Non-B subtypes account for a large proportion of clustered primary HIV-1 infections in Italy

Giorgio Bozzi, Lavinia Fabeni, Isabella Abbate, Giulia Berno, Antonio Muscatello, Lucia Taramasso, Massimiliano Fabbiani, Silvia Nozza, Giuseppe Tambussi, Stefano Rusconi, Andrea Giacomelli, Emanuele Focà, Carmela Pinnetti, Gabriella d’Ettorre, Cristina Mussini, Vanni Borghi, Benedetto Maurizio Celesia, Giordano Madeddu, Antonio Di Biagio, Diego Ripamonti, Nicola Squillace, Andrea Antinori, Andrea Gori, Maria Rosaria Capobianchi, Alessandra Bandera

ABSTRACT

Objectives and design Using pol sequences obtained for routine resistance testing, we characterised the molecular patterns of HIV-1 transmission and factors associated with being part of a transmission cluster among individuals who in 2008–2014 presented with primary HIV-1 infection (PHI) at 11 urban centres across Italy.

Methods Pol sequences were obtained by Sanger sequencing. Transmission clusters were identified by phylogenetic analysis (maximum likelihood method, confirmed by Bayesian analysis). Multivariable logistic regression explored factors associated with a participant being part of a transmission cluster.

Results The PHI cohort comprised 186 participants (159/186, 85.5% males) with median age 44 years, median CD4 count 464 cells/mm³ and median plasma HIV-1 RNA 5.6 log₁₀ copies/mL. Drug resistance associated mutations were found in 16/186 (8.6%). A diversity of non-B subtypes accounted for 60/186 (32.3%) of all infections. A total of 17 transmission clusters were identified, including 44/186 (23.7%) participants. Each cluster comprised 2–6 sequences. Non-B subtypes accounted for seven clusters and 22/44 (50%) of clustered sequences. In multivariable logistic regression analysis, factors associated with being part of a transmission cluster comprised harbouring a non-B subtype (adjusted OR 2.28, 95% CI 1.03 to 5.05; p=0.04) and showing a lower plasma HIV-1 RNA (adjOR 0.80, 95% CI 0.64 to 0.99; p=0.04).

Conclusions There was a large contribution of diverse non-B subtypes to transmission clusters among people presenting with acute or recent HIV-1 infection in this cohort, illustrating the evolving dynamics of the HIV-1 epidemic in Italy, where subtype B previously dominated.

INTRODUCTION

Despite great improvements in prevention and treatment, 4.2 new HIV-1 diagnoses per 100 000 residents were made in Italy in 2019. Individuals in the acute and recent phases of the infection play an important role in sustaining HIV transmission, making early diagnosis and prompt initiation of antiretroviral treatment (ART) critical. Italian Network of Acute HIV Infection is a nationwide, multicentre cohort study that up to 2014 collected data and viral genome sequences from patients diagnosed with acute or early HIV-1 infection across clinical centres in Italy. Here, we performed a retrospective analysis to obtain a molecular characterisation of the infection among patients who were diagnosed in 2008–2014 and related the findings to demographic and clinical parameters. By using sequences obtained within early HIV infection, we expected to obtain a clearer snapshot of the HIV epidemic in Italy.

METHODS

Study population

We studied 186 treatment-naïve participants with primary HIV-1 Infection (PHI), defined as detectable plasma HIV-1 RNA plus either: (1) a negative third or fourth generation HIV immunoassay or (2) an indeterminate or positive western blot or RIBA, without reactivity of p31 band. PHI was classified according to Fiebig stages I–V. Symptoms compatible with acute retroviral infection were investigated and reported.

Pol sequences (containing the full-length protease and the first 330–335 reverse transcriptase (RT) codons), obtained for routine baseline drug resistance testing, were performed at each centre by Sanger sequencing. All HIV-1 pol genotype analyses were performed on plasma samples. Transmitted drug resistance (TDR) was evaluated according to the WHO 2009 list, with the additional RT mutations K65E/N, E138G/K/Q/R, V179L, G190Q, T215N, H221Y, F227C and M230I.

