doi.org/10.3346/jkms.1999.14.6.593) PMID: [10642935](https://pubmed.ncbi.nlm.nih.gov/10642935/)
17. Li J, Mei J, Wang X, Hu L, Lin Y, Yang P. Human papillomavirus type-specific prevalence in women with cervical intraepithelial neoplasm in Western China. *J Clin Microbiol*. 2012; 50: 1079–1081. doi: [10.1128/JCM.06214-11](https://doi.org/10.1128/JCM.06214-11) PMID: [22170939](https://pubmed.ncbi.nlm.nih.gov/22170939/)
18. Sun ZR, Ji YH, Zhou WQ, Zhang SL, Jiang WG, Ruan Q. Characteristics of HPV prevalence among women in Liaoning province, China. *Int J Gynaecol Obstet*. 2010; 109: 105–109. doi: [10.1016/j.ijgo.2009.11.026](https://doi.org/10.1016/j.ijgo.2009.11.026) PMID: [20138618](https://pubmed.ncbi.nlm.nih.gov/20138618/)
19. Zhao R, Zhang WY, Wu MH, Zhang SW, Pan J, Zhu L, et al. Human papillomavirus infection in Beijing, People's Republic of China: a population-based study. *Br J Cancer*. 2009; 101: 1635–1640. doi: [10.1038/sj.bjc.6605351](https://doi.org/10.1038/sj.bjc.6605351) PMID: [19862002](https://pubmed.ncbi.nlm.nih.gov/19862002/)
20. Chen Z, Wang Q, Ding X, Li Q, Zhong R, Ren H. Characteristics of HPV prevalence in Sichuan Province, China. *Int J Gynaecol Obstet*. 2015; 131: 277–280. doi: [10.1016/j.ijgo.2015.06.027](https://doi.org/10.1016/j.ijgo.2015.06.027) PMID: [26391672](https://pubmed.ncbi.nlm.nih.gov/26391672/)
21. Li N, Franceschi S, Howell-Jones R, Snijders PJ, Clifford GM. 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. doi: [10.1002/ijc.25396](https://doi.org/10.1002/ijc.25396) PMID: [20473886](https://pubmed.ncbi.nlm.nih.gov/20473886/)
22. de Sanjose S, Quint WG, Alemany L, Geraets DT, Klaustermeier JE, Lloveras B, et al. Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study. *Lancet Oncol*. 2010; 11: 1048–1056. doi: [10.1016/S1470-2045\(10\)70230-8](https://doi.org/10.1016/S1470-2045(10)70230-8) PMID: [20952254](https://pubmed.ncbi.nlm.nih.gov/20952254/)

23. Calleja-Macias IE, Villa LL, Prado JC, Kalantari M, Allan B, Williamson AL, et al. Worldwide genomic diversity of the high-risk human papillomavirus types 31, 35, 52, and 58, four close relatives of human papillomavirus type 16. *J Virol*. 2005; 79: 13630–13640. doi: [10.1128/JVI.79.21.13630-13640.2005](https://doi.org/10.1128/JVI.79.21.13630-13640.2005) PMID: [16227283](https://pubmed.ncbi.nlm.nih.gov/16227283/)
24. Chan PK, Zhang C, Park JS, Smith-McCune KK, Palefsky JM, Giovannelli L, et al. Geographical distribution and oncogenic risk association of human papillomavirus type 58 E6 and E7 sequence variations. *Int J Cancer*. 2013; 132: 2528–2536. doi: [10.1002/ijc.27932](https://doi.org/10.1002/ijc.27932) PMID: [23136059](https://pubmed.ncbi.nlm.nih.gov/23136059/)
25. Buchan DW, Minneci F, Nugent TC, Bryson K, Jones DT. Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Res*. 2013; 41: W349–357. doi: [10.1093/nar/gkt381](https://doi.org/10.1093/nar/gkt381) PMID: [23748958](https://pubmed.ncbi.nlm.nih.gov/23748958/)
26. Tamura K, Stecher G, Peterson D, Filipinski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol*. 2013; 30: 2725–2729. doi: [10.1093/molbev/mst197](https://doi.org/10.1093/molbev/mst197) PMID: [24132122](https://pubmed.ncbi.nlm.nih.gov/24132122/)
27. Chen Z, Schiffman M, Herrero R, Desalle R, Anastos K, Segondy M, 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. doi: [10.1371/journal.pone.0020183](https://doi.org/10.1371/journal.pone.0020183) PMID: [21673791](https://pubmed.ncbi.nlm.nih.gov/21673791/)
