RESEARCH ARTICLE

Editorial Process: Submission:01/20/2022 Acceptance:06/17/2022

Human Papillomavirus Type 16 *L2* Gene Sequence Variation Analysis in Indonesian Cervical Cancer Specimens

Syandrez Prima Putra[1,2], Andani Eka Putra[2]*, Indwiani Astuti[3], Widya Asmara[4]

Abstract

Background: Human papillomavirus type 16 (HPV16) is the most prevalent etiology of cervical cancer in Indonesian women. The L2 minor capsid protein has considerable potential as a broad-protective antigen target of the cervical cancer vaccine strategies, yet the data on *L2* gene variation is still minimal. In this research, we determined the variations of the *HPV16 L2* gene sequences in Indonesian cervical cancer specimens. **Methods:** We cross-sectionally observed 23 DNA isolates of HPV16 positive cervical cancer specimens stored in the laboratory of the Center for Diagnostic and Research on Infectious Diseases (PDRPI Lab), Faculty of Medicine, Universitas Andalas, Padang, Indonesia. We detected and amplified the *HPV16 L2* gene sequences in the samples, followed by sequencing, DNA alignment, single nucleotide polymorphisms (SNPs) analysis, and phylogenetic tree reconstruction. **Results:** As many as 35 SNPs were found, consist of 18 synonymous SNPs (sSNPs) and 17 non-synonymous SNPs (nsSNPs). Amino acid variations were mostly detected at S269P (100%) and L330F (43.48%) with no variation in the immuno-protective region near L2 N-terminus. A total of 5 HPV16 phylogenetic sub-lineages were found closely related to A1 (n=5), A2 (n=12), A3 (n=2), A4 (n=3), and C (n=1). **Conclusion:** The variations of *HPV16 L2* gene sequences are mainly located in the central region of the L2 sequences, and the cross-protective region near the L2 N-terminus is remarkably conserved. This study should enhance the information about *HPV16 L2* gene variation in Indonesia.

Keywords: Human papillomavirus- HPV16- *L2* gene- genetic variation- phylogeny

Asian Pac J Cancer Prev, **23 (6)**, 2009-2016

Introduction

Cervical cancer is one of the leading causes of morbidity and mortality in women (Torre et al., 2016). The incidence and the death rates of cervical cancer reached 15.1 and 8.2, respectively, per 100,000 global population in 2018, the third-highest of all cancer cases (Bruni et al., 2019). At the same time, Indonesia recorded 32,469 new cases and 18,279 deaths annually and had the highest incidence rates among South-East Asia countries (23.4 per 100,000 women) (Bruni et al., 2018). The etiology of the disease is human papillomavirus (HPV) infection, transmitted through sexual contact (Tommasino, 2014; Oyervides-Muñoz et al., 2018). HPV can be found in 8.6% of women with cancers globally (de Martel et al., 2017). The infection has remained a global health problem that affects more than 24% population in many countries, including Indonesia, especially in rural areas (Sabeena et al., 2017; Serrano et al., 2018; Bruni et al., 2019).

HPV belongs to the Papillomaviridae family that infects the skin and mucosal epithelial tissue. This virus is a circular double-stranded DNA virus with a genome of around 8,000 bp and consists of early (E), late (L), and non-coding (NCR) regions (Araldi et al., 2018). HPV carcinogenicity is mainly due to the activity of E5, E6, and E7 proteins in the tumor suppressor genes *p53* and *pRb* of host cells (Faridi et al., 2011). However, the L2 protein is crucial in viral entry into the host cell, avoiding the immune system and integrating viral DNA into the host genome. L2 is the HPV minor capsid protein with around 67-78 kDa molecular weight and approximately 500 amino acids. This protein is also necessary to release the virus from the endosomal vesicles after entering the host cell, transporting the viral genome into the nucleus, and assembling viral components (Wang and Roden, 2013). Nevertheless, some studies have shown that the conserved region near the N-terminus of L2 protein may induce cross-neutralizing antibodies against highly divergent HPV types and can be single-expressed in bacteria. (Karanam et al., 2009; Schellenbacher et al., 2017; Olczak and Roden, 2020). Thus, this finding proposed the HPV L2 become a potential antigen candidate for the broad

*[1]Master Program of Tropical Medicine, Faculty of Medicine, Public Health, and Nursing, Universitas Gadjah Mada, Yogyakarta, Indonesia. [2]Department of Microbiology, Faculty of Medicine, Universitas Andalas, Padang, Indonesia. [3]Department of Pharmacology and Therapy, Faculty of Medicine, Public Health, and Nursing, Universitas Gadjah Mada, Yogyakarta, Indonesia. [4]Department of Microbiology, Faculty of Veterinary Medicine, Universitas Gadjah Mada, Yogyakarta, Indonesia. *For Correspondence: andani_ep@yahoo.com*

HPV vaccine.

