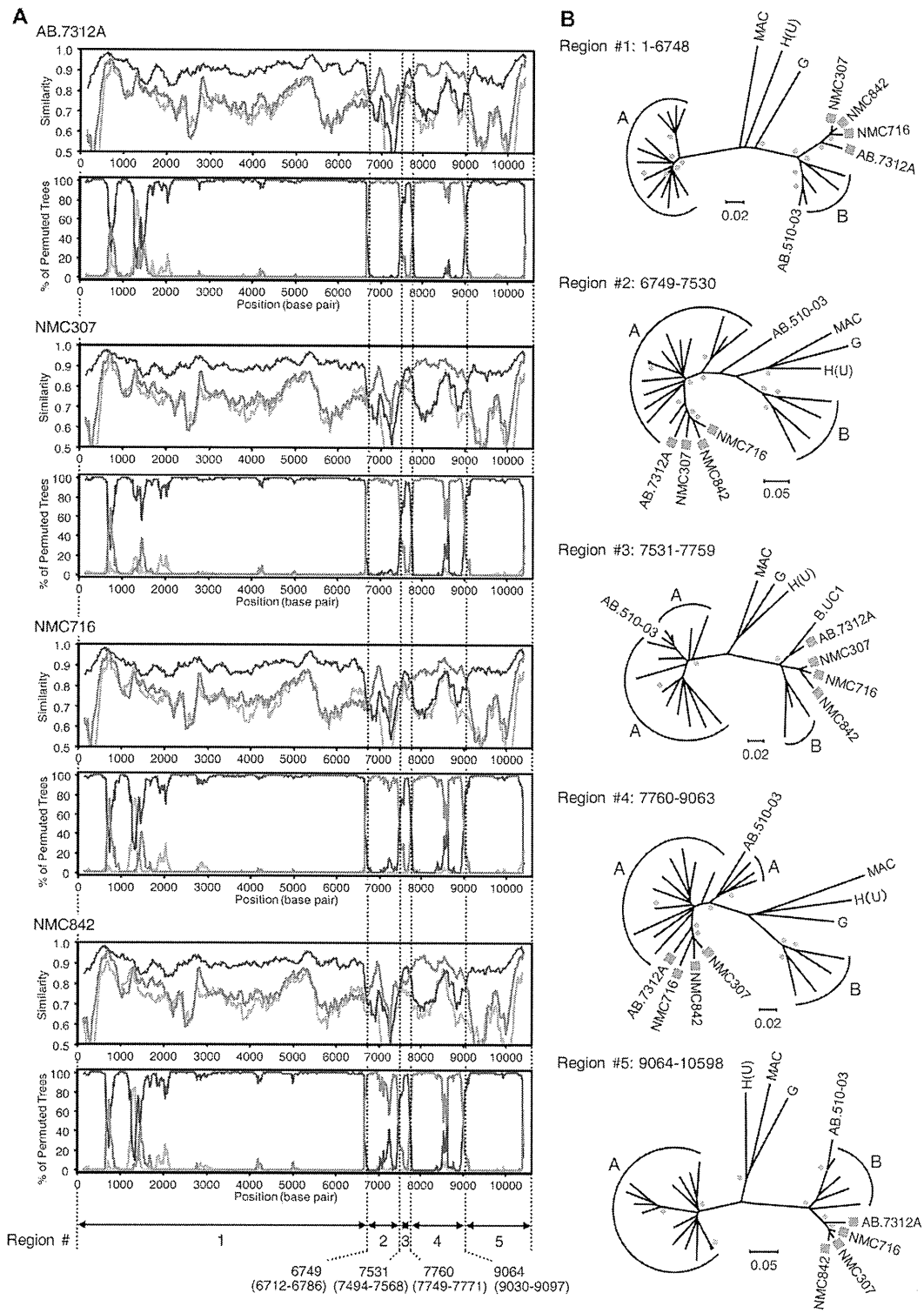


FIGURE 2. Determination of mosaic genome structures of HIV-2 AB recombinants. A, Similarity plotting (top) and bootscanning (bottom) data for each case of AB.7312A, NMC307, NMC716, and NMC842. Plots for consensus group A, consensus group B, and SIVmac239 are shown in red, blue, and gray, respectively. Both similarity plotting and bootscanning were performed with window and step sizes of 300 and 20 nucleotides, respectively. Bootscanning was performed using the neighbor-joining algorithm with 500 replicates. Each position of the 4 recombinant breakpoints is represented in the aligned sequence data set as the midpoint and

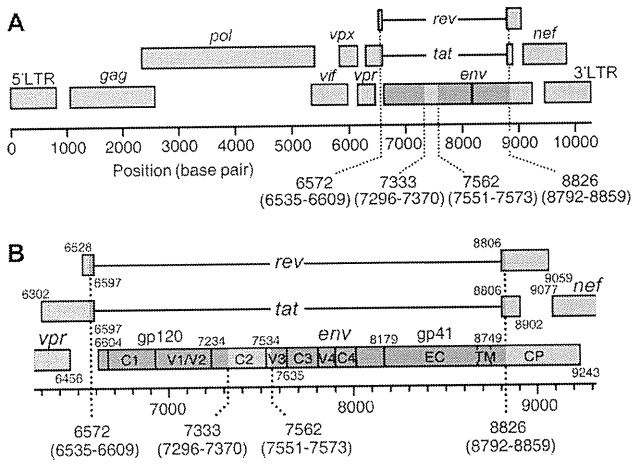


FIGURE 3. Schematic drawings for the genomic structure of HIV-2 CRF01_AB. A, Whole genomic structure; and B, Details around the *env* gene are represented. Regions belonging to group A and B are shown in red and blue, respectively. Numbering positions were adjusted to the reference SIVmac239 sequence.[35,36] Each position of 4 recombinant breakpoints is represented as the midpoint and range. C, constant region; CP, cytoplasmic domain; EC, extracellular domain; gp, glycoprotein; TM, transmembrane domain; V, variable region.

requirements are perfectly fulfilled with full-length genomic sequence information for 4 cases independently infected on different occasions with the AB recombinant identified by us and others.[12,13,19] Our data were carefully reviewed by editors of the Los Alamos HIV sequence database and confirmed as the first CRF discovered in HIV-2. They decided that the least confusing and most consistent way to name this new strain was to call it HIV-2 CRF01_AB.

The genomic structure of CRF01_AB is shown in Fig. 3. Interestingly, all 4 recombinant breakpoints of the CRF were located near or within the *env* gene (Fig. 3A). Further detailed analysis revealed that CRF01_AB possessed a chimeric gp120 containing a backbone of group A and a partial C2V3 fragment of group B and a chimeric gp41 containing extracellular and transmembrane domains of group A and a cytoplasmic domain of group B (Fig. 3B).

CRF01_AB Emerged Approximately in the Mid 20th Century

To estimate the time of CRF01_AB emergence, the time of the most recent common ancestor (tMRCA) of the recombinant was calculated by the Bayesian MCMC method. The mean substitution rates per year for the group A and B regions were estimated as 2.22×10^{-3} and 1.64×10^{-3}, respectively (Table 2), and the mean tMRCAs for groups A and B were estimated from 1921 to 1929, and from 1909 to 1948, respectively (Table 3). Similar results[31,37] validate our estimations. Finally, the mean tMRCA of CRF01_AB was estimated from 1964 to 1973. As the emergent times for groups A and B were estimated in the early 20th century, several decades seem to have been required for CRF01_AB to emerge. Concerning the geographical origin of the recombinant form, 3 of 4 isolates (7312A, NMC307, and NMC716) were identified in West Africans from Côte d'Ivoire and Nigeria. As these 2 countries were reported as sites of an epidemic in HIV-2 group A and B strains,[38,39] the most likely geographical origin of CRF01_AB is the south coastal area of West Africa.

