New Kaposi's sarcoma-associated herpesvirus variant in men who have sex with men 3 associated with severe pathologies Running title: New KSHV variant identified in MSM

Aude Jary, Valentin Leducq, Nathalie Désiré, Héloïse Petit, Romain Palich, Véronique Joly, Ana Canestri, Adélie Gothland, Sidonie Lambert-Niclot, Laure Surgers, et al.

Category: Major article

Title: New Kaposi's sarcoma-associated herpesvirus variant in men who have sex with men associated with severe pathologies

Running title: New KSHV variant identified in MSM

Aude JARY[1], Valentin LEDUCQ[1], Nathalie DESIRE[1], Héloïse PETIT[1], Romain PALICH[2], Véronique JOLY[3], Ana CANESTRI[4], Adélie GOTHLAND[1], Sidonie LAMBERT-NICLOT[5], Laure SURGERS[6], Corinne AMIEL[7], Diane DESCAMPS[8], Jean-Philippe SPANO[9], Christine KATLAMA[2], Vincent CALVEZ[1], Anne-Geneviève MARCELIN[1]

[1]Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique (iPLESP), AP-HP, Pitié Salpêtrière Hospital, Department of Virology, F-75013 Paris, France

[2]Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique (iPLESP), AP-HP, Pitié Salpêtrière Hospital, Department of Infectious Diseases, F-75013 Paris, France

[3]IAME, UMR 1137, INSERM, Université Paris Diderot, Sorbonne Paris Cité, AP-HP, Service de Maladies Infectieuses et Tropicales, Hôpital Bichat, AP-HP, Paris, France

[4]Service de Maladies Infectieuses et Tropicale, AP-HP Hôpital Tenon, Paris, France

[5]Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique (iPLESP), AP-HP, Hôpital Saint Antoine, Service de Virologie, Paris, France

[6]Sorbonne Université, INSERM, Centre d'Immunologie et des Maladies Infectieuses (CIMI), AP-HP, Hôpital Saint Antoine, Service de Maladies Infectieuses et Tropicales, Paris, France

[7]Service de Virologie, AP-HP Hôpital Tenon, Paris, France

[8]IAME, UMR 1137, INSERM, Université Paris Diderot, Sorbonne Paris Cité, AP-HP, Service de Virologie, Hôpital Bichat, AP-HP, Paris, France

[9]Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique (iPLESP), AP-HP, Hôpital Pitié Salpêtrière, Service d'Oncologie Médicale, Paris, France

Summary

We identified a new Kaposi sarcoma-associated herpesvirus F variant in 5 Caucasian men who have sex with men. Careful screening may be required in this population, given the severe clinical presentation of associated diseases in the context of immunosuppression.

Abstract (word count=200)

Background: Kaposi's sarcoma-associated herpesvirus (KSHV) subtype depends mostly on patient origin. This study aimed to assess KSHV-diversity in a population of MSM living in France.

Methods: We included 264 patients; (i) 65 MSM: 57 HIV-infected with Kaposi's sarcoma (KS), multicentric Castleman disease (MCD) or primary effusion lymphoma (PEL), and 8 HIV-uninfected on HIV-pre-exposure prophylaxis (PrEP) to perform KSHV-typing by ORF-K1 Sanger and KSHV-whole-genome sequencing, (ii) 199 other patients for real-time PCR screening for the new variant.

Results: We found that 51% of KSHV-strains were subtype C (85% C3) and 33% were subtype A. Four patients with severe KSHV-disease (2 visceral KS, 1 MCD, 1 PEL) and one asymptomatic PrEP user had a new variant resembling the Ugandan subtype F, but with different ORF-K1 and KSHV-whole-genome sequences and a different epidemiological context (MSM *versus* African population). Its prevalence was 4.5% in Caucasian MSM and absent in other epidemiological groups.

Conclusions: Subtype C predominated among MSM living in France. The new F-variant was identified in Caucasian MSM and associated with severe KSHV-disease, suggesting that subtype F could be split into F1 and F2 variants. Careful screening for this variant may be required in MSM given the severe clinical presentation of associated diseases.

Keywords: KSHV, MSM, new variant, ORF-K1, phylogenetic analysis, whole-genome sequencing

Word count = 2671

INTRODUCTION

Kaposi's sarcoma-associated herpesvirus (KSHV) is the recognized etiologic agent of all epidemiological forms of Kaposi's sarcoma (KS), including classic, endemic, post-transplant and epidemic (HIV-associated) forms. This virus is also involved in the development of two lymphoid malignancies: primary effusion lymphoma (PEL) and multicentric Castleman disease (MCM), mostly in immunocompromised patients.

The K1 open reading frame (ORF-K1) encodes a transmembrane glycoprotein with an amino-acid sequence varying by 20% to 44% between KSHV subtypes and by about 10% within subtypes (genotype variant) [1]. A molecular epidemiological analysis of ORF-K1 led to the identification of seven KSHV subtypes (A, B, C, D, E, F and Z), the distribution of which worldwide depends on patient origin. Subtypes A and C were found in Europe, North America, the Middle East, the Mediterranean and Asia [2–4]; subtypes B and A5 were found predominantly in sub-Saharan Africa [5]; subtype D were found on Pacific islands and in Taiwan [6]; subtype E was found in Brazilian Indians [7,8], subtype F in individuals from Uganda [9] and subtype Z was identified in a small cohort of Zambian children [10].

Among men who have sex with men (MSM), KSHV-seroprevalence is higher than in the general population of Western Europe (<5%) [11]. A recent systemic review and meta-analysis reported a pooled KSHV-seroprevalence in both HIV-infected and uninfected MSM of 33% [12]. The C and A subtypes of KSHV are the most prevalent in European and Asian MSM with KSHV-associated diseases [13–15], consistent with the prevalence of KSHV subtypes in the general populations of these regions [1–3]. However, although the KSHV subtype distribution depends mostly on the patient's region of origin, several studies have suggested that it may also be affected by clinical presentation or progression, particularly for different forms of KS. For example, the A5 variant has been associated with extensive disease

in epidemic KS form [5], the A subtype has been associated with rapidly evolving classic KS form [16] and the A and B' subtypes have been shown to contrast with the C subtype by occurring at extracutaneous sites in post-transplant KS form [17].

We conducted a retrospective study analyzing HIV-infected MSM living in France with KS, MCD or PEL, to describe KSHV subtype diversity and potential associations with the severity of clinical presentation. As it was clearly demonstrated that KSHV-prevalence infection correlated with the number of MSM partners [18] and considering the fact that this population could be exposed to KSHV-transmission by frequenting dense sexual network, we also included HIV-seronegative MSM under HIV pre-exposure prophylaxis (PrEP) program to compare KSHV strains.

METHODS

Study population

We studied 65 MSM: (i) 57 HIV-infected patients with KSHV-associated diseases diagnosed between 2012 and 2017 in France and (ii) 8 MSM HIV-uninfected participants on HIV-PrEP with positive KSHV-antibodies and displaying KSHV-DNA shedding in a buccal swab sample. Demographic and medical data were collected, including age, country of origin, CD4 count, HIV-RNA viral load, KSHV-DNA viral load and clinical presentation of KSHV-associated diseases. Severe clinical presentation was defined as visceral KS or lymphoid malignancies.

We subsequently included 199 patients testing positive for KSHV-DNA since 2013, diagnosed in our department, to estimate the prevalence of the new variant identified. For each patient, demographic data were collected, including sex, age, sexual orientation, country of origin and KSHV-DNA viral load.

KSHV typing

For HIV-infected patients, we obtained a whole-blood sample at the time of KSHV disease diagnosis. For PrEP-using MSM, we obtained a buccal swab sample during standard medical follow-up. DNA was extracted from this samples and subjected to real-time PCR to amplify both ORF-73 (encoding the latency-associated nuclear antigen, LANA) and the albumin gene, as previously described [19].

