Virus Genes (2007) 34:111–116
DOI 10.1007/s11262-006-0054-z

Genetic analysis of the long terminal repeat (LTR) promoter region in HIV-1-infected individuals with different rates of disease progression

Eva Ramírez de Arellano · Cristina Martín · Vincent Soriano · José Alcamí · África Holguín

Received: 23 August 2006 / Accepted: 26 October 2006 / Published online: 9 December 2006

Abstract The long terminal repeat (LTR) region of HIV-1 promotes and modulates proviral transcription. LTR genetic variability might influence viral replication and disease progression. Proviral LTR sequences from 32 HIV-1-infected individuals showing different rates of disease progression were examined. Non-progressors (NP, $n = 11$) were individuals with high and stable CD4 counts and persistently low or undetectable plasma HIV-RNA. Slow progressors (SP, $n = 6$) were subjects with minimal CD4 decays over time and low plasma HIV-RNA. Typical progressors (TP, $n = 15$) were individuals with chronic infection showing CD4 counts repeatedly below 500 cells/μl. The mutation frequency within distinct LTR functional regions involved in HIV-1 transcription were compared in these three groups of patients. No significant differences were observed in the mutation frequency in most LTR regulatory sites when comparing the three groups. However, changes in USF regulatory binding sites were more frequent in TP than in SP/NP, while changes in Sp1 binding sites were less common in the former. This is the first study examining the genetic variability of the HIV-1 LTR region in long-term non-progressors showing further divergent outcomes.

E. Ramírez de Arellano · C. Martín · V. Soriano · Á. Holguín (✉)
Department of Infectious Diseases, Hospital Carlos III, C/ Sinesio Delgado, 10, 28029 Madrid, Spain
e-mail: aholguin.hciii@salud.madrid.org

J. Alcamí
AIDS Immunopathogenesis Unit, Instituto de Salud Carlos III, Majadahonda, Spain

Keywords HIV-1 · Long-term non-progressors · LTR · Transcription · Genetic variability

Introduction

Individuals with HIV-1 infection without evidence of progressive immunodeficiency for more than a decade represent a unique subset of patients, particularly of interest for studying HIV-1 pathogenesis. Long-term non-progressors (LTNP) are generally defined as a heterogeneous group of clinically healthy individuals, with normal CD4+ T cell counts and very low or undetectable viral load, after being infected with HIV-1 for longer than 10 years, in the absence of any antiretroviral treatment. Multiple factors, both from the virus and the host, seem to account for the lack of disease progression in these subjects [1].

Mutations in cellular HIV-1 co-receptor genes, including a 32 bp deletion in the chemokine receptor gene CXCR5, have been associated with slow disease progression [2]. In other studies, high titers of neutralizing antibodies [3], more effective cellular immune responses [4], and the presence of certain HLA alleles [5] have all been claimed to be involved in the benign course of HIV-1 disease in LTNP.

The genetic characterization of viruses infecting LTNP has occasionally revealed defects in several regions, such as *vif, vpr, vpu*, *rev* and *nef* [6–9]. However, the association between structural or functional alterations in LTR and their influence on disease progression is controversial. While some studies could not find any association between LTR specific genetic changes and the speed of disease progression [10, 11], others have shown that multiple alterations in certain LTR

regions might be associated with increased virus replication [12] or with impaired viral transcription, as recently described in one individual with slow disease progression due to a large deletion in the *nef*/LTR region [13].

Two identical LTR sequences located in both extremes of all HIV-1 proviruses exist. They are essential for viral transcription, integration and gene expression. Each LTR is divided into three regions: U3, R and U5 (Fig. 1). U3 is further subdivided into three elements: promoter, enhancer and modulator regions, which contains important binding sites for cellular proteins such as Sp1 (Specific protein 1), NF-kB (nuclear factor kappa B), NF-AT (nuclear factor of activated T cells), USF (upstream stimulatory factor), TCF-1α (T cell factor alpha) and RBF-2 (Ras response element binding factor 2) [14, 15]. The R region contains the TAR hairpin for binding the viral protein Tat (transcription trans-activator), required for the production of viable viral transcripts. The U5 region contains other binding sites for cellular transcription factors such as AP-1 (activating protein-1), AP3-like (corresponding to a NF-AT site) and Sp1, which are important for virus infectivity [16]. Viruses from a few LTNP have shown mutations in Sp1 sites I and III, which down-regulated LTR transcription in vitro [17, 18]. Others could not find detectable levels of viral transcription in LTNP carrying G → A hypermutated viruses throughout the entire LTR region [19].

