Multiple Evolutionary Mechanisms Drive Papillomavirus Diversification

Marc Gottschling, Alexandros Stamatakis,† Ingo Nindl,* Eggert Stockfleth,* Ángel Alonso,‡ and Ignacio G. Bravo‡*

*Skin Cancer Center Charité, University Hospital of Berlin, Berlin, Germany; †École Polytechnique Fédérale de Lausanne, School of Computer & Communication Sciences, Laboratory for Computational Biology and Bioinformatics, Lausanne, Switzerland; and ‡Deutsches Krebsforschungszentrum/German Cancer Research Centre, Infection and Cancer, Heidelberg, Germany

The circular, double-stranded 8-kb DNA genome of papillomaviruses (PVes) consists mainly of 4 large genes, E1, E2, L2, and L1. Approximately 150 papillomavirus genomes have been sequenced to date. We analyzed a representative sample of 53 PVes genomes using maximum likelihood, Bayesian inference, maximum parsimony, and distance-based methods both on nucleotide (nt) and on amino acid (aa) alignments. When the 4 genes were analyzed separately, aa-inferred phylogenies contradicted each other less than nt-inferred trees (judged by partition homogeneity tests). In particular, gene combinations including the L2 gene generated significant incongruence ($P < 0.001$). Combined analyses of the remaining genes E1–E2–L1 produced a well-supported phylogeny including supertaxon $\beta + \gamma + \pi + \xi$-PVes (infecting Artiodactyla, Carnivora, Primates, and Rodentia) and supertaxon $\kappa + \lambda + \mu + \nu + \sigma$-PVes (infecting Carnivora, Lagomorpha, Primates, and Rodentia). Based on the tree topology, host-linked evolution appears plausible at shallow, rather than deeper, taxonomic levels. Diversification within PVes may also involve adaptive radiation establishing different niches (within a single-host species) and recombination events (within single-host cells). Heterogeneous groups of closely related PVes infecting, for example, humans and domestic animals such as hamster, dog, and cattle suggest multiple infections across species borders. Additional evolutionary phenomena such as strong codon usage preferences, and computational biases including reconstruction artifacts and insufficient taxon sampling, may contribute to the incomplete resolution of deep phylogenetic nodes. The molecular data globally supports a complex evolutionary scenario for PVes, which is driven by multiple mechanisms but not exclusively by coevolution with corresponding hosts.

Introduction

Papillomaviruses (PVes) infect stratified squamous epithelia of warm-blooded animals. Targets of the infection are undifferentiated keratinocytes in the basal cell layer. The progression of the virus infection depends on keratinocyte differentiation (Bedell et al. 1991; Doorbar 2005; Egawa 2005). A major interest for papillomavirus (PV) research arises from the causal association of individual types with cervical cancer and their potential for malignant transformation in mucosal tissue. Moreover, some PVes are associated with benign cutaneous lesions and probably with nonmelanoma skin cancer (zur Hausen 2000; Pfister 2003; Nindl et al. forthcoming).

In the past years, the available number of complete PV genome sequences has increased substantially and comprises nearly 150 GenBank entries (November 2006). The PV genome is a single molecule of double-stranded DNA and comprises approximately 8,000 bp. Eight well-defined open reading frames (ORFs) are encoded, which are all transcribed from the same DNA strand with the same orientation. The translated proteins are classified as "early" (E) and "late" (L) based on their temporal expression. They include 3 regulatory genes involved in transcription and replication (E1, E2, and E4), 3 oncogenes (E5, E6, and E7), and 2 genes coding for self-assembling proteins that give rise to the viral capsid (L1 and L2; Münger and Howley 2002). The complete L1 gene, or fragments of it, is commonly used for detecting PV infections and for typing PVes. For this reason, PV systematics have traditionally been inferred from the L1 gene, defining clear-cut nucleotide (nt) identity thresholds for the delimitation of higher taxonomic units such as "species" and "genera" (de Villiers et al. 2004; Bernard 2005).

PVes have been isolated from birds, marsupials, and placental mammals and are generally considered to be highly specific for their hosts. However, bovine PVes are able to cause nonproductive infections in horses and other only distantly related mammals (Thomas et al. 1964; Lancaster et al. 1977; Pfister et al. 1981; Trenfield et al. 1985; Otten et al. 1993; Chambers et al. 2003). Many viral taxa such as *Alphapapillomavirus* (α-PVes), *Deltapapillomavirus* (δ-PVes), and *Lambdapapillomavirus* (λ-PVes) roughly correspond to their mammalian host taxa, namely Primates, Artiodactyla, and Carnivora (Bernard et al. 1994; Myers et al. 1994; Chan et al. 1995; Farmer et al. 1995; de Villiers et al. 2004; García-Vallvé et al. 2005). Furthermore, phylogenetic clusters of PV variants are congruent with the geographic origin, at least in some human PV (HPV) types such as HPV-16 and HPV-18 (Chan et al. 1992; Ho et al. 1993; Ong et al. 1993; Yamada et al. 1997; Arias-Pulido et al. 2005; Prado et al. 2005). This has led to the general assumption that "host-linked evolution" (Chan et al. 1995, 1997) is the driving force for the diversification of PVes (Halpern 2000; Bernard et al. 2006).

However, the evolutionary mechanisms of PVes are more complex. For example, infections across species borders termed zoonoses (WHO Expert Committee 1982) may have contributed to the evolution of PVes (Myers et al. 1996; Rector, Van Doorslaer, et al. 2005; Gottschling et al. 2007). In addition, phylogenetic inconsistencies between early and late genes have been identified for some groups in the α-PVes (García-Vallvé et al. 2005; Narechania et al. 2005). This group includes cervical cancer-associated human PVes (HPVes) and accounts for more

Key words: adaptive radiation, coevolution, high performance computing, host, interspecies transmission, recombination, virology, zoonosis.

E-mail: i.bravo@dkfz.de.

Mol. Biol. Evol. 24(5):1242–1258. 2007
doi:10.1093/molbev/msm039
Advance Access publication March 6, 2007

than half of the complete PV genomes in GenBank. Evolutionary incongruence might arise from singular events in the past such as recombination, the establishment of new ecological niches, and/or asymmetric genome convergences driven by intense selection (Narechania et al. 2005; Varsani et al. 2006). However, well-supported contradicting tree topologies between early and late genes have not been found for *Betapapillomaviruses* (β-PVes) (Gottschling et al. 2007), which represent another important and diverse PV clade. This may indicate that concerted evolution of early and late genes is the rule in PVes.

Knowledge about viral evolution is still relatively poor compared with living organisms. However, a broad range of bioinformatics tools has been applied to analyze the complete PV genome (or at least properly alignable regions of it) during the past 2 years. The computation of confidence values for internal nodes allows for well-substantiated phylogenetic conclusions (Chen et al. 2005; Rector, Van Doorslaer, et al. 2005; Schiffman et al. 2005). Appropriate outgroup choice enables the evaluation of evolutionary polarity in PVes (García-Vallvé et al. 2005; Narechania et al. 2005; Gottschling et al. 2007). Nonetheless, a comprehensive scenario of evolution and phylogenetic relationships within PVes has not yet been developed, especially with respect to the basal nodes of the tree. The usage of high performance computing techniques and platforms in combination with advanced maximum likelihood (ML) search algorithms such as RAxML (Stamatakis et al. 2005; Stamatakis 2006b) enables thorough ML-based phylogenetic analyses including a sufficiently large amount of 1,000 bootstrap replicates.

In this study, we aim to identify those PV sequences that perturb the reconstruction of a concerted phylogeny and to choose the optimal set of suitable genes for phylogenetic inference. We have calculated ML bootstrap values and compared them with alternative phylogenetic methods and criteria including Bayesian inference, maximum parsimony (MP), and distance-based methods. Partition homogeneity tests (PHTs) quantify, how and whether distinct individual genes can be combined into multigene alignments in order to infer a consensus phylogeny. We have applied various techniques to achieve the best-possible reduction of reconstruction artifacts. By application of these techniques, we provide the best-supported phylogenetic tree of PVes so far. It might serve as a basis for improved classifications, outgroup choice for internal phylogenetic analyses, and critical time estimates in future studies. Our results support a multicausal scenario of PV evolution including host-linked evolution, recombination, possible transmission across species borders, and potential adaptive radiation events acting together under mutual influence.

Materials and Methods

For each of the genes (E1, E2, L2, and L1), a representative set of 53 sequences covering the currently known diversity of PVes (table 1) was manually aligned at the amino acid (aa) level and back-translated into codon-aligned nt sequences with Se-Al v2.0a72 (Rambaut 2001). In order to eliminate positions that may not be homologous, or that may have been saturated by multiple substitutions, the alignments of the 4 genes were separately processed with GBlocks (Castresana 2000; supplementary table S1, Supplementary Material online) using the following settings: maximum number of contiguous nonconserved positions, 50; minimum length of a block, 10; allowed gap positions, "half" ("all" for the highly divergent L2 gene sequences). The 3rd-codon position was excluded from all nt analyses in order to avoid evolutionary bias and random phylogenetic clusters and to minimize perturbing effects by convergent evolution at the codon level (Ong et al. 1997). The region of the E4 gene that overlaps with the E2 gene was also excluded from the analysis because it was not possible to reliably align this gene. Final data matrices are available at http://icwww.epfl.ch/~stamatak/index-Dateien/material/Alignment-Data.zip.

The "complete genome" matrix comprising the concatenated E1–E2–L2–L1 sequences was partitioned into the 4 genes (supplementary table S1, Supplementary Material online) in order to investigate previously reported divergent gene evolution in PVes (Bravo and Alonso 2004; García-Vallvé et al. 2005; Narechania et al. 2005). PHTs (Farris et al. 1994) as implemented in PAUP* version 4.0b10 (Swofford 2002) were performed under the MP criterion with 1,000 replicates and heuristic search by random sequence in addition with 10 replicates. The tests investigated the support for the null hypothesis of congruence, and values $P \leq 0.001$ were considered as indicators for significant incongruence between the partitions (Cunningham 1997). We calculated PHT values using both aa and nt data for all 6 possible combinations of gene pairs. For each partition, an individual phylogenetic analysis was performed. Trees were rooted using the 2 known complete bird PV sequences based on a previous E1 tree topology (García-Vallvé et al. 2005).

ML-based phylogenetic analyses were conducted using the parallel Message Passing Interface (MPI) version of RAxML-VI-HPC (Stamatakis 2006b; freely available at http://icwww.epfl.ch/~stamatak). The analyses were executed on the Infiniband cluster at the Technical University of Munich (www.lrr.in.tum.de/Par/arch/infiniband), which is equipped with 136 AMD Opteron processors. Initially, the best-scoring aa substitution model was determined by optimizing branch lengths and model parameters on a fixed random stepwise addition sequence MP RAxML starting tree under the 20 distinct aa substitution models currently implemented in the program. Parameters were optimized on a fixed MP tree because ML model parameters do not change significantly when optimized on a reasonable (i.e., nonrandom) tree (Yang 1996). For L2, the best-scoring aa model was WAG + F + Γ (WAG with empirical base frequencies and the Γ model of rate heterogeneity; Whelan and Goldman [2001]) and rtREV + F + Γ (rtREV with empirical base frequencies and the Γ model of rate heterogeneity; Dimmic et al. [2002]) for the E1, E2, and L1 genes (supplementary table S2, Supplementary Material online).

