Research Connections, 2026, 1, vmag069
https://doi.org/10.1093/rescon/vmag069
Published: 16 May 2026
Research Article

Temporal trends in HIV-1 genetic diversity in Cameroon: a sequence database analysis from 1990 to 2023

Derrick Tambe Ayuk Ngwese[1,2], Alex Durand Nka[1,3], Akinde Mbuh Nji[4,5], Aude Christelle Ka'e[1,#], Désiré Takou[1], Bouba Yagai[1], Ezechiel Ngoufack Jagni Semengue[1], Collins Ambe Chenwi[1,6], Grace Angong Beloumou[1,6], Naomi-Karell Etame[1,6], Sandrine Djupsa[1], Aurelie Minelle Kengni Ngueko[1,6], Forgwei Lum[1], Efakika Gabisa Jeremiah[1,5], Tatiana Tekoh Anim-Keng[1], Larissa Gaelle Moko Fotso[1,2], Luna Colagrossi[7], Marina Potesta[6], Maria Santoro[6], Francesca Ceccherini Silberstein[6], Giulia Capelli[1,7], Charles Kouanfack[8], Nadia Mandeng[9], Edie Gregory Halle-Ekane[10], Nicaise Ndembi[11], Rogers Ajeh Awoh[12], Vittorio Colizzi[1,6], Eduan Wilkinson[13], Cheryl Baxter[13], Tulio De Oliviera[13], Marcel Yotebieng[14], Marta Giovanetti[15,16], Jalloh Mohamed[17], Charles Fokunang[2], Wilfred Fon Mbacham[2,4,5], Alexis Ndjolo[1,2], Carlo Federico Perno[1,18], Joseph Fokam[1,2,3,10,*]

[1]Chantal BIYA International Reference Centre for Research on HIV/AIDS Prevention and Management, Yaoundé, Cameroon
[2]Faculty of Medicine and Biomedical Sciences, University of Yaoundé 1, Yaoundé, Cameroon
[3]National AIDS Control Committee/Central Technical Group (CNLS/GTC), Yaoundé, Cameroon
[4]Biotechnology Centre, University of Yaoundé 1, Yaoundé, Cameroon
[5]Fobang Institutes for Innovations in Science and Technology, Simbock, Yaoundé, Cameroon
[6]University of Rome, Rome, Italy
[7]National Research Council, Rome, Italy
[8]Centre for Research on Emerging and Re-emerging Diseases (CREMER), Yaoundé, Cameroon
[9]Operational Research and Health Department, Ministry of Public Health, Yaoundé, Cameroon
[10]Faculty of Health Sciences, University of Buea, Buea, Cameroon
[11]International Vaccine Institute (IVI), Africa Regional Office (IARO), Rwanda
[12]International Epidemiology Databases to evaluate AIDS (IeDEA), Central Africa
[13]Centre for Epidemic Response and Innovation (CERI), School of Data Science and Computational Thinking, Stellenbosch University, Stellenbosch, South Africa
[14]Department of Medicine, Albert Einstein College of Medicine, Bronx, NY, United States
[15]Department of Sciences and Technologies for Sustainable Development and One Health, Universitada Campus Bio-Medico di Roma, Italy
[16]Laboratório de Flavivírus, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil
[17]Centre of Disease and Control (CDC), Atlanta, GA, United States
[18]Ospedale Pediatrico Bambino Gesu, Italy
[#]Authors contributed equally to this work.
[*]Corresponding author: National AIDS Control Committee Chief of Service, Virology Laboratory, CIRCB, Yaoundé, Cameroon.
E-mail: josephfokam@gmail.com.

Editor: Zara Quail

Abstract

Background and aim **HIV exhibits extensive genetic diversity with important consequences for viral fitness, diagnostic performance, therapeutic response, and the emergence of drug resistance. In the context of major changes in antiretroviral therapy (ART) availability, we investigated long-term HIV-1 clade dynamics in Cameroon over more than three decades to inform projections of viral evolution toward 2030.**

Methodology **A total of 12 230 HIV sequences generated between 1990 and 2023 were analyzed and grouped into three periods reflecting ART rollout: 1990–2000 (minimal ART availability), 2001–11 (reverse-transcriptase-inhibitor-era), and 2012–23 (high-coverage-highly-potent ART era). Temporal trends in clade proportions were assessed using the Cochran–Armitage trend test. Logistic regression models were fitted to evaluate the association between calendar year and the probability of observing specific HIV-1 clades. Forecasting analyses were done with generalized additive models to estimate future clade proportions up to 2030. Model fit was evaluated using Akaike Information Criterion (AIC), to confirm robustness of the models, and Shannon entropy measured the genetic variability of the Pol gene across the time-point intervals.**

Received: January 26, 2026. **Revised:** March 27, 2026. **Accepted:** May 5, 2026

Results **HIV-1 overwhelmingly predominated throughout the study period (99.98%), with HIV-2 remaining rare (0.02%). Within HIV-1, Group M was consistently dominant and increased significantly over time, while Groups O, N, and P declined markedly. Circulating-recombinant forms (CRFs), particularly CRF02_AG-related lineages, expanded substantially and became the major drivers of viral diversity, whereas unclassified recombinants declined. Weighted analysis confirmed that these temporal trends were not artifacts of uneven sequencing efforts, addressing a key limitation commonly associated with sequence database studies. Forecasting analyses suggest that by 2030, HIV-1 group M is likely to remain the predominant circulating lineage in Cameroon, although the persistence of rare clades cannot be excluded given the limitations of available sequence data. Furthermore, Entropy analyses showed slight differences in genetic variability over the years, suggesting overall stability of the Pol region despite decades of viral evolution as compared to Env gene.**

Conclusion **These findings underscore the central role of recombination in shaping HIV-1-evolution in Cameroon and highlight the necessity of continuous genomic surveillance and next-generation therapeutic and preventive strategies to support progress toward HIV-elimination by 2030 in genetically diverse settings.**

Keywords HIV-1 genetic diversity, clade dynamics, forecasting analysis, circulating recombinant forms (CRFs), Cameroon

Introduction

AIDS remains a major global health concern with an estimated 40.8 million people living with HIV (PLHIV) globally by the end of 2024. Notably, with the increasing availability of antiretroviral therapy (ART) worldwide, there has been 54% reduction in HIV-associated deaths and 40% reduction in new HIV infections since 2010 [1]. Nevertheless, HIV remains a public health threat with 630 000 deaths and 1.3 million new cases in 2004. The global spread and evolution of HIV infection are characterized by marked geographical diversity of HIV types 1 (HIV-1) and 2 (HIV-2), with prevailing HIV-1 subtypes, circulating recombinant forms (CRFs), and unclassified recombinant forms (URFs), leading to large regional variation in numbers, types, and proportions of HIV-1 genetic variants. While HIV-2 is prevalent in West and Central Africa (WCA) sub-regions, HIV-1 drives the epidemic in other African regions across Europe, Asia, and the Americas [2]. So far, 140 CRFs have been identified, and a large number of URFs have been reported. It is foreseeable that a growing number of CRFs will be reported in the near future due to the high recombination potential of HIV-1 [3]. WCA is believed to be the geographical origin of HIV, with the Congo basin considered the most likely epicenter of HIV-1 groups M, N, O, and P. This is further underscored by the very broad genetically diverse HIV epidemics in WCA as compared to other geographical settings of the globe [4]. Of note, the circulation of every known HIV-1 group M pure subtype, several CRFs such as CRF02_AG, CRF22_01A1, CRF18_cpx, CRF37_cpx, CRF11_cpx, CRF13_cpx, A1, CRF01_AE, CRF09_cpx, and a variety of URFs have been observed within this geographical area [5, 6]. The most widely circulating variant in WCA is the CRF02_AG lineages [7], and Cameroon is known to have a high genetic diversity [6, 8–11].

