Distribution of the CD4 Alleles in Sus scrofa Demonstrates the Genetic Profiles of Western Breeds and Miniature Pigs

Manuscript for *Animal Biotechnology* (Article)

Distribution of the *CD4* alleles in *Sus scrofa* demonstrates the genetic profiles of Western breeds and miniature pigs

Tomoko Eguchi-Ogawa[*,1,2], Tatsuya Matsubara[†,1], Daisuke Toki[‡,3], Naohiko Okumura[‡], Asako Ando[§], Hitoshi Kitagawa[†] and Hirohide Uenishi[*,¶]

[*]Animal Genome Research Unit, Agrogenomics Research Center, National Institute of Agrobiological Sciences, 2 Ikenodai, Tsukuba, Ibaraki 305-8602, Japan

[†]United Graduate School of Veterinary Sciences, Gifu University, 1-1 Yanagido, Gifu 501-1193, Japan.

[‡]Animal Research Division, Institute of Japan Association for Techno-innovation in Agriculture, Forestry and Fisheries, 446-1 Ippaizuka, Kamiyokoba, Tsukuba, Ibaraki 305-0854, Japan.

[§]Department of Molecular Life Science, Division of Basic Medical Science and Molecular Medicine, Tokai University School of Medicine, 143 Shimokasuya, Isehara, Kanagawa 259-1193, Japan.

[¶]Animal Bioregulation Unit, Division of Animal Sciences, Institute of Agrobiological Sciences, National Agriculture and Food Research Organization, 1-2 Owashi, Tsukuba, Ibaraki 305-8634, Japan

[1] These two authors contributed equally to this work.

[2] Present address: Office of Evaluation, National Agriculture and Food Research Organization, 3-1-1 Kannondai, Tsukuba, Ibaraki 305-8517, Japan.

[3] Present address: National Livestock Breeding Center, 1 Odakurahara, Odakura, Nishigo, Fukushima 961-8511, Japan

27 Corresponding author: Hirohide Uenishi
28 Animal Bioregulation Unit, Division of Animal Sciences, Institute of Agrobiological
29 Sciences, National Agriculture and Food Research Organization, 1-2 Owashi, Tsukuba,
30 Ibaraki 305-8634, Japan
31 TEL/FAX: 029-838-6292; E-mail: huenishi@affrc.go.jp
32
33 Running title: Distribution of *CD4* alleles in *Sus scrofa*

Abstract

Widely used anti-pig CD4 monoclonal antibodies (mAbs) fail to recognize CD4 alleles characteristic of miniature pig lines such as NIH miniature pigs and microminipigs. We surveyed polymorphisms in the coding sequence of the porcine *CD4* gene among Western and Oriental pig breeds and Japanese wild boars and investigated their distribution. Of the 13 alleles that we identified among the 47 animals, two in group I and three in group II were found exclusively in Western breed pigs. Group IV alleles, which included mAb-nonbinding alleles, were found frequently in Oriental breed pigs, suggesting that the mAb-nonbinding allele arose from the gene pool of Oriental pigs. Group IV alleles were also found in Duroc and Large White pigs, suggesting genetic inflow from Oriental pig breeds into Western breeds. Comparison of the *CD4* sequences of species in Cetartiodactyla suggested that the group IV alleles in *Sus scrofa* occurred before the divergence of this species from the other artiodactyls. The different antibody specificities of the various CD4 alleles may facilitate the discrimination of T-cell populations in transplantation studies using miniature pigs. The significance of the preservation of CD4 polymorphisms to immune function in pigs warrants further investigation.

Keywords: allele frequency; CD4; genetic polymorphisms; miniature pig; swine

INTRODUCTION

The CD4 molecule characterizes various T-cell populations, including helper and regulatory T cells (1). The CD4 molecule interacts with MHC class II molecules, and most CD4-postive T cells are MHC-II restricted.

Miniature pigs, generated by crossbreeding of pigs with a small body size, have become popular research models, given the similarity of their organ size and immune-related genes to those of humans and their ease of handling. NIH miniature pigs, a representative line of miniature pigs, have a characteristic *CD4* allele (*CD4.2*) that fails to react with widely used anti-pig CD4 monoclonal antibodies (mAbs). Compared with the major *CD4* allele (*CD4.1*), *CD4.2* has nucleotide substitutions within exons 3 and 4, thus preventing the molecule from binding to the mAbs (2).

Microminipigs (MMPs) are a unique miniature pig line whose characteristic extremely small body size (<10 kg in mature animals) provides cost benefits in terms of husbandry and materials during preclinical testing (3). The major *CD4* allele in MMPs (*CD4.A*; LC064059.1) shows good agreement with the pig *CD4* reference sequence (NM_001001908.1). In addition, MMPs carry an alternative allele (*CD4.B*; LC064060.1) that contains multiple altered bases in exons 3 and 4 and that is highly similar to *CD4.2* in NIH miniature pigs (4). T cells derived from MMPs homozygous for *CD4.B* fail to bind to currently available anti-pig CD4 mAbs, indicating that the CD4.B molecule has lost the epitope recognized by the mAbs, again similar to CD4.2 (5). To clarify the origin and functional relevance of the mAb-nonbinding CD4 alleles, we investigated the distribution of the *CD4* alleles in populations of pigs and wild boars and analyzed the origin of the mAb-nonbinding *CD4* allele.

