Virology Journal

CrossMark

L1 and *L2* gene polymorphisms in HPV-58 and HPV-33: implications for vaccine design and diagnosis

Zuyi Chen[1,2†], Yaling Jing[1,2†], Qiang Wen[1,2], Xianping Ding[1,2*], Shun Zhang[1,2], Tao Wang[1,2], Yiwen Zhang[1,2] and Jianhui Zhang[1,2]

Abstract

Background: Cervical cancer is associated with infection by certain subtypes of human papillomavirus (HPV). The *L1* protein comprising HPV vaccine formulations elicits high-titre neutralizing antibodies and confers protection against specific HPV subtypes. HPV *L2* protein is an attractive candidate for cross-protective vaccines. HPV-33 and HPV-58 are very prevalent among Chinese women.

Methods: To study the gene intratypic variations and polymorphisms of HPV-33 and HPV-58 *L1/L2* in Sichuan China, HPV-33 and HPV-58 *L1* and *L2* genes were sequenced and compared with other genes submitted to GenBank. Phylogenetic trees were constructed by maximum-likelihood and the Kimura 2-parameters methods (MEGA 6). The secondary structure was analyzed by PSIPred software, and HPV-33 and HPV-58 L1 homology models were created by SWISS-MODEL software. The selection pressures acting on the *L1/L2* genes were estimated by PAML 4.8.

Results: Among 124 HPV-33 *L1* sequences 20 single nucleotide mutations were observed included 8/20 non-synonymous and 12/20 synonymous mutations. The 101 HPV-33 *L2* sequences included 12 single nucleotide mutations comprising 7/12 non-synonymous and 5/12 synonymous mutations. The 223 HPV-58 *L1* sequences included 32 single nucleotide mutations comprising 9/32 non-synonymous and 23/32 synonymous mutations. The 201 HPV-58 *L2* sequences comprised 26 single nucleotide mutations including 9/26 non-synonymous and 17/26 synonymous mutations. Selective pressure analysis showed that most of the common non-synonymous mutations showed a positive selection. HPV-33 and HPV-58 *L2* were more stable than HPV-33 and HPV-58 *L1*.

Conclusions: HPV-33 and HPV-58 *L2* were better candidates as clinical diagnostic targets compared with HPV-33 and HPV-58 *L1*. Clinical diagnostic probes and second-generation polyvalent vaccines should be designed on the basis of the unique sequence of HPV-33 and 58 *L1/L2* variations in Sichuan, to improve the accuracy of clinical detection and the protective efficiency of vaccines.

Keywords: HPV-58, HPV-33, L1 and L2, Genetic diversity

* Correspondence: brainding@scu.edu.cn
†Equal contributors
[1]Key Laboratory of Bio-Resources and Eco-Environment, Ministry of Education; Institute of Medical Genetics, College of Life Science, Sichuan University, Chengdu, China
[2]Bio-resource Research and Utilization Joint Key Laboratory of Sichuan and Chongqing, Sichuan and Chongqing, China

Chen *et al. Virology Journal* (2016) 13:167

Page 2 of 12

Background

Human papillomavirus (HPV) infection plays a critical role in the development of cervical cancer [1]. The risk of developing cervical cancer in HPV-infected patients is 50-fold higher than in uninfected women [1, 2]. Approximately 500000 new cases of cervical cancer are diagnosed every year, with 250000 deaths, more than 85 % of all patients belong to low-income countries [3, 4].

Genital HPV subtypes are typically divided into two groups according to their presumed oncogenic potential. High-risk oncogenic HPV subtypes −16, −18, −58, −33, −52 and −45 are more common in Asia than elsewhere [5]. HPV-33 and HPV-58 are prevalent among Chinese women, only after HPV-16 [6–8]. HPV-33 and HPV-58 account for approximately 5 and 2 % of cervical cancer cases world-wide [9, 10]. Nevertheless, the carcinoma in situ and cervical cancer cases caused by HPV-58 and HPV-33 outnumber HPV-18, ranking second and third in China, respectively [6, 11]. Compared with the high-risk HPV-18, vaccine design in China is focused predominantly on high-risk subtypes HPV-58, and HPV-33.

The HPV genome is packaged within the major capsid late protein *L1* and the minor capsid proteins L2 [12, 13]. Five *L1* proteins form a pentamer, and 72 pentamers constitute the virus capsid. The *L1* and *L2* proteins self-assemble into virus-like particles (VLPs) that induce high levels of neutralizing antibodies and are highly protective [14, 15]. The L1-VLPs are the components used in the design of specific prophylactic vaccines. Vaccine targeting *L1* only prevents infection by specific HPV subtypes because of the lack of cross-protective epitopes in different HPV subtypes. HPV *L2* protein also induces neutralizing antibodies, the N-terminal of *L2* protein contains cross-protective epitopes and represents the target of neutralizing antibodies [14]. Therefore, targeting *L2* may be an attractive approach for a candidate vaccine.

Vaccines against HPV-16, HPV-18, HPV-6 and HPV-11 *L1* are available [15]. The two HPV vaccines Gardasil and Cervarix target only two high-risk HPV subtypes. Prevention of more than 90 % of HPV infections requires targeting at least 5 additional high-risk HPV subtypes HPV-31, HPV-33, HPV-45, HPV-52 and HPV-58 [16]. Vaccines targeting *L2* and other subtypes of *L1* are currently under investigation [14].

The data supporting HPV-58 *L2* and HPV-33 *L1/L2* in China are limited. The molecular variants of HPV-33 *L1/L2* worldwide are not widely reported. Ethno-geographical variations are observed in distribution of HPV subtypes. Among different subtypes of HPV, there are subtypes and variants that can acquire biological advantages through fixed mutations in their genomes, and even small variations could result in small adaptive improvements that could alter the composition of an HPV-infected population [17]. Altered amino acid composition affects the host immune response, and in such cases, intra-type protection may be less effective [18]. Ideally, the diagnosis and treatment of vaccine constructs needs to be developed locally.

This study investigated the HPV-33 and HPV-58 *L1/L2* gene polymorphism and intratypic variations in Sichuan, China. This study can provide essential data for future research on viral prevention and therapeutics. Above all, our study provides critical data facilitating the development of diagnostic probes and design of vaccines based on HPV-33 and HPV-58 *L1/L2*.

Methods

Study population and specimen collection

Cervical specimens were collected from 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 other institutions. Between January 1, 2009, and December 31, 2014, women presenting for cervical screening underwent histology and cytology evaluations for cervical disease. Women over 14 years of age and with visible cervical lesions and/or HPV-related diseases (e.g., cervicitis, cervical intraepithelial neoplasia) were eligible for inclusion. Cervical specimens were collected from participants and placed in a preservative buffer and stored at −20 °C.

