

Conservation of Major L1 and Variability of Minor L2 Capsid Late Protein Genes in Human Papillomavirus of Indonesia Variants

Sukma Nuswantara^{1*}, Titik K Prana¹, Dwi Wulandari¹, Henni Widyowati¹, Vera Anzela², Dea Levy², Petrus Cahyadi³, and Lukas D Tjandra³

¹Research Centre for Biotechnology, Indonesian Institute of Sciences (LIPI), Indonesia

²Department of Biology, Padjadjaran University, Jatinangor, Sumedang, Indonesia

³Santosa Bandung International Hospital, Sandia Biotech Diagnostic Centre, Bandung, Indonesia

Abstract

Human Papilloma Virus (HPV) has an outstanding feature for its vast intraspecies variability. Of all known 100 types or more, 15 types of them are classified as high risk because of their occurrence in more than 95% of cervical cancer cases. Among all genes in their genome, E6 and E7 genes are considered oncogenes and have close relevance with their pathogenicity, whilst L1 and L2 genes produce capsid proteins that directly interact with their host receptors. Considering the importance of L1 and L2 in host-receptor relationship, we tried to investigate their molecular variability thereby uncover their specificity as Indonesian variants. Here we reported about the conservation of L1 minor capsid protein and variability of L2 capsid protein among high-risk types Human Papilloma Virus (HPV). The results indicated that L1 DNA was relatively more conserved than its L2 counterpart. Also it was indicated that the middle part of either L1 or L2 CDS' showed more DNA variability than those at their upstream sequences. It is concluded that L2 middle sequences are important factors for intraspecific variations found in HPV of Indonesian variants.

Keywords: Human Papilloma Virus (HPV), L1 gene, L2 gene, cervical cancer, high risk HPV

*Corresponding author

Jl. Raya Bogor Km. 46, Cibinong 16911, Indonesia

Tel. +62-21-8754587, Fax. +62-21-8754588

E-mail. sukma.nuswantara@lipi.go.id

Introduction

Cervical cancer is an “infectious” cancer caused by persistent infection of Human Papilloma Virus (HPV) of oncogenic types which could potentially lead to cancer. More than 95% of cervical cancer cases are caused by several types of high-risk HPV (Bekkers *et al*, 2004; Motoyama *et al*, 2004). In Indonesia, cervical cancer is the leading cause of death in women followed by breast cancer (Novel *et al*, 2010).

Since the advent of molecular methods, over 100 HPV genotypes have been isolated and characterized. Among these more than 40 have been shown to infect the genital tract, and 15 of them (genotypes 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68, 73, and 82) have been found to be associated with cervical cancer or High-Grade Cervical Intraepithelial Neoplasia (CIN) and therefore classified as high-risk types (Munoz, 2000; Munoz *et al*, 2003).

For decades, early detection method of cervical cancer has been relied on cytological technique known as Papsmear developed by Dr. Papanicolaou (Novel *et al*, 2010). Recently HPV DNA tests have been developed that make use of molecular biology techniques. The FDA approved HPV detection kit of Digene/Hybrid Capture 2 (HC2) manufactured by Digene (MD, USA) was presently the best example of HPV-DNA test available in the market. The HC2 procedure includes several steps using genotype-specific probes that are mixed in a high-risk or a low-risk RNA cocktail, and RNA-DNA hybrids are recognized by an antibody, used both for the capture step and for a signal amplification detection method. This technique is used to screen high-risk types from low-risk and other types; however it cannot discriminate among different genotypes (Burd, 2003).

HPV belongs to the Papovaviridae family, has a diameter of 55 µm with isohedral protein structure that are arranged from 72

capsomeres. It has circular double-stranded DNA with a length of approximately 8000 base pairs (de Villiers *et al*, 2004). The linearized representation of HPV 16 genome is shown in Figure 1.

Studies of HPV genome variation are very important to provide biological component as diagnostic or immunogenic materials originated from diverse sources of local

variants. This study investigated molecular variation of L1 and L2 DNA sequence of HPV type 16 of Indonesia origin. At the end, this study may provide genetic information about genomic variations of L1 and L2 genes of HPV type 16 of Indonesian variants so that it opens up opportunities in the design and manufacture of recombinant vaccine as well as diagnostic kits based on the local database.

