

## Short Sequence in L1 Region of Human Papillomaviruses Correlates with Clinical Pictures and Grouping by Cross-hybridization

Aikichi Iwamoto,<sup>1,4</sup> Hiroyuki Yoshikawa,<sup>2</sup> Kyoko Kitagawa,<sup>2</sup> Hiroko Igarashi,<sup>1</sup> Takashi Kawana<sup>3</sup> and Hiroshi Yoshikura<sup>1</sup>

Departments of <sup>1</sup>Bacteriology and <sup>2</sup>Obstetrics and Gynecology, Faculty of Medicine, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113 and <sup>3</sup>Department of Obstetrics and Gynecology, Tokyo University Branch Hospital, 3-28-6 Mejirodai, Bunkyo-ku, Tokyo 112

A consensus primer-mediated polymerase chain reaction devised to amplify a short sequence in L1 region (L1-PCR) efficiently detected genital human papillomaviruses (HPVs) in clinical materials. Nucleotide sequencing of the amplified fragment showed that L1-PCR is also applicable to hitherto unsequenced HPVs. By comparing the amplified 210 nucleotides HPVs were classified into six groups, which are consistent with clinical pictures and a grouping based on cross-hybridization under the stringent condition.

Key words: Human papillomavirus — Cervical carcinoma — Polymerase chain reaction

Human papillomaviruses (HPVs) have been classified according to nucleic acid homology. By definition, if there is less than 50% cross-hybridization when tested by reassociation in the liquid phase, the DNAs are assigned to different types.<sup>1)</sup> Sixty-six different types have been identified so far<sup>2)</sup> and about one third of them are so-called genital HPVs. For simplifying the diagnosis of genital HPV infection, we and others have developed assay systems employing consensus (general) primer-mediated polymerase chain reaction (PCR).<sup>3-7)</sup> The PCR for L1 region (L1-PCR) that we described previously could amplify and type at least nine genital HPVs (HPV types 6, 11, 16, 18, 31, 33, 42, 52 and 58).<sup>7)</sup> We could identify the types in 87% of HPV-positive clinical specimens by restriction mapping but the remaining 13% could not be assigned to any sequenced type and was tentatively grouped as type X. Table I shows the clinical status of the patients from whom type X DNA fragments were amplified. All these type X DNAs could be amplified by the combination of L1C1 (5'-CGTAAA-CGTTTTCCCTATTTTTT-3') and L1C2 (5'-TACCCTAAATACTCTGTATTG-3') primers.<sup>7)</sup> L1-PCR' using a combination of L1C1 and modified L1C2 (5'-TACCCTAAATACCCTATATTG-3')<sup>7)</sup> could also amplify all these, although the amplification of X04 was slightly weaker (data not shown). The amplified fragments were subcloned into Bluescript SK(+) (Stratagene) at the *Sma* I site. Double-stranded DNA sequencing was performed with sequenase (USB) on both strands using T7 or T3 primers (Stratagene). Fig. 1 shows the sequences sandwiched by the primers. The

pairs X01 and X05, X02 and X10, and X06 and X07 each shared identical nucleotide sequences; each pair of sequences will be referred to as X01, X02 or X06. X08 was different from X06 at only 4 out of 211 bases (98.1% identity), all of which were silent mutations. Since the frequency of mismatched base is much higher than the estimated error frequency of Taq polymerase,<sup>8)</sup> and since the highest nucleotide identity in this region among sequenced HPVs is 90.4% (see below), X08 is probably a variant of X06. X04 had one base difference from type 35.<sup>9)</sup> Identity of 207 out of 208 bases (99.5% identity) suggested that X04 is type 35, possibly a variant. Consequently, X01, X02, X03, X06, X09 and X11 were isolates which could not be assigned to any HPVs so far sequenced.