- *Sanger sequencing of ORF-K1 or VR1*

A 679 bp fragment of ORF-K1 including the two hypervariable regions, VR1 (amino acids (aa) 54 to 93) and VR2 (aa 191 to 228), was amplified by nested-PCR as previously described [20]. If ORF-K1 amplification was unsuccessful, a second nested-PCR was performed to amplify a 363 bp fragment including only VR1, as previously described [2]. Bidirectional sequencing was performed with BigDye Terminator chemistry (Thermo Fisher Scientific®), with analysis of the reaction products on an ABI sequencer.

- *KSHV whole-genome sequencing and assembly*

Sample libraries preparations and target enrichment were performed according to the *SeqCap EZ HyperCap workflow* (Roche®). The DNA sample was fragmented mechanically and specific adapters were added. DNA libraries were pooled and enriched in KSHV sequences by two rounds of hybridization with 100 bp overlapping DNA probes designed in conjunction with Roche® on the KSHV GK18 (AF148805.2) sequence (excluding repeat regions) and on 28 ORF-K1 reference sequences available from the NCBI database (see Supplementary Table 1), to give 5x coverage. Finally, next-generation sequencing was

performed with paired-end reads (Mid Output Kit v2, 2x75 bp) on the NextSeq500 Illumina® system.

Reads were trimmed with Trimmomatic, using a quality (Q) threshold such that only bases with Q>30 were retained and reads of less than 50 bp were filtered out. Paired-end reads were first mapped onto the KSHV reference sequence (GK18) with Bowtie 2.3.4.3. They were then assembled *de novo* with Spades3.12.0 and Mira1.1.1 to generate two other sequences. Finally, we used Mauve1.1.1 in Geneious11.1.4 to align the three sequences and establish the consensus whole-genome KSHV sequence (KSHV-WG).

- *Phylogenetic analysis*

We performed a phylogenetic analysis of ORF-K1 amino-acid sequences and KSHV-WG nucleotide sequences by the maximum likelihood method. Multiple sequence alignments were generated with Mafft7 [21], and a phylogenetic analysis was then performed with PhyML3.0 [22] and 1000 bootstraps resampling. Pairwise genetic distances between ORF-K1 amino-acid sequences were calculated with Mega7.0.14 [23], using the JTT model of substitutions and a gamma distribution with 4 parameters.

Screening for the new F-variant by real-time PCR

Whole-blood samples with positive KSHV-DNA viral load (n=199) were screened for the new variant with a specific real-time PCR. We designed specific primers and probes with Geneious 11.1.4, to amplify a fragment of about 120 bp in length encompassing the VR1 region and specific for the new KSHV variant (see Supplementary Table 2). The specificity of that PCR was assessed by testing 25 different subtypes/genotype variants and the sensitivity at 25 copies/10^6 cells.

Statistical analysis

Continuous variables were expressed as the median and interquartile range [IQR] and discrete variables were expressed as numbers and percentages. GraphPad was used to perform non-parametric tests, specifically Mann-Whitney U tests for quantitative data and Fisher's exact test for qualitative data.

Ethics statement

The study was carried out in accordance with the Declaration of Helsinki. It was a retrospective non-interventional study with no addition to standard care procedures. Reclassification of biological remnants into research material after completion of the ordered virological tests was approved by the local interventional review board of Pitié-Salpêtrière hospital. According to the French Public Health Code (CSP Article L.1121-1.1) such protocols are exempted from individual informed consent.

RESULTS

Patients' characteristics

In total, 41 of the HIV-infected men had been diagnosed with KS, 12 with MCD and 4 with PEL. Most were Caucasian (50/57, 88%); 46/57(81%) originated from France, 6/57 (10%) from the Mediterranean Basin, 4/57 (7%) from South America and 1/57 (2%) from Asia (Table 1). KSHV-DNA viral load in whole blood was lower in patients with KS, than in those with MCD or PEL ($p<0.0001$). Most of the PrEP users were also Caucasian (7/8, 88%); 4 (50%) originated from France, one from Switzerland, two from Eastern Europe and one from South America. None of the PrEP users were infected with HIV or had KSHV-associated diseases.

ORF-K1 phylogenetic analysis

ORF-K1 amplification was successful for 34 of the 57 MSM with KSHV diseases (60%), and related to the rate of KSHV-DNA viral load. For patients with all KSHV-related diseases considered together, subtype C was the most prevalent [18/34 (53%)], followed by subtype A [11/34 (32%)] and finally subtype B [1/34 (3%)]. For subtype C, 15/18 cases (83%) were classified as the C3 variant. Subtype classification was possible for five of the eight PrEP users (63%): the C3 variant in two cases and the A4 variant in two cases (Figure 1) (see Supplementary Table 3). The genetic distance (GD) between subtypes B and A was 34%, that between subtypes B and C was 39%, and the ORF-K1 amino-acid sequences of subtypes A and C differed by 19%. Within subtypes, genotype variant differed from 9% (subtype A) and from 13% (subtype C) at amino-acid level.

Interestingly, 5 KSHV strains (from P030, P035, P075, P076 and PrEP004) were closely related to the F subtype described in an African population from Uganda (AAX55469_F) [9]. However, their amino-acid sequences differed from those of the F subtype by 11% and were separated with a bootstrap confidence of 54% (Figure 1).

Overall, we found 3 clusters, corresponding to the C3 variant (17/39), the A4 variant (7/39) and the "F subtype" (5/39). In most cases, the geographic origin of the MSM was consistent with the subtype detected. Subtypes A and C were the most prevalent in patients originating from France, the Mediterranean Basin and Asia (see Supplementary table 3).

KSHV subtype and clinical presentation

The proportions of patients with KS and MCD did not differ significantly between subtypes: 26% of KS patients and 36% of MCD patients had subtype A viruses ($p=0.69$) and 63% of KS and 55% of MCD patients had subtype C viruses ($p=0.71$), suggesting that the subtypes

involved in these two diseases may reflect the prevalence of the various subtypes in this population.

Among patients with KS: (i) KSHV-DNA viral load tended to be higher for subtype A than for subtype C (*p=0.051*), regardless of immunovirological status (Figure 2), (ii) the C3 variant was more associated with purely cutaneous and/or oral-mucous involvement than the other subtypes (odds ratio=11.6, 95% CI: 1.1-214.2, *p=0.023*), regardless of immunovirological status ($CD4^+$ T-cell counts: C3 variant: 70/mm^3 [21-325] *vs.* other subtypes: 82.5/mm^3 [68-213], *p=0.78* - HIV-RNA viral load: C3 variant: 5.62 log_{10} copies/ml [1.60-5.89] *vs.* other subtypes: 5.38 log_{10} copies/ml [4.71-5.65], *p=0.75*).

New KSHV F variant

A new genotype variant closely resembling the F subtype was detected in five MSM, four of whom had severe clinical presentation: two cases of visceral KS (P030 with cutaneous and pulmonary involvement and P035 with cutaneous, pulmonary and ganglionic involvement), one case of MCD associated with visceral KS (P075), and one case of PEL associated with cutaneous KS (P076). The remaining individual was an asymptomatic PrEP user (PrEP004). All these individuals were Caucasian (four from France and one from Peru) (see Supplementary Table 4).

- *ORF-K1 Sanger sequencing*

Partial ORF-K1 sequences, 191 aa long (P030, P075, P076) and encompassing VR1 and VR2, or 142 and 141 aa long (from P035 and PrEP004, respectively) and encompassing only VR1, were generated; all sequences were strictly identical. The longest fragment differed by 18 aa from the reference subtype F sequence from Uganda (AAX55469_F), and an insertion of 5 aa was detected in VR2 (Figure 3). Otherwise, these sequences were closest (GD=10^{-6}) to the KSHV sequence described in 2000 from a French HIV-positive MSM patient with PEL

(AAG01610_F) [15], from which they differed by only one amino acid, in position 202 (Figures 1 and 3).