Human papillomavirus type 16 (HPV16) is the most carcinogenic type of HPV (Johnson et al., 2019). A large retrospective study in Iran revealed that HPV16 infection is dominant in most patients with high-risk HPV types (30.5%), followed by HPV53 (17.3%) and HPV39 (13.3%) (Chalabiani et al., 2017). In Indonesia, HPV16 infection is higher, found in 47% of cervical cancer cases, and may worsen the quality of life and decrease one's five-year life expectancy by 66.9% (Bruni et al., 2018; Hallowell et al., 2018; Setiawan et al., 2018). This HPV type belongs to the Alphapapilloma genus and alpha-9 species group. Based on phylogenetic analysis, HPV16 consists of four lineages (A, B, C, D lineage) and ten sub-lineages, namely: A1-A3 (European/ E); A4 (Asian/ As); B1 (African-1a/ Afr-1a) and B2 (African-1b/ Afr-1b); C (African-2a/ Afr-2b); D1 (North American-1 / NA), D2 (Asian-American-1/ AA1), and D3 (Asian-American-2/ AA2). The non-European (NE) lineages (B/C/D) are known to have a higher risk of infection persistence, progression to the pre-cancerous lesion, and development of cancer than the European (E) lineage (A) (Burk et al., 2013).

HPV16 L2 genetic variation and its phylogenetic analysis are vital to understanding viral infectivity and pathogenicity. These analyses are also crucial for the clinical setting of cervical cancer management, especially in diagnostic and vaccine strategies (Yue et al., 2013). Although L1 protein is widely known to induce neutralizing antibodies and be used in viral-like particle (VLP) vaccines worldwide, the complexity and high cost of manufacturing multivalent L1-VLP vaccines limit the number of protection spectrum and leave many low-resource countries unvaccinated (Schellenbacher et al., 2017). Meanwhile, the L2-based HPV vaccine should protect women from the most carcinogenic HPV types and increase the possibility of mass vaccination in developing countries. However, research on the *L2* gene, especially in HPV16 as Indonesia's most prevalent oncogenic HPV type, is still limited. Here we determine the *HPV16 L2* genetic variation and phylogenetic lineage from cervical cancer specimens in Indonesian women.

Materials and Methods

Study design and sample collection

This study was a cross-sectional observational study. The samples obtained in this study were as many as 23 HPV16 positive DNA isolates of cervical cancer specimens stored in the Center for Diagnostic and Research on Infectious Disease, Faculty of Medicine, Universitas Andalas, Padang, Indonesia (PDRPI Lab). This study has been approved by the ethics committee of the Faculty of Medicine, Public Health, and Nursing, Universitas Gadjah Mada, Yogyakarta, Indonesia, Ref. No.: KE/FK/1273/EC/2019.

DNA Amplification

We detected the *HPV16 L2* gene (NCBI reference sequence: NC_001526.4, nucleotide position: 3,373-4,794) in the samples with the conventional PCR method. We used three HPV16 L2 targeted primers: L2.1, L2.2, and L2.3 (Table 1), designed using AmplifX software and confirmed with NCBI-BLAST to amplify the DNA target. The amplifications were performed separately for each primer set with a concentration of 10 µM (1 µL), DNA template (3 µL), and using TopTaq master mix kit Qiagen with a total PCR mix volume of 50 µL. We performed the PCR as follows: (1) initial denaturation (94°C for 2 min); (2) 35 cycles of denaturation (94°C for 45 s), annealing (different temperature of each primer for 45 s), and extension (72°C for 1 min); and (3) final extension (72°C for 10 min). The DNA amplicons were electrophoresed at 7 µL each in 1% agarose gell with 0.5 µL Gel Red staining with a voltage of 100 V for 30 minutes. The indicator was a DNA ladder of 1 kbp or 100 bp. PCR results were observed under UV transillumination or a Gel Doc machine and documented.

Gene sequencing and variation analysis

To acquire the gene sequences, the amplicons of the *HPV16 L2* gene (40 µL of each sample) were sequenced in the 1st Base Singapore (41 Science Park Road #04-08, The Gemini, Singapore 117610) using the Bigdye® Terminator V3.1 Cycle Sequencing Kit and the three HPV16 L2 primer sets. The sequence products of each sample with different primer sets were combined using MEGA X v10.0.5 software to obtain a maximum length of *HPV16 L2* gene sequence (full size 1,422 bp, nucleotide position 3,373-4,794). The combined HPV16 L2 sequences were aligned with the HPV16 genome (NCBI reference sequence: NC_001526.4) to observe the single nucleotide polymorphisms (SNPs) and amino acid variations.

Phylogenetic analysis

To recognize the HPV16 L2 sub-lineages, the phylogenetic analysis was performed by aligning all the 23 HPV16 L2 sequences with the NCBI reference sample (NC_001526.4) and 10 HPV16 genomes representing different sub-lineages of variant prototypes (A1, K02718; A2, AF526179; A3, HQ644236; A4, AF534061; B1, AF536180; B2, HQ644298; C, AF472509; D1, HQ644257; D2, AY686579; D3, AF402678). We reconstructed the phylogenetic tree using MEGA X v10.0.5 software with the evolutionary analysis by maximum likelihood method and Tamura-Nei model with the number of bootstrap replications was 1,000.

Nucleotide sequence accession number

The *HPV16 L2* gene sequences obtained were submitted to the GenBank with accession numbers from MW810431 to MW810453.

