HIV-1 subtype diversity and phylogenetic insight into non-B subtype transmission in Slovenia, 1989-2013

Jana Mlakar[1], Maja M. Lunar[1], Ana B. Abecasis[2], Anne-Mieke Vandamme[2,3], Janez Tomažič[4], Tomaž D. Vovko[4], Blaž Pečavar[4], Gabriele Turel[4], Mario Poljak[1]✉

[1]*Institute of Microbiology and Immunology, Faculty of Medicine, University of Ljubljana, Ljubljana, Slovenia.* [2]*Global Health and Tropical Medicine, Instituto de Higiene e Medicina Tropical, Universidade Nova de Lisboa, Lisbon, Portugal.* [3]*Clinical and Epidemiological Virology, Institute for the Future, Rega Institute for Medical Research, KU Leuven, Leuven, Belgium.* [4]*Department of Infectious Diseases, University Medical Center Ljubljana, Ljubljana, Slovenia.*

Abstract

Introduction: Disease progression, drug resistance mutations, and treatment strategies may vary by HIV-1 subtype. This study determined HIV-1 subtypes circulating in Slovenia, a Central European country with an HIV-1 epidemic driven by men who have sex with men, focusing on molecular epidemiology of non-B subtypes.
Methods: A total of 367 HIV-1 sequences were included. Subtype was assigned by employing eight different HIV subtyping tools coupled with maximum likelihood phylogenetic analyses.
Results: The subtyping tools COMET, jpHMM, and REGA 3.0 exhibited the best performance on the dataset studied. Phylogenetic analyses showed a 14.7% prevalence of non-B subtypes, with subtype A detected most frequently (4.9%), followed by CRF02_AG (2.4%), subtype C (1.1%), subtypes D, G, and CRF01_AE (0.8% each), and subtypes F and CRF22_01A1 (0.3% each). A subtype could not be assigned to 12 sequences (3.3%), indicating potential unique recombinant forms. Non-B subtypes were significantly associated with a heterosexual route of transmission and infection acquired in Eastern Europe, Africa, or Asia.
Conclusions: In a country where subtype B is predominant, non-B subtypes were observed in one out of seven patients, a non-negligible proportion, which underlines the importance of systematic surveillance of HIV subtype diversity and the corresponding molecular epidemiology.

Keywords: HIV-1, non-B subtype, subtyping, phylogeny, molecular epidemiology

Received: 19 August 2023 | Returned for modification: 24 August 2023 | Accepted: 2 September 2023

Introduction

The HIV pandemic is still a global health problem, with an estimated 38.4 million people living with HIV in 2021 (1). However, the number of newly diagnosed HIV-1 infections is continuously declining (from 3.4 million in 2001 to 1.5 million in 2021), as well as the number of HIV-related deaths (from 2.3 million in 2005 to 650,000 in 2021), due to effective prevention efforts and high treatment coverage (1).

HIV-1 exists in many different subtypes and recombinant forms, which are divided into four groups: M (major), O (outlier), N (non-M, non-O), and P. Most of the subtypes circulating worldwide belong to group M, which currently includes nine subtypes, 132 circulating recombinant forms (CRFs; https://www.hiv.lanl.gov/components/sequence/HIV/crfdb/crfs.comp), and numerous unique recombinant forms (URFs). HIV-1 group M has been shown to originate from the Democratic Republic of the Congo (2), which is supported by the fact that the greatest diversity of subtypes is found in Central Africa, where all nine subtypes and most CRFs and URFs have been identified. Globally, the most prevalent subtype is subtype C, which accounts for nearly half (46.6%) of all HIV-1 infections, especially in southern Africa, East Africa, Ethiopia, and India. Other most prevalent subtypes are subtype B (12.1%), subtype A (10.3%), CRF02_AG (7.7%), CRF01_AE (5.3%), subtype G (4.6%), and subtype D (2.7%) (3). In Western and Central Europe and North America, subtype B predominates, at 83.3%. Other common non-B subtypes and CRFs are subtype C, CRF02_AG, subtype A, and subtype F (3).

Differences in transmission, identification of infection, disease progression, pathogenesis, development of resistance, clinical management of HIV infection, response to treatment, and drug and vaccine design can be observed between subtypes (4–8). One explanation for these differences is that distinct biological properties of HIV-1 variants play a role in their spread. For example, it was observed that subtype A virus in Uganda is transmitted at a higher rate heterosexually than subtype D, resulting in an increase in the proportion of subtype A and a decrease in the proportion of subtype D in East Africa and worldwide (4, 7, 9–11). Subtype C has been found to be both congenitally and sexually more transmissible compared to subtypes A and D (12, 13). In terms of disease progression, it has been observed that, in patients infected with subtype D, the CD4+ T-cell count declines faster, disease progression is faster, and dementia and advanced immunosuppression occur to a greater extent than in patients infected with subtype A (14, 15). On the other hand, individuals infected with HIV-1 subtypes A and C have been found to have a longer asymptomatic period and slower progression to AIDS together with an increased possibility of transmission compared to other subtypes (8, 16–18). Viral load suppression during highly active antiretroviral treatment (HAART) occurred most rapidly in patients infected with subtype A, followed by subtype C and subtype B (19). A study by Geretti et al. found that CD4+ T-cell recovery was similar across subtypes, and so the immunological efficacy of HAART should be similar (19). However, a study by Chaix et al. showed that patients infected with CRF02_AG or other

✉ Corresponding author: mario.poljak@mf.uni-lj.si

non-B subtypes had better immunological response to HAART than patients infected with subtype B (20). In addition, some drug resistance mutations due to naturally occurring polymorphisms are more common in certain HIV-1 subtypes, and so information about the subtype should be considered when initiating treatment (5, 7). Because it is still unknown how various HIV-1 subtypes respond to different antiretroviral therapy regimens, HIV-1 subtype is more frequently considered in clinical trials (5, 21).

In Slovenia, a small country in Central Europe, fewer than one in 1,000 inhabitants are estimated to be infected with HIV-1. Most HIV-1 infections occur among men who have sex with men (MSM). Other identified modes of transmission in Slovenia in the past decade (2012-2021) were heterosexual contact, injection drug use, and one case of mother-to-child transmission. According to national registry data, 759 people were living with HIV-1 in Slovenia at the end of 2021. Of these, 95% (718/759) are receiving antiretroviral treatment, and 97% (698/718) have an undetectable viral load (<40 copies/ ml) (22). In the last decade, the annual number of HIV-1 newly diagnosed individuals in Slovenia ranged from 27 in 2020 to 62 in 2016. Previous studies have shown that the majority of patients diagnosed with HIV-1 in Slovenia (82%–89%) are infected with subtype B (23–28). Subtype B was introduced in Slovenia at several time points from the beginning of the Slovenian epidemic, and the local transmission links of most patients indicate a closed community. Long-lasting infections have been identified as the driving force of the epidemic in this region. Lunar et al. previously studied the characteristics and transmission patterns of the subtype B epidemic in Slovenia (28). A later study characterized the potential HIV-1 URFs identified in Slovenia (29). However, the epidemic dynamics of non-B subtypes in Slovenia has not been comprehensively studied yet.

The aim of this study was to better understand the epidemic of non-B HIV-1 subtypes in Slovenia, with particular emphasis on transmission patterns, in order to adapt HIV-1 prevention strategies. To better distinguish the epidemics of the different lineages, we meticulously determined the HIV-1 subtypes circulating in Slovenia by employing different HIV subtyping tools and in-depth manual phylogenetic analyses.

Methods

Study population

For the purpose of the study, data and partial HIV-1 *pol* sequences (> 900 bp) were obtained from previous studies in Slovenia (23–28). Together with sequences obtained for routine genotypic drug resistance testing, the study sample consisted of 367 patients (324 drug-naive and 43 treated patients) with a confirmed HIV-1 diagnosis between the years 1989 and 2013. CD4+ T-cell count, HIV-1 viral load, and epidemiological data (nationality, country of infection, route of infection, and risk factors) were routinely collected by clinicians for each newly diagnosed patient in the form of a blinded questionnaire approved by the national Medical Ethics Committee of the Republic of Slovenia (consent number: 126/12/03) and in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

Subtyping

Nucleotide sequences were analyzed to determine the subtype by employing the following subtyping tools: REGA HIV-1 & 2 Automat-

ed Subtyping Tool Version 2.0 (REGA 2.0) (30, 31), REGA HIV Subtyping Tool Version 3.0 (REGA 3.0) (32), COMET HIV-1 1.0 (COMET) (33), jpHMM (34), SCUEAL (35), HIVdb: Genotypic Resistance Interpretation Algorithm (HIVdb) (36), STAR (37), and Geno2pheno [resistance] 3.3 (Geno2pheno) (38). The subtype of each sequence was confirmed by phylogenetic analyses (see below) if the sequence formed a monophyletic group together with the subtype-specific reference sequences or whole genome sequences.

