The Challenge of HIV-1 Subtype Diversity

Barbara S. Taylor, M.D., Magdalena E. Sobieszczyk, M.D., M.P.H., Francine E. McCutchan, Ph.D., and Scott M. Hammer, M.D.

From the Department of Medicine and the Division of Infectious Diseases, Columbia University Medical Center, New York. Address reprint requests to Dr. Taylor at Columbia University Medical Center, Department of Medicine, Division of Infectious Diseases, 630 W. 168th St., P&S Box 82, New York, NY 10032, or at bs2026@columbia.edu

Nearly 27 years after the first reported cases of the acquired immunodeficiency syndrome (AIDS) and 25 years after the discovery of the etiologic agent, effective control of the AIDS pandemic remains elusive. At the root of this challenge is the molecular pathogenesis of human immunodeficiency virus (HIV) type 1 (HIV-1), a virus that has evolved a number of mechanisms to elude immune control. Among the most prominent of these are the heavy glycosylation of the external glycoprotein, which protects neutralization epitopes; the virus’ direct targeting of the CD4 molecule expressed by the key T lymphocyte in immune orchestration; integration into the host-cell genome, which implies that cells that are not killed are infected permanently; and the potential of the virus to mutate and therefore evade the host immune system (mutational escape).1,2 This last mechanism results in a remarkable degree of viral diversity within HIV-1 and its rapid adaptation, in response to both immune activity and antiretroviral therapy. Over the past decade, advances in sequencing technology and expanded disease surveillance have allowed researchers to characterize the variation in HIV-1 within individual patients and around the world.

The initial view that the virus is classifiable into distinct subtypes or clades now needs to reflect the reality of a dynamic genetic evolutionary process, through which new HIV-1 strains are constantly emerging. The resultant viral diversity has implications for possible differential rates of disease progression, responses to antiretroviral therapy (including the development of resistance), and vaccine development.

ORIGIN OF HIV AND MECHANISMS OF HIV DIVERSITY

The origin of HIV-1 among nonhuman primates has been traced to a simian virus, SIVcpz, which infected several geographically isolated chimpanzee communities in southern Cameroon. This HIV-1 progenitor probably was passed from chimpanzees to human hunters through bloodborne transmission. Phylogenetic analysis of HIV-1 and related viruses from nonhuman primates suggests that three independent transmission events early in the 20th century spawned three HIV-1 groups: major (M, between 1915 and 1941), outlier (O), and nonmajor and nonoutlier (N).3,4 Although strains related to the M and N groups have been found in chimpanzees, recent evidence suggests that group O HIV-1 may have originated in gorillas, in which the closest relatives of this group have been identified.5 It is speculated that the virus then spread among humans along the Congo River into Kinshasa, Zaire, where the earliest documented case of HIV-1 infection (with group M strain) in humans has been traced to a blood sample from 1959.6

Dr. Hammer reports receiving consulting fees from Merck, Progenics, and Wyeth and serving on a data and safety monitoring board for a clinical trial sponsored by Bristol-Myers Squibb. No other potential conflict of interest relevant to this article was reported.
HIV has several intrinsic mechanisms that ensure rapid viral evolution. The reverse transcriptase of HIV lacks proofreading activity, the ability to confirm that the DNA transcript it makes is an accurate copy of the RNA code, and confers a mutation rate of approximately $3.4 	imes 10^{-5}$ mutations per base pair per replication cycle. Since the HIV genome is an estimated $10^4$ base pairs in length and the baseline rate of viral production is approximately $10^{10}$ virions per day, millions of viral variants are produced within any infected person in a single day.\(^7\) HIV-1 recombination can lead to further viral diversity and occurs when one person is coinfected with two separate strains of the virus that are multiplying in the same cell (Fig. 1).\(^8,9\)

**CLASSIFICATION AND MOLECULAR EPIDEMIOLOGY OF HIV**

Group M is the predominant circulating HIV-1 group. It has been divided into subtypes, denoted with letters, and sub-subtypes, denoted with numerals. Subtypes A1, A2, A3, A4, B, C, D, F1, F2, G, H, J, and K are currently recognized. HIV-1 subtypes, also called clades, are phylogenetically linked strains of HIV-1 that are approximately the same genetic distance from one another; in some cases, subtypes are also linked geographically or epidemiologically. Genetic variation within a subtype can be 15 to 20%, whereas variation between subtypes is usually 25 to 35%.\(^10\) Over the past decade, advances in full-genome sequencing of HIV have led to the identification of circulating and unique recombinant forms (CRFs and URFs, respectively). These are the result of recombination between subtypes within a dually infected person, from whom the recombinant forms are then passed to other people. The recombinant progeny are classified as circulating recombinant forms if they are identified in three or more people with no direct epidemiologic linkage; otherwise they are described as unique recombinant forms (Table 1).\(^11\)