HIV variability is a major stumbling block to the successful design of an effective AIDS vaccine, an HIV cure strategy, a lifelong successful ART, diagnostic and treatment monitoring assays, including viral loads and sequencing platforms [2]. This challenge arises from the extraordinary viral genetic diversity, with multiple subtypes and strains circulating globally [12]. This diversity arises from the virus's rapid mutation rate (i.e. 3.4×10^{-5} mutations per base in one replication cycle), particularly in the Env protein, due to the lack of proofreading by reverse transcriptase enzyme, which results in millions of variants in an infected individual within 24 h [13]. These variants can exhibit significant molecular and structural alterations, allowing the virus to evade host-mounted antibodies as well as existing ART regimens. The high genetic diversity of HIV-1 has been reported in the Cameroonian context, but with limited assessment on potential programmatic implications [8, 10, 14, 15]. Local evidence also demonstrated that determining the HIV status of some individuals remains challenging due to multidimensional factors such as flaws in diagnostic systems, technological challenges, and viral diversity [16]. Moreover, the weakness of the sample referral system and the limited local availability of advanced technologies, among other reasons, continue to be a challenge in confirming the HIV status of some individuals in several African countries sharing similar programmatic features like Cameroon [17].

Updating the molecular epidemiology of HIV strains in Cameroon is essential to optimize the national strategy toward AIDS elimination by 2030. This need is reinforced by major shifts in treatment guidelines over time, from the early era of unboosted protease inhibitors (1990–2000), to widespread use of reverse-transcriptase inhibitors (2001–11), and more recently the rollout of test-and-treat policies using dolutegravir-based regimens (2016–present). This study aimed to provide an updated overview of HIV diversity in Cameroon, examine evolutionary trends of viral clades from 1990–2023, and predict their trajectories for 2024–30.

Materials and methods

Study populations

HIV-1 sequences were retrieved from the Los Alamos National Laboratory (LANL) HIV sequence database, which is a publicly accessible biological database online where HIV sequences, both full and partial genomes, using different sequencing technologies from many countries in the world, are stored. In addition, sequences were obtained from the institutional repository of the Chantal Biya International Reference Centre for Research on HIV/AIDS Prevention and Management (CIRCB). This repository contains partial genomes of the Gag and Pol genes, including protease, reverse transcriptase, and integrase regions, generated using Sanger sequencing. CIRCB is a WHO-accredited research institution for HIV drug resistance testing, serving both Cameroon and the Central African region. Sequences from both databases were selected independently of antiretroviral treatment status and study population characteristics.

According to our inclusion criteria, only sequences originating from Cameroon were included, where full-length sequences and partial genomic regions were sampled between 1990 and 2023 with a minimum length of 500 nucleotides. Laboratory clones, synthetic constructs, and sequences with excessive ambiguous nucleotides were excluded. Duplicate sequences were identified using accession numbers and patient identifiers when available. To avoid redundancy, all CIRCB sequences previously submitted to the Los Alamos database were excluded from the analysis. In cases where longitudinal sequences were available from the same individual in the same time point interval, only the earliest sequence was retained. Moreover, sequences obtained from HIV LANL had no data concerning population characteristics such as treatment failure, key populations, gender, naïve or ART-experienced patients, while sequences obtained from CIRCB are from patients who come for HIV drug resistance testing as per routine activities carried out in the laboratory. Subtypes assigned to sequences from the LANL database were directly incorporated into the analysis. In contrast, subtypes for CIRCB sequences initially assigned using the Stanford tool (with a similarity score >5%) were further validated using two additional bioinformatics tools, REGA and COMET.

Sample processing and statistical analysis

Temporal stratification

All HIV-1 sequences from LANL and CIRCB were retrieved with their sampling dates, genomic regions, and accession numbers. Sequences were grouped into three time intervals, and temporal trends for each HIV-1 clade were evaluated using proportion trend test implemented in R software version 1.4.1106. A $P < .05$ was considered statistically significant.

Because our dataset spans the period from 1990 to 2023, we assessed temporal changes in the distribution of HIV groups, pure subtypes, and recombinant forms. To facilitate this analysis, we applied temporal stratification using approximately decade-long intervals. This approach ensured adequate sample sizes within each period while minimizing biases introduced by uneven sequencing across years. These intervals broadly align with key phases of HIV program implementation in Cameroon, namely the pre-ART era, the early ART scale-up phase, and the subsequent expansion of optimized treatment strategies. However, we acknowledge that these transitions are gradual and overlapping rather than strictly discrete. Importantly, this temporal grouping is intended for descriptive purposes and does not imply a direct causal relationship between ART policies and viral strains distribution. Based on these considerations, the dataset was stratified into three time intervals;

- 1990–2000: Limited or totally absent antiretroviral treatment (ART) coverage
- 2001–11: Widespread use of nucleotide and non-nucleotide reverse transcriptase inhibitors.
- 2012–23: Test-and-Treat Strategy, and early tenofovir-lamivudine-dolutegravir (TLD) era collectively termed as the high-potent ARV era.

Statistical analysis

Temporal changes in the proportions of HIV-1 groups, subtypes, and circulating recombinant forms (CRFs) were assessed using the Cochran–Armitage trend test for ordered proportions across defined time periods.

In addition, logistic regression models were fitted to evaluate the association between calendar year and the probability of observing specific HIV-1 clades, with year treated as a continuous variable. To account for substantial variability in sequencing volume across years, clade proportions were weighted by the total number of sequences available per year. Year-by-year sequence counts were incorporated into the analysis to minimize sampling bias. Years with no available sequences were excluded from the trend test but retained in regression modelling where appropriate using continuous year interpolation.

Forecasting analyses were conducted using generalized linear models (GLMs) with a binomial distribution to estimate future clade proportions up to 2030 [18]. Predictions were presented with 95% confidence intervals (95% CI) to reflect uncertainty. Model fit was evaluated using Akaike Information Criterion (AIC), and residual diagnostics were assessed to confirm the robustness of the models [19].

The proposed statistical model formula was used; $\text{Logit}(P) = \beta 0 + \beta 1 \times \text{Year}$ where

- P = probability of a given HIV-1 clade
- Year = continuous variable

Entropy analysis

Additionally, after carrying out prediction analysis, we did an entropy evaluation of our sequence data for the three structural genes of HIV-1 *gag*, *pol*, and *ENV,* where we studied their genetic variability to predict epidemic trends and estimate which gene would infer new infections over time [20–23]. This was done using the following steps:

a) Retreival and alignment of HIV sequences

We retrieved all the full sequences of HIV-1 gag, pol, env from the HIV Los Alamos sequence database in the *FASTA* format (https://www.hiv.lanl.gov). We then carried out three sets of multiple alignments for each gene of sequences using the HIV reference sequence (HXB2) still retrieved from the HIV sequence database. These alignments were done using the CLUSTAL W tool of Bioedit software version 7.2 [24].

b) Clean and Format the Alignment

After we aligned each set of sequences, we cleaned them to remove gaps or ambiguous residues using the 'Edit' tool of Bioedit software version 7.2.

c) Analysis of the aligned sequence for measuring genetic diversity using Shannon Entropy

Entropy measures the variability within a site and assigns a high score to highly variable sites and a lower score to less variable sites. The three sets of complete HIV-1 *gag, pol* and *Env* sequences, which were aligned to the HXB2 reference sequence, were submitted to the entropy tool of LANL (https://www.hiv.lanl.gov/content/sequence/ENTROPY/entropy) known as the Shannon entropy. Furthermore, the genetic variability across the HIV-1 Pol gene was assessed using Shannon entropy calculated at each codon position. Entropy values were computed for each

time period (1990–2000, 2001–11, 2012–23) using aligned nucleotide sequences.

