Basic Medical Research

Phylogenetic analysis and predicted functional effect of protein mutations of E6 and E7 HPV16 strains isolated in Indonesia

Dwi Wulandari,[1,3] Lisnawati Rachmadi,[2,4] Tjahjani M. Sudiro[3,4]

[1] Research Center for Biotechnology, Indonesian Institute of Sciences-LIPI, Bogor, Indonesia

[2] Department of Anatomical Pathology, Faculty of Medicine, Universitas Indonesia, Cipto Mangunkusumo Hospital, Jakarta, Indonesia

[3] Department of Microbiology, Faculty of Medicine, Universitas Indonesia, Cipto Mangunkusumo Hospital, Jakarta, Indonesia

[4] Center for Research and Integrated Development of Tropical Health and Infectious Diseases, Universitas Indonesia, Jakarta, Indonesia

ABSTRAK

Latar belakang: *Protein E6 dan E7 merupakan onkoprotein HPV. Variasi alami protein E6 HPV16 dapat mengubah potensi karsinogeniknya. Penelitian ini bertujuan untuk menganalisis filogenetik gen dan protein E6 dan E7 HPV16 dari Indonesia dan mempredikasi efek substitusi asam amino tunggal terhadap fungsi protein. Analisis ini digunakan sebagai penapis awal dalam pemilihan protein untuk diuji coba secara in vitro ataupun in vivo sehingga dapat menghemat waktu, tenaga, dan biaya penelitian.*

Metode: *Data sekuen gen E6 dan E7 yang digunakan berasal dari 12 sampel isolat Indonesia yang dibandingkan dengan HPV16R (prototipe) dan 6 isolat standar yang merupakan varian HPV16 dengan kategori cabang Eropa (E), Asia (As), Asia-Amerika (AA), Afrika-1 (Af-1), Afrika-2 (Af-2), dan Amerika utara (NA) yang diperoleh dari* Genbank*. Sekuen DNA dianalisis menggunakan BioEdit v.7.0.0 untuk mendapatkan komposisi dan substitusi asam amino tunggal. Analisis filogenetik gen dan protein E6 dan E7 menggunakan* software Clustal X *(1.81) dan NJPLOT. Analisis efek subtitusi asam amino tunggal menggunakan metode SNAP.*

Hasil: *Isolat Java dan ui66* termasuk dalam cabang Eropa sedangkan yang lain termasuk cabang Asia dan Afrika. Terdapat 12 mutasi asam amino pada protein E6 dan 1 mutasi pada protein E7. Analisis SNAP menghasilkan 2 mutasi* non netral*, yaitu R10I dan C63G pada protein E6. Mutasi R10I ditemukan pada AF472509 (Af-2) dan isolat Indonesia yaitu Afrika-2 (Af2*). Mutasi C63G ditemukan hanya pada isolat Af2*.*

Kesimpulan: *Protein E6 lebih bervariasi dibanding dengan protein E7. Analisis metode SNAP menunjukkan hanya protein E6 dari cabang Afrika-2 yang diprediksi memiliki perbedaan fungsional protein dibandingkan dengan HPV16R.*

ABSTRACT

Background: E6 and E7 are oncoproteins of HPV16. Natural amino acid variation in HPV16 E6 can alter its carcinogenic potential. The aim of this study was to analyze phylogenetically E6 and E7 genes and proteins of HPV16 from Indonesia and predict the effects of single amino acid substitution on protein function. This analysis could be used to reduce time, effort, and research cost as initial screening in selection of protein or isolates to be tested *in vitro* or *in vivo*.

Methods: In this study, E6 and E7 gene sequences were obtained from 12 samples of Indonesian isolates, which were compared with HPV16R (prototype) and 6 standard isolates in the category of European (E), Asian (As), Asian-American (AA), African-1 (Af-1), African-2 (Af-2), and North American (NA) branch from Genbank. Bioedit v.7.0.0 was used to analyze the composition and substitution of single amino acids. Phylogenetic analysis of E6 and E7 genes and proteins was performed using Clustal X (1.81) and NJPLOT softwares. Effects of single amino acid substitutions on protein function of E6 and E7 were analysed by SNAP.

Results: Java variants and isolate ui66* belonged to European branch, while the others belonged to Asian and African branches. Twelve changes of amino acids were found in E6 and one in E7 proteins. SNAP analysis showed two non neutral mutations, i.e. R10I and C63G in E6 proteins. R10I mutations were found in Af-2 genotype (AF472509) and Indonesian isolates (Af2*), while C63G mutation was found only in Af2*.

