

RESEARCH ARTICLE

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



Elena Delgado¹, María Teresa Cuevas¹, Francisco Domínguez^{1a}, Yolanda Vega¹, Marina Cabello^{1b}, Aurora Fernández-García¹, Marcos Pérez-Losada^{2c}, María Ángeles Castro³, Vanessa Montero¹, Mónica Sánchez¹, Ana Mariño⁴, Hortensia Álvarez⁴, Patricia Ordóñez⁵, Antonio Ocampo⁶, Celia Miralles⁶, Sonia Pérez-Castro⁷, María José López-Álvarez⁸, Raúl Rodríguez⁹, Matilde Trigo¹⁰, Julio Diz-Arén¹¹, Carmen Hinojosa¹², Pablo Bachiller¹³, Silvia Hernández-Crespo¹⁴, Ramón Cisterna¹⁴, Eugenio Garduño¹⁵, Lucía Pérez-Álvarez¹, Michael M Thomson^{1*}

OPEN ACCESS

Citation: Delgado E, Cuevas MT, Domínguez F, Vega Y, Cabello M, Fernández-García A, et al. (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. doi:10.1371/journal.pone.0143325

Editor: Chiyu Zhang, Institut Pasteur of Shanghai, Chinese Academy of Sciences, CHINA

Received: January 26, 2015

Accepted: November 3, 2015

Published: November 24, 2015

Copyright: © 2015 Delgado et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Newly obtained sequences are deposited in GenBank under accessions KJ883030-KJ883089, KJKJ883091-KJ883108, KJ883110-KJ883152, KT982428-KT982463.

Funding: This work received support from the Dirección General de Farmacia, Ministerio de Sanidad, Servicios Sociales e Igualdad, Government of Spain, grant EC11-272; European Network of Excellence EUROPRISE (Rational Design of HIV

1 HIV Biology and Variability Unit, Centro Nacional de Microbiología, Instituto de Salud Carlos III, Majadahonda, Madrid, Spain, **2** Centro de Investigação em Biodiversidade e Recursos Genéticos (CIBIO-InBIO), Vairão, Portugal, **3** Department of Internal Medicine, Complejo Hospitalario Universitario de A Coruña, A Coruña, Spain, **4** Infectious Diseases Unit, Department of Internal Medicine, Complejo Hospitalario Universitario Arquitecto Marcide, Ferrol, A Coruña, Spain, **5** Department of Microbiology, Complejo Hospitalario Universitario Arquitecto Marcide, Ferrol, A Coruña, Spain, **6** Department of Internal Medicine, Complejo Hospitalario Universitario de Vigo, Vigo, Pontevedra, Spain, **7** Department of Microbiology, Complejo Hospitalario Universitario de Vigo, Vigo, Pontevedra, Spain, **8** Infectious Diseases Unit, Hospital Universitario Lucus Augusti, Lugo, Spain, **9** Department of Internal Medicine, Complejo Hospitalario Universitario de Ourense, Ourense, Spain, **10** Department of Microbiology, Complejo Hospitalario Provincial de Pontevedra, Pontevedra, Spain, **11** Department of Internal Medicine, Complejo Hospitalario Provincial de Pontevedra, Pontevedra, Spain, **12** Department of Internal Medicine, Hospital Clínico Universitario de Valladolid, Valladolid, Spain, **13** Department of Internal Medicine, Hospital Universitario Río Hortega, Valladolid, Spain, **14** Department of Clinical Microbiology and Infection Control, Hospital Universitario de Basurto, Bilbao, Vizcaya, Spain, **15** Department of Microbiology, Hospital Infanta Cristina, Badajoz, Spain

^a Current address: School of Biology, University of St. Andrews, St. Andrews, United Kingdom

^b Current address: Laboratório de AIDS e Imunologia Molecular, Instituto Oswaldo Cruz, FIOCRUZ, Rio de Janeiro, Brazil,

^c Current address: Computational Biology Institute, The George Washington University, Washington, DC, United States of America

* mthomson@isciii.es

Abstract

We recently reported the rapid expansion of an HIV-1 subtype F cluster among men who have sex with men (MSM) in the region of Galicia, Northwest Spain. Here we update this outbreak, analyze near full-length genomes, determine phylogenetic relationships, and estimate its origin. For this study, we used sequences of HIV-1 protease-reverse transcriptase and *env* V3 region, and for 17 samples, near full-length genome sequences were obtained. Phylogenetic analyses were performed via maximum likelihood. Locations and times of most recent common ancestors were estimated using Bayesian inference. Among samples

Vaccines and Microbicides), grant LSHP-CT-2006-037611; European Research Infrastructures for Poverty Related Diseases (EURIPRED). Seventh Framework Programme: FP7-Capacities-infrastructures-2012-1, grant agreement 312661; Instituto de Salud Carlos III, Subdirección General de Evaluación, and Fondo Europeo de Desarrollo Regional (FEDER), Plan Nacional I + D + I, through project RD12/0017/0026; Consellería de Sanidade, Government of Galicia, Spain (MVI 1291/08); and the Osakidetza-Servicio Vasco de Salud, Basque Country, Spain (MVI-1255-08). Marcos Pérez-Losada was supported by a DC D-CFAR Research Award from the District of Columbia Developmental Center for AIDS Research (P30AI087714) and by an University Facilitating Fund award from George Washington University. Aurora Fernández-García is supported by CIBER in Epidemiology and Public Health, Instituto de Salud Carlos III, Madrid, Spain.

Competing Interests: The authors have declared that no competing interests exist.

analyzed by us, 100 HIV-1 F1 subsubtype infections of monophyletic origin were diagnosed in Spain, including 88 in Galicia and 12 in four other regions. Most viruses ($n = 90$) grouped in a subcluster (Galician subcluster), while 7 from Valladolid (Central Spain) grouped in another subcluster. At least 94 individuals were sexually-infected males and at least 71 were MSM. Seventeen near full-length genomes were uniformly of F1 subsubtype. Through similarity searches and phylogenetic analyses, we identified 18 viruses from four other Western European countries [Switzerland ($n = 8$), Belgium ($n = 5$), France ($n = 3$), and United Kingdom ($n = 2$)] and one from Brazil, from samples collected in 2005–2011, which branched within the subtype F cluster, outside of both Spanish subclusters, most of them corresponding to recently infected individuals. The most probable geographic origin and age of the Galician subcluster was Ferrol, Northwest Galicia, around 2007, while the Western European cluster probably emerged in Switzerland around 2002. In conclusion, a recently expanded HIV-1 subtype F cluster, the largest non-subtype B cluster reported in Western Europe, continues to spread among MSM in Spain; this cluster is part of a larger cluster with a wide geographic circulation in diverse Western European countries.

Introduction

The HIV-1 epidemic among men who have sex with men (MSM) has experienced a notable upsurge in recent years in many countries [1] associated with increased high risk behavior in this population [1,2]. This has been frequently accompanied by the emergence of local HIV-1 transmission clusters [3–16], whose expansion is mostly driven by onward transmission from individuals with recent infection who are unaware of their HIV status [3,4,8,10–13]. In Western Europe and North America, HIV-1 clusters associated with MSM are usually of subtype B, the clade initially introduced and largely predominant in this population [17]. However, several non-subtype B clusters associated with HIV-1 transmission among MSM have been recently reported in Western countries [14,18–21]. The largest one is an F1 subsubtype cluster of Brazilian ancestry which has rapidly spread among MSM in the region of Galicia, Northwest Spain [21]. Since the original report, covering samples collected up to April 2011, the size of this cluster has increased considerably and its geographic range has expanded. Here we update the information on this cluster, analyze near full-length genomes, determine phylogenetic relationships with viruses from other countries, and estimate its geographic and temporal origin.

Materials and Methods

Samples

Plasma samples were collected from 2009 to 2013 from HIV-1-infected individuals residing in all four Galician provinces attended at public hospitals in the cities of A Coruña, Ferrol, Vigo, Pontevedra, Lugo, and Ourense, as well as from individuals attended at 5 hospitals located in four cities of four other Spanish regions: Valladolid (Castilla y León, Central Spain), Bilbao (Basque Country, North Spain), Badajoz (Extremadura, Southwest Spain), and Madrid (Central Spain).

The study was approved by the Bioethics and Animal Well-being Committee of Instituto de Salud Carlos III, Majadahonda, Madrid, Spain. Written informed consent was obtained from all participants in the study.

RNA extraction, RT-PCR amplification and sequencing

RNA was extracted using Nuclisens EasyMAG kit (bioMérieux, Marcy l'Etoile, France) from 1 ml of plasma, following manufacturer's instructions. The HIV-1 protease-reverse transcriptase (PR-RT) *pol* segment and the C2-V3-C3 *env* segment were amplified by RT-PCR followed by nested PCR. Primers for PR-RT amplification were (sequences and HXB2 positions are indicated) RP1-S (GAAAAAGGGCTGTTGGAAATGTGGAA, 2016–2041) and RP-1-A (AAATTTAG GAGTCTTTCCCATATTACTATGC, 3685–3716) in RT-PCR, and PR-O-S2 (GCTAATTTTTT AGGGAARATYTTGGCCTT, 2080–2107) and RT-O-A (TGCTCTGTTAATTGTTTTACATCATT AGTGTG, 3630–3662) in nested PCR, and those used for amplification of the V3 region were described previously [22]. Near full-length genome amplification was done in four overlapping segments, as described [23,24], using RNA extracted either from plasma or from the primary isolate's culture supernatant grown from plasma using a previously described protocol [25].

Sequence electropherograms were viewed and assembled with Seqman (DNASTAR, Madison, WI, USA). For amplicons from suspected dual B/F1 infections, as indicated by frequent mixed peaks in sequence electropherograms at positions of B/F1 discordance, or, in one case, by sequences of different subtypes depending on sequencing primers used, TA cloning of the PCR products with subsequent clone sequencing was performed. In one sample with suspected dual B/F1 infection, nested PCR with PR-RT F1-specific primers (whose 3' ends coincide with positions of B/F1 discordance) was performed. Sequences and HXB2 positions of these primers were PR-RT-nested-F1-S, GAAAAGAAGGACACCAAATGAAAGAATGC (2039–2067) and PR-RT-nested-F1-A, GTTAATTGTTTTACATCATTAGTGTGGGCAG (3625–3655).

Newly obtained sequences are deposited in GenBank under accessions KJ883030-KJ883089, KJ883091-KJ883108, KJ883110-KJ883152 and KT982428-KT982463.

Phylogenetic sequence analyses

Sequences were aligned with MAFFT v.7 [26]. Phylogenetic trees were constructed via maximum likelihood with RAxML v.7.2.7 [27], applying the general time reversible substitution model with CAT approximation for among-site rate heterogeneity, with assessment of node support by bootstrapping. The possibility of intersubtype recombination was analyzed by bootscanning using Simplot v3.5 [28].