DISCUSSION

In this study, we identified 3 HIV-2 AB recombinants with the same recombination pattern as 7312A, an isolate reported in Côte d'Ivoire in 1990.[12,13,19] These 4 isolates are determined as the first CRF of HIV-2, named CRF01_AB. It is noteworthy that all 3 of our cases infected with CRF01_AB were found at the AIDS stage. Considering that more than 75% of HIV-2–infected cases have a prognosis of remaining asymptomatic throughout their lifetimes[4] and that few HIV-2–seropositive cases were reported in Japan in the last 2 decades, 3 HIV-2 cases in the AIDS stage infected with the same CRF and identified in the past 5 years is highly unusual. Regarding the incubation periods for AIDS development in the 3 cases, not much information was available except for NMC842. This case was found to be seronegative for HIV-1/2 when tested in 2000. Thus, this case seems to have developed AIDS at most within 8 years, same as the median incubation period for AIDS development in HIV-1 infections (7.7–12.3 years).[40–45] As for the other 2 cases (NMC307 and NMC716), they developed AIDS at 28 and 36 years old (Table 1), which is significantly younger than age 65, reported as the peak of death by HIV-2 infections.[46,47] Though the number of cases identified is still small, we are concerned that the CRF01_AB might have acquired higher pathogenicity through recombination and adaptation to humans. As shown in Figure 3B, CRF01_AB has a recombination in the C2V3 region, the site of the major determinant for anti-envelope host immune responses and a functional domain for the chemokine receptor–binding site. The chimeric structure in the C2V3 region may confer advantages in host immune escape and viral replication capacity.

According to tMRCA analysis of the 4 isolates, CRF01_AB is estimated to have emerged sometime between 1964 and 1973. Interestingly, the mean tMRCA of the 3 isolates collected at NMC was estimated from 1982 to 1995 (Table 3), a later estimate than that of the 4 isolates, suggesting ongoing selection and evolution of CRF01_AB through transmission which has been taking place from the era of the 7312A isolate to the NMC isolates.

In conclusion, we report here the first CRF of HIV-2, CRF01_AB. Although national borders worldwide have

range (bottom). B, Subregion phylogenetic tree analyses. Phylogenetic trees were individually constructed by the neighbor-joining method using 5 subregion sequences. The HIV-2 isolates identified in this study (NMC307, NMC716, and NMC842) and AB.7312A are shown by green filled squares. Bootstrap values were calculated from 1000 analyses, and values greater than 95% are shown as orange dots at tree nodes. Scale bar represents 0.02 or 0.05 nucleotide substitutions per site. MAC, SIVmac239.

TABLE 2. Parameters in Bayesian MCMC Analysis for HIV-2/SIV Phylogenetic Inferences

	Substitution Rate Per Year		Coefficient of Variation		Population Size	
Data Set	Mean	95% HPD	Mean	95% HPD	Mean	95% HPD
Group A Region	2.22×10^{-3}	6.86×10^{-4}–3.68×10^{-3}	0.173	0.076–0.293	405.2	98.3–830.2
Group B Region	1.64×10^{-3}	5.99×10^{-4}–2.87×10^{-3}	0.269	0.170–0.395	341.2	93.3–668.9
Combined*	1.87×10^{-3}	6.39×10^{-4}–3.32×10^{-3}	0.235	0.088–0.382	357.9	93.3–709.2

*Combined data were produced from the 2 subsets, "group A region" and "group B region," using a LogCombiner program.
HPD, highest posterior density.

TABLE 3. Estimated TMRCAs of Monophyletic Clades in the HIV-2/SIV Lineage

	Group A region		Group B region		Combined	
Data set	Mean	95% HPD	Mean	95% HPD	Mean	95% HPD
Clade						
NMC isolates*	1982	1960–1996	1995	1987–2002	1990	1974–2002
CRF01_AB†	1964	1933–1985	1973	1956–1986	1971	1949–1986
Group A	1921	1864–1963	1929	1882–1964	1927	1879–1964
Group B	1909	1837–1962	1948	1915–1973	1934	1879–1973
HIV-2/SIV	1818	1670–1923	1821	1697–1930	1822	1693–1926

*This clade consisted of our 3 CRF01_AB isolates: NMC307, NMC716, and NMC842.
†This clade consisted of all 4 CRF01_AB isolates: 7312A, NMC307, NMC716, and NMC842.
HPD, highest posterior density;
SIV, simian immunodeficiency virus.

become more porous than ever, it is still surprising that the same recombinant strain was harvested in Japan, an island nation remote from the original endemic area, West Africa. This ectopic observation of the virus outside its endemic area suggests an ongoing global spread of HIV-2 CRF01_AB.

ACKNOWLEDGMENTS

We thank Dr. Thomas Leitner and editors of the Los Alamos HIV sequence database for discussing our data and naming the new HIV-2 circulating recombinant form. We thank Dr. Koya Ariyoshi for critical reading of our article and Claire Baldwin for her help in preparing the article.

REFERENCES

1. Marlink R. Lessons from the second AIDS virus, HIV-2. *AIDS*. 1996;10: 689–699.
2. Schim van der Loeff MF, Aaby P. Towards a better understanding of the epidemiology of HIV-2. *AIDS*. 1999;13:S69–S84.
3. Bock PJ, Markovitz DM. Infection with HIV-2. *AIDS*. 2001;15:S35–S45.
4. de Silva TI, Cotten M, Rowland-Jones SL. HIV-2: the forgotten AIDS virus. *Trends Microbiol*. 2008;16:588–595.
5. Whittle H, Morris J, Todd J, et al. HIV-2-infected patients survive longer than HIV-1-infected patients. *AIDS*. 1994;8:1617–1620.
6. Marlink R, Kanki P, Thior I, et al. Reduced rate of disease development after HIV-2 infection as compared to HIV-1. *Science*. 1994;265: 1587–1590.
7. Kanki PJ, Travers KU, MBoup S, et al. Slower heterosexual spread of HIV-2 than HIV-1. *Lancet*. 1994;343:943–946.
8. Adjorlolo-Johnson G, De Cock KM, Ekpini E, et al. Prospective comparison of mother-to-child transmission of HIV-1 and HIV-2 in Abidjan, Ivory Coast. *JAMA*. 1994;272:462–466.
9. Ota MO, O'Donovan D, Alabi AS, et al. Maternal HIV-1 and HIV-2 infection and child survival in The Gambia. *AIDS*. 2000;14:435–439.
10. O'Donovan D, Ariyoshi K, Milligan P, et al. Maternal plasma viral RNA levels determine marked differences in mother-to-child transmission rates of HIV-1 and HIV-2 in The Gambia. MRC/Gambia Government/ University College London Medical School working group on mother-child transmission of HIV. *AIDS*. 2000;14:441–448.
11. Schim van der Loeff MF, Jaffar S, Aveika AA, et al. Mortality of HIV-1, HIV-2 and HIV-1/HIV-2 dually infected patients in a clinic-based cohort in The Gambia. *AIDS*. 2002;16:1775–1783.
12. Gao F, Yue L, White AT, et al. Human infection by genetically diverse SIV_{SM}-related HIV-2 in West Africa. *Nature*. 1992;358:495–499.
13. Gao F, Yue L, Robertson DL, et al. Genetic diversity of human immunodeficiency virus type 2: evidence for distinct sequence subtypes with differences in virus biology. *J Virol*. 1994;68:7433–7447.
14. Chen Z, Luckay A, Sodora DL, et al. Human immunodeficiency virus type 2 (HIV-2) seroprevalence and characterization of a distinct HIV-2 genetic subtype from the natural range of simian immunodeficiency virus-infected sooty mangabeys. *J Virol*. 1997;71:3953–3960.
15. Yamaguchi J, Devare SG, Brennan CA. Identification of a new HIV-2 subtype based on phylogenetic analysis of full-length genomic sequence. *AIDS Res Hum Retroviruses*. 2000;16:925–930.
16. Damond F, Worobey M, Campa P, et al. Identification of a highly divergent HIV type 2 and proposal for a change in HIV type 2 classification. *AIDS Res Hum Retroviruses*. 2004;20:666–672.
17. Ndembi N, Abraha A, Pilch H, et al. Molecular characterization of human immunodeficiency virus type 1 (HIV-1) and HIV-2 in Yaoundé, Cameroon: evidence of major drug resistance mutations in newly diagnosed patients infected with subtypes other than subtype B. *J Clin Microbiol*. 2008;46:177–184.
18. Yamaguchi J, Vallari A, Ndembi N, et al. HIV type 2 intergroup recombinant identified in Cameroon. *AIDS Res Hum Retroviruses*. 2008; 24:86–91.
19. Robertson DL, Hahn BH, Sharp PM. Recombination in AIDS viruses. *J Mol Evol*. 1995;40:249–259.
20. Kusagawa S, Imamura Y, Yasuoka A, et al. Identification of HIV type 2 subtype B transmission in East Asia. *AIDS Res Hum Retroviruses*. 2003; 19:1045–1049.
21. Utsumi T, Nagakawa H, Uenishi R, et al. An HIV-2-infected Japanese man who was a long-term nonprogressor for 36 years. *AIDS*. 2007;21: 1834–1835.
22. Kato S, Hanabusa H, Kaneko S, et al. Complete removal of HIV-1 RNA and proviral DNA from semen by the swim-up method: assisted

