doi: 10.1111/j.1744-313X.2007.00658.x

A three-base-deletion polymorphism in the upstream non-coding region of human interleukin 7 (IL-7) gene could enhance levels of IL-7 expression

H. Song,* E. E. Nakayama,* S. Likanonsakul,† C. Wasi,‡ A. Iwamoto§ & T. Shioda*

Summary

Interleukin 7 (IL-7) is a key factor in the survival, development and proliferation of B and T lymphocytes. Elevation of plasma IL-7 has been reported in several lymphopenia cases such as HIV-1 patients. After patients started to receive antiretroviral drugs and their CD4$^+$ cell counts had recovered, IL-7 in plasma decreased to normal levels. There are considerable variations in the levels of plasma IL-7 as well as the rate of CD4$^+$ T-cell restoration. Although pre-treatment plasma IL-7 levels have been shown to be prognostic for the rate of post-treatment CD4$^+$ T-cell restoration, the mechanisms responsible for the variations in plasma IL-7 and rate of CD4$^+$ T-cell restoration are still completely unknown. In the study here, we searched for genetic polymorphisms that might affect levels of IL-7 gene expression. For this purpose, we used 1658-bp PCR-amplified fragments of the IL-7 gene containing 1470 bp of the upstream non-coding region obtained from 151 Japanese and 234 Thai subjects. We found two novel human genetic polymorphisms in the upstream non-coding region of the IL-7 gene. The luciferase reporter assay demonstrated that one of those polymorphisms could increase the gene expression of IL-7. We speculate that this polymorphism, a three base ATC deletion just upstream of an out-of-frame ATG codon in the upstream non-coding region of the IL-7 gene, reduces the efficiency of translation from the upstream, out-of-frame ATG, resulting in increased translation efficiency from the authentic ATG of IL-7. Although the frequency of this allele is very low, it would be interesting to analyse this polymorphism in HIV-1-infected individuals with different rates of immune reconstitution after treatment with a highly active antiretroviral therapy.

* Department of Viral Infections, Research Institute for Microbial Diseases, Osaka University, Osaka, Japan, † Bamrasnaradura Institute, Nonthaburi, Thailand, ‡ Department of Microbiology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand and § Division of Infectious Diseases, Institute of Medical Science, University of Tokyo, Tokyo, Japan

Received 4 July 2006; revised 17 October 2006; accepted 26 November 2006

Correspondence: Tatsuo Shioda, Department of Viral Infections, Research Institute for Microbial Diseases, Osaka University, 3–1 Yamada-Oka, Suita, Osaka 565-0871, Japan. Tel: +81 6 6879 8346; Fax: +81 6 6879 8347; E-mail: shioda@biken.osaka-u.ac.jp

This work was supported by grants from the Human Health Foundation, the Ministry of Education, Culture, Sports, Science, and Technology, and the Ministry of Health, Labour and Welfare, Japan.
There is no conflict of interest.
Haihan Song, E-mail: hhsong@biken.osaka-u.ac.jp
Emi E. Nakayama, E-mail: emien@biken.osaka-u.ac.jp
Sirirat Likanonsakul, E-mail: siratlik@health.moph.go.th
Chantapong Wasi, E-mail: sicws@mahidol.ac.th
Aikichi Iwamoto, E-mail: aikichi@ims.u-tokyo.ac.jp

Introduction

Human interleukin 7 (IL-7) is a cytokine produced by stromal cells of the thymus and bone marrow (Wolf & Cohen, 1992; Heufler *et al.*, 1993; Sudo *et al.*, 1993) and has the capacity to induce growth of immature B lymphocytes (Namen *et al.*, 1988). Similarly, IL-7 contributes to the development, proliferation and homeostatic maintenance of T cells (Grabstein *et al.*, 1990; Plum *et al.*, 1996; Schluns *et al.*, 2000; Fry *et al.*, 2001). Human IL-7 gene located on chromosome 8q12–13, has six exons that distributed to more than 33-Kb of genomic DNA (Lupton *et al.*, 1990; Fry & Mackall, 2002). It is known that the IL-7 gene has no canonical core promoter sequence in the 5′ upstream region (Lupton *et al.*, 1990; Oshima *et al.*, 2004). Recently, it has been reported, however, that transcription start sites of the IL-7 gene are clustered within two distinct regions that are approximately 515 bp to 600 bp and 130 bp to 217 bp upstream from the translation initiation ATG codon (Oshima *et al.*, 2004). Moreover, the region −282 to −251 upstream from the initiation ATG codon contains an interferon regulatory factor element (IRF-E) and could thus up-regulate the transcription of the IL-7 gene upon stimulation with gamma interferon (IFN-γ) in human intestinal epithelial cells (Oshima *et al.*, 2004). This study also revealed the presence of several out-of-frame ATG codons with unknown function in the upstream non-coding region of the IL-7 gene (Oshima *et al.*, 2004).

With respect to HIV-1 infection, there is a reverse correlation between CD4$^+$ T-cell numbers and IL-7 plasma levels in HIV-1-infected patients (Llano *et al.*, 2001; Beq *et al.*, 2004; Kopka *et al.*, 2005). After these patients started to receive antiretroviral drugs and their CD4$^+$ T-cell counts had recovered, the elevated IL-7 in the plasma decreased to normal levels (Llano *et al.*, 2001). Furthermore, it is well known that there are considerable variations in the levels of plasma IL-7 as well as the rate of CD4 T-cell restoration after HIV-1 patients started to

receive antiretroviral drugs, and pre-treatment plasma IL-7 levels have been shown to be prognostic for the rate of post-treatment CD4 T-cell restoration (Beq *et al.*, 2004). However, knowledge of the molecular mechanisms controlling IL-7 gene expression remains very limited, and the mechanisms responsible for the variations in plasma IL-7 levels and rate of CD4 T-cell restoration among individuals are still completely unknown.

Human genetic polymorphisms have recently been shown to affect expression of the corresponding genes and to consequently modify the clinical course of several human diseases such as HIV-1 infection (Dean *et al.*, 1996; Michael *et al.*, 1997; Kostrikis *et al.*, 1998; Liu *et al.*, 1999; Nakayama *et al.*, 2000). We aimed to know the molecular mechanisms controlling variations in IL-7 gene expression among individuals. For this purpose, we searched for genetic polymorphisms that might affect levels of IL-7 gene expression in 1658-bp PCR-amplified fragments of the IL-7 gene containing 1470 bp of the upstream non-coding region, 9 bp of the first coding exon and 179 bp of the downstream intron, although there was no previous report on human genetic polymorphisms that alter the levels of IL-7 gene expression. We found two novel human genetic polymorphisms in the upstream non-coding region of the IL-7 gene, one of which could enhance IL-7 expression probably by reducing the efficiency of translation from an upstream, out-of-frame ATG that would result in diminished efficiency of translation from the downstream initiation ATG.

Materials and methods

Genotyping of IL-7 gene

Human genomic DNA was obtained from peripheral blood mononuclear cells of 52 unrelated non-HIV-1-infected and 99 HIV-1-infected Japanese, as well as 122 non-HIV-1-infected and 112 HIV-infected Thais, who provided written informed consent. Genomic regions of 1658 nucleotides containing 1470 nucleotides of the upstream non-coding region and the first exon and part of the intron of IL-7 were amplified by using the primer pair P1: 5′-TCCCTCCTCTTCCTTGTTC-3′ and P2: 5′-GGTTCAAGTGGCTATGTGC-3′. Polymerase chain reaction (PCR) was run for 40 cycles of denaturation at 94 °C for 30 s, annealing at 53 °C for 30 s and extension at 72 °C for 2 min. Fluorescence-based automated cycle sequencing of the PCR products was then carried out by an ABI 3100 using P1, P2 (mentioned previously), P3: 5′-TGCTGCATTTGGGCTGTAGA-3′, P4: 5′-TGGTTTTTCCTGCGGTGAT-3′ and P5: 5′-GGTCTGCAGGTTCAATCT-3′ as sequencing primers.

Luciferase reporter gene assays

NheI and NcoI-tagged DNA fragments, corresponding to the sequences spanning positions −632 to +3, −632 to −67 and −297 to +3 from the initiation ATG of the IL-7 gene, were inserted into the corresponding restriction enzyme cleavage sites of the pGL3-Basic Vector in order to fuse ATGs in the IL-7 gene directly to the firefly luciferase open reading frame (Promega, Madison, WI). Constructs carrying an ATC deletion at position −29 to −27 from the initiation ATG of IL-7 were generated by PCR-based in vitro mutagenesis using P6: 5′-GGCTAGCAGACGACTTGGCATCGTCC-3′ and P8: 5′-TGGACCATGGTCTGCGGGAGGCGGGCGTAGTCATGACCGC-3′ or P7: 5′-GGCTAGCAGATTGAACCTGCAGACCA-3′ and P8 (mentioned previously) as the respective primer pairs for the −632 to +3 or −297 to +3 upstream region of the IL-7 gene with ATC deletion. All constructs were verified for sequence authenticity. Four micrograms of the resultant constructs was transfected with DMRIE-C (Gibco/BRL, Gaithersburg, MD) into Jurkat ($CD4^+$ T-lymphocyte cell line) and U937 cells (monocytic cell line). Transfection efficiency was normalized by cotransfection with 0.2 μg of pRL-CMV vector, which expresses *Renilla* luciferase under the control of the cytomegalovirus immediate early promoter. When necessary, INF-γ (Peprotech, Rocky Hill, NJ) was added to the transfected cell culture at a final concentration of 50 ng mL^{-1} 5 h after transfection. The cells were harvested 40 h after transfection, and firefly and *Renilla* luciferase activities were determined according to the manufacturer's instructions (Dual-Luciferase Reporter Assay System, Promega) with a Luminometer Centro LB960 (Berthold, Bad Wildbad, Germany). Relative luciferase expression (fold increase) was calculated with the following equation: fold increase = (firefly luciferase activity of upstream region of IL-7 gene construct/*Renilla* luciferase activity) / (firefly luciferase activity of promoterless vector pGL3-Basic/*Renilla* luiferase activity).

Statistical analysis

The unpaired *t*-test was used.

Results

Polymorphisms in the upstream non-coding region of the IL-7 gene

We sequenced a 1658-bp PCR-amplified fragment of the IL-7 gene containing 1470 bp of the upstream non-coding region, 9 bp of the first coding exon and 179 bp of the downstream intron. Samples were obtained from 52 unrelated non-HIV-1-infected and 99 HIV-1-infected Japanese, as well as from 122 non-HIV-1-infected and 112 HIV-infected Thais. Polymorphisms were identified at two positions: an A to G substitution at position −485 and an ATC deletion at a position from −29 to −27 upstream from the open frame ATG codon of the IL-7 gene (Fig. 1). Frequencies of these two polymorphisms are summarized in Table 1. As for the A to G mutation at position −485, there was no difference in frequency of the G allele between HIV-1-infected and non-HIV-1-infected individuals. For the allele of the ATC deletion, two of the 99 HIV-1-infected Japanese carried this allele, but none of the Thais. There was no linkage disequilibrium between these two mutations.

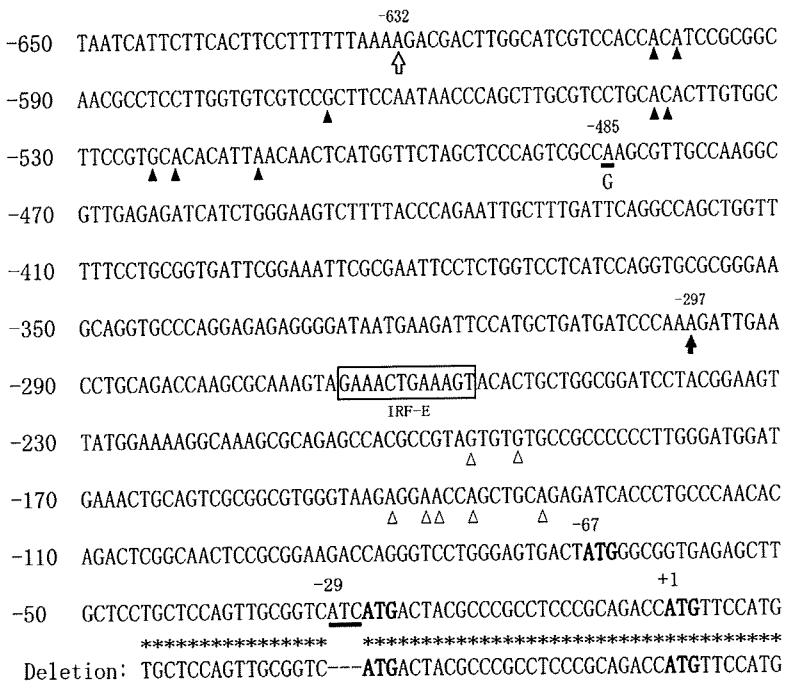


Figure 1. Fragment containing 650-bp of the upstream non-coding region and a part of the coding region of the IL-7 gene. Two polymorphisms, A to G at −485 and ATC deletion at −29 to −27 are underlined. The sequence with the ATC deletion is shown below the sequence without the ATC deletion. Asterisks denote sequence identity. Numbers denote positions from the initiation ATG of IL-7. An open arrow at −632 and a closed arrow at −297 denote the 5′ ends of the IL-7 upstream non-coding region inserted into reporter plasmids (see Figs 3 and 4). Triangles denote multiple transcription start sites that are clustered within the two distinct regions reported by Oshima *et al.* (2004). Open triangles denote transcription start sites specifically activated by IFN-γ. An open square denotes IRF-E (Oshima *et al.*, 2004).

In addition, calculation of nucleotide diversity in the 1470-bp fragment of the upstream non-coding region of IL-7 genes in all Japanese and Thai subjects showed 1.4×10^{-5} in Japanese and 0.9×10^{-5} in Thais. These results suggested that the human IL-7 gene has a highly conserved upstream non-coding region.