We also compared these sequences with that for a subtype F KSHV strain from our database that was obtained from a Congolese woman with MCD (P072). The ORF-K1 sequence from this woman differed by 23 aa from our virus and by 20 aa (15 aa and the insertion described above) from AAX55469_F. However, P072 and AAX55469_F clustered together in the phylogenetic tree and this sequence was considered to correspond to an "African" F subtype (Figures 1 and 3).

- *KSHV whole-genome sequencing*

For the formal identification of a new F variant, we performed KSHV-WG sequencing on the four KSHV strains newly identified from patients with KSHV-related diseases, and compared the sequences obtained with that for the P072 virus identified as belonging to the "African" F subtype, because no KSHV-WG F subtype reference sequence was available in the NCBI database. A consensus KSHV-WG sequence was obtained for all patients other than P035 (see Supplementary Table 4).

As in the phylogenetic analysis for ORF-K1, the phylogenetic tree generated for KSHV-WG sequences presented a separation of subtype B from subtype A/C/F, with a bootstrap confidence level of 100%. However, subtypes A and C were not clearly separated, suggesting the probably involvement of other genes in KSHV variability, as well as recombination events, as previously reported [24,25]. Phylogenetic analysis also confirmed that the two F variants (Caucasian MSM *versus* African) differed and were separated with a bootstrap confidence level of 61% (Figure 4). Although ORF-K1 amino-acid sequences were strictly identical between P030, P075 and P076, KSHV-WG phylogenetic tree showed that new F variant from KS patients differed from that in patients with lymphoid malignancies.

- *Prevalence of the new F variant*

Finally, we screened all available samples testing positive for KSHV-DNA since 2013 (*n*=199) with a new F variant-specific real-time PCR. The median KSHV-DNA viral load was 1.92 [1.45-2.69] $\log_{10}$ copies/10^6 cells. The patients tested had a median age of 54 [42-62.5] years and 78% were men (156/199); 47% were Caucasian (73/156), and 77% (56/73) of these Caucasians were MSM. None of the 199 samples tested was positive for the new F variant. Overall, this new F variant was only described in Caucasian MSM patients with a prevalence of 4.5% (5/113) in this population.

DISCUSSION

This study aimed to assess KSHV diversity in a population of MSM with and without KSHV-associated disease living in France and to evaluate the possible correlation between subtype and clinical presentation.

The global distribution of KSHV subtypes and variants in our MSM population was consistent with the findings of previous studies conducted in France [15] and other countries (Germany and Taiwan) [13,14], with KSHV subtypes C (*n*=18) and A (*n*=11) the most prevalent.

We also found that the C3 variant was the most prevalent KSHV strain in a population of MSM living in France, and that this variant was associated with a less severe clinical presentation of the epidemic form of KS. In patients with epidemic KS, KSHV-DNA viral load tended to be higher for subtype A than for subtype C. These results are consistent with those of previous studies reporting a more aggressive clinical presentation in patients with subtype A than in those with subtype C, for the classic and post-transplantation forms of KS [16,17].

Although subtype F has been described only a few times in the literature [9,26,27], we identified a new KSHV variant closely related to the first subtype F virus described in Uganda. This new F variant was identified by Sanger sequencing and confirmed by whole-genome sequencing on various samples from five patients. Based on the results of whole-genome sequencing, we were able to classify this KSHV strain as a new F variant. Finally, based on our results, we propose the subdivision of subtype F into two variants: F1 variant for the KSHV described in Uganda and F2 variant for the KSHV identified in Caucasian MSM.

All the patients harboring this new F2 variant were Caucasian MSM living in Paris. The screening of all our available samples since 2013 did not detect this new variant in other patients, which suggests that it is present in a small, restrictive population. The overall prevalence of the F2 variant was 4.5% (5/113) in the Caucasian MSM population, and this variant was absent from other epidemiological groups. The patients harboring the new F2 variant included one immunocompetent PrEP user who was found to be merely a carrier, with oral shedding but with no KSHV-related disease. The other four patients harboring the new variant were immunocompromised (AIDS) and had severe forms of KSHV disease. Further investigations are required to confirm these results and to determine whether specific subtypes or viral determinants of virulence involved in KSHV tumor-associated processes can lead to severe or persistent KSHV-related diseases, as recently reported for KS in HIV-infected patients on effective antiretroviral therapies [28,29].

In conclusion, careful screening of the MSM population may be required for this new F2 variant, which is circulating in Caucasian MSM living in Paris, given the severe clinical presentation of associated diseases in the context of immunosuppression.

ACKNOWLEDGMENT

Funding. This work was funded by the *Agence Nationale de Recherche sur le SIDA et les Hépatites Virale*s (ANRS) AC43 "sexually transmitted infections" working group.

Conflicts of interest. AJ, VL, ND, HP, RP, VJ, AC, AG, SLN, LS, CA, JPS, CK and VC have nothing to disclose. DD reports personal fees from Gilead Sciences, personal fees from MSD, personal fees from ViiV Healthcare, personal fees from Janssen-Cilag, outside the submitted work. AGM reports grants and personal fees from VIIV Healthacare, grants and personal fees from Gilead, grants and personal fees from MSD, personal fees from Janssen-Cilag, outside the submitted work.

Meeting presentation. This work was presented as a poster discussion at the Conference on Retroviruses and Opportunistic Infections (CROI), in March 2019 (Themed discussion: KSHV virology and pathogenesis, poster no. 273).

Corresponding author: Aude Jary, 47-83 boulevard de l'hôpital, 75013 Paris, +33142177406, aude.jary@aphp.fr

Alternate corresponding author: Anne-Geneviève Marcelin, anne-genevieve.marcelin@aphp.fr

REFERENCES

1. Zong JC, Ciufo DM, Alcendor DJ, et al. High-level variability in the ORF-K1 membrane protein gene at the left end of the Kaposi's sarcoma-associated herpesvirus genome defines four major virus subtypes and multiple variants or clades in different human populations. J Virol. **1999**; 73(5):4156–4170.

2. Ouyang X, Zeng Y, Fu B, et al. Genotypic analysis of Kaposi's sarcoma-associated herpesvirus from patients with Kaposi's sarcoma in Xinjiang, China. Viruses. **2014**; 6(12):4800–4810.

3. Ötvös R, Juhasz A, Szalai E, et al. Molecular typing of human herpesvirus 8 isolates from patients with Kaposi's sarcoma in Hungary. Anticancer Res. **2014**; 34(2):893–898.

4. Varmazyar S, Shoja Z, Kakavand-Ghalehnoei R, Shahmahmoodi S, Marashi SM, Jalilvand S. Molecular typing of human herpesvirus 8 among HIV positive in comparison to HIV-negative individuals in Iran. J Med Virol. **2017**; 89(4):703–709.

5. Isaacs T, Abera AB, Muloiwa R, Katz AA, Todd G. Genetic diversity of HHV8 subtypes in South Africa: A5 subtype is associated with extensive disease in AIDS-KS. J Med Virol. **2016**; 88(2):292–303.

6. Cassar O, Charavay F, Bassot S, et al. Divergent KSHV/HHV-8 subtype D strains in New Caledonia and Solomon Islands, Melanesia. J Clin Virol Off Publ Pan Am Soc Clin Virol. **2012**; 53(3):214–218.

7. Whitby D, Marshall VA, Bagni RK, et al. Genotypic characterization of Kaposi's sarcoma-associated herpesvirus in asymptomatic infected subjects from isolated populations. J Gen Virol. **2004**; 85(Pt 1):155–163.