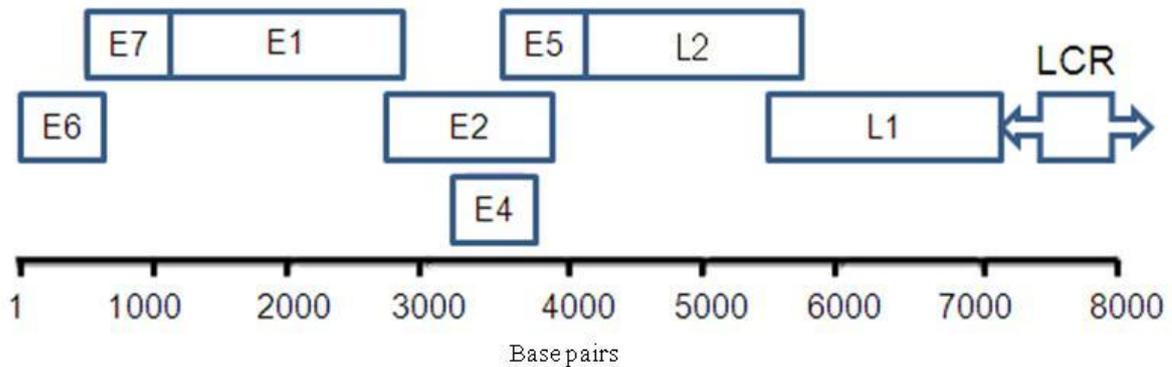


Figure 1. Linearized representation of HPV genome consists of Early (E1-E7), and Late (L) genes (L1 = 1421 bp & L2 = 1595 bp) genes and long controlling region (LCR).

Table 1. Oligonucleotide primers for PCR and DNA sequencing for the amplification of partial L1 and L2 regions.

No.	Primer name	5'-3' sequence	Length
1.	HPV 16 F	GGT CGG TGG ACC GGT CGA TG	20
2.	HPV 16 R	GCA ATG TAG GTG TAT CTC CA	20
3.	L1 LEFT F	TGC AGC AGG TGA CTT TTA TTT ACA TCC	27
4.	L1 LEFT R	ACC ACG ACC TAC CTC ACC TAC	21
5.	L1 MIDF	GCT GGT TTG GGC CTG TGT AG	20
6.	L1 MIDR	TTG GCC TTT AAT CCT GCT TGT AG	23
7.	L2 LEFT F	TTT TTG GTG GGT TAG GA	17
8.	L2 LEFT R	AAA GTG GGA TTA TTA TGT GTA GTA	24
9.	L2 MIDF	AGA CCC CCT TTA ACA GTA GAT	21
10.	L2 MIDR	TAG GCA GCC AAA GAG ACA	18

Table 2. Temperature program for real time PCR followed by HRM analysis of L1 and L2 PCR fragments. LP= Left primer; M= Middle primer. *=denotes sec/increment

Cycle Step	Temperature (°C)		Time (min:sec)		Cycle
	L1	L2	L1	L2	
Enzyme activation	98	98	03:00	03:00	1
Denaturation	98	98	00:05	00:05	39
Annealing/ extension	61.4 (LP) 64.9 (M)	55-60	00:10	00:05	
HRM	75-95	75-95	00:05*	00:05*	1

Materials & Methods

Sampling & HPV DNA preparation. Cervical swab sample suspensions in the Specimen Transport Medium (Digene™, USA) were spotted on filter paper (Whatman FTA™

Cards) to immobilize their DNA and then air-dried for long term storage afterwards. DNA template was prepared from at least 3 punched paper discs by elution with FTA elution buffer (Whatman™, USA). DNA

concentration was measured using a spectrophotometer.

