原　著　麻布大学雑誌　第 31 巻　1 － 10

Single Nucleotide Polymorphisms and Conserved Transcription Factor Binding Site of *CD14* Gene in Yorkshire Terrier

Haruka SUGINO[1], Shinpei KAWARAI[1], Yukihiro FUJITA[2], Miyoko SAITO[2], Hideki KAYANUMA[3], Yoshihiro FUJIMORI[5], Takuya MARUO[3], Hiroo MADARAME[1], Masaru MURAKAMI[4] and Nobuyuki KANEMAKI[1]

[1] Laboratory of Small Animal Clinics, Veterinary Teaching Hospital and
[2] Laboratory of Small Animal Surgery,
[3] Laboratory of Veterinary Radiology and
[4] Laboratory of Molecular Biology, Department of Veterinary Medicine, School of Veterinary Medicine, Azabu University, 1-17-71 Fuchinobe, Chuo-ku, Sagamihara, Kanagawa 252-5201, Japan
[5] Fujimori Animal General Hospital, 56-2 Miyanose, Minoyama, Yahata-shi, Kyoto 614-8277, Japan

Abstract: In humans, single nucleotide polymorphisms (SNPs) of *CD14* promoter region are associated with allergic diseases. There are no reports of canine *CD14* gene cloning, nor any reports studying the pathogenesis of *CD14* polymorphism in dogs. In this study, canine *CD14* gene was cloned by 5' and 3' RACE method and its SNPs were analyzed. The presence of two SNPs were identified. SNP1 was located at predicted AP-1 binding site of *CD14* promoter region (-187C>T) in the Yorkshire terrier. SNP2 (c.930G>A), silent mutation of Glu, identified in 5 dog breeds. The relative transcriptional activity of a general breed type *CD14* promoter (-187C/luc) did not differ from those of Yorkshire terrier type (-187T/luc); however, the relative activity was diminished when the AP-1 site was mutated. In conclusion, *CD14* gene structure was highly conserved in dog breeds and its expression could be regulated by the transcription factor AP-1.

Keywords: Canine atopic dermatitis, Cloning, Luciferase reporter assay, Transcription factor

1. Introduction

The gene encoding CD14 was first sequenced in humans in 1988 [1]. The structure of CD14 was revealed as a bent solenoid, which is typical of leucine-rich repeat (LRR) proteins with an N-terminal pocket that binds lipopolysaccharide (LPS) [2]. CD14 exists in two forms: a membrane form and a soluble protein form. Membrane CD14 is a glycosylphosphatidylinositol (GPI)-anchored glycoprotein expressed on monocytes/macrophages and granulocytes [3]. The recognition of LPS by CD14 is catalyzed by LPS-binding protein, and it is then shuttled to the toll-like receptor 4 and myeloid differentiation-2 complex [4]. The signal transduction induces the differentiation of progenitor cells to monocytes, macrophages, and/or dendritic cells [5]. A previous study has demonstrated that CD14 plays a role in pattern recognition receptors for innate immune responses [6]. According to the hygiene hypothesis stating that a lack of exposure to LPS in early childhood increases susceptibility to allergic diseases, the expression of CD14, the receptor of LPS associates with disease development

Corresponding author: Shinpei Kawarai (e-mail: kawarai@azabu-u.ac.jp)

COI: There are no conflicts of interest to declare.

and adaptive immunity [7]. In humans, a single nucleotide polymorphism (SNP) of the *CD14* promoter at position −159 from the transcription start site (TSS) is associated with concentrations of serum soluble CD14 and total serum IgE [8] are reportedly correlated. Moreover, this SNP is associated with Kawasaki disease [9], while an SNP at position −260 is associated with myocardial infarction [10]. In dogs, *CD14* polymorphisms and its relation has not been reported yet. Here we examined the full length sequence of the canine *CD14* gene and identified two SNPs within 16 dog breeds. We focused the one of SNPs at the position prior to TSS in the dog *CD14* gene, which was identified particularly Yorkshire terrier within 16 dog breeds.

2. Material and Methods

2.1. Dogs

This investigation was approved by The Azabu University Ethics committee (approval 170621-7). For cloning, a total of 26 samples of frozen-stored clot and serum or whole blood from 16 dog breeds were obtained from the veterinary teaching hospital at Azabu University (Table S1). And for genotyping, a total of 62 samples of frozen-stored clot and serum or whole blood from Yorkshire terriers were obtained from primary animal hospitals and the veterinary teaching hospital at Azabu University. The frozen-stored samples left after laboratory examinations were used in this study. The whole blood samples were obtained after informed consent to pet owners. For case-control analysis, 27 Yorkshire Terriers (age, 0–18 years) were divided into groups with or without canine atopic dermatitis (CAD), which was diagnosed if the dog satisfied more than five criteria of set [11] (Table 1).

