

Identification of a Novel Second-Generation Circulating Recombinant Form (CRF48_01B) in Malaysia: A Descendant of the Previously Identified CRF33_01B

Yue Li, BSc,*† Kok Keng Tee, PhD,*‡¶ Huanan Liao, PhD,* Saiki Hase, PhD,* Rie Uenishi, PhD,*
Xiao-Jie Li, MD, PhD,*§ Takayo Tsuchiura, BSc,* Rongge Yang, MD, PhD,†
Sumathi Govindasamy, BA,‡ Yean Kong Yong, MMedSc,‡ Hong Yien Tan, MMedSc,‡
Oliver G. Pybus, PhD,^{||} Adeeba Kamarulzaman, MBBS, FRACP,‡ and Yutaka Takebe, MD, PhD*†

Abstract: A molecular epidemiological investigation conducted among injecting drug users in eastern Peninsular Malaysia in 2007 identified a cluster of sequences (n = 3) located outside any known HIV-1 genotype. Analyses of near full-length nucleotide sequences of these strains from individuals with no recognizable linkage revealed that they have an identical subtype structure comprised of CRF01_AE and subtype B', distinct from any known circulating recombinant forms (CRFs). This novel CRF, designated CRF48_01B, is closely related to CRF33_01B, previously identified in Kuala Lumpur. Phylogenetic analysis of multiple CRF48_01B genome regions showed that CRF48_01B forms a monophyletic cluster within CRF33_01B, suggesting that this new recombinant is very likely a descendant of CRF33_01B. CRF48_01B thus represents one of the first examples of a "second-generation" CRF, generated by additional crossover with pre-existing CRFs. Corroborating these results, Bayesian molecular clock analyses indicated that

CRF48_01B emerged in ~2001, approximately ~8 years after the emergence of CRF33_01B.

Key Words: Asia, Bayesian coalescent analysis, circulating recombinant form, CRF33_01B, CRF48_01B, injection drug user, Malaysia, phylogeny, time of the most recent common ancestor

(*J Acquir Immune Defic Syndr* 2010;54:129–136)

INTRODUCTION

The genetic diversity of the HIV-1 epidemic in Southeast Asia is becoming increasingly complex. In the 2 decades since the HIV-1 epidemic in Thailand began in 1988–1989, 2 major circulating strains, CRF01_AE and subtype B' (Thai variant of subtype B),^{1,2} have been cocirculating in the region, leading to the generation of various recombinant strains,^{3–6} some of which have spread widely in populations, becoming circulating recombinant forms (CRFs). To date, 3 CRFs descended from CRF01_AE and subtype B' lineages are recognized in Asia: CRF15_01B^{3,7} and CRF34_01B⁸ from Thailand; and CRF33_01B from Malaysia.^{6,9}

The Joint United Nations Program on HIV/AIDS estimated that approximately 69,000 people in Malaysia were living with HIV in 2005 (<http://www.unaids.org/en/CountryResponses/Regions/Asia.asp>). Injecting drug use is the major mode of transmission in Malaysia, accounting for more than two-thirds of HIV infection.¹⁰ According to a previous study conducted in Kuala Lumpur, the capital city of Malaysia, CRF33_01B (42%) and subtype B' (40%) predominated among injecting drug users (IDUs), whereas CRF01_AE accounts for 12% of infections in this risk group.⁶ In addition to these 3 major strains, Tee et al⁶ identified diverse unique recombinant forms (URFs) in about 6% of infections, whose genomes were comprised of CRF01_AE and subtype B' regions. This suggests that cocirculation of CRF01_AE and subtype B' has repeatedly led to the generation of new recombinant strains in Malaysia.^{6,11}

Due to the development of improved genetic analysis methods and the growing size of HIV sequence databases, we can now trace the transmission history of HIV,^{12–15} estimate the time of the most recent common ancestor (tMRCA) of various epidemics.^{16,17} Molecular, epidemiological, and evolutionary analysis of different HIV-1 subtypes and CRFs enhances our

Received for publication September 25, 2009; accepted February 5, 2010.

From the *Laboratory of Molecular Virology and Epidemiology, AIDS Research Center, National Institute of Infectious Diseases, Tokyo, Japan; †The State Key Laboratory of Virology, Wuhan Institute of Virology, Chinese Academy of Sciences, Xiaohongshan, Wuhan, China; ‡Center of Excellence for Research in AIDS, University of Malaya, Kuala Lumpur, Malaysia; §Department of Dermatology, The First People's Hospital, Shanghai Jiao Tong University, Shanghai, China; ||Department of Zoology, University of Oxford, South Parks Road, Oxford, United Kingdom; and ¶Center for Infection and Immunity, Mailman School of Public Health, Columbia University, New York, NY.

Supported by grants from the Ministry of Health, Labor, and Welfare (H18-AIDS-General-016), Ministry of Education, Culture, Sports, Science, and Technology (Overseas Research Fund), the Japanese Foundation for AIDS Prevention (JFAP), the Malaysian Ministry of Science, Technology, and Innovation (eScienceFund 02-01-03-SF0379) to A.K.; and a Royal Society International Project grant.