Hybrid Capture Technology. Second generation Hybrid capture technology (HC-II; The Digene Corporation, USA) was used to screen high risk HPV types to detect nucleic acid targets directly. This screening discriminates the following high-risk types (as initially defined by Digene and supported by epidemiological studies): HPV types 16, 18, 31, 33, 35, 45, 51, 52, and 56. The HC-II procedure includes: DNA denaturation (1 hr), Mixing with RNA probe cocktail followed by hybridization at 65°C (1.5 hr), Hybrid Capture (1.5 hr), Detection for labeling (1 hr) and finally Data collection, validation and interpretation (1 hr).

PCR primers. PCR and sequencing primers were designed using a standalone primer design software (PrimerSelect™, DNA Star™, Hitachi, Japan) generated from L1 (bases 5560-7155) and L2 (bases # 4236-5657) CDS' of Human Papillomavirus of East Asia Type as reference (GenBank Accession number AF534061.1). Several alternative primers covering at least the first 300 bp upstream regions of both L1 and L2 genes are shown in. (Table 1).

Standard PCR, Real Time PCR and High Resolution Melting (HRM) Analysis. Standard PCR was done using Master Gradient Thermal Cycler (Eppendorf™, Germany). Real Time PCR was performed on CFX96 machine attached to its C1000 base unit (BioRad, USA). PCR mix (GoTaq Green, BioRad USA) was used throughout the experiment to prepare DNA for genotyping of HPV 16 and DNA sequencing. The typical PCR cycle was as follows: Initial denaturation 95°C, 5 min; Denaturation 95°C, 30 sec; Annealing 54°C, 30 sec; Extension 72°C, 30 sec. The denaturation, annealing and extension steps were repeated to 39 cycles followed by final extension at 72°C for 10 min. High Resolution Melting (HRM) Analysis was done in a Real Time PCR setting using dsDNA specific dye (Sso Fast™ EvaGreen® Supermix, BioRad USA) which was chosen as a preferred alternative to SYBR Green dye (Mao *et al*, 2007). HRM temperature scheme was immediately

attached following the PCR temperature protocol as shown in Table 2.

DNA Sequencing and analysis. Analysis of the L1 and L2 were done using modified dideoxy Sanger Method in the PCR cycle sequencing to determine the polymorphism in detail, with Automated DNA Sequencer ABI 3730 XLs (Applied Biosystem, USA). Sequence analysis such as multiple alignments and phylogenetic clustering were performed using Bio Edit or MegAlign™ DNA Star (Hitachi, Japan). The Clustal V method groups sequences into clusters by examining sequence distances between all pairs. Clusters are aligned as pairs then collectively as sequence groups to produce the overall alignment. After finishing the multiple alignment, a neighbor-joining method was employed to reconstruct phylogeny for the putative alignment.

Results

L1 HPV DNA Analysis.

High resolution melting (HRM) analyses were performed following Realtime PCR amplification of L1 HPV. The results indicated that there were 12 different types of polymorphisms grouped into 12 clusters. (Figure 2A) on the first 300 bp upstream regions. The same analysis was performed on L1 middle regions and finally grouped them into 8 polymorphic clusters (Figure 2B). All clusters were used as basis for choosing sequencing samples thereby avoiding too many samples to sequence.

Phylogenetic analysis of L1 HPV DNA sequences

Based on the multiple alignment analyses, DNA distances were grouped to construct a phylogenetic tree of upstream region of L1 HPV DNA sequence (Figure 3). As seen from the tree construct, upstream region of L1 DNAs were relatively homogenous with tight grouping.

These results coincide with DNA sequence analyses shown on Figure 4 where Indonesian variants had relatively conserved L1 sequences at their upstream region. However Indonesian variants grouped together with the majority of them were separated from L1 HPV from other countries.