2.2. DNA extraction and sequencing

Genomic DNA was extracted using QuickGene DNA whole blood kit S (Kurabo Industries LTD., Tokyo, Japan). Total RNA was extracted from peripheral blood mononuclear cells using an RNeasy Plus Mini Kit (Qiagen K.K., Tokyo, Japan) and reverse-transcribed to first strand cDNA using a Prime Script 1st Strand cDNA Synthesis Kit (Takara Bio Inc., Shiga, Japan). Genomic DNA (200 ng) and cDNA (1 μL) were amplified by polymerase chain reaction (PCR) using the Ex Taq Hot-Start Version (Takara Bio Inc.). The primers were constructed on the basis of a predicted sequence in the NCBI database (accession number XM_843653) (Table S2). Ethidium bromide electrophoresis was used to visualize the PCR products. The 3′ and 5′ end sequences of RNA transcripts were isolated by the rapid amplification of cDNA ends (RACE) method using the 3′-Full and 5′-Full RACE Core Set (Takara Bio Inc.), respectively. PCR products were sequenced at a commercial laboratory (Fasmac CO., Ltd., Kanagawa, Japan), and the DNA sequence was analyzed using MEGA6 software. The potential signal peptide cleavage site was searched using SignalP4.0. N-glycosylation site prediction was performed using NetNGlyc1.0 software, and a GPI-anchor site was predicted using the PredGPI prediction server. The predicted transcription factors binding sites (TFBS) in the promoter sequence were searched by Match-1.0 Public using weight matrices from TRANSFAC Public 6.0. The URL for the *in silico* database is available in Table S3.

Table 1 Genotype and clinical condition of 27 Yorkshire Terriers enrolled in case-control association study

Diagnosis of canine atopic dermatitis	Genotypes	Number of dogs	Age (years)*	Total scores of Favrot's set 1 criteria (Favrot *et al.*, 2010)*
Case	TT	6	9.5 ± 2.7	6.8 ± 0.9
Case	TC	3	9.0 ± 2.8	6.7 ± 1.2
Case	CC	0	ND**	ND**
Control	TT	13	7.6 ± 5.0	2.8 ± 0.9
Control	TC	4	4.3 ± 4.1	2.3 ± 1.3
Control	CC	1	1	1

* Mean ± standard deviation, ** Not detected.

2.3. Plasmid and luciferase assays

Plasmid construction and luciferase assays were performed as previously described [12]. A DNA sequence spanning nt –373 to nt +73 (nt +1 was defined as the TSS) of the sequenced dog *CD14* promoter was amplified and cloned into a pGL4 basic vector containing the firefly luciferase reporter (Promega, Tokyo, Japan). Reporter plasmids with the general breed type SNP (TGA<u>C</u>TCA) and Yorkshire Terrier type SNP (TGA<u>T</u>TCA) and the mutated control sequence (TGA<u>TTTT</u>) at the predicted TFBS were prepared using a PCR-based method and verified by DNA sequencing (as a template).

The human acute monocytic leukemia cell line THP-1 (DS Pharma Biomedical, Osaka, Japan) was cultivated and maintained in an RPMI-1640 medium containing 10% fetal calf serum, penicillin (100 U/mL), and streptomycin (0.1 mg/mL) at 37°C in a 5%CO_2 incubator. THP-1 cells (1.0×10^6) were transfected with the constructed reporter plasmids and a plasmid expressing Renilla luciferase under the control of the thymidine kinase promoter (pRL-tk) using polyethylenimine (PEI) “Max (Polysciences, Warrington, PA). At 48-h post-transfection, cells (2.0×10^5) were harvested onto 48-well culture plates and incubated with 100 ng/mL 12-*O*-tetradecanoylphorbol-13-acetate (TPA) (Sigma-Aldrich Japan, Tokyo, Japan) for 48 h. Firefly luciferase activity was normalized to the Renilla luciferase activity. The resulting data are reported as the mean value of triplicate wells.

2.4. Statistical analysis

Chi-squared test was used to assess whether the genotype frequencies were in Hardy–Weinberg equilibrium (HWE) and for the case-control analysis of CAD. Differences in the luciferase activity between the transfected cells were examined using paired *t*-tests. Differences of $p < 0.05$ were considered statistically significant.

