JOURNAL OF CLINICAL MICROBIOLOGY, Dec. 2008, p. 3896–3899 Vol. 46, No. 12
0095-1137/08/$08.00+0 doi:10.1128/JCM.00515-08

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▽

África Holguín,[1,2]* Marisa López,[1] and Vincent Soriano[1]

Department of Infectious Diseases. Hospital Carlos III,[1] and HIV-1 Molecular Epidemiology Laboratory, Department of Microbiology, Hospital Universitario Ramón y Cajal and CIBER-ESP,[2] Madrid, Spain

Received 17 March 2008/Returned for modification 17 May 2008/Accepted 30 September 2008

Human immunodeficiency virus type 1 (HIV-1) subtyping is often estimated on the basis of *pol* sequences by using online websites instead of phylogenetic analysis (phy). We evaluated the reliability of distinct rapid subtyping tools versus phy with a large panel of HIV-1 non-B subtypes and circulating recombinant forms (CRF). *pol* sequences (277 protease [PR] and 171 reverse transcriptase [RT] sequences) previously assigned by phy to eight distinct HIV-1 non-B subtypes were obtained from 277 HIV-infected patients. Phy was run again to identify CRF. Subtyping was then performed using three rapid tools (the Stanford, NCBI, and REGA online tools). Thirty-three additional clade B sequences were tested as controls. New phylogenetic analyses reclassified two-thirds of *pol* sequences previously assigned to HIV-1 non-B clades as CRF. CRF02_AG variants were correctly assigned by the Stanford and NCBI tools for 92 to 97% and 96 to 99% of PR-RT sequences, respectively, while they were correctly assigned by the REGA tool for only 18 to 32% of PR-RT sequences. The Stanford, NCBI, and REGA tools failed to assign pure non-B clades correctly for 24 to 33%, 35%, and 57 to 64% of PR-RT sequences, respectively. For PR-RT sequences from CRF other than CRF02_AG, discrepancies occurred in 98 to 100%, 18 to 43%, and 80 to 87% of sequences, respectively. The concordance between those tools and phy was almost complete for subtype B assignment. Rapid subtyping tools show relatively low agreement with phy in identifying HIV-1 non-B clades and CRF other than CRF02_AG. The Stanford tool shows the best concordance with phy for the assignment of pure non-B clades, while the NCBI tool performs better at identifying CRF. Before entering routine clinical use, rapid subtyping tools should be optimized and updated periodically. Larger numbers of different non-B subtypes and CRF sequences should be included.

Both the high error rate of reverse transcriptase (RT) and the occurrence of recombination events contribute to the expansion of the genetic heterogeneity of human immunodeficiency virus type 1 (HIV-1) in vivo (11). Phylogenetic analysis (phy) permits the identification within HIV-1 group M of nine subtypes (A, B, C, D, F, G, H, J, and K) (13), at least 43 major circulating recombinant forms (CRF) (http://www.hiv.lanl.gov/content/sequence/HIV/CRFs/CRFs.html), and multiple unique recombinant forms. CRF are defined as intersubtype recombinants for which at least three epidemiologically unlinked variants are monophyletic and share an identical genetic structure. Unique recombinant form variants are widely distributed worldwide, with recombination breakpoints different from those of CRF, and they have been found in fewer than three infected persons. Recombination events result from coinfection and/or superinfection of cells with different HIV variants; they occur especially in populations where multiple subtypes cocirculate, such as in sub-Saharan Africa (18). Among the HIV-1 recombinant viruses, the CRF02_AG variant accounts for more than half of the infections in West and Central Africa (16, 17). This variant infects several million people and is the predominant strain among the population from these African countries living in Southwest Europe but infected in their native countries. An increasing number of non-B subtypes have been reported throughout Europe and in areas of North America with large numbers of immigrants. Local transmission of non-B subtypes and CRF have already been documented (1, 5, 14). In Spain, almost all HIV-1 subtypes have been identified (7, 15), and multiple recombinants, mostly CRF02_AG variants, are currently in circulation (6, 8, 9).

