JOURNAL OF VIROLOGY, Mar. 1997, p. 2463–2472 Vol. 71, No. 3
0022-538X/97/$04.00+0

Human Papillomavirus Type 16 Sequence Variation in Cervical Cancers: a Worldwide Perspective

TAKASHI YAMADA,[1,2] M. MICHELE MANOS,[3] JULIAN PETO,[4] CATHERINE E. GREER,[5] NUBIA MUÑOZ,[6] F. XAVIER BOSCH,[7] AND COSETTE M. WHEELER[2,8,9]*

Departments of Medicine,[1] Cell Biology,[8] and Epidemiology and Cancer Control Program,[9] and the New Mexico Cancer Research and Treatment Center,[2] University of New Mexico School of Medicine, Albuquerque, New Mexico 87131; Department of Molecular Microbiology and Immunology, School of Public Health, Johns Hopkins University, Baltimore, Maryland 21205[3]; Institute of Cancer Research, Belmont, England[4]; Chiron Corporation, Emeryville, California 94608[5]; Unit of Field and Intervention Studies, International Agency for Research on Cancer, Lyon, France[6]; and Servei d'Epidemiologica i Resistre del Cancer, Institute Catala d' Oncologia, Hospital Duran i Reynals, CSUB, E-08907 L'Hospitalet del Llobregat, Barcelona, Spain[7]

Received 25 September 1996/Accepted 26 November 1996

We examined intratype human papillomavirus type 16 (HPV-16) sequence variation in tumor samples that were collected and analyzed in an international study of invasive cervical cancer. The collection included tumors from 22 countries in five continents. Using our recently developed E6 and L1 PCR-based hybridization systems to distinguish HPV-16 variant lineages, we analyzed material from tumors previously found to contain HPV-16 DNA. Of 408 specimens analyzed in the E6 hybridization assay, 376 (92.2%) belonged to previously reported HPV-16 variant lineages. The remaining 32 specimens (7.8%) harbored HPV-16 variants with novel hybridization patterns, novel nucleotide changes, or both. Nucleotide sequences (1,203 bp) were determined for the E6, the MY09/11 region of L1, and the long control region of each novel variant and representative specimens from each hybridization pattern observed. Based on E6 hybridization patterns, most of the variants from European and North American samples were phylogenetically classified as European prototype (E) while samples from Africa contained primarily African 1 (Af1) or African 2 (Af2) variants. The majority of Asian (As) variants were observed in Southeast Asia, and almost all Asian American (AA) variants were from Central and South America or Spain. A single North American 1 (NA1) variant was detected in a tumor from Argentina. Nucleotide changes previously shown to covary between the MY09/11 region of L1 and the E6 coding region were examined in a subset of 249 specimens. We observed 22 combined E6-L1 hybridization patterns, of which 11 (in 21 samples) were novel. No unanticipated nucleotide covariation was observed between the E class and the AA-Af1-Af2-NA1 classes, suggesting the absence or rarity of genomic recombination between HPV-16 lineages. This extensive description of HPV-16 variants forms a basis for further examining the relationship between intratype variation and basic functional differences in biological activities. HPV-16 variants may prove important for the determination of the risk of cervical neoplasia and for the design of HPV-16 vaccine strategies.

Infection with specific types of genital human papillomavirus (HPV) is the main risk factor for both cervical cancer (8, 22) and its precursor lesion, cervical intraepithelial neoplasia (19, 27). The International Biological Study of Cervical Cancer (IBSCC) reported that HPV DNA is detectable in more than 93% of invasive cervical cancers worldwide (4). Of the over 20 different HPV types found associated with tumors in the IBSCC, HPV-16 was the most prevalent and represented about 50% of infections.

HPVs vary genetically not only between but also within types. Intratype variants are defined as HPVs that vary by 2% or less in specified regions of the genome (3). Our recent report (30) included the IBSCC tumor specimens in an evaluation of global intratype variation in the less prevalent HPV types. Several studies have reported intratype sequence variations of the most prevalent type, HPV-16, and have included specimens from diverse geographic locations (2, 7, 9–11, 14, 16, 17, 25, 29, 33, 36).

Recent investigations have evaluated associations of specific HPV-16 variants with viral persistence and with the development of high-grade cervical lesions (20, 32, 34, 35). Given its prevalence in cervical carcinoma, HPV-16 is the focus of most prophylactic and therapeutic vaccine development efforts that are under way. Intratype HPV-16 amino acid variations may be relevant to the generation of specific immune responses, especially in the context of rational vaccine strategies.

Ho et al. analyzed the long control region (LCR) from a worldwide collection of cervical samples and demonstrated that HPV-16 sequence variants form five distinct phylogenetic clusters (13). A subsequent study conducted by our group expanded and complemented these classifications of HPV-16 variants through sequence analysis of HPV-16 E6, L1, L2, and LCR variation (36). These data demonstrated a strong intergene sequence covariation within single HPV-16 isolates and suggested that nucleotide variations in one genomic region can be used to distinguish distinct variant virus lineages. This observation was further examined through the development and application of simplified PCR dot blot hybridization methods that are capable of distinguishing major HPV-16 class and

* Corresponding author. Mailing address: Department of Cell Biology, University of New Mexico Cancer Research and Treatment Center, 900 Camino de Salud N.E., Albuquerque, NM 87131-5226. Phone: (505) 277-9151. Fax: (505) 277-9494. E-mail: cwheeler@medusa.unm.edu.

subclass variants (33). These accurate HPV-16 variant assays target the HPV-16 E6 coding region and the MY09/11 region of L1 and are applicable to large-scale epidemiologic investigations. Here we describe the characterization of HPV-16 variants by these newly developed methods in a worldwide sample of invasive cervical cancers. In this report, the major groupings of HPV-16 variants are designated European (E), Asian (As), Asian American (AA), African 1 (Af1), African 2 (Af2), and North American 1 (NA1). Variants of these groups are denoted by HPV-16 E6 nucleotide position and substituted nucleotide. Thus, AA-G183 denotes an Asian American variant with a G at E6 nucleotide (nt) 183. An additional final letter denotes the nucleotide at E6 nt 350, a frequently varied position among European variants. This abbreviation is used to specify alterations at two nt positions one of which is E6 nt 350 (e.g. E-A176G).

MATERIALS AND METHODS

Clinical specimens. Crude DNA preparations from HPV-16 DNA-containing cervical cancer specimens were obtained from the IBSCC (4). DNA was available from 432 of the original 465 IBSCC tumors that were positive for HPV-16.

PCR. *Thermus aquaticus* (*Taq*)-based PCR amplifications were performed essentially as described previously (33, 36). Reaction mixtures contained 10 mM Tris (pH 8.5), 50 mM KCl, 200 μM each deoxynucleoside triphosphate, 2.5 or 4.0 mM $MgCl_2$, 1 nM to 0.2 μM each sense- and antisense-strand oligonucleotide primers, and 2.5 U of *Taq* DNA polymerase (Perkin-Elmer, Foster City, Calif.). A one-tube nested PCR method (33) was used to amplify the E6 coding region. Additionally, sequential two-tube nested PCR systems were applied to amplify the MY09/11 L1 region for subsequent hybridization assays and the E6, MY09/11 L1, and LCR segments for DNA sequence analysis (33). The crude DNA preparation was used at 2 μl in both one- and two-tube nested amplifications. We used 5′-biotinylated primers in the inner nested amplification reactions that were intended for direct DNA sequence analysis. Multiple independent PCR amplifications were conducted for hybridization and supplementary DNA sequence analysis to avoid the detection of *Taq*-derived nucleotide changes. PCR conditions and oligonucleotide primer sequences for all PCR amplification systems have been reported elsewhere (33).

Dot-blot hybridization. Dot-blot hybridization was performed as described previously (33). Briefly, PCR amplicons (80 μl) were denatured with 1.0 ml of 0.4 N NaOH–25 mM EDTA for 10 min at room temperature. Aliquots (50 μl) of the denatured amplicons were applied to a Biodyne-B nylon membrane (ICN, Irvine, Calif.) with a 96-well dot blot manifold. Following the application of the denatured samples, each well was rinsed with 20× SSPE (3.6 M NaCl, 0.2 M $NaH_2PO_4 \cdot H_2O$, 0.11 N NaCl, 0.02 M disodium $EDTA_2 \cdot H_2O$ [pH 7.4]). Bound DNA was cross-linked to the membranes with a Stratlinker UV cross-linker on the autolink setting (Stratagene, San Diego, Calif.).

