Supplementary Materials

Supplementary Table S1 - Human polyomaviruses classification, isolation origin, and associated human diseases.

Virus	Genus	Species	Abbreviation	Genotypes/Subtypes/Subgroups[a]	Isolation sites	Associated clinical diseases
BKPyV[1]	*Betapolyomavirus*	*Human polyomavirus 1*	HPyV1	I (Ia, Ib1, Ib2, Ic), II, III, IV (IVa1, IVa2, IVb1, IVb2, IVc1, IVc2)	Urine, feces	Haemorrhagic cystitis, nephropathy
JCPyV[2]	*Betapolyomavirus*	*Human polyomavirus 2*	HPyV2	1 (1A, 1B), 2 (2A/C [2A1, 2A2], 2B, 2D [2D1, 2D2, 2D3], 2E), 3 (3A, 3B), 4, 6, 7 (7A, 7B [7B1, 7B2], 7C [7C1, 7C2]), 8 (8A, 8B), Eu-c	Brain tissue, urine	Progressive multifocal leukoencephalopathy
KIPyV[3]	*Betapolyomavirus*	*Human polyomavirus 3*	HPyV3	n.d.	Respiratory tract	None
WUPyV[4]	*Betapolyomavirus*	*Human polyomavirus 4*	HPyV4	I (Ia, Ib, Ic, Id), II, III (IIIa, IIIb, IIIc)	Respiratory tract	None
MCPyV[5]	*Alphapolyomavirus*	*Human polyomavirus 5*	HPyV5	Europe/North America, Africa, Asia, Oceania, South America	Skin cancer	Merkel cell carcinoma
HPyV6[6]	*Deltapolyomavirus*	*Human polyomavirus 6*	HPyV6	Worldwide, Asia	Healthy skin, skin tumors	Pruritic dermatitis
HPyV7[7]	*Deltapolyomavirus*	*Human polyomavirus 7*	HPyV7	n.d.	Healthy skin, skin tumors	Pruritic dermatitis
TSPyV[8]	*Alphapolyomavirus*	*Human polyomavirus 8*	HPyV8	n.d.	Skin lesions	Trichodysplasia spinulosa
HPyV9[9]	*Alphapolyomavirus*	*Human polyomavirus 9*	HPyV9	n.d.	Serum, urine	None
HPyV10[10]	*Deltapolyomavirus*	*Human polyomavirus 10*	HPyV10	n.d.	Diarrheic and non-diarrheic feces	None
STLPyV[11]	*Deltapolyomavirus*	*Human polyomavirus 11*	HPyV11	n.d.	Diarrheic and non-diarrheic feces	None
HPyV12[12,b]	n.d.	*Human polyomavirus 12*	HPyV12	n.d.	Gastrointestinal tract, liver	None
NJPyV[13]	*Alphapolyomavirus*	*Human polyomavirus 13*	HPyV13	n.d.	Muscle tissue	None
LIPyV[14]	*Alphapolyomavirus*	*Human polyomavirus 14*	HPyV14	n.d.	Blood, skin	None
QPyV[15,c]	n.d.	n.d.	HPyV15	n.d.	Feces	None

"n.d." stands for "non determined".

