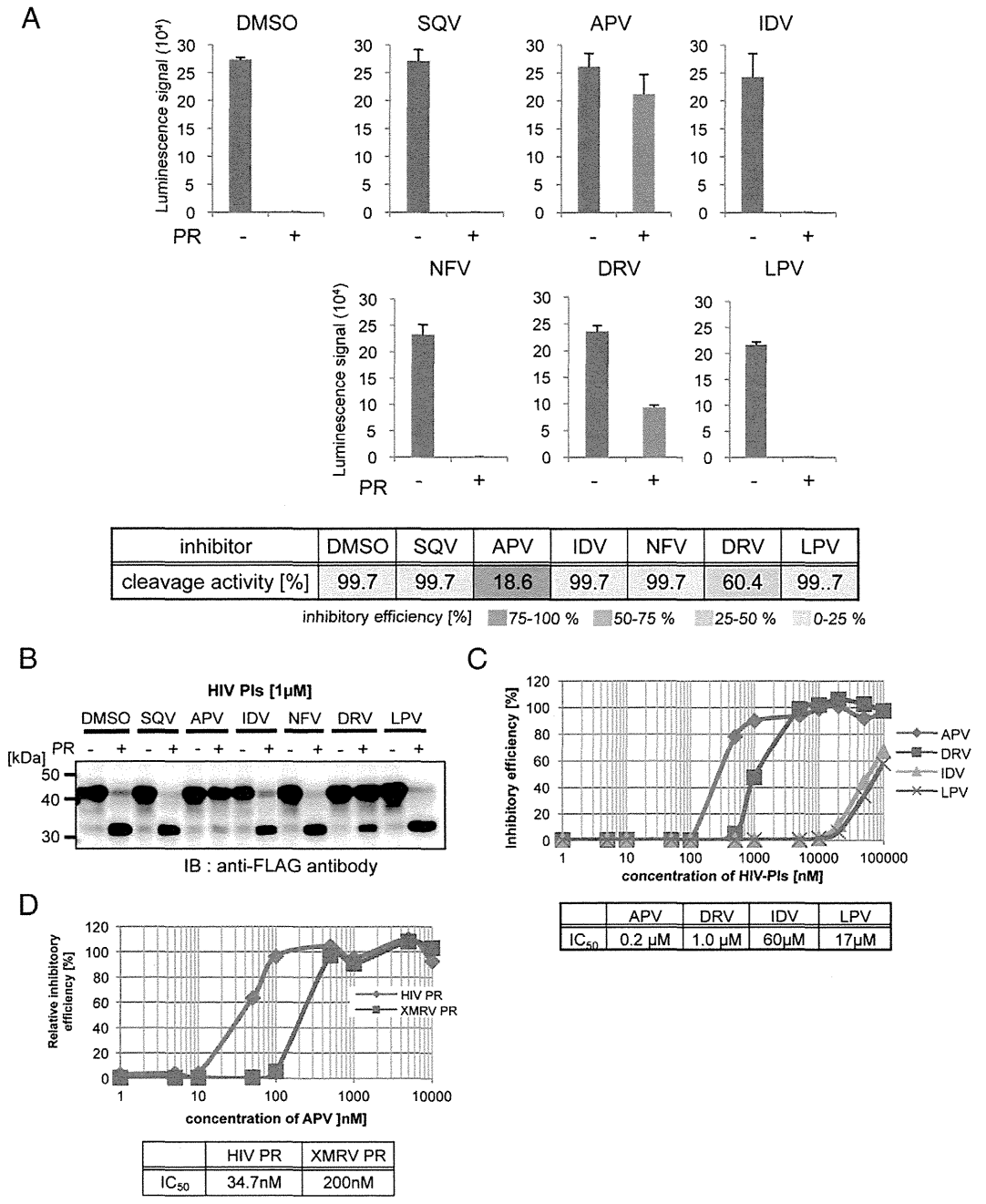
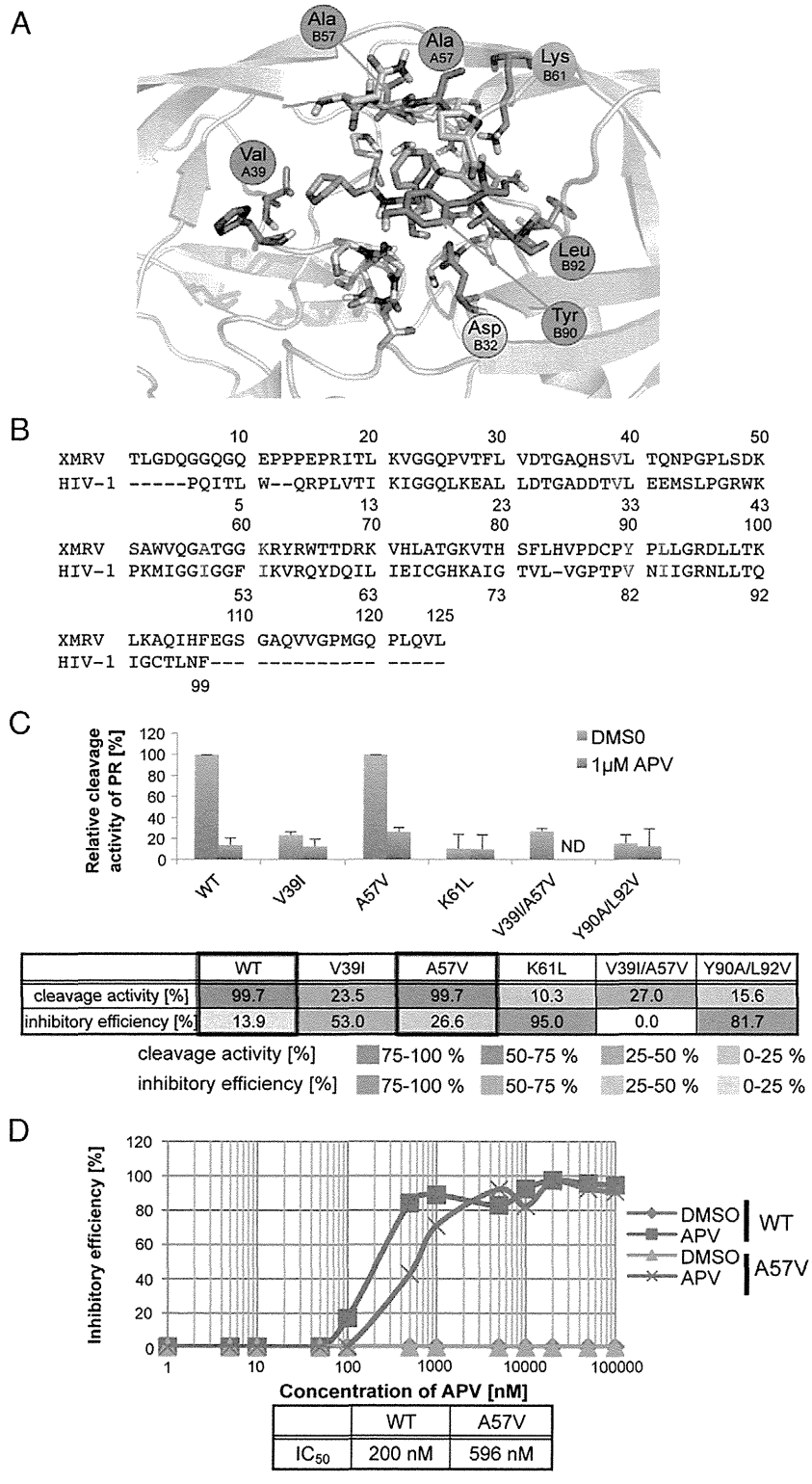


Fig. 3 – Drug screening for XMRV PR based on the cleavage activity. A,B. XMRV protease (+) or DHFR (–) was pre-incubated with indicated HIV PIs (SQV, saquinavir; APV, amprenavir; IDV, indinavir; NFV, nelfinavir ; DRV, darunavir; LPV, lopinavir; 1 μM each) and then subjected to AlphaScreen. Luminiscent AlphaScreen signal (upper panel) and relative enzymatic activity (lower panel) were listed. C. Conformation of the cleavage of the tester polypeptide by immunoblot analysis with anti-FLAG antibody. D. Dose–response curve of XMRV PR with HIV PIs using AlphaScreen (upper panel). IC_{50} values were calculated for each inhibitor (lower panel). E. Dose–response curve of XMRV PR and HIV-1 PR with APV using AlphaScreen (upper panel). IC_{50} values were calculated for each protease (lower panel).

formed between XRMV PR and several protease inhibitors, including APV [24,35]. In the current study we moved a step closer to clarifying the molecular interactions between XMRV PR and APV during drug-resistance, by developing an effective cell-free in vitro protease assay for XMRV PR. This assay revealed that an Ala57Val substitution induced significant drug-resistance to APV regardless of the integrity of the protease activity. The data indicates that this cell-free assay is useful for analyzing the drug-resistance properties of retroviral proteases.

Proteases often modify the activities of their target substrates [36]. Identification of the specific substrates cleaved by viral PR is of great significance for understanding the molecular etiology of virus infection. Proteomic studies with mass spectrometry could, theoretically, exhaustively identify the cellular proteins cleaved



	WT	V39I	A57V	K61L	V39I/A57V	Y90A/L92V
cleavage activity [%]	99.7	23.5	99.7	10.3	27.0	15.6
inhibitory efficiency [%]	13.9	53.0	26.6	95.0	0.0	81.7

	WT	A57V
IC_{50}	200 nM	596 nM

Fig. 4 – Prediction of the amino acid residues of XMRV PR interacting with APV. A. The predicted 3D-structure for the interaction between XMRV PR and APV. This homology modeling was based on the HIV-1 PR and APV complex as a template. B. Sequence alignment of XMRV PR and HIV-1 PR. The amino acids related to interaction of APV with HIV-1 PR and the corresponding amino acids in XMRV PR are highlighted with red letters. C. Cleavage activity of XMRV PR-WT and its mutants in the presence of 1 μM APV or equivalent amount of DMSO (control). Lower panel is cleavage activity and inhibitory efficiency (APV value/DMSO value) for each XMRV PR. D. Dose–response curve of the inhibitory rate of PR-WT or PR-A57V by APV.

A

No.	Symbol	Gene bank accession No.	aa	Da	Cleavage by XMRV PR
1	ANXA1	BC001275	346	38714	No
2	GJB2	BC017048	226	26215	No
3	ARL11	BC013150	196	21391	Yes
4	RND3	BC012513	244	27368	No
5	FHIT	BC032336	147	16858	No
6	KCNRG	BC032343	272	31048	No
7	SFRP1	BC036503	314	35386	No
8	WNT5A	BC064694	380	42339	No
9	SMAD3	BC050743	425	48081	No
10	p21	BC000312	164	18119	No
11	NME1	BC000293	152	17149	No
12	MAL	BC000458	153	16714	No
13	SPINT2	BC001668	252	28228	No
14	SLC39A1	BC003152	324	34250	No
15	CHEK2	BC004207	543	60915	No
16	DKK3	BC007660	350	38291	Yes
17	TIMP3	BC014277	211	24145	No
18	BAX	BC014175	192	21184	Yes
19	ALPL	BC021289	524	57305	No
20	CST6	BC031334	149	16510	No
21	KLF5	BC042131	457	50792	No
22	CAV1	BC082246	178	20472	No
23	PTEN	BC005821	403	47166	Yes
24	TP53	BC003596	393	43653	No

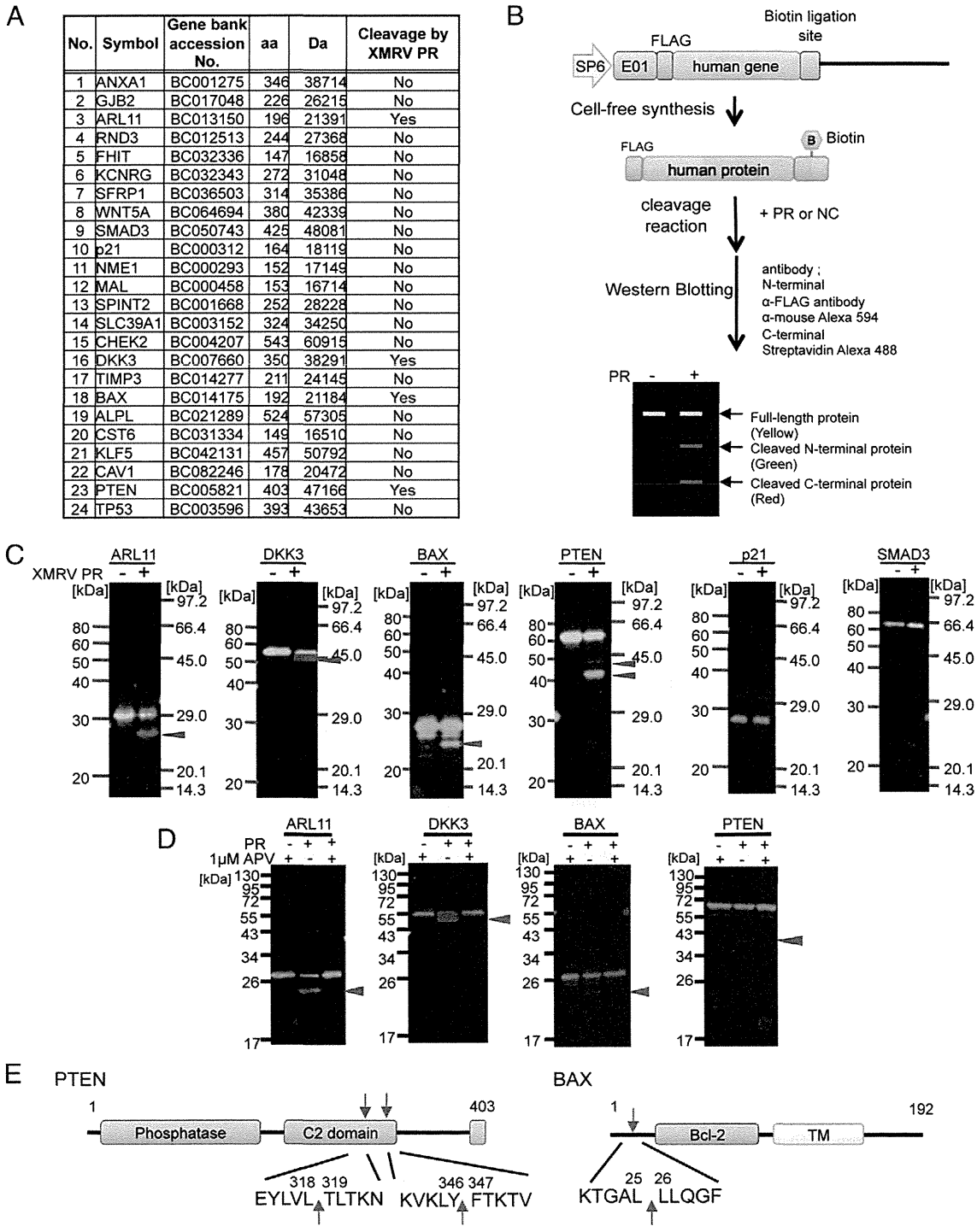


Fig. 5 – Screening of host proteins cleaved by XMRV PR in vitro. A. The list of human tumor suppressor proteins tested in this study. B. Scheme of the tester proteins construction and the cleavage assay system by immunoblotting. The genes were amplified by PCR with primer sets containing either FLAG or biotin ligation site (bls) in the flanking sequence, respectively. The recombinant host proteins flanking FLAG and biotin (FLAG-X-biotin) were incubated with XRMV PR at 37 °C for 2 h followed by SDS-PAGE. The proteins were detected using anti-FLAG-Alexa592 antibody (green) and Alexa488-conjugated streptavidin (red). Full-length protein is seen as a yellow band. C. Tester proteins were treated with XMRV PR or carrier. 2-color immunoblot analysis was performed as in Materials and methods. D. Tester proteins were treated with XMRV PR in the absence or presence of amprenavir. Immunoblot analysis was performed as in C. E. Identification of the cleavage site in the XMRV PR amino acid sequence.