Phylogenetic analyses

Phylogenetic analysis was performed to confirm the subtyping of the sequences and to investigate the molecular epidemiology of the non-B subtypes in Slovenia. Sequences were aligned to the 2010 HIV-1 subtype reference sequence set of group M subtypes (A–K and recombinants) obtained from the Los Alamos HIV Database using ClustalW multiple alignment incorporated in BioEdit version 7.0.9 (39). The Find Best DNA/Protein Models tool (ML) in Mega 5.05 (40) using the maximum likelihood statistical method and neighbor-joining tree was employed to select the best fitted evolutionary model. Finally, maximum likelihood (ML) phylogenetic trees were created by using PhyML 3.0 (41) and visualized in FigTree v1.4.2, available at http://tree.bio.ed.ac.uk/software/figtree/ (42). Transmission clusters were identified at this point according to approximate likelihood ratio test (aLRT) branch support values (> 0.90). The initial analysis was performed on all Slovenian sequences and subtype reference sequences, and a common tree was created. All non-B clusters containing Slovenian sequences were identified and were reanalyzed separately with up to 10 most similar sequences per Slovenian sequence selected from those available in GenBank using the HIV BLAST search tool (43–45), Standard Nucleotide BLAST (46), and Sequence Search Interface of the HIV sequence database (http://www.hiv.lanl.gov/). The maximum likelihood phylogenetic trees of subtype A and subtype G and their recombinants were inferred from two separate alignments for protease and reverse transcriptase combined into one.

Sequence data are available in GenBank; accession numbers: AJ971091–AJ971144, AM113750, GQ398934, GQ399003, GQ399157, GQ399167, GQ399210, GQ399318, GQ399406, GQ399433, GQ399494, GQ399553, GQ399574, GQ399677, GQ399709, GQ399721, GQ399731, GQ399787, GQ399882, GQ399950, GQ399979, GQ400015, GQ400033, GQ400039, GQ400057, GQ400283, GQ400355, GQ400410, GQ400411, GQ400442, GQ400452, GQ400472, JX028303–JX028406, JX046402–JX046417, KF753699–KF753751, KP013639–KP013747.

Phylodynamic analysis

The previously identified clusters of Slovenian non-B sequences were further analyzed along with up to 10 most similar control sequences used in the phylogenetic analyses to determine the time of the most recent common ancestor (tMRCA) by using the Monte Carlo Markov Chain (MCMC) method available in the software BEAST 2.1.3 (47). A relaxed clock log normal distribution and the Bayesian skyline coalescent model were employed on the data set (47–49), and a reversible jump-based substitution model (RB) was used as the substitution model. Sequences were split into two codon partitions (1+2 codon, 3 codon). The chain was run for 300,000,000 generations until the effective sample size exceeded 200. Tracer 1.5 (50) was used to analyze the output results. TreeAn-

notator 1.8.0 (51) was used to annotate the trees obtained with 10% burn-in, and the trees were viewed in FigTree v1.4.2 (42). The clusters identified in the ML analysis were confirmed if the posterior probability values obtained were > 0.990.

Drug resistance mutation determination

Sequences obtained from treatment-naive individuals were analyzed for the presence of surveillance drug resistance mutations (SDRM) using the Stanford University HIV Drug Resistance Database, HIVdb Program: Sequence Analysis (36), and Drug Resistance Mutations for Surveillance of Transmitted HIV-1 Drug-Resistance: 2009 Update (52).

Statistical analysis

Statistical significance tests were performed using Fisher's exact test for categorical data and a t-test for continuous data available in the open-source program OpenEpi 2.3.1 (http://www.openepi.com/Menu/OE_Menu.htm). Multivariate logistic regression was performed in IBM SPSS Statistics 27. Because of highly correlated variables, only route of transmission, country of infection, and CD4+ T-cell count < 200 cells/mm³ were included in the multivariate analysis. Values of $p \leq 0.05$ were considered significant.

Results

Subtyping of partial *pol* sequences using eight different subtyping tools and ML analyses

Of 367 sequences included in the study, all eight subtyping tools and ML analyses yielded congruent results for 80.9% (297/367) of the sequences: namely, 289 subtype B (this corresponds to 92.3% of subtype B sequences identified by ML analyses), two subtype A (11.1% of subtype A), four subtype C (100% of subtype C), and two CRF01_AE (66.7% of subtype CRF01_AE; Table 1, Fig. 1). Subtyping results obtained for the remaining 70 sequences for which at least one of the subtyping analyses gave discordant results are shown in Table 2.

As shown in Table 1, the only subtype for which the results of all subtyping tools matched for all sequences was subtype C. Subtype A was concordantly assigned by all tools except the similarity-based tools HIVdb and Geno2pheno, which subtyped part of the sequence or whole sequence as CRF01_AE, and STAR, which was unable to assign the subtype to many of the sequences (Table 2). Subtyping tools and phylogenetic analyses did not include ref-

erence sequences for all sub-subtypes of subtypes A and F, and so the assigned sub-subtype may not be correct and only the subtype is considered in the analysis. Only 8% (24/313) of the sequences classified as subtype B in the ML analyses were not concordantly assigned by one or more tools. HIVdb, jpHMM, and Geno2pheno assigned all these sequences to subtype B, whereas other tools classified them as BD, BF, and BU recombinants, B-like, complex recombinants, or unassigned. Three Slovenian sequences were classified as subtype D by REGA 2.0, REGA 3.0, COMET, Geno2pheno, and STAR, whereas jpHMM, SCUEAL, and HIVdb suggested BD recombinants or CRF19_cpx. When the subtype ML phylogenetic tree was examined (Fig. 1C), these sequences formed a separate cluster (aLRT > 0.9) and were clustered together with subtype D control sequences with an aLRT support of 0.875. Similar to subtype D, the ML analyses, REGA 2.0, REGA 3.0, and Geno2pheno were able to assign three sequences to subtype G that other tools were unable to assign or assigned them as GJ, AG, or AEG recombinants. Most tools had no problems identifying CRF02_AG, except REGA 2.0 and SCUEAL. Only SCUEAL was able to subtype a sequence as CRF22_01A1, as confirmed by ML analyses (Fig. 1A, cluster 2), whereas other tools classified it as A or as unassigned. One sequence determined as CRF22_01A1 by SCUEAL remained unassigned by phylogenetic analysis.

As shown in Table 2, and assuming that our ML analyses correspond to the correct assignment, the subtyping tools COMET, jpHMM, and REGA 3.0 provided the highest number of correctly subtyped sequences and thus showed the best performance for the Slovenian sequence set.

For 12 sequences, the subtype could not be determined even by manual ML analyses. The subtyping tools classified six of these sequences as subtype A, CRF01_AE, and/or their recombinants. The ML analyses also indicated recombinant forms for six other sequences, for which part of a sequence belonged to a subtype with which they formed a phylogenetic cluster. In addition, four sequences appeared to be markedly complex, indicating unique recombinant forms (Table 2).

Overall, according to the classification of ML analyses, the majority of patients (85.3%) were infected with subtype B (313/367), whereas subtype A was present in 4.9% (18/367), CRF02_AG in 2.4% (9/367), subtype C in 1.1% (4/367), subtypes D, G, and 01_AE in 0.8% (three patients each), and subtypes F and CRF22_01A1 in 0.3% (one patient each). In 12 sequences (3.3%), the subtype could not be assigned. Thus, a total of 54 patients (14.7%) were determined to be infected with non-B subtypes, including potential recombinants.

Table 1 | Number of subtyped sequences obtained with eight different subtyping tools available (REGA HIV-1 & 2 Automated Subtyping Tool version 2.0, REGA HIV Subtyping Tool version 3.0, COMET HIV-1 1.0, jpHMM, SCUEAL, HIVdb: Genotypic Resistance Interpretation Algorithm, STAR, Geno2pheno 3.3) and phylogenetic analyses on 367 Slovenian sequences.

Subtype	REGA 2.0	REGA 3.0	COMET	jpHMM	SCUEAL	HIVdb	Star	Geno2pheno 3.3	Phylogeny	Concordant
A	26	24	19	24	17	5	5	7	18	2
B	302	304	311	314	307	313	309	313	313	289
C	4	4	4	4	4	4	4	5	4	4
D	3	3	3	2	2	3	3	4	3	0
F	1	1	1	1	1	0	1	2	1	0
G	4	4	1	2	2	2	0	3	3	0
CRF01_AE	2	4	4	5	3	7	3	19	3	2
CRF02_AG	0	8	9	8	0	9	9	11	9	0
CRF09_cpx	0	0	0	/	0	/	/	3	0	0
CRF19_cpx	1	1	0	/	1	/	/	0	0	0
CRF22_01A1	/	/	0	/	2	/	/	0	1	0
NA	24	14	15	7	28	24	33	0	12	0

