Article

Comprehensive Analysis of HIV-1 Integrase Resistance-Related Mutations in African Countries

Francesco Branda [1,†], Marta Giovanetti [2,3,4,†], Leonardo Sernicola [5], Stefania Farcomeni [5], Massimo Ciccozzi [1] and Alessandra Borsetti [5,*]

1   Unit of Medical Statistics and Molecular Epidemiology, University Campus Bio-Medico of Rome, 00128 Rome, Italy; f.branda@unicampus.it (F.B.); m.ciccozzi@unicampus.it (M.C.)
2   Sciences and Technologies for Sustainable Development and One Health, Università Campus Bio-Medico di Roma, 00128 Rome, Italy; giovanetti.marta@gmail.com
3   Instituto Rene Rachou, Fundação Oswaldo Cruz, Belo Horizonte 30190-002, Brazil
4   Climate Amplified Diseases and Epidemics (CLIMADE), Brasilia 70070-130, Brazil
5   National HIV/AIDS Research Center (CNAIDS), Istituto Superiore di Sanità, 00161 Rome, Italy; leonardo.sernicola@iss.it (L.S.); stefania.farcomeni@iss.it (S.F.)
*   Correspondence: alessandra.borsetti@iss.it
†   These authors contributed equally to this work.

Abstract: The growing emergence of non-nucleoside reverse transcriptase inhibitor (NNRTI) HIV drug resistance in sub-Saharan Africa (SSA) led to the World Health Organization (WHO) recommending, in 2018, a transition to dolutegravir (DTG) as a first-line antiretroviral therapy (ART) in SSA. The broad HIV-1 genetic diversity in SSA could shape DTG effectiveness and the pattern of drug resistance mutations (DRMs) in this region. This study evaluated HIV-1 integrase (IN) DRMs and conserved regions among published groups M, N, O, and P HIV-1 sequences spanning forty years of the HIV epidemic during the transition of DTG-based ART. Overall, we found low levels of integrase strand transfer inhibitor (INSTI)-DRMs (<1%) across HIV groups between the years 1983 and 2023; however, it was unexpected to detect DRMs at statistically significantly higher frequencies in pre-INSTI (1983–2007) than in the INSTI (2008–2023) era. The variability of accessory INSTI-DRMs depended on the HIV subtypes, with implications for susceptibility to DTG. Our findings provide new perspectives on the molecular epidemiology and drug resistance profiles of INSTIs in SSA, emphasizing the need for ongoing surveillance and customized treatment approaches to address the continent's varied HIV subtypes and changing resistance patterns.

Keywords: HIV-1; integrase; drug resistance; viral evolution; African countries

1. Introduction

At the close of 2023, approximately 39 million individuals globally were living with HIV (PLWH) [WHO], and Africa, with 25.6 million infected people, accounting for the highest HIV-1 genetic diversity globally, exhibits different distributions of HIV-1 subtypes and recombinant forms [1,2]. This includes subtype B in Northern Africa, subtype C in the South, CFR02_AG in the West and Central, and subtype C and D in the East [1–4]. The extensive genetic diversity of HIV, especially in West and Central Africa, potentially contributes to the emergence of natural polymorphisms associated with resistance and varying susceptibilities to combined antiretroviral therapy (cART) [5]. While cART has significantly reduced transmission, morbidity, and mortality among People Living with HIV (PLWH), the rise of drug resistance mutations (DRMs) in the virus increases the risk of virological failure and challenges the effectiveness of cART [5].

In many countries of Sub-Saharan Africa (SSA), high levels of pretreatment drug resistance to nucleotide reverse transcriptase inhibitors (NRTIs), non-nucleoside reverse transcriptase inhibitors (NNRTIs), and protease inhibitors (PIs) have led, in 2018, the

World Health Organization (WHO) to recommend a transition to dolutegravir (DTG)—an integrase strand transfer inhibitor (INSTI) with a strong resistance barrier—as a first-line ART component in resource-limited settings [6]. INSTIs, targeting the integrase (IN) enzyme crucial for HIV replication, have shown high efficacy in treatment-naïve and -experienced individuals [7]. The first INSTI approved by the United States Food and Drug Administration (US FDA) for clinical use was raltegravir (RAL) in 2007, followed by elvitegravir (EVG) in 2012, dolutegravir (DTG) in 2013, bictegravir (BIC) in 2018, and cabotegravir (CAB) in 2021 [6–8].

In African countries with prevalent NRTI resistance, a transition to (DTG)-based ART is underway. Optimized use of DTG is vital for cART success in settings lacking genotyping testing or drug resistance surveillance to prevent the emergence of resistant variants. The Stanford HIV Drug Resistance Database (HIVdb) (http://hivdb.stanford.edu, accessed on 25 November 2023). identifies key IN mutations across HIV-1 subtypes at positions T66, E92, G118, E138, G140, Y143, S147, Q148, N155 and R263 [9–12]. Additionally, mutations that reduce susceptibility in combination with major DRMs include polymorphic M50, L74, T97 and nonpolymorphic H51, Q95, F121, Q146, G149, V151, S153, G163, S230, and D232 accessory mutations. The pattern of DRMs varies among HIV-1 subtypes, like the G118R mutation in non-B subtypes and R263K in subtype B [9–12]. Given the significant INSTI susceptibility differences among HIV subtypes and the genetic diversity in SSA, it is crucial to monitor and investigate the pattern of INSTI-DRMs through a surveillance program. This study aims to provide updated insights into the molecular epidemiology and drug resistance mutation profiles of INSTIs in published IN sequences from 1983 to 2023 of the four phylogenetic M, N, O, and P HIV-1 groups in African countries.

2. Materials and Methods

2.1. Spatio-Temporal HIV-1 Subtype Epidemic Trend

We retrieved 7906 HIV-1 IN sequences, covering all 864 base pairs, collected from both ART-naïve and ART-treated individuals between 1983 and 2022 from a public database maintained by Los Alamos National Laboratory https://www.hiv.lanl.gov/content/index (accessed on 1 October 2023). HIV-1 group M included sequences belonging to subtypes A, B, C, D, F, G, H, J, K, L, and the circulating recombinant forms (CRFs) 01 to 143. A set of 96 sequences referring to HIV-1 groups N, O, P, and untyped "U" were assembled separately from the M group. The temporal distribution of HIV-1 IN subtypes across the SSA was meticulously analyzed through a comprehensive study of these 7906 HIV-1 samples. The selection criteria for samples were stringent, prioritizing those with detailed epidemiological metadata, including the date of sample collection and the country of origin. These HIV-1 strains represent a diverse set of geographical origins, categorized into four major African macro-regions: Western, Eastern, Central, and Southern. This regional classification followed the framework provided by UN-AIDS. The study further explored the nuances of subtype distribution within these regions, revealing patterns and trends indicative of both the biological characteristics of the virus and the socio-epidemiological factors at play.