8. Biggar RJ, Whitby D, Marshall V, Linhares AC, Black F. Human herpesvirus 8 in Brazilian Amerindians: a hyperendemic population with a new subtype. J Infect Dis. **2000**; 181(5):1562–1568.

9. Kajumbula H, Wallace RG, Zong J-C, et al. Ugandan Kaposi's sarcoma-associated herpesvirus phylogeny: evidence for cross-ethnic transmission of viral subtypes. Intervirology. **2006**; 49(3):133–143.

10. Kasolo FC, Monze M, Obel N, Anderson RA, French C, Gompels UA. Sequence analyses of human herpesvirus-8 strains from both African human immunodeficiency virus-negative and -positive childhood endemic Kaposi's sarcoma show a close relationship with strains identified in febrile children and high variation in the K1 glycoprotein. J Gen Virol. **1998**; 79 (Pt 12):3055–3065.

11. Mesri EA, Cesarman E, Boshoff C. Kaposi's sarcoma and its associated herpesvirus. Nat Rev Cancer. **2010**; 10(10):707–719.

12. Liu Z, Fang Q, Zuo J, et al. Global Epidemiology of Human Herpesvirus 8 in Men Who Have Sex with Men: A Systematic Review and Meta-analysis. J Med Virol. **2017**; .

13. Kouri V, Marini A, Doroudi R, et al. Molecular epidemiology of Kaposi's sarcoma herpesvirus (KSHV) in Cuban and German patients with Kaposi's sarcoma (KS) and asymptomatic sexual contacts. Virology. **2005**; 337(2):297–303.

14. Lee Y-M, Hung P-S, Lin C-W. Seroepidemiology and phylogenetic analysis of human herpesvirus type 8 in injection drug users and men who have sex with men in northern Taiwan. J Int Med Res. **2018**; :300060518764747.

15. Lacoste V, Judde JG, Brière J, et al. Molecular epidemiology of human herpesvirus 8 in africa: both B and A5 K1 genotypes, as well as the M and P genotypes of K14.1/K15 loci, are frequent and widespread. Virology. **2000**; 278(1):60–74.

16. Mancuso R, Biffi R, Valli M, et al. HHV8 a subtype is associated with rapidly evolving classic Kaposi's sarcoma. J Med Virol. **2008**; 80(12):2153–2160.

17. Barete S, Calvez V, Mouquet C, et al. Clinical features and contribution of virological findings to the management of Kaposi sarcoma in organ-allograft recipients. Arch Dermatol. **2000**; 136(12):1452–1458.

18. Martin JN, Ganem DE, Osmond DH, Page-Shafer KA, Macrae D, Kedes DH. Sexual transmission and the natural history of human herpesvirus 8 infection. N Engl J Med. **1998**; 338(14):948–954.

19. Lallemand F, Desire N, Rozenbaum W, Nicolas JC, Marechal V. Quantitative analysis of human herpesvirus 8 viral load using a real-time PCR assay. J Clin Microbiol. **2000**; 38(4):1404–1408.

20. Cordiali-Fei P, Trento E, Giovanetti M, et al. Analysis of the ORFK1 hypervariable regions reveal distinct HHV-8 clustering in Kaposi's sarcoma and non-Kaposi's cases. J Exp Clin Cancer Res CR. **2015**; 34:1.

21. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. **2013**; 30(4):772–780.

22. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol. **2010**; 59(3):307–321.

23. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol. **2016**; 33(7):1870–1874.

24. Olp LN, Jeanniard A, Marimo C, West JT, Wood C. Whole-Genome Sequencing of Kaposi's Sarcoma-Associated Herpesvirus from Zambian Kaposi's Sarcoma Biopsy Specimens Reveals Unique Viral Diversity. J Virol. **2015**; 89(24):12299–12308.

25. Sallah N, Palser AL, Watson SJ, et al. Genome-Wide Sequence Analysis of Kaposi Sarcoma-Associated Herpesvirus Shows Diversification Driven by Recombination. J Infect Dis. **2018**; 218(11):1700–1710.

26. Tozetto-Mendoza TR, Ibrahim KY, Tateno AF, et al. Genotypic distribution of HHV-8 in AIDS individuals without and with Kaposi sarcoma: Is genotype B associated with better prognosis of AIDS-KS? Medicine (Baltimore). **2016**; 95(48):e5291.

27. Tornesello ML, Biryahwaho B, Downing R, et al. Human herpesvirus type 8 variants circulating in Europe, Africa and North America in classic, endemic and epidemic Kaposi's sarcoma lesions during pre-AIDS and AIDS era. Virology. **2010**; 398(2):280–289.

28. Yanik EL, Achenbach CJ, Gopal S, et al. Changes in Clinical Context for Kaposi's Sarcoma and Non-Hodgkin Lymphoma Among People With HIV Infection in the United States. J Clin Oncol Off J Am Soc Clin Oncol. **2016**; 34(27):3276–3283.

29. Palich R, Veyri M, Valantin M-A, et al. Recurrence and Occurrence of Kaposi's Sarcoma in Patients Living With Human Immunodeficiency Virus (HIV) and on Antiretroviral Therapy, Despite Suppressed HIV Viremia. Clin Infect Dis Off Publ Infect Dis Soc Am. **2019**; .

Table 1: Epidemiological and medical characteristics of the 57 HIV-infected MSM with KSHV-associated diseases and the eight PrEP users

	HIV-infected			PrEP users
	KS	**MCD**	**PEL**	**Positive KSHV-antibodies**
No. of participants (%)	41/57 (63)	12/57 (19)	4/57 (6)	8/8 (100)
No. of MSM (%)	41/41 (100)	12/12 (100)	4/4 (100)	8/8 (100)
Median age [IQR], years	42 [35-53]	41 [37.75-53.75]	52 [46.75-53.75]	39 [35-42.5]
Country of origin, No. of participants (%)				
France	31/41 (76)	11/12 (92)	4/4 (100)	4/8 (50)
Western Europe (other than France)	0/41 (0)	0/12 (0)	0/4 (0)	1/8 (12.5)
Eastern Europe	0/41 (0)	0/12 (0)	0/4 (0)	2/8 (25)
Mediterranean Basin	5/41 (12)	1/12 (8)	0/4 (0)	0/8 (0)
South America	4/41 (10)	0/12 (0)	0/4 (0)	1/8 (12.5)
Asian	1/41 (2)	0/12 (0)	0/4 (0)	0/8 (0)
HIV infection				
No. (%)	41/41 (100)	12/12 (100)	4/4 (100)	0/8 (0)
Median $CD4^+$ T-cell count [IQR], cells/mm^3	82.5 [34-443.5]	382.5 [95.5-497.25]	219 [126-319.5]	NA
Median HIV viral load [IQR], log_{10} copies/ml	5.36 [2.55-5.73]	3.23 [1.31-5.57]	4.93 [4.58-5.36]	NA
KSHV infection				
Median KSHV viral load [IQR], log_{10} copies/10^6 cells in whole-blood sample	1.94 [1.16-2.85]	4.29 [4.09-4.71]	3.73 [3.70-4.56]	-
Median KSHV viral load [IQR], log_{10} copies/10^6 cells in buccal swab sample	-	-	-	3.71 [3.51-4.01]

HIV: Human immunodeficiency virus; KS: Kaposi's sarcoma; KSHV: Kaposi's sarcoma-associated herpesvirus; MCD: multicentric Castleman disease; MSM: men who have sex with men; NA: not applicable; No.: number; PEL: Primary effusion lymphoma; PrEP: Pre-exposure prophylaxis

Figure 1: Amino-acid maximum-likelihood phylogenetic tree constructed with PhyML (3.0) of ORF-K1 patients and references sequences for subtype A (A1, A2, A3, A4 and A5), B (B1, B2 and B3), C (C1, C2 and C3) and F (AAX55469_F and AAG01610_F) available from the NCBI database. *The patients' sequences are shown in blue, the reference sequences are shown in black and the 3 clusters found are shown in red on the tree. Nodes presenting a branch support > 70% (bootstrap analysis with 1000 replicates) are indicated by an asterisk.*

GenBank reference sequence accession numbers: A1: ACS74793, ACS74803 and AAD30529; A2: AAO86800 and AAD26415; A3: AAB71616; A4: AAD30530; A5: ACS74801, AAG01621 and AAG01597; B1: AAG01622, AAG01601 and AAD30531; B2: AAK72680, AAD26369 and AAG01617; B3:AAK72674; C1: AAD26377 and AAD30532; C2: ABD52266; C3: AAD30533 and ACY00482; F: AAX55469 and AAG01610.

