Glycoconj J (2012) 29:335–345
DOI 10.1007/s10719-012-9415-0

Structural and functional conservation of CLEC-2 with the species-specific regulation of transcript expression in evolution

Lan Wang · Shifang Ren · Haiyan Zhu · Dongmei Zhang · Yuqing Hao · Yuanyuan Ruan · Lei Zhou · Chiayu Lee · Lin Qiu · Xiaojing Yun · Jianhui Xie

Received: 6 February 2012 / Revised: 5 June 2012 / Accepted: 6 June 2012 / Published online: 28 June 2012

Abstract CLEC-2 was first identified by sequence similarity to C-type lectin-like molecules with immune functions and has been reported as a receptor for the platelet-aggregating snake venom toxin rhodocytin and the endogenous sialoglycoprotein podoplanin. Recent researches indicate that CLEC-2-deficient mice were lethal at the embryonic stage associated with disorganized and blood-filled lymphatic vessels and severe edema. In view of a necessary role of CLEC-2 in the individual development, it is of interest to investigate its phylogenetic homology and highly conserved functional regions. In this work, we reported that CLEC-2 from different species holds with an extraordinary conservation by sequence alignment and phylogenetic tree analysis. The functional structures including *N*-linked oligosaccharide sites and ligand-binding domain implement a structural and functional conservation in a variety of species. The glycosylation sites (N120 and N134) are necessary for the surface expression CLEC-2. CLEC-2 from different species possesses the binding activity of mouse podoplanin. Nevertheless, the expression of CLEC-2 is regulated with a species-specific manner. The alternative splicing of pre-mRNA, a regulatory mechanism of gene expression, and the binding sites on promoter for several key transcription factors vary between different species. Therefore, CLEC-2 shares high sequence homology and functional identity. However the transcript expression might be tightly regulated by different mechanisms in evolution.

Keywords CLEC-2 · Homology · N-glycosylation · Ligand-binding · Alternative splicing

Lan Wang and Shifang Ren contributed equally to this work.

L. Wang · S. Ren · H. Zhu · D. Zhang · Y. Hao · Y. Ruan · L. Zhou · C. Lee · X. Yun (✉) · J. Xie (✉)
Department of Biochemistry and Molecular Biology,
Shanghai Medical College of Fudan University,
Box 103, Yixueyuan Road 138,
Shanghai 200032, China
e-mail: xjyun@fudan.edu.cn
e-mail: jhxie@fudan.edu.cn

L. Wang
Gene research center, Institutes of Biomedical Sciences,
Fudan University,
Shanghai, China

L. Qiu
Institute for Nutritional Sciences, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences,
Shanghai, Peoples Republic of China

Introduction

C-type lectin receptors (CLRs) are a superfamily of proteins containing conserved C-type lectin-like domain (CTLD) and have diverse functions including cell adhesion, regulation of natural killer function, complement activation, tissue remodeling, platelet activation, endocytosis and innate immunity [1, 2]. The members of CLRs encoded by the natural killer (NK) gene complex and expressed primarily on myeloid cells have been the focus of recent researches, especially the Dectin-1 cluster, which includes C-type lectin domain family 12, member A (MICL), C-type lectin domain family 1, member A (CLEC-1), C-type lectin domain family 1, member B (CLEC-2), C-type lectin domain family 12, member B (CLEC12B), C-type lectin domain family 9, member A (CLEC9A), Dectin-1 and lectin-like oxidized

LDL receptor-1 (LOX-1) [3, 4]. In contrast to its NK counterparts that primarily control cellular activation through recognition of major histocompatibility antigen class I and related molecules [5], the Dectin-1 cluster appears to have a wider range of functions and ligands, including those of exogenous origin.

Since the Dectin-1 cluster mediates important functions in innate immunity and immune homeostasis, it is of interest to investigate the evolutionary relationship between the genes of Dectin-1 cluster. Recent reports indicate that the highest similarity of Dectin-1 is seen with LOX-1(44 %), while CLEC-1 and CLEC-2 are somewhat more distantly related (37 %) [3]. Also, members of this family from different species share high sequence homology. Clearly, human Dectin-1 has a 72 % homology to mouse Dectin-1 [6] and LOX-1 between human and rat suggests a 71 % homology [7]. However, some structural variations between different species are observed. For instance, the position and number of *N*-linked glycosylation modifications are distinct between mouse and human Dectin-1 [8]. Additionally, rat LOX-1 possesses a triple-repeat structure within the neck region but there is a lack of this structure in human LOX-1 [9]. So far, little is known about the functional significance of the structural variants.

CLEC-2 is highly expressed on the surface of platelets and megakaryocytes and at a lower level on several other hematopoietic lineages of myeloid origin, including monocytes and dendritic cells [10, 11]. The cytosolic tail of CLEC-2 contains a novel sequence YxxL known as hemi immunoreceptor tyrosine-based activatory motif (hemITAM). Within ligand engagement, the hemITAM sequence would provide docking sites for the tandem-SH2 domains of the tyrosine kinase Syk allowing for tyrosine phosphorylation across a CLEC-2 receptor dimmer [11–15]. Rhodocytin, isolated from the venom of *Calloselasma rhodostoma* by affinity chromatography assay [11, 16] and podoplanin, a type-I transmembrane sialomucin-like glycoprotein expressed in lymphatic endothelial cells, were identified as the ligands of CLEC-2 [17, 18]. Interaction between CLEC-2 with rhodocytin, podoplanin, or specific antibodies elicits powerful platelet aggregation and secretion, which signals through Src- and Syk-dependent tyrosine kinases, leading to phosphorylation of a series of adapter and effector proteins that culminate in activation of phospholipase-Cγ2 (PLCγ2) and platelet activation [11]. In mouse, CLEC-2 is also expressed on peripheral blood neutrophils and mediates phagocytosis of antibody-coated beads and the production of proinflammatory cytokines [19]. Recently, we also have demonstrated that mouse CLEC-2 could be cleaved into a soluble form by proteases sensitive to aprotinin and phenylmethanesulfonyl fluoride (PMSF) [20].

Observation from Dectin-1-deficient and LOX-1-deficient mice shows no phenotypic abnormalities [21, 22]. However, recent reports indicated that CLEC-2-deficient mice displayed abnormal phenotype with blood/lymphatic vessel misconnections, which suggests that CLEC-2 is crucial to the development of embryonic vasculature and formation of lymphatic vessels and implies a necessary role of CLEC-2 on ontogenesis in contrast to other members of Dectin-1 cluster [23, 24]. To better understand the functional significance of CLEC-2, the investigation of CLEC-2 in the structural conservation and the expression regulation is important. Here, we investigate the structural and functional conservation of CLEC-2 and the regulation of CLEC-2 expression in evolution. We demonstrate that CLEC-2 holds with an extraordinary structural and functional conservation among species. However, the expression of CLEC-2 might be regulated with a species-specific manner.

