the Cα atom of Arg144 in the γδ resolvase (Figure S5 of the Supporting Information) and that of Tyr at the N-terminus of the zinc finger domain is ~13.2 Å (Figure 6B,C). In polypeptides in the extended conformation, the distance between Cα atoms of sequential amino acids is 3.8 Å. Thus, a linker consisting of three amino acids (clone 4 or clone 14) should allow the protein to bind to both DNA regions, although these ZFRs had very low recombination efficiencies. In the complex with DNA, the amino groups at positions 145 and 146 of the main chain in γδ resolvase interact with the phosphate backbone of DNA and amino acids of these positions are involved in the folding of the catalytic domain (Figure S5 of the Supporting Information). In the case of clone 4, the Lys-Pro residues at the C-terminus of linker residues are involved in the folding of the zinc finger domain. Thus, these amino acids are considered to be members of both domains, not of the linker sequences. With this reasoning, the six and nine amino acids in the linkers for clones 4 and 5, respectively, are shorter than the theoretically optimal length. Moreover, in the sequences of the six- and nine-amino acid linkers, the amino acid at position 146 is Pro, which could disrupt the interaction with DNA phosphate, thus lowering the recombination efficiency. Consistent with these estimations, the Gly-Ser linker with six or nine amino acids (clones 15 and 16, respectively) showed the best recombination ratio. This evidence indicates that the residues at the C-terminus of the catalytic domain and the N-terminus of the zinc finger domain are involved in domain folding because Lys-Pro residues at the N-terminus of the zinc finger domains are not included in these clones. Variants around this optimal linker length, especially those with 12 and 15 amino acids, had similar recombination efficiencies. These results show that the flexibility of the linker is not necessary when the linker length is optimal. In mammalian cells, the variant with a linker of six amino acids (clone 5) showed the best recombination and the zero-amino acid linker (clone 3) showed better recombination than the variants with longer linkers of more than 12 amino acids. The reason for this effect is unclear, but it could be due to differences in the structures of target sites on the plasmid DNA compared to the genomic DNA. Additionally, the distances between the binding sites in these systems are different. In the genomic target, the binding sites are separated by sequences of more than 2500 bp.

In this study, a newly developed recombination system allowed measurement of recombination efficiencies of ZFRs in *E. coli* and in mammalian cells. In mammalian cells, recombination with genomic targets was evaluated within 48 h of the transient expression of recombinases. Artificial enzymes such as ZFN and ZFR have been studied mainly by using viral vector systems to deliver their genes into mammalian genomes. In a report describing utilization of the retrovirus vectors for gene delivery, the recombination efficiency was as high as ~18%.[8] In our study, we also observed up to 18% recombination in cells. This system could be utilized in future studies to evaluate function of ZFRs on specific targets.

■ ASSOCIATED CONTENT

Supporting Information

Details of subcloning, experimental results of plasmid digestion and sequencing, results of FACS analyses, and a description of key interactions in γδ resolvase. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: nomura.mr@tmd.ac.jp or tamamura.mr@tmd.ac.jp. Phone: +81-3-5280-8036. Fax: +81-3-5280-8039.

Funding

This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology, Japan (20790060), Health and Labor Sciences Research Grants from the Japanese Ministry of Health, Labor, and Welfare, and a grant from the Mochida Memorial Foundation for Medical and Pharmaceutical Research to W.N.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Beerli, R. R., Dreier, B., and Barbas, C. F. III (2000) Positive and negative regulation of endogenous genes by designed transcription factors. *Proc. Natl. Acad. Sci. U.S.A. 97*, 1495–1500.

(2) Pabo, C. O., Peisach, E., and Grant, R. A. (2001) Design and selection of novel Cys2His2 zinc-finger proteins. *Annu. Rev. Biochem. 70*, 313–340.

(3) Beerli, R. R., and Barbas, C. F. III (2002) Engineering polydactyl zinc-finger transcription factors. *Nat. Biotechnol. 20*, 135–141.

(4) Jamieson, A. C., Miller, J. C., and Pabo, C. O. (2003) Drug discovery with engineered zinc-finger proteins. *Nat. Rev. Drug Discovery 2*, 361–368.

(5) Blancafort, P., Segal, D. J., and Barbas, C. F. III (2004) Designing transcription factor architectures for drug discovery. *Mol. Pharmacol. 66*, 1361–1371.

(6) Carroll, D. (2008) Progress and prospects: Zinc-finger nucleases as gene therapy agents. *Gene Ther. 15*, 1463–1468.

(7) Akopian, A., He, J., Boocock, M. R., and Stark, W. M. (2003) Chimeric recombinases with designed DNA sequence recognition. *Proc. Natl. Acad. Sci. U.S.A. 100*, 8688–8691.

(8) Gordley, R. M., Smith, J. D., Gräslund, T., and Barbas, C. F. III (2007) Evolution of programmable zinc-finger-recombinases with activity in human cells. *J. Mol. Biol. 367*, 802–813.

(9) Gordley, R. M., Gersbach, C. A., and Barbas, C. F. III (2009) Synthesis of programmable integrases. *Proc. Natl. Acad. Sci. U.S.A. 106*, 5053–5058.

(10) Gersbach, C. A., Gaj, T., Gordley, R. M., and Barbas, C. F. III (2010) Directed evolution of recombinase specificity by split gene reassembly. *Nucleic Acids Res. 38*, 4198–4206.

(11) Gaj, T., Mercer, A. C., Gersbach, C. A., Gordley, R. M., and Barbas, C. F. III (2011) Structure-Guided Reprogramming of Serine Recombinase DNA Sequence Specificity. *Proc. Natl. Acad. Sci. U.S.A. 108*, 498–503.

(12) Gersbach, C. A., Gaj, T., Gordley, R. M., Mercer, A. C., and Barbas, C. F. III (2011) Targeted plasmid integration into the human genome by an engineered zinc-finger recombines. *Nucleic Acids Res. 39*, 7868–7878.

(13) Xu, G.-L., and Bestor, T. H. (1997) Cytosine methylation targeted to predetermined sequences. *Nat. Genet. 17*, 376–378.

(14) McNamara, A. R., Hurd, P. J., Smith, A. E., and Ford, K. G. (2002) Characterisation of site-biased DNA methyltransferases: Specificity, affinity and subsite relationships. *Nucleic Acids Res. 30*, 3818–3130.

(15) Carvin, C. D., Parr, R. D., and Kladde, M. P. (2003) Site-selective in vivo targeting of cytosine-5 DNA methylation by zinc-finger proteins. *Nucleic Acids Res. 31*, 6493–6501.

(16) Minczuk, M., Papworth, M. A., Kolasinska, P., Murphy, M. P., and Klug, A. (2006) Sequence-specific modification of mitochondrial DNA using a chimeric zinc-finger methylase. *Proc. Natl. Acad. Sci. U.S.A. 103*, 19689–19694.

(17) Li, F., Papworth, M., Minczuk, M., Rohde, C., Zhang, Y., Ragozin, S., and Jeltsch, A. (2007) Chimeric DNA methyltransferases

target DNA methylation to specific DNA sequences and repress expression of target genes. *Nucleic Acids Res. 35*, 100–112.
(18) Smith, A. E., and Ford, K. G. (2007) Specific targeting of cytosine methylation to DNA sequences in vivo. *Nucleic Acids Res. 35*, 740–754.
(19) Smith, A. E., Hurd, P. J., Bannister, A. J., Kouzarides, T., and Ford, K. G. (2008) Heritable Gene Repression through the Action of a Directed DNA Methyltransferase at a Chromosomal Locus. *J. Biol. Chem. 283*, 9878–9885.
(20) Nomura, W., and Barbas, C. F. III (2007) In vivo site-specific DNA methylation with a designed sequence-enabled DNA methylase. *J. Am. Chem. Soc. 129*, 8676–8677.
(21) Grindley, N. D., Whiteson, K. L., and Rice, P. A. (2006) Mechanisms of site-specific recombination. *Annu. Rev. Biochem. 75*, 567–605.
(22) Yang, W., and Steitz, T. A. (1995) Crystal structure of the site-specific recombinase gamma delta resolvase complexed with a 34 bp cleavage site. *Cell 82*, 193–207.
(23) Arnold, P. H., Blake, D. G., Grindley, N. D., Boocock, M. R., and Stark, W. M. (1999) Mutants of Tn3 resolvase which do not require accessory binding sites for recombination activity. *EMBO J. 18*, 1407–1414.
(24) Li, W., Kamtekar, S., Xiong, Y., Sarkis, G. J., Grindley, N. D., and Steitz, T. A. (2005) Structure of a synaptic $\gamma\delta$ resolvase tetramer covalently linked to two cleaved DNAs. *Science 309*, 1210–1215.
(25) Olorunniji, F. J., He, J., Wenwieser, S. V., Boocock, M. R., and Stark, W. M. (2008) Synapsis and catalysis by activated Tn3 resolvase mutants. *Nucleic Acids Res. 36*, 7181–7191.
(26) Gonzalez, B., Schwimmer, L. J., Fuller, R. P., Ye, Y., Asawapornmongkol, L., and Barbas, C. F. III (2010) Modular system for the construction of zinc-finger libraries and proteins. *Nat. Protoc. 5*, 791–810.
(27) Mandell, J. G., and Barbas, C. F. III (2006) Zinc Finger Tools: Custom DNA-binding domains for transcription factors and nucleases. *Nucleic Acids Res. 34*, W516–W523.
(28) Kim, C. A., and Berg, J. M. (1996) A 2.2 Angstroms resolution crystal structure of a designed zinc finger protein bound to DNA. *Nat. Struct. Mol. Biol. 3*, 940–945.
(29) Segal, D. J., Dreier, B., Beerli, R. R., and Barbas, C. F. III (1999) Toward controlling gene expression at will: Selection and design of zinc finger domains recognizing each of the 5′-GNN-3′ DNA target sequences. *Proc. Natl. Acad. Sci. U.S.A. 96*, 2758–2763.
(30) Dreier, B., Segal, D. J., and Barbas, C. F. III (2000) Insights into the molecular recognition of the 5′-GNN-3′ family of DNA sequences by zinc-finger domains. *J. Mol. Biol. 303*, 489–502.
(31) Dreier, B., Beerli, R. R., Segal, D. J., Flippin, J. D., and Barbas, C. F. III (2001) Development of zinc finger domains for recognition of the 5′-ANN-3′ family of DNA sequences and their use in the construction of artificial transcription factors. *J. Biol. Chem. 276*, 29466–29478.
(32) Dreier, B., Fuller, R. P., Segal, D. J., Lund, C., Blancafort, P., Huber, A., Koksch, B., and Barbas, C. F. III (2005) Development of zinc finger domains for recognition of the 5′-CNN-3′ family DNA sequences and their use in the construction of artificial transcription factors. *J. Biol. Chem. 280*, 35588–35597.
(33) Kamiuchi, T., Abe, E., Imanishi, M., Kaji, T., Nagaoka, M., and Sugiura, Y. (1998) Artificial nine zinc-finger peptide with 30 base pair binding sites. *Biochemistry 37*, 13827–13834.
(34) Guo, J., Gaj, T., and Barbas, C. F. III (2010) Directed evolution of an enhanced and highly efficient FokI cleavage domain for zinc finger nuclease. *J. Mol. Biol. 400*, 96–107.

JVI
Journals.ASM.org

Impact of Vaccination on Cytotoxic T Lymphocyte Immunodominance and Cooperation against Simian Immunodeficiency Virus Replication in Rhesus Macaques

Hiroshi Ishii,[a,b] Miki Kawada,[b] Tetsuo Tsukamoto,[b] Hiroyuki Yamamoto,[a] Saori Matsuoka,[a] Teiichiro Shiino,[a] Akiko Takeda,[a] Makoto Inoue,[c] Akihiro Iida,[c] Hiroto Hara,[c] Tsugumine Shu,[c] Mamoru Hasegawa,[c] Taeko K. Naruse,[d] Akinori Kimura,[d] Masafumi Takiguchi,[e] and Tetsuro Matano[a,b]

AIDS Research Center, National Institute of Infectious Diseases, Tokyo, Japan[a]; Institute of Medical Science, University of Tokyo, Tokyo, Japan[b]; DNAVEC Corporation, Tsukuba, Japan[c]; Department of Molecular Pathogenesis, Medical Research Institute, Tokyo Medical and Dental University, Tokyo, Japan[d]; and Center for AIDS Research, Kumamoto University, Kumamoto, Japan[e]

Cytotoxic T lymphocyte (CTL) responses play a central role in viral suppression in human immunodeficiency virus (HIV) infections. Prophylactic vaccination resulting in effective CTL responses after viral exposure would contribute to HIV control. It is important to know how CTL memory induction by vaccination affects postexposure CTL responses. We previously showed vaccine-based control of a simian immunodeficiency virus (SIV) challenge in a group of Burmese rhesus macaques sharing a major histocompatibility complex class I haplotype. $Gag_{206-216}$ and $Gag_{241-249}$ epitope-specific CTL responses were responsible for this control. In the present study, we show the impact of individual epitope-specific CTL induction by prophylactic vaccination on postexposure CTL responses. In the acute phase after SIV challenge, dominant $Gag_{206-216}$-specific CTL responses with delayed, naive-derived $Gag_{241-249}$-specific CTL induction were observed in $Gag_{206-216}$ epitope-vaccinated animals with prophylactic induction of single $Gag_{206-216}$ epitope-specific CTL memory, and vice versa in $Gag_{241-249}$ epitope-vaccinated animals with single $Gag_{241-249}$ epitope-specific CTL induction. Animals with $Gag_{206-216}$-specific CTL induction by vaccination selected for a $Gag_{206-216}$-specific CTL escape mutation by week 5 and showed significantly less decline of plasma viral loads from week 3 to week 5 than in $Gag_{241-249}$ epitope-vaccinated animals without escape mutations. Our results present evidence indicating significant influence of prophylactic vaccination on postexposure CTL immunodominance and cooperation of vaccine antigen-specific and non-vaccine antigen-specific CTL responses, which affects virus control. These findings provide great insights into antigen design for CTL-inducing AIDS vaccines.