For DNA analyses, we used the GTR + Γ model of nt substitution (with 4 discrete rate categories) because RAxML only provides GTR + Γ and the GTR + CAT approximation (Stamatakis 2006a) of rate heterogeneity for nt data. The rationale for this is that the shape of the topology

Table 1
List of PVes and Vouchers (EV: Epidermodysplasia verruciformis)

Name	Taxonomy (de Villiers et al. 2004)	Host	Biopsy	Country	GenBank Accession Numbers	References
HPV-2	A4	*Homo sapiens* (Linnaeus, 1758) (Hominidae, Primates)	Pooled DNA from warts of different patients	Not specified	NC_001352 (X55964)	Orth et al. (1977); Hirsch-Behnam et al. (1990)
HPV-18	A7	*H. sapiens* (Linnaeus, 1758) (Hominidae, Primates)	Cervical cancer	Brazil	NC_001357 (X05015)	Boshart et al. (1984); Cole and Danos (1987)
HPV-16	A9	*H. sapiens* (Linnaeus, 1758) (Hominidae, Primates)	Invasive cervical carcinoma	Europe	NC_001526 (K02718)	Seedorf et al. (1985)
HPV-6	A10	*H. sapiens* (Linnaeus, 1758) (Hominidae, Primates)	Condyloma acuminata	Not specified	NC_001355 (X00203)	Gissmann and zur Hausen (1980); Schwarz et al. (1983)
CCPV	A10	*Pan troglodytes* (Blumenbach, 1775) (Hominidae, Primates)	Focal epithelial hyperplasia-like disease	Not specified	NC_001838 (AF020905)	Scinicariello F, Soza I, Brasky KM and Hilliard JK (unpublished data)
PCPV	A10	*Pan paniscus* (Schwartz, 1929) (Hominidae, Primates)	Focal epithelial hyperplasia-like disease in oral cavity	Zoological garden, not specified	X62844	Van Ranst et al. (1991, 1992)
RhPV-1	A12	*Macaca mulatta* (Zimmermann, 1780) (Cercopithecidae, Primates)	Penile squamous cell carcinoma	United States	NC_001678 (M60184)	Kloster et al. (1988); Ostrow et al. (1991)
HPV-54	A13	*H. sapiens* (Linnaeus, 1758) (Hominidae, Primates)	Penile Buschke-Löwenstein tumour coexisting with Condylomata acuminata of 50-year-old man	not specified	NC_001676 (U37488)	Favre et al. (1990)
HPV-5	B1	*H. sapiens* (Linnaeus, 1758) (Hominidae, Primates)	Flat wart from EV patient	Poland	NC_001531 (M17463)	Ostrow et al. (1982); Zachow et al. (1987)
HPV-9	B2	*H. sapiens* (Linnaeus, 1758) (Hominidae, Primates)	Flat wart from EV patient	Poland	NC_001596 (X74464)	Kremsdorf et al. (1982); Delius and Hofmann (1994)
HPV-49	B3	*H. sapiens* (Linnaeus, 1758) (Hominidae, Primates)	Pooled flat warts	Poland	NC_001591 (X74480)	Favre et al. (1989a); Delius and Hofmann (1994)
HPV-92	B4	*H. sapiens* (Linnaeus, 1758) (Hominidae, Primates)	Basal cell carcinoma of 89-year-old man	Australia	NC_004500 (AF531420)	Forslund et al. (2003)
HPV-4	G1	*H. sapiens* (Linnaeus, 1758) (Hominidae, Primates)	Wart from EV patient	Not specified	NC_001457 (X70827)	Heilman et al. (1980); Egawa et al. (1993)
HPV-48	G2	*H. sapiens* (Linnaeus, 1758) (Hominidae, Primates)	Squamous cell carcinoma of the hand of immunosuppressed 36-year-old woman	Not specified	NC_001690 (U31789)	Müller et al. (1989)
HPV-50	G3	*H. sapiens* (Linnaeus, 1758) (Hominidae, Primates)	Actinic keratosis from EV patient	Poland	NC_001691 (U31790)	Kremsdorf et al. (1984); Favre et al. (1989b)
HPV-60	G4	*H. sapiens* (Linnaeus, 1758) (Hominidae, Primates)	Keratinous plantar cyst	Japan	NC_001693 (U31792)	Matsukura et al. (1992)
EEPV	D1	*Alces alces* (Linnaeus, 1758) (Cervidae, Artiodactyla)	Epithelial layer of cutaneous warts	Sweden	NC_001524 (M15953)	Moreno-Lopéz et al. (1981); Ahola et al. (1986)
RPV	D1	*Rangifer tarandus* (Linnaeus, 1758) (Cervidae, Artiodactyla)	Epithelial layer of a cutaneous fibropapilloma	Sweden	AF443292	Moreno-Lopéz et al. (1987); Terai et al. (2002)
DPV	D2	*Odocoileus virginianus* (Zimmermann, 1780) (Cervidae, Artiodactyla)	Pooled fibromas of females	CT, United States	NC_001523 (M11910)	Lancaster and Sundberg (1982); Groff and Lancaster (1985)

Table 1
Continued

Name	Taxonomy (de Villiers et al. 2004)	Host	Biopsy	Country	GenBank Accession Numbers	References
OPV-1	D3	*Ovis aries* (Linnaeus, 1758) (Bovidae, Artiodactyla)	Not specified	Not specified	NC_001789 (U83594)	Karlis J, Delius H, Baired PJ, Meischke HRC, Burrel CJ and Higgins GD (unpublished data)
OPV-2	D3	*O. aries* (Linnaeus, 1758) (Bovidae, Artiodactyla)	Not specified	Not specified	U83595	Karlis J, Delius H, Baired PJ, Meischke HRC, Higgins GD and Burrel CJ (unpublished data)
BPV-1	D4	*Bos taurus* (Linnaeus, 1758) (Bovidae, Artiodactyla)	Fibropapilloma of skin	Not specified	NC_001522 (X02346)	Lancaster and Olson (1978); Chen et al. (1982)
BPV-2	D4	*B. taurus* (Linnaeus, 1758) (Bovidae, Artiodactyla)	Fibropapilloma of skin	Not specified	M20219	Lancaster and Olson (1978); Groff DE, Mitra R and Lancaster WD (unpublished data)
BPV-5	E1	*B. taurus* (Linnaeus, 1758) (Bovidae, Artiodactyla)	"Rice grain" papilloma of benign tumor of teat	Scotland, United Kingdom	NC_004195 (AF457465)	Campo et al. (1981); Terai et al. (2002)
EcPV	Z1	*Equus caballus* (Linnaeus, 1758) (Equidae, Perissodactyla)	Pooled from cutaneous lesions of muzzle and external "nares" of 7 yearling ponies and 1 horse	IL, United States	NC_003748 (AF498323)	O'Banion et al. (1986); Terai et al. (2002)
FPV	H1	*Fringilla coelebs* (Linnaeus, 1758) (Fringillidae, Passerida)	Epithelial warts on tarsus and feet	The Netherlands	NC_004068 (AY057109)	Osterhaus et al. (1977); Terai et al. (2002)
PePV	Th1	*Psittacus erithacus timneh* (Fraser, 1844) (Psittacidae, Psittaciformes)	Cutaneous lesion at head	Western Africa	NC_003973 (AF502599)	O'Banion et al. (1992); Tachezy et al. (2002)
MnPV-1	I1	*Mastomys coucha* (Smith, 1834) (Muridae, Rodentia)	Benign and malignant proliferations of adult animals	Southern Africa	NC_001605 (U01834)	Müller and Gissmann (1978); Tan et al. (1994)
CRPV	K1	*Sylvilagus floridanus* (J. A. Allen, 1890) (Leporidae, Lagomorpha)	Pooled papillomas	KS, United States	NC_001541 (K02708)	Favre et al. (1982); Giri et al. (1985)
CRPVb	K1	*S. floridanus* (J. A. Allen, 1890) (Leporidae, Lagomorpha)	Pooled papillomas	KS, United States	AJ243287	Salmon et al. (1997, 2000)
ROPV	K2	*Oryctolagus cuniculus* (Linnaeus, 1758) (Leporidae, Lagomorpha)	Pooled from lesions at underside of tongue	PA, United States	NC_002232 (AF227240)	Christensen et al. (1996, 2000)
COPV-1	L1	*Canis familiaris* (Linnaeus, 1758) (Canidae, Carnivora)	Papilloma of 5-month-old female beagle	Japan	NC_001619 (D55633)	Isegawa et al. (1995)
FdPV-1	L2	*Felis silvestris* (Schreber, 1775) (Felidae, Carnivora)	Sessile hyperkeratotic skin lesions of Persian cat	United States	AF377865	Carney et al. (1990); Terai and Burk (2002)
"PlPV" (name already occupied)	L, not classified	*Procyon lotor* (Linnaeus, 1758) (Procyonidae, Carnivora)	Papillomatous skin lesions of adult	Toronto Zoo (Ontario, Canada)	NC_007150 (AY763115)	Rector, Van Doorslaer, et al. (2005)
HPV-1	M1	*H. sapiens* (Linnaeus, 1758) (Hominidae, Primates)	Plantar wart	Not specified	NC_001356 (V01116)	Favre et al. (1975), Danos et al. (1982)
HPV-63	M2	*H. sapiens* (Linnaeus, 1758) (Hominidae, Primates)	Pooled punctuate keratotic lesion	Japan	NC_001458 (X70828)	Egawa et al. (1993)
HPV-41	N1	*H. sapiens* (Linnaeus, 1758) (Hominidae, Primates)	Disseminated facial, perianal, and foot warts from 15-year-old girl	Not specified	NC_001354 (X56147)	Grimmel et al. (1988), Hirt et al. (1991)
BPV-3	X1	*B. taurus* (Linnaeus, 1758) (Bovidae, Artiodactyla)	Hyperplasic epithelial warts	Australia	NC_004197 (AF486184)	Pfister et al. (1979); Terai et al. (2002)

Table 1
Continued

Name	Taxonomy (de Villiers et al. 2004)	Host	Biopsy	Country	GenBank Accession Numbers	References
BPV-4	X1	*B. taurus* (Linnaeus, 1758) (Bovidae, Artiodactyla)	Esophagus papilloma	Scotland, United Kingdom	X05817	Jarrett et al. (1978); Patel et al. (1987)
BPV-6	X1	*B. taurus* (Linnaeus, 1758) (Bovidae, Artiodactyla)	Frond epithelial papillomas of udder	Scotland, United Kingdom	AJ620208	Jarrett et al. (1984); Jackson et al. (1991)
PsPV-1	O1	*Phocoena spinipinnis* (Burmeister, 1865) (Phocoenidae, Cetacea)	Genital wart	Peru	NC_003348 (AJ238373)	Van Bressem MF, Cassonet P, Rector A, Desaintes C, van Waerebeek K, Alfaro Shigeto J, van Ranst M and Orth G. (forthcoming)
HaOPV	P1	*Mesocricetus auratus* (Waterhouse, 1839) (Muridae, Rodentia)	Lesions in lingual mucosa	Syria	E15111	Iwasaki et al. (1997)
TmPV	R1	*Trichechus manatus latirostris* (Harlan, 1824) (Trichechidae, Sirenia)	Sessile papillomatous skin lesion of female	FL, United States	NC_006563 (AY609301)	Rector et al. (2004)
EdPV	S1	*Erethizon dorsatum* (Linnaeus, 1758) (Erethizontidae, Rodentia)	Epidermal hyperplasia, with acanthosis and orthokeratotic hyperkeratosis, from multiple white to light brown lobulated, raised, firm masses on foot pads	New York Bronx Zoo, NY, United States	NC_006951 (AY684126)	Rector et al. (2005)
BPV-7	Not classified	*B. taurus* (Linnaeus, 1758) (Bovidae, Artiodactyla)	Teat	Japan	NC_007612 (DQ217793)	Ogawa et al. (2004)
ChPV	Not classified	*Capra hircus*, (Linnaeus, 1758) (Bovidae, Artiodactyla)	Healthy skin of 7-year-old female	Belgium	NC_008032 (DQ091200)	Van Doorslaer et al. (2006)
CfPV-2	Not classified	*C. familiaris* (Linnaeus, 1758) (Canidae, Carnivora)	Foot pad papilloma of a Golden retriever	United States	NC_006564 (AY722648)	Yuan et al. (forthcoming)
CPV-3	Not classified	*C. familiaris* (Linnaeus, 1758) (Canidae, Carnivora)	Skin lesions from 7-year-old Rhodesian ridgeback with canine EV and in situ squamous cell carcinoma	Switzerland	NC_008297 (DQ295066)	Tobler et al. (2006)
HPV-101	Not classified	*H. sapiens* (Linnaeus, 1758) (Hominidae, Primates)	Cervicovaginal cells from a 34-year-old woman with intraepithelial neoplasia grade 3	Costa Rica	NC_008189 (DQ080081)	Chen Z, Schiffman M, Herrero R, DeSalle R and Burk RD. (2007)
HPV-103	Not classified	*H. sapiens* (Linnaeus, 1758) (Hominidae, Primates)	Cervicovaginal cells from 30-year-old woman with normal cytology	Costa Rica	NC_008188 (DQ080078)	Chen Z, Schiffman M, Herrero R, DeSalle R and Burk RD. (2007)
McPV-2	Not classified	*M. coucha* (Smith, 1834) (Muridae, Rodentia)	Anal lesion	Southern Africa	DQ664501	Nafz J, Ibberson M, Bravo I, Nindl I, Stockfleth E and Roesl F. (unpublished data)
RaPV	Not classified	*Rousettus egyptiacus* (E. Geoffroy, 1810) (Pteropodidae, Chiroptera)	Basosquamous carcinoma on the left wing membranes	Egypt	NC_008298 (DQ366842)	Rector et al. (2006)
TtPV-2	Not classified	*Tursiops truncatus* (Montagu, 1821) (Delphinidae, Cetacea)	Genital Condylomata	Off SC, United States	NC_008184 (AY956402)	Rehtanz et al. (2006)

