Papillomavirus subtypes are natural and old taxa

Journal of Virology

Papillomavirus Subtypes Are Natural and Old Taxa: Phylogeny of Human Papillomavirus Types 44 and 55 and 68a and -b

Itzel E. Calleja-Macias, Mina Kalantari, Bruce Allan, Anna-Lise Williamson, Lap-Ping Chung, Robert J. Collins, Rosemary E. Zuna, S. Terence Dunn, Rocio Ortiz-Lopez, Hugo A. Barrera-Saldaña, Heather A. Cubie, Kate Cuschieri, Luisa L. Villa and Hans-Ulrich Bernard

J. Virol. **2005, 79(10):6565. DOI: 10.1128/JVI.79.10.6565-6569.2005.**

JOURNAL OF VIROLOGY, May 2005, p. 6565–6569 Vol. 79, No. 10
0022-538X/05/$08.00+0 doi:10.1128/JVI.79.10.6565–6569.2005

Papillomavirus Subtypes Are Natural and Old Taxa: Phylogeny of Human Papillomavirus Types 44 and 55 and 68a and -b

Itzel E. Calleja-Macias,[1] Mina Kalantari,[1] Bruce Allan,[2] Anna-Lise Williamson,[2,3] Lap-Ping Chung,[4] Robert J. Collins,[4] Rosemary E. Zuna,[5] S. Terence Dunn,[5] Rocio Ortiz-Lopez,[6] Hugo A. Barrera-Saldaña,[6] Heather A. Cubie,[7] Kate Cuschieri,[7] Luisa L. Villa,[8] and Hans-Ulrich Bernard[1]*

Department of Molecular Biology and Biochemistry, University of California Irvine, Irvine, California 92697[1]; Institute of Infectious Disease and Molecular Medicine, Faculty of Health Sciences,[2] and National Health Laboratory Service,[3] University of Cape Town, Cape Town, South Africa; Queen Mary Hospital and the University of Hong Kong, Hong Kong[4]; Department of Pathology, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma 73104[5]; Departamento de Bioquimica, Facultad de Medicina, Universidad Autonoma de Nueva Leon, Monterrey, Mexico[6]; Royal Infirmary of Edinburgh, Edinburgh, Scotland[7]; and Ludwig Institute for Cancer Research, Sao Paulo, Brazil[8]

Received 16 November 2004/Accepted 18 January 2005

A human papillomavirus (HPV) type is defined as an HPV isolate whose L1 gene sequence is at least 10% different from that of any other type, while a subtype is 2 to 10% different from any HPV type. In order to analyze the phylogeny behind the subtype definition, we compared 49 isolates of HPV type 44 (HPV-44) and its subtype HPV-55, previously misclassified as a separate type, and 41 isolates of the subtype pair HPV-68a and -b, sampled from cohorts in four continents. The subtypes of each pair are separated by deep dichotomic branching, and three of the four subtypes have evolved large phylogenetic clusters of genomic variants forming a "star" phylogeny, with some branches specific for ethnically defined cohorts. We conclude that subtypes of HPV types are natural and old taxa, equivalent to types, which either diverged more recently than types or evolved more slowly.

Papillomaviruses (PVs) form a separate family of viruses, which can be hierarchically subgrouped into PV genera, species, and types, the last term traditionally being used to describe novel PV isolates. A PV type is defined as a complete viral genome with a sequence organization typical of PVs and an L1 gene sequence that is at least 10% dissimilar from that of any other PV type (4). Each human PV (HPV) type can be reisolated in the form of genomic variants, as found during systematic worldwide searches and comparisons (2, 5, 8, 10). These variants differ from one another in nucleotide exchanges by maximally 2%, and they form intra-HPV-type phylogenetic trees without discontinuities. Evolution of HPV types apparently predated the evolution of *Homo sapiens*, while variation of each HPV type was linked to the evolution of humans (2), occurring over tens to hundreds of thousands of years.

