Current trends of HIV recombination worldwide

Katherine A. Lau,1 Justin J.L. Wong2
1Retroviral Genetics Division, Centre for Virus Research, Westmead Millennium Institute, Westmead Hospital, The University of Sydney; 2Gene and Stem Cell Therapy Program, Centenary Institute, Royal Prince Alfred Hospital, The University of Sydney, Sydney, Australia

Abstract

One of the major characteristics of HIV-1 is its high genetic variability and extensive heterogeneity. This characteristic is due to its molecular traits, which in turn allows it to vary, recombine, and diversify at a high frequency. As such, it generates complex molecular forms, termed recombinants, which evade the human immune system and so survive. There is no sequence constraint to the recombination pattern as it appears to occur at inter-group (between groups M and O), as well as intra- and intra-subtype within group M. Rapid emergence and active global transmission of HIV-1 recombinants, known as circulating recombinant forms (CRFs) and unique recombinant forms (URFs), requires urgent attention. To date, 55 CRFs have been reported around the world. The first CRF01_AE originated from Central Africa but spread widely in Asia. The most recent CRF, CRF55_01B is a recombinant form of CRF01_AE and subtype B, although its origin is yet to be publicly disclosed. HIV-1 recombination is an ongoing event and plays an indispensable role in HIV epidemics in different regions. Africa, Asia and South America are identified as recombination hot-spots. They are affected by continual emergence and co-circulation of newly emerging CRFs and URFs, which are now responsible for almost 20% of HIV-1 infections worldwide. Better understanding of recombinants is necessary to determine their biological and molecular attributes.

HIV-1 molecular epidemiological trends in worldwide

Since the first clinical observation of human immunodeficiency virus (HIV) in the United States in 1981, it has become a worldwide epidemic. During this spread, HIV started to diversify and undergo molecular evolution. As a result, a plethora of HIV types have been identified, which prompted further genotypic classifications. To date, HIV is divided into HIV-1 and HIV-2.1,2 HIV-1 is driving the global HIV pandemic and is further subdivided into three different groups (M, O and N). In contrast, HIV-2 is less transmissible than HIV-1 and has largely remained confined to West Africa.3 Within HIV-1 group M, nine genetically distinct subtypes are recognised (A-D, F-H, J and K), along with 55 circulating recombinant forms (CRFs), which together are accountable for more than 90% of HIV infections worldwide.4,5 The assignment of other strains of HIV as group M sub-subtypes (A1, A2, A3, F1 and F2) is still pending, since these sub-subtypes are frequently regarded as independent subtypes due to their distinct phylogenetic topology.4 The HIV-1 subtype classification takes into account all genomic regions although it was initially based on gag and/or env sequences. The intra-subtype genetic variation ranges from 15-20% whereas the variation between subtypes is approximately 25-35%, depending on the subtypes and genome regions examined.7 Intra-subtype variation continues to increase over time and epidemic rates may differ according to subtype.8 Although the global availability of sequencing techniques has increased, the classification of HIV-1 strains into subtypes and CRFs remains a complex issue. Definitions of different HIV-1 subtypes are subject to change as new viral sequences are identified.

Despite the complexity of HIV-1 subtype classification, different HIV-1 subtypes have distinct global distribution patterns. Prior to 2005, HIV-1 subtype C or recombinant forms containing at least the env gene of subtype C were responsible for almost all HIV-1 infections worldwide.9,10 This was partly a result of the pandemics in Southern Africa, South America and Asia.9,11,12 However, in recent years, HIV-1 subtype B has become the most predominant strain worldwide, with 60.1% infections (based on the subtype distribution available from the Los Alamos HIV Database; http://www.hiv.lanl.gov/). Other main subtypes such as C, A, D, G, and F have accounted for 13.8, 6.9, 3.7, 1.1, and 1.0% of HIV-1 infections, respectively. In terms of worldwide distribution of the recombinant forms, three most important CRFs (CRF01_AE, CRF_02_AG and CRF07_BC) have contributed towards 4.6, 2.9, and 0.9% of total HIV-1 infections (Figure 1A).

