

High prevalence of diverse forms of HIV-1 intersubtype recombinants in Central Myanmar: geographical hot spot of extensive recombination

Yutaka Takebe^a, Kazushi Motomura^a, Masashi Tatsumi^b,
Hla Htut Lwin^c, Myint Zaw^c and Shigeru Kusagawa^a

Objectives: To investigate the molecular epidemiology and genetic structure of HIV-1s causing the epidemic in Central Myanmar and to explore the genesis of HIV epidemic in this area.

Design: A molecular epidemiological investigation was conducted in 1999–2000 in the city of Mandalay among high-risk populations and the structural features of circulating HIV-1s were analyzed.

Methods: HIV-1 genotypes of 59 specimens were screened based on *gag* (p17) and *env* (C2/V3) regions. Near full-length nucleotide sequences of HIV-1 isolates with subtype discordance were determined and their recombinant structures were characterized.

Results: Three lineages of HIV-1 strains, including CRF01_AE (27, 45.8%), subtype B' (Thailand variant of subtype B) (15, 25.4%) and subtype C (8, 13.6%), were distributed in Mandalay, while substantial portions (9, 15.3%) of specimens showed various patterns of subtype discordance in different regions of HIV-1 genomes. The study on six HIV-1 isolates with subtype discordance revealed that they were highly diverse types of unique recombinant forms (URFs) comprised of various combinations of three circulating subtypes. One URF was a particularly complex mosaic that contained 13 recombination breakpoints between three HIV-1 subtypes. Approximately half of recombinants showed 'pseudotype' virion structures, in which the external portions of envelope glycoproteins were exchanged with different lineages of HIV-1 strains, suggesting the potential selective advantage of 'pseudotype' viruses over parental strains.

Conclusion: The study revealed the unique geographical hot spot in Central Myanmar where extensive recombination events appeared to be taking place continually. This reflects the presence of highly exposed individuals and social networks of HIV-1 transmission.

© 2003 Lippincott Williams & Wilkins

AIDS 2003, **17**:2077–2087

Keywords: HIV-1, recombination, unique recombinant forms (URFs), molecular epidemiology, injecting drug users, Myanmar

From the ^aLaboratory of Molecular Virology and Epidemiology, AIDS Research Center and ^bDepartment of Veterinary Science, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan and ^cAIDS Prevention and Control Programme, Department of Health, 36, Theinbyu St., Yangon, Myanmar.

Correspondence to Yutaka Takebe, Laboratory of Molecular Virology and Epidemiology, AIDS Research Center, National Institute of Infectious Diseases, Toyama 1-23-1, Shinjuku, Tokyo 162-8640, Japan.

Tel: +81 3 5285 1111; fax: +81 3 5285 1258; e-mail: takebe@nih.go.jp

Received: 26 September 2002; revised: 6 March 2003; accepted: 19 March 2003.

DOI: 10.1097/01.aids.0000076350.42412.a5

Introduction

The extensive genetic diversity is a hallmark of human immunodeficiency viruses (HIVs). The high error rate of reverse transcriptase (3×10^{-5} sites/genome/replication cycle) [1,2], coupled with an *in vivo* virus production rate exceeding 10^9 per day and persistent nature of infections [3–5], provides tremendous scope of generation of viral diversity. Second mechanism for acquiring sequence diversity is genetic recombination. This can occur when a cell that is dually infected with different viruses produces progeny virions with genomic RNAs from each virus, and strand-switching takes place during the next round of reverse transcription [6–8]. Unlike the incremental accumulation of sequence changes by copying errors, recombination has the potential to introduce large numbers of genetic changes simultaneously, accelerating the genetic diversity and the adaptive evolution of HIVs [9]. As increasing numbers of full-length viral sequences become available, the number of recombinant or mosaic viruses is being recognized more frequently. Some recombinant strains disseminate widely in the human populations, becoming circulating recombinant forms (CRFs) [10] (<http://hiv-web.lanl.gov/CRFs/CRFs.html>).

In addition to CRFs that play a major role in global or regional HIV spread, a variety of HIV-1 unique recombinant forms (URFs), that are so far identified only in a single individual without evidence of epidemic spread, have been reported in various regions in the world, including Eastern Africa (A/D and A/C recombinants) [11–15], India (A/C) [16], Thailand (B/CRF01_AE) [17], and Yunnan Province of China (B/C) [18]. URFs are often recognized among high-risk individuals who acquired HIV relatively recently in the regions where multiple HIV-1 subtypes are co-circulating.

In Myanmar, HIV-1 infection was first reported in 1989 among injecting drug users (IDUs) and subsequently spread rapidly into various risk populations [19–21]. Approximately 30 % of HIV cases are attributed to injecting drug use and 68% to heterosexual transmission. The 1999 UNAIDS estimates indicated approximately 530 000 HIV cases in Myanmar, the second largest number in South-east Asia, only after Thailand. Of Asian countries, Myanmar has one of the highest prevalence of HIV-1 [22], especially in the city of Mandalay in Central Myanmar [20,23]. The epidemic has included the spread of three HIV-1 strains, including subtypes B' (Thailand variant of subtype B) and C and CRF01_AE, that are likely to have originated in the surrounding regions [23]. To gain a comprehensive picture of the HIV-1 diversity in Myanmar, we have characterized the genetic structure of HIV-1 strains from Central Myanmar and found that substantial portions of HIV-1 isolates were URFs that

were distinct from any known recombinants. The present study describes the identification of a unique geographical hot spot in Central Myanmar where the extensive recombination events are taking place continually, leading to the generation of diverse forms of HIV-1 intersubtype chimeras with unique structural features.

Materials and methods

Study subjects and specimens

EDTA-treated blood samples were collected from 59 asymptomatic HIV-positive consenting patients from various risk populations, including 21 male IDUs, 16 female commercial sex workers (fCSW), 12 sexually transmitted disease (STD) patients (11 male and one female), and 10 heterosexuals (seven male and three female), in the city of Mandalay and the vicinity in Central Myanmar during the period between December 1999 and December 2000. The participants included 39 males with an age range of 21–46 years (mean: 31.2 ± 6.2 years old) and 20 females with an age range of 16–32 years (mean: 21.0 ± 4.0 years old). All 59 specimens were serologically determined as HIV-1 infections. No HIV-2 infections were detected. The peripheral blood mononuclear cells (PBMCs) were separated on Ficoll-Hypaque (Pharmacia, Piscataway, New Jersey, USA) density gradient centrifugation. For virus isolation, PBMCs from HIV-1-positive individuals were co-cultured with phytohemagglutinin (PHA, 1 $\mu\text{g}/\text{ml}$)-stimulated CD8+ T-cell-depleted PBMCs from HIV-negative healthy donors in RPMI 1640 containing 10% fetal calf serum and interleukin-2 (20 U/ml). Virus production was detected by virion-associated reverse transcriptase (RT) assay as described previously [24]. Plasma were saved for genotype screening based on the nucleotide sequence determination of virion HIV-1 RNAs [18].

