



HAL
open science

High HIV burden and recent transmission chains in rural forest areas in southern Cameroon, where ancestors of HIV-1 have been identified in ape populations

Ginette Edoul, Julius Ebua Chia, Nicole Vidal, Emilande Guichet, Celine Montavon, Eric Delaporte, Eitel Mpoudi Ngole, Ahidjo Ayouba, Martine Peeters

► To cite this version:

Ginette Edoul, Julius Ebua Chia, Nicole Vidal, Emilande Guichet, Celine Montavon, et al.. High HIV burden and recent transmission chains in rural forest areas in southern Cameroon, where ancestors of HIV-1 have been identified in ape populations. *Infection, Genetics and Evolution*, 2020, 84, pp.104358 -. 10.1016/j.meegid.2020.104358 . hal-03490940

HAL Id: hal-03490940

<https://hal.science/hal-03490940>

Submitted on 22 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

High HIV burden and recent transmission chains in rural forest areas in southern Cameroon, where ancestors of HIV-1 have been identified in ape populations.

Ginette Edoul^{1*}, Julius Eboa Chia^{1*}, Nicole Vidal², Emilande Guichet¹, Celine Montavon², Eric Delaporte², Eitel Mpoudi Ngole¹, Ahidjo Ayouba², Martine Peeters²

1. Centre de Recherche sur les Maladies Emergentes et Reemergentes (CREMER), Virology Laboratory IMPM-IRD, IMPM, Yaoundé, Cameroon.
2. TransVIHMI, Institut de Recherche pour le Développement (IRD), INSERM, Université de Montpellier, Montpellier, France.

Corresponding author : Martine Peeters ; IRD, TransVIHMI, 911 avenue Agropolis, 34394 Montpellier Cedex5, France. Tel : + 33- 4 67 41 61 61 ; Fax : + 33 4 67 41 61 46 ; e-mail : martine.peeters@ird.fr.

*these authors contributed equally to the study

Highlights

- HIV prevalence is high in rural forest areas where HIV-1 M originated
- 18.6% of HIV-1 strains in rural forest form 12 recent transmission chains
- New zoonotic pathogens could rapidly establish and spread in rural forest areas
- Today rural forest areas where HIV-1 originated are hot-spots for HIV infection

Abstract

We studied HIV prevalence and genetic diversity in rural forest areas in Cameroon, where chimpanzee and gorilla populations infected with the ancestors of the different HIV-1 groups have been identified and transmitted to humans during the 20th century.

A total of 2,812 individuals were studied, 924 from south-central, 1116 from south-east and 772 from south-west Cameroon. Of 208 (7.4%) samples that were confirmed for HIV-1 infection all belong to HIV-1 group M. In all sites and in all age categories, HIV-1 prevalence was higher in women (160/1599 (10.0%)) as compared to men (48/1213 (4.0%)) with the highest prevalence in women aged between 25 and 34 years (>17%). For 188/208 (92.3%) HIV-1 positive individuals, a fragment of the *pol* gene was successfully amplified and sequenced. Phylogenetic analysis showed predominance of CRF02_AG (58%), a large diversity of subtypes (A, D, F2 and G), seven different CRFs and more than 12% URFs. Interestingly, 35/188 (18.6%) HIV-1 strains form 12 recent transmission chains. The majority of the clusters are composed of two (n=8) or three (n=3) sequences but one cluster included ten HIV-1 strains from women living in four different villages on a major road for logging concessions in the south-east (60 km distance).

In the three regions of Cameroon where the ancestors of the four HIV-1 groups have been transmitted to humans, we observed a high HIV prevalence, especially in the southeast where HIV-1 M originated. Many factors allowing rapid establishment in the human population and subsequent rapid spread to urban areas of a new retrovirus or other pathogens of zoonotic origin are now present. Our study shows clearly that some rural areas should also be considered as hot-spots for HIV infection. Prevention efforts together with growing access to HIV diagnosis and antiretroviral treatment are urgently needed in these remote areas.

1. Introduction.

Cross-species transmissions of simian immune deficiency viruses (SIV), SIVcpz from chimpanzees and SIVgor from gorillas, are at the origin of the four groups of HIV-1 in humans. The zoonotic reservoir of these strains are ape populations from Cameroon; the ancestors of HIV-1 group M and N have been traced to wild chimpanzee populations from south-east and south-central Cameroon, respectively, and the ancestors of HIV-1 group O and P have been identified in wild gorilla populations from south-central and south-west Cameroon, respectively (D'arc et al. 2015, Keele et al. 2006)). Humans became most likely infected when they were exposed to infected blood, secretions or tissues through hunting and butchering non-huma primates (NHP) as bushmeat. However, the four independant cross-species transmissions that are the origin of the different HIV-1 groups have different epidemiological histories (Peeters, Locatelli 2014). Factors associated with single cross-species events have thus to be differentiated from those associated with the epidemic spread of a virus, the latter resulting from a combination of several factors like viral adaptation to the new host and presence of environmental and socio-demographic factors fostering efficient human-to-human transmission (Peeters, Locatelli 2014). Only HIV-1 group M has spread worldwide and is responsible for the global pandemic, with more than 60 million infections worldwide since the description of the first AIDS cases in the 1980's. HIV-1 group O was able to spread to some extent in West Central Africa and represents actually about 0.5% of HIV-1 infections in Cameroon (Villabona-Arenas et al. 2015, Mourez et al. 2013). HIV-1 group N and P have only been observed in a handful of individuals in Cameroon. The absence of epidemiological opportunities most likely played a role in the fact that group O viruses have not spread more widely because the most recent common ancestors (MRCA) are estimated around 1920 for both HIV-1 M and O (Faria et al. 2014). HIV-1 group N and P are more recent transmissions, the MRCA for group N is estimated around 1960 and for group P

the MRCA cannot be calculated with good precision because only two strains have been reported yet (Wertheim et al. 2009, Sauter et al. 2011).

Recent reports also show ongoing transmission of other simian retroviruses to humans in Central Africa, i.e. SFV (simian foamy viruses) and new HTLV variants, closely related to viruses in co-habiting non-human primates (NHPs), have been observed in humans involved in primate hunting and butchering or living in rural forest areas in Cameroon (Calattini et al. 2007; Wolfe et al. 2004; Wolfe et al. 2005; Zheng et al. 2010, Switzer et al. 2012 and 2016). Moreover, the description of HIV-1 group P in 2009 and a new HIV-2 group in 2013 also suggests that our knowledge on HIV diversity and possible cross-species transmissions is still incomplete (Ayouba et al. 2013, Plantier et al. 2009). Zoonotic emergence of new lentiviruses cannot be excluded given the high prevalence of SIVs in some primate species, the increased contact between humans and wild primates through hunting and bushmeat preparation, together with the actual socio-economic and demographic factors predisposing global expansion of viral infections.

Because the four HIV-1 groups originated in rural forest areas in Cameroon, it is important to continue surveillance of HIV-1 diversity in these regions and to evaluate potential risk for spread of a new lentivirus. Here we studied HIV prevalence and genetic diversity of HIV strains that circulate in rural forest areas in Cameroon, especially in areas where wild chimpanzee or gorilla populations infected with the ancestors of the different HIV-1 groups have been identified and transmitted an SIV to humans during the 20th century.

