

Sequence Variation Analysis of HPV-18 Isolates in Southwest China

Mengjie Shen^{1,2}, Xianping Ding^{1,2*}, Tianjun Li^{1,2}, Gangyi Chen^{1,2}, Xiao Zhou¹

1 Bio-resource Research and Utilization Joint Key Laboratory of Sichuan and Chongqing, Institute of Medical Genetics, College of Life Science, Sichuan University, Chengdu, People's Republic of China, **2** Institute of Medical Genetics, College of Life Science, Sichuan University, Chengdu, People's Republic of China

Abstract

Intratypic variations of HPV-18 are known to differ in the persistence of the infection, frequency of carcinogenesis and the progression of precursor lesions to advanced cervical cancer. This study was designed to analyze sequence variations of HPV-18 isolates in order to discover novel HPV-18 variants and to evaluate the variations among infected women in southwest China. Cervical biopsies from 56 HPV-18-positive women with cervical neoplasia were assayed by PCR amplification and sequencing of all eight genes (E1, E2, E4, E5, E6, E7, L1, L2) of the HPV-18 genome. The most frequently observed variation was a C to G transversion at nucleotide 287 of E6, which was found in 48.2% of samples. Analysis of E7 revealed only one specimen as having sequence variations. In addition, we have identified several novel variations: A551C in E6, G6906A in L1, and C4915T and C5147A in L2. The mutations in E6 and L2 are silent, while the E7 mutation results in a single amino acid change. This study complements and expands on previous descriptions of HPV-18 variants. The sequence variation data presented here provides a foundation for future research on HPV-induced oncogenesis and may prove valuable for developing diagnostic probes and in the design of HPV vaccines for targeted populations.

Citation: Shen M, Ding X, Li T, Chen G, Zhou X (2013) Sequence Variation Analysis of HPV-18 Isolates in Southwest China. *PLoS ONE* 8(2): e56614. doi:10.1371/journal.pone.0056614

Editor: Michael Nevels, University of Regensburg, Germany

Received: June 25, 2012; **Accepted:** January 15, 2013; **Published:** February 25, 2013

Copyright: © 2013 Shen et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by funds from the Science and Technology Department of Sichuan Province, China. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

* E-mail: xianping2011@yeah.net

Introduction

Cervical cancer is the second leading cause of cancer-related deaths in women worldwide [1]. Every year, more than 500,000 new cases of cervical cancer and roughly 250,000 deaths are recorded [2]. Nearly all cervical cancers are caused by human papillomavirus (HPV) infection [1].

HPV is a double-stranded circular DNA virus with a genome size of about 8000 bp that encodes early proteins (E1, E2, E5, E6 and E7) and late proteins (L1, L2 and E4) [3,4,5]. As described by de Villiers et al. [6], the sequence of the highly conserved L1 ORF is used to classify HPV isolates. Isolates with a L1 sequence more than 10% different than the nearest HPV type are a distinct “type”, while isolates with differences of less than 2% are termed “variants.” There are 120 different HPV types identified to date [6,7], of which 15 have been termed “high-risk” types due to their significant association with cervical cancers [8]. There are 25 HPV types with strong, sufficient, or limited evidence of causing cervical cancer [9]. The types were classified and assessed by the IARC Monograph Working Group. Among them, the HPV-16 was the most oncogenic HPV type, known to cause cancer at 7 sites. HPV-18,31,33,35,45,52,58 were frequently found in cancer, while HPV-39,51,56,59 appeared less frequently [10]. Taking everything into account, HPV-16 and HPV-18 represent the most commonly identified high-risk HPV genotypes, which cause 40–60% and 10–20%, respectively, of all cervical cancers [11,12].

In addition to the diversity among types of HPV, there are many natural intratypic variants, some of which contain

alterations that lead to changes in the amino acid (AA) residues of functional and/or antigenic domains. Some of these changes have been shown to influence the persistence of the infection, morbidity of carcinogenesis, and the progression of precursor lesions to cancer [13,14,15]. At present, analyses of sequence variations of HPV-16, which occur in early genes, late genes, and the upstream regulatory region (URR or long control region, LCR), have been relatively comprehensive [16,17,18,19]. In contrast, data on HPV-18 variants is limited, and sequence variations of HPV-18 in the Asian population have been evaluated even less than those in European and American populations. The present study examined a collection of 56 different isolates of HPV-18 from cervical cancer patients in the southwest China and analyzed the sequence variations in the E1, E2, E4, E5, E6, E7, L1 and L2 viral genes. The data presented here is useful for future research on viral persistence, transmission and oncogenic potential, and may provide critical information for developing diagnostic probes and as well as designing vaccines for a specific population.

Materials and Methods

2.1 Ethics Statement

The study was approved by the Education and Research Committee and the Ethics Committee of Sichuan University (approval # SCU20100056494). Written informed consent was obtained from each patient, who granted us permission to use the data obtained in subsequent studies.

