



FIGURE 226-5 Organization of the genome of the HIV provirus together with a summary description of its 9 genes encoding 15 proteins. (Adapted from WC Greene, BM Peterlin: *Nat Med* 8:673, 2002.)

involved in gene expression (Fig. 226-5). The major difference between the genomes of HIV-1 and HIV-2 is the fact that HIV-2 lacks the *vpu* gene and has a *vpx* gene not contained in HIV-1.

MOLECULAR HETEROGENEITY OF HIV-1

Molecular analyses of HIV isolates reveal varying levels of sequence diversity over all regions of the viral genome. For example, the degree of difference in the coding sequences of the viral envelope protein ranges from a few percent (very close, among isolates from the same infected individual) to more than 50% (extreme diversity, between isolates from the different groups of HIV-1: M, N, O, and P). The changes tend to cluster in hypervariable regions. HIV can evolve by several means, including simple base substitution, insertions and deletions, recombination, and gain and loss of glycosylation sites. HIV sequence diversity arises directly from the limited fidelity of the reverse transcriptase. The balance of immune pressure and functional constraints on proteins influences the regional level of variation within proteins. For example, Envelope, which is exposed on the surface of the virion and is under immune selective pressure from both antibodies and cytolytic T lymphocytes, is extremely variable, with clusters of mutations in hypervariable domains. In contrast, reverse transcriptase, with important enzymatic functions, is relatively conserved, particularly around the active site. The extraordinary variability of HIV-1 contrasts markedly with the relative stability of HTLV-1 and -2.

The four groups (M, N, O and P) of HIV-1 are the result of four separate chimpanzee-to-human (or possibly gorilla-to-human for groups O and P) transfers. Group M (major), which is responsible for most of the infections in the world, has diversified into subtypes and intersubtype recombinant forms, due to “sub-epidemics” within humans after one of those transfers.

Among primate lentiviruses, HIV-1 is most closely related to viruses isolated from chimpanzees and gorillas (Fig. 226-1). The chimpanzee subspecies *Pan troglodytes troglodytes* has been established to be the natural reservoir of the HIV-1 M and N groups. The rare viruses of the HIV-1 O and P groups are most closely related to viruses found in Cameroonian gorillas. The M group comprises nine subtypes, or *clades*, designated A, B, C, D, F, G, H, J, and K, as well as more than 60 known circulating recombinant forms (CRFs) and numerous unique recombinant forms. Intersubtype recombinants are generated by infection of an individual with two subtypes that then recombine and create a virus with a selective advantage. These CRFs range from highly prevalent forms such as CRF01_AE, common in southeast Asia, and CRF02_AG from west and central Africa, to a large number of CRFs that are relatively rare, either because they are of a more recent origin (newly recombined) or because they have not broken out into a major population. The subtypes and CRFs create the major lineages of the M group of HIV-1. HIV-1 M group subtype C dominates the global pandemic, and there is much speculation that it is more transmissible than other subtypes, but solid data on transmissibility variations between subtypes are rare. Human population densities, access to prevention and treatment, prevalence of genital ulcers, iatrogenic transmissions, and other confounding host factors are all possible reasons why one subtype has spread more than another.

Figure 226-6 schematically diagrams the worldwide distribution of HIV-1 subtypes by region. Seven strains account for the vast majority of HIV infections globally: HIV-1 subtypes A, B, C, D, G and two of the CRFs, CRF01_AE and CRF02_AG. Subtype C viruses (of the M group) are by far the most common form worldwide, likely accounting for ~50% of prevalent infections worldwide. In sub-Saharan Africa, home to approximately two-thirds of all individuals living with HIV/AIDS, the majority of infections are caused by