The frame starting at the second base in Fig. 1 is the coding frame of L1. About 210 bases (from the second to the last base of each shown in Fig. 1) of type X and the corresponding region of sequenced HPVs were compared using a computer program, GAP (Genetics Computer Group), which is based on an algorithm of Needleman and Wunsch.<sup>10)</sup> Table II shows the percent nucleotide identity. Among the sequenced HPVs the lowest nucleotide identity was 50.7% between types 1a and 35. The highest was 90.4% between types 6b and 11. By setting an arbitrary borderline at around 75% nucleotide identity, the sequenced HPVs could be classified into six groups as follows, type 1a; types 5, 8 and 47; types 2a and 57; types 6b and 11; types 16, 31, 33, 35 and 58; types 18 and 39. Type 1a is associated with cutaneous lesion (verrucae plantares).<sup>11)</sup> Types 5, 8 and 47 have been grouped into cutaneous type associated with a rare hereditary disease, epidermodysplasia verruciformis.<sup>12-15)</sup>

<sup>4</sup> To whom correspondence should be addressed.

Table I. Clinical Status of the Patients from whom Type X HPVs Were Amplified

HPV types	Clinical status	
	Site of origin	Diagnosis
X01	Cervix	Moderate dysplasia
X02	Cervix	Carcinoma <i>in situ</i>
X03	Cervix	Carcinoma <i>in situ</i>
X04	Cervix	Invasive carcinoma <sup>a)</sup>
X05	Cervix	Normal cervix <sup>b)</sup>
X06	Cervix	Normal cervix
X07	Cervix	Normal cervix
X08	Cervix	Normal cervix
X09	Vulva	Invasive carcinoma
X10	Cervix	Moderate dysplasia
X11	Cervix	Invasive carcinoma

a) Invasive squamous cell carcinoma.

b) Cytologically normal cervix.

Types 2a and 57, highly homologous to each other,<sup>2)</sup> have been grouped as ambivalent HPVs causing both cutaneous and mucosal lesions.<sup>16)</sup> Others are genital HPVs: Types 6b and 11<sup>17, 18)</sup> have been grouped into benign condyloma-associated type, while types 16, 31, 33, 35 and 58<sup>19-24)</sup> and types 18 and 39<sup>25-27)</sup> have been shown to be associated with invasive carcinomas. Thus, grouping on the basis of the nucleotide identity in such a short region in L1 was consistent with the grouping by cross-hybridization under stringent conditions using whole genomes.<sup>28)</sup> Although type 33 and type 39, respectively, have been classified as separate from others by cross-hybridization, nucleotide sequencing of the whole genome suggested similarity between types 33 and 16<sup>20)</sup> and between types 39 and 18.<sup>29)</sup> Therefore, grouping by the 210 nucleotide sequence in L1 region accords better with the results of nucleotide sequencing of the whole genome than the cross-hybridization.

Table II also shows a comparison of type X DNAs. X03, X02 and X11 had nucleotide identity higher than 75% with type 39, while the latter two were also highly homologous to type 18. X01, X06 and X09 had around 60-70% identity with each other and with sequenced HPVs listed in Table II. Type X DNAs probably represent new or cloned but yet unsequenced HPVs. Fig. 2 shows the tree representation of the nucleotide sequence relationship obtained by using the computer programs PILEUP (which is essentially GAP) and FIGURE (Genetics Computer Group) (similar analysis based on amino acid sequences gave essentially the same results). Twenty-one types of HPV were again classified into six groups. L1-PCR is applicable to four genital HPV groups. X01 which was detected in moderate dysplasia

