JOURNAL OF VIROLOGY, May 1996, p. 3127–3136 Vol. 70, No. 5
0022-538X/96/$04.00+0

Intratype Variation in 12 Human Papillomavirus Types: a Worldwide Perspective

ANN-CHARLOTTE M. STEWART,[1] ANNIKA M. ERIKSSON,[1] M. MICHELE MANOS,[2] NUBIA MUÑOZ,[3] F. XAVIER BOSCH,[4] JULIAN PETO,[5] AND COSETTE M. WHEELER[1]*

Department of Cell Biology and the New Mexico Cancer Research and Treatment Center, University of New Mexico School of Medicine, Albuquerque, New Mexico 87131[1]; Johns Hopkins School of Public Health, Baltimore, Maryland 20205[2]; International Agency for Research on Cancer, Lyon, France[3]; Servei d'Epidemiología i Registre del Càncer, Barcelona, Spain[4]; and Institute of Cancer Research, Belmont, United Kingdom[5]

Received 26 October 1995/Accepted 31 January 1996

In this study, we have examined intratype human papillomavirus (HPV) sequence variation in a worldwide collection of cervical specimens. Twelve different HPV types including HPV-18, HPV-33, HPV-35, HPV-39, HPV-45, HPV-51, HPV-52, HPV-58, HPV-59, HPV-68 (ME180), MM9/PAP238A (recently designated HPV-73), and a novel partial genomic HPV sequence designated MM4/W13B were analyzed in this study. Cervical specimens were collected as part of epidemiological investigations conducted in New Mexico and an international study of invasive cervical cancer (IBSCC). Specimens from several countries including Argentina, Brazil, Bolivia, Benin, Cuba, Colombia, Chile, Germany, Mali, Panama, Paraguay, Spain, Algeria, Uganda, Guinea, Tanzania, Indonesia, Philippines, Thailand, and the United States were evaluated. Specimen DNAs were subjected to amplification with the MY09/11 L1 consensus PCR system. The PCR products were cloned, and an approximately 410-bp region in the L1 open reading frame was sequenced from 146 specimens (~60,000 bp). Within a single HPV type, nucleotide diversity varied between 0.2 and 2.9% (i.e., between any pair of variants) and the majority of nucleotide changes were synonymous (amino acid conserving). These data provide information pertinent to HPV diagnostic probe development and are potentially relevant to future rational vaccine strategies. Similarly, amino acid diversity varied between 0 and 5.1%. Some of these amino acid changes may represent markers of intertype evolutionary relationships. Presuming that HPVs have evolved under the same constraints as their corresponding hosts, the limited genetic diversity observed for all HPVs studied to date may reflect an evolutionary bottleneck occurring in both virus and host populations.

Human papillomaviruses (HPVs) constitute a group of viruses associated with benign and malignant neoplasia of cutaneous and mucosal epithelia (33, 67). To date, more than 70 different HPV types have been identified (16). Evidence from partial sequences suggests the existence of at least 13 additional types that would qualify as novel HPVs (4, 5, 51), and it is likely that unidentified HPVs remain. About half of the reported HPVs are associated with mucosal lesions, including cervical neoplasias (15, 61). A strong and consistent association has been found between infection with certain types of HPV (e.g., HPV-16) and invasive cervical carcinoma (20, 36). HPV-16 represents about 50% of cervical cancer-associated HPV infections worldwide, although regional variations have been reported (6).

Papillomaviruses (PVs) are defined by genomic sequence similarities rather than by classical serology. An HPV genome is defined as a new type if it is separated by a Hamming distance or dissimilarity of more than 10% in its nucleotide sequence compared with other known HPV types in the E6, E7, and L1 open reading frames (ORFs) combined (15, 61). Isolates within the same type differing by 0 to 2% in their nucleotide sequences compared with the reference sequence are referred to as variants, and those differing by 2 to 10% are referred to as subtypes.

* 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. Phone: (505) 277-9151. Fax: (505) 277-9494. Electronic mail address: cwheeler@medusa.unm.edu.

PVs are ubiquitous and have been detected in all human populations studied to date (6, 28). They have also been identified in many animals including cows, horses, dogs, birds, deer, and rabbits (55). Recent systematic sequencing efforts have made PVs one of the best studied DNA virus families and a resource for studies of viral evolution (9, 14). Studies of phylogenetic relationships and taxonomic grouping of these viruses have, however, still not clarified how the many different types of PVs have evolved. The wide range of host-specific PVs and the adaptation to cutaneous or mucosal epithelia support the role of selection for HPV types and lineages specifically adapted to different ecological niches. The large number of types occupying similar ecological niches, on the other hand, suggests that mutation and/or random extinction of lineages must have played an important role in the development of the current range of PV types as well.

PVs are host specific, and most evidence suggests that recombination between viral genomes is either nonexistent or extremely rare event (30, 66). In addition, PVs are largely dependent upon the host cellular machinery and use the host DNA polymerase for the replication of their genomes. Therefore, it is reasonable to consider that resultant genomic diversity may be constrained similarly for both virus and host (10, 39, 50). As has been suggested in previous studies, PVs and other persistently infecting small DNA viruses may represent useful tools for studying the coevolution of virus and host (50). An extension of this suggestion is that the limited polymorphism observed in our data from 12 different HPV types and novel sequences reflects a previously proposed population bottleneck (45) occurring in the host population.

Limited information is available concerning HPV intratype variation except, for HPV-6, HPV-11, HPV-16, and HPV-18 (10, 21, 26, 28–30, 32, 39, 42, 52, 65, 66). Sequence analyses of ORFs, such as E5, E6, E7, L1, and L2, and of the upstream regulatory region (URR) in more prevalent HPVs (e.g., HPV-16 and HPV-18) have indicated that intratype polymorphism is very limited. Intratype variation observed in these HPV types is commonly less than 2.5% intragenically (within ORFs) and 5% extragenically (within the URR). While most sampling and sequencing efforts have been biased toward a search for more divergent HPV genomes, recent investigations have failed to identify intermediate levels of divergence. HPV subtypes have been encountered very rarely. Five examples have been reported; they include subtypes of HPV-5 and HPV-8, the HPV-34–HPV-64 pair, the HPV-44–HPV-55 pair, and HPV-68 and the genome identified in the ME180 cell line (12, 13, 16, 43).

Analysis of intratype genomic diversity of HPV sequences has several objectives. It is important to establish databases for (i) development of efficient diagnostic tools and vaccines, (ii) application to epidemiological studies in which sequence variations can be used as markers for monitoring HPVs in defined populations, (iii) studies of relationships between genotype and phenotype (i.e., certain variants may be positively or negatively associated with cervical disease, depending on their different biological and functional properties), and (iv) evolutionary and taxonomic studies.

During our previous efforts to identify novel HPVs, we observed intratype variation in the MY09/11 region of the L1 ORF, especially in less prevalent HPVs (40). To further consider these observations and to assess the intratype variation in these HPV types and novel sequences, we have characterized HPVs detected in cervical specimens that were collected during epidemiologic investigations of HPV-associated cervical disease. Specimens were obtained from a collection of cervical cancers in patients from 22 countries throughout the world (6) and from studies performed in New Mexico (3, 64). In addition to evolutionary, diagnostic, and epidemiological aspects, this study afforded the opportunity to consider geography and, potentially, ethnicity as factors in HPV genomic diversity for some of the HPV types and sequences under study.

MATERIALS AND METHODS

Clinical specimens, DNA preparation, and reference sequences. Specimens included invasive cervical tumor biopsy specimens ($n = 135$) (6) or cervical cells ($n = 11$) (3, 64). Crude proteinase digests were prepared and specific HPV types were determined as previously described for each epidemiological investigation. This study included the analysis of HPV-18, HPV-33, HPV-35, HPV-39, HPV-45, HPV-51, HPV-52, HPV-58, HPV-59, HPV-68 (ME180), MM4/W13B (novel partial genomic sequence), and MM9/PAP238A (recently designated HPV-73). The number of specimens analyzed for each HPV type and novel sequence is shown in Table 1. Sequences for all reference HPVs were obtained from GenBank and have recently been published as part of a compendium (37).

