Molecular Phylogenetics and Evolution 126 (2018) 210–220

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Phylodynamics of Merkel-cell polyomavirus and human polyomavirus 6: A long-term history with humans

Carolina Torres[a,b,*], Melina Elizabeth Barrios[a,b], Robertina Viviana Cammarata[a,b], Matías Victoria[c], Xavier Fernandez-Cassi[d], Silvia Bofill-Mas[d], Rodney Colina[c], María Dolores Blanco Fernández[a,b], Viviana Andrea Mbayed[a,b]

[a] Universidad de Buenos Aires, Facultad de Farmacia y Bioquímica, Departamento de Microbiología, Inmunología y Biotecnología, Cátedra de Virología, Buenos Aires, Argentina
[b] CONICET, Buenos Aires, Argentina
[c] Laboratorio de Virología Molecular, CENUR Litoral Norte, Sede Salto, Universidad de la República, Uruguay
[d] Laboratory of Virus Contaminants of Water and Food, Department of Genetics, Microbiology and Statistics, University of Barcelona, Barcelona, Catalonia, Spain

ARTICLE INFO

Keywords:
Merkel Cell polyomavirus
Human polyomavirus 6
phylodynamic analysis, South America
Evolution

ABSTRACT

New human polyomaviruses have been described in the last years, including the Merkel-cell polyomavirus (MCPyV; *Human polyomavirus 5*) and the *Human polyomavirus 6* (HPyV6). Although their infection is usually asymptomatic, in immunocompromised host can cause life-threatening pathologies, such as the Merkel cell carcinoma, an aggressive skin neoplasia associated to the MCPyV. Despite being prevalent viruses in population, epidemiological data from South America are scarce, as well as the characterization of the viral types circulating and their origin. The aims of this work were to describe MCPyV and HPyV6 from environmental samples with different geographical origin and to analyze their phylogenetic and evolutionary histories, particularly for MCPyV.

Partial and complete genome sequences were obtained from sewage samples from Argentina, Uruguay and Spain. A total number of 87 sequences were obtained for MCPyV and 33 for HPyV6. Phylogenetic analysis showed that MCPyV sequences distributed according to their geographic origin in Europe/North America, Africa, Asia, South America and Oceania groups, suggesting that viral diversification might have followed human migrations across the globe. In particular, viruses from Argentina associated with Europe/North America and South America genotypes, whereas those from Uruguay and Spain also grouped with Africa genotype, reflecting the origin of the current population in each country, which could arrive not only during ancient human migration but also during recent migratory events. In addition, the South American group presented a high level of clusterization, showing internal clusters that could be related to specific locations, such as French Guiana and Brazil or the Southern region into South America, such as Argentina and Uruguay, suggesting a long term evolutionary process in the region.

Additionally, in this work, we carried out the first analysis about the evolutionary history of MCPyV trough the integration of phylogenetic, epidemiological and historical data. Since a strong association is observed between the phylogenetic relationships and the origin of the sampled population, this analysis was based on the hypothesis of co-divergence between the virus and human populations. This analysis resulted in a substitution rate of 5.1×10^{-8} s/s/y (~5.1% of divergence per million years) for the complete genome of MCPyV, which is in the range of those estimated for other double-stranded DNA viruses.

Regarding HPyV6, a South American group with clusterization was observed (sequences from Uruguay). Meanwhile, sequences from Argentina grouped with European ones (France and Spain) and remained separated from those isolated in China, USA or Australia.

The analysis of viruses from the environment allowed us to deep characterize prevalent infections in different geographic regions, reveling that viruses circulating in each population reflected its origin and that there are specific lineages associated with South America.

* Corresponding author at: Universidad de Buenos Aires, Facultad de Farmacia y Bioquímica, Departamento de Microbiología, Inmunología y Biotecnología. Cátedra de Virología, Junín 954, 4° piso, Ciudad Autónoma de Buenos Aires C1113AAD, Argentina.
E-mail address: ctorres@ffyb.uba.ar (C. Torres).

https://doi.org/10.1016/j.ympev.2018.04.025
Received 26 October 2017; Received in revised form 6 April 2018; Accepted 16 April 2018
Available online 20 April 2018
1055-7903/

1. Introduction

New human polyomaviruses have been discovered in the last years, including the Merkel-cell (MCPyV, *Human polyomavirus 5*) and the *Human polyomavirus 6* (HPyV6) (Feng et al., 2008; Schowalter et al., 2010).

These viruses have been described to be part of the skin microbioma (Foulongne et al., 2012; Schowalter et al., 2010). Viral DNA has been detected in ~80% (MCPyV) and ~28% (HPyV6) of skin swabs of healthy individuals (Foulongne et al., 2010; Wieland et al., 2014), although they were also found in samples from other tissues such as respiratory tract, urine or feces (Dalianis and Hirsch, 2013; Spurgeon and Lambert, 2013). Both viruses have been detected in sewage reinforcing the idea of a high prevalence in the general population (Bofill-Mas et al., 2010; Cantalupo et al., 2011; Di Bonito et al., 2014; Torres et al., 2016).

For both viruses, it was proposed that primary infection occurs during childhood with a significant increase in seroprevalence with age, becoming life-long components of the skin flora (Chen et al., 2011; Van Der Meijden et al., 2013). Particularly, seropositivity ranged between ~ 10–40% in the early childhood to ~ 60–80% or higher values in adulthood (Martel-Jantin et al., 2013; Nicol et al., 2013; Vahabpour et al., 2016; Van Der Meijden et al., 2013).

Although their infection is usually asymptomatic, as for other human polyomaviruses, in immunocompromised host can cause life-threatening pathologies. Particularly, the MCPyV has been associated to the Merkel cell carcinoma (MCC), an aggressive skin cancer arising in the elderly and in chronically immunosuppressed individuals (Dalianis and Hirsch 2013; IARC, 2014).

In regards to HPyV6, it has been proposed to be a contributing agent of epithelial proliferations in patient receiving BRAF inhibitor therapy (Schrama et al., 2014) and it was frequently detected in keratoacanthomas (a non-melanoma skin cancer) and in bile samples of patients with malignant biliary obstruction (Beckervordersandforth et al., 2016; Chan et al., 2017). Besides, it was also found in the cerebrospinal fluid of an HIV-positive individuals with leukoencephalopathy, in cervical samples and in tonsillar tissues (Delbue et al., 2015; Kolia-Diafouka et al., 2016; Salakova et al., 2016).

Molecular characterization allows to deepen the study of molecular signatures associated with pathologies or severity, in particular, tumor-specific mutations were described for MCPyV genome (Shuda et al., 2008; Kassem et al., 2008; Laude et al., 2010; Sastre-Garau et al., 2009; among others). In addition, this characterization is also useful to describe viral lineages associated with specific human populations, as was shown for other polyomaviruses (Kitchen et al., 2008; Krumbholz et al., 2008; Mes et al., 2010; Shackelton et al., 2006).

However, despite being prevalent viruses in population, epidemiological data from South America are scarce, as well as the knowledge about the viral types circulating and their origin.

Thus, the aims of this work were: i. To describe MCPyV and HPyV6 -two of the most prevalent polyomaviruses scarcely studied so far - from environmental samples with different geographical origin; and ii. To analyze the phylogenetic and evolutionary profiles to study their spatio-temporal dispersion patterns, particularly for MCPyV.

2. Material and methods

2.1. Samples

Raw sewage samples from different countries were analyzed. They were collected from Buenos Aires city and suburbs in Argentina (2013, n = 2 and 2016, n = 1); Fray Bentos, Bella Unión, Salto and Paysandú cities in Uruguay (2011 and 2012, n = 48); and Barcelona city in Spain (2011, n = 10).

2.2. Sample processing

Briefly, samples from Argentina and Uruguay were concentrated by ultracentrifugation based on protocols by Pina et al. (1998), as previously described (Torres et al., 2016; Victoria et al., 2014), whereas samples from Spain were concentrated by flocculation based on protocols by Calgua et al. (Calgua et al., 2013), as previously described (Rusiñol et al., 2015). These concentration protocols allow retrieving virions, but free-circulating DNA can also be present in the concentrate.

Nucleic acids from concentrates were extracted with the High Pure Viral Nucleic Acid Kit (Roche, Germany) and the QIAamp® Viral RNA Mini Kit (Qiagen, USA) for samples from Argentina and Spain, or using guanidinium/silica method (Boom et al., 1990) for samples from Uruguay.

2.3. Polyomavirus detection and sequencing

Samples from Argentina were previously characterized as positive for MCPyV and HPV6 by Torres et al. (2016) while samples from Barcelona correspond to those with the highest viral load for MCPyV (≥10.000 genome copies/L) analyzed by Rusiñol et al. (2015). Samples from Uruguay were screened for MCPyV in this work, by a nested PCR directed to the overlapping genetic region that encodes for VP2 and VP1 proteins (244 nt length for the second round amplicon) (Torres et al., 2016). A subset of the MCPyV positive samples from Uruguay and Spain were amplified and sequenced in the VP1 gene (20 from Uruguay and six from Spain), according to previously published protocols (Torres et al., 2016), whereas the two VP1 direct sequences from Argentina were previously obtained (Torres et al. 2016) (Table 1).

One VP1 amplicon from each country was cloned into a pGEM-T Easy Vector System II (Promega, USA) and the clones were sequenced. In addition, the two samples from Argentina were amplified and cloned in the complete genome by a nested PCR described in Table S1. Different clones were sequenced along the complete or partial genomes (Table 1).