More than 20 years ago the term "subtype" was introduced for HPV isolates whose genomes hybridized with a known HPV type but differed by restriction analysis. Recently, subtypes became redefined as HPV genomes with L1 genes differing by 2 to 10% from that of the original HPV type, placing the taxon "subtype" in between types and variants. While all quantitative definitions in PV taxonomy were arbitrary, they apparently describe natural taxonomic discontinuities, as there are hundreds of different HPV types (1, 4) and no genomes intermediate to any two HPV types. It is not clear how the continuous process of evolution created these discontinuities. Variants, on the other hand, form a continuous spectrum; i.e., they differ from one another by one or a few individual nucleotide exchanges. As a result, they form a star phylogeny rather than deep dichotomic branches. There are numerous variants of each type, while subtypes are extraordinarily rare. The reason for this is not understood. Only four HPV genomes presently fulfill the definition of a subtype: HPV type 55 (HPV-55; originally a separate HPV type, reclassified as a subtype of HPV-44 with 5% genome diversity), HPV-46 and HPV-64 (reclassified as subtypes of HPV-20 and HPV-34, respectively) (5), and the two subtypes of HPV-68, HPV-68a and -b (6, 9).

In order to explore the phylogeny of HPV types that gave rise to subtypes, we examined the diversity of isolates of HPV-44 and its subtype HPV-55 and of HPV-68a and -b by amplifying and sequencing a segment of the long control region (LCR) of these PVs in samples from six geographically and ethnically remote cohorts (Brazil, Hong Kong, Mexico, Scotland, South Africa, and the United States). LCR sequences are typically more diverse than genes and therefore phylogenetically more informative (2, 5, 8, 10). The samples that we studied had been detected during ongoing clinical research of genital smears unrelated to the objectives of our project.

HPV-44 and -55 are "low-risk" HPV types related to HPV-6 and -11 and therefore normally not found in carcinomas. HPV-44 was originally isolated from a vulvar condyloma (7), and HPV-55, when isolated from a Bowenoid papulosis (reviewed in reference 3) was originally described as a separate HPV type (3) and later reclassified as a close relative of

* Corresponding author. Mailing address: Department of Molecular Biology and Biochemistry, University of California, Irvine, CA 92697-3900. Phone: (949) 824-5162. Fax: (949) 824-8551. E-mail: hbernard@uci.edu.