Results

HPV16 L2 sequence variations

From a total of 23 samples, the sequencing process generated as many as nine complete and 14 partial *HPV16 L2* gene sequences. All partial sequences were obtained at best from HPV16 L2 primer sets of L2.2 and L2.3 only, with a combined sequence size of 1,132 bp in a nucleotide position of 3,663-4,794. From these available sequences, we detected a total of 35 SNPs with 17 non-

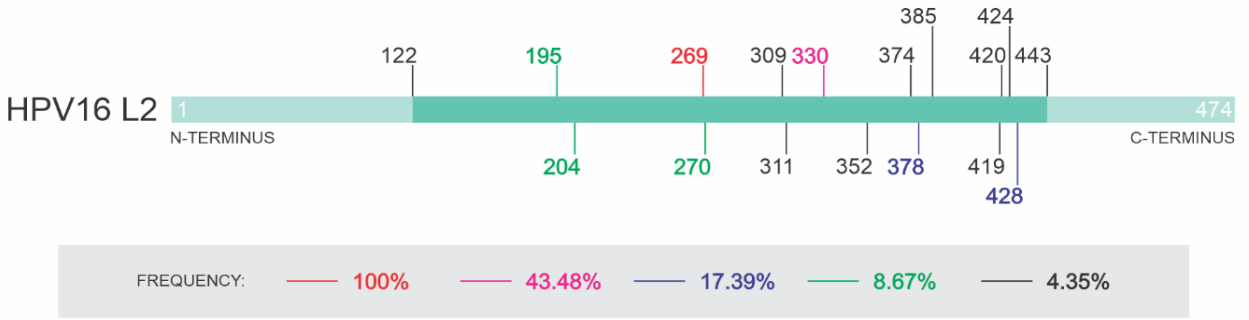

Figure 1. Position of Amino Acid Variations in HPV16 L2 Open Reading Frames. An illustrative green bar represents the L2 protein. The amino acid variation position site is shown in line with the number. The variation frequencies are described in percentages representing color (red, pink, blue, green, and black). Most of these variations were detected and figured in the central region (deep green). There is no variation near the N-terminus and C-terminus region of L2 (light green).

synonymous SNPs (nsSNPs) and 18 synonymous SNPs (sSNPs) (Table 2).

The most variable sites of nucleotide were G4074A (100%), T4177C (100%), T4362G (52.2%) and T4362C (43.5%) compared to NCBI reference. According to these sequences, we recognized at least ten unique genetic variations or variants. Most amino acid variations were detected at the center region of HPV16 L2 protein, and the

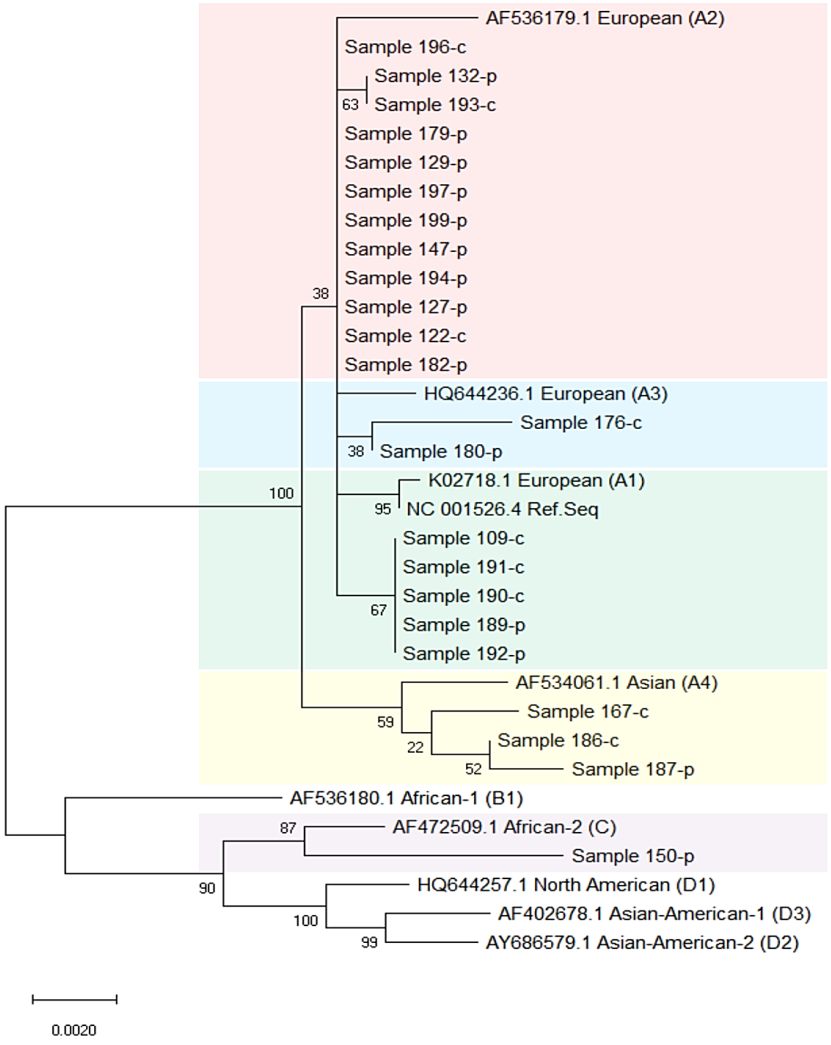

Figure 2. Phylogeny of HPV16 L2 sequences. The phylogenetic tree was constructed from 23 HPV16 L2 sequences compared with the whole-genome sequences of NCBI reference (NC_001526.4) and 10 intra-typic variants representing the HPV16 sub-lineages (A1, A2, A3, A4, B1, B2, C, D1, D2, D3). The analysis was performed with the maximum likelihood method and the Tamura-Nei model with 1,000 bootstrap replications. The colored box represented the HPV16 sub-lineage: red (A2), blue (A3), green (A1), yellow (A4), and purple (C). Scale bar was nucleotide substitution per site.