Conclusion: E6 proteins of HPV16 variants were more variable than E7. SNAP analysis showed that only E6 protein of African-2 branch had functional differences compared to HPV16R.

Keywords: cervical cancer, E6 & E7 gene, E6 & E7 proteins, HPV16 variant, phylogenetic, SNAP

pISSN: 0853-1773 • eISSN: 2252-8083 • http://dx.doi.org/10.13181/mji.v24i4.1197 • Med J Indones. 2015;24:197–205
• Received 02 Feb 2015 • Accepted 27 Nov 2015

Correspondence author: **Dwi Wulandari**, wulandedy@yahoo.com

Human papillomaviruses (HPVs) can induce hyperproliferation of the skin and mucosal epithelia causing benign and malignant tumors. HPVs belong to the family *Papillomaviridae*, a deoxyribose nucleic acid (DNA) virus with a genome size 7900 bp. HPV genome consists of early genes (E1, E2, E4, E5, E6, and E7), late genes (L1 and L2) and the long control region (LCR) for regulatory functions. Among the HPV proteins, E6 and E7 proteins of HR-HPV are oncoproteins required to induce malignancy. E6 and E7 are small proteins disrupting important cellular pathway, p53 and pRB, respectively, involved in apoptosis and cell cycle regulation.[1-3]

Factors related to the virus contribute to progression of the infection to cancer, such as HPV type involved in the infection, viral variants, persistance and viral load. Currently, more than 120 different types of HPV have been identified featuring distinct tropisms for different body sites (skin, mouth, and genitalia).[4,5] About 40 genotypes are involved in anogenital tract infection, causing various diseases ranging from genital warts to invasive cancer. HPV types 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68, 73, and 82 have been considered as high risk (HR-HPV). HPV type 16 and 18 are the most common cause of cervical cancer. HPV16 is related to 50% of cervical cancer and cancer precursor lesions in the world and 38% in Indonesia.[6,7]

Human papillomavirus isolates historically were classified into types by comparing their DNA with a set of HPV reference genomes. Isolates that belong to the same HPV type are closely related to one another on the basis of nucleotide sequence and represent a discrete phylogenetic group.[3,8] Today, different types of HPV are defined as having more than 10% variation in L1 gene. Viruses differing by 2–10% are called subtypes and are infrequently observed, while those differing by <2% are more frequently seen and called intratypic variants.[9]

The intratypic evolutionary variant studies were initially inferred from the partial upstream regulatory region (URR) and E6 sequences, and have been recently expanded to include the complete genomes. HPV16 variants are geographically distributed. Yamada et al[10] have categorized HPV16 variants into six phylogenetic branches namely: European (E), Asian (As), Asian-American (AA), African-1 (Af-1), African-2 (Af-2), and North American (NA). The prototype variant named HPV16R (K02718), which was reported by Seedorf et al[8] was isolated from Germany is included in the European branch.

The description and understanding of HPV genome variants is an important area for molecular pathogenesis and for the development of molecular diagnostics for HPV, vaccines and other therapeutic approaches to control and/or eliminate virus-induced diseases. Despite phylogenetic relatedness, HPV variants can differ in pathogenicity. There is a three-fold or greater risk of cervical cancer for Asian-American (AA) or African (Af) HPV16 variants compared to European (E) variants; and, non-European variants of HPV16 may be more common in cancer tissues and high-grade cervical lesions.[11-13]

HPV16 variants' data in Indonesia are still limited. de Boer et al[14] reported the presence of a variant of Java from samples taken at Cipto Mangunkusumo Hospital, Jakarta, and found that Java variant (JV) was dominant compared to other variants (15 of 22 samples). JV having mutation in the E6 open reading frame (ORF; 276G), gave the predicted amino acid change N58S.[14]

The changes in gene sequences may cause changes in biological function of the protein encoded by these genes, which may affect the course of the disease. Phylogenetic analysis of genetic and protein variation especially E6 and E7 of HPV16 variants is the essential factor to determine differences in biological properties of the virus and to be used for studies of infectivity and pathogenicity.[15,16]

Screening for non-acceptable polymorphisms (SNAP) method has been developed by de Boer et al[14] to determine the predictive effect of a single non-synonymous nucleotide polymorphism to the function of the protein compared with the wild-type protein. Non-synonymous single nucleotide polymorphisms (SNPs) are 'neutral' if the resulting point-mutated protein is not functionally discernible from the wild type and 'non-neutral' otherwise. The ability to identify non-neutral substitutions could significantly aid in targeting disease causing detrimental mutations, as well as SNPs that increase the fitness of particular phenotypes.[17-19] SNAP analysis is useful to guide selection of isolates

to be tested experimentally *in vitro*, especially in resource-limited laboratory.

