Genetic diversity of HIV-1: the moving target

Martine Peetersa and Paul M. Sharpb

AIDS 2000, 14 (suppl 3):S129-S140

Keywords: HIV-1, genetic subtypes, circulating recombinant forms, recombination, classification, vaccines, epidemiology, superinfection

Introduction

One of the major characteristics of HIV is the extremely high genetic variability. This extensive heterogeneity is the result of two factors. First, there have been multiple introductions of genetically diverse simian viruses into humans [1]. Numerous African primates have now been found to be infected with SIV, and the two major viral types infecting humans, HIV-1 and HIV-2, represent zoonotic transmissions from two different sources, namely chimpanzees (Pan troglodytes) [2] and sooty mangabeys (Cercocebus atys) [3,4], respectively (Fig. 1). Second, since these simian viruses entered the human population, they have rapidly accumulated even more genetic diversity because of the high error rate of reverse transcriptase [5] and the fast turnover of virions in HIV-infected individuals [6,7]. On top of this, reverse transcriptase is known to be highly recombinogenic [8], so that radically different genomic combinations may be generated in individuals infected by genetically diverse viruses.

The initial epicentres of HIV-1 and HIV-2 infection appear to have been Central Africa and West Africa, respectively, reflecting the natural habitats of chimpanzees and sooty mangabeys. From these origins, different lineages of viruses have spread to very different extents around the world. For example, the original survey describing HIV-2 infection in commercial sex workers in West Africa in 1985 revealed no HIV-1 infection in the same population [9]. HIV-2 is still primarily found in West Africa, and HIV-2 prevalences are currently stable or even decreasing [10]. In contrast, HIV-1 has spread throughout Africa, including West Africa, and some lineages of HIV-1 have dispersed around the world, so that HIV-1 is predominant globally.

This review will focus on the current state of knowledge of HIV-1 diversity and its geographical distribution, as well as the implications of this genetic variability for diagnosis, treatment, and the development of vaccines.

Classification of HIV-1 strains

HIV-1 groups

Phylogenetic analyses of numerous strains of HIV-1, isolated from diverse geographic origins, have revealed three distinct clades of viruses, which have been termed groups M (main), N (new, or non-M, non-O) and O (outlier). Since SIVcpz lineages are interspersed among these three HIV-1 lineages (Fig. 1; see also [1,2,11]), each of the groups must have arisen from a separate cross-species transmission event [2]. All of the HIV-1 strains described in the 1980s, and still the vast majority of strains found worldwide and responsible for the pandemic, belong to just one of these lineages, group M. Within group M, there is further phylogenetic structure, which has allowed the classification of strains into numerous subtypes (see later).

Group O isolates were first described around 1990 [12]. These viruses are highly divergent from group M (Fig. 1), exhibiting only about 50% amino acid sequence identity to group M in the env gene [13,14]. Group O seems to be endemic to Cameroon and neighbouring countries in West Central Africa, but...
Fig. 1. Evolutionary relationships among the primate lentiviruses HIV-1, HIV-2 and SIV from various species of simians. Five major clades are evident: the HIV-1/SIVcpz lineage, which includes three groups of HIV-1 plus examples of SIVcpz isolated from chimpanzees; the HIV-2/SIVsm lineage, including several subtypes of HIV-2 plus SIVsm isolated from sooty mangabeys and SIVst from a captive macaque; the SIVagm lineage, which includes isolates from four species of African green monkeys (viruses from tantalus, grivet and vervet monkeys are shown); the SIVlhoest lineage, including isolates from L'Hoest monkeys, sun-tailed monkeys, and mandrills; and the SIVsyk lineage from Sykes' monkeys. This unrooted phylogenetic tree was derived by maximum likelihood analysis of Pol protein sequences. (Data from Hahn et al. [1].)

