While studying the virome of the skin surface of a patient with a Merkel cell carcinoma (MCC) by using unbiased, high-throughput sequencing, we identified a human polyomavirus nearly identical to human polyomavirus 9, a virus recently reported in blood and urine of renal transplantation patients and closely related to the African green monkey lymphotropic polyomavirus. Specific PCR analysis further identified this virus in 2/8 patients with MCC but in only 1/111 controls without MCC. This virus was shed for >20 months by the MCC index patient and was on the skin of the spouse of the index patient. These results provide information on the viral ecology of human skin and raise new questions regarding the pathology of virus-associated skin disorders.

Polyomaviridae is a family of nonenveloped viruses with a circular double-stranded DNA genome. Natural hosts for Polyomaviridae are primates, including humans and monkeys, cattle, rabbits, rodents, and birds (1). Currently, viruses in this family that infect humans are the opportunistic JC polyomavirus (JCPyV) associated with progressive multifocal leukoencephalopathy in immunocompromised person; BK polyomavirus (BKPyV) associated with interstitial nephritis and hemorrhagic cystitis; KI polyomavirus (KIPyV) identified in respiratory secretions of patients with respiratory symptoms at the Karolinska Institute (Stockholm, Sweden); WU polyomavirus (WUPyV), isolated from patients with similar symptoms at Washington University (St. Louis, MO, USA); Merkel cell polyomavirus (MCPyV), associated with the rare but aggressive Merkel cell carcinoma (MCC); and trichodysplasia spinulosa–associated polyomavirus (TSPyV), associated with a rare skin condition in immunocompromised persons (2).

Two recently identified polyomaviruses, human polyomaviruses 6 and 7 (HPyV6 and HPyV7) have been detected in cutaneous swab specimens of healthy persons (3). HPyV9 has been identified by consensus PCR in blood and urine of asymptomatic renal transplant recipients (4). In this study of the skin virome of a patient with MCC, using high-throughput sequencing (HTS) and comparing sequences from a patient with MCC with sequences from healthy controls, we identified a human polyomavirus strain nearly identical to HPyV9, a virus species closely related to the lymphotropic polyomavirus (LPV).

Methods

Patients and Sample Collection
For analysis by HTS, 6 DNA samples extracted from cutaneous swabs obtained from the skin surface of facial areas (forehead and eyebrows) of patients previously studied by PCR for MCPyV sequences were selected (5). These samples included 1 from an index patient with an MCC on his elbow and 5 from the skin of 5 healthy persons. For investigation of prevalence by specific nested PCR, 120 skin specimens were similarly obtained from 120 volunteers. The median age of these persons was 48 years (range 19–96 years); 30 persons were 57–96 years of age (median age 71 years). This group of 120 volunteers was composed of 40 patients hospitalized or attending outpatient
The entire sequence of the Institut Pasteur polyomavirus (IPyV) strain genome was analyzed and annotated by using CLC Genomics Workbench (CLC Bio, Aarhus, Denmark). GenBank reference sequences of other members of the family *Polyomaviridae* used were JCPyV (NC_001699), BKPyV (NC_001538), KIPyV (NC_009238), WuPyV (NC_009539), MCPyV (NC_010277), SV40 (NC_001669), TSPyV (NC_014361), and LPV (M30540). Protein structures were visualized by using Pymol (Delano Scientific LLC, San Francisco, CA, USA).

**Phylogenetic Analysis**

Phylogenetic reconstructions were based on separate analyses of nucleotide sequences from viral protein 1 (VP1) and large T antigen (LT). A 974-nt region of monkey B-lymphotropic papovavirus (reference sequence M30540.1 from the VP1 coding sequence) was aligned with corresponding regions from the polyomaviruses available in GenBank. For the LT matrix, a 1,453-nt region (same reference sequence as for VP1) was used for analysis. Sequences were aligned by using SeaView version 4 (6) and the Muscle algorithm (7). Only partial but contiguous parts of each alignment were included in final matrices because alignment in some regions was not possible. The 742–1715 and 2904–4356 nt regions were included in matrices. The matrices were based on monkey B-lymphotropic papovavirus M30540 and African green monkey polyomavirus NC_004763 sequence annotations for VP1 and LT genes, respectively.

