Infection, Genetics and Evolution 11 (2011) 1805–1810

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

Short communication

Pre-vaccination genomic diversity of human papillomavirus genotype 6 (HPV 6): A comparative analysis of 21 full-length genome sequences

Boštjan J. Kocjan, Mateja M. Jelen, Polona J. Maver, Katja Seme, Mario Poljak *

Institute of Microbiology and Immunology, Faculty of Medicine, University of Ljubljana, Ljubljana, Slovenia

ARTICLE INFO

Article history:
Received 12 April 2011
Received in revised form 8 June 2011
Accepted 27 June 2011
Available online 22 July 2011

Keywords:
Human papillomavirus
HPV 6
Complete genome
Genomic diversity
Genomic variant

ABSTRACT

Comparative analysis of 21 full-length genome sequences of human papillomavirus genotype 6 (HPV 6): 18 determined in this study and three sequences available in nucleotide sequence databases, revealed more than 98% nucleotide similarity to the HPV 6 prototype isolate. The minimum and maximum genomic distance between the full-length genomic variants and the prototype sequence was three nucleotide substitutions, and 122 nucleotide substitutions and three insertions, respectively. Detailed sequence analysis of early viral genes E7, E1, E2 and E4, late viral gene L2, and three non-classic non-coding genomic regions (NNCR) revealed the existence of at least four E7, twelve E1, eleven E2, six E4, eleven L2, two NNCR1, two NNCR2, and three NNCR3 genomic variants. In addition, several novel, potentially important amino acid mutations were identified. A phylogenetic tree calculated from viral genome sequences was dichotomic, separating all isolates into HPV 6b (prototypic) and HPV 6a/6vc (non-prototypic) genetic lineages. This study, which contributed the largest number of full-length HPV 6 genome sequences to date, confirmed that HPV 6 diversifies virtually equally across the entire genome by nucleotide (amino acid) exchanges in coding regions and additional nucleotide insertions/deletions in non-coding regions. However, this diversification trend was more evident in non-coding regions LCR and NNCR3 and early viral genes E4, E5a and E5b.

1. Introduction

Human papillomavirus genotype 6 (HPV 6) is classified taxonomically in the alpha papillomavirus genus – species 10 or "HPV 6 species", together with four closely related HPVs: HPV 11, HPV 13, HPV 44 and HPV 74 (Bernard et al., 2010; de Villiers et al., 2004). Due to its exclusive association with benign neoplasm, HPV 6 is generally regarded as a low-risk genotype (Bouvard, 2009; Li et al., 2010). Next to HPV 11, it is the etiological agent of at least 90% of all genital warts and laryngeal papillomas (Aubin et al., 2008; Gale et al., 1994; Potočnik et al., 2007), and at least 12.1% of low-grade squamous cervical intraepithelial lesions (Clifford et al., 2005). HPV 6 is the most prevalent low-risk HPV genotype, with a worldwide prevalence of 0.5% in HPV positive women with normal cytology (Bruni et al., 2010). HPV 6 has been included in the recently developed quadrivalent HPV vaccine, together with three other HPV genotypes: HPV 16, HPV 18 and HPV 11 (Garland et al., 2007).

Although the level of genomic diversity has been well established for some of the high-risk HPV genotypes, notably HPV 16 and HPV 18, identifying some viral variants that differ profoundly in pathogenicity, transmission and persistence, and progression of the clinical disease (Gagnon et al., 2004; Lee et al., 2008; Pista et al., 2007; Schiffman et al., 2010; Sichero et al., 2007; Xi et al., 2006), similar knowledge about low-risk HPVs such as HPV 6 is scarce. The few genomic diversity studies performed so far have investigated a limited number of HPV 6 isolates and most of them focused on a single genomic region or even only a short part of it (reviewed in Kocjan et al., 2009a). Recently, an extensive study of the genomic diversity of HPV 6 was performed by our group, in which approximately half of the viral genome of 77 clinically important isolates was analyzed, including the 3′ half of the E2 [E2(3′)] and the entire L1, LCR, E6, E5a and E5b genomic regions (Kocjan et al., 2009a). The study showed a substantial degree of HPV 6 genetic diversity and identified several novel genomic variants, some harboring potentially important mutations. In addition, the presence of major sequence errors in HPV 6 sequences deposited in sequence repositories was identified or confirmed.