2. Materials and Methods

2.1. Study population

Within the framework of health campaigns organized by the Ministry of Public Health around general health issues (ex. ocular problems, high blood pressure, malaria, helminths ..) an HIV

test was also proposed to participants. These cross-sectional surveys were performed between October 2011 and May 2013 in remote villages located close to sites where HIV-1 group M, N, O and P ancestors have been identified : i.e. Moloundou/Mambele (distance of 20 km) in south-east Cameroon for HIV-1 group M ancestors, Messok (distance of 100 km)for HIV-1 group N and O ancestors in south-central Cameroon and Bipindi (distance of 20 km) in the southwest for HIV-1 group P ancestors (**Figure 1**). The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the National Ethics Committee of Cameroon (N°201/CNE/SE/2011). After participants gave informed consent, they were interviewed by using a verbal standard questionnaire that included the village name, date of birth or age, sex, ethnic group, marital status, level of education, occupation. All individuals who were diagnosed as HIV positive received counselling and were referred to a specialized center for confirmation of HIV infection and access to antiretroviral treatment.

2.2. HIV serology

Whole blood was drawn on EDTA tubes. Plasma and buffy-coats were separated and aliquoted in the field. Serologic screening for HIV infection was done on site with two rapid tests i.e. Alere Determine HIV-1/2 (Abbott Diagnostics) followed by INSTI HIV-1/HIV-2 antibody test (BioLytical Laboratories, Richmond, Canada) when the Determine test was positive. Plasma and buffy coats were stored frozen the same day in local health facilities and shipped at the end of the health campaigns to the central laboratory CREMER, in Yaoundé, for additional testing and subsequent storage at -80°C. Samples reactive on both rapid tests were retested for confirmation of HIV infection and HIV type (HIV-1 or -2) with a line immunoassay (INNO-LIA HIV I/II Score, Innogenetics, Ghent, Belgium). All HIV-1 positive

samples were further typed (HIV-1 group M, N, O or P) with an in-house ELISA based on V3 loop peptides as previously described (Villabona-Arenas et al. 2015).

2.3.PCR amplification and sequencing

Viral nucleic acids were extracted from 140 µl buffycoat with the Nuclisens EasyMag method (BioMérieux, 0Lyon, France). Protease (300 bp) and part of the reverse transcriptase (RT) (800 bp) in the *pol* gene were amplified as a single fragment with an in-house method as previously described (Vergne et al. 2000) and alternatively with the protocol from the Agence Nationale de Recherche sur le SIDA et les Hépatites in France (ANRS) (www.hivfrenchresistance.org) when initial PCR attempts were unsuccessful. PCR products were directly sequenced by using BigDye Terminator v3.1 cycle sequencing kit (LifeTechnologies, France). Electrophoresis and data collection were performed on Applied Biosystems 3500 genetic analyzers. Sequences were assembled using SeqMan Pro v8.1.5 from the DNASTar package (Madison, USA).

Samples that were negative in the above described PCR method, were also tested by nested RT-PCR to amplify partial *env* (315 bp) and *pol* (285 bp) sequences with previously described universal HIV-1/SIVcpz/SIVgor primers to evaluate the presence of divergent HIV-1 strains (Keele et al, 2007; D’Arc et al, 2015). The new sequences have been submitted to Genbank with the following accession numbers: MT425251 to MT425438.

2.4.Phylogenetic analyses

The newly obtained sequences were aligned with reference sequences representing the overall genetic diversity of HIV-1 in West and Central Africa, by using MAFFT version 7 (<https://mafft.cbrc.jp/alignment/server/>). Poorly aligned positions and divergent regions were removed with G-Blocks implemented in Seaview v5.0 (Gouy et al., 2010). Phylogenetic trees

were constructed with GTR+I+G model using Maximum Likelihood (ML) methods implemented in PhyML v3.0 (Guindon et al. 2009). In addition, bootscanning analysis was done to explore any eventual mosaic structure for each strain with Simplot software and confirmed by phylogenetic analysis of the corresponding sub-segments in case of mosaic viruses (Lole et al. 1999, Guindon et al. 2009). In case of mosaic structures, the profiles will be analyzed in detail to check whether they are similar to those described in published circulating recombinant forms (CRFs) and the strains will be assigned as belonging to the matching CRF, if not, the new strains will be classified as unique recombinant Forms (URFs).

Recent transmission clusters were investigated using the statistical robustness of ML topologies based on high bootstrap values (98%) with 1000 re-samplings and short branch lengths following criteria previously defined (Hué et al. 2005). The HIV clusters were also inferred with HIV-TRACE, establishing putative transmission links between HIV sequences with a genetic distance of <1.5% after removing all of the positions associated with major drug resistance mutations from the sequences to reduce their impact on the genetic distance comparison (Kosakovsky Pond SL et al. 2018).

2.5. Drug resistance testing

To evaluate the presence of transmitted HIV-1 drug resistance mutations, *pol* sequences were submitted to the calibrated population resistance (CPR) tool (<http://cpr.stanford.edu/cpr.cgi>) and predicted drug resistance was analyzed using Stanford HIVdb algorithm (version 8.9.1), scores of intermediate or high-level resistance were combined and designated as resistant.

3. Results

3.1. Study population and HIV prevalence

A total of 2,812 individuals were studied, 924 from 8 villages around Messok (south-central), 1116 from 7 villages on the axis Moloundou-Mambele (south-east) and 772 from 7 villages around Bipindi (south-west) (**Figure 1**). The majority were women, 1599 (56.8%), and the age distribution is shown in **Table 1**. HIV confirmatory testing showed that all 208 samples that were confirmed for HIV infection were HIV-1 and serotyping of HIV-1 positive samples did not identify group HIV-1 O, N or P infections. Overall, HIV-1 prevalence was higher in women as compared to men, 160/1599 (10.0%) versus 48/1213 (4.0%) (X^2 test; $p < 10^{-4}$) respectively. This trend is observed in all sites and in all age categories, with the highest prevalence (>17%) in women aged between 25 and 34 years. HIV prevalence varies among the three sites and is highest in the south-east (Moloundou-Mambele) (**Table 1**).

3.2. HIV-1 variants and drug resistance mutations.

For 188 (92.3%) out of the 208 HIV-1 positive individuals, HIV-1 strains have been successfully amplified and sequenced in the *pol* region: 39/42 (92.9%) in Bipindi, 60/65 (92.3%) in Messok and 89/101 (88.1%) in Moloundou/Mambele. Samples that could not be amplified were also negative in PCR attempts using universal HIV-1/SIVcpz/SIVgor primers to amplify partial *pol* and *env* fragments. Phylogenetic analysis was done to identify the HIV-1 subtypes and circulating recombinant forms (CRFs). CRF02_AG predominates and represents 58% of the strains, ranging between 51.7% to 61.8% according to the geographic region. Importantly, the five major CRF02_AG african lineages CRF02_{WA}, CRF02_{CM-I}, CRF02_{CM-II}, CRF02_{CM-III} and CRF02_{CM-IV} as described by Mir and colleagues were represented in nearly balanced proportions (23.8%, 26.6%, 20.2%, 22%, 7.4% respectively) (Mir et al. 2016). The remaining strains represent a large diversity of four subtypes (A, D, F2 and G) and eight CRFs (CRF09_cpx, CRF11_cpx, CRF13_cpx, CRF18_cpx, CRF22_01A1, CRF25_cpx, CRF37_cpx, and CRF45_cpx) (**Table 2**). Importantly, more than 12% of the new HIV-1

strains represent unique recombinant forms (URFs), involving mainly actual circulating CRFs and subtypes. Interestingly two sets of strains among these URFs could potentially represent each a new CRF : (i) three strains have a similar mosaic structure involving CRF02_AG and subtype F2 and (ii) three other strains have a similar mosaic pattern involving subtype D, CRF25_cpx and unclassified fragments (**Figures 2a, 2b**). These two sets of three strains were observed in epidemiologically unlinked individuals from different villages in the extreme south-east.

