HIV-1 Genetic Diversity in the Republic of Congo: Seventeen Years in Review

Laure Stella Ghoma Linguissi1,2, Nicaise Ndembi3, Céline Nguefeu Nkenfou4,5, Pierre Poulain1,6,9 and Francine Ntoumi1,10,11#*

1Congolese Foundation for Medical Research, Brazzaville, Republic of Congo
2Département of Biomolecular Research, University of Ouagadougou, Burkina Faso
3Department of Human Virology, University of Maryland, USA
4Department of Biological Sciences, University of Yaoundé I, Cameroon
5Chantal Biya International Reference Centre (CIRCB) for Research on HIV/AIDS Prevention and Management, Cameroon
6National Institutes of Health and Medical Research, France
7Department of System Dynamics and Interactions Biological Macromolecules, Université Paris Diderot, France
8Institut National de la Transfusion Sanguine, DSIMB, France
9Laboratory of Excellence GR-Ex, France
10Departments of Sciences and Techniques, Marien Ngouabi University, Republic of Congo
11Department of Tropical Medicine, University of Tübingen, Germany

#Both authors contributed equally to the work.

Abstract

**Background**: The Republic of Congo is at the epicenter of HIV emergence and it is characterized by a high genetic diversity. Previous studies tried to understand the genetic diversity and strain distribution of HIV-1 since 1990, but all of them were based on small sample sizes and limited to urban areas.

**Objectives**: The main objective of this review was to provide a comprehensive overview and pooled prevalence estimate of different HIV-1 strains circulating in the Republic of Congo between 1999 and 2015.

**Methodology**: We conducted a literature search using the Pub Med database and retrieved research articles related to the genetic diversity of HIV-1 in the Republic of Congo. The results of these published papers were analyzed and the findings are presented in this review.

**Findings**: Subtype a remains the most common strain followed by subtypes C, D, E, G, and H. Several circulating recombinants: CRF01_AE, CRF02_AG, CRF11_cpx, CRF37_cpx, CRF18_cpx, unique recombinant forms: A/CRF01_AE, A/H, A/I, A/G, G/H as well as unclassified strains have been documented.

**Conclusion**: Overall, the high number of HIV-1 subtypes and recombinant viruses in the Republic of Congo suggests the need for a continuous viral surveillance to ensure diagnostic tests and HIV research keep pace with these rapidly evolving viruses.

**INTRODUCTION**

HIV pandemic originated in Kinshasa in the Democratic Republic of Congo (DRC) in the 1920s [1]. Central Africa was the focus of early transmission and the source of pre-1960 pandemic viruses [2]. As a consequence, the greatest genetic diversity of human immunodeficiency virus type 1 (HIV-1) is observed in Africa [3]. Emergence of HIV-1 in human resulted from at least four cross-species transmissions of simian immunodeficiency viruses (SIVs) from chimpanzees and gorillas [4] and the most recent common ancestor is dated around 1908 [5].

The transmission of SIVs from chimpanzees (SIVcpz) to humans placed the origin of the disease in Central Africa [6]. Congolese SIVcpz genomes are mosaic, probably due to a recombinational event in the recent past, and it provides evidence for a rather recently occurring cross-species transmission between humans and chimpanzees [7]. Former study suggested...
that HIV-1 has been introduced into Pygmies through their neighboring Bantu rather than directly from nonhuman primates [8]. Strains from Pygmies and Bantu were similar to those found in the general population.

A viral sequence from 1959 (ZR59) is the oldest HIV-1 infection known so far [9,10]. This viral sequence presented near the ancestral node of subtypes B and D in the major group, indicating that these HIV-1 subtypes, and perhaps all major group viruses, may have evolved from a single introduction into the African population not long before 1959 [9].

In Central Africa, HIV-1 groups M, N, O and P co-circulate in human populations and chimpanzees are infected with genetically closely related viruses [11-19].

