

を決定できたのは、これまでに 61 名中 54 名である。それらを感染経路別に示したものが、表 2 である。血液製剤による感染例では 17 例全てが、男性同性間性的接触による感染でも 24 例全てがサブタイプ B であった。異性間性的接触による感染でも、9 件中 5 件がサブタイプ B であった。また、サブタイプ AE の 2 例は、どちらも異性間性的接触による感染例であった。また、日本人では、その他に異性間性的接触による感染に 1 例、サブタイプ C の例があった。

外国人では、サブタイプ A が 1 例、C/A、C/E(gag/env)の様なりコンビナントが疑われる例が各 1 例見られた。

初診年別にサブタイプの分布を表 3 に示す。1900 年から 2007 年の期間を 3 時期に分けて比較したが、1995 年から 2000 年の時期に 2 例のサブタイプ AE を含む非 B 型が 4 例見られた以外は、サブタイプ B が主流であることに大きな変化は見られなかった。

D. 考察

1990 年以降の HIV-1 感染者 54 名について、サブタイプを決定することができた。その結果、48 名(89%)がサブタイプ B に感染していることがわかった。サブタイプ AE 感染例は異性間性的接触による 2 例だけであった。関東地区ではサブタイプ AE が同様な経路で急増しているが、福岡県におい

ては大きな増加は認められていない。これまでに、外国人においては非 B 型が見られているが、今後も新たなサブタイプが国内に持ち込まれる可能性があり、さらに分子疫学的なサーベイランスが必要である。

E. 結論

福岡県内では、1990 年代から現在に至るまで、サブタイプ B による感染が主流であり、いわゆるサブタイプ AE の増加は見られていない。また外国人においては、非 B 型の感染例が 3 例見られた。

参考文献

- 1)Ou,C.-Y., Takebe,Y., Weniger,B.G., Luo,c.-c., Kalish,M.L., Auwanit,W., Yamazaki,S., Gayle,H.D., Young,N.L., and Schochetman,G., (1993). Independent introduction of two major HIV-1 genotypes into distinct high-risk populations in Thailand. Lancet 341 : 1171-1174.
- 2)Holmes,E.C., Zhang,L.Q., Robertson,P., Cleland,A., Harvey,E., Simmonds,P., and Brown,A.L., (1995). The molecular Epidemiology of human immunodeficiency virus type 1 in Edinburgh. J.Infect.Dis. 171 : 45-53.

表1.対象者の初診年、及び推定感染経路

初診年	血液製剤	男性同性間	異性間	薬物濫用	不明	合計
1990-1994	12	9	1	0	1	23
1995-2000	5	10	5	1	1	22
2001-2007	1	10	4	0	1	16
合計	18	29	10	1	3	61

表2. 感染経路別サブタイプ(qaq/env)

感染経路	B/B	B/ND*	ND/B	A/E	A/A	C/C	C/A	C/E
血液製剤	11	4	2	0	0	0	0	0
男性同性間	12	8	4	0	0	0	0	0
異性間	3	1	1	2	1**	1	0	0
薬物濫用	1**	0	0	0	0	0	0	0
不明	0	0	1	0	0	0	1**	1**
合計	27	13	8	2	1	1	1	1

* not determined

**foreigner

表3. 初診年別サブタイプ (qaq/env)

初診年	B/B	B/ND*	ND/B	A/E	A/A	C/C	C/A	C/E
1990-1994	17	3	1	0	0	0	1	0
1995-2000	10	3	3	2	1	0	0	1
2001-2007	0	7	4	0	0	1	0	0
合計	27	13	8	2	1	1	1	1

* not determined

北海道における HIV 感染の現状とサブタイプ

分担研究者 澤田幸治（北海道立衛生研究所）

協力研究者 長野秀樹、地主勝、伊木繁雄（北海道立衛生研究所微生物部）
工藤伸一（北海道立衛生研究所生物科学部）

研究概要

日本における HIV 感染者数の増加傾向は継続しているが、北海道においても、エイズ患者とあわせた新規 HIV 感染者については、1998 年以降は 10 名前後で推移していたが、2005 年に 20 名を超え、2006 年では 28 名であった。2006 年の全感染者数のうちエイズ発症者は 11 名で 39%であった。また、感染経路別では、同性間性的接触が 14 名で 50%であったが、HIV 感染者だけでは 65%であった。また、年齢についてはエイズ患者、HIV 感染者ともに 30 代が多く、全体の 68%であった。一方、入手可能な材料について HIV の塩基配列によるサブタイプについて調べたところ、全例サブタイプ B であった。

A. 研究目的

我が国における HIV 感染者・エイズ患者の総数は 1 万人を超え、また、年間の感染者・患者報告数が 1,000 人を上回るなど、HIV 感染症の拡大傾向がますます加速されつつある。北海道においても近年、増加傾向が続いている。行政的には平成 16 年 4 月から、道立保健所の無料匿名 HIV 検査において、迅速検査法を導入することによって、検査当日に結果が判明する即日検査システムを構築した。これによって HIV 抗体検査受検者の利便性が向上し、年間の受検者数は増加傾向を示している。しかしながら、感染の拡大を阻止するには至っていない。本研究は、エイズ流行の形成メカニズムおよびそれに関連する諸要因を理解し、それによりエイズ流行の予防制圧に向けた研究の推進を目的としている。なお本研究は、「北海道立衛生研究所ヒトを対象とする医学研究に関する規定」

に基づき北海道立衛生研究所倫理審査委員会の承諾を得て実施している。

B. 研究方法

1. 北海道におけるエイズ流行の現状把握

感染症法に基づく発生届けによるデータについて集計解析した。

2. HIV のサブタイプ分析

HIV 感染者の血清からウイルス RNA を分離精製し、鋳型とした。*env* 遺伝子内の C2/V3 領域、プロテアーゼ領域、逆転写酵素領域の 5'末端側とそれよりも 3'末端の 4 領域について RT-PCR 法を用いて増幅した。PCR 産物を精製し、塩基配列を決定した。当該塩基配列について、既知のサブタイプ HIV の代表的な株との比較においてサブタイプを決定した。得られた塩基配列のデータについて ClustalX を用いてアライメントをかけ、MEGA 3 の近隣接合法を用い、系統樹を

作成した。系統樹の信頼性評価のためにブートストラップ値を1,000とした。

C. 研究結果

1. 北海道におけるエイズ流行の現状把握

図1に北海道における HIV 感染者・エイズ患者数の年次推移を示した。1997年までは10人以下で推移していたが、1998年に10名を超えてからは数年間10名前後であった。しかし、2005年には20名を超過し、2006年は28名であった。このように、2005、2006年では急増しており、2006年では2004年の2倍の感染者患者数であった。エイズ患者の割合では、2003年は若干高く66.7% (8/12)であったが、ここ3年(2004年から2006年)では30.0%から46.7%の間で推移し、50%を超えることはなかった。しかし、エイズを発症してから見つかる患者数には減少傾向は見られず、潜在的な HIV 感染者の存在を裏付けるものと思われる。