[1]BK polyomavirus; [2]John Cunningham polyomavirus; [3]Karolinska Institute polyomavirus.; [4]Washington University polyomavirus; [5]Merkel Cell Polyomavirus; [6]Human polyomavirus 6; [7]Human polyomavirus 7; [8]Trichodysplasia spinulosa polyomavirus; [9]Human polyomavirus 9; [10]Human polyomavirus 10; [11]Saint Louis polyomavirus; [12]Human polyomavirus 12; [13]New Jersey polyomavirus; [14]Lyon-IARC polyomavirus; [15]Quebec polyomavirus.
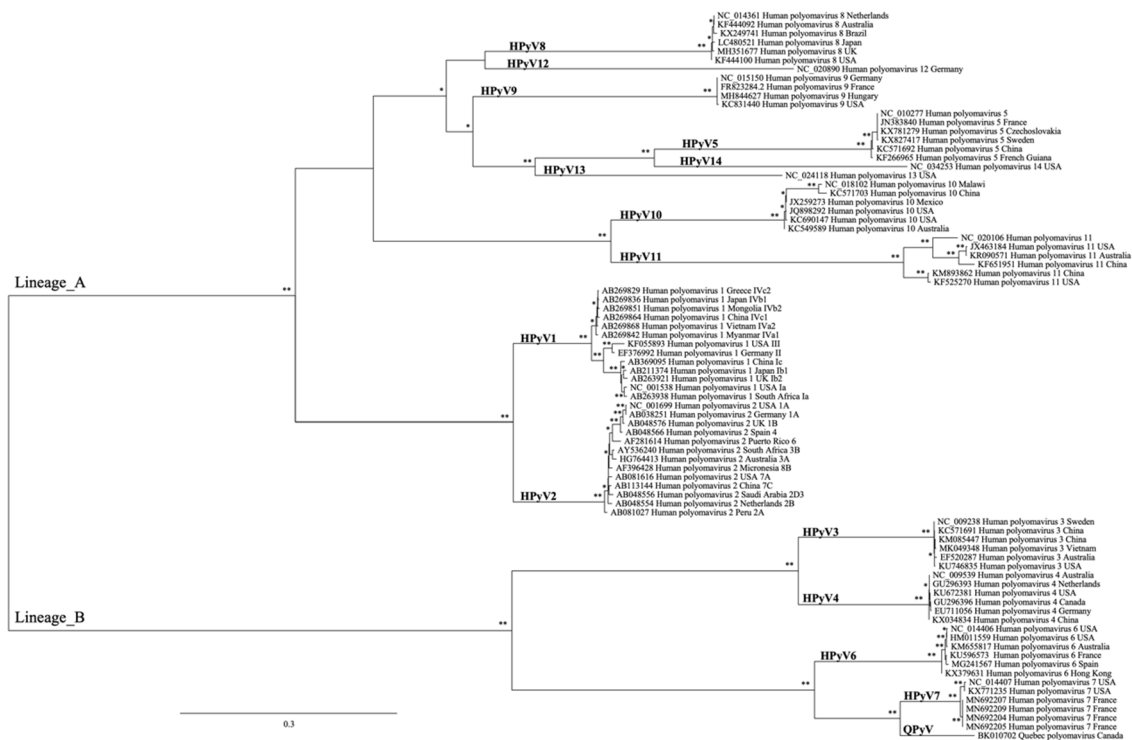
[a] Subtypes are indicated inside the parentheses while the respective subgroups are indicated inside square brackets.

[b] Excluded from the family on the ICTV 2019 report.

[c] Recently described, not yet featured on the current family *Polyomaviridae* taxonomy report (ICTV 2019).

Supplementary Table S2 - Analysis of the performance of the primers used for amplification of HPyV sequences in a singleplex vs multiplex format.

|	1st-round					
	PCR-A				PCR-B	
	HPyV1/2	HPyV5	HPyV8/9	HPyV10/11	HPyV3/4	HPyV6/7
Site 2[wastewater]	+	-	-	-	-	+
Site 7[wastewater]	+	+	-	-	-	+
Site 10[environmental]	-	-	-	-	-	-
Site 14[environmental]	-	-	-	-	-	-
	2nd-round					
	PCR-A				PCR-B	
	HPyV1/2	HPyV5	HPyV8/9	HPyV10/11	HPyV3/4	HPyV6/7
Site 2[wastewater]	+	-	-	-	-	+
Site 7[wastewater]	+	+	-	-	-	+
Site 10[environmental]	+	-	-	-	-	-
Site 14[environmental]	-	-	-	-	-	-