These definitions have been evolving over the past decade. Nomenclature in the published literature varies, and certain subtypes were found to be more complex after their full genomes have been sequenced. For example, what was previously described as subtype “E,” circulating in Southeast Asia, proved to be a circulating recombinant form containing components of subtype A and was redefined as CRF01_AE in 1998.\(^11–13\)

The global distribution of subtypes and circulating recombinant forms reflects the complexity of the molecular epidemiology of HIV-1 (Fig. 2). The CRF01_AE virus was first identified in Thailand in the late 1980s.\(^12,14\) This strain and its close relatives in Central Africa have had very different fates. CRF01_AE dominates in Southeast Asia, whereas in Africa, this circulating recombinant form remains relatively rare.\(^15,16\) Two new circulating recombinant forms, combining the Thai B and Indian C strains in related but distinct mosaic structures, emerged in southern China among injection-drug users and spread, along various drug-trafficking routes, across the country.\(^17,18\) A subtype A strain of low diversity, and a new circulating recombinant form derived from it, CRF03_AB, emerged in the former Soviet Union.\(^19,20\) More recently, CRF14_BG arose among injection-drug users in Spain and Portugal and has continued to spread.\(^21\) Subtype F, rare in Central Africa, emerged in South America in the form of BF recombinant strains; the subtype F parent of these recombinants has never gained significant prevalence.\(^22,23\) Finally, the HIV-1 epidemics in Afghanistan and Iran, fueled principally by the use of injection drugs, were shown to be linked: the newly emerged CRF35_AD strain was implicated in both.\(^24,25\)

**CORECEPTOR USE BY HIV-1 SUBTYPES, TRANSMISSION, AND DISEASE PROGRESSION**

Differential characteristics of viral subtypes and their interactions with the human host may influence HIV transmission and disease progression. The HIV strains capable of using the
chemokine coreceptor CCR5 (R5 viruses) are more frequently transmitted than strains that use the CXCR4 co-receptor (X4 viruses); X4 viruses emerge later in infected patients and are associated with more rapid disease progression. All HIV-1 subtypes can use both coreceptors, but subtype D may be dualtropic (i.e., an R5X4 virus) most frequently. The percentage of X4 virus appears to be lower in subtype C than in subtype B, even when the viruses are obtained from patients with advanced AIDS.

There are suggestions in the published literature that HIV-1 subtype or CRF may affect efficiency of transmission. Early data on mother-to-child transmission implied that subtype C was transmitted more frequently than subtype B. Pregnant women in Kenya infected with subtype C were more likely than those infected with subtype A or D to shed HIV-1–infected vaginal cells, implying that sexual transmission may be more likely with this subtype. A study in a longitudinal cohort of injection-drug users in Thailand conducted from 1995 through 1998 found an increased probability of transmission of CRF01_AE as compared with subtype B, though it was unclear whether epidemiologic, virologic, or host factors were affecting viral spread. An examination of subtype distribution between 1986 and 2000 in Kenya did not show an increase in the prevalence of subtype C; instead, an increase in the number of recombinant viruses was found.

Another important question is whether subtype differences result in variable rates of disease progression. There have been several prospective, observational studies of the course of HIV-related disease in cohorts infected with various subtypes. An early study was published in 1999 by Kanki et al., who had examined subtypes in 54 female sex workers in Senegal who were infected with HIV-1. Subtypes A, C, D, and G were represented, and the likelihood of AIDS developing was increased by a factor of eight among women infected with a non-A subtype as compared with those infected with subtype A. The subtypes were determined on the basis of the standard at the time, envelope-gene subtyping only; thus, further genetic complexity may have been present.

Subsequent studies have reported discordant results. In a large Swedish study of patients infected with subtype A, B, C, or D, disease progression did not differ significantly according to subtype or ethnic group. A second prospective multicenter study in Western and West-Central Africa did not show a significant difference in survival or clinical disease progression among people infected with CRF02_AG, as compared with those infected with other CRFs or subtypes. A survival study of 836 Thai heterosexual men and women infected with CRF01_AE showed a shorter time from HIV-1 infection to death than among those in Western populations.