北海道における HIV 感染者・エイズ患者を感染経路別にみると、性的接触による感染が多かった。HIV 感染者・エイズ患者をあわせると、2006年では同性間性的接触が50%であったが、HIV 感染者では64.7%であった(図2)。また、年齢別では、HIV 感染者・エイズ患者ともに30代が多かった。2004年以降は20代にもエイズ患者が認められ、感染の低年齢化が北海道においても進んでいることが示された(図3)。さらに、届け出保健所を基に発生状況を地域別にみると、2004年、2006年では都市部での届け出件数がほぼ80%であった(図4)。

2. HIV のサブタイプ分析

2006年に当研究所で得られた試料は28例中3例であったので、この試料についてサブタイプ分析を行った。近年、HIVの組み換え型が多数報告されていることから、*env* 領域(C2/V3)のみならず、*pol* 遺伝子についても解析を行った。C2/V3 領域では系統樹解析からサブタイプBであることがわかった(図5)。さらに、*pol* 遺伝子のプロテアーゼ領域、逆転写領域の5‘末端側、及びその下流領域の3領域について系統樹解析を実施したところ、3領域すべてサブタイプBであった(図6-8)。このように今回の3試料については、解析した領域において矛盾はみられず、組み換え型ではないことが示された。

E. 考察

北海道では、HIV 感染者・エイズ患者数が2005年に20名を超え、2006年には30名近くになっている。2004年4月から迅速検査を保健所に導入し、受検者の利便性の向上を図っており、受検者数が増加傾向にあるなど一定の成果は得られている。また、感染経路、年齢構成などのエイズ流行の諸要素を分析することにより、より効果的なアプローチが必要になる。

サブタイプ分析においては、今回の試料はすべてBであったが、全体の1割のみの知見であることから、北海道全体のサブタイプを反映しているとは言い難い。また技術的にもアウトグループの設定などの問題もあり、今後の検討課題である。

F. 結論

1. 北海道では、2005 年以降、HIV 感染者・エイズ患者数が増え、2006 年には28名であった、
2. 感染経路別では同性間性的接触が多く、50%を占めた。
3. 年齢構成では30代が多かった。
4. 地域別では都市部が全体の約8割であった。
5. サブタイプでは3試料ともBであった。

図1 新規HIV/AIDS感染患者の年次推移
(北海道)