HIV-1 group M subtypes
Since 1992, phylogenetic analyses of env and gag sequences have been used to classify the prevalent viruses observed in the global AIDS epidemic [20,21]. Subtypes were proposed because most sequences were found to fall into a limited number of discrete clades. Initially, five subtypes were identified, but more extensive global sampling has revealed additional subtypes: at the latest count, there are 11, designated A–K [22]. The subtypes are approximately equidistantly related (Fig. 2), exhibiting 25–35% amino acid sequence difference in their Env proteins, and up to 20% difference within subtypes [23]. This ‘starburst’ radiation at the root of the group M evolutionary tree is consistent with an exponentially growing population size [24]; in this case, meaning an exponential growth of the number of infected individuals at the start of the M group radiation.

It is important to distinguish between pure subtypes and recombinant viruses (see later). To be considered as a subtype, isolates should resemble each other, and no other existing subtype, across the entire genome. In this light, there are only nine subtypes of HIV-1 group M, since the prototypic viruses of subtypes E and I have been found to be recombinants, in which substantial fractions of the genome were derived from
Fig. 2. Evolutionary relationships among non-recombinant HIV-1 group M subtypes, based on neighbour-joining phylogenetic analysis of near-full-length genome sequences. Each of the internal branches defining a subtype or sub-subtype is supported by 100% of bootstraps.

other subtypes. In addition, in the case of subtype G, there is some ambiguity about the origins of the accessory gene region, which resemble subtype A [25,26]; however, most of the subtype G genome is phylogenetically distinct.

A number of HIV-1 isolates remain unclassified, because they neither fall within any of the designated subtypes nor are recognizable inter-subtype recombinants, but have not been assigned to a new subtype because they are unique. Since full-length genome sequencing has become technically easier, and with the development of a variety of tools to detect the existence of mosaic (i.e. recombinant) genomes, the criteria to propose a novel subtype have been revised recently [27]. Three near-full-length genomic sequences, preferably, or two complete genomes in conjunction with partial sequences, are needed to define a new subtype. These strains must be identified in at least three individuals with no direct epidemiological linkage.

Within some subtypes, further phylogenetic structure can be identified, leading to a classification into sub-clades. Subtype F is subdivided into two sub-clades, F1 and F2. F1 viruses have been found in Brazil, Romania and Central Africa, whereas the F2 subclade consists exclusively of viruses from Cameroon [22]. Interestingly, the extent of divergence between sub-clades F1 and F2, while hardly any greater than the diversity seen within some other subtypes, is similar to that seen between subtypes B and D. With the benefit of hindsight, it is clear that for consistency, subtypes B and D would be better considered as sub-clades of a single subtype but, for historical reasons, it is difficult to change these designations. Subtype B gained its subtype status because it is the common form of HIV-1 in North America and Europe, and was very heavily represented among the first strains characterized.

It seems likely that the various subtypes, and sub-clades within subtypes, have been generated by founder effects. Thus, subtype B arose from a single strain that emerged from Africa. The phylogenetic structure at the root of subtype B is again a starburst, mirroring the initial radiation of group M subtypes, and again most likely due to exponential growth in the number of infected individuals at the start of the epidemic in Central and North America, and Europe.

Recombinant HIV-1 viruses
Following the designation of group M subtypes, it was realized that certain isolates seem to belong to different subtypes, depending on the region of the genome used for phylogenetic analysis [28]. Such discordant phylogenetic positions most likely indicate that the virus is an inter-subtype recombinant. One of the first African HIV-1 isolates to be characterized, MAL from Zaire [29], had been identified as a probable recombinant [30], but it was not until 1994 that the first multiply-infected individual was identified [31]. Subsequent reports have identified many more mosaic viruses, and confirmed their frequency at about 10-20% of newly characterized strains [28,32,33]. That these mosaic viruses are indeed recombinants is supported by the fact that discrete breakpoints can be identified between the genomic regions with differing phylogenetic affinities [25,26], and is further reinforced by the observation that most of these strains can be traced to geographical areas where the relevant subtypes co-circulate [25,34]. For example, numerous inter-subtype mosaic combinations have been reported from Africa [35-37], where all subtypes are found, while B/F recombinants have been found in Brazil, where subtypes B and F are both common [34].

