Short Communication: Nucleotide Variation and Positively Selected Sites in HIV Type 1 Reverse Transcriptase Among Heterosexual Transmission Pairs

Uma Shanmugasundaram,¹ Suniti Solomon,¹ Kailapuri G. Murugavel,¹ Kumarasamy Nagalingeswaran,¹ Sunil S. Solomon,¹ Kenneth H. Mayer,² and Balakrishnan Pachamuthu¹

Abstract

Mutations in the *env* gene of HIV-1 have been the primary focus in most epidemiologically related cohort studies of virus evolution and very limited studies have focused on the reverse transcriptase (RT) region, the primary target of antiretroviral therapy (ART). Hence, we measured the selection pressure and searched for the positively selected sites in the RT sequences amplified from HIV-1-infected heterosexual transmission pairs. Married couples ($n = 10$) who were ART naive were included in this study. Phylogenetic analysis, the measurement of synonymous and nonsynonymous ratio (dN/dS) and the interpatient nucleotide variation, was done. Phylogenetic analysis demonstrated distinct subclusters of the RT sequences from heterosexual transmission pairs and the median (IQR) nucleotide variation between the epidemiologically related transmission pairs was significantly ($p < 0.001$) lower [0.01% (0.01–0.02%)] compared to the epidemiologically unrelated transmission pairs [0.04% (0.03–0.04%)]. The ratio of dN/dS was < 1 and codons 135, 162, 166, 207, and 211 were positively selected in $> 50\%$ of the donor and recipient RT sequences. Purifying selection pressure and low nucleotide variation in the RT sequences between epidemiologically related transmission pairs highlight its essential role in HIV-1 replication. The effect of the RT positively selected mutations that persist over time following transmission between individuals needs to be studied to determine the fitness cost of the mutations *in vivo*, which may possibly represent good targets for inclusion in HIV-1 vaccines.

ONE OF THE MOST IMPORTANT and difficult to combat characteristics of human immunodeficiency virus type 1 (HIV-1) is its large adaptive potential, which allows the virus to escape from the host's immune system and from antiretroviral drugs. Viral characteristics such as short generation time,^{1,2} high copy numbers,^{3,4} high mutation rate,⁵ and recombination^{6–8} contribute to the adaptive potential of HIV-1. In addition, the host cellular immune response can also significantly influence the adaptives potential of the virus.⁹ Based on these factors, it may be predicted that the most common viral population is likely to represent the genotype best adapted to the most frequently encountered disease-modifying major histocompatibility alleles. Mutations in the *env* gene of HIV-1 have been the primary focus in most epidemiologically related cohort studies of virus evolution^{10–15} and very few studies have focused on the reverse transcriptase (RT) region of polymerase (*pol*), which is the primary target of antiretroviral

therapy (ART).^{16–18} In this study, we measured the selection pressure on HIV-1 RT and searched for the positively selected codons that may play an important role in the escape from host immunity.

Retrospective HIV-seropositive samples that were collected from ART-naïve transmission pairs (married couples, $n = 10$), during their visit to YRG CARE Medical Centre, were obtained to diagnose HIV. Follow-up samples were also collected after a period of 10–12 months from 12 subjects. Women (recipient) included in the study self-reported to have acquired HIV-1 infection through their husbands (donor). The mean (standard deviation) age of males and females was 34 ± 8.8 and 29 ± 8.6 years, respectively, and their median (interquartile range, IQR) CD4⁺ T cell count was 386 (295–595) and 437 (366–751) cells/ μ l. The study protocol was approved by YRG CARE's Institutional Review Board and written informed consent was obtained from all the participants included in the study.

¹Y.R. Gaitonde Centre for AIDS Research and Education, Chennai, India.

²Warren Alpert School of Medicine, Brown University/Miriam Hospital, Providence, Rhode Island.