Table 2
Partition Homogeneity Tests (of amino acid alignments if not otherwise specified; test rejections are indicated in bold; note that analyses including the L2 gene render predominantly the weakest values)

	E1–E2	E1–L2	E1–L1	E2–L2	E2–L1	L1–L2
All taxa, nt sequences (3rd-codon position removed)	0.020	**0.001**	**0.001**	0.013	0.010	**0.001**
All taxa	0.735	**0.001**	0.162	0.004	0.304	0.009
Excluding HPV-16, μ-PVes, "PlPV"	0.563	0.005	0.482	0.017	0.312	0.121
α-PVes	0.644	0.280	0.383	0.197	0.270	0.271
κ-, λ-, μ-, ν-, σ-PVes	0.802	0.444	0.636	0.025	0.381	0.086
λ-PVes	0.950	0.824	0.859	0.773	0.868	0.338

has a significantly higher impact on final likelihood values than model details. Therefore, RAxML implements a technically highly optimized GTR likelihood function which allows for a more exhaustive exploration of the huge tree search space, and thus yields better results than competing ML programs on real data (Stamatakis et al. 2005; Stamatakis 2006b). Nonetheless, the usage of rate heterogeneity has a significant impact on final tree shapes. An estimate of the proportion of invariant sites is not implemented in RAxML due to statistical concerns regarding the simultaneous usage of Γ- and P-Invar, which are discussed in the RAxML manual. Finally, we did not use the significantly faster GTR + CAT approximation of rate heterogeneity because the alignments were relatively small with respect of the number of taxa, and thus, we were concerned about insufficient data for the estimation of per-site evolutionary rates. Moreover, trees inferred under GTR + CAT scored on average slightly worse (1–2 log likelihood units) under GTR + Γ than trees inferred entirely under GTR + Γ.

We analyzed all multigene alignments under both plain (one set of ML substitution parameters was estimated over the entire alignment) and mixed models (ML model parameters were estimated separately for each gene). In order to determine the best-known ML tree for each alignment/model combination, we executed 127 tree searches from distinct random stepwise addition sequence MP starting trees on 128 processors of the Infiniband cluster. Therefore, each central processing unit (CPU) executed one tree inference on a distinct starting tree, whereas the 128th CPU acted as master process for work distribution as previously described (Stamatakis 2006b). Thereafter, we executed 1,000 nonparametric bootstraps for each alignment with RAxML, and the bootstrap values were drawn on the best-scoring ML-tree using the respective RAxML program option (see RAxML manual for details). In total, we executed over 10,000 nonparametric bootstraps and over 1,270 ML searches for best-known trees.

Bayesian phylogenetic analyses were performed with BEAST version 1.3 (Drummond et al. 2002; Drummond and Rambaut 2003; freely distributed by the authors at http://evolve.zoo.ox.ac.uk/beast/). For the WAG + Γ aa substitution model (Whelan and Goldman 2001) with 4 discrete Γ rate categories as well as for the HKY + Γ nt substitution model (Hasegawa et al. 1985) with 4 discrete categories, we used an uncorrelated relaxed clock. In such models, the rate for each branch of the tree is drawn independently and identically from the underlying exponential distribution (Drummond et al. 2006). Parameter values were optimized via Markov Chain Monte Carlo methodology after repetitive short heuristic searches (50,000 iterations with 10,000 burn-in cycles). The unweighted pair group method with arithmetic mean was used to construct a starting tree for the BEAST analyses, and the final topology was estimated based on 1,000,000 iterations using 100,000 burn-in cycles and sampling every 1,000 iteration.

MP and distance-based calculations were run in PAUP*. Trees were generated by performing heuristic searches with tree bisection reconnection and starting trees obtained via random taxon addition with 10 replicates (parsimony) or Neighbor-Joining (distance measure: mean character difference), respectively. No upper limit for the number of equally parsimonious trees was specified. In addition, we assessed the performance of the parsimony ratchet (Nixon 1999) in order to search all most-parsimonious tree (MPT) islands, despite the fact that the number of MPTs was low during heuristic searches with PAUP*. We used perlRat v.1.9a (Bininda-Emonds 2006) to generate batch files for parsimony ratchet runs with PAUP*. Nonparametric bootstrap support was estimated based on 1,000 replicates using the same search strategy as in the tree searches. The best-fit substitution model for nt data was selected based on the Akaike Information Criterion as implemented in Modeltest 3.7 (Posada and Crandall 1998) and was used for distance-based analyses (with ML settings). Gaps were treated as missing data in all analyses.

Results

The E1–E2–L1 Open Reading Frames of PVes Are Phylogenetically Congruent

Data on length and number of informative sites of the aa and nt alignments (calculated with the best-fit model, GTR + Γ + I; number of substitution types, 6; number of distinct data patterns under this model, 4003 using the complete data matrix in PAUP* analyses) used in this study is provided in supplementary table S1 (Supplementary Material online). Overall, PHT values were low between gene pairs of nt sequence data ($P \leq 0.020$; table 2) but were consistently higher for aa sequence data. With respect to aa sequence data, each gene pair that included the L2 gene yielded low PHT values ($P < 0.010$ taking into consideration the entire taxon sampling), whereas all other pairs rendered values above the threshold. PHT values increased, even in analyses including the L2 gene, when PVes with

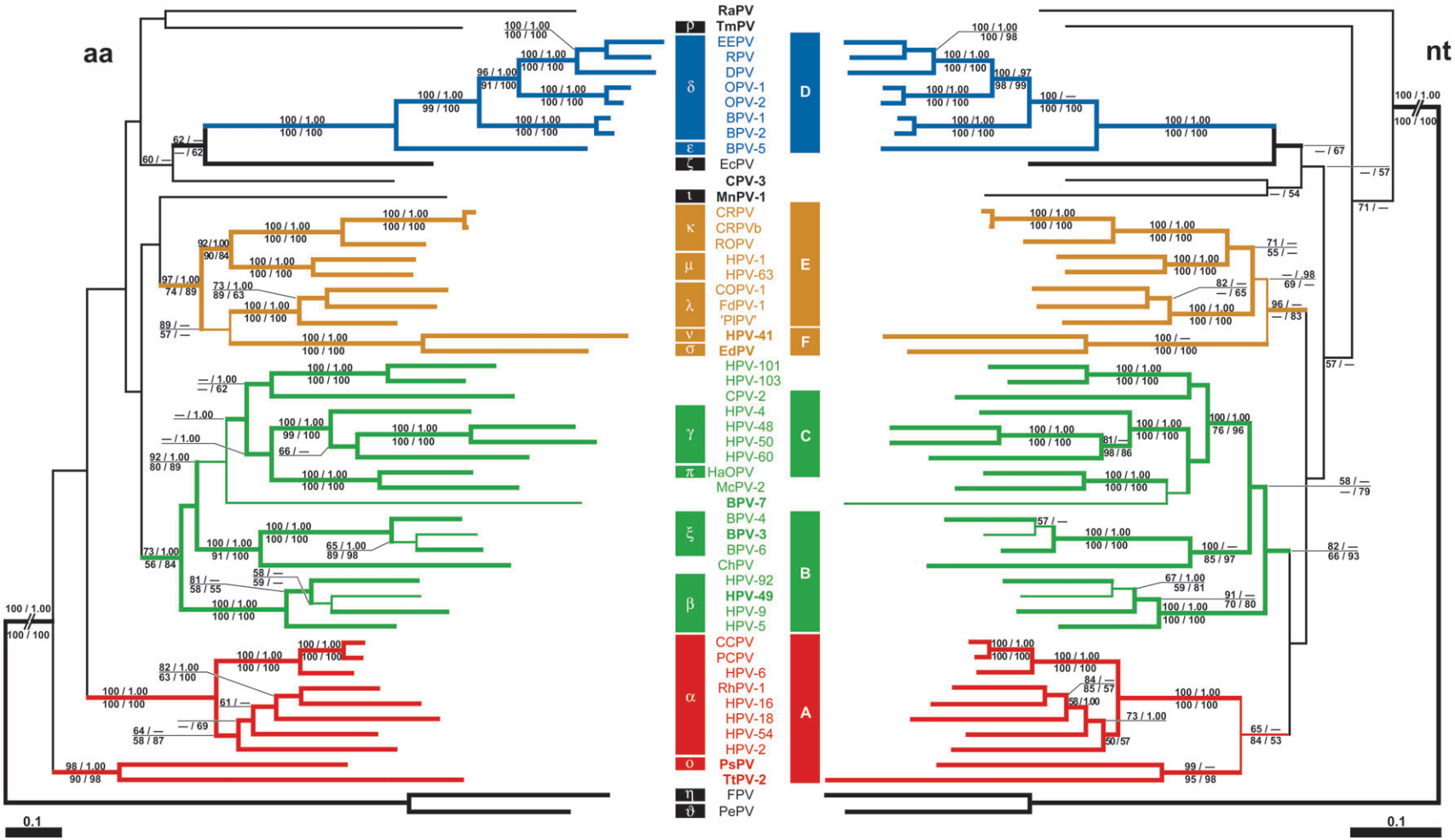

FIG. 1.—Phylogenetic comparison of amino acid (aa: 1,293 parsimony-informative positions) and nt sequence data (nt: 2,407 parsimony-informative positions) of 53 phylogenetically representative PVes. All available non-HPVes and 18 representative HPV types were used for analyses. Genera PV clades are indicated by Greek lettering, the supertaxa are colored blue (δ+ ε), ocher (κ+ μ+ λ+ ν+ σ), green (γ+ π+ ξ+ β), and red (α+ ο), respectively. Branch lengths are drawn to scale with the scale bar indicating the number of amino acid substitutions per site. Numbers on branches are bootstrap support values to clusters on the right of them (above: criteria = ML/Bayesian probabilities; below: criteria = MP/distance; values under 50 are not shown). Bold branches indicate congruence between aa and nt; note that tree topologies do not show significant contradictions.

