DOI: 10.35772/ghm.2024.01047

Identification of new circulating recombinant form of HIV-1 CRF139_02B in Japan, and search for the origin

Tsunefusa Hayashida, Kiyoto Tsuchiya, Shinichi Oka, Hiroyuki Gatanaga*

AIDS Clinical Center, National Center for Global Health and Medicine, Tokyo, Japan.

Abstract: Many circulating recombinant forms (CRFs) of HIV-1 have been reported, resulting in complex molecular epidemiology of HIV-1 infection. In this study, we newly identified CRF139_02B in Japan from 4 cases of antiretroviral therapy naïve people living with HIV. Near full-length genome sequences of CRF139_02B were determined using Illumina MiSeq. Basic Local Alignment Search Tool (BLAST) revealed that there were several sequences having the same breakpoints as CRF139_02B in the UK and Nepal, though its full-length genome sequences were not available. Maximum clade credibility tree analysis using the region of protease and reverse transcriptase of HIV-1 estimated that the time to the most recent common ancestor of CRF139_02B variants found in Japan was 2017.6 (95% highest posterior density interval: 2015.9-2019.3), and that among the UK, Nepal, and Japan was 2010.4 (2007.8-2012.5). These results suggested that CRF139_02B circulated in Japan recently and domestically. Furthermore, the origin of CRF139_02B could be in the UK. Because there is a possibility that further international circulation of CRF139_02B may be observed in the near future, continuous monitoring of HIV-1 molecular epidemiology will be needed.

Keywords: HIV, Japan, circulating recombinant form, phylogenetic tree analysis, time to the most recent common ancestor

Introduction

Development of antiretroviral therapy (ART) averted approximately 20.8 million AIDS-related deaths in the past three decades (1). However, the AIDS epidemic is still a public health threat and should be ended (2). In 2022, people living with HIV (PLWH) were 39 million, people acquiring HIV were 1.3 million, and people dying from HIV-related causes were 630,000 (3). HIV-1 group M is the majority of the global epidemic of HIV infections (4) and has diverged into 10 subtypes (A-D, F-H, J-L) (5). Furthermore, many intersubtype recombinants of HIV-1 have been identified (4). Typically, recombinants found in 3 or more epidemiologically unrelated PLWH were distinguished as circulating recombinant forms (CRFs) and were numbered sequentially. In May 2024, over 150 CRFs were reported to the HIV Sequence Database of Los Alamos National Laboratory (https://www.hiv.lanl. gov/content/index). Recombination of HIV-1 makes viruses diversified (6,7) and molecular epidemiology complicated (8, 9).

In Japan, newly diagnosed cases of HIV-1 infection were 884 in 2022, which was the lowest number in the past 20 years (10). On the other hand, the percentage

of people who were diagnosed with HIV infection per PLWH was estimated at 85% (11,12) and was below the 95% targeted by WHO (13). The dominant strain of HIV-1 was subtype B (82.8%), followed by CRF 01_ AE (7.2%), among people newly diagnosed with HIV in Japan from 2003 to 2019 (14). Two CRFs have been identified in Japan and were composed of subtype B and CRF01_AE (15,16). Recently, we found a candidate for a new CRF in PLWH visiting our outpatient clinic. The objective of this study was to determine whether these HIV-1 strains were new CRFs and to explore their molecular epidemiology.

Materials and Methods

Participants and sample collection

In AIDS Clinical Center, National Center for Global Health and Medicine (NCGM), pre-treatment HIV-1 drug resistance testing was performed before the introduction of ART by determining protease and reverse transcriptase sequences in 2,482 ART-naïve PLWH from January 2003 to June 2022. During the analysis of these sequences, we noticed that 4 cases had a similar unique recombination pattern in protease and reverse transcriptase of HIV-1

composed of subtype B and CRF02_AG. Their stocked plasma or serum samples and clinical information at (or nearby) the first visit were used for analysis anonymously. The study was approved by the Human Ethics Committee of NCGM (#NCGM-A-000172-01), and each participant also signed informed consent in accordance with the Declaration of Helsinki.