Supplementary Figure S1 - Analysis of the phylogenetic relationships of 15 HPyV lineages (maximum likelihood) using the complete structural protein coding-sequence of references downloaded from the public genomic databases. Each sequence is identified by its accession number, HPyV type, country of origin, and, when possible, viral genotype and/or subtype. At the main tree branches ** indicates the support revealed by relevant aLRT and bootstrap values (≥75%). The bar indicates the average number of substitutions per site.
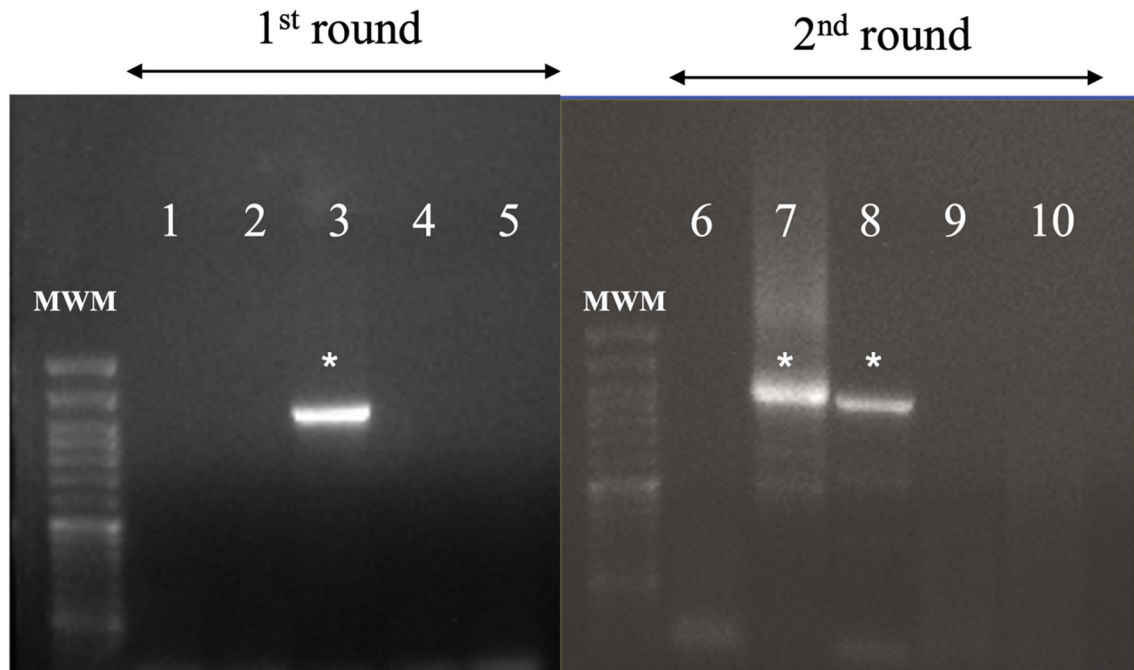
Supplementary Figure S2 - Analysis of the sensitivity/specificity of amplification using the PCR-A and B touchdown protocols. MWM indicates the NZYtech Molecular Weight marker VI (NZYtech, Portugal). Lanes 1 and 6: 1st and 2nd round negative controls for PCR-A; lanes 2 and 7: analysis of the 1st- (lane 2) and 2nd-round (lane 7) amplification reactions of PCR-A using a positive control [a HPyV2 DNA extract (diluted 1:10,000) was used in the 1st-round (lane 2) and a 1:50 dilution