K.K.T., H.L., R.U. and X-J. L. are recipients of the JFAP research resident fellowship. S.H. is a recipient of the research resident fellowship from the Human Sciences Foundation.

Y.L. and K.K.T. contributed equally to the study.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.jaids.com).

Correspondence to: Yutaka Takebe, MD, PhD, Laboratory of Molecular Virology and Epidemiology, AIDS Research Center, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan (e-mail: takebe@nih.go.jp).

Copyright © 2010 by Lippincott Williams & Wilkins

understanding of the genesis, transmission, and potential public health impact of local HIV-1 epidemics.

Here, we identify a novel CRF candidate among IDUs in Malaysia and find it to be a close relative of the CRF33_01B previously identified in Kuala Lumpur. By employing new dating strategy, we elucidate the likely timeline of emergence of these 2 closely related Malaysian CRFs.

MATERIALS AND METHODS

Study Subjects and Specimens

HIV-1-seropositive IDUs (n = 17) recruited in August 2007 from a fishing village located near the city of Kuantan in the eastern Peninsular Malaysia. All study participants are male, whose ages at the time of sampling ranged from 26 to 55 (mean age = 35.5). The ethnic groups in this study were Malay (100%). In our previous study, we identified 35 persons infected with CRF33_01B among a total of 184 Malay-majority HIV-1 infected.⁶ Although we do not have a complete data for the patients ethnicity, especially from those samples collected from prison, sequences did not cluster according to ethnic group. Specimens were serologically determined to be HIV-1 positive. No HIV-2 infection was detected. This study was approved by the Medical Ethics Committee of the University of Malaya Medical Center. Informed consent was obtained from all study participants before sample collection.

Determination of HIV-1 Genotype, Viral Isolation, and Near Full-Length HIV-1 Nucleotide Sequencing

Peripheral blood mononuclear cells (PBMCs) were separated by Ficoll-Hypaque density-gradient centrifugation (Amersham Biosciences AB, Uppsala, Sweden) according to the manufacturer's protocol. HIV-1 RNA was extracted from plasma and used for the initial screening of HIV-1 genotype. For virus isolation, PBMCs were cocultured with phytohemagglutinin (1 μ g/mL)-stimulated CD8⁺ T-cell-depleted PBMCs (Miltenyi Biotec GmbH, Bergisch Gladbach, Germany) from HIV-negative healthy donors in RPMI 1640 containing 10% fetal calf serum and interleukin-2 (20 U/mL) for 30 days. Virus production was monitored by the virion-associated reverse transcriptase (RT) assay described previously.¹⁸ HIV-infected PBMCs were harvested, and proviral DNA was isolated with guanidine detergent (Invitrogen, Carlsbad, CA). HIV-1 genotypes were screened on the basis of the *pol* (protease-RT) region (HXB2: 2199-3425 nucleotides), using plasma HIV-1 RNA. HIV-1 RNA was extracted using the column purification method (QIAamp Viral Mini Kit, Qiagen, Hilden, Germany) and used in a nested polymerase chain reaction (PCR). The first-round PCR was performed according to the QIAGEN OneStep RT-PCR protocol (QIAGEN OneStep RT-PCR Kit, Qiagen, Hilden, Germany), with cycling conditions of reverse transcription at 50°C for 30 minutes, followed by initial PCR activation at 94°C for 2 minutes, and 30 cycles of denaturation at 94°C for 30 seconds, annealing at 55°C for 30 seconds, extension at 72°C for 3 minutes, and a final extension at 72°C for 7 minutes. Subsequently, a 5- μ L aliquot of the first-round PCR product was subjected to a second-round

PCR. For *pol* (protease-RT) region, primers 509A (5'-AGA CAG GCT AAT TTT TTA GGG A-3', HXB2: 2074-2095 nucleotides) and 326B (5'-CTG TAC TTC TGC TAC TAA GTC TTT TGA TGG G-3', HXB2: 3539-3509 nucleotides) were used in the first-round PCR, and 510A (5'-AGA GCC AAC AGC CCC ACC AG-3', HXB2: 2148-2167 nucleotides) and 328B (5'-CTG CCA ACT CTA ATT CTG CTT C-3', HXB2: 3462-3441 nucleotides) were used for the second-round PCR.