Near full-length genome sequencing of HIV-

Near full-length genome (NFLG) sequencing of HIV-1 (HXB2 position: 706-9531) was performed using the previously published method (17, 18) with little modification. Briefly, viral RNA was extracted from plasma or serum samples using QIAamp Viral RNA Mini Kit (Qiagen., Venro, Nederland). Reverse-transcription followed by polymerase chain reaction (PCR) and nested PCR was performed to amplify two (or four) segments of NFLG of HIV-1 using PrimeScript II High Fidelity One Step RT-PCR Kit and PrimeSTAR GXL DNA Polymerase (Takara Bio Inc., Shiga, Japan) (19,20). Nextera XT DNA Library Preparation Kit and MiSeq (Illumina, San Diego, CA, USA) were used for library preparation and next generation sequencing (NGS). Vicuna was used for the de novo assembly of short read NGS data (21). Short read NGS data were mapped into a tentative NFLG sequence made by Vicuna using BWA-MEM (22). Mapped NFLG sequence data was visualized using IGV (23) and polished to get final consensus sequence. Anti-HIV drug resistant mutation was checked by HIV DRUG RESISTANCE DATABASE of Stanford University (https://hivdb.stanford.edu). Viral tropism of coreceptor usage was checked with geno2pheno [coreceptor] (https://coreceptor.geno2pheno.org).

Analysis of breakpoints and phylogenetic tree

Breakpoints of recombination were estimated using jpHMM (24) and RIP (25). Reference sequences for phylogenetic tree analysis were obtained from the Nucleotide Basic Local Alignment Search Tool (BLAST) (*https://blast.ncbi.nlm.nih.gov/Blast.cgi*), and from the HIV Sequence Database of Los Alamos National Laboratory. Phylogenetic tree analysis was conducted by MEGA7 using the neighbor-joining method with Kimura two-parameter model and 1,000 bootstrap replications (26). Time to the most recent common ancestor (tMRCA)

was estimated by maximum clade credibility (MCC) tree analysis using the Bayesian Markov chain Monte Carlo method. The analysis was conducted by BEAST v2.4.6 with 100,000,000 states using GTR +G +I, relaxed clock log normal, and Coalescent Bayesian Skyline model (27).

Results and Discussion

The clinical information of the 4 participants is shown in Table 1. They were all men who have sex with men. Two of them were Japanese, and the remaining 2 cases were non-Japanese. A certain epidemiological link was not observed among them. We succeeded in NFLG sequencing of HIV-1 in all participants. No sign of dual infection was found in the result of the de novo assembly made by Vicuna and in the mapping images depicted by IGV. We registered 4 NFLG sequences of HIV-1 to DDBJ. Accession numbers are as follows: ACCRF1, LC762505; ACCRF2, LC762506; ACCRF3, LC762507; and ACCRF4, LC762508. Next, we analyzed the recombination pattern of these sequences. First, jpHMM analysis suggested that these 4 sequences had the same recombination pattern composed of subtype A, B, G, and CRF01 AE (Figure 1A). Second, RIP analysis suggested that these 4 sequences were more likely composed of subtype B and CRF02_AG, rather than subtype A, G, and CRF01 AE (Figure 1B). To confirm the recombination pattern suggested by RIP, we conducted phylogenetic tree analysis in 3 regions; region I (HXB2 position:706-3141), region II (3142-4855), and region III (4856-9531) (Figure 1C). It revealed that these 4 sequences made significant clusters (bootstrap score 99) and belonged to CRF02 AG in region I (Figure 1D), to subtype B in region II (Figure 1E), and to CRF02 AG in region III (Figure 1F).

Nucleotide BLAST with our NFLG sequences revealed that even the most similar sequence to ACCRF1 had only 91.11% identity, suggesting that no sequence was close to ACCRF1. We also checked known CRFs composed of subtype B and CRF02_AG in the HIV DATABASE of Los Alamos National Laboratory. Three CRFs (CRF56_cpx, CRF94_cpx, and CRF95_02B) were found, though their recombination patterns were different from our 4 sequences. Thus, we contacted HIV DATABASE to resister our new CRF. After that, HIV DATABASE assigned the following code to the new CRF: CRF139_02B. This was the third CRF of HIV-

Table 1. Information of study participants at (or nearby) first visit

ID	Sex	Transmission route	Age	Nationality	Last HIV negative	Sampling	CD4 (/µL)	Viral load (/mL)
ACCRF1	Male	MSM	20s	China	2020/6	2020/9	326	789,000
ACCRF2	Male	MSM	30s	Japan	NA	2020/10	322	22,800
ACCRF3	Male	MSM	30s	Mexico	2019	2021/10	263	140,000
ACCRF4	Male	MSM	20s	Japan	2021/8	2022/1	430	405,000

MSM: men who have sex with men; NA: not available.