Human immunodeficiency virus (HIV) infection induces chronic, persistent viral replication leading to AIDS onset in humans. Virus-specific cytotoxic T lymphocyte (CTL) responses play a central role in the resolution of acute peak viremia (3, 4, 13, 22, 28) but mostly fail to contain viral replication in the natural course of HIV infection. Vaccination resulting in more effective CTL responses after viral exposure than in natural HIV infections would contribute to HIV control (30, 33). CTL memory induction by prophylactic vaccination may lead to efficient secondary CTL responses, but naive-derived primary CTL responses specific for viral nonvaccine antigens can also be induced after viral exposure. It is important to know how CTL memory induction by vaccination affects these postexposure CTL responses.

Cumulative studies on HIV-infected individuals have shown association of HLA genotypes with rapid or delayed AIDS progression (5, 14, 31, 34). For instance, most of the HIV-infected individuals possessing *HLA-B*57* have been indicated to show a better prognosis with lower viral loads, implicating HLA-B*57-restricted epitope-specific CTL responses in this viral control (1, 8, 23, 24). Indian rhesus macaques possessing certain major histocompatibility complex class I (MHC-I) alleles, such as *Mamu-A*01*, *Mamu-B*08*, and *Mamu-B*17*, tend to show simian immunodeficiency virus (SIV) control (19, 25, 36). This implies possible HIV control by induction of particular effective CTL responses (2, 7, 12, 16, 27).

Recent trials of prophylactic T-cell-based vaccines in macaque AIDS models have indicated the possibility of reduction in post-challenge viral loads (6, 15, 17, 21, 35). We previously developed a prophylactic AIDS vaccine consisting of a DNA prime and a boost with a Sendai virus (SeV) vector expressing SIVmac239 Gag (SeV-Gag) (20). Our trial showed vaccine-based control of an SIVmac239 challenge in a group of Burmese rhesus macaques sharing the MHC-I haplotype *90-120-Ia* (21). Animals possessing *90-120-Ia* dominantly elicited Mamu-A1*043:01 (GenBank accession number AB444869)-restricted $Gag_{206-216}$ (IINEEAADWDL) epitope-specific and Mamu-A1*065:01 (AB444921)-restricted $Gag_{241-249}$ (SSVDEQIQW) epitope-specific CTL responses after SIV challenge and selected for viral *gag* mutations, GagL216S (leading to a leucine [L]-to-serine [S] substitution at amino acid [aa] 216 in Gag) and GagD244E (aspartic acid [D]-to-glutamic acid [E] at aa 244), resulting in escape from CTL recognition with viral fitness costs in the chronic phase (9, 26). Vaccinees possessing *90-120-Ia* failed to control a challenge with a mutant SIV carrying these two CTL escape mutations, indicating that $Gag_{206-216}$-specific and $Gag_{241-249}$-specific CTL responses play a crucial role in the vaccine-based control of wild-type SIVmac239 replication

Received 5 September 2011 Accepted 31 October 2011
Published ahead of print 9 November 2011
Address correspondence to Tetsuro Matano, tmatano@nih.go.jp.

doi:10.1128/JVI.06226-11

TABLE 1 Animals analyzed in this study

Group	No. of animals	Vaccination[a]	SIV-specific CTL response postboost
I	6	None	None
II	5	Gag (pCMV-SHIVdEN DNA prime, SeV-Gag boost)	Gag-specific CTL
III	6	$Gag_{241-249}$-specific (p$Gag_{236-250}$-EGFP-N1 DNA prime, SeV-$Gag_{236-250}$-EGFP boost)	$Gag_{241-249}$-specific CTL
IV	5	$Gag_{206-216}$-specific (p$Gag_{202-216}$-EGFP-N1 DNA prime, SeV-$Gag_{202-216}$-EGFP boost)	$Gag_{206-216}$-specific CTL

[a] All animals were challenged with SIVmac239.

(10). Furthermore, in an SIVmac239 challenge experiment with *90-120-Ia*-positive rhesus macaques that received a prophylactic vaccine expressing the $Gag_{241-249}$ epitope fused with enhanced green fluorescent protein (EGFP), this single-epitope vaccination resulted in control of SIVmac239 replication with dominant induction of $Gag_{241-249}$-specific CTL responses in the acute phase postchallenge (32).

Thus, it is hypothesized that induction of single $Gag_{206-216}$ or $Gag_{241-249}$ epitope-specific CTL responses by vaccination may result in different patterns of CTL immunodominance and viral replication after SIV challenge. In the present study, we analyzed the impact of prophylactic vaccination inducing single $Gag_{206-216}$ epitope-specific CTL responses on SIV control in *90-120-Ia*-positive macaques and compared the results with those of vaccination inducing single $Gag_{241-249}$ epitope-specific CTL responses. This analysis revealed differences in CTL responses and patterns of viral control after SIV challenge between these vaccinated groups, indicating significant effects of prophylactic vaccination on postexposure CTL immunodominance and cooperation of vaccine antigen-specific and non-vaccine antigen-specific CTL responses.

MATERIALS AND METHODS

Animal experiments. Animal experiments were conducted through the Cooperative Research Program at Tsukuba Primate Research Center, National Institute of Biomedical Innovation, with the help of the Corporation for Production and Research of Laboratory Primates. Blood collection, vaccination, and virus challenge were performed under ketamine anesthesia. All animals were maintained in accordance with the Guideline for Laboratory Animals of the National Institute of Infectious Diseases.

Five Burmese rhesus macaques (*Macaca mulatta*) possessing the MHC-I haplotype *90-120-Ia* (26) (group IV) received a DNA-prime/SeV-boost vaccine eliciting $Gag_{206-216}$-specific CTL responses followed by an SIVmac239 challenge and were compared with three groups (I, II, and III) of *90-120-Ia*-positive animals reported previously (10, 32) (Table 1). Group I animals ($n = 6$) received no vaccination, while group II animals ($n = 5$) received a DNA-prime/SeV-boost vaccine eliciting Gag-specific CTL responses. The DNA, CMV-SHIVdEN, used for the vaccination was constructed from a simian/human immunodeficiency virus ($SHIV_{MD14YE}$) molecular clone DNA with *env* and *nef* deleted (29) and has the genes encoding SIVmac239 Gag, Pol, Vif, and Vpx; SIVmac239–HIV-1 chimeric Vpr; and HIV-1 Tat and Rev (21). In group II animals, CTL responses were undetectable after DNA prime but Gag-specific CTL responses became detectable after SeV-Gag boost. Group III animals ($n = 6$) received a DNA-prime/SeV-boost vaccine eliciting $Gag_{241-249}$-specific CTL responses. A p$Gag_{236-250}$-EGFP-N1 DNA and an SeV-$Gag_{236-250}$-EGFP vector, both expressing an SIVmac239 $Gag_{236-250}$ (IAGTTSSVDEQ IQWM)-EGFP fusion protein, were used for the group III vaccination. After the SeV-$Gag_{236-250}$-EGFP boost, group III animals induced $Gag_{241-249}$-specific CTL responses; the animals showed no $Gag_{236-250}$-specific $CD4^+$ T-cell responses but elicited SeV/EGFP-specific $CD4^+$ T-cell responses (32). For the group IV vaccination, A p$Gag_{202-216}$-EGFP-N1 DNA and an SeV-$Gag_{202-216}$-EGFP vector, both expressing an SIVmac239 $Gag_{202-216}$ (IIRDIINEEAADWDL)-EGFP fusion protein, were used (Fig. 1). Approximately 3 months after the boost, all animals were challenged intravenously with 1,000 50% tissue culture infective doses of SIVmac239 (11). In our previous study (32), the unvaccinated and the control-vaccinated

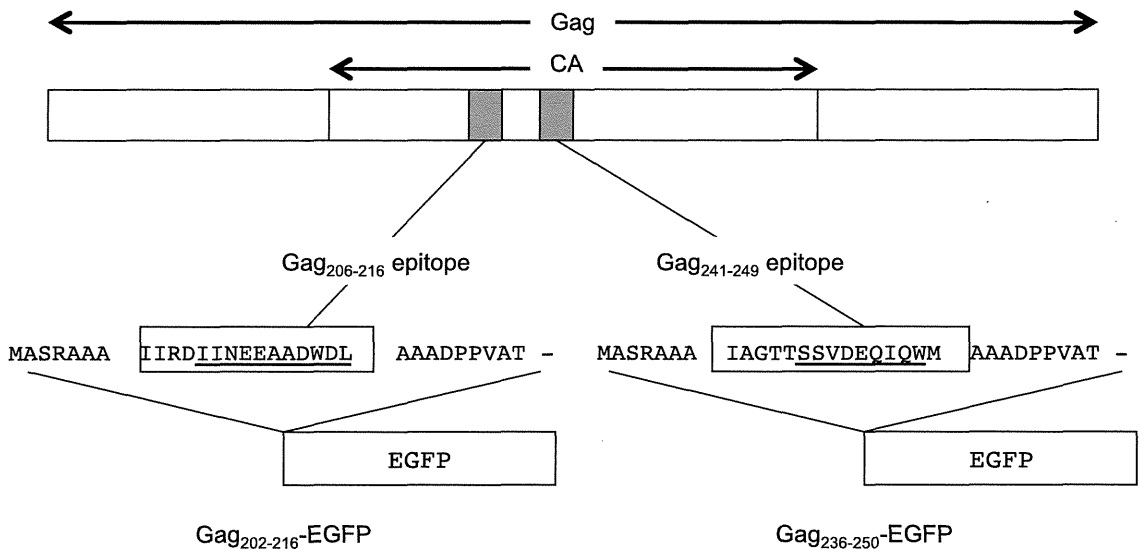

FIG 1 Schema of the cDNA constructs encoding $Gag_{202-216}$-EGFP and $Gag_{236-250}$-EGFP fusion proteins. A DNA fragment that encodes a 31-mer peptide (boxes) including the $Gag_{202-216}$ or $Gag_{236-250}$ sequence (underlining) was introduced into the 5′ end of the EGFP cDNA.

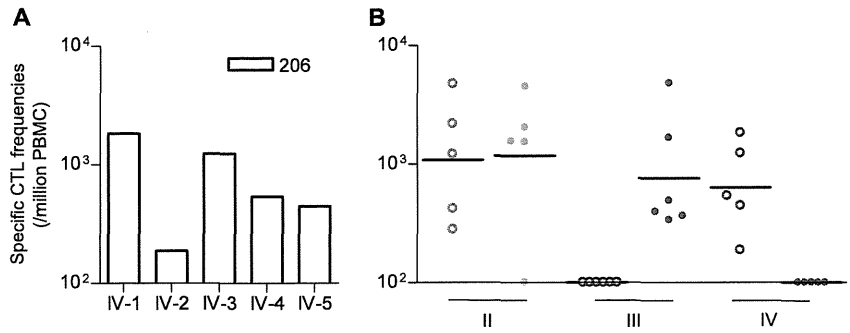

FIG 2 $Gag_{206-216}$-specific and $Gag_{241-249}$-specific CTL responses after prophylactic vaccination. (A) $Gag_{206-216}$-specific $CD8^+$ T-cell frequencies 1 week after SeV-$Gag_{202-216}$–EGFP boost in group IV macaques (open boxes). (B) $Gag_{206-216}$-specific (open circles) and $Gag_{241-249}$-specific (closed circles) $CD8^+$ T-cell frequencies 1 week after boost in group II (green), III (blue), and IV (red) macaques. The bars indicate the geometric mean of each group. No animal showed detectable Gag-specific CTL responses before the boost.

animals receiving a DNA and an SeV expressing EGFP showed no significant differences in viral loads after SIV challenge.

Analysis of antigen-specific CTL responses. We measured virus-specific $CD8^+$ T-cell levels by flow cytometric analysis of gamma interferon (IFN-γ) induction after specific stimulation, as described previously (21). Peripheral blood mononuclear cells (PBMCs) were cocultured with autologous herpesvirus papioimmortalized B-lymphoblastoid cell lines pulsed with 1 μM SIVmac239 $Gag_{206-216}$ (IINEEAADWDL), $Gag_{241-249}$ (SSVDEQIQW), or $Gag_{367-381}$ (ALKEALAPVPIPFAA) peptide for $Gag_{206-216}$-specific, $Gag_{241-249}$-specific, or $Gag_{367-381}$-specific stimulation. Intracellular IFN-γ staining was performed with a CytofixCytoperm kit (BD, Tokyo, Japan) and fluorescein isothiocyanate-conjugated anti-human CD4 (BD), peridinin chlorophyll protein-conjugated anti-human CD8 (BD), allophycocyanin (APC)-Cy7-conjugated anti-human CD3 (BD), and phycoerythrin (PE)-conjugated anti-human IFN-γ (Biolegend, San Diego, CA) monoclonal antibodies. Specific T-cell levels were calculated by subtracting nonspecific IFN-γ T-cell frequencies from those after peptide-specific stimulation. Specific T-cell levels lower than 100 per million PBMCs were considered negative.

Sequencing of the viral genome. Plasma RNA was extracted using the High Pure viral RNA kit (Roche Diagnostics, Tokyo, Japan). Fragments corresponding to nucleotides from 1231 to 2958 (containing the entire *gag* region) in the SIVmac239 genome (GenBank accession number M33262) were amplified by nested reverse transcription (RT)-PCR. The PCR products were sequenced using dye terminator chemistry and an automated DNA sequencer (Applied Biosystems, Tokyo, Japan).

Statistical analysis. Statistical analyses were performed using R software (R Development Core Team). Differences in geometric means of plasma viral loads were examined by one-way analysis of variance (ANOVA) and Tukey-Kramer's multiple-comparison test. Plasma viral loads at week 3 were examined for differences between group III and groups II and IV by analysis of covariance (ANCOVA) with week 5 viral loads as a covariate.

RESULTS

CTL responses after prophylactic vaccination. We previously reported the efficacy of vaccination eliciting whole Gag-specific or single $Gag_{241-249}$ epitope-specific CTL memory against SIVmac239 challenge (10, 32). In the present study, we examined the efficacy of prophylactic induction of single $Gag_{206-216}$ epitope-specific CTL memory against SIVmac239 challenge and compared the results with those of the previous experiments.