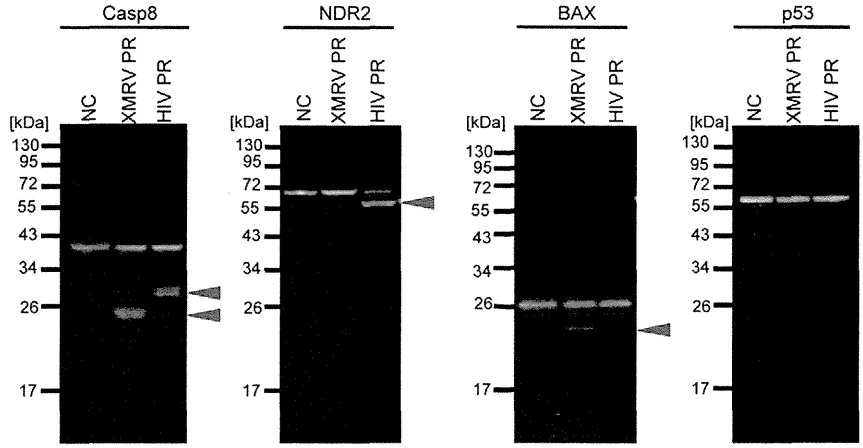


Fig. 6 – Comparative analysis of host proteins cleaved by XMRV PR and HIV-1 PR. The recombinant host proteins flanking FLAG and biotin (FLAG-X-biotin) were incubated with either XRMV PR or HIV-1 PR at 37 °C for 2 h followed by SDS-PAGE. The proteins were detected using anti-FLAG-Alexa592 antibody (green) and Alexa488-conjugated streptavidin (red). Full-length protein is seen as a yellow band. Arrows depict the cleavage products.

by retroviral proteases in infected cells. However, this cell-based will run into difficulty identifying individual substrates if several host proteases act simultaneously on the substrate. To circumvent this potential problem we developed the cell-free in vitro method for the identification of substrates cleavable by XMRV PR. Wheat extracts purified rarely include endogenous proteases that can interfere with the proteolytic reaction, making them suitable for the cell-free protease assay.

Tumor suppressor proteins play a major role in preventing tumor initiation. Our current results demonstrate that XMRV PR can cleave PTEN and BAX tumor suppressors as well as the intrinsic substrate XMRV Gag. It has been reported that the C-terminal region of PTEN is important for the protein's stability, and the C-terminal deletion mutant is degraded rapidly in cells [37]. Since XMRV cleaves within the C-terminal region, the native function and stability of PTEN might be abrogated by XMRV infection. The N-terminal region of BAX has been demonstrated to mediate its activity in apoptosis [38]. We demonstrated in the present study that XMRV PR can cleave the N-terminal region of BAX, suggesting that XMRV infection might affect the activity of BAX protein.

A biochemical approach to the evaluation of PR-inhibitor susceptibility has been attempted previously using several related methods [39,40]. The essence of each of these procedures is the synthesis of catalytically-active PR and substrate peptide and inhibitor in vitro, and measurement of the amount of substrate cleavage. The advantage of this approach is that it can directly detect the catalytic activity of PR. However, it is often difficult to produce sufficient quantities of enzymatically active viral PR in conventional cell-based protein expression systems such as *E. coli* or insect cells. In our current study, we successfully created catalytically-active XMRV PR in a cell-free system that, when mixed with a reporter substrate flanked with N- and C-terminal fluorophores, substrate cleavage could be assayed by AlphaScreen or 2-color IB. This approach directly evaluates the cleavage activity of the PR and, in addition, cleavage sites can be estimated by the size of cleavage products. The current availability of full-length cDNA libraries, derived from higher eukaryotes, will facilitate the in vitro synthesis of full-length proteins, making this cell-free system approach could further be applicable to the assay of a broad range of, not only viral, but also host proteases.

5. Conclusion

We have delineated the molecular and enzymatic characteristics of XMRV PR by utilizing wheat-germ cell-free protein synthesis and AlphaScreen. Furthermore, we have developed an in vitro cleavage assay for drug screening based on the enzymatic activity. Our results suggest that XMRV-protease cleavage of certain host proteins and inhibited by APV. Further in vivo studies with XMRV-infected cells will be necessary to confirm a molecular link between XMRV and human diseases.

Acknowledgments

We thank Drs. G. Quinn, Y. Kojima and A. Kudo for the discussion and comments. This work was supported in part by grants from the Ministry of Education, Culture, Sports, Science and Technology of Japan and Research Grants on HIV/AIDS Health Labour Sciences Research Grant from The Ministry of Health Labour and Welfare of Japan to A.R. MK was supported by grants from MEXT, JST, Sumitomo-Denko and Iwatani.

REFERENCES

[1] Urisman A, Molinaro RJ, Fischer N, Plummer SJ, Casey G, Klein EA, et al. Identification of a novel Gammaretrovirus in prostate tumors of patients homozygous for R462Q RNASEL variant. PLoS Pathog 2006;2:e25.

[2] Schlaberg R, Choe DJ, Brown KR, Thaker HM, Singh IR. XMRV is present in malignant prostatic epithelium and is

associated with prostate cancer, especially high-grade tumors. Proc Natl Acad Sci U S A 2009;106:16351–6.
[3] Knox K, Carrigan D, Simmons G, Teque F, Zhou Y, Hackett Jr J, et al. No evidence of murine-like gammaretroviruses in CFS patients previously identified as XMRV-infected. Science 2011;333:94–7.
[4] van Kuppeveld FJ, van der Meer JW. XMRV and CFS–the sad end of a story. Lancet 2012;379:e27–8.
[5] Simmons G, Glynn SA, Komaroff AL, Mikovits JA, Tobler LH, Hackett Jr J, et al. Failure to confirm XMRV/MLVs in the blood of patients with chronic fatigue syndrome: a multi-laboratory study. Science 2011;334:814–7.
[6] Paprotka T, Delviks-Frankenberry KA, Cingoz O, Martinez A, Kung HJ, Tepper CG, et al. Recombinant origin of the retrovirus XMRV. Science 2011;333:97–101.
[7] Groom HC, Yap MW, Galao RP, Neil SJ, Bishop KN. Susceptibility of xenotropic murine leukemia virus-related virus (XMRV) to retroviral restriction factors. Proc Natl Acad Sci U S A 2010;107:5166–71.
[8] Dong B, Silverman RH. Androgen stimulates transcription and replication of xenotropic murine leukemia virus-related virus. J Virol 2010;84:1648–51.
[9] Abudu A, Takaori-Kondo A, Izumi T, Shirakawa K, Kobayashi M, Sasada A, et al. Murine retrovirus escapes from murine APOBEC3 via two distinct novel mechanisms. Curr Biol 2006;16:1565–70.
[10] Ventoso I, Blanco R, Perales C, Carrasco L. HIV-1 protease cleaves eukaryotic initiation factor 4G and inhibits cap-dependent translation. Proc Natl Acad Sci U S A 2001;98:12966–71.
[11] Zaragoza C, Saura M, Padalko EY, Lopez-Rivera E, Lizarbe TR, Lamas S, et al. Viral protease cleavage of inhibitor of kappaBalpha triggers host cell apoptosis. Proc Natl Acad Sci U S A 2006;103:19051–6.
[12] Takai K, Sawasaki T, Endo Y. Practical cell-free protein synthesis system using purified wheat embryos. Nat Protoc 2010;5:227–38.
[13] Kamura N, Sawasaki T, Kasahara Y, Takai K, Endo Y. Selection of 5′-untranslated sequences that enhance initiation of translation in a cell-free protein synthesis system from wheat embryos. Bioorg Med Chem Lett 2005;15:5402–6.
[14] Tadokoro D, Takahama S, Shimizu K, Hayashi S, Endo Y, Sawasaki T. Characterization of a caspase-3-substrate kinome using an N- and C-terminally tagged protein kinase library produced by a cell-free system. Cell Death Dis 2010;1:e89.
[15] Akagi T, Shimizu K, Takahama S, Iwasaki T, Sakamaki K, Endo Y, et al. Caspase-8 cleavage of the interleukin-21 (IL-21) receptor is a negative feedback regulator of IL-21 signaling. FEBS Lett 2011;585:1835–40.
[16] Sakuma R, Sakuma T, Ohmine S, Silverman RH, Ikeda Y. Xenotropic murine leukemia virus-related virus is susceptible to AZT. Virology 2010;397:1–6.
[17] Sawasaki T, Kamura N, Matsunaga S, Saeki M, Tsuchimochi M, Morishita R, et al. Arabidopsis HY5 protein functions as a DNA-binding tag for purification and functional immobilization of proteins on agarose/DNA microplate. FEBS Lett 2008;582:221–8.
[18] Takahashi H, Nozawa A, Seki M, Shinozaki K, Endo Y, Sawasaki T. A simple and high-sensitivity method for analysis of ubiquitination and polyubiquitination based on wheat cell-free protein synthesis. BMC Plant Biol 2009;9:39.
[19] Sawasaki T, Gouda MD, Kawasaki T, Tsuboi T, Tozawa Y, Takai K, et al. The wheat germ cell-free expression system: methods for high-throughput materialization of genetic information. Methods Mol Biol 2005;310:131–44.
[20] Sawasaki T, Hasegawa Y, Tsuchimochi M, Kamura N, Ogasawara T, Kuroita T, et al. A bilayer cell-free protein synthesis system for high-throughput screening of gene products. FEBS Lett 2002;514:102–5.
[21] Matsuoka K, Komori H, Nose M, Endo Y, Sawasaki T. Simple screening method for autoantigen proteins using the N-terminal biotinylated protein library produced by wheat cell-free synthesis. J Proteome Res 2010;9:4264–73.
[22] Baker D, Sali A. Protein structure prediction and structural genomics. Science 2001;294:93–6.
[23] Labute P. The generalized Born/volume integral implicit solvent model: estimation of the free energy of hydration using London dispersion instead of atomic surface area. J Comput Chem 2008;29:1693–8.
[24] Li M, Dimaio F, Zhou D, Gustchina A, Lubkowski J, Dauter Z, et al. Crystal structure of XMRV protease differs from the structures of other retropepsins. Nat Struct Mol Biol 2011;18:227–9.
[25] Johnson VA, Brun-Vezinet F, Clotet B, Gunthard HF, Kuritzkes DR, Pillay D, et al. Update of the drug resistance mutations in HIV-1. Top HIV Med 2008;16:138–45.
[26] Alvarez E, Castello A, Menendez-Arias L, Carrasco L. HIV protease cleaves poly(A)-binding protein. Biochem J 2006;396:219–26.
[27] Bellecave P, Sarasin-Filipowicz M, Donze O, Kennel A, Gouttenoire J, Meylan E, et al. Cleavage of mitochondrial antiviral signaling protein in the liver of patients with chronic hepatitis C correlates with a reduced activation of the endogenous interferon system. Hepatology 2010;51:1127–36.
[28] Nie Z, Phenix BN, Lum JJ, Alam A, Lynch DH, Beckett B, et al. HIV-1 protease processes procaspase 8 to cause mitochondrial release of cytochrome c, caspase cleavage and nuclear fragmentation. Cell Death Differ 2002;9:1172–84.
[29] Devroe E, Silver PA, Engelman A. HIV-1 incorporates and proteolytically processes human NDR1 and NDR2 serine-threonine kinases. Virology 2005;331:181–9.
[30] Knouf EC, Metzger MJ, Mitchell PS, Arroyo JD, Chevillet JR, Tewari M, et al. Multiple integrated copies and high-level production of the human retrovirus XMRV (xenotropic murine leukemia virus-related virus) from 22Rv1 prostate carcinoma cells. J Virol 2009;83:7353–6.
[31] Stieler K, Schindler S, Schlomm T, Hohn O, Bannert N, Simon R, et al. No detection of XMRV in blood samples and tissue sections from prostate cancer patients in Northern Europe. PLoS One 2011;6:e25592.
[32] Rodriguez JJ, Goff SP. Xenotropic murine leukemia virus-related virus establishes an efficient spreading infection and exhibits enhanced transcriptional activity in prostate carcinoma cells. J Virol 2010;84:2556–62.
[33] Kim S, Kim N, Dong B, Boren D, Lee SA. Das Gupta J, et al. Integration site preference of xenotropic murine leukemia virus-related virus, a new human retrovirus associated with prostate cancer. J Virol 2008;82:9964–77.
[34] Metzger MJ, Holguin CJ, Mendoza R, Miller AD. The prostate cancer-associated human retrovirus XMRV lacks direct transforming activity but can induce low rates of transformation in cultured cells. J Virol 2010;84:1874–80.
[35] Li M, Gustchina A, Matuz K, Tozser J, Namwong S, Goldfarb NE, et al. Structural and biochemical characterization of the inhibitor complexes of xenotropic murine leukemia virus-related virus protease. FEBS J 2011;278:4413–24.
[36] Etlinger JD, Gu M, Li X, Weitman D, Rieder RF. Protease/inhibitor mechanisms involved in ATP-dependent proteolysis. Rev Biol Cell 1989;20:197–216.
[37] Georgescu MM, Kirsch KH, Akagi T, Shishido T, Hanafusa H. The tumor-suppressor activity of PTEN is regulated by its carboxyl-terminal region. Proc Natl Acad Sci U S A 1999;96:10182–7.
[38] Toyota H, Yanase N, Yoshimoto T, Moriyama M, Sudo T, Mizuguchi J. Calpain-induced Bax-cleavage product is a more potent inducer of apoptotic cell death than wild-type Bax. Cancer Lett 2003;189:221–30.
[39] Dreyer GB, Metcalf BW, Tomaszek Jr TA, Carr TJ, Chandler 3rd AC, Hyland L, et al. Inhibition of human immunodeficiency virus 1 protease in vitro: rational design of substrate analogue inhibitors. Proc Natl Acad Sci U S A 1989;86:9752–6.
[40] Hoffmann D, Buchberger B, Nemetz C. In vitro synthesis of enzymatically active HIV-1 protease for rapid phenotypic resistance profiling. J Clin Virol 2005;32:294–9.