MATERIALS AND METHODS

Animals and Genomic DNA

We used 47 animals that had no obvious kinship with each other: 32 Western-breed subjects (Berkshire, 8; Duroc, 8; Landrace, 8; Large White, 8), 8 pigs representing Oriental breeds (Meishan, 4; Jinhua, 4), and 7 Japanese wild boars. Sperm, blood, and tissue samples (tail ends, ear auricles, and kidney) were commercially purchased or kindly provided by the Miyagi Livestock Experimental Station (Osaki, Miyagi, Japan); the Chiba Prefecture Livestock Research Center (Yachimata, Chiba, Japan); the Shizuoka Swine and Poultry Research Center (Kikugawa, Shizuoka, Japan); the Tokushima Agriculture, Forestry, and Fisheries Technology Support Center (Kamiita, Tokushima, Japan); and the Ibaraki Farm of the National Livestock Breeding Center (Chikusei, Ibaraki, Japan). Samples for DNA extraction were obtained according to the regulations regarding animal experiments at the Institute of Livestock and Grassland Science, National Agriculture and Food Research Organization, and the Institute of Japan Association for Techno-innovation in Agriculture, Forestry and Fisheries, and at the institutes that provided the samples. Genomic DNA was obtained from sperm or tissue samples by using phenol extraction followed by ethanol precipitation (6).

Sequencing

Primers were designed and used to PCR-amplify exons 2 through 10 of the porcine *CD4* gene. For each reaction, 20 ng of genomic DNA was used as a template. PCR amplification was performed by using AmpliTaq Gold polymerase (Thermo Fisher Scientific, Waltham, MA, USA) in 35 cycles consisting of 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 1 min after preincubation at 94 °C for 9 min. The amplified products were kept at 4 °C until further use. Residual primers in the amplified products were removed by treating with Exo-SAP IT solution (Affymetrix, Santa Clara, CA, USA) according to the manufacturer's instructions. The products were sequenced with the

same primers used for the PCR amplification and BigDye Terminator v3.1 by using an ABI 3730xl Genetic Analyzer (Thermo Fisher Scientific). Primers used in the PCR and sequencing reactions are shown in Table 1. The sequences thus obtained were processed as previously described (7). Polymorphisms were detected by using the PolyPhred program (8). The pig *CD4* mRNA sequence (NM_001001908.1) was used as a reference; it is 2715 bp long, comprises 11 exons, and contains a 1374-bp coding sequence (CDS) corresponding to 457 amino acids. The locations of SNPs and the exon–intron structure were determined by comparing the reference and the genomic sequence of pig chromosome 5 (NC_010447.4). Phylogenetic analysis according to the neighbor-joining method (9) was performed with ClustalX2 (10). Haplotypes of SNPs within pig *CD4* genes were predicted by using PHASE with default parameters (11).

RESULTS AND DISCUSSION

We identified 40 SNPs, which were located in exons 3 through 9 within the CDS of the reference pig *CD4* gene (Table 2). Of these, 14 SNPs were in exon 3, and 10 of these SNPs were nonsynonymous. We then estimated 13 *CD4* alleles in the surveyed animals according to computational reconstruction of the SNP haplotypes (Tables 3 and 4). The alleles with reconstructed haplotypes were derived from computational estimations. Furthermore, some of the alleles (*CD4*02*, *CD4*05*, *CD4*08*, *CD4*09*, and *CD4*13*) were not observed as homozygotes in the population thus examined. The actual combination of SNPs in single alleles was difficult to confirm because of chimerism between alleles during PCR amplification. However, because alleles that were not found as homozygotes (except for *CD4*09* and *CD4*13*) were observed as heterozygotes with other major alleles, we think that the estimated alleles are reliable. All of the *CD4*09* and *CD4*13* alleles were found in combination with each other, thus leading to uncertainty in their estimation; consequently the locations of these alleles in

131 the phylogenetic tree may be slightly inaccurate.

132 One of the 13 alleles, *CD4*06* was identical to *CD4.A*, which was the major
133 allele in MMPs. These alleles were classified into four groups through phylogenetic
134 analysis (Figure 1A). The most frequently observed allele, *CD4*01*, accounted for
135 approximately two-thirds of the alleles in Western breeds. *CD4*01*, as well as the other
136 alleles in group I including *CD4.1* of NIH miniature pigs, had the identical sequence to
137 the reference *CD4* sequence in the extent of *CD4.1*; *CD4*01* was therefore considered
138 to be the representative allele in Western breeds (Figure 1B). Group II alleles were
139 common in Western breeds and shared several characteristic nucleotides in exons 3, 4,
140 5, and 8 with group I alleles; overall their similarity to group I alleles was higher than
141 that of other groups (Table 3 and Figure 1A). Alleles in groups I and II were found
142 exclusively in Western breed pigs in this study. Duroc pigs examined here did not have
143 any group II alleles. Although our small sample size might have accounted for this
144 finding, the genetic population size of the Duroc breed may be smaller than that for
145 other Western breeds, such as Landrace and Large White (12), and pigs with group II
146 alleles might have had no, or very limited, contribution to the establishment of the
147 Duroc breed. The representative alleles of Oriental breeds composed a particular clade
148 in group III, including *CD4.A*/*CD4*06*. *CD4* alleles derived from Japanese wild boars
149 were also included in group III (Table 4 and Figure 1).