Five Burmese rhesus macaques possessing MHC-I haplotype *90-120-Ia* received a DNA-prime/SeV-boost vaccine eliciting single $Gag_{206-216}$ epitope-specific CTL responses. A plasmid DNA (p$Gag_{202-216}$-EGFP-N1) and an SeV (SeV-$Gag_{202-216}$-EGFP) vector, both expressing an SIVmac239 $Gag_{202-216}$-EGFP fusion pro-

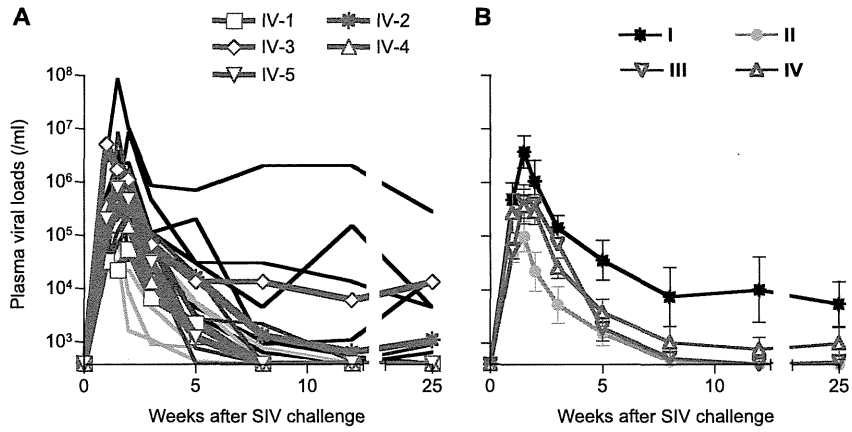

FIG 3 Plasma viral loads after SIVmac239 challenge. The plasma viral loads in group I, group II, group III, and group IV animals were determined as described previously (21). The lower limit of detection was approximately 4×10^2 copies/ml. (A) Changes in plasma viral loads (SIV *gag* RNA copies/ml plasma) after challenge. (B) Changes in geometric means of plasma viral loads after challenge. Groups II and III (but not group IV) showed significantly lower set point viral loads than group I ($P = 0.0390$ between groups I and II, $P = 0.0404$ between groups I and III, and $P > 0.05$ between groups I and IV at week 25 by one-way ANOVA and Tukey-Kramer's multiple-comparison test).

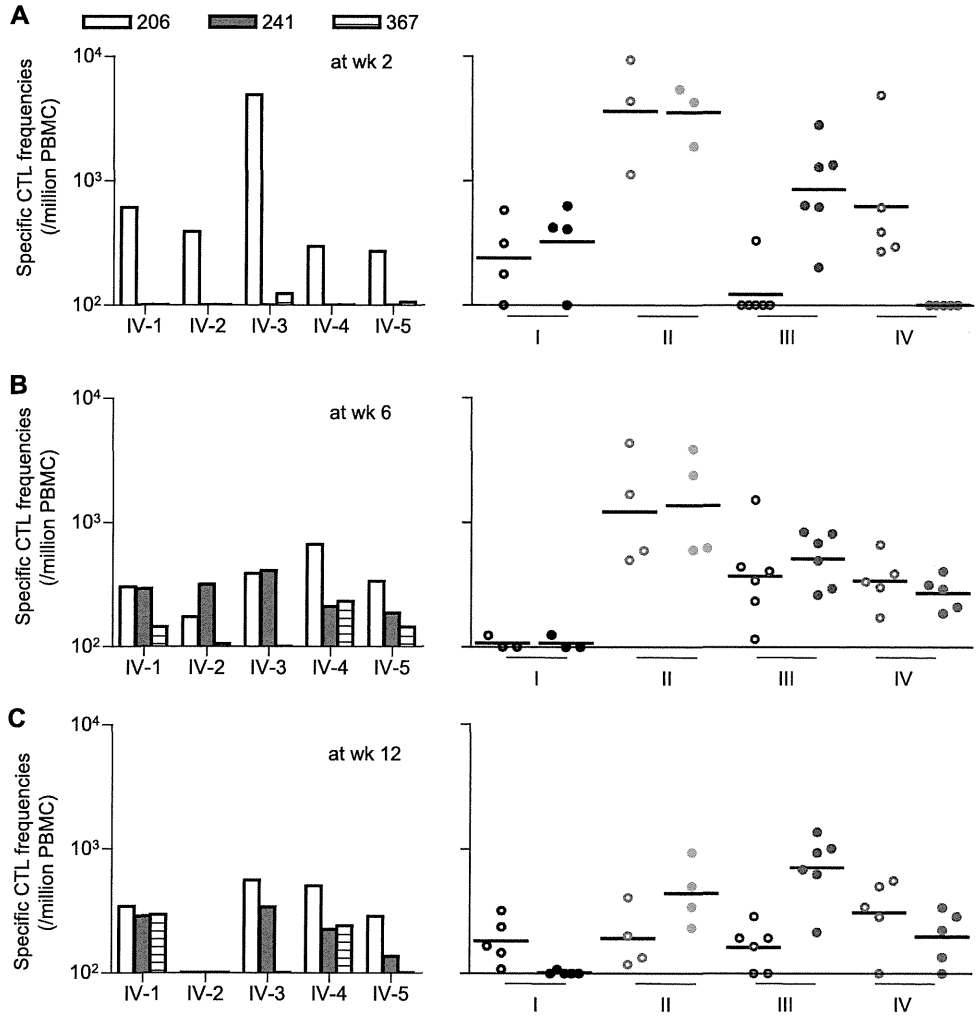

FIG 4 $Gag_{206-216}$-specific and $Gag_{241-249}$-specific CTL responses after SIVmac239 challenge. CTL responses at week 2 (A), week 6 (B), and week 12 (C) are shown. In the graphs on the left, $Gag_{206-216}$-specific (open boxes), $Gag_{241-249}$-specific (closed boxes), and $Gag_{367-381}$-specific (striped boxes) $CD8^+$ T-cell frequencies in group IV macaques are shown. On the right, $Gag_{206-216}$-specific (open circles) and $Gag_{241-249}$-specific (closed circles) $CD8^+$ T-cell frequencies in group I (black), II (green), III (blue), and IV (red) macaques are shown. The bars indicate the geometric mean of each group. Samples from macaques I-1, I-6, II-1, and II-3 at week 2; macaques I-1, I-2, I-6, and II-5 at week 6; and macaques I-1 and II-5 at week 12 were unavailable for this analysis. Statistical analyses among four groups at week 12 revealed significant differences in $Gag_{241-249}$-specific CTL levels (I and III, $P < 0.0001$; I and II, and III and IV, $P < 0.01$; I and IV, II and III, and II and IV, $P > 0.05$ by one-way ANOVA and Tukey-Kramer's multiple-comparison test) but not in $Gag_{206-216}$-specific CTL levels ($P > 0.05$ by one-way ANOVA).

tein, were used for the vaccination (Fig. 1). We confirmed $Gag_{206-216}$-specific CTL responses 1 week after SeV-$Gag_{202-216}$-EGFP boost in all five animals (Fig. 2A). As expected, no $Gag_{241-249}$-specific CTL responses were detected in these animals. No $Gag_{202-216}$-specific $CD4^+$ T-cell responses were detected in the animals except for one (IV-5) showing marginal levels of responses (data not shown).

Plasma viral loads after SIV challenge. We compared these five animals (referred to as group IV) with other groups (I, II, and III) of *90-120-Ia*-positive macaques reported previously (Table 1). Group I animals ($n = 6$) received no vaccination, group II ($n = 5$) received a DNA-prime/SeV-boost vaccine eliciting whole Gag-specific CTL responses, and group III ($n = 6$) received a DNA-prime/SeV-boost vaccine eliciting single $Gag_{241-249}$ epitope-specific CTL responses. Both $Gag_{206-216}$-specific and $Gag_{241-249}$-specific CTL responses were detectable after SeV-Gag boost in four of five group II animals except for one animal (II-3), in which $Gag_{206-216}$-specific, but not $Gag_{241-249}$-specific, CTL responses were detected. In all group III animals, $Gag_{241-249}$-specific CTL responses were confirmed, while no $Gag_{206-216}$-specific CTL responses were detected after SeV-$Gag_{236-250}$-EGFP boost (Fig. 2B).

After SIVmac239 challenge, all animals were infected and showed plasma viremia during the acute phase. Plasma viremia was maintained in five of six unvaccinated animals in group I but became undetectable in one animal (I-2) at week 12. In contrast, all animals in groups II and III contained SIV replication with significantly reduced plasma viral loads compared to group I at the set point. In group IV, however, vaccine efficacy was not so clear; while three out of five animals contained SIV replication, the remaining two (IV-2 and IV-3) failed to control viral replication with persistent plasma viremia (Fig. 3).

Gag-specific CTL responses after SIV challenge. We then measured $Gag_{206-216}$-specific and $Gag_{241-249}$-specific CTL responses after SIVmac239 challenge by detection of peptide-

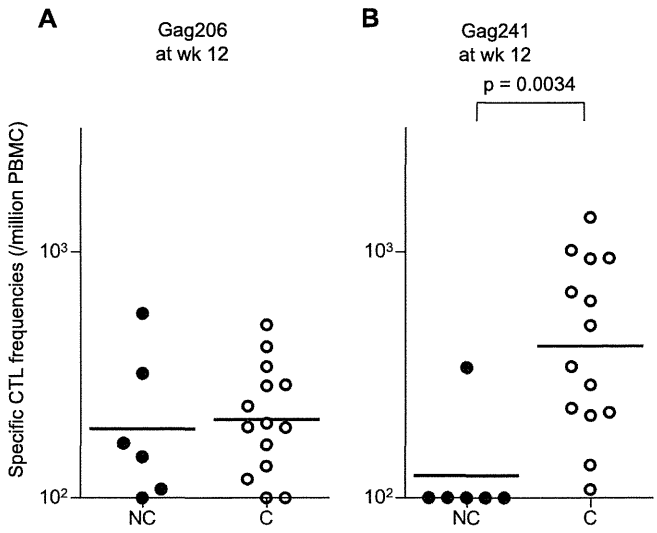

FIG 5 Comparison of $Gag_{206-216}$-specific or $Gag_{241-249}$-specific CTL responses in noncontrollers and controllers at week 12. (A) $Gag_{206-216}$-specific $CD8^+$ T-cell frequencies in noncontrollers (NC; closed circles) and controllers (C; open circles). (B) $Gag_{241-249}$-specific $CD8^+$ T-cell frequencies in noncontrollers and controllers. $Gag_{241-249}$-specific CTL levels in controllers were significantly higher than those in noncontrollers ($P = 0.0034$ by Mann-Whitney test). The bars indicate the geometric mean of each group. Data on a noncontroller (I-1) and a controller (II-5) were unavailable.

specific IFN-γ induction. At week 2 (Fig. 4A), most animals in groups I and II elicited both $Gag_{206-216}$-specific and $Gag_{241-249}$-specific CTL responses, whereas group III animals induced $Gag_{241-249}$-specific CTL responses dominantly. Remarkably, all animals in group IV showed efficient $Gag_{206-216}$-specific CTL responses without detectable $Gag_{241-249}$-specific CTL responses at week 2. These results indicate dominant $Gag_{206-216}$-specific CTL responses with delayed induction of $Gag_{241-249}$-specific CTL responses postchallenge in group IV animals with prophylactic $Gag_{206-216}$-specific CTL induction, and vice versa in group III animals.

At week 6 (Fig. 4B), efficient $Gag_{206-216}$-specific and $Gag_{241-249}$-specific CTL responses were observed in all vaccinated animals in groups II, III, and IV, but not in group I. $Gag_{206-216}$-specific and $Gag_{241-249}$-specific CTL responses were induced equivalently even in groups III and IV. We also examined subdominant $Gag_{367-381}$ epitope-specific CTL responses, which were undetectable at week 2 but became detectable at week 6 in most group IV animals (Fig. 4, graphs on left). At week 12 (Fig. 4C), however, different CTL immunodominance patterns were observed among the groups. $Gag_{241-249}$-specific CTL levels were higher than $Gag_{206-216}$-specific levels in groups II and III but were reduced in groups I and IV. Interestingly, comparison between the animals with persistent viremia (referred to as noncontrollers) and those controlling SIV replication (referred to as controllers) revealed significant differences in $Gag_{241-249}$-specific CTL levels, but not in $Gag_{206-216}$-specific levels, at week 12 ($P = 0.0034$ by Mann-Whitney test) (Fig. 5).

Selection of a CTL escape mutation. Next, we examined viral genome *gag* sequences at weeks 5 and 12 after challenge to determine whether CTL escape mutations were selected in these animals (Table 2). At week 5, a mutation leading to an L-to-S substitution at the 216th residue in Gag (L216S) was selected in all the group II animals. This GagL216S change results in escape from $Gag_{206-216}$-specific CTL recognition, as described previously (21). All the group IV animals with $Gag_{206-216}$-specific CTL induction also showed rapid selection of this CTL escape mutation at week 5. Analysis at week 3 found the GagL216S mutation dominant in two (II-2 and II-5) group II and two (IV-1 and IV-3) group IV animals (data not shown). However, animals in group III showed no *gag* mutations at week 5, except for one animal (III-5) selecting a mutation leading to an L-to-F substitution at the 216th residue. Later, at week 12, the $Gag_{206-216}$-specific CTL escape mutation, GagL216S, was selected even in group III animals. No animals showed mutations around the $Gag_{241-249}$ epitope-coding region even at week 12. These results indicate that selection of this $Gag_{206-216}$-specific CTL escape mutation may be accelerated by prophylactic vaccination inducing $Gag_{206-216}$-specific CTL responses. On the other hand, in group III animals with single $Gag_{241-249}$ epitope-specific CTL induction, selection of a $Gag_{206-216}$-specific CTL escape mutation was delayed but was observed before selection of a $Gag_{241-249}$-specific CTL escape mutation, suggesting strong selective pressure by delayed $Gag_{206-216}$-specific CTL responses after SIV challenge.