2.2. Analysis of Polymorphisms and Integrase DRMs

All sequences were submitted to the Stanford HIV Drug Resistance Database version 7.0 for quality analysis. Sequence quality control was ensured by excluding any sequence having two or more of the following features: internal stop codons, frame shifts, more than 1% ambiguous nucleotide, hypermutations, insertions, or deletions. To assess the DRM$_s$, we used the Calibrated Population Resistance (CPR) tool version 8.1, available at Stanford University HIV Drug Resistance Database (HIVdb) (https://hivdb.stanford.edu/, (accessed on 1 October 2023)) and Stanford HIVDB 8.9-1 "https://hivdb.stanford.edu/dr-summary/comments/INSTI/ (last updated 25 November 2023)". The analysis also employed the 2022 IAS–USA DRM list using the HIV-1 subtype B reference sequence (GenBank accession number: K03455).

2.3. Examination of Differences in Sequence Composition

The analysis involved examining each position along the sequences to determine the presence of mutations. In this study, the polymorphisms were defined as mutations present in ≥0.5% of group M sequences. By comparing each base to a reference sequence, we calculated the frequency of amino acid changes at specific positions as follows:

$$f(i) = \frac{\sum_{j=1}^{N} I\left(x_j \neq a_i\right)}{L}$$

where $f(i)$ represents the mutation frequency at the position i, N is the total number of sequences examined, x_j is the amino acid at the position i in the sequence j, $I(x_j \neq a_i)$ is a function that returns 1 if $x_j \neq a_i$ and 0 otherwise, and L is the length of the reference sequence. This allowed us to identify positions of increased mutational activity, providing insight into the molecular evolution of the sequence. In addition to isolated positions, we constructed cumulative mutation profiles, illustrating the cumulative frequency of mutations over the entire length of the sequences. This facilitated the identification of regions characterized by higher mutation density, offering cues for potential functional domains or regions under selective pressure.

Next, divergence analysis aimed to quantify dissimilarity between the reference sequence and those within specific subtypes, excluding ambiguous or non-contributing elements, to provide a comprehensive perspective on the genetic variability present. For each subtype, the divergence was calculated as:

$$D = \frac{1}{N \times M} \sum_{i=1}^{N} \sum_{j=1}^{M} d_{ij} \times 100$$

where D represents the average divergence, N and M are the total number of sequences, and d_{ij} represents the divergence between the sequences i and j.

By comparing each sequence to a reference, we calculated the number of amino acid variations, describing the specific nature of genetic divergence at the protein level. The results were saved in a divergence matrix, which systematically captures the dissimilarity between each sequence pair. This matrix facilitated the identification of clusters of sequences with similar divergence profiles, contributing to a more granular understanding of genetic relationships. Finally, to test the statistical significance of the observed mutation frequencies, hypothesis tests, such as the chi-square test, were applied to determine whether the observed mutation patterns deviated significantly from what would be expected by chance alone. This step added a level of robustness to our results, ensuring that the identified mutation frequencies were not simply stochastic fluctuations. All statistical analyses were performed using R (version 4.3.1). A *p*-value of <0.05 was considered statistically significant.

**3. Results**

3.1. Distribution of Published HIV Integrase Sequences in SSA

Figure 1 shows the distribution of published sequences of HIV IN across SSA collected over forty years and included in the study.

We examined a total of 7906 HIV IN gene samples, representing 3% (7906/238, 924) of the HIV-1 sequences available in public databases. These samples were specifically selected for analysis because they corresponded to the gene of interest, with the remaining sequences being excluded. Our analysis encompassed integrase sequences across various HIV-1 subtypes from 1983 to 2023, including subtype A (24.4%, 1932/7906), B (0.17%, 14/7906), C (57.23%, 4525/7906), D (4.32%, 342/7906), F (0.59%, 47/7906), G (1.2%, 95/7906), H, J, K, L (0.21%, 17/7906), CRFs (10.62%, 840/7906), and the untyped groups N, O, P, and U (1.21%, 96/7906). While Figure 1 does not report the distribution of HIV subtypes and recombinant forms within the sub-region of SSA over time, it does depict the integrase

samples included in the study. The sampling of HIV subtypes in Figure 1 is in agreement with previous reports concerning the HIV genetic diversity in SSA [1–4].
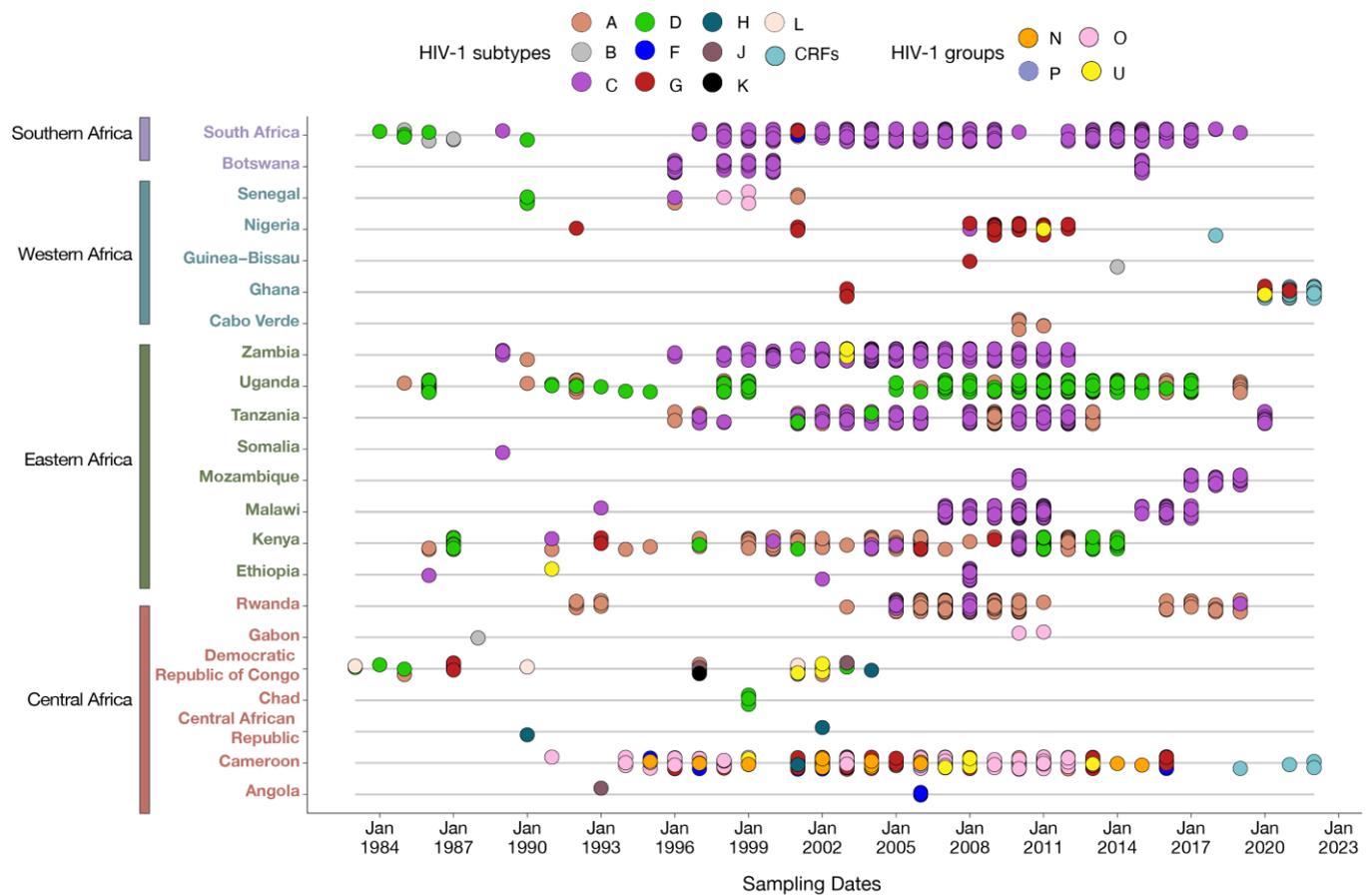
Figure 1. Temporal sampling of HIV integrase sequences in different African regions (1983–2022). Sampling of all available HIV IN sequences in different African countries through time, annotated according to their subtype assignment. Each dot represents an individual sample, with overlapping subtypes indicating isolation within the same time frame.