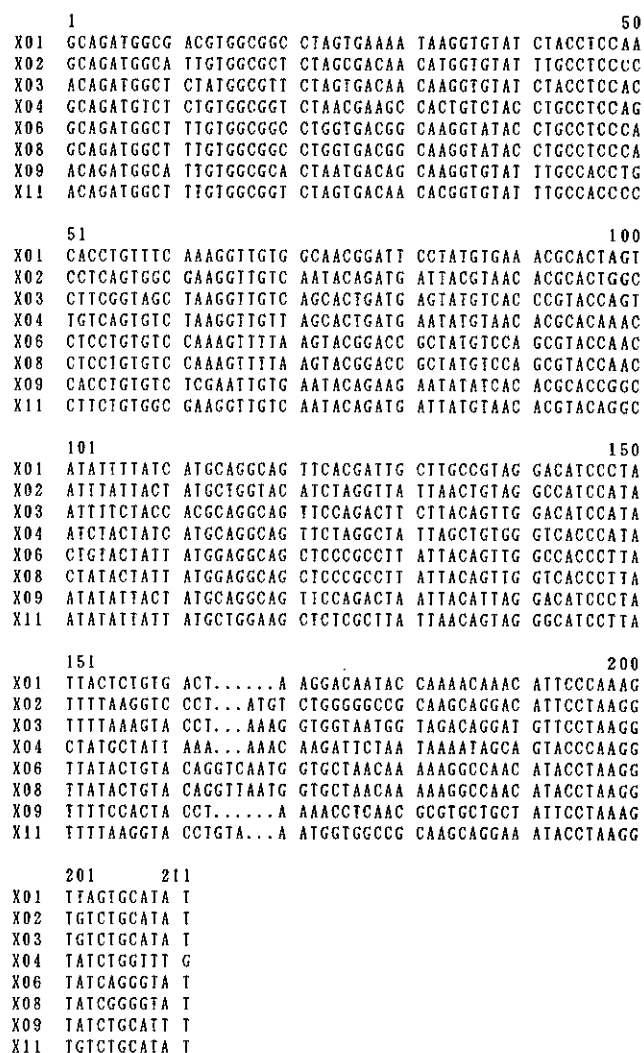


Fig. 1. Alignment of the nucleotide sequences of type X DNAs. Dots were inserted to maximize the homology.

and normal cervix was grouped with types 6b and 11. X06 which was amplified from normal cervix was grouped with types 2a and 57. X02, X03, X09 and X11 which were from carcinomas were grouped with types 18 and 39. Thus, L1-PCR is not only an efficient and simple method for detecting and typing HPVs but also its products can be used for grouping.

In order to examine whether this correlation is particular to this region or not, we used PILEUP and compared respective nucleotide sequences corresponding to the open reading frames (ORFs) E1, E6, E7, L1 or L2 of HPVs types 1a, 2a, 5, 6b, 8, 11, 16, 18, 31, 33, 39, 47 and 57. These HPV types were grouped into six groups roughly as above for all the ORFs examined (data not