To identify viruses from other geographical areas available in public databases related to the subtype F cluster, we used BLAST searches followed by phylogenetic analyses. For this, we downloaded all HIV-1 F1 subsubtype sequences at the Los Alamos HIV Sequence Database [29] ($n = 3,660$) and examined similarity to all available F1 subsubtype near full-length genome sequences ($n = 39$), including the 17 obtained in this study, with local BLAST searches using BioEdit v.7.1.3.0 (Tom Hall, www.mbio.ncsu.edu/BioEdit/bioedit.html). Database sequences with the highest similarity scores to ≥ 2 viruses of the subtype F cluster were selected for phylogenetic analysis with RAxML.

Antiretroviral drug resistance determination

Antiretroviral (ARV) drug resistance in PR-RT sequences was analyzed with the Calibrated Population Resistance Tool [30].

Phyldynamic and phylogeographic analyses

To estimate emergence times of the most recent common ancestors (tMRCA) of clades and their geographical locations, and to analyze the demographic growth of the subtype F cluster, we used a Bayesian Markov Chain Monte Carlo (MCMC) coalescent method as implemented

in BEAST v1.7.5 [31]. For this analysis, we used all PR-RT sequences of the subtype F cluster ≥ 1 kb. Since the evolutionary rate could not be inferred directly from the sequences, due to the narrow time span of sample collection, we estimated it from 68 F1 subsubtype PR-RT sequences from viruses collected along 22 years retrieved from the HIV Sequence Database or obtained by us [21]. Estimated substitution rates were then used as prior means (normal distribution) for the analysis of the subtype F cluster. We chose an HKY substitution model with gamma-distributed among-site rate heterogeneity and two partitions in codon positions (1st +2nd; 3rd) [32]; we also used an uncorrelated lognormal relaxed clock model and a Bayesian skyline plot demographic model [33]. Each MCMC chain was run for 150 million generations, sampling every 5,000. MCMC convergence and effective sample sizes (ESS) were checked with Tracer v.1.5 (<http://tree.bio.ed.ac.uk/software/tracer/>), ensuring that the ESS of each parameter was >200 . Results were summarized with a maximum clade credibility (MCC) tree, using TreeAnnotator v1.5.3, after removal of a 50% burn-in. The MCC tree was visualized with Fig-Tree v1.3.1. (<http://tree.bio.ed.ac.uk/software/figtree/>). Parameter uncertainty was summarized in the 95% highest posterior density (HPD) intervals.

Results

Among HIV-1 infections diagnosed in Spain in 2009–2013, we identified 87 viruses of subtype F, mainly from Galicia, that formed a clade for the PR-RT sequence (Fig 1a). None of the 1,660 HIV-1 samples collected during 1999–2008 in Galicia and previously sequenced by us branched in this cluster. Three viruses of the subtype F cluster (one from Galicia and two from Valladolid, in central Spain) branched off basally to the rest, which formed a subcluster. We will refer to this subcluster as the Galician subcluster. Similarly, we will refer to the entire subtype F cluster, comprising viruses from Spain and other Western European countries and one from Brazil [identified in a previous [21] or in this study (see below)] as the Western European subtype F cluster. For one virus from Galicia, X3049, whose PR-RT bulk sequence was of subtype B (Fig 1a), the sequence electropherogram showed numerous mixed positions with minor peaks overlapping major peaks, most of them corresponding to positions where B subtype and F cluster consensus differ. Suspecting dual B/F1 infection, we obtained 19 clones for the amplified PR-RT fragment, all of which were of subtype B; however, nested PCR with F1 subsubtype-specific primers yielded an amplicon whose sequence branched within the Galician F subcluster (Fig 1a), indicating the presence of a dual B subtype/F cluster infection. Another virus from Galicia, X3461, also showed numerous mixed electropherogram peaks only in protease at positions differing between B subtype and the F cluster. Nineteen PR-RT clones of this virus, analyzed by bootscanning, revealed an infection with diverse variants, with 18 clones being BF1 recombinant and one of subtype B (S1 Fig). Most recombinant clones had 5'-F/B-3' structures, with breakpoints near the protease-RT junction in 12 and at around position 120 of protease in 7, and one had a B/F/B structure, with breakpoints at around positions 120 and 290 of protease (S1 Fig); the subtype B fragment of the recombinant clones derived from the subtype B strain represented by a single clone (S1 and S2 Figs). A phylogenetic tree of protease of the recombinant clones of X3461 having the largest subtype F fragment showed branching in the subtype F cluster (S3 Fig). Three additional viruses from Galicia (X2890, X2955, and X3253) that could not be RT-PCR-amplified for PR-RT but were amplified for protease, also branched within the subtype F cluster for this segment (S3 Fig). Assignment of protease sequences to the Galician subcluster was uncertain because the subcluster was not well supported in the protease tree (S3 Fig). Analysis of ARV drug resistance mutations only revealed a minor population with protease M46I, associated with low level resistance to Nelfinavir, in a patient who was under treatment with atazanavir/ritonavir.

(a) PR-RT

(b) V3

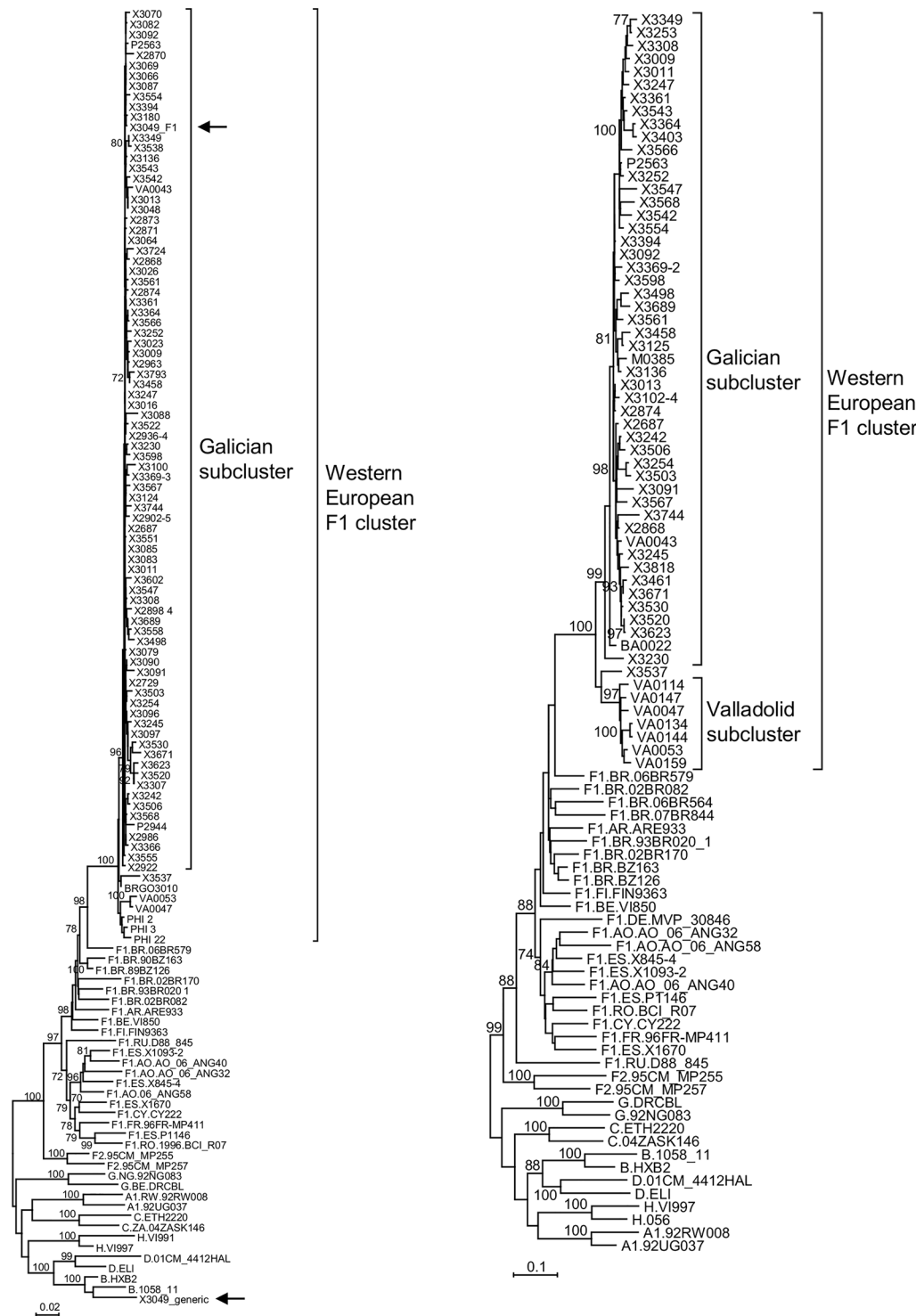


Fig 1. Maximum likelihood trees for the PR-RT (a) and V3 (b) segments of viruses of the subtype F cluster. The PR-RT tree, in addition to the sequences from Spain, also includes three sequences from Lausanne, Switzerland (PHI_2, PHI_3, PHI_22) and one from Goiania, Brazil (BRG0310) previously reported by us to be closely related to the Spanish viruses of the subtype F cluster [21]. In Fig 1(a), both sequences derived from the X3049 sample (one obtained with generic primers and the other with F1-specific primers) (see [Methods](#) for details) are signaled with arrows. Both trees include F1 subsubtype references corresponding to viruses with near full-length genome sequences available in databases. Country of sample collection of these

viruses is indicated with the ISO two-letter country code. Among viruses of this study collected in Spain, those with names starting with X, P, VA, BA, and M are from Galicia, Bilbao, Valladolid, Badajoz, and Madrid, respectively. Only bootstrap values $\geq 70\%$ are shown.

doi:10.1371/journal.pone.0143325.g001

In phylogenetic analyses of the *env* V3 region of viruses from Spain, we identified 59 sequences that branched within the subtype F cluster, 48 corresponding to viruses also analyzed for PR-RT or protease, and 11 to viruses only sequenced for V3. In the V3 tree, 8 viruses (7 from Valladolid and 1 from Galicia) fell outside the Galician subcluster, with the viruses from Valladolid forming a well-supported subcluster (Fig 1b). The V3 region sequence of X3461 was of subtype F or B depending on the primers (forward or reverse) used for sequencing. Therefore, we performed TA cloning of the PCR products with subsequent clone sequencing; this allowed us to confirm the dual nature of the infection in the V3 region of X3461, with some clones branching with viruses of the subtype F cluster and others with subtype B references (S4 Fig). The subtype F V3 sequences of X3461 and that of X3253 fell in the Galician subcluster. In total, 100 viruses of the subtype F cluster were identified in Spain, of which 90 corresponded to the Galician subcluster.

Of the 100 viruses of the subtype F cluster sequenced by us, 88 were collected in Galicia (44 in the city of A Coruña), 8 in Valladolid, 2 in Bilbao, 1 in Badajoz, and 1 in Madrid (Fig 2). All individuals harboring them were men, and all 94 with reported transmission routes were infected sexually. Of these, at least 71 were MSM, according to their self-reported sexual behavior, although 11 were reportedly heterosexual, and for 12 the sexual transmission mode was not specified. The countries of origin, reported for 91 individuals, were Spain ($n = 80$), a Latin American country ($n = 9$), Israel ($n = 1$), and Equatorial Guinea ($n = 1$). At least 97 samples were from new HIV-1 diagnoses and at least 12 were from recent seroconversions (less than one year from last seronegative to first seropositive samples).