To date, only three full-length genome sequences of HPV 6 have been available in nucleotide sequence databases: HPV 6b (GenBank accession number X00203), HPV 6a (L41216), and HPV 6vc (AF092932). These isolates, showing a substantial degree of a genetic variability across the entire genome, form the basis for HPV 6 genomic variant classification. It has been shown, namely, usually by comparing a single genomic region only, that all HPV 6 isolates

* Corresponding author. Address: University of Ljubljana, Institute of Microbiology and Immunology, Faculty of Medicine, Zaloška 4, 1105 Ljubljana, Slovenia. Tel.: +386 1 543 7453; fax: +386 1 543 7418.
E-mail address: mario.poljak@mf.uni-lj.si (M. Poljak).

1567-1348/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.meegid.2011.06.022

can be divided into two genetically distinct variant groups: one containing prototype HPV 6b-related (prototypic) and the other containing HPV 6a/6vc-related (non-prototypic) genomic variants (Caparrós-Wanderley et al., 1999; Grassmann et al., 1996; Heinzel et al., 1995; Kocjan et al., 2009a; Kovelman et al., 1999). Non-prototypic HPV 6 genomic variants predominate in genital warts and laryngeal papillomas (Kocjan et al., 2009a), thus giving this particular variant group a more important role in the etiopathogenesis of genital warts and other HPV 6-related tumours.

In order to investigate HPV 6 genetic diversity further, representative isolates of the 18 most divergent HPV genomic variants determined in our previous study (Kocjan et al., 2009a) were selected for full-length genome amplification and sequence variation analysis. This allowed robust analysis of the HPV 6 phylogeny for the first time and provided new important data concerning previously non-characterized regions of the viral genome, including the 5′ half of E2 [E2(5′)], the entire E7, E1, E4 and L2 ORF, and three non-classic non-coding regions (NNCR) positioned between E7 and E1, E2 and E5a, and E5b and L2 ORF. All coding regions were examined in detail for the presence of specific amino acid (aa) substitutions that might have divergent biological or/and pathological properties.

2. Materials and methods

2.1. HPV 6 isolates

A total of 18 HPV 6 isolates from our previous study on HPV 6 genomic diversity (Kocjan et al., 2009a), representing the 18 most divergent HPV 6 genomic variants with specific aa changes in E2(3′), L1, LCR, E6 and E5 ORF and/or nucleotide sequence alternations in the LCR genomic region, were included in the study. Six isolates belonged to the HPV 6b genomic variant group, one to the HPV 6a, and 11 isolates belonged to the HPV 6vc genomic variant group. Eleven isolates were originally obtained from the same number of male patients with exophytic genital warts and eight isolates were obtained from four male and four female patients with laryngeal papillomas.

2.2. Full-genome amplification and sequencing of HPV 6 isolates

In order to sequence the complete genome of each HPV 6 isolate included in the study, two overlapping DNA fragments were generated using long-template PCR. A 4511 bp fragment containing the complete E6, E7, E1, E2, E4 and E5 ORF was amplified using the primer pair HPV6-E6S and HPV6-E5R (Table S1), and a 4908 bp fragment containing the complete L2 ORF, L1 ORF and LCR region was amplified using the primer pair HPV6-E5S and HPV6-E6R (Table S2). All primers used in the study were designed according to the corrected full-length genome sequence of the HPV 6b prototype isolate (X00203) (Heinzel et al., 1995) and full-length genome sequences of HPV 6a (L41216) and HPV 6vc (AF092932) isolates, as described previously (Kocjan et al., 2009a).