In order to see the effect of rapid selection of the $Gag_{206-216}$-specific CTL escape mutation on SIV control, we compared plasma viral loads at weeks 3 and 5 between groups II and IV (referred to as group II+IV) with rapid selection of the GagL216S

TABLE 2 Selection of a CTL escape mutation

Group	Macaque ID	Amino acid change for Gag residues[b]: 206–216 Wk 5	206–216 Wk 12	241–249 Wk 5	241–249 Wk 12
I	I-1	None	ND	None	ND
	I-2[a]	None	L216S	None	None
	I-3	None	L216S	None	None
	I-4	None	None	None	None
	I-5	None	None	None	None
	I-6	None	None	None	None
II	II-1[a]	L216S	ND	None	ND
	II-2[a]	L216S	ND	None	ND
	II-3[a]	L216S	ND	None	ND
	II-4[a]	L216S	ND	None	ND
	II-5[a]	L216S	ND	None	ND
III	III-1[a]	None	L216S	None	None
	III-2[a]	None	L216S	None	None
	III-3[a]	None	NA	None	NA
	III-4[a]	None	NA	None	NA
	III-5[a]	L216F	L216S	None	None
	III-6[a]	None	L216S	None	None
IV	IV-1[a]	L216S	L216S	None	None
	IV-2	L216S	L216S	None	None
	IV-3	L216S	L216S	None	None
	IV-4[a]	L216S	L216S	None	None
	IV-5[a]	L216S	NA	None	NA

[a] Animals that controlled SIV replication at week 12 (controllers).
[b] Plasma viral gag genome mutations were examined at weeks 5 and 12. Amino acid substitutions in $Gag_{206-216}$ and $Gag_{241-249}$ epitope regions are shown. L216S results in viral escape from $Gag_{206-216}$-specific CTL recognition. It remains undetermined whether L216F results in CTL escape. ND, not determined; NA, not determined because Gag fragments were unable to be amplified from plasma RNA.

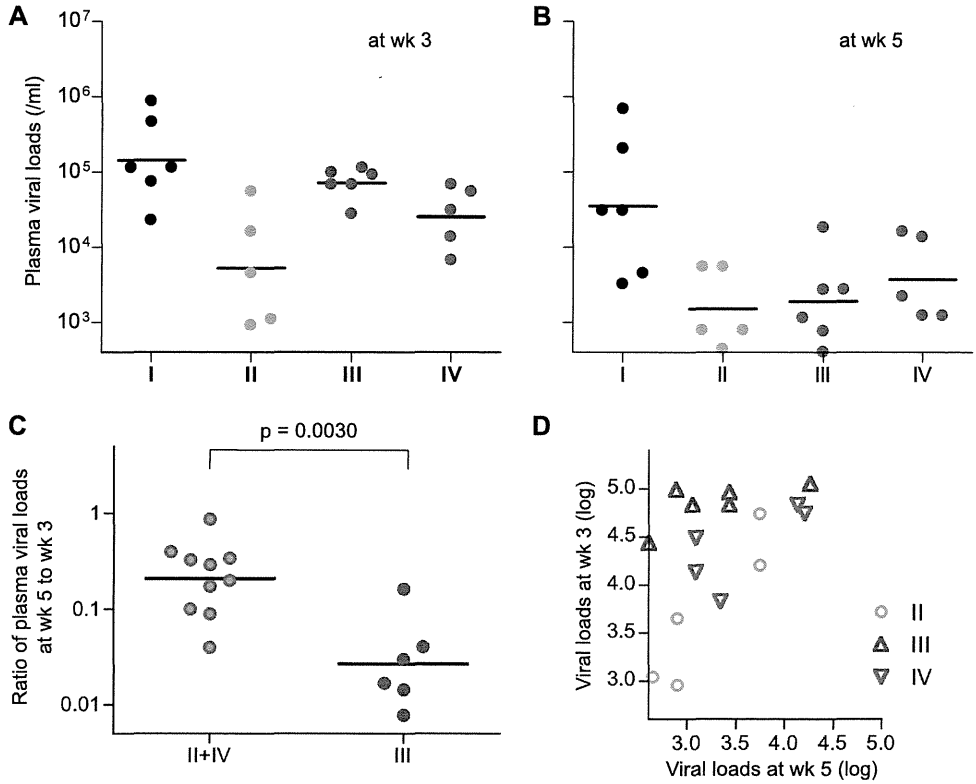

FIG 6 Comparison of plasma viral loads at weeks 3 and 5 among four groups. (A) Plasma viral loads at week 3 in group I, II, III, and IV animals. (B) Plasma viral loads at week 5 in group I, II, III, and IV animals. (C) Comparison of ratios of plasma viral loads at week 5 to week 3 in group II+IV animals and group III animals. The ratios in group III were significantly lower than those in group II+IV ($P = 0.0030$ by Mann-Whitney test). The bars indicate the geometric mean of each group. (D) Scatter plots between plasma viral loads at weeks 3 and 5 in group II, III, and IV animals.

mutation and group III without the mutation at week 5 (Fig. 6). Ratios of plasma viral loads at week 5 to week 3 in group III were significantly lower than those in group II+IV ($P = 0.0030$ by Mann-Whitney test) (Fig. 6C). To confirm this result, we examined the difference in week 3 viral loads between groups III and II+IV by ANCOVA, with week 5 viral loads as a covariate. This analysis revealed that week 3 viral loads controlled for by week 5 viral loads were significantly higher in group III than those in group II+IV (Fig. 6D and Table 3); i.e., the decline in viral loads from week 3 to week 5 was significantly sharper in group III than in group II+IV, possibly reflecting viral escape from suppressive pressure by $Gag_{206\text{-}216}$-specific CTL responses in the latter group during this period (from week 3 to week 5).

DISCUSSION

In the present study, we analyzed the impact of vaccination inducing single $Gag_{206\text{-}216}$ epitope-specific CTL memory on postchallenge CTL responses and SIV control in *90-120-Ia*-positive macaques and then compared the results with those of vaccination inducing single $Gag_{241\text{-}249}$ epitope-specific CTL responses. Our results indicate that these prophylactic vaccinations result in different patterns of $Gag_{206\text{-}216}$-specific and $Gag_{241\text{-}249}$-specific CTL immunodominance and cooperation after SIVmac239 challenge.

Unvaccinated *90-120-Ia*-positive macaques (group I) showed both $Gag_{206\text{-}216}$-specific and $Gag_{241\text{-}249}$-specific CTL responses after SIV challenge. In group IV animals with prophylactic induc-

TABLE 3 ANCOVA on week 3 viral loads with week 5 viral loads as a covariate between groups III and II+IV

ANOVA	Parameter	SS[a]	df[b]	MS[c]	*F*	*P* value
Homogeneity of slopes of regression	Group × slope	0.304	1	0.304	2.099	0.173
	Residual	1.735	12	0.145		
	Total	2.038	13	0.157		
Difference in week 3 viral loads with week 5 viral loads as a covariate between groups III and II+IV	Effect and group	1.106	1	1.106	7.052	0.020
	Residual	2.038	13	0.157		
	Total	3.144	14	0.225		

[a] SS, sum of squares.
[b] df, degrees of freedom.
[c] MS, mean squares.

tion of single $Gag_{206-216}$ epitope-specific CTL responses, $Gag_{206-216}$-specific CTL responses were induced dominantly but $Gag_{241-249}$-specific CTL responses were undetectable at week 2. In contrast, $Gag_{241-249}$-specific CTL responses were induced dominantly at week 2 in group III. Both groups showed $Gag_{206-216}$-specific and $Gag_{241-249}$-specific CTL responses equivalently at week 6. It may be difficult to compare these results with those in group II animals inducing whole Gag antigen-specific CTL and $CD4^+$ T-cell responses before challenge; the group II animals elicited $Gag_{206-216}$-specific and $Gag_{241-249}$-specific CTL responses equivalently at week 2. Our results indicate that prophylactic vaccination results in dominant induction of vaccine antigen-specific CTL responses and may delay CTL responses specific for viral antigens other than vaccine antigens (referred to as nonvaccine antigens) after viral exposure.

A significant difference between groups III and IV is the pattern of selection of CTL escape mutation. All group IV animals showed rapid selection of a $Gag_{206-216}$-specific CTL escape mutation, while most group III animals showed no *gag* mutation at week 5 but selection of the $Gag_{206-216}$-specific CTL escape mutation later, at week 12. Thus, prophylactic vaccination may affect the patterns of viral genome diversification, possibly accelerating selection of CTL escape mutations. Interestingly, $Gag_{241-249}$-specific CTL mutations were not detected even at week 12 in group III animals, although a previous study observed not only the $Gag_{206-216}$-specific CTL escape mutation (GagL216S), but also a $Gag_{241-249}$-specific CTL escape mutation (GagD244E) in the chronic phase of SIV infection in *90-120-Ia*-positive macaques (9). These results indicate that delayed, naive-derived $Gag_{206-216}$-specific CTL responses, as well as preceding $Gag_{241-249}$-specific CTL responses, exert strong suppressive pressure on SIV replication in group III animals, implying cooperation between vaccine antigen-specific and non-vaccine antigen-specific CTL responses for virus control.

Rapid selection of the $Gag_{206-216}$-specific CTL escape mutation (GagL216S) in group II and delayed selection of this mutation without a detectable $Gag_{241-249}$-specific CTL escape mutation (GagD244E) in group III suggest that the virus with GagL216S (SIVmac239Gag216S) replicates more efficiently than the virus with GagD244E (SIVmac239Gag244E) under both $Gag_{206-216}$-specific and $Gag_{241-249}$-specific CTL responses. Our previous competition assay did not find a significant difference in viral fitness between these mutant viruses. Possibly, escape of SIVmac239Gag216S from $Gag_{206-216}$-specific CTL pressure may be more efficient than that of SIVmac239Gag244E from $Gag_{241-249}$-specific CTL pressure.

Our analysis revealed that the decline of plasma viral loads from week 3 to week 5 in group II+IV with rapid selection of the GagL216S mutation was significantly less than that in group III without the mutation at week 5, possibly reflecting viral escape from suppressive pressure by $Gag_{206-216}$-specific CTL responses in the former groups around weeks 3 to 5. Even the comparison between groups II and III, both showing dominant $Gag_{241-249}$-specific CTL responses at week 2, revealed a significantly sharper decline in the latter ($P = 0.0087$). Thus, our results suggest three patterns of $Gag_{206-216}$-specific and $Gag_{241-249}$-specific CTL cooperation for virus control after SIVmac239 challenge. First, as observed in group II, dominantly induced $Gag_{206-216}$-specific and $Gag_{241-249}$-specific CTL responses both work against wild-type SIV replication around week 2, but then a mutant virus escaping from the former CTL responses is selected, and the responses work against this mutant virus replication. Second, as observed in group III, dominantly induced $Gag_{241-249}$-specific CTL responses work against wild-type SIV replication around week 2 and then contribute to virus control, together with delayed, naive-derived $Gag_{206-216}$-specific CTL responses. Third, as observed in group IV, dominantly induced $Gag_{206-216}$-specific CTL responses work against wild-type SIV replication around week 2, but then a mutant virus escaping from $Gag_{206-216}$-specific CTL responses is selected, and delayed, naive-derived $Gag_{241-249}$-specific CTL responses instead work against this mutant virus replication. Viral loads at week 3 in group III looked higher than those in group IV, implying that $Gag_{206-216}$-specific CTL responses may exert a stronger suppressive effect on SIV replication in the acute phase than $Gag_{241-249}$-specific CTL responses. However, viral loads at week 5 in group III looked lower than those in group IV, and the comparison between the two groups showed significantly less decline in the latter ($P = 0.0303$). It is speculated that the third pattern observed in group IV is prone to failure in virus control. Indeed, two of five animals in group IV failed to control SIV replication. Even if vaccines are designed to express multiple antigens, of the vaccine-induced CTLs generated, only several epitope-specific cells may recognize the incoming HIV because of viral diversity and host MHC polymorphisms (18), and cooperation of these vaccine antigen-specific and non-vaccine antigen-specific CTL responses would be required for viral control. Thus, our results may imply a rationale of inducing escape-resistant, epitope-specific CTL memory by prophylactic AIDS vaccines.

In summary, this study showed dominant induction of vaccine antigen-specific CTL responses and delay in non-vaccine antigen-specific CTL responses in the acute phase of SIV infection, clearly describing the impact of prophylactic vaccination on CTL immunodominance and cooperation after virus exposure. Our results indicate that the patterns of cooperation of vaccine antigen-specific and non-vaccine antigen-specific CTL responses affect virus control and selection of CTL escape mutations. These findings provide great insights into antigen design in the development of a CTL-inducing AIDS vaccine.

ACKNOWLEDGMENTS

This work was supported by grants-in-aid from the Ministry of Education, Culture, Sports, Science, and Technology; grants-in-aid from the Ministry of Health, Labor, and Welfare; and a grant from the Takeda Science Foundation, Japan.

We thank F. Ono, K. Oto, K. Komatsuzaki, M. Hamano, Y. Emoto, H. Akari, and Y. Yasutomi for their assistance in animal experiments.