150 Groups I through III of pig *CD4* alleles were highly similar to each other within
151 the region covering exon 3. However, the remaining *CD4* alleles, group IV, had many
152 polymorphic sites within the CDR2-like regions of exons 3 and 4 (13), compared with
153 those in the other groups (Table 3). Group IV included *CD4.2* and *CD4.B*, which are
154 regarded as characteristic alleles in miniature pigs, as well as alleles from Oriental
155 breeds: *CD4*12* in Jinhua pigs and *CD4*13* in Meishan pigs. Furthermore, *CD4*12*,
156 which comprised three-eighths of the alleles of Jinhua pigs, also occurred in Duroc and

157 Large White pigs, albeit at low frequency (Table 4). The deduced amino-acid sequence
158 *CD4*12* was completely identical to those of *CD4.2* and *CD4.B* in the region
159 represented in the registered *CD4.2* sequence (X65630.1). Taken together, our findings
160 indicate that the mAb-nonbinding pig *CD4* alleles that were first noted in miniature pig
161 lines occur widely throughout various pig populations, including representative Oriental
162 and Western breeds.

163 We revealed that the atypical *CD4* alleles (group IV) were not specific to
164 miniature pig lines but occurred relatively frequently in Oriental breeds and
165 occasionally in some Western breeds, thus showing good agreement with the genetic
166 influx from Oriental breeds into Western breeds, such as Duroc (14). Oriental-breed
167 pigs are frequently used as founders for miniature pig lines: for example, the generation
168 of MMPs included Vietnamese Pot-bellied pigs (3). The relatively high frequency of
169 group IV *CD4* alleles in Oriental breeds and the influence of these pigs in the
170 establishment of Western breeds and miniature pig lines suggest that the group IV
171 alleles originated in Oriental-breed pigs. However, none of the wild boars that we
172 analyzed contained any group IV *CD4* alleles, suggesting that the origin of the group IV
173 alleles might have been restricted to animals in a particular area of Asia. In addition,
174 cannot rule out the possibility that the group IV alleles in Duroc and Large White breeds
175 were derived from minor alleles in European wild boars. The existence and frequency of
176 group IV alleles in Japanese and European wild boars should be investigated to answer
177 this question.

178 The CDR2-like region of CD4, which corresponds to the boundary of exons 3
179 and 4, protrudes to the outer side of the molecule (13). This region of the human CD4
180 molecule is required for binding to gp120 of the human immunodeficiency virus (15,
181 16), and its interaction with MHC class II molecules is debated (17, 18). Although
182 polymorphisms in the CDR2-like region might be hypothesized to influence T-cell

function, studies have failed to reveal any difference in immunologic traits between pigs with typical and group IV *CD4* alleles (2, 5). The CDR2-like region of CD4 is monomorphic in most species (13); however, among species of the order Cetartiodactyla, the CD4 sequences corresponding to polymorphic sites in the CDR2-like region of porcine CD4 showed high rates of polymorphism. Furthermore, four of the amino acids of group IV CD4 at the 10 polymorphic sites between deduced amino acid sequences of *CD4.2* and the pig *CD4* reference sequence are identical to the residues in other Cetartiodactyla species, suggesting that the CD4.2/CD4.B allele arose before the divergence of pigs from other artiodactyls (Figure 2).

CONCLUSIONS

Here, we have shown the distribution of *CD4* alleles—particularly those arising from polymorphisms in CDR2-like regions—in several *Sus scrofa* populations. The functional influence of these polymorphisms remains unclear, but the differences in antibody specificity among these *CD4* alleles may facilitate the discrimination of T-cell subpopulations in transplantation studies using miniature pigs. The significance of the preservation of polymorphisms of porcine CD4 in terms of their interaction with MHC molecules and other immune functions warrants further investigation.

Acknowledgments

We thank Takako Suzuki, Toshie Iioka, and Ikuyo Nakamori for their technical assistance. This work was supported in part by the Research Program on Innovative Technologies for Animal Breeding, Reproduction and Vaccine Development (AGB1002) of the Ministry of Agriculture, Forestry and Fisheries of Japan.

Figure legends

Figure 1. Phylogeny of pig *CD4* alleles. Sequences corresponding to (A) the complete coding sequences and (B) the regions determined in a previous study (X65629.1 and X65630.1; positions 99 to 397 in the coding sequence of NM_001001908.1) (13) were analyzed by using the neighbor-joining method (9). Identical sequences were analyzed as single sequences. The full-length coding sequence of pig *CD4* (NM_001001908.1), as well as alleles in the MMP population and in NIH miniature pigs, was included in the phylogenetic analysis. The coding sequence of cattle *CD4* (NM_001103225.1) was used as an outgroup. Bootstrap trials were conducted 10,000 times, and the bootstrap values are shown at the nodes. The lengths of branches with omissions are indicated in parentheses.

Figure 2. Alignment of CD4 sequences of Cetartiodactyla species corresponding to the polymorphic region between the pig reference sequence and group IV CD4 alleles. The deduced amino-acid sequences of the porcine CD4.2 allele (X65630.1) and the corresponding regions of the pig reference sequence (NP_001001908.1) and of the CD4 proteins of representative Cetartiodactyla species (*Bos taurus*, NP_001096695.1; *Ovis aries*, NP_001123374.1; *Capra hircus*, ACG76115.1; *Orcinus orca*, XP_004279107.1; *Tursiops truncatus*, NP_001267583.1) were aligned. Triangles indicate polymorphic sites between the pig reference sequence and CD4.2; fringed letters indicate amino acids in non-pig species that are identical to those of the polymorphic sites in porcine CD4.2; dots indicate residues identical to the pig reference sequence; and dashes indicate deletions.

