

Fig. 1. Clinical course of patient-treatment protocols. Open squares and solid circles indicate plasma viral load and CD4⁺ cell count, respectively. Each treatment regimen is indicated in the upper part of the graph by horizontal dashed lines between arrows. Sample collection points are labeled A to H, and numbers of sequences analyzed at each point are shown in parentheses. AZT, zidovudine; d4T, stavudine; ddl, didanosine; 3TC, lamivudine; NVP, nevirapine; NFV, nelfinavir; RTV, ritonavir; SQV, saquinavir; IDV, indinavir.

viral RNA from patient plasma must be reverse-transcribed to cDNA, and the target regions are then amplified by PCR. Since viral cDNA includes multiple viral populations, using such samples as a template for the PCR step followed by sequencing results in amplification of predominantly the major variants (Gunthard et al., 1998; Hance et al., 2001), which may not represent the diversity of the original population. Furthermore, there is always a risk of artificial recombination of viral cDNAs, which are estimated to occur at rates of 4–70% during the PCR step (Meyerhans et al., 1990) and may disturb the linkage of mutation sites. Thus, this artificial linkage information may significantly affect the results of co-evolution analyses. To circumvent this technical problem, we employed the single-genome sequencing (SGS) technique based on limiting-dilution assays (Palmer et al., 2005). With SGS, artificial recombination can be avoided during PCR because cDNA samples are diluted after the reverse transcriptase reaction to a single cDNA molecule, which is used as the PCR template. Therefore, the viral sequence information obtained by SGS is not only more sensitive for analyzing HIV population diversity and mutation linkages than that obtained by standard genotype analysis, but also more precise for identifying co-evolving sites.

Second, many analytical methods have been developed for detecting co-evolving sites in molecular sequences. However, early methods did not accommodate the evolutionary history among sequences, and risked generating false-positive predictions (Altschuh et al., 1987; Gutell et al., 1992; Martin et al., 2005; Neher, 1994; Pollock and Taylor, 1997; Tillier and Lui, 2003). In recent years, the accuracy of estimating co-evolving sites has greatly improved due to several computational algorithms that incorporate the phylogenetic relationships among molecular data (Bhattacharya et al., 2007; Dutheil et al., 2005; Poon et al., 2008; Tuff and Darlu, 2000; Wollenberg and Atchley, 2000; Yeang and Haussler, 2007). In particular, Spidermonkey analysis (Poon et al., 2007a,b, 2008) not only considered the phylogenetic relationships but also implemented two empirical HIV-1 subtype B amino acid-substitution models for describing between- and within-host HIV evolution (Nickle et al., 2007).

In this study, we clarified the impact of Gag and PR co-evolution in the acquisition of drug resistance by inferring co-evolving sites

between Gag and PR in a dataset of HIV-1 Gag-PR linkage sequences from a single patient who had undergone highly active antiretroviral therapy (HAART) for a long period. We used the SGS method for sequencing viral samples to avoid artificial recombination, and applied Spidermonkey analysis to our data by using its option for the within-host HIV substitution model. Virological significance of the estimated co-evolving sites was analyzed *in vitro* by constructing recombinant viruses on a pNL4-3 backbone. The replication kinetics, susceptibility to anti-HIV drugs, and the Gag processing efficiency were evaluated for each clone. We also conducted homology modeling analysis to examine Gag-PR interactions and database analysis to confirm the universality of the co-evolving sites.

2. Materials and methods

2.1. Sample collection and gag-PR-coding region sequencing by single-genome sequencing

From all HIV/AIDS patients monitored from April 1998 to August 2002 at the National Institute of Infectious Diseases in Japan, we selected a virological failure case with a history of multiple antiretroviral treatments. Of all cases, this one had been followed for the longest period and had enough detailed clinical information to perform our analysis. For the selected case, we collected plasma samples and clinical information, such as treatment regimens, changes in viral load, and CD4 counts (Fig. 1).

The single-genome sequencing (SGS) method was used as described (Palmer et al., 2005). Briefly, HIV-1 RNA was extracted from plasma samples (containing a minimum of 1000 copies of HIV-1 RNA) by guanidinium isothiocyanate treatment. cDNA was synthesized using Superscript II RT kit (Invitrogen, Carlsbad, CA) and a random hexamer. cDNA was serially diluted and amplified by PCR and nested PCR using Platinum Taq DNA Polymerase High Fidelity (Invitrogen). The endpoint of reverse-transcribed cDNA was determined as a single clone by the Poisson distribution, with cDNA dilutions yielding PCR products in 3 out of 10 reactions. The following primer sets were used in the first and nested PCR amplifications: first PCR-WGPF1 (5'-CTCTCTCGACGACGACTCG-

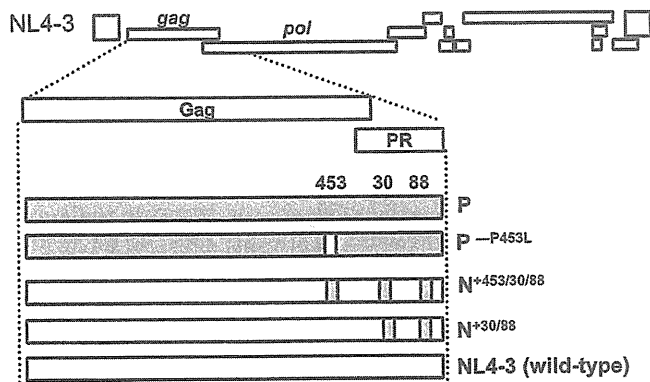


Fig. 2. Construction of recombinant viruses. Four different types of recombinant viruses were constructed according to their mutation patterns: (1) P, NL4-3 backbone with patient-derived Gag and protease (PR); (2) P⁻⁴⁵³, NL4-3 backbone with patient-derived Gag and PR but P453L^{Gag} was converted to wild-type; (3) N^{+453/30/88}, NL4-3 backbone including the substitutions P453L^{Gag}, D30N^{PR}, and N88D^{PR}; (4) N^{+30/88}, NL4-3 backbone including the substitutions D30N^{PR} and N88D^{PR}.

3') and 3500- (5'-CTATTAAGTATTTTATGGGTCATAA-3'), nested PCR-WGPF2 (5'-TTGCTGAAGCGCGCACGGCAAGA-3') and 3410- (5'-CAGTTAGTGGTATTACTTCTGTAGTGCTT-3'). These primer sets amplified a 2.7 kbps fragment containing *gag*, the PR-coding region of *pol*, and the first 900 nucleotides of the reverse transcriptase-coding region of *pol*. Positive PCR products were determined by agarose gel electrophoresis and sequenced using ABI Prism BigDye Terminator version 3.1 dideoxyterminator cycle sequencing (Applied Biosystems, Carlsbad, CA). Sequences, including the entire *gag* gene (1500 bp) and PR-coding region (297 bp), were aligned on the HXB2 reference using Sequence Navigator software (Applied Biosystems). Sequences containing mixtures at any position were excluded from analysis. PI-resistant mutations and other PR mutations in our data set were determined using the Calibrated Population Resistance tool Version 4.3 beta (<http://cpr.stanford.edu/cpr/>).

2.2. Detecting co-evolution using Spidermonkey analysis

Co-evolving sites between Gag and PR were inferred using Spidermonkey analysis (Poon et al., 2007a,b, 2008) to analyze all sequences (data type = protein), including the entire Gag (500 amino acids) and PR (99 amino acids) sequences. A phylogenetic tree was first reconstructed using the neighbor-joining method (Saitou and Nei, 1987). To infer co-evolving sites, we selected the within-host HIV protein-substitution model (frequencies: model-defined), based on the maximum likelihood method (Nickle et al., 2007). Amino acid sequences at MA/CA, CA/p2, p2/NC, NC/p1, and p1/p6 Gag cleavage-sites (residues P5 to P5') and the complete PR sequence were selected for BGM analysis. We used the default options for calculations and inferred co-evolving sites if the estimated posterior probability for a pair of positions exceeded the cutoff value of 0.5.

2.3. Construction of recombinant HIV with a patient-derived gag-PR-coding region

To confirm the contributions of P453L^{Gag} in the context of patient-derived *gag*-PR-coding region, three types of recombinant viruses were constructed (Fig. 2). Recombinant viruses were constructed using the pNL4-3 molecular clone of HIV-1 (GenBank accession No. AF324493). A patient-derived *gag*-PR-coding region containing P453L^{Gag}, D30N^{PR}, and N88D^{PR} was chosen from sampling point F. The recombinants were (1) P, NL4-3 backbone with patient-derived *gag*-PR-coding region; (2) P^{-P453L}, NL4-3 backbone

with patient-derived *gag*-PR-coding region but P453L^{Gag} was converted to wild-type.

Details of the recombinant virus construction were as follows. The patient-derived 1.9 kbp *gag*-PR-coding region was subcloned into pCR4-TOPO (Invitrogen) (designated pCR4-TOPO^{Patient}). Using pCR4-TOPO^{Patient} as a template, P453L^{Gag} was substituted with wild-type P453. Each substitution was introduced using the following primer sets: For the Gag-P453 substitution, forward: 5'-TTCAGAACAGACCAGAGCCATCAGCT-3', reverse: 5'-AGCTGATGGCTCTGGTCTGTTCTGAA-3'. Subsequently, the substituted *gag*-PR-coding region was amplified using the following primers: WGPF2: 5'-TTGCTGAAGCGCGCACGGCAAGA-3' and DRPRO6: 5'-ACTTTTGGCCATCCATTCCTGGCTTT-3'. The NL4-3-derived RT-coding region was amplified using the following primers: RT-63F: 5'-TAAACAATGGCCATTGACAGAAG-3' and RT-898R: 5'-CTGCTTCTTCTGTTAGTGGTACTAC-3'. Primers DRPRO6 and RT-63F are phosphorylated at their 5' ends. Both resulting fragments were ligated and digested with BssHII and SbfI, and cloned back into pNL4-3. All pNL4-3-based recombinant DNAs (3.75 µg) were transfected into 2 × 10⁵ HeLa cells using Fugene6 (Roche, Indianapolis, IN), and culture supernatants were harvested at 72 h after transfection, filtered through a 0.45 µm membrane, assayed for reverse transcriptase (RT) activity (Willey et al., 1988), and kept as virus stocks at -80 °C until use. Each virus stock (5 × 10⁶ ³²P cpm of RT activity) was used for replication kinetics analyses.

2.4. Construction of pNL43 with P453L^{Gag}, D30N^{PR}, and N88D^{PR} by site-directed mutagenesis

In addition to constructing patient-derived *gag*-PR-coding region recombinant viruses, we evaluated the effect of P453L^{Gag}, D30N^{PR}, and N88D^{PR} interference by constructing three types of pNL4-3-based recombinant viruses (Fig. 2). These were (1) N^{+453/30/88}, NL4-3 including P453L^{Gag}/D30N^{PR}/N88D^{PR}; and (2) N^{+30/88}, NL4-3 including D30N^{PR}/N88D^{PR}. These recombinant viruses were constructed as follows. pCR4-TOPO (Invitrogen) including the Apal-SbfI fragment of pNL4-3, which contained the complete p1/p6 of *gag* and PR-coding region, was constructed as a template for further mutagenesis. This construct was designated as pCR4-TOPO/NL^{Apal-SbfI}. Using pCR4-TOPO/NL^{Apal-SbfI} as a template, we introduced three mutations (P453L^{Gag}, D30N^{PR}, and N88D^{PR}) by site-directed mutagenesis. The following primer sets were used to introduce each mutation: For P453L^{Gag}, forward: 5'-TTCAGAGCAGACTAGAGCCAACAGCC-3' and reverse: 5'-GGCTGTTGGCTCTAGTCTGCTCTGAA-3'. For D30N^{PR}, forward: 5'-CAGGAGCAGATAATACAGTATTAGAAG-3' and reverse: 5'-CTTCTAATACTGTATTATCTGCTCCTG-3'. For N88D^{PR}, forward: 5'-CATAATTGGAAGAGATCTGTTGACTC-3' and reverse: 5'-GAGTCAACAGATCTTCCAATTATG-3'. After each mutagenesis reaction, the entire sequences were verified and cloned back into pNL4-3 using Apal and SbfI restriction enzymes. All pNL4-3-based recombinant DNAs (3.75 µg) were transfected into 2 × 10⁵ HeLa cells using Fugene6 (Roche), and culture supernatants were harvested at 72 h after transfection, filtered through a 0.45 µm membrane, assayed for RT activity (Willey et al., 1988), and kept as virus stocks at -80 °C until use. Each virus stock (4 × 10⁵ ³²P cpm of RT activity) was used for replication kinetics analysis.

2.5. Replication kinetics of recombinant viruses

Replication kinetics of recombinant viruses was evaluated as described (Matsuoka-Aizawa et al., 2003). Briefly, 5 × 10⁴ MT-2 cells were infected with each virus stock in the absence or presence of 0.1 µM nelfinavir at 37 °C for 16 h. Cells were then washed once and resuspended in 0.5 ml culture medium with the same concen-

tration of nelfinavir, and cultures were maintained for 12–22 days, changing half of the medium every 2 or 3 days. The titer of each virus was evaluated by both RT activity and p24 amount; as both measures demonstrated good correlation, p24 amount was used to adjust the virus inoculum. Culture supernatants were collected, and residual supernatants were kept at -80°C until use. Replication kinetics was independently analyzed two times.

2.6. Evaluation of nelfinavir susceptibilities of recombinant viruses

Recombinant viruses were evaluated for nelfinavir susceptibility using an in-house drug susceptibility assay with the MaRBLE cell line (MaRBLE assay) as described (Chiba-Mizutani et al., 2007). Briefly, 1×10^5 MaRBLE cells were infected with 100 CCID₅₀ of each recombinant virus, and virus replication was monitored in serial dilutions of nelfinavir from 1.28×10^{-13} to 1.0×10^{-6} M for 7 days in triplicate. At day 7, cells were harvested and lysed in luciferase assay reagent. Firefly (FF) and renilla luciferase (RF) activities produced by the cells were quantified using a Dual-luciferase Reporter Assay System (Promega, Madison, WI). The relative virus replication rate (% replication) at each drug concentration was calculated by the following formula: % replication = (observed FF luciferase activity with drug – background [mock] FF luciferase activity) / (observed FF luciferase activity without drug – background [mock] FF luciferase activity) \times 100. IC₅₀ values were calculated with 95% confidence intervals using GraphPad Prism software and nonlinear regression analysis fitted with a sigmoidal dose–response curve with variable slope.

2.7. Analysis of Gag processing patterns in recombinant viruses

The Gag processing patterns of recombinant viruses were analyzed as described (Sugiura et al., 2002) with minor modifications. In brief, NL4-3-based recombinant DNAs (18 μg) was transfected into 5×10^6 HeLa cells using Fugene6 (Roche). The culture supernatants were harvested at 48 h after transfection with a culture medium change at 12 h post transfection to the absence or presence of 0.1 μM nelfinavir. For pelleting virus through a sucrose cushion, 25 ml of cell culture medium was layered onto 10 ml of 20% sucrose (wt/vol, in PBS) before centrifugation at 30,000 rpm in a swing rotor for 1.5 h. The medium and cushion were discarded, and the virus pellet was dissolved in 200 μl Laemmli sample buffer (Bio-Rad, Hercules, CA). Viral supernatants were normalized using PETRO-TEK HIV-1 p24 Antigen ELISA (ZeptoMetrix Corporation, Buffalo, NY) and subjected to sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE). Proteins in the gels were passively transferred to PVDF membranes (Bio-Rad). The membranes were incubated with a chicken anti-p6 polyclonal antibody (Sigma–Aldrich Corporation, St. Louis, MO) and a murine anti-p24 monoclonal antibody (ZeptoMetrix Corporation) for 2 h, followed by incubation with secondary antibody, a chicken HRP-conjugated, IgG antibody (Bethyl Laboratories Inc., Montgomery, TX), a mouse HRP-conjugated, IgG antibody (Thermo Fisher Scientific, Yokohama, JP), respectively. Finally, proteins were visualized using SuperSignal West Dura Extended Duration Substrate (Thermo Fisher Scientific).

2.8. Molecular modeling of Gag p1/p6 peptide–PR complexes

We constructed three-dimensional models of PR in complex with peptide representing Gag-p1/p6 substrate by a homology modeling method (Baker and Sali, 2001; Marti-Renom et al., 2000; Shirakawa et al., 2008) using Molecular Operating Environment (MOE) ver. 2008.10 (<http://www.chemcomp.com/>, Chemical Computing Group Inc., Montreal, Quebec, Canada). For the modeling

template, we used an X-ray crystal structure of inactive D25N PR in complex with the p1/p6 substrate at 2 Å resolution (PDB code: 1KJF) (Prabu-Jeyabalan et al., 2002), as it had the highest similarity (94.2% identity) to the HIV-1 NL4-3 strain among HIV-1 protease-p1/p6 peptide structures in the protein data bank, even though this is an inactive model. Furthermore, the active-site D25N^{PR} mutation has been reported to hardly influence the structure of the protease in complex with ligands (Sayer et al., 2008). In the models, nine-amino-acid-length peptides corresponding to the Gag p1/p6 of the NL4-3, N^{+30/88}, and N^{+453/30/88} strains were bound to the catalytic sites of PRs of the same strains. We considered effects of a water molecule (HOH11) that mediates important hydrogen bonds between the Gag p1/p6 peptide and the PR. AMBER ff99 force field (Wang et al., 2000) and generalized Born/volume integral (GB/VI) implicit solvent model (Labute, 2008) were applied for intra- and inter-molecular energy calculations.