Phylogenetic analysis

To determine HIV-1 subtype, pol sequences were aligned with full-length reference sequences retrieved from the LANL database (https://www.hiv.lanl.gov/content/sequence/NewALIGN/align.html) using Clustal X and ≥ 10 reference sequences for each subtype, as previously described. Transmission clusters were first deduced through the neighbour-joining method using all sequences (MEGA v6, Kimura 2-parameter model, bootstrap analysis of 1000 replicates). To avoid influence of convergent evolution, sequences were stripped at drug
resistance positions. Transmission clusters supported by a bootstrap value ≥90% and average genetic distance ≤0.020 substitution/site were selected. Robustness was further tested using the maximum likelihood method inferred through GTR+I+Γ5, and 1000 bootstrap replicates.6 7 The tree was rooted through FigTree V.1.4.4. The GTR+I+Γ model was considered the best one (MEGA 6 model test) due to showing the lowest Bayesian Information Criterion score. Finally, a Bayesian phylogenetic tree was reconstructed with MrBayes, using GTR+I+Γ5. The Monte Carlo Markov Chain search was run for 5×10⁶ generations with trees sampled every 100th generation (10% burn-in). Posterior probability of each monophyletic clade was calculated for statistical support, and a posterior consensus tree was generated after 10% burn-in. Transmission clusters showed inferred posterior probabilities ≥0.90.

**Statistical analysis**

Comparisons between participants belonging or not to clusters were evaluated using the Mann-Whitney test for quantitative variables and χ² or Fisher’s exact test for categorical variables. Factors potentially associated with being part of a transmission cluster (based on p value at univariable analysis or previous scientific reports) were evaluated using multivariable logistic regression analysis. The following factors were included in the model: gender, year of diagnosis, age, Fiebig stage, transmission group, HIV-1 subtype, plasma HIV-1 RNA load and CD4 cell count at the time of diagnosis, and presence of TDR. All analyses were performed using SPSS V.23.

**RESULTS**

Among 186 individuals with PHI, 153 (82.3%) were from cities in the north of Italy (Milan: 54, Brescia: 38, Monza: 20, Bergamo: 16, Modena: 22, Genoa: 3); 26 (14.0%) from Rome; and 7 (3.8%) from Sardinia (Sassari 2) and Sicily (Catania 5). Their characteristics are shown in table 1. TDR was present in 16/186 (8.6%) participants and included 2/44 (4.5%) of clustered sequences and 14/142 (9.9%) of non-cluster sequences. Molecular subtyping showed that clade B