Biochemical and Biophysical Research Communications 425 (2012) 284–289

Contents lists available at SciVerse ScienceDirect

Biochemical and Biophysical Research Communications

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

NMR study of xenotropic murine leukemia virus-related virus protease in a complex with amprenavir

Ayako Furukawa [a], Hideyasu Okamura [a], Ryo Morishita [b,c], Satoko Matsunaga [c], Naohiro Kobayashi [d], Takahisa Ikegami [d], Tsutomu Kodaki [a,e], Akifumi Takaori-Kondo [f], Akihide Ryo [c], Takashi Nagata [a,e,*], Masato Katahira [a,e,g,*]

[a] Institute of Advanced Energy, Kyoto University, Gokasho, Uji, Kyoto 611-0011, Japan
[b] CellFree Sciences Co. Ltd., Ehime University, Venture Business Laboratory, Matsuyama 790-8577, Japan
[c] Department of Microbiology, Yokohama City University School of Medicine, 3-9 Fukuura, Kanazawa-ku, Yokohama 236-0004, Japan
[d] Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan
[e] Graduate School of Energy Science, Kyoto University, Gokasho, Uji, Kyoto 611-0011, Japan
[f] Department of Hematology and Oncology, Graduate School of Medicine, Kyoto University, Sakyo-ku, Kyoto 606-8507, Japan
[g] CREST, JST, 5-3 Yonban-cho, Chiyoda-ku, Tokyo 102-8666, Japan

ARTICLE INFO

Article history:
Received 14 July 2012
Available online 25 July 2012

Keywords:
XMRV
Protease
Cell-free protein synthesis
NMR

ABSTRACT

Xenotropic murine leukemia virus-related virus (XMRV) is a virus created through recombination of two murine leukemia proviruses under artificial conditions during the passage of human prostate cancer cells in athymic nude mice. The homodimeric protease (PR) of XMRV plays a critical role in the production of functional viral proteins and is a prerequisite for viral replication. We synthesized XMRV PR using the wheat germ cell-free expression system and carried out structural analysis of XMRV PR in a complex with an inhibitor, amprenavir (APV), by means of NMR. Five different combinatorially ^{15}N-labeled samples were prepared and backbone resonance assignments were made by applying Otting's method, with which the amino acid types of the [^{1}H, ^{15}N] HSQC resonances were automatically identified using the five samples (Wu et al., 2006) [14]. A titration experiment involving APV revealed that one APV molecule binds to one XMRV PR dimer. For many residues, two distinct resonances were observed, which is thought to be due to the structural heterogeneity between the two protomers in the APV:XMRV PR = 1:2 complex. PR residues at the interface with APV have been identified on the basis of chemical shift perturbation and identification of the intermolecular NOEs by means of filtered NOE experiments. Interestingly, chemical shift heterogeneity between the two protomers of XMRV PR has been observed not only at the interface with APV but also in regions apart from the interface. This indicates that the structural heterogeneity induced by the asymmetry of the binding of APV to the XMRV PR dimer is transmitted to distant regions. This is in contrast to the case of the APV:HIV-1 PR complex, in which the structural heterogeneity is only localized at the interface. Long-range transmission of the structural change identified for the XMRV PR complex might be utilized for the discovery of a new type of drug.

1. Introduction

Xenotropic murine leukemia virus-related virus (XMRV) has been implicated in prostate cancer [1] and chronic fatigue syndrome [2]. These reports have attracted much attention since a causal relationship between a retrovirus and a human disease had only been known for HIV and AIDS [3]. After several years of controversial research, a recent study has indicated that XMRV was generated through the recombination of two proviruses during the passaging of human tumors in mice [4]. Thus, XMRV is as not yet regarded as an etiological agent for human diseases, however, it can be considered as a rational example of a gammaretrovirus that can infect human cells and replicate within them.

Proteases (PRs) of viruses such as those of the HIV-1 and hepatitis C ones have been prime targets for antiviral drug development. Although hundreds of inhibitors have been discovered for HIV-1 PR for example, only a few of them are potent antiprotease drugs that are in clinical use [5]. Additionally, long-term administration of these drugs causes the emergence of drug-resistant mutants [6]. It is expected that a detailed description of the interaction between a newly found protease of the retrovirus

* Corresponding authors. Address: Institute of Advanced Energy, Kyoto University, Gokasho, Uji, Kyoto 611-0011, Japan (M. Katahira).
E-mail addresses: nagatat@iae.kyoto-u.ac.jp (T. Nagata), katahira@iae.kyoto-u.ac.jp (M. Katahira).

0006-291X/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.bbrc.2012.07.083

family and its cognate substrate or inhibitor may provide valuable information.

XMRV PR is a homodimer, each subunit (protomer) comprising 125 amino acids with a single catalytic Asp residue [7,8]. XMRV PR processes viral polyproteins to yield mature proteins required in the viral life cycle, as is the case for HIV-1 PR. The sequence identity between and similarity of these two proteins are 21% and 27%, respectively, over 99 amino acid residues with a single gap (sequences aligned with ClustalW2 [9,10]). The crystal structures of XMRV PR have been solved in both the free form and complex forms with several inhibitors [7,8]. The fold of XMRV PR is similar to that of HIV-1 PR. The overall structural similarity between the protomer of XMRV PR (PDB ID: 3NR6; [7]) and that of HIV-1 PR (99 amino acids; PDB ID: 3HVP; [11]) in the free form was found to be 1.04 Å of RMSD over 53 C^{α} atoms (calculated using CHIMERA [12]).

In the present study, we used our established wheat germ protein production system to obtain otherwise cell-toxic XMRV PR [13]. This system allowed us to obtain not only a large amount of free and amprenavir (APV)-bound XMRV PR, but also uniformly and combinatorially isotope-labeled samples. The latter isotope-labeled samples made it possible to apply Otting's rapid assignment strategy [14] in combination with conventional protocols [15]. It was critical to utilize this strategy for accomplishment of the assignments. We found that the asymmetry of the binding for the APV:XMRV PR = 1:2 complex results in the structural heterogeneity between the two protomers of XMRV PR at the interface, and, but more importantly, in regions distant from the interface. In the case of the binding of APV to HIV-1 PR, in contrast, the asymmetry causes structural heterogeneity only at the interface.

2. Materials and methods

2.1. Sample preparation

XMRV PR comprising 1–125 residues with an additional threonine residue at the C-terminal end, was synthesized by means of our wheat germ cell-free expression system as an N-terminal glutathione-S-transferase (GST) fusion protein. The XMRV PR fragment was cloned by PCR from the XMRV VP62 clone [16,17] (UniProt ID: A1Z651) and subcloned into pEU-E01-GW [18] with the DNA sequences of the GST gene and Tobacco Etch Virus Protease (TEV) cleavage site. This plasmid DNA was subjected to *in vitro* transcription and cell-free protein synthesis with the wheat germ protein production system [13]. Cell-free protein production was carried out using the ENDEXT.AN. Wheat Germ Expression Kit and according to the instructions provided by the supplier (Cell-Free Sciences Co. Ltd., Matsuyama, Japan). The cell-free protein production and cleavage purification were also carried out using an automatic robot, Protemist DTII (CellFree Sciences Co. Ltd.), basically according to manufacturer's instructions. GST-XMRV PR was synthesized either in the absence or presence of APV. After the synthesis, the protein solution was loaded onto a Glutathione-Sepharose 4B column (GE Healthcare), and the bound GST-XMRV PR was washed thoroughly with PBS buffer. Subsequently, GST-fused TEV protease (Nacalai Tesque) was applied to cleave the XMRV PR off on-column. The XMRV PR was then collected and dialyzed against either 20 mM sodium phosphate (pH 6.5), 50 mM NaCl and 10 μM APV for the APV-bound XMRV PR or 20 mM Tris–HCl (pH 8.0) and 50 mM NaCl for APV-free XMRV PR. As substrates for protein synthesis, isotropically labeled amino acids were used to synthesize uniformly ^{15}N- and ^{15}N, ^{13}C-labeled ([U-^{15}N] and [U-^{13}C, U-^{15}N], respectively) XMRV PRs. Five different combinatorially ^{15}N-labeled samples of XMRV PR were prepared following Otting's labeling scheme, as shown in Fig. 1C [14].

2.2. NMR spectroscopy

For XMRV PR in a complex with APV, NMR spectra were acquired at 22.9 °C using Bruker 600 and 950 MHz spectrometers each equipped with a cryoprobe. Data were processed and analyzed using NMRPipe [19] and Kujira [20]. The sample solutions comprised 0.1–0.3 mM protein in a complex with APV, 20 mM sodium phosphate (pH 6.5), 50 mM NaCl, 10 μM APV, and 5% D_2O. The amino acid in [^{1}H, ^{15}N] HSQC spectra were identified using the five different combinatorially ^{15}N-labeled samples according to Wu et al. [14]. Sequential assignments of the main chain resonances were made by means of conventional triple resonance experiments [15,21,22]. The main chain $^{1}H^{N}$ and ^{15}N chemical shift heterogeneity between the two protomers in the APV:XMRV PR = 1:2 complex was defined as $\{(\Delta\delta^{1}H^{N})^2 + (\Delta\delta^{15}N/5)^2\}^{1/2}$, where $\Delta\delta^{1}H^{N}$ and $\Delta\delta^{15}N$ are the chemical shift differences between the two protomers for $^{1}H^{N}$ and ^{15}N resonances, respectively. The secondary structures of XMRV PR in a complex with APV was identified based on $^{13}C^{\alpha}$, $^{13}C^{\beta}$ and $^{13}C'$ chemical shift values [23]. The intermolecular NOEs between XMRV PR and APV were obtained from 2D [F1, F2] ^{13}C, ^{15}N-filtered NOESY (120 ms mixing time), 2D [F2] ^{13}C, ^{15}N-filtered NOESY (120 ms mixing time), 3D [F1] ^{13}C, ^{15}N-filtered, [F2] ^{13}C-edited NOESY (120 ms mixing time) [24], and 3D ^{13}C-edited NOESY (78 ms mixing time) spectra.

For free XMRV PR, NMR spectra were acquired at 15.0 °C using a Bruker 600 MHz spectrometer equipped with a cryoprobe. The sample solutions comprised ~0.1 mM protein, 20 mM Tris–HCl (pH 8.0), 50 mM NaCl and 5% D_2O. Partial assignments of the main chain $^{1}H^{N}$ and ^{15}N resonances were made using 3D ^{15}N-edited NOESY spectra utilizing on the assignments for the XMRV PR–APV complex as a reference. The main chain $^{1}H^{N}$ and ^{15}N chemical shift perturbations between the free and APV-bound forms of XMRV PR were defined as $\{(\Delta\delta^{1}H^{N})^2 + (\Delta\delta^{15}N/5)^2\}^{1/2}$, where $\Delta\delta^{1}H^{N}$ and $\Delta\delta^{15}N$ are the chemical shift differences between the free and APV-bound forms for $^{1}H^{N}$ and ^{15}N resonances, respectively. For the APV-bound form, the average of the chemical shift values for the two protomers was used to calculate the differences.

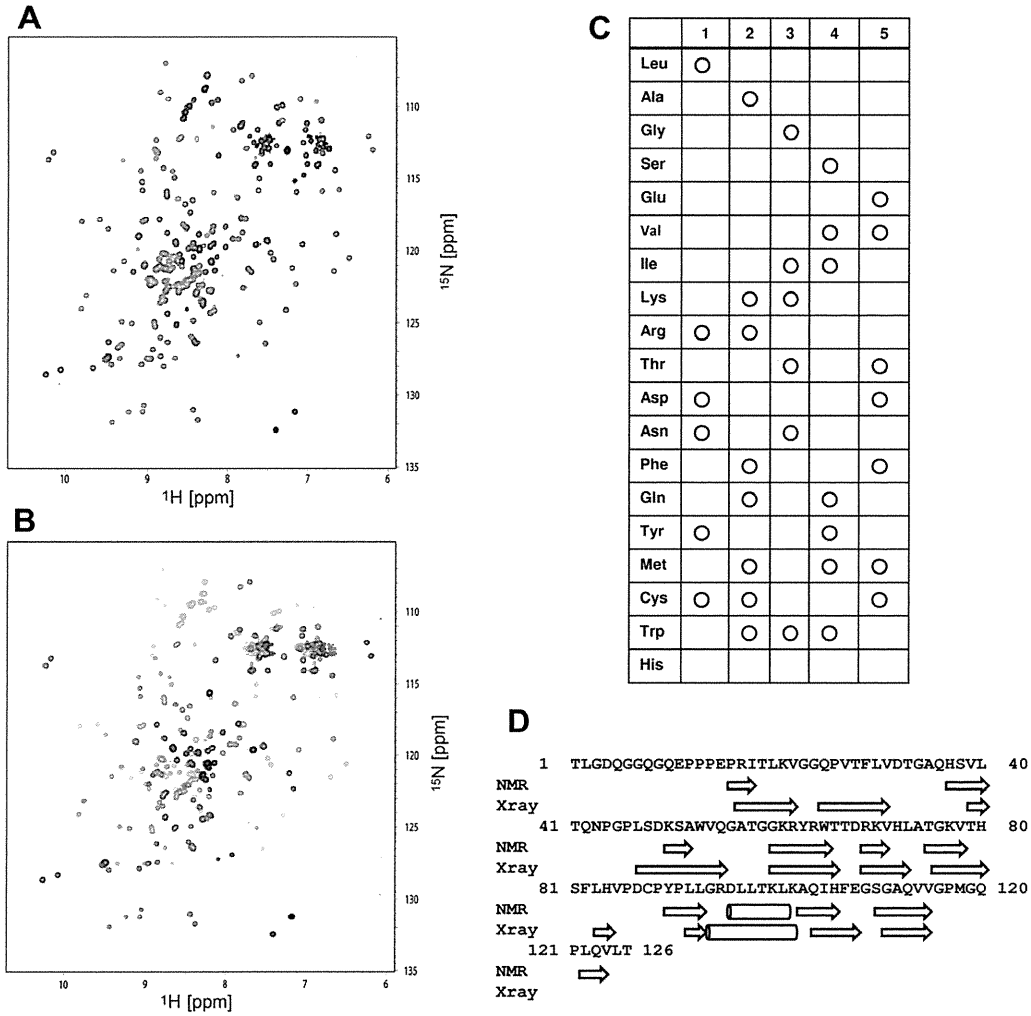
3. Results and discussion

3.1. Characteristics of the [^{1}H, ^{15}N] HSQC spectrum of the XMRV PR in complex with APV

We have successfully obtained cell-toxic XMRV PR by using a cell-free protein synthesis system in both the presence and absence of APV. Moreover, we were able to prepare [U-^{15}N] and [U-^{13}C, U-^{15}N] labeled, and five different combinatorially ^{15}N-labeled samples of XMRV PR, all of which critically facilitated our NMR analyses.