2.9. Database analysis

To confirm the universality of P453L^{Gag}/D30N^{PR}/N88D^{PR}, we obtained 3249 sequences of HIV-1 subtype B gag-PR-coding region (positions 2146–2516) from the Los Alamos National Laboratory HIV sequence database (<http://www.hiv.lanl.gov/>). For the dataset, we applied Fisher's exact test and investigated associations between the D30N/N88D mutations in PR and the P453L mutation in Gag and between the 30/88 mutations in PR and the P453L mutation in Gag, respectively. Fisher's exact test was implemented using the R 2.6.1 statistical package.

3. Results

3.1. The majority of patient-derived Gag cleavage-site mutations are in p2/NC and p1/p6

Twenty-three plasma samples were serially collected from a patient who had received HAART for 52 months. The patient's clinical history, changes in viral load, and CD4 counts are depicted in Fig. 1. At 8 sampling points (A–H), 129 gag-PR-coding region sequences were obtained, with p2/NC and p1/p6 Gag cleavage-site mutations observed at each point in more than 60% of clones (Table 1). PI-resistant mutations and other PR mutations are also summarized in Table 1. None of the clones had CA/p2 cleavage-site mutations, and only a few clones had MA/CA or NC/p1 cleavage-site mutations. The Y132F^{Gag} mutation in MA/CA was found in 1.6% of the clones, and E428D^{Gag} and R429K^{Gag} within NC/p1 had prevalences of 1.6% and 5.4%, respectively. On the other hand, cleavage-site mutations were frequently observed within p2/NC and p1/p6. In the p2/NC site, we observed 11 mutations: S373Q^{Gag} (83.7%), S373P^{Gag} (15.5%), A374G^{Gag} (79.8%), A374V^{Gag} (16.3%), A374S^{Gag} (2.3%), A374R^{Gag} (0.8%), T375N^{Gag} (100%), I376V^{Gag} (3.9%), M378L^{Gag} (0.8%), G381S^{Gag} (0.8%), and N382Y^{Gag} (0.8%). In the p1/p6 site, we observed three mutations: L449F^{Gag} (0.8%), S451N^{Gag} (100%), and P453L^{Gag} (65.1%).

3.2. Ten Gag–protease co-evolving sites are inferred by Spidermonkey analysis

Among the 129 gag-PR-coding region sequences analyzed by Spidermonkey analysis (Poon et al., 2007a,b, 2008), ten co-evolving sites were inferred. These sites were identified using the default cutoff posterior probability (pp) value of 0.5. The ten co-evolving pairs identified with pp > 0.5 are shown in Table 2. Four pairs, R429K^{Gag}/M36V^{PR} (pp = 0.87), S373Q^{Gag}/T12A^{PR} (pp = 0.83), P453L^{Gag}/D30N^{PR} (pp = 0.63), and P453L^{Gag}/N88D^{PR} (pp = 0.61), represented Gag/PR inter-molecular co-evolution;

Table 1
Gag cleavage-site mutations and PR mutations from a HAART-treated case.

Sampling point	Gag cleavage-site mutation (%) ^a		PR mutation (%) ^a	
	p2/NC	p1/p6	PI-resistant mutations	Other PR mutations
A to H (n = 129)	T375N	S451N	–	I62V, L63P, A71T, I93L
A (n = 19)	S373Q(100), A374G (89),	–	–	M36I(95), I72V(11), V77I(100)
B (n = 18)	S373Q(100), A374G (94)	P453L(100)	D30N(100), M46I(6), N88D(100)	E35D(94), M36I(100), V77I(100),
C (n = 15)	S373Q(100), A374G(100)	P453L(100)	D30N(100), N88D(100)	L10F(80), E35D(100), M36I(100), K45R(27),
D (n = 13)	S373P (85), A374V (85)	P453L (8)	D30N (8), N88D (8)	I72T(7), V77I (100),
E (n = 14)	S373P (64), A374V(71)	–	–	L10F(8), V11I (8), T12A(8), K20R(8), E35D(8),
F (n = 14)	S373Q(100), A374G(100)	P453L(100)	D30N(100), N88D(100)	M36I(100), H69Y(8), V77I(8)
G (n = 19)	S373Q(100), A374G(100)	P453L(100)	D30N(100), N88D(100)	T12A(7), M36I(100), K55N (7), V77I(7)
H (n = 17)	S373Q(100), A374G(100)	P453L(100)	D30N(100), I54V(76), N88D(100),	L10F(100), I13V(79), E34G(7), E35D(100), M36I
			L90M (47)	(100), N37T (86), K45R (14), Q58E (86), I72T
				(7), V77I (100)
				L10F(100), I13V (68), E35D(100), M36I(68),
				M36V(32), N37T(100), Q58E(100), V77I(100)
				L10F(100), I13V (94), K20R (76), E35D (100),
				M36I (94), M36V (6), N37T (100), Q58E (100),
				I72T (53), V77I (100), G78R (6)

Gag mutations refer to HXB2 and PI-resistant mutations and other PR mutations were determined using the Calibrated Population Resistance tool Version 4.3 beta.

^a Numbers in parentheses are the percentages of mutations at each sampling point.

one pair, S373Q^{Gag}/A374G^{Gag}, represented Gag/Gag intra-molecular co-evolution; and the other five pairs, N37T^{PR}/Q58E^{PR} (pp = 0.94), E35D^{PR}/M46I^{PR} (pp = 0.89), K20R^{PR}/I54V^{PR} (pp = 0.88), V11I^{PR}/K20R^{PR} (pp = 0.86), and D30N^{PR}/N88D^{PR} (pp = 0.62), represented PR/PR intra-molecular co-evolution.

We focused on three pairs, P453L^{Gag}/N88D^{PR}, P453L^{Gag}/D30N^{PR}, and D30N^{PR}/N88D^{PR}, because D30N^{PR} and N88D^{PR} are well-known major and minor nelfinavir-resistant mutations, respectively (Johnson et al., 2008), and P453L^{Gag} is the P5' position of the p1/p6 cleavage-site mutation. Although D30N^{PR} has been associated with N88D^{PR} (Rhee et al., 2007; Wu et al., 2003), the interactions among P453L^{Gag}, D30N^{PR}, and N88D^{PR} have not been investigated. Since P453L^{Gag}/D30N^{PR}/N88D^{PR} was frequently observed in the presence of nelfinavir (Fig. 1 and Table 1), we conducted *in vitro* experiments to confirm whether the co-existence of P453L^{Gag}/D30N^{PR}/N88D^{PR} has a virological advantage in the presence of nelfinavir.

3.3. P453L^{Gag} improves the replication capacity of viruses with D30N^{PR}/N88D^{PR} in both patient- and NL4-3-derived genetic backgrounds

To evaluate the virological impact of P453L^{Gag} in the patient-derived genetic background, we constructed two types of patient-derived gag-PR-coding region viruses, P and P^{-P453L} (Fig. 2). These two recombinant viruses and the wild-type virus (NL4-3) were cultured independently in the absence or presence of nelfinavir, and their replication kinetics was monitored by measuring RT activity in culture supernatants. Assays for replication kinetics were independently performed twice, confirming identical orders of replication kinetics.

Table 2
Positions of coevolving pairs inferred by Spidermonkey analysis.

Position	Position	Expected posterior probability	Total number of sequences
429 ^{Gag}	36 ^{PR}	0.866	7
373 ^{Gag}	12 ^{PR}	0.825	2
453 ^{Gag}	30 ^{PR}	0.632	84
453 ^{Gag}	88 ^{PR}	0.607	84
373 ^{Gag}	374 ^{Gag}	0.895	22
37 ^{PR}	58 ^{PR}	0.938	48
35 ^{PR}	46 ^{PR}	0.887	1
20 ^{PR}	54 ^{PR}	0.877	13
11 ^{PR}	20 ^{PR}	0.858	1
30 ^{PR}	88 ^{PR}	0.618	84

Significant coevolving sites (pp value > 0.5) are shown.

In the absence of nelfinavir (Fig. 3A), the RT activities of the NL4-3 and P viruses peaked at 6 days after infection, and the RT activity of NL4-3 was higher than that of the P virus, whereas viral replication was delayed in the P^{-P453L} virus (i.e., P virus without P453L^{Gag}), and its RT activity peaked at 10 days after infection. The order of replication kinetics in the absence of nelfinavir was wild-type (NL4-3) > P > P^{-P453L}. On the other hand, in the presence of nelfinavir (Fig. 3B), replication of the wild-type virus was completely suppressed, and replication of P virus was the most active, demonstrating peak RT activity at 6 days after infection. The P^{-P453L} virus showed lower replication capacity than the P virus. Thus, the order of replication kinetics in the presence of nelfinavir was P > P^{-P453L} > wild-type (NL4-3).

To assess whether the complex P453L^{Gag}/D30N^{PR}/N88D^{PR} conferred an advantage not only in the patient-derived genetic background but also in the HIV-1 molecular clone (NL4-3)-derived genetic background, we constructed two types of NL4-3-based gag-PR-coding region recombinant viruses, N^{+453/30/88} and N^{+30/88} (Fig. 2). The results of independent culture studies are shown in Fig. 3C and D. In the absence of nelfinavir, the RT activities of NL4-3, NL4-3 with P453L^{Gag}/D30N^{PR}/N88D^{PR} (N^{+453/30/88}) and NL4-3 with D30N^{PR}/N88D^{PR} (N^{+30/88}) viruses peaked at 11 days, 14 days, and 16 days after infection, respectively, and the order of replication kinetics was NL4-3 > N^{+453/30/88} > N^{+30/88}. On the other hand, in the presence of nelfinavir, the N^{+453/30/88} virus grew the most actively with its peak RT activity at 14 days after infection. The N^{+30/88} virus was the second most actively replicating, and the wild-type did not replicate. Thus, the order of replication kinetics in the presence of nelfinavir was N^{+453/30/88} > N^{+30/88} > NL4-3.

In both studies, not only in the patient-derived genetic background but also in the NL4-3-derived genetic background, the virus with D30N^{PR}/N88D^{PR} showed lower replication capacity than the virus with P453L^{Gag}/D30N^{PR}/N88D^{PR}, suggesting that P453L^{Gag} significantly contributes to the fitness recovery of virus with D30N^{PR}/N88D^{PR}.

3.4. P453L^{Gag} does not influence susceptibility to nelfinavir

To clarify whether P453L^{Gag} affects nelfinavir susceptibility, the IC₅₀s of nelfinavir for NL4-3-based gag-PR-coding region recombinants were determined using MARBLE cells. Both N^{+453/30/88} and N^{+30/88} recombinants showed 14.4 and 15.8-fold greater resistance to nelfinavir than wild-type virus, respectively (Table 3). However, the difference in IC₅₀ between the two recombinants was

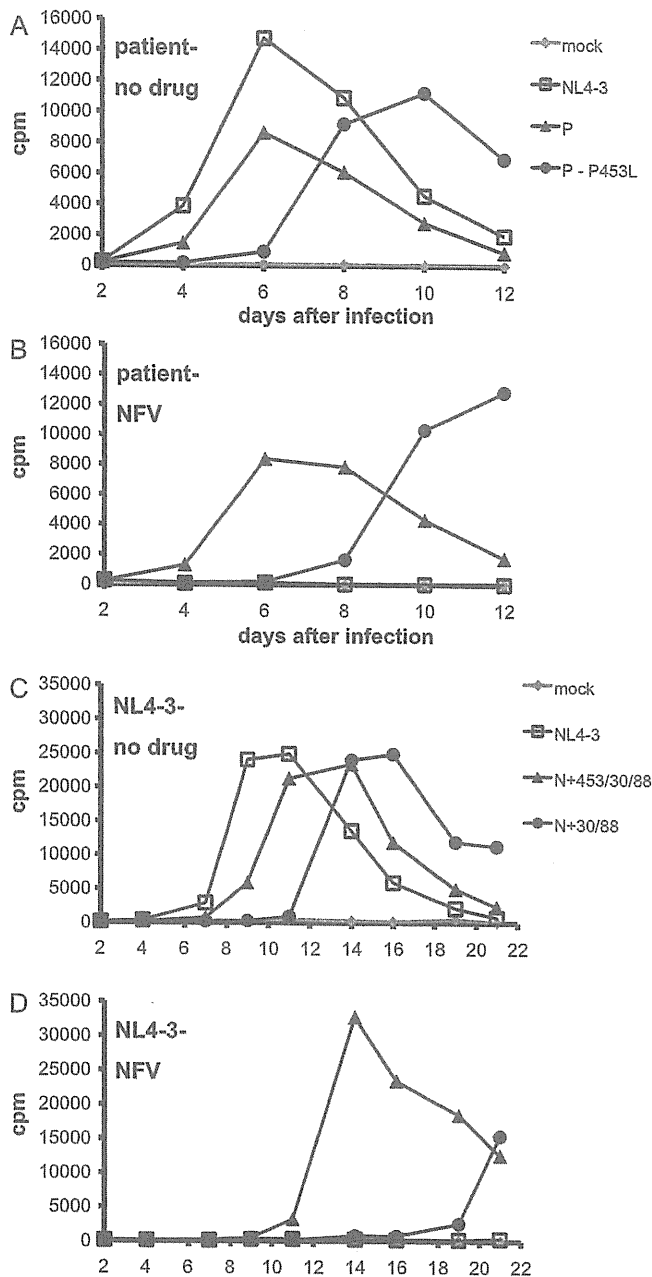


Fig. 3. Replication kinetics of recombinants. MT-2 cells were infected with patient-derived gag-PR-coding region, (A) in the absence of, and (B) in the presence of 0.1 μM nelfinavir. Open squares, solid triangles, and solid circles indicate wild-type NL4-3, NL4-3 with patient gag-PR-coding region insert, and NL4-3 with patient insert without P453L substitution, respectively. MT-2 cells were also infected with NL4-3-based recombinant, (C) in the absence of, and (D) in the presence of 0.1 μM nelfinavir. Open squares, solid triangles, and solid circles indicate wild type NL4-3; NL4-3 with P453L, D30N, and N88D; and NL4-3 with D30N and N88D, respectively. Diamonds indicate mock-infected controls. Assays were independently performed twice, and one representative set of results is shown.

not significant, suggesting that P453L^{Gag} had no effect on nelfinavir susceptibility.

3.5. P453L^{Gag} improves Gag p1/p6 processing in virus with D30N^{PR}/N88D^{PR}

To gain further insights into the virological effects of P453L^{Gag}, we examined Gag processing patterns in the absence and presence of 0.1 μM nelfinavir by Western blot analysis with an anti-p6 polyclonal antibody and an anti-p24 monoclonal antibody. The amount

Table 3
Nelfinavir susceptibilities of recombinant viruses.

HIV-1	IC ₅₀ (nM)	95% confidence interval	Fold-resistance
NL4-3	1.3	0.8–2.4	1.0
NL ^{+453/30/88}	18.7	9.2–37.8	14.4
N ^{+30/88}	20.6	12.6–33.6	15.8

of sample loaded in each lane was normalized by p24 antigen content (600 ng for each lane as determined by ELISA) (Fig. 4A). In the absence of nelfinavir, the partially cleaved Gag intermediate p15 (including NC, p1, and p6) of NL4-3 was efficiently cleaved to the p6 peptide (Fig. 4B, lane 3). In contrast, the processing of NL4-3 was less efficient in the presence of nelfinavir, as indicated by the accumulation of p15 (Fig. 4B, lane 4). N^{+30/88} showed defects in cleavage at the p1/p6 site, as demonstrated by the accumulation of p15 and p7 (p1/p6) (Fig. 4B, lanes 5 and 6). On the other hand, lesser p15 and p7 accumulated in N^{+453/30/88} than in N^{+30/88} (Fig. 4, lanes 7 and 8). Interestingly, a 8–9 kDa band, which is neither p6 nor p7, was observed in N^{+30/88} (Fig. 4B, lanes 5 and 6).

3.6. Impaired Gag–PR affinity in the N^{+30/88} strain is recovered by new L453^{Gag} interactions with M46^{PR} and F53^{PR}

To elucidate the structural impact of the mutations described above on interactions between Gag-p1/p6 substrate and PR, we generated three-dimensional models of the PR in complex with peptide representing Gag-p1/p6 substrate by homology modeling (Baker and Sali, 2001; Marti-Renom et al., 2000; Shirakawa et al., 2008) using Gag and PR sequences of the NL4-3, N^{+30/88}, and N^{+453/30/88} strains. Comparison of the thermodynamically optimized models showed obvious differences in interactions between side chains of PR and p1/p6 (Fig. 5). First, D30N^{PR} mutation resulted in fewer hydrophilic interactions between side chains of the 30th PR and 452nd p1/p6 residues; NL4-3 had two hydrogen bonds between the side chains of D30^{PR} and R452^{Gag} (Fig. 5A), while N^{+30/88} and N^{+453/30/88} had only a single hydrogen bond between the side chains of N30^{PR} and R452^{Gag} (Fig. 5B and C). Second, the P453L^{Gag} mutation in the p1/p6 substrate of the N^{+453/30/88} strain led to new hydrophobic interactions between side chains of the 46th PR and 453rd p1/p6 residues, as well as of the PR 53F and p1/p6 P453L^{Gag} residues (Fig. 5C).