Genomic DNA extraction

HPV-DNA was extracted and examined using a Human Papillomavirus Genotyping Kit (Yaneng Bio, Shenzhen, China) according to the manufacturer's instructions. This kit enabled the classification of 23 HPV subtypes (18 types of high-risk HPV 16, 18, 31, 33, 35, 39, 45, 51, 52, 53, 56, 58, 59, 66, 68, 73, 83, MM4 and five types of low-risk HPV genotypes 6, 11, 42, 43, and 44). In total, 478 HPV-58 and 273 HPV-33 positive samples were collected.

PCR amplification and sequencing

The entire *L1/L2* genes of HPV-33 and HPV-58 were amplified using primer pairs. The primers were designed according to the GeneBank reference sequences for HPV-33 (GenBank: M12732.1) and HPV-58 (GenBank: D90400). The primer sequences were listed in Table 1. Each 50 μL PCR reaction 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 PCR conditions were 95 °C for 10 min; 35 cycles of 50 s each at 94 °C, 54 °C (difference for each gene) for 60 s, 72 °C for 60 s, 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 ultra

Chen et al. Virology Journal (2016) 13:167

Page 3 of 12

Table 1 Primers used for the molecular characterization of HPV-33 and HPV-58 L1 and L2

Gene	Primer name	Sequence 5′ to 3′	Primer position	Product size (bp)	Annealing Temperature
33 L1	HPV33 L1 1 F	ATTTGTTCCTATTTCGCCT	5468	921	54 °C
	HPV33 L1 1R	CAGCCCTATTAAAAAAGTGTC	6388		
	HPV33 L1 2 F	CAGTACATGCAAATATCCAGATTA	6271	913	54 °C
	HPV33 L1 2R	ACATACACAAAAAACAAACAACA	7183		
33 L2	HPV33 L2 1 F	GCACATGGTGGTGTTTTAAC	4116	935	54 °C
	HPV33 L2 1R	AGTCAGGATCAGGAGCAGGT	5050		
	HPV33 L2 2 F	AATGTAACATCAAGCACGCC	4834	994	55 °C
	HPV33 L2 2R	TAAACGGACCCTAAAAACCC	5827		
58 L1	HPV58 L1 1 F	TCCATTTATTCCTATATCTCCACTA	5515	876	56 °C
	HPV58 L1 1R	GAAAAAAGAACAAACTATCCCCA	6390		
	HPV58 L1 2 F	ATGGTAGATACAGGGTTTGGAT	6240	1044	55 °C
	HPV58 L1 2R	AAACAGGAAACTGACAAAGACA	7283		
58 L2	HPV58 L2 1 F	GTATACTGGTTATGCACATGGT	4125	847	54 °C
	HPV58 L2 1R	AAAAAAGCAGGGTCAACAAC	4971		
	HPV58 L2 2 F	AATGGATACCTTTGTTATTTCTAC	4834	992	54 °C
	HPV58 L2 2R	TACTTTTTTATTGTTATTGGGACT	5825		

violet light WFH-202. Target products were sequenced by Sangon Biotech.

Variant identification and analysis

The sequences and variations were analyzed by NCBI Blast, and DNAMAN version 5.2.2. The nucleotide positions were numbered according to the GeneBank reference sequences of HPV-33 (GenBank: M12732.1) and HPV-58 (GenBank: D90400). All the data were confirmed by repeating PCR amplification and sequence analysis at least twice.

Phylogenetic trees analysis

Then, phylogenetic trees of respective HPV-33 *L1/L2* and HPV-58 *L1/L2* variation patterns were constructed with the maximum-likelihood trees using MEGA (Molecular Evolutionary Genetics Analysis Version) 6 software and Kimura's two-parameter model. The tree topologies were evaluated using bootstrap resampled 1,000 times [19].

Analysis of the selection pressures and secondary structure

To estimate the positive selections at particular sites of the HPV-33 and HPV-58 *L1/L2* gene sequences, the *codeml* program in the PAML (Phylogenetic Analyses by Maximun Likelihood) version 4.8 package was used to perform the likelihood ratio tests (LRTs) to infer non-synonymous and synonymous nucleotide divergence for coding regions by the method of Nei and Gojobor [20–23]. The secondary structure of the reference sequences were analyzed by PSIPred servers at (http://bioinf.cs.ucl.ac.uk/psipred).

PSIPred is a simple and accurate secondary structure prediction method, incorporating two feed-forward neural networks, which enable the analysis of output obtained from PSI-BLAST (Position Specific Iterated-BLAST). A very stringent cross validation of the method indicated that PSIPred 3.2 attained an average Q3 score of 81.6 % [24].

Homology models analysis

HPV33 and HPV58 L1 homology models were created by SWISS-MODEL (http://swissmodel.expasy.org/) based on the crystal structure of the HPV16 L1 pentamer [25, 26], and then visualized using Swiss-PDP viewer V4.0 software.

Results

Of all the HPV-58 and HPV-33 samples, only 223 sequences of HPV-58 *L1* gene, 201 sequences of the HPV-58 *L2* gene, 124 sequences of the HPV-33 *L1* gene, and 101 sequences of the HPV-33 *L2* gene were obtained owing to the small number of copies of infected HPV in some women and limited amplicons obtained for sequencing, and there maybe a potential sampling bias against integrated HPV genomes resulting in lost capsid genes.

Gene polymorphism of HPV-33 L1

Compared with the HPV-33 reference sequence (GenBank: M12732.1), the nucleotide variation rate of HPV-33 *L1* was 68.55 % (85/124) in the 124 HPV-33 *L1* sequences studied. We identified 20 single nucleotide changes among the 124 sequences studied. Specifically, 12/20 (60.00 %) were synonymous and 8/20 (40.00 %)

Chen *et al. Virology Journal* (2016) 13:167

Page 4 of 12

were non-synonymous mutations. Only 1 non-synonymous mutation was observed in sequence encoding the helix. The detected mutations are summarized in Table 2. The maximum-likelihood phylogenetic tree can be seen in Fig. 1a. The secondary structure predicting result of the HPV-33 *L1* was showed in Additional file 1: Figure S1 (A).

Gene polymorphism of HPV-33 *L2*

Compared with the HPV-33 reference sequence (GenBank: M12732.1), the nucleotide variation rate of HPV-33 *L2* was 100.0 % (101/101) in the 101 HPV-33 *L2* sequences studied. We identified 12 single nucleotide changes among the 101 *L2* sequences studied. Specifically, 5/12 (41.7 %) were synonymous mutations and 7/12 (58.3 %) were non-synonymous mutations. The detected mutations are summarized in Table 3. The maximum-likelihood phylogenetic tree is represented in Fig. 1b. The secondary structure predicting result of the HPV-33 *L2* was showed in Additional file 1: Figure S1 (B).

HPV-58 *L1* gene polymorphism

Compared with the HPV-58 reference sequence (GenBank: D90400), the nucleotide variation rate of HPV-58 *L1* was 96.86 % (216/223) in the 223 HPV-58 L1 sequences studied. We identified 32 single nucleotide changes among the 223 sequences studied.