A cohort of 1045 Ugandans did show a faster progression to death among people infected with subtype D than among those infected with subtype A, even after controlling for CD4+ count at enrollment. Data from the Rakai cohort in Uganda also suggest that HIV-1 disease progresses more rapidly, and that the risk of death is greater, among persons infected with subtype D, with recombinant forms, or with multiple subtypes than among those infected with subtype A. A recent study of a Kenyan cohort showed that 21 patients infected with subtype D had a higher mortality rate and a faster decline in CD4+ count than those infected with subtype A or C. The propensity of subtype D to exhibit a greater degree of dualtropism than other subtypes may help to explain the observation that subtype D appears to be associated with a more rapid rate of disease progression than other HIV-1 subtypes. The notable caveat relevant to all these studies of disease progression is that confounders such as access to medical care, nutritional status, host genetic factors, and mode of viral transmission (e.g., sexual, injection-drug, or vertical) may contribute to the divergent results.
Interactions between the host and HIV-1 that vary according to subtype may also be important. The known differences in HIV-1 transmission and disease progression in hosts carrying specific HLA class I types may vary according to infecting HIV-1 subtype (Table 2). In an infected person, T cells specific to HIV-1 can exhibit cross-subtype specificity and recognize viral epitopes within subtypes other than the one that generated the initial response.\textsuperscript{52,53} CD8+ T cells obtained from persons infected with subtype B recognize viral epitopes within conserved regions of the consensus sequences from genomes of subtypes A, B, and C. However, the immune response tends to be greatest against the infecting subtype, and CD8+ T-cell responses can wane over time. Despite some similarities in T-cell responses, there may be intersubtype differences in the plasma HIV-1 viral load levels after in vivo infection, though data thus far are conflicting.\textsuperscript{54,55} This is an important area of investigation because of the well-described link between viral load and transmission and the rate of disease progression.\textsuperscript{56}

**RESPONSE TO THERAPY**

Does HIV-1 subtype influence the response to antiretroviral treatment? This question is urgent, since only 12% of global infections are caused by the most studied subtype, B; and 50% of prevalent HIV infections and 47% of all new HIV-1 infections are with subtype C.\textsuperscript{10} This discrepancy in the availability of clinical data for non-B subtypes is exacerbated by the fact that, until the past few years, antiretroviral treatment had been largely unavailable in many countries with non-B subtypes of HIV-1.

Initial data from treatment cohorts in Africa raise two concerns: first, that certain subtypes of HIV-1 might spread or progress more rapidly than others, making treatment decisions more challenging;\textsuperscript{37} and second, that the data on baseline antiretroviral susceptibility derived from studies of subtype B may not be applicable to non-B subtypes.\textsuperscript{57} This concern is illustrated by HIV type 2 and group O strains of HIV-1, which possess intrinsic resistance to nonnucleoside reverse-transcriptase inhibitors.\textsuperscript{58,59}

Though there are potential problems with comparing responses to therapy among persons infected with group M, non-B–subtype strains — who frequently live in settings with limited resources — and those infected with subtype B, the data available thus far are encouraging. In 2002, Alexander et al.\textsuperscript{60} published data from a retrospective cohort of 485 patients receiving antiretroviral treatment in British Columbia, Canada, 4.4% of whom were infected with non-B subtypes of HIV-1. Though initial CD4+ counts were lower in the patients infected with non-B subtypes than in those infected with subtype B, the proportion of patients with an HIV-1 RNA viral load of less than 400 copies per milliliter at 18 months did not differ significantly between the two groups. A French cohort study of 416 patients, 24% of whom carried non-B subtypes of HIV-1, showed that at 3, 6, and 12 months after initiation of antiretroviral therapy, HIV-1 subtype did not affect clinical progression, CD4+ count, or viral load in response to treatment.\textsuperscript{61}