To examine whether these changes in interactions influenced the binding affinity of the p1/p6 substrate to PR, we analyzed

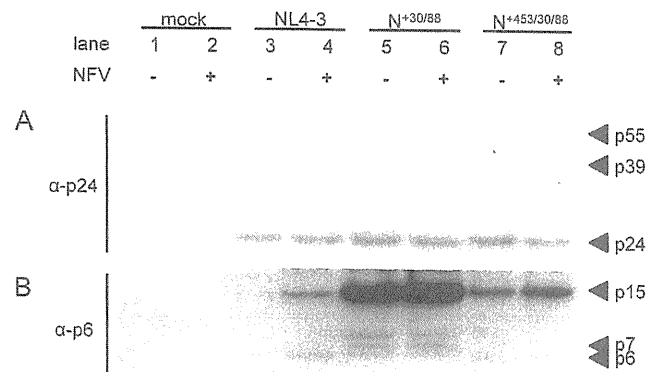


Fig. 4. Western blot analysis of Gag processing in the absence or presence of nelfinavir. Western blot analyses in the absence or presence of nelfinavir. HeLa cells were transfected by each recombinant clones and cultured in the absence or presence of NFV (0.1 μM). At 48 h post-transfection, virions in culture supernatants were harvested and subjected to Western blot analysis with anti-p24 monoclonal antibody (A) and an anti-p6 polyclonal antibody (B). Each lane was normalized by p24 antigen content (5 ng for each lane as determined by ELISA). Lanes 1 and 2, mock, lanes 3 and 4, NL4-3, lanes 5 and 6, N^{+30/88}, lanes 7 and 8, N^{+453/30/88}.

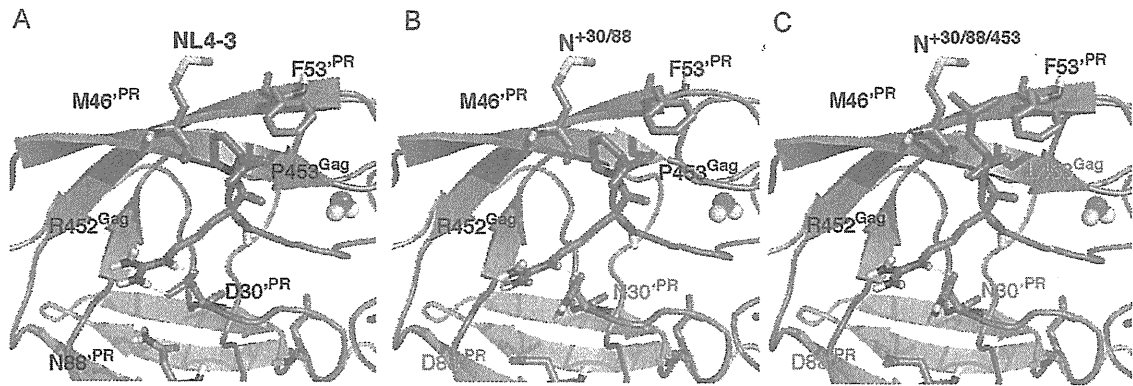


Fig. 5. Structural models of the Gag–PR complexes. Catalytic sites of the NL4-3 (A), N^{+30/88} models (B), and N^{+453/30/88} models (C) are highlighted. Green cartoons and sticks represent main and side chains of PR, respectively. Cyan sticks represent 452nd and 453rd residues in Gag corresponding to the p1/p6 region. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

their binding energies using the PR–p1/p6 peptide complex models. The predicted binding energies of the NL4-3, N^{+30/88} and N^{+453/30/88} models were -137.6 kcal/mol, -133.3 kcal/mol, and -137.1 kcal/mol, respectively, suggesting that p1/p6 substrate has lower affinity with N^{+30/88} PR than with NL4-3 and N^{+453/30/88} PRs. Taken together, these data suggest that PR mutations in the N^{+30/88} strain can reduce Gag–PR affinity primarily via loss of the hydrogen bond between N30^{PR} and R452^{Gag} and that the Gag p1/p6 mutation in the N^{+453/30/88} strain (P453L^{Gag}) can recover affinity by generating new hydrophobic interactions of L453^{Gag} with M46^{PR} and F53^{PR}.

3.7. P453L^{Gag}/D30N^{PR}/N88D^{PR} association is commonly observed in a large database

To confirm the prevalence of the P453L^{Gag}/D30N^{PR}/N88D^{PR} association in other HIV-infected individuals, we investigated 3249 sequences of HIV-1 subtype B gag-PR-coding region from the Los Alamos National Laboratory HIV sequence database (<http://www.hiv.lanl.gov/>). We found that the P453L^{Gag} mutation was significantly associated with D30N^{PR}/N88D^{PR} (Table 4; $p < 0.001$, Fisher's exact test). These data support the virological advantage of the P453L^{Gag}/D30N^{PR}/N88D^{PR} association.

4. Discussion

In this study, we analyzed mechanisms of anti-HIV drug-resistant mutation acquisition by investigating crosstalk between Gag and PR mutations. We traced the clinical course and sequence changes in gag and the PR-coding region of a virological-failure case heavily treated with multiple regimens, including different protease inhibitors. To focus on the quality of sequence data and accuracy of analysis, we used SGS and Spidermonkey analysis, respectively.

Among the ten co-evolving Gag–Protease pairs inferred by Spidermonkey analysis, we confirmed a linkage between P453L^{Gag} and D30N^{PR}/N88D^{PR}. D30N^{PR} has been reported to associate with N88D^{PR} (Rhee et al., 2007; Wu et al., 2003), P453L (Verheyen

et al., 2006), and positively correlate with p1/p6 cleavage-site mutations (Kolli et al., 2006). We were interested in the association between P453L^{Gag} and D30N^{PR}/N88D^{PR} nelfinavir-resistant mutations, as P453L^{Gag} is located at the P5' position of the p1/p6 cleavage site, and expected to physically interact with the protease. Thus, we sought to clarify the biological advantage of interference among P453L^{Gag}, D30N^{PR}, and N88D^{PR} in recombinant viruses with patient- and NL4-3-derived genetic backgrounds. Virological advantage was evaluated in three aspects: (1) virological superiority in replication competency, (2) resistance to antiretroviral selective pressure, and (3) Gag processing pattern in virions. Our results indicated that the P453L^{Gag} cleavage-site mutation has the potential to improve the replication capacity and Gag processing of viruses with D30N^{PR}/N88D^{PR}, but has little effect on nelfinavir susceptibility. This latter finding is of interest since Gag cleavage-site mutations have been suggested as a mechanism for protease to develop drug resistance (Dam et al., 2009; Kolli et al., 2009). We also need to consider not only antiretroviral selective pressure but also immune selective pressure. Several of the positions we noted have been described as associated with human leukocyte antigen (HLA) escape mutations. For example, PR codons 12, 35, and 36, and Gag codons 373 and 374 are all potentially HLA-related (Brumme et al., 2009). To confirm the contributions of HLA and immune pressure, further study is required.

Although samples were collected chronologically at multiple times, the order of P453L^{Gag}, D30N^{PR} and N88D^{PR} acquisitions was unclear as all three mutations were detected at the same time. However, a plausible order seems to be the selection of D30N^{PR}/N88D^{PR} followed by P453L^{Gag} acquisition. Nelfinavir appears to select D30N^{PR}/N88D^{PR} mutations for resistance because these mutations obviously increase drug resistance to nelfinavir (Johnson et al., 2009), whereas the P453L^{Gag} mutation without any PR mutations has been reported to have almost no effect on susceptibility to PIs and on viral replication capacity (Maguire et al., 2002). Although D30N^{PR} is known as one of the most unstable PI-resistant mutations (Martinez-Picado et al., 1999) and viruses with this mutation have lower PR activity than the wild-type, the impaired replication caused by D30N^{PR} has been reported to be compensated by N88D^{PR} (Mitsuya et al., 2006; Sugiura et al., 2002). Furthermore, the fitness of virus with D30N^{PR}/N88D^{PR} was recovered in our study by an additional P453L^{Gag} mutation (Fig. 3). However, the P453L^{Gag} mutation was not introduced into NL4-3 carrying D30N^{PR}/N88D^{PR} (N^{+30/88}) and patient D30N^{PR}/N88D^{PR} clones (P^{-P453L}) during *in vitro* culture with nelfinavir, suggesting that the P453L^{Gag} mutation is a sufficient condition for D30N^{PR}/N88D^{PR} clones to replicate efficiently.

Virus with D30N^{PR}/N88D^{PR} was suggested by results of our Western blot analyses to process p1/p6 cleavage inefficiently, as

Table 4
Los Alamos National Laboratory HIV sequence database analysis.

453 ^{Gag}	30 ^{PR} /88 ^{PR}	
	D30/N88	N30N/D88D
P (n = 2801)	n = 2743	n = 8
L (n = 237)	n = 223	n = 8
p-value	<0.001	

demonstrated by the accumulation of p15 and p7 non-cleaved precursors, but addition of P453L^{Gag} improved the processing (Fig. 4B). Accumulation of p15 and p7 products has been reported in previous studies using different PR mutant viruses (Doyon et al., 1996; Maguire et al., 2002). In these studies, the additional mutations at NC/p1 or p1/p6 cleavage sites also resulted in efficient processing of these precursors. Interestingly, an aberrant band, which did not match either p6 or p7, was observed in N^{30/88} in our study (Fig. 4, lanes 5 and 6). Although further analyses will be required to determine the exact mechanisms, the band suggests inaccurate or alternative recognition of the cleavage site by virus with D30N^{PR}/N88D^{PR}, and P453L^{Gag} may confer an advantage by adjusting the protease to recognize and cleave the right site. As the HIV-1 p6 protein is important for efficient particle budding (von Schwedler et al., 2003), the defect in p1/p6 cleavage may affect viral maturation, which in turn may reduce viral infectivity and replication capacity.

To understand the relevance of P453L^{Gag} from a structural viewpoint, we used homology modeling with the published X-ray crystal structure of the PR–p1/p6 substrate complex as a template (Fig. 5). Although P453L^{Gag} is located at the P5' position and does not directly interfere with the protease active site or subsites, the modeling demonstrated that P453L^{Gag} can compensate for the binding affinity of PR and p1/p6. This mechanism is interesting because it suggests that a mutation outside the cleavage site interferes with the PR–Gag interaction. Indeed, Prabu-Jeyabalan et al. (2004) documented that a Gag mutation (A431V^{Gag}) compensates for a PR mutation (V82A^{PR}), which is not in direct contact with A431V^{Gag}. Thus, our data confirm the virological and structural advantages of P453L^{Gag} in viruses possessing D30N^{PR}/N88D^{PR}. Furthermore, this association appeared to be quite common as the frequency of P453L^{Gag} is 7.3% in the Los Alamos National Laboratory HIV sequence database. Though nelfinavir is no longer recommended as a first-line antiretroviral in the guidelines of developed countries, many cases previously exposed to nelfinavir have acquired D30N/N88D mutations. Indeed, the prevalence of the D30N mutation in PI-treated persons infected with subtype B viruses ($n = 7396$) and in nelfinavir-treated persons ($n = 1128$) is 7.9% and 28.1%, respectively, in the Stanford HIV drug resistance database (<http://hivdb.stanford.edu/>).

Regarding the bioinformatics analysis strategy, we selected the within-host substitution model in Spidermonkey analysis to infer the co-evolving sites (Nickle et al., 2007) as the data were sequences serially collected over 5 years from a single patient under anti-HIV treatment. One disadvantage of this program is that it does not account for the number of descendant clones. Often, mutation pairs on few viral clones might be determined as co-evolving pairs. In our study, 46^{PR} and 35^{PR} mutations were determined as a co-evolving pair with high posterior probability, but only one clone with this pair was observed among 129 sequences (Table 2), suggesting this co-mutation pair could not become “fixed” in a viral population. Thus, it is important to confirm the significance of the program output.

5. Conclusions

In conclusion, we successfully determined the Gag-protease associated sites P453L^{Gag}/D30N^{PR}/N88D^{PR} by applying single-genome sequencing, suggesting the usefulness of this method. However, as SGS is a more expensive method than direct sequencing, researchers need to consider which method has the best advantage for their samples. Extrapolating from our data, the relationships between the major mutations found by SGS may not differ significantly from direct sequencing results, but we might have a greater chance of seeing a variety of minority clones with minor mutations. In addition, our observation of higher variation at later

sampling points suggests that cases with longer treatment histories are good sample candidates.

We found that the viruses acquiring P453L^{Gag}/D30N^{PR}/N88D^{PR} distinctly showed biological advantages. From results obtained using both viral experiments and bioinformatics, we speculate that the P453L^{Gag} mutation does not necessarily occur in the presence of nelfinavir, but if it occurs with D30N^{PR}/N88D^{PR} mutations, viral fitness can be improved, which may eventually lead to worse clinical outcomes of anti-HIV therapy. We believe that the findings of this study provide new insight into the mechanism of within-patient HIV-1 co-evolution and into the acquisition of resistance to anti-HIV drugs.

Acknowledgements

The authors thank Dr. Akira Shirahata and Mr. Yuki Kitamura for their support. We thank patients who contributed to our study. We also thank Ms. Claire Baldwin for her help in preparing the manuscript. This study was supported by a Grant-in-Aid for AIDS research from the Ministry of Health, Labor and Welfare of Japan (H19-AIDS-007), and also by Scientific Research from the Ministry of Education, Culture, Sports, and Technology of Japan (Project number: 19510208).

References

- Altschuh, D., Lesk, A.M., Bloomer, A.C., Klug, A., 1987. Correlation of co-ordinated amino acid substitutions with function in viruses related to tobacco mosaic virus. *J. Mol. Biol.* 193, 693–707.
- Baker, D., Sali, A., 2001. Protein structure prediction and structural genomics. *Science* 294, 93–96.
- Bally, F., Martinez, R., Peters, S., Sudre, P., Telenti, A., 2000. Polymorphism of HIV type 1 gag p7/p1 and p1/p6 cleavage sites: clinical significance and implications for resistance to protease inhibitors. *AIDS Res. Hum. Retroviruses* 16, 1209–1213.
- Bhattacharya, T., Daniels, M., Heckerman, D., Foley, B., Frahm, N., Kadie, C., Carlson, J., Yusim, K., McMahon, B., Gaschen, B., Mallal, S., Mullins, J.I., Nickle, D.C., Herbeck, J., Rousseau, C., Learn, G.H., Miura, T., Brander, C., Walker, B., Korber, B., 2007. Founder effects in the assessment of HIV polymorphisms and HLA allele associations. *Science* 315, 1583–1586.
- Brumme, Z.L., John, M., Carlson, J.M., Brumme, C.J., Chan, D., Brockman, M.A., Swenson, L.C., Tao, I., Szeto, S., Rosato, P., Sela, J., Kadie, C.M., Frahm, N., Brander, C., Haas, D.W., Riddler, S.A., Haubrich, R., Walker, B.D., Harrigan, P.R., Heckerman, D., Mallal, S., 2009. HLA-associated immune escape pathways in HIV-1 subtype B Gag, Pol and Nef proteins. *PLoS One* 4, e6687.
- Chiba-Mizutani, T., Miura, H., Matsuda, M., Matsuda, Z., Yokomaku, Y., Miyauchi, K., Nishizawa, M., Yamamoto, N., Sugiura, W., 2007. Use of new T-cell-based cell lines expressing two luciferase reporters for accurately evaluating susceptibility to anti-human immunodeficiency virus type 1 drugs. *J. Clin. Microbiol.* 45, 477–487.
- Dam, E., Quercia, R., Glass, B., Descamps, D., Launay, O., Duval, X., Krausslich, H.G., Hance, A.J., Clavel, F., 2009. Gag mutations strongly contribute to HIV-1 resistance to protease inhibitors in highly drug-experienced patients besides compensating for fitness loss. *PLoS Pathog.* 5, e1000345.
- Doyon, L., Croteau, G., Thibeault, D., Poulin, F., Pilote, L., Lamarre, D., 1996. Second locus involved in human immunodeficiency virus type 1 resistance to protease inhibitors. *J. Virol.* 70, 3763–3769.
- Dutheil, J., Pupko, T., Jean-Marie, A., Galtier, N., 2005. A model-based approach for detecting coevolving positions in a molecule. *Mol. Biol. Evol.* 22, 1919–1928.
- Gallego, O., de Mendoza, C., Corral, A., Soriano, V., 2003. Changes in the human immunodeficiency virus p7-p1-p6 gag gene in drug-naive and pretreated patients. *J. Clin. Microbiol.* 41, 1245–1247.
- Gunthard, H.F., Wong, J.K., Ignacio, C.C., Havlir, D.V., Richman, D.D., 1998. Comparative performance of high-density oligonucleotide sequencing and dideoxynucleotide sequencing of HIV type 1 pol from clinical samples. *AIDS Res. Hum. Retroviruses* 14, 869–876.
- Gutell, R.R., Power, A., Hertz, G.Z., Putz, E.J., Stormo, G.D., 1992. Identifying constraints on the higher-order structure of RNA: continued development and application of comparative sequence analysis methods. *Nucleic Acids Res.* 20, 5785–5795.
- Hance, A.J., Lemiale, V., Izopet, J., Lecossier, D., Joly, V., Massip, P., Mammano, F., Descamps, D., Brun-Vézinet, F., Clavel, F., 2001. Changes in human immunodeficiency virus type 1 populations after treatment interruption in patients failing antiretroviral therapy. *J. Virol.* 75, 6410–6417.
- Ho, S.K., Coman, R.M., Bunger, J.C., Rose, S.L., O'Brien, P., Munoz, I., Dunn, B.M., Sleasman, J.W., Goodenow, M.M., 2008. Drug-associated changes in amino acid residues in Gag p2, p7(NC), and p6(Gag)/p6(Pol) in human immunodeficiency virus type 1 (HIV-1) display a dominant effect on replicative fitness and drug response. *Virology*.