The long terminal repeat (LTR) region of HIV-1 promotes and modulates proviral transcription. Since LTR genetic variability might influence viral replication and disease progression, the genetic characterization of the HIV-1 LTR was undertaken in 32 HIV-1-infected patients with different rates of disease progression (non-progressors, slow progressors and typical progressors). This is the first study showing different sequence variability within two different regulatory sites (Sp1 and USF) in HIV-1+ subjects showing further divergent outcomes.

Materials and methods

Proviral DNA consensus sequences from the 5′ LTR (630 bp) region were obtained by PCR amplification from HIV-1 proviruses integrated in peripheral blood mononuclear cells (PBMC) genome in a cohort of 17 LTNP established in 1997 at our institution. All these individuals had more than 16 years of confirmed HIV-1 infection, despite which all were asymptomatic, and had more than 500 CD4+ T cells per microliter in the absence of any antiretroviral therapy [20]. Interestingly, on prolonged follow-up, some of these subjects showed slow CD4 declines and mild viral load increases while others remained unchanged. The former were re-defined as slow progressors (SP) while the latest were considered as truly non-progressors (NP). The 11 subjects fulfilling the criteria for NP had high and stable CD4 counts and persistently low or undetectable plasma HIV-RNA (<50 copies/ml) for the whole study period, while this was not the case for the 6 SP. The main virological and immunological features of these individuals are in Table 1. Briefly, 11 out of 17 were male; their mean age was 40.4 years and the mean length of proven HIV-1 infection was 14.7 years in 1997, when the cohort was established, and of 22 years when the current study was conducted.

As controls, we used previously reported LTR sequences deposited at Genbank and derived from 15

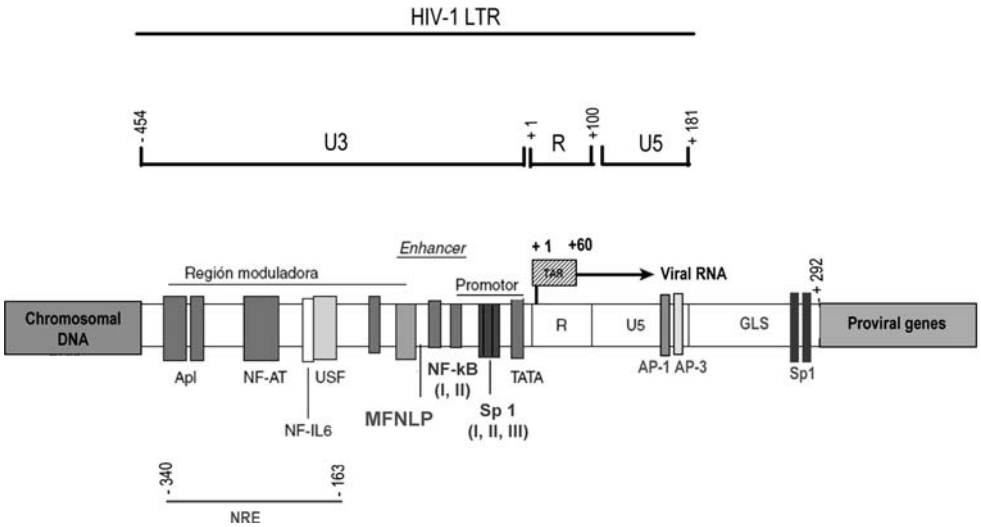

Fig. 1 HIV-1 LTR structure of proviral DNA and binding sites for cellular transcription factors in the promoter, enhancer and modulatory regions of U3, R and U5. Numbering the boxed region is relative to the transcription start site nucleotide +1