Screening of HIV-1 genotypes

The nucleotide sequences of 432-bp *gag* (*p17*) and 336-bp *env* (*C2/V3*) regions were determined for the primary screening of HIV-1 genotypes, as described previously [18,23]. All the nucleotide sequences obtained in the present study were screened by the BLAST 2.0 program (National Center For Biotechnology Information, USA.) to search for sequence similarities to previously reported sequences in the databases, and to rule out potential laboratory errors.

Isolation of near full-length HIV-1 molecular clones

DNAs were extracted from CD8-depleted PHA-stimulated PBMCs infected with respective HIV-1 isolates. The near full-length (approximately 9.1 kb) HIV-1 genomes were amplified by polymerase chain reaction

(PCR) and cloned as described previously [25,26]. Positive clones with near full-length HIV-1 inserts were selected and the nucleotide sequences of HIV-1 genomes were determined on both strands by direct sequencing method with fluorescent dye terminators in an automated ABI PRISM310 DNA sequencer (Applied Biosystems, Inc., Foster City, California, USA), using the primer-walking approach.

Data analyses

The near full-length nucleotide sequences were aligned with the HIV-1 reference strains (<http://hiv-web.lanl.gov/HTML/alignments.html>), using CLUSTAL W version 1.4 [27], and corrected manually to ensure that gaps did not alter the reading frame. Phylogenetic trees were constructed by the neighbor-joining method [28] based on Kimura's two-parameter distance matrix with 100 bootstrap replicates [29] using PHYLIP, version 3.573 [30]. Bootscanning analysis [31] were performed on neighbor-joining trees for a window of 500 nucleotides moving along the alignment in increments of 100 nucleotides, using two different sets of reference sequences: A_92UG037, B'_RL42, C_95IN21068, D_NDK and CRF01_93TH253; A_Q23, B_HXB2, C_92BR025, D_94UG1141, and CRF01_90CF402. The analyses were implemented by the SIMPLOT 2.5 program [32]. To further define the precise boundaries of intersubtype recombination, informative site analysis was performed as described previously [33,34]. Confirmatory tree analyses were then carried out to explore the subtype origins of respective HIV-1 segments. Briefly, using the co-ordinates for the recombination breakpoints estimated by informative site analysis and bootscanning plots, HIV-1 genomes of the putative recombinants were divided into segments. Each segment was subjected to separate phylogenetic analyses using the neighbor-joining method to confirm the subtype or CRF origin of each segment. The 'predicted' parental sequences (B'_RL42, C_95IN21068

and CRF01_93TH253) for bootscanning and informative site analyses were selected based on the data obtained from confirmatory tree analysis.

Nucleotide sequence accession numbers

The near full-length nucleotide sequences reported in this article are available under the database accession numbers AB097865 to AB097873.

Results

High prevalence of HIV-1 strains with subtype discordance in Central Myanmar

HIV-1 genotypes of 59 specimens from Central Myanmar were determined by phylogenetic tree analyses based on the nucleotide sequences of *gag* (p17) and *env* (C2/V3) regions, using plasma HIV-1 RNAs. The distribution of HIV-1 genotypes in this population was as follows: 27 CRF01_AE (45.8%); 15 subtype B' (Thailand variant of subtype B) (25.4%); eight subtype C (13.6%). The remaining nine specimens (15.3%) showed the discordance between *gag* (p17) and *env* (C2/V3) subtypes (Table 1). The samples with subtype discordance were found most frequently among IDUs (six of 21, 28.6%) and fCSWs (two of 16, 12.5%), whereas none was detected among STD patients (none of 12) (Table 1).

To investigate the detailed structural features of HIV-1 strains in Mandalay, a total of 27 HIV-1 strains were isolated from these 59 HIV-1 samples. The relatively low isolation rate of HIV-1 (27 of 59, 45.8%) was due to the long-distance transportation of specimens. The genotype screening of these 27 isolates based on *gag* (p17) and *env* (C2/V3) regions identified 11 CRF01_AE (40.7%), seven subtype B' (26.0%) and four subtype C (14.8%), whereas the remaining five

Table 1. Summary of primary screening of HIV-1 genotypes^a in Central Myanmar.

Risk factor	Number subtyped	RF ^b								
		B'	C	CRF01_AE	RF ^b	B'/C	B'/01	C/B'	C/01	01/B'
IDU	21	7 (33.3%)	4 (19.0%)	4 (19.0%)	6 (28.6%)	1	1	1	2	1
Person with sexual exposure	38	8 (21.1%)	4 (10.5%)	23 (60.5%)	3 (7.9%)	0	1	1	0	1
CSW	16	2	0	12	2	0	0	1	0	1
STD	12	3	2	7	0	0	0	0	0	0
Hetero	10	3	2	4	1	0	1	0	0	0
Total	59	15 (25.4%)	8 (13.6%)	27 (45.8%)	9 (15.2%)	1	2	2	2	2

^aHIV-1 genotypes: HIV-1 subtype B', B'; subtype C, C; CRF01_AE, 01. ^b RF indicates the HIV-1 samples that show the discordance between *gag* (p17) and *env* (C2/V3) subtypes. The columns under RF show the breakdowns of the specimens with different profiles of subtype discordance. For example, B'/C indicates the HIV-1 sample that belongs to subtype B' in *gag* (p17) region and subtype C in *env* (C2/V3) region. IDU, injecting drug user; CSW, commercial sex worker; STD, sexually transmitted diseases; Hetero, heterosexual.

strains (18.5%) showed the diverse patterns of discordance between *gag* (p17) and *env* (C2/V3) subtypes, including *gag* subtype B' and *env* subtype C (99MM-mIDU106), B' and CRF01_AE (00MM-mIDU502), C and B' (00MM-mCSW503), C and CRF01_AE (99MM-mIDU107), and CRF01_AE and B' (99MM-mCSW104). Moreover, the additional genotyping based on nucleotide sequences of 3'-long terminal repeats (LTRs) found that one strain (99MM-mIDU103) assigned as subtype B' in the primary screening harbored the LTR of subtype C origin (see below). Taken together, our genotype screenings identified a total of six strains with subtype discordance (22.2%) among 27 isolates from Central Myanmar.