Table II. % Nucleotide Identity

	1a	2a	5	6b	8	11	16	18	31	33	35	39	47	57	58	X01	X02	X03	X06	X09	X11
1a		53.2	58.7	57.7	59.7	56.6	55.2	51.7	57.2	53.5	50.7	51.2	60.7	53.7	53.2	54.7	51.7	59.2	52.2	54.2	53.2
2a	53.2		61.3	68.7	61.8	70.8	65.2	66.2	68.1	69.1	71.1	66.7	63.7	89.7	69.1	69.1	62.7	69.1	72.1	62.7	63.2
5	58.7	61.3		57.1	81.4	57.6	58	62.8	60.4	60.4	59.4	61.4	83.3	58.5	59.4	57.4	59.4	61.4	61.4	57.4	59.4
6b	57.7	68.7	57.1		58.6	90.4	72.3	67.7	73.8	72.3	66.7	70.8	58.1	68.2	70.3	68.7	63.6	67.7	68.7	65.1	66.7
8	59.7	61.8	81.4	58.6		62.1	63.3	55.6	61.4	63.3	59.4	60.9	85.7	60.4	61.4	60.3	58.9	61.4	60.9	61.4	55.6
11	56.6	70.8	57.6	90.4	62.1		72.8	65.6	73.3	71.8	71.3	69.7	59.6	69.2	70.3	73.3	64.1	67.7	66.7	65.6	64.6
16	55.2	65.2	58	72.3	63.3	72.8		61.8	80.7	82.1	76.8	63.7	61.8	68.1	82.6	66.2	60.9	65.7	64.3	66.2	61.8
18	51.7	66.2	62.8	67.7	55.6	65.6	61.8		65.7	64.7	63.3	78.7	56.5	64.7	67.2	66.2	75.8	72.9	65.7	66.2	81.2
31	57.2	68.1	60.4	73.8	61.4	73.3	80.7	65.7		74.8	77.8	64.7	62.3	69.6	75.2	67.2	64.7	69.6	63.8	67.2	65.7
33	53.5	69.1	60.4	72.3	63.3	71.8	82.1	64.7	74.8		75.8	64.3	62.8	71.4	89.5	62.7	61.8	67.2	63.8	66.2	62.3
35	50.7	71.1	59.4	66.7	59.4	71.3	76.8	63.3	77.8	75.8		65.2	60.9	69.1	77.8	62.3	61.8	64.7	64.7	61.3	63.3
39	51.2	66.7	61.4	70.8	60.9	69.7	63.7	78.7	64.7	64.3	65.2		60.4	67.2	67.2	68.6	86.0	76.3	68.1	71.1	88.9
47	60.7	63.7	83.3	58.1	85.7	59.6	61.8	56.5	62.3	62.8	60.9	60.4		61.8	61.4	60.8	59.4	60.9	63.8	58.8	58.5
57	53.7	89.7	58.5	68.2	60.4	69.2	68.1	64.7	69.6	71.4	69.1	67.2	61.8		70.5	68.6	61.8	67.6	72.5	62.7	65.2
58	53.2	69.1	59.4	70.3	61.4	70.3	82.6	67.2	75.2	89.5	77.8	67.2	61.4	70.5		62.7	63.8	62.7	64.3	65.7	63.8
X01	54.7	69.1	57.4	68.7	60.3	73.3	66.2	66.2	67.2	62.7	62.3	68.6	60.8	68.6	62.7		63.2	68.1	64.2	64.7	64.7
X02	51.7	62.7	59.4	63.6	58.9	64.1	60.9	75.8	64.7	61.8	61.8	86.0	59.4	61.8	63.8	63.2		72.5	62.8	69.6	86.5
X03	59.2	69.1	61.4	67.7	61.4	67.7	65.7	72.9	69.6	67.2	64.7	76.3	60.9	67.6	62.7	68.1	72.5		66.7	69.6	73.9
X06	52.2	72.1	61.4	68.7	60.9	66.7	64.3	65.7	63.8	63.8	64.7	68.1	63.8	72.5	64.3	64.2	62.8	66.7		63.7	71.0
X09	54.2	62.7	57.4	65.1	61.4	65.6	66.2	66.2	67.2	66.2	61.3	71.1	58.8	62.7	65.7	64.7	69.6	69.6	63.7		70.6
X11	53.2	63.2	59.4	66.7	55.6	64.6	61.8	81.2	65.7	62.3	63.3	88.9	58.5	65.2	63.8	64.7	86.5	73.9	71.0		70.6

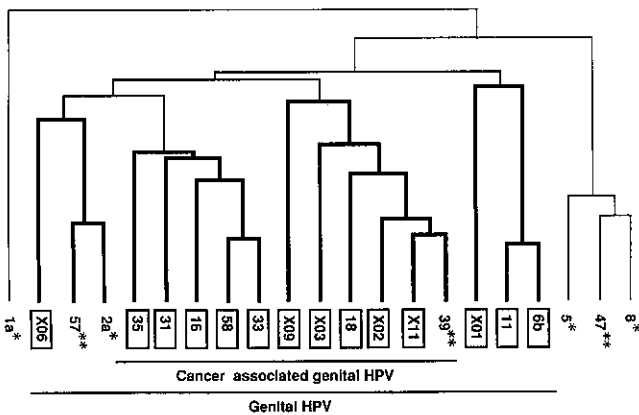


Fig. 2. A tree representation of clustering relationships of the nucleotide sequences. Numbers represent the corresponding HPV types. Boxed: HPV types which are amplified by L1-PCR. \*: HPV types which are not amplified by L1-PCR. \*\*: HPV types for which L1-PCR has not been tested. Although there may be an argument that type 2a should not be included among genital HPVs, it was included here to show the close relationship with type 57. The nucleotide numbers used for comparison are as follows, type 1a, 1887–2087; type 2a, 5787–5990; type 5, 5914–6123; type 6b, 5786–5983; type 8, 5848–6057; type 11, 5768–5965; type 16, 5634–5840; type 18, 5610–5816; type 31, 5549–5758; type 33, 5591–5800; type 35, 5571–5777; type 39, 5640–5846; type 47, 5900–6109; type 57, 5748–5954; and type 58.<sup>30)</sup> The nucleotide numbers are those used in the GENBANK and EMBL data bases.

shown). Therefore, it appears that the similarity of nucleotide composition among the members of the same group is widespread over the genome. These results are consistent with the hypothesis that a primordial HPV diverged first into several prototypes which diverged further thereafter.

