HIV-1 subtype C in Ethiopia: genotypic and phenotypic variation
Abebe, A.

Citation for published version (APA):
CHAPTER VII

Intra-subtype recombination of human immunodeficiency virus type 1 subtype C: evidence for advantage of a gp120 envelope subgroup

Georgios Pollakis, Almaz Abebe, Aletta Kliphuis, Bitew Fisseha, Belete Tegbaru, Girma Tesfaye, Hailu Negassa, Yohannes mengistu, Arnaud L. Fontanet, Tobias F. Rinke de Wit, Marion Cornelissen, and Jaap Goudsmit

Submitted for publication
ABSTRACT
We studied thirty randomly selected HIV-1 subtype C infected sera collected from several Ethiopian towns. The phylogenetic analysis of the 5’gag (including the p17 and 134 amino-acids of the p24 protein) and their respective C2V3 envelope regions showed evidence that five isolates are intra-subtype recombinants between the two groups of viruses, C and C’, co-circulating in Ethiopia. To substantiate this finding 16 isolates, including those suspected to be recombinants, were selected for further analysis. A 2600 nucleotide fragment comprising the gag and part of the pol (including protease and the 5’half of RT) genes was sequenced along with the respective V1V2/C2V3 envelope regions. By phylogenetic and bootscan analysis we have identified six intra-subtype recombinant viral isolates indicating that the recombination rate between the two groups of subtype C viruses is approximately 20%. All six recombinant viruses share the envelope regions of the C’ group strongly suggesting that the C’ group envelope has a biological advantage over the envelope of the C group. The neighboring-joining phylogenetic analysis of the 30 5’gag sequences showed that in Ethiopia two different groups of subtype C viruses are circulating as previously shown by the C2V3 sequence analysis. Furthermore, maximum likelihood sequence analysis of the 2600 nucleotide fragment together with the V1V2/C2V3 envelope regions established the genetic distinction between the two groups.

INTRODUCTION
Africa has been heavily affected by the human immunodeficiency virus Type 1 (HIV-1) pandemic with many of the known virus subtypes identified, including group M (subtypes A – K), group N and group O viruses, with group M subtype C being most prevalent. UNAIDS/WHO-2000 has reported that in sub-Saharan Africa there are approximately 24 million individuals infected with HIV-1 of which ten million are from countries where subtype C dominates. Recently the high incidence of subtype C infection has been reported in countries where non-subtype C viruses previously dominated. In Ethiopia the HIV/AIDS epidemic is believed to have originated in the early-1980s with the earliest HIV positive sera samples being collected in 1984, and this incidence has increased dramatically ever since. The prevalence of HIV-1 infection was estimated in 1988 at 17% amongst commercial sex workers and 13% amongst truck drivers, at 11% amongst pregnant women in Addis Ababa in 1991 and at 18% in 1997. In 1998 the prevalence amongst the commercial sex workers of Addis Ababa had risen dramatically to 74% (Aklilum, M.; personal communication). Unlike many other African countries the Ethiopian HIV-1 epidemic
is dominated exclusively by subtype C viruses with at least two distinct strains co-circulating, C and C'\textsuperscript{18,19}. The co-circulation of different viral strains within the same population can lead to co-infection and viral genome recombination and there have been many descriptions of recombinant viral genotypes world-wide \textsuperscript{20-27}. A report by Salminen et al. has provided evidence that recombination is likely to be an ongoing process since variant recombinant viruses could be identified at different time points from an individual infected with two different strains of HIV-1 \textsuperscript{21}. It has also been reported that differences in the RNA dimmerization signal of the HIV-1 genome does not impede the co-packaging of two heterologous RNA strands into one virion\textsuperscript{28}, suggesting that recombination during co-infection is likely. Approximately 10\% of all sequenced HIV-1 strains are composed of genetic material derived from the recombination between different subtypes \textsuperscript{27,29-33} and in some areas the incidence is estimated to be as high as 20-25\% \textsuperscript{21,22,27,34-36}.