PCR and DNA sequencing. Amplification of a ~450-bp HPV-specific segment from the L1 ORF was performed as described previously, with the MY09/11 L1 consensus primer system (2). This region corresponds in the HPV-16 genome to nucleotides (nt) 6580 to 7030. PCR products were subjected to electrophoresis in 1% NuSieve agarose using 1× Tris-acetate-EDTA buffer and visualized under UV light. Amplimers of the expected 450-bp size were excised from the gels and extracted with Gelase (Epicentre Technologies, Madison, Wis.) as described by the manufacturer. Gel-purified amplimers were ligated to a pGEM-T vector (Promega, Madison, Wis.). Individual transformants were subjected to direct colony PCR with M13 forward and reverse primers. Resultant PCR products were subsequently screened by hybridization with HPV type-specific probes (2, 27).

Plasmid minipreps of cloned HPVs with correct insert size and positive by hybridization assays were prepared as previously described (54). At least two clones from each clinical specimen were subsequently subjected to dideoxy sequencing (46) with ^{35}S-dATP and Sequenase 2.0 (U.S. Biochemical, Cleveland, Ohio). Both forward- and reverse-strand sequences were obtained for each clone. Specific sequencing primers in the pGEM-T vector, pGEM-F (5′-ACGGCCAGTGAATTGTA-3′) and pGEM-R (5′-ACGCGTTGGGAGCTCTC-3′) were used to facilitate sequencing efforts (54). pGEM-F and pGEM-R are closer to the pGEM-T insertion site than are standard M13 sequencing primers.

TABLE 1. Comparison of intratypic nucleotide and amino acid sequence variation in the MY09/11 L1 region for 13 HPV types

HPV type	No. of specimens examined	Total no. of nt differences from reference[a]	Minimum and maximum nt difference from reference[b]	Maximum difference between any pair of variants for:	
				nt [no. (%)]	aa [no. (%)]
HPV-16[c]	29	11	0, 7	9 (2.2)	3 (2.1)
HPV-18	44	11[d]	0, 7	9 (2.2)	4 (2.9)
HPV-33	10	4	0, 2	3 (0.7)	1 (0.7)
HPV-35	8	1	0, 1	1 (0.2)	1 (0.7)
HPV-39	8	7	1, 5	5 (1.2)	2 (1.5)
HPV-45	21	16	0, 9	9 (2.2)	7 (5.1)
HPV-51	6	1	0, 1	1 (0.2)	0 (0)
HPV-52	13	19	0, 10	12 (2.9)	3 (2.1)
HPV-58	9	10	0, 8	9 (2.2)	5 (3.6)
HPV-59	11	4	0, 3	3 (0.7)	2 (1.5)
HPV-68	6	3	0, 3	3 (0.7)	0 (0)
MM4	5	7	0, 5	6 (1.5)	4 (2.9)
MM9	5	8	4, 6	6 (1.5)	4 (2.9)

[a] All varied nt positions in all HPV variants observed for each HPV type are included.
[b] Per ~410 nt sequenced per specimen.
[c] HPV-16 data from reference 66.
[d] Two sequencing errors detected compared with the HPV-18 reference sequence are not included.

Phylogenetic analysis. Phylogenetic analyses were performed over the MY09/11 L1 region with 53 distinct HPV sequences. These 53 HPV sequences were obtained from 53 separate clinical specimens. Twelve different HPV types and novel sequences and the corresponding reference sequences were represented in the analyses. For each type, an alignment of the distinct MY09/11 sequences (including the sequence of the reference clone from GenBank) was constructed by using the complete MY09/11 region except for the primer-binding sites. The most parsimonious phylogenetic trees were determined with PAUP 3.1.1 software (56) on a Power Macintosh 7100, using the branch-and-bound search strategy, which is guaranteed to find the most parsimonious tree(s). In some cases, several equally parsimonious trees were obtained.

Nucleotide sequence accession numbers. Nucleotide sequences have been submitted to GenBank and will be released upon publication. Accession numbers have been assigned as follows: U45889 to U45894 to the HPV-18 sequences, U45895 to U45897 to the HPV-33 sequences, U45898 to the HPV-35 sequence, U45899 to U45905 to the HPV-39 sequences, U45906 to U45916 to the HPV-45 sequences, U45917 to the HPV-51 sequence, U45918 to U45923 to the HPV-52 sequences, U45924 to U45929 to the HPV-58 sequences, U45930 to U45933 to the HPV-59 sequences, U45934 to the HPV-68 (ME180) sequence, U45935 to U45937 to the MM9 sequences, and U45939 to the MM4 sequence, IS1019. Additional MM4 sequences including IS766, IS887, and IS1016 were previously assigned GenBank accession numbers U12483, U12484, and U12482, respectively (40).

RESULTS

L1 ORF intratype nucleotide sequence variation for HPV types and novel sequences HPV-18, HPV-33, HPV-35, HPV-39, HPV-45, HPV-51, HPV-52, HPV-58, HPV-59, HPV-68, MM4, and MM9. We observed sequence polymorphism in the HPV L1 ORF by restriction fragment length polymorphism techniques during our previous efforts to identify novel HPV sequences, especially in less prevalent HPV types (40, 41). We have now determined the intratype nucleotide variation in the MY09/11 L1 fragment from 11 different HPV types and one novel sequence (MM4/W13B).

Table 1 and Fig. 1 present a nucleotide comparison with the corresponding reference HPV sequences of the 12 HPVs under study. For purposes of orientation, we have used the first nucleotide of the specific HPV sequence (37) as nt 1 for each individual HPV type. When considering all nucleotide substi-

tutions in all MY09/11 variants (excluding 5′ and 3′ primer regions), variation within a single HPV type or novel sequence ranged from no changes (HPV-33, HPV-35, HPV-45, HPV-51, HPV-52, HPV-58, HPV-59, HPV-68, and MM4) to a maximum of 19 changes (HPV-52). Nucleotide changes between two individual specimens ranged from no changes (HPV-18, HPV-33, HPV-35, HPV-45, HPV-51, HPV-52, HPV-58, HPV-59, HPV-68, and MM4) to a maximum of 9 changes in IS761 (an HPV-45 variant) and 10 changes in IS464 (an HPV-52 variant). In some of the HPV types analyzed, isolates could be classified into variant clusters.

In the 44 samples of HPV-18, two main groups were observed, one prototype or prototype-like and one with variants containing six or more changes compared with the reference sequence (Fig. 1A). In the following text, the word "prototype" or "prototype-like" refers to the corresponding reference clone and published sequence of the original HPV isolate. The majority (77%) of the HPV-18 specimens analyzed fell into the prototype group. One specimen from Chile, IS326, had two substitutions at nt positions 6677 and 6697, that were not observed in any of the other HPV 18 specimens analyzed. Another specimen, IS664, had two nucleotide changes not detected in any of the other specimens analyzed, a T-to-C substitution at position 6877 and an A-to-G substitution at position 6943. These two HPV-18 variants do not seem to form intermediate classes between the two major HPV-18 groups but may be representative of unique groups of HPV-18 variants. The nine remaining HPV-18 specimens (21%) all had six or more changes, and all but one (IS227 from Cuba) were obtained in Africa.

None of the HPV-18 specimens analyzed contained the previously reported HPV-18 reference sequence. All 44 of the HPV-18 specimens analyzed had two nucleotide changes, a C-to-G substitution at position 6625, resulting in a substitution (proline to arginine) at amino acid (aa) 399 in the L1 ORF, and a synonymous C-to-G substitution at nt 6842. The nucleotide substitution at position 6625 has been previously reported (38), whereas the nucleotide change at position 6842 has not. These changes could reflect the extreme rarity of the reference sequence that originally was obtained from a subject in northeastern Brazil (11, 17). Alternatively, these changes could be due to cloning artifacts or sequencing errors in the original HPV-18 reference clone. Sequence analysis in our laboratory of the reference clone obtained from the Deutsches Krebsforshungszentrum, Papillomavirus Referenzzentrum, Heidelberg, Germany, revealed that this clone also had the two nucleotide changes at positions 6625 and 6842. Consequently, we conclude from our data that the previously reported HPV reference has two sequencing errors at nt 6625 and 6842.