GenBank new sequence accession numbers: P149_K1_PEL: MK840448; P076_K1_PEL: MK840449; P075_K1_MCD: MK840450; P072_K1_MCD: MK840451; P042_K1_KS: MK840452; P100_K1_PEL: 840453; P030_K1_KS: MK840454; P062_K1_MCD: MK840455; P044_K1_PEL: MK840456; P073_K1_KS: MK840457; P020_K1_KS: MK840458; PrEP_002_K1: MK840459; P095_K1_KS: MK840460; P147_K1_MCD: MK840461; P006_K1_KS: MK840462; P077_K1_KS: MK840463; P055_K1_KS: MK840464; P007_K1_KS: MK840465; P132_K1_KS: MK840466; P012_K1_KS: MK840467; P078_K1_MCD: MK840468; P107_K1_MCD: MK840469; P151_K1_KS: MK840470; P133_K1_MCD: MK840471; P106_K1_KS: MK840472; P152_K1_MCD: MK840473; PrEP_001_K1: MK840474; P003_K1_KS: MK840475; P009_K1_KS: MK840476.

Figure 2: Comparison, at the time of Kaposi's sarcoma diagnosis, of age (A), KSHV-DNA viral load (B) and immunovirological status (C and D) between patients with KSHV subtype A and patients with KSHV subtype C. **Horizontal lines**: median; **cross:** mean; **boxes:** quartiles 1 and 3; **whiskers:** 95% confidence intervals; **Mann-Whitney *U* test** (*p*).

Figure 3: Multiple alignment, with Mafft, of partial ORF-K1 amino-acid (aa) sequences (from aa 37 to aa 228) identified as subtype F. *The sequence from Uganda (AAX55469_F; shown entirely in color) is compared with sequence described in 2000 from a French MSM HIV-infected with PEL (AAG01610_F), and with our five sequences isolated from MSM living in France with or without KSHV-associated diseases (P076, P075, P030, P035 and PrEP004) and the KSHV sequence of a woman from Congo with multicentric Castleman disease (P072). Mismatches are highlighted in various colors, according to the amino-acid concerned, and the insertions are indicated by dashes in the other sequences. Hypervariable regions 1 and 2 are located between aa 54 and 93 and between aa 191 and 228, respectively.*

GenBank new sequence accession numbers: P072_K1_MCD: MK840451; P076_K1_PEL: MK840449; P075_K1_MCD: MK840450; P030_K1_KS: MK840454; P035_VR1_KS: MK840478; PrEP_004_VR1: MK840477

Figure 4: Nucleotides maximum-likelihood phylogenetic tree constructed with PhyML (3.0) of KSHV-WG patients newly sequenced and reference sequences for subtypes A (A, A3 and A5), B (B1, B3 and B4) and variant C3, available from the NCBI database. *Sanger sequencing of ORF-K1 assigned the sequences from P044 and P100 to variants A4 and A1, respectively, that from P133 to variant C3, those of P030, P075 and P076 to "Caucasian subtype F" and that of P072 to "African subtype F". The patients' sequences are shown in blue and the reference sequences are shown in black. Nodes presenting a branch support > 70% (bootstrap analysis with 1000 replicates) are indicated by an asterisk.*

GenBank reference sequence accession numbers: A: AP017458; A3: HQ404500 and KX189629; A5: JQ619843; B1: KT271465 and KT271458; B3: KT271460; B4: KT271461 and KT271462; C3: NC_009333_C3_GK18, KF588566 and GQ994935.

GenBank new sequence accession numbers: BC3: MK876731; P030_KS: MK876732; P044_PEL: MK876733; P072_MCD: MK876734; P075_MCD: MK876735; P076_PEL: MK876736; P100_PEL: MK876737; P133_MCD: MK876738

Supplementary Table 1: ORF-K1 reference sequences available on NCBI and used to design DNA-probes for KSHV-enrichment before high throughput sequencing.

Reference number on NCBI	ORFK1 subtype or genotype variant
AF133038	A1
AF130305	A2
U75698	A2
U86667	A3
AF133039	A4
FJ884625	A4
AF178823	A5
AY329023	A5
FJ884624	A5
AF133040	B1
AY042947	B2
AF178818	B2
AY042941	B3
AF178825	B3
DQ309754	B4
AF133041	C1
AF130267	C1
DQ394048	C2
DQ394058	C2
AF133042	C3
GQ994935	C3
AF130268	C3
AF133043	D1
AF133044	D2

AF220292	E
AF220293	E
AY329028	E2
AY953882	F

NCBI: National Center for Biotechnology Information; ORF-K1: open reading frame K1

Supplementary Table 2: Experimental conditions for the amplification of the new KSHV genotype variant by real-time PCR

DNA: Desoxyribonucleic acid; KSHV: Kaposi's sarcoma-associated herpesvirus; Min: minute; PCR: polymerase chain reaction; Sec: second

	Volume (µl) 1 reaction	Final concentration	PCR program			
TaqMan PCR Master Mix 2x	10	1X	**Hold**	2 min	50°C	1 cycle
Primer Forward (10µM) 5'-TCCTGGTATTGCAACGGAAC-3'	1.8	900nM	**Polymerase activation**	10 min	95°C	1 cycle
Primer Reverse (10µM) 5'-TCCAAATGCTGTGTGAAGGC-3'	1.8	900nM	**Denaturation**	15 sec	95°C	40 cycles
Probe (10µM) 5'-GCAGCGATCAGTACCTGTTT-3'	0.4	200nM	**Annealing**	1 min	60°C	
Extracted DNA	6	-	**Hold**	∞	4°C	
Total	20	-				

Supplementary Table 3: KSHV-typing results by pathologies and for each patient with KSHV-associated diseases or PrEP users

