Infection, Genetics and Evolution 66 (2018) 43–47

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

Research paper

Genome-scale analysis of evolutionary rate and selection in a fast-expanding Spanish cluster of HIV-1 subtype F1

Juan Á. Patiño-Galindo[a,b,1], Francisco Domínguez[c], María T. Cuevas[c], Elena Delgado[c], Mónica Sánchez[c], Lucía Pérez-Álvarez[c], Michael M. Thomson[c], Rafael Sanjuán[d,e], Fernando González-Candelas[a,b,d,e], José M. Cuevas[d,e,*]

[a] Joint Research Unit "Infection and Public Health" FISABIO-Universitat de València, València, Spain
[b] CIBER in Epidemiology and Public Health, Madrid, Spain
[c] HIV Biology and Variability Unit, Centro Nacional de Microbiología, Instituto de Salud Carlos III, Majadahonda, Madrid, Spain
[d] Institute for Integrative Systems Biology (I2SysBio), Universitat de València, València, Spain.
[e] Department of Genetics, Universitat de València, València, Spain

ARTICLE INFO

Keywords:
HIV-1
Subtype F1
Transmission cluster
Men who have sex with men
Selection

ABSTRACT

This work is aimed at assessing the presence of positive selection and/or shifts of the evolutionary rate in a fast-expanding HIV-1 subtype F1 transmission cluster affecting men who have sex with men in Spain. We applied Bayesian coalescent phylogenetics and selection analyses to 23 full-coding region sequences from patients belonging to that cluster, along with other 19 F1 epidemiologically-unrelated sequences. A shift in the overall evolutionary rate of the virus, explained by positively selected sites in the cluster, was detected. We also found one substitution in Nef (H89F) that was specific to the cluster and experienced positive selection. These results suggest that fast transmission could have been facilitated by some inherent genetic properties of this HIV-1 variant.

1. Introduction

In Western Europe, the HIV-1 epidemic is dominated by subtype B, especially among men who have sex with men (MSM) (Abecasis et al., 2013). However, a large HIV-1 subtype F1 transmission cluster, affecting > 100 MSM in Spain and other Western European countries, was recently reported (Delgado et al., 2015; Thomson et al., 2012). This cluster was identified in an HIV-1 molecular epidemiological surveillance study in the region of Galicia, Northwest Spain, with the first cases diagnosed in 2009 and subsequent rapid expansion resulting in an increase in the prevalence of subtype F1 infections among new HIV-1 diagnoses from 1.1% in 2008 to 8.3% in 2009 and 25.5% in 2010. Viruses belonging to this cluster were also identified in several other Spanish regions and in four Western European countries (Delgado et al., 2015; Thomson et al., 2012). This subtype is rare in Western Europe, displaying a prevalence of < 2% (Abecasis et al., 2013). The transmission cluster presented no major resistance-associated mutations, but was characterized by a rapid expansion among MSM in Spain, which could be accounted for by the epidemiological scenario of HIV-1 in Europe, which is characterized by an increase in risk behavior among MSM (Bezemer et al., 2008; ECDC, 2013). Nonetheless, it has also been suggested that its efficient transmission might be linked to some intrinsic genetic properties of the viral lineage, given its unusual large size (Delgado et al., 2015; Thomson et al., 2012) and the fact that patients included in this outbreak presented significantly higher viral loads and poorer response to antiretroviral treatment than subtype B variant (Cid-Silva et al., 2018; Pernas et al., 2014).

The hypothesis of positive selection acting on HIV in fast-expanding transmission clusters contrasts with previous works suggesting that adaptive selection is weaker during early infection, when most infections occur due to the lack of awareness of HIV serological status (Maljkovic Berry et al., 2007). Furthermore, mutations that are adaptive in one individual are possibly maladaptive in other individuals and, consequently, HIV transmitted to a new host undergoes reversions of mutations that adaptively occurred in the donor. Altogether, these facts would slow down the evolutionary rate of a viral lineage involved in a fast-expanding epidemic (Maljkovic Berry et al., 2007; Lythgoe and Fraser, 2012). However, fast transmissions have also been associated