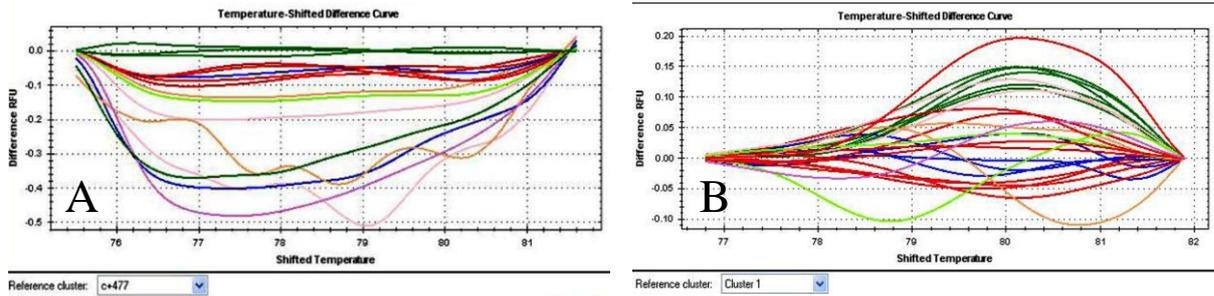


Figure 2. L1 HPV-DNA clusters generated by *High Resolution Melting* (HRM) analysis resulting in 12 clusters of L1 upstream regions (A) and 8 clusters of L1 middle regions (B) respectively.

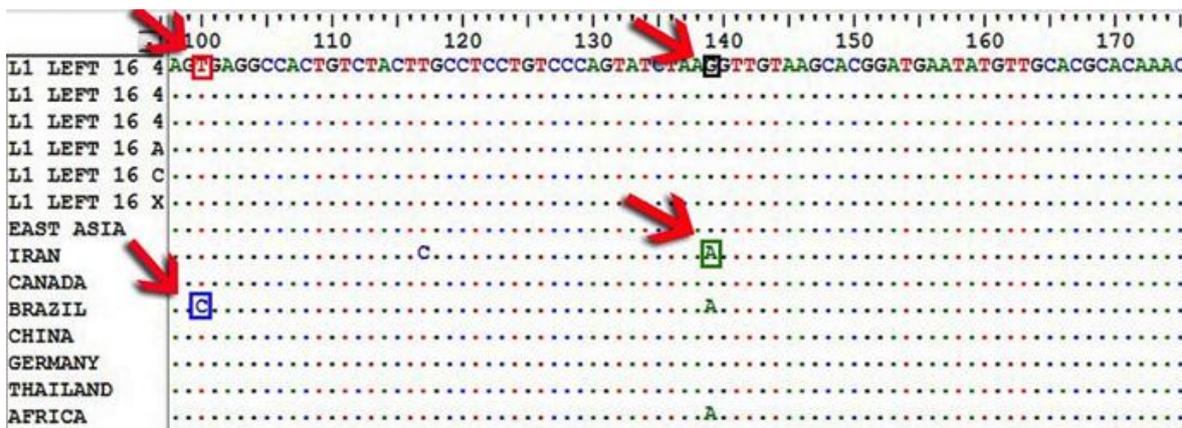


Figure 4. Multiple alignment of L1 HPV DNA

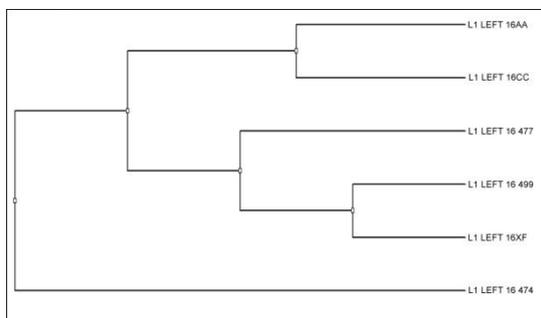


Figure 3. Phylogenetic analyses of L1 HPV-DNA of upstream sequence region.

In contrast, the L1 middle regions were more variable than the upstream regions as seen from the phylogenetic construct on Figure 5. Again, Indonesian variants grouped together to form a distinctive major clusters with only HPV from China was the closest relative. Surprisingly, the Indonesian variants had farthest genetic distance with those of East Asia type.