NA = subtype not assigned, / = corresponding subtype could not be assigned because it is not included in the reference set of the subtyping tool.
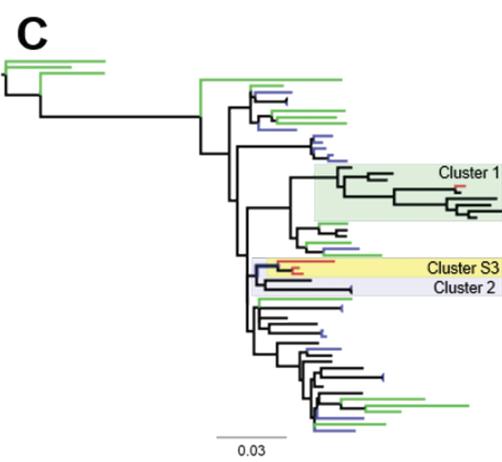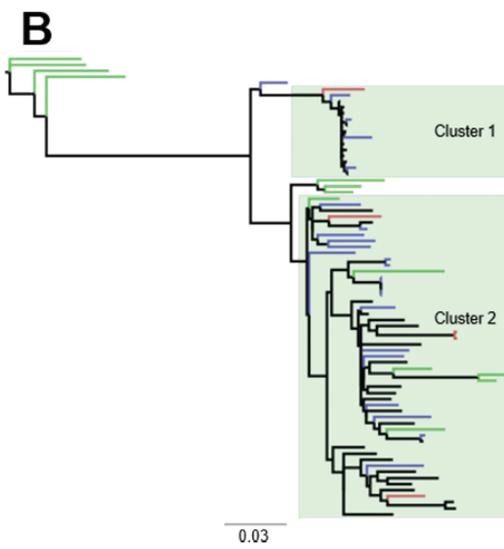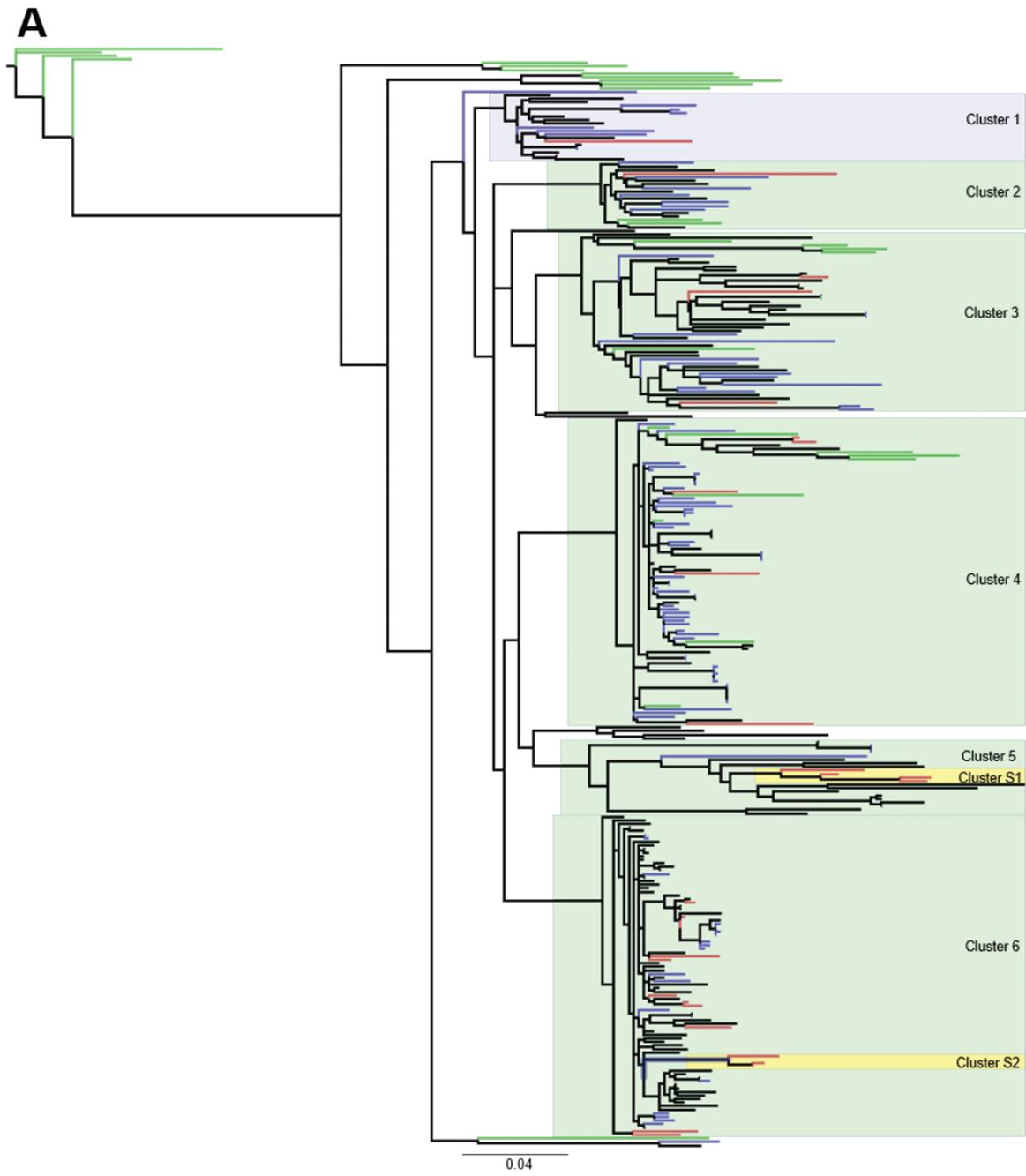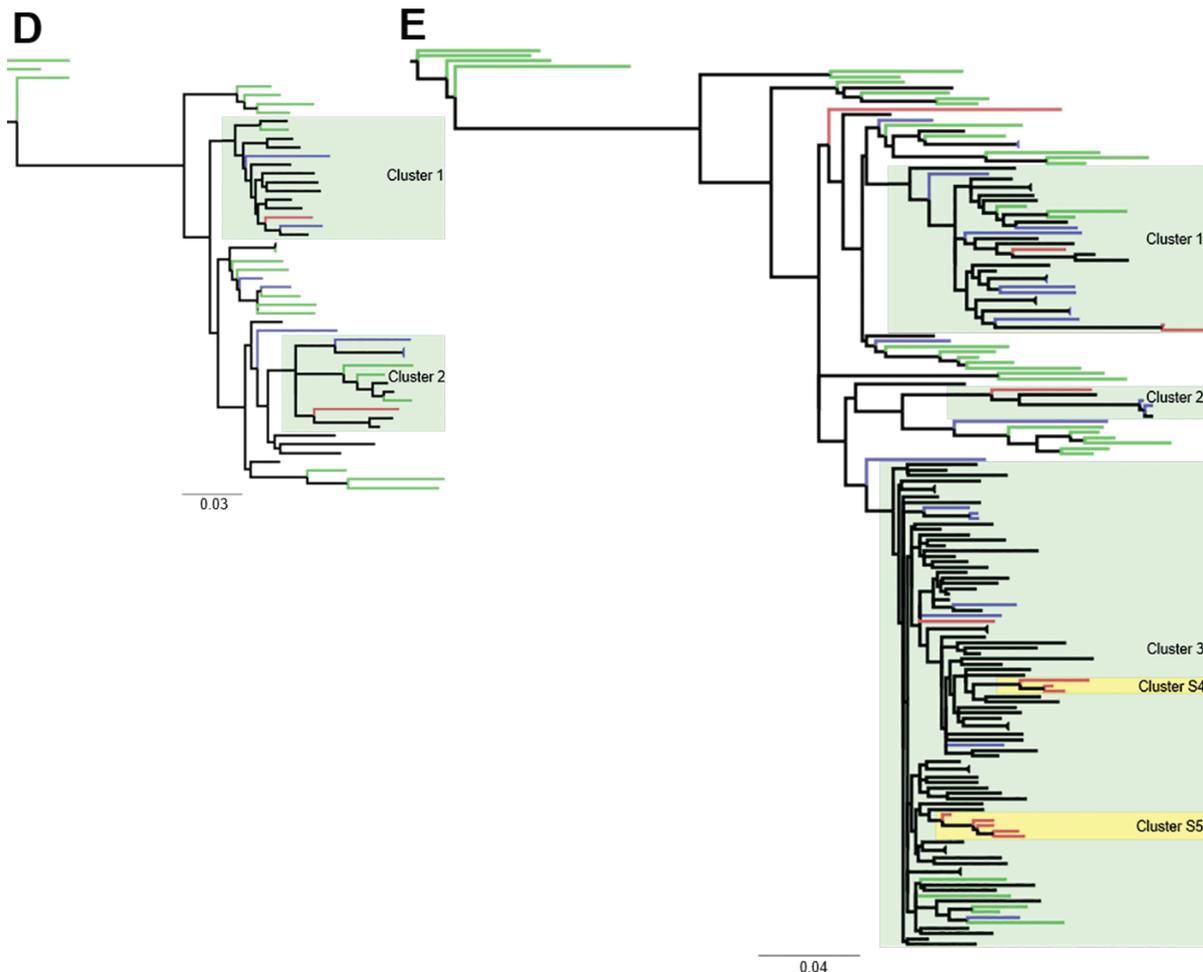
Figure 1 | The maximum likelihood phylogenetic trees of the different subtypes and their recombinants: (A) subtype A (relative to HXB2: 2262–2549, 2661–3290), (B) subtype C (HXB2: 2253–3554), (C) subtype D (HXB2: 2253–3554), (D) subtype F (HXB2: 2253–3554), (E) subtype G (HXB2: 2262–2549, 2661–3290). Slovenian sequences are colored red; reference sequences from 2010 HIV-1 subtype reference sequence set of group M subtypes (A–K and recombinants) obtained from the Los Alamos HIV Database are colored green; full genome control sequences (> 8500 bp) are colored blue; other similar control sequences are colored black (most similar sequences per Slovenian sequence selected from those available in GenBank using the HIV BLAST search tool (43–45), Standard Nucleotide BLAST (46), and Sequence Search Interface of the HIV sequence database). Clusters with aLRT > 0.9 are highlighted in green, clusters with aLRT > 0.8 are purple, and Slovenian clusters are yellow.

