REVIEW

OPEN

HIV subtype diversity worldwide

Nicholas Bbosa[a], *Pontiano Kaleebu*[a,b], *and Deogratius Ssemwanga*[a,b]

Purpose of review
To provide a summary of the current data on the global HIV subtype diversity and distribution by region. HIV is one of the most genetically diverse pathogens due to its high-mutation and recombination rates, large population size and rapid replication rate. This rapid evolutionary process has resulted in several HIV subtypes that are heterogeneously globally distributed.

Recent findings
Subtype A remains the most prevalent strain in parts of East Africa, Russia and former Soviet Union countries; subtype B in Europe, Americas and Oceania; subtype C in Southern Africa and India; CRF01_AE in Asia and CRF02_AG in Western Africa. Recent studies based on near full-length genome sequencing highlighted the growing importance of recombinant variants and subtype C viruses.

Summary
The dynamic change in HIV subtype distribution presents future challenges for diagnosis, treatment and vaccine design and development. An increase in recombinant viruses suggests that coinfection and superinfection by divergent HIV strains has become more common necessitating continuous surveillance to keep track of the viral diversity. Cheaper near full-length genome sequencing approaches are critical in improving HIV subtype estimations. However, missing subtype data and low sequence sampling levels are still a challenge in some geographical regions.

Video abstract
http://links.lww.com/COHA/A14.

Keywords
diversity, HIV, subtype

INTRODUCTION

As of 2017, there were 36.9 million people living with HIV worldwide with a majority of these infections in sub-Saharan Africa [1]. AIDS is a global pandemic caused by two genetically diverse lentiviruses, HIV-1 and HIV-2 [2], that were introduced through multiple cross-species transmissions of simian immunodeficiency viruses from nonhuman primates to humans. These distinct zoonotic viral transmissions gave rise to separate HIV-1 groups M (Major), O (Outlier), N (non-M, non-O) and the most recent group P [2]. The origin of HIV-1 has been documented around Kinshasa in the present day Democratic Republic of Congo around the 1920s from where it spread along a transport network to other areas in sub-Saharan Africa, West Africa, Europe and the rest of the world [3]. This global spread was marked by a geographically defined distribution of several genetically distinct viruses. For instance, subtype B became prevalent in almost all parts of Europe and the Americas while a variety of subtypes and intersubtype recombinants are found in Africa with the highest diversity reported in West Central Africa. Whereas group M viruses have dominated the global HIV pandemic since its inception, other group N, O and P viruses have not been disseminated widely. Group M viruses are further subdivided into nine subtypes (A-D, F-H, J, K) [2]. The genetic distance within a subtype can be between 15 and 20%, whereas genetic distances across subtypes are usually 25 and 35% [4]. HIV-2 remains largely restricted to the

[a]Medical Research Council (MRC)/Uganda Virus Research Institute (UVRI) and London School of Hygiene and Tropical Medicine (LSHTM) Uganda Research Unit and [b]Uganda Virus Research Institute, Entebbe, Uganda

Correspondence to Deogratius Ssemwanga, Medical Research Council (MRC)/Uganda Virus Research Institute (UVRI) and London School of Hygiene and Tropical Medicine (LSHTM) Uganda Research Unit, P.O. Box 49, Entebbe, Uganda. Tel: +256 (0) 417 704170;
e-mail: Deogratius.Ssemwanga@mrcuganda.org

Curr Opin HIV AIDS 2019, 14:153–160

DOI:10.1097/COH.0000000000000534

KEY POINTS

- Subtypes A, C and CRF02-AG are the most prevalent on the African continent, whereas subtype B predominates in Europe/the Americas and CRF01_AE dominate in Asia.
- There is an increase in the prevalence of recombinant viruses particularly in geographical areas where multiple subtypes cocirculate such as Asia and Africa.
- The emerging and dynamic HIV genetic diversity needs constant monitoring due to the challenges it presents towards diagnosis, phylogenetic reconstruction, treatment and vaccine development.
- The classification of HIV variants based on near full-length genomes is still limited but needs to be scaled up in a cost-effective way.
- The problem of missing HIV subtype data and low sequence sampling levels needs to be addressed.

western part of Africa although viral introductions have been reported elsewhere in Europe (Portugal and France), India and the United States of America [5]. HIV-2 has also been documented to be less infectious than HIV-1 [6,7]; it is comprised of at least nine groups (formerly referred to as subtypes; A to I) of which groups A and D are presently circulating [5]. HIV-2 subtype data is still limited and only a few recombinants have been described [5,8].

CIRCULATING RECOMBINANT FORMS

Recombination is an important phenomenon for viral diversification that allows the virus to evade the host immune system and antiretroviral treatment [9]. A recombinant is a viral sequence that has regions from two or more distinct parental strains [10]. A circulating recombinant form or circulating recombinant form (CRF) is a combination of viral genomes of different subtypes in dually infected people that results in a mosaic genome composed of regions from each of the different subtypes. When these recombinants are transmitted and spread within a population, they are recognized as circulating strains in the HIV epidemic and are classified as CRFs [11]. The virus must be isolated from at least two unlinked individuals and fully sequenced [12]. The designation 'cpx' (complex) is used when more than three subtypes are present [13]. To date, 98 CRFs have been identified and updated in the Los Alamos National Laboratory HIV sequence database [14]. These are designated with sequential numbering as they are discovered and reported.

