academicJournals

Vol. 15(50), pp. 2744-2750, 14 December 2016
DOI: 10.5897/AJB2016.15554
Article Number: 5629E5E62025
ISSN 1684-5315

http://www.academicjournals.org/AJB

African Journal of Biotechnology

Full Length Research Paper

Performance of rapid subtyping tools used for the classification of HIV type 1 recombinants isolated from selected countries in west and central Africa

Bimela Jude Saber[1], Judith Ndongo Torimiro[2,3]*, Atogho Tiedeu Barbara[1], Tchapda Arlette Laure[1], Joseph Fokam[3], Desire Takou[3] and Wilfred Fon Mbacham[1,2,4]

[1]Faculty of Science, University of Yaoundé I, Yaoundé, Cameroon.
[2]Faculty of Medicine and Biomedical Sciences, University of Yaoundé I, Yaoundé, Cameroon.
[3]Centre International de Référence "Chantal Biya", Yaoundé, Cameroon.
[4]Biotechnology Centre, University of Yaounde I, Yaoundé, Cameroon.

Received 1 July, 2016; Accepted 7 September, 2016

HIV-1 genetic diversity in sub-Saharan Africa is broad and the AIDS epidemic is driven predominantly by recombinants in Central and West Africa. The classification of HIV-1 strains is therefore necessary to understand diagnostic efficiency, individual treatment responses as well as options for designing vaccines and antiretroviral (ARV) treatment guidelines. More so, accurate subtyping of a partial or full genome would represent the population dynamics of HIV and provide evidence for designing surveillance strategies within a geographic region. Evaluating the performance of rapid subtyping tools with options that incorporate phylogeny could be fast, more user-friendly and of high performance. A total of 570 HIV-1 partial sequences from Cameroon, Angola, Democratic Republic of Congo, Gabon and Senegal were obtained from the Los Alamos National Library (LANL) HIV Sequence Database. Phylogeny was performed using MEGA v6 and the results were used to evaluate the performance of eleven different rapid HIV-1 subtyping tools: REGA v2, REGA v3, NCBI, Stanford HIVDB, SUDI, Geno2Pheno, Euresist, STAR, jpHMM, COMET and SCUEAL. The performance of these subtyping tools differed among HIV-1 clades and across different viral genes. NCBI and SUDI showed the highest performance in subtyping. The discordance observed between the rapid subtyping tools and phylogeny implies that phylogenetic analysis is still the more suitable method for HIV-1 classification. However, the need to update the reference datasets of the subtyping tools, and validate algorithms for rapid subtyping and quality control is imperative as this information is relevant for clinical use and policy-making to the AIDS response.

Key words: HIV, phylogeny, performance, subtyping tools, algorithm.

INTRODUCTION

Globally, an estimated 36.9 million people were living with human immunodeficiency virus (HIV) at the end of 2014. About 70% of people living with HIV are in sub-Saharan Africa; a region with only 12% of the world's

population (UNAIDS Global Report, 2014). The genetic diversity of HIV continues to increase due to its high replication rate (Ho et al., 1995; Michael, 1999), host selective pressure (Temin, 1993), recombination events in dually infected patients (Op de Coul et al., 1997) and the inefficient proofreading capacity of reverse transcriptase (MacNeil et al., 2007). Genetic diversity has numerous implications at various levels in the health care system, notably, diagnosis, rates of disease induction and progression, and response to treatment (Henguell et al., 2008; Drylewicz et al., 2008). The impact of HIV genetic diversity was reported in the initial failure to diagnose HIV-1 Group O infections in individuals who showed clinical signs and symptoms of AIDS. In terms of antiretroviral treatment, HIV-1 Group O and HIV-2 viruses are naturally resistant to the non-nucleoside reverse transcriptase inhibitors (NNRTIs) (Menéndez-Arias and Tözsér, 2008; Carr et al., 2001). Similarly, the efficacy of protease inhibitors against HIV-2 is poor (Menéndez-Arias and Tözsér, 2008; Peeters, 2001). One of the major obstacles in the development of an effective vaccine is the broad diversity of HIV-1. Subtyping is in particular important for subsequent prediction of HIV drug resistance. Genetic analysis and classification of HIV-1 are therefore considered important in understanding the dynamics of the virus, the relevance for developing diagnostic assays and vaccines, planning treatment strategies, and strategies for surveillance and monitoring of new viral variants. Riemenschneider et al. (2016) have demonstrated that prediction of co-receptor is not possible for some subtypes and recombinants with existing tools at the moment. Thus, subtyping is very important to get reliable predictions and effective treatment of the patient (Riemenschneider et al., 2016). Africa has the highest HIV-1 genetic diversity and remains a suitable geographical region for studies on antiretrovirals and vaccine development. HIV-1 Groups M, N, O and P (Plantier et al., 2009; Torimiro et al., 2009; Ceccarelli et al., 2012), several circulating recombinants forms (CRF) and unique recombinant forms (URF) (Zhang et al., 2010; Fokam et al., 2011) have been identified in Cameroon and in other countries in sub-Saharan Africa. Its identification and classification are therefore public health challenges.