Phylogenetic analyses were performed by using a probabilistic (Bayesian) approach implemented in BEAST (8). Matrixes were tested against 88 substitution models by using jModelTest software (9). On the basis of results obtained with jModelTest software, the generalized time reversible substitution model (with invariant sites and a gamma site heterogeneity distribution) was used for analysis. The 3-codon partition model of evolution and the Yule speciation process were also specified as priors. A Markov chain Monte Carlo method was used to set 30,000,000 states to obtain an adequate posterior effective sample size >300. Pertinence of nodes was evaluated by using posterior probabilities. Sequences used in phylogenetic reconstructions were obtained from the National Center for Biotechnology Information (NCBI) database (www.ncbi.nlm.nih.gov/nucleotide). Confidence intervals for proportions were calculated according to the efficient-score method (corrected for continuity) (10) (http://dogsbody.psych.mun.ca/VassarStats).

**PCR**

For sequencing the IPyV by the Sanger method, 9 primer pairs were designed to amplify the entire genome by reference to the contigs assembled from HTS data.
acquired in the first phase (see Results). Primer sequences and protocols are available upon request. After the genome was sequenced, we developed a specific nested PCR for detection of IPPyV in samples by using primers based on the IPPyV genome sequence and designed by using PrimerPro 3.4 software (www.changbioscience.com): VP1_354F (5′-ACCATATCAGTAGGATAGGTA-3′) and VP1_354R (5′-TGAATTGTATGGCTACAGTGC-3′) for the outer PCR, and VP1_198F (5′-CACTGGGATAGTTCCTGAGG-3′) and VP1_198R (5′-CTATGCTACTACCCTCCTC-3′) for the inner PCR. These primers were designed to avoid amplification of other known human polyomaviruses.

Ethical Approval
The study was reviewed and approved by the Institut Pasteur Comité de Recherche Clinique and the French Commission Nationale Informatique et Libertés (09.465). Consent was provided by participants for obtaining human samples according to French regulations.

Results

Identification of the IPPyV Strain of HPyV9
Using 8,052,770 Illumina reads obtained from DNA extracted from the skin surface of the MCC index patient, we assembled the complete genome of MCPyV. We found numerous papillomavirus contigs and contigs covering more than half of the genomes of HPyV6 and HPyV7. Additionally, 14 other contigs were assembled that showed a better homology with LPV (NCBI accession no. M30540, version M30540.1, GI:333282) than with any other virus present at that time in the NCBI database, including other human or animal members of the family Polyomaviridae. On the basis of sequence of 8 of the 14 obtained contigs, which were distributed along the LPV genome, we defined a set of 9 primer pairs encompassing the entire target genome. These primers enabled amplification of the entire genome by PCR and analysis of its sequence of 5,028 nt by using the Sanger method, which confirmed the circular nature of the genome.

Whole genome organization of the IPPyV strain, which exhibits general molecular characteristics of polyomaviruses, is shown in Figure 1. It encodes analogs of small T antigen, LT, and structural proteins VP1, VP2, and VP3 and does not appear to encode an agnoprotein. Pairwise amino acid identity was 100% between IPPyV and HPyV9 proteins, 72%–80% between IPPyV and LPV proteins, and much lower for other known family Polyomaviridae members (Table). Because the nucleotide sequence of IPPyV is nearly identical to that of HPyV9, with a difference of only 2 nt in a noncoding region at nt 4449, it appears that IPPyV should be considered a strain of HPyV9. Its sequence have been deposed in GenBank (accession no. FR823284).