Materials and methods

Antibodies and reagents

Mouse anti-Myc antibody was purchased from Invitrogen, rabbit anti-Myc antibody was from Cell Signaling Technology and goat anti-mouse-IgG HRP was from Santa Cruz Biotechnology (Santa Cruz, CA). Goat anti-mouse IgG antibody conjugated to DyLightTM488 and goat anti-rabbit IgG antibody conjugated to DyLightTM549 were from Jackson ImmunoResearch. Mouse podoplanin-Fc chimera (mPod-Fc) was from R&D Systems.

Amino acid sequence alignments

The deduced amino acid alignments were performed with CLUSTAL W at the European Molecular Biology Laboratory European Bioninformatics Institute (http://www.ebi.ac.uk/clustalw/) with default parameters. The alignments were then edited and shaded using GeneDoc. These cDNAs of CLEC-2 are from NM_001191978 (Rat); XM_002712905 (Rabbit); XM_001251938 (Bovine); NM_016509 (Human) and NM_019985 (Mouse).

Phylogenetic analysis

The cDNA sequences of CLEC-2 from different species were obtained through Entrez (http://www.ncbi.nlm.nih.gov/sites/entrez/) and gene prediction web server (http://www.ensembl.org/info/about/species.html). The retrieved sequences were subjected to further analysis by using various bioinformatics tools including similarity search BLAST (http://www.ncbi.nlm.nih.gov/blast) and Lasergene DNASTar program (EditSeq and MegAlign). Phylogenetic trees were built with Mega 4.0 software package [25] based on

Table 1 Primers of CLEC-2 in different species

Species	Primers (5′→3′)
Mouse	F: CAACTCGAGATGCAGGATGAAGATG
	R: CAAGGATCCAAGCAGTTGGTCCACT
Human	F: CAACTCGAGATGCAGGATGAAGATG
	R: CAAGGATCCAGGTAGTTGGTCCACC
Rat	F: ACACTCGAGATGCAGGATGAAGATG
	R: CAAGGATCCAAGCAGTTGTTCCACT
Rabbit	F: ACACTCGAGATGCAGGCTGAAGACG
	R: CAAGGATCCAAGTAGTTGGTCCAGC
Bovine	F: AACCTCGAGATGCAGGATGAAGATG
	R: CAAGGATCCAAGTAGATGTTCCACC

F forward primer, *R* reverse primer

the multiple alignments of sequences by ClustalX program and BioEdit. The statistical reliability of the tree was assessed by the bootstrap test and branches supported by a bootstrap <60 are collapsed.

Cell culture and transfection

Cells were cultured in Dulbecco's modification of Eagle's medium (Sigma) supplemented with 10 % fetal bovine serum, 100units/ml penicillin and streptomycin at 37 °C with 5 % CO_2. All transfections were carried out using LipofectAMINE2000 (Invitrogen) according to the manufacturer's instructions.

RT-PCR analysis

To determine the expression level of CLEC-2, total RNA was isolated from the spleen and liver of mouse (BALB/c), rat (Sprague-Dawley), rabbit (Newzealand white), bovine (Chinese yellow cattle) using TRIzol (Invitrogen) according to the manufacturer's protocol. Then, first strand cDNA synthesis was performed using an oligo (dT) primer and RT-PCR kit (TaKaRa). Subsequently, cDNA was amplified using the primers in Table 1. PCR was performed at 94 °C for 30 s, 60 °C for 30 s, 72 °C for 40 s for 30 cycles followed by a final extension at 72 °C for 10 min. Then the products were separated and viewed. To determine the

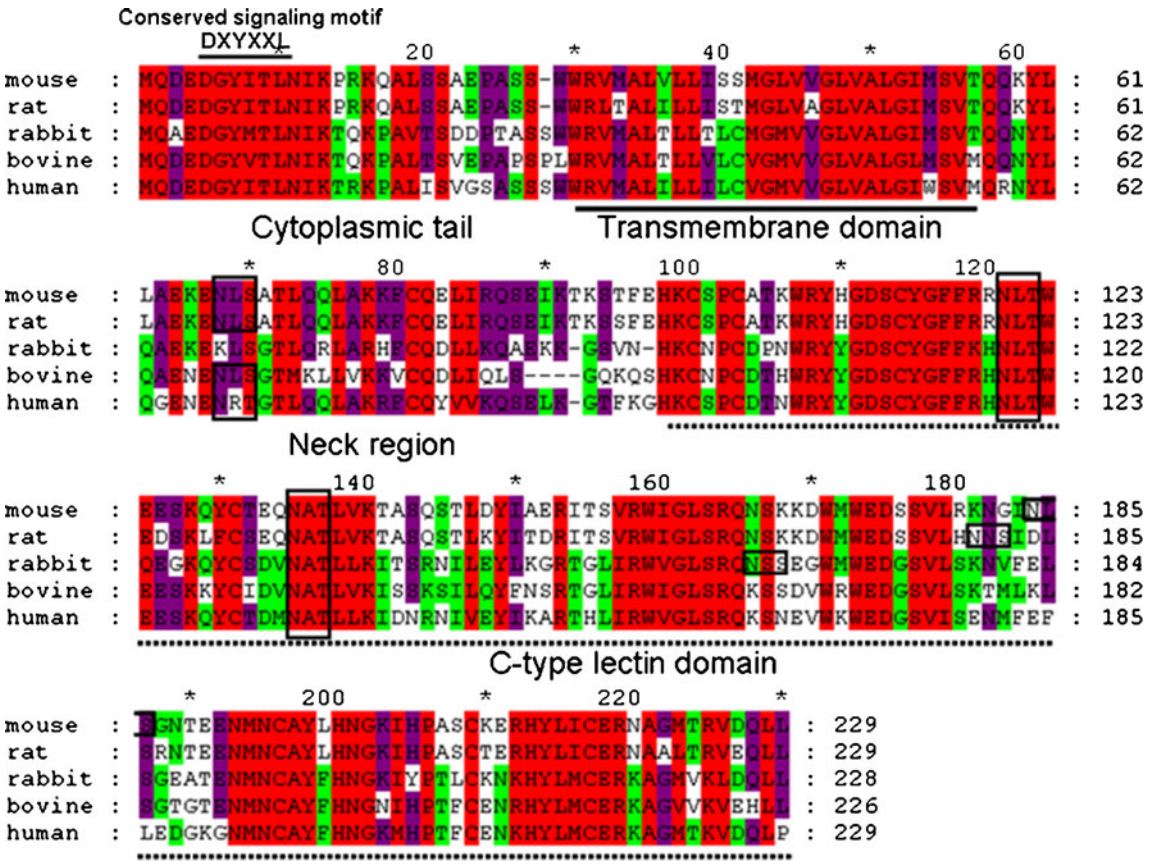

Fig. 1 Amino acid sequence alignments of CLEC-2 from mouse, rat, rabbit, bovine and human. The data was done using CLUSTAL W and edited by administrating GeneDoc. Shading is based on conservation, with the red shading representing 100 % conservation, purple representing 80 % and green 60 %. Predicted N-glycosylation sites are indicated in a black box

sequence, the DNA bands were cut, purified, and cloned into pMD19-T vector, and sequenced.