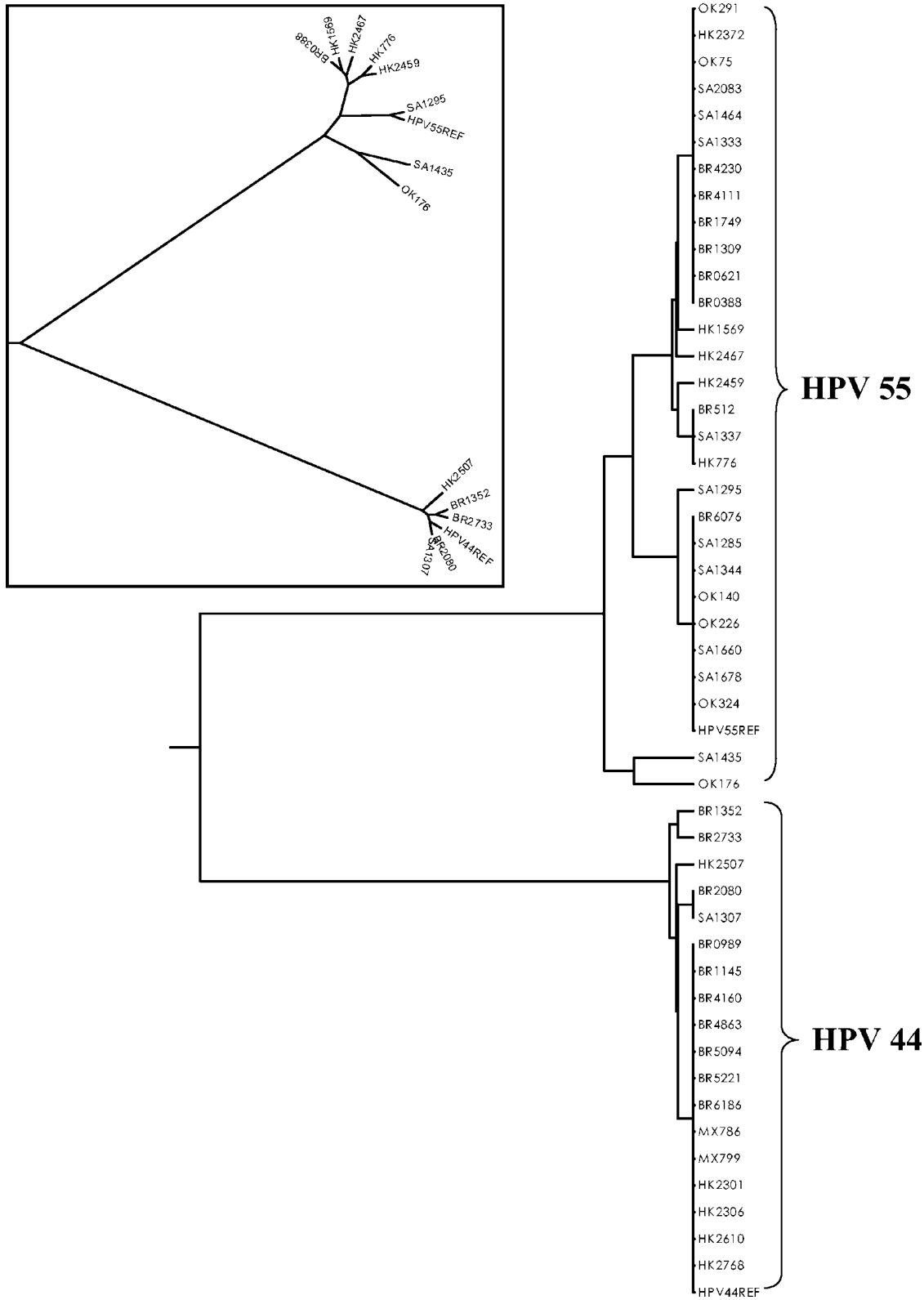

FIG. 1. (A) Phylogenetic tree of 49 isolates of HPV-44 and its subtype HPV-55, which has been originally misclassified as an independent type and therefore numbered differently. The samples came from five cohorts of patients living in Sao Paulo, Brazil (BR), Hong Kong (HK), Monterrey, Mexico (MX), Cape Town, South Africa (SA), and Oklahoma City, Oklahoma (OK). The numerical codes following these abbreviations correlate to patients, whose identities remained unknown to the principal investigators of this study. The major phylogenetic tree was generated by the UPGMA (unweighted pair group method with arithmetic average) algorithm. The small inset was calculated by the neighbor-joining method and included in this figure to increase the visual impact of a star phylogeny as opposed to evolution by deep dichotomic branching.

