Sequence-based structural analysis and evaluation of polymorphism in buffalo Nod-like receptor-1 gene

S. K. Mishra¹ · P. K. Dubey¹ · Asmita Dhiman¹ · Shubham Dubey¹ · Deepu Verma² · A. C. Kaushik³ · Ravinder Singh¹ · S. K. Niranjani¹ · V. Vohra¹ · K. L. Mehrara⁴ · R. S. Kataria¹

Received: 22 February 2018 / Accepted: 14 December 2018 / Published online: 2 January 2019
© King Abdulaziz City for Science and Technology 2019

Abstract

In this study, we have sequence characterized and analyzed the polymorphism in buffalo *NOD1* (nucleotide-binding oligomerization domain 1) gene as well as its expression analysis. Full-length sequence analysis of *NOD1* revealed this gene in buffalo being conserved with respect to the domain structures, similar to other species. Alternate splice variants having exon3 skipping also identified for the first time in the gene expressed in buffalo-purified peripheral blood mononuclear cells (PBMCs). Phylogenetically ruminant species were found to be clustering together and buffalo displaying maximum similarity with cattle. Sequencing of *NOD1* across 12 Indian buffalo breeds identified 23 polymorphic sites within coding region, among which 16 were synonymous and 7 changes found to be non-synonymous. Four SNPs (single nucleotide polymorphisms) of them were genotyped in 393 animals belonging to 12 riverine, swamp and hybrid (riverine × swamp) buffalo populations of diverse phenotypes and utilities, showing variable allelic frequencies. Principal component analysis revealed, riverine and swamp buffaloes being distinctly placed with the distribution of breeds within the group based on the geographical isolation. Further, quantitative real-time PCR detected *NOD1* expression in multiple tissues with PBMCs and lungs showing highest expression among the tissues examined. Structural analysis based on the translated amino acid sequence of buffalo *NOD1* identified four protein interaction motifs LxxLL important for ligand binding. Molecular interaction analysis of iE-DAP and *NOD1*-LRR and their complex stability and binding-free energy studies indicated variable binding energies in buffalo and cattle *NOD1*. Overall, the study reveals unique structural features in buffalo *NOD1*, important for species-specific ligand interaction.

Keywords *NOD1* · Buffalo · Polymorphism · Alternate splicing · Ligand interaction

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13205-018-1534-2>) contains supplementary material, which is available to authorized users.

✉ R. S. Kataria
katarianji@yahoo.co.in

¹ ICAR-National Bureau of Animal Genetic Resources, GT Road Bypass, Karnal 132001, Haryana, India

² ICAR-National Dairy Research Institute, Karnal 132001, Haryana, India

³ Present Address: School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai 200240, China

⁴ ICAR-Central Institute for Research on Buffaloes, Nabha Campus, Bir-Dosanjh, Nabha, Punjab, India

Introduction

Among higher eukaryotes, the conserved microbial structures known as microbe-associated molecular patterns (MAMPs) are recognized by germ line-encoded pattern recognition molecules (Koller et al. 2009). Four kinds of germ line-encoded receptors are involved in pathogen control and elimination, which include the Toll-like receptors (TLRs), the Nod-like receptors (NLRs), the retinoic acid-inducible gene-1 (RIG-1)-like receptors (RLRs) and the C-type lectin receptors (CLRs). Individual molecules recognize pathogen-derived molecular components from all major classes of microbes, including bacteria, viruses, yeast and parasites. These microbial components include lipopolysaccharides (LPS), lipopeptide, peptidoglycan (PGN), flagellin, dsRNA, ssRNA and CpG DNA, among others (Takeda and Akira 2005). Identification of these MAMPs by pattern

recognition molecules triggers the activation of signaling pathways including the nuclear factor κ B (NF- κ B), mitogen-activated protein kinases (MAPKs) and type I interferon (IFN) molecules, which promote inflammatory and antimicrobial responses (Kawai and Akira 2009).

Nod-like receptors are implicated in the induction of inflammatory responses through activation of nuclear factor κ B (NF- κ B), upon stimulation by other cellular danger signals (Ogura et al. 2001). Characteristic of NLR proteins is the presence of a NACHT (NAIP-neuronal apoptosis inhibitory protein), which is flanked by highly variable leucine-rich repeat (LRR) domain, play a critical role in the sensing of Gram-negative peptidoglycan. NOD molecules, a family of intracellular proteins including *NOD1* (CARD4) and *NOD2* (CARD15), represents a new group of intracytoplasmic pattern recognition molecules, thus allowing detection of intracellular invasive bacteria. *NOD1* recognizes *g*-D-glutamyl-meso-diaminopimelic acid (iE-DAP) which is found to be part of PGN in all gram-negative bacteria as well as some gram-positive bacteria (Chamaillard et al. 2003). In particular, *NOD1* expressed in epithelial cells considered as the first line of defense against invasive pathogens has been reported to play an important role in sensing the intracellular Gram-negative bacterial cell wall components (Girardin et al. 2003). A study has shown that *Helicobacter pylori* introduces iE-DAP (combat with *NOD1*) into the cytosol using its type IV secretion system (Viala et al. 2004). Inflammatory cytokines and other antimicrobial genes which contribute to host defense, are induced by signaling via *NOD1* and *NOD2* (reviewed by Mogensen 2009). Mutations in *NOD1* gene are strongly associated with ischemic stroke (Tiszlavicz et al. 2009) and with susceptibility to bowel disease (McGovern et al. 2005) in humans. Various splice variants of *NOD1* have also been identified in human beings with the truncated protein lacking number of LRR domains across tissues (Hysi et al. 2005).

The dairy industry of India largely depends on the 108.7 million buffaloes (Livestock Census, 2012, <http://dahd.nic.in/dahd/bahs-2012.aspx>), which account for 56.7% of the world's buffalo population, contributing about 52.5% to total milk production in the country. Besides milk, buffaloes are an important source of meat, leather and draft power. Association of polymorphism in immune response receptor genes with several infectious diseases has been reported in bovines, including mastitis in Indian dairy buffaloes (Dhiman et al. 2017). Hence, it is expected that the investigation of genetic diversity in immune genes, particularly pathogen-binding receptors that play a key role in innate immune defense, would shed focus on the disease susceptibility or resistance in riverine and swamp buffaloes.

Among the microbial-sensing receptor molecules, NOD-like receptors have emerged as potential candidates for investigating the genetic structure and polymorphism

analysis. While the work focused on the TLRs' role in innate resistance of buffalo is sparse (Dubey et al. 2012), very less information is available on riverine and swamp-type buffaloes related to characterization of immune response elements like *NOD1*. In this study, we report cloning-based sequence characterization of buffalo *NOD1* gene as well as tried exploring the polymorphism and generating gene expression profile across tissues. Sequence-based structural analysis of *NOD1* data available for bovine species has also been attempted to investigate the possible role of this gene in species-specific immune response.