According to HIV-1 molecular epidemiology studies in the African continent [1–4], the majority of IN samples from 22 countries in SSA since the onset of the epidemic were of subtypes A, B, C, D, G, H, and L. From 1993 onwards, sequences of subtypes F, J, K, CRFs, and the untyped groups N, O, and U were also documented. Subtype C samples have been consistently significant throughout the entire period, especially in Southern and Eastern Africa and, to a lesser extent, in Central Africa. In Eastern Africa, subtypes A and D samples were predominant from 1990 to 2020. In contrast, Western Africa had limited representation, mainly with a few sequences from subtype G and CRFs. Central Africa showcased a diverse range of sequences from all subtypes and HIV groups throughout the period, with a notable expansion of CRFs between 2019 and 2022.

3.2. HIV Integrase Amino Acid Diversity

Table 1 shows the extent of IN amino acid diversity by groups and subtypes for sequences from two time periods, 1983–2007 and 2008–2023, using subtype B (HXB2) as the reference strain.

Table 1. The inter-group–subtype diversity of IN sequences (%).

Divergence	(1983–2007)		(2008–2023)
Subtype A (618) vs. B	0.06 ± 0.03	Subtype A (1314) vs. B	0.07 ± 0.04
Subtype B (11) vs. B	0.05 ± 0.03	Subtype B (3) vs. B	0.04 ± 0.03
Subtype C (2099) vs. B	0.06 ± 0.02	Subtype C (2024) vs. B	0.05 ± 0.02
Subtype D (148) vs. B	0.04 ± 0.02	Subtype D (194) vs. B	0.06 ± 0.05
Subtype F (32) vs. B	0.05 ± 0.02	Subtype F (15) vs. B	0.05 ± 0.03
Subtype G (51) vs. B	0.05 ± 0.01	Subtype G (44) vs. B	0.05 ± 0.02
Subtypes (H, J, K, L) (17) vs. B	0.06 ± 0.03	Subtypes (H, J, K, L) vs. B	-
CRF$_s$ (453) vs. B	0.06 ± 0.02	CRF$_s$ (387) vs. B	0.06 ± 0.02
Groups (N, O, P, U) (62) vs. B	0.12 ± 0.07	Groups (N, O, P, U) (34) vs. B	0.17 ± 0.19
All groups (M, N, O, P, U) (3491) vs. B	0.07 ± 0.04	All groups (M, N, O, P, U) (4415) vs. B	0.07 ± 0.04

Divergence: the mean proportion of amino acid difference between all sequence pairs. The number of sequences compared are in brackets.

Statistical analysis did not show significant differences in IN amino acid diversity between the years 1983 and 2007 and 2008 and 2023. However, an increase in amino acid diversity was found in subtypes A (0.06 ± 0.03 vs. 0.07 ± 0.04) and D (0.04 ± 0.02 vs. 0.06 ± 0.05), and notably in the NOPU group (0.12 ± 0.07 vs. 0.17 ± 0.19), while a decrease was noted in subtypes B (0.05 ± 0.03 vs. 0.04 ± 0.03) and C (0.06 ± 0.02 vs. 0.05 ± 0.02) between the years 1983 and 2007 and 2008 and 2023. For subtypes H, J, K, and L, amino acid diversity (0.06 ± 0.03) was estimated only for the 1983–2007 period. No variation was found in subtypes F, G, CRFs, and all HIV groups (M, N, O, P) and U between the two sampling periods. Overall, these results were in agreement with previous estimates of mean amino acid diversity, around 7% between HIV sub-subtypes and 0.11–0.18% between HIV groups [9]. Regarding subtypes A and D, the observed divergences between the years 1983 and 2007 and 2008 and 2023 suggested that there may have been a trend toward a genetic diversification over time, which was also much more evident in the NOPU group, that could shape INSTI effectiveness and INSTI DRM in SSA.

3.3. Natural Polymorphism Patterns and Conservation Analysis

To evaluate the conservation degree in INSTI-DRMs and polymorphisms of HIV-IN amino acidic sequences in SSA over time, we conducted a chronological analysis of the overall contribution of the entire study population. The dataset was divided into two time periods, 1983–2007 (3491 samples) vs. 2008–2023 (4415 samples), with consideration given to the time before and after approval of INSTI-based treatment regimens [13], and the amino acidic sequences were first aligned and compared to the reference strain B (HXB2) and then compared to each other (Figure 2).

The polymorphisms, defined as mutations present in ≥5% of sequences, were found in 49/288 (17.0%) and 54/288 (18.75%) amino acid positions in the sequences from the 1983–2007 and 2008–2023 periods, respectively. Highly polymorphic positions, defined as positions with substitutions detected in ≥20% of sequences analyzed, were lower in samples from 1983–2007, accounting for 7.29% (21/288) of positions, compared to those from 2008–2023, where they accounted for 9.02% (26/288). Some positions (K14, D25, V31, M50, F100, L101, T112, T124, T125, G134, K136, D167, V201, T206, T218, L234, A265, R269, D278, and S283) remained highly polymorphic across both periods, while other positions (E11, S17, I113, S119, and D256) were highly polymorphic only from 2008–2023.