Near full-length nucleotide sequences of HIV-1 were obtained from either provirus DNA from PBMCs or from plasma virus RNA, as previously described.^{6,19}

Phylogenetic Analysis

Nucleotide sequences of partial HIV-1 genomes (HXB2: 2199-3425 for 1.2-kilobase *pol* region; 735-3425 nucleotides for 2.7-kilobase *gag-pol* region) and near full-length nucleotide sequences (HXB2: 790-9655) were aligned with the HIV-1 reference subtypes and CRFs obtained from the Los Alamos HIV database (<http://hiv-web.lanl.gov/>), using the ClustalX 1.83.²⁰ Alignments were further adjusted by hand whilst taking codon structure into account. Phylogenetic trees were constructed using the neighbor-joining method²¹ based on the Kimura 2-parameter model with a transition-to-transversion ratio of 2.0.²² Statistical support for phylogeny nodes was obtained by bootstrap analysis with 1000 replicates.^{23,24} Analysis was implemented in MEGA 4.0 software platform.²⁵

Recombination Analysis

Gaps (insertions/deletions) were stripped before recombination analyses. SimPlot, version 3.5²⁶ was used to identify recombination breakpoints in samples with unclassified subtypes or CRFs, using a neighbor-joining algorithm with a sliding window 200 nucleotides long, overlapping by 50 nucleotides, as previously described.²⁷ Reference sequences for subtype A (92UG037), subtype B' (RL42), subtype B (RF), subtype C (95IN21068), subtype D (NDK), CRF01_AE (CM240), subtype F (VI850), subtype G (SE6165), subtype H (VI991), and subtype J (SE7022) and K (MP535) were used in the Simplot analysis and subtype B' (CN.RL42 or 96TH.M041), CRF01_AE (90TH.CM240) and C (95IN21068) were used in the Bootscanning analysis. Subregion confirmatory tree analyses were performed to confirm the subtype within each region, and a bootstrap value of $\geq 70\%$ was considered to be definitive.

Estimation of The Time of the Most Recent Common Ancestor

Investigation of the evolutionary history of CRF33_01B and CRF48_01B was carried out using BEAST v1.4,¹⁷ thereby estimating the date of the most recent common ancestor of each phylogenetic cluster.^{28,29} Reference nucleotide sequences belonging to CRF01_AE (n = 37) and subtype B (n = 55) with known sampling dates (ranging from 1990 to 2006 and from 1983 to 2005, respectively) were retrieved from the Los Alamos HIV Sequence Database (<http://www.hiv.lanl.gov/content/index>). A posterior distribution of phylogenies was obtained using a Bayesian Markov chain Monte Carlo

(MCMC) approach, as implemented in BEAST v1.4.¹⁷ Dates were estimated using Bayesian MCMC inference under both the general time-reversal (GTR) and Hasegawa-Kishino-Yano (HKY) nucleotide substitution models, with a gamma-distribution model of among site rate heterogeneity (with 4 rate categories)^{30–32} with a relaxed clock model,³³ using 2 different coalescent models: constant population size and the Bayesian skyline model. Each analysis was computed for 10 million states sampled every 10000 states. The MCMC output was tested for convergence and effective sampling sizes were computed using Tracer v1.4 (available from <http://beast.bio.ed.ac.uk>). Rates of evolution were estimated from the abovementioned CRF01_AE and subtype B datasets, which were sampled over a wide range of times. The estimated evolutionary rates were then incorporated as fixed prior values during the analysis of the CRF33_01B and CRF48_01B sequences.

Nucleotide Sequence Accession Numbers

The sequences reported in this study have been deposited in the GenBank database (accession numbers: GQ175881-GQ175903).

RESULTS

Identification of a New CRF (CRF48_01B) in Eastern Peninsular Malaysia

We carried out a molecular epidemiological investigation of HIV-1-seropositive IDUs ($n = 17$) recruited in August 2007 in a fishing village in Kuantan city in the eastern coast of Peninsular Malaysia. Initial characterization of HIV-1 genotype was based on phylogenetic analysis of the nucleotide sequence of the protease and RT regions (HXB2: 2199–3425 nt), which classified 76.5% (13 of 17) of strains as CRF33_01B and 5.9% (1 of 17) of strains as subtype B'. In addition, we identified a cluster of sequences (3 of 17, 17.6%) that were located outside any known HIV-1 genotype on the neighbor-joining tree and were supported by a high bootstrap score (100%) (Fig. 1; for full phylogenetic tree, see **Supplemental Digital Content 1**, <http://links.lww.com/QAI/A41>).

To define the subtype structure of these 3 strains, we determined near full-length HIV-1 nucleotide sequences from either proviral DNA (07MYKT016 and 07MYKT021) or plasma RNA (07MYKT014) from 3 individuals, among whom there was no apparent epidemiological linkage. Bootscanning analyses²⁶ revealed that these 3 strains shared a common recombinant structure: a short subtype B' region approximately 1.0-kilobase long (HXB2: 2064–3102) was found within a backbone of CRF01_AE (Fig. 2A). The subtype structure of these new CRF01_AE/B recombinants is distinct from the previously identified CRF15_01B, CRF33_01B and CRF34_01B, although considerable structural similarity to CRF33_01B, previously identified in Kuala Lumpur, was observed (Fig. 2A). Because these 3 strains from Kuantan have identical recombinant genome structures, they fulfill the criteria for the designation of a new "CRF"³⁴ (<http://www.hiv.lanl.gov/content/sequence/HIV/CRFs/>). They constitute the forty-sixth CRF identified in the global pandemic, and we therefore propose the designation CRF48_01B.

