Gai *et al. Chem. Biol. Technol. Agric.* (2023) 10:34
https://doi.org/10.1186/s40538-023-00406-7

Chemical and Biological Technologies in Agriculture

RESEARCH

Open Access

Whole genome level analysis of the DEATH protein superfamily in sheep (*Ovis aries*) and their coordination relationship in regulating lactation

Zhongchao Gai[1], Songhao Hu[1], Yujiao He[1], Guoli Gong[1*] and Jieqiong Zhao[2*]

Abstract

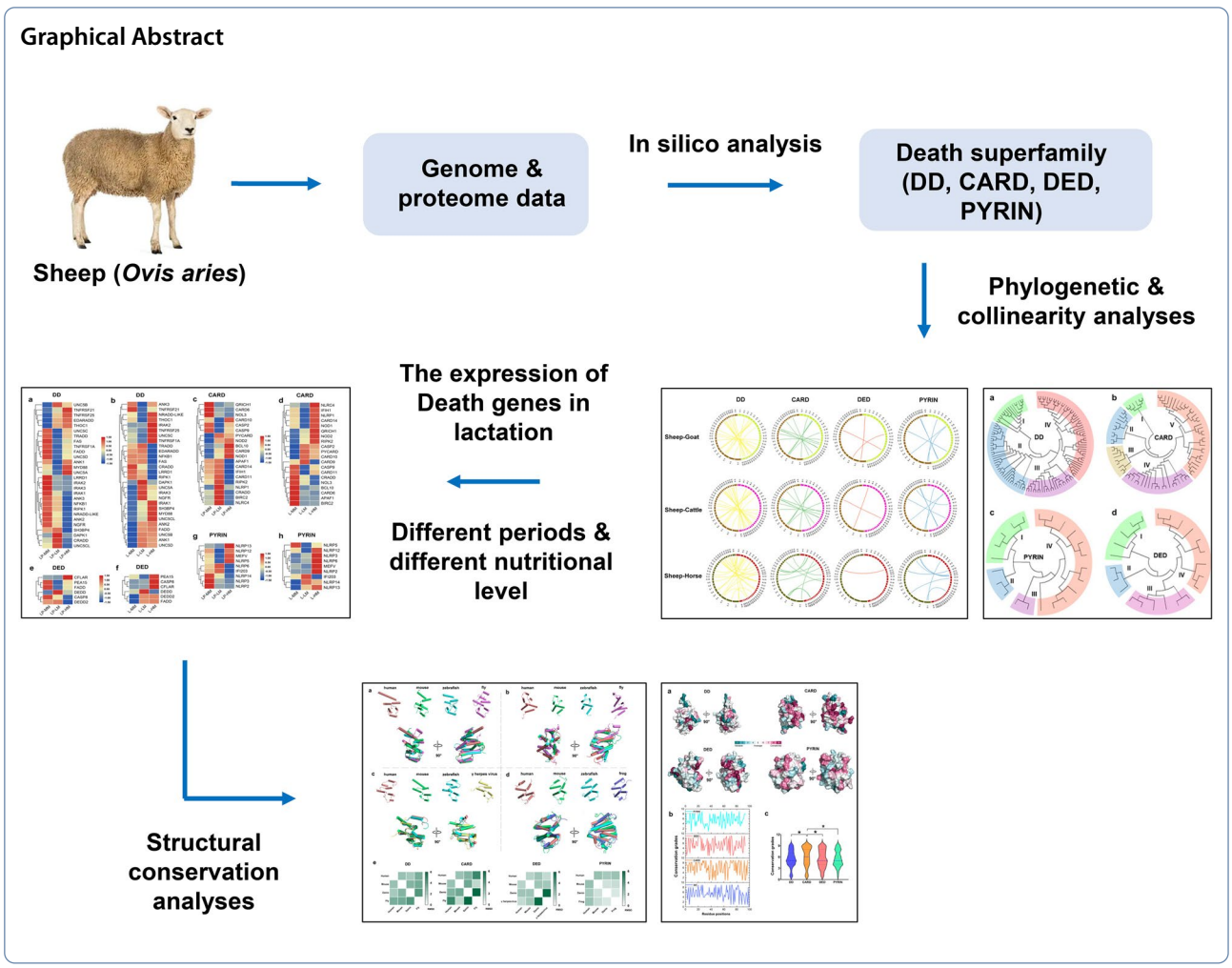
Background Sheep milk is a nutritional and health-promoting food source for humans. The DEATH superfamily is a conserved protein family, and some of its members are closely related to lactation. Systematic studies of the members of the DEATH superfamily are important for further understanding its functions in the mammary gland during lactation; however, there studies are currently lacking.

Results Herein, 74 members of the DEATH superfamily were identified in sheep, and phylogenetic analyses indicated that four subfamilies were strongly correlated in evolution. The Ka/Ks calculations demonstrated that negative selection was the primary pressure acting on DEATH members; however, the immune-related gene *IFI203* was undergoing strong positive selection in sheep. Furthermore, in the late pregnancy and lactation period, these DEATH genes exhibited similar expression patterns under different nutritional conditions in the mammary gland, and four subfamilies were positively correlated in expression patterns. Additionally, half or more DEATH genes were upregulated in the lactation period, which implied their crucial roles in the lactation of sheep.