Membranes were pretreated for 1 h at 65°C in 0.1× SSPE–0.5% sodium dodecyl sulfate (SDS) with vigorous shaking. Hybridization with biotinylated probes (1 pmol/ml) was conducted overnight at 42.0°C in 5× SSPE–0.1% SDS. The membranes were rinsed with Blot Wash solution (BW; 2× SSPE–0.1% SDS), washed twice for 10 min with BW at the desired temperatures (33) and incubated in BW containing 30 ng of streptavidin-horseradish peroxidase conjugate (Vector Laboratories, Burlingame, Calif.) per ml under gentle agitation for 20 min at room temperature. After two room temperature washes with BW for 10 min each, the membranes were incubated individually in the enhanced chemiluminescence substrate reagents (Amersham, Arlington Heights, Ill.) for 1 min. Exposures were conducted for 10 min followed by 2 h. Control PCR amplicons used in our previous study (33) were included on each membrane for both E6 and L1 hybridizations and included HPV-16 E-P, E-G131, E-C109, As, AA, Af1, and Af2 variants.

E6 and L1 hybridization assays were interpreted from chemiluminescent signals recorded on X-ray films. Scoring, data entry, and analyses were performed as described previously (33).

After these studies, we developed additional oligonucleotide probes to distinguish the most common novel nucleotide variations at nt 176, 178, and 188 (data not shown). Sequences for these probes are as follows: A176, 5′CAACTATAC AT(A)ATATAATAT3′; A178, 5′TATACATGA(A)ATAATATTAGA3′; and C188, 5′TATAATATTA(C)AATGTGTGT3′. Parentheses mark positions that vary from the reference HPV-16 sequence. For these probes, hybridization and washes were conducted at 42.0 and 45°C, respectively, as described above for the probes used in this study.

Nucleotide sequence analysis. Single-stranded DNAs were prepared from PCR amplicons by using streptavidin-coated magnetic beads (Dynabeads; Dynal, Lake Success, N.Y.). The nucleotide sequence was determined directly from the single-stranded DNAs by the dideoxy termination method (26) with Sequenase version 2.0 kits (United States Biochemical, Cleveland, Ohio) and ^{35}S.

HPV-16 DNA nucleotide positions are numbered according to the HPV-16R sequence published in a recent HPV DNA sequence compendium (24). This numbering differs from that in our previous report (36) in that we have accounted for an additional nucleotide insertion at nt 1139 (21). This results in an increase of one nucleotide in the L1 and LCR numbering presented.

Phylogenetic analysis. Phylogenetic analyses were performed over the E6, MY09/11, and LCR segments of 49 HPV-16 samples. All sequences from a given sample were combined, and the combined sequences were used to construct an alignment. Parsimony analyses were performed by the Phylogenetic Analysis Using Parsimony (PAUP) package version 3.1.1 (31) on a Macintosh Quadra 800, running system 7.1.

Statistical analyses. Fisher's exact test (see Table 1) for row and column independence was used to test whether the frequencies of HPV-16 variant classes differ only randomly across the five continents. Since the frequencies did show some significant differences ($P < 0.0001$), Fisher's exact test was also used on 2 × 2 tables to contrast the frequencies of the variants of interest (e.g., Af1, Af2, and E-350T in Africa with respect to other continents, E-350G and AA in Central and South America with respect to other continents, and As in Southeast Asia with respect to other continents). By using a Bonferroni procedure to correct for multiple testing, the results were considered significant in cases where P values were less than 0.0017 = 0.005/30. All tests were carried out in StatXact Turbo (Cytel Software Corp., Cambridge, Mass.). The test on the full 5 × 6 table was based on an asymptotic approximation; tests on the 2 × 2 tables used exact calculations.

RESULTS

HPV-16 E6 hybridizations. We evaluated 432 cervical cancer specimens previously characterized in the IBSCC (4) and found to contain HPV-16 DNA by the MY09/11 L1 consensus PCR system (1). Crude DNA preparations were amplified in a one-tube nested E6 PCR method described in our previous report (33). Both one- and two-tube nested E6 PCR amplification systems have demonstrated single-copy sensitivity at the level of ethidium bromide staining following agarose gel electrophoresis. Similarly, the evaluation of 24 randomly selected tumor samples in this set also demonstrated equivalent sensitivity for both one- and two-tube E6 nested PCR systems (data not shown). Immobilized HPV-16 E6 PCR amplicons were hybridized with 23 E6-specific oligonucleotide probes capable of detecting all class- and subclass-specific HPV-16 signature nucleotide changes identified in our previous investigations (33, 36).

Of the 432 specimens evaluated, 21 failed to hybridize with any of the E6 probes applied. Agarose gel electrophoresis of corresponding PCR amplicons revealed no appropriately sized bands, suggesting that these specimens were insufficient for amplification. Of the 432 specimens, 2 hybridized strongly with both reference and variant probes at a single position and 1 hybridized with both reference and variant probes at several positions, suggesting the possible presence of two HPV-16 variants in these samples. The inference that these three specimens contained two HPV-16 variants was confirmed in L1 hybridization experiments. Due solely to time constraints, these samples were excluded from subsequent consideration by DNA sequence analysis and variant designation. The remaining 408 cervical cancer samples were analyzed further by various methods.

All ambiguous E6 hybridization patterns observed in this study are shown in Fig. 1. Of the 408 specimens analyzed, ambiguous E6 hybridizations were observed in 45 specimens at 52 total probe positions. Nucleotide sequences were determined for all 52 E6 probe positions in the 45 samples demonstrating ambiguous E6 hybridization patterns. We found that 19 ambiguities were due to technical hybridization problems and 33 ambiguities were due to novel nucleotide changes within the probe binding sites subsequently detected by DNA sequence analyses (Fig. 1). These included a G-to-A change at nt 176 in 11 samples (A176), a G-to-C change at nt 188 in 6 samples (C188), a T-to-A change at nt 178 in 3 samples (A178), T-to-A and C-to-G changes at nt 137 and nt 285 in 2

A. E6 REGION

AMBIGUOUS HYBRIDIZATION PATTERNS	DNA SEQUENCE			
1111111223345 0334478883503 9123583695032 TAGCGTTTACTAA	1111111223345 0334478883503 9123583695032 TAGCGTTTACTAA	EXPLANATION OF FAILURE	AMINO ACID CHANGES	NO.OF ISOLATES
*----------	----------	PROBE		2
-----*-------	----------	PROBE		1
-**---*---G--	----------G--	PROBE		1
-**-------G--	----------G--	PROBE		1
-**-T--AGTG-G	----T--AGTG-G	PROBE		1
---??--AGTG-G	----T--AGTG-G	PROBE		1
-??GT--AGT---	--CGT--AGT---	PROBE		1
-**GT--AGT---	--CGT--AGT---	PROBE		1
C??GT--AGT-G-	C-TGT--AGT-G-	PROBE		5
C**GT--AG*-*-	C-TGT--AGT-G-	PROBE		1
*---------G--	----------G--	C to T at nt 115,	SILENT	1
-**----------	----------	A to C at nt 134,	Lys to Gln at aa 11	1
C**GT--AGT-G-	C-TGT--AGT-G-	T to G at nt 137,	Leu to Val at aa 12	1
-**GT--**T---	---GT--AGT---	T to A at nt 137,	Leu to Ile at aa 12	
		C to G at nt 285,	Ala to Gly at aa 61	2
-****--------	----------	A to G at nt 139,	SILENT	1
-----*----G--	----------G--	C to T at nt 173,	His to Tyr at aa 24	1
-----**------	----------	G to A at nt 176,	Asp to Asn at aa 25	6
-----**---G--	----------G--	G to A at nt 176,	Asp to Asn at aa 25	4
----T**AGTG-G	----T--AGTG-G	G to A at nt 176,	Asp to Asn at aa 25	1
-----**------	----------	T to A at nt 178,	Asp to Glu at aa 25	3
-----**---G--	----------G--	A to G at nt 184,	SILENT	1
------*--?---	----------	A to G at nt 187,	SILENT	
		PROBE		1
------*------	----------	G to C at nt 188,	Glu to Gln at aa 30	2
------*---G--	----------G--	G to C at nt 188,	Glu to Gln at aa 30	3
-G-GT-*AGT---	-G-GT--AGT---	G to C at nt 188,	Glu to Gln at aa 30	1
----T--AG*G-G	----T--AGTG-G	G to A at nt 333,	Arg to Lys at aa 77	1

