
Association of Human Papillomavirus Type 58 Variant With the Risk of Cervical Cancer

Paul K. S. Chan, Ching-Wan Lam, Tak-Hong Cheung, William W. H. Li, Keith W. K. Lo, May Y. M. Chan, Jo L. K. Cheung, Augustine F. Cheng

Human papillomavirus (HPV) type 58 has been found to be prevalent among Chinese patients with cervical cancer. This study examined the oncogenic risk of HPV58 variants in Hong Kong, a southern part of China. Altogether, 1924 women were studied: 42.8% with a normal cervix, 16.2% with cervical intraepithelial neoplasia (CIN) I, 12.7% with CIN II, 20.8% with CIN III, and 7.6% with invasive cervical cancer (ICC). The overall prevalence of HPV58 was 11.4% (220) and increased statistically significantly with the severity of neoplasia ($P_{\text{trend}} < .001$, χ^2 test for trend). Among HPV58-positive women, the occurrence of E7 632C→T (T20I) and E7 760G→A (G63S) variants (T20I/G63S) showed a positive trend of association with the severity of neoplasia ($P_{\text{trend}} < .001$, χ^2 test for trend). HPV58 variants carrying these two substitutions showed an odds ratio (OR) for ICC of 26.79 (95% confidence interval = 10.14 to 74.72), and this OR was 6.9-fold higher than the ORs of variants without these substitutions. Patients with CIN III or ICC who were also infected with T20I/G63S variants had a statis-

Affiliations of authors: P. K. S. Chan, J. L. K. Cheung, A. F. Cheng (Department of Microbiology), C.-W. Lam (Department of Chemical Pathology), T.-H. Cheung, K. W. K. Lo (Department of Obstetrics and Gynaecology), The Chinese University of Hong Kong, Prince of Wales Hospital, Hong Kong SAR, China; W. W. H. Li, M. Y. M. Chan, Department of Obstetrics and Gynaecology, Queen Elizabeth Hospital, Hong Kong SAR, China.

Correspondence to: Paul K. S. Chan, MRC Path, Department of Microbiology, The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, New Territories, Hong Kong SAR, China (e-mail: paulkschan@cuhk.edu.hk).

See "Notes" following "References."

© Oxford University Press

tically significant younger age at diagnosis than those infected with other variants (median age = 37 years versus 48 years; $P = .038$, two-sided Mann-Whitney U test). Thus, HPV58 variants carrying E7 T20I/G63S substitutions may be associated with an increased risk for cervical cancer. [J Natl Cancer Inst 2002;94:1249-53]

The E6 and E7 proteins of human papillomaviruses (HPVs) have important functions in the development of cervical cancer, but the implications of sequence variations in E6 and E7 open reading frames (ORFs) are still being disputed (1). Most studies on sequence variation have been based on HPV16, the most prevalent high-risk type of HPV (2-5). Because our previous studies suggest that HPV58, a rare HPV variant worldwide, is prevalent in Hong Kong (6,7), we examined sequence variations in E6 and E7 genes of HPV58 and their risk association with the development of cervical neoplasia.

Subjects were recruited from a colposcopy clinic and were classified according to biopsy examination findings. All subjects gave their written informed consent. Women found to have a normal cervix by colposcopy, hence not receiving a biopsy examination, were regarded as normal, provided that their subsequent Papanicolaou (Pap) smear taken 4-6 months later did not reveal neoplasia. Those women with abnormal or missing follow-up smears were excluded. Histologic findings of inflammatory changes or HPV effects but without neoplasia were also classified as normal. The study was approved by the local Institutional Ethics Committee.

