

1 **Transmission clusters, predominantly associated with men who have**
2 **sex with men, play a main role in the propagation of HIV-1 in**
3 **Northern Spain (2013-2018)**

4
5 Horacio Gil¹†*, Elena Delgado¹†, Sonia Benito¹, Leonidas Georgalis¹, Vanessa
6 Montero¹, Mónica Sánchez¹, Javier E. Cañada-García¹, Elena García-Bodas¹, Asunción
7 Díaz², Michael M. Thomson^{1*}, the Spanish Group for the Study of New HIV Diagnoses.

8
9 1. HIV Biology and Variability Unit. Centro Nacional de Microbiología, Instituto de
10 Salud Carlos III, Majadahonda, Madrid, Spain.

11 2. HIV Surveillance and Behavioral Monitoring Unit. Centro Nacional de Epidemiología,
12 Instituto de Salud Carlos III, Madrid, Spain.

13

14 †These authors have contributed equally to this work and share first authorship.

15 **Running title:** HIV-1 transmission clusters in Spain

16

17 **Keywords:** Spain, HIV-1, transmission clusters, molecular epidemiology, men who have
18 sex with men, migrants

19 **Word count:** 4207

20 **Tables:** 8 (2 of them supplementary)

21 *** Corresponding authors:**

22 Michael M. Thomson (email: mthomson@isciii.es; Phone: +34 91 822 3900)

23 Horacio Gil (email: hgil@isciii.es; Phone +34 91 822 3592)

24

25 HIV Biology and Variability Unit

26 Centro Nacional de Microbiología. Instituto de Salud Carlos III

27 28220 Majadahonda, Madrid, Spain.

28 Fax: +34 91 509 7966

29

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

1 **ABSTRACT**

2 Viruses of HIV-1-infected individuals whose transmission is related group
3 phylogenetically in transmission clusters (TCs). The study of the phylogenetic relations
4 of these viruses and the factors associated with these individuals is essential to analyze
5 the HIV-1 epidemic. In this study, we examine the role of TCs in the epidemiology of
6 HIV-1 infection in Galicia and the Basque County, two regions of Northern Spain. A
7 total of 1158 newly HIV-1-diagnosed patients (NDs) from both regions diagnosed in
8 2013-2018 were included in the study. Partial HIV-1 *pol* sequences were analyzed
9 phylogenetically by approximately-maximum-likelihood with FastTree 2. In this
10 analysis, 10,687 additional sequences from samples from HIV-1-infected individuals
11 collected in Spain in 1999-2019 were also included to assign and determine the TCs'
12 sizes. TCs were defined as those which included viruses from ≥ 4 individuals, at least
13 50% of them Spaniards, and with ≥ 0.95 Shimodaira-Hasegawa-like node support in the
14 phylogenetic tree. Factors associated to TCs were evaluated using odds ratios (OR) and
15 their 95% confidence intervals (CI). Fifty one percent of NDs grouped in 162 TCs. Male
16 patients (OR: 2.6; 95% CI: 1.5-4.7) and men having sex with men (MSM) (OR: 2.1; 95%
17 CI: 1.4-3.2) had higher odds of belonging to a TC compared to female and heterosexual
18 patients, respectively. Individuals from Latin America (OR: 0.3; 95% CI: 0.2-0.4), North
19 Africa (OR: 0.4; 95% CI: 0.2-1.0), and especially Sub-Saharan Africa (OR: 0.02; 95% CI:
20 0.003-0.2) were inversely associated to belonging to TCs compared to native
21 Spaniards. Differences in distribution and sizes of local TCs were observed in both
22 regions reflecting a different epidemic pattern. Our results show that TCs play an
23 important role in the spread of HIV-1 infection in the two Spanish regions studied,
24 where transmission between MSM is predominant. The majority of migrants were

- 1 infected with viruses not belonging to TCs that expand in Spain. Molecular
- 2 epidemiology is essential to identify local peculiarities of HIV-1 propagation. The early
- 3 detection of TCs and prevention of their expansion, implementing effective control
- 4 measures, could reduce HIV-1 infections.
- 5

1 INTRODUCTION

2 The HIV-1 epidemic is still one of the major public health problems in Spain. Around
3 3,500 - 4,000 new diagnoses of HIV-1 infection (NDs) are reported every year, with an
4 estimated incidence of NDs of 7.5 per 100,000 population in 2019 (HIV, STI and
5 hepatitis surveillance unit, 2020). A decreasing trend in the incidence of NDs has been
6 observed since 2010, although it is still higher than the average rate found in the
7 EU/EEA (European Centre for Disease Prevention and Control and WHO Regional
8 Office for Europe, 2020). In 2019, most (56%) of the NDs were diagnosed among men
9 who had sex with men (MSM) and 36% of all reported NDs were in people born
10 outside of Spain (HIV, STI and hepatitis surveillance unit, 2020).

11 Molecular epidemiology is an important tool for describing the HIV-1 epidemic (Board
12 et al., 2020; Brenner et al., 2013; Paraskevis et al., 2016; Wertheim et al., 2017).

13 Individuals whose transmission is related group phylogenetically in clades named
14 transmission clusters (TCs). Due to the high genetic variability of HIV-1, phylogenetic
15 analysis allows reconstructing transmission events through the identified TCs and infer
16 the history of the HIV-1 epidemic (Hué et al., 2004; Hué et al., 2005).

17 National databases of protease and reverse transcriptase (Pr-RT) sequences, primarily
18 obtained for antiretroviral drug resistance testing, contain valuable data about HIV-1
19 expansion and have been used in molecular epidemiology studies (Fabeni et al., 2021;
20 Oster et al., 2018; Parczewski et al., 2017; Petersen et al., 2018; Pineda-Peña et al.,
21 2019; Verhofstede et al., 2018). The phylogenetic analyses performed in these studies
22 combined with clinical and epidemiological data of the patients provide relevant public
23 health information for the implementation of control measures and for monitoring the

1 HIV-1 epidemic (Board et al., 2020; Brenner et al., 2013; Campbell et al., 2020;
2 Paraskevis et al., 2019; Vasylieva et al., 2016; Vasylieva et al., 2020).
3 Spanish clinical guidelines recommend to perform a genotypic drug resistance test
4 before starting antiretroviral therapy (ART) in all HIV-1-diagnosed patients and in ART-
5 failing cases (AIDS study group (GESIDA) of the Spanish Society of Infectious Diseases
6 and Clinical Microbiology and the national AIDS Plan., 2020). Pr-RT sequences obtained
7 for these tests have been analyzed in different molecular epidemiology studies,
8 describing the genetic features of the HIV-1 epidemic in different regions of Spain
9 (Cuevas et al., 2009a; Cuevas et al., 2009b; González-Alba et al., 2011; González-
10 Domenech et al., 2020; Holguín et al., 2007; Patiño-Galindo et al., 2016; Patiño-Galindo
11 et al., 2017b; Pérez-Parra et al., 2016; Pérez-Parra et al., 2015; Thomson et al., 2001;
12 Yebra et al., 2013). The phylogenetic studies have also allowed the identification of
13 large TCs among MSM which are actively growing in Spain (Delgado et al., 2015;
14 Delgado et al., 2019; González-Domenech et al., 2018; Patiño-Galindo et al., 2017a), as
15 well as transmitted drug resistance mutations which are spreading in TCs (Cuevas et
16 al., 2009b; González-Domenech et al., 2018; González-Domenech et al., 2020; Vega et
17 al., 2015; Viciano et al., 2016), highlighting the value of molecular epidemiology as a
18 tool for HIV surveillance.

19

20 The Basque Country and Galicia are two regions located in northern Spain, comparable
21 in terms of population and HIV-1 diagnose rates. During 2013-2018 the mean resident
22 population and rates of new HIV diagnoses were estimated in 2,2 million and 6,9 per
23 100,000 inhabitants for the Basque Country and 2,7 million and 5,6 per 100,000

1 inhabitants for Galicia (HIV, STI and hepatitis surveillance unit, 2020; Instituto Nacional
2 de Estadística, 2021). In this study, we analyze the role of TCs in the transmission of
3 HIV-1 infection among patients diagnosed in 2013-2018 in these northern Spanish
4 regions, identifying the different features associated with TCs in the HIV-1 epidemics in
5 these regions.

6

7 **MATERIAL AND METHODS**

8 **Patients**

9 Samples from individuals newly diagnosed of HIV-1 infection in the northern
10 Spanish regions of Galicia and Basque Country during 2013-2018, sent to the HIV Biology
11 and Variability Unit, Centro Nacional de Microbiología, Instituto de Salud Carlos III, were
12 included in the study.