The HPV-39 specimens ($n = 10$) analyzed did not form any obvious discrete groups. Four of the specimens were prototype-like and had either one change each at nt 6638 (T to A) and 6903 (C to T) (IS073 and IS270, respectively) or both of these changes (IS114 and IS281) (Fig. 1D). The remaining six specimens had three to six changes each. All but one (IS270) HPV-39 specimens had the C-to-T substitution at nt 6903, and none of the specimens were identical to the HPV-39 reference sequence.

HPV 45 specimens ($n = 21$) could be classified into at least three major groups (Fig. 1E). Five of the specimens (23%) were classified as prototype or prototype-like. All variants within this group, except one that had a G-to-A change at nt 6861, had no substitutions when compared with the HPV-45 reference sequence. The largest (36%) group of HPV-45 variants had six to eight substitutions each, with six of these changes (at nt 6677, 6687, 6705, 6816, 6837, and 6861) being common to all of these variants. Four HPV-45 variants (18%) that were observed solely in African specimens shared two "marker" substitutions at nt 6665 and 6914. The remaining five specimens appeared as an intermediate group between the prototype/prototype-like group and the group containing African specimens, sharing some of the marker substitutions from both groups. It was not possible to make any geographical assignments to any of the variants except for those of African origin, since the two other groups included specimens from various continents.

In this study, HPV-52 specimens ($n = 13$) showed the greatest number of intratype nucleotide substitutions (Fig. 1G). A total of 19 substitutions were observed in all HPV-52 variants combined. Of the HPV-52 specimens, 60% could be classified as prototype or prototype-like; all these specimens had no nucleotide changes ($n = 7$) or two changes ($n = 1$) (IS705). The remaining five specimens had between 4 and 10 changes each. Only one substitution, C to A at nt 6917, was shared among these six HPV-52 sequences.

Specimens within the HPV-58 group ($n = 9$) did not form any natural groups, with the exception of prototype or prototype-like (Fig. 1H). The majority of the samples (56%) could be considered prototype or prototype-like, whereas the remaining sequences (44%) had three to eight nucleotide changes and shared only one nucleotide substitution at position 240 (C to A).

HPV-59 specimens ($n = 11$) showed very limited polymorphism (Fig. 1I). Two specimens (18%) had no substitutions compared with the reference sequence. Four HPV-59 specimens had two substitutions at nt 6653 (T to C) and 6950 (C to A) and formed the largest group of HPV-59 variants. Another four HPV-59 specimens shared the C-to-A substitution with the former HPV-59 group but had a G rather than the C substitution at nt 6653. One of these specimens, IS583, had an additional T-to-C change at nt 6616 that was not detected in any of the other HPV-59 specimens analyzed. One specimen, IS598, also shared the C-to-A nucleotide change observed in the two former HPV-59 groups but had an additional nucleotide change at position 6652 (T to G) that was not observed in any of the other HPV-59 specimens analyzed.

MM4 ($n = 5$) and MM9 ($n = 5$) specimens showed a relatively large number of nucleotide substitutions; however, the sample sizes for these HPV sequences were small (Fig. 1K and L). One MM4 specimen (23981) was classified as prototype, while the remainder had one to five nucleotide changes each. None of the MM9 specimens were identical to the reference sequence, and most ($n = 4$) of the specimens had five or six nucleotide substitutions. The remaining specimen, IS324, had four nucleotide substitutions, two (at nt 102 and 303) that were not observed in any of the other specimens analyzed (Fig. 1L).

HPV-33 specimens ($n = 10$), HPV-35 ($n = 8$), HPV-51 ($n = 6$), and HPV-68 ($n = 6$) showed limited total nucleotide variation, with a single nucleotide substitution identified in HPV-35 ($n = 6$) and HPV-51 ($n = 1$; t273), respectively (Fig. 1C and F). Three nucleotide substitutions were observed in HPV-68 ($n = 1$; IS362), and four nucleotide substitutions were observed in HPV-33 ($n = 4$) (Fig. 1J and B). The remaining specimens containing these four HPV types were all identical to the reference sequence.

Figure 2 shows the intratype phylogenetic relationships of variants of all 12 HPV types and novel sequences analyzed in this report plus an HPV-16 phylogenetic tree based on data obtained elsewhere (66). All trees were determined as described in Materials and Methods.

Intratype nucleotide substitutions were not distributed evenly over the MY09/11 region. The positions at which nu-