In their study of patients of African origin who were infected with a non-B subtype of HIV-1 and were living in London, Frater et al.\textsuperscript{41} found no significant difference in the response to therapy among patients infected with subtype A, those infected with subtype C, and those infected with subtype D. In the Paediatric European Network for Treatment of AIDS (PENTA) 5 trial,\textsuperscript{62} there was no significant difference according to HIV-1 subtype in the virologic response to treatment or in the frequency of development of resistance among children. Overall, it appears that HIV-1 subtypes do not effect major differences in the response to antiretroviral therapy. What has emerged is a growing body of evidence that polymorphisms found in various subtypes before antiretroviral therapy is begun may affect genetic pathways of resistance.
EMERGENCE OF RESISTANCE TO ANTIRETROVIRAL THERAPY

HIV resistance to antiretroviral therapy can be divided into two categories: primary resistance, which reflects acquisition of a drug-resistant strain of HIV by a newly infected person; and secondary, or acquired, resistance, which develops after a period of HIV treatment. Not surprisingly, studies of non-B subtypes in patients who have never received antiretroviral therapy reveal that protease and reverse transcriptase sequences vary from those of the subtype B reference strains in a somewhat predictable manner. \(^6^3\) For example, polymorphisms in the reverse transcriptase gene do not typically occur in known sites of resistance to nucleoside reverse-transcriptase inhibitors. \(^6^4,6^5\) In contrast, data from several studies indicate that protease sequences from non-B subtypes in patients who have never received antiretroviral therapy contain amino acid substitutions associated with mutations in subtype B known to contribute to secondary resistance, including K20 → R, M36 → I, and H69 → K/Q. \(^6^6,6^7\) However, these genotypic changes do not confer consistently decreased susceptibility by themselves when viral strains are subjected to phenotypic testing (Fig. 3 and Table 3). \(^6^5,6^6\)

For the fusion inhibitor enfuvirtide, substantial differences in resistance-associated mutations between B and non-B subtypes have not been found, but some polymorphisms, such as N42 → S in the heptadrepeat region 1 of glyco-protein 41, appear to be more common in non-B subtypes. \(^7^9\) Data on in vitro resistance against maraviroc, the recently approved entry inhibitor and CCR5 antagonist, do not suggest that there are differences among subtypes, \(^8^0\) but the differences in coreceptor tropism noted above lead to concern that intersubtype variation in response to therapy could exist in vivo.

Studies of resistance patterns that emerge in non-B subtypes in patients receiving antiretroviral therapy indicate that polymorphisms present in these subtypes before therapy may provide a background for the emergence of subtype-specific pathways to secondary resistance. Mutations leading to resistance appear to be similar among subtypes, but certain mutations seem to occur more frequently in non-B subtypes in particular. In a study of patients in Botswana infected with subtype C, in those without a response to a didanosine- or stavudine-based regimen, virus containing K65 → R mutations developed within 8 months, more rapidly than is seen in patients infected with subtype B. \(^4^1\) When exposed to tenofovir in culture, subtype C developed K65 → R mutations more rapidly than other subtypes, \(^4^7\) but data from clinical trials remain inconclusive. \(^4^8\) Subtype C viruses also develop resistance against nonnucleoside reverse-transcriptase inhibitors through either the K103 → N or V106 → M mutations, whereas subtype B viruses rarely develop V106 → M mutations. \(^4^9\) Nelfinavir resistance appears to occur primarily through L90 → M mutations in subtypes G and C and other non-B subtypes, whereas subtype B acquires either D30 → N or L90 → M nelfinavir-resistance mutations. \(^8^1,8^2\) Overall, it appears that most antiretroviral resistance in non-B subtypes is accounted for within the current resistance databases. \(^8^3\) Further studies of treated cohorts infected with non-B HIV-1 are needed to determine whether other subtype-specific pathways to resistance exist (Table 3).

IMPLICATIONS FOR VACCINE DEVELOPMENT

Ultimate control of the HIV-1 pandemic is dependent on the development of an effective, preventive vaccine (see the recent review by Johnston and Fauci\(^8^4\)). Among the many challenges to achieving this goal, one of the greatest is HIV-1 diversity — reflected by the presence of HIV-1 subtypes, circulating recombinant forms, and continuous viral evolution within populations and individual hosts. Hosts infected with HIV-1 have cellular and humoral immune responses to their infecting strains, but there is evidence of mutational escape by viruses from responses by CD8+ cytotoxic T cells and neutralizing antibodies over time. \(^8^5,8^6\) Although cross-reactive responses to other viral subtypes have been shown, \(^5^3,8^7\) the strength and breadth of these responses are typically limited. \(^5^2\) Several researchers have developed multisubtype consensus sequences or expression cassettes and have coadministered...
vaccines and selected cytokines in attempts to increase the breadth and strength of the cyto-
toxic T-cell response. Critical to the development of a successful HIV-1 vaccine will be our ability to decipher the genetic diversity of the virus, elicit broadly neutralizing antibodies, and generate strong CD4+ and CD8+ T-cell responses.