HIV-1 group M is divided into nine subtypes (A,B,C,D,F,G,H,J,K) with at least 72 circulating recombinant forms (CRFs) currently identified and thousand unique recombinant forms (URFs) [5,20]. More than 90% of HIV-1 infections worldwide are caused by non-B clades of group M [21].

Epidemiological studies have provided data of HIV-1 distribution and patterns in sub-Saharan Africa [22]. Distribution of HIV-1 subtypes is very heterogeneous [23] and is the result of population migrations [2,24].

The ROC remains a highly endemic area with HIV-1 prevalence at 2.5% [25]. This prevalence is comparable to that of neighboring countries: 1.1% in DRC, 3.9% in Gabon, 2.4% in Angola and 4.3% in Cameroon [26]. The purpose of this review is to provide an overview of published data on HIV-1 genetic diversity, and recombinant forms in the ROC.

METHODOLOGY

Presently, no comprehensive report on the genetic diversity of HIV in the ROC has been done. To collect research data on this topic, research articles on HIV-1 genetic diversity in the ROC have been retrieved from the U.S. National Institutes of Health’s National Library of Medicine (NIH/NLM/Pub Med) database. The query was: (Congo* [Title/Abstract]) NOT (democratic republic [Title/Abstract]) AND (HIV [Title/Abstract]) AND ((recombinant* [Title/Abstract]) OR (subtype* [Title/Abstract])).

A total of 277 nucleotide sequences from the ROC were retrieved from the Los Alamos Database (http://www.hiv.lanl.gov) Due to down-sampling, we focused our phylogenetic analysis on HIV-1 pol gene sequences (HXB2: 2253-3468), which are typically obtained for HIV drug resistance testing. We included 39 HIV-1 pol reference strain and 148 CG sequences in the ClustalW [27] alignment. We performed maximum likelihood phylogenetic reconstruction using PhyML based on General Time-Reversible model with gamma distributed rate variation among nucleotides [28].

RESULTS

Fifteen research articles have been retrieved from the Pub Med database. Seven of them were excluded because their content did not mention any subtype or recombinants of HIV-1. From 1999 to 2015, only 8 HIV molecular characterization studies were conducted in the ROC [29–36].

Study area

Most of the studies were conducted in Brazzaville and Pointe-Noire, mostly for convenience sampling [29–36]. Brazzaville is the capital of the ROC and Pointe-Noire is the largest industrial city. Both cities include about 52% of the Congolese population. Few studies were also conducted in other cities such as Gamboma and Ouesso [32–34] (See Figure 1 for the geographic catchment of cities and neighboring countries of the ROC).

Epidemiology of HIV-1 in the ROC

In the ROC, HIV-related mortality is high with 37% of the deaths due to AIDS in 2008 [37]. The main route of transmission is almost exclusively through heterosexual vaginal intercourse [35,38]. In 2000, HIV prevalence in Pointe-Noire was 14% and 5% in Brazzaville [39]. In 2009, a marked decrease of the prevalence has been noticed, reaching 6.2% and 3.5% in Pointe-Noire and in Brazzaville, respectively [40]. In 2013, the national prevalence rate dropped to 2.5% [26].

Cohort characteristics

Most published studies have been conducted on small sample sizes including pregnant women (see Table 1 for further details): 29 Congolese AIDS patients enrolled in 1996 and 1997 [29], 32 HIV-1 infected patients living in Brazzaville and Pointe-Noire recruited in 1988 and 1992 [32], 28 HIV-1 strains isolated from Congolese AIDS patients used for the study carried out in 1996 and 1997 [36], 114 HIV-1 positive persons enrolled in 2003 and 2004 [34], 30 seropositive pregnant women in Pointe-Noire in 2005 and 2008 [41], 100 patients in Brazzaville recruited in 2011 [35] and finally 95 HIV-1-positive naïve pregnant women in Pointe-Noire enrolled between 2005 and 2008 [31].