Specifically, 23/32 (71.88 %) were synonymous mutations and 9/32 (28.12 %) were non-synonymous mutations. 3 non-synonymous mutations were observed in sequences encoding the helix, 1 non-synonymous mutation was observed in sequence encoding the sheet. The detected mutations are summarized in Table 4. The maximum-likelihood phylogenetic tree is shown in Fig. 2c. The secondary structure predicting result of the HPV-58 *L1* was showed in Additional file 2: Figure S2 (C).

HPV-58 *L2* gene polymorphism

Compared with the HPV-58 reference sequence (GenBank: D90400), the nucleotide variation rate of HPV-58 *L2* was 68.55 % (168/201) in the 124 HPV-58 *L2* sequences studied. We identified 26 single nucleotide

Table 2 Nucleotide sequence mutation at *L1* of 13 HPV-33 isolates

Sequence pattern	5760	5960	5990	5997	6385	6390	6396	6463	6478	6487	6520	6544	6613	6664	6673	6694	6748	6951	7044	7063	n	Sub-lineages
M12732	C	T	G	A	T	C	G	T	T	A	G	A	T	A	A	A	A	A	C	G		
KU550663	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	39	A1
KU550664	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4	A1
KU550665	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-	1	A1
KU550666	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	1	A1
KU550667	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	2	A1
KU550668	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	54	A1
KU550669	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	A	-	1	A1
KU550670	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	A	-	2	A1
KU550671	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	1	A1
KU550672	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	1	A1
KU550673	-	-	-	G	-	A	A	-	-	-	-	G	-	G	-	-	-	-	-	-	4	A1
KU550674	-	-	-	-	-	A	A	C	-	-	-	G	-	G	-	-	-	-	-	-	1	A1
KU550675	A	-	A	-	-	A	-	-	C	-	-	C	G	G	-	C	-	-	-	-	13	A2
		G	K			T	G										E	K	P			
	T	1	1			2	2										3	4	4			
aa mutation	5	-	3	3	-	6	6	-	-	-	-	-	-	-	-	-	8	5	8	-	124	
	6		3	5		6	8										5	3	4			
	N		S	R		K	E										D	T	H			
second struture	-	-	-	-	-	-	-	-	-	-	H	-	-	H	H	-	H	-	-	-		

Note: M12732 was used as reference. The nucleotides conserved with respect to the reference sequence are marked with a dash(-), whereas a variation position was indicated by a letter. Predicted amino acid changes were also shown. The "S" in the last row of the table means Sheet, the "H" means Helix
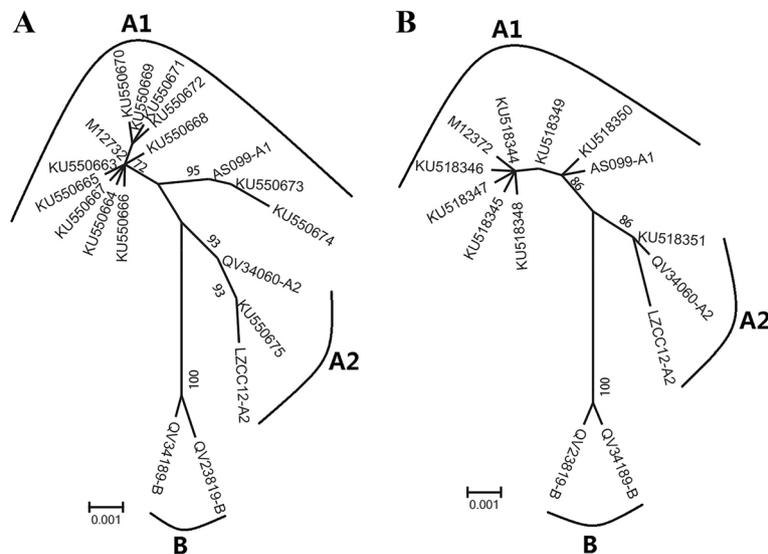
Fig. 1 Maximum-likelihood trees of HPV-33 variants. *Note*: Phylogenetic trees based on the Maximum-likelihood method of representing the relationships between variants of HPV-33, (**a**) Tree based on gene *L1*, (**b**) Tree based on gene *L2*. Sub-lineage classification based on full genomes from a reported research [27], and one or two representative sequences were chosen from each branch of the tree, respectively (AS099-A1/HQ537697, INCC0137-A2/HQ537699, QV34060-A2/HQ537698, QV23819-B/HQ537705, QV34189-B/HQ537707). Numbers above the branches indicate the bootstrap values that are greater than 70 %

Table 3 Nucleotide sequence mutation at *L2* of 8 HPV-33 isolates

Sequence pattern	HPV33 L2 Nucleotide sequence												*n*	Sub-lineages
	4	4	4	4	5	5	5	5	5	5	5	5		
	4	6	8	9	0	0	2	2	2	3	3	4		
	3	7	0	0	2	5	2	5	8	1	2	5		
	8	1	9	5	6	1	0	7	7	4	4	7		
M12732	G	T	C	C	A	T	C	G	G	G	A	C		
KU518344	-	-	-	-	-	-	T	-	-	-	-	-	73	A1
KU518345	-	-	-	-	-	-	T	-	-	A	-	-	3	A1
KU518346	-	-	-	-	-	-	T	-	-	-	-	T	1	A1
KU518347	-	-	-	-	G	-	T	-	-	-	-	-	1	A1
KU518348	-	-	-	-	-	A	T	-	-	-	-	-	8	A1
KU518349	A	-	-	-	-	-	T	-	-	-	-	-	6	A1
KU518350	A	-	A	T	-	-	T	-	-	-	-	-	3	A1
KU518351	A	C	-	T	-	-	T	C	A	-	C	-	6	A2
					I	F		D	D	D	N			
	D				2	2		3	3	3	3			
aa mutation	7	-	-	-	7	8	-	5	6	6	7	-	101	
	7				3	1		0	0	9	2			
	N				V	Y		H	N	N	T			
second struture	-	S	-	S	-	-	-	-	-	-	-	-		

Note: M12732 was used as reference. The nucleotides conserved with respect to the reference sequence are marked with a dash(-), whereas a variation position was indicated by a letter. Predicted amino acid changes were also shown. The "S" in the last row of the table means Sheet, the "H" means Helix