Materials and methods

Sample collection

Genomic DNA isolated from blood samples of 393 buffaloes was used in the study, which comprised 221 riverine types—Murrah (30), Nili Ravi (34), Toda (47), Marathwadi (46), Jaffarabadi (30) and Chilika (34) and 48 Assamese hybrid buffaloes and 124 pure swamps belonging to Assamese Silchar (27), Assamese Dibrugarh (12), Nagaland (36), Mizoram (15), Manipuri (34). Blood samples collected from 12 different buffalo breeds, including 6 riverine buffaloes and 6 swamps, representing different utilities and geographical locations from India, were utilized for polymorphism detection. Necessary approval from the Institute Animal Ethic Committee was taken for the blood samples' collection. The genomic DNA was isolated using the standard SDS-proteinase K method (Sambrook and Russell 2001) and the quality/quantity of DNA was assessed using a Nanodrop (ND-1000).

Sequencing and SNP detection in buffalo *NOD1* gene

For the characterization of buffalo *NOD1* gene, a total of 2865 nucleotides comprising ORF region including 11 exons were amplified, and 15 sets of overlapping primers were designed from gene sequence of cattle sequence available in Ensembl genome browser (ENSBTAT000000061183), for PCR amplification (Table 1-SI in supplementary information). PCR was performed on DNA samples of panel representing 12 pooled samples of 3 animals from each breed above, in a total volume of 20 μ l containing approximately 80 ng genomic DNA, 10 $\times$ PCR buffer having 15 mM $MgCl_2$, 0.5 μ l of 10 mM dNTPs, 10 pmol of each primer and 1 U of Taq DNA polymerase (Bangalore Genei, India). Following the initial denaturation step (95 $^{\circ}$ C for 3 min), samples were subjected to 32 cycles of PCR consisting of 94 $^{\circ}$ C for 30 s, primer-specific annealing temperature for 30 s (Table 1-SI in supplementary information); and 72 $^{\circ}$ C for 1 min, followed by a final extension for 10 min at 72 $^{\circ}$ C. Amplified products

were checked on 1.5% agarose gel and sequenced from both ends after column purification. Sequencing of purified PCR products was carried out using BigDye[®] Terminator v3.1 Cycle Sequencing Kit on ABI PRISM[®] 3100 Genetic Analyzer (Applied Biosystems, CA, USA). Data was further analyzed using SeqMan and MegAlign program of the LaserGene software (DNASTAR Inc., USA). Polymorphic sites were documented by sequence alignment and crosschecking the chromatograms of individual animal.

Genotyping of polymorphic nucleotide by PCR-RFLP and tetra-ARMS PCR

Identified SNPs, were further analyzed for their position, structural domain location and nucleotide substitution status. A PCR-RFLP genotyping protocol was developed for the three non-synonymous SNPs—(g.338C>T, g.797A>G and g.1482G>C), using the NEBcutter online software (<http://tools.neb.com/NEBcutter2/>). Primers were designed and a tetra-primer Amplification Refractory Mutation System (ARMS)-PCR procedure (Table 1-SI in supplementary information), was also used for genotyping of one synonymous SNP-(g.2394A>C) across different buffalo breeds. PCR-RFLP genotyping of these three SNPs-g.338C>T, g.797A>G and g.1482G>C were carried out using AciI, MspI and PvuII restriction enzymes (NEW ENGLAND BioLabs[®] Inc.), respectively. Amplified PCR products were digested with the respective restriction enzyme and resolved in 3% agarose gel and genotypes were recorded. Genotypic frequency and genetic distance (DA, Nei's standard genetic distances) were calculated using GenAlex6.2 program (Peakall and Smouse 2006, <http://www.anu.edu.au/BoZo/GenAlEx/>). Genetic distance was determined based on allelic frequencies at three loci in buffalo *NOD1* gene. Principal component analysis (PCA) was performed using GenAlex6.2 program, to determine population relationships directly based on the variability in allele frequencies of different genotypes taken as the largest possible variance (Manly 1986).

Sequencing and characterization of buffalo *NOD1* cDNA

To amplify buffalo *NOD1* transcript, PCR primers specific to amplify complete ORF were designed and got synthesized (Table 1-SI in supplementary information). Total RNA was extracted from PBMCs of Murrah buffalo, using TRIzol method (Invitrogen, Bangalore, India), and subjected to further purification and on-column DNase digestion using RNeasy mini kit (Qiagen India). Concentration of isolated RNA was determined spectrophotometrically using the NanoDrop ND-1000 (Thermo Scientific). Good-quality samples with optical density OD260/OD280 ratios within the range

of 1.9–2.1 were used further for reverse transcription-PCR. cDNA was synthesized from 1 µg of total RNA using RevertAid First Strand cDNA Synthesis kit (Fermentas, Thermo Fisher Scientific, India), with the help of Oligo (dT) primers, as prescribed by the manufacturer. PCR was performed using 3 µl of cDNA in a 50 µl reaction volume under the conditions of one cycle of initial denaturation at 95 °C for 3 min followed by 32 cycles of 94 °C for 30 s, 60 °C for 30 s, 72 °C for 75 s and a final extension at 72 °C for 10 min.

PCR product was analyzed in 1.5% agarose gel, the single specific 2.9 kb band was gel purified by agarose gel purification kit (Roche, Germany) and cloned using InsTAclone PCR Cloning Kit (Thermo Scientific), following manufacturer protocol. Both strand sequencing was carried out with M13 forward and reverse as well as gene-specific internal (Table 1-SI in supplementary information) primers (ABI prism 3000). Overlapping sequences of different fragments were converted into single ORF coding and analyzed further. Molecular weight and an isoelectric point (pI) of buffalo *NOD1* was analyzed using ExPASy software on the Proteomics Server (https://expasy.org/tools/pi_tool.html) and the cellular location was predicted by pSORT (<https://www.worfpSORT.org>). An online program SMART was used for the analysis of protein domain architecture of buffalo *NOD1* gene. Online software PolyPhen-2 (<https://genetics.bwh.harvard.edu/pph2/>) was used to predict the possible impact of non-synonymous SNPs on the protein structure. PCR amplification of cDNA synthesized from RNA isolated from PBMCs of one animal showed amplification of two fragments of 2.9 kb and 1.2 kb of equal intensity, supposed to be splice variants. To confirm the lower size band to be from a splice variant, cloning and sequencing of both the fragments was performed.

Tissue distribution and expression analysis of buffalo *NOD1* gene

Quantitative real-time PCR was performed for comparative analysis of *NOD1* gene expression across ten adult (4–6 years) buffalo tissues—small intestine, lung, liver, heart, mammary gland, stomach, kidney, spleen, skeletal muscle, and PBMCs, collected from three animals each representing Murrah breed. RNA was subsequently isolated and checked for concentration and purity as described above using Nanodrop ND-1000 (Thermo Fisher Scientific). Quality of RNA samples was further checked by Experion Bioanalyzer (BioRad Laboratories, CA, USA). RNA samples with RIN (RNA integration number) value of more than 8.0 were used further for gene expression analysis. An amount of 1 µg of total RNA from each tissue sample was reverse transcribed into cDNA using RevertAid First Strand cDNA Synthesis Kit (Fermentas, Thermo Fisher Scientific, India), using Oligo(dT) primers as described previously.