Table 1. Primer for Amplification of *HPV16 L2* Gene

Primer*	Nucleotide sequences	Position	Product size (bp)	Tm (°C)	Ta (°C)
L2.1 F	5'-CAG CCT CTG CGT TTA GGT GTT T-3'	3143-3754	612	61.39	62.8
L2.1 R	5'-GGG GAA TGG AAG GTA CAG ATG TTG-3'			61.17	
L2.2 F	5'-TGT GGG CCC TTC TGA TCC TTC TAT-3'	3663-4281	619	63.00	59.6
L2.2 R	5'-GTA CCT AAT GCC AGT ACG CCT AGA-3'			61.52	
L2.3 F	5'-CTA GGC GTA CTG GCA TTA GGT ACA-3'	4259-4919	661	61.76	58.7
L2.3 R	5'-AGG GAT GTC CAA CTG CAA GTA G-3'			60.03	

*Primer: F, forward; R, reverse; bp, base pair; Tm, melting temperature; Ta, annealing temperature.

variations predominantly were Serine to Proline (n=23/23; T4177C/ S269P) and Leucine to Phenylalanine (n=10/23; A4362C/ L330F) (Figure 1). We detected neither SNPs nor amino acid variations near the N-terminus region of HPV16 L2 in all complete sequences. There were also no deletions or insertions alongside the HPV16 L2 region.

HPV16 L2 phylogeny

We reconstructed the phylogenetic tree of the *HPV16 L2* gene sequences between the samples, the GenBank reference, and ten intra-typic variants representing the sub-lineage of HPV16 with the Tamura-Nei model (Figure 2). There were five phylogenetic sub-lineages of HPV16 L2 found closely related to sub-lineage A1/European (n=5), A2/European (n=12), A3/European (n=2), A4/ Asian (n=3), and C/African-2a (n=1). There was no sample close to the B and D sub-lineages. These results indicated that the A lineage (European-Asian) was the most prevalent variant in this study, with the highest dominance being the A2 sub-lineage (European).

Discussion

HPV16 infection remains the leading etiological cause of cervical cancer worldwide, including in developing countries like Indonesia (Kombe Kombe et al., 2021). The L2 minor capsid protein is vital in HPV infection, especially for entering the host's cell, vesicular trafficking, nuclear entry, and supporting nuclear activities. L2 is also proposed as a broadly neutralizing epitope, unlike L1, especially in the region near the N-terminus at amino acid residues 17-36, 58-64, 64-73, and some other segments beyond the first 120 residues (Day et al., 2010; Zhai et al., 2017; Huber et al., 2021). A study on *HPV16 L2* gene variation should be essential for future research about their differences in pathogenicity in cervical cancer and recognizing their phylogenetic sub-linages. It is also essential in a clinical setting, especially for vaccine strategies against HPV broad types and therapeutic modalities targeted at L2 (Yue et al., 2013; Yan et al., 2019; Huber et al., 2021; Xie et al., 2021). However, the *HPV16 L2* gene variation data in Indonesia is still limited. Here we find most SNPs in Indonesia's 23 HPV16 L2 sequences at the middle to C-terminus. However, the amino acid near N-terminus is notably conserved, particularly at the first 120 residues.

Sequencing analysis reveals as many as 35 SNPs and results in ten unique variants representing their specificity in the SNPs. For example, Variant 1, which is phylogenetically related to A1/European sub-lineage, has a distinctive SNP at T4548G, different from the NCBI reference sequence, another A1 sub-lineage member. The entire *L2* sequence of Variant 1 is 100% similar to the Japan variant (NCBI accession number LC456632.1 (Hirose et al., 2019). The most prevalent variant is Variant 4 from the A2/European sub-lineage (39.13%), homolog with variants of Netherland (KY549169.1) and Japan (LC511106.1) (van der Weele, Meijer and King, 2017; Hirose et al., 2019). Nevertheless, a unique variant is Variant 10 from the C/African-2 sub-lineage, which is, to the best of our knowledge, the 100% similarity is not available. Together, we assume that there should be a historical relationship between Indonesian variants and these variants. However, further research needs to be done with larger sample size.