First, titration with APV was carried out. Disappearance of the resonances of XMRV PR in the free state and simultaneous appearance of new resonances for the APV-bound form were observed in the [^{1}H, ^{15}N] HSQC spectrum during the titration. This indicates that the free and bound forms of XMRV PR are in a slow exchange regime on an NMR time scale. Such spectral changes for XMRV PR were seen until APV:XMRV PR = 1:2, but no further change was seen on further addition of APV. This indicates that the stoichiometry is one APV molecule per one XMRV PR dimer (two XMRV PR monomers).

Each protomer comprises 126 amino acid residues, including 12 proline ones. If the homodimer exhibits symmetry, maximally 113 backbone resonances are expected to be observed, because those of the N-terminus and proline residues do not appear. However, the [^{1}H, ^{15}N] HSQC spectrum contained more than 1.5-times as many



	1	2	3	4	5
Leu	O				
Ala		O			
Gly			O		
Ser				O	
Glu					O
Val				O	O
Ile			O	O	
Lys		O	O		
Arg	O	O			
Thr			O		O
Asp	O				O
Asn	O		O		
Phe		O			O
Gln		O		O	
Tyr	O			O	
Met		O		O	O
Cys	O	O			O
Trp		O	O	O	
His					

Fig. 1. (A) [^{1}H, ^{15}N] HSQC spectrum of uniformly ^{13}C, ^{15}N-labeld XMRV PR in a complex with APV. (B) Superposition of the [^{1}H, ^{15}N] HSQC spectra of five different combinatorially ^{15}N-labeled XMRV PRs (1–5 of (C)) in a complex with APV, colored red, cyan, green, magenta and blue, respectively. (C) Five different combinatorial labeling schemes [14]. (D) Comparison of the secondary structures of XMRV PR in a complex with APV between the crystal [8] and solution (this work) states. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

backbone resonances as expected (Fig. 1A). This observation reveals that the APV-bound XMRV PR exhibits asymmetry.

3.2. Sequence-specific resonance assignments and identification of the secondary structure of the XMRV PR in a complex with APV

We applied Otting's strategy, with which amino acids could be automatically identified on the basis of the comparison of five spectra [14]. Assignments of the $^{1}H^{N}$, ^{15}N, $^{13}C^{\alpha}$, $^{13}C^{\beta}$, and $^{13}C'$ backbone resonances of XMRV PR in complex with APV were carried out with the knowledge of the amino acid. Almost all of the observed resonances were successfully assigned. It was found that two distinct backbone resonances were observed for many residues of XMRV PR in a complex with APV, which resulted in more than 1.5 times as many backbone resonances as expected. The observation of two distinct peaks for many residues must be due to that the two protomers of the XMRV PR dimer are not identical to each other in the APV:XMRV PR = 1:2 complex.

The secondary structure of the XMRV PR in a complex with APV was identified according to the $^{13}C^{\alpha}$, $^{13}C^{\beta}$ and $^{13}C'$ chemical shift values [23] (Fig. 1D). It was noticed that the two protomers exhibited almost identical secondary structures. The solution structure of XMRV PR in a complex with APV turned out to be basically the same as the crystal structure, which was solved recently [8], as shown in Fig. 1D.

3.3. Binding mode of APV as to XMRM PR in solution

The quality of the [^{1}H, ^{15}N] HSQC spectrum of APV-free XMRV PR appeared to be rather poor due to line broadening of resonances in comparison with in the case of APV-bound XMRV PR. By comparing the spectral patterns of the 3D ^{15}N-edited NOESY spectra between the APV-bound and APV-free forms of XMRV PR, we achieved partial assignment of the main chain $^{1}H^{N}$ and ^{15}N resonances of APV-free XMRV PR. It was found on this analysis that only one set of resonances was observed for APV-free XMRV PR. This indicates that the two protomers are symmetric as to each other in the APV-free form. We then calculated the chemical shift perturbations of the backbone resonances of XMRV PR upon APV binding, and mapped them onto the crystal structure (Fig. 2A). The largest perturbations were observed for the residues located close to APV in the crystal structure (Fig. 2A). This strongly suggests that in solution APV resides at a similar position to as observed in the crystal.

In order to further confirm the interface of XMRV PR with APV, identification of intermolecular NOEs was attempted. First, the $^{1}H^{N}$

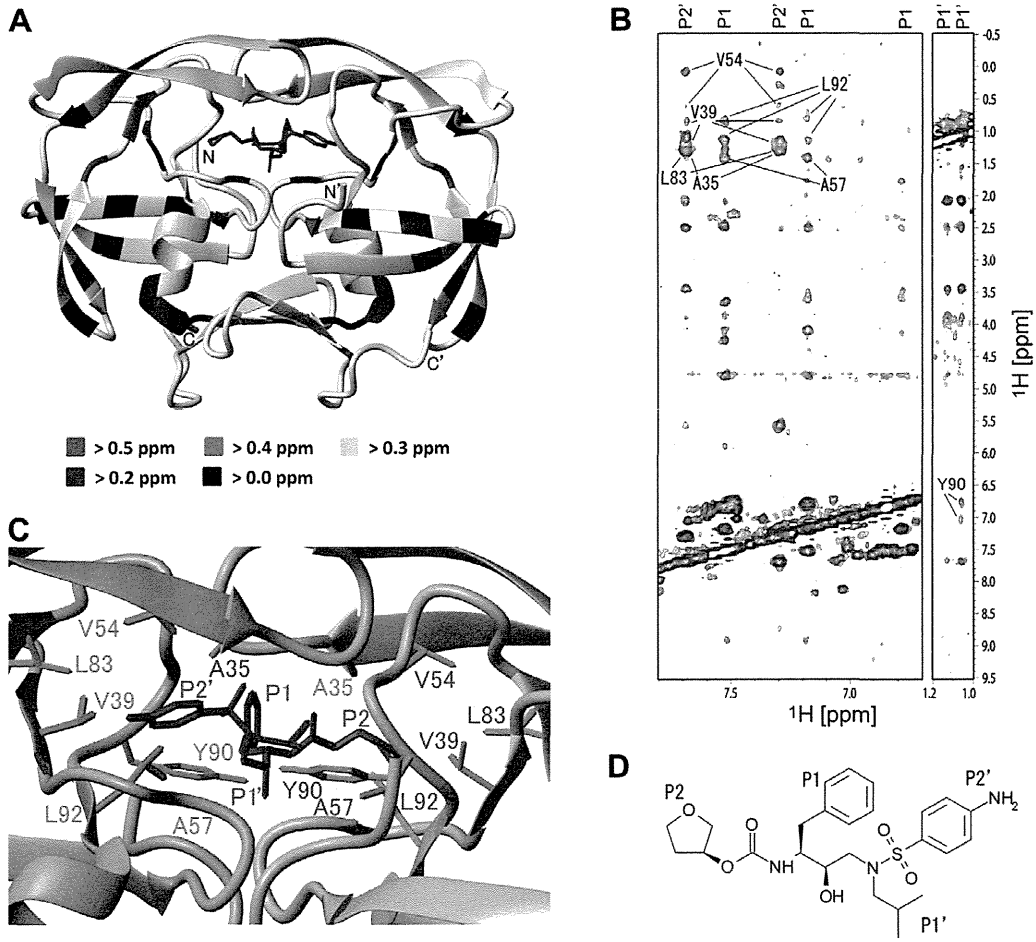


Fig. 2. (A) Mapping of the $^{1}H^{N}$, ^{15}N chemical shift perturbations between the APV-free and APV-bound forms of XMRV PR onto the crystal structure of the APV:XMRV PR complex (PDB accession number 3SM2). The N- and C-termini of the two protomers are denoted by N and C, and N′ and C′, respectively. (B) 2D [F2] ^{13}C, ^{15}N-filtered NOESY (120 ms mixing time) (blue) and 2D [F1, F2] ^{13}C, ^{15}N-filtered NOESY (120 ms mixing time) (red) of XMRV PR in a complex with APV. The assigned intermolecular NOE peaks for XMRV PR and APV are labeled. (C) APV bound to XMRV PR in the crystal (PDB accession number 3SM2). The protomers of XMRV PR are shown in yellow and gray, respectively. APV, is colored magenta. (D) Chemical structure of APV. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

resonances of APV bound to XMRV PR were assigned using the 2D [F1, F2] ^{13}C, ^{15}N-filtered NOESY spectrum (Fig. 2B). Secondly, the intermolecular NOEs between XMRV PR and APV were discriminated by comparing the 2D [F1, F2] ^{13}C, ^{15}N-filtered NOESY and 2D [F2] ^{13}C, ^{15}N-filtered NOESY spectra (Fig. 2B). That is, NOEs observed in the latter spectrum but not in the former one were judged to be intermolecular NOEs. Finally, these intermolecular NOEs were assigned by combinational use of the 3D [F1] ^{13}C, ^{15}N-filtered, [F2] ^{13}C-edited NOESY, 3D ^{13}C-edited NOESY, and 3D ^{15}N-edited NOESY spectra. Consequently, we were able to identify and assign the intermolecular NOEs between APV (P1, P1′, P2 and P2′), and either the methyl groups of A35, V39, V54, A57, L83 and L92 or the aromatic ring of Y90 (Fig. 2B). These intermolecular NOEs turned out to be fully consistent with the crystal structure of the APV:XMRV PR complex. All of the listed amino acid residues are indeed located close to APV in the crystal structure (Fig. 2C). This revealed that in solution APV binds to XMRV in the same manner as observed in the crystal.

3.4. Long-range transmission of the structural heterogeneity between two protomers in the APV:XMRV PR complex

The $^{1}H^{N}$, ^{15}N chemical shift heterogeneity of the backbone resonances between the two protomers of XMRV PR in a complex with APV was calculated and mapped onto the crystal structure of the XMRV PR:APV complex (Fig. 3A). It turned out that chemical shift heterogeneity was observed not only at the APV-binding site (D32, A35, Q36 and L92), but also at positions distant from the APV-binding site (S38, L73, H80, S81, L97 and L122). This suggests that the structural heterogeneity occurs not only at the interface with APV but also in these regions in the APV:XMRV PR = 1:2 complex.

Then, we examined the structural heterogeneity of the crystal structure of the APV:XMRV PR complex. The backbone atoms of the two protomers were superimposed and the differences in position between the C^{α} atoms of the two protomers were calculated to estimate the structural heterogeneity in the crystal. The C^{α}–C^{α} distance was mapped on the crystal structure of the APV:XMRV PR complex (Fig. 3B). It is found that structural heterogeneity is present not only at the interface with APV but also in regions distant in the crystal structure too. Thus, it was found that the structural heterogeneity exhibits long-range transmission in the APV:XMRV PR = 1:2 complex in both solution and crystal states.

3.5. Long-range transmission of the structural heterogeneity is specific to the APV:XMRV PR complex, not being observed in the APV:HIV-1 PR complex

Next, we examined if the long-range transmission of the structural heterogeneity occurs for PR of another virus, HIV-1. The $^{1}H^{N}$, ^{15}N chemical shift heterogeneity between the two protomers of

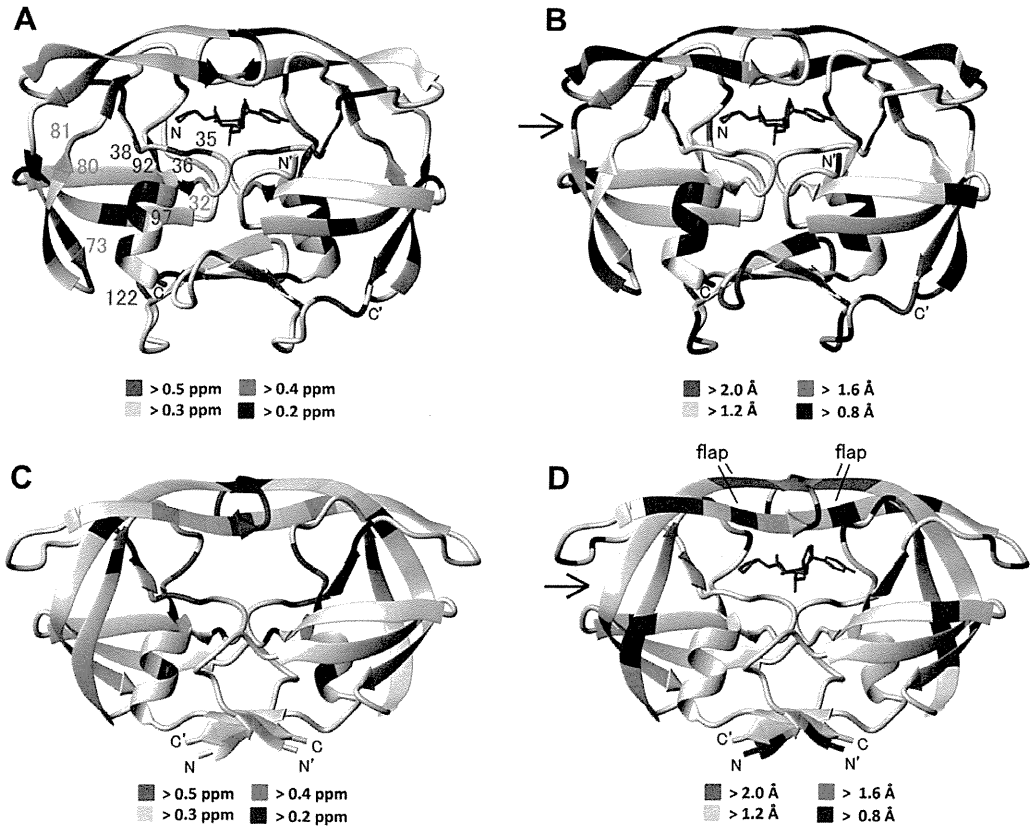


Fig. 3. (A) Mapping of the $^{1}H^{N}$, ^{15}N chemical shift heterogeneity between the two protomers of XMRV PR in a complex with APV onto the crystal structure of the XMRV PR–APV complex (PDB accession number 3SM2). (B) Mapping of the C^{α}–C^{α} distance between the two superimposed protomers of the XMRV PR–APV crystal structure. The shallow groove is indicated by an arrow. (C) Mapping of the $^{1}H^{N}$, ^{15}N chemical shift heterogeneity between the two protomers of the HIV PR (D25N) in a complex with a peptide substrate [25] onto the crystal structure of the HIV PR (D25N)–APV complex. The structure of APV has been deleted. (D) Mapping of the C^{α}–C^{α} distance between the two superimposed protomers in the HIV PR (D25N)–APV crystal structure. The shallow groove is indicated by an arrow. The flap is also indicated.