- Johnson, V.A., Brun-Vézinet, F., Clotet, B., Gunthard, H.F., Kuritzkes, D.R., Pillay, D., Schapiro, J.M., Richman, D.D., 2008. Update of the drug resistance mutations in HIV-1: Spring 2008. *Top. HIV Med.* 16, 62–68.
- Johnson, V.A., Brun-Vézinet, F., Clotet, B., Gunthard, H.F., Kuritzkes, D.R., Pillay, D., Schapiro, J.M., Richman, D.D., 2009. Update of the drug resistance mutations in HIV-1: December 2009. *Top. HIV Med.* 17, 138–145.
- Koch, N., Yahi, N., Fantini, J., Tamalet, C., 2001. Mutations in HIV-1 gag cleavage sites and their association with protease mutations. *AIDS* 15, 526–528.
- Kolli, M., Laster, S., Schiffer, C.A., 2006. Co-evolution of nelfinavir-resistant HIV-1 protease and the p1-p6 substrate. *Virology* 347, 405–409.
- Kolli, M., Stawiski, E., Chappey, C., Schiffer, C.A., 2009. Human immunodeficiency virus type 1 protease-correlated cleavage site mutations enhance inhibitor resistance. *J. Virol.* 83, 11027–11042.
- Labute, P., 2008. The generalized Born/volume integral implicit solvent model: estimation of the free energy of hydration using London dispersion instead of atomic surface area. *J. Comput. Chem.* 29, 1693–1698.
- Maguire, M.F., Guinea, R., Griffin, P., Macmanus, S., Elston, R.C., Wolfram, J., Richards, N., Hanlon, M.H., Porter, D.J., Wrin, T., Parkin, N., Tisdale, M., Furfine, E., Petropoulos, C., Snowden, B.W., Kleim, J.P., 2002. Changes in human immunodeficiency virus type 1 Gag at positions L449 and P453 are linked to I50V protease mutants in vivo and cause reduction of sensitivity to amprenavir and improved viral fitness in vitro. *J. Virol.* 76, 7398–7406.
- Mahalingam, B., Louis, J.M., Reed, C.C., Adomat, J.M., Krouse, J., Wang, Y.F., Harrison, R.W., Weber, I.T., 1999. Structural and kinetic analysis of drug resistant mutants of HIV-1 protease. *Eur. J. Biochem.* 263, 238–245.
- Malet, I., Roquebert, B., Daiban, C., Wirdein, M., Amellal, B., Agher, R., Simon, A., Katlama, C., Costagliola, D., Calvez, V., Marcelin, A.G., 2007. Association of Gag cleavage sites to protease mutations and to virological response in HIV-1 treated patients. *J. Infect.* 54, 367–374.
- Marti-Renom, M.A., Stuart, A.C., Fiser, A., Sanchez, R., Melo, F., Sali, A., 2000. Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* 29, 291–325.
- Martin, L.C., Gloor, G.B., Dunn, S.D., Wahl, L.M., 2005. Using information theory to search for co-evolving residues in proteins. *Bioinformatics* 21, 4116–4124.
- Martinez-Picado, J., Savara, A.V., Sutton, L., D'Aquila, R.T., 1999. Replicative fitness of protease inhibitor-resistant mutants of human immunodeficiency virus type 1. *J. Virol.* 73, 3744–3752.
- Matsuoka-Aizawa, S., Sato, H., Hachiya, A., Tsuchiya, K., Takebe, Y., Gatanaga, H., Kimura, S., Oka, S., 2003. Isolation and molecular characterization of a nelfinavir (NFV)-resistant human immunodeficiency virus type 1 that exhibits NFV-dependent enhancement of replication. *J. Virol.* 77, 318–327.
- Meyerhans, A., Vartanian, J.P., Wain-Hobson, S., 1990. DNA recombination during PCR. *Nucleic Acids Res.* 18, 1687–1691.
- Mitsuya, Y., Winters, M.A., Fessel, W.J., Rhee, S.Y., Hurley, L., Horberg, M., Schiffer, C.A., Zolopa, A.R., Shafer, R.W., 2006. N88D facilitates the co-occurrence of D30N and L90M and the development of multidrug resistance in HIV type 1 protease following nelfinavir treatment failure. *AIDS Res. Hum. Retroviruses* 22, 1300–1305.
- Myint, L., Matsuda, M., Matsuda, Z., Yokomaku, Y., Chiba, T., Okano, A., Yamada, K., Sugiura, W., 2004. Gag non-cleavage site mutations contribute to full recovery of viral fitness in protease inhibitor-resistant human immunodeficiency virus type 1. *Antimicrob. Agents Chemother.* 48, 444–452.
- Neher, E., 1994. How frequent are correlated changes in families of protein sequences? *Proc. Natl. Acad. Sci. U. S. A.* 91, 98–102.
- Nickle, D.C., Heath, L., Jensen, M.A., Gilbert, P.B., Mullins, J.I., Kosakovsky Pond, S.L., 2007. HIV-specific probabilistic models of protein evolution. *PLoS One* 2, e503.
- Nijhuis, M., Schuurman, R., de Jong, D., Erickson, J., Gustchina, E., Albert, J., Schipper, P., Gulnik, S., Boucher, C.A., 1999. Increased fitness of drug resistant HIV-1 protease as a result of acquisition of compensatory mutations during suboptimal therapy. *AIDS* 13, 2349–2359.
- Palmer, S., Kearney, M., Maldarelli, F., Halvas, E.K., Bixby, C.J., Bazmi, H., Rock, D., Falloon, J., Davey Jr., R.T., Dewar, R.L., Metcalf, J.A., Hammer, S., Mellors, J.W., Coffin, J.M., 2005. Multiple, linked human immunodeficiency virus type 1 drug resistance mutations in treatment-experienced patients are missed by standard genotype analysis. *J. Clin. Microbiol.* 43, 406–413.
- Pollock, D.D., Taylor, W.R., 1997. Effectiveness of correlation analysis in identifying protein residues undergoing correlated evolution. *Protein Eng.* 10, 647–657.
- Poon, A.F., Kosakovsky Pond, S.L., Richman, D.D., Frost, S.D., 2007a. Mapping protease inhibitor resistance to human immunodeficiency virus type 1 sequence polymorphisms within patients. *J. Virol.* 81, 13598–13607.
- Poon, A.F., Lewis, F.I., Frost, S.D., Pond, S.L., 2008. Spidermonkey: rapid detection of co-evolving sites using Bayesian graphical models. *Bioinformatics*.
- Poon, A.F., Lewis, F.I., Pond, S.L., Frost, S.D., 2007b. An evolutionary-network model reveals stratified interactions in the V3 loop of the HIV-1 envelope. *PLoS Comput. Biol.* 3, e231.
- Prabu-Jeyabalan, M., Nalivaika, E., Schiffer, C.A., 2002. Substrate shape determines specificity of recognition for HIV-1 protease: analysis of crystal structures of six substrate complexes. *Structure* 10, 369–381.
- Prabu-Jeyabalan, M., Nalivaika, E.A., King, N.M., Schiffer, C.A., 2004. Structural basis for coevolution of a human immunodeficiency virus type 1 nucleocapsid-p1 cleavage site with a V82A drug-resistant mutation in viral protease. *J. Virol.* 78, 12446–12454.
- Rhee, S.Y., Liu, T.F., Holmes, S.P., Shafer, R.W., 2007. HIV-1 subtype B protease and reverse transcriptase amino acid covariation. *PLoS Comput. Biol.* 3, e87.
- Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425.
- Sayer, J.M., Liu, F., Ishima, R., Weber, I.T., Louis, J.M., 2008. Effect of the active site D25N mutation on the structure, stability, and ligand binding of the mature HIV-1 protease. *J. Biol. Chem.* 283, 13459–13470.
- Shirakawa, K., Takaori-Kondo, A., Yokoyama, M., Izumi, T., Matsui, M., Ito, K., Sato, T., Sato, H., Uchiyama, T., 2008. Phosphorylation of APOBEC3G by protein kinase A regulates its interaction with HIV-1 Vif. *Nat. Struct. Mol. Biol.* 15, 1184–1191.
- Sugiura, W., Matsuda, Z., Yokomaku, Y., Hertogs, K., Larder, B., Oishi, T., Okano, A., Shiino, T., Tatsumi, M., Matsuda, M., Abumi, H., Takata, N., Shirahata, S., Yamada, K., Yoshikura, H., Nagai, Y., 2002. Interference between D30N and L90M in selection and development of protease inhibitor-resistant human immunodeficiency virus type 1. *Antimicrob. Agents Chemother.* 46, 708–715.
- Tillier, E.R., Lui, T.W., 2003. Using multiple interdependency to separate functional from phylogenetic correlations in protein alignments. *Bioinformatics* 19, 750–755.
- Tuff, P., Darlu, P., 2000. Exploring a phylogenetic approach for the detection of correlated substitutions in proteins. *Mol. Biol. Evol.* 17, 1753–1759.
- Verheyen, J., Litau, E., Sing, T., Daumer, M., Balduin, M., Oette, M., Fatkenheuer, G., Rockstroh, J.K., Schuldenzucker, U., Hoffmann, D., Pfister, H., Kaiser, R., 2006. Compensatory mutations at the HIV cleavage sites p7/p1 and p1/p6-gag in therapy-naïve and therapy-experienced patients. *Antivir. Ther.* 11, 879–887.
- von Schwedler, U.K., Stuchell, M., Muller, B., Ward, D.M., Chung, H.Y., Morita, E., Wang, H.E., Davis, T., He, G.P., Cimbora, D.M., Scott, A., Krausslich, H.G., Kaplan, J., Morham, S.G., Sundquist, W.I., 2003. The protein network of HIV budding. *Cell* 114, 701–713.
- Wang, J., Cieplak, P., Kollman, P.A., 2000. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *Journal of Computational Chemistry* 21, 1049–1074.
- Wiley, R.L., Smith, D.H., Lasky, L.A., Theodore, T.S., Earl, P.L., Moss, B., Capon, D.J., Martin, M.A., 1988. In vitro mutagenesis identifies a region within the envelope gene of the human immunodeficiency virus that is critical for infectivity. *J. Virol.* 62, 139–147.
- Wollenberg, K.R., Atchley, W.R., 2000. Separation of phylogenetic and functional associations in biological sequences by using the parametric bootstrap. *Proc. Natl. Acad. Sci. U. S. A.* 97, 3288–3291.
- Wu, T.D., Schiffer, C.A., Gonzales, M.J., Taylor, J., Kantor, R., Chou, S., Israelski, D., Zolopa, A.R., Fessel, W.J., Shafer, R.W., 2003. Mutation patterns and structural correlates in human immunodeficiency virus type 1 protease following different protease inhibitor treatments. *J. Virol.* 77, 4836–4847.
- Yeang, C.H., Haussler, D., 2007. Detecting coevolution in and among protein domains. *PLoS Comput. Biol.* 3, e211.
- Zhang, Y.M., Imamichi, H., Imamichi, T., Lane, H.C., Falloon, J., Vasudevachari, M.B., Salzman, N.P., 1997. Drug resistance during indinavir therapy is caused by mutations in the protease gene and in its Gag substrate cleavage sites. *J. Virol.* 71, 6662–6670.

Outbreak of Infections by Hepatitis B Virus Genotype A and Transmission of Genetic Drug Resistance in Patients Coinfected with HIV-1 in Japan[∇]

Seiichiro Fujisaki,¹ Yoshiyuki Yokomaku,¹ Teiichiro Shiino,² Tomohiko Koibuchi,³ Junko Hattori,¹ Shiro Ibe,¹ Yasumasa Iwatani,^{1,4} Aikichi Iwamoto,³ Takuma Shirasaka,⁵ Motohiro Hamaguchi,⁶ and Wataru Sugiura^{1,4*}

Department of Infectious Diseases and Immunology, Clinical Research Center, National Hospital Organization, Nagoya Medical Center, Nagoya, Japan¹; Infectious Disease Surveillance Center, National Institute of Infectious Diseases, Tokyo, Japan²; Institute of Medical Science, The University of Tokyo, Tokyo, Japan³; Department of AIDS Research, Nagoya University Graduate School of Medicine, Nagoya, Japan⁴; AIDS Medical Center, National Hospital Organization, Osaka National Hospital, Osaka, Japan⁵; and Aichi Blood Center, Japanese Red Cross Society, Nagoya, Japan⁶

Received 24 October 2010/Returned for modification 2 December 2010/Accepted 8 January 2011

The major routes of hepatitis B virus (HBV) infection in Japan has been mother-to-child transmission (MTCT) and blood transfusion. However, HBV cases transmitted through sexual contact are increasing, especially among HIV-1-seropositive patients. To understand the molecular epidemiology of HBV in HBV/HIV-1 coinfection, we analyzed HBV genotypes and HIV-1 subtypes in HBV/HIV-1-coinfected patients at Nagoya Medical Center from 2003 to 2007. Among 394 HIV-1-infected Japanese men having sex with men (MSM) who were newly diagnosed during the study period, 31 (7.9%) tested positive for the hepatitis B virus surface antigen. HBV sequence analyses were successful in 26 cases, with 21 (80.7%) and 5 (19.3%) cases determined as genotypes A and C, respectively. Our finding that HBV genotype A was dominant in HIV-1-seropositive patients alerts clinicians to an alternative outbreak of HBV genotype A in the HIV-1-infected MSM population and a shift in HBV genotype from C to A in Japan. The narrow genetic diversity in genotype A cases suggests that genotype A has been recently introduced into the MSM population and that sexual contacts among MSM were more active than speculated from HIV-1 tree analyses. In addition, we found a lamivudine resistance mutation in one naïve case, suggesting a risk of drug-resistant HBV transmission. As genotype A infection has a higher risk than infection with other genotypes for individuals to become HBV carriers, prevention programs are urgently needed for the target population.

The number of hepatitis B virus (HBV)-infected persons in Japan is estimated to be 1 million, or 0.8% of the total population (31). HBV is classified into eight genotypes, A to H, by their differences in genome sequences (11, 12, 22). Circulating genotypes in Japan differ according to geographical region, with the prevalent genotypes in 2001 being C (84.7%) and B (12.2%), while A (1.7%) and D (0.4%) were less frequent (17). HBV infection in Japan has been transmitted mainly by two routes, mother-to-child transmission (MTCT) and blood transfusion, which have been targeted by prevention programs still being operated today (13, 15–17, 25).

Regarding MTCT, all pregnant women are screened for HBV antigen and antibody. Mothers who are HBV infected are prohibited from breast-feeding, and their newborns are vaccinated against HBV. Regarding infection by blood transfusion, all donated blood is tested by anti-hepatitis B surface antibody (HBsAb) testing and PCR to exclude HBV-contaminated blood from the supply. These prevention programs have

been successful, and the risks of HBV infection by these two routes have been reduced dramatically.

However, HBV infection by sexual contact has recently become a prevailing alternative transmission route of HBV in Japan (30, 36). In particular, coinfection with HBV and human immunodeficiency virus type 1 (HIV-1), the causative agent of AIDS, has been increasing among men who have sex with men (MSM), and the incidence of HBV infection associated with HIV-1-seropositive cases appeared to be 8.8%, which is higher than that in the general population (5). Thus, the epidemiology of HBV infection in Japan is quickly shifting. Here we report the most recent molecular epidemiologic status of HBV/HIV-1 coinfection.

MATERIALS AND METHODS

Sample. HIV/AIDS patients newly diagnosed at Nagoya Medical Center from 2003 to 2007 were tested for hepatitis B surface antigen (HBsAg), and HBsAg-positive patients were enrolled in the study. Clinical data (age, gender, suspected route of HIV-1 infection, aspartate aminotransferase [AST] and alanine aminotransferase [ALT] plasma levels, CD4-positive T cell count, and HIV viral load) were obtained from medical records. Plasma HBV viral load was measured with COBAS TaqMan (Roche Diagnostics, Basel, Switzerland), and plasma HBe IgM titer was measured with Lumipulse (Fujirebio, Tokyo, Japan). The time of HBV infection was estimated by patient interview and HBe IgM titer results. This study was conducted according to the principles expressed in the Declaration of Helsinki. The study was approved by the Institutional Review Boards of the National Institute of Infectious Diseases and Nagoya Medical Center. All pa-

* Corresponding author. Mailing address: Department of Infection and Immunology, Clinical Research Center, Nagoya Medical Center, 4-1-1 Sannomaru, Nakaku, Nagoya 4600001, Japan. Phone: 81-52-951-1111. Fax: 81-52-963-3970. E-mail: wsugiura@nnh.hosp.go.jp.