13 The representativeness of the NDs of our cohort was estimated comparing them
14 to the NDs reported to the Spanish information system on new HIV diagnoses (SINIVIH
15 by its Spanish acronym)(HIV, STI and hepatitis surveillance unit, 2020), from the same
16 period and studied regions. As patients in both databases cannot be linked, gender, age
17 group, transmission route and region of birth frequencies were compared between both
18 groups of individuals using the chi-squared test.

19 **Nucleic acid extraction, amplification and sequencing**

20 Nucleic acid was extracted from plasma or whole blood samples. RNA was
21 extracted from 1 ml plasma using NUCLISENS® easyMAG® (BioMérieux, Marcy l'Etoile,
22 France) and DNA was extracted from 200 µl whole blood using QIAamp® DNA DSP

1 blood mini kit (Qiagen, Hilden, Germany), following the manufacturer's instructions. A
2 Pr-RT fragment of *pol* (HXB2 positions 2253-3629) was amplified by RT-PCR followed
3 by nested PCR from RNA or by nested PCR from DNA. Reagents, PCR thermal profiles,
4 and primers are described in supplementary tables 1 and 2.

5 Population sequencing was performed with ABI Prism BigDye Terminator Cycle
6 Sequencing kit and ABI 3730 XL sequencer (Applied Biosystems, Foster City, CA, U.S.A.)
7 in the Genomic Unit of Instituto de Salud Carlos III. Sequences were assembled with
8 SeqMan Pro v.12.2.1 (DNA STAR Lasergene, Madison, WI, USA) and edited with BioEdit
9 v.7.2.5 (Hall, 1999) (<http://www.mbio.ncsu.edu>).

10 **Phylogenetic analyses and transmission cluster identification**

11 Partial *pol* sequences from the patients included in this study and sequences
12 obtained by us from another 10,687 HIV-1-infected individuals attended in clinical
13 centers from Spain whose samples were collected in 1999-2019 were included in the
14 analyses. Sequences were analyzed phylogenetically by an approximately maximum-
15 likelihood method using FastTree2 (Price et al., 2010). In these analyses, the general
16 time reversible model of nucleotide substitution with CAT approximation to account
17 for among-site heterogeneity in substitution rates was used, and the reliability of
18 nodes was assessed with Shimodaira-Hasegawa(SH)-like local support values.
19 Classification of sequences in subtypes and recombinant forms was based on
20 clustering with clade references in approximately maximum-likelihood trees.
21 Sequences suspected of intersubtype recombination, were subsequently analyzed by
22 bootscanning with SimPlot v3.5 (Lole et al., 1999)

1 TCs were defined as those comprising viruses from four or more individuals, at
2 least 50% of them Spanish, and whose sequences grouped in the phylogenetic tree
3 with a SH-like node support value ≥ 0.95 .

4 **Statistical analysis**

5 To evaluate the association between variables, chi-squared and Fisher's exact
6 tests were used, with associations being considered statistically significant at a p -value
7 < 0.05 . A multivariable logistic regression model was performed to identify factors
8 associated to TCs. The model was adjusted by sex, age group, transmission route,
9 country of origin of the patient, Spanish region of sample collection, and HIV-1 genetic
10 form. Associations were measured using the odds ratio (OR) and its 95% confidence
11 interval (CI). Data analyses were performed using the STATA statistical software
12 package Version 16 (Stata Corporation, College Station, TX, US).

13

14 **RESULTS**

15 **Study Population**

16 A total of 1,158 HIV-1 NDs from Galicia and Basque Country diagnosed in 2013-
17 2018 were included in the study (Table 1). The majority of the patients were male
18 (82%), with MSM being the most common transmission route (46%). Heterosexuals
19 represented 34%, and 15% of the individuals were male whose transmission route was
20 reported as sexual without specifying whether they were in the MSM or heterosexual
21 category. Subtype B (63%) was the most frequent genetic form among the individuals

1 included in the study, followed by unique recombinant forms (URFs) (8.2%), CRF02_AG
2 (7.7%), and subtype F (7.6%) (Table 1).

3 Seventy percent of the patients were native Spanish (Table 1), followed by Latin
4 Americans (17%) and Sub-Saharan Africans (8.0%). There were important differences
5 in gender and transmission route proportions between these two migrant populations,
6 being higher in Sub-Saharan Africans than in Latin Americans: female [52% (44/85) vs.
7 26% (37/180)] and heterosexual transmission [88% (65/74) vs. 33%. (54/165)],
8 respectively.

9 The distribution of patients by gender and transmission route was similar in
10 both studied regions (Table 1), although the frequency of people who inject drugs
11 (PWID) was slightly higher in Galicia (5.3% vs. 3.2%). The percentage of patients of
12 non-Spanish origin was 35% in the Basque Country (13% corresponding to Africans), a
13 proportion which is statistically higher ($p < 0.001$) compared with the 17% found in
14 Galicia (2.2% Africans). Statistical differences in the distribution of genetic forms were
15 also observed between both regions ($p < 0.001$), with the frequencies of subtype F
16 (13%) and CRF02_AG (9.9%) being higher in Galicia and the Basque Country,
17 respectively (Table 1).

18 **Representativeness of the patients included in the study**

19 Our cohort of newly diagnosed HIV-1-infected patients represents 64% of the
20 notified cases in 2013-2018 (Table 2), assuming that all NDs included in our study have
21 been reported to the SINIVIH. This percentage was higher in the Basque Country,
22 where it reached 84%, than in Galicia, where it was 41% of the notified cases. No
23 statistical differences were found between both patient groups regarding the

1 distribution of individuals by gender, transmission route, region of origin or age group
2 (Table 2).

3 **TCs identified in the study**

4 Fifty one percent (594/1158) of the individuals belonged to 162 different TCs.
5 Regarding the regions, 45% (351/776) of the patients were included in 98 TCs in the
6 Basque Country and 64% (243/382) belonged to 82 TCs in Galicia (Table 3). Eighteen
7 (11%) TCs comprised patients from both regions. The sizes of the TCs ranged from 4 to
8 204 patients. MSM was the main transmission route associated with 59% of the TCs
9 (Table 3). Transmission routes other than MSM were more frequently associated with
10 TCs in Galicia (44%) than in the Basque Country (35%) (Table 3), although the difference
11 was not statistically significant ($p=0.207$).

12 A total of 15 large TCs (here defined as those with ≥ 30 individuals) were
13 identified in the study (Table 4), 10 of them of subtype B. MSM was the associated
14 transmission route in 14 (93%) of them. Most of these large TCs have spread mainly in
15 a specific region. Thus, TCs B07, B08, B09, B10, B12, B70, and F1_3 were found mainly
16 in the Basque Country, while TCs F1_1, B13, BG_2, and B05 were spread mainly in
17 Galicia (Table 4). In addition, A1_1, CRF02_1, B50 and B31 have a global expansion in
18 Spain, with a high number of patients outside of both studied regions. Four TCs had
19 greater than 50% increase in NDs during the period 2013-2018: F1_3 (94%), A1_1
20 (76%), B70 (72%) and CRF02_1 (53%) (Table 4).

21 **Factors associated with TCs**

1 The percentage of patients in TCs was higher in males (57%) than in females
2 (22%) and in MSM (67%) compared to heterosexuals (31%). This percentage was also
3 higher in Spaniards (64%) than in Latin Americans (35%) or in Sub-Saharan Africans
4 (1.2%) and in individuals infected with viruses of subtypes F (76%) or B (58%)
5 compared to those infected with CRF02_AG viruses (17%). The different distribution of
6 TCs according to gender, mode of HIV-1 transmission, country of origin of the patient
7 and HIV-1 genetic form were statistically significant (Table 5). The age group was also
8 associated to TCs ($p=0.036$). The age range of 20-29 years showed an increased
9 percentage of patients in TCs (59%), which was especially high in Galicia (81%,
10 $p=0.003$) but was not statistical significant in the Basque Country, and no statistical
11 differences were found in the distribution of main variables

12 In the multivariate analysis, males (OR: 2.6; 95% CI: 1.5-4.7) and MSM (OR: 2.1;
13 95% CI: 1.4-3.2) had higher odds of belonging to a TC compared to females and
14 heterosexuals, respectively (Table 6). Patients from Galicia (OR: 1.5; 95% CI: 1.1-2.2)
15 and subtype F infections (OR: 2.5; 95% CI: 1.3-5.0) were also factors associated with
16 TCs (Table 6)

17 Individuals from Latin America (OR: 0.3; 95% CI: 0.2-0.4), North Africa (OR: 0.4;
18 95% CI: 0.2-1.0) and especially from Sub-Saharan Africa (OR: 0.02; 95% CI: 0.003-0.2)
19 were inversely associated to belonging to TCs compared to Spanish patients (Table 6).
20 Similarly, patients over 29 years old (OR: 0.6; 95% CI: 0.4-0.9) and infections with a
21 virus of a CRF_BF (OR: 0.1; 95% CI: 0.03-0.4) or of a genetic form classified in the
22 “other” category (OR 0.3 95% CI: 0.2-0.8) had lower odds of belonging to TCs than

1 patients in the age range of 20-29 years and with subtype B infections, respectively
2 (Table 6).