REFERENCES

1. **Altfeld M, et al.** 2003. Influence of HLA-B57 on clinical presentation and viral control during acute HIV-1 infection. AIDS 17:2581–2591.
2. **Berger CT, et al.** 2011. High functional avidity CTL responses to HLA-B-restricted Gag-derived epitopes associate with relative HIV control. J. Virol. 85:9334–9345.
3. **Borrow P, Lewicki H, Hahn BH, Shaw GM, Oldstone MB.** 1994. Virus-specific CD8 cytotoxic T-lymphocyte activity associated with control of viremia in primary human immunodeficiency virus type 1 infection. J. Virol. 68:6103–6110.
4. **Goulder PJ, Watkins DI.** 2004. HIV and SIV CTL escape: implications for vaccine design. Nat. Rev. Immunol. 4:630–640.
5. **Goulder PJ, Watkins DI.** 2008. Impact of MHC class I diversity on immune control of immunodeficiency virus replication. Nat. Rev. Immunol. 8:619–630.
6. **Hansen SG, et al.** 2009. Effector memory T cell responses are associated

with protection of rhesus monkeys from mucosal simian immunodeficiency virus challenge. Nat. Med. **15**:293–299.
7. **Julg B, et al.** 2010. Enhanced anti-HIV functional activity associated with Gag-specific CD8 T-cell responses. J. Virol. **84**:5540–5549.
8. **Kaslow RA, et al.** 1996. Influence of combinations of human major histocompatibility complex genes on the course of HIV-1 infection. Nat. Med. **2**:405–411.
9. **Kawada M, et al.** 2006. Involvement of multiple epitope-specific cytotoxic T-lymphocyte responses in vaccine-based control of simian immunodeficiency virus replication in rhesus macaques. J. Virol. **80**:1949–1958.
10. **Kawada M, et al.** 2008. Gag-specific cytotoxic T-lymphocyte-based control of primary simian immunodeficiency virus replication in a vaccine trial. J. Virol. **82**:10199–10206.
11. **Kestler HW III, et al.** 1991. Importance of the nef gene for maintenance of high virus loads and for development of AIDS. Cell **65**:651–662.
12. **Kiepiela P, et al.** 2007. CD8+ T-cell responses to different HIV proteins have discordant associations with viral load. Nat. Med. **13**:46–53.
13. **Koup RA, et al.** 1994. Temporal association of cellular immune responses with the initial control of viremia in primary human immunodeficiency virus type 1 syndrome. J. Virol. **68**:4650–4655.
14. **Leslie A, et al.** 2010. Additive contribution of HLA class I alleles in the immune control of HIV-1 infection. J. Virol. **84**:9879–9888.
15. **Letvin NL, et al.** 2006. Preserved CD4+ central memory T cells and survival in vaccinated SIV-challenged monkeys. Science **312**:1530–1533.
16. **Li F, et al.** 2011. Mapping HIV-1 vaccine induced T-cell responses: bias towards less-conserved regions and potential impact on vaccine efficacy in the Step study. PLoS One **6**:e20479.
17. **Liu J, et al.** 2009. Immune control of an SIV challenge by a T-cell-based vaccine in rhesus monkeys. Nature **457**:87–91.
18. **Liu Y, McNevin JP, Holte S, McElrath MJ, Mullins JI.** 2011. Dynamics of viral evolution and CTL responses in HIV-1 infection. PLoS One **6**:e15639.
19. **Loffredo JT, et al.** 2008. Patterns of CD8 immunodominance may influence the ability of Mamu-B*08-positive macaques to naturally control simian immunodeficiency virus SIVmac239 replication. J. Virol. **82**:1723–1738.
20. **Matano T, Kano M, Nakamura H, Takeda A, Nagai Y.** 2001. Rapid appearance of secondary immune responses and protection from acute CD4 depletion after a highly pathogenic immunodeficiency virus challenge in macaques vaccinated with a DNA prime/Sendai virus vector boost regimen. J. Virol. **75**:11891–11896.
21. **Matano T, et al.** 2004. Cytotoxic T lymphocyte-based control of simian immunodeficiency virus replication in a preclinical AIDS vaccine trial. J. Exp. Med. **199**:1709–1718.
22. **Matano T, et al.** 1998. Administration of an anti-CD8 monoclonal antibody interferes with the clearance of chimeric simian/human immunodeficiency virus during primary infections of rhesus macaques. J. Virol. **72**:164–169.
23. **Migueles SA, et al.** 2000. HLA B*5701 is highly associated with restriction of virus replication in a subgroup of HIV-infected long term nonprogressors. Proc. Natl. Acad. Sci. U. S. A. **97**:2709–2714.
24. **Miura T, et al.** 2009. HLA-B57/B*5801 human immunodeficiency virus type 1 elite controllers select for rare gag variants associated with reduced viral replication capacity and strong cytotoxic T-lymphocyte [corrected] recognition. J. Virol. **83**:2743–2755.
25. **Mothé BR, et al.** 2003. Expression of the major histocompatibility complex class I molecule Mamu-A*01 is associated with control of simian immunodeficiency virus SIVmac239 replication. J. Virol. **77**:2736–2740.
26. **Naruse TK, et al.** 2010. Diversity of MHC class I genes in Burmese-origin rhesus macaques. Immunogenetics **62**:601–611.
27. **Ndhlovu ZM, et al.** 2011. Mosaic HIV-1 Gag antigens can be processed and presented to human HIV-specific CD8+ T cells. J. Immunol. **186**:6914–6924.
28. **Schmitz JE, et al.** 1999. Control of viremia in simian immunodeficiency virus infection by CD8+ lymphocytes. Science **283**:857–860.
29. **Shibata R, et al.** 1997. Infection and pathogenicity of chimeric simian-human immunodeficiency viruses in macaques: determinants of high virus loads and CD4 cell killing. J. Infect. Dis. **176**:362–373.
30. **Streeck H, et al.** 2009. Human immunodeficiency virus type 1-specific CD8+ T-cell responses during primary infection are major determinants of the viral set point and loss of CD4+ T cells. J. Virol. **83**:7641–7648.
31. **Tang J, et al.** 2002. Favorable and unfavorable HLA class I alleles and haplotypes in Zambians predominantly infected with clade C human immunodeficiency virus type 1. J. Virol. **76**:8276–8284.
32. **Tsukamoto T, et al.** 2009. Impact of cytotoxic-T-lymphocyte memory induction without virus-specific CD4+ T-cell help on control of a simian immunodeficiency virus challenge in rhesus macaques. J. Virol. **83**:9339–9346.
33. **Turnbull EL, et al.** 2009. Kinetics of expansion of epitope-specific T cell responses during primary HIV-1 infection. J. Immunol. **182**:7131–7145.
34. **Wang YE, et al.** 2009. Protective HLA class I alleles that restrict acute-phase CD8+ T-cell responses are associated with viral escape mutations located in highly conserved regions of human immunodeficiency virus type 1. J. Virol. **83**:1845–1855.
35. **Wilson NA, et al.** 2006. Vaccine-induced cellular immune responses reduce plasma viral concentrations after repeated low-dose challenge with pathogenic simian immunodeficiency virus SIVmac239. J. Virol. **80**:5875–5885.
36. **Yant LJ, et al.** 2006. The high-frequency major histocompatibility complex class I allele Mamu-B*17 is associated with control of simian immunodeficiency virus SIVmac239 replication. J. Virol. **80**:5074–5077.

Immunogenetics
DOI 10.1007/s00251-011-0556-2

ORIGINAL PAPER

Positive selection of Toll-like receptor 2 polymorphisms in two closely related old world monkey species, rhesus and Japanese macaques

Akiko Takaki · Akiko Yamazaki · Tomoyuki Maekawa · Hiroki Shibata · Kenji Hirayama · Akinori Kimura · Hirohisa Hirai · Michio Yasunami

Received: 7 January 2011 / Accepted: 21 June 2011

Abstract Toll-like receptor 2 (TLR2) plays an important role in the recognition of a variety of pathogenic microbes. In the present study, we compared polymorphisms of *TLR2* locus in two closely related old world monkey species, rhesus monkey (*Macaca mulatta*) and Japanese monkey (*Macaca fuscata*). By nucleotide sequencing of the third exon of *TLR2* gene from 21 to 35 respective individuals, we could assign 17 haplotype combinations of 17 coding SNPs of ten non-synonymous and seven synonymous substitutions. A non-synonymous substitution at codon position 326 appeared to be differentially fixed in each species, asparagine for *M. mulatta* whereas tyrosine for *M. fuscata*, and may contribute to certain functional properties because it locates in the region contributing to ligand binding and interaction with dimerization partner of TLR2-TLR1 heterodimeric complex. Although *TLR2* alleles have diverged to similar extent in both species, they have evolved in significantly different ways; *TLR2* of *M. fuscata* has undergone purifying selection while the membrane-proximal part of the extracellular domain of *M. mulatta TLR2* exhibits higher rates of non-synonymous substitutions, indicating a trace of Darwinian positive selection.

Keywords Innate immunity · TLR · Polymorphism · Nonhuman primate · Molecular evolution · Reporter gene assay

Electronic supplementary material The online version of this article (doi:10.1007/s00251-011-0556-2) contains supplementary material, which is available to authorized users.

A. Takaki · A. Yamazaki · H. Shibata · K. Hirayama · M. Yasunami (✉)
Department of Immunogenetics, Institute of Tropical Medicine, Nagasaki University,
1-12-4 Sakamoto,
Nagasaki 852-8523, Japan
e-mail: yasunami@nagasaki-u.ac.jp

A. Takaki · T. Maekawa · M. Yasunami
Department of Clinical Medicine, Institute of Tropical Medicine, Nagasaki University,
Nagasaki, Japan

A. Kimura
Department of Molecular Pathogenesis,
Medical Research Institute, Tokyo Medical and Dental University,
Tokyo, Japan

H. Hirai
Department of Cellular and Molecular Biology,
Primate Research Institute, Kyoto University,
Kyoto, Japan

Introduction

The rhesus macaque, *Macaca mulatta*, is one of the best known old world monkeys and has been used for various biomedical researches as a nonhuman primate model including infections of simian immunodeficiency virus (Ling et al. 2002; Matano et al. 2004) and *Mycobacterium tuberculosis* (McMurray 2000; Huang et al. 2007). *M. mulatta* belongs to the primate family Cercopithecidae that shares the last common ancestor of approximately 25 million years ago (Mya) with human and hominoids (Kumar and Hedges 1998). According to this fact, nucleotide sequence similarity between humans and *M. mulatta* has been maintained as high as 93% in average (Rhesus Macaque Genome Sequencing and Analysis Consortium 2007). Analyses of molecular evolution of mitochondrial and nuclear DNA among species of genus *Macaca* estimate the divergence between rhesus and

Published online: 09 July 2011

cynomolgus macaque, *Macaca fascicularis* at 0.9–2.5 Mya (Hayasaka et al. 1996; Blancher et al. 2008; Osada et al. 2008; Stevison and Kohn 2009). Japanese monkey, *Macaca fuscata* and Taiwanese monkey, *Macaca cyclopis* had been derived from the rhesus lineage relatively recently, and the geographical isolation is fundamental for the diversification of these species (Smith et al. 2007). *M. mulatta* has a relatively broad geographical distribution from Afghanistan across Asia to the Chinese shore of the Pacific Ocean (Rhesus Macaque Genome Sequencing and Analysis Consortium 2007), where the tropical infectious agents such as *Plasmodium* species and arthropod-borne viruses are circulating and may influence the host genome as selection pressure. To date, a number of comparative studies on primate genome including *Macaca* species have been conducted. Among them, immune-related genes such as *TNFA* (Baena et al. 2007), eosinophil cationic protein (*ECP*), and eosinophil-derived neurotoxin (*EDN*) ribonuclease (Zhang et al. 1998) are shown to be evolved under positive Darwinian selection (or diversifying selection), suggesting the presence of a differential pressure between primate species.

Toll-like receptors (TLRs) are the first line of the host defense mechanisms against pathogenic microorganisms as pattern-recognition receptors (PRRs) (Akira et al. 2006). A number of single nucleotide polymorphisms (SNPs) have been found in the components of the TLR signaling pathway in humans. Alterations of the structure of these signaling molecules are often associated with susceptibility to various infectious diseases. In the case of human TLR2 that binds to peptidoglycan and lipoteichoic acid derived from Gram-positive bacteria as well as a variety of macromolecules from other microbes (Texereau et al. 2005), a non-synonymous substitution of glutamine for arginine at amino acid position 753 (Arg753Gln) was shown to be a risk factor of developing tuberculosis (Bochud et al. 2003; Ogus et al. 2004) and septic shock (Lorenz et al. 2000), whereas, have a protective effect against the development of Lyme disease (Schröder et al. 2005). It is hypothesized that TLR2 is also variable in *Macaca* species and exhibits a trace of molecular evolution events which is conferred by possible differences in individual response to various infectious agents. Previous reports on molecular evolution of TLRs of primates, which support the functional conservation of intra-cellular signal transduction machinery involving a common functional cytoplasmic Toll/Interleukin-1 receptor (TIR) domain during mammalian evolution, also suggested the presence of the positive Darwinian selection on the extracellular domain of TLR4 in *Macaca* species (Sanghavi et al. 2004; Nakajima et al. 2008).

In the present study, SNPs in the coding sequence of TLR2 from two closely related *Macaca* species, *M. mulatta* and *M. fuscata* were investigated. By nucleotide sequencing of 21 *M. mulatta* and 35 *M. fuscata* individuals, ten non-synonymous and seven synonymous substitutions were identified in the coding region. These 17 SNPs compose 17 haplotype combinations (or alleles) existing in the examined population. The ratio of rates of non-synonymous versus synonymous substitutions suggested the conservation of overall amino acid sequence except for a part of extracellular domain of *M. mulatta* TLR2, where multiple amino acid substitutions have taken place to give rise different alleles. Although the functional difference of TLR2 alleles was not evident by in vitro transfection study of expression vector to HEK293 cells, molecular modeling suggested that modified receptor-ligand interaction conferred by amino acid substitution in *M. mulatta* TLR2 is a driving force of diversifying evolution.

Materials and methods

Genomic DNA samples

Individual genomic DNA of *M. mulatta* were prepared from B lymphoblastoid cell lines which had been established from peripheral blood of mutually unrelated founders of breeding colony originated from wild population in Myanmar and Laos as described before (Takahashi-Tanaka et al. 2007). Samples of peripheral blood of *M. fuscata* were collected from three isolated colonies originated from different wild populations in Japan maintained at Primate Research Institute, Kyoto University, after the institutional review of experimental procedures. Genomic DNA of leukocytes was isolated using Wizard® Genomic DNA Purification Kit (Promega).