Figure 1. Analysis of recombination in CRF139_02B. (A) Recombination pattern of ACCRF1 analyzed by jpHMM; (B) Recombination pattern of ACCRF1 analyzed by RIP; (C) Recombination pattern of ACCRF1 depicted by Recombinant HIV-1 Drawing Tool; (D) Phylogenetic tree analysis of region I (HXB2 position: 706-3141); (E) Phylogenetic tree analysis of region III (4856-9531). The closed circle showed CRF139_02B in Japan. The number of nearby nodes showed a bootstrap score ≥ 90 .

1 identified in Japan, to our knowledge. Previously identified 2 CRFs of HIV-1 in Japan (CRF69_01B (15) and CRF76_01B (16)) were composed of 2 major subtypes in Japan (B and CRF01_AE) (14). While, CRF139_02B was composed of subtype B and CRF02_ AG, the latter was a minor subtype (1.2%) in Japan (14). All variants of CRF139_02B did not have any anti-HIV drug resistant mutation in Protease, Reverse transcriptase, Integrase, and Capsid region. Additionally, these viral tropism of coreceptor usage were CCR5.

To search for a similar sequence to CRF139_02B, we used Nucleotide BLAST with 2 regions; region IV was protease and reverse transcriptase (HXB2 position: 2253-3509) and region V was integrase (4230-5093) (Figure 2A). There were 3 reasons why we chose these regions. First, a part of CRF139_02B genome might have a similar sequence, though we could not find any similar sequence to NFLG of CRF139 02B. Second, substantial reference sequence data were available because drug resistance testing of HIV-1 had been conducted in all over the world. Third, these regions contained the breakpoint of recombination in CRF139_02B. The top 15 sequences of max score in BLAST were added to phylogenetic tree analysis in each region. In region IV, 9 sequences from the UK (28,29) and 1 sequence from Nepal (30) were found near CRF139 02B sequences with 100 bootstrap

score (Figure 2B). In region V, 7 sequences from the UK (28) were found near CRF139 02B sequences with 100 bootstrap score (Figure 2C). It turned out that some sequences in region IV and region V were obtained from the same PLWH (e.g. MT571106 and MT570777) (28). Thus, we united paired sequences and analyzed recombination patterns with RIP (Figure 3). It turned out that 6 cases (MT570597, MT571106, MT571148, MT571077, MT571085, and MT571086) had the 2 same breakpoints as CRF139 02B. Additionally, 2 cases (MZ538466 and MT570773) had the same breakpoint as CRF139 02B, though only one region sequence data was available in each case. It suggested that these 8 cases can be CRF139 02B, though their NFLG sequences were not available. The other 3 cases (MT571265, MF109637, and MT571025) also had the same breakpoint as CRF139 02B in protease and reverse transcriptase region, at the same time, however, they had subtype F recombination. Thus, these 3 sequences were not CRF139 02B, but were a related recombinant to CRF139 02B. According to Yebra G, et al., a part of the MF109637 genome derived from subtype F was close to the pure subtype F sequence observed in the UK, suggesting that MF109637 was generated in the UK (29).

To search for the origin of CRF139_02B, MCC tree analysis was performed. To eliminate the effect of



Figure 2. Search for similar sequences to CRF139_02B. (A) Two regions were used for BLAST; (B) Phylogenetic tree analysis of region IV (HXB2 position: 2253-3509); (C) Phylogenetic tree analysis of region V (4230-5093). The closed circle showed CRF139_02B in Japan. The number of nearby nodes showed a bootstrap score ≥ 90 .