Percentages of newly diagnosed HIV-1 infections caused by viruses of the subtype F cluster in Galicia and in the city of A Coruña, where the greatest number of infections of this cluster was diagnosed, are shown in Fig 3. From 2009 to 2012 percentages in Galicia were 4.9%, 27.3%, 14.2%, and 19.6%, respectively; while in A Coruña they were 22.2%, 46.3%, 17.5%, and 42.9%, respectively. Considering only MSM, percentages in Galicia were 13.8%, 46%, 24.3%, and 39.6%, respectively; while in A Coruña they were 50%, 70.8%, 23.8%, and 61.9%, respectively. (Data from 2013 were not included because samples from A Coruña diagnosed that year were not available for sequencing.)

We obtained near full-length genome sequences for 17 viruses of the subtype F cluster. All were uniformly of F1 subsubtype in the bootscan analyses (Fig 4). In the phylogenetic tree (Fig 5) they formed a cluster, with the virus from Valladolid branching in a basal position and the virus 06BR579, from Sao Paulo, Brazil, as its closest relative.

Through local BLAST searches with all F1 subsubtype viruses downloaded from the Los Alamos Database, using all available near full-length genome F1 subsubtype sequences as references, followed by phylogenetic analyses, we identified 21 partially-sequenced database viruses branching within the Western European subtype F cluster, which included 5 previously reported by us [21]. Of those viruses, 2 were collected in Madrid [34], 8 in Switzerland [35], 5 in Belgium [36], 3 in France [12], 2 in the United Kingdom (UK), and 1 in Goiania, West-Central Brazil [37] (Fig 6). A second virus from Goiania (BRGO3042) was related to the Western European cluster, branching outside of it (Fig 6c). Both viruses from Madrid branched within the Galician subcluster, and all non-Spanish viruses branched outside of both Galician and Valladolid subclusters. All 4 Swiss viruses analyzed for p24^{gag} grouped in a subcluster (Fig 6b). Similarly, three viruses from Belgium and one from Galicia (X3537) analyzed for RT grouped in a subcluster (Fig 6c). The common ancestry of X3537 and the Belgian viruses was also

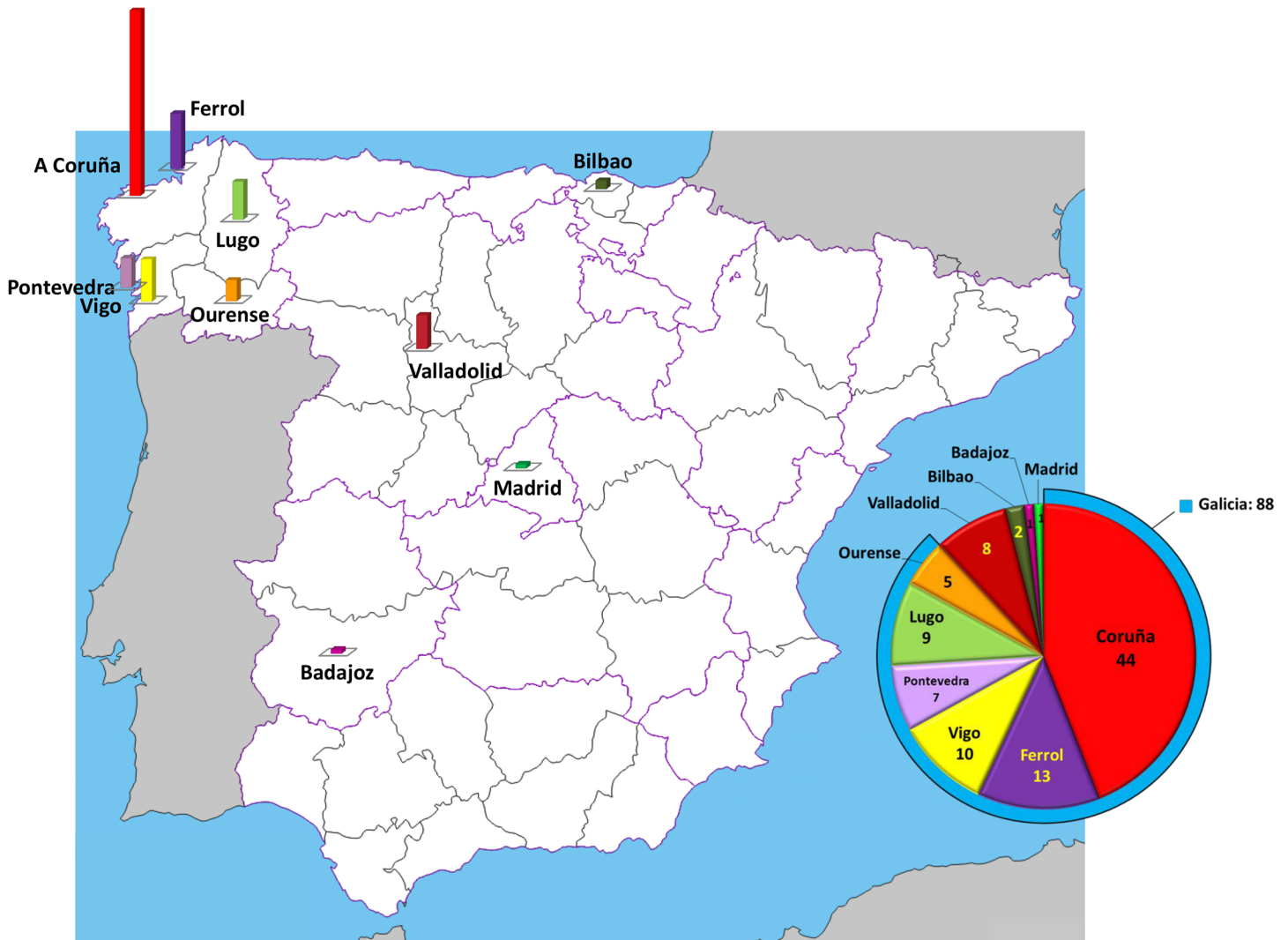


Fig 2. Distribution of viruses of the subtype F cluster sequenced by us according to city of sample collection.

doi:10.1371/journal.pone.0143325.g002

supported for the V3 region (Fig 6d). According to the information available in databases or in published studies [12,35], at least 9 infections with non-Spanish viruses of the Western European subtype F cluster correspond to primary or recent infections. Data on transmission routes were available for 5 viruses, 3 from Switzerland [35] and 2 from Belgium [36]; all were sexually transmitted, with 4 corresponding to MSM. Sample collection dates ranged from 2005 to 2011, with the earliest dates in 2005 corresponding to 4 viruses from Switzerland.

To estimate the origins of the Western European cluster and of the Galician subcluster, we analyzed, using a Bayesian approach, PR-RT sequences of the Western European subtype F cluster, which, in addition to the Galician viruses, included viruses from Valladolid (n = 3), Madrid (n = 2), Bilbao (n = 2), Switzerland (n = 3), Belgium (n = 2), and Goiania, Brazil (n = 1). We also included all F1 subsubtype viruses of the Brazilian variant with available near full-length sequences lacking drug resistance-associated mutations, and two other viruses from Brazil (BR03SP179, from Sao Paulo, and BRGO3042, from Goiania) which were found to be related to the Western European subtype F cluster, branching outside of it, in phylogenetic

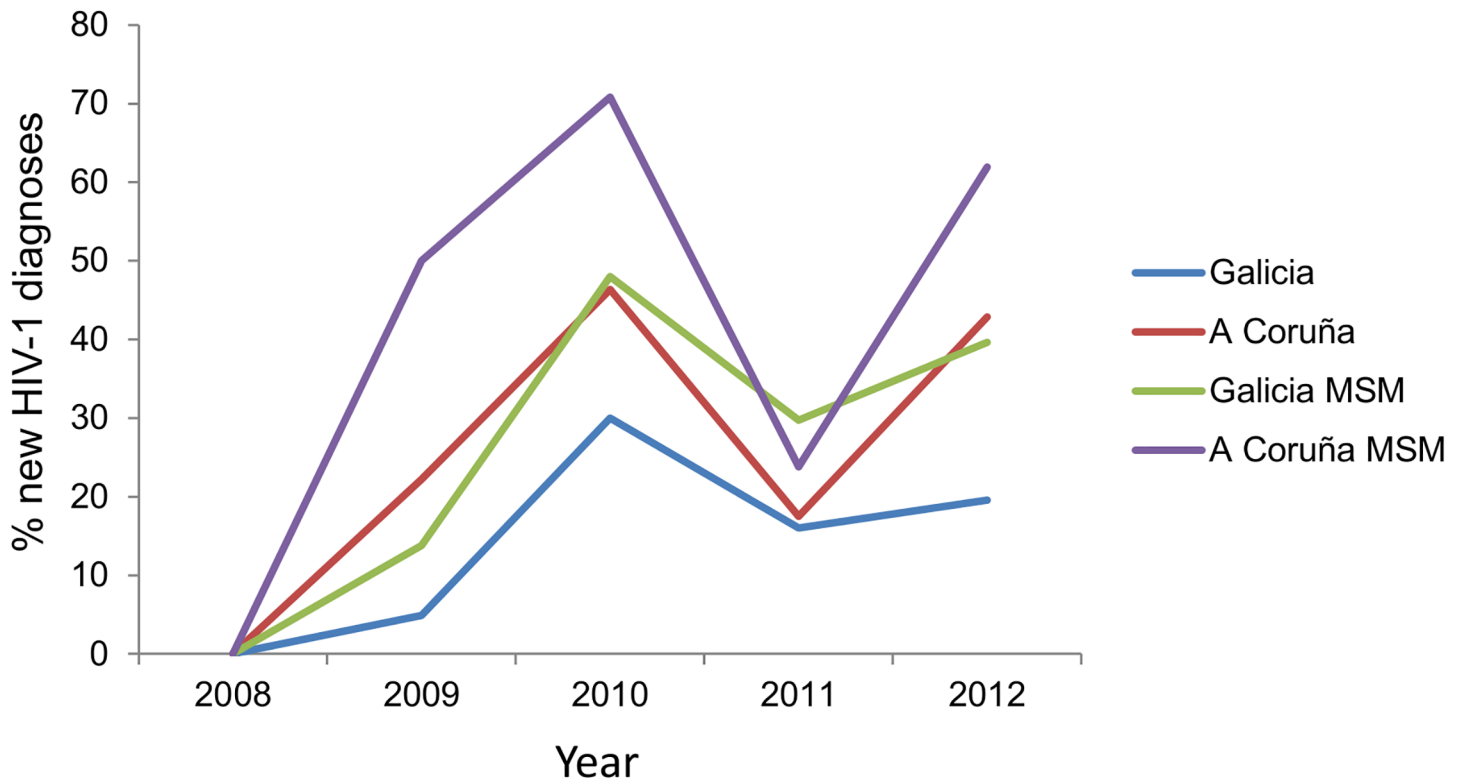


Fig 3. Prevalence of infections with viruses of the subtype F cluster in 2009–2012 among all new HIV-1 diagnoses and among new HIV-1 diagnoses in MSM in Galicia and A Coruña city (Northwest Galicia).

doi:10.1371/journal.pone.0143325.g003

analyses (Fig 6c). The results, summarized in the MCC tree (Fig 7), estimated the tMRCA of the Galician subcluster in 2007.5 (95% HPD 2006.2–2008.6) and that of the Western European cluster in 2002.2 (95% HPD 2000.1–2003.9). The most probable origin of the Galician subcluster was estimated in Ferrol [0.57 posterior probability (PP)], followed by A Coruña (distant 20 km by sea and 55 km by road from Ferrol) (PP = 0.16). The most probable origin of the Western European cluster was Switzerland (PP = 0.72) followed by Goiania (PP = 0.1). Since Ferrol and A Coruña are geographically close, as are Vigo and Pontevedra (distant 30 km), we repeated the Bayesian analysis combining the samples from Ferrol and A Coruña and those from Vigo and Pontevedra. In this new analysis, the most probable location for the origin of the Galician subcluster was Ferrol-Coruña (PP = 0.88).