Highly diverse forms of unique intersubtype recombinants in Central Myanmar

To analyze the genome structures of these six HIV-1 isolates, we cloned and determined the near full-length nucleotide sequences of their proviral genomes. Three representative HIV-1 isolates from Central Myanmar that are presumably non-recombinant forms of subtypes B' (99MM-mSTD101) and C (99MM-mIDU101) and CRF01_AE (99MM-mCSW105) were analyzed in parallel for comparison. The 9.1 kb near full-length clones of six HIV-1 strains with subtype discordance were designated 99MM-mIDU106.18, 99MM-mCSW104.16, 99MM-mIDU107.34, 00MM-mIDU502.6, 00MM-mCSW503.2, and 99MM-mIDU103.10, and those of putative non-recombinant forms of HIV-1 subtypes B' and C and CRF01_AE strains are termed 99MM-mSTD101.8, 99MM-mIDU101.3 and 99MM-mCSW105.18, respectively. Whereas 99MM-mIDU101.3 carried a frame-shift mutation in the *gag* gene, other eight molecular clones had intact open reading frames for all nine HIV-1 genes.

The near full-length nucleotide sequences of these Mandalay isolates were subjected to recombination identification programs. The bootscanning analyses, using the reference strains of HIV-1 subtypes A (92UG037), B' (RL42), C (95IN21068), and D (NDK) and CRF01_AE (93TH253), revealed that all six HIV-1 strains with subtype discordance were highly diverse forms of unique recombinants comprised of various combinations of the segments derived from HIV-1 subtypes B' and C and CRF01_AE (Fig. 1a–f), whereas 99MM-mSTD101.8, 99MM-mIDU101.3 and 99MM-mCSW105.18 were indeed non-recombinant forms of HIV-1 subtypes B' and C and CRF01_AE, respectively (Fig. 1g–i). Similar results were obtained with an alternative set of reference sequences (A_Q23, B_HXB2, C_92BR025, D_94UG1141, 01_90CF402) (data not shown). To further define the recombination breakpoints, the informative site analysis was performed with predicted parental strains, including B'_RL42, C_95IN21068, and CRF01_93TH253. The informa-

tive site analysis gave high statistical support ($P < 0.001$) for all breakpoints (data not shown).

Next, we carried out the confirmatory tree analysis to examine the phylogenetic position of these newly identified recombinants in the respective HIV-1 segments and to estimate the origin of the viruses that have been involved in the recombinations (Fig. 2). The exact coordinates for the breakpoints relative to HXB2 were given in parentheses in Fig. 2. As shown in Fig. 2, the subtype structure deduced from confirmatory tree analysis was consistent with the results obtained from bootscanning (Fig. 1) and informative site analyses (data not shown). The deduced subtype structures are illustrated at the bottom of each panel in Figure 1. The confirmatory tree analysis also revealed that most of subtype B parts in these recombinants indeed belonged to subtype B' cluster and that the subtype C parts were closely-related phylogenetically to subtype C sequences from India (Fig. 2). However, the sequences in some fragments in 99MM-mIDU107.34 (Fig. 2e) and 00MM-mCSW503.2 (Fig. 2f) did not clearly cluster with the reference sequences for Indian subtype C (C_{IN}) and for CRF01_AE of Thai origin (01_{TH}). This is probably due to the poor resolution in the tree analysis because the fragments were too short (Fig. 2e and f). Alternatively, a few breakpoints may not be correctly identified. In fact, informative site analysis indicated that small patches of different subtypes appeared to be inserted in some segments (data not shown), although the analyses gave high statistical support for estimated subtype boundaries ($P < 0.001$). Of note, 00MM-mCSW503.2 was a particularly complex chimera that contained at least 13 recombination breakpoints between subtypes B' and C and CRF01_AE (Figs 1f and 2f). This would be one of the most complex chimeras ever reported.

Mandalay URFs with “pseudotype” virion structures

Intriguingly, some of recombinant strains identified in the present study showed ‘pseudotype’ virion structures, in which the external portions of the envelope glycoproteins were exchanged with that of different lineage of HIV-1 strains. The most distinct examples were a pair of 00MM-mIDU502.6 (Fig. 1b) and 99MM-mCSW104.16 (Fig. 1c), that showed reciprocal patterns of recombination between subtype B' and CRF01_AE. The 00MM-mIDU502.6 was a subtype B' strain recombined with CRF01_AE in the external portion of the envelope glycoprotein (Fig. 1b). In contrast, the 99MM-mCSW104.16 was a CRF01_AE strain recombined with the external portion of the envelope glycoprotein of subtype B' (Fig. 1c). Similarly, the 99MM-mIDU106.18, that was comprised mostly of subtype B', was recombined with the external portion of the envelope glycoprotein of subtype C origin (Fig. 1d). Although the subtype