Chen *et al. Virology Journal* (2016) 13:167

Page 6 of 12

Table 4 Nucleotide sequence mutation at *L1* of 37 HPV-58 isolates

Sequence pattern	5747	5798	5801	5807	5809	5822	5828	5834	5873	5984	6014	6406	6414	6437	6438	6450	6475	6480	6503	6507	6511	6529	6530	6561	6648	6685	6697	6798	6799	6820	6821	6869	6889	6999	n	Sub-lineages
D90400	T	C	A	C	A	A	A	C	T	G	A	G	A	T	A	G	T	T	C	A	T	A	A	A	G	C	A	G	A	C	A	A	A	A		
KU550602	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	7	A1
KU550603	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	13	A1
KU550604	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4	A1
KU550605	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	A1
KU550606	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	31	A1
KU550607	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	G	-	-	-	1	A1
KU550608	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	A	-	-	-	-	-	-	-	1	A1
KU550609	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	A1
KU550610	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4	A1
KU550611	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	5	A1
KU550612	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	4	A1
KU550613	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	A	-	-	-	-	-	-	-	1	A1
KU550614	C	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	A	-	-	-	-	-	-	-	2	A1
KU550615	-	-	-	-	-	-	G	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	A	-	-	-	-	-	-	-	1	A1
KU550616	-	T	-	-	C	-	-	A	-	-	C	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	A	-	-	-	-	-	-	-	3	A1
KU550617	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	A	-	-	-	-	-	-	-	2	A1
KU550618	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	G	-	A	-	-	-	-	-	-	-	3	A1
KU550619	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	A	-	-	-	A	-	-	-	2	A1
KU550620	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	A	-	-	-	-	-	-	-	1	A1
KU550621	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	A	-	-	-	-	-	-	-	36	A1
KU550622	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	G	-	A	-	-	-	-	-	-	-	6	A1
KU550623	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	G	-	A	-	-	-	-	-	-	-	3	A1
KU550624	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	A	-	-	-	-	C	-	-	2	A1
KU550625	-	-	-	-	-	-	-	-	-	-	-	-	-	G	C	-	-	-	-	-	-	-	-	-	G	-	A	-	-	-	-	-	-	-	1	A2
KU550626	-	-	-	-	-	-	-	-	-	-	C	-	G	C	-	-	-	C	-	-	-	-	-	-	G	-	A	-	-	-	-	-	-	-	6	A2
KU550627	-	-	-	-	-	-	-	-	-	-	-	-	-	G	C	-	-	C	-	-	-	-	-	-	G	-	A	-	-	-	-	-	-	-	1	A2
KU550628	-	-	-	-	-	-	-	-	-	-	C	-	G	C	-	-	-	-	-	-	-	-	-	-	G	G	A	-	-	-	-	-	-	-	2	A2
KU550629	-	-	-	-	-	-	-	-	-	-	C	-	G	C	-	A	-	-	-	-	-	-	-	-	G	-	A	-	-	-	-	-	-	-	4	A2
KU550630	-	-	-	-	-	-	G	-	-	-	C	-	G	C	-	-	-	-	-	-	-	-	-	G	G	-	A	-	-	-	-	-	-	-	4	A2
KU550631	-	-	-	-	-	-	-	-	-	-	C	-	G	C	-	-	-	-	-	-	G	-	-	-	G	-	A	-	-	-	-	-	-	-	3	A2
KU550632	-	-	-	-	-	-	-	-	G	-	C	-	G	C	-	-	-	-	-	-	G	-	-	-	G	-	A	-	-	-	-	-	-	-	2	A2
KU550633	-	-	G	-	-	-	-	-	-	-	C	-	G	C	-	-	-	-	-	-	-	-	-	-	G	-	A	-	-	-	-	-	-	-	5	A2
KU550634	-	-	-	-	-	-	-	-	-	-	C	-	G	C	-	-	-	-	-	-	-	-	-	-	G	-	A	-	-	-	-	-	-	-	28	A2
KU550635	-	-	-	-	-	-	-	-	-	-	C	-	G	C	-	-	-	-	-	-	-	-	-	-	G	-	A	-	-	-	-	-	-	C	9	A2
KU550636	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	A	-	-	-	-	5	A3
KU550637	C	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	A	-	-	-	-	8	A3

Page 7 of 12

Chen *et al. Virology Journal* (2016) 13:167

Page 7 of 12

Table 4 Nucleotide sequence mutation at *L1* of 37 HPV-58 isolates *(Continued)*

KU550638	C	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	T	-	-	-	G	-	-	-	-	-	G	A	-	-	G	-	10	A3
										L							F				I				T	E				I		N		K
				N						1						3			3				3	3		4			4			4		
aa mutation	-	-	-	-	8	-	-	-	-	5	-	-	-	-	-	-	-	-	1	-	2	-	-	7	7	-	1	-	2	-	-	7	223	
					2					0									8		5			5	7		2		2			9		
					T					F									Y		M			N	D		V		T			Q		
second structure	S	S	S	-	-	-	S	-	S	-	-	H	H	-	-	-	-	-	-	H	S	S	-	S	-	-	H	H	-	-	-	-	-	

Note: D90400 was used as reference. The nucleotides conserved with respect to the reference sequence are marked with a dash(-), whereas a variation position was indicated by a letter. Predicted amino acid changes were also shown. The "S" in the last row of the table means Sheet, the "H" means Helix

changes among the 201 *L2* sequences studied. Specifically, 17/26 (65.38 %) were synonymous and 9/26 (34.62 %) were non-synonymous mutations. 3 non-synonymous mutations were observed in sequences encoding the sheet. The detected mutations are summarized in Table 5. The maximum-likelihood phylogenetic tree is displayed in Fig. 2d. The secondary structure predicting result of the HPV-58 *L2* was showed in Additional file 2: Figure S2 (D).

Selective pressure analysis

The variable dN/dS rate ratios were tested among the various lineages using the PAML4.8 [10]. HPV-33 *L1*-positive selection was seen with T56N, G133S, K135R, T226K, G268E and E385D. HPV-33 L2-positive selection was seen in D77N, D350H, D360N and N372T. The HPV-58 *L1*-positive selection included L150F, F318Y, I325M, T375N and E377D. The HPV-58 *L2*-positive

lineages included N231T, A416T and M446L. Results of the selective pressure analysis of HPV-58 and HPV-33 *L1* and *L2* genes (*P*-value, 0.1) are summarized in Tables 6, 7, 8 and 9.