Concerning highly conserved amino acid positions (with <1% variability), sequences between 1983 and 2007 and 2008 and 2023 showed similar levels of variability, accounting for 65.62% (189/288) and 64.63% (187/288), respectively. Among the 288 IN amino acids, there was no significant difference in variability concerning major or accessory INSTI DRMs, except for position G163, which was less conserved in samples from 2008–2023.
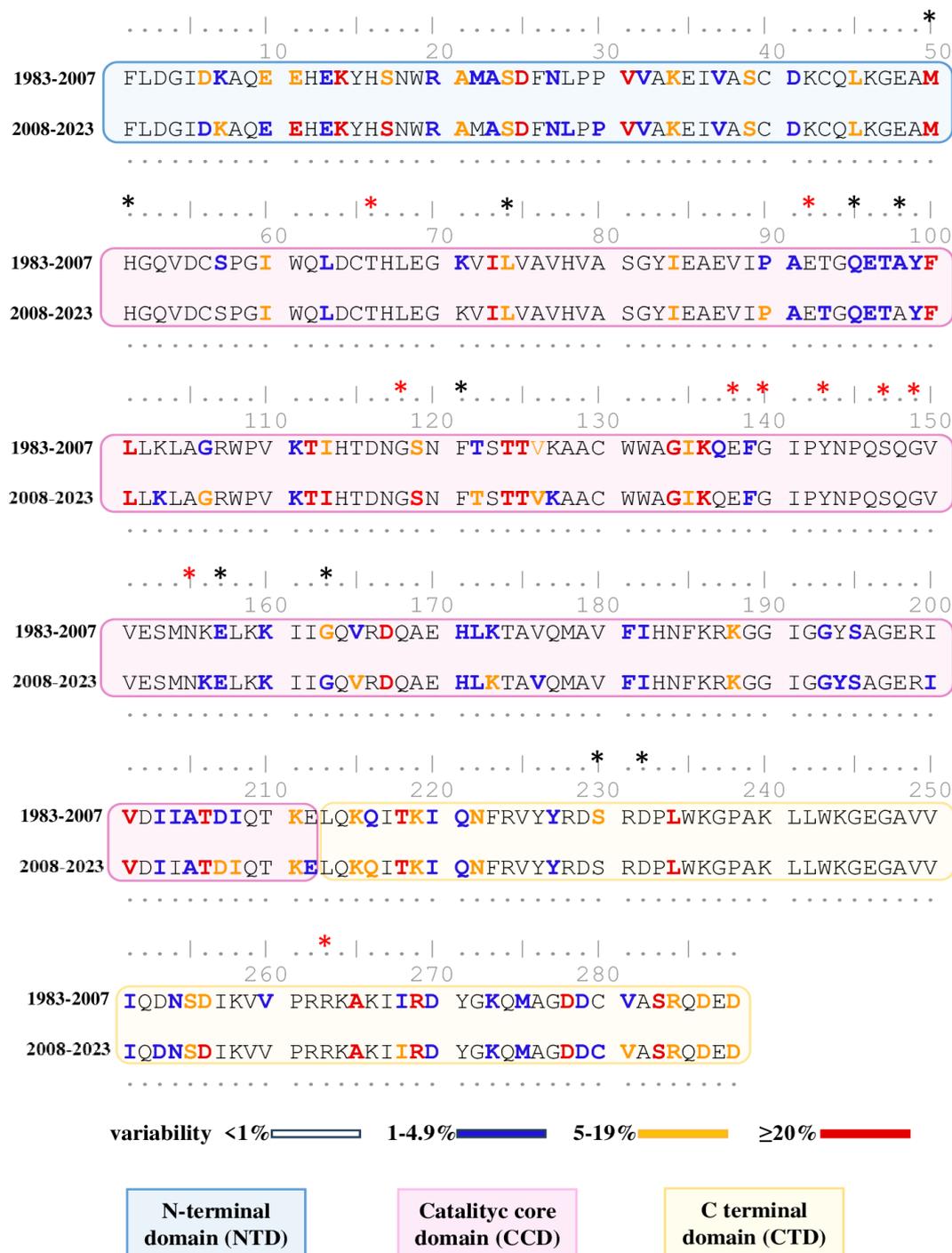
Figure 2. Comparison of degree variability between HIV integrase sequences from 1983–2007 vs. 2008–2023. The reference strain B (HXB2), used for the definition of mutations, is colored according to the frequency rate of mutations present in the 1983–2007 and 2008–2023 groups. Domains of IN are colored boxes. Red asterisk indicates major INSTI DRMs, and black asterisk indicates INSTI accessory polymorphic and nonpolymorphic mutations from the latest version of Stanford's list (https://hivdb.stanford.edu/, accessed on 25 November 2023).

3.4. Prevalence of INSTI Major DRMs

The overall estimated prevalence for major INSTI-DRMs over forty years was 0.35% (28/7906), comprising 0.51% (18/3491) in the period of 1983–2007 and 0.22% (10/4415) in the period of 2008–2023. Notably, there was a statistically significant decrease ($p = 0.03$) observed in 2008–2023 as compared to 1983–2007. Among the major INSTI-DRMs identified, the most common was T66A, accounting for 0.08% (6/7906), followed by Y143H and E138K at 0.05% (4/7906) each, G118R and S147G with 0.04% (3/7906) each, and N155D at 0.03% (2/7906). Less frequently E92G, E138A, G140R, Q148H, N155H, and N155T were observed at 0.01% (1/7906) each. No major DRMs were found in subtypes B, D (1983–2007), F, G, HJKL, CFR$_S$ (2008–2023), or the NOPU group. Further details regarding the prevalence of major DRMs over the two time periods, as per the Stanford HIVDB 8.9-1, can be found in Table 2.

Table 2. Distribution of major DRMs across subtypes.

	A 1983–2007 [618]	A 2008–2023 [1314]	C 1983–2007 [2099]	C 2008–2023 [2426]	D 1983–2007 [148]	CFR$_S$ 1983–2007 [453]
T66A	0.32 A(2)		0.1 A(2)	0.04 A(1)		0.22 A(1)
E92G			0.08 G(1)			
G118R				0.12 R(3)		
E138K/A		0.15 K(2)	0.05 A(1)	0.04 K(1)	0.6 K(1)	
G140R				0.12 R(1)		
Y143H			0.1 H(2)	0.04 H(1)	0.68 H(1)	
S147G			0.1 G(2)	0.04 G(1)		
Q148H			0.05 H(1)			
N155H/T/D			0.05 D(2)		0.68 H(1)	0.22 T(1)

Frequency of major DRMs, n (%), is shown. Round brackets: number of samples harboring the variant. Square brackets: number of samples tested for each subtype.

