農研機構
NARO 国立研究開発法人 農業・食品産業技術総合研究機構

Pig lacks functional NLRC4 and NAIP genes

メタデータ	言語: eng 出版者: 公開日: 2020-01-06 キーワード (Ja): キーワード (En): 作成者: 佐久間, 智理, 新開, 浩樹, 竹之内, 敬人, 佐藤, 充, 木谷, 裕, 上西, 博英 メールアドレス: 所属:
URL	https://repository.naro.go.jp/records/3339

This work is licensed under a Creative Commons Attribution-NonCommercial 3.0 International License.

Manuscript for *Immunogenetics* (Brief communication)

Pig lacks functional *NLRC4* and *NAIP* genes

Chisato Sakuma,[1,3] Daisuke Toki,[2] Hiroki Shinkai,[1] Takato Takenouchi,[1] Mitsuru Sato,[1] Hiroshi Kitani,[1] Hirohide Uenishi[1,*]

[1]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

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

[3]Present address: National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku, Tokyo 162-8640, Japan

[*]Address correspondence to: Hirohide Uenishi
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
E-mail: huenishi@affrc.go.jp; Tel./fax.: +81-29-838-6292

Abstract: The NLRC4 inflammasome, which recognizes flagellin and components of the type III secretion system, plays an important role in the clearance of intracellular bacteria. Here, we examined the genomic sequences carrying two genes encoding key components of the NLRC4 inflammasome—NLR family, CARD-containing 4 (NLRC4), and NLR apoptosis inhibitory protein (NAIP)—in pigs. Pigs have single loci encoding NLRC4 and NAIP. Comparison of the sequences thus obtained with the corresponding regions in humans revealed the deletion of intermediate exons in both pig genes. In addition, the genomic sequences of both pig genes lacked valid open reading frames encoding functional NLRC4 or NAIP protein. Additional pigs representing multiple breeds and wild boars also lacked the exons that we failed to find through genome sequencing. Furthermore, neither the *NLRC4* nor the *NAIP* gene was expressed in pigs. These findings indicate that pigs lack the NLRC4 inflammasome, an important factor involved in monitoring bacterial proteins and contributing to the clearance of intracellular pathogens. These results also suggest that genetic polymorphisms affecting the molecular functions of TLR2, TLR4, TLR5, and other pattern recognition receptors associated with the recognition of bacteria have a more profound influence on disease resistance in pigs than in other species.
Keywords: Genome sequencing; Inflammasome; Pattern recognition receptors; Swine

Opportunistic infections, which can result in great economic loss, are serious problems for modern pork production. The defense response of hosts against pathogens is accomplished by a variety of immune factors, including pattern recognition receptors (PRRs). PRRs recognize pathogen-associated molecular patterns derived from infectious agents and induce diverse immune responses, including the production of inflammatory cytokines to eradicate the agents. Mutations in PRR genes may hamper ligand recognition, thus increasing susceptibility to infection. Studies in humans and mice have revealed that single-nucleotide polymorphisms (SNPs) that occur in PRRs and impair the molecular functions of these receptors also lead to increases in host susceptibility to disease (Schröder and Schumann 2005). We have characterized many polymorphisms in porcine PRR genes, some of which influence ligand recognition (Uenishi and Shinkai 2009; Uenishi et al. 2012). For instance, in NOD-like receptors (NLRs; one subgroup of PRRs), nucleotide-binding oligomerization domain 1 (NOD1), which recognizes γ-d-glutamyl-meso-diaminopimelic acid (iE-DAP), has two nonsynonymous SNPs in the leucine-rich repeat region; these SNPs impair iE-DAP–induced activation of NF-κB in pigs (Shinkai et al. 2015). Such results suggest that studying genetic polymorphism in genes associated with immune responses, such as those for PRRs, will provide useful information for breeding in regard to disease resistance.

A large number of NLRs, such as the NLRC and NLRP families, are components of inflammasomes, which are cytoplasmic protein complexes that recognize microbial components and endogenous danger signals. The NLRC4 inflammasome contains the NLR family, CARD-containing 4 (NLRC4), and NLR apoptosis inhibitory protein (NAIP) as critical components. The NLRC4

inflammasome responds to bacterial flagellin and proteins in the type III secretion system (T3SS) (Franchi et al. 2006; Miao et al. 2010). Mice express multiple NAIP paralogs; for example, murine NAIP1 and NAIP2 respond to the needle proteins and inner rod proteins, respectively, of T3SS, and murine NAIP5 and NAIP6 recognize flagellin (Kofoed and Vance 2011; Zhao et al. 2011). Humans express a single NAIP homolog, which binds to the CprI needle proteins of T3SS (Zhao et al. 2011). NAIP interacts with flagellin or components of T3SS and then activates NLRC4 to form an inflammasome complex. Caspase-1 is activated by the NLRC4 inflammasome and cleaves the proinflammatory form of cytokine IL-1β into its active form (Poyet et al. 2001).

In many mammalian species, the NLRC4 inflammasome has a pivotal role in innate immune responses to microorganisms and contributes to resistance to infectious diseases. To investigate the function of the NLRC4 inflammasome and its component molecules in pigs, we examined the structures of the *NLRC4* and *NAIP* genes in the pig genome.

Because sequencing of the pig genome had not completely clarified the regions encompassing *NLRC4* and *NAIP*, we screened bacterial artificial chromosome (BAC) clones carrying porcine *NLRC4* and *NAIP* genes. We used a polymerase chain reaction (PCR)-based system (Suzuki et al. 2000) and a library constructed by using the pCC1BAC vector (Epicentre, Madison, WI, USA) and DNA obtained from the kidney of a Landrace boar (Tanaka et al. 2006). Two clones were isolated from the library by using specific primers designed within the estimated porcine *NLRC4* and *NAIP* exonic sequences (Table 1).