The study included 1924 Chinese women, most of whom were descendants of families from Canton province in the southern part of China. Their mean age was 40.2 years (range = 16-88 years; standard deviation = 10.7 years). Diagnoses among these 1924 women were as follows: 823 (42.8%) with a normal cervix, 311 (16.2%) with cervical intraepithelial neoplasia (CIN)

I, 244 (12.7%) with CIN II, 400 (20.8%) with CIN III, and 146 (7.6%) with invasive cervical cancer (ICC)—including 137 with squamous cell carcinoma, seven with adenocarcinoma, and two with adenosquamous carcinoma. Overall, 220 (11.4%) samples were positive for HPV58 by a type-specific polymerase chain reaction (PCR) that used primers E7-P1/E7-P2 (5'-CTGTAA CAACGCCATGAGAG-3' and 5'-TCAGGGTCATCCATTGCAGA-3'). The positive rate for HPV58 showed a statistically significant trend of increase with the degree of neoplasia (5.1% for normal cervix, 12.2% for CIN I, 17.2% for CIN II, 13.0% for CIN III, and 31.5% for ICC; $P_{\text{trend}} < .001$ by the χ^2 test for trend; Epi Info 2000; Centers for Disease Control and Prevention, Atlanta, GA).

Sequences of E7 and E6 regions were obtained by PCR-based cycle sequencing with E7-P1/E7-P2 primers and E6-P1/E6-P2 primers (5'-GACCGAA ACCGGTGCATATA-3' and 5'-TCT CATGGCGTTGTTACAGG-3'), respectively. Sequencing reactions were

HPV58 variant	No. of isolates	Open reading frame														Predicted amino acid substitution									
		E6							E7							E6		E7							
		1	2	3	3	3	5	5	6	6	7	7	7	7	7	7	7	7	8	8	8	8			
Prototype-like	26	C	G	C	C	A	A	G	C	G	T	T	T	G	G	A	C	C	T	C	T				
E6/E7-HK-1	68	-	-	T	-	-	-	-	-	A	-	G	-	-	A	-	-	-	-	-	-				G41R / G63D
E6/E7-HK-2	58	-	-	T	-	-	-	-	T	-	-	G	-	A	-	-	-	-	-	-	-				T20I / G63S
E6/E7-HK-3	12	-	-	-	-	C	-	-	-	-	-	G	-	-	-	-	-	-	C	-	-	K93N			V77A
E6/E7-HK-4	12	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	C	-	-	K93N			V77A
E6/E7-HK-5	10	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	K93N			
E6/E7-HK-6	6	-	-	-	-	-	-	-	T	-	-	G	-	A	-	-	-	-	-	-	-				T20I / G63S
E6/E7-HK-7	4	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				
E6/E7-HK-8	4	-	C	T	A	C	-	-	-	-	-	G	-	A	-	G	T	A	-	-	-	E32Q / D86E / K93N			G63S / T74A / D76E
E6/E7-HK-9	2	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-				
E6/E7-HK-10	2	-	-	T	-	-	-	A	T	-	-	G	-	A	-	-	-	-	-	-	-				R9K / T20I / G63S
E6/E7-HK-11	2	-	-	T	-	-	-	-	-	A	-	G	C	-	A	-	-	-	-	-	-				G41R / G63D
E6/E7-HK-12	2	-	-	T	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-				
E6/E7-HK-13	2	-	-	T	-	-	-	-	-	A	-	G	-	C	A	-	-	-	-	-	-				G41R / G63H
E6/E7-HK-14	2	-	-	-	-	C	-	-	-	A	-	G	-	-	A	-	-	-	-	-	-	K93N			G41R / G63D
E6/E7-HK-15	2	-	-	T	-	-	C	-	-	A	-	G	-	-	A	-	-	-	-	-	-				G41R / G63D
E6/E7-HK-16	2	-	-	-	-	-	-	-	-	A	-	G	-	-	A	-	-	-	-	-	-				G41R / G63D
E6/E7-HK-17	2	-	-	T	A	-	-	-	-	-	C	G	-	-	G	T	A	-	-	-	-	D86E			T74A / D76E
E6/E7-HK-18	2	T	-	T	A	-	-	-	-	-	-	G	-	-	G	T	A	-	T	C	-	D86E			T74A / D76E