REFERENCES

1. Castellino F, Germain RN. Cooperation between CD4+ and CD8+ T cells: when, where, and how. Annu Rev Immunol. 2006; 24: 519-540.
2. Sundt TM, 3rd, LeGuern C, Germana S, Smith CV, Nakajima K, Lunney JK, Sachs DH. Characterization of a polymorphism of CD4 in miniature swine. J Immunol. 1992; 148: 3195-3201.
3. Kaneko N, Itoh K, Sugiyama A, Izumi Y. Microminipig, a non-rodent experimental animal optimized for life science research: preface. J Pharmacol Sci. 2011; 115: 112-114.
4. Matsubara T, Nishii N, Takashima S, Takasu M, Imaeda N, Aiki-Oshimo K, Yamazoe K, Kakisaka M, Takeshima SN, Aida Y, Kametani Y, Kulski JK, Ando A, Kitagawa H. Identification and characterization of two CD4 alleles in Microminipigs. BMC Vet Res. 2016; 12: 222.
5. Matsubara T, Nishii N, Takashima S, Takasu M, Imaeda N, Aiki-Oshimo K, Yamazoe K, Kametani Y, Ando A, Kitagawa H. Identification of a CD4 variant in Microminipigs not detectable with available anti-CD4 monoclonal antibodies. Vet Immunol Immunopathol. 2015; 168: 176-183.
6. Sambrook J, Russell DW. Molecular Cloning, 3rd ed. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press; 2001.
7. Eguchi-Ogawa T, Toki D, Uenishi H. Genomic structure of the whole D-J-C clusters and the upstream region coding V segments of the TRB locus in pig. Dev Comp Immunol. 2009; 33: 1111-1119.
8. Nickerson DA, Tobe VO, Taylor SL. PolyPhred: automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing. Nucleic Acids Res. 1997; 25: 2745-2751.
9. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 1987; 4: 406-425.
10. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. Clustal W and Clustal X version 2.0. Bioinformatics. 2007; 23: 2947-2948.
11. Stephens M, Scheet P. Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. Am J Hum Genet. 2005; 76: 449-462.
12. Grossi DA, Jafarikia M, Brito LF, Buzanskas ME, Sargolzaei M, Schenkel FS. Genetic diversity, extent of linkage disequilibrium and persistence of gametic phase in Canadian pigs. BMC Genet. 2017; 18: 6.
13. Gustafsson K, Germana S, Sundt TM, 3rd, Sachs DH, LeGuern C. Extensive allelic polymorphism in the CDR2-like region of the miniature swine CD4 molecule. J Immunol. 1993; 151: 1365-1370.
14. Giuffra E, Kijas JM, Amarger V, Carlborg Ö, Jeon JT, Andersson L. The origin of the domestic pig: independent domestication and subsequent introgression. Genetics. 2000; 154: 1785-1791.

15. Arthos J, Deen KC, Chaikin MA, Fornwald JA, Sathe G, Sattentau QJ, Clapham PR, Weiss RA, McDougal JS, Pietropaolo C, et al. Identification of the residues in human CD4 critical for the binding of HIV. Cell. 1989; 57: 469-481.
16. Landau NR, Warton M, Littman DR. The envelope glycoprotein of the human immunodeficiency virus binds to the immunoglobulin-like domain of CD4. Nature. 1988; 334: 159-162.
17. Fleury S, Lamarre D, Meloche S, Ryu SE, Cantin C, Hendrickson WA, Sekaly RP. Mutational analysis of the interaction between CD4 and class II MHC: class II antigens contact CD4 on a surface opposite the gp120-binding site. Cell. 1991; 66: 1037-1049.
18. Moebius U, Clayton LK, Abraham S, Diener A, Yunis JJ, Harrison SC, Reinherz EL. Human immunodeficiency virus gp120 binding C'C" ridge of CD4 domain 1 is also involved in interaction with class II major histocompatibility complex molecules. Proc Natl Acad Sci U S A. 1992; 89: 12008-12012.

Table 1 Primers sequences and procedures for PCR of the porcine *CD4* gene and preparation of sequencing templates.

Target exon	Location (chromosome 5)[a]	Sequence (5′→3′)	Amplicon (bp)
2	66,346,741−66,346,718	GTACCTGTGGGTGTCAGTTTAGAG	383
	66,346,359−66,346,382	CTTACCCAGCACCAGATATTTTTC	
3	66,346,515−66,346,492	CTCAGACTCAAACTGGGATGATTG	366
	66,346,150−66,346,173	GATCCCAGAGTTTACTAGGAGCTG	
4	66,333,536−66,333,512	AATGAGCAACTCAGATCAGAAGAGT	363
	66,333,174−66,333,195	CTTATCCATCTCTGGACGGTTG	
5	66,332,952−66,332,931	CTTCTCCTTGGGGATAGTGCAT	358
	66,332,595−66,332,615	ACACTACAGCCACGAGCAGAG	
6	66,331,799−66,331,775	GCCTAGAGCTAGATGGGAATTTAAG	617
	66,331,183−66,331,204	GATTCCAGCCTCAGTTCAAACC	
7	66,329,886−66,329,862	CTTTAGAGCAGACAAGTGCTAGGAA	373
	66,329,514−66,329,537	ACCATACCCATAACCCACTGACTC	
8	66,328,935−66,328,912	AGCATAAGGATCAGACCCAAGTGT	400
	66,328,536−66,328,559	TAACTCTGTGGCTTCTTGTCTCTC	
9	66,328,539−66,328,516	GTTAATTCTGGGACAGATGGCTTC	238
	66,328,302−66,328,323	CTCTCTTCACCCCTCCTCTTTG	
10	66,328,153−66,328,132	CCATCTCTGTGCAGGAAAAGTC	698
	66,327,456−66,327,476	AGCTGAGCTGCTTGGGTGATA	
10	66,327,556−66,327,537	ACTGACGGAGCCACAGACTC	668
	66,326,889−66,326,908	GGCTATCAACTTTCGCAGGA	

Table 2 SNPs in the porcine *CD4* gene.