Homology models analysis

HPV-16, 33, and 58 are known to be closely related and belong to the α-9 species. The whole protein sequences of HPV-16, 33, and 58 *L1* were aligned (Additional file 3: Figure S3) and the BC-loop, DE-loop, EF-loop, FG-loop, and HI-loop of HPV-33 and HPV-58 *L1* were predicted in the present study (Fig. 3). G4438A (T56N) of HPV-33 L1 and A5809C (N82T) of HPV-58 L1 were found in the BC-loop; G5990A (G133S) and G5997A (K135R) of HPV-33 were found in the DE-loop; C6390A (T266K) and G6396A (G268E) of HPV-33 L1 were found in the FG-loop; and C6688A (T375N) and
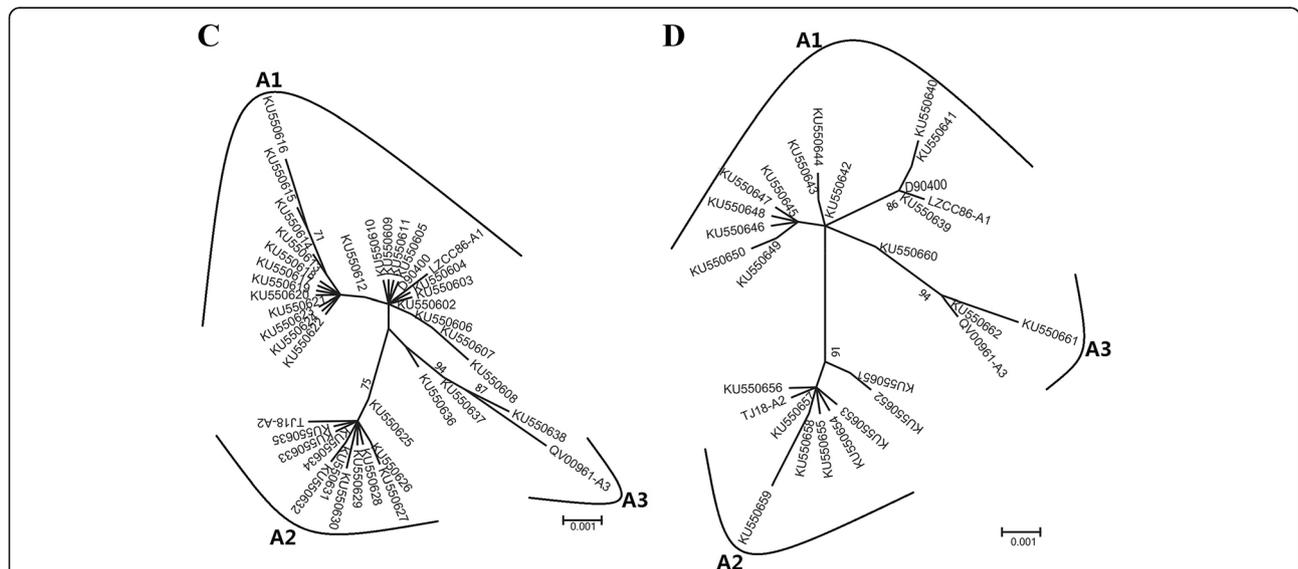
Fig. 2 Maximum-likelihood trees of HPV-58 variants. *Note*: Phylogenetic trees based on the Maximum-likelihood method of representing the relationships between variants of HPV-58, (**c**) Tree based on gene *L1*, (D) Tree based on gene *L2*. Sub-lineage classification based on full genomes from a roported research [27], and a representative sequence were chosen from the tree, respectively (LZCC86-A1/EU918765, TJ18-A2/ GQ472850, QV00961-A3/HQ537758). Numbers above the branches indicate the bootstrap values that are greater than 70 %

Chen *et al. Virology Journal* (2016) 13:167

Page 8 of 12

Table 5 Nucleotide sequence mutation at *L2* of 24 HPV-58 isolates

Sequence pattern	HPV58 L2 Nucleotide sequence																									n	Sub-lineages
	4	4	4	4	4	4	4	4	4	4	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5		
	4	4	5	5	5	6	6	6	9	9	0	0	0	1	1	2	2	3	3	4	4	4	4	5	5		
	5	7	4	5	7	0	2	2	3	4	0	0	1	1	4	0	6	2	9	3	6	8	8	4	7		
	2	0	9	8	0	9	1	7	5	5	0	0	4	9	3	6	6	5	7	7	2	9	0	0	9		
D90400	G	G	T	A	G	A	A	T	A	A	G	G	A	A	T	A	G	A	G	A	A	A	G	A	A		
KU550639	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	33	A1
KU550640	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	4	A1
KU550641	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	18	A1
KU550642	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	C	-	-	-	A	-	-	-	5	A1
KU550643	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	T	C	-	-	-	A	-	-	-	1	A1
KU550644	-	-	-	-	-	C	C	-	-	-	-	-	-	-	-	-	T	C	-	-	-	A	-	-	-	1	A1
KU550645	-	-	-	-	-	C	C	-	-	-	-	-	-	-	-	-	-	C	-	-	-	A	-	-	-	36	A1
KU550646	-	A	-	-	-	C	C	-	-	-	-	-	-	-	-	-	-	C	-	-	-	A	-	-	-	6	A1
KU550647	-	-	-	-	-	C	C	-	-	-	C	-	-	-	-	-	-	C	-	-	-	A	-	-	-	1	A1
KU550648	A	-	-	-	-	C	C	-	-	-	-	-	-	-	-	-	-	C	-	-	-	A	-	-	-	4	A1
KU550649	-	-	-	G	-	C	C	-	-	-	-	-	-	-	-	-	-	C	-	-	-	A	-	-	-	1	A1
KU550650	-	-	G	G	-	C	C	-	-	-	-	-	-	-	-	-	-	C	-	-	-	A	-	-	-	4	A1
KU550651	-	-	-	-	-	-	C	-	G	-	-	G	G	G	-	A	-	C	-	C	-	A	-	-	-	10	A2
KU550652	-	-	-	-	-	C	C	-	G	-	-	G	G	G	-	A	-	C	-	C	-	A	-	-	-	2	A2
KU550653	-	-	-	-	-	C	C	-	G	-	-	G	G	G	-	A	-	C	-	-	-	A	C	-	-	1	A2
KU550654	-	-	-	-	G	-	C	-	G	-	-	G	G	G	-	A	-	C	-	-	-	A	C	-	-	2	A2
KU550655	-	-	-	-	-	-	C	-	G	-	-	G	G	G	-	A	C	C	-	-	-	A	C	-	-	8	A2
KU550656	-	-	-	-	-	-	C	-	G	-	-	G	G	G	-	A	-	C	-	-	T	A	C	-	-	7	A2
KU550657	-	-	-	-	-	-	C	-	G	-	-	G	G	G	-	A	-	C	-	-	-	A	C	-	-	25	A2
KU550658	-	-	-	-	-	-	C	-	G	-	-	G	G	G	-	A	-	C	G	-	-	A	C	-	-	6	A2
KU550659	-	-	-	-	-	-	C	-	G	T	-	G	G	G	-	A	-	C	G	-	-	A	C	-	-	1	A2
KU550660	-	-	-	-	A	G	-	C	-	-	-	-	-	-	-	-	-	C	-	-	-	A	-	-	-	1	A3
KU550661	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	A	-	C	-	-	-	A	-	-	C	2	A3
KU550662	-	-	-	-	A	G	-	C	C	-	-	-	-	-	-	A	-	C	-	-	-	A	-	-	C	22	A3
aa mutation	G70E	S76N	-	-	-	-	-	-	-	N231T	-	D253Y	D253H	-	-	-	-	-	-	-	-	A416T	T433P	M446L		201	
second structure	-	-	-	-	-	-	-	-	-	S	S	S	-	S	S	-	-	-	S	-	-	-	S	-	-		