of the 1st-round product as input in the 2nd-round (lane 7)]; lanes 3 and 8: 1st- (lane 3) and 2nd- (lane 8) round of a singleplex reaction using primers HPyV1/2Fo and HPyV1/2Ro or HPyV1/2Fi and HPyV1/2Fi, for the 1st and 2nd round, respectively. The input DNA was the same as the one mentioned for the reactions in lanes 2 and 7; lanes 4 and 9: negative control (1st and 2nd-round, respectively) for PCR-B; lanes 5 and 10: analysis of 1st (lane 5) and 2nd-round (lane 10) amplification reactions of PCR-B using as input for the 1st-round a diluted (1:10,000) HPyV2 DNA extract and as 1:50 dilution of the product obtain in the 1st-round as input in the 2nd-round (lane 10). The "*" indicates the expected band with approximately 1100 bp (1st-round) and 990 (2nd-round).
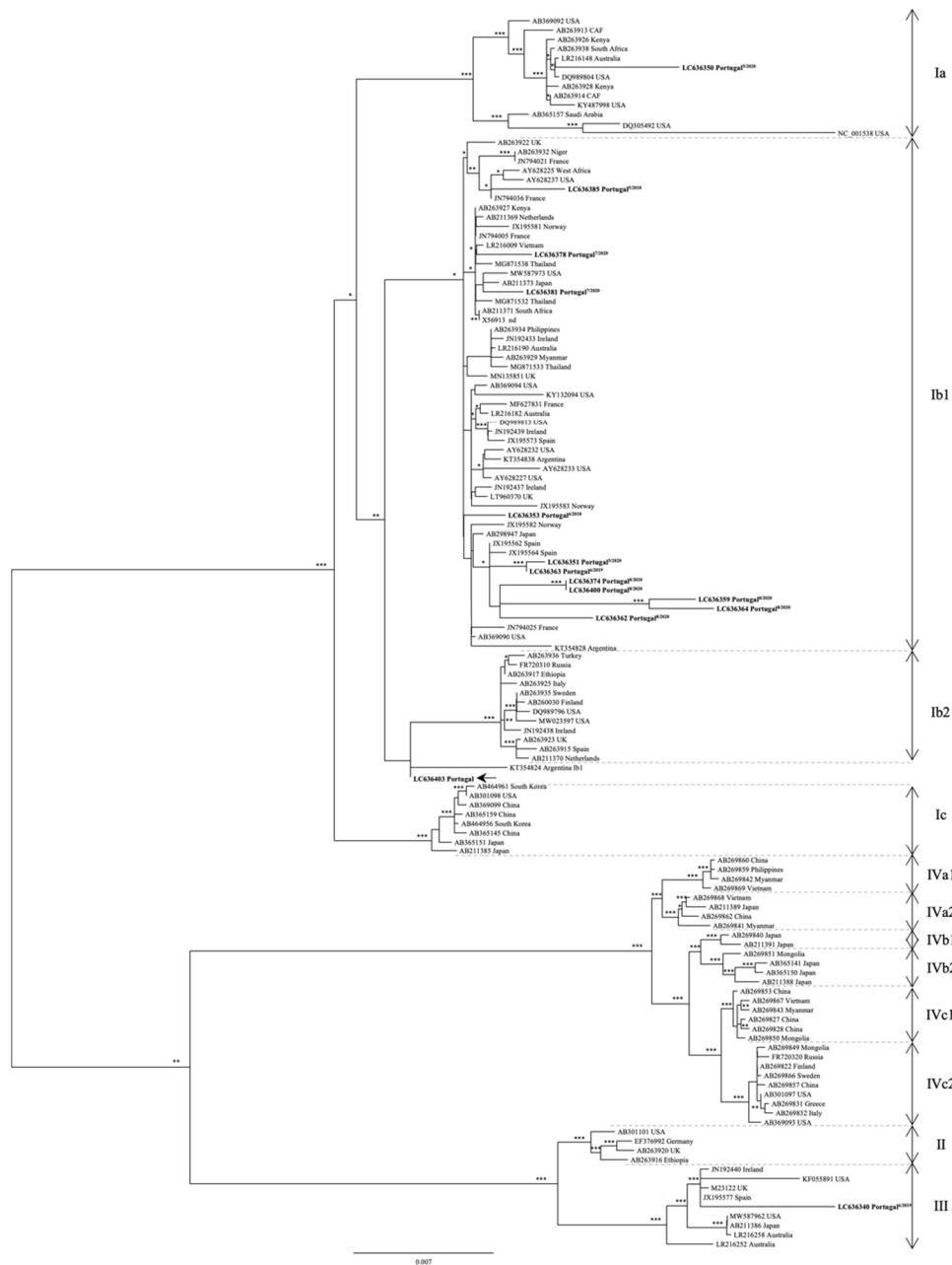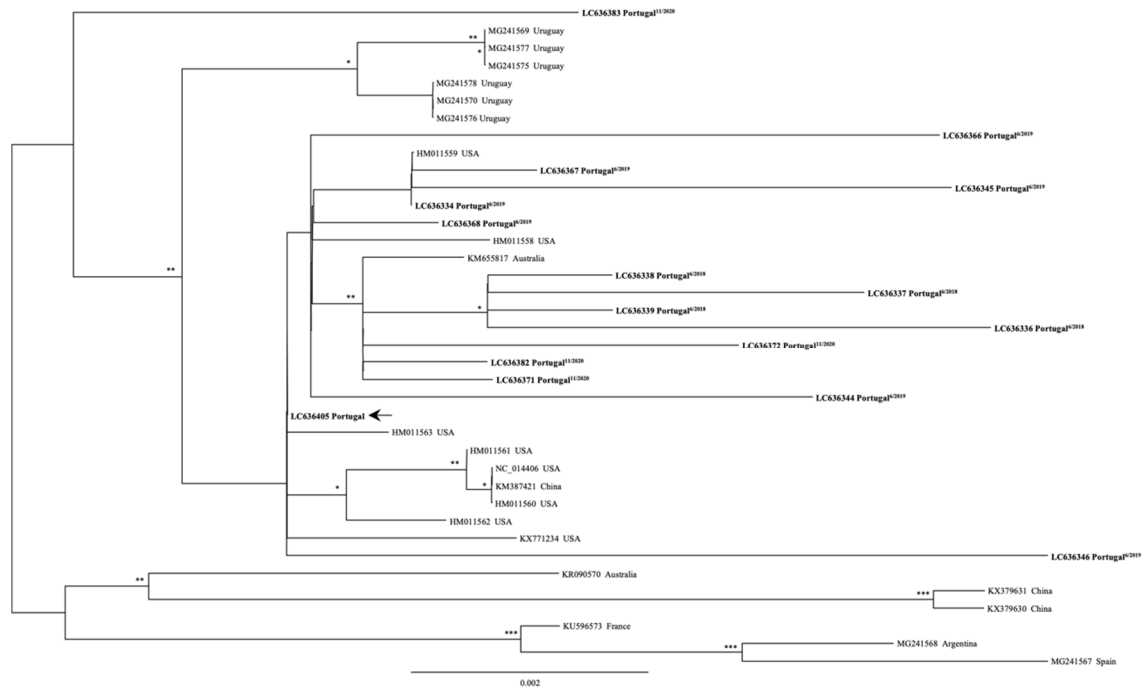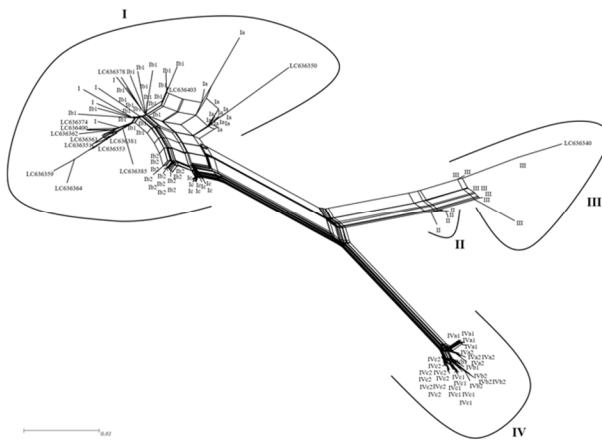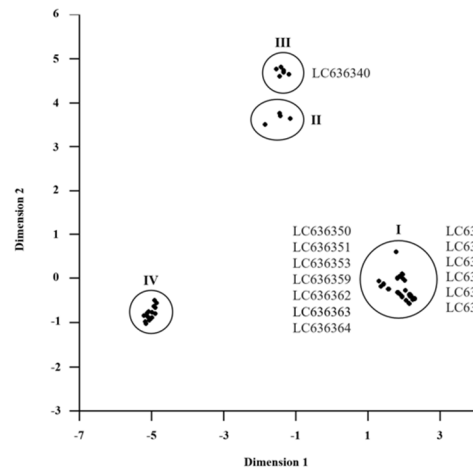
Sup_Fig_3a
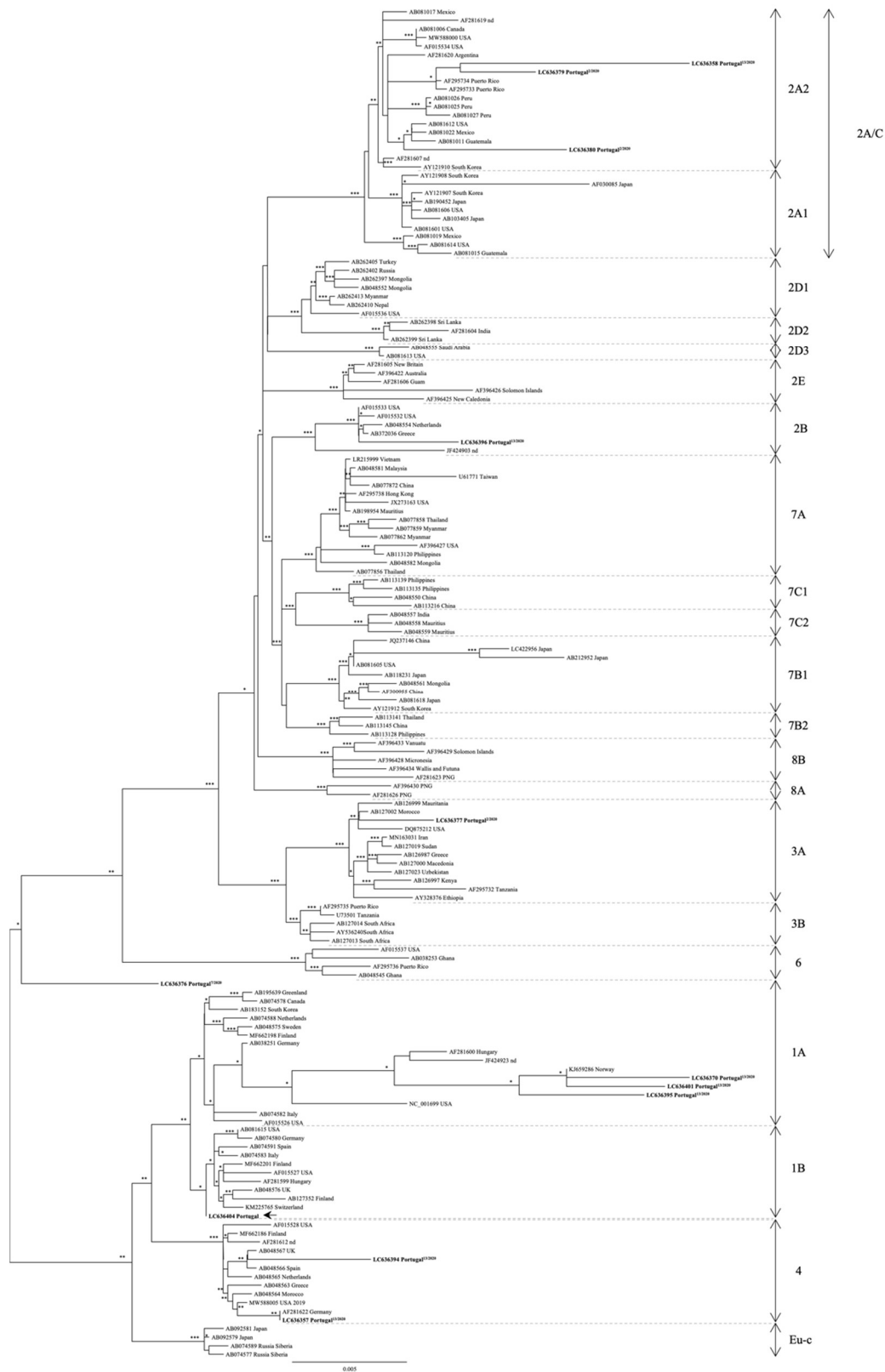
Sup_Fig_3b
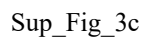
Sup_Fig_3c