According to the Joint United Nations Programme on HIV and AIDS (UNAIDS), 34 million people was estimated to be living with HIV around the world in 2011. In the same year, there were an estimated 2.5 million new HIV infections. In comparison to 1999, there was a decline of 19% in the number of newly infected individuals with HIV. Sub-Saharan Africa remains the most affected area of the world; being home to 67.6% of all individuals living with HIV and accounted for 72% of Acquired Immunodeficiency Syndrome (AIDS) deaths in 2009. Within the sub-region, South Africa remains the country with the world largest HIV epidemic with 5.6 million of HIV infections and 17.2% of the total AIDS mortality in 2009. In other regions such as Asia, the number of people living with HIV remains stable and is estimated to be around 4.9 million. South and Southeast Asia are predicted to have the highest HIV infection prevalence (estimated to be at 4.1 million in 2009), due to their disparate epidemic trends. Other regions, for instance North America, Central and South America, as well as Eastern Europe and Central Asia, all accounted for 1.5, 1.4 and 1.4 million HIV infections, respectively. Overall, the global growth of AIDS epidemic appears to have stabilised. This trend is attributed to a combination of factors, including the significant scale-up of antiretroviral therapy over the past few years, the impact of HIV prevention efforts and the natural course of HIV epidemics.

HIV recombination

HIV possesses high recombination rates, due to the highly recombinogenic reverse transcriptase (RT) enzyme.12 RT has a high error rate and low binding affinity,13 which are necessary for strand transfers of reverse transcrip-
Recombination is not limited by sequence similarity. It has been demonstrated to occur inter-group (between HIV-1 group M and O),19 and inter- and intra-strain within HIV group M.20 These new recombinants display discrete breakpoints that can be identified between genomic regions with different phylogenetic associations.21,22 Given that the physical sites of recombination are distributed randomly along the viral genome, preferred sites for HIV-1 recombination remain unknown. Recombination has become a common occurrence among different HIV-1 strains, and the inter-subtype recombination is the most frequently observed, although intra-subtype recombination is possible.

At the molecular level, HIV-1 recombination occurs at two different stages: during the synthesis of the minus strand DNA or plus strand DNA using different mechanisms. It is a result of the placement of multiple selection factors, which are summarised into three major processes. The first factor includes the reverse transcription step that generates recombinant genomes, dependent on the mechanism of copy choice. The second factor is the selection for the functional forms, while the last factor is the selection for the replicable forms within the host, which subsequently transmits from one host to another. All these factors are fundamental in shaping the HIV-1 predominance in the population. Thus, it becomes vital to gain further understanding and knowledge into the virus-host interplay.

**HIV recombinants in the epidemic**

The highly unequal geographic distribution of viral variants is the result of global variation in the HIV-1 strains, the dynamic nature of the HIV-1 epidemic, and accidental epidemiologic transmissions. The recombinant HIV-1 strains are emerging at a high frequency due to co-circulation of multiple HIV-1 subtypes in almost all geographic regions of the globe and co- or super-infection of individuals with multiple subtypes. These factors result in HIV CRFs, defined as characteristic full-length or near full-length HIV sequence that are found in three epidemiologically unlinked individuals, or URFs, when these criteria are not met. Novel and newly identified CRFs are named in the order in which they are reported and described, and thus do not represent their historical evolutionary or chronological order. When there are three or more different HIV strains forming a CRF, it is given the extension _cpx_ for complex. CRFs that consist of five or more HIV-1 subtypes, for instance CRF18_cpx, CRF27_cpx, CRF37_cpx, and CRF49_cpx can be true mosaics, as their sequence segments are composed of many different HIV-1 subtypes or other unidentifed sequences.