	KAPOSI SARCOMA				MULTICENTRIC CASTLEMAN DISEASE					PRIMARY EFFUSION LYMPHOMA						PrEP USERS		
No of KSHV-typing (%)	19/41 (46%)				11/12 (92)					4/4 (100)						5/8 (63)		
KSHV-typing	**No (%)**	**Cut and/ or oral lesions**	**Visceral lesions**	**Native country**	**No (%)**	**Simultaneous or previously KS-diagnosis** *No*	*Cut and /or oral lesions*	*Visceral lesions*	**Native country**	**No (%)**	**Simultaneous or previously KS-diagnosis** *No*	*Cut and/ or oral lesions*	*Visceral lesions*	**PEL localization**	**Native country**	**No (%)**	**Native country**	**Total**
A subtype																		
A	-	-	-	-	1/11 (9)	1/1	0	1	F	-	-	-	-	-	-	-	-	
A1	1/19 (5)	1	0	F	2/11 (18)	0/2	-	-	2 F	1/4 (25)	0/1	-	-	ascitis	F	-	-	
A3	1/19 (5)	1	0	F	-	-	-	-	-	-	-	-	-	-	-	-	-	
A4	3/19 (16)	0	3	2 M. Ba 1 F	1/11 (9)	1/1	1	0	M. Ba	1/4 (25)	1/1	0	1	pleural	F	2/5 (40)	1 F 1 S. Am	
Total A subtype	**5/19 (26)**	2	3	-	**4/11 (36)**	2/4	1	1	-	**2/4 (50)**	1/2	0	1	-	-	**2/5 (40)**	**-**	**13/39 (33)**
B subtype																		
B1	**0/19 (0)**	-	-	-	**0/11 (0)**	-	-	-	-	**1/4 (25)**	MD	MD	MD	MD	F	**-**	**-**	**1/39 (3)**
C subtype																		
C1	1/19 (5)	0	1	M. Ba	1/11 (9)	0/1	-	-	F	-	-	-	-	-	-	-	-	
C1-C2	1/19 (5)	0	1	F	-	-	-	-	-	-	-	-	-	-	-	-	-	
C3	10/19 (53)	8	2	7 F 1 M. Ba 1 S. Am 1 As	5/11 (46)	4/5	3	1	5F	-	-	-	-	-	-	2/5 (40)	1 F 1 MD	
Total C subtype	**12/19 (63)**	8	4	-	**6/11 (55)**	4/5	3	1	-	**0/4 (0)**	-	-	-	-	-	**2/5 (40)**	**-**	**20/39 (51)**
"F subtype"	**2/19 (11)**	0	2	1 F 1 S. Am	**1/11 (9)**	1/1	0	1	F	**1/4 (25)**	1/1	1	0	pleural	F	1/5 (20)	F	**5/39 (13)**

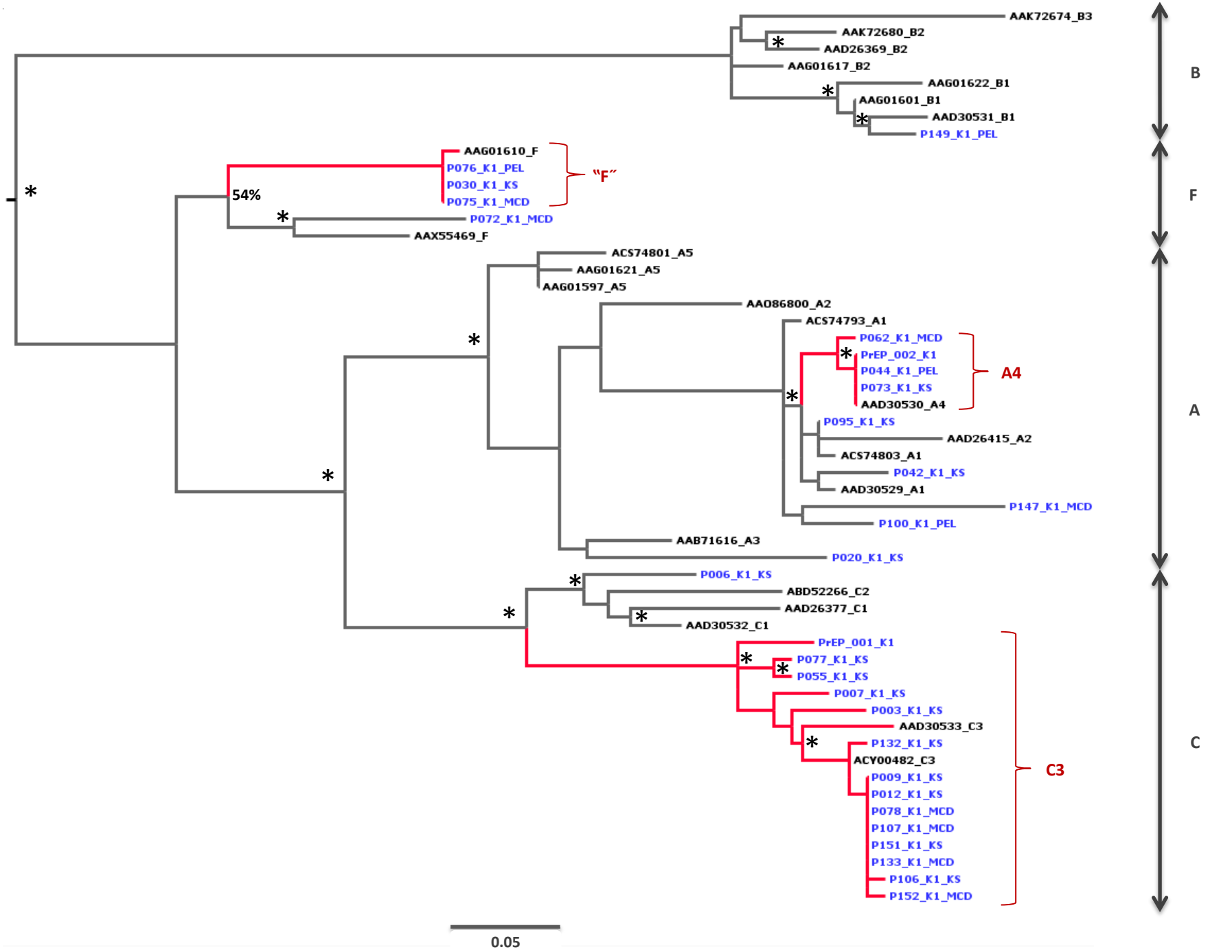
As: Asian; Cut: cutaneous; F: France; KS: Kaposi sarcoma; M. Ba: Mediterranean basin; MD: Missed data; No: Number; PEL: Primary effusion lymphoma; S. Am: South America; Sw: Swiss