Clinical and epidemiologic characteristics of patients infected with non-B subtypes compared to patients infected with subtype B

Patient characteristics are shown in Table 3. Completed questionnaires from the SPREAD surveillance program were available for 75% (275/367) of patients, all treatment naive. Most patients were men (87.2%), and the most common route of infection was homo/bisexual contact (68.9%). Mean viral load was 4.94 ± 0.96 log copies/ml, and 34.3% of patients had a CD4+ T-cell count < 200 cells/mm³ at the time of diagnosis. Patients infected with subtype B were predominantly men (93.3%), whereas both sexes were equally represented in those not infected with subtype B, with a statistically significant difference between sex in subtype B and non-B subtype–infected patient groups ($p < 0.001$). Other statistically significant characteristics of patients found to be infected with non-B subtypes were heterosexual contact ($p < 0.001$), infection acquired abroad ($p < 0.001$), mainly in Eastern Europe, Africa, and Asia ($p < 0.001$, $p < 0.001$, $p = 0.014$, respectively), patients originated from abroad ($p < 0.001$), mainly from Eastern Europe and Africa ($p < 0.001$, $p = 0.031$, respectively), and more frequent diagnosis with a CD4+ T-cell count of < 200 cells/mm³ ($p = 0.029$). Multivariate logistic regression confirmed the statis-

tical significance of heterosexual transmission and infection acquired in Eastern Europe, Africa, and Asia ($p < 0.001$, $p = 0.023$, $p = 0.017$, $p = 0.049$, respectively; Table 3). SDRM were detected in 3.5% sequences obtained from treatment naive patients (Table 3). In addition, 21/43 treated patients had drug resistance mutations identified. Amongst, 2/21 were infected with non-B subtypes, one was identified as CRF22_01A1 and one as CRF02_AG.

Molecular epidemiology of non-B subtypes in Slovenia

ML analysis showed that non-B subtypes were introduced at least 34 times into Slovenia (Fig. 1). In the phylogenetic trees of non-B subtypes, only five small clusters (three to five sequences) and six transmission pairs of Slovenian non-B sequences were formed, which means that only 57% (31/54) of patients were found to have a transmission link within the country.

The most prevalent non-B subtype in Slovenia was subtype A, with 18 sequences identified. The phylogenetic tree obtained by ML analysis of subtype A and subtype A–related recombinants that fully or partially correspond to subtype A in part of the HIV-1 genome studied (e.g., CRF01_AE) showed that Slovenian sequences were part of six major international clusters (Fig. 1A). Interestingly, a sequence obtained from a patient originating from

Table 2 | Frequency table of discordant results obtained with the following subtyping tools: REGA HIV-1 & 2 Automated Subtyping Tool version 2.0, REGA HIV Subtyping Tool version 3.0, COMET HIV-1 1.0, jpHMM, SCUEAL, HIVdb: Genotypic Resistance Interpretation Algorithm, Geno2pheno 3.3, STAR, and maximum likelihood phylogenetic analyses on a set of 70 Slovenian sequences.

No	No	REGA 2.0	REGA 3.0	COMET	jpHMM	SCUEAL	HIVdb	Geno2pheno 3.3	Star	Phylogeny
1	9	A (A1)	A (A1)	A1	A1	A1	A/CRF01_AE	01_AE	Unassigned	A
2	2	A (A1)	A (A1)	A1	A1	A1	A/A	A1	Unassigned	A
3	1	A (A1)	A (A1)	A2	A1	A1	CRF01_AE/ CRF01_AE	01_AE	Unassigned	A
4	1	A (A1)	A (A1)	A1	A1	A1	CRF01_AE/ CRF01_AE	01_AE	Unassigned	A
5	1	A (A1)	A (A1)	A1	A1	A1	A/CRF01_AE	01_AE	A	A
6	1	A (A1)	A (A1)	A1	A1	A-ancestral, A1 rec	A/CRF01_AE	01_AE	A	A
7	1	A (A1)	A (A1)	A1	A1	A1	CRF01_AE/A	A1	A	A
8	4	NA	B	B	B	B	B/B	B	B	B
9	3	B	B	B	B	B	B/B	B	Unassigned	B
10	3	B	Rec. of B, D	B	B	B	B/B	B	B	B
11	3	NA	Rec. of B, D	B	B	B	B/B	B	B	B
12	2	B	B	B	B	B, U rec.	B/B	B	B	B
13	1	B	B	B (check for 29_BF)	B	B	B/B	B	B	B
14	1	B	B	B	B	B, D rec.	B/B	B	B	B
15	1	B	B	Unassigned_1;D, 02_AG, B	B	B	B/B	B	B	B
16	1	B	B	B	B	B, D rec.	B/B	B	B	B
17	1	B	Rec. of B, D	B	B	Complex	B/B	B	B	B
18	1	NA	B-like	B	B	B	B/B	B	B	B
19	1	NA	B-like	B	B	B, D rec.	B/B	B	B	B
20	1	NA	B	B	B	B	B/B	B	Unassigned	B
21	1	NA	B, pot. rec.	B	B	B	B/B	B	B	B
22	1	D	D	D	B, D	D	D/D	D	D	D
23	1	D	D	D	B	D	D/D	D	D	D
24	1	D	D	D	D	CRF19	B/D	D	D	D
25	1	F (F1)	F (F1)	F1	F1	F1	D/F	F1	F	F
26	1	G	G	Unassigned_1;D, J, G	G, J	AE, G rec.	A/G	G	Unassigned	G
27	1	G	G	G	G	G	G/G	G	Unassigned	G
28	1	G	G	Unassigned_1;11_cpx, J, G	G, J	G, J, rec.	J/G	G	Unassigned	G
29	1	A (01_AE)	A (01_AE)	01_AE	01_AE	AE	CRF01_AE/ CRF01_AE	01_AE	CRF01	CRF01_AE
30	4	NA	CRF02_AG	02_AG	A1, G	Complex	CRF02_AG/CRF02_AG	02_AG	CRF02	CRF02_AG
31	4	NA	CRF02_AG	02_AG	A1, G	CRF02-like	CRF02_AG/CRF02_AG	02_AG	CRF02	CRF02_AG
32	1	NA	G (02_AG)	02_AG	A2, G	CRF18-like	CRF02_AG/CRF02_AG	02_AG	CRF02	CRF02_AG
33	1	A (A1)	A (A1)	Unassigned_1;02_AG, A2	A1	CRF22	K/A	A1	Unassigned	CRF22
34	1	A (A1)	A (A1)	Unassigned_1;D, 10_CD, 01_AE, B	01_AE	A1, AE rec.	CRF01_AE/ CRF01_AE	01_AE	Unassigned	NA
35	1	A (A1)	A (01_AE)	01_AE	01_AE	CRF15	CRF01_AE/ CRF01_AE	01_AE	Unassigned	NA
36	1	A (A1)	A (A1)	A1	A1	Complex	A/CRF01_AE	01_AE	Unassigned	NA
37	1	A (A1)	A (A1)	Unassigned_1;09_cpx, A1, 45_cpx	A1	Complex	A/CRF01_AE	09_CPX	Unassigned	NA
38	1	A (A1)	A (A1)	Unassigned_1;09_cpx, A1	A1	CRF22	A/CRF01_AE	09_CPX	Unassigned	NA
39	1	A (A1)	A (A1)	Unassigned_1;09_cpx, A1	A1	A, A-ancestral rec.	A/A	A1	Unassigned	NA
40	1	G	G	Unassigned_1;18_cpx, 02_AG, G	G	G	G/G	02_AG	Unassigned	NA
41	1	CRF 19_cpx	CRF 19_cpx	Unassigned_1;D, 11_cpx, G, 20_BG	D	Complex	D/D	D	Unassigned	NA
42	1	NA	NA (G, B)	G (check for 02_AG)	A1, B, G	Complex	CRF02_AG/B	02_AG	Unassigned	NA
43	1	NA	C, pot. rec.	Unassigned_1;F1, A1, 01_AE, B, C	A1, B, C	Complex	C/D	C	Unassigned	NA
44	1	NA	Rec. of F1, 05_DF	Unassigned_1;B, F1	B, F1	B, F1 rec.	B/F	F1	Unassigned	NA
45	1	NA	Rec. of 09_cpx, G	Unassigned_2;G, 09_cpx	A1	Complex	CRF02_AG/K	09_CPX	Unassigned	NA
CA		36	49	53	53	44	39	44	37	

No = number of sequences with the obtained discordant results, Rec. = recombinant, pot. rec. = potential recombinant, NA = not assigned, CA = number of correctly assigned sequences according to phylogeny.

Table 3 | Characteristics of patients included in the study and univariate and multivariate analysis comparing patients infected with non-B HIV-1 subtypes with patients infected with subtype B virus as determined by phylogenetic analyses.