In addition, samples from Uruguay and Spain were screened for HPyV6/HPyV7 using a nested PCR directed against the VP2/VP1 partial region (Torres et al., 2016). Identity of viruses was confirmed by sequencing. The complete genome of HPyV6 was amplified from one

Table 1
Clones and sequences obtained for MCPyV positive samples.

Origin	Positive/Total[a] (%)	N° of VP1 sequences	N° of samples cloned in VP1 (n° of clones sequenced)	N° samples cloned in complete genome (n° of clones sequenced)
Argentina	2/24 (8.3%)	2	1 (10 clones)	2 (15 clones: 4 clones in complete genome, 9 in VP1 and LT genes, and 2 only in VP1 gene)
Uruguay	37/48 (77.1%)	20	1 (16 clones)	–
Spain	21/24 (87.5%)	6	1 (10 clones)	–

[a] Total of samples corresponding to the original sampling from which MCPyV positive samples were obtained. Samples from Uruguay have been fully analyzed for MCPyV detection and sequencing in this work. Samples from Spain were screened for detection in Rusiñol et al. (2015) but cloned and sequenced in this work. The two VP1 sequences from Argentina were obtained in Torres et al. (2016) but cloned in this work, along with the amplification, cloning and sequencing of the complete genomes.

sample of each of the three countries using a nested PCR described in Table S1. Products were purified, cloned into a pGEM-T Easy Vector System II (Promega, USA) and sequenced. Different clones were sequenced along the complete or partial genomes.

2.4. Phylogenetic analysis

Phylogenetic analysis was performed on MCPyV VP1 sequences from Argentina, Uruguay and Spain, along with all available sequences longer than 500 bp in GenBank at September 30th, 2017. In addition, phylogenetic trees were also built using the LT genetic region, the concatenate VP1-LT genetic regions or the complete genome sequences. Only one of identical sequences were retained for the analyses, except for sequences reported in this work or from Argentina (see Figure's captions for details).

Sequences were aligned with ClustalX v2.1 (Larkin et al., 2007) and edited with BioEdit v7.0 (Hall, 1999). Phylogenetic trees were built using Bayesian inference with MrBayes v3.2.6 (Ronquist et al., 2012). For each dataset, analyses were performed using an appropriate substitution model according to the Bayesian Information Criterion, estimated with the jModelTest v2.1 software (Darriba et al., 2012). Substitution models selected for datasets were: TPM3uf + G (VP1), TPM2uf + G (VP1-LT), TIM2 + G (LT), TVM + I (complete genome). Analyses were run up to convergence, assessed by effective sample size values higher than 200 using the Tracer v1.6 software (Rambaut et al., 2014), and 10% of the sampling was discarded as burn-in. Majority rule consensus trees were used to represent the most common clades of the posterior distribution of trees obtained. In addition, all the analyses were also carried out by Maximum Likelihood with the IQ-Tree software v1.5 (Nguyen et al., 2015), implemented in the IQ-Tree Server (Trifinopoulos et al., 2016). Confidence was evaluated through the Ultrafast Bootstrap Approximation (1000 replicates) (Hoang et al., 2017).

In addition, for HPyV6, phylogenetic trees were built for both methods using the VP2/VP1 partial region, the concatenate VP1 and LT genetic regions or the complete genome, using sequences reported in this work and all available sequences from GenBank at September 30th, 2017. Substitution models selected for datasets were: JC (VP2/VP1 partial region), HKY + G (VP1-LT) and HKY + G (complete genome).

2.5. MCPyV phylodynamics: Geographical structure and evolutionary history

The degree of association between the phylogenetic clustering and the geographical origin of MCPyV sequences has been tested and quantified estimating association statistics and pairwise F_{ST} values. On one hand, the Parsimony Score (PS), the Association Index (AI) and the Maximum Exclusive Single-state Clade Size (MC) statistics were estimated with the BaTS v0.9 software (Parker et al., 2008). The phylogeny-discrete trait correlation is evaluated considering uncertainty from phylogenetic analysis, by integrating over the posterior distribution of trees obtained by Bayesian methods. For this analysis, VP1 dataset was analyzed labeling sequences according to the country of origin reported for the sequences or the geographic region (Africa, Asia, Europe, North America, Oceania and South America). Statistics were computed for a null distribution with 1000 replicates (expected values under a random phylogeny-trait association) and compared to those obtained for the posterior distribution of trees from the Bayesian analysis (observed values) (Section 2.4). P-values < 0.05 were considered as significant.

On the other hand, pairwise F_{ST} distances were estimated for the VP1 dataset with groups (subpopulations) defined by the country of origin or the geographic region of the sequences, using TrN + G evolutionary model (the most similar model, implemented in the program, to the estimated with jModeltest) with the Arlequin software v.3.5.2.2 (Excoffier and Lischer, 2010). The significance of estimates was determined using 10,000 randomizations of sequences between populations.

The evolutionary history of MCPyV was studied using a Bayesian coalescent analysis on complete genome sequences obtained in this work and from GenBank. Analysis was carried out on a complete dataset (n = 49) and on a reduced dataset (n = 33) that excluded sequences obtained from MCC samples, which could present specific signature mutations that may bias ancestral dates and long-term evolutionary rates estimations. These datasets included only one sequence per individual (when more than one was available) and neither consensus sequences nor those obtained from cell lines were included. The analysis was carried out using an appropriate substitution model according to the Bayesian Information Criterion, estimated with the jModelTest v2.1 software. Substitution models selected for datasets were HKY + I (complete genome, n = 33) and TVM + I (complete genome, n = 49). As MCPyV phylogenetic groups showed association with different populations and geographic regions that suggest co-divergence with humans, temporal calibration was performed based on previously proposed dispersal dates for human populations in the Out-of-Africa model (Forster, 2004; Mellars et al., 2013; Pagani et al., 2016) and for the peopling of Europe (Goebel, 2007; Mellars, 2006) and South America (Rothhammer and Dillehay, 2009). Uniform priors on the time to the most recent common ancestors (tMRCA) were used: 50,000–100,000 years for MCPyV, 40,000–45,000 years for the Europe/North America group and 13,500–15,000 years for the South America group. The uncorrelated lognormal (UCLN) molecular clock model (Drummond et al., 2006) and the GMRF Bayesian Skyride coalescent model (Minin et al., 2008) implemented in the BEAST v1.8.4 software package, in the CIPRES Science Gateway server (Miller et al., 2010) were used. Convergence was assessed by effective sample size (ESS) values higher than 200 and 10% of the sampling was discarded as burn-in. The results obtained were annotated in the maximum clade credibility tree of each run (the tree of the posterior distribution that shows the maximum product of posterior clade probabilities). Uncertainty in parameter estimates was evaluated in the 95% highest posterior density (HPD95%) interval. Two independent runs were combined.

In addition, for the reduced dataset, different evolutionary hypotheses were compared through Bayes Factor using the Marginal Likelihood values estimated by the Generalized Stepping-Stone method (Baele et al., 2016). The three models analyzed included a non-constrained analysis, Africa as a basal group of the tree (fitting the Out-of-Africa model) and the Africa group as sister group of the Europe/North America genotype (fitting the topology obtained in the most clade credibility tree for a non-constrained analysis).

2.6. GenBank accession numbers

Nucleotide sequences reported in this work have been deposited in GenBank under the accession numbers MG241567-MG241665. Sequences shorter than 200 bp were included in Supplementary Material.

3. Results

Sewage samples from Argentina, Uruguay and Spain were analyzed to characterize MCPyV and HPyV6. These viruses have been selected for a deep analysis given their reported high prevalence and scarce characterization, especially in South America.

3.1. MCPyV detection and sequencing

The 77.1% (37/48) of the samples from Uruguay were positive for MCPyV. Direct sequences of VP1 gene were obtained from 20 samples from Uruguay and six samples from Spain, and they were analyzed along with the two previous sequences from Argentina. Besides, one sample from each country was cloned in the VP1 gene and 10 clones