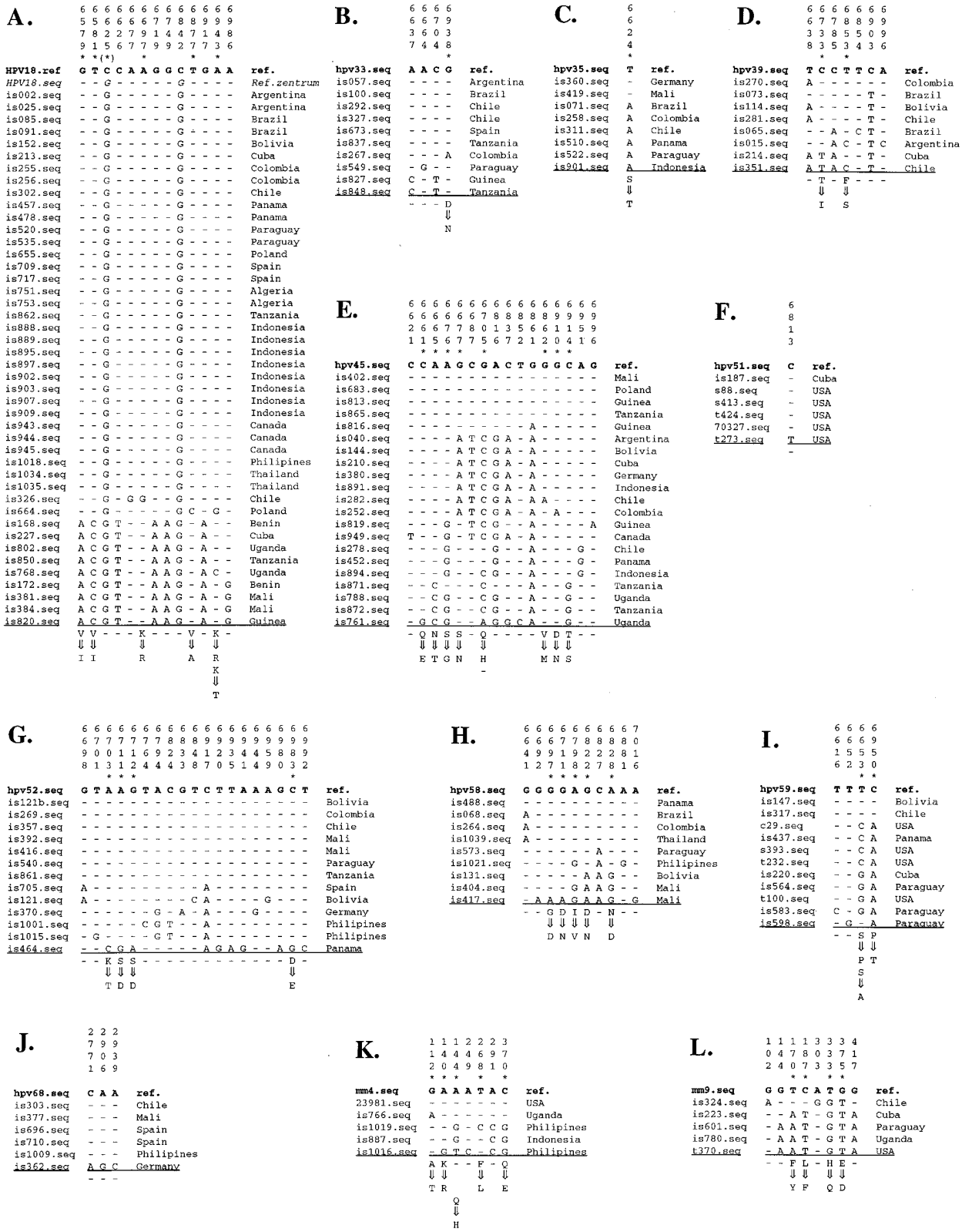

FIG. 1. Nucleotide and amino acid sequence variation in isolates of HPV types and sequences HPV-18, HPV-33, HPV-35, HPV-39, HPV-45, HPV-51, HPV-52, HPV-58, HPV-59, HPV-68, MM4, and MM9. ORF L1 nucleotide positions at which variations were observed are given across the top of the figure in each panel. The numbering refers to the first nucleotide of each specific HPV genome for all HPV types except HPV-68, MM4, and MM9, for which the numbering refers to the first nucleotide of the sequence for each type as presented in GenBank and in human papillomaviruses 1994 compendium (37). The reference sequence used for each type is indicated as ref. Each row in each panel of the figure indicates, from left to right, the following information: the study specimen identification codes, the nucleotide sequence alignment compared with reference sequence, and the origin of each specimen on the right. Specimens obtained from the IBSCC begin with the letters IS; all remaining specimens were obtained from studies performed in Albuquerque, N.Mex. For each variant sequence, positions that do not vary relative to the HPV reference nucleotide or amino acid sequence are marked with a dash in the alignment. Nucleotide changes resulting in nonsynonymous amino acid changes are indicated above the reference sequence for each HPV type with the symbol *, and the corresponding amino acid change is indicated below the line at the bottom of each HPV grouping of like sequences. The symbol (*) in panel A indicates a corrected nonsynonymous amino acid change as a result of a sequencing error detected in the HPV-18 reference sequence.

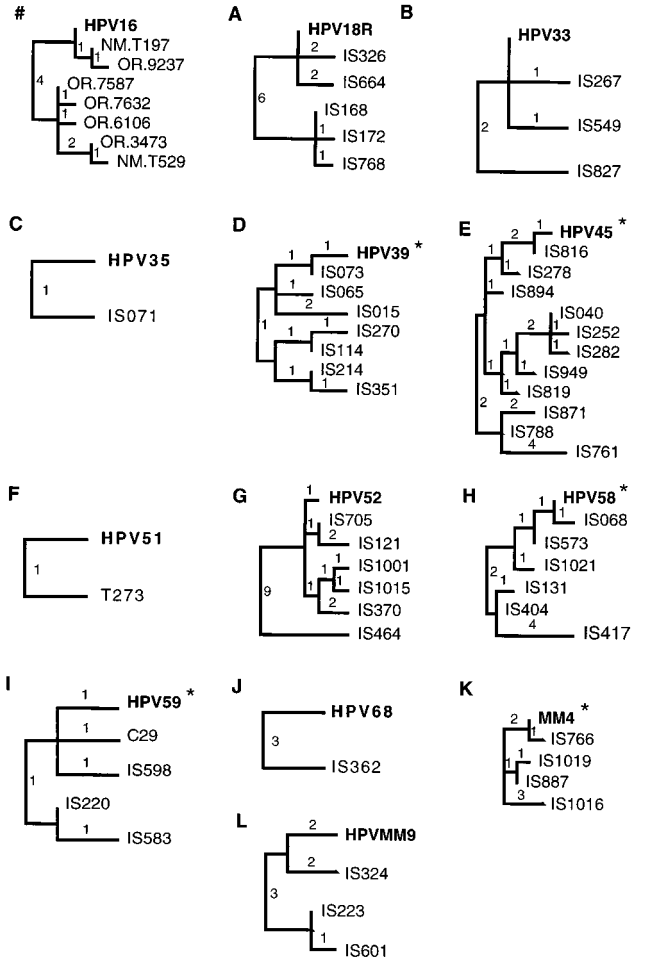

FIG. 2. Phylogenetic trees of variants of HPV types and novel sequences HPV-16, HPV-18, HPV-33, HPV-35, HPV-39, HPV-45, HPV-51, HPV-52, HPV-58, HPV-59, HPV-68, MM4, and MM9. Shown are phylogenetic trees of variants for 11 HPV types and 1 partial genomic HPV sequence. These designations for each HPV correspond to those assigned in Fig. 1. A previously reported phylogenetic tree (66) for HPV-16 is also provided for purposes of comparison. These analyses are based on parsimony evaluation of the MY09/11 L1 region (primer sites excluded). Branch lengths are proportional to the numbers of steps between reconstructed bifurcations. Trees with an asterisk next to the reference locus (e.g., HPV-39*) are representatives of several equally parsimonious trees.

cleotide and amino acid intratype variations were observed are shown in Fig. 3. The ratio of synonymous to nonsynonymous changes varies greatly for individual HPV types (0 to 100%), although this is most probably an effect of the limited sample numbers under study. The variable regions V1 and V2 (Fig. 3) appear to contain a larger than average number of nonsynonymous nucleotide substitutions. V1 and V2 roughly overlap with the more variable portions of the MY09/11 region between types. This would be expected if the selective pressures are similar for all HPV types.

Intratype amino acid variation in HPV types and sequences HPV-18, HPV-33, HPV-35, HPV-39, HPV-45, HPV-51, HPV-52, HPV-58, HPV-59, HPV-68, MM4, and MM9. For purposes of orientation, we have designated the first amino acid of the presumed L1 coding region as aa 1 for each individual HPV type. MM4 and MM9 are exceptions to this numbering, because available sequence information for these is limited to the MY09/11 L1 region (35). For MM4 and MM9, the numbering refers to the first amino acid of the MY09/11 region.

Nucleotide changes observed in the MY09/11 L1 region resulted in fewer amino acid changes (Fig. 1 and 3). The pairwise amino acid variation within a single HPV type ranged from no changes (e.g., HPV-51 and HPV-68) to a maximum of seven changes in HPV-45 (Table 1). Amino acid variation compared with the reference in individual specimens ranged from no changes in HPV-33, HPV-35, HPV-39, HPV-45, HPV-51, HPV-52, HPV-58, and HPV-68 to a maximum of five changes in IS417, an HPV-58 variant. None of the HPV variants examined showed insertions or deletions relative to the sequence of the corresponding reference clone.

All HPV-18 specimens analyzed ($n = 44$) had a proline-to-arginine change at aa 399. This results from a C-to-G nucleotide sequencing error in the reference sequence observed at nt 6625 (this amino acid change is not shown in Fig. 1A). Most of the HPV-18 specimens (75%) were identical to the reference sequence and had only this corrected amino acid change. The remaining 11 specimens had two or three additional amino acid changes each, with specimens IS326 and IS664 possibly forming two separate HPV-18 variant groups.

HPV-33, HPV-35, and HPV-39 showed limited variation, with one amino acid position substituted in HPV-33 and HPV-35 at aa 449 and 351, respectively (Fig. 1B through D). Whereas most (75%) of the HPV-35 specimens had a serine-to-threonine substitution at aa 351, only one (10%) HPV-33 specimen had an aspartic acid-to-asparagine substitution at aa 449. Two positions (aa 364 and 404) were substituted in HPV-39 variants. Most (62.5%) of the HPV-39 specimens were identical to the reference peptide sequence. The remain-

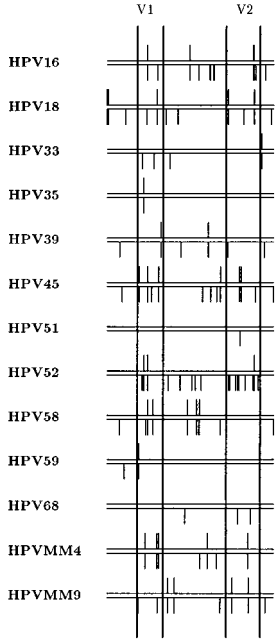

FIG. 3. Distribution of nucleotide and amino acid substitutions among HPV types and novel sequences HPV-16, HPV-18, HPV-33, HPV-35, HPV-39, HPV-45, HPV-51, HPV-52, HPV-58, HPV-59, HPV-68, MM4, and MM9 within the MY09/11 segment of the L1 ORF. The sequenced MY09/11 portion of the L1 ORF (primer sites excluded) is represented by horizontal lines for each HPV type. Vertical bars below the line represent the positions of nucleotide substitutions, while vertical bars above the lines represent predicted amino acid changes. The regions V1 and V2 represent regions with apparently larger number of nonsynonymous mutations than average. V1 and V2 were defined by visual inspection.