3.5. Prevalence of INSTI Accessory DRMs

Overall, the prevalence of INSTI accessory resistance mutations over the course of forty years was 46.61% (2665/7906), with a rate of 22.26% (1760/7906) in 1983–2007 samples vs. 24.29% (1921/7906) in 2008–2023 samples (as shown in Table 2). Among the identified INSTI accessory mutations, the most common were M50I at 33.70% (2665/7906), followed by L74I at 6.89% (544/7906), T97A at 2.77% (219/7906), E157Q at 1.59% (126/7906), and L74M at 1.23% (98/7906). Less-prevalent mutations included Q95K at 0.20% (16/7906), G163R at 0.1% (8/7906), D232N at 0.088% (7/7906), and G163K at 0.02% (2/7906).

The prevalence of residue substitutions in the 1983–2007 and 2008–2023 periods was as follows: M50I 35.40% (1236/3491) vs. 32.36% (1429/4415), L74I 6.87% (240/3491) vs. 6.88% (304/4415), Q95 0.22% (8/3491) vs. 0.18% (8/4415), T97A 4.2% (147/3491) vs. 1.63% (72/4415), E157Q 1.91% (67/3491) vs. 1.33% (59/4415), L74M 1.48% (52/3491) vs. 1.0% (46/4415), D232N 0.2% (7/3491), G163R 0.09% (3/3491) vs. 0.02% (1/4415), and G163K 0.04% (2/4415). In the NOPU group, M50I and L74I were statistically significantly more prevalent in the years 2008–2023 than in 1983–2007 ($p < 0.05$), while in subtype A, mutation T97A was statistically significantly more prevalent in the years 1983–2007 ($p < 0.05$). Mutation D232N was detected only in samples from the years 1983–2007 (see Table 3).

Table 3. Distribution of INSTI accessory resistance across subtypes.

	A 1983-2007 [618]	A 2008-2023 [1314]	C 1983-2007 [2099]	C 2008-2023 [2426]	D 1983-2007 [148]	D 2008-2023 [194]	F 1983-2007 [32]	F 2008-2023 [15]	G 1983-2007 [51]	G 2008-2023 [44]	HJIL 1983-2007 [17]	CFR$_S$ 1983-2007 [453]	CFR$_S$ 2008-2023 [387]	NOPU 1983-2007 [62]	NOPU 2008-2023 [34]
M50I	18.28 I(113)	9.36 I(123)	47.97 I(1007)	48.8 I(1184)	0.67 I(1)	2.58 I(5)	18.75 I(6)	26.6 I(4)	11.74 I(6)	18.18 I(8)		17.66 I(80)	19.9 I(77)	* 37.09 **I(23)**	* 85.29 **I(28)**

Table 3. *Cont.*

	A	A	C	C	D	D	F	F	G	G	HJIL	CFR$_S$	CFR$_S$	NOPU	NOPU
L74M/I	5.82 I(36) 0.97 M(6)	3.50 I(46) 1.67 M(22)	6.29 I(83) 0.76 M(16)	6.00 I(146) 0.08 M(2)	4.00 I(6) 0.67 M(1)	2.58 I(5)	3.10 M(1)	6.60 I(1)	15.69 I(8) 1.96 M(1)	18.18 I(8) 2.27 M(1)		15.45 I(70) 6.00 M(27)	17.57 I(68) 5.42 M(21)	* 59.67 **I(37)**	* 88.24 **I(30)**
Q95K	0.32 K(2)	0.15 K(2)	0.14 K(3)	0.16 K(4)		0.51 K(1)				2.27 K(1)	5.88 K(1)	0.44 K(2)			
T97A	* 15.04 **A(93)**	* 3.27 **A(43)**	0.86 A(18)	0.28 A(7)	2.68 A(4)		12.50 A(4)	6.60 A(1)	5.88 A(3)	6.82 A(3)	5.88 A(1)	3.97 A(18)	4.65 A(18)	9.60 A(6)	
E157Q	1.13 Q(7)	2.28 Q(30)	2.00 Q(42)	0.28 Q(7)	2.68 Q(4)					2.27 Q(1)		3.00 Q(14)	5.42 Q(21)		
G163R/K	0.16 R(1)	0.07 K(1)	0.05 R(1)	0.16 R(4)	0.67 R(1)								0.25 K(1) 0.25 R(1)		
D232N			0.19 N(4)		1.35 N(2)							0.22 N(1)			

Frequency of accessory DRMs, n (%), is shown. Round brackets: number of samples harboring the variant. Square brackets: number of samples tested for each subtype; * NOPU group M50I and L74I (1983–2007) vs. (2008–2023), subtype A, T97 (1983–2007) vs. (2008–2023) ($p < 0.05$).

4. Discussion

Limited information is available regarding the evolution of INSTI-DRM resistance in SSA, underscoring the presence of pre-existing DRMs before transitioning to dolutegravir-based ART regimens, which could significantly impact virological outcomes [14,15]. Several groups have reported contrasting data on INSTI-DRMs in SSA, with many showing a lack of major DRMs and a low frequency of accessory INSTI-DRMs [15–17], while others reported frequencies ranging from 0.8% up to 5.4% of major INSTI-DRMs in INSTI-naïve populations, suggesting a potential threat to the effectiveness of the DTG-based ART programs in SSA [18–20].

Our large-scale analysis provided a comprehensive mapping of polymorphisms and conserved amino acid positions within the HIV IN gene across major HIV-1 subtypes and HIV groups, considering both INSTI-naïve and experienced individuals. This analysis also examined the baseline INSTI resistance, encompassing major and accessory DRMs, with implications for the use of DTG as a first-line regimen.

HIV-1 genetic diversity in SSA plays a crucial role in susceptibility to INSTIs and the selection of resistance mutations [18–22]. In our study, the majority of the viruses analyzed were subtypes C and A, followed by CRFs, D, F, and G, as well as groups N, O, P, and U, whereas subtypes H, J, F, L, and B were poorly represented, reflecting the distribution of HIV infection in SSA. Despite a significant increase in the number of CRFs and changes in the geographic distribution of M HIV-1 subtypes and groups N, O, P, and U forms over the past forty years, the diversity estimates for IN inter-subtype diversity levels remained consistently low at <0.07, as did inter-group diversity at <0.18, during both the 1983–2007 and 2008–2023 periods. This suggested that the IN region is a conserved drug target [9].

Analysis of the conservation within the genetically conserved key function enzyme amino acids revealed similar levels of variability (<1%) in both the 1983–2007 and 2008–2023 periods, consistent with findings in HIV-1 group M subtypes and previous studies [9,18]. In contrast, a higher number of polymorphic positions were detected at a frequency of ≥5% or ≥20% in samples from the 2008–2023 period. Notably, residue G163, where substitutions have been reported to act synergistically with other mutations to reduce susceptibility to INSTIs, appeared to be more polymorphic (≥5% variability) in samples from 1983–2007 [23].