Table 1 Main virological and immunological features of SP and NP patients

Patient[a]	Year of diagnosis	Age (years)	Gender	Mean HIV-1 RNA (copies/ml)	Mean CD4 count ($\times 10^6$ cells/l)
NP1	1989	41	M	7728	837
NP2	1987	35	M	2159	652
NP3	1986	38	M	UN	640.5
NP4	1985	40	F	UN	903
NP5	1987	44	M	78	966.5
NP6	1985	44	F	92	1100
NP7	1988	41	M	116.5	1194
NP9	1989	40	M	UN	918
NP10	1987	40	M	967	586
NP11	1987	38	M	159.5	746
NP12	1989	35	F	UN	1395
SP13	1985	42	F	1171.5	740
SP14	1986	36	M	500	810
SP16	1985	39	F	325	525
SP18	1989	33	M	675	798
SP19	1979	43	M	2664.5	689
SP20	1987	50	M	477	537

[a] Patients numbered according to Ref [20]

M, male; F, female; UN, undetectable viraemia at all times

individuals considered as typical progressors (TP) of HIV-1 disease. All were drug-naive at the time they were analyzed, and had been infected with HIV-1 for at least 8 years. Other features have been previously described [11].

To obtain the consensus sequences from the 5′ LTR (630-bp) region of each LTNP, a nested PCR was carried out. Outer primers were LTR-54D (5′-TAC CAC ACA CAA GGC TAC TTC CCT GA-3′) and LTR-718R (5′-GTG CGC GCT TCA GCA AGC CGA GTC C-3′). The inner 5′ LTR amplification was performed with primers LTR-74D (5′-CCC TGA TTG GCA GAA YTA CAC ACC AGG-3′) and LTR-708R (5′-CAG CAA GCC GAG TCC TGC GTC GAG AG-3′). The standard reaction conditions for the outer PCR were incubation at 94°C for 3 min, 35 cycles at 94°C for 40 s, 55°C for 40 s, and 72°C for 1 min, with a final extension at 72°C for 10 min. For the inner PCR, the same conditions were used with primer hybridization at 50°C. The PCR fragments were sequenced directly using an automatic sequencer (ABI PRISM, Applied Biosystems, Foster City, CA). Only one consensus sequence was obtained for each PCR product, which includes the most frequently found nucleotide at each position in the individual clones of the HIV-1 quasiespecies population.

For HIV-1 subtype characterization, sequences belonging to HIV-1 group M obtained from the GenBank were used as references for phylogenetic analyses. The tree topology was obtained using the Neighbor-Joining program. After alignment of consensus sequences from viruses infecting TP, SP and NP using the Clustal W method, columns containing gaps were eliminated. The distance matrix was estimated using the Kimura two-parameter model within the DNADIST program, as implemented in the PHYLIP software package. Bootstrap (1,000 data sets) of the multiple alignment was performed to test the statistical robustness of the tree.

The LTR nucleotide sequence data reported in this paper have been submitted to the GenBank nucleotide sequence database (accession numbers AF851669 to AF851684).

The frequency of mutation (number of mutations found in each consensus sequence with respect to the HXB2 reference sequence divided by the total number of nucleotides sequenced) was calculated for the most important LTR regions known to interact with cellular or viral proteins relevant for HIV-1 transcription regulation. The chi-square or Fisher's exact tests were used to detect significant differences between groups. Comparisons were analyzed using Epi Info v6.02 (CDC, Atlanta, GA). All P values were two-sided. Only P values below 0.05 were considered as significant.

Results

All LTR sequences from viruses infecting LTNP and controls were clustered with HIV-1 subtype B. No significant differences were observed in the mutation frequency in most LTR regulatory sites when comparing the three groups of patients. The most conserved nucleotide sequences were observed in RBF, NF-kB and the viral trans-activator Tat binding sites within LTR (Table 2). However, the genetic variability in the USF (located within the Core-NRE) and Sp1 host proteins binding sites differed significantly when comparing these populations. USF binding sequences

Table 2 LTR genetic variability at different binding sites relevant for virus transcription regulation in HIV-1-infected individuals with different rates of disease progression