Roles of ATGs in the upstream non-coding region of IL-7 gene

In the upstream non-coding region, there are several out-of-frame ATGs (Fig. 1), with even the shortest transcript starting from position −130 containing two out-of-frame ATGs in the upstream non-coding region (Fig. 1). Because one of those out-of-frame ATGs occurred just downstream of the ATC deletion described previously, we then investigated roles of these upstream ATGs in IL-7 gene expression. For this purpose, we constructed a reporter plasmid in which the luciferase open reading frame was fused with the upstream ATG under the control of the upstream region of the IL-7 gene. As shown in Fig. 2, when a reporter plasmid carrying the region from position −632 to the authentic translation initiation ATG codon was transiently transfected into Jurkat or U937 cells, a significant increase in luciferase activity was observed, compared with the pGL3-basic vector employed as a control reporter plasmid in either cell, confirming a previous observation (Oshima *et al.*, 2004). When ATG at position −69 to −67 was fused with the luciferase open reading frame, luiferase activity became greatly enhanced (Fig. 2). These results indicated that the upstream AUG in IL-7 mRNA was more efficiently used for expression than the authentic AUG, and suggested that presence of upstream AUGs in IL-7 mRNA can be expected to reduce IL-7 translation levels.

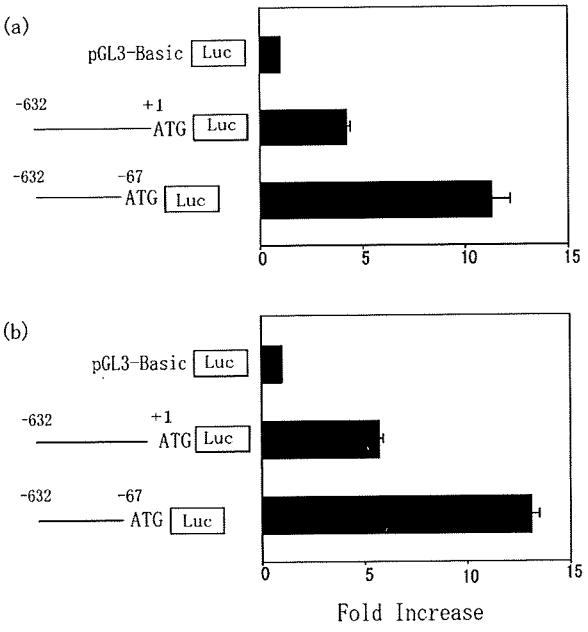


Figure 2. Luciferase activity mediated by the upstream non-coding region of the IL-7 gene. Jurkat (a) or U937 cells (b) were transfected with the plasmids indicated. The fold increase of each construct is represented by a bar. Data represent three independent experiments with similar results. Error bars show actual fluctuations among measurements of fold increase in four clones of each construct.

ATC deletion could affect the gene expression of IL-7

As mentioned previously, the ATC deletion occurred just upstream of the ATG located at position −26 to −24 (Fig. 1). Kozak previously reported that ACCATGG is

Table 1. Allele and genotype frequencies of A to G mutation at −485 and deletion mutation at −29 to −27 in HIV-1-infected and non-HIV-1-infected Japanese and Thai people

	Japan				Thailand			
	HIV-1-Infected *n* (%)	Non-HIV-1-Infected *n* (%)	HIV-1-Infected *n* (%)	Non-HIV-1-Infected *n* (%)	HIV-1-Infected *n* (%)	Non-HIV-1-Infected *n* (%)	HIV-1-Infected *n* (%)	Non-HIV-1-Infected *n* (%)
Allele			Genotype		Allele		Genotype	
A	196 (99.0)	102 (98.1)	AA	97 (98.0)	50 (96.2)	A	221 (98.7)	241 (98.8)	AA	109 (97.3)	119 (97.5)
G	2 (1.0)	2 (1.9)	AG	2 (2.0)	2 (3.8)	G	3 (1.3)	3 (1.2)	AG	3 (2.7)	3 (2.5)
			GG	0 (0)	0 (0)				GG	0 (0)	0 (0)
Total	198	104		99	52	Total	224	244		112	122
Allele			Genotype			Allele			Genotype		
W[a]	196 (99.0)	104 (100)	WW	97 (98.0)	52 (100)	W[a]	224 (100)	244 (100)	WW	112 (100)	122 (100)
D[b]	2 (1.0)	0 (0)	WD	2 (2.0)	0 (0)	D[b]	0 (0)	0 (0)	WD	0 (0)	0 (0)
			DD	0 (0)	0 (0)				DD	0 (0)	0 (0)
Total	198	104		99	52	Total	224	244		112	122

[a] W denotes the wild type at −29 to −27.

[b] D denotes deletion at −29 to −27.

the optimal sequence for translation initiation of pre-proinsuline by eukaryotic ribosomes and that substitution of G for A at position −3 (3-bp upstream from the ATG codon) reduced translation efficiency (Kozak, 1986). The A to G substitution at position −3 of an upstream, out-of-frame ATG condon also reportedly diminished translation from the corresponding upstream ATG and consequently increased translation from the authentic downstream ATG (Kozak, 1986). In the case of the human IL-7 gene, the sequence surrounding the ATG at position −26 to −24 is ATCATG but the ATC deletion observed in our study converted it into GTCATG (Fig. 1). These data indicate that the ATC deletion altered the A at position −3 into G (Fig. 1), thus hypothetically reducing translation efficiency from the upstream ATG at position −26 to −24 and increasing translation from the authentic IL-7 ATG. We therefore decided to test experimentally whether the ATC deletion polymorphism actually affected levels of expression from the authentic IL-7 ATG.

We constructed a reporter plasmid containing the upstream non-coding region from −632 to +3 with the ATC deletion and compared its luciferase activity with that of the wild-type version. As shown in Fig. 3(a), the reporter activity of the deletion mutant was approximately 30% higher than that of the wild-type plasmid. We also generated shorter versions of the wild type as well as mutant constructs carrying the upstream non-coding region from −297 to +3, which spans the minimal promoter region containing IRF-E (−268 to −257) (Oshima *et al.*, 2004). Again, the reporter activity of the deletion mutant was approximately 30% higher than that of the wild-type plasmid (Fig. 3b). We repeated the same experiments by using monocytic U937 cells. Here too, luciferase activity in the deletion mutant was approximately 25% higher than that in the corresponding wild-type plasmid when the upstream non-coding region of −632 to +3 was used (Fig. 4a). An approximately 20% increase in luciferase activity was observed in the deletion mutant when the upstream non-coding region −297 to +3 was used (Fig. 4c). It is known that INF-γ is capable of up-regulating the gene expression of IL-7 in intestinal epithelial cells through the IRF-E in the region −268 to −257 from the initiation ATG codon (Oshima *et al.*, 2004). As shown in Fig. 4(b,d), the addition of INF-γ to the transfected cells in fact did augment luciferase activity in U937 cells. Moreover, the deletion mutant exhibited significantly higher luciferase activity than the wild-type constructs (Fig. 4b,d). These results clearly indicate that ATC deletion in the upstream non-coding region resulted in higher expression from the authentic IL-7 ATG.

Discussion

In the study reported here, we demonstrated that an out-of-frame ATG in the upstream non-coding exon of IL-7 gene was more efficiently used for expression than the authentic ATG of IL-7 gene. We also found a naturally occurring ATC deletion polymorphism at position −29 to

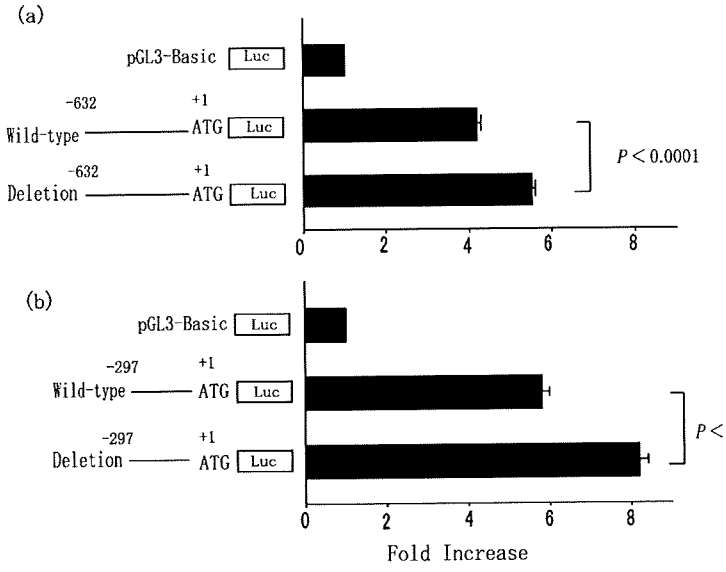


Figure 3. Luciferase activity mediated by the upstream non-coding region of the wild-type and ATC deletion in Jurkat cells. (a) A region from position −632 to +3 of the upstream non-coding region of the wild-type and ATC deletion. (b) A region from position −297 to +3 of the upstream non-coding region of the wild-type and ATC deletion. Data represent three independent experiments with similar results. Error bars show actual fluctuations among measurements of fold increase in four clones of each construct. *P* values for differences in fold increase are shown.

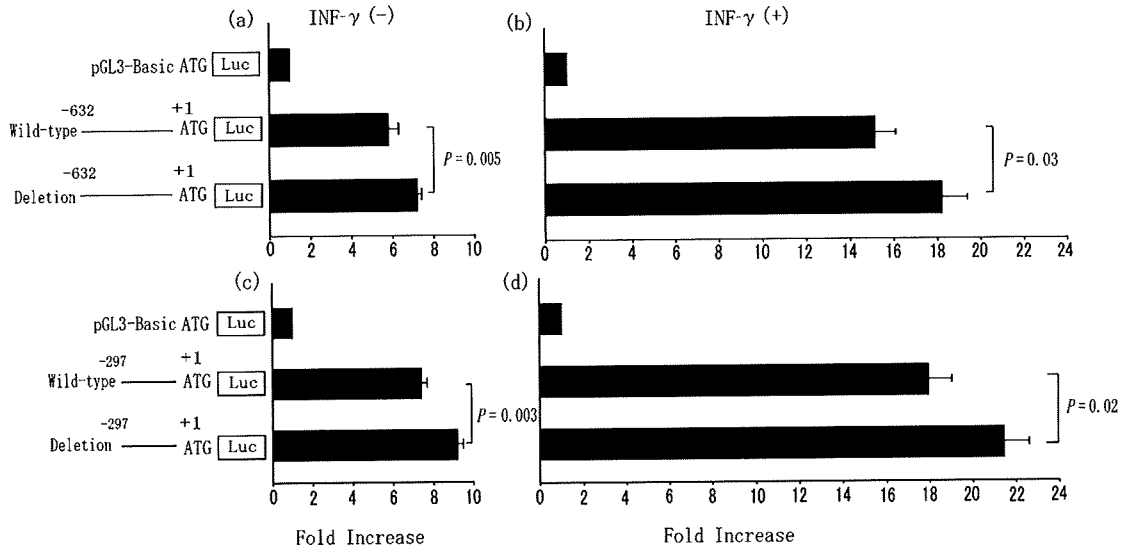


Figure 4. Luciferase activity mediated by the upstream non-coding region of the wild-type and ATC deletion in U937 cells. (a) A region from position −632 to +3 of the upstream non-coding region of the wild-type and ATC deletion without INF-γ. (b) A region from position −632 to +3 of the upstream non-coding region of the wild-type and ATC deletion with INF-γ. (c) A region from position −297 to +3 of the upstream non-coding region of the wild-type and ATC deletion without INF-γ. (d) A region from position −297 to +3 of the upstream non-coding region of the wild-type and ATC deletion with INF-γ. Data represent three independent experiments with similar results. Error bars show actual fluctuations among measurements of fold increase in four clones of each construct. *P* values for differences in fold increase are shown.

−27 in the upstream non-coding exon next to one of the upstream ATGs. This polymorphism was found to be capable of increasing the expression from the authentic IL-7 ATG in Jurkat T cell and U937 monocytic cell lines. This is the first time human genetic polymorphism has been identified that is supposed to affect expression of a protein by changing the translation efficiency from the out-of-frame AUG in the upstream non-coding region of mRNA.

There are a few precedents for a human genetic polymorphism near the initiation ATG codon affecting translation efficiency. A single nucleotide polymorphism (SNP) that switches C to T at position −1 upstream from the open frame ATG codon in the human annexin V gene has been found to increases translation efficiency and plasma levels of annexin V, and to decreases the risk of early myocardial infarction (Gonzalez-Conejero *et al.*, 2002). Also, an SNP that switches G to T at position −3 upstream from the open frame ATG codon of the BRCA1 gene in sporadic breast cancer causes down-modulation of translation efficiency (Signori *et al.*, 2001). Moreover, a mutation of G into A at +4 downstream from the open frame ATG codon of the human androgen receptor gene observed in

a family with partial androgen insensitivity syndrome can reduce the efficiency of protein translation (Choong *et al.*, 1996). Finally, it has been reported recently that a Graves'-disease–associated SNP that substitutes T for C at position −1 upstream from the open frame ATG codon of the CD40 gene enhances translation and could predispose to disease (Jacobson *et al.*, 2005). However, all these SNPs are located near the authentic translation initiation ATG codon of the proteins and directly affect the translation efficiency from the open frame ATG codon. In the case of the IL-7 gene, however, the mutation is located at −29 to −27 upstream from the open frame ATG codon, rather than near the authentic translation initiation codon. Nevertheless, it could up-regulate the IL-7 gene expression probably by changing the translation efficiency from the upstream, out-of-frame ATG codon. Our data showed that the consensus sequence for translation initiation is important, not only for the open-frame initiation ATG codon, but also for the upstream, out-of-frame ATG that is thought to reduce translation efficiency from the downstream initiation ATG. Similar to mutations within the consensus sequence of the open-frame initiation ATG codon, nucleotide substitution within the consensus sequence of the upstream, out-of-frame ATG can also modulate translation efficiency.