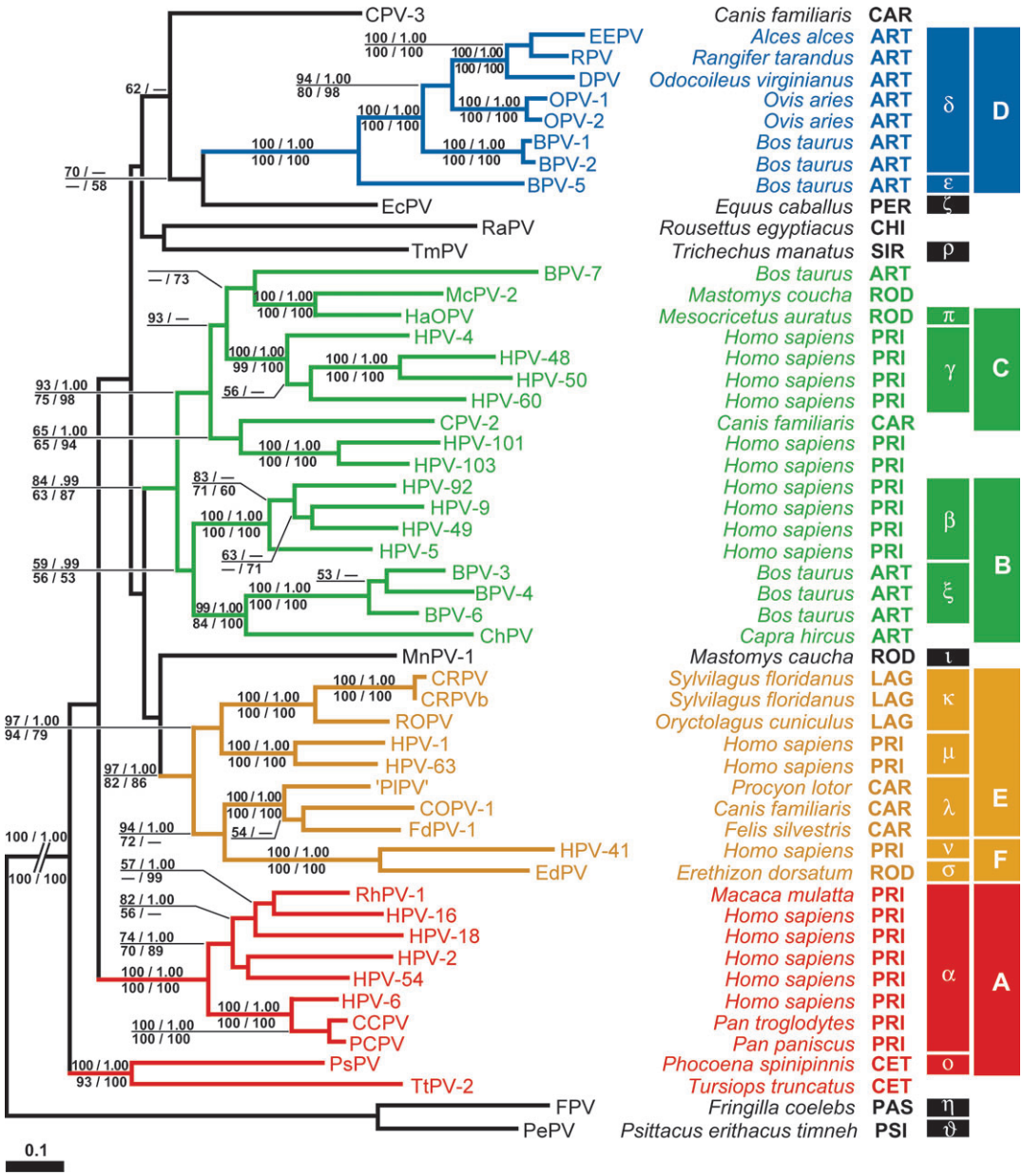

FIG. 2.—ML tree of 53 phylogenetically representative PVes as inferred from a combined E1–E2–L1 amino acid sequence analysis (1,082 parsimony-informative positions) justified by PHTs (table 2). All non-human PVes and 18 representative HPV types were used for analyses. PV genera (de Villiers et al. 2004) are indicated by Greek lettering, upper case lettering follow an alternative E1–E2 classification of Bravo and Alonso (2007). Higher order host taxa are abbreviated as follows: ART, Artiodactyla; CAR, Carnivora; CET, Cetacea; CHI, Chiroptera; LAG, Lagomorpha; PAS, Passeriformes; PER, Perissodactyla; PRI, Primates; PSI, Psittaciformes; ROD, Rodentia; and SIR, Sirenia. The supertaxa are colored blue (δ+ ε), ocher (κ+ μ+λ+ ν+ σ), green (π+ γ+ β+ ξ), and red (α+ ο), respectively. Branch lengths are drawn to scale, with the scale bar indicating the number of amino acid substitutions per site. Numbers on branches are bootstrap support values to clusters on the right of them (above: criteria = ML/Bayesian probabilities; below: criteria = MP/distance; values under 50 are not shown).

well-supported, contradicting phylogenetic positions (i.e., HPV-16, *Mupapillomaviruses* [μ-PVes], "PlPV"; see below) were excluded from the analyses and when well-defined, taxonomic subsets were separately investigated (e.g., α-, λ-PVes, and the supergroup comprising *Kappapapillomaviruses* [κ-PVes], *Nupapillomaviruses* [ν-PVes], *Sigmapapillomaviruses* [σ-PVes], λ-, and μ-PVes). The results shown in supplementary table S2 (Supplementary Material online) support the combination of the E1–E2–L1 ORFs at the aa level for a simultaneous phylogenetic analysis, but not the combination of the L2 ORF with any other gene. Therefore, we performed thorough phylogenetic analyses with each single gene and with the E1–E2–L1 gene combination.

Phylogenetic Relationships in PVes Have Largely Reliable Support

The various phylogenetic approaches explored in this study comprising different sequence data (nt vs. aa

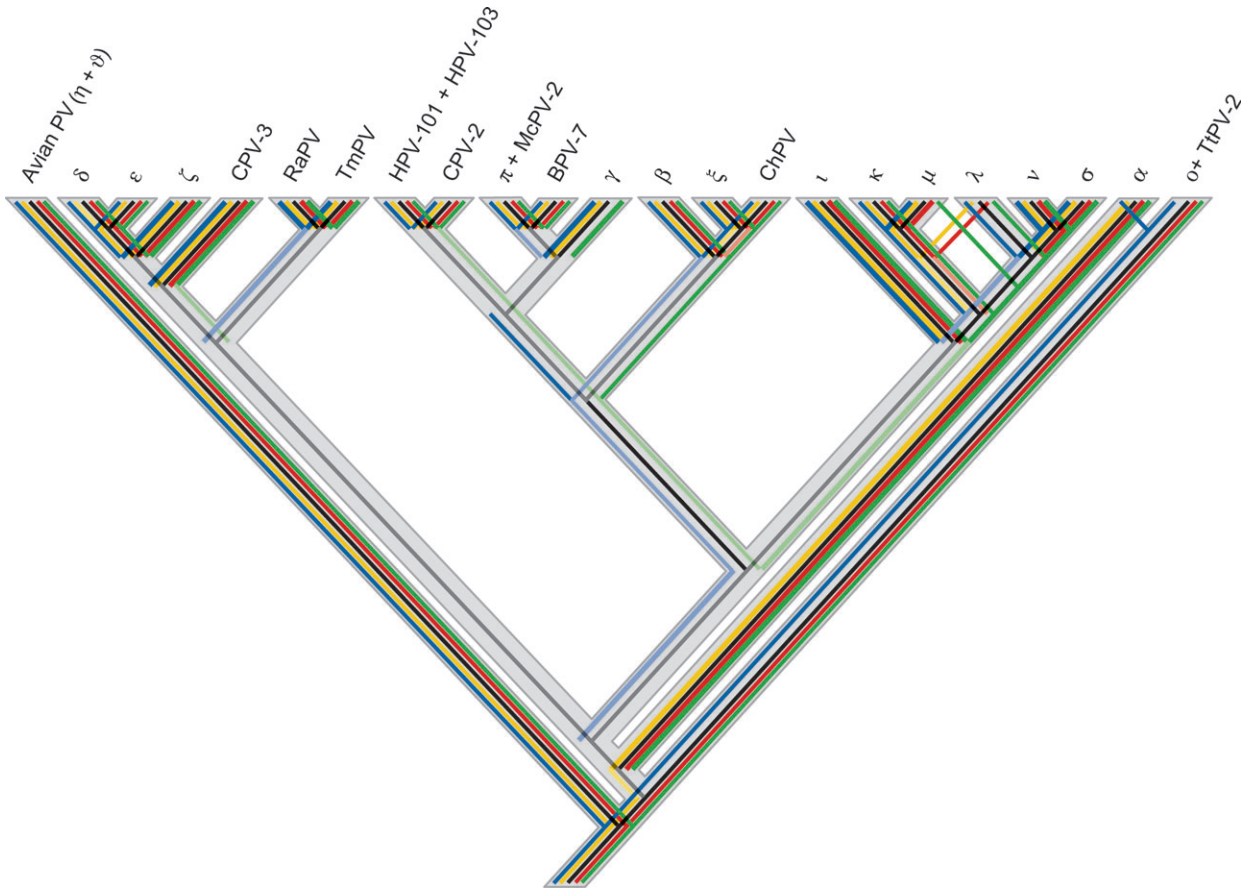

FIG. 3.—Cladogram of PVes summarizing the ML results. The 3-gene analysis provides the phylogenetic backdrop (black), on which congruent branches of the E1 (blue), E2 (ocher), L2 (red), and L1 (green) phylogenies are projected. Branches with high bootstrap support (BS >75) are dark colored and those with lower values (BS < 75) are light colored.

sequences), different partitions (separated genes vs. combined analyses), different models (mixed models vs. plain models), and different methodological criteria (ML, Bayesian inference, MP and distance) did not render overall congruent phylogenies. Nonetheless, the statistic support for many internal nodes was extraordinarily high. A series of major monophyletic assemblages could clearly be distinguished, 1) independently of whether the data was analyzed at aa or nt level (fig. 1); 2) independently of whether the data was analyzed simultaneously or in separate partitions (figs. 1 and 3 and Supplementary fig. S1, Supplementary Material online); and 3) independently of the alternative phylogenetic methods (figs. 1–3 and Supplementary fig. S2, Supplementary Material online).

The PV genera, including *Xipapillomaviruses* (ξ-PVes), *Gammapapillomaviruses* (γ-PVes), α-, β-, δ-, κ-, λ-, and μ-PVes were each confidently monophyletic under the different approaches investigated (bootstrap support values from ML analyses, maximum likelihood bootstrap support (MLBS); MP analyses, maximum parsimony bootstrap support (MPBS); and distance analyses, distance bootstrap support (DBS) each > 75 and Bayesian posterior probabilities, bayesian posterior probability (BPP) > 0.95), with the exceptions of γ-PVes that were not well supported in the separate L2 and L1 analyses and a few other nodes in nt and in separate E2 and L2 aa analyses (fig. 1 and Supplementary fig. S1, Supplementary Material online). Monophyly of PV clades therefore corresponded to monophyly of some infected host taxa such as Primates (α-, β-, γ-, μ-PVes), Artiodactyla (δ-, ξ-PVes), Lagomorpha (κ-PVes), and Carnivora (λ-PVes).

Deeper phylogenetic nodes showed similarly high confidence values in the various analyses (MLBS, MPBS, DBS > 75 and BPP > 0.95): δ + ε-PVes (both infecting Artiodactyla), HPV-101 + HPV-103 (infecting Primates), π + McPV-2 (infecting Rodentia, only 71 MPBS in the L1 analysis), and ν + σ-PVes (infecting Primates and Rodentia). Furthermore, the 2 groupings super-ξ-PVes (including ChPV) and super-γ-PVes (comprising γ- and π-PVes, BPV-7, CfPV-2, HPV-101 and HPV-103) were well supported by the multigene analyses (figs. 1–2).

Figure 2 shows the best-scoring likelihood tree of the combined genes E1–E2–L1 that was calculated using the best-fit model (rtREV + F + Γ; supplementary table S2, Supplementary Material online) with the statistical support values for each of the 4 phylogenetic approaches used. The relationships between the major monophyletic assemblages described above and the remaining PVes were not fully resolved, but the following 7 groupings and individual viruses could be confidently stated at highest taxonomic level

(fig. 2 and Supplementary fig. S2, Supplementary Material online):

1) A diverse and heterogeneous clade (with respect to the hosts) comprising the groups κ, λ, μ, ν, and σ (97 MLBS, 1.00 BPP, 82 MPBS, and 86 DBS). λ-, ν-, and σ-PVes clustered together (94 MLBS, 1.00 BPP, and 72 MPBS) and were the sister group of κ- and μ-PVes (97 MLBS, 1.00 BPP, 94 MPBS, and DBS 79). PVes in this κ + λ + μ + ν + σ-supertaxon were isolated from Carnivora, Lagomorpha, Primates, and Rodentia.
2) Another diverse and heterogeneous clade (with respect to the hosts) comprised super-γ-, super-ξ-, and β-PVes (84 MLBS, 0.99 BPP, 86 MPBS, and 56 DBS), whereas the latter 2 appeared to be closely related by moderate statistical support (59 MLBS, 0.99 BPP, 56 MPBS, and 53 DBS). PVes in this β + γ + π + ξ-supertaxon infected Artiodactyla, Carnivora, Primates, and Rodentia.
3) Close relationship of the 2 PVes isolated from Cetacea: PsPV-1 and TtPV-2 (100 MLBS, 1.00 BPP, 93 MPBS, and 100 DBS).
4) Close relationship between the artiodactylan δ + ε-supertaxon with the equine PV (70 MLBS and 58 DBS), and both were closely allied to CPV-3 (62 MLBS).

However, relationships between these 4 groups and 5) α-PVes, 6) RaPV, and 7) TmPV only showed low statistical support. α-PVes and the 2 PVes isolated from Cetacea were closely related as inferred from the separate E1 gene analysis (α + o-supertaxon: 87 MLBS, 1.00 BPP, 70 MPBS, and 75 DBS; Supplementary fig. S1, Supplementary Material online).