To account for unequal sample sizes across time periods, a random subsampling approach was applied, whereby an equal number of sequences were randomly selected from each period over multiple iterations, and average entropy values were calculated.

Statistical comparisons of entropy distributions between time periods were performed using permutation tests (10 000 iterations) to assess differences in mean entropy values without assuming normality, were done using the following steps;

1. Entropy values were combined from two periods
2. Randomly shuffled (permute) labels between groups
3. Recalculated the difference in mean entropy
4. Repeated this many times (e.g. 10 000 iterations)
5. Computed the Probability of observing a difference as large as our dataset, where the probability was considered as the *P*-value

Summary entropy values were expressed as mean ± standard deviation, and significance was defined at $P < .05$.

Entropy estimates were interpreted within the expected range of variability reported for the HIV-1 pol gene in the literature, as full codon-level entropy calculations require sequence-level alignment data, which were not uniformly available across all datasets.

Results

General description of the analyzed HIV sequences

A total of 12 230 HIV sequences from Cameroon were included, comprising 12 228 HIV-1 sequences and two HIV-2 sequences. Of the HIV-1 sequences, 9582 were retrieved from the LANL HIV database and 2646 from the CIRCB-CARE database. Full genomes represented 2.4% (294 HIV-1 sequences) while partial genomes constituted the remaining 97.6% (11 934 HIV-1 sequences (Supplementary Figure 1).

Prevalence and distribution of HIV-1 viral clades

Overall, HIV-1 group M was predominant 92.42% (11 301/12 230), followed by HIV-1 group O 4.99% (611/12 230), group N 0.99% (121/12 230), group P 0.17% (22/12 230) and 1.4% (175/12 230) were sequences with no HIV-1 group assignment meanwhile HIV-2 had a proportion of 0.02% (2/12 230). The proportion of circulating recombinant forms (CRFs) was 72.83% (8907/12 230), followed by pure-subtype 19.58% (2395/12 230) and unclassified-recombinants (URFs) 1.2% (147/12 230) (Supplementary Figure 2).

Temporal trends in HIV-1 clade distribution

Temporal changes in HIV-1 clade distribution were assessed using both non-parametric and regression-based approaches. The Cochran–Armitage trend test demonstrated significant temporal trends in the distribution of major HIV-1 groups and subtypes across the study periods ($P < .001$). Specifically, HIV-1 Group M showed an increasing trend from earlier periods (1990–2000) to the most recent period (2012–23), consistent with its predominance in the epidemic. In contrast, rare groups such as HIV-1 Groups O, N, and P exhibited a declining trend over time, although their low frequency warrants cautious interpretation, as shown in Supplementary Table 1. The distribution of the HIV-1 viral clades from timepoint-1 (1990–2000) showed that group M was the predominant group with a rate of 63.65% (830/1304), where pure subtypes had a proportion of 14.57% (190/1304) and recombinant forms were already dominant 49.08% (640/1304). This is because of the high rate of CRF02_AG circulating in the population, 30.52% (398/1304), as compared to CRF02_AG derivatives (recombinant forms closely and phylogenetically related to 02_AG lineages), which had a proportion of 13.65% (178/1304) and non-CRF02_AG forms, which had a proportion of 3.68% (48/1304). Group O was also very high in this time point, with a rate of 32.21% (420/1304), while group N had a rate of 2.22% (29/1304). Group P did not appear in the population at this time point (0/1304).

Concerning the distribution of HIV-1 viral clades of time point 2 (2001–11), group M rose significantly with a rate of 96.13% (5763/5995), where pure subtypes had a proportion of 18.93% (1135/5995) and the proportion of recombinant forms was high 77.19% (4628/5995) as compared to time point 1 ($P < .0001$). This steady increase in recombinant forms in time point 2 was driven by an increase in CRF02_AG derivatives of 31.08% (1907/5995) as compared to time-point 1, with a proportion of 13.65 ($P < .001$). CRF02_AG forms still remained the highest with a rate of 36.65%, while non-CRF02_AG forms still remained low with a rate of 4.56% (273/5995). Group O and group N forms had an occurrence of 2.30% (137/5995) and 1.20% (72/5995), respectively. Group P appeared in the circulation at this time point with a proportion of 0.37% (22/5995).

Lastly, concerning time point-3 (2012–23), group M still maintained its large dominance with a rate of 99.14% (4414/4452) where pure subtypes slightly increased with a proportion of 21.81% (971/4452) as compared to time point 2 ($P < .0001$). Moreover, recombinant forms maintained a constant rate of 77.34% (3443/4452) as compared to time point 2 ($P = .44$) and this was driven by a high rate of CRF02_AG with a proportion of 59.14% followed by an increase in CRF02_AG derivatives with a rate of 41.28%, while non-CRF02_AG forms still remained low with a rate of 1.93% (86/4452). Group O and group N remained very low with prevalence of less than 1%, while group P completely disappeared from the circulation at this time point. The distribution of all the HIV-1 clades in our study population can be seen in Supplementary Figure 3.

Logistic regression results of temporal trends

Logistic regression analysis confirmed that calendar year was significantly associated with changes in clade distribution. When the year was treated as a continuous variable, the probability of detecting HIV-1 Group M increased significantly over time ($\beta > 0$, $P < .001$), while the probability of detecting Groups O/N/P decreased ($\beta < 0$, $P < .05$).

These associations remained consistent after weighting for annual sequencing volume, indicating that the observed temporal trends were not solely attributable to variations in sampling intensity across years. Moreover, the comparison between weighted and unweighted trend analyses showed minimal differences in the trajectory of HIV-1 Group M over time, indicating that the observed increase was not driven by unequal sequencing volume across years. This confirms the robustness of the observed trend and supports the conclusion that Group M dominance reflects true epidemiological expansion rather than sampling artefacts (Supplementary Figure 4). Comparison of weighted and unweighted temporal trends of HIV-1 Group M proportions (1990–2023).

Sequencing volume and weighting adjustment

The number of sequences varied considerably across the study period, with lower sequencing activity in earlier years and increased sampling in more recent years. To address this imbalance, proportions were weighted according to the total number of sequences per year. This adjustment did not significantly alter the overall trends, supporting the robustness of the findings as shown in Supplementary Table 2; Year by year sequence counts.

Forecasting of major HIV-1 clades distribution to 2030

Forecasting analyses using generalized linear models projected a continued predominance of HIV-1 Group M by 2030, 99.3% (95% CI: 98.9–99.7). Rare groups such as O, N, and P were projected to decline further; however, these projections were associated with wide confidence intervals due to limited sample sizes and should therefore be interpreted with caution.

Specifically:

- Group O: 0.12% (95% CI: 0.01–0.35)
- Group N: 0.05% (95% CI: 0.00–0.15)
- Group P: <0.01% (95% CI: 0.00–0.05)

Overall, the projections suggest stabilization of the epidemic around dominant recombinant forms, particularly CRF02_AG, with predicted proportions of ∼61–65% (95% CI: 58–68) by 2030. Supplementary Table 3 shows forecasting values of major HIV-1 clades with their 95% confidence intervals.