Nucleotide sequencing of coding region of TLR2

The coding sequence of TLR2 was amplified by PCR using KOD FX DNA polymerase (TOYOBO) with a pair of oligonucletides, TLR2 exon 3-forward (5′-ATTAGAAT TACGATATGCTGTC-3′) and TLR2 exon 3-reverse (5′-ATGACGGTACATCCACGTAG-3′) as primers, essentially according to manufacturer's recommendations. Sequencing reaction was performed using BigDye® Terminators v1.1 Cycle Sequencing Kit (Applied Biosystems) and analyzed by ABI3730xl automated DNA sequencer (Applied Biosystems).

Transfection of TLR2 expression vectors

PCR products including the entire coding sequence of TLR2 were ligated to pGEM®-TEasy (Promega) with T4 DNA ligase after treatment with A-attachment Mix (TOYOBO). After the confirmation of nucleotide sequence

of plasmid clones, the DNA fragment containing entire coding region for TLR2 was isolated by cleavage with restriction enzymes and inserted to pcDNA3.1/Hygro(+) expression vector (Invitrogen). Luciferase reporter for nuclear factor kappa B (NF-κB) activity, pGL4.22-6xκB was constructed by inserting DNA fragment containing 6 times tandem repeats of NF-κB binding sequence and TATA box (Shibata et al. 2006) in pGL4.22[luc2CP/Puro] promoterless firefly luciferase reporter vector (Promega). All expression plasmid DNA and reporter plasmid DNA were prepared by Illustra™ Plasmid Prep Midi Flow kit (GE Healthcare) and confirmed virtually endotoxin free.

HEK293 cells were cultured in Dulbecco's MEM (Wako) supplemented with 10% fetal bovine serum (GIBCO) and 100 U/ml penicillin–100 μg/ml streptomycin (Invitrogen) at 37°C in the presence of 5% CO_2. Synthetic bacterial lipoproteins as TLR2 agonists, Pam2CSK4 and Pam3CSK4 were purchased from InvivoGen and applied at 10 and 100 ng/ml, respectively. Transfection of DNA was performed as follows: HEK293 cells were inoculated in 35-mm tissue culture dishes at 4×10^5 cells/dish on the previous day, transfected with a mixture of 250 ng of each pcDNA3.1-TLR2 expression plasmid, 100 ng of pGL4.22-6xκB, and 20 ng pRL-TK internal control (Promega) with FuGENE® HD Transfection Reagent (Roche), then at 42 h after the addition of DNA, stimulated with TLR2 agonists for 6 h. Cells were washed with 1 ml of PBS and suspended in 100 μl of 1× Passive Lysis Buffer (Promega). Luciferase activities were determined using Dual-Luciferase® Reporter Assay System (Promega) by Wallac 1420 Multilabel Counter ARVO MA (PerkinElmer). Transfection was set up in triplicate to evaluate experimental variations and repeated at least twice to confirm the results. Difference in luciferase activity was statistically examined by ANOVA and Student's *t* test.

Analysis of molecular evolution

We estimate values of the number of segregating sites (S), number of haplotypes (H), haplotype diversity (Hd), nucleotide diversities (π), nucleotide polymorphism (θ), K_a, K_s, K_a/K_s ratio, and Tajima's D (Tajima 1989) by using DnaSP ver.5.10.00 (Librado and Rozas 2009). Genetic distances for non-synonymous substitution (d_N) and synonymous substitution (d_S) and the d_N-d_S difference were calculated according to Nei and Gojobori's method (Nei and Gojobori 1986) with Jukes and Canter's correction by MEGA4 (Kumar et al. 2008). Z tests for positive and purifying selections were also carried out on MEGA4, by which standard errors were produced with 500 bootstrap replications. Results were considered statistically significant when $p<0.05$. A neighbor-joining tree (Saitou and Nei 1987) was constructed on the basis of genetic distances from alignments of the coding sequence of TLR2 alleles, estimated by Kimura's two-parameter method (Kimura 1980). Difference in frequencies of non-synonymous and synonymous substitutions was evaluated by Fisher's exact test (Zhang et al. 1998). Further, codon-wise evaluation for positive selection was conducted by employing statistical test implemented in PAML 4 package (Yang 2007) based on maximum likelihood estimation method.

Homology modeling of molecular structure of *Macaca* TLR2-TLR1

The optimal structures of macaque TLR2-TLR1 variants were determined by means of homology modeling algorithm implemented in Molecular Operation Environment (MOE) 2010.10 (Chemical Computing Group Inc, http://www.chemcomp.com/) using a template structure of human TLR2-TLR1 heterodimer bound with Pam3CSK4 (PDB accession # 2Z7X). The PDB data contained structural data for human TLR2 (27–506) and human TLR1 (25–475) fused to 67 amino acid-long peptide derived from inshore hagfish *Eptatretus* VLRB.61. The structure of ternary complex of human TLR2 (27–506)-human TLR1 (25–475) dimer and a lipopeptide ligand Pam3CSK4 was optimized after trimming of C-terminal fusion peptides. Molecular stability was expressed by the sum of potential energy between atoms in kcal/mol. Amino acid sequence of macaque TLR1 (25–475) was retrieved from GenBank accession # NM_001130424. London dG score, an index for affinity between receptor and ligand within the binding pocket, was obtained for each optimized complex structure by Lig-X algorithm suite implemented in MOE 2010.10.

Results

Identification of SNPs in TLR2 coding sequence from *M. mulatta* and *M. fuscata*

In the present study, nucleotide sequence of the 2,355-bp long entire coding region of TLR2 encoded within the third exon was determined for of 21 *M. mulatta* and 35 *M. fuscata* individuals. By the comparison of *M. mulatta* sequences, six non-synonymous and three synonymous substitutions were identified, while three non-synonymous and five synonymous substitutions were found in *M. fuscata* (Table 1; Fig. 1). One synonymous substitution, GTG(Val)>GTT(Val) at codon 373 was common to these two species. Further, all 21 *M. mulatta* individuals were homozygous for AAT(Asn) while all *M. fuscata* examined were homozygous for TAT(Tyr) at codon 326. Collectively, ten non-synonymous and seven synonymous substitutions were found in this population.

Table 1 Haplotype of 17 SNPs found in 21 *Macaca mulatta* and 35 *Macaca fuscata* individuals in the present study

Haplotype	Nucleotide position																
	93 Codon	292	570	664	703[a]	976[a]	1119[a]	1176	1214	1246	1359[a]	1507	1633	1667	1743	1908	2344
	31	98	190	222	235[a]	326[a]	373[a]	392	405	416	453[a]	503	545	556	581	636	782
M. mulatta (frequency in 2*N*=42)																	
Mamu-Hap1 (17)	GAC (Asp)	TCC (Ser)	GAG (Glu)	AGT (Ser)	AAC (Asn)	AAT (Asn)	GTG (Val)	TTA (Leu)	ACC (Thr)	ACT (Thr)	GGC (Gly)	CTG (Leu)	ACT (Thr)	GCT (Ala)	TCA (Ser)	AAC (Asn)	ATA (Ile)
Mamu-Hap2 (10)	–	–	–	–	–	–	–	TTG (Leu)	–	–	–	GTG (Val)	–	–	–	–	–
Mamu-Hap3 (2)	–	–	–	–	–	–	–	–	–	GCT (Ala)	–	–	–	–	–	–	–
Mamu-Hap4 (2)	GAT (Asp)	–	–	–	–	–	–	TTG (Leu)	–	–	–	GTG (Val)	GCT (Ala)	–	–	–	–
Mamu-Hap5 (5)	–	–	–	–	–	–	–	–	–	–	–	GTG (Val)	–	–	–	–	–
Mamu-Hap6 (2)	–	–	–	–	–	–	–	–	–	–	–	GTG (Val)	–	GTT (Val)	–	–	–
Mamu-Hap7 (1)	–	–	–	–	–	–	GTT (Val)	–	ATC (Ile)	–	–	GTG (Val)	–	GTT (Val)	–	–	–
Mamu-Hap8 (1)	–	CCC (Pro)	–	–	–	–	–	–	–	–	–	GTG (Val)	–	–	–	–	–
Mamu-Hap9 (1)	–	–	–	–	–	–	–	–	–	–	–	GTG (Val)	GCT (Ala)	–	–	–	–
Mamu-Hap10 (1)	–	–	–	–	–	–	GTT (Val)	–	–	GCT (Ala)	–	–	–	–	–	–	–
M. fuscata (frequency in 2*N*=70)																	
Mafu-Hap1 (50)	GAC (Asp)	TCC (Ser)	GAG (Glu)	AGT (Ser)	GAC (Asp)	TAT (Tyr)	GTT (Val)	TTA (Leu)	ACC (Thr)	ACT (Thr)	GGT (Gly)	CTG (Leu)	ACT (Thr)	GCT (Ala)	TCA (Ser)	AAC (Asn)	ATA (Ile)
Mafu-Hap2 (4)	–	–	GAA (Glu)	–	AAC (Asn)	–	GTG (Val)	–	–	–	GGC (Gly)	–	–	–	–	–	–
Mafu-Hap3 (1)	–	–	–	–	AAC (Asn)	–	GTG (Val)	–	–	–	GGC (Gly)	–	–	–	–	–	–
Mafu-Hap4 (2)	–	–	–	GGT (Gly)	–	–	–	–	–	–	–	–	–	–	TCG (Ser)	–	–
Mafu-Hap5 (5)	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	AAT (Asn)	–
Mafu-Hap6 (7)	–	–	–	GGT (Gly)	–	–	–	–	–	–	–	–	–	–	–	–	–
Mafu-Hap7 (1)	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	TTA (Leu)
Human[b]	GAC (Asp)	TCC (Ser)	GAG (Glu)	AGT (Ser)	GAC (Asp)	TAT (Tyr)	GTT (Val)	TTA (Leu)	ACC (Thr)	ACT (Thr)	GGT (Gly)	GTG (Leu)	ACT (Thr)	ATT (Ile)	TCA (Ser)	AGC (Ser)	ATA (Ile)

Polymorphic nucleotide in each codon is underscored. Frequency of haplotype appeared in the population is shown in parenthesis

[a] The positions where the major haplotypes for *M. mulatta* (Mamu-Hap1) and *M. fuscata* (Mafu-Hap1) are different from each other

[b] Only the human sequence (GenBank accession #NM_003264) corresponding to polymorphic codons among *Macaca* species are shown to tell whether is more probable to be ancestral between *Macaca* alleles

The extracellular domain of TLR2 is mainly composed with the leucine-rich repeat (LRR) modules and can be divided into three sub-domains; N-terminal, central, and C-terminal sub-domains with their unique β-sheet conformations (Fig. 1). The structure of extracellular domains of triacylated lipopeptide-bound TLR2-TLR1 heterodimer and those of diacylated lipopeptide-bound TLR2-TLR6 heterodimer has been revealed by X-ray diffraction analysis (Jin et al. 2007; Kang et al. 2009). Amino acid residues which contribute to the binding of ligands and the interaction to dimerization partner were mapped in the middle of extracellular domain. The distribution of polymorphic sites in terms of structural sub-domains was shown in Fig. 1. It is of note that the abovementioned inter-species substitution at codon 326, Asn in *M. mulatta* versus Tyr in *M. fuscata*, is located within ligand binding and dimerization domain for TLR2-TLR1 heterodimer which may contribute to certain functional properties (Fig. 1b).

Upon the inference of the combination of SNPs by PHASE program implemented in DnaSP package (see below), a major allele of *M. mulatta* population (Mamu-Hap1 in Table 1) was obtained. Mamu-Hap1 had four nucleotide substitutions, GTT(Val)>GTG(Val) at codon 373, ATC (Ile)>ACC(Thr) at codon 405, GTG(Val)>CTG (Leu) at codon 503 and GTT(Val)>GCT(Ala) at codon 556, from the *M. mulatta TLR2* sequence reported previously, GenBank accession #AB445630 (Nakajima et al. 2008), which is identical to Mamu-Hap7 in the present study. The homology between *M. mulatta* and humans is 97.2% (2,289/2,355 bp) at nucleotide level and 95.8% (751/784) at amino acid level. Because general homology in nucleotide sequences between *M. mulatta* and humans is 93.94% identical in average (Rhesus Macaque Genome Sequencing and Analysis Consortium 2007), TLR2 is one of the well-conserved genes during primate evolution. A major allele of *M. fuscata* (Mafu-Hap1 in Table 1) was also inferred (GenBank accession, #AB607964) and had four nucleotide substitutions from Mamu-Hap1: AAC(Asn)>GAC(Asp) at codon 235, AAT(Asn)>TAT(Tyr) at codon 326, GTG(Val)>GTT(Val) at codon 373 and GGC(Gly)>GGT(Gly) at codon 453. Another less-frequent Mafu-Hap3 had only one substitution, AAT(Asn)>TAT(Tyr) at codon 326 from Mamu-Hap1. Figure 2 explains genealogy of allelic variants. Intra-specific variants are connected by as many as five and six mutation events in *M. mulatta* and *M. fuscata*, respectively, while Mamu-Hap1 and Mafu-Hap3 differ by one nucleotide substitution as mentioned above. In addition, a synonymous substitution at codon 373 was

