AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Clinical Microbiology®

VIROLOGY

Comparative Evaluation of Subtyping Tools for Surveillance of Newly Emerging HIV-1 Strains

Lavinia Fabeni,[a] Giulia Berno,[a] Joseph Fokam,[b] Ada Bertoli,[c] Claudia Alteri,[c] Caterina Gori,[a] Federica Forbici,[a] Desiré Takou,[b] Alessandra Vergori,[a] Mauro Zaccarelli,[a] Gaetano Maffongelli,[d] Vanni Borghi,[e] Alessandra Latini,[f] Alfredo Pennica,[g] Claudio Maria Mastroianni,[h] Francesco Montella,[i] Cristina Mussini,[e] Massimo Andreoni,[d] Andrea Antinori,[a] Carlo Federico Perno,[a] Maria Mercedes Santoro,[c] The Resistance Study Group

National Institute for Infectious Diseases L. Spallanzani, IRCCS, Rome, Italy[a]; Chantal Biya International Reference for Research on HIV/AIDS Prevention and Management, Yaoundé, Cameroon[b]; University of Rome Tor Vergata, Rome, Italy[c]; University Hospital Tor Vergata, Rome, Italy[d]; Modena University Hospital, Modena, Italy[e]; San Gallicano Dermatological Institute, IRCCS, Rome, Italy[f]; S. Andrea Hospital, Rome, Italy[g]; Polo Pontino-Sapienza University, Latina, Italy[h]; S. Giovanni Addolorata Hospital, Rome, Italy[i]

ABSTRACT HIV-1 non-B subtypes/circulating recombinant forms (CRFs) are increasing worldwide. Since subtype identification can be clinically relevant, we assessed the added value in HIV-1 subtyping using updated molecular phylogeny (Mphy) and the performance of routinely used automated tools. Updated Mphy (2015 updated reference sequences), used as a gold standard, was performed to subtype 13,116 HIV-1 protease/reverse transcriptase sequences and then compared with previous Mphy (reference sequences until 2014) and with COMET, REGA, SCUEAL, and Stanford subtyping tools. Updated Mphy classified subtype B as the most prevalent (73.4%), followed by CRF02_AG (7.9%), C (4.6%), F1 (3.4%), A1 (2.2%), G (1.6%), CRF12_BF (1.2%), and other subtypes (5.7%). A 2.3% proportion of sequences were reassigned as different subtypes or CRFs because of misclassification by previous Mphy. Overall, the tool most concordant with updated Mphy was Stanford-v8.1 (95.4%), followed by COMET (93.8%), REGA-v3 (92.5%), Stanford-old (91.1%), and SCUEAL (85.9%). All the tools had a high sensitivity (≥98.0%) and specificity (≥95.7%) for subtype B. Regarding non-B subtypes, Stanford-v8.1 was the best tool for C, D, and F subtypes and for CRFs 01, 02, 06, 11, and 36 (sensitivity, ≥92.6%; specificity, ≥99.1%). A1 and G subtypes were better classified by COMET (92.3%) and REGA-v3 (98.6%), respectively. Our findings confirm Mphy as the gold standard for accurate HIV-1 subtyping, although Stanford-v8.1, occasionally combined with COMET or REGA-v3, represents an effective subtyping approach in clinical settings. Periodic updating of HIV-1 reference sequences is fundamental to improving subtype characterization in the context of an effective epidemiological surveillance of non-B strains.

KEYWORDS HIV-1, subtypes, circulating recombinant forms, phylogeny, subtyping automated tools, genetic diversity

Human immunodeficiency virus type 1 (HIV-1) is characterized by extensive genetic diversity due to various mechanisms driven by its evolution within an infected individual, thus leading to a broad viral heterogeneity (1–3).

HIV-1 has been divided into four groups: M, O, N, and P (1, 3, 4). The HIV-1 pandemic has been mainly caused by group M (1–3, 5), which is subdivided into 9 subtypes (A to D, F to H, J, and K) and at least 79 circulating recombinant forms (CRFs) (http://www.hiv.lanl.gov/content/sequence/HIV/CRFs/CRFs.html, accessed 25 February 2017) and multiple unique recombinant forms (URFs) widely spread across the globe.

Received 2 May 2017 **Returned for modification** 28 May 2017 **Accepted** 7 July 2017

Accepted manuscript posted online 12 July 2017

Citation Fabeni L, Berno G, Fokam J, Bertoli A, Alteri C, Gori C, Forbici F, Takou D, Vergori A, Zaccarelli M, Maffongelli G, Borghi V, Latini A, Pennica A, Mastroianni CM, Montella F, Mussini C, Andreoni M, Antinori A, Perno CF, Santoro MM, The Resistance Study Group. 2017. Comparative evaluation of subtyping tools for surveillance of newly emerging HIV-1 strains. J Clin Microbiol 55:2827–2837. https://doi.org/10.1128/JCM.00656-17.

Editor Angela M. Caliendo, Rhode Island Hospital

Address correspondence to Lavinia Fabeni, lavinia.fabeni@gmail.com.

L.F., G.B., and J.F. contributed equally to this work.

The dominant HIV-1 strain in North America, Western Europe, and Australia is subtype B. As a result, the great majority of HIV-1 clinical research has been conducted in populations where subtype B predominates. However, this subtype represents only 11% of global HIV-1 infections (6). Of note, among non-B strains, subtypes A, C, and F and CRFs 01_AE and 02_AG are responsible for over 70% of all infections, and there are increasing trends of non-B subtypes and newly emerging CRFs reported in the Western world (7–11), including Italy (12, 13).

Geographical patterns in subtype distribution are changing over time, due to migration and the mixing of populations (9).

In addition to the epidemiological impact, the spread of HIV-1 subtypes is clinically relevant: HIV-1 clades show differences in pathogenesis and resistance pathways, with implications for clinical outcomes, diagnosis, viral quantification, and vaccine development (14–19). With continuous discovery of new HIV-1 strains (https://www.hiv.lanl.gov/content/sequence/HIV/CRFs/CRFs.html) coupled with the increasing phylogeography and phylodynamics of non-B subtypes spreading into Western countries (6–8; see also WHO data at http://www.who.int/gho/hiv/en/), it would be of paramount importance to accurately identify newly emerging strains, both for efficient molecular epidemiological surveillance and for optimal clinical management of patients infected with diverse HIV-1 strains, particularly in Italy, where the migration rate is high (12, 13).

Proper detection and description of clinical HIV-1 samples remain challenging, particularly in the frame of increasing recombinant forms (CRFs and URFs), since algorithms are designed mainly for the B strains (20–22).

Practically, HIV-1 subtyping can be performed through several approaches, among which automated tools are commonly used for clinical purposes (23–25), while molecular phylogeny (Mphy) is commonly used for epidemiological surveillance. To date, Mphy is the gold standard for both epidemiological surveillance and clinical practice (23). However, this gold standard is not widely used in routine practice due to its complexity (it is manually performed, cumbersome, and time-consuming and requires skills in data interpretation). In addition, Mphy might not be updated regularly with reference sequences.

Routine subtyping is often based on automated tools because they are user friendly, speedy, and free of charge (24, 26). However, they have considerable limitations compared to Mphy especially in assigning non-B variants: outputs of different tools are usually in disagreement (26–28), their algorithms are not regularly updated, and they have only a limited number of CRFs in the reference data set. In routine clinical practice, the most commonly used automated tools are statistically based use of partial matching compression algorithms (*co*ntext-based *mo*deling for *e*xpeditious *t*yping [COMET] (29), a similarity-based tool (Stanford HIV drug resistance database [Stanford]), and phylogenetics-based tools (REGA and *s*ubtype *c*lassification *u*sing *e*volutionary *al*gorithms [SCUEAL]).

We thus aimed at assessing the added value of updated Mphy (using recently available HIV-1 reference sequences) and determining the performance of four commonly used automated subtyping tools (COMET, Stanford, REGA, and SCUEAL) against updated Mphy, in order to propose a highly reliable approach for subtyping in routine clinical practice.

(This work was presented in part as a poster at the VIIIth Italian Conference on AIDS and Antiviral Research Workshop, 6 to 8 June 2016, Milan, Italy [30].)