Our data also showed that the 1470-bp upstream non-coding region of the IL-7 gene exhibited extremely low levels of diversity in both Japanese and Thai populations. Also, no non-synonymous polymorphism has yet been identified in the IL-7 coding region. The reason for the low levels of diversity of the IL-7 gene is not clear at present, but is probably the result of its importance for the survival, development and proliferation of B and T cells. Experiments with IL-7 deficient mice proved that IL-7 is a non-redundant cytokine (von Freeden-Jeffry *et al.*, 1995). It is therefore reasonable to assume that low levels of diversity of the upstream non-coding region of the IL-7 gene that regulates the transcription of this gene are needed to provide a stable condition for IL-7 production.

In conclusion, we have identified a polymorphism in the upstream non-coding region of the IL-7 gene that could up-regulate gene expression. Although the frequency of this allele is very low in Japan and Thailand, it would be interesting to analyse this polymorphism in HIV-1-infected individuals with different rates of immune reconstitution after treatment with a highly active antiretroviral therapy. It would be important to analyse this polymorphism in other ethnic groups. On the other hand, an elevation of plasma IL-7 has also been reported in lymphopenia cases, including patients undergoing bone marrow transplantation or chemotherapy for cancer, or patients with idiopathic $CD4^+$ lymphopenia (Fry & Mackall, 2005; Bolotin *et al.*, 1999). It would also be of interest to analyse this polymorphism in those patients.

Acknowledgements

We thank J Sakuragi and S Sakuragi for their helpful discussions, S Bandou for technical assistance and N Teramoto for help. We also thank the Japanese and Thais who kindly donated blood for this study.

References

Beq, S., Rannou, M.T., Fontanet, A., Delfraissy, J.F., Theze, J. & Colle, J.H. (2004) HIV infection: pre-highly active antiretroviral therapy IL-7 plasma levels correlate with long-term CD4 cell count increase after treatment. *AIDS*, **18**, 563.

Bolotin, E., Annett, G., Parkman, R. & Weinberg, K. (1999) Serum levels of IL-7 in bone marrow transplant recipients: relationship to clinical characteristics and lymphocyte count. *Bone Marrow Transplantation*, **23**, 783.

Choong, C.S., Quigley, C.A., French, F.S. & Wilson, E.M. (1996) A novel missense mutation in the amino-terminal domain of the human androgen receptor gene in a family with partial androgen insensitivity syndrome causes reduced efficiency of protein translation. *Journal of Clinical Investigation*, **98**, 1423.

Dean, M., Carrington, M., Winkler, C., Huttley, G.A., Smith, M.W., Allikmets, R. *et al.* (1996) Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. *Science*, **273**, 1856.

von Freeden-Jeffry, U., Vieira, P., Lucian, L.A., McNeil, T., Burdach, S.E. & Murray, R. (1995) Lymphopenia in interleukin (IL) -7 gene-deleted mice identifies IL-7 as a nonredundant cytokine. *Journal of Experimental Medicine*, **181**, 1519.

Fry, T.J. & Mackall, C.L. (2002) Interleukin-7: from bench to clinic. *Blood*, **99**, 3892.

Fry, T.J. & Mackall, C.L. (2005) The many faces of IL-7: from lymphopoiesis to peripheral T cell maintenance. *Journal of Immunology*, **174**, 6571.

Fry, T.J., Connick, E., Falloon, J., Lederman, M.M., Liewehr, D.J., Spritzler, J. *et al.* (2001) A potential role for interleukin-7 in T-cell homeostasis. *Blood*, **97**, 2983.

Gonzalez-Conejero, R., Corral, J., Roldan, V., Martinez, C., Marin, F., Rivera J. *et al.* (2002) A common polymorphism in the annexin V Kozak sequence (−1C>T) increases translation efficiency and plasma levels of annexin V, and decreases the risk of myocardial infarction in young patients. *Blood*, **100**, 2081.

Grabstein, K.H., Namen, A.E., Shanebeck, K., Voice, R.F., Reed, S.G. & Widmer, M.B. (1990) Regulation of T cell proliferation by IL-7. *Journal of Immunology*, **144**, 3015.

Heufler, C., Topar, G., Grasseger, A., Stanzl, U., Koch, F., Romani, N. *et al.* (1993) Interleukin 7 is produced by murine and human keratinocytes. *Journal of Experimental Medicine*, **178**, 1109.

Jacobson, E.M., Concepcion, E., Oashi, T. & Tomer, Y. (2005) A Graves' disease-associated Kozak sequence single-nucleotide polymorphism enhances the efficiency of CD40 gene translation: a case for translational pathophysiology. *Endocrinology*, **146**, 2684.

Kopka, J., Mecikovsky, D., Aulicino, P.C., Mangano, A.M., Rocco, C.A. & Bologna, R. (2005) High IL-7 plasma levels may induce and predict the emergence of HIV-1 virulent strains in pediatric infection. *Journal of Clincal Virology*, **33**, 237.

Kostrikis, L.G., Huang, Y., Moore, J.P., Wolinsky, S.M., Zhang, L., Guo, Y. *et al.* (1998) A chemokine receptor CCR2 allele delays HIV-1 disease progression and is associated with a CCR5 promoter mutation. *Nature Medicine*, **4**, 350.

Kozak, M. (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell*, **44**, 283.

Liu, H., Chao, D., Nakayama, E.E., Taguchi, H., Goto, M., Xin, X. *et al.* (1999) Polymorphism in RANTES chemokine promoter affects HIV-1 disease progression. *Proceedings of the National Academy of Sciences USA*, **96**, 4581.

Llano, A., Barretina, J., Gutierrez, A., Blanco, J., Cabrera, C. & Clotet, B. (2001) Interleukin-7 in plasma correlates with CD4 T-cell depletion and may be associated with emergence of syncytium-inducing variants in human immunodeficiency virus type 1-positive individuals. *Journal of Virology*, **75**, 10319.

Lupton, S.D., Gimpel, S., Jerzy, R., Brunton, L.L., Hjerrild, K.A., Cosman, D. *et al.* (1990) Characterization of the human and murine IL-7 genes. *Journal of Immunology*, **144**, 3592.

Michael, N.L., Louie, L.G., Rohrbaugh, A.L., Schultz, K.A., Dayhoff, D.E., Wang, C.E. *et al.* (1997) The role of CCR5 and CCR2 polymorphisms in HIV-1 transmission and disease progression. *Nature Medicine*, **3**, 1160.

Nakayama, E.E., Hoshino, Y., Xin, X., Liu, H., Goto, M., Watanabe, N. *et al.* (2000) Polymorphism in the interleukin-4 promoter affects acquisition of human immunodeficiency virus type 1 syncytium-inducing phenotype. *Journal of Virology*, 2000 (74), 5452.

Namen, A.E., Lupton, S., Hjerrild, K., Wignall, J., Mochizuki, D.Y., Schmierer, A. *et al.* (1988) Stimulation of B-cell progenitors by cloned murine interleukin-7. *Nature*, **333**, 571.

Oshima, S., Nakamura, T., Namiki, S., Okada, E., Tsuchiya, K., Okamoto, R. *et al.* (2004) Interferon regulatory factor 1 (IRF-1) and IRF-2 distinctively up-regulate gene expression and production of interleukin-7 in human intestinal epithelial cells. *Molecular and Cellular Biology*, **24**, 6298.

Plum, J., De Smedt, M., Leclercq, G., Verhasselt, B. & Vandekerckhove, B. (1996) Interleukin-7 is a critical growth factor in early human T-cell development. *Blood*, **88**, 4239.

Schluns, K.S., Kieper, W.C., Jameson, S.C. & Lefrancois, L. (2000) Interleukin-7 mediates the homeostasis of naive and memory CD8 T cells in vivo. *Nature Immunology*, **1**, 426.

Signori, E., Bagni, C., Papa, S., Primerano, B., Rinaldi, M., Amaldi, F. *et al.* (2001) A somatic mutation in the 5′UTR of BRCA1 gene in sporadic breast cancer causes down-modulation of translation efficiency. *Oncogene*, **20**, 4596.

Sudo, T., Nishikawa, S., Ohno, N., Akiyama, N., Tamakoshi, M. & Yoshida, H. (1993) Expression and function of the interleukin 7 receptor in murine lymphocytes. *Proceedings of the National Academy of Sciences, USA*, **90**, 9125.

Wolf, S.S. & Cohen, A. (1992) Expression of cytokines and their receptors by human thymocytes and thymic stromal cells. *Immunology*, **77**, 362.

Immunogenetics (2007) 59:511–515
DOI 10.1007/s00251-007-0217-7

BRIEF COMMUNICATION

Wild type and *H43Y* variant of human *TRIM5α* show similar anti-human immunodeficiency virus type 1 activity both in vivo and in vitro

Emi E. Nakayama · Wassila Carpentier · Dominique Costagliola · Tatsuo Shioda · Aikichi Iwamoto · Patrice Debre · Kazuhisa Yoshimura · Brigitte Autran · Shuzo Matsushita · Ioannis Theodorou

Received: 29 November 2006 / Accepted: 21 March 2007 / Published online: 4 April 2007

Abstract Polymorphisms in human genes have been shown to affect the rate of disease progression to acquired immune deficiency syndrome in human immunodeficiency virus type 1 (HIV-1)-infected individuals. Recently, tripartite motif 5α (TRIM5α) was identified as a factor that confers resistance to HIV-1 infection in Old World monkey cells. Subsequently, Sawyer et al. (Curr Biol 16:95–100, 2006) reported a single nucleotide polymorphism (H43Y) in the human *TRIM5α* gene and TRIM5α protein with 43Y was found to lose its ability to restrict HIV-1. In the present study, we reevaluated effects of this allele on in vitro anti-HIV-1 activity as well as on HIV-1 disease progression in European and Asian cohorts of HIV-1-infected individuals. Our epidemiological and molecular biological findings clearly indicate H43Y has a very minor effect on anti-HIV-1 activity of TRIM5α, suggesting that this allele is immaterial, at least in HIV-1-infected Europeans and Asians.

Keywords TRIM5α · H43Y · RING domain · Polymorphism · HIV-1 disease progression · Anti-HIV-1 activity

E. E. Nakayama · T. Shioda (✉)
Department of Viral Infections, Research Institute for Microbial Diseases, Osaka University,
3-1 Yamada-oka,
Suita-shi, Osaka 565-0871, Japan
e-mail: shioda@biken.osaka-u.ac.jp

W. Carpentier · P. Debre · B. Autran · I. Theodorou
Laboratoire Central d'Immunologie Cellulaire et Tissulaire, Hôpital Pitié Salpetrière et INSERM UR543 Bâtiment CERVI, Paris, France

D. Costagliola
INSERM U 720, Epidemiologie clinique et therapeutique de l'infection a VIH, Université Pierre et Marie Curie-Paris 6, Paris, France

A. Iwamoto
Division of Infectious Diseases, Institute of Medical Science, University of Tokyo,
Tokyo, Japan

K. Yoshimura · S. Matsushita
Division of Clinical Retrovirology and Infectious Diseases, Center for AIDS Research, Kumamoto University,
Kumamoto, Japan

Human immunodeficiency virus type 1 (HIV-1) has a very narrow host range limited to humans and chimpanzees. In experiments, HIV-1 does not infect Old World monkeys, such as rhesus and cynomolgus monkeys. Recently, the screening of a rhesus monkey cDNA library identified tripartite motif 5 (TRIM5) as a factor that confers resistance to HIV-1 infection (Stremlau et al. 2004). Shortly afterwards, TRIM5α of the African green monkey (AGM), another Old World monkey, was also shown to restrict HIV-1 infection (Hatziioannou et al. 2004; Keckesova et al. 2004; Nakayama et al. 2005), while human TRIM5α reportedly restrict HIV-1 only weakly but potently restrict N-tropic murine leukemia virus (N-MLV; Hatziioannou et al. 2004; Keckesova et al. 2004; Perron et al. 2004; Yap et al. 2004). TRIM5α is composed of two zinc-finger (RING and B-box), coiled-coil, and SPRY (B30.2) domains.

