Genomic Diversity of Human Immunodeficiency Viruses

Lt Col RM Gupta*, Lt Col AK Sahni[†], Col J Jena[#], Col SK Nema^{**}

Abstract

Background: Globally circulating strains of human immunodeficiency virus type one (HIV-1) exhibit an extraordinary degree of genetic diversity. Sequences derived from HIV-1 strains have historically been classified on the basis of their phylogenetic relationships. The viruses have been classified into groups, subtypes or clades and circulating recombinant forms (CRFs). Groups were originally named M for main, O for outlier and N for Non-M-Non-O. The identification of subtypes and CRFs provides a means of tracking the dissemination of the pandemic.

Methods: Various methods to study the molecular epidemiology of HIV-1 are virus isolation, cloning, DNA sequencing, restricted fragment length polymorphism of the molecularly cloned and amplified genome (PCR- RFLP), RNase mismatch cleavage analyses of RNA, RNA heteroduplexes derived from culture amplified virus, primer mismatch sensitive PCR to identify specific mutations, single strand conformational polymorphism (SSCP) to localize mutations arising over short regions of env gene, denaturing gradient gel electrophoresis and serological assays based on V3 peptide. Except for PCR-RFLP and denaturing gradient gel electrophoresis, these techniques do not easily allow simultaneous analyses of multiple sequence variants and include the laborious and selective process of virus co-cultivation or molecular cloning prior to analyses. The extensive DNA sequence analyses remain the gold standard for epidemiological investigations.

Conclusions: Both HIV-1 and HIV-2 are present in India. The Indian strains of HIV1 also show diverse subtypes with HIV1 subtype C predominance. Tracking the genetic diversity has implications towards understanding the evolution of the epidemic, immunopathogenesis, natural course of infection, response to therapy and most importantly vaccine design.

MJAFI 2005; 61 : 267-270

Key words : Genetic diversity; Human immunodeficiency virus type one (HIV-1); Subtypes

Introduction

Globally circulating strains of human immunodeficiency virus type one (HIV-1) exhibit an extraordinary degree of genetic diversity, which may influence aspects of their biology such as infectivity, transmissibility and immunogenicity. Molecular analyses of various HIV isolates reveal sequence variations over many parts of the viral genome.

The Genesis of viral diversity

One of the major characteristics of HIV is the extremely high genetic variability. This results from two major factors. Firstly, there have been multiple introductions of genetically diverse simian viruses into humans [1]. Secondly, this diversity is a result of the extremely high error rate of reverse transcriptase (RT) [2]. The rapid turnover of virions in HIV infected individuals is another important factor [3]. Recombination is the fundamental property of retroviruses because of their dimeric RNA genome and a RT that can switch between templates during proviral DNA synthesis. RT is known to be highly recombinogenic [4]. This entails that radically different genomic combinations may be generated in individuals

infected by genetically different viruses.

Current Nomenclature

Sequences derived from these HIV-1 strains have been classified on the basis of their phylogenetic relationships. The groups were originally named M for main, O for outlier and N for Non-M-Non-O [5,6,7]. Group M includes the viruses that dominate the global epidemic. Subtypes refer to the major clades within group M. Sub-subtype refers to a distinctive lineage that is very closely related to a particular subtype lineage and is not genetically distant enough to justify calling a new subtype. Circulating Recombinant Form (CRF) describes a recombinant lineage that plays an important role in the HIV-1 pandemic. General requirements for defining a new subtype, sub-subtype or CRF are that representative strains must be identified in at least three individuals with no direct epidemiological linkage. Three near full-length genomic sequences are preferred.

HIV-1 genetic subtypes

Group M viruses are by far the most widespread [8,9], being the variants of HIV-1 that are responsible for more than 99% infections worldwide. The M-group viruses

*Associate Professor, [†]Professor and Head, Department of Microbiology, ^{**}Professor and Head, Department of Pathology, Armed Forces Medical College, Pune-40, [‡]Classified Specialist, Pathology (Microbiology), DRDE Gwalior.

have been divided into distinct genetic subtypes or clades. Currently ten different subtypes and two sub-subtypes have been identified (A through K) within group M [10,11]. Prototype viruses representing the genetic subtypes E and I have not yet been found. The viruses originally identified as subtype E (the predominant group of viruses involved in heterosexual transmission in Thailand) and subtype I (a small group of viruses from the Mediterranean region) are now considered intersubtype recombinants.