* Corresponding author at: Institute for Integrative Systems Biology (I2SysBio), Universitat de València, Calle Catedratico José Beltrán, 46980 Paterna, València, Spain.
E-mail address: cuevast@uv.es (J.M. Cuevas).
[1] Present address: Department of Systems Biology, Columbia University, New York, NY, USA

https://doi.org/10.1016/j.meegid.2018.09.008
Received 29 June 2018; Received in revised form 6 September 2018; Accepted 12 September 2018
Available online 13 September 2018

with higher evolutionary rates (Pybus and Rambaut, 2009; Salemi et al., 1999). These studies argued that, if within-host rates slow down during infection, faster transmissions should result in higher long-term evolutionary rates. Also, faster transmission rates could lead mutations to be more likely to occur, which may increase the evolutionary rate.

For this study, the complete coding sequences of 24 publicly available Spanish isolates from a fast-expanding HIV-1 subtype F1 transmission cluster were retrieved. In addition, sequences of 25 HIV-1 subtype F1, worldwide-distributed isolates were included in the analysis. First, we evaluated the impact of recombination in our samples in order to remove its effect in subsequent analyses. Then, we performed Bayesian coalescent analyses to compare genomic evolutionary rates of the Spanish transmission cluster and the epidemiologically unrelated sequences. Finally, we assessed the effect of selection on the evolution of the transmission cluster.

2. Materials and methods

2.1. Dataset

Twenty four full coding sequences (CDS) from different patients belonging to the Spanish transmission cluster (Delgado et al., 2015), along with 25 HIV-1 subtype F1 sequences from epidemiologically unrelated patients, were retrieved from the Los Alamos National Laboratory (LANL) HIV Sequence Database (http://www.hiv.lanl.gov). Accession numbers and information on these sequences are provided in Supplementary Table 1. The correct subtype assignment was corroborated with the COMET HIV-1 subtyping tool (http://comet.retrovirology.lu). A recombination analysis was performed with RDP4, using five different methods of recombination detection: RDP, Geneconv, Bootscan, Maxchi and Chimaera (Martin et al., 2015). Only those breakpoints detected by at least two of the five methods implemented in RDP4 were accepted. Consequently, recombinant sequences, as well as those lacking information on the collection date, were excluded from subsequent analyses. An alignment consisting of the concatenated non-overlapping regions from all HIV-1 genes (considering the HIV-1 reference sequence HXB2, accession number K03455) was obtained using MAFFT v7 (Katoh and Standley, 2013), and regions with poor homology ("gappy" sites) were trimmed using TrimAl (Capella-Gutiérrez et al., 2009). The final alignment is available upon request.

2.2. Inference and comparison of the genomic evolutionary rates

The selected set of HIV-1 F1 subtype sequences was subjected to a Bayesian coalescent analysis in order to estimate the genomic evolutionary rate of the transmission cluster and to compare it with that of the HIV-1 lineages not included in the Spanish cluster. The presence of molecular clock signal was assessed with a root-to-tip divergence *versus* sample time correlation analysis, performed with Path-O-Gen (now renamed as TempEst; Rambaut et al., 2016), using as input a maximum-likelihood (ML) phylogenetic tree obtained with PhyML (Guindon et al., 2010) under the GTR + Γ (4 CAT) model. The Bayesian coalescent analysis was performed with BEAST v1.8.2 (Drummond et al., 2012), under a nonparametric demographic model (Bayesian Skyline Plot), combined with the GTR + Γ (4 CAT) nucleotide model and a random local molecular clock model. This clock model proposes and compares a series of alternative local molecular clocks, which can arise on any branch of the phylogeny and then extend along adjacent lineages (Drummond and Suchard, 2010), and has previously been reported to detect evolutionary rate shifts in specific lineages of a phylogeny (Fourment and Holmes, 2014). The comparison of the evolutionary rates estimated from lineages belonging to the transmission cluster and those falling outside was performed by means of a randomization test: rates inferred at branches from both groups (transmission cluster and outside) are randomly sampled with replacement 1000 times, and compared. A *P* value is obtained by counting the number of comparisons where the rate from the transmission cluster is higher than that from branches outside it (Abecasis et al., 2009; Patiño-Galindo and González-Candelas, 2017).