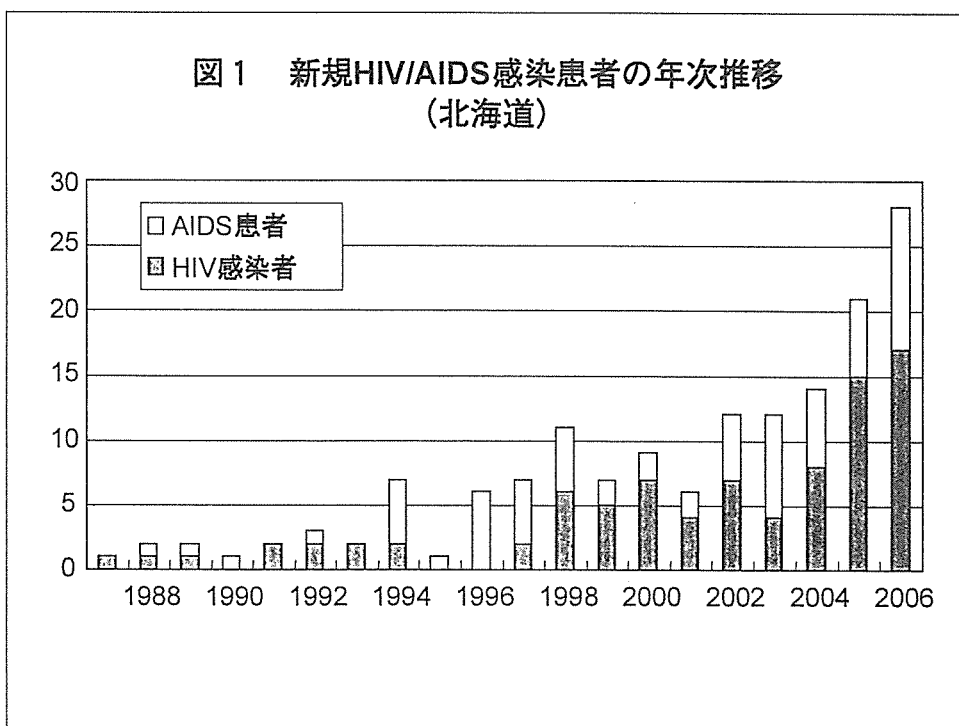


図2 HIV/AIDS感染患者の感染経路

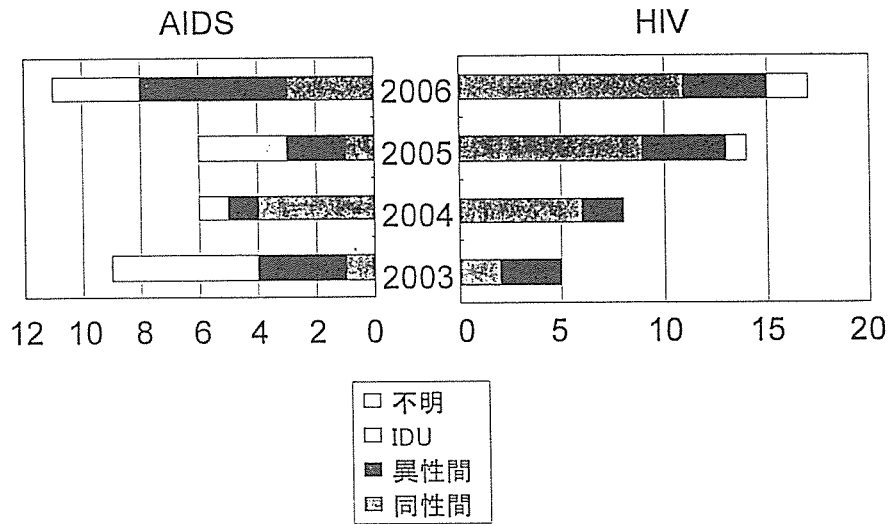


図3 HIV/AIDS感染患者の年齢構成

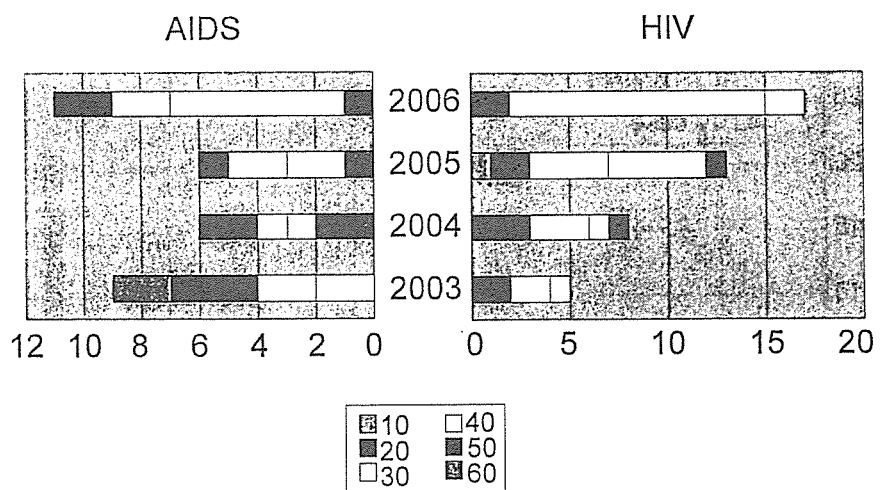


図4 HIV感染の地域性

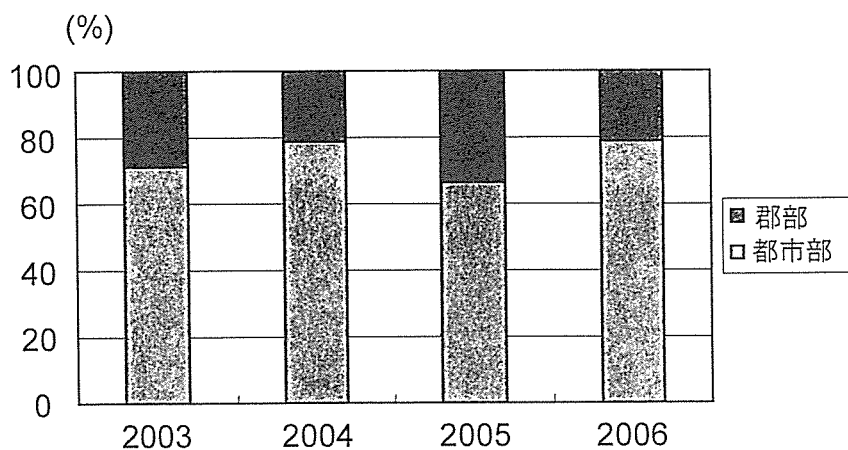
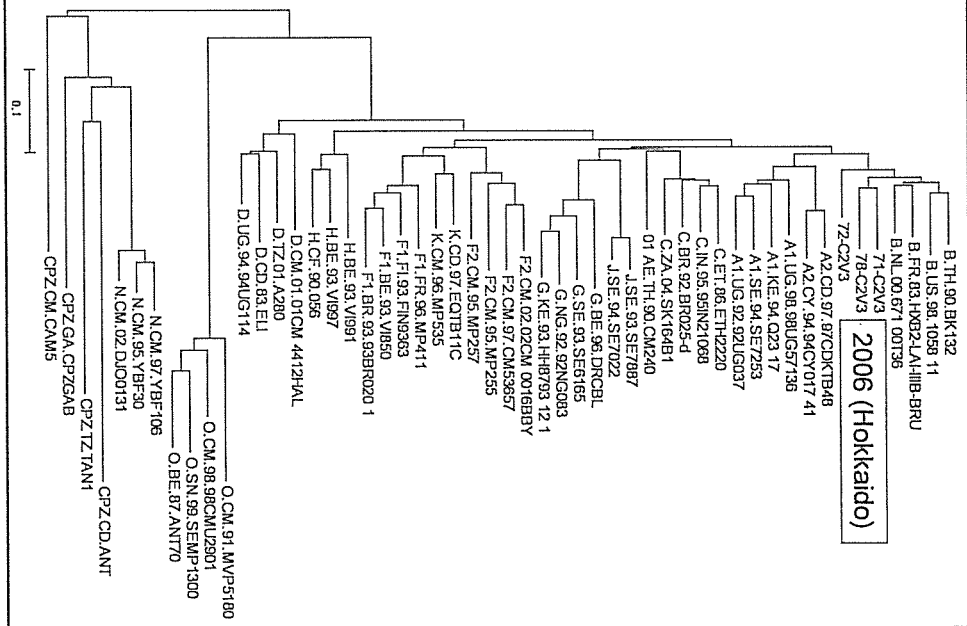


図5 Neighbor-joining法による系統樹解析 env (C2/V3)



平成 18 年度厚生労働省科学研究費補助金

分担研究報告書

長野県における HIV 陽性者の診断契機と免疫不全進行度

分担研究者 齊藤博 長野県立須坂病院 院長

協力研究者

四本美保子 長野赤十字病院 副部長

高山義浩 佐久総合病院 科長

山崎善隆 信州大学医学部付属病院 助手

北野喜良 国立病院機構松本病院 副院長

田中征雄 国立病院機構長野病院

小平宏 諏訪赤十字病院 部長

柳川宗平 飯田市立病院 医長

研究要旨

長野県下 8 エイズ拠点病院に受診した 125 例の HIV 感染者につき診療録を検討した。日本人が 78.4%、そのうちの男性患者は 88.8%であった。年齢別では 40 歳代が 31.2%で最も多く、20 歳代は 11.2%であった。感染者の 58.8%は既にエイズを発症し、死亡率は CD₄ 数 200/uL 以下の症例に限られていた。感染経路の 84%は異性間性的接触によるものであった。感染が偶然に判明したきっかけは術前検査や妊婦検査が多かった。既往歴からは性病の既往や帯状疱疹含めた難治性皮膚疾患が多いことも分かり、一般診療の中で早期発見の手がかりになるものと期待される。

A. 研究目的

増え続ける長野県の HIV 感染者対策に必要な臨床的、疫学的指標を探ることを目的とした。

B. 研究方法

対象は 2001 年 1 月から 2005 年 8 月までに長野県下 8 拠点病院を受診した 125 例。診療録から年齢、性別、国籍、住居地、初診時の CD₄ 値、病期、診断のきっかけとなった症状、既往歴を調査検討した。

(倫理面への配慮)

データは全て、病院毎の番号で登録され個人の判定はできない仕組みになっている。

C. 研究結果

対象症例 125 例中、年齢は 20 歳代 11.2%、30 歳代 28.8%、40 歳代 31.2%、50 歳代 20.0%、60 歳代 7.2%であった。20 歳代、30 歳代の若年層が 40%、一方 40 歳代以上の中年世代が併せて 51.2%であった。国籍別では 78.4%が日本人でそのうちの 88.8%が男性であり、男女比は 8 : 1 であった、外国人の男女比は 1 : 2 と女性が優位であった。感染経路別では、84%が異性間性的接触で同性間性的接触は 8%であった。診断時の病気は、エイズ発症が 58.4%であった。CD₄陽性 T 細胞数の中央値は 70.5/uL であった。エイズ診断後の死亡率は 11.2%であったが、死亡した症例は全例が CD₄ 値 200/uL 以下の症例であった。HIV 陽性と判明したきっかけは「無症状」では「パートナー HIV 陽性」が最も多く、次いで術前検査、妊婦検診の順であった。「症状あり」ではカリニー肺炎が最も多く、口腔カンジダ症、体重減少、CMV 感染症の順であった。既往歴では帯状疱疹、B 型肝炎含めた性感染症、帯状疱疹以外の皮膚疾患の順であった。