Identification of non-B subtypes and CRF is important for confronting potential problems with the use of antiretroviral drugs and diagnostic tools, such as viral load and resistance tests, which have been designed mainly based on clade B sequences (3, 10, 19). Moreover, pathogenicity can differ across HIV-1 clades, and several studies have shown that subtype D may lead to more-rapid disease progression than other subtypes (2, 22). Thus, HIV-1 subtype characterization is becoming an important aspect of the adequate clinical management of HIV-infected persons (4).

Since phylogenetic analyses are not widely implemented in clinical settings due to their complexity, most clinicians perform HIV-1 subtyping by interpreting *pol* sequences obtained following drug resistance testing by using online websites instead of phy. Here we evaluate the reliability of the three rapid subtyping tools available (the Stanford, NCBI, and REGA

* Corresponding author. Mailing address: HIV-1 Molecular Epidemiology Laboratory, Department of Microbiology, Hospital Ramón y Cajal, Crta. Colmenar Viejo Km. 9.100, 28034, Madrid, Spain. Phone: 34-91-3368832. Fax: 34-91-3368809. E-mail: aholguin.hciii@salud.madrid.org.

▽ Published ahead of print on 8 October 2008.

TABLE 1. Agreement of the three rapid subtyping tools with phylogenetic analysis for the study population

Subtype defined by phy of PR[a]	No. of PR sequences	Concordance[b] with phy (%) for PR			No. of RT sequences	Concordance[b] with phy (%) for RT		
		Stanford	NCBI	REGA[c]		Stanford	NCBI	REGA[c]
A	27	63	44.4	52	18	22.2	33.3	44.4
B[d]	2	100	100	50	7	100	100	71.4
C	16	62.5	56.3	12.5	9	77.8	77.8	55.6
D	5	100	80	20	2	100	100	0
F	4	75	100	50	3	100	100	66.7
G	32	87.5	72	37.5	20	80	65	25
H	4	100	100	25	1	100	100	100
J	1	0	100	0	0			
CRF01_AE	2	50	50	0	0			
CRF02_AG	141	97.2	99.3	18.4	81	92.5	96.3	32
CRF03_AB	2	0	100	0	0			
CRF06_cpx	11	*0*	*81.8*	18.2	5	0	60	40
CRF10_CD	2	0	0	0	3	0	0	0
CRF11_cpx	5	*0*	*100*	40	4	*0*	*75*	50
CRF12_BF	9	*0*	*66.7*	22.2	7	*0*	*57.1*	28.6
CRF13_cpx	1	0	100	0	1	0	0	0
CRF14_BG	13	*0*	*100*	0	10	*0*	*100*	0
Pure subtypes	91	75.8	64.8	36.2	60	66.7	65	43.3
All CRFs	186	*74.2*	*95.1*	17.5	111	*67.6*	*85.6*	28.8
CRFs without CRF02_AG	45	*2.2*	*82.2*	13.3	30	*0*	*56.7*	20
Total	277	*74.7*	*85.2*	*23.5*	171	*67.2*	*78.4*	33.9
Clade B control[e]	33	100	100	90.9	33	100	100	90.9

[a] Subtypes in the *pol* coding region are as follows: CRF01_AE, subtype A; CRF02_AG, subtypes A and G; CRF03_AB, subtypes A and B; CRF06_cpx, subtypes G, K, A, and J; CRF10_CD, clades C and D; CRF11_cpx, subtypes J, G, and A; CRF12_BF, subtypes F and B; CRF13_cpx, subtypes G and J; CRF14_BG, subtype G.
[b] Italicized values indicate significant differences ($P < 0.05$) between Stanford and NCBI results.
[c] Including all the sequences not assigned to a subtype by the REGA tool (194 PR and 95 RT sequences).
[d] Specimens carrying non-B sequences at the *pol* coding region.
[e] Specimens carrying clade B sequences at both PR and RT.

tools) versus phy with a large panel of HIV-1 non-B subtypes and CRF from HIV-1-infected patients living in Spain.