CRF48_01B Is a Descendant of CRF33_01B Previously Identified in Kuala Lumpur

To confirm the subtype structure of CRF48_01B and to estimate likely parental lineages, we performed subgenomic phylogenetic analyses, in which the HIV-1 genome was divided into 3 regions (denoted I, II, and III) as illustrated in Fig. 2A. Of note, 1 recombination breakpoint (site 1, HXB2: 2058 \pm 5 basepairs) between segments I and II was found to be identical in both CRF33_01B and CRF48_01B (Fig. 2A), suggesting a close evolutionary relationship between these 2 recombinants.

As shown in Fig. 2B, regions I and III of CRF48_01B are most closely related to CRF01_AE of Thai origin and form a well supported monophyletic cluster within CRF33_01B. In contrast, the CRF01_AE subgenomic regions of CRF15_01B and CRF34_01B (both identified in Thailand) were more distantly related to CRF33_01B and CRF48_01B. Therefore, although the overall structure of the 5' part of the CRF34_01B genome resembles that of CRF33_01B and CRF48_01B, subgenomic phylogenies reveal that the 2 pairs of recombinant forms have different origins. Region II of CRF48_01B is classified as subtype B', which is commonly associated with IDUs in Southeast Asian countries^{1,35} (Fig. 2B).

To further investigate the evolution of the subtype B' regions of CRF48_01B and CRF33_01B, we performed an additional subgenomic phylogenetic analysis of the concatenated subtype B' regions (IIa + IIb) that are common to both CRF33_01B and CRF48_01B. The subtype B' segments of CRF48_01B form a well-supported monophyletic cluster which, in the neighbor-joining tree, forms a sister lineage to CRF33_01B. However, this sister-relationship has no bootstrap support and relies on a single short branch (Fig. 2B). Therefore, the most parsimonious evolutionary explanation is that CRF33_01B is paraphyletic with respect to CRF48_01B³⁶ in this region, just as it is in regions I and III. Taken together, these results strongly suggest that CRF48_01B arose as a separate lineage from the CRF33_01B; in other words, CRF48_01B is very likely a descendant of CRF33_01B.

Timeline of Emergence of CRF33_01B and CRF48_01B in Malaysia

Next, we investigate the possible timeline of emergence of these 2 closely related CRFs (CRF33_01B and CRF48_01B) in Malaysia, using the Bayesian relaxed molecular clock approach implemented in BEAST v1.4. The rates of evolution (μ) of the subgenomic regions I, I + IIIa, and IIa + IIb (Fig. 3A) were estimated (Table 1). As shown in Table 1 and Figure 3, when the GTR + Γ_4 Constant-size relaxed model was used, the tMRCA of CRF33_01B and CRF48_01B were dated to around 1993 and 2001, respectively. In contrast, the dates of origin of African and Thailand CRF01_AE were significantly older and estimated to be around 1974 and 1982, respectively (Fig. 3, Table 1), essentially consistent to the results of previous studies.^{12,37} The choices of different models or genome regions (see **Figure, Supplemental Digital Content 2**, <http://links.lww.com/QAI/A42>) have no significant effect on the estimated dates.