According to the type of point mutation, there are only transversion and translation detected in the *HPV16 L2* sequence without any deletion or insertion. This finding is similar to most studies involving whole-genome sequencing of HPV16 in Asia, Europe, and America (Chen et al., 2005; Yue et al., 2013; Makowsky et al., 2016; Hirose et al., 2019; Lagström et al., 2019). However, deletions in *L2* gene sequences have been reported in Central China in three sites per 51 samples (Liu et al., 2017). Although there is no deletion or insertion in our study, amino acid (AA) variations are generated from nsSNPs as 17/35 (48.57%). The most prevalent AA variation is S269P of L2 269 (100%), which still has an unknown effect on the L2 activities. However, AA variations in L2 428 (17.4%) and L2 424 (4.3%) may be related to the L1-L2 bond, which is responsible for maintaining HPV virion (Finnen et al., 2003). We also detect AA variation in Variant 10, such as L2 419 and L2 420, important locations for virus interactions with the ND10 subdomain in the host cell nucleus (Becker et al., 2003). Other AA variations, L2 309 and L2 311 are found in all variants of the A4 sub-lineage that regulates the viral genome complex to accumulate in the cell nucleus in vivo (Mamoor et al., 2013).

Based on nine complete *HPV16 L2* sequences obtained from the study, there is no AA variation near the N-terminus region (AA 1-121). This finding likely supports the evidence that the region near the N-terminus of HPV L2 is highly conserved (Alphs et al., 2008; Wu et al., 2015; Namvar et al., 2019). This region is essential for L2 infection mechanisms such as cell surface exposure site, furin cleavage sites right before entering the host's cell, DNA binding domains, and localization of cell

Table 2. HPV16 L2 Gene Sequence Variations

Sample*	GenBank Accession	Nucleotide Position***																						
		3546	3661	3736	3780	3831	3861	3864	3873	3957	3983	3990	4023	4074	4177	4181	4298	4303	4362	4395	4407	4426	4446	4492
Ref	NC001526.4	A	C	T	T	T	C	T	T	A	G	C	A	G	T	G	A	A	A	A	A	A	T	A
109-c	MW810431	.	.	.	.	.	.	.	.	.	.	.	.	A	C	.	.	.	C	.	.	.	.	.
189-p	MW810445	.	.	.	.	.	.	.	.	.	.	.	.	A	C	.	.	.	C	.	.	.	.	.
190-c	MW810446	.	.	.	.	.	.	.	.	.	.	.	.	A	C	.	.	.	C	.	.	.	.	.
191-c	MW810447	.	.	.	.	.	.	.	.	.	.	.	.	A	C	.	.	.	C	.	.	.	.	.
192-p	MW810448	.	.	.	.	.	.	.	.	.	.	.	.	A	C		.	.	C	.	.	.	.	.
132-p	MW810435	.	.	.	.	.	.	.	.	.	C	.	.	A	C	.	.	.	G	.	.	.	.	.
193-c	MW810449	.	.	.	.	.	.	.	.	.	C	.	.	A	C	.	.	.	G	.	.	.	.	.
122-c	MW810432	.	.	.	.	.	.	.	.	.	.	.	.	A	C	.	.	.	C	.	.	.	.	.
127-p	MW810433	.	.	.	.	.	.	.	.	.	.	.	.	A	C	.	.	.	G	.	.	.	.	.
129-p	MW810434	.	.	.	.	.	.	.	.	.	.	.	.	A	C	.	.	.	G	.	.	.	.	.
147-p	MW810436	.	.	.	.	.	.	.	.	.	.	.	.	A	C	.	.	.	G	.	.	.	.	.
179-p	MW810440	.	.	.	.	.	.	.	.	.	.	.	.	A	C	.	.	.	G	.	.	.	.	.
182-p	MW810442	.	.	.	.	.	.	.	.	.	.	.	.	A	C	.	.	.	G	.	.	.	.	.
194-p	MW810450	.	.	.	.	.	.	.	.	.	.	.	.	A	C	.	.	.	G	.	.	.	.	.
196-c	MW810451	.	.	.	.	.	.	.	.	.	.	.	.	A	C	.	.	.	G	.	.	.	.	.
197-p	MW810452	.	.	.	.	.	.	.	.	.	.	.	.	A	C	.	.	.	G	.	.	.	.	.
199-p	MW810453	.	.	.	.	.	.	.	.	.	.	.	.	A	C	.	.	.	G	.	.	.	.	.
176-c	MW810439	G	.	.	.	.	.	C	.	C	.	.	.	A	C	.	.	.	C	.	G	.	.	C
180-p	MW810441	.	.	.	.	.	.	.	.	C	.	.	.	A	C	.	.	.	G	.	.	.	.	.
167-c	MW810438	.	T	.	.	.	T	.	.	.	.	.	.	A	C	A	T	G	C	.	.	.	.	.
186-c	MW810443	.	.	.	.	.	T	.	.	.	.	.	.	A	C	A	.	.	C	.	G	.	.	.
187-p	MW810444	.	.	.	.	.	.	.	.	.	.	.	.	A	C	.	.	.	C	.	G	.	.	.
150-p	MW810437	.	.	C	A	A	T	.	C	.	.	T	G	A	C	.	.	.	.	G	.	C	C	.
Amino Acid Variation**	Reference			S						E	S				S	S	K	T	L			T		I
	AA Var			P						D	T				P	N	I	A	F			P		L
	Nucl Var			C						C	C				C	A	T	G	C			C		C
	AA Residue			122						195	204				269	270	309	311	330			352		374

*sample: c, complete sequence; p, partial sequence; **amino acid variation: AA, amino acid; Var, variation; Nucl, nucleotide; ***the unique SNPs that differentiate the sub-lineages are highlighted with the light gray background.

