Virology 393 (2009) 304–310

Contents lists available at ScienceDirect

Virology

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

Genomic diversity and interspecies host infection of α12 *Macaca fascicularis* papillomaviruses (MfPVs)

Zigui Chen [a,*], Koenraad van Doorslaer [a], Rob DeSalle [b], Charles E. Wood [c], Jay R. Kaplan [c], Janice D. Wagner [c], Robert D. Burk [a,d,*]

[a] *Department of Microbiology and Immunology, Albert Einstein College of Medicine, Bronx, New York 10461, USA*
[b] *Sackler Institute of Comparative Genomics, American Museum of Natural History, New York, New York 10024, USA*
[c] *Department of Pathology/Section on Comparative Medicine, Wake Forest University School of Medicine, Winston-Salem, North Carolina 27157, USA*
[d] *Departments of Pediatrics, Epidemiology and Population Health, Obstetrics, Gynecology and Woman's Health, and Albert Einstein Cancer Center, Albert Einstein College of Medicine, Bronx, New York 10461, USA*

ARTICLE INFO

Article history:
Received 18 June 2009
Accepted 14 July 2009
Available online 28 August 2009

Keywords:
Alpha papillomavirus
Macaca fascicularis
Novel PVs
Genomic diversity
Evolution

ABSTRACT

Alpha human papillomaviruses (HPVs) are among the most common sexually transmitted agents of which a subset causes cervical neoplasia and cancer in humans. Alpha-PVs have also been identified in non-human primates although few studies have systematically characterized such types. We cloned and characterized 10 distinct types of PVs from exfoliated cervicovaginal cells from different populations of female cynomolgus macaques (*Macaca fascicularis*) originating from China and Indonesia. These include 5 novel genotypes and 5 previously identified genotypes found in rhesus (*Macaca mulatta*) (RhPV-1, RhPV-a, RhPV-b and RhPV-d) and cynomolgus macaques (MfPV-a). Type-specific primers were designed to amplify the complete PV genomes using an overlapping PCR method. Four MfPVs were associated with cervical intraepithelial neoplasia (CIN). The most prevalent virus type was MfPV-3 (formerly RhPV-d), which was identified in 60% of animals with CIN. In addition, the complete genomes of variants of MfPV-3 and RhPV-1 were characterized. These variants are 97.1% and 97.7% similar across the L1 nucleotide sequences with the prototype genomes, respectively. Sequence comparisons and phylogenetic analyses indicate that these novel MfPVs cluster together within the α12 PV species closely related to the α9 (e.g., HPV16) and α11 species (e.g., HPV34), and all share a most recent common ancestor. Our data expand the molecular diversity of non-human primate PVs and suggest a recent expansion of alpha-PV species groups. Moreover, identification of an overlapping set of MfPVs in rhesus and cynomolgus macaques indicates that non-human primate alpha-PVs might not be strictly species-specific and may represent past interspecies infection.

Introduction

Cervical cancer is the most frequent gynecologic malignancy and one of the leading causes of cancer mortality in women worldwide (Parkin and Bray, 2006; Parkin et al., 2005). Genital human papillomavirus (HPV) infection is one of the most prevalent sexually transmitted infections and the principal cause of cervical cancer and its precursor, cervical intraepithelial neoplasia (CIN). Papillomaviruses (PVs) comprise a diverse group of small epitheliotropic viruses with circular double-stranded DNA genomes about 8 kb in size. PVs are highly species-specific pathogens that cause proliferation of predominantly epithelial cells in a wide range of host species. For a PV to be recognized as a distinct type, the DNA sequence of the L1 open reading frame (ORF) of the cloned genome should be no more than 90% similar to previously typed PVs; a HPV genome whose L1 nucleotide sequence is greater than 2% and less than 10% different from that of the closest type is defined as "subtype" (de Villiers et al., 2004). Currently, over 100 types of HPVs and 47 types of non-human PVs have been fully characterized. Based on the L1 ORF, PVs have recently been classified into alpha-PV (mucosal/genital), beta-PV (cutaneous), and gamma-PV (cutaneous), in addition to at least 13 other genera (de Villiers et al., 2004). All genital HPVs associated with cervical cancer are members of the alpha-PV (α-PV) genus, which collectively contains the largest proportion of currently identified HPV types (de Villiers et al., 2004; Munoz et al., 2003).

