Virology 427 (2012) 189–197

Contents lists available at SciVerse ScienceDirect

Virology

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

Evidence of recombination and positive selection in cetacean papillomaviruses

Refugio Robles-Sikisaka [a,1], Rebecca Rivera [a], Hendrik H. Nollens [a,b,c], Judy St. Leger [c], Wendy N. Durden [d], Megan Stolen [d], Jennifer Burchell [a], James F.X. Wellehan Jr. [b,*]

[a] Hubbs-SeaWorld Research Institute, Center for Marine Veterinary Virology, 2595 Ingraham Street, San Diego, CA 92109, USA
[b] College of Veterinary Medicine, University of Florida, PO Box 110885, Gainesville, FL 32611, USA
[c] SeaWorld San Diego, 500 SeaWorld Drive, San Diego, CA 92109, USA
[d] Hubbs-SeaWorld Research Institute, 3830 South Highway A1A #4-181, Melbourne Beach, FL 32951, USA

ARTICLE INFO

Article history:
Received 26 September 2011
Returned to author for revision
21 November 2011
Accepted 26 January 2012
Available online 3 March 2012

Keywords:
Bottlenose dolphin
Cetacean
Papillomavirus
Genital papilloma
Virus evolution
Genetic diversity

ABSTRACT

Papillomaviruses (PVs) are small DNA viruses that have been associated with increased epithelial proliferation. Over one hundred PV types have been identified in humans; however, only three have been identified in bottlenose dolphins (*Tursiops truncatus*) to date. Using rolling circle amplification and degenerate PCR, we identified four novel PV genomes of bottlenose dolphins. TtPV4, TtPV5 and TtPV6 were identified in genital lesions while TtPV7 was identified in normal genital mucosa. Bayesian analysis of the full-length L1 genes found that TtPV4 and TtPV7 group within the *Upsilonpapillomavirus* genus while TtPV5 and TtPV6 group with *Omikronpapillomavirus*. However, analysis of the E1 gene did not distinguish these genera, implying that these genes may not share a common history, consistent with recombination. Recombination analyses identified several probable events. Signals of positive selection were found mostly in the E1 and E2 genes. Recombination and diversifying selection pressures constitute important driving forces of cetacean PV evolution.

Introduction

Papillomaviruses (PVs) constitute a large family of circular non-enveloped DNA viruses of approximately 8 kb that infect the skin and mucosal membranes of numerous amniote host species. The clinical significance of PVs has been widely recognized due to the potential of some types to induce neoplasia in their hosts, although they can also be detected in healthy epithelial tissues. The family *Papillomaviridae* is currently classified into 29 different genera (Bernard et al., 2010) and over one hundred PV types have been identified in humans. Although both oral and genital papillomatous lesions are common in cetaceans (Rehtanz et al., 2010; Van Bressem et al., 2009), only a limited number of PV genomes, from genital and esophageal lesions, have been sequenced and characterized. Three PVs have been isolated from bottlenose dolphins (*Tursiops truncatus*) (TtPV1, TtPV2 and TtPV3) (Rector et al., 2008; Rehtanz et al., 2006), one from the Burmeister's porpoise (*Phocoena spinnipinis*) (PsPV1) (Van Bressem et al., 2007), one from the short-beaked common dolphin (*Delphinus delphis*) (DdPV1), one from the Atlantic white-sided dolphin (*Lagenorhynchus acutus*) (TtPV3 variant), and three from the harbor porpoise (*Phocoena phocoena*) (PphPV1, PphPV2, and PphPV3) (Gottschling et al., 2011). According to their phylogenetic relationships, cetacean PVs have been classified into the genera *Omikronpapillomavirus* and *Upsilonpapillomavirus*. The clustering of these two cetacean genera in a monophyletic group indicates a shared common ancestry (Bernard et al., 2010).

The understanding of PV evolution has advanced on several fronts in recent years. Host-linked codivergence has been considered the most significant factor in PV diversification (Gottschling et al., 2007b). However, there is increasing evidence suggesting that various evolutionary mechanisms such as inter-specific transmission, adaptive radiation, recombination, and positive selection also play important roles in PV evolution (Angulo and Carvajal-Rodríguez, 2007; Carvajal-Rodríguez, 2008; Gottschling et al., 2007a,b; Gottschling et al., 2011). Although evidence of positive or diversifying selection has been found on human PVs (Carvajal-Rodríguez, 2008; Chen et al., 2005; DeFilippis et al., 2002), no significant signals of positive selection have been found in cetacean PV's to date (Rector et al., 2008).