Some of these mosaic HIV-1 genomes are unique, or restricted to small transmission clusters. However, others have been identified in several, apparently unlinked, individuals and play a major role in the global AIDS epidemic: these are now designated as 'circulating recombinant forms' (CRF) [38]. Members of a CRF should resemble each other over the entire genome, with similar breakpoints reflecting common ancestry from the same recombination event(s). There are currently several CRF of HIV-1: under new nomenclature proposals, each will be designated by an identifying number, with letters
indicating the subtypes involved [27]. If the genome contains sequences originating from more than three subtypes, the letters will be replaced by 'cpx', denoting 'complex'.

All known representatives of what was initially described as subtype E appear, in fact, to be recombinants of subtypes A and E [39,40], and are now designated CRF01_AE [27]. The only non-subtype A sequences are found within (most of) the env gene, parts of vif, vpr and nef, and the LTR (Fig. 3). Subtype E was first designated on the basis of the distinct phylogenetic position of these viruses in env trees. However, no virus has yet been described that represents a full-length non-recombinant, subtype E sequence, perhaps simply because that lineage has become extinct. This absence of one of the 'parental' lineages leads to some technical difficulties in formally proving the recombinant nature of these viruses.

The IbNg strain from Ibadan, Nigeria, was initially described as a divergent lineage within subtype A, based on partial gp120 sequences [41]. However, after the determination of a full-length sequence, IbNg was recognized as a complex mosaic of alternating subtype A and subtype G sequences [26]. Since a number of similar viruses have been reported from countries in both West and East Africa, this clade is now designated CRF02_AG (Fig. 3).

An epidemic among intravenous drug users (IVDU) in Kaliningrad, Russia, involves viruses that are mosaics of subtypes A and B [42,43], and which are therefore termed CRF03_AB.

Isolate 94CY032 from Cyprus was designated as the prototype of subtype I based on gp120 sequences [44]. However, full genome sequencing revealed this virus to be a complex mosaic with multiple break-
HIV-1 genetic diversity Peeters and Sharp

Fig. 4. Distribution of HIV-1 groups, subtypes and circulating recombinant forms (CRF) in different parts of the world. Only the predominant forms in each region are shown. Due to international travel and intermingling of populations, many subtypes other than those indicated are reported in each region.

Worldwide distribution of HIV-1 variants

Subtype designations have been powerful molecular epidemiological markers to track the course of the HIV-1 pandemic. Extensive efforts have been made to collect and characterize HIV-1 isolates from around the world, and a broad picture of the distribution of subtypes and CRFs have emerged (Fig. 4). Globally, the predominant viral forms are subtypes A and C, followed by subtype B, and the recombinants CRF01_AE and CRF02_AG [33,49].

The greatest genetic diversity of HIV-1 has been found in Africa. Group M subtypes A and C are most common, but all groups and subtypes are found, consistent with this continent being the source of the epidemic. As expected, given the presence of numerous co-circulating subtypes, a high frequency and a wide variety of recombinants have also been reported in Africa [36,37]. In South and East Africa, subtype C predominates. In West and Central Africa, as judged by env sequences, subtype A-like viruses are most common [50]. Full-length genome sequences of viruses resembling IbNg, representing CRF02_AG, have recently been reported from different locales in Senegal, Cote d'Ivoire, and Cameroon, and it seems likely that the majority of viruses with subtype A gag and/or env sequences in West Africa and West Central Africa belong to this circulating recombinant form [33,51]. In contrast, in East Africa, the subtype A viruses are predominantly non-recombinant [52]. Subtype D is present at frequencies of 5-40% in Central and East Africa [50,53], while subtype G has been documented in many West and Central African countries. Subtypes F, H, J and K, as well as CRF01_AE, are mainly seen in Central Africa [22,36,37,50]. We have already noted that HIV-1 group O is only common in West Central Africa, and group N has only been reported in the same region.