DNAs of the BAC clones thus isolated were prepared by using the QIAGEN Large-Construct Kit (Qiagen, Hilden, Germany) and subjected to shotgun sequencing as described previously (Eguchi-Ogawa et al. 2007) by using an ABI 3730xl sequencer (Applied Biosystems, Foster City, CA, USA) and Big Dye Terminator version 3.1 cycle sequencing kit (Applied Biosystems). Sequence data were processed by using the Phred base-calling program and assembled into contiguous sequences by using the Phrap assembler (Ewing and Green 1998; Ewing et al. 1998). Bases with low-quality values determined by Phred and those considered to be errors or misassembled were verified by using the Consed program (Gordon et al. 1998). The integrity of the assembly of the sequences of the BAC clone was confirmed by digestion of the BAC DNA with the restriction enzymes *BssH*II, *Pvu*I, and *Not*I and by pulsed-field gel electrophoresis. Sequencing of BAC clone L442I04 yielded 193,453 bp of porcine genomic sequence (GenBank: LC144946), which contained a putative *NLRC4* gene located between the *SLC30A6* and *YIPF4* genes (Fig. 1a). BAC clone L147A24 contained 171,988 bp of porcine genomic sequence (GenBank: LC144947), which possessed a putative *NAIP* gene between the *GTF2H2* and *SMN1* genes (Fig. 1b).

Exons of porcine *NLRC4* and *NAIP* were estimated through comparison of the porcine genomic sequences with human cDNA sequences coding *NLRC4* (NM_021209.4) and *NAIP* (NM_004536.2),

respectively. In addition, nucleotide sequences were compared between pigs and humans by dot-plot analysis using PipMaker (Schwartz et al. 2000), with default parameters. Whereas the coding sequence (CDS) for the human *NLRC4* transcript has 8 exons (exons 2 through 9), porcine *NLRC4* lacked exons corresponding to exons 5, 6, and 7 of human *NLRC4*. In addition, an inversion was detected approximately 5 kb upstream of porcine *NLRC4* (Fig. 1a).

The number of loci of *NAIP* genes varies among species: for example, mice have four functional *Naip* genes (*Naip1*, *Naip2*, *Naip5*, and *Naip6*) in their genome, whereas humans have only one functional locus for *NAIP*. According to the synteny between the genomic region carrying *NAIP* in pigs and those of other species, including humans, mice, and cattle, pigs have a single *NAIP* locus in the genome, as do humans and cattle. The human *NAIP* transcript has 17 exons, and its CDS is coded by exons 4 through 17. However, porcine *NAIP* lacked exons corresponding to human exons 5 through 8, 10, 11, and 15 (Fig. 1b, Table 2). In addition, the porcine exons corresponding to exons 12 and 17 of human *NAIP* were collapsed, and clear correspondence between the human and pig sequences was not observed (Supplementary Fig. 1). Although most putative exons of porcine *NLRC4* gene were flanked by introns with canonical GT/AG splicing sites, we did not detect canonical GT/AG splicing sites at the end of introns upstream of the porcine *NAIP* exon corresponding to human *NAIP* exon 14 (Table 2). For both *NLRC4* and *NAIP* exons, because of the presence of in-frame stop codons we were unable to reconstruct putative mRNAs (data not shown).

Although public nucleotide databases (DDBJ/EMBL/GenBank) contain more than 1,680,000 expressed sequence tags (ESTs) of pigs, none of these ESTs showed clear similarity to any putative exon of *NLRC4* and *NAIP* in the pig genome (data not shown). To obtain additional evidence, we prepared porcine macrophages isolated from the kidney; these cells showed classic characteristics of macrophages, such as cell-surface antigens and phagocytosis (Takenouchi et al. 2014). By using RT-PCR analysis and primers designed to bind within putative exons of porcine *NLRC4* and *NAIP* (Supplementary Fig. 2), we failed to obtain any *NLRC4* or *NAIP* amplicons from either unstimulated porcine macrophages or those stimulated with lipopolysaccharide or flagellin. Taken together, our findings imply that porcine NLRC4 and NAIP are not produced, and we conclude that *NLRC4* and *NAIP* are pseudogenes in the pig genome.

Because the deletion of the exons in genes associated with the NLRC4 inflammasome might have been restricted to particular individuals or populations of pigs, we conducted PCR amplification around the porcine genomic regions corresponding to the missing exons of *NLRC4* and *NAIP*. We designed PCR primers within flanking regions adjacent to the missing porcine exons corresponding to exons 5 through 7 of human *NLRC4* and exons 5 through 8 of human *NAIP*. The PCR primers for *NLRC4* were designed to amplify five short, overlapping fragments (less than 700 bp each) because the intervening sequence between the porcine *NLRC4* exons corresponding to human *NLRC4* exons 4 and 8 was longer than 2 kb (Table 1; Supplementary Fig. 3). As templates for the PCR amplifications, genomic DNAs were prepared

from pigs representing major commercial breeds of Japanese livestock (three Duroc, four Landrace, four Large White, and three Berkshire breed pigs), two Chinese breed pigs (Jinhua), and two Japanese wild boars and were used with the primers and AmpliTaq Gold (Applied Biosystems). Genomic DNA of all of the individuals thus examined, which generated a correct size of the PCR fragment for *TLR2* as a positive control (Table 1), generated fragments with sizes corresponding to those of the sequences of the BAC clones. Sequencing of the ends of the PCR fragments showed that the PCR correctly amplified the genomic regions carrying the missing exons of *NLRC4* and *NAIP* on the pig genome, demonstrating that the loss of function in the NLRC4 inflammasome occurs commonly in the species *Sus scrofa* (data not shown).