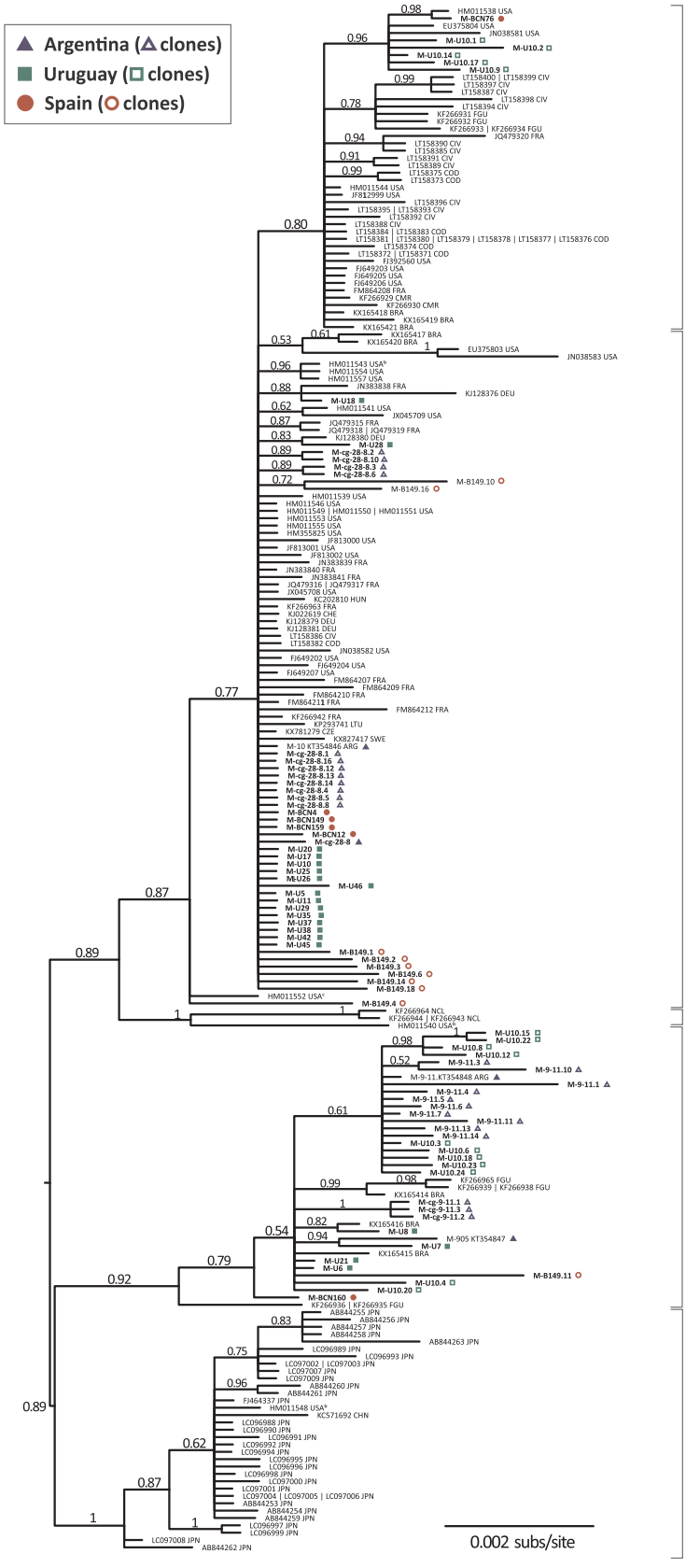

Fig. 1. Majority rule consensus tree from the Bayesian analysis of VP1 sequences of MCPyV (1147 nt) (mid-point rooted). Posterior probabilities values higher than 0.5 are shown at nodes for relevant groups. Genotypes proposed by Martel-Jantin et al. (2014) are indicated in squared brackets. Sequences reported in this work are shown in bold. Country of origin is indicated when available, abbreviated in uppercase. BRA: Brazil, CIV: Ivory Coast, CHE: Switzerland, CHN: China, CMR: Cameroon, COD: Democratic Republic of the Congo, CZE: Czech Republic, DEU: Germany, FGU: French Guiana, FRA: France, HUN: Hungary, JPN: Japan, LTU: Lithuania, NCL: New Caledonia, SWE: Sweden, USA: the United States of America. Genbank accession numbers for identical sequences are indicated among vertical lines, except for those reported in this work: M-cg-9-11.1 | M-cg-9-11.3, M-cg-28-8.1 | M-cg-28-8.16, M-cg-28-8.2 | M-cg-28-8.10, M-U42 | M-U45. [a]Monophyletic in the analysis of complete genome sequences (Fig. S2). [b]Sequence reported in USA from individuals born in Asia (Schowalter et al., 2010). [c]Sequence reported in USA from an individual born in Europe (Schowalter et al., 2010).

(from Argentina and Spain) or 16 clones (from Uruguay) were sequenced. In addition, the complete genome of MCPyV was cloned from two samples from Argentina; 15 of the clones obtained were sequenced along the complete genome or in different regions (Table 1). These sequences were included in phylogenetic and evolutionary analyses (Sections 3.2 and 3.3).

3.2. Phylogenetic analyses of MCPyV

The phylogenetic analyses showed that MCPyV sequences were distributed according to their geographic origin in Europe/North America, Africa, Asia, South America and Oceania groups or genotypes (Fig. 1 and Figs. S1–S4), as was previously suggested (Martel-Jantin et al., 2014). For complete genome sequences, these groups were monophyletic, whereas for VP1, LT or VP1-LT regions, the Europe/North America genotype became paraphyletic owing to the location of the Africa genotype. In particular, the analysis of VP1 region allowed a deep description of internal grouping, especially within the South America genotype, given that few sequences are available for the other genetic regions. Similar topologies were observed for Bayesian and Maximum likelihood methods (Fig. 1 and Figs. S1–S4).

The analysis of VP1 sequences from Argentina included direct sequences obtained for the two positive samples, molecular clones of VP1 region (from one sample) and molecular clones of the complete genomes (from the two samples). The direct sequence of one of the Argentinean samples grouped with samples from Europe/North America. All clones obtained from this sample distributed within the Europe/North America genotype, and seven of them formed a monophyletic cluster in the analyses of VP1-LT and LT regions (Figs. S3 and S4). The direct sequence of the other Argentinean sample grouped with sequences from South America (Fig. 1 and Fig. S1). Its molecular clones clustered with sequences from Uruguay, forming a monophyletic group within the South American genotype.

The analysis of MCPyV from Uruguay comprised the VP1 direct sequences from 20 samples. Most of them (80.0%; 16/20) grouped with samples from the Europe/North America genotype and 20.0% (4/20) with samples from South America. One sample related to the Europe/North America genotype that showed mixed nucleotide positions was cloned in VP1 region and 16 clones were obtained. Analysis showed that 11 variants related with the South American genotype and four of them formed an inner group with high posterior probability and bootstrap value, whereas five sequences associated with sequences from Afro-descendant individuals (Africa genotype) (Fig. 1 and Fig. S1).

Finally, the six positive MCPyV samples from Barcelona were sequenced in VP1. Most of them (4/6) grouped with sequences from Europe/North America, whereas one was related with Afro-descendant individuals (Africa genotype) and one with sequences from South America. One of the samples related with the Europe/North America genotype that presented mixed nucleotide positions was cloned and ten clones were obtained. Most of the variants (9/10) were related with the Europe/North America genotype while one sequence clustered with the South America genotype.

Therefore, the South American group presented a high level of clusterization, showing a first split into sequences from French Guiana and a subgroup formed by sequences from Argentina, Uruguay, Brazil, French Guiana and Spain (Fig. 1 and Fig. S1). This last subgroup also showed internal clusters that could be related to specific locations, such as French Guiana and Brazil or the Southern region into South America, such as Argentina and Uruguay (Fig. 1 and Fig. S1).

This analysis also revealed that sequences from samples collected in South America (Uruguay, Brazil and French Guiana) and classified within the African genotype were not monophyletic, probably corresponding to introductions in South America from different countries.

3.3. Phylodynamics of MCPyV

The geographical structure of MCPyV was evaluated using the AI, PS and MC statistics and pairwise F_{ST} distances, estimated in the VP1 dataset. On overall, a non-random distribution of sequences was observed when sequences were labeled according to the country of origin or the geographic region (Europe, North America, Africa, Asia, Oceania, South America) ($p < 0.001$ for AI and PS statistics for both analyses) (Table 2). In addition, all the regions and most of the countries showed significant geographical association ($p < 0.05$ for the MC statistics) (Table 2). On the other hand, the analysis of pairwise F_{ST} distances among groups defined by geographic regions showed a significant population subdivision for all values ($p < 0.05$), except for the Europe and North America groups (Fig. 2), as expected, given the sequence intermingling observed in phylogenetic trees (Fig. 1 and Fig. S1). Regarding the F_{ST} pairwise distances among sequences grouped by the country of origin, in general, subdivision was observed among countries belonging to the different geographic regions and low or null subdivision was obtained between the countries of the same region (Europe, North America, Africa, Asia, Oceania, South America) (Table S2 and Fig. S5).

Table 2
Association between the phylogenetic clustering and the geographical origin of MCPyV VP1 sequences.

Discrete trait	Statistic	Observed mean (95% confidence interval)	Expected mean (95% confidence interval)	p-value[a]
Geographic region	**AI**	6.9 (5.6–8.3)	15.2 (14.4–16.1)	**< 0.001**
	PS	63.2 (58.0–68.0)	115.3 (111.9–118.6)	**< 0.001**
	MC (Africa)	3.4 (3.0 – 5.0)	1.5 (1.2–2.0)	**0.002**
	MC (Asia)	16.2 (10.0–27.0)	1.9 (1.5–2.4)	**0.001**
	MC (Europe)	3.0 (2.0–5.0)	2.1 1.8–3.0)	**0.046**
	MC (North America)	3.1 (3.0–4.0)	1.9 (1.5–2.2)	**0.014**
	MC (Oceania)	2.0 (2.0–2.0)	1.0 (1.0–1.0)	**0.001**
	MC (South America)	27.0 (20.0–35.0)	3.2 (2.7–3.7)	**0.001**
Country of origin	**AI**	**10.1 (8.6–11.6)**	**18.1 (17.4–18.7)**	**< 0.001**
	PS	**97.0 (92.0–102.0)**	**147.5 (144.0–150.8)**	**< 0.001**
	MC (ARG)	3.5 (3.0–5.0)	1.8 (1.4–2.2)	**0.015**
	MC (BRA)	1.6 (1.0–2.0)	1.1 (1.0–1.2)	**0.005**
	MC (CHE)	1.0 (1.0–1.0)	1.0 (1.0–1.0)	1.000
	MC (CHN)	1.0 (1.0–1.0)	1.0 (1.0–1.0)	1.000
	MC (CIV)	3.2 (3.0–4.0)	1.2 (1.0–1.9)	**0.001**
	MC (CMR)	1.1 (1.0–2.0)	1.0 (1.0–1.0)	1.000
	MC (COD)	2.1 (2.0–3.0)	1.1 (1.0–1.2)	**0.005**
	MC (CZE)	1.0 (1.0–1.0)	1.0 (1.0–1.0)	1.000
	MC (DEU)	1.0 (1.0–1.0)	1.0 (1.0–1.0)	1.000
	MC (ESP)	2.0 (1.0–3.0)	1.3 (1.1–2.0)	**0.032**
	MC (FGU)	2.0 (2.0–2.0)	1.0 (1.0–1.1)	**0.002**
	MC (FRA)	2.0 (1.0–3.0)	1.3 (1.1–2.0)	**0.035**
	MC (HUN)	1.0 (1.0–1.0)	1.0 (1.0–1.0)	1.000
	MC (JPN)	14.6 (10.0–25.0)	1.8 (1.5–2.2)	**0.001**
	MC (LTU)	1.0 (1.0–1.0)	1.0 (1.0–1.0)	1.000
	MC (NCL)	2.0 (2.0–2.0)	1.0 (1.0–1.0)	**0.001**
	MC (SWE)	1.0 (1.0–1.0)	1.0 (1.0–1.0)	1.000
	MC (URY)	4.2 (4.0–5.0)	2.0 (1.6–2.3)	**0.001**
	MC (USA)	3.1 (3.0–4.0)	1.9 (1.5–2.2)	**0.019**

