HIV Diversity, Molecular Epidemiology, and the Role of Recombination

Gustavo H. Kijak, PharmD, PhD, and Francine E. McCutchan, PhD*

Address
*US Military HIV Research Program, Henry M. Jackson Foundation, 1600 East Gude Drive, Rockville, MD 20850, USA.
E-mail: gkijak@hivresearch.org

Current Infectious Disease Reports 2005, **7:**480–488
Current Science Inc. ISSN 1523-3847

The magnitude of the HIV pandemic and its extensive genetic variation may earn it a unique place among infectious agents. A high mutation rate and a rampant recombination are driving HIV's evolution. Nine subtypes and a variety of recombinant forms of HIV now exist. The source of recombinant forms is the multiple infection of target cells, which becomes highly significant when individuals become infected with two or more divergent strains. In the current paper, we re-examine the role of dual infection and recombination in the generation of HIV-1 diversity, both in individuals and on a global scale. The current molecular epidemiology of HIV-1 is reviewed, emphasizing the latest reports from regional epidemics.

Introduction

Since the description of the first cases of HIV/AIDS in the early 1980s, the HIV epidemic has reached more than 150 countries on six continents in a continuing, relentless expansion [1•]. More than 40 million people are currently living with HIV—men and women in nearly equal numbers and including as many as one in three adults in some populations. Children are significantly affected through perinatal transmission of the virus and by orphanhood. The pandemic is being addressed by many interventions and treatments, but the ultimate solution to HIV/AIDS will be a globally effective vaccine.

The genetic diversity of HIV is among the main obstacles for vaccine development. The global HIV/AIDS epidemic is distinguished from those of most other infectious diseases by the presence of multiple subtypes, or clades, each with a specific geographic distribution [1•]. A myriad of intersubtype recombinant forms exist, some of which circulate widely and are as important to the pandemic as the subtypes from which they are derived. This genetic and geographic complexity is mirrored in miniature within the infected individual, where a swarm of highly related but nonidentical virus variants make up a "quasispecies," whose many forms can be unevenly distributed among blood and tissue compartments [2]. The genetic variability of HIV also contributes to its capacity to evade host immune responses and gain resistance to antiretroviral drugs [3].

The epidemiology of HIV is constantly changing. Viral evolution is only one contributor. Human travel and migration; cultural, social, and political change; trade and trafficking of goods and people; and human genetics can also influence the global distribution and prevalence of HIV subtypes.

A striking feature of the HIV pandemic is the profusion of intersubtype recombinant forms [1•]. This fact constitutes *prima facie* evidence that individuals can become co-infected with two or more HIV subtypes, because only this can provide the opportunity to generate such recombinants. Cases of co-infection are now being documented with regularity, especially in high-risk populations where multiple HIV subtypes co-circulate. The development of new tools and approaches has permitted a deeper understanding of the role of co-infection in the generation of recombinant forms, both in the individual and at the population level.

HIV Diversity

HIV diversity is the combined result of three forces acting independently: a high replication rate, frequent introduction of mutations, and recombination [2,3]. Within an infected individual who is not receiving antiretroviral therapy, up to 10 billion viral particles can be generated every day. The HIV reverse transcriptase, which copies the viral RNA into a double-stranded DNA copy ready for integration into the host cell genome, is highly error-prone, introducing about one mutation per replication cycle. The net result of the high replication and mutation rates is that every one of the 9200 nucleotides that make up the HIV-1 genome can be mutated every day. This virus exists not as a defined sequence but as a quasispecies where no two molecules may be precisely alike [2,3].

When the pressure to escape host immune responses or antiretroviral drugs comes to bear on HIV, this arsenal of variants is a key defense. It is probable that many escape mutants already pre-exist, albeit in infinitesimal levels, in the chronically infected individual, only to be brought to light by selection [3]. This seemingly limitless potential for HIV to vary sounds an ominous note for the durability of vaccines and therapies.

HIV may even have exploited the very structure of the genetic code to further adapt to selection. Triplet codons that differ in nucleotide sequence but specify the same amino acid are termed "synonymous." Nucleotide substitutions that substitute one synonymous codon for another are generally considered to be silent, in the evolutionary sense, because they do not alter the structure of the encoded proteins, or phenotype. Brenner *et al.* [4] observed that within the reverse transcriptase, HIV strains that shared similar amino acid sequences but encoded them with different synonymous codons could develop antiviral drug resistance at different rates. The concept of "quasi-synonymy" was elaborated by Kijak *et al.* [5], whose main thesis is that synonymous codons do not necessarily lie on identical evolutionary pathways; how an amino acid is encoded can determine which amino substitutions can readily occur. For example, a valine important in drug resistance is encoded quasi-synonymously, either GTG or GTA, in HIV; GTG-containing strains can substitute the resistance-conferring methionine, ATG, with one mutational step, whereas GTG-encoding strains require two steps.

We estimate that a substantial part of the genetic variability that distinguishes HIV subtypes represents quasi-synonymous codon usage, possibly setting the stage for important genetic changes under drug or immune selection and, at the same time, leaving the pathway to reversion open if selective pressure changes. This somewhat heretical concept of selection acting on "silent" mutations, if further validated, may set HIV apart among living organisms for its ability to adapt.

Mutation drives HIV diversity [2], but its contribution may be overshadowed by yet another mechanism. Two copies of the RNA genome of HIV are packaged into each viral particle. Reverse transcriptase, which generates the double-stranded DNA form or provirus, copies the two RNAs alternately as it traverses the genome. The DNA provirus is, thus, a mosaic of the initial RNAs and, after integration into the cellular genome, becomes the template for the RNA packaged into progeny virions. The extent of variation produced through recombination depends on the degree of similarity of the two RNA genomes in the initial viral particle.