FIG. 1. Phylogenetic analysis of HIV-1 RT sequences from heterosexual transmission pairs (D, donor; R, recipient) along with the consensus reference sequence (HIV-1 subtypes A, B, and C) downloaded from the Los Alamos National Laboratory HIV Molecular Immunology Database. The neighbor joining tree is based on the distance calculated between the nucleotide sequences from the transmission pairs. The numbers on the branch points indicate the percent occurrence of branches over 1000 bootstrap resamplings of the data set.

HIV-1 RNA was isolated using the QIAamp viral RNA kit (QIAGEN, Inc., USA). The RT region (20–240) was amplified from cDNA using nested PCR as described earlier¹⁹ with appropriate controls. Bidirectional population sequencing of purified products was done using ABI 3100-*Avant* genetic analyzer (Applied Biosystems, USA). Sequences were edited using Seqscape (Applied Biosystems, USA, v. 2.5) software. Sequence alignment was carried out on the translated amino acid sequence in Clustal W,²⁰ as implemented in MEGA version 3.1.²¹ Maximum likelihood, minimum evolution, neighbor-joining phylogenetic analyses were used to explore the heterosexual relationships among the donor and recipient RT sequences. The robustness of each tree was evaluated by bootstrap analysis of 1000 replicas.

Interpatient nucleotide distance was measured by comparing the donor and recipient RT sequences using the Kimura two-parameter nucleotide distance method as implemented in MEGA version 3.1.²¹ The RT sequences amplified from primary and follow-up samples were classified into distinct groups as donor and the recipient and compared with the consensus C reference sequence to calculate the ratio of nonsynonymous and synonymous substitutions (dN/dS) and the analysis was done using Syn-SCAN.²² Drug resistance mutations were identified using the Stanford HIV-1 Drug

Resistance Database (<http://hivdb.stanford.edu/>). The donor and recipient RT sequences were compared with the consensus C reference sequence to identify the codon sites evolving under the influence of positive Darwinian selection and the analysis was performed using HyPhy with codon substitution model MG94.²³ The Mann-Whitney U test was used to compare the RT nucleotide variation between the epidemiologically related and unrelated transmission pairs.

The phylogenetic analysis demonstrated distinct subclusters of the RT sequences from heterosexual transmission pairs (Fig. 1). The median (IQR) nucleotide distance between the epidemiologically related transmission pairs was significantly ($p < 0.001$) lower [0.01% (0.01–0.02)] than the epidemiologically unrelated transmission pairs [0.04% (0.03–0.04)]. A previous study related the level of nucleotide variation among epidemiologically linked mother and infant transmission pairs with viral pathogenesis.¹⁸ Two transmission pairs (Pair 2 and 8) showed no sequence variation (0.00%), however, clustered separately, confirming that they belong to the transmission chain. The remaining transmission pairs exhibited nucleotide distances ranging from 0.01% to 0.02%. The extent of HIV-1 nucleotide variation observed between the transmission pairs could be associated with duration of infection, selective pressure imposed by the host immune system, and the rate of disease

FIG. 2. Occurrence of positively selected sites in HIV-1 RT among the transmission pairs. The dark and light colored bars represent the prevalence of positively selected codons in the RT sequences amplified from the donor and the recipient, respectively.

progression.^{24–28} Owing to the cross-sectional analysis and lack of information on the duration of HIV-1 infection, these aspects could not be delineated from the level of nucleotide variation.

The ratio of dN/dS in the donor [0.2 (0.1–0.2)] and recipient sequences [0.2 (0.12–0.23)] was <1 , which indicates the purifying selection pressure²⁹ over the RT region in parallel with other studies.^{30–32} Selective transmission of a few viral variants has been proposed in horizontal as well as vertical transmission due to the detection of highly homogeneous *env* sequences in recently infected individuals and due to the detection of greater *env* gene diversity in chronically infected individuals.^{33–38} The higher dN/dS estimates have been linked to selection for changes in the virus by the host adaptive immune response, which includes neutralizing antibodies, cytotoxic T cells, and/or T helper cell responses.^{39,40} These changes will enable the virus to escape from the immune response and are most likely to be associated with longer survival times.^{27,41} However, synonymous nucleotide changes were usually observed to predominate among the surviving variant sequences and dN/dS tends to be <1 when there is a requirement to maintain protein function.^{30–32}