TABLE 2. Sequences of HPV type-specific probes

HPV probe[a]	Sequence, 5′ to 3′[b]	nt changed
Probe 18	TGCTTCTACACAGTCTCCTGTACCTGGGCA	None
HPV 18 (MY130)	GGGCA**A**TATGATGCTACCAA***T***[c]	G, 6676; *A, 6692*[c]
(WD74)	GGATGCTGCACC**G**GCTGA	6917
Probe 33	TTTATGCACACAAGT**A**ACTAGTGACAGTAC	C, 6637
HPV 33 (MY16)	CACACAAGT**A**ACTAGTGACAG	C, 6637
(MY64)	TCCTTTGGAGGTACTGTTTTT	None
Probe 35	GTCTGTGTGTTCTGCTGTGTCT**T**CTAGTGA	A, 6624
HPV 35 (MY115)	CTGCTGTGTCT**T**CTAGTGACAG	A, 6624
(MY117)	ATCATCTTTAGGTTTTGGTGC	None
Probe 39	TCTACCTCTATAGAGTCTTCCATACCTTCT	None
HPV 39 (MY89)	TAGAGTCTTCCATACCTTCTAC	None
(MY90)	AGACACTTACAGATAC**C**TACAG	T, 6903
Probe 45	ACACAAAA**A**TCCTGTGCCA**AG**TACATATGA**C**	C, 6665; G, 6676; A, 6677; T, 6687
HPV 45 (MY69)	ATA**C**TACACCTCCAGAAAAGC	G, 6914
(MY129)	GCACAGGA**T**TTTGTGTAGAG	G, 6665
Probe 51	AGCACTGCCACTGCTGCGGTTTCCCCAACA	None
HPV 51 (MY87)	TATTAGCACTGCCACTGCTG	None
(MY88)	CCCAACATTTACTCCAAGTAAC	None
Probe 52	TGCTGA**G**GT**T**TA**A**AAAGGAA**AG**CACATATAA	A, 6698; G, 6701; C, 6703; G, 6711; A, 6712
HPV 52 (MY81)	**C**AC**T**TCTACTGCTATAAC**T**TGT	A, 6917; G, 6920; A, 6935
(MY82)	ACACACCACCTAA**A**GGAAAGG	G, 6959
Probe 58	ATTATGCACTGAAGTAACTAA**G**GAAG**G**TAC	A, 6692; A, 6697
HPV 58 (MY94)	AGCACCCCCTAAAGAAAAGGA	None
(MY179)	GACATTATGCACTGAAGTAACTAA**G**	A, 6692
HPV 59 (MY123)	GCCAGTTAAACAGGACCC	None
(MY162)	CCTAATGAWATACACACCTACCAG	None
HPV 68 (MY191)	CATACCGCTATCTGCAATCAG	None
(MY194)	CTACTACTGAATCAGCTGTACC	None
MM4 (MY164)	CTCAATCTGTT**G**CACAAACA	A, 112
(MY165)	TAACCTTGCCCCCCTC**A**G	C, 291
MM9 (MY104)	GTAGGTACACA**G**GCTAGTAGCTC	A, 102
(MY106)	AGTTGCCAACGTCCTCAAC	None

[a] Type-specific probes named Probe XX are from reference 31. Probes named HPV XX or MYX are from reference 27.
[b] Nucleotides in boldface type and underlined represent mismatched nucleotide positions found in the present study.
[c] The T-to-A nucleotide substitution at position 6692 (marked in italics) represents an error found in the published sequence (27) of this oligonucleotide.

ing three specimens had either a threonine-to-isoleucine substitution at position 364 (IS214), a phenylalanine-to-serine substitution at position 404 (IS015), or both of these substitutions (IS351).

HPV-45 specimens showed intratype amino acid variation ranging from zero to four changes within the analyzed MY09/11 region (Fig. 1E). Five (24%) of the HPV-45 specimens showed no variation compared with the reference sequence and could be classified as prototype. Another five specimens (24%) had two amino acid substitutions at aa 383 and 392, serine to asparagine and glutamine to histidine, respectively. Two other specimens, IS252 and IS282, shared these substitutions but had additional amino acid substitutions at positions 461 and 445, respectively. Four specimens of African origin all shared substitutions at aa 379 and 462. These two amino acid substitutions were not observed in any of the other specimens. Along with these two amino acid changes, these specimens had one (IS871) or two (IS788, IS872, and IS761) other amino acid changes. The remaining specimens (24%) had either a single substitution at aa 383 or an additional substitution at aa 392.

The majority of HPV-52 specimens (92%) showed no amino acid changes compared with the reference sequence. A single specimen obtained in Panama (IS464) had three amino acid substitutions at positions 380, 383, and 473 (Fig. 1G).

The majority (55%) of HPV-58 specimens contained no amino acid substitutions compared with the reference sequence. The remaining four specimens had one, two, three, or five amino acid changes at aa 378, 383, 412, 420, and 422 (Fig. 1H).

Two HPV-59 specimens (IS147 and IS317) had no substitutions compared with the HPV-59 reference sequence (Fig. 1I). Most (82%) of the HPV-59 specimens had one (IS598) or two amino acid substitutions at aa 350 and 449. The proline-to-threonine substitution at aa 449 was shared by all these specimens. Four of these specimens had a serine-to-alanine substitution at aa 350, and four others had a serine-to-proline substitution at the same position.

Amino acid variation in MM4 specimens ranged from no

amino acid changes in one specimen (23981) to a maximum of three changes in specimen IS1016 (Fig. 1K). The remaining three specimens had one amino acid substitution, a glutamine-to-glutamic-acid substitution at aa 124 (IS887 and IS766), or two amino acid changes (IS1019). Three of the MM4 specimens (IS1019, IS887, and IS1016) shared the glutamine-to-glutamic-acid substitution at aa 124.

MM9 specimens formed two groups based on amino acid variation (Fig. 1L). One specimen, IS324, had two amino acid substitutions at positions 111 and 125. The remaining four specimens (80%) had, in addition, two substitutions at aa 57 and 63. Consequently, none of the MM9 specimens were identical to the MM9 reference sequence.

HPV-51 and HPV-68 specimens had no amino acid substitutions compared with the reference sequences (Fig. 1F and J).

Evaluation of oligonucleotide probe systems previously developed for HPV diagnostics. Table 2 lists HPV type-specific oligonucleotide probes from two widely used L1-based PCR detection systems. The MY09/11 (1, 27, 47–49) and GP5+/6+ (31, 44, 53, 59, 60) assays have been previously used in the detection of several HPV types. The GP5+/6+ system targets a 150-bp fragment internal to the MY09/11 primers. PCR amplification in both systems is often followed by type-specific hybridization detection with oligonucleotide probes in pairs (27), as singlets (31), or as cocktails (31). We have compared the published oligonucleotide probe sequences used in these systems with sequence variants observed in the 12 different HPV types and novel sequences included in our current study. The data presented in Table 2 suggest that many of the oligonucleotide probes used in these PCR-based systems may not hybridize or may hybridize only weakly to HPV variants identified in our present study. One or more mismatched nucleotide positions were observed in several cases. Type-specific, "single-use" probes developed for HPV-45, HPV-52, and HPV-58 (31) (Table 2, probes 45, 52, and 58) have three, five, or two mismatched nucleotide positions when compared with variant sequences of HPV-45, HPV-52, and HPV-58, respectively. In the system using pairs of oligonucleotides (27), both of the primers used for detection of HPV-18, HPV-52, and MM4 (Table 2, probes 18 [MY130 and WD74] 52 [MY81 and MY82] and MM4 [MY164 and MY165]) have mismatched nucleotide positions compared with some of the variant sequences observed in our present study. For the remaining HPV types, only one or none of the oligonucleotide probes within a pair has mismatches.

The previously reported novel HPV sequence IS039 (40) would not be detected by either MM4 (W13B) or HPV-51 probes, since it is variant from these two types by 9.2 and 16.4% in the MY09/11 region, respectively.