Fig. 1. Nucleotide sequence variations in the E6 and E7 open reading frames (ORFs) of human papillomavirus (HPV) 58 detected in 220 Hong Kong women. HPV58 prototype (GenBank accession number NC_001443) was used as the reference. Nucleotide positions where variations were detected are written vertically across the top. Positions at which no variation was found are marked with dashes. **Lightface type** = a silent nucleotide variation; **boldface type** = a missense variation. The one-letter amino acid code

is used. The position of amino acid change is stated numerically. The letter preceding this number refers to the reference amino acid, and the letter following it refers to its substitution. Prototype-like isolates refer to those without nucleotide sequence deviation from prototype over the E6 and E7 ORFs. GenBank accession numbers for sequences reported in this paper are AF478150-AF478167 for E6 sequences and AF478132-AF478149 for E7 sequences.

performed on the ABI Prism 310 Genetic Analyzer (Applied Biosystems, Foster City, CA), according to the manufacturer's protocol (BigDye Terminator Cycle sequencing kit; Applied Biosystems). All sequences were confirmed by sequencing from both directions.

The HPV58 prototype (GenBank accession no. NC_001443), isolated from a Japanese cervical cancer patient, was used as the reference (8). Six nucleotide sequence variations were found over the 450-base-pair (bp) E6 ORF, and 14 were found over the 297-bp E7 ORF. E6 variations were scattered, whereas E7 variations were mainly found within the carboxyl-terminal residues (amino acids 41–89). Sequence variability was higher for E7 than for E6. Only 10 (4.5%) isolates showed one or more variations at E6, whereas 176 (80.0%) isolates

showed one or more variations at E7. The maximum number of variations for any variant was four (0.9%) for E6 and six (2.0%) for E7. The average chance of having a nucleotide sequence deviation from prototype was two substitutions per 1000 bp for E6 and eight substitutions per 1000 bp for E7. Thus, three of six E6 substitutions and nine of 14 E7 substitutions were nonsynonymous (Fig. 1). Mutations generating a frameshift or a premature stop codon were not observed.

Of the 220 HPV58 isolates, 26 (11.8%) showed complete E6 and E7 sequence homology with the prototype and were thus referred to as "prototype-like." The remaining 194 isolates were grouped into 18 different variants and named according to their frequency of occurrence as E6/E7-HK-1 to E6/E7-HK-18, respectively (Fig. 1).

The 388A→C (K93N) E6 nucleotide substitution showed a statistically significant negative trend of association with the severity of neoplasia ($P_{\text{trend}} < .001$ by the exact test for trend; StatXact 4.0.1; Cytel Software Corp., Cambridge, MA), whereas 632C→T (T20I) and 760G→A (G63S) E7 nucleotide substitutions showed a statistically significant positive trend of association with the severity of neoplasia ($P_{\text{trend}} < .001$ by the exact test for trend) (Table 1, A). E7 T20I and E7 G63S amino acid substitutions were apparently linked, with both found in 66 isolates. Only four isolates from subjects with a normal cervix had the G63S substitution in the absence of T20I. No isolate carried the E7 T20I substitution in the absence of E7 G63S. The HPV58 E6/E7-HK-2 variant that carried the E7 T20I and G63S (T20I/G63S) substitu-

Table 1. Distribution of human papillomavirus (HPV) 58 variant infection status according to the degree of cervical neoplasia*