Western blot analysis

Proteins from total cell lysates were separated under reducing condition by SDS-PAGE and transferred onto PVDF membranes. The expression of CLEC-2 was detected with indicated antibodies according to Santa Cruz Biotechnology's protocol.

Plasmid construction and site mutagenesis

The RT-PCRs were performed using the mRNA from the tissue samples as described above and the productions of mouse, rat, rabbit, bovine and human CLEC-2 cDNAs were obtained. Then, these productions were in frame cloned into pcDNA3.1/Myc-His vector (Invitrogen) to generate Myc-tag construct pcDNA3.1/CLEC-2, respetively. Point mutant was generated using Product Mutan™-Super Express Km Mutagenesis Kit (TaKaRa) according to the manufacturer's protocol.

Confocal microscopy

The wild CLEC-2 and mutant CLEC-2 were transfected into HeLa cells grown on the glass coverslips. After 24~36 h, the transfected cells were washed twice with PBS and fixed with 4 % paraformaldehyde for 20 min at room temperature. After washed three times with PBS, the nonpermeabilized cells were detected for cell surface molecules with anti-Myc antibody and were stained with goat anti-mouse IgG-FITC antibody. To detect intracellular CLEC-2, cells were permeabilized with 0.1 % Triton X-100/PBS for 5 min on ice. Then, cells were viewed with laser confocal scanning microscope (Leica Microsystems Heidelberg GmbH, Germany).

Ligand binding assay

For the binding assay, the transfected cells seeded on coverslips were washed in PBS and preincubated with 50ug/ml

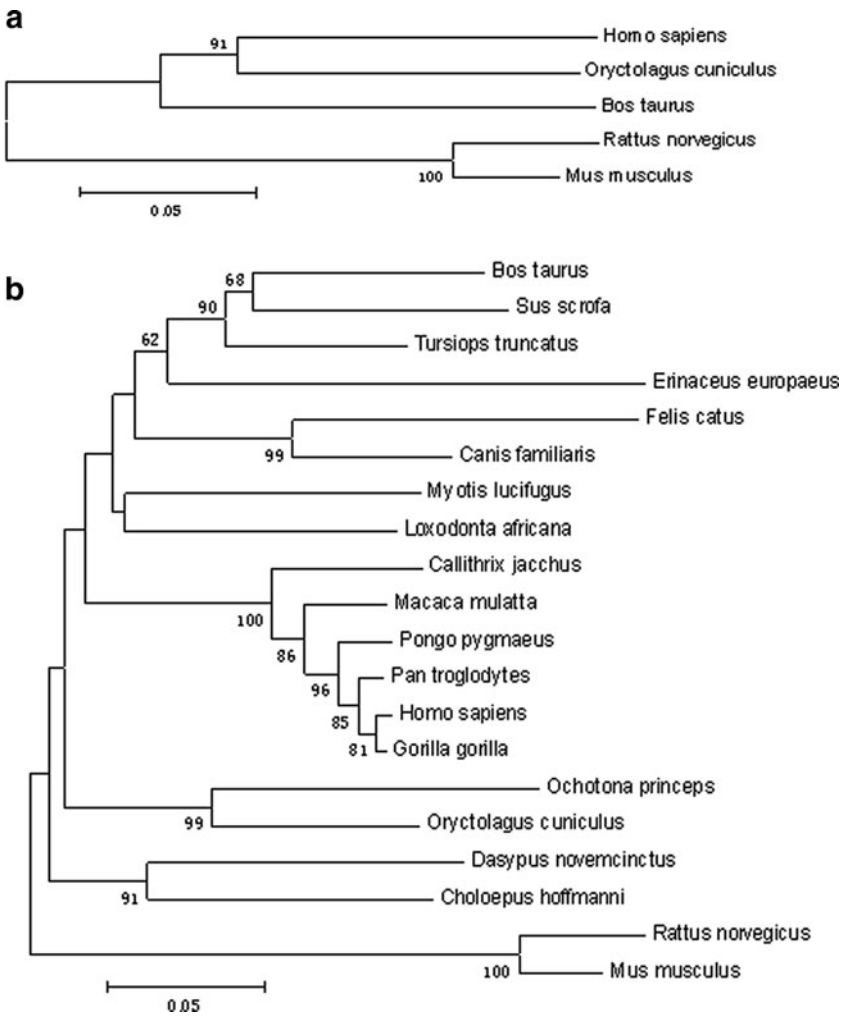

Fig. 2 Phylogenetic tree of CLEC-2 based on nucleotide sequences. **a** Phylogenetic tree of CLEC-2 was constructed between mouse, rat, rabbit, bovine and human. **b** Phylogenetic tree was built across the 20 predicted CLEC-2 nucleotide sequences. Phylogenetic trees were produced using alignments of complete or nearly complete CLEC-2 nucleotide sequences, from which part of the codons were removed. The bootstrap values, which show as a percentage calculated from 1,000 data sets, are shown at the nodes

mPod-Fc proteins and rabbit anti-Myc antibody at 4 °C for 45 min. Cells were extensively washed and fixed with 4 % paraformaldehyde for 10 min. Cells were stained with goat anti-mouse IgG antibody conjugated to DyLightTM488 to mark the binding mPod-Fc and goat anti-rabbit IgG antibody conjugated to DyLightTM549 for marking CLEC-2 expression. Finally, cell nucleus was stained with DAPI. Labeled cells were visualized using Leica confocal microscope.

Computer analysis of the transcription factors

To search putative binding sites for transcription factors in the promoter region, 2000 bp upstream sequences of the transcription start site of CLEC-2 genes were acquired from UCSC Genome Bioinformatics (http://genome.ucsc.edu) and the TFSEARCH program was performed to obtain the binding sites for transcription factors (http://www.cbrc.jp/research/db/TFSEARCH.html).

Results

The amino acid sequence and structure of CLEC-2 is highly conservative

To compare CLEC-2 sequences between different organisms, CLEC-2 cDNA from rabbit and bovine were obtained according to their genome and the amino acid sequence was deduced. Human, mouse and rat CLEC-2 was obtained via previous reports. Alignment of all the amino acid sequences with the software CLUSTAL W revealed that there are 129 positions with 100 % sequence conservation, 36 positions with 80 % sequence conservation, and 44 positions with 60 % sequence conservation (Fig. 1). Results from the report of Genedoc showed 58 %~90 % sequence identity of CLEC-2 among the five species. Phylogenetic tree based on the five CLEC-2 genes was constructed to further investigate the homology of CLEC-2. By using neighbor-joining method as implemented in Mega software, we observed that