### Table 1  Characteristics of the study population stratified by clustering

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Overall (n=186)</th>
<th>In cluster (n=44; 23.7%)</th>
<th>Out of cluster (n=142; 76.3%)</th>
<th>P value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>159 (85.5)</td>
<td>40 (91.0)</td>
<td>119 (83.8)</td>
<td>0.33</td>
</tr>
<tr>
<td>Female</td>
<td>27 (14.5)</td>
<td>4 (9.0)</td>
<td>23 (16.2)</td>
<td></td>
</tr>
<tr>
<td>Age (years), median (IQR)</td>
<td>44 (36–53)</td>
<td>42 (34–49)</td>
<td>45 (37–53)</td>
<td>0.19</td>
</tr>
<tr>
<td>Fiebig stages</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I–IV</td>
<td>128 (68.8)</td>
<td>27 (61.4)</td>
<td>101 (71.1)</td>
<td>0.22</td>
</tr>
<tr>
<td>V</td>
<td>58 (31.2)</td>
<td>17 (38.6)</td>
<td>41 (28.9)</td>
<td></td>
</tr>
<tr>
<td>Transmission group, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterosexual</td>
<td>57 (30.6)</td>
<td>12 (27.3)</td>
<td>45 (31.7)</td>
<td>0.94</td>
</tr>
<tr>
<td>Men having sex with men</td>
<td>118 (63.4)</td>
<td>29 (65.9)</td>
<td>89 (62.7)</td>
<td></td>
</tr>
<tr>
<td>Intravenous drug users</td>
<td>7 (3.8)</td>
<td>2 (4.5)</td>
<td>5 (3.5)</td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>4 (2.2)</td>
<td>1 (2.3)</td>
<td>3 (2.1)</td>
<td></td>
</tr>
<tr>
<td>CD4 cell count (cells/mm³), median (IQR)</td>
<td>464 (335–603)</td>
<td>437 (316–548)</td>
<td>476 (340–617)</td>
<td>0.30</td>
</tr>
<tr>
<td>HIV-1 RNA (log10 copies/mL), median (IQR)</td>
<td>5.6 (4.9–6.4)</td>
<td>5.4 (4.3–6.1)</td>
<td>5.7 (5.0–6.4)</td>
<td>0.03</td>
</tr>
<tr>
<td>Subtype, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>126 (67.7)</td>
<td>22 (50.0)</td>
<td>104 (73.2)</td>
<td>0.003</td>
</tr>
<tr>
<td>Non-B</td>
<td>60 (32.3)</td>
<td>22 (50.0)</td>
<td>38 (26.8)</td>
<td></td>
</tr>
<tr>
<td>CRF02_AG</td>
<td>21 (11.3)</td>
<td>10 (22.7)</td>
<td>11 (7.7)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>FI</td>
<td>17 (9.1)</td>
<td>9 (20.5)</td>
<td>8 (5.6)</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>7 (3.8)</td>
<td>0 (0.0)</td>
<td>7 (4.9)</td>
<td></td>
</tr>
<tr>
<td>CRF18_cpx</td>
<td>3 (1.6)</td>
<td>3 (6.8)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>Others†</td>
<td>12 (6.5)</td>
<td>0 (0.0)</td>
<td>12 (8.5)</td>
<td></td>
</tr>
<tr>
<td>Country of birth</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Italian</td>
<td>106 (57.0)</td>
<td>26 (59.1)</td>
<td>80 (56.3)</td>
<td>0.66</td>
</tr>
<tr>
<td>Non-Italian‡</td>
<td>9 (4.8)</td>
<td>1 (2.3)</td>
<td>8 (5.6)</td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>71 (38.2)</td>
<td>17 (38.6)</td>
<td>54 (38.1)</td>
<td></td>
</tr>
<tr>
<td>ART initiation§</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;3 months from diagnosis</td>
<td>43 (24.7)</td>
<td>11 (28.2)</td>
<td>32 (23.7)</td>
<td>0.57</td>
</tr>
<tr>
<td>≥3 months from diagnosis</td>
<td>131 (75.3)</td>
<td>28 (71.8)</td>
<td>103 (76.3)</td>
<td></td>
</tr>
<tr>
<td>Transmitted drug resistance, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>170 (91.4)</td>
<td>42 (95.5)</td>
<td>128 (90.1)</td>
<td>0.27</td>
</tr>
<tr>
<td>Any class</td>
<td>16 (8.6)</td>
<td>2 (4.5)</td>
<td>14 (9.9)</td>
<td></td>
</tr>
<tr>
<td>NNRTI</td>
<td>9 (4.8)</td>
<td>1 (2.3)</td>
<td>8 (5.6)</td>
<td>0.37</td>
</tr>
<tr>
<td>NRTI</td>
<td>5 (2.7)</td>
<td>0 (0.0)</td>
<td>5 (3.5)</td>
<td></td>
</tr>
<tr>
<td>PI</td>
<td>3 (1.6)</td>
<td>1 (2.3)</td>
<td>2 (1.4)</td>
<td></td>
</tr>
</tbody>
</table>

*By Mann-Whitney test (for quantitative variables) and χ² test or Fisher’s exact test (for categorical variables), as appropriate.
†Others: CRF01_AE, G, A1, CRF12_BF, CRF19_cpx and CRF28_BF.
‡Non-Italian area of origin was as follows: Europe for 3/9, Sub-Saharan Africa for 3/9, Northern Africa for 2/9, Asia for 1/9.
§Data available for 174/186 individuals.

ART, antiretroviral treatment; NNRTI, non-nucleoside reverse transcriptase inhibitors; NRTI, nucleoside reverse transcriptase inhibitors; PI, protease inhibitors.
accounted for 67.7% sequences. The most common non-B clades were CRF02_AG (21/186, 11.3%), F1 (17/186, 9.1%) and C (7/186, 3.8%). Other non-B clades were CRF18_cpx, CRF01_AE, G, A1, CRF12_BF, CRF19_cpx and CRF28_BF.