DISCUSSION

The aim of this study was to analyze HPV intratype variation of less prevalent genital HPV types in a worldwide collection of specimens. Sequence variation within the less prevalent HPV types is important for the development of diagnostics and vaccines, for epidemiological investigations, and for studies of phylogeny and taxonomy of HPVs. Previous studies of intratype HPV variation have generally been performed with regions of the genome with higher intratype variation, such as the URR or the E6 or the L2 ORFs (10, 26, 29, 30, 32, 42, 65, 66). Our choice of the MY09/11 L1 region was based on the targeting of this region for HPV diagnostic purposes and the range of intertype variation observed in this region that allows the distinction and assessment of known and novel HPV types (5). The polymorphism observed in the MY09/11 L1 coding region is not as great as that observed in the noncoding URR. Our earlier (66) and present investigations do, however, indicate that the variation observed in this region is sufficient for assessment of the major intratype HPV variants and groupings. Moreover, our previous studies (66) demonstrate that typical variant sequence patterns are linked in different regions of the HPV genome, which is evidence for the absence (or rarity) of recombination. This observed linkage makes it possible to use very short regions of an HPV genome to assess the classification of types, subtypes, or variants. Linkage between intratype variant patterns in different parts of the genome should not be taken as evidence for a causal connection, since mutations in different parts of the genome may be independent events. As a consequence of this, HPV variants defined by intratype sequence patterns in the MY09/11 L1 region may be further subclassified after recording of variation elsewhere in the particular genome.

In the MY09/11 L1 region, the maximal intratype pairwise evolutionary distances observed in our study were 2.9% (i.e., 12 nucleotide substitutions in 410 bp). This percentage was observed and then maintained after analysis of rather few specimens, indicating that the intratype variation for less prevalent HPV types mimics that observed previously for more prevalent HPV types such as HPV-16. Intratype variation previously reported for genital HPV types has varied from approximately 1.0 to 5.0% within ORFs depending on the ORF under consideration (21, 24, 30, 38, 42, 52, 66) and approximately 5% in the URR (26, 28, 38, 39, 66). Intratype variation did not appear to be random, and the relationship among variants within an HPV type was notable.

It is unlikely that the observed and confirmed HPV sequence variations presented are due to errors produced during the amplification process. In all cases, multiple clones were analyzed for each specimen, sometimes from independent PCRs. Our PCR protocol was not optimized for the detection of low-copy-number HPV genomes. Therefore, PCR errors occurring in an early amplification cycle would most probably remain outnumbered by the amplification of correct genomes. The frequency of unconfirmed changes in our complete data set was 9×10^{-4} (one presumed nucleotide error per 1,100 bp or approximately one error per 2.5 clones). This error frequency is well within the limits (2×10^{-4} to 30×10^{-4}) reported for standard *Thermus aquaticus* (*Taq* polymerase)-based PCR as reviewed by Eckert and Kunkel (18). In addition, most (60%) of the unconfirmed changes observed were transitions (T→C or A→G and G→A or C→T) which are common *Taq* polymerase errors.

The limited intratype genomic diversity observed in HPVs is surprising. Evolution is a continuous process, but common ancestors to separate types of HPVs have so far not been observed and must have been eradicated by processes like genetic drift or reduction/expansion events, such as evolutionary bottlenecks.

Recent studies have examined human mitochondrial DNA (mtDNA) in attempts to distinguish selection from past reduction/expansion events in the evolution of modern humans. mtDNA has a higher rate of mutation than most regions of the human genome (7), does not recombine, and is, for all practical purposes, maternally inherited (23). Conclusions from two studies with mtDNA suggest that modern humans arose in Africa about 100,000 to 200,000 years ago from interpretations of genealogical trees (8, 62). The fact that several of these trees have their deepest branches within the African population has been interpreted as support for an African origin of modern humans. This hypothesis has, however, been challenged by recent work which suggests that the observed pattern could be

an artifact from the order in which observations were entered into the computer programs used (22, 25, 34, 57, 58). Recently, a new statistical method (45) was used to analyze the mtDNA data previously reported (8, 62). These analyses have provided an alternative hypothesis of modern human origins. The hypothesis, called the Weak Garden of Eden, suggests that a small ancestral population of ~10,000 breeding individuals separated into several partially isolated groups approximately 100,000 years ago. These ancestral groups further underwent either simultaneous bottlenecks or simultaneous expansions about 30,000 years later (45).

The puzzle in the observed pattern of limited HPV diversity may find an explanation in an evolutionary bottleneck and subsequent expansion in the HPV population resulting from a hypothesized bottleneck and expansion event in the host population as described above. Although data described in our present investigation are limited for many of the HPV types analyzed because of small sample sizes, this and other investigations have presented similar values for the intratype pairwise distances for many types of HPVs. Notable is the relative lack of comparisons in the 3 to 10% range and the sharp drop-off around 2.5%. A bottleneck in the human population would have limited the number of HPV lineages that would survive, possibly causing massive extinction of lineages. It is possible that the relative lack of HPV polymorphism in the range between 2.5 and 10% reflects a severe pruning of the HPV evolutionary tree as a result of such a reduction event in the human population approximately 100,000 years ago. Moreover, the relative abundance of variants within the 0 to 2.5% range may reflect subsequent expansion events occurring in the host population about 30,000 years later. The reduction/expansion hypothesis would provide a connection between specific variants and the continents where they originated and expanded. Many of the HPV variants found in this and other studies are, however, not restricted to specific continents, but on closer observation they do predominate by continent and lineage. The observed deviations are most probably a result of more recent migrations of the host populations. A complete continental restriction would be observed only in nonmigrating, totally isolated cultures of modern humans.

Naturally, there are other explanations that cannot be ruled out by existing evolutionary HPV data. Effects such as genetic drift, for example, could provide an alternative explanation of the observed range of HPV polymorphism described here and by others. We believe that HPVs and possibly other types of viruses may be important materials for the study of modern human evolution. Larger HPV data sets from distinct, isolated human populations will be needed for further evaluation of the hypotheses presented.

In this study, the majority of HPV specimens were collected as part of the IBSCC, a worldwide study of cervical cancer (6). Therefore, our data are limited by a bias toward HPVs associated with invasive cervical cancer. It is likely, however, that our data reflect the HPV variation present in the normal population. Cancers are presumably "dead ends" for HPVs and most probably the secondary result of persistent disruption of cell cycle regulatory elements. An HPV infection resulting in a cancer often exhibits integration of the viral genome into the host chromosome such that most of the episomal virus forms are lost and few or no viral particles are shed from such lesions. This suggests that the viral variants found in cancers most probably represent those found in normal individuals and therefore reflect the actual range of HPV variation. Moreover, the range of intragene variation seen in women with normal cervices infected with HPV-16 (66) is similar to that observed in this study and does not exceed 2.9% in the MY09/11 region. This does not rule out the possibility that some HPV variants are more likely than others to be associated with cancers. The possibility that certain variants are disease associated has recently been suggested (19, 24, 63). A combination of the genetic/immunological host background and the particular HPV variant infecting the host may explain why certain HPV infections are associated with invasive cervical cancers.

Additional limitations of this study included the numbers of specimens under investigation and the geographic representation within individual HPV types and novel sequences. It was not possible to evaluate potential geographic distributions of particular HPV variants because of our small sample sizes, with the exception of HPV-18 and HPV-45. For these two HPV types, we clearly observed nucleotide marker substitutions in African specimens that were not present in specimens obtained from other continents. We observed only one exception to this in an HPV-18 specimen, IS227, from Cuba. The observation that certain variants are more prevalent in certain geographic regions is not surprising and may be explained by founder effects. It may also reflect a potential origin for certain HPV types in certain geographical regions. This hypothesis has previously been reported for HPV-18, for which an African origin has been suggested (39).

The two most widely used PCR-based systems for detection of HPVs use all or part of the MY09/11 L1 region in combination with oligonucleotide probe hybridization. Our present data indicate that neither of these systems may readily detect some of the HPV intratype variants observed in our study. Modification of some of the previously described oligonucleotide probes may improve the detection of variants and obviate the need to use multiple probes. Because of sample size limitations, it is likely that not all potential variants of specific HPV types were represented. The use of probes in pairs, or the use of alternate hybridization protocols (e.g., with lower stringency), may in these cases facilitate the detection of such variant HPV types.