Molecular characterization of HIV-1

HIV-1 sequence diversity varies across genes with a difference of 35% in the envelope glycoprotein’s (env) and between HIV-1 groups and sub-subtypes [42,43].

Studies done on the HIV subtypes distribution in the ROC were based on: part of the env region including the V3 loop [29], part of the 59 tat–env (vpu) and env sequences [36] ; the p24 gag region and V3–V5 env region [34] ; env and gag regions and a short segment encoding the gag p7/p9 protein have also been successfully used for phylogenetic analysis [32]. Other analyses were based on the envC2V3 and/or the pol integrate regions [33], pol sequences [41], full protease and partial reverse transcriptase sequencing [31].

In 2006, Niama et al. found that 4.8% (from gag sequences) and 6.3% (from env sequences) of strains could not be classified [34]. Full-length genomic sequences are necessary to identify these unknown strains [44] and to optimally characterize them as potential CRFs, or distinct subtype or sub-subtype [45,46]. In 2012, Pircher et al observed 58% of URFs occurred in the Congolese population. The presence of many URFs may be due to a likely high level of multiple infections (super infections or dual infections) [35].

Time and emergence of different HIV-1 genetic variants

Studies have shown that intra-subtype genetic diversity...
### Table 1: Key elements of HIV studies conducted in the Republic of Congo. U: unclassified subtypes.

<table>
<thead>
<tr>
<th>Area of study</th>
<th>Date of data collection</th>
<th>Cohort size</th>
<th>HIV genes studied</th>
<th>Subtypes (%)</th>
<th>Recombinants (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mboudjeka et al., 1999</td>
<td>Cameroon, ROC, 1995</td>
<td>57</td>
<td>env C2V3 and/or the pol integrase regions</td>
<td>D, F, G and H</td>
<td></td>
</tr>
<tr>
<td>Bikandou et al., 2000</td>
<td>Brazzaville, Pointe-Noire, 1996 and 1997</td>
<td>29</td>
<td>Part of the env region including the V3 loop</td>
<td>A (41), D (3), G (21), H (21), J (7), U (7)</td>
<td></td>
</tr>
<tr>
<td>Taniguchi et al., 2002</td>
<td>Brazzaville, Pointe-Noire, 1996 and 1997</td>
<td>28</td>
<td>part of the 5′ tat-env (vpu) and env</td>
<td>gag: A (3.5), D (3.5), G (60), H (17.8), U (14.2) env: A (39.2), D (3.5), G (17.8), H (21.4), U (7.1), J (7.1), G (3.5)</td>
<td></td>
</tr>
<tr>
<td>Bikandou et al., 2004</td>
<td>Brazzaville, Pointe-Noire, 1998 and 1999</td>
<td>29</td>
<td>env</td>
<td>G (20.4)</td>
<td></td>
</tr>
<tr>
<td>Niama et al., 2006)</td>
<td>Pointe-Noire, Gambia, Ouesso, 2003 and 2004</td>
<td>114</td>
<td>p24 gag and V3-V5 env region</td>
<td>gag: A (36.5), G (30.8), D (12.5) and C, F, H, I, K (15 for all) env: A (32.5) and G (21.3), D (12.5) and C, F, H, I, K (15 for all)</td>
<td>CRF_01, CRF_02, CRF_05, CRF_06, CRF_18</td>
</tr>
<tr>
<td>Pircher et al., 2013</td>
<td>Brazzaville, 2011</td>
<td>100</td>
<td>env</td>
<td>G, A1, B, D, H, F1, A2, C</td>
<td>CRF02_AG, CRF37_cpx, CRF13_cpx, CRF11_cpx, CRF20_BG, CRF21_A2D, CRF33_01BG, CRF02_A2D, CRF37_cpx</td>
</tr>
<tr>
<td>Bruzzone et al., 2015</td>
<td>Pointe-Noire, 2005 to 2012</td>
<td>95</td>
<td>env</td>
<td>G (88), A3 (88), D (4.4), B (2.9), H (2.9), A1, C, F1, F2 (each 1.45)</td>
<td>URF (35), CRF45_cpx (10.3), CRF37_cpx (7.4), CRF18_cpx (5.9), CRF02_A2D (2.9), CRF25_cpx (1.5)</td>
</tr>
</tbody>
</table>