Despite the prevalence of HPV16, analysis of E6 and E7 proteins of HPV16 variants from Indonesia has not been reported. Therefore in this work, we analyzed phylogenetic of E6 and E7 gene and protein and used SNAP to predict the effect of a single non-synonymous nucleotide polymorphisms on the function of the protein of HPV16 isolates from Indonesia compared to the existing HPV reference in GenBank. This analysis could be used to reduce time, effort, and cost research as initial screening in selection of protein or isolates to be tested *in vitro* or *in vivo*.

METHODS

The laboratory work was done for six months in 2012 at Microbiology Department, Faculty of Medicine, Universitas Indonesia, Jakarta, Indonesia. The protocol of this study has been approved by Medical Ethics Committee of Universitas Indonesia (No.278/PT02.FK/ETIK/2010).

Data collection

Sequence data of E6 and E7 genes in this paper used 12 samples of Indonesian isolates which were 11 samples (Java1*, Java2*, Java276G1*, Java276G2*, Java276G3*, Af2*, As647G1*, As647G2*, E647G*, E350G*, E109C*) obtained from data published by de Boer et al[14] and one sample of ui66*. Indonesian isolates were compared with the data of HPV16R (prototype) and six standard variant isolates of HPV16 in the category of European (E), Asian (As), Asian-American (AA), African-1 (Af-1), African-2 (Af-2), and North American (NA) branches obtained from the Genbank (Table 1).

Sample and DNA isolation

ui66 DNA was isolated from cytology samples from cervical swab of a patient with cervical intraepithelial neoplasia (CIN) two obtained from Department of Pathological Anatomy, Faculty of Medicine, Universitas Indonesia, Jakarta in 2012. The sample ui66* was tested by Linear Array® HPV Genotyping Test (Roche) method, and was found containing HPV16. DNA was isolated by QIamp DNA Blood Mini Kit (Qiagen) method.

PCR procedure

Primers HPV16–7531F 5'GTA CGT TTC CTG CTT GCC AT 3' and HPV16–2450R 5'CAA CAG GGC ACT GTA GCA TC 3' were used to amplify fragments of ui66* DNA. Polymerase chain reaction (PCR) mixture contained: 1 µl DNA isolate, 1 x KOD Plus Neo DNA polymerase (Toyobo) buffer, 1.5 mM $MgSO_4$, 0.2 mM dNTPs, 0.3 U enzyme KOD Plus Neo DNA polymerase (Toyobo) and 0.2 mM each primer for total a volume of 15 µl. Amplifications were performed with the following reaction condition: 95°C five minutes, followed by 40 cycles of 30 second denaturation at 95°C, 30 second annealing at 52°C, and a two minutes 30 second elongation at 68°C. The last cycle was followed by a final extension step of seven minutes at 68°C.

DNA sequencing

Primers HPV16–7811F 5'GTT AGT CAT ACA TTG TTC ATT TGT 3' and HPV16–552R 5'CAA CAA GAC ATA CAT CGA CCG3', HPV16–368F 5'ACA GCA ATA CAA CAA ACC GTT 3', and HPV16–1664R 5'GCA GCA ATA CAC CAA TCG CA 3' were used for DNA sequencing. The sequencing was done at Eijkman

Table 1. List of HPV16 reference strains representing variants of HPV16 used as standard in phylogenetic and SNAP analysis

Variants	Genbank accesion number	Year of discovery	Isolated from
Reference (HPV16R)	K02718	1985	Germany
Africa 1 (Af1)	AF472508	2003	Africa
Africa 2 (Af-2)	AF472509	2003	Africa
Europe (E)	U118173	2007	Germany
Asia (As)	AF534061	2003	East Asian
Asia-america (AA)	AY686579	2005	Costa Rica
North America (NA)	AF125673	2000	Unknown

(Modified from Lurchachaiwong et al)[20]

Institute for Molecular Biology, Jakarta, by the dye-deoxy terminator method on a 310 ABI Prism Sequencer (PE Applied Biosystems) using 5 pmol of either forward or back primers.