The global acquired immune deficiency syndrome (AIDS) pandemic in different geographic regions is driven by different genetic variants, thus, the need for describing country HIV epidemiology regularly. In this study we compared the performance of eleven rapid subtyping tools (designed for specific purposes) to give instant classification of HIV-1 partial genome of which some are used for the interpretation of drug resistance profile. The Stanford HIV Drug Resistance Database (Stanford HIV DB) for example provides both classification and prediction of antiretroviral resistance while the National Center for Biotechnology Information (NCBI) provides predictions of the genotypic variants only. However, these subtyping tools all utilize algorithms to classify HIV-1. The AIDS pandemic is evolving into a complex which makes classification especially of the partial genome challenging. A more accurate and updated molecular epidemiology of HIV in a given geographic region or country could guide the development of the national diagnostic and treatment guidelines.

METHODS

Study population

HIV-1 gag, pol and env sequences obtained from five countries in sub-Saharan Africa notably, Angola, Cameroon, Democratic Republic of Congo (DRC), Gabon and Senegal were downloaded from the HIV Sequence Database of the Los Alamos National Library (HIV DB LANL) and used for *in silico* analyses and classification.

HIV-1 sequence download, trimming and alignment of partial HIV-1 gag, pol and env sequences

HIV-1 sequences were downloaded in FASTA format from the HIV DB LANL, imported into MEGA v6, trimmed with reference to the HXB2 (Genbank Acc No: K03455) genome landmarks. A total of 661 full length or partial HIV-1 genome sequences were downloaded from Cameroon. After trimming, a total of 39 *gag*, 66 *pol* and 65 *env* partial sequences were obtained. The *gag* sequences were within 790 to 2292 nucleotides, while the *pol* and *env* regions were within 2085 to 5096 and 6225 to 8795, respectively, on the HXB2 genome landmark. Similarly, 101 HIV-1 recombinant gag sequences from Angola (12), DRC (25), Gabon (25) and Senegal (39) in addition to 216 HIV-1 recombinant pol sequences from Angola (42), DRC (27), Gabon (50) and Senegal (97) and 83 HIV-1 recombinant env sequences from Angola (0), DRC (27), Gabon (22) and Senegal (34) were also obtained. Reference sequences CM240 for CRF01_AE, IbNG for CRF02_AG, BFP90 for CRF06_cpx, and GR17 for CRF11_cpx, subtypes A and B (B.FR.1983.HXB2-LAI-IIIB-BRU.K03455) were downloaded from the LANL HIV Sequence Database. Other reference sequences were selected by downloading many subtypes and recombinants found in the LANL HIV DB from different parts of the world that would be representative of all the recombinant forms that drive the HIV epidemic in Central and West Africa. A first phylogenetic tree was then constructed using these sequences and one or two sequences were selected from each cluster and used in the phylogenetic tree as reference sequences in identifying other sequences of interest.

Multiple alignments were performed using ClustalW in the

*Corresponding author. E-mail: jn.torimiro@gmail.com.

Table 1. HIV-1 sequences used for phylogenetic analysis for the partial gag, pol and env region sequences from Cameroon, Angola, DRC, Gabon and Senegal.