### Table 2: Summary of ARVs resistance-associated mutations observed in the ROC Compilation of results from Pircher et al. [35] and Bruzzone et al. [31]. ARVs are categorized into non-nucleoside reverse-transcriptase inhibitors (NNRTIs) and protease inhibitors (PIs).

<table>
<thead>
<tr>
<th>Patients</th>
<th>Subtypes</th>
<th>Mutations conferring resistance</th>
<th>ARV class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naive</td>
<td>A1</td>
<td>115F</td>
<td>NRTIs</td>
</tr>
<tr>
<td>URF FU</td>
<td>M184V</td>
<td>101E, 103N, 190A, V90I</td>
<td>NRTIs, NNRTIs</td>
</tr>
<tr>
<td>G</td>
<td>90M</td>
<td>E138G</td>
<td>PIs, NNRTIs</td>
</tr>
<tr>
<td>CRF13_cpx</td>
<td>46L</td>
<td>PIs</td>
<td></td>
</tr>
<tr>
<td>A1</td>
<td>115F, 46L</td>
<td>PIs</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>V32I, I54M, I84V, M41L, M184V, L210W, T215Y</td>
<td>PIs</td>
<td>NRTIs, NNRTIs</td>
</tr>
<tr>
<td>C</td>
<td>M184V</td>
<td>G190A, H221Y</td>
<td>NRTIs, NNRTIs</td>
</tr>
<tr>
<td>CRF45_CPX</td>
<td>L210W, T215S</td>
<td>D30N, F53Y, G37S</td>
<td>NRTIs, PIs</td>
</tr>
<tr>
<td>H</td>
<td>K101E</td>
<td>NNRTIs</td>
<td></td>
</tr>
<tr>
<td>CRF02_AG</td>
<td>K65E</td>
<td>V90I</td>
<td>NNRTIs</td>
</tr>
<tr>
<td>F2</td>
<td>V179D</td>
<td>NNRTIs</td>
<td></td>
</tr>
<tr>
<td>J</td>
<td>K65E</td>
<td>NRTIs</td>
<td></td>
</tr>
<tr>
<td>Treated</td>
<td>184V</td>
<td>NRTIs</td>
<td></td>
</tr>
<tr>
<td></td>
<td>215Y/F, 41L, 67N, 70R, 219Q/E, and 210W 69D/N/S, 74V/L, 44D, 75M/A, 215I/N, and 70E</td>
<td>thymidine-associated mutations (TAMs)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>151M</td>
<td>Nucleotide-associated mutation (NAM)</td>
<td></td>
</tr>
</tbody>
</table>
increases with time [47,48]. The temporal multiplicity of passages of SIVcpz or VISsmm to humans may be one of the reasons for the existence of these different groups of HIV-1. In this evolutionary process, HIV is extremely fast in its replication, which leads to a large number of variants currently identified in Central Africa, particularly in DRC and in ROC. This high genetic variability of HIV is also due to several causes including transcription errors in the reverse transcriptase [49,50], either by the large number of virions produced that can carry different mutations [51–53], either by selection pressure on the virus that affects the population level [51,52,54,55], either by new genetic recombinants derived from risk behaviors [51,56–59], that increase the likelihood of multiple infections in the same person [60–62]. This last point is very important in Africa, where several studies have observed risk behaviors for HIV infection in the population [63–67].