L2 HPV DNA analysis

As with L1 DNA, the L2 DNA was first analyzed by High Resolution Melting following Realtime PCR and the resulting

melting curves were converted into temperature shifted difference curves to see clearer polymorphic characteristics. The L2 DNA upstream regions were grouped into 3 clusters on the basis of melting temperature differences (Figure 6A). The middle sequences, however, showed more variable sequences with 14 clusters (Figure 6B).

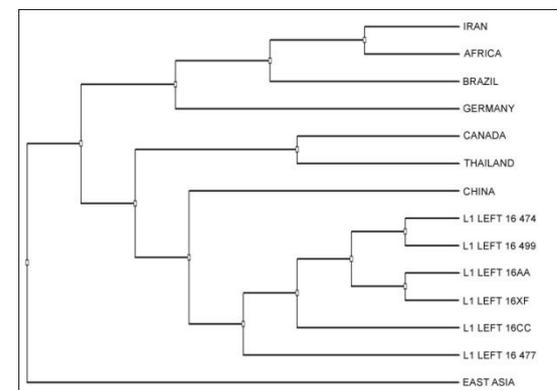


Figure 5. Phylogenetic of HPV DNA of middle L1 regions

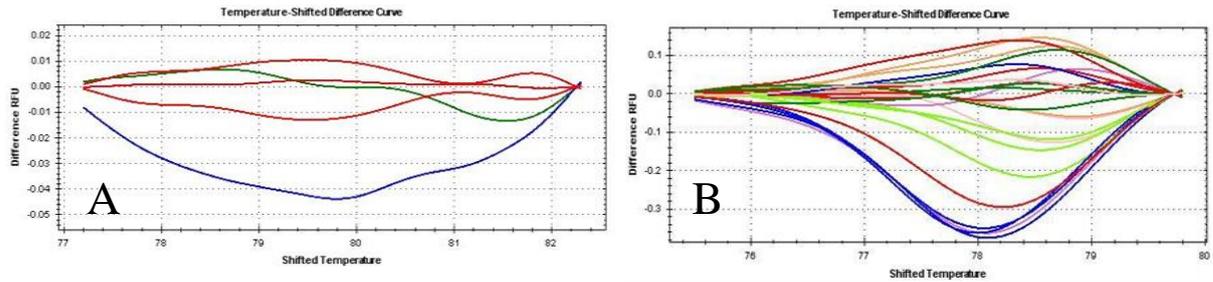


Figure 6. Temperature-shifted difference curve on HRM analysis showing clusters on upstream (A) and middle (B) regions

Multiple alignments of the first 300 bp L2 upstream sequences is shown in Figure 7. Here, several sequence variations were seen as single bases differences. Overall, L2 upstream regions showed more variable than L1 upstream region (Figure 8).

The middle part of L2-HPV DNAs were also positioned in separate clusters from other HPV from East Asia, Africa, Brazil, Thailand, Germany as well as from China and USA (Figure 9).