In many regions around the world, there is a significant displacement of the existing HIV-1 subtype by other new strains, for instance sub-type B and CRF01_AE in Southern China,12,13 Southeast Asia,21,22 and subtype B and F in Brazil.24-25 It is likely that the initial predominance of a certain HIV subtype is attributable to its specific mode of transmission within a given geographical region, for instance displacement of subtype C by subtype B.12 However, there is no substantial evidence to suggest that HIV-1 strains such as CRF01_AE and subtypes A, B, C and D are more or less transmissible by a specific route, in a specific ethnic group, or in a specific cell type.33-39 As the prevalence of variants of HIV-1 recombinant viruses, either CRFs or URFs are becoming more frequent in certain regions around the world, subtype displacement is no longer due to the emergence of another pure HIV-1 subtype. HIV-1 recombinants are most prevalent in areas where multiple subtypes co-circulate and therefore play a major role in the global AIDS epidemic. To date, 55 CRFs have been identified based on the Los Alamos HIV Database, and the global pandemic of HIV-1 is likely to be affected by widespread of the

---

**Figure 1. Worldwide distribution of HIV-1 recombinant forms. (A) Frequency of recombinant forms that contributed towards total HIV infections as of the year 2012. (B) Frequency of different recombinant forms that contributed towards 55 CRFs identified to date.**

---

[page 16] [Infectious Disease Reports 2013; 5:s1e4]
recombinant forms. Their incidence in global epidemics has been estimated around 18-20%, as reported by Hemelaar et al. and Buonaguro et al. in 2007. There is a gradual replacement and phasing-out of the initial predominant HIV-1 strains (the pure subtypes) by the increasingly-epidemiologically important CRFs, as evident from the epidemics in some countries or regions, for instance Malaysia, Brazil, and West Africa.

Identifications of the initial CRFs, namely CRF01_AE and CRF02_AG were controversial due to the constraint of HIV sequence availability and the difficulty in determining recombination breakpoints. CRF01_AE is defined as a HIV recombinant with insertion of subtype E segment at the env region within the backbone of subtype A. Some have suggested that CRF01_AE is in fact subtype E, due to the fact that pure subtype E lineage has never been found. However, it is possible that CRF01_AE is a result of an early event of HIV-1 recombination, which is certainly plausible owing to frequent incidence of unique and unclassifiable viral sequences found in West and Central Africa. Likewise, CRF02_AG was proposed to be a pure subtype, although it was later confirmed as a CRF composed of subtype A and G.

CRF01_AE and CRF02_AG were found to form and emerge at the early stage of HIV-1 epidemic in Central Africa. Subsequently, they spread and expanded to other regions and thus play a crucial role in bringing changes in the regional and global HIV-1 epidemic trend. Today, CRF01_AE has become the most predominant CRF (Figure 1A). It is circulating mostly in Southeast Asia and is responsible for almost 5% of the total HIV-1 infection worldwide. It has also recombined with many other HIV-1 strains and contributes to a total of 9% of all currently identified CRFs, as documented in the Los Alamos HIV Database. The most prevalent pure HIV-1 subtype known to recombine with CRF01_AE is subtype B, which resulted in CRF15_01B, CRF33_01B, CRF34_01B, CRF48_01B, CRF51_01B, CRF32_01B, CRF53_01B, CRF54_01B, and CRF55_01B. There is also an inter-CRF recombination between CRF01_AE and CRF02_AG, as evident in CRF36_cpx and CRF37_cpx. Both of them are complex CRFs and consist of at least four HIV-1 strains. At present, CRF02_AG accounts for 2.9% of total global HIV-1 infections and has a relevant epidemiological prevalence. It is not confined within a geographical location and has been reported to co-circulate with many other HIV-1 strains within the same region. However, CRF02_AG is only found in 3% of all 55 CRFs that have been reported in the Los Alamos HIV Database. This could be explained by its better capacity to prevent its virion from recombining with another HIV-1 strains, in comparison to CRF01_AE. The factor(s) underlying the ability of HIV-1 strains to recombine is an interesting topic to explore further, in order to understand molecular and biological factors that are involved within certain HIV-1 strains to undergo inter-strain recombination.