RESULTS

Mphy subtype assignment. Of the total 13,116 HIV-1 *pol* sequences analyzed and based on new Mphy, B was the most prevalent subtype (73.4%), followed by CRF02_AG (7.9%), C (4.6%), F1 (3.4%), A1 (2.2%), G (1.6%), CRF12_BF (1.2%), and other subtypes (5.7%).

By comparing old Mphy with new Mphy, the overall concordance was 97.7% between the two approaches, with a subtyping agreement of 99.8% and 88.2% for B and non-B subtypes, respectively. The 297 (2.3%) discrepant sequences reassigned as

TABLE 1 Reassigned HIV-1 subtypes using new Mphy[a]

Old Mphy subtype/CRF (*n*)	New Mphy subtype(s)/CRF(s) (*n*)
F1 (52)	CRF71_BF (27), CRF72_BF (12), CRF40_BF (4), CRF70_BF (3), CRF05_DF (2), CRF12_BF (1), CRF39_BF (1), CRF42_BF (1), CRF47_BF (1)
CRF02_AG (43)	CRF36_cpx (18), CRF37_cpx (10), CRF06_cpx (8), CRF09_cpx (2), CRF01_AE (1), CRF20_BG (1), CRF25_cpx (1), CRF43_02G (1), CRF63_02A1 (1)
B (41)	CRF39_BF (9), CRF03_AB (4), CRF12_BF (4), CRF42_BF (4), CRF51_01B (4), CRF44_BF (3), CRF17_BF (2), CRF28_BF (2), CRF38_BF (2), CRF08_BC (1), CRF15_01B (1), CRF23_BG (1), CRF40_BF (1), CRF46_BF (1), CRF47_BF (1), D (1)
CRF12_BF (27)	CRF05_DF (6), CRF40_BF (6), CRF72_BF (6), CRF29_BF (2), CRF42_BF (2), CRF03_AB (1), CRF28_BF (1), CRF38_BF (1), CRF39_BF (1), CRF47_BF (1)
A1 (18)	CRF22_01A1 (11), A2 (3), CRF35_AD (3), CRF02_AG (1)
CRF17_BF (14)	CRF12_BF (5), CRF40_BF (3), CRF38_BF (2), CRF47_BF (2), CRF39_BF (1), CRF72_BF (1)
G (13)	CRF43_02G (7), CRF02_AG (3), CRF06_cpx (1), CRF36_cpx (1), CRF32_06A1 (1)
CRF01_AE (12)	F22_01A1 (11), CRF02_AG (1)
CRF28_BF (11)	CRF29_BF (5), CRF40_BF (3), CRF12_BF (1), CRF42_BF (1), CRF60_BC (1)

[a]The table reports subtypes obtained by using old Mphy and reclassified by new Mphy. Only reassigned subtypes with a prevalence of >3% are reported.

different subtypes were due to newly available CRF reference sequences in the HIV sequence databases. Of note, new Mphy allowed the identification of more CRFs than old Mphy. The most prevalent reclassified subtypes, earlier assigned by old Mphy, were F1 (n = 52 [17.5%], 50 of them reclassified as BF recombinants), CRF02_AG (n = 43 [14.5%], 41 of them reclassified as more complex recombinants), and B (n = 41 [13.8%], 29 of them reclassified as BF recombinants) (Table 1).

Concordance between automated subtyping tools and new Mphy. HIV-1 subtypes defined by new Mphy were compared with those provided by the four rapid subtyping tools (COMET, SCUEAL, REGA-v3, and Stanford [Stanford-old and Stanford-v8.1]). Overall, the tool most concordant with new Mphy was Stanford-v8.1 (95.4%), followed by COMET (93.8%), REGA-v3 (92.5%), Stanford-old (91.1%), and SCUEAL (85.9%), as shown in Fig. 1A.

In particular, concordance with new Mphy was excellent for subtype B with all tools (both Stanford versions, 99.8%; COMET, 99.3%; REGA-v3, 98.8%; and SCUEAL, 98.1%) but was lower for non-B subtypes (Stanford-v8.1, 84.0%; COMET, 78.5%; REGA-v3, 74.9%; Stanford-old, 67.0%; and SCUEAL, 52.6%) (Fig. 1B and C). Thus, for both B and non-B subtypes, the highest concordance rates were reported with Stanford-v8.1.

Sensitivity and specificity of automated subtyping tools for HIV-1 pure subtypes. Comparing new Mphy HIV-1 subtypes with those provided by the four rapid subtyping tools, all the tools had a high sensitivity (≥98.0%) for subtype B (n = 9,627) (Table 2). In particular, the highest sensitivity was observed for Stanford-v8.1 (99.6%), followed by COMET (99.5%), Stanford-old (99.0%), and SCUEAL and REGA-v3 (both 98.0%). The specificity of each tool was ≥97.9%, except for SCUEAL (95.7%) (Table 2).

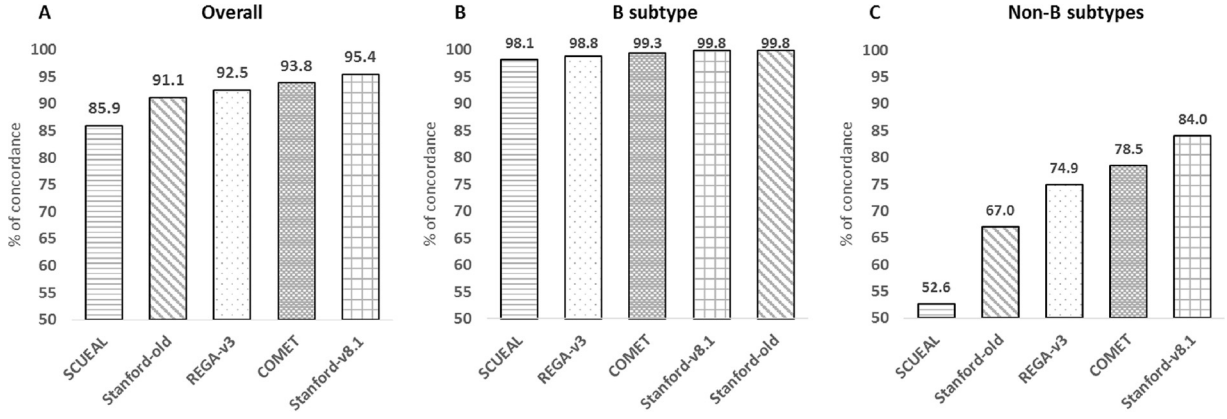

FIG 1 Concordance between HIV-1 automated subtyping tools and new Mphy in the overall population (A), for B subtype (B), and for non-B subtypes (C).

TABLE 2 Performance of HIV-1 subtyping tools[a]