[a] The p-value corresponds to the proportion of trees from the null distribution (expected distribution in a random phylogeny-trait association) equal to, or more extreme than, the posterior estimate of the statistic from the posterior distribution of trees (observed distribution). Significant values are shown in bold.

The timescale of MCPyV evolutionary history was studied through Bayesian coalescent analyses. Briefly, previously estimated dispersal dates for human population were used to calibrate the tMRCA of MCPyV, Europe/North America and South America groups, for the dataset containing complete genome sequences under different hypothesis.

This analysis showed that the tMRCA for MCPyV dated ~72,443 years ago (in a non-constrained analysis), and that the genotypes diversified in the range of ~14,297–42,148 years ago, although high uncertainty in estimations were obtained (Table 3 and Fig. 3a).

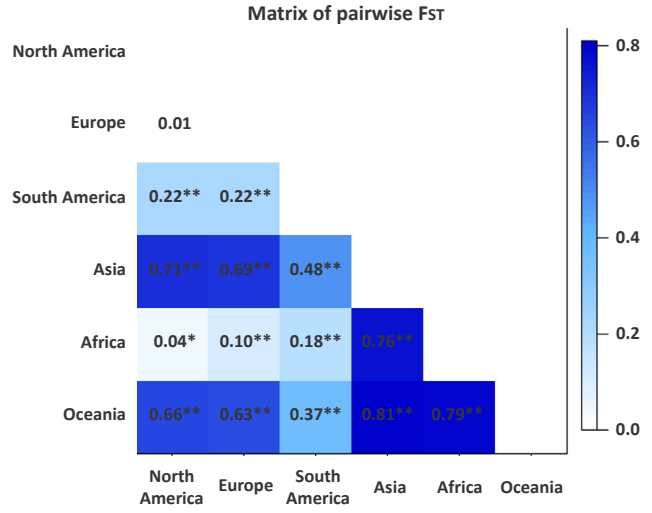

Fig. 2. Graphical visualization and pairwise F_{ST} values for MCPyV VP1 sequences grouped by geographic regions. $^{*}p < 0.05$, $^{**}p < 0.01$.

Table 3
Time to the most recent common ancestors for MCPyV complete genome sequences.

	Non-constrained analysis		Constrained analysis (Africa as basal group)	
Group	Median (years)	HPD95% (years)	Median (years)	HPD95% (years)
Europe/NA	42,148	40,041–44,724	42,482	40,045–44,785
Africa	33,191	11,564–53,932	39,278	10,688–66,150
Oceania	34,768	13,063–67,043	23,908	8231–40,108
South America	14,297	13,579–14,995	14,278	13,572–14,993
Asia	22,700	6019–44,857	17,918	5561–32,801
MCPyV ancestor	72,443	52,852–97,697	58,119	50,004–80,723

Furthermore, a global evolutionary rate for MCPyV genome was estimated as 5.1×10^{-8} s/s/y (HPD95% = $3.7 \times 10^{-8} - 6.6 \times 10^{-8}$). Notably, estimates for the dataset that included sequences from MCC samples resulted in more recent ancestral dates (~55,652 years ago for MCPyV) and in a slightly faster evolutionary rate (5.8×10^{-8} s/s/y) (Table S3 and Fig. S7).

Under the hypothesis of co-divergence with humans, the Africa group is expected to be a basal group in the MCPyV phylogenetic tree, however, as was observed in the midpoint-rooted phylogeny, the most clade credibility tree obtained for the non-constrained analysis showed, instead, that this group shared a common ancestor with the Europe/North America genotype (Fig. 3a). Then, the hypothesis of co-divergence with humans with Africa as a basal group in the phylogeny (Fig. 3b) was compared with a non-constrained topology hypothesis using Bayes Factor (BF). This comparison did not significantly favor any of these models (lnBF = 1.1, Table S4). However, when a deliberate hypothesis of Africa and Europe/North America groups sharing a common ancestor was compared to the hypothesis with Africa as a basal group, the former was favored by BF (lnBF = 11.5, Table S4).

Under the hypothesis of Africa as the basal group, an evolutionary rate of 5.5×10^{-8} s/s/y (HPD95% = $4.1 \times 10^{-8} - 7.1 \times 10^{-8}$ s/s/y) is obtained, whereas the time for the ancestor of MCPyV was estimated as 58,119 years (HPD95%= 50,004–80,723) (Table 3 and Fig. 3b). In Fig. S6 are shown the maximum clade credibility trees in substitutions per site scale for the coalescent analyses under a non-constrained hypothesis and forcing Africa as a basal group.

3.4. HPyV6 detection and phylogenetic analysis

HPyV6 detection resulted positive in 41.7% (20/48) of samples from Uruguay and in 20.0% (2/10) of samples from Barcelona. Remarkably, in one HPyV6 positive sample from Uruguay also HPyV7 was found by sequencing (Fig. 4b and Fig. S8b). A further characterization of HPyV6 viruses was carried out through amplification of complete genome sequences, followed by molecular cloning and sequencing. Analysis of VP1-LT concatenate regions and complete genome sequences showed that viruses from Argentina and Spain grouped with the only sequence from Europe available so far (from France) and that sequences from Uruguay grouped together with high posterior probability and bootstrap support (Fig. 4a and Figs. S8a–S9).

4. Discussion

In this work, MCPyV and HPyV6 viruses from samples with different geographical origin were analyzed through phylogenetic and phylodynamic analysis, particularly for MCPyV, a prevalent virus associated with a neoplastic disease but scarcely studied so far in South America. Then, samples from two South American countries, Argentina and Uruguay, together with a European country, Spain, were analyzed in the context of the known worldwide genetic information of these viruses.

For this analysis, environmental samples were used, which is a strategy that has previously shown to be useful to analyze epidemiological patterns of viral infections in the general population. In this study, a high level of detection was observed in sewage samples from Uruguay, reaching 77.1% and 41.7% for MCPyV and HPyV6, respectively, suggesting a significant level of circulation in general population. The distribution of MCPyV reported for the samples from Spain reached similar levels (75.0%) (Rusiñol et al. 2015). However, in Argentina, the detection of MCPyV and HPyV6 in raw sewage was 8.3% for both viruses (Torres et al. 2016). These differences could be attributed to different protocols or processing methods used. Other studies have also shown the viral detection in sewage from Spain and Italy (Bofill-Mas et al., 2010; Cantalupo et al., 2011; Di Bonito et al., 2014), supporting a possible excretion by the fecal or urinary routes or because of skin peeling (Bofill-Mas et al., 2010; Torres et al., 2016). It is worth noting that two negative samples from Uruguay tested positive after an enrichment step using a Rolling Circle Amplification (RCA) technique (data not shown), suggesting an even increased viral presence in sewage; in addition, the RCA might be a useful technique to enhance the amplification of circular double-stranded DNA viruses.

Based on the molecular characterization of MCPyV, a distribution of phylogenetic groups (genotypes) associated to their geographic origin was observed. The clustering pattern defined the Europe/North America, Africa, Asia, South America and Oceania genotypes, as previously proposed (Martel-Jantin et al., 2014). This distribution allowed us to propose that viral diversification might have follow human migrations across the globe, as was suggested for other DNA viruses associated with persistent infections and high host specificity (Kitchen et al., 2008; Krumbholz et al., 2008; Torres et al., 2011; Yogo et al., 2004; Zheng et al., 2007; among others).