Phylogenetic analysis

Deoxy-ribonucleic acid (DNA) sequence of ui66* was compared to data of E6 and E7 gene sequences of HPV16 from Indonesia published by de Boer et al.[14] HPV16 reference strains representing variants of HPV16 was used as standard in phylogenetic. SNAP analysis referred to research of Lurchachaiwong et al[20] (table 1), and the sequence data were obtained from the national center for biotechnology (NCBI). DNA sequences were analyzed using BioEdit v.7.0.0 to get the amino acid composition and single amino acid substitution.[21] A multiple sequenced alignment of genes and proteins was perform using ClustalX1.81.[22] Then dnd. file were exported to obtain phylogenetic tree by NJPLOT V 2.3.[23]

SNAP analysis

Effects of single amino acid substitutions on protein function of E6 and E7 were analyzed by SNAP feature.[17]

RESULTS

Phylogenetic tree analysis in figure 1 and 2 showed the branches of HPV16 variants. Results of analysis of both E6 and E7 genes and proteins showed different patterns of phylogenetic branch. For example E6 gene analysis results showed Java 276G1*, 2*, 3* were sublineages of HPV16R, while the E6 protein analysis results were on a different branch of HPV16R.

Table 2. Nucleotide sequence variations among the HPV16 isolates relative to reference HPV16 (HPV16R)

Variant	Nucleutide position of E6																									Nucleotide position of E7						
	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	3	3	3	4	4	4	4	5	5	6	6	6	7	7	8	8
	0	3	3	3	4	4	5	7	7	7	8	1	5	7	8	8	1	3	5	0	0	3	8	2	3	1	4	6	8	9	4	4
	9	1	2	5	3	5	3	3	6	8	8	9	6	6	6	9	0	5	0	2	3	3	8	2	2	8	7	6	9	5	3	6
HPV16R (E) (Ref)	T	A	G	A	C	G	C	C	G	T	G	G	C	A	T	A	T	C	T	A	A	G	A	A	A	A	A	G	T	T	T	T
ui66*																																
AY686579 (AA)						T				A					A	G		T	G					G								
AF125673 (NA)													T		A				G													
AF472508 (AF-1)			C		G	T									A	G		T								G						
AF472509 (AF-2)	C		T	G	T								T				G			G												
EU118173 (E)																																
AF534061 (AS)										G																						C
Java1*	C																											A				
Java2*																												A				
Java276G1*									A					G														A				
Java276G 2*														G														A				
Java276G3*														G													G	A				
Af2*	C		T		G	T									A	G		T		G	G						G		C	G		
As647G1*										G																	G					C
As647G2*										G																	G				C	C
E647G*																											G					C
E350G*																			G													
E109C*	C										C																					

Note: E6 and E7 nucleotide positions at which variations observed are written vertically across. E: European, As: Asian, Java: Javanese, Af1: African1, Af2: African-2, NA: North-American, and AA: Asian–American with their accession number. Indonesian isolates coded with symbol (*)

Table 3. Analysis E6 and E7 using SNAP method

SNAP analysis mutation	E6 protein												E7 protein
	R10T	R10I	Q14H	Q14D	D25E	D25N	L27I	E29Q	N58S	C63G	H78Y	L83V	N29S
	Net	Non	Net	Net	Net	Net	Net	Net	Net	Non	Net	Net	Net
AY686579 (AA)			√				√				√	√	
AF125673 (NA)							√					√	
AF472508 (AF-1)	√			√			√				√		
AF472509 (AF-2)		√		√			√				√		√
EU118173 (E)							√					√	
AF534061 (AS)					√		√						√
Java1*													
Java2*													
Java276G1*						√			√				
Java276 G2*									√				
Java276G3*									√				√
Af2*		√	√							√	√		√
As647G1*					√								√
As647G2*					√								√
E647G*													√
E350G*												√	
E109C*								√					
ui66*													

Note: E6 and E7 amino acid mutation comparison to HPV16R at which the variations were observed are written vertically across the top and the SNAP analysis result, net: neutral, non: non-neutral

Comparison analysis between E6 to E7 genes produces different patterns of branching. Phylogenetic tree of E6 and E7 genes showed ui66* and Java isolates were included in the European branches except Java276G3* in figure 2 that was showed in a separate branch. This suggests the possibility that the isolates derived from Europe.

While based on phylogenetic tree E6 and E7 proteins showed ui66*, Java1*, Java2* into the European branch, Java276G1*, 2*, 3* were closer to Asia. Phylogenetic analysis of E6 genes and proteins revealed genetic divergence, while there was only one nucleotide mutation in the E7 gene that cause amino acid changes. Thus the phylogenetic tree of the E7 proteins revealed into two branches only.

