Identification and Characterization of 3 Novel Genital Human Papillomaviruses by Overlapping Polymerase Chain Reaction: *cand*HPV89, *cand*HPV90, and *cand*HPV91

Masanori Terai and Robert D. Burk

Three novel human papillomaviruses (HPVs), *cand*HPV89, *cand*HPV90, and *cand*HPV91, that were previously identified from short polymerase chain reaction (PCR) fragments AE6/CP6108, JC9710, and JC9813, respectively, were cloned and characterized from cervicovaginal cells by use of an overlapping PCR method. The complete nucleotide sequences of *cand*HPV89 (8078 bp), *cand*HPV90 (8033 bp), and *cand*HPV91 (7966 bp) were determined by sequence walking. *cand*HPV89 and *cand*HPV91 were closely related to HPV83 and HPV7 and were placed in the HPV genome homology groups A3 and A8, respectively, by phylogenetic analyses. The genome of *cand*HPV90 was most closely related to HPV71, although these HPV genomes do not seem to form a single lineage, because of the disproportionate divergence of the HPV71 L1 region. On the basis of phylogenetic analyses and available clinical data, these 3 novel HPV genomes appear to have a low oncogenic risk and expand the heterogeneity of HPVs detected in the lower genital tract.

Human papillomaviruses (HPVs) are a heterogeneous group of viruses with an ~8000-bp, closed-circular, double-stranded DNA genome. They are causally involved in the etiology of neoplastic lesions of mucosal and skin epithelium [1]. Genital HPVs are sexually transmitted, and a subset of these are considered to be the sexually transmitted etiologic agent of cervical neoplasia and cancer [2]. HPVs are classified by the homology of their genomes. A novel type is defined as a cloned HPV genome with an L1 open reading frame (ORF) that displays >90% similarity to previously designated types [3]. More than 100 putative HPV types have been described, including >86 cloned and officially designated types and others identified from the sequences of polymerase chain reaction (PCR) products [3]. HPVs cloned and characterized from PCR products are termed “candidatus HPVs” (*cand*HPVs) by convention and in accordance with guidelines of the International Committee on the Taxonomy of Viruses [3]. When the *cand*HPV genomes are cloned directly (i.e., without the use of PCR), the isolate is referred to by the same name and/or number as the *cand*HPV but without the prefix “cand.” Priority for the isolation of that specific type will remain with the PCR-amplified genome [3]. Recently, with the availability and analysis of complete papillomavirus sequences, it was determined that HPV46, HPV55, and HPV64 are actually subtypes and not distinct HPVs [3].

We have isolated and characterized, using the overlapping PCR method, 3 novel genital HPV types that were amplified and cloned from cervicovaginal cells from young women with normal cytology [4, 5]. The overlapping PCR method is a technique that facilitates the complete characterization of HPV DNA from specimens with low copy numbers. Here, we describe the *cand*HPV89, *cand*HPV90, and *cand*HPV91 novel genomes, which were previously identified from short PCR fragments AE6/CP6108, JC9710, and JC9813, respectively, and present the complete nucleotide sequence, genome organization, predicted proteins, and phylogenetic analyses [3, 6–8].

**Materials and Methods**

*Species.* Clinical samples originated from a 19-year-old African American, non-Hispanic woman (*cand*HPV89), a 20-year-old Hispanic woman (*cand*HPV90), and an 18-year-old Asian woman (*cand*HPV91) who were participating in a study on the natural history of HPV infection [9, 10]. All subjects had normal Pap smears when cervicovaginal lavage samples were obtained for HPV analyses.

