HIV Diversity and Classification, Role in Transmission

Duri Kerina¹*, Stray-Pedersen Babill², F. Muller³

¹Department of Immunology, University of Zimbabwe, Harare, Zimbabwe; ²Division of Women and Children, Oslo University Hospital, Rikshospitalet and Institute of Clinical Medicine, University of Oslo, Oslo, Norway; ³Department of Microbiology, University of Oslo and Oslo University Hospital, Rikshospitalet, Oslo, Norway.

Email: *kerina.duri@yahoo.com

Received November 1st, 2012; revised December 18th, 2012; accepted January 19th, 2013

Copyright © 2013 Duri Kerina et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT

The hallmark of HIV-1 is its extensive genetic diversity that emanates mainly from high mutations. Phylogenetically, HIV can be classified into geographically confined groups, types, subtypes and circulating recombinant forms (CRFs) that are however subject to change over time. HIV genetic diversity may partially explain the observed heterogeneity in HIV prevalence and has also been reported to impact on viral transmissibility and differential rates of disease progression. The aim of this review is to present a simple overview of the principles and concepts of HIV diversity and classification. Tracking the presence of new HIV strains is not only important for surveillance purposes but is also critical in facilitating personalized targeted therapy as well as forming the basis for development of the much anticipated effective vaccines against this scourge.

Keywords: HIV; Diversity; Classification; Transmission

1. Introduction

1.1. HIV-1 Genetic Diversity

The unique characteristics of human immunodeficiency virus type 1 (HIV-1) are its extensive genetic diversity [1-4]. Diversification is due to errors encountered during viral replication including host immune response selection pressures. Diversity is manifested as sequence variability particularly within the env variable (V) regions [5]. Variability not only makes it difficult for the immune system to identify the virus but it also facilitates the rapid viral immune escape. The high level of genetic diversity has important implications in screening, diagnostic testing, disease monitoring and treatment outcomes [6-14]. Questions have been raised on whether diversity may also affect viral transmissibility and pathogenicity [15-20]. Sadly, genetic diversity has been the major impediment in the effective vaccine design and development since the human immune response to HIV is strain-specific [21]. Four factors vis-a-vis, the infidelity of reverse transcriptase (RT), recombination, superinfection and high replication rate of the virus have been shown to contribute to the development of the extensive HIV genetic variation [22,23].

1.2. Properties of RT Enzyme and Recombination

The infidelity of HIV RT enzyme confers mutations at an approximate rate of one error per genome per replication cycle [24]. RT also accounts for genomic heterogeneity in progeny viruses through its role in recombination. Genetic recombination is an evolutionary strategy for survival in a changing environment for viral variants with superior fitness at an average of $1.38 \times 10^{-4}$ recombination events/adjacent sites/generation in vivo [25]. It occurs when an individual is co-infected with at least two different HIV strains that are multiplying concurrently in the same cell [26,27]. It is caused by high selection pressure from either the natural host immune response or antiretroviral drugs [28]. Recombinants between highly similar HIV-1 strains occur at highest frequencies while that between distant HIV-1 strains happen at very low frequencies [29]. Dual and even triple HIV-1 infections have been reported [30,31]. HIV superinfections also allow a mechanism for genetic recombinants between distant variants [32-37]. Superinfection and co-infection which both involve reinfection by at least two genetically distinct viral variants differ based on whether the secondary infection is contracted prior to or after the primary host immune response has been mounted [38]. These are as-
associated with high viral loads and accelerated rates of disease progression [39,40]. HIV-1 superinfection presents an additional concern to the already challenging problem of HIV-1 vaccine design in the face of the virus’s rapid evolution [41].

1.3. High Turnover Rates of HIV-1 in Vivo

HIV-1 virions are produced and cleared at an extremely rapid pace. Since the HIV-1 genome is about $10^3$ base pairs in length, then the baseline rate of viral production is approximately $10^{10}$ virions per day [42]. This rapid turnover has been considered as the major factor underlying the pathogenesis of HIV/AIDS alongside with the destruction of CD4+ T-helper lymphocytes [42]. Besides the viral RT, host RNA polymerase II makes minimal contributions to retroviral frame shift mutations [43].