The analysis of the phylogeographic distribution observed is consistent with the proposal of viral diversification following migrations. On one hand, Argentina population (City of Buenos Aires) is mainly composed by Native American and European descendants. Europeans arrived and settled in the current Argentine territory in several waves since the conquest of America with a maximum of immigration between 1880 and 1930, which was reactivated to a lesser extent in the mid-XXth century (Devoto and Benencia, 2003). Whereas, Uruguay population includes, in addition, individuals with African ancestry – descendants of people enslaved and transferred from Sub-Saharan Africa to the Americas in the XVIth-XIXth centuries - that reach 8.1% of the current population of the country. The Afro-descendant population is composed

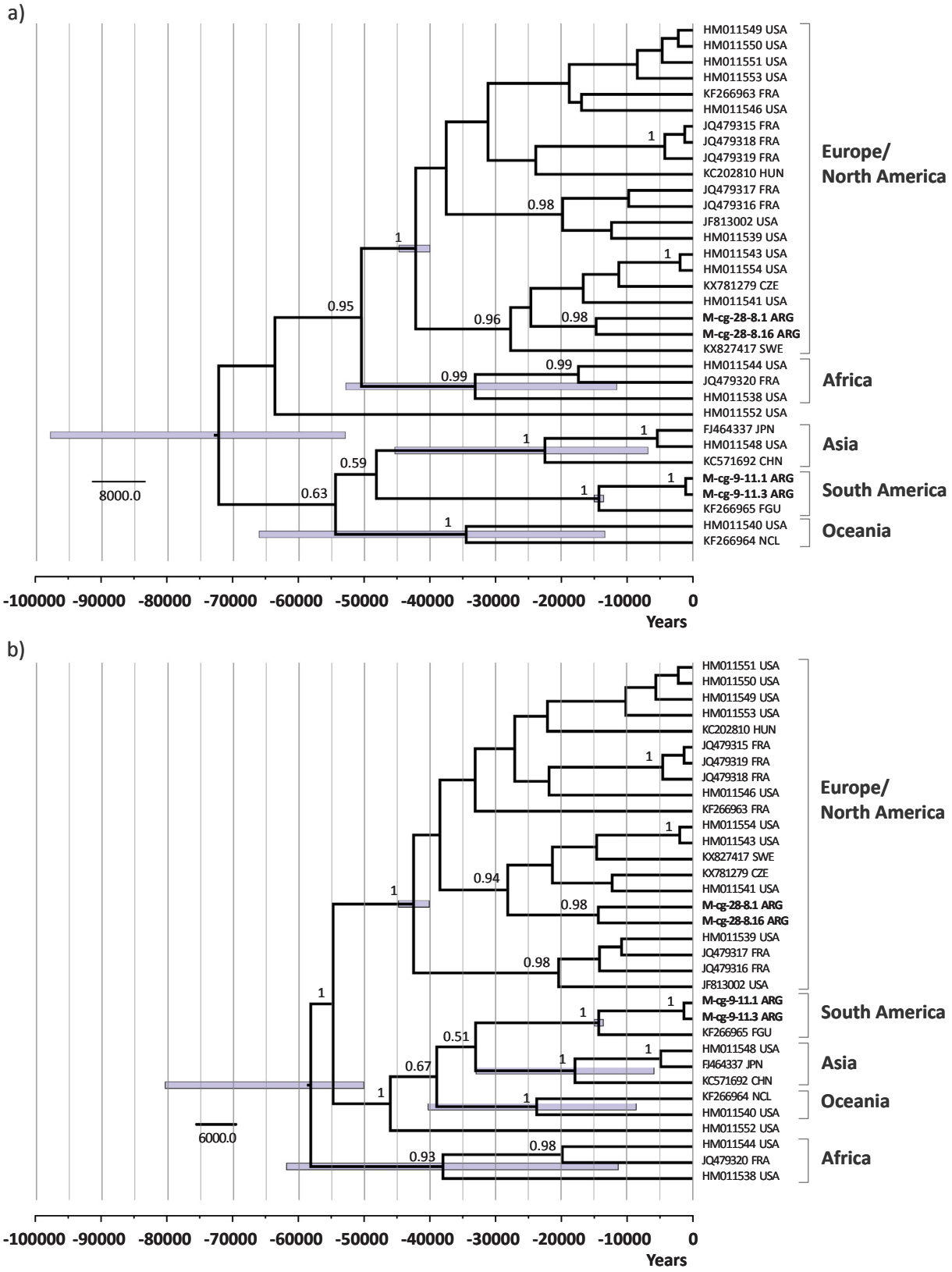

Fig. 3. Maximum clade credibility trees of (a) the non-constrained coalescent analysis, and (b) the constrained analysis with Africa as basal group, carried out on MCPyV complete genome sequences. Trees are in temporal scale, negative values indicate years before the present. Posterior probability values are shown at nodes for relevant groups. Colored bars represent the HPD95% interval for the tMRCA. Sequences reported in this work are shown in bold. ARG: Argentina, CHN: China, CZE: Czech Republic, FGU: French Guiana, FRA: France, HUN: Hungary, JPN: Japan, NCL: New Caledonia, SWE: Sweden, USA: the United States of America. [a]Sequence reported in USA from individuals born in Asia (Schowalter et al., 2010). [b]Sequence reported in USA from an individual born in Europe (Schowalter et al., 2010). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

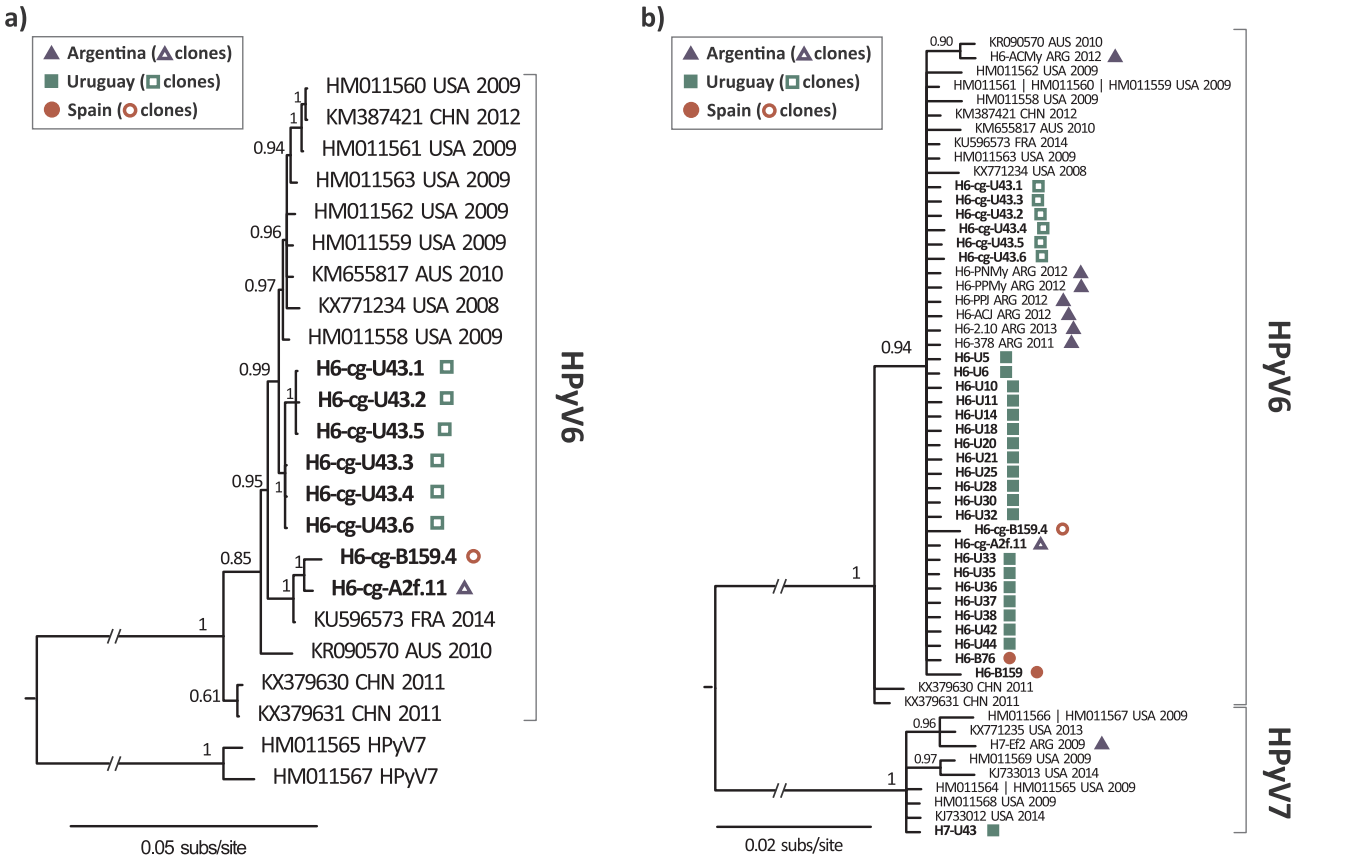

Fig. 4. Majority rule consensus trees from the Bayesian analyses of (a) the concatenate VP1-LT region sequences of HPyV6 (2082 nt) (mid-point rooted), and (b) the VP2/VP1 fragment (165 nt) of HPyV6 and HPyV7 (mid-point rooted). Posterior probabilities values higher than 0.5 are shown at nodes. Sequences reported in this work are shown in bold. Year of isolation and country of origin is indicated when available, abbreviated in uppercase. AUS: Australia, CHN: China, FRA: France, USA: the United States of America. Genbank accession numbers for identical sequences are indicated among vertical lines, except for those from Argentina for the VP2/VP1 fragment: H6-cg-U43.1 | H6-cg-U43.3, H6-PNMy ARG 2012 | H6-PPMy ARG 2012 | H6-PPJ ARG 2012, H6-U6 | H6-U10 | H6-U11 | H6-U14 | H6-U18 | H6-U20 | H6-U21 | H6-U25 | H6-U28 | H6-U30, H6-U33 | H6-U35, H6-U37 | H6-U38 | H6-U42 | H6-U44 | H6-B76.

by individuals that directly arrived to Uruguay from Africa but also by more recent migrants from Brazil, Central America and the Caribbean (Cabella et al., 2011). Therefore, in South America, the circulation of at least three MCPyV genotypes was demonstrated (Europe-North America, South America and Africa genotypes), reinforcing previous descriptions made for clinical samples from Brazil and French Guiana (Baez et al., 2016; Martel-Jantin et al., 2014). Notably, different variants and clusters within the genotypes were observed. In particular, supported clusters found within the South American group might be related to diversification events occurred in particular regions of South America associated with the early settlement of the continent.

