Virology 378 (2008) 151–161

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

Genomic characterization of novel dolphin papillomaviruses provides indications for recombination within the *Papillomaviridae*

Annabel Rector [a,*], Hans Stevens [a], Géraldine Lacave [b,c], Philippe Lemey [a,d], Sara Mostmans [a], Ana Salbany [c], Melissa Vos [a], Koenraad Van Doorslaer [a], Shin-Je Ghim [e], Manuela Rehtanz [e,f], Gregory D. Bossart [f,g], A. Bennett Jenson [e], Marc Van Ranst [a]

[a] *Laboratory of Clinical Virology, Department of Microbiology and Immunology, Rega Institute for Medical Research, University of Leuven, Minderbroedersstraat 10, 3000 Leuven, Belgium*
[b] *Marine Mammal Veterinary Services, Daverlostraat 186, 8310 Brugge — Assebroek, Belgium*
[c] *Mundo Aquatico Zoomarine, Albufeira, Portugal*
[d] *Department of Zoology, University of Oxford, Oxford OX1 3PS, UK*
[e] *The Brown Cancer Center, University of Louisville, Louisville, KY 40202, USA*
[f] *Center for Coastal Research–Marine Mammal Research and Conservation Program, Harbor Branch Oceanographic Institution, Fort Pierce, FL 34946, USA*
[g] *Florida Atlantic University, Boca Raton, FL 33431, USA*

ARTICLE INFO

Article history:
Received 27 March 2008
Returned to author for revision
29 April 2008
Accepted 16 May 2008
Available online 24 June 2008

Keywords:
Papillomaviridae
Cetacea
Bottlenose dolphin (*Tursiops truncatus*)
Condylomata
Evolution
Phylogeny
Bootscanning
Recombination
Selection

ABSTRACT

Phylogenetic analysis of novel dolphin (*Tursiops truncatus*) papillomavirus sequences, TtPV1, -2, and -3, indicates that the early and late protein coding regions of their genomes differ in evolutionary history. Sliding window bootscan analysis showed a significant change in phylogenetic clustering, in which the grouped sequences of TtPV1 and -3 move from a cluster with the *Phocoena spinipinnis* PsPV1 in the early region to a cluster with TtPV2 in the late region. This provides indications for a possible recombination event near the end of E2/beginning of L2. A second possible recombination site could be located near the end of L1, in the upstream regulatory region. Selection analysis by using maximum likelihood models of codon substitutions ruled out the possibility of intense selective pressure, acting asymmetrically on the viral genomes, as an alternative explanation for the observed difference in evolutionary history between the early and late genomic regions of these cetacean papillomaviruses.

Introduction

The *Papillomaviridae* are a large family of epitheliotropic viruses that can cause benign and malignant proliferations of the stratified squamous epithelium in various vertebrate species. A wide genetic diversity of different human papillomavirus (HPV) types has been described, and is associated with a broad range of genotype-specific clinical conditions, ranging from genital and skin warts to invasive cervical carcinoma and skin cancer (Van Ranst et al., 1992). The limited number of non-human papillomaviruses (PVs) that has been characterized to date covers a broad range of host species, including mostly domestic and wild mammals but also two bird species (Sundberg et al., 1997; Sundberg, 1987; Tachezy et al., 2002). Taken together with the numerous partial sequences of putative novel non-human PVs that have been reported (Antonsson and Hansson, 2002; Chan et al., 1997; Ogawa et al., 2004), these data suggest that every vertebrate species could carry a set of species-specific PVs.

Papillomatous lesions including gastric papillomas and cutaneous and lingual papillomatosis, with PV infection as a suspected cause based on clinical, histopathological, transmission electron microscopic and immunohistochemical findings, have been described in several members of the order Cetacea (Bossart et al., 1996, 2005; De Guise et al., 1994; Geraci et al., 1987; Van Bressem et al., 1996, 1999). Furthermore, genital papillomatosis has been reported in a number of cetacean species, such as the sperm whale, the Blainville's beaked whale, the killer whale, the Atlantic white-sided dolphin, the long-beaked common dolphin, the Dusky dolphin, Burmeister's porpoise, and the bottlenose dolphin (Bossart et al., 2005; Flom et al., 1980; Geraci et al., 1987; Greenwood et al., 1974; Lambertsen et al., 1987; Van Bressem et al., 1996). Whereas cutaneous papillomatosis has been described in many orders of placental mammals, cetaceans represent the only non-primate mammalian order in which genital PV infection

* Corresponding author. Fax: +32 16 332131.
E-mail address: annabel.rector@uz.kuleuven.be (A. Rector).

0042-6822/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2008.05.020

has been confirmed. In some cetacean species genital warts are highly prevalent (e.g. up to 66% in Dusky dolphins) (Van Bressem et al., 1996). In a recent survey, a high occurrence of related benign and malignant orogenital neoplastic epithelial lesions was found in bottlenose dolphins, indicating that these could represent an emerging infectious disease (Bossart et al., 2005). The same study provided arguments for an infectious etiology of these tumors, with an orogenital transmission route (Bossart et al., 2005). Although these reports indicate that PV infection might be widespread in cetaceans, PV etiology had until recently only been unequivocally demonstrated in the case of genital lesions in the Burmeister's porpoise (*Phocoena spinipinnis*), caused by the *Phocoena spinipinnis* PV (PsPV1, GenBank accession number NC_003348) (Van Bressem et al., 2007). It therefore remains to be determined whether genital papillomatosis within single species of cetaceans is caused by a single PV type, or whether multiple types can infect the same species, as is the case in humans.

Papillomaviruses use the high-fidelity DNA replication machinery of their eukaryotic host cells. Together with their slow replication, which is linked to the division of infected epithelial cells, this results in an evolutionary rate of only 1.95×10^{-8} nucleotide substitutions per site per year (Rector et al., 2007). The apparent absence of inter-type recombination further contributes to the high degree of conservation among PVs. Recently however, data indicating that recombination could have occurred in the evolutionary history of PVs have accumulated (Angulo and Carvajal-Rodriguez, 2007; Bravo and Alonso, 2007; Gottschling et al., 2007b; Narechania et al., 2005; Varsani et al., 2006), but no proven examples of recombination between established PV types have been reported so far.

We have isolated the genomes of three distinct PV types from genital mucosal lesions in bottlenose dolphins (*Tursiops truncatus*). Phylogenetic comparison of these dolphin PV types with PsPV1 indicates that the different parts of their genomes have different evolutionary histories, and provides a strong indication that recombination could have occurred between different (ancestral) types of the *Papillomaviridae*.

Results

TtPV1, TtPV2 and TtPV3 complete genomic sequence

The complete genomes of three distinct *Tursiops truncatus* PVs have been isolated: TtPV1 and TtPV3, isolated from a single condylomatous lesion on the penis of a captive bottlenose dolphin from the Mundo Aquatico Zoomarine in Portugal (described in this paper), and TtPV2, which was isolated from a genital condyloma of a free-ranging bottlenose dolphin from Charleston Harbor, South Carolina, USA (described elsewhere) (Rehtanz et al., 2006). The complete nucleotide sequence of TtPV1 counts 8089 base pairs (bp) and has a GC content of 42.1%, TtPV2 counts 7866 bp with a 46.0% GC content, and TtPV3 counts 7915 bp with 44.2% CG.

Like all other PVs that have been characterized to date, TtPV1 and -3 have all their ORFs on the same coding strand of their dsDNA. The location of the ORFs and the size of the corresponding protein sequences, with comparison to the corresponding ORF data from TtPV2 and PsPV1, are summarized in a table that is provided as Supplementary file 1. The dolphin PV genomes contain the classical PV ORFs E6, E1, and E2 in their early region, and the major and minor capsid proteins L1 and L2 in their late region. An additional E5 ORF, which is also present in the E2–L2 region of the genome of the mucosal HPVs belonging to the genus *Alphapapillomavirus* (α-PV) and the cutaneous (fibro-) PVs of the genus *Deltapapillomavirus* (Garcia-Vallve et al., 2005), as well as in TtPV2, could be identified in TtPV1 and -3.

Typical features of the early region of the PVs of Cetacea

The E6 ORFs of TtPV1 and TtPV3 contain two zinc-binding domains with the sequence C-X-X-C-X_{29}-C-X-X-C, separated by 36 amino acids. This motif is identical to those found in the other cetacean PVs, TtPV2 and PsPV1 (Rehtanz et al., 2006; Van Bressem et al., 2007). Similarly to TtPV2 and PsPV1, the genomes of TtPV1 and TtPV3 encode E6 proteins which are larger than those generally found in PVs: 224 and 220 amino acids for TtPV1 and TtPV3 respectively, in comparison to 206 amino acids in TtPV2, 211 in PsPV1, and about 150 in most other PVs (Howley and Lowy, 2001). The predicted E6 proteins of TtPV1 and -3 do not contain a PDZ domain binding motif in their extreme C-terminus, which would enable E6 to bind to proteins containing a PDZ domain and direct them to proteolysis. Such a motif, with the consensus sequence X-S/T-X-V/L (in which X represents any amino acid), is present in the very C-terminus of the high risk but not of the low risk genital HPVs (Gardiol et al., 1999; Jing et al., 2007; Kiyono et al., 1997). A PDZ domain binding motif has also been identified in the E6 of TtPV2 E6 (Rehtanz et al., 2006), but not in the PsPV1 E6 (Van Bressem et al., 2007).