Specific distributions of the various subtypes are seen in other continents, although sporadic cases of many different variants of HIV-1 have now appeared in most parts of the world. In North America, Europe and Australia, subtype B is by far the most common. However, various other group M subtypes, and even group O viruses, have been reported in the United States [54-56] and several European countries [57-60]. For example, a study of blood donors in France showed a 16% increase between 1984 and 1995 in the prevalence of individuals infected by non-subtype B viruses [61], and other varieties are also on the increase in Sweden and Germany [62,63].

In South America, subtype B predominates, but subtypes F and C are also found. Several instances of B/F recombinants have been reported from Brazil and Argentina [49,64,65].

Across Asia, different forms of HIV-1 have taken hold in different regions. In India, subtype C predomi-
nates, but subtypes A and B co-circulate, and A/C recombinants are also found [66,67]. In contrast, in Thailand, first reports indicated two separate epidemics, of subtype B and 'subtype E' (i.e. CRF01_AE) viruses, spreading in different sections of the population: IVDU were infected with subtype B viruses, whereas those infected heterosexually predominantly carried subtype E (now CRF01_AE) viruses. More recent data show that the proportion of CRF01_AE has increased in almost all population groups, so that 35% of IVDU in Bangkok are now infected with this variant, and it is also the predominant form among younger individuals [68–70]. CRF01_AE is also increasing in frequency in other South-East Asian countries [71]. This has led to speculation that CRF01_AE is more transmissible than subtype B, but it is difficult to discount that the differential spread merely reflects epidemiological factors (see later).

The example of Thailand, in particular, clearly illustrates that the global distribution of different forms of HIV-1 is a dynamic process. As more HIV-1 variants inevitably intermix in different parts of the world, the likelihood of generating new recombinant viruses will increase. Indeed, the pattern of mosaicism will become even more complex, since recombination involving viruses that are already recombinant will occur. Mosaics involving CRF02_AG have already been observed in various African countries [33,72]. Continued monitoring is necessary to determine the future role of non-subtype B viruses in North America and Europe, and to chart the emergence of new predominant subtypes and CRF around the world.

**Methods of classification**

Sequence determination is the most accurate approach to characterize virus genomes, but there is a clear need for continued development of techniques for typing that are less cumbersome, and less expensive. Serological methods, as well as polymerase chain reaction (PCR)-based techniques, are continually being refined. These approaches can usually define the subtype affinity of a particular region of the viral genome, but they will not detect whether other regions are mosaic or belong to a different subtype.

**Serological methods for subtyping**

Serological assays are greatly preferable, because they are easy to use and large numbers of samples can be tested. Enzyme-linked immunosorbent assays detecting antibodies against the V3 loop of the envelope protein have been used successfully to discriminate among HIV-1 group M, N and O infections [15,19,73]. However, V3 serotyping is less useful in determining group M subtypes. These enzyme-linked immunosorbent assays are only useful in areas where a limited number of major variants co-circulate, as with subtype B and CRF01_AE in Thailand, or subtypes B and C in South Africa [68,69,74]. The technique has also proved useful in France, to discriminate between subtype B and non-subtype B infections [58]. However, in regions where many subtypes have been co-circulating for a long period, serotyping can lead to misidentification of genotypes [75,76].

**PCR-Based approaches**

The heteroduplex mobility assay (HMA) is a PCR-based method that allows rapid subtyping using a set of reference reagents representing different subtypes [77]. Less sophisticated equipment is required for this technique than for sequencing, and HMA has in fact been widely used in developing countries to identify env gene subtypes. One limitation of HMA is that, as HIV-1 sequences continue to diverge and as new geographic regions are sampled, the primers used are failing more often, and a greater number of samples remain indeterminate. Another is that env HMA cannot discriminate between subtype A and CRF02_AG viruses, which are also subtype A in the env region analyzed. More recently, HMA has been developed to characterize gag genotypes, which can distinguish between subtype A and CRF02_AG. Simultaneous use of gag and env HMA can give a preliminary estimation of the frequency of recombinant viruses [78].