Phylogenetic Analysis
Reconstructions of VP1 and LT phylogenies on the basis of nucleotide sequences clustered the mammalian polyomaviruses and placed the species Avipolyomavirus in basal position when rooting with the oldest known Avipolyomavirus (Budgerigar fledging virus) (Figure 2). Despite the divergences described below, VP1 and LT of HPyV9 were closely related to those of the monkey B-lymphotropic polyomavirus. Moreover, VP1 and LT phylogenies consistently identify several monophyletic groups among mammalian polyomaviruses (Figure 2). Nevertheless, the incongruence of VP1 and LT signals induce notable differences in the topology of these 2 phylogenies. For instance, LT of bovine polyomavirus is closely related to one of the common ancestors of all other mammalian polyomaviruses, whereas its VP1 is closely related to VP1 of the sea lion polyomavirus. Incongruence between VP1 and T phylogenies has been observed for HPyV6 and HPyV7. Nomenclature described in proposals of the International Committee on Taxonomy of Viruses is shown in Figure 2, even though the species Orthopolyomavirus is not monophyletic and therefore should be considered cautiously.

Comparison of VP1 from LPV and HPyV9
We compared the secondary structure of VP1 from LPV and HPyV9 because the external capsid protein of Polyomaviridae is known to interact with the cell
Human Polyomavirus Related to Monkey Polyomavirus

Receptor and because antibodies cross-reacting with LPV VP1 have been detected in a large proportion of humans. Overall amino acid identity was 87.1%. The VP1 monomer consists of antiparallel \(\beta\)-strands folded into a jelly roll \(\beta\)-barrel structure. Three outer loops (BC, DE, and HI) are exposed outside the pentamer core and are most likely recognized by antibodies. Using the crystal structure of SV40 VP1 (3BWQ), we mapped the amino acids that differed between the 2 proteins. Polymorphic residues are present in the 3 VP1 loops, and the BC and HI loops appear more conserved than DE loop, which thus shows the major differences (Figure 3).

Detection of HPyV9 in Human Samples

We first confirmed by specific nested PCR the presence of HPyV9 in the skin swab specimen of the index patient in which the virus had been identified by HTS. We also identified this virus by nested PCR in a second cutaneous sample from the same index patient case obtained 20 months after the first sample. Because HPyV9 was identified in a patient who had MCC, we analyzed the skin surface of 7 other MCC patients. HPyV9 was detected in 1 patient 80 years of age. The overall prevalence in the MCC group was thus 2/8 (25%, 95% confidence interval 4.4%–64%; \(p = 0.05\)).

Table. Amino acid identity between putative proteins encoded by IPPyV and proteins of Polyomaviridae deduced by using pairwise sequence alignment

<table>
<thead>
<tr>
<th>Protein</th>
<th>Putative open reading frame</th>
<th>Frame</th>
<th>No. amino acids</th>
<th>JCV</th>
<th>BKV</th>
<th>KIV</th>
<th>WuV</th>
<th>MCyV</th>
<th>TSV</th>
<th>SV40</th>
<th>LPV</th>
<th>HPyV9</th>
</tr>
</thead>
<tbody>
<tr>
<td>VP1</td>
<td>1443–2558</td>
<td>+3</td>
<td>371</td>
<td>53.9</td>
<td>53.2</td>
<td>28.3</td>
<td>28.3</td>
<td>54.8</td>
<td>60.6</td>
<td>52.9</td>
<td>87.1</td>
<td>100.0</td>
</tr>
<tr>
<td>VP2</td>
<td>503–1561</td>
<td>+2</td>
<td>352</td>
<td>32.3</td>
<td>32.6</td>
<td>23.8</td>
<td>20.8</td>
<td>26.1</td>
<td>43.5</td>
<td>33.1</td>
<td>74.9</td>
<td>100.0</td>
</tr>
<tr>
<td>VP3</td>
<td>860–1561</td>
<td>+2</td>
<td>233</td>
<td>34.1</td>
<td>35.9</td>
<td>24.5</td>
<td>20.3</td>
<td>15.1</td>
<td>41.3</td>
<td>33.7</td>
<td>72.5</td>
<td>100.0</td>
</tr>
<tr>
<td>ST antigen</td>
<td>5028–4459</td>
<td>–1</td>
<td>189</td>
<td>35.1</td>
<td>34.0</td>
<td>39.5</td>
<td>34.6</td>
<td>40.1</td>
<td>42.5</td>
<td>31.8</td>
<td>81.0</td>
<td>100.0</td>
</tr>
<tr>
<td>LT antigen</td>
<td>5028–4792</td>
<td>–1</td>
<td>680</td>
<td>40.4</td>
<td>41.2</td>
<td>44.2</td>
<td>42.0</td>
<td>39.9</td>
<td>49.3</td>
<td>40.0</td>
<td>80.5</td>
<td>100.0</td>
</tr>
</tbody>
</table>