Genomic region	Country					Total
	Angola	Cameroon	DRC	Gabon	Senegal	
Gag	12	39	25	25	39	140
Pol	42	66	27	50	97	282
Env	00	65	27	22	34	148
Total	54	170	79	97	170	570

MEGA v6 Software (Tamura et al., 2007). The alignment was manually edited and improved where possible with the use of the BioEdit v7.1 sequence editing programme (Hall, 2001).

Construction of phylogenetic trees for HIV-1 gag, pol and env sequences

Neighbor-Joining phylogenetic trees were constructed in MEGA v6 (Saitou and Nei, 1987) based on the Kimura two parameter model (K2P) (Kimura, 1980). The reliability of the branching orders was assessed by bootstrap analysis of 1,000 replicates (Thompson et al., 1997; Felsenstein, 1985) and bootstrap values >70% were considered significant. All positions containing gaps and missing data were eliminated from the dataset with the help of the complete deletion option within MEGA v6 tool. HIV-1 subtypes were identified by phylogenetic analysis of partial *gag* sequences from Cameroon (39), Angola (12), Democratic Republic of Congo (25), Gabon (25) and Senegal (39). *Pol* sequences from Cameroon (66) and 216 recombinant pol sequences from Angola (42), Democratic Republic of Congo (27), Gabon (50) and Senegal (97) together with 148 HIV-1 partial env sequences from Cameroon (65), DRC (27), Gabon (22) and Senegal (34) were also analyzed using recombinant reference sequences. There were not enough recombinant env sequences from Angola available on the LANL HIV sequence database for a robust analysis. Circulating recombinant forms other than CRF01_AE and CRF02_AG identified by phylogeny with low bootstrap values (<70%) were further studied to identify recombination events using the Recombination Identification Programme (RIP).

HIV-1 subtyping using rapid subtyping tools

HIV-1 subtyping using eleven online rapid subtyping tools REGA v2, REGA v3, NCBI, Stanford HIVDB 6.0.10, SUDI, Gene2pheno 3.3, EuResist 2009, STAR, jpHMM, COMET version 2 and SCUEAL (Yebra et al., 2011; Pineda-Peña et al., 2013) were tested using partial *gag*, *pol* and *env* sequences.

Statistical analysis

The phylogenetic trees of the *gag*, *pol* and *env* regions were constructed and analyzed within MEGA v6 using bootstrap values. Recombinants other than CRF01_AE and CRF02_AG and clusters with bootstrap values less than 70% were verified using the RIP tool. Sensitivity was calculated using the formula (Banoo et al., 2010): TP/(TP+FN), where TP = true positives and FN = false negatives.

RESULTS

Phylogenetic analysis of the partial HIV-1 gag, pol and env sequences

Phylogenetic analysis of the HIV-1 gag sequences from Cameroon, Angola, Democratic Republic of Congo, Gabon and Senegal

Phylogenetic analysis of 140 HIV-1 gag sequences from Cameroon (39) and HIV-1 gag recombinants from Angola (12), Democratic Republic of Congo (25), Gabon (25) and Senegal (39) was carried out using recombinant reference sequences (Table 1).

Each colour represents a cluster. The scale at the bottom indicates 0.05 substitutions/site. Numbers above branches indicate the statistical robustness and reliability of the branching order, estimated with a total bootstrap support of 1000 replicates for each dataset and with a cut-off at 700 on a Neighbor-Joining tree. Only bootstrap values ≥70 are shown in red dots on the tree.

Phylogenetic analysis of the HIV-1 pol sequences from Cameroon, Angola, Democratic Republic of Congo, Gabon and Senegal

Phylogenetic analysis of 282 HIV-1 pol sequences from Cameroon (66), Angola (42), Democratic Republic of Congo (27), Gabon (50) and Senegal (97) was performed.

Phylogenetic analysis of the HIV-1 env sequences from Cameroon, Angola, Democratic Republic of Congo, Gabon and Senegal

Phylogenetic analysis of 148 env sequences of HIV-1 recombinants from Cameroon (65), Democratic Republic of Congo (27), Gabon (22) and Senegal (34) was carried out using recombinant reference sequences. There were not enough env sequences from Angola available on the LANL HIV sequence database for a robust analysis.

Table 2. Performance of 11 rapid subtyping tools for the HIV-1 gag, pol and env sequences from Cameroon, Angola, DRC, Gabon and Senegal.