HIV-1 infection in humans is generally characterized by a long-term, chronic disease course gradually progressing to acquired immune deficiency syndrome (AIDS). Polymor-

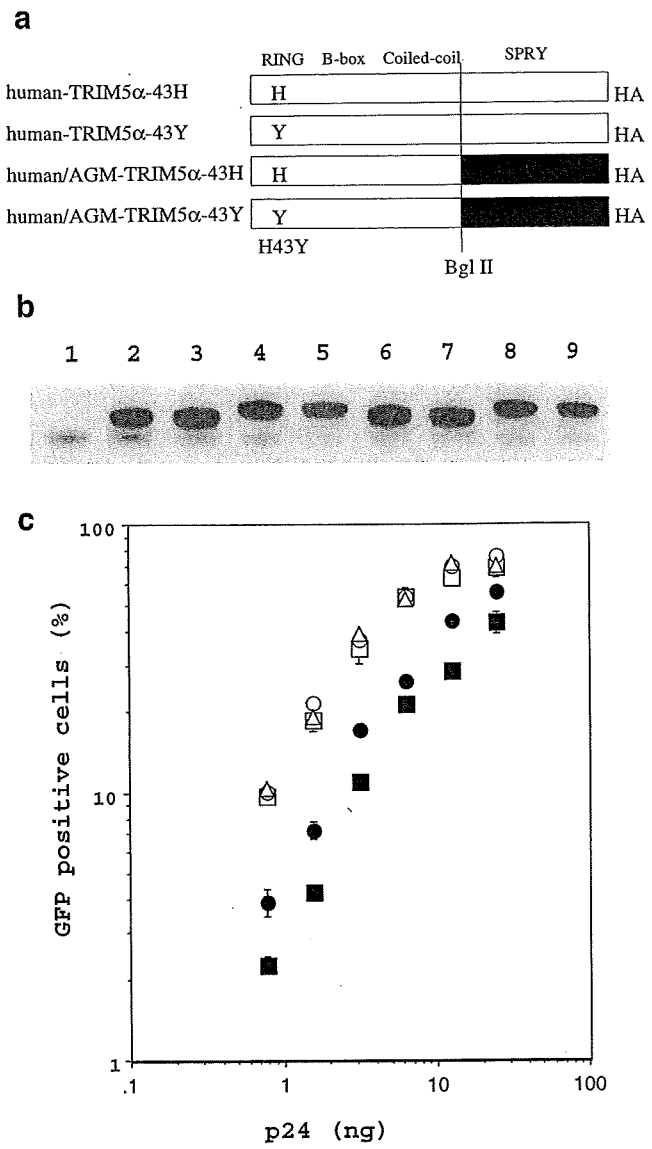


Fig. 1 **a** Schematic representation of TRIM5α fused with HA-tag. Four domains of TRIM5α are shown *at the top*. *Black* and *white bars* denote human and AGM sequences, respectively. A Bgl II site was used to swap SPRY domains between human and AGM TRIM5α. *H* or *Y* denotes the amino acid residue at the 43rd position. **b** Expression levels of HA-tagged TRIM5α proteins. C143 cells were transfected with an empty pCEP4 plasmid (*lane 1*) or pCEP4 carrying human-TRIM5α-43Y (*lanes 2* and *3*), human/AGM-TRIM5α-43Y (*lanes 4* and *5*), human-TRIM5α-43H (*lanes 6* and *7*), human/AGM-TRIM5α-43H (*lanes 8* and *9*) protein and hygromycin-resistant cells were selected for 14 days. One million cells of each transformant were lysed for immunoprecipitation with an anti-HA antibody. Two independent clones for each construct were evaluated for expression levels. **c** C143 cells expressing human-TRIM5α-43H (*open squares*), human-TRIM5α-43Y (*open circles*), human/AGM-TRIM5α-43H (*closed squares*), human/AGM-TRIM5α-43Y (*closed circles*), or cells transfected with an empty vector (*open triangles*) were exposed to the indicated p24 amounts of GFP-expressing HIV-1 vector. GFP-positive cells were counted with a flowcytometry (FACScan, Beckton Dickinson). *Error bars* indicated actual fractures of two independent cell cultures derived from independent clones. Representative data from three independent experiments are shown

phisms in human *CCR5* and other genes reportedly affect the rate of disease progression to AIDS (Kasper et al. 2005). Regarding the human *TRIM5α* gene, Sawyer et al. (2006) reported a common histidine-to-tyrosine polymorphism at the 43rd amino acid residue (H43Y) of human *TRIM5α*. This single nucleotide polymorphism (SNP) locates in the RING domain, and TRIM5α protein with H43Y was found to lose its ability to restrict HIV-1. This SNP was also shown to greatly reduce the ability of TRIM5α to restrict N-MLV. On the other hand, Speelmon et al. (2006) sequenced the *TRIM5α* gene from 110 HIV-1-infected and 96 exposed-seronegative European Americans and found 48 SNPs in their *TRIM5α* genes. However, they did not observe any association between H43Y polymorphism in HIV-1-infected subjects and their set-point viral load after acute infection. Furthermore, they detected no difference in in vitro HIV-1 susceptibility of CD4+ cells between 43Y homozygote and the wild type. Sawyer et al. (2006) and Speelmon et al. (2006) thus came to opposite conclusions; the former suggested that 43Y incapacitates even the modest human TRIM5α resistance to HIV-1 infection, while the latter showed no difference between 43H and 43Y.

After that, two more groups published their observations. Goldschmidt et al. (2006) have published their analysis on Caucasians in Swiss cohort and reported lack of association of H43Y with rapid progression to AIDS in HIV-1-infected individuals. In this report, they showed that HeLa cells stably transduced with human TRIM5α with 43Y do not differ from those with 43H in susceptibility to HIV-1 infection, whereas the 43Y variant failed to restrict N-MLV. Javanbakht et al. (2006) have published their analysis on large number of European Americans and African Americans and reported lack of any significant associations of TRIM5α SNPs with different rate of disease progression in HIV-1-infected individuals, although they found a controversial protective effect of H43Y against HIV-1 transmission only in African Americans but not in European Americans. They showed the 43Y human TRIM5α was less effective in restricting HIV-1 as well as N-MLV infections in vitro. Those reports agreed that 43Y variant failed to restrict N-MLV. However, the effect of H43Y substitution on HIV-1 restriction was not consistent among four reports. To make an addition to this debate, we conducted molecular biological and epidemiological studies in H43Y allele.

To reevaluate the effects of H43Y on in vitro anti-HIV-1 activity of TRIM5α, we first established stable cell lines expressing recombinant TRIM5α proteins. An expression plasmid carrying a hygromycin-resistant gene (pCEP4, Invitrogen) and hemagglutinin (HA)-tagged TRIM5α genes with 43H or 43Y (Fig. 1a) were introduced into CD4-negative human osteosarcoma cell line C143 and hygromycin-resistant cells were selected. Equal levels of HA-tagged TRIM5α expression were detected in those cell lines by

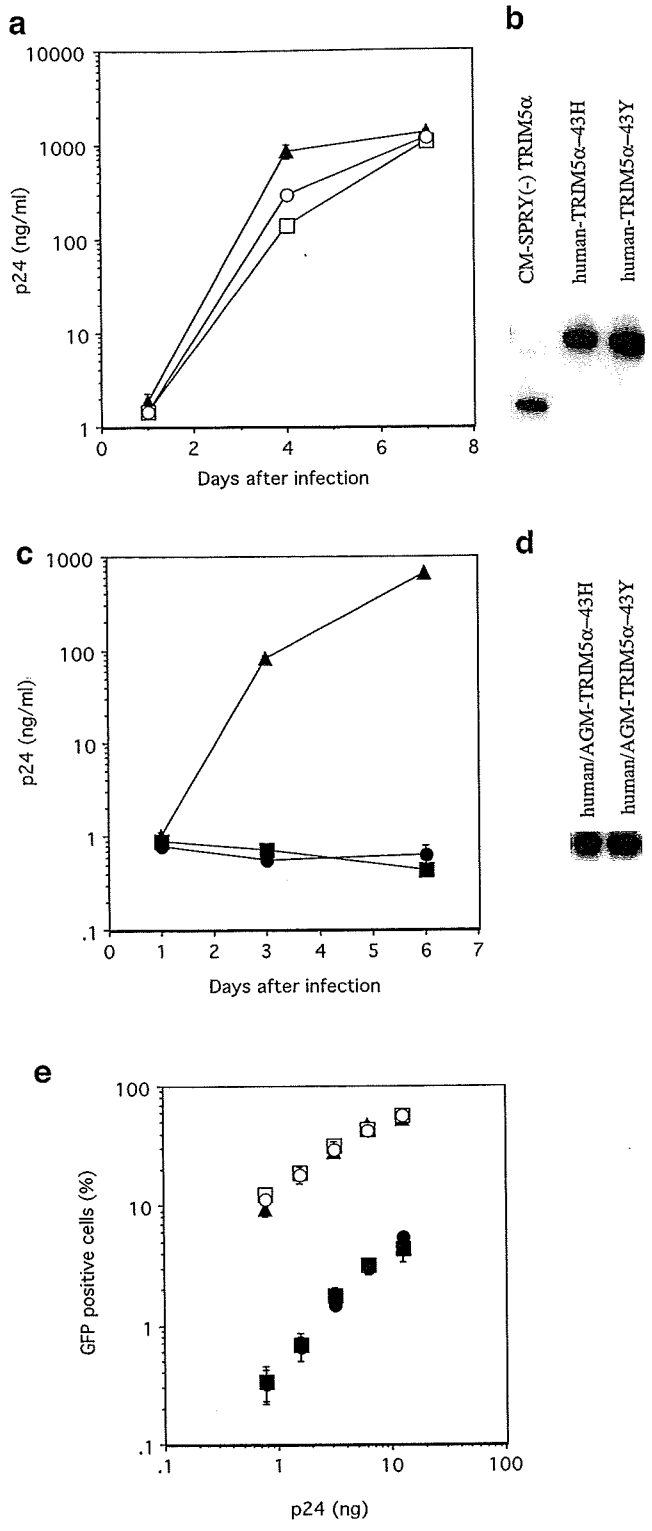
using immunoprecipitation followed by Western blot analysis (Nakayama et al. 2005; Fig. 1b). Serially diluted vesicular stomatitis virus (VSV)-pseudotyped HIV-1 vectors encoding green fluorescent protein (GFP) were then inoculated into the TRIM5α expressing cells, and infected cells were counted by flow-cytometry 40 h after infection. As shown in Fig. 1c, there was no difference in anti-HIV-1 activity among empty vector, human TRIM5α with 43H and that with 43Y, probably because human TRIM5α did not show any anti-HIV-1 effect in C143 cells.

Several recombinant studies of human and monkey TRIM5α revealed that the determinant of the species-specific restriction of HIV-1 lies in the SPRY domain of monkey TRIM5α (Nakayama et al. 2005, 2006; Perez-Caballero et al. 2005; Sawyer et al. 2005; Stremlau et al. 2005; Yap et al. 2005). To enhance the weak anti-HIV-1 activity of human TRIM5α, we introduced H43Y SNP in a chimeric version of TRIM5α, which carried part of the SPRY domain of AGM-TRIM5α and RING, B-box, and coiled-coil domains of human TRIM5α (Fig. 1a). Equal levels of TRIM5α expression were detected (Fig. 1b), and both chimeric TRIM5αs with 43H and 43Y showed potent anti-HIV-1 activity. Although there was a small increase in HIV-1-infected cells in transfectants with 43Y TRIM5α compared with those with 43H, we did not observe any complete loss of anti-HIV-1 activity for this variant.

To evaluate the effect of H43Y on multiple replication of HIV-1, we constructed recombinant Sendai viruses (SeVs) expressing human TRIM5α with 43H or that with 43Y, or their chimeric versions. There were 10^5 cells of human T cell line MT4 infected with recombinant SeV expressing human TRIM5α with 43H or that with 43Y at a multiplicity of infection at 10 plaque forming units per cell. Nine hours after infection, 20 ng of p24 of HIV-1 NL43 strain was challenged, and culture supernatants were periodically assayed for the levels of p24 by enzyme-linked immunosorbent assay (ZeptoMetrix). In this assay, we can observe multiple replications of HIV-1 and weak anti-HIV-1 activity of human TRIM5α can be amplified. As expected, both human TRIM5αs with 43H and 43Y showed weak but apparent anti-HIV-1 activity, although there was a small increase in HIV-1 titer in cells infected with SeV expressing human TRIM5α with 43Y compared with those infected with SeV expressing human TRIM5α with 43H (Fig. 2a). Equal levels of TRIM5α expressions were detected in those SeV-infected cells (Fig. 2b). When we used SeVs expressing human/AGM chimeric TRIM5α, both TRIM5α with 43H and 43Y completely suppressed HIV-1 replication (Fig. 2c). Again, equal levels of TRIM5α expressions were detected in those SeV infected cells (Fig. 2d).

To exclude the possible effect of endogenous human TRIM5α, we then used TK-ts13 cells, a derivative of baby hamster kidney cell. As HIV-1 cannot complete the late step of its replication in rodent cells, serially diluted VSV-pseudotyped HIV-1 vectors encoding GFP were inoculated into the cells infected with SeV expressing TRIM5αs. GFP-positive cells were counted 40 h after infection. As shown in Fig. 2c, there was no difference in anti-HIV-1 activity among cynomolgus monkey TRIM5α lacking SPRY domain [CM-SPRY(−)TRIM5α], human TRIM5α with 43H, and that with 43Y. Both chimeric TRIM5αs with 43H and 43Y showed potent anti-HIV-1 activity, and there was no difference in anti-HIV-1 activity between chimeric TRIM5α with 43H and that with 43Y. These results indicate that H43Y exerts only a minor effect on the anti-HIV-1 activity of TRIM5α protein. They are in contrast with those reported by Sawyer et al. (2006), who found the 43Y SNP completely abolished the anti-HIV-1 activity of human TRIM5α. Although the reason for this discrepancy is not clear at present, differences in the expression systems used may be involved.