Non-human primate alpha-PVs have been detected from a wide range of apes and monkeys including rhesus (*Macaca mulatta*) and cynomolgus macaques (*Macaca fascicularis*), common (*Pan troglodytes*) and pygmy chimpanzees (*Pan paniscus*), and gorillas (*Gorilla gorilla*) (Antonsson and Hansson, 2002; Chan et al., 1997; Van Ranst et al., 1991). At least 30 non-human primate PV types have been identified; however, only 3 have been completely characterized. Two

* Corresponding authors. Fax: +1 718 430 8975.
E-mail addresses: zchen@aecom.yu.edu (Z. Chen), burk@aecom.yu.edu (R.D. Burk).

0042-6822/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2009.07.012

chimpanzee PV subtypes, PtPV-1 (Common chimpanzee PV type 1) and PpPV-1 (Pygmy chimpanzee PV type 1), the latter from an oral focal epithelial hyperplasia in the pygmy chimpanzee, cluster into the α10 PV species group together with HPV6, HPV11 and HPV13 (Van Ranst et al., 1991). RhPV-1, isolated from a rhesus macaque metastatic penile squamous cell carcinoma (Kloster et al., 1988; Ostrow, LaBresh, and Faras, 1991), is contained within the α12 species group and is phylogenetically closely related to HPV16 (α9 species) and HPV34 (α11 species) (de Villiers et al., 2004). In addition, 13 RhPVs previously identified by sequencing small PCR fragments from cervical samples from rhesus monkeys formed three new clades/species within the alpha-PV group, which do not include any HPV types (Chan et al., 1997). In a recent survey of naturally occurring PVs associated with cervical intraepithelial neoplasia in female cynomolgus macaques, Wood et al. (2007) identified a PV type (RhPV-d) that was closely related to HPV16, experimentally transmissible, and capable of inducing early cervical neoplasia (Wood et al., 2007).

In this study, we characterized macaque PV genomes identified in genital samples from female cynomolgus macaques (*Macaca fascicularis*). Ten distinct macaque PVs were identified, including 5 novel types and 5 types previously identified in rhesus macaques and cynomolgus macaques (Chan et al., 1997). The complete genomes were cloned and characterized to investigate the molecular diversity and the evolutionary history of primate PVs associated with cervical lesions.

Results and discussion

Genital PV and cervical neoplasia in macaques

PV DNA was detected in 24.9% (45/181) of adult female cynomolgus macaques originating from Indonesia or China (Table 1) (Wood et al., 2007). Characterization of these genomes revealed 5 novel macaque PVs and 5 previously identified types, four in rhesus macaques (RhPV-1, RhPV-a, RhPV-b and RhPV-d) and one in a cynomolgus macaque (MfPV-a) (Chan et al., 1997; Ostrow, LaBresh, and Faras, 1991). Except for RhPV-1 that was previously characterized (Ostrow, LaBresh, and Faras, 1991), we cloned and sequenced the complete genomes of these putative novel types that are named *Macaca fascicularis* PV-number (MfPV-3 to -11) (Table 1). All macaque samples were negative for over 40 common HPVs by oligonucleotide hybridization. Four different macaque PV types were associated with histologically-confirmed CIN (MfPV-3, -4, -5, and -8) (Table 1). Prevalence of different PVs in the collection of samples confirmed that MfPV-3 (formerly RhPV-d in Wood et al. 2007) was the most common macaque PV, and we infer that in these animals it was both persistent (based on animals with no breeding contact for over 3.5 years) and oncogenic (Wood et al., 2007).

Characterizations of the novel MfPV complete genomes

Analysis of novel MfPV genome sequences revealed sizes of 7920–8063 bp, with a GC content between 46.02% and 50.35% (Table 1). They all contain the classical six early (E6, E7, E1, E2, E4 and E5) and two late (L2 and L1) ORFs (Supplement 1).