Table 2. Continued

Sample*	GenBank Accession	Nucleotide Position***												Variant ID	Sub-lineage
		4505	4515	4525	4526	4539	4548	4623	4627	4631	4642	4654	4700		
Ref	NC001526.4	C	G	G	T	T	T	C	A	T	G	A	C	Ref	A1
109-c	MW810431	.	.	.	.	.	G	.	.	.	.	.	.	1	
189-p	MW810445	.	.	.	.	.	G	.	.	.	.	.	.		
190-c	MW810446	.	.	.	.	.	G	.	.	.	.	.	.		
191-c	MW810447	.	.	.	.	.	G	.	.	.	.	.	.		
192-p	MW810448	.	.	.	.	.	G	.	.	.	.	.	.		
132-p	MW810435	.	.	.	.	.	.	.	.	.	.	.	.	2	A2
193-c	MW810449	.	.	.	.	.	.	.	.	.	.	.	.		
122-c	MW810432	.	.	.	.	.	.	.	.	.	.	.	.	3	
127-p	MW810433	.	.	.	.	.	.	.	.	.	.	.	.	4	
129-p	MW810434	.	.	.	.	.	.	.	.	.	.	.	.		
147-p	MW810436	.	.	.	.	.	.	.	.	.	.	.	.		
179-p	MW810440	.	.	.	.	.	.	.	.	.	.	.	.		
182-p	MW810442	.	.	.	.	.	.	.	.	.	.	.	.		
194-p	MW810450	.	.	.	.	.	.	.	.	.	.	.	.		
196-c	MW810451	.	.	.	.	.	.	.	.	.	.	.	.		
197-p	MW810452	.	.	.	.	.	.	.	.	.	.	.	.		
199-p	MW810453	.	.	.	.	.	.	.	.	.	.	.	.		
176-c	MW810439	.	.	.	.	.	.	.	.	.	.	C	.	5	A3
180-p	MW810441	.	.	.	.	.	.	.	.	.	.	.	.	6	
167-c	MW810438	T	.	.	.	.	.	.	.	.	.	C	.	7	A4
186-c	MW810443	T	.	.	.	.	.	.	G	.	.	C	.	8	
187-p	MW810444	T	.	.	.	.	.	.	G	.	.	C	.	9	
150-p	MW810437	T	A	A	C	C	.	T	C	C	A	.	G	10	C
Amino Acid Variation**	Reference	S		V					N	I	A	I	A		
	AA Var	F		T					H	T	T	L	G		
	Nucl Var	T		A					C	C	A	C	G		
	AA Residue	378		385					419	420	424	428	443		

*sample: c, complete sequence; p, partial sequence; **amino acid variation: AA, amino acid; Var, variation; Nucl, nucleotide; ***the unique SNPs that differentiate the sub-lineages are highlighted with the light gray background.

nucleus signals (Wang and Roden, 2013). Several studies have shown that the L2 N-terminus region potentially induces cross-neutralizing antibodies and promises to be a broad HPV vaccine epitope candidate (Tumban et al., 2012; Wu et al., 2015; Zhai et al., 2017; Olczak and Roden, 2020; Yang et al., 2020). The low-cost L2-based vaccine strategy may become a possible alternative compared to the high cost of the L1-based vaccine available (Roden and Stern, 2018). We hypothesize that L2-based vaccine strategies should effectively lower the morbidity and mortality of cervical cancer in developing countries like Indonesia, yet further research is necessary.

To address the variants and sub-lineages of Indonesian *HPV16 L2* sequences, we conduct a phylogenetic analysis using the maximum likelihood method and Tamura-Nei model with a number of bootstrap replications of 1,000. As many as ten unique sequence variants are revealed, with European sub-lineages (A) dominance, particularly A2. Our finding follows a sizeable phylogenetic study of HPV16 whole genome-sequencing, highlighting the A sub-lineage as the world's most prevalent phylogeny (78.8%), except for the A2 sub-lineage domination (Clifford et al., 2019). The limitations of this study are due to using stored HPV16 DNA positive isolates, observing few samples available, and inadequate original patients' clinical information. However, this research should be one of the preliminary studies that reveal the variation of *HPV16 L2* gene sequence in Indonesia and provide supporting evidence to further research on HPV-related disease.

In conclusion, our study determines the *HPV16 L2* gene sequence variation and phylogeny of Indonesian isolates. We find the most variations in the central region of the L2 sequences, and the cross-protective region near the L2 N-terminus is notably conserved. This finding should support the opportunity of a cost-effective HPV L2-based vaccine strategy to deal with cervical cancer, especially in developing countries. Future research is necessary to explore HPV16 genetic variation and enhance our understanding of genetic-based strategies to reduce cervical cancer-related morbidity and mortality.

Author Contribution Statement

SPP performed the research experiments, analyzed the data, and drafted the manuscript; AEP collected the specimens, prepared the HPV DNA isolates, designed the experiments, provided laboratory equipment, and drafted the manuscript; WA and IA contributed to reviewing the manuscript and guiding the research. All authors read and approved the final manuscript.