	No. of women (%)					Total	OR (95% CI)†	$P_{\text{trend}}‡$
	Normal	CIN I	CIN II	CIN III	ICC			
A) HPV58-positive subgroup§								
Predicted amino acid substitution								
E6 ORF, E32Q	4 (9.5)	0	0	0	0	4 (1.8)		—
E6 ORF, D86E	4 (9.5)	2 (5.3)	2 (4.8)	0	0	8 (3.6)		—
E6 ORF, K93N	12 (28.6)	8 (21.1)	6 (14.3)	2 (3.8)	0	28 (12.7)		<.001
E7 ORF, R9K	0	0	0	2 (3.8)	0	2 (0.9)		—
E7 ORF, T20I	6 (14.3)	6 (15.8)	6 (14.3)	24 (46.2)	24 (52.2)	66 (30.0)		<.001
E7 ORF, G41R	20 (47.6)	8 (21.1)	18 (42.9)	16 (30.8)	16 (34.8)	78 (35.5)		.456
E7 ORF, G63D	20 (47.6)	8 (21.1)	18 (42.9)	16 (30.8)	14 (30.4)	76 (34.5)		.251
E7 ORF, G63H	0	0	0	0	2 (4.3)	2 (0.9)		—
E7 ORF, G63S	10 (23.8)	6 (15.8)	6 (14.3)	24 (46.2)	24 (52.2)	70 (31.8)		<.001
E7 ORF, T74A	4 (9.5)	2 (5.3)	2 (4.8)	0	0	8 (3.6)		—
E7 ORF, D76E	4 (9.5)	2 (5.3)	2 (4.8)	0	0	8 (3.6)		—
E7 ORF, V77A	2 (4.8)	6 (15.8)	6 (14.3)	6 (11.5)	4 (8.7)	24 (10.9)		.820
HPV58 variant								
E6/E7-prototype-like¶	4 (9.5)	8 (21.1)	8 (19.0)	4 (7.7)	2 (4.3)	26 (11.8)		.139
E6/E7-HK-1	18 (42.9)	6 (15.8)	18 (42.9)	14 (26.9)	12 (26.1)	68 (30.9)		.280
E6/E7-HK-2	6 (14.3)	6 (15.8)	6 (14.3)	22 (42.3)	18 (39.1)	58 (26.4)		<.001
E6/E7-HK-3	2 (4.8)	4 (10.5)	6 (14.3)	0	0	12 (5.5)		—
E6/E7-HK-4	0	2 (5.3)	0	6 (11.5)	4 (8.7)	12 (5.5)		—
E6/E7-HK-5	4 (9.5)	4 (10.5)	0	2 (3.8)	0	10 (4.5)		—
B) Whole study population#								
HPV58-negative	781 (94.9)	273 (87.8)	202 (82.8)	348 (87.0)	100 (68.5)	1704 (88.6)	Not done	
HPV58-positive, all variants	42 (5.1)	38 (12.2)	42 (17.2)	52 (13.0)	46 (31.5)	220 (11.4)	8.55 (5.23 to 14.02)	
HPV58 E7 T20I/G63S variants	6 (0.7)	6 (1.9)	6 (2.5)	24 (6.0)	24 (16.4)	66 (3.4)	26.79 (10.14 to 74.72)	
HPV58 non-E7 T20I/G63S variants	36 (4.4)	32 (10.3)	36 (14.8)	28 (7.0)	22 (15.1)	154 (8.0)	3.88 (2.13 to 7.05)	
HPV58 E6 K93N variants	12 (1.5)	8 (2.6)	6 (2.5)	2 (0.5)	0	28 (1.5)	0.00 (0.00 to 2.42)	

*CIN I, II, or III = cervical intraepithelial neoplasia grade I, II, or III, respectively; ICC = invasive cervical cancer; OR = odds ratio; CI = confidence interval; ORF = open reading frame.

†OR (95% CI) for ICC relative to normal group by the two-sided χ^2 test or the two-sided Fisher's exact test, as appropriate.

‡By the two-sided exact test (StatXact 4.0.1) for trend in proportion with increasing severity of cervical neoplasia. A dash indicates that numbers are too low for statistical analysis.

§Total number of women per group: Normal = 42; CIN I = 38; CIN II = 42; CIN III = 52; ICC = 46; Total = 220.

||Women with isolates having more than one missense nucleotide variation are counted more than once. Position of amino acid change is stated numerically. The one-letter amino acid code is used. The letter preceding this number refers to the reference amino acid derived from the prototype sequence (GenBank accession number NC_001443), and the letter following it refers to its substitution.

¶Isolates without nucleotide sequence deviation from prototype over the E6 and E7 ORFs.

#Total number of women per group: Normal = 823; CIN I = 311; CIN II = 244; CIN III = 400; ICC = 146; Total = 1924.

tions was found to have a statistically significant positive trend of association with the severity of neoplasia ($P_{\text{trend}} < .001$ by the exact test for trend; Table 1, A).