B. L1 MY09/11 REGION

666666 678889 920569 513424 AGACTG	666666 678889 920569 513424 AGACTG			
C*-TCA	CA-TCA	PROBE		1
*-----	------	T to G at nt 6689,	Ser to Ala at aa 351	1
C*-T-A	CC-T-A	G to C at nt 6721,	Lys to Asn at aa 361	1

FIG. 1. Ambiguous hybridization patterns, explanation of hybridization failure, and novel amino acid changes. All ambiguous HPV-16 hybridization patterns observed are presented. The DNA sequence was obtained for HPV-16 isolates demonstrating ambiguous patterns. The HPV-16 DNA reference sequence (28), revised to include multiple corrections (24), is indicated below the nucleotide position number. For each variant sequence, positions that do not vary relative to the HPV-16 reference sequence are marked with dashes. Explanations for ambiguous results are shown to the right of the sequences. PROBE indicates a technical hybridization failure. When specific nucleotide changes were identified within the probe binding site, the nucleotide change and positions are designated. Hybridization ambiguities were defined with the following symbols as previously used (33): i) * and ** denote negative for all paired or grouped probes, respectively, targeting a specified nucleotide position(s); ii) ? and ?? denote a concurrent strong and weak hybridization signal, respectively, at more than one of a pair or group of probes targeting a specified nucleotide position(s).

samples (A137/G285), and sporadic changes that were observed only once in a single sample.

An overall frequency of HPV-16 E6 probe patterns following selected supplementary nucleotide sequence analysis is shown in Fig. 2. After excluding nucleotide changes that were detected only once in the 408 samples, 25 distinct HPV-16 E6 probe patterns were obtained. The large majority of specimens (376 of 408, 92.2%) fell into the typical (major) HPV-16 variant patterns (E-P-350T, E-P-350G, E-G131T, E-G131G, E-C109G, As, AA, AA-G183, NA1, Af1, and Af2) which were observed in our previous studies conducted in Portland, Oreg. (33, 36). The remaining 32 samples (7.8%) included HPV-16 variants with novel E6 hybridization patterns and/or novel nucleotide changes (e.g., A176 and C188, described above). Of these, 18 novel variants were members of the E class and 14 were members of the AA-Af1-Af2-NA1 classes. These variants could be assigned to major HPV-16 variant classes (E class or AA-Af1-Af2-NA1 classes) without DNA sequence clarification (Fig. 2). As a result, 283 of the 408 samples were classified as members of the HPV-16 E class and the remaining 125 sam-

1111111223345 0334478883503 9123583695032 TAGCGTTTACTAA	ADDITIONAL NUCLEOTIDE CHANGES(#)	CLASS-SUBCLASS($)	n (%)
----------		E-P-350T	101 (24.8)
----------	A176	E-A176T	6 (1.5)
----------	A178	E-A178T	3 (0.7)
----------	C188	E-C188T	2 (0.5)
----------G--		E-P-350G	138 (33.8)
----------G--	A176	E-A176G	3 (0.7)
----------A--	A176	E-A176A	1 (0.2)
----------G--	C188	E-C188G	3 (0.7)
C---------G--		E-C109G	3 (0.7)
-G----------		E-G131T	4 (1.0)
-G--------G--		E-G131G	8 (2.0)
-----G-------		As	11 (2.7)
--CGT--AGT---		Af1	40 (9.8)
--CGT--AGT--G		Af1	1 (0.2)
-G-GT--AGT---	C188	Af1	1 (0.2)
-G-GT--AGTG--		Af1	2 (0.5)
C-TGT--AGT-G-		Af2	18 (4.4)
---GT--AGTG--		Af2	1 (0.2)
---GT--AGT---		Af2	5 (1.2)
---GT--AGT---	A137/G285	Af2	2 (0.5)
----T--AGTG--		NA1	1 (0.2)
----T--AGTG-G		AA	42 (10.3)
----T--AGTG-G	A176	AA	1 (0.2)
----T--AG-G-G		AA	1 (0.2)
----T-GAGTG-G		AA	10 (2.5)
		TOTAL	408

FIG. 2. Frequencies of HPV-16 variants defined by E6 hybridization and supplementary sequence analysis. # denotes DNA sequence determinations of specific additional nucleotide variations with respect to the reference HPV-16 sequence. $ denotes the class and subclass. The major HPV-16 variant designation (i.e., European prototype [E-P], other European [E] variants, Asian [As], African [Af1 and Af2], North American [NA1], and Asian American [AA]) is the first letter assigned. Minor sequence variants observed in single specimens that are detailed in Fig. 1 are assigned categorically to the most relevant phylogenetic group. A letter immediately following a dash and preceding a nucleotide position number represents the variation at the specific nucleotide position designated compared to the reference sequence. Final letters represent E6 nt 350 variations in addition to the specified variation. This distinction is provided because this is a frequently varied nucleotide position.

E6 1111111223345 0334478883503 9123583695032 TAGCGTTTACTAA	L1:MY09/11 666666 678889 920569 513424 AGACTG	CLASS-SUBCLASS	n
----------	------	E-P-350T	44
-----A-------	------	E-A178T	3
----------G--	------	E-P-350G	48
----------G--	-----A	E-m	2
C---------G--	------	E-C109G	3
-G----------	------	E-G131T	4
-G--------G--	------	E-G131G	8
-----G-------	------	As	11
--CGT--AGT---	-A-T-A	Af1-a	39
--CGT--AGT---	CA-T-A	Af1-a	1
--CGT--AGT--G	-A-T-A	Af1-b	1
-G-GT--AGT---	-A-T-A	Af1-c	1
-G-GT--AGTG--	-A-T-A	Af1-d	2
C-TGT--AGT-G-	CA-T-A	Af2-a	17
C-TGT--AGT-G-	CC-T-A	Af2-a	1
---GT--AGT---	CA-T-A	Af2-b	7
---GT--AGTG--	-A-T-A	Af2-b	1
----T--AGTG--	CA-T-A	NA1	1
----T--AGTG-G	CATT-A	AA-a	41
----T--AG-G-G	C--T-A	AA-b	1
----T--AGTG-G	CA-TCA	AA-b	1
----T-GAGTG-G	CA-TCA	AA-c	10
		TOTAL	247

FIG. 3. Distribution of combined E6 and L1 hybridization patterns. Combined E6 and L1 hybridization patterns are summarized. Patterns in this figure are determined only based on 13 E6 and 6 L1 MY09/11 nucleotide positions; additional nucleotide changes found in other nucleotide positions (e.g., A176 and C188) are not considered. The class and subclass designations are similar to those presented in Fig. 2, but minor phylogenetic designations have been added.

ples were classified as members of the HPV-16 AA-Af1-Af2-NA1 classes.

Technical hybridization failures ($n = 19$) were observed with probes for the following positions: 2 at nt 109, 11 at nt 131/2, 1 at nt 143/5, 2 at nt 178/183, 2 at nt 335, and 1 at nt 403. Of 9,384 hybridizations to individual HPV-16-containing specimens (408 specimens with 23 probes), approximately 0.2% hybridization or probe failures were observed.

HPV-16 L1 hybridization. Nucleotide changes in the MY09/11 L1 region (nt 6617 to 6999), previously shown to covary with specific nucleotide changes in the E6 region (33, 36), were examined in 249 of the 408 HPV-16-positive specimens. These represent 61% of the specimens evaluated in the E6 hybridization assays above. Based on E6 hybridization results, the selected 249 specimens comprised all specimens within the novel E-class variants ($n = 18$), all specimens classified as AA-Af1-Af2-NA1-class variants ($n = 125$), and all As, E-G131, and E-C109 specimens ($n = 26$). Samples that had demonstrated typical E-P-350T and E-P-350G E6 hybridization patterns were randomly selected, and only 36 and 44 specimens were analyzed, respectively. A modified two-tube nested L1 PCR amplification described in our previous study was applied (33). This modification shifts the original primer sequence 1 nt upstream and employs a 5′-sense primer beginning at nt 6598 and ending at nt 6616 for the inner nested-amplification reaction. As detailed previously (33), this modification was intended to eliminate the amplification of several other genital HPV genotypes. A battery of 12 probes capable of distinguishing single nucleotide changes within the MY09/11 region was utilized in these L1 analyses.