To evaluate the effects of H43Y polymorphism on anti-HIV-1 activity of human TRIM5α in Asian population, 49 HIV-1-infected Japanese subjects with different rates of disease progression were analyzed. Of the 49 patients, 21 were long-term non-progressors (LTNPs). They are all hemophiliacs and infected through contaminated blood products before 1985, and their CD4 counts were over 200 cells/μl without highly active anti-retroviral therapy until 2001. The LTNPs included ten cases with undetectable viral load without any kinds of anti-retroviral therapy by 2001. The remaining 28 cases were standard progressors (SPs) comprising 15 hemophiliacs, 9 homosexual, and 4 heterosexual cases. The homosexual and heterosexual cases were infected with HIV-1 after 1985. Among 28 SPs, 13 died of AIDS before 1999, and 15 developed AIDS before 1996. RING and B-box region was polymerase chain reaction (PCR)-amplified from genomic DNA by using primer pair forward (5′-TCAGGTCTATCATGACAAGG CAG-3′) and reverse (5′-GGCAGGAGCAGTGGAATA GC-3′). Genotypes of the 43rd position were determined by direct sequencing of the resultant 542-bp PCR product with forward primer. Among 21 LTNPs, one subject was homozygous for 43Y allele, five were heterozygous, and 15 were homozygous for the wild type. Of the 28 SPs, one subject was homozygous for 43Y, five were heterozygous, and 22 were homozygous for the wild type. There was no statistically significant difference in 43Y allele frequencies between LTNPs and SPs (0.167, seven out of 42 chromosomes vs 0.125, seven out of 56 chromosomes $p=0.77$, Yates chi square test). As the number of HIV-1 infected patients studied here was relatively small, we performed statistical simulation with ten times more subject numbers. Nevertheless, difference did not reach statistical significance ($p=0.065$). To exclude possible confounding effect of *CCR2-64I*, which is known as a protective genetic factor



◀ **Fig. 2** **a**, **c** MT4 cells were infected with recombinant SeV expressing human-TRIM5α-43H (*open squares*), human-TRIM5α-43Y (*open circles*), human/AGM-TRIM5α-43H (*closed squares*), human/AGM-TRIM5α-43Y (*closed circles*) or a truncated form of CM-SPRY(−) TRIM5α as a negative control (*closed triangles*). Nine hours after infection, cells were inoculated with 20 ng of p24 of HIV-1 NL43, and culture supernatants were periodically assayed for levels of p24. *Error bars* showed actual fluctuations between measurements of p24 in duplicate samples. Representative data from two independent experiments are shown. **b**, **d** One million cells of each recombinant SeV-infected cells were lysed for immunoprecipitation with an anti-HA antibody. **e** TK-ts13 cells were infected with SeV expressing human-TRIM5α-43H (*open squares*), human-TRIM5α-43Y (*open circles*), human/AGM-TRIM5α-43H (*closed squares*), human/AGM-TRIM5α-43Y (*closed circles*), or CM-SPRY(−)TRIM5α (*closed triangles*) and then exposed to the indicated p24 amounts of GFP-expressing HIV-1 vector. GFP-positive cells were counted with a flowcytometry. *Error bars* indicated standard deviation of triplicate samples

against AIDS progression, we excluded patients with this allele from the analysis. There was still no statistically significant difference of the 43Y allele frequency between 9 LTNPs and 15 SPs (0.111, two out of 18 chromosomes, vs 0.166, five out of 30 chromosomes, $p=0.69$, Fisher's exact test). Statistical simulation with ten times more subject numbers again failed to show statistically significant difference ($p=0.09$)

We also analyzed HIV-1-infected individuals in two well-characterized French cohorts: 45 subjects from the Agence Nationale de Recherches sur le Sida CO16 Long Term Non-Progressors Cohort (ALT) and 62 from the cohort of SPs (IMMUNOCO). The patients in ALT cohort were with no AIDS symptoms at the time of recruit in 1996 without any kinds of treatment, and their CD4 counts were more than 600/μl during last 5 years. The patients in IMMUNOCO cohort were at any stage, with or without treatment, and their CD4 counts were more than 150/μl at entry in 1991 to 1992 (Magierowska et al. 1999). Thirteen of the ALT subjects and 14 of the IMMUNOCO subjects were heterozygous for 43Y. Again, there was no statistically significant difference in the ratio of 43Y heterozygotes between ALT and IMMUNOCO subjects (0.29 vs 0.23, $p=0.460$, chi square test). The odds ratio was 0.72 with a 95% confidence interval of 0.30–1.73, indicating that H43Y did not exert any strong effect on HIV-1 disease progression in the French subjects. Our findings for Asians and Europeans indicated that the effects of H43Y SNP of the human *TRIM5α* gene on HIV-1 disease progression are minor, if any. Our results are consistent with the previous observation that 43Y does not have a protective effect against HIV-1 replication or disease progression in European Americans (Goldschmidt et al. 2006; Javanbakht et al. 2006; Speelmon et al. 2006).

In conclusion, the results of our epidemiological and molecular biological studies clearly indicate that H43Y SNP in the human *TRIM5α* gene has a minor effect on the anti-HIV-1 activity of TRIM5α. Although we did not evaluate the effects of H43Y on the anti-MLV activity of human TRIM5α, this allele is immaterial, at least in cases with HIV-1 subtype B such as found in Europeans, European Americans, and Japanese hemophiliacs. It might be important to test the restriction capability of human TRIM5α in other subtypes of HIV-1 or HIV-2.

Acknowledgment We thank Dr. S. Osa and Dr. H. Hanabusa for the supply of patients' samples and valuable discussions. We thank S. Bando for her skillful assistance. This work was supported by grants from the Human Health Foundation, the Ministry of Education, Culture, Sports, Science, and Technology, and the Ministry of Health, Labour and Welfare, Japan.

References

Goldschmidt V, Bleiber G, May M, Martinez R, Ortiz M, Telenti A (2006) Role of common human TRIM5alpha variants in HIV-1 disease progression. Retrovirology 3:54

Hatziioannou T, Perez-Caballero D, Yang A, Cowan S, Bieniasz PD (2004) Retrovirus resistance factors Ref1 and Lv1 are species-specific variants of TRIM5alpha. Proc Natl Acad Sci USA 101:10774–10779

Javanbakht H, An P, Gold B, Petersen DC, O'Huigin C, Nelson GW, O'Brien SJ, Kirk GD, Detels R, Buchbinder S, Donfield S, Shulenin S, Song B, Perron MJ, Stremlau M, Sodroski J, Dean M, Winkler C (2006) Effects of human TRIM5alpha polymorphisms on antiretroviral function and susceptibility to human immunodeficiency virus infection. Virology 354:15–27

Kasper DL, Fauci AS, Longo DL, Braunwald E, Hauser SL, Jameson JL (2005) Harrison's Principles of Internal Medicine. McGraw-Hill, New York

Keckesova Z, Ylinen LM, Towers GJ (2004) The human and African green monkey TRIM5alpha genes encode Ref1 and Lv1 retroviral restriction factor activities. Proc Natl Acad Sci USA 101:10780–10785

Magierowska M, Theodorou I, Debre P, Sanson F, Autran B, Riviere Y, Charron D, Costagliola D (1999) Combined genotypes of CCR5, CCR2, SDF1, and HLA genes can predict the long-term nonprogressor status in human immunodeficiency virus-1-infected individuals. Blood 93:936–941

Nakayama EE, Miyoshi H, Nagai Y, Shioda T (2005) A specific region of 37 amino acid residues in the SPRY (B30.2) domain of African green monkey TRIM5alpha determines species-specific restriction of simian immunodeficiency virus SIVmac infection. J Virol 79:8870–8877

Nakayama EE, Maegawa H, Shioda T (2006) A dominant-negative effect of cynomolgus monkey tripartite motif protein TRIM5alpha on anti-simian immunodeficiency virus SIVmac activity of an African green monkey orthologue. Virology 350:158–163

Perez-Caballero D, Hatziioannou T, Yang A, Cowan S, Bieniasz PD (2005) Human tripartite motif 5alpha domains responsible for retrovirus restriction activity and specificity. J Virol 79:8969–8978

Perron MJ, Stremlau M, Song B, Ulm W, Mulligan RC, Sodroski J (2004) TRIM5alpha mediates the postentry block to N-tropic murine leukemia viruses in human cells. Proc Natl Acad Sci USA 101:11827–11832

Sawyer SL, Wu LI, Emerman M, Malik HS (2005) Positive selection of primate TRIM5alpha identifies a critical species-specific retroviral restriction domain. Proc Natl Acad Sci USA 102:2832–2837

Sawyer SL, Wu LI, Akey JM, Emerman M, Malik HS (2006) High-frequency persistence of an impaired allele of the retroviral defense gene TRIM5alpha in humans. Curr Biol 16:95–100

Speelmon EC, Livingston-Rosanoff D, Li SS, Vu Q, Bui J, Geraghty DE, Zhao LP, McElrath MJ (2006) Genetic association of the antiviral restriction factor TRIM5alpha with human immunodeficiency virus type 1 infection. J Virol 80:2463–2471

Stremlau M, Owens CM, Perron MJ, Kiessling M, Autissier P, Sodroski J (2004) The cytoplasmic body component TRIM5alpha restricts HIV-1 infection in Old World monkeys. Nature 427:848–853

Stremlau M, Perron M, Welikala S, Sodroski J (2005) Species-specific variation in the B30.2(SPRY) domain of TRIM5alpha determines the potency of human immunodeficiency virus restriction. J Virol 79:3139–3145

Yap MW, Nisole S, Lynch C, Stoye JP (2004) Trim5alpha protein restricts both HIV-1 and murine leukemia virus. Proc Natl Acad Sci USA 101:10786–10791

Yap MW, Nisole S, Stoye JP (2005) A single amino acid change in the SPRY domain of human Trim5alpha leads to HIV-1 restriction. Curr Biol 15:73–78

doi: 10.1111/j.1744-313X.2007.00694.x

Polymorphisms in CCR5 chemokine receptor gene in Japan

H. Liu,*† E. E. Nakayama,* I. Theodorou,‡ Y. Nagai,§ S. Likanonsakul,¶ C. Wasi, P. Debre,‡ A. Iwamoto†† & T. Shioda***

Summary

Mutations in the human CC chemokine receptor 5 (*CCR5*) gene may alter the expression or function of the protein product, thereby altering chemokine binding/signalling or human immunodeficiency virus type 1 (HIV-1) infection of the cells that normally express CCR5 protein. We performed a systematic survey of natural sequence variations in an 8.1-kb region of the entire *CCR5* gene as well as *CCR2V64I* in 50 Japanese subjects and evaluated the effects of those variations on *CCR5* promoter activity. We also analysed *CCR5* promoters and *CCR2V64I* in 80 more Japanese and 186 Thais. There was no 32-bp deletion observed in Caucasians, but two types of non-synonymous substitutions were found in *CCR5* genes of Japanese. Our results showed several novel characteristics of the *CCR2-CCR5* haplotype structure that were not reported from studies on Caucasians and African-Americans. Specifically, we were able to show that the G allele at position –2852 from the *CCR5* open reading frame in Japanese and Thais is the representative of the CCR5 promoter haplotype that was reported to be associated with rapid progression to acquired immune deficiency syndrome (AIDS) in HIV-1-infected individuals. Furthermore, nearly all non-synonymous polymorphisms in Japanese *CCR5* occurred in haplotypes with elevated promoter activity. We thus hypothesized that there was a certain selective pressure favouring low levels of *CCR5* expression during human evolution.

* Department of Viral Infections, Research Institute for Microbial Diseases, Osaka University, Osaka, Japan, † Miller School of Medicine, University of Miami, Miami, FL, USA, ‡ Hôpital Pitié SalPetrière et INSERM UR543 Bâtiment CERVI, Paris, France, § Center of Research Network for Infectious Diseases, Riken, Yokohama, Japan, ¶ Bamrasnaradura Institute, Nonthaburi, Thailand, ** Department of Microbiology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand and †† Division of Infectious Diseases, Institute of Medical Science, University of Tokyo, Tokyo, Japan

Authors H. Liu and E. E. Nakayama contributed equally to this work.

Received 5 May 2007; revised 5 May 2007; accepted 17 May 2007

Correspondence: Tatsuo Shioda, Department of Viral Infections, Research Institute for Microbial Diseases, Osaka University, 3-1 Yamada-oka, Suita, Osaka 565-0871, Japan. Tel: +81 6 6879 8346; Fax: +81 6 6879 8347; E-mail: shioda@biken.osaka-u.ac.jp

Introduction

Human CC chemokine receptor 5 (CCR5) mediates the activation of cells by the CC chemokines macrophage inflammatory protein-1α and -1β (MIP-1α or CCL3, and MIP-1β or CCL4), and regulated on activation normal T cells expressed and secreted (RANTES or CCL5). Identification of CCR5 as an essential co-receptor for the cellular entry of human immunodeficiency virus type 1 (HIV-1) R5 strains (Alkhatib *et al.*, 1996; Deng *et al.*, 1996; Dragic *et al.*, 1996; Feng *et al.*, 1996), which is preferentially transmitted between individuals (Zhu *et al.*, 1993), has led to many studies on CCR5 and its ligands. Mutations in the *CCR5* gene may alter the expression or function of the protein product, thereby altering chemokine binding/signalling or HIV-1 infection of the cells that normally express the CCR5 protein. Indeed, a 32-base pair (bp) deletion in the *CCR5* coding region (*CCR5Δ32*), which results in a premature termination codon, confers marked resistance to HIV-1 infection in homozygotes (Liu *et al.*, 1996; Samson *et al.*, 1996), and delays progression to AIDS and death by 2–3 years in patients heterozygous for this allele (Dean *et al.*, 1996; Huang *et al.*, 1996). The delayed progression to AIDS and death in individuals heterozygous for *CCR5Δ32* has been attributed to reduced cell surface expression of CCR5 (Wu *et al.*, 1997), which is speculated to result in a slower rate of replication and spread of the virus. Although heterozygosity for *CCR5Δ32* is associated with a small reduction in surface expression of CCR5, cells from individuals with the wild-type *CCR5* genotype showed a wide range in surface expression (Wu *et al.*, 1997), raising the possibility that polymorphisms, other than *CCR5Δ32*, exert significant effects on CCR5 expression. For example, the much more rarely occurring *CCR5 m303A* is a nonsense mutation of the *CCR5* coding region that exerts effects similar to those of *CCR5Δ32* (Quillent *et al.*, 1998). Moreover, Asian-specific *CCR5 893(–)* is a single-nucleotide deletion in the *CCR5* coding region, and the levels of CCR5 expression on the surface of CD4 positive cells are greatly reduced in individuals bearing this allele (Shioda *et al.*, 2001). The CCR2 mutation, *CCR2 64I*, which is in strong linkage disequilibrium with another mutation *CCR5 –1835T* (*CCR5 927T* in numbering system C, see Materials and Methods for definitions of numbering systems A, B, and C) in the second intron of the *CCR5*

gene, is also associated with a delay in HIV-1 disease progression (Smith *et al.*, 1997; Kostrikis *et al.*, 1998; Mummidi *et al.*, 1998).