D. 考察

長野県の HIV 感染者・エイズ患者数は長野県衛生部報告によると、2005 年末現在累積数で 352 名にのぼる。2004 年の 100 万人あたりの人口比では、新規 HIV 陽性者 17.6 人と全国 2 位、新規 AIDS 患者は 9.47 人と全国 1 位の状態である。長野県内の 8 拠点病院の症例 125 例の検討から、日本人男性が 69.9%を占めており、年齢的には 40 歳以上が 58.4%であった。20 歳代は 11.2%であったが長野県衛生部発表の 20 歳代感染者届け出報告数は 38.6%と、実診療患者数との間に乖離がみられた。この原因としては長野県では 20 歳代の感染者は外国人女性が主体で、彼らの拠点病院での定期受診率が低いことを示しているものと思われる。病期ではエイズでの初診が 58.4%と極めて高く、全国平均の 33.9%を遥かに上回った、所謂「いきなりエイズ」の症例が多いことが明らかになった。また初診時 CD₄ 値が極めて低い事実はエイズ症例の中でも重症で見つかる事を示している。このことは早期発見のための検査カウンセリング体制の不備、医師の HIV 感染症に対する知識不足、更には県民意識の低さに起因するものと

思われる。感染経路で異性間性的接触が主体であることは、わが国農村部の標準的傾向と思われる。「無症状」で HIV 陽性と判明したきっかけのうち、「パートナーHIV 陽性」が最も多いが、術前検査と妊婦検診を併せた所謂「偶然の診断」例は 42 例中 19 例（46.3%）と半数近くを占め、術前検査や検診の重要性も再認識された。既往歴でみられた帯状疱疹や B 型肝炎含めた性感染症、慢性難治性皮膚疾患は、HIV 感染症の診断の目安の 1 つとなるものと思われた。

E. 結論

長野県の拠点病院に受診した HIV 感染者の分析から、感染者の 7 割は日本人男性であり、多くがエイズを発症して受診していた。長野県においては、HIV エイズ対策は地域住民の性感染症対策の 1 つとして、より積極的な介入が必要である事を示唆している。感染者の早期発見のためには、性感染の危険が高いと思われるハイリスク集団を疫学的に特定し、エイズを発症する前に治療へ結び付ける方策を確立する戦略作りが必要である。他方、一般住民に対しては、HIV 検査の機会を増やすための環境作りを進めるべきである。一般医師に向けての対応として、HIV 陽性者に見られた既往歴を手がかりに、一般診療の場で早期診断につながる臨床症状を、教育・啓発していくことが重要である。

G. 研究発表

1. 論文発表

四本美保子、北野喜良、斉藤博：長野県における HIV 陽性者の診断契機と免疫不全進行度. 信州医学雑誌 54(4) : 183~187,2006

Ⅲ. 研究成果の刊行に関する一覧表

平成18年度 研究成果の刊行に関する一覧表

発表論文リスト (2006-2007 年度)

主任研究者：武部 豊

分担研究者：駒野 淳・花房秀次・草川 茂・椎野禎一郎・近藤真規子・加藤真吾・貞升健志・小島洋子・千々和勝己・澤田幸治

協力研究員：白阪琢磨・上遠野賢之助・齋藤 博・四本美保子・高山義浩

論文発表

1. Naito, Y., Ui-Tei, K., Nishikawa, T., Takebe, Y., Saigo, K. (2006). siVirus: web-based antiviral siRNA design software for highly divergent viral sequences. Nucl. Acid Res. (Web Server issue) W448-W450.

2. Takebe, Y. and Telesnitsky, A. (2006). Evidence for the acquisition of multidrug resistance by an HIV clinical isolate via human gene transduction. Virology 351: 1-6.

3. Murakami, Y., Yamagoe, S. Noguchi, K., Takebe, Y., Uehara, Y. and Fukazawa, H. (2006). Ets-1-dependent expression of vascular endothelial growth factor receptors is activated by Latency-associated nuclear antigen of Kaposi's sarcoma-associated herpesvirus through interaction with Daxx. J. Biol. Chem. 281(38): 28113-28121.

4. Tee, K. K., Li, X.-J., Nohtomi, K., Ng, K. P., Kamarulzaman, A., and Takebe, Y. (2006). Identification of a novel circulating recombinant form (CRF33_01B) disseminating widely among various risk populations in Kuala Lumpur, Malaysia. J. AIDS 43(5): 523-9.

5. Futahashi Y, *Komano J, Urano E, Aoki T, Hamatake M, Miyauchi K, Yoshida T, Koyanagi Y, Matsuda Z, Yamamoto N.(2006). Separate elements are required for ligand-dependent and -independent internalization of metastatic potentiator CXCR4. Cancer Science. (In press)

6. Shimizu S, Urano E, Futahashi Y, Miyauchi K, Isogai M, Matsuda Z, Nohtomi K, Onogi T, Takebe Y, Yamamoto N, Komano J. (2006). Inhibiting lentiviral replication by HEXIM1, a cellular inhibitor of cdk9/cyclinT complex. AIDS. (In press)

7. Miyauchi K, *Komano J, Myint L, Futahashi Y, Urano E, Matsuda Z, Chiba T, Miura H, Sugiura W, Yamamoto N. (2006). Rapid propagation of low-fitness drug resistant mutants of human immunodeficiency virus type 1 by a streptococcal metabolite sparsomycin. Antivir Chem Chemother. 17(4): 167-174.

8. Miyauchi K, Curran R, Matthews E, Komano J, Hoshino T, Engelman DM, Matsuda Z. (2006). Mutations of conserved glycine residues within the membrane-spanning

domain of human immunodeficiency virus type 1 gp41 can inhibit membrane fusion and incorporation of Env onto virions. *Jpn J Infect Dis.* Apr;59(2):77-84.

9. Kato, S., Hanabusa, H., Kaneko, S., Takakuwa, K., Suzuki, M., Kuji, N., Jinno, M., Tanaka, R., Kojima, K., Iwashita, M., Yoshimura, Y., and Tanaka, K. (2006). 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* 20(7):967-973.

10. 須藤弘二, 嶋貴子, 近藤真規子, 加藤真吾, 今井光信. (2007). Real-time PCR を用いた HIV-1 RNA 測定キットの基礎的検討. *感染症学雑誌* 81(1), 1-5.

11. Hamatake, M., Nishizawa, M., Yamamoto, N., Kato, S., and Sugiura, W. A simple competitive RT-PCR assay for quantitation of HIV-1 subtype B and non-B RNA in plasma. *J. Virol. Methods* (in press)

12. 宇宿秀三、野口有三、坂本光男、足立拓也、相楽裕子、須藤弘二、西澤雅子、近藤真規子、柄久保修、今井光信. (2006). Analysis of a long term discrepancy in drug-targeted genes in plasma HIV-1 RNA and PBMC HIV-1 DNA in the same patient. *Jpn. J.Infect.Dis.*, 59: 122-125.

13. 嶋貴子、一色ミユキ、近藤真規子、塚田三夫、潮見重毅、今井光信. (2006). 保健所における HIV 即日検査導入の試みとその効果. *日本公衆衛生雑誌*、53, 167-177.