Sup_Fig_3d

Supplementary Figure S3 - Phylogenetic analysis of HPyV1 (a), HPyV2 (b), HPyV5 (c), and HPyV6 (d) sequences (maximum likelihood) considering a structural protein-coding genome region. At specific branches, the identity of the HPyV is indicated by their accession number and country of origin. For the analysis of HPyV1, HPyV2 and HPyV5, the different viral genotypes/subtypes, as defined previously [14,20], are indicated. The sequences obtained during this study are indicated in boldface. In the latter, the numbers indicated in superscript refer to the identity of the sample collection sites, as they are indicated in Fig.1. The consensus sequence generated by NGS is indicated by the arrow. At the main tree branches, the number of "*" indicates the support revealed by the different phylogenetic reconstruction methods used, assuming as relevant aLRT and bootstrap values ≥75%, as well as posterior probability values ≥0.80 when a Bayesian approach was further used to confirm the proposed phylogenetic clustering. The bars indicate the average number of substitutions per site.
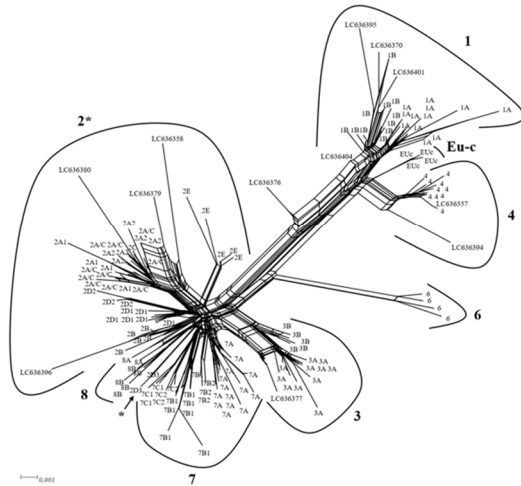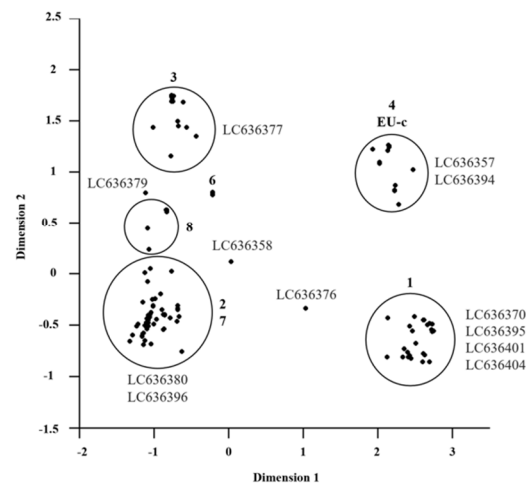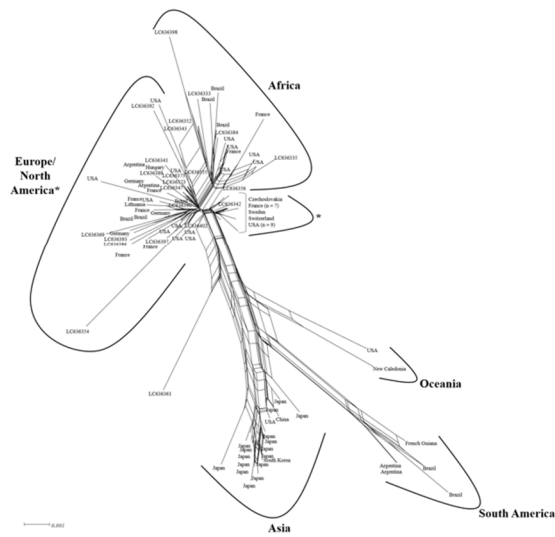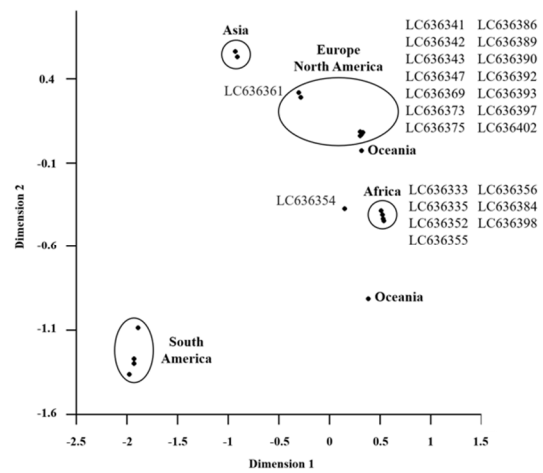
A1.