MATERIALS AND METHODS

pol sequences (277 protease [PR] and 171 RT sequences) previously assigned to eight distinct HIV-1 non-B subtypes by phy (5, 7, 15) were obtained from 277 HIV-1-infected patients. Most of them were on regular follow-up at our institution, an HIV/AIDS and Tropical Medicine Reference Center. All patients were under epidemiological suspicion of carrying non-B clades, because they were either (i) foreigners coming from areas where non-B clades are endemic, (ii) natives who had traveled to countries where HIV-1 non-B subtypes are endemic and who admitted to having engaged in high-risk behaviors there, or (iii) natives who had sexual contacts with people coming from those areas.

The study population had a mean age of 40 years, and almost half (51.5%) were male. Among the patients in known risk groups, 89.5% and 1.9% had acquired HIV-1 infections through heterosexual and homosexual relationships, respectively, while 5.8% admitted intravenous drug use, 1.9% were infected by a vertical route, and 0.9% were infected by blood transfusion. This study was part of a project approved by a review board and the ethical committee from our institution. It was designed to protect the rights of all subjects involved under the appropriate local regulations. To maintain subject confidentiality, a unique identification number was assigned to each specimen.

Sequences were obtained from plasma samples by direct sequencing of purified nested-PCR products of HIV-1 PR and/or RT coding regions using an automatic sequencer (ABI Prism; Applied Biosystems, Foster City, CA) as previously described (5–7, 15). To identify which non-B subtypes were in fact CRF, new phylogenetic analyses were run again on all PR and RT sequences for recombinant identification, including reference *pol* sequences from the 34 CRF variants available in the GenBank database when the study was performed in addition to the 9 pure subtypes. The tree topology was obtained using the neighbor-joining method. DNA sequences were aligned using the ClustalX program. The pairwise distance matrix was estimated using the Kimura two-parameter model within the DNADIST program, as implemented in the PHYLIP software package. Bootstrap resampling (1,000 data sets) of the multiple alignment was performed to test the statistical robustness of trees. Sequences from 33 randomly selected native Spanish individuals infected with HIV-1 subtype B strains, defined by performing phy of both PR and RT coding regions, were used as controls and tested in parallel using the same rapid subtyping tools.

HIV-1 *pol* gene subtyping was done for all sequences by using three rapid tools: the HIV-1 Drug Resistance Database (Stanford University, Palo Alto, CA), available at http://hivdb.stanford.edu/; the NCBI genotyping tool (National Center for Biotechnology Information, National Institutes of Health, Bethesda, MD), available at http://www.ncbi.nlm.nih.gov/projects/genotyping/formpage.cgi; and the REGA HIV Subtyping Tool (REGA Institute at the Catholic University of Leuven, Leuven, Belgium), available at http://www.bioafrica.net/subtypetool/html/. We considered discrepancies in HIV-1 subtype and CRF identification between rapid subtyping tools and phy when the former assigned an HIV-1 variant different from that provided by phy and when sequences were not assigned to a type, as frequently happened when the REGA tool was used. All statistical analyses were performed using SPSS software, version 13.0 (SPSS Inc., North Chicago, IL).

RESULTS

Phylogenetic characterization. The new phylogenetic analyses reassigned as CRF two-thirds (67% of PR and 65% of RT sequences) of *pol* sequences previously assigned as pure HIV-1 non-B clades by analysis of PR and/or RT (5, 7, 15). These variants belonged to nine different CRF, and the rest belonged to eight different HIV-1 subtypes (Table 1). As expected, among the previously reported clade G PR and/or RT sequences, the new phy allowed us to identify a large number of CRF02_AG strains.