On the other hand, as expected, most of the sequences from Barcelona were related to the Europe/North America genotype, although viruses associated to South America and Africa genotypes were also found. In this case, epidemiological records indicate that these variants might be the result of recent migration events from South America and Africa to Spain. These records indicate that in the period from 1991 to 2014, the proportion of individuals from South America and Africa in the population of the Metropolitan area of Barcelona increased from 0.3 to 4.0% and from 0.2 to 2.0%, respectively (AMB, 2014).

Additionally, in this work, we carried out the first analysis about the evolutionary history of MCPyV trough the integration of phylogenetic, epidemiological and historical data. This analysis was based on the hypothesis of co-divergence between the virus and human populations since a strong association is observed between the phylogenetic relationships and the origin of the samples. Calibration of the analysis was based on estimated dates of human migrations across the globe, according to which anatomically modern humans spread out of Africa into Eurasia 50,000–100,000 years ago (Forster, 2004; Mellars et al., 2013; Pagani et al., 2016), and the dispersal of humans from Eurasia to the Mediterranean and temperate Europe could have occurred about 40,000–45,000 years ago (Goebel, 2007; Mellars, 2006). Furthermore, many studies about the peopling of America suggest that a small population entered the Americas from Beringia around 16,000 years ago (Llamas et al., 2016), and after a rapid migration to the south, entered South America around 13,500–15,000 years ago (Rothhammer and Dillehay, 2009).

As expected, the temporal estimation for Europe/North America and South America genotypes fitted the calibration used in the analysis. According to our hypothesis testing results, no significant differences between the non-constrained and the constrained (Africa as basal group) models was found, indicating that both scenarios may be plausible for the evolution of this virus, at least with the data available so far. The Africa-as-basal model would be the best fitting scenario for the hypothesis of viral dispersion following human migration, however, in the other hypothesis (non-constrained analysis), an inconsistency with the Out-of-Africa model (Nielsen et al., 2007) is observed for the location of the Africa group in the coalescent tree, and for both hypotheses, in the estimation date (median) for its ancestor that resulted more recent than the European one. If, as proposed, MCPyV co-diverged with the human populations during migrations across the globe, our results might indicate that current sampled lineages of the Africa group probably represent only a small part of the actual diversity of viruses from Africa, resulting in an underestimation of the time to the most recent common ancestor. It is worth noting that no information is

available about the ethnical origin of individuals from which the MCPyV complete genome of the Africa group were sequenced, thus, their inclusion was based on the observed intermingling with other sequences with a known ethnical origin in the phylogenetic tree of VP1 region. Wider information of complete genome sequences of MCPyV, especially from sub-Saharan Africa will clarify the first events in the evolution of this virus.

In addition, questions arise about the position into the tree of a sequence isolated in USA from an individual born in Europe without MCC (HM011552) (Schowalter et al., 2010). Although the lack of information about the country of origin or the ethnical background of the individual do not allow solid speculations, one possibility is that it could be the unique sequence available so far from a group of viruses possibly not described yet, which could not be related to viruses that migrated to North America. Also, other sequences showed an unexpected pattern such as those isolated from individuals born in Asia, particularly in China (C. Buck, personal communication). One of them grouped with sequences from New Caledonia (Oceania genotype) and the other, with sequences from USA (Europe/North America genotype). These cases emphasize the importance of having a more detailed record of samples to be used in phylogeographic studies, since the geographic origin assignment could not be enterally useful in some cases. Even though other processes than co-divergence with human could be involved in the evolution of MCPyV, we think that it might explain a great part of the current geographical distribution of this virus. Further speculations and testing could be done when more sequences and information are available.

In addition, a median long-term substitution rate of 5.1×10^{-8} s/s/y (~5.1% of divergence per million years) was estimated for the non-constrained analysis, whereas a rate of 5.5×10^{-8} s/s/y (HPD95% = 4.1×10^{-8} s/s/y – 7.1×10^{-8} s/s/y) was estimated for the analysis that forced Africa as a basal group. These rates are in the range of those estimated for other double-stranded DNA viruses (Aiewsakun and Katzourakis, 2016). However, they are faster than the rates estimated for the PyV Monominor clade -that includes MCPyV- (5×10^{-9} s/s/y for the 1st + 2nd codon position and 2×10^{-8} s/s/y for the 3rd codon position for VP1 and LT regions) (Buck et al., 2016). This discrepancy could be explained by the fact that evolutionary rate estimates are negatively correlated with the measurement timescale (time-dependent rate phenomenon) (Aiewsakun and Katzourakis, 2016). Therefore, the phylodynamic analysis calibration using the time of diversification of different host species will estimate slower substitutions rates that doing so with diversification events within a single host species. We consider that this last approach is more appropriate for an accurate evaluation of the evolutionary behavior of MCPyV lineages in the human host.

Regarding HPyV6, even though it would be a prevalent virus, still few sequences are available. Similar to MCPyV, a South American group with clusterization was observed (sequences from Uruguay). European sequences separated from those isolated in China, USA or Australia. The only sequence from Argentina grouped with the European ones, which could represent only part of the viral diversity in the country. As many sequences became available from distinct locations, more analyses will allow understanding the evolutionary behavior and the geographic structure of this virus.

Finally, the analysis of viruses from the environment allowed us to deep characterize prevalent infections in different geographic regions, reveling that viruses circulating in each population reflected its origin and that there are specific lineages associated with South America. Further studies are needed to analyze the clinical and epidemiological impact of these genetic variants.

Funding

This work was supported by grants from Universidad de Buenos Aires (SECyT-UBA 20020120100345BA), Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT; PICT 2012-2679 and PICT 2014-2435) and Fundación “Alberto J. Roemmers”, Argentina; CSIC I + D 2010-Universidad de la República, Uruguay. During the development of this study, Xavier Fernandez-Cassi was a fellow of the Catalan Government “AGAUR” (FI-DGR).

Funding sources had not been involved in the study design, the collection, analysis and interpretation of data, in the writing of the manuscript or in the decision to submit the article for publication.

Conflicts of interest

None.

Acknowledgements

Authors thank Dr. Brian Foley and Dr. Ignacio Bravo for critical review and constructive comments regarding the manuscript, greatly improving this work. Also, authors thank Dr. Christopher Buck for data and helpful discussion.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.ympev.2018.04.025.

References

Aiewsakun, P., Katzourakis, A., 2016. Time-dependent rate phenomenon in viruses. J. Virol. 90, 7184–7195. http://dx.doi.org/10.1128/JVI.00593-16.

Area Metropolitana de Barcelona, 2014. AMB. Population structure [WWW Document]. URL http://www.amb.cat/en/web/area-metropolitana/dades-estadistiques/demografia/estructura-de-la-poblacio

Baele, G., Lemey, P., Suchard, M.A., 2016. Genealogical working distributions for bayesian model testing with phylogenetic uncertainty. Syst. Biol. 65, 250–264. http://dx.doi.org/10.1093/sysbio/syv083.

Baez, C.F., Diaz, N.C., Venceslau, M.T., Luz, F.B., Guimarães, M.A.A.M., Zalis, M.G., Varella, R.B., 2016. Phylogenetic and structural analysis of merkel cell polyomavirus VP1 in Brazilian samples. Virus Res. 221, 1–7. http://dx.doi.org/10.1016/j.virusres.2016.05.004.

Beckervordersandforth, J., Pujari, S., Rennspiess, D., Speel, E.J.M., Winnepenninckx, V., Diaz, C., Weyers, W., Haugg, A.M., Kurz, A.K., zur Hausen, A., 2016. Frequent detection of human polyomavirus 6 in keratoacanthomas. Diagn. Pathol. 11, 58. http://dx.doi.org/10.1186/s13000-016-0509-z.

Bofill-Mas, S., Rodriguez-Manzano, J., Calgua, B., Carratala, A., Girones, R., 2010. Newly described human polyomaviruses Merkel cell, KI and WU are present in urban sewage and may represent potential environmental contaminants. Virol. J. 7, 141. http://dx.doi.org/10.1186/1743-422X-7-141.

Boom, R., Sol, C., Salimans, M., Jansen, C., Wertheim-van Dillen, P., van der Noordaa, J., 1990. Rapid and simple method for purification of nucleic acids. J. Clin. Microbiol. 28, 495–503. http://dx.doi.org/10.2116/analsci.25.941.

Buck, C.B., Van Doorslaer, K., Peretti, A., Geoghegan, E.M., Tisza, M.J., An, P., Katz, J.P., Pipas, J.M., McBride, A.A., Camus, A.C., McDermott, A.J., Dill, J.A., Delwart, E., Ng, T.F.F., Farkas, K., Austin, C., Kraberger, S., Davison, W., Pastrana, D.V., Varsani, A., 2016. The ancient evolutionary history of polyomaviruses. PLoS Pathog. 12, 1–26. http://dx.doi.org/10.1371/journal.ppat.1005574.

Cabella, W., Nathan, M., Tenenbaum, M., 2011. Atlas sociodemográfico y de la desigualdad del Uruguay.