[∇] Published ahead of print on 19 January 2011.

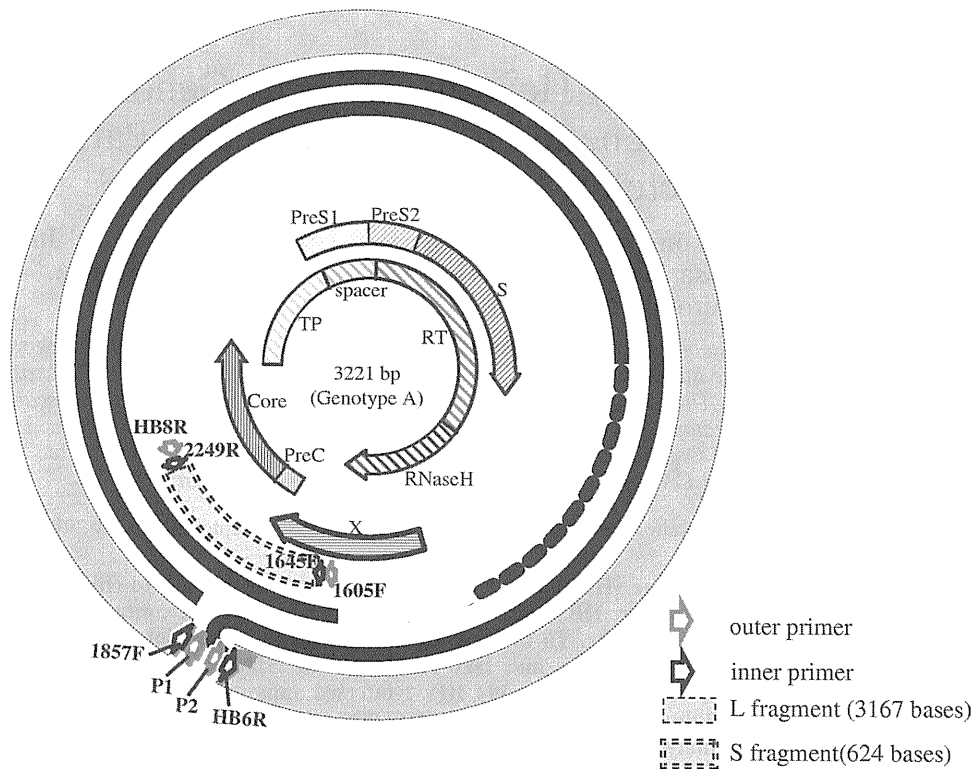


FIG. 1. Genetic regions of HBV and HIV-1 used for phylogenetic tree analyses. The whole HBV genome was amplified in two fragments, L and S, and assembled. L and S fragments are indicated by single and double dashed lines, respectively.

tients provided written informed consent for collection of samples and subsequent analysis.

Amplification of HBV and HIV DNA fragments and determination of DNA sequences. HBV nucleic acid was extracted from plasma using a MagNA Pure Compact Nucleic Acid Isolation Kit I (Roche Diagnostics). As shown in Fig. 1, the full-length HBV genome was amplified in two fragments, L (3,167 bases) and S (624 bases). The primers used for amplifying HBV DNA were both newly designed and have been published previously (27). Details of these primers are summarized in Table 1. The DNA polymerases used for the first and nested

PCRs were LA *Taq* (Takara, Shiga, Japan) and Prime Star HS (Takara) polymerase, respectively. The HBV genotypes were also determined using a commercial kit (Institute of Immunology, Tokyo, Japan) based on enzyme immunoassay to confirm that the results did not differ from those based on phylogenetic tree analysis.

The HIV-1 *gag p17* (396 bp [bp 790 to 1185]), *pol* (1,117 bp [bp 2253 to 3369]), and *env C2V3* (222 bp [bp 6996 to 7217]) regions were amplified from extracted plasma HIV-1 RNA by reverse transcription-PCR (RT-PCR) using the SuperScript one-step RT-PCR system for long templates (Invitrogen, Carlsbad, CA)

TABLE 1. Primers for amplifying the HBV and HIV-1 genomes

Name	Direction ^a	Sequence (5' → 3')	Region
P1	F	TTTTACCTCTGCCTAATCA	First PCR, HBV L fragment
P2	R	AAAAAGTTGCATGGTGCTGG	First PCR, HBV L fragment
1605F	F	CGCATGGAGACCACCGTGAA	First PCR, HBV S fragment
HB8R	R	ATAGGGGCATTTGGTGGTCT	First PCR, HBV S fragment
1857F	F	CTACTGTTCAAGCCTCCAAG	Nested PCR, HBV L fragment
HB6R	R	AACAGACCAATTTATGCCTA	Nested PCR, HBV L fragment
1645F	R	AGGTCTTGCATAAGAGGACT	Nested PCR, HBV S fragment
2249R	F	CCAAAAGACACCAAATAYTC	Nested PCR, HBV S fragment
172A	F	ATCTCTAGCAGTGGCGCCGAAACAG	RT-PCR, HIV-1 <i>gag</i> fragment
173B	R	CTGATAATGCTGAAAACATGGGTAT	RT-PCR, HIV-1 <i>gag</i> fragment
174A	F	CTCTCGACGCAGGACTCGGCTTGCT	Nested PCR, HIV-1 <i>gag</i> fragment
175B	R	CCCATGCATTCAAAGTTCTAGGTGA	Nested PCR, HIV-1 <i>gag</i> fragment
K1	F	AAGGGCTGTTGGAAATGTGG	RT-PCR, HIV-1 <i>pol</i> fragment
U13	R	CCCCTCAGGAATCCAGGT	RT-PCR, HIV-1 <i>pol</i> fragment
K4	F	GAAAGGAAGGACACCAAATGA	nested PCR, HIV-1 <i>pol</i> fragment
U12	R	CTCATTCTTGCATATTTCTCTGTT	Nested PCR, HIV-1 <i>pol</i> fragment
106A	F	CATACATTATTGTGCCCCGGCTGG	RT-PCR, HIV-1 <i>env</i> fragment
17B	R	AGAAAAATCCCCTCTACAATTA	RT-PCR, HIV-1 <i>env</i> fragment
14A	F	AATGTCAGCTCAGTACAATGCACAC	Nested PCR, HIV-1 <i>env</i> fragment
10B	R	ATTTCTGGTCCCCTCCTGAGG	Nested PCR, HIV-1 <i>env</i> fragment

^a F, forward; R, reverse.

TABLE 2. HBV genotype reference sequences collected from the DNA Database of Japan (DDBJ) for tMRCA analysis

Genotype	DDBJ accession no.
A.....	FJ692588, GQ325786, GQ477503, GQ477496, GQ486599, EU414132
B.....	FJ751547, GQ924611
C.....	GQ924615, GQ486096, EU939589, GQ486684
D.....	GQ486337, FJ349228, GQ924652, EU414124, GQ922001, GQ486586
E.....	GQ486756, GQ161830, FJ349237
F.....	GQ486537, GQ486515, GQ486570
G.....	GQ486843
H.....	GQ486592, AB266536

followed by a second PCR using LA *Taq* polymerase. The primers used for HIV-1 sequencing are also summarized in Table 1. The amplicons were purified using a MultiScreen PCR filter plate (Millipore, Billerica, MA), and the sequencing reaction was performed using the BigDye Terminator v3.1 cycle sequencing kit (Applied Biosystems, Carlsbad, CA) and analyzed with the ABI PRISM 3130 (Applied Biosystems) autosequencer. Electropherograms were edited and verified by SeqScape v2.5 software (Applied Biosystems).

Phylogenetic tree analyses and genotype determination. HBV genotypes were determined by phylogenetic tree analysis with reference sequences. HBV sequences were aligned with 23 reference sequences from the National Center for Biotechnology Information (NCBI) database by using the CLUSTAL W program and analyzed by Kimura two-parameter methods. Genetic distances were calculated by the maximum composite likelihood, and phylogenetic trees were constructed by the neighbor-joining method using MEGA version 4 software. The reliabilities of branches were evaluated by bootstrap analysis with 1,000 replicates.

Phylogenetic trees of the HIV-1 *gag*, *pol*, and *env* regions were also constructed with 62 HIV-1 reference sequences obtained from the HIV-1 sequence database (Los Alamos National Laboratory).

Estimated tMRCA. Evolutionary rates, chronological phylogenies, and other evolutionary parameters of HBV genotypes were estimated from heterochronous data for the HBV genomic sequences collected in our study, together with reference sequences from public databases (Table 2), using the Bayesian Markov chain Monte Carlo (MCMC) method. The nucleotide substitution model was evaluated by the hierarchical likelihood ratio test using PAUP v4.0 (29) with MrModeltest (14) and the general time-reversible (GTR) model with both invariable site (I) and gamma-distributed site (G) heterogeneity for four rate categories showing maximum likelihood. Bayesian MCMC analyses were performed with BEAST v1.4.8 (4) using the substitution model of GTR + I + G, three partitions into codon positions, and a relaxed molecular clock model (the uncorrelated log normal-distributed model) (3). Four different population dynamic models (exponential growth, logistic growth, constant population, and Bayesian skyline plot [BSP]) were tested in the analyses. According to BSP properties, constant-growth models were adopted for the HBV genome sequences. Each Bayesian MCMC analysis was run for 40 million states and sampled every 10,000 states. Posterior probabilities were calculated with a burn-in of 4 million states and checked for convergence using Tracer v1.4 (21). The maximum clade credibility tree for analyzing the MCMC data set was annotated by TreeAnnotator in the BEAST package. The posterior distribution of the substitution rate obtained from the heterochronous sequences was subsequently incorporated as a prior distribution for the mean evolutionary rate of the HBV genome, thereby adding a time scale to the phylogenetic histories of the given viruses and enabling estimation of the time of the most recent common ancestor (tMRCA) (19).

Determination of HBV drug resistance mutations. HBV cases resistant to nucleoside analogue reverse transcriptase inhibitors (NRTI) were determined by analyzing amino acid sequences of the RT region. The approved anti-HBV drugs in Japan are lamivudine, adefovir, and entecavir. In cases of HBV/HIV-1 coinfection, tenofovir and emtricitabine are also used. We studied whether the viruses have drug resistance mutations against these antiretroviral drugs with or without a history of antiretroviral treatments and confirmed the following resistance mutations: lamivudine/emtricitabine resistance mutations V173L, L180M, and M204I/V; adefovir resistance mutations A181V, I233V, and N236T; entecavir resistance mutations I169T, L180M, T184G, S202I, M204I/V, and M250V; and tenofovir resistance mutation A194T (1, 2, 24, 32, 34, 35). Furthermore, major drug resistance mutations in HIV-1 were defined according to the criteria

of the International AIDS Society (IAS)-USA and Stanford HIV drug resistance database (7, 23).

RESULTS

The major HBV genotype circulating among Japanese MSM is genotype A. During the study period, 394 cases were newly diagnosed as HIV/AIDS, and 31 cases were determined as HBsAg positive. Thus, the average prevalence of HBV/HIV-1 coinfection in our study population was 7.9%. Analysis of the coinfection prevalence in each year showed increases from 2.8 to 3.3% in 2003 to 2004 and from 7.4 to 13.2% in 2005 to 2007 (Fig. 2). As the suspected route of HIV-1 infections in all 31 cases was MSM, HBV appears to be quickly spreading among the MSM population. Of these HBV/HIV-1-coinfected cases, 26 isolates were successfully sequenced for both HBV and HIV-1, and their subtypes and genotypes were determined. Regarding the five cases for which the HBV genome could not be sequenced, plasma HBV DNA copies were undetectable in four cases, and low ($10^{3.3}$ copies/ml) in one case.

The median age of the patients was 34 years (interquartile range [IQR], 29.5 to 37.0) (Table 3). The median plasma viral loads of HBV and HIV-1 were 4.4×10^8 (IQR, $4.9 \times 10^4 - 6.3 \times 10^8$) and 6.4×10^4 (IQR, $2.0 \times 10^4 - 2.0 \times 10^5$) copies/ml, respectively. Hepatitis B core antigen (HBcAg) IgM was positive in nine patients, of which two were suspected to harbor acute HBV infection according to their HBsAg positivity, AST and ALT plasma levels, and patient interviews. The other 7 HBcAg-positive patients were categorized as having acute hepatitis or exacerbated chronic hepatitis, and 17 HBcAg-negative patients were determined as being in the chronic hepatitis stage.

According to phylogenetic tree analysis, 26 cases were classified into two genotypes, either A or C. As shown in Fig. 3, 21 and 5 cases were classified as genotypes A and C, respectively. The subgenotypes of the 21 genotype A cases were all A2, the predominant subgenotype in Europe and North America, whereas the subgenotypes of the 5 genotype C cases were all C1, the most prevalent subgenotype in eastern Asia, including Japan, South Korea, and northern China. Genotype B, the

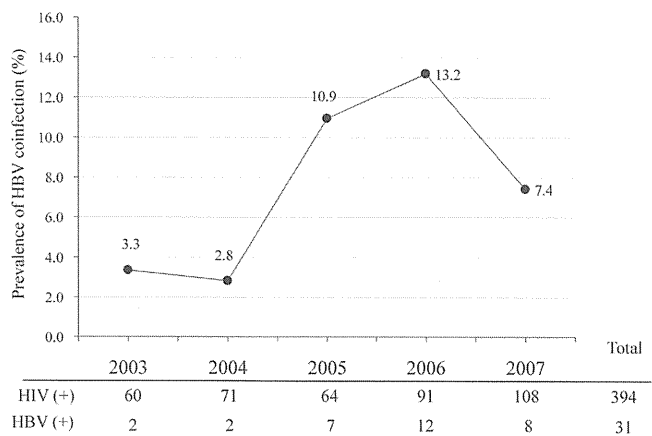


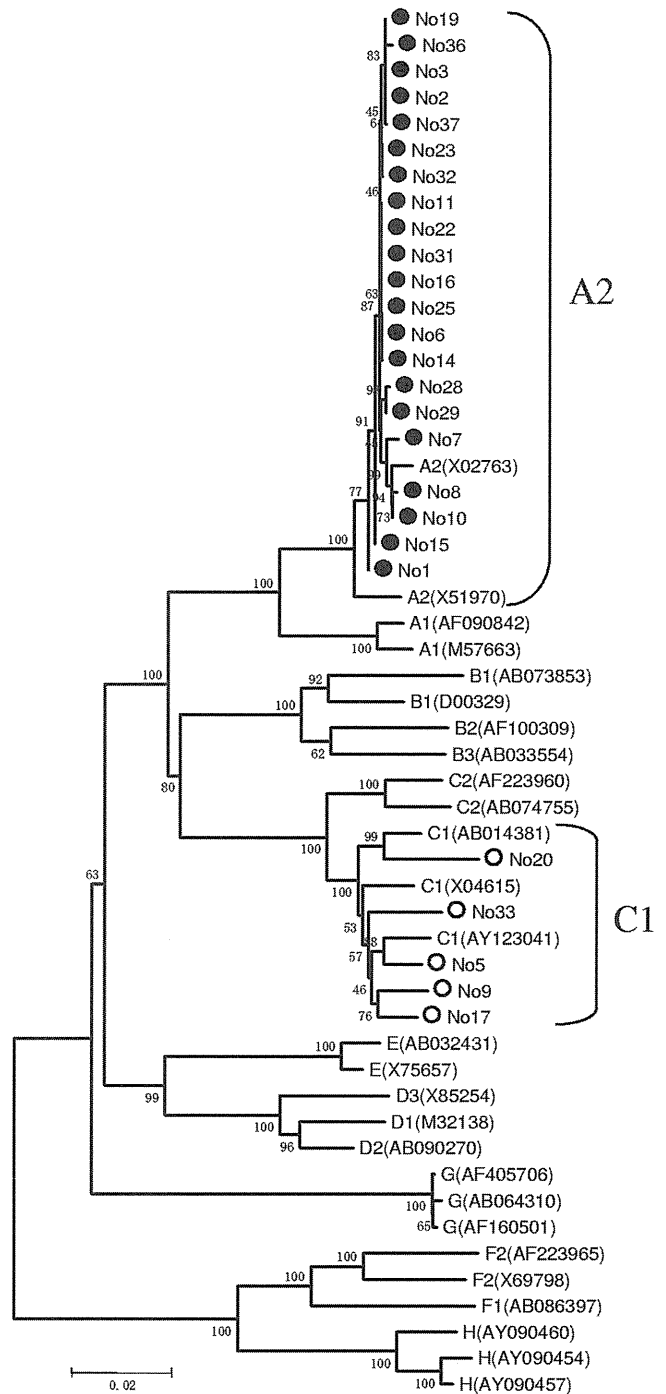
FIG. 2. Transitions in HBV infection rates in HBV/HIV-1-coinfected patients. HBV infection rates are plotted versus year, with the numbers of HIV-1-infected and HBV/HIV-1-coinfected patients shown below the x axis.