HIV-1 subtypes defined by phy were compared with those

provided by the three rapid subtyping tools (Stanford, NCBI, and REGA). In more detail, subtyping by the three rapid tools was performed for 277 PR and 171 RT sequences derived from 277 HIV-1 infected subjects previously determined by phy to be infected by non-B subtypes. At the PR the subtypes were 27A, 2B, 16C, 5D, 4F, 32G, 4H, 1J, 2CRF01_AE, 141CRF02_AG, 2CRF03_AB,11CRF06_cpx, 2CRF10_CD, 5CRF11_cpx, 9CRF12_BF, 1CRF13_cpx, and 13CRF14_BG. At the RT they were 18A, 7B, 9C, 2D, 3F, 20G, 1H, 81CRF02_AG, 5CRF06_cpx, 3CRF10_CD, 4CRF11_cpx, 7CRF12_BF, 1CRF13_cpx, and 10CRF14_BG (Table 1). Clade B sequences were derived from recombinant viruses showing distinct clade assignment at the PR and RT. In addition, 33 clade B sequences at both genes were tested as controls.

Assignment of subtype B sequences using rapid subtyping tools. All tested subtype B variants, which are the most prevalent in the United States and Western Europe, were correctly assigned by the Stanford and NCBI tools. Both methods detected all PR and RT clade B sequences derived from CRF, as well as all 33 PR and RT clade B sequences used in pure clade B controls (Table 1). In contrast, the REGA tool failed to identify one PR and two RT clade B sequences from CRF carrying sequences from another subtype at *pol* (Table 1). Moreover, 3 of the 33 pure clade B PR/RT control sequences were not assigned by REGA.

Assignment of CRF02_AG sequences using rapid subtyping tools. Among the CRF tested in this study, 76% (141/186) of PR and 73% (81/111) of RT sequences belonged to CRF02_AG variants (Table 1). CRF02_AG sequences were correctly identified by the Stanford and NCBI tools for 92 to 97% and 96 to 99% of the PR-RT sequences, respectively, but only for 18 to 32% of the sequences by the REGA tool (Table 1). Overall, the REGA tool did not provide subtype assignment for 79% of PR and 60.5% of RT sequences derived from CRF02_AG strains, as was shown by phy.

Assignment of pure non-B HIV-1 subtypes. The overall accuracy of the Stanford, NCBI, and REGA tools in the identification of pure non-B clades of HIV-1 and CRF other than CRF02_AG carrying sequences from distinct subtypes at *pol* was relatively low compared to that of phy. As shown in Table 1, rapid subtyping tools showed discrepant results when interpreting sequences from non-B subtypes and CRF defined by phy. The Stanford, NCBI, and REGA tools failed to correctly assign pure non-B clades for 24 to 33%, 35%, and 57 to 64% of PR-RT sequences, respectively. Thus, the Stanford database showed the best concordance with phy for the assignment of pure non-B clades.

Of note, the PR sequence belonging to clade J was detected only by the NCBI tool. The REGA tool showed a similar concordance with other rapid tools for the assignment of clade A at PR, identifying a higher number of subtype A RT sequences (8/18) than the other two methods (Table 1). However, the REGA tool did not assign any subtype for 194 (70%) PR sequences and 95 (55.5%) RT sequences, which were known to belong to CRF in nearly 70% of the PR and RT sequences not assigned a subtype (data not shown).

Assignment of CRF other than CRF02_AG by using rapid subtyping tools. Overall, rapid subtyping tools showed limitations in the assignment of CRF other than CRF02_AG variants for *pol* sequences (45 PR and 30 RT specimens) when phy was taken as the reference. For instance, the Stanford, NCBI, and REGA tools failed to assign CRF correctly in our study for 98 to 100%, 18 to 43%, and 80 to 87% of PR-RT sequences, respectively (Table 1). The NCBI tool demonstrated the best concordance with phy for these CRF variants. In more detail, the concordance between NCBI and phy was complete for CRF03_AB, CRF11_cpx, CRF13_cpx, and CRF14_BG at the PR and for CRF14_BG at the RT, as well as for 75% and 57% of CRF11_cpx and CRF12_BF sequences, respectively, at the RT. Finally, the NCBI tool provided the best concordance with phy for CRF06_cpx, with correct assignment for nearly 80% of PR sequences, while REGA was correct in only 18% and 40% of PR and RT sequences, respectively (Table 1).