3 **DISCUSSION**

4 The phylogenetic analysis of almost 12,000 HIV-1-infected patients from our database
5 has allowed us to investigate the transmission events occurring during the 2013-2018
6 period in two regions of Northern Spain, Galicia and Basque Country, identifying TCs
7 and their associated factors. Our ND cohort had a good representativeness,
8 corresponding to 64% of the NDs reported to the SINIVIH from these regions during
9 that period, especially in the Basque Country where it reached 84%.

10 Our analysis has assigned 51% of the NDs from Galicia and Basque Country to 162
11 different TCs, indicating that TCs are playing an important role in the spread of HIV-1
12 infections in both regions. There is no consensus on the criteria for defining a TC, as
13 these should depend on the aim of the study (Hassan et al., 2017). We considered only
14 TCs including at least 4 documented infections. Therefore, our more strict TC
15 definition, compared to some other studies, has allowed us to identify the structure
16 and factors associated to HIV-1 strains whose transmission is established in Spain.
17 TCs were found associated with male gender and MSM. In fact, almost all the largest
18 TCs (≥ 30 patients) identified in our study were associated with MSM. HIV-1 molecular
19 epidemiological studies in different countries, independently of the methodology used,
20 have found a similar association (Duran-Ramirez et al., 2021; Parczewski et al., 2017;
21 Wertheim et al., 2017). Also, studies in other Spanish regions have found this
22 association of TCs with MSM (González-Domenech et al., 2020; Pérez-Parra et al.,
23 2015). These findings indicate that MSM TCs are playing a main role in spreading HIV-1

1 infections and are consistent with the fact that MSM is the main frequent mode of
2 HIV-1 transmission in Spain since the early 2000 (Nuñez et al., 2018).

3 Spaniards showed higher odds than Latin Americans or Africans of belonging to TCs.
4 However, there were important differences between these migrant populations, as
5 35% of Latin Americans were included in TCs, compared to only 1.2% of Sub-Saharan
6 Africans, suggesting that HIV-1 transmission between Latin Americans and Spaniards is
7 relatively common, presumably derived from cultural and linguistic affinities. Similarly
8 to our results, a low percentage of Sub-Saharan Africans in TCs have been also
9 reported in Italy (Fabeni et al., 2020) and Belgium (Verhofstede et al., 2018). In our
10 study, TCs are mainly MSM-driven. In contrast, we have found high frequencies of
11 females (52%) and heterosexuals (88%) in Sub-Saharan Africans, which, in addition to
12 the likely HIV-1 acquisition in their countries of origin, can explain the low proportion
13 of individuals in this migrant population belonging to TCs. Indeed, we have observed
14 that Sub-Saharan Africans are infected frequently with genetic forms circulating in
15 Africa.

16 In Spain, 36% of HIV-1 infections diagnosed in 2019 were in migrants (HIV, STI and
17 hepatitis surveillance unit, 2020). Determining the origin of HIV-1 infections in the
18 migrant population would allow proper allocation of resources for preventive
19 measures (Álvarez del Arco et al., 2017; Fakoya et al., 2015). A previous study based on
20 models using clinical and behavioral data found a high level of post-migration HIV
21 acquisition, up to 71% of the migrants in Spain (Álvarez del Arco et al., 2017). However,
22 phylogenetic analyses can provide additional information on the geographical area of
23 HIV-1 acquisition, decreasing the bias due to data collection or failure to consider the

1 infections acquired by migrants when visiting their countries (Álvarez del Arco et al.,
2 2017), as migrants' mobility was associated with increased risk for HIV acquisition (Dias
3 et al., 2020). In our study, the lower percentage of foreigners found in TCs could be
4 due to infections with virus strains which are not circulating in Spain, suggestive of
5 imported infections, which would be especially frequent in Sub-Saharan Africans.
6 Further studies using a combined phylogenetic and epidemiological approach can
7 address the accuracy of the origin of the HIV-1 infections in the migrant populations in
8 Spain.

9 Subtype B was the most frequent genetic form associated to TCs. Only subtype F
10 infections (76% of NDs in TCs) showed a higher probability than subtype B infections
11 (56% of NDs in TCs) of belonging to TCs. This was mainly due to the presence of two
12 large TCs of F1 subsubtype, designated F1_1 (205 individuals) and F1_3 (36
13 individuals), which have spread among MSM mainly in Galicia and Basque Country,
14 respectively (Delgado et al., 2015; Delgado et al., 2019). Interestingly, two other large
15 non-subtype B TCs associated with MSM, of CRF02_AG and A1 subsubtype, are also
16 present in these regions (Delgado et al., 2019). In Western Europe, the HIV-1 epidemic
17 among MSM is dominated by subtype B. However, the epidemic among MSM in Spain
18 is becoming increasingly diverse through the expansion of multiple non-subtype B TCs,
19 comprising or related to viruses circulating in other countries (Delgado et al., 2019).
20 This phenomenon has been also documented in other European countries (Dauwe et
21 al., 2015; Duran Ramirez et al., 2021; Fabeni et al., 2019; Ragonnet-Cronin et al., 2016;
22 Verhofstede et al., 2018), where viruses introduced from abroad have expanded
23 successfully among the MSM population.

1 A different distribution of TCs is responsible for the spread of HIV-1 in Galicia and the
2 Basque Country. Large MSM TCs have spread locally in both regions and the majority
3 of the identified TCs were associated with this population group. However, we found a
4 higher frequency of TCs associated with PWID or heterosexual transmission in Galicia.
5 Galician patients have 1.5 higher odds of belonging to TCs than individuals from the
6 Basque Country. This can be related to the low percentage of migrants in Galicia, who
7 have a reduced association to TCs, and the high number of patients belonging to the
8 large F1_1 TC, currently comprising 205 individuals, which has successfully spread in
9 Galicia and other Spanish regions (Delgado et al., 2015), even in other European
10 countries (Delgado et al., 2015; Vinken et al., 2019). We have found an association of
11 young patients (age range 20-29 years) to TCs in Galicia, similarly to the findings in
12 other studies (Brenner et al., 2017; Dennis et al., 2018; Fabeni et al., 2019; González-
13 Domenech et al., 2020; Paraskevis et al., 2019; Verhofstede et al., 2018), but not in the
14 Basque Country. Public health interventions for reducing high risk behavior and control
15 measures in this age group in Galicia should be implemented to decrease HIV-1
16 transmissions in the region. Moreover, our results highlight that even regions with
17 comparable demographic features can show a different HIV-1 epidemic pattern, likely
18 due to specific social and cultural differences at regional level.

19 In this study we have shown that MSM TCs are a keystone of the HIV-1 epidemic at
20 regional and country levels. Consequently, reinforcement of public health measures
21 for preventing HIV-1 infections in MSM, such as behavioral interventions to reduce
22 risky practices, pre-exposure prophylaxis (Grant et al., 2010; McCormack et al., 2016;
23 Riddell et al., 2018), and early diagnosis and treatment of HIV-1 infections (European
24 Centre for Disease Prevention and Control, 2015; United Nations Population Fund,

1 Global Forum on MSM & HIV, United Nations Development Programme, World Health
2 Organization, United States Agency for International Development, World Bank,
3 2015) are recommended. In addition, new strategies and other tools for reaching the
4 MSM population, such as the apps for sex dating suggested in other studies (Duncan et
5 al., 2018; Wirden et al., 2019) could be considered for preventing HIV-1 infections.

6 Implementation of measures targeting MSM could reduce indirectly the spread of HIV-
7 1, which is mainly MSM-driven, in other population groups. National surveillance data
8 in the USA suggest that infections among heterosexual women predominantly
9 originate from MSM (Oster et al., 2015). Moreover, the presence of female and self-
10 declared heterosexual male individuals in MSM TCs is frequent (Esbjörnsson et al.,
11 2016; Hué et al., 2014; Ragonnet-Cronin et al., 2018; Verhofstede et al., 2018), as we
12 have also observed in large TCs in Spain (Delgado et al., 2019). Analyses to evaluate
13 the contribution of MSM in HIV-1 transmission to other population groups are needed
14 in Spain.

15 The high representativeness of our cohort, especially in the Basque Country, suggests
16 that the largest and most spread TCs in these regions have been identified, which
17 supports the conclusions of the study. However, our data may have some limitations
18 as some of the NDs from Basque Country and Galicia during the studied period were
19 not analyzed and not all the TCs propagating in Spain were detected in our
20 phylogenetic analysis.