2.2 Specimens

Specimens were obtained from the cervical biopsies of 56 women with cervical cancer who tested positive for HPV-18 at the Cancer Hospital of Sichuan Province and the 4th People's Hospital of Chongqing. Medical background information was recorded from all patients.

2.3 Genomic DNA Extraction

DNA was extracted using a commercial kit (U-gene DNA kit) according to manufacturer's instructions (AnHui U-gene Biotechnology Co., Ltd. AnHui 242000, PR. China).

DNA was extracted and stored in a designated area free from potential DNA contaminants.

2.4 PCR

PCR reagents were mixed under a biosafety-hood in a designated room that was free from potential DNA contaminants. Amplification of the integrated HPV genes was performed with primers that were designed according to the HPV-18 reference sequence published in GenBank (accession number NC_001357) (Table 1).

The PCR amplification was performed under the following conditions. An initial 5 min denaturation step at 95°C was followed by 35 amplification cycles, with each cycle including a 45 s denaturation step at 94°C, a 45 s annealing step at 55°C, and a 60 s elongation step at 72°C. Amplification was complete after a final 10 min elongation step at 72°C. Each 50 µl PCR reaction contained 4 mmol/L MgCl₂, 400 µmol/L dNTPs, 3 U Taq DNA polymerase and 4 pmol primers.

Amplicons were visualized on 1.0% agarose gels stained with ethidium bromide under UV transillumination (data not shown).

2.5 Sequencing and Variations Analysis

PCR products were automatically sequenced using the Big Dye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) according to the manufacturer's instructions. The sequences were subsequently analyzed by NCBI Blast and DNAMAN version 5.2.2. HPV-18 DNA nucleotide positions were numbered according to the HPV-18 reference sequence (NC_001357).

All data were confirmed by repeating PCR amplification and sequence analysis at least twice.

Results

In this study, nucleotide (nt) sequences of E1 (nt 914 to 2887), E2 (nt 2817 to 3914), E4 (nt 3418 to 3684), E5 (nt 3936 to 4157), E6 (nt 105 to 581), E7 (nt 590 to 907), L1 (nt 5430 to 7136) and L2 (nt 4244 to 5632) were determined. Sequence variations were identified in 33 for the 56 specimens assayed (Table 2). In 23 of 56 (41.1%) specimens, the nucleotide sequences corresponded exactly to the HPV-18 reference sequence (NC_001357).

3.1 HPV-18 E1 Sequence Variations

As shown in Fig. 1A, DNA sequence analysis of the HPV-18 E1 region revealed the following four variations: a T to G transversion at nt2856 and a G to C transversion at nt2857 leading to a L648C AA substitution (n = 18, 32.1%), and a C to G transversion at nt2858 and a G to T transversion at nt2859 leading to a R649V AA change (n = 18, 32.1%). The four variations were shared in E1 and E2.

3.2 HPV-18 E2 Sequence Variations

DNA sequence analysis of the HPV-18 E2 region revealed seven variations: a T to G transversion at nt2856, G to C transversion at nt2857 and C to G transversion at nt2858, leading to a C14A AA substitution (n = 18, 32.1%); a G to T transversion at nt2859, leading to a V15L AA substitution (n = 18, 32.1%) (Fig. 1B); a C to G transversion at nt3084 and a G to C transversion at nt3085, leading to a R90A AA change (n = 11, 19.6%) (Fig. 1C); and a C to G transversion at nt3275, which does not lead to an AA change (n = 7, 12.5%) (Fig. 1D). Among these variations, T2856C, G2857C, C2858G and G2859T were simultaneous, as were C3084G and G3085C.

3.3 HPV-18 E6 Sequence Variations

DNA sequence analysis of the HPV-18 E6 region revealed two variations that did not lead to AA changes. The most frequent variation was a C to G transversion at nt287 (n = 27, 48.2%) (Fig. 1E), with a less frequent A to C transversion occurring at nt551 (n = 9, 16.1%) (Fig. 1F).

3.4 HPV-18 E7 Sequence Variations

DNA sequence analysis of the HPV-18 E7 region revealed two variations, both of which occurred only once and in the same specimen: a C to T transition at nt640, which did not result in an AA change (Fig. 1G), and an A to G transition at nt864, leading to a N92S AA substitution (n = 1, 1.79%) (Fig. 1H).

3.5 HPV-18 L1 Sequence Variations

DNA sequence analysis of the HPV-18 L1 region revealed six variations: a G to A transition at nt5503, leading to a R25Q AA substitution (n = 23, 41.1%) (Fig. 1I); a C to G transversion at nt5701 with an AA change of P91R (n = 23, 41.1%) (Fig. 1J); a C to G transversion at nt6460 with an AA change of P344R (n = 24, 42.9%) (Fig. 1K); a C to G transversion at nt6625 with an AA change of P399R (n = 25, 44.6%) (Fig. 1L); a C to G transversion at nt6842 with an AA change of P471R (n = 23, 41.1%) (Fig. 1M); and G to A transition at nt6906 with an AA change of D493N (n = 12, 21.4%) (Fig. 1N).