Other PCR–based methods have been developed, including a restriction fragment length polymorphism technique to determine protease region genotypes [79], and oligonucleotide assays that have been used to distinguish between env sequences of subtypes B and E in Thailand [80]. Subtype A-specific PCR of the env gene has been developed for areas where that clade is highly prevalent [81].

**Sequence determination**

Nevertheless, sequencing remains the most accurate approach to identify HIV-1 variants. Even only partial gag and/or env sequences give more precise information than serotyping or HMA with regard to the presence of subclades or recombinant viruses. Also, partial genome characterization allows a preliminary selection of important strains for full sequencing. Full-length genome sequencing is needed for the continued surveillance of HIV-1 global variation. A full-length sequence is also necessary to exclude the possibility that a new isolate is recombinant, or to determine the pattern of mosaicism within an isolate that is recombinant.

A variety of complementary approaches have been developed to identify sequences that are recombinants,
and to map the positions of breakpoints within mosaic sequences. General methods to identify recombinant sequences examine whether the evolutionary relationships, or more simply relative distances between sequences, vary among different windows along a sequence alignment [82–84]. Breakpoints between sequence regions with different evolutionary histories can be identified by mapping the linear distribution of divergent and/or phylogenetically incompatible sites [85,86]. In the specific case of HIV-1 sequences, because the different subtypes of the M group have been well defined, potential intersubtype recombinants can be analyzed in a more-or-less automated fashion. Using moving window analysis, diversity plots [25,67] can be used to display the extent of similarity of a new sequence to representatives of other subtypes; the recombinant identification program [87] uses distance measures to assign subtypes to regions within the new sequence, and the strength of bootstrap support for the phylogenetic placement of the new sequence with any subtype representative can be assessed [88]. Fine-scale mapping of recombination breakpoints has been performed using informative site analysis [25,89]. Much of the software used to perform these analyses is freely available and/or can be accessed online.

Studies of HIV-1 variants in Nigeria, the most populous African country, illustrate the value of rapid screening assays that can be performed locally, and also the need for more sophisticated analysis of a subset of samples in order to completely understand the situation. In Nigeria, two env subtypes, A and G, were found to be predominant by HMA. However, partial gag sequencing revealed that more than 30% of samples had discordant gag and env subtype designations. Phylogenetic tree analysis of gag and env sequences showed several subclusters within subtypes A and G, and numerous A/G recombinants with different genomic structures. Thus, despite the limited number of common HIV-1 subtypes in Nigeria, there is a complex mix of viruses circulating in that country [90,91].

Implications of HIV-1 variability

Whether the various groups, subtypes and recombinant forms of HIV-1 have biological differences, for example with respect to transmissibility and the course of disease progression, is not yet well known. Clearly, the extent of genetic divergence among subtypes (let alone among groups) is more than sufficient to cause such differences, but adequately controlled data from in vivo studies may take a long time to emerge. Importantly, diagnostic tests, antiretroviral drugs, and HIV-1 vaccines have so far mainly been developed only for subtype B viruses.

Serological and molecular diagnostic tests

Examples of HIV-1 group O were initially identified because they were not detected by certain commercial serological screening assays, and group O sera can give indeterminate results in confirmatory tests such as Western blots [92,93]. The majority of commercial assays are now capable of detecting group O antibodies, either through the use of broadly cross-reactive group M antigens, or by addition of specific group O antigens. However, continued surveillance is necessary, since even subclade changes in antigenic structure of some HIV-1 variants may affect the sensitivity of these tests. During seroconversion, a significantly lower sensitivity was observed for the detection of non-subtype B infections using screening assays with subtype B–derived antigens [94]. Several studies have reported that commercial viral load assays were either not able to detect or incorrectly quantified viral RNA in plasma from patients infected with non-subtype B strains of HIV-1 [95]. These assays have been re-evaluated, and other assays are under development, but as yet no assay is commercially available that can detect and quantify HIV-1 group O or HIV-2 strains. As even greater variability of HIV-1 emerges, these tests will need to be continuously monitored to evaluate the efficiency of the primers or probes used.