NOD1 gene-specific primers were designed and used for real-time-based transcriptional profiling (Table 1-SI in supplementary information). Primer pair was tested for specificity through amplification of cDNA synthesized from PBMCs through normal PCR amplification. Real-time PCR was performed using SYBRGreen-based detection on a LightCycler480 (Roche) machine, with each 20 µl reaction containing 2 µl of cDNA synthesized from 1 µg of RNA, 0.5 µl (5 pmol) of sense and antisense primers, 2×SYBR-Green qPCR Master Mix (USB, OH, USA). The mixture was subjected to the following real-time PCR conditions: 95 °C for 10 min; followed by 40 cycles of 95 °C for 15 s and 60 °C for 1 min. Each reaction was performed in duplicate, using two internal control genes ribosomal protein S9 (RPS9) and ubiquitously expressed prefoldin-like chaperone (UXT) for normalization. PCR efficiency was determined from the slope of a standard curve generated using a tenfold dilution series of the cDNA template. The data of target *NOD1* gene was normalized against geometric mean of two reference genes. Fold changes in target transcript levels were determined using the delta Ct method (Pfaffl 2001). $\Delta\Delta C_t$ values were calculated using kidney tissue, showing lowest expression (highest ΔC_t) as control.

Sequence-based homology modeling

To find suitable templates for the modeling of consensus, river, swamp and cow *NOD1* protein sequence, BLASTp search was performed against the protein databank (PDB) (<https://www.rcsb.org/pdb/>). Best scoring template 5IRL resulting from fold recognition servers were considered for model construction using MODELLER 9.12. Both single-template and multiple-templates approaches were adopted for building *NOD1* protein of each species. The 3D models were also constructed using the automated modeling programs, SWISS-MODEL, BioSerf, and Easy Pred 3D for comparison. The MODELLER-derived 3D models of buffalo *NOD1*-LRR were evaluated based on discrete optimized protein energy (DOPE) scores. Five models with lowest DOPE scores were chosen for further refinement using AMBER force field.

Comparative modeling, validation and annotation of NOD1

The complete 3D structure of 954 amino acids sequence of river, swamp and cow *NOD1* protein, (NCBI: gbl AJC01625.1, AJC01625.1 and NP_001243492 respectively) was predicted. The predicted structure was further visualized in PyMol program. The stereochemical quality of the final energy-optimized model was verified using SAVE (<http://nihserver.mbi.ucla.edu/SAVE/>), bond length and bond angle analyses of the model which was performed in MolProbity.

Furthermore, the Z score of hydrogen bond (H-bond) energy, packing defects, bump score, radius of gyration and deviation of Y angles of the refined model was tested in VADAR, GeNMR and PROSESS web servers. Furthermore, the PROCHECK analysis, used for the validation of the modeled structure, yielded the corresponding Ramachandran plot with 95% residues being noticed in the most favored regions.

Molecular docking of NOD1-LRR with iE-DAP

The 2D structure of iE-DAP (CID_45480617) was obtained from the PubChem database (<http://pubchem.ncbi.nlm.nih.gov>). The Automated Topology Builder (ATB) server was used to generate the 3D coordinates of iE-DAP employing energy minimization with the Gromos 96-53a6 force field. Molecular docking was carried out using AutoDock 4.2 to identify key ligand-binding residues. The *NOD1* model was assigned with Kollman charges and Geister partial charges, that were applied for iE-DAP using Autodock Tools 1.5.

Results and discussion

The innate immune system represents one of the first lines of defense against microorganisms, through the specific detection of highly conserved structures. NOD-like receptor1 is one of the best known cytosolic innate immune receptors of the NLR family in vertebrates (Franchi et al. 2009). NLR1 in human, piscine and pig is very well-characterized (Inohara et al. 1999; Tohno et al. 2008; Sha et al. 2009; Xie et al. 2013), but the polymorphism detection and its association with disease incidences in buffalo *NOD1* gene have not been reported so far.

Detection of polymorphism in buffalo NOD1

Sequencing of 15 overlapping PCR fragments ranging from 400 to 700 bp revealed 2.9 kb nucleotide sequence of buffalo *NOD1* gene. The ORF of buffalo *NOD1* gene consisting of 2865-nucleotides long region comprising 11 exons, was found to be highly conserved across livestock species. The intron–exon boundaries were assessed by comparing the genomic sequence data obtained with buffalo *NOD1* mRNA sequence (Accession no. KP280085) as well as predicted *NOD1* gene sequence of Mediterranean buffalo from whole genome shotgun data (accession no. NW_005784709), available in GenBank. By comparative analysis, buffalo *NOD1* nucleotide and amino acid sequence, respectively, showed the highest identity with that of cow, followed by goat, sheep, pig, horse, human and mice with being the most distant (Table 1). To examine the nucleotide genetic diversity for single nucleotide polymorphisms (SNPs) across breeds of buffalo *NOD1*, sequencing of different fragments

Table 1 Comparative analysis of buffalo *NOD1* structural features and homology with different species

Species	Accession no.	CARD	NACHT	No. of LRR	No. of AA	Length of ORF	Percent homology nucleotide	Percent homology amino acid
Buffalo	KP280085	85	173	8	954	2865	—	—
Cow	NM_001256563	85	173	8	954	2865	98.2	99.2
Goat	XM_005679253	85	173	8	954	2865	97.4	97.6
Sheep	XM_012112248	85	173	8	954	2865	97.4	97.3
Pig	AB187219	85	173	7	953	2862	89.2	88.3
Horse	XM_001499566	85	173	7	953	2862	87.7	86.1
Mice	AY160222	86	173	7	953	2862	79.4	77.8
Human	AF113925	85	173	7	953	2862	86.2	83.1

generated from various buffalo breeds were sequenced. Comparative analysis of 2865 nucleotides long buffalo *NOD1* sequences covering the entire coding region, acquired for a panel of 12 animals of 6 swamp and riverine buffaloes, revealed the presence of 23 polymorphic sites, including 16 transitions and 7 transversions, among which 16 SNPs were synonymous, and 7 were non-synonymous (Table 2; Fig. 1). Shinkai et al. (2015) tried to assess the polymorphism in pigs *NOD1* and identified 21 SNPs deciphered 9

non-synonymous SNPs localized throughout coding region and also in sequences encoding domains important for ligand recognition. Recently in buffaloes, 42 SNPs in the *NOD2* transcript (homolog of CARD4/*NOD1*) have been reported by Dhiman et al. (2017).