3.6 HPV-18 L2 Sequence Variations

DNA sequence analysis of HPV-18 L2 region revealed two variations that did not lead to AA changes: a C to T transition at nt4915 (n = 11, 19.6%) (Fig. 1O) and a C to A transversion at nt5147 (n = 10, 17.9%) (Fig. 1P).

3.7 HPV-18 E4 and E5 Sequence Variations

No variations were detected in the HPV-18 E4 and E5 regions.

Discussion

A study by Angulo et al. [20] highlighted the fact that co-infection with more than one HPV type is not a rare occurrence, and this situation can lead to HPV recombination and the generation of new HPV types. Therefore, it is of the utmost importance to gather information on HPV variants and analyze different genomic regions and cases of multiple infections for developing HPV diagnostics, vaccines and other therapeutic approaches to manage viral-induced cancer [16]. HPV-18 variants have been shown to have different biological as well as biochemical effects, as some variants can result in greater transforming efficiency, more aggressive clinical outcome, higher probability of cancer recurrence and worse prognosis [21,22,23,24]. In this study, we described the HPV-18 intratype diversity over a span of

Table 1. Primers used for PCR.

HPV-18 region	primer	Primer position	primer nucleotide sequence	Product size (bp)
E1	E1-1F	867	5'- CCCTGTCCTTTGTGTGCCG-3'	612
	E1-1R	1478	5'-TATTGCTATTGTCACTGTACCGTC-3'	
	E1-2F	1436	5'- GGCAACAACAGCAGTGTAGACGGTA-3'	661
	E1-2R	2096	5'-TGCTGTTGCTGTCTGCTAATAAGGC-3'	
	E1-3F	2038	5'- GCTGACAGATGAAAGCGATATG -3'	433
	E1-3R	2470	5'-CGGTTCCAACCAAAAATGACTAGTG-3'	
	E1-4F	2361	5'- GTGGACCAGCAAATACAGGAAAATC -3'	561
	E1-4R	2921	5'-CTGTATTGGCTGTCTATGTC-3'	
E2	E2-1F	2814	5'-AGATGCAGACACCGAAGGAAAC-3'	632
	E2-1R	3445	5'-TCGTCCTGGTACTGCACATAGA-3'	
	E2-2F	3420	5'- GACTCTATGTGCAGTACCAGTGACG-3'	599
	E2-2R	4018	5'-ATACAGACAGATGGCAAAAGCGGGA-3'	
E4	E4F	3296	5'-GGGATTGTATTATGTAAGGAAGGG-3'	456
	E4R	3751	5'-TGGTGCTATGTTTTCGCAATCTGT-3'	
E5	E5F	3885	5'- CAAATTTGGTGGGATACATGAC-3'	380
	E5R	4264	5'- TGCGGCACGGTGGGATACCATACTT-3'	
E6	E6F	85	5'- AGAAACACACCACAATACTATGGCG -3'	662
	E6R	746	5'-GTCGGGCTGGTAAATGTTGAT-3'	
E7	E7F	540	5'- CGACAGGAACGACTCCAACGA-3'	431
	E7R	970	5'- ATAAAACACGCCGTTACAACCCGTG-3'	
L1 ^a	L1-1F	5402	5'- GTAACGGTCCCTTTAACTCCTC-3'	650
	L1-1R	6051	5'- 5'- CATTGTCCTAACGTCTCAG-3'	
	L1-2F	6002	5'- AAGTCCCATGCCGCCAGTCTAAT-3'	505
	L1-2R	6506	5'- AGAGCCACTGGAGAGGGAGAATAC-3'	
	L1-3F	6502	5'- GCTCTATTGTTACTCTGACTCC-3'	636
	L1-3R	7137	5'- ATTACTTCTGGCACGTACACGCAC-3'	
L2	L2-1F	4239	5'- AAAGTATGGTATCCCACCGTGCCGC-3'	518
	L2-1R	4756	5'-TGGAACCTCAATAATGGACGGA-3'	
	L2-2F	4755	5'- CACAACTGGGGAGGTGGCAGGTAA-3'	527
	L2-2R	5281	5'-GTCATTGTCTCCGTGGCAGATACT-3'	
	L2-3F	5259	5'- TATCTGCCACGGAGGACAATGACTT-3'	398
	L2-3R	5656	5'-CTGTTATGGCATATAGAAGGTGG-3'	

Because of the gene is too long, E1, E2, L1 and L2 were divided into some subunits to design primers for future sequencing and variation analysis. F, upstream primer; R, downstream primer.

doi:10.1371/journal.pone.0056614.t001

7000 nucleotides among isolates from southwest China. Our data set on E1, E2, E4, E5, E6, E7, L1 and L2 sequence variations of HPV-18 complements and expands on previous descriptions of HPV-18 variants, which were based on a targeted analysis of E6, LCR-E6, E2, L1 and URR [23,25,26].

Among the early gene products, the E1 viral protein plays a critical role in controlling viral replication and load, and it requires the interaction with the E2 protein to bind to the LCR [27]. Whereas E1 viral helicase is the replication initiator protein, E2 is the major viral regulator of viral transcription and replication, repressing transcription of the E6/E7 oncogenes, and together with the E1 helicase, activating viral DNA replication [28].