This work was supported in part by a Grant-in-Aid for Cancer Research from the Ministry of Education, Science and Culture of Japan, and by a Grant-in-Aid from the Ministry of Health and Welfare for the Comprehensive 10-Year Strategy for Cancer Control, Japan.

(Received December 5, 1991/Accepted February 7, 1992)

REFERENCES

- 1) Coggins, J. R., Jr. and zur Hausen, H. Workshop on papillomaviruses and cancer. *Cancer Res.*, **39**, 545-546 (1979).
- 2) Hirsch-Behnam, A., Delius, H. and de Villiers, E.-M. A comparative sequence analysis of two human papillomavirus (HPV) types 2a and 57. *Virus Res.*, **18**, 81-98 (1990).
- 3) Gregoire, L., Arella, M., Campione-Piccardo, J. and Lancaster, W. D. Amplification of human papillomavirus DNA sequences by using conserved primers. *J. Clin. Microbiol.*, **27**, 2660-2665 (1989).
- 4) Manos, M. M., Ting, Y., Wright, D. K., Lewis, A. J., Broker, T. R. and Wolinsky, S. M. Use of polymerase chain reaction amplification for the detection of genital human papillomaviruses. *Cancer Cells*, **7**, 209-214 (1989).
- 5) van den Brule, J. C. A., Snijders, P. J. F., Gordijn, R. L. J., Bleker, O. P., Meijer, C. J. L. M. and Walboomers, J. M. M. General primer-mediated polymerase chain reaction permits the detection of sequenced and still unsequenced human papillomavirus genotypes in cervical scrapes and carcinomas. *Int. J. Cancer*, **45**, 644-649 (1990).
- 6) Yoshikawa, H., Kawana, T., Kitagawa, K., Mizuno, M., Yoshikura, H. and Iwamoto, A. Amplification and typing of multiple cervical cancer-associated human papillomavirus DNAs using a single pair of primers. *Int. J. Cancer*, **45**, 990-992 (1990).
- 7) Yoshikawa, H., Kawana, T., Kitagawa, K., Mizuno, M., Yoshikura, H. and Iwamoto, A. Detection and typing of multiple genital human papillomaviruses by DNA amplification with consensus primers. *Jpn. J. Cancer Res.*, **82**, 524-531 (1991).
- 8) Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S., Higuchi, R. H., Horn, G. T., Mullis, K. B. and Erlich, H. A. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, **239**, 487-491 (1988).
- 9) Marich, J. M., Pontsler, A. P., Rice, S. R., McGraw, K. A., Clark, P. R., Fleurbaaij, G. A. and Dubensky, T. W. Sequence analysis and propagation in cells of human papillomavirus type 35 DNA. In "Abstract of Papillomavirus Workshop," p. 338 (1991).
- 10) Needleman, S. B. and Wunsch, C. D. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.*, **48**, 443-453 (1970).
- 11) Favre, M., Orth, G., Croissant, O. and Yaniv, M. Human papillomavirus DNA: physical map. *Proc. Natl. Acad. Sci. USA*, **72**, 4810-4814 (1975).
- 12) Orth, G., Jablonska, S., Favre, M., Croissant, O., Jarzabek-Chorzelska, M. and Rzesa, G. Characterization of two types of human papillomaviruses in lesions of epidermodysplasia verruciformis. *Proc. Natl. Acad. Sci. USA*, **75**, 1537-1541 (1978).
- 13) Orth, G., Favre, M., Breitbart, F., Croissant, O., Jablonska, S., Obalek, S., Jarzabek-Chorzelska, M. and Rzesa, G. Epidermodysplasia verruciformis: a model for the role of papillomaviruses in human cancer. In "Viruses in Naturally Occurring Cancers," ed. M. Essex, pp. 259-281 (1980). Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- 14) Adachi, A., Yasue, H., Ohashi, M. and Ishibachi, M. A novel type of human papilloma virus DNA from the lesion of epidermodysplasia verruciformis. *Jpn. J. Cancer Res.*, **77**, 978-984 (1986).