In mice, NAIP2 activates NLRC4 in response to conserved components of bacterial T3SS, whereas NAIP5 interacts directly with flagellin (Miao et al. 2010; Zhao et al. 2011). Such stimulatory signals activate caspase 1 through the NLRC4 inflammasome, resulting in maturation of IL-1β/IL-18. These findings indicate that NLRC4–NAIP works as an intracellular sensing system for bacterial proteins. In comparison, TLR5 detects extracellular flagellin and induces the expression of inflammatory cytokines (Hayashi et al. 2001). Furthermore, T3SS induces inflammatory responses through TLR2 and TLR4 (Jessen et al. 2014). These redundant recognition systems for bacterial proteins may help to prevent the otherwise critical influence of mutated forms of receptors involved in bacterial recognition. Here, we showed that the NLRC4 inflammasome is absent in pigs, suggesting their potentially increased susceptibility to genetic polymorphisms that affect the functions of TLR2, TLR4, or TLR5.

Both porcine *NLRC4* and *NAIP* lacked several exons corresponding to those in the human CDSs. However, porcine *NLRC4* had canonical exon–intron boundaries, whereas precise exon–intron boundaries in porcine *NAIP* could hardly be determined (Table 2). This result suggests that pseudogenization occurred earlier in NAIP than in NLRC4. Jabir *et al.* showed that intracellular mitochondrial DNA directly or indirectly bound to NLRC4 and activated the NLRC4 inflammasome after infection with *Pseudomonas aeruginosa*; recognition of mitochondrial DNA by NLRC4 is not dependent on NAIP (Jabir et al. 2015). These findings suggest that components different from NAIP contribute to activation of the NLRC4 inflammasome. The possible time lag in pseudogenization between *NLRC4* and *NAIP* in pigs suggests that *NLRC4* retained various functions after the pseudogenization of *NAIP*. Compared with those of humans and other species examined, the porcine genomic sequence upstream of NLRC4 contained an inversion (Fig. 1a). This inversion might prevent the expression of porcine NLRC4, leading to collapse of the genomic structure of *NLRC4*.

Several differences in functional PRRs have been reported among vertebrate species. For example, chickens lack the intracellular nucleotide receptor RIG-I (*DDX58*), which is involved in RNA recognition (Barber et al. 2010). Although they have functional MDA5 (*IFIH1*)—another PRR for RNAs—chickens are highly susceptible to RNA viruses, including highly pathogenic influenza (Karpala et al. 2011).

Whereas *NLRC4* and *NAIP* of pigs are nonfunctional, as shown here, the genome of cattle carries *NLRC4* and *NAIP*, the CDSs of which are estimated to encode functional molecules. Therefore, our accumulated immunological knowledge of animals such as humans and mice cannot necessarily be applied to livestock, as shown by our results and those of these other studies.

In conclusion, here we demonstrated that *NLRC4* and *NAIP* in the pig genome are pseudogenes, resulting in a nonfunctional NLRC4 inflammasome in this species. This finding suggests that pigs lack a key means of monitoring bacterial proteins and clearing intracellular pathogens. This result also suggests that genetic polymorphisms in TLR2, TLR4, TLR5, or other PRRs involved in the recognition of bacteria have a more profound influence on disease resistance in pigs than in other species.

Fig. 1 Comparison of the genomic structures of porcine *NLRC4* and *NAIP* with their human counterparts by dot-plot analysis. Horizontal lines indicate the genomic structures of the genes in the BAC clones isolated (a, L442I04; b, L147A24), which contain the *NLRC4* and *NAIP* genes, respectively. Putative exons corresponding to CDSs of porcine *NLRC4* and *NAIP* and genes in the flanking regions were presumed owing to their similarity to their human counterparts and are indicated by black rectangles. Arrows indicate the transcriptional directions of the genes. Vertical lines indicate the human genomic sequences corresponding to those of the BAC clones. The porcine and human genome sequences were aligned with each other by using the PipMaker program (Schwartz et al. 2000). Gray boxes indicate possible deletions in the porcine genomic sequences in comparison with their human counterparts.

Acknowledgments

This study was supported in part by grants from the Project of Animal Genomics for Innovative Breeding Technology (AGB1002) of the Ministry of Agriculture, Forestry and Fisheries of Japan.

References

Barber MR, Aldridge JR, Jr., Webster RG, Magor KE (2010) Association of RIG-I with innate immunity of ducks to influenza Proc Natl Acad Sci U S A 107:5913−5918 doi:10.1073/pnas.1001755107

Eguchi-Ogawa T et al. (2007) Analysis of the genomic structure of the porcine *CD1* gene cluster Genomics 89:248−261 doi:10.1016/j.ygeno.2006.10.003

Ewing B, Green P (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities Genome Res 8:186−194

Ewing B, Hillier L, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment Genome Res 8:175−185

Franchi L et al. (2006) Cytosolic flagellin requires Ipaf for activation of caspase-1 and interleukin 1β in salmonella-infected macrophages Nat Immunol 7:576−582 doi:10.1038/ni1346

Gordon D, Abajian C, Green P (1998) Consed: a graphical tool for sequence finishing Genome Res 8:195−202