Among the seven non-synonymous SNPs, three were predicted to lie in functionally important domains: aspartic acid to asparagine at position 66 (Asp66Asn) was in CARD domain, glycine to glutamic acid at position 266

Table 2 Polymorphic nucleotides detected in buffalo *NOD1* with their position and amino acid changes

Location	SNP	Nucleotide position as per accession no. KP280085	Transition/transversions	Syn/non-syn	Amino acid change	Domain region
Exon-1	A/G	113	Transition	Syn	Gln38Gln	CARD
	G/A	196	Transition	Non-syn	Asp66Asn	
Exon-2	G/C	312	Transversion	Syn	Leu104Leu	Unknown
	C/T	338	Transition	Non-syn	Pro113Leu	
Exon-3	A/G	556	Transition	Non-syn	Thr186Ala	Unknown
	C/A	606	Transversion	Syn	Gly202Gly	
	C/T	795	Transition	Syn	Pro265Pro	NBD
	G/A	797	Transition	Non-syn	Gly266Glu	
	C/T	990	Transition	Syn	Leu330Leu	Unknown
	G/C	1137	Transversion	Non-syn	Leu379Asp	
	C/T	1275	Transition	Syn	Phe425Phe	Unknown
	C/T	1362	Transition	Syn	Arg454Arg	
	C/T	1365	Transition	Syn	Ala455Ala	Unknown
	C/T	1428	Transition	Syn	Phe476Phe	
	G/C	1482	Transversion	Non-syn	Gln494His	Unknown
	G/A	2016	Transition	Syn	Ala672Ala	
	C/T	2058	Transition	Syn	Cys686Cys	LRR
	C/T	2316	Transition	Syn	Val772Val	
Exon-6	A/C	2394	Transversion	Syn	Thr798Thr	LRR
Exon-7	C/T	2475	Transition	Syn	Ile825Ile	
Exon-9	C/T	2649	Transition	Syn	Asp883Asp	LRR
	G/C	2661	Transversion	Non-syn	Glu887Asp	
Exon-11	A/C	2832	Transversion	Syn	Val944Val

Fig. 1 Schematic representation showing structural placement of non-synonymous SNPs identified in the buffalo *NOD1*

(Gly266Glu) in the NACHT and glutamic acid to aspartic acid at position 887 (Glu887Asp) was found in LRR7 domain, while four SNPs proline to leucine at position 113 (Pro113Leu), threonine to alanine at position 186 (Thr186Ala), leucine to aspartic acid at position 379 (Leu379Asp) and glutamine to histidine at position 494 (Gln494His) were detected in the region of unknown function. PolyPhen-2 analysis was performed on translated amino acids to comprehend the probable impact of nsSNPs on the function and structure of buffalo *NOD1* protein (Fig. 2). PolyPhen-2 analysis suggested that two

amino acid changes, Asp66Asn and Glu887Asp, possibly damaged the protein structure and function of functionally important CARD and LRR domains, respectively. These changes may impact the overall structure and biological function of the protein, which needs to be explored further. Earlier, polymorphism within or adjacent to the LRR muramyl dipeptide recognition domain of *NOD2* has been found to associated with altered NF- κ B activation and this increased the susceptibility to Crohn's disease (Hugot et al. 2001).

Fig. 2 Non-synonymous SNPs in buffalo *NOD1* analyzed by Polyphen-2 for their sensitivity and specificity score and the possible/probable damaging effect on protein structure and function

Genotyping and allele frequency distribution

Genotyping of three polymorphic sites, SNPs—g.338C>T (AciI) transition (Pro113Leu) in *NOD1* exon-2, g.797A>G (MspI) transition (Gly266Glu) and g.1482G>C (PvuII) transversion (Gln494His) in exon 3 was carried out using restriction fragment length polymorphism (RFLP), while one synonymous SNP—g.2394A>C found in functionally important leucine-rich repeat-4 (LRR4) domain analyzed by tetra-ARMS PCR in a total of 393 buffaloes was represented by different breeds/populations. These SNPs were selected based on the type of amino acid changes (mostly hydrophobic to hydrophilic or charged) resulting in the non-synonymous nature of SNPs and the possibility of designing simple RFLP or tetra-ARMS PCR genotyping protocols. Variable allelic frequencies were observed across riverine and swamp buffaloes at all the four loci. SNP variation g.1482G>C was observed only in swamp buffaloes with an overall frequency of 0.65 for G and 0.35 for C allele, whereas in riverine and hybrid animals, this polymorphic site was fixed for G allele (Table 3). Though premature to say, these variable allelic frequencies could be responsible for the variation in immune response to diseases as reported earlier (Mingala et al. 2009). Individual frequencies of the genotypes were in the Hardy–Weinberg equilibrium for most of the SNPs, as determined by Chi square test. The calculated Chi square values ranged from 0.02 to 34.00 (Table 2-SI in supplementary information). *NOD1* polymorphism has been demonstrated

to modulate pro-inflammatory cytokine responses induced through Toll-like receptor or Nod-like receptor ligands in rheumatoid arthritis-affected human beings (Plantinga et al. 2013). Principal component analysis based on genetic distance among haplotypes of four SNPs examined also showed classification of different riverines, swamps and their hybrid buffaloes into distinct quadrants. Within riverine and swamp types, classification was in correspondence with the geographical distribution of breeds (Fig. 3). The results further supported the distinct allelic distribution among riverines,

Fig. 3 Genetic distance-based principal component analysis on the basis of haplotypes of four SNPs (g.338C>T, g.2394A>C, g.797A>G, g.1482G>C) genotyped and examined between buffalo breeds [ASW—Assam (riverine and hybrid), DB—Dibrugarh (Assam swamp), SL—Silchar (Assam swamp), NGL—Nagaland (swamp), MN—Manipur (swamp), MZ—Mizoram (swamp). Riverine buffaloes include—MU—Murrah, JAFF—Jaffarabadi, MTW—Marathwadi, CH—Chilika, TD—Toda, NR—Nili Ravi]

Table 3 Allelic frequencies of the genotyped SNPs (g.338C>T, g.797A>G, g.1482G>C, g.2394A>C) in the coding region of *NOD1* gene of Indian riverine and swamp buffaloes

Breed ^a	<i>N</i> (animals)	g.338C>T		g.797A>G		g.1482G>C		g.2394A>C	
		C	T	A	G	G	C	A	C
Swamp									
DB	12	0.96	0.04	0.63	0.38	0.96	0.04	0.96	0.04
SL	27	0.94	0.06	0.82	0.19	0.88	0.11	0.87	0.13
NGL	36	0.97	0.03	0.83	0.17	0.74	0.26	0.86	0.14
MN	34	1.00	0.00	0.96	0.04	0.38	0.62	0.99	0.02
MZ	15	1.00	0.00	0.83	0.17	0.30	0.70	1.00	0.00
Swamp overall	124	0.97	0.03	0.81	0.19	0.65	0.35	0.93	0.07
Hybrids									
ASW	48	0.83	0.17	0.47	0.53	1.00	0.00	0.27	0.73
Riverine									
MU	30	0.78	0.22	0.53	0.47	1.00	0.00	0.43	0.57
JAFF	30	0.65	0.35	0.48	0.52	1.00	0.00	0.13	0.87
MTW	46	0.72	0.28	0.30	0.69	1.00	0.00	0.54	0.46
CH	34	0.82	0.18	0.43	0.57	1.00	0.00	0.46	0.54
TD	47	0.85	0.15	0.20	0.79	1.00	0.00	0.59	0.40
NR	34	0.85	0.15	0.60	0.39	1.00	0.00	0.44	0.56
Riverine overall	221	0.77	0.23	0.42	0.58	1.00	0.00	0.43	0.57