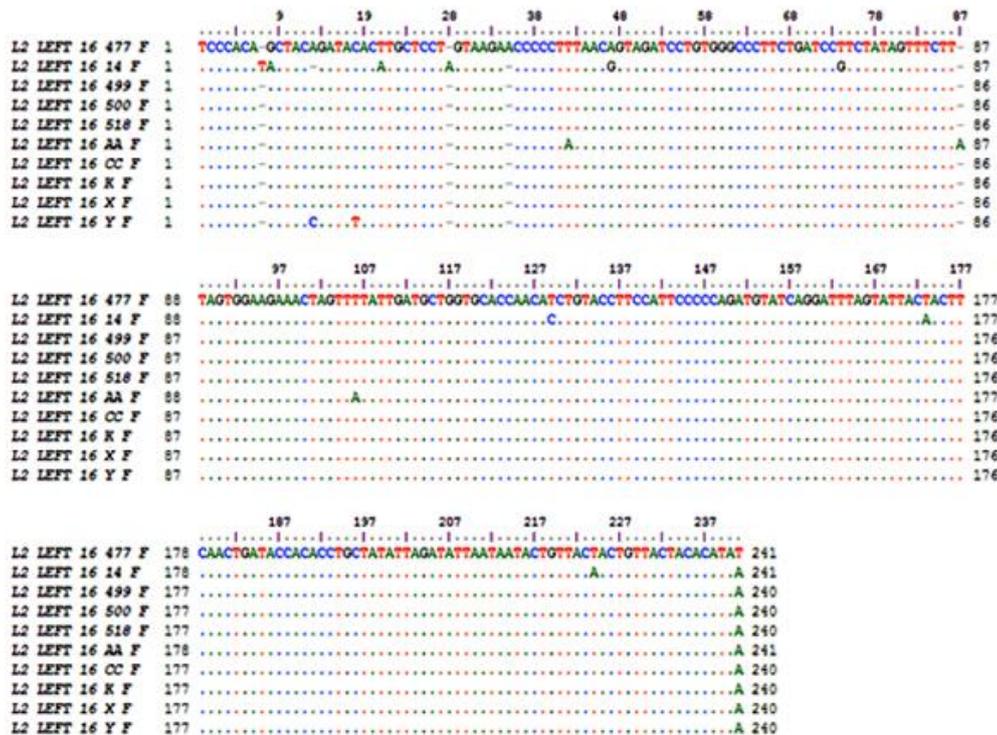


Figure 7. Multiple alignment analysis of L2 upstream sequence

Discussion

HPV genome consists of early proteins genes (E1, E2, E4, E5, E6, E7), in which the E6 and E7 plays an important role in cancer formation, hence oncogene term applied. The other 2 genes are late proteins (L1, L2) which are immunogenic and have been recently much explored in the design of HPV bivalent and quadrivalent vaccines (Harper *et al*,

2004; Koutsky *et al*, 2002; Schiller *et al*, 2004; Villa *et al*, 2005; Zheng *et al*, 2006).

It was reported that nucleotide variation rate of 2% was found in HPV type 16 (Bernard *et al*, 2006). The fact that mutations occurred in L1 and L2 genes lead to the change of amino acid sequence of the capsid protein. This also supports the possibility of an immunogenic and oncogenic diversity between the variants of all types of HPV (Prado *et al*, 2005). L2 minor capsid protein

has a more reactive antibody than the L1 major capsid protein, making it a suitable candidate for the development of modern vaccines against HPV infection. Virus-like particle (VLP) is a DNA-less capsid protein used in the construction of prophylactic bivalent and quadrivalent vaccines against both high risk and low risk HPV marketed as Gardasil (Merck & Co) and Cervarix (Glaxo Smith Kline) which finally gained FDA approval (Lowy & Schiller, 2006; FDA, 2008).

L2 HPV plays an important role in capsid assembly by introducing the HPV DNA to the virus particles. It is also a binding protein that can interact with any DNA without any requirement of specific DNA sequences which helps in the formation of virions with viral DNA and kapsomer linking. HPV L2 also has an important role in viral entry into cells, localization of viral components into the core, capsid formation, and stability of DNA binding (Ishii *et al*, 2005; Zhou *et al*, 1994).

Compared with L1, L2 is much smaller but the DNA sequence is more variable. L2 has been used as sources of antigen for type-specific HPV antibody (Motoyama *et al*, 2004). L2 HPV types 11 and 16 require L1 in the process of virus penetration into the nucleus (Bordeaux *et al*, 2006; Kieback *et al*, 2006). L2 can function in the uncoating of virions and virus delivery to the core genome (Richard *et al*, 2001).

Previously, it was reported that there were three mutations found in L2 sequence of HPV types 6 (Wang *et al*, 2001). Mutation from A to G at position 7081 in the genome of HPV types 6 and from G to A in 7099 represent the regional characteristics of L2.

The result of this study indicated that L2 sequences are more varied than the L1, either the upstream or the middle regions. This means that larger possibility of variations could be expected from this region including those found in the Indonesian variants. In general, Indonesian variants formed separate clusters on the basis of single base differences found in both L1 and L2 genes. Characterization of the L1 & L2 sequence variation from the high-risk HPV group of the Indonesian variants can provide further information of their immunogenic properties, biomarkers potential and future development

of novel therapeutic agents against cervical cancer.