This source RNA for viral particles is more complex than first thought. Jung *et al.* [6••], using elegant fluorescent techniques on HIV-1–infected cells, showed that 80% of cells harbored at least two proviruses, and some contained many more. Multiple infections at the cellular level set the stage for copackaging of diverse viral genomes, coming from different proviruses, into progeny viruses, and recombining them at the next replication cycle. More recently, a series of experiments with cleverly constructed fluorescent reporter viruses designed to measure the parameters of recombination in vitro showed that co-infection of cells is quite frequent; the recombinant forms increased as a linear function of co-infection, which, in turn, was an exponential function of the infection rate [7••]. Crossovers per replication cycle ranged from three to 30, depending on the cell type.

The implications of this work are profound. The concept of infection of a cell with a single virus and production of progeny that differ, at most, by a few point mutations, must seemingly be replaced by one of multiple infection, random co-packaging of viral genomes, and generation of multiple recombinants. Moreover, the virus particles produced from multiply infected cells may themselves be mosaic, for example, expressing envelope glycoproteins from diverse strains or even envelope trimers whose component monomers are divergent. Cell tropism, receptor and co-receptor interactions, and susceptibility to antibody-mediated neutralization could be modulated by a viral swarm of immense phenotypic diversity. This potential for pseudo-typing, or dislinkage of the phenotype of viral particles from the genotype they contain, renders them into a plethora of tiny Trojan horses, yet another aspect of HIV's biology that may earn it a unique place in nature.

Multiple HIV Infection In Vivo

Laboratory studies only define HIV's possibilities, not the realities it faces in human populations. Are co-infection and recombination a frequent occurrence in populations and individuals? Molecular epidemiology provides compelling evidence that they are. In populations where multiple HIV subtypes co-circulate, recombinants among these subtypes are commonly found [1•]. New recombinant forms have been detected quite frequently in high-risk groups, such as injecting drug users (IDUs) [8], who are repeatedly exposed to many forms of HIV (sometimes of different subtypes) through needle-sharing, or among commercial sex workers. In some epidemic regions, recombinant forms of HIV constitute more than half of the circulating strains [9•].

Reports of dual-infected individuals have begun to accumulate in the last few years, reflecting a growing awareness of their importance. Infections with two or more subtypes or strains of the same subtype have been documented [10•]. To study dual infection more systematically, many laboratory screening techniques and strategies for characterizing the viral quasispecies in dual-infected individuals have been developed [11]. Populations exposed to multiple subtypes and with a high proportion of recombinant forms, each with a unique structure, were the subject of study. In a high-risk female population in Tanzania where about half of the strains are unique recombinants, dual infections were easily detected, and many have been characterized in detail [12,13••]. Dually, or even triply, infected individuals harbor many different recombinant forms whose proportions seem to vary considerably over time [12]. Over the long clinical course of HIV infection, a dually infected individual could be a continual source of recombinant forms, perhaps transmitting different strains over time as they wax and wane in abundance. The link

between multiple-exposure risk, dual infection, and recombinant strains is strongly forged by this study [9•]. A broader study in the same population, including low and high risk groups, urban and rural populations, showed lower rates of HIV-1 infection, fewer recombinant strains, and less dual infection in rural versus urban settings [9•].

Does the immune response generated by HIV infection fail to prevent re-infection? This question was posed and debated, spurring efforts to determine whether re-infection is actually occurring after full antiviral immunity develops or earlier, soon after initial infection or even at the time of transmission. Superinfection, or the sequential acquisition of HIV strains, has been definitively demonstrated, and the intervals between initial and re-infection range from 3 to 32 months [10•]. In one case, an HIV superinfection was actually caught in the act. Using follow-up every 3 months in a high-risk cohort in Tanzania, this individual was enrolled within days of infection, before the development of a detectable HIV antibody response, and followed for 30 months [13••]. At first, a single HIV strain was present, with little development of diversity. At the 9-month follow-up, another strain and many recombinants derived from the initial and second strain abruptly appeared. Using the most sensitive tests available, no trace of the second strain could be found before 9 months, but from that point forward, multiple strains were consistently present. It would be a mistake to extrapolate too far from a few cases, but it is becoming clear that there is much more to be learned about the effectiveness and breadth of the human immune response to HIV infection.

Dual infection may have additional implications, if the observations by Gottleib *et al.* [14] and by Grobler *et al.* [15] develop into generalities. These investigators demonstrated a link between dual HIV infection and either a higher concentration of HIV in the bloodstream, or a more rapid progression of disease. A recent review of superinfection highlights another detrimental consequence of re-infection, the acquisition of a drug-resistant strain by individuals whose initial, drug-sensitive infection had been responding to therapy [10•,16]. Preventing dual infection may have benefit, both to the individual, in light of these possible adverse outcomes, and to society, by limiting the genetic complexity of HIV strains that must be controlled by therapies and vaccines.