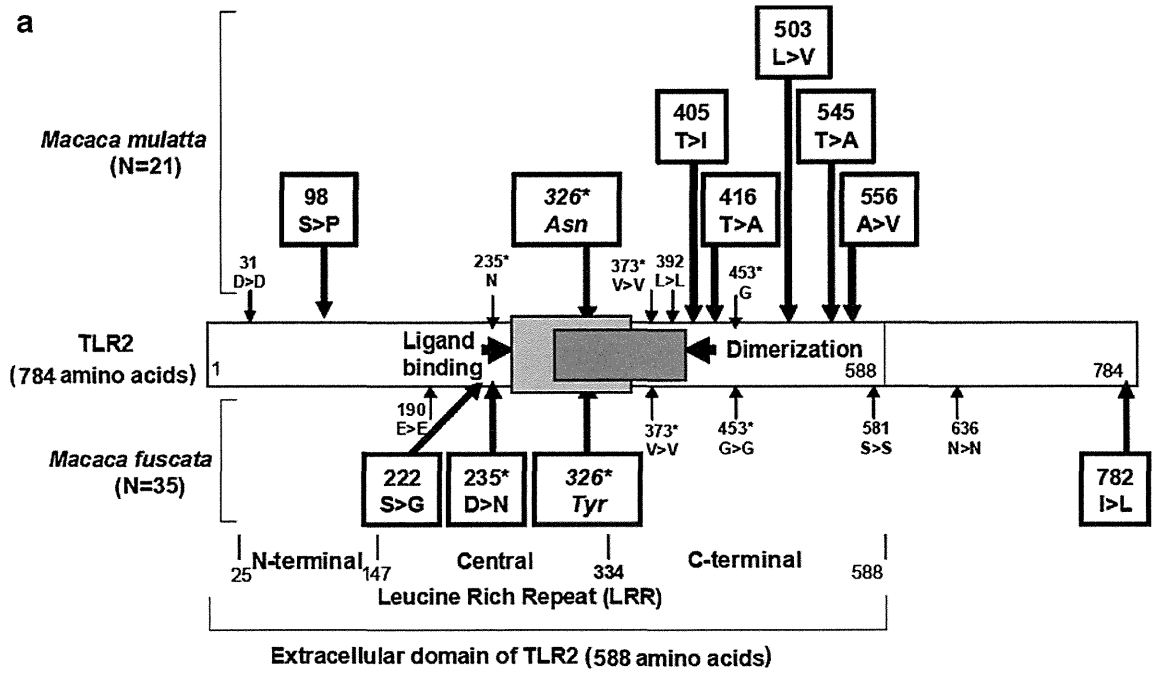

Fig. 1 The distribution of SNPs found in *M. mulatta* and *M. fuscata* populations. **a** The position of codons with synonymous (*plain*) and non-synonymous (*boxed*) SNPs is shown in schematic representation of TLR2; SNPs in *M. mulatta* and *M. fuscata* populations are the upper and the lower halves, respectively. Domain and sub-domain structures are also illustrated. The regions contributing to ligand binding (codon 266–359) and interaction with dimerization partner (codon 318–398 for TLR2-TLR1 and codon 318–404 for TLR2-TLR6) are depicted as overlapping *filled boxes* in the middle of extracellular domain (Jin et al. 2007; Kang et al. 2009). **b** Structure and amino acid sequence of TLR2. Leucine-rich domain (LRR)s at N terminus, #1 to #20 and at C terminus, and transmembrane and cytoplasmic regions (*TM+CY*) are shown above the sequence. Amino acid residues involved in ligand binding (*l*) and dimerization interface (*i*) are indicated. Sequence variations between human and *M. mulatta* (*asterisk*), among *M. mulatta* and among *M. fuscata* (var. site) are also indicated

b

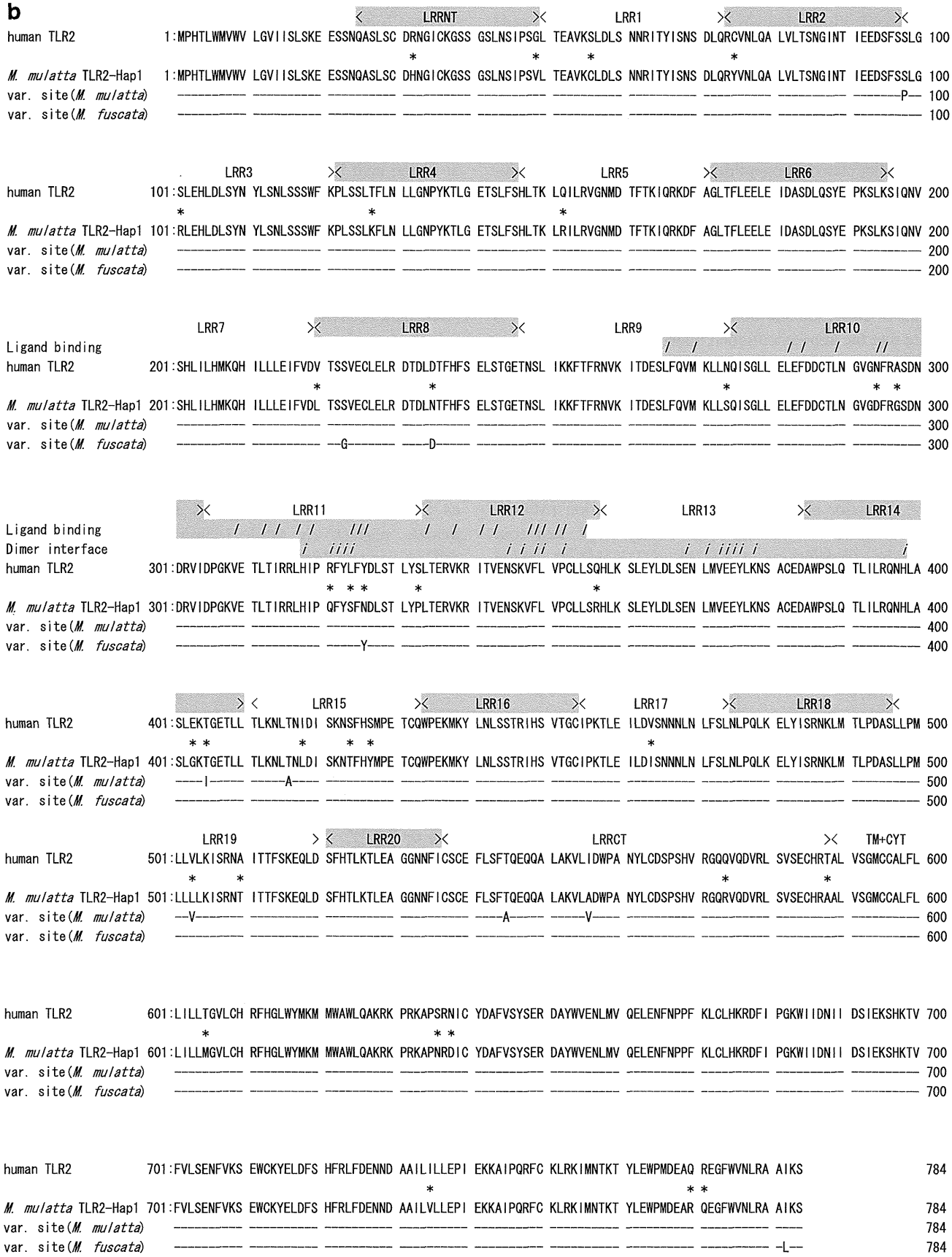

Fig. 1 (continued)

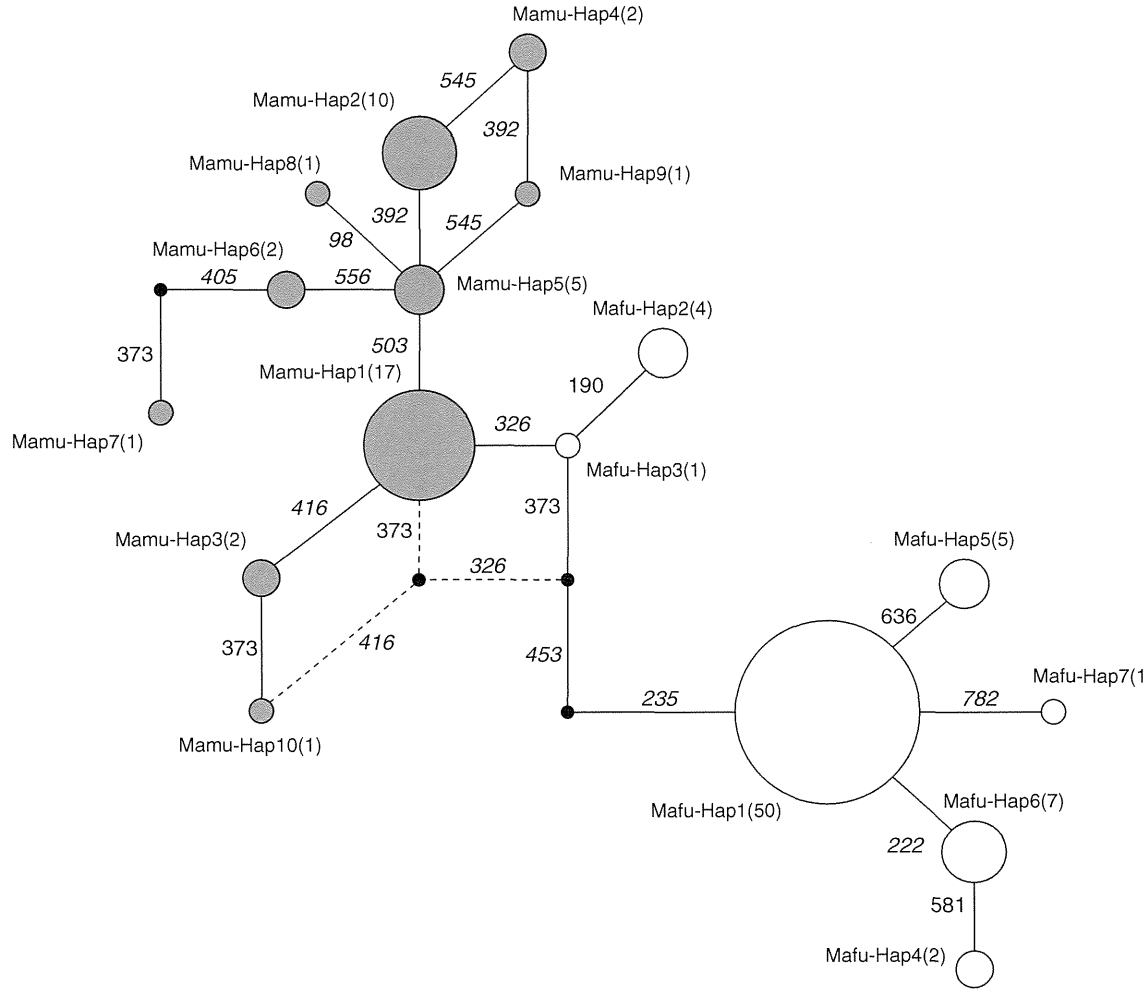

Fig. 2 Genealogical relationship of *Macaca* TLR2 alleles. A genealogical pathway of allelic variants is illustrated by network. *Gray* and *open circles* are allelic variants of *M. mulatta* and *M. fuscata* TLR2, respectively. Two variants differ by one nucleotide substitution are connected by *bar*. *Numbers near the bars* are codon positions; non-synonymous substitution is shown in *italic* style. Intra-specific variants are connected by as many as five and six mutation events in *M. mulatta* and *M. fuscata*, respectively, while Mamu-Hap1 and Mafu-Hap3 differ by one nucleotide substitution. It is of note that a synonymous substitution at codon 373 is commonly found in two species

commonly found in two species. Thus the difference between variants across the species is not more than the maximal difference within a species for certain combinations.

Molecular evolution of TLR2

In order to evaluate the allelic divergence of *Macaca TLR2* locus, we applied PHASE program to infer the haplotypes of 17 SNPs observed in two *Macaca* species. Consequently, ten haplotypes for *M. mulatta* and seven haplotypes for *M. fuscata* were identified (Table 1). Presence of these haplotypes was confirmed by nucleotide sequencing of plasmid clones. A phylogenetic tree (Fig. 3) was drawn by neighbor-joining method according to sequence alignments shown in Table 1. Because of amino acid substitution at codon 326, alleles of *M. mulatta* and *M. fuscata* were positioned in two separate lineages.

S, *H*, *Hd*, π, θ, average codon-based evolutionary divergence (d_S, d_N, and d_N, d_S), and Tajima's D were calculated for overall and sub-domains of *M. mulatta* and *M. fuscata* TLR2 (Table 2). Sliding window plots for Tajima's D value for each species and K_a/K_s ratio between two species are also presented (Figs. 4 and 5). When the statistical significance was examined by *Z* test, the entire TLR2 coding sequence of *M. fuscata* gave a significant result with a codon-based test (d_N, d_s) for purifying selection (p=0.0435), whereas became not significant if the extracellular domain was separately examined (p= 0.061). But further breaking down of the extracellular domain into three parts, C-terminal LRR domain underwent purifying selection by codon-based test (p=0.0406). In contrast, a part of the extracellular domain C-terminal to ligand binding and dimerization domain of *M. mulatta* TLR2 was rich in non-synonymous substitutions with a