Geographical distribution of HIV-1 subtypes

The contribution of travel to the spread of HIV-1 around the world is well illustrated by the first documented HIV-1 infection in Europe, which was retrospectively identified in a Norwegian family. The father, a seaman probably infected by HIV-1 group O in the early 1960s through heterosexual contacts in a West African seaport, transmitted the infection to his wife who transmitted the infection to her daughter. They all died in 1976. This observation also underlines that depending on life style, HIV-1 spread can be limited to a few individuals and need not lead to a local epidemic.

Globally, group M viruses predominate. Of these, HIV-1 subtype A is found primarily in Central Africa, subtype B in North America, Western Europe, Australia and Japan, subtype C in South Africa and India, subtype D in Central Africa, subtype E in Thailand and subtype F includes isolates from Brazil and Romania. Other sequence subtypes (G, H, I) include viruses from Africa, Russia and Taiwan. Subtype K, has been reported from Cameroon [12]. The geographical distribution of various subtypes of HIV is shown in Fig. 1.

Since most studies concern western countries where HIV-1 B subtype is predominant, the global view on the contribution of travel to HIV-1 spread usually derives from the prevalence of non-B subtypes in these countries [13]. The prevalence of non-B infection has indeed markedly increased in recent years in several European countries. There is no doubt that behavior and not occupation, determines HIV-1 infection risk. However, some groups of individuals, particularly travelers,

contribute to the initiation of local epidemics or in the increase in HIV-1 diversity worldwide. They include, in particular, immigrants, intravenous drug abusers, tourists, truck drivers, military troops and seamen. In large countries such as Russia, China, India, Brazil and South Africa, internal migrants contribute largely to the spread of HIV-1 infection and resultant diversity.

Globally, subtypes A and C account for most current infections followed by subtype B. Subtype B is dominant in Europe, the Americas and Australia (which accounts for the emphasis that was placed on this subtype in the early-to-middle years of the AIDS pandemic) [14]. Subtype C is probably the subtype that currently infects more people worldwide than any other and it is common in South Africa and India. Sub-types A and D infect large numbers of people in central and eastern Africa. In China, inter-subtype recombinants between subtypes C and B are becoming common. Of course, as HIV-1 continues to spread globally, the geographical restrictions are increasingly breaking down.

Recombinant forms of HIV1

A study of isolates from the Democratic Republic of Congo indicates Central Africa as the epicenter of HIV-1 diversity, with a large number of different genetic subtypes and subtype recombinants circulating. The prevalence of intersubtype recombinant strains is increasing and creates even more HIV-1 antigenic diversity [15]. Several recombinant viruses have now spread epidemically to establish distinct lineages.

Recombination events between segments from two different viral stains in the same individual have resulted in the emergence of CRFs. Some CRFs are today the predominant virus in West Africa and Thailand. Overall, CRFs constitute 10–20% of newly characterized strains, and recombination between CRFs has also been reported. 14 CRFs have so far been identified. The definition of a CRF is restrictive since, to be accepted as a recombinant-circulating lineage, a recombinant form has to be fully sequenced and identified in three epidemiologically unrelated individuals [16]. The entire M group subtypes and the CRFs derived from them, can be traced back to a single successful natural transfer of HIV-1 to a human from a chimpanzee infected with simian immunodeficiency virus SIVcpz. This event occurred sometime in the first half of the 20th century, somewhere in Central Africa.

The Indian scenario

As per UNAIDS, there are localized epidemics within high-risk groups in India, and the virus is spreading to the general population [17]. Nearly one percent (or 4.5 million persons) of the adult population is estimated to be infected with HIV. In six states, Andhra Pradesh,

Fig. 1 : Geographical distribution of HIV-1 subtypes

Karnataka, Maharashtra, Manipur, Nagaland and Tamil Nadu more than one percent of the general population is HIV positive based on antenatal clinic surveys. Almost 90 percent of the cases reported, fall within the most productive age group of 15 to 44 years. Every one in four cases of HIV in India is seen among women. HIV sentinel surveillance surveys indicate infection rates between 1 to 2 per cent among antenatal mothers. Given India's large population, a mere 0.1 percent increase in the prevalence rate would increase the number of adults living with HIV/AIDS by over half a million persons. HIV/AIDS is already affecting children.