2.3. Positive selection analysis

The presence of positive selection in the transmission cluster was tested with MEDS, as implemented in HyPhy, a ML method used to detect independent sites under directional selection (adaptive evolution in which mutations to a particular amino acid are selected) (Murrell et al., 2012). A ML phylogenetic tree previously obtained with PhyML (GTR + Γ_4) for the molecular clock signal analysis was used as input, specifying *a priori* that all lineages included in the transmission cluster were susceptible of being under directional selection (foreground branches). Only positively selected sites meeting two conditions were considered. First, in order to minimize the presence of false positives (*i.e.*, sites evolving under neutral evolution or sites in which a deleterious mutation had occurred but had not been removed from the viral population at the time of sampling), those sites in which amino acid variability was caused only by singletons (affecting only one sequence in the phylogeny) and/or only in one step of change along the phylogeny were excluded from the list of positively selected sites. Second, given that the goal of these analyses was to find evidence of selection specific to the transmission cluster, we only considered those positively selected sites in which the consensus amino acid of the transmission cluster differed from that of the other F1 sequences, as performed with VESPA; (Korber and Myers, 1992). Additionally, we also required that the amino acid distribution differed significantly between both groups of sequences according to Fisher's exact tests, with *P* values corrected by using the false discovery rate (Benjamini and Hochberg, 1995).

Positively selected sites were mapped onto the HXB2 reference genome. A search was performed in LANL database to identify those sites located in human epitopes (antibody, $CD8^+$ and/or $CD4^+$ T cells), as well as those sites where the transmission cluster presented a consensus amino acid associated with $CD8^+$ and/or $CD4^+$ T cell immune escape.

3. Results

Recombination analyses performed with RDP4 found evidence of recombination in 3 of the 49 full-genome sequences, including one sequence from the transmission cluster (sequence name VA0043). These sequences were removed from the dataset, as well as those with no collection date information. Consequently, the final HIV-1 subtype F1 dataset consisted of 23 CDS sequences from the Spanish transmission cluster and 19 epidemiologically unrelated HIV-1 F1 sequences (Supplementary Table 1). The resulting alignment of concatenated HIV-1 genes, in their correct coding frame, was 7962 nt long (> 90% of the full HIV-1 CDS).

The dataset of 42 HIV-1 F1 genomic sequences displayed sufficient clock-like signal as to proceed with the Bayesian coalescent analysis ($R^2 = 0.72$). The resulting dated Bayesian phylogenetic tree displayed evidence for the existence of an acceleration in the evolutionary rate in the transmission cluster, compared with the rest of lineages in the phylogeny (lineages outside the transmission cluster: median = 2.00×10^{-3} substitutions per site and year-s/s/y-, 95%HPD = 1.60×10^{-3}–2.30×10^{-3} s/s/y; lineages within the cluster: median = 3.20×10^{-3} s/s/y, 95%HPD = 2.10×10^{-3}–3.90×10^{-3} s/s/y; Fig. 1a), although the difference was not statistically significant ($P = .09$).

MEDS detected 89 sites significantly associated with directional selection, but only 19 sites met the inclusion criteria: 1 in *pol* (0.1% of the amino acid sequence analyzed), 2 in *vif* (1.3%), 2 in *vpu* (4.0%), 11 in *env* (1.4%) and 3 in *nef* (1.5%). Thus, most sites detected to be under selection were located in the *env* gene. Most positively selected sites

