The clinical importance of the nomenclature, evolution and taxonomy of human papillomaviruses

Hans-Ulrich Bernard *

Department of Molecular Biology and Biochemistry, 114 Sprague Hall, University of California, Irvine, CA 92697-3900, USA

Received 22 June 2004; accepted 14 October 2004

Abstract

Human papillomaviruses (HPVs) are formally described by isolation of their circular double-stranded DNA genomes and establishment and comparison of the nucleotide sequence of these genomes. Alternatives such as serological diagnosis and maintenance of HPVs in culture are neither clinically useful nor consistently feasible. Novel HPV isolates have traditionally been described as "types". The analysis of specific HPV types is of medical importance, because HPV types typically induce type-specific lesions, i.e. they may be specific for cutaneous or mucosal epithelia, or give rise to benign warts or malignant carcinomas. Recently, it was formally decided that papillomaviruses are a virus family separate from the polyomaviruses. Within the papillomavirus family, closely or remotely related types form species or genera. These formal agreements were important as they brought the taxonomy of papillomaviruses in line with that of other viruses, bacteria and higher organisms, although their impact on medical practice and terminology used in clinical studies is limited. Notably, however, HPV types that are closely related (i.e. form "species") are associated with similar lesions. Confusion of the terms "type" and "subtype" should be avoided, as the latter term refers to some specific but rare taxonomic assemblages. In contrast to many RNA viruses, HPV types evolve very slowly, and diverged since the origin of humans only by about 2%. These divergent isolates are called "variants". HPVs evolved together with humankind and Homo sapiens was never without HPVs, and consequently never without warts and cervical cancer. Variants of the same HPV type may have different pathogenicity and may account for part of the worldwide disparities in the occurrence of genital cancers.

Keywords: Papillomaviridae; Virus nomenclature; Taxonomy of HPV; Phylogeny

Contents

1. Introduction .......................................................... S2
2. Papillomaviruses, a separate virus family .................. S2
3. Nomenclature ......................................................... S2
   3.1. Papillomavirus types ......................................... S2
   3.2. Subtypes ....................................................... S3
   3.3. Variants .......................................................... S3
4. Taxonomy above the level of HPV types ..................... S3
5. The relationship between taxonomy, biology, and pathology ..................................................... S3
6. Taxonomy and diagnosis ........................................... S5
7. Taxonomy, vaccination, and treatment ...................... S5
References ............................................................ S5

Abbreviations: HPV, human papillomaviruses; PCR, polymerase chain reaction; ICTV, International Council on Taxonomy of Viruses; EV, epidermodysplasia verruciformis
* Tel.: +1 949 824 5162; fax: +1 949 824 8551.
E-mail address: hbernard@uci.edu.

1386-6532/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
1. Introduction

Until the late 1970s, papillomaviruses attracted little interest as they were only known as the causal agents of “warts”, benign cutaneous lesions in some mammals and humans. Since warts are normally only a cosmetic problem but not a major threat to public health, they were of only academic interest as human viruses inducing neoplasia. In the 1980s, the newly developed powerful techniques of molecular biology led to the detection of dozens of human papillomaviruses (HPVs) in benign and malignant mucosal lesions (Gissmann et al., 1977; zur Hausen, 2002), such as cervical cancer and its precursor lesions, as well as in genital and laryngeal warts, setting the stage for a major expansion of papillomavirus research. Present data support the existence of more than 200 HPV types. The whole genomes of about 100 HPV types have been isolated and completely sequenced, while we have only indirect evidence for the existence of other types, most often through the sequence of polymerase chain reaction (PCR) amplicons, which unambiguously identify HPV derivatives. Description of HPV types has changed several times in parallel with technical progress and phylogenetic analyses. An HPV type is now defined as a complete genome, whose L1 gene sequence is at least 10% dissimilar to that of any other HPV type. This definition was introduced somewhat arbitrarily, but subsequent research had the fortunate outcome that the 10% distance criterion describes natural taxa (de Villiers et al., 2004). HPV genomes evolve as slowly as the genomes of their hosts. Since sequence distances between HPV types exceed those of the distances of related mammals, even closely related HPV types should be viewed as molecularly quite divergent, for example as distant as humans and chimpanzees.