The E2 protein contains three distinct domains: a transactivation domain (TAD), an unstructured hinge domain, and a DNA binding domain (DBD) [28,29]. We discovered that all of the E2 sequence variations in our set of samples are located in the TAD

(2856–2859) and the unstructured hinge domain (3084, 3085 and 3275). None were found in the DBD domain, which is not unexpected given that the sequence of the DBD is highly conserved due to its role in binding a specific consensus sequence (ACCN₆GGT) [28,30,31] in the LCR of the viral genome.

Mutagenesis analysis of the E2 TAD indicated that most of the conserved AAs in this domain are required for both transcriptional activation and viral DNA replication [25,32]. Moreover, Cooper et al. [33] proposed that the N-terminus of E2 is tightly configured, as single AA substitutions have been found to impair its ability to support replication. Our study identified four variations in the TAD domain at the N-terminus. These changes resulted in two AA changes: C14A and V15L. Both the C14A and V15L amino acid substitutions occur in the first amino-terminal α -helix of the TAD domain [34]. In a study by Harris and Botchan, both structural and mutagenesis data indicate that this region is more important for HPV transcription than replication [35]. To

Table 2. Sequence variations of HPV-18.

regions	E6		E7		E1	E2		E2			L2		L1						
	287	551	640	864	2856	2857	2858	2859	3084	3085	3275	4915	5147	5503	5701	6460	6625	6842	6906
reference	C	A	C	A	T	G	C	G	C	G	C	C	C	G	C	C	C	C	G
variation	G	C	T	G	G	C	G	T	G	C	G	T	A	A	G	G	G	G	A
Amino acid sequence	P	R	P	N92S	L648C	C14A	R649V	V15L	R 9 0 A	T	P	R	R25Q	P91R	P344R	P399R	P		D493N
cc-03	G													A	G	G	G	G	
cc-05	G													A	G	G	G	G	A
cc-06	G													A	G				A
cc-07	G				G	C	G	T	G	C		T	A	A	G	G	G	G	
cc-10	G	C			G	C	G	T	G	C	G		A			G	G	G	A
cc-12												T		A	G	G	G	G	
cc-15	G													A	G	G	G		A
cc-16					G	C	G	T				T	A				G	G	
cc-18	G	C			G	C	G	T	G	C	G	T	A			G			
cc-19					G	C	G	T						A	G	G	G	G	
cc-20					G	C	G	T	G	C									
cc-21	G													A	G	G	G	G	A
cc-22	G								G	C	G		-	A	G	G	G	G	
cc-25	G		T	G										A	G		G	G	
cc-26	G	C			G	C	G	T	G	C	G	T		A	G	G			
cc-27	G				G	C	G	T	G	C	G	T		A	G	G			
cc-29	G	C			G	C	G	T					A				G	G	
cc-30	G													A	G	G	G	G	A
cc-33					G	C	G	T				T	A			G	G		
cc-35	G	C												A	G				
cc-36	G	C			G	C	G	T	G	C	G	T	A	A	G		G	G	A
cc-37					G	C	G	T	G	C	G			A	G	G	G	G	A
cc-39	G								G	C						G			
cc-43	G													A	G	G	G	G	
cc-44	G				G	C	G	T				T	A				G	G	
cc-45	G				G	C	G	T	G	C				A	G	G	G	G	A
cc-46	G				G	C	G	T											
cc-47	G													A	G	G	G	G	
cc-48	G	C			G	C	G	T							G	G	G	G	A
cc-49	G	C												A	G	G	G	G	A
cc-52	G										T	A	A	A	G	G	G		
cc-53	G				G	C	G	T			T	A	A	G	G	G	G	G	A
cc-56	G	C			G	C	G	T						A	G	G	G	G	

Sequence variations of HPV-18 E1, E2, E6, E7, L1 and L2 were based on the HPV-18 reference sequence (NC_001357). The identification codes of the specimens are indicated along the far left column. The number and the letter represent the nucleotide variation at that position. Blanks indicate no variation. In the 'reference' row, the letters represent the nucleotides in the reference sequence. In the 'amino acid sequence' row, the number and the letter represent the AA change at that position. doi:10.1371/journal.pone.0056614.t002

the best of our knowledge, no functional assignment of these residues has been made at this time.

E1 shares the sequence variations at nt2856–2859 with E2 in TAD, through which E2 interacts with the E1 helicase [5]. These variations result in L648C and R649V AA substitutions in the E1 coding region, especially, in the C-terminal helicase/ATPase region, which unwinds the HPV genome to allow for replication [36]. The importance of these residues to the ATP-dependent DNA helicase activity is unknown at this time, and elucidating

functional changes brought about by these substitutions is an important area of future investigation.