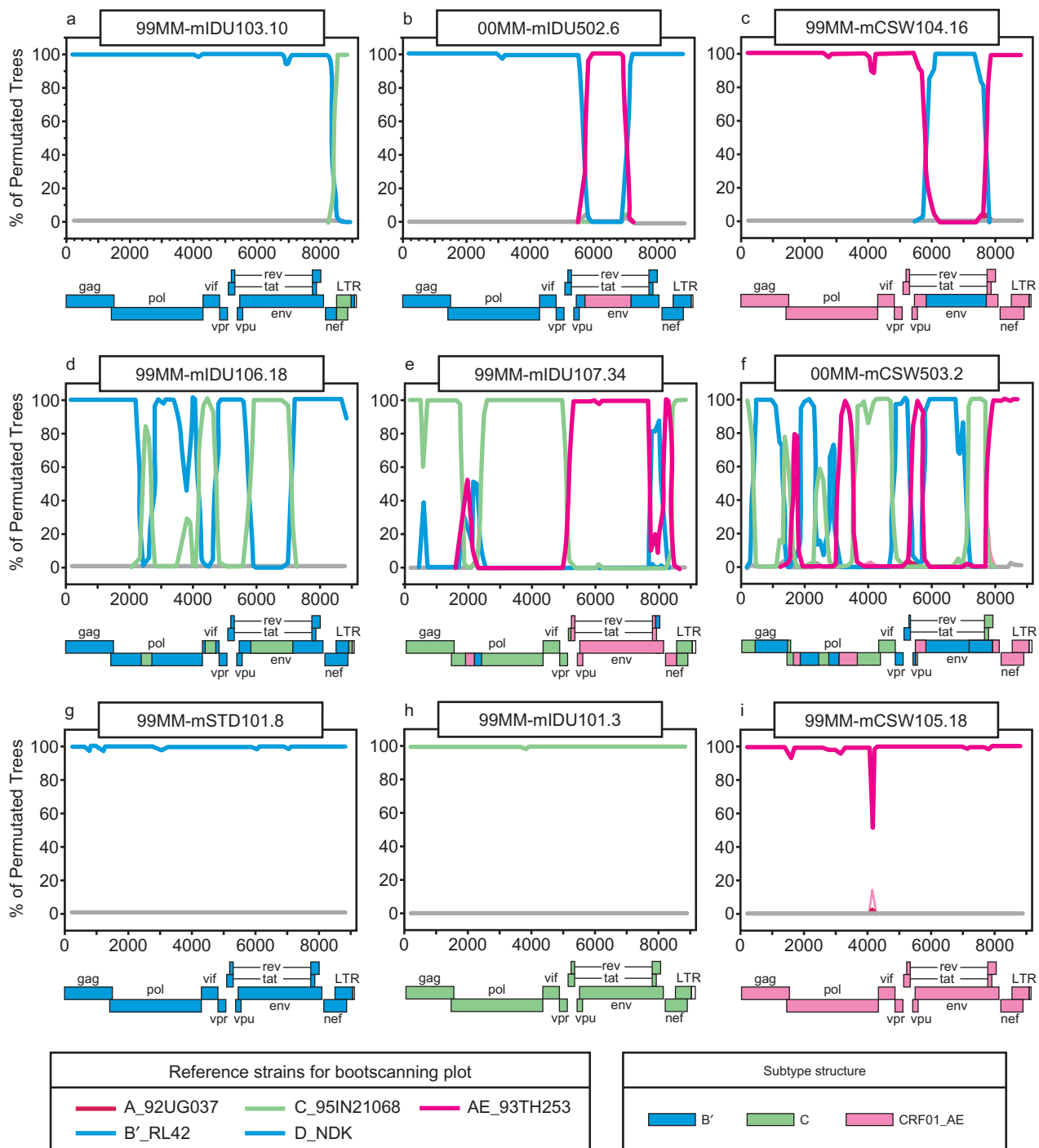


Fig. 1. Bootscanning plots and deduced subtype structures of near full-length genome sequence of HIV-1 isolates from Mandalay (Central Myanmar). The bootscanning plots, depicting the relationship of Mandalay isolates to the reference strains of HIV-1 subtypes A (92UG037), B' (RL42), C (95IN21068), and D (NDK) and CRF01_AE (93TH253) (indicated in the inset at the bottom left) are shown. The bootstrap values are plotted for a window of 500 bp moving in increments of 100 bp along the alignment. The deduced subtype structure is illustrated at the bottom of each panel. The regions of the respective subtypes are shown in different colors indicated at the bottom right. The long terminal repeat (LTR) region of 99MM-mIDU106.18 was found to contain small subtype C segment in the enhancer–promoter region (see text). The exact coordinates for the recombination breakpoints are given in the parenthesis in Fig. 2.

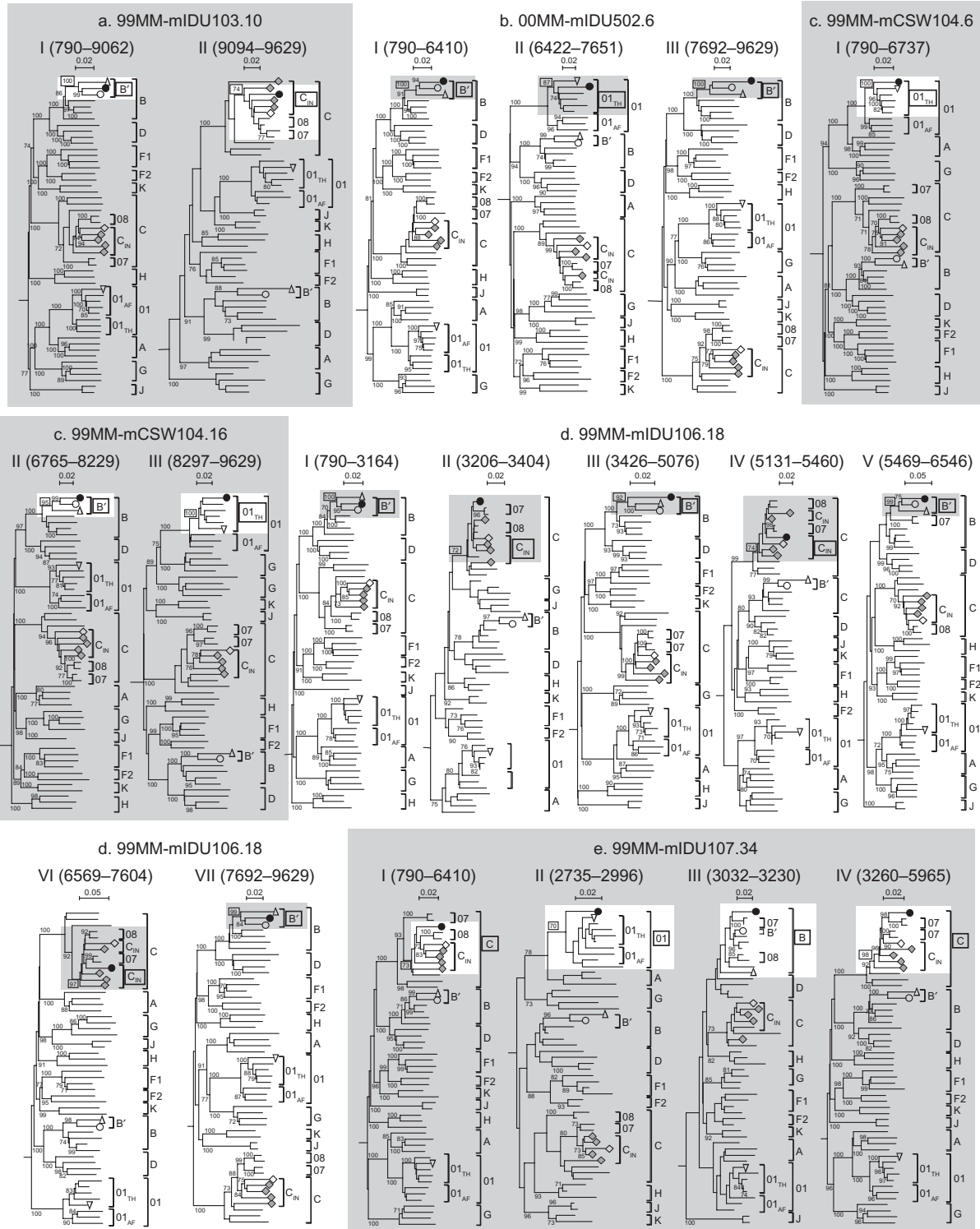
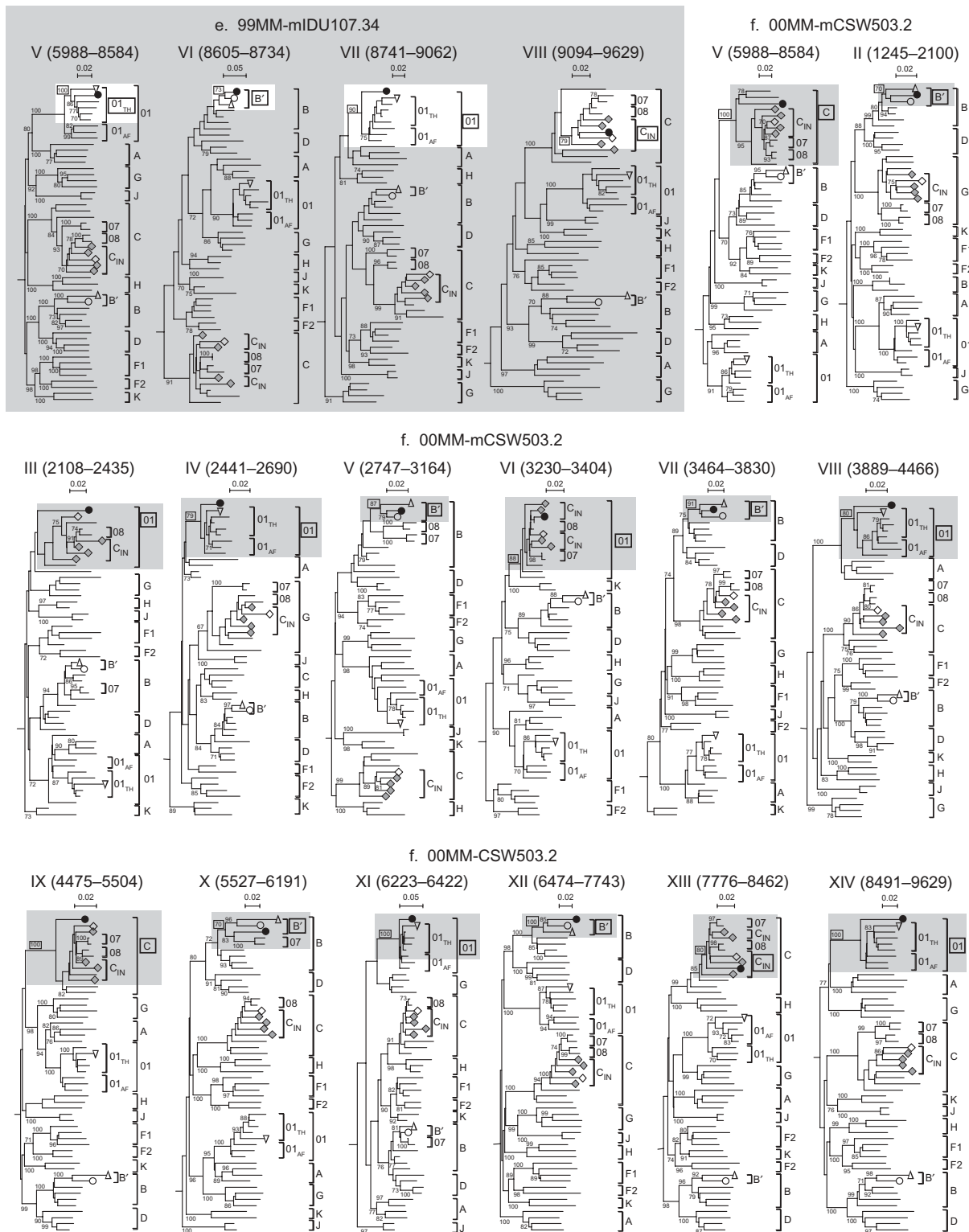