Exon[a]	Pig genome			Coding sequence		Frequency of the reference allele in each breed (%)[d]							
						Western breeds				Oriental breeds			
	Position[a]	Allele[b]		Position[c]	Amino acid	LR (8)	LW (8)	DR (8)	BS (8)	MS (4)	JH (4)	WB (7)	Overall (47)
3	66,346,335	G	A	126	Syn	100.0	93.8	93.8	100.0	75.0	62.5	100.0	92.6
	66,346,306	C	T	155	P→L	100.0	100.0	100.0	100.0	100.0	100.0	0.0	85.1
	66,346,300	A	G	161	N→S	100.0	93.8	93.8	100.0	75.0	62.5	100.0	92.6
	66,346,291	A	G	170	N→S	93.7	100.0	75.0	100.0	25.0	37.5	100.0	83.0
	66,346,286	A	G	175	Syn	100.0	93.8	93.8	100.0	75.0	62.5	100.0	92.6
	66,346,279	C	T	182	T→I	100.0	93.8	93.8	100.0	75.0	62.5	100.0	92.6
	66,346,268	G	A	193	G→S	100.0	93.8	93.8	100.0	75.0	62.5	100.0	92.6
	66,346,266	C	G	195	Syn	100.0	93.8	93.8	100.0	75.0	62.5	100.0	92.6
	66,346,265	G	A	196	G→S	100.0	93.8	93.8	100.0	75.0	62.5	100.0	92.6
	66,346,260	T	C	201	Syn	100.0	93.8	93.8	100.0	75.0	62.5	100.0	92.6
	66,346,259	G	C	202	G→R	100.0	93.8	93.8	100.0	75.0	62.5	100.0	92.6
	66,346,255	G	A	206	S→N	100.0	93.8	93.8	100.0	75.0	62.5	100.0	92.6
	66,346,251	C	G	210	F→L	100.0	93.8	93.8	100.0	75.0	62.5	100.0	92.6
	66,346,243	C	A	218	T→K	50.0	56.3	93.8	43.8	75.0	62.5	0.0	53.2
4	66,333,448	C	T	231	Syn	100.0	100.0	100.0	100.0	50.0	100.0	100.0	95.7
	66,333,447	G	C	232	E→K	100.0	100.0	100.0	100.0	75.0	100.0	42.9	89.4
	66,333,440	C	G	239	T→S	50.0	56.3	93.8	43.8	100.0	62.5	100.0	70.2
5	66,332,898	A	G	385	T→A	100.0	100.0	100.0	100.0	50.0	100.0	100.0	95.7
	66,332,728	T	A	555	Syn	43.7	62.5	75.0	43.7	50.0	100.0	0.0	51.1
	66,332,683	T	C	600	Syn	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0[e]
6	66,331,624	C	T	665	A→V	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0[e]
	66,331,602	C	T	687	Syn	100.0	100.0	100.0	68.8	100.0	100.0	100.0	94.7
	66,331,599	G	A	690	Syn	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0[e]
	66,331,420	C	T	869	P→L	56.3	43.8	31.3	56.3	100.0	100.0	100.0	63.8
	66,331,411	C	T	878	A→V	100.0	93.8	100.0	100.0	100.0	100.0	100.0	98.9
	66,331,410	G	A	879	Syn	93.7	100.0	75.0	100.0	25.0	37.5	57.1	76.6
	66,331,380	T	C	909	Syn	56.3	43.8	31.3	56.3	100.0	100.0	100.0	63.8
	66,331,341	G	A	948	M→I	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0[e]
7	66,329,815	C	T	953	A→V	50.0	56.3	100.0	43.8	100.0	100.0	100.0	74.5
	66,329,814	G	T	954	Syn	100.0	100.0	100.0	100.0	25.0	100.0	42.9	85.1
	66,329,781	T	G	987	Syn	0.0	6.2	6.2	0.0	0.0	37.5	0.0	5.3
	66,329,724	G	C	1044	M→I	56.3	43.8	31.3	56.3	100.0	100.0	100.0	63.8
	66,329,707	A	G	1061	Q→R	6.3	6.3	31.3	0.0	25.0	100.0	42.9	24.5
	66,329,619	C	T	1149	Syn	50.0	62.5	100.0	43.8	100.0	100.0	100.0	75.5
8	66,328,791	A	G	1185	Syn	93.7	93.7	68.7	100.0	0.0	0.0	0.0	60.6
	66,328,790	G	A	1186	E→K	93.7	93.7	68.7	100.0	0.0	0.0	0.0	60.6
	66,328,777	C	T	1199	S→F	93.7	93.7	68.7	100.0	0.0	0.0	0.0	60.6
	66,328,726	C	T	1250	A→V	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0[e]
	66,328,713	C	T	1263	Syn	93.7	93.7	68.7	100.0	0.0	0.0	0.0	60.6
9	66,328,348	G	A	1335	Syn	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0[e]

Bases different from those in the reference sequence (NM_001001908.1) were regarded as SNPs. BS, Berkshire; DR, Duroc; JH, Jinhua; LR, Landrace; LW, Large White; MS, Meishan; Syn, synonymous; WB, Japanese wild boar

[a] Positions of SNPs on the pig genome are indicated relative to p-ter of chromosome 5 (NC_010447.4).