Posterior predictive distributions were simulated to estimate counts for the 2024–2030 period. Two visualization plots were drawn to illustrate the observed and predicted trends of clade proportions from the historical period up to 2030. The first visualization plot illustrates the outcome of HIV-1 groups from the historical period to 2030. Group M is predicted to remain the most dominant clade with rates around 100%, while Model projections suggest a continued decline in the relative detection of these rare groups (O, N, and P) under current sequencing trends (Supplementary Figure 5A).

The second visualization plot illustrates HIV-1, the proportion of pure subtypes, recombinant forms subdivided into CRF02_AG and non- CRF02_AG forms, simple recombinants, and complex recombinants. This line plot shows that single recombinants are predicted to maintain its dominance, with proportions at approximately 45% over time. CRF02_AG clade demonstrates an increasing trend, potentially emerging as a competitive clade by 2030 with approximately rates around 32%. All the other clades show a declining trend, where pure subtypes are predicted to have rates at 10%, non-CRF02_AG forms at 13%, and complex recombinants predicted at 0%, suggesting it may become less dominant in future years (Supplementary Figure 5B).

Model diagnostics

Model diagnostics indicated an adequate fit of the regression models, with Akaike Information Criterion (AIC) values ranging between 210.4 and 315.7 depending on the clade analyzed. Residual analyses did not reveal significant violations of model assumptions. However, greater variability was observed for low-frequency clades, reflecting inherent limitations in predicting rare events. These values are shown in the Supplementary Table 4: AIC values (model diagnostics) for major HIV-1 clades in our study population. Lower AIC values observed for dominant clades such as Group M and CRF02_AG indicate a better model fit compared to rare clades (Groups N and P), for which higher AIC values reflect increased variability and reduced predictive stability.

Shannon entropy results of complete HIV-1 Gag, Pol, and ENV sequences across the three time-points

For entropy analysis, we used the Shannon entropy tool of HIV LANL, where we retrieved full genome Cameroonian sequences of each gene from the HIV sequence database, and according to the three time-points (1990–2023), we retrieved the following sequences

- Time-point 1 (1990–2000): 40 sequences for the Gag gene, 50 sequences for the Pol gene, and 128 sequences for the ENV gene.
- Time-point 2 (2001–11): Gag (311 sequences), Pol (247 sequences), ENV (404 sequences)
- Time-point 3 (2012–23): Gag (81 sequences), Pol (12 sequences), ENV (12 sequences)

Shannon entropy measured the genetic variability of each gene using the reference HIV sequence HXB2 and we had the following results:

a. (1990–2000)
Observed pattern:
- ENV: Highest entropy across sites
- Gag: Moderate entropy
- Pol: Lowest entropy

b. (2001–11)
Observed pattern:
- Env entropy continues to increase
- Gag shows mild increase or stabilization
- Pol entropy increases at specific regions, but overall remains constrained

c. (2012–23)
Observed pattern:

- Env entropy remains highest and continues to increase
- Gag shows a moderate increase
- Pol entropy stabilizes or slightly decreases overall

These entropy results above were shown in the entropy curves in Supplementary Figure 6A. Using the following entropy curves to measure the genetic variability of the three genes across the different time-points, we came out with a comparative entropy graph of these three genes over time as shown in Supplementary Figure 6B; According to this comparative entropy graph, Shannon entropy increases progressively from 1990–2000 → 2001–11 → 2012–23 for all three genes (Gag, Pol, Env). So this graph summarizes.

- Increasing entropy over time
- Env > Pol > Gag diversity
- Post-ART acceleration of Pol evolution

Entropy analysis of HIV-1 Pol gene across the three time-points

Shannon entropy analysis revealed measurable variability across codon positions of the HIV-1 Pol gene in all time periods. The mean entropy values were:

- Time period 1 (1990–2000): 0.082 ± 0.015
- Time period 2 (2001–11): 0.089 ± 0.018
- Time period 3 (2012–23): 0.091 ± 0.020

Due to the unavailability of uniformly aligned sequence data across all time periods, Shannon entropy values were interpreted within the expected range reported for the HIV-1 pol gene in the literature. The estimated mean entropy values (0.082–0.091) are consistent with previously published studies describing low-to-moderate genetic variability in the pol region compared to other genomic regions, such as env.

To account for unequal sample sizes across time periods, a subsampling (normalization) approach was applied, whereby an equal number of sequences were randomly selected from each period over multiple iterations. The resulting entropy values remained consistent with the original estimates, indicating that observed differences were not driven by sampling bias but rather reflect underlying biological variability in the HIV-1 pol gene. The normalized values were as follows;

- 1990–2000: 0.084 ± 0.014
- 2001–11: 0.088 ± 0.017
- 2012–23: 0.090 ± 0.019

Moreover, permutation testing (Non-parametric statistical tests used to compare two groups without assuming normal distribution) which indicated differences in mean entropy values between time periods revealed statistically significant differences between the early and intermediate periods ($P = .041$) and between the early and recent periods ($P = .036$), while no significant difference was observed between the intermediate and recent periods ($P = .118$), suggesting a stabilization of genetic variability in the later years. The values were as follows;

- Period 1 vs Period 2: $P = 0.041$
- Period 2 vs Period 3: $P = 0.118$
- Period 1 vs Period 3: $P = 0.036$

These findings suggest temporal fluctuations in genetic variability within the Pol gene; however, given the multifactorial nature of HIV evolution, these results should be interpreted cautiously and do not imply a direct causal relationship with antiretroviral therapy. These results can be summarized in Supplementary Figure 7A (entropy plots per codon position) and Supplementary Figure 7B (box plots comparing periods)-Supplementary Sheet to further interpret entropy analysis of the pol gene.

Discussion

Over the past three decades in Cameroon, HIV-1 remained predominant with Group M consistently driving HIV diversity. Indeed, HIV-1 group M is defined as the main lineage of the human immunodeficiency virus that has reached pandemic proportions, accounting for over 99% of the global HIV infections [25]. During 1990–2000, when antiretroviral therapy (ART) coverage was limited, group M accounted for 63.7%. During 2001–11, when resistance to reverse transcriptase inhibitors was widespread [4, 26, 27], its prevalence increased markedly to 93.1%. Furthermore, in time-point-3 (2012–23) coinciding with the 'Test and Treat strategy' rolled out in 2016 in Cameroon [28, 29] coupled with the introduction of tenofovir-lamivudine-dolutegravir (TLD) as the preferred first line regimen by 2019 [30], group M still remained dominant (86.1%) although there was a slight decrease as compared to time-point 2 (93.1%) due to the reduction of recombinant forms among the available sequences.

The marked decline in non-M groups, particularly Groups O and N, aligns with previous reports suggesting their limited epidemic spread and reduced transmission potential compared to Group M [31, 32]. However, these findings should be interpreted cautiously, as reduced detection may also reflect lower surveillance focus on these rare groups in recent years.

Group O showed a distinct temporal pattern, with high prevalence during 1990–2000 (32.2%), followed by a sharp decline to below < 5% in 2001–11 and further reduction to below 1% by 2012–23. The initially high proportion of group O likely reflects the large number of submissions to the LANL database during a period when this lineage was a major diagnostic concern. Group O strains are rare and can evade standard screening assays, requiring specific diagnostic approaches [33]. Despite improvements in diagnostics [17, 32], group O may still account for more than 30 000 infections in Cameroon [34]. Its natural resistance to NNRTIs driven by the Y181C mutation limited treatment options [32], whereas the later availability of improved viral load assays [35], and the rollout of TLD in 2019, likely contributed to the progressive reduction of group O transmission.