Global Molecular Epidemiology of HIV

This field of study, now more than 20 years old, seeks to describe and explain the origins of the pandemic and its current status in terms of HIV genetic variation and the distribution and prevalence of strains. HIV-1 and HIV-2 are thought to have originated by cross-species transmission from chimpanzees and sooty-mangaby monkeys, respectively [1•,2]. The subsequent fates of the two types of HIV have been quite different: HIV-2 accounts for a small fraction of the pandemic and is largely confined to West Africa, whereas HIV-1 sparked a global pandemic of monumental proportions. Moreover, of the three separate chimpanzee-to-human transmissions that apparently introduced Groups M, N, and O of HIV-1, only Group M strains are of epidemic importance; groups N and O infections are rare and mostly found in Cameroon. Group M HIV-1 has diversified into nine genetic subtypes, namely subtypes A, B, C, D, F, G, H, J, and K, apparently during spread in human populations [1•]. Two strains, initially identified as subtypes E and I, turned out to be recombinants and were later re-classified. Indeed, intersubtype recombinant strains, combining the different subtypes in a variety of mosaic structures, constitute a very substantial fraction of the pandemic. When recombinant strains with identical structure are retrieved from at least three unlinked individuals, they are called circulating recombinant forms (CRFs). So far 17 CRFs have been described, and more are awaiting publication. A very large number of recombinants, with different structures and subtype compositions, have been found but only in single individuals; these are designated unique recombinant forms (URFs).

Because intersubtype recombination is very frequent in HIV-1, genetic characterization of strains is challenging. Complete genome sequencing is the only unequivocal way to classify strains. Since the publication of the first complete genome sequence of HIV-1 in 1985, full-length sequences from viruses infecting almost 900 different individuals have been reported. Table 1 shows the number reported for each HIV clade or recombinant form and their country of origin. The relative frequency of the different strains in the database does not necessarily reflect their abundance in the pandemic, but the six that are at highest prevalence in the world HIV epidemic—subtypes A, B, C, D, CRF01_AE, and CRF02_AG—are highly represented, accounting for 65% of reported sequences. A much larger database of partial sequences of the *pol* gene exists, mostly generated in antiretroviral-drug resistance studies. The recombinogenic nature of HIV precludes the extrapolation of the genotype from a single gene to the whole genome; however, *pol* gene studies, and others based on partial sequences or genotypes derived by other means, have contributed to overall knowledge of the pandemic.

With the exception of Central Africa (where all the subtypes several of the CRFs, and URFs cocirculate), all other parts of the globe have a much smaller and defined genetic diversity [1•]. In much of the world, a single subtype or CRF of HIV-1 dominates the epidemic. In many regions, however, several clades cocirculate, and inter-subtype recombinants can be detected among them, mainly in the form of URFs [8,9•,17–21]. Some of these URFs expand into pre-established social networks, leading to explosive CRF epidemics, especially among IDUs. Thin barriers among risk groups provide the new CRFs access to the general community. Systematic epidemiologic surveillance has permitted direct observation of the rise of CRF epidemics in Russia (CRF03_AB), China (CRF07_BC and CRF08_BC), Spain (CRF14_BG), and Thailand (CRF15_01B) [1•,22].

Table 1. The database of published HIV full-genome sequences: Clades and their countries of origin*

Clades	Number	Country of origin
HIV-1		
Group M		
A1	55	Kenya, Uganda, Tanzania, Uzbekistan, Byelorussia, Ukraine
A2	2	Democratic Republic of Congo, Cyprus
A3	3	Senegal
B	143	North and South America, Western Europe, Australia, Russia, China, Thailand, Yemen
C	230	South Africa, Botswana, Tanzania, India, Ethiopia, Brazil
D	49	Uganda, South Africa, Chad, Cameroon, Kenya, Democratic Republic of Congo
F1	6	Brazil, Democratic Republic of Congo
F2	4	Cameroon
G	9	Nigeria, Cameroon, Senegal, Spain
H	3	Central African Republic
J	2	Democratic Republic of Congo, Cyprus
K	2	Chad, Cameroon
CRFs		
01_AE	57	Thailand, China, Japan, Central African Republic
02_AG	38	Cameroon, Senegal, Djibouti, Nigeria, Ghana, Uzbekistan
03_AB	3	Russia, Byelorussia
04_cpx	3	Greece, Cyprus
05_DF	3	Democratic Republic of Congo
06_cpx	5	Burkina Faso, Mali, Senegal
07_BC	3	China
08_BC	4	China
09_cpx	3	Senegal
10_CD	3	Tanzania
11_cpx	10	Cameroon
12_BF	7	Argentina, Uruguay
13_cpx	3	Cameroon
14_BG	6	Spain
15_01B	4	Thailand
16_A2D	1	Kenya, South Korea, Argentina
18_cpx	3	Cameroon, Cuba
URFs		
A1/C	16	Tanzania, Kenya
A1/D	26	Uganda, Kenya, Tanzania
B/C	10	China, Myanmar
B/F1	35	Argentina, Brazil, Chile
CRF01_AE/B	14	Thailand, Myanmar
Other	65	Multiple countries
Unclassified	4	Chad
Group N	3	Cameroon
Group O	21	Cameroon, Senegal
HIV-2	20	Guinea Bissau, Mali, Ivory Coast, Ghana, Senegal, Gambia

*Only one reported sequence per individual is considered.
CRF—circulating recombinant form; URF—unique recombinant form.

The regional distribution of HIV-1 in the epidemic has been described in considerable detail. The general outlines of this knowledge are reviewed below, highlighting the most recent new information.

Southern Africa

This region is home to a homogeneous clade C epidemic and the highest HIV-1 prevalence in the world [1•,23]. Are the subtype C strains in the region recombining? Grobler *et al.* [15], in a study of female commercial sex workers in South Africa, found 25% to be infected with more than one subtype C strain. Dual infections were associated with a significantly higher set point viral load, which has lead to an accelerated disease progression rate in other studies.

East Africa

This is a mixed-subtype epidemic with clades A, C, and D and their recombinants [1•]. The distribution varies by country, with a predominance of clade A in Kenya, clade C in Tanzania, and clade D in Uganda. Each of these countries also harbors 30% to 50% URFs; however, CRFs are rare. The proportions of the different strains appear stable

over time. In a study of 460 women attending an antenatal clinic in western Kenya, the molecular landscape (mostly subtype A and URFs) remained unchanged between 1996 and 2000 [24].