Of the 249 specimens evaluated, two failed to hybridize with any of the L1 probes applied. PCR amplicons from these two specimens were subjected to ethidium bromide agarose gel analysis. Hybridization failure was attributed to inadequate amplification, since no amplicons were observed. These two samples were excluded from subsequent variant classification and DNA sequence analyses. The remaining 247 samples were examined further. These samples consisted of 123 E-class variants and 124 AA-Af1-Af2-NA1 variants based on E6 hybridization patterns determined.

Ambiguous patterns obtained with the battery of 12 L1-specific oligonucleotide probes are shown in Fig. 1. Ambiguous L1 hybridizations were observed at three probe positions. Sub-

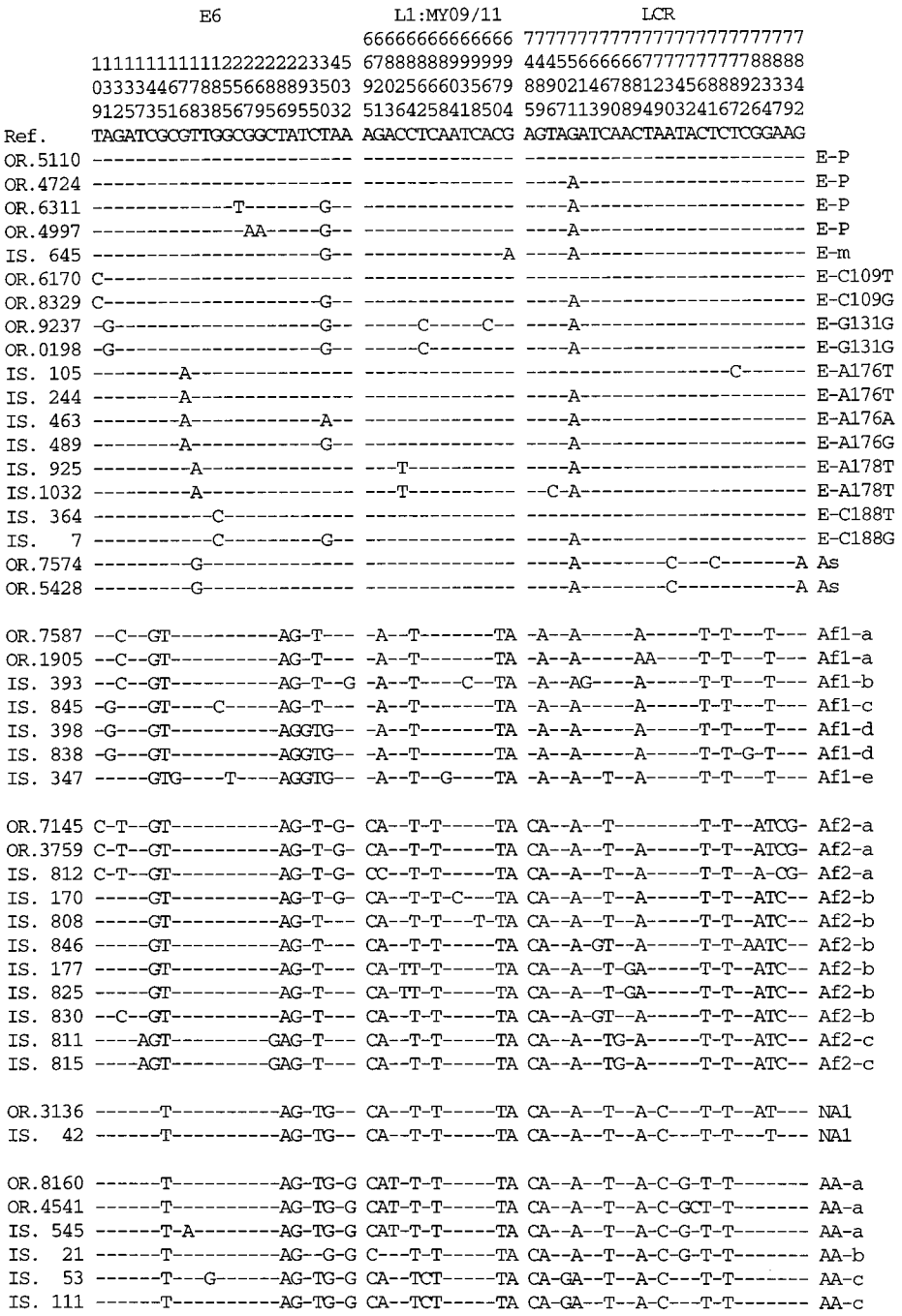

FIG. 4. E6, partial L1, and LCR sequences of representative HPV-16 variants. A total of 28 clinical samples representative of the novel HPV-16 E6-L1 variation patterns (sample designations begin with IS) were sequenced across the E6 region (nt 104 to 559), the MY09/11 L1 region (nt 6618 to 6999), and the LCR (nt 7480 to 7843). For reference comparison, 17 samples that represent the major HPV-16 variants described in our previous studies (33, 36) are included (designations begin with OR.). Novel minor variants were classified as members of the major variant classes according to phylogenetic analyses (right-hand column, Af1-b, Af1-c, etc.).

sequent sequence analyses showed that one ambiguity was due to a hybridization failure with one probe (the 6721-A probe) and two ambiguities resulted from a novel nucleotide change identified within the probe binding region (T to G at nt 6689 and G to C at nt 6721 occurring within either the 6695 or 6721 probe binding sites). In the 2,952 individual hybridizations to HPV-16-containing specimens (246 specimens with 12 probes), probe and hybridization failure in the L1 hybridization system was approximately 0.03%. For comparison, conventional (not nested) MY09/11 PCR amplicons from 37 randomly selected samples were hybridized with five probes (6654/62CC, 6654/62TT, 6654/62CT, 6994A, and 6994G). Of the 185 individual hybridizations (37 specimens with 5 probes), the probe and hybridization failure rate was approximately 16.2% (30 of 185), demonstrating that the nested PCR strategy had greatly improved the sensitivity of the system (data not shown).

The frequencies of the combined E6 and L1 hybridization patterns are shown in Fig. 3. The nucleotide changes that

TABLE 1. Distribution of HPV-16 variant classes by continent

Variant class[a]	No. (%) of specimens from:					Total no.
	Europe	North America	Central and South America	Southeast Asia	Africa	
E-350T	20 (40.0)	16 (53.3)	56 (24.6)	19 (54.3)	5 (7.8)[c]	116
E-350G	22 (44.0)	12 (40.0)	119 (52.2)[b]	2 (5.7)	1 (1.6)	156
As	1 (2.0)	1 (3.3)	0 (0.0)	9 (25.7)[b]	0 (0.0)	11
AA	7 (14.0)	0 (0.0)	45 (19.7)[b]	2 (5.7)	0 (0.0)	54
Af1	0 (0.0)	0 (0.0)	5 (2.2)	0 (0.0)	39 (60.9)[b]	44
Af2	0 (0.0)	1 (3.3)	3 (1.3)	3 (8.6)	19 (29.7)[b]	26
Total	50	30	228	35	64	407

[a] The single NA 1 specimen, from Central and South America, is not included in this table.
[b] Significant increase, $P < 0.0001$.
[c] Significant decrease, $P < 0.0001$.

required DNA sequence analysis for identification (e.g., A176, A178, and C188 in Fig. 1 and 2) were not considered in these combined patterns. Therefore, only variants without these nucleotide changes are included in this figure. The A176, A178, and C188 nucleotide changes did not, however, affect the DNA sequence covariation between E6 and L1 segments within the major HPV-16 classes. From 247 samples, we observed 22 combined E6-L1 hybridization patterns, of which 11 patterns were novel. The 11 novel patterns were present in only 21 (8.5%) of the specimens. Of these 21 samples, 16 were assigned to the AA-Af1-Af2-NA1 variant classes. Between the E class and the AA-Af1-Af2-NA1 classes, no unanticipated covariation was observed. All 123 of the E class E6 variants showed L1 sequences consistent with E class variants, and all 124 of the AA-Af1-Af2-NA1 class E6 variants contained L1 nucleotide sequences consistent with the AA-Af1-Af2-NA1 classes.