The Bayesian skyline plot (Fig 8) showed an exponential growth in the effective number of infections starting at the end of 2008 and continuing through the first half of 2009, with slower growth rates in 2010 and 2011 and stabilization in 2012.

Discussion

The most remarkable finding of this study is probably that the Galician HIV-1 subtype F cluster previously reported by us [21] is part of a larger cluster that has independently spread in at least three other Western European geographic areas, as evidenced by local clustering of sequences from Valladolid, central Spain (7 infections) (Fig 1b), Switzerland (7 infections) (Fig 5b and 5c) (although clustering of 3 Swiss viruses is not strongly supported in PR-RT, they share K103N drug resistance mutation [35], and in p24^{gag} the 4 Swiss viruses group with a 96% bootstrap value), and Belgium (5 infections, with a sixth virus from Galicia clustering with

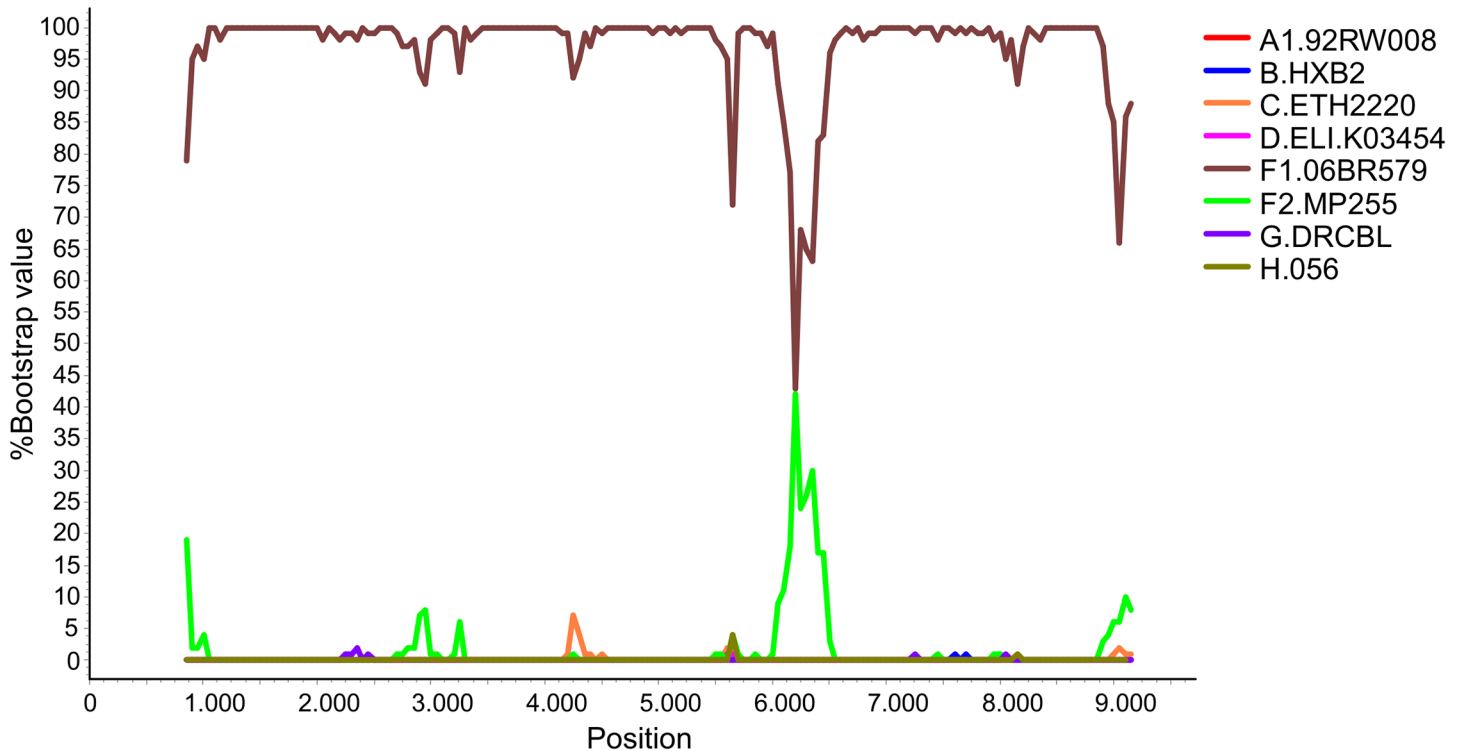


Fig 4. Bootscan analysis of the near full-length genome sequence of X3364. The analysis was done with Simplot v.3.5.1, using a window of 500 nucleotides, moving in 50 nucleotide increments. Phylogenetic trees were constructed using the neighbor-joining algorithm based on Kimura 2-parameter distances, with Tv:Ti ratios estimated from the dataset.

doi:10.1371/journal.pone.0143325.g004

them) (Fig 6c and 6d). Additionally, 2 newly diagnosed infections from the UK branching within the Western European subtype F cluster are closely related to each other (Fig 6a), and 3 viruses from France, with uncertain mutual relations, also branch within the Western European subtype F cluster (Fig 6c and 6d). The fact that most of the viruses are from recent diagnoses, with 7 infections from Switzerland [35] and 2 from France [12] being from primary or recent infections, also indicates that viruses of the Western European subtype F cluster are propagating throughout multiple local networks. The available data indicate that transmission takes place among MSM, since all Spanish cases studied by us are sexually-infected men, at least 71 MSM, as are 3 infections from Switzerland (2 from MSM) [35] and 2 from Belgium (both from MSM) [36]. Considering that all 100 infections of the cluster studied by us were detected in men, we strongly suspect that those cases from Spain reported to be transmitted via heterosexual or unspecified sexual contact also correspond to MSM.

Since our initial report [21] comprising samples collected up to April 2011, the number of viruses of the Galician subtype F subcluster identified by us has increased from 38 individuals to 90, including individuals diagnosed outside of Galicia in the cities of Bilbao (n = 2), Valladolid (n = 1), Badajoz (n = 1), and Madrid (n = 1). It is important to note that this study does not derive from a national survey. Outside of Galicia, we have only analyzed samples from all public hospitals in the Basque Country and in the cities of Valladolid, Badajoz, León, and Toledo, and from two hospitals in the province of Madrid. Therefore, it is most likely that the real size of the subtype F cluster in Spain is considerably larger than here described. Comparing with other non-subtype B clusters reported in Western Europe [7,18–20,38–40], the Galician subtype F subcluster is substantially larger than all of them, followed by a CRF01_AE cluster

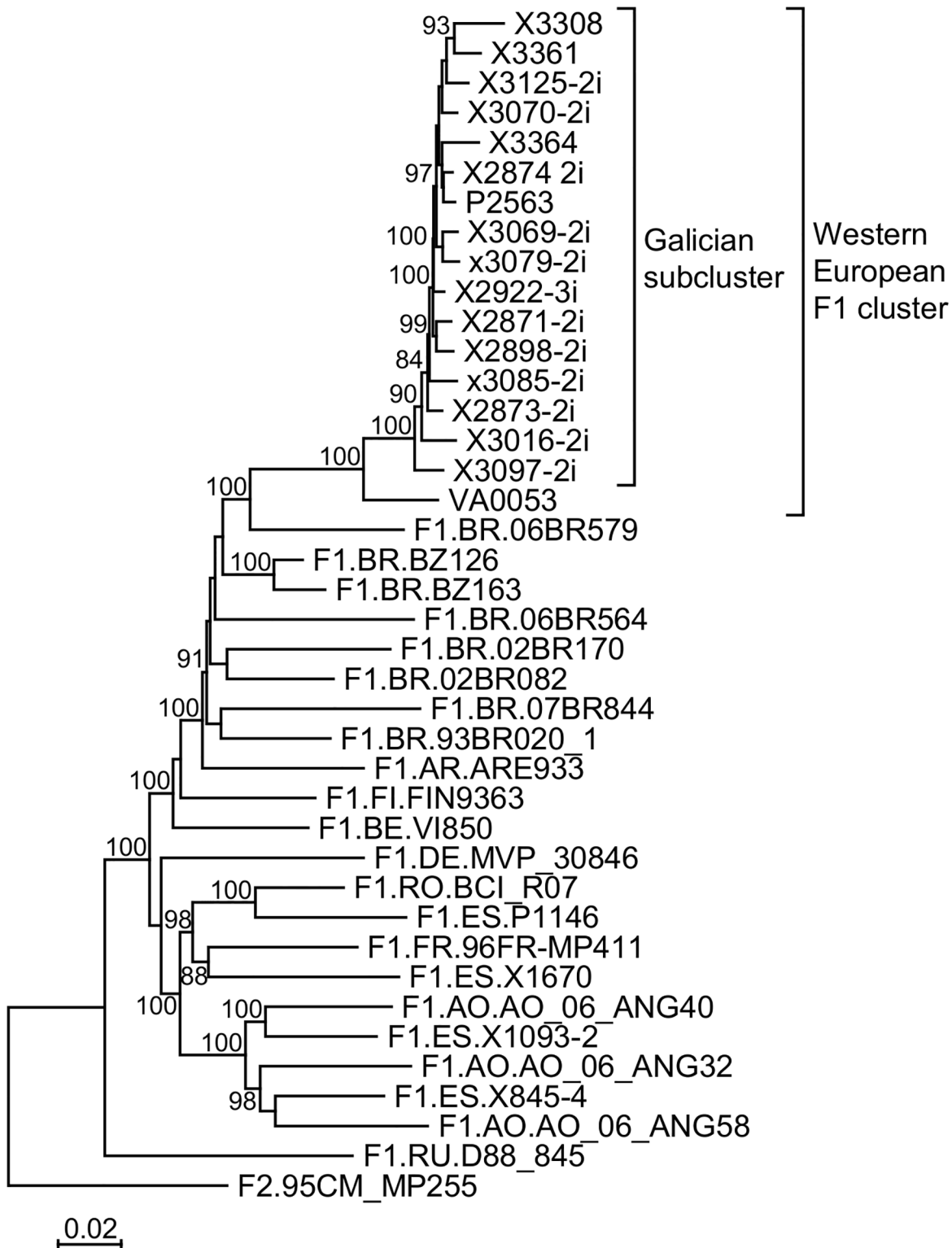


Fig 5. Maximum likelihood tree of near full-length genome sequences of viruses of the subtype F cluster. F1 subsubtype references are included in the analysis, with country of sample collection indicated with the ISO two-letter country code. The tree is rooted with an F2 subsubtype sequence. Only bootstrap values $\geq 80\%$ are shown.