3. Results

3.1. Sequence of CD14 gene

The nucleotide and deduced amino acid sequence of dog CD14, which has registered NCBI database (LC147081), is shown in Figs. 1 and 2. The dog *CD14* gene comprises two exons and a short intron (80 bp) located next to the ATG translation start codon. The ATG initiation codon was identified and is followed by an open reading frame at +1119 and a TGA stop codon at +1341. The open reading frame was flanked by 132-bp 5′ untranslated sequence and 545-bp 3′ untranslated sequence. Four consensus polyadenylation signals were located at +513–534 of the termination codon, the latter of which was followed 9 bp downstream by a polyadenylate tail. As deduced from cDNA, the primary structure of dog CD14 protein comprises 373 amino acids. The alignment of the amino acid sequence of dog CD14 with the human, cow, pig, mouse, and rabbit sequences revealed 71%, 75%, 73%, 65%, and 72% amino acid identity, respectively. The first methionine in canine CD14 is followed by a stretch of nearly hydrophobic and/or neutral residues, which is typical of a eukaryotic signal peptide, followed by a potential signal peptide cleavage site between positions 17 and 18. Canine CD14 contains 12 cysteine residues, 8 of which are conserved in the human, cow, pig, mouse, and rabbit sequences, suggesting that these might be important in determining the conformation of dog CD14 through the formation of specific disulfide bridges. Eleven LRRs representing the LXXLXLX motif were observed, consistent with previous reports [2]. There were five potential N-linked glycosylation sites (Asn-X-Thr/Ser), three of which are conserved in the human, cow, pig, mouse, and rabbit sequences. The GPI-anchor site was unique as the dog sequence contained threonine. The signal peptide, motif, and GPI-anchor sites demonstrated that the structure of dog CD14 protein is very similar to that of human, cow, pig, mouse, and rabbit CD14 proteins.

3.2. SNPs of CD14

The transcription factor activating protein-1 (AP-1), CHOP-C/EBPalpha, and STATx binding site were identified at the position prior to TSS in the dog *CD14* gene by Match-1.0 Public (Fig. 3). Next, we analyzed the full length of the *CD14* gene sequence from 16 dog breeds, and identified two SNPs (Fig. 4). SNP1 (–187C>T) was found in the four of four Yorkshire Terriers and was located in AP-1 binding sequence (Fig. 3). SNP2 (c.930G>A), a silent

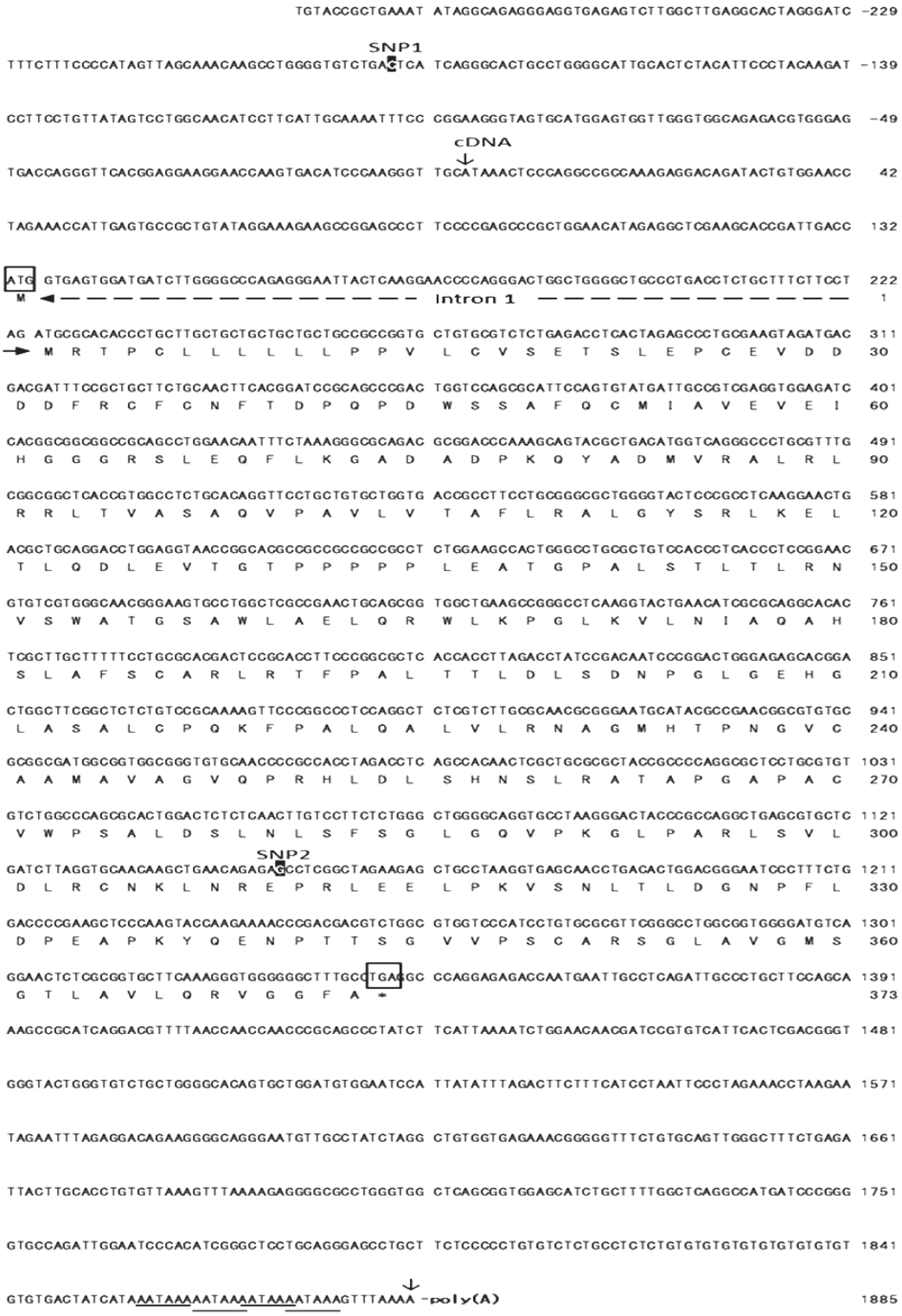

Fig. 1 The nucleotide and deduced amino acid sequences of dog CD14.