Mutations associated with drug resistance were observed in 13/188 (6.9%) HIV-1 strains; 2/39 (5.1%) in Bipindi, 4/60 (6.7%) in Messok and 7/89 (7.9%) in Moloundou-Mambele (**Table 2**). NRTI (nucleoside reverse transcriptase inhibitors) mutations were present in 4/188 (2.1%) strains (M41L (n=1), M184V (n=2) and D67N + T69D + M184V + T215Y + K219R (n=1)); NNRTI (non-nucleoside reverse transcriptase inhibitors) mutations were seen in 7/188 (3.7%) strains (Y188L (n=2), K101E (n=2), K103N (n=2) and K103N + P225H (n=1)) and PI (protease inhibitors) mutations were seen in 5/188 (2.7%) strains (M46I/L (n=4) and L90M (n=1)). The PI mutations, especially M46I/L most likely represent natural polymorphisms observed in non-B HIV-1 variants and HIV drug resistance algorithm did not predict resistance to PI in these patients.

3.3.Recent transmission chains

Interestingly, 35 (18.6%) of the 188 HIV-1 strains fall in 12 clusters that correspond to recent transmission chains because clusters are supported with bootstrap values >98% and show less than 1.5% divergence as confirmed by phylogenetic tree analysis with PhyML (**Figure 3**). Similar results were observed by using HIV-Trace. The majority of the clusters (n = 8) are composed of two sequences, three clusters contained three sequences and one larger cluster included ten HIV-1 strains. The highest number of recent transmission chains were seen in the

area of Messok with 17/60 (28.3%) of strains that were present in seven clusters. Three transmission chains included HIV-1 strains from three individuals, one male and two females in each cluster and with the male individual having links with a female individual from the same and a neighboring village. Five other clusters contained each only two strains, with three clusters that included only female individuals and two clusters including one male and one female individual. For one of these clusters, the individuals were from neighboring villages. In Moloundou-Mambele, 16/89 (18%) of the strains were present in four clusters, including one large transmission cluster of ten HIV-1 strains from women living in four different villages on the axis Moloundou-Mambele (60 km distance). The three other clusters in that area were all from individuals living in Mambele (n=2) or Moloundou (n=1) and included one female and one male individual for two of them and two females for the third one. Finally, in the Bipindi area, we observed one cluster with two strains isolated from one male and one female individual from the same village. Moreover, one new strain obtained from an individual around Messok, forms a recent transmission chain with a previously described CRF22_01A1 strain that has been reported in previous study on retroviruses in rural forest areas, and that was sampled around Lomie at 50 km distance from Messok (Carr et al. 2010). Another new strain from Messok also fell very close to a previously described CRF25_cpx strain from the same study (Carr et al. 2010).

4. Discussion

We observed high HIV prevalence, especially in women (10.0%), high numbers of recent transmission clusters and unique recombinant forms (URFs) in rural forest areas in southern Cameroon where HIV-1 ancestors have been transmitted from apes to humans. The high HIV prevalence observed in our study is close to the prevalence observed in the national demographic health survey (DHS) conducted in 2011; i.e. between 8.8% and 10.6% in the

south (includes Bipindi) and east (includes Messok and Moloundou-Mambebe) versus 7.4% to 9.9% in females in our study, and between 3.3% and 3.8% in the DHS versus 2.9% to 5.2% in males from our study. (<https://dhsprogram.com> > *pubs* > *pdf*). The high genetic diversity observed corresponds to what was reported in a previous study conducted in rural areas from Cameroon, where CRF02_AG predominated with 66.5% and co-circulates with subtypes D, F2 and G and numerous CRFs (CRF01_AE, CRF11_cpx, CRF13_cpx, CRF18_cpx, CRF22_01A1, CRF25_cpx, and CRF37_cpx) together with 10% of unique recombinant forms (URFs) (Carr et al. 2010). Within the CRF02_AG radiation, the five major lineages previously described in Cameroon are present (Mir et al. 2016). Overall, the number of patients with predicted drug resistance was still low during the study period, i.e. below 5% for NRTI, NNRTIs and PIs, and is close to what is reported for rural areas in a recent nationwide study conducted in 2015 (Tchouwa et al. 2019).

The areas of our study are those where the wild ape populations that were at the origin of the four HIV-1 groups have been identified (D'Arc et al. 2015). The ancestor of HIV-1 group M, responsible for the global AIDS pandemic, has been transmitted to humans in southeastern Cameroon which is different from the area that is considered as the epicenter of the global HIV epidemic, i.e. Kinshasa, the capital city of the Democratic Republic of Congo (DRC) (Keele et al. 2006, Vidal et al. 2000, Rambaut et al. 2001). HIV-1 M virus got established in Kinshasa during its initial development as a city most likely because of the highly impaired gender balance (5 to 10 males for 1 female) and the high rates of sexually transmitted infections which can favor HIV transmission (Pepin J 2011). The initial growth of HIV-1 M matched closely with population growth and the expansion of HIV-1 M in Central Africa was apparently associated with the development of transportation networks that allowed increased mobility of human populations in DRC (Faria et al. 2014 and 2019). The rapid increase of HIV-1 M around 1960 was associated with iatrogenic interventions and

changes of sexual behavior in DRC and expanding connections between countries (Faria et al. 2014 and 2019, Pepin et al. 2008, Pepin 2011).

This unique combination of circumstances leading to the establishment of HIV-1 M in the human population and allowing its rapid spread were not present at the time of the initial cross-species transmission events in southern Cameroun around early 1900 and during the major part of the 20th century. However today, more than a century later, the socio-economic environment changed in southern Cameroun. Population density increased in forest areas, especially around logging/mining concessions, as well as the mobility of human populations between rural forest and urban areas, related to the development of road infrastructure allowing better and faster connections from these remote areas with urban centers. As illustrated by the data on HIV prevalence, human behavior associated with high rates of HIV transmission are now present. The HIV-1 strains that circulate today in these areas have been introduced more recently and have spread successfully in these areas that were initially characterized by the absence or low levels of HIV (Faria et al. 2012).

Successful cross-species transmissions depend on several factors, frequency and modes of contacts with infected animals and prevalence of the pathogen in the animal host species (Peeters, Locatelli 2014). The growing number of people in rural forest areas leads to more hunting of wildlife for bushmeat and thus offers increased opportunities for cross-species transmission. Today, many factors associated with the initial establishment of HIV in Kinshasa and the rapid spread to other areas are now present in rural forest areas in southern Cameroon, suggesting that introduction of a new pathogen could become established faster in these areas and spread rapidly to other densely populated areas. The rapid spread of a virus has been illustrated by the large transmission chain on the axis Moloundou-Mambele, where ten women from four different villages on this road have been infected with the same index case. This road is a major axis for the logging concessions to the harbor of Douala. The

additional small clusters also illustrate the high number of sexual partners, because five of the eight clusters with two individuals included only two female individuals, and the clusters with three individuals included two female and one male. These clusters also show absence of prevention and information campaigns in these rural areas.