In neighboring Uganda, a study on drug resistance among antiretroviral-treated patients in Kampala showed equivalent frequencies of subtypes A and D and a slightly higher prevalence of drug resistance subtype D infections. Subtype C infections, although still rare in Uganda, appeared to be more recent [25]. A recent review of 10 years of epidemiologic research on HIV-1 in Uganda reiterates, however, the stability of subtype percentages in the country [26].

In the Mbeya region of southwestern Tanzania, a community-based cohort with over 3,000 participants was studied in preparation for HIV vaccine trials [9•]. The authors used a fluorescent genotyping assay, capable of distinguishing subtypes, recombinants, and dual infections, to study 507 HIV seropositive samples [11]. The predominant clades were subtype C (43%), C-containing recombinants (32%), subtype A (18%), and subtype D (3%). Because of the large sample size, the authors were able to observe subtle complexities in the epidemic: the proportions of subtypes, recombinants, and dual infections differed among genders, sites, and enrollment strategies. Geographic isolation of rural villages and urban population access to the Trans-African highway connecting Tanzania with neighboring countries has probably played an important role in shaping the HIV epidemics in these communities.

Ethiopia has a homogeneous subtype C HIV epidemic, unlike its east African neighbors to the south [1•]. Djibouti, in the Horn of Africa, has a predominance of subtype C (73%) alongside a strain prevalent in West Africa, CRF02_AG (18%) [27].

West and West Central Africa

In West and West Central Africa, CRF02_AG, a recombinant between subtypes A and G with a specific pattern of breakpoints across the genome, is predominant, with a declining prevalence from West to East [1•]. In Ivory Coast, for example, approximately 95% of HIV-1 strains are CRF02_AG. In Cameroon, the CRF02_AG prevalence falls to 50%, and it cocirculates with a very complex mixture of subtype G, CRF01_AE, CRF11_cpx, and CRF13_cpx, rare subtypes, and URFs.

Another dimension of the epidemic has been revealed by studies in Africa, where some of the HIV-1 subtypes have become further subdivided. Sub-subtypes F1 and F2, and A1 and A2, represent distinct lineages within their respective subtypes, each with a distinctive geographic distribution. Recently, a third lineage within subtype A, now termed A3, was described in Senegal [28]. This strain has already spread through much of West and West Central Africa, accounts for an increasing fraction of new infections in Senegal, and has generated multiple A3/CRF02_AG recombinant strains [29].

A particularly high genetic complexity of strains appears to circulate in Cameroon and Nigeria. In the north of Cameroon, CRF02_AG, CRF02_AG-containing recombinants, and the complex recombinant CRF09_cpx have been found [17]. A larger, multicenter study in Cameroon, including nearly 3500 HIV samples, found that the overwhelming majority were Group M HIV-1 [30]. CRF02_AG accounted for more than one half of the specimens, and about 25% were CRF02_AG-containing recombinants. Group O HIV-1 was found in 0.4% of the HIV cases, and showed a remarkably high between-strain genetic diversity. Three cases of Group O/Group M dual infection were documented, leading to recombination in at least one individual. Two persons were dually infected with Group M HIV-1 and HIV-2. A recent report from Nigeria showed a stronger presence of HIV-2 than in Cameroon [31]. Among 420 samples, 4.3% were reactive for HIV-2, mostly in the setting of HIV-1/HIV-2 dual infection. The authors also observed a bias in HIV-2 subtype distribution in single versus dual infections.

Central Africa

This region is considered by many investigators to be the cradle of HIV-1, from which all subtypes and some CRFs radiated to the rest of the world [1•]. The genetic diversity is extensive, with no predominant clade. Moreover, many rare subtypes have been found only here, and some of the sequences from the region remain partly or wholly unclassified. A recent report from Likasi, concerning the southeastern region of the Democratic Republic of Congo, showed that half of the sequences were recombinants, mostly involving subtype A [32]. Even though this region borders with subtype C-dominated Zambia, this clade showed a very small contribution to the epidemic.

Around 250,000 people are estimated to be infected with HIV in Angola, but genetic data from this country has been very scarce. Bartolo *et al.* [33] studied 48 HIV-1 samples and found that subtypes A1, C, and H were the most frequent clade, and that 17% of the strains were recombinants.

North America

The main HIV-1 clade in the United States and Canada is subtype B, but other subtypes are present at low levels, mostly introduced through immigration and travel. In a study in New York City among 53 HIV-1 seropositive individuals that had emigrated from non-European countries, the main risk factor was heterosexual contact, and the clade distribution was concordant with the strains circulating in the countries of origin: CRF02_AG in West and West Central Africa, subtype C in Southern Africa, subtype A in East Africa, CRF01_AE in Thailand and Myanmar, and subtype B in Central America and the Caribbean [34]. The state of Minnesota is home to many African immigrants, and sequences from a group of 87 showed again that the HIV-1 clades fully matched the ones circulating in the countries of origin [35]. United States military personnel

can be exposed to non-subtype B strains during overseas deployment. Among 520 recent HIV-1 seroconverters (1997 to 2000), the majority were infected with subtype B, whereas 5.4% harbored non-B strains reflecting their regions of deployment: CRF01_AE from Southeast Asia and CRF02_AG from West/West Central Africa. Full-genome sequencing of a subset revealed the presence of a unique B/CRF01_AE recombinant from Southeast Asia, a strain clustering with CRF09_cpx West African sequences, and an A/D URF likely proceeding from East Africa [36].