Numbers on the right of each row represent the nucleotide and amino acid position, respectively. Arrowheads above 1bp and 1885bp indicate the first and last bases of the dog *CD14* cDNA, respectively. Sequences in the boxes represent the start codon (ATG) and termination codon (TGA), respectively. Eukaryotic polyadenylation signals AATAAA are underlined. The identified SNPs are shown in darkened boxes.

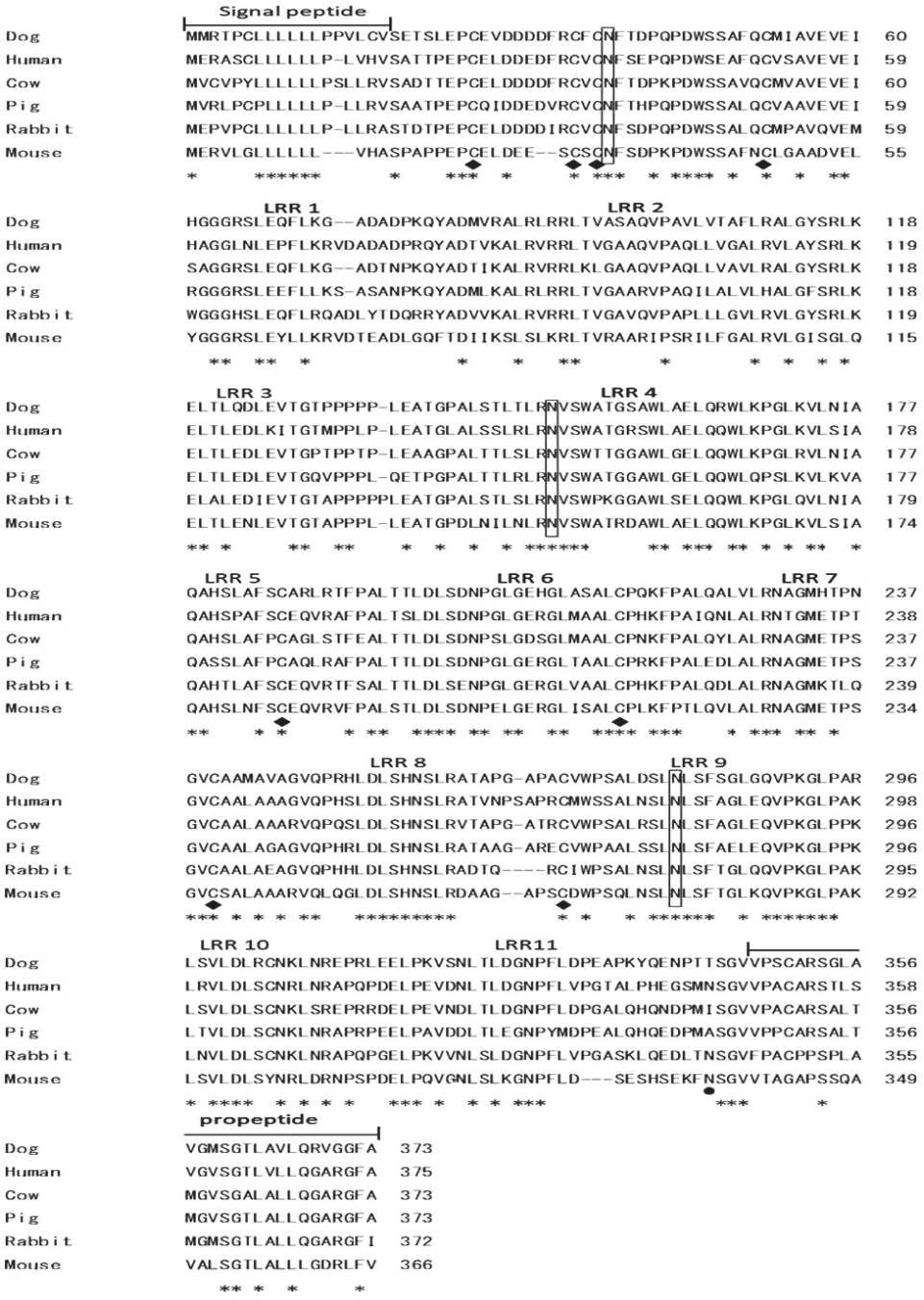

Fig. 2 Alignment of CD14 amino acid sequences of dog, human, cow, pig, rabbit and mouse.