E6 and E7 encode the main viral proteins involved in human epithelial cell immortalization and transformation. The oncoproteins are able to increase the expression of genes related to DNA synthesis and coordinate cell cycle induction, thereby increasing the productive life cycle of HPV [37]. The E6 protein induces cell proliferation by interacting with numerous cellular proteins that are involved with tumor suppression, including p53 [38,39]. In our data set, the most frequent variation was the C to G silent

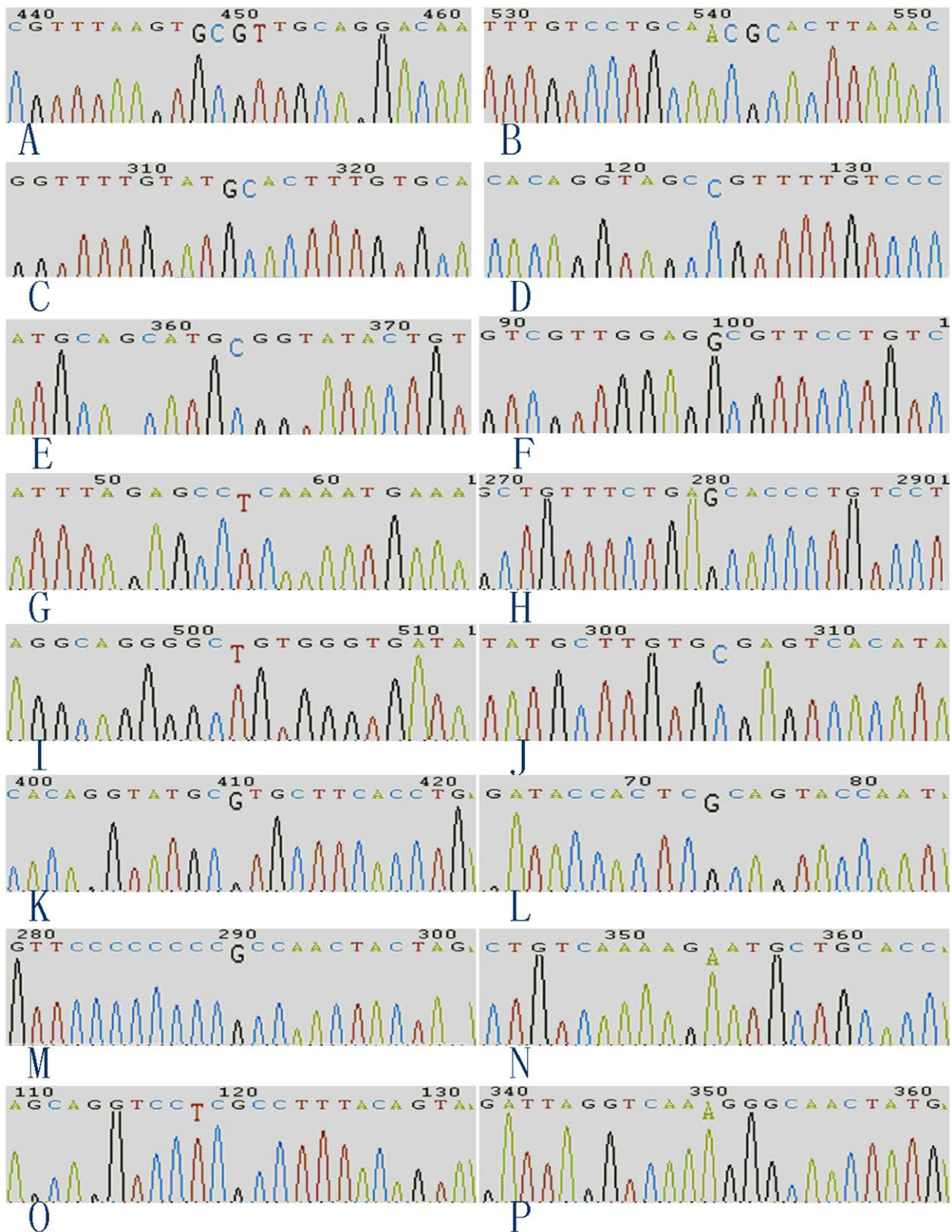


Figure 1. Chromatograms of sequence variations. The sequence variations were enlarged to facilitate observation. In B, D, E, F, I and J, the reverse complement of the sequence variations are displayed because the sequences were obtained using downstream primers. G, guanine; C, cytosine; A, adenine; and T, thymine.
doi:10.1371/journal.pone.0056614.g001

mutation at nt287. Furthermore, the guanine at E6 coding region nt183 is represented as HPV-18 prototype (AY262282). These results are in good agreement with a study by Chang et al. [26] in Taiwan, who reported the identical nucleotide variation, and suggests that the C to G variation at nt287 may be a common mutation worldwide. Since this frequent variant is a silent mutation, it may be used as a biomarker to track and study viral transmission [26].