Inter-subtype/CRF recombinants are the result of co- or super-infection with different subtypes/CRFs and are more frequently observed in populations with high-risk behaviour. The high number of URFs in our study is an additional evidence of high-risk behavior in rural areas and are close to those observed in high-risk populations such as MSM (men having sex with men) in urban settings in Africa that can range between 4% in Senegal to 19% in Togo (Konou et al. 2016, Leye et al. 2013, Ndiaye et al. 2009). High-risk populations are also characterized by presence of high numbers of recent transmission chains, for example 40% to 60% of HIV-1 strains in MSM populations in Senegal, Togo and Kenya belong to such clusters (Konou et al. 2016, Ndiaye et al. 2009, Hassan et al. 2018). In our study, up to 18.6% of the HIV-1 strains are part of recent transmission chains. Interestingly, high HIV prevalence and high numbers of transmission clusters are also observed in populations of fishing communities around lake Victoria in Uganda, showing that certain populations in rural areas should also be considered as at high-risk for HIV infection (Kiwuwa-Muyingo S et al. 2017, Bbosa et al. 2019). Interestingly two sets of strains among these URFs could potentially represent each a new CRF, however this should be further confirmed by full-length genome sequencing.

As a consequence of the high HIV prevalence, a significant proportion of the population in rural forest areas have also to some extent immune deficiency which could be in favor for the adaptation of a new pathogen transmitted from an animal reservoir. Recombination occurs frequently among closely related HIV-1 M viruses, i.e. between different HIV-1 subtypes, but also between more divergent strains like HIV-1 group O and M

(Peeters et al. 1999, De Oliveira F. et al 2017). No recombination has been documented yet between HIV-1 and HIV-2, but studies on SIVs in wild NHP showed several examples of recombination between distant SIV lineages (Peeters M, Locatelli 2014). The presence of an existing HIV-1 virus that recombines with a new SIV could also lead to faster adaptation of a new virus to human host. Finally, the growing population density in forest areas leads to more hunting of wildlife and thus increased exposure to a wide diversity of animal pathogens. Because a significant part of the population has immune deficiency, this could not only lead to faster adaptation of a new simian lentivirus, but also for many other pathogens from animal origin to which humans are exposed either by hunting and butchering wild life, or by transmission via insects or other vectors (Lebreton et al 2007).

5. Conclusions

In the three regions of Cameroun where the ancestors of the four HIV-1 groups have been transmitted to humans, we observed a high HIV prevalence, especially in the southeast where HIV-1 M originated. We observed also high numbers of recent transmission chains and a high genetic diversity. With the increased number of individuals with immune deficiency having frequent contacts with wildlife and exposure to animal pathogens together with high risk behavior for HIV transmission and high mobility in these areas, many factors allowing rapid establishment in the human population and subsequent rapid spread to urban areas of a new retrovirus or other pathogens of zoonotic origin are now present. Therefore it remains important to continue the surveillance for HIV diversity in these areas. Our study shows clearly that today, some rural areas should also be considered hot-spots for HIV infection. These remote regions should become a priority for prevention and treatment, and countries should not only focus on urban areas and high-risk populations. Prevention efforts together with growing access to HIV diagnosis and antiretroviral treatment are urgently needed in

these remote areas in order to reach the 90-90-90 UNAIDS objectives to end the AIDS epidemic.

Declarations of interest : The authors declare that they have no conflict of interest.

Acknowledgements

We thank Dr Etoundi Balla Serges Alain from the Ministry of Health, Pr Zoung-Kanyi Bissek Anne-Cecile, BISSEK Anne_Cécile from the Directorate of the Fight against Disease, Epidemics and Pandemics, the medical staff, the mayors and population from Messok, Moloundou, Mambele and Bipindi. We also thank the staff from CREMER for logistical support in the field and the laboratory.

Funding: This work was supported in part by grants from the National Institute of Health (RO1 AI 50529), the Agence Nationale de Recherches sur le SIDA (ANRS 12125/12182/12325) and IRD (Institut de Recherche pour le Développement).

References

- Ayoub A, Akoua-Koffi C, Calvignac-Spencer S, Esteban A, Locatelli S, Li H, Li Y, Hahn BH, Delaporte E, Leendertz FH, Peeters M. Evidence for continuing cross-species transmission of SIVsmm to humans: characterization of a new HIV-2 lineage in rural Côte d'Ivoire. *AIDS*. 2013;27(15):2488-91. doi: 10.1097/01.aids.0000432443.22684.50.
- Bbosa N, Ssemwanga D, Nsubuga RN, Salazar-Gonzalez JF, Salazar MG, Nanyonjo M, Kuteesa M, Seeley J, Kiwanuka N, Bagaya BS, Yebra G, Leigh-Brown A, Kaleebu P. Phylogeography of HIV-1 suggests that Ugandan fishing communities are a sink for, not a source of, virus from general populations. *Sci Rep*. 2019;9(1):1051. doi: 10.1038/s41598-018-37458-x.
- Calattini S, Betsem EB, Froment A, Maucière P, Tortevoeye P, Schmitt C, Njouom R, Saib A, Gessain A. Simian foamy virus transmission from apes to humans, rural Cameroon. *Emerg Infect Dis*. 2007;13(9):1314-20. doi: 10.3201/eid1309.061162.
- Carr JK, Wolfe ND, Torimiro JN, Tamoufe U, Mpoudi-Ngole E, Eyzaguirre L, Birx DL, McCutchan FE, Burke DS. HIV-1 recombinants with multiple parental strains in low-prevalence, remote regions of Cameroon: evolutionary relics? *Retrovirology*. 2010;7:39. doi: 10.1186/1742-4690-7-39.
- D'arc M, Ayoub A, Esteban A, Learn GH, Boué V, Liegeois F, Etienne L, Tagg N, Leendertz FH, Boesch C, Madinda NF, Robbins MM, Gray M, Cournil A, Ooms M, Letko M, Simon VA, Sharp PM, Hahn BH, Delaporte E, Mpoudi Ngole E, Peeters M. Origin of the HIV-1 group O epidemic in western lowland gorillas. *Proc Natl Acad Sci U S A*. 2015;112(11):E1343-52. doi: 10.1073/pnas.1502022112.
- De Oliveira F, Mourez T, Vessiere A, Ngoupo PA, Alessandri-Gradt E, Simon F, Rousset D, Plantier JC. Multiple HIV-1/M + HIV-1/O dual infections and new HIV-1/MO inter-group recombinant forms detected in Cameroon. *Retrovirology*. 2017;14(1):1. doi: 10.1186/s12977-016-0324-3.
- Faria NR, Suchard MA, Abecasis A, Sousa JD, Ndemi N, Bonfim I, Camacho RJ, Vandamme AM, Lemey P. Phylodynamics of the HIV-1 CRF02_AG clade in Cameroon. *Infect Genet Evol*. 2012;12(2):453-60. doi: 10.1016/j.meegid.2011.04.028.
- Faria NR, Rambaut A, Suchard MA, Baele G, Bedford T, Ward MJ, Tatem AJ, Sousa JD, Arinaminpathy N, Pèpin J, Posada D, Peeters M, Pybus OG, Lemey P. HIV epidemiology. The early spread and epidemic ignition of HIV-1 in human populations. *Science*. 2014;346(6205):56-61. doi: 10.1126/science.1256739.
- Faria NR, Vidal N, Lourenco J, Raghwani J, Sigaloff KCE, Tatem AJ, van de Vijver DAM, Pineda-Peña AC, Rose R, Wallis CL, Ahuka-Mundeke S, Muyembe-Tamfum JJ, Muwonga J, Suchard MA, Rinke de Wit TF, Hamers RL, Ndemi N, Baele G, Peeters M, Pybus OG, Lemey P, Dellicour S. Distinct rates and patterns of spread of the major HIV-1 subtypes in Central and East Africa. *Plos Pathogens* 2019. 15(12): e1007976. <https://doi.org/10.1371/journal.ppat.1007976>

Gouy M, Guindon S, Gascuel O. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol.* 2010;27(2):221-4. doi: 10.1093/molbev/msp259.