[b] The first allele in each pair is identical to the reference sequence (NM_001001908.1).

[c] Exons and positions of SNPs in the coding sequence are shown relative to the nucleotide sequence NM_001001908.1.

[d] The number of animals for each breed is indicated in parentheses.

[e] Monomorphic in the animals used in this study.

Table 3. Alleles of pig *CD4* genes inferred by SNP haplotypes in the coding sequence.

Position in the coding sequence of the reference pig *CD4* mRNA (NM_001001908.1)

Group	Allele	Exon 3													
		126	**155**	**161**	**170**	175	**182**	**193**	195	**196**	201	**202**	**206**	**210**	**218**
I	Pig RefSeq CD4 (NM_001001908.1)	G	**C**	**A**	**A**	A	**C**	**G**	C	**G**	T	**G**	**G**	**C**	**C**
	CD4.1 (X65629.1)	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	*CD4*01*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	*CD4*02*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
II	*CD4*03*	.	.	.	.	.	.	.	.	.	.	.	.	.	**A**
	*CD4*04*	.	.	.	.	.	.	.	.	.	.	.	.	.	**A**
	*CD4*05*	.	.	.	.	.	.	.	.	.	.	.	.	.	**A**
III	*CD4.A* (LC064059.1) / *CD4*06*	.	.	.	**G**	.	.	.	.	.	.	.	.	.	.
	*CD4*07*	.	.	.	**G**	.	.	.	.	.	.	.	.	.	.
	*CD4*08*	.	.	.	**G**	.	.	.	.	.	.	.	.	.	.
	*CD4*09*	.	.	.	.	.	.	.	.	.	.	.	.	.	**A**
	*CD4*10*	.	**T**	.	.	.	.	.	.	.	.	.	.	.	**A**
	*CD4*11*	.	**T**	.	.	.	.	.	.	.	.	.	.	.	**A**
IV	*CD4.2* (X65630.1)	A	.	**G**	.	G	**T**	**A**	G	**A**	C	**C**	**A**	**G**	**A**
	CD4.B (LC064060.1)	A	.	**G**	.	G	**T**	**A**	G	**A**	C	**C**	**A**	**G**	**A**
	*CD4*12*	A	.	**G**	.	G	**T**	**A**	G	**A**	C	**C**	**A**	**G**	**A**
	*CD4*13*	A	.	**G**	**G**	G	**T**	**A**	G	**A**	C	**C**	**A**	**G**	.

Group	Allele	Exon 4			Exon 5			Exon 6								Exon 7						Exon 8					Exon 9
		231	**232**	**239**	385	**555**	600	**665**	687	690	**869**	**878**	879	909	**948**	**953**	954	987	**1044**	**1061**	1149	1185	**1186**	**1199**	**1250**	1263	1335
I	Pig RefSeq CD4 (NM_001001908.1)	C	**G**	**C**	A	**T**	T	**C**	C	G	**C**	**C**	G	T	**G**	**C**	G	T	**G**	**A**	C	A	**G**	**C**	**C**	C	G
	CD4.1 (X65629.1)	.	.	.	.																						
	*CD4*01*	.	.	.	.	.	C	**T**	.	A	**T**	.	.	C	.	.	.	G	**C**	**G**	.	.	.	.	**T**	.	A
	*CD4*02*	.	.	.	.	.	C	**T**	.	A	**T**	.	.	C	.	**T**	.	G	**C**	**G**	.	.	.	.	**T**	.	A
II	*CD4*03*	.	.	**G**	.	**A**	C	**T**	.	A	.	.	.	.	.	**T**	.	G	.	**G**	T	.	.	.	**T**	.	A
	*CD4*04*	.	.	**G**	.	**A**	C	**T**	T	A	.	.	.	.	.	**T**	.	G	.	**G**	T	.	.	.	**T**	.	A
	*CD4*05*	.	.	**G**	.	**A**	C	**T**	.	A	.	**T**	.	.	.	**T**	.	G	.	**G**	T	.	.	.	**T**	.	A
III	*CD4.A* (LC064059.1) / *CD4*06*	T	.	.	G	.	C	**T**	.	A	.	.	A	.	.	.	T	G	.	**G**	.	G	**A**	**T**	**T**	T	A
	*CD4*07*	.	.	.	.	.	C	**T**	.	A	.	.	A	.	.	.	.	G	.	.	.	G	**A**	**T**	**T**	T	A
	*CD4*08*	.	.	.	.	**A**	C	**T**	.	A	.	.	A	.	.	.	.	G	.	.	.	G	**A**	**T**	**T**	T	A
	*CD4*09*	.	.	.	.	**A**	C	**T**	.	A	.	.	A	.	.	.	.	G	.	.	.	G	**A**	**T**	**T**	T	A
	*CD4*10*	.	.	.	.	**A**	C	**T**	.	A	.	.	A	.	.	.	.	G	.	.	.	G	**A**	**T**	**T**	T	A
	*CD4*11*	.	**A**	.	.	**A**	C	**T**	.	A	.	.	.	.	.	.	T	G	.	**G**	.	G	**A**	**T**	**T**	T	A
IV	*CD4.2* (X65630.1)	.	.	**G**	.																						
	CD4.B (LC064060.1)	.	.	**G**	.	.	C	**T**	.	A	.	.	.	.	**A**	.	.	.	.	.	.	G	**A**	**T**	**T**	T	A
	*CD4*12*	.	.	**G**	.	.	C	**T**	.	A	.	.	.	.	.	.	.	.	.	.	.	G	**A**	**T**	**T**	T	A
	*CD4*13*	.	**A**	.	.	**A**	C	**T**	.	A	.	.	.	.	.	.	T	G	.	**G**	.	G	**A**	**T**	**T**	T	A