**Subtypes, sub-subtypes and recombinants**

Based on the studies listed above, the genotypes observed in the ROC are presented in Table 1. Most HIV infections reported in the ROC are caused by HIV subtypes A, C, D, G, H and inter-subtype recombinants [29–32,34–36,41,44,68]. The emergence of subtypes and unclassified strains is recent and must be closely monitored [29,31]. Figure 2 represents the phylogenetic tree of HIV sequences reported in the ROC with highlight on major subtypes and recombinants.

High frequencies of CRFs were observed in the ROC such as CRF18, CRF19, CRF02_AG, CRF11_cpx, CRF20_BG, CRF21_A2D, CRF37_cpx, CRF25_cpx and CRF45_cpx (see Table 1 for a comprehensive list) [31,35]. Recombinant CRF37_cpx was also found in Cameroon [69]. Recombinant CRF02_AG is the predominant molecular form of HIV-1 found in Kumasi, Ghana [70]. This CRF was also reported predominant (47% frequency) in Gabon [71] in 2 cross sectional surveys performed in 9 cities in Gabon in 2005 and 2008. Bruzzone et al. reported the presence of CRF25_cpx and CRF45_cpx in Pointe-Noire, these recombinants were also found in Angola, the DRC and Gabon [72–74], probably due to the long trading history between these countries.

Studies revealed a high prevalence of URFs: 35% in 2006 [34] and 20% in 2015 [31]. Among all these known recombinants, a study in Pointe-Noire reported a proportion of 57% of putative URFs [75].

The genetic variability of subtype F strains observed in different Central African countries had been studied in depth [71,76]. Phylogenetic analysis of env sequences (V3–V5 region and complete gp160) and of partial gag sequences revealed that subtype F sequences can be divided into three sub-subtypes: F1, F2, and F3 [31,35,77]. The F3 subgroup was composed of strains originating from several Central African countries [78]. F1 et F2 sub-subtypes were reported in the ROC, with 1.5% frequency each [31,35].

The genetic variability of subtype A strains has been observed in different Central African countries and this subtype Ais divided into sub-subtypes A1–A5 [79,80]. The sub-subtypes A1, A2, A4 and A5 were observed in the DRC [81–83]. The sub-subtypes A1, A2 and A3 were also reported in the ROC, at 3, 1 and 8.8% frequency, respectively [31,35]. The sub-subtype A3 was found in Senegal and Guinea-Bissau, as well as in a few neighboring countries in West Africa [84,85]. Bruzzone et al. reported the presence of this sub-subtype in the ROC [31].

<table>
<thead>
<tr>
<th>Subtypes</th>
<th>NNRTIs</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>46L¹</td>
</tr>
<tr>
<td></td>
<td>K70T²</td>
</tr>
<tr>
<td></td>
<td>V106I²</td>
</tr>
</tbody>
</table>

**Figure 1** Map of the Republic of Congo with the cities where studies on HIV took place.
In Central Africa, HIV-1 group O (HIV-O) was found in Cameroon and exhibited a very high genetic diversity [86]. HIV-2 was reported in Angola [87,88]. However, HIV-2 and HIV-1 group O were not detected in isolates from Congolese population [32].

**HIV-1 detection and diagnosis**

Serological diagnosis of HIV-1 infection in sub-Saharan Africa is mostly done with rapid tests such as ImmunoComb HIV-1/2 (Alere/Organics Ltd, Yavne, Israel), Hema-Strip HIV-1/2 (Saliva Diagnostic Systems, SDS, 11719 NE 95th Street, Vancouver, WA 98682, USA), OraQuick HIV-1 (OraSure Technologies Inc, Bethlehem, USA), Determine HIV-1/2 (Alere Medical Co. Ltd, Chiba, Japan), and UniGold Recombigen HIV (Trinity Biotech plc, IDA Business Park, Bray, Co., Wicklow, Ireland) [89]. However, some of these assays have shown limitations in detecting HIV-1 subtypes D, F, H, and recombinant CRF02_AG, HIV-1 group O and HIV-2 [90–95]. It has been reported that some fourth-generation assays presented low sensitivity for the detection of p24 antigen from some non-subtype B HIV-1 strains (A, C, F, H, CRF01_AE) and group O [96,97]. Van Heuverswyn et al. pointed out that HIV-1 genetic diversity has an impact not only on serological but also on nucleic acid-based diagnostics [98].