Diversity may also be enhanced by different genetic factors including HLA in patients from different regions of the world. Viral genetic factors include proteins such as Tat, Vif and Rev that interact with human genetic factors such as APOBEC, langerin, tetherin and CCR5 and HLA B27, B57, DRB1*1303, KIR and PARD3B [44]. The inability of Vif to counteract host APOBEC3 proteins leads to the deamination of cytidine to uridine consequently causing viral guanosine to adenosine hypermutations [45]. Some error causing mechanisms contributing to HIV-1 variation are shown in Figure 1.

Genetic diversity helps the virus evade the immune system and consequently this viral heterogeneity allows for a quick adaptation to the human immune system, antiretroviral drugs, or both leading to viral fitness/positive selection in the face of pharmacologic or immunologic selection pressures [47]. Every day millions of genetic variants accumulate in latently infected cells only to be reactivated at some time in the future [48]. Thus, the extensive diversity of HIV resulting in a myriad of HIV variants has necessitated the need for its classification. This taxonomy facilitates better utilization of the ever growing viral sequence database through comparison with previously published work.

1.4. Classification of HIV

HIV-1 strains are not randomly distributed across the globe but they display a distinctive geographical distribution [49]. Prior to 1992, HIV-1 strains were classified into two main classes on the basis of their respective geographical origin being then, the North American and African variants [50]. Thus, HIV variation is highest among viruses from different geographical locations, higher among isolates from different individuals. However, variants present as relatively similar quasi-species within the same individual [51,52]. A quasi-species is a cloud or swarm of genetically diverse variants that are linked through mutations that interact cooperatively on a functional level but collectively contributing to the characteristics of a population [53].

With the advent of phylogenetic analysis the env gene has revealed the existence of multiple phylogenetic clusters that were used in the compilation of the 1992 HIV classification compendium based on envelope viral sequence similarities [54,55]. As the env sequence database increased over the years the gag and pol gene sequences were also incorporated in the classification procedure consequently, identifying HIV types, groups, subtypes, sub-subtypes and circulating recombinant forms (CRFs) [4,56-68], as summarised in Figure 2.

1.4.1. HIV Types

Phylogenetically HIV can be classified into two types; type 1 (HIV-1) and type 2 (HIV-2) [65]. Both viral types cause AIDS. HIV-1 is the first in the class of human retroviruses and accounts for most of the world’s HIV in-

---

**Figure 1.** A schematic sketch of error-causing machinery causing HIV-1 genetic diversity [46].
Infections. Its origin can be traced back to a Simian Immunodeficiency Virus (SIV) isolated from a Chimpanzee (cpz) sub-species, *Pan troglodytes troglodytes* (SIVcpz) cross species transmission to humans [70-72]. HIV-2 is the second in the same class of human retroviruses but is largely confined to West Africa. The primate reservoir of HIV-2 is sooty mangabey, *Cercocebus atys* (green monkey) [73,74]. Thus, (SIVCPZ) is closely related to HIV-1, while SIV from sooty mangabey (SIVSM) is closest to HIV-2 [75]. HIV-1 and HIV-2 are closely related viruses with nucleotide sequence homology of 58%, 59% and 39% in the *gag*, *pol* and *env* genes, respectively [76]. Despite similar modes of transmission, HIV-2 is not as efficient in transmission horizontally and vertically [65,77,78]. Relative to HIV-1, HIV-2 has a reduced rate of disease development and has shown natural resistance to readily available non-nucleoside reverse transcriptase inhibitors (NNRTIs) [79]. Genetic recombination between HIV types-1 and-2 has been reported [80]. Distinction of HIV types is essential for accurate surveillance, diagnosis as well as administration of appropriate antiretroviral therapies.

1.4.2. HIV Groups

Phylogenetic analysis of HIV-1 suggested that zoonosis occurred on at least three independent cross species transmission events from chimpanzees, *Pan troglodytes* (pts), *Pan troglodytes troglodytes* (pts) or gorilla (gor) resulting in four distinct HIV groups called the major (M), outlier (O), non-M/non-O (N) as shown in Figure 3. Studies have estimated the timing for the zoonosis of each lineage of groups M, O, and N at around 1931, 1920 and 1963, respectively [82]. Group M is responsible for more than 90% of the world HIV infections [83]. Interestingly, the genetic analysis of sequences from clinical materials obtained from members of a Norwegian family infected earlier than 1971 showed that they carried viruses of the group O mainly restricted to West Africa [84]. HIV groups have genetic sequence differences of >40% in some coding regions [85]. More recently, a new putative group, designated P, was reported in France from a Cameroonian female immigrant [86]. Group P viral sequences have been shown to form a distinct HIV-1 lineage with SIV sequences from western gorillas (SIVgor; *Gorilla gorilla gorilla*), suggesting that group P originated from gorillas [87] Reports have indicated that HIV-1 group P infections are rare, accounting for only 0.06% of HIV infections in Cameroon [88]. Unlike groups O, N or P, group M has been classified into subtypes.