Guindon S, Delsuc F, Dufayard JF, Gascuel O. Estimating maximum likelihood phylogenies with PhyML. *Methods Mol Biol.* 2009;537:113-37. doi: 10.1007/978-1-59745-251-9_6.

Hassan AS, Esbjörnsson J, Wahome E, Thiong'o A, Makau GN, Price MA, Sanders EJ. HIV-1 subtype diversity, transmission networks and transmitted drug resistance amongst acute and early infected MSM populations from Coastal Kenya. *PLoS One.* 2018;13(12): e0206177. doi: 10.1371/journal.pone.0206177.

Hué S, Pillay D, Clewley JP, Pybus OG. 2005. Genetic analysis reveals the complex structure of HIV-1 transmission within defined risk groups. *Proc Natl Acad Sci U S A.* 102(12):4425-4429.

Keele BF, Van Heuverswyn F, Li Y, Bailes E, Takehisa J, Santiago ML, Bibollet-Ruche F, Chen Y, Wain LV, Liegeois F, Loul S, Ngole EM, Bienvenue Y, Delaporte E, Brookfield JF, Sharp PM, Shaw GM, Peeters M, Hahn BH. Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. *Science.* 2006;313(5786):523-6.

Kiwuwa-Muyingo S, Nazziwa J, Ssemwanga D, Ilmonen P, Njai H, Ndembi N, Parry C, Kitandwe PK, Gershim A, Mpendo J, Neilsen L, Seeley J, Seppälä H, Lyagoba F, Kamali A, Kaleebu P. HIV-1 transmission networks in high risk fishing communities on the shores of Lake Victoria in Uganda: A phylogenetic and epidemiological approach. *PLoS One.* 2017;12(10):e0185818. doi: 10.1371/journal.pone.0185818.

Konou AA, Vidal N, Salou M, Anato S, Singo-Tokofäi A, Ekouevi DK, Pitché P, Prince-David M, Delaporte E, Peeters M, Dagnra AY. Genetic diversity and transmission networks of HIV-1 strains among men having sex with men (MSM) in Lomé, Togo. *Infect Genet Evol.* 2016;46:279-285. doi: 10.1016/j.meegid.2016.05.030.

Kosakovskiy SL, Weaver S, Leigh Brown AJ, Wertheim JO. HIV-TRACE (TRANSMISSION Cluster Engine): a Tool for Large Scale Molecular Epidemiology of HIV-1 and Other Rapidly Evolving Pathogens. *Mol Biol Evol.* 2018;35(7):1812-1819. doi: 10.1093/molbev/msy016.

LeBreton M, Yang O, Tamoufe U, Mpoudi-Ngole E, Torimiro JN, Djoko CF, Carr JK, Tassy Prosser A, Rimoin AW, Birx DL, Burke DS, Wolfe ND. Exposure to wild primates among HIV-infected persons. *Emerg Infect Dis.* 2007;13(10):1579-82. doi: 10.3201/eid1310.070338

Leye N, Vidal N, Ndiaye O, Diop-Ndiaye H, Wade AS, Mboup S, Delaporte E, Toure-Kane C, Peeters M. High frequency of HIV-1 infections with multiple HIV-1 strains in men having sex with men (MSM) in Senegal. *Infect Genet Evol.* 2013;20:206-14. doi: 10.1016/j.meegid.2013.09.002.

Locatelli S, Peeters M. Cross-species transmission of simian retroviruses: how and why they could lead to the emergence of new diseases in the human population. *AIDS.* 2012;26(6):659-73. doi: 10.1097/QAD.0b013e328350fb68. Review.

Lole, K.S., Bollinger, R.C., Paranjape, R.S., Gadkari, D., Kulkarni, S.S., Novak, N.G., Ingersoll, R., Sheppard, H.W., Ray, S.C., 1999. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J. Virol.* 73, 152–160.

Mourez T, Simon F, Plantier JC. Non-M variants of human immunodeficiency virus type 1. *Clin Microbiol Rev.* 2013;26(3):448-61. doi: 10.1128/CMR.00012-13. Review.

Ndiaye HD, Toure-Kane C, Vidal N, Niama FR, Niang-Diallo PA, Dièye T, Gaye-Diallo A, Wade AS, Peeters M, Mboup S. Surprisingly high prevalence of subtype C and specific HIV-1 subtype/CRF distribution in men having sex with men in Senegal. *J Acquir Immune Defic Syndr.* 2009;52(2):249-52. doi: 10.1097/QAI.0b013e3181af70a4

Peeters M, Liegeois F, Torimiro N, Bourgeois A, Mpoudi E, Vergne L, Saman E, Delaporte E, Saragosti S. Characterization of a highly replicative intergroup M/O human immunodeficiency virus type 1 recombinant isolated from a Cameroonian patient. *J Virol.* 1999;73(9):7368-75.

Pépin J. *Origin of AIDS.* Cambridge University Press, Cambridge, UK, 2011 (305 pages).

Pépin J. The expansion of HIV-1 in colonial Leopoldville, 1950s: driven by STDs or STD control? *Sex Transm Infect.* 2012;88(4):307-12. doi: 10.1136/sextrans-2011-050277.. Review.

Pépin J, Labbé AC. Noble goals, unforeseen consequences: control of tropical diseases in colonial Central Africa and the iatrogenic transmission of blood-borne viruses. *Trop Med Int Health.* 2008;13(6):744-53. doi: 10.1111/j.1365-3156.2008.02060.x.

Plantier JC, Leoz M, Dickerson JE, De Oliveira F, Cordonnier F, Lemée V, Damond F, Robertson DL, Simon F. A new human immunodeficiency virus derived from gorillas. *Nat Med.* 2009;15(8):871-2. doi: 10.1038/nm.2016.

Rambaut A, Robertson DL, Pybus OG, Peeters M, Holmes EC. Human immunodeficiency virus. Phylogeny and the origin of HIV-1. *Nature.* 2001;410(6832):1047-8. DOI: 10.1038/35074179.

Sauter D, Hué S, Petit SJ, Plantier JC, Towers GJ, Kirchhoff F, Gupta RK. HIV-1 Group P is unable to antagonize human tetherin by Vpu, Env or Nef. *Retrovirology.* 2011;8:103. doi: 10.1186/1742-4690-8-103.

Switzer WM, Bhullar V, Shanmugam V, Cong ME, Parekh B, Lerche NW, Yee JL, Ely JJ, Boneva R, Chapman LE, Folks TM, Heneine W. Frequent simian foamy virus infection in persons occupationally exposed to nonhuman primates. *J Virol.* 2004;78(6):2780-9. DOI: 10.1128/jvi.78.6.2780-2789.2004

Switzer WM, Tang S, Zheng H, Shankar A, Sprinkle PS, Sullivan V, Granade TC, Heneine W. Dual Simian Foamy Virus/Human Immunodeficiency Virus Type 1 Infections in Persons from Côte d'Ivoire. *PLoS One.* 2016;11(6):e0157709. doi: 10.1371/journal.pone.0157709. eCollection

Switzer WM, Tang S, Ahuka-Mundeke S, Shankar A, Hanson DL, Zheng H, Ayouba A, Wolfe ND, LeBreton M, Djoko CF, Tamoufe U, Esteban A, Heneine W, Peeters M, Wright LL, Muyembe-Tamfum JJ, Wemakoy EO, Mulembakani P, Hoff NA, Rimoin AW. Novel simian foamy virus infections from multiple monkey species in women from the Democratic Republic of Congo. *Retrovirology*. 2012;9:100. doi: 10.1186/1742-4690-9-100.