A challenge will be to determine whether intratype HPV variants are relevant to HPV vaccine strategies. Such considerations might be facilitated by complete sequencing of variant capsid genes followed by the development of immunologic test systems. Single amino acid substitutions within the L1 and L2 capsid genes may be important in virus escape from neutralizing antibodies. In the absence of a virus neutralization assay, such determinations may only be suggested by characterization of HPV variants infecting unprotected vaccinees.

ACKNOWLEDGMENTS

We thank Aaron Halpern for discussions of HPV intratype diversity, including genetic drift and population bottleneck theories, as well as for performing the phylogenetic analyses; the members of the IBSCC Study Group for their support in providing specimens to this investigation; and E.-M. de Villiers, DKFZ, Heidelberg, Germany, for providing the HPV-18 reference clone. Additionally, we are grateful for the continuing support of Joe Icenogle and Bill Reeves, Centers for Disease Control and Prevention; Jay Stoerker, AT Biochem; Catherine Greer, Chiron Corp.; and Patti Gravitt and Tracy Zhang, Roche Molecular Systems. We are especially grateful to our students John (Bob) Herron, Vanessa Padilla, and April Martinez, University of New Mexico (UNM), and Angela Jansen, Johns Hopkins, for their efforts in helping to complete this study. This investigation utilized the UNM Molecular Analysis Shared Facility.

Funding for these studies was provided in part by the National Institutes of Health (grant R01-AI32917) to C.M.W. and was additionally supported by Research Training Fellowships awarded to A.-C.M.S. by the International Agency for Research on Cancer and to A.M.E. by the Swedish Institute. The IBSCC was funded by NCI contract MA-5623-41 to M.M.M., the European Community (contract

C11-0371-F [CD]), and the Cancer Research Campaign Endowment Fund.

REFERENCES

1. **Bauer, H., Y. Ting, C. E. Greer, J. C. Chambers, C. J. Tashiro, J. Chimera, A. Reingold, and M. M. Manos.** 1991. Genital human papillomavirus infection in female university students as determined by a PCR-based method. JAMA **265:**472–477.
2. **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.), Diagnostic molecular pathology: a practical approach. Oxford University Press, Oxford.
3. **Becker, T. M., C. M. Wheeler, N. S. McGough, C. A. Parmenter, S. W. Jordan, C. A. Stidley, R. S. McPherson, and M. H. Dorin.** 1994. Sexually transmitted diseases and other risk factors for cervical dysplasia among southwestern Hispanic and non-Hispanic white women. JAMA **271:**1181–1188.
4. **Berkhout, R. J. M., L. M. Tieben, H. L. Smits, J. N. Bouwes Bavinck, B. J. Vermeer, and J. ter Schegget.** 1995. Nested PCR approach for detection and typing of epidermodysplasia verruciformis-associated human papillomavirus types in cutaneous cancers from renal transplant recipients. J. Clin. Microbiol. **33:**690–695.
5. **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.
6. **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.
7. **Brown, W. M., M. George, and A. C. Wilson.** 1979. Rapid evolution of animal mitochondrial DNA. Proc. Natl. Acad. Sci. USA **76:**1967–1971.
8. **Cann, R., M. Stoneking, and A. C. Wilson.** 1987. Mitochondrial DNA and human evolution. Nature (London) **325:**31–36.
9. **Chan, S.-Y., H. Delius, A. L. Halpern, and H.-U. Bernard.** 1995. Analysis of genomic sequences of 95 papillomavirus types: uniting typing, phylogeny, and taxonomy. J. Virol. **69:**3074–3083.
10. **Chan, S.-Y., L. Ho, C.-K. Ong, V. Chow, B. Drescher, M. Dürst, J. ter Meulen, L. Villa, J. Luande, H. N. Mgaya, and H.-U. Bernard.** 1992. Molecular variants of human papillomavirus type 16 from four continents suggest ancient pandemic spread of the virus and its coevolution with humankind. J. Virol. **66:**2057–2066.
11. **Cole, S. T., and O. Danos.** 1987. Nucleotide sequence and comparative analysis of the human papillomavirus type 18 genome. J. Mol. Biol. **193:**599–608.
12. **Deau, M.-C., M. Favre, S. Jablonska, L.-A. Rueda, and G. Orth.** 1993. Genetic heterogeneity of oncogenic human papillomavirus type 5 (HPV5) and phylogeny of HPV5 variants associated with epidermodysplasia verruciformis. J. Clin. Microbiol. **31:**2918–2926.
13. **Deau, M.-C., M. Favre, and G. Orth.** 1991. Genetic heterogeneity among human papillomaviruses (HPV) associated with epidermodysplasia verruciformis: evidence for multiple allelic forms of HPV5 and HPV8 E6 genes. Virology **184:**492–503.
14. **Delius, H., and B. Hofmann.** 1994. Primer-directed sequencing of human papillomavirus types. Curr. Top. Microbiol. Immunol. **186:**13–31.
15. **de Villiers, E.-M.** 1989. Heterogeneity of the human papillomavirus group. J. Virol. **63:**4898–4903.
16. **de Villiers, E.-M.** 1994. Human pathogenic papillomavirus types: an update. Curr. Top. Microbiol. Immunol. **186:**1–12.
17. **Dürst, M., L. Gissman, H. Ikenberg, and H. zur Hausen.** 1983. 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.
18. **Eckert, K. A., and T. A. Kunkel.** 1991. DNA polymerase fidelity and the polymerase chain reaction. PCR Methods Applic. **1:**17–24.
19. **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 HPV 16 oncogene variant with HLA-B7 has implications for vaccine design in cervical cancer. Nat. Med. **1:**464–470.
20. **Eluf-Neto, J., M. Booth, N. Muñoz, F. X. Bosch, C. J. L. M. Meijer, and J. M. M. Walboomers.** 1994. Human papillomavirus and invasive cervical cancer in Brazil. Br. J. Cancer. **69:**114–119.
21. **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.
22. **Gibbons, A.** 1992. Mitochondrial Eve: wounded, but not dead. Science **257:**873–875.
23. **Giles, R. E., H. Blanc, H. M. Cann, and D. C. Wallace.** 1980. Maternal inheritance of human mitochondrial DNA. Proc. Natl. Acad. Sci. USA **77:**6715–6719.
24. **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.
25. **Hedges, S. B., S. Kumar, K. Tamura, and M. Stoneking.** 1992. Human origins and analysis of mitochondrial DNA sequences. Science **255:**737–739.
26. **Heinzel, P. A., S.-Y. Chan, L. Ho, M. O'Connor, P. Balaram, M. S. Campo, K. Fujinaga, N. Kiviat, J. Kuypers, H. Pfister, B. M. Steinberg, S.-K. Tay, L. L. Villa, and H.-U. Bernard.** 1995. Variation of human papillomavirus type 6 (HPV-6) and HPV-11 genomes sampled throughout the world. J. Clin. Microbiol. **33:**1746–1754.
27. **Hildesheim, A., M. H. Schiffman, P. E. Gravitt, A. G. Glass, C. E. Greer, T. Zhang, D. R. Scott, B. B. Rush, P. Lawler, M. E. Sherman, R. J. Kurman, and M. M. Manos.** 1994. Persistence of type-specific human papillomavirus infection among cytologically normal women. J. Infect. Dis. **169:**235–240.
28. **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, B. Norrild, M. R. Pillai, J. Stoerker, K. Syrjaenen, S. Syrjaenen, 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.
29. **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.
30. **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.
31. **Jacobs, M. V., A. M. de Roda Husman, A. J. C. van den Brule, P. J. F. Snijders, C. J. L. M. Meijer, and J. M. M. Walboomers.** 1995. Group-specific differentiation between high- and low-risk human papillomavirus genotypes by using the general primer-mediated polymerase chain reaction and two cocktails of oligonucleotide probes. J. Clin. Microbiol. **33:**901–905.
32. **Kitasato, H., H. Delius, H. zur Hausen, K. Sorger, F. Rösl, and E.-M. de Villiers.** 1994. Sequence rearrangements in the upstream regulatory region of human papillomavirus type 6: are these involved in malignant transition? J. Gen. Virol. **75:**1157–1162.
33. **Lowy, D. R., R. Kirnbauer, and J. T. Schiller.** 1994. Genital human papillomavirus infection. Proc. Natl. Acad. Sci. USA **91:**2436–2440.
34. **Maddison, D. R., M. Ruvolo, and D. L. Swofford.** 1992. Geographic origins of human mitochondrial DNA: phylogenetic evidence from control region sequences. Syst. Biol. **41:**111–124.
35. **Manos, M. M., J. Waldman, T. Y. Zhang, C. E. Greer, G. Eichinger, M. H. Schiffman, and C. M. Wheeler.** 1994. Epidemiology and partial nucleotide sequence of four novel genital human papillomaviruses. J. Infect. Dis. **170:**1096–1099.
36. **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.
37. **Myers, G., H. U. Bernard, H. Delius, M. Favre, J. Icenogle, M. van Ranst, and C. Wheeler (ed.).** 1994. Human papillomaviruses. A compilation and analysis of nucleic and amino acid sequences. Los Alamos National Laboratory, Los Alamos, N.Mex.
38. **Ong, C. K.** 1994. Molecular perspectives of the epidemiology, evolution and taxonomy of human papillomaviruses. Ph.D. thesis. Institute of Molecular and Cell Biology, National University of Singapore, Singapore.
39. **Ong, C.-K., S.-Y. Chan, M. S. Campo, K. Fujinanga, P. Mavromara, H. Pfister, S.-K. Tay, J. Ter Meulen, L. L. Villa, and H.-U. Bernard.** 1993. Evolution of human papillomavirus type 18: an ancient phylogenetic root in Africa and intratype diversity reflect coevolution with human ethnic groups. J. Virol. **67:**6424–6431.
40. **Peyton, C. L., A. M. Jansen, C. M. Wheeler, A.-C. Stewart, J. Peto, F. X. Bosch, N. Muñoz, A. R. Teyssie, M. Torroella, H. R. Wabinga, Sarjadi, C. Ngelangel, and M. M. Manos.** 1994. A novel human papillomavirus sequence from an international cervical cancer study. J. Infect. Dis. **170:**1093–1095.
41. **Peyton, C. L., and C. M. Wheeler.** 1994. Identification of five novel human papillomaviruses in the New Mexico triethnic population. J. Infect. Dis. **170:**1089–1092.
42. **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.
43. **Reuter, S., H. Delius, T. Kahn, B. Hofmann, H. zur Hausen, and E. Schwarz.** 1991. Characterization of a novel papillomavirus DNA in the cervical carcinoma cell line ME180. J. Virol. **65:**5564–5568.
44. **Roda Husman, A.-M., J. M. M. Walboomers, C. J. L. M. Meijer, E. K. J. Risse, M. E. I. Schipper, T. M. Helmerhorst, O. P. Bleker, H. Delius, A. J. C. van den Brule, and P. J. F. Snijders.** 1994. Analysis of cytomorphologically