In addition it has been shown that some recombinant viral isolates are mosaics of several subtypes \textsuperscript{31}. Several inter-subtype recombinant viruses have also been shown to cause local epidemics like the AE (CM240) recombinant virus identified in Southeast Asia \textsuperscript{22,28} the AG (IbNG) virus from West and West Central Africa \textsuperscript{37,38}, the AGI(CYO32) virus \textsuperscript{21} in Cyprus and the AB virus in Russia \textsuperscript{39}.

In this study we analysed HIV-1 amino acid sequences in sera isolated from 30 infected individuals across seven Ethiopian towns of broad geographic distribution. The presence of the two distinct co-circulating viruses in Ethiopia, C and C', was confirmed by the phylogenetic analysis of their 5' gag and the corresponding C2V3 envelope fractions of the genome. Intra-subtype recombination between the two co-circulating strains of virus was demonstrated. The frequency and pattern of recombination was found to be similar to that previously reported between two different subtypes. All the intra-subtype C/C' recombinants identified shared the envelope of the C' group indicating a transmission advantage for the C' envelope.

**MATERIALS AND METHODS**

**Study subjects and virus selection:** Stored sera samples collected from 30 HIV-1 infected individuals residing in different Ethiopian towns from 1988 and from 1996/7 were used in this study. The 1988 samples were collected from commercial sex workers and the 1996/7 samples were from blood donors. All samples were analysed for the C2V3 envelope and Gag region DNA sequences with sixteen samples selected for further analysis.
Isolation and PCR amplification of Gag, Pol and envelope regions: RNA was extracted from 100μl of serum by the Boom method and the cDNA amplified by a nested PCR protocol. The outer primers used for amplification of the C2-V3 region of the envelope were; 5'V3 NOT, 5'-GCAGGTCAGCAGACATGTACATCAGTGG-3' (HxB2 positions 6995-7020) and 3'-V3 NOT, 5'-GCCGGCGCAGCAGCCCTCTACAATTAATAGGTAAGT-3' (HxB2 positions 7390-7412). The inner primers were 5'V3-SP6, 5'-AATGGAAGCTACAGGAA-3' (HxB2 positions 7050-7070) and 3'V3-T7 5'-AATTTCTIGITCCCTCCG-3' (HxB2 positions 7368-7388), using the conditions previously described. The outer primers for the V1V2 region of the envelope were; 5'-GAGGATATAAATCAGTTTATGGG-3' (HxB2 positions 6476-6498) and 5'-ATTCCATGTGTACATTGTACTG-3' (HxB2 positions 6991-7023) and the inner primers were; 5'-GAGGATATAAATCAGTTTATGGG-3' (HxB2 positions 6538-6592) and 5'-GAGGATATAAATCAGTTTATGGG-3' (HxB2 positions 6833-6855) using the conditions previously described. The outer primers used for the amplification of 5' fragment of the Gag gene (including the p17 and 134 amino-acids of the p24 protein) were; 3'SK39, 5'-GCATTCTGGACATAAGACAAGGACCAAA-3' (HxB2 positions 1630-1658) and 5' Louw-1GAG, 5'-GCGAGAGCGTCAGTATTAAGC-3' (HxB2 positions 795-815). The inner Gag primers were; 5'GAG-2-SP6, 5'-GGGAAAAATTCGGTTAAAGGCC-3' (HxB2 positions 835-856) and 3'GAGAE3-T7, 5'-TAGGACCCTAATTTATTTATCA-3' (HxB2 positions 1587-1612) using the conditions previously described. For the amplification of 3' fragment of the Gag gene (including the 3' end of the p24, p7 and p6 proteins, that has 49 nt. overlap with the 5'Gag fragment and 49 nt. with the Pol gene) the outer primers were 5'-GGGGAAGTGACATAGCAGAACTA-3' (HxB2 positions 1483-1506) and 5'-TAATACTGTATCATCTGCTC-3' (HxB2 positions 1561-1583) and 5'-TACTGTGACAAGGTCGTGGCCA-3' (HxB2 positions 2267-2290) using the conditions described previously. For amplification of the Pol region (including the protease and 297 amino acids of the RT gene) the DNA was amplified by three nested PCR reactions (A, B and C) covering the full length of the region following the first PCR reaction. The outer primers were; 5'ProtFM, 5'-CAAGGGAAGGCGCAGGAAATTT-3' (HxB2 positions 2111-2130) and 3'halfRT, 5'-TGACCCCATAAAAGACTTATAAGCAGAAATA-3' (HxB2 positions 3505-3535). The inner primers used were; for A 5'SP6-prot/RT, 5'-CTTACCTCCCTCAGATC-3' (HxB2 positions 2242-2263) and 3'T7prot, 5'-CCTATTGAAACTGTACCAGTA-3' (HxB2 positions 2558-2578), for B 5'SP6P66/out, 5'-
GACCTACACCTGTCAACATAAT-3' (HxB2 positions 2484-2505) and 3'endprotT7, 5'-TGGAAGGATCACCAGCAATATT-3' (HxB2 positions 3005-3027) for C 5' SP6P66, 5'-AGATATCAGTACAATGTGTT-3' (HxB2 positions 2975-2994) and 3' halfPol, 5'-AAGCAGAGCTAGAACTGGCAGA-3' (HxB2 positions 3441-3462) using the conditions previously described 24.