2. Papillomaviruses, a separate virus family

In the mid-1950s to 1960s, when papillomaviruses and polyomaviruses became amenable to observation by electron microscopy and basic nucleic acid analyses, these two groups were found to be the only viruses that had (1) double-stranded circular DNA genomes, and (2) non-enveloped particles consisting of icosahedral capsids. As a consequence, they were considered closely related and were placed into a common family, the papovaviridae (Papovaviridae). Sequence and functional studies of the 1980s (Danos et al., 1982) showed that these similarities were too superficial to establish relationship. All papovaviruses have genome sizes around 5 kb, while those of papillomaviruses are close to 8 kb. Papillomaviruses have two transcription units that read toward one another, while papillomavirus transcription occurs only in one direction. And lastly, and most importantly, polyomaviruses and papillomaviruses do not share any substantial amount of nucleotide or amino acid sequence similarity, with the exception of a small homologous segment in their T-antigen and E1 genes, respectively (Rebrinkov et al., 2002). Since taxonomic classifications should reflect natural relationships, it was concluded that these viruses form two separate families. Only fairly recently, the family “papillomaviruses” (Papillomaviridae) became officially recognized by the International Council on Taxonomy of Viruses (ICTV) (de Villiers et al., 2004).

3. Nomenclature

3.1. Papillomavirus types

Papillomaviruses are identified by the abbreviation PV and one or two letters indicating the host species. This can be derived from an English word, for example “HPV” for human papillomaviruses and “CRPV” for cottontail rabbit papillomavirus, or the scientific name of the host, e.g. MnPV for Mastomys natalensis papillomavirus, which infects an African rat. HPV types are identified by numbers in the historic sequence of their description, e.g. HPV-1, HPV-2, etc. (Coggin and zur Hausen, 1979). Presently, and for the last 20 years, E.M. de Villiers at the Reference Center for Papillomaviruses at the German Cancer Research Center in Heidelberg has integrated this process (de Villiers, 2001) and new HPV types have to be registered with this center to confirm (i) completeness of the genomic isolate, and to show (ii) 10% nucleotide sequence diversity of the L1 gene of the isolate from all known HPV types. Thereafter, assignment of a new number and publication of this new HPV type is possible. HPV types, whose genome was generated by PCR rather than traditional cloning techniques, are identified by addition of the abbreviation “cand” (for candidate) before their number, e.g. candHPV-86 (de Villiers et al., 2004). Presumed HPV types that are only identified by partial PCR amplicons do not qualify as a numbered type, but are described with an abbreviation identifying the laboratory of origin (e.g. LVX100, for Luisa Villa X100) (de Villiers et al., 2004). All sequence information about papillomaviruses is not only available through central databases like GenBank, but also, in an annotated form in a special database that is publicly available on the internet (HPV Sequence Data Base; Farmer et al., 1995).

3.2. Subtypes

The term “subtypes” was used in the 1980s to identify isolates of HPV types with divergent restriction patterns. Subsequently, this term became redefined as referring to an isolate whose L1 sequence is 2–10% different from that of any known type. A consequence of this redefinition was that the traditional subtypes (e.g. HPV-6a, HPV-6b and HPV-6c) had to be eliminated, as they showed less than 2% sequence diversity. Surprisingly, today only three HPV isolates are known that fulfill this subtype definition. HPV-46, HPV-55, HPV-64 had been originally described as separate types, but their type status has now been cancelled, as they are subtypes...
of HPV-20, HPV-44, and HPV-34, respectively (de Villiers et al., 2004). It remains a mystery of academic interest why "subtype" genomes, which once must have existed in the process of evolution, are today apparently virtually absent.

3.3. Variants

“Variants” of HPV types differ by about 2% of the original isolate, which is in this context referred to as “prototype” or “reference genome”, while the new variants can be designated by arbitrary abbreviations. In order to identify a variant, one typically sequences a small part of the HPV genome, for example 400 bp of the long control region or the 450 bp of the E6 gene, and calls a variant every isolate that differs from the prototype by at least one nucleotide. This strategy has been applied to numerous HPV types (Ho et al., 1993; Ong et al., 1993; Heinzl et al., 1995; Stewart et al., 1996; Yamada et al., 1997; Chan et al., 1997) from isolates throughout the world. The two principal observations from these studies were that (i) there is apparently only a limited number (for example 20–100) of common variants of each HPV type, and (ii) that variants showed maximal divergence when they were sampled from ethnic groups that evolved for a long time without contact, for example Africans and American Indians. From this, one has to conclude that each HPV type was with the human species since our origins, and evolved and spread together with the infected cohort (Bernard, 1994). There was never a time or an ethnic group without warts, genital warts, and cervical cancer, and HPVs did not infect humans from an animal reservoir such as some other viruses, like HIV-1, Ebola, the SARS coronavirus, or influenza. There are indications that variants of the same HPV type differ biologically and etiologically (Xi et al., 1998; Villa et al., 2000). Such differences may contribute to the disparities in the incidence of cervical cancer throughout the world, although this question still requires substantial research before conclusions can be drawn (Calleja-Macias et al., 2004).