In the early region of the PV genome, the E6 ORF is typically followed by an ORF coding for the papillomaviral E7 protein, which contains a zinc-binding domain and in some cases a pRb binding domain (containing the L-X-C-X-E motif, critical for binding of pRb). In TtPV1 and TtPV3 however, no E7 ORFs can be identified. A similar genomic organization, with the absence of a classical E7, has previously been reported for PsPV1 (Van Bressem et al., 2007) and TtPV2 (Rehtanz et al., 2006). Whereas the PsPV1 genome still has the capacity to encode a small peptide (26 amino acids) harbouring the pRb binding motif, through a small ORF overlapping with the 5′ end of the PsPV1 E1 (Narechania et al., 2004), this is not the case in TtPV1 and -3, nor in TtPV2 (Rehtanz et al., 2006). In TtPV3 however, an L-X-C-X-E motif was identified in the E1 ORF, but since this was located upstream of the E1 start codon it is not likely that this motif is effectively translated. The absence of a canonical E7 ORF is highly uncommon among PVs, and so far seems to be a unique feature of the PVs of cetaceans.

Sequence similarity to other PVs and classification of dolphin PVs

The similarities between the dolphin PVs mutually, and with the porpoise PsPV1 (which is the only other cetacean PV characterized to date), the manatee TmPV1 (representing a sea mammal PV isolated from a member of the Sirenia, an order of aquatic mammals which is only distantly related to the Cetacea order), the human low risk mucosal HPV11, and the high risk mucosal HPV16, were investigated for the different ORFs by pairwise nucleotide and amino acid sequence alignments. For all ORFs, the highest similarities were noted between TtPV1 and TtPV3, and the similarities to TmPV1, HPV11, and HPV16 were lower and in the same range (Table 1). For the ORFs in the early region, TtPV1 and TtPV3 show high similarities to PsPV1. In contrast, the late region of TtPV1 and TtPV3 displayed higher similarity to TtPV2 as to PsPV1. This was even more apparent when the dolphin and porpoise PV genomes were analyzed in Maizel–Lenk dot matrix plots (Fig. 1). The plot illustrates the genome-wide high similarity between TtPV1 and TtPV3. Between TtPV1 and TtPV2, a high degree of similarity is still visible across the late protein region, especially in the L1 ORF, but not in the early region. With PsPV1 on the other hand, TtPV1 shared higher similarity across the early region compared to the late region.

For classification purposes, only the sequence of the L1 ORF is taken into account. According to the current definition, PV types belong to the same genus when they share at least 60% nucleotide sequence identity across the entire L1 ORF (de Villiers et al., 2004). This places TtPV1, TtPV2, and TtPV3 in the same genus, with identities between 66 and 80% (Table 1). Since the similarities of the dolphin PVs to PsPV1, which is the next closest related PV type, are below 60%, TtPV1, -2, and -3 cannot be assigned to the genus *Omikronpapillomavirus* (ο-PV), of which PsPV1 is to date the only member, but have to be classified in a novel, yet unnamed genus.

Table 1
Percentage nucleotide (amino acid) sequence similarity of the open reading frames of TtPV1, -2, and -3 mutually and to corresponding ORFs of PsPV1, TmPV1, HPV11 and HPV16

ORF		TtPV2	TtPV3	PsPV1		TmPV1	HPV11		HPV16
E6	TtPV1	40 (30)	70 (61)	58 (53)		27 (16)	31 (24)		29 (19)
	TtPV2		38 (30)	41 (32)		27 (20)	30 (23)		33 (23)
	TtPV3			55 (49)		26 (17)	35 (28)		29 (20)
E1	TtPV1	55 (46)	69 (63)	68 (62)		49 (40)	54 (43)		54 (42)
	TtPV2		54 (43)	55 (45)		48 (39)	50 (40)		49 (40)
	TtPV3			68 (62)		50 (40)	53 (43)		54 (43)
E2	TtPV1	43 (32)	59 (51)	55 (47)		39 (26)	40 (28)		40 (26)
	TtPV2		42 (32)	44 (32)		38 (30)	45 (32)		44 (29)
	TtPV3			56 (49)		38 (28)	40 (30)		38 (29)
E5	TtPV1	11 (8)	38 (30)	30 (22)[a]	44 (30)[b]	no E5	18 (12)[a]	17 (8)[b]	17 (8)
	TtPV2		14 (7)	12 (11)[a]	9 (7)[b]	no E5	28 (22)[a]	16 (8)[b]	23 (16)
	TtPV3			29 (20)[a]	32 (30)[b]	no E5	20 (12)[a]	26 (14)[b]	22 (10)
L2	TtPV1	55 (52)	79 (80)	40 (30)		38 (31)	38 (32)		36 (27)
	TtPV2		55 (52)	39 (30)		37 (30)	35 (29)		37 (27)
	TtPV3			39 (30)		37 (29)	37 (30)		36 (27)
L1	TtPV1	67 (69)	80 (80)	55 (49)		50 (44)	51 (42)		49 (41)
	TtPV2		66 (69)	54 (48)		50 (44)	50 (42)		48 (42)
	TtPV3			55 (48)		49 (44)	51 (44)		50 (41)

[a,b]For PsPV1 and HPV11, the percentages similarity were calculated for the two E5 sequences (a and b) that are available in Genbank (under accession numbers NC_003348 for PsPV1 and M14119 for HPV11), with "a" representing the most upstream (most 5′) E5 ORF and "b" representing the more downstream E5 ORF.

Phylogenetic analysis

Separate maximum likelihood (ML) phylogenetic trees were constructed for the E1 and L1 ORFs, based on multiple sequence alignments including all PV type species and the novel dolphin PVs. A list of all sequences included in the phylogenetic analysis is provided in Supplementary file 2. For both the L1 and the E1 sequence alignment, phylogenetic reconstruction using the neighbor joining (NJ) method resulted in the same clustering as was obtained by the ML analysis. In the L1 (Fig. 2A) as well as in the E1 tree (Fig. 2B), the PVs cluster according to the previously defined genera, and PV types belonging to the same genus are found in monophyletic branches. In the L1 tree, the dolphin PVs TtPV1, -2, and -3 also cluster in a monophyletic branch, excluding all other PV types, and this clustering is supported by a high bootstrap value (100%). This supports the classification of TtPV1, -2, and -3 in one, separate, genus. The most closely related non-dolphin PV is the porpoise PsPV1, although bootstrap support for this clustering was much lower (65%). The ML phylogenetic analysis based on the E1 sequence alignment however resulted in a different clustering of the dolphin PVs as compared to the L1 tree, since in the E1 tree TtPV1 and -3 are more closely related to PsPV1 as to TtPV2, which branches off before PsPV1. So although TtPV1 and TtPV3 are consistently most closely related in both the E1 and the L1 ML tree, the monophyletic clustering of all TtPVs in the L1 tree is not reproduced in the tree of the E1 sequences.