With respect to the promoter region, Martin *et al.* (1998) identified 10 polymorphic nucleotide positions in the 517 bp region of Caucasian *CCR5* promoters, and described 10 haplotypes in this region. Of these, the *P1* haplotype has G, C, and A at the 208th, 627th, and 676th positions, respectively, from the transcription start site of the *CCR5* gene. Although both *CCR2 64I* and *CCR5Δ32* were consistently found in the *CCR5 P1* haplotype, the homozygotes for (*CCR2 64V/CCR5 P1*/lacking *CCR5Δ32*) haplotype exhibited an epidemiological association with rapid progression to AIDS (Martin *et al.*, 1998). McDermott *et al.* (1998) reported that the presence of homozygous for *CCR5 59029G* lacking *CCR5Δ32* and *CCR2 64I* caused a delay by 3.8 years in the progression to AIDS compared to that seen in patients homozygous for *59029 A*. Position 59029 in McDermott's report corresponds to position 303 from the transcription start site. Since different reports have adopted different numbering systems for the *CCR5* gene (Moriuchi *et al.*, 1997; Guignard *et al.*, 1998; Kostrikis *et al.*, 1998; Martin *et al.*, 1998; McDermott *et al.*, 1998; Mummidi *et al.*, 1998), Carrington *et al.* (1999) proposed a new numbering system starting from the translation start codon. Using this numbering system, Mummidi *et al.* (2000) reported six human haplogroups (*HHA, -B, -C, -D, -E, -F* and -G) according to *CCR2 64*, *CCR5Δ32*, and eight polymorphic positions (–2733, –2554, –2459, –2135, –2132, –2086, and –1835) in the 926 bp region of Caucasian and African-American *CCR5* promoters.

It is well known that genetic polymorphisms and haplotype structures can vary among ethnic groups. *CCR5Δ32* is extremely rare in Asian populations (Martinson *et al.*, 1997). To explore the *CCR2-CCR5* haplotype structure in an Asian population, we performed a systematic survey of natural sequence variations in an 8.1-kb region of the entire *CCR5* gene as well as of *CCR2V64I* in 50 Japanese individuals and evaluated effects of those variations on *CCR5* promoter activity. Our results show several novel characteristics of the *CCR2-CCR5* haplotype structure that were not reported from studies on Caucasians and African-Americans.

Materials and methods

Clinical samples

Blood samples were collected with anticoagulant from randomly selected, non-related 80 non-HIV-1-infected and 50 HIV-1-infected Japanese at the outpatient clinic of the Institute of Medical Science, the University of Tokyo, Tokyo, Japan, after their written informed consent had been obtained. Blood samples were also collected with anticoagulant from randomly selected, non-related 97 non-HIV-1-infected and 89 HIV-1-infected Thai subjects at the Bamrasnaradura Institute, Nonthaburi, Thailand, after their written informed consent had been obtained. Peripheral blood mononuclear cells (PBMC) were obtained from blood with the Ficoll–Histopaque method, and DNA was extracted from the PBMC with a previously described method (Shioda *et al.*, 1994). The six French samples with *CCR5Δ32* (one homozygote and five heterozygotes) used for this study were described previously (Meyer *et al.*, 1997; Magierowska *et al.*, 1999).

Genotyping of CCR5 and CCR2

The 8.1-kb *CCR5* target region was amplified from each of the genomic DNA samples in seven (R5-1 to R5-7) overlapping segments using the amplification primers shown in Table 1. Polymerase chain reaction (PCR) was performed in a 50-μL reaction mixture containing 1 μg of DNA. Thermal cycling was performed with an initial 94 °C for 3 min followed by 40 cycles at 94 °C for 30 s, primer annealing for 30 s and primer extension at 72 °C and a final extension at 72 °C for 7 min. The PCR conditions of annealing temperature and extension time were shown in Table 2. The positions specified in *CCR5* were based on the numbering system proposed by Carrington *et al.* (1999). We defined this system as numbering system A and used this system in the present study. The positions in *CCR2* were based on the sequence with GenBank accession number U95626. McDermott *et al.* (1998) used this system, and we defined this system as numbering system B. The numbering system starting from the transcription start site (Martin *et al.*, 1998) was defined as numbering system C in the present study. Table 3 shows polymorphic positions in those three different numbering systems. Sequencing reactions were performed according to the dideoxy-chain-termination method using the ABI PRISM 377 (Applied Biosystems, Foster City, CA, USA) automated DNA sequencer. The sequences of primers used for sequencing reaction are shown in Table 1. The ABI sequence software (version 2.1.2) was used for lane tracking and first pass base-calling (PerkinElmer, Wellesley, MA, USA). Variant sites identified by both software and visual inspection were scored and entered into a database for subsequent analysis. Each variant site was confirmed by reamplifying and resequencing the variant site from the opposite strand. The linkage disequilibrium coefficient was calculated by using a software (ARLEQUIN version 2.01, Gennetica and Biometry Laboratory, Geneva, Switzerland).

Haplotype determination

Haplotypes were inferred from the samples by using a heuristic algorithm based on population genetic principles (Clark, 1990). Direct molecular haplotyping was used to confirm linkage relationships in compound heterozygotes using allele-specific PCR combined with direct sequencing and limited cloning of the specific products. All inferred haplotypes were confirmed with these molecular haplotyping techniques.

Allele-specific PCR

For this procedure, the amplification refractory mutation system (ARMS) was used (Newton *et al.*, 1989). An

Table 1. Primer sequences

Sense primer	5′ to 3′ sequence	Position[a]	Position[b]	Fragment	Antisense primer	5′ to 3′ sequence	Position[a]		Position[b]	Fragment
1. Primers used for PCR amplification and sequencing										
CCR2F	ATGCTGTCCACATCTCGTTC		46106 to 46125	R2-1	CCR2R	GGAAATTATTCCATCCTCGTG			46709 to 46689	R2-1
P61	AGCAAGGAGACAGCAAAGCA	–4796 to –4777	56696 to 56715	R5-1	P63	CCTTGTATTCACTGGTAATCC	–3924 to –3944		57568 to 57548	R5-1
P66	AGAACCAGCAATTGCCACAAACAGA	–4012 to –3988	57480 to 57504	R5-2	P68	GCTTCCGTGACCTTGGCTCTAG	–2941 to –2962		58547 to 58526	R5-2
LK81	AATTTTGCTGTTTGGGGTCTC	–3022 to –3002	58466 to 58486	R5-3	LK83	CTGATTATCTTAAGAGTTGC	–2369 to –2388		59119 to 59100	R5-3
LK84	AAGTCCAGGATCCCCCTCTA	–2443 to –2424	59045 to 59064	R5-4	LK87	CATTCCAAACTGTGACCCTTTCC	–1756 to –1778		59732 to 59710	R5-4
CR53F	TCCAGTGAGAAAAGCCCGTAAATA	–2162 to –2139	59326 to 59349	R5-5	CR53R	TGCCCACAAAACCAAAGATGAACA	142 to 119		61624 to 61601	R5-5
CKR5a +	CAGTTTGCATTCATGGAGGG	–84 to –65	61399 to 61418	R5-6	CKR5a–	CTAAGCCATGTGCACAACTC	1110 to 1091		62592 to 62573	R5-6
CR52F	GGAAATATCTGTGGGCTTGT	1038 to 1057	62520 to 62539	R5-7, IIIa, IIIb	CR52R	CTGCTTATAAAATGCTCTGG	3383 to 3364		64865 to 64846	R5-7
2. Primers used for sequencing										
P62F	TGAACCCCCCTGTCCTGAG	–4425 to –4407	57067 to 57085	R5-1, Ia, Ib	P62R	GCTAGATGGGAAACAAGGTGAGGA	–4295 to –4318		57197 to 57174	R5-1
P67F	GATTCGCTGTCCTTGTTATG	–3533 to –3514	57955 to 57974	R5-2, Ia, Ib	P67R	CCCAGCGATCAAGACACCC	–3436 to –3454		58052 to 58034	R5-2, Ia, Ib
LK82F	AGAACCTGAACTTGACCAT	–2811 to –2793	58677 to 58695	R5-3, Ia, Ib	LK82R	TTTTAACTATGGGCTCACG	–2573 to –2591		58915 to 58897	R5-3, Ia, Ib
LK85	GTGTAGTGGGATGAGCAGAGA	–2198 to –2178	59290 to 59310	R5-4, II a, II b	LK85R	TTTGAAGGAGGGTGGAGTT	–2019 to –2037		59469 to 59451	R5-4
433	ACACCAAGTGCTCATACAAT	–1825 to –1806	59663 to 59682	R5-5	1552R	TAGAGTTAGCCCACAAAGAA	–685 to –704		60801 to 60782	R5-5, IIa, IIb
CR53S1	TTTTCTAGGGGCTTCTCTCA	–1367 to –1348	60121 to 60140	R5-5, IIa, IIb	2046 R	CTGCTAGCTTCCCTGTCCACT	–187 to –207		61296 to 61276	R5-5, IIIa, IIIb
CR53S2	GCAGGCTTCCCGCATTCAAA	–938 to –919	60548 to 60567	R5-5, IIa, IIb	CKR5e–	GCTGCGATTTGCTTCAC	89 to 73		61571 to 61555	R5-5
1770	ACTTGGAGGGGTGAGGTGAG	–486 to –467	61000 to 61019	R5-5, IIa, IIb	CKR5d–	GAAGATTCCAGAGAAGAAGCC	351 to 331		61833 to 61813	R5-6
CKR5b +	GAGCATGACTGACATCTACC	186 to 205	61668 to 61687	R5-6	322	GTGAAGATAAGCCTCACAGCC	716 to 696		62198 to 62178	R5-6
CKR5c +	CTGTGTTTGCGTCTCTCC	467 to 484	61949 to 61966	R5-6, IIIa, IIIb	CR5R1	TGTCTTTTCTCCCCATAGCA	1527 to 1508		63009 to 62990	R5-7
CKR5d +	CAGTAGCTCTAACAGGTTGG	807 to 824	62289 to 62308	R5-6	CR5R4	TCACCACTATAGGGACCCCTT	2578 to 2559		64060 to 64041	R5-7
CR5S1	GCTGATTCTTGAGTTTAGTG	1387 to 1406	62869 to 62888	R5-7						
CR5S2	CATGGGGAGGAAGGACAAGG	1768 to 1787	63250 to 63269	R5-7, IIIa, IIIb, IVa						
CR5S3	GGAGGAGGAGGTTTAGGTCA	2168 to 2188	63650 to 63670	R5-7						
CR5S4	AAGGGGTCCCATAGAGGTGA	2559 to 2578	64041 to 64060	R5-7						
CR5S5	TGAATTTGGGGGATGGCTAA	2954 to 2973	64436 to 64455	R5-7						
3. Primers used for allele-specific PCR amplification										
627UA	TCCTATGGGGTGTCCGAATGT	–4638 to –4618	56854 to 56874	Ia, Ib	627N	GAATAGATCTCTGGTCTGtAA	–2115	–2135	59373 to 59353	Ia
303 N	GAGAGTGGAGAAAAAGGtGG	–2478 to –2459	59010 to 59029	IIa	627M	GAATAGATCTCTGGTCTGtAG	–2115	–2135	59373 to 59353	Ib
303 M	GAGAGTGGAGAAAAAGGtGA	–2478 to –2459	59010 to 59029	IIb	303DA	TAAGAAACTGGGTCAAGCAT	–316	–335	61167 to 61148	IIa, IIb
2398 N	CTGTCTCACAACAACAtCAG	–381 to –362	61105 to 61124	IIIa	2398DA	AACCAGACCCATCCTTTTAC	2226	2207	63708 to 63689	IIIa, IIIb
2398 M	CTGTCTCACAACAACAtCAA	–381 to –362	61105 to 61124	IIIb	5765M	ATGTGCACAATCATATGtGAC	2939	2919	64421 to 64401	IVa
5765UA	GACTTAGAACCAGGCGAGAG	1602 to 1621	63084 to 63103	IVa						

Underline; allele-specific nucleotide.

Lower case t; additional internal (position –3) mismatch that increases specificity of allele-specific PCR.

[a]Position according to Mummidi *et al.* (2000) (numbering system A).

[b]Position in the baseline sequence (GenBank accession number U95626, numbering system B).

Table 2. Amplification protocol

Fragment	Sense primer	Antisense primer	Annealing temp (C)	Extension time (s)	Length (bp)	Position[a]	Position[b]
R2-1	CCR2F	CCR2R	60	30	604		46106 to 46709
R5-1	P61	P63	60	30	873	−4796 to −3924	56696 to 57568
R5-2	P66	P68	60	30	1072	−4012 to −2941	57480 to 58547
R5-3	LK81	LK83	60	30	654	−3022 to −2369	58466 to 59119
R5-4	LK84	LK87	60	30	688	−2443 to −1756	59045 to 59732
R5-5	CR53F	CR53R	58	120	2303	−2162 to 142	59326 to 61624
R5-6	CKR5a +	CKR5a −	58	60	1194	−84 to 1110	61399 to 62592
R5-7	CR52F	CR52R	56	120	2346	1038 to 3383	62520 to 64865
Ia	627UA	627N	55	120	2520	−4637 to −2115	56854 to 59373
Ib	627UA	627M	55	120	2520	−4637 to −2115	56854 to 59373
IIa	303N	303DA	61	120	2158	−2478 to −318	59010 to 61167
IIb	303M	303DA	61	120	2158	−2478 to −318	59010 to 61167
IIIa	2398N	2398DA	63	120	2607	−382 to 2226	61105 to 63708
IIIb	2398M	2398DA	63	120	2607	−382 to 2226	61105 to 63708
IVa	5765UA	5765M	60	70	1338	1602 to 2939	63084 to 64421

[a]Position according to Mummidi *et al.* (2000) (numbering system A).
[b]Position in the baseline sequence (GenBank accession number U95626, numbering system B).