Acknowledgments

The authors would like to thank the Dean of Faculty of Medicine, Universitas Andalas, for supporting and funding the research. We also thank all members of the Microbiology Department and the laboratory of the Center for Diagnostic and Research on Infectious Diseases (PDRPI Lab) for the support and facilitation of the experiments. We appreciate all staff in the Master Program of Tropical Medicine of Universitas Gadjah Mada who contributed to the study.

Funding statement

The funding was received from the Faculty of Medicine, Universitas Andalas, allocated to the "Budget Implementation List (DIPA) of Fundamental Scheme" project, with grant number 1097/UN16.02.D/PP/2020.

Ethics approval

This research was approved by The Medical and Health Research Ethics Committee (MHREC) of the Faculty of Medicine, Public Health, and Nursing, Universitas Gadjah Mada, Ref.No.: KE/FK/1273/EC/2019 as a part of student thesis. Written informed consent was previously derived from each patient to use their specimen data for laboratory studies. The research was conducted following the relevant ethical guidelines and regulations.

Availability of data and materials

The data is available from the corresponding author and PDRPI Lab, but restrictions apply to the availability of these data. Data are available directly from the author (syandrez@med.unand.ac.id) and PDRPI (divisi_diagnostik_infeksi@med.unand.ac.id).

Conflict of interest

The authors declare that there is no conflict of interest in this study.

References

Alphs HH, Gambhira R, Karanam B, Roberts JN, Jagu S, et al (2008). Protection against heterologous human papillomavirus challenge by a synthetic lipopeptide vaccine containing a broadly cross-neutralizing epitope of L2. *Proc Natl Acad Sci U S A*, **105**, 5850–5.

Araldi RP, Sant'Ana TA, Módolo DG, et al (2018). The human papillomavirus (HPV)-related cancer biology: An overview. *Biomed Pharmacother*, **106**, 1537–56.

Becker KA, Florin L, Sapp C, Sapp M (2003). Dissection of human papillomavirus type 33 L2 domains involved in nuclear domains (ND) 10 homing and reorganization. *Virology*, **314**, 161–7.

Bruni L, Albero G, Serrano B, Mena M, Gómez D, et al (2018). ICO/IARC Information Centre on HPV and Cancer (HPV Information Centre). Human Papillomavirus and Related Diseases in Indonesia. Summary Report 10 December 2018. Accessed from http://www.hpvcentre.net/statistics/reports/IDN.pdf on 04 April 2021.

Bruni L, Albero G, Serrano B, et al (2019). ICO/IARC Information Centre on HPV and Cancer (HPV Information Centre). Human Papillomavirus and Diseases in the World. Summary Report 22 January 2019. Accessed from https://hpvcentre.net/statistics/reports/XWX.pdf on 04 April 2021.

Burk RD, Harari A, Chen Z (2013). Human papillomavirus genome variants. *Virology*, **445**, 232–43.

Chalabiani S, Nazari MK, Shabani M, et al (2017). Retrospective Analysis of Prevalence of High-risk and Low-risk Human Papillomavirus (HPV) Genotypes in Iranian Women During 2013-2016. *Asian Pac J Cancer Biol*, **2**, 85-90.

Chen Z, Terai M, Fu L, et al (2005). Diversifying selection in human papillomavirus type 16 lineages based on complete genome analyses. *J Virol*, **79**, 7014–23.

Clifford GM, Tenet V, Georges D, et al (2019). Human papillomavirus 16 sub-lineage dispersal and cervical cancer risk worldwide: Whole viral genome sequences from 7116 HPV16-positive women. *Papillomavirus Res*, **7**, 67–74.

Day PM, Kines RC, Thompson CD, et al (2010). In vivo mechanisms of vaccine-induced protection against HPV infection. *Cell Host Microbe*, **8**, 260–70.

de Martel C, Plummer M, Vignat J, Franceschi S (2017). Worldwide burden of cancer attributable to HPV by site, country and HPV type. *Int J Cancer*, **141**, 664–70.

Faridi R, Zahra A, Khan K, Idrees M (2011). Oncogenic potential of human papillomavirus (HPV) and its relation with cervical cancer. *Virol J*, **8**, 269.

Finnen RL, Erickson KD, Chen XS, Garcea RL (2003). Interactions between Papillomavirus L1 and L2 Capsid Proteins. *J Virol*, **77**, 4818–26.

Hallowell BD, Saraiya M, Thompson TD, et al (2018). Population-based assessment of HPV genotype-specific cervical cancer survival: CDC Cancer Registry Sentinel Surveillance System. *JNCI Cancer Spectr*, **2**, 1–8.

Hirose Y, Mori S, Onuki M, et al (2019). Whole-Genome Analysis of Human Papillomavirus Type 16 Prevalent in Japanese Women with or without Cervical Lesions. *Viruses*, **11**, 1–14.

Huber B, Wang JW, Roden RBS, Kirnbauer R (2021). RG1-VLP and Other L2-Based, Broad-Spectrum HPV Vaccine Candidates. *J Clin Med*, **10**, 1044.

Johnson CA, James D, Marzan A, Armaos M (2019). Cervical Cancer: An Overview of Pathophysiology and Management. *Semin Oncol Nurs*, **35**, 166–74.