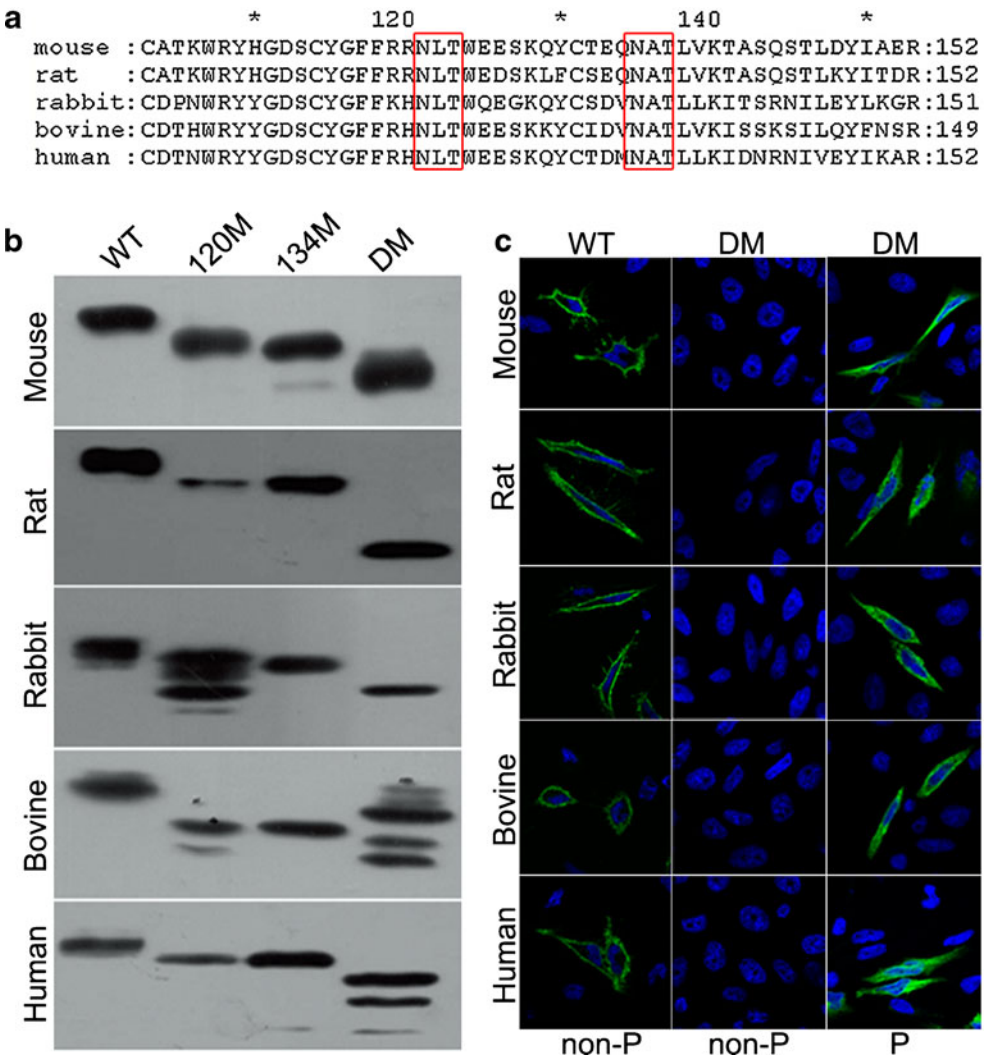

Fig. 3 Analysis of conservative *N*-glycosylation sites in CLEC-2. **a** Two conservative *N*-linked glycosylation sites are indicated in red box. **b** Analysis of the *N*-glycosylation sites of CLEC-2. Myc-tagged wild-type and mutant CLEC-2 were introduced into HEK293T cells. The whole cell lysates were subjected to 10 % SDS-PAGE. The proteins on the gel were immunoblotted with anti-Myc antibody (WT: wild CLEC-2; 120 M: mutant CLEC-2 at site Asn 120; 134 M: mutant CLEC-2 at site Asn 120; DM: mutant CLEC-2 at both sites of Asn 120 and Asn 134). **c** *N*-linked glycosylation modification impairs translocation of CLEC-2. The expression of wild CLEC-2 (WT) and mutant CLEC-2 (DM) on the cell surface (non-P) and inside cells (P) is shown in green. Nuclear stain (DAPI) is in blue

human CLEC-2 was more closely related to rabbit CLEC-2 than the other vertebrate CLEC-2 nucleotide sequence, while mouse CLEC-2 was more closely to rat, and the neighbor-joining bootstrap values supporting the relationship were very high, which is consistent with the results from sequence alignment that human CLEC-2 is 70 % identical to rabbit CLEC-2 and 62 % identical to mouse CLEC-2 (Fig. 2a). Gene predictions were also carried out on website Species List (http://www.ensembl.org/info/about/species.html) in order to build a more integrated phylogenetic tree including 20 CLEC-2 genes of mammals (Fig. 2b). The Neighbor- Joining trees suggested first that pig is closer to the cow than to the human, rabbit or mouse. Additionally, CLEC-2 in primate species formed separate but closely related clusters with a high degree of reliability (*i.e.*, bootstraps >80 %), indicating that they arose by a gene duplication event in the common ancestor to this group.

Furthermore, the prediction of *N*-glycosylation sites showed that there are three or four potential *N*-glycosylation sites on CLEC-2 from different species. Of note, there are two conservative sites (N120, N134) in these species (Fig. 3a), which implies a necessary role in the functional significance of CLEC-2. Previous studies demonstrate that the covalent and hydrophobic interactions within the CTLD of CLEC-2 determine its space conformation [26]. Except that three conserved disulfide bonds in the CTLD of C-type lectin-like receptors including C102-C113, C130-C216, and C195-C208 have a completely identical location, three core hydrophobic regions including two stacked tyrosine rings Y109 and Y148, two perpendicular tryptophan rings W123 and W158 and a conserved Trp-x-Trp motif with a leucine at position 161, are completely identical. These features suggest that the CTLD of CLEC-2 from different species should be of a highly resembled space conformation. Additionally, sequence alignment revealed that CLEC-2 from different species has a highly conservative cytoplasmic tail with an identical hemITAM motif, which implies that CLEC-2 employs a similar mechanism to propagate its signals in different species.

Site N120 together with N134 is necessary for the expression of CLEC-2 on cells

To determine whether N120 and N134 could be glycosylated, both sites N120 and N134 were replaced with amino acid Ala. Wild and mutant construct of CLEC-2 were transfected into HEK293T cells, respectively, and western blot analysis was performed with anti-Myc antibody. As shown in Fig. 3b, mutant at sites N120 and N134 increased the electrophoretic mobility of CLEC-2 in SDS-PAGE gel, compared with wild type CLEC-2. In fact, an increase of electrophoretic mobility of CLEC-2 was also observed when only site N120 or site N134 was replaced with Ala (Fig. 3b). These data suggested that both sites N120 and N134 of CLEC-2 from different species are conservative and could be glycosylated. *N*-linked oligosasaccharides at Asn-X-Ser/Thr sequence are crucial modifications for the folding, stability, localization and ligand binding of transmembrane proteins [27–29]. Subsequently, we determined whether mutant at both sites N120 and N134 would impair the translocation of CLEC-2 to the cell surface. HeLa cells were transfected with wild or mutant CLEC-2 construct and immunofluorescence analysis was performed with anti-Myc antibody. As illustrated in Fig. 3c, the expression of mutant