HIV infection is spreading in India by multiple modes of transmission. Heterosexual transmission remains the predominant mode of transmission. An alarming aspect of the present day scenario in India is that the infection is rapidly spreading to the remote rural areas through migrant workers and truck drivers. Though the actual scenario of HIV infection in India is extremely grim, the molecular epidemiological information is still limited. Till date, a few preliminary studies that have been conducted on these lines have all suffered from some drawbacks like small sample size, lack of proper representations from various geographical regions and inclusion of different risk groups and modes of transmission. Initial studies have indicated that both HIV-1 and HIV-2 are present in India. Later studies have indicated the presence and predominance of HIV-1 subtype C in India [18,19].

Implications of molecular heterogeneity of HIV-1

It is important to emphasize that the genetic subtypes or recombinant lineages of HIV-1 are not analogous to classic viral serotypes. The HIV-1 genetic diversity currently present in the human population dwarfs anything that has been described for other human viral infections. In HIV-1, replicating viruses can differ as much as 10% in sequence even within a single individual. Therefore, even within a genetic subtype, the extent of HIV-1 genetic and antigenic diversity is simply enormous when compared to the diversity found for viruses for which effective vaccines have been developed. The degree of genetic and hence antigenic diversity therefore is daunting from the perspective of HIV-1 vaccine development.

Knowing that genetic variants can escape immunosurveillance, information on genetic subtypes becomes important. Nonetheless, if phylogenetically defined groups broadly correspond to antigenic groups, vaccination using the most appropriate (antigenically related) strain will require information regarding the genetic subtypes of the most likely challenge strains found

regionally. Secondly, distribution of subtypes helps us understand the global HIV epidemic.

The subtype prevalence and degree of inter-isolate variation in any region or group also provides us with clues about the origin and spread of HIV. Thirdly, in face of this global variation of HIV-1, little is known about the genetic diversity correlation to the biological properties of the virus. For example, in Thailand, HIV-1 subtype B and E has caused two parallel epidemics. The subtype B sequence, predominating in the intravenous drug user population is different from the North-European subtype B strains and is generally referred to as subtype B. The subtype E strains which are actually recombinants of subtype A and E with the *gag* gene of subtype A and the *env* gene of subtype E are associated mostly with heterosexual transmission.

Reasons, why genetic subtypes and inter subtype recombinant genomes need to be analyzed are compelling. These include molecular epidemiological studies to elucidate the geographic distribution of strains co-circulating in a particular region, studies of historical evolution of strains in a population group; detection of divergent strains being introduced, developing subtype specific serological techniques such as synthetic peptide enzyme immunoassays (EIAs) for accurate diagnosis and improved sensitivity and specificity, studies of patterns of transmission and degree of transmissibility amongst variant strains to delineate the host range and specificity, viral phenotypes like NSI and SI strains and their relation to disease progression, studies of drug resistance patterns amongst subtypes and amongst quasiespecies in an individual over a period of time and last but not the least in vaccine designs based on individual sequence subtypes. Moreover, information on the presence of HIV multiple infections caused by distinct viral subtypes circulating within a population and the long-term observation of patients dually infected with distinct HIV strains may contribute to a better understanding of pathogenesis of HIV infections. To study the magnitude of HIV-1 epidemic in India, more extensive molecular epidemiological analysis will be required from high-risk population from different geographical regions of the country, to substantiate the breadth of subtype representation in individual regions. Sequencing of full-length genes, rather than sub genomic fragments of the viral genome, is the order of the day.

References

1. Hahn BH, Shaw GM, De Cock KM, Sharp PM. AIDS as a zoonosis: scientific and public health implications. *Science* 2000; 287(5453): 607-14.
2. Preston BD, Poiesz BL, Loeb LA. Fidelity of HIV-1 reverse transcriptase. *Science* 1988; 242: 1168-71.
3. Ho DD, Neumann AU, Perelson AS, Chen W, Leonard J M,