The E7 protein interacts with another cellular tumor suppressor, pRb, and disrupts its normal function in regulating the cell cycle, thereby stimulating expression of many E2F-regulated transcripts [38]. E7 encodes a protein of about 100 AAs and contains three conserved regions: CD1, CD2 and CD3 [38]. Our sequencing results showed that HPV-18 E7 is highly conserved, with only one specimen detected as having sequence variations at nt640 and nt864; these variations were also detected in all African variants by Arroyo et al. [40]. The transition C to T at nt640 is a silent mutation, and A to G transition at nt864, leading to a N92S AA substitution. The N92S AA substitution is in the region linking CD2 and CD3. There is an optimal zinc-sensitive tertiary structure of HPV-18 E7, which modulates its interaction with target cellular proteins [41]. Indeed, the E7 domain of HPV-16 and HPV-18 is highly conserved [15,16,17,38], which strongly implies that it encodes the major transforming function [42,43], and that the conserved domains are critical for the transforming activities of each of these viral oncoproteins [44,45]. A recent study by Todorovic and colleagues suggests that the N92S substitution may have functional significance [46]. This residue was predicted to be exposed on the surface of the CR3 domain of E7 protein, and its mutation in HPV-16 (P92A) led to reduced dimerization of E7 (according to yeast two-hybrid analysis) as well as increased transformation potential. Thus, it is possible that this variant increases the incidence and severity of cervical cancer in infected women.

Changes in the L1/L2 regions of the HPV genome may be important in determining the infectious potential of different variants, as well as in defining epitopes relevant to vaccine design [47,48]. The L1 protein is the main component of the viral icosahedron capsid and is capable of self-assembly, thus giving rise to virus-like particles [49]. The analysis of the L1 gene is of immense importance because of its high diagnostic value [16], as the range of intratypic variations observed in this region allows the distinction and assessment of known or novel HPV types [7,50].

In the L1 coding region, one A to G transition at nt5503 and four C to G transversions were found. These five variations were also observed in a study by Arias-Pulido et al. [25] in the United

States. Furthermore, the four guanine variations are represented in the HPV-18 prototype sequence (AY262282), indicating that these variations are wide-spread. In addition, we also found a novel G to A transition at nt6906. All five sequence variations near C-terminus described above result in AA changes which may affect immune responses to HPV-18 capsid protein [25,47]. The HPV L1 model postulates that the L1 C-terminal domain is exposed on the viral surface and expected to be highly antigenic [51].

Like L1, the L2 protein plays an important role in the initial steps of papillomavirus infection [52]. We found two sequence variations in the L2 ORF that did not result in AA changes. These changes could reflect the extreme rarity of the reference sequence that was originally obtained from a subject in southwest China.

Interestingly, no sequence variation was found in E4 and E5, in contrast to a report by Arias-Pulido et al. [25] in the United States, in which there were seven sequence variations found in E4. This discrepancy may be attributed to differences in sample size as well as the geographic and racial characteristics of the study populations.

Conclusions

In summary, this study reports sequence variations in the E1, E2, E6, E7, L1 and L2 genes of HPV-18 isolates from southwest China. The sequence variations in the genes examined may contribute to HPV oncogenesis. Our data will provide a solid foundation for further biological and clinical studies, and may also be employed in epidemiological studies where sequence variations are used as markers for monitoring HPV infections in target populations. Elucidating the functional and pathological differences between intratypic variants of HPV, as well as determining the molecular basis for their propensity to induce different histological types of cervical cancer, are important areas of future investigation. Further studies of HPV sequence variation will also need to include different geographic regions with distinct and isolated human populations.

Acknowledgments

We thank the Cancer Hospital of Sichuan Province and the 4th People's Hospital of Chongqing for providing specimens.

Author Contributions

Conceived and designed the experiments: MS XD. Performed the experiments: MS TL. Analyzed the data: MS GC. Contributed reagents/materials/analysis tools: XZ.

References

1. Thun MJ, DeLancey JO, Center MM, Jemal A, Ward EM (2010) The global burden of cancer: priorities for prevention. *Carcinogenesis* 31(1): 100–110.
2. Paavonen J (2007) Human papillomavirus infection and the development of cervical cancer and related genital neoplasias. *International Journal of Infectious Diseases*. (Suppl 2): S3–S9.
3. Garcia-Vallve S, Alonso A, Bravo IG (2005) Papillomaviruses: different genes have different histories. *Trends Microbiol* 13(11): 514–521.
4. Munger K, Howley PM (2002) Human papillomavirus immortalization and transformation functions. *Virus Res* 89(2): 213–228.
5. Wu MH, Chan JYH, Liu PY, Liu ST, Huang SM (2007) Human papillomavirus E2 protein associates with nuclear receptors to stimulate nuclear receptor- and E2-dependent transcriptional activations in human cervical carcinoma cells. *Int. J. Biochem. Cell B* 39(2): 413–425.
6. de Villiers EM, Fauquet C, Broker TR, Bernard HS, Hausen H (2004) Classification of papillomaviruses. *Virology* 324(1): 17–27.
7. Bernard HU, Burk RD, Chen Z, van Doorslaer K, Hausen H, et al. (2010) Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology* 401(1): 70–79.
8. Munoz N, Bosch FX, Sanjose S (2003) Epidemiologic classification of human papillomavirus types associated with cervical cancer. *New Engl J Med* 348(6): 518–527.
9. Bouvard V, Baan R, Straif K, Grosse Y, Secretan B, et al. (2009) A review of human carcinogens—Part B: biological agents. *The Lancet Oncology* 10: 321–322.
10. Monographs on the Evaluation of Carcinogenic Risks to Humans. Volume 100. A Review of Human Carcinogens. Part B: Biological Agents, IARC 2011.
11. Bosch FX, Lorincz A, Munoz N, Meijer CJLM, Shah KV (2002) The causal relation between human papillomavirus and cervical cancer. *J Clin. Pathol* 55(4): 244–265.
12. Clifford GM, Smith JS, Plummer M, Munoz N, Franceschi S (2003) Human papillomavirus types in invasive cervical worldwide: a meta-analysis. *Brit J Cancer* 88(1): 63–73.
13. Cai HB, Chen CC, Ding XH (2010) Human papillomavirus type 16 E6 gene variations in Chinese population. *EJSO* 36(2): 160–163.
14. Lonsdaleborough P, Ho L, Terry G, Cuzick J, Wheeler C, et al. (1996) Human papillomavirus genotype as a predictor of persistence and development of high-grade lesions in women with minor cervical abnormalities. *Int J Cancer* 69(5): 364–368.