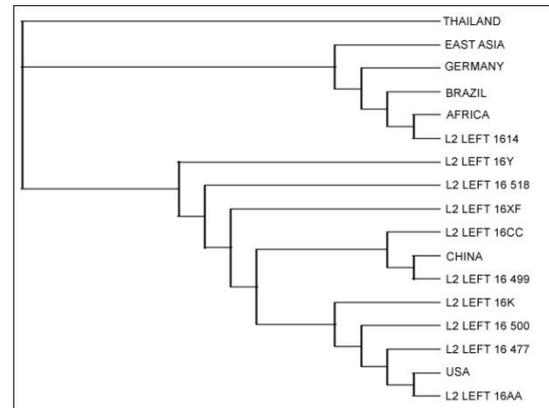


Figure 8. Phylogenetic analysis of upstream L2 HPV-DNA sequence

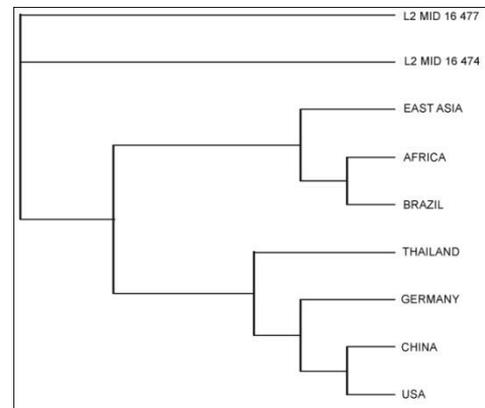


Figure 9. Phylogenetic analysis based on SNP of the middle L2 sequences

Conclusion

It was understood that L1 gene represented the best conserved part of HPV genome because of its major contribution in the physical structure of the capsid protein so that the main integrity of the virus was maintained. L2 gene on the other hands, showed intraspecific DNA variability which was found in most HPV type 16 of Indonesian variants. These facts opens up possibilities of diagnostic methods as well as novel vaccine containing immunogenic determinants which are developed based on L1 and L2 of HPV of Indonesian variants. The new generation of HPV vaccine has also been prepared from modified L1 & L2 regions generating pseudovirions or virus-like particle (VLP).

Acknowledgements

We are gratefully acknowledged the research grant from Competitive Project (2010) and Research Centre for Biotechnology, the Indonesian Institute of Sciences (LIPI). We also thank to Drs. Ratu Safitri and Supartini Sjarif, for fruitful discussion.