doi:10.1371/journal.pone.0143325.g005

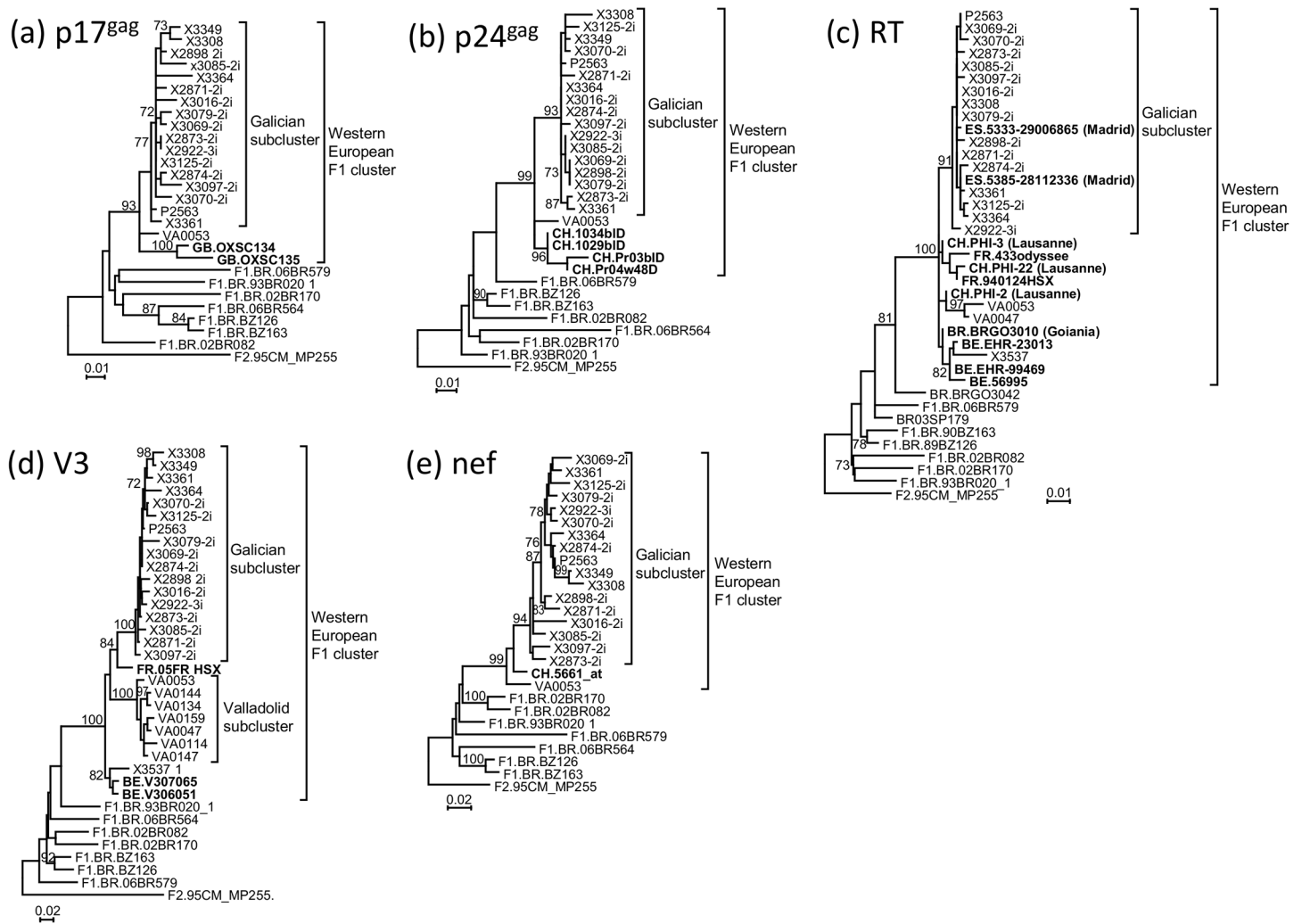


Fig 6. Maximum likelihood trees of sequences from databases belonging to the Western European subtype F cluster. The trees correspond to genome segments in (a) p17^{gag} (HXB2 positions 790–1185), (b) p24^{gag} (HXB2 positions 1575–2042), (c) reverse transcriptase (HXB2 positions 2637–3278), (d) *env* V3 region (HXB2 positions 6709–7448), and (e) *nef* (HXB2 positions 8797–9417). References for the subtype F cluster are near full-length genome sequences obtained by us, with names beginning with X, P or VA, with ‘i’ ending denoting those derived from cultured isolates. Database viruses branching in the subtype F cluster are in bold type. Countries of sample collection of database viruses are indicated with the ISO two-letter codes: BE, Belgium; BR, Brazil; CH, Switzerland; ES, Spain; FR, France; GB, United Kingdom. Cities of sample collection of database viruses of the subtype F cluster, when known, are in parentheses. Only bootstrap values $\geq 70\%$ are shown.

doi:10.1371/journal.pone.0143325.g006

comprising 46 injecting drug users reported in Sweden [40], and is part of an even larger cluster comprising viruses from other regions of Spain and at least four other Western European countries and Brazil. In a recent study [41] analyzing HIV-1 transmission networks using 84,757 PR-RT sequences available at the Los Alamos HIV Sequence Database, with clustering based on genetic distances $\leq 1\%$, the originally described Galician subtype F cluster [21] represented the largest HIV-1 network in Western Europe and one of the largest in the world.

Using a Bayesian phylogeographic approach, the most probable origin of the Galician subtype F subcluster was estimated in Ferrol, province of A Coruña, Northwest Spain (PP = 0.57), with the second most probable location being the city of A Coruña, near Ferrol (PP = 0.16) (Fig 7). This is consistent with the places of the earliest diagnosed infections with viruses of the Galician subcluster studied by us: the first two cases were diagnosed in Ferrol in March and

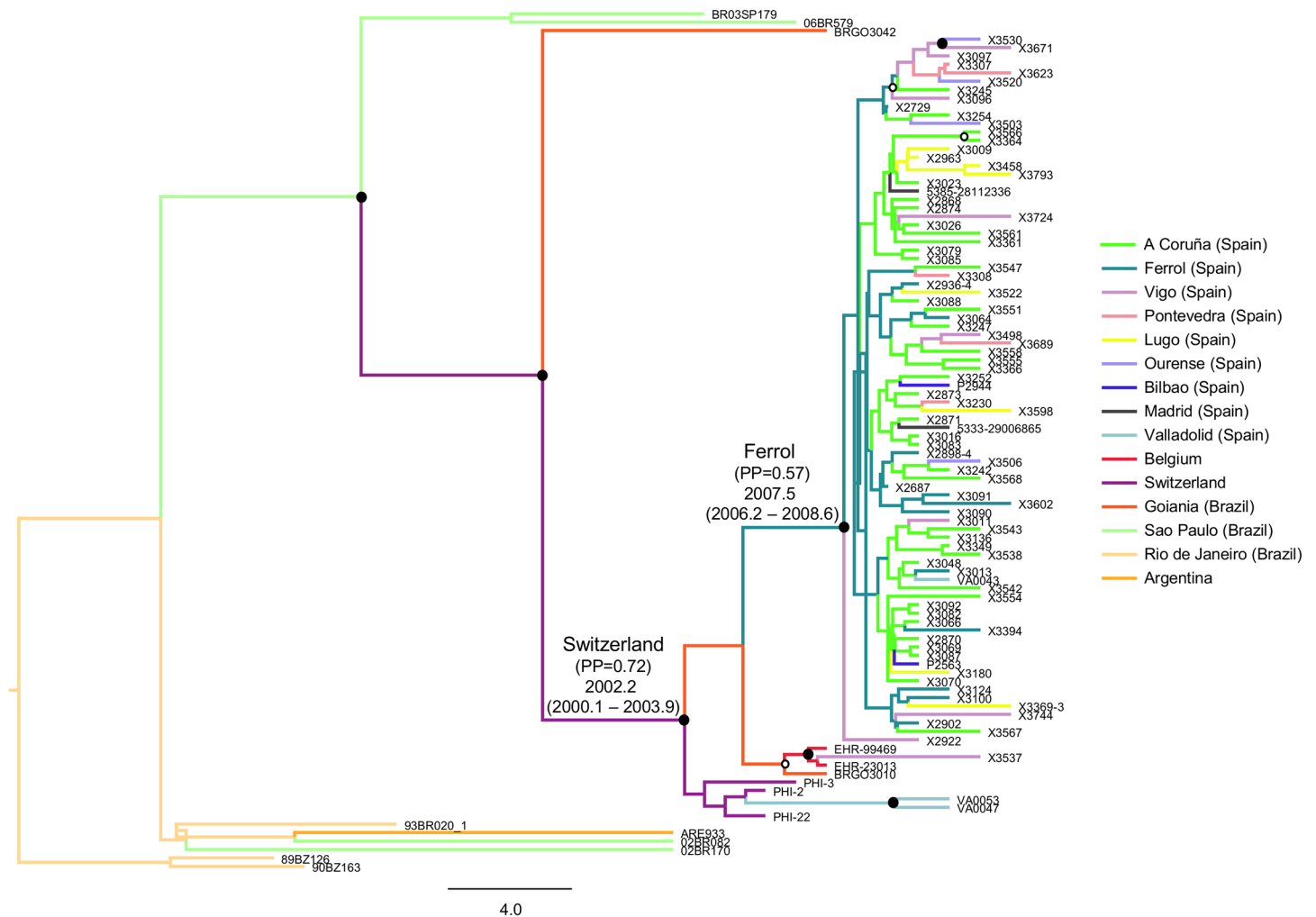


Fig 7. Maximum clade credibility tree of PR-RT sequences of the subtype F Western European cluster and Galician subcluster. Nodes supported by PP = 1 and PP = 0.95–0.99 are marked with filled and unfilled circles, respectively. Colors of terminal and internal branches represent sampling locations and most probable locations of the corresponding nodes, respectively, according to the legend on the right. For the nodes corresponding to the Galician subcluster and the Western European cluster, the posterior probabilities for the most probable locations and the tMRCAs are indicated above the subtending branches (95% HPD intervals are in parentheses).

doi:10.1371/journal.pone.0143325.g007

May 2009 and the next six in A Coruña, two in November-December 2009 and four in January-March 2010. Our Bayesian analysis placed the origin of the Western European subtype F cluster in Switzerland (PP = 0.72). This is also consistent with the years of sample collection, which among infections diagnosed in Switzerland were 2005 for four and 2006 for one, all from primary infections [35], while for Belgian infections they were 2006 through 2010, and for the Goianian infection it was 2007. The spread of the subtype F cluster in Valladolid appears to be more recent, with no diagnoses dating earlier than 2011. It is important to note that the conclusions drawn from the phylogeographic analyses about the ancestral locations of clades should be interpreted with caution, since geographic sampling is incomplete. However, we are confident that the Galician subtype F outbreak probably originated in Ferrol, since we have analyzed most of the new HIV-1 diagnoses from Galicia. In contrast, the estimated origin of the Western European cluster (Switzerland) is less certain, considering the relative paucity of available HIV-1 sequences from new infections.