Isolates were further grouped according to the presence of E6 K93N and E7 T20I/G63S substitutions. E6 K93N-positive isolates included variants HPV58 E6/E7-HK-3, -5, -8, and -14. E7 T20I/G63S-positive isolates included variants HPV58 E6/E7-HK-2, -6, and -10. Among the study subjects (all had an abnormal Pap smear), the presence of HPV58 (prototype-like or any variant) carried an odds ratio for ICC of 8.55 (95% CI = 5.23 to 14.02). The corresponding odds ratio for E7 T20I/G63S-positive variants was 26.79 (95% CI = 10.14 to 74.72), whereas that for variants without these substitutions was 3.88 (95% CI = 2.13 to 7.05; Table 1, B). The association was 6.9-fold higher for E7 T20I/G63S-positive variants than for variants without these substitutions. The E6 K93N substitution appeared to be more prevalent among women with a normal cervix or low-grade lesions (Table 1, A). However, when the whole study population was considered, variants carrying this substitution did not show any statistically significant positive or negative association with ICC (Table 1, B).

Among patients with CIN III or ICC, infection with E7 T20I/G63S variants was associated statistically significantly with a younger age at diagnosis than was infection with other variants (median age [interquartile range]: 37 years [30.5–59.3 years] versus 48 years [36.8–63.0 years], $P = .038$ by two-sided Mann–Whitney U test; SPSS 10.1.0; SPSS Inc., Chicago, IL). This observation is in line with a higher oncogenicity of E7 T20I/G63S.

Worldwide, HPV58 has been found in only 2% of cervical cancers (9). In contrast, one third of the patients with cancer in the current study were positive for HPV58. This unusually high prevalence has also been reported in Chinese populations living in Shanghai (10), Jiangxi (11), and Taiwan (12). HPV58 is closely related to HPV33 and is grouped with HPV16, -31, -33, -35, -52, and -67 under the same branch of an HPV phylogenetic tree (13). We found that sequence variability of HPV58 variants was fourfold higher for E7 than for E6, in contrast with HPV16, where E6 is

reported to be more variable (2–5). The E7 protein has three domains: Conserved region (CR)-1 (amino-terminal 20 amino acids) and CR-2 (amino acids 21–40) are similar to adenovirus E1A (14,15), and CR-3 (carboxyl-terminal amino acids 41–98) contains two zinc-binding motifs (Cys-Xaa-Xaa-Cys) that are involved in dimerization and protein stability (16). We found two substitutions that were associated with a higher oncogenic risk. The first substitution (E7 T20I) is located at amino acid 20, close to the Leu-Xaa-Cys-Xaa-Glu domain that mediates association with the retinoblastoma protein and its related proteins, p107 and p130. The second substitution (E7 G63S) results in a change from glycine to serine. The E7 protein is phosphorylated on serine residues by casein kinase II (17), and a positive association between phosphorylation rate and oncogenic potential has been found (18). Thus, we suspected that the E7 G63S substitution might have created an additional phosphorylation site that confers increased transforming activity. However, our hypothesis was not supported by an artificial neural network-based computer prediction (NetPhos 2.0 [http://www.cbs.dtu.dk/services/NetPhos/]) (19). The NetPhos prediction gave a very low score for Ser-63 compared with Ser-31 and Ser-32, which are phosphorylated by casein kinase II (phosphorylation scores determined by NetPhos were 0.002 for Ser-63, 0.977 for Ser-31, and 0.992 for Ser-32). In the NetPhos prediction, a score closer to unity indicates a higher chance of phosphorylation.

In summary, our study provides epidemiologic evidence that HPV58 variants carrying E7 T20I/G63S substitutions are associated with an increased risk for cervical cancer. The exact mechanism for the increased oncogenicity needs further investigation. Cosegregating sequence variations in other parts of the virus genome may also play an important role. Finally, because this is a cross-sectional study, the magnitude of the risk should be interpreted cautiously.

REFERENCES

(1) Giannoudis A, Herrington CS. Human papillomavirus variants and squamous neoplasia of the cervix. *J Pathol* 2001;193:295–302.

(2) Zehbe I, Wilander E, Delius H, Tommasino M. Human papillomavirus 16 E6 variants are more prevalent in invasive cervical carcinoma than the prototype. *Cancer Res* 1998;58:829–33.

(3) Stephen AL, Thompson CH, Tattersall MH, Cossart YE, Rose BR. Analysis of mutations in the URR and E6/E7 oncogenes of HPV 16 cervical cancer isolates from central China. *Int J Cancer* 2000;86:695–701.