No. of samples	Mutation frequency ($\times 10^{-2}$)[a]													
	U3												R	
	U3/Modulatory region						U3/Region enhancer		U3/Core				R/TAR	
	Core-NRE		RBF-1 (–148 to –145)	LEF-1 (–136 to –131)	RBF-2 (–130 to –122)		NF-kB(II) (–109 to –99)	NF-kB(I) (–94 to –84)	SP1(III) (–81 to –70)	SP1(II)[c] (–69 to –58)	SPI(I)[c] (–57 to –46)	TATAA (–28 to –24)	TAR (+1 to +60)	
	Total (–174 to –163)	USF[c] (–168 to –163)												
SP (6)	5.1	5.5	0	33	0		0	0	8.3	4.5	1.6	0	0.3	
NP (11)	6.3	4.5	0	30	2.2		2.7	3.3	8	9	5.5	0	1.8	
Total LTNP (17)	5.8	5	0	31	1.4		1.6	2.1	8.2	7.5	4.1	0	1.3	
TP (15)[b]	2	14.4	0	12.2	3.3		2	0.6	7.3	2.4	0.6	0	1.2	

[a] Mutant frequencies have been calculated by dividing the number of mutations with respect to HXB2 sequence by the total number of nucleotides sequenced in each LTR region

[b] Calculations including data from 15 sequences from clade B viruses derived from typical progressors [11]

The LTR position is indicated according to HXB2 sequence

[c] Regulatory sites with significant differences in mutation frequency among TP and LTNP

(LTR positions –168 to –163) were nearly 3-fold more conserved in the whole group of LTNP than in TP (5×10^{-2} vs. 14.4×10^{-2} substitutions/nt; $P = 0.02$) (Table 2). This site contains a characteristic E-Box motif, CAC(A/G)TG and in vivo USF has been found to repress transcription in epithelial cells, although it activates the same transcription in T cell lines [21].

Two out the three LTR binding sites for the Sp1 cellular protein [Sp1 (I) and (II)] (LTR positions –46 to –69), are responsible for LTR promoter activity in vitro and transcription activation. In contrast, the third Sp1 (III) site (LTR positions –70 to –81) has minimal effects on viral transcription [22]. These three Sp1 sites are rich in G/C nucleotides, which are structurally essential for the binding of the Sp1 cellular factor [22, 23]. Interestingly, the mutation frequency of the first two Sp1 binding sites was significantly greater in viruses infecting LTNP (both NP and SP) compared to those infecting TP (4.1×10^{-2} vs. 0.6×10^{-2} substitutions/nt for Sp1 (I), $P = 0.04$; and 7.5×10^{-2} vs. 2.4×10^{-2} substitutions/nt for Sp1 (II), $P = 0.03$) (Table 2). Moreover, the number of subjects with viruses harboring A/T nucleotide changes in Sp1 (I, II) was higher among NP compared to TP [36.4% vs. 6.6%, for Sp1 (I); and 63.6% vs. 13.3%, in Sp1 (II)].

The genetic variability in NP and SP was similar along most LTR sites, except in Sp1 (I, II). Viruses derived from NP tended to show a higher mutation frequency than those from SP. This increased genetic variability found within Sp1 (I, II) in NP was mostly due to a higher A/T nucleotide content at the Sp1 (I, II) sites (Fig. 2).

The most frequent naturally occurring length polymorphism (MFNLP) within the LTR region is located upstream of the NF-kB site at the LTR HXB2 position –120. Insertions of 15 to 34 bp long have been previously described [24, 25]. We found insertions of 6–15 nucleotides in the MFNLP region in viruses derived from 27% of TP and 17.5% of LTNP (P = ns).

Discussion

This is the first study examining the genetic variability of the HIV-1 LTR region in LTNP showing further divergent outcomes (NP or SP individuals). Overall LTR variability in most regulatory sites did not differ when comparing TP and LTNP, although different mutation frequency in motifs involved in the binding of Sp1 and USF cellular proteins was seen in LTNP compared to TP. These sites are essential for HIV-1 transcription regulation. As a consequence, genetic changes in them could influence the interactions with