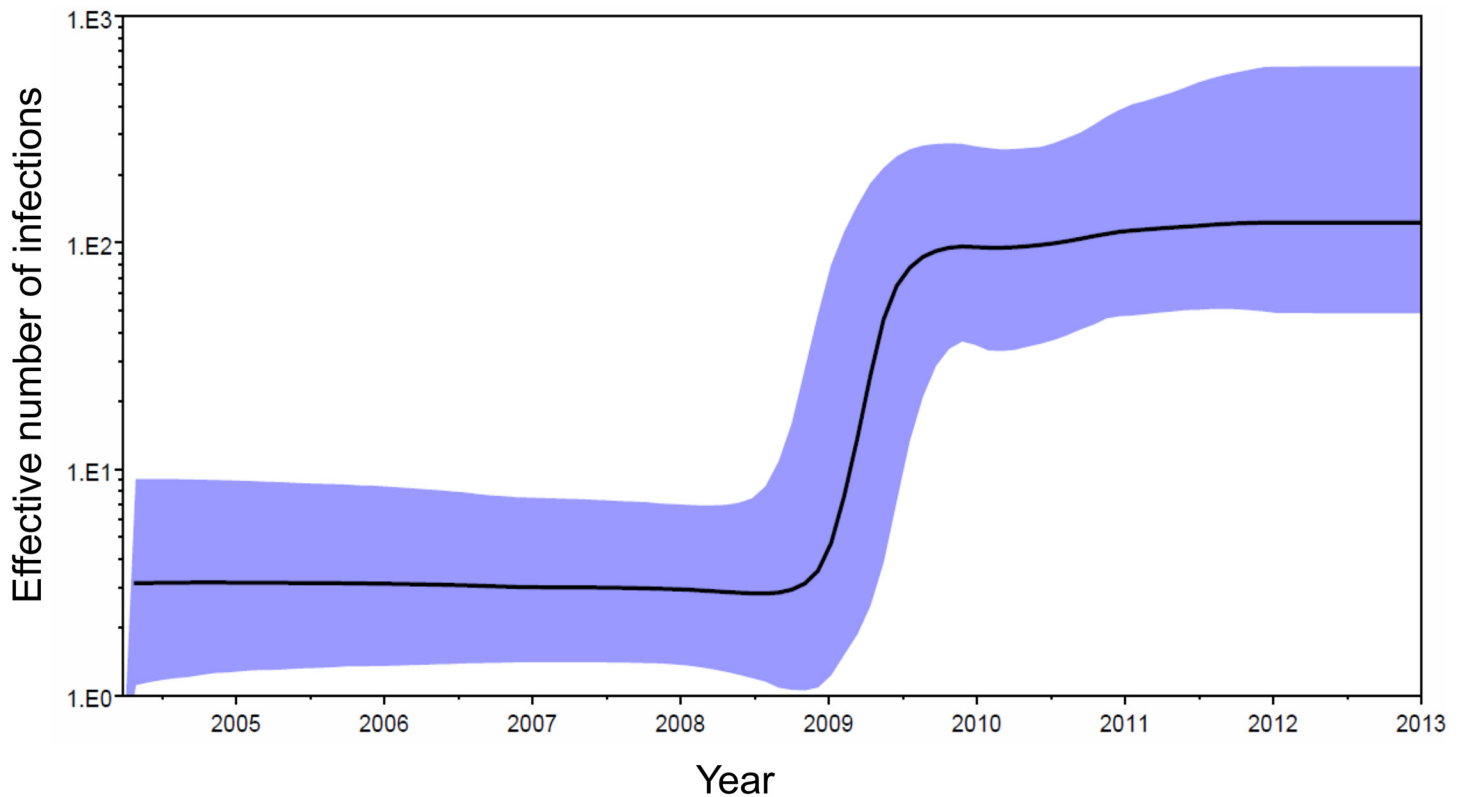


Fig 8. Bayesian skyline plot of the population growth of the subtype F cluster. The black line represents the median estimate of the effective number of infections through time (logarithmic scale) and the shaded area represents the 95% HPD credibility interval. The horizontal axis represents calendar years.

doi:10.1371/journal.pone.0143325.g008

The tMRCA estimated for the Galician subtype F subcluster was around 2007, which is consistent with the date previously estimated by us using a smaller dataset [21], but differs widely from a younger estimate (1993), recently proposed by other authors based on samples from two cities in Galicia [42]. The sequences used in the latter study are currently unavailable in public databases. However we find difficult to reconcile the 1993 estimate with (1) the total absence of viruses of the Western European subtype F cluster in our surveillance analyses of 1,660 HIV-1 samples collected across Galicia from 1999 to 2008; and with (2) the date (2002) and geographical origin (Switzerland) estimated here for this cluster using all available sequences. It is possible either that the basal sequence presumably responsible for the estimated tMRCA in [42] may not represent the origin of the Western European subtype F cluster, but rather had a separate ancestry in a Brazilian subtype F variant related to the MRCA of the mentioned cluster; or that the mentioned sequence is of recombinant origin, since recombinant sequences have a tendency to branch basally to one of the parental clades, mimicking an ancestral split [43].

The rapid spread of the subtype F cluster in Galicia, where its prevalence among new HIV-1 diagnoses grew from 0% in 2008 to 27% (46% among MSM) in 2010, is remarkable (Fig 3). The increase was even more prominent in the city of A Coruña, where the percentage reached 46% (71% among MSM) in 2010. The rapid growth of the subtype F cluster is also supported by the Bayesian skyline plot analysis, which estimates that most of its growth occurred in less than one year from the end of 2008 (Fig 8). Its rapid spread in Galicia, together with its propagation in multiple geographically distant locations, suggests the existence of some inherent

biological property of the viruses of this cluster favoring transmission. It is difficult to ascribe its rapid growth only to unusually risky and promiscuous sexual behaviors among a group of MSM in Galicia, since we have detected multiple other clusters among MSM in this region and none of them has approached the growth dynamics of the subtype F cluster. The second largest cluster detected in Galicia comprises 44 subtype B viruses, most of them from MSM, diagnosed during an 8 year period (compared to 90 in less than 5 years for the Galician subtype F subcluster). At least 3 individuals of the subtype F cluster had dual infections (2 detected initially and 1 during the follow-up), and in 2 of them the second virus belonged to a Galician subtype B cluster, one of them being the largest of subtype B in this region. This suggests that the subtype F network overlaps with other sexual networks in Galicia, which makes difficult to explain solely on epidemiological grounds why no other clusters circulating in Galicia have spread at a similar speed as the subtype F cluster. Moreover, its spread in local networks in at least four distant areas from three countries, a phenomenon unreported for other Western European HIV-1 clusters, also supports the idea of an unusual high transmissibility of the subtype F cluster.

In summary, the HIV-1 subtype F cluster here described is the largest and most widely geographically spread nonsubtype B cluster reported in Western Europe, exhibiting an unusually rapid expansion and involving multiple separate sexual networks in distant geographical areas. These epidemiological features suggest some inherent biological properties favoring sexual transmission, whose study will require further work. In this regard, distinct *in vivo* biological features of viruses of the Galician subtype F cluster have been reported recently, with significantly higher plasma viral loads and poorer virologic responses to antiretroviral therapy compared to subtype B viruses [44]. The rapid expansion of this cluster in Galicia is reminiscent of that of other genetic forms in established HIV-1 epidemics, such as subtype C in southern Brazil [45], BG recombinant viruses in Cuba [46], subtype A in Greece [47], CRF01_AE among MSM in China [48], or CRF63_02A1 in Siberia [49]. This highlights the need for continued molecular epidemiological surveillance of HIV-1 variants, which may have multiple implications, including the design of vaccine immunogens [50,51]. Finally, it is important to note that the vast majority of infections of the subtype F cluster were newly diagnosed and that its rapid expansion in Galicia took place in the context of a public health care system covering all HIV-1-infected individuals. This indicates that the spread of the Galician subtype F cluster is mostly driven by transmission from recently infected individuals unaware of their HIV status, as reported for other clusters transmitted among MSM [3,4,8,10–13], and that, contrary to some beliefs, universal access to effective antiretroviral therapies may not result in substantial control of the epidemic spread of HIV-1 [2,52]. Therefore, it will be very important to implement public health measures aimed at reducing high risk sexual behaviors and to support efforts for the development of an effective vaccine and other methods for prevention of HIV-1 transmission in order to contain the epidemic.

Supporting Information

S1 Fig. Bootscan analyses of four PR-RT clones of X3461. The analysis was done using a window of 300 nucleotides for clone 18 and of 150 nucleotides for all other clones, moving in 20 nucleotide increments. Phylogenetic trees were constructed using the neighbor-joining algorithm based on Kimura 2-parameter distances, with Tv:Ti ratios estimated from the dataset. The analyses were done using subtype references, with a Galician F1 cluster virus used as F1 reference. For clones showing a BF recombinant structure, new bootscan analyses (shown in the figure) were done incorporating the subtype B clone 18 as reference strain. Eleven other clones had bootscan plots virtually identical to that of clone 7 and seven to that of clone 4. (TIF)

S2 Fig. Maximum likelihood tree for the RT gene of X3461 clones. Only bootstrap values $\geq 70\%$ are shown.
(TIF)

S3 Fig. Maximum likelihood tree of the protease gene for X2890, X2955, and X3253, and for 12 F^{PR}/B^{RT} recombinant clones of X3461 (labeled with circles), analyzed together with representative viruses of the F1 cluster and F1 subtype references. Only bootstrap values $\geq 70\%$ are shown.
(TIF)

S4 Fig. Maximum likelihood tree of X3461 clones of the V3 region. Sequences obtained by bulk sequencing of the PCR product using forward [X3461_bulk(a)] or reverse [X3461_bulk(b)] primers are also included in the analysis. Only bootstrap values $\geq 70\%$ are shown.
(TIF)

Acknowledgments

We thank Dr. José Antonio Taboada, Consellería de Sanidade da Xunta de Galicia and Dr. Daniel Zulaica, Unidad de Coordinación del Plan de Prevención y Control del SIDA, Osakidetza-Servicio Vasco de Salud, for their support of this study. We also thank Elena García Bodas and the personnel at the Genomic Unit at Centro Nacional de Microbiología, Instituto de Salud Carlos III, Majadahonda, Madrid, Spain, for technical assistance.