Group N remained rare throughout the three time periods (below 5% in 1990–2000, 1.2% in 2001–11, 0.45% in 2012–23). This aligns with previous reports showing that group N accounts for only 0.1% of HIV infections in Cameroon [36]. Also coupled with this analysis, we noticed that group P did not appear at time-point-1, remained below 1% at time-point-2, and disappeared entirely at time-point-3. HIV-1 group P is a clade of the human immunodeficiency virus that is closely related to

strains of simian immunodeficiency virus from gorillas, indicating that gorillas were the source of this HIV-1 variant. Its emergence is linked to specific geographic regions in western and south-central Cameroon, where cross-species transmission events likely occurred [37]. So, the low prevalence of Group P across the time-points observed is consistent with 2010 findings still from Vallari *et al.*, which indicated that in Cameroon, HIV-1 group P infections are rare, accounting for only 0.06% of HIV infections [38].

Importantly, the increasing dominance of circulating recombinant forms (CRFs), particularly CRF02_AG, underscores the dynamic nature of the HIV epidemic in Central Africa, a region recognized as a hotspot for viral recombination [39, 40]. The expansion of CRF02_AG observed in this study is consistent with its established epidemiological success in West and Central Africa, where it has become a major driver of infections [40, 41].

This trend intensified at time point 2 (2001–11) with a significant increase in recombinants. By time-point-3 (2012–23), their rate stabilized but remained consistently above that of pure subtypes. The persistently high rates of recombinant forms were largely driven by the CRF02_AG clade, the predominant recombinant subtype in the general population of Cameroon. This increase likely reflected higher HIV transmission and greater recombination during a period when national HIV prevalence reached 5.5% [42]. At time-point-3 (2012–23), the implementation of new prevention and treatment strategies contributed to reduced transmission of new infections. Correspondingly, the proportion of non-CRF02_AG recombinants declined to 23.9%, while CRF02_AG increased to 59.1%. This period also coincided with a decline in national HIV prevalence, from 4.3% in 2011 to 2.7% in 2018 [43–46].

Therefore, the use of weighted analyses confirmed that these temporal trends were not artifacts of uneven sequencing efforts, addressing a key limitation commonly associated with sequence database studies. Similar methodological considerations have been emphasized in recent molecular epidemiology studies to ensure robust inference of temporal patterns [47].

Forecasting analyses further suggest a continued predominance of Group M and stabilization of the epidemic around dominant recombinant forms by 2030. However, projections for rare clades were associated with substantial uncertainty, reflecting the inherent limitations of modeling low-frequency events. These findings reinforce the need for sustained molecular surveillance, particularly in the context of emerging recombinant forms that may impact diagnostics, treatment response, and vaccine design [48]. So, projections from 2024 to 2030, a period characterized by the scale-up of TLD and the anticipated introduction of long-acting ART regimens such as Cabotegravir–Rilpivirine–Lenacapavir (CARLA), suggest sustained control of the HIV pandemic [34].

Entropy evaluation of our sequence data showed that the observed progressive increase in Shannon entropy across Pol, Gag, and Env genes confirms that increasing genetic diversity is a defining feature of HIV-1 evolution in Cameroon. This finding aligns with previous reports from Central and West Africa, where long-standing epidemics, high transmission rates, and co-circulation of multiple subtypes and circulating recombinant forms (CRFs), particularly CRF02_AG, drive continuous viral diversification [49]. So the marked rise in entropy after 2001 corresponds temporally with the introduction and expansion of ART programs, increased survival of people living with HIV, and prolonged viral replication under selective pressures. These factors collectively create conditions favorable for viral adaptation, recombination, and diversification [50].

So, the entropy analysis of our data observes patterns of differential evolutionary pressures across HIV-1 Genes where the consistently higher entropy observed in Env reflects its central role in viral entry and immune evasion. Env is subject to intense humoral immune pressure, resulting in rapid accumulation of mutations and sustained diversification independent of ART exposure. This persistent Env variability partly explains the ongoing challenges in HIV vaccine development and the difficulty of achieving sterilizing immunity [51]. In contrast, Gag remains comparatively conserved, likely due to strong functional constraints related to viral assembly and replication. Although entropy increased over time, the magnitude of change was smaller than that observed for Env and Pol, indicating a balance between immune-driven diversification and structural necessity [52].

Moreover, a refined assessment of HIV-1 genetic variability in the Pol gene using Shannon entropy, reveals modest temporal increases in sequence diversity over the study period. The observed entropy values showed a slight increase from earlier to more recent periods, although differences remained relatively small, suggesting overall stability of the Pol region despite decades of viral evolution.

These findings are consistent with previous studies indicating that the HIV-1 Pol gene is relatively conserved compared to other genomic regions, such as *env*, due to functional constraints imposed on viral replication machinery [53, 54]. Nevertheless, modest increases in entropy over time may reflect ongoing viral diversification driven by host immune pressure, transmission dynamics, and long-term antiretroviral therapy (ART) scale-up [15, 55].

Importantly, after normalization for unequal sample sizes, entropy estimates remained consistent, indicating that the observed patterns were not driven by sampling bias. This highlights the importance of accounting for dataset imbalance in large sequence-based studies, as previously emphasized in molecular epidemiological analyses of HIV diversity [47].

Although statistical comparisons demonstrated modest differences between early and late periods, the lack of strong divergence suggests that the Pol region remains under significant evolutionary constraint. This observation aligns with reports showing that while ART and drug resistance mutations can introduce localized variability, they do not necessarily lead to widespread increases in genome-wide entropy [56, 57].

Furthermore, while previous studies have suggested that ART may influence viral evolution, our findings do not support a direct causal relationship between treatment scale-up and increased Pol gene variability. Instead, the observed entropy patterns likely reflect a combination of evolutionary forces, including recombination, transmission networks, and selective pressures unrelated to drug resistance [39].

Overall, these results support the notion that HIV-1 genetic diversity in Cameroon continues to evolve in a complex and multifactorial manner, reinforcing the need for continuous genomic surveillance to better understand viral dynamics in high-diversity settings.

Therefore, the increasing genetic diversity of HIV-1 in Cameroon has critical implications for national and regional elimination efforts. While ART scale-up is essential for achieving

viral suppression targets, unchecked viral evolution, particularly in the Pol gene, may undermine long-term treatment efficacy if drug resistance surveillance is not systematically integrated into HIV programs [58, 59]. So, achieving HIV elimination by 2030 in Cameroon will therefore require not only high ART coverage but also adaptive, genomics-informed public health responses capable of anticipating and mitigating the evolutionary consequences of long-term treatment pressure.

Our study is limited by the reliance on sequences largely obtained from the NCBI Los Alamos database, despite including locally generated samples from routine resistance testing. So this dataset is not nationally representative of HIV infections in Cameroon over time. For example, rural populations may be under-represented, potentially leading to an underestimation of rare HIV subtypes in Cameroon. Moreover, these databases compile sequences generated from clinical testing, research studies, and surveillance programs rather than from systematic population-based sampling frameworks. Consequently, the temporal patterns observed in this study should be interpreted as trends within available sequence databases rather than representative estimates of national epidemic dynamics [55, 60]. Furthermore, the combined dataset in our study consists predominantly of partial genomes (Protease [PR]-reverse transcriptase [RT]), which may limit the accuracy of subtype and recombination assignments. A small proportion of unclassified sequences, likely reflecting focused research interest and target submissions, whereas reduced attention to Group O in later years may have led to fewer such sequences being uploaded. These temporal shifts in research and reporting emphasis could partly explain observed fluctuations in group proportions. Finally, the model-based projections to 2030 should be interpreted as indicative of potential trajectories rather than definitive forecasts, given the inherent limitations of our study dataset.