Some Nodes Show Well-Supported Phylogenetic Contradiction

The comparison of the 4 separate E1-, E2-, L2-, and L1-phylogenies identified 3 (groups of) PVes with highly supported, contradicting phylogenetic positions under the ML criterion (MLBS > 75; frequently supported also by alternative methods), namely HPV-16 within α-PVes, "PlPV" within λ-PVes, and μ-PVes within the κ + λ + μ + ν + σ-supertaxon (Supplementary fig. S1 and table S3, Supplementary Material online). In addition, a series of nodes contradicted close relationships identified in the 3-genes ML tree (fig. 2) with high statistical support from alternative methods (DBS, MPBS > 75 and BPP > 0.95; see supplementary table S4 [Supplementary Material online] for details).

Phylogenies of PVes and Their Hosts Are Partly Incongruent

Some of the PV groups recovered were largely congruent to the corresponding taxa of the mammals they infected (see above). However, a series of PV types did not cluster accordingly to the phylogeny of their hosts:

1. PVes infecting Primates: α-, β-, γ-, μ-, and ν-PVes never constituted a monophyletic group in any of our analyses, and not even 2 of them showed a close relationship.
2. Non-human PVes infecting Primates: RhPV-1 (rhesus monkey) and CCPV + PCPV (chimpanzees) nested within α-PVes and appeared independently derived from the paraphyletic HPVes of this group.
3. PVes infecting Bovidae: OPV-1, OPV-2, BPV-1, BPV-2, and BPV-5 were paraphyletic, and not monophyletic, as their mammalian host taxon; to the contrary, OPV-1 + OPV-2 showed a well-supported close relationship with PVes infecting Cervidae (i.e., DPV, EEPV, RPV; 94 MLBS, 1.00 BPP, 80 MPBS, 98 DBS).
4. Other PVes infecting Artiodactyla: BPV-7 and ChPV + ξ-PVes were each only distantly related with the core artiodactylan δ + ε-PVes.
5. PVes infecting Cetartiodactyla: Cetacean PVes (PsPV-1 + TtPV-2) did not cluster with any of the artiodactylan PVes, not even with the core δ + ε-PVes.
6. PVes infecting Carnivora: neither CfPV-2 nor CPV-3 was closely related to the core carnivoran λ-PVes, but alternatively with HPV-101 + HPV-103 (65 MLBS, 1.00 BPP, 65 MPBS, 94 DBS) and δ + ε + ζ-PVes (62 MLBS), respectively.
7. PVes infecting Rodentia: McPV-2 + π-PVes, MnPV-1, and TmPV did not constitute an own clade.

Discussion

The E1–E2–L1 Gene Combination is Suitable for the Reconstruction of Papillomavirus Evolution

The importance of knowledge about the evolution of pathogens such as PVes has been increasingly acknowledged during the last years (Bernard 2005). Information about the PV phylogeny will effectively contribute to their classification, which is of diagnostic and therapeutic relevance. For example, phylogenetic inference helps differentiate between PV risk groups with respect to benign or malignant lesions (Van Ranst et al. 1992; Bible et al. 2000; Muñoz et al. 2003; Bravo and Alonso 2004, 2007; Chen et al. 2005; Schiffman et al. 2005). Furthermore, the development of evolutionary scenarios might reveal currently elusive relationships between the viruses and their microenvironment (i.e., the single infected cell) and/or macroenvironment (i.e., the skin tissue), which are the basis for hypotheses that can be experimentally verified.

So far, previous molecular studies have focused on particular ORFs such as the E1–E2 genes (Bravo and Alonso 2007), the E5 gene (Bravo and Alonso 2004), the E6 gene (Van Ranst et al. 1995), the E7 gene (Van Ranst et al. 1992), and the L1 gene (Chan et al. 1995; de Villiers et al. 2004). Only few approaches have incorporated as much genetic information as possible to investigate various PV genes separately (García-Vallvé et al. 2005) and/or particular subordinate PV groups (e.g., HPVes: Van Ranst et al. [1995], α-PVes: Narechania et al. [2005], β-PVes: Gottschling et al. [2007]). Thus, we present the first comprehensive analysis on the internal phylogenetic relationships of PVes in this study using the 4 large genes and covering the currently known diversity.

We have aimed to minimize reconstruction artifacts by manual refinement of the alignment. The PHTs (table 2) indicate that the E1–E2–L1 ORF combination at aa level is well suited for simultaneous phylogenetic inference of the entire PV sequence data set. However, the inclusion of the L2 gene in analyses appears to be only justified when the reconstruction of PVes at a lower taxonomic level is addressed. Our study therefore identifies those parts of the PV genome that can confidently be combined to minimize the degree of data-inherent perturbation for phylogenetic analyses of PVes.

Based on extraordinarily high statistical support for many nodes, we confirm the existence of a series of PV clades that have been previously ranked as "genera" based on L1 gene analyses (de Villiers et al. 2004; Bernard et al. 2006). Our results reliably expand the knowledge about basic relationships between such groups of PVes, particularly by identifying the supertaxa $\beta + \gamma + \pi + \xi$- and $\kappa + \lambda + \mu + \nu + \sigma$-PVes. This is a clear advantage with respect to the present formal listing of more than 15 equally ranked groups and may be of importance for future PV classification. Despite our extensive phylogenetic analyses, the precise positions of the supertaxa α + o- and $\delta + \varepsilon$-PVes as well as of a few isolated PVes including CPV-3, EcPV, MnPV-1, RaPV, and TmPV remain currently unresolved.

Host-linked Evolution Alone Cannot Explain the Molecular Trees of PVes

The apparent congruence between phylogenies of both PVes and their mammalian hosts has initially led to the assumption that host-linked evolution is the driving force for virus diversification (Bernard et al. 1994; Myers et al. 1994; Chan et al. 1995; Halpern 2000; de Villiers et al. 2004; García-Vallvé et al. 2005; Bernard et al. 2006). Despite their proven importance, in-depth investigations of the role of coevolutionary interactions in phylogenetic diversification of pathogens and host lineages are remarkably limited (Nunn 2004). Coevolution is plausible if the phylogeny of a group of hosts is congruent with the phylogeny of a group of corresponding parasites, organelles, or pathogens. The presence of α-PVes on Primates, δ-PVes on Artiodactyla, and λ-PVes on Carnivora makes, for example, such an assumption plausible at a first glance.

However, our results challenge the view that host-linked evolution fully explains the molecular PV trees without alternative. Viral phylogeny is frequently incongruent to hominid phylogeny (Purvis 1995) at a broad scale, and molecular data for PVes also contradicts the coevolutionary hypothesis: non-human PVes infecting Primates (RhPV-1 and CCPV+PCPV) do not have basal, but highly derived and polyphyletic positions within α-PVes, and are closely related to different HPV types (fig. 2). Concomitantly, the numerous HPVes are not monophyletic, though this would have been expected if strict coevolution between hominids and PVes had occurred. This is in agreement with previous studies that showed a large diversity of non-human PVes nesting within α- and β-PVes in a polyphyletic pattern (Chan et al. 1997; Antonsson and Hansson 2002; Gottschling et al. 2007).

Another instructive example for evolutionary incongruence between host- and PV-phylogenies is given by the monophyly of the Bovidae (Hernández Fernández and Vrba 2005) and the paraphyly of the corresponding PVes from the δ-group (fig. 2). Furthermore, bovine PVes infecting the same host (*Bos taurus*) are found in at least 3 only distantly related lineages. This is in agreement with a previous study that found a broad spectrum of only distantly related bovine PVes (Ogawa et al. 2004). Other cases of incongruence between PV- and host-phylogenies comprise PVes that infect Cetartiodactyla, Rodentia, and Carnivora (see results section), for which the hypothesis of exclusively host-linked evolution in PVes is likewise rejected by the molecular trees.

Various Putative Interferences Including Ancient Recombination Events May Perturb the Reconstruction of Papillomavirus Evolution

The question arises whether phylogenetic incongruence between PVes and hosts reflects the natural history of the viruses or whether it is due to reconstruction artifacts, as suggested for metazoan phylogeny (Baurain et al. 2006). Long branch attraction by rate heterogeneity among different parts of the tree (Philippe et al. 2005) is frequently discussed as a reason for phylogenetic interferences, but may have played a minor role in our reconstructions. The branches of the trees are largely well balanced, with the only exception of some isolated PVes showing uncertain phylogenetic positions (e.g., CPV-3, EcPV, MnPV-1, RaPV, and TmPV). Particularly, the L2 phylogeny exhibits some prime outliers (e.g., RaPV and TtPV-2; Supplementary fig. S1, Supplementary Material online), and we have excluded this gene from our simultaneous analysis based on the PHT results. The limitation arising from gene exclusion might become negligible in phylogenomics because it is possible to discard more than half of the data, although still recovering highly supported and plausible trees substantially devoid of tree reconstruction artifacts (Delsuc et al. 2005; Jeffroy et al. 2006).

Evolutionary disturbance may account for conflicting tree topologies when different genetic regions are investigated separately (Bravo and Alonso 2004; García-Vallvé et al. 2005; Gottschling et al. 2007). Recently, recombination, which necessarily had to occur within single-host cells, has been investigated more rigorously, and up to 7 such events have been reconstructed in PVes by bioinformatics approaches (Narechania et al. 2005; Varsani et al. 2006). Five of the possible recombination sites are located in the L2 gene, which is in agreement with our PHT results. They clearly indicate the phylogenetic incongruence between L2 and the remaining genes. The potential for evolutionary signal perturbation of L2 is underlined by the large amount of positions that may not be homologous, or may have been saturated by multiple substitutions, and have therefore been removed from our analyses (only 35% of the original length after GBlocks processing; supplementary table S1 [Supplementary Material online]). Derived both from the relatively low number of well-supported phylogenetic conflicts (Supplementary fig. S1 and tables S3–S4, Supplementary Material online) and

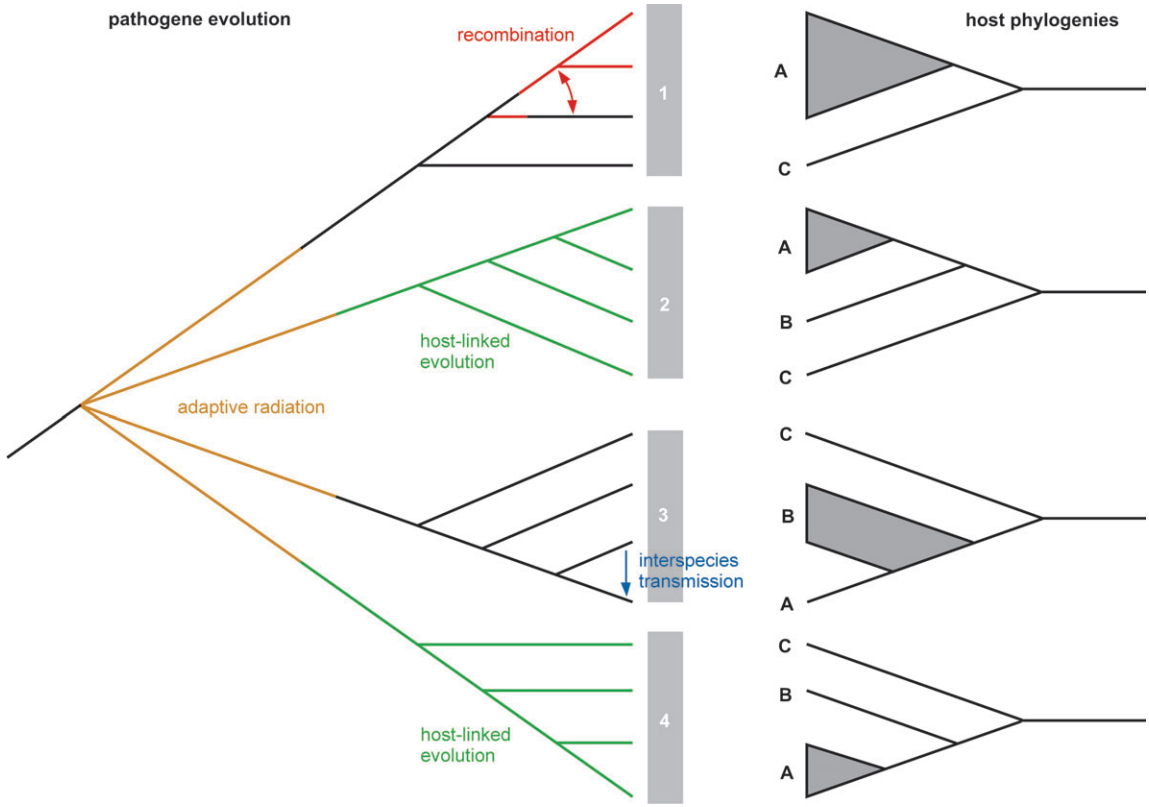

FIG. 4.—Exemplification of 4 evolutionary mechanisms that might drive evolution of PVes, namely adaptive radiation, host-linked evolution, recombination, and lateral gene transfer by, for example, interspecies transmission. Model PV species 1, 2, 3, and 4 and host species A, B, and C are indicated. Note the possible absence of a corresponding PV type on host B (uppermost cladogram) due to, for example, extinction.

from the relative stability of the trees using multigene matrices, ancient recombination should be considered rather rare events in PVes.