Calgua, B., Rodriguez-Manzano, J., Hundesa, A., Suñen, E., Calvo, M., Bofill-Mas, S., Girones, R., 2013. New methods for the concentration of viruses from urban sewage using quantitative PCR. J. Virol. Methods 187, 215–221. http://dx.doi.org/10.1016/j.jviromet.2012.10.012.

Cantalupo, P.G., Calgua, B., Zhao, G., Hundesa, A., Wier, A.D., Katz, J.P., Grabe, M., Hendrix, R.W., Girones, R., Wang, D., Pipas, J.M., 2011. Raw sewage harbors diverse viral populations. MBio 2, 1–11. http://dx.doi.org/10.1128/mBio. 00180-11.

Chan, J.F.W., Tee, K.-M., Choi, G.K.Y., Zhu, Z., Poon, R.W.S., Ng, K.T.P., Chan, K.-H., Hung, I.F.N., Man, K., Yuen, K.-Y., 2017. First detection and complete genome sequence of a phylogenetically distinct human polyomavirus 6 highly prevalent in human bile samples. J. Infect. 74, 50–59. http://dx.doi.org/10.1016/j.jinf.2016.11.002.

Chen, T., Hedman, L., Mattila, P.S., Jartti, T., Ruuskanen, O., Söderlund-Venermo, M., Hedman, K., 2011. Serological evidence of Merkel cell polyomavirus primary infections in childhood. J. Clin. Virol. 50, 125–129. http://dx.doi.org/10.1016/j.jcv.2010.10.015.

Dalianis, T., Hirsch, H.H., 2013. Human polyomaviruses in disease and cancer. Virology 437, 63–72.

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9, 772–772. https://doi.org/10.

1038/nmeth.2109.
Delbue, S., Elia, F., Signorini, L., Bella, R., Villani, S., Marchioni, E., Ferrante, P., Phan, T.G., Delwart, E., 2015. Human polyomavirus 6 DNA in the cerebrospinal fluid of an HIV-positive patient with leukoencephalopathy. J. Clin. Virol. 68, 24–27. http://dx.doi.org/10.1016/j.jcv.2015.04.016.
Devoto, F., Benencia, R., 2003. History of immigration in Argentina. Chapter 6. Mass immigration (in Spanish).
Di Bonito, P., Libera, S. Della, Petricca, S., Iaconelli, M., Accardi, L., Muscillo, M., La Rosa, G., 2014. Frequent and abundant merkel cell polyomavirus detection in urban wastewaters in Italy. Food Environ. Virol. 7, 1–6. http://dx.doi.org/10.1007/s12560-014-9168-y.
Drummond, A.J., Ho, S.Y.W., Phillips, M.J., Rambaut, A., 2006. Relaxed phylogenetics and dating with confidence. PLOS Biol. 4, e88.
Excoffier, L., Lischer, H.E.L., 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Resour. 10, 564–567. http://dx.doi.org/10.1111/j.1755-0998.2010.02847.x.
Feng, H., Shuda, M., Chang, Y., Moore, P.S., 2008. Clonal integration of a polyomavirus in human Merkel cell carcinoma. Science (80-.) 319, 1096–1100.
Forster, P., 2004. Ice Ages and the mitochondrial DNA chronology of human dispersals: a review. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 359, pp. 255–264 (discussion 264). https://doi.org/10.1098/rstb.2003.1394
Foulongne, V., Kluger, N., Dereure, O., Mercier, G., Molès, J., Guillot, B., Segondy, M., 2010. Polyomavirus Cutaneous Swabs 16, 685–687. http://dx.doi.org/10.3201/eid1604.091278.
Foulongne, V., Sauvage, V., Hebert, C., Dereure, O., Cheval, J., Manuguerra, J., Gouilh, A., Pariente, K., Segondy, M., Burguie, A., Eloit, M., 2012. Human skin microbiota: high diversity of dna viruses identified on the human skin by high throughput sequencing 7. https://doi.org/10.1371/journal.pone.0038499
Goebel, T., 2007. ANTHROPOLOGY: The Missing Years for Modern Humans. Science (80-.). 315, 194–196. https://doi.org/10.1126/science.1137564
Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.
Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., Le, S.V., 2017. UFBoot2: improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35, 518–522. http://dx.doi.org/10.1093/molbev/msx281.
IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, 2014. Malaria and some polyomaviruses (Sv40, Bk, Jc, and Merkel Cell Viruses). IARC Monogr. Eval. Carcinog. Risks Hum. 104, 9–350.
Kassem, A., Schopflin, A., Diaz, C., Weyers, W., Stickeler, E., Werner, M., Zur Hausen, A., Schöpflin, A., Diaz, C., Weyers, W., Stickeler, E., Werner, M., Zur Hausen, A., 2008. Frequent detection of Merkel cell polyomavirus in human Merkel cell carcinomas and identification of a unique deletion in the VP1 gene. Cancer Res. 68, 5009–5013. http://dx.doi.org/10.1158/0008-5472.CAN-08-0949.
Kitchen, A., Miyamoto, M.M., Mulligan, C.J., 2008. Utility of DNA viruses for studying human host history: case study of JC virus. Mol. Phylogenet. Evol. 46, 673–682. http://dx.doi.org/10.1016/j.ympev.2007.09.005.
Kolia-Diafouka, P., Foulongne, V., Boulle, N., Ngou, J., Kelly, H., Sawadogo, B., Delany-Moretlwe, S., Mayaud, P., Segondy, M., Group, H.S., 2016. Detection of four human polyomaviruses (MCPyV, HPyV6, HPyV7 and TSPyV) in cervical specimens from HIV-infected and HIV-uninfected women. Sex. Transm. Infect. 92, 492–494.
Krumbholz, A., Bininda-Emonds, O.R.P., Wutzler, P., Zell, R., 2008. Evolution of four BK virus subtypes. Infect. Genet. Evol. 8, 632–643. http://dx.doi.org/10.1016/j.meegid.2008.05.006.
Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23, 2947–2948.
Laude, H.C., Jonchere, B., Maubec, E., Carlotti, A., Marinho, E., Couturaud, B., Peter, M., Sastre-Garau, X., Avril, M.F., Dupin, N., Rozenberg, F., 2010. Distinct merkel cell polyomavirus molecular features in tumour and non tumour specimens from patients with merkel cell carcinoma. PLoS Pathog. 6, e1001076.
Llamas, B., Fehren-Schmitz, L., Valverde, G., Soubrier, J., Mallick, S., Rohland, N., Nordenfelt, S., Valdiosera, C., Richards, S.M., Rohrlach, A., Romero, M.I.B., Espinoza, I.F., Cagigao, E.T., Jimenez, L.W., Makowski, K., Reyna, I.S.L., Lory, J.M., Torrez, J.A. B., Rivera, M.A., Burger, R.L., Ceruti, M.C., Reinhard, J., Wells, R.S., Politis, G., Santoro, C.M., Standen, V.G., Smith, C., Reich, D., Ho, S.Y.W., Cooper, A., Haak, W., 2016. Ancient mitochondrial DNA provides high-resolution time scale of the peopling of the Americas. Sci. Adv. 2, e1501385–e1501385. https://doi.org/10.1126/sciadv.1501385.
Martel-Jantin, C., Filippone, C., Tortevoye, P., Afonso, P.V., Betsem, E., Descorps-Declere, S., Nicol, J.T.J., Touzé, A., Coursaget, P., Crouzat, M., Berthet, N., Cassar, O., Gessain, A., Touze, A., Coursaget, P., Crouzat, M., Berthet, N., Cassar, O., Gessain, A., 2014. Molecular epidemiology of merkel cell polyomavirus: evidence for geographically related variant genotypes. J. Clin. Microbiol. 52, 1687–1690. http://dx.doi.org/10.1128/JCM.02348-13.
Martel-Jantin, C., Pedergnana, V., Nicol, J.T.J., Leblond, V., Trégouët, D.A., Tortevoye, P., Plancoulaine, S., Coursaget, P., Touzé, A., Abel, L., Gessain, A., 2013. Merkel cell polyomavirus infection occurs during early childhood and is transmitted between siblings. J. Clin. Virol. 58, 288–291. http://dx.doi.org/10.1016/j.jcv.2013.06.004.
Mellars, P., 2006. Going east: new genetic and archaeological perspectives on the modern human colonization of Eurasia. Nat. Hist. 796–800.
Mellars, P., Gori, K.C., Carr, M., Soares, P.a., Richards, M.B., 2013. Genetic and archaeological perspectives on the initial modern human colonization of southern Asia. Proc. Natl. Acad. Sci. USA 110, 10699–10704. http://dx.doi.org/10.1073/pnas.1306043110.
Mes, T.H.M., van Doornum, G.J.J., Schutten, M., 2010. Population genetic tests suggest that the epidemiologies of JCV and BKV are strikingly different. Infect. Genet. Evol. 10, 397–403. http://dx.doi.org/10.1016/j.meegid.2009.04.007.
Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. 2010 Gatew. Comput. Environ. Work. GCE 2010. https://doi.org/10.1109/GCE.2010.5676129.
Minin, V.N., Bloomquist, E.W., Suchard, M.A., 2008. Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. Mol. Biol. Evol. 25, 1459–1471. http://dx.doi.org/10.1093/molbev/msn090.
Nguyen, L.T., Schmidt, H.A., Von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274. http://dx.doi.org/10.1093/molbev/msu300.
Nicol, J.T.J., Robinot, R., Carpentier, A., Carandina, G., Mazzoni, E., Tognon, M., Touzé, A., Coursageta, P., 2013. Age-specific seroprevalences of merkel cell polyomavirus, human polyomaviruses 6, 7, and 9, and trichodysplasia spinulosa-associated polyomavirus. Clin. Vaccine Immunol. 20, 363–368. http://dx.doi.org/10.1128/CVI.00438-12.
Nielsen, R., Akey, J.M., Jakobsson, M., Pritchard, J.K., Tishkoff, S., Willerslev, E., 2017. Tracing the peopling of the world through genomics. Nature 541, 302–310. http://dx.doi.org/10.1038/nature21347.
Pagani, L., Lawson, D.J., Jagoda, E., Mörseburg, A., Eriksson, A., Mitt, M., Clemente, F., Hudjashov, G., DeGiorgio, M., Saag, L., Wall, J.D., Cardona, A., Mägi, R., Sayres, M.A. W., Kaewert, S., Inchley, C., Scheib, C.L., Järve, M., Karmin, M., Jacobs, G.S., Antao, T., Iliescu, F.M., Kushniarevich, A., Ayub, Q., Tyler-Smith, C., Xue, Y., Yunusbayev, B., Tambets, K., Mallick, C.B., Saag, L., Pocheshkhova, E., Andriadze, G., Muller, C., Westaway, M.C., Lambert, D.M., Zoraqi, G., Turdikulova, S., Dalimova, D., Sabitov, Z., Sultana, G.N.N., Lachance, J., Tishkoff, S., Momynaliev, K., Isakova, J., Damba, L. D., Gubina, M., Nymadawa, P., Evseeva, I., Atramentova, L., Utevska, O., Ricaut, F.-X., Brucato, N., Sudoyo, H., Letellier, T., Cox, M.P., Barashkov, N.A., Škaro, V., Mulahasanović, L., Primorac, D., Sahakyan, H., Mormina, M., Eichstaedt, C.A., Lichman, D. V., Abdullah, S., Chaubey, G., Wee, J.T.S., Mihailov, E., Karunas, A., Litvinov, S., Khusainova, R., Ekomasova, N., Akhmetova, V., Khidiyatova, I., Marjanović, D., Yepiskoposyan, L., Behar, D.M., Balanovska, E., Metspalu, A., Derenko, M., Malyarchuk, B., Voevoda, M., Fedorova, S.A., Osipova, L.P., Lahr, M.M., Gerbault, P., Leavesley, M., Migliano, A.B., Petraglia, M., Balanovsky, O., Khusnutdinova, E.K., Metspalu, E., Thomas, M.G., Manica, A., Nielsen, R., Villems, R., Willerslev, E., Kivisild, T., Metspalu, M., 2016. Genomic analyses inform on migration events during the peopling of Eurasia. Nature 538, 238–242. https://doi.org/10.1038/nature19792.
Parker, J., Rambaut, A., Pybus, O.G., 2008. Correlating viral phenotypes with phylogeny: accounting for phylogenetic uncertainty. Infect. Genet. Evol. 8, 239–246.
Pina, S., Jofre, J., Emerson, S.U., Robert H., Girones, R., Purcell, R.H., 1998. Characterization of a Strain of Infectious Hepatitis E Virus Isolated from Sewage in an Area where Hepatitis E Is Not Endemic These include: Characterization of a Strain of Infectious Hepatitis E Virus Isolated from Sewage in an Area where Hepatitis E Is 64, 4485 4488 https://doi.org/10.1097/00006454-199903000-00030.
Rambaut, A., Suchard, M., Drummond, A., 2014. Tracer v1.6.
Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.a., Huelsenbeck, J.P., 2012. Mrbayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542. http://dx.doi.org/10.1093/sysbio/sys029.
Rothhammer, F., Dillehay, T.D., 2009. The late pleistocene colonization of South America: an interdisciplinary perspective. Ann. Hum. Genet. 73, 540–549. http://dx.doi.org/10.1111/j.1469-1809.2009.00537.x.
Rusiñol, M., Fernandez-Cassi, X., Timoneda, N., Carratala, A., Abril, J.F., Silvera, C., Figueras, M.J., Gelati, E., Rodo, X., Kay, D., Wyn-Jones, P., Bofill-Mas, S., Girones, R., 2015. Evidence of viral dissemination and seasonality in a Mediterranean river catchment: implications for water pollution management. J. Environ. Manage. 159, 58–67. http://dx.doi.org/10.1016/j.jenvman.2015.05.019.
Salakova, M., Koslabova, E., Vojtechova, Z., Tachezy, R., Sroller, V., 2016. Detection of human polyomaviruses MCPyV, HPyV6, and HPyV7 in malignant and non-malignant tonsillar tissues. J. Med. Virol. 88, 695–702.
Sastre-Garau, X., Martine, P., Avril, M.-F., Laude, H., Couturier, J., Rozenberg, F., Almeida, A., Boitier, F., Carlotti, A., Couturaud, B., Dupin, N., 2009. Merkel cell carcinoma of the skin: pathological and molecular evidence for a causative role of MCV in oncogenesis. J. Pathol. 218, 48–56.
Schowalter, R.M., Pastrana, D.V., Pumphrey, K.a., Moyer, A.L., Buck, C.B., 2010. Merkel cell polyomavirus and two previously unknown polyomaviruses are chronically shed from human skin. Cell Host Microbe 7, 509–515. http://dx.doi.org/10.1016/j.chom.2010.05.006.
Schrama, D., Groesser, L., Ugurel, S., Hafner, C., Pastrana, D.V., Buck, C.B., Cerroni, L., Theiler, A., Becker, J.C., 2014. Presence of human polyomavirus 6 in mutation-specific BRAF inhibitor-induced epithelial proliferations. JAMA Dermatol. 150, 1180–1186. http://dx.doi.org/10.1001/jamadermatol.2014.1116.
Shackelton, L.A., Rambaut, A., Pybus, O.G., Holmes, E.C., 2006. JC virus evolution and its association with human populations. J. Virol. 80, 9928–9933.
Shuda, M., Feng, H., Kwun, H.J., Rosen, S.T., Gjoerup, O., Moore, P.S., Chang, Y., 2008. T antigen mutations are a human tumor-specific signature for Merkel cell polyomavirus. Proc. Natl. Acad. Sci. USA 105, 16272–16277.
Spurgeon, M.E., Lambert, P.F., 2013. Merkel cell polyomavirus: a newly discovered human virus with oncogenic potential. Virology 435, 118–130. http://dx.doi.org/10.1016/j.virol.2012.09.029.
Torres, C., Barrios, M.E., Cammarata, R.V., Cisterna, D.M., Estrada, T., Martini Novas, S., Cahn, P., Blanco Fernández, M.D., Mbayed, V.A., 2016. High diversity of human polyomaviruses in environmental and clinical samples in Argentina: Detection of JC, BK, Merkel-cell, Malawi, and human 6 and 7 polyomaviruses. Sci. Total Environ. 542, 192–202. http://dx.doi.org/10.1016/j.scitotenv.2015.10.047.