ARMS primer contained allele-specific nucleotide at its 3′ end. Therefore, we could specifically amplify DNA sequence on the chromosome carrying the allele. Allele-specific products ranged in size from 1.3 to 2.6 kb. Primers used for allele-specific PCR are shown in Table 1. PCR and sequencing were performed in the same manner as described above (Table 1 and Table 2).

PCR-restriction fragment length polymorphism analysis

The region spanning CCR5 −2852 was amplified by PCR with primer pair P67F and LK83 (Table 1). PCR was performed for 40 cycles at 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 1 min. Digestion of the 1164-bp amplified products with Aci I yielded 680- and 464-bp fragments when position −2852 was G.

Promoter assay

Luciferase reporter gene assays were performed as described previously (Liu *et al.*, 1999). Briefly, the promoter region of each haplotype was PCR-amplified and cloned into a pGL3-Basic vector carrying the firefly luciferase gene (Promega, Madison, WI, USA). All constructs were verified for sequence authenticity by sequencing the entire insert, and 5 μg of the resultant constructs was transfected with DMRIE-C (Invitrogen, Carlsbad, CA, USA) into monocytic U937 cells. Transfection efficiency was normalized by cotransfecting 0.2 μg of pRL-TK, which expresses *Renilla* luciferase under the control of the herpes simplex virus thymidine kinase promoter. Cells were harvested 40 h after transfection, and firefly and *Renilla* luciferase activities were determined with the Dual-Luciferase Reporter Assay System (Promega) according to manufacturer's instructions. Relative luciferase expression was derived with the following equation: (firefly luciferase activity of CCR5 promoter construct/*Renilla* luciferase activity)/(firefly luciferase activity of promoterless vector pGL3-Basic/*Renilla* luciferase activity). Data points are means derived from measurements of two independent clones of each of the constructs.

Results

Sequence polymorphisms in human *CCR5* gene

The human *CCR5* gene, located on chromosome 3, has four exons distributed across 6 kb of genomic DNA (Mummidi *et al.*, 1997). We amplified seven overlapping DNA fragments covering an 8.1-kb genomic region spanning 2031 bp of the upstream non-coding region, the exons and introns in their entirety, and 81 bp of the 3′ untranscribed region of *CCR5* gene. The fragments were obtained from 50 randomly selected Japanese subjects comprising 38 non-HIV-1-infected and 12 HIV-1-infected individuals. Direct sequencing of the PCR fragments allowed us to identify 25 polymorphic positions in this region (Table 3). Among the 25 variable positions, four were insertions or deletions, and 21 were single-nucleotide substitutions. There was one additional C-to-T substitution in the 4-base insertion at position −3887 (Table 3). Except for this tri-allelic site, all the other variations in these samples were di-allelic. No *CCR5Δ32* or *CCR5 m303* was found in these Japanese samples, which confirms previously reported observations (Martinson *et al.*, 1997; Quillent *et al.*, 1998).The genotype frequencies in 50 individuals analysed in this study were consistent with the Hardy–Weinberg equilibrium, suggesting the absence of any tendency towards an excess or deficiency of any particular genotype. Overall, nucleotide diversity was 0.00108, which is the equivalent of approximately one variant in every 926 bp. The sequence diversity in coding-region was lower (0.00012) than in non-coding region (0.00122). In addition to the *CCR5* gene, we analysed the *CCR2* gene of the same 50 individuals for the *CCR2 64I* allele, which has been reported to be associated with delayed HIV-1

Table 3. Sequence variants identified in *CCR2* and *CCR5* genes and used to define haplotypes

						Allele frequency			
Site[a]	Position[b]	Position[c]	Position[d]	Variant	NCBI SNP ID	Japanese[j]	Thais[k]	Caucasian	African-American
1		46295		G, A (CCR2 V64I)	rs1799864	0.67, 0.33	0.83, 0.17	0.90, 0.10[l]	0.84, 0.16[l]
2	–4358	57134		G, A	rs7637813	0.20, 0.80			
3	–3900	57592		A, C	rs2856757	0.57, 0.43			
4[e]	–3888 ~ –3884	57604		GCTAT, G, GTTAT	rs10577983	0.56, 0.43, 0.01			
5	–3458	58030		G, T	rs2734225	0.57, 0.43			
6	–2852	58636		G, A	rs2227010	0.21, 0.79	0.18, 0.82		
7	–2554	58934	208	G, T	rs2734648	0.57, 0.43			
8	–2459	59029	303	A, G	rs1799987	0.56, 0.44		0.57, 0.43[m]	0.43, 0.57[m]
9	–2135	59353	627	C, T	rs1799988	0.55, 0.45	0.35, 0.65		
10	–2086	59402	676	A, G	rs1800023	0.60, 0.40	0.39, 0.61		
11	–1835	59653	927	C, T	rs3181036	0.66, 0.34	0.83, 0.17	0.90, 0.10[n]	0.80, 0.20[n]
12[f]	–1132 ~ –1130	60356		CAG, C	rs3054375	0.57, 0.43			
13	–1060	60426		C, T	rs2856762	0.99, 0.01			
14	–976	60510		C, T	rs2254089	0.57, 0.43			
15	–651	60835		C, T	rs2856764	0.57, 0.43			
16	–451	61035		C, T		0.97, 0.03			
17	–444	61042		G, A	rs2856765	0.57, 0.43			
18	–362	61124		A, G		0.57, 0.43			
19[g]	–361 ~ –359	61125		CAAC, C		0.57, 0.43			
20[h]	668	62150		G, A	rs1800452	0.95, 0.05			
21[h,i]	893	62375		C, –		0.99, 0.01			
22	1171	62653		C, G		0.98, 0.02			
23	1823	63305		C, T	rs17765882	0.99, 0.01			
24	2077	63559		G, T	rs1800874	0.57, 0.43			
25	2150	63632		G, C		0.98, 0.02			
26	2919	64401		G, T	rs746492	0.56, 0.44			

[a]Site number assigned to a variable character used to define haplotypes in order 5′ to 3′.
[b]Position according to Mummidi *et al.* (2000) (numbering system A).
[c]Position in the baseline sequence (GenBank accession number U95626, numbering system B).
[d]Position according to Martin *et al.* (1998) (numbering system C).
[e]There was one additional C-to-T substitution within the duplication of CTAT at position –3887.
Except for this tri-allelic site, all the other variations were di-allelic in these samples.
[f]Insertion of AG.
[g]Duplication of CAA.
[h]Variant in coding region.
[i]Deletion of C.
[j]Allele frequencies of site1,6,9,10 and 11 in 130 Japanese. Allele frequencies of other sites in 50 Japanese.
[k]Allele frequencies of site1,6,9,10 and 11 in 186 Thais.
[l]Allele frequencies reported by Martin *et al.* (1998).
[m]Allele frequencies reported by McDermott *et al.* (1998).
[n]Allele frequencies reported by Mummidi *et al.* (1998).

disease progression (Kostrikis *et al.*, 1998; Smith *et al.*, 1997). The *CCR2* gene was found to be located 14-kb upstream of the *CCR5* gene. The *CCR2* and *CCR5* genotypes of the 50 individuals analysed were sorted and renumbered according to their genotypes and are shown in Fig. 1. Using the method described by Clark *et al.* (1990) we were able to infer all the haplotypes that were subsequently confirmed by allele-specific PCR as well as limited cloning and sequencing of the PCR-amplified fragments. As shown in Fig. 2, we were able to identify 12 independent haplotypes that showed a high degree of linkage disequilibrium of multiple variable sites. *CCR2 64I* was found in *H6*, *H7*, and *H8*, and *CCR5* –1835T in *H6*, *H7*, *H8*, *H9*, and *H10*. As reported previously, *CCR2 64I* is in strong linkage disequilibrium with *CCR5 –1835T* (D′ = 1). *CCR5 P1* (*–2554G*, *–2135C* and *–2086A*, which correspond to 208G, 627C, and 676 A, respectively, in numbering system C) was found in *H6*, *H7*, *H8*, *H9*, *H10* and *H11*. The *–2459A*, which corresponds to 59029 A in numbering system B, was in complete linkage disequilibrium with *CCR5 P1* and found in *H6*, *H7*, *H8*, *H9*, *H10* and *H11*. All the samples showed C in nucleotide position –2132, where T is located in approximately 11.3% of African-American chromosomes (Bamshad *et al.*, 2002).

Twelve haplotypes were subsequently categorized into four major haplotype groups based on the polymorphisms at positions –2135, –2086, and –1835 (Fig. 2). Haplotype group I, comprising *H1*, *H2*, *H3*, *H4*, and *H5*, seems to

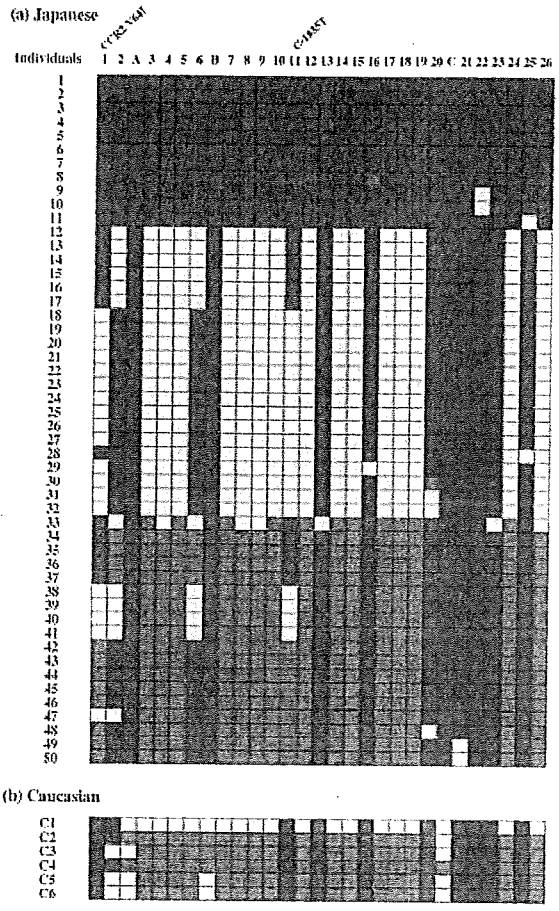


Figure 1. *CCR2* and *CCR5* genotypes of the 50 Japanese (a) and six French (b) subjects at each of the polymorphic sites. Individual samples were sorted and re-numbered according to their genotype, and sample identifiers are shown on the left side of the array. The 26 polymorphic sites observed in Japanese are numbered consecutively across the top. A, B, and C denote polymorphisms specifically observed in the French samples. Positions of *CCR2 V64I* and *CCR5 –1835T* are indicated. Genotypes for each individual were assigned directly based on the fluorescence sequencing trace at each position. At every site, individuals homozygous for the baseline allele (U95626) are shown in red, heterozygotes in yellow and those homozygous for the variant in blue. The heterozygote for the four-base insertion and the additional substitution at position –3887 is shown in white.

correspond to *HHC* according to Mummidi *et al.* (2000). *CCR5 –1835T* was found only in haplotype group II (*H6*, *H7*, *H8*, *H9*, and *H10*), which corresponds to *HHF*. *CCR2 64I* and the Asian-specific disrupting mutation in the *CCR5* open reading frame, *CCR5 893 (–)* (Shioda *et al.*, 2001), was found in haplotype group II. Haplotype group III (*H11*) lacked the *CCR2 64I* but showed sequences identical to the consensus sequence of haplotype group II except for nucleotides at positions *–2852G* and *–1835C*. Haplotype groups III and IV (*H12*) correspond to *HHE* and *HHA*, respectively. None of the single mutations was associated with *CCR2 64I*, except for *CCR5 –1835T*. No *HHB*, *HHD*, *HHG*1* or *HHG*2* was observed in the 100 Japanese chromosome 3. The frequencies of *HHA*, *HHC*, *HHE*, and *HHF* in these chromosomes were similar to those reported for 27 Asians (Bamshad *et al.*, 2002). Mummidi *et al.* (2000) previously pointed out that disease-accelerating genotypes (*CCR5 P1* or *59029 A* lacking *CCR2 64I* and *CCR5Δ32*) are a mixture of haplotypes *HHE*, *HHF*1*, and *HHG*1*. In our study, those disease-accelerating alleles were found in *H9* (*HHF*1*), *H10* (*HHF*1*), and *H11* (*HHE*). On the other hand, the –2852 G allele was found in *H11* (*HHE*), but not in *H9* (*HHF*1*) or *H10* (*HHF*1*). Therefore, the *–2852G* was found to be associated with 90% (19/21) of the disease-accelerating alleles in those Japanese subjects (Fig. 2).

To determine the sequence polymorphisms associated with Caucasian-specific *CCR5Δ32*, we also analysed six French individuals carrying *CCR5Δ32*. The results showed that the nucleotide sequence of the haplotype with *CCR5Δ32* (*H13* in Fig. 2) was very similar to that of the haplotypes containing *CCR2 64I*, *CCR5 –1835T*, or *CCR5 P1*. Phylogenetic analysis of all the haplotypes defined in Fig. 2 showed that there were two major groups (H1-H5 and H6-H13) (Fig. 3), one of which contained all the haplotypes with *CCR2 64I*, *CCR5 –1835T*, or *CCR5 P1*. These data indicate that all the HIV-1 disease-modifying *CCR5* haplotypes were in fact very similar to each other regardless of the direction of their effects on HIV-1 diseases.