Country	Sensitivity (Percentage)										
	Rega 2.0	**Rega 3.0**	**NCBI 2009**	**Stanford 6.0.10**	**SUDI**	**Geno2pheno 3.3**	**EuResist 2009**	**STAR**	**jpHMM**	**COMET 2**	**SCUEAL**
Gag											
Cameroon (39)	20 (51.3)	29 (74.4)	34 (87.2)	NA	37 (94.9)	23 (58.9)	28 (71.8)	28 (71.8)	NA	32 (82.1)	NA
Angola (12)	6 (50)	8 (66.7)	12 (100)	NA	12 (100)	8 (66.7)	8 (66.7)	5 (41.7)	NA	9 (75)	NA
DRC (25)	13 (52)	21 (84)	22 (88)	NA	25 (100)	16 (64)	16 (64)	14 (56)	NA	18 (72)	NA
Gabon (25)	20 (80)	22 (88)	24 (96)	NA	24 (96)	21 (84)	22 (88)	18 (72)	NA	20 (80)	NA
Senegal (39)	29 (74.4)	32 (82.1)	37 (94.9)	NA	39 (100)	29 (74.4)	27 (69.2)	24 (61.5)	NA	25 (64.1)	NA
Subtotal (140)	**88 (62.9)**	**112 (80)**	**129 (92.1)**	**NA**	**137 (97.9)**	**97 (69.3)**	**101 (72.1)**	**89 (63.6)**	**NA**	**86 (61.4)**	**NA**
Pol											
Cameroon (66)	26 (39.4)	37 (56.1)	54 (81.8)	25 (37.9)	52 (78.8)	35 (53.01)	36 (54.5)	32 (48.5)	41 (62.1)	49 (74.2)	44 (66.7)
Angola (42)	25 (59.5)	34 (80.9)	35 (83.3)	28 (66.7)	40 (95.2)	40 (95.2)	22 (52.4)	21 (50.0)	22 (52.4)	34 (80.9)	25 (59.9)
DRC (27)	11 (40.7)	22 (81.5)	26 (96.3)	16 (59.3)	27 (100)	22 (81.5)	15 (55.6)	14 (51.9)	21 (77.8)	24 (88.9)	22 (81.5)
Gabon (50)	21 (42.0)	31 (62.0)	50 (100)	36 (72.0)	50 (100)	27 (54.0)	27 (54.0)	24 (48.0)	30 (60.0)	35 (70)	37 (74)
Senegal (97)	37 (38.1)	63 (64.9)	88 (90.7)	61 (62.9)	93 (95.9)	63 (64.9)	65 (67.0)	43 (48.5)	69 (71.1)	80 (82.5)	72 (74.2)
Subtotal (282)	**120 (42.6)**	**187 (66.3)**	**253 (89.7)**	**166 (58.9)**	**262 (92.9)**	**187 (66.3)**	**165 (58.5)**	**134 (47.5)**	**183 (64.9)**	**222 (78.7)**	**200 (70.9)**
Env											
Cameroon (65)	27 (41.5)	43 (66.2)	54 (83.1)	NA	56 (86.2)	36 (55.4)	34 (52.3)	31 (47.7)	49 (75.4)	53 (81.5)	46 (70.8)
Angola (00)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
DRC (27)	15 (55.6)	20 (74.1)	25 (82.6)	NA	26 (96.3)	18 (66.7)	16 (59.3)	14 (51.9)	20 (70.1)	20 (70.1)	18 (66.7)
Gabon (22)	16 (72.7)	18 (81.8)	21 (95.5)	NA	22 (100)	19 (86.4)	16 (72.7)	17 (77.3)	18 (81.8)	19 (86.4)	18 (81.8)
Senegal (34)	16 (47.1)	27 (79.4)	33 (97.1)	NA	32 (94.1)	22 (64.7)	23 (67.6)	19 (55.9)	25 (73.5)	29 (85.3)	20 (58.8)
Subtotal (148)	**74 (50)**	**108 (72.9)**	**133 (89.9)**	**NA**	**136 (91.9)**	**95 (64.2)**	**89 (60.1)**	**81 (54.7)**	**112 (75.7)**	**121 (81.8)**	**102 (68.9)**
Grand total (570)	**282 (49.5)**	**407 (71.4)**	**515 (90.3)**	**166 (58.9)**	**535 (93.9)**	**379 (66.5)**	**355 (62.3)**	**304 (53.3)**	**295 (51.8)**	**429 (75.3)**	**302 (52.9)**

NA, Not applicable.