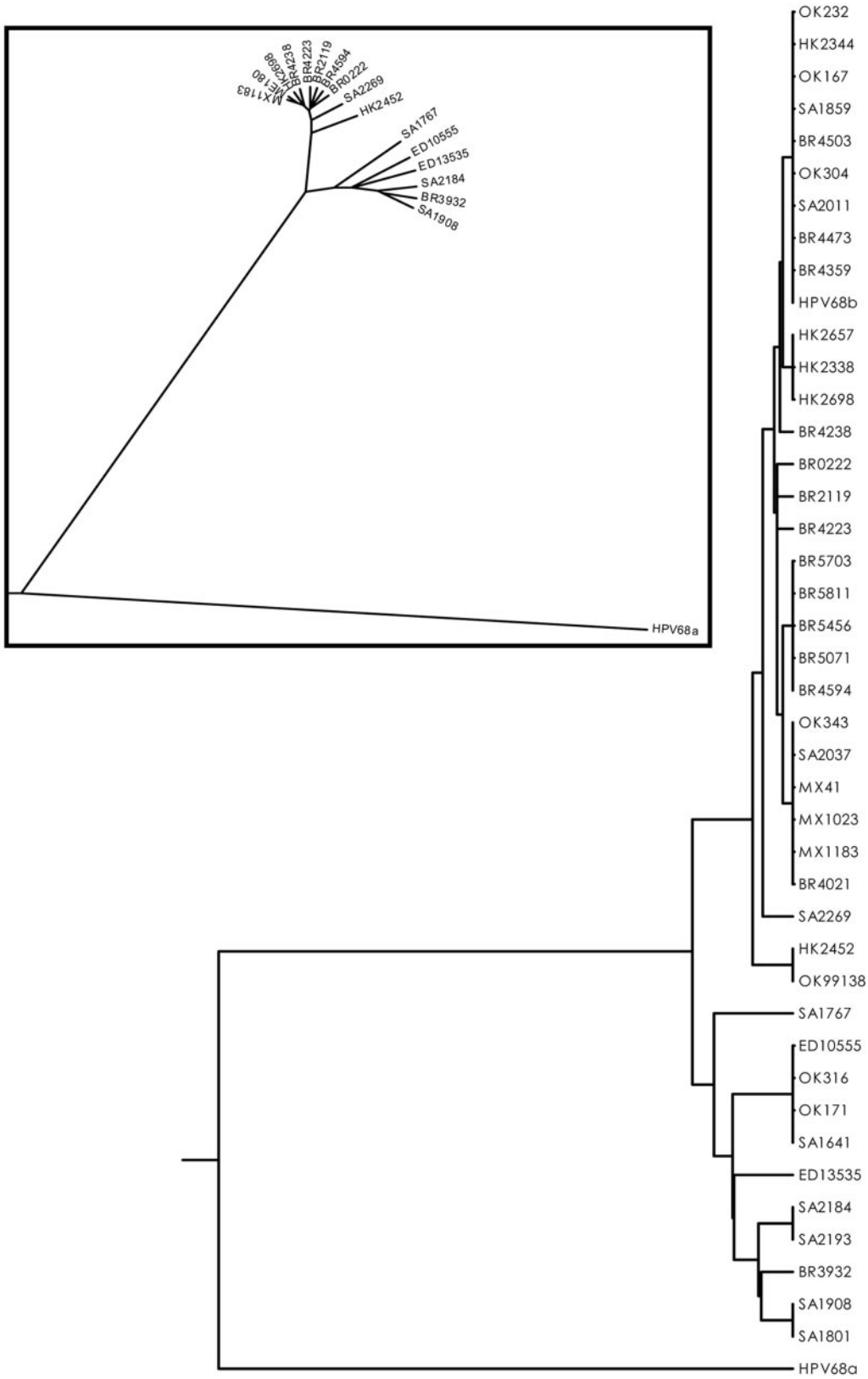

FIG. 2. Phylogenetic tree of 41 samples of HPV-68a and -b. Abbreviations and procedures were as described in the legend to Fig. 1. In addition to the samples from the above-named regions, there are samples from Edinburgh, Scotland (ED).