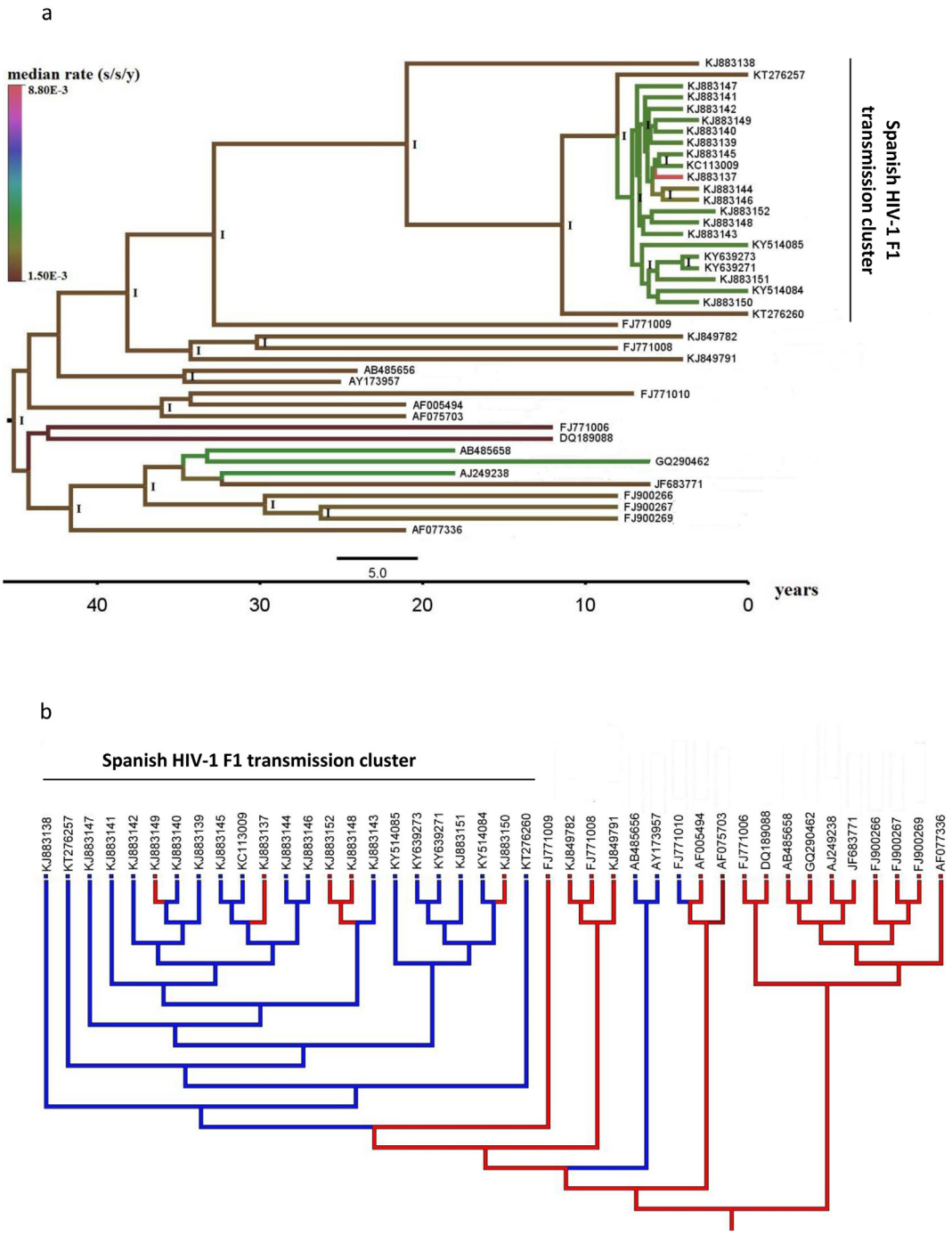

Fig. 1. a) Dated phylogenetic tree obtained with BEAST (time measured in years) for the HIV-1 subtype F1 analyzed using the random local clock model for the full CDS. "I" labels on nodes represent clades with posterior probability = 1.0, and branches are colored depending on their median evolutionary rate. b) Dendrogram displaying the evolutionary history of the *nef* H89F mutation (blue color) in the transmission cluster. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

were included in $CD8^+$ or $CD4^+$ T cell epitopes (Table 1).