Fig. 2. Confirmatory tree analysis. Using the coordinates for the recombination breakpoints identified by the informative site analysis, HIV-1 genomes were divided into segments numbered by Roman numerals. Each segment was subjected to separate phylogenetic analyses based on neighbor-joining method to confirm the subtype or circulating recombinant form (CRF) origin of segment. The stability of the nodes was assessed by using maximum parsimony [30] with the bootstrap value of 100 replications [29]. Bootstrap values 70% or higher of key nodes are shown. *SIV_{CPZ}GAB* was used as an outgroup but is not shown for simplicity. The coordinates of each segment (shown in the parentheses at the top of each tree) are given relative to the HXB2 genome (<http://hiv-web.lanl.gov/content/hiv-db/NUM-HXB2/HXB2.MAIN.html>). The analysis starts from *gag* open reading



frame. The bootstrap values, with which the cluster containing each query strain (closed circle) is supported, are marked within the squares at the corresponding nodes. The genotype assignments of each query strain in the respective phylogenetic analyses are indicated in the box at the right-hand side of each panel. Non-recombinant forms of subtypes B' (99MM-mSTD101.8, open triangle) and C (99MM-mIDU101.3, open diamond) and CRF01_AE (99MM-mCSW105.18, open inverted triangle) identified in the present study were also included. B', subtype B' (Thailand variant of subtype B; B'_CN.RL42, open circle); C_{IN}, India cluster of subtype C (93IN905, 95IN21068, 98IN012 and 98IN022; striped diamond); 01, CRF01_AE; 01_{TH}, Thailand CRF01; 01_{AF}, African CRF01; 07, CRF07_BC; 08, CRF08_BC.

structure was less simple, 99MM-mIDU107.34, that was comprised mainly of subtype C, was recombined with CRF01_AE in *env* region (Fig. 1e). Even in the most complex chimera, 00MM-mCSW503.2, we saw a similar trend. In 00MM-mCSW503.2, the external portion of the envelope glycoprotein of subtype B' origin appeared to be recombined with a chimeric genome with highly complex subtype composition (Fig. 1f and 2f). One exception was 99MM-mIDU103.10, that was a chimera comprised mostly of subtype B' with a small segment (approximately 200 bp) derived from subtype C in LTR region (Fig. 1a and 2a, see below).

Phylogenetic relationship of HIV-1 strains from Central Myanmar

The neighbor-joining tree analysis based on near full-length nucleotide sequences revealed that HIV-1 intersubtype chimeras were the 'outliers' placed outside the clusters of known HIV-1 subtypes or CRFs, whereas three putative non-recombinant forms of HIV-1 strains formed monophyletic clusters with subtypes B' and C and CRF01_AE, respectively (Fig. 3). The locations of HIV-1 recombinants in the phylogenetic tree appeared

to reflect the proportion of the length of each subtype segment in respective HIV-1 chimera. For instance, 00MM-mIDU502.6 and 99MM-mIDU107.34 and 99MM-mCSW104.16 which were made up mainly of subtype B', C and CRF01_AE (Fig. 1) were located near the clusters of subtype C and CRF01_AE, respectively (Fig. 3). In contrast, the highly complex chimera, 00MM-mCSW503.2, that was made up of multiple segments of subtypes B' and C and CRF01_AE, branched out from the central part of the tree (Fig. 3). Of note, a subtype C strain from Central Myanmar (99MM-mIDU101.3) was most closely related to those of Indian origin (95IN21068 and 93IN905) (Fig. 3) [23].