UNIQUE RECOMBINANT FORMS

Unique recombinant forms (URFs) consist of a mixture of subtypes, but unlike the CRFs, they were sampled only once from a single multiply-infected individual [13]. Hence, all sequences corresponding to a CRF have the same recombination break points in the genome while each URF shows unique breakpoints. Intrasubtype recombination does occur but as it does not lead to subtype recombination, these viruses are not described as URFs. Dual infections with HIV-1 and HIV-2 have frequently been reported in regions where both viruses circulate [15] although no recombinants between them have been described. Several AD and AC intersubtype recombinants have been described in Eastern Africa [16,17], where these three subtypes cocirculate. BF intersubtype recombinants have been found in Brazil and Argentina, where subtypes B and F are both common [18,19]. Subtypes A and C cocirculate in India with AC recombinants [20].

LITERATURE SEARCH

We performed a structured literature search in PubMed using MeSH terms. A combination of keywords related to the subject of 'HIV genetic diversity worldwide' were used with Boolean operators (OR and AND) in PubMed's advanced search builder. The results generated in the search builder at each level were combined to output all articles that overlapped across all the search words for the specified review period (Fig. 1). This article focused mainly on HIV-1 as there was limited data on HIV-2 during the review period.

GLOBAL DISTRIBUTION OF HIV-1 BY REGION

Africa

The greatest genetic diversity of HIV-1 has been found in Africa especially in West Central Africa [21]; however, other parts of the continent show an assortment of diverse viral strains. In Southern Africa, subtype C collectively predominates [22,23], whereas in Eastern Africa subtype A is the most dominant [24▪▪] although other subtype D and C viruses cocirculate as well [25]. An increase in subtype C and a decrease in subtype D have been reported in Kenya together with extensive intersubtype recombination [24▪▪]. A recent study in Uganda reported an increased prevalence of HIV-1 intersubtype recombinants of about 46% based on near full length HIV sequences [26▪▪]. A separate study by the Phylogenetic And Networks for Generalized HIV Epidemics in Africa (PANGEA)-HIV consortium [27▪] showed a high proportion (about 50%) of

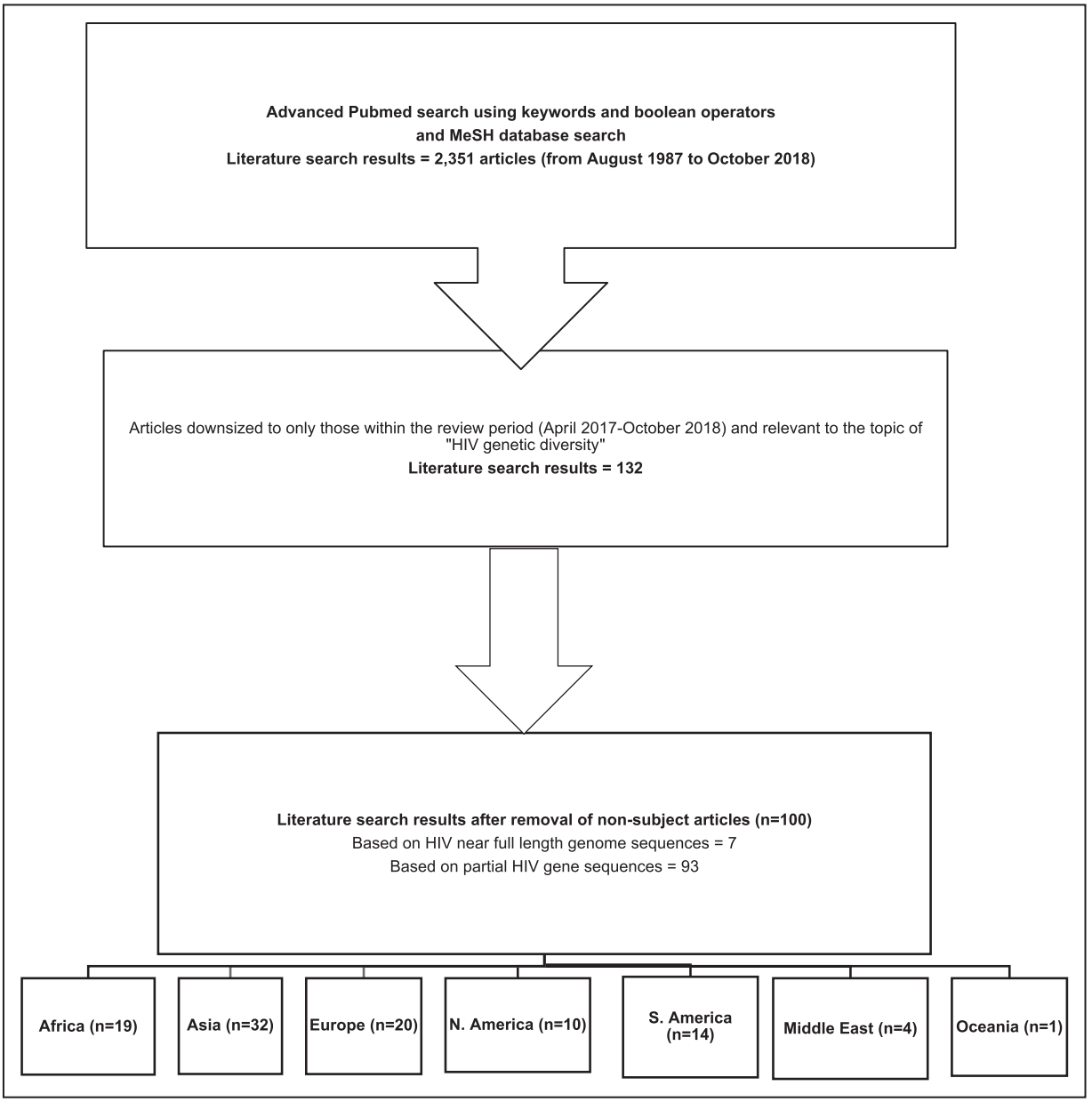

FIGURE 1. Literature search. Diagrammatic representation of the literature search and selection of articles.