Antiretroviral treatment

Antiretroviral drugs are not yet widely used in developing countries, where non-subtype B strains of HIV-1 predominate. However, this situation could change in the coming years, as several governments and international organizations have recently made antiretroviral treatments available to certain developing countries, and antiretroviral drugs are being introduced to prevent perinatal transmission. In vitro data have indicated that HIV-1 group O viruses, as well as HIV-2, are naturally resistant to non-nucleoside reverse transcriptase inhibitors [96,97]. There are also indications of variation in drug susceptibility within group M. Thus, some subtype F samples are less susceptible to 8-chloro–tetrahydro-imidazo(4,5,1-jk)(1,4)-benzodiazepin-2(1H)thione, a non-nucleoside reverse transcriptase inhibitor [98], while some subtype G strains are less susceptible to protease inhibitors [99]. Many amino acid mutations associated with secondary resistance to protease inhibitors have been reported as natural variants in treatment-naïve patients infected with non-subtype B HIV-1 strains, but the biological consequences of this remain to be studied [100–103].

Biological differences, transmission and disease progression

Differences between HIV-1 and HIV-2 with respect to transmissibility and disease progression have been clearly documented: HIV-2 exhibits a lower rate of
heterosexual transmission, a near absence of vertical transmission, and a longer incubation period before development of AIDS [104–106]. The question remains whether similar differences exist among the various forms of HIV-1.

There have been some indications that there may be biological differences among HIV-1 groups and subtypes, but also some apparently contradictory results. A prospective cohort study among female sex workers in Senegal showed that subtype A-infected women were eight times less likely to develop AIDS [107]. The CXCR4-positive phenotype, previously known as rapid/high or syncytium inducing, is rare among subtype C-infected patients [108–110]. With subtype B, the presence of CXCR4-positive viruses is associated with a more rapid progression to AIDS, but it remains to be determined whether subtype C viruses are less virulent or are associated with different clinical outcomes. No difference in disease progression was found between patients infected with subtypes B and C in Israel [111], or among patients infected with subtypes A, B, C and D in Sweden [112]. However, in a cohort of pregnant women in Kenya, higher viral loads and lower CD4 cell counts were seen in subtype C patients than in those infected with subtypes A or D [113].

The rate of heterosexual transmission for CRF01_AE in Thailand is fivefold higher than for subtype B in the United States [114]. More prospective epidemiological studies from different regions are needed to determine whether this is really due to intrinsic properties of the virus rather than population-related factors. No difference was found between CRF01_AE and subtype B in rate of disease progression during a 3-year follow-up study in Thailand [115].

The rate of disease progression varies from person to person and is influenced by multiple virological and host factors. The evaluation of associations between subtypes and phenotypes will be very difficult, and it will not be easy to determine whether any association is causal or due to bias or chance. Nevertheless, it would be premature to conclude that there are no subtype-specific differences in virus biology, transmission or disease development.

Vaccines
Most candidate HIV-1 vaccines currently in development or in clinical trials are based on subtype B antigens, and the degree to which these will elicit cross-protection against other subtypes is poorly understood. In one study, in Thailand, neutralizing antibodies appeared to be subtype specific [116]. However, several other studies have suggested that there is no correlation between HIV-1 subtypes and neutralizing serotypes: some sera appear to be broadly cross-reactive, and some isolates highly susceptible to neutralization by many sera, but other isolates seem neutralization resistant [117].

Similarly, while one study reported a subtype-specific cytotoxic T-lymphocyte response [118], several others have shown cross-clade cytotoxic T-lymphocyte reactivities in individuals infected with non-B subtypes, as well as in recipients of a subtype B vaccine [119,120]. These data suggest that some cytotoxic T-lymphocyte epitopes may be conserved among different subtypes, and emphasize the importance of including some of the more conserved proteins in vaccine constructs.