Conclusions The current research contributes to a better understanding of the evolutionary characteristics of the DEATH superfamily and their roles in sheep lactation, and it also provides potential target genes for the molecular breeding of dairy sheep.

Keywords Lactation, DEATH superfamily, Sheep milk

*Correspondence:
Guoli Gong
gongguoli@sust.edu.cn
Jieqiong Zhao
zhaojieqiong1234@163.com
Full list of author information is available at the end of the article



Introduction

Ovis aries (sheep) is an important domestic animal throughout the world. Sheep milk can provide more protein (~5.5 g/100 g milk) than cow milk (3.4 g/100 g milk) to human beings and can lower cholesterol levels. In addition, sheep milk can serve as an excellent source of protein and provide a wide range of minerals and vitamins for humans [1, 2]. Various hormones are closely related to the lactation process in sheep, such as prolactin (PRL), growth hormones (GH), and estrogen (E) [3]. After giving birth, some of these lactation-related hormones are greatly inhibited [4, 5]. These lactation-related hormones regulate the synthesis of lipids and proteins, as well as milk secretion, through different approaches. Numerous studies have suggested that lactation processes are regulated by hormones through various intracellular signaling pathways, such as apoptosis signaling, the JAK-STAT pathway, and the mTOR pathway [6–8]. It is known that a large proportion of mammary epithelial cells (MECs) suffer from apoptotic events throughout the lactation process, and the amount and level of apoptotic MECs directly affect the milk production of animals [9].

The DEATH domain superfamily is composed of the death domain (DD), death effector domain (DED), PYRIN and caspase recruitment domain (CARD) families, most of which are involved in the regulation of apoptosis and inflammation. DNA damage is a major cause of apoptotic events. Ankyrin-1 (ANK1), the DD subfamily member, is markedly upregulated under DNA damage conditions, and this upregulation is mediated by the p53 protein [10]. Besides, P53-induced death domain-containing protein 1 (PIDD1), the DD subfamily member, can interact with components of death receptor signaling pathways and plays crucial roles in p53-dependent apoptosis in mammals [11, 12]. Moreover, uncoordinated-5 (UNC-5) proteins, such as UNC5 A/B/C/D, are receptors for secreted netrins, both of which are transmembrane proteins with an intracellular DD domain and involved in the regulation of cell apoptosis [13, 14]. Additionally, many members of the tumour necrosis factor

receptor (TNFRSF) superfamily, including TNFRSF1A, TNFRSF11A, TNFRSF21, and TNFRSF25, promote apoptosis and regulate inflammatory reactions [15, 16].

Caspases are aspartate-specific cysteine proteases that are closely related to apoptosis. Caspase-2 is a CARD domain-containing protein and plays important roles in mediating stress-induced apoptosis and cell cycle regulation [17]. Additionally, caspase-9 contains the CARD domain and functions as an essential initiator required for apoptosis through the mitochondrial pathway; furthermore, the failure of caspase-9 activation has profound pathological consequences [18]. Different from caspase-2 and caspase-9, caspase-8 is a DED domain-containing protein that can trigger extrinsic apoptotic pathways, and emerging evidence suggests that caspase-8 controls the comprehensive crosstalk between necroptosis, pyroptosis, and apoptosis, and further determines the type of cell death induced by cell death signaling [19, 20]. In addition, some studies have also suggested that caspase-8 has some nonapoptotic functions, such as promoting cell growth [21]. On the other hand, the NLR family, PYRIN domain-containing protein (NLRP), is involved in the formation of inflammasomes and further regulates inflammasome-induced apoptosis [22]. For example, NLRP3 inflammatory signaling mediates angiotensin II -induced INS-1 cell apoptosis [23]. Moreover, NLRP2 interacts with IKKa and promotes the DNA binding activity of NF-κB, and the overexpression of NLRP2 decreases the apoptotic cell rate [24].

Genome-wide analysis can provide a systematic understanding of the characteristics of a specific gene family, including member classification, family expansion, and molecular evolution [25, 26]. To systematically identify the molecular functions of the DEATH superfamily and to explore their potential roles in the regulation of lactation, it is necessary to analyse the DEATH superfamily at the whole genome level. In this study, we identified 74 members of the DEATH superfamily in the sheep genome by using in silico methods. Phylogenetic analyses suggested that these four subfamilies of the DEATH superfamily have a similar phylogenetic pattern, which implied a potential coevolutionary relationship between them. Moreover, we identified an inflammatory response and the lipid homeostasis-related gene *IFI203*, which was undergoing genetic positive selection and evolved rapidly, by using the Ka/Ks calculation method. High-throughput RNA sequence data showed that more than half of DEATH members were upregulated in the sheep mammary gland during the lactation period compared to the pregnancy period. Furthermore, this expressional upregulation of DEATH members could not be affected by nutritional conditions. Our results suggest that members of the sheep DEATH superfamily may play crucial roles in lactation and can serve as potential targets in the genetic breeding of dairy sheep breeds.