4. Taxonomy above the level of HPV types

Modern taxonomy aims to understand the relationship of biological objects such as viruses based on their evolutionary history rather than on phenotypic characteristics. This concept is relevant for HPVs, since unexpectedly, it was found that different HPV types associated with similar lesions are sometimes only very distinctly related to one another. For example, HPV-1, HPV-2, HPV-4 and HPV-41 are all found in wart-like cutaneous lesions, but are on remote branches of evolutionary trees. HPV-16 and HPV-18, the two HPV types that have become paradigms in the research on cervical carcinogenesis, are less related to one another than to some HPV types that are never found in cervical malignancies. Sequence studies showed that groupings originally based on pathology ("genital HPVs" such as HPV-16) contained, based on sequence similarity, closely related HPV types that occurred in cutaneous lesions, such as HPV-2. Research in numerous laboratories over the last decade confirmed that the sequence comparison of whole papillomavirus genomes, individual genes or even short segments of genes leads to phylogenetic trees that summarize the evolutionary history of papillomaviruses as a prerequisite to establish taxonomic groups (Chan et al., 1995). A complete list and the taxonomic relationship of all known and completely described papillomaviruses together with a phylogenetic tree that shows the relationship of all papillomavirus types, species, and genera can be accessed in a recent publication from de Villiers et al. (2004).

5. The relationship between taxonomy, biology, and pathology

It was noted more than 10 years ago (van Ranst et al., 1992) that related HPV types lead to similar lesions, but problems with interpretation of phylogenetic trees required a refinement of the original proposal (Chan et al., 1995). A consensus was reached by the working group on the taxonomy of papillomaviruses and the ICTV (de Villiers et al., 2004). Table 1 is a shortened list extracted from that publication, which summarizes those papillomavirus types and their taxonomy that are relevant to understand most clinical, epidemiological or molecular publications.

In the new taxonomy of papillomaviruses, higher phylogenetic assemblages of papillomavirus types are termed “genus” and lower assemblages “species”. The presently known human and animal papillomaviruses form 16 genera, which are identified by Greek letters. Five of these genera are composed exclusively of HPV types and papillomaviruses identified in some apes and monkeys, all other genera contain types found in various mammals and birds. The clinically most important genus is referred to as the alpha-papillomaviruses. It contains all HPV types associated with mucosal and genital lesions and emerged from a grouping originally referred to as “genital” or “mucosal” HPVs. The neutral designation “alpha-papillomaviruses” is superior to these two terms, which led to many confusing contradictions and were in need of revision. Examples of these contradictions include:

- HPV-6 and HPV-11 are typically found in genital warts or condylomata acuminata of the cervix, and they were therefore considered “genital” HPVs. However, they are also found in non-genital sites, for example in papillomas of the larynx.
- Genital warts can be of mucosal as well as cutaneous origin, making the term “mucosal” inappropriate.
- Based on phylogenetic criteria, the former “genital” and “mucosal” HPV types also contained HPV-2, HPV-27, and HPV-57, which are the cause of most common warts, and HPV-7, traditionally known as an HPV type causing
Table 1
The most frequently studied papillomavirus types and their biological and clinical properties