In the ROC, similar conclusion could be drawn based on this review. Bruzzone et al. reported a 5.5% overestimation of HIV seroprevalence when Determine, instead of Vironostika, was used as second-line test [99]. Studies suggest that HIV-1 genetic diversity may affect the ability of commercially available assays. For instance, the NucliSens Easy Q v.1.2 assay (BioMérieux, Laval, Canada) had difficulty sequencing subtype C, A1, AG, G, and CRF02_AG templates while the Versant HIV-1 RNA 3.0 assay (Siemens Medical Solutions, Mississauga, Canada) had difficulty sequencing B, C, D A1, AG, F1, K, CRF02_AG and non-B subtypes [100,101]. Awareness of any clinical or laboratory differences between the common subtype B of HIV-1 group M and the new HIV-1 strains being seen in practice is therefore increasingly important [102]. For a reliable detection and classification of HIV-1 strains, and in order to minimize the risk of mis-treatment, appropriate reference sequences are needed [41].

**HIV-1 and pathologies**

HIV-1 diversity may influence the course of HIV infection [103–105]. A number of studies have shown that there is a potential association between HIV-1 subtypes and HIV-1 transmission [106–108]. HIV-1 diversity has an impact on the disease progression through the viral replication and the virus pathogenicity [21,38,109,110]. In a longitudinal study in Uganda, Kiwanuka et al. reported that the progression of infection to AIDS disease was shorter in subtype D patients [111]. As this particular subtype is present between 3% and 12.5% in the ROC [29,32,34,36], we can speculate that similar observations
HIV and ARV resistance mutations

In the ROC, two studies had been conducted so far on HIV resistance mutations to ARV. Pircher et al. showed that the resistance to ARV is the major viral causes of treatment failure [35]. Bruzzone et al. reported that Lopinavir-boosted (LPVr) has a very low bio-availability [128]. Table 2 summarizes the results of these studies, which showed a significant resistance to non-nucleoside reverse-transcriptase inhibitors (NNRTIs) and protease inhibitors (PIs) [35].

In Gabon and Indonesia, studies also reported natural ARVs resistance-associated mutations [71,129]. The knowledge of these mutations matters particularly for the introduction of new ARVs or switch therapy for HIV-infected patients [71].

In the ROC, administration of poor quality ARVs may increase the risk of mutations conferring resistance to ARVs [128]. We strongly believe a comprehensive study on ARVs resistance-associated mutations should be conducted in this country.

CONCLUSION

These findings suggest a high genetic diversity and extensive heterogeneity in the ROC. The majority of HIV-1 group M subtypes found in the Central Africa was also detected in the ROC, suggesting a local and historical co-circulation of subtypes A, G, B, C, D, E, and F [32]. Based on a limited number of investigations conducted in only two cities, subtype A was reported to be dominant and many recombinants were also observed. Because results were limited to two cities, we may speculate that it is not a picture of the reality, justifying why extensive studies should be carried out in the ROC to accurately depict the representation of HIV-1 diversity.

ACKNOWLEDGEMENTS

We thank Sylvia Nkombo Nkoula for proofreading the manuscript. PP received financial support from Total E&P Congo. This work has been supported through the Central Africa Network on Tuberculosis, HIV/AIDS and Malaria (CANTAM), which is a network of excellence supported by The European & Developing Countries Clinical Trials Partnership (EDCTP).

REFERENCES


39. IFC Against AIDS - Partnership lists available and reviewed.2014


evidence of extreme diversity at the origin of the HIV-1 group M