No drug resistance mutations were observed in the RT sequences amplified from the transmission pairs. Positive selection of codons 135, 162, 166, 207, and 211 occurred in $>50\%$ of the donor and the recipient RT sequences (Fig. 2). These codons were positively selected in the 12 sequences collected during follow-up, of which eight were from the transmission pairs ($n=4$). Codons positively selected in the donor sequences were also observed to be positively selected in the recipient sequences. Repeated positive selection of the same amino acid mutations that are located in the regions of the cytotoxic T lymphocyte epitope (aa 128–135, 158–166, and 202–212) is likely to be a characteristic of the epitope subjected to the strongest host immune selective pressure.^{9,42–47} Baseline polymorphisms in treatment-naïve subjects within RT have been linked to resistance to RT inhibitors.^{30,48–50} It appears that CD8⁺ T cell escape variants could become fixed and propagated through populations⁴⁷ in the same way viruses resistant to ART are transmitted.⁵¹

An earlier study⁵² proposed two rationales for the lack of reversion of CD8⁺ T cell escape variants: the absence of a

significant influence on viral fitness and the existence of CD8⁺ T cell-mediated selective pressure restricted by a human leukocyte antigen (HLA) allele. Since no evidence was reported to support the above affirmed rationale, future studies are required to identify associations between an HLA allele and specific amino acid polymorphism and also to evaluate dominant immune responses to the antigenic peptides of subtype C *pol*, which will be needed to gain insight into the viral pathogenic mechanisms as well as the design of a vaccine.

In conclusion, the observed purifying selection pressure and low RT nucleotide variation between epidemiologically related transmission pairs highlight its essential role in HIV-1 replication. The effect of RT positively selected mutations that persist over time following transmission between individuals needs to be studied to determine the fitness cost of the mutations *in vivo*, which may possibly represent good targets for inclusion in HIV-1 vaccines.

Nucleotide Sequence Accession Numbers

The GenBank accession numbers of the HIV-1 RT described here are EU429988 through EU430023 and EU597182 through EU597197.

Author Disclosure Statement

No competing financial interests exist.

References

- Rodrigo AG, Shpaer EG, Delwart EL, Iversen AK, Gallo MV, Brojtsch J, Hirsch MS, Walker BD, and Mullins JI: Coalescent estimates of HIV-1 generation time in vivo. *Proc Natl Acad Sci USA* 1999;96:2187–2191.
- Perelson AS, Neumann AU, Markowitz M, Leonard JM, and Ho DD: HIV-1 dynamics in vivo: Virion clearance rate, infected cell life-span, and viral generation time. *Science* 1996;271:1582–1586.
- Ho DD, Neumann AU, Perelson AS, Chen W, Leonard JM, and Markowitz M: Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* 1995;373:123–126.