1.4.3. HIV-1 Subtypes

Subtypes are phylogenetically linked strains of HIV-1 that are approximately the same genetic distance from one another. Group M has been classified into nine distinct subtypes, also called clades or genotypes, denoted with letters, A, B, C, D, F, G, H, J and K, thus making the development of effective blanket diagnostic and monitoring tests or vaccine a challenge [84,89]. Intersubtype variation is about 30% with respect to the *env* gene sequence and 15% for both the *gag* and *pol* genes sequences [90]. Different risk groups for HIV infection are associated with specific subtypes with intravenous drug users (IDUs) including the gay communities and heterosexual population generally acquiring subtype B and non-B subtypes, respectively [91-93].

1.4.4. HIV-1 Sub-Subtypes

Within each subtype numerous HIV-1 variants exist that exhibit minor intra-subtype genetic diversity of within 10% called sub-subtypes [94]. These are distinctive HIV-1 lineages that are closely related to a particular subtype lineage, but are not genetically distant enough to justify calling them new subtypes. Sub-subtypes are denoted by numerals for instance in the case for subtype A.
these have been named A1, A2 or A3 [95]. Recent studies have demonstrated the need for HIV classification using full-length genomic sequences if new distinctive subtypes are to be accurately identified rather than relying on sequencing of different viral gene fragments as has been the standard.

1.4.5. HIV Recombinants

Full genome sequencing of HIV has resulted in the discovery of circulating and unique recombinant forms (CRFs) and URFs, respectively. Recombinants are unique in the sense that they may be described in isolated individuals without any evidence of epidemic spread. To be classified as a CRF, a virus strain must be detected in at least three epidemiologically unlinked individuals and must be capable of establishing an epidemic on its own. Thus, these mosaic HIV-1 strains reflecting a mixture of subtypes circulating in different populations may have altered pathogenic and/or transmissibility properties [96]. CRFs are referred to by their number that is assigned according to the order of their discovery and the respective subtypes involved for example CRF02_AG or by their number(s) followed by the letters “cpx” (for complex), when more than two subtypes are involved for example CRF04_cpx or CRF06_cpx [97]. One of the most common group M CRFs is A/E, previously described as subtype E in Southeast Asia but was later renamed CRF01_AE following full HIV genome sequencing [98]. To date about 21 CRFs and several URFs have been described [99]. All CRFs together account for 18% of the world’s HIV-1 infections [100]. HIV-1 subtypes and recombinants may differ with respect to viral load levels [101], transcriptional activation levels, disease progression and response to antiretroviral therapy including drug induced/natural resistance patterns [102-105].

1.5. Distribution of HIV-1 Subtypes and Recombinants

Over 50 different subtypes and CRFs have been described [106,107]. Subtype B is geographically confined to North America, Western Europe and Australia. See Figure 4. Paradoxically subtype B is quite rare in Africa, the purported origin of HIV. Global proportions of HIV-1 subtypes and recombinants have shown that subtype C accounts for more than 50% of world’s infections followed by 12%, 10%, 6% and 3% for subtypes A, B, G and D respectively whilst subtypes F, H, J and K together account for about 0.94% of all the infections [90]. CRF01_AE and CRF02_AG are each responsible for 5% of infections while CRF03_AB is responsible for 0.1% of global infections with the other recombinants contributing to the remaining 8% of all HIV infections [90].

In most Southern African nations subtype C predominates, contributing 93% - 100% of the HIV-1 infections amongst individual countries [90,109]. Interestingly, the greatest diversity of subtypes and recombinants is present in Central Africa, Central African Republic, Gabon, Angola and Chad harboring only about 5% of the world’s infected individuals [109]. Thus a general observation is that a higher diversity of subtypes is associated with relatively slower epidemics whilst explosive epidemics generally have only one prevalent subtype.
1.5.1. Subtypes Trends and Distribution in Zimbabwe

Previous Zimbabwean studies in the 1990s and early 2000s have observed a predominant subtype C [110-112]. The origins and evolutionary history of HIV-1 subtype C in Zimbabwe with respect to the pol sequence data sets generated from four sequential cohorts of antenatal women in Harare, from 1991-2006 has demonstrated increasing sequence divergence over the 15-year period. This data also indicates a most recent common ancestor date of around 1973 with three epidemic growth phases: an initial slow phase (1970s) followed by exponential growth (1980s), and a linearly expanding epidemic to the present day [113]. However, current HIV subtype(s) distribution in Zimbabwe remain elusive in view of the influence of the population movements in the past decade as result of the economic meltdown which could have facilitated subtype inter mixing. Generally, subtype specific variations may exist that influence differential transmissibility in different regions [114-116].