14. 長島真美, 貞升健志, 新開敬行, 秋場哲哉, 吉田勲, 吉田靖子, 矢野一好, 甲斐明美, 諸角聖. (2005). 東京都における HIV 検査成績 (1999 年-2004 年) 東京都健康安全研究センター年報, 56. 41-44.

15. 川畑拓也、小島洋子、森治代、大竹徹、大國剛. (2007). 当所にて HIV 感染を確認した、2 例のイムノクロマトグラフィー法陰性の感染初期例、*感染症学雑誌*、80:76-77.

16. 四本美保子、北野喜良、斎藤博. (2006). 長野県における HIV 陽性者の診断契機と免疫不全進行度. *信州医学雑誌* 54(4) : 183~187.

IV. 研究成果の刊行物・別刷

siVirus: web-based antiviral siRNA design software for highly divergent viral sequences

Yuki Naito^{1,*}, Kumiko Ui-Tei^{1,2}, Toru Nishikawa³, Yutaka Takebe⁴ and Kaoru Saigo¹

¹Department of Biophysics and Biochemistry, Graduate School of Science, ²Undergraduate Program for Bioinformatics and Systems Biology, School of Science and, ³Department of Computer Science, Graduate School of Information Science and Technology, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan and ⁴Laboratory of Molecular Virology and Epidemiology, AIDS Research Center, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan

Received February 14, 2006; Revised and Accepted March 24, 2006

ABSTRACT

siVirus (<http://siVirus.RNAi.jp/>) is a web-based online software system that provides efficient short interfering RNA (siRNA) design for antiviral RNA interference (RNAi). siVirus searches for functional, off-target minimized siRNAs targeting highly conserved regions of divergent viral sequences. These siRNAs are expected to resist viral mutational escape, since their highly conserved targets likely contain structurally/functionally constrained elements. siVirus will be a useful tool for designing optimal siRNAs targeting highly divergent pathogens, including human immunodeficiency virus (HIV), hepatitis C virus (HCV), influenza virus and SARS coronavirus, all of which pose enormous threats to global human health.

500 entries of near full-length sequences of HIV-1 group M, which is largely responsible for global pandemic, are stored in the sequence databases, but it proved impossible to select a common 21mer from among all of them. Moreover, RNAi-resistant viral mutants achieved through point mutation or deletion emerge rapidly when targeting viruses in cell culture. These problems suggest a strong need to select highly conserved target sites for designing antiviral siRNAs. Furthermore, the off-target silencing effects of siRNA are also a serious problem that could affect host gene expression (9). Off-target silencing effects arise when an siRNA has sequence similarities with unrelated genes. In antiviral RNAi, it is desirable to minimize off-target effects against human genes.

Consequently, only a limited fraction of 21mers is suitable for use as antiviral siRNAs. In this study, we developed a novel web-based online software system, siVirus, which provides functional, off-target minimized siRNAs targeting highly conserved regions of divergent viral sequences.

INTRODUCTION

RNA interference (RNAi) is now widely used to knockdown gene expression in a sequence-specific manner, making it a powerful tool not only for studying gene function, but also for therapeutic purposes, including antiviral treatments (1–4). Currently, the replication of a wide range of viruses can be inhibited successfully using RNAi, with both short interfering RNAs (siRNAs) and siRNA expression vectors (5).

In mammalian RNAi, the efficacy of each siRNA varies widely depending on its sequence; only a limited fraction of randomly designed siRNAs is highly effective. Many experiments have been conducted to clarify possible sequence requirements of functional siRNAs. Of these, our work incorporates guidelines from three major studies (6–8) of selecting functional siRNAs. However, designing functional siRNAs that target viral sequences is problematic because of their extraordinarily high genetic diversity. For example, about

METHODS

Selection of highly conserved siRNA target sites

Highly conserved siRNA sequences are selected based on their *degree of conservation*, defined as the proportion of viral sequences that are targeted by the corresponding siRNA, with complete matches (i.e. 21/21 matches). All possible siRNA candidates targeting every other position of user-selected viral sequences are generated and their degrees of conservation are computed. Users can arbitrarily specify a set of viral sequences for the computation; e.g. sequences can be selected from a specific geographic region(s) or a specific genotype(s) to design the best siRNAs tailored to specific user needs. siVirus also accepts user's own sequences in a multi-FASTA format and shows whether each siRNA can target the posted sequences.

*To whom correspondence should be addressed. Tel: +81 3 5841 4404; Fax: +81 3 5841 4400; Email: y-naito@RNAi.jp