HIV-1 (D25N) PR in a complex with a peptide substrate was calculated using the previously reported data [25]. Here, heterogeneity in solution was analyzed for the peptide-bound form, because NMR data on HIV-1 in a complex with APV are not available. It has been confirmed that the mutant (D25N) and wild-type HIV-1 PRs exhibited essentially identical backbone chemical shift values, except for those of mutated residue 25 and adjacent residues [25]. Since the peptide substrate is larger than APV in size and has a highly asymmetric structure, this system could be used instead to compare the heterogeneity between the XMRV and HIV-1 PRs. The chemical shift heterogeneity values were mapped onto the crystal structure of the HIV-1 PR (D25N) in a complex with APV (Fig. 3C). The chemical shift heterogeneity was mostly observed for residues located close the substrate/inhibitor-binding site. This indicates that the structural heterogeneity is localized around the substrate binding site and that the structural heterogeneity is not transmitted to distant regions in the HIV-1 PR, which is in contrast to the case of the XMRV PR.

In order to further confirm the localization of the structural heterogeneity for HIV-1 PR, the crystal structure of the APV:HIV-1 PR (D25N) complex was examined. The differences in the positions of the C^{α} atoms of the two protomers were calculated and mapped onto the crystal structure (Fig. 3D). The structural heterogeneity was mostly observed for the residues that are close to the substrate-binding site. This is consistent with the conclusion derived from NMR data in solution for HIV-1 PR. Thus, the long-range transmission of the structural heterogeneity does not occur in HIV-1 PR, but is specific to XMRV-PR.

3.6. Implications for discovery of new types of drug on the basis of the identification of long-range transmission of the structural heterogeneity

It has been pointed out for HIV-1 PR that opening and closing of the flap of the binding site results in compression and expansion of the shallow groove indicated by the arrow in Fig. 3D, respectively [26,27]. Recently, the peripheral surface of HIV-1 PR, which is relatively apart from the binding site, was suggested to be a good target for the discovery of allosteric inhibitors. A fragment-based screen against HIV-1 PR in a complex with TL-3, a universal inhibitor of retroviral PRs, has been employed using the Active Sight fragment library and X-ray crystallography [28]. The authors of that study discovered that 2-methylcyclohexanol binds to the 'exo site' just by the shallow groove [28]. This compound binds to the shallow groove and restricts its compression, and thus prevents the flap from opening. Therefore, it was proposed that the compound can be used as a starting molecule for the development of high-affinity allosteric inhibitors [28]. This example illustrates that a drug that binds to a position distant from the binding site can affect the binding affinity, being a candidate inhibitor. We have shown that the structural heterogeneity is transmitted away in XMRV PR. This means that the heterogeneous structural changes on binding of APV are transmitted to distant regions. This implies conversely that a drug that binds to a distant site may affect the structure of the substrate-binding site and thus inhibit binding of a substrate. This new type of a drug might be discovered more easily for XMRV PR than for HIV-1 PR. Then, in some cases, the discovered drug might

bind to only one protomer of the dimer due to structural heterogeneity, as is the case for 2-methylcyclohexanol [28].

Acknowledgments

This work was supported by Grants from MEXT (23570146, 23657072, 24121714, and 24113710), JST-SENTAN, JST-CREST, the Sumitomo-Denko Foundation and the Iwatani Foundation. This work was performed using the NMR spectrometer with the ultra-high magnetic field under the Cooperative Research Program of Institute for Protein Research, Osaka University.

References

[1] R. Schlaberg, D.J. Choe, K.R. Brown, H.M. Thaker, I.R. Singh, XMRV is present in malignant prostatic epithelium and is associated with prostate cancer, especially high-grade tumors, Proc. Natl. Acad. Sci. USA 106 (2009) 16351–16356.
[2] V.C. Lombardi, F.W. Ruscetti, J. Das Gupta, M.A. Pfost, K.S. Hagen, D.L. Peterson, S.K. Ruscetti, R.K. Bagni, C. Petrow-Sadowski, B. Gold, M. Dean, R.H. Silverman, J.A. Mikovits, Detection of an infectious retrovirus, XMRV, in blood cells of patients with chronic fatigue syndrome, Science 326 (2009) 585–589.
[3] S. Broder, R.C. Gallo, A pathogenic retrovirus (HTLV-III) linked to AIDS, N. Engl. J. Med. 311 (1984) 1292–1297.
[4] T. Paprotka, K.A. Delviks-Frankenberry, O. Cingoz, A. Martinez, H.J. Kung, C.G. Tepper, W.S. Hu, M.J. Fivash Jr., J.M. Coffin, V.K. Pathak, Recombinant origin of the retrovirus XMRV, Science 333 (2011) 97–101.
[5] I.R. Singh, J.E. Gorzynski, D. Drobysheva, L. Bassit, R.F. Schinazi, Raltegravir is a potent inhibitor of XMRV, a virus implicated in prostate cancer and chronic fatigue syndrome, PLoS ONE 5 (2010) e9948.
[6] E.J. Arts, D.J. Hazuda, HIV-1 antiretroviral drug therapy, Cold Spring Harb. Perspect. Med. 2 (2012) a007161.
[7] M. Li, F. Dimaio, D. Zhou, A. Gustchina, J. Lubkowski, Z. Dauter, D. Baker, A. Wlodawer, Crystal structure of XMRV protease differs from the structures of other retropepsins, Nat. Struct. Mol. Biol. 18 (2011) 227–229.
[8] M. Li, A. Gustchina, K. Matuz, J. Tozser, S. Namwong, N.E. Goldfarb, B.M. Dunn, A. Wlodawer, Structural and biochemical characterization of the inhibitor complexes of xenotropic murine leukemia virus-related virus protease, FEBS J. 278 (2011) 4413–4424.
[9] M. Goujon, H. McWilliam, W. Li, F. Valentin, S. Squizzato, J. Paern, R. Lopez, A new bioinformatics analysis tools framework at EMBL-EBI, Nucleic Acids Res. 38 (2010) W695–699.
[10] M.A. Larkin, G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T.J. Gibson, D.G. Higgins, Clustal W and Clustal X version 2.0, Bioinformatics 23 (2007) 2947–2948.
[11] A. Wlodawer, M. Miller, M. Jaskolski, B.K. Sathyanarayana, E. Baldwin, I.T. Weber, L.M. Selk, L. Clawson, J. Schneider, S.B. Kent, Conserved folding in retroviral proteases: crystal structure of a synthetic HIV-1 protease, Science 245 (1989) 616–621.
[12] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera – a visualization system for exploratory research and analysis, J. Comput. Chem. 25 (2004) 1605–1612.
[13] S. Matsunaga, T. Sawasaki, H. Ode, R. Morishita, A. Furukawa, R. Sakuma, W. Sugiura, H. Sato, M. Katahira, A. Takaori-Kondo, N. Yamamoto, A. Ryo, Molecular and enzymatic characterization of XMRV protease by a cell-free proteolytic analysis, J. Proteomics (2012).
[14] P.S. Wu, K. Ozawa, S. Jergic, X.C. Su, N.E. Dixon, G. Otting, Amino-acid type identification in ^{15}N-HSQC spectra by combinatorial selective ^{15}N-labelling, J. Biomol. NMR 34 (2006) 13–21.
[15] J. Cavanagh, W.J. Fairbrother, A.G. Palmer, N.J. Skelton, M. Rance, Protein NMR Spectroscopy: Principles and Practice, second ed., Academic Press, San Diego, CA, 2007.
[16] R. Sakuma, T. Sakuma, S. Ohmine, R.H. Silverman, Y. Ikeda, Xenotropic murine leukemia virus-related virus is susceptible to AZT, Virology 397 (2010) 1–6.
[17] A. Urisman, R.J. Molinaro, N. Fischer, S.J. Plummer, G. Casey, E.A. Klein, K. Malathi, C. Magi-Galluzzi, R.R. Tubbs, D. Ganem, R.H. Silverman, J.L. DeRisi, Identification of a novel Gammaretrovirus in prostate tumors of patients homozygous for R462Q RNASEL variant, PLoS Pathog. 2 (2006) e25.
[18] H. Takahashi, A. Nozawa, M. Seki, K. Shinozaki, Y. Endo, T. Sawasaki, A simple and high-sensitivity method for analysis of ubiquitination and polyubiquitination based on wheat cell-free protein synthesis, BMC Plant Biol. 9 (2009).
[19] F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, A. Bax, NMRPipe: a multidimensional spectral processing system based on UNIX pipes, J. Biomol. NMR 6 (1995) 277–293.
[20] N. Kobayashi, J. Iwahara, S. Koshiba, T. Tomizawa, N. Tochio, P. Guntert, T. Kigawa, S. Yokoyama, KUJIRA, a package of integrated modules for systematic and interactive analysis of NMR data directed to high-throughput NMR structure studies, J. Biomol. NMR 39 (2007) 31–52.
[21] A. Furukawa, T. Nagata, A. Matsugami, Y. Habu, R. Sugiyama, F. Hayashi, N. Kobayashi, S. Yokoyama, H. Takaku, M. Katahira, Structure, interaction and real-time monitoring of the enzymatic reaction of wild-type APOBEC3G, EMBO J. 28 (2009) 440–451.
[22] T. Nagata, E. Niyada, N. Fujimoto, Y. Nagasaki, K. Noto, Y. Miyanoiri, J. Murata, K. Hiratsuka, M. Katahira, Solution structures of the trihelix DNA-binding domains of the wild-type and a phosphomimetic mutant of Arabidopsis GT-1: mechanism for an increase in DNA-binding affinity through phosphorylation, Proteins 78 (2010) 3033–3047.
[23] D.S. Wishart, B.D. Sykes, The ^{13}C chemical-shift index: a simple method for the identification of protein secondary structure using ^{13}C chemical-shift data, J. Biomol. NMR 4 (1994) 171–180.
[24] C. Zwahlen, P. Legault, S.J.F. Vincent, J. Greenblatt, R. Konrat, L.E. Kay, Methods for measurement of intermolecular NOEs by multinuclear NMR spectroscopy: application to a bacteriophage lambda N-peptide/boxB RNA complex, J. Am. Chem. Soc. 119 (1997) 6711–6721.
[25] E. Katoh, J.M. Louis, T. Yamazaki, A.M. Gronenborn, D.A. Torchia, R. Ishima, A solution NMR study of the binding kinetics and the internal dynamics of an HIV-1 protease–substrate complex, Protein Sci. 12 (2003) 1376–1385.
[26] V. Hornak, A. Okur, R.C. Rizzo, C. Simmerling, HIV-1 protease flaps spontaneously open and reclose in molecular dynamics simulations, Proc. Natl. Acad. Sci. USA 103 (2006) 915–920.
[27] A.L. Perryman, J.H. Lin, J.A. McCammon, Restrained molecular dynamics simulations of HIV-1 protease: the first step in validating a new target for drug design, Biopolymers 82 (2006) 272–284.
[28] A.L. Perryman, Q. Zhang, H.H. Soutter, R. Rosenfeld, D.E. McRee, A.J. Olson, J.E. Elder, C.D. Stout, Fragment-based screen against HIV protease, Chem. Biol. Drug Des. 75 (2010) 257–268.

Immunogenetics (2012) 64:669–678
DOI 10.1007/s00251-012-0628-y

ORIGINAL PAPER

Lineage-specific evolution of T-cell immunoglobulin and mucin domain 1 gene in the primates

Hitoshi Ohtani · Taeko K. Naruse · Yuki Iwasaki · Hirofumi Akari · Takafumi Ishida · Tetsuro Matano · Akinori Kimura

Received: 30 December 2011 / Accepted: 6 June 2012 / Published online: 19 June 2012

Abstract T-cell immunoglobulin domain and mucin domain containing protein 1 (TIM1), also known as a cellular receptor for hepatitis A virus (HAVCR1) or a molecule induced by ischemic injury in the kidney (KIM1), is involved in the regulation of immune responses. We investigated a natural selection history of *TIM1* by comparative sequencing analysis in 24 different primates. It was found that *TIM1* had become a pseudogene in multiple lineages of the New World monkey. We also investigated T cell lines originated from four different New World monkey species and confirmed that *TIM1* was not expressed at the mRNA level. On the other hand, there were ten amino acid sites in the Ig domain of TIM1 in the other primates, which were suggested to be under positive natural selection. In addition, mucin domain of TIM1 was highly polymorphic in the Old World monkeys, which might be under balanced selection. These data suggested that *TIM1* underwent a lineage-specific evolutionary pathway in the primates.

Keywords Natural selection · Molecular evolution · Pseudogene · TIM1 · Primate

Electronic supplementary material The online version of this article (doi:10.1007/s00251-012-0628-y) contains supplementary material, which is available to authorized users.

H. Ohtani · T. K. Naruse · A. Kimura (✉)
Department of Molecular Pathogenesis, Medical Research Institute, Tokyo Medical and Dental University,
1-5-45 Yushima, Bunkyo-ku,
Tokyo 113-8510, Japan
e-mail: akitis@mri.tmd.ac.jp

Y. Iwasaki · H. Akari
Center for Human Evolution Modeling Research, Primate Research Institute, Kyoto University,
Inuyama, Japan

T. Ishida
Unit of Human Biology and Genetics, Graduate School of Science, The University of Tokyo,
Tokyo, Japan

T. Matano
AIDS Research Center, National Institute of Infectious Diseases,
Tokyo, Japan

Introduction

Comparative genomics is a useful tool for understanding the gene function from the view point of evolution. It has recently been reported that genes involved in regulation of immune system may have undergone the control of positive selection (Gibbs et al. 2007; Kosiol et al. 2008). The accelerated evolution may be due to a direct consequence of complex selection pressure exerted by infectious reagents including microbes and viruses (Barreiro and Quintana-Murci 2010). The known cases include genes of defensin family, which play crucial roles in antibacterial activity (Hollox and Armour 2008), and genes of APOBEC family, which are known to function as specific inhibitors against the infection of human immunodeficiency virus-1 (HIV-1) (Sawyer et al. 2004).

We previously performed a comparative genome analysis of primates and reported that genes encoding the immunoglobulin superfamily (IgSF) were classified into 11 functional categories based on the Gene Ontology (GO) database. The IgSF genes in three functional categories, immune system process (GO:0002376), defense response (GO:0006952), and multi-organism process (GO:0051704), had more chance to be under the positive natural selection than the IgSF genes in the other categories (Ohtani et al. 2011). In our previous comparative genome analysis, we focused on the orthologous

IgSF genes that appeared to be functional in all of human, chimpanzee, orangutan, rhesus macaque, and common marmoset. In other words, we excluded several genes of which an ortholog was considered to be non-functional, i.e., deleted gene, grossly rearranged gene, or pseudogene, in any of the five primate species. Such a lineage-specific destruction of IgSF genes, especially those involved in the immune response, may be interesting in view of the natural selection occurred during the evolution of primates. One of the excluded genes in our previous analysis was a gene for T-cell Ig domain and mucin domain containing protein 1 (TIM1), which was suggested to be a pseudogene due to an insertion in the common marmoset, while it should be functional in the other primates. TIM1 tightly linked to immune system, playing an important role in generation and/or maintenance of the balance between T helper 1 (Th1) cells and T helper 2 (Th2) cells, and it is up-regulated in Th2 cells after activation and interacts with its ligand expressed on antigen-presenting cells (de Souza and Kane 2006). *TIM1* can be found in the non-primate mammals including mouse and rat. However, *TIM1* orthologs are not found in the non-mammalian vertebrates such as chiken and zebrafish, implying that it might be involved in the mammalian-specific function. In addition, it was reported that *TIM1* is highly polymorphic in humans, but quite less polymorphic in chimpanzees, especially around the mucin domain (Nakajima et al. 2005). These observations suggested a unique evolutionary feature of *TIM1* in the primates.