Based on the Los Alamos HIV Database, HIV-1 subtype B is the strain that has the highest recombination frequency (19%) to form novel CRF, in comparison to other HIV-1 group M viruses (Figure 1B). This is strongly indicated by its predominance worldwide. Surprisingly, subtype C, which accounts for 13.8% of total global HIV-1 infections, only have a recombination occurrence of 4% of the total 55 CRFs. On the other hand, subtype G which only accounts for 1.1% HIV-1 infection worldwide contributes up to 12% of total CRFs identified so far. The ability of subtype G to combine could be due to the formation of heterozygous virions or functional chimeras with another HIV-1 strain to form a new and novel CRF. This also provides evidence for the possible generation of new viruses that possess biological properties distinct to parental strains, which may in turn lead to new epidemiological trend.

**Distribution of HIV recombinant forms**

HIV-1 recombination usually takes place in regions where different subtypes and CRFs co-circulate. These regions are named recombi-

---

**Figure 2.** Circulating recombinant forms (CRFs) of HIV-1 identified globally in the past 20 years. (A) Cumulative number of CRFs reported yearly in major geographical regions (Africa, Asia, South America) and globally (Overall). (B) Cumulative percentage of novel HIV-1 CRFs discovered from 1993 to 2012. In all geographical regions and overall, marked increase in the percentage of newly discovered CRFs was observed in the last 10 years (years 2003-2012). Data were extracted from the Los Alamos HIV Database. CRF5501B was excluded from this analysis since data related to this CRF is not publicly available.

[Infectious Disease Reports 2013; 5:s1e4]
nation hot-spots. Recombination is an ongoing process while HIV-1 recombinants continue to evolve, emerge and disseminate into world population in an aggressive manner, to give rise to new CRF or URF.