15. Zehbe I, Wilander E, Delius H, Tommasino M (1998) Human papillomavirus 16 E6 variants are more prevalent in invasive cervical carcinoma than the prototype. *Cancer Res* 58(4): 829–833.
16. Pande S, Jain N, Prusty BK, Bhambhani S, Gupta S, et al. (2008) Human papillomavirus type 16 variant analysis of E6, E7, and L1 genes and long control region in biopsy samples from cervical cancer patients in north India. *J Clin Microbiol* 146(3): 1060–1066.
17. Sathish N, Abraham P, Reddicayil A, Sridharan G, Chandy G (2005) HPV 16 E6 sequence variations in Indian patients with cervical neoplasia. *Cancer Lett* 229(1): 93–99.
18. Stephen AL, Thompson CH, Tattersall MH, Cossart YE, Rose BR (2000) Analysis of mutations in the URR and E6/E7 oncogenes of HPV 16 cervical cancer isolates from central China. *Int J Cancer* 86(5): 695–701.
19. Wu YP, Chen YL, Li LY, Yu GF, He Y, et al. (2006) Analysis of mutations in the E6/E7 oncogenes and L1 gene of human papillomavirus 16 cervical cancer isolates from China. *J Gen Virol* 87(5): 1181–1188.
20. Angulo M, Carvajal-Rodríguez A (2007) Evidence of recombination within human alpha-papillomavirus. *Virol J* 4: 33.
21. Yumin W, Jie C, Wenhui Z, Wangdong H, Fangyou Y (2012) Study of the prevalence of human Papillomavirus infection in Chinese women with cervical cancer. *African Journal of Microbiology Research* 6(5): 1048–1053.
22. Cullen AP, Reid R, Campio M, Lorincz AT (1991) Analysis of the physical state of different human papillomavirus DNAs in intraepithelial and invasive cervical neoplasm. *J Virol* 65(2): 606–612.
23. Rose BR, Thompson CH, Zhang JW, Stoeter M, Stephen A, et al. (1997) Sequence variation in the upstream regulatory region of HPV 18 isolates from cervical cancers. *Gynecol Oncol* 66(2): 282–289.
24. Walker J, Bloss JD, Liao SY (1989) Human papillomavirus genotype as a prognostic indicator in carcinoma of the uterine cervix. *Obstet Gynecol* 74(5): 781–785.
25. Arias-Pulido H, Peyton CL, Torrez-Martinez N, Anderson DN, Wheeler CM (2005) Human papillomavirus type 18 variant lineages in United States populations characterized by sequence analysis of LCR-E6, E2, and L1 regions. *Virology* 338(1): 22–34.
26. Chang CH, Chen TH, Hsu RC, Chou PH, Yang JJ, et al. (2005) The prevalence of HPV-18 and variants of E6 gene isolated from cervical cancer patients in Taiwan. *J Clin Virol* 32(1): 33–37.
27. Amador-Molina A, Gonzalez-Montoya JL, Garcia-Carranca A, Mohar A, Lizano M (2012) Intratype changes of the E1 gene and the Long Control Region affect ori function of HPV type 18 variants. *J Gen Virol*.
28. Bellanger S, Tan CL, Xue YZ, Teissier S, Thierry F (2011) Tumor suppressor or oncogene? A critical role of the human papillomavirus (HPV) E2 protein in cervical cancer progression. *J Cancer Res Am* 1(3): 373–389.
29. Giri I, Yaniv M (1988) Structural and mutational analysis of E2 trans-activating proteins of papillomaviruses reveals three distinct functional domains. *EMBO J* 7(9): 2823–2929.
30. Androphy EJ, Lowy DR, Schiller J (1987) Bovine papillomavirus E2 trans-activating gene product binds to specific sites in papillomavirus DNA. *Nature* 325(1): 70–73.
31. Desaintes C, Demeret C (1996) Control of papillomavirus DNA replication and transcription. *Cancer Biol Semin* 7(6): 339–347.
32. Demeret C, Goyat S, Yaniv M, Thierry F (1998) The human papillomavirus type 18 (HPV18) replication protein E1 is a transcriptional activator when interacting with HPV18 E2. *Virology* 242(2): 378–386.
33. Cooper CS, Upmeyer SN, Winokur PL (1998) Identification of single amino acids in the human papillomavirus 11 E2 protein critical for the transactivation or replication functions. *Virology* 241(2): 312–322.