HIV-1 intersubtype recombinants in Uganda based on deep sequencing of near full-length viral genomes (unpublished data). In West and West Central Africa, the majority of viruses are CRF02_AG [21,28]. Subtypes A and G cocirculate in some countries such as Nigeria [21]. Mosaic viruses involving CFR02_AG have been reported in several African countries [29] and an increase of subtype F2 and other recombinant forms in Cameroon [30] (Fig. 2).

Europe and North America

In North America [31], Europe [32–36] and Australia [37], subtype B is still the most widespread viral strain. Other subtypes have also been introduced. For example, in Spain, subtype F1 was identified in a cohort of MSM [38] or CRF02_AG among immigrant populations [39].

South America

Subtype B predominates in South America [40,41]. Other intersubtype BF recombinants have also been recently reported [42]. An increase in pure subtype C compared with subtype B has also been reported among HIV/viral hepatitis coinfected patients in Southern Brazil [43].

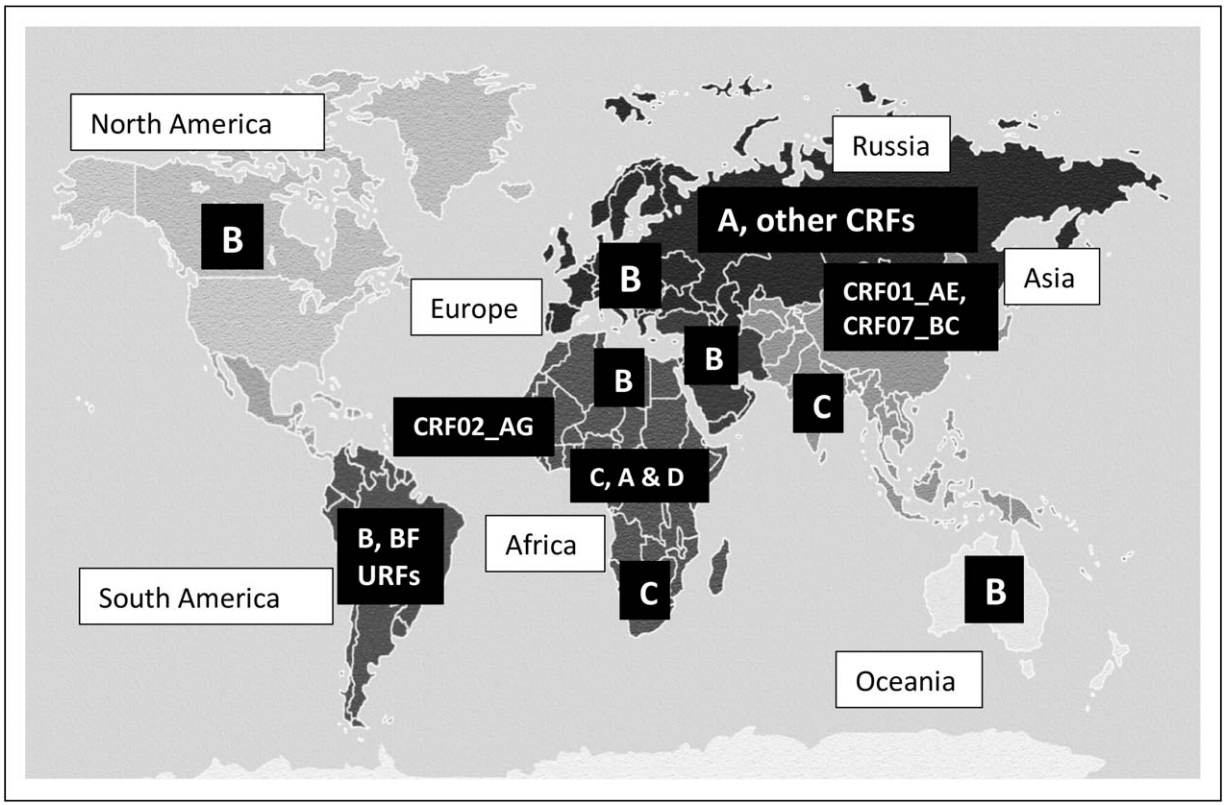

FIGURE 2. Global distribution of major HIV subtypes. Map showing the global distribution of the major HIV subtypes and circulating recombinant forms from the review. *Source:* Map of the world modified to show HIV-1 subtype diversity worldwide. Map-menu.com.

Asia

Several subtypes circulate in Asia and it has been described as the 'hotbed' of recombinant viruses with several CRFs [44]. Subtype A predominates in Russia and other former Soviet Union countries although CRF02_AG is common in Kyrgyzstan [45]. CRF01_AE and CRF07_BC predominate in China [46–48]; new CRFs have been found in China more than in any other country [49–52]. In southeast Asia, CRF01-AE is the most prevalent [53,54] while subtype C predominates in India with reports of increased URFs in the north eastern part of the country [55].