Effects on promoter activity by sequence polymorphisms in the regulatory sequences of the CCR5 gene

To compare *CCR5* promoter activity among the three major haplotype groups, I, II, and III, we constructed a series of firefly luciferase reporter fusions containing various lengths of the 5′-non-coding region of *CCR5*, and analysed their promoter activity in monocytic U937 cells. Two distinct promoters for the *CCR5* gene, upstream (Pu) and downstream (Pd), have been identified (Mummidi *et al.*, 1997). Constructs labelled with Pu (Pu-I, Pu-IIa, Pu-IIb, and Pu-III) contained both Pu and Pd, while those labelled with Pd (Pd-I and Pd-III) contained only Pd (Fig. 4). The IIa promoter corresponded to *H6*, *H7*, and *H8*, while its IIb counterpart corresponded to *H9*.

Our results showed that promoters of haplotype groups II and III expressed higher luciferase activity than those of haplotype group I in both constructs (Fig. 4). These results are especially relevant in view of the fact that studies of HIV-1-infected individuals homozygous for a haplotype analogous to haplotype group III (*the CCR5 P1 lacking CCR2 64I and CCR5Δ32*) found that they progressed to AIDS more rapidly than those with other *CCR5* promoter genotypes (Martin *et al.*, 1998; McDermott *et al.*, 1998). One of these studies also reported that a *CCR5* promoter bearing *–2459 A* (*59029 A* in numbering system C), specific for haplotype groups II and III, expressed higher promoter activity than one bearing *–2459G* (McDermott *et al.*, 1998). In addition, Mummidi *et al.* (2000) demonstrated

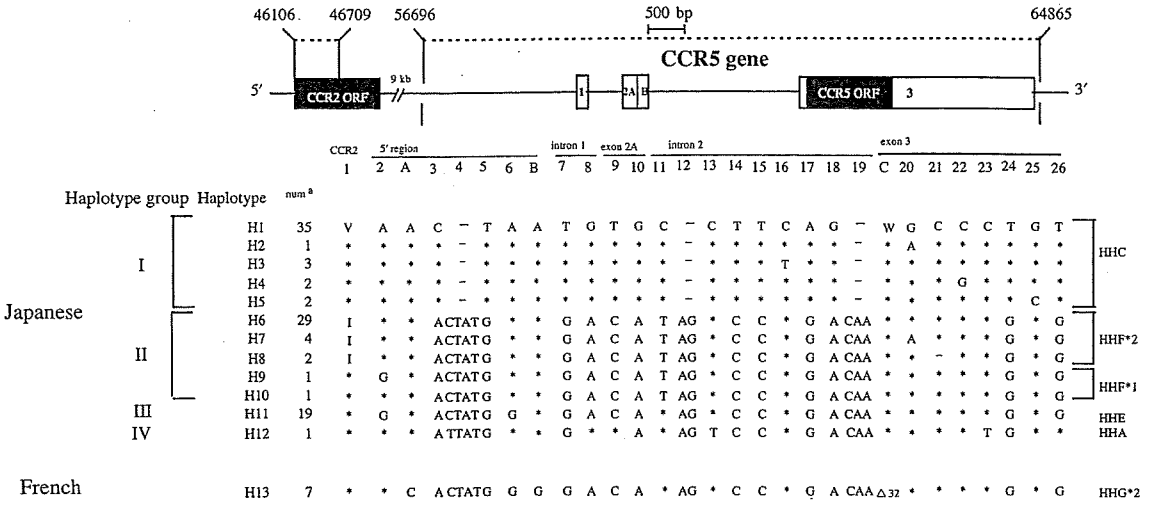


Figure 2. Map of *CCR2, CCR5* and nucleotide polymorphisms. White and black boxes indicate non-coding exons and open reading frames (ORF), respectively. Dotted lines signify the sequenced regions. PCR and sequence screening of eight segments of the *CCR2* and *CCR5* gene identified 29 nucleotide positions with genetic polymorphisms. Twenty-six polymorphic sites observed in Japanese are numbered consecutively across the top. A, B, and C denote polymorphisms specifically observed in the French samples. Combining the 29 polymorphic positions indicated 13 *CCR2-CCR5* haplotype alleles designated H1 through H13, with H1 being exactly the same as the GenBank U95626 sequence. Asterisks indicate nucleotides identical to those of H1; dashes represent deletion sites. W shows *CCR5* coding sequences without 32 bp deletion, while Δ32 indicates 32 bp deletion of the *CCR5* coding region. V and I in position 1 indicate *CCR2 64V* and *CCR2 64I*, respectively. [a]Number in this column represents the actual numbers of haplotypes identified in the 100 sequenced chromosomes. The human haplogroups according to Mummidi *et al.* (2000) are shown on the right.

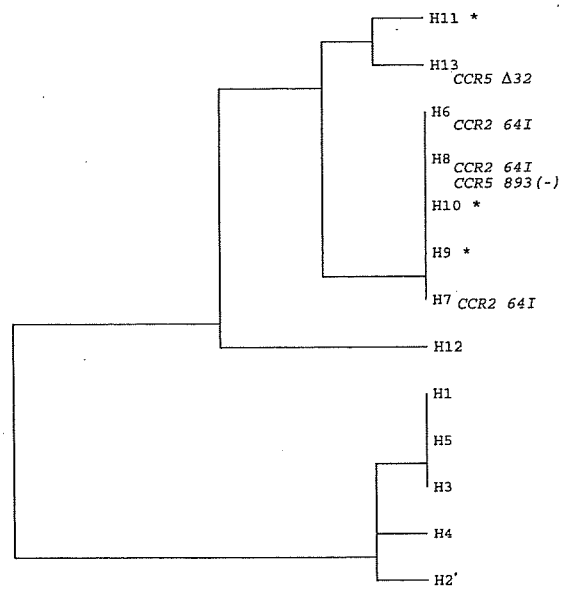


Figure 3. A phylogenetic tree of 13 *CCR2-CCR5* haplotypes defined in Fig. 2. The tree was produced with the UPGMA (unweighted pair-group method with arithmetic averages) method. Haplotypes with *CCR5 Δ32, CCR2 64I, CCR5 893(-)* are shown. Asterisks denote haplotypes with *CCR5 P1 lacking CCR5 Δ32 and CCR2 64I.*

that the *HHC* (haplotype group I) promoter construct demonstrated lower promoter activity than that of *HHF* (haplotype group II) and of *HHE* (haplotype group III). Our study further identified *–2135C* and *–2086A*, which are linked to *–2459A*, as mutations responsible for elevated promoter activity, since Pd-III was shown to express higher luciferase activity than Pd-I (Fig. 4).

To examine the effects of *CCR5 –1835T* on expression level, we generated two series of constructs with part of the intron containing position –1835 and the splicing acceptor site located immediately upstream of *CCR5* exon 4 (LSA and PdSA in Fig. 4). Again, promoters from haplotype groups II and III expressed higher luciferase activity than the promoter from haplotype group I in both the LSA and the PdSA constructs, and since there were no other differences between haplotype groups II and III, this suggests that *CCR5 –1835T* has no effect on CCR5 expression.

Novel method using PCR-restriction fragment length polymorphism for detecting HIV-1 disease-accelerating haplotype

As described above, in Japanese subjects, *–2852G* was associated with 90% of the disease-accelerating haplotype, this is, *CCR5 P1 lacking CCR5Δ32* and *CCR2 64I*, since no *CCR5Δ32* was identified in these individuals. The remaining 10% were accounted for by individuals with *CCR5 –1835T* without *CCR2 64I*. To confirm the strong linkage disequilibrium between *–2852G* and the disease-accelerating allele observed in 50 Japanese, we established a PCR-RFLP method to detect *–2852G* (see Materials and Methods) and used it to genotype 80 additional (38 HIV-1-infected and 42 non-HIV-1-infected) Japanese. *CCR5 C-2135T, G-2086A*, and *C-1835T* were genotyped by direct-sequencing of PCR product R5-4

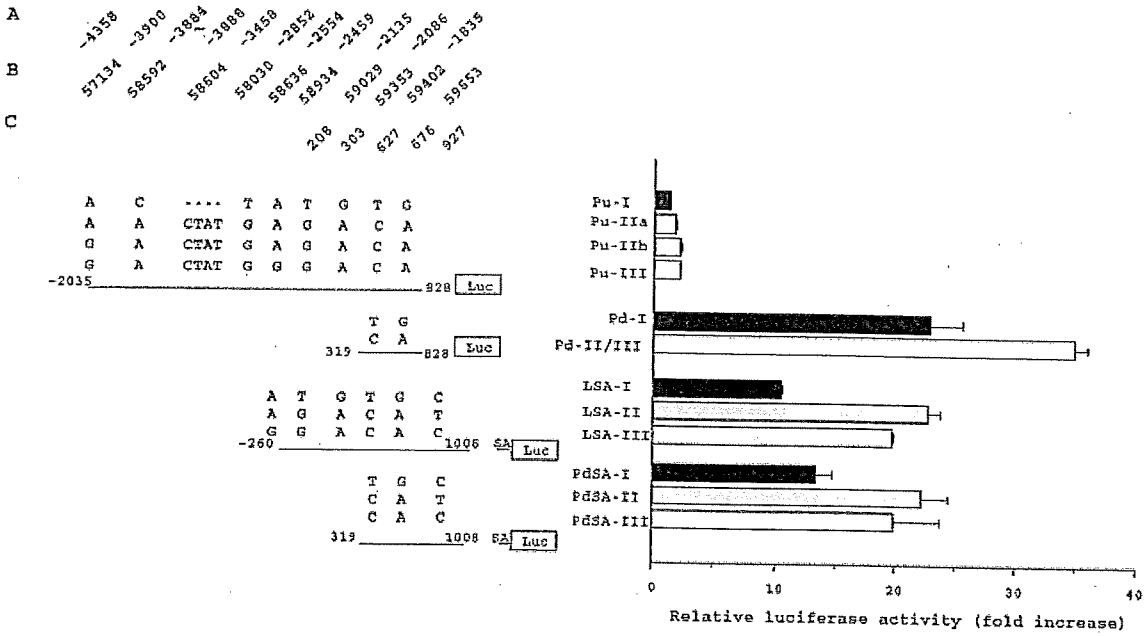


Figure 4. Effect of sequence polymorphism in CCR5 regulatory region on promoter activity. The promoter regions inserted into the pGL3-basic vector are shown by solid lines with the first and last nucleotides enumerated according to numbering system C. SA indicates 0.25 kb of the 5'-non-coding region containing a splice acceptor immediately upstream of the CCR5 open reading frame. Nucleotides at the polymorphic positions are marked. Boxes represent firefly luciferase open reading frames. The relative luciferase activity of each construct is represented by a solid bar (haplotype group I), a grey bar (haplotype group II), and an open bar (haplotype group III). Data shown are representative of five independent experiments with similar results. Error bars indicate fluctuations between measurements of relative luciferase activity in two independent clones of a construct.

Table 4. Linkage disequilibrium between *CCR5 –1835T* and *CCR2 64I* in Japanese and Thai subjects

		CCR2 V64I		
Subjects	*CCR5 C-1835T*	VV	VI	II
Japanese	CC	58	0	0
	CT	2	54	0
	TT	0	1	15
Thais	CC	123	0	0
	CT	0	61	0
	TT	0	0	2

(Table 2), and *CCR2 V64I* was genotyped by PCR-RLFP according to the method described by Smith *et al.* (1997). Frequencies of *CCR2 64I* and *CCR5 –1835T* of 130 Japanese subjects were 33% and 34%, respectively, and were higher than those observed in Caucasians and African-Americans (Table 3). As shown in Table 4 and Table 5, we confirmed the strong linkage disequilibrium between *CCR2 64I* and *CCR5 –1835T* (D′ = 1). We also found the strong linkage disequilibrium between *CCR5 –2852G* and the disease-accelerating haplotype (D′ = 1) since *–2852G* was found in 55 out of 58 (95%) of the disease-accelerating haplotype in 130 Japanese subjects.

In Thai subjects, we previously observed that *CCR2 64I* alleles were 100% associated with *CCR5 –1835T* (unpublished results). We therefore examined whether *–2852G* was associated with the disease-accelerating haplotype in 186 Thai subjects (97 non-HIV-1-infected and 89 HIV-1-infected individuals). Frequencies of *CCR2 64I*, *CCR5 –2852G*, *–2135T*, *–2086 A*, and *–1835T* of 186 Thai subjects were shown in Table 3. As expected, *CCR2 64I* alleles were associated with *CCR5 –1835T* without any exceptions (Table 4). Furthermore, *CCR5 –2852G* was also 100% associated with the disease-accelerating haplotype, *CCR5 P1* lacking *CCR2 64I* and *CCR5Δ32* (Table 5). These results indicated that only one allele (*–2852G*) could be used as a representative of the disease-accelerating haplotype, *CCR5 P1* lacking *CCR2 64I* and *CCR5Δ32* in the Thai population.

Discussion

Polymorphisms in human *CCR5* genes were initially studied for their effects on susceptibility to HIV-1 infection and rate of disease progression to AIDS in HIV-1-infected individuals. Subsequently, these polymorphisms were evaluated for their roles in other human phenotypes such as those accounting for differences in renal transplantation outcomes (Fischereder *et al.*, 2001), myocardial infarction (Gonzalez *et al.*, 2001; Valdes *et al.*, 2002), and autoimmune diseases (Garred *et al.*, 1998; Gomez-Reino *et al.*, 1999; Spagnolo *et al.*, 2005). Our systematic survey of natural sequence variations in an 8.1-kb region of the