*Overlapping PCR.* The clinical samples were positive for HPV DNA, as determined by MY09/MY11 PCR and/or Southern blot hybridization during the course of a clinical study [9, 10]. Clinical samples containing HPVs *cand*HPV90 and *cand*HPV91 were negative for HPV DNA by MY09/MY11 PCR but positive by low-
stringency Southern blot analysis, whereas the sample containing candHPV89 was strongly positive using the MY09/MY11 primers. The initial overlapping PCR primers were designed by alignment of closely related HPV genomes determined by basic local alignment sequence tool (BLAST) analysis using the sequences of the partial L1 region amplified by MY09/MY11 (candHPV89 or GP5+/6+) (candHPV90 and candHPV91) [4, 5]. Additional primers were designed to amplify the entire genome in fragments. The primer sequences are available from the authors. Amplification was performed using either Gold Tag DNA polymerase (Perkin-Elmer Applied Biosystems) or an equal mixture of Gold Taq (Perkin-Elmer) and Pwo DNA polymerase (Platinum Tag DNA Polymerase High Fidelity; Gibco BRL). Pwo polymerase has an inherent 3'→5' exonuclease proofreading activity. The PCR products were separated by electrophoresis on agarose gels, stained with ethidium bromide, and visualized under UV illumination. After confirmation of appropriate product size, each PCR product was purified (Qiagen Gel Extraction kit; Qiagen) and ligated into the HPV insert, and additional primers were designed by sequence walking. Sequencing was performed in the Albert Einstein College of Medicine DNA sequencing core facility (Bronx, NY). The sequences of the overlapping fragments was assembled manually and confirmed by sequencing the complementary strand. Several additional primers were designed and used to clarify sequence ambiguities. Once assembled, the sequence was analyzed for homology to other HPVs by use of the BLAST software. The same software was used to determine protein sequence homologies. Phylogenetic trees were created using published HPV sequences available on-line from the Human Papillomaviruses 1997 Compendium (http://hpv-web.lanl.gov/stdgen/virus/hpv/compendium/htdocs/HTML_FILES/HPVcompintro4.html#comp97) and GenBank. Phylogenetic trees were derived, using public domain software [11], from individual ORFs and long control regions to determine the relationships of candHPV89, candHPV90, and candHPV91 to the available HPV sequences.

### Results

#### Characterization of 3 novel HPVs by overlapping PCR.

The genomes of candHPV89, candHPV90, and candHPV91 were amplified and cloned as 4 (i.e., 0.7, 1.1, 2.0, and 5.1 kb), 3 (i.e., 1.7, 2.5, and 4.2 kb), and 4 (i.e., 1.4, 1.5, 2.5, and 3.2 kb) overlapping PCR fragments, respectively. These genomes had been previously identified from short PCR fragments AE6/CP6108 (candHPV89), JC9710 (candHPV90), and JC9813 (candHPV91) [3, 6–8]. The assembled sequences of the viral genomes revealed a total size of 8078 bp for candHPV89, 8033 bp for candHPV90, and 7966 bp for candHPV91. The GenBank accession numbers of the sequences are AF436128 (candHPV89), AY057438 (candHPV90), and AF419318 (candHPV91). The nucleotide sequences of the candHPV89, candHPV90, and candHPV91 L1 ORFs were <90% related to other HPV types and thus were designated as distinct types. The DNA clones and sequences were submitted to the Human Papillomavirus Reference Laboratory (Heidelberg, Germany). The genomes of candHPV89, candHPV90, and candHPV91 displayed the characteristic ORF distribution found in other sequenced HPV types. Table 1 shows the homology of putative proteins of candHPV89, candHPV90, and candHPV91 to the analogous proteins of several closely related HPV types. The candHPV89, candHPV90, and candHPV91 ORFs were most closely related to HPV83, HPV71, and HPV7 ORFs, respectively. To determine whether 2 genomes showed similar degrees of homology across different regions of their genomes, ratios of relatedness were calculated. These ratios provide evidence for a similar degree of relatedness across the genomes of the A3 and A8 groups. candHPV90 and HPV71 showed strikingly disparate ratios in the L1 ORF region (http://hpv-web.lanl.gov/stdgen/virus/hpv/compendium/htdocs/HTML_FILES/HPVcompintro4.html#comp97). To determine whether this difference was driven by an unequal accumulation of change throughout the L1 ORF region of one or both of these HPVs, relative ratios were calculated using the L1 ORF of