TABLE 3. Characteristics of HBV/HIV-1-coinfected patients

Characteristic	Value ^a for genotype:			P
	All (n = 26)	A (n = 21)	C (n = 5)	
Age (yr)	34 (30-37)	33 (29-37)	56 (46-57)	<0.01
Suspected route of HIV-1 infection	MSM	MSM	MSM	
AST (IU/liter)	31 (26-63)	29 (26-48)	54 (20-74)	<0.01
ALT (IU/liter)	43 (33-90)	42 (32-85)	44 (34-99)	
No. HBcAg IgM positive	9	9	0	
CD4 (μl)	293 (91-492)	300 (94-484)	202 (9-494)	
HIV-1viral load (copies/ml)	6.4×10^4 (2.0×10^4 - 2.0×10^5)	6.8×10^4 (2.4×10^4 - 2.1×10^5)	2.4×10^4 (2.4×10^3 - 9.7×10^4)	
HBV viral load (copies/ml)	4.4×10^8 (4.9×10^4 - 6.3×10^8)	6.3×10^8 (4.7×10^4 - 6.3×10^8)	2.0×10^8 (4.7×10^5 - 6.3×10^8)	

^a Median values are shown. Numbers in parentheses represent interquartile ranges.

second most predominant HBV genotype in Japan, was not detected in our study. Interestingly, the genotype A and C populations showed obvious differences in genetic diversity. The 21 group A2 samples (Fig. 3) formed a cluster with little



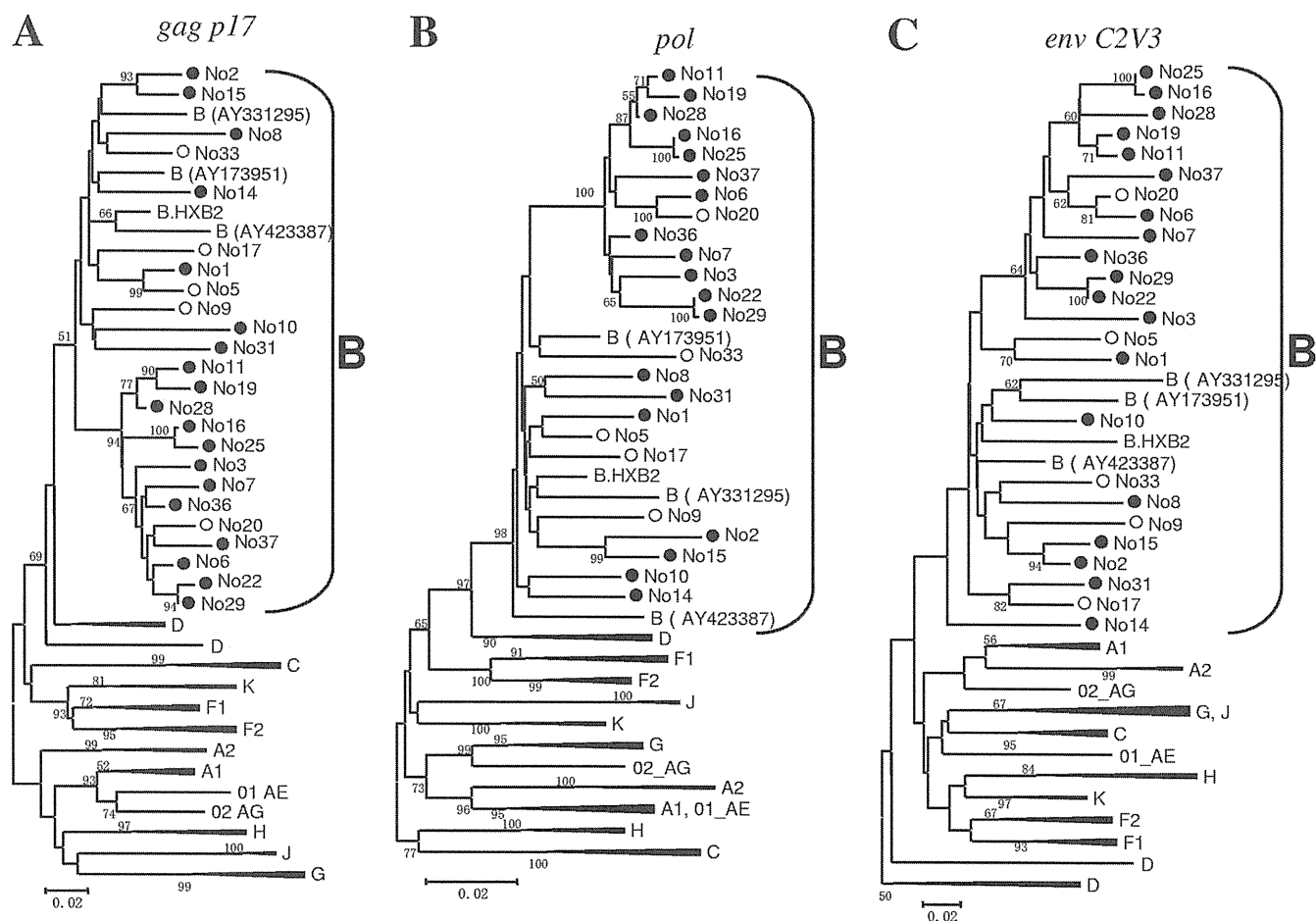


FIG. 4. Phylogenetic tree analyses of HIV-1 isolated from HBV/HIV-1-coinfected patients. Phylogenetic trees were constructed using the 25 HIV-1 sequences obtained in this study and 62 HIV-1 reference sequences from the Los Alamos National Laboratory database. The nucleotide base sequences of *gag p17* (A), *pol* PR to RT (B), and *env C2V3* (C) gene regions were analyzed. In all analyses, all the HIV-1-positive cases detected in Nagoya (both solid and open circles) were distributed in the subtype B cluster. Cases of coinfection with genotype C HBV are shown with open circles.

or no genetic distance between each other, indicating their extremely close genetic relationships. In contrast, the five group C1 cases did not form a single cluster and had longer branches than those of group A2.

Patients with genotypes A and C also differed significantly in age (Table 3). The median age of the genotype A patients was 33 years (IQR, 29 to 37), whereas that of the genotype C patients was 56 (IQR, 46 to 57) ($P < 0.01$). Furthermore, all nine HBcAg IgM-positive cases, including five suspected cases of acute infection, were categorized in genotype A2, suggesting ongoing active transmission of the virus among the Japanese MSM population. Thus, the genotype A2 population appeared to be younger, with more acute cases, and infected with an almost genetically identical HBV strain. These two genotypes did not differ significantly in regard other clinical data, such as AST and ALT levels, CD4⁺ T cell count, and HBV and HIV-1 viral loads.

To clarify the detailed epidemiological features of HBV/HIV-1-coinfected patients, the HIV-1 subtypes and their genetic distances were determined by phylogenetic analyses of three genome regions, *gag p17*, *pol*, and *env C2V3*. All 26 samples were determined as subtype B (Fig. 4A, B, and C), and

interestingly, branch patterns and relationships among cases were different from those for HBV. There were six paired cases, demonstrating a significantly close genetic relationship (>50% bootstrap value) in more than two regions. These paired cases were cases 1 and 5, 2 and 15, 6 and 20, 11 and 19, 16 and 25, and 22 and 29, and these connections were not evident in the HBV phylogeny, suggesting different origins of sexual partner between the two pathogens in each pair. An alternative explanation could be that little genetic variation in HBV made it difficult to clarify the genetic relationships between cases. However, there was one discordant pair (cases 1 and 5); one case had HBV genotype A, and the other case had HBV genotype C2. Furthermore, the other four HBV genotype C2 cases (cases 9, 17, 20, and 33) were scattered among HIV-1 phylogenies within genotype B HIV-1-infected patients.

HBV strains detected in HIV-1-infected patients from Nagoya are the same viruses found in other parts of Japan. To clarify whether the dominance of genotype A HBV in HIV-1-infected MSM is a regional issue in the Nagoya urban area or a more nationwide epidemic, we reconstructed an HBV phylogenetic tree of 26 cases together with HBV sequences col-

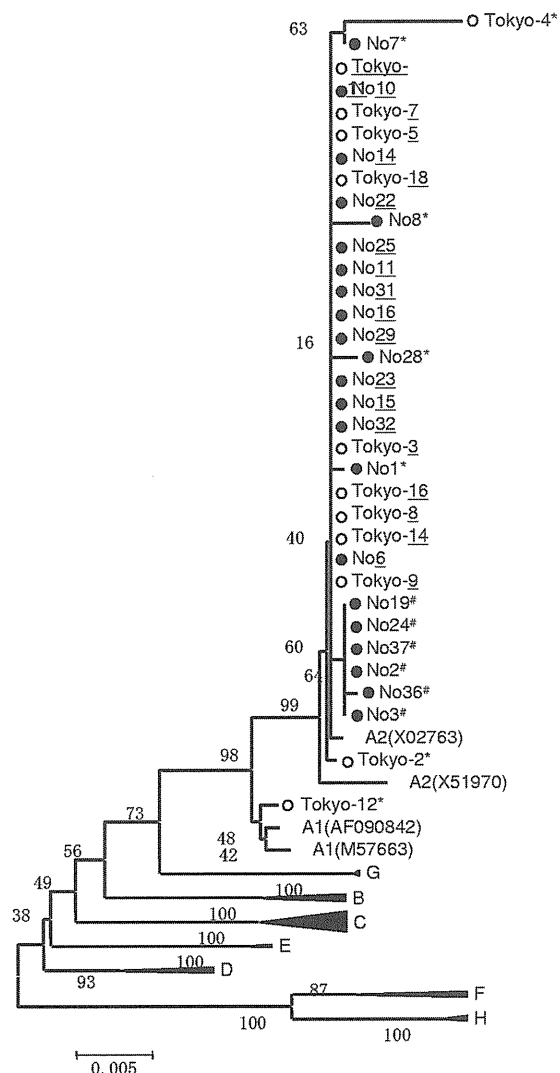


FIG. 5. Phylogenetic tree analysis of 35 HBV region S sequences, 22 from Nagoya (solid circles) and 13 from Tokyo (open circles). Three genetically different groups are indicated by asterisks, pound signs, and underlining.

lected at a different time and from a different area of Japan, i.e., 12 genotype A sequences from HBV/HIV-1-coinfected patients collected in Tokyo about 10 years before this study (8). As no full genome sequences were available for the Tokyo cases, only the S gene (681 bp [bp 155 to 835]) was analyzed. From the phylogenetic tree pattern, genotype A was classified into three groups (Fig. 5). The first is a group of 21 identical sequences (underlined in Fig. 5). As this group had the largest number of cases and included sequences from both Nagoya and Tokyo, this strain appears to be prevailing nationwide. The second group is a cluster of cases, i.e., cases 2, 3, 19, 24, 36, and 37. As all six cases were from Nagoya, this isolate still seems to be in an endemic status. The third group comprises isolates with longer branches (noted by asterisks), i.e., Tokyo-2, -4, and -12 and Nagoya-1, -7, -8, and -28. These isolates appear to be quite distinct from the others, suggesting that their origin may not be sexual contact but another route, such as MTCT or transfusions.

The prevailing HBV genotype A2 emerged more recently than most other genotypes. To estimate the emergence time of the prevailing genotype A2 strain, we estimated its mutation rate per year and tMRCA. First, the median mutation rate per year was calculated as 3.23×10^{-5} (5.62×10^{-8} to 9.01×10^{-5}), which is close to those previously reported (10, 18). Next, the median tMRCAs of all A strains, A1, A2, and C were determined to be 370.8, 88.9, 184.3, and 494.9 years ago, respectively (Table 4; Fig. 6). Thus, the A2 genotype is one of the youngest HBV genotypes.

A lamivudine resistance amino acid HBV mutation detected in an antiretroviral therapy-naïve patient. We clarified not only HBV genotypes but also the incidence of transmitted drug-resistant HBV among the study patients. Analysis of the amino acid sequence of the HBV RT region showed a combined triple amino acid mutation, rtV173L + rtL180M + rtM204V, which was a mutation causing resistance against lamivudine and its 5-fluoro analogue (2',3'-dideoxy-3'-thia-5-fluorocytidine), in two patients (patients 5 and 8). However, one patient (patient 5) had been treated with stavudine-lamivudine-efavirenz at the time of sample collection, and thus only one case (case 8) was suspected to be a transmitted HBV drug-resistant case. No HIV-1 drug-resistant virus transmission was detected in the study sample.

DISCUSSION

This molecular epidemiological study of HBV infection in HIV-1-seropositive patients revealed epidemiological characteristics that were unique compared to those of the general population in Japan. All HBV/HIV-1-coinfected patients were MSM, they had a 10-fold-higher prevalence (7.9%) than that of the general population, and genotype A was the predominant HBV genotype (31). This distinct HBV epidemic in MSM was first reported in 2001 in other regions of the country (9, 36), a decade before our study. Furthermore, phylogenetic analysis of sequences from the two studies, collected in different regions and years, revealed that an identical genotype A strain prevails among the MSM population nationwide.

Considering the status of HBV epidemiology in the general population of Japan, genotypes C and B must have an equal or greater chance to disseminate among the HIV-1-seropositive

TABLE 4. Estimated times of the most recent ancestor (tMRCAs) for HBV genotypes

Genotype	Mean tMRCA (yr before)		95% HPD ^a	
	Mean	Median	L	H
A	1,294.2	370.8	27.1	4,046.4
A1	306.6	88.9	12.4	976.4
A2	597.4	184.3	18.8	1,886.2
B	345.8	88.5	4.2	1,069.3
C	1,655.3	494.9	36.6	5,124.7
C2	1,062.4	308.6	20.8	3,296.6
D	827.2	226.6	11.2	2,469.4
E	163.7	38.9	4.5	539.7
F	1,060.8	308.2	13.7	3,277.4
H	433.8	110.1	5.6	1,303.0

^a HPD, highest posterior density.

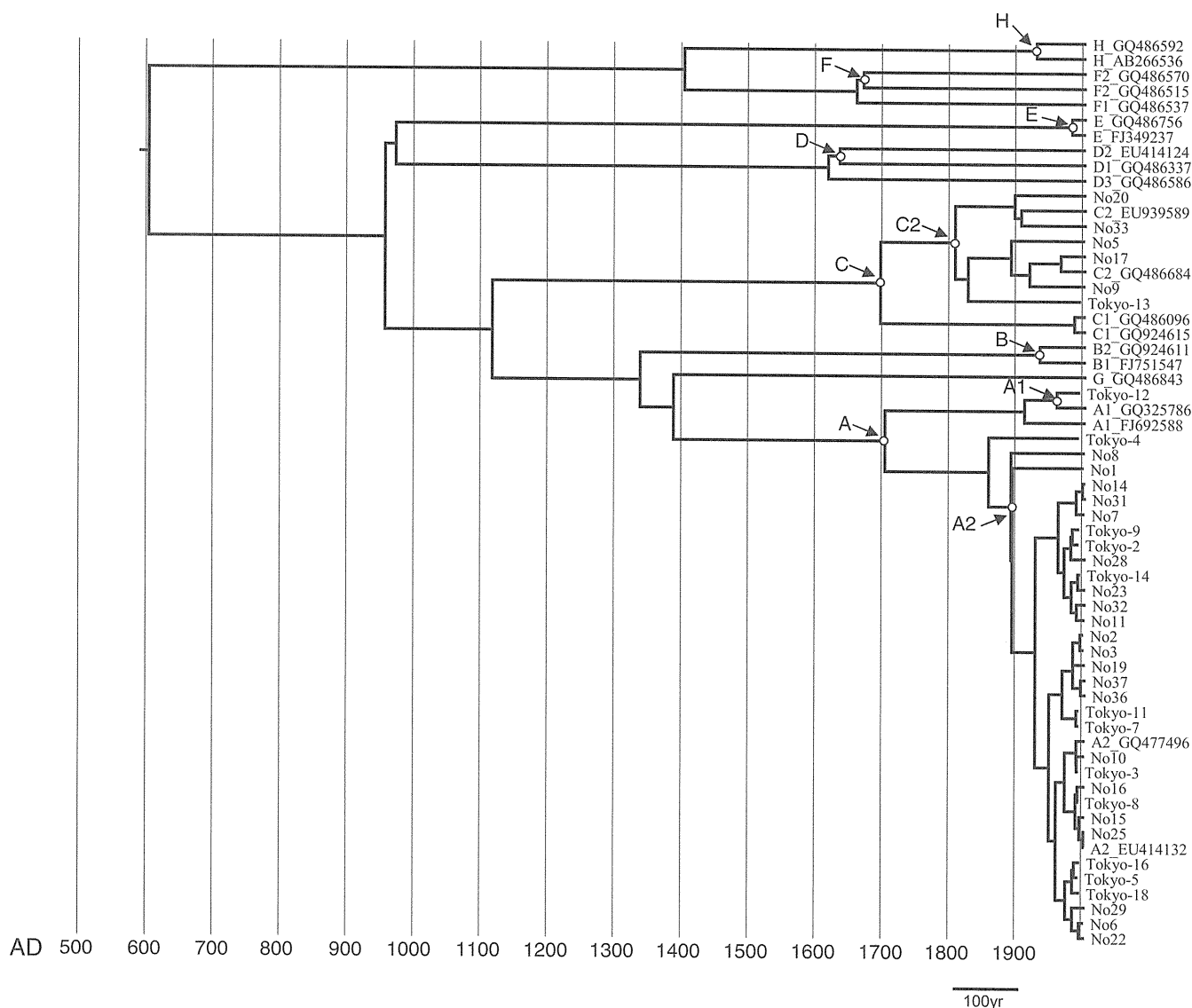


FIG. 6. Maximum clade credibility tree depicted according to median tMRCA. Nodes with open circles are evaluated points for each genotype summarized in Table 2.