21 Our study has provided valuable data on the HIV-1 epidemic in Spain, identifying
22 specific features at geographical and population levels for tailoring more specific public
23 health interventions. New approaches and strategies are needed to reach the UNAIDS

1 90-90-90 target proposed by 2020 (UNAIDS, 2014). To achieve this goal, the
2 implementation of a molecular HIV-1 surveillance system to monitor the evolution of
3 the epidemic in Spain could be of great help. Such a surveillance system could allow to
4 promptly detect rapidly expanding TCs needing urgent investigation and the
5 implementation of additional public health measures to prevent their spread (Centers
6 for Disease Control and Prevention, 2018), as well as to readjust the control measures
7 in place according to the evolution of the TCs.

8

9 **MEMBERS OF THE SPANISH GROUP FOR THE STUDY OF NEW HIV DIAGNOSES**

10 **Basque Country:** Hospital Universitario de Araba, Vitoria: Andrés Canut-Blasco, José
11 Joaquín Portu, Carmen Gómez-González; Hospital Universitario de Basurto, Bilbao:
12 Josefa Muñoz, M^a Carmen Nieto, María Zuriñe Zubero, Silvia Hernández-Crespo, Estibaliz
13 Ugalde; Hospital Universitario de Cruces, Barakaldo: Luis Elorduy, Elena Bereciartua,
14 Leyre López Soria; Hospital de Galdakao: M^a José López de Goicoechea, José Mayo;
15 Hospital Universitario Donostia: Gustavo Cilla, Julio Arrizabalaga, José Antonio
16 Iribarren, M^a Jesús Bustinduy, M^a Julia Echevarría. M^a Yolanda Salicio, David
17 Grandioso. **Galicia:** Área sanitaria de Ferrol: Ana Mariño, Patricia Ordóñez, Hortensia
18 Álvarez, Nieves Valcarce; Complejo Hospitalario Universitario de A Coruña: Ángeles
19 Cañizares, M^a Ángeles Castro. Hospital Universitario Lucus Augusti, Lugo: Ramón
20 Rabuñal-Rey, María José García-País, M^a José Gude-González, Pilar Alonso-García,
21 Antonio Moreno-Flores; Complejo Hospitalario Universitario de Ourense: Juan García
22 Costa, Ricardo Fernández-Rodríguez, Raúl Rodríguez-Pérez, Jorge Guitián, María
23 Dolores Díaz-López, María Genoveva Naval-Calviño; Complejo Hospitalario

- 1 Universitario de Vigo: Celia Miralles, Antonio Ocampo, Sonia Pérez-Castro, Jorge Julio
- 2 Cabrera; Complejo Hospitalario Universitario de Pontevedra: Julio Díz-Arén, Matilde
- 3 Trigo, M^a Ángeles Pallarés. **Navarra:** Complejo Hospitalario de Navarra, Pamplona:
- 4 Carmen Ezpeleta Baquedano, Carmen Martín Salas, Irati Arregui García, María Gracia
- 5 Ruiz de Alda. **Madrid:** Centro Sanitario Sandoval, Madrid: Jorge del Romero, Carmen
- 6 Rodríguez, Mar Vera, Óskar Ayerdi, Eva Orviz; Hospital Universitario de Fuenlabrada:
- 7 María Isabel García-Arata, Santiago Prieto-Menchero; Hospital Clínico Universitario San
- 8 Carlos, Madrid: Esther Culebras, Icíar Rodríguez-Avial; Hospital Universitario Fundación
- 9 Jiménez Díaz, Madrid: Raquel Téllez-Pérez, Olalla Calabia-González, Alfonso Cabello-
- 10 Úbeda, Miguel Górgolas Hernández-Mora; Hospital Universitario Severo Ochoa,
- 11 Leganés: Sara María Quevedo, Lucía Puente, Manuel del Álamo; Hospital Fundación
- 12 Alcorcón, Madrid: Carolina Campelo Gutiérrez, María José Goyanes Galán; **Castilla y**
- 13 **León:** Hospital Clínico Universitario de Valladolid: Carmen Hinojosa, Carlos Dueñas,
- 14 Begoña Monteagudo, Edita Sánchez; Hospital Río Hortega, Valladolid: Jessica Abadía,
- 15 Belén Lorenzo Vidal; Hospital Virgen de la Concha, Zamora: Teresa Martín-Domínguez,
- 16 Rosa Martínez-González. **La Rioja:** Hospital San Pedro, Logroño: José Ramón Blanco,
- 17 Miriam Blasco. **Aragón:** Hospital Universitario Miguel Servet, Zaragoza: Ana María
- 18 Martínez-Sapiña, Piedad Arazo. **Castilla-La Mancha:** Hospital Universitario de
- 19 Guadalajara: Alejandro González Praetorius, Complejo Hospitalario de Toledo: César
- 20 Gómez-Hernando, José Largo-Pau; **Comunitat Valenciana:** Hospital Universitari Sant
- 21 Joan d'Alacant: Fernando Buñuel, Ana Infante.

22 **FUNDING**

1 This work was funded through Acción Estratégica en Salud Intramural (AESI), Instituto
2 de Salud Carlos III, Project “Estudios sobre vigilancia epidemiológica molecular del VIH-
3 1 en España,” PI16CIII/00033 and Project “Epidemiología molecular del VIH-1 en
4 España y su utilidad para investigaciones biológicas y en vacunas” PI19CIII/0042; Red
5 de Investigación en SIDA (RIS), Instituto de Salud Carlos III, Subdirección General de
6 Evaluación y Fondo Europeo de Desarrollo Regional (FEDER), Plan Nacional ICDCI,
7 project RD16ISCIII/0002/0004; and scientific agreements with Consellería de Sanidade,
8 Government of Galicia (MVI 1004/16) and Osakidetza-Servicio Vasco de Salud,
9 Government of Basque Country (MVI 1001/16).

10 **ETHICS STATEMENT**

11 Sequences were derived from antiretroviral resistance genotypic test. The use of
12 anonymized, de-identified clinical/demographic and sequence data was reviewed and
13 approved under an exempt protocol by the Bioethics and Animal Well-being
14 Committee of Instituto de Salud Carlos III, with report numbers CEI PI 38_2016-v3
15 (dated 20 June 2016) and CEI PI 31_2019-v5 (dated 6 November 2019). This study did
16 not require written informed consent by study participants, except for those
17 participants who required additional samples different from the ones obtained in the
18 routine clinical practice.

19

20 **AUTHOR CONTRIBUTIONS**

21 HG, ED, and MT conceived the study and supervised the experimental work. ED, HG,
22 MT, and JC processed sequences and performed phylogenetic analyses. HG and ED
23 performed data curation. HG, AD, and LG performed statistical analyses. SB, MS, VM

1 and EG-B performed experimental work. The members of the Spanish Group for the
2 Study of New HIV Diagnoses recruited patients and obtained epidemiological and
3 clinical data. HG wrote the manuscript draft, with contributions to the text by AD, MT
4 and ED. All authors read and approved the text.

5

6 **ACKNOWLEDGEMENTS**

7 We would like to thank José Antonio Taboada, from Consellería de Sanidade, Xunta de
8 Galicia, and Daniel Zulaika, from Osakidetza-Servicio Vasco de Salud, for their support
9 of this study, and the personnel at the Genomic Unit, Centro Nacional de
10 Microbiología, Instituto de Salud Carlos III, for technical assistance in sequencing.