<table>
<thead>
<tr>
<th>Novel HPV, related HPV type (group)*</th>
<th>E6</th>
<th>E7</th>
<th>E1</th>
<th>E2</th>
<th>E4</th>
<th>L2</th>
<th>L1</th>
<th>L1:E1 ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>candHPV89</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPV83 (A3)</td>
<td>67.1</td>
<td>58.2</td>
<td>77.7</td>
<td>62.2</td>
<td>58.4</td>
<td>71.1</td>
<td>82.7</td>
<td>1.064</td>
</tr>
<tr>
<td>HPV61 (A3)</td>
<td>60.7</td>
<td>62.2</td>
<td>76.4</td>
<td>63.7</td>
<td>65.1</td>
<td>64.2</td>
<td>79.0</td>
<td>1.034</td>
</tr>
<tr>
<td>HPV72 (A3)</td>
<td>63.1</td>
<td>59.6</td>
<td>76.9</td>
<td>62.6</td>
<td>61.2</td>
<td>67.1</td>
<td>79.0</td>
<td>1.027</td>
</tr>
<tr>
<td>candHPV86 (A3)</td>
<td>61.0</td>
<td>55.1</td>
<td>73.0</td>
<td>60.5</td>
<td>53.7</td>
<td>70.3</td>
<td>79.1</td>
<td>1.084</td>
</tr>
<tr>
<td>HPV84 (A3)</td>
<td>62.6</td>
<td>53.1</td>
<td>73.7</td>
<td>62.4</td>
<td>56.4</td>
<td>69.9</td>
<td>78.8</td>
<td>1.069</td>
</tr>
<tr>
<td>candHPV87 (A3)</td>
<td>58.5</td>
<td>56.1</td>
<td>74.1</td>
<td>63.1</td>
<td>56.3</td>
<td>69.9</td>
<td>78.3</td>
<td>1.057</td>
</tr>
<tr>
<td>candHPV90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPV71 (A15)</td>
<td>65.5</td>
<td>71.1</td>
<td>80.1</td>
<td>63.6</td>
<td>58.5</td>
<td>70.1</td>
<td>69.2</td>
<td>0.864</td>
</tr>
<tr>
<td>candHPV91</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPV7 (A8)</td>
<td>68.4</td>
<td>54.5</td>
<td>74.2</td>
<td>58.5</td>
<td>38.2</td>
<td>66.5</td>
<td>79.6</td>
<td>1.073</td>
</tr>
<tr>
<td>HPV40 (A8)</td>
<td>65.8</td>
<td>53.6</td>
<td>73.9</td>
<td>60.5</td>
<td>41.4</td>
<td>66.1</td>
<td>78.2</td>
<td>1.058</td>
</tr>
</tbody>
</table>

* See reference [3].
HPV71 and candHPV90, compared with HPVs of other groups (data not shown). The relative homologies for the HPV71 L1 ORF, but not for candHPV90, consistently gave scores lower than expected.

Phylogenetic analysis of 3 novel HPVs. To investigate the relationship between the 3 novel HPVs and related HPV genomes, the predicted amino acid sequences (i.e., E1, E2, E4, E6, E7, L1, and L2) were aligned with the corresponding sequences of HPVs from homology groups A3, A8, and A15 [3]. The resulting phylogenetic tree was calculated and based on the compiled amino acid sequences from HPV7, HPV40, HPV61, HPV71, HPV72, HPV83, HPV84, candHPV86, and candHPV87 [11]. Representative trees are shown in figure 1. The trees indicated that candHPV89 was most closely related to HPV83 and was placed into group A3 [3]. The genome of candHPV91 was most closely related to HPV77 and HPV90 and was placed into group A8 [3]. Although the compiled genome analysis placed candHPV90 and HPV71 into a single lineage (figure 1A), phylogenetic analysis restricted to the L1 ORF did not place them in a single lineage (figure 1B).