In order to assess whether positive selection could explain the acceleration of the evolutionary rate in the transmission cluster, we repeated the BEAST analyses using an unlinked molecular clock model, in which two different partitions (all positively selected sites detected and the rest of the alignment) were allowed to have different molecular clock models. We found a significant acceleration in the evolutionary rate of the transmission cluster occurring in positively selected sites (lineages outside the transmission cluster: median = 4.90×10^{-3} s/s/y, 95%HPD = 3.90×10^{-3}–6.0×10^{-3} s/s/y; lineages within the cluster: median = 12.70×10^{-3} s/s/y, 95%HPD = 9.60×10^{-3} –16.20×10^{-3} s/s/y; $P < .001$. Supplementary Fig. 1A) but not in the rest of the alignment (lineages outside the transmission cluster: median = 1.80×10^{-3} s/s/y, 95%HPD = 1.50×10^{-3}–2.20×10^{-3} s/s/y; lineages within the cluster: median = 1.70×10^{-3} s/s/y, 95% HPD = 1.40×10^{-3}–2.00×10^{-3} s/s/y; $P = .18$; Supplementary Fig. 1B). In order to validate this finding, we repeated the MEDS positive selection analyses choosing the unrelated F1 sequences as foreground lineages, and performed another BEAST analysis partitioning the selected sites found for this subset (233 codons) and the rest of the genome. No acceleration in the evolutionary rate was found.

Regarding positively selected sites, Nef H89F substitution (Table 1),

Table 1
Positively selected positions in the transmission cluster that met the inclusion criteria. Signature amino acid (AA) inside and outside the cluster, location in human antibody (AB), and $CD8^+$ and/or $CD4^+$ epitopes is specified. For each position, codon is indicated at the corresponding gene (reference sequence HXB2, Genbank accession number K03455).

Position	Signature AA (cluster)	Signature AA (no cluster)	AB epitope	CD8+ epitope (A-list, best defined CTLs)	CD4+ epitope
pol 95	T	P	no	yes	no
vif 50	K	R	no	yes	no
vif 127	Y	H	no	no	no
vpu 14	L	V	no	no	yes
vpu 19	V	A	no	no	yes
env 94	N	D	no	no	yes
env 173	S	H (< 50%)	yes	no	yes
env 271	V	I	no	no	yes
env 297	S	T	no	no	yes
env 389	K	G	no	no	yes
env 442	H	N	yes	no	yes
env 467	I	T	yes	no	yes
env 565	M	L	yes	yes	yes
env 621	Q	E	no	no	yes
env 750	S	N	no	no	no
env 797	W	G	no	yes	no
nef 89*	F	H	no	yes	yes
nef 120	F	Y	no	yes	yes
nef 176	A	E	no	no	yes

* Mutation associated with escape to immune response.

which was located in $CD8^+$ and $CD4^+$ T cell human epitopes (Nef positions 84–92, wild type sequence AVDLSHFLK), has been previously associated with a decreased $CD8^+$ T cell response for different human HLA alleles (Fukada et al., 2002; Hoof et al., 2010) (Table 1; Fig. 1b). It is noteworthy that this epitope has been studied in subtypes A, B, C, D and CRF01_AE but not in subtype F1. However, in our dataset the consensus amino acid of all the positions was wild type with the exception of Nef 89.

4. Discussion

This work was aimed at assessing the evolutionary dynamics of an unusually large, fast-expanding HIV-1 F1 subtype transmission cluster, which affected > 100 MSM in Spain, and has been associated with higher plasma viral loads and poorer virological response to antiretroviral treatment than other HIV-1 variants (Cid-Silva et al., 2018; Pernas et al., 2014). Despite the relatively small number of sequences analyzed here, a genome-scale analysis has allowed to detect an acceleration of the evolutionary rate of this HIV-1 F1 cluster associated with fast transmissions, which could be explained by the effect of positive selection in this cluster. This is in agreement with previous studies that associated the evolutionary rates with the speed at which transmissions occur in an epidemic (Salemi et al., 1999; Pybus and Rambaut, 2009). At this point, our results must be taken cautiously because it is worth recalling that the acceleration detected in our study was only significant for the partition of positively selected sites, but not for the whole genome analysis. Anyway, the analysis of this specific partition is a way of confirming that positively selected sites evolve faster, which is true by definition.