reproduction technique using spermatozoa free from HIV-1. *AIDS*. 2006; 20:967–973.
23. Kinai E, Hanabusa H, Kato S. Prediction of the efficacy of antiviral therapy for hepatitis C virus infection by an ultrasensitive RT-PCR assay. *J Med Virol*. 2007;79:1113–1119.
24. Damond F, Loussert-Ajaka I, Apetrei C, et al. Highly sensitive method for amplification of human immunodeficiency virus type 2 DNA. *J Clin Microbiol*. 1998;36:809–811.
25. Tamura K, Dudley J, Nei M, et al. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol*. 2007;24: 1596–1599.
26. Lole KS, Bollinger RC, Paranjape RS, et al. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J Virol*. 1999;73: 152–160.
27. Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol*. 2007;7:214.
28. Drummond AJ, Ho SY, Phillips MJ, et al. Relaxed phylogenetics and dating with confidence. *PLoS Biol*. 2006;4:e88.
29. Wilgenbusch JC, Swofford D. Inferring evolutionary trees with PAUP*. *Curr Protoc Bioinformatics*. 2003;Chapter 6:Unit 6.4.
30. Rodríguez F, Oliver JL, Marín A, et al. The general stochastic model of nucleotide substitution. *J Theor Biol*. 1990;142:485–501.
31. Lemey P, Pybus OG, Wang B, et al. Tracing the origin and history of the HIV-2 epidemic. *Proc Natl Acad Sci U S A*. 2003;100:6588–6592.
32. Pybus OG, Drummond AJ, Nakano T, et al. The epidemiology and iatrogenic transmission of hepatitis C virus in Egypt: a Bayesian coalescent approach. *Mol Biol Evol*. 2003;20:381–387.
33. Robertson DL, Anderson JP, Bradac JA, et al. HIV-1 nomenclature proposal. In: Kuiken CL, Foley B, Hahn B, et al, eds. *Human Retroviruses and AIDS 1999*. Los Alamos, NM: Los Alamos National Laboratory; 1999:492–505.
34. Robertson DL, Anderson JP, Bradac JA, et al. HIV-1 nomenclature proposal. *Science*. 2000;288:55–56.
35. Calef C, Mokili J, O'Connor DH, et al. Numbering positions in SIV relative to SIVMM239. In: Kuiken C, Foley B, Hahn B, et al, eds. *HIV Sequence Compendium 2001*. Los Alamos, NM: Los Alamos National Laboratory; 2001:171–181.
36. Lin G, Bertolotti-Ciarlet A, Haggarty B, et al. Replication-competent variants of human immunodeficiency virus type 2 lacking the V3 loop exhibit resistance to chemokine receptor antagonists. *J Virol*. 2007;81: 9956–9966.
37. Wertheim JO, Worobey M. Dating the age of the SIV lineages that gave rise to HIV-1 and HIV-2. *PLoS Comput Biol*. 2009;5:e1000377.
38. Pieniazek D, Ellenberger D, Janini LM, et al. Predominance of human immunodeficiency virus type 2 subtype B in Abidjan, Ivory Coast. *AIDS Res Hum Retroviruses*. 1999;15:603–608.
39. Zeh C, Pieniazek D, Agwale SM, et al. Nigerian HIV type 2 subtype A and B from heterotypic HIV type 1 and HIV type 2 or monotypic HIV type 2 infections. *AIDS Res Hum Retroviruses*. 2005;21:17–27.
40. Hessol NA, Koblin BA, van Griensven GJ, et al. Progression of human immunodeficiency virus type 1 (HIV-1) infection among homosexual men in hepatitis B vaccine trial cohorts in Amsterdam, New York City, and San Francisco, 1978–1991. *Am J Epidemiol*. 1994;139:1077–1087.
41. Veugelers PJ, Page KA, Tindall B, et al. Determinants of HIV disease progression among homosexual men registered in the Tricontinental Seroconverter Study. *Am J Epidemiol*. 1994;140:747–758.
42. UK Register of HIV Seroconverters Steering Committee. The AIDS incubation period in the UK estimated from a national register of HIV seroconverters. *AIDS*. 1998;12:659–667.
43. Pezzotti P, Galai N, Vlahov D, et al. Direct comparison of time to AIDS and infectious disease death between HIV seroconverter injection drug users in Italy and the United States: results from the ALIVE and ISS studies. *J Acquir Immune Defic Syndr Hum Retrovirol*. 1999;20:275–282.
44. Collaborative Group on AIDS Incubation and HIV Survival including the CASCADE EU Concerted Action. Time from HIV-1 seroconversion to AIDS and death before widespread use of highly-active antiretroviral therapy: a collaborative re-analysis. *Lancet*. 2000;355:1131–1137.
45. Morgan D, Mahe C, Mayanja B, et al. HIV-1 infection in rural Africa: is there a difference in median time to AIDS and survival compared with that in industrialized countries? *AIDS*. 2002;16:597–603.
46. Poulsen AG, Aaby P, Larsen O, et al. 9-year HIV-2-associated mortality in an urban community in Bissau, west Africa. *Lancet*. 1997;349:911–914.
47. Berry N, Jaffar S, Schim van der Loeff M, et al. Low level viremia and high CD4% predict normal survival in a cohort of HIV type-2-infected villagers. *AIDS Res Hum Retroviruses*. 2002;18:1167–1173.

第23回日本エイズ学会シンポジウム記録

HIV細胞進入とその防御機序

HIV Entry : New Insights of the Molecular Mechanism

松下　修三[1]，横山　勝[2]，宮内　浩典[3,4]，松田　善衛[5,6]，俣野　哲朗[7]，岩谷　靖雅[8]

Shuzo MATSUSHITA[1], *Masaru YOKOYAMA*[2], *Kosuke MIYAUCHI*[3,4], *Zene MATSUDA*[5,6], *Tetsuro MATANO*[7] *and Yasumasa IWATANI*[8]

[1]熊本大学エイズ学研究センター
[2]国立感染症研究所病原体ゲノム解析研究センター
[3]国立感染症研究所エイズ研究センター
[4]エイズ予防財団
[5]東京大学医科学研究所アジア感染症研究拠点
[6]中国科学院生物物理研究所
[7]東京大学医科学研究所感染症国際研究センター
[8]国立病院機構名古屋医療センター臨床研究センター

はじめに

AIDSの主因であるHIV感染症は，ウイルスがリンパ球を中心としたCD4陽性細胞に慢性感染することに起因する。さらに，ウイルスの細胞指向性の変容が病態進行と深く関連していると考えられている。これらHIV感染症の特徴を決定するものがHIVエンベロープタンパク（Env）である。一方で，細胞外に露出しているEnvは個体におけるウイルス感染防御の格好な標的でもありうる。そのため，Envは，ウイルス感染において特異的な細胞へ吸着し進入する機能が維持されつつ，免疫（特に中和抗体）による防御から逃れるためにその抗原性が変化し続けている。Envの複雑な機能と多様性を解明するには，Env多量体構造を踏まえた構造学的解析が必要不可欠になっており，HIV研究の中でも非常に複雑な研究分野になっている。1996年のケモカインレセプター（Fusin）の発見以来，構造学研究者を取り込んで，今再び，Env研究が脚光を浴びはじめている。本稿では，HIVのEnvとその機能について，第23回日本エイズ学会学術集会シンポジウムにおいて発表された内容を中心に，研究の最前線で活躍している研究者の成果を概説する。