Figure 3. Recombination pattern of similar sequences to CRF139_02B. Each recombination pattern was analyzed by RIP and depicted by Recombinant HIV-1 Drawing Tool. LC762505 (at the top) was ACCRF1.



Figure 4. Search for the origin of CRF139_02B. (A) Two regions were used for maximum clade credibility (MCC) tree analysis; (B) MCC tree analysis of region VI (HXB2 position: 2253-3141); (C) MCC tree analysis of region VII (4230-4855). The closed circle showed CRF139_02B in Japan. The open circle showed CRF139_02B in the UK or Nepal. The dotted circle showed a node of the common ancestor. The number near the dotted circle showed posterior probability.

recombination, we used regions composed of a single subtype or CRF; region VI (HXB2 position: 2253-3141), and region VII (4230-4855) (Figure 4A). MCC tree analysis was performed with only CRF02 AG or CRF139_02B sequences in region VI (Figure 4B). Four sequences of CRF139 02B in Japan formed a cluster, and tMRCA was 2017.6 (95% highest posterior density (95% HPD) interval: 2015.9-2019.3). At the upstream, CRF139 02B sequences in the UK and Nepal were located and made a cluster with CRF139 02B sequences in Japan, and tMRCA was 2010.4 (2007.8-2012.5). Further upstream, pure CRF02 AG sequences in the UK were located. These results revealed that CRF139 02B in Japan, the UK and Nepal had the same origin. Furthermore, the origin of CRF139 02B could be in the UK. In Japan, CRF139 02B seemed to be circulating recently and domestically, rather than migrating independently from elsewhere to Japan. This suggestion did not conflict with the clinical information of the participants at (or nearby) the first visit; their CD4 counts were maintained, and 3 of 4 participants had an HIV negative history recently (Table 1). In the same way, MCC tree analysis was performed with only subtype B or CRF139 02B sequences in region VII (Figure 4C). Four sequences of CRF139 02B in Japan formed a cluster, and tMRCA was 2017.2 (2014.2-2019.7). Upstream,

CRF139_02B sequences in the UK were located and made clusters with CRF139_02B sequences in Japan, and tMRCA was 2001.7 (1992.3-2009.7). However, no close sequence was found further upstream. These results revealed that CRF139_02B in Japan and the UK had the same origin, though no information was found about the parental subtype B strain of CRF139_02B. That may be the reason why tMRCA of CRF139_02B in region VII had a longer 95% HPD interval than that in region VI.

Our study has 4 limitations to be mentioned. First, this study was conducted in a single center. We had no more information about current circulation of CRF139_02B in Japan. Second, the number of study participants was small. The clinical or virological feature of CRF139_02B was unclear. Third, no epidemiological link among the participants was observed. We could not confirm epidemiological migration or circulation of CRF139_02B in Japan. Fourth, the putative parental subtype B of CRF139_02B was unknown, which weakens our hypothesis that CRF139_02B originated in the UK.

In conclusion, we newly identified HIV-1 CRF139_02B from 4 PLWH in Japan. It turned out that CRF139_02B has existed in at least 8 PLWH outside of Japan. Furthermore, the origin of CRF139_02B could be in the UK. Because there is a possibility that further international circulation of CRF139_02B may be observed in the near future, continuous monitoring of HIV-1 molecular epidemiology will be needed.

Acknowledgements

We thank the study participants who generously donated blood for our study. We also thank all staff of AIDS Clinical Center, National Center for Global Health and Medicine because this study was based on their faithful work.

Funding: This study was financially supported by Grantsin Aid for AIDS research from the Ministry of Health, Labor and Welfare (Research for the development of medical systems for HIV infection (23HB2001)).

Conflict of Interest: SO and HG have received research grants from Gilead Sciences, MSD K.K., and ViiV Healthcare, and honorarium for lectures from ViiV Healthcare and Gilead Sciences. The other authors declare no conflicts of interest.