Recombination analysis with bootscan and Simplot

The previous analyses indicated that for the dolphin PVs, the evolutionary history of the early region might be different from that of the late region, which could be the result of a recombination event. To further investigate this possibility, we performed a bootscan analysis using the Simplot program. Since the PV genomes of the different genera are too divergent to construct a reliable complete genome alignment, we constructed a concatenated alignment of the unambiguously alignable parts of the E6, E1, E2, L2, and L1 ORFs, including all PV type species that contain an E6. For the analysis, all PV types were grouped according to their genus, and the grouped sequences of TtPV1 and -3, which are closely related across the entire genome, were

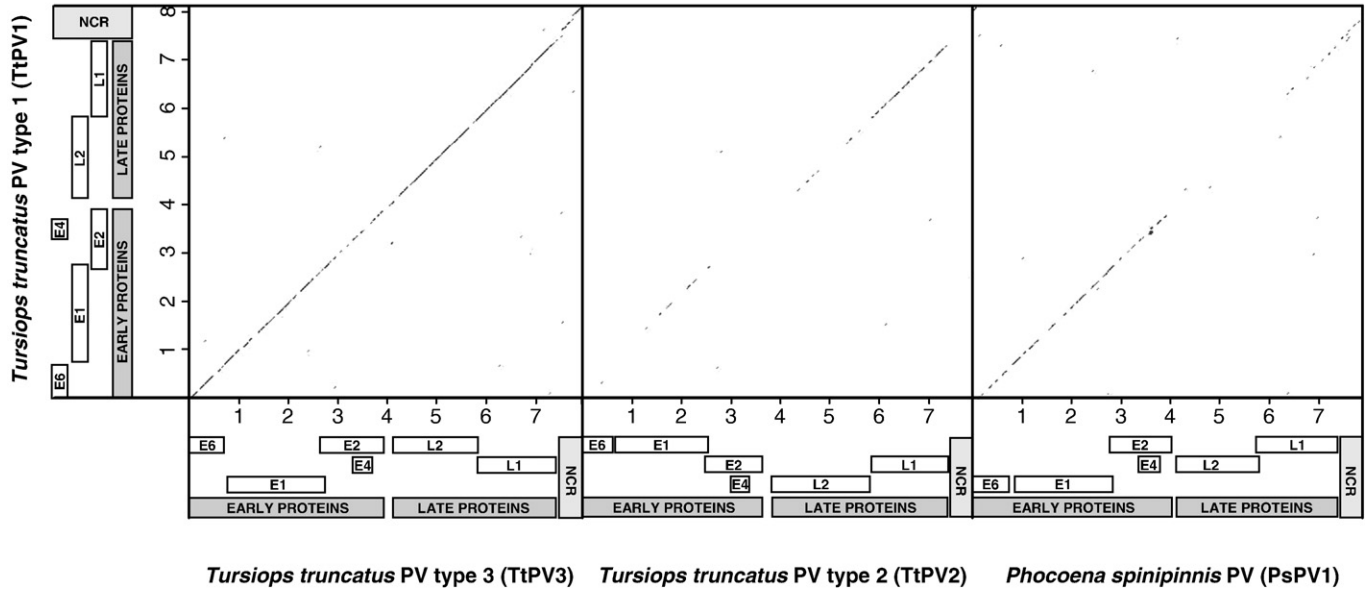

Fig. 1. Maizel–Lenk dot matrix plots aligning the complete genome of TtPV1 with that of TtPV3, TtPV2, and PsPV1 (window size: 19 nucleotides, mismatch allowance: 3/19).

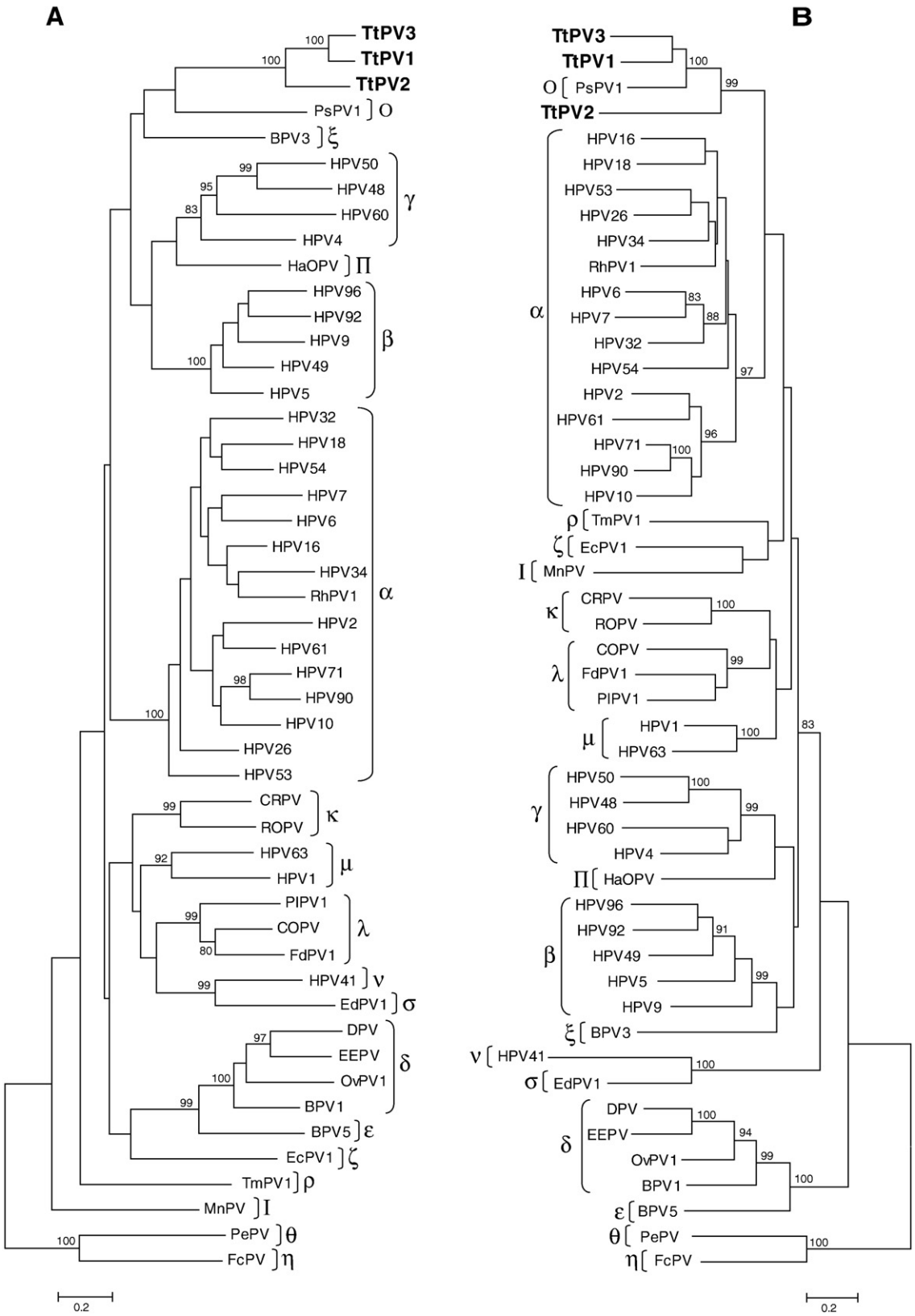

Fig. 2. Maximum likelihood phylogenetic tree based on an alignment of the L1 (A) and E1 (B) sequences of the dolphin TtPVs and all PV type species. The names of the different papillomavirus genera are indicated by their Greek letter. Numbers at internal nodes represent the percentage of bootstrap probabilities as determined by the neighbor-joining method for 10 000 iterations. Only bootstrap values greater than 80% are shown. The scale bar indicates the genetic distance (nucleotide substitutions per site).

used as a query for bootscanning (Fig. 3A). In the first part of the alignment the TtPV1/TtPV3 group clusters with the porpoise PsPV1, with high bootstrap support. As the window moves down the alignment however, the phylogenetic branching pattern changes, and in the distal part the TtPV1/TtPV3 group clusters with TtPV2, again with high bootstrap support. No monophyletic clustering with