Subtype or CRF	Total no.	% sensitivity (95% CI[b])					% specificity (95% CI)				
		COMET	SCUEAL	REGA-v3	Stanford-v8.1	Stanford-old	COMET	SCUEAL	REGA-v3	Stanford-v8.1	Stanford-old
B	9,627	**99.5 (99.3–99.6)**	**98.0 (97.9–98.2)**	**98.0 (97.9–98.2)**	**99.6 (99.4–99.6)**	**99.0 (98.8–99.1)**	**98.4 (98.0–98.6)**	**95.7 (95.2–96.2)**	**98.2 (97.7–98.5)**	**98.4 (98.1–98.7)**	**97.9 (97.5–98.2)**
CRF02_AG	1,037	**90.3 (89.1–91.1)**	18.1 (17.1–18.8)	62.6 (61.5–63.3)	**98.5 (97.6–99.0)**	**95.4 (94.5–96.1)**	**99.7 (99.6–99.8)**	**99.8 (99.8–99.9)**	**99.8 (99.7–99.9)**	**99.6 (99.6–99.7)**	**99.8 (99.7–99.9)**
C	606	**91.9 (90.8–92.5)**	**92.2 (91.0–93.0)**	**92.7 (91.5–93.5)**	**92.6 (91.5–93.2)**	**98.7 (97.6–99.4)**	**100.0 (99.9–100.0)**	**99.9 (99.9–100.0)**	**99.9 (99.9–100.0)**	**100.0 (99.9–100.0)**	**99.8 (99.7–99.8)**
F1	446	84.7 (82.1–87.1)	87.9 (85.3–90.1)	**91.2 (88.7–93.4)**	**99.1 (97.6–99.7)**	70.4 (67.2–73.2)	**99.5 (99.4–99.6)**	**99.5 (99.4–99.5)**	**99.2 (99.1–99.3)**	**99.1 (99.1–99.2)**	**99.4 (99.2–99.5)**
A1	283	**92.3 (89.4–94.6)**	76.9 (72.2–81.1)	**90.6 (87.2–93.2)**	81.8 (78.9–84.0)	29.4 (27.4–29.9)	**99.7 (99.7–99.8)**	**98.7 (98.6–98.8)**	**99.6 (99.5–99.7)**	**99.9 (99.8–99.9)**	**100.0 (99.9–100.0)**
G	214	**96.2 (93.2–98.2)**	88.3 (83.8–91.9)	**98.6 (95.9–99.6)**	59.8 (55.6–63.0)	**93.0 (89.2–95.7)**	**99.8 (99.7–99.8)**	**99.4 (99.399.5)**	**99.7 (99.6–99.7)**	**99.8 (99.8–99.9)**	**99.6 (99.5–99.6)**
CRF12_BF	162	61.1 (58.1–61.7)	22.2 (19.2–22.8)	48.1 (45.4–48.1)	78.4 (73.8–81.7)	NA	**100.0 (100.0–100.0)**	**100.0 (100.0–100)**	**100.0 (100.0–100.0)**	**99.9 (99.8–99.9)**	NA
CRF01_AE	98	**92.9 (87.1–96.5)**	75.5 (66.7–82.8)	83.7 (77.2–83.3)	**98.0 (92.5–99.6)**	**99.0 (93.9–99.9)**	**99.9 (99.9–99.9)**	**99.4 (99.4–99.5)**	**99.9 (99.9–99.9)**	**99.6 (99.6–99.6)**	**99.5 (99.5–99.5)**
CRF06_cpx	65	81.5 (75.0–83.0)	37.0 (29.4–40.3)	89.2 (81.8–93.6)	**95.4 (88.3–98.7)**	NA	**100.0 (100.0–100.0)**	**100.0 (100.0–100.0)**	**100.0 (99.9–100.0)**	**99.9 (99.9–100.0)**	NA
D	39	89.7 (79.7–93.9)	74.4 (62.1–81.9)	56.4 (45.3–60.6)	**94.9 (87.5–99.9)**	**92.3 (81.5–97.7)**	**100.0 (100.0–100.0)**	**100.0 (99.9–100)**	**100.0 (100.0–100.0)**	**100.0 (99.9–100)**	**100.0 (99.9–100.0)**
CRF22_01A1	35	11.4 (4.6–11.4)	68.6 (56.0–74.8)	0.0 (0.0–0.0)	82.9 (68.8–92.1)	NA	**100.0 (100.0–100.0)**	**100.0 (99.9–100.0)**	**100.0 (100.0–100.0)**	**99.9 (99.9–99.9)**	NA
CRF40_BF	31	0.0 (0.0–0.0)	3.3 (0.2–11.2)	0.0 (0.0–0.0)	3.23 (0.2–6.3)	NA	**100.0 (100.0–100.0)**	**100.0 (100.0–100.0)**	**100.0 (100.0–100.0)**	**100.0 (100.0–100.0)**	NA
CRF11_cpx	28	78.6 (66.4–78.6)	75.0 (59.7–84.0)	89.3 (74.4–96.8)	**92.9 (80.7–96.2)**	NA	**100.0 (100.0–100.0)**	**100.0 (99.9–100)**	**100.0 (99.9–100)**	**100.0 (100.0–100.0)**	NA
F2	25	72.0 (56.4–78.5)	88.8 (73.3–94.4)	60.0 (45.2–63.8)	**100.0 (87.1–100)**	NA	**100.0 (100.0–100.0)**	**100.0 (100.0–100.0)**	**100.0 (100.0–100.0)**	**100.0 (100.0–100.0)**	NA
CRF49_cpx	23	0.0 (0.0–0.0)	NA	NA	34.8 (20.6–38.9)	NA	**100.0 (100.0–100.0)**	NA	**100.0 (100.0–100.0)**	NA	NA
CRF29_BF	21	57.1 (36.8–74.5)	14.3 (4.0–29.0)	0.0 (0.0–7.6)	**57.1 (37.3–73.3)**	NA	**99.9 (99.9–99.9)**	**100.0 (99.9–100.0)**	**100.0 (100.0–100.0)**	**99.9 (99.9–100)**	NA
CRF36_cpx	20	0.0 (0.0–0.0)	5.0 (0.3–5.0)	NA	**100.0 (100.0–100.0)**	NA	**100.0 (100.0–100.0)**	**100.0 (100.0–100.0)**	**100.0 (100.0–100.0)**	NA	NA

[a]The table reports sensitivity and specificity of subtypes present in at least 20 sequences analyzed. The values with automated tool performances of >90% are shown in bold. The not-applicable (NA) values indicate the absence of the reference sequences in the automated tool algorithm.

[b]CI, confidence interval.

Considering subtype C ($n = 606$), all the tools had a high sensitivity (≥91.9%) and a high specificity (≥99.8%). For the other non-B pure subtypes (F1, F2, D, A1, and G), Stanford-v8.1 showed the highest sensitivity for F1, F2, and D subtypes (99.1%, 100.0%, and 97.4%, respectively) but not for A1 and G subtypes (81.8% and 59.8%, respectively), which were better assigned by COMET (92.3%) and REGA-v3 (98.6%), respectively. Of note, the specificity for these pure non-B subtypes was ≥99.1% for each tool (Table 2).

Sensitivity and specificity of automated subtyping tools for HIV-1 CRFs. The Stanford-v8.1 tool had the best sensitivity in recognizing CRFs (Table 2). In particular, the sensitivity was ≥92.9% on five different recombinants: CRF01_AE (98.0%), CRF02_AG (98.5%), CRF06_cpx (95.4%), CRF11_cpx (92.9%), and CRF36_cpx (100.0%). COMET also showed a desirable sensitivity on CRF02_AG (90.3%) and CRF01_AE (92.9%). The specificity of each tool was ≥99.6% for all the CRFs (Table 2).

Performance of Stanford-v8.1 versus Stanford-old versions. To evaluate the added value in subtyping performance of the new version of the Stanford tool (Stanford-v8.1), potential subtype differences between the tool versions were evaluated. Overall, concordance between Stanford-old versions and Stanford-v8.1 was 90.2%; in particular, the concordance was 99.0% and 66.2% on B and non-B subtypes, respectively. In comparing the sensitivities of the different versions, Stanford-v8.1 significantly improved the capacity in assigning subtypes B (from 99.0% to 99.6%, $P < 0.0001$), F1 (from 70.4% to 99.1%, $P < 0.0001$), A1 (from 29.4% to 81.8%, $P < 0.0001$), and CRF02_AG (from 95.4% to 98.5%, $P < 0.0001$). However, this capacity significantly decreased in assigning subtypes C (from 98.7% to 92.6%, $P < 0.0001$) and G (from 93.0% to 59.8%, $P < 0.0001$), which were wrongly classified as recombinant forms. There were no significant differences in sensitivity between the two tools for assigning CRF01_AE (from 99.0% to 98.0%, $P = 1$) (Table 2). Of note, as described in Materials and Methods, the majority of CRFs are now more accurately identified by Stanford-v8.1 than Stanford-old versions, due to the addition of CRF reference sequences in the Stanford HIVdb.

Proposal for practical subtyping. Based on our results, we propose the following algorithm for subtyping in routine clinical practice (Fig. 2). This strategy is based on the desirable performance (≥90%) of automated subtyping tools with respect to new Mphy. Briefly, use Stanford-v8.1 as a screening tool for subtyping. If the output leads to pure subtypes (B, C, D, or F) or CRFs (01_AE, 02_AG, 06_cpx, 11_cpx, or 36_cpx), consider assigning the respective strain. If the output leads to subtype A1 or G, confirm with COMET or REGA-v3, respectively. If the output is different from the above, consider performing Mphy for a subtype inference. If the output leads to all other non-B strains, consider performing Mphy for a subtype inference (Fig. 2).