E6, L1, and LCR nucleotide sequence and phylogeny of representative HPV-16 molecular variants. Specimens that were representative of the novel HPV-16 E6-L1 variation pat-

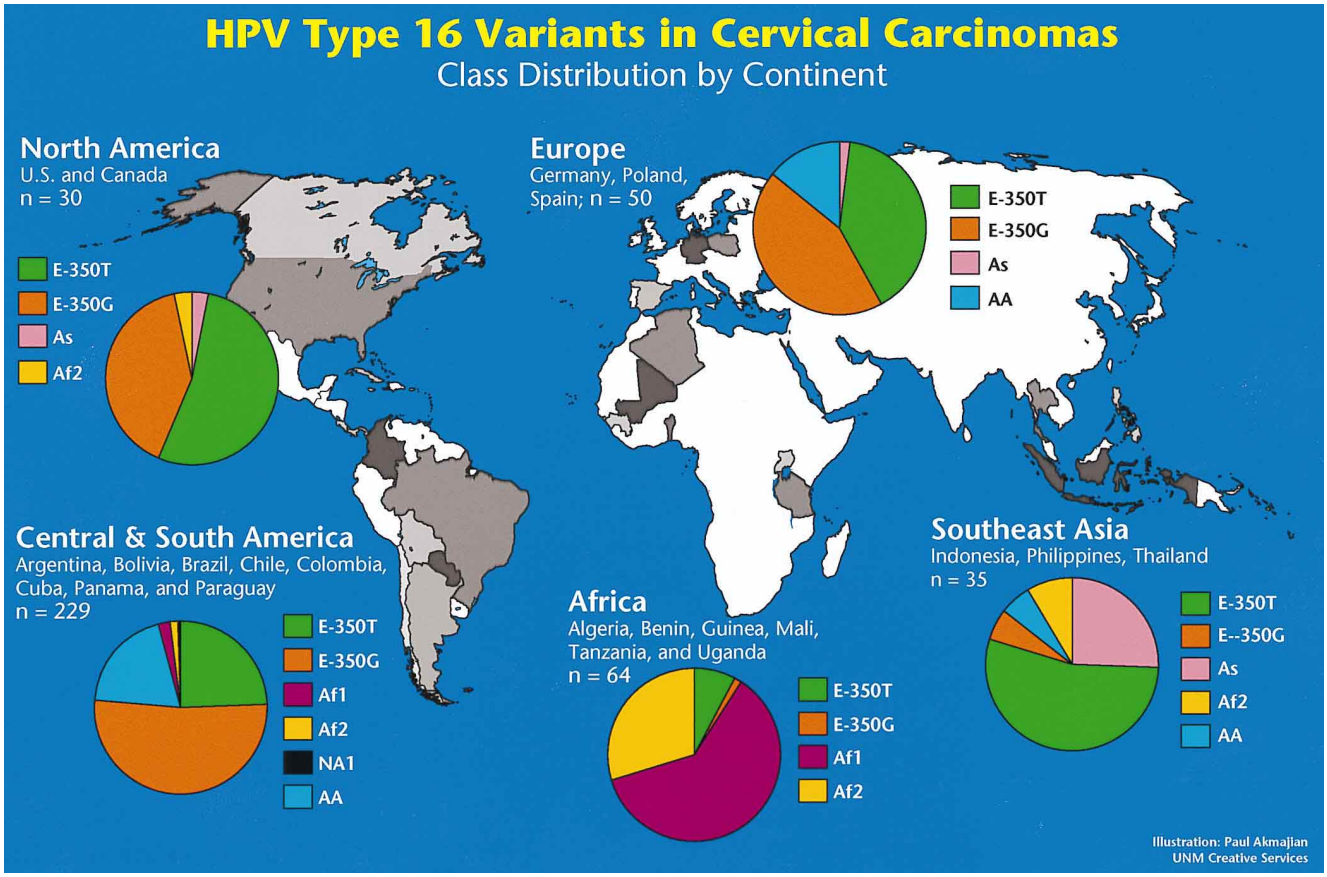

FIG. 5. Global distribution of major HPV-16 classes and subclasses. The assignments of major HPV-16 variants including E, As, Af-1, Af-2, and AA are shown. The European prototype (E) has been divided into variants demonstrating a nucleotide variation within the E6 coding region at nt 350. This E class distinction is represented by E-350T and E-350G categories and excludes As variants. HPV-16 variants sampled in this study were from countries that are shaded.

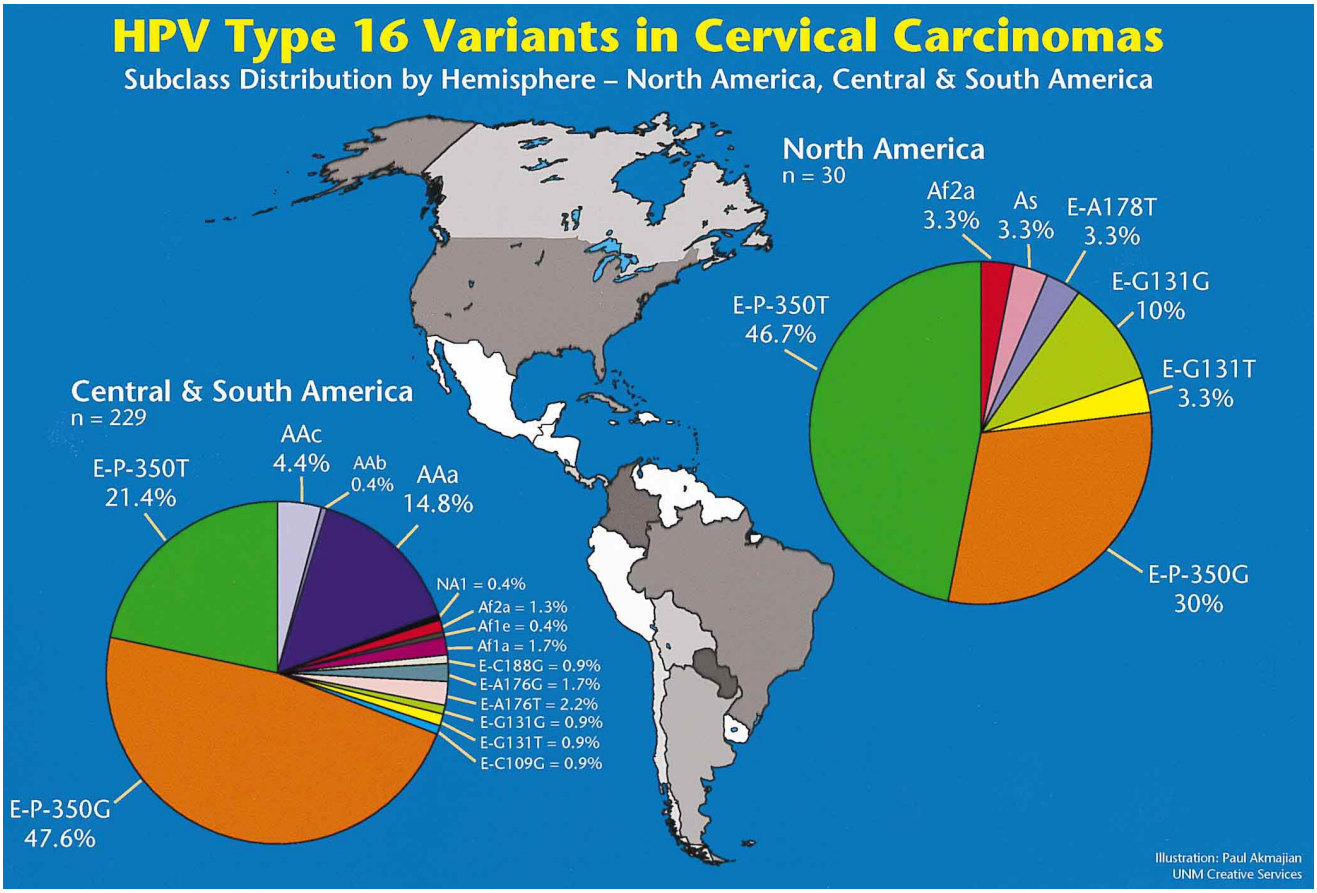

FIG. 6. Distribution of minor HPV-16 classes and subclasses in the Western Hemisphere. The specific class-subclass designations follow the criteria used in Fig. 3 to 5. The major HPV-16 variant designation (i.e., European prototype (E-P), other European (E) variants, Asian (As), African (Af1 and Af2), North American (NA1), and Asian American (AA) is given as the first letters assigned. A letter immediately following a dash and preceding a nucleotide position number represents the variation at the specific nucleotide position designated compared to the reference sequence. Final letters represent E6 nt 350 variations. HPV-16 variants sampled in this study were from countries that are shaded.

terns observed (n = 31) were sequenced across the E6 coding region (nt 104 to 559), the MY09/11 region of L1 (nt 6618 to 6999), and the LCR (nt 7480 to 7843). Sequence differences in the combined 1,203 bp are summarized in Fig. 4, which includes 28 samples from this study (sample designations begin with IS.) and 17 samples that represent the major HPV 16 variants described in our previous studies (33, 36) (designations begin with OR).