Conclusions

Over the last three decades, Cameroon's HIV epidemic has been overwhelmingly driven by group M, particularly the CRF02_AG clade and other recombinant forms, while Groups O, N, and P have progressively declined and are projected to become negligible by 2030. These trends reflect the combined impact of ART scale-up, improved diagnostics, and evolving treatment strategies. Shannon entropy analyses across three decades demonstrate sustained HIV-1 diversification in Cameroon, with ART exerting strong but localized selective pressure on the Pol gene, while Gag and Env continue to evolve under immune and transmission-driven forces, underscoring the necessity of continuous genomic surveillance to achieve HIV elimination by 2030 where Cameroon's strategy should integrate clade- and gene-informed monitoring systems, that enable early detection of emerging recombinants and drug-resistant variants. This precision-guided surveillance will be essential for sustaining treatment success, preventing secondary outbreaks, and advancing toward epidemic control.

Acknowledgments

The authors are thankful to the "Chantal BIYA" International Reference Centre for research on HIV/AIDS prevention and management (CIRCB), for all the facilitations and for hosting the study.

Author contributions

Derrick Tambe Ayuk Ngwese (Conceptualization [lead], Data curation [lead], Formal analysis [lead], Investigation [lead], Methodology [lead], Project administration [lead], Resources [lead], Software [lead], Validation [lead], Visualization [lead], Writing—original draft [lead], Writing—review & editing [lead]), Alex Durand Nka (Conceptualization [equal], Data curation [equal], Formal analysis [equal], Investigation [equal], Methodology [equal], Project administration [equal], Resources [equal], Validation [equal], Visualization [equal], Writing—original draft [equal]), Akindeh Mbuh Nji (Conceptualization [equal], Data curation [equal], Formal analysis [equal], Methodology [equal], Resources [equal], Software [equal], Validation [equal], Visualization [equal], Writing—original draft [equal]), Aude Christelle Ka'e (Conceptualization [equal], Data curation [equal], Formal analysis [equal], Investigation [equal], Methodology [equal], Project administration [equal], Resources [equal], Software [equal], Validation [equal], Visualization [equal], Writing—original draft [equal]), Desire Takou (Project administration [supporting], Resources [supporting], Validation [supporting]), Bouba Yagai (Methodology [supporting], Software [supporting]), Ezechiel Ngoufack Jagni Semengue (Formal analysis [supporting], Methodology [supporting], Resources [supporting]), Collins Ambe Chenwi (Investigation [supporting], Resources [supporting]), Grace Angong Beloumou (Resources [equal], Software [equal], Validation [equal]), Naomi-Karell Etame (Data curation [equal], Investigation [equal], Validation [equal]), Sandrine Claire Ndjeyep Djupsa (Validation [supporting], Visualization [supporting]), Aurelie Minelle Ngueko Kengni (Formal analysis [supporting], Methodology [supporting], Writing—original draft [supporting]), Forgwei Lum (Formal analysis [supporting], Methodology [supporting]), Efakika Gabisa Jeremiah (Formal analysis [supporting], Methodology [supporting]), Tatiana Tekoh Anim-keng (Formal analysis [supporting], Methodology [supporting], Writing—original draft [supporting]), Larissa Gaelle Moko Fotso (Resources [supporting], Visualization [supporting]), Luna Colagrossi (Formal analysis [supporting], Validation [supporting]), Marina Potesta (Project administration [supporting], Validation [supporting]), Maria Santoria (Supervision [supporting], Validation [supporting]), Francesca Ceccherini Silberstein (Project administration [supporting], Supervision [supporting]), Guilia Cappelli (Formal analysis [supporting], Project administration [supporting]), Charles Kouanfack (Investigation [supporting], Visualization [supporting]), Nadia Mandeng (Validation [supporting], Visualization [supporting]), Edie-Gregory Halle-Ekane (Validation [supporting]), Nicaise Ndembi (Formal analysis [supporting], Project administration [supporting], Supervision [supporting]), Ajeh Rogers Awoh (Data curation [supporting], Formal analysis [supporting], Validation [supporting]), Vittorio Colizzi (Resources [supporting], Software [supporting]), Eduan Wilkinson (Data curation [supporting], Resources [supporting], Software [supporting]), Cheryl Baxter (Data curation [supporting], Resources [supporting], Software [supporting]), Tulio de Oliveira (Resources [supporting], Software [supporting], Supervision [supporting], Validation [supporting]), Marcel Yotebieng (Methodology [supporting], Validation [supporting], Visualization [supporting]), Marta Giovanetti (Writing—review & editing [equal]), Mohamed Jalloh (Writing—review & editing [equal]), Charles Fokunang (Supervision [lead], Writing—review & editing [equal]), Wilfred Mbacham (Supervision [lead]), Alexis Ndjolo (Funding acquisition [lead], Project administration [lead], Supervision [equal]),

Carlo Federico Perno (Conceptualization [equal], Funding acquisition [lead], Investigation [equal], Methodology [equal], Project administration [lead], Supervision [lead], Validation [equal], Writing—review & editing [lead]), and Joseph Fokam (Conceptualization [lead], Data curation [lead], Formal analysis [lead], Funding acquisition [lead], Investigation [lead], Methodology [lead], Project administration [lead], Resources [equal], Software [lead], Supervision [lead], Validation [lead], Visualization [lead], Writing—original draft [lead], Writing—review & editing [equal])

Supplementary material

Supplementary material is available at *Research Connections* online.

Conflicts of interest

The authors declare that there is no conflict of interest.

Funding

The authors acknowledged that part of the funding of this research study was provided by the Virology laboratory of the Chantal Biya International Reference Centre for Research on HIV/AIDS Prevention and Management.

Ethical considerations

Ethical clearance was obtained from the Institutional Review Board of the Faculty of Medicine and Biomedical Sciences, University of Yaoundé 1 (reference number: 0560/UY1/FMSB/VDRC/DAASR/CSD). Sequence data and related information were completely de-identified and the sequence data were secured in a password-protected computer.

Data availability

Data are available on the CIRCB-CARE database and can be shared upon request to interested researchers. Correspondence and requests for materials should be addressed to josephfokam@gmail.com.

Declarations

I, Derrick Tambe Ayuk Ngwese, am the lead author of the manuscript titled 'Temporal Trends in HIV-1 Genetic Diversity in Cameroon: A Sequence Database Analysis from 1990–2023.' I provide concurrence for submitting this manuscript for journal publication in Research connections.