The presence of positive selection in transmission clusters has been rarely reported. By performing diversifying selection analyses on three different HIV-1 CRF01_AE clusters affecting MSM occurring in China, Peng et al. (2015) detected that positive selection was present in approximately 5% of the amino acid sites in the *env* gene. In the Spanish F1 transmission cluster, we also found significant evidence of directional positive selection, although there were no coincidences between sites detected in our study and those observed by Peng et al. This result suggests that, although transmission events occurred fast and probably during the first months after infection, this viral lineage had enough time to adapt to the hosts' immune systems. Indeed, most positively selected sites were located in *env*, and a majority of them fell within $CD8^+$ and/or $CD4^+$ T cell epitopes. A mutation in one of these positions, Nef H89F, has been reported to be associated with immune escape from $CD8^+$ T cells (Fukada et al., 2002; Hoof et al., 2010).

The rapid expansion of this transmission cluster could be explained by the epidemic scenario characterized by an increased risky sexual behavior among MSM (Bezemer et al., 2008; Diez et al., 2014; González-Domenech et al., 2018; Patiño-Galindo et al., 2017). However, the positively selected positions detected, and particularly Nef H89F, might be associated with the already reported high plasma viral loads in the infected patients, that could have facilitated transmission. At this point, it is important to note that variants from this cluster have already been detected in seven Spanish regions, the last cases described in Catalonia (Bes et al., 2017). Also, unpublished results have recently shown the rapid expansion of a new F1 subcluster in Belgium, closely related with the Spanish variant, presenting 188 reported cases between MSM (Vinken et al., 2017). Altogether, these results remark the increasing prevalence of F1 subtype in geographically distant areas from Europe and supports the hypothesis that the introduction and dissemination of these variants strongly rely on the combination of the epidemic scenario and the viral biological properties.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.meegid.2018.09.008.

Declarations of interest

None.

Author contributions

FD, MTC, ED, MS, LPA, and MMT participated in the sequencing of some of the analyzed samples. JAPG, MMT, RS, FGC and JMC conceived and designed the study. JAPG did the analyses. JAPG and JMC conceived the analyses and drafted the manuscript. MMT, RS, and FGC provided critical review and editing of the manuscript. All authors have seen and approved the paper.

Acknowledgments

This work was supported by the Spanish AIDS Research Network, funded by the Instituto de Salud Carlos III, Spanish Health Ministry (Grant n° RD12/0017/0033 to RS). JMC was funded by a Ramón y Cajal contract and the MICIU (project SAF2017-82287-R). RS and FGC were funded by Generalitat Valenciana (project PROMETEO/2016/122). FGC was funded by Spanish MINECO (projects BFU2014-58656-R and BFU2017-89594-R). RS was funded by MINECO (project BFU2017-84762-R) and the European Research Council (grant 724519-Vis-à-vis). MMT and ED are funded by Instituto de Salud Carlos III through Acción Estratégica en Salud Intramural (AESI) (project PI16CIII/00033) and Red de Investigación en SIDA (RIS) (project RD16ISCIII/0002/000C).

References

Abecasis, A.B., Vandamme, A.-M., Lemey, P., 2009. Quantifying differences in the tempo of human immunodeficiency virus type 1 subtype evolution. J. Virol. 83 (24), 12917–12924. https://doi.org/10.1128/JVI.01022-09.

Abecasis, A.B., Wensing, A.M.J., Paraskevis, D., Vercauteren, J., Theys, K., Van de Vijver, D.A.M.C., ... Vandamme, A.-M., 2013. HIV-1 subtype distribution and its demographic determinants in newly diagnosed patients in Europe suggest highly compartmentalized epidemics. Retrovirology 10, 7. https://doi.org/10.1186/1742-4690-10-7.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. B 57 (1), 289–300. https://doi.org/10.2307/2346101.

Bes, M., Piron, M., Casamitjana, N., Gregori, J., Esteban, J.I., Ribera, E., ... Sauleda, S.,

2017. Epidemiological trends of HIV-1 infection in blood donors from Catalonia, Spain (2005-2014). Transfusion 57 (9), 2164–2173. https://doi.org/10.1111/trf.14195.

Bezemer, D., de Wolf, F., Boerlijst, M.C., van Sighem, A., Hollingsworth, T.D., Prins, M., ... Fraser, C., 2008. A resurgent HIV-1 epidemic among men who have sex with men in the era of potent antiretroviral therapy. AIDS 22 (9), 1071–1077. https://doi.org/10.1097/QAD.0b013e3282fd167c.