South America and the Caribbean

The Cuban HIV-1 epidemic stands out in Latin America for its high genetic diversity; half of the strains are subtype B and others are recombinants involving subtypes A, D, and H [18]. The recent analysis of full-genomes revealed that two Cuban sequences of very complex recombinant structure mirrored a strain previously reported in Cameroon, and the three were, therefore, designated CRF18_cpx. This new CRF or related strains seem to account for a significant fraction of Cuban HIV-1 cases [37]. In the rest of the Caribbean region, subtype B circulates almost exclusively.

Two different HIV-1 epidemics occur in South America. The Pacific rim of the continent resembles the North American epidemic, with a dominance of subtype B, whereas Brazil and the Southern cone contain a more complex mixture [19]. In Brazil, subtypes B, C, F, and recombinants co-circulate. In Sao Paulo, subtype B accounts for 90% of the infections, and subtype F for 4% [20]. B/F recombinants, some of which share common breakpoints in the *pol* gene but differ in their structure in other genes, were also found, in concordance with previous studies that found several B/F URFs but failed to detect a CRF in Brazil. A subsequent full-genome study, this time on samples from a heterosexual cohort from the nearby port of Santos, revealed two new B/F1 CRFs related by descent [38].

In the Southern cone, B and B/F recombinants are the principal strains [19]. The recombinants include CRF12_BF and many more B/F URFs. Strains seem to segregate according to risk factor; mostly subtype B in men who have sex with men, but many recombinants among heterosexuals and IDUs [39]. Subtype C sequences are beginning to be detected in Argentina, Uruguay, and Paraguay at low levels and also in southern Brazil [40,41].

Western Europe

At the onset of the HIV-1 epidemic in Western Europe, the predominant clade was subtype B, as in North America. Continuing links to Africa through travel and immigration have resulted in an increased genetic complexity. In France, up to 25% of new HIV-1 infections are non-B, mainly CRF02_AG [42]. In Belgium, 50% of prevalent infections are now from non-B subtypes, mostly clades A, C, and G [43]. In selected sexually transmitted infections clinics in England and Wales, subtypes B and C each account for about 30% of the infections among heterosexuals [44]. Women were more likely to be infected with non-B clades. In Spain, the introduction of subtype G in an IDU network, where subtype B circulated, led to the emergence of a novel recombinant that rapidly spread, becoming CRF14_BG.

Eastern Europe and Former Soviet Republics

The political changes during the 1990s in former Soviet Republics and in Eastern Europe destabilized the social milieu. Intravenous drug use and sexually transmitted infections increased, and both fostered the spread of HIV-1. The summer of 1996 saw a 100-fold increase in new HIV-1 infections among IDU in Kaliningrad, a Russian enclave on the Baltic Sea. The epidemic was highly homogeneous, involving a novel A/B recombinant strain, designated CRF03_AB. Now the wider spread of this CRF as well as the parental subtype A strain, also of low diversity, are spreading further among IDUs in many former Soviet Republics [45]. Subtype A has also permeated to the heterosexual population [46].

The wider Baltic region also has a dynamic HIV-1 epidemic, mostly subtype A in Latvia, for example, especially among IDUs [47]. A rare West African strain, CRF06_cpx, was introduced into Estonia and is rapidly replacing the local A strain among IDU there, and multiple, novel A/CRF06_cpx unique recombinants have emerged [21,48].

In Uzbekistan, a survey among IDUs again showed a predominance of subtype A, but 10% of the strains clustered with West/West Central African CRF02_AG [49]. Once more, co-circulation of local and newly introduced strains led to the emergence of A/CRF02_AG recombinants. In Albania, the influence of East European epidemics is also evidenced. A study conducted among patients infected through heterosexual risk showed that half of infections are due to subtype A, and one third are subtype B [50].

Asia and the Pacific

HIV-1 affects more than 5 million people in Asia. In China, India, and Southeast Asia, the main circulating strains are subtypes B, C, and CRF01_AE, respectively, and in each of these areas, the different strains co-circulate in high-risk cohorts, generating multiple inter-subtype recombinants [22]. Most of these recombinants arose among IDUs in the "golden triangle," an opium cultivation hub positioned between Myanmar, Thailand, and Laos and were later introduced into local networks through different drug-trafficking routes. In Thailand, subtypes initially segregated according to risk factors: CRF01_AE through heterosexual transmission and subtype B in IDU. However, the barriers between the groups are permeable, and, in a few years, the epidemics changed. CRF01_AE now accounts for the majority of new infections in both risk groups, and only few new infections are due to subtype B. However, an increasing number of CRF01_AE/B recombinants have been documented, constituting 15% of the strains in IDUs in the Northern province of Chiang Mai [8], and 8% in a

community-based cohort from the Eastern provinces of Rayong and Chon Buri [51]. Even though the majority of these recombinants are URFs, a specific one was found in four individuals with different risk factors and separated by 800 kilometers; this recombinant was designated CRF15_01B. The generation of recombinants seems to be an ongoing process, driven by a significant rate of dual infection [52]. A similar landscape was reported in Malaysia, where 20% of partial genome sequences were B/CRF01_AE recombinants [53].

In Vietnam, the predominant strain is CRF01_AE, as recently shown among IDUs from Hanoi [54]. A group of Vietnamese IDUs living in Melbourne, Australia, had a mixed epidemic, with equal proportions of CRF01_AE and the locally circulating subtype B [55].