Knowledge about diversity is still extremely sparse, not only but also especially for PVes, and insufficient taxon sampling can have a major impact on phylogenetic analyses (Jeffroy et al. 2006). Thus, the present scattered and fragmentary collection of both human and non-human PVes might influence any reconstruction of PV evolution. The number of complete PV genomes available is extremely biased toward human sources due to the clinical focus on the association between HPV infections and various types of cancer. The insufficient sampling of non-human PVes is best illustrated by the historical distinction between "HPVes" and "animal PVes" as were they two distinct and unrelated entities (Bernard 2005). Thus, increasing the taxon sampling by isolating and sequencing novel PVes from systematically selected hosts is of crucial importance. This may shed light on viral evolution and on the interactions with their hosts by breaking the long branches that lead to isolated viruses in the molecular trees.

Additional evolutionary phenomena at the molecular level such as strong codon usage bias between PVes and their hosts have been reported (Ong et al. 1997; Zhou et al. 1999; Zhao et al. 2003; Mossadegh et al. 2004; Bravo and Müller 2005). Codon preferences might contribute to the weak PHT values when investigating PV sequence data at the nt level, even under exclusion of the 3rd-codon position in our analyses. We have aimed to avoid such phylogenetic interferences by the usage of sequence data at the aa level. However, the impact on the trees by persistence of ancestral polymorphisms ("incomplete lineage sorting": Maddison [1997]; Maddison and Knowles [2006]) remains to be determined in future studies.

Infections across Species Borders and Adaptive Radiations May Have Additionally Contributed to Papillomavirus Diversification

Given the assumption that the evolutionary incongruence essentially reflects the natural history of both PVes and their hosts, alternative explanations for the PV tree topologies should be discussed. Our molecular data suggest that PVes infect groups of organisms (lineages) rather than particular species, at least in geological times. As previously suggested (Myers et al. 1996; Rector, Van Doorslaer, et al. 2005; Gottschling et al. 2007), lateral gene transfer by infections across species borders may be relatively frequent within those groups of close relationship. The more the phylogenetic distance grows between the native and the putative new hosts, the more such zoonoses will become unlikely. For closely related PVes, the main obstacle for infection of novel hosts might usually be the absence of physical contact, exemplified by the exceptional case of a zookeeper who temporarily tested positive for

a chimpanzee PV (Antonsson and Hansson 2002). Furthermore, the bovine PV-1 is able to infect horses and to produce sarcoids (Pfister et al. 1981; Otten et al. 1993; Chambers et al. 2003), and even roughly half of the healthy horses in contact with infected fellows carry bovine viruses in their skin (Bogaert et al. 2005). Finally, the perpetuation of host specificity might have particular importance in the α + o-supertaxon, where sexual intercourse is required for contagion.

The increasing proximity of human and animal populations has generally led to the increase of zoonotic transmission events, but the factors causing them are still poorly understood (Mahy and Brown 2000). Multiple invasions of only distantly related mammals may explain the existence of clades such as the super-γ-PVes infecting humans and domestic animals such as hamster, dog, and cattle. PV establishment on a new host has not yet been experimentally verified (Halpern 2000; Bernard et al. 2006), but endothermy of the hosts might be one of the licenses that allows the viruses to cross the species barrier. Furthermore, a low immune status may facilitate invasions into comparable ecological environments provided by putative new hosts as shown for influenza viruses (Weiss 2003; Fislova and Kostolansky 2005; Kaye and Pringle 2005). However, the underlying mechanisms of host invasion have not been seriously addressed for PVes to date. To the contrary, their presence in humans has exclusively been regarded as old primate inheritance (Bernard et al. 2006), without considering alternative explanations and without accounting for the topological inconsistencies described above.

The split between CCPV + PCPV and HPV-13 has been considered to reflect the speciation between *Pan* and *Homo* (Van Ranst et al. 1995; Halpern 2000), but such an assumption ignores the derived phylogenetic position of chimpanzee PVes within the α-PVes (García-Vallvé et al. 2005; Bravo and Alonso 2007). A similar polyphyletic pattern of non-human PVes nested within various HPV species has also been observed for the β-PVes (Gottschling et al. 2007). Despite the inferred importance of the evolutionary mechanisms discussed above, adaptive radiation in a PV ancestor (e.g., by establishment of new ecological niches) followed by temporally close, host-linked evolution (García-Vallvé et al. 2005; Bravo and Alonso 2007) may also explain the present tree topologies of α- and β-PVes. Initial analyses for the identification of PVes in the normal skin of different animals have recovered hundreds of partial sequences from putative novel viruses (Forslund et al. 1999; Antonsson and Hansson 2002; Ogawa et al. 2004). These results suggest that a puzzling diversity of PVes within the same host is the rule rather than the exception. However, this implies that host-linked evolution can be primarily reconstructed at shallow (such as the L1 gene-based "species," de Villiers et al. [2004]) rather than at deeper taxonomic level ("genera").

Conclusions

Our data shows that nt alignments harbor more sequence heterogeneity than aa alignments, and we propose to exclusively use aa sequence data in future PV phylogenetic analyses despite the significantly larger computational cost. Comparing the single genes in separate analyses, only few nodes show well-supported phylogenetic contradictions. Particularly, the L2 gene appears to exhibit a high potential of biasing phylogeny reconstructions, which is a strong argument to use the E1–E2–L1 ORF combination for multigene analyses. Based on well-resolved molecular phylogenies using this gene combination, diversification within PVes cannot be explained monocausally but rather results from multiple evolutionary mechanisms. The relative frequencies of host-linked evolution, adaptive radiation, recombination, and lateral gene transfer (fig. 4) must therefore be quantified and their potential for reciprocal interaction be analyzed. Each single case may typically include components of each of those mechanisms. The most plausible explanation may challenge traditional views about the interactions between warm-blooded vertebrates and their colonizing PVes. Thus, we recommend the development and improvement of phylogenetic methods that detect and remove those parts of the data containing a high level of perturbing signals. Finally, the generation of phylogenetically representative full-genome PV sequences especially from nonhuman hosts is necessary in order to fill the numerous gaps in the current knowledge about PV evolution.

Supplementary Material

Supplementary figures S1 and S2 and tables S1–S4 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Literature Cited

Ahola H, Bergman P, Ström AC, Moreno-Lopéz J, Pettersson U. 1986. Organization and expression of the transforming region from the European elk papillomavirus (EEPV). Gene. 50:195–205.

Antonsson A, Hansson BG. 2002. Healthy skin of many animal species harbors papillomaviruses which are closely related to their human counterparts. J Virol. 76:12537–12542.

Arias-Pulido H, Peyton CL, Torrez-Martínez N, Anderson DN, Wheeler CM. 2005. Human papillomavirus type 18 variant lineages in United States populations characterized by sequence analysis of LCR-E6, E2, and L1 regions. Virology. 338:22–34.

Baurain D, Brinkmann H, Philippe H. 2007. Lack of resolution in the animal phylogeny: closely spaced cladogeneses or undetected systematic errors? Mol Biol Evol. 24:6–9.

Bedell MA, Hudson JB, Golub TR, Turyk ME, Hosken M, Wilbanks GD, Laimins LA. 1991. Amplification of human papillomavirus genomes in vitro is dependent on epithelial differentiation. J Virol. 65:2254–2260.

Bernard H-U. 2005. The clinical importance of the nomenclature, evolution and taxonomy of human papillomaviruses. J Clin Virol. 32:1–6.

Bernard H-U, Calleja-Macias IE, Dunn ST. 2006. Genome variation of human papillomavirus types: phylogenetic and medical implications. Int J Cancer. 118:1071–1076.

Bernard H-U, Chan S-Y, Manos MM, Ong C-K, Villa LL, Delius H, Peyton CL, Bauer HM, Wheeler CM. 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.
Bible JM, Mant C, Best JM, et al. (11 co-authors). 2000. Cervical lesions are associated with human papillomavirus type 16 intratypic variants that have high transcriptional activity and increased usage of common mammalian codons. J Gen Virol. 81:1517–1527.
Bininda-Emonds ORP. 2006. perlRat.pl v.1.0.9a. Distributed by the author. Jena (Germany): Friedrich-Schiller-Universität.
Bogaert L, Martens A, De Baere C, Gasthuys F. 2005. Detection of bovine papillomavirus DNA on the normal skin and in the habitual surroundings of horses with and without equine sarcoids. Res Vet Sci. 79:253–258.
Boshart M, Gissmann L, Ikenberg H, Kleinheinz A, Scheurlen W, zur Hausen H. 1984. A new type of papillomavirus DNA, its presence in genital cancer biopsies and in cell lines derived from cervical cancer. EMBO J. 3:1151–1157.
Bravo IG, Alonso Á. 2004. Mucosal human papillomaviruses encode four different E5 proteins whose chemistry and phylogeny correlate with malignant or benign growth. J Virol. 78:13613–13626.
Bravo IG, Alonso Á. 2007. Phylogeny and evolution of papillomaviruses based on the E1 and E2 proteins. Virus Genes. 34:249–262.
Bravo IG, Müller M. 2005. Codon usage in papillomavirus genes. Papillomavirus Rep. 16:63–72.
Campo MS, Moar MH, Laird HM, Jarrett WF. 1981. Molecular heterogeneity and lesion site specificity of cutaneous bovine papillomaviruses. Virology. 113:323–335.
Carney HC, England JJ, Hodgin EC, Whiteley HE, Adkison DL, Sundberg JP. 1990. Papillomavirus infection of aged Persian cats. J Vet Diagn Invest. 2:294–299.
Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol. 17:540–552.
Chambers G, Ellsmore VA, O'Brien PM, Reid SWJ, Love S, Campo MS, Nasir L. 2003. Association of bovine papillomavirus with the equine sarcoid. J Gen Virol. 84:1055–1062.
Chan S-Y, Bernard H-U, Ratterree M, Birkebak TA, Faras AJ, Ostrow RS. 1997. Genomic diversity and evolution of papillomaviruses in rhesus monkeys. J Virol. 71:4938–4943.
Chan S-Y, Delius H, Halpern AL, Bernard H-U. 1995. Analysis of genomic sequences of 95 papillomavirus types: uniting typing, phylogeny, and taxonomy. J Virol. 69:3074–3083.
Chan S-Y, Ho L, Ong C-K, et al. (11 co-authors). 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.
Chen EY, Howley PM, Levinson AD, Seeburg PH. 1982. The primary structure and genetic organization of the bovine papillomavirus type 1 genome. Nature. 299:529–534.
Chen Z, Schiffman M, Herrero R, DeSalle R, Burk R. 2007. Human papillomavirus (HPV) types 101 and 103 isolated from cervicovaginal cells lack an E6 open reading frame (ORF) and are related to gamma-papillomaviruses. Virology. 360:447–453.
Chen Z, Terai M, Fu L, Herrero R, DeSalle R, Burk RD. 2005. Diversifying selection in human papillomavirus type 16 lineages based on complete genome analyses. J Virol. 79:7014–7023.
Christensen ND, Cladel NM, Reed CA, Budgeon LR, Welsh PA, Patrick SD, Kreider JW. 1996. Laboratory production of infectious stocks of rabbit oral papillomavirus. J Gen Virol. 77:1793–1798.
Christensen ND, Cladel NM, Reed CA, Han R. 2000. Rabbit oral papillomavirus complete genome sequence and immunity following genital infection. Virology. 269:451–461.
Cole ST, Danos O. 1987. Nucleotide sequence and comparative analysis of the human papillomavirus type 18 genome. Phylogeny of papillomaviruses and repeated structure of the E6 and E7 gene products. J Mol Biol. 193:599–608.
Cunningham CW. 1997. Can three incongruence tests predict when data should be combined? Mol Biol Evol. 14:733–740.
Danos O, Katinka M, Yaniv M. 1982. Human papillomavirus 1a complete DNA sequence: a novel type of genome organization among papovaviridae. EMBO J. 1:231–236.
Delius H, Hofmann B. 1994. Primer-directed sequencing of human papillomavirus types. Curr Top Microbiol Immunol. 186:13–31.
de Villiers E-M, Fauquet C, Broker TR, Bernard H-U, zur Hausen H. 2004. Classification of papillomaviruses. Virology. 324:17–27.
Delsuc F, Brinkmann H, Philippe H. 2005. Phylogenomics and the reconstruction of the tree of life. Nat Rev Genet. 6:361–375.
Dimmic MW, Rest JS, Mindell DP, Goldstein RA. 2002. rtREV: an amino acid substitution matrix for inference of retrovirus and reverse transcriptase phylogeny. J Mol Evol. 55:65–73.
Doorbar J. 2005. The papillomavirus life cycle. J Clin Virol. 32:7–15.
Drummond AJ, Ho SY, Phillips MJ, Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol. 4:e88.
Drummond AJ, Nicholls GK, Rodrigo AG, Solomonc W. 2002. Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. Genetics. 161:1307–1320.
Drummond AJ, Rambaut A. 2003. BEAST 1.3. Bayesian evolutionary analysis sampling trees. Oxford: Department of Zoology, University of Oxford.
Egawa K. 2005. Eccrine-centred distribution of human papillomavirus 63 infection in the epidermis of the plantar skin. Br J Dermatol. 152:993–996.
Egawa K, Delius H, Matsukura T, Kawashima M, de Villiers E-M. 1993. Two novel types of human papillomavirus, HPV 63 and HPV 65: comparisons of their clinical and histological features and DNA sequences to other HPV types. Virology. 194:789–799.
Farmer AD, Calef SE, Millman K, Myers GL. 1995. The human papillomavirus database. J Biomed Sci. 2:90–104.
Farris JS, Källersjö M, Kluge AG, Bult C. 1994. Testing significance of incongruence. Cladistics. 10:315–319.
Favre M, Jibard N, Orth G. 1982. Restriction mapping and physical characterization of the cottontail rabbit papillomavirus genome in transplantable VX2 and VX7 domestic rabbit carcinomas. Virology. 119:298–309.
Favre M, Kremsdorf D, Jabłońska S, Obalek S, Pehau-Arnaudet G, Croissant O, Orth G. 1990. Two new human papillomavirus types (HPV54 and 55) characterized from genital tumours illustrate the plurality of genital HPVs. Int J Cancer. 45:40–46.
Favre M, Obalek S, Jabłońska S, Orth G. 1989a. Human papillomavirus type 49, a type isolated from flat warts of renal transplant patients. J Virol. 63:4909–4909.
Favre M, Obalek S, Jabłońska S, Orth G. 1989b. Human papillomavirus (HPV) type 50, a type associated with epidermodysplasia verruciformis (EV) and only weakly related to other EV-specific HPVs. J Virol. 63:4910.
Favre M, Orth G, Croissant O, Yaniv M. 1975. Human papillomavirus DNA: physical map. Proc Natl Acad Sci USA. 72:4810–4814.
Fislova T, Kostolansky F. 2005. The factors of virulence of influenza a virus. Acta Virol. 49:147–157.