Numbers on the right of each row represent amino acid position. The predicted signal peptide, LRRs, and propeptide are labeled above the alignment according to Kelley *et al.* (2013). Predicted N-glycosylation sites are in boxes. Below the alignment, the GPI-anchor amidated asparagine is indicated by "●", and the cysteine residues are indicated by "◆". Asterisks denote the amino acids that are entirely identical among all species. The amino acid sequences of dog (BAX01969), human (NP_000582), cow (BAA21517), pig (ABC84487), rabbit (NP_001075664), mouse (NP_033971) were obtained from the NCBI database.

mutation of Glu (Fig. 4), was confirmed in the two of three Shiba dogs, the one of three American Cocker Spaniels, and in one Maltese, Pekingese and Labrador Retriever (all of them were heterozygous mutants). As a result of examining the frequency of SNP 1 in 62 Yorkshire Terriers, 41 Yorkshire terriers were TT, 12 were TC and 9 were CC. In

Fig. 3 Predicted TFBS of dog *CD14* gene by Match 1.0 Public.

Rightward arrows indicate plus strand and leftward arrows indicate minus strand, respectively. Numbers on the right and left of each row represent nucleotide position of Figure 1A. Top sequence is a general breed type (C allele), and bottom sequence is a Yorkshire terrier type (T allele). Polymorphism positions are shown by darkend box.

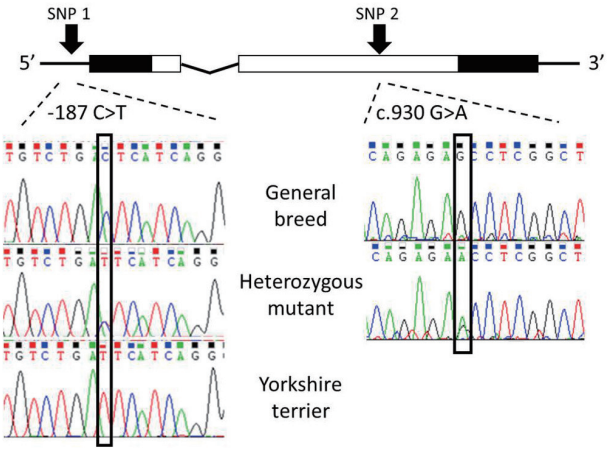

Fig. 4 Schematic representation of SNP1 (–187 C>T) and SNP 2 (c.930 G>A) located at CD14 structure.

Yorkshire terrier, the TT, TC, and CC genotype frequencies for SNP1 were 0.672, 0.172, and 0.155, respectively. The T and C allelic frequencies were 0.759 and 0.241, respectively. The genotype frequencies indicated a significant departure from HWE ($p < 0.01$; $x2 = 16.24 > 6.63$). Yorkshire Terriers which visited the veterinary teaching hospital at Azabu University each have some sort of disease, and since there is a possibility that it might have been biased, we excepted those and examined. As a result, genotype frequencies were 0.704, 0.259, and 0.037, respectively. The T and C allelic frequencies were 0.833 and 0.167, respectively. The genotype frequencies indicated consistent with HWE.

In humans, *CD14* polymorphisms in the promoter region are associated with allergic diseases [8]. To assess whether SNP1 was associated with the pathogenesis of CAD in Yorkshire Terriers, a case-control analysis was performed. The numbers of CC, TC, and TT dogs in the CAD group were 0, 4, and 6, respectively, and in the control group the numbers were 1, 3, and 13. The genotype frequencies for SNP1 did not differ significantly between these two groups ($p = 0.67$).

3.3. Transcriptional activity of AP-1 binding sequence

The C to T substitution of SNP1 in Yorkshire Terriers did not show an AP-1 binding site in Match-1.0 Public

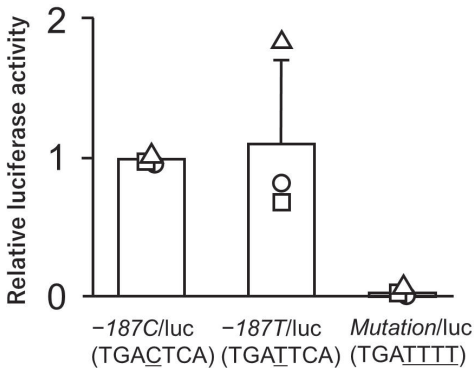

Fig. 5 The comparison of luciferase activities between general breed type CD14 promoter (*−187C*/luc) and Yorkshire Terrier type (*−187T*/luc).

THP-1 cells were transfected with the constructed reporter plasmids of the general breed type SNP (TGACTCA), Yorkshire Terrier type SNP (TGATTCA), or the mutated control sequence (TGATTTT) at the predicted AP-1 binding site and the plasmid expressing Renilla luciferase under the control of the pRL-tk using PEI "Max". At 48-h post-transfection, cells were incubated with TPA (100 ng/mL) for 48 h. Firefly luciferase activity was normalized to the Renilla luciferase activity. Data represent the mean value of triplicate wells. The three independent experiments were performed and each experiment was shown as different shapes.