4. Wei X, Ghosh SK, Taylor ME, Johnson VA, Emini EA, Deutsch P, Lifson JD, Bonhoeffer S, Nowak MA, Hahn BH, *et al.*: Viral dynamics in human immunodeficiency virus type 1 infection. *Nature* 1995;373:117–126.
5. Mansky LM and Temin HM: Lower in vivo mutation rate of human immunodeficiency virus type 1 than predicted from the fidelity of purified reverse transcriptase. *J Virol* 1995;69:5087–5094.
6. Levy D, Aldrovandi G, Kutsch O, and Shaw G: Dynamics of HIV-1 recombination in its natural target cells. *Proc Natl Acad Sci USA* 2004;101:4204–4209.
7. Charpentier C, Nora T, Tenaillon O, Clavel F, and Hance A: Extensive recombination among human immunodeficiency virus type 1 quasispecies makes an important contribution to viral diversity in individual patients. *J Virol* 2006;80:2472–2482.
8. Ramirez BC, Liorie ES, Galetto R, and Negroni M: Implications of recombination for HIV diversity. *Virus Res* 2008;134:64–73.
9. Ahlenstiel G, Roomp K, Däumer M, Nattermann J, Vogel M, Rockstroh JK, Beerenwinkel N, Kaiser R, Nischalke HD, Sauerbruch T, Lengauer T, and Spengler U; Kompetenznetz HIV/AIDS: Selective pressures of HLA genotypes and antiviral therapy on human immunodeficiency virus type 1 sequence mutation at a population level. *Clin Vaccine Immunol* 2007;14:1266–1273.
10. Mikhail M, Wang B, Lemey B, Beckthold B, Vandamme A, Gill JM, and Saksena NK: Role of viral evolutionary rate in HIV-1 disease progression in a linked cohort. *Retrovirology* 2005;2(1):41.
11. Song JZ, Wang B, Ge YC, Dwyer DE, Cunningham AL, and Saksena NK: Significance of plasma and peripheral blood mononuclear cell derived HIV-1 sequences in establishing epidemiologic linkage between two individuals multiply exposed to HIV-1. *Microb Pathog* 1999;26(6):287–298.
12. Shin BG, Nam JG, Kim GJ, Lee JS, and Kim SS: Limited genetic variation during sexual transmission of HIV-1 Korean clade B. *International Conference AIDS*, August 13–18, 2006. Abstract no. WePe0028.
13. Gray L, Fiscus S, and Shugars D: HIV-1 variants from a perinatal transmission pair demonstrate similar genetic and replicative properties in tonsillar tissues and peripheral blood mononuclear cells. *AIDS Res Hum Retroviruses* 2007;23:1095–1104.
14. Hoffmann FG, He X, West JT, Lemey P, Kankasa C, and Wood C: Genetic variation in mother-child acute seroconverter pairs from Zambia. *AIDS* 2008;22:817–824.
15. Campbell MS, Mullins JI, Hughes JP, Wong KG, Raugi DN, Sorensen S, Stoddard JN, Zhao H, Deng W, McCutchan FE, Albert J, Leitner T, Celum C, and Lingappa JR: Determination of HIV transmission linkage in the Partners in Prevention Study. *International AIDS Society*, 2009. Abstract No. LBPEC07.
16. Bao L, Vidal N, Fang H, Deng W, Chen S, Guo W, Qin C, Peeters M, Delaporte E, Andrieu JM, and Lu W: Molecular tracing of sexual HIV type 1 transmission in the southwest border of China. *AIDS Res Hum Retroviruses* 2008;24:733–742.
17. Han Z, Leung TW, Zhao J, Wang M, Fan L, Li K, Pang X, Liang Z, Lim WW, and Xu H: A HIV-1 heterosexual transmission chain in Guangzhou, China: A molecular epidemiological study. *Virol J* 2009;25(6):148.
18. Sundaravaradan V, Hahn T, and Ahmad N: Conservation of functional domains and limited heterogeneity of HIV-1 reverse transcriptase gene following vertical transmission. *Retrovirology* 2005;26(2):36.