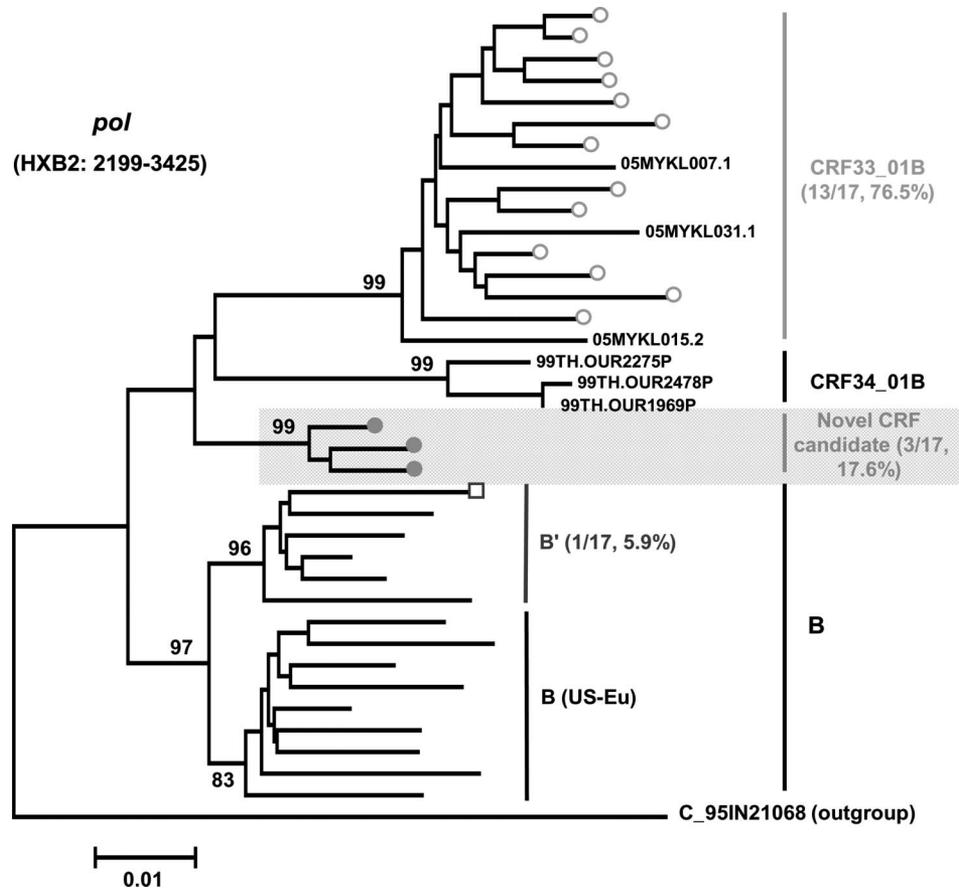


FIGURE 1. Identification of novel CRF candidate. The neighbor-joining tree is constructed based on the nucleotide sequence of protease-RT region (HXB2: 2199-3425 nt). Tree was rooted with subtype C (95IN21068). HIV-1 CRF01_AE and subtype B/B' reference strains retrieved from the Los Alamos HIV Sequence Database (<http://www.hiv.lanl.gov/content/index>). The HIV-1 sequences (n = 17) obtained from this study were indicated as follows: CRF33_01B (open circle) (n = 13); novel CRF candidate (CRF48_01B) (solid circle) (n = 3); subtype B' (open square) (n = 1). Bootstrap values $\geq 70\%$ are shown.

DISCUSSION

The present study identified a novel CRF (CRF48_01B) comprised of CRF01_AE and subtype B' among IDUs in eastern Peninsular Malaysia (Fig. 2). CRF48_01B is closely related to CRF33_01B previously identified in Kuala Lumpur,⁶ with which it shares 1 recombination breakpoint (site 1 in Fig. 2A). Because recombination breakpoints are very likely to be conserved during viral evolution, the presence of a shared recombination breakpoint suggests a direct parental/progenitor relationship between these 2 CRFs. The phylogenetic analyses of multiple CRF48_01B genome regions (regions I, IIa + IIb, and III in Fig. 2A) showed that CRF48_01B forms a monophyletic cluster within CRF33_01B (Fig. 2B). In other words, CRF33_01B has paraphyletic relationship with respect to CRF48_01B.³⁶ Taken together, these observations strongly suggest that CRF48_01B is indeed a descendant of CRF33_01B, created by additional crossing-over of a subtype B' strain in the protease-RT region (Fig. 2A).

Because the genome structure of CRF33_01B seems more complicated than that of CRF48_01B and that multiple genome cross overs are required for the generation of CRF48_01B, it is tempting to hypothesize that CRF48_01B was generated through recombination between CRF01_AE and subtype B' rather than between CRF33_01B and subtype B'. However, as previously mentioned, the phylogenetic analyses point toward CRF48_01B being nested within

CRF33_01B, which in turn is nested with CRF01_AE (Fig. 2), indicating that CRF33_01B is more ancestral than CRF48_01B. Of course, if HIV-1 recombination is common, then we should not expect the most parsimonious number of breakpoints to always represent the true evolutionary history.

CRF48_01B is thus one of the first examples of a "second-generation" CRF generated by recombination between pre-existing CRFs and other HIV-1 strains. In this respect, CRF48_01B may be more accurately represented by the label CRF48_33B. Other possible examples of "second-generation" CRFs so far reported include CRF30_0206 (Niger, West Africa),³⁸ CRF32_06A1 (Estonia)³⁹ and CRF43_02G (Saudi Arabia)⁴⁰.

The phylogenetic and Bayesian coalescent analysis performed using BEAST v1.4 (<http://beast.bio.ed.ac.uk>) revealed that CRF48_01B emerged around 1993, approximately 8 years after the emergence of CRF33_01B in Malaysia (Fig. 3, Table 1). This is consistent with the notion that CRF33_01B is a parental strain of CRF48_01B (CRF48_33B). To obtain more accurate estimates regarding the time of origin, movement and phylogenetic relationships of these 2 new CRFs, we intend to extend this study in future to a wider range of sampling times and locations in Malaysia.

The prevalence and public importance of the newly identified CRF48_01B remain to be evaluated. On-going large-scale surveillance of HIV-1 drug resistance mutations in Malaysia has recently identified 1 IDU case infected with