著者連絡先：岩谷靖雅（〒460-0001　愛知県名古屋市中区三の丸4-1-1　国立病院機構名古屋医療センター臨床研究センター）

2010年4月8日受付

SY4-1.　エンベロープの進化と中和抗体

松下　修三（熊本大学・エイズ学研究センター　病態制御分野）

Shuzo Matsushita（Center for AIDS Research, Kumamoto University）

HIV-1のエンベロープ蛋白は標的細胞のCD4分子及びCCR5やCXCR4などのケモカインレセプターと相互作用するgp120（SU）と，gp120とともに3量体を形成し，膜融合から侵入の過程において重要な役割を果たすgp41（TM）からなる。関連する霊長類ウイルスの系統樹解析から，HIV-1はgroup M，O及びNに分けられているが，世界各地での流行株には特徴があり，group Mは13以上のsubtype（またはclades）に分類されている。また，これらのsubtype間のキメラウイルスの流行が認められ，疫学的解析を複雑なものにしている。一方，HIV-1は感染個体内では互いに似ているが遺伝子配列の異なるクアシスピーシスを形成し，細胞性・液性の免疫選択圧の下に，刻々と変化（進化）すると考えられる。我々は中和単クローン抗体を用いて，ウイルスの中和エスケープ変異を研究し，標的エピトープの変化ばかりでなく，V2領域による強固な3量体構造の形成や，糖鎖による遮蔽，さらにこれらの変異によって起こる増殖性の低下を補う変異がほぼ同時に起こってくることを見出した。これは，*in vivo*で感染初期に起こる現象を説明するとともに，エンベロープ進化のプロセスの一部を明らかにするものと考えられる。また，図1と2に示すように，CCR5と相互作用するV3やCD4 induced（CD4i）epitopesはCD4-gp120の結合後に露出し，中和抗体の標的となるが，遮蔽が強力であればある程感染力は弱くなるこ

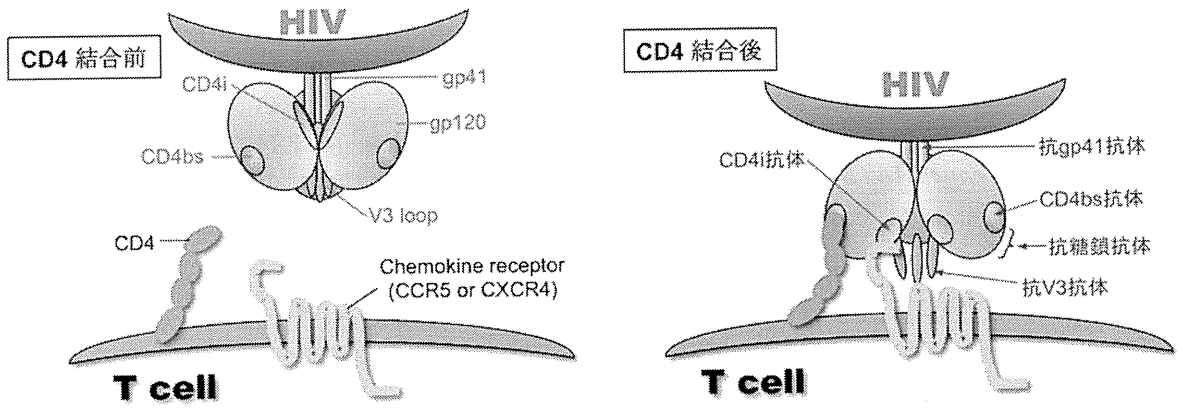


図 1 HIV エンベロープ蛋白と中和抗体反応部位

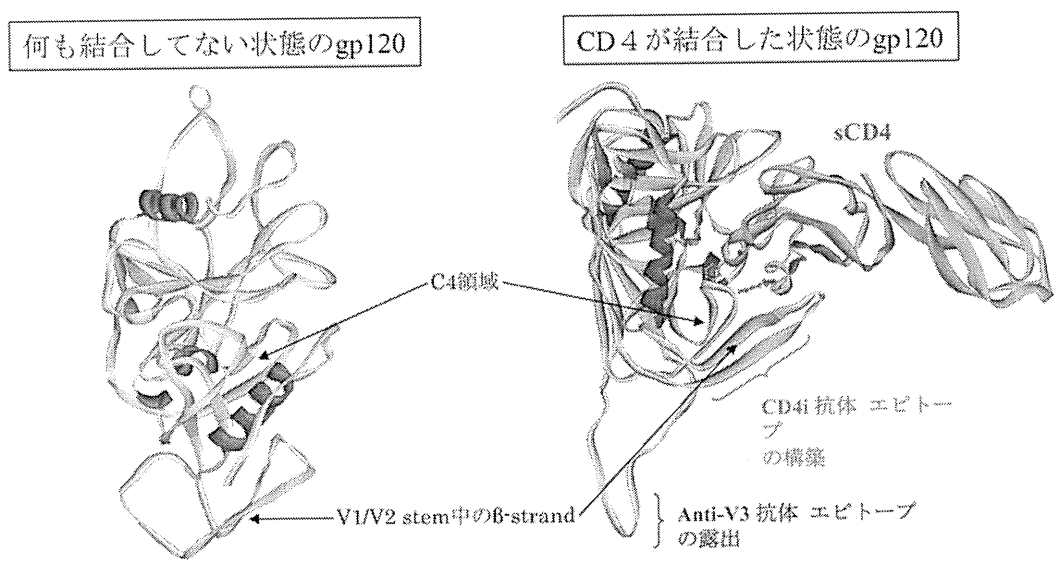


図 2 結晶構造解析による CD4 結合前後の gp120 の立体構造変化と中和抗体反応エピトープの構築および露出の模式図

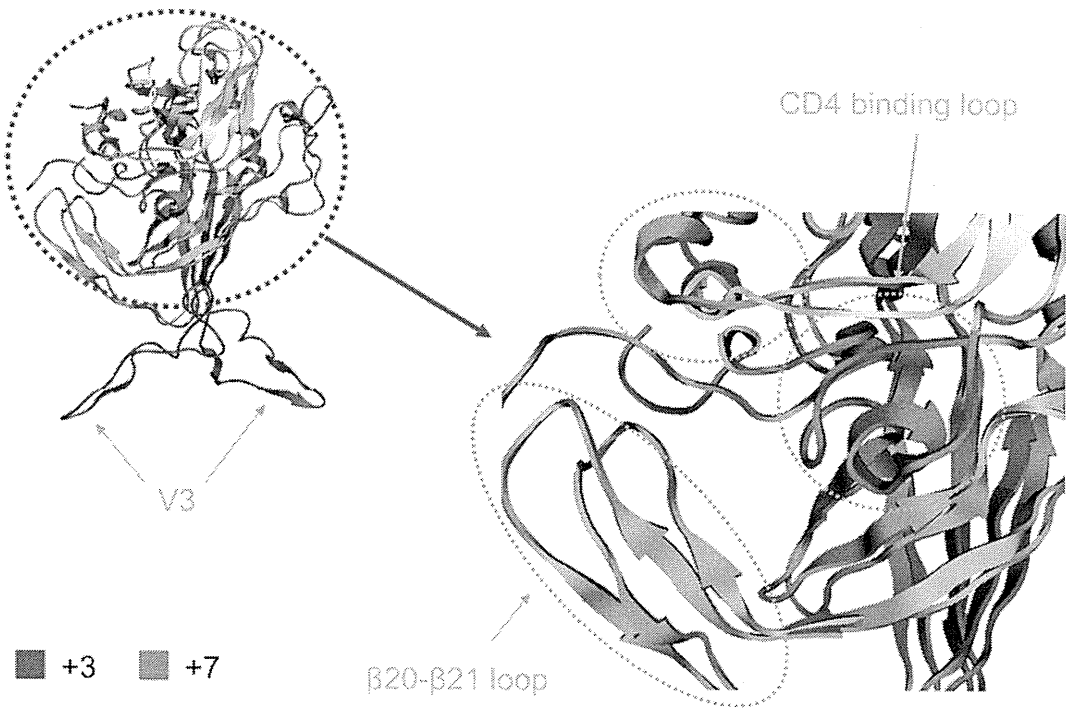


図 3 HIV-1 gp120 の V3 と CD4 結合ループの配置。
ホモロジーモデリング法および分子動力学計算により構築した HIV-1 gp120 分子モデル。V3 配列の荷電量が＋3，荷電量が＋7 の構造を重ね合わせている。リボン表示は gp120 を表し，V3 と CD4 結合ループのみ青または赤で表している。青は V3 配列の荷電量が＋3，赤は荷電量が＋7 である。