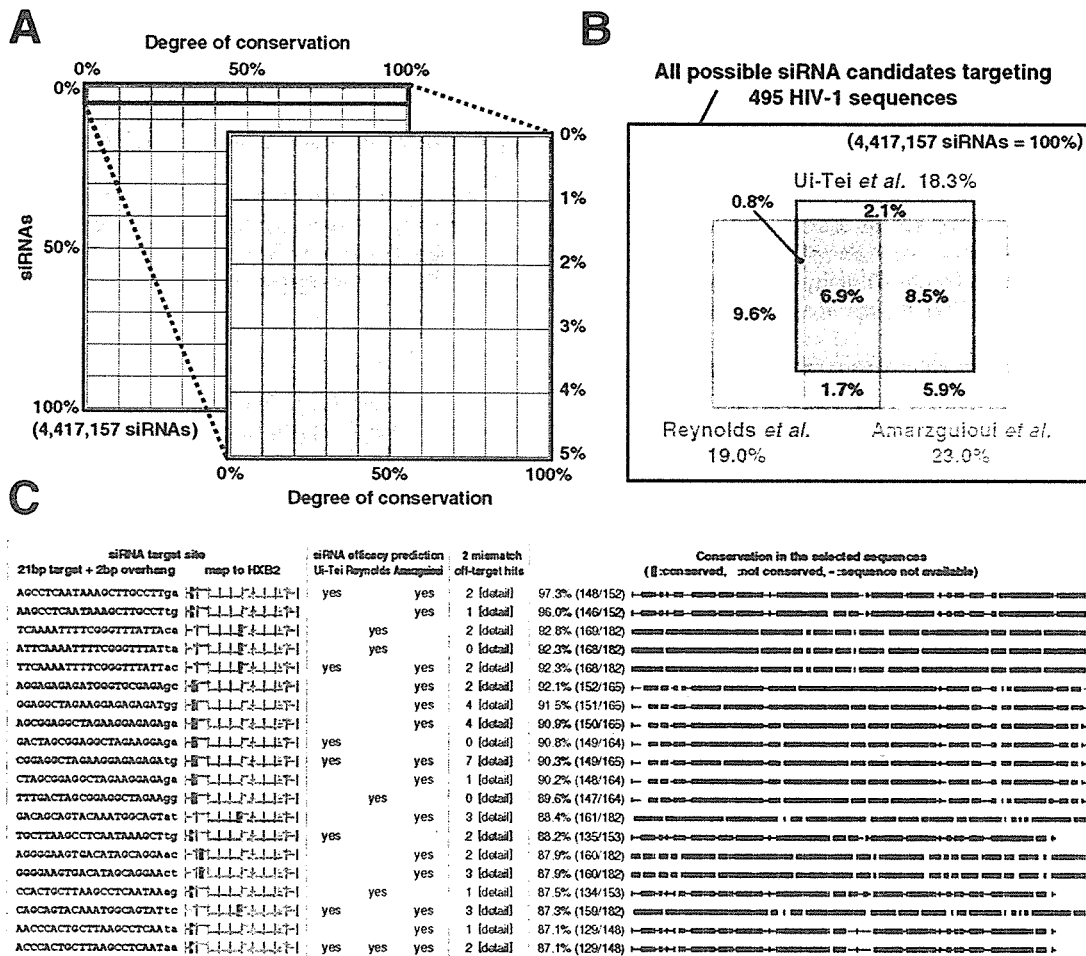


Figure 1. (A) The degree of conservation is calculated for all possible siRNA candidates (total 4 417 157) targeting every other position of 495 HIV-1 sequences. (B) The efficacy predictions of these 4 417 157 siRNA candidates based on three different guidelines: Ui-Tei *et al.* (6), Reynolds *et al.* (7) and Amarzguioui *et al.* (8). (C) Typical output of siVirus for designing anti-HIV siRNAs. Sequence information, efficacy predictions, off-target search results and the degrees of conservation are shown.

siRNA efficacy prediction

In mammalian RNAi, the efficacy of each siRNA varies markedly depending on its sequence; hence, several groups have reported guidelines for selecting functional siRNAs. siVirus incorporates the guidelines of Ui-Tei *et al.* (6), Reynolds *et al.* (7) and Amarzguioui *et al.* (8) and shows whether each siRNA satisfies these guidelines.

Off-target searches

Off-target searches were performed for each siRNA using siDirect (10,11). siVirus shows the number of off-target hits within two mismatches against the non-redundant database of human transcripts (10).

Database maintenance

Currently, siVirus incorporates viral genome sequences of HIV-1, HCV, influenza A virus and SARS coronavirus.

These sequences were downloaded from the Los Alamos HIV Sequence Database (<http://hiv-web.lanl.gov/>), the Los Alamos HCV Sequence Database (12), the NCBI Influenza Virus Sequence Database (<http://www.ncbi.nlm.nih.gov/genomes/FLU/FLU.html>), and NCBI GenBank (13), respectively. siVirus will be updated continuously as these databases are revised. We also plan to incorporate other viruses if sufficient numbers of their sequences are available.

RESULTS AND DISCUSSION

To design anti-HIV siRNA, we analyzed the 495 near full-length HIV-1 sequences listed in Supplementary Table 1. A total of 4 417 157 possible siRNA candidates (i.e. substrings of length 21) targeting every other position of the HIV-1 sequences were produced from the 495 viral sequences. The analysis of these siRNA candidates revealed that highly

conserved siRNAs constituted only 0.3% of the possible siRNAs if >90% conservation is expected (Figure 1A). The fraction is still as small as 0.8% even if the threshold of the conservation is relaxed to 80%. On the other hand, siRNAs predicted to be functional by one or more guidelines (6–8) constituted 35.5% of the 4 417 157 siRNAs (Figure 1B). Taken together, siRNAs that are >80% conserved, and satisfy at least one guideline constitute only 0.2% of the siRNAs. In this condition, 20–30 siRNAs can be designed for each full-length sequence of HIV-1. These indicate that most of the randomly designed siRNAs are not suited for targeting HIV-1 efficiently.

Figure 1C shows typical output from siVirus for designing anti-HIV siRNAs. A total of 182 sequences from HIV-1 subtypes B, C and CRF01_AE, which are the most prevalent HIV-1 genotypes circulating in Asia, were selected. The results were sorted by their degree of conservation, and filtered to display siRNAs that satisfy at least one efficacy guideline. The off-target search results against human genes are also shown. It is desirable to select an siRNA that has less off-target hits.

To test the validity of siVirus, 35 siRNAs satisfying the guideline by Ui-Tei *et al.* (6) were designed against the conserved regions of HIV-1 genomes using siVirus and were assayed for inhibition of viral replication. Among them, 31 siRNAs effectively inhibited HIV-1 replication by >80% when each siRNA duplex was transfected at 5 nM (Y. Naito, K. Ui-Tei, K. Saigo and Y. Takebe, unpublished data).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

This work was supported in part by grants from the Ministry of Education, Culture, Sports, Science and Technology of Japan to K.S., K.U.-T. and Y.T., and by grants from the Ministry of

Health, Labour and Welfare of Japan to Y.T. Funding to pay the Open Access publication charges for this article was provided by the Ministry of Education, Culture, Sports, Science and Technology of Japan. Y.N. is a Research Fellow of the Japan Society for the Promotion of Science.

Conflict of interest statement. None declared.

REFERENCES

1. Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E. and Mello, C.C. (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*, **391**, 806–811.
2. Mello, C.C. and Conte, D. Jr (2004) Revealing the world of RNA interference. *Nature*, **431**, 338–342.
3. Hannon, G.J. and Rossi, J.J. (2004) Unlocking the potential of the human genome with RNA interference. *Nature*, **431**, 371–378.
4. Voinnet, O. (2005) Induction and suppression of RNA silencing: insights from viral infections. *Nature Rev. Genet.*, **6**, 206–220.
5. Leonard, J.N. and Schaffer, D.V. (2006) Antiviral RNAi therapy: emerging approaches for hitting a moving target. *Gene Ther.*, **13**, 532–540.
6. Ui-Tei, K., Naito, Y., Takahashi, F., Haraguchi, T., Ohki-Hamazaki, H., Juni, A., Ueda, R. and Saigo, K. (2004) Guidelines for the selection of highly effective siRNA sequences for mammalian and chick RNA interference. *Nucleic Acids Res.*, **32**, 936–948.
7. Reynolds, A., Leake, D., Boese, Q., Scaringe, S., Marshall, W.S. and Khvorovova, A. (2004) Rational siRNA design for RNA interference. *Nat. Biotechnol.*, **22**, 326–330.
8. Amarzguioui, M. and Prydz, H. (2004) An algorithm for selection of functional siRNA sequences. *Biochem. Biophys. Res. Commun.*, **316**, 1050–1058.
9. Jackson, A.L. and Linsley, P.S. (2004) Noise amidst the silence: off-target effects of siRNAs? *Trends Genet.*, **20**, 521–524.
10. Naito, Y., Yamada, T., Ui-Tei, K., Morishita, S. and Saigo, K. (2004) siDirect: highly effective, target-specific siRNA design software for mammalian RNA interference. *Nucleic Acids Res.*, **32**, W124–W129.
11. Yamada, T. and Morishita, S. (2005) Accelerated off-target search algorithm for siRNA. *Bioinformatics*, **21**, 1316–1324.
12. Kuiken, C., Yusim, K., Boykin, L. and Richardson, R. (2005) The Los Alamos hepatitis C sequence database. *Bioinformatics*, **21**, 379–384.
13. Benson, D.A., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J. and Wheeler, D.L. (2006) GenBank. *Nucleic Acids Res.*, **34**, D16–D20.



