Phylogenetically related, clinically different: human papillomaviruses 6 and 11 variants distribution in genital warts and in laryngeal papillomatosis

J. M. Godínez[1,2,†], S. Nicolás-Párraga[1,2,†], V. N. Pimenoff[1,2], B. Mengual-Chuliá[1,2], N. Muñoz[3], F. X.Bosch[1,2], G. I.Sánchez[4], J. McCloskey[5] and I. G. Bravo[1,2]

1) *Infections and Cancer Laboratory, Cancer Epidemiology Research Program, Catalan Institute of Oncology (ICO),* 2) *Bellvitge Institute of Biomedical Research (IDIBELL), Barcelona, Spain,* 3) *Cancer Institute of Colombia, Bogota, Colombia,* 4) *Infection and Cancer Group, School of Medicine, University of Antioquia, Medellin, Colombia and* 5) *Sexual Health Clinic, Royal Perth Hospital, School of Laboratory Medicine, University of WA, Perth, Australia*

Abstract

Genital warts (GWs) and laryngeal papillomatosis (LP) are two usually benign pathologies related to infection with human papillomaviruses (HPVs), mainly HPV6 and HPV11. The aim of this work was to describe the genetic diversity of HPV6 and HPV11 isolates found in GWs and LPs, and to analyse the differential involvement of viral variants in either lesion. A total of 231 samples diagnosed as GWs (n = 198) or LP (n = 33) and caused by HPV6 or HPV11 monoinfections were analysed. The phylogenetic relationships of the retrieved viral sequences were explored. We have identified the long control region and the intergenic E2–L2 region as the two most variable regions in both HPV6 and HPV11 genomes. We have generated new HPV6 (n = 166) or HPV11 (n = 65) partial sequences from GWs and LPs lesions spanning both regions and studied them in the context of all available sequences of both types (final n = 412). Our results show a significant ($p < 0.01$) differential presence of HPV6 variants among both pathologies, with HPV6 B variants being preferentially found in GW versus LP samples. No differential involvement of HPV11 variants was observed. Our findings suggest that different HPV6 variants may either show differential tropism or have different potential to induce lesions in different epithelia.

Keywords: Genital warts, human papillomaviruses, laryngeal papillomatosis, phylogeny, recurrent respiratory papillomatosis, tissue tropism, variants

Original Submission: 25 June 2013; **Revised Submission:** 2 September 2013; **Accepted:** 30 September 2013

Editor: G. Antonelli

Article published online: 9 December 2013

Clin Microbiol Infect 2014; **20:** O406–O413

doi: 10.1111/1469-0691.12420

Corresponding author: J. M. Godínez, Infections and Cancer Laboratory, Cancer Epidemiology Research Programme, Institut Català d' Oncologia, L'Hospitalet de Llobregat, ICO-IDIBELL, Gran Via 199-203, 08907 Barcelona, Spain

E-mail: jmgodinez@iconcologia.net

[†]Both authors contributed equally to this work.

Introduction

Papillomaviruses are small, non-enveloped viruses with a circular double-stranded DNA genome of around 8000 bp [1]. More than 250 complete papillomavirus genomes have been described, infecting human and non-human hosts (http://pave.niaid.nih.gov/#home). Human papillomaviruses (HPVs) are the causative agents of cancer of the cervix, and are also involved in cancers of the penis, anus, vagina, vulva, and head and neck, as well as in other benign, wart-like lesions [2]. Based on this association to cervical cancer, HPVs have been epidemiologically stratified into three risk groups: carcinogenic, probably and possibly carcinogenic, and not carcinogenic to humans [3]. Alphapapillomaviruses HPV6 and HPV11 belong to the non-carcinogenic group, being the most common non-oncogenic HPVs found in cervical specimens in the general population [4].

HPV6 and HPV11 are the causative agents in some conspicuous lesions, namely anogenital warts (GWs) and

laryngeal papillomatosis (LP). GWs are benign tumours of the epithelium caused by papillomavirus infection, mainly with HPV6 and HPV11 (85% of the cases) [5]. Co-infections by oncogenic and non-oncogenic types are commonly detected in a high proportion of anogenital warts (45%), which have been proposed as a partial explanation of the increased risk of cervical intraepithelial neoplasia and invasive cervical carcinoma in women with GWs [6]. GWs are closely associated with sexual behaviour, with number of sexual partners being the main risk factor [7]. The highest incidence rate for GWs in women is at 20–24 years, which correlates well with the peak of papillomavirus infection in the female genital tract [4]. In men, the incidence peak occurs at 20–29 years of age [2].