Both PCR reactions were carried out in 0.2 ml reaction tubes, each containing up to 200 ng of template DNA, 25 μl of High Fidelity PCR Master (Roche Diagnostics, Mannheim, Germany), 0.3 μl of each primer and water up to 50 μl. The cycling conditions consisted of an initial denaturation at 94 °C for 2 min, 10 cycles of 10 s at 94 °C, 70 s at 56 °C and 4 min at 68 °C, and an additional 25 cycles of 15 s at 94 °C, 30 s at 56 °C and 4 min at 68 °C; for the later cycles the extension step increased by 5 s for each successive cycle. The final extension step was performed at 72 °C for 7 min and reactions were cooled at 4 °C. All PCR amplifications were carried out on a GeneAmp® PCR instrument type 9700 (PE Applied Biosystems, Foster City, USA).

The PCR products were analyzed on 0.8% agarose gel (A9539-506 Agarose For Routine Use, Sigma–Aldrich, St. Louis, USA) using E-Gel 96 High Range DNA Marker (Invitrogen, Carlsbad, USA) and purified with a QIAquick PCR purification kit (Qiagen, Hilden, Germany). Concentrations of purified amplicons were estimated on 0.8% agarose gel using a High DNA Mass Ladder (Invitrogen) and set to 50 ng/μl. Sequencing of the long-template PCR products was performed at Macrogen, Ltd. (Seoul, Korea) with primers listed in Supplementary Tables S1 and S2. The acquired nucleotide sequences were analyzed and assembled with Vector NTI Advance v11.0 (Invitrogen).

2.3. Identification of HPV 6 genomic variants and phylogenetic analyses

The identification of HPV 6 genomic variants and HPV 6 intergenomic co-variation analysis was done with the ClustalW Multiple alignment program of the BioEdit Sequence Alignment Editor v7.0.9.0 program package (North Carolina State University, Raleigh, NC), using the corrected full-length genome sequence of the HPV 6b prototype isolate (X00203) (Heinzel et al., 1995) as a standard for comparisons and nucleotide position numbering.

Phylogenetic analyses of the HPV 6 full-length genome sequences were performed using the Mega program package (v4.0.1), as described previously (Tamura et al., 2007). A neighbor-joining method was used to construct a phylogenetic tree and a bootstrapping method was used to assess the reliability of the obtained tree.

2.4. Nucleotide sequence accession numbers

The HPV 6 nucleotide sequences reported in this paper are deposited in the DDBJ, EMBL and GenBank databases under the following accession numbers: FR751320, FR751321, FR751322, FR751323, FR751324, FR751325, FR751326, FR751327, FR751328, FR751329, FR751330, FR751331, FR751332, FR751334, FR751335, FR751336, FR751337 and FR751338.

3. Results and discussion

In order to perform the first phylogenetic analysis of full-length HPV 6 genome sequences and to determine the level of genetic diversity in previously non-analyzed genomic regions, one representative isolate of each of the six prototypic and one representative isolate of each of the 12 most divergent non-prototypic HPV 6 genomic variants, all containing specific aa changes in L1, E6, E2(3′) and/or E5 ORF and/or nucleotide sequence alternations in the LCR genomic region, were selected from our previous study (Kocjan et al., 2009a) for full-length genome sequencing. In addition, the genetic variation of the three non-classic non-coding regions of the HPV 6 genome (NNCR1-3) was investigated for the first time in this study.

All 18 HPV 6 isolates were successfully amplified and sequenced across the entire genome; the sizes of the full-length genome sequences ranged from 7996 to 8051 bp and differed in terms of specific insertions in the LCR genomic region (Fig. 1). Overall, among 18 full-length HPV 6 genomic variants, mutations were observed in 203 genomic positions and included 197 nucleotide substitutions, one 6 bp and one 1 bp deletion, and two 20 bp, one 14 bp and one 1 bp insertions. The minimum genomic distance between the full-length HPV 6 genomic variants and the HPV 6 prototype sequence was three nucleotide substitutions (Fig. 1, isolate FR751325) and the maximum genomic distance was 122 nucleotide substitutions and three insertions (Fig. 1, isolate FR751328). The most divergent full-length HPV 6 genomic variant revealed