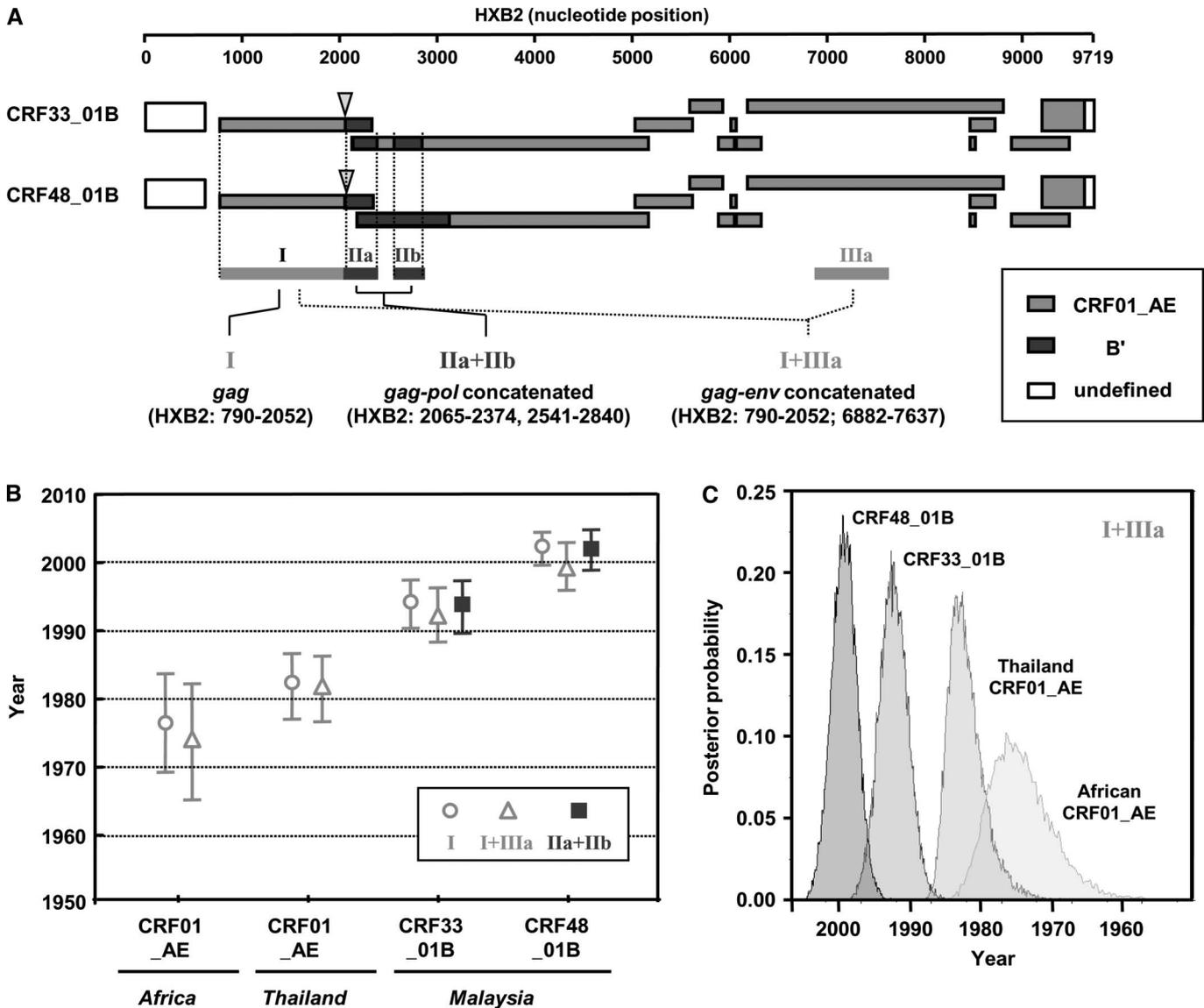


FIGURE 3. Estimated dates of origin of CRF33_01B and CRF48_01B in comparison with African and Thailand CRF01_AE. **A**, Molecular clock analyses were performed using BEAST v1.4 (see methods for details) for different genome regions: Region I, gag CRF01_AE (HXB2: 790-2052); Region I+IIIa, gag-env CRF01_AE (concatenated) (HXB2: 790-2052; 6882-7637); Region IIa+IIb, gag-pol subtype B' (concatenated) (HXB2: 2065-2374; 2541-2840). **B**, Estimated dates of the tMRCA of four sequence clusters, obtained using analysis of different subgenomic regions. Estimates were calculated under a GTR substitution model with a relaxed molecular clock and a constant size coalescent model. Vertical bars represent the 95% credible region (confidence limits) for each estimate. ○ = Region I; □ = Region I+IIIa; ■ = Region IIa+IIb. **C**, The full posterior probability distributions of the of each cluster. This plot represents the distributions obtained for the concatenated gag-env CRF01_AE region (Region I+III). The tMRCA means are as follows: CRF33_01B ancestor = ~1993; CRF48_01B ancestor = ~2001; Thailand CRF01_AE ancestor = ~1982; African CRF01_AE ancestor = ~1974 (Table 1).

CRF48_01B among 120 HIV-1-infected individuals in Kuala Lumpur (Ong et al, BSc, unpublished data, June 2009). This suggests that CRF48_01B may be spreading in at least 2 cities (Kuantan and Kuala Lumpur) in Malaysia.