^aASW Assam (hybrids), DB Dibrugarh (Assam-swamp), SL Silchar (Assam swamp), NGL Nagaland (swamp), MN Manipur (swamp), MZ Mizoram (swamp). Riverine buffaloes include—MU Murrah, JAFF Jaffarabadi, MTW Marathwadi, CH Chilika, TD Toda, NR Nili Ravi

		CARD										
Buffalo	MEKHGCSKMEIVPSESYSFIKLLKVNREHLVTHIRNTQCCLLDNLLQNDYFSTEDAEIVCACTPQDPKVRRLILDV											75
Cattle	H.....N.....											75
Goat	...R.E.M...PH.....											75
Sheep	...R.E...PH...S.....											75
Human	..EQ.H.E...I...HPH.Q...S...L.....V...K...A.....K.....											75
Buffalo	QSKGEEVSEFFLHVIQQLADAYVDLRPWLSEIGFSPSPLIQSKAVVNTDPVSRYSQKLQQQLGRDSKFILCYAQK										150	
Cattle	G.....											150
Goat	Q.....											150
Sheep	S...K.....Q.....											150
Human	YLL...L...L.T...V...T.Q.RHH...V.....											150
Buffalo	EELLEQMYTDTTVVELVNFNSNESLGSLSLACLILDASTGVLINEQGEIIFVFGDAGMGKSMLLQRLQSLWAAGQLD											225
Cattle	T.....											225
Goat	T.....											225
Sheep	EI.M...IM...G...N...HT...I...T...IL...V.....T.R.....											225
Human	EI.M...IM...G...N...HT...I...T...IL...V.....T.R.....											225
Buffalo	PGIKFFFHFRCTLSCFKKSAALCLQDLFLKHYCYPEQDPPEEVFAFLRRFPHTALFTFDGLDELHSDFDLSSEPD											300
Cattle	G.....											300
Goat	G.....											300
Sheep	G.....											300
Human	A.V.....MF...E.DR.....R.....V.....L...RV.....											300
Buffalo	TSSPWEPAHPLALLASLLSGKLLKGARKLLTARTGVVPRHLLWKKVLLRGFSPSHLRAYTGRMPPEHAVRQRL											375
Cattle	A.....R.....											375
Goat	A.....R.M.....											375
Sheep	S.C.....V...N...S...I...QF.R.....AR...R.LQD.....											375
Human	S.C.....V...N...S...I...QF.R.....AR...R.LQD.....											375
Buffalo	DQLEANPNLCSLCAVPLFCWIIIFRCFQHFHDAFESSQLPDLTVTLTDIFLLVTEVHLNRTQPTSLVQNTSRST											450
Cattle	T.....											450
Goat	T.....											450
Sheep	S.....S.....RA...G...C.M...V.....M...S...R...PV.....											450
Human	S.....S.....RA...G...C.M...V.....M...S...R...PV.....											450
Buffalo	ETLRAGRDTLCSLGRVAHRGMEKSLFVFNQDEVQASEVREGDLQLGFLRTMPELGLGGEQQSYEFFHLLTLQAFFA											525
Cattle	V...E.....											525
Goat	V...E.....											525
Sheep	H.....Q.....T.E...GLQ.R.M...AL...P...D.....T.....											525
Human	H.....Q.....T.E...GLQ.R.M...AL...P...D.....T.....											525
Buffalo	ALFLVVDKVGTRRELLRFFQEWTPPEEAVATSTCYPPFLPVQCLGGHGLHGDEPFKNKDHFFHTNLFCLGLLSKG											600
Cattle	H...D.....											600
Goat	F.....R.....											600
Sheep	F.....R.....											600
Human	..F...L...R...Q.....M...-AGA...TS...F...Q.S.PAR...L...Q.....A.....											599
Buffalo	KQKLLQHLVPTAALRRRKALWTHLFASLRSHLKSLEPRVQSGGFNOVQALPTFIWMLRCIYETQSEKVGQLAARG											675
Cattle	R.....N...S...F...A.....											675
Goat	R...T...N...S...F...A.....											675
Sheep	R...A...A...S...GY...VES...M...Q.....											675
Human	R...A...A...S...GY...VES...M...Q.....											674
Buffalo	ICANYLKLTFCNACSADCSALSFLVHHLRKRLALDLNNDNFVRELQPCFSFLTVIRLSVNQITDSGVKVL											750
Cattle	Y.....FP.....Y...L...G...S.....											750
Goat	Y.....FP.....Y...L...G...S.....											750
Sheep	Y.....FP.....Y...L...G...S.....											750
Human	Y.....FP.....Y...L...G...S.....											749
Buffalo	EELTKYKILTFGLYNNQITDVGARYIARILDKCKGLTHLKLGNKITSEGGKSLALAVKNSKSI FEVGMWGNQI											825
Cattle	G...E.....											825
Goat	G...E.....											825
Sheep	V...Y...VTK...F...Y...S...V.....											825
Human	V...Y...VTK...F...Y...S...V.....											824
Buffalo	GDEGAKAFAEALRNHPSLTNLSLAFNGISTEGGKSLAWALQCNASLRIFWLTKNELDDEVAESFAEMLKVNQTLK											900
Cattle	T...S...R...T...E...L...Q...N...L.....											900
Goat	T...S...R...T...E...L...Q...N...L.....											900
Sheep	T...S...R...T...E...L...Q...N...L.....											900
Human	T...S...R...T...E...L...Q...N...L.....											899
Buffalo	HLWLIQNQITAKGIAQLAEALQKNTGIMEICLNGNLKPEEAKVFEDEKRIVCF											954
Cattle	I.....											954
Goat	I.....											954
Sheep	T...D...S...T...Y...I.....											954
Human	T...D...S...T...Y...I.....											953

Fig. 4 Alignment and architecture of trans-species conservation of deduced amino acid sequences of *NOD1*. Amino acid sequences of buffalo, cow, goat, sheep, and human *NOD1* were aligned. Shaded sequences indicate LxxLL motifs, while numbers indicate the amino acid position. Similar residues are marked with dots. Gaps were introduced to optimize alignment. The domain architecture is highlighted under boxes with relevant stretch of sequence as follows: CARD (green strip), NACHT (blue strip), LRRs (boxed). Within NACHT domain, WalkerA (GKS) at position 207–209 and WalkerB (DGLDE) 284–288 sequences are shown in boxes

swamps and their hybrids, at these four loci, which could be responsible for the variable immune response to pathogens. Similar findings have been reported with other innate immune response genes also (Dubey et al. 2013a).