DNA sequencing and analysis: The amplified DNA was directly sequenced on an ABI 373A automated sequencer (Applied Biosystems, Foster City, CA) using the Thermo Sequenase fluorescence-labelled primer cycle-sequencing kit (Amersham International, Little Chalfont, England) according to the manufacturers instructions. All nested primers were extended with the SP6 and T7 sequences and the PCR products were sequenced with the dye-labelled Primers SP6 (5'GATTTAGGTGACATATAG 3') and T7 (5' TAATACGACTCACTATAGGG 3'). The alignment of the sequences was performed manually based on the alignment of the Los Alamos database reference sequences for subtyping. Phylogenetic analysis of the aligned sequences was performed using the neighbor-joining method of MEGA 44 and the DNADIST, NEIGHBOR and DRAWTREE options of the PHYLIP software package 45. The distance matrix was generated by Kimura’s two-parameter estimation 44. A bootstrap value equal to or greater than 70% was considered significant based on one hundred replications 46,47. Other sequences obtained from the Los Alamos database were included as reference-sequences. Phylogenetic trees were constructed using the neighbor-joining and maximum likelihood options of the PHYLIP package programs 45. The bootscanning method was used to study the recombinant viruses as implemented in the SIMPLOT program. Our analysis was performed by calculating the distances for a sliding window of 200 nucleotides of the test sequences moving along the alignment of a panel of reference sequences by increments of 20 bp 48. One hundred replications were generated by the bootstrap method for each window and the percent bootstrap values were plotted with the nucleotide position of the sequence of the reference panel.

RESULTS

Phylogenetic analysis of the Gag and C2V3 envelope sequences. Based on the phylogenetic analysis of the C2V3 envelope region we have previously identified in Ethiopia the presence of two different subtype C groups, which we termed C and C'. No associations 108
were found between the geographic location or the different risk groups into which the individuals were divided\textsuperscript{18}.