Estimates of INSTI-DRMs are limited within SSA [24–38]. Most reports indicate that primary INSTI DRMs are rare, and very few INSTI accessory resistances can be detected prior to the initiation of INSTI therapy, possibly due to transmitted INSTI resistances and the cumulative natural occurrence of mutations without selective drug pressure [27,39]. DRMs such as G118R, E138K/A/T, G140S/A/C/R, Q148H/R/K, S153 F/Y, N155H, and R263K have been identified as being associated with DTG resistance [34]. In our study, the

examination of INSTI resistance-associated mutations showed low levels of major DRMs, such as T66A, E92G, G118R, E138K/A, G140R, Y143H, S147G, Q148H, and N155H/T/D, over forty years, and as mentioned above, some of these have been reported to have an effect on DTG. Notably, it was unexpected to detect major INSTI-DRMs at statistically significantly higher frequencies in 1983–2007 compared to 2008–2023. The most common major DRM was T66A, found in Cameroon, Rwanda, and South Africa [38]. The rare mutation E92G was present in South Africa; G118, previously described as a DTG-resistance pathway in non-B subtypes, was present in South Africa; the extremely rare mutation G140R was present in South Africa; and the less common mutation Y143H was present in Uganda and Botswana. The appearance of INSTI resistance mutations in the pre-INSTI era, as found in our study, might be explained by low retention in care and poor treatment adherence in SSA, while the decline in the frequency of DRMs in the INSTI era likely reflects the availability of ART regimens and improved adherence in treatment-experienced populations [37,39].

Regarding accessory DRMs, in spite of the broad genetic diversity over forty years, no relevant difference in mutation frequency between 1983–2007 and 2008–2023 was detected, and more than 46% of the study population harbored at least one mutation. Seven accessory INSTI-DRMs that may contribute to drug resistance when combined with major DRMs were found, with M50I, L74I/M, T97A, and E157Q being the most frequent. In particular, M50I, L74I/M, and T97A, in combination with other INSTI drug resistance mutations, may contribute to reduced DTG susceptibility [37,40]. These mutations, which have been shown to occur before initiation of cART with different prevalences depending on the subtypes, could facilitate viral evolution of resistance under drug-selective pressure. M50I has been shown to increase DTG resistance in combination with R263K [39]. Here, we show a high prevalence of the M50I in HIV subtypes circulating in SSA over forty years, particularly in subtype C and in HIV groups N, O, P, and U, with a higher increase in groups N, O, P, and U in 2008–2023. L74 is located in the catalytic core domain (amino acids 51–212) and is involved in the catalytic processes of the enzyme. L74M, in combination with T79A, may reduce susceptibility to RAL and/or EVG in the absence of primary INI-resistance mutations [40]. It has also been reported that L74M combined with G118R or E138K mutations increases resistance to DTG [41]. To date, the role of L74I substitution is still unclear [42,43], although it has been shown to enhance integrase inhibitor resistance in combination with additional major INSTI mutations [8] and is also being evaluated as a putative cause of DRMs for ART using DTG [41]. Our analysis indicated a low prevalence of L74I polymorphisms in subtypes A, D, and F, a relatively high prevalence in subtypes C, G, and CRFs, and a high prevalence in groups N, O, P, and U, particularly in 2008–2023, while L74M remained relatively low except in CFRs, where it reached a prevalence of about 6%.

T97A is a relatively common polymorphism-associated INSTI resistance and is co-selected in the presence of major INSTI-DRMs by RAL and EVG and by DTG in INSTI-experienced individuals [44]. Our analysis revealed that in subtype A, mutation T97A had a statistically significantly higher prevalence in the years 1983–2007 compared with 2008–2023, and it had a relatively high prevalence in subtype F and the NOPU group, while it was absent in the NOPU group in the years 2008–2023. Our results suggest the existence of major and accessory DRMs prior to INSTI-based therapy that poses a minimal risk of compromised virologic response for patients enrolling in INSTI-based therapy.

This study is limited in that (a) our analysis did not distinguish ART-experienced from treatment-naïve individuals in the years 1983–2007 and 2008–2023, (b) the heterogeneity in the number of sequences analyzed for different African regions depends on the number of sequences deposited in the Los Alamos HIV database, (c) the limited number of available IN sequences for subtypes F and G and NOPU groups may have affected the frequency of DRMs and polymorphisms at specific positions.

5. Conclusions

In summary, our study presented a comprehensive analysis of HIV integrase gene variability over forty years. Our data showed a higher prevalence of major DRMs in the years 1983–2007, prior to INSTI-based therapy, than in the years 2008–2023. This could be explained by the low retention in care and the poor treatment adherence in SSA, which favor genetic variability and the high evolution rate of HIV-1. No significant difference in mutation frequency was detected in accessory DRMs in the years 1983–2007 and 2008–2023. Natural HIV-1 variation could explain this because accessory DRMs occur in the absence of therapy, but similarly to major DRMs, other factors, such as poor adherence support and poor retention in care, may also support the level of resistance reported here. The results of this study suggest the need to expand and strengthen surveillance through an HIV drug resistance national survey.

References

1. World Health Organization (WHO). Available online: https://www.afro.who.int/health-topics/hivaids (accessed on 10 October 2023).
2. Global Report 2022. UNAIDS Report on the Global AIDS Epidemic 2020. UNAIDS Web Site. Available online: https://www.unaids.org/en/resources/fact-sheet (accessed on 10 October 2023).
3. Hemelaar, J.; Loganathan, S.; Elangovan, R.; Yun, J.; Dickson-Tetteh, L.; Kirtley, S. WHO-UNAIDS Network for HIV Country-level diversity of the HIV-1 pandemic between 1990 and 2015. *J. Virol.* **2020**, *21*, JVI.01580-20.
4. Giovanetti, M.; Ciccozzi, M.; Parolin, C.; Borsetti, A. Molecular Epidemiology of HIV-1 in African Countries: A Comprehensive Overview. *Pathogens* **2020**, *9*, 1072. [CrossRef] [PubMed]
5. WHO. HIV Drug Resistance Report 2021. 2021. Available online: https://www.who.int/publications/i/item/9789240038608 (accessed on 25 November 2023).
6. WHO. Updated Recommendations on First-Line and Second-Line Antiretroviral Regimens and Post-Exposure Prophylaxis and Recommendations on Early Infant Diagnosis of HIV—Interim Guidance. 2018. Available online: https://www.who.int/hiv/pub/guidelines/ARV2018update/en/ (accessed on 10 October 2023).
7. Zhao, A.V.; Crutchley, R.D.; Guduru, R.; Ton, K.; Lam, T.; Min, A.C. A clinical review of HIV integrase strand transfer inhibitors (INSTIs) for the prevention and treatment of HIV-1 infection. *Retrovirology* **2022**, *19*, 22. [CrossRef] [PubMed]
8. Smith, S.J.; Zhao, X.Z.; Passos, D.O.; Lyumkis, D.; Burke, T.R., Jr.; Hughes, S.H. Integrase Strand Transfer Inhibitors Are Effective Anti-HIV Drug. *Viruses* **2021**, *13*, 205. [CrossRef] [PubMed]
9. Rhee, S.Y.; Liu, T.F.; Kiuchi, M.; Zioni, R.; Gifford, R.J.; Holmes, S.P.; Shafer, R.W. Natural variation of HIV-1 group M integrase: Implications for a new class of antiretroviral inhibitors. *Retrovirology* **2008**, *5*, 74. [CrossRef] [PubMed]
10. Thierry, E.; Deprez, E.; Delelis, O. Different Pathways Leading to Integrase Inhibitors Resistance. *Front. Microbiol.* **2017**, *7*, 2165. [CrossRef]
11. Tang, M.W.; Liu, T.F.; Shafer, R.W. The HIVdb System for HIV-1 Genotypic Resistance Interpretation. *Intervirology* **2012**, *55*, 98–101. [CrossRef] [PubMed]
12. Low, A.; Prada, N.; Topper, M.; Vaida, F.; Castor, D.; Mohri, H.; Hazuda, D.; Muesing, M.; Markowitz, M. Natural Polymorphisms of Human Immunodeficiency Virus Type 1 Integrase and Inherent Susceptibilities to a Panel of Integrase Inhibitors. *Antimicrob. Agents Chemother.* **2009**, *53*, 4275–4282. [CrossRef]