Middle East and North Africa

Middle East and North Africa countries are grouped together because subtype B is the predominant subtype in the region and these countries show relatively low prevalence and limited sampling of infections [56▪,57].

CONCLUSION

Previous reviews have shown that subtype C accounts for more than 50% of all infections worldwide with a concentration in East and Southern Africa and India, CRF01_AE as the major strain throughout Southeast Asia, CRF02_AG as the dominant virus in West Africa and about 30% of strains in East Africa to consist of URFs [13,21]. However, there is increased and extensive HIV-1 genetic diversity in Asia [47,49–52,58] and Eastern Africa [24▪▪,26▪▪]. By 2010, there were about 45 reported CRFs [59], this figure has more than doubled within a 7-year period with at least 98 CRFs that have currently been described [14]. This fast rate of increase in the prevalence of CRFs and the fact that several subtypes cocirculate globally suggests that these numbers will likely continue to grow in the future. The largest source of HIV sequence data (subtype B) is from Europe and the Americas with less data coming from the Middle East (Table 1). As a result, HIV sequence databases like Los Alamos [60] tend to represent a biased sampling and therefore do not necessarily reflect the distribution of subtypes worldwide. For instance, subtype B accounted for 55% of HIV sequences in the Los Alamos database [60] which is about five times its frequency in the world [61]. Information on subtype diversity based on near full-length genomes is still relatively scarce with majority of the available data still based on partial HIV fragments. In East Africa, classification of HIV based on near full-length sequences showed a prevalence

Table 1. Global prevalence and geographical distribution of HIV-1 subtypes and recombinant forms by studies reviewed

Major subtype	Global Prevalence based on the level of geographic spread and available data in the HIV sequence database	Description	Countries	Geographical location	Study (year)	HIV genomic region/available sequences
A	High	Includes intrasubtypes like A1 and A6	Kenya, Pakistan, Armenia, Azerbaijan, Belarus, Estonia, Georgia, Kazakhstan, Kyrgyzstan, Latvia, Lithuania, Moldova, Russia, Ukraine and Uzbekistan	East Africa, Asia, EuroAsia	Gounder *et al.* (2017), Khan *et al.* (2018), Aibekova *et al.* (2018), Lapovok *et al.* (2017), Kazennova *et al.* (2017)	Gag and env ($n=386$); gag, pol and env ($n=755$); pol ($n=3022$), integrase ($n=506$)
B	High	Includes reports of other multiple nonsubtype B virus importations	United states of America, Canada, France, United Kingdom, Belgium, Bulgaria, Cyprus, Spain, Iceland, Mexico, Brazil, Guatemala, Honduras, Panama, Nicaragua, Belize, El Salvador, Amazonas, Australia, Middle East and North Africa that include Algeria, Bahrain, Djibouti, Egypt, Iran, Iraq, Jordan, Kuwait, Lebanon, Libya, Morocco, Palestinian Territories, Oman, Qatar, Saudi Arabia, Somalia, Sudan, Syria, Tunisia, United Arab Emirates, and Yemen	North America, South America, Australia, Europe	Dapp *et al.* (2017), Pérez-Losada *et al.* (2017), Vrancken *et al.* (2017), Dennis *et al.* (2017), Oster *et al.* (2017), Tumiotto *et al.* (2018), Volz *et al.* (2018), Hebberecht *et al.* (2018), Alexiev *et al.* (2018), Yebra *et al.* (2018), Pineda-Peña *et al.* (2018), Pérez-Parra *et al.* (2017), Rojas Sánchez *et al.* (2017), Sallam *et al.* (2017), Hernandez-Sanchez *et al.* (2018), Alves *et al.* (2017), Martin *et al.* (2017), Chaillon *et al.* (2017), Villanova *et al.* (2017), Delatorre *et al.* (2017), Andrade *et al.* (2017), Barral *et al.* (2017), Machado *et al.* (2017), Lima *et al.* (2017), Castley *et al.* (2017), Sallam *et al.* (2017), Daw *et al.* (2017)	Gag ($n=1790$); env ($n=1755$); pol and integrase ($n=63\,917$); pol ($n=23\,196$), Near full length ($n=450$), gag & pol ($n=66$), vif ($n=317$)
C	High	Consists of half of all HIV infections worldwide. Higher prevalence has been found among coinfected patients in Brazil	South Africa, Tanzania, India, Southern Brazil	Southern Africa, East Africa, Asia	Sivay *et al.* (2018), Ngcapu *et al.* (2017), Ngandu *et al.* (2017), Bauer *et al.* (2017), Naidoo *et al.* (2017), Billings *et al.* (2017), Sharma *et al.* (2018), Sharma *et al.* (2017), Avanzi *et al.* (2017)	pol ($n=200$); RNase H ($n=140$); gag ($n=91$); gag-protease ($n=28$); Near full length ($n=170$), vpu ($n=80$); nef ($n=57$); pol/tat-vpu-env ($n=110$); pol and env ($n=38$)
A/D	Moderate	Intersubtype recombinants of subtypes A1 and D	Uganda	East Africa	Lee *et al.* (2017)	Near full length ($n=200$)

Table 1 (*Continued*)