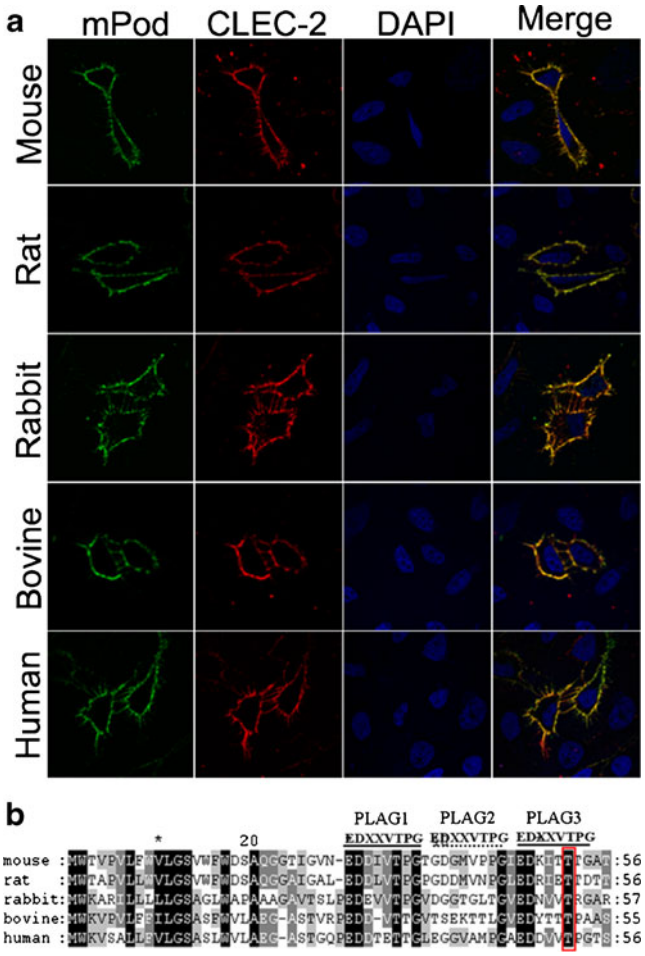

Fig. 4 Podoplanin binds to CLEC-2 on Hela cells. **a** Hela cells were transiently transfected with CLEC-2 and confocal microscopy analysis was performed as described in Materials and methods. Confocal microscopy assay was performed and the localization of CLEC-2 (red) and mPod (green) and their colocalization (yellow) are shown (original magnification 364). Nuclear staining (DAPI) is in blue. **b** Amino acid sequence alignments of podoplanin PLAG domain. Completely conserved residues are shaded black; 80 % is shaded dark gray; 60 % is shaded light grey. Conserved Thr52 in the PLAG domain is indicated in red box

CLEC-2 (N120A and N134A) could not be detected on the cell surface while the expression of wild type CLEC-2 could be seen on the cell surface. There is no expression of CLEC-2 because the expression of CLEC-2 was significantly viewed inside the cells. These data suggested that modification of *N*-linked glycosylation at both sites N120 and N134 of CLEC-2 from different species has an identical role and is necessary for the translocation of CLEC-2 to the cell surface.

Similar conformation of C-type lectin-like domain of CLEC-2

Since CLEC-2-deficient mice displayed abnormal phenotype with blood/lymphatic vessel misconnections, we speculate that the binding of Podoplanin to CLEC-2 should be conservative to maintain the development of embryonic vascular and formation of lymphatic vessel. To test our speculation, the binding of recombinant mPod-Fc fusion protein to CLEC-2 was performed. Results from confocal microscopy analysis showed that mPod-Fc could bind to the cell surface. CLEC-2 from different species were also detected on the cell surface and extensively colocalized with mPod-Fc, as revealed by merged images, which confirms the results from sequence alignment that the CTLD of CLEC-2 from different species should be of a highly resembled space conformation (Fig. 4a). Previous studies have demonstrated that a disialyl-core1 on Thr52 plus the platelet aggregation-stimulating domain 3 (PLAG3) are critical for the CLEC-2-binding activity of Podoplanin [30]. Sequence alignment revealed that Thr52 together with PLAG3 is highly conserved across species (Fig. 4b), which suggests that this region of podoplanin remains a proper conformation to bind to CLEC-2 in evolution.

The alternative splicing of the pre-mRNA of CLEC-2 is differently regulated

We subsequently investigate whether the expression of CLEC-2 from different species is regulated by a similar mechanism in evolution. To determine the expression of CLEC-2, RT-PCR was performed on RNA from liver and spleen. As shown in Fig. 5a, a different expression profile of CLEC-2 in liver from different species was observed. The observable bands were purified, cloned into pMD19-T vector and the sequences were determined. The isoforms of CLEC-2 from different species were indicated in Fig. 5c. The coding sequences of rabbit and bovine CLEC-2 confirmed the deduced sequences from their genomes and submitted to GenBank (Accession nos. HQ738490 and HQ738491). Results from sequencing show that the pre-mRNA of CLEC-2 could be alternatively spliced, which implies that the expression of CLEC-2 could be regulated by pre-mRNA splicing in different species. Furthermore, these splicing variants of CLEC-2 generated by alternative splicing differ from each other in different species, which suggests that the alternative splicing of the pre-mRNA of CLEC-2 is regulated in a species-specific manner. In fact, similar results were obtained in spleen, which confirms our observation (Fig. 5b).