11

12

REFERENCES

- 1
- 2 AIDS study group (GESIDA) of the Spanish Society of Infectious Diseases and Clinical
3 Microbiology and the national AIDS Plan. (2020). *Consensus Document on*
4 *antiretroviral therapy in adults infected by the Human Immunodeficiency Virus*
5 *(Updated 2020)*. Available at: [https://gesida-seimc.org/wp-](https://gesida-seimc.org/wp-content/uploads/2020/07/TAR_GUIA_GESIDA_2020_COMPLETA_Julio.pdf)
6 [content/uploads/2020/07/TAR_GUIA_GESIDA_2020_COMPLETA_Julio.pdf](https://gesida-seimc.org/wp-content/uploads/2020/07/TAR_GUIA_GESIDA_2020_COMPLETA_Julio.pdf)
- 7 Álvarez del Arco, D., Fakoya, I., Thomadakis, C., Pantazis, N., Touloumi, G., Gennotte, A.F. et al.
8 (2017). High levels of postmigration HIV acquisition within nine European countries.
9 *AIDS* 31, 1979-1988. doi: 10.1097/QAD.0000000000001571
- 10 Board, A.R., Oster, A.M., Song, R., Gant, Z., Linley, L., Watson, M. et al. (2020). Geographic
11 distribution of HIV transmission networks in the United States. *J. Acquir. Immune.*
12 *Defic. Syndr.* 85, e32-e40. doi: 10.1097/QAI.0000000000002448
- 13 Brenner, B., Wainberg, M.A. and Roger, M. (2013). Phylogenetic inferences on HIV-1
14 transmission: implications for the design of prevention and treatment interventions.
15 *AIDS* 27, 1045-1057. doi: 10.1097/QAD.0b013e32835cfd9
- 16 Brenner, B.G., Ibanescu, R.I., Hardy, I., Stephens, D., Otis, J., Moodie, E. et al. (2017). Large
17 cluster outbreaks sustain the HIV epidemic among MSM in Quebec. *AIDS* 31, 707-717.
18 doi: 10.1097/QAD.0000000000001383
- 19 Campbell, E.M., Patala, A., Shankar, A., Li, J.F., Johnson, J.A., Westheimer, E. et al. (2020).
20 Phylodynamic analysis complements partner services by identifying acute and
21 unreported HIV transmission. *Viruses*. 12. doi: 10.3390/v12020145
- 22 Centers for Disease Control and Prevention. (2018). *Detecting, investigating, and responding to*
23 *HIV transmission clusters*. Available at:
24 [https://www.cdc.gov/hiv/pdf/funding/announcements/ps18-1802/CDC-HIV-PS18-](https://www.cdc.gov/hiv/pdf/funding/announcements/ps18-1802/CDC-HIV-PS18-1802-AttachmentE-Detecting-Investigating-and-Responding-to-HIV-Transmission-Clusters.pdf)
25 [1802-AttachmentE-Detecting-Investigating-and-Responding-to-HIV-Transmission-](https://www.cdc.gov/hiv/pdf/funding/announcements/ps18-1802/CDC-HIV-PS18-1802-AttachmentE-Detecting-Investigating-and-Responding-to-HIV-Transmission-Clusters.pdf)
26 [Clusters.pdf](https://www.cdc.gov/hiv/pdf/funding/announcements/ps18-1802/CDC-HIV-PS18-1802-AttachmentE-Detecting-Investigating-and-Responding-to-HIV-Transmission-Clusters.pdf)
- 27 Cuevas, M.T., Fernández-García, A., Sánchez-García, A., González-Galeano, M., Pinilla, M.,
28 Sánchez-Martínez, M. et al. (2009a). Incidence of non-B subtypes of HIV-1 in Galicia,
29 Spain: high frequency and diversity of HIV-1 among men who have sex with men. *Euro.*
30 *Surveill.* 14. doi: 10.2807/ese.14.47.19413-en
- 31 Cuevas, M.T., Muñoz-Nieto, M., Thomson, M.M., Delgado, E., Iribarren, J.A., Cilla, G. et al.
32 (2009b). HIV-1 transmission cluster with T215D revertant mutation among newly
33 diagnosed patients from the Basque Country, Spain. *J. Acquir. Immune. Defic. Syndr.*
34 51, 99-103. doi: 10.1097/QAI.0b013e318199063e
- 35 Dauwe, K., Mortier, V., Schauvliege, M., Van Den Heuvel, A., Franssen, K., Servais, J.Y. et al.
36 (2015). Characteristics and spread to the native population of HIV-1 non-B subtypes in
37 two European countries with high migration rate. *BMC. Infect. Dis.* 15, 524. doi:
38 110.1186/s12879-015-1217-0
- 39 Delgado, E., Benito, S., Montero, V., Cuevas, M.T., Fernández-García, A., Sánchez-Martínez, M.
40 et al. (2019). Diverse large HIV-1 non-subtype B clusters are spreading among men
41 who have sex with men in Spain. *Front. Microbiol.* 10, 655. doi:
42 10.3389/fmicb.2019.00655

- 1 Delgado, E., Cuevas, M.T., Domínguez, F., Vega, Y., Cabello, M., Fernández-García, A. et al.
2 (2015). Phylogeny and phylogeography of a recent HIV-1 subtype F outbreak among
3 men who have sex with men in Spain deriving from a cluster with a wide geographic
4 circulation in Western Europe. *PLoS. One.* 10, e0143325. doi:
5 10.1371/journal.pone.0143325
- 6 Dennis, A.M., Volz, E., Frost, A.S.M.S., Hossain, M., Poon, A.F.Y., Rebeiro, P.F. et al. (2018). HIV-
7 1 transmission clustering and phylodynamics highlight the important role of young
8 men who have sex with men. *AIDS Res. Hum. Retroviruses* 34, 879-888. doi:
9 10.1089/AID.2018.0039
- 10 Dias, S., Gama, A., Loos, J., Roxo, L., Simões, D. and Nöstlinger, C. (2020). The role of mobility in
11 sexual risk behaviour and HIV acquisition among sub-Saharan African migrants residing
12 in two European cities. *PLoS. One.* 15, e0228584. doi: 10.1371/journal.pone.0228584
- 13 Duncan, D.T., Park, S.H., Hambrick, H.R., Dangerfield li, D.T., Goedel, W.C., Brewer, R. et al.
14 (2018). Characterizing geosocial-networking app use among young black men who
15 have sex with men: A multi-city cross-sectional survey in the southern United States.
16 *JMIR. Mhealth. Uhealth.* 6, e10316. doi: 10.2196/10316
- 17 Duran-Ramirez, J.J., Ballouz, T., Nguyen, H., Kusejko, K., Chaudron, S.E., Huber, M. et al. (2021).
18 Increasing frequency and transmission of HIV-1 non-B subtypes among men who have
19 sex with men in the Swiss HIV cohort study. *J. Infect. Dis.* doi: 10.1093/infdis/jiab360
- 20 Esbjörnsson, J., Mild, M., Audelin, A., Fonager, J., Skar, H., Bruun-Jørgensen. et al. (2016). HIV-1
21 transmission between MSM and heterosexuals, and increasing proportions of
22 circulating recombinant forms in the Nordic Countries. *Virus Evol.* 2, vew010. doi:
23 10.1093/ve/vew010
- 24 European Centre for Disease Prevention and Control. (2015). *ECDC Guidance: HIV and STI*
25 *prevention among men who have sex with men.* Available at:
26 [https://ecdc.europa.eu/sites/portal/files/media/en/publications/Publications/hiv-sti-](https://ecdc.europa.eu/sites/portal/files/media/en/publications/Publications/hiv-sti-prevention-among-men-who-have-sex-with-men-guidance.pdf)
27 [prevention-among-men-who-have-sex-with-men-guidance.pdf](https://ecdc.europa.eu/sites/portal/files/media/en/publications/Publications/hiv-sti-prevention-among-men-who-have-sex-with-men-guidance.pdf)
- 28 European Centre for Disease Prevention and Control and WHO Regional Office for Europe.
29 (2020). *HIV/AIDS surveillance in Europe 2020 – 2019 data.* Available at:
30 [https://www.ecdc.europa.eu/sites/default/files/documents/hiv-surveillance-report-](https://www.ecdc.europa.eu/sites/default/files/documents/hiv-surveillance-report-2020.pdf)
31 [2020.pdf](https://www.ecdc.europa.eu/sites/default/files/documents/hiv-surveillance-report-2020.pdf)
- 32 Fabeni, L., Alteri, C., Berno, G., Scutari, R., Orchi, N., De, C.G. et al. (2019). Characterisation of
33 HIV-1 molecular transmission clusters among newly diagnosed individuals infected
34 with non-B subtypes in Italy. *Sex Transm. Infect.* 95, 619-625. doi: 10.1136/sextrans-
35 2019-054017
- 36 Fabeni, L., Rozera, G., Berno, G., Giombini, E., Gori, C., Orchi, N. et al. (2021). Molecular
37 transmission dynamics of primary HIV infections in Lazio region, Years 2013-2020.
38 *Viruses.* 13. doi: 10.3390/v13020176
- 39 Fabeni, L., Santoro, M.M., Lorenzini, P., Rusconi, S., Gianotti, N., Costantini, A. et al. (2020).
40 Evaluation of HIV transmission clusters among natives and foreigners living in Italy.
41 *Viruses.* 12. doi: 10.3390/v12080791