\textbf{Figure 1:} Gag (A) and C2V3 envelope (B) phylogenetic tree analysis of Ethiopian HIV-1 sequences using the neighbor-joining method of the DNADIST, NEIGHBOR and DRAWTREE options of the PHYLIP software package. The sub-cluster is indicated as sub-cluster C'. The sequences from the different towns have been indicated by codes: AA = Addis Ababa, AM = Arba Minch, AS = Assab, DD = Dire Dawa, DE = Dessie, JM = Jima and GO = Gondar. The first two digit number following these codes indicate the year of sample collection, the next three digits indicate sample number. The subtype A (UG455, KEQ2317), subtype D (94UG114, 84ZR085) and subtype B (HxB2R, USJRFL) references are used. Numbers by the branches represent bootstrap values out of one 100 replications.
To further study this observation we randomly selected thirty HIV-1 infected individuals from seven Ethiopian towns and used their stored sera samples to analyse the 5' fragment of the Gag gene (p17 plus 134 amino acids of the p24 protein) together with the respective C2V3 envelope region of their viruses. For both regions studied the neighboring-joining/kimura-2-parameter phylogenetic analysis demonstrated that there is indeed two genetically distinct subtype C groups co-circulating in Ethiopia, illustrated in figure 1. The presence of the genetic subcluster C' is supported by significant bootstrap values of 93 for the Gag region and 97 for the C2V3 envelope region. Analysis of the C2V3 region shows fifteen sequences belonging to the C group and fifteen to the C' group whilst the analysis of the gag region reveals nineteen viruses belonging to the C group with only ten in the C' group, with one gag sequence, GO88052, interspersed between the two.

**Intra subtype C/C' recombination:** Since four isolates (AM96146, GO96009, JM96111, DE96043) belong to the C' group with respect to their C2V3 envelope sequence (figure 1B) but switch to the C group for their gag sequence (figure 1A) we believed that recombination between the two groups of viruses may be occurring. The isolate GO88052 (figure 1A) was believed to be a recombinant virus with a crossover event somewhere within the region of the genome analysed. The majority of the isolates we expected to be recombinants were collected in 1996 opposed to 1988 which is in accordance with the fact that recombination is more likely to have occurred as the HIV-1 prevalence increases. Based on the phylogenetic analysis of the 5'gag and C2V3 sequences we found evidence of C/C' recombination for five out of thirty isolates (16.6%) indicating a high recombination incidence.

In order to substantiate the presence of viral recombination and to further strengthen the genotypic distinction between the C and C' groups we selected sixteen isolates for further analysis. We included the predicted recombinant isolates AM96146, GO88052, JM96111, GO96009 and DE96043, the C group isolates JM88420, DE96054, DD88379, DD88477, DE88404 and AS88651 and the C' isolates for both their 5'gag and C2V3 envelope regions JM96125, JM96102, AM96148, AA97202 and AA88055. A 2600 bp fragment of the HIV-1 genome comprising the gag gene, the protease and the 5' half of the RT genes together with the respective V1V2 and C2V3 regions of the envelope was sequenced and analysed. In figure 2 we illustrate how the phylogenetic analysis indicates recombination for four of the studied viruses.

The analysis of the Gag region (p17 and 134 amino acids of p24) show that the isolates DE96043, JM96111 and AM96146 are grouped with the main group C viruses while
GO88052 is indeed interspersed. For the Pol region (including the protease and 293 amino acids of RT), the isolates GO88052 and DE96043 switched sequence to the subcluster C’ whilst for the C2V3 envelope region of all four viruses are grouped with the C’ subcluster.

Additionally if the phylogenetic analysis is extended to larger fragment of the genome, the genetic distinction between C and C’ becomes more apparent. The maximum likelihood tree in figure 3A illustrates the analysis of the 2600 nt fragment and thereby strengthens the evidence for the presence of the two different C and C’ strains in Ethiopia. Concurrently the neighboring joining /Kimura-2-parameter analysis demonstrates that both virus-groups C and C’ are separated into two distinct clusters by statistically significant bootstrap values of one hundred. The same result was obtained for the V1V2/C2V3 region (fig 3B). The strains found to be only C or C’ for the 2600 gag/pol fragment but also for their respective V1V2 and C2V3 envelope regions were considered the parental sequences for the boot scan analysis of isolates believed to be recombinants.