The Chinese HIV-1 epidemic, mainly driven through intravenous drug use, is also diverse, with different subtype distributions across the country [22]. The highest HIV-1 heterogeneity lies in the southwestern provinces of Yunnan and Guangxi, where subtypes B, C, CRF01_AE, their recombinants, and also CRF08_BC circulate. Although CRF01_AE predominates in Guangdong, subtype B circulates in the central provinces of Hubei and Henan and in Shaghai. In the northwestern province of Xinjiang another B/C recombinant, CRF07_BC, was found. CRF07_BC and CRF08_BC are descendents of a common recombinant ancestor and spread to different provinces through divergent drug-trafficking routes [22]. A recent report has shown that the HIV-1 epidemic in Hong Kong is due, mainly, to CRF01_AE and subtype B and is seemingly unlinked to mainland China [56].

The majority of the HIV-1 infections in Asia have occurred in India. The predominant circulating strain, both in Mumbai in the West [57] and in Chennai in the South [58], is subtype C. In Manipur, near China and Myanmar, four of 14 specimens were B/C recombinants [59].

Conclusions

This brief review provides only a glimpse of an immensely complex and rapidly changing pandemic. The mechanisms of variation available to HIV-1, coupled with its rapid replication rate in the individual and high prevalence in many populations, have played out on the global scale to generate what is arguably the most variability of any human pathogen. The regional distributions of subtypes and recombinant forms are subject to rapid shifts, both through newly generated strains and through introduction of established strains into new populations, countries, and regions. High risk and multiple exposures are the engine driving dual infection and recombination, and recombinant strains are as important in the pandemic as the subtypes from which they derive. The interplay between host and virus seems to have played out, at least so far, to the advantage of the virus, and it will be the great task of another generation of researchers to tame HIV diversity and turn the tables back to the advantage of humankind.

Disclosure

The views and opinions expressed in the previous article do not necessary reflect those of the United States Army or Department of Defense.

Acknowledgment

Support for this work came from the United States Military HIV Research Program (cooperative agreement W81XWH-04-2-005) and the European Commission. The United States Military HIV Research Program is a Department of Defense tri-service program executed through the Medical Research and Materiel Command, Walter Reed Army Institute of Research, and Henry M. Jackson Foundation, in direct partnership with the Division of AIDS (DAIDS)/National Institute of Allergy and Infectious Diseases/United States Department of Health and Human Services and jointly planned and coordinated through DAIDS leadership.

References and Recommended Reading

Papers of particular interest, published recently,
have been highlighted as:

- • Of importance
- •• Of major importance

1.• McCutchan FE: **Understanding the genetic diversity of HIV-1.** *AIDS* 2000, **14(Suppl 3):**S31–44.

A comprehensive review of global distribution of HIV subtypes and recombinants.

2. Coffin JM, Hughes SH, Varmus HE, eds.: *Retroviruses.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 1997.
3. Coffin JM: **HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy.** *Science* 1995, **267:**483–489.
4. Brenner B, Turner D, Oliveira M, *et al.*: **A V106M mutation in HIV-1 clade C viruses exposed to efavirenz confers cross-resistance to non-nucleoside reverse transcriptase inhibitors.** *AIDS* 2003, **17:**F1–5.
5. Kijak GH, Currier JR, Tovanabutra S, *et al.*: **Lost in translation: implications of HIV-1 codon usage for immune escape and drug resistance.** *AIDS Rev* 2004, **6:**54–60.
6.•• Jung A, Maier R, Vartanian JP, *et al.*: **Multiply infected spleen cells in HIV patients.** *Nature* 2002, **418:**144.

Groundbreaking report on in vivo infection of target cells with multiple HIV-1 proviruses.

7.•• Levy DN, Aldrovandi GM, Kutsch O, Shaw GM: **Dynamics of HIV-1 recombination in its natural target cells.** *Proc Natl Acad Sci U S A* 2004, **101:**4204–4209.

A compilation of elegant in vitro experiments showing the relationship between rates of HIV-1 infection, dual infection, and recombination

8. Tovanabutra S, Beyrer C, Sakkhachornphop S, *et al.*: **The changing molecular epidemiology of HIV type 1 among northern Thai drug users, 1999 to 2002.** *AIDS Res Hum Retroviruses* 2004, **20:**465–475.
9.• Arroyo MA, Hoelscher M, Sateren W, *et al.*: **HIV-1 diversity and prevalence differ between urban and rural areas in the Mbeya region of Tanzania.** *AIDS,* In press.

Cohort analysis from epidemiologic and molecular perspectives showed association between increased risk, HIV-1 dual infection and inter-clade recombination.

10.• Smith DM, Richman DD, Little SJ: **HIV superinfection.** *J Infect Dis* 2005, **192:**438–444.

Up-to-date review of HIV-1 superinfection and its clinical implications.

11. Hoelscher M, Dowling WE, Sanders-Buell E, *et al.*: **Detection of HIV-1 subtypes, recombinants, and dual infections in east Africa by a multi-region hybridization assay.** *AIDS* 2002, **16:**2055–2064.
12. Gerhardt M, Mloka D, Tovanabutra S, *et al.*: **In-depth, longitudinal analysis of viral quasispecies from an individual triply infected with late-stage human immunodeficiency virus type 1, using a multiple PCR primer approach.** *J Virol* 2005, **79:**8249–8261.
13.•• FE, Hoelscher M, Tovanabutra S, *et al.*: **First in-depth analysis of a heterosexually acquired HIV-1 superinfection: evolution, temporal fluctuation, and inter-compartment dynamics from the window period through 30 months post infection.** *J Virol,* In press.