MSM population. In fact, we found five genotype C patients in our study sample, and all five patients were MSM. However, because these five genotype C patients were older and their isolates had longer branches in phylogenetic analysis than the prevailing genotype A isolates, they appear to have been independently infected through either MTCT or blood transfusion events rather than sexual contact. Furthermore, as all five cases were singletons without any genetically close isolates among the samples analyzed, this genotype appeared to be less efficiently transmitted by sexual contact.

Interestingly, predominant genotype A HBV coinfection in HIV-seropositive MSM populations has also been reported in European and South American countries (20, 26, 33), suggesting that the prevailing genotype A in HIV-seropositive MSM has become a worldwide HBV epidemic. Regarding HBV genotypes in the HIV-negative population in Japan, genotype A has been increasing, but the major HBV

genotype is still C, with genotype A remaining at 3.5% nationwide and 2.1% in the Tokai area, which includes Nagoya city (9). Therefore, the prevalence of genotype A HBV in the MSM population is significantly higher than in the rest of the population.

Thus, it is interesting to discuss the virological advantages disposing this genotype A isolate to become the major player in HBV/HIV-1 coinfection among MSM. One such advantage might be the higher progression rate (16 to 23%) to chronicity of genotype A than of genotype C (28, 30), enhancing its capacity to serve as a source of new infections. As 9 of 26 genotype A-infected patients (35%) were HBcAg IgM positive and 2 had acute hepatitis, it is obvious that genotype A infections are actively ongoing among the MSM population. Though further studies are needed, considering the tMRCA of the prevailing strain A2, the younger age of patients infected with this strain than of those infected with other genotypes,

and its high prevalence among MSM, this strain may have acquired higher infectivity and efficient transmission through sexual contact.

Another issue we wanted to clarify in this study was the transmission of antiviral drug resistance. We found no antiretroviral resistance in the 26 sequenced cases. On the other hand, we detected two cases with a mutation combination of rtV173L + rtL180M + rtM204V in HBV reverse transcriptase, demonstrating resistance against lamivudine-emtricitabine. One patient was antiretroviral therapy naïve; thus, transmission of drug-resistant HBV is strongly suspected. It is peculiar that the isolate harboring the drug-resistant mutations in HBV was a singleton, considering that genetically identical isolates were prevailing, that there were very low mutation rates that suggest few chances of reverting to wild type, and that there were actively ongoing *de novo* infections. This finding might be due to resistant viruses being masked by wild-type viruses under untreated conditions, as reported in the case of HIV-1 drug resistance (6). The possibility of minority resistance populations of HBV could be verified by detection with a highly sensitive method.

In conclusion, we clarified the molecular epidemiology of HBV/HIV-1 coinfection in Japan. Our data suggest that ongoing HBV infections lie outside prevention programs targeting the MTCT and blood transfusion infection routes, and they suggest the urgent need for new prevention strategies focusing on the high-risk group of the HIV-1-seropositive MSM population.

ACKNOWLEDGMENTS

This study was supported by a Research Grant for Research on HIV/AIDS from the Ministry of Health, Labor, and Welfare of Japan (no. H19-AIDS-007, H21-AIDS-005, and H22-AIDS-004).

We thank Yasuhito Tanaka, Nagoya City University Graduate School of Medical Sciences, for helpful discussion and Claire Baldwin for help in preparing the manuscript.

REFERENCES

- Allen, M. I., et al. 1998. Identification and characterization of mutations in hepatitis B virus resistant to lamivudine. Lamivudine Clinical Investigation Group. *Hepatology* **27**:1670–1677.
- Angus, P., et al. 2003. Resistance to adefovir dipivoxil therapy associated with the selection of a novel mutation in the HBV polymerase. *Gastroenterology* **125**:292–297.
- Drummond, A. J., S. Y. Ho, M. J. Phillips, and A. Rambaut. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* **4**:e88.
- Drummond, A. J., and A. Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**:214.
- Gatanaga, H., et al. 2007. Drug-resistant HIV-1 prevalence in patients newly diagnosed with HIV/AIDS in Japan. *Antiviral Res.* **75**:75–82.
- Harrigan, P. R., S. Bloor, and B. A. Larder. 1998. Relative replicative fitness of zidovudine-resistant human immunodeficiency virus type 1 isolates in vitro. *J. Virol.* **72**:3773–3778.
- Johnson, V. A., et al. 2009. Update of the drug resistance mutations in HIV-1: December 2009. *Top. HIV Med.* **17**:138–145.
- Koibuchi, T., et al. 2001. Predominance of genotype A HBV in an HBV-HIV-1 dually positive population compared with an HIV-1-negative counterpart in Japan. *J. Med. Virol.* **64**:435–440.
- Matsuura, K., et al. 2009. Distribution of hepatitis B virus genotypes among patients with chronic infection in Japan shifting toward an increase of genotype A. *J. Clin. Microbiol.* **47**:1476–1483.
- Michitaka, K., et al. 2006. Tracing the history of hepatitis B virus genotype D in western Japan. *J. Med. Virol.* **78**:44–52.
- Miyakawa, Y., and M. Mizokami. 2003. Classifying hepatitis B virus genotypes. *Intervirology* **46**:329–338.
- Norder, H., et al. 2004. Genetic diversity of hepatitis B virus strains derived worldwide: genotypes, subgenotypes, and HBsAg subtypes. *Intervirology* **47**:289–309.
- Noto, H., et al. 2003. Combined passive and active immunoprophylaxis for preventing perinatal transmission of the hepatitis B virus carrier state in Shizuoka, Japan during 1980–1994. *J. Gastroenterol. Hepatol.* **18**:943–949.
- Nylander, J. 2004. MrModeltest v2. Uppsala University, Uppsala, Sweden.
- Oda, T. 2000. Further decline of hepatitis B surface antigen (HBsAg) prevalence in Japan. *Jpn. J. Cancer Res.* **91**:361.
- Okada, K., I. Kamiyama, M. Inomata, M. Imai, and Y. Miyakawa. 1976. e antigen and anti-e in the serum of asymptomatic carrier mothers as indicators of positive and negative transmission of hepatitis B virus to their infants. *N. Engl. J. Med.* **294**:746–749.
- Orito, E., et al. 2001. Geographic distribution of hepatitis B virus (HBV) genotype in patients with chronic HBV infection in Japan. *Hepatology* **34**:590–594.
- Osiowy, C., E. Giles, Y. Tanaka, M. Mizokami, and G. Y. Minuk. 2006. Molecular evolution of hepatitis B virus over 25 years. *J. Virol.* **80**:10307–10314.
- Pybus, O. G., A. J. Drummond, T. Nakano, B. H. Robertson, and A. Rambaut. 2003. The epidemiology and iatrogenic transmission of hepatitis C virus in Egypt: a Bayesian coalescent approach. *Mol. Biol. Evol.* **20**:381–387.
- Quarleri, J., et al. 2007. Hepatitis B virus genotype distribution and its lamivudine-resistant mutants in HIV-coinfected patients with chronic and occult hepatitis B. *AIDS Res. Hum. Retroviruses* **23**:525–531.
- Rambaut, A., and A. J. Drummond. 2007. Tracer v1.4. Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, Scotland. <http://tree.bio.ed.ac.uk>.
- Schaefer, S. 2007. Hepatitis B virus taxonomy and hepatitis B virus genotypes. *World J. Gastroenterol.* **13**:14–21.
- Shafer, R. 2010, posting date. Stanford drug resistance database. <http://hivdb.stanford.edu/>.
- Sheldon, J., et al. 2005. Selection of hepatitis B virus polymerase mutations in HIV-coinfected patients treated with tenofovir. *Antivir. Ther.* **10**:727–734.
- Shiraki, K. 2000. Perinatal transmission of hepatitis B virus and its prevention. *J. Gastroenterol. Hepatol.* **15**(Suppl.):E11–E15.
- Soriano, V., et al. 2010. Predictors of hepatitis B virus genotype and viraemia in HIV-infected patients with chronic hepatitis B in Europe. *J. Antimicrob. Chemother.* **65**:548–555.
- Sugauchi, F., et al. 2001. A novel variant genotype C of hepatitis B virus identified in isolates from Australian Aborigines: complete genome sequence and phylogenetic relatedness. *J. Gen. Virol.* **82**:883–892.
- Suzuki, Y., et al. 2005. Persistence of acute infection with hepatitis B virus genotype A and treatment in Japan. *J. Med. Virol.* **76**:33–39.
- Swafford, D. 2003. PAUP. Phylogenetic analysis using parsimony (and other methods), version 4. Sinauer Associates, Sunderland, MA.
- Takeda, Y., et al. 2006. Difference of HBV genotype distribution between acute hepatitis and chronic hepatitis in Japan. *Infection* **34**:201–207.
- Tanaka, J., et al. 2004. Sex- and age-specific carriers of hepatitis B and C viruses in Japan estimated by the prevalence in the 3,485,648 first-time blood donors during 1995–2000. *Intervirology* **47**:32–40.
- Tenney, D. J., et al. 2004. Clinical emergence of entecavir-resistant hepatitis B virus requires additional substitutions in virus already resistant to lamivudine. *Antimicrob. Agents Chemother.* **48**:3498–3507.
- Trimoulet, P., et al. 2007. Hepatitis B virus genotypes: a retrospective survey in southwestern France, 1999–2004. *Gastroenterol. Clin. Biol.* **31**:1088–1094.
- Westland, C. E., et al. 2003. Week 48 resistance surveillance in two phase 3 clinical studies of adefovir dipivoxil for chronic hepatitis B. *Hepatology* **38**:96–103.
- Yang, H., et al. 2002. Resistance surveillance in chronic hepatitis B patients treated with adefovir dipivoxil for up to 60 weeks. *Hepatology* **36**:464–473.
- Yotsuyanagi, H., et al. 2005. Distinct geographic distributions of hepatitis B virus genotypes in patients with acute infection in Japan. *J. Med. Virol.* **77**:39–46.



Trends in transmitted drug-resistant HIV-1 and demographic characteristics of newly diagnosed patients: Nationwide surveillance from 2003 to 2008 in Japan

Junko Hattori^a, Teiichiro Shiino^b, Hiroyuki Gatanaga^c, Shigeru Yoshida^d, Dai Watanabe^e, Rumi Minami^f, Kenji Sadamasu^g, Makiko Kondo^h, Haruyo Moriⁱ, Mikio Ueda^j, Masao Tateyama^k, Atsuhisa Ueda^l, Shingo Kato^m, Toshihiro Itoⁿ, Masayasu Oie^o, Noboru Takata^p, Tsunefusa Hayashida^c, Mami Nagashima^g, Masakazu Matsuda^q, Shiro Ibe^a, Yasuo Ota^r, Satoru Sasakiⁿ, Yoshiaki Ishigatsubo^l, Yoshinari Tanabe^o, Ichiro Koga^r, Yoko Kojima^l, Masahiro Yamamoto^f, Jiro Fujita^k, Yoshiyuki Yokomaku^a, Takao Koike^s, Takuma Shirasaka^e, Shinichi Oka^c, Wataru Sugiura^{a,b,t,u}

^a National Hospital Organization, Nagoya Medical Center, Clinical Research Center, 4-1-1 Sannomaru, Naka-ku, Nagoya 4600001, Japan

^b National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 1628640, Japan

^c AIDS Clinical Center, National Center for Global Health and Medicine, 1-21-1 Toyama, Shinjuku-ku, Tokyo 1628655, Japan

^d Faculty of Health Sciences, Hokkaido University, N12-W5, Kita-ku, Sapporo 0600812, Japan

^e National Hospital Organization, Osaka Medical Center, 2-1-14 Houenzaka, Chuo-ku, Osaka 5400006, Japan

^f National Hospital Organization, Kyushu Medical Center, 1-8-1 Jigyohama, Chuo-ku, Fukuoka 8108563, Japan

^g Tokyo Metropolitan Institute of Public Health, 3-24-1 Hyakunin-cho, Shinjuku-ku, Tokyo 1690073, Japan

^h Kanagawa Prefectural Institute of Public Health, 1-3-1 Shimomachiya, Chigasaki, Kanagawa 2530087, Japan

ⁱ Osaka Prefectural Institute of Public Health, 1-3-69 Nakamichi, Higashinari-ku, Osaka 5370025, Japan

^j Ishikawa Prefectural Central Hospital, 2-1 Angetsu-higashi, Kanazawa 9208530, Japan

^k University of the Ryukyus, 207 Uehara, Nishihara, Okinawa 9030125, Japan

^l Yokohama City University Graduate School of Medicine, 3-9 Fukuura, Kanazawa-ku, Yokohama 2360004, Japan

^m Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 1608582, Japan

ⁿ National Hospital Organization, Sendai Medical Center, 2-8-8 Miyagino, Miyagino-ku, Sendai 9830045, Japan

^o Niigata University Graduate School of Medical and Dental Sciences, 1-757 Asahimachi-dori, Chuo-ku, Niigata 9518510, Japan

^p Hiroshima University Hospital, 1-2-3 Kasumi, Minami-ku, Hiroshima 7348551, Japan

^q Mitsubishi Chemical Medience Corporation, 3-30-1 Shimura, Itabashi-ku, Tokyo 1748555, Japan

^r Teikyo University School of Medicine, 2-11-1 Kaga, Itabashi-ku, Tokyo 1738605, Japan

^s Hokkaido University Graduate School of Medicine, N15-W7, Kita-ku, Sapporo 0608638, Japan

^t Nagoya University Graduate School of Medicine, 65 Tsurumai, Showa-ku, Nagoya 4668550, Japan

ARTICLE INFO

Article history:

Received 16 April 2010

Received in revised form 12 July 2010

Accepted 28 July 2010

Keywords:

Drug-resistant HIV-1

Prevalence

Newly diagnosed HIV/AIDS patients

Treatment-naïve

BED assay

ABSTRACT

The emergence and transmission of drug-resistant human immunodeficiency virus-1 (HIV-1) compromises antiretroviral treatment for HIV-1. Thus, testing for drug resistance is recommended at diagnosis and before initiating highly active antiretroviral treatment. We conducted an epidemiological study enrolling newly diagnosed patients between 2003 and 2008 in our nationwide surveillance network. In the 6-year study period, the prevalence of drug-resistant HIV-1 among 2573 patients, consisting mainly of Japanese men in their late-30s and infected through male-to-male sexual contacts, followed an increasing trend from 5.9% (16/273) in 2003 to 8.3% (50/605) in 2008. Nucleoside reverse transcriptase inhibitor-associated mutations predominated in each year, with T215 revertants being the most abundant. The predictive factor for drug-resistant HIV-1 transmission was subtype B (OR = 2.36; $p = 0.004$), and those for recent HIV-1 infection were male gender (OR = 3.79; $p = 0.009$), MSM behavior (OR = 1.67; $p = 0.01$), Japanese nationality (OR = 2.31; $p = 0.008$), and subtype B (OR = 5.64; $p < 0.05$). Continued activities are needed to raise awareness of the risks of HIV-1 infection and complications of drug-resistant strains. Continued surveillance is also needed to understand trends in the HIV-1 epidemic.

© 2010 Elsevier B.V. All rights reserved.

Abbreviations: HIV-1, human immunodeficiency virus type 1; HAART, highly active antiretroviral therapy; PI, protease inhibitor; HBV, hepatitis B virus; HCV, hepatitis C virus; PR, protease; RT, reverse transcriptase; RT-PCR, reverse transcription polymerase chain reaction; CRF, circulating recombinant form; NRTI, nucleoside RT inhibitor; NNRTI, non-nucleoside RT inhibitor; OR, odds ratio; CI, confidence interval; MSM, men who have sex with men; IDU, intravenous drug user.

* Corresponding author at: Clinical Research Center, National Nagoya Medical Center, 4-1-1 Sannomaru Naka-ku, Nagoya, Aichi 4600001, Japan.

Tel.: +81 52 951 1111; fax: +81 52 963 3970.

E-mail address: wsugiura@nih.go.jp (W. Sugiura).

0166-3542/\$ - see front matter © 2010 Elsevier B.V. All rights reserved.

doi:10.1016/j.antiviral.2010.07.008

1. Introduction

The emergence of drug-resistant human immunodeficiency virus type 1 (HIV-1) among patients under highly active antiretroviral therapy (HAART) limits the successful suppression of HIV-1 replication. Several years after the introduction of HAART, drug-resistant strains are being detected among newly diagnosed HAART-naïve patients, suggesting the transmission of drug-resistant HIV-1 from the treatment-exposed population. Thus, treatment-naïve patients have been recommended by the US Department of Health and Human Services, International AIDS Society-USA, and other drug-resistance testing guidelines to undergo drug resistance testing at diagnosis and before initiation of HAART (DHHS, 2009; Hirsch et al., 2000, 2008). Indeed, choosing effective antiretrovirals according to the results obtained from this testing has led to successful control of HIV-1 infection. Furthermore, the drug resistance testing at diagnosis helps to understand transmission of drug-resistant HIV-1 in HAART-naïve individuals which in turn may help prevent transmission events.