Nucleotides that differ from the reference pig *CD4* mRNA sequence (NM_001001908.1) are shown; dots indicate nucleotides identical to the reference. Nonsynonymous substitutions leading to changes in amino acids (see Table 2) are bolded. Alleles detected in NIH miniature pigs (*CD4.1* and *CD4.2*) (13) and MMPs (*CD4.A* and *CD4.B*) (5) are included also; the determined sequences of the alleles in NIH miniature pigs correspond to positions 99 through 397 of the coding sequence of pig *CD4* (NM_001001908.1).

Table 4. Pig *CD4* alleles in each population.

		Western breeds (n = 8 each)				Oriental breeds (n = 4 each)			
Group	Allele	Berkshire	Duroc	Landrace	Large White	Meishan	Jinhua	Japanese wild boar (n = 7)	Total alleles
I	*CD4*01*	7	11	7	8				33
	*CD4*02*				1				1
II	*CD4*03*	4		8	5				17
	*CD4*04*	5							5
	*CD4*05*				1				1
III	*CD4*06*					4			4
	*CD4*07*						5		5
	*CD4*08*		4	1					5
	*CD4*09*					2			2
	*CD4*10*							6	6
	*CD4*11*							8	8
IV	*CD4*12*		1		1		3		5
	*CD4*13*					2			2
Total alleles		16	16	16	16	8	8	14	94

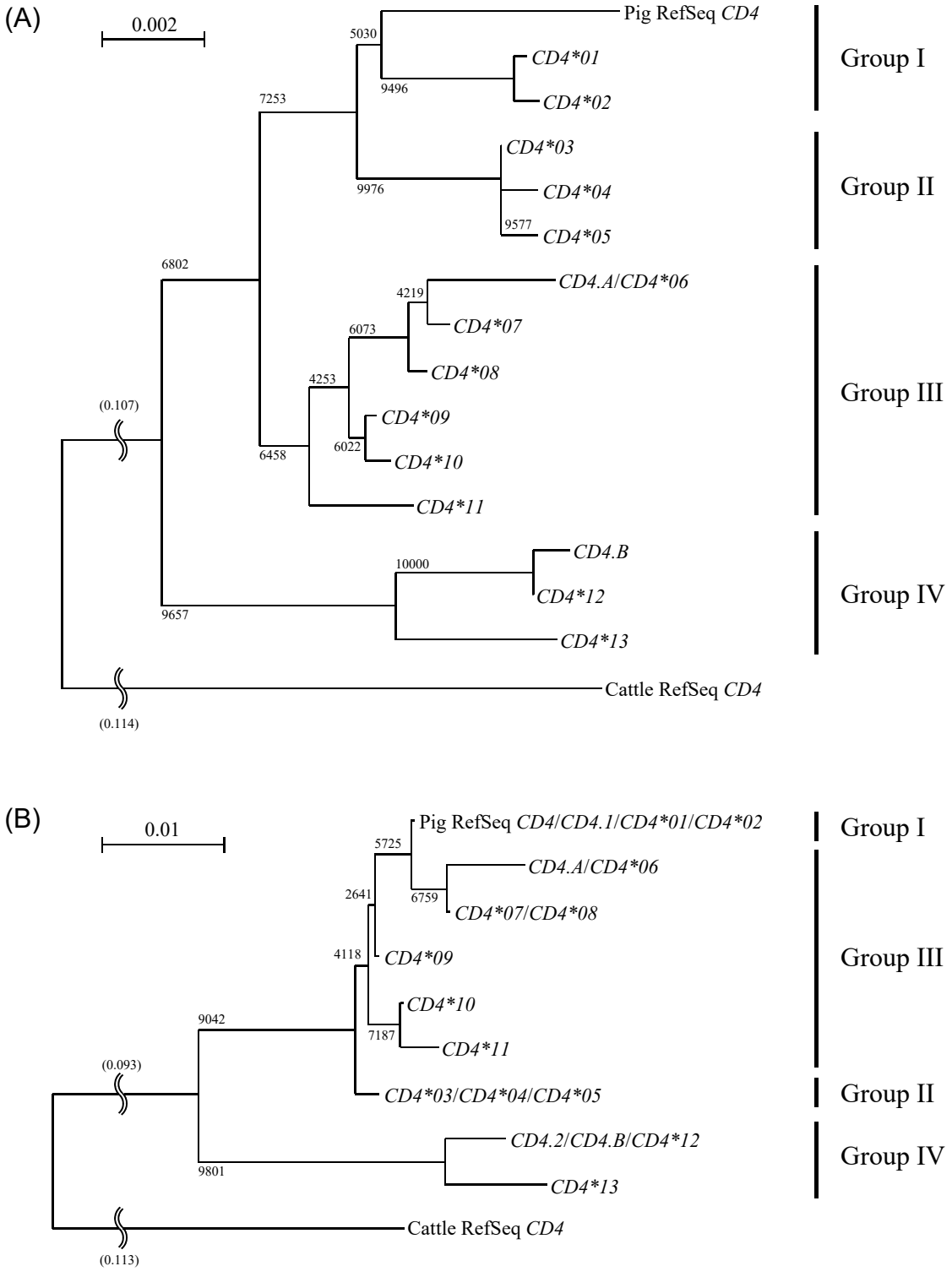
(A)
0.002
Pig RefSeq CD4
5030
CD4*01
9496
CD4*02
7253
CD4*03
9976
CD4*04
9577
CD4*05
6802
CD4.A/CD4*06
4219
CD4*07
6073
CD4*08
4253
(0.107)
CD4*09
6022
CD4*10
6458
CD4*11
CD4.B
10000
CD4*12
9657
CD4*13
Cattle RefSeq CD4
(0.114)
Group I
Group II
Group III
Group IV
(B)
0.01
Pig RefSeq CD4/CD4.1/CD4*01/CD4*02
5725
CD4.A/CD4*06
6759
CD4*07/CD4*08
2641
4118
CD4*09
CD4*10
9042
7187
CD4*11
(0.093)
CD4*03/CD4*04/CD4*05
CD4.2/CD4.B/CD4*12
9801
CD4*13
Cattle RefSeq CD4
(0.113)
Group I
Group III
Group II
Group IV