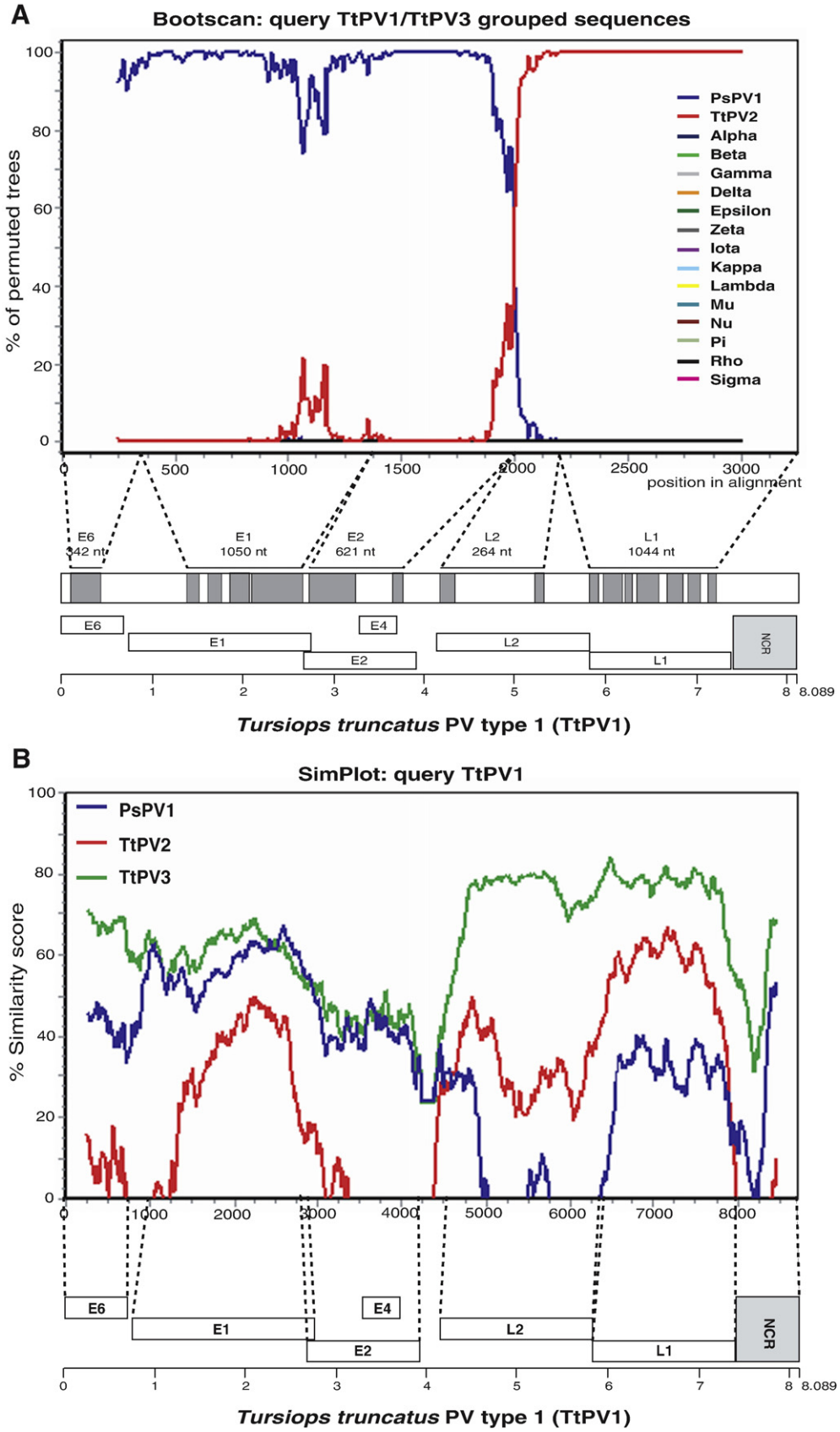

Fig. 3. Recombination analysis of the TtPV genomes. (A). Bootscan analysis performed on a concatenated nucleotide sequence alignment of the E6, E1, E2, L2, and L1 ORFs of all PV type species and TtPV1, -2, and -3. The positions in the concatenated alignment are correlated to the corresponding positions in the TtPV1 genome. The PV type species were grouped according to their PV genera, which are indicated by their Greek letter in the legend (except for PsPV1, representing the sole member of the *Omikronpapillomavirus*, which is indicated with its type name). The grouped sequences of TtPV1 and TtPV3 were used as query in the bootscan analysis, which was performed with a window size of 500 nucleotides, and a step size of 10 nucleotides. (B) Simplot analysis performed on a complete genome alignment of the bottlenose dolphin PVs and PsPV1. The TtPV1 sequence was used as query, and the analysis was performed with a window size of 500 nucleotides and step size of 10 nucleotides. The positions in the alignment are correlated to the positions on the TtPV1 genome.

Table 2
The d_N/d_S ratios and presence of positively selected sites in the E6, E1, E2, L2 and L1 ORFs, calculated with different models for codon substitution, with the log likelihood of each model

ORF	Model	d_N/d_S	Positively selected sites[a]	Log likelihood[b]	Parameters
E6	M0	0,18853		−16162,193701	ω=0.18853
	M1	0,91239		−16309,864583	p_0=0.0876139; p_1=0.9123861
	M2	0,73129	40 (0.7758); 61 (0.7821)	−16010,195456	p_0=0.3366245; p_1=0.6470543; p_2=0.0163212; ω_2=2.09659249
	M3	0,21611	No rate class with ω>1	**−15622,257732**	p_0=0.1050977; p_1=0.2465489; p_2=0.6483534; ω_0=0.00182095; ω_1=0.09486448; ω_2=0.29694989
	M7	0,22207		−15634,713534	beta p=0.632511; beta q=2.21574
	M8	0,22349	No sites with ω>1	−15634,715399	beta p=0.631678; beta q=2.19473; p=1; ω=1.01924
E1	M0	0,10600		−37963,643354	ω=0.10600
	M1	0,88571		−40276,413739	p_0=0.1142922; p_1=0.8857078
	M2	0,30085	No rate class with ω>1	−37498,627748	p_0=0.7955631; p_1=0.1013131; p_2=0.1031237; ω_2=1.00000000
	M3	0,12362	No rate class with ω>1	−36320,714846	p_0=0.2289199; p_1=0.4026392; p_2=0.3684408; ω_0=0.00739304; ω_1=0.07298550; ω_2=0.25115855
	M7	0,12230		−36242,165691	beta p=0.640624; beta q=4.59758
	M8	0,12233	No rate class with ω>1	**−36242,165385**	beta p=0.641106; beta q=4.59986; p=1; ω=1
E2	M0	0,17994		−26618,723071	ω=0.17994
	M1	0,95667		−27510,732911	p_0=0.0433321; p_1=0.9566679
	M2	0,46175	No rate class with ω>1	−26088,062621	p_0=0.6543610; p_1=0.1626076; p_2=0.1830314; ω_2=1.00000000
	M3	0,21973	No rate class with ω>1	−25592,776975	p_0=0.3044109; p_1=0.3991928; p_2=0.2963962; ω_0=0.03818005; ω_1=0.18386436; ω_2=0.45449993
	M7	0,19701		−25543,785018	beta p=0.727674; beta q=2.96588
	M8	0,20214	No rate class with ω>1	**−25543,021928**	beta p=0.744973; beta q=3.11547; p=0.988649; ω=1
L2	M0	0,07223		−9449,886371	ω=0.07223
	M1	0,87537		−10357,972738	p_0=0.1246294; p_1=0.8753706
	M2	0,10542	No rate class with ω>1	−9418,074129	p_0=0.9659257; p_1=0.0187240; p_2=0.0153503; ω_2=1.00000000
	M3	0,07966	No rate class with ω>1	−9147,494274	p_0=0.2284724; p_1=0.5002116; p_2=0.2713160; ω_0=0.00329732; ω_1=0.06040459; ω_2=0.17946503
	M7	0,12476		−9149,350886	beta p=0.583745; beta q=4.09531
	M8	0,08004	No sites with ω>1	**−9143,179022**	beta p=0.719436; beta q=8.26903; p=1; ω=1
L1	M0	0,06380		−33097,377130	ω=0.063780
	M1	0,84728		−36862,552675	p_0=0.1527165; p_1=0.8472835
	M2	0,10346	No rate class with ω>1	−32985,200165	p_0=0.9594510; p_1=0.0230246; p_2=0.0175244; ω_2=1.00000
	M3	0,07067	No rate class with ω>1	−31939,281406	p_0=0.3666243; p_1=0.4370988; p_2=0.1962769; ω_0=0.00877;ω_1=0.06722; ω_2=0.19399
	M7	0,08000		**−31895,335271**	beta p=0.636414; beta q=7.31902
	M8	0,08002	No sites with ω>1	−31895,336071	beta p=0.635787; beta q=7.30917; p=1; ω=5.46683
Concatenated	M0	0,11554		−126962,982179	ω=0.11554
	M1	0,88895		−133881,000524	p_0=0.1110458; p_1=0.8889542
	M2	0,34387	No rate class with ω>1	−124791,364532	p_0=0.7481161; p_1=0.1257516; p_2=0.1261322; ω_2=1.00000000
	M3	0,13859	No rate class with ω>1	−121192,957283	p_0=0.2681900; p_1=0.4111098; p_2=0.3207003; ω_0=0.01120225;ω_1=0.09418531; ω_2=0.30205532
	M7	0,14512		−120899,557449	beta p=0.623811; beta q=3.67476
	M8	0,14392	40 (0.9443); 61 (0.9687); 663 (0.6295)	**−120895,320717**	beta p=0.629741; beta q=3.83253; p=0.99675; ω=1

[a]For the models allowing for sites with ω>1, the presence of such a rate class is indicated. If a rate class with ω>1 exists, the codons positions that have a posterior probability of more than 0.5 (value indicated between the parentheses) to belong to the ω>1 rate class (= positively selected sites) according to this model are indicated.
[b]The most likely model (highest log likelihood value) is indicated in bold.

any of the other PV genera could be detected along the entire alignment (the % permuted trees for clustering with any of the other PV genera does not reach a detectable level). The significant change in phylogenetic relationship along the alignment indicates a possible recombination event. By correlating the positions in the input alignment to the locations in the TtPV1 genome, the position of this possible recombination could be estimated to be near the end of the E2 — beginning of L2 ORF: this is where the clustering suddenly changes, and the TtPV1/TtPV3 group moves from a clade with PsPV1 to a clade with TtPV2. A second recombination in the circular genome then possibly also occurred in the NCR sequence, which is not included in this concatenated alignment. We tried to locate this putative recombination more accurately by performing a Simplot analysis on a complete genome alignment of the dolphin and porpoise PVs (Fig. 3B). The position where the similarity with PsPV1 decreases and that with TtPV2 increases, coincides with a small noncoding sequence between the E2 and L2 ORF of TtPV1. Near the end of the alignment the similarity with TtPV2 decreases again, below that with PsPV1, and this second putative recombination point is located around the end of L1 and the beginning of the noncoding region. We carefully examined the alignment at these places, but the sequences are too divergent to allow a more accurate determination of these probable recombination points in the alignment.