<table>
<thead>
<tr>
<th>Family: papillomaviruses (Papillomaviridae)</th>
<th>Genus</th>
<th>Species</th>
<th>Type(s)</th>
<th>Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Alpha-papillomaviruses</strong></td>
<td>4</td>
<td>HPV-2, HPV-27, HPV-57</td>
<td>Common skin warts, frequently in genital warts of children</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>HPV-26, HPV-51, HPV-69, HPV-82, HPV-84</td>
<td>High-risk malignant and benign mucosal lesions</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>HPV-33, HPV-35, HPV-56, HPV-66</td>
<td>High-risk malignant and benign mucosal lesions</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>HPV-18, HPV-39, HPV-45, HPV-59, HPV-68, HPV-70</td>
<td>High-risk malignant mucosal lesions, some (e.g., HPV-18) more frequent in adenocarcinoma of the cervix</td>
<td></td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>HPV-7, HPV-46, HPV-43</td>
<td>Low-risk mucosal and cutaneous lesions, HPV-7 known as &quot;butcher's wart virus&quot;, often in lesions of HIV infected patients</td>
<td></td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>HPV-16, HPV-31, HPV-33, HPV-35, HPV-52, HPV-58, HPV-65</td>
<td>High-risk malignant mucosal lesions, some (e.g., HPV-16) more frequent in squamous than in adenocarcinoma of the cervix, HPV-16 most prevalent HPV type in cervical malignancies</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>HPV-6, HPV-11, HPV-13, HPV-44, HPV-74</td>
<td>Benign mucosal lesions. HPV-6 and HPV-11 in male and female genital warts, condylomata acuminata of cervix, laryngeal papillomas. Some of these lesions can progress malignant</td>
<td></td>
</tr>
</tbody>
</table>

| **Beta-papillomaviruses**                  | 1     | HPV-5, HPV-8 (selected from a very type-rich genus) | Cutaneous benign and malignant lesions in EV and immune-suppressed patients |
|                                           | 2     | HPV-6, HPV-65 (selected from a very type-rich genus) | Cutaneous benign lesions |

| **Gamma-papillomaviruses**                | 1     | HPV-4, HPV-65 (selected from a very type-rich genus) | Fibropapillomas in cattle, sarcoïds in horses. An important cell culture model. |

| **Delta-papillomaviruses**                | 1     | Bovine papillomavirus-1 (BPV-1) (selected from a type-rich genus) | Cutaneous lesions. An important animal model |

| **Kappa-papillomaviruses**                | 1     | Cottontail rabbit papillomavirus (CRPV) | Cutaneous lesions, frequently in footworts. |
| **Mu-papillomaviruses**                  | 1.2   | HPV-1, HPV-63 | Cutaneous lesions. |
| **Nu-papillomaviruses**                  | 1     | HPV-41 (unrelated to any other HPV type) | Papillomas of the alimentary canals. BPV-4 is an important model for multi-step carcinogenesis and vaccination research. |
| **Xi-papillomaviruses**                  | 1     | BPV-1, BPV-4, BPV-4 | Papillomas of the alimentary canal. BPV-4 is an important model for multi-step carcinogenesis and vaccination research. |

The papillomaviruses were recently recognized by the International Council on the Taxonomy of Viruses (ICTV) as an independent virus family, papillomaviridae, which also introduced the taxonomic levels "genus" and "species" between the familiar description of these viruses as "types".

In summary, no generalizations can be made regarding pathological associations of the genus alpha-papillomaviruses, and terms like "genital" HPV are useful in vernacular language, but scientifically and clinically fairly useless.

On the phylogenetic level below the genus, closely related HPV types are now lumped into species, and here one finds a clear correlation between taxonomy and pathogenicity, as recognized earlier (van Ranst et al., 1992; Chan et al., 1995). The term "species" will likely not become broadly used by the community of HPV researchers, but it had to be introduced following a rule of the ICTV that a virus can only be considered a species if it has biologically distinct properties. On the phylogenetic level below the genus, closely related HPV types are now lumped into species, and here one finds a clear correlation between taxonomy and pathogenicity, as recognized earlier (van Ranst et al., 1992; Chan et al., 1995). The term "species" will likely not become broadly used by the community of HPV researchers, but it had to be introduced following a rule of the ICTV that a virus can only be considered a species if it has biologically distinct properties. Based on their genomic sequences, all HPV types of the genus alpha-papillomavirus form 15 species, and for each of these there are data about functional distinctions. Twelve of the 15 HPV types that are classified by a recent epidemiological meta-study as "high-risk" HPV types due to their association with cervical carcinogenesis (Munoz et al., 2003) are members of two species, HPV-species 7 (HPV-18, HPV-39, HPV-45, HPV-59, and HPV-68) and HPV-species 9 (HPV-16, HPV-31, HPV-33, HPV-35, HPV-52, HPV-58, and HPV-67). The overlap is not complete, since the prevalence of some HPV types in most epidemiological and etiological studies is not sufficiently high to generate sufficient statistical power to elevate all HPV types in species 7 and 9 to "high-risk" types, nor to raise yet other species and types into the "high-risk" category (Matsukura and Sugase, 2001) (for some details, see Table 1).