- Markowitz M. Rapid turnover of plasma virion and CD4 lymphocytes in HIV-1 infection. *Nature* 1995; 373(6510): 123-6.
4. Hu W-S, Temin HM. Genetic consequences of packaging two RNA genomes in one retroviral particle: pseudodiploidy and high rate of genetic recombination. *Proc Natl Acad Sci* 1990; 87: 1556-60.
 5. De Leys R, Vanderborght B, Vanden Haesevelde M, et al. Isolation and partial characterization of an unusual human immunodeficiency retrovirus from two persons of West-Central African origin. *J Virol* 1990; 64: 1207-16.
 6. Charneau P, Borman AM, Quillent C, et al. Isolation and envelope sequence of a highly divergent HIV-1 isolate: definition of a new HIV-1 group. *Virology* 1994; 15: 205(1): 247-53.
 7. Simon F, Mauclore P, Roques P, et al. Identification of a new human immunodeficiency virus type 1 distinct from group M and group O. *Nat Med* 1998; 4:1032-7.
 8. Robertson DL, Anderson JP, Bradac JA, et al. HIV-1 nomenclature proposal. *Science* 2000; 288:55-6.
 9. Peeters M and Sharp PM. Genetic diversity of HIV-1: the moving target. *AIDS* 2000; 14 Suppl 3:S129-40.
 10. Perrin L, Kaiser L, Yerly S. Travel and the spread of HIV-1 genetic variants. *Lancet Infect Dis* 2003; 3(1): 22-7.
 11. Thomson MM, Perez-Alvarez L, Najera R. Molecular epidemiology of HIV-1 genetic forms and its significance for vaccine development and therapy. *Lancet Infect Dis* 2002; 2:461-71.
 12. Delgado E, Thomson MM, Villahermosa ML, et al. Identification of a newly characterized HIV-1 BG intersubtype circulating recombinant form in Galicia, Spain, which exhibits a pseudotype-like virion structure. *J Acquir Immune Defic Syndr* 2002; 5:536-43.
 13. Roques P, Menu E, Narwa A, et al. An unusual HIV type 1 env sequence embedded in a mosaic virus from Cameroon: identification of a new env clade. *AIDS Res Hum Retroviruses* 1999; 15: 15851-9.
 14. Thomson MM, Najera R. Travel and the introduction of human immunodeficiency virus type 1 non-B subtype genetic forms into Western countries. *Clin Infect Dis* 2001; 32: 1732-7.
 15. Kuiken C, Thakallapalli R, Esklid A, Ronde A de. Genetic analysis reveals epidemiologic patterns in the spread of human immunodeficiency virus. *Am J Epidemiol* 2000; 152:814-22.
 16. Heyndrickx L, Janssens W, Zekeng L, et al. Simplified strategy for detection of recombinant human immunodeficiency virus type 1 group M isolates by Gag/Env heteroduplex mobility assay. *J Virol* 2000; 74:363-70.
 17. UNAIDS 2003. Report on Global HIV/AIDS epidemic. The Barcelona Report July 2003. (www.UNAIDS.org/barcelona/content.html).
 18. Maitra A, Singh B, Banu S, et al. Subtypes of HIV Type 1 circulating in India: Partial envelope sequences. *AIDS Res Hum Retroviruses* 1999; 15: 941-4.
 19. Sahni AK, Prasad VV, Seth P. Genomic diversity of human immunodeficiency virus type-1 in India. *Int J STD AIDS* 2002; 13(2): 115-8.
 20. Mandal D, Jana S, Bhattacharya SK, Chakrabarti S. HIV type 1 subtypes circulating in eastern and northeastern regions of India. *AIDS Res Hum Retroviruses* 2002; 18(16): 1219-27.

ANNOUNCEMENT

BEST ARTICLE AWARD 2004 : MJAFI

With effect from 1994 all Original Articles published in MJAFI are being screened for selection of the best two articles. These articles receive the 'Best Article Award' and the 'Second Best Article Award'. They carry a cash prize of Rs. 2000/- and Rs. 1000/- respectively to be shared by all authors. Articles are judged for their originality and research content.

So all those who believe that they have original work, not yet published, please send it in fast.

The following articles received the award for 2004 :

Best Article Award

Maj MP Cariappa, Maj Gen J Jayaram, Col Rajvir Bhalwar, Col AK Praharaj, Col VK Mehta, Lt Col LK Kapur "*Epidemiological Differentials of Hepatitis B carrier State in the Army : A Community Based Sero-epidemiological Study*". MJAFI;2004;60(3):251-254.

Second Best Article Award

Maj SK Roy Chowdhury, Brig S Murali Mohan "*Maxillofacial Injuries in Counter Proxy War Posture of the Armed Forces*". MJAFI;2004; 60(4): 357-362.