(Fig. 3). Thus, to determine whether SNP1 modulated CD14 transcriptional activity, the luciferase activities of the general breed type CD14 promoter (*−187C*/luc) and Yorkshire Terrier type (*−187T*/luc) were compared. The relative activity of *−187T*/luc was not significantly different from that of *−187C*/luc (mean ± SD = 1.11 ± 0.62, $n = 3$). The relative activity of the mutated negative control sequence could not be detected (Fig. 5).

4. Discussion

To the best of our knowledge, the full length of complete sequence and its polymorphisms have not been previously elucidated in any dog breeds. Similar to other mammals, almost 70% of the amino acid sequences and characteristic structure of the disulfide bridges, LRR, N-linked glycosylation sites, and GPI-anchor sites were conserved in dog CD14 (Fig. 2). These results indicate that dog CD14 can recognize LPS and has similar functions in innate immune responses.

In addition, we revealed two SNPs in the dog *CD14* gene, namely SNP1 located at the AP-1 binding site and SNP2 at the second exon. Because the silent mutation in SNP2 does not affect gene regulation, we focused on the correlation between SNP1 allele frequency and the prevalence of CAD, as reported in humans [8]. The genotype frequency of SNP1 was not associated with the prevalence of CAD. In humans, correlations of the CD14 promoter at position −159 from the TSS and concentrations of serum soluble CD14 and total serum IgE are shown [8]. In the present study, we could not examine the association of SNP1 with severity of CAD because we didn't evaluate degree of the severity and sensitization such as clinical scores, allergen-specific serum IgE test, and intradermal skin test in the healthy controls. Thus, further research is warranted to examine SNP1 in the pathogenesis of CAD.

In the dog *CD14* gene, the AP-1 binding site was located at position −190 from the TSS (Fig. 1). This location is similar to that reported in the human (−220), rat (−270), and mouse (−255) *CD14* genes [13-15]. THP-1 cells were used for the luciferase assay of the dog *CD14* promotor containing SNP1. AP-1 is known to bind the TPA response element (TRE), 5′-TGACTCA-3′ and induce the following transcript in response to TPA. THP-1 cells can be induced by TPA to differentiate into macrophage-like cells and express the scavenger receptor gene, the transcript of which is regulated by AP-1 [16].

In humans, the SNP at position −159 in the *CD14* promotor is associated with increased disease risk; this region binds to the transcription factor Sp family. The C to T substitution at position −159 in *CD14* causes a slight increase (32%) in luciferase activity in human monocytic Mono Mac 6 cells compared with that in the wild type [17]. Conversely, a significant difference in the relative luciferase activity between the general breed type CD14 promoter (*−187C*/luc) and Yorkshire Terrier type (*−187T*/luc) was not observed in the present study. AP-1, which comprises the constituent proteins Jun and Fos, binds to the TGACTCA consensus motif.

In previous work, the effect of all possible SNPs within the TGACTCA motif on AP-1 and DNA interaction was evaluated *in vitro* isothermal titration calorimetry using synthetic double-stranded DNA oligos and *E. coli* expressed

proteins [18]. Single nucleotide substitutions within the TGACTCA motif did not abrogate the binding of the Jun–Fos heterodimer, and the relative binding affinities to the Jun–Fos heterodimer were not markedly different between the TGACTCA and TGATTCA variants [18]. The AP-1 binding site of the TGATTCA motif has also been reported in rodent *CD14* promoters [13,14]. Our data (Fig. 5) may indicate that both SNP1 alleles result in functional TRE in dogs because the TRE mutation has completely lost the basal transcription of *CD14*.

Limitation of the present study is we could not show the direct evidence of canine AP-1 activation on canine CD14 promoter. Our results suggested SNP1 could bind to human AP-1, but we could not examine to canine AP-1. In the preliminary experiments, canine hepatocellular carcinoma cell line (CHKS) [19] failed to transfect with the constructed reporter plasmids and a plasmid expressing Renilla luciferase (data not shown). Because AP-1 binding activity of rat apical sodium-dependent bile acid transporter was examined in human and canine cell lines [20] and the presence of abundant AP-1 activity was suggested in THP-1 cells [16], we used THP-1 cells in the present study. If the binding affinity to TGATTCA of the Jun–Fos heterodimer were different between humans and dogs, SNP1 might modulate the binding ability to canine AP-1. Thus, further researches are required to examine canine AP-1 activation on canine CD14 promoter by transcription factor binding assay such as electrophoretic mobility shift assay, as reported in humans [17].