Development of a vaccine that induces neutralizing antibodies that bind to the trimeric envelope on the surface of the virus remains a great challenge. In two large, phase 3 trials of a monomeric form of the external glycoprotein 120, conducted in the United States and Thailand, the protein failed to protect healthy subjects from HIV infection. Some current approaches to the design of a neutralizing immunogen are to mimic glyco-protein 120–glycoprotein 41 envelope trimers on the virion surface, to produce envelope molecules with enhanced expression of neutralizing epitopes and thereby improve their relative immunogenicity, and to remove or mask the variable regions and expose conserved epitopes to focus the immune response.

There has been progress as well as challenges in the development of vaccines that induce T-cell–mediated immune responses that may not prevent HIV-1 infection but may modulate the viral load and subsequent disease progression in people who do become infected. The diversity of T-cell epitopes makes it unlikely that the use of one natural isolate for a vaccine will protect against other viral subtypes or variants within the same subtype. Many strategies are being pursued to confront this problem, including the use of consensus sequences, the deployment of a combination of immunogens from different subtypes, the creation of mosaic immunogens assembled through computational optimization from pieces of natural sequences, and the construction of multisubtype immunogens derived from conserved regions of the HIV-1 consensus proteome.

Two advanced approaches to the development of T-cell vaccines, both of which use a recombinant adenovirus type 5 vector, illustrate different strategies to address the challenge of HIV-1 subtype diversity. The first, developed by Merck Research Laboratories, involved the immunization of HIV-1–seronegative persons with a recombinant adenovirus type 5 vector containing gag, pol, and nef genes from subtype B. An unexpected development for the HIV vaccine field occurred in September 2007: a data and safety monitoring board review for the phase 2B, test-of-concept, efficacy trial of this product in the Americas (the Step study [HIV Vaccine Trials Network study 502, Merck protocol 023]) recommended that vaccinations in this trial be stopped, since statistical criteria for futility had been met. Vaccinations were also discontinued in a sister, phase 2B study, with partial enrollment, of the same product in South Africa (HIV Vaccine Trials Network study 503). Data released subsequently also raised the issue of whether there was an increased risk of HIV-1 acquisition conferred by the vaccine in persons with preexisting immunity to adenovirus type 5.

The second approach is illustrated by the preventive vaccine regimen being developed by the Vaccine Research Center of the National Institute of Allergy and Infectious Diseases. This regimen involves a DNA prime, recombinant adenovirus type 5 boost with gag, pol, and nef genes derived from subtype B and envelope genes from subtypes A, B, and C; thus, it is a multigene, multisubtype vaccine. The regimen is currently in phase 1–2 testing, and phase 2B, preliminary efficacy testing is under consideration. Ultimately, a fully effective, preventive vaccine regimen will probably need to induce strong, cross-subtype HIV-specific T-cell immunity as well as broadly reactive, neutralizing antibody activity to overcome the challenge of HIV diversity.

CONCLUSIONS

With the continuing spread of HIV, the world faces a pandemic of unprecedented genetic and geographic complexity. Five subtypes and two circulating recombinant forms have each
established a global prevalence greater than 2.5%, a level that virtually ensures their continued presence in the decades to come. Factors that influence the spread of particular subtypes or circulating recombinant forms in different geographic regions are incompletely understood. Mutation and recombination, both essential features of the HIV replication cycle, are major forces driving diversity. Only through a deeper understanding of this diversity and its implications for HIV prevention, vaccine development, and antiretroviral therapy will we be able to end the pandemic.

Acknowledgements

Supported by grants from the National Institutes of Health (training grant T32 AI49821-06, to Dr. Taylor; and Clinical Trials Unit Grant U01 AI069470-01, to Drs. Sobieszczyk and Hammer).

References


35. Laurent C, Bourgeois A, Faye MA, et al. No difference in clinical progression between patients infected with the predominant human immunodeficiency virus type 1 circulating recombinant form...