	Total	%	Non-B	%	B	%	Univariate analysis OR (95% CI)	P-value*	Multivariate analysis OR (95% CI)	P-value*
Total	367		54	14.7	313	85.3				
Sex										
Male	320	87.2	28	51.9	292	93.3	0.077 (0.039, 0.155)	**<0.001**		
Female	47	12.8	26	48.1	21	6.7				
Route of infection										
Heterosexual	80	21.8	41	75.9	39	12.5	38.9 (16.3, 92.8)	**<0.001**	22.0 (7.1, 68.0)	**<0.001**
Homo/bisexual	253	68.9	6	11.1	247	78.9	0.030 (0.010, 0.076)	**<0.001**	Ref	
IDU	6	1.6	0	0.0	6	1.9	0 (0, 5.33)	0.811	/	
Mother-to-child transmission	4	1.1	1	1.9	3	1.0	2.09 (0.039, 26.6)	0.903	/	
Hemophiliac	3	0.8	0	0.0	3	1.0	0 (0, 15.2)	>0.999	/	
Unknown	21	5.7	6	11.1	15	4.8			/	
Country of infection										
Slovenia	176	48.0	13	24.1	163	52.1	0.208 (0.086, 0.494)	**<0.001**	Ref	
Europe (other)	36	9.8	3	5.6	33	10.5	0.586 (0.108, 2.08)	0.589	1.44 (0.327, 6.37)	0.629
Eastern Europe	8	2.2	6	11.1	2	0.6	25.6 (4.17, 265)	**<0.001**	11.1 (1.39, 88.3)	**0.023**
Africa	6	1.6	5	9.3	1	0.3	41.2 (4.24, 1,960)	**<0.001**	30.9 (1.85, 517)	**0.017**
Asia	4	1.1	3	5.6	1	0.3	22.9 (1.72, 1,210)	**0.014**	22.2 (1.01, 487)	**0.049**
America	7	1.9	0	0.0	7	2.2	0 (0, 4.88)	0.765	0 (0, 0)	0.999
Unknown	130	35.4	24	44.4	106	33.9			/	
Country of origin										
Slovenia	240	65.4	21	38.9	219	70.0	0.133 (0.053, 0.342)	**<0.001**		
Europe (other)	17	4.7	3	5.6	14	4.5	1.54 (0.269, 5.97)	0.723		
Eastern Europe	6	1.6	6	11.1		0.0	Undefined (9.22, undefined)	**<0.001**		
Africa	2	0.5	2	3.7		0.0	Undefined (1.33, undefined)	**0.031**		
Asia	3	0.8	2	3.7	1	0.3	14.8 (0.734, 873)	0.085		
America	3	0.8	0	0.0	3	1.0	0 (0, 17.1)	>0.999		
Unknown	96	26.2	20	37.0	76	24.3				
Viral load (log ± SD)	4.94 ± 0.96		4.88 ± 0.94		4.96 ± 0.96			0.587		
CD4+ T-cell count (cells/mm³ ± SD)	313 ± 250		252 ± 294		324 ± 241			0.054		
<200 cells/mm³	126	34.3	27	50.0	99	31.6	2.02 (1.07, 3.82)	**0.029**	0.647 (0.191, 2.19)	0.483
≥200 cells/mm³	219	59.7	26	48.1	193	61.7			Ref	
Unknown	22	6.0	1	1.9	21	6.7				
SDRMs										
Yes	13	3.5	1	1.9	12	3.8	0.492 (0.011, 3.49)	0.852		
No	311	84.7	45	83.3	266	85.0				
Unknown	43	11.7	8	14.8	35	11.2				

OR = odds ratio, CI = confidence interval, SD = standard deviation, Ref = reference, IDU = intravenous drug user, SDRMs = surveillance drug resistance mutations.
*Values of p ≤ 0.05 are bolded and considered significant.

Slovenia that reported HIV-1 infection in Kenya was indeed found among the sequences of a cluster originating from Africa (Fig. 1A, Cluster 1). The second cluster included CRF22 sequences, in which a Slovenian sequence was found (Fig. 1A, Cluster 2). The third cluster included A sequences from Africa, Australia, and several European countries, including three from Slovenia (Fig. 1A, Cluster 3). The fourth cluster included CRF01_AE sequences and its recombinants, with five Slovenian sequences (three with subtype CRF01_AE determined with the subtyping tools, two without an assigned subtype; Fig. 1A, Cluster 4). Many of the sequences in this cluster originated from Thailand and other Asian countries, and in fact three Slovenian sequences belonged to Slovenian patients presumably infected in Southeast Asia. The fifth cluster contained A recombinant sequences originating from Africa and Cyprus, with four originating from Slovenia, one of which became infected in Slovenia, whereas the country of infection of the other was unknown (Fig. 1A, Cluster 5). The sixth cluster consisted of 14 Slovenian sequences and other sequences from Eastern Europe and Central Asia (Fig. 1A, Cluster 6). Six of the Slovenian patients were from Ukraine (five reported being infected in Ukraine) and three were from Bosnia and Herzegovina and reported being infected in Russia, Bosnia and Herzegovina, and Slovenia. Two clusters were identified that consisted of only Slovenian sequences, and so additional MCMC analysis was performed. The results of a Slovenian cluster with four subtype A sequences showed a posterior probability value of 1 and a tMRCA determined to be between 1985 and 2008, with a mean in 2001 (Fig. 1, Cluster S2 and Fig. 2). Bayesian analysis for the second cluster of Slovenian sequences with an unknown subtype did not reach convergence, probably due to possible recombinant sequences (Fig. 1A, Cluster S1).

Similarly, a cluster of five Slovenian sequences and a cluster of three Slovenian sequences were observed in a CRF02_AG cluster in the phylogenetic tree of subtype G and its recombinant forms, which included a total of 14 Slovenian sequences (Fig. 1E, Cluster S4 and S5). The control sequences most similar to the Slovenian ones were mainly from West and Central Africa. An additional Bayesian analysis was performed for this cluster, but it did not reach convergence.

The third Slovenian cluster with a sufficient aLRT was that of subtype D (three sequences; Fig. 1C, Cluster S3). Bayesian analysis was not performed for this cluster because it was too small.

Discussion

The aim of this study was to investigate the distribution of HIV-1 subtypes in Slovenia and to identify the characteristics of patients associated with specific subtypes. In order to accurately determine the subtype of infection of all patients, eight different subtyping tools and phylogenetic analyses were used on all available Slovenian sequences for subtype determination. All tools congruently assigned subtypes to 80.9% of Slovenian sequences; however, this was true for only 14.8% of non-B sequences. We found that disagreements between tools were more frequent for: a) subtypes and recombinants for which the analyzed region (*pol*) originated from the same subtype (e.g., subtype A and CRF01_AE; CRF01_AE and CRF15_01B; D and CRF19_cpx; A or CRF01_AE and CRF22_01A1); and b) closely related but epidemiologically distinct groups (e.g., subtypes B and D; A and G), as previously reported (53). Our approach has been to consider manual maximum-likelihood phylogenetic analyses as the gold standard for comparing different subtyping tools, which has its limitations, especially for

recombinant forms. However, in some cases, especially when epidemiologically distinct groups are studied, phylogenetic analyses can be used to clarify misclassifications, as was the case for the majority of Slovenian sequences. Phylogenetic analysis yielded no results for only 12 Slovenian sequences, indicating possible URFs or rarer CRFs. However, because only part of the *pol* region was studied, the diversity among Slovenian sequences could be much higher. The unresolved sequences were further investigated with detailed whole genome analyses (29).

Our assessment of the performance of different subtyping tools is broadly consistent with the results of other studies (32,53–55). COMET, jpHMM, and REGA 3.0 performed best, as previously noted and recommended for detailed epidemiological analyses (5,32). A useful feature of REGA 3.0, a phylogeny-based tool, is that it displays recombination breakpoints, whereas COMET and jpHMM, statistically based methods, perform better when analyzing sequences with a low phylogenetic signal. On the other hand, Geno2pheno has been shown to perform well in identifying pure non-B pure subtypes, a result not observed in our study (55). Indeed, this tool had difficulty identifying both subtype A and complex recombinants. SCUEAL was the only tool that correctly identified CRF22_01A1. Misclassifications occurred largely because the tools did not include CRF22_01A1 in their reference sequence set because this strain had not yet been confirmed in bootscanning analysis at that time. The Slovenian sequence clustered among full genome sequences and several non-full genome ones that were previously characterized as CRF22_01A1 or its recombinants. This cluster also contained partial genome sequences classified as A or CRF01_AE that may have been misclassified because they were analyzed prior to the identification and characterization of CRF22_01A1 in 2010 (56, 57). Many of the control sequences included in our phylogenetic analyses were classified by the original author and using different subtyping methods, and so we cannot fully rely on the subtype indicated in the database when examining phylogenetic trees. Because the Slovenian sequence clustered together with full genomes of this recombinant, we determined that this sequence in this genomic region is indeed CRF22_01A1, whereas we cannot make any claims for the other genomic regions that were not sequenced (http://www.hiv.lanl.gov/content/sequence/HelpDocs/classification.html).

REGA 2.0 yielded a higher number of unassigned sequences because it was designed to emphasize specificity over sensitivity and because a smaller number of CRFs were included in the reference set (32, 55, 58). In contrast, HIVdb performed poorly because of shorter sequences (the protease and reverse transcriptase region sequences are subtyped separately). The subtyping tool STAR left many of the sequences unclassified, as previously reported in another study (55). However, when STAR determined the subtype, it was the only tool that assigned all the subtypes concordantly with manual phylogenetic assignment of the subtype. Our results illustrate that more than one subtyping tool should be used when determining the subtype of sequences because these tools vary in performance when analyzing different subtypes. Performing additional phylogenetic analyses can help resolve divergent results, but this proves to be a tedious and time-consuming task.