Selective pressure analysis

The sequence dataset used for bootscanning (ORF alignments of the dolphin TtPVs and all PV type species that contain an E6) was investigated for the presence of positively selected sites with a ML method in which probabilistic models of codon substitution, that allow for variable non-synonymous/synonymous substitution rate ratios (ω=d_N/d_S) among sites, were applied to the data. A likelihood ratio test (LRT) was used to determine whether allowing for sites with ω of >1 significantly improved the fit of the model to the data. If the d_N/d_S ratio for any site class was above 1, the Bayes theorem was used to calculate the posterior probability that each site, given its data, belonged to such a site class. The d_N/d_S ratios for the E6, E1, E2, L2 and L1 ORFs, as well as for the concatenated alignment of all these ORFs, calculated with the different models for codon substitution, together with the log likelihood values of those models, are indicated in Table 2. In the E6 ORF, the discrete model M3 showed the highest log likelihood (significantly better than the nested neutral model M0 according to a LRT), but no rate class with ω>1 (positively selected sites) was estimated from the data under this model. For E1 as well as E2, the beta and ω model M8 had the highest log likelihood, although not significantly better than the beta model M7 according to LRT, and again no rate class with sites under positive selection (ω>1) was estimated. In

the L2 ORF, model M8 was significantly better than its nested model M7, and although a rate class with $\omega > 1$ was determined no sites were attributed to this rate class with a posterior probability of more than 0.5. For the L1 ORF, the beta model M7, which does not allow for positively selected sites, best fitted the data, although not significantly better than the more complex model M8. No sites with a d_N/d_S ratio of >1 (posterior probability cutoff value 0.5) could be identified in any of the separate ORF alignments with the models that best fitted the data. For the concatenated alignment however, the beta and ω model (M8) best fitted the data (significantly better than model M7 according to LRT), and in the alignment of 1107 codon sites in total, 3 sites under positive selection were identified under this model, with a posterior probability of at least 0.5. These sites were at codon positions 40 (posterior probability 0.944), 61 (posterior probability 0.969) and 663 (posterior probability 0.629) of the concatenated alignment, all three of which were located in the early region of the PV genome.

Although no major selective pressure was found, the bootscan analysis for detection of recombination was repeated on the concatenated alignment from which the sites under positive selection were removed, resulting in the same bootscan results as with the original alignment (data not shown). This was done to rule out the possibility that the presence of positively selected sites in the early but not in the late genomic region, thus asymmetric selective pressure, could account for the observed difference in evolutionary clustering between the early and late genomic regions of the cetacean PVs.

Discussion

Distinct PV types in genital condylomas of dolphins

Three different PV types were discovered in genital lesions of bottlenose dolphins by examining biopsy material from two different animals. Together with the high prevalence of genital papillomatosis in dolphins this indicates that dolphins, like humans, could harbour a multitude of different genital PV types. Some mucosal HPV types are known to cause genital lesions as well as infections of the respiratory tract and of the oral mucosa (Syrjänen and Syrjänen, 2000). In Atlantic bottlenose dolphins, lingual papillomas and squamous cell carcinomas have been described, with pathological evidence suggesting a viral etiology and the possibility of orogenital transmission (Bossart et al., 2005). Whether the genital PV types that we have characterized now are also responsible for (some of the) oral lesions in bottlenose dolphins still needs to be determined.

In humans, co-infection of the mucosa with different HPV types is frequent (21.3% of cases) (Mendez et al., 2005). The fact that the genomes of TtPV1 and TtPV3 were amplified from a single biopsy specimen of a penile condylomatous lesion proves that also in dolphins multiple PV types can co-infect the genital mucosa.

A number of HPV types have the potential to induce neoplastic lesions of the human genital tract, most importantly cervical cancer. There are major concerns that also in dolphins, PV infection of the genital mucosa could result in malignant lesions. The simultaneous presence of benign oral papillomas and oral squamous cell carcinomas in bottlenose dolphins is highly suggestive of such a progressive pathologic process involving malignant transformation (Bossart et al., 2005). Whether there are certain PV types associated with an increased risk of progression of lesions to malignancy in dolphins, as is the case in their human counterparts which can be classified in high and low risk mucosal HPVs, remains to be determined.

Recombination in Papillomaviridae

A difference in evolutionary history between the early and late protein coding regions of the cetacean PVs was consistently detected by a number of different phylogenetic methods. The percentages pairwise identity (Table 1) and the Maizel–Lenk dot matrix plots (Fig. 1) show that the dolphin PVs TtPV1 and -3 are most closely related to each other and this close relationship is maintained along the entire genome. In the early region, PsPV1 is more similar to this TtPV1/TtPV3 group than TtPV2, whereas in the late region the similarities of TtPV2 to TtPV1 and -3 are higher than those of PsPV1. These findings are confirmed by the NJ and ML phylogenetic analyses (Fig. 2): TtPV1 and -3 are found in a monophyletic branch in both the phylogenetic tree based on the E1 sequences and the phylogenetic tree based on L1, PsPV1 is most closely related to this TtPV1/TtPV3 cluster in the E1 tree, and in the L1 tree TtPV2 is more closely related to the TtPV1/TtPV3 cluster than PsPV1. Sliding window bootscan analysis on an E6–E1–E2–L2–L1 concatenated alignment of the dolphin PVs and all PV type species that contain an E6 in their genome (Fig. 3A) shows that this change in phylogenetic branching pattern occurs near the end of the E2 or the beginning of the L2 ORF. This position could be narrowed down to a small noncoding sequence between the E2 and L2 ORFs by performing a sliding window Simplot analysis on a complete genome alignment of the cetacean PV genomes (Fig. 3B). This moreover indicated that near the end of the alignment (around the end of L1 or in the noncoding region) the phylogenetic clustering reverted to the situation found in the early region. From these analyses, we deduce the hypothesis that recombination has occurred between ancestors of the extant cetacean PVs, in which the early region of an ancestor of PsPV1 recombined with the late region of an ancestor of TtPV2, and TtPV1 and -3 are descendants of this recombinant virus.

The construction of phylogenetic trees, and hence also the bootscan analysis for detection of possible sites of recombination in sequences, is based on the assumption of stochastic evolution of the sequences under analysis. Since positive selection in nucleotide sequences is a non-stochastic event, the presence of positively selected sites can result in phylogenetic incongruence for different parts of the sequences, possibly creating a false positive signal for recombination. If nucleotide positions under positive selection are present, these should therefore be removed from the alignment, and phylogenetic analyses should be performed on a corrected dataset for which the assumption of stochastic evolution is valid. We therefore examined our dataset for the presence of codon positions under positive selection. Although only three sites in the concatenated alignment of 1107 codon sites were possibly positively selected (Table 2), we have repeated the bootscan and Simplot analyses on an alignment from which these sites were stripped. We obtained the same results when analyzing this "corrected" alignment as with the original alignment, indicating that the difference in phylogenetic clustering between the early and late genomic regions of these cetacean PVs cannot be attributed to selective pressure acting asymmetrically on the viral genomes, but is most likely the consequence of an ancient recombination event.

Despite the large number of complete PV genomes that have been sequenced to date, no proven recombinations between extant PV types have been reported, and PV evolution is thought to proceed through slow accumulation of point mutations, insertions, and deletions. Recently however, evidence for a non-monophyletic mode of evolution within at least the genus α-PV has accumulated. Qualitative indications for a different evolutionary pattern of the early and late genes of these PVs were provided by Bravo and Alonso (Bravo and Alonso, 2004), and statistical analysis of topology differences in the phylogenetic trees inferred from different ORF alignments of the α-PVs provided evidence for a phylogenetic incongruence at the putative high risk node in the evolutionary tree (Narechania et al., 2005). These findings support the occurrence of one or more early recombination events in the evolutionary history of this PV lineage, although alternative explanations such as asymmetric genome convergence driven by intense selection and/or ecological niche changes could not be ruled out. The possibility of recombination within the genus α-PV was later confirmed with independent phylogeny-based statistical tests, which also provided phylogenetic support for three additional ancient recombinations in the predecessors of PsPV1, HPV41 and HPV75 (Varsani et al., 2006). Furthermore, examination of human α-PV sequences by using a

model-based composite likelihood method provided additional evidence that recombination may have occurred in this PV lineage (Angulo and Carvajal-Rodriguez, 2007). More generally, a recent large scale analysis of PV genomes belonging to different established genera showed that PV evolution is driven by multiple evolutionary mechanisms, including the rare occurrence of ancient recombination events (Gottschling et al., 2007b). The fact that no well-supported contradictions in the tree topologies of the early and late PV genes were found when examining the phylogeny of the genus *Betapapillomavirus*, the second largest PV genus with 26 distinct HPV types, also indicates that recombination would occur only exceptionally and concerted evolution of the early and late genes should be considered the rule in the PV evolutionary history (Gottschling et al., 2007a).