Previous studies, including ours,^{6,9,11,41} have suggested that new URFs seem to be continuously emerging through ongoing coinfection and recombination and thus contribute to shaping the nature of the HIV-1 epidemic in Malaysia. It is thus likely that as

FIGURE 2. (continued) between CRF33_01B and CRF48_01B. Inset: map of Southeast Asia, with the locations of Kuantan (KT) and Kuala Lumpur (KL) in Malaysia (shadowed) indicated. **B**, Subgenomic phylogenies, estimated using the neighbor-joining method from alignments representing regions I, II, III and the concatenated IIa+IIb region. Bootstrap scores greater than 70% are indicated at corresponding nodes. "01TH" = Thailand CRF01_AE; "01AF" = African CRF01_AE; open diamond = CRF15_01B; open circle = CRF33_01B; solid diamond = CRF34_01B; solid circle = CRF48_01B. The HIV-1 genotype to which CRF48_01B belongs to is shadowed.

TABLE 1. Estimated Substitution Rates and Dates Origin for CRF33_01B and CRF48_01B in Comparison With African and Thailand CRF01_AE*

	Genome Region	Rate of Evolution†	Coefficient of Variation	Date of tMRCAs (yr)‡			
				CRF33_01B	CRF48_01B	African CRF01_AE	Thai CRF01_AE
I	gag CRF01_AE (HXB2: 790-2052)	2.8 (2.1, 3.7)	0.5 (0.2, 0.7)	1994.0 (1990.4, 1997.3)	2002.2 (1999.7, 2004.5)	1976.4 (1969.1, 1983.6)	1982.3 (1976.9, 1986.8)
I + IIIa	gag-env CRF01_AE (concatenated) (HXB2: 790-2052; 6882-7637)	3.8 (2.8, 4.9)	0.5 (0.3, 0.7)	1992.3 (1988.3, 1996.2)	1999.4 (1995.9, 2002.9)	1974.2 (1965.1, 1982.2)	1982.0 (1976.7, 1986.4)
IIa + IIb	gag-pol subtype B' (concatenated) (HXB2: 2065-2374; 2541-2840)	4.1 (2.9, 5.4)	0.5 (0.3, 0.8)	1993.7 (1989.5, 1997.5)	2001.9 (1998.8, 2004.8)	N/A	N/A
	Mean	—	—	1993.3 (1989.4, 1997.0)	2001.2 (1998.1, 2004.1)	1974.1 (1967.1, 1982.9)	1982.2 (1976.8, 1986.6)

*Based on BEAST analysis under a GTR substitution model with a relaxed molecular clock and a constant size coalescent model.

†Estimates of the mean evolutionary rate (μ : $\times 10^{-3}$ nucleotide substitutions/site/year) for either CRF01_AE or subtype B'.

‡Mean time of the most common ancestor (tMRCAs: year) for the CRF01_AE dataset [95% highest posterior density (HPD) in parentheses].

yet undefined new CRFs could evolve from such URFs, spreading more widely through the population, especially through IDU transmission networks which are often characterized by high rates of coinfection. The present study emphasizes the importance of comprehensive and continued surveillance of HIV-1 strains in Malaysia and neighboring countries.

ACKNOWLEDGMENTS

We thank Naoki Yamamoto for his support and encouragement and Midori Kawasaki for preparation of manuscript. We also thank study participants for their cooperation.

REFERENCES

- Ou CY, Takebe Y, Weniger BG, et al. Independent introduction of two major HIV-1 genotypes into distinct high-risk populations in Thailand. *Lancet*. 1993;341:1171-1174.
- Kalish ML, Baldwin A, Raktham S, et al. The evolving molecular epidemiology of HIV-1 envelope subtypes in injecting drug users in Bangkok, Thailand: implications for HIV vaccine trials. *AIDS*. 1995;9:851-857.
- Tovanabutra S, Polonis V, De Souza M, et al. First CRF01_AE/B recombinant of HIV-1 is found in Thailand. *AIDS*. 2001;15:1063-1065.
- Kijak GH, Currier JR, Tovanabutra S, et al. Lost in translation: implications of HIV-1 codon usage for immune escape and drug resistance. *AIDS Rev*. 2004;6:54-60.
- Matsuoka-Aizawa S, Sato H, Hachiya A, et al. Isolation and molecular characterization of a nelfinavir (NFV)-resistant human immunodeficiency virus type 1 that exhibits NFV-dependent enhancement of replication. *J Virol*. 2003;77:318-327.
- Tee KK, Li XJ, Nohtomi K, et al. Identification of a novel circulating recombinant form (CRF33_01B) disseminating widely among various risk populations in Kuala Lumpur, Malaysia. *J Acquir Immune Defic Syndr*. 2006;43:523-529.
- Tovanabutra S, Watanaveeradej V, Viputtikul K, et al. A new circulating recombinant form, CRF15_01B, reinforces the linkage between IDU and heterosexual epidemics in Thailand. *AIDS Res Hum Retroviruses*. 2003;19:561-567.
- Tovanabutra S, Kijak GH, Beyrer C, et al. Identification of CRF34_01B, a second circulating recombinant form unrelated to and more complex than CRF15_01B, among injecting drug users in northern Thailand. *AIDS Res Hum Retroviruses*. 2007;23:829-833.
- Tee KK, Pon CK, Kamarulzaman A, et al. Emergence of HIV-1 CRF01_AE/B unique recombinant forms in Kuala Lumpur, Malaysia. *AIDS*. 2005;19:119-126.