DISCUSSION

In this study, we aimed at assessing the necessity of updating MPhy, as well as the current performance of commonly used automated subtyping tools, in an era of changing molecular epidemiology in Western countries, including Italy (7, 12, 13, 15, 16, 28). Of note, the Italian clinical context is experiencing a rapid change in circulating HIV-1 strains, mainly due to a diversified migration system, ongoing infection in the high-risk populations (mostly men having sex with men [MSM]), and specific transmission clusters (12, 13). The large number of PR/reverse transcriptase (RT) sequence data generated routinely from HIV-1-infected patients ($n = 13,116$) and used for the present assessment offers a greater representativeness of the present findings for settings with closely related HIV epidemiological features. Of note, though subtype B remains highly prevalent in our data set, there is a growing rate of non-B subtypes and recombinants, thus confirming the need for regular surveillance (7, 8, 12). Based on updated phylogeny (up to 79 CRFs at the moment), the 2.3% overall discordance reported from old Mphy highlights the importance of continuously updating Mphy algorithms for an accurate surveillance of newly emerging strains in countries sharing similar challenges as Italy as well as in countries where non-B strains are predominant (6, 13, 15, 24, 25, 31). The high discrepancy observed with non-B subtypes (11.8% versus only 0.2% with

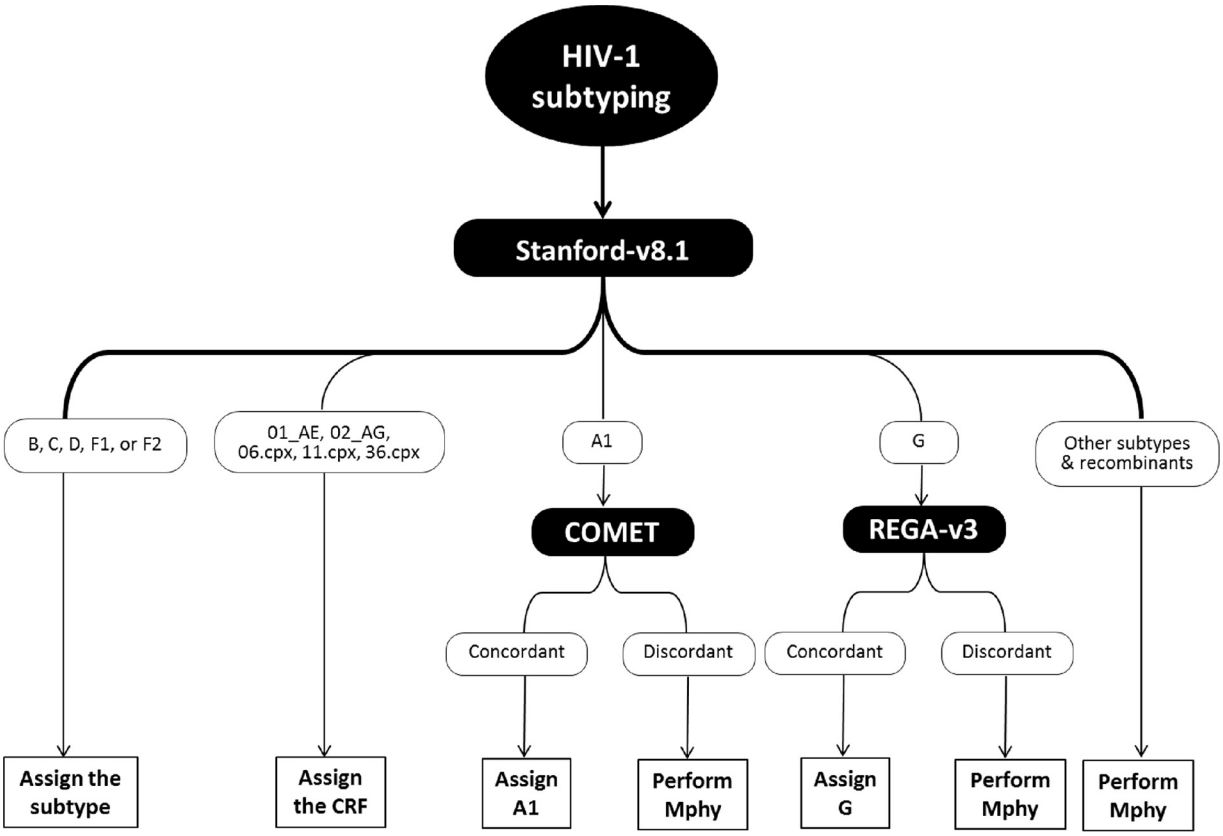

FIG 2 Flow chart of a practical HIV-1 subtyping approach. This flow chart represents the HIV-1 subtyping practical approach through online subtyping tools, proposed in the case that Mphy cannot be used in routine practice due to its complexity.

B) underscores the need for greater considerations of non-B surveillance (7, 8, 12, 13). This discrepancy observed between old Mphy and new Mphy highlights the need for periodic updating of reference sequences for phylogenetic analyses (23, 24). In particular, new Mphy included 530 new reference sequences with respect to old Mphy, of which 25.3% belonged to CRFs, and this could explain the 2.3% overall discordance between old Mphy and new Mphy.

HIV-1 subtypes defined by new Mphy were compared with those provided by the four rapid subtyping tools (COMET, SCUEAL, REGA, and Stanford). These tools have different characteristics in inferring subtypes (26). In particular, because REGA-v3 and SCUEAL are phylogenetics-based tools, their computation times are longer than those of other tools that are based on a statistical (COMET) or similarity (Stanford HIVdb) approach (24, 26, 28, 29).

The overall concordance between automated subtyping tools and new Mphy reveals an acceptable performance of all tools (≥90%), except for SCUEAL (85.9%). This finding highlights the limited performance of the current SCUEAL for subtype assignment in the context of a growing HIV-1 molecular epidemic in Western countries (24, 27).

Regarding B versus non-B subtype assignment, an excellent concordance with all automated subtyping tools (≥98.1%, including SCUEAL) was reported for the B subtype. This observation confirms the suitability in B subtyping, largely due to an initial design of algorithms based on subtype B reference sequences (15, 32). In contrast to the B subtype, none of the automated tools achieved the desirable concordance for non-B subtypes (ranging from 84.0% with Stanford-v8.1 down to 52.6% with SCUEAL). However, although phylogenetic analysis remains the gold standard for subtyping, our findings highlight that HIVdb in its updated version (Stanford-v8.1) offers the highest accuracy compared to the other routinely used subtyping tools, for both B and non-B

subtypes, thus serving as a convenient approach for subtype screening. The higher number of CRF reference sequences used by Stanford-v8.1 for subtyping and thus the better representation of the CRF circulation worldwide may explain the higher accuracy in subtyping of this tool with respect to the others. Interestingly, all the CRF reference sequences used by Stanford-v8.1 are included in our Mphy analysis, thus explaining the high concordance between these two methodologies and confirming Stanford-v8.1 as the best tool for subtype assignment.

Moreover, these findings confirm that regular updating improves sensitivity in assigning HIV-1 strains, in an era of ongoing recombination events. This is in contrast with our preliminary analyses conducted on about one-third of sequences and before the release of the new version of Stanford HIVdb (30) and with other previous reports on the performance of several subtyping tools (24, 26, 28), which is normal because our analysis was performed on recently updated tools and with more reference sequences.

A high sensitivity for subtype B assignment ($\geq$98.0%) was reported with all tools. This is in accordance with previous findings and underlines the consistent reliability for subtype B surveillance using current automated tools (24, 26, 28). Though the sensitivities on subtype B were highly similar for all automated tools, SCUEAL showed the lowest specificity (95.7%) by wrongly assigning some CRFs to subtype B. The current overestimation of subtype B suggests caution when assigning this viral strain with SCUEAL, thus giving preference for B subtyping to the other algorithms (Stanford HIVdb, COMET, and REGA-v3).