A phylogenetic tree was constructed encompassing all 186 sequences (online supplemental figure 1). A total of 17 transmission clusters were detected each including two to six sequences and encompassing 44/186 individuals (23.7%). Two clusters comprising six and four participants, respectively, involved subtype CRF02_AG, one cluster (three participants) involved CRF18_cpx and four clusters (two to three participants) involved subtype F1. The remaining 10 clusters involved two to three participants with subtype B. When comparing clustered and non-cluster participants, 22/44 (50.0%) and 38/142 (26.8%) (p = 0.003), respectively, had a non-B subtype, whereas the median HIV-1 RNA load was 5.4 log10 copies/mL and 5.7 log10 copies/mL, respectively (p = 0.03) (table 1). In multivariable logistic regression analysis, factors associated with being part of a transmission cluster comprised harbouring a non-B subtype (adjusted OR (adjOR) 2.28; 95% CI 1.03 to 5.05; p = 0.04) and showing a lower plasma HIV-1 RNA (adjOR 0.80, 95% CI 0.64 to 0.99; p = 0.04) (online supplemental table 1). In contrast, gender, year of diagnosis, age, Fiebig stage, HIV subtype and non-B subtype were born in Italy. This suggests that non-B strains formerly associated with migration are being transmitted and have become endemic in Italy. In fact, a substantial proportion of migrants living with HIV in Europe acquired HIV, including non-B strains, postmigration. Characterising subtype diversity is of epidemiological interest, but the clinical implications are not fully understood. In example, altered drug susceptibility to NRTI and NNRTI has been widely described for subtype C and a faster CD4 cell decline for subtype D, even though it is generally accepted that current antiretroviral regimens can be used reliably to treat patients with both B and non-B subtypes. Multivariable logistic regression confirmed that harbouring a non-B subtype and showing a lower viral load were associated with being in a cluster, whereas other characteristics including HIV transmission group were not. The median viral load was lower among participants within clusters than those outside of clusters. This might indicate that most cluster participants were in the later phases of PHI and therefore having had more time to become part of transmission networks, although we did not detect an association between Fiebig stage and clustering, or early treatment and clustering. Such finding might also be a proxy for other factors, that is, lack of symptoms (neurological symptoms are hinted to correlate to higher viral loads in PHI and might hamper the opportunity of transmission) or less accurate HIV-RNA quantification for certain non-B subtypes (challenges in accuracy for CRF02_AG viral load have been described).

In most molecular epidemiology studies, the high proportion of sequences belonging to patients with unknown duration of HIV infection poses a major limitation, hampering identification of features relevant to the time of analysis. The strength of our study is the use of sequences obtained within the first few months of infection, allowing a clearer snapshot of epidemic patterns. This is also a limitation of the study, as the methods employed only allow the detection of recent transmission clusters as the cohort does not encompass individuals with chronic infection. Other major limitations are that we could not build separate phylogenetic trees per each subtype. We also lacked important data for a proportion of participants, especially about country of origin. Nonetheless, our data show that HIV-1 diversity is increasing within the Italian epidemic and provide evidence of the indigenous transmission of non-B subtypes.

### Key messages
- By phylogenetic analysis of viral pol sequences obtained from 186 individuals who in 2004–2018 were diagnosed with primary HIV-1 infection at centres across Italy, we found that 24% of infections occurred within transmission clusters.
- There was a diversity of non-B subtypes that accounted for half of the transmission clusters. In this population, harbouring a non-B subtype was associated with being part of a cluster.
- Non-B subtypes are being transmitted indigenously and are becoming established in Italy.

### Author affiliations
1. Infectious Diseases Unit, Fondazione IRCCS Ca’ Granda Ospedale Maggiore Policlinico – Milan, Milan, Lombardia, Italy
2. Virology and Biosafety Laboratories Unit and HIV/AIDS Clinical Unit, National Institute for Infectious Diseases, “Lazzaro Spallanzani” - IRCCS Rome, Rome, Italy
3. Infection Diseases Unit, IRCCS Policlinico San Martino Hospital, Genoa, Italy
4. Infectious and Tropical Diseases Unit, Azienda Ospedaliero-Universitaria Senese, Siena, Italy
5. Clinic of Infectious Diseases, San Raffaele Hospital, University Vita Salute, Milan, Italy
6. Infectious Diseases Unit, Department of Biomedical and Clinical Sciences “Luigi Sacco” Hospital, University of Milan, Milan, Italy
7. Infectious Diseases Unit, Ospedale Civile di Legnano, ASST Ovest Milanese, Legnano, Italy
8. Division of Infectious and Tropical Diseases, University of Brescia, ASST Spedali Civili Hospital, Brescia, Italy
9. Infectious Diseases Unit, Umberto I Hospital, La Sapienza University, Rome, Rome, Italy
10. Clinic of Infectious Diseases, University of Modena and Reggio Emilia, Modena Hospital, Modena, Italy
11. Unit of Infectious Diseases, Garibaldi Hospital, Catania, Italy
12. Department of Medical, Surgical and Experimental Sciences, Unit of Infectious Diseases, University of Sassari, Sassari, Italy
13. Department of Health Sciences (DiSal), University of Genoa, Italy, Genoa, Italy
14. Division of Infectious Diseases, ASST Papa Giovanni XXIII, Bergamo, Italy
15. Infectious Diseases Unit, Department of Internal Medicine, ASST San Gerardo, Monza, Italy
16. Department of Pathophysiology and Transplantation, University of Milan, Milan, Italy