Methods

Genome-wide identification of DEATH superfamily members

All of the sheep (*Ovis aries*) protein sequences were retrieved from the NCBI database, and the HMM files of the DD domain (PF00531), DED domain (PF01335), CARD domain (PF00619), and PYRIN domain (PF02758) were obtained from the Pfam database. Subsequently, the members of the DEATH superfamily (DD, CARD, DED and PYRIN) were identified by using the hmmsearch program with an E-value $\leq 1.0\times10^{-5}$. Additionally, all of these candidates were further checked in the InterPro database. Similarly, the DEATH superfamily members of 15 other species (*Homo sapiens*, *Mus musculus*, *Columba livia*, *Gallus gallus*, *Danio rerio*, *Xenopus laevis*, *Octopus bimaculoides*, *Hydra vulgaris*, *Tribolium castaneum*, *Drosophila hydei*, *Penaeus vannamei*, *Osmia Taurus*, *Limulus polyphemus*, *Apis mellifera*, and *Caenorhabditis elegans*) were identified through the same procedure.

Phylogenetic analysis

The amino acid sequences of the DEATH superfamily members from five organisms (*Ovis aries*, *Gallus gallus*, *Danio rerio*, *Octopus bimaculoides*, and *Trichoplusia ni*) were aligned via the *MUSCLE* method in MEGA X software. Afterwards, phylogenetic trees of DD, CARD, DED, and PYRIN members were separately constructed by using the neighbor-joining (N-J) method with 1,000 bootstrap replicates.

Genome-wide collinearity and Ka/Ks analysis

Analyses of homologous gene pairs of DDs, CARDs, DEDs and PYRINs from sheep-goat, sheep-cattle, and sheep-horse were conducted as previously reported [2]. Briefly, whole-genome FASTA sequences and GTF files of *Ovis aries*, *Capra hircus*, *Bos taurus*, and *Equus caballus* were downloaded from the ENSEMBL database. DNA and protein sequences were aligned via the Diamond program, and the collinearity files were visualized in the jcvi package in Python [27]. The Ka, Ks, and Ka/Ks values were separately calculated for the DD, CARD, DED, and PYRIN family members between *Ovis aries*, *Capra hircus*, *Equus Caballus* and *Camelus dromedaries*. All of these protein pairs were aligned by using the BLASTP method, and we subsequently calculated the Ka/Ks values through the KaKs_Calculator program (version 2.0) [28].

Orthologue analysis

Orthologues of sheep protein coding genes in yeast (*Saccharomyces cerevisiae*), Arabidopsis (*Arabidopsis thaliana*), Drosophila (*Drosophila melanogaster*), and zebrafish (*Danio rerio*) were identified by using the InParanoid 4 program [29, 30]. Protein-coding genes in sheep were divided into five different groups according to the existence of orthologues. The status of the orthologues was named "none of the other species" (labelled by ****), "only in zebrafish" (#***), "only in zebrafish and Drosophila" (##**), "existed in zebrafish, Drosophila and Arabidopsis" (###*), and "existed both in the other four species" (####). The different phylogenetic patterns represented different evolutionary stages. For the orthologous analysis of all of the sheep protein-coding genes, we first downloaded five full protein datasets from the NCBI database, including 40,658 sheep sequences, 5,405 yeast sequences, 27,334 Arabidopsis sequences, 30,717 Drosophila sequences, and 57,100 zebrafish sequences. The sequence pairs of sheep-yeast, sheep-arabidopsis, sheep-drosophila, and sheep-zebrafish were aligned by using the blastp algorithm. Afterwards, the orthologous genes were identified by using the Inparanoid 4 program. Finally, the percentages of orthologous genes were calculated through the number of identified orthologues divided by 40,658 (the quantity of sheep protein sequences).

Gene expression analyses of DEATH superfamily members

It is known that nutrition levels in the late pregnancy and lactation periods can exert a significant influence on lactation performance. Herein, we analysed the expression patterns of DEATH superfamily genes in the mammary gland during the late pregnancy (LP) and lactation (L) periods under different nutritional conditions. First, we retrieved the gene expression data from the NCBI GEO website (accession: GSE71424), which contained the RNA-Seq data of 15 mammary gland samples (nine samples from the late pregnancy period and six samples from the lactation period). In addition, all 15 samples were divided into three groups according to their respective nutritional conditions: ad libitum nutrition (HM), maintenance nutrition (MM), and submaintenance nutrition (LM). Expressional heatmaps of members of the DEATH superfamily were generated by using the *pheatmap* package in the R suite (version 4.1).

Three-dimensional structural analysis

The 3D structural alignments were performed as previously described [26]. Briefly, the 3D structures of the DD, CARD, DED, and PYRIN domains of human, mouse, zebrafish, fly, and gamma herpesvirus were downloaded from the PDB database or AlphaFold Protein Structure Database [31]. Structural alignments and RMSD calculations were performed by using PyMOL software (version 2.0). The amino acid conservation of each domain and the residues were analysed by using the ConSurf online server.

Results