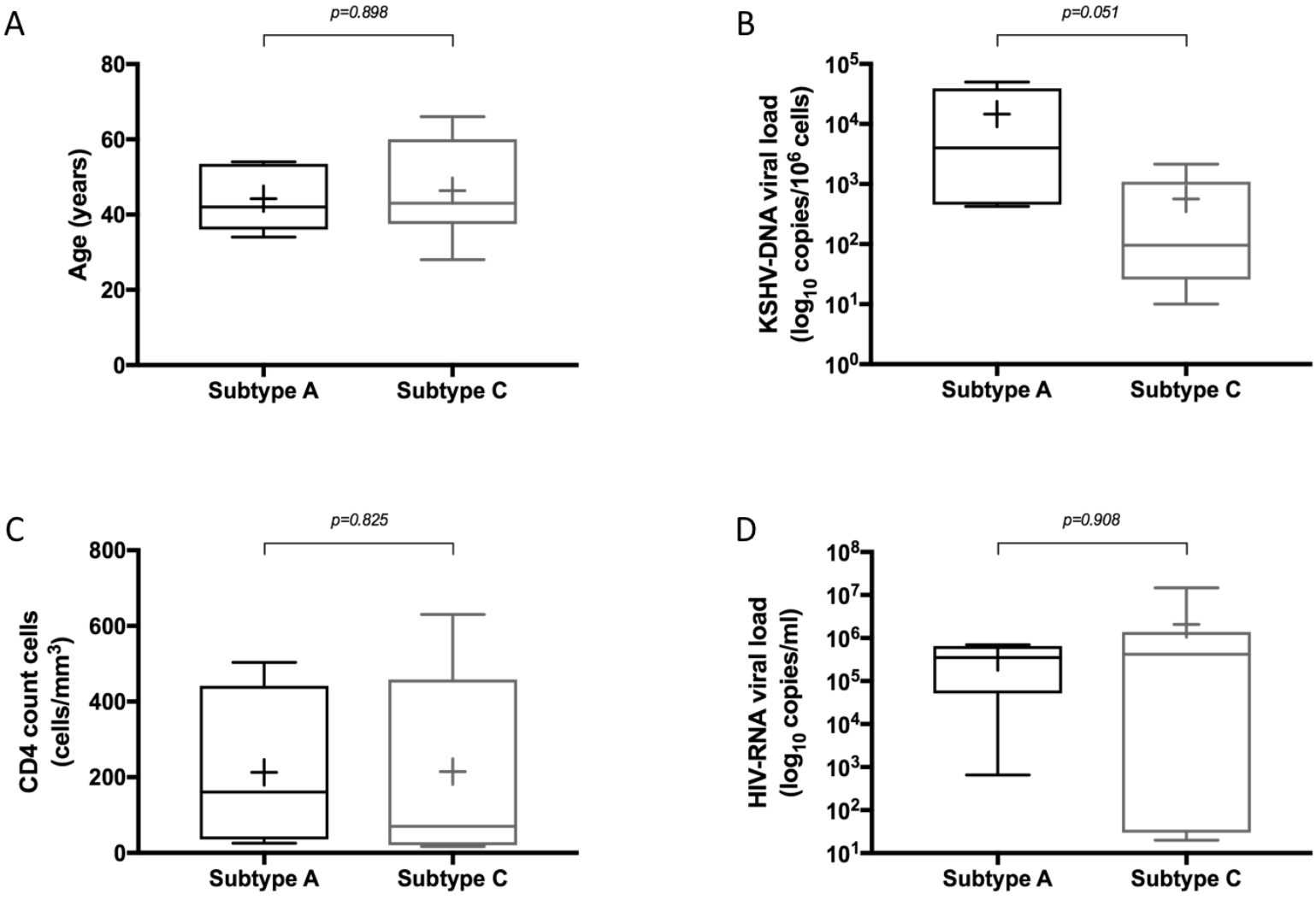
Supplementary Table 4

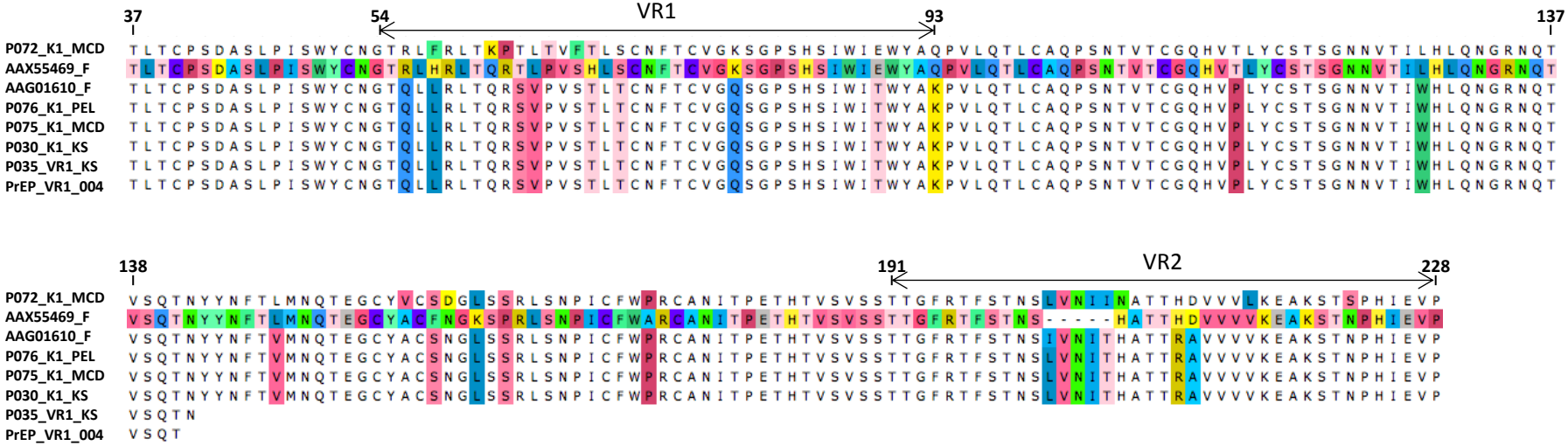
New genotype variant F characteristics

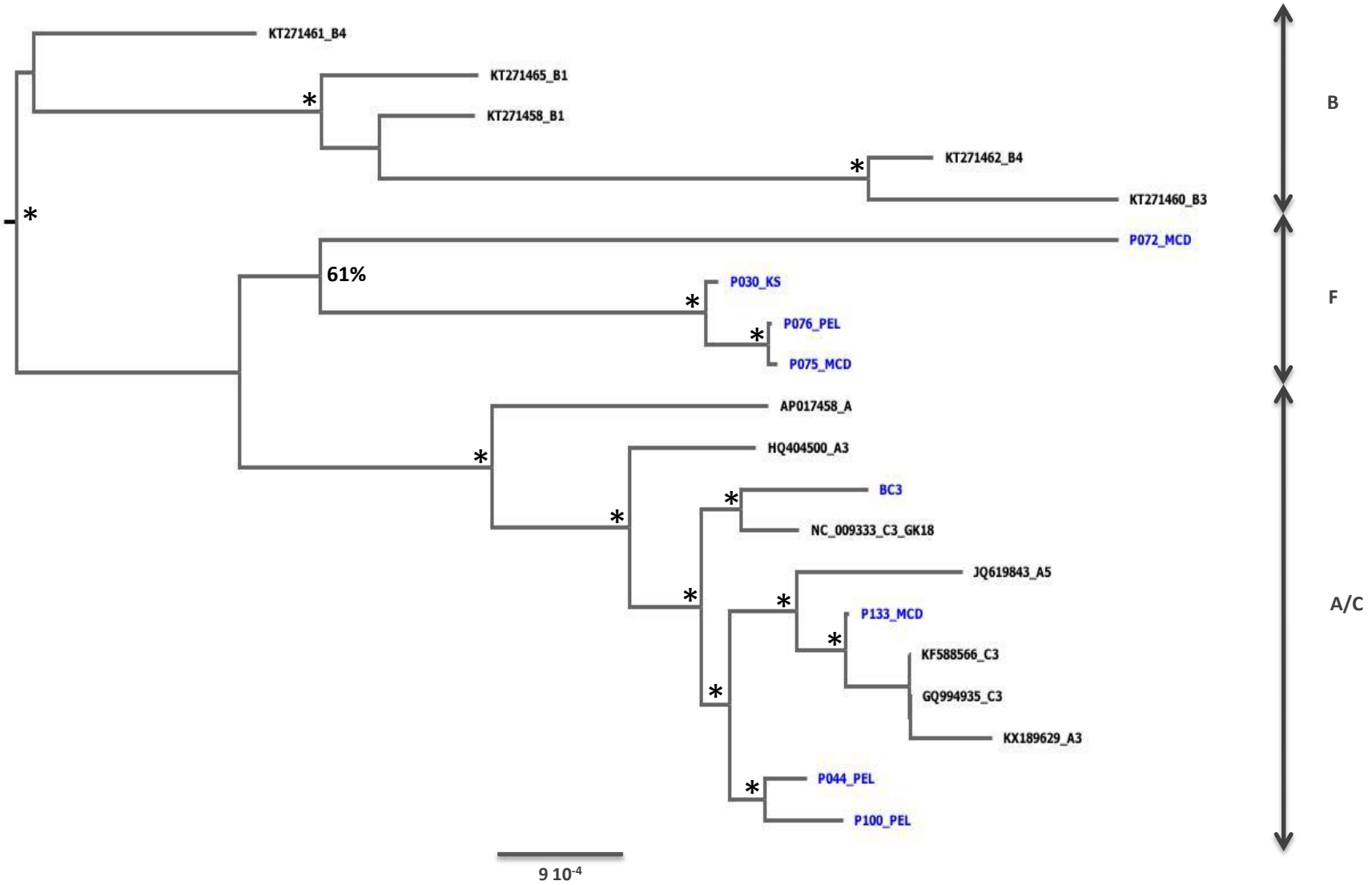
	P030	P035	P075	P076	PrEP004
Age, years old	34	33	50	52	47
Native country	France	Peru	France	France	France
HIV-infection					NA
CD4 count, cells /mm^3	85	80	45	162	
HIV-RNA viral load, log$_{10}$ copies/ml	5.43	1.69	5.69	5.07	
KSHV-infection					
Clinical presentation	Visceral KS: Cutaneous lesions Pulmonary involvement	Visceral KS: Cutaneous lesions Pulmonary involvement Ganglionic involvement	MCD Visceral KS: Digestive involvement Ganglionic involvement	PEL: Pleural involvement KS: Cutaneous lesions	Asymptomatic
KSHV-DNA viral load, log$_{10}$ copies/10^6 cells					
- ***Whole blood sample***	5.07	2.83	4.70	4.83	
- ***Respiratory sample***	4.91	4.97			
- ***Oral swab***					5.61
KSHV-typing					
Gene sequenced	ORF-K1	ORF-K1 (VR1)	ORF-K1	ORF-K1	ORF-K1 (VR1)
Subtype	F	F	F	F	F
KSVH whole genome sequencing					
Genome size without repeat region (bp)	137 139	~ 136 000	135 278	137 438	NR
Percentage of coverage	99.36	80.6	98.16	99.01	NR
Depth of coverage (fold)	52	7	108	292	NR