Mutations in E6 and E7 proteins as can be seen in table 2 was obtained by Clustal X software analysis. There were 12 mutations of amino acids of protein E6 and one mutation of E7 proteins. These mutation has been analyzed using SNAP method and showed two non-neutral mutations, R10I and C63G in E6 proteins. R10I mutations were found in AF472509 (Af-2) and in the Indonesian isolates belonged to African-2 branch (Af2*). The C63G mutation was found only in Indonesian isolate which belonged to Af2* (table 3). Thus suggests that only E6 proteins of the African-2 branch had functional differences compared to HPV16R. It is necessary to do laboratory tests *in vitro* or *in vivo* to determine the implications as carcinogenic agents.

DISCUSSION

Phylogenetic analysis is the study of evolutionary relationships among molecules, phenotypes, and organisms.[11] The physicochemical properties of nucleic acids or amino acids are crucial factors that affect their structures or functions, thus it may provide function prediction, and in relation with pathogens, it may relate to pathogenicity.[21]

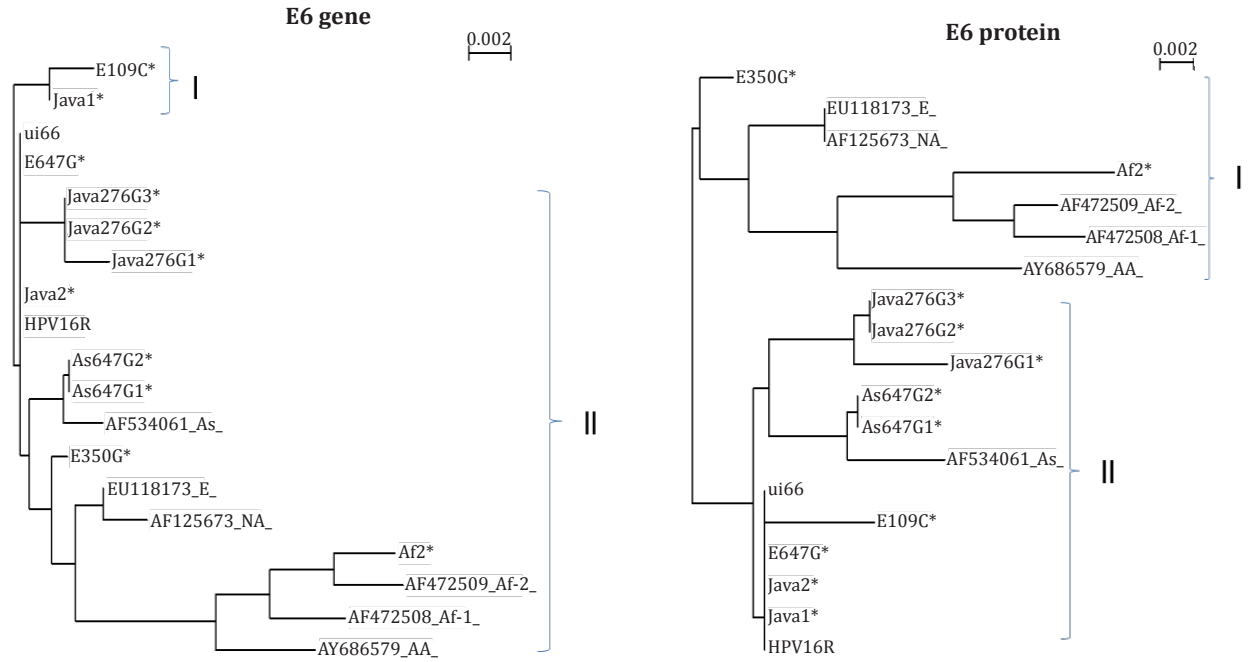

Figure 1. Comparative phylogenetic trees of E6 gene and protein of different HPV16 isolates. E6 genes and proteins formed branching with different pattern and member