In this study, subtype B was the most prevalent subtype, determined in 85% of the sequences. This is consistent with the well-documented high proportion of subtype B in Western and Central Europe and North America (83.3%) (3). In a pan-European study, the lowest prevalence of subtype B was found in Israel (27.9%) and Portugal (39.2%) and the highest in Poland (96.2%) and Slo-
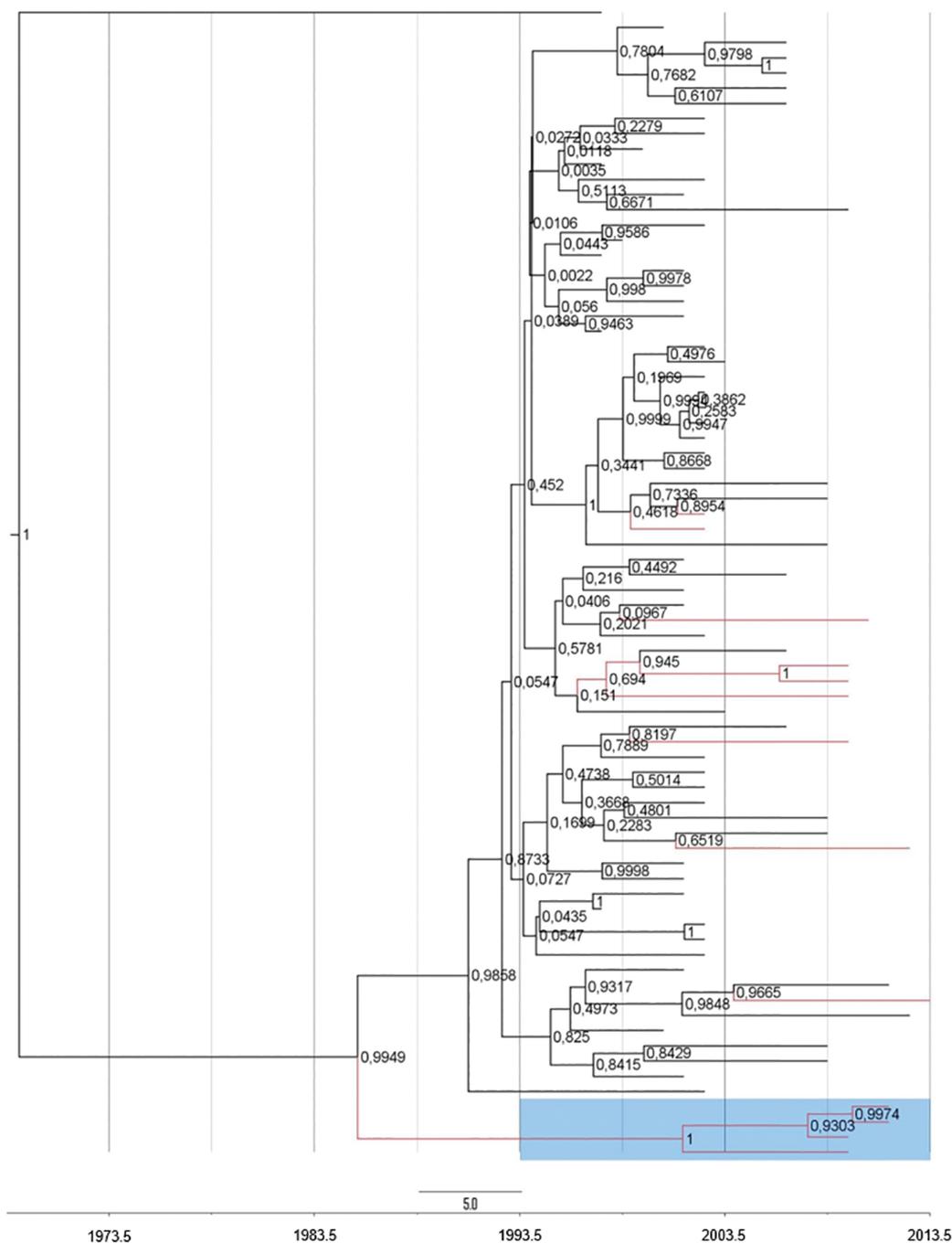
Figure 2 | The maximum clade credibility tree of subtype A sequences (relative to 2253–3554 nucleotides of the HXB2 genome) constructed using a relaxed clock log normal distribution and Bayesian skyline coalescent model in BEAST 2.1.3. Branches of Slovenian sequences are colored red; branches of other similar control sequences are colored black. The cluster of Slovenian sequences with posterior probability value > 0.990 is highlighted in blue. The scale axis is set at 5-year intervals from right to left, starting with the sampling time of the latest sequence (2013.5). Posterior probabilities are shown at the nodes.

venia (93.6%) (59). High prevalence of subtype B was also observed in neighboring countries and some Balkan countries (Croatia, Hungary, Serbia, and Montenegro), in contrast to countries where non-B subtypes are prevalent (e.g., subtype A in Albania and subtype F in Romania) (60).

Of the non-B subtypes, subtype A was most prevalent in Slovenia, accounting for 5% of patients. This subtype was also detected in low numbers in neighboring countries (from 0.5% in Austria to 5% in Croatia), whereas subtype A is highly prevalent worldwide (52.8%–53.4%) in East Africa, Eastern Europe, and Central Asia (59–62). Analysis of transmission clusters revealed that subtype A cases in Slovenia originated from four different countries. Five patients reported being infected in Ukraine, one in Russia, one in Bosnia and Herzegovina, and one in Kenya, which mostly coin-

cides with regions where subtype A is highly prevalent.

In this study, infection with non-B subtypes was found to be significantly associated with heterosexual transmission and infection outside the country (from Eastern Europe, Africa, and Asia), compared to infection with subtype B. Both sexes were almost equally represented in the cohort of non-B–infected patients, whereas subtype B was significantly associated with male sex. This finding is due to the fact that the most at-risk group for HIV-1 infection in Slovenia is MSM. Phylogeny confirmed this because more than half of the large subtype B clusters consisted only of male patients and, in addition, two large clusters contained only patients that reported homosexual route of transmission (28).

Our study shows that at least 31% of non-B subtypes were introduced into the country by immigrants and Slovenians traveling

abroad (Eastern Europe, Africa, and Asia). This could still be an understatement because in 44% of non-B–infected patients the country of infection was unknown. Studies conducted in other countries have also shown that the non-B epidemic is maintained by frequent introduction from many geographic sources—namely, immigration from countries with high HIV-1 prevalence (sub-Saharan Africa, Eastern Europe, South America, and Southeast Asia) and travel throughout the world—resulting in increasing spread of non-B variants (6, 8, 59, 61, 63–69). Interestingly, a study conducted in Croatia, a country neighboring Slovenia, found that in particular seamen (labor migrants) and their steady sexual partners were introducing non-B subtypes into the country (65). Consistent with this, the present study found a significant association between non-B subtypes and infection acquired abroad. Thus, 24% of non-B infections were attributable to immigrants, mainly from Eastern Europe (subtype A), Africa (subtypes C and G), and Asia (CRF01_AE). Subtype A was found predominantly among immigrants from Ukraine, where this subtype is most prevalent (70, 71). We expect to see a significant increase of subtype A as a result of the influx of war immigrants to other European countries in the last few years. Subtype C has also been identified in patients originating from a region where this subtype is prevalent; namely, East Africa (Tanzania) (3). Many studies have reported that immigrants from countries with high HIV-1 prevalence are increasing its diversification in regions where subtype B is most prevalent. For example, in Sicily, Italy, a very low proportion (9.7%) of natives were found to be infected with non-B subtypes (66).

When phylogenetic trees were examined, only five small clusters (consisting of three to five patients) and six transmission pairs of Slovenian non-B sequences were observed, meaning that 43% of Slovenian non-B–infected patients were found without an identified transmission link to another Slovenian patient. On the other hand, a study of HIV-1 patients infected with subtype B in Slovenia showed that the majority of patients were found in large clusters of 10 or more, leaving only 19.3% of patients without an epidemiological link within the country (28). This reconfirms that non-B subtypes are not transmitted further, except to regular sexual partners. This was also observed in a study from Serbia, in which no non-B subtypes were identified within transmission clusters (72). However, some of the studies reported increasing frequency and transmission of HIV-1 non-B subtypes in non-migrant European populations and among MSM (69, 73, 74). Moreover, the countries in which patients reported having acquired HIV infection were confirmed by phylogenetic analysis because Slovenian sequences were found in clusters with GenBank control sequences from the same parts of the world.

Another patient characteristic found to be significantly associated with non-B infection in univariate logistic regression analysis, but not in multivariate logistic regression, was a CD4+ T-cell count of < 200 cells/mm³ at the time of HIV diagnosis. This suggests that patients infected with non-B subtypes are either less aware of the risk of HIV-1 infection and are less likely to be tested regularly or are immigrants from countries where HIV testing is not widely available. This important finding identifies a specific population at higher risk of being diagnosed late in infection. Given the high prevalence of late presenters in many European countries and its importance for epidemic spread, these specific populations should be targeted by campaigns to promote HIV testing.

This study has some limitations. First, the sample size is relatively small, which could reduce the statistical power and flexibility of the effect size. This is mainly due to the small number of

newly diagnosed individuals in the country, as our study sample represented 58% of all HIV patients diagnosed in Slovenia by the end of 2013. Despite the relatively good sample coverage, it is possible that we have missed the less frequent HIV-1 variants in circulation. Next, analyses were performed on partial HIV-1 *pol* sequences > 900 bp long obtained from routine genotypic drug resistance testing and generated for our previous studies. This genomic region has previously been shown to provide sufficient phylogenetic signal for confident subtype identification by manual phylogenetic analysis (75). However, the gold standard for HIV-1 classification remains phylogenetic analysis of full-length genome sequences, which was out of the scope of our study (32). Indeed, some of the unresolved sequences from this study were further investigated with detailed whole genome analyses in another study (29). However, we cannot rule out the possibility that the sequences resolved as specific subtypes or CRFs in this study belong to other subtypes in other regions of the genome that were not sequenced. Sequencing additional genes such as *gag*, *env*, or even the full genome would allow us to more accurately determine the subtype and identify recombination breakpoints. In addition, the subtype assignment of the control sequences obtained from databases and used in the phylogenetic analyses may not have been correct and could have impacted our results. For this reason, we color-coded the sequences derived from the sequenced whole genomes and assumed their subtype assignments with greater confidence. However, because of the different methods used by authors to assign subtypes, even these assignments could have been inaccurate. Considerable diversity was observed within subtypes A and F, highlighting the importance of sub-subtypes. Unfortunately, in this study it was not possible to compare results at the sub-subtype level because subtyping tools used employ different reference sequence sets for analysis. In the example of subtype A, many of them did not include sub-subtypes other than A1 in their reference sets at the time of the analysis, even though at least five other sub-subtypes have been recognized to date (A2, A3, A4, A6 and A7) (Los Alamos HIV Sequence Database). Finally, due to logistical, financial, and COVID-19-related reasons, this study presents relatively old data collected up to 2013. Despite this important limitation, the present study still considerably contributes to the knowledge of the diversity of HIV-1 subtypes. The results of this study demonstrate the importance of HIV-1 subtype surveillance, which may prove to be particularly important due to the significant migration flows to Europe in recent years. As soon as the situation of our research group improves, we aim to fill this research gap and continue surveillance of subtypes in our country.