Implications of recombination
Retroviral recombination requires the simultaneous infection of a cell with two different proviruses, allowing the encapsidation of one RNA transcript from each provirus into a heterozygous virion. The discovery of large numbers of recombinant viruses clearly implies that co-infection with divergent HIV-1 strains is not as rare as once thought. Indeed, dual infections with different subtypes have been reported in regions where multiple variants co-circulate [31,65,72,121]. In any geographic region, the proportion of recombinant viruses will depend on a number of factors, including the prevalence rates of different subtypes, the probability that certain population groups acquire multiple infections and transmit their viruses further, and the fitness of any mosaic viruses generated. However, the frequency of recombinant viruses is almost certain to increase; recombination, once it has occurred, cannot be undone, and so the frequency of 'pure' subtype strains is likely to decrease.

It remains to be determined whether there are constraints, during the course of HIV infection, on when superinfection can occur. In macaques, it has recently been shown that superinfection with divergent strains of HIV-2 is possible only during certain periods when antibodies are not yet efficiently expressed [122]. In contrast, in a chimpanzee, superinfection with a CRF01_AE virus 32 weeks after experimental infection with a subtype B strain led to a dual infection with the rapid appearance of recombinant viruses [123]. The latter result suggests that superinfection is not restricted to the early phase of infection, and implies that the humoral and cellular antibody response are not efficient against divergent strains.

It was initially suspected that homologous recombination between group M and group O viruses may not be possible because of their high degree of divergence. However, two recent reports have documented intergroup recombinants in two different patients from Cameroon [72,124]. One study showed...
that these M/O mosaic viruses replicated well in vivo and in vitro, and indeed became the predominant variants within the patient’s viral population [72]. Recombination between such divergent strains could contribute substantially to the emergence of new HIV-1 variants, and would have important implications. Both for diagnosis by serological and molecular tests, and for treatment. This observation also suggests the possibility that HIV could recombine with SIV, for example in individuals who are exposed to SIV by cross-species transmission, and so by this means additional SIV sequences could spread into the human population.

While dual infections with HIV-1 and HIV-2 have frequently been reported in regions where both viruses circulate [37], no recombinants between them have as yet been described. In this case, the level of genetic divergence may be too high for successful recombination, although its possibility cannot be entirely excluded.

Clearly, superinfection and consequent recombination can affect the biological properties, and pathogenesis, of HIV strains. Several studies have found that, under the selective pressure imposed by antiretroviral drugs, recombination between strains with different drug sensitivities occurred, resulting in new HIV-1 variants with dual drug resistance [125]. In vitro experiments with feline and murine retroviruses have demonstrated that, under appropriate selection pressures, mixed infections can generate recombinant viruses with altered tissue tropism, pathogenicity, and host range, or with changes in antigenic epitopes [126,127]. Finally, recombination also has important implications for vaccine strategies based on live-attenuated viruses, since these could recombine with infecting strains, even if the two are quite divergent.

Conclusions and perspectives

It is increasingly evident that the distribution of HIV-1 genetic subtypes is a dynamic and unpredictable process. The geographical distribution of subtypes is evolving, and intermixing of HIV-1 variants is inevitable. Recombinant viruses are already contributing substantially to the global pandemic, and the frequency and complexity of recombinant viruses will increase as the various subtypes spread to all continents, and viruses that are already mosaic are involved in further recombination events.

If there are biological differences among the various subtypes and CRFs, it is important that these should be clearly defined, and it remains important to track the molecular epidemiology of HIV-1. In future studies, both ‘pure’ subtypes and CRFs will need to be monitored.

More studies are needed to understand the role and the implications of recombinant viruses in the global HIV evolution. It is important to study in more detail the impact of viral recombination on viral properties, since recombination may introduce genetic and biological consequences that are far greater than those resulting from the steady accumulation of single mutations. In order to develop an efficient vaccine, it remains to be determined when superinfection can occur during the course of HIV infection, and to what extent humoral and cellular antibody response are efficient against divergent strains.

References

18. Mas A, Quines-Baen ME, Domingo E, Soria Y. Phylogeny of HIV type 1 group O isolates based on env gene sequences.