- 15) Pfister, H., Neurnberger, F., Gissmann, L. and zur Hausen, H. Characterization of a human papillomavirus from epidermodysplasia verruciformis lesions of a patient from Upper Volta. *Int. J. Cancer*, **27**, 645-650 (1981).
- 16) de Villiers, E.-M. Heterogeneity of the human papillomavirus group. *J. Virol.*, **63**, 4898-4903 (1989).
- 17) de Villiers, E.-M., Gissmann, L. and zur Hausen, H. Molecular cloning of viral DNA from human genital warts. *J. Virol.*, **40**, 932-935 (1981).
- 18) Gissmann, L., Diehl, V., Schultz-Coulon, H.-J. and zur Hausen, H. Molecular cloning and characterization of human papillomavirus DNA derived from a laryngeal papilloma. *J. Virol.*, **44**, 393-400 (1982).
- 19) Beaudenon, S., Kremsdorf, D., Croissant, O., Jablonska, S., Wain-Hobson, S. and Orth, G. A novel type of human papillomavirus associated with genital neoplasias. *Nature*, **321**, 246-249 (1986).
- 20) Cole, S. T. and Streeck, R. E. Genome organization and nucleotide sequence of human papillomavirus type 33, which is associated with cervical cancer. *J. Virol.*, **58**, 991-995 (1986).
- 21) Durst, M., Gissmann, L., Ikenberg, H. and zur Hausen, H. A papillomavirus DNA from a cervical carcinoma and its prevalence in cancer biopsy samples from different geographic regions. *Proc. Natl. Acad. Sci. USA*, **80**, 3812-3815 (1983).
- 22) Lorincz, A. T., Lancaster, W. D. and Temple, G. F. Cloning and characterization of the DNA of a new human papillomavirus from a woman with dysplasia of the uterine cervix. *J. Virol.*, **58**, 225-229 (1986).
- 23) Lorincz, A. T., Quinn, A. P., Lancaster, W. D. and Temple, G. F. A new type of papillomavirus associated with cancer of the uterine cervix. *Virology*, **159**, 187-190 (1987).
- 24) Matsukura, T. and Sugase, M. Molecular cloning of a novel human papillomavirus (type 58) from an invasive cervical carcinoma. *Virology*, **177**, 833-836 (1990).
- 25) Beaudenon, S., Kremsdorf, D., Obalek, S., Jablonska, S., Pehau-Arnaudet, G., Croissant, O. and Orth, G. Plurality of genital human papillomaviruses: characterization of two new types with distinct biological properties. *Virology*, **161**, 374-384 (1987).
- 26) Boshart, M., Gissmann, L., Ikenberg, H., Kleinheinz, A.,

- Scheurlen, W. and zur Hausen, H. A new type of papillomavirus DNA, its presence in genital cancer biopsies and in cell lines derived from cervical cancer. *EMBO J.*, **3**, 1151–1157 (1984).
- 27) Cole, S. T. and Danos, O. Nucleotide sequence and comparative analysis of the human papillomavirus type 18 genome. Phylogeny of papillomaviruses and repeated structure of the E<sub>6</sub> and E<sub>7</sub> gene products. *J. Mol. Biol.*, **193**, 599–608 (1987).
- 28) Pfister, H. and Fuchs, P. G. Papillomaviruses: particles, genome organization and proteins. In "Papillomaviruses and Human Disease," ed. K. Syrjaenen, L. Gissmann and L. G. Koss, pp. 1–18 (1988). Springer-Verlag, Berlin.
- 29) Volpers, C. and Streeck, R. E. Genome organization and nucleotide sequence of human papillomavirus type 39. *Virology*, **181**, 419–423 (1991).
- 30) Kirii, Y., Iwamoto, S. and Matsukura, T. Human papillomavirus type 58 DNA sequence. *Virology*, **185**, 424–427 (1991).