Tchouwa GF, Eymard-Duvernay S, Cournil A, Lamare N, Serrano L, Butel C, Bertagnolio S, Mpoudi-Ngole E, Raizes E, Aghokeng AF; Members of the EHRICA Study Group. Prevalence of pretreatment HIV drug resistance in Cameroon following a nationally representative WHO survey. *J Antimicrob Chemother*. 2018;73(9):2468-2474. doi: 10.1093/jac/dky221.

Vergne L, Peeters M, Mpoudi-Ngole E, Bourgeois A, Liégeois F, Toure-Kane C, M'Boup S, Mulanga-Kabeya C, Saman E, Jourdan J, Reynes J, Delaporte E. Genetic diversity of protease and reverse transcriptase in non-subtype-B human immunodeficiency virus type 1 strains: Evidence of many minor drug resistance mutations in treatment-naïve patients. *J Clin Microbiol*. 2000;38:3919–3925.

Vidal N, Peeters M, Mulanga-Kabeya C, Nzilambi N, Robertson D, Ilunga W, Sema H, Tshimanga K, Bongo B, Delaporte E. Unprecedented degree of human immunodeficiency virus type 1 (HIV-1) group M genetic diversity in the Democratic Republic of Congo suggests that the HIV-1 pandemic originated in Central Africa. *J Virol*. 2000;74(22):10498-507.

Villabona-Arenas CJ, Domyeum J, Mouacha F, Butel C, Delaporte E, Peeters M, Mpoudi-Ngole E, Aghokeng AF. HIV-1 group O infection in Cameroon from 2006 to 2013: Prevalence, genetic diversity, evolution and public health challenges. *Infect Genet Evol*. 2015;36:210-216. doi: 10.1016/j.meegid.2015.09.006.

Wertheim JO, Worobey M. Dating the age of the SIV lineages that gave rise to HIV-1 and HIV-2. *PLoS Comput Biol*. 2009;5(5):e1000377. doi: 10.1371/journal.pcbi.1000377.

Wolfe ND, Heneine W, Carr JK, Garcia AD, Shanmugam V, Tamoufe U, Torimiro JN, Prosser AT, Lebreton M, Mpoudi-Ngole E, McCutchan FE, Birx DL, Folks TM, Burke DS, Switzer WM. Emergence of unique primate T-lymphotropic viruses among central African bushmeat hunters. *Proc Natl Acad Sci U S A*. 2005;102(22):7994-9. 23.

Zheng H, Wolfe ND, Sintasath DM, Tamoufe U, Lebreton M, Djoko CF, Dikko Jle D, Pike BL, Heneine W, Switzer WM. Emergence of a novel and highly divergent HTLV-3 in a primate hunter in Cameroon. *Virology*. 2010;401(2):137-45. doi: 10.1016/j.virol.2010.03.010.

Figure legends

Figure 1: Study sites (white dots for main city and dotted line for the area covered), areas where HIV-1 ancestors (M, N, O, and P) have been identified in ape populations are shown with red and yellow stars.

Figures 2: Mosaic structure of recombinant strains identified as putative new CRFs (circulating recombinant forms): recombinant form 1 (RF1) (**Fig 2a**) and recombinant form 2 (RF2) (**Fig 2b**). The mosaic structure was determined by boot-scanning and confirmed with ML phylogenetic tree analysis of the different sub-fragments, as indicated in Materials and Methods.

Figure 3: Phylogenetic relationships of 1730 unambiguously aligned nucleotide sequences of *pol* gene (protease and partial reverse transcriptase) for 188 samples from the study. The maximum likelihood tree was inferred with PhyML v3.0 under the GTR+I+Γ4 model of nucleotide substitution. The best of SPR and NNI heuristic options was selected and branch support was estimated with the approximate likelihood ratio (aLRT) test. Branches are colored according to the geographic origin of each sequence, orange for Moloundou-Mambele, pink for Messok and blue for Bipindi. HIV-1 reference sequences are in black. The recent transmission chains, identified as indicated in Materials and Methods, are highlighted with two types of colored boxes (light yellow for patients from the same village, light grey for individuals coming from different villages). The aLRT support value of each identified subtype/CRF clade is >0.85. Asterisks refer to unique recombinant strains (URFs). Potential new recombinant clades are indicated (RF1 and RF2).

Table 1: HIV prevalence in the three different study sites, shown in Figure 1, according to gender (Female (F), Male (M) and age.

Age category (years)	gender	Bipindi		Messok		Molondou/ Mambele		Total	
		N	(%)	N	(%)	N	(%)	N	(%)
18-24	F	3/81	(3.7%)	7/123	(5.7%)	20/154	(13.0%)	30/358	(8.4%)
	M	0/47	(0.0%)	3/83	(3.6%)	2/85	(2.4%)	5/215	(2.3%)
25-34	F	12/103	(11.7%)	25/116	(21.5%)	33/189	(17.5%)	70/408	(17.2%)
	M	0/71	(0.0%)	2/96	(2.1%)	9/133	(6.8%)	11/300	(3.7%)
35-44	F	6/83	(7.2%)	9/97	(9.3%)	11/125	(8.8%)	26/305	(8.5%)
	M	0/65	(0.0%)	5/96	(5.2%)	9/100	(9.0%)	14/261	(5.4%)
>44	F	11/164	(6.7%)	10/181	(5.5%)	13/183	(7.1%)	34/528	(6.4%)
	M	10/158	(6.3%)	4/132	(3.0%)	4/147	(2.7%)	18/437	(4.1%)
Total	F	32/431	(7.4%)	51/517	(9.9%)	77/651	(11.8%)	160/1599	(10.0%)
	M	10/341	(2.9%)	14/407	(3.4%)	24/465	(5.2%)	48/1213	(4.0%)
Total	F+M	42/772	(5.4%)	65/924	(7.0%)	101/1116	(9.1%)	208/2812	(7.4%)

Table 2 : Details on distribution of subtypes, circulating recombinant forms (CRF) and unique recombinant forms (URFs) in the three different study sites.

Subtype/CRF	Messok		Moloundou/ Mambele		Bipindi		Total	
	N=60		N=89		N=39		N=188	
	n	(%)	n	(%)	n	(%)	n	(%)
A	0	-	3	(3.4)	2	(5.1)	5	(2.7)
CRF02_AG	31	(51.7)	55	(61.8)	23	(59.0)	109	(58.0)
CRF09_cpx	3	(5.0)	0	-	0	-	3	(1.6)
CRF11_cpx	6	(10.0)	6	(6.7)	1	(2.6)	13	(6.9)
CRF13_cpx	2	(3.3)	1	(1.1)	0	-	3	(1.6)
CRF18_cpx	0	-	2	(2.2)	1	(2.6)	3	(1.6)
CRF22_01A1	2	(3.3)	1	(1.1)	3	(7.7)	6	(3.2)
CRF25_cpx	1	(1.7)	1	(1.1)	0	-	2	(1.1)
CRF37_cpx	1	(1.7)	3	(3.4)	1	(2.6)	5	(2.7)
CRF45_cpx	1	(1.7)	0	-	1	(2.6)	2	(1.1)
D	2	(3.3)	1	(1.1)	0	-	3	(1.6)
F2	0	-	3	(3.4)	1	(2.6)	4	(2.1)
G	3	(5.0)	2	(2.2)	1	(2.6)	6	(3.2)
URFs	8	(13.3)	11	(12.4)	5	(12.8)	24	(12.8)

Table 3 : Details on levels of drug resistance to protease (PI) and reverse transcriptase inhibitors (NRTI and NNRTI) and the different mutations observed in the 188 newly described HIV-1 strains from this study.