In human, *TIM1* located on chromosome 5 at band q33 contains two distinct domains (Ig domain and mucin domain) (Khademi et al. 2004). It is known that TIM1 is a cellular receptor for hepatitis A virus (HAVCR1) in human (Feigelstock et al. 1998). TIM1 is also known to be induced in the kidney by ischemic injury and is called as kidney injury molecule 1 (KIM1) (Ichimura et al. 1998). It has been reported that *TIM1* polymorphisms are associated with various immune-related diseases and infectious diseases, including asthma, allergic rhinitis, atopic dermatitis, multiple sclerosis, type 1 diabetes, rheumatoid arthritis, AIDS, and cerebral malaria (Khademi et al. 2004; Kuchroo et al. 2003; McIntire et al. 2004; Meyers et al. 2005b; Su et al. 2008; Wichukchinda et al. 2010). These data implied that variations in *TIM1* might have been more or less selected during the evolution of humans.

In the present study, we determined nucleotide sequences of exons or equivalent regions of *TIM1* from 24 different primate species, including eight hominoids, six Old World monkeys, nine New World monkeys, and one prosimian, to investigate an evolutionary history of *TIM1*.

Materials and methods

Subjects

DNA samples from 24 primate species including human (*Homo sapiens*), chimpanzee (*Pan troglodytes*), bonobo (*Pan paniscus*), western gorilla (*Gorilla gorilla*), Bornean orangutan (*Pongo pygmaeus*), western black-crested gibbon (*Nomascus concolor*), lar gibbon (*Hylobates lar*), siamang (*Symphalangus syndactylus*), rhesus macaque (*Macaca mulatta*), long-tailed macaque (*Macaca fascicularis*), Hamadryas baboon (*Papio hamadryas*), mantled Guereza colobus (*Colobus guereza*), dusky leaf monkey (*Trachypithecus obscurus*), silver leaf monkey (*Trachypithecus cristatus*), Geoffroy's spider monkey (*Ateles geoffroyi*), white-fronted spider monkey (*Ateles belzebuth*), tufted capuchin (*Cebus apella*), common squirrel monkey (*Saimiri sciureus*), white-lipped tamarin (*Saguinus labiatus*), golden-handed tamarin (*Saguinus midas*), cotton-top tamarin (*Saguinus oedipus*), golden lion tamarin (*Leonthopithecus rosalia),* common marmoset (*Callithrix jacchus*), and Sunda slow loris (*Nycticebus coucang*) were the subjects.

Polymerase chain reaction (PCR) and sequencing analysis

Sequence information for homologous regions to the coding regions of human *TIM1* was obtained from 24 primate species by direct sequencing of PCR products from the genomic DNA samples. Primers used for PCR and direct sequencing were designed by referring the human, chimpanzee, rhesus macaque, common marmoset gene sequences, and whole-genome shotgun sequences from prosimians deposited in the UCSC Genome Browser and NCBI BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) (Supplemental Table S1). PCR condition was composed of a denaturing step (94 °C for 2 min), 35 cycles of chain reaction (94 °C for 30 s, 56 °C for 30 s, and 72 °C for 1 min), and a final extention step (72 °C for 5 min). The PCR products were then purified and sequenced by the BigDye Terminator cycling system using an ABI3130x automated DNA sequencer (Applied Biosystems, Foster City, CA, USA). Editing and assembly of sequences were done by using SEQUENCHER (Gene Codes, Ann Arbor, MI, USA). When sequence variations (heterozygous sequences) in a specific species were detected, the sequences which were more conserved among 24 primate species were considered as ancestral sequences and used for statistical analyses. The *TIM1* sequences determined in this study were deposited in DNA Data Bank of Japan (DDBJ) (Supplemental Table S2).

Expression analysis of *TIM1*

Total RNA was extracted from a human T cell line, Jurkat, and four different T cell lines originated from the New World monkeys, HSF-10 (tufted capuchin), HSQ-115 (common squirrel monkey), HST-3 (white-lipped tamarin) (Akari et al., manuscript in preparation), and HSCj-109 (common marmoset) (Hohjoh et al. 2009), by using RNAiso (TaKaRa Bio Inc., Shiga, Japan). Extracted RNAs (500 ng) were subjected to reverse transcription (RT) by using RT reagent Kit (TaKaRa Bio Inc., Shiga, Japan). Aliquots of RT products were used for the expression analysis of *TIM1*. Primers for PCR were designed in the highly conserved regions of *TIM1* among human and the New World monkeys (Supplemental Table S3). PCR condition was the same as that described in the previous section.

Diversity of TIM1 mucin domain in the Old World monkeys

Nucleotide sequences for the mucin domain from eight samples of rhesus macaques were determined by sequencing of PCR products, which were cloned into pT7Blue Blunt vector and transformed Nova Blue Single Competent cells using the Perfectly Blunt cloning kit (Novagen Inc., Madison, WI). Colony PCR was used to identify positive clones, and at least 20 positive clones from each sample were subsequently sequenced as described previously. We also examined length variations of exon 4 from 16 rhesus macaques and 10 crab-eating macaques by direct sequencing of the PCR products.

Statistical analyses

We used both Bn-Bs program and PAML program as described previously (Ohtani et al. 2011). In brief, the Bn-Bs program was used to investigate the presence of branch-specific positive selection (the branch model). The Bn-Bs program estimates the values of non-synonymous substitution rate (dn) and synonymous substitution rate (ds) based on the modified Nei–Gojobori method, where a phylogenetic tree is given (Zhang et al. 1998). The value of ω, an abbreviation for the value of dn/ds, is a criterion of natural selective pressure acting on the gene, and the modified Nei–Gojobori method has been used for estimating non-synonymous/synonymous substitution rates (Nei and Gojobori 1986). Statistical significance of the difference between dn and ds was examined by *Z*-test (Chatterjee et al. 2009). An ordinary least-squares method was used to estimate branch lengths and variances for *Z*-test. The least-squares method gives estimates for evolutionary distances among the analyzed sequences (Rzhetsky and Nei 1993). The PAML program version 4.7 was used to investigate the presence of site-specific positive selection (the site model). The site model treats ω allowing the variance among codons (Yang 2005; Yang and Nielsen 2000), and the following null and alternative models were implemented in the site model: M0 (null), M1a (nearly neutral), M2a (positive selection) (Wong et al. 2004), M3 (discrete), M7 (beta), and M8 (beta and ω) (Yang and Nielsen 2000). The likelihood ratio tests (LRT) of three pairwise comparisons, i.e., comparisons of M1a vs. M2a, M1a vs. M3, and M7 vs. M8, determined whether particular models would provide a significantly better fit. When the LRT suggested positive selection, the Bayes empirical Bayes (BEB) method was used to detect the sites under the positive selection (Yang et al. 2005). To investigate a possible selection operated on exon 4 region of *TIM1* alleles in rhesus macaques, we calculated Tajima's *D* (Tajima 1989; Tamura et al. 2011).

Results

TIM1 is non-functional in several lineages of New World monkey

TIM1 is a member of *TIM* gene family composing of *TIM1*, *TIM3*, and *TIM4*, in the human genome (Khademi et al. 2004). In the previous comparative genome analysis, we searched for orthologous genes for human *TIM1*, *TIM3*, and *TIM4* in the genome of chimpanzee, orangutan, rhesus macaque, and common marmoset by using the UCSC/MULTIZ alignment program. It was found that there was an insertion of 205 bp in exon 2-equivalent region in the common marmoset gene, which would generate multiple frameshift/nonsense mutations in the coding sequence and/or destroy the splicing junction.

To confirm the presence of deleterious insertion in *TIM1* in the genome of common marmoset and possibly in other primate genomes, we determined nucleotide sequences for exons or equivalent regions of *TIM1* from 24 primate species including human, chimpanzee, orangutan, rhesus macaque, and common marmoset. For this purpose, we designed primers by referring the known *TIM1* sequences (Supplemental Table 1). The sequencing analysis of the genomic gene for *TIM1* revealed the deleterious insertions of 206–212 bp in several New World monkeys, i.e., golden lion tamarin (212 bp), cotton-top tamarin (206 bp), white-lipped tamarin (207 bp), and golden-handed tamarin (210 bp) (Supplemental Figure S1). It was speculated that the insertion had been occurred within a sequence stretch of 13 bp, AGCCTCATCCTAC, corresponding to codons 9–13, because these sequences were repeated and flanked the insertion in the genomes of common marmoset and cotton-top tamarin, and there were a few substitutions in this sequence stretch from the other New World monkeys (Supplemental Figure S1). The inserted sequences belong to the

LINE/L1 (L1PA7) repeat, which contain a poly A stretch at one end in a reverse orientation to the TIM1 coding sequences, and homologous sequences can be found as multiple copies in the marmoset genome. On the other hand, we found nucleotide substitutions in exon 3-equivalent regions, which resulted in termination mutations in three other New World monkeys not carrying the insertion, common squirrel monkey (S84X), tufted capuchin (V22X), and Geoffroy's spider monkey (C36X) (Fig. 1). Among the New World monkey species investigated in this study, only the white-fronted spider monkey appeared to carry a functional gene for TIM1.

To investigate whether *TIM1* was non-functional in the New World monkey lineages, we performed RT-PCR analysis of mRNA expression in T cell lines originated from human, tufted capuchin, squirrel monkey, white-lipped tamarin, and common marmoset. As illustrated in Fig. 2a, two pairs of primers were used in the RT-PCR analysis, where forward primers were designed in exon 3 and junction of exon 5–exon 6, while reverse primers were designed in exon 9 and junction of exon 5–exon 6 (Supplemental Table S3). The *TIM1* expression was confirmed in the human T cell line, but could not be detected in the T cell lines from the New World monkeys carrying either the insertions (common marmoset and white-lipped tamarin) or the nonsense mutations (common squirrel monkey and tufted capuchin) (Fig. 2). No expression of *TIM3* and *TIM4* was observed in the T cell lines from human and the New World monkeys (data not shown).

Positive selection sites of *TIM1* in the primates

In the other primate species than the New World monkey, *TIM1* appeared to be functional, and there were many substitutions. When we calculated the dn and ds values in each primate lineage, it was found that the dn values were higher

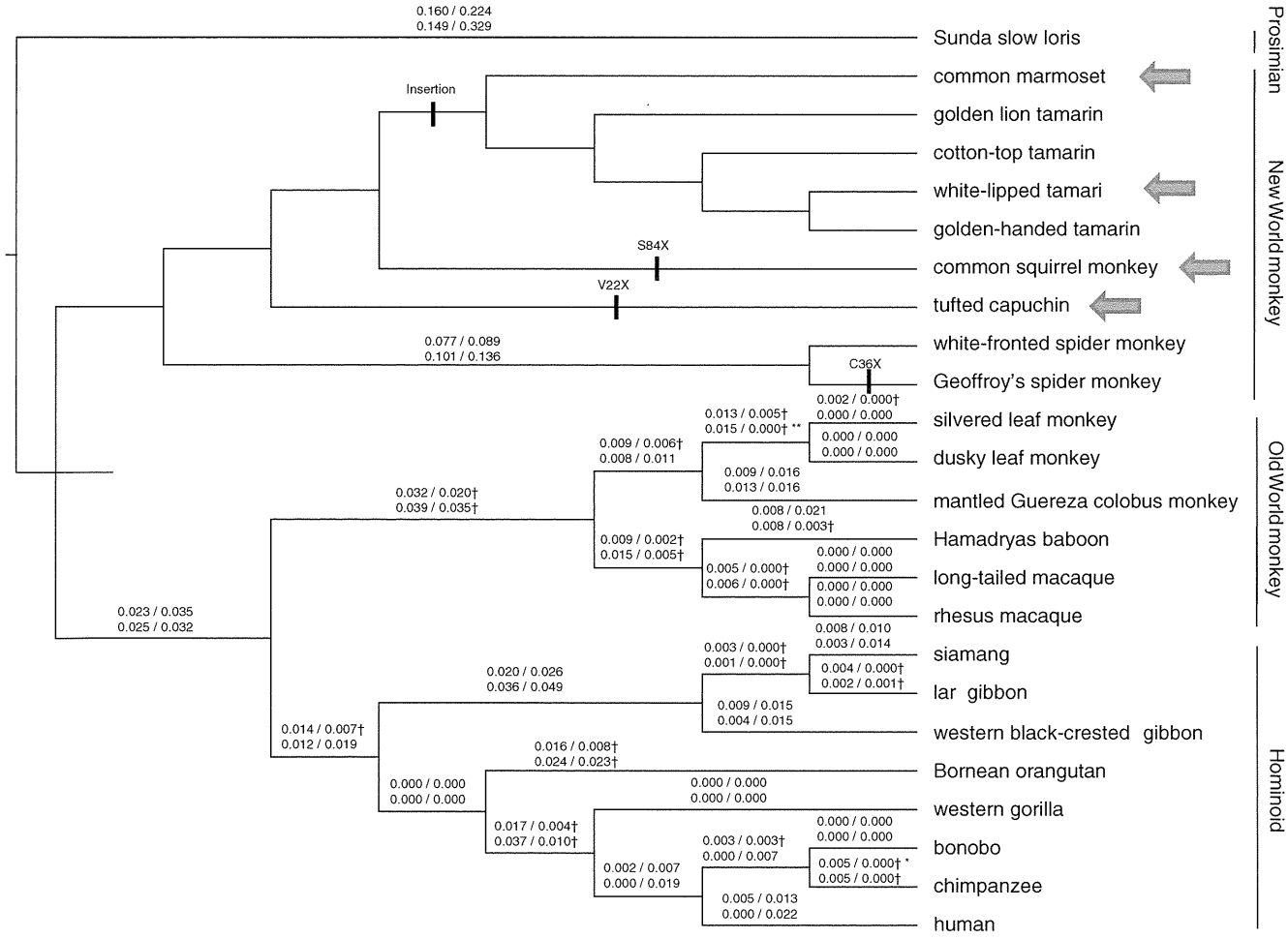


Fig. 1 Phylogenetic trees for TIM1 in the primate evolution. Values above branches indicate estimated values of dn and ds per lineage by using the Bn-Bs program. Upper values are for the entire coding region, while lower values are for the Ig domain. *Daggers* indicate that the dn value was higher than the ds value. *Asterisks* indicate that there is a significant difference between the dn and ds values ($^{**}p<0.01$; $^{*}p<0.05$; *Z*-test). *Vertical lines* indicate that *TIM1* had become pseudogene in the specific lineage. *Arrows* indicate the species for which the mRNA expression of *TIM1* in T cell line was investigated