Torres, C., Leone, F.G.P.Y, Pezzano, S.C., Mbayed, V.A., Campos, R.H., 2011. New perspectives on the evolutionary history of hepatitis B virus genotype F. Mol. Phylogenet. Evol. 59, 114–122. https://doi.org/10.1016/j.ympev.2011.01.010.

Trifinopoulos, J., Nguyen, L.T., von Haeseler, A., Minh, B.Q., 2016. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res. 44, W232–W235. http://dx.doi.org/10.1093/nar/gkw256.

Vahabpour, R., Aghasadeghi, M.R., Salehi-Vaziri, M., Mohajel, N., Keyvani, H., Nasimi, M., Esghaei, M., Monavari, S.H., 2016. Prevalence of Merkel cell polyomavirus in Tehran: an age-specific serological study. Iran. Red Crescent Med. J. 18. http://dx.doi.org/10.5812/ircmj.26097.

Van Der Meijden, E., Bialasiewicz, S., Rockett, R.J., Tozer, S.J., Sloots, T.P., Feltkamp, M.C.W., 2013. Different serologic behavior of MCPyV, TSPyV, HPyV6, HPyV7 and HPyV9 polyomaviruses found on the skin. PLoS One 8, 1–9. http://dx.doi.org/10.1371/journal.pone.0081078.

Victoria, M., Tort, L.F.L., García, M., Lizasoain, A., Maya, L., Leite, J.P.G., Miagostovich, M.P., Cristina, J., Colina, R., 2014. Assessment of gastroenteric viruses from wastewater directly discharged into Uruguay River, Uruguay. Food Environ. Virol. 6, 116–124. http://dx.doi.org/10.1007/s12560-014-9143-7.

Wieland, U., Silling, S., Hellmich, M., Potthoff, A., Pfister, H., Kreuter, A., 2014. Human polyomaviruses 6, 7, 9, 10 and trichodysplasia spinulosa-associated polyomavirus in HIV-infected men. J. Gen. Virol. 95, 928–932. http://dx.doi.org/10.1099/vir.0.061259-0.

Yogo, Y., Sugimoto, C., Zheng, H.-Y., Ikegaya, H., Takasaka, T., Kitamura, T., 2004. JC virus genotyping offers a new paradigm in the study of human populations. Rev. Med. Virol. 14, 179–191.

Zheng, H.Y., Nishimoto, Y., Chen, Q., Hasegawa, M., Zhong, S., Ikegaya, H., Ohno, N., Sugimoto, C., Takasaka, T., Kitamura, T., Yogo, Y., 2007. Relationships between BK virus lineages and human populations. Microbes Infect. 9, 204–213. http://dx.doi.org/10.1016/j.micinf.2006.11.008.