New and novel URFs are highly prevalent, and are markedly evident in Africa, South America, Cuba, China, and Southeast Asia. They can continue to spread in the population, and lead to the emergence of new CRFs. This phenomenon has been demonstrated by the emergence of several CRFs that are derived from unique HIV-1 subtype B and F-like URFs in South America, Africa, Central Africa, and are markedly evident in Africa, Asia, and South America. New and novel CRFs are highly prevalent, and are markedly evident in Africa, South America, Cuba, China, and Southeast Asia. They can continue to spread in the population, and lead to the emergence of new CRFs. This phenomenon has been demonstrated by the emergence of several CRFs that are derived from unique HIV-1 subtype B and F-like URFs in South America, Africa, Central Africa, and are markedly evident in Africa, Asia, and South America. New and novel CRFs are highly prevalent, and are markedly evident in Africa, South America, Cuba, China, and Southeast Asia. They can continue to spread in the population, and lead to the emergence of new CRFs. This phenomenon has been demonstrated by the emergence of several CRFs that are derived from unique HIV-1 subtype B and F-like URFs in South America, Africa, Central Africa, and are markedly evident in Africa, Asia, and South America. New and novel CRFs are highly prevalent, and are markedly evident in Africa, South America, Cuba, China, and Southeast Asia. They can continue to spread in the population, and lead to the emergence of new CRFs. This phenomenon has been demonstrated by the emergence of several CRFs that are derived from unique HIV-1 subtype B and F-like URFs in South America, Africa, Central Africa, and are markedly evident in Africa, Asia, and South America. New and novel CRFs are highly prevalent, and are markedly evident in Africa, South America, Cuba, China, and Southeast Asia. They can continue to spread in the population, and lead to the emergence of new CRFs. This phenomenon has been demonstrated by the emergence of several CRFs that are derived from unique HIV-1 subtype B and F-like URFs in South America, Africa, Central Africa, and are markedly evident in Africa, Asia, and South America. New and novel CRFs are highly prevalent, and are markedly evident in Africa, South America, Cuba, China, and Southeast Asia. They can continue to spread in the population, and lead to the emergence of new CRFs. This phenomenon has been demonstrated by the emergence of several CRFs that are derived from unique HIV-1 subtype B and F-like URFs in South America, Africa, Central Africa, and are markedly evident in Africa, Asia, and South America. New and novel CRFs are highly prevalent, and are markedly evident in Africa, South America, Cuba, China, and Southeast Asia. They can continue to spread in the population, and lead to the emergence of new CRFs. This phenomenon has been demonstrated by the emergence of several CRFs that are derived from unique HIV-1 subtype B and F-like URFs in South America, Africa, Central Africa, and are markedly evident in Africa, Asia, and South America. New and novel CRFs are highly prevalent, and are markedly evident in Africa, South America, Cuba, China, and Southeast Asia. They can continue to spread in the population, and lead to the emergence of new CRFs. This phenomenon has been demonstrated by the emergence of several CRFs that are derived from unique HIV-1 subtype B and F-like URFs in South America, Africa, Central Africa, and are markedly evident in Africa, Asia, and South America. New and novel CRFs are highly prevalent, and are markedly evident in Africa, South America, Cuba, China, and Southeast Asia. They can continue to spread in the population, and lead to the emergence of new CRFs. This phenomenon has been demonstrated by the emergence of several CRFs that are derived from unique HIV-1 subtype B and F-like URFs in South America, Africa, Central Africa, and are markedly evident in Africa, Asia, and South America. New and novel CRFs are highly prevalent, and are markedly evident in Africa, South America, Cuba, China, and Southeast Asia. They can continue to spread in the population, and lead to the emergence of new CRFs. This phenomenon has been demonstrated by the emergence of several CRFs that are derived from unique HIV-1 subtype B and F-like URFs in South America, Africa, Central Africa, and are markedly evident in Africa, Asia, and South America. New and novel CRFs are highly prevalent, and are markedly evident in Africa, South America, Cuba, China, and Southeast Asia. They can continue to spread in the population, and lead to the emergence of new CRFs. This phenomenon has been demonstrated by the emergence of several CRFs that are derived from unique HIV-1 subtype B and F-like URFs in South America, Africa, Central Africa, and are markedly evident in Africa, Asia, and South America. New and novel CRFs are highly prevalent, and are markedly evident in Africa, South America, Cuba, China, and Southeast Asia. They can continue to spread in the population, and lead to the emergence of new CRFs. This phenomenon has been demonstrated by the emergence of several CRFs that are derived from unique HIV-1 subtype B and F-like URFs in South America, Africa, Central Africa, and are markedly evident in Africa, Asia, and South America. New and novel CRFs are highly prevalent, and are markedly evident in Africa, South America, Cuba, China, and Southeast Asia. They can continue to spread in the population, and lead to the emergence of new CRFs. This