A2.

Sup_Fig_4a

B1.

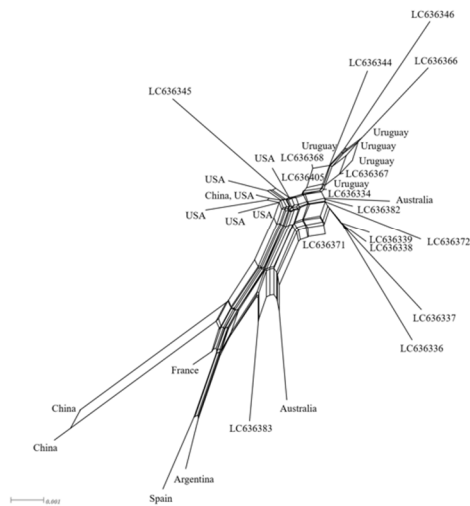**B2.**

Sup_Fig_4b

C1.

Sup_Fig_4c

C2.

D1.
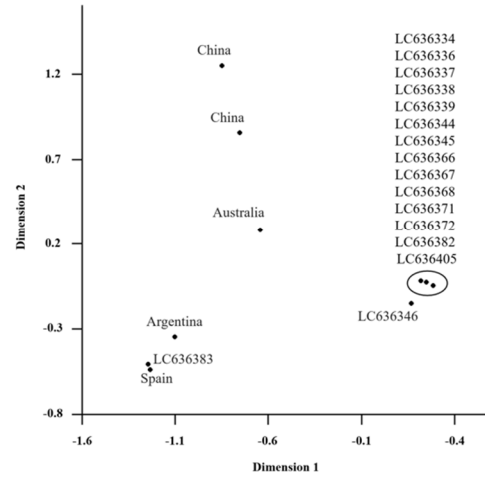

Sup_Fig_4d

D2.

Supplementary Figure S4 - NeighborNet networks (A) and PCOORD (B) analyses of HPyV1 (A1/B1), HPyV2 (A2/B2), HPyV5 (A3/B3), and HPyV6 (A4/B4) sequences. The sequences obtained in the course of this study are indicated by their accession number. In the PCOORD graphs, the two first dimensions accounts for 50.30% (B1), 47.02% (B2), 73.95% (B3), and 40.20% (B4) if the total nucleotide differences, while the first 10 axes cover 90.60% (B1), 78.78% (B2), 90.88% (B3), and 67.75% (B4) of the dataset sequence variation.