Regarding subtype C, the excellent and similar sensitivities and specificities reported for all the automatic subtyping tools indicate the suitability of all tools in discriminating subtype C viruses (24, 33) and their possible routine use in settings with subtype C predominance like southern Africa, eastern Africa, India, Nepal, and part of China (2, 5, 34). On one hand, Stanford-v8.1 showed excellent performance in sensitivity (97.4% to 100%) on other pure subtypes (D, F1, and F2) in contrast to other rapid tools (reporting $<$90% sensitivity); this indicates that Stanford-v8.1 might be acceptable in assigning D or F strains. On the other hand, only COMET or REGA-v3 might be acceptable in assigning viruses as subtype A1 or G, respectively (27, 29). Regarding the sensitivity to recombinant forms, Stanford-v8.1 appeared highly reliable in assigning CRF01_AE, CRF02_AG, CRF06_cpx, CRF11_cpx, and CRF36_cpx. COMET showed an acceptable (though lower than Stanford-v8.1) performance only on CRF01_AE and CRF02_AG. Thus, Stanford-v8.1 possesses the most reliable algorithm on major CRFs compared with the other evaluated rapid tools. As this trend is consistent even with other CRFs below target performance, the revolutionary update in Stanford-v8.1 now makes this tool the reference automated subtyping tool, coupled with its wide use in clinical practice. As described in Materials and Methods, the majority of CRFs are now more accurately identified by Stanford-v8.1 than by older Stanford versions, due to the addition of CRF reference sequences in HIVdb. Thus, the significant improvements in Stanford-v8.1 underscore the relevance of regular updating of both automated tools and advanced phylogenetic approaches (23, 27, 28). In spite of its great performance, Stanford-v8.1 has some limitations: misclassification of CRFs 01_AE, 14_BG, 15_01B, and 46_BF (due to the absence of recombination breakpoints in the *pol* region), which are subtypes A, G, AE, and F in the *pol* region, respectively (26). Moreover, Stanford-v8.1 wrongly classifies 40.0% of G subtype strains as CRF43_02G (90.7%), CRF02_AG (7.0%), or CRF06_cpx (2.3%) and 16.7% of A subtype strains as CRF01_AE (61.8%), CRF22_01A1 (19.1%), CRF09_cpx (8.5%), or CRF02_AG (10.6%), important considerations for those areas where non-B strains are prevalent.

Through an evidence-based approach, we propose an algorithm that could be used to facilitate the subtyping process during routine clinical practice. Such a subtyping strategy becomes more relevant for settings where a large number of sequences are routinely generated, with less expertise in advanced phylogeny, and limited resources (Fig. 2).

There are some limitations observed with this study. First, we used *pol* sequences for subtyping (35), instead of the full genome, which could provide a better assignment.

This could be a problem especially for those CRFs with recombination breakpoints outside the *pol* region, as reported above.

However, it should be considered that *pol* is the most extensively investigated region for clinical and diagnostic purposes, and the more suitable way to assign a correct subtype/CRF is to trim the *pol* sequences from full-length genomes of pure subtypes and CRFs from the HIV sequence database, as previously reported by Pineda-Peña et al. (26).

Moreover, the underrepresentativeness of some non-B strains (though with a higher number than previous reports) warrants further investigations in those settings. Finally, only the most commonly used rapid tools were included in the present analysis, thus missing possible revisions of other existing tools.

In conclusion, Mphy remains the gold standard method for an accurate HIV-1 subtyping. In the case that Mphy cannot be used in routine practice due to its complexity, online subtyping tools can be a valid option for the subtype characterization. Though rapid subtyping tools have various performances with respect to Mphy, Stanford-v8.1 appears most reliable for rapid subtyping of both pure and recombinant strains. Thus, for practical use in routine clinical practice, the usage of Stanford-v8.1, occasionally combined with COMET or REGA-v3, represents an effective subtyping approach in clinical settings.

Periodic updating of algorithms together with the latest HIV-1 reference sequences is fundamental to improve HIV-1 subtype characterization in the context of effective epidemiological surveillance of non-B HIV-1 strains.

MATERIALS AND METHODS

Study population. The present study was conducted on a data set of 13,116 HIV-1 *pol* sequences (containing the full-length protease [PR] and the first 300/335 reverse transcriptase [RT] codons), performed for routine clinical purposes from 1997 to 2015 in three reference laboratories in Italy (National Institute for Infectious Diseases L. Spallanzani, IRCCS, Rome; University of Rome Tor Vergata, Rome; and Modena University Hospital, Modena) and in one in Cameroon (Chantal Biya International Reference for Research on HIV/AIDS Prevention and Management, Yaoundé) and then collected in an anonymous database. HIV-1 *pol* genotype analyses were performed on plasma samples, as previously described (36–38).

Mphy analysis. For each sequence, HIV-1 subtype was determined by molecular phylogeny (Mphy). In particular, *pol* sequences were aligned by using Clustal X, with full-length reference sequences of HIV-1 subtypes and CRFs retrieved from the HIV sequence databases, available at https://www.hiv.lanl.gov/content/sequence/NEWALIGN/align.html, using at least 10 reference sequences for each subtype/CRF, for a total of 3,923 sequences. The alignment type chosen was the "all complete sequences" in the fasta format, and only one sequence per patient was included. Then, the complete alignment was manually trimmed from full-length genomes to the PR/RT region, and gaps were removed from the final alignment, by using BioEdit software version 7.2.5.

Two different Mphy analyses were conducted. (i) The first Mphy (referred to as old Mphy) was conducted with a data set of reference sequences available until 2014; (ii) the second one (referred to as new Mphy) was performed with a data set of reference sequences updated in 2015, to evaluate possible discrepancies with previous subtype assignments. In particular, in the 2015 analysis we had 530 more reference sequences than those available in the 2014 analysis. Subtype or CRF assignments were achieved by constructing phylogenetic trees using the neighbor-joining (NJ) method (39). Regarding old Mphy, until 2009, the F84 substitution model with both NJ and maximum likelihood (ML) tree building methods was used (40), performed by PAUP software (http://paup.sc.fsu.edu/). From 2010 until now, phylogenetic analyses were conducted using MEGA (version 5, from 2010 to 2013; version 6, from 2014 until now), based on the Kimura 2-parameter (K2P) model (41). The reliability of the branching orders was assessed by bootstrap analysis of 1,000 replicates. To confirm subtype classification, an ML tree with 1,000 bootstrap replicates, a general time-reversible (GTR) nucleotide substitution model with gamma distribution among site heterogeneity, and a proportion of invariable sites ($G+I+\Gamma$) were inferred. A sequence that clustered monophyletically inside a clade with a bootstrap support value of $\geq$70% was assigned to that clade; otherwise, the sequence was analyzed for recombination using RDP4 software. Recombination events detected are displayed graphically, with statistical evidence provided, and recombination events are also drawn on phylogenetic trees constructed from proposed recombinant regions. For the sequences without any signal for recombination, the sequence was assigned the clade with the highest similarity in RDP4 with a bootstrap support value of $\geq$70%; differently, in the presence of a recombination signal, the sequence was identified as a unique recombinant form (URF) (Fig. 3) (26, 42). The trees were rooted using midpoint rooting by FigTree software version 1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/).