Major subtype	Global Prevalence based on the level of geographic spread and available data in the HIV sequence database	Description	Countries	Geographical location	Study (year)	HIV genomic region/available sequences
F1 and B/F recombinants	Low	Increasing prevalence among MSM in Spain and Argentina	Spain, Argentina	Europe, South America	Patiño-Galindo *et al.* (2017), Cevallos *et al.* (2017)	Near ful length ($n=49$), Near ful length ($n=25$), pol ($n=312$)
CRF01_AE	High	Intersubtype recombinant of subtypes A and E. Previously classified as subtype E	Indonesia, China, Vietnam, Myanmar	Asia	Ueda *et al.* (2018), Li *et al.* (2018), Xu *et al.* (2018), Chen *et al.* (2018), Chikata *et al.* (2018), Yang *et al.* (2018), Zhang *et al.* (2017), Sun *et al.* (2017), Li *et al.* (2017)	pol ($n=837$); Near full length ($n=335$); partial gag, pol and env ($n=638$); gag, pol and nef ($n=388$), env and gag ($n=70$); gag-RT ($n=187$)
CRF02_AG	High	Intersubtype recombinants of subtype A and G although may show more extensive diversity	Ghana, Cameroon	West Africa	Nii-Trebi *et al.* (2017), Nanfack *et al.* (2017), Rodgers *et al.* (2017)	pol ($n=99$); pol & env ($n=555$); NFL ($n=1$)
CRF07_BC	Moderate	Intersubtype recombinants of subtypes B and C	China	Asia	Yang *et al.* (2017), Zhang *et al.* (2017), Wang *et al.* (2017), Li *et al.* (2017)	pol ($n=276$); gag ($n=610$); Near full length ($n=1$); gp160 ($n=749$)
CRF08_BC	Low	Recombinant between subtyes B and C and commonly cocirculate with CRF07_BC and CRF01_AE in parts of Yunnan, China	China	Asia	Chen *et al.* (2018)	gag, pol and env ($n=291$)
CRF35_AD	Low	Recombinant between subtypes A and D	Iran (Tehran)	Middle East	Vahabpour *et al.* (2017)	pol ($n=42$)
Mixed subtypes and new CRFs or URFs	Low	CRF01_AE, CRF02_BC, subtypes A–H, J and K. Also include newly-recommended strains like CRF55_01B, CRF67_01B and CRF69_01B, CRF96_cpx, CRF01/CRF07_BC, CRF07_BC/CRF55_01B, CRF87_cpx/CRF88_BC	Several countries within congo Basin, China, African countries,Canada, South Africa, Uganda and Kenya	Central and West Africa, Asia, N. America, South and East Africa	Tongo *et al.* (2018) and Tongo *et al.* (2018), Fan *et al.* (2018), Miao *et al.* (2018), Lu *et al.* (2018), Kong *et al.* (2017), Li *et al.* (2017), Wu *et al.* (2017), Hu *et al.* (2017), Kiguoya *et al.* (2017)	Near full length ($n=305$); pol ($n=18$); env ($n=48$); gag/pol/env ($n=183$), Gag-protease ($n=1342$)

CRF, circulating recombinant form; NFL, near full-length; RT, reverse transcritpase; URF, unique recombinant form.

of 46% of recombinant viruses in Uganda [26▪▪]. The classification of HIV variants based on near full-length genomes is still relatively low (Table 1) although some consortia have generated data that spans entire HIV genomes [62]. One such consortium is the PANGEA HIV [27▪] which is a multinational project funded by the Bill and Melinda Gates Foundation to study the transmission dynamics of HIV within generalized epidemics in Africa. PANGEA aimed at generating close to 20 000 HIV near full-length sequences mainly from sub-Saharan Africa which have been valuable in characterizing HIV variants and understanding the subtype diversity in several African countries. Preliminary results have reported close to 50% of HIV-1 intersubtype recombinants in Uganda based on near full-length genomes (unpublished data from PANGEA-HIV, [27▪]).

Recent data suggest a significant increase in the prevalence of recombinants in some countries [26▪▪,55,63▪▪] and several newly discovered CRFs/URFs in parts of Asia and Africa [49–52,58,59]. This suggests that coinfection or superinfection by divergent HIV-1 strains has become more common in regions where multiple subtypes cocirculate [26▪▪,42,49,52,64]. Viral sequences from recombinant forms have been shown to affect the accuracy of phylogenetic reconstructions [65,66]. This is due to the different regions of the aligned recombinant genome having distinct evolutionary relationships that affect tree topology and branch lengths [67]. Furthermore, such sequences introduce errors in phylodynamic inferences [68] by biasing ancestral state reconstruction and estimations for the most recent common ancestor [65]. As the prevalence of recombinant forms increases, these challenges are expected to become more important for molecular epidemiological-based investigations. The scale up of cheaper near full-length genome sequencing approaches is critical in improving the accuracy of viral classification. It is important to consider how the fast-increasing emergence of HIV recombinant forms will impact accurate diagnosis, phylogenetic reconstruction, antiretroviral treatment and future vaccine development.

Acknowledgements

None.

Financial support and sponsorship

The work was supported by the UK MRC and the UK Department of International Development (DFID) under the MRC/DFID Concordat agreement. Further support is from a Career Development Fellowship (grant number TMA 2015 CDF-982) to Deogratius Ssemwanga from the European and Developing Countries Clinical Trials Partnership.