HPV-44 (4), by definition a subtype of HPV-44. Figure 1 shows the phylogenetic tree of all samples originally diagnosed as HPV-44 or HPV-55. The genomic segments were PCR amplified with primers that recognized the sequences of both viruses (HPV-44SF, 5′-ACCCCATGAGTAAGTGTTGTAGTG-3′; HPV-44SR, 5′-AATTCGGTTCCTCCTCTTTTT-3′; HPV-55SF, 5′-ACCCCATGAGTAAGTGTTGTAGTAT-3′; and HPV-55SR, 5′-AATTCGGTCCCTCCTCTTTTCT-3′) and generated 524- and 511-bp segments, respectively, between the genomic positions 7354 and 44 (HPV-44) and 7353 and 41 (HPV-55), respectively. The tree contains the two HPV-44 and HPV-55 reference sequences as well as 18 and 29, respectively, new isolates of each type, together representing a total of 15 variant genomes: six in an HPV-44 cluster and nine in an HPV-55 cluster. The reference genomes of HPV-44 and -55 differed in the analyzed LCR segment by 42 nucleotide exchanges and eight insertions and deletions, respectively. As expected, these numbers exceed the criterion of 2 to 10% dissimilarity, which was defined for the conserved L1 gene, while the LCR is hyperdiverse. Among the HPV-44 variants, we observed mutations in five positions, and maximal distances of 4 nucleotides. Among the HPV-55 variants, there were 11 nucleotide exchanges, one insertion (33 bp), and maximal distances of six mutations. These clusters form a tree with two deep dichotomic branches representing the two HPV subtypes, while at the tips of these branches the variants form star-like clusters of minor branches (inset of Fig. 1). Several variants from Hong Kong and South Africa were not represented in other cohorts, which may indicate the restriction of some variants to specific ethnic groups. Unfortunately, since most of the cohorts that we analyzed were multiethnic, a detailed linkage between the evolution of HPV-44 and -55 and humans is difficult to determine, unlike that of other HPV types (2, 5, 8, 10).

HPV-68 has been originally observed as a partial genome integrated into a chromosome of the carcinoma cell line ME180 (9). It was later isolated as a full-size genome from a low-grade cervical intraepithelial neoplasia (6). In view of a 7% divergence of the sequence of the two isolates, the latter was termed HPV-68a and the former HPV-68b. In our research, we first completed the published partial sequence of the LCR of the original clone of HPV-68a (kindly supplied by G. Orth and M. Favre) by amplifying a segment of this clone, whose sequence has not been published. This was accomplished with primers modeled according to two sequences that were conserved between the HPV-68a and -b genomes, namely, HPV-68F (5′-CAGGCAGGTGTTCGCAGAC-3′), corresponding to positions 1607 to 1625 of the published L1 sequence (GenBank accession number X67161), and HPV-68R (5′-GCAAT TTGTATGGCCGTTCCTCA-3′), corresponding to positions 300 to 322 of the published E6 sequence (GenBank accession number X67160). Within the resulting 963-bp segment, we amplified and phylogenetically evaluated the 547-bp segment based on the HPV-68a genes with the primers 68-2AF (5′-G ACTGCAACATTTCCTACAT-3′) and 68-2ABF (5′-CGTTT TCGGTCACTCCCTTTAT-3′). The corresponding segments of HPV-68a and -b differed in 52 bp from one another. We then amplified and sequenced the same LCR segment of 41 clinical samples from the same five cohorts as described above. Surprisingly, HPV-68a was never found again, as all 41 clinical samples contained 1 of 15 different HPV-68b variants. Just like HPV-44 and -55, HPV-68a and -b formed two deep dichotomic branches, and all HPV-68b variants formed a tight cluster at one end of these two branches (inset in Fig. 2). Among the HPV-68b variants, we observed mutations in 29 positions and maximal distances of 12 nucleotides. Correlation with geography was better than with HPV-44 and -55, as three of five samples from Hong Kong formed a minor branch without genomes from other localities, and four out of eight samples from South Africa formed a tight cluster that included one Brazilian clone.