HIV-1 automated subtyping assignment. The entire *pol* sequence data set was analyzed using four automatic tools: COMET (https://comet.lih.lu/), REGA-v3 (http://dbpartners.stanford.edu:8080/RegaSubtyping/stanford-hiv/typingtool/), SCUEAL (http://www.datamonkey.org/dataupload_scueal.php),

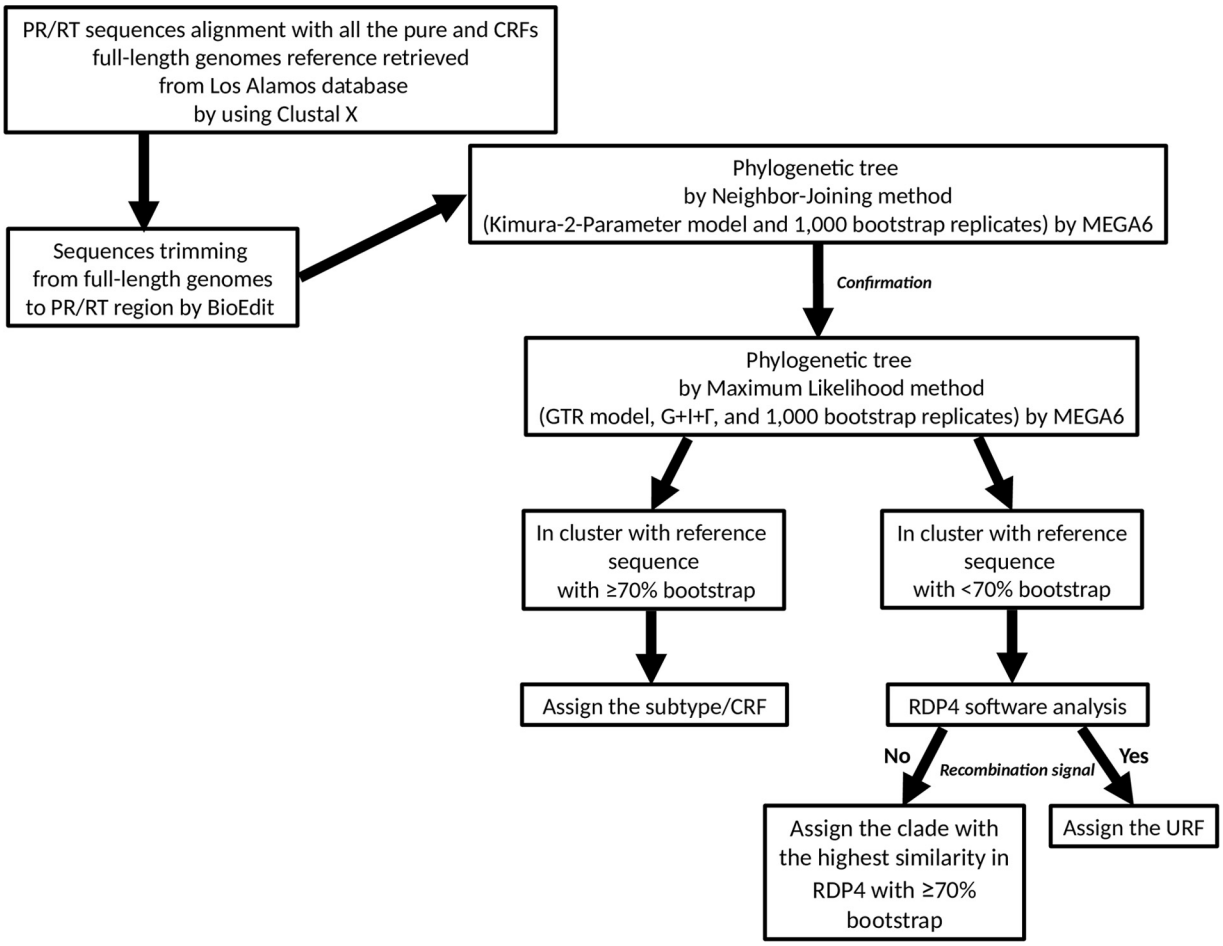

FIG 3 Flow chart of Mphy. Shown are the main steps of HIV molecular phylogeny (Mphy) used for subtype assignment.

and Stanford HIVdb (versions 1.0 to 7.0, referred to as Stanford-old, or Stanford HIVdb 8.1 version, referred to as Stanford-v8.1). Regarding Stanford HIVdb, until 14 September 2016, the only reference sequences present in Stanford-old were the pure subtypes A, B, C, D, F, and G and the two CRFs 01_AE and 02_AG. The Stanford algorithm has now been updated, with the addition of HIV-1 group M reference sequences for CRFs up to 65. In addition to the current update in Stanford HIVdb, the *pol* gene is now concatenated to encompass the entire PR/RT/IN sequence. Features of the updated reference sequences in Stanford-v8.1, used in the present subtyping assessment, are now available in the open-source project for subtyping (https://github.com/hivdb/hiv-genotyper/blob/master/src/main/resources/HIVGenotypeReferences.json).

Concordance, sensitivity, and specificity. Since new Mphy was considered the gold standard, the performance of automated tools versus new Mphy was evaluated in terms of subtyping agreement (concordance), sensitivity, and specificity. When the automated tool reported the same HIV-1 subtype/CRF as the new Mphy did, the result was considered concordant. Regarding Stanford HIVdb, because old versions separated PR and RT subtype results, different results between PR and RT were considered discordant. Sensitivity and specificity were calculated using the formula available at http://statpages.info/ctab2x2.html. Desirable target performance was set at ≥90%.

Statistical analysis. Statistical significance of the differences between subtyping tools was evaluated with McNemar's test, using R version 3.3.1; P values of <0.05 were considered statistically significant.

ACKNOWLEDGMENTS

This work was financially supported by the Italian Ministry of Health (Progetto Ricerca Corrente 2016, line n.2, project n.2, subproject n.2d); the Italian Ministry of Education, University and Research (MIUR) (Bandiera InterOmics Protocollo PB05 1°); and an unrestricted grant from the AVIRALIA foundation.

We have no conflicts of interest directly related to this work.

The Resistance Study Group includes members at the following institutions: University Hospital Tor Vergata, Rome, Italy—Claudia Alteri, Massimo Andreoni, Daniele Armenia,

Maria Concetta Bellocchi, Ada Bertoli, Andrea Biddittu, Massimiliano Bruni, Anna Rita Buonomini, Luca Carioti, Francesca Ceccherini-Silberstein, Carlotta Cerva, Novella Cesta, Domenico Di Carlo, Luca Dori, Luca Foroghi, Elisa Gentilotti, Sara Giannella, Tania Guenci, Gaetano Maffongelli, Vincenzo Malagnino, Alessandra Ricciardi, Marzia Romani, Romina Salpini, Maria Mercedes Santoro, Loredana Sarmati, Rossana Scutari, Valentina Serafini, Pasquale Sordillo, Francesca Stazi, Cristof Stingone, Valentina Svicher, Elisabetta Teti, and Magdalena Viscione; INMI Lazzaro Spallanzani, IRCCS, Rome, Italy—Isabella Abbate, Rosa Acinapura, Lucia Alba, Andrea Antinori, Adriana Ammassari, Franco Baldini, Rita Bellagamba, Evangelo Boumis, Giulia Berno, Maria Rosaria Capobianchi, Stefania Carta, Stefania Cicalini, Fabio Continenza, Gabriella De Carli, Roberta D'Arrigo, Gianpiero D'Offizi, Lavinia Fabeni, Valentina Fedele, Vincenzo Galati, Alberto Giannetti, Enrico Girardi, Caterina Gori, Susanna Grisetti, Raffaella Libertone, Giuseppina Liuzzi, Patrizia Lorenzini, Rita Maddaluno, Andrea Mariano, Assunta Navarra, Emanuele Nicastri, Giuseppina Nurra, Nicoletta Orchi, Antonio Palummieri, Carlo Federico Perno, Carmela Pinnetti, Silvia Pittalis, Daniele Pizzi, Vincenzo Puro, Alessandro Sampaolesi, Maria Rosaria Sciarrone, Paola Scognamiglio, Catia Sias, Ubaldo Visco-Comandini, and Mauro Zaccarelli; San Gallicano Institute, IRCCS, Rome, Italy—Manuela Colafigli, Antonio Cristaudo, Massimo Giuliani, Alessandra Latini, and Anna Pacifici; S. Andrea Hospital, Rome, Italy—Alfredo Pennica; S. Giovanni Addolorata Hospital, Rome, Italy—Fiorella Di Sora, Filippo Iebba, and Francesco Montella; Polo Pontino-Sapienza University, Latina, Italy—Miriam Lichtner, Raffaella Marocco, and Claudio Maria Mastroianni; Children's Hospital Bambino Gesù, Rome, Italy—Stefania Bernardi; CRAIDS Hospital, Frosinone, Italy—Enza Anzalone; CRAIDS Hospital, Rieti, Italy—Maria Elena Bonaventura, Mauro Marchili, and Antonella Pitorri; Teramo Hospital, Teramo, Italy—Luigi Falconi Di Francesco and Dante Di Giammartino; CRAIDS Hospital, Viterbo, Italy—Antonio Caterini and Orlando Armignacco; Avezzano Hospital, Avezzano, Italy—Rinalda Mariani and Maurizio Paoloni; Pescara Hospital, Pescara, Italy—Giustino Parruti, Alessandro Pieri, and Federica Sozio; S. Salvatore Hospital ASL Abruzzo, L'Aquila, Italy—Antonio Cellini, Alessandro Grimaldi, Maurizio Mariani, and Giovanna Picchi; Modena University Hospital, Modena, Italy—Vanni Borghi, Cristina Mussini, and William Gennari; and Chantal Biya International Reference for Research on HIV/AIDS Prevention and Management, Yaoundé, Cameroon—Joseph Fokam, Aubin J. Nanfack, Alexis Ndjolo, Desiré Takou, and Judith N. Torimiro.