DISCUSSION

HIV-1 subtype characterization is becoming an important aspect of the clinical management of infected persons. Thus, updated information about circulating HIV-1 variants in distinct countries is helpful for epidemiological purposes as well as for confronting potential problems with the use of diagnostic assays (3, 18), viral load measurements (19), drug resistance testing, and the prescription of antiretroviral drugs (10). Since phylogenetic analyses are not widely used in clinical settings, most clinicians base subtype characterization on the interpretation of *pol* sequences obtained following drug resistance testing by using online tools. Here we have performed the first comparison of the reliabilities of different rapid HIV-1 subtyping tools (the Stanford, NCBI, and REGA tools), taking as a reference the results obtained using phy by testing a relatively large group of clinical specimens obtained from patients infected with HIV-1 non-B subtypes and CRF.

The number of HIV-1 strains that have been fully sequenced has increased within the past few years, revealing a high prevalence of recombinant variants among strains previously considered to represent pure non-B subtypes, based on information derived from single viral coding regions. For instance, viruses carrying subtype G sequences at *pol* could be pure G subtypes or recombinants, and interestingly, 18 of the CRF described (http://www.hiv.lanl.gov/content/sequence/HIV/CRFs/CRFs.html) harbor clade G sequences in partial or complete *pol* coding regions. Therefore, it is not surprising that CRF may constitute at least 10 to 20% of new HIV-1 infections worldwide and that recombination involving different non-B subtypes is rising where diverse variants cocirculate, particularly in areas where distinct subtypes are highly prevalent (8, 12, 21, 24).

The concordance between rapid subtyping tools and phy for clade B viruses was excellent in this study. These results reinforce the good performance of rapid subtyping tools for this subset of patients. Our results also revealed that the Stanford database, the tool most widely used by clinicians for routine clinical characterization of HIV-1 subtypes, showed the best concordance with phy for the recognition of pure non-clade-B subtypes. However, the Stanford database failed to identify nearly 25% of PR/RT sequences belonging to non-B clades, as well as 98% of PR and all RT sequences belonging to CRF variants other than CRF02_AG. The NCBI tool showed the best concordance with phy for CRF identification, although it

failed to assign 35% of pure non-B subtypes correctly in our study. The REGA tool failed to assign either pure non-B subtypes or CRF correctly for an unexpectedly large number of samples. However, it was the only tool that identified CRF06_cpx variants, and it showed the best performance for identifying clade A sequences at the RT gene. Taken together, our results point out that misclassification of HIV-1 non-B subtypes and CRF strains other than CRF02_AG may occur frequently when rapid subtyping tools are used, a limitation that may have epidemiological and clinical consequences. However, the concordance between the Stanford tool, the NCBI tool, and phy was relatively good, with complete agreement for nearly two-thirds of *pol* sequences.

It is reassuring that CRF02_AG strains, which are spreading rapidly in West and Central Africa (16, 17), as well as in European countries with strong links to, and immigrants from, these regions, are well represented in subtyping databases. Correspondingly, rapid subtyping tools performed quite well with these sequences. However, CRF variants other than CRF02_AG were less well represented in these databases and often were misclassified by rapid subtyping tools. Since these strains are similarly spreading worldwide, it may be worth highlighting the importance of reporting these non-B sequences and stressing the need for improved accuracy of rapid subtyping tools for these variants. Other potential bioinformatics tools have been developed, or are in development, to deal with the unreliability of subtyping tools for correctly identifying HIV-1 subtypes, subsubtypes, CRF, and other HIV variants using viral sequences (20, 23).