19. Larder BA, Kellam P, and Kemp SD: Zidovudine resistance predicted by direct detection of mutations in DNA from HIV-infected lymphocytes. *J Acquir Immune Defic Syndr* 1991;5:37–44.
20. Thompson JD, Higgins DG, and Gibson TJ: Clustal W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl Acids Res* 1994;22:4673–4680.
21. Kumar S, Tamura K, and Nei M: MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform* 2004;5:150–163.
22. Gonzales MJ, Dugan JM, and Shafer RW: Synonymous–non-synonymous mutation rates between sequences containing ambiguous nucleotides (Syn-SCAN). *Bioinformatics* 2002;18(6):886–887.
23. Pond SLK, Frost SDW, and Muse SV: HyPhy: Hypothesis testing using phylogenies. *Bioinformatics* 2005;21:676–679.
24. Lee JS, Kim SS, Koo BK, Choi BS, Seong BM, Suh SD, Kim OJ, and Cho HW: Adaptive evolution of human immunodeficiency virus type 1 having common origin within different individuals. *International Conference on AIDS* July 9–14, 2000. Abstract no. MoOrA164.
25. Zhang H, Hoffmann F, He J, He X, Kankasa C, West JT, *et al.*: Characterization of HIV-1 subtype C envelope glycoproteins from perinatally infected children with different courses of disease. *Retrovirology* 2006;3:73.
26. Ganesan S, Dickover RE, Korber BT, Bryson YJ, and Wolinsky SM: Human immunodeficiency virus type 1 genetic evolution in children with different rates of development of disease. *J Virol* 1997;71:663–677.
27. Williamson S: Adaptation in the *env* gene of HIV-1 and evolutionary theories of disease progression. *Mol Biol Evol* 2003;20:1318–1325.
28. Halapi E, Leitner T, Jansson M, Scarlatti G, Orlandi P, Plebani A, Romiti L, Albert J, Wigzell H, and Rossi P: Correlation between HIV sequence evolution, specific immune response and clinical outcome in vertically infected infants. *AIDS* 1997;11(14):1709–1717.
29. Nielsen R and Yang Z: Likelihood models for detecting positively selected amino acid sites and application to the HIV-1 envelope gene. *Genetics* 1998;148:929–936.
30. Ceccherini-Silberstein F, Gago F, Santoro M, Gori C, Svicher V, Rodríguez-Barrios F, d'Arrigo R, Ciccozzi M, Bertoli A, d'Arminio Monforte A, Balzarini J, Antinori A, and Perno CF: High sequence conservation of human immunodeficiency virus type 1 reverse transcriptase under drug pressure despite the continuous appearance of mutations. *J Virol* 2005;79:10718–10729.
31. Chen L and Lee C: Distinguishing HIV-1 drug resistance, accessory, and viral fitness mutations using conditional selection pressure analysis of treated versus untreated patient samples. *Biol Direct* 2006;1:14.
32. Seibert SA, Howell CY, Hughes MK, and Hughes AL: Natural selection on the gag, pol, and env genes of human immunodeficiency virus 1 (HIV-1). *Mol Biol Evol* 1995;12:803–813.
33. Burger HB, Flaherty WK, Gulla J, Nguyen PN, and Gibbs RA: Evolution of human immunodeficiency virus type 1 nucleotide sequence diversity among close contacts. *Proc Natl Acad Sci USA* 1991;88:11236–11240.
34. Karlsson AC, Lindback C, Gaines H, and Sonnerborg A: Characterization of the viral population during primary HIV-1 infection. *AIDS* 1998;12:839–847.