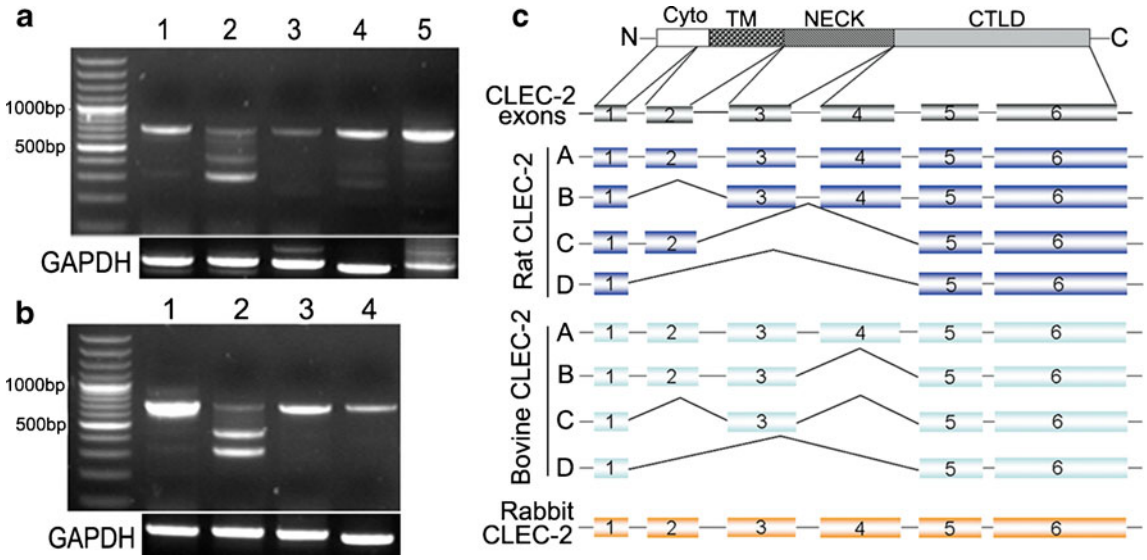

Fig. 5 Expression profiles of CLEC-2 transcripts. **a** Expression profile of CLEC-2 transcripts in liver of different species by RT-PCR analysis, and human from peripheral blood mononuclear cell (PBMC). Lane 1: mouse; lane 2: rat; lane 3: rabbit; lane 4: bovine; lane 5: human PBMC. **b** Expression profile of CLEC-2 transcripts in spleen of different species by RT-PCR analysis. Lane 1: mouse; lane 2: rat; lane 3: rabbit; lane 4: bovine. **c** Genomic structure of CLEC-2 gene. Rat isofoms and bovine isoforms are shown with putative length of encoding peptides (Cyto, cytoplamic tail; TM, transmembrane domain; NECK, neck region; CTLD, C-type lectin-like domain)

The promoter of CLEC-2 gene from different species displays a disparity

The promoter acts as an important regulator for the level of gene expression and we subsequently determined whether the transcription of CLEC-2 gene from different species is regulated in a similar manner. The transcription factors-binding sites in the promoter region of CLEC-2, 2000 bp upstream of the transcription initiation site, were predicted. Shown in Table 2 are the binding sites for several key transcription factors in hematopoietic development including NF-κB, activator protien-1 (AP-1), Ikaros, Oct-1, GATA family members GATA-1/2/3, AML1, MZF1 and C/EBP [31]. Significantly, the binding sites in the promoter region of CLEC-2 gene for several key transcription factors display a wide disparity and differ from each other, which suggest that the initiation of transcription of CLEC-2 might be under the control of different transcription factors in different species (Table 2).

Discussion

In this work, we investigate the structural and functional conservation of CLEC-2 and the regulation of CLEC-2 expression in evolution. CLEC-2 from different species holds with an extraordinary conservation and the functional structures including *N*-linked oligosaccharide sites and C-type lectin-like domain implement a structural and functional conservation. Nevertheless, the alternative splicing of pre-mRNA and the binding sites on promoter region for several key transcription factors vary between different species.

The sequence alignment of CLEC-2 from five species in this study displays an amazing homology. Proteins with a highly-homology manner are usually reported such as heat shock proteins [32]. These molecules might have an essential role in the individual development and cell survival. CLEC-2-deficient mice are lethal at the embryonic and neonatal stage with blood/lymphatic misconnections, which suggests that CLEC-2 plays an important role in the regulation of blood-lymphatic vascular separation. In fact, many genes have a functional conservation and need not a highly sequence homology in evolution, and the reason that CLEC-2 shows an amazing homology remains to be investigated. Therefore, further understanding of the function of CLEC-2 might be helpful to clarify this issue.

Some pattern motifs and domains are widely employed in the regulation of protein function. N-linked oligosasaccharides at Asn-X-Ser/Thr sequence are crucial modifications for the folding, stability, localization and ligand binding of transmembrane proteins. Unlike dectin-1 with different N-linked carbohydrate sites between human and mouse [8], CLEC-2 possesses two identical N-linked glycosylation sites (N120 & N134). Our results indicate that N120 and N134 of CLEC-2 could be glycosylated and mutant of both N120 and N134 abolished the translocation of CLEC-2 to the cell surface, which implies that N-linked oligosasaccharides at N120 and N134 of CLEC-2 are essential for the surface expression of CLEC-2. Furthermore, our results show that mouse Podoplanin could bind to CLEC-2 from different species, which suggests CLEC-2 from different species has a highly similar space conformation to recognize mouse Podoplanin. Additionally, sequence alignment reveals that the site Thr52 plus the PLAG3 of Podoplanin, critical for the CLEC-2-binding activity, is a conservative site in evolution, which implies the binding of Podoplanin to CLEC-2 is necessary and conservative to maintain the development of embryonic vascular and formation of lymphatic vessels in evolution. However, Kaneko *et al.* reported that the bovine podoplanin has a sporadic deletion mutation in the first PLAG domain and did not induce the aggregation of platelets [33], which is consistent with recent reports that a disialyl-core1 on Thr52 plus the PLAG3 are necessary, but not enough to induce the aggregation of platelets [30]. These observations imply that bovine Podoplanin could bind to CLEC-2 with the inability to induce aggregation of platelets. It is possible that the activity of Podoplanin to maintain the development of embryonic vascular and formation of lymphatic vessels in bovine could be compensated by other mechanisms.

Alternative splicing is one of the most important mechanisms to generate a large number of mRNA and protein isoforms from the surprisingly low number of human genes [34]. Unlike promoter activity, which primarily regulates the amount of transcripts, alternative splicing changes the structure of transcripts and their encoded proteins. For instance, LOXIN, a gene variant of LOX-1, was deficient in ligand binding, but interacted with the full-length LOX-1 receptors and blocked their surface expression, ox-LDL binding activity and its uptake [35]. Although CLEC-2 shares a high sequence-homology, the splicing variants of CLEC-2 generated by alternative splicing differ from each other in the detected tissues. Furthermore, different tissues have distinct expression profile of these isoforms of CLEC-2 in the same species. Moreover, analysis of the promoter region of CLEC-2 gene shows that the binding sites for several key transcription factors differ from each other, which suggests that the initiation of transcription of CLEC-2 might be under the control of different transcription factors in different species. We do not have suitable cell lines from these species to test whether these elements have a role in initiating the transcription of CLEC-2 while we observe that the promoter region of CLEC-2 contains a number of binding sites for GATA-1, important transcription factors for megakaryocytes differentiation [36, 37], which implies that the prediction of the binding sites for transcription factors is to a

Table 2 Transcription factors binding sites identified by using TFSEARCH (threshold score 87) within predicted promoter regions of CLEC-2 in mouse, rat, rabbit, bovine and human