- Reid G, Kamarulzaman A, Sran SK. Malaysia and harm reduction: the challenges and responses. *Int J Drug Policy*. 2007;18:136-140.
- Wang B, Lau KA, Ong LY, et al. Complex patterns of the HIV-1 epidemic in Kuala Lumpur, Malaysia: evidence for expansion of circulating recombinant form CRF33_01B and detection of multiple other recombinants. *Virology*. 2007;367:288-297.
- Korber B, Muldoon M, Theiler J, et al. Timing the ancestor of the HIV-1 pandemic strains. *Science*. 2000;288:1789-1796.
- Rambaut A, Posada D, Crandall KA, et al. The causes and consequences of HIV evolution. *Nat Rev Genet*. 2004;5:52-61.
- Lemey P, Rambaut A, Pybus OG. HIV evolutionary dynamics within and among hosts. *AIDS Rev*. 2006;8:125-140.
- Pybus OG, Rambaut A. Evolutionary analysis of the dynamics of viral infectious disease. *Nat Rev Genet*. 2009;10:540-550.
- Drummond AJ, Rambaut A, Shapiro B, et al. Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol Biol Evol*. 2005;22:1185-1192.
- Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol*. 2007;7:214.
- 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.
- Salminen MO, Koch C, Sanders-Buell E, 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.
- Thompson, JD, Gibson, TJ, Plewniak, F, et al. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res*. 1997;25:4876-4882.
- Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol*. 1987;4:406-425.
- Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol*. 1980;16:111-120.
- Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*. 1985;39:783-791.
- Felsenstein J. *PHYLIP (Phylogeny Inference Package)* [computer program]. Version 3.6a3. Seattle, WA: Department of Genome Sciences, University of Washington. 2002.
- Tamura K, Dudley J, Nei M, et al. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol*. 2007;24:1596-1599.
- Lole KS, Bollinger RC, Paranjape RS, 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.

27. Carr JK, Salminen MO, Koch C, et al. Full-length sequence and mosaic structure of a human immunodeficiency virus type 1 isolate from Thailand. *J Virol*. 1996;70:5935–5943.
28. Pybus OG, Drummond AJ, Nakano T, et al. The epidemiology and iatrogenic transmission of hepatitis C virus in Egypt: a Bayesian coalescent approach. *Mol Biol Evol*. 2003;20:381–387.
29. Drummond AJ, Nicholls GK, Rodrigo AG, et al. Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. *Genetics*. 2002;161:1307–1320.
30. Hasegawa M, Kishino H, Yano T. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J Mol Evol*. 1985;22:160–174.
31. Rodriguez F, Oliver JL, Marin A, et al. The general stochastic model of nucleotide substitution. *J Theor Biol*. 1990;142:485–501.
32. Yang Z. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *J Mol Evol*. 1994;39:306–314.
33. Drummond AJ, Ho SY, Phillips MJ, et al. Relaxed phylogenetics and dating with confidence. *PLoS Biol*. 2006;4:e88.
34. Robertson DL, Anderson JP, Bradac JA, et al. HIV-1 nomenclature proposal. *Science*. 2000;288:55–56.
35. Weniger BG, Takebe Y, Ou CY, et al. The molecular epidemiology of HIV in Asia. *AIDS* 1994;8(Suppl 2):S13–S28.
36. Abecasis AB, Lemey P, Vidal N, et al. Recombination confounds the early evolutionary history of human immunodeficiency virus type 1: subtype G is a circulating recombinant form. *J Virol*. 2007;81:8543–8551.
37. Liao H, Tee KK, Hase S, et al. Phylodynamic analysis of the dissemination of HIV-1 CRF01_AE in Vietnam. *Virology*. 2009.
38. Mamadou S, Vidal N, Montavon C, et al. Emergence of complex and diverse CRF02-AG/CRF06-cpx recombinant HIV type 1 strains in Niger, West Africa. *AIDS Res Hum Retroviruses*. 2003;19:77–82.
39. Adojaan M, Kivisild T, Mannik A, et al. Predominance of a rare type of HIV-1 in Estonia. *J Acquir Immune Defic Syndr*. 2005;39:598–605.
40. Yamaguchi J, Badreddine S, Swanson P, et al. Identification of new CRF43_02G and CRF25_cpx in Saudi Arabia based on full genome sequence analysis of six HIV type 1 isolates. *AIDS Res Hum Retroviruses*. 2008;24:1327–1335.
41. Lau KA, Wang B, Kamarulzaman A, et al. Near full-length sequence analysis of a Unique CRF01_AE/B recombinant from Kuala Lumpur, Malaysia. *AIDS Res Hum Retroviruses*. 2007;23:1139–1145.