40. Laeyendecker, O.; Li, X.; Arroyo, M., et al. The effect of HIV subtype on rapid disease progression in Rakai, Uganda. Presented at the 13th Conference on Retroviruses and Opportunistic Infections; Denver. February 5–8, 2006; abstract


44. de Brito A, Komminakiss SC, Novoa P, et al. Women infected with HIV type 1 Brazilian variant, subtype B (B′-GWGR motif) have slower progression to AIDS, compared with patients infected with subtype B (B-GPGR motif). Clin Infect Dis 2006;43:1476–81. [PubMed: 17083025]


N Engl J Med. Author manuscript; available in PMC 2009 January 6.


Figure 1. Evolution of Diversity in HIV-1 during the Typical Viral Life Cycle and Creation of Unique Recombinant Forms in the Context of Coinfection with Two Subtypes

RT denotes reverse transcriptase.
Figure 2.
Current Global Distribution of HIV-1 Subtypes and Recombinant Forms.
Figure 3. Predominant Amino Acid (AA) Changes Conferred by Polymorphisms in HIV-1 Protease, According to Subtype

AA sequences for HIV-1 protease were compared with a consensus subtype B sequence. Polymorphism data were obtained in May 2007 with the use of the HIVseq program from the Stanford HIV Drug Resistance Database (http://hivdb.stanford.edu/pages/algs/HIVseq.html). Polymorphisms in AAs 10 through 99 of the HIV-1 protease are shown for a given subtype if 10% or more of the sequences in the database were polymorphic at that site; the percentage of sequences that were polymorphic appear within the colored bars. The subtype is shown in the key, followed by the number of sequences in the analysis in parentheses. Black rectangles indicate the sites of major protease resistance mutations, and white rectangles indicate the sites of minor protease resistance mutations, according to the International AIDS Society–USA in 2007. The predominant AA changes at each site are shown on the left-hand side.

Polymorphisms in HIV-1 reverse transcriptase AAs 40 through 240 are shown in the Supplementary Appendix, available with the full text of this article at www.nejm.org.
**Table 1**

Phylogenetic Classifications of HIV-1.

<table>
<thead>
<tr>
<th>Classification</th>
<th>Definition</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subtypes or clades</td>
<td>Genetically related HIV-1 strains that are essentially phylogenetically equidistant, generating a starlike, rather than a treelike, phylogeny</td>
<td>Subtypes A, B, C, D, F, G, H, I, and K are currently known; A through D are highly prevalent, others have low prevalence and limited geographic distributions</td>
</tr>
<tr>
<td>Subtypes</td>
<td>Distinct lineages within a subtype; genetic distance between sub-types is smaller than that between sub-types</td>
<td>Subtypes A and F are subdivided into sub-types A1 through A4 and F1 and F2, respectively; mostly these circulate in Central and West Africa</td>
</tr>
<tr>
<td>Intersubtype recombinant forms</td>
<td>Mosaic strains with segments from two or more subtypes alternating across the genome</td>
<td>Common in mixed-subtype epidemics; thought to result from infection of a person with more than one HIV-1 subtype</td>
</tr>
<tr>
<td>Circulating recombinant forms</td>
<td>Specific recombinant forms that are spreading in a population; new forms are defined when three people without direct epidemiologic linkage are found to be infected; the assigned name reflects sequence of discovery and subtype composition, with “cpx” indicating forms containing three or more subtypes</td>
<td>Currently, 43 forms are described; CRF01_AE and CRF02_AG are found principally in Southeast Asia and West Africa, respectively; others have more limited distributions</td>
</tr>
<tr>
<td>Unique recombinant forms</td>
<td>Intersubtype recombinant forms recovered from only a single person</td>
<td>Hundreds of forms have been described on the basis of partial or complete genome sequences; their potential for epidemic spread is unknown</td>
</tr>
<tr>
<td>Geographically distinct lineages</td>
<td>Lineages, often country-specific, that are distinguishable phylogenetically; unlike sub-subtypes, they are not phylogenetically equidistant within subtypes</td>
<td>Thai B, Indian C, West vs. East African D, and Former Soviet Union A (FSU-A)</td>
</tr>
<tr>
<td>Subtype or CRF</td>
<td>Location</td>
<td>Global Prevalence</td>
</tr>
<tr>
<td>---------------</td>
<td>----------</td>
<td>-------------------</td>
</tr>
<tr>
<td>A Subtype</td>
<td>East and Central Africa, Central Asia, Eastern Europe</td>
<td>12.3%</td>
</tr>
<tr>
<td>B Subtype</td>
<td>Americas, Western Europe, East Asia, Oceania</td>
<td>10.2%</td>
</tr>
<tr>
<td>C Subtype</td>
<td>India, Eastern and Southern Africa</td>
<td>49.9%</td>
</tr>
<tr>
<td>D Subtype</td>
<td>East Africa</td>
<td>2.5%</td>
</tr>
<tr>
<td>G Subtype</td>
<td>West Africa</td>
<td>6.3%</td>
</tr>
<tr>
<td>F, H, J, and K CRF</td>
<td>Various</td>
<td>Each &lt;1.0%</td>
</tr>
<tr>
<td>CRF01_AE CRF</td>
<td>Southeast Asia</td>
<td>4.7%</td>
</tr>
<tr>
<td>CRF02_AG CRF</td>
<td>West Africa</td>
<td>4.8%</td>
</tr>
<tr>
<td>Other CRF</td>
<td>Various</td>
<td>Each &lt;0.1%</td>
</tr>
</tbody>
</table>