Laryngeal papillomatosis, or recurrent respiratory papillomatosis, is a neoplastic disease of the airways mainly caused by HPV6 and HPV11, although HPV16 has also been identified in a few cases [8]. It represents the most common benign tumour of the larynx in infants and children [9]. Some studies have identified infection with HPV11 as being associated with more aggressive disease and higher recurrence of lesions [8,10], and malignant transformation of lesions has been described in approximately 5% of cases [11]. The clinical complications of this pathology include dysphonia, dyspnoea and, in serious cases, complete obstruction of the airways [12].

Papillomavirus variants are defined as viral sequences sharing >98% identity in the nucleotide sequence in the L1 gene [13]. Based on this criterion, HPV6 and HPV11 variant lineages have been described [14]. Several studies have addressed the genetic diversity of HPV6 and HPV11 [15–17], and some of them have aimed to establish a link between genetic variation and differential outcome of the infection [8,18].

The aim of this study was to analyse first the genetic diversity of HPV6 and HPV11 sequences retrieved from two different but related pathologies, namely GWs and LP. Further, the phylogenetic relationship of all HPV6 and HPV11 sequences and tissue-dependent distribution of the variants were analysed.

Methods

Samples

Samples analysed in this project originate from two different formalin-fixed paraffin-embedded (FFPE) sample repositories. GWs were obtained from the Surgical Genital Wart Biobank established in 1995 at the Sexual Health Clinic at Royal Perth Hospital, Perth, Australia. These samples include FFPE surgery specimens excised from patients who required surgical resection of anal and/or perianal GWs [19]. One hundred and forty-three HPV6 and sixty-four HPV11 single infected samples from the first surgical event of each patient were included.

Laryngeal papillomatosis samples originated from a multicentre study of cases diagnosed between 1985 and 2009, in the cities of Cali and Medellin, Colombia [10]. Forty-one HPV6 and eleven HPV11 single-infected samples, each from a different patient, were included. Detailed information about the samples included is shown in the Supplementary material, Table S1.

Presence of HPV DNA in the samples was assessed by using the SPF_{10}-DEiA-LiPA protocol (version 1; Laboratory Biomedical Products, Rijswijk, the Netherlands). The SPF_{10} system targets a 65-base pair region of L1 gene of a broad spectrum of Alphapapillomaviruses. HPV-positive samples were identified and genotyped by amplicon hybridization (DEiA) and reverse hybridization line probe assay, LiPA25. The detected viruses were HPV6, 11, 16, 18, 31, 33, 34, 35, 39, 40, 42, 43, 44, 45, 51, 52, 53, 54, 58, 59, 56, 66, 68, 70 and 74.

Selection of the most informative genomic regions

Fragmentation of genetic material in the FFPE samples prevented us from obtaining full-genome sequences. To select the most informative regions for the study, the variability of different regions of the viral genomes was assessed.

HPV6 and HPV11 complete genome unique sequences were obtained via GenBank. The different open reading frames (ORFs: E6, E7, E1, E2, L1 and L2), the long control region (LCR), and the intergenic E2–L2 region (IntE2L2) were extracted and aligned. This intergenic region spans the E5a and E5b ORFs of HPV6 and HPV11 [20]. All sequences were aligned at amino acid level (except the non-coding LCR), back-translated and concatenated to obtain full-genome reference alignments. For each of the alignments, phylogenetic relationships were inferred under a maximum likelihood framework using RAxML v7.2.8 (http://www.exelixis-lab.org/) [21], using the GTR+Γ4 model, and the number of required bootstrap cycles was determined with the *–autoMRE* command [22]. The well-resolved phylogenetic trees obtained were further employed to compute tree-guided, model-based pairwise genetic distances between taxons (*–f x* command in RAxML).

PCR and sequencing

DNA was extracted by incubation of the material with 250 μL of proteinase K buffer (10 mg/mL proteinase K in 50 mM Tris–HCl, pH 8.0) overnight at 56°C. The samples were later incubated at 95°C for 8 min to inactivate proteinase K and were stored at −20°C until use.

Based on the pairwise distance results, the LCR and the IntE2L2 were chosen as amplification targets. Different type-specific PCR systems were designed for the amplification