Performance of rapid subtyping tools for HIV-1 gag, pol and env classification

The performance of eleven rapid subtyping tools (REGA v2, REGA v3, NCBI, Stanford HIVDB, SUDI, Geno2Pheno, EuResist, STAR, jpHMM, COMET and SCUEAL) for HIV-1 classification differed between sub genomic regions analyzed (Table 2). A discrepancy was recorded when the rapid subtyping tool assigned no subtype or a different HIV-1 subtype or recombinant from phylogeny.

DISCUSSION

HIV/AIDS pandemic is growing in complexity (Zhang et al., 2010) by the increasing prevalence and spread of CRFs, URFs and other complex recombinants. In Cameroon, the AIDS epidemic

is driven by the genetically diverse recombinants (Peeters, 2001; Ceccarelli et al., 2012) with the CRF02_AG being most predominant (Torimiro et al., 2009; Fokam et al., 2011). Such is the case in some countries in West and Central Africa and thus the need to correctly identify the genetic variants. Phylogeny is often considered to be of very high performance in HIV subtyping. It could be laborious involving the use of multiple software applications as well as well-trained personnel. In this regard subtyping tools have been developed which are faster and more user-friendly.

The performance of these subtyping tools depends on the composition of the reference dataset and method of estimating the relationship between sequences. Most of these tools use a high proportion of subtype B variants in their reference datasets and therefore not efficient for non-B predominant AIDS epidemic. In West and Central Africa where the AIDS epidemic is driven by recombinants, subtyping tools with a high representation of several recombinants are recommended. For pure non-B subtypes, REGA v2 showed a high sensitivity though there was no significant difference with REGA v3 in identifying pure non-B subtypes. NCBI falsely identified some pure non-B variants as CRFs, probably due to the high representation of CRFs in its database. The Stanford HIV Drug Resistance database was unable to identify *gag* and *env* sequences because it is a database built for HIV drug resistance analysis of *pol* sequences as they are targets for antiretroviral drugs. The Stanford HIVDB could not identify any CRF other than CRF01_AE and CRF02_AG reflecting the lack of sufficient sequences of other CRFs in its reference dataset (Yebra et al., 2011).

Apart from identifying CRF01_AE and CRF02_AG, REGA v2 could also identify CRF11_cpx and CRF25_cpx. The performance of REGA v2 to classify other CRFs was poor. REGA v3 however demonstrated improved performance (Pineda-Peña et al., 2013) to CRFs other than CRF01_AE, CRF02_AG, CRF11_cpx and CRF25_cpx. REGA v3 could correctly assign CRFs such as CRF14_BG and CRF47_BF. The sensitivity of REGA v3 to CRFs almost doubled that of REGA v2 showing a great improvement in its performance. The NCBI and SUDI correctly assigned the majority of the CRFs that were available. It is worth noting that NCBI showed the best performance in identifying rare CRFs and URFs as well as recombinant breakpoints (Figure 1). Geno2pheno identified some CRFs as pure subtypes and assigned CRF06_cpx as CRF09_cpx. The jpHMM tool had more of CRF01_AE in its reference dataset such that it identified most CRF02_AG and other CRFs as CRF01_AE. The STAR tool uses Z-scoring and an accumulated identity difference to assign a particular subtype or recombinants and thus wrongly assigning recombinants as pure subtypes.