Transcription factor	Position(Strand)				
	Mouse	Rat	Rabbit	Bovine	Human
GATA-1	−1977(+);−1938(+); −1864(−); −1226(+); −550(+); −65(+)	−1989(−); −1943(−); −1807(−); −1802 (−); −1761(−);−1652(+); −1681(+); −1281(+); −1199(+); −960(+); −606(−)	−1957(−);−1121(+); −1140(+); −993(+); −879(+); −626(−); −530(+); −492(−)	−1741(−);−1670(+); −1338(−); −1076(−); −914(+); −806(+); −696(−); −628(−); −326(−)	−1770(−);−1516(+); −1450(−); −1136(−); −1012(+); −725(−); −407(−); −273(−); −44(−)
AML1	−1932(+);−1714(+); −371(+)	−1503(−); −1113(+); −737(−); −572(−); −77(−)	−1853(−);−1715(+); −957(−); −820(+); −658(+); −602(+); −482(−); −390(−); −203(−); −162(+)	−1952(−);−1881(−); −1145(+); −1021(+); −923(−); −915(−); −846(+); −261(+); −148(−);	−1891(−);−1467(−); −1456(−); −1098(−); −34(−)
C/EBP	−1660(−); −1110(−); −827(+)	−1945(−); −940(+); −949(+)	−1970(−);−1972(+); −1909(+); −1771(−); −1602(−)	−1470(+); −597(+); −19(−)	−1905(+);−1141(−); −814(−); −653(−); −357(+); −173(−); −114(+)
GATA-2	−1237(−);−1034(+); −550(−)	−1807(−); −1841(−); −1802(−); −1724(−); −606(−)	−1957(−);−1771(−); −1140(+); −1121(+; −993(+); −626(−); −530(+); −494(−); −339(+)	−1741(−); −806(+); 696(−); −628(−)	−1771(−);−1012(+); −946(−); −407(−); −275(−)
Oct-1	−1608(−);−1547(+); −1526(−); −1511(−); −1408(−);−1363(+); −761(+); −746(+); −727(+); −633(+); −165(−)	−1483(−); −1385(−); −689(+); −676(−)	−1801(+);−1838(−); −1077(+); −938(−); −949(+)	−1823(+);−1461(−); −1187(+); −823(−); −434(+); −302(−)	−1938(+); −935(+); −370(+); −339(+); −250(+)
MZF1	−1559(−);−1234(−); −547(−)	−1824(−); −1839(−); −1811(−); −1727(−); −1206(−); −994(+); −603(−); −453(−)	−1255(−);−1090(−)	−1684(−); −635(−)	−748(−)
GATA-3	−1937(+)	−1990(−); −1808(−); −1279(+)	−1958(−);−528(+); −495(−)	−1682(−); −161(−)	−1451(−);−1304(+); −1333(+);−1011(+); −276(−); −152(+); −47(−)
Ikaros	−1269(−)	−1727(−); −542(−)	No detection	−690(−)	−1331(+)
AP-1	−299(+)	No detection	−58(+)	−1142(+); −1036(−)	−1355(+)
NF-κB	No detection	−1814(−); −138(−)	−1543(−); −1015(+)	−499(+)	−93(−)

certain extent reliable. As a result, the biological function of CLEC-2 might be tightly controlled and differently regulated in species although CLEC-2 from different species shares a high sequence-homology and functional identity in evolution.

Acknowledgements This work was supported by Shanghai Leading Academic Discipline Project (B110), National fund program of personnel training in basic science (J0730860), National Natural Science Foundation of China (30900266), Doctoral Fund of Ministry of Education of China for Young Teachers (20090071120053), Major Program of National Natural Science Foundation of China (1991483) and Projects of International Cooperation and Exchanges NSFC (31010103906).

References

1. Zelensky, A.N., Gready, J.E.: The C-type lectin-like domain superfamily. FEBS J **272**(24), 6179–6217 (2005)
2. Osorio, F., Reis, E.S.C.: Myeloid C-type lectin receptors in pathogen recognition and host defense. Immunity **34**(5), 651–664 (2011)
3. Sobanov, Y., Bernreiter, A., Derdak, S., Mechtcheriakova, D., Schweighofer, B., Duchler, M., Kalthoff, F., Hofer, E.: A novel cluster of lectin-like receptor genes expressed in monocytic, dendritic and endothelial cells maps close to the NK receptor genes in the human NK gene complex. Eur J Immunol **31**(12), 3493–3503 (2001)
4. Huysamen, C., Brown, G.D.: The fungal pattern recognition receptor, Dectin-1, and the associated cluster of C-type lectin-like receptors. FEMS Microbiol Lett **290**(2), 121–128 (2009)
5. Ljunggren, H.G., Karre, K.: In search of the 'missing self': MHC molecules and NK cell recognition. Immunol Today **11**(7), 237–244 (1990)
6. Grunebach, F., Weck, M.M., Reichert, J., Brossart, P.: Molecular and functional characterization of human Dectin-1. Exp Hematol **30**(11), 1309–1315 (2002)
7. Nagase, M., Hirose, S., Fujita, T.: Unique repetitive sequence and unexpected regulation of expression of rat endothelial receptor for oxidized low-density lipoprotein (LOX-1). Biochem J **330**(Pt 3), 1417–1422 (1998)
8. Kato, Y., Adachi, Y., Ohno, N.: Contribution of N-linked oligosaccharides to the expression and functions of beta-glucan receptor, Dectin-1. Biol Pharm Bull **29**(8), 1580–1586 (2006)
9. Aoyama, T., Sawamura, T., Furutani, Y., Matsuoka, R., Yoshida, M.C., Fujiwara, H., Masaki, T.: Structure and chromosomal assignment of the human lectin-like oxidized low-density-lipoprotein receptor-1 (LOX-1) gene. Biochem J **339**(Pt 1), 177–184 (1999)
10. Colonna, M., Samaridis, J., Angman, L.: Molecular characterization of two novel C-type lectin-like receptors, one of which is selectively expressed in human dendritic cells. Eur J Immunol **30**(2), 697–704 (2000)
11. Suzuki-Inoue, K., Fuller, G.L., Garcia, A., Eble, J.A., Pohlmann, S., Inoue, O., Gartner, T.K., Hughan, S.C., Pearce, A.C., Laing, G.D., Theakston, R.D., Schweighoffer, E., Zitzmann, N., Morita, T., Tybulewicz, V.L., Ozaki, Y., Watson, S.P.: A novel Syk-dependent mechanism of platelet activation by the C-type lectin receptor CLEC-2. Blood **107**(2), 542–549 (2006)
12. Fuller, G.L., Williams, J.A., Tomlinson, M.G., Eble, J.A., Hanna, S.L., Pohlmann, S., Suzuki-Inoue, K., Ozaki, Y., Watson, S.P., Pearce, A.C.: The C-type lectin receptors CLEC-2 and Dectin-1, but not DC-SIGN, signal via a novel YXXL-dependent signaling cascade. J Biol Chem **282**(17), 12397–12409 (2007)
13. Hughes, C.E., Pollitt, A.Y., Mori, J., Eble, J.A., Tomlinson, M.G., Hartwig, J.H., O'Callaghan, C.A., Futterer, K., Watson, S.P.: CLEC-2 activates Syk through dimerization. Blood **115**(14), 2947–2955 (2010)
14. Sada, K., Zhang, J., Siraganian, R.P.: SH2 domain-mediated targeting, but not localization, of Syk in the plasma membrane is critical for FcepsilonRI signaling. Blood **97**(5), 1352–1359 (2001)
15. Watson, A.A., Christou, C.M., James, J.R., Fenton-May, A.E., Moncayo, G.E., Mistry, A.R., Davis, S.J., Gilbert, R.J., Chakera, A., O'Callaghan, C.A.: The platelet receptor CLEC-2 is active as a dimer. Biochemistry **48**(46), 10988–10996 (2009)
16. Shin, Y., Morita, T.: Rhodocytin, a functional novel platelet agonist belonging to the heterodimeric C-type lectin family, induces platelet aggregation independently of glycoprotein Ib. Biochem Biophys Res Commun **245**(3), 741–745 (1998)
17. Christou, C.M., Pearce, A.C., Watson, A.A., Mistry, A.R., Pollitt, A.Y., Fenton-May, A.E., Johnson, L.A., Jackson, D.G., Watson, S.P., O'Callaghan, C.A.: Renal cells activate the platelet receptor CLEC-2 through podoplanin. Biochem J **411**(1), 133–140 (2008)
18. Suzuki-Inoue, K., Kato, Y., Inoue, O., Kaneko, M.K., Mishima, K., Yatomi, Y., Yamazaki, Y., Narimatsu, H., Ozaki, Y.: Involvement of the snake toxin receptor CLEC-2, in podoplanin-mediated platelet activation, by cancer cells. J Biol Chem **282**(36), 25993–26001 (2007)
19. Kerrigan, A.M., Dennehy, K.M., Mourao-Sa, D., Faro-Trindade, I., Willment, J.A., Taylor, P.R., Eble, J.A., Reis e Sousa, C., Brown, G.D.: CLEC-2 is a phagocytic activation receptor expressed on murine peripheral blood neutrophils. J Immunol **182**(7), 4150–4157 (2009)
20. Xie, J., Wu, T., Guo, L., Ruan, Y., Zhou, L., Zhu, H., Yun, X., Hong, Y., Jiang, J., Wen, Y., Gu, J.: Molecular characterization of two novel isoforms and a soluble form of mouse CLEC-2. Biochem Biophys Res Commun **371**(2), 180–184 (2008)
21. Saijo, S., Fujikado, N., Furuta, T., Chung, S.H., Kotaki, H., Seki, K., Sudo, K., Akira, S., Adachi, Y., Ohno, N., Kinjo, T., Nakamura, K., Kawakami, K., Iwakura, Y.: Dectin-1 is required for host defense against Pneumocystis carinii but not against Candida albicans. Nat Immunol **8**(1), 39–46 (2007)
22. Mehta, J.L., Sanada, N., Hu, C.P., Chen, J., Dandapat, A., Sugawara, F., Satoh, H., Inoue, K., Kawase, Y., Jishage, K., Suzuki, H., Takeya, M., Schnackenberg, L., Beger, R., Hermonat, P.L., Thomas, M., Sawamura, T.: Deletion of LOX-1 reduces atherogenesis in LDLR knockout mice fed high cholesterol diet. Circ Res **100**(11), 1634–1642 (2007)
23. Bertozzi, C.C., Schmaier, A.A., Mericko, P., Hess, P.R., Zou, Z., Chen, M., Chen, C.Y., Xu, B., Lu, M.M., Zhou, D., Sebzda, E., Santore, M.T., Merianos, D.J., Stadtfeld, M., Flake, A.W., Graf, T., Skoda, R., Maltzman, J.S., Koretzky, G.A., Kahn, M.L.: Platelets regulate lymphatic vascular development through CLEC-2-SLP-76 signaling. Blood **116**(4), 661–670 (2010)
24. Suzuki-Inoue, K., Inoue, O., Ding, G., Nishimura, S., Hokamura, K., Eto, K., Kashiwagi, H., Tomiyama, Y., Yatomi, Y., Umemura, K., Shin, Y., Hirashima, M., Ozaki, Y.: Essential *in vivo* roles of the C-type lectin receptor CLEC-2: embryonic/neonatal lethality of CLEC-2-deficient mice by blood/lymphatic misconnections and impaired thrombus formation of CLEC-2-deficient platelets. J Biol Chem **285**(32), 24494–24507 (2010)
25. Tamura, K., Dudley, J., Nei, M., Kumar, S.: MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol **24**(8), 1596–1599 (2007)
26. Watson, A.A., Brown, J., Harlos, K., Eble, J.A., Walter, T.S., O'Callaghan, C.A.: The crystal structure and mutational binding analysis of the extracellular domain of the platelet-activating receptor CLEC-2. J Biol Chem **282**(5), 3165–3172 (2007)