Figure 1

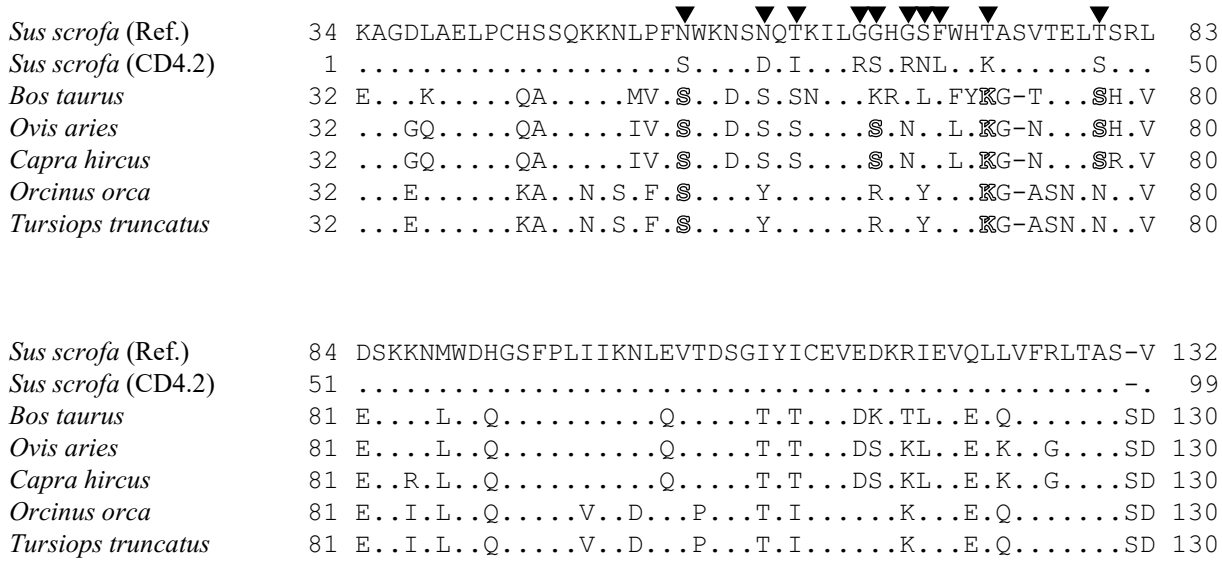
Sus scrofa (Ref.) 34 KAGDLAELPCHSSQKKNLPFNWKNSNQTKILGGHGSFWHTASVTELTSRL 83
Sus scrofa (CD4.2) 1 .....S....D.I...RS.RNL..K......S... 50
Bos taurus 32 E...K.....QA.....MV.S..D.S.SN...KR.L.FYKG-T...SH.V 80
Ovis aries 32 ...GQ.....QA.....IV.S..D.S.S....S.N..L.KG-N...SH.V 80
Capra hircus 32 ...GQ.....QA.....IV.S..D.S.S....S.N..L.KG-N...SR.V 80
Orcinus orca 32 ...E......KA..N.S.F.S....Y......R..Y...KG-ASN.N..V 80
Tursiops truncatus 32 ...E......KA..N.S.F.S....Y......R..Y...KG-ASN.N..V 80
Sus scrofa (Ref.) 84 DSKKNMWDHGSFPLIIKNLEVTDSGIYICEVEDKRIEVQLLVFRLTAS-V 132
Sus scrofa (CD4.2) 51 .....-. 99
Bos taurus 81 E....L..Q.....Q.....T.T...DK.TL..E.Q.....SD 130
Ovis aries 81 E....L..Q.....Q.....T.T...DS.KL..E.K..G....SD 130
Capra hircus 81 E..R.L..Q.....Q.....T.T...DS.KL..E.K..G....SD 130
Orcinus orca 81 E..I.L..Q.....V..D...P...T.I......K...E.Q.....SD 130
Tursiops truncatus 81 E..I.L..Q.....V..D...P...T.I......K...E.Q.....SD 130

Figure 2