It will take some time for scientists and clinicians, to embrace the concept of HPV "species" in scientific publications or clinical practice, but the notion of the relationship of these HPV types is certainly useful: Only few HPV types, for example HPV-16, HPV-18 and HPV-31, have been intensively researched. They can be considered paradigms, and their molecular and medical properties can be extrapolated to their relatives.

The beta-papillomaviruses include all HPV types associated with epidermodysplasia verruciformis (EV), a cutaneous neoplastic disease with a genetic component (Orth et al., 1978; Rumooz et al., 2002). In carriers who are not genetically predisposed, beta- and related gamma-papillomaviruses normally seem to establish asymptomatic infections, or, at worst induce small benign cutaneous neoplastic lesions. They are apparently very widespread and contain many HPV types...
that are not yet formally described (Antonsson et al., 2000). Some of the viruses in these two genera have also been found in association with skin cancer in immune suppressed individuals, but a potential causality between HPV infection and malignant neoplasia is still under investigation (de Villiers, 1998).

While the alpha-, beta- and gamma-papillomaviruses are extraordinarily rich in different types, no close relatives of HPV-1, HPV-41, and HPV-63 have been found, which form two separate genera (Table 1). All other genera are formed by mammal and bird papillomaviruses and since papillomaviruses have evolved in linkage with their host species, a papillomavirus identified in any little studied animal group normally forms a new genus. As only a small number of mammals and birds have been investigated for the occurrence of papillomaviruses, one can predict that the search for papillomaviruses in novel host species will vastly expand the number of genera and create a challenge for the use of the Greek alphabet in the present taxonomy.

6. Taxonomy and diagnosis

For details of the role of HPV testing in cervical cancer screening and techniques for detection and typing, see Cuschieri and Cubie (2005) and Molijn et al. (2005). In brief, the Papanicolaou test, a cytological staining technique developed prior to the detection of HPVs, continues to be the diagnostic gold standard to detect cells derived from cervical carcinomas and its precursors. However, cytology can now be complemented with powerful procedures that detect HPV DNA, and epidemiological studies are defining criteria where these molecular techniques may enter standard gynecological practice. Some of the emerging commercial tests have their roots in efforts to amplify HPV genomes selectively by the polymerase chain reaction (PCR) with “consensus primers”, together with oligonucleotides complementary to a diverse range of HPV types and which were originally designed to study the genomic diversity of HPVs (Bernard et al., 1994; Jacobs et al., 1995). As HPV types that cause genital neoplasia were always central to the search for new HPV types, these techniques were biased to detect alpha-papillomaviruses. Today’s commercial kits refine this approach either by combining amplification with type-specific DNA hybridization to discriminate between the major HPV types that infect the genital mucosa or by distinguishing “high-risk” HPV types which are normally found in cervical carcinoma (found in HPV-species 7 and 9) from those that rarely progress to malignancy (included in HPV-species 10).

7. Taxonomy, vaccination, and treatment

The extensive study of papillomaviruses over the last two decades had a remarkable impact on our understanding of the etiology of HPV associated lesions. Unfortunately, it did much less to improve the clinical practice than one could have hoped for, based on the success of studying infections by other viruses, which were more amenable to vaccination and to the development of antiviral drugs (Waugh et al., 2002; Bernard, 2004). Improvement of DNA diagnosis and recent breakthroughs in prophylactic vaccination is changing this situation (Koutsky et al., 2002). However, as one weighs the tremendous molecular diversity against the few similarities between HPV types, it is not easy to recognize an “Achilles heel” of this virus group. For example, the lack of immunological cross-reaction among the capsids of the numerous high-risk HPV types may make the formulation of broad-spectrum vaccination protocols a challenge. The same lack of sequence similarity may make it difficult to target the early proteins by therapeutic vaccination and drug treatment. While sequence diversity is clearly an obstacle to combat these viruses, functions and structures that are shared across all HPV types may help to focus on shared targets, such as the unique structure of the zinc fingers of the E6 and E7 oncoproteins (Beerheide et al., 1999; Munger and Howley, 2002) and the helicase function of the E1 replication protein (Underwood et al., 2000).

References