- 1 Fakoya, I., Álvarez del Arco, D., Woode-Owusu, M., Monge, S., Rivero-Montesdeoca, Y.,
2 Delpech, V. et al. (2015). A systematic review of post-migration acquisition of HIV
3 among migrants from countries with generalised HIV epidemics living in Europe:
4 implications for effectively managing HIV prevention programmes and policy. *BMC*
5 *Public Health* 15, 561. doi: 10.1186/s12889-015-1852-9
- 6 González-Alba, J.M., Holguín, A., García, R., García-Bujalance, S., Alonso, R., Suárez, A. et al.
7 (2011). Molecular surveillance of HIV-1 in Madrid, Spain: a phylogeographic analysis. *J.*
8 *Virol.* 85, 10755-10763. doi: 10.1128/JVI.00454-11
- 9 González-Domenech, C.M., Sena-Corrales, G., Viciano-Ramos, I., Palacios-Muñoz, R., Mora-
10 Navas, L., Clavijo-Frutos, E. et al. (2020). High prevalence of sequences included in
11 transmission clusters within newly diagnosed HIV-1 patients in southern Spain (2004-
12 2015). *Microb. Drug Resist.* doi: 10.1089/mdr.2019.0344
- 13 González-Domenech, C.M., Viciano, I., Delaye, L., Mayorga, M.L., Palacios, R., de la Torre, J. et
14 al. (2018). Emergence as an outbreak of the HIV-1 CRF19_cpx variant in treatment-
15 naïve patients in southern Spain. *PLoS. One.* 13, e0190544. doi:
16 10.1371/journal.pone.0190544
- 17 Grant, R.M., Lama, J.R., Anderson, P.L., McMahan, V., Liu, A.Y., Vargas, L. et al. (2010).
18 Preexposure chemoprophylaxis for HIV prevention in men who have sex with men. *N.*
19 *Engl. J. Med.* 363, 2587-2599. doi: 10.1056/NEJMoa1011205
- 20 Hall, T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis
21 program for Windows 95/98/NT. *Nucl. Acids. Symp. Ser.* 41, 95-98. doi:
- 22 Hassan, A.S., Pybus, O.G., Sanders, E.J., Albert, J. and Esbjornsson, J. (2017). Defining HIV-1
23 transmission clusters based on sequence data. *AIDS* 31, 1211-1222. doi:
24 10.1097/QAD.0000000000001470
- 25 HIV, STI and hepatitis surveillance unit. (2020). *Vigilancia del VIH y sida en España: Sistema de*
26 *información sobre nuevos diagnósticos de VIH y registro nacional de casos de sida. Plan*
27 *Nacional sobre el Sida.* Available at:
28 [https://www.msbs.gob.es/ca//ciudadanos/enfLesiones/enfTransmisibles/sida/vigilan](https://www.msbs.gob.es/ca//ciudadanos/enfLesiones/enfTransmisibles/sida/vigilancia/Informe_VIH_SIDA_20201130.pdf)
29 [cia/Informe_VIH_SIDA_20201130.pdf](https://www.msbs.gob.es/ca//ciudadanos/enfLesiones/enfTransmisibles/sida/vigilancia/Informe_VIH_SIDA_20201130.pdf)
- 30 Holguín, A., Pena, M.J., Troncoso, F. and Soriano, V. (2007). Introduction of non-B subtypes
31 among Spaniards newly diagnosed with HIV type 1 in the Canary Islands. *AIDS Res.*
32 *Hum. Retroviruses* 23, 498-502. doi: 10.1089/aid.2006.0191
- 33 Hué, S., Brown, A.E., Ragonnet-Cronin, M., Lycett, S.J., Dunn, D.T., Fearnhill, E. et al. (2014).
34 Phylogenetic analyses reveal HIV-1 infections between men misclassified as
35 heterosexual transmissions. *AIDS* 28, 1967-1975. doi:
36 10.1097/QAD.0000000000000383
- 37 Hué, S., Clewley, J.P., Cane, P.A. and Pillay, D. (2004). HIV-1 pol gene variation is sufficient for
38 reconstruction of transmissions in the era of antiretroviral therapy. *AIDS* 18, 719-728.
39 doi: 10.1097/00002030-200403260-00002
- 40 Hué, S., Pillay, D., Clewley, J.P. and Pybus, O.G. (2005). Genetic analysis reveals the complex
41 structure of HIV-1 transmission within defined risk groups. *Proc. Natl. Acad. Sci. U. S. A*
42 102, 4425-4429. doi: 10.1073/pnas.0407534102

- 1 Instituto Nacional de Estadística. (2021). *Población residente*. Available at:
2 <https://www.ine.es/jaxiT3/Datos.htm?t=9681>
- 3 Lole, K.S., Bollinger, R.C., Paranjape, R.S., Gadkari, D., Kulkarni, S.S., Novak, N.G. et al. (1999).
4 Full-length human immunodeficiency virus type 1 genomes from subtype C-infected
5 seroconverters in India, with evidence of intersubtype recombination. *J. Virol.* 73, 152-
6 160. doi: 10.1128/JVI.73.1.152-160.1999
- 7 McCormack, S., Dunn, D.T., Desai, M., Dolling, D.I., Gafos, M., Gilson, R. et al. (2016). Pre-
8 exposure prophylaxis to prevent the acquisition of HIV-1 infection (PROUD):
9 effectiveness results from the pilot phase of a pragmatic open-label randomised trial.
10 *Lancet* 387, 53-60. doi: 10.1016/S0140-6736(15)00056-2
- 11 Nuñez, O., Hernando, V. and Díaz, A. (2018). Estimating the number of people living with HIV
12 and the undiagnosed fraction in Spain in 2013. *AIDS* 32, 2573-2581. doi:
13 10.1097/QAD.0000000000001989
- 14 Oster, A.M., France, A.M., Panneer, N., Bañez-Ocfemia, M.C., Campbell, E., Dasgupta, S. et al.
15 (2018). Identifying clusters of recent and rapid HIV transmission through analysis of
16 molecular surveillance data. *J. Acquir. Immune. Defic. Syndr.* 79, 543-550. doi:
17 10.1097/QAI.0000000000001856
- 18 Oster, A.M., Wertheim, J.O., Hernandez, A.L., Ocfemia, M.C., Saduvala, N. and Hall, H.I. (2015).
19 Using molecular HIV surveillance data to understand transmission between
20 subpopulations in the United States. *J. Acquir. Immune. Defic. Syndr.* 70, 444-451. doi:
21 10.1097/QAI.0000000000000809
- 22 Paraskevis, D., Beloukas, A., Stasinou, K., Pantazis, N., de, M.C., Bannert, N. et al. (2019). HIV-1
23 molecular transmission clusters in nine European countries and Canada: association
24 with demographic and clinical factors. *BMC. Med.* 17, 4. doi: 10.1186/s12916-018-
25 1241-1
- 26 Paraskevis, D., Nikolopoulos, G.K., Magiorkinis, G., Hodges-Mameletzis, I. and Hatzakis, A.
27 (2016). The application of HIV molecular epidemiology to public health. *Infect. Genet.*
28 *Evol.* 46, 159-168. doi: 10.1016/j.meegid.2016.06.021
- 29 Parczewski, M., Leszczyszyn-Pynka, M., Witak-Jedra, M., Szetela, B., Gasiorowski, J., Knysz, B.
30 et al. (2017). Expanding HIV-1 subtype B transmission networks among men who have
31 sex with men in Poland. *PLoS. One.* 12, e0172473. doi: 10.1371/journal.pone.0172473
- 32 Patiño-Galindo, J.A., Thomson, M.M., Pérez-Álvarez, L., Delgado, E., Cuevas, M.T., Fernández-
33 García, A. et al. (2016). Transmission dynamics of HIV-1 subtype B in the Basque
34 Country, Spain. *Infect. Genet. Evol.* 40, 91-97. doi: 10.1016/j.meegid.2016.02.028
- 35 Patiño-Galindo, J.A., Torres-Puente, M., Bracho, M.A., Alastrué, I., Juan, A., Navarro, D. et al.
36 (2017a). Identification of a large, fast-expanding HIV-1 subtype B transmission cluster
37 among MSM in Valencia, Spain. *PLoS. One.* 12, e0171062. doi:
38 10.1371/journal.pone.0171062
- 39 Patiño-Galindo, J.A., Torres-Puente, M., Bracho, M.A., Alastrué, I., Juan, A., Navarro, D. et al.
40 (2017b). The molecular epidemiology of HIV-1 in the Comunidad Valenciana (Spain):
41 analysis of transmission clusters. *Sci. Rep.* 7, 11584. doi: 10.1038/s41598-017-10286-1