Figure 2: Gag (A), Pol (B) and Envelope (C) phylogenetic tree analysis of Ethiopian HIV-1 sequences using the neighbor-joining method of the MEGA software package. The sub-cluster is indicated as sub-cluster C’. The C/C’ recombinant HIV-1 isolates are underlined. DE96043, JM96111 and AM96146 are grouped with C in Gag while GO88052 is outlier. JM96111 and AM96146 are grouped with C in Pol while DE96043 and GO88052 are grouped with C’. In the envelope all four are grouped with C’. Sequences codes have been indicated in figure 1. Numbers by the branches represent bootstrap values out of 100 replications.
Figure 3: Maximum likelihood tree of the gag/pol 2600 nucleotide fragment (A) and the V1V2/C2V3 envelope region (B) showing the two subtype C subgroups (C and C') of HIV-1 viruses circulating in Ethiopia. The bootstrap values are based on one hundred replications estimated by the Kimura-2-parameter method using the neighbor-joining method of MEGA.

Recombination pattern analysis: While Phylogenetic analysis can lead to strong evidence for recombination, alone it does not provide information about the recombination profile that is achieved by bootscan analysis. In the figure 4A we illustrate the bootscan analysis of the 2600 nucleotide fragment for the isolate JM96111 clearly demonstrating that recombination has occurred. The nucleotide sequences coding for the p17 and the RT proteins belong to C group whilst the sequence coding for the p24 and p7 proteins belongs to the C' group with two crossover points, one between the p17 and p24 proteins and one at the 5'end of the Pol gene. The isolates D88477 (of the C group) and JM96102 (of the C' group) together with U455 (Subtype A), RF (Subtype B) and SE6165 (Subtype G) were used as reference viruses and the results of the analysis were identical with any other C and C' reference sequences (data not shown). For a 250 nucleotide segment at the 3' end of the 2600 nt sequence the bootscan analysis was inconclusive and we could not determine whether this isolate belonged to the C or the C' group previously with the (figure 4A).
For some isolates this region, together with one between the p17 and the p24 proteins and one at the common sequence of the gag and pol genes it was difficult in determining their

<table>
<thead>
<tr>
<th>Nucleotide Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
</tr>
<tr>
<td>24</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>Pol</td>
</tr>
<tr>
<td>Gag</td>
</tr>
<tr>
<td>Prot</td>
</tr>
<tr>
<td>1/2RT</td>
</tr>
</tbody>
</table>

**Figure 4:** Analysis of C/C' mosaic isolates by bootscanning, based on the neighbor-joining tree and Kimura-2-parameter methods with bootstrapping. The bootstrap values that support the clustering of the sample sequences with the references are plotted. The chosen window size is 200 nt, moving in steps of 20 nt along the alignment. (A) Plot analysis of the isolate JM96111 illustrating two crossover points, one between p17 and p24 and the other at the 5' end of the Pol gene. The region noted by a solid line belongs to C and the region noted by a broken line belongs to C'. (B) Plot analysis of the isolate AM96148 shown as C'. (C) Plot analysis of the isolate DD88379 are shown as C. The isolates DD88477 of the C group and AA96102 of the C group together with U455 (subtype A), RF (subtype B) and SE6165 (subtype G) were used as references.
affiliation to the C or C' groups (figure 5). This same difficulty has been encountered analysis of recombination between different subtype viruses. Figures 4B and 4C illustrate the analysis of two of the parental strains, DD88379 of group C and AM96148 of group C'. The strong affiliation to their respective groups along the whole length of the DNA sequence analysed is demonstrated, supporting the results from the phylogenetic study. Combining the phylogenetic and bootscan analysis further demonstrated that six out of the thirty sera analysed contained C/C' recombinant viruses (AM96146, GO88052, JM96111, GO96009, DE96043 and AA88055). The Isolate AA88055 was found to be recombinant only after a longer DNA fragment was sequenced and the bootscan analysis was performed. Even though the full length of the viral genome was not analysed our results indicate that the incidence of recombination between the two co-circulating groups is equal or greater than 20%.