Rapid Communication

Evidence for the acquisition of multi-drug resistance in an HIV-1 clinical isolate via human sequence transduction

Yutaka Takebe^a, Alice Telesnitsky^{b,*}^a *Laboratory of Molecular Virology and Epidemiology, AIDS Research Center, National Institute of Infectious Diseases, Tokyo 162-8640, Japan*^b *Department of Microbiology and Immunology, University of Michigan Medical School, Ann Arbor, MI 48109-0620, USA*

Received 29 March 2006; returned to author for revision 3 May 2006; accepted 9 May 2006

Available online 14 June 2006

Abstract

Insertions in HIV-1 reverse transcriptase's fingers subdomain can enhance chain terminator excision and confer resistance to multiple nucleoside analogs. Inserts that resemble flanking sequences likely arise by local sequence duplication. However, a remarkable variety of non-repeat fingers insertions have been observed. Here, molecular epidemiology, sequence analyses and mechanistic modeling were employed to show that one Japanese isolate's RT fingers insert likely resulted from non-homologous recombination between virus and host sequences and the transductive copying of 37 nucleotides from human chromosome 17. These findings provide evidence that human sequence transduction can, at least rarely, contribute to genetic and phenotypic variation in pandemic HIV.

© 2006 Elsevier Inc. All rights reserved.

Keywords: Genetic recombination; Retroviruses; Reverse transcriptase; Drug resistance**Introduction**

Much of HIV-1's genetic variation arises by stepwise – albeit at times hypermutation-accelerated – accumulation of point mutations (Harris et al., 2003; Leitner and Albert, 1999). Genetic recombination also contributes to HIV genetic diversity and occurs about 10-fold more frequently than base substitution (An and Telesnitsky, 2002; Jetzt et al., 2000). Because fewer molecular events are required to introduce panels of mutations by recombination than by serial mutation, clustered genome changes are generally believed to reflect recombination or related rearrangements (Malim and Emerman, 2001; Wain-Hobson et al., 2003). Retroviral recombination results from template switching during reverse transcription and generally occurs in regions of high sequence similarity between the two intact genomes each retrovirus co-packages (An and Telesnitsky, 2002). Non-homologous recombination, often guided by microhomology between donor and acceptor templates, can

also occur (Hajjar and Linial, 1993; Zhang and Temin, 1993). Deletions and insertions can arise via non-homologous recombination between discontinuous portions of the viral genome (Parthasarathi et al., 1995; Temin, 1993). Duplications result if template switch occurs from one RNA position to a locus further downstream on the co-packaged RNA, while deletions arise when reverse transcriptase “jumping” bypasses sequences and terminates upstream of the point of template departure (Parthasarathi et al., 1995). Insertion-in-a-deletion or insertion-in-a-duplication mutations can result from a series of non-homologous crossovers (Lobato et al., 2002; Parthasarathi et al., 1995; Pathak and Temin, 1990). Either virus or host sequences can template insertions, as postulated by models for oncogene transduction (Muriaux and Rein, 2003). Although whole gene transduction is rare, incorporation of short host segments into defective viral genomes is observed fairly frequently (Dunn et al., 1992; Fang and Pincus, 1995; Hajjar and Linial, 1993; Mikkelsen and Pedersen, 2000; Pulsinelli and Temin, 1991; Sun et al., 2001). Thus, experimentally, the use of host sequences to bridge non-homologous recombination junctions and the insertion of host segments at strong stop or non-homologous crossover sites is a well-established

* Corresponding author. Fax: +1 734 764 3562.

E-mail addresses: takebe@nih.gov (Y. Takebe), ateles@umich.edu (A. Telesnitsky).

phenomenon. Possible contributions of host sequences to HIV clinical isolates has been less clear: both because pre- and post-mutation sequence data are seldom available and because inserts maintained in replication-competent virus are generally too short to implicate a specific template (Winters and Merigan, 2005).

Results

In this report, we examined an unusually lengthy insertion mutation in an HIV-1 isolate from a Japanese child, with the goal of elucidating its likely origins (Sato et al., 2001). Epidemiologic investigation provided sequence information for the viral populations observed in the patient, referred to as NH3, during the 6 years prior to and at the onset of clinical drug resistance. These included the multi-drug-resistant isolate itself (designated 99JP-NH3-II) and several co-circulating insertion region variants (Fig. 1A). Additional sequence

information was available for isolates from NH3's father, NH1, who had been infected in Thailand, and from his mother, NH2, who had contracted HIV-1 from NH1 and subsequently infected NH3 through maternal transmission. This clinical history confirmed that the insertion arose during highly active antiretroviral treatment in NH3 (Sato et al., 2001). A comparison of virus with and without the insert demonstrated that the 11-amino-acid insertion in the reverse transcriptase of this CRF001_AE circulating recombinant form variant contributed to its high-level resistance to multiple nucleoside analogs.

The region encompassing the 33 nucleotides inserted in NH3's RT gene was evaluated for similarity to experimentally described retroviral recombination products. The insertion did not mirror flanking sequences and thus was not a simple duplication (Parthasarathi et al., 1995). The entire sequence of the 99JP-NH3-II genome was available and none of it resembled the 33-nucleotide insert. Thus, sequential template