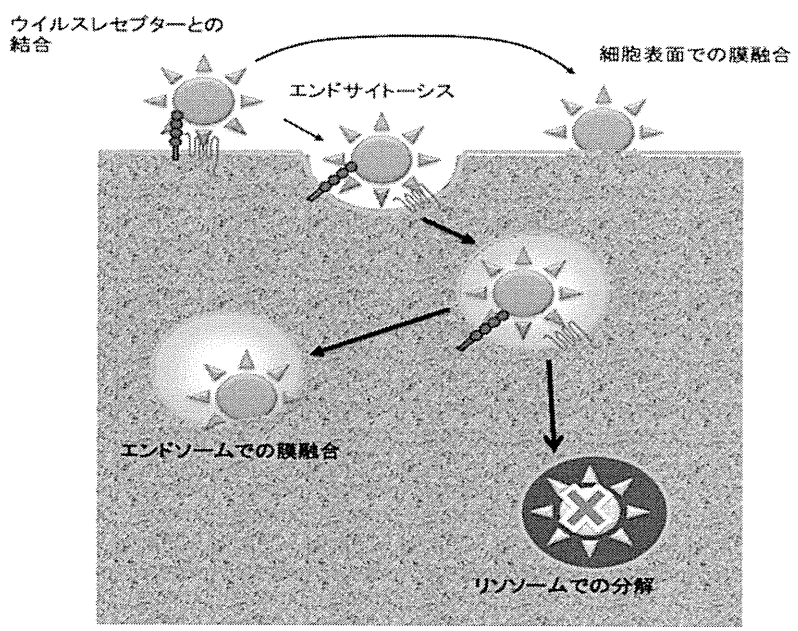


図 4 HIV はどこで膜融合するのか？

とから，初感染時には，むしろ感染性がよく，中和感受性も保たれているウイルスが増殖するという説がある。我々は，中和抗体に対する感受性の検討のため，最近の感染症例からウイルスを分離し，エンベロープ配列を調べているが，古典的なsubtype Bではなく，重感染によって起こるキメラ化によってエンベロープに進化が見られるウイルスが増えている可能性が観察された。これらのデータは，HIV-1エンベロープの進化は，CTLや中和抗体からのエスケープによるエピトープ変化の蓄積ばかりでなく，重感染によるキメラ化が重要な因子であることを示す。

SY4-2. 計算科学によるHIV-1 Gp120の構造解析

横山 勝（国立感染症研究所・病原体ゲノム解析研究センター）
Masaru Yokoyama (Pathogen Genomics Center, National Institute of Infectious Diseases)

HIV-1 g120のV3は感染受容体との相互作用に中心的役割を担う。そのため本来は機能的制約が強く作用し，アミノ酸変異は抑制されるはずである。ところがV3は高変異領域として知られる。これは，V3は免疫原性が高く，持続感染には抗原変異を必要とするため，とされる。本研究では，計算科学によるHIV-1 gp120の構造解析を行い，HIVの液性免疫逃避能を制御するgp120の構造要因を特定した。

これまである種の受容体結合特異性（CCR5指向性）を規定するV3配列は，感染者体内で比較的均一な集団として維持されることが知られており，他の細胞指向性V3に比べて塩基性アミノ酸が少ない。HIV sequence database (http : //www.hiv.lanl.gov/) より得たアミノ酸配列を用いてエントロピー解析を行うと，自然界で多く見られるV3配列は荷電量が＋3かつ糖鎖付加部位を持ち，荷電量の大きいV3配列に比べアミノ酸が保存されていた。この結果より，荷電量の低下によってV3配列の多様性が低下すると考えられる。次に，荷電量の抗V3抗体中和感受性への影響を知るために中和アッセイを行った。V3配列の荷電量が＋2，＋3，＋4かつ糖鎖付加部位を持つとき，ウイルスは抗V3抗体に低感受性であった。また，同じ荷電量でもCCR5指向性であるとき中和抵抗性であった。

ウイルスが抗V3抗体中和を逃避するメカニズムを知るため，HIV-1 gp120の分子モデルを構築しV3配列の荷電量のgp120構造への影響を調べた。分子モデルはV3配列の荷電量の影響を検討するため，V3配列のみが異なる2つの分子モデル（V3配列の荷電量が＋3と＋7）を，ホモロジーモデリング法および分子動力学計算により構築した。荷電量が＋3のgp120分子モデルと＋7のgp120分子モデルを比較した（図3）。荷電量が＋3ではV3先端部がβ20-β21ループとは反対方向に向き，荷電量が＋7ではV3先端部はβ20-β21ループの方向に向いた。したがって，gp120コアが同じアミノ酸配列であっても，V3配列の荷電量によりV3の立体配置が変化する。クライオ電子顕微鏡の像をもとにgp120三量体におけるV3の配置を検討すると，V3配列の荷電量の低下により中和エピトープの露出度が低下する可能性が示唆された。

次に，V3配列の荷電量がCD4結合ループの配置に及ぼす影響を調べた。V3配列の荷電量が＋7であるgp120のCD4結合ループは，＋3であるgp120よりもV4ループの方向にシフトしていた。そのため，V3配列の荷電量が＋7のgp120は＋3のgp120よりもCD4結合部位が広い。この結果より，V3配列の荷電量が大きいとCD4結合部位を認識する中和抗体による感受性が大きいと考えられる。そこで，gp120の多様性をShannonの式を用いたエントロピー解析により調べると，V3配列の荷電量が＋7では，荷電量が＋3と比べてCD4結合ループ周辺のアミノ酸がより多様である。これはCD4結合部位近傍が抗体の選択圧を受けていることを示唆し，構造解析の結果を支持する。荷電量が＋7では＋3よりも中和感受性であることを確かめるために中和アッセイを行うと，＋7のウイルスのCD4結合部位を認識する中和抗体の感受性は，＋3のウイルスよりも大きいことが分かった。

以上より，V3の荷電量によってgp120コアのアミノ酸配列が同じでも，V3配列の荷電量によってV3の配置が変わることが示された。また，V3配列の荷電量はV3の配置だけではなく，gp120コアの構造にも影響を与えた。したがって，HIV-1 gp120 V3は荷電量に基づきウイルスの中和感受性と細胞指向性を司る機能領域と考えられる。

SY4-3. エンドサイトーシスを介したHIV-1の細胞侵入

宮内 浩典（国立感染症研究所・エイズ研究センター/エイズ予防財団）
Kosuke Miyauchi (AIDS Research Center, National Institute of Infectious Diseases)

エンベロープウイルスは脂質二重層からなるウイルス膜を有し，宿主細胞の細胞質内に侵入するためにウイルス膜と細胞膜との膜融合の過程を必要とする。この膜融合はウイルス粒子の表面に存在するフュージョンタンパク質の構造変化によって誘導される。ウイルスフュージョンタンパク質の構造変化はウイルスによって異なった刺激により誘導され，この構造変化は一般的に不可逆的であることから，ウイルスにとってどのタイミングでこの構造変化をスタートさせるのかは感染を成立させるために重要なポイントとなる（図4）。インフルエンザウイルスのような低pH

依存性ウイルスはpHの低下をフュージョンタンパク質の構造変化のきっかけとしている。一方，ヒト免疫不全ウイルス（HIV）のような低pH非依存性ウイルスはウイルスレセプターとの結合を構造変化のきっかけとする。低pH依存性ウイルスはエンドソームの低pH環境を利用して膜融合するため，エンドサイトーシスはウイルス感染において必須である。HIVは低pH非依存性ウイルスであり，その感染成立にエンドソームの低pH環境を必要とせず，HIVのウイルスレセプターは細胞表面に存在するためHIVは細胞表面で膜融合を行うものと一般には考えられてきた。