Conclusions

Among 367 HIV-1 sequences from patients diagnosed in Slovenia, 14.7% belonged to non-B subtypes, with the most prevalent non-B subtypes being subtype A and CRF02_AG. Consistent with previous studies in other European countries, non-B subtype infection was found to be significantly associated with heterosexual route of transmission and infection outside of the country, in contrast to subtype B infection.

The distribution, diversity, and prevalence of HIV-1 variants are important factors in the HIV-1 pandemic and may impact HIV diagnostics, clinical management, and preventive measures. It is essential to continue with systematic surveillance of the HIV-1 subtype distribution and corresponding molecular epidemiology throughout the world.

Funding

References

1. UNAIDS. Global HIV & AIDS statistics – Fact sheet 2022 [Internet]. 2022. [cited 2023 Jan 9 2023]. Available from: https://www.unaids.org/en/resources/fact-sheet.

2. Faria NR, Rambaut A, Suchard MA, Baele G, Bedford T, Ward MJ, et al. HIV epidemiology. The early spread and epidemic ignition of HIV-1 in human populations. Science. 2014;346:56–61.

3. Hemelaar J, Elangovan R, Yun J, Dickson-Tetteh L, Fleminger I, Kirtley S, et al. WHO–UNAIDS Network for HIV Isolation Characterisation. Global and regional molecular epidemiology of HIV-1, 1990–2015: A systematic review, global survey, and trend analysis. Lancet Infect Dis. 2019;19:143–55. Erratum in: Lancet Infect Dis. 2020;20,e27.

4. Hemelaar J. The origin and diversity of the HIV-1 pandemic. Trends Mol Med. 2012;18:182–92.

5. Lessells RJ, Katzenstein DK, de Oliveira T. Are subtype differences important in HIV drug resistance? Curr Opin Virol. 2012;2:636–43.

6. Peeters M, Aghokeng AF, Delaporte E. Genetic diversity among human immunodeficiency virus-1 non-B subtypes in viral load and drug resistance assays. Clin Microbiol Infect. 2010;16:1525–31.

7. Santoro MM, Perno CF. HIV-1 genetic variability and clinical implications. ISRN Microbiol. 2013;481314.

8. Tebit DM, Arts EJ. Tracking a century of global expansion and evolution of HIV to drive understanding and to combat disease. Lancet Infect Dis. 2011;11:45–56.

9. Conroy SA, Laeyendecker O, Redd AD, Collinson-Streng A, Kong X, Makumbi F, et al. Changes in the distribution of HIV type 1 subtypes D and A in Rakai District, Uganda between 1994 and 2002. AIDS Res Hum Retroviruses 2010;26:1087–91.

10. Kiwanuka N, Laeyendecker O, Quinn TC, Wawer MJ, Shepherd J, Robb M, et al. HIV-1 subtypes and differences in heterosexual HIV transmission among HIV-discordant couples in Rakai, Uganda. AIDS. 2009;23:2479–84.

11. Rainwater S, DeVange S, Sagar M, Ndinya-Achola J, Mandaliya K, Kreiss JK, et al. No evidence for rapid subtype C spread within an epidemic in which multiple subtypes and intersubtype recombinants circulate. AIDS Res Hum Retroviruses. 2005;21:1060–5.

12. John-Stewart GC, Nduati RW, Rousseau CM, Mbori-Ngacha DA, Richardson BA, Rainwater S, et al. Subtype C is associated with increased vaginal shedding of HIV-1. J Infect Dis. 2005;192:492–6.

13. Renjifo B, Gilbert P, Chaplin B, Msamanga G, Mwakagile D, Fawzi W, et al. Preferential in-utero transmission of HIV-1 subtype C as compared to HIV-1 subtype A or D. AIDS. 2004;18:1629–36.

14. Easterbrook PJ, Smith M, Mullen J, O'Shea S, Chrystie I, de Ruiter A, et al. Impact of HIV-1 viral subtype on disease progression and response to antiretroviral therapy. J Int AIDS Soc. 2010;13:4.

15. Sacktor N, Nakasujja N, Skolasky RL, Rezapour M, Robertson K, Musisi S, et al. HIV subtype D is associated with dementia, compared with subtype A, in immunosuppressed individuals at risk of cognitive impairment in Kampala, Uganda. Clin Infect Dis. 2009;49:780–6.

16. Abraha A, Nankya IL, Gibson R, Demers K, Tebit DM, Johnston E, et al. CCR5- and CXCR4-tropic subtype C human immunodeficiency virus type 1 isolates have a lower level of pathogenic fitness than other dominant group M subtypes: Implications for the epidemic. J Virol. 2009;83:5592–605.

17. Venner CM, Nankya I, Kyeyune F, Demers K, Kwok C, Chen PL, et al. Infecting HIV-1 subtype predicts disease progression in women of Sub-Saharan Africa. EBioMedicine. 2016;13:305–14.

18. Kiwanuka N, Laeyendecker O, Robb M, Kigozi G, Arroyo M, McCutchan F, et al. Effect of human immunodeficiency virus type 1 (HIV-1) subtype on disease progression in persons from Rakai, Uganda, with incident HIV-1 infection. J Infect Dis 2008;197:707–13.

19. Geretti AM, Harrison L, Green H, Sabin C, Hill T, Fearnhill E, et al. Effect of HIV-1 subtype on virologic and immunologic response to starting highly active antiretroviral therapy. Clin Infect Dis. 2009;48:1296–305.

20. Chaix ML, Seng R, Frange P, Tran L, Avettand-Fenoel V, Ghosn J, et al. Increasing HIV-1 non-B subtype primary infections in patients in France and effect of HIV subtypes on virological and immunological responses to combined antiretroviral therapy. Clin Infect Dis. 2013;56:880–7.

21. Häggblom A, Svedhem V, Singh K, Sönnerborg A, Neogi U. Virological failure in patients with HIV-1 subtype C receiving antiretroviral therapy: An analysis of a prospective national cohort in Sweden. Lancet HIV. 2016;3:e166–74.

22. Klavs I, Kustec T, Berlot L, Kastelic Z, Mrzel M, Kunšič N. Okužba s HIV v Sloveniji: letno poročilo 2021. Nacionalni inštitut za javno zdravje, 2022 [Internet]. [cited 2023 Jan 25]. Available from: https://nijz.si/wp-content/uploads/2022/11/letno_porocilo_hiv_2021_koncno.pdf. Slovenian.

23. Babič DZ, Poljak M, Seme K, Tomažič J, Vidmar L. Molecular epidemiology of HIV-1 subtypes based on analysis of pol sequences in Slovenia, 1996–2005. J Med Virol. 2006;78:997–1002.

24. Babič DZ, Seme K, Tomažič J, Vidmar L, Poljak M. HIV-1 subtype B epidemic and transmission patterns in Slovenia. Coll. Antropol. 2006;30 Suppl 2:25–31.

25. Babič DZ, Zelnikar M, Seme K, Vandamme AM, Snoeck J, Tomažič J, et al. Prevalence of antiretroviral drug resistance mutations and HIV-1 non-B subtypes in newly diagnosed drug-naive patients in Slovenia, 2000–2004. Virus Res. 2006; 118:156–163.

26. Lunar MM, Hošnjak L, Tomažič J, Vidmar L, Karner P, Vovko TD, et al. Prevalence of HIV-1 transmitted drug resistance among patients diagnosed in 2011–2012 in Slovenia (Abstract). Rev Antivir Ther Infect Dis. 2013;2:42–3.

27. Lunar MM, Židovec Lepej S, Abecasis AB, Tomažič J, Vidmar L, Karner P, et al. Short communication: Prevalence of HIV type 1 transmitted drug resistance in Slovenia: 2005–2010. AIDS Res Hum Retroviruses. 2013;29:343–9.

28. Lunar MM, Vandamme AM, Tomažič J, Karner P, Vovko TD, Pečavar B, et al. Bridging epidemiology with population genetics in a low incidence MSM-driven HIV-1 subtype B epidemic in Central Europe. BMC Infect Dis. 2015;15:65.

29. Lunar MM, Mlakar J, Zorec TM, Poljak M. HIV-1 unique recombinant forms identified in Slovenia and their characterization by near full-length genome sequencing. Viruses. 2020;12:63.

30. Alcantara LC, Cassol S, Libin P, Deforche K, Pybus OG, Van Ranst M, et al. A standardized framework for accurate, high-throughput genotyping of recombinant and non-recombinant viral sequences. Nucleic Acids Res. 2009;37:W634–42.

31. de Oliveira T, Deforche K, Cassol S, Salminen M, Paraskevis D, Seebregts C, et al. An automated genotyping system for analysis of HIV-1 and other microbial sequences. Bioinformatics. 2005;21:3797–800.