Here, we cloned the full length of the dog *CD14* gene for the first time and examined polymorphisms among different breeds. Similar to other mammals, the amino acid sequence and the characteristic structures were conserved in dog *CD14*. Two SNPs were identified. SNP1 (−187C>T) located at a transcription factor AP-1 binding prediction site in Yorkshire Terriers. SNP2 (c.930G>A), which was identified in several dog breeds, was a silent mutation of Glu. The correlation between SNP1 genotype and CAD in Yorkshire Terriers was not significant. In addition, a significant difference in the relative luciferase activity between the general dog breed *CD14* promoter (*−187C*/luc) and Yorkshire Terrier type (*−187T*/luc) was not observed. The reporter plasmid harboring mutated sequence (TGATTTT) lacked the relative luciferase activity. These results suggest that both SNP1 alleles were functional TREs in dogs. *CD14* was highly conserved in dog breeds, and its expression may be regulated by transcription factor AP-1.

Acknowledgements

The authors thank Masaki Masuda, Masanari Iwata, and Ryuji Hashigami for assisting with experiments, Takeshi Yamaguchi for assistance with sample collection and Shinobu Matsuura and Merran Govendir for their advice. This work was partially supported by The Young Scientist Research Training Award (2015) funded by the Azabu University Research Service Division, Grant-in-Aid for Young Scientists (B) by The Japan Society for the Promotion of Science (17K18189), MEXT-Supported Program for the Strategic Research Foundations (2011-2015) and the Science Research Promotion Fund for Private University.

References

1. Goyert, S.M., Ferrero, E., Rettig, W.J., Yenamandra, A.K., Obata, F., and Le Beau, M.M. The CD14 monocyte differentiation antigen maps to a region encoding growth factors and receptors. *Science.* **239**: 497–500 (1988).
2. Kelley, S.L., Lukk, T., Nair, S.K., and Tapping, R.I. The crystal structure of human soluble CD14 reveals a bent solenoid with a hydrophobic amino-terminal pocket. *J Immunol.* **190**: 1304–1311 (2013).
3. Haziot, A. The monocyte differentiation antigen, CD14, is anchored to the cell membrane by a phosphatidylinositol linkage. *J Immunol.* **141**: **547**–552 (1988).
4. Da Silva, Correia, J., Soldau, K., Christen, U., Tobias, P.S., and Ulevitch, R.J. Lipopolysaccharide is in close proximity to each of the proteins in its membrane receptor complex. transfer from CD14 to TLR4 and MD-2. *J Biol Chem.* 276: 21129–21135(2001).
5. Nagai, Y., Garrett, K.P., Ohta, S., Bahrun, U., Kouro, T., Akira, S., Takatsu, K., and Kincade, P.W. Toll-like receptors on hematopoietic progenitor cells stimulate innate immune system replenishment. *Immunity.* **24**: 801–812 (2006).
6. Wright, S.D., Ramos, R.A., Tobias, P.S., Ulevitch, R.J., and Mathison, J.C. CD14, a receptor for complexes of lipopolysaccharide (LPS) and LPS binding protein. *Science.* **249**: 1431–1433 (1990).
7. Lauener, R.P., Birchler, T., Adamski, J., Braun-Fahrländer,

C., Bufe, A., Herz, U., von Mutius, E., Nowak, D., Riedler, J., Waser, M. and Sennhauser, F.H. Expression of CD14 and Toll-like receptor 2 in farmers' and non-farmers' children. *The Lancet.* **360**: 465–466 (2002).

