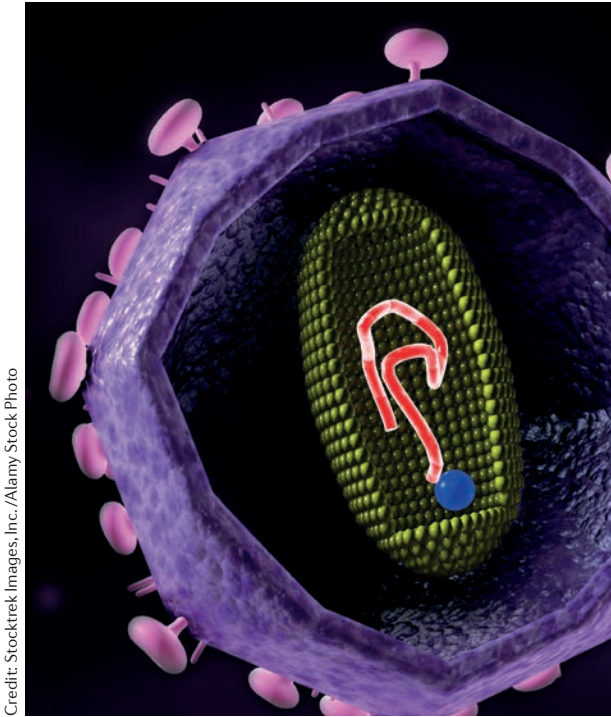


MILESTONE 5

From sequence to proteins



Credit: Stocktrek Images, Inc./Alamy Stock Photo

1985 marked the year when the full nucleotide sequence of HIV-1 was reported by three groups, a development that was instrumental to further understanding of the genetics and molecular biology of the virus. Ratner et al., Sanchez-Pescador et al. and Wain-Hobson et al. were the first to describe the full DNA sequence and genome organization of viral isolates. The complete sequences (>9,000 kb in length) were derived from proviral DNA and circular unintegrated viral DNA, and they encompassed the long terminal repeats (LTRs), which have crucial roles in the regulation of transcription of viral genes and integration. It is now established that the viral genome encodes the capsid proteins (Gag), viral enzymes (Pol) and the envelope glycoprotein (Env), as well as six additional open reading frames. By determining the locations and sizes of the viral open reading frames, it was revealed that the fundamental genetic structure is similar to that of other retroviruses, but that HIV-1 not only has distinctive genetic complexity but

“ Understanding the genetic structure of the virus led to new insights into the regulation of viral gene expression and RNA export ”

also encodes genes with features not previously recognized in biology.

Understanding the genetic structure of the virus led to new insights into the regulation of viral gene expression and RNA export.

An important discovery following rapidly on the heels of the full sequence information was the finding that HIV-1 encodes a trans-acting factor, termed Tat, which was shown to be vital for the transactivation of viral gene expression from the 5' LTR. During the following years, several studies revealed the mechanism of Tat-mediated transactivation: early during infection, low levels of viral transcripts are generated, which are subsequently spliced and translated to make Tat. Tat binds to an RNA stem-loop structure, the trans-activation response element (TAR), a regulatory element located downstream of the transcriptional initiation site at the 5' end of nascent viral transcripts. Following binding to TAR, Tat recruits the positive transcription elongation factor b (P-TEFb), a host factor that comprises cyclin-dependent kinase 9 (CDK9) and cyclin T1 as well as other elongation factors. Cyclin T1 binds directly to Tat and CDK9 phosphorylates the C-terminal domain of RNA polymerase II, thus promoting efficient transcriptional elongation. Smaller fully spliced messages, such as those encoding Tat, are exported readily from the nucleus to the cytoplasm and are translated, whereas unspliced and incompletely spliced mRNAs require the action of Rev (regulator of expression of virion

proteins), a regulatory HIV-1 protein that is also expressed during the early phase of infection.

The mechanism of Rev-dependent export of HIV-1 mRNA species became apparent in the late 1980s. Rev induces the sequence-specific nuclear export of late-phase HIV-1 mRNA species and promotes the cytoplasmic expression of HIV-1 mRNAs that encode viral accessory and structural proteins, including Gag and Env. The initial step in this pathway involves binding of Rev to the Rev-response element (RRE; a stem-loop structure that is present in intron-retaining viral mRNAs) in a highly cooperative manner. An important finding was that RRE-bound Rev forms a complex with cellular nuclear export factor CRM1 through its nuclear export signal. This interaction enables CRM1 to transport the mRNA–Rev complex into the cytoplasm for ensuing translation.

The description of the full nucleotide sequence enabled remarkable discoveries that revealed how gene expression in HIV-1 is controlled by the HIV-1 RNA-binding proteins Tat and Rev and how the virus hijacks the core molecular machinery of the host during viral replication, using mechanisms that were unprecedented at that time. The sequencing work also set the stage for further discoveries regarding the origins and diversity of the virus (MILESTONE 7).

Andrea Du Toit,
Nature Reviews Microbiology

ORIGINAL ARTICLES Ratner, L. et al. Complete nucleotide sequence of the AIDS virus, HTLV-III. *Nature* **313**, 277–284 (1985) | Sanchez-Pescador, R. et al. Nucleotide sequence and expression of an AIDS-associated retrovirus (ARV-2). *Science* **227**, 484–492 (1985) | Wain-Hobson, S., Sonigo, P., Danos, O., Cole, S. & Alizon, M. Nucleotide sequence of the AIDS virus, LAV. *Cell* **40**, 9–17 (1985)

FURTHER READING Arya, S. K., Guo, C., Josephs, S. F. & Wong-Staal, F. Trans-activator gene of human T-lymphotropic virus type III (HTLV-III). *Science* **229**, 69–73 (1985) | Sodroski, J., Patarca, R., Rosen, C., Wong-Staal, F. & Haseltine, W. Location of the trans-activating region on the genome of human T-cell lymphotropic virus type III. *Science* **229**, 74–77 (1985) | Feinberg, M. B., Aldovini, A., Gallo, R. C. & Wong-Staal, F. HTLV-III expression and production involve complex regulation at the levels of splicing and translation of viral RNA. *Cell* **46**, 807–817 (1986) | Kao, S. Y., Calman, A. F., Luciw, P. A. & Peterlin, B. M. Anti-termination of transcription within the long terminal repeat of HIV-1 by *tat* gene product. *Nature* **330**, 489–493 (1987) | Wei, P., Garber, M. E., Fang, S. M., Fischer, W. H. & Jones, K. A. A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA. *Cell* **92**, 451–462 (1998) | Malim, M. H., Hauber, J., Fenrick, R. & Cullen, B. R. Immunodeficiency virus Rev trans-activator modulates the expression of the viral regulatory genes. *Nature* **335**, 181–183 (1988) | Malim, M. H., Hauber, J., Le, S. Y., Maizel, J. V. & Cullen, B. R. The HIV-1 Rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. *Nature* **338**, 254–257 (1989) | Malim, M. H., Böhnlein, S., Hauber, J. & Cullen, B. R. Functional dissection of the HIV-1 Rev trans-activator—derivation of a trans-dominant repressor of Rev function. *Cell* **58**, 205–214 (1989) | Fukuda, M. et al. CRM1 is responsible for intracellular transport mediated by the nuclear export signal. *Nature* **390**, 308–311 (1997) | Stade, K., Ford, C. S., Guthrie, C. & Weis, K. Exportin 1 (Crm1p) is an essential nuclear export factor. *Cell* **90**, 1041–1050 (1997)