34. Wang Y, Coulombe R, Cameron DR, Thauvette L, Massariol M, et al. (2004) Crystal Structure of the E2 Transactivation Domain of Human Papillomavirus Type 11 Bound to a Protein Interaction Inhibitor. *Journal of Biological Chemistry* 279(8): 6976–6985.
35. Harris SF, Botchan MR (1999) Crystal structure of the human papillomavirus type 18 E2 activation domain. *Science* 284(5420): 1673–1677.
36. Abbate EA, Berger JM, Botchan MR (2004) The X-ray structure of the papillomavirus helicase in complex with its molecular matchmaker E2. *Genes and Development* 18(16): 1981–1996.
37. Frago-Ontiveros V, Alvarez-Garcia RM, Contreas-Paredes A, Vaca-Paniagua F, Herrera LA, et al. (2012) Gene expression profiles induced by E6 from non-European HPV18 variants reveals a differential activation on cellular processes driving to carcinogenesis. *Virology* 432(1): 81–90.
38. Yugawa T, Kiyono T (2009) Molecular mechanisms of cervical carcinogenesis by high-risk human papillomaviruses: novel functions of E6 and E7 oncoproteins. *Reviews in Medical Virology* 19(2): 97–113.
39. Blachon S, Demeret C (2003) The regulatory E2 proteins of human genital papillomaviruses are pro-apoptotic. *Biochimie* 85(8): 813–819.
40. Arroyo SL, Basaras M, Arrese E, Hernaez S, Andia D, et al. (2012) Human Papillomavirus (HPV) genotype 18 variants in patients with clinical manifestations of HPV related infections in Bilbao. *Virol J Spain* 9: 258.
41. Liu SQ, Tian YX, Greenaway FT, Sun MZ (2010) A C-terminal hydrophobic, solvent-protected core and a flexible N-terminus are potentially required for human papillomavirus 18 E7 protein functionality. *Biochimie* 92(7): 901–908.
42. Phelps WC, Yee CL, Munger K, Howley PM (1988) The human papillomavirus type 16 E7 gene encodes transactivation and transformation functions similar to those of adenovirus E1A. *Cell* 53(4): 539–547.
43. Yatsudo M, Okamoto Y, Hakura A (1988) Functional dissociation of transforming genes of human papillomavirus type 16. *Virology* 166(2): 594–597.
44. Dyson N, Howley PM, Munger K (1989) The human papillomavirus-16 E7 oncoprotein is able to bind the retinoblastoma gene product. *Science* 243: 934–937.
45. Whyte P, Buchkovich KJ, Horowitz JM, Friend SH, Raybuck M, et al. (1988) Association between an oncogene and an anti-oncogene: the adenovirus E1A proteins bind to the retinoblastoma gene product. *Nature* 334(6178): 124–129.
46. Todorovic B, Massimi P, Hung K, Shaw GS, Banks L (2011) Systematic analysis of the amino acid residues of human papillomavirus type 16 E7 conserved region 3 involved in dimerization and transformation. *J virol* 85(19): 10048–10057.
47. Frati E, Bianchi S, Colzani D, Zappa A, Orlando G, et al. (2011) Genetic variability in the major capsid L1 protein of human papillomavirus type 16 (HPV-16) and 18 (HPV-18). *Infect Genet Evol* 11(8): 2119–2124.
48. Villa LL, Sichero L, Rahal P, Caballero O, Ferenczy A, et al. (2000) Molecular variants of human papillomavirus types 16 and 18 preferentially associated with cervical neoplasia. *J Gen Virol* 81(12): 2959–2968.
49. Pfister H (2003) Human papillomavirus and skin cancer in: Kramer, B.S. (Ed.), *J Natl Cancer Inst Monogr Oxford* 52–56.
50. Stewart AC, Eriksson AM, Manos MM, Bosch FX, Peto J, et al. (1996) Intra-type variation in 12 human papillomavirus types: a worldwide perspective. *J Virol* 70(5): 3127–3136.
51. Modis Y, Trus BL, Harrison SC (2002) Atomic model of the papillomavirus capsid. *EMBO J* 21(8): 4754–4762.
52. Rhoads CK, Cynthia DT, Thompson DL, Lowy JS, Patricia MD (2009) The initial steps leading to papillomavirus infection occur on the basement membrane prior to cell surface binding. *J PNAS* 106(48): 20458–20463.