8. Baldini, M., Lohman, I.C., Halonen, M., Erickson, R.P., Holt, P.G., and Martinez, F.D. A polymorphism in the 5' flanking region of the CD14 gene is associated with circulating soluble CD14 levels and with total serum immunoglobulin E. *Am J Respir Cell Mol Biol.* **20**: 976–983 (1999).
9. Nishimura, S., Zaitsu, M., Hara, M., Yokota, G., Watanabe, M., Ueda, Y., Imayoshi, M., Ishii, E., Tasaki, H., and Hamasaki, Y. A polymorphism in the promoter of the CD14 gene (CD14/-159) is associated with the development of coronary artery lesions in patients with Kawasaki disease. *J Pediatr.* **143**: 357–362 (2003).
10. Hubacek, J.A., Rothe, G., Pit'ha, J., Skodová, Z., Staněk, V., Poledne, R. and Schmitz, G. C(-260)-->T polymorphism in the promoter of the CD14 monocyte receptor gene as a risk factor for myocardial infarction. *Circulation.* 99: 3218–3220 (1999).
11. Favrot, C., Steffan, J., Seewald, W., and Picco, F. A prospective study on the clinical features of chronic canine atopic dermatitis and its diagnosis. *Vet Dermatol.* **21**: 23–30 (2010).
12. Katakawa, Y., Funaba, M., and Murakami, M. Smad8/9 Is Regulated Through the BMP Pathway. *J Cell Biochem.* **117**: 1788–1796 (2016).
13. Liu, S., Shapiro, R.A., Nie, S., Zhu, D., Vodovotz, Y., and Billiar, T.R. Characterization of rat CD14 promoter and its regulation by transcription factors AP1 and Sp family proteins in hepatocytes. *Gene*. **250**: 137-147 (2000).
14. Matsuura, K., Ishida, T., Setoguchi, M., Higuchi, Y., Akizuki, S., and Yamamoto, S. Identification of a tissue-specific regulatory element within the murine CD14 gene. *J Biol Chem*. **267**: 21787-21794 (1992).
15. Zhang, D.E., Hetherington, C.J., Tan, S., Dziennis, S.E., Gonzalez, D.A., Chen, H.M., and Tenen, D.G. Sp1 is a critical factor for the monocytic specific expression of human CD14. *J Biol Chem*. **269**: 11425-11434 (1994).
16. Wu, H., Moulton, K., Horvai, A., Parik, S., and Glass, C.K.. Combinatorial interactions between AP-1 and ets domain proteins contribute to the developmental regulation of the macrophage scavenger receptor gene. *Mol Cell Biol*. **14**: 2129-2139 (1994).
17. LeVan, T.D., Bloom, J.W., Bailey, T.J., Karp, C.L., Halonen, M., Martinez, F.D., and Vercelli, D. A common single nucleotide polymorphism in the CD14 promoter decreases the affinity of Sp protein binding and enhances transcriptional activity. *J Immunol*. **167**: 5838-5844 (2001).
18. Seldeen, K.L., McDonald, C.B., Deegan, B.J., Bhat, V., and Farooq, A. DNA plasticity is a key determinant of the energetics of binding of Jun-Fos heterodimeric transcription factor to genetic variants of TGACGTCA motif. *Biochemistry*. **48**: 12213-12222 (2009).
19. Kawarai, S., Hashizaki, K., Kitao, S., Nagano, S., Madarame, H., Neo, S., Ishikawa, T., Furuichi, M., Hisasue, M., Tsuchiya, R., Tsujimoto, H., and Yamada, T. Establishment and characterization of primary canine hepatocellular carcinoma cell lines producing alpha-fetoprotein. *Vet Immunol Immunopathol.* **113**: 30-36 (2006).
20. Chen, F., Ma, L., Al-Ansari, N., and Shneider, B. The role of AP-1 in the transcriptional regulation of the rat apical sodium-dependent bile acid transporter. *J Biol Chem*. **276**: 38703-38714 (2001).

Table S1 Enrolled dog breeds and the number of samples for full sequences

Dog breed	Number
Yorkshire Terrier	4
Beagle	3
Shiba	3
American Cocker Spaniel	3
Miniature Schnauzer	2
Italian Greyhound	1
Cavalier King Charles Spaniel	1
Saluki	1
Papillon	1
Brussels Griffon	1
Bulldog	1
Pekingese	1
Pomeranian	1
Maltese	1
Miniature Pinscher	1
Labrador Retriever	1

Table S2 Primers sequence used in this study and the product lengths

Primer name	Forward Primer	Reverse Primer	Product(bp)
Primer1	GGCTTGCAATCTCTTCCAAA	GGTGCTTCGAGCCTCTATGT	531
Primer2	GAGTCTTGGCTTGAGGCACT	TGCGCCCTTTAGAAATTGTT	698
Primer3	CACACCCTGCTTGCTGCT	CGCGATGTTCAGTACCTTGA	523
Primer4	GGGCCTCAAGGTACTGAACA	CAGCCCAGAGAAGGACAAGT	352
Primer5	TCTCTGTCCGCAAAAGTTCC	GGCAATCTGAGGCAATTCAT	517
Primer6	GGGATGTCAGGAACTCTCGC	CCTGTATTTTCCTAAGCATGTTTACCA	658
5’RACE RT Primer	AGCGTCAGTTCCTTG		497
5’RACE Primer 1st	CGCGGACCCAAAGCAGTA	CCACCTCGACGGCAATCATA	448
5’RACE Primer 2nd	TCACCGTGGCCTCTGCAC	TCGTCGTCATCTACTTCGCAG	315
3’RACE Primer	GGTACCCGGTGGGGATGTCAGGAA		624
3’RACE vector specific Primer	GAGCGGATAACAATTTCACACAGG	CGCCAGGGTTTTCCCAGTCACGAC	805

Table S3 In silico database used in the study

Database	URL
MEGA6 software	http://megasoftware.net/
SignalP 4.0	http://www.cbs.dtu.dk/services/SignalP-4.0/
NetNGlyc 1.0 software	http://www.cbs.dtu.dk/services/NetNGlyc/
PredGPI prediction server	http://gpcr.biocomp.unibo.it/predgpi/index.htm
TRANSFAC Public 6.0	http://www.gene-regulation.com/cgi-bin/pub/programs/match/bin/match.cgi
White Paper on Household Animals 2015	https://www.anicom-page.com/hakusho/book/pdf/book_201512.pdf