Similarity of the L1 capsid protein has historically been used as a yardstick to determine levels of similarity and evolutionary relationships among different PVs (de Villiers et al., 2004). Nucleotide sequence similarities of the L1 protein less than 60% have indicated different genera, whereas species within a genus have been defined by nucleotide identities of 60 to 70% (Bernard et al., 2010; Howley and Lowy, 2007). In recent years, data has been presented suggesting

* Corresponding author.
E-mail addresses: refugio.robles1@gmail.com (R. Robles-Sikisaka), RRivera@hswri.org (R. Rivera), Hendrik.Nollens@SeaWorld.com (H.H. Nollens), Judy.St.Leger@SeaWorld.com (J. St. Leger), WNoke@hswri.org (W.N. Durden), MStolen@hswri.org (M. Stolen), JBurchell@hswri.org (J. Burchell), WellehanJ@ufl.edu (J.F.X. Wellehan).
[1] Current address: Department of Pathology, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0612, USA.

0042-6822/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2012.01.039

that this standard may be suboptimal, since L1 may not accurately reflect biological and whole genome relationships. As supporting evidence, previous studies have shown that the evolutionary relationships of PVs differ depending on the gene analyzed (Bravo and Alonso, 2007; García-Vallvé et al., 2005; Gottschling et al., 2007a,b; Rector et al., 2008; Van Bressem et al., 2007).

In this study, we report full genomes of four novel bottlenose dolphin papillomaviruses, tentatively named TtPV4, TtPV5, TtPV6 and TtPV7. Genomes TtPV4-6 were identified in genital lesions while TtPV7 was identified in normal genital mucosa. The availability of these new full-length genomes allowed us to conduct an in-depth analysis of the evolutionary processes that have influenced the diversification of this group. Our analysis suggests that recombination and diversifying selection pressures are prevalent across most genes on cetacean PVs and constitute important driving forces of their evolution.

Results

TtPV4, TtPV5, TtPV6 and TtPV7 genome structure

Four complete PV genomes were detected and amplified from genital mucosa demonstrating papillomatous lesions as well as from normal genital mucosa of three male bottlenose dolphins. The genomes were named TtPV4, TtPV5, TtPV6 and TtPV7 based on the chronological order in which they were found. Full-genome nucleotide sequences were deposited in the GenBank database with accession numbers (JN709469, JN709470, JN709471 and JN709472). TtPV4 and TtPV5 were found in penile lesions from the same dolphin using rolling circle amplification (RCA) and sequencing of cloned fragments. TtPV6 was detected in a genital papilloma lesion from a different dolphin also using the RCA method while TtPV7 was detected using degenerate primers E1F2-E1R4 in normal penile mucosa of a dolphin.

The complete genomes of TtPV4-TtPV7 range from 7783 bp (TtPV7) to 7895 bp (TtPV6) and present the typical organization observed in other characterized cetacean PVs (TtPV1, TtPV2, TtPV3, PsPV1, DdPV1, TtPV3-Variant, PphPV1, PphPV2 and PphPV3). Details of the genomic structure of these novel PVs are summarized in Table 1 and Fig. 1. Seven main ORFs encoded in the same strand and orientation can be identified, with E6, E1, E2, and E4 (which is nested within E2) clustered in the early region of the genome followed by L2, L1, and L3 (nested within L1) in the late region. Like all cetacean PVs except TtPV2, the predicted E4 ORFs lack an ATG start codon. A conserved ORF with homology to one known as L3 in BPV3 and Bovine PV4 (BPV4) was identified in all cetacean PVs except TtPV1, although TtPV5 and PsPV1 lacked traditional start codons. The non-coding region (NCR) ranges from 611 bp in TtPV7 to 710 bp in TtPV6 and is well within the range of other published cetacean PVs (546 bp in TtPV3 and 916 bp in TtPV2) (Table 1; Fig. 1).