1 Hayashi F et al. (2001) The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5 Nature
2 410:1099−1103 doi:10.1038/35074106
3 Jabir MS et al. (2015) Mitochondrial damage contributes to Pseudomonas aeruginosa activation of the inflammasome
4 and is downregulated by autophagy Autophagy 11:166−182 doi:10.4161/15548627.2014.981915
5 Jessen DL, Osei-Owusu P, Toosky M, Roughead W, Bradley DS, Nilles ML (2014) Type III secretion needle proteins
6 induce cell signaling and cytokine secretion via Toll-like receptors Infect Immun 82:2300−2309
7 doi:10.1128/iai.01705-14
8 Karpala AJ, Stewart C, McKay J, Lowenthal JW, Bean AG (2011) Characterization of chicken Mda5 activity:
9 regulation of IFN-β in the absence of RIG-I functionality J Immunol 186:5397−5405
10 doi:10.4049/jimmunol.1003712
11 Kofoed EM, Vance RE (2011) Innate immune recognition of bacterial ligands by NAIPs determines inflammasome
12 specificity Nature 477:592−595 doi:10.1038/nature10394
13 Miao EA et al. (2010) Innate immune detection of the type III secretion apparatus through the NLRC4 inflammasome
14 Proc Natl Acad Sci U S A 107:3076−3080 doi:10.1073/pnas.0913087107
15 Poyet J-L, Srinivasula SM, Tnani M, Razmara M, Fernandes-Alnemri T, Alnemri ES (2001) Identification of Ipaf, a
16 Human Caspase-1-activating Protein Related to Apaf-1 J Biol Chem 276:28309−28313
17 doi:10.1074/jbc.C100250200
18 Schröder NW, Schumann RR (2005) Single nucleotide polymorphisms of Toll-like receptors and susceptibility to
19 infectious disease Lancet Infect Dis 5:156−164 doi:10.1016/S1473-3099(05)01308-3
20 Schwartz S et al. (2000) PipMaker—a web server for aligning two genomic DNA sequences Genome Res 10:577−586
21 Shinkai H, Matsumoto T, Toki D, Okumura N, Terada K, Uenishi H (2015) Porcine NOD1 polymorphisms with
22 impaired ligand recognition and their distribution in pig populations Mol Immunol 63:305−311
23 doi:10.1016/j.molimm.2014.07.020
24 Suzuki K et al. (2000) Construction and evaluation of a porcine bacterial artificial chromosome library. Anim Genet
25 31:8−12
26 Takenouchi T, Suzuki S, Shinkai H, Tsukimoto M, Sato M, Uenishi H, Kitani H (2014) Extracellular ATP does not
27 induce P2X7 receptor-dependent responses in cultured renal- and liver-derived swine macrophages Results
28 Immunol 4:62−67 doi:10.1016/j.rinim.2014.07.002
29 Tanaka M et al. (2006) Genomic structure and gene order of swine chromosome 7q1.1→q1.2 Anim Genet 37:10−16
30 doi:10.1111/j.1365-2052.2005.01362.x
31 Uenishi H, Shinkai H (2009) Porcine Toll-like receptors: the front line of pathogen monitoring and possible
32 implications for disease resistance Dev Comp Immunol 33:353−361 doi:10.1016/j.dci.2008.06.001
33 Uenishi H, Shinkai H, Morozumi T, Muneta Y (2012) Genomic survey of polymorphisms in pattern recognition
34 receptors and their possible relationship to infections in pigs Vet Immunol Immunopathol 148:69−73
35 doi:10.1016/j.vetimm.2011.07.019
36 Zhao Y et al. (2011) The NLRC4 inflammasome receptors for bacterial flagellin and type III secretion apparatus
37 Nature 477:596−600 doi:10.1038/nature10510

Table 1. Primers used in this study.