In figure 5 we summarise the recombination patterns of the six C/C' recombinant strains and the recombination crossover points that we have identified. Although several crossover points can be found and some isolates are mosaics, we found the crossover sites more frequently at the region 3'p17/5'p24 of the Gag gene and at the 5'end of the Pol gene. Among the numerous recombination studies between different subtypes these crossover points have been frequently documented. Interestingly, all the recombinant viruses carry the V1V2 and V3 sequences from the C' group suggesting that the envelope of the C' subcluster may have some phenotypic advantage over the envelope of the main C group. Based on the analysis of the C2V3 envelope region, both here and in a previous study, the prevalence of the two subtype C groups of viruses co-circulating in Ethiopia appears to be similar. Nevertheless when larger numbers of sequences are considered to include all the Ethiopian strains isolated over the years there is evidence that the prevalence of the C' strains is increasing comparatively to that of the C strains. During the period 1984-92 8,18,19,49,50 55% of the viruses analysed (n=109) belong to the sub-cluster C' while for the period 1995-97 8,18,19,51 65% (n=180) were found to be C'.
Figure 5: (A) Map of the sequence of the HIV-1 genome. (B) Summary of the bootscan and phylogenetic analysis of the C/C' recombinant isolates. Regions clustering with the C group are shown as ■ and regions clustering with the C' group are shown as ▲. Regions that could not be determined are in white boxes. The numbers at the axis indicate the nucleotide positions in HxB2.
The genetic variation observed with retroviruses and their capacity for rapid adaptation poses a serious challenge for the successful development of anti-HIV-1 chemotherapeutic and vaccine reagents. For this reason the identification and surveillance of the viruses that are either transmitted or circulating in the population is necessary and especially in areas with a high incidence of HIV-1 infection. Subtype C HIV-1 viruses require special attention because it remains unclear as to why this subtype has gained such a dominant foothold in certain regions of the world. Subtype C viruses account for one third of total infections and 50% of all new infections in Africa and therefore a better understanding of its biology and the means to slow down its spread are urgently required. The HIV-1 epidemic in Ethiopia has been dominated by subtype C whilst in other east-African countries the epidemic has been dominated by the subtypes A and D which have been found to be co-circulating at variant proportions \(10:13:52:53\). However, the monophyletic nature of the Ethiopian epidemic is only apparent since we have shown that although subtype C alone is responsible for the HIV-1 epidemic there are two genetically distinct subtype C viruses co-circulating in the country without either geographic or risk-group inference \(^{18}\). The distinction between C and C' was shown by the neighbor joining / Kimura-2-parameter phylogenetic analysis of the Gag region of the genome as well as the C2V3 envelope regions. Furthermore, the maximum likelihood tree based on the 2600 nucleotide sequence including the Gag and half of the Pol genes confirmed that indeed in Ethiopia there are two distinct strains which belong to the C subtype. In support of our data it has been reported based upon the sequence analysis of the envelope V3V4 sequence that in Nepal two Subtype C viruses are co-circulating \(^{54}\). We have shown in this study that even though the two co-circulating subtype C strains in Ethiopia can evolve independently creating two sub-epidemics they can also result in the appearance of new recombinant viruses. Recombination within the same subtype is often difficult to detect because the genetic distance among the circulating strains in one population is often not significant. However, our data demonstrates that recombination can occur not only among the different subtypes of HIV-1 but also within one subtype. In this study we identified six Ethiopian individuals carrying C/C' recombinant viruses amongst the 30 individuals that we studied. The bootscan analysis demonstrated that all the six recombinant strains had more than one recombination crossover point per genome. One in Gag (between the p17 and the p24 peptides) and one at the 5'end of the Pol gene and, although we have not proven it, the data indicates that there may be one at the 5' end of the envelope gene as well.

116
This intra-subtype recombination pattern is in agreement with previously reported inter-subtype recombination indicating that recombination follows the same rules whether it happens between different subtypes or within the same subtype.