*IPPyV, Institut Pasteur polyomavirus; JCV, JC polyomavirus; BKV, BK polyomavirus; KIV, KI polyomavirus; WuV, Wu polyomavirus; MCyV, Merkel cell polyomavirus; TSV, trichodysplasia spinulosa–associated polyomavirus; SV40, Simian virus 40; LPV, lymphotrophic polyomavirus; HPyV9, human polyomavirus 9; VP, viral protein; ST antigen; small T antigen; LT antigen, large T antigen. Pairwise sequence alignment was performed by using EMBOSs Needle Software (http://emboss.sourceforge.net/apps/release/5.0/emboss/apps/needle.html).
Because we were interested in possible interhuman transmission of HPyV9, a skin swab specimen from the wife of the first index patient was analyzed and found to be positive for this virus. We sampled 111 skin swab specimens from healthy persons or non-MCC patients who had no known contact with MCC patients and screened them with the same nested PCR. Only 1 healthy 30-year-old person had HPyV9, which demonstrated low prevalence in this control group of 1/111 (0.9%, 95% confidence interval 0.05%–5.6%; p = 0.05). Age range of the MCC population was 57–86 years (median 75 years), and because we were unable to detect HPyV9 among the 30 older controls (age range 57–96 years, median 71 years), this finding suggests that the rate of detection in MCC samples was not biased by older ages of these patients. None of the 92 respiratory and 92 stool specimens was positive for HPyV9.

**Discussion**

There is much evidence indicating that healthy human skin harbors numerous viruses. This finding has been extensively reported for cutaneous human papillomavirus (HPV), which is commonly present on the superficial layers of the skin of most persons (11). The recent description of new human viruses belonging to the family **Polyomaviridae** suggests that some of these viruses share the cutaneous tropism of β-HPV and γ-HPV. MCPyV associated with MCC has also detected on the surface of healthy skin of most persons (5,12), and HPyV6 and HPyV7 have been identified on the surface of the skin of healthy persons (13). Detection of an additional human polyomavirus in cutaneous samples reinforces the perception of the skin as a complex microecosystem colonized by many viruses, and polyomaviruses represent part of this viral microbiota.

The existence in humans of a polyomavirus closely related to LPV, whose natural host is the African green monkey, has been anticipated (14). An African green monkey polyomavirus, also known as monkey B-lymphotropic papovavirus, or LPV, was isolated 30 years ago (15) from a lymphoblastoid cell line derived from the African green monkey. The existence of a virus closely related to LPV has been suspected for >30 years because of cross-reacting antibodies in humans who had no known contact with monkeys (14,16). PCR amplification of short sequences matching those of LPV has been reported, but the length and overlap of these sequences were insufficient for characterizing this virus (13,17). The phylogenetic position of HPyV9 we described makes it the sister group of LPV. Moreover, we have shown that the sequence of this virus makes it a good candidate to be the target of antibodies found in humans.

The BC and HI loops of VP1 appear conserved between HPyV9 and LPV; the DE loop shows major changes. Thus, human antibodies cross-reacting with LPV may preferentially recognize the BC and HI loops, and antibodies specific for the DE loop probably cross-react only weakly with the LPV VP1. With the availability of HPyV9-specific reagents, analysis of the specificity of antibodies found in humans will provide useful information.

HPyVs are ubiquitous viruses that infect most of a given population, remain latent, and show frequent reactivations that lead to asymptomatic shedding (1). The VP1 and T proteins of HPyV6, HPyV7, and several polyomaviruses from wild animals show different pathways of evolution (Figure 2). The close relationship between the HPyV6–HPyV7 group and the WUPyV–KIPyV group could result from convergence caused by common host pressure on
VP1 domains (involved in virus–host interactions) or from horizontal gene transfer (recombination).