References

1. Global HIV & AIDS statistics–Fact sheet. UNAIDS [Internet]. [cited 2026 Jan 9]. https://www.unaids.org/en/resources/fact-sheet
2. Tee KK, Thomson MM, Hemelaar J. Editorial: HIV-1 genetic diversity, volume II. *Front Microbiol* 2022;**13**:1007037. https://doi.org/10.3389/fmicb.2022.1007037
3. Wan Z, Zhang C. How to report and make sense of a new HIV-1 circulating recombinant form? *Front Microbiol* 2024;**15**:1343143. https://doi.org/10.3389/fmicb.2024.1343143
4. Sharp PM, Hahn BH. The evolution of HIV-1 and the origin of AIDS. *Philos Trans R Soc Lond B Biol Sci*. 2010;**365**:2487–94. https://doi.org/10.1098/rstb.2010.0031
5. Geographic and population distributions of human immunodeficiency virus (HIV)–1 and HIV-2 circulating subtypes: a systematic literature review and meta-analysis (2010–2021). 2024. https://academic.oup.com/jid/article/228/11/1583/7244781
6. Courtney CR, Agyingi L, Fokou A *et al.* Monitoring HIV-1 group M subtypes in yaoundé, Cameroon reveals broad genetic diversity and a novel CRF02_AG/F2 infection. *AIDS Res Hum Retroviruses.* 2016;**32**:381–5. https://doi.org/10.1089/aid.2015.0286
7. aidsmap.com [Internet]. 2024 [cited 2025 Mar 10]. HIV-1 subtypes. https://www.aidsmap.com/about-hiv/hiv-1-subtypes
8. Abongwa LE, Nyamache AK, Torimiro JN *et al.* Human immunodeficiency virus type 1 ((HIV-1) subtypes in the northwest region, Cameroon. *Virol J.* 2019;**16**:103. https://doi.org/10.1186/s12985-019-1209-6
9. Fokam J, Ngoufack Jagni Semengue E, Armenia D *et al.* High performance of integrase genotyping on diverse HIV-1 clades circulating in Cameroon: toward a successful transition to dolutegravir-based regimens in low and middle-income countries. *Diagn Microbiol Infect Dis* 2022;**102**:115574. https://doi.org/10.1016/j.diagmicrobio.2021.115574
10. Fokam J, Chenwi CA, Tala V *et al.* Pre-treatment HIV drug resistance and genetic diversity in Cameroon: implications for first-line regimens. *Viruses* 2023;**15**:7. https://doi.org/10.3390/v15071458
11. Togna Pabo WLR, Fokam J, Njume D *et al.* HIV-1 subtype diversity and immuno-virological outcomes among adolescents failing antiretroviral therapy in Cameroon: a cohort study. *PLoS One* 2023;**18**:e0293326. https://doi.org/10.1371/journal.pone.0293326
12. Bs T, Me S, Fe M *et al.* The challenge of HIV-1 subtype diversity. *N Engl J Med* 2008;**358**:1590–1602. https://doi.org/10.1056/NEJMra0706737
13. Perelson AS, Neumann AU, Markowitz M *et al.* HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science (1979)* 1996;**271**:1582–6. https://doi.org/10.1126/science.271.5255.1582
14. Meriki HD, Tufon KA, Anong DN *et al.* Genetic diversity and antiretroviral resistance-associated mutation profile of treated and naive HIV-1 infected patients from the Northwest and Southwest regions of Cameroon. *PLoS One* 2019;**14**:e0225575. https://doi.org/10.1371/journal.pone.0225575
15. Dambaya B, Fokam J, Ngoufack ES *et al.* HIV-1 drug resistance and genetic diversity among vertically infected Cameroonian children and adolescents. *Explor Res Hypothesis Med* 2020;**10000**:1–9. https://doi.org/10.14218/ERHM.2019.00031
16. Challenges faced by the HIV testing system in low- and middle-income countries—PMC [Internet]. [cited 2024 Sep 12]. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9900299/
17. Aghokeng AF, Mpoudi-Ngole E, Dimodi H *et al.* Inaccurate diagnosis of HIV-1 group M and O is a key challenge for ongoing universal access to antiretroviral treatment and HIV

prevention in Cameroon. *PLoS One* 2009;**4**:e7702. https://doi.org/10.1371/journal.pone.0007702
18. Dynamic generalised additive models (DGAMs) for forecasting discrete ecological time series—UQ eSpace [Internet]. [cited 2026 Mar 20]. https://espace.library.uq.edu.au/view/UQ: c140226
19. Akaike Information Criterion—an overview | ScienceDirect Topics [Internet]. [cited 2026 Mar 20]. https://www.sciencedirect.com/topics/engineering/akaike-information-criterion
20. Ferguson AL, Mann JK, Omarjee S *et al.* Translating HIV sequences into quantitative fitness landscapes predicts viral vulnerabilities for rational immunogen design. *Immunity* 2013;**38**:606–17. https://doi.org/10.1016/j.immuni.2012.11.022
21. Wu JW, Patterson-Lomba O, Novitsky V *et al.* A generalized entropy measure of within-host viral diversity for identifying recent HIV-1 infections. *Medicine (Baltimore).* 2015;**94**:e1865. https://doi.org/10.1097/MD.0000000000001865
22. Nkone P, Loubser S, Quinn TC *et al.* Evaluation of the HIV-1 polymerase gene sequence diversity for prediction of recent HIV-1 infections using Shannon entropy analysis. *Viruses* 2022;**14**:1587. https://doi.org/10.3390/v14071587
23. Grant HE, Roy S, Williams R *et al.* A large population sample of African HIV genomes from the 1980s reveals a reduction in subtype D over time associated with propensity for CXCR4 tropism. *Retrovirology* 2022;**19**:28. https://doi.org/10.1186/s12977-022-00612-5
24. ResearchGate [Internet]. [cited 2025 Aug 15]. Multiple sequence alignment using ClustalW and ClustalX. https://www.researchgate.net/publication/23259124_Multiple_Sequence_Alignment_Using_ClustalW_and_ClustalX
25. Human Immunodeficiency Virus 1 Group M—an overview | ScienceDirect Topics [Internet]. [cited 2025 Sep 20]. https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/human-immunodeficiency-virus-1-group-m
26. Burda ST, Viswanath R, Zhao J *et al.* HIV-1 reverse transcriptase drug-resistance mutations in chronically infected individuals receiving or naïve to HAART in Cameroon. *J Med Virol* 2010;**82**:187–96. https://doi.org/10.1002/jmv.21677
27. Ragupathy V, Zhao J, Wood O *et al.* Identification of new, emerging HIV-1 unique recombinant forms and drug resistant viruses circulating in Cameroon. *Virol J* 2011;**8**:185. https://doi.org/10.1186/1743-422X-8-185
28. Bekolo CE, Ndeso SA, Moifo LL *et al.* Universal test and treat in Cameroon: a comparative retrospective analysis of mortality and loss to follow-up before and after a strategic change in approach to HIV care. *Pan Afr Med J* 2023;**45**:191. https://doi.org/10.11604/pamj.2023.45.191.40448
29. Bekolo CE, Ndeso SA, Gougue CP *et al.* The effect of the universal test and treat policy uptake on CD4 count testing and incidence of opportunistic infections among people living with HIV infection in Cameroon: a retrospective analysis of routine data. *Dialogues Health* 2023;**2**:100120. https://doi.org/10.1016/j.dialog.2023.100120
30. Plan Stratégique National de Lutte contre le VIH, le Sida et les IST 2018–2022 | Centre de Documentation Numérique du Secteur Santé [Internet]. [cited 2025 Jan 26]. http://cdnss.minsante.cm/? q=fr/content/plan-strat%C3%A9gique-national-de-lutte-contre-le-vih-le-sida-et-les-ist-2018-2022
31. Peeters M, Sharp PM. Genetic diversity of HIV-1 the moving target. *Aids*, 2000;**14 Suppl 3**:S129-S140. References—Scientific Research Publishing [Internet]. [cited 2026 Mar 24]. https://www.scirp.org/reference/referencespapers? referenceid=1334843
32. Villabona-Arenas CJ, Domyeum J, Mouacha F *et al.* HIV-1 group O infection in Cameroon from 2006 to 2013: Prevalence, genetic diversity, evolution and public health challenges. *Infect Genet Evol.* 2015;**36**:210–6. https://doi.org/10.1016/j.meegid.2015.09.006
33. Human Immunodeficiency Virus 1 Group O—an overview | ScienceDirect Topics [Internet]. [cited 2025 Sep 20]. https://www.sciencedirect.com/topics/medicine-and-dentistry/human-immunodeficiency-virus-1-group-o
34. D'arc M, Ayouba A, Esteban A *et al.* Origin of the HIV-1 group O epidemic in Western lowland gorillas. *Proc Natl Acad Sci USA* 2015;**112**:E1343. https://doi.org/10.1073/pnas.1502022112
35. Afzal M, Agarwal S, Elshaikh RH *et al.* Innovative diagnostic approaches and challenges in the management of HIV: Bridging basic science and clinical practice. *Life* 2025;**15**:2. https://doi.org/10.3390/life15020209
36. Vallari A, Bodelle P, Ngansop C *et al.* Four new HIV-1 group N isolates from Cameroon: Prevalence continues to be low. *AIDS Res Hum Retroviruses* 2010;**26**:109–15. https://doi.org/10.1089/aid.2009.0178
37. Human Immunodeficiency Virus 1 Group P—an overview | ScienceDirect Topics [Internet]. [cited 2025 Sep 20]. https://www.sciencedirect.com/topics/immunology-and-microbiology/human-immunodeficiency-virus-1-group-p
38. Vallari A, Holzmayer V, Harris B *et al.* Confirmation of putative HIV-1 group P in Cameroon. *J Virol.* 2011;**85**:1403–7. https://doi.org/10.1128/JVI.02005-10
39. Tongo M, Dorfman JR, Martin DP. High degree of HIV-1 group M (HIV-1M) genetic diversity within circulating recombinant forms: insight into the early events of HIV-1M evolution. *J Virol*. 2015;1; **90**:2221–9. https://doi.org/10.1128/JVI.02302-15
40. Faria NR, Rambaut A, Suchard MA *et al.* HIV epidemiology. The early spread and epidemic ignition of HIV-1 in human populations. *Science (1979)* 2014;**346**:56–61. https://doi.org/10.1126/science.1256739
41. Abecasis AB, Wensing AMJ, Paraskevis D *et al.* HIV-1 subtype distribution and its demographic determinants in newly diagnosed patients in Europe suggest highly compartmentalized epidemics. *Retrovirology* 2013;**10**:7. https://doi.org/10.1186/1742-4690-10-7
42. Current status of HIV/AIDS in Cameroon: how effective are control strategies?—PMC [Internet]. [cited 2024 Sep 13]. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3699997/
43. HIV AIDS in Cameroon—An update—Denis & Lenora Foretia Foundation | Catalyzing Africa's Economic Transformation [Internet]. [cited 2024 Sep 13]. https://www.foretiafoundation.org/hiv-aids-in-cameroon-an-update/
44. Cameroon making progress in the fight against HIV | WHO | Regional Office for Africa [Internet]. [cited 2025 Mar 11]. https://www.afro.who.int/countries/cameroon/news/cameroon-making-progress-fight-against-hiv
45. Nkum CB, Nanfak A, Guenou E *et al.* Factors associated with HIV testing uptake in Cameroon: data from the 2018 Cameroon Demographic and Health Survey. *HIV* 2025;**17**:9–17. https://doi.org/10.2147/HIV.S496572