Author Contributions

Conceived and designed the experiments: MMT LPA. Performed the experiments: FD AFG VM MS. Analyzed the data: MMT ED MTC FD YV MC AFG MPL. Contributed reagents/materials/analysis tools: AM HA PO AO CM SPC MJLA RR MT JDA CH PB SHC RC EG MAC. Wrote the paper: MMT. Made contributions to the writing of the manuscript: ED AM HA PO AO CM SPC MJLA RR MT JDA CH PB SHC RC EG MAC LPA FD AFG VM MS MTC YV MC MPL.

References

1. Beyrer C, Baral SD, van Griensven GF, Goodreau SM, Chariyalertsak S, Wirtz AL, et al. Global epidemiology of HIV infection in men who have sex with men. *Lancet* 2012; 380: 367–377. doi: [10.1016/S0140-6736\(12\)60821-6](https://doi.org/10.1016/S0140-6736(12)60821-6) PMID: [22819660](https://pubmed.ncbi.nlm.nih.gov/22819660/)
2. Bezemer D, de Wolf F, Boerlijst MC, van Sighem A, Hollingsworth TD, Prins M, et al. A resurgent HIV-1 epidemic among men who have sex with men in the era of potent antiretroviral therapy. *AIDS* 2008; 22: 1071–1077. doi: [10.1097/QAD.0b013e3282fd167c](https://doi.org/10.1097/QAD.0b013e3282fd167c) PMID: [18520351](https://pubmed.ncbi.nlm.nih.gov/18520351/)
3. Brenner BG, Roger M, Routy JP, Moisi D, Ntemgwa M, Matte C, et al. High rates of forward transmission events after acute/early HIV-1 infection. *J Infect Dis* 2007 195: 951–959. PMID: [17330784](https://pubmed.ncbi.nlm.nih.gov/17330784/)
4. Lewis F, Hughes GJ, Rambaut A, Pozniak A, Leigh Brown AJ (2008) Episodic sexual transmission of HIV revealed by molecular phylodynamics. *PLOS Med* 2008; 5:e50. doi: [10.1371/journal.pmed.0050050](https://doi.org/10.1371/journal.pmed.0050050) PMID: [18351795](https://pubmed.ncbi.nlm.nih.gov/18351795/)
5. Cuevas MT, Muñoz-Nieto M, Thomson MM, Delgado E, Iribarren JA, Cilla G et al. HIV-1 transmission cluster with T215D revertant mutation among newly diagnosed patients from the Basque Country, Spain. *J Acquir Immune Defic Syndr* 2009; 51: 99–103. doi: [10.1097/QAI.0b013e318199063e](https://doi.org/10.1097/QAI.0b013e318199063e) PMID: [19282784](https://pubmed.ncbi.nlm.nih.gov/19282784/)
6. Cuevas M, Fernández-García A, Sánchez-García A, González-Galeano M, Pinilla M, Sánchez-Martínez M, et al. Incidence of non-B subtypes of HIV-1 in Galicia, Spain: high frequency and diversity of HIV-1 among men who have sex with men. *Euro Surveill* 2009; 14:pii: 19413.
7. Chalmet K, Staelens D, Blot S, Dinakis S, Pelgrom J, Plum J, et al. Epidemiological study of phylogenetic transmission clusters in a local HIV-1 epidemic reveals distinct differences between subtype B and non-B infections. *BMC Infect Dis* 2010; 10: 262. doi: [10.1186/1471-2334-10-262](https://doi.org/10.1186/1471-2334-10-262) PMID: [20822507](https://pubmed.ncbi.nlm.nih.gov/20822507/)

8. Bezemer D, van Sighem A, Lukashov VV, van der Hoek L, Back N, Schuurman R, et al. Transmission networks of HIV-1 among men having sex with men in the Netherlands. *AIDS* 2010; 24: 271–282. doi: [10.1097/QAD.0b013e328333ddee](https://doi.org/10.1097/QAD.0b013e328333ddee) PMID: [20010072](https://pubmed.ncbi.nlm.nih.gov/20010072/)
9. Zehender G, Ebranati E, Lai A, Santoro MM, Alteri C, Giuliani M, et al. Population dynamics of HIV-1 subtype B in a cohort of men-having-sex-with-men in Rome, Italy. *J Acquir Immune Defic Syndr* 2010; 55: 156–160. doi: [10.1097/QAI.0b013e3181eb3002](https://doi.org/10.1097/QAI.0b013e3181eb3002) PMID: [20703157](https://pubmed.ncbi.nlm.nih.gov/20703157/)
10. Brenner BG, Roger M, Stephens D, Moisi D, Hardy I, Weinberg J, et al. Transmission clustering drives the onward spread of the HIV epidemic among men who have sex with men in Quebec. *J Infect Dis* 2011; 204: 1115–1119. doi: [10.1093/infdis/jir468](https://doi.org/10.1093/infdis/jir468) PMID: [21881127](https://pubmed.ncbi.nlm.nih.gov/21881127/)
11. Ambrosioni J, Junier T, Delhumeau C, Calmy A, Hirschel B, Zdobnov E, et al. Impact of highly active antiretroviral therapy on the molecular epidemiology of newly diagnosed HIV infections. *AIDS* 2012; 26: 2079–2086. doi: [10.1097/QAD.0b013e32835805b6](https://doi.org/10.1097/QAD.0b013e32835805b6) PMID: [23052354](https://pubmed.ncbi.nlm.nih.gov/23052354/)
12. Frange P, Meyer L, Deveau C, Tran L, Goujard C, Ghosn J, et al. Recent HIV-1 infection contributes to the viral diffusion over the French territory with a recent increasing frequency. *PLOS One* 2012; 7: e31695. doi: [10.1371/journal.pone.0031695](https://doi.org/10.1371/journal.pone.0031695) PMID: [22348121](https://pubmed.ncbi.nlm.nih.gov/22348121/)
13. Audelin AM, Cowan SA, Obel N, Nielsen C, Nielsen C, Jørgensen LB, Gerstoft. Phylogenetics of the Danish HIV epidemic: the role of very late presenters in sustaining the epidemic. *J Acquir Immune Defic Syndr* 2013; 62: 102–108. doi: [10.1097/QAI.0b013e318276becc](https://doi.org/10.1097/QAI.0b013e318276becc) PMID: [23075917](https://pubmed.ncbi.nlm.nih.gov/23075917/)
14. Antoniadou ZA, Kousiappa I, Skoura L, Pilalas D, Metallidis S, Nicolaidis P, et al. Short communication: molecular epidemiology of HIV type 1 infection in northern Greece (2009–2010): evidence of a transmission cluster of HIV type 1 subtype A1 drug-resistant strains among men who have sex with men. *AIDS Res Hum Retroviruses* 2014; 30: 225–232. doi: [10.1089/AID.2013.0166](https://doi.org/10.1089/AID.2013.0166) PMID: [24059291](https://pubmed.ncbi.nlm.nih.gov/24059291/)
15. Ye J, Xin R, Yu S, Bai L, Wang W, Wu T, et al. Phylogenetic and temporal dynamics of human immunodeficiency virus type 1 CRF01_AE in China. *PLOS One* 2013; 8: e54238. doi: [10.1371/journal.pone.0054238](https://doi.org/10.1371/journal.pone.0054238) PMID: [23365653](https://pubmed.ncbi.nlm.nih.gov/23365653/)
16. Han X, An M, Zhang M, Zhao B, Wu H, Liang S, et al. Identification of 3 distinct HIV-1 founding strains responsible for expanding epidemic among men who have sex with men in 9 Chinese cities. *J Acquir Immune Defic Syndr* 2013; 64: 16–24. doi: [10.1097/QAI.0b013e3182932210](https://doi.org/10.1097/QAI.0b013e3182932210) PMID: [23542640](https://pubmed.ncbi.nlm.nih.gov/23542640/)
17. Kuiken C, Thakallapalli R, Esklid A, de Ronde A. Genetic analysis reveals epidemiologic patterns in the spread of human immunodeficiency virus. *Am J Epidemiol* 2000; 152: 814–822. PMID: [11085392](https://pubmed.ncbi.nlm.nih.gov/11085392/)
18. Gifford RJ, de Oliveira T, Rambaut A, Pybus OG, Dunn D, Vandamme AM, et al. Phylogenetic surveillance of viral genetic diversity and the evolving molecular epidemiology of human immunodeficiency virus type 1. *J Virol* 2007; 81: 13050–13056. PMID: [17898057](https://pubmed.ncbi.nlm.nih.gov/17898057/)
19. de Oliveira T, Pillay D, Gifford RJ. The HIV-1 subtype C epidemic in South America is linked to the United Kingdom. *PLOS One* 2010; 5: e9311. doi: [10.1371/journal.pone.0009311](https://doi.org/10.1371/journal.pone.0009311) PMID: [20174561](https://pubmed.ncbi.nlm.nih.gov/20174561/)
20. Monno L, Brindicci G, Lai A, Punzi G, Altamura M, Simonetti FR, et al. An outbreak of HIV-1 BC recombinants in Southern Italy. *J Clin Virol* 2012; 55: 370–373. doi: [10.1016/j.jcv.2012.08.014](https://doi.org/10.1016/j.jcv.2012.08.014) PMID: [22981618](https://pubmed.ncbi.nlm.nih.gov/22981618/)
21. Thomson MM, Fernández-García A, Delgado E, Vega Y, Díez-Fuertes F, Sánchez-Martínez M, et al. Rapid expansion of a HIV-1 subtype F cluster of recent origin among men who have sex with men in Galicia, Spain. *J Acquir Immune Defic Syndr* 2012; 59: e49–e51. doi: [10.1097/QAI.0b013e3182400fc4](https://doi.org/10.1097/QAI.0b013e3182400fc4) PMID: [22327248](https://pubmed.ncbi.nlm.nih.gov/22327248/)
22. Delwart EL, Shpaer EG, Louwagie J, McCutchan FE, Grez M, Rubsamen-Waigmann H, et al. Genetic relationships determined by a DNA heteroduplex mobility assay: analysis of HIV-1 env genes. *Science* 1993; 262: 1257–1261. PMID: [8235655](https://pubmed.ncbi.nlm.nih.gov/8235655/)
23. Delgado E, Thomson MM, Villahermosa ML, Sierra M, Ocampo A, Miralles C, et al. Identification of a newly characterized HIV-1 BG intersubtype circulating recombinant form in Galicia, Spain, which exhibits a pseudotype-like virion structure. *J Acquir Immune Defic Syndr* 2002; 29: 536–543. PMID: [11981372](https://pubmed.ncbi.nlm.nih.gov/11981372/)
24. Sierra M, Thomson MM, Ríos M, Casado G, Ojea de Castro R, Delgado E, et al. The analysis of near full-length genome sequences of human immunodeficiency virus type 1 BF intersubtype recombinant viruses from Chile, Venezuela and Spain reveals their relationship to diverse lineages of recombinant viruses related to CRF12_BF. *Infect Genet Evol* 2005; 5: 209–217. PMID: [15737911](https://pubmed.ncbi.nlm.nih.gov/15737911/)
25. Delgado E, Fernández-García A, Vega Y, Cuevas T, Pinilla M, García V, et al. Evaluation of genotypic tropism prediction tests compared with in vitro co-receptor usage in HIV-1 primary isolates of diverse subtypes. *J Antimicrob Chemother* 2012; 67: 25–31. doi: [10.1093/jac/dkr438](https://doi.org/10.1093/jac/dkr438) PMID: [22010208](https://pubmed.ncbi.nlm.nih.gov/22010208/)
26. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 2013; 30: 772–780. doi: [10.1093/molbev/mst010](https://doi.org/10.1093/molbev/mst010) PMID: [23329690](https://pubmed.ncbi.nlm.nih.gov/23329690/)