Handling editor Apostolos Beloukas

Acknowledgements The authors are grateful to Elena Cappelletti, Valeria Pastore, Valentina Ferroni, Ilaria Beretta and Teresa Itri for their help in data collection. We would like to thank Effeti Srl (Milan) for administrative support to the INACTION network. We also acknowledge Associazione Italiana per la Lotta contro l’AIDS (ANLAIM) Onlus for supporting our work.
Provenance and peer review

It was then approved by the EC of each participating center. It was conducted on 12 June 2014, with reference no. 1893 (authorised by Decree 93, 19 June 2014).

4

Patient consent for publication

Boeringher, Gilead, Janssen, MSD, Novartis, Pfizer, Roche, ViiV and no other conflict of interest. Speaker’s bureau compensation and/or travel support from Abbvie, Astellas, BMS, Gilead Sciences, Janssen, Viiv Healthcare, MSD and no other conflict of interest. AG reports membership or consultancy fees, speaker’s honoraria or travel support) with Gilead Sciences, Janssen, Viiv Healthcare and no other conflict of interest. CP reports past (within 36 months) financial relationships (advisory board membership or consultancy fees, speaker’s honoraria or travel support) with Gilead Sciences, Janssen, Viiv Healthcare and no other conflict of interest. SR reports past (within 36 months) financial relationships (advisory board membership or consultancy fees, speaker’s honoraria or travel support) with Gilead Sciences, Janssen, Viiv Healthcare, BMS, MSD and no other conflict of interest. AG reports past (within 36 months) financial relationships (advisory board membership or consultancy fees, speaker’s honoraria or travel support) with Gilead Sciences, Janssen, Viiv Healthcare, BMS, MSD and no other conflict of interest.

Funding

The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests

LT reports past (within 36 months) financial relationships (advisory board membership or consultancy fees, speakers’ honoraria or travel support) with Gilead Sciences, Janssen, ViiV Healthcare and no other conflict of interest. MF reports past (within 36 months) financial relationships (advisory board membership or consultancy fees, speakers’ honoraria or travel support) with Gilead Sciences, Janssen, ViiV Healthcare, BMS, MSD and no other conflict of interest. SR reports past (within 36 months) financial relationships (advisory board membership or consultancy fees, speakers’ honoraria or travel support) with Gilead Sciences, Janssen, ViiV Healthcare and no other conflict of interest. AG reports past (within 36 months) financial relationships (educational support) with Mylan and no other conflict of interest. EF reports past (within 36 months) financial relationships (advisory board membership or consultancy fees, speakers’ honoraria or travel support) with Gilead Sciences, Janssen, ViiV Healthcare, BMS, MSD and no other conflict of interest. GM reports past (within 36 months) financial relationships (advisory board membership or consultancy fees, speakers’ honoraria or travel support) with Gilead Sciences, Janssen, ViiV Healthcare, BMS, MSD and no other conflict of interest. AG reports past (within 36 months) grant/research supports, honoraria or consultation fees, speaker’s bureau compensation and/or travel support from Abbvie, Astellas, BMS, Boeringher, Gilead, Janssen, MSD, Novartis, Pfizer, Roche, ViiV and no other conflict of interest.

Patient consent for publication

Not applicable.

Ethics approval

The study was approved by the Ethics Committee of Monza-Brianza province, Lombardy, Italy, for the coordinating center, San Gerardo Hospital, on 12 June 2014, with reference no. 1893 (authorised by Decree 93, 19 June 2014). It was then approved by the EC of each participating center. It was conducted according to the declaration of Helsinki principles. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review

Not commissioned; externally peer reviewed.

Supplemental material

This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

ORCID iDs

Giorgio Bozzi http://orcid.org/0000-0002-4064-4017
Lucia Taramasso http://orcid.org/0000-0002-6622-6358
Stefano Rusconi http://orcid.org/0000-0002-0375-9990
Gabriella d’Ettorre http://orcid.org/0000-0002-3571-5677

REFERENCES