Our analyses of the cetacean PV sequences indicate that TtPV1 and -3 share a common ancestor that was generated through recombination of an ancestor of PsPV1, which delivered the early genes, and an ancestor of TtPV2, which provided the late genomic region to the recombinant genome. This is in accordance with our recent discovery of a new virus prototype, the Bandicoot papillomatosis carcinomatosis virus type 1 (BPCV1), in which genomic features of both the *Papillomaviridae* and *Polyomaviridae* are combined (Woolford et al., 2007). We suggested that this could represent the product of a recombination event between members of these two virus families, in which a PV provided the late region structural genes cassette, and a polyomavirus delivered the early region non-structural genes. Together with the data that we present here, this points towards an evolutionary mechanism in which the early and late genes cassettes of PV genomes are relatively independent entities that can be interchanged by recombination, albeit that such events are extremely rare.

Materials and methods

Clinical material

In 2002, condylomatous lesions were noticed in the genital area of two female bottlenose dolphins, living in captivity in the Mundo Aquatico Zoomarine at Albufeira, Portugal. Biopsies of these lesions showed histopathological features suggestive of PV infection. Also, immunohistochemical staining with polyclonal antibodies raised against bovine PV type 1 was positive. At the same facility, condyloma-like lesions were also identified in two male bottlenose dolphins. Lesion material from one of these penile condylomata was obtained by performing a voluntary biopsy procedure.

Amplification, cloning and sequencing of papillomaviral genomic DNA

The tissue biopsy material was finely minced with a scalpel and digested overnight at 56 °C in 500 µl digestion buffer (10 mM Tris, 0.5% SDS, pH 7.4) with 500 µg PCR-grade proteinase K. The digest was deproteinized twice by phenol/chloroform/isoamylalcohol extraction, followed by chloroform extraction, and DNA was recovered by ethanol precipitation, then air-dried and resuspended in 50 µl TE buffer (10 mM Tris–HCl, 1 mM EDTA, pH 8.0). For amplification of papillomaviral complete genomic DNA, multiply primed rolling-circle amplification (RCA) was performed on this extracted DNA by using the TempliPhi 100 Amplification kit (Amersham Biosciences) following a protocol which we have optimized for the amplification of papillomaviral episomal DNA (Rector et al., 2004b,a). One microliter of extracted DNA or water (negative control) was transferred into a 0.5 ml tube with 5 µl of TempliPhi sample buffer, containing exonuclease-protected random hexamers. The samples were denatured at 95 °C for 3 min, and afterwards placed on ice. A premix was prepared on ice by mixing for each sample 5 µl of TempliPhi reaction buffer, 0.2 µl of TempliPhi enzyme mix containing the ϕ29 DNA polymerase and exonuclease-protected random hexamers in 50% glycerol, and 0.2 µl of extra dNTPs (at a concentration of 25 mM of each dNTP). After mixing by vortexing, 5 µl of this premix was added to the cooled samples. The reactions were incubated overnight (approximately 16 h) at 30 °C. Afterwards, the reactions were placed on ice, subsequently heated to 65 °C for 10 min to inactivate the ϕ29 DNA polymerase, and stored at −20 °C awaiting further analysis. To investigate whether PV DNA was amplified, 2 µl of the RCA products was digested with 10 U of BamHI, EcoRI, HindIII, SalI, and XbaI. After digestion, the products were run on a 0.8% agarose gel to check for the presence of a DNA band consistent with full length PV DNA (~8 kb), or multiple bands with sizes adding up to this length. To clone the genomes of the *Tursiops truncatus* papillomavirus type 1 (TtPV1) and type 3 (TtPV3), 10 µl of the RCA product was digested with 100 U of restriction enzyme, and run on a 0.8% agarose gel, after which the DNA fragments were cut out from the gel, and extracted from the gel slices using GeneClean (BIO 101 Systems/Qbiogene). The complete TtPV1 genomic DNA was cloned into pUC18 via a unique BamHI restriction site, located in the L1 open reading frame (ORF) of the TtPV1 genome, whereas TtPV3 was cloned using EcoRI. In TtPV3, EcoRI sites are located in the E1 and the L2 ORF, and two EcoRI fragments (of approximately 5.2 and 2.5 kb respectively) were cloned separately in pUC18. After transformation of One Shot MAX Efficiency DH5α-T1R competent cells (Invitrogen), the bacteria were incubated for blue-white colony screening on agar plates containing X-gal, and white colonies were checked by restriction digestion of miniprep DNA. The EZ::TN™ <KAN-2> Insertion Kit (Epicentre, Landgraaf, The Netherlands) was used to sequence the cloned viral sequences. This kit uses the Tn5 transposase to randomly insert primer binding sites and a kanamycin resistance selection marker into target DNA in vitro. The reaction was performed according to the manufacturer's protocol. The reaction product was used to transform One Shot MAX Efficiency DH5α-T1R competent cells (Invitrogen). For each of the different cloned PV genomes or fragments, between 12 and 32 colonies were selected and the provided primers were used to sequence the insertion clones bidirectionally from primer binding sites at the 5′- and 3′-ends of the inserted transposon. The remaining gaps in the sequences were determined by primer-walking. Sequencing was performed on an ABI Prism 3100 Genetic Analyzer (Perkin-Elmer Applied Biosystems, Foster City, CA, USA) at the Rega Institute sequencing facility. Chromatrogram sequencing files were inspected with Chromas 2.2 (Technelysium, Helensvale, Australia), and contigs were compiled using the SeqMan software (DNASTAR, Madison, WI, USA). Preliminary sequence analysis indicated that a short partial sequence of the E1 ORF of TtPV3 was lost during cloning. To recover this sequence, conventional PCR with specific primers for the E1 of TtPV3 was carried out on the original RCA product, and the PCR product was sequenced with the same primers as used for PCR. The resulting additional sequence (276 bp) was integrated into the TtPV3 contig.

The cloning and sequencing of the TtPV2 genome (GenBank accession number NC_008184), isolated from a genital condyloma specimen collected from a free-ranging bottlenose dolphin in the coastal waters of Charleston Harbor, SC, USA, is described elsewhere (Rehtanz et al., 2006).

Sequence analysis

The ORF analysis was performed using the ORF Finder tool on the NCBI server (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). Similarity searches were performed using the NCBI Basic Local Alignment Search Tool (BLASTN 2.2.13) server on GenBank DNA database release 151.0 (Altschul et al., 1997). The molecular weight of the putative proteins was calculated using the ExPASy (Expert Protein Analysis System) Compute pI/Mw tool (http://au.expasy.org/tools/pi_tool.html). Pairwise sequence alignments were calculated using the DAMBE software package version 4.2.7 (Xia and Xie, 2001), in which pairwise alignments were performed at the amino acid level with the ClustalW program (Thompson et al., 1994), after which the corresponding nucleotide sequences were aligned accordingly. Maizel–Lenk dot matrix plots were used to investigate the genome-wide similarity of

the TtPV1 sequence to the sequences of TtPV2, TtPV3 and PsPV1. These were calculated with a window size of 19 nucleotides and a mismatch allowance of 3/19 via the Dot Matrix module on the Molecular Toolkit server of the Colorado State University (http://arbl.cvmbs.colostate.edu/molkit/dnadot/index.html).

Phylogenetic analysis

Phylogenetic analysis was performed on the nucleotide sequences of TtPV1, TtPV2, TtPV3, and 46 other PV types (type species of the different PV genera and species). The PV types included (with their GenBank accession numbers) are listed in Supplementary file 2. The different protein coding regions of these sequences were aligned at the amino acid level with ClustalW (Thompson et al., 1994) and translated back to nucleotide sequences by using the DAMBE software package version 4.2.7 (Xia and Xie, 2001). The alignments were then corrected manually in the GeneDoc Multiple Sequence Alignment Editor and Shading Utility software package version 2.6.002 (Nicholas et al., 1997). After removal of the sequences of PV types that do not contain an E6 ORF (BPV3, FcPV and PePV), the corrected alignments of the E6, E1, E2, L2, and L1 ORFs were concatenated, resulting in an alignment of a total of 3321 nucleotides or 1107 codon sites.