In-depth molecular characterization of HIV-1 within-patient evolution during superinfection

14. Gottlieb GS, Nickle DC, Jensen MA, *et al.*: **Dual HIV-1 infection associated with rapid disease progression.** *Lancet* 2004, **363:**619–622.
15. Grobler J, Gray CM, Rademeyer C, *et al.*: **Incidence of HIV-1 dual infection and its association with increased viral load set point in a cohort of HIV-1 subtype C-infected female sex workers.** *J Infect Dis* 2004, **190:**1355–1359.
16. Yang OO, Daar ES, Jamieson BD, *et al.*: **Human immunodeficiency virus type 1 clade B superinfection: evidence for differential immune containment of distinct clade B strains.** *J Virol* 2005, **79:**860–868.
17. Burda ST, Konings FA, Williams CA, *et al.*: **HIV-1 CRF09_cpx circulates in the North West Province of Cameroon where CRF02_AG infections predominate and recombinant strains are common.** *AIDS Res Hum Retroviruses* 2004, **20:**1358–1363.
18. Cuevas MT, Ruibal I, Villahermosa ML, *et al.*: **High HIV-1 genetic diversity in Cuba.** *AIDS* 2002, **16:**1643–1653.
19. Hierholzer J, Montano S, Hoelscher M, *et al.*: **Molecular epidemiology of HIV type 1 in Ecuador, Peru, Bolivia, Uruguay, and Argentina.** *AIDS Res Hum Retroviruses* 2002, **18:**1339–1350.
20. Sa Filho DJ, Sanabani S, Diaz RS, *et al.*: **Analysis of full-length human immunodeficiency virus type 1 genome reveals a variable spectrum of subtypes B and F recombinants in Sao Paulo, Brazil.** *AIDS Res Hum Retroviruses* 2005, **21:**145–151.
21. Zetterberg V, Ustina V, Liitsola K, *et al.*: **Two viral strains and a possible novel recombinant are responsible for the explosive injecting drug use-associated HIV type 1 epidemic in Estonia.** *AIDS Res Hum Retroviruses* 2004, **20:**1148–1156.
22. Tovanabutra S, Birx DL, McCutchan FE: *Molecular Epidemiology of HIV in Asia and the Pacific.* New York, New York: Kluwer Academic; 2004.
23. Bellocchi MC, Forbici F, Palombi L, *et al.*: **Subtype analysis and mutations to antiviral drugs in HIV-1-infected patients from Mozambique before initiation of antiretroviral therapy: results from the DREAM program.** *J Med Virol* 2005, **76:**452–458.
24. Yang C, Li M, Shi YP, *et al.*: **Genetic diversity and high proportion of intersubtype recombinants among HIV type 1-infected pregnant women in Kisumu, western Kenya.** *AIDS Res Hum Retroviruses* 2004, **20:**565–574.
25. Richard N, Juntilla M, Abraha A, *et al.*: **High prevalence of antiretroviral resistance in treated Ugandans infected with non-subtype B human immunodeficiency virus type 1.** *AIDS Res Hum Retroviruses* 2004, **20:**355–364.
26. Yirrell DL, Kaleebu P, Morgan D, *et al.*: **HIV-1 subtype dynamics over 10 years in a rural Ugandan cohort.** *Int J STD AIDS* 2004, **15:**103–106.
27. Maslin J, Rogier C, Berger F, *et al.*: **Epidemiology and genetic characterization of HIV-1 isolates in the general population of Djibouti (Horn of Africa).** *J Acquir Immune Defic Syndr* 2005, **39:**129–132.
28. Meloni ST, Kim B, Sankale JL, *et al.*: **Distinct human immunodeficiency virus type 1 subtype A virus circulating in West Africa: sub-subtype A3.** *J Virol* 2004, **78:**12438–12445.
29. Meloni ST, Sankale JL, Hamel DJ, *et al.*: **Molecular epidemiology of human immunodeficiency virus type 1 sub-subtype A3 in Senegal from 1988 to 2001.** *J Virol* 2004, **78:**12455–12461.
30. Yamaguchi J, Bodelle P, Vallari AS, *et al.*: **HIV infections in northwestern Cameroon: identification of HIV type 1 group O and dual HIV type 1 group M and group O infections.** *AIDS Res Hum Retroviruses* 2004, **20:**944–957.
31. Zeh C, Pieniazek D, Agwale SM, *et al.*: **Nigerian HIV type 2 subtype A and B from heterotypic HIV type 1 and HIV type 2 or monotypic HIV type 2 infections.** *AIDS Res Hum Retroviruses* 2005, **21:**17–27.
32. Kita K, Ndembi N, Ekwalanga M, *et al.*: **Genetic diversity of HIV type 1 in Likasi, southeast of the Democratic Republic of Congo.** *AIDS Res Hum Retroviruses* 2004, **20:**1352–1357.
33. Bartolo I, Epalanga M, Bartolomeu J, *et al.*: **High genetic diversity of human immunodeficiency virus type 1 in Angola.** *AIDS Res Hum Retroviruses* 2005, **21:**306–310.
34. Achkar JM, Burda ST, Konings FA, *et al.*: **Infection with HIV type 1 group M non-B subtypes in individuals living in New York City.** *J Acquir Immune Defic Syndr* 2004, **36:**835–844.
35. Sides TL, Akinsete O, Henry K, *et al.*: **HIV-1 subtype diversity in Minnesota.** *J Infect Dis* 2005, **192:**37–45.
36. Tovanabutra S, Brodine SK, Mascola JR, *et al.*: **Characterization of complete HIV type 1 genomes from non-B subtype infections in U.S. military personnel.** *AIDS Res Hum Retroviruses* 2005, **21:**424–429.
37. Thomson MM, Casado G, Posada D, *et al.*: **Identification of a novel HIV-1 complex circulating recombinant form (CRF18_cpx) of Central African origin in Cuba.** *AIDS* 2005, **19:**1155–1163.
38. Sa Filho DJ, Sucupira MCA, Sabino EC, *et al.*: **Identification of two HIV-1 circulating recombinant forms in Brazil brings more complexity to the South American epidemic.** *AIDS Res Hum Retroviruses,* In press.
39. Avila MM, Pando MA, Carrion G, *et al.*: **Two HIV-1 epidemics in Argentina: different genetic subtypes associated with different risk groups.** *J Acquir Immune Defic Syndr* 2002, **29:**422–426.
40. Carrion G, Eyzaguirre L, Montano SM, *et al.*: **Documentation of subtype C HIV Type 1 strains in Argentina, Paraguay, and Uruguay.** *AIDS Res Hum Retroviruses* 2004, **20:**1022–1025.
41. Aulicino PC, Kopka J, Mangano AM, *et al.*: **Circulation of novel HIV type 1 A, B/C, and F subtypes in Argentina.** *AIDS Res Hum Retroviruses* 2005, **21:**158–164.
42. Descamps D, Chaix ML, Andre P, *et al.*: **French national sentinel survey of antiretroviral drug resistance in patients with HIV-1 primary infection and in antiretroviral-naive chronically infected patients in 2001–2002.** *J Acquir Immune Defic Syndr* 2005, **38:**545–552.
43. Snoeck J, Van Laethem K, Hermans P, *et al.*: **Rising prevalence of HIV-1 non-B subtypes in Belgium: 1983–2001.** *J Acquir Immune Defic Syndr* 2004, **35:**279–285.
44. Tatt ID, Barlow KL, Clewley JP, *et al.*: **Surveillance of HIV-1 subtypes among heterosexuals in England and Wales, 1997–2000.** *J Acquir Immune Defic Syndr* 2004, **36:**1092–1099.
45. Roudinskii NI, Sukhanova AL, Kazennova EV, *et al.*: **Diversity of human immunodeficiency virus type 1 subtype A and CRF03_AB protease in Eastern Europe: selection of the V77I variant and its rapid spread in injecting drug user populations.** *J Virol* 2004, **78:**11276–11287.
46. Bobkov AF, Kazennova EV, Selimova LM, *et al.*: **Temporal trends in the HIV-1 epidemic in Russia: predominance of subtype A.** *J Med Virol* 2004, **74:**191–196.
47. Balode D, Ferdats A, Dievberna I, *et al.*: **Rapid epidemic spread of HIV type 1 subtype A1 among intravenous drug users in Latvia and slower spread of subtype B among other risk groups.** *AIDS Res Hum Retroviruses* 2004, **20:**245–249.
48. Adojaan M, Kivisild T, Mannik A, *et al.*: **Predominance of a rare type of HIV-1 in Estonia.** *J Acquir Immune Defic Syndr* 2005, **39:**598–605.
49. Carr JK, Nadai Y, Eyzaguirre L, *et al.*: **Outbreak of a West African recombinant of HIV-1 in Tashkent, Uzbekistan.** *J Acquir Immune Defic Syndr* 2005, **39:**570–575.
50. Ciccozzi M, Gori C, Boros S, *et al.*: **Molecular diversity of HIV in Albania.** *J Infect Dis* 2005, **192:**475–479.