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

Kombe Kombe AJ, Li B, Zahid A, et al (2021). Epidemiology and Burden of Human Papillomavirus and Related Diseases, Molecular Pathogenesis, and Vaccine Evaluation. *Front Public Heal*, **8**, 1–19.

Lagström S, van der Weele P, Rounge TB, et al (2019). HPV16 whole genome minority variants in persistent infections from young Dutch women. *J Clin Virol*, **119**, 24–30.

Liu Y, Pan Y, Gao W, Ke Y, Lu Z (2017). Whole-genome analysis of human papillomavirus types 16, 18, and 58 isolated from cervical precancer and cancer samples in Chinese women. *Sci Rep*, **7**, 1–9.

Makowsky R, Lhaki P, Wiener HW, et al (2016). Genomic diversity and phylogenetic relationships of human papillomavirus 16 (HPV16) in Nepal. *Infect Genet Evol J Mol Epidemiol Evol Genet Infect Dis*, **46**, 7–11.

Mamoor S, Onder Z, Karanam B, et al (2013). Genetic variability in minor capsid protein (L2 gene) of human papillomavirus type 16 among Indian women. *NIH Public Access*, **422**, 413–24.

Namvar A, Bolhassani A, Javadi G, Noormohammadi Z (2019). In silico/In vivo analysis of high-risk papillomavirus L1 and L2 conserved sequences for development of cross-subtype prophylactic vaccine. *Sci Rep*, **9**, 1–22.

Olczak P, Roden RBS (2020). Progress in l2-based prophylactic vaccine development for protection against diverse human papillomavirus genotypes and associated diseases. *Vaccines*, **8**, 1–22.

Oyervides-Muñoz MA, Pérez-Maya AA, Rodríguez-Gutiérrez HF, et al (2018). Understanding the HPV integration and its progression to cervical cancer. *Infect Genet Evol,* **61**, 134–44.

Roden RBS, Stern PL (2018). Opportunities and challenges for human papillomavirus vaccination in cancer. *Nat Rev Cancer*, **18**, 240–54.

Sabeena S, Bhat PV, Kamath V, et al (2017). Community-based prevalence of genital human papilloma virus (HPV) infection: A systematic review and meta-analysis. *Asian Pacific J Cancer Prev*, **18**, 145–54.

Schellenbacher C, Roden RBS, Kirnbauer R (2017). Developments in L2-based human papillomavirus (HPV) vaccines. *Virus Res*, **231**, 166–75.

Serrano B, Brotons M, Bosch FX, Bruni L (2018). Epidemiology and burden of HPV-related disease. *Best Pract Res Clin Obstet Gynaecol*, **47**, 14–26.

Setiawan D, Dusafitri A, Galistiani GF, van Asselt ADI, Postma MJ (2018). Health-Related Quality of Life of Patients with HPV-Related Cancers in Indonesia. *Value Heal Reg Issues*, **15**, 63–9.

Tommasino M (2013). The human papillomavirus family and its role in carcinogenesis. *Semin Cancer Biol*, **26**, 13–21.

Torre LA, Siegel RL, Ward EM, Jemal A (2016). Global cancer incidence and mortality rates and trends - An update. *Cancer Epidemiol Biomarkers Prev*, **25**, 16–27.

Tumban E, Peabody J, Tyler M, Peabody DS, Chackerian B (2012). VLPs Displaying a Single L2 Epitope Induce Broadly Cross-Neutralizing Antibodies against Human Papillomavirus. *PLoS One*, **7**.

van der Weele P, Meijer CJLM, King AJ (2017). Whole-Genome Sequencing and Variant Analysis of Human Papillomavirus 16 Infections. *J Virol*, **91**.

Wang JW, Roden RBS (2013). L2, the minor capsid protein of papillomavirus. *Virology*, **445**, 175–86.

Wu WH, Alkutkar T, Karanam B, et al (2015). Capsid display of a conserved human papillomavirus L2 peptide in the adenovirus 5 hexon protein: A candidate prophylactic hpv vaccine approach. *Virol J*, **12**, 1–11.

Xie J, Zhang P, Crite M, Lindsay CV, DiMaio D (2021). Retromer stabilizes transient membrane insertion of L2 capsid protein during retrograde entry of human papillomavirus. *Sci Adv*, **7**.

Yan H, Foo SS, Chen W, et al (2019). Efficient inhibition of human papillomavirus infection by l2 minor capsid-derived lipopeptide. *MBio*, **10**, 1–18.

Yue Y, Yang H, Wu K, et al (2013). Genetic Variability in L1 and L2 Genes of HPV-16 and HPV-58 in Southwest China. *PLoS One*, **8**.

Yang F, Mariz FC, Zhao X, et al (2020). Broad Neutralization Responses Against Oncogenic Human Papillomaviruses Induced by a Minor Capsid L2 Polytope Genetically Incorporated Into Bacterial Ferritin Nanoparticles. *Front Immunol*, **11**, 1–16.

Zhai L, Peabody J, Pang Y-YS, et al (2017). A novel candidate HPV vaccine: MS2 phage VLP displaying a tandem HPV L2 peptide offers similar protection in mice to Gardasil-9. *Antiviral Res*, **147**, 116–23.