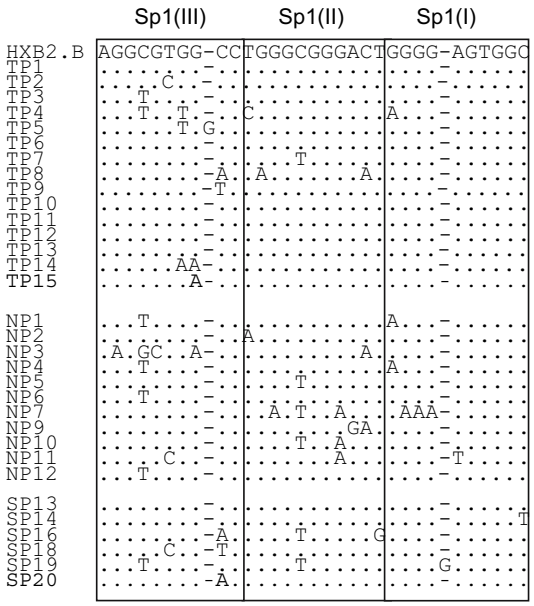

Fig. 2 Sequences of Sp1 sites (I, II, III) from viruses infecting TP, NP and SP individuals with respect to HXB2 consensus reference. TP, patients numbered according to Ref. [11]. NP and SP, patients numbered according to Ref. [20]. Dot, identical nucleotide than HXB2

cellular transcription factors modifying HIV-1 transcriptional levels [22].

The observed difference on genetic variability in Sp1 sites in viruses infecting NP versus SP individuals could determine different biological properties between both groups. Since a high G/C nucleotide content is required for the binding of the cellular factor Sp1 to LTR [23], the elevated A/T nucleotide content at the Sp1 (I, II) sites we found in viruses derived from LTNP (particularly in NP) could decrease the binding affinity for Sp1 in these patients, contributing to a reduced viral transcription in them.

The low size of the study population most likely precluded obtaining statistical significance. However, it has been reported that only in a minority of individuals without disease progression (1–5%), HIV-1 seems to be less pathogenic. These LTNP remain asymptomatic with CD4 $>500 \times 10^6$ cell/l and show low or undetectable viral load in the absence of antiretroviral therapy. In fact, only 17 out of a total of 2,300 HIV-1 individuals on regular follow-up at our institution fit the criteria for LTNP, being the target population in this study.

The recognition of insertions within the MFNLP region in 17–27% of individuals, without significant differences between LTNP and TP, is of interest. Although the MFNLP region is a binding site for cellular factors such as TCF-1α, which activates LTR activity [24], others have shown that MFNLP can bind a specific nuclear factor that appears to be identical to RBF-2, producing a negative effect on viral transcription [25].

The fact that previous studies did not find any relationship between LTR genetic variability and the rate of disease progression [10, 11] could be related to the LTNP population studied. While other authors studied viral LTR variability in LTNP, our study is the first examining the LTR genetic variability in LTNP individuals further stratified into two groups: SP and NP, showing different clinical outcome after prolonged follow up. Therefore, we observed that NP tended to show a higher mutation frequency in Sp1 (I, II) sites than those from SP or TP subjects.

Although a greater genetic variability at the LTR region could be expected for LTNP compared to TP, given their longer length of infection, some regulatory sites, such as USF, showed a lower genetic variability in LTNP than in TP. Functional studies will be required to finally prove that these LTR genetic differences may lead to differences in HIV-1 transcription regulation in larger series of patients with different rates of disease progression.

Acknowledgements This work was supported in part by grants from Fundación Investigación y Educación en SIDA (IES) and Red de Investigación en SIDA (RIS project G03/173). A.H. was supported by Fundación para la Investigación Biomédica del Hospital Ramón y Cajal and Fondo de Investigaciones Sanitarias. V.S. was supported by Instituto de Salud Carlos III and Agencia Laín Entralgo.