32. Pineda-Pena AC, Faria NR, Imbrechts S, Libin P, Abecasis AB, Deforche K, et al. Automated subtyping of HIV-1 genetic sequences for clinical and surveillance purposes: Performance evaluation of the new REGA version 3 and seven other tools. Infect Genet Evol. 2013;19:337–48.

33. Struck D, Lawyer G, Ternes AM, Schmit JC, Perez Bercoff D. COMET: Adaptive context-based modeling for ultrafast HIV-1 subtype identification. Nucleic Acids Res. 2014;42:e144.

34. Schultz AK, Zhang M, Leitner T, Kuiken C, Korber B, Morgenstern B, et al. A jumping profile Hidden Markov Model and applications to recombination sites in HIV and HCV genomes. BMC Bioinformatics. 2006;7:265.

35. Kosakovsky Pond SL, Posada D, Stawiski E, Chappey C, Poon AF, Hughes G, et al. An evolutionary model-based algorithm for accurate phylogenetic breakpoint mapping and subtype prediction in HIV-1. PLoS Comput. Biol. 2009;5:e1000581.

36. Liu TF, Shafer RW. Web resources for HIV type 1 genotypic-resistance test interpretation. Clin Infect Dis. 2006;42:1608–18.

37. Gale CV, Myers R, Tedder RS, Williams IG, Kellam P. Development of a novel human immunodeficiency virus type 1 subtyping tool, Subtype Analyzer (STAR): Analysis of subtype distribution in London. AIDS Res Hum Retroviruses. 2004; 20:457–64.

38. Beerenwinkel N, Daumer M, Oette M, Korn K, Hoffmann D, Kaiser R, et al. Geno2pheno: Estimating phenotypic drug resistance from HIV-1 genotypes. Nucleic Acids Res. 2003;31:3850–5.

39. Hall TA. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser. 1999;41:95–8.

40. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol. 2011;28:2731–9.

41. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. Syst Biol. 2010;59:307–21.

42. Rambaut A. Molecular evolution, phylogenetics and epidemiology: FigTree. 2009.

43. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res. 1997;25:3389–402.

44. Karlin S, Altschul SF. Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. Proc Natl Acad Sci U S A. 1990;87:2264–8.

45. Karlin S, Altschul SF. Applications and statistics for multiple high-scoring segments in molecular sequences. Proc Natl Acad Sci U S A. 1993;90:5873–7.

46. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215:403–10.

47. Bouckaert R, Heled J, Kuhnert D, Vaughan T, Wu CH, Xie D, et al. BEAST 2: A software platform for Bayesian evolutionary analysis. PLoS Comput Biol. 2014;10: e1003537.

48. Bouckaert R, Alvarado-Mora MV, Pinho JR. Evolutionary rates and HBV: Issues of rate estimation with Bayesian molecular methods. Antivir Ther. 2013;18:497–503.

49. Drummond AJ, Ho SY, Phillips MJ, Rambaut A. Relaxed phylogenetics and dating with confidence. PLoS Biol. 2006;4:e88.

50. Rambaut A, Drummond AJ. Tracer v1.4. 2007.

51. Rambaut A, Drummond AJ. TreeAnnotator 1.8.0 2013.

52. Bennett DE, Camacho RJ, Otelea D, Kuritzkes DR, Fleury H, Kiuchi M, et al. Drug resistance mutations for surveillance of transmitted HIV-1 drug-resistance: 2009 update. PLoS One. 2009;4:e4724.

53. Gifford R, de Oliveira T, Rambaut A, Myers RE, Gale CV, Dunn D, et al. Assessment of automated genotyping protocols as tools for surveillance of HIV-1 genetic diversity. AIDS. 2006;20:1521–9.

54. Holguin A, Lopez M, Soriano V. Reliability of rapid subtyping tools compared to that of phylogenetic analysis for characterization of human immunodeficiency virus type 1 non-B subtypes and recombinant forms. J Clin Microbiol. 2008;46: 3896–9.

55. Yebra G, de Mulder M, Martin L, Perez-Cachafeiro S, Rodriguez C, Labarga P, et al. Sensitivity of seven HIV subtyping tools differs among subtypes/recombinants in the Spanish cohort of naive HIV-infected patients (CoRIS). Antiviral Res. 2011;89:19–25.

56. Zhao J, Tang S, Ragupathy V, Carr JK, Wolfe ND, Awazi B, et al. Identification and genetic characterization of a novel CRF22_01A1 recombinant form of HIV type 1 in Cameroon. AIDS Res Hum Retroviruses. 2010;26:1033–45.

57. Zhao J, Tang S, Ragupathy V, Gaddam D, Wang X, Zhang P, et al. CRF22_01A1 is involved in the emergence of new HIV-1 recombinants in Cameroon. J Acquir Immune Defic Syndr. 2012;60:344–50.

58. Abecasis AB, Wang Y, Libin P, Imbrechts S, de Oliveira T, Camacho RJ, et al. Comparative performance of the REGA subtyping tool version 2 versus version 1. Infect Genet Evol. 2010;10:380–5.

59. Abecasis AB, Wensing AM, Paraskevis D, Vercauteren J, Theys K, Van de Vijver DA, et al. HIV-1 subtype distribution and its demographic determinants in newly diagnosed patients in Europe suggest highly compartmentalized epidemics. Retrovirology. 2013;10:7.

60. Stanojevic M, Alexiev I, Beshkov D, Gokengin D, Mezei M, Minarovits J, et al. HIV-1 molecular epidemiology in the Balkans: A melting pot for high genetic diversity. AIDS Rev. 2012;14:28–36.

61. Falkensammer B, Doerler M, Kessler HH, Puchhammer-Stoeckl E, Parson W, Duftner C, et al. Subtype and genotypic resistance analysis of HIV-1 infected patients in Austria. Wien Klin Wochenschr. 2007;119:181–5.

62. Grgic I, Lepej SZ, Lunar MM, Poljak M, Vince A, Vrakela IB, et al. The prevalence of transmitted drug resistance in newly diagnosed HIV-infected individuals in Croatia: The role of transmission clusters of men who have sex with men carrying the T215S surveillance drug resistance mutation. AIDS Res Hum Retroviruses 2013;29:329–36.

63. Neogi U, Haggblom A, Santacatterina M, Bratt G, Gisslen M, Albert J, et al. Temporal trends in the Swedish HIV-1 epidemic: Increase in non-B subtypes and recombinant forms over three decades. PLoS One. 2014;9:e99390.

64. Pyne MT, Hackett JJr, Holzmayer V, Hillyard DR. Large-scale analysis of the prevalence and geographic distribution of HIV-1 non-B variants in the United States. J Clin Microbiol. 2013;51:2662–9.

65. Ramirez-Piedad MK, Židovec Lepej S, Yerly S, Begovac J. High prevalence of non-B HIV-1 subtypes in seamen and their sexual partners in Croatia. J Med Virol. 2009;81:573–7.

66. Tramuto F, Maida CM, Bonura F, Perna AM, Vitale F. Dynamics and molecular evolution of HIV-1 strains in Sicily among antiretroviral naive patients. Infect Genet Evol. 2013;16:290–7.

67. von Wyl V, Kouyos RD, Yerly S, Boni J, Shah C, Burgisser P, et al. The role of migration and domestic transmission in the spread of HIV-1 non-B subtypes in Switzerland. J Infect Dis. 2011;204:1095–103.

68. Yebra G, de Mulder M, Martin L, Rodriguez C, Labarga P, Viciana I, et al. Most HIV type 1 non-B infections in the Spanish cohort of antiretroviral treatment-naive HIV-infected patients (CoRIS) are due to recombinant viruses. J Clin Microbiol. 2012;50:407–13.

69. Beloukas A, Psarris A, Giannelou P, Kostaki E, Hatzakis A, Paraskevis D. Molecular epidemiology of HIV-1 infection in Europe: An overview. Infect Genet Evol. 2016;46:180–9.

70. Saad MD, Shcherbinskaya AM, Nadai Y, Kruglov YV, Antonenko SV, Lyullchuk MG, et al. Molecular epidemiology of HIV type 1 in Ukraine: Birthplace of an epidemic. AIDS Res Hum Retroviruses. 2006;22:709–14.

71. Aibekova L, Foley B, Hortelano G, Raees M, Abdraimov S, Toichuev R, et al. Molecular epidemiology of HIV-1 subtype A in former Soviet Union countries. PLoS One. 2018;13:e0191891.

72. Siljic M, Salemovic D, Jevtovic D, Pesic-Pavlovic I, Zerjav S, Nikolic V, et al. Molecular typing of the local HIV-1 epidemic in Serbia. Infect Genet Evol. 2013;19:378–85.

73. Duran Ramirez JJ, Ballouz T, Nguyen H, Kusejko K, Chaudron SE, Huber M, et al. Swiss HIV Cohort Study. Increasing frequency and transmission of HIV-1 non-B subtypes among men who have sex with men in the Swiss HIV cohort study. J Infect Dis. 2022;225:306–16.

74. Bozzi G, Fabeni L, Abbate I, Berno G, Muscatello A, Taramasso L, et al. INACTION study group. Non-B subtypes account for a large proportion of clustered primary HIV-1 infections in Italy. Sex Transm Infect. 2023;99:53–6.

75. Gifford RJ, de Oliveira T, Rambaut A, Pybus OG, Dunn D, Vandamme AM, et al. UK Collaborative Group on HIV Drug Resistance. Phylogenetic surveillance of viral genetic diversity and the evolving molecular epidemiology of human immunodeficiency virus type 1. J Virol. 2007;81:13050–6.