51. Watanaveeradej V, Carr JK, Hoelscher M, *et al.*: **HIV genetic diversity in the communities prepared for phase III vaccine trial in Thailand and a new essential assay to study HIV genetic variety in Asia.** In press.
52. Hu DJ, Subbarao S, Vanichseni S, *et al.*: **Frequency of HIV-1 dual subtype infections, including intersubtype superinfections, among injection drug users in Bangkok, Thailand.** *AIDS* 2005, **19:**303–308.
53. Tee KK, Pon CK, Kamarulzaman A, Ng KP: **Emergence of HIV-1 CRF01_AE/B unique recombinant forms in Kuala Lumpur, Malaysia.** *AIDS* 2005, **19:**119–126.
54. Tran TT, Maljkovic I, Swartling S, *et al.*: **HIV-1 CRF01_AE in intravenous drug users in Hanoi, Vietnam.** *AIDS Res Hum Retroviruses* 2004, **20:**341–345.
55. Ryan CE, Elliott JH, Middleton T, *et al.*: **The molecular epidemiology of HIV type 1 among Vietnamese Australian injecting drug users in Melbourne, Australia.** *AIDS Res Hum Retroviruses* 2004, **20:**1364–1367.
56. Lim WL, Xing H, Wong KH, *et al.*: **The lack of epidemiological link between the HIV type 1 infections in Hong Kong and Mainland China.** *AIDS Res Hum Retroviruses* 2004, **20:**259–262.
57. Deshpande A, Recordon-Pinson P, Deshmukh R, *et al.*: **Molecular characterization of HIV type 1 isolates from untreated patients of Mumbai (Bombay), India, and detection of rare resistance mutations.** *AIDS Res Hum Retroviruses* 2004, **20:**1032–1035.
58. Balakrishnan P, Kumarasamy N, Kantor R, *et al.*: **HIV type 1 genotypic variation in an antiretroviral treatment-naive population in southern India.** *AIDS Res Hum Retroviruses* 2005, **21:**301–305.
59. Tripathy SP, Kulkarni SS, Jadhav SD, *et al.*: **Subtype B and subtype C HIV type 1 recombinants in the northeastern state of Manipur, India.** *AIDS Res Hum Retroviruses* 2005, **21:**152–157.