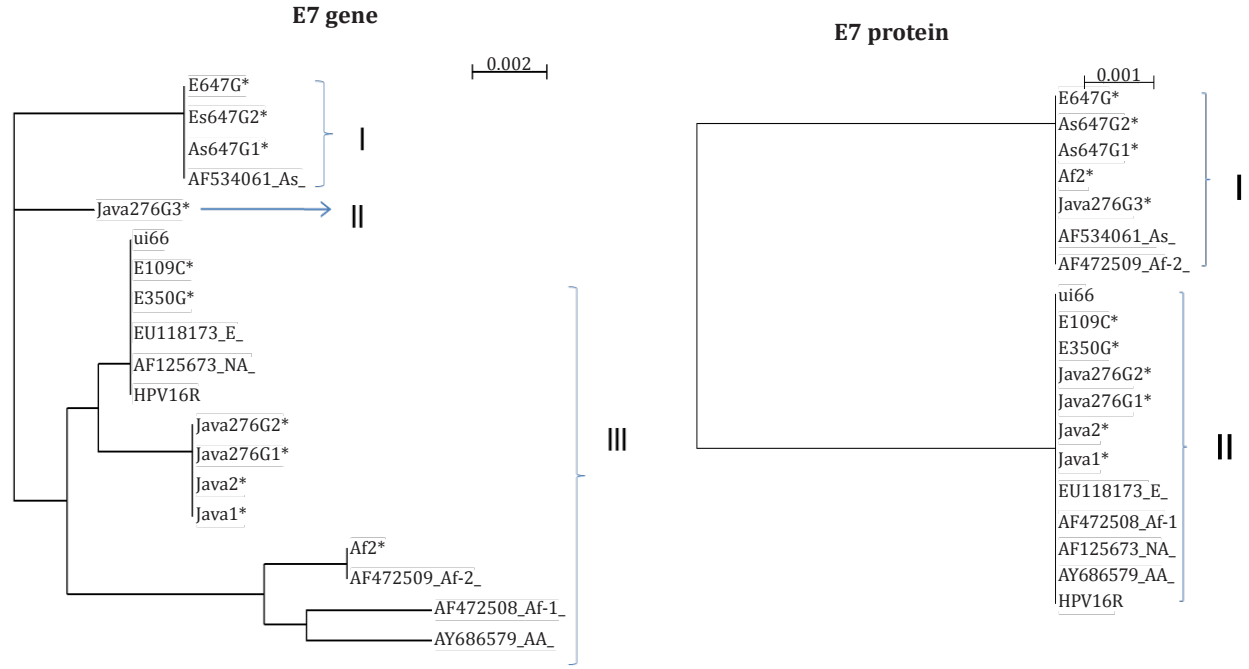

Figure 2. Comparative phylogenetic trees of E7 gene and protein of different HPV16 isolates. E7 genes and proteins formed branching with different pattern and member

Relationship analysis of some isolates, as could be seen from phylogenetic tree, could give different results when different genes or proteins were used as point of view. Phylogenetic tree result of genes and proteins of E6 and E7 showed different pattern relationship for example Java276G3* showed inconsistency of which branch it belong to. de Boer et al[14] revealed that Javanese variant would belong to European branch but it was argumented to define these variants as a

separate branch. So it is difficult to choose which phylogenetic tree can be used as a benchmark. In this paper, variants other than E were also found i.e. As, Af-1 and Af-2 although in small amounts.

However, for functional studies it may be more appropriate to use protein sequences for such analysis. The fundamental building blocks of life are proteins. The catalysts of virtually all of the chemical transformations in the cell, are proteins. The functional properties of proteins is determined by the sequence of the 20 amino acids. For protein-encoding genes, the object on which natural selection acts is the protein.[25]

Determination of related phylogenetic classification or naming isolates should be based on one of the most conserved gene. The result of International Workshop of Papillomavirus in Quebec in 1995 have agreed to use the L1 gene to determine the types of HPV.[4] Further studies discovered that there are variants in one type of HPV that related to their pathogenicity and distribution.[1,26]

Human papilomavirus variant data are important in developing HPV diagnostics, vaccines, and other therapeutic approaches to control virus-induced diseases. HPV16 variants have been shown to have different biological as well as biochemical effects, resulting in altered oncogenic potentials. The oncogenicity of distinct HPV variants may also differ between geographical regions because of population history and host genetics, such as the difference in the distribution of HLA in the population.[27,28]

In this study, branching pattern of phylogenetic trees did not match with the results of Yamada et al.[1] For example, the NA should be closer to AA than E. It is possible that our analysis used only E6 or E7 gene alone, while Yamada et al[1] used E6, L1, L2 and LCR to determine HPV16 variant branches.

The results of phylogenetic analysis of HPV16 variants from Indonesia showed that Java and ui66* isolates were included in the European branches. de Boer et al[14] revealed that Java isolate was found dominantly (15 of 22 samples) and Java variants belonged to European branch, which was not in accordance with Hildesheim and Wang,[9] stating that there was greater risk of cervical cancer for non-European variants of HPV16.[9] It could be because of the lack of sample number. The majority of isolates analysis in this study were isolated in year 2004. Since there are not many published data on recent HPV isolates from Indonesia, to know the recent distribution of HPV variants in Indonesia needs further study.