Sequencing and characterization of buffalo *NOD1* transcript

Sequencing of full-length cDNA revealed buffalo *NOD1* gene comprising of an open reading frame (ORF) of 2865 nucleotides, encoding a polypeptide of 954 amino acids found to similar in size to that of other ruminants. The buffalo *NOD1* protein had a theoretical molecular weight of 107 kDa and an isoelectric point (pI) of pH 6.86, and the cellular location was predicted to be cytoplasmic as reported in other species (Kufer et al. 2008). Structurally, buffalo *NOD1* comprised of one caspase recruitment domain (CARD, residues 20–104) at NH₂ terminal, involved in interaction with RIP2, leading to the downstream signals for the activation of NF- κ B and MAPKs pathways (Viala et al. 2004; Maharana et al. 2017). Centrally located NACHT domain (residues 196–368) and eight leucine-rich repeat (LRR) motifs at C-terminal were conserved across all ruminant species including buffalo. Further, within NACHT domain, two functionally important motifs WalkerA (GKS) at position 207–209 and WalkerB (DGLDE) 284–288 were found to be conserved across species (Fig. 4). Mutations within these regions have been found to be associated with several genetic diseases in human (Zurek et al. 2012).

Buffalo *NOD1* cDNA, amplification from buffalo PBMCs revealed, apart from amplification of expected 2.9 kb full-length transcript, the generation of a smaller product of 1.2 kb in length, with equal intensity in agarose gel (Fig. 5a). Cloning and sequencing of both products divulged a smaller product being generated due to synthesis of a splice variant transcript, having exon3 (1084 nucleotides of amplified cDNA, 1037 nucleotides considering only the ORF) spliced out (Fig. 5b). Expression of this novel splice variant of *NOD1* has been identified in ruminants for the first time. This variant with a deletion of exon3 region lacks the functionally important NACHT as well as downstream LRR domains and had an early stop codon resulting in the expression of a truncated protein (Fig. 5c). In human beings, a total

of eight *NOD1* splice variants has been reported. Cryptic splice sites contribute to resulting in altered protein translation, stability, and expression of *NOD1* isoforms, which might be the link between several *NOD1*-associated inflammatory diseases (Fan et al. 2016). Functions specific to tissues have also been attributed to alternative splicing mechanisms in various genes (Yeo et al. 2004; Zhang et al. 2013), confirmed recently by protein–protein interaction studies of several alternate spliced genes in human (Yang et al. 2016). Keeping in view the multiple functions assigned to *NOD1*, this functional attribute cannot be ruled out in buffalo also.

Apart from documentation of splice variants, sequencing also helped in identification of indels to be present in the exon 3 of buffalo *NOD1*. Alignment of buffalo full-length transcript revealed, one insertion of three nucleotides (1533–1535) in ruminants, as compared to human (Fig. 1). McGovern et al. (2005) found that indels polymorphism in functional *NOD1* is associated with inflammatory bowel disease (IBD). However, the role of splice variants and indels (insertion and deletion) observed in buffalo needs to be explored further to ascertain their role if there is any in immune response.

The phylogenetic analysis, based on nucleotides sequence of *NOD1* of different species, revealed buffalo having the closest phylogenetic relationship with other ruminant species, cow, goat and sheep (Fig. 6). All the ruminant species clearly branched off in a separate cluster from non-ruminant species. However, among non-ruminants, pig was comparatively more close to ruminant species. Mice *NOD1* sequence was found to most diverged among all the mammalian species. Our findings on sequence and phylogenetic analysis of *NOD1* gene are similar to other reports for different species and other PRR genes (Dhiman et al. 2017; Dubey et al. 2013b; Goyal et al. 2014), which indicates co-evolution and conserved nature of this receptor protein across closely related ruminant species. Conserved residues found within functional domains across NOD-like receptors are speculated to contribute to the intramolecular interaction of the LRR regions to regulate NLR activation (Proell et al. 2008).

Tissue distribution by quantitative real-time assay of buffalo *NOD1*

The mRNA expression of buffalo *NOD1* quantified in ten different buffalo tissues was obtained by real-time PCR assay. The *NOD1* transcript was found to be expressed in all the ten buffalo tissues, however, the expression varied between tissues. Our results suggest that buffalo *NOD1* transcript was most abundantly expressed in PBMCs. Substantial expression was also found in lung, small intestine and liver tissues, whereas, the spleen, stomach, and mammary gland showed moderate expression. Although, a low expression level was seen in heart, skeletal muscle

Fig. 5 Splice variant identification in buffalo *NOD1* gene: **a** gel picture showing amplification of buffalo *NOD1* cDNA revealing amplification of two fragments of 1086 bp and 2914 bp of equal intensity. **b** Buffalo *NOD1* cDNA sequence analysis giving evidence of alternate splicing. **c** Domain structure prediction of the splice variant, showing spliced out region

Fig. 6 Phylogenetic analysis based on *NOD1* nucleotide sequence alignment, showing close clustering of ruminant species, buffalo, cow, sheep and goat. Nucleotide sequences were aligned using the ClustalW program and the phylogenetic tree was constructed using MEGA 6 software

Fig. 7 Basal expression of *NOD1* gene across buffalo organs by quantitative real-time PCR, taking kidney showing the least expression, as the normalizer tissue. Relative expression of *NOD1* in other tissues were represented as fold change from the normalizer (kidney)

and kidney tissues (Fig. 7). As kidney tissues showed the lowest expression (highest Ct value), it was included as a calibrator (Table 3-SI in supplementary information). Overall expression analysis across tissues indicates functional significance of *NOD1* in these organs, showing variability in transcript levels of this important multifunctional protein. This is the first report on the expression of the *NOD1* gene in buffalo tissues, and it is important given the fact that it performs variable functions across tissues (Zhang et al. 2013).

Trans-species comparison of key residues in NOD1

NOD1 amino acid residues of buffalo, cow, goat, sheep and human were aligned and evolutionary tracing was used to examine the amino acid conservation (Fig. 4). Comparison among Bovidae *NOD1* amino acid residues was found to be highly similar in CARD, NACHT and LRR domains. Disulphide bond formation in buffalo *NOD1* through Cys39 residue was found to be conserved in ruminants as well as humans as previously studied by Mayle et al. (2014). In spite of that, buffalo *NOD1* shows varying degrees of conservation of the protein interaction motif LxxLL found in the nuclear receptors. Buffalo *NOD1* has four LxxLL motifs L41 (CARD), L250, L314 (NACHT) and L593. Of these, two motifs LQDLL (41) and LCGLL (593) were found conserved across examined species. Residues involved in dimeric interface, implicated in RICK binding and signal transduction of *NOD1*-CARD were conserved across the examined species with exceptions of Leu41Val, Ala94Ser, and Arg101Lys. Residues that contributed to ubiquitination were found conserved across the ruminants with exception of Ala94Ser in sheep, while His88, Val89, Ile90 were substituted for Tyr88 and Lys89, 90 in humans (Fig. 8).