From our study of the genomic diversity of two HPV types from which subtypes have been identified, we conclude that subtypes originate through phylogenetic processes similar to those of HPV types, namely, through a dichotomic split in some ancestor population rather than representing fortuitous isolates of a large diverse cluster of viral genomes with extensive continuities between nucleotide sequences. Three of the four subtypes that we investigated had given rise to clusters of genomic variants with indications of the geographic and ethnic group specificity of some variants. While subtypes are quantitatively defined by having less diversity than HPV types, these two properties—namely, dichotomic division and separately evolved star-like clusters of variants—make them taxa resembling HPV types. Our research aimed to understand the phylogenetic pathway of the origins of subtypes, but we abstain from a proposal to change the "type" and "subtype" definition, which would be confusing without serving a scientific or clinical need. The evolutionary origins of HPV subtypes, just like those of HPV types, are not well understood but may lie in genetic bottlenecks that separated populations of previously continuous viral nucleotide sequences, possibly created by the extinction of prehuman host species. This view is supported by the close subtype-like relationship of the chimpanzee and the pigmy chimpanzee PVs (4, 11), which may have diverged at the time that their closely related ape hosts diverged. These two ape PVs are closely related with HPV-6, -11, -13, -44, -55, and 74, with whom they form HPV species 10 (4). This supports the views that PV type diversification even between closely related types (and the subtype pair -44 and -55) took place over a time frame of up to a few million years (the age of the common ancestor of humans and chimpanzees) and that PV types originated both by splits between host species and by divergent evolution of HPVs within their human host.

Nucleotide sequence accession numbers. The sequences of all variants are deposited with the following GenBank accession numbers: for HPV-44, AY829114 to AY829131; for HPV-55, AY829174 to AY829202; and for HPV-68, AY829132 to AY829173).

This research was supported by NIH grant ROI CA-91964 and by funds from the Chao Family Comprehensive Cancer Center of the University of California Irvine to H.-U.B., by a postdoctoral fellowship of the Cancer Center of the University of California—Irvine to M.K., and by a collaborative research grant from UC MEXUS-CONACYT to I.E.C.-M., H.A.B.-S., and H.-U.B.

REFERENCES

1. **Antonsson, A., O. Forslund, H. Ekberg, G. Sterner, and B. G. Hansson.** 2000. The ubiquity and impressive genomic diversity of human skin papillomaviruses suggest a commensalic nature of these viruses. J. Virol. **74:**11636–11641.
2. **Bernard, H. U.** 1994. Coevolution of papillomaviruses and human populations. Trends Microbiol. **2:**140–143.

3. **de Villiers, E. M.** 1989. Heterogeneity of the human papillomavirus group. J. Virol. **63:**4898–4903.
4. **de Villiers, E. M., C. Fauquet, T. R. Broker, H. U. Bernard, and H. zur Hausen.** 2004. Classification of papillomaviruses. Virology **324:**17–27.
5. **Ho, L., S. Y. Chan, R. D. Burk, B. C. Das, K. Fujinaga, J. P. Icenogle, T. Kahn, N. Kiviat, W. Lancaster, P. Mavromara, 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 movement of ancient human populations. J. Virol. **67:**6413–6414.
6. **Longuet, M., S. Beaudenon, and G. Orth.** 1996. Two novel genital human papillomavirus (HPV) types, HPV68 and HPV70, related to the potentially oncogenic HPV39. J. Clin. Microbiol. **34:**738–755.
7. **Lörincz, A. T., A. P. Quinn, M. D. Goldsborough, B. J. Schmidt, and G. F. Temple.** 1989. Cloning and partial DNA sequencing of two new human papillomavirus types associated with condylomas and low-grade cervical neoplasia. J. Virol. **63:**2829–2834.
8. **Ong, C. K., S. Y. Chan, M. S. Campo, K. Fujinaga, P. Mavromara, V. Labropoulou, 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.
9. **Reuter, S., H. Delius, T. Kahn, B. Hofmann, H. zur Hausen, and E. Schwarz.** 1991. Characterization of a novel human papillomavirus DNA in the cervical carcinoma cell line ME180. J. Virol. **65:**5564–5568.
10. **Stewart, A. C., A. M. Eriksson, M. M. Manos, N. Munoz, 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.
11. **Van Ranst, M., A. Fuse, P. Fiten, E. Beuken, H. Pfister, R. D. Burk, and G. Opdenakker.** 1992. Human papillomavirus type 13 and pigmy chimpanzee papillomavirus type 1: comparison of the genome organization. Virology **190:**587–596.