Screening for non-acceptable polymorphisms (SNAP) prediction of E7 proteins from all isolates showed no change in protein functional properties. The non-neutral (R10I & C63G) mutations were found only in E6 protein of the African-2 branch. It is estimated that these mutations alter protein function, but actually to determine how this mutation change the carcinogenicity of the protein is still unknown. Of the two HPV16 E6 and E7 oncogenes, E6 has been found to show more variation than E7, which is relatively conserved.[17] Due to the fact that several codons can code a same amino-acid, the amino-acid sequence is usually more conserved than the nucleotide sequence

Genetic studies had reported correlation between specific HPV16 variants and persistent viral infection, followed by the development of malignant lesions.[29-31] Non-European variants were found to be associated with an excess risk of cervical cancer.[32,33] Lichtig, et al[32] experiment showed that protein E6 of HPV16 variants naturally had a different activity on the carcinogenic potential. The observations were carried out about the activity of p53 degradation, Bax degradation, inhibition of p53 transactivation, binding to E6BP, and hDIg. The L83V variant, previously associated with the risk of cancer development in some populations, had the same level of activity as prototype in some functional tests. On the other hand, L83V showed more efficient degradation of Bax and binding E6BP, but lower binding to hDlg.[34]

To determine the implications of oncoproteins, HPV research needs laboratory experiment both *in vitro* and *in vivo*. The use of this software analysis is only to predict mutations that may cause functional changes in protein so it can be as early screening for the selection of the sample to be tested *in vitro* and *in vivo*.

In conclusion, E6 genes and proteins of HPV16 variants have been found more varied than E7. SNAP method of analysis results showed that only

the E6 protein of African-2 branch had functional differences compared to HPV16R. This prediction has to be proved by laboratory experiments to determine the effects of the mutation as carcinogen.

Acknowledgment

The authors would like to acknowledge Ulima Darmania Amanda, S.Si at the Department of Pathological Anatomy, Faculty of Medicine, Universitas Indonesia, Jakarta for maintaining material sample. Financial support for this study was provided by PHKI CRID-TROPHID FKUI 2012.

Conflict of interest

The authors affirm no conflict of interest in this study.