Homology modeling and NOD1-iE-DAP interaction

The 3D model of riverine and swamp type buffaloes and cattle *NOD1* was generated to have an insight into the structural basis of its function using rabbit *NOD2* (5IRL) as the template that shared 35% sequence identity with 85% query coverage. The best models were selected and validated by SAVE server on the basis of core region in Ramachandran plot and Z score value. Riverine, swamp and cow had almost similar DOPE scores, $-98,428.617$, $-97,959.078$ and $-98,886.242$, respectively, altogether having 13 amino acid variations between them. The amino acid variations and ligand interaction of *NOD1* of two buffalo types and cattle are shown in Fig. 1-SI (in supplementary

Fig. 8 Amino acid conservation in CARD domain of *NOD1* across species (buffalo, cow, sheep, goat and human). Conservation patterns observed are around residues implicated in the ubiquitination. Residues are colored according to hydrophobicity (green—hydrophobic, blue—hydrophilic)

Fig. 10 *NOD1* receptor ligand interaction. **a** *NOD1*-LRR-iE-DAP cattle complex, **b** *NOD1*-LRR-iE-DAP riverine buffalo complex, **c** *NOD1*-LRR-iE-DAP swamp buffalo complex

Acknowledgements The work reported has been carried out under the financial support received from National Agricultural Innovation Project of ICAR, Government of India under the project scheme C2153, which is thankfully acknowledged.

Compliance with ethical standards

Conflict of interest On behalf of all the authors, the corresponding author states that there is no conflict of interest for publishing this manuscript.

References

- Boyle JP, Mayle S, Parkhouse R, Monie TP (2013) Comparative genomic and sequence analysis provides insight into the molecular functionality of NOD1 and NOD2. *Front Immunol* 4:317. <https://doi.org/10.3389/fimmu.2013.00317>
- Brahma B, Kumar S, De BC, Mishra P, Patra MC, Gaur D, Chopra M, Gautam D, Mahanty S, Malik H, Malakar D (2015) Comparative genomic analysis of buffalo (*Bubalus bubalis*) *NOD1* and *NOD2* receptors and their functional role in in-vitro cellular immune response. *PLoS One* 10:e0119178. <https://doi.org/10.1371/journal.pone.0119178>
- Chamaillard M, Hashimoto M, Horie Y, Masumoto J, Qiu S, Saab L, Ogura Y, Kawasaki A, Fukase K, Kusumoto S, Valvano MA