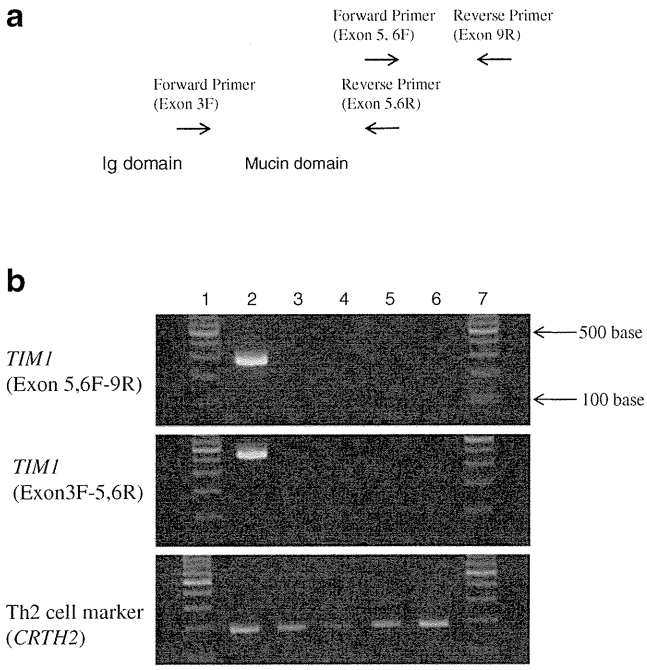


Fig. 2 Expression of *TIM1* in T cell lines. **a** TIM1 cDNA is schematically shown. *Arrows* indicate the regions of PCR primers which were designed in the highly conserved coding regions of *TIM1* among human and New World monkeys. **b** From *left* to *right*: *lane 1* 100-bp ladder size marker, *lane 2* Jurkat (human), *lane 3* HSF-10 (tufted capuchin), *lane 4* HSQ-115 (common squirrel monkey), *lane 5* HST-3 (white-lipped tamarin), *lane 6* HSCj-109 (common marmoset), and *lane 7* 100-bp ladder size marker. A marker gene (*CRTH2*) was used as a positive control of gene expression because it is known to be expressed in T cells especially in Th2-type cells

than the ds values in several lineages, especially in the Old World monkeys (Fig. 1). Sequence alignment of TIM1 at the amino acid (AA) level in the primates is shown in Fig. 3. To identify possible target sites for positive selection, we analyzed the TIM1 sequences by the BEB method. Sixteen TIM1 sequences from primate species, human, chimpanzee, bonobo, western gorilla, Bornean orangutan, western black-crested gibbon, lar gibbon, siamang, rhesus macaque, long-tailed macaque, Hamadryas baboon, mantled Guereza colobus, dusky leaf monkey, silver leaf monkey, white-fronted spider monkey, and Sunda slow loris, were used in the statistical test. It was revealed that 14 AA sites, at 23, 25, 30, 37, 51, 54, 58, 59, 72, 93, 102,120, 125, and 288 positions equivalent to the human TIM1, were the positively selected sites. These 14 AA sites were highly variable among the 16 primate sequences, e.g., AA site at 23 was lysine in hominoid; asparagine in rhesus macaque, long-tailed macaque, Hamadryas baboon, and mantled Guereza colobus; tyrosine in dusky leaf monkey and silver leaf monkey; serine in white-fronted spider monkey; and glutamine in Sunda slow loris. Most of the positive selection sites (10/14: 71.4 %) were found in the Ig domain (Fig. 3a, b).

Diversity of TIM1 in the Old World monkey

As shown in Fig. 3, a large number of deletion/insertion events were observed in the mucin domain (Fig. 3b, c). It was reported that human TIM1 exhibited a high degree of amino-acid variability in the mucin domain (Nakajima et al. 2005). Because the mucin domain of TIM1 encoded by exon 4 might be under the positive selection in the Old World monkey (Supplemental Table S4), we investigated the diversity of TIM1 in rhesus macaques by determining nucleotide sequences for the mucin domain from eight samples (16 haplotypes). As shown in Fig. 4, a high level of sequence diversity with multiple insertion/deletion of 18-bp sequences, A(T/C)GACAAC(G/A)(A/G)C(T/C)CT(T/G)CCA forming a part of AA stretch Thr-Thr-Thr-Thr-Leu-Pro (TTTTLP), was observed in the mucin domain of rhesus TIM1. We then investigated a possible selection by using the Bn-Bs program, but no statistically significant data were obtained, presumably because the compared sequences were not long enough to give a definite conclusion. However, when we calculated Tajima's *D* for these *TIM1* alleles, a value of 0.607 was obtained, which suggested a balanced selection of polymorphisms in the mucin domain of TIM1 in the Old World monkey.

On the other hand, because the diversity of mucin domain in the rhesus macaques could be detectable as a length diversity of exon 4, we examined length variations of *TIM1* exon 4 in additional samples of rhesus and long-tailed macaques. It was found that the length polymorphism was due to the repeat number polymorphisms or insertion/deletion polymorphisms of 18-bp unit and its components of 3- and 6-bp repeats (Supplemental Table S5).

Discussion

In this study, we investigated sequence diversity in the protein coding exons of *TIM1* from various species by direct sequencing method. Although there were a few sites with heterozygous sequences, we used "evolutionary conserved" sequences obtained from each sample in the statistical analyses so that the substitutions were underestimated in this study. Even though there was an underestimation, we demonstrated that the Ig domain of TIM1 has been under the positive selection during the course of primate evolution. Another interesting finding was that TIM1 has undergone pseudogene evolution in the New World monkey. It was suggested that the generation of pseudogene had occurred several times over the New World monkey lineages, by the insertion of deleterious sequences or base substitutions leading to a termination codon. Similar natural selection pattern was reported for type 1 vomeronasal receptors (*V1RL*), in which the pseudogene generation had independently

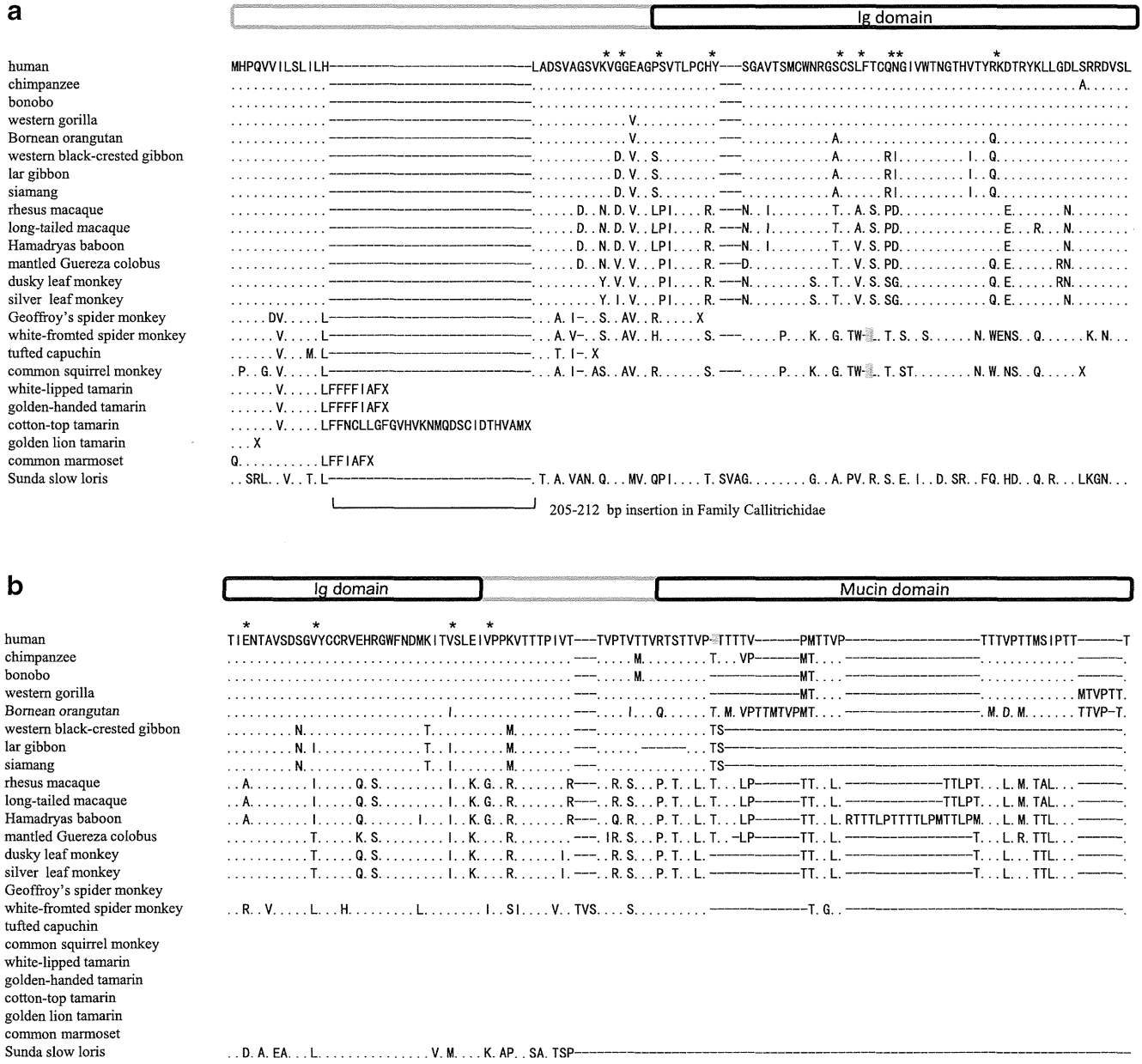


Fig. 3 Alignments of TIM1 amino acid sequences from 24 primate species. *Dots* indicate identities to the human reference sequence, while *dashes* indicate alignment gaps. *Asterisks* indicate AA sites identified as being under the significant positive selection ($p<0.05$). **a**, **b**, **c**, and **d** represent quarter parts of TIM1 from N-terminus to C-terminus. Ig and mucin domains are schematically indicated

occurred in the human and other primate lineages, implying that the function of *V1RL* might be highly lineage specific (Mundy and Cook 2003).

TIM1 plays an important role in generation and/or maintenance of the balance between Th1 and Th2 cells (Su et al. 2008). It is known that a natural ligand for TIM1 is T-cell Ig domain and mucin domain containing protein 4 (TIM4) (Meyers et al. 2005a), and the interaction of TIM1 with TIM4 plays a crucial role in sustaining the polarization status of Th2 cells (Khademi et al. 2004; Mariat et al. 2005). Because TIM1 expressed on the surface of Th2 cells regulated the immune response by modulating cytokine production in mammals, the Th1/Th2 balance might be skewed or affected by the lack of TIM1 in the most lineages of New World monkey. Previous study demonstrated that expression of Th1-type cytokines, IFN-γ and TNF-β, was considerably lower than that of Th2-type cytokines, IL-4 and IL-10, in a New World monkey, owl monkey (Pico de Coana et al. 2004). This observation may support the disturbance of Th1/Th2 balance in the owl monkey lacking TIM1, although we cannot exclude a possibility that some other molecules than TIM1 might regulate Th1/Th2 balance

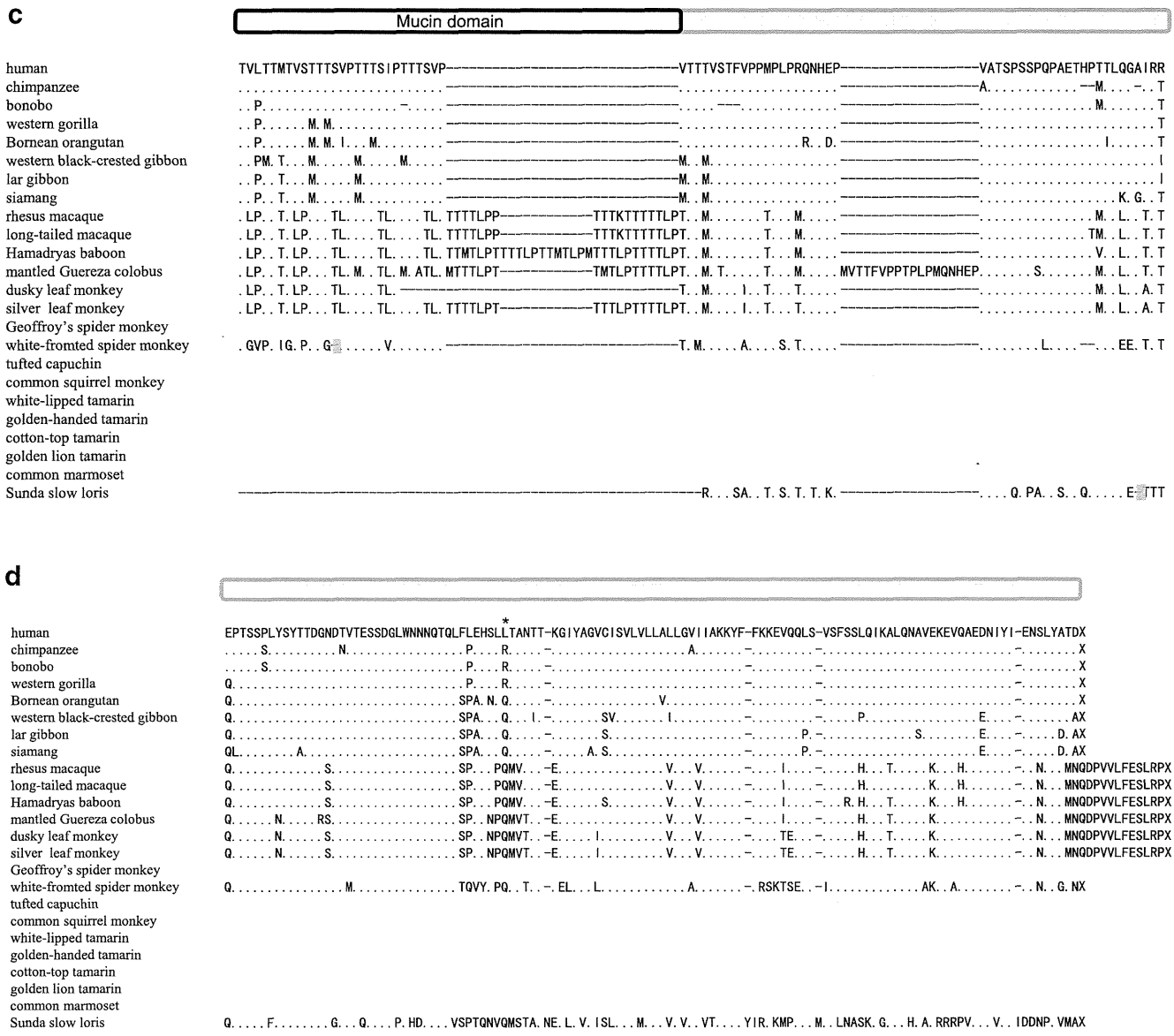


Fig. 3 (continued)

in the primates including the TIM1-deficient New World monkey. It is interesting to note that there is a *TIM1*-like gene on chromosome 13 in the marmoset genome, which might replace the function of *TIM1*. However, this TIM1-like gene lacks introns and contains several in-frame stop codons, when it is aligned with human *TIM1* (Supplemental Figure S2).