Note: D90400 was used as reference. The nucleotides conserved with respect to the reference sequence are marked with a dash(-), whereas a variation position was indicated by a letter. Predicted amino acid changes were also shown. The "S" in the last row of the table means Sheet, the "H" means Helix

Table 6 Site-specific tests for positive selection on HPV-33 L1

| Models	InL	Estimates of parameters	2Δl	Positively selected sites
M7	−2415.234	$p = 0.005$ q = 0.048		NA
M8	−2395.143	$p0 = 0.982$ $p = 0.114$ q =6.861 (p1 = 0.018) $\omega = 16.384$	40.183 $p < 0.01$	56 T**, 133G**, 135 K*, 266 T**, 268G**, 385E*

Chen *et al. Virology Journal* (2016) 13:167

Page 9 of 12

Table 7 Site-specific tests for positive selection on HPV-33 L2

Models	lnL	Estimates of parameters	2Δl	Positively selected sites
M7	−2185.718	$p = 0.005$ $q = 0.020$		NA
M8	−2174.779	$p0 = 0.985$ $p = 0.551$ $q = 5.040$ ($p1 = 0.015$) $\omega = 20.381$	21.878 $p < 0.01$	77D**, 350D**, 360D*, 372 N**

A6695C (E377D) of HPV-58 L1 were found in the HI-loop.

Discussion

In our previous study of 10682 patients, 3370 (31.5 %) were positive for HPV infection. High-risk subtypes were as follows: HPV16 ($n = 791$; 27.1 %), HPV58 ($n = 476$; 16.3 %), HPV33 ($n = 273$; 9.4 %), HPV52 ($n = 244$; 8.4 %), HPV18 ($n = 201$; 6.9 %), HPV56 ($n = 191$; 6.6 %), HPV66 ($n = 134$; 4.6 %), HPV31 ($n = 123$; 4.2 %), HPV59 ($n = 93$; 3.2 %), HPV68 ($n = 78$; 2.7 %), HPV51 ($n = 72$; 2.5 %), HPV35 ($n = 63$, 2.2 %), HPV53 ($n = 58$; 2.0 %), HPV45 ($n = 34$; 1.2 %), HPV73 ($n = 29$; 1.0 %), HPV39 ($n = 30$; 1.0 %), HPV83 ($n = 20$; 0.7 %), and MM4 ($n = 5$; 0.2 %). Low-risk genotypes included HPV6 ($n = 703$; 40.8 %), HPV11 ($n = 505$; 29.3 %), HPV43 ($n = 403$; 23.4 %), HPV42 ($n = 112$; 6.5 %), and HPV44 ($n = 1$; 0.1 %) [7]. Furthermore, 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 [7]. In addition to the analysis of the treatment and prevention in China, we also obtained data related to the DNA of HPV-33 and HPV-58.

Chen et al showed that HPV-33 manifested two viral lineages (A and B). The HPV-58 was classified into four variant lineages (A, B, C and D) [27]. In the 124 HPV-33 *L1*, our study samples matched with clade A, 111 samples in A1, and 13 samples in A2. All the 101 HPV-33 *L2* study samples matched clade A, 88 samples in A1, and 13 samples in A2. All the 223 HPV-58 *L1* study samples matched clade A, 135 samples in A1, 65 samples in A2, and 23 samples in A3. In case of 201 HPV-58 *L2*, all our study samples matched clade A, 114 samples in A1, 62 samples in A2, and 25 samples in A3. The sequence patterns of HPV-33 *L1* and *L2* were 13 and 8, and for HPV-58, they were 37 and 24, respectively.

We demonstrated that HPV-58 *L1* and *L2* variation frequencies were higher than those of HPV-33 *L1* and *L2*. Among these variations, C5807A, A5822G, G5984A, A6437G, T6470C, T6485C and A6695C (E377D, which is a positive selection variation) represented novel HPV-58 *L1* mutations, which were found until now only in

Sichuan, China [15, 17, 28–32]. In HPV-58 *L2*, mutations other than A4621C and A5206G, were newly reported [15]. In HPV-33 *L1*, T5960C, A5997G (K135R), T6385G, G6396A (G268E), T6463C, G6520A, T6613C, A6694G, A6951C (K453T), C7044A (P484H) and G7063A were reported for the first time, these newly reported mutations were only found in China in reports related to HPV-33 *L1* [30, 33, 34]. We reported the HPV-33 L2 mutations for the first time.

The most common non-synonymous mutations of HPV-33 *L1* included T266K (18/124), T56N (13/124), G133S (13/124) and E385D (13/124). The most common non-synonymous mutations of HPV-33 *L2* were D77N (15/101), F281Y (8/101), D350H (6/101), D360N (6/101), and N372T (6/101). The most common non-synonymous mutations of HPV-58 *L1* comprised L150F (88/223), I135M (88/223), and T375N (62/223). The most common non-synonymous mutations of HPV-58 *L2* were A416T (146/201), N231T (24/201), and M446L (24/201). Results of the selective pressure analysis suggest that most positive selection mutations of HPV-33 and HPV-58 *L1/L2* were common non-synonymous mutations, indicating that the positive selection variations beneficial for HPV-33 and HPV-58 adapted to their environments widely. The other variations in the positive selection mutations were uncommon. However, with the inheritance of the virus, positive variations may become increasingly common and gain importance in the future. The most common synonymous mutations of HPV-33 *L1* were A6664G (18/124), T6613C (14/124), T6478C (13/124) and A6673G (13/124). The most common synonymous mutations of the HPV-33 *L2* included C5220T (101/101) and C4905T (9/101). The most common synonymous mutations of HPV-58 *L1* were A6416G (65/223), T6434C (65/223), G6641A (65/223), and A6560G (64/223); and those of the HPV-58 *L2* were T4627C (148/201), G5395C (146/201), and G5266A (86/201). Due to the high diagnostic value of *L1*, and its variability [35], *L1* is often selected as a clinical diagnostic target. We considered the most sites in common mutations to design clinical diagnostic probes targeting HPV *L1* and

Table 8 Site-specific tests for positive selection on HPV-58 L1

Models	lnL	Estimates of parameters	2Δl	Positively selected sites
M7	−2849.987	$p = 0.005$ $q = 0.048$		NA
M8	−2821.953	$p0 = 0.995$ $p = 0.006$ $q = 0.096$ ($p1 = 0.005$) $\omega = 22.821$	56.068 $p < 0.01$	150 L**, 318 F, 325I**, 375 T**, 377E

Chen *et al. Virology Journal* (2016) 13:167

Page 10 of 12

Table 9 Site-specific tests for positive selection on HPV-58 L2

Models	lnL	Estimates of parameters	2Δl	Positively selected sites
M7	−2432.604	$p = 0.005$ q = 0.048		NA
M8	−2424.886	$p0 = 0.978$ $p = 2.63$ q = 99.000 ($p1 = 0.022$) ω = 5.255	15.436 $p < 0.01$	231 N*,416A*,446 M*

Note: Tables 6, 7, 8 and 9: ln L, the log-likelihood difference between the two models; 2Δl, 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 ne 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

L2 genes. The intratypic variations observed in *L1* and *L2* enabled the analysis of known and novel HPV subtypes [36, 37]. In our study, we observed that the sequence patterns and single nucleotide changes of HPV-33 *L2* were less frequent than those of HPV-33 *L1*. The sequence and single nucleotide changes of HPV-58 *L2* were less frequent than those of HPV-58 *L1*, while those of HPV-33 and HPV-58 *L2* were more conserved than those of *L1*, suggesting that HPV-33 and HPV-58 *L2* were better candidates as clinical diagnostic targets compared with HPV-33 and HPV-58 *L1*.