Drug resistance mutations (DRM)	N	%	predicted drug Resistance ^a
Any PI DRM		5	2.7
M46L/I		4	2.1
L90M		1	0.5
Any NRTI DRM		4	2.2
M41L		1	0.5
M184V		2	1.1
D67N, T69D, M184V, T215Y, K219R		1	0.5
Any NNRTI DRM		7	3.7
K101E		2	1.1
K103N		2	1.1
Y188L		2	1.1
K103N, P225L		1	0.5
To any of PI, NRTI or NNRTI		13	6.9

^a : only intermediate and high level drug resistance were considered
abbreviations of the drugs : 3TC (Lamvidine), FTC (Emtricitabine), ABC (Abacavir), AZT (Zidovudine), TDF (Tenofovir), EFV (Efavirenz), NVP (Nevirapine), ETR (Etravirine), DOR (Doravirine)

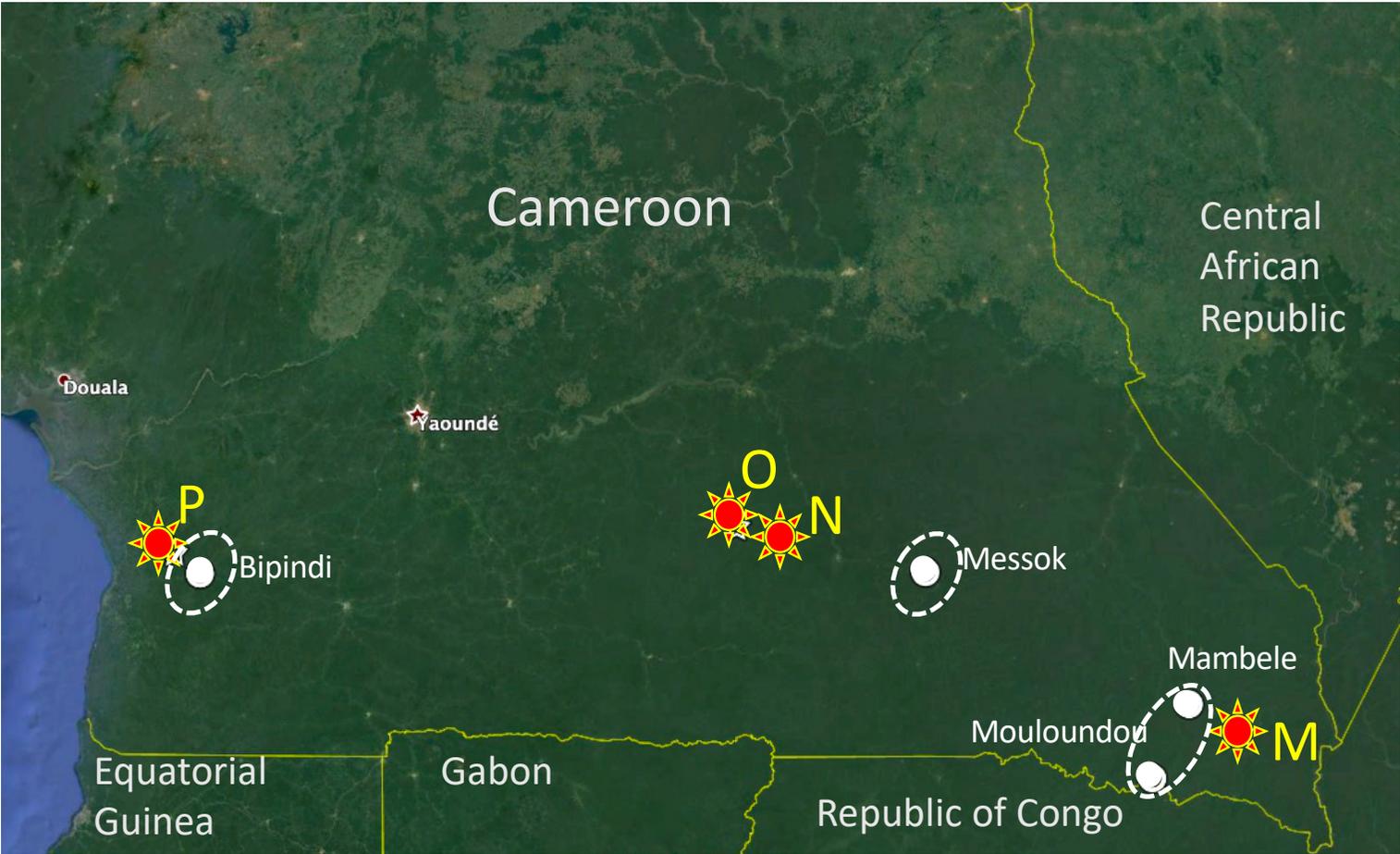


Figure 2a – RF1

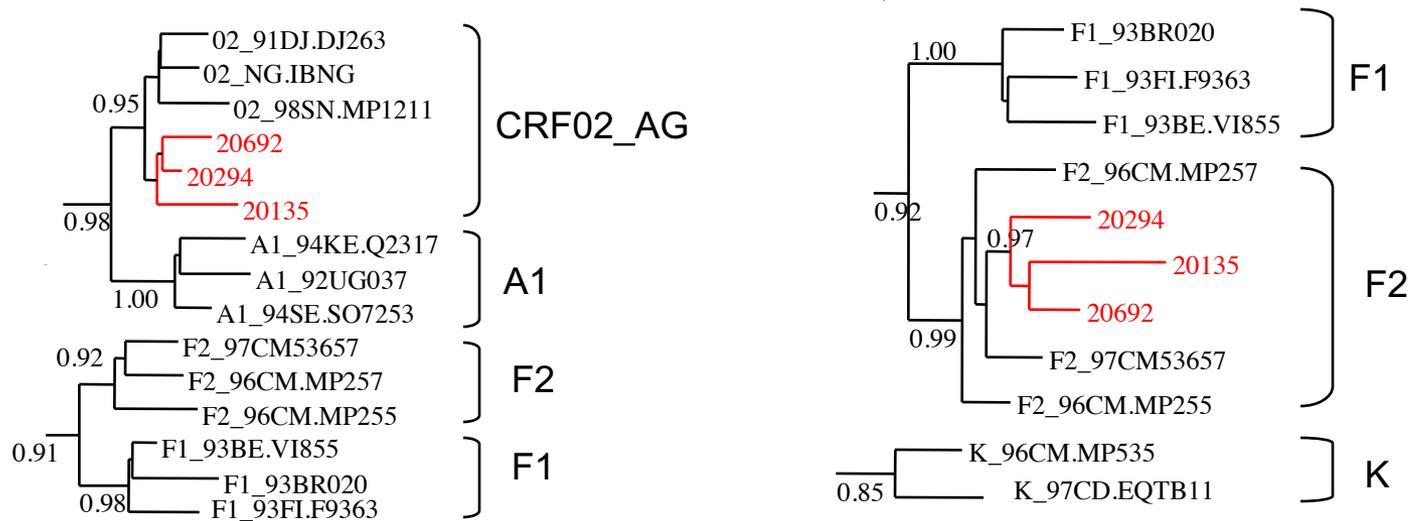
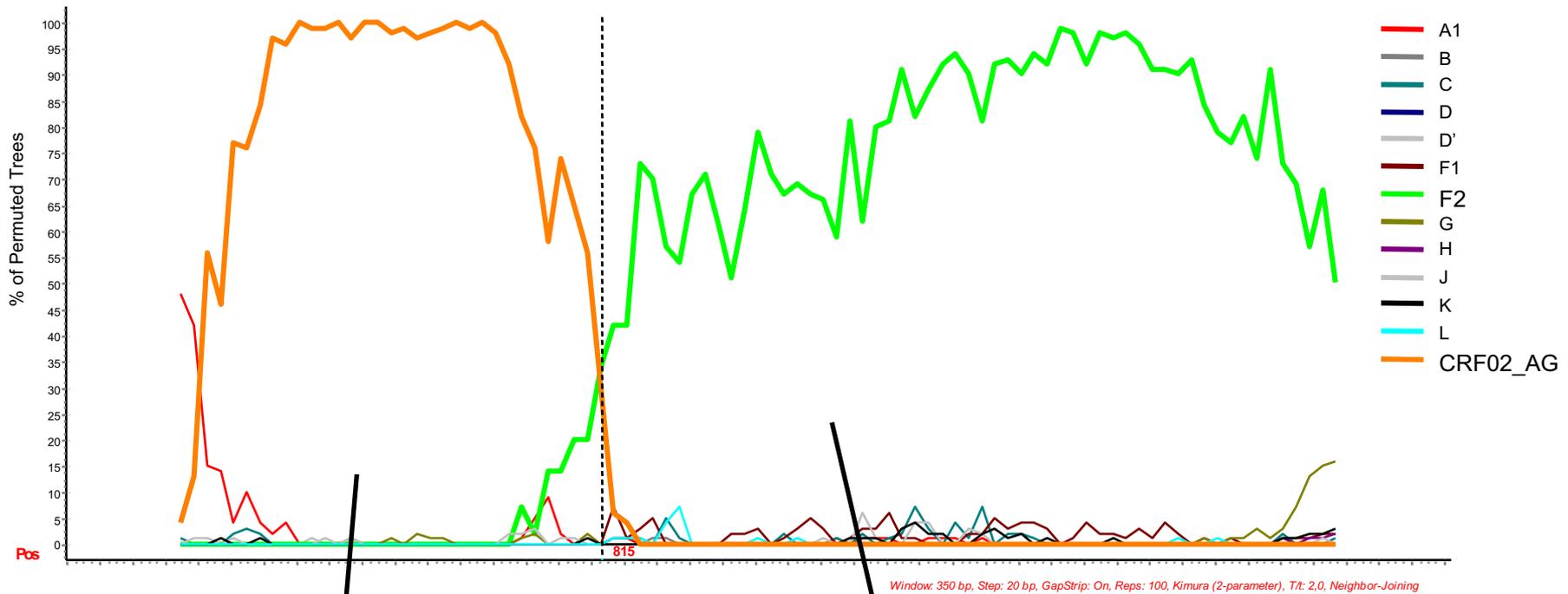


Figure 2b – RF2

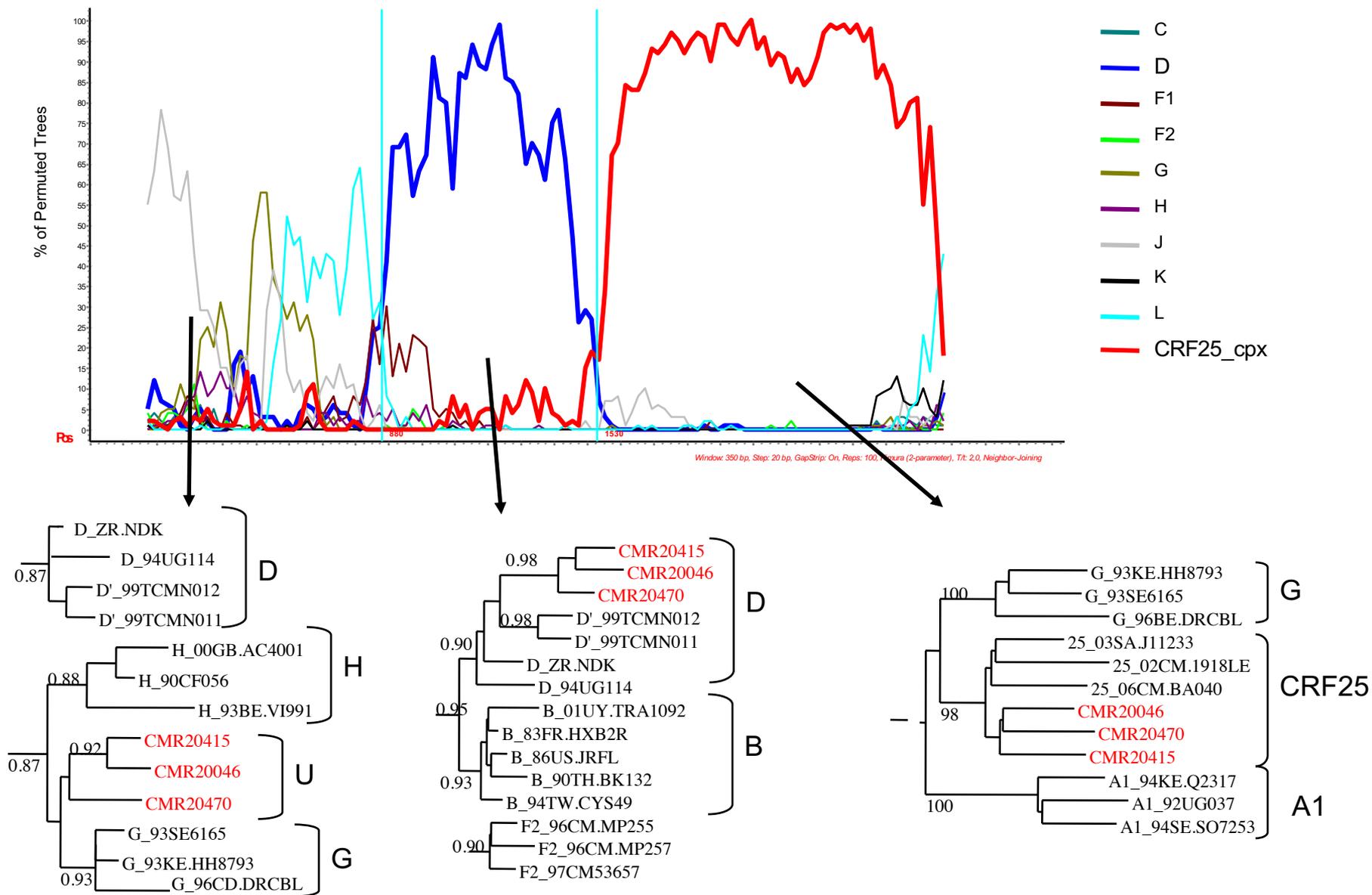


Figure 3

- Messok
- Bipindi
- Moloundou