27. Li, L.B., Chen, N., Ramamoorthy, S., Chi, L., Cui, X.N., Wang, L.C., Reith, M.E.: The role of N-glycosylation in function and surface trafficking of the human dopamine transporter. J Biol Chem **279**(20), 21012–21020 (2004)
28. Tanaka, K., Xu, W., Zhou, F., You, G.: Role of glycosylation in the organic anion transporter OAT1. J Biol Chem **279**(15), 14961–14966 (2004)
29. Cai, G., Salonikidis, P.S., Fei, J., Schwarz, W., Schulein, R., Reutter, W., Fan, H.: The role of N-glycosylation in the stability, trafficking and GABA-uptake of GABA-transporter 1. Terminal N-glycans facilitate efficient GABA-uptake activity of the GABA transporter. FEBS J **272**(7), 1625–1638 (2005)
30. Kato, Y., Kaneko, M.K., Kunita, A., Ito, H., Kameyama, A., Ogasawara, S., Matsuura, N., Hasegawa, Y., Suzuki-Inoue, K., Inoue, O., Ozaki, Y., Narimatsu, H.: Molecular analysis of the pathophysiological binding of the platelet aggregation-inducing factor podoplanin to the C-type lectin-like receptor CLEC-2. Cancer Sci **99**(1), 54–61 (2008)
31. Orkin, S.H.: Transcription factors and hematopoietic development. J Biol Chem **270**(10), 4955–4958 (1995)
32. Lindquist, S., Craig, E.A.: The heat-shock proteins. Annu Rev Genet **22**, 631–677 (1988)
33. Kaneko, M.K., Kato, Y., Kitano, T., Osawa, M.: Conservation of a platelet activating domain of Aggrus/podoplanin as a platelet aggregation-inducing factor. Gene **378**, 52–57 (2006)
34. Stamm, S., Ben-Ari, S., Rafalska, I., Tang, Y., Zhang, Z., Toiber, D., Thanaraj, T.A., Soreq, H.: Function of alternative splicing. Gene **344**, 1–20 (2005)
35. Biocca, S., Filesi, I., Mango, R., Maggiore, L., Baldini, F., Vecchione, L., Viola, A., Citro, G., Federici, G., Romeo, F., Novelli, G.: The splice variant LOXIN inhibits LOX-1 receptor function through hetero-oligomerization. J Mol Cell Cardiol **44**(3), 561–570 (2008)
36. Visvader, J., Adams, J.M.: Megakaryocytic differentiation induced in 416B myeloid cells by GATA-2 and GATA-3 transgenes or 5-azacytidine is tightly coupled to GATA-1 expression. Blood **82**(5), 1493–1501 (1993)
37. Visvader, J.E., Elefanty, A.G., Strasser, A., Adams, J.M.: GATA-1 but not SCL induces megakaryocytic differentiation in an early myeloid line. EMBO J **11**(12), 4557–4564 (1992)