(4) Nindl I, Rindfleisch K, Lotz B, Schneider A, Durst M. Uniform distribution of HPV 16 E6 and E7 variants in patients with normal histology, cervical intra-epithelial neoplasia and cervical cancer. *Int J Cancer* 1999;82:203–7.

(5) Bontkes HJ, van Duin M, de Gruijl TD, Duggan-Keen MF, Walboomers JM, Stukart MJ, et al. HPV 16 infection and progression of cervical intra-epithelial neoplasia: analysis of HLA polymorphism and HPV 16 E6 sequence variants. *Int J Cancer* 1998;78:166–71.

(6) Chan PK, Li WH, Chan MY, Ma WL, Cheung JL, Cheng AF. High prevalence of human papillomavirus type 58 in Chinese women with cervical cancer and precancerous lesions. *J Med Virol* 1999;59:232–8.

(7) Chan PK, Chang AR, Cheung JL, Chan DP, Xu LY, Tang NL, et al. Determinants of cervical human papillomavirus infection: differences between high- and low-oncogenic risk types. *J Infect Dis* 2002;185:28–35.

(8) Kirii Y, Iwamoto S, Matsukura T. Human papillomavirus type 58 DNA sequence. *Virology* 1991;185:424–7.

(9) Bosch FX, Manos MM, Munoz N, Sherman M, Jansen AM, Peto J, et al. Prevalence of human papillomavirus in cervical cancer: a worldwide perspective. International biological study on cervical cancer (IBSCC) Study Group. *J Natl Cancer Inst* 1995;87:796–802.

(10) Huang S, Afonina I, Miller BA, Beckmann AM. Human papillomavirus types 52 and 58 are prevalent in cervical cancers from Chinese women. *Int J Cancer* 1997;70:408–11.

(11) Lin QQ, Yu SZ, Qu W, Cruz Y, Burk RD. Human papillomavirus types 52 and 58. *Int J Cancer* 1998;75:484–5.

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

(13) Bernard HU, Chan SY, Manos MM, Ong CK, Villa LL, Delius H, et al. Identification and assessment of known and novel human papillomaviruses by polymerase chain reaction amplification, restriction fragment length polymorphisms, nucleotide sequence, and phylogenetic algorithms. *J Infect Dis* 1994;170:1077–85.

(14) Phelps WC, Yee CL, Munger K, Howley PM. The human papillomavirus type 16 E7 gene encodes transactivation and transformation functions similar to those of adenovirus E1A. *Cell* 1988;53:539–47.

- (15) Moran E, Mathews MB. Multiple functional domains in the adenovirus E1A gene. *Cell* 1987;48:177-8.
- (16) McIntyre MC, Frattini MG, Grossman SR, Laimins LA. Human papillomavirus type 18 E7 protein requires intact Cys-X-X-Cys motifs for zinc binding, dimerization, and transformation but not for Rb binding. *J Virol* 1993;67:3142-50.
- (17) Firzlaff JM, Luscher B, Eisenman RN. Negative charge at the casein kinase II phosphorylation site is important for transformation but not for Rb protein binding by the E7 protein of human papillomavirus type 16. *Proc Natl Acad Sci U S A* 1991;88:5187-91.
- (18) Barbosa MS, Edmonds C, Fisher C, Schiller JT, Lowy DR, Vousden KH. The region of the HPV E7 oncoprotein homologous to adenovirus E1a and Sv40 large T antigen contains separate domains for Rb binding and casein kinase II phosphorylation. *EMBO J* 1990;9:153-60.
- (19) Blom N, Gammeltoft S, Brunak S. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J Mol Biol* 1999;294:1351-62.

NOTES

Supported by the Lee Hysan Foundation Research Grant and the Endowment Fund Research Grant (CA10830, CA10990) from United College, The Chinese University of Hong Kong. C. Lam is a Croucher Foundation Fellow.

Part of the results have been presented at the 18th International Papillomavirus Conference, Barcelona 2000 (abstract Nos. 095 and 096).

Manuscript received February 5, 2002; revised May 21, 2002; accepted June 18, 2002.