Forslund O, Antonsson A, Higgins G, Ly H, Delius H, Hunziker A, de Villiers E-M. 2003. Nucleotide sequence and phylogenetic classification of candidate human papilloma virus type 92. Virology. 312:255–260.

Forslund O, Antonsson A, Nordin P, Stenquist B, Hansson BG. 1999. A broad range of human papillomavirus types detected with a general PCR method suitable for analysis of cutaneous tumours and normal skin. J Gen Virol. 80:2437–2443.

García-Vallvé S, Alonso Á Bravo IG. 2005. Papillomaviruses: different genes have different histories. Trends Microbiol. 13:514–521.

Giri I, Danos O, Yaniv M. 1985. Genomic structure of the cottontail rabbit (Shope) papillomavirus. Proc Natl Acad Sci USA. 82:1580–1584.

Gissmann L, zur Hausen H. 1980. Partial characterization of viral DNA from human genital warts (*Condylomata acuminata*). Int J Cancer. 25:605–609.

Gottschling M, Köhler A, Stockfleth E, Nindl I. 2007. Phylogenetic analysis of beta-papillomaviruses as inferred from nucleotide and amino acid sequence data. Mol Phylogenet Evol. 42:213–222.

Grimmel M, de Villiers E-M, Neumann C, Pawlita M, zur Hausen H. 1988. Characterization of a new human papillomavirus (HPV 41) from disseminated warts and detection of its DNA in some skin carcinomas. Int J Cancer. 41:5–9.

Groff DE, Lancaster WD. 1985. Molecular cloning and nucleotide sequence of deer papillomavirus. J Virol. 56:85–91.

Halpern AL. 2000. Comparison of papillomavirus and immunodeficiency virus evolutionary patterns in the context of a papillomavirus vaccine. J Clin Virol. 19:43–56.

Hasegawa M, Kishino H, Yano T. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol. 22:160–174.

Heilman CA, Law MF, Israel MA, Howley PM. 1980. Cloning of human papilloma virus genomic DNAs and analysis of homologous polynucleotide sequences. J Virol. 36:395–407.

Hernández Fernández M, Vrba ES. 2005. A complete estimate of the phylogenetic relationships in Ruminantia: a dated species-level supertree of the extant ruminants. Biol Rev Camb Philos Soc. 80:269–302.

Hirsch-Behnam A, Delius H, de Villiers E-M. 1990. A comparative sequence analysis of two human papillomavirus (HPV) types 2a and 57. Virus Res. 18:81–97.

Hirt L, Hirsch-Behnam A, de Villiers E-M. 1991. Nucleotide sequence of human papillomavirus (HPV) type 41: an unusual HPV type without a typical E2 binding site consensus sequence. Virus Res. 18:179–189.

Ho L, Chan S-Y, Burk RD, et al. (12 co-authors). 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.

Isegawa N, Ohta M, Shirasawa H, Tokita H, Yamaura A, Simizu B. 1995. Nucleotide sequence of a canine oral papillomavirus containing a long noncoding region. Int J Oncol. 7:155–159.

Iwasaki T, Maeda H, Kameyama Y, Moriyama M, Kanai S, Kurata T. 1997. Presence of a novel hamster oral papillomavirus in dysplastic lesions of hamster lingual mucosa induced by application of dimethylbenzanthracene and excisional wounding: molecular cloning and complete nucleotide sequence. J Gen Virol. 78:1087–1093.

Jackson ME, Pennie WD, McCaffery RE, Smith KT, Grindlay GJ, Campo MS. 1991. The B subgroup bovine papillomaviruses lack an identifiable E6 open reading frame. Mol Carcinog. 4:382–387.

Jarrett WF, Campo MS, O'Neil BW, Laird HM, Coggins LW. 1984. A novel bovine papillomavirus (BPV-6) causing true epithelial papillomas of the mammary gland skin: a member of a proposed new BPV subgroup. Virology. 136: 255–264.

Jarrett WF, Murphy J, O'Neil BW, Laird HM. 1978. Virus-induced papillomas of the alimentary tract of cattle. Int J Cancer. 22:323–328.

Jeffroy O, Brinkmann H, Delsuc F, Philippe H. 2006. Phylogenomics: the beginning of incongruence? Trends Genet. 22:225–231.

Kaye D, Pringle CR. 2005. Avian influenza viruses and their implication for human health. Clin Infect Dis. 40:108–112.

Kloster BE, Manias DA, Ostrow RS, Shaver MK, McPherson SW, Rangen SR, Uno H, Faras AJ. 1988. Molecular cloning and characterization of the DNA of two papillomaviruses from monkeys. Virology. 166:30–40.

Kremsdorf D, Favre M, Jabłońska S, Obalek S, Rueda LA, Lutzner MA, Blanchet-Bardon C, Van Voorst Vader PC, Orth G. 1984. Molecular cloning and characterization of the genomes of nine newly recognized human papillomavirus types associated with Epidermodysplasia verruciformis. J Virol. 52:1013–1018.

Kremsdorf D, Jabłońska S, Favre M, Orth G. 1982. Biochemical characterization of two types of human papillomaviruses associated with Epidermodysplasia verruciformis. J Virol. 43:436–447.

Lancaster WD, Olson C. 1978. Demonstration of two distinct classes of bovine papilloma virus. Virology. 89:372–379.

Lancaster WD, Olson C, Meinke W. 1977. Bovine papilloma virus: presence of virus-specific DNA sequences in naturally occurring equine tumors. Proc Natl Acad Sci USA. 74:524–528.

Lancaster WD, Sundberg JP. 1982. Characterization of papillomaviruses isolated from cutaneous fibromas of white-tailed deer and mule deer. Virology. 123:212–216.

Maddison WP. 1997. Gene trees in species trees. Syst Biol. 46:523–536.

Maddison WP, Knowles LL. 2006. Inferring phylogeny despite incomplete lineage sorting. Syst Biol. 55:21–30.

Mahy BW, Brown CC. 2000. Emerging zoonoses: crossing the species barrier. Rev Sci Tech. 19:33–40.

Matsukura T, Iwasaki T, Kawashima M. 1992. Molecular cloning of a novel human papillomavirus (type 60) from a plantar cyst with characteristic pathological changes. Virology. 190: 561–564.

Moreno-Lopéz J, Ahola H, Eriksson A, Bergman P, Pettersson U. 1987. Reindeer papillomavirus transforming properties correlate with a highly conserved E5 region. J Virol. 61:3394–3400.

Moreno-Lopéz J, Pettersson U, Dinter Z, Philipson L. 1981. Characterization of a papilloma virus from the European elk (EEPV). Virology. 112:589–595.

Mossadegh N, Gissmann L, Müller M, Zentgraf H, Alonso Á, Tomakidi P. 2004. Codon optimization of the human papillomavirus 11 (HPV 11) L1 gene leads to increased gene expression and formation of virus-like particles in mammalian epithelial cells. Virology. 326:57–66.

Müller H, Gissmann L. 1978. *Mastomys natalensis* papilloma virus (MnPV), the causative agent of epithelial proliferations: characterization of the virus particle. J Gen Virol. 41: 315–323.

Müller M, Kelly G, Fiedler M, Gissmann L. 1989. Human papillomavirus type 48. J Virol. 63:4907–4908.

Münger K, Howley PM. 2002. Human papillomavirus immortalization and transformation functions. Virus Res. 89:213–228.

Muñoz N, Bosch FX, de Sanjose S, Herrero R, Castellsague X, Shah KV, Snijders PJLM, Meijer CJF. 2003. Epidemiologic classification of human papillomavirus types associated with cervical cancer. N Engl J Med. 348:518–527.