REFERENCES

1. Yamada T, Manos MM, Peto J, Greer CE, Munoz N, Bosch FX, et al. Human papillomavirus type 16 sequence variation in cervical cancers: a worldwide perspective. J Virol. 1997;71(3):2463–72.
2. Narisawa-Saito M, Kiyono T. Basic mechanisms of high-risk human papillomavirus-induced carcinogenesis: roles of E6 and E7 proteins. Cancer Sci. 2007;98(10):1505–11.
3. Smith B, Chen Z, Reimers L, van Doorslaer K, Schiffman M, Desalle R, et al. Sequence imputation of HPV16 genomes for genetic association studies. PLoS One. 2011;6(6):e21375.
4. de Villiers EM, Fauquet C, Broker TR, Bernard HU, zur Hausen H. Classification of papillomaviruses. Virology. 2004;324(1):17–27.
5. 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.
6. Muñoz N, Bosch FX, de Sanjosé S, Herrero R, Castellsagué X, Shah KV, et al. Epidemiologic classification of human papillomavirus types associated with cervical cancer. N Engl J Med. 2003;348(6):518–27.
7. hpvcentre.net [internet]. Spain: ICO Information Centre on HPV and Cervical Cancer (HPV Information Centre), [update 2014 Dec 18]. Available from: http://www.hpvcentre.net/statistics/reports/XWX.pdf
8. Seedorf K, Krämmer G, Dürst M, Suhai S, Röwekamp WG. Human papillomavirus type 16 DNA sequence. Virology. 1985;145(1):181–5.
9. Hildesheim A, Wang SS. Host and viral genetics and risk of cervical cancer: a review. Virus Res. 2002;89(2):229–40.
10. Yamada T, Wheeler CM, Halpern AL, Stewart AC, 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(12):7743–53.
11. 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.
12. Burk RD, Chen Z, Harari A, Smith BC, Kocjan BJ, Maver PJ, et al. Classification and nomenclature system for human alphapapillomavirus variants: general features, nucleotide landmarks and assignment of HPV6 and HPV11 isolates to variant lineages. Acta Dermatovenerol Alp Pannonica Adriat. 2011;20(3):113–23.
13. Cornet I, Gheit T, Franceschi S, Vignat J, Burk RD, Sylla BS, et al. Human papillomavirus type 16 genetic variants: phylogeny and classification based on E6 and LCR. J Virol. 2012.86(12):6855–61.
14. de Boer MA, Peters LA, Aziz MF, Siregar B, Cornain S, Vrede MA, et al. Human papillomavirus type 16 E6, E7, and L1 variants in cervical cancer in Indonesia, Suriname, and The Netherlands. Gynecol Oncol. 2004;94(2):488–94.
15. Daf S, Jena L, Kumar S. Comparative phylogenetic analysis of E6 and E7 proteins of different 42 strains of HPV. JK Science. 2010;12(1):6–10.
16. Pientong C, Wongwarissara P, Ekalaksananan T, Swangphon P, Kleebkaow P, Kongyingyoes B, et al. Association of human papillomavirus type 16 long control region mutation and cervical cancer. Virol J. 2013;10(30):1–9.
17. Bromberg Y, Rost B. SNAP: predict effect of non-synonymous polymorphisms on function. Nucleic Acids Res. 2007;35(11):3823–35.
18. Katsonis P, Koire A, Wilson SJ, Hsu TK, Lua RC, Wilkins AD, et al. Single nucleotide variations: biological impact and theoretical interpretation. Protein Sci. 2014;23(12):1650–66.
19. Adeel MM, Qasim M, Ashfaq UA, Masoud MS, ur Rehman M, ul Qamar MT, et al. Modelling and simulation of mutant alleles of breast cancer metastasis suppressor 1 (BRMS1) gene. Bioinformation. 2014;10(7):454–9.
20. Lurchachaiwong W, Junyangdikul P, Payungporn S, Chansaenroj J, Sampathanukul P, Tresukosol D, et al. Entire genome characterization of human papillomavirus type 16 from infected Thai women with different cytological findings. Virus Genes. 2009;39(1):30–8.
21. Bioedit software. (internet).available from http://bioedit.software.informer.com
22. ClustalX software. (internet).available from http://www.clustal.org/download/1.X/ftp-igbmc.u-strasbg.fr/pub/ClustalX
23. NJPLOT software. (internet). available from http://en.freedownloadmanager.org/Windows-PC/NJPLOT-FREE.html
24. Zhang S, Wang T. Phylogenetic analysis of protein sequences based on conditional LZ complexity. MATCH Commun. Math. Comput. Chem. 2010;63:701–16.
25. Oppedoes F. Phylogenetic analysis using protein sequences In: Salemi M, Vandamme A-M, editors. The phylogenetic handbook: a practical approach to DNA and protein phylogeny. Cambridge: Cambridge University Press. 2003; p. 210–43.

26. Giannoudis A, Herrington CS. Human papillomavirus variants and squamous neolasia of the cervix. J Pathol. 2001;193(3):295–302.
27. 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.
28. Metcalfe S, Roger M, Faucher MC, Coutlée F, Franco EL, Brassard P. The frequency of HLA alleles in a population of Inuit women of northern Quebec. Int J Circumpolar Health. 2013;5(72):1–4.
29. Berumen J, Ordoñez RM, Lazcano E, Salmeron J, Galvan SC, Estrada RA, et al. Asian-American variants of human papillomavirus 16 and risk for cervical cancer: a case-control study. J Natl Cancer Inst. 2001;93(17):1325–30.
30. Hildesheim A, Schiffman M, Bromley C, Wacholder S, Herrero R, Rodriguez A, et al. Human papillomavirus type 16 variants and risk of cervical cancer. J Natl Cancer Inst. 2001;93(4):315–8.
31. Xi LF, Koutsky LA, Hildesheim A, Galloway DA, Wheeler CM, Winer RL, et al. Risk for high-grade cervical intraepithelial neoplasia associated with variants of human papillomavirus types 16 and 18. Cancer Epidemiol Biomarkers Prev. 2007;16(1):4–10.
32. Sichero L, Ferreira S, Trottier H, Duarte-Franco E, Ferenczy A, Franco EL, et al. High grade cervical lesions are caused preferentially by non-European variants of HPVs16 and 18. Int J Cancer. 2007;120(8):1763–8.
33. Freitas LB, Chen Z, Muqui EF, Boldrini NA, Miranda AE, Spano LC, et al. Human papillomavirus 16 non-European variants are preferentially associated with high-grade cervical lesions. 2014. PLoS One. 2014;9(7):e100746.
34. Lichtig H, Algrisi M, Botzer LE, Abadi T, Verbitzky Y, Jackman A, et al. HPV16 E6 natural variants exhibit different activities in functional assays relevant to the carcinogenic potential of E6. Virology. 2006;350(1):216–27.