phenomenon has been demonstrated by the emergence of several CRFs that are derived from unique HIV-1 subtype B and F-like URFs in South America, Africa, Central Africa, and are markedly evident in Africa, Asia, and South America. New and novel CRFs are highly prevalent, and are markedly evident in Africa, South America, Cuba, China, and Southeast Asia. They can continue to spread in the population, and lead to the emergence of new CRFs. This phenomenon has been demonstrated by the emergence of several CRFs that are derived from unique HIV-1 subtype B and F-like URFs in South America, Africa, Central Africa, and are markedly evident in Africa, Asia, and South America. New and novel CRFs are highly prevalent, and are markedly evident in Africa, South America, Cuba, China, and Southeast Asia. They can continue to spread in the population, and lead to the emergence of new CRFs. This phenomenon has been demonstrated by the emergence of several CRFs that are derived from unique HIV-1 subtype B and F-like URFs in South America, Africa, Central Africa, and are markedly evident in Africa, Asia, and South America. New and novel CRFs are highly prevalent, and are markedly evident in Africa, South America, Cuba, China, and Southeast Asia. They can continue to spread in the population, and lead to the emergence of new CRFs. This phenomenon has been demonstrated by the emergence of several CRFs that are derived from unique HIV-1 subtype B and F-like URFs in South America, Africa, Central Africa, and are markedly evident in Africa, Asia, and South America. New and novel CRFs are highly prevalent, and are markedly evident in Africa, South America, Cuba, China, and Southeast Asia. They can continue to spread in the population, and lead to the emergence of new CRFs. This phenomenon has been demonstrated by the emergence of several CRFs that are derived from unique HIV-1 subtype B and F-like URFs in South America, Africa, Central Africa, and are markedly evident in Africa, Asia, and South America. New and novel CRFs are highly prevalent, and are markedly evident in Africa, South America, Cuba, China, and Southeast Asia. They can continue to spread in the population, and lead to the emergence of new CRFs. This phenomenon has been demonstrated by the emergence of several CRFs that are derived from unique HIV-1 subtype B and F-like URFs in South America, Africa, Central Africa, and are markedly evident in Africa, Asia, and South America. New and novel CRFs are highly prevalent, and are markedly evident in Africa, South America, Cuba, China, and Southeast Asia. They can continue to spread in the population, and lead to the emergence of new CRFs. This phenomenon has been demonstrated by the emergence of several CRFs that are derived from unique HIV-1 subtype B and F-like URFs in South America, Africa, Central Africa, and are markedly evident in Africa, Asia, and South America. New and novel CRFs are highly prevalent, and are markedly evident in Africa, South America, Cuba, China, and Southeast Asia. They can continue to spread in the population, and lead to the emergence of new CRFs. This phenomenon has been demonstrated by the emergence of several CRFs that are derived from unique HIV-1 subtype B and F-like URFs in South America, Africa, Central Africa, and are markedly evident in Africa, Asia, and South America. New and novel CRFs are highly prevalent, and are markedly evident in Africa, South America, Cuba, China, and Southeast Asia. They can continue to spread in the population, and lead to the emergence of new CRFs. This phenomenon has been demonstrated by the emergence of several CRFs that are derived from unique HIV-1 subtype B and F-like URFs in South America, Africa, Central Africa, and are markedly evident in Africa, Asia, and South America. New and novel CRFs are highly prevalent, and are markedly evident in Africa, South America, Cuba, China, and Southeast Asia. They can continue to spread in the population, and lead to the emergence of new CRFs. This phenomenon has been demonstrated by the emergence of several CRFs that are derived from unique HIV-1 subtype B and F-like URFs in South America, Africa, Central Africa, and are markedly evident in Africa, Asia, and South America. New and novel CRFs are highly prevalent, and are markedly evident in Africa, South America, Cuba, China, and Southeast Asia. They can continue to spread in the population, and lead to the emergence of new CRFs. This phenomenon has been demonstrated by the emergence of several CRFs that are derived from unique HIV-1 subtype B and F-like URFs in South America, Africa, Central Africa, and are markedly evident in Africa, Asia, and South America. New and novel CRFs are highly prevalent, and are markedly evident in Africa, South America, Cuba, China, and Southeast Asia. They can continue to spread in the population, and lead to the emergence of new CRFs. This phenomenon has been demonstrated by the emergence of several CRFs that are derived from unique HIV-1 subtype B and F-like URFs in South America, Africa, Central Africa, and are markedly evident in Africa, Asia, and South America. New and novel CRFs are highly prevalent, and are markedly evident in Africa, South America, Cuba, China, and Southeast Asia. They can continue to spread in the population, and lead to the emergence of new CRFs. This phenomenon has been demonstrated by the emergence of several CRFs that are derived from unique HIV-1 subtype B and F-like URFs in South America, Africa, Central Africa, and are marked...
30. Guimaraes ML, Eyer-Silva WA, Couto-