13. Ngoufack Jagni Semengue, E.; Santoro, M.M.; Ndze, V.N.; Ka'e, A.C.; Yagai, B.; Nka, A.D.; Dambaya, B.; Takou, D.; Teto, G.; Fabeni, L.; et al. HIV-1 integrase resistance associated mutations and the use of dolutegravir in Sub-Saharan Africa: A systematic review and meta-analysis. *PLoS Glob. Public Health* **2022**, *10*, e0000826. [CrossRef]

14. Schramm, B.; Temfack, E.; Descamps, D.; Nicholas, S.; Peytavin, G.; Bitilinyu-Bangoh, J.E.; Storto, A. Viral suppression and HIV-1 drug resistance 1 year after pragmatic transitioning to dolutegravir first-line therapy in Malawi: A prospective cohort study. *Lancet HIV* **2022**, *9*, e544–e553. [CrossRef]

15. Giovanetti, M.; Farcomeni, S.; Sernicola, L.; Virtuoso, S.; Fontanelli Sulekova, L.; Maggiorella, M.T.; Buttò, S.; Taliani, G.; Ciccozzi, M.; Borsetti, A. Analysis of HIV-1 integrase genotypes and polymorphisms among integrase inhibitors-based antiretroviral treatment naïve patients in South Sudan. *J. Med. Virol.* **2022**, *94*, 3320–3327. [CrossRef] [PubMed]

16. Mabeya, S.; Nyamache, A.; Ngugi, C.; Nyerere, A.; Lihana, R. Characterization of HIV-1 integrase gene and resistance associated mutations prior to roll out of integrase inhibitors by Kenyan National HIV-Treatment Program in Kenya. *Ethiop. J. Health Sci.* **2020**, *30*, 37–44. [PubMed]

17. Arimide, D.A.; Szojka, Z.I.; Zealiyas, K.; Gebreegziabxier, A.; Adugna, F.; Sasinovich, S.; Björkman, P.; Medstrand, P. Pre-Treatment Integrase Inhibitor Resistance and Natural Polymorphisms among HIV-1 Subtype C Infected Patients in Ethiopia. *Viruses* **2022**, *14*, 729. [CrossRef] [PubMed]

18. Semengue, E.N.J.; Armenia, D.; Inzaule, S.; Santoro, M.M.; Dambaya, B.; Takou, D.; Teto, G.; Nka, A.D.; Yagai, B.; Fabeni, L.; et al. Baseline integrase drug resistance mutations and conserved regions across HIV-1 clades in Cameroon: Implications for transition to dolutegravir in resource-limited settings. *J. Antimicrob. Chemother.* **2021**, *76*, 1277–1285. [CrossRef] [PubMed]

19. Mikasi, S.G.; Gichana, J.O.; van der Walt, C.; Brado, D.; Obasa, A.E.; Njenda, D.; Messembe, M.; Lyonga, E.; Assoumou, O.; Cloete, R.; et al. HIV-1 integrase diversity and resistance-associated mutations and polymorphisms among integrase strand transfer inhibitor-naive HIV-1 patients from Cameroon. *AIDS Res. Hum. Retrov.* **2020**, *36*, 450–455. [CrossRef] [PubMed]

20. Mikasi, S.G.; Isaacs, D.; Ikomey, G.M.; Shimba, H.; Cloete, R.; Jacobs, G.B. HIV-1 drug resistance mutation analyses of Cameroon-derived integrase sequences. *AIDS Res. Hum. Retrov.* **2021**, *37*, 54–56. [CrossRef] [PubMed]

21. Han, Y.S.; Mesplède, T.; Wainberg, M.A. Differences among HIV-1 subtypes in drug resistance against integrase inhibitors. *Infect. Genet. Evol.* **2016**, *46*, 286–291. [CrossRef] [PubMed]

22. Inzaule, S.C.; Hamers, R.L.; Noguera-Julian, M.; Casadellà, M.; Parera, M.; de Wit, T.F.R.; Paredes, R. Primary resistance to integrase strand transfer inhibitors in patients infected with diverse HIV-1 subtypes in sub-Saharan Africa. *J. Antimicrob. Chemother.* **2018**, *73*, 1167–1172. [CrossRef]

23. Fofana, D.B.; Diarra, H.; Guindo, I.; Savadogo, M.K.; d'Almeida, M.; Diallo, F.I.; Baldé, A.; Soulié, C.; Kone, A.; Marcelin, A.-G.; et al. Prevalence of HIV-1 Natural Polymorphisms and Integrase-Resistance-Associated Mutations in African Children. *Viruses* **2023**, *15*, 546. [CrossRef]

24. *Update on the Transition to Dolutegravir-Based Antiretroviral Therapy: Report of a WHO Meeting. 29–30 March 2022*; World Health Organization: Geneva, Switzerland, 2022.