- 1 Pérez-Parra, S., Chueca, N., Álvarez, M., Pasquau, J., Omar, M., Collado, A. et al. (2016).
2 Phylodynamic and Phylogeographic Profiles of Subtype B HIV-1 Epidemics in South
3 Spain. *PLoS. One.* 11, e0168099. doi: 10.1371/journal.pone.0168099
- 4 Pérez-Parra, S., Chueca-Porcuna, N., Álvarez-Estevez, M., Pasquau, J., Omar, M., Collado, A. et
5 al. (2015). Los estudios de resistencia a antirretrovirales como herramienta para el
6 análisis de los clusters de transmisión del virus de la inmunodeficiencia humana. *Enferm.*
7 *Infec. Microbiol. Clin.* 33, 603-608. doi: 10.1016/j.eimc.2014.11.016
- 8 Petersen, A., Cowan, S.A., Nielsen, J., Fischer, T.K. and Fonager, J. (2018). Characterisation of
9 HIV-1 transmission clusters and drug-resistant mutations in Denmark, 2004 to 2016.
10 *Euro. Surveill* 23. doi: 10.2807/1560-7917.ES.2018.23.44.1700633
- 11 Pineda-Peña, A.C., Pingarilho, M., Li, G., Vrancken, B., Libin, P., Gomes, P. et al. (2019). Drivers
12 of HIV-1 transmission: The Portuguese case. *PLoS. One.* 14, e0218226. doi:
13 10.1371/journal.pone.0218226
- 14 Price, M.N., Dehal, P.S. and Arkin, A.P. (2010). FastTree 2--approximately maximum-likelihood
15 trees for large alignments. *PLoS. One.* 5, e9490. doi: 10.1371/journal.pone.0009490
- 16 Ragonnet-Cronin, M., Hué, S., Hodcroft, E.B., Tostevin, A., Dunn, D., Fawcett, T. et al. (2018).
17 Non-disclosed men who have sex with men in UK HIV transmission networks:
18 phylogenetic analysis of surveillance data. *Lancet HIV.* 5, e309-e316. doi:
19 10.1016/S2352-3018(18)30062-6
- 20 Ragonnet-Cronin, M., Lycett, S.J., Hodcroft, E.B., Hué, S., Fearnhill, E., Brown, A.E. et al. (2016).
21 Transmission of Non-B HIV Subtypes in the United Kingdom Is Increasingly Driven by
22 Large Non-Heterosexual Transmission Clusters. *J. Infect. Dis.* 213, 1410-1418. doi:
23 10.1093/infdis/jiv758
- 24 Riddell, J., Amico, K.R. and Mayer, K.H. (2018). HIV Preexposure Prophylaxis: A Review. *JAMA*
25 319, 1261-1268. doi: 10.1001/jama.2018.1917
- 26 Thomson, M.M., Delgado, E., Manjón, N., Ocampo, A., Villahermosa, M.L., Mariño, A. et al.
27 (2001). HIV-1 genetic diversity in Galicia Spain: BG intersubtype recombinant viruses
28 circulating among injecting drug users. *AIDS* 15, 509-516. doi: 10.1097/00002030-
29 200103090-00010
- 30 UNAIDS. (2014). *90-90-90 An ambitious treatment target to help end the AIDS epidemic.*
31 Available at:
32 [https://files.unaids.org/en/media/unaids/contentassets/documents/unaidspublication](https://files.unaids.org/en/media/unaids/contentassets/documents/unaidspublication/2014/90-90-90_en.pdf)
33 [/2014/90-90-90_en.pdf](https://files.unaids.org/en/media/unaids/contentassets/documents/unaidspublication/2014/90-90-90_en.pdf)
- 34 United Nations Population Fund, Global Forum on MSM & HIV, United Nations Development
35 Programme, World Health Organization, United States Agency for International
36 Development, World Bank. (2015). *Implementing comprehensive HIV and STI*
37 *programmes with men who have sex with men: practical guidance for collaborative*
38 *interventions.* Available at: <https://www.unfpa.org/es/node/13155>
- 39 Vasylyeva, T.I., Friedman, S.R., Paraskevis, D. and Magiorkinis, G. (2016). Integrating molecular
40 epidemiology and social network analysis to study infectious diseases: Towards a
41 socio-molecular era for public health. *Infect. Genet. Evol.* 46, 248-255. doi:
42 10.1016/j.meegid.2016.05.042

- 1 Vasyljeva, T.I., Zarebski, A., Smyrnov, P., Williams, L.D., Korobchuk, A., Liulchuk, M. et al.
2 (2020). Phylodynamics helps to evaluate the impact of an HIV prevention intervention.
3 *Viruses*. 12. doi: 10.3390/v12040469
- 4 Vega, Y., Delgado, E., Fernández-García, A., Cuevas, M.T., Thomson, M.M., Montero, V. et al.
5 (2015). Epidemiological Surveillance of HIV-1 transmitted drug resistance in Spain in
6 2004-2012: Relevance of transmission clusters in the propagation of resistance
7 mutations. *PLoS. One*. 10, e0125699. doi: 10.1371/journal.pone.0125699
- 8 Verhofstede, C., Dauwe, K., Fransen, K., Van, L.K., Van den Wijngaert, S., Ruelle, J. et al. (2018).
9 Phylogenetic analysis of the Belgian HIV-1 epidemic reveals that local transmission is
10 almost exclusively driven by men having sex with men despite presence of large
11 African migrant communities. *Infect. Genet. Evol.* 61, 36-44. doi:
12 10.1016/j.meegid.2018.03.002
- 13 Viciano, I., González-Domenech, C.M., Palacios, R., Delgado, M., Del, A.A., Tellez, F. et al.
14 (2016). Clinical, virological and phylogenetic characterization of a multiresistant HIV-1
15 strain outbreak in naive patients in southern Spain. *J. Antimicrob. Chemother.* 71, 357-
16 361. doi: 10.1093/jac/dkv332
- 17 Vinken, L., Fransen, K., Cuypers, L., Alexiev, I., Balotta, C., Debaisieux, L. et al. (2019). Earlier
18 Initiation of Antiretroviral Treatment Coincides With an Initial Control of the HIV-1
19 Sub-Subtype F1 Outbreak Among Men-Having-Sex-With-Men in Flanders, Belgium.
20 *Front Microbiol.* 10, 613. doi: 10.3389/fmicb.2019.00613 [doi]
- 21 Wertheim, J.O., Kosakovsky Pond, S.L., Forgiione, L.A., Mehta, S.R., Murrell, B., Shah, S. et al.
22 (2017). Social and genetic networks of HIV-1 transmission in New York City. *PLoS.*
23 *Pathog.* 13, e1006000. doi: 10.1371/journal.ppat.1006000
- 24 Wirden, M., De, O.F., Bouvier-Alias, M., Lambert-Niclot, S., Chaix, M.L., Raymond, S. et al.
25 (2019). New HIV-1 circulating recombinant form 94: from phylogenetic detection of a
26 large transmission cluster to prevention in the age of geosocial-networking apps in
27 France, 2013 to 2017. *Euro. Surveill* 24. doi: 10.2807/1560-
28 7917.ES.2019.24.39.1800658
- 29 Yebra, G., Holguín, A., Pillay, D. and Hué, S. (2013). Phylogenetic and demographic
30 characterization of HIV-1 transmission in Madrid, Spain. *Infect. Genet. Evol.* 14, 232-
31 239. doi: 10.1016/j.meegid.2012.12.006
32
33
34

1 **Table 1.** Characteristics of the HIV-1 newly diagnosed individuals included in the study.

		BC*		GA†		All		p-value
		N	%	N	%	N	%	
Gender	Male	626	82	306	83	932	82	0.71
	Female	133	17	65	17	198	17	
	Transexual	5	0.65	0	0	5	0.44	
	Unknown	12		11		23		
Transmission route	MSM ‡	317	47	142	44	459	46	0.61
	Heterosexual	233	34	112	35	345	34	
	MNSST §	103	15	48	15	151	15	
	PWID ¶	22	3.2	17	5.3	39	3.9	
	Other	6	0.88	3	0.93	9	0.90	
	Unknown	95		60		155		
Region of origin	Spain	477	65	268	83	745	70	<0.001
	Latin America	147	20	37	11	184	17	
	Sub-Saharan Africa	79	11	6	1.9	85	8.0	
	North Africa	14	1.9	1	0.3	15	1.4	
	Europe	16	2.2	12	3.7	28	2.6	
	Other	5	0.68	1	0.31	6	0.56	
	Unknown	38		57		95		
Genetic form	B	489	63	235	62	724	63	<0.001
	F	39	5.0	49	13	88	7.6	
	A	21	2.7	12	3.1	33	2.9	
	C	19	2.5	20	5.2	39	3.4	
	G	12	1.6	8	2.1	20	1.7	
	CRF02_AG	77	9.9	12	3.1	89	7.7	
	CRF_BF	20	2.6	8	2.1	28	2.4	
	URF	65	8.4	30	7.9	95	8.2	
	Other	34	4.4	8	2.1	42	3.6	
	Total		776		382		1158	

2 *BC: Basque Country; †GA: Galicia; ‡MSM: Men who have sex with men; § MNSST: Men who
3 have non-specified sexual transmission; ¶ PWID: Person who injects drugs.

4

1 **Table 2.** New HIV-1 diagnoses from Basque Country and Galicia (2013-18) included in the study
2 and reported to the HIV National Surveillance System.