27. Stamatakis A, Hoover P, Rougemont J. A rapid bootstrap algorithm for the RAxML Web servers. *Syst Biol* 2008; 57: 758–771. doi: [10.1080/10635150802429642](https://doi.org/10.1080/10635150802429642) PMID: [18853362](https://pubmed.ncbi.nlm.nih.gov/18853362/)
28. Lole KS, Bollinger RC, Paranjape RS, Gadhari D, Kulkarni SS, Novak NG, et al. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J Virol* 1999; 73: 152–160. PMID: [9847317](https://pubmed.ncbi.nlm.nih.gov/9847317/)
29. HIV Sequence Database 2014. Available: <http://www.hiv.lanl.gov/content/sequence/HIV/mainpage.html>. Accessed 28 November 2014.
30. Gifford RJ, Liu TF, Rhee SY, Kiuchi M, Hue S, Pillay D, et al. The calibrated population resistance tool: standardized genotypic estimation of transmitted HIV-1 drug resistance. *Bioinformatics* 2009; 25: 1197–1198. doi: [10.1093/bioinformatics/btp134](https://doi.org/10.1093/bioinformatics/btp134) PMID: [19304876](https://pubmed.ncbi.nlm.nih.gov/19304876/)
31. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol* 2012; 29: 1969–1973. doi: [10.1093/molbev/mss075](https://doi.org/10.1093/molbev/mss075) PMID: [22367748](https://pubmed.ncbi.nlm.nih.gov/22367748/)
32. Shapiro B, Rambaut A, Drummond AJ. Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. *Mol Biol Evol* 2006; 23: 7–9. PMID: [16177232](https://pubmed.ncbi.nlm.nih.gov/16177232/)
33. Drummond AJ, Rambaut A, Shapiro B, Pybus OG. Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol Biol Evol* 2005; 22: 1185–1192. PMID: [15703244](https://pubmed.ncbi.nlm.nih.gov/15703244/)
34. Yebra G, Delgado R, Pulido F, Rubio R, Galán JC, Moreno S, et al. Different trends of transmitted HIV-1 drug resistance in Madrid, Spain, among risk groups in the last decade. *Arch Virol* 2014; 159: 1079–1087. doi: [10.1007/s00705-013-1933-y](https://doi.org/10.1007/s00705-013-1933-y) PMID: [24297490](https://pubmed.ncbi.nlm.nih.gov/24297490/)
35. Castro E, Khonkarly M, Ciuffreda D, Burgisser P, Cavassini M, Yerly S, et al. HIV-1 drug resistance transmission networks in southwest Switzerland. *AIDS Res Hum Retroviruses* 2010; 26: 1233–1238. doi: [10.1089/aid.2010.0083](https://doi.org/10.1089/aid.2010.0083) PMID: [20860534](https://pubmed.ncbi.nlm.nih.gov/20860534/)
36. Chalmet K, Dauwe K, Foquet L, Baatz F, Seguin-Devaux C, Van der Gucht B, et al. Presence of CXCR4-using HIV-1 in patients with recently diagnosed infection: correlates and evidence for transmission. *J Infect Dis* 2012; 205: 174–184. doi: [10.1093/infdis/jir714](https://doi.org/10.1093/infdis/jir714) PMID: [22147802](https://pubmed.ncbi.nlm.nih.gov/22147802/)
37. Cardoso LP, Queiroz BB, Stefani MM. HIV-1 pol phylogenetic diversity and antiretroviral resistance mutations in treatment naive patients from Central West Brazil. *J Clin Virol* 2009; 46: 134–139. doi: [10.1016/j.jcv.2009.07.009](https://doi.org/10.1016/j.jcv.2009.07.009) PMID: [19682948](https://pubmed.ncbi.nlm.nih.gov/19682948/)
38. Ciccozzi M, Montieri S, Salemi M, De Oliveira T, Dorrucchi M, Sinicco A, et al. An outbreak of HIV-1 subtype G among Italian injecting drug users. *AIDS* 2007; 21: 1213–1215. PMID: [17502735](https://pubmed.ncbi.nlm.nih.gov/17502735/)
39. Hughes GJ, Fearhill E, Dunn D, Lycett SJ, Rambaut A, Leigh Brown AJ. Molecular phylogenetics of the heterosexual HIV epidemic in the United Kingdom. 2009; *PLOS Pathog* 5: e1000590. doi: [10.1371/journal.ppat.1000590](https://doi.org/10.1371/journal.ppat.1000590) PMID: [19779560](https://pubmed.ncbi.nlm.nih.gov/19779560/)
40. Skar H, Axelsson M, Berggren I, Thalme A, Gyllensten K, Liitsola K, et al. Dynamics of two separate but linked HIV-1 CRF01_AE outbreaks among injection drug users in Stockholm, Sweden, and Helsinki, Finland. *J Virol* 2011; 85: 510–518. doi: [10.1128/JVI.01413-10](https://doi.org/10.1128/JVI.01413-10) PMID: [20962100](https://pubmed.ncbi.nlm.nih.gov/20962100/)
41. Wertheim JO, Leigh Brown AJ, Hepler NL, Mehta SR, Richman DD, Smith DM, et al. The global transmission network of HIV-1. *J Infect Dis* 2014; 209: 304–313. doi: [10.1093/infdis/jit524](https://doi.org/10.1093/infdis/jit524) PMID: [24151309](https://pubmed.ncbi.nlm.nih.gov/24151309/)
42. Paraskevis D, Kostaki E, Beloukas A, Cañizares A, Aguilera A, Rodríguez J, et al. Molecular characterization of HIV-1 infection in Northwest Spain (2009–2013): Investigation of the subtype F outbreak. *Infect Genet Evol* 2014; 30: 96–101. doi: [10.1016/j.meegid.2014.12.012](https://doi.org/10.1016/j.meegid.2014.12.012) PMID: [25527396](https://pubmed.ncbi.nlm.nih.gov/25527396/)
43. Thomson MM, Fernández-García A. Phylogenetic structure in African HIV-1 subtype C revealed by selective sequential pruning. *Virology* 2011; 415: 30–38. doi: [10.1016/j.virol.2011.03.021](https://doi.org/10.1016/j.virol.2011.03.021) PMID: [21507449](https://pubmed.ncbi.nlm.nih.gov/21507449/)
44. Pernas B, Grandal M, Mena A, Castro-Iglesias A, Cañizares A, Wyles DL, et al. High prevalence of subtype F in newly diagnosed HIV-1 persons in northwest Spain and evidence for impaired treatment response. *AIDS* 2014; 28: 1837–1840. doi: [10.1097/QAD.0000000000000326](https://doi.org/10.1097/QAD.0000000000000326) PMID: [24871456](https://pubmed.ncbi.nlm.nih.gov/24871456/)
45. Soares EA, Martinez AM, Souza TM, Santos AF, Da Hora V, Silveira J, et al. HIV-1 subtype C dissemination in southern Brazil. *AIDS* 2005; 19(Suppl 4): S81–S86. PMID: [16249660](https://pubmed.ncbi.nlm.nih.gov/16249660/)
46. Pérez L, Thomson MM, Bleda MJ, Aragonés C, González Z, Pérez J, et al. HIV type 1 molecular epidemiology in Cuba: high genetic diversity, frequent mosaicism, and recent expansion of BG intersubtype recombinant forms. *AIDS Res Hum Retroviruses* 2006; 22: 724–733. PMID: [16910827](https://pubmed.ncbi.nlm.nih.gov/16910827/)
47. Paraskevis D, Magiorkinis E, Magiorkinis G, Sypsa V, Pappas V, Lazanas M, et al. Increasing prevalence of HIV-1 subtype A in Greece: estimating epidemic history and origin. *J Infect Dis* 2007; 196: 1167–1176. PMID: [17955435](https://pubmed.ncbi.nlm.nih.gov/17955435/)
48. Wang W, Xu J, Jiang S, Yang K, Meng Z, Ma Y, et al. The dynamic face of HIV-1 subtypes among men who have sex with men in Beijing, China. *Curr HIV Res* 2011; 9: 136–139. PMID: [21361866](https://pubmed.ncbi.nlm.nih.gov/21361866/)

49. Shcherbakova NS, Shalamova LA, Delgado E, Fernández-García A, Vega Y, Karpenko LI, et al. Molecular epidemiology, phylogeny, and phylodynamics of CRF63_02A1, a recently originated HIV-1 circulating recombinant form spreading in Siberia. *AIDS Res Hum Retroviruses* 2014; 30: 912–919. doi: [10.1089/AID.2014.0075](https://doi.org/10.1089/AID.2014.0075) PMID: [25050828](https://pubmed.ncbi.nlm.nih.gov/25050828/)
50. Thomson MM, Pérez-Álvarez L, Nájera R. Molecular epidemiology of HIV-1 genetic forms and its significance for vaccine development and therapy. *Lancet Infect Dis* 2002; 2: 461–471. PMID: [12150845](https://pubmed.ncbi.nlm.nih.gov/12150845/)
51. Hraber P, Korber BT, Lapedes AS, Bailer RT, Seaman MS, Gao H, et al. Impact of clade, geography and age of the epidemic on HIV-1 neutralization by antibodies. *J Virol* 2014; 88: 12623–12643. doi: [10.1128/JVI.01705-14](https://doi.org/10.1128/JVI.01705-14) PMID: [25142591](https://pubmed.ncbi.nlm.nih.gov/25142591/)
52. Phillips AN, Cambiano V, Nakagawa F, Brown AE, Lampe F, Rodger A, et al. Increased HIV incidence in men who have sex with men despite high levels of ART-induced viral suppression: analysis of an extensively documented epidemic. *PLOS One* 2013; 8: e55312. doi: [10.1371/journal.pone.0055312](https://doi.org/10.1371/journal.pone.0055312) PMID: [23457467](https://pubmed.ncbi.nlm.nih.gov/23457467/)