25. Kouamou, V.; Inzaule, S.; Manasa, J. Dolutegravir drug-resistance monitoring in Africa. *Lancet* **2021**, *8*, e664–e666. [CrossRef]

26. Taylor, G. Rolling out HIV antiretroviral therapy in sub-Saharan Africa: 2003–2017. *Can. Commun. Dis. Rep.* **2018**, *44*, 68–70. [CrossRef] [PubMed]

27. Tzou, P.L.; Rhee, S.Y.; Descamps, D.; Clutter, D.S.; Hare, B.; Mor, O.; Grude, M.; Parkin, N.; Jordan, M.R.; Bertagnolio, S.; et al. Integrase strand transfer inhibitor (INSTI)-resistance mutations for the surveillance of transmitted HIV-1 drug resistance. *J. Antimicrob. Chemother.* **2020**, *75*, 170–182. [CrossRef]

28. Rhee, S.-Y.; Tzou, P.; Shafer, R. Temporal Trends in HIV-1 Mutations Used for the Surveillance of Transmitted Drug Resistance. *Viruses* **2021**, *13*, 879. [CrossRef] [PubMed]

29. Depatureaux, A.; Mesplède, T.; Quashie, P.; Oliveira, M.; Moisi, D.; Plantier, J.C.; Brenner, B.; Wainberg, M.A. HIV-1 group O resistance against integrase inhibitors. *J. Acquir. Immune Defic. Syndr.* **2015**, *70*, 9–15. [CrossRef] [PubMed]

30. Lataillade, M.; Chiarella, J.; Kozal, M.J. Natural Polymorphism of the HIV-1 Integrase Gene and Mutations associated with Integrase Inhibitor Resistance. *Antivir. Ther.* **2007**, *12*, 563–570. [CrossRef] [PubMed]

31. Boerma, R.S.; Sigaloff, K.C.E.; Akanmu, A.S.; Inzaule, S.; van Hensbroek, M.B.; de Wit, T.F.R.; Calis, J.C. Alarming increase in pretreatment HIV drug resistance in children living in sub-Saharan Africa: A systematic review and meta-analysis. *J. Antimicrob. Chemother.* **2016**, *72*, 365–371. [CrossRef] [PubMed]

32. Madyadi, A.; Dhoro, M.; Shamu, T.; Washaya, T.; Kouamou, M.V.; Chimukangara, B.; Katzenstein, D.; Manasa, J. HIV-1 Genetic Diversity and Natural Polymorphisms of the Integrase Gene in Integrase Inhibitor-Naive Patients in Harare, Zimbabwe. *AIDS Res. Hum. Retrov.* **2021**, *37*, 954–961. [CrossRef]

33. Ndashimye, E.; Avino, M.; Olabode, A.S.; Poon, A.F.Y.; Gibson, R.M.; Li, Y.; Meadows, A.; Tan, C.; Reyes, P.S.; Kityo, C.M.; et al. Accumulation of integrase strand transfer inhibitor resistance mutations confers high-level resistance to dolutegravir in non-B subtype HIV-1 strains from patients failing raltegravir in Uganda. *J. Antimicrob. Chemother.* **2020**, *75*, 3525–3533. [CrossRef]

34. Brado, D.; Obasa, A.; Ikomey, G.M.; Cloete, R.; Singh, K.; Engelbrecht, S.; Neogi, U.; Jacobs, G.B. Analyses of HIV-1 integrase sequences prior to South African national HIV-treatment program and available of integrase inhibitors in Cape Town, South Africa. *Sci. Rep.* **2018**, *8*, 4709. [CrossRef]

35. Kiros, M.; Tefera, D.A.; Andualem, H.; Geteneh, A.; Tesfaye, A.; Woldemichael, T.S.; Kidane, E.; Alemayehu, D.H.; Maier, M.; Mihret, A.; et al. Low level of HIV-1 C integrase strand transfer inhibitor resistance mutations among recently diagnosed ART-naive Ethiopians. *Sci. Rep.* **2023**, *13*, 6546. [CrossRef]

36. Abdullahi, A.; Kida, I.M.; Maina, U.A.; Ibrahim, A.H.; Mshelia, J.; Wisso, H.; Adamu, A.; Onyemata, J.E.; Edun, M.; Yusuph, H.; et al. Limited emergence of resistance to integrase strand transfer inhibitors (INSTIs) in ART-experienced participants failing dolutegravir-based antiretroviral therapy: A cross-sectional analysis of a Northeast Nigerian cohort. *J. Antimicrob. Chemother.* **2023**, *8*, 2000–2007. [CrossRef] [PubMed]

37. Abram, M.E.; Ram, R.R.; Margot, N.A.; Barnes, T.L.; White, K.L.; Callebaut, C.; Miller, M.D. Lack of impact of pre-existing T97A HIV-1 integrase mutation on integrase strand transfer inhibitor resistance and treatment outcome. *PLoS ONE* **2017**, *12*, e0172206. [CrossRef] [PubMed]

38. Wensing, A.M.; Calvez, V.; Ceccherini-Silberstein, F.; Charpentier, C.; Günthard, H.F.; Paredes, R.; Shafer, R.W.; Richman, D.D. Update of the drug resistance mutations in HIV-1. *Top. Antivir. Med.* **2022**, *30*, 559–574. [PubMed]

39. Gil, H.; Delgado, E.; Benito, S.; Moreno-Lorenzo, M.; Thomson, M.M.; Spanish Group for the Study of Antiretroviral Drug Resistance. Factors associated with HIV-1 resistance to integrase strand transfer inhibitors in Spain: Implications for dolutegravir-containing regimens. *Front. Microbiol.* **2022**, *13*, 1051096. [CrossRef] [PubMed]

40. Wares, M.; Mesplède, T.; Quashie, P.K.; Osman, N.; Han, Y.; Wainberg, M.A. The M50 polymorphic substitution in association with the R263K mutation in HIV-1 subtype B integrase increases drug resistance but does not restore viral replicative fitness. *Retrovirology* **2014**, *11*, 7. [CrossRef] [PubMed]

41. Venter, W.D.F.; Moorhouse, M.; Sokhela, S.; Fairlie, L.; Mashabane, N.; Masenya, M.; Serenata, C.; Akpomiemie, G.; Qavi, A.; Chandiwana, N.; et al. Dolutegravir plus two different prodrugs of tenofovir to treat HIV. *N. Engl. J. Med.* **2019**, *381*, 803–815. [CrossRef]

42. El Bouzidi, K.; Kemp, S.A.; Datir, R.P.; Murtala-Ibrahim, F.; Aliyu, A.; Kwaghe, V.; Frampton, D.; Roy, S.; Breuer, J.; Sabin, A.C.; et al. High prevalence of integrase mutation L74I in West African HIV-1 subtypes prior to integrase inhibitor treatment. *J. Antimicrob. Chemother.* **2020**, *75*, 1575–1579. [CrossRef]

43. Parolin, C.; Borsetti, A.; Choe, H.; Farzan, M.; Kolchinsky, P.; Heesen, M.; Ma, Q.; Gerard, C.; Palú, G.; Dorf, M.E.; et al. Use of murine CXCR-4 as a second receptor by some T-cell-tropic human immunodeficiency viruses. *J. Virol.* **1998**, *72*, 1652–1656. [CrossRef]

44. Hardy, I.; Brenner, B.; Quashie, P.; Thomas, R.; Petropoulos, C.; Huang, W.; Moisi, D.; Wainberg, M.A.; Roger, M. Evolution of a novel pathway leading to dolutegravir resistance in a patient harbouring N155H and multiclass drug resistance. *J. Antimicrob. Chemother.* **2015**, *70*, 405–441. [CrossRef]