我々はHIVの宿主細胞への侵入経路を明らかにするために，HIVの膜融合過程をベータラクタマーゼを用いたウイルス-細胞間膜融合アッセイと蛍光標識した単一ウイルスのトラッキングアッセイの二つの方法によって解析した。ベータラクタマーゼを用いたウイルス-細胞間膜融合アッセイでは，細胞膜不透過性の膜融合阻害ペプチドで得られた膜融合キネティクスと温度を低下させる方法で測定した膜融合キネティクスは大きく異なっており，HIVはエンドサイトーシスで細胞内に取り込まれた後，エンドソームで膜融合を完了することが示唆された。また蛍光標識した単一ウイルスのトラッキングアッセイにおいても，ほとんどのHIVがエンドソームで膜融合を起こしている事を示す結果が得られた。さらに，各種のエンドサイトーシス阻害剤を用いたところ，HIVの膜融合は著しく阻害され，HIVが膜融合の完了にエンドサイトーシスを必要とする事が明らかとなった。ペプチド型の膜融合阻害剤や一部の中和抗体はHIVのフュージョンタンパク質の構造中間体を認識するため，HIVがエンドソーム内で膜融合を完了することにより，これらの分子によるHIV阻害効果を減弱させる可能性が考えられる。

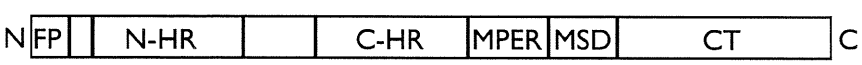
SY4-4. HIV-1エンベロープタンパク質gp41サブユニットの機能

松田 善衛（東京大学医科学研究所アジア感染症研究拠点/中国科学院生物物理研究所日中連携構造ウイルス学・免疫学研究室）

Zene Matsuda (Research Center for Asian Infectious Diseases, Institute of Medical Science, The University of Tokyo/China-Japan Joint Laboratory of Structural Virology and Immunology, Institute of Biophysics, Chinese Academy of Sciences)

HIV-1エンベロープタンパク質は前駆体gp160として生成され，宿主域を決定するgp120と膜融合を司るgp41のサブユニットに切断される。gp41はクラス1型の融合タンパク質に属し，N末に融合ペプチド（FP），その下流にコイルドコイル形成領域（N-及びC-HR）と広域中和抗体エピトープが存在するmembrane-proximal external region（MPER）を有する細胞外ドメインと，脂質二重膜を通過する膜貫通部分（MSD），それ以降の細胞内部分（CT）から構成されC末端に至る（図5）。近年膜融合機構の概略が明らかにされてきたが，関連分子の必要分子数，時間的，空間的変化の詳細はまだ明らかではない。

膜融合に際してはもともと脂質二重膜で隔てられた二つの画分（A, B）が，融合によってひとつの画分（A+B）となる。したがって融合に際して生じる二つの画分間の交通を測定することによって膜融合のモニタリングを行うことができる（図6）。最近われわれは膜融合の経時的解析を簡易に行える細胞ベースのアッセイ系を確立した。本法は膜融合のレポーターとしてウミシイタケルシフェラーゼ（RL）を分割酵素（split enzyme）として用いるものである。すなわち，それぞれ単独では活性を有しないRLの2分割体を作成し，それらを2種類の細胞（エンベロープ発現細胞（Env+）及び，受容体発現細胞（Receptor+））それぞれに別個に発現させておき，分割酵素の活性の復元をもって膜融合を測定するものである（図6）。RLには膜透過性の基質が存在するので，膜融合の過程を生きた細胞のままモニターすることができる。膜融合の際の移動分子の差異による既存の方法との比較を表1に示す。転写因子の移動によ



FP: Fusion Peptide
HR: Heptad Repeat
MPER:Membrane-Proximal External Region
MSD:Membrane-Spanning Domain
CT:Cytoplasmic Tail

図 5 gp41のドメイン構造

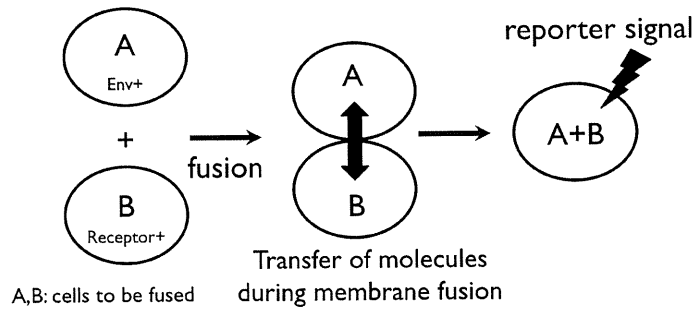


図 6　アッセイの概要　膜融合に際して AB 間で起こる分子の移動をマーカーとする

表 1　Comparison of different assays

Transferred molecules	Generation of signals	Detection Time	Quantification	Real-time assay
Transcription factor	only when fusion(+)	Slow	Easy	No
Dye	throughout	Quick	Hard	Yes
Split enzyme	only when fusion(+)	Quick	Easy	Yes

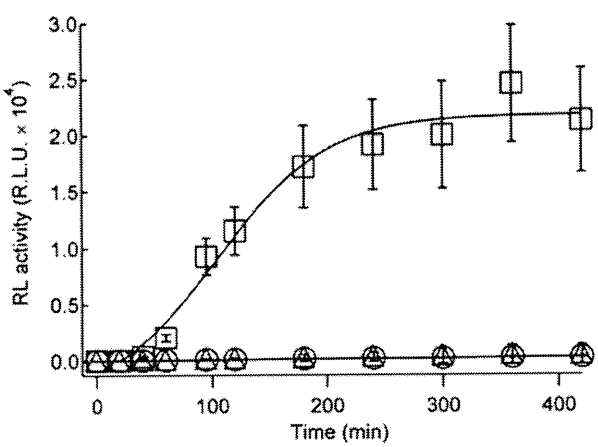


図 7　分割 RL による膜融合のモニタリング。(□は野生型,
○と△はコントロール)

るレポーター遺伝子の活性化による方法は定量的である一方，検出までの時間が長くリアルタイムアッセイには適さない。また色素の移動による方法は感度もよくリアルタイム測定も可能であるが，色シグナルの出現ではなく光学機器によるその位置変化を拠り所とするためアッセイの簡便さの面では劣っている。われわれの方法ではルシフェラーゼシグナルは融合後のみに発生し，迅速，簡便かつ定量的に測定可能である。野生型 HIV-1 エンベロープ蛋白質を用

いた実際の測定例を図7に示す。これは分割RLの一方とエンベロープ蛋白質とを発現する293FT細胞と，もう一方の分割RLを発現させた293CD4細胞の共培養の系においてRL活性の回復を膜透過性基質を用いて測定したものである。

われわれはこの方法を用いてgp41膜貫通部分変異体の解析を行った。膜貫通部分は約20アミノ酸残基からなり各クレード間でも非常によく保存されている。この配列の高い保存性に反して，それぞれのアミノ酸の置換変異に対する許容性は極めて高い。しかしその許容度には明らかに限界があり幾つかの膜貫通部分の異種膜貫通部分による完全置換体においては膜融合活性の著しい低下が見られる。われわれの以前の解析ではCD4結合過程以降に障害があることが示唆されていたが，そのメカニズムは不明であった。今回これらの変異体による膜融合をこのアッセイ系で経時的に測定したところ，野生型に比べて膜融合の進行が大きく遅れていることが示唆された（Kondo *et al.*, J. Biol. Chem., 285 : 14681-14688, 2010）。

今回開発された膜融合アッセイ系はHIV-1のエンベロープ蛋白質の指向性の決定や，変異体の膜融合の解析に有用であると考えられる。膜融合の判断が簡便に且つ定量的に測定できることから，膜融合阻害剤のスクリーニングにも応用可能であると考えられる。さらにこのシステムはHIV-1以外のエンベロープ蛋白質による膜融合のアッセイや，ウイルスエンベロープ以外による膜融合過程の解析にも応用可能である。