Discussion

We describe the complete characterization of 3 novel genital HPV genomes. candHPV89 includes sequences homologous to those derived from MY09/MY11 fragments AE6 and CPE608 [3, 6]. Since candHPV89 is well amplified by the MY09/MY11 primers, it has been detected in a large number of epidemiologic studies [10, 12–14]. The highest prevalence (1.8%) was detected in women from New Mexico [14], whereas in other populations it tended to be quite rare and was never associated with high-grade squamous intraepithelial lesions. The candHPV89-containing phylogenetic branch within the A3 clade includes HPV84 and candHPV86. HPV84 is a highly prevalent type detected in normal and human immunodeficiency virus (HIV) type 1–infected women [5]. Thus, on the basis of phylogenetic tree analysis and clinical data, it appears that candHPV89 is associated with a low risk of cervical cancer. candHPV90 and candHPV91 include sequences homologous to HPV PCR fragments JC9710 and JC9813, respectively [7, 8]. There are few reports about the prevalence and disease association of candHPV90 and candHPV91 since, similar to candHPV86, they are not amplified with the MY09/MY11 PCR system [4]. All 3 viruses were identified from samples that were negative for HPV DNA by MY09/MY11 PCR but positive by low-stringency Southern blot analysis. By use of direct PCR sequencing, JC9710 (i.e., candHPV90) was detected in condylomatosus specimens from an HIV-positive French-Canadian woman, whereas JC9813 (i.e., candHPV91) was detected in low-grade squamous intraepithelial lesion specimens from an HIV-positive English-Canadian woman [7, 8]. In this report, candHPV90 and candHPV91 were isolated from cervicovaginal cells from young women with normal cytology [9, 10]. candHPV91 is closely related to HPV77 and was enigmatically positioned in the A8 group. Although HPV91 was initially reported to be associated with cutaneous benign butcher’s warts in meat and fish handlers, it has also been isolated from unusual oral mucosal lesions in HIV-positive patients [15].

To evaluate the relationship of related HPVs in a group, we calculated the ratio of relatedness in L1:E1 ORFs, using the homology of candHPV89, candHPV90, and candHPV91 amino acid sequences to their related HPVs. A ratio of >1.0 in A3 and A8 HPVs is consistent with the L1 ORF being the most conserved region. Although the ratios of A3 and A8 HPVs appear to be similar, the ratio between candHPV90 and HPV71 indicated a more distant relationship and raised the question of whether they form a single taxonomic group. To evaluate whether HPV71 and/or candHPV90 might have some unique characteristics, we calculated additional L1:E1 ratios with these viruses, compared with those for other HPVs (data not shown). In all cases, candHPV90 gave scores >1.0, which is similar to scores for other HPVs; however, HPV71 gave consistently lower scores. These data indicate that the HPV71 L1 ORF is relatively more divergent than other HPV71 ORFs, when compared with divergence among other HPV genomes. We interpret these results to indicate that the L1 ORF of HPV71 has undergone natural selection.

Figure 1. Phylogenetic trees based on the alignment of the amino acid sequences of the compiled open reading frames (ORFs; E1, E2, E4, E6, E7, L1, and L2) (A) and the amino acid sequences of the L1 ORF (B) of the indicated human papillomavirus (HPV) genomes.
Novel HPV types continue to emerge, and the clinical significance of each type needs to be evaluated empirically. In this paper, candHPV89, candHPV90, and candHPV91 were cloned from clinical specimens by use of the overlapping PCR method [4, 5]. Detailed phylogenetic analyses revealed that these HPVs cluster with other low-risk HPVs. candHPV89 is detected relatively infrequently in the female genital tract, but little is known about candHPV90 and candHPV91, since they are not amplified by MY09/MY11 primers.

Acknowledgment

Assignment of the HPV type number was kindly performed by Ethel-Michele de Villiers (Human Papillomavirus Reference Laboratory, Heidelberg, Germany).

References