Capella-Gutiérrez, S., Silla-Martínez, J.M., Gabaldón, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics (Oxford, England) 25 (15), 1972–1973. https://doi.org/10.1093/bioinformatics/btp348.

Cid-Silva, P., Margusino-Framiñán, L., Balboa-Barreiro, V., Martín-Herranz, I., Castro-Iglesias, Á., Pernas-Souto, B., ... Poveda, E., 2018. Initial treatment response among HIV subtype F infected patients who started antiretroviral therapy based on integrase inhibitors. AIDS (London, England) 32 (1), 121–125. https://doi.org/10.1097/QAD.0000000000001679.

Delgado, E., Cuevas, M.T., Domínguez, F., Vega, Y., Cabello, M., Fernández-García, A., ... Thomson, M.M., 2015. Phylogeny and Phylogeography of a recent HIV-1 Subtype F Outbreak among men who have sex with men in Spain Deriving from a Cluster with a Wide Geographic Circulation in Western Europe. PLoS One 10 (11), e0143325. https://doi.org/10.1371/journal.pone.0143325.

Diez, M., Bleda, M. J., Varela, J. R., Ordonana, J., Azpiri, M. A., Vall, M., … EPI-VIH Study Group. (2014). Trends in HIV testing, prevalence among first-time testers, and incidence in most-at-risk populations in Spain: the EPI-VIH Study, 2000 to 2009. Euro Surveillance, 19(47), 20971. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/25443036

Drummond, A.J., Suchard, M.A., 2010. Bayesian random local clocks, or one rate to rule them all. BMC Biol. 8, 114. https://doi.org/10.1186/1741-7007-8-114.

Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29 (8), 1969–1973. https://doi.org/10.1093/molbev/mss075.

ECDC, 2013. Thematic Report : Men Who Have Sex with Men. Monitoring Implementation of the Dublin Declaration on Partnership to Fight HIV/AIDS in Europe and Central Asia: 2012 Progress Report.

Fourment, M., Holmes, E., 2014. Novel non-parametric models to estimate evolutionary rates and divergence times from heterochronous sequence data. BMC Evol. Biol. 14 (1), 163. https://doi.org/10.1186/s12862-014-0163-6.

Fukada, K., Tomiyama, H., Wasi, C., Matsuda, T., Kusagawa, S., Sato, H., ... Takiguchi, M., 2002. Cytotoxic T-cell recognition of HIV-1 cross-clade and clade-specific epitopes in HIV-1-infected Thai and Japanese patients. AIDS (London, England) 16 (5), 701–711. Retrieved from. http://www.ncbi.nlm.nih.gov/pubmed/11964526.

González-Domenech, C.M., Viciana, I., Delaye, L., Mayorga, M.L., Palacios, R., de la Torre, J., ... Santos, J., 2018. Emergence as an outbreak of the HIV-1 CRF19_cpx variant in treatment-naïve patients in southern Spain. PLoS One 13 (1), e0190544. https://doi.org/10.1371/journal.pone.0190544.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59 (3), 307–321. https://doi.org/10.1093/sysbio/syq010.

Hoof, I., Perez, C.L., Buggert, M., Gustafsson, R.K.L., Nielsen, M., Lund, O., Karlsson, A.C., 2010. Interdisciplinary Analysis of HIV-Specific CD8+ T Cell responses against Variant Epitopes reveals Restricted TCR Promiscuity. J. Immunol. 184 (9), 5383–5391. https://doi.org/10.4049/jimmunol.0903516.

Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30 (4), 772–780. https://doi.org/10.1093/molbev/mst010.

Korber, B., Myers, G., 1992. Signature pattern analysis: a method for assessing viral sequence relatedness. AIDS Res. Hum. Retrovir. 8 (9), 1549–1560. https://doi.org/10.1089/aid.1992.8.1549.