In immunocompromised persons, reactivation of HPyV often results in specific diseases such as progressive multifocal leukoencephalopathy caused by JCPyV and tubular nephropathy and hemorrhagic cystitis caused by BKPyV. Most newly identified HPyVs are also detected frequently in healthy persons. In addition, most HPyVs (KIPyV, WUPyV, HPyV6, and HPyV7) have not been clearly associated with any human diseases; however, MCPyV has been associated with MCC (18) and TSPyV has been associated with a rare cutaneous disorder (2); both occurred in immunocompromised patients.

HPyV9 was detected in skin of healthy persons who did not have any obvious immunodeficiency and thus might have asymptomatically shed HPyVs. However, among the 3 persons in whom we detected HPyV9, 2/8 had MCC in cutaneous samples, yielding an apparent higher rate of detection (2/8) in this subset of patients, compared with 1/111 in a control and heterogeneous population. This difference in prevalence rate was not biased by age within each group (0/30 had HPyV in the age-matched part of the control group). Although this unexpected result should be interpreted with caution, given the small number of MCC samples and the retrospective nature of our analysis, MCC patients might be prone to reactivation of HPyV9 shedding from their skin or HPyV9 might be involved in the pathophysiology of MCC. We are currently exploring expression of HPyV9 LT antigen in MCC and other cancer tissues. Long-term carriage of HPyV9 for ≥20 months observed in the index patient with MCC is consistent with this hypothesis.

Chronic shedding of HPyV from skin is similar to a well-known feature of cutaneous HPVs that replicate in keratinocytes and are likely to be transmitted environmentally or through person-to-person contact. In our study, detection of HPyV9 in skin of the wife of the index patient suggests a similar route of transmission.

It has been proposed that MCPyV and HPyV6 or HPyV7 may infect superficial cells of the epidermis and that production of virions may be, as for HPVs, linked to keratinocytes and are likely to be transmitted environmentally or through person-to-person contact. In our study, detection of HPyV9 in skin of the wife of the index patient suggests a similar route of transmission.

HPyV9 was detected in cutaneous samples but not in respiratory and fecal samples, and the rate of detection appeared lower than that reported for MCPyV (19) or HPyV6 and HPyV7 (3). The sampling site was chosen because of methods useful for detection of HPyVs on the skin (3,12). Furthermore, we have observed a specific pattern for MCPyV shedding because face swab specimens yielded a higher rate of viral detection than limb specimens (5). However, because shedding of HPyV was not as extensively studied as that of HPVs, we cannot rule out a similar pattern of excretion, which results in underestimating detection of HPyV9 on face swab specimens. The exact prevalence of HPyV9 should be investigated by using serologic and PCR assays, notably to investigate the relevance of published data suggesting that ≈30% of humans have antibodies specific for an LPV-like virus.

Because clinical manifestations associated with HPyV infections dramatically increase in immunocompromised patients, clinical manifestations caused by HPyV9, if they exist, are also more likely to occur in these patients. Furthermore, HPyV9 infection might not remain restricted to the cutaneous area in immunocompromised patients, and reactivation might lead to systemic dissemination and in some cases clinical signs. This hypothesis is supported by identification of HPyV9 in blood and urine of asymptomatic renal transplant recipients (4). Further investigations will be required to decipher the potential role of HPyV9 in homeostatic and pathologic processes.

Acknowledgment

We thank Felix Rey for help with analysis of VP1.

The project “Genotyping of Pathogens and Public Health” was supported in part by the Institut de Veille Sanitaire, Saint-Maurice, France. This study was supported in part by grants from region Île de France and the Programme Hospitalier de Recherche Clinique of the Montpellier University Hospital (AOI 2008, Protocole UF8425).

Dr Sauvage is a postdoctoral research associate at the Institut Pasteur, Paris. Her research focuses on pathogens involved in infectious diseases of unknown etiology.

References


Address for correspondence: Marc Eloit, Department of Virology, Institut Pasteur, 28 Rue du Docteur Roux, F-75015 Paris, France; email: marc.eliot@pasteur.fr

All material published in Emerging Infectious Diseases is in the public domain and may be used and reprinted without special permission; proper citation, however, is required.