35. Verhofstede C, Demecheleer E, De Cabooter N, Gaillard P, Mwanjumba F, Claeys P, Chohan V, Mandaliya K, Temmerman M, and Plum J: Diversity of the human immunodeficiency virus type 1 (HIV-1) env sequence after vertical transmission in mother-child pairs infected with HIV-1 subtype A. *J Virol* 2003;77:3050–3057.
36. Zhu T, Mo H, Wang N, Nam DS, Cao Y, Koup RA, and Ho DD: Genotypic and phenotypic characterization of HIV-1 patients with primary infection. *Science* 1993;261:1179–1181.
37. Ahmad N, Baroudy BM, Baker RC, and Chappey C: Genetic analysis of human immunodeficiency virus type 1 envelope V3 region isolates from mothers and infants after perinatal transmission. *J Virol* 1995;69:1001–1012.
38. Scarlatti G, Leitner T, Halapi E, Wahlberg J, Marchisio P, Clerici-Schoeller MA, Wigzell H, Fenyo EM, Albert J, Uhlen M, *et al.*: Comparison of variable region 3 sequences of human immunodeficiency virus type 1 from infected children with the RNA and DNA sequences of the virus populations of their mothers. *Proc Natl Acad Sci USA* 1993;90:1721–1725.
39. Huber M and Trkola A: Humoral immunity to HIV-1: Neutralization and beyond. *J Intern Med* 2007;262:5–25.
40. Ostrowski MA, Yu Q, Yue FY, Liu J, Jones B, Gu XX, Loutfy M, Kovacs CM, and Halpenny R: Why can't the immune system control HIV-1? Defining HIV-1-specific CD4⁺ T cell immunity in order to develop strategies to enhance viral immunity. *Immunol Res* 2006;35:89–102.
41. Ross HA and Rodrigo AG: Immune-mediated positive selection drives human immunodeficiency virus type 1 molecular variation and predicts disease duration. *J Virol* 2002;76:11715–11720.
42. Doolittle RF: Convergent evolution: The need to be explicit. *Trends Biochem Sci* 1994;19:15–18.
43. Hughes AL: *Adaptive Evolution of Genes and Genomes*. Oxford University Press, New York, 1999.
44. Piontkivska H and Hughes AL: Patterns of sequence evolution at epitopes for host antibodies and cytotoxic T-lymphocytes in human immunodeficiency virus type 1. *Virus Res* 2006;116:98–105.
45. Brumme ZL, Brumme CJ, Heckerman D, Korber BT, Daniels M, Carlson J, Kadie C, Bhattacharya T, Chui C, Szinger J, Mo T, Hogg RS, Montaner JS, Frahm N, Brander C, Walker BD, and Harrigan PR: Evidence of differential HLA class I-mediated viral evolution in functional and accessory/regulatory genes of HIV-1. *PLoS Pathogens* 2007;3(7):e94.
46. Allen TM, Altfeld M, Geer SC, Kalife ET, Moore C, O'Sullivan M, Desouza KI, Feeney ME, Eldridge RL, Maier EL, Kaufmann DE, Lahaie MP, Reytor L, Tanzi G, Johnston MN, Brander C, Draenert R, Rockstroh RK, Jessen H, Rosenberg ES, Mallal SA, and Walker BD: Selective escape from CD8⁺ T-cell responses represents a major driving force of human immunodeficiency virus type 1 (HIV-1) sequence diversity and reveals constraints on HIV-1 evolution. *J Virol* 2005;79:13239–13249.
47. Moore CB, John M, James IR, Christiansen FT, Witt CS, and Mallal SA: Evidence of HIV-1 adaptation to HLA-restricted immune responses at a population level. *Science* 2002;296:1439–1443.
48. Kantor R, Katzenstein DA, Efron B, Patricia Carvalho A, Wynhoven B, Cane P, Clarke J, Sirivichayakul S, Soares MA, and Snoeck J: Impact of HIV-1 subtype and antiretroviral therapy on protease and reverse transcriptase genotype: Results of a global collaboration. *PLoS Med* 2005;2(4):e112.
49. Margot NA, Waters JM, and Miller MD: In vitro human immunodeficiency virus type 1 resistance selections with combinations of tenofovir and emtricitabine or abacavir and lamivudine. *Antimicrob Agents Chemother* 2006;50:4087–4095.
50. Svicher V, Sing T, Santoro MM, Forbici F, Rodríguez-Barrios F, Bertoli A, Beerenwinkel N, Bellocchi MC, Gago F, d'Arminio Monforte A, Antinori A, Lengauer T, Ceccherini-Silberstein F, and Perno CF: Involvement of novel human immunodeficiency virus type 1 reverse transcriptase mutations in regulation of resistance to nucleoside inhibitors. *J Virol* 2006;80:7186–7198.
51. Luzuriaga K, McManus M, Mofenson L, Britto P, Graham B, and Sullivan JL: A trial of three antiretroviral regimens in HIV-1-infected children. *N Engl J Med* 2004;350:2471–2480.
52. Merino VS, Nie S, and Luzuriaga K: HIV-1-specific CD8⁺ T cell responses and viral evolution in women and infants. *J Immunol* 2005;175:6976–6986.

Address correspondence to:

Balakrishnan Pachamuthu
Y.R. Gaitonde Centre for AIDS Research and Education
Voluntary Health Services
Taramani, Chennai-600113, India

E-mail: bala@yrgcare.org