Based on these alignments, phylogenetic trees were constructed for the different ORFs by using the NJ method in MEGA version 3.1 (Kumar et al., 2004). For the E1 and L1 ORFs, phylogenetic reconstruction was also performed by using the ML method in the phylogenetic analysis using parsimony (PAUP) package version 4.0b 10 (http://paup.csit.fsu.edu) (Swofford, 1998). The GTR substitution model (with invariable sites I and taking into account rate heterogeneity over sites by using a discretized gamma distribution) was selected according to the Akaike Information Criterion in the Modeltest 3.066 PPC software package (http://darwin.uvigo.es/software/modeltest.html) (Posada and Crandall, 1998). With this model and using a NJ tree as starting tree, a branch swapping ML heuristic search implementing both nearest neighbor interchange and tree bisection reconnection was performed in PAUP.

Bootscanning

The possibility of recombination was investigated with the bootscanning method (Salminen et al., 1995) implemented in the Simplot software version 3.5.1 (http://sray.med.som.jhmi.edu/SCRoftware/simplot/) (Lole et al., 1999). Bootscanning uses a sliding window approach to detect recombination by performing a phylogenetic analysis with bootstrapping at each window position, and the bootscan value indicates the bootstrap support for the clade that includes the query sequence. Recombination may change the phylogenetic branching pattern, and can thus potentially move the query sequence to another clade. Probable recombination events are therefore indicated by significant changes in phylogenetic relationships along the examined alignment, reflected in changes in bootscan values. Since there is so far no application available to perform this analysis on amino acid sequences, and given the extensive nucleotide sequence diversity of the examined sequences, we could only perform the bootscan analysis on the concatenated nucleotide sequence alignment described above, which only includes the conserved parts of the different PV ORFs. Bootscanning was carried out with the Kimura 2 parameter model, and using a window size of 500 nucleotides and a step size of 10 nucleotides. The grouped sequences of TtPV1 and TtPV3 were applied as query, and the other PV sequences were grouped according to their genus.

Similarity analysis

The Simplot software version 3.5.1 was used to plot similarity versus position (Lole et al., 1999). Simplot analysis with a window size of 500 nucleotides and step size of 10 nucleotides was performed on a complete genome alignment of the TtPV1, TtPV2, TtPV3, and PsPV1 genomic sequences, constructed with the MUSCLE algorithm on the MUSCLE web server of the Berkeley Phylogenomics Group (http://phylogenomics.berkeley.edu/cgi-bin/muscle/input_muscle.py) (Edgar, 2004).

Selective pressure analysis

Natural selection acting at the amino acid level can be detected by comparison of relative fixation rates of non-synonymous (amino acid altering) substitutions and synonymous (silent) substitutions. The non-synonymous/synonymous rate ratio ($\omega = d_N/d_S$) is therefore an important indicator of selective pressure, with an ω value = 1 indicating neutral evolution, ω values of <1 indicating purifying or negative selection, and ω values of >1 indicating diversifying or positive selection. Since in many proteins a high proportion of amino acids are largely invariable, and most proteins are under overall purifying selection, the average ω value for all codon sites within a gene will in most cases be <1. Different amino acid sites in the gene can however be under different selective pressure, with different underlying ω values (Nielsen and Yang, 1998; Yang et al., 2000). The presence of positively selected sites in the different PV genes was investigated under probabilistic models of codon substitutions that allow for variable non-synonymous/synonymous rate ratios among sites. The likelihood ratio test (LRT) was used to determine whether allowing for sites with ω value of >1 significantly improved the fit of the model to the data. If the d_N/d_S ratio for any site class was above 1, the Bayes theorem was used to calculate the posterior probability for each site, given its data, to belong to this site class. Six codon substitution models where applied to the data: the one-ratio model M0 (allowing for one ω for all sites), the neutral model M1 (assuming a proportion of p_0 of negative or conserved sites with ω_0 of 0 and a proportion $p_1 = 1 - p_0$ of neutral sites with ω_1 of 1), the selection model M2 (allowing for an extra class of sites under diversifying selection with frequency $p_2 = 1 - p_0 - p_1$ and ω_2 of >1 estimated from the data, but not allowing for sites with $0 < \omega < 1$), the discrete model M3 (with a discrete distribution of three site classes with proportions p_0, p_1, and p_2 and ω_0, ω_1, and ω_2, estimated from the data), the beta model M7 (incorporating a beta distribution with parameters p and q to account for variable ω_0 in the interval between 0 and 1 among neutral or negatively selected sites), and the beta and ω model M8 (to which compared to M7 an extra component allowing for positively selected sites is added, with proportion p_2 and ω_2) (Yang et al., 2000). Nested models that were compared by LRT were model M1 with M2, M0 with M3, and M7 with M8. In case the more general model indicates the presence of sites with $\omega > 1$, the comparison constitutes an LRT of positive selection. For the LRT, twice the log likelihood difference between the two models was compared with a χ^2 distribution with a number of degrees of freedom equal to the difference in the number of parameters estimated by the two models (equaling 2 between M1 and M2, 4 between M0 and M3, and 2 between M7 and M8). All calculations were performed separately on the different ORF alignments and on the concatenated alignment of TtPV1, TtPV2, TtPV3, PsPV1 and all PV type species containing an E6 ORF. The analyses were performed using the Hyphy software package version 0.99 beta (Kosakovski Pond et al., 2005).

To investigate whether our selection analysis was sufficiently sensitive to detect positively selected sites in this dataset of highly divergent sequences, we performed a Monte Carlo simulation analysis. Sequences were simulated along a tree, with branch lengths inferred in codon substitutions units from the real data set, according to M3 using three classes of sites with proportions $p_0 = 0.6$, $p_1 = 0.3$ and $p_2 = 0.1$ and $\omega_0 = 0.1$, $\omega_1 = 0.6$ and $\omega_2 = 3.0$ with a length corresponding to the concatenated alignment. Twenty replicate data sets were subsequently investigated with the same methodology as used for the

true dataset. The fact that we were able to consistently retrieve accurate parameters for the three classes and hence detect positively selected sites in all simulated dataset illustrates the robustness of the methodology for detection of positive selection in our PV dataset (graphical representation of the results provided in Supplementary File 3).

Nucleotide sequence accession numbers

The nucleotide sequence data reported in this article were deposited in GenBank by using the NCBI BankIt v3.0 submission tool (http://www.ncbi.nlm.nih.gov/BankIt/) under the accession numbers EU240894 for TtPV1, and EU240895 for TtPV3.

Acknowledgments

We thank the colleagues of the laboratory of Clinical Virology, Rega Institute for Medical Research, University of Leuven, for their helpful comments and critical reading of the manuscript. This work was supported by the Flemish Fund for Scientific Research (Fonds voor Wetenschappelijk Onderzoek, FWO) grant G.0513.06 and by a post-doctoral fellowship of the Research Fund K.U.Leuven to Annabel Rector.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.virol.2008.05.020.

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Angulo, M., Carvajal-Rodriguez, A., 2007. Evidence of recombination within human alpha-papillomavirus. J. Virol. 4, 33.

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

Bossart, G.D., Cray, C., Solorzano, J.L., Decker, S.J., Cornell, L.H., Altman, N.H., 1996. Cutaneous papillomaviral-like papillomatosis in a killer whale (*Orcinus orca*). Marine Mamm. Sci. 12, 274–281.

Bossart, G.D., Ghim, S.J., Rehtanz, M., Goldstein, J.D., Varela, R.A., Ewing, R.Y., Fair, P.A., Lenzi, R., Joseph, B., Hicks, C., Schneider, L., McKinnie, C.J., Reif, J.S., Sanchez, R., Lopez, A., Novoa, S., Bernal, J., Goretti, M., Rodriguez, M., Defran, R.H., Jenson, A.B., 2005. Orogenital neoplasia in Atlantic bottlenose dolphins (*Tursiops truncatus*). Aquat. Mamm. 31, 473–480.

Bravo, I.G., Alonso, A., 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, I.G., Alonso, A., 2007. Phylogeny and evolution of papillomaviruses based on the E1 and E2 proteins. Virus Genes 34, 249–262.

Chan, S.Y., Bernard, H.U., Ratterree, M., Birkebak, T.A., Faras, A.J., Ostrow, R.S., 1997. Genomic diversity and evolution of papillomaviruses in rhesus monkeys. J. Virol. 71, 4938–4943.