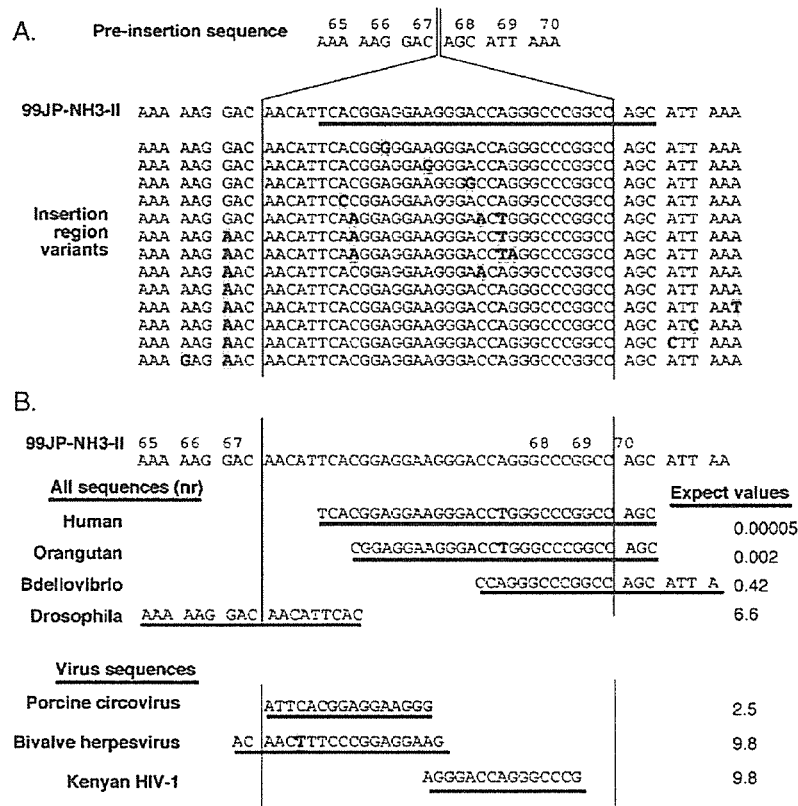


Fig. 1. Sequence analyses of the 99JP-NH3-II insertion region. (A) Deduced preinsertion sequence, sequence insertion in 99JP-NH3-II, and insertion region variants. The NH3 isolates shown are GenBank accession nos. AB053002, AB053074, AB053071, AB053005, AB053038, AB053073, AB053065, AB053035, AB053084, AB053003, AB053077, AB053068, AB053058, and AB053083. Numbers at top represent amino acid residue numbers within reverse transcriptase. Line under 99JP-NH3-II is region of greatest human match. Shaded sequences are positions that differ from 99JP-NH3-II. (B) Comparisons of the 99JP-NH3-II insertion region to other sequences in GenBank. Underlined sequences indicate greatest match to the 51-base insertion region of 99JP-NH3-II. At the top ("All sequences"): blastn results from nr GenBank. The human match encompasses an intron/exon junction in AC135178 (Chromosome 17 open reading frame 68; also found in NC_00017); the orangutan match lies within CR857769; *Bdellovibrio*: BX842654, *Drosophila*: AY183918. Expect values at left are those assigned by blastn, as described in Materials and Methods. Note that although human and chimpanzee sequences are highly related in the 3' regions of human chromosome 17 orf 68, the orf 68 5' region differs markedly between human and chimpanzee, and thus, no match between the 99JP-NH3-II insert and GenBank chimpanzee sequences is detectable. At the bottom of panel B: blastn results from nr GenBank limited to viruses, after manual removal of NH3 sequences. Accession number for the circovirus: AY321985; for herpesvirus: AY509253; for Kenyan HIV-1: AF457073.

switches among viral segments could not account for NH3's RT insert.

Support for possible non-viral origins of the 33-base insertion came from its striking nucleotide composition. Whereas typical HIV sequences have a <40% G + C content (Berkhout et al., 2002), this insert was 67% G + C. Because abnormal G + C content is a hallmark of horizontal gene transfer (Hacker and Kaper, 2000), we explored possible human origins for the insert. If microhomology guided template switch had contributed to insert generation, viral sequences flanking the insert would be predicted to retain human sequence homology. To accommodate this, a 51-base segment comprised of the 33-base insert plus nine nucleotides from either flank was used in the database searches summarized in Fig. 1B.

When this 51-base segment was used to query unrestricted databases in GenBank, the best match was to 99JP-NH3-II itself, followed closely by the insert variants isolated from NH3 (Fig. 1A). The similarity of these NH3 isolates to one another and the differences between their inserts and all other HIV sequences (see below) makes it reasonable to assume that a single insertion event gave rise to all insert variants in patient NH3: an interpretation also supported by our previous studies on viral molecular evolution in NH3 (Sato et al., 2001).

Excluding NH3's HIV strains, the closest GenBank match to this 51-base sequence was a 30/31-base match to a non-repetitive sequence on human chromosome 17 (Fig. 1B). This was followed by a 27/28 match to the orthologous locus in orangutan, with the remaining top matches including a 20/20 match to the genome of the predatory bacterium *Bdellovibrio*, followed by less extensive matches to sequences from *Drosophila*, mouse, dog, and rice, but notably not from any non-NH3 HIV-1 isolates.

To assess similarity of the 99JP-NH3-II insert to sequences in other HIV isolates, the 51-base segment was subsequently BLASTed against all virus entries in GenBank ('Virus sequences' in Fig. 1B). Besides 99JP-NH3-II and the co-circulating isolates from NH3, no viral sequences in GenBank yielded expect values of <1. Because expect values are measures of probability roughly equivalent to *P* values, the absence of <1 matches in any isolate of any type of virus suggests that the insert in 99JP-NH3-II is even less similar to any known virus sequence than would be predicted by random chance. Whereas BLAST had assigned the 30/31 human genome match an expect score of 0.00005, the only HIV match assigned an expect score <10 was a 15/15-base match (expect score 6.6) to a portion of *env* in a clinical isolate from Kenya (Fig. 1B).

That the virus-human match included a few bases downstream of the insertion was consistent with the possibility that microhomology guided recombination between HIV-1 and the BLAST-identified human sequence generated this insertion. These observations and the mutation's structure suggested that the NH3 insertion was generated via the splinted non-homologous recombination model shown in Fig. 2. Briefly, an HIV-1 provirus was established on chromosome 17 just upstream of the putative insert-encoding sequences (Fig. 2A). Viral polyadenylation signal read-through generated a chimeric

HIV-human RNA that became encapsidated. During subsequent reverse transcription, microhomology-guided template switching between portions of the RT gene and human sequences on the read-through RNA generated the observed insertion-in-a-duplication structure (Fig. 2B).

This model is based on experimental outcomes of non-homologous recombination, models for retroviral transduction, and properties of the putative human bridging template (An and Telesnitsky, 2002; Mikkelsen and Pedersen, 2000; Muriaux and Rein, 2003). Specifically, the putative human template straddles an intron/exon junction that lies in the antisense orientation of a mapped and verified mRNA encoding a putative protein of unknown function (Gao et al., 2005; Ota et al., 2004). Thus, although recombination between viral and unlinked host sequences has been reported previously (Sun et al., 2001), for the unspliced antisense sequences in this case, host sequences were more likely to have become encapsidated on a read-through transcript than as a free RNA (Muriaux and Rein, 2003).

Because it differs at three nucleotide positions from the founder strain predicted by our model (Fig. 2C), the postulated recombination events alone cannot explain the JP-NH3-II RT insertion mutation. However, the spectra of insert variants isolated from NH3 differed from one another at up to 4 positions within the examined sequence interval (Fig. 1A). Thus, these variants' sequence heterogeneity demonstrates that the extent of viral diversification required to generate JP-NH3-II's insert from the putative founder strain indisputably did occur within patient NH3 after initial insert acquisition. This supports the notion that JP-NH3-II arose via the mechanism outlined in Figs. 2A and B, followed by the introduction of point mutations at the positions boxed in Fig. 2C.

Discussion

The findings here analyzed a drug resistance-associated sequence insertion that is not closely related to sequences in any other HIV-1 isolate in GenBank. The closest match to this insert among all sequences in GenBank was to a portion of human chromosome 17. The structure of this mutation resembles insertion-within-a-duplication mutations that are well represented among defective retroviral replication products in the experimental literature. This insert and its flanking sequences are so dissimilar from one another that the alternate possibility for insert generation – local sequence duplication followed by mutation – can in large part be ruled out because the number of rare events required to generate the observed structure would far exceed those required by the postulated splinted recombination mechanism. Because RT is not known to polymerize more than a single nucleotide or two without a template, it is likely that all retroviral insertions longer than a couple of bases are synthesized using some form of template (Pathak and Temin, 1990; Preston and Dougherty, 1996). Thus, there is no precedence for *de novo* generation of a heteropolymeric insert of this length.

The insert examined in this study was located in RT's $\beta 3$ - $\beta 4$ hairpin, a region where multiple drug resistance-associated