Nearly all conformational epitopes are located on one or more of the outwardly facing surface-exposed loops of BC, DE, EF, FG, and HI [38]. Sites 54, 55, 135–139, 141–143, 181, 182, 184, 267, 269, 270, 273, 278, 280, 282–287, 348, 354, 358, and 361 were previously reported to be important for the Fab interaction [38]. Although we found no mutations at these Fab interaction sites, we did find several mutations (T56N, K135R, T266K, and G268E of HPV-33 *L1*; N82T of HPV-58 *L1*) next to these sites. We believe mutations occurred on the outwardly facing surface-exposed loops
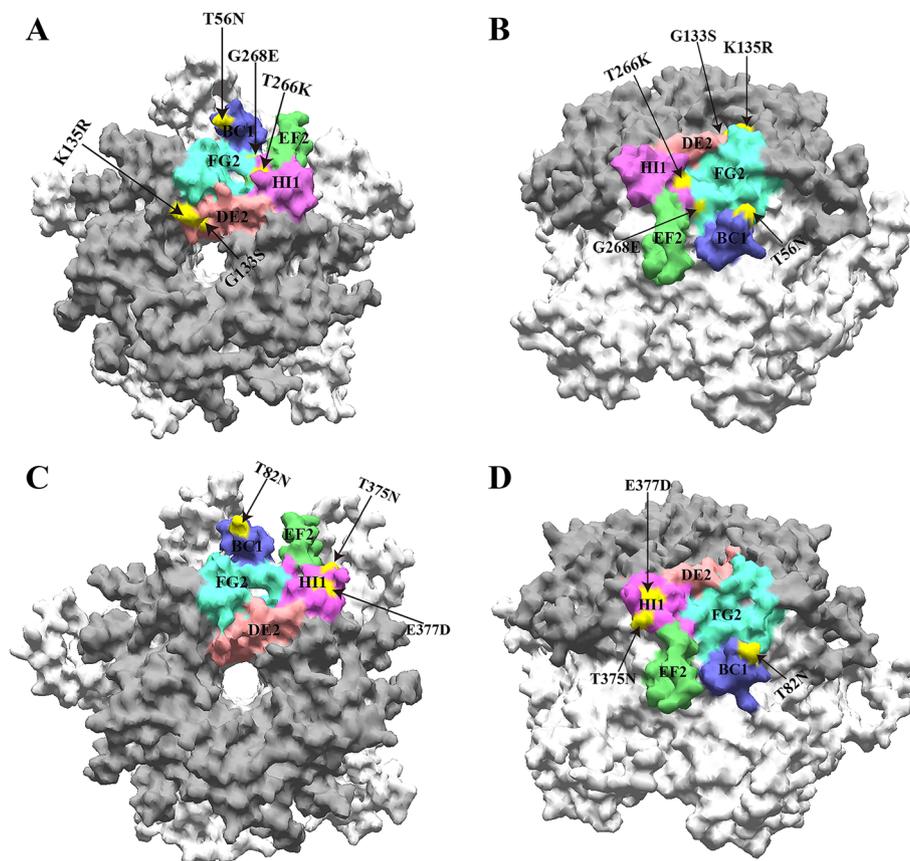
Fig. 3 Homology modeling of HPV-33 and HPV-58 variants. *Note*: Homology models based on the HPV-16 L1 pentamer crystal (PDB code: 2R5H). **a** top and (**b**) side surface-filled views shown for HPV-33 and (**c**) top and (**d**) side surface-filled views shown for HPV-58 with external loops indicated by dark grey shading. While all five monomers of the capsomer are pictured, adjacent external loops are indicated (purple, HI1; blue, BC1; red, DE2; cyan FG2; green, EF2), and amino acid mutation is shown with black arrows and marked with yellow color in one of the five copies of a residue

Chen *et al. Virology Journal* (2016) 13:167

Page 11 of 12

deserve research attention and should be studied for vaccine design targeting HPV-33 and 58 *L1*.

Amino acid residues 69–81 and 108–120 of L2 protein are highly conserved and contain cross-reacting epitopes that play an important role in inducing neutralizing antibodies [14]. G4438A (D77N) was discovered at residues 69–81 and 108–120 of HPV-33 *L2*. G4452A (G70E) and G4470A (S76N) were identified at residues 69–81 and 108–120 of HPV-58 *L2*. Amino acid residues 33–52, 73–84, 89–100, and 121–140 of *L2* contain non-neutralizing antibody epitopes [14]. G4438A (D77N) of HPV-33 *L2* and The G4470A (S76N) of HPV-58 *L2* were discovered at residues 73–84. These mutations must be considered during vaccine design targeting HPV-33 and 58 *L2*.

This is the first study examining the role of *L1/L2* proteins of HPV-58 variants in Sichuan and that of the *L1/L2* proteins of HPV-33 in China. Because of limitations related to sample size, sample copies, and sequencing technology, the present study may have had a sampling bias against integrated HPV genomes. The data presented in this study have significant implications for the understanding of intrinsic geographical and biological differences in HPV-33 and HPV-58 *L1/L2*, as well as contribute to the design of clinical diagnostic probes and second-generation polyvalent vaccine based on HPV-33 and HPV-58 *L1/L2*.

Conclusions

Mutations in HPV *L1* and *L2* may alter the virulence of variants, and also define altered epitopes in vaccine design. The reference sequences of HPV-33 and 58 only represent minor sequence patterns of HPV-33 and 58 *L1/L2*. Further, the distribution of HPV-33 and 58 *L1/L2* variations in Sichuan has its own peculiarities. Therefore, clinical diagnostic probes and second-generation polyvalent vaccines should be designed on the basis of the unique sequence of HPV-33 and 58 *L1/L2* variations in Sichuan, whereby the accuracy of clinical detection and the protective efficiency of vaccines can be improved.