Myers G, Bernard H-U, Delius H. 1994. Human papillomaviruses 1994. A compilation and analysis of nucleic and amino acid sequences. Los Alamos (NM): Los Alamos National Laboratory.
Myers G, Lu H, Calef C, Leitner T. 1996. Heterogeneity of papillomaviruses. Semin Cancer Biol. 7:349–358.
Narechania A, Chen Z, DeSalle R, Burk RD. 2005. Phylogenetic incongruence among oncogenic genital alpha human papillomaviruses. J Virol. 79:15503–15510.
Nindl I, Gottschling M, Stockfleth E. Forthcoming (2007). Human papillomaviruses and non-melanoma skin cancer: basic virology and clinical manifestations. Dis Markers.
Nixon KC. 1999. The parsimony ratchet, a new method for rapid parsimony analysis. Cladistics. 15:407–414.
Nunn CL. 2004. Parasites and the evolutionary diversification of primate clades. Am Nat. 164(Suppl 5):S90–S103.
O'Banion MK, Jacobson ER, Sundberg JP. 1992. Molecular cloning and partial characterization of a parrot papillomavirus. Intervirology. 33:91–96.
O'Banion MK, Reichmann ME, Sundberg JP. 1986. Cloning and characterization of an equine cutaneous papillomavirus. Virology. 152:100–109.
Ogawa T, Tomita Y, Okada M, Shinozaki K, Kubonoya H, Kaiho I, Shirasawa H. 2004. Broad-spectrum detection of papillomaviruses in bovine teat papillomas and healthy teat skin. J Gen Virol. 85:2191–2197.
Ong C-K, Chan S-Y, Campo MS, et al. (11 co-authors). 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.
Ong C-K, Nee S, Rambaut A, Bernard H-U, Harvey PH. 1997. Elucidating the population histories and transmission dynamics of papillomaviruses using phylogenetic trees. J Mol Evol. 44:199–206.
Orth G, Favre M, Croissant O. 1977. Characterization of a new type of human papillomavirus that causes skin warts. J Virol. 24:108–120.
Osterhaus ADME, Ellens DJ, Horzinek MC. 1977. Identification and characterization of a papillomavirus from birds (Fringillidae). Intervirology. 8:351–359.
Ostrow RS, Bender M, Niimura M, Seki T, Kawashima M, Pass F, Faras AJ. 1982. Human papillomavirus DNA in cutaneous primary and metastasized squamous cell carcinomas from patients with Epidermodysplasia verruciformis. Proc Natl Acad Sci USA. 79:1634–1638.
Ostrow RS, LaBresh KV, Faras AJ. 1991. Characterization of the complete RhPV 1 genomic sequence and an integration locus from a metastatic tumor. Virology. 181:424–429.
Otten N, von Tscharner C, Lazary S, Antczak DF, Gerber H. 1993. DNA of bovine papillomavirus type 1 and 2 in equine sarcoids: PCR detection and direct sequencing. Arch Virol. 132:121–131.
Patel KR, Smith KT, Campo MS. 1987. The nucleotide sequence and genome organization of bovine papillomavirus type 4. J Gen Virol. 68:2117–2128.
Pfister H. 2003. Chapter 8: human papillomavirus and skin cancer. J Natl Cancer Inst Monogr. 31:52–56.
Pfister H, Fink B, Thomas C. 1981. Extrachromosomal bovine papillomavirus type 1 DNA in hamster fibromas and fibrosarcomas. Virology. 115:414–418.
Pfister H, Linz U, Gissmann L, Huchthausen B, Hoffmann D, zur Hausen H. 1979. Partial characterization of a new type of bovine papilloma viruses. Virology. 96:1–8.
Philippe H, Zhou Y, Brinkmann H, Rodrigue N, Delsuc F. 2005. Heterotachy and long-branch attraction in phylogenetics. BMC Evol Biol. 5:50.
Posada D, Crandall KA. 1998. Modeltest: testing the model of DNA substitution [Application Note]. Bioinformatics. 14:817–818.
Prado JC, Calleja-Macias IE, Bernard H-U, et al. (19 co-authors). 2005. Worldwide genomic diversity of the human papillomaviruses-53, 56, and 66, a group of high-risk HPVs unrelated to HPV-16 and HPV-18. Virology. 340:95–104.
Purvis A. 1995. A composite estimate of primate phylogeny. Philos Trans R Soc Lond B Biol Sci. 348:405–421.
Rambaut A. 2001. Se-Al. Sequence alignment program v2.0a72. Oxford.
Rector A, Bossart GD, Ghim SJ, Sundberg JP, Jenson AB, Van Ranst M. 2004. Characterization of a novel close-to-root papillomavirus from a Florida manatee by using multiply primed rolling-circle amplification: *Trichechus manatus latirostris* papillomavirus type 1. J Virol. 78:12698–12702.
Rector A, Mostmans S, Van Doorslaer K, McKnight CA, Maes RK, Wise AG, Kiupel M, Van Ranst M. 2006. Genetic characterization of the first chiropteran papillomavirus, isolated from a basosquamous carcinoma in an Egyptian fruit bat: the *Rousettus aegyptiacus* papillomavirus type 1. Vet Microbiol. 117:267–275.
Rector A, Tachezy R, Van Doorslaer K, MacNamara T, Burk RD, Sundberg JP, Van Ranst M. 2005. Isolation and cloning of a papillomavirus from a North American porcupine by using multiply primed rolling-circle amplification: the *Erethizon dorsatum* papillomavirus type 1. Virology. 331:449–456.
Rector A, Van Doorslaer K, Bertelsen M, Barker IK, Olberg RA, Lemey P, Sundberg JP, Van Ranst M. 2005. Isolation and cloning of the raccoon (*Procyon lotor*) papillomavirus type 1 by using degenerate papillomavirus-specific primers. J Gen Virol. 86:2029–2033.
Rehtanz M, Ghim S-J, Rector A, Van Ranst M, Fair PA, Bossart GD, Jenson AB. 2006. Isolation and characterization of the first American bottlenose dolphin papillomavirus: *Tursiops truncatus* papillomavirus type 2. J Gen Virol. 87:3559–3565.
Salmon J, Nonnenmacher M, Cazé S, Flamant P, Croissant O, Orth G, Breitburd F. 2000. Variation in the nucleotide sequence of cottontail rabbit papillomavirus a and b subtypes affects wart regression and malignant transformation and level of viral replication in domestic rabbits. J Virol. 74:10766–10777.
Salmon J, Ramoz N, Cassonnet P, Orth G, Breitburd F. 1997. A cottontail rabbit papillomavirus strain (CRPVb) with strikingly divergent E6 and E7 oncoproteins: an insight in the evolution of papillomaviruses. Virology. 235:228–234.
Schiffman M, Herrero R, DeSalle R, et al. (18 co-authors). 2005. The carcinogenicity of human papillomavirus types reflects viral evolution. Virology. 337:76–84.
Schwarz E, Dürst M, Demankowski C, Lattermann O, Zech R, Wolfsperger E, Suhai S, zur Hausen H. 1983. DNA sequence and genome organization of genital human papillomavirus type 6b. EMBO J. 2:2341–2348.
Seedorf K, Krämmer G, Dürst M, Suhai S, Röwekamp WG. 1985. Human papillomavirus type 16 DNA sequence. Virology. 145:181–185.
Stamatakis A. 2006a. Phylogenetic models of rate heterogeneity: a high performance computing perspective. IPDPS 2006. Proceedings of 20th IEEE/ACM International Parallel and Distributed Processing Symposium; 2006 April 25–29; Rhodos (Greece): IEEE/ACM
Stamatakis A. 2006b. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 22:2688–2690.
Stamatakis A, Ludwig T, Meier H. 2005. RAxML-III: a fast program for maximum likelihood-based inference of large phylogenetic trees. Bioinformatics. 21:456–463.
Swofford DL. 2002. PAUP*: phylogenetic analysis using parsimony (*and other methods). Version 4. Sunderland (MA): Sinauer Associates.

Tachezy R, Rector A, Havelkova M, Wollants E, Fiten P, Opdenakker G, Jenson AB, Sundberg JP, Van Ranst M. 2002. Avian papillomaviruses: the parrot *Psittacus erithacus* papillomavirus (PePV) genome has a unique organization of the early protein region and is phylogenetically related to the chaffinch papillomavirus. BMC Microbiol. 2:19.

Tan CH, Tachezy R, Van Ranst M, Chan S-Y, Bernard H-U, Burk RD. 1994. The *Mastomys natalensis* papillomavirus: nucleotide sequence, genome organization, and phylogenetic relationship of a rodent papillomavirus involved in tumorigenesis of cutaneous epithelia. Virology. 198: 534–541.

Terai M, Burk RD. 2002. *Felis domesticus* papillomavirus, isolated from a skin lesion, is related to canine oral papillomavirus and contains a 1.3 kb non-coding region between the E2 and L2 open reading frames. J Gen Virol. 83:2303–2307.

Terai M, DeSalle R, Burk RD. 2002. Lack of canonical E6 and E7 open reading frames in bird papillomaviruses: *Fringilla coelebs* papillomavirus and *Psittacus erithacus timneh* papillomavirus. J Virol. 76:10020–10023.

Thomas M, Boiron M, Levy JP, Tanzer J, Bernard J. 1964. In vitro transformation of mice cells by bovine papilloma virus. Nature. 202:709–710.

Tobler K, Favrot C, Nespeca G, Ackermann M. 2006. Detection of the prototype of a potential novel genus in the family Papillomaviridae in association with canine Epidermodysplasia verruciformis. J Gen Virol. 87:3551–3557.

Trenfield K, Spradbrow PB, Vanselow B. 1985. Sequences of papillomavirus DNA in equine sarcoids. Equine Vet J. 17:449–452.

Van Bressem MF, Cassonet P, Rector A, Desaintes C, van Waerebeek K, Alfaro-Shigeto J, van Ranst M, Orth G. Forthcoming. Genital warts in Burmeister's porpoises: characterization of Phocoena spinipinnis papillomavirus type 1 (PsPV-1) and evidence for a second, distantly related PsPV1. J Gen Virol.

Van Doorslaer K, Rector A, Vos P, Van Ranst M. 2006. Genetic characterization of the *Capra hircus* papillomavirus: a novel close-to-root artiodactyl papillomavirus. Virus Res. 118: 164–169.

Van Ranst M, Fuse A, Fiten P, Beuken E, Pfister H, Burk RD, Opdenakker G. 1992. Human papillomavirus type 13 and pygmy chimpanzee papillomavirus type 1: comparison of the genome organizations. Virology. 190:587–596.

Van Ranst M, Fuse A, Sobis H, De Meurichy W, Syrjänen SM, Billiau A, Opdenakker G. 1991. A papillomavirus related to HPV type 13 in oral focal epithelial hyperplasia in the pygmy chimpanzee. J Oral Pathol Med. 20:325–331.

Van Ranst M, Kaplan JB, Burk RD. 1992. Phylogenetic classification of human papillomaviruses: correlation with clinical manifestations. J Gen Virol. 73:2653–2660.

Van Ranst M, Kaplan JB, Sundberg JP, Burk RD, Gibbs A, Calisher CH, García-Arenal F. 1995. Molecular evolution of the human papillomaviruses. In: Gibbs AJ, Calisher CH, García-Arenal F, editors. Molecular basis of virus evolution. Cambridge: Cambridge University Press. p. 455–476.

Varsani A, van der Walt E, Heath L, Rybicki EP, Williamson AL, Martin DP. 2006. Evidence of ancient papillomavirus recombination. J Gen Virol. 87:2527–2531.

Weiss RA. 2003. Cross-species infections. Curr Top Microbiol Immunol. 278:47–71.

Whelan S, Goldman N. 2001. A general empirical model of protein evolution derived from multiple protein families using a Maximum-Likelihood approach. Mol Biol Evol. 18: 691–699.

WHO Expert Committee. 1982. Bacterial and viral zoonoses. Report of a WHO Expert Committee with the participation of FAO. World Health Organ Tech Rep Ser. 682:1–146.

Yamada T, Manos MM, Peto J, Greer CE, Muñoz N, Bosch FX, Wheeler CM. 1997. Human papillomavirus type 16 sequence variation in cervical cancers: a worldwide perspective. J Virol. 71:2463–2472.

Yang ZH. 1996. Among-site rate variation and its impact on phylogenetic analyses. Trends Ecol Evol. 11:367–372.

Yuan H, Ghim S-J, Newsome J, Apolinario T, Olcese V, Martin M, Delius H, Felsburg P, Jenson B, Schlegel R. Forthcoming. An epidermotropic canine papillomavirus with malignant potential contains an E5 gene and establishes a unique genus. Virology. 359:28–36.

Zachow KR, Ostrow RS, Faras AJ. 1987. Nucleotide sequence and genome organization of human papillomavirus type 5. Virology. 158:251–254.

Zhao KN, Liu WJ, Frazer IH. 2003. Codon usage bias and A+T content variation in human papillomavirus genomes. Virus Res. 98:95–104.

Zhou J, Liu WJ, Peng SW, Sun XY, Frazer IH. 1999. Papillomavirus capsid protein expression level depends on the match between codon usage and tRNA availability. J Virol. 73:4972–4982.

zur Hausen H. 2000. Papillomaviruses causing cancer: evasion from host-cell control in early events in carcinogenesis. J Natl Cancer Inst. 92:690–698.

Aoife McLysaght, Associate Editor

Accepted February 28, 2007