The difference we observed in this study on the performance of subtyping tools to different HIV-1 variants support the fact that specific subtyping tools can be developed and used for a particular sub genomic region. The reported discrepancies between various rapid subtyping tools and phylogeny could be explained by differences in both the subtyping method applied and the reference sequence dataset (Yebra et al., 2011). HIV-1 variants were identified by phylogenetic analysis using the Neighbor-Joining (NJ) method. Among the rapid subtyping tools evaluated, REGA and SUDI were based on phylogeny applying NJ method as well. Although REGA uses phylogeny combined with bootstrapping, this tool has a threshold which prevents the assignment of a subtype or CRF when it does not obtain enough statistical support. Although SUDI uses phylogeny, its performance is restricted because it does not use bootstrapping and is unable to better assign subtypes or CRFs when bootstrap values are low and with little or no support. Its relatively high performance is due to the Phylogeny Inference Package (PHYLIP) in-built software within its database. NCBI uses a sliding-window along the query sequence and each window is compared to the references by BLAST (Rozanov et al., 2004). However, this sliding-window causes an over-estimation of recombination and further magnifies when more reference sequences were included. With this method, the inclusion of multiple references with shared similarity confounded the results. Some CRFs are genetically very closely related and difficult to discriminate in the studied sub genomic region and thus difficult to obtain reliable results with systems based on BLAST analysis. Stanford HIVDB, however, had multiple pure subtypes, CRF01_AE and CRF02_AG but lacked other CRFs in its reference dataset. The jpHMM, COMET and STAR are statistical-based tools that use position-specific scoring matrices of each subtype to perform profile subtype alignments. Of note, NCBI, Geno2pheno, Euresist and Stanford HIV DB are similarity-based tools while SUDI, REGA v2, REGA v3 and SCUEAL are phylogeny based tools.

Few studies have compared the performance of rapid subtyping tools for HIV classification. Following *in silico* analysis of HIV sequences from Spain, Yebra and colleagues reported that the efficacy of the seven subtyping tools (Stanford HIV DB, Geno2pheno, REGA, NCBI, EuResist, STAR and Therapy Edge) which they analyzed dropped when identifying recombinants other than CRF02_AG. They also reported that only NCBI, REGA and STAR tools could identify URFs but with very low sensitivity (Yebra et al., 2011). These findings are similar to what we report here. These show that the current performance of rapid subtying tools depends on the prevalence of non-B variants and recombinants in the population. In other words, both Yebra and colleagues' and our analyses show that geographic-specific HIV distribution depends on the tool used for subtyping and classification.