Objective	Gene	Primer name	Location [a]		Primer sequence (5′→3′)
			Sscrofa10.2	BAC clone	
BAC clone screening	*NLRC4*	NLRC4-2F	3:114,575,376−114,575,355	L442I04:125,291−125,312	GCTGCCTGATAATTTCTGGTTC
		NLRC4-2R	3:114,575,224−114,575,248	L442I04:125,443−125,419	TCTGAGTAAAAGCTTTCCAGTTTTC
		NLRC4-9F	3:114,520,930−114,520,949	L442I04:149,179−149,198	GGATTTGGTGGGAAATTGTG
		NLRC4-9R	3:114,521,080−114,521,061	L442I04:149,329−149,310	GGATAATGCATGGCCAAGTT
	NAIP	NAIP-4F	16:51,643,443−51,643,462	L147A24:33,299−33,318	GCGAAGCAAAAAGGTTGAAG
		NAIP-4R	16:51,643,656−51,643,637	L147A24:33,512−33,493	TTGCCCAAAAGAAACACACA
		NAIP-17F	16:51,677,503−51,677,522	L147A24:67,295−67,314	CAGGCTTTCAGGCTAACCAC
		NAIP-17R	16:51,677,732−51,677,713	L147A24:67,524−67,505	TGGCAGTGGATAAAGGGAAG
RT-PCR analysis	*NLRC4*	NLRC4-RTF	3:114,571,552−114,571,533	L442I04:129,109−129,128	ATTCTGCGAGAGGGGTCATT
		NLRC4-RTR	3:114,569,963−114,569,984	L442I04:130,695−130,674	TGTCAATATCTTCTCCCAGGGG
	NAIP	NAIP-RTF	16:51,643,749−51,643,768	L147A24:33,605−33,624	AGGCCAGACTCGAATCCTTC
		NAIP-RTR	16:51,655,133−51,655,114	L147A24:44,881−44,862	AGGACTTGGGAGATTGAGCA
	NLRP3	NLRP3-RTF	2:57,856,199−57,856,218	N.A.	CTGGATCTGAGCCACAATGC
		NLRP3-RTR	2:57,857,904−57,857,885	N.A.	GGGAATGATTGCTGCTCAGG
	IL1B2	IL1B2-RTF	3:45,182,013−45,182,032	N.A.	CATCCAGCTGCAAATCTCCC
		IL1B2-RTR	3:45,183,367−45,183,349	N.A.	TGGCATGCTTTTCAAGGACA
Analysis regarding missing exons	*NLRC4*	NLRC4-G1F	N.D.	L442I04:132,426−132,445	GCCTCAGTTCAGTCCTCAGC
		NLRC4-G1R	N.D.	L442I04:132,896−132,875	CTTCGCCTACTGTCTACGTGTG
		NLRC4-G2F	N.D.	L442I04:132,868−132,890	GTGTCCGCACACGTAGACAGTAG
		NLRC4-G2R	N.D.	L442I04:133,381−133,357	ATCACTTGATGATGAAAACAAATCC
		NLRC4-G3F	N.D.	L442I04:133,347−133,371	CTGTAATCTGGGATTTGTTTTCATC
		NLRC4-G3R	3:114,576,758−114,576,734	L442I04:133,919−133,895	ATGTAGGTCAAATTCAGCACATAGG
		NLRC4-G4F	3:114,576,732−114,576,756	L442I04:133,893−133,917	CACCTATGTGCTGAATTTGACCTAC
		NLRC4-G4R	3:114,577,395−114,577,370	L442I04:134,558−134,533	CTCTGCCCTCTCAAGGTACAATAATG
		NLRC4-G5F	3:114,577,362−114,577,389	L442I04:134,525−134,552	AAGTAGATCATTATTGTACCTTGAGAGG
		NLRC4-G5R	3:114,577,635−114,577,608	L442I04:134,798−134,771	CTGTATCTGTGAGTCTCCAGTTTTTCAG
	NAIP	NAIP-GF	16:51,643,715−51,643,734	L147A24:33,571−33,590	AGGGTGGATGACAAAGCAAG
		NAIP-GR	16:51,644,306−51,644,325	L147A24:34,162−34,181	TTTGGTGTGTTTTTCCAACG
	TLR2	TLR2-GF	8:79,826,766−79,826,737	N.A.	TCCTATAATCACTTGTCTAACTTATCATCC
		TLR2-GR	8:79,824,750−79,824,774	N.A.	CTCTCAAATTTAACCAAAACCCTTC

N.A., not applicable; N.D., not detected on the reference genomic sequence.

[a] Locations of the primers on the reference pig genomic sequence (Sscrofa10.2) and sequences of the BAC clones isolated in this study are indicated. Locations on the genomic sequence are shown as chromosome numbers, with base positions relative to the p-ter. Locations on BAC clone sequences are shown with the clone name (L442I04 or L147A24).

Table 2. Nucleotide sequences of splicing sites of *NLRC4* and *NAIP*

Gene	Human exon					Corresponding pig exons			
	Acceptor	Exon number	Location on human chromosome (Sscrofa10.2)		Donor	Acceptor	Location on BAC clone		Donor
			Start	End			Start	End	
			Chromosome 2				Clone L442I04		
NLRC4	TTAC**AG**	2	32,256,893	32,256775	**GT**AAGT	TTAC**AG**	125,330	125,447	**GT**AAAT
	CCTC**AG**	3	32,252,679	32,252,419	**GT**AAGT	CTCC**AG**	128,969	129,225	**GT**ATTA
	TAAT**AG**	4	32,251,601	32,249,607	**GT**ATTG	TAAT**AG**	130,582	132,575	**GT**ACTG
	TGAC**AG**	5	32,241,125	32,241,033	**GT**CAGA	N.D.			
	TTTC**AG**	6	32,238,302	32,238,132	**GT**AACT	N.D.			
	CCCC**AG**	7	32,236,339	32,236,248	**GT**AAGA	N.D.			
	TTCC**AG**	8	32,235,568	32,235,401	**GT**AGGT	TTTC**AG**	134,645	134,813	**GT**AGGT
	ACCC**AG**	9	32,224,765	32,224,449		AATC**AG**	149,136	149,452	
			Chromosome 5				Clone L147A24		
NAIP	TTTC**AG**	4	71,012,918	71,012,348	**GT**ACCT	CACC**AG**	33,094	33,694	**GT**AAGT
	ATAT**AG**	5	71,011,375	71,011,275	**GT**AAGT	N.D.			
	TTAC**AG**	6	71,003,857	71,003,776	**GT**AATG	N.D.			
	CCTC**AG**	7	71,002,143	71,002,092	**GT**AAAA	N.D.			
	TTCA**AG**	8	71,001,817	71,001,698	**GT**GAGT	N.D.			
	TTTT**AG**	9	70,998,823	70,998,724	**GT**GAGT	GTCT**AG**	34,095	34,192	**GT**GAGT
	ACTT**AG**	10	70,987,874	70,987,793	**GT**AATG	N.D.			
	TCAA**AG**	11	70,986,802	70,986,745	**GT**AAGA	N.D.			
	CCAC**AG**	12	70,985,886	70,983,775	**GT**ATAC	ACAA**AG**	44,832	46,705	**GT**AAAC
	TTCC**AG**	13	70,980,036	70,979,869	**GT**AAGG	CAAC**AG**	62,204	62,374	**GT**AAGG
	TTAC**AG**	14	70,977,058	70,976,906	**GT**AAGA	CATCTT	63,492	63,639	**GT**AAGA
	TCAT**AG**	15	70,976,246	70,976,160	**GT**ATGT	N.D.			
	TGAC**AG**	16	70,974,294	70,974,129	**GT**GAGC	TGAT**AG**	65,354	65,513	**GT**GAGC
	TTGT**AG**	17	70,970,468	70,968,483		TTAT**AG**	66,378	68,495	

N.D., The corresponding porcine exon could not be determined.