In summary, this study revealed that HIV-1 rapid subtyping tools can be useful for identifying clade B sequences. However, phy of *pol* sequences continues to be the only reliable method for correctly assigning HIV-1 non-B subtypes and CRF, which are growing in number and complexity. Clearly, the number of sequences from non-B and CRF strains in reference databases must increase, and rapid tools should improve their prediction of subtyping for the distinct non-B variants, including CRF, before being used in routine clinical settings.

ACKNOWLEDGMENTS

This work was supported in part by grants from the Red de Investigación en SIDA (RIS, project ISCIII-RETIC RD06/006) and the Fondo de Investigaciones Sanitarias (FIS PI06/0925). A.H. is supported by FIS and the Agencia Laín Entralgo.

The authors do not have any conflict of interest.

REFERENCES

1. **Aggarwal, I., M. Smith, I. D. Tatt, S. Murad, N. Osner, A. M. Geretti, and P. J. Easterbrook.** 2006. Evidence for onward transmission of HIV-1 non-B subtype strains in the United Kingdom. J. Acquir. Immune Defic. Syndr. **41:**201–209.
2. **Baeten, J. M., B. Chohan, L. Lavreys, V. Chohan, R. S. McClelland, L. Certain, K. Mandaliya, W. Jaoko, and J. Overbaugh.** 2007. HIV-1 subtype D infection is associated with faster disease progression than subtype A in spite of similar plasma HIV-1 loads. J. Infect. Dis. **195:**1177–1180.
3. **Baldrich-Rubio, E., S. Anagonou, K. Stirrups, E. Lafia, D. Candotti, H. Lee, and J. P. Allain.** 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.
4. **Geretti, A. M.** 2006. HIV-1 subtypes: epidemiology and significance for HIV management. Curr. Opin. Infect. Dis. **19:**1–7.
5. **Holguín, A., A. Álvarez, and V. Soriano.** 2002. High prevalence of subtype G and spectrum of natural polymorphisms at the protease gene among HIV-infected immigrants attended in Madrid. AIDS **16:**1163–1170.
6. **Holguín, A., A. Álvarez, and V. Soriano.** 2005. Heterogeneous nature of HIV-1 recombinants spreading in Spain. J. Med. Virol. **75:**374–380.
7. **Holguín, A., A. Álvarez, and V. Soriano.** 2005. Differences in the length of Gag proteins among different HIV type 1 subtypes. AIDS Res. Hum. Retrovir. **21:**886–893.
8. **Holguín, A., E. Lospitao, M. López, E. Ramírez de Arellano, M. J. Pena, J. del Romero, C. Martín, and V. Soriano.** 2008. Genetic characterization of complex inter-recombinant HIV-1 strains circulating in Spain and reliability of distinct rapid subtyping tools. J. Med. Virol. **80:**383–391.
9. **Holguín, A., M. J. Pena, F. Troncoso, and V. Soriano.** 2007. Introduction of non-B subtypes among Spaniards newly diagnosed with HIV-1 in the Canary Islands. AIDS Res. Hum. Retrovir. **23:**498–502.
10. **Holguín, A., E. Ramírez de Arellano, P. Rivas, and V. Soriano.** 2006. Efficacy of antiretroviral therapy in individuals infected with HIV-1 non-B subtypes. AIDS Rev. **8:**98–107.
11. **Jetzt, A. E., H. Yu, G. J. Klarmann, Y. Ron, B. D. Preston, and J. P. Dougherty.** 2000. High rate of recombination throughout the HIV-1 genome. J. Virol. **74:**1234–1240.