De Guise, S., Lagace, A., Beland, P., 1994. Gastric papillomas in eight St. Lawrence beluga whales (*Delphinapterus leucas*). J. Vet. Diagn. Invest. 6, 385–388.

de Villiers, E.M., Fauquet, C., Broker, T.R., Bernard, H.U., zur Hausen, H., 2004. Classification of papillomaviruses. Virology 324, 17–27.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797.

Flom, J.O., Brown, R.J., Jones, R.E., Schonewald, J., 1980. Vaginal fibromas in a beaked whale, *Mesoplodon densirostris*. J. Wildl. Dis. 16, 99–102.

Garcia-Vallve, S., Alonso, A., Bravo, I.G., 2005. Papillomaviruses: different genes have different histories. Trends Microbiol. 13, 514–521.

Gardiol, D., Kuhne, C., Glaunsinger, B., Lee, S.S., Javier, R., Banks, L., 1999. Oncogenic human papillomavirus E6 proteins target the discs large tumour suppressor for proteasome-mediated degradation. Oncogene 18, 5487–5496.

Geraci, J.R., Palmer, N.C., St. Aubin, D.J., 1987. Tumors in cetaceans: analysis and new findings. Can. J. Fish. Aquat. Sci. 44, 1289–1300.

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

Gottschling, M., Stamatakis, A., Nindl, I., Stockfleth, E., Alonso, A., Bravo, I.G., 2007b. Multiple evolutionary mechanisms drive papillomavirus diversification. Mol. Biol. Evol. 24, 1242–1258.

Greenwood, A.G., Harrison, R.J., Whitting, H.W., 1974. Functional and pathological aspects of the skin of marine mammals. In: Harrison, R.J. (Ed.), Functional Anatomy of Marine Mammals. Academic Press, London, pp. 73–110.

Howley, P.M., Lowy, D.R., 2001. Papillomaviruses and their replication. In: Knipe, D.M., Howley, P.M. (Eds.), Fields Virology. Lippincott Williams & Wilkins, Philadelphia, pp. 2197–2229.

Jing, M., Bohl, J., Brimer, N., Kinter, M., Vande Pol, S.B., 2007. Degradation of tyrosine phosphatase PTPN3 (PTPH1) by association with oncogenic human papillomavirus E6 proteins. J. Virol. 81, 2231–2239.

Kiyono, T., Hiraiwa, A., Fujita, M., Hayashi, Y., Akiyama, T., Ishibashi, M., 1997. Binding of high-risk human papillomavirus E6 oncoproteins to the human homologue of the *Drosophila* discs large tumor suppressor protein. Proc. Natl. Acad. Sci. U. S. A. 94, 11612–11616.

Kosakovski Pond, S.L., Frost, S.D., Muse, S.V., 2005. HyPhy: hypothesis testing using phylogenies. Bioinformatics 21, 676–679.

Kumar, S., Tamura, K., Nei, M., 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief. Bioinform. 5, 150–163.

Lambertsen, R.H., Kohn, B.A., Sundberg, J.P., Buergelt, C.D., 1987. Genital papillomatosis in sperm whale bulls. J. Wildl. Dis. 23, 361–367.

Lole, K.S., Bollinger, R.C., Paranjape, R.S., Gadkari, D., Kulkarni, S.S., Novak, N.G., Ingersoll, R., Sheppard, H.W., Ray, S.C., 1999. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. J. Virol. 73, 152–160.

Mendez, F., Munoz, N., Posso, H., Molano, M., Moreno, V., van den Brule, A.J., Ronderos, M., Meijer, C., Munoz, A., 2005. Cervical coinfection with human papillomavirus (HPV) types and possible implications for the prevention of cervical cancer by HPV vaccines. J. Infect. Dis. 192, 1158–1165.

Narechania, A., Chen, Z., DeSalle, R., Burk, R.D., 2005. Phylogenetic incongruence among oncogenic genital alpha human papillomaviruses. J. Virol. 79, 15503–15510.

Narechania, A., Terai, M., Chen, Z., DeSalle, R., Burk, R.D., 2004. Lack of the canonical pRB-binding domain in the E7 ORF of artiodactyl papillomaviruses is associated with the development of fibropapillomas. J. Gen. Virol. 85, 1243–1250.

Nicholas, K.B., Nicholas, H.B., Deerfield, D.W., 1997. GeneDoc: analysis and visualization of genetic variation. Embnet News 4, 14.

Nielsen, R., Yang, Z., 1998. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. Genetics 148, 929–936.

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.

Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Rector, A., Bossart, G.D., Ghim, S.J., Sundberg, J.P., Jenson, A.B., Van Ranst, M., 2004a. 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., Lemey, P., Tachezy, R., Mostmans, S., Ghim, S.J., Van Doorslaer, K., Roelke, M., Bush, M., Montali, R.J., Joslin, J., Burk, R.D., Jenson, A.B., Sundberg, J.P., Shapiro, B., Van Ranst, M., 2007. Ancient papillomavirus-host co-speciation in Felidae. Genome Biol. 8, R57.

Rector, A., Tachezy, R., Van Ranst, M., 2004b. A sequence-independent strategy for detection and cloning of circular DNA virus genomes by using multiply primed rolling-circle amplification. J. Virol. 78, 4993–4998.

Rehtanz, M., Ghim, S.J., Rector, A., Van Ranst, M., Fair, P.A., Bossart, G.D., Jenson, A.B., 2006. Isolation and characterization of the first American bottlenose dolphin papillomavirus: *Tursiops truncatus* papillomavirus type 2. J. Gen. Virol. 87, 3559–3565.

Salminen, M.O., Carr, J.K., Burke, D.S., McCutchan, F.E., 1995. Identification of breakpoints in intergenotypic recombinants of HIV type 1 by bootscanning. AIDS Res. Hum. Retrovir. 11, 1423–1425.

Sundberg, J.P., 1987. Papillomavirus infections in animals. In: Syrjänen, K., Gissmann, L., Koss, L.G. (Eds.), Papillomaviruses and Human Disease. Springer-Verlag, Berlin, pp. 40–103.

Sundberg, J.P., Van Ranst, M., Burk, R.D., Jenson, A.B., 1997. The nonhuman (animal) papillomaviruses: host range, epitope conservation, and molecular diversity. In: Gross, G., von Krogh, G. (Eds.), Human papillomavirus infections in dermatovenereology. CRC Press, Boca Raton, pp. 47–68.

Swofford, D.L., 1998. PAUP* 4.0 — Phylogenetic Analysis Using Parsimony. Sinauer Associates, Sunderland.

Syrjänen, K.J., Syrjänen, S.M., 2000. Papillomavirus Infections in Human Pathology. John Wiley and Sons Ltd, Chichester, West Sussex, UK.

Tachezy, R., Rector, A., Havelkova, M., Wollants, E., Fiten, P., Opdenakker, G., Jenson, B., Sundberg, J., 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, 19–27.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Van Bressem, M.F., Cassonnet, P., Rector, A., Desaintes, C., Van Waerebeek, K., Alfaro-Shigueto, J., Van Ranst, M., Orth, G., 2007. Genital warts in Burmeister's porpoises: characterization of *Phocoena spinipinnis* papillomavirus type 1 (PsPV-1) and evidence for a second, distantly related PsPV. J. Gen. Virol. 88, 1928–1933.

Van Bressem, M.F., Kastelein, R.A., Flamant, P., Orth, G., 1999. Cutaneous papillomavirus infection in a harbour porpoise (*Phocoena phocoena*) from the North Sea. Vet. Rec. 144, 592–593.

Van Bressem, M.F., Van Waerebeek, K., Piérard, G.E., Desaintes, C., 1996. Genital and lingual warts in small cetaceans from coastal Peru. Dis. Aquat. Org. 26, 1–10.

Van Ranst, M., Kaplan, J.B., Burk, R.D., 1992. Phylogenetic classification of human papillomaviruses: correlation with clinical manifestations. J. Gen. Virol. 73, 2653–2660.
Varsani, A., van der Walt, E., Heath, L., Rybicki, E.P., Williamson, A.L., Martin, D.P., 2006. Evidence of ancient papillomavirus recombination. J. Gen. Virol. 87, 2527–2531.
Woolford, L., Rector, A., Van Ranst, M., Ducki, A., Bennett, M.D., Nicholls, P.K., Warren, K.S., Swan, R.A., Wilcox, G.E., O'Hara, A.J., 2007. A novel virus detected in papillomas and carcinomas of the endangered western barred bandicoot (*Perameles bougainville*) exhibits genomic features of both the *Papillomaviridae* and *Polyomaviridae*. J. Virol. 81, 13280–13290.
Xia, X., Xie, Z., 2001. DAMBE: software package for data analysis in molecular biology and evolution. J. Heredity 92, 371–373.
Yang, Z., Nielsen, R., Goldman, N., Pedersen, A.M., 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics 155, 431–449.