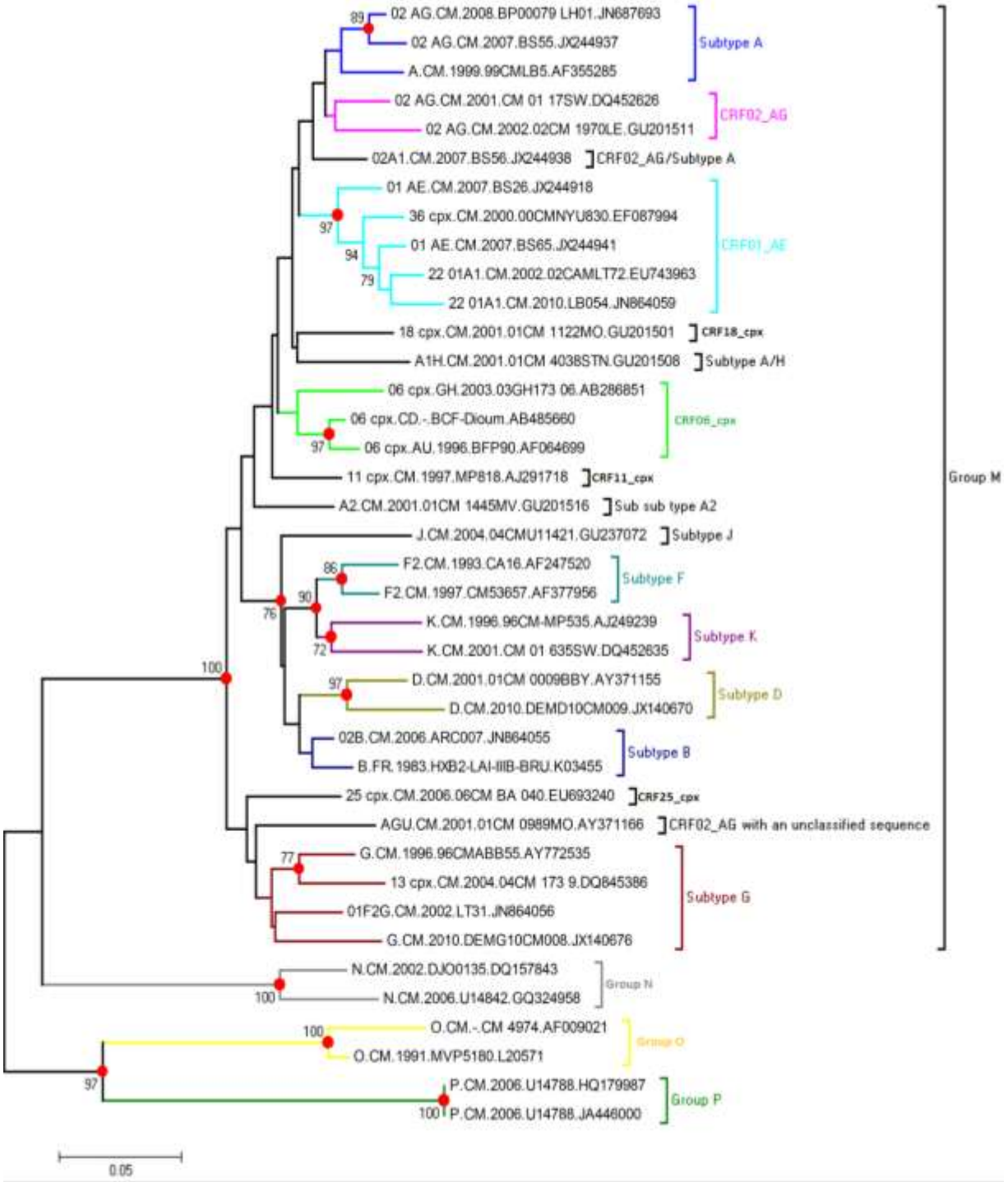

Figure 1. Neighbor-Joining phylogenetic analysis of 39 HIV-1 gag nucleotide sequences from Cameroon.

From these results, Stanford HIVDR database for example cannot be recommended as the only subtyping tool for the interpretation of drug resistance profile. With the growing use of rapid subtyping tools in clinical studies, there is need to define an algorithm using a combination of two or more rapid tools which integrate

phylogeny, statistical-based and similarity-based methods for classification and quality control of results. We also recommend that before using the tools for routine clinical interpretation and patient care, they should be optimized, validated alongside clinical outcome and updated periodically to include larger number of pure subtypes and CRF sequences in the reference databases. To reinforce this, scientists are encouraged to submit to the Genbank or any other virus sequence databases, the sequences published that could be eventually used for *in silico* analyses to better understand the dynamics and evolution of HIV, a virus with the propensity to recombine and mutate to 10^{10} particles daily (Perelson et al., 1996). These findings suggest a revision of method and composition of datasets of subtyping tools for specific clinical and/or epidemiological purposes.

Conflict of interests

The authors have not declared any conflict of interests.

REFERENCES

Banoo S, Bell D, Bossuyt P, Herring A, Mabey D (2010). Evaluation of diagnostic tests for infectious diseases: general principles. Nat. Rev. Microbiol. 8:17-29.

Carr JK, Torimiro JN, Wolfe ND, Eitel MN, Kim B (2001). The AG recombinant IbNG and novel strains of group M HIV-1 are common in Cameroon. Virol. 286:168-181.

Ceccarelli L, Salpini R, Moudourou S, Cento V, Santoro MM (2012). Characterization of drug resistance mutations in naïve and ART-treated patients infected with HIV-1 in Yaoundé, Cameroon. J. Med. Virol. 84:721-727.

Drylewicz J, Matheron S, Lazaro E, Damond F, Bonnet F (2008). Comparison of viro-immunological marker changes between HIV-1 and HIV-2- infected patients in France. AIDS 22:457-468.

Felsenstein J (1985). Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39:783-791.

Fokam J, Salpini R, Santoro MM, Cento V, Perno CF (2011). Drug resistance among drug-naive and first-line antiretroviral treatment-failing children in Cameroon. Pediatr. Infect. Dis. J. 30:1062-1068.

Hall T (2001). BioEdit, Biological sequence alignment for Windows 95/98/NT.

Henguell C, Jacomet C, Antoniotti O, Chaib A, Regagnon C (2008). Difficulties in diagnosing group O Human immunodeficiency virus type 1 acute primary infection. J. Clin. Microbiol. 46:2453-2456.