References

- Bekkers RL, Massuger LF, Bulten J, & Melchers W J. 2004. Epidemiological and clinical aspects of human papillomavirus detection in the prevention of cervical cancer. *Rev Med Virol* 14: 95–105.
- Bernard HU, Calleja IE, & Dunn ST. 2006. Genome Variation of Human Papillomavirus Types: Phylogenetics and Medical Implications. Department of Molecular Biology and Biochemistry, University of California, USA.
- Burd EM. 2003. Human papillomavirus and cervical cancer. *Clin Microbiol Rev* 16: 1–17.
- de Villiers EM, Fauquet C, Broker TR, Bernard HU, & Zur HH. 2004. Classification of Papillomaviruses. *Virology* 324: 17–27.
- U.S. Food and Drug Administration (FDA) (2008-09-12). FDA Approves Expanded Uses for Gardasil to Include Preventing Certain Vulvar and Vaginal Cancer. Press release <http://www.fda.gov/NewsEvent/Newsroom/PressAnnouncements/2008/ucm116945.htm> Retrieved 2009-11-11.
- Harper DM, Franco EL, Wheeler C, Ferris DG, Jenkins D, Schuind A, Zahaf T, Innis B, Naud P, De Carvalho NS, Roteli-Martins CM, Teixeira J, Blatter MM, Korn AP, Quint W, & Dubin G. 2004. Efficacy of A Bivalent L1 Virus-Like Particle Vaccine in Prevention of Infection with Human Papillomavirus Types 16 and 18 in Young Women: A Randomised Controlled Trial. *Lancet* 364: 1757–1765
- Ishii Y, Ozaki S, Tanaka K, & Kanda T. 2005. Human Papillomavirus 16 Minor Capsid Protein L2 Helps capsomeres assemble Independently of Intercapsomeric Disulphide Bonding. *J Virol* 162: 8640.
- Kieback E & Muller M. 2006. Factors Influencing Subcellular Localization of the Human Papillomavirus L2 Minor Structural Protein. *J Virology* 47(1): 9-16.
- Koutsky LA, Ault KA, Wheeler CM, Brown DR, Barr E, Alvarez FB, Chiacchierini LM, & Jansen KU. 2002. A Controlled Trial of a Human Papillomavirus Type 16 Vaccine. *N Engl J Med* 347: 1645–1651.
- Lowy DR & Schiller JT. 2006. Prophylactic Human Papillomavirus Vaccines. *J Clin Invest* 116 (5): 1167–73.
- Mao F, Leung WY, & Xin X. 2007. Characterization of EvaGreen and the Implication of Its Physicochemical Properties for qPCR Applications. *J BMC Biotechnol* 7: 76.
- Motoyama S, Ldines-Ilave CC, Villanueva SL, & Maruo T. 2004. Role of Human Virus in the Molecular Biology of Cervical Carcinogenesis. *J Med Sci* 50(1): 9-19.
- Munoz N. 2000. Human Papilloma Virus and Cancer: Epidemiological Evidence. *Clin Virol*. 19: 15.
- Munoz N, Bosch FX, de Sanjose S, Herrero R, Castellsague X, Shah KV, Snijders PJ, and Meijer CJ. 2003. Epidemiologic classification of human papillomavirus types associated with cervical cancer. *N Engl J Med* 348: 518–527.
- Novel SS, Safitri R, & Nuswantara, S. 2010. Kanker Serviks dan Infeksi Human Papillomavirus (HPV). Javamedia Network, Jakarta. xxii+206 pp
- Pereira R, Hitzeroth I, & Rybicki EP. 2009. Insights into the Role and Function of L2, the Minor Capsid Protein of Papillomaviruses. *Arch Virol* 154(2): 187-97.
- Prado JC, Calleja-Macias IE, Bernard HU, Kalantari M, Macay SA, Allan B, Williamson AL, Chung LP, Collins RJ, Zuna RE, Dunn ST, Ortiz-Lopez R, Barrera-Saldana HA, Cubie HA, Cuschieri K, Von Knebel-Doeberitz M, Sanchez GI, Bosch FX, & Villa FL. 2005. Worldwide Genomic Diversity of the Human Papillomaviruses-53, 56 and 66, a Group of High Risk HPVs unrelated to HPV-16 and HPV-18. *Virology* 340:95-104.
- Schiller JT & Davies P. 2004. Delivering on the promise: HPV vaccines and cervical cancer. *Nat Rev Microbiol* 2: 343–347.
- Villa LL, Costa RL, Petta CA, Andrade RP, Ault KA, Giuliano AR, Wheeler CM, Koutsky LA, Malm C, Lehtinen M, Skjeldestad FE, Olsson SE, Steinwall M, Brown DR, Kurman RJ, Ronnett BM, Stoler MH, Ferenczy A, Harper DM, Tamms GM, Yu J, Lupinacci L, Railkar R, Taddeo FJ, Jansen KU, Esser MT, Singhs HL, Saah AJ, & Barr E. 2005. Prophylactic quadrivalent human papillomavirus (types 6, 11, 16, and 18) L1 virus-like particle vaccine in young women: a randomised double-blind placebo-controlled multicentre phase II efficacy trial. *Lancet Oncol* 6: 271–278.

Wang M, Zhang MC, Xu XM, Chen LF, & Song GX. 2001. Sequence Analysis of the Late Region of Human Papillomavirus type 6 Genome. *J Biophysics* 7(1): 7-20.

Zheng ZM & Baker CC. 2006. Papillomavirus Genome Structure, Expression, and Post-Transcriptional Regulation. *Front Biosci* 11: 2286–2302.