It is known that TIM1 is a cellular receptor for hepatitis A virus (HAV) in human (Feigelstock et al. 1998). The results in this study suggested that the New World monkey lineages

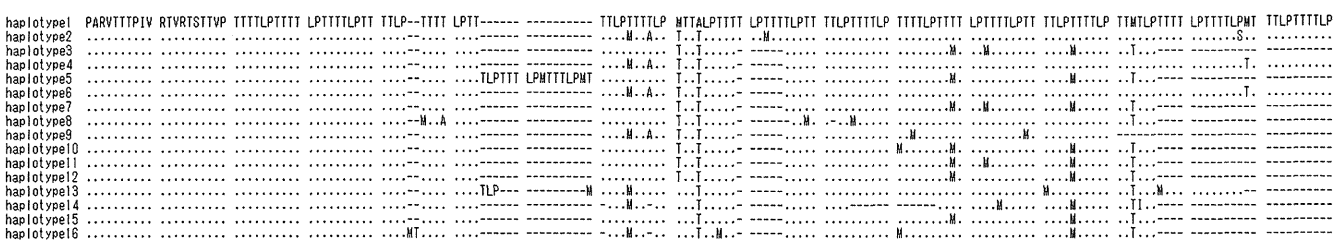


Fig. 4 Diversity of the mucin domain in rhesus macaques. *Dots* indicate identities to the haplotype 1 sequence, while *dashes* indicate alignment gaps

might lack a cellular receptor for HAV, TIM1. However, because the New World monkey is susceptible to HAV infection (Mathiesen et al. 1980), further studies are needed to clarify or find other cellular receptors for HAV in the New World monkey and a cause or reason of pseudogene generation, which had occurred in several lineages of the New World monkey. In addition, it has been reported that TIM1 polymorphisms are associated with resistance to autoimmune diseases including multiple sclerosis, which are associated with the imbalance between Th1 and Th2 cells (Khademi et al. 2004). Nevertheless, common marmosets are used as an animal model for multiple sclerosis, experimental autoimmune encephalomyelitis (Uccelli et al. 2003). Thus, susceptibilities of the New World monkey to the autoimmune diseases should be investigated in relation to the non-functional TIM1.

In this study, 14 AA sites of TIM1 were identified as positively selected sites in the evolutional course of primates other than the New World monkey. It was suggested that the Ig domain of TIM1 was a binding site for HAV (Feigelstock et al. 1998). In addition, structural conformation of the mucin domain is required for the efficient viral entry (McIntire et al. 2003). In this study, TIM1 was considered to be under the significant positive natural selection in the Ig domain, prompting us to investigate the three dimensional (3D) structures of the Ig domains using SWISS-MODEL, an Automated Comparative Protein Modeling Server (http://swissmodel.expasy.org/SWISS-MODEL.html) (Bordoli et al. 2009). As shown in Fig. 5, it was suggested that most of the target sites for the positive selection accumulated on the surface of Ig domain. These observations support a hypothesis that the evolution of TIM1 in the primates might be driven by exogenous pathogens. The positively selected sites in the Ig domain of TIM1 in the primates other than the New World monkey and a high level diversity in the mucin domain of TIM1 in the Old World monkey might be a direct consequence of a selection pressure exerted by HAV. However, because TIM1 polymorphisms are also associated with other infectious diseases including HIV/AIDS and cerebral malaria, further functional studies are required to clarify the mechanism of natural selection at specific sites of Ig and mucin domains of TIM1. For example, we are now investigating whether the TIM1 repeat polymorphism would influence the production level of neutralizing antibodies against challenging Simian Immunodeficiency Virus (SIV) in experimental models of SIV vaccination in rhesus macaques (Sugimoto et al. 2010; Ishii et al. 2012; Nomura et al. 2012).

Because TIM1 is known to interact with TIM4 and both Ig and mucin domains of TIM1 are involved in this interaction (Meyers et al. 2005a), one might speculate a co-evolution of TIM1 and TIM4. It should be noted here that human TIM family includes three members, TIM1, TIM3, and TIM4, while mouse TIM family includes eight members, TIM1 toTIM8. Although we searched for orthologs of mouse TIM2, TIM5, TIM6, TIM7, and TIM8 in the common marmoset genome by using Blat program (http://genome.ucsc.edu/cgi-bin/hgBlat), we could not detect any orthologous genes. Therefore, TIM family in the New World monkey consists of only two functional members, TIM3 and TIM4. Then, we investigated a possible evolutionary selection of TIM3 and TIM4. However, no significant positive selection appeared to operate on the evolution of TIM3 and TIM4 in the primates (Supplemental Table S6). A marginal and non-significant positive selection for TIM4 in chimpanzee was observed, but it may not correlate with co-evolution of TIM1 and TIM4, because the mucin domain of TIM1 is virtually non-polymorphic in chimpanzee (Nakajima et al. 2005). Nevertheless, the observations in this study suggest that the diversity of TIM family is widely ranged among mammalian species. It may be of interest to investigate whether the binding affinity of TIM1 and TIM4 would be affected by the TIM1 variations in future experiments. On the other hand, it may be noteworthy that TIM1, TIM3, and TIM4 can independently serve as receptors for phosphatidylserine to mediate uptake of apoptotic cells (Kobayashi et al. 2007; Freeman et al. 2010), implying that their cooperation would be dispensable in some functional aspects.

In conclusion, we investigated the molecular evolution of TIM1 in 24 primate species. TIM1 had become pseudogenes in most lineages of the New World monkey, while it

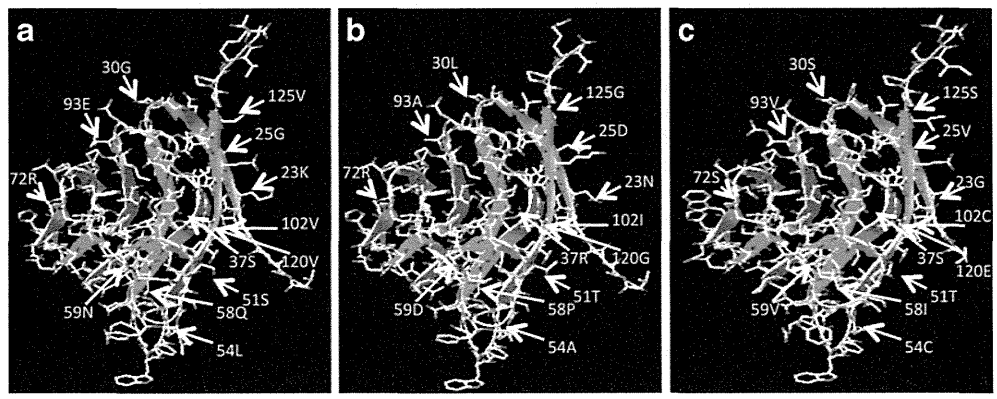


Fig. 5 Three-dimensional structures of TIM1 modeled by SWISS-MODEL. *Arrows* indicate AA sites identified as being under the positive selection by using the BEB method in the PAML program. **a** human TIM1, **b** rhesus macaque TIM1, **c** long-haired spider monkey TIM1

was under the positive selection in the other primates, especially in the Old World monkey. TIM1 might undergo a selection pressure exerted by infectious disease and autoimmune disease.

Acknowledgments We would like to thank Dr. Toshiaki Nakajima for his critical comments and contributions in the initial course of this study. This work was supported in part by research grants from the Ministry of Health, Labor and Welfare, Japan; by Grant-in-Aids for scientific research from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT), Japan; by grants for India–Japan Cooperative Science Program from Japan Society for the Promotion of Science (JSPS), Japan and Department of Science and Technology (DST), India; and Joint Usage/Research Programs of Research Institute for Microbial Diseases, Osaka University and Medical Research Institute Tokyo Medical and Dental University. This work was also supported by a program of support for women researchers from the Tokyo Medical and Dental University.

References

Barreiro LB, Quintana-Murci L (2010) From evolutionary genetics to human immunology: how selection shapes host defence genes. Nat Rev Genet 11:17–30

Bordoli L, Kiefer F, Arnold K, Benkert P, Battey J, Schwede T (2009) Protein structure homology modeling using SWISS-MODEL workspace. Nat Protoc 4:1–13

Chatterjee HJ, Ho SY, Barnes I, Groves C (2009) Estimating the phylogeny and divergence times of primates using a supermatrix approach. BMC Evol Biol 9:259

de Souza AJ, Kane LP (2006) Immune regulation by the TIM gene family. Immunol Res 36:147–155

Feigelstock D, Thompson P, Mattoo P, Zhang Y, Kaplan GG (1998) The human homolog of HAVcr-1 codes for a hepatitis A virus cellular receptor. J Virol 72:6621–6628

Freeman GJ, Casasnovas JM, Umetsu DT, DeKruyff RH (2010) TIM genes: a family of cell surface phosphatidylserine receptors that regulate innate and adaptive immunity. Immunol Rev 235:172–189

Gibbs RA et al (2007) Evolutionary and biomedical insights from the rhesus macaque genome. Science 316:222–234

Hohjoh H, Akari H, Fujiwara Y, Tamura Y, Hirai H, Wada K (2009) Molecular cloning and characterization of the common marmoset huntingtin gene. Gene 432:60–66

Hollox EJ, Armour JA (2008) Directional and balancing selection in human beta-defensins. BMC Evol Biol 8:113

Ichimura T, Bonventre JV, Bailly V, Wei H, Hession CA, Cate RL, Sanicola M (1998) Kidney injury molecule-1 (KIM-1), a putative epithelial cell adhesion molecule containing a novel immunoglobulin domain, is up-regulated in renal cells after injury. J Biol Chem 273:4135–4142

Ishii H, Kawada M, Tsukamoto T, Yamamoto H, Matsuoka S, Shiino T, Takeda A, Inoue M, Iida A, Hara H, Shu T, Hasegawa M, Naruse TK, Kimura A, Takiguchi M, Matano T (2012) Impact of vaccination on cytotoxic T lymphocyte immunodominance and cooperation against simian immunodeficiency virus replication in rhesus macaques. J Virol 86:738–745

Khademi M, Illes Z, Gielen AW, Marta M, Takazawa N, Baecher-Allan C, Brundin L, Hannerz J, Martin C, Harris RA, Hafler DA, Kuchroo VK, Olsson T, Piehl F, Wallstrom E (2004) T cell Ig- and mucin-domain-containing molecule-3 (TIM-3) and TIM-1 molecules are differentially expressed on human Th1 and Th2 cells and in cerebrospinal fluid-derived mononuclear cells in multiple sclerosis. J Immunol 172:7169–7176

Kobayashi N, Karisola P, Peña-Cruz V, Dorfman DM, Jinushi M, Umetsu SE, Butte MJ, Nagumo H, Chernova I, Zhu B, Sharpe AH, Ito S, Dranoff G, Kaplan GG, Casasnovas JM, Umetsu DT, Dekruyff RH, Freeman GJ (2007) TIM-1 and TIM-4 glycoproteins bind phosphatidylserine and mediate uptake of apoptotic cells. Immunity 27:927–940

Kosiol C, Vinar T, da Fonseca RR, Hubisz MJ, Bustamante CD, Nielsen R, Siepel A (2008) Patterns of positive selection in six mammalian genomes. PLoS Genet 4:e1000144

Kuchroo VK, Umetsu DT, DeKruyff RH, Freeman GJ (2003) The TIM gene family: emerging roles in immunity and disease. Nat Rev Immunol 3:454–462

Mariat C, Sanchez-Fueyo A, Alexopoulos SP, Kenny J, Strom TB, Zheng XX (2005) Regulation of T cell dependent immune responses by TIM family members. Philos Trans R Soc Lond B Biol Sci 360:1681–1685

Mathiesen LR, Moller AM, Purcell RH, London WT, Feinstone SM (1980) Hepatitis A virus in the liver and intestine of marmosets after oral inoculation. Infect Immun 28:45–48

McIntire JJ, Umetsu SE, Macaubas C, Hoyte EG, Cinnioglu C, Cavalli-Sforza LL, Barsh GS, Hallmayer JF, Underhill PA, Risch NJ, Freeman GJ, DeKruyff RH, Umetsu DT (2003) Immunology: hepatitis A virus link to atopic disease. Nature 425:576

McIntire JJ, Umetsu DT, DeKruyff RH (2004) TIM-1, a novel allergy and asthma susceptibility gene. Springer Semin Immunopathol 25:335–348

Meyers JH, Chakravarti S, Schlesinger D, Illes Z, Waldner H, Umetsu SE, Kenny J, Zheng XX, Umetsu DT, DeKruyff RH, Strom TB, Kuchroo VK (2005a) TIM-4 is the ligand for TIM-1, and the TIM-1-TIM-4 interaction regulates T cell proliferation. Nat Immunol 6:455–464

Meyers JH, Sabatos CA, Chakravarti S, Kuchroo VK (2005b) The TIM gene family regulates autoimmune and allergic diseases. Trends Mol Med 11:362–369

Mundy NI, Cook S (2003) Positive selection during the diversification of class I vomeronasal receptor-like (V1RL) genes, putative pheromone receptor genes, in human and primate evolution. Mol Biol Evol 20:1805–1810

Nakajima T, Wooding S, Satta Y, Jinnai N, Goto S, Hayasaka I, Saitou N, Guan-Jun J, Tokunaga K, Jorde LB, Emi M, Inoue I (2005) Evidence for natural selection in the HAVCR1 gene: high degree of amino-acid variability in the mucin domain of human HAVCR1 protein. Genes Immun 6:398–406

Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol 3:418–426

Nomura T, Terahara K, Yamamoto H, Shiino T, Takahashi N, Nakane T, Iwamoto N, Ishii H, Tsukamoto T, Kawada M, Matsuoka S, Takeda A, Terahara K, Tsunetsugu-Yokota Y, Iwata-Yoshikawa N, Hasegawa H, Sata T, Naruse TK, Kimura A, Matano T (2012) Association of major histocompatibility complex class I haplotypes with disease progression after simian immunodeficiency virus challenge in Burmese rhesus macaques. J Virol 86:6481–6490

Ohtani H, Nakajima T, Akari H, Ishida T, Kimura A (2011) Molecular evolution of immunoglobulin superfamily genes in primates. Immunogenetics 63:417–418

Pico de Coana Y, Barrero C, Cajiao I, Mosquera C, Patarroyo ME, Patarroyo MA (2004) Quantifying Aotus monkey cytokines by real-time quantitative RT-PCR. Cytokine 27:129–133

Rzhetsky A, Nei M (1993) Theoretical foundation of the minimum-evolution method of phylogenetic inference. Mol Biol Evol 10:1073–1095

Sawyer SL, Emerman M, Malik HS (2004) Ancient adaptive evolution of the primate antiviral DNA-editing enzyme APOBEC3G. PLoS Biol 2:E275