28. Yang L, Yang H, Chen J, Huang X, Pan Y, Li D, et al. Genetic variability of HPV-58 E6 and E7 genes in Southwest China. *Infect Genet Evol*. 2014; 21: 395–400. doi: [10.1016/j.meegid.2013.12.009](https://doi.org/10.1016/j.meegid.2013.12.009) PMID: [24368255](https://pubmed.ncbi.nlm.nih.gov/24368255/)
29. Chen AA, Heideman DA, Boon D, Chen Z, Burk RD, De Vuyst H, et al. Human papillomavirus 33 worldwide genetic variation and associated risk of cervical cancer. *Virology*. 2014; 448: 356–362. doi: [10.1016/j.virol.2013.10.033](https://doi.org/10.1016/j.virol.2013.10.033) PMID: [24314666](https://pubmed.ncbi.nlm.nih.gov/24314666/)
30. Godínez JM, Heideman DA, Gheit T, Alemany L, Snijders PJ, Tommasino M, et al. Differential presence of Papillomavirus variants in cervical cancer: an analysis for HPV33, HPV45 and HPV58. *Infect Genet Evol*. 2013; 13: 96–104. doi: [10.1016/j.meegid.2012.09.011](https://doi.org/10.1016/j.meegid.2012.09.011) PMID: [23022714](https://pubmed.ncbi.nlm.nih.gov/23022714/)
31. Burk RD, Harari A, Chen Z. Human papillomavirus genome variants. *Virology*. 2013; 445: 232–243. doi: [10.1016/j.virol.2013.07.018](https://doi.org/10.1016/j.virol.2013.07.018) PMID: [23998342](https://pubmed.ncbi.nlm.nih.gov/23998342/)
32. Wu X, Zhang C, Feng S, Liu C, Li Y, Yang Y, et al. Detection of HPV types and neutralizing antibodies in Gansu province, China. *J Med Virol*. 2009; 81: 693–702. doi: [10.1002/jmv.21435](https://doi.org/10.1002/jmv.21435) PMID: [19235880](https://pubmed.ncbi.nlm.nih.gov/19235880/)
33. Wu EQ, Zha X, Yu XH, Zhang GN, Wu YG, Fan Y, et al. Profile of physical status and gene variation of human papillomavirus 58 genome in cervical cancer. *J Gen Virol*. 2009; 90: 1229–1237. doi: [10.1099/vir.0.008227-0](https://doi.org/10.1099/vir.0.008227-0) PMID: [19264634](https://pubmed.ncbi.nlm.nih.gov/19264634/)
34. Garbuglia AR, Carletti F, Minosse C, Piselli P, Zaniratti MS, Serraino D, et al. Genetic variability in E6 and E7 genes of human papillomavirus -16, -18, -31 and -33 from HIV-1-positive women in Italy. *New Microbiol*. 2007; 30: 377–382. PMID: [18080672](https://pubmed.ncbi.nlm.nih.gov/18080672/)
35. S Khouadri S, Villa LL, Gagnon S, Koushik A, Richardson H, Ferreira 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. doi: [10.1086/507431](https://doi.org/10.1086/507431) PMID: [16960775](https://pubmed.ncbi.nlm.nih.gov/16960775/)
36. Chan PK, Lam CW, Cheung TH, Li WW, Lo KW, Chan MY, et al. Association of human papillomavirus type 58 variant with the risk of cervical cancer. *J Natl Cancer Inst*. 2002; 94: 1249–1253. PMID: [12189229](https://pubmed.ncbi.nlm.nih.gov/12189229/)
37. Singh H, Raghava GP. ProPred1: prediction of promiscuous MHC Class-I binding sites. *Bioinformatics*. 2003; 19: 1009–1014. PMID: [12761064](https://pubmed.ncbi.nlm.nih.gov/12761064/)
38. Singh H, Raghava GP. ProPred: prediction of HLA-DR binding sites. *Bioinformatics*. 2001; 17: 1236–1237. PMID: [11751237](https://pubmed.ncbi.nlm.nih.gov/11751237/)
39. Saha S, Raghava GP. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins*. 2006; 65: 40–48. doi: [10.1002/prot.21078](https://doi.org/10.1002/prot.21078) PMID: [16894596](https://pubmed.ncbi.nlm.nih.gov/16894596/)
40. Nei M, Gojobori T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol*. 1986; 3: 418–426. PMID: [3444411](https://pubmed.ncbi.nlm.nih.gov/3444411/)
41. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol*. 2007; 24: 1586–1591. doi: [10.1093/molbev/msm088](https://doi.org/10.1093/molbev/msm088) PMID: [17483113](https://pubmed.ncbi.nlm.nih.gov/17483113/)