Additional files

Additional file 1: Figure S1. Predicted of HPV-33 L1 and L2 proteins Secondary Structure by PSIPred. Note: A) Secondary structure within the reference sequence of HPV-33 L1 protein, B) Secondary structure within the reference sequence of HPV-33 L2 protein. Black arrow indicates corresponding mutation is a non-synonymous mutation. (TIF 948 kb)

Additional file 2: Figure S2. Predicted of HPV-58 L1 and L2 proteins secondary structure by PSIPred. Note: C) Secondary structure within the reference sequence of HPV-58 L1 protein, D) Secondary structure within the reference sequence of HPV-58 L2 protein. Black arrow indicates corresponding mutation is a non-synonymous mutation. (TIF 922 kb)

Additional file 3: Figure S3. Sequence alignment of L1 from the HPV types HPV16, HPV33, and HPV58. Note: The residues conserved across four HPV types are shown in capital letters, whereas the nonconserved residues are given in lowercase letters. The five loops displayed on the surface of the virus particle are marked and labeled. (TIF 1448 kb)

Authors' contributions
ZC and XD conceived and designed the study. ZC, YJ, QW, SZ, TW YZ and JZ performed the experiments. ZC and YJ analyzed the data. ZC and YJ wrote the paper.. All authors read and approved the final manuscript.

References

1. Bosch FX, Manos MM, Munoz 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(11):796–802.
2. Bosch FX, Lorincz A, Muñoz N, Meijer CJLM, Shah KV. The causal relation between human papillomavirus and cervical cancer. J Clin Pathol. 2002;55(4):244–65.
3. Walboomers JMM, V Jacobs M, 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(1):12–9.
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(2):e56614.
5. de Sanjosé S, Diaz M, Castellsagué X, Clifford G, Bruni L, Muñoz N, et al. Worldwide prevalence and genotype distribution of cervical human papillomavirus DNA in women with normal cytology: a meta-analysis. Lancet Infect Dis. 2007;7(7):453–9.
6. 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(2):105–9.
7. Chen Z, Wang Q, Ding X, Li Q, Zhong R, Ren H. Characteristics of HPV prevalence in Sichuan Province, China. Int J Gynecol Obstet. 2015;131(3):277–80.
8. 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(9):1635–40.
9. 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;448:356–62.
10. de Sanjose S, Quint WGV, Alemany L, Geraets DT, Klaustermeier JE, Lloveras B, et al. Human papillomavirus genotype attribution in invasive cervical

Chen *et al. Virology Journal* (2016) 13:167

Page 12 of 12

cancer: a retrospective cross-sectional worldwide study. Lancet Oncol. 2010;11(11):1048–56.

11. Lai HC, Sun CA, Yu MH, Chen HJ, Liu HS, Chu TY. Favorable clinical outcome of cervical cancers infected with human papilloma virus type 58 and related types. Int J Cancer. 1999;84:553–7.

12. Trus BL, Roden RB, Greenstone HL, Vrhel M, Schiller JT, Booy FP. Novel structural features of bovine papillomavirus capsid revealed by a three-dimensional reconstruction to 9 A resolution. Nat Struct Biol. 1997;4(5):413–20.

13. Buck CB, Cheng N, Thompson CD, Lowy DR, Steven AC, Schiller JT, et al. Arrangement of L2 within the papillomavirus capsid. J Virol. 2008;82(11):5190–7.

14. Karanam B, Jagu S, Huh WK, Roden RB. Developing vaccines against minor capsid antigen L2 to prevent papillomavirus infection. Immunol Cell Biol. 2009;87(4):287–99.

15. 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(1):e55204.

16. Muñoz N, Bosch FX, 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(6):518–27.

17. 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(21):13630–40.

18. Fleury MJ, Touze A, Coursaget P. Human papillomavirus type 16 pseudovirions with few point mutations in L1 major capsid protein FG loop could escape actual or future vaccination for potential use in gene therapy. Mol Biotechnol. 2014;56(5):479–86.

19. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol. 2013;30(12):2725–9.

20. Nei M, Gojoborit T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol. 1986;3:418–26.

21. Yamada T, Wheeler CM, Halpern AL, Stewart ACM, Hildesheim A, Jenison SA. 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;69:7743–53.

22. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol. 2007;24(8):1586–91.

23. 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(3):183–90.

24. Buchan DW, Minneci F, Nugent TC, Bryson K, Jones DT. Scalable web services for the PSIPRED Protein Analysis Workbench. Nucleic Acids Res. 2013;41(Web Server issue):W349–57.

25. Bishop B, Dasgupta J, Klein M, Garcea RL, Christensen ND, Zhao R, et al. Crystal structures of four types of human papillomavirus L1 capsid proteins: understanding the specificity of neutralizing monoclonal antibodies. J Biol Chem. 2007;282(43):31803–11.

26. Bordoli L, Kiefer F, Arnold K, Benkert P, Battey J, Schwede T. Protein structure homology modeling using SWISS-MODEL workspace. Nat Protoc. 2009;4(1):1–13.

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(5):e20183.

28. Gurgel AP, Chagas BS, do Amaral CM, Nascimento KC, Leal LR, Silva Neto Jda C, et al. Prevalence of human papillomavirus variants and genetic diversity in the L1 gene and long control region of HPV16, HPV31, and HPV58 found in North-East Brazil. Biomed Res Int. 2015;2015:130828.

29. 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(Pt 5):1229–37.

30. Raiol T, Wyant PS, de Amorim RM, Cerqueira DM, Milanezi N, Brigido Mde M, et al. Genetic variability and phylogeny of the high-risk HPV-31, -33, -35, -52, and -58 in central Brazil. J Med Virol. 2009;81(4):685–92.

31. Bae J-H. Distribution of human papillomavirus type 58 variants in progression of cervical dysplasia in Korean women. J Microbiol Biotechnol. 2009;19(9):1051–4.

32. 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(10):1752–61.

33. Cornut G, Gagnon S, Hankins C, Money D, Pourreaux K, Franco EL, et al. Polymorphism of the capsid L1 gene of human papillomavirus types 31, 33, and 35. J Med Virol. 2010;82(7):1168–78.

34. Ntova CK, Kottaridi C, Chranioti A, Spathis A, Kassanos D, Paraskevaidis E, et al. Genetic variability and phylogeny of high risk HPV type 16, 18, 31, 33 and 45 L1 gene in Greek women. Int J Mol Sci. 2012;13(1):1–17.

35. Pande S, Jain N, Prusty BK, Bhambhani S, Gupta S, Sharma R, et al. Human papillomavirus type 16 variant analysis of E6, E7, and L1 genes and long control region in biopsy samples from cervical cancer patients in north India. J Clin Microbiol. 2008;46(3):1060–6.

36. Stewart AC, Eriksson AM, Manos MM, Bosch FX, Peto J, Wheeler CM. Intratype variation in 12 human papillomavirus types: a worldwide perspective. J Virol. 1996;70(5):3127–36.

37. 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(1):70–9.

38. Guan J, Bywaters SM, Brendle SA, Lee H, Ashley RE, Makhov AM, et al. Structural comparison of four different antibodies interacting with human papillomavirus 16 and mechanisms of neutralization. Virology. 2015;483:253–63.