Discussion

The present study revealed that substantial portions of HIV-1 strains circulating in the city of Mandalay in Central Myanmar were URFs with diverse profiles of intersubtype recombinations (Fig. 1, Table 1). They were not related each other nor to any known

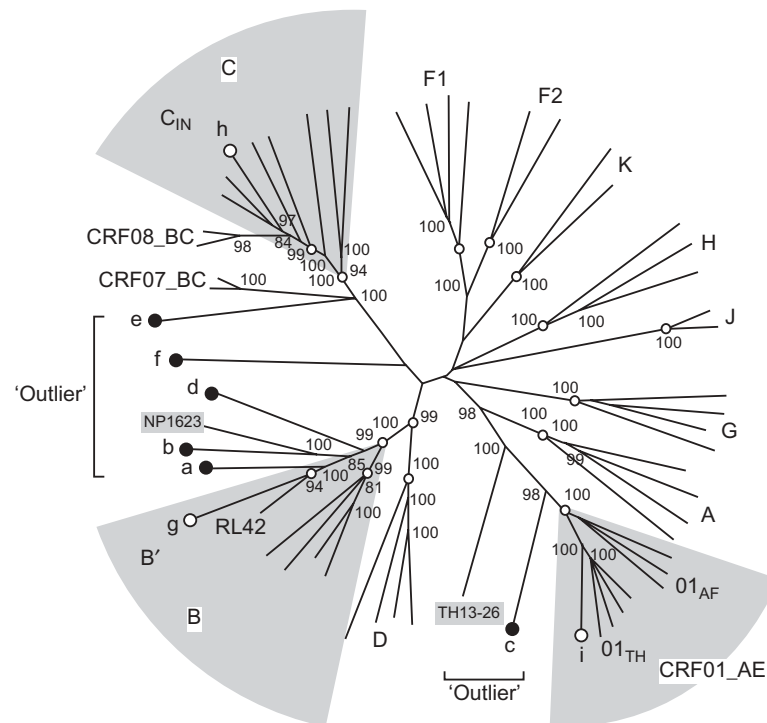


Fig. 3. Phylogenetic relationship of the diverse forms of unique HIV-1 intersubtype recombinants isolated in Central Myanmar. Neighbor-joining tree based on near full-length nucleotide sequences is shown with the reference strains for HIV-1 group M (subtypes A–D, F–H, J, and K) and circulating recombinant forms (CRFs) as well as the recently reported recombinants between subtype B' and CRF01_AE identified in Thailand (97TH.NP1623 [17] and TH13-26, shadowed). SIV_{CPZ}GAB was used as an outgroup but is not shown for simplicity. Subtype and CRF designations are indicated outside the tree. Bootstrap values greater than 90 are shown at corresponding nodes. HIV-1 unique recombinant forms (closed circle): a, 99MM-mIDU103.10; b, 00MM-mIDU502.6; c, 99MM-mCSW104.16; d, 99MM-mIDU106.18; e, 99MM-mIDU107.34; f, 00MM-mCSW503.2. Non-recombinant forms of subtype B' (g, 99MM-mSTD101.8), subtype C (h, 99MM-mIDU101.3) and CRF01_AE (i, 99MM-mCSW105.18) strains from Central Myanmar (open circle). HIV-1 genotype notations are same as in Figure 2.

recombinants, indicating that they had arisen independently. The prevalence of URFs among high risk population in Central Myanmar is approximately 10 % in CSWs and 20–30 % in IDUs (Table 1) [20,23,35]. Since our screening system for HIV-1 genotypes relied on the nucleotide sequencing of *gag* (p17) and *env* (C2/V3) regions, the actual numbers of HIV-1 intersubtype recombinants in Central Myanmar could be greater than we reported in the present study. Indeed, by additional genotype screening based on 3'-LTR, we identified one B'/C recombinant (99MM-mIDU103.10) (Fig. 1a) among four HIV-1 isolates that were originally assigned as subtype B'. Obviously, the more HIV-1 segments are analyzed, the more recombinants are likely to be recognized in this study site. As seen in 00MM-mCSW503.2 (with 13 recombination breakpoints) (Fig. 1f) and 99MM-mIDU107.34 (with seven recombination breakpoints) (Fig. 1e), some Mandalay URFs were highly complex chimeras between multiple lineages of HIV-1 strains. These findings, taken together, suggest that extensive recombination events are taking place in an ongoing fashion and new recombinants appeared to be arising continually in this area. Similar phenomenon was observed in Central Africa, where the proportion of discordant *gag/env* samples account for up to 40% [36].

This unusually high rate of detection of intersubtype recombinants is not due to the technical artifacts, including the possibility of the template switch during the long PCR procedure. We scrutinized the recombination breakpoints by determining the nucleotide sequences of original virus stocks and plasma samples in separate experiments. The results were consistent with the data obtained from near full-length DNA clones in the present study.

The exploratory tree analysis gave further information on the origin of the viruses that have been involved in the recombinations. As shown in Figure 2, exploratory analyses demonstrated that most of subtypes B and C and CRF01_AE parts of recombinant genomes belonged to the clusters of subtype B' (Thailand variant of subtype B), subtype C of India origin, and Thailand CRF01_AE, respectively, indicating that newly emerged recombinants in Central Myanmar have been generated by the mixing of subtype B' and CRF01_AE from Thailand and the subtype C strains from India (Fig. 4).

A geographical hot spot of extensive HIV-1 intersubtype recombination events has been recently identified in the western part (Dehong Prefecture) of Yunnan Province in China, near the border to Myanmar, where approximately two-thirds of circulating strains are URFs mainly comprised of subtypes B' and C [18] (Fig. 4). The high prevalence of URFs in Central Myanmar and the western part of Yunnan Province of

China are likely to reflect the presence of highly exposed individuals and social networks of HIV-1 transmissions in these areas [18]. The areas, including western Yunnan and Central Myanmar, thus appear to be the 'melting pots' where the diverse forms of HIV-1 recombinants are continually generated (Fig. 4).