These DNA sequence data sets were subjected to phylogenetic analysis. Phylogenetic trees (data not shown) were not more informative than those constructed in our previous study (36) and were even less stable because much less total nucleotide sequence information was analyzed (1,203 versus 3,690 bp). However, fine split branches of novel minor variants enabled us to identify their relationships with major variant classes. We classified minor AA-Af1-Af2-NA1 variants as members of major variant classes accordingly (Fig. 4, right-hand column, Af1-b, Af1-c, etc.).

In total, we identified 13 novel nucleotide changes in the E6 region and 2 novel nucleotide changes in the MY09/11 region. The resulting predicted amino acid changes are shown in Fig. 1. Amino acid substitutions are predicted to result from nine of the E6 and two of the MY09/11 nucleotide changes.

Geographical distribution of the HPV-16 variants. Based on the E6 hybridization patterns and phylogenetic designations, the 408 invasive cervical cancer specimens were classified to the E-350T, E-350G, As, AA, Af1, Af2, and NA1 classes. E-350T and E-350G included all E class samples with T at nt 350 and G at nt 350, respectively, and excluded As variants. The distribution of these HPV-16 variant classes within five continents is shown in Table 1 and Fig. 5 to 7. Particular classes were found to cluster geographically; highly significant increases or decreases in prevalence are noted in Table 1. African variants (Af1 and Af2) accounted for 92% of African specimens and 0 to 9% of those in other regions, while European variants (E) were found in the majority of specimens in all regions other than Africa, ranging from 60% in Southeast Asia to 93% in North America. Asian variants (As) constituted 26% of Southeast Asian specimens but were rare or absent in other continents. Apart from two specimens from Southeast Asia, Asian American (AA) variants occurred only in Central and South America (20% of specimens) and in Europe (14% of specimens; all AA variants in Europe were from Spain). Geographical variation was also observed within the European variant. The E-350G variant was common in Europe and throughout the Americas, but in Southeast Asia and Africa, 89% (24 of 27) of European variants were the E-350T variant.

DISCUSSION

We analyzed HPV-16 sequence variation in a worldwide collection of cervical cancer specimens by using our recently

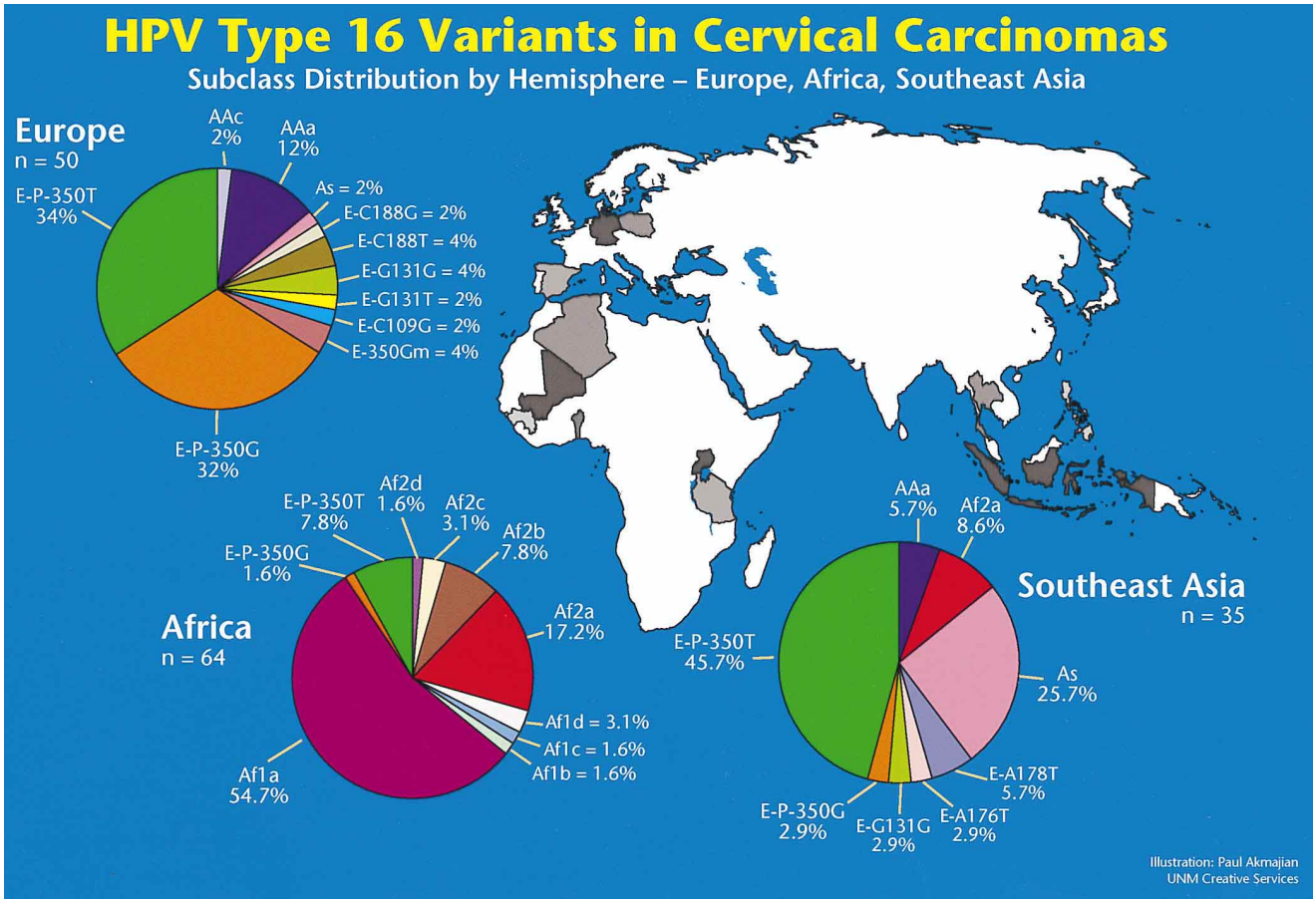

FIG. 7. Distribution of minor HPV-16 classes and subclasses in the Eastern Hemisphere. For details, see the legend to Fig. 6.

developed PCR-based oligonucleotide dot hybridization system (33, 36). Hybridization data were complemented by direct nucleotide sequencing of the targeted E6 and L1 regions, and phylogenetic relationships were examined through additional sequence analyses of the LCR.

The HPV-16 variants identified in our previous study (33, 36) represented the majority of HPV-16 variants detected in this investigation. Although we identified several novel HPV-16 variants in each variant class, 92.2% of the tumor samples examined were shown to contain typical major variants described in previous investigations (9, 13, 17, 25, 29, 31, 33, 36). Intergene covariation between E6 and L1 regions was observed here as well as in our previous investigations conducted in U.S. populations (36), strongly suggesting that recombination events between HPV-16 variants are nonexistent or very rare.

Direct DNA sequence analysis of PCR amplicons was applied in these studies to obviate the detection of *Taq*-induced nucleotide changes. The majority of novel nucleotide variations (22 of 33) reported here were observed in multiple specimens from distinct geographic locations. Some of the more common novel variations were also observed in multiple HPV-16 lineages (i.e., E and AA). These data indicate that the observed nucleotide variations are biological entities not derived from PCR artifacts. Unique nucleotide variations (i.e., those observed once in a single specimen) were present in only 11 of 408 specimens analyzed. It should be noted that multiple independent amplifications of individual specimens were conducted for purposes of hybridization and supplementary DNA sequence.