Lythgoe, K.A., Fraser, C., 2012. New insights into the evolutionary rate of HIV-1 at the within-host and epidemiological levels. Proc. R. Soc. B Biol. Sci. 279 (1741), 3367–3375. https://doi.org/10.1098/rspb.2012.0595.

Maljkovic Berry, I., Ribeiro, R., Kothari, M., Athreya, G., Daniels, M., Lee, H.Y., ... Leitner, T., 2007. Unequal evolutionary rates in the human immunodeficiency virus type 1 (HIV-1) pandemic: the evolutionary rate of HIV-1 slows down when the epidemic rate increases. J. Virol. 81 (19), 10625–10635. https://doi.org/10.1128/JVI.00985-07.

Martin, D.P., Murrell, B., Golden, M., Khoosal, A., Muhire, B., 2015. RDP4: Detection and analysis of recombination patterns in virus genomes. Virus Evolution 1 (1), vev003. https://doi.org/10.1093/ve/vev003.

Murrell, B., de Oliveira, T., Seebregts, C., Kosakovsky Pond, S.L., Scheffler, K., Southern African Treatment and Resistance Network-SATuRN Consortium, on behalf of the S. A. T. and R. N. (SATuRN), 2012. Modeling HIV-1 drug resistance as episodic directional selection. PLoS Comput. Biol. 8 (5), e1002507. https://doi.org/10.1371/journal.pcbi.1002507.

Patiño-Galindo, J.Á., González-Candelas, F., 2017. The substitution rate of HIV-1 subtypes: a genomic approach. Virus Evolution 3 (2). https://doi.org/10.1093/ve/vex029.

Patiño-Galindo, J.Á., Torres-Puente, M., Bracho, M.A., Alastrué, I., Juan, A., Navarro, D., ... González-Candelas, F., 2017. Identification of a large, fast-expanding HIV-1 subtype B transmission cluster among MSM in Valencia, Spain. PLoS One 12 (2), e0171062. https://doi.org/10.1371/journal.pone.0171062.

Peng, X., Wu, H., Peng, X., Jin, C., Wu, N., 2015. Heterogeneous Evolution of HIV-1 CRF01_AE in men who have sex with men (MSM) and Other Populations in China. PLoS One 10 (12), e0143699. https://doi.org/10.1371/journal.pone.0143699.

Pernas, B., Grandal, M., Mena, A., Castro-Iglesias, A., Cañizares, A., Wyles, D.L., ... Poveda, E., 2014. High prevalence of subtype F in newly diagnosed HIV-1 persons in Northwest Spain and evidence for impaired treatment response. AIDS 28 (12), 1837–1840. https://doi.org/10.1097/QAD.0000000000000326.

Pybus, O.G., Rambaut, A., 2009. Evolutionary analysis of the dynamics of viral infectious disease. Nat. Rev. Genet. 10 (8), 540–550. https://doi.org/10.1038/nrg2583.

Rambaut, A., Lam, T.T., Max Carvalho, L., Pybus, O.G., 2016. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). Virus Evolution 2 (1), vew007. https://doi.org/10.1093/ve/vew007.

Salemi, M., Lewis, M., Egan, J. F., Hall, W. W., Desmyter, J., & Vandamme, A. M. (1999). Different population dynamics of human T cell lymphotropic virus type II in intravenous drug users compared with endemically infected tribes. Proc. Natl. Acad. Sci. of the United States of America, 96(23), 13253–13258. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10557307

Thomson, M.M., Fernández-García, A., Delgado, E., Vega, Y., Díez-Fuertes, F., Sánchez-Martínez, M., ... Pérez-Álvarez, L., 2012. Rapid expansion of a HIV-1 subtype F cluster of recent origin among men who have sex with men in Galicia, Spain. J. Acquir. Immun. Def. Syndrome (1999) 59 (3), e49–e51. https://doi.org/10.1097/QAI.0b013e3182400fc4.

Vinken, L., Fransen, K., Pineda-Peña, A.C., Alexiev, I., Balotta, C., Debaisieux, L., ... Van Laethem, K., 2017. A21 HIV-1 sub-subtype F1 outbreak among MSM in Belgium. Virus Evolution 3 (suppl_1). https://doi.org/10.1093/ve/vew036.020.