As typically seen in CRF01_AE (*env* E in subtype A backbone), and CRF14_BG (*env* B in subtype G backbone), some CRFs exhibited 'pseudotype' virion structures, in which the external portions of the envelope glycoprotein were exchanged with that of different lineage of HIV-1 strain. A similar tendency was observed in other CRFs with more complex recombinant structure, including CRF02_AG (*env* A in subtypes A/G backbone), CRF06_cpx (*env* G in subtypes A/J/K backbone), CRF011_cpx (*env* A in subtypes A/E/G/J backbone), and CRF13_cpx (*env* A in subtypes A/E/G/J/U backbone). Intriguingly, approximately half of Mandalay URFs (four of six) identified in the present study displayed such 'pseudotype' virion structures (Fig. 1).

Tovanabutra *et al.* has identified one recombinant (97TH.NP1623), in which the external portion of *env* gene of CRF01_AE were exchanged with subtype B' genome, in a multiply exposed individual in Thailand [17]. Although the subtype structure of 97TH.NP1623 resembles that of 00MM-mIDU502.6, the precise locations of recombination breakpoints are not identical, indicating that both recombinants have been generated independently. This may also suggest that there are the mechanism(s) to allow the convergent evolution of recombinants with virtually identical configuration, albeit they seem to be generated independently in two epidemiologically-unrelated individuals. It is thus tempting to speculate that such 'pseudotype' virion structures might confer potential selective advantage on recombinant viruses over parental viruses (e.g. to escape from host immune surveillance), leading to the evolution of the recombinants with nearly identical subtype structures.

In summary, the present study identified the unique geographical hot spot in Central Myanmar where the extensive recombinations appear to be taking place continually. The presence of highly exposed individuals and social networks of HIV-1 transmission could quickly lead to the generation of highly diverse forms of unique intersubtype recombinants. This provides insights into the understanding the genesis of HIV-1 epidemic in this particular area in Asia.

Acknowledgements

We would like to thank Pe Thet Htoon, Min Thwe, Soe Lwin, Khin Yi Oo, the staff at the Ministry of

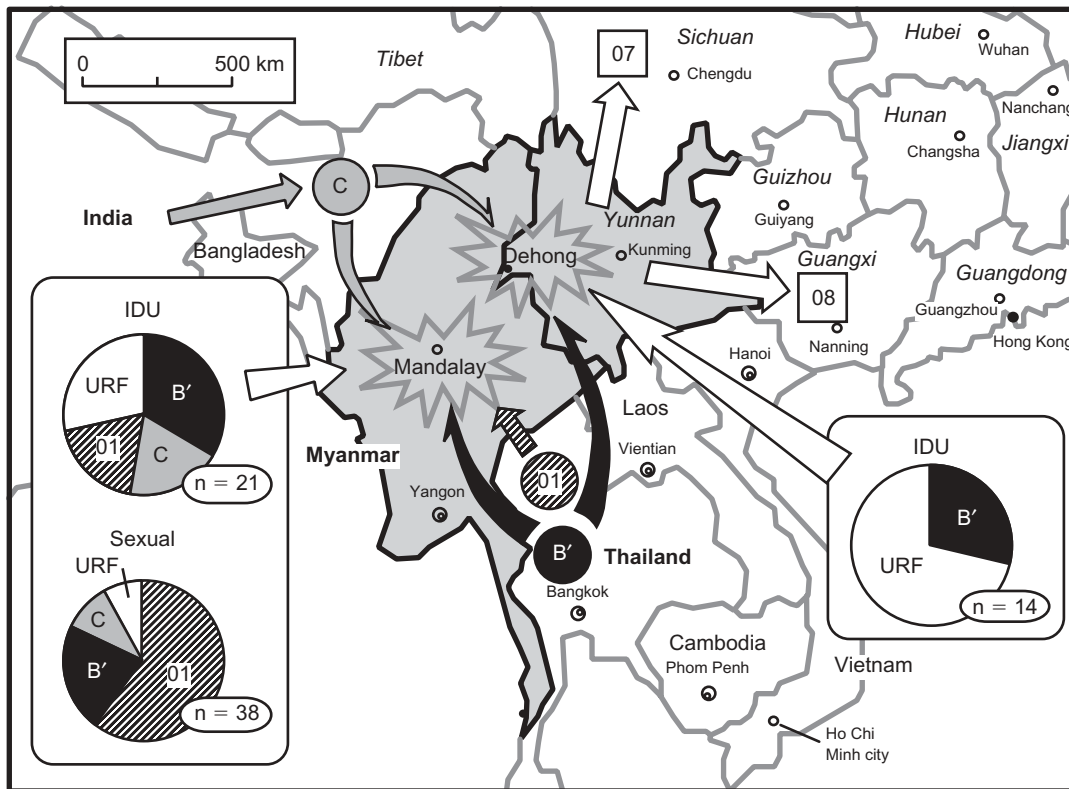


Fig. 4. Geographical hot spots of extensive recombinations between circulating HIV-1 subtypes in Southeast Asia. Western Yunnan (Dehong Prefecture) and Central Myanmar (Mandalay) (marked with 'explosion' symbols) are the 'melting pots' where extensive recombinations between different lineages of HIV-1 strains appear to be taking place continually. The pies show the prevalence of respective HIV-1 genotypes at the indicated study sites. The data for western Yunnan Province of China were adopted from the study reported by Yang *et al.* [18]. Plausible routes of HIV-1 spread are schematically illustrated. B', subtype B'; C, subtype C; IDU, injecting drug users; URF, unique recombinant form; 01, CRF01_AE; 07, CRF07_BC; 08, CRF08_BC.

Health of Myanmar, Naoki Yamamoto and Yoshiyuki Nagai for their support and Tim Mastro for his critical reading of the manuscript. We also thank Rongge Yang, Teichiro Shiino, Yuko Imamura and Kayoko Kato for their help and discussion, and Midori Kawasaki for preparation of the manuscript.

Sponsorship: This study was supported by the grants from the Ministry of Health, Labour and Welfare, the Ministry of Education, Science and Technology, and the Japanese Foundation for AIDS Prevention.