The geographic distribution of the major HPV-16 variants in our study strongly resembles that previously reported by others (13). These previous studies evaluated specimens from a variety of cervical disease diagnostic categories. The similarity in these data suggests that the HPV-16 variations reported here reflect the variation present in the normal population. In Europe and North America, E class variants predominated; Af1 and Af2 made up the large majority (92%) identified in African tumors; and the As variant was found mainly in Southeast Asia. The AA variant clustered in Central and South America and in Europe, but within Europe AA variants occurred only in Spain. It is tempting to posit that AA variants entered Spain through the Spanish conquistadors, who may have brought them from the Americas. However, the origin of AA variants remains unclear, and thus the proof of such a postulate is elusive. The geographic distribution of HPV types and of specific variants may be influenced by several factors including founder effects, coevolution of HPVs with human races, human migration patterns, and viral fitness measures such as transmissibility.

We found few examples of multiple HPV-16 variants within a single tumor specimen. By the methods we applied, only 3 of 411 invasive cervical samples contained two different HPV-16 variants. Compared with reports by other investigators (15, 34), the prevalence of multiple HPV-16 variant infection in our study was significantly lower (0.7%). This could be due to methodologic bias but may simply reflect the clonality of HPV-16-positive cervical cancers.

The extensive hybridization data from our study confirmed that overall hybridization and probe failure rates for both the E6 and MY09/11 hybridization systems are extremely low (0.2% in the E6 system and 0.03% in the MY09/11 system). This supports the general utility of our PCR hybridization systems (33) to detect HPV-16 variants. It should be noted that these probe-based variant assays are limited in that nucleotide changes are distinguished only within the probe binding sites. While it is probable that novel HPV-16 variants remain, they are likely to represent a small minority of HPV-16 infections worldwide.

HPV protein sequence variation affects virus assembly (18) and may also affect complex characteristics such as carcinogenic potential. Differences in immortalizing activity and in p53 degradation have been reported for HPV-16 variants (6). Host immunologic recognition, perhaps in association with specific HLA haplotypes, may be influenced by amino acid changes in HPV proteins. Some investigations have suggested the association of particular HPV molecular variants and the risk of cervical neoplasia (2, 7, 12, 20, 33, 35). Ellis et al. reported that a particular HPV-16 variant was potentially associated with the HLA B-7 allele and cervical cancer risk (7). We previously identified this E6-G131 sequence variation within an HPV-16 E-class phylogenetic group designated E-G131 (36). In the present study, we detected the E-G131 variant and also found an E6-G131 variant within the Af-1 class ($n = 3$). In our 408 HPV-16-positive cervical cancer specimens, E-G131 variants and Af-G131 variants represented 2.9 and 0.7%, respectively. Given its rarity in this global collection of tumors, it is unlikely that the E6-G131 variation contributes significantly to cervical cancer risk worldwide.

It is clear that some variations in HPV genomes may be determinants of biological behaviors. To date, however, variant identification and sequence characterizations have been limited. The virologic and clinical significance of the observed variations remain unknown. Almost all basic and applied research efforts have utilized HPV-16 clones that represent European variants. Few investigations have included E subclass variants or any of the prevalent non-E classes (e.g., Af1 and AA). This is worrisome, given the high global proportion of these minimally studied variants in cervical cancers.

Diverse and comprehensive research efforts are required to elucidate a biological understanding of HPV intratype variation. We do not yet understand whether amino acid variation in HPV antigens influences host immune response and recognition. Even more complex is the question whether specific virus variants pose distinct challenges in specific host immunogenetic backgrounds (e.g., HLA haplotypes). Investigators have just begun to study serologic cross-reactivity between HPV-16 variants (5), but it is still not known whether immunity to one HPV-16 variant can protect against infection with another variant.

Although this study represents the largest data set of HPV-16 variation reported to date, our investigation was still limited by sample size. Reliable estimates of the prevalence of particular variants in cervical cancers can be obtained only for the Latin America region. Greater numbers of specimens must be analyzed to provide further regional prevalences and country-specific data. Future case-control and longitudial studies will be necessary to estimate the individual disease risk of each molecular variant and to establish the prevalence of HPV variants in specific populations. This study may form a basis for those further investigations. The simplified detection of HPV-16 variants will also facilitate the distinction of recurrent and new infection more reliably in epidemiological studies of infection and persistence.

Numerous investigators are pursuing diverse approaches to the development of HPV-16 vaccines, and some clinical trials are imminent (for a review, see reference 23). It seems prudent to consider the potential biological significance of HPV-16 sequence variation in both the formulation of vaccine strategies and efficacy studies of candidate vaccines.

ACKNOWLEDGMENTS

We thank members of the IBSCC Study Group for providing specimens for this investigation; Keerti V. Shah, in whose laboratory the original IBSCC HPV typing was performed; J. R. (Bob) Herron, Cheri Goodall, and Nicole Wheeler (University of New Mexico) and Angela Jansen (Johns Hopkins University) for persistent technical assistance; J. Icenogle for helpful advice; A. Halpern for the phylogenetic analysis and helpful discussions; Victor Moreno for critical manuscript review; K. Altobelli and K. Trinkaus for computer program development and data analyses; P. Akmajian for graphic creativity, and Lee Fernando for patience in preparing the manuscript. We also thank SLONNA, Inc., for providing depth and breadth to our thoughts about virus phylogeny.

This work was supported by Public Health Service grants AI32917 to C.M.W. and CA57975 to Steven A. Jenison from the National Institutes of the Health and utilized the University of New Mexico Cancer Research and Treatment Center (UNM-CRTC) Molecular Analysis Facility. The IBSCC was funded by NCI contract MA-5623-41 to M.M.M. and K. V. Shah (Johns Hopkins University), the European Community (contract C11-0371-F[CD]), and the Cancer Research Campaign Endowment Fund. T.Y. was a postdoctoral fellow sponsored by the UNM-CRTC and the UNM Department of Medicine.

REFERENCES

1. **Bauer, H. M., C. E. Greer, and M. M. Manos.** 1992. Determination of genital HPV infection using consensus PCR, p. 131–152. *In* C. S. Herrington and J. O. D. McGee (ed.), Diagnositic molecular pathology: a practical approach. Oxford University Press, Oxford, United Kingdom.
2. **Bavin, P. J., P. G. Walker, and V. C. Emery.** 1993. Sequence microheterogeneity in the long control region of clinical isolates of human papillomavirus type 16. J. Med. Virol. **39:**267–272.
3. **Bernard, H.-U., S.-Y. Chan, M. M. Manos, C.-K. Ong, L. L. Villa, H. Delius, C. L. Peyton, H. M. Bauer, and C. M. Wheeler.** 1994. 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. **170:**1077–1085.
4. **Bosch, F. X., M. M. Manos, N. Muñoz, M. Sherman, A. M. Jansen, J. Peto, M. H. Schiffman, V. Moreno, R. Kurman, K. V. Shah, and I. B. S. C. C. Study Group.** 1995. Prevalence of human papillomavirus in cervical cancer: a worldwide perspective. J. Natl. Cancer Inst. **87:**796–802.
5. **Cheng, G., J. P. Icenogle, R. Kirnbauer, N. L. Hubbert, M. E. St Louis, C. Han, E. I. Svare, S. K. Kjaer, D. R. Lowy, and J. T. Schiller.** 1995. Divergent human papillomavirus type 16 variants are serologically cross-reactive. J. Infect. Dis. **172:**1584–1587.
6. **Conrad-Stoppler, M., K. Ching, H. Stoppler, K. Clancy, F. X. Bosch, N. Muñoz, K. Shah, R. Schlegel, and J. Icenogle.** 1996. Natural variants of the HPV 16 E6 protein differ in immortalizing activity and ability to degrade p53. J. Virol. **70:**6987–6993.
7. **Ellis, J. R. M., P. J. Keating, J. Baird, E. F. Hounsell, D. V. Renouf, M. Rowe, D. Hopkins, M. F. Duggan-Keen, J. S. Bartholomew, L. S. Young, and P. L. Stern.** 1995. The association of an HPV16 oncogene variant with HLA-B7 has implications for vaccine design in cervical cancer. Nat. Med. **1:**464–470.
8. **Eluf-Neto, J., M. Booth, N. Muñoz, F. X. Bosch, C. J. L. M. Meijer, and Walboomers.** 1994. Human papillomavirus and invasive cervical cancer in Brazil. Br. J. Cancer **69:**114–119.
9. **Eschle, D., M. Dürst, J. ter Meulen, J. Luande, H. C. Eberhardt, M. Pawlita, and L. Gissmann.** 1992. Geographical dependence of sequence variation in the E7 gene of human papillomavirus type 16. J. Gen. Virol. **73:**1829–1832.
10. **Fujinaga, Y., K. Okazawa, A. Nishikawa, Y. Yamakawa, M. Fukushima, I. Kato, and K. Fujinaga.** 1994. Sequence variation of human papillomavirus type 16 E7 in preinvasive and invasive cervical neoplasias. Virus Genes **9:**85–92.
11. **Fujinaga, Y., K. Okazawa, Y. Ohashi, Y. Yamakawa, M. Fukushima, I. Kato, and K. Fujinaga.** 1990. Human papillomavirus type 16 E7 gene sequence in human cervical carcinoma analysed by polymerase chain reaction and direct sequencing. Tumor Res. **25:**85–91.
12. **Hecht, J. L., A. S. Kadish, G. Jiang, and R. D. Burk.** 1995. Genetic characterization of the human papillomavirus (HPV) 18 E2 gene in clinical specimens suggests the presence of a subtype with decreased oncogenic potential. Int. J. Cancer **60:**369–376.
13. **Ho, L., S.-Y. Chan, R. D. Burk, B. C. Das, K. Fujinaga, J. P. Icenogle, T. Kahn, N. Kiviat, W. Lancaster, P. Mavromara-Nazos, V. Labropoulou, S. Mitrani-Rosenbaum, M. Norrild, M. R. Pillai, J. Stoerker, K. Syrjanen, S. Syrjanen, S.-K. Tay, L. L. Villa, C. M. Wheeler, A.-L. Williamson, and H.-U.**