The putative E6 ORFs of these MfPVs contain two zinc-binding domains ($CxxC(x)_{29}CxxC$), separated by 36 amino acids. An additional upstream in frame ATG was identified in the MfPV-10 and RhPV-1 E6 ORFs (Supplement 2) (Ostrow, LaBresh, and Faras, 1991). However, these upstream nucleotides (MfPV-10, nt 7840–7920; RhPV-1, 7948–8028) overlap a classical TATAA box and potentially encode amino acids not related to the canonical E6 genes of the other genital MfPV types, making it unlikely to be the E6 translation initiation codon. Accordingly, we suggest that a conserved downstream ATG of the E6 ORF is the translational start site and we used it to position the first nucleotide of the complete genome (Supplement 1). All MfPV E6s have a splice donor/acceptor pair that allows them to generate a spliced E6 (E6*) (data not shown). Alpha HPVs associated with malignant progression (e.g., HPV16, HPV18) have splice donor/acceptor sites and generate a small protein corresponding to E6*, which appears to be critical in the generation of an mRNA for E7 expression. In contrast, viruses lacking this splice donor/acceptor site (e.g., HPV11, HPV6) transcribe a major collinear E6–E7 mRNA, responsible for generation of E7 from a promoter located within the E6 gene (Sedman et al., 1991; Smotkin, Prokoph, and Wettstein, 1989). Interestingly, a PDZ-binding motif (x-T/S-x-V/L), which is highly conserved in α5–7, α9, α11 species HPV E6s, many of which are associated with cancer, is present in the carboxy terminus of MfPV-3, -4, -6, -7, -8, -10 and -11, but absent in that of MfPV-5, -9 and RhPV-1. PDZ-domain-containing proteins are involved in a variety of cellular functions such as cell signaling and cell adhesion. High-risk HPV E6s (e.g., HPV16, HPV18) but not low-risk HPV E6s (e.g., HPV6, HPV11) have been shown to interact with these proteins through the C-terminal motif, leading to their degradation (Kiyono et al., 1997). However, RhPV-1 isolated from a penile squamous cell carcinoma (Schneider et al., 1991) lacks a PDZ-binding motif indicating that this motif may not be essential to E6-induced carcinogenesis and host cell transformation. Alternatively, other virus protein functions may complement the lack of a PDZ-binding domain.

The putative E7 ORFs of these genital MfPVs contain a conserved zinc-binding domain, $CxxC(x)_{29}CxxC$, and a pRB binding motif (LxCxE). Similar to genital HPV types, the carboxy-terminal region of the E1 proteins contains the highly conserved binding site of the ATP-dependent helicase (GPP/ANTGKS). Although both the putative E4 and E5 ORFs show significant homology with known HPV E4 and E5 genes (e.g., HPV16, HPV34), no initiation/start codon was identified, suggesting that these ORFs are translated from a spliced transcript. These macaque PVs contain E5 homologues to carcinogenic alpha HPVs within the α5–7, α9 and α11 species, and the members of the α1, α8 and α10 species causing venereal warts (Schiffman et al., 2005). A polyadenylation consensus sequence (AATAAA) for processing of early viral mRNA transcripts is present at the beginning of the L2 gene. The major (L1) and minor (L2) capsid proteins of these types contain a nuclear localization signal at their 3′ end.

The upstream regulatory regions (URRs) of these MfPVs are located between the stop codon of L1 and the first start codon of the E6 ORF consisting of 764–803 bp, similar in length to the URRs of alpha HPVs, and contain many *cis*-acting regulatory sequences that control viral transcription and replication (Supplement 3). Four canonical palindromic E2-binding sites [$ACC(N)_6GGT$] are present within the URR regions of all these MfPVs, with an E1 binding site located between two E2-binding sites. Binding sites for transcriptional regulatory factors including AP-1 (Chan et al., 1990), NF-1 (Apt et al., 1993), SP-1 (Gloss and Bernard, 1990), transcriptional enhancer factors (TEF)-1 (Ishiji et al., 1992), and YY-1 (Dong et al., 1994) were also present within the URR regions. Regulatory sites include a polyadenylation site (AATAAA) processing L1/L2 capsid mRNA transcripts at the URR 5′ end (except for RhPV-1), and a TATA box representing the E6/E7 promoter at the 3′ end (Supplement 3).

Phylogeny of genital MfPVs

In order to assess the phylogenetic relationships of these novel primate PVs, trees inferred from the concatenated amino acids and nucleotide sequences of early ORFs (E6, E7, E1, and E2) and late ORFs (L2 and L1) were constructed using multiple algorithms (Fig. 1). Representative types within the alpha HPVs were selected for analysis. Three cutaneous types, HPV4, HPV5 and MfPV-1, were set as the "referent out-group". The topologies strongly support the existence of a monophyletic clade grouping MfPV-3 to -11 together with RhPV-1, expanding the genomic diversity of the α12 species (Fig. 1a). The members of the α12 species are closely related to the members of α9