Characteristic DNA and amino acid motifs of TtPV4, TtPV5, TtPV6 and TtPV7

The number of DNA and amino acid motifs found in the four characterized TtPVs varied in each of the genomes and is given in Table 2. In all four TtPVs, the E6 ORF codes for a protein ranging from 215 aa to 219 aa, which is similar in size to other cetacean PVs, but larger in comparison to human PVs. In all cases, the predicted E6 contains two metal-binding motifs ($CX_2CX_{29}CX_2C$) separated by 36 aa, identical to those found in other cetacean PVs (Table 2). Interestingly, a pRb-binding motif (LXCXE), involved in permitting cell cycle dysregulation and usually found in E7, was found in the predicted E6 protein of TtPV6. The only other cetacean PV where a pRb motif has been found is in the E1 protein of PsPV1. A modified ATP-dependent helicase motif (GPANTGKS) was found in the predicted E1 protein of all four TtPVs (Table 2).

The number of E2 binding sites (E2BS; ACC-N_6-GGT) ranged from three in TtPV4 to six in TtPV7. The number of polyA sites (AATAAA) also varied from two in TtPV4 and TtPV7 to five in TtPV5. An SP1 binding site (GGGCGG) was identified within the E1 gene in TtPV4 and TtPV7, but not in TtPV5 or TtPV6. The number of NF1 binding sites (TTGGC) ranged from four in TtPV7 to nine in TtPV6. Numbers of the TATA signal (TATAAA) ranged from one in TtPV5 to five in TtPV6.

Sequence homology among cetacean PVs and phylogenetic analysis

Based on the L1 sequence comparisons, the most homologous pair of cetacean PV L1 ORFs is TtPV3 and TtPV3-Variant with 98.23% similarity, followed by TtPV4 and TtPV7 with 89.21% similarity, and the least homologous are the pairs of PhPV3 and TtPV2 and PhPV1 and TtPV2, both with 56.78% similarity (Table 3). The E1 protein showed less structured relationships, with most of the similarities being higher than 60%. Sequence similarity below 60% was found mostly in comparisons involving TtPV2 (Table 3).

Different clustering patterns of the cetacean PVs were observed in the El and L1 Bayesian phylogenetic analyses. Based on the L1 tree, the cetacean PVs are paraphyletic; all cetacean PVs group in one clade except for PhPV3, which groups in a strongly supported clade containing *Alphapapillomavirus*, *Omegapapillomavirus*, and *Dyodeltapapillomavirus* instead of the main cetacean clade (Fig. 2), in agreement with Gottschling et al. (2011). The L1 cetacean cluster is split into two highly supported subsets, with *Omikronpapillomavirus* including TtPV5, TtPV6 PsPV1 and PhPV1, and *Upsilonpapillomavirus* including DdPV1, TtPV1, TtPV4, TtPV7, TtPV3 and TtPV3-Variant with PhPV2 and TtPV2, as the most basal taxa of the cluster. Additionally, in the L1 analysis the PV cetacean clusters clearly represent the three genera suggested by the sequence similarity comparisons (*Upsilonpapillomavirus*, *Omikronpapillomavirus* and one yet undefined genus). Based on the L1 phylogenetic results, the ruminant

Table 1
Details of the sequence analysis of TtPV4, TtPV5, TtPV6 and TtPV7 that includes previously described cetacean PV genomes for comparison. Length in nucleotides for predicted genes is given in each column. ORFs marked with * lack traditional start codons. NI = not identified.

Virus	Sample origin	Length (bp)	E6	E7	E1	E2	E4	E5	L2	L1	L3	NCR
TtPV1[a]	Genital papilloma	8089	672	NI	1917	1215	402*	NI	1623	1521	207*	737
TtPV2[b]	Genital papilloma	7866	618	NI	1872	1140	339	309	1632	1536	288	916
TtPV3[a]	Genital papilloma	7915	660	NI	1926	1245	411*	NI	1626	1512	279	546
TtPV4	Genital papilloma	7792	654	NI	1935	1185	366*	NI	1686	1515	348	612
TtPV5	Genital papilloma	7853	657	NI	1932	1176	357*	NI	1704	1560	282	658
TtPV6	Genital papilloma	7895	654	NI	1926	1191	363*	NI	1686	1560	348	710
TtPV7	Normal penile mucosa	7783	648	NI	1929	1191	354*	NI	1626	1518	348	611
PsPV1[c]	Genital papilloma	7879	636	NI	1929	1245	426*	NI	1665	1575	228	621

[a] From Rector et al. (2008).
[b] From Rehtanz et al. (2006).
[c] From Van Bressem et al. (2007).