Conflicts of interest

There are no conflicts of interest.

REFERENCES AND RECOMMENDED READING

Papers of particular interest, published within the annual period of review, have been highlighted as:

- ▪ of special interest
- ▪▪ of outstanding interest

1. UNAIDS. Global HIV and AIDS statistics-2018 fact sheet [Online] 2018; Available from: http://www.unaids.org/en/resources/fact-sheet/. [Accessed 29 September 2018]
2. Sharp PM, Hahn BH. Origins of HIV and the AIDS pandemic. Cold Spring Harb Perspect Med 2011; 1:a006841.
3. Faria NR, Rambaut A, Suchard MA, *et al.* The early spread and epidemic ignition of HIV-1 in human populations. Science 2014; 346:56–61.
4. Hemelaar J, Gouws E, Ghys PD, Osmanov S. Global and regional distribution of HIV-1 genetic subtypes and recombinants in. AIDS 2006; 20:W13–W23.
5. Visseaux B, Damond F, Matheron S, *et al.* Hiv-2 molecular epidemiology. Infect Genet Evol 2016; 46:233–240.
6. Gilbert PB, McKeague IW, Eisen G, *et al.* Comparison of HIV-1 and HIV-2 infectivity from a prospective cohort study in Senegal. Stat Med 2003; 22:573–593.
7. Kanki PJ, Travers KU, MBoup S, *et al.* Slower heterosexual spread of HIV-2 than HIV-1. Lancet 1994; 343:943–946.
8. Ibe S, Yokomaku Y, Shiino T, *et al.* HIV-2 CRF01_AB: first circulating recombinant form of HIV-2. J Acquir Immune Defic Syndr 2010; 54:241–247.
9. Ng KT, Ong LY, Takebe Y, *et al.* Genome sequence of a novel HIV-1 circulating recombinant form 54_01B from Malaysia. J Virol 2012; 86:11405–11406.
10. Song H, Giorgi EE, Ganusov VV, *et al.* Tracking HIV-1 recombination to resolve its contribution to HIV-1 evolution in natural infection. Nat Commun 2018; 9:1928.
11. Peeters M. Recombinant HIV sequences: Their role in the global epidemic. HIV Sequence Compendium 2000. 2000; 54–72. Available from: https://www.hiv.lanl.gov/content/immunology/pdf/2000/1/Peeters.pdf.
12. Carr J, Foley B, Leitner T, *et al.* Reference sequences representing the principal genetic diversity of HIV-1 in the pandemic. Los Alamos, NM: Los Alamos National Laboratory; 1998.
13. McCutchan FE. Global epidemiology of HIV. J Med Virol 2006; 78(S1):S7–S12.
14. HIV circulating recombinant forms (CRFs) [Online]. Available from: https://www.hiv.lanl.gov/content/sequence/HIV/CRFs/CRFs.html. [Accessed 25 November 2018]
15. Kanki PJ, Peeters M, Guéye-Ndiaye A. Virology of HIV-1 and HIV-2: implications for Africa. AIDS 1997; 11(Suppl B):S33–S42.
16. Ssemwanga D, Lyagoba F, Ndembi N, *et al.* Multiple HIV-1 infections with evidence of recombination in heterosexual partnerships in a low risk Rural Clinical Cohort in Uganda. Virology 2011; 411:113–131.
17. McCutchan FE. Understanding the genetic diversity of HIV-1. AIDS 2000; 14(Suppl 3):S31–S44.
18. Sabino EC, Shpaer EG, Morgado MG, *et al.* Identification of human immunodeficiency virus type 1 envelope genes recombinant between subtypes B and F in two epidemiologically linked individuals from Brazil. J Virol 1994; 68:6340–6346.
19. Marquina S, Leitner T, Rabinovich RD, *et al.* Coexistence of subtypes B, F, and as B/F env recombinant of HIV type 1 in Buenos Aires Argentina. AIDS Res Hum Retroviruses 1996; 12:1651–1654.
20. Lole KS, Bollinger RC, Paranjape RS, *et al.* Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. J Virol 1999; 73:152–160.
21. Lihana RW, Ssemwanga D, Abimiku A, Ndembi N. Update on HIV-1 diversity in Africa: a decade in review. AIDS Rev 2012; 14:83–100.
22. Ngcapu S, Theys K, Libin P, *et al.* Characterization of nucleoside reverse transcriptase inhibitor-associated mutations in the RNAse H region of HIV-1 subtype c infected individuals. Viruses 2017; 9. doi: 10.3390/v9110330.
23. Sivay MV, Hudelson SE, Wang J, *et al.* HIV-1 diversity among young women in rural South Africa: HPTN 068. PLoS One 2018; 13:e0198999.
24. ▪▪ Gounder K, Oyaro M, Padayachi N, *et al.* Complex subtype diversity of HIV-1 among drug user in major Kenyan cities. AIDS Res Hum Retroviruses 2017; 33:500–510.

A study that highlights extensive HIV-1 intersubtype genetic diversity in East Africa and changing dynamics in the prevalence of pure viral subtypes.

25. Billings E, Sanders-Buell E, Bose M, *et al.* HIV-1 genetic diversity among incident infections in Mbeya, Tanzania. AIDS Res Hum Retroviruses 2017; 33:373–381.