Ho DD, Neumann AU, Perelson AS, Chen W, Leonard JM (1995). Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. Nature 373:123-126.

Kimura M (1980). A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16:111-120.

MacNeil A, Sarr AD, Sankalé JL, Meloni ST, Mboup S (2007). Direct evidence of lower viral replication rates in vivo in human immunodeficiency virus type 2 (HIV-2) infections than in HIV-1 infection. J. Virol. 81(10):5325-5330.

Menéndez-Arias L, Tözsér J (2008). HIV-1 protease inhibitors: effects on HIV-2 replication and resistance. Trends Pharmacol. Sci. 29:42-49.

Michael NL (1999). Host genetic influences on HIV-1 pathogenesis. Curr. Opin. Immunol. 11:466-474.

Op de Coul ELM, Prins M, Cornelissen M, van der Schoot A, Boufassa F (1997). Using phylogenetic analysis to trace HIV-1 migration among western European infecting drug users seroconverting from 1984 to 1997. AIDS 15:257-266.

Peeters M (2001). The genetic variability of HIV-1 and its implications. Tansfus. Clin. Biol. 8:222-225.

Perelson AS, Neumann AU, Markowitz M, Leonard JM, Ho DD (1996). HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. Science 271:1582-1586.

Pineda-Peña AC, Faria NR, Imbrechts S, Libin P, Abecasis AB (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.

Plantier JC, Leoz M, Dickerson JE, De Oliveira F, Cordonnier F (2009). A new human immunodeficiency virus derived from gorillas. Nat. Med. 15:871-872.

Riemenschneider M, Cashin KY, Budeus B, Sierra S, Elham SD, Bayanolhagh S, Kaiser R, Gorry PR, Heider D (2016). Genotypic Prediction of Co-receptor Tropism of HIV-1 Subtypes A and C. Nat. Sci. Rep. 6:24883

Rozanov M, Plikat U, Chappey C, Kochergin A, Tatusova T (2004). A web-based genotyping resource for viral sequences. Nucl. Acids Res. 32:654-659.

Saitou N, Nei M (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4:406-425.

Tamura K, Dudley J, Nei M, Kumar S (2007). MEGA 4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24:1596-1599.

Temin HM (1993). Retrovirus variation and reverse transcription: abnormal strand transfers result in retrovirus genetic variation. Proc. Natl. Acad. Sci. USA 22:6900-6903.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997). The CLUSTAL X windows interface: Flexible strategies for multiple-sequence alignment aided by quality analysis tools. Nucl. Acids Res. 25:4876-4882.

Torimiro JN, D'Arrigo R, Takou D, Nanfack A, Pizzi D (2009). Human immunodeficiency virus type 1 intersubtype recombinants predominate in the AIDS epidemic in Cameroon. New Microbiol. 32:325-331.

UNAIDS Global report (2014). UNAIDS Report on the global AIDS epidemic 2014. Geneva: UNAIDS; 2014.

Yebra G, de Muldera M, Leticia M, Pérez-Cachafeirob S, Carmen R (2011). Sensitivity of seven HIV subtyping tools differs among subtypes/recombinants in the Spanish cohort of naïve HIV-infected patients (CoRIS). Antivir. Res. 89:19-25.

Zhang M, Foley B, Schultz AK, Macke JP, Bulla I (2010). The role of recombination in the emergence of a complex and dynamic HIV epidemic. Retrovirol. 7:25.

http://www.bioafrica.net/subtype tool/html
http://bioafrica.mrc.ac.za:8080/rega-genotype3.0.2/hiv/typing tool
http://www.ncbi.nlm.nih.gov/projects/genotyping/formpage.cgi
http://sierra2.stanford.edu/sierra/servlet/JSierra?action=sequenceInput
http://www.hiv.lanl.gov/content/sequence/SUDI /sudi.html
http://www.geno2pheno.org
http://engine.euresist.org/data_analysis /viral_sequence/new
http://www.vgb.ucl.ac.uk/starn.shtml
http://jphmm.gobics.de/submission_hiv.html
http://comet.retrovirology.lu
http://www.datamonkey.org/dataupload_scueal.php
http://ww.hiv.lanl.gov/