References

- Preston BD, Poiesz BJ, Loeb LA. **Fidelity of HIV-1 reverse transcriptase.** *Science* 1988, **242**:1168–1171.
- Mansky LM. **Retrovirus mutation rates and their role in genetic variation.** *J Gen Virol* 1998, **79**:1337–1345.
- Ho DD, Neumann AU, Perelson AS, Chen W, Leonard JM, Markowitz M. **Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection.** *Nature* 1995, **373**:123–126.
- Wei X, Ghosh SK, Taylor ME, Johnson VA, Emini EA, Deutsch P, *et al.* **Viral dynamics in human immunodeficiency virus type 1 infection.** *Nature* 1995, **373**:117–122.
- Perelson AS, Neumann AU, Markowitz M, Leonard JM, Ho DD. **HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time.** *Science* 1996, **271**:1582–1586.
- Goodrich DW, Duesberg PH. **Retroviral recombination during reverse transcription.** *Proc Natl Acad Sci USA* 1990, **87**:2052–2056.
- Hu WS, Temin HM. **Retroviral recombination and reverse transcription.** *Science* 1990, **250**:1227–1233.
- Stuhlmann H, Berg P. **Homologous recombination of copackaged retrovirus RNAs during reverse transcription.** *J Virol* 1992, **66**:2378–2388.
- Malim MH, Emerman M. **HIV-1 sequence variation: drift, shift, and attenuation.** *Cell* 2001, **104**:469–472.
- Carr JK, Salminen MO, Albert J, Sanders-Buell E, Gotte D, Birk DL, *et al.* **Full genome sequences of human immunodeficiency virus type 1 subtypes G and A/G intersubtype recombinants.** *Virology* 1998, **247**:22–31.
- Carr JK, Laukkanen T, Salminen MO, Albert J, Alaeus A, Kim B, *et al.* **Characterization of subtype A HIV-1 from Africa by full genome sequencing.** *AIDS* 1999, **13**:1819–1826.
- Kuiken CL, Foley B, Hahn B, Marx P, McCutchan F, Mellor J, *et al.* (eds). *Human Retroviruses and AIDS.* Los Alamos, NM: Los Alamos National Laboratory; 1999.
- McCutchan FE. **Understanding the genetic diversity of HIV-1.** *AIDS* 2000, **14** (suppl 3):S31–44.
- Salminen MO, Carr JK, Robertson DL, Hegerich P, Gotte D, Koch C, *et al.* **Evolution and probable transmission of intersubtype recombinant human immunodeficiency virus type 1 in a Zambian couple.** *J Virol* 1997, **71**:2647–2655.
- Neilson JR, John GC, Carr JK, Lewis P, Kreiss JK, Jackson S, *et al.* **Subtypes of human immunodeficiency virus type 1 and disease stage among women in Nairobi, Kenya.** *J Virol* 1999, **73**:4393–4403.

16. Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, *et al.* **Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination.** *J Virol* 1999, **73**:152–160.
17. Tovanabutra S, Polonis V, De Souza M, Trichavaroj R, Chanbancherd P, Kim B, *et al.* **First CRF01_AE/B recombinant of HIV-1 is found in Thailand.** *AIDS* 2001, **15**:1063–1065.
18. Yang R, Xia, X., Kusagawa, S., Zhang, C., Ben, K., Takebe, Y. **Ongoing generation of multiple forms of HIV-1 intersubtype recombinants in the Yunnan Province of China.** *AIDS* 2002, **16**:1401–1407.
19. Htoon MT, Lwin HH, San KO, Zan E, Thwe M. **HIV/AIDS in Myanmar.** *AIDS* 1994, **8**:S105–109.
20. Kusagawa S, Sato H, Watanabe S, Nohtomi K, Kato K, Shino T, *et al.* **Genetic and serologic characterization of HIV type 1 prevailing in Myanmar (Burma).** *AIDS Res Hum Retroviruses* 1998, **14**:1379–1385.
21. UNAIDS. *United Nations response to HIV/AIDS in Myanmar: the United Nations joint plan of action 2001–2002.* Geneva, Switzerland: UNAIDS; 2001.
22. UNAIDS/WHO. *AIDS Epidemic Update December 2001.* Geneva: UNAIDS/WHO; 2001.
23. Motomura K, Kusagawa S, Kato K, Nohtomi K, Lwin HH, Tun KM, *et al.* **Emergence of new forms of human immunodeficiency virus type 1 intersubtype recombinants in central Myanmar.** *AIDS Res Hum Retroviruses* 2000, **16**:1831–1843.
24. Kato K, Sato H, Takebe Y. **Role of naturally occurring basic amino acid substitutions in the human immunodeficiency virus type 1 subtype E envelope V3 loop on viral coreceptor usage and cell tropism.** *J Virol* 1999, **73**:5520–5526.
25. Salminen MO, Koch C, Sanders-Buell E, Ehrenberg PK, Michael NL, Carr JK, *et al.* **Recovery of virtually full-length HIV-1 provirus of diverse subtypes from primary virus cultures using the polymerase chain reaction.** *Virology* 1995, **213**:80–86.
26. Kusagawa S, Takebe Y, Yang R, Motomura K, Ampofo W, Brandful J, *et al.* **Isolation and characterization of a full-length molecular DNA clone of Ghanaian HIV type 1 intersubtype A/G recombinant CRF02_AG, which is replication competent in a restricted host range.** *AIDS Res Hum Retroviruses* 2001, **17**:649–655.
27. Thompson JD, Higgins DG, Gibson TJ. **CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice.** *Nucl Acids Res* 1994, **22**:4673–4680.
28. Saitou N, Nei M. **The neighbor-joining method: a new method for reconstructing phylogenetic trees.** *Mol Biol Evol* 1987, **4**:406–425.
29. Felsenstein J. **Confidence limits on phylogenies: an approach using the bootstrap.** *Evolution* 1985, **39**:783–791.
30. Felsenstein J. *PHYLIP (Phylogeny Inference Package) version 3.5c.* Seattle: Department of Genetics, University of Washington; 1993.
31. Salminen MO, Carr JK, Burke DS, McCutchan FE. **Identification of breakpoints in intergenotypic recombinants of HIV type 1 by bootscanning.** *AIDS Res Hum Retroviruses* 1995, **11**:1423–1425.
32. Ray SC. **SIMPLOT FOR WINDOWS, version 2.5.** (distributed by author via <http://www.welch.jhu.edu/~sray/download>) Johns Hopkins Medical Institutions, Baltimore, MD, USA. 1999.
33. Gao F, Robertson DL, Carruthers CD, Morrison SG, Jian B, Chen Y, *et al.* **A comprehensive panel of near-full-length clones and reference sequences for non-subtype B isolates of human immunodeficiency virus type 1.** *J Virol* 1998, **72**:5680–5698.
34. Robertson DL, Sharp PM, McCutchan FE, Hahn BH. **Recombination in HIV-1.** *Nature* 1995, **374**:124–126.
35. Motomura K, Kusagawa S, Lwin HH, Thwe M, Kato K, Oishi K, *et al.* **Different subtype distributions in two cities in Myanmar: Evidence for independent clusters of HIV-1 transmission.** *AIDS* 2002, **16**:633–636.
36. Peeters M. **Recombinant HIV sequences: their role in the global epidemic.** In *Human Retroviruses and AIDS: a compilation and Analysis of Nucleic Acid and Amino Acid sequences.* Edited by Kuiken C, Foley B, Hahn B, Marx P, McCutchan F, Mellors J, *et al.* Los Alamos, NM: Los Alamos National Laboratory; 2000: 54–72.