1.5.2. HIV Diversity, Transmission and Disease Progression

Following sexual transmission of HIV the virus initially replicates locally in the vaginal or rectal mucosa [117]. Genetic diversity of HIV is lost during horizontal transmission and the virus gradually evolves towards a common ancestral sequence once in the new host [15,118]. Newly infected subjects acquire a subset of the viruses that were circulating in the transmitting partner; transmitted variants have less diversity and divergence [119]. Studies have correlated high HIV replication capacity with increased transmission rates [120]. Understanding the quantitative relationship between plasma HIV-1 RNA and HIV-1 transmission risk has been the cornerstone for ART preventive interventions that strive to reduce plasma HIV-1 levels that in turn reduce the risk of HIV-1 transmission [121]. Interestingly, Langerhans cells have shown minimal susceptibility to infection with subtype B virus but substantially greater sensitivity for infections by subtype C [122]. In the Rakai, Ugandan study, subtype A viruses have been shown to have a significantly higher rate of heterosexual transmission relative to subtype D viruses [123]. Differential subtype transmission efficiency may be important for HIV vaccine evaluation especially for the subtype-specific HIV epidemics in SSA. HIV-1-discordant couples are increasingly viewed as a valuable source of participants for HIV vaccine and prevention trials [124]. Curiously HIV-1 subtype C has been found to be the predominant subtype in sero-discordant couples followed by subtypes B and A, respectively [125].

Increasing HIV-1 replication efficiency has also been related to a concomitant increase in HIV-1 diversity, which in turn has been the determining factor in disease progression [126,127]. Non-A subtype infections have been shown to progress to AIDS faster than those infected with subtype A [128]. More so subtype D has been associated with the most rapid disease progression relative to subtypes A, C and CRFs [129-131]. Pregnancy has been shown to increase the risk of female-to-male HIV-1 transmission by two folds [132]. Pregnant women infected with subtype C have been shown to significantly shed more HIV-1-infected vaginal cells than were those infected with subtype A or D [133]. Increased HIV-1 shedding has been correlated with a more complex population of HIV-1 quasi-species in the genital tracts of parturient women, which may increase the probability of transmission of fetotropic strains [134]. Identifying the specific genetics characteristics of successfully transmitted variants is also paramount in the development of an effective vaccine. Subtype and CRFs determination is generally done using the gag/env heteroduplex mobility assay (HMA) originally developed by Delwart [135] and later modified by Heyndrick et al., in the year 2000 [136]. Whole genome sequencing remains the gold standard although partial sequencing also gives good results at reasonable cost.

2. Conclusion

The geographic distribution of subtypes is subject to constant change. Recombinant forms of the virus will continue to appear as long as the different subtypes of HIV-1 continue to circulate between continents and recombination continues to occur. With the world fast becoming a global village new HIV strains are emerging in areas where they were originally non-existent. Thus importation and exportation of new types, subtypes and even CRFs of HIV is possible. The risky behaviour of military personnel plus high HIV-1 sero-prevalence within this group may have facilitated the introduction of new HIV types, subtypes or recombinants within the armed forces themselves and to the general population both at home and abroad. Tracking the presence of new HIV strains is important for surveillance purposes, effective chemotherapy, diagnosis and disease monitoring including vaccine design and development. In the absence of effective prophylactic HIV vaccines, behavior change remains the key to successful prevention efforts.

3. Acknowledgements

We are grateful to Letten Foundation of Oslo, Norway for the sponsorship.

REFERENCES


HIV Diversity and Classification, Role in Transmission

pp. 153-159. doi:10.1097/01.aids.00002030-200401230-00003


[51] J. L. Sankale, R. S. De La Tour, R. G. Marlink, R. Scheib,
HIV Diversity and Classification, Role in Transmission


Copyright © 2013 SciRes.


Copyright © 2013 SciRes.