26. Lee GQ, Bangsberg DR, Mo T, *et al.* Prevalence and clinical impacts of HIV-1 ■■ intersubtype recombinants in Uganda revealed by near-full-genome population and deep sequencing approaches. AIDS 2017; 31:2345–2354.
The study shows a high prevalence of HIV-1 intersubtype recombinants in Uganda based on deep sequencing and near full length coverage of the viral genome.

27. Pillay D, Herbeck J, Cohen MS, *et al.* PANGEA-HIV: phylogenetics for ■ generalised epidemics in Africa. Lancet Infect Dis 2015; 15:259–261.
The study is a good example of what is needed to generate a large number of HIV near full genomes across multiple sites/countries through a consortium-based approach.

28. Nii-Trebi NI, Brandful JAM, Ibe S, *et al.* Dynamic HIV-1 genetic recombination and genotypic drug resistance among treatment-experienced adults in northern Ghana. J Med Microbiol 2017; 66:1663–1672.

29. Janssens W, Salminen MO, Laukkanen T, *et al.* Near full-length genome analysis of HIV type 1 CRF02. AG subtype C and CRF02. AG subtype G recombinants. AIDS Res Hum Retroviruses 2000; 16:1183–1189.

30. Nanfack AJ, Redd AD, Bimela JS, *et al.* Multimethod longitudinal HIV drug resistance analysis in antiretroviral-therapy-naive patients. J Clin Microbiol 2017; 55:2785–2800.

31. Oster AM, Switzer WM, Hernandez AL, *et al.* Increasing HIV-1 subtype diversity in seven states, United States, 2006–2013. Ann Epidemiol 2017; 27:244–251.e1.

32. Sallam M, Esbjörnsson J, Baldvinsdóttir G, *et al.* Molecular epidemiology of HIV-1 in Iceland: early introductions, transmission dynamics and recent outbreaks among injection drug users. Infect Genet Evol 2017; 49:157–163.

33. Tumiotto C, Bellecave P, Recordon-Pinson P, *et al.* Diversity of HIV-1 in Aquitaine, Southwestern France, 2012–2016. AIDS Res Hum Retroviruses 2018; 34:471–473.

34. Volz EM, Le Vu S, Ratmann O, *et al.* Molecular epidemiology of HIV-1 subtype B reveals heterogeneous transmission risk: implications for intervention and control. J Infect Dis 2018; 217:1522–1529.

35. Hebberecht L, Vancoillie L, Schauvliege M, *et al.* Frequency of occurrence of HIV-1 dual infection in a Belgian MSM population. PLoS One 2018; 13:e0195679.

36. Alexiev I, Lo Presti A, Dimitrova R, *et al.* Origin and spread of HIV-1 subtype B among heterosexual individuals in Bulgaria. AIDS Res Hum Retroviruses 2018; 34:244–253.

37. Castley A, Sawleshwarkar S, Varma R, *et al.* A national study of the molecular epidemiology of HIV-1 in Australia. PLoS One 2017; 12:e0170601.

38. Patiño-Galindo JÁ, Domínguez F, Cuevas MT, *et al.* Genome-scale analysis of evolutionary rate and selection in a fast-expanding Spanish cluster of HIV-1 subtype F1. Infect Genet Evol 2018; 66:43–47.

39. Pérez-Parra S, Chueca N, Álvarez M, *et al.* High prevalence and diversity of HIV-1 non-B genetic forms due to immigration in southern Spain: a phylogeographic approach. PLoS One 2017; 12:e0186928.

40. Lima K, Leal É, Cavalcanti AMS, *et al.* Increase in human immunodeficiency virus 1 diversity and detection of various subtypes and recombinants in North-Eastern Brazil. J Med Microbiol 2017; 66:526–535.

41. Hernandez-Sanchez PG, Guerra-Palomares SE, Ramirez-GarciaLuna JL, *et al.* Prevalence of drug resistance mutations in protease, reverse transcriptase, and integrase genes of North Central Mexico HIV isolates. AIDS Res Hum Retroviruses 2018; 34:498–506.

42. Cevallos CG, Jones LR, Pando MA, *et al.* Genomic characterization and molecular evolution analysis of subtype B and BF recombinant HIV-1 strains among Argentinean men who have sex with men reveal a complex scenario. PLoS One 2017; 12:e0189705.

43. Avanzi VM, Vicente BA, Beloto NCP, *et al.* Profile of HIV subtypes in HIV/HBV- and HIV/HCV-coinfected patients in Southern Brazil. Rev Soc Bras Med Trop 2017; 50:470–477.

44. Saksena NK, Dwyer DE, Wang D. HIV recombination and pathogenesis – biological and epidemiological implications. In: Dumais N, editor. HIV and AIDS – updates on biology, immunology, epidemiology and treatment strategies. London: IntechOpen; 2011. Available from: https://www.intechopen.com/books/hiv-and-aids-updates-on-biology-immunology-epidemiology-and-treatment-strategies/hiv-recombination-and-pathogenesis-biological-and-epidemiological-implications. [Accessed 30 September 2018]; ISBN: 978-953-307-665-2.

45. Aibekova L, Foley B, Hortelano G, *et al.* Molecular epidemiology of HIV-1 subtype A in former Soviet Union countries. PLoS One 2018; 13:e0191891.

46. Xiao P, Li J, Fu G, *et al.* Geographic distribution and temporal trends of HIV-1 subtypes through heterosexual transmission in China: a systematic review and meta-analysis. Int J Environ Res Public Health 2017; 14. doi: 10.3390/ijerph14070830.