Human exons comprising the CDSs (*NLRC4*: NM_021209.4; *NAIP*: NM_004536.2) and their corresponding exons in pigs are indicated. Six-nucleotide sequences of the introns at the exon–intron junctions of the human *NLRC4* and *NAIP* genes are shown; the corresponding exons and splicing sites are also indicated. The two highly conserved bases (AG/GT) at the splice sites in the introns are shown in bold.

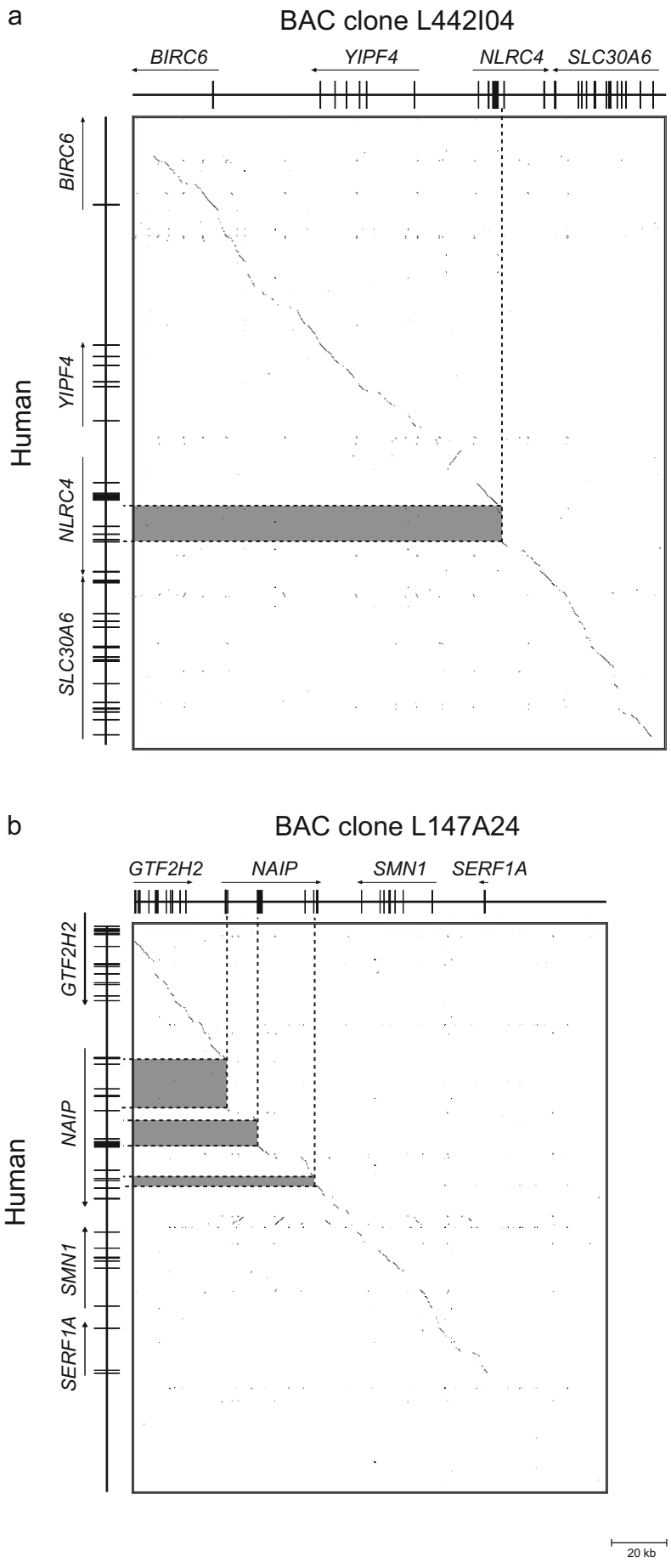

a
BAC clone L442I04
BIRC6
YIPF4
NLRC4
SLC30A6
Human
BIRC6
YIPF4
NLRC4
SLC30A6
b
BAC clone L147A24
GTF2H2
NAIP
SMN1
SERF1A
Human
GTF2H2
NAIP
SMN1
SERF1A
20 kb

Figure 1

Method for supplementary materials

Expression analysis of porcine *NLRC4* and *NAIP* by using semi-quantitative RT-PCR

Macrophages (5×10^5 per well) isolated from a mixed culture of porcine kidney tissue as previously described (Takenouchi et al., 2014) were seeded into 35-mm dishes (Sumitomo Bakelite, Tokyo, Japan) and stimulated with 100 ng/mL flagellin (tlrl-pstfla; InvivoGen, San Diego, CA, USA) or 1 μg/mL lipopolysaccharide (L8274; Sigma-Aldrich, St. Louis, MO, USA) for 8 h. The stimulated macrophages were harvested by treatment with TrypLE Express (Gibco/Thermo Fisher Scientific, Grand Island, NY, USA), and total RNA (1.3−2.1 μg) was extracted from the cells by using RNeasy Mini Kit (Qiagen, Hilden, Germany). In addition, total RNA was prepared from unstimulated macrophages.