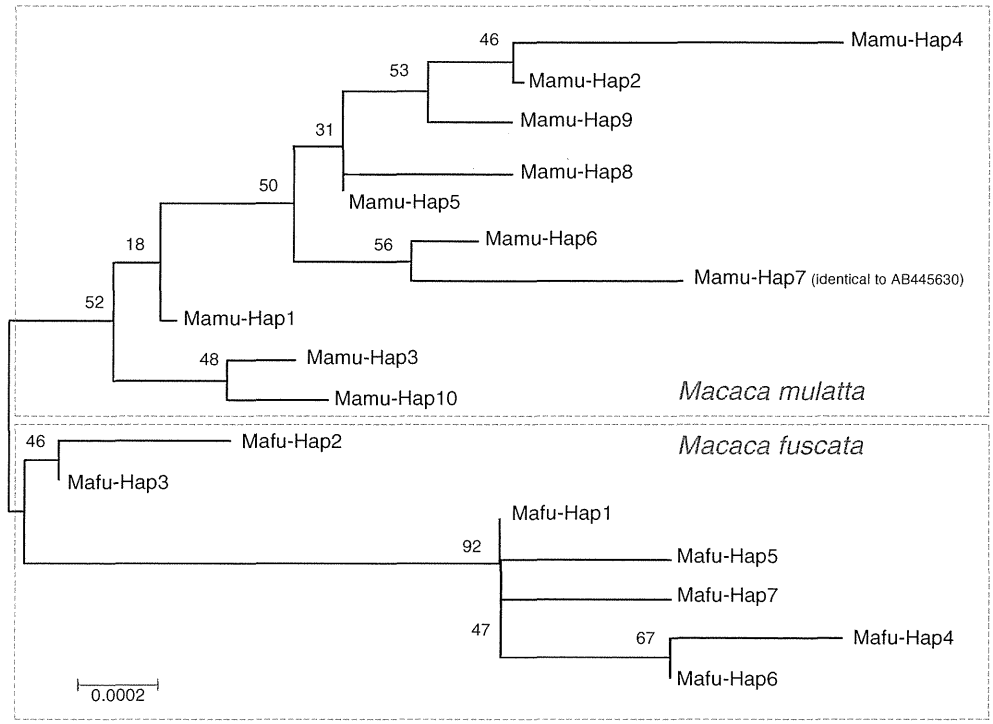

Fig. 3 Phylogenetic tree for *Macaca* TLR2 alleles. A phylogenetic tree is drawn by neighbor-joining method based on a distance matrix of Kimura's two-parameter evolutional model using MEGA4.0.2. Bootstrap percent probability for 500 replications is shown at each node. The *scale bar* represents a genetic distance of 0.0002

significant result for codon-based test for positive or diversifying selection (p=0.0362). Sliding window plot of Tajima's D for *M. mulatta* TLR2 (Fig. 4a) shows clearly different pattern from that of *M. fuscata* TLR2 (Fig. 4b): the N terminus of *M. mulatta* TLR2 and the membrane-proximal region of extracellular domain of *M. fuscata* TLR2 exhibit weak tendency of purifying selection because of the presence of synonymous substitutions, while the membrane-proximal region of extracellular domain of *M. mulatta* TLR2 shows positive D value suggesting the operation of positive selection. When all sequence of two species ($2N$=42 for ten *M. mulatta* variants and $2N$=70 for seven *M. fuscata* variants) were taken into account, average codon-based evolutionary divergence indicated by the value of K_a/K_s ratio is highest in the part of central LRR, mostly attributed to amino acid substitution at codon 326 (Table 3) while two peaks of K_a/K_s ratio corresponding to the central LRR and the amino acid substitutions in the membrane-proximal part of extracellular domain of *M. mulatta* TLR2 appear when the sliding window is narrowed to 350-bp range (Fig. 5). When the frequency of non-synonymous substitution in possible sites (five in 387) within in the membrane-proximal part of extracellular domain (codon 401 to 566) of *M. mulatta* TLR2 was compared with that of synonymous substitution (0 in 111), they were not significantly different (p=0.358; Table 4, comparison A), but the number of non-synonymous substitutions is over-represented in *M. mulatta* than *M. fuscata* (5 vs. 0) in comparison to that of synonymous substitutions (0 vs.2) within the same region (p=0.048; Table 4, comparison B), suggesting the presence of differential selection pressure on these two species. Further, codon-wise evaluation for non-synonymous and synonymous substitutions was conducted for extracellular domain of *M. mulatta* TLR2 by employing statistical test implemented in PAML 4 package based on maximum likelihood estimation method. Likelihood statistics under different molecular evolution models were presented in Table 5. By comparison of likelihoods between neutral and selection site models, we could obtain a weak but significant signal for positive selection at codon position 545. This position is not known to be involved in ligand binding or dimerization interface in the diacylated lipoprotein-bound structure of TLR2-TLR6 dimer or the triacylated lipoprotein-bound structure of TLR2-TLR1 dimer (Fig. 1b).

Functional analysis for TLR2 alleles

In order to clarify whether amino acid substitutions in *Macaca* TLR2 alleles can bring any functional alteration, an experimental system was designed to measure the downstream transcriptional activation of NF-κB-dependent synthetic enhancer-promoter upon the stimulation with TLR2 agonists. It is of importance to examine functional integrity of variants to rule out pseudogenization of TLR-2 in *Macaca* lineage as one of possible cause of rapid accumulation of non-synonymous mutations. We chose HEK293 cells as recipients of DNA transfection assay because it has been proven to be lower intrinsic TLR2

Table 2 Polymorphism and nucleotide diversity in the coding sequence of TLR2 locus

Domain[a]	Range (bp)	*S*	*H*	*Hd*	π ($\times 10^{-4}$)	θ ($\times 10^{-4}$)	d_S (±SE)	d_N (±SE)	d_N-d_S (±SE)	Tajima's D
Macaca mulatta (n=21)										
Total	1–2,355	9	10	0.775	6.9	8.9	0.0011±0.0008	0.0006±0.0003	−0.0006±0.0008	−0.65546
Extracellular	1–1,764	9	10	0.775	9.2	11.9	0.0015±0.0010	0.0007±0.0004	−0.0008±0.0011	−0.65546
N-terminal LRR	73–438	2	3	0.138	3.8	12.7	0.0010±0.0010	0.0002±0.0002	−0.0009±0.0011	−1.30048
Central LRR	439–999	0	1	0.000	0.0	n.d.	0.0000±0.0000	0.0000±0.0000	0.0000±0.0000	n.d.
C-terminal LRR	1,000–1,764	7	9	0.769	19.3	21.3	0.0029±0.0024	0.0016±0.0009	−0.0013±0.0027	−0.25176
N-terminal part	1–795	2	3	0.138	1.8	5.8	0.0005±0.0005	0.0001±0.0001	−0.0004±0.0005	−1.30048
Lig-bind/dimer	796–1,194	2	3	0.483	12.8	11.6	0.0057±0.0048	0.0000±0.0000	−0.0057±0.0046	0.18543
C-terminal part	1,195–1,764	5	6	0.695	17.0	20.4	0.0000±0.0000	0.0022±0.0012	0.0022±0.0012*	−0.42381
TM-Cyt	1,765–2,355	0	1	0.000	0.0	n.d.	0.0000±0.0000	0.0000±0.0000	0.0000±0.0000	n.d.
Macaca fuscata (n=35)										
Total	1–2,355	8	7	0.477	4.1	7.0	0.0011±0.0005	0.0002±0.0001	−0.0009±0.0005**	−1.08369
Extracellular	1–1,764	6	5	0.351	4.5	7.1	0.0011±0.0005	0.0003±0.0002	−0.0008±0.0006	−0.85999
N-terminal LRR	73–438	0	1	0.000	0.0	n.d.	00000±0.0000	0.0000±0.0000	0.0000±0.0000	n.d.
Central LRR	439–999	3	4	0.345	8.4	11.1	0.0009±0.0009	0.0008±0.0006	0.0000±0.0010	−0.46715
C-terminal LRR	1,000–1,764	3	3	0.187	4.3	8.1	0.0019±0.0011	0.0000±0.0000	−0.0019±0.0011***	−0.91696
N-terminal part	1–795	3	4	0.345	5.9	7.8	0.0006±0.0006	0.0006±0.0004	0.0000±0.0007	−0.46715
Lig-bind/dimer	796–1,194	1	2	0.135	3.4	5.2	0.0015±0.0015	0.0000±0.0000	−0.0015±0.0015	−0.43152
C-terminal part	1,195–1,764	2	3	0.187	3.3	7.3	0.0015±0.0011	0.0000±0.0000	−0.0015±0.0012	−0.89232
TM-Cyt	1,765–2,355	2	3	0.161	2.8	7.0	0.0011±0.0010	0.0001±0.0001	−0.0010±0.0011	−1.00276

S number of variable sites, *H* number of haplotypes, *Hd* haplotype diversity, π nucleotide diversity (per site), θ nucleotide polymorphism (per site), d_S number of synonymous substitutions per site, d_N number of non-synonymous substitutions per site, d_N-d_S difference between the non-synonymous and synonymous distances per site; Tajima's D: Tajima's D statistic

*p=0.0362 upon Z test for positive (or diversifying) selection; *p=0.0435; *p=0.0406 upon Z test for purifying selection

[a] 2,355-bp of total open reading frame for TLR2 is divided to extracellular domain and the rest (TM-Cyt: transmembrane and cytoplasmic). Extracellular domain is further broken down to three sub-domains in two ways: one according to the differential motifs of leucine-rich repeats (LRR) and the other according to structural analysis of TLR2-TLR1 heterodimeric complex (Jin et al. 2007) delineating the region contributes to ligand binding and interaction to dimerization partner (Lig-bind/Dimer)

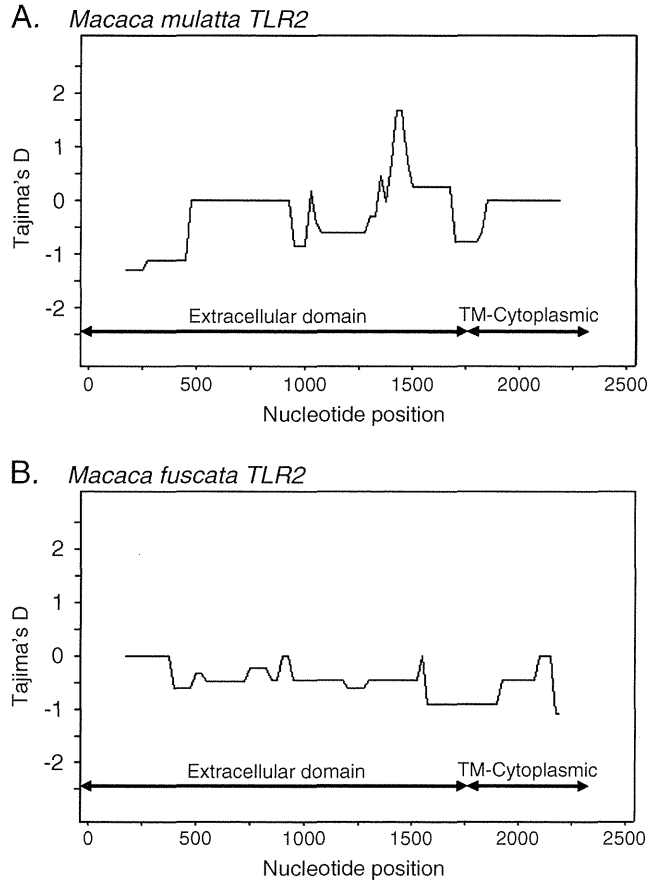

Fig. 4 Distribution of polymorphisms evaluated by Tajima's D statistic. Sliding window plots of Tajima's D with 350 bp window and 25-bp steps are drawn for *M. mulatta* (**a**) and *M. fuscata* (**b**) *TLR2* using DnaSP ver5.10.00. Although Tajima's D does not reach to significant levels, it tends to be below zero in general

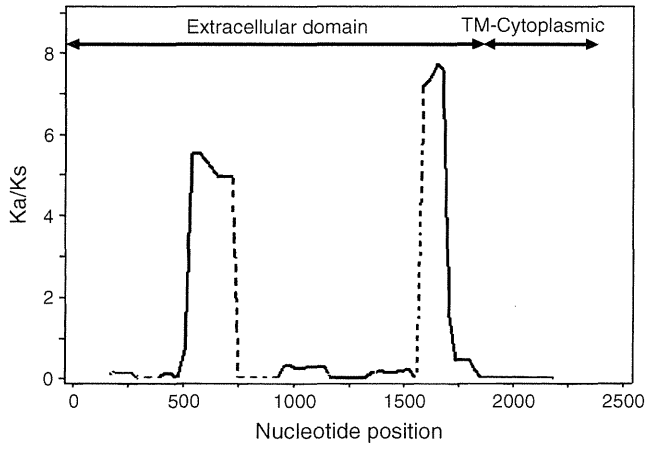

Fig. 5 Distribution of non-synonymous substitutions in two *Macaca* species. A sliding window plot of K_a/K_s ratio between *M. mulatta* and *M. fuscata* alleles with 350 bp window and 30-bp steps is drawn using DnaSP ver5.10.00. *Dotted lines* show the parts where the K_a/K_s ratio is not obtained because of the absence of synonymous substitutions

Table 3 Estimation of rate of synonymous and non-synonymous substitutions between *Macaca mulatta* and *Macaca fuscata*

Domain[a]	Range (bp)	K_s	K_a	K_a/K_s
Total	1–2,355	0.00435	0.00157	0.359
Extracellular	1–1,764	0.00553	0.00209	0.377
N-terminal LRR	73–438	0.00053	0.00009	0.164
Central LRR	439–999	0.00046	0.00471	10.312
C-terminal LRR	1,000–1,764	0.01221	0.00129	0.105
TM-Cyt	1,765–2,355	0.00057	0.00003	0.054

K_s synonymous nucleotide divergence, K_a non-synonymous nucleotide divergence, K_a/K_s the ratio of K_a to K_s

[a] See footnote of Table 2

activity and to possess the other components necessary for TLR2-mediated, ligand-induced transcriptional activation (Schwandner et al. 1999). When the expression plasmid for a major allele of *M. mulatta* TLR2, Mamu-Hap1 was introduced to HEK293 cells, two synthetic TLR2 ligands, Pam2CSK4 and Pam3CSK4, which are relatively specific to TLR2-TLR6 and TLR2-TLR1 heterodimers, respectively, could induce very strong luciferase activity to almost equivalent extent to human TLR2 expression plasmid (Fig. 6a). Therefore, the function of *M. mulatta* TLR2 in combination with HEK293-intrinsic TLR6 and TLR1 remains intact in the presence of mutations taken place after the species diversification, indicating the higher rate of amino acid substitutions observed in *M. mulatta* lineage are not attributable to pseudogenization. Further, the effect of amino acid substitution at codon 326 was evaluated by the comparison between Mamu-Hap1 and Mafu-Hap3, obtaining no evidence for functional relevance to the substitution by this experimental system (Fig. 6b). Further, the effect of all six non-synonymous substitutions found in *M. mulatta* population was examined by using four additional expression plasmids: Mamu-Hap3 carrying Ala416, Mamu-Hap7 carrying Ile405-Val503-Val556, Mamu-Hap8 carrying Pro98-Val503, and Mamu-Hap9 carrying Val503-Ala545 (Fig. 6c). Although Pam3CSK4-induced luciferase activity was higher for Mamu-Hap8 than Mamu-Hap7 and Mamu-Hap9, statistic obtained by ANOVA,. by which the inflation of statistical error by multiplicity of comparison is taken into account, did not reach to significance. Thus, we could not demonstrate any functional augmentation or deterioration accompanied with seven non-synonymous amino acid substitutions identified in the present study.

Structural analysis by computational chemistry

The non-synonymous substitution at codon position 326 appeared to be differentially fixed in each species, asparagine for *M. mulatta* whereas tyrosine for *M. fuscata*.