Despite the fact that the number of virus sequences analysed was low and full-length genomes were not studied in our analysis we have shown that the recombination frequency is approximately equal to 20%. Frequencies of approximately 20-25% have been calculated for recombination between different subtypes, much in accordance with what we see for the intra-subtype recombination. If we speculate that recombination can happen amongst all the circulating viruses the occurrence of recombination in areas of high prevalence must be a common event with the only prerequisites the diploid character of the HIV-1 genome and the geographic co-localisation of the parental strains. Indeed, in vitro studies showing that recombination events can occur two to three times during each replication cycle would support this conclusion. The high incidence of intra-subtype recombination may well challenge the classic understanding that the high divergence rate observed in HIV-1 infection primarily results from a high replication rate in combination with the error prone nature of the reverse transcriptase protein with significant implications for the comprehension of the evolution of the virus. Intra-subtype recombination provides a mechanism whereby two replicating viruses within the same host and each carrying a different selective advantage may be able to recombine to generate progeny virus that carry both the advantageous phenotypes.

A striking observation in this study is that all the recombinant viruses identified carried the envelope of the C' virus suggesting that this envelope may be more efficient for virus transmission than the envelope of subgroup C. This is supported by the finding that the prevalence of the C' group envelope in comparison to the C group is increasing over time and that neither the C nor the C' group is restricted by geography or confined to one risk group. Concomitantly inter-subtype recombination can also be unidirectional. In Tanzania where a high number of A/C and D/C inter-subtype recombinant viruses were identified the ones with the subtype C envelopes were only found, suggesting an increased fitness of the subtype C enveloped viruses circulating in that area. The (Circulating Recombinant Form) CRF-AE virus carrying the subtype E envelope was found to have created the epidemic in Southeast Asia whilst the parental subtype E or AE recombinants with the subtype A envelope were never found. While this must be further investigated, evolution under the light of recombination can lead to the hypothesis that a recombinant progeny if robust could spread faster and more efficiently than the parental viruses, thereby heightening the epidemic.
The subtype C HIV-1 pandemic is spreading alarmingly fast with no explanation to-date as to why. Several studies have linked the three Nf-Kb sites of the subtype C long terminal repeat sequence with rapid replication rates and it has been reported that the viral loads of individuals infected with subtype C are in general higher than individuals infected with the other subtypes. A high incidence of intra-subtype C recombination may be another factor to be taken into consideration when monitoring the spread of this subtype and when considering HIV-1 vaccine development.
References


44. Kumar, S., K. Tamura, and M. Nei. 1993. Molecular Evolutionary Genetics Analysis (MEGA), version 1.01, Institute of Molecular Evolutionary Genetics, Pennsylvania State University, USA.


the branched DNA and nucleic acid sequence-based amplification assays of HIV-1

human immunodeficiency virus type 1 epidemic among injecting drug users of
Kathmandu, Nepal, is caused by a subtype C virus of restricted genetic diversity. J.
Virol. 74:1149-1157.


High rate of recombination throughout the human immunodeficiency virus type 1
genome. J. Virol. 74:1234-1240.


Hunter, W. Fawzi, and M. Essex. 1999. Emerging recombinant human
immunodeficiency viruses: uneven representation of the envelope V3 region. AIDS
13:1613-1621.

Retroviruses 15:1293-1303.

Berkhout. 2000. Functional differences between the long terminal repeat transcriptional
promoters of human immunodeficiency virus type 1 subtypes A through G. J. Virol.
74:3740-3751.

Acknowledgements
This study was supported by the Ethio-Netherlands AIDS Research Programme (ENARP), a collaborative project of the Ethiopian Health and Nutrition Research Institute (EHNRI), the Amsterdam Municipal Health Service (GG/GD), the Central Laboratory of the Netherlands Red Cross Blood Transfusion Service (CLB) and the Academic Medical Center of the University of Amsterdam (AMC). ENARP is financed by the Netherlands Ministry of Foreign Affairs and the Ethiopian Ministry of Health (MOH) as a bilateral project. The authors would like to thank Dr William Paxton for critical review of the manuscript.