The extracted RNA (180 ng per reaction) underwent cDNA synthesis by using Superscript III First-Strand Synthesis System for RT-PCR (Invitrogen, Grand Island, NY, USA) as previously described (Shinkai et al., 2015). The cDNA transcribed from 3.6 ng total RNA was amplified by using primers designed to amplify boundaries of putative exons of porcine *NLRC4* and *NAIP*; in addition, primers for porcine *NLRP3* and *IL1B2* were designed as positive controls (Table 1). Each amplification reaction contained the designed primers (5 μM each), dNTPs (0.2 μM each), AmpliTaq Gold polymerase and the supplied buffer (Life Technologies/Thermo Fisher Scientific, Foster City, CA, USA) according to the manufacturer's instructions. Amplification conditions comprised an initial denaturation at 94 °C for 15 min, followed by various numbers (30 to 50) of cycles of 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 30 s; the reactions underwent a final elongation step at 72 °C for 7 min.

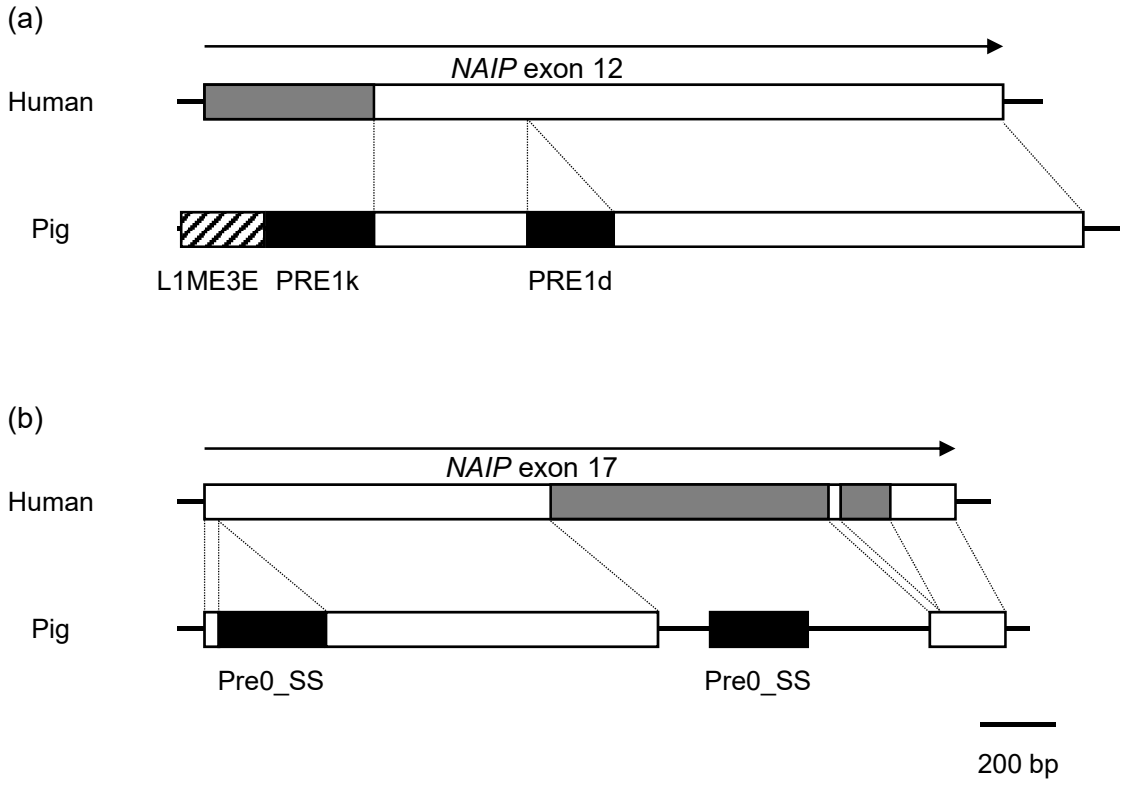
To confirm PCR amplification from the designed primers, we had synthesized (Integrated DNA Technologies, Coralville, IA, USA) double-stranded DNA equivalent to the putative transcripts that would be generated from the genes if they were expressed (*NLRC4*: positions 125,330−125,447, 128,969–129,225, and 130,582–130,706 on BAC clone L442I04 [GenBank: LC144946]; *NAIP*: positions 33,494−33,694, 34,095−34,192, and 44,832−45,032 on BAC clone L147A24 [GenBank: LC144947]; *NLRP3*: exons corresponding to positions 2163−2327, 2328−2498, and 2499−2662 on the pig mRNA sequence [GenBank: NM_001256770.1]; and *IL1B2*: exons corresponding to positions 278−329, 330−525, 526−690, and 691−777 on the pig mRNA sequence [GenBank: NM_001302388.1]). These synthetic transcripts were inserted individually into plasmids (pUCIDT (Amp), Integrated DNA Technologies) for use as positive controls; each PCR reaction contained 100 copies of the positive-control plasmid.