42. Hamza AA, Robene-Soustrade I, Jouen E, Lefeuvre P, Chiroleu F, Fisher-Le Saux M, et al. MultiLocus Sequence Analysis- and Amplified Fragment Length Polymorphism-based characterization of xanthomonads associated with bacterial spot of tomato and pepper and their relatedness to *Xanthomonas* species. *Syst Appl Microbiol*. 2012; 35: 183–190. doi: [10.1016/j.syapm.2011.12.005](https://doi.org/10.1016/j.syapm.2011.12.005) PMID: [22336775](https://pubmed.ncbi.nlm.nih.gov/22336775/)
43. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res*. 1997; 25: 4876–4882. PMID: [9396791](https://pubmed.ncbi.nlm.nih.gov/9396791/)

44. Schiffman M, Rodriguez AC, Chen Z, Wacholder S, Herrero R, Hildesheim A, et al. A population-based prospective study of carcinogenic human papillomavirus variant lineages, viral persistence, and cervical neoplasia. *Cancer Res.* 2010; 70: 3159–3169. doi: [10.1158/0008-5472.CAN-09-4179](https://doi.org/10.1158/0008-5472.CAN-09-4179) PMID: [20354192](https://pubmed.ncbi.nlm.nih.gov/20354192/)
45. Stewart AC, Eriksson AM, Manos MM, Muñoz N, Bosch FX, Peto J, et al. Intratype variation in 12 human papillomavirus types: a worldwide perspective. *J Virol.* 1996; 70: 3127–3136. PMID: [8627792](https://pubmed.ncbi.nlm.nih.gov/8627792/)
46. Bernard HU, Burk RD, Chen Z, van Doorslaer K, zur Hausen H, de Villiers EM. Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology.* 2010; 401: 70–79. doi: [10.1016/j.virol.2010.02.002](https://doi.org/10.1016/j.virol.2010.02.002) PMID: [20206957](https://pubmed.ncbi.nlm.nih.gov/20206957/)
47. Chang YJ, Chen HC, Lee BH, You SL, Lin CY, Pan MH, et al. Unique variants of human papillomavirus genotypes 52 and 58 and risk of cervical neoplasia. *Int J Cancer.* 2011; 129: 965–973. doi: [10.1002/ijc.25724](https://doi.org/10.1002/ijc.25724) PMID: [20949622](https://pubmed.ncbi.nlm.nih.gov/20949622/)
48. Iyengar VK, Reeve HK, Eisner T. Paternal inheritance of a female moth's mating preference. *Nature.* 2002; 419: 830–832. doi: [10.1038/nature01027](https://doi.org/10.1038/nature01027) PMID: [12397356](https://pubmed.ncbi.nlm.nih.gov/12397356/)
49. Liu JH, Lu ZT, Wang GL, Zhou WQ, Liu C, Yang LX, et al. Variations of human papillomavirus type 58 E6, E7, L1 genes and long control region in strains from women with cervical lesions in Liaoning province, China. *Infect Genet Evol.* 2012; 12: 1466–1472. doi: [10.1016/j.meegid.2012.05.004](https://doi.org/10.1016/j.meegid.2012.05.004) PMID: [22659102](https://pubmed.ncbi.nlm.nih.gov/22659102/)
50. Cento V, Rahmatalla N, Ciccozzi M, Perno CF, Ciotti M. Intratype variations of HPV 31 and 58 in Italian women with abnormal cervical cytology. *J Med Virol.* 2011; 83: 1752–1761. doi: [10.1002/jmv.22201](https://doi.org/10.1002/jmv.22201) PMID: [21837791](https://pubmed.ncbi.nlm.nih.gov/21837791/)
51. Bae JH, Cheung JL, Lee SJ, Luk AC, Tong SY, Chan PK, et al. Distribution of Human Papillomavirus Type 58 Variants in Progression of Cervical Dysplasia in Korean Women. *J Microbiol Biotechnol.* 2009; 19: 1051–1054. PMID: [19809265](https://pubmed.ncbi.nlm.nih.gov/19809265/)
52. Chemes LB, Camporeale G, Sánchez IE, de Prat-Gay G, Alonso LG. Cysteine-rich positions outside the structural zinc motif of human papillomavirus E7 provide conformational modulation and suggest functional redox roles. *Biochemistry.* 2014; 53: 1680–1696. doi: [10.1021/bi401562e](https://doi.org/10.1021/bi401562e) PMID: [24559112](https://pubmed.ncbi.nlm.nih.gov/24559112/)
53. Yue Y, Yang H, Wu K, Yang L, Chen J, Huang X, et al. Genetic Variability in L1 and L2 Genes of HPV-16 and HPV-58 in Southwest China. *PLoS One.* 2013; 8: e55204. doi: [10.1371/journal.pone.0055204](https://doi.org/10.1371/journal.pone.0055204) PMID: [23372836](https://pubmed.ncbi.nlm.nih.gov/23372836/)