* Location and prevalence data are from Hemelaar et al. Other CRFs include CRF03 through CRF43, and this category is expanding. CTL denotes cytotoxic T lymphocyte, and NA not available.
Table 3
Details of Selected Studies of Mutations Conferring Primary Resistance to Antiretroviral Therapy.*

<table>
<thead>
<tr>
<th>Study</th>
<th>Location and Year</th>
<th>No. of Subjects</th>
<th>Prevalence of Subtypes or CRFs</th>
<th>Prevalence of Known Resistance Mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nkengafac et al. 68</td>
<td>Cameroon, 2005–2006</td>
<td>180</td>
<td>B (42%), G (29%), other (29%)</td>
<td>8%</td>
</tr>
<tr>
<td>Palma et al. 69</td>
<td>Portugal, 2003</td>
<td>180</td>
<td>B (42%), G (29%), other (29%)</td>
<td>8%</td>
</tr>
<tr>
<td>Tee et al. 70</td>
<td>Malaysia, 2003–2004</td>
<td>100</td>
<td>CRF01_AE (65%), AE/B (22%), B (12%)</td>
<td>1%</td>
</tr>
<tr>
<td>Paraskevis et al. 71</td>
<td>Greece, 2002–2003</td>
<td>101</td>
<td>B (48%), A (33%)</td>
<td>9%</td>
</tr>
<tr>
<td>Ly et al. 72</td>
<td>Cambodia, 2003–2004</td>
<td>144</td>
<td>CRF01_AE (97%)</td>
<td>5%</td>
</tr>
<tr>
<td>Descamps et al. 73</td>
<td>France, 2001–2002</td>
<td>662</td>
<td>B (71%), non-B (29%)</td>
<td>12% in patients with acute infection, 2% in those with chronic B infection, lower in those with non-B infection</td>
</tr>
<tr>
<td>Wensing et al. 74</td>
<td>Europe, 1996–2002</td>
<td>2208</td>
<td>B (70%), C (10%)</td>
<td>13% in patients infected with B, 5% in those infected with non-B</td>
</tr>
<tr>
<td>Vazquez de Parga et al. 75</td>
<td>Former Soviet Union, 1997–2004</td>
<td>278</td>
<td>A (80%)</td>
<td>13% for reverse transcriptase mutations, 4% for protease mutations None detected</td>
</tr>
<tr>
<td>Roudinskii et al. 76</td>
<td>Former Soviet Union, 1995–2003</td>
<td>119</td>
<td>A (97%)</td>
<td>50% in patients infected with C, 10% in those infected with B</td>
</tr>
<tr>
<td>Grossman et al. 49</td>
<td>Israel, 1999–2003</td>
<td>117</td>
<td>C (75%), B (25%)</td>
<td>&lt;2% in those infected with B</td>
</tr>
<tr>
<td>Deshpande et al. 77</td>
<td>India, 2003</td>
<td>128</td>
<td>C (96%)</td>
<td>9%</td>
</tr>
<tr>
<td>Maljkovic et al. 78</td>
<td>Sweden, 1998–2001</td>
<td>100</td>
<td>B (55%), C (29%)</td>
<td>5%</td>
</tr>
</tbody>
</table>

* Studies that included at least 100 isolates were selected during a literature review in May 2007. Minor protease mutations were not included in the calculation of the prevalence of known resistance mutations. CRF denotes circulating recombinant form.