REFERENCES

1. Gao F, Bailes E, Robertson DL, Chen Y, Rodenburg CM, Michael SF, Cummins LB, Arthur LO, Peeters M, Shaw GM, Sharp PM, Hahn BH. 1999. Origin of HIV-1 in Pan troglodytes troglodytes. Nature 397:436–441. https://doi.org/10.1038/17130.
2. Sharp PM, Bailes E, Chaudhuri RR, Rodenburg CM, Santiago MO, Hahn BH. 2001. The origins of acquired immune deficiency syndrome viruses: where and when? Philos Trans R Soc Lond B Biol Sci 356:867. https://doi.org/10.1098/rstb.2001.0863.
3. Plantier JC, Leoz M, Dickerson JE, De Oliveira F, Cordonnier F, Lemée V, Damond F, Robertson DL, Simon F. 2009. A new human immunodeficiency virus derived from gorillas. Nat Med 15:871–872. https://doi.org/10.1038/nm.2016.
4. Sharp PM, Hahn BH. 2011. Origins of HIV and the AIDS pandemic. Cold Spring Harb Perspect Med 1:a006841. https://doi.org/10.1101/cshperspect.a006841.
5. Van Heuverswyn F, Li Y, Neel C, Bailes E, Keele BF, Liu W, Loul S, Butel C, Liegeois F, Bienvenue Y, Ngolle EM, Sharp PM, Shaw GM, Delaporte E, Hahn BH, Peeters M. 2006. Human immunodeficiency viruses: SIV infection in wild gorillas. Nature 444:164. https://doi.org/10.1038/444164a.
6. Hemelaar J, Gouws E, Ghys PD, Osmanov S, WHO-UNAIDS Network for HIV Isolation and Characterisation. 2011. Global trends in molecular epidemiology of HIV-1 during 2000–2007. AIDS 25:679–689. https://doi.org/10.1097/QAD.0b013e328342ff93.
7. Vachot L, Ataman-Onal Y, Terrat C, Durand PY, Ponceau B, Biron F, Verrier B. 2004. Short communication: retrospective study to time the introduction of HIV type 1 non-B subtypes in Lyon, France, using env genes obtained from primary infection samples. AIDS Res Hum Retroviruses 20:687–691. https://doi.org/10.1089/0889222041524607.
8. Holguín A, de Mulder M, Yebra G, López M, Soriano V. 2008. Increase of non-B subtypes and recombinants among newly diagnosed HIV-1 native Spaniards and immigrants in Spain. Curr HIV Res 6:327–334. https://doi.org/10.2174/157016208785132455.
9. Fox J, Castro H, Kaye S, McClure M, Weber JN, Fidler S, UK Collaborative Group on HIV Drug Resistance. 2010. Epidemiology of non-B clade forms of HIV-1 in men who have sex with men in the UK. AIDS 24:2397–2401. https://doi.org/10.1097/QAD.0b013e32833c703e.
10. Lai A, Bozzi G, Franzetti M, Binda F, Simonetti FR, Micheli V, Meraviglia P, Corsi P, Bagnarelli P, De Luca A, Ciccozzi M, Zehender G, Zazzi M, Balotta C. 2014. Phylogenetic analysis provides evidence of interactions between Italian heterosexual and South American homosexual males as the main source of national HIV-1 subtype C epidemics. J Med Virol 86:729–736. https://doi.org/10.1002/jmv.23891.
11. Dauwe K, Mortier V, Schauvliege M, Van Den Heuvel A, Fransen K, Servais JY, Bercoff DP, Seguin-Devaux C, Verhofstede C. 2015. Characteristics and spread to the native population of HIV-1 non-B subtypes in two European countries with high migration rate. BMC Infect Dis 15:524. https://doi.org/10.1186/s12879-015-1217-0.
12. Santoro MM, Alteri C, Ronga L, Flandre P, Fabeni L, Mercurio F, D'Arrigo R, Gori C, Palamara G, Bertoli A, Forbici F, Salpini R, Boumis E, Tozzi V, Visco-Comandini U, Zaccarelli M, Van Houtte M, Pattery T, Narciso P, Antinori A, Ceccherini-Silberstein F, Perno CF. 2012. Comparative analysis of drug resistance among B and the most prevalent non-B HIV type 1subtypes (C, F, and CRF02_AG) in Italy. AIDS Res Hum Retroviruses 28:1285–1293. https://doi.org/10.1089/aid.2011.0142.
13. Fabeni L, Alteri C, Orchi N, Gori C, Bertoli A, Forbici F, Montella F, Pennica A, De Carli G, Giuliani M, Continenza F, Pinnetti C, Nicastri E, Ceccherini-Silberstein F, Mastroianni CM, Girardi E, Andreoni M, Antinori A, Santoro MM, Perno CF. 2015. Recent transmission clustering of HIV1 C and