3

		Laboratory		Surveillance*		<i>p</i> -value
		N†	%	N	%	
Gender	Male	932	82	1477	81	0.461
	Female	198	18	340	19	
	Unknown & Other‡	28		0		
Transmission route	MSM §	459	54	957	57	0.096
	Heterosexual	345	41	631	37	
	PWID ¶	39	4.6	104	6.1	
	MNSST ¶	151		NA‡		
	Unknown & Other‡	164		125		
Age group	<20	25	2.2	5	1.5	0.131
	20-29	256	23	23	21	
	30-39	396	35	151	33	
	≥40	456	40	280	44	
	Unknown‡	25		0		
Region of origin	Spain	745	70	1259	70	0.881
	Latin America	184	17	301	17	
	Sub-Saharan Africa	85	8.0	159	8.9	
	North Africa	15	1.4	22	1.2	
	Europe	28	2.6	54	3.0	
	Unknown & Other‡	101		22		
Total	Both Regions	1158	64	1817	100	
	Basque Country	759	84	907	100	
	Galicia	371	41	910	100	

4 * Data from the Spanish Information Systems on HIV New Diagnoses (SINIVIH); † N: number of
5 patients; ‡ Sexual male, unknown and others categories were not included in the percentage
6 and *p*-value calculations; § MSM: men who have sex with men; ¶ PWID: person who injects
7 drugs; ¶ MNSST: Men who have non-specified sexual transmission; ‡ NA: Not applicable,
8 MNSST is not a category collected by SINIVIH.

1 **Table 3:** Characteristic of the identified TCs of the study.
2

		Basque Country		Galicia		Both regions	
		N	%	N	%	N	%
Transmission route*	MSM†	64	65	46	56	95	59
	Heterosexual	10	10	14	17	24	15
	Sexual‡	16	16	10	12	23	14
	PWID §	8	8.2	12	15	20	12
Total		98	100	82	100	162	100

3 * Main mode of HIV transmission among individuals in the TC; †MSM: men who have sex with
4 men; ‡ There is not predominance of MSM or heterosexual transmission; § PWID: Persons who
5 inject drugs.

1 **Table 4.** Characteristics of the transmission clusters comprising ≥ 30 patients in the Basque Country and Galicia (2013-2018).

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22

Cluster	Genetic form	No. patients	Regional distribution						% Spanish	% Male	Main trasmission route	Patients (2013-18)‡		Period§
			BC*		GA†		Others					N	%	
			N	%	N	%	N	%						
F1_1	F1	205	9	4.4	138	67	58	28	85	97	MSM	78	38	2000-2019
CRF02_1	CRF02_AG	83	34	41	3	3.6	46	55	79	92	MSM	44	53	2008-2019
B50	B	74	15	20	3	4.1	56	76	85	100	MSM	34	46	2007-2019
A1_1	A1	63	8	13	24	38	31	49	71	95	MSM	48	76	2006-2019
B70	B	61	56	92			5	8.2	88	98	MSM	44	72	2008-2019
B08	B	60	46	77	6	10	8	13	92	98	MSM	13	22	2006-2019
B09	B	57	55	97			2	3.5	88	98	MSM	10	18	1991-2019
B13	B	55			54	98	1	1.8	89	94	MSM	16	29	2004-2017
B31	B	49	19	39	2	4.1	28	57	77	98	MSM	23	47	2001-2019
B07	B	38	36	95			2	5.3	91	95	MSM	7	18	2000-2019
B10	B	36	29	81			7	19	90	97	MSM	7	19	1991-2019
F1_3	F1	36	29	81			7	19	76	100	MSM	34	94	2014-2018
B12	B	35	25	71	1	2.9	9	26	63	100	MSM	11	31	2009-2019
BG_2	URF_BG	33			33	100			85	80	PWID	2	6	1994-2019
B05	B	31			31	100			93	94	MSM	10	32	2003-2019

23 1. *BC: Basque Country; †GA: Galicia; ‡ Newly-diagnosed HIV-1 patients during 2013-2018; §Year of HIV-1 diagnosis of the patients in the TC.

1 **Table 5:** Characteristics of the new HIV diagnoses included in TCs (2013-2018).
2

Variables	Categories	Basque Country					Galicia					Both regions				
		No		Yes		p-value	No		Yes		p-value	No		Yes		p-value
		N*	%	N	%		N	%	N	%		N	%	N	%	
Gender (n= 1135, no data=23)	Female	115	86	18	14	<0.001	40	62	25	38	<0.001	155	78	43	22	<0.001
	Male	303	48	323	52		95	31	211	69		398	43	534	57	
	Transexual	3	60	2	40		0	0	0	0		3	60	2	40	
Transmission route (n=1003, no data=155)	Heterosexual	178	76	55	24	<0.001	60	54	52	46	<0.001	238	69	107	31	<0.001
	MSM†	120	38	197	62		33	23	109	77		153	33	306	67	
	MNSST‡	48	47	55	53		18	37	30	63		66	43	85	56	
	PWID§	15	68	7	32		8	47	9	53		23	59	16	41	
	Other	4	67	2	33		3	100	0	0		7	78	2	22	
Region of origin (n=1063, no data=95)	Spain	189	40	288	60	<0.001	82	31	186	69	0.001	271	37	474	64	<0.001
	Latin America	102	69	45	31		18	49	19	51		120	65	64	35	
	Sub-Saharan Africa	79	100	0	0		5	83	1	17		84	99	1	1.2	
	North Africa	9	64	5	36		1	100	0	0		10	67	5	33	
	Europe	11	69	5	31		7	58	5	42		18	64	10	36	
	Other	4	80	1	20		1	100	0	0		5	83	1	17	
Age group (n=1133, no data=25)	<20	10	53	9	47	0.44	3	50	3	50	0.003	13	52	12	48	0.036
	20-29	92	51	90	49		14	19	60	81		106	41	150	59	
	30-39	160	58	115	41		49	41	72	59		209	53	187	47	
	≥40	161	55	137	45		68	43	90	57		229	50	227	50	
Genetic form (n=1158, no data=0)	A	14	67	7	33	<0.001	3	25	9	75	0.002	17	52	16	48	<0.001
	B	224	46	265	54		83	35	152	65		307	42	417	58	
	C	11	58	8	42		8	40	12	60		19	49	20	51	
	F	10	26	29	74		11	23	38	78		21	24	67	76	
	G	10	83	2	17		5	63	3	37		15	75	5	25	
	CRF02_AG	63	82	14	18		11	92	1	8.3		74	83	15	17	
	CRF_BF	18	90	2	10		4	50	4	50		22	79	6	21	
	URF	47	72	18	28		10	33	20	67		57	60	38	40	
	Other	28	82	6	18		4	50	4	50		32	76	10	24	

3 *N: Number of patients; † MSM: men who have sex with men; ‡ MNSST: Men who have non-
4 specified sexual transmission; §PWID: People who inject drugs.

5

1 **Table 6:** Factors associated with transmission clusters, univariate/multivariate analysis.
2

Variables		Univariate analysis		Multivariate analysis	
		OR*	95% CI†	Adjusted OR	95% CI
Gender	Female	Reference		Reference	
	Male	4.8	3.3-7.1	2.6	1.5-4.7
	Transexual	2.4	0.4-15	NA‡	NA
Transmission route	Heterosexual	Reference		Reference	
	MSM§	4.5	3.2-6.1	2.1	1.4-3.2
	MNNSST¶	2.9	1.9-4.3	1.4	0.8-2.3
	PWIDⓇ	1.6	0.8-3.1	NA	NA
	Other	0.6	0.1-3.1	NA	NA
Region of origin	Spain	Reference		Reference	
	Latin America	0.3	0.2-0.4	0.3	0.2-0.4
	Sub-Saharan Africa	<0.01	<0.001-0.06	0.02	0.003-0.2
	North Africa	0.3	0.1-0.9	0.4	0.2-1.0
	Europe	0.3	0.1-0.7	0.3	0.03-3.5
	Other	0.1	0.1-1.0	0.4	0.1-1.6
Age group	<20	0.7	0.3-1.5	NA	NA
	20-29	Reference		Reference	
	30-39	0.6	0.5-0.9	0.6	0.4-0.9
	≥40	0.7	0.5-1.0	0.6	0.4-0.9
Spanish Region	Basque Country	Reference		Reference	
	Galicia	2.1	1.6-2.7	1.5	1.1-2.2
Genetic form	A	0.7	0.3-1.4	NA	NA
	B	Reference		Reference	
	C	0.8	0.4-1.5	NA	NA
	F	2.4	1.4-3.9	2.5	1.3-5.0
	G	0.3	0.1-0.7	0.4	0.1-1.5
	CRF02_AG	0.2	0.1-0.3	0.6	0.3-1.2
	CRF_BF	0.2	0.1-0.5	0.1	0.03-0.4
	URF	0.5	0.3-0.8	0.8	0.4-1.4
Other	0.4	0.3-0.7	0.3	0.2-0.8	

3
4 *OR: Odds ratio; † 95% CI: 95% confidence interval; ‡NA: Not applicable. Adjusted OR were
5 not calculated for categories with p<0.1 in the univariate analysis; § MSM: Men having sex
6 with men; ¶ MNNSST: Men who have non-specified sexual transmission; Ⓡ PWID: People who
7 inject drugs.

1

2