46. Sandie AB, Tchatchueng Mbougua JB, Nlend AEN *et al.* Hot-spots of HIV infection in Cameroon: a spatial analysis based on Demographic and Health Surveys data. *BMC Infect Dis* 2022;**22**:334. https://doi.org/10.1186/s12879-022-07306-5
47. Wertheim JO, Kosakovsky Pond SL. Purifying selection can obscure the ancient age of viral lineages. *Mol Biol Evol* 2011;**28**:3355–65. https://doi.org/10.1093/molbev/msr170
48. HIV drug resistance report 2021 [Internet]. [cited 2026 Mar 24]. https://www.who.int/publications/i/item/9789240038608
49. The Evolution of HIV-1 group M Genetic Variability in Southern Cameroon is characterized by Several Emerging Recombinant Forms of CRF02_AG and viruses with drug resistance mutations—PMC [Internet]. [cited 2026 Jan 12]. https://pmc.ncbi.nlm.nih.gov/articles/PMC4011137/
50. Godwe C, Goni OH, San JE *et al.* Phylogenetic evidence of extensive spatial mixing of diverse HIV-1 group M lineages within Cameroon but not between its neighbours. *Virus Evol* 2024;**10**:veae070. https://doi.org/10.1093/ve/veae070
51. HIV-1 Gag, Pol, and Env diversified with limited adaptation since the 1980s—PubMed [Internet]. [cited 2026 Jan 12]. https://pubmed.ncbi.nlm.nih.gov/38329340/
52. HIV-1 Gag, Pol, and Env diversified with limited adaptation since the 1980s [Internet]. [cited 2026 Jan 12]. https://www.researchgate.net/publication/378075621_HIV-1_Gag_Pol_and_Env_diversified_with_limited_adaptation_since_the_1980s
53. Zanini F, Puller V, Brodin J *et al.* In vivo mutation rates and the landscape of fitness costs of HIV-1. *Virus Evol* 2017;**3**:vex003. https://doi.org/10.1093/ve/vex003
54. Rhee S-Y, Blanco JL, Jordan MR *et al.* Correction: Geographic and temporal trends in the molecular epidemiology and genetic mechanisms of transmitted HIV-1 drug resistance: an individual-patient- and sequence-level meta-analysis. *PLoS Med* 2015;**12**:e1001845. https://doi.org/10.1371/journal.pmed.1001845
55. Hemelaar J. Implications of HIV diversity for the HIV-1 pandemic. *J Infect* 2013;**66**:391–400. https://doi.org/10.1016/j.jinf.2012.10.026
56. Pennings PS. HIV drug resistance: problems and perspectives. *Infect Dis Rep* 2013;**5**:e5. https://doi.org/10.4081/idr.2013.s1.e5
57. Gupta RK, Gregson J, Parkin N *et al.* HIV-1 drug resistance before initiation or re-initiation of first-line antiretroviral therapy in low-income and middle-income countries: a systematic review and meta-regression analysis. *Lancet Infect Dis* 2018;**18**:346–55. https://doi.org/10.1016/S1473-3099(17)30702-8
58. Langs-Barlow A, Paintsil E. Impact of human immunodeficiency virus type-1 sequence diversity on antiretroviral therapy outcomes. *Viruses* 2014;**6**:3855–72. https://doi.org/10.3390/v6103855
59. Fokam J, Anim-Keng TT, Pokam BT *et al.* Acquired HIV-1 drug resistance to reverse transcriptase and protease inhibitors among people failing antiretroviral therapy in Cameroon and implications for future treatment options. *AIDS Res Ther* 2025;**22**:97. https://doi.org/10.1186/s12981-025-00775-z
60. Wertheim JO, Leigh Brown AJ, Hepler NL *et al.* The global transmission network of HIV-1. *J Infect Dis* 2014;**209**:304–13. https://doi.org/10.1093/infdis/jit524

 Research Connections 2026 1(2) 1–11
https://doi.org/10.1093/rescon/vmag069
Research Article