References

- UNAIDS. The path that ends AIDS: UNAIDS Global AIDS Update 2023. https://www.unaids.org/sites/default/ files/media_asset/2023-unaids-global-aids-update_en.pdf (accessed June 24, 2024).
- UNAIDS. Fast-Track Ending the AIDS epidemic by 2030. https://www.unaids.org/sites/default/files/media_ asset/JC2686_WAD2014report_en.pdf (accessed June 24, 2024).
- WHO. HIV statistics, globally and by WHO region, 2023. https://iris.who.int/bitstream/handle/10665/376793/WHO-UCN-HHS-SIA-2023.01-eng.pdf (accessed June 24, 2024).
- Bbosa N, Kaleebu P, Ssemwanga D. HIV subtype diversity worldwide. Curr Opin HIV AIDS. 2019; 14:153-160.
- Yamaguchi J, Vallari A, McArthur C, Sthreshley L, Cloherty GA, Berg MG, Rodgers MA. Brief report: Complete genome sequence of CG-0018a-01 establishes HIV-1 subtype L. J Acquir Immune Defic Syndr. 2020; 83:319-322.
- Ramirez BC, Simon-Loriere E, Galetto R, Negroni M. Implications of recombination for HIV diversity. Virus Res. 2008; 134:64-73.
- Song H, Giorgi EE, Ganusov VV, *et al.* Tracking HIV-1 recombination to resolve its contribution to HIV-1 evolution in natural infection. Nat Commun. 2018; 9:1928.
- Hemelaar J, Elangovan R, Yun J, Dickson-Tetteh L, Fleminger I, Kirtley S, Williams B, Gouws-Williams E, Ghys PD; WHO–UNAIDS Network for HIV Isolation Characterisation. Global and regional molecular epidemiology of HIV-1, 1990–2015: A systematic review, global survey, and trend analysis. Lancet Infect Dis. 2019; 19:143-155.
- Switzer WM, Shankar A, Jia H, *et al.* High HIV diversity, recombination, and superinfection revealed in a large outbreak among persons who inject drugs in Kentucky and Ohio, USA. Virus Evol. 2024; 10:veae015.

- Ministry of Health, Labour And Welfare. Annual report for HIV/AIDS in Japan *https://api-net.jfap.or.jp/status/ japan/nenpo.html* (accessed June 24, 2024). (in Japanese)
- Iwamoto A, Taira R, Yokomaku Y, Koibuchi T, Rahman M, Izumi Y, Tadokoro K. The HIV care cascade: Japanese perspectives. PLoS One. 2017; 12:e0174360.
- Matsuoka S, Nagashima M, Sadamasu K, Mori H, Kawahata T, Zaitsu S, Nakamura A, de Souza MS, Matano T. Estimating HIV-1 incidence in Japan from the proportion of recent infections. Prev Med Rep. 2019; 16:100994.
- Frescura L, Godfrey-Faussett P, Feizzadeh A A, El-Sadr W, Syarif O, Ghys PD; on and behalf of the 2025 testing treatment target Working Group. Achieving the 95 95 95 targets for all: A pathway to ending AIDS. PLoS One. 2022; 17:e0272405.
- Otani M, Shiino T, Hachiya A, et al. Association of demographics, HCV co-infection, HIV-1 subtypes and genetic clustering with late HIV diagnosis: A retrospective analysis from the Japanese Drug Resistance HIV-1 Surveillance Network. J Int AIDS Soc. 2023; 26:e26086.
- 15. Hosaka M, Fujisaki S, Masakane A, Hattori J, Shiino T, Gatanaga H, Shigemi U, Okazaki R, Hachiya A, Matsuda M, Ibe S, Iwatani Y, Yokomaku Y, Sugiura W; Japanese Drug Resistance HIV-1 Surveillance Network Team. HIV-1 CRF01_AE and Subtype B Transmission Networks Crossover: A New AE/B Recombinant Identified in Japan. AIDS Res Hum Retroviruses. 2016; 32:412-419.
- 16. Ogawa S, Hachiya A, Hosaka M, Matsuda M, Ode H, Shigemi U, Okazaki R, Sadamasu K, Nagashima M, Toyokawa T, Tateyama M, Tanaka Y, Sugiura W, Yokomaku Y, Iwatani Y. A novel drug-resistant HIV-1 circulating recombinant form CRF76_01B identified by near full-length genome analysis. AIDS Res Hum Retroviruses. 2016; 32:284-289.
- 17. Tran GV, Hayashida T, Dang AL-D, Nagai M, Matsumoto S, Tran LK, Le HN-M, Van TD, Tanuma J, Pham TN, Oka S. Prevalence of transmitted drug resistance and phylogenetic analysis of HIV-1 among antiretroviral therapy-naïve patients in Northern Vietnam from 2019 to 2022. Glob Health Med. 2024; 6:117-123.
- Hayashida T, Tran LK, Dang AL-D, Nagai M, Matsumoto S, Le HN-M, Van TD, Tran GV, Tanuma J, Pham TN, Oka S. Identification of new circulating recombinant form of HIV-1 CRF127_07109 in Northern Vietnam. AIDS Res Hum Retroviruses. 2024; doi: 10.1089/AID.2024.0022. Epub ahead of print.
- Ode H, Matsuda M, Matsuoka K, Hachiya A, Hattori J, Kito Y, Yokomaku Y, Iwatani Y, Sugiura W. Quasispecies analyses of the HIV-1 near-full-length genome with illumina MiSeq. Front Microbiol. 2015; 6:1258.
- Mori M, Ode H, Kubota M, Nakata Y, Kasahara T, Shigemi U, Okazaki R, Matsuda M, Matsuoka K, Sugimoto A, Hachiya A, Imahashi M, Yokomaku Y, Iwatani Y. Nanopore sequencing for characterization of HIV-1 recombinant forms. Microbiol Spectr. 2022; 10:e0150722.
- Yang X, Charlebois P, Gnerre S, Coole MG, Lennon NJ, Levin JZ, Qu J, Ryan EM, Zody MC, Henn MR. De novo assembly of highly diverse viral populations. BMC Genomics. 2012; 13:475.
- 22. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009; 25:1754-1760.
- 23. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman

M, Lander ES, Getz G, Mesirov JP. Integrative genomics viewer. Nat Biotechnol. 2011; 29:24-26.

- Schultz AK, Bulla I, Abdou-Chekaraou M, Gordien E, Morgenstern B, Zoulim F, Deny P, Stanke M. jpHMM: Recombination analysis in viruses with circular genomes such as the hepatitis B virus. Nucleic Acids Res. 2012; 40:W193-W198.
- 25. Siepel AC, Halpern AL, Macken C, Korber BT. A computer program designed to screen rapidly for HIV type 1 intersubtype recombinant sequences. AIDS Res Hum Retroviruses. 1995; 11:1413-1416.
- Kumar S, Stecher G, Tamura K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol. 2016; 33:1870-1874.
- Bouckaert R, Heled J, Kuhnert D, Vaughan T, Wu CH, Xie D, Suchard MA, Rambaut A, Drummond AJ. BEAST
 A software platform for Bayesian evolutionary analysis. PLoS Comput Biol. 2014; 10:e1003537.
- Mbisa JL, Ledesma J, Kirwan P, Bibby DF, Manso C, Skingsley A, Murphy G, Brown A, Dunn DT, Delpech V, Geretti AM. Surveillance of HIV-1 transmitted integrase strand transfer inhibitor resistance in the UK. J Antimicrob Chemother. 2020; 75:3311-3318.

- Yebra G, Frampton D, Gallo Cassarino T, *et al.* A high HIV-1 strain variability in London, UK, revealed by fullgenome analysis: Results from the ICONIC project. PLoS One. 2018; 13:e0192081.
- 30. Deuba K, Panta G, Rajbhandari RM, Kunwar R, Pokhrel TN, Pandey LR, Changsom D, Saeng-aroon S, Thakur SK. Prevalence of viral load suppression and acquired drug resistance among people living with HIV in Nepal: a nationally representative surveillance study. J Glob Antimicrob Resist. 2023; 35:122-127.

Received June 27, 2024; Revised August 16, 2024; Accepted September 11, 2024.

Released online in J-STAGE as advance publication September 13, 2024.

*Address correspondence to:

Hiroyuki Gatanaga, AIDS Clinical Center, National Center for Global Health and Medicine, 1-21-1 Toyama, Shinjuku-ku, Tokyo 162-8655, Japan.

E-mail: higatana@acc.ncgm.go.jp