References

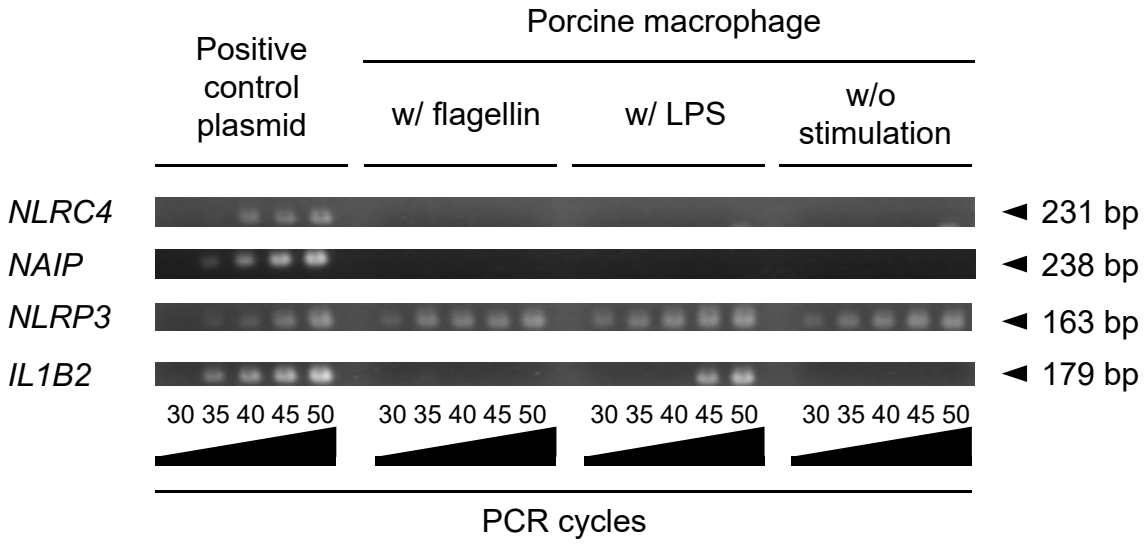
Jurka J (2000) Repbase update: a database and an electronic journal of repetitive elements. Trends Genet 16:418−420

Shinkai H, Matsumoto T, Toki D, Okumura N, Terada K, Uenishi H (2015) Porcine NOD1 polymorphisms with impaired ligand recognition and their distribution in pig populations. Mol Immunol 63:305−311

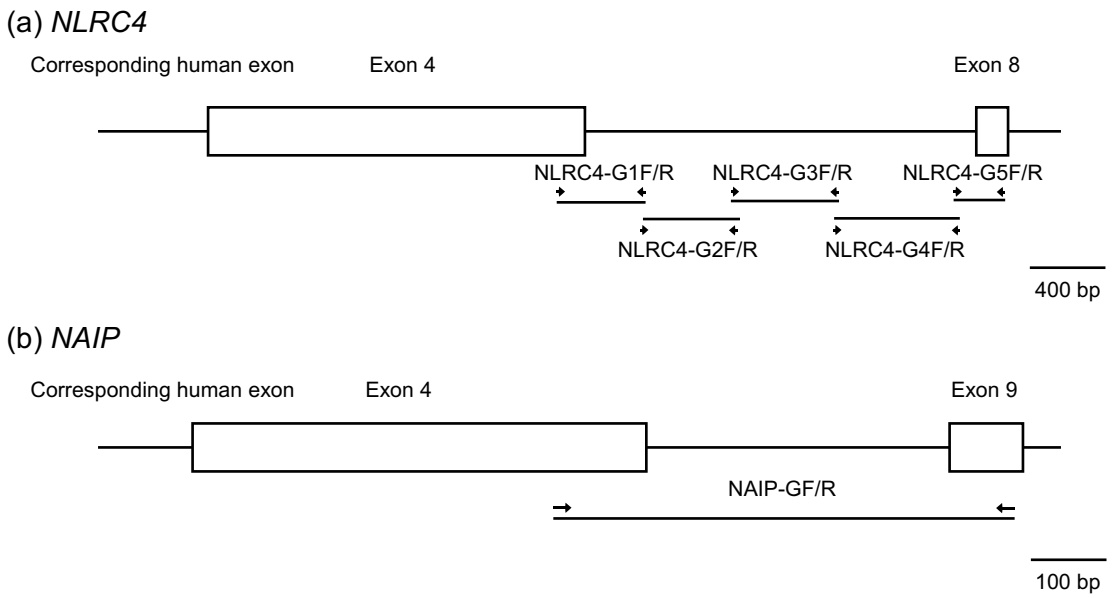
Takenouchi T, Suzuki S, Shinkai H, Tsukimoto M, Sato M, Uenishi H, Kitani H (2014) Extracellular ATP does not induce P2X7 receptor-dependent responses in cultured renal- and liver-derived swine macrophages. Results Immunol 4:62−67



Supplementary Fig. 1. Genomic structure of exons corresponding to human *NAIP* exons 12 and 17. Gray rectangles indicate human exonic sequences that lack marked similarity in the pig genomic sequence. Short- and long-interspersed nucleotide elements (SINEs and LINEs) on the pig genome are indicated by black and diagonal-hatched rectangles, respectively. The SINEs and LINEs were detected by using RepeatMasker (http://www.repeatmasker.org) with the RepBase (Jurka et al., 2000) database. Arrows indicate the transcriptional directions of the human transcripts. (a) More than 400 bp of the 5′-end of the human *NAIP* exon 12 did not have a corresponding sequence in the pig exon; the remaining pig exonic sequence was divided by a SINE sequence. (b) The pig exonic sequence corresponding to human *NAIP* exon 17 was divided by a SINE sequence. Most of the latter half of the human *NAIP* exon 17 did not have correspondence in the pig exon.



Supplementary Fig. 2. Semi-quantitative RT-PCR analysis of expression of porcine *NLRC4* and *NAIP*. Gene expression was investigated by using cDNA synthesized from total RNA harvested from stimulated (flagellin and lipopolysaccharide [LPS]) and unstimulated porcine macrophages. Plasmid DNA (100 copies) carrying the expected amplicons underwent PCR amplification in parallel as positive controls. The number of amplification cycles was varied to estimate the approximate amounts of cDNA by comparison with positive control reactions.



Supplementary Fig. 3. PCR primers designed to examine the genomic structure of the region corresponding to the missing exons of porcine (a) *NLRC4* and (b) *NAIP*. Arrows indicate the designed primers (precise locations are presented in Table 1). Estimated PCR amplicons are shown as lines. All of the pigs that we examined generated PCR fragments of the expected estimated size. Both ends of all fragments were sequenced to confirm correct amplification.