Bernard. 1993. The genetic drift of human papillomavirus type 16 is a means of reconstructing prehistoric viral spread and the movement of ancient human populations. J. Virol. **67:**6413–6423.
14. **Ho, L., S.-Y. Chan, V. Chow, T. Chong, S.-K. Tay, L. L. Villa, and H.-U. Bernard.** 1991. Sequence variants of human papillomavirus type 16 in clinical samples permit verification and extension of epidemiological studies and construction of a phylogenetic tree. J. Clin. Microbiol. **29:**1765–1772.
15. **Ho, L., S.-K. Tay, S.-Y. Chan, and H.-U. Bernard.** 1993. Sequence variants of human papillomavirus type 16 from couples suggest sexual transmission with low infectivity and polyclonality in genital neoplasia. J. Infect. Dis. **168:**803–809.
16. **Icenogle, J. P., M. Laga, D. Miller, R. A. Tucker, and W. C. Reeves.** 1992. Genotypes and sequence variants of human papillomavirus DNAs from human immunodeficiency virus type 1-infected women with cervical intraepithelial neoplasia. J. Infect. Dis. **166:**1210–1216.
17. **Icenogle, J. P., P. Sathya, D. L. Miller, R. A. Tucker, and W. E. Rawls.** 1991. Nucleotide and amino acid sequence variation in the L1 and E7 open reading frames of human papillomavirus type 6 and type 16. Virology **184:** 101–107.
18. **Kirnbauer, R., J. Taub, H. Greenstone, R. Roden, M. Durst, L. Gissmann, D. R. Lowy, and J. T. Schiller.** 1993. Efficient self-assembly of human papillomavirus type 16 L1 and L1-L2 into virus-like particles. J. Virol. **67:**6929–6936.
19. **Koutsky, L., K. K. Holmes, C. W. Critchlow, C. E. Stevens, J. Paavonen, A. M. Beckmann, T. A. DeRouen, D. A. Galloway, D. Vernon, and N. B. Kiviat.** 1992. A cohort study of the risk of cervical intrepithelial neoplasia grade 2 or 3 in relation to papillomavirus infection. N. Engl. J. Med. **327:** 1272–1278.
20. **Londesborough, P., L. Ho, G. Terry, J. Cuzick, C. Wheeler, and A. Singer.** 1996. Human papillomavirus (HPV) genotype as a predictor of persistence and development of high grade lesions in women with minor cervical abnormalities. Int. J. Cancer **69:**364–368.
21. **Matsukura, T., T. Kanda, A. Furuno, H. Yoshikawa, T. Kawana, and K. Yoshiike.** 1986. Cloning of monomer human papillomavirus type 16 DNA integrated within cell DNA from a cervical carcinoma. J. Virol. **58:**979–982.
22. **Muñoz, N., F. X. Bosch, S. de Sanjosé, L. Tafur, I. Izarzugaza, M. Gili, P. Viladiu, C. Navarro, C. Martos, N. Ascunce, L. C. Gonzalez, J. M. Kaldor, E. Guerrero, A. Lörincz, M. Santamaria, P. Alonso de Ruiz, N. Aristizabal, and K. Shah.** 1992. The causal link between human papillomavirus and cervical cancer: a population-based case-control study in Colombia and Spain. Int. J. Cancer **52:**743–749.
23. **Muñoz, N., L. Crawford, and P. Coursaget.** 1995. HPV vaccines and their potential use in the prevention and treatment of cervical neoplasia. Papillomavirus Rep. **6:**54–55.
24. **Myers, G., H.-U. Bernard, H. Delius, C. Baker, J. Icenogle, A. Halpern, and C. Wheeler.** 1995. Human papillomavirus. A compilation of analysis of nucleic acid and amino acid sequences. Publication LA-UR 95-3675. Los Alamos National Laboratory, Los Alamos, N.M.
25. **Pushko, P., T. Sasagawa, J. Cuzick, and L. Crawford.** 1994. Sequence variation in the capsid protein genes of human papillomavirus type 16. J. Gen. Virol. **75:**911–916.
26. **Sanger, F., S. Nicklen, and A. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
27. **Schiffman, M. H., H. M. Bauer, R. N. Hoover, A. G. Glass, D. M. Cadell, B. B. Rush, D. R. Scott, M. E. Sherman, R. J. Kurman, S. Wacholer, C. K. Stanton, and M. M. Manos.** 1993. Epidemiologic evidence showing that human papillomavirus infection causes most cervical intraepithelial neoplasia. J. Natl. Cancer Inst. **85:**958–964.
28. **Seedorf, K., G. Krämmer, M. Dürst, S. Suhai, and W. G. Röwenkamp.** 1985. Human papillomavirus type 16 DNA sequence. Virology **145:**181–185.
29. **Smits, H. L., K. F. Traanberg, M. R. L. Krul, P. R. Prussia, C. L. Kuiken, M. F. Jebbink, J. A. F. W. Kleyne, R. H. van den Berg, B. Capone, A. de Bruyn, and J. ter Schegget.** 1994. Identification of a unique group of human papillomavirus type 16 sequence variants among clinical isolates from Barbados. J. Gen. Virol. **75:**2457–2462.
30. **Stewart, A.-C. M., A. M. Eriksson, M. M. Manos, N. Muñoz, F. X. Bosch, J. Peto, and C. M. Wheeler.** 1996. Intratype variation in 12 human papillomavirus types: a worldwide perspective. J. Virol. **70:**3127–3136.
31. **Swofford, D. L.** 1993. Phylogenetic analysis using parsimony: PAUP, version 3.1. Illinois Natural History Survey, Champaign, Ill.
32. **Wheeler, C. M., A. Hildesheim, and M. Schiffman.** Human papillomavirus (HPV) type 16 E6 nucleotide variations as cervical disease correlates. Submitted for publication.
33. **Wheeler, C. M., T. Yamada, A. Hildesheim, and S. A. Jenison.** 1997. Human papillomavirus type 16 sequence variants: identification by E6 and L1 lineage-specific hybridization. J. Clin. Microbiol. **35:**11–19.
34. **Xi, L. F., G. W. Demers, L. A. Koutsky, N. B. Kiviat, J. Kuypers, H. Watts, K. K. Holmes, and D. A. Galloway.** 1995. Analysis of human papillomavirus type 16 variants indicates establishment of persistent infection. J. Infect. Dis. **172:**747–755.
35. **Xi, L. F., L. A. Koutsky, J. Kuypers, D. A. Galloway, J. P. Hughes, C. M. Wheeler, K. K. Holmes, and N. B. Kiviat.** Risk of cervical intraepithelial neoplasia grade 2-3 in relation to human papillomavirus type 16 variants. Submitted for publication.
36. **Yamada, T., C. M. Wheeler, A. L. Halpern, A.-C. M. Stewart, A. Hildesheim, and S. A. Jenison.** 1995. Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence ananlysis of the E6, L2, and L1 coding segments. J. Virol. **69:**7743–7753.