AIDS: acquired immune deficiency syndrome; bp: base pair; HIV: Human immunodeficiency virus; KSHV: Kaposi's sarcoma-associated herpesvirus; NA: not applicable; NR: not realized; ORF-K1: open reading frame K1; VR1: variable region 1


AAK72674_B3
AAK72680_B2
AAD26369_B2
AAG01617_B2
AAG01622_B1
AAG01601_B1
AAD30531_B1
P149_K1_PEL
B
AAG01610_F
P076_K1_PEL
P030_K1_KS
P075_K1_MCD
"F"
54%
P072_K1_MCD
AAX55469_F
F
ACS74801_A5
AAG01621_A5
AAG01597_A5
AAO86800_A2
ACS74793_A1
P062_K1_MCD
PrEP_002_K1
P044_K1_PEL
P073_K1_KS
AAD30530_A4
A4
P095_K1_KS
AAD26415_A2
ACS74803_A1
P042_K1_KS
AAD30529_A1
P147_K1_MCD
P100_K1_PEL
AAB71616_A3
P020_K1_KS
A
P006_K1_KS
ABD52266_C2
AAD26377_C1
AAD30532_C1
PrEP_001_K1
P077_K1_KS
P055_K1_KS
P007_K1_KS
P003_K1_KS
AAD30533_C3
P132_K1_KS
ACY00482_C3
P009_K1_KS
P012_K1_KS
P078_K1_MCD
P107_K1_MCD
P151_K1_KS
P133_K1_MCD
P106_K1_KS
P152_K1_MCD
C3
C
0.05


A
p=0.898
Age (years)
80
60
40
20
0
Subtype A
Subtype C
B
p=0.051
KSHV-DNA viral load
(log10 copies/10^6 cells)
10^5
10^4
10^3
10^2
10^1
10^0
Subtype A
Subtype C
C
p=0.825
CD4 count cells
(cells/mm^3)
800
600
400
200
0
Subtype A
Subtype C
D
p=0.908
HIV-RNA viral load
(log10 copies/ml)
10^8
10^7
10^6
10^5
10^4
10^3
10^2
10^1
Subtype A
Subtype C


37
54
VR1
93
137
P072_K1_MCD TLTCPSDASLPISWYCNGTRLFRLTKPTLTVFTLSCNFTCVGKSGPSHSIWIEWYAQPVLQTLCAQPSNTVTCGQHVTLYCSTSGNNVTILHLQNGRNQT
AAX55469_F TLTCPSDASLPISWYCNGTRLHRLTQRTLPVSHLSCNFTCVGKSGPSHSIWIEWYAQPVLQTLCAQPSNTVTCGQHVTLYCSTSGNNVTILHLQNGRNQT
AAG01610_F TLTCPSDASLPISWYCNGTQLLRLTQRSVPVSTLTCNFTCVGQSGPSHSIWITWYAKPVLQTLCAQPSNTVTCGQHVPLYCSTSGNNVTIWHLQNGRNQT
P076_K1_PEL TLTCPSDASLPISWYCNGTQLLRLTQRSVPVSTLTCNFTCVGQSGPSHSIWITWYAKPVLQTLCAQPSNTVTCGQHVPLYCSTSGNNVTIWHLQNGRNQT
P075_K1_MCD TLTCPSDASLPISWYCNGTQLLRLTQRSVPVSTLTCNFTCVGQSGPSHSIWITWYAKPVLQTLCAQPSNTVTCGQHVPLYCSTSGNNVTIWHLQNGRNQT
P030_K1_KS TLTCPSDASLPISWYCNGTQLLRLTQRSVPVSTLTCNFTCVGQSGPSHSIWITWYAKPVLQTLCAQPSNTVTCGQHVPLYCSTSGNNVTIWHLQNGRNQT
P035_VR1_KS TLTCPSDASLPISWYCNGTQLLRLTQRSVPVSTLTCNFTCVGQSGPSHSIWITWYAKPVLQTLCAQPSNTVTCGQHVPLYCSTSGNNVTIWHLQNGRNQT
PrEP_VR1_004 TLTCPSDASLPISWYCNGTQLLRLTQRSVPVSTLTCNFTCVGQSGPSHSIWITWYAKPVLQTLCAQPSNTVTCGQHVPLYCSTSGNNVTIWHLQNGRNQT
138
191
VR2
228
P072_K1_MCD VSQTNYYNFTLMNQTEGCYVCSDGLSSRLSNPICFWPRCANITPETHTVSVSSTTGFRTFSTNSLVNIINATTHDVVVLKEAKSTSPHIEVP
AAX55469_F VSQTNYYNFTLMNQTEGCYACFNGKSPRLSNPICFWARCANITPETHTVSVSSTTGFRTFSTNS-----HATTHDVVVVKEAKSTNPHIEVP
AAG01610_F VSQTNYYNFTVMNQTEGCYACSNGLSSRLSNPICFWPRCANITPETHTVSVSSTTGFRTFSTNSIVNITHATTRAVVVVKEAKSTNPHIEVP
P076_K1_PEL VSQTNYYNFTVMNQTEGCYACSNGLSSRLSNPICFWPRCANITPETHTVSVSSTTGFRTFSTNSLVNITHATTRAVVVVKEAKSTNPHIEVP
P075_K1_MCD VSQTNYYNFTVMNQTEGCYACSNGLSSRLSNPICFWPRCANITPETHTVSVSSTTGFRTFSTNSLVNITHATTRAVVVVKEAKSTNPHIEVP
P030_K1_KS VSQTNYYNFTVMNQTEGCYACSNGLSSRLSNPICFWPRCANITPETHTVSVSSTTGFRTFSTNSLVNITHATTRAVVVVKEAKSTNPHIEVP
P035_VR1_KS VSQTN
PrEP_VR1_004 VSQT


KT271461_B4
KT271465_B1
KT271458_B1
KT271462_B4
KT271460_B3
P072_MCD
61%
P030_KS
P076_PEL
P075_MCD
AP017458_A
HQ404500_A3
BC3
NC_009333_C3_GK18
JQ619843_A5
P133_MCD
KF588566_C3
GQ994935_C3
KX189629_A3
P044_PEL
P100_PEL
9 10-4
B
F
A/C