47. Wang X, Zhang M, Li J, *et al.* Genetic characterization of a unique recombinant strain identified in Yunnan with genome comprising B and C. AIDS Res Hum Retroviruses 2017; 33:614–620.

48. Chen M, Ma Y, Chen H, *et al.* HIV-1 genetic transmission networks among men who have sex with men in Kunming, China. PLoS One 2018; 13:e0196548.

49. Li K, Ou W, Feng Y, *et al.* Near Full-Length Genomic Characterization of a Novel HIV Type 1 Recombinant Form (CRF01_AE/B) Identified from Anhui, China. AIDS Res Hum Retroviruses 2018; doi: 10.1089/AID.2018.0144. [Epub ahead of print]

50. Wu Y, Ren X, Yin D, *et al.* Characterization of a novel HIV-1 unique recombinant form between CRF07_BC and CRF55_01B in men who have sex with men in Guangzhou, China. PLoS One 2017; 12:e0175770.

51. Miao J, Ran J, Song Y, *et al.* Characterization of a novel HIV-1 circulating recombinant form, CRF01_AE/B'/C (CRF96_cpx), in Yunnan, China. AIDS Res Hum Retroviruses 2018; 34:393–397.

52. Kong D, Wang Y, Wang C, *et al.* Characterization of a new HIV-1 CRF01_AE/CRF07_BC recombinant virus in Guangxi, China. AIDS Res Hum Retroviruses 2017; 33:1166–1170.

53. Ueda S, Witaningrum AM, Khairunisa SQ, *et al.* Genetic Diversity and Drug Resistance of HIV-1 Circulating in North Sulawesi, Indonesia. AIDS Res Hum Retroviruses 2018; doi: 10.1089/AID.2018.0221. [Epub ahead of print]

54. Zhang L, Wang B, Liang Y, *et al.* Phylogenetic characteristics of HIV-1 among travelers entering China from Myanmar: a retrospective study. J Med Virol 2017; 89:1404–1411.

55. Sharma AL, Singh TR, Devi KR, Singh LS. Molecular epidemiology of HIV-1 among the HIV infected people of Manipur, Northeastern India: emergence of unique recombinant forms. J Med Virol 2017; 89:989–999.

56. Sallam M, ahin GÖ, Ingman M, *et al.* Genetic characterization of human ■ immunodeficiency virus type 1 transmission in the Middle East and North Africa. Heliyon 2017; 3:e00352.
The study reports on HIV-1 subtype diversity in the Middle East and North Africa, geographical regions with a relatively lower sampling coverage.

57. Daw MA, El-Bouzedi A, Ahmed MO, Dau AA; The Libyan Study Group of Hepatitis & HIV. Molecular and epidemiological characterization of HIV-1 subtypes among Libyan patients. BMC Res Notes 2017; 10:170.

58. Hu Y, Wan Z, Zhou Y-H, *et al.* Identification of two new HIV-1 circulating recombinant forms (CRF87_cpx and CRF88_BC) from reported unique recombinant forms in Asia. AIDS Res Hum Retroviruses 2017; 33:353–358.

59. Zhao J, Tang S, Ragupathy V, *et al.* Identification and genetic characterization of a novel CRF22_01A1 recombinant form of HIV type 1 in Cameroon. AIDS Res Hum Retroviruses 2010; 26:1033–1045.

60. HIV sequence database. Distribution of all HIV-1 sequences: WORLD [Online]. 2019. Available from https://www.hiv.lanl.gov/components/sequence/HIV/geo/geo.comp [Accessed 15 February 2019].

61. Hemelaar J, Gouws E, Ghys PD, *et al.* Global trends in molecular epidemiology of HIV-1 during. AIDS 2011; 25:679–689.

62. Ratmann O, Wymant C, Colijn C, *et al.* HIV-1 full-genome phylogenetics of generalized epidemics in sub-Saharan Africa: impact of missing nucleotide characters in next-generation sequences. AIDS Res Hum Retroviruses 2017; doi: 10.1089/AID.2017.0061. [Epub ahead of print]

63. Neogi U, Siddik AB, Kalaghatgi P, *et al.* Recent increased identification and ■■ transmission of HIV-1 unique recombinant forms in Sweden. Sci Rep 2017; 7:6371.
The study shows the important role near full length sequencing of HIV genomes plays in identifying recombinant variants and assessing recent trends of the viral diversity.

64. Li F, Li Y, Feng Y, *et al.* Four closely related HIV-1 CRF01_AE/CRF07_BC recombinant forms identified in East China. AIDS Res Hum Retroviruses 2017; 33:740–744.

65. Posada D, Crandall KA. The effect of recombination on the accuracy of phylogeny estimation. J Mol Evol 2002; 54:396–402.

66. Ruths D, Nakhleh L. Recombination and phylogeny: effects and detection. Int J Bioinform Res Appl 2005; 1:202–212.

67. Maljkovic Berry I, Athreya G, Kothari M, *et al.* The evolutionary rate dynamically tracks changes in HIV-1 epidemics: application of a simple method for optimizing the evolutionary rate in phylogenetic trees with longitudinal data. Epidemics 2009; 1:230–239.

68. Frost SDW, Pybus OG, Gog JR, *et al.* Eight challenges in phylodynamic inference. Epidemics 2015; 10:88–92.