- (2003) An essential role for *NOD1* in host recognition of bacterial peptidoglycan containing diaminopimelic acid. *Nat Immunol* 4:702–707. <https://doi.org/10.1038/ni945>
- Dhiman A, Mishra SK, Dubey PK, Goyal S, Sehgal M, Niranjana SK, Sodhi M, Mishra BP, Kataria RS (2017) Identification of genetic variation in NOD-like receptor 2 gene and influence of polymorphism on gene structure and function in buffalo (*Bubalus bubalis*). *Res Vet Sci* 115:43–50. <https://doi.org/10.1016/j.rvsc.2017.01.017>
- Dubey PK, Goyal S, Aggarwal J, Gahlawat SK, Kathiravan P, Mishra BP, Kataria RS (2012) Sequence and topological characterization of Toll-like receptor 8 gene of Indian riverine buffalo (*Bubalus bubalis*). *Trop Anim Health Prod* 45:91–99. <https://doi.org/10.1007/s11250-012-0178-1>
- Dubey PK, Goyal S, Kumari N, Mishra SK, Arora R, Kataria RS (2013a) Genetic diversity within 5' upstream region of Toll-like receptor 8 gene reveals differentiation of riverine and swamp buffaloes. *Meta Gene* 1:24–32. <https://doi.org/10.1016/j.mgene.2013.08.001>
- Dubey PK, Goyal S, Kathiravan P, Mishra BP, Gahlawat SK, Kataria RS (2013b) Sequence characterization of river buffalo Toll-like receptor genes 1–10 reveals distinct relationship with cattle and sheep. *Int J Immunogenet* 40:140–148. <https://doi.org/10.1111/j.1744-313X.2012.01135.x>
- Fan YH, Roy S, Mukhopadhyay R, Kapoor A, Duggal P, Wojcik GL, Pass RF, Arav-Boger R (2016) Role of nucleotide-binding oligomerization domain 1 (*NOD1*) and its variants in human cytomegalovirus control in vitro and in vivo. *Proc Natl Acad Sci USA* 113:E7818–E7827. <https://doi.org/10.1073/pnas.1611711113>
- Franchi L, Warner N, Viani K, Nunez G (2009) Function of Nod-like receptors in microbial recognition and host defense. *Immunol Rev* 227:106–128. <https://doi.org/10.1111/j.1600-065X.2008.00734.x>
- Girardin SE, Boneca IG, Carneiro LA, Antignac A, Jéhanho M, Viala J, Tedin K, Taha MK, Labigne A, Zathringer U, Coyle AJ (2003) Nod1 detects a unique muropeptide from gram-negative bacterial peptidoglycan. *Science* 300:1584–1587. <https://doi.org/10.1126/science.1084677>
- Goyal S, Dubey PK, Kumari N, Niranjana SK, Kathiravan P, Mishra BP, Mahajan R, Kataria RS (2014) Caprine Toll-like receptor 8 gene sequence characterization reveals close relationships among ruminant species. *Int J Immunogenet* 41:81–89. <https://doi.org/10.1111/iji.12075>
- Hugot JP, Chamaillard M, Zouali H, Lesage S, Cézard JP, Belaiche J, Almer S, Tysk C, O'morain CA, Gassull M, Binder V (2001) Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* 411:599–603. <https://doi.org/10.1038/35079107>
- Hysi P, Kubesch M, Moffatt MF, Schedel M, Carr D, Zhang Y, Boardman B, Von Mutius E, Weiland SK, Leupold W, Fritzsche C (2005) NOD1 variation, immunoglobulin E and asthma. *Hum Mol Gen* 14:935–941. <https://doi.org/10.1093/hmg/ddi087>
- Inohara N, Koseki T, del Peso L, Hu Y, Yee C, Chen S, Carrio R, Merino J, Liu D, Ni J, Núñez G (1999) Nod1, an Apaf-1-like activator of caspase-9 and nuclear factor- κ B. *J Biol Chem* 274:14560–14567
- Kawai T, Akira S (2009) The roles of TLRs, RLRs and NLRs in pathogen recognition. *Int Immunol* 21:317–337. <https://doi.org/10.1093/intimm/dxp017>
- Koller B, Bals R, Roos D, Kortling HC, Griesse M, Hartl D (2009) Innate immune receptors on neutrophils and their role in chronic lung disease. *Eur J Clin Invest* 39:535–547. <https://doi.org/10.1111/j.1365-2362.2009.02145.x>
- Kufer TA, Kremmer E, Alexander C, Adam AC, Philpott DJ, Sansonetti PJ (2008) The pattern-recognition molecule Nod1 is localized at the plasma membrane at sites of bacterial interaction. *Cell Microbiol* 10:477–486. <https://doi.org/10.1111/j.1462-5822.2007.01062.x>
- Maharana J, Pradhan SK, De S (2017) NOD1^{CARD} might be using multiple interfaces for RIP2-mediated CARD–CARD interaction: insights from molecular dynamics simulation. *PLoS One* 12:e0170232. <https://doi.org/10.1371/journal.pone.0170232>
- Manly BF (1986) Randomization and regression methods for testing for associations with geographical, environmental and biological distances between populations. *Res Popul Ecol* 28:201–218. <https://doi.org/10.1007/BF02515450>
- Mayle S, Boyle JP, Sekine E, Zurek B, Kufer TA, Monie TP (2014) Engagement of nucleotide-binding oligomerization domain-containing protein 1 (NOD1) by receptor-interacting protein 2 (RIP2) is insufficient for signal transduction. *J Biol Chem* 289:22900–22914. <https://doi.org/10.1074/jbc.M114.557900>
- McGovern DP, Hysi P, Ahmad T, van Heel DA, Moffatt MF, Carey A, Cookson WO, Jewell DP (2005) Association between a complex insertion/deletion polymorphism in NOD1 (CARD4) and susceptibility to inflammatory bowel disease. *Hum Mol Genet* 14:1245–1250. <https://doi.org/10.1093/hmg/ddi135>
- Mingala CN, Konnai S, Cruz LC, Onuma M, Ohashi K (2009) Comparative molecule-immunological analysis of swamp- and riverine-type water buffaloes responses. *Cytokine* 46:273–282. <https://doi.org/10.1016/j.cyto.2009.02.006>
- Mogensen TH (2009) Pathogen recognition and inflammatory signaling in innate immune defenses. *Clin Microbiol Rev* 22:240–273. <https://doi.org/10.1128/CMR.00046-08>
- Ogura Y, Inohara N, Benito A, Chen FF, Yamaoka S, Núñez G (2001) Nod2, a Nod1/Apaf-1 family member that is restricted to monocytes and activates NF- κ B. *J Biol Chem* 276:4812–4818. <https://doi.org/10.1074/jbc.M008072200>
- Peakall ROD, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Resour* 6:288–295. <https://doi.org/10.1111/j.1471-8286.2005.01155.x>
- Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 29:e45
- Plantinga TS, Fransen J, Knevel R, Netea MG, Zwerina J, Helsen MM, van der Meer JW, van Riel PL, Schett G van der, van den Berg WB, Helm-van Mil AH (2013) Role of NOD1 polymorphism in susceptibility and clinical progression of rheumatoid arthritis. *Rheumatology* 52:806–814. <https://doi.org/10.1093/rheumatology/kes404>
- Proell M, Riedl SJ, Fritz JH, Rojas AM, Schwarzenbacher R (2008) The Nod-like receptor (NLR) family: a tale of similarities and differences. *PLoS One* 3:e2119. <https://doi.org/10.1371/journal.pone.0002119>
- Sambrook J, Russell RW (2001) Molecular cloning: a laboratory manual, 3rd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Sha Z, Abernathy JW, Wang S, Li P, Kucuktas H, Liu H, Peatman E, Liu Z (2009) NOD-like subfamily of the nucleotide-binding domain and leucine-rich repeat containing family receptors and their expression in channel catfish. *Dev Comp Immunol* 33:991–999. <https://doi.org/10.1016/j.dci.2009.04.004>
- Shinkai H, Matsumoto T, Tokic 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. <https://doi.org/10.1016/j.molimm.2014.07.020>
- Takeda K, Akira S (2005) Toll-like receptors in innate immunity. *Int Immunol* 17:1–14. <https://doi.org/10.1093/intimm/dxh186>
- Tiszlavicz Z, Somogyvári F, Kocsis AK, Szolnoki Z, Sztrihai LK, Kis Z, Vécsei L, Mándi Y (2009) Relevance of the genetic polymorphism of NOD1 in *Chlamydia pneumoniae* seropositive stroke patients. *Eur J Neurol* 16:1224–1229. <https://doi.org/10.1111/j.1468-1331.2009.02698.x>

- Tohno M, Shimazu T, Aso H, Uehara A, Takada H, Kawasaki A, Fujimoto Y, Fukase K, Saito T, Kitazawa H (2008) Molecular cloning and functional characterization of porcine nucleotide-binding oligomerization domain-1 (NOD1) recognizing minimum agonists, meso-diaminopimelic acid and meso-lanthionine. *Mol Immunol* 45:1807–1817. <https://doi.org/10.1016/j.molimm.2007.09.029>
- Viala J, Chaput C, Boneca IG, Cardona A, Girardin SE, Moran AP, Athman R, Mémet S, Huerre MR, Coyle AJ, DiStefano PS (2004) Nod1 responds to peptidoglycan delivered by the *Helicobacter pylori* cag pathogenicity island. *Nat Immunol* 5:1166–1174. <https://doi.org/10.1038/ni1131>
- Xie J, Hodgkinson JW, Katzenback BA, Kovacevic N, Belosevic M (2013) Characterization of three Nod-like receptors and their role in antimicrobial responses of goldfish (*Carassius auratus* L.) macrophages to *Aeromonas salmonicida* and *Mycobacterium marinum*. *Dev Comp Immunol* 39:180–187. <https://doi.org/10.1016/j.dci.2012.11.005>
- Yang X, Coulombe-Huntington J, Kang S, Sheynkman GM, Hao T, Richardson A et al (2016) Widespread expansion of protein interaction capabilities by alternative splicing. *Cell* 164:805–817. <https://doi.org/10.1016/j.cell.2016.01.029>
- Yeo G, Holste D, Kreiman G, Burge CB (2004) Variation in alternative splicing across human tissues. *Genome Biol* 5:R74. <https://doi.org/10.1186/gb-2004-5-10-r74>
- Zhang ZB, Zhang W, Li RL, Li JB, Zhong JF, Zhao ZS, Huang JM (2013) Novel splice variants of the bovine PCK1 gene. *Genet Mol Res* 12:4028–4035. <https://doi.org/10.4238/2013.September.27.4>
- Zurek B, Proell M, Wagner RN, Schwarzenbacher R, Kufer TA (2012) Mutational analysis of human NOD1 and NOD2 NACHT domains reveals different modes of activation. *Innate Immun* 18:100–111. <https://doi.org/10.1177/1753425910394002>