The prevalence of drug-resistant HIV-1 among treatment-naïve patients has been closely monitored and reported from many countries. Before and early in the HAART era, when only mono or dual therapy was available, the prevalence was as high as 10–20% (Boden et al., 1999; Gómez-Cano et al., 1998; Tambussi et al., 1998). However, after the introduction of antiretrovirals with better pharmacokinetics, such as ritonavir-boosted protease inhibitor (PI), the emergence of drug-resistant viruses seemed to decrease (Gallego et al., 2001; Maia Teixeira et al., 2006).

Furthermore, despite the great number of HIV-1-infected patients, the prevalence tended to be low in developing countries where patients had limited or no access to antiretroviral drugs, e.g., 0–4.2% in Africa (Bártolo et al., 2009; Mintsu-Ndong et al., 2009; Ndembu et al., 2008; Pillay et al., 2008), 1.5% in Cambodia (Nouhin et al., 2009), and 2.6% in Vietnam (Ishizaki et al., 2009). In contrast, in countries where antiretroviral drugs are more accessible, the prevalence has been higher, e.g., 5.2% in Thailand (Apisarnthanarak et al., 2008), 9.4% in Taiwan (Chang et al., 2008), 10.0% in India (Lall et al., 2008), 7.8% in Portugal (Palma et al., 2007), 9.0% in Germany (Sagir et al., 2007), 9.5% in Belgium (Vercauteren et al., 2008), 10.9% in France (Chaix et al., 2009), and 15.9% in the US (Eshleman et al., 2007).

In Japan, since the first HIV-1-infected case was identified in 1985, the annual number of reported cases has been increasing every year, reaching 15 451 by the end of 2008. With more people getting infected, larger numbers of patients are starting anti-HIV-1 treatment and the risk of emerging drug-resistant HIV-1 is increasing. To understand the trends in drug-resistant HIV-1 in Japan, a nationwide surveillance project has been in effect since 2003. In our previous report of surveillance results from 2003 to 2004, the prevalence of drug-resistant HIV-1 in newly diagnosed patients was 4.0% (Gatanaga et al., 2007). We have continued collecting and analyzing data from newly diagnosed HIV-1-infected patients at participating clinical and research facilities in Japan. We report here the prevalence of drug-resistant HIV-1 among newly diagnosed therapy-naïve patients between 2003 and 2008.

2. Materials and methods

2.1 Sample

The study population included all the HIV-1-infected patients newly diagnosed between January 2003 and December 2008 at any of the participating HIV/AIDS clinics. Drug resistance genotypic tests were performed at 12 laboratories including 8 clinical laboratories at HIV/AIDS clinics, 3 public health laboratories, and

the National Institute of Infectious Diseases. After patients agreed to participate in our surveillance project and gave informed consent, peripheral blood was drawn with EDTA added, and their demographic and clinical information were collected. Demographic information included age, gender, nationality, and risk behavior. Clinical data included HIV-1 viral loads, CD4⁺ T cell counts, status of hepatitis B and C virus (HBV, HCV) co-infection, baseline sequence data, and drug-resistant amino acid mutations.

This study was conducted according to the principles in the Declaration of Helsinki, and was approved by the ethical committee of the National Institute of Infectious Diseases, Japan. By Japanese law, HIV-1-infected patients must be reported to the Japanese Ministry of Health, Labour, and Welfare upon diagnosis. The numbers reported to the Ministry are considered the “official numbers” of newly diagnosed HIV/AIDS cases, and were used as comparison controls to evaluate our study population.

2.2 Drug resistance genotypic testing

Drug resistance genotypic testing was performed using in-house protocols. Briefly, viral RNA was extracted from patient plasma samples. HIV-1 protease (PR, 1–99 amino acids) and the N-terminal region of reverse transcriptase (RT, 1–240 amino acids) were amplified in reverse transcription polymerase chain reaction (RT-PCR) followed by nested PCR using in-house primer sets. Subsequently, the amplified PCR products were purified and their sequences were analyzed by direct sequencing method using an automated sequencer. The resulting electropherograms were analyzed using commercially available software. The quality of testing methods used at each participating facility was assessed and confirmed for detection of drug-resistant mutations (Fujisaki et al., 2007). Thus, detection of drug-resistant mutations was consistent among facilities.

2.3 Determination of HIV-1 subtypes and drug-resistant HIV-1

HIV-1 subtypes were determined using the sequences of HIV-1 PR and RT genes obtained in the drug resistance genotypic testing explained above. Each sequence was aligned with the reference sequences of HIV-1 subtypes A through K, and circulating recombinant forms (CRFs), all of which were obtained from the Los Alamos HIV Databases (Los Alamos, 2010), using ClustalW, and phylogenetic trees were constructed using the neighbor-joining method with bootstrap value of 1000.

The resulting sequences were compared to that of HXB2 to judge the presence of amino acid mutations. The drug-resistant mutations were determined according to criteria of the HIV Drug Resistance Database of Stanford University (Bennett et al., 2009). Thus, a sample was considered to harbor drug-resistant HIV-1 if it possessed any of the following mutations: in the PR gene, L23I, L24I, D30N, V32I, M46I/L, I47V/A, G48V/M, I50V/L, F53L/Y, I54V/L/M/A/T/S, G73S/T/C/A, L76V, V82A/T/F/S/C/M/L, N83D, I84V/A/C, I85V, N88D/S, and L90M (indicating PI resistance); in the RT gene, M41L, K65R, D67N/G/E, T69D/insertion, K70R/E, L74V/I, V75M/T/A/S, F77L, Y115F, F116Y, Q151M, M184V/I, L210W, T215Y/F/I/S/C/D/V/E, K219Q/E/N/R (indicating nucleoside RT inhibitor [NRTI] resistance), and L100I, K101E/P, K103N/S, V106M/A, V179F, Y181C/I/V, Y188L/H/C, G190A/S/E, P225H, M230L (indicating non-nucleoside RT inhibitor [NNRTI] resistance).

2.4 BED assay

The time of HIV-1 seroconversion was estimated in randomly selected samples as recent (within 155 days) or not recent using the BED assay (Calypate HIV-1 BED Incidence EIA, BioRad) according to the Manufacturer's instruction. Briefly, 5 µL of plasma was diluted

with 500 μ L of sample diluent in the kit, and the proportion of anti-HIV-1 IgG to a total IgG in the sample was measured by optical density.

2.5. Statistical analysis

Statistical analyses were performed using R software (SAS Institute). Chi-square or Fisher's exact probability tests were used to determine associations among patients' demographic characteristics, nationality, BED assay results, and transmission of drug resistance. The odds ratio (OR) and 95% confidence intervals (CI) were calculated for all the variables. Recent and not-recent seroconversion groups were examined for differences in HIV-1 viral loads by analysis of covariance (ANCOVA), with CD4⁺ T cell count as the covariate.

3. Results

3.1. Majority of treatment-naïve patients are Japanese men who have sex with men (MSM) in mid-30s

The demographics of the 2573 newly diagnosed HIV-1-infected patients enrolled between 2003 and 2008 are summarized in Table 1. Male ($n = 2397$, 93.2%), Japanese (90.1%), and those infected through male-to-male sexual contact (68.9%) predominated, and the median age was 35. For the female cases ($n = 170$), high-risk heterosexual contact was the major risk factor ($n = 152$, 89.4%), and approximately half were non-Japanese ($n = 63$, 41.4%). Further analysis showed a significant association between the transmission route and nationality, i.e., most Japanese patients were infected through male-to-male sexual contact, while non-Japanese patients were infected by other routes (OR = 5.60; 95% CI 4.14–7.63; $p < 0.01$) (Table 2). It should be noted that sexual contacts (92.1%) are the major risk factor for HIV-1 infection in Japan. On the other hand, injecting drug usage, one of the high risk factors in other countries, accounts for only 0.4%.

HBV and/or HCV co-infection, an important clinical factor affecting prognosis and treatment of HIV infection (Ockenga et al., 1997; Piroth et al., 2000), was found to have a prevalence of 8.4% of 2101 patients, and 4.7% of 2071, respectively (Table 1). These prevalence rates did not change significantly throughout the study period (supplementary Table 1). HBV co-infection was found to be significantly associated with subtype B (OR = 2.04; $p < 0.05$) or infection through male-to-male sexual contact (OR = 1.66; $p < 0.05$).

3.2. Subtype B HIV-1 predominates in Japan

Of 2573 plasma samples collected during the study period, the sequences of PR and RT genes were successfully amplified and analyzed in 2536 (98.6%) and 2534 (98.5%) samples, respectively. Of these, we examined sequences of the PR-RT region from 2496 cases by phylogenetic tree analysis to determine the distribution of HIV-1 subtypes in Japan. Subtype B HIV-1 was found to predominate among the study population ($n = 2194$, 87.9%). The remaining non-B subtypes included 210 (8.4%) CRF01_AE, 30 (1.2%) C, 19 (0.8%) CRF02_AG, 18 (0.7%) A, 9 (0.4%) G, 7 (0.3%) F, 5 (0.2%) D, and 1 (0.04%) CRF08_BC (Table 1). In addition, 1 recombinant case of K/C, A/K, and D/B was detected in 2005, 2006, and 2007, respectively. These non-B subtype viruses were found mostly among the heterosexually infected population (223/302, 73.8%). In contrast, subtype B HIV-1 was found in the vast majority of MSM (1700/1773, 95.9%). In terms of nationality, Japanese patients, most of whom were MSM, were infected with subtype B HIV-1. On the other hand, only about a half of non-Japanese patients harbored subtype B HIV-1, and the remaining half were infected with non-B HIV-1, such as CRF01_AE

Table 1
Demographic characteristics of newly diagnosed HIV/AIDS patients.

	6-Year total (2573)	
Age		
Average	37.4	
Median	35	
Mode	35	
Quartile (Q1, Q3)	29, 43	
Nationality	<i>n</i>	(%)
Japanese	2319	(90.1)
Non-Japanese	225	(8.7)
Asian	83	(3.2)
Oceanian	4	(0.2)
North American	17	(0.7)
South American	58	(2.3)
European	10	(0.4)
African	26	(1.0)
Unspecified ^a	27	(1.0)
Unknown	29	(1.1)
Transmission category		
Male	2397	(93.2)
Male-to-male sexual contact	1773	(68.9)
High-risk heterosexual contact	369	(14.3)
Sexual contact	75	(2.9)
IDU	8	(0.3)
Other ^b	26	(1.0)
Unidentified	146	(5.7)
Female	170	(6.6)
High-risk heterosexual contact	152	(5.9)
IDU	3	(0.1)
Other ^b	5	(0.2)
Unidentified	11	(0.4)
Unknown	6	(0.2)
Unidentified	6	(0.2)
Hepatitis co-infection ^c		
HBV		
(+)	176	(8.4)
(-)	1925	(91.6)
Unknown	472	
HCV		
(+)	98	(4.7)
(-)	1973	(95.3)
Unknown	502	
HIV-1 subtype ^c		
B	2194	(87.9)
non-B	302	(12.1)
AE	210	(8.4)
C	30	(1.2)
AG	19	(0.8)
A	18	(0.7)
G	9	(0.4)
F	7	(0.3)
D	5	(0.2)
Other	4	(0.2)
Unidentified	77	

^a Unspecified individuals in the nationality category were identified only as of non-Japanese origin.

^b Other transmission categories include mother-to-child, blood products, transfusion, and needle stick.

^c Prevalence of subtypes, HBV, and HCV was calculated after omitting the unidentified or unknown data. DU, intravenous drug user; HBV, hepatitis B virus; HCV, hepatitis C virus; HIV-1, human immunodeficiency virus type 1.

(OR = 8.85; 95% CI 6.46–12.1; $p < 0.01$) (Table 2). This result is reasonable considering that the predominant HIV-1 subtype differs by country, and our study population included many Thais and Malaysians. In addition, this result suggests that subtype B HIV-1 is transmitted in a closed community of MSM, while non-B subtype strains are spread in wider areas among those infected through high-risk heterosexual contacts.

3.3. Prevalence of drug-resistant HIV-1 is increasing in Japan

A total of 194 cases (7.7%) in the 6-year study period were found to harbor HIV-1 strains with at least one major drug-resistant muta-

Table 2
Characteristics of newly diagnosed Japanese and non-Japanese HIV/AIDS patients.

	Nationality (n)			Odds ratio
	Japanese	Non-Japanese	Unknown	
Gender				
Male	2224	151	22	11.45 ^c
Female	95	74	1	
Unknown ^b			6	
Transmission category				
Male-to-male sexual contact	1691	73	9	5.60 ^{a,c}
High-risk heterosexual contact	399	114	7	
Sexual contact	72	4	0	
Other	29	10	2	
Unidentified ^b	128	24	11	
Subtype				
B	2051	118	25	8.85 ^c
Non-B	198	101	3	
Unidentified ^b	70	6	1	
BED assay (n = 640)				
Recent	220	13	0	2.31
Not recent	351	48	8	
Drug-resistant HIV-1				
Detected	173	16	5	1.05
Not detected	2146	209	24	

^a Odds ratios for the transmission category were calculated between male-to-male sexual contact and other categories which include high-risk heterosexual contact, sexual contact, and other.

^b Unknown and Unidentified cases were omitted in calculation of odds ratio.
^c p < 0.01.

tion conferred by PIs, NRTIs, or NNRTIs. The annual prevalence of drug-resistant mutations shown in Fig. 1 had an overall tendency to increase from 5.9% (16/273) in 2003 to 8.3% (50/605) in 2008. The most prevalent mutation in each year was NRTI-associated resistance, with 11 (4.0%), 12 (4.0%), 21 (5.0%), 23 (5.2%), 28 (5.9%), and 23 (3.7%) cases, followed by PI- and NNRTI-associated mutations. PI-resistant major mutations were detected in 63 cases (2.5%), and NNRTI-associated mutations were detected only in 20 cases (0.8%). These data reflect the type of antiretrovirals being prescribed in treated population. In other words, NRTIs have a long history of being prescribed including the period of mono and dual therapy; thus, NRTIs have been more frequently used. As a consequence, NRTI-resistant HIV-1 has emerged and been transmitted

more frequently to treatment-naïve patients. Regarding the drug-resistant mutations shown in Table 3, T215X (3.2%), M184I/V (0.5%), K103N (0.6%), and M46I/L (1.7%) accounted for the majority of detected mutations in contrast to other muta-

Table 3
Drug-resistant mutations in newly diagnosed HIV/AIDS patients, by class of antiretroviral drugs.

	6-Year total (2573)	
	n	(%)
NRTI ^a		
M41L	11	(0.4)
K65R	1	(0.0)
D67N/G/E	7	(0.3)
T69D	8	(0.3)
69INS	1	(0.0)
K70R/E	2	(0.1)
L74V/I	3	(0.1)
V75A/M	2	(0.1)
Y115F	3	(0.1)
M184V/I	12	(0.5)
L210W	5	(0.2)
T215X	81	(3.2)
K219Q/E/N/R	4	(0.2)
NNRTI ^a		
L100I	1	(0.0)
K101E	2	(0.1)
K103N	14	(0.6)
V106A/M	1	(0.0)
Y181C/I/V	3	(0.1)
P225H	1	(0.0)
P236L	1	(0.0)
PI ^a		
L24I	1	(0.0)
D30N	5	(0.2)
V32I	3	(0.1)
M46I/L	44	(1.7)
I47V/A	2	(0.1)
V82A/L	2	(0.1)
I85V	5	(0.2)
N88D/S	7	(0.3)
L90M	4	(0.2)

^a Numbers of cases and the proportions in parentheses are listed.

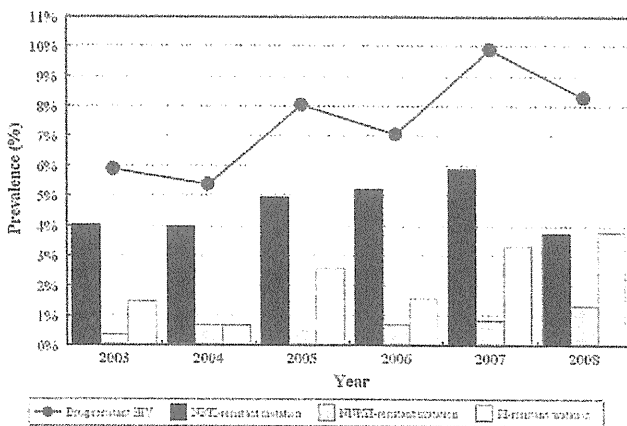


Fig. 1. Annual overall prevalence of drug-resistant HIV-1 (solid circles) in Japan increased in treatment-naïve patients in Japan from 2003 to 2008. The most prevalent mutation in each year was associated with resistance to nucleoside reverse transcriptase inhibitor (NRTI) treatment. Annual prevalence of drug-resistance mutations was categorized by antiretroviral drug class (NRTIs, solid black bars; non-nucleoside reverse transcriptase inhibitors [NNRTIs], horizontally striped bars; protease inhibitors [PIs], solid white bars). Drug-resistant HIV-1 was counted once even when the strain contained multiple drug-resistant mutations. Each drug-resistant mutation was counted even when multiple mutations were detected in one patient.