12. **Konings, F. A., G. R. Haman, Y. Xue, M. M. Urbanski, K. Hertzmark, A. Nanfack, J. M. Achkar, S. T. Burda, and P. N. Nyambi.** 2006. Genetic analysis of HIV-1 strains in rural eastern Cameroon indicates the evolution of second-generation recombinants to circulating recombinant forms. J. Acquir. Immune Defic. Syndr. **42:**331–341.
13. **Leitner, T., B. Korber, M. Daniels, C. Calef, and B. Foley.** 2005. HIV-1 subtype and circulating recombinant form (CRF) reference sequences, 2005, p. 41–48. *In* T. Leitner, B. Foley, B. Hahn, P. Marx, F. McCutchan, J. W. Mellors, S. Wolinsky, and B. Korber (ed.), HIV sequence compendium 2005. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM. LA-UR 06-0680. http://www.hiv.lanl.gov/content/sequence/HIV/COMPENDIUM/2005/partI/leitner.pdf.
14. **Lin, H. H., B. K. Gaschen, M. Collie, M. El-Fishaway, Z. Chen, B. T. Korber, S. T. Beatrice, and L. Zhang.** 2006. Genetic characterization of diverse HIV-1 strains in an immigrant population living in New York City. J. Acquir. Immune Defic. Syndr. **41:**399–404.
15. **Lospitao, E., A. Álvarez, V. Soriano, and A. Holguín.** 2005. HIV-1 subtypes in Spain: a retrospective analysis from 1995 to 2003. HIV Med. **6:**313–320.
16. **McCutchan, F.** 2000. Understanding the genetic diversity of HIV-1. AIDS **14**(Suppl. 3)**:**S31–S44.
17. **Montavon, C., C. Toure-Kane, F. Liegeois, E. Mpoudi, A. Bourgeois, L. Vergne, J. L. Perret, A. Boumah, E. Saman, S. Mboup, E. Delaporte, and M. Peeters.** 2000. Most env and gag subtype A HIV-1 viruses circulating in West and Central Africa are similar to the prototype AG recombinant virus IBNG. J. Acquir. Immun Defic. Syndr. **23:**363–374.
18. **Peeters, M., C. Toure-Kane, and J. N. Nkengasong.** 2003. Genetic diversity of HIV in Africa: impact on diagnosis, treatment, vaccine development and trials. AIDS **17:**2547–2560.
19. **Rouet, F., M.-L. Chaix, E. Nerrienet, N. Ngo-Giang-Huong, J. C. Plantier, M. Burgard, M. Peeters, F. Damond, D. K. Ekouevi, P. Msellati, L. Ferradini, S. Rukobo, V. Maréchal, N. Schvachsa, L. Wakrim, C. Rafalimanana, B. Rakotoambinina, J. P. Viard, J. M. Seigneurin, and C. Rouzioux.** 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.
20. **Schultz, A. K., M. Zhang, T. Leitner, C. Kuiken, B. Korber, B. Morgenstern, and M. Stanke.** 2006. A jumping profile Hidden Markov Model and applications to recombination sites in HIV and HCV genomes. BMC Bioinformatics **7:**265.
21. **Takebe, Y., K. Motomura, M. Tatsumi, H. H. Lwin, M. Zaw, and S. Kusagawa.** 2003. High prevalence of diverse forms of HIV-1 intersubtype recombinants in Central Myanmar: geographical hot spot of extensive recombination. AIDS **17:**2077–2087.
22. **Vasan, A., B. Renjifo, E. Hertzmark, B. Chaplin, G. Msamanga, M. Essex, W. Fawzi, and D. Hunter.** 2006. Different rates of disease progression of HIV type 1 infection in Tanzania based on infecting subtype. Clin. Infect. Dis. **42:**843–852.
23. **Wu, X., Z. Cai, X. F. Wan, T. Hoang, R. Goebel, and G. Lin.** 2007. Nucleotide composition string selection in HIV-1 subtyping using whole genomes. Bioinformatics **23:**1744–1752.
24. **Yang, R., S. Kusagawa, C. Zhang, X. Xia, K. Ben, and Y. Takebe.** 2003. Identification and characterization of a new class of HIV-1 recombinants comprised of two circulating recombinant forms, CRF07_BC and CRF08_BC, in China. J. Virol. **77:**685–695.