References

1. Y. Cao, L. Qin, J. Zhang, J. Safrit, D. Ho, N. Engl. J. Med. **332**, 201 (1995)
2. M. Samson, F. Libert, B.J. Doranz, J. Rucker, C. Liesnard, C.M. Farber, S. Saragosti, C. Lapoumeroulie, J. Cognaux, C. Forceille, G. Muyldermans, C. Verhofstede, G. Burtonboy, M. Georges, T. Imai, S. Rana, Y. Yi, R.J. Smyth, R.G. Collman, R.W. Doms, G. Vassart, M. Parmentier, Nature **382**, 722 (1996)
3. A.K. Pilgrim, G. Pantaleo, O.J. Cohen, L.M. Fink, J.Y. Zhou, J.T. Zhou, D.P. Bolognesi, A.S. Fauci, D.C. Montefiori, J. Infect. Dis. **176**, 924 (1997)
4. D. Blackbourn, C. Mackewicz, E. Barker, T. Hunt, B. Herndier, A. Haase, A. Levy, Proc. Natl. Acad. Sci. USA, **93**, 13125 (1996)
5. M. Westby, F. Manca, A. Dalgleish, Immunol. Today **17**, 120 (1996)
6. N.J. Deacon, A. Tsykin, A. Solomon, K. Smith, M. Ludford-Menting, D.J. Hooker, D.A. McPhee, A.L. Greenway, A. Ellett, C. Chatfield, V.A. Lawson, S. Crowe, A. Maerz, S. Sonza, J. Learmont, J.S. Sullivan, A. Cunningham, D. Dwyer, D. Dowton, Mills J. Sci. **270**, 988 (1995)
7. A.K. Iversen, E.G. Shpaer, A.G. Rodrigo, M.S. Hirsch, B.D. Walker, H.W. Sheppard, T.C. Merigan, J.I. Mullins, J. Virol. **69**, 5743 (1995)

8. N.L. Michael, G. Chang, L.A. d'Arcy, P.K. Ehrenberg, R. Mariani, M.P. Busch, D.L. Birx, D.H. Schwartz, J. Virol. 69, 4228 (1995)
9. J.J. Lum, O.J. Cohen, Z. Nie, J.G. Weaver, T.S. Gomez, X.J. Yao, D. Lynch, A.A. Pilon, N. Hawley, J.E. Kim, Z. Chen, M. Montpetit, J. Sanchez-Dardon, E.A. Cohen, A.D. Badley, J. Clin. Invest. **145**, 1547 (2003)
10. M.E. Quinones-Mateu, A. Mas, T. Lain de Lera, V. Soriano, J. Alcami, M.M. Lederman, E. Domingo, Virus Res. **57**, 11 (1998)
11. U. Visco-Comandini, Z. Yun, A. Vahlne, A. Sonnerborg, AIDS Res. Hum. Retroviruses **15**, 609 (1999)
12. K. Hiebenthal-Millow, T.C. Greenough, D.B. Bretttler, M. Schindler, S. Wildum, J.L. Sullivan, F. Kirchhoff, Virology **317**, 109 (2003)
13. M. Kondo, T. Shima, M. Nishizawa, K. Sudo, S. Iwamuro, T. Okabe, Y. Takebe, M. Imai, J. Infect. Dis. **192**, 56 (2005)
14. A. El Kharroubi, M. Martin, Mol. Cell Biol. **16**, 2958 (1996)
15. L. Pereira, K. Bentley, A. Peeters, M. Churchill, N. Deacon, Nucleic Acids Res. **28**, 663 (1999)
16. C. Vant Lint, C. Ann Amella, S. Emiliani, M. John, T. Jie, E. Verdin, J. Virol. **71**, 6113 (1997)
17. F. Kirchhoff, T. Greenough, M. Hamacher, J. Sullivan, R. Desrosiers, Virology **232**, 319 (1997)
18. G. Fang, H. Burger, C. Chappey, S. Rowland-Jones, A. Visosky, C.H. Chen, T. Moran, L. Townsend, M. Murray, B. Weiser, AIDS Res. Hum. Retroviruses **17**, 1395 (2001)
19. L. Zhang, Y. Huang, H. Yuan, B. Chen, D. Ho, J. Virol. **71**, 5608 (1997)
20. B. Rodés, C. Toro, E. Paxinos, E. Poveda, M. Martinez-Padial, J.M. Benito, V. Jiménez, T. Wrin, S. Basan, V. Soriano, AIDS **18**, 1109 (2004)
21. M. Naghavi, M. Estable, S. Schwartz, R. Roeder, A. Vahlne, J. Gen. Virol. **82**, 547 (2001)
22. K.A. Jones, J.T. Kadonaga, P.A. Luciw, R. Tjian, Science **232**, 755 (1986)
23. J.T. Kadonaga, K.R. Carner, F.R. Masiarz, R. Tjian, Cell **51**, 1079 (1987)
24. L. Zhang, Y. Huang, H. Yuan, B. Chen, J. Ip, D. Ho, J. Virol. **71**, 1651 (1997)
25. M.C. Estable, B. Bel, M. Hirst, I. Sadowski, J. Virol. 72, 6465 (1998)