abnormal cervical scrapes for the presence of 27 mucosotropic human papillomavirus genotypes, using polymerase chain reaction. Int. J. Cancer **56:** 802–806.
45. **Rogers, A. R., and H. Harpending.** 1992. Population growth makes waves in the distribution of pairwise genetic differences. Mol. Biol. Evol. **9:**552–569.
46. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
47. **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. Wacholder, 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.
48. **Schiffman, M. H., H. M. Bauer, A. T. Lorincz, M. M. Manos, J. C. Byrne, A. G. Glass, D. M. Cadell, and P. M. Howley.** 1991. Comparison of Southern blot hybridization and polymerase chain reaction methods for the detection of human papillomavirus DNA. J. Clin. Microbiol. **29:**573–577.
49. **Schiffman, M. H., N. B. Kiviat, R. D. Burk, K. V. Shah, R. W. Daniel, R. Lewis, J. Kuypers, M. M. Manos, D. R. Scott, M. E. Sherman, R. J. Kurman, M. H. Stoler, A. G. Glass, B. B. Rush, I. Mielzynska, and A. T. Lorincz.** 1995. Accuracy and interlaboratory reliability of human papillomavirus DNA testing by hybrid capture. J. Clin. Microbiol. **33:**545–550.
50. **Shadan, F. F., and L. P. Villarreal.** 1993. Coevolution of persistently infecting small DNA viruses and their hosts linked to host-interactive regulatory domains. Proc. Natl. Acad. Sci. USA **90:**4117–4121.
51. **Shamanin, V., M. Glover, C. Rausch, C. Proby, I. M. Leigh, H. zur Hausen, and E.-M. de Villiers.** 1994. Specific types of human papillomavirus found in benign proliferations and carcinomas of the skin in immunosuppressed patients. Cancer Res. **54:**4610–4613.
52. **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.
53. **Snijders, P. J. F., A. J. C. van den Brule, H. F. J. Schrijnemakers, G. Snow, C. J. L. M. Meijer, and J. M. M. Walboomers.** 1990. The use of general primers in the polymerase chain reaction permits the detection of a broad spectrum of human papillomavirus types. J. Gen. Virol. **71:**173–181.
54. **Stewart, A.-C. M., P. E. Gravitt, S. Cheng, and C. M. Wheeler.** 1995. Generation of entire human papillomavirus genomes by long PCR: frequency of errors produced during amplification. Genome Res. **5:**79–88.
55. **Sundberg, J. P.** 1987. Papillomavirus infections in animals, p. 40–103. *In* K. Syrjaenen, L. Gissmann, and L. G. Koss (ed.), Papillomaviruses and human disease. Springer-Verlag KG, Berlin.
56. **Swofford, D. L.** 1993. Phylogenetic analysis using parsimony: PAUP, version 3.1. Illinois Natural History Survey, Champaign.
57. **Templeton, A.** 1993. The "Eve hypothesis": a genetic critique and reanalysis. Am. Anthropol. **95:**51–72.
58. **Templeton, A.** 1992. Human origins and analysis of mitochondrial DNA sequences. Science **255:**737.
59. **van den Brule, A. J. C., C. J. L. M. Meijer, V. Bakels, P. Kenemans, and J. M. M. Walboomers.** 1990. Rapid detection of human papillomavirus in cervical scrapes by combined general primer-mediated and type-specific polymerase chain reaction. J. Clin. Microbiol. **28:**2739–2743.
60. **van den Brule, A. J. C., P. J. F. Snijders, P. M. C. Raaphorst, H. F. J. Schrijnemakers, H. Delius, L. Gissman, C. J. L. M. Meijer, and J. M. M. Walboomers.** 1992. General primer polymerase chain reaction in combination with sequence analysis for identification of potentially novel human papillomavirus genotypes in cervical lesions. J. Clin. Microbiol. **30:**1716–1721.
61. **Van Ranst, M. A., R. Tachezy, H. Delius, and R. D. Burk.** 1993. Taxonomy of the human papillomaviruses. Papillomavirus Rep. **4:**61–65.
62. **Vigilant, L., M. Stoneking, H. Harpending, K. Hawkes, and A. C. Wilson.** 1991. African populations and the evolution of human mitochondrial DNA. Science **253:**1503–1507.
63. **Wheeler, C. M., A. Hildesheim, and M. H. Schiffman.** HPV 16 E6 nucleotide sequence variants as cervical disease correlates. Submitted for publication.
64. **Wheeler, C. M., C. A. Parmenter, W. C. Hunt, T. M. Becker, C. E. Greer, A. Hildesheim, and M. M. Manos.** 1993. Determinants of genital human papillomavirus infection among cytologically normal women attending the University of New Mexico Student Health Center. Sex. Transm. Dis. **20:**286–289.
65. **Xi, L. F., G. W. Demers, N. B. Kiviat, J. Kuypers, A. M. Beckmann, and D. A. Galloway.** 1993. Sequence variation in the noncoding region of human papillomavirus type 16 detected by single-strand conformation polymorphism analysis. J. Infect. Dis. **168:**610–617.
66. **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 analysis of the E6, L2, and L1 coding segments. J. Virol. **69:**7743–7753.
67. **zur Hausen, H.** 1991. Viruses in human cancers. Science **254:**1167–1173.