CRF17_BF strains characterized by NNRTI-related mutations among newly diagnosed men in central Italy. PLoS One 10:e0135325. https://doi.org/10.1371/journal.pone.0135325.
14. Lessells RJ, Katzenstein DK, de Oliveira T. 2012. Are subtype differences important in HIV drug resistance? Curr Opin Virol 2:636–643. https://doi.org/10.1016/j.coviro.2012.08.006.
15. Santoro MM, Perno CF. 2013. HIV-1 genetic variability and clinical implications. ISRN Microbiol 2013:481314. https://doi.org/10.1155/2013/481314.
16. Armenia D, Di Carlo D, Gori C, Pérez AB, Borghi V, Bertoli A, Alvarez M, Latini A, Sterrantino G, Lambert S, Callegaro AP, Milesi M, Ghisetti V, Fabeni L, Coppola N, Scognamiglio P, Bellagamba R, Ammassari A, Cicalini S, Cerva C, Maggiolo F, Di Perri G, Cristaudo A, Girardi E, Marcelin AG, Calvez V, Andreoni M, Mussini C, Garcia F, Antinori A, Ceccherini-Silberstein F, Perno CF, Santoro MM. 2016. The co-presence of specific HIV-1 CRF02_AG polymorphisms correlates with a lower response to PI-based first line HAART, abstr 31, p 34–35. XXV Int HIV Drug Resist Workshop, Boston, MA.
17. Kouri V, Khouri R, Alemán Y, Abrahantes Y, Vercauteren J, Pineda-Peña AC, Theys K, Megens S, Moutschen M, Pfeifer N, Van Weyenbergh J, Pérez AB, Pérez J, Pérez L, Van Laethem K, Vandamme AM. 2015. CRF19_cpx is an evolutionary fit HIV-1 variant strongly associated with rapid progression to AIDS in Cuba. EBioMedicine 28:244–254. https://doi.org/10.1016/j.ebiom.2015.01.015.
18. Safrit JT, Fast PE, Gieber L, Kuipers H, Dean HJ, Koff WC. 2016. Status of vaccine research and development of vaccines for HIV-1. Vaccine 34:2921–2925. https://doi.org/10.1016/j.vaccine.2016.02.074.
19. Tatarelli P, Taramasso L, Di Biagio A, Sticchi L, Nigro N, Barresi R, Viscoli C, Bruzzone B. 2016. HIV-1 RNA quantification in CRF02_AG HIV-1 infection: too easy to make mistakes. New Microbiol 39:150–152.
20. Baldrich-Rubio E, Anagonou S, Stirrups K, Lafia E, Candotti D, Lee H, Allain JP. 2001. A complex HIV type 1 A/G/J recombinant virus isolated from a seronegative patient with AIDS from Benin, West Africa. J Gen Virol 82:1095–1106. https://doi.org/10.1099/0022-1317-82-5-1095.
21. Holguín A, Ramírez de Arellano E, Rivas P, Soriano V. 2006. Efficacy of antiretroviral therapy in individuals infected with HIV-1 non-B subtypes. AIDS Rev 8:98–107.
22. Rouet F, Chaix ML, Nerrienet E, Ngo-Giang-Huong N, Plantier JC, Burgard M, Peeters M, Damond F, Ekouevi DK, Msellati P, Ferradini L, Rukobo S, Marechal V, Schvachsa N, Wakrim L, Rafalimanana C, Rakotoambinina B, Viard JP, Seigneurin JM, Rouzioux C. 2007. Impact of HIV-1 genetic diversity on plasma HIV-1 RNA quantification: usefulness of the Agence Nationale de Recherches sur le SIDA second generation long terminal repeat-based real-time reverse transcriptase polymerase chain reaction test. J Acquir Immune Defic Syndr 45:380–388. https://doi.org/10.1097/QAI.0b013e3180640cf5.
23. Lemey P, Salemi M, Vandamme AM (ed). 2009. The phylogenetic handbook. A practical approach to phylogenetic analysis and hypothesis, testing, 2nd ed. Cambridge University Press, New York, NY.
24. Yebra G, de Mulder M, Martín L, Pérez-Cachafeiro S, Rodríguez C, Labarga P, García F, Tural C, Jaén A, Navarro G, Holguín A, Cohort of Spanish AIDS Research Network (CoRIS). 2011. Sensitivity of seven HIV subtyping tools differs among subtypes/recombinants in the Spanish cohort of naïve HIV-infected patients (CoRIS). Antiviral Res 89:19–25. https://doi.org/10.1016/j.antiviral.2010.10.008.
25. Vrbik I, Stephens DA, Roger M, Brenner BG. 2015. The gap procedure: for the identification of phylogenetic clusters in HIV-1 sequence data. BMC Bioinformatics 16:355. https://doi.org/10.1186/s12859-015-0791-x.
26. Pineda-Peña AC, Faria NR, Imbrechts S, Libin P, Abecasis AB, Deforche K, Gómez-López A, Camacho RJ, de Oliveira T, Vandamme AM. 2013. 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 19:337–348. https://doi.org/10.1016/j.meegid.2013.04.032.
27. Gifford R, de Oliveira T, Rambaut A, Myers RE, Gale CV, Dunn D, Shafer R, Vandamme AM, Kellam P, Pillay D. 2006. Assessment of automated genotyping protocols as tools for surveillance of HIV-1 genetic diversity. AIDS 20:1521–1529. https://doi.org/10.1097/01.aids.0000237368.64488.ae.
28. Holguín A, López M, Soriano V. 2008. 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 46:3896–3899. https://doi.org/10.1128/JCM.00515-08.
29. Struck D, Lawyer G, Ternes AM, Schmit JC, Perez Bercoff D. 2014. COMET: adaptive context-based modeling for ultrafast HIV-1 subtype identification. Nucleic Acids Res 42:e144. https://doi.org/10.1093/nar/gku739.
30. Fabeni L, Berno G, Fokam J, Alteri C, Carioti L, Bertoli A, Gori C, Forbici F, Carta S, Fedele V, Cicalini S, Pinnetti C, Bellagamba R, Ammassari A, Borghi V, Giuliani M, De Carli G, Orchi N, Scognamiglio P, Pennica A, Mastroianni CM, Montella F, Cristaudo A, Mussini C, Girardi E, Andreoni M, Antinori A, Ceccherini-Silberstein F, Perno CF, Santoro MM. 2016. Regular update of HIV-1 subtyping is highly needed, mainly for newly emerging strains in western countries, abstr P226. VIII Ital Conf AIDS Antiviral Res, Milan, Italy.
31. Robertson DL, Anderson JP, Bradac JA, Carr JK, Foley B, Funkhouser RK, Gao F, Hahn BH, Kalish ML, Kuiken C, Learn GH, Leitner T, McCutchan F, Osmanov S, Peeters M, Pieniazek D, Salminen M, Sharp PM, Wolinsky S, Korber B. 2000. HIV-1 nomenclature proposal. Science 288:55–56. https://doi.org/10.1126/science.288.5463.55d.
32. Ciccozzi M, Santoro MM, Giovanetti M, Andrissi L, Bertoli A, Ciotti M. 2012. HIV-1 non-B subtypes in Italy: a growing trend. New Microbiol 35:377–386.
33. Yahi N, Fantini J, Tourres C, Tivoli N, Koch N, Tamalet C. 2001. Use of drug resistance sequence data for the systematic detection of non-B human immunodeficiency virus type 1 (HIV-1) subtypes: how to create a sentinel site for monitoring the genetic diversity of HIV-1 at a country scale. J Infect Dis 183:1311–1317. https://doi.org/10.1086/319859.
34. Wang Y, Rawi R, Wilms C, Heider D, Yang R, Hoffmann D. 2013. A small set of succinct signature patterns distinguishes Chinese and non-Chinese HIV-1 genomes. PLoS One 8:e58804. https://doi.org/10.1371/journal.pone.0058804.
35. Pasquier C, Millot N, Njouom R, Sandres K, Cazabat M, Puel J, Izopet J. 2001. HIV-1 subtyping using phylogenetic analysis of pol gene sequences. J Virol Methods 94:45–54. https://doi.org/10.1016/S0166-0934(01)00272-5.
36. Fokam J, Salpini R, Santoro MM, Cento V, D'Arrigo R, Gori C, Perno CF, Colizzi V, Nanfack A, Gwom LC, Cappelli G, Takou D. 2011. Performance evaluation of an in-house human immunodeficiency virus type-1 protease-reverse transcriptase genotyping assay in Cameroon. Arch Virol 156:1235–1243. https://doi.org/10.1007/s00705-011-0982-3.
37. Santoro MM, Fabeni L, Armenia D, Alteri C, Di Pinto D, Forbici F, Bertoli A, Di Carlo D, Gori C, Carta S, Fedele V, D'Arrigo R, Berno G, Ammassari A, Pinnetti C, Nicastri E, Latini A, Tommasi C, Boumis E, Petrosillo N, D'Offizi G, Andreoni M, Ceccherini-Silberstein F, Antinori A, Perno CF. 2014. Reliability and clinical relevance of the HIV-1 drug-resistance test in patients with low viremia levels. Clin Infect Dis 58:1156–1164. https://doi.org/10.1093/cid/ciu020.
38. Armenia D, Di Carlo D, Maffongelli G, Borghi V, Alteri C, Forbici F, Bertoli A, Gori C, Giuliani M, Nicastri E, Zaccarelli M, Pinnetti C, Cicalini S, D'Offizi G, Ceccherini-Silberstein F, Mussini C, Antinori A, Andreoni M, Perno CF, Santoro MM. 2017. Virological response and resistance profile in HIV-1-infected patients starting darunavir-containing regimens. HIV Med 18:21–32. https://doi.org/10.1111/hiv.12388.
39. Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425.
40. Alteri C, Svicher V, Gori C, D'Arrigo R, Ciccozzi M, Ceccherini-Silberstein F, Selleri M, Bardacci SA, Giuliani M, Elia P, Scognamiglio P, Balzano R, Orchi N, Girardi E, Perno CF, SENDIH Study Group. 2009. Characterization of the patterns of drug-resistance mutations in newly diagnosed HIV-1 infected patients naïve to the antiretroviral drugs. BMC Infect Dis 9:111. https://doi.org/10.1186/1471-2334-9-111.
41. Kimura M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:111–120. https://doi.org/10.1007/BF01731581.
42. Martin DP, Murrell B, Golden M, Khoosal A, Muhire B. 2015. RDP4: detection and analysis of recombination patterns in virus genomes. Virus Evol 1:vev003. https://doi.org/10.1093/ve/vev003.