Antiviral Research 90 (2011) 33–41

Contents lists available at ScienceDirect

Antiviral Research

journal homepage: www.elsevier.com/locate/antiviral

Within-host co-evolution of Gag P453L and protease D30N/N88D demonstrates virological advantage in a highly protease inhibitor-exposed HIV-1 case

Junko Shibata[a,b,1], Wataru Sugiura[b,c,d,*], Hirotaka Ode[e], Yasumasa Iwatani[b,c,d], Hironori Sato[e], Hsinyi Tsang[b], Masakazu Matsuda[d,f], Naoki Hasegawa[a], Fengrong Ren[a], Hiroshi Tanaka[a]

[a] School of Biomedical Sciences, Tokyo Medical and Dental University, Tokyo, Japan
[b] Clinical Research Center, National Hospital Organization, Nagoya Medical Center, Nagoya, Japan
[c] Nagoya University Graduate School of Medicine, Nagoya, Japan
[d] AIDS Research Center, National Institute of Infectious Diseases, Tokyo, Japan
[e] Pathogen Genomics Center, National Institute of Infectious Diseases, Tokyo, Japan
[f] Mitsubishi Chemical Medience Corporation, Tokyo, Japan

ARTICLE INFO

Article history:
Received 7 September 2010
Received in revised form
28 December 2010
Accepted 11 February 2011
Available online 19 February 2011

Keywords:
HIV
Protease
Gag
Drug resistance
Co-evolution

ABSTRACT

To better understand the mechanism of HIV group-specific antigen (Gag) and protease (PR) co-evolution in drug-resistance acquisition, we analyzed a drug-resistance case by both bioinformatics and virological methods. We especially considered the quality of sequence data and analytical accuracy by introducing single-genome sequencing (SGS) and Spidermonkey/Bayesian graphical models (BGM) analysis, respectively. We analyzed 129 HIV-1 Gag–PR linkage sequences obtained from 8 time points, and the resulting sequences were applied to the Spidermonkey co-evolution analysis program, which identified ten mutation pairs as significantly co-evolving. Among these, we focused on associations between Gag-P453L, the P5′ position of the p1/p6 cleavage-site mutation, and PR-D30N/N88D nelfinavir-resistant mutations, and attempted to clarify their virological significance *in vitro* by constructing recombinant clones. The results showed that $P453L^{Gag}$ has the potential to improve replication capacity and the Gag processing efficiency of viruses with $D30N^{PR}/N88D^{PR}$ but has little effect on nelfinavir susceptibility. Homology modeling analysis suggested that hydrogen bonds between the 30th PR residue and the $R452^{Gag}$ are disturbed by the $D30N^{PR}$ mutation, but the impaired interaction is compensated by $P453L^{Gag}$ generating new hydrophobic interactions. Furthermore, database analysis indicated that the $P453L^{Gag}/D30N^{PR}/N88D^{PR}$ association was not specific only to our clinical case, but was common among AIDS patients.

1. Introduction

Major mutations in the human immunodeficiency virus-1 (HIV-1) protease (PR)-coding region selected by protease inhibitors (PIs) are mainly located within the active sites of the PR, and these mutations significantly reduce PR activity and viral replication capacity, i.e., viral fitness, compared to that of wild-type strains (Mahalingam et al., 1999). However, PI-resistant viruses have the potential to undergo further selection and evolution to recover their impaired PR activity and viral fitness by acquiring additional mutations not only in the PR but also in its natural substrate, the Gag protein (Nijhuis et al., 1999). In particular, mutations in Gag cleavage sites can improve the replication capacity of PI-resistant viruses. Indeed, tight associations have been demonstrated between PI-resistant mutations and Gag cleavage-site mutations, such as $S373Q^{Gag}$ (Malet et al., 2007) and $I376V^{Gag}$ (Ho et al., 2008) at p2/NC, $A431V^{Gag}$ (Bally et al., 2000; Doyon et al., 1996; Gallego et al., 2003; Koch et al., 2001; Maguire et al., 2002; Malet et al., 2007; Verheyen et al., 2006; Zhang et al., 1997) at NC/p1, $L449F^{Gag}$ (Doyon et al., 1996), and $P453L^{Gag}$ (Bally et al., 2000; Maguire et al., 2002; Verheyen et al., 2006) at p1/p6. Gag mutations at non-cleavage sites have also been reported to improve the fitness of PI-resistant viruses (Myint et al., 2004). Thus, the selection and evolution of Gag and PR are accepted to significantly interfere with each other. This phenomenon is known as "Gag–PR co-evolution."

However, previous reports of Gag–PR co-evolution appear to have two technical limitations related to data quality and analytical method. First, the standard population-based genotyping method commonly used for determining HIV-1 variants has limited accuracy for technological reasons. To sequence the viral genome,

* Corresponding author at: Clinical Research Center, National Hospital Organization, Nagoya Medical Center, 4-1-1, San-no-maru, 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).
[1] Current address: Inserm U941, IUH, Université Paris Diderot, Hôpital Saint Louis, 75475 Paris Cedex 10, France.

0166-3542/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.antiviral.2011.02.004

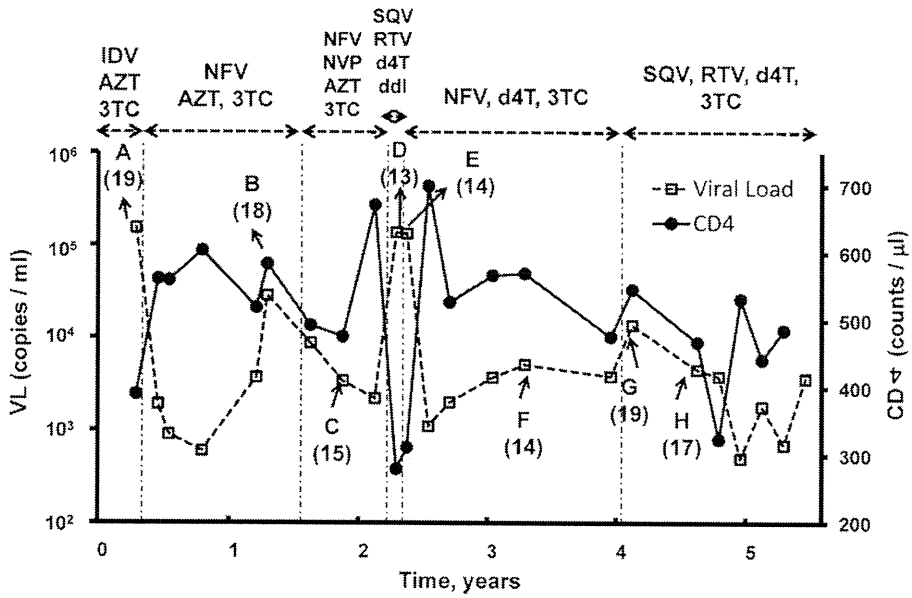


Fig. 1. Clinical course of patient-treatment protocols. Open squares and solid circles indicate plasma viral load and CD4+ cell count, respectively. Each treatment regimen is indicated in the upper part of the graph by horizontal dashed lines between arrows. Sample collection points are labelled A to H, and numbers of sequences analyzed at each point are shown in parentheses. AZT, zidovudine; d4T, stavudine; ddI, 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′-CTCTCTCGACGCAGGACTCG-

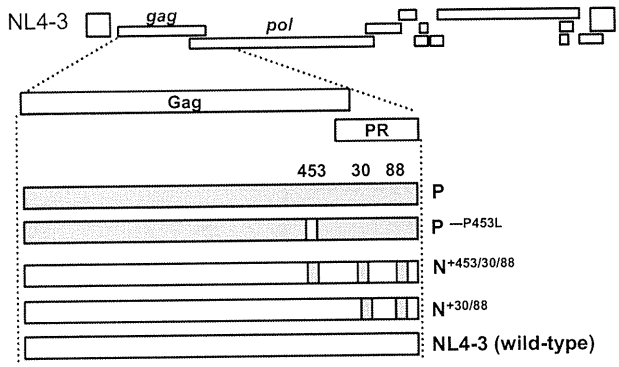


Fig. 2. Construction of recombinant viruses. Four different types of recombinant viruses were constructed according to their mutation patterns: (1) P, NL4-3 backbone with patient-derived Gag and protease (PR); (2) P^{-453}, NL4-3 backbone with patient-derived Gag and PR but $P453L^{Gag}$ was converted 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′-CTATTAAGTATTTTGATGGGTCATAA-3′), nested PCR-WGPF2 (5′-TTGCTGAAGCGCGCACGGCAAGA-3′) and 3410- (5′-CAGTTAGTGGTATTACTTCTGTTAGTGCTT-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 $pCR4\text{-}TOPO^{Patient}$). Using $pCR4\text{-}TOPO^{Patient}$ 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′-ACTTTTGGGCCATCCATTCCTGGCTTT-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 2×10^5 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 (5×10^6 ^{32}P 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 ApaI-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 $pCR4\text{-}TOPO/NL^{ApaI\text{-}SbfI}$. Using $pCR4\text{-}TOPO/NL^{ApaI\text{-}SbfI}$ 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′-GAGTCAACAGATCTCTTCCAATTATG-3′. After each mutagenesis reaction, the entire sequences were verified and cloned back into pNL4-3 using ApaI and SbfI restriction enzymes. All pNL4-3-based recombinant DNAs (3.75 μg) were transfected into 2×10^5 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 (4×10^5 ^{32}P 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, 5×10^4 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 $CCID_{50}$ 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) × 100. IC_{50} 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), I72T(7), V77I (100),
D (n = 13)	S373P (85), A374V (85)	P453L (8)	D30N (8), N88D (8)	L10F(8), V11I (8), T12A(8), K20R(8), E35D(8), M36I(100), H69Y(8), V77I(8)
E (n = 14)	S373P (64), A374V(71)	–	–	T12A(7), M36I(100), K55N (7), V77I(7)
F (n = 14)	S373Q(100), A374G(100)	P453L(100)	D30N(100), N88D(100)	L10F(100), I13V(79), E34G(7), E35D(100), M36I (100), N37T (86), K45R (14), Q58E (86), I72T (7), V77I (100)
G (n = 19)	S373Q(100), A374G(100)	P453L(100)	D30N(100), N88D(100)	L10F(100), I13V (68), E35D(100), M36I(68), M36V(32), N37T(100), Q58E(100), V77I(100)
H (n = 17)	S373Q(100), A374G(100)	P453L(100)	D30N(100), I54V(76), N88D(100), L90M (47)	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 IC_{50}s 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 IC_{50} between the two recombinants was

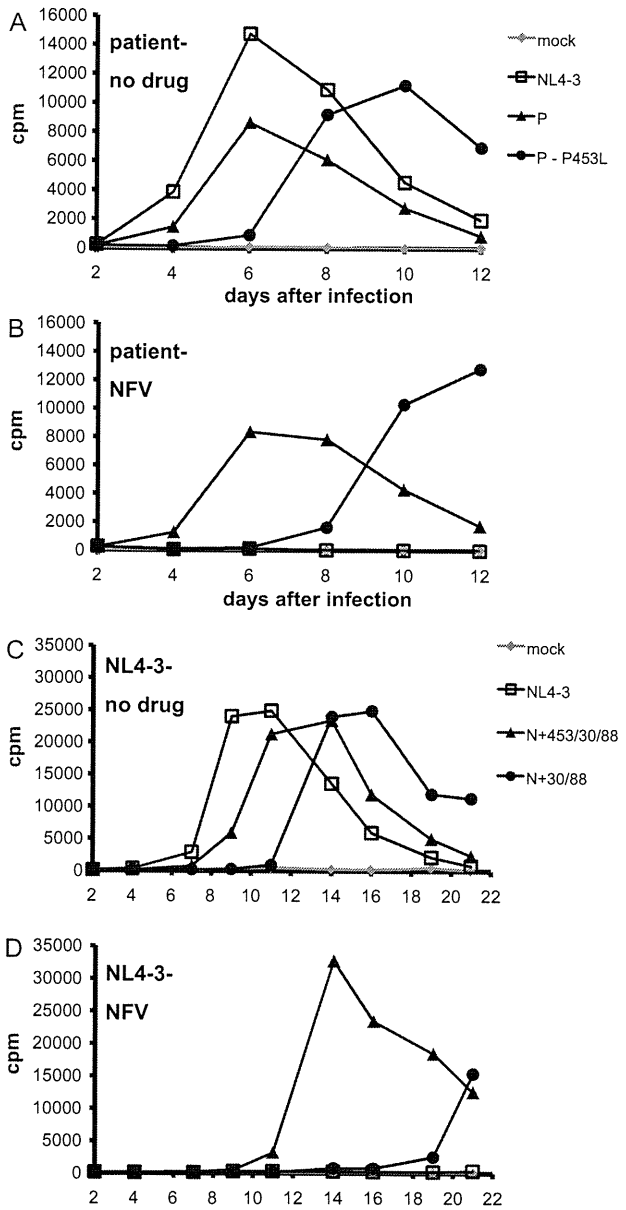


Fig. 3. Replication kinetics of recombinants. MT-2 cells were infected with patient-derived *gag*-PR-coding region, (A) in the absence of, and (B) in the presence of 0.1 μM nelfinavir. Open squares, solid triangles, and solid circles indicate wild-type NL4-3, NL4-3 with patient *gag*-PR-coding region 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

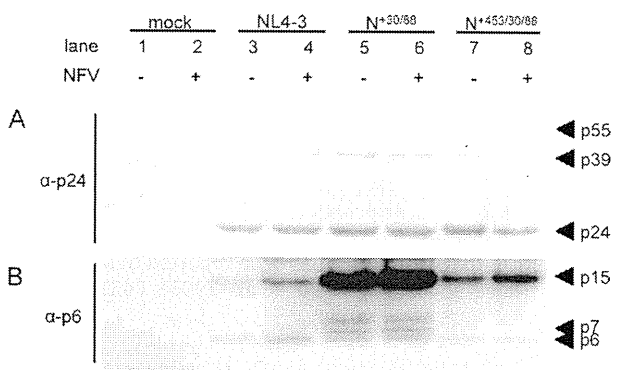


Fig. 4. Western blot analysis of Gag processing in the absence or presence of nelfinavir. Western blot analyses in the absence or presence of nelfinavir. HeLa cells were transfected by each recombinant clones and cultured in the absence or presence of NFV (0.1 μM). At 48 h post-transfection, virions 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$^{+453/30/88}$, lanes 7 and 8, N$^{+30/88}$.

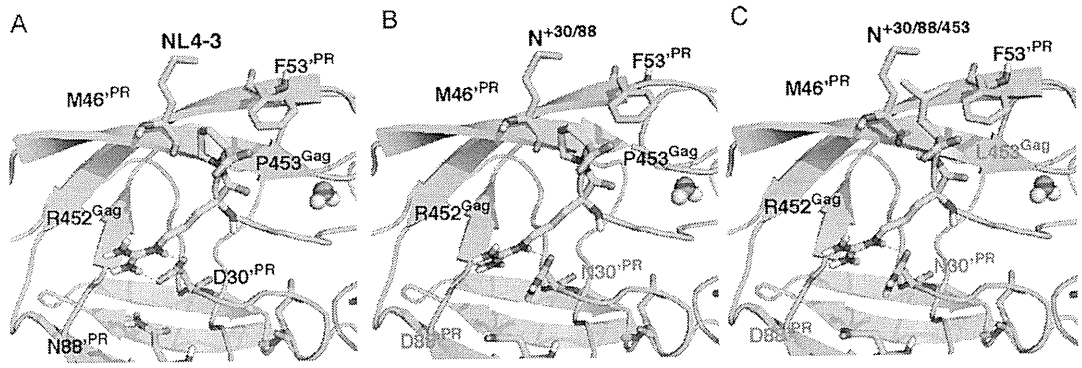


Fig. 5. Structural models of the Gag–PR complexes. Catalytic sites of the NL4-3 (A), $N^{+30/88}$ models (B), and $N^{+453/30/88}$ models (C) are highlighted. Green cartoons and sticks represent main and side chains of PR, respectively. Cyan sticks represent 452nd and 453rd residues in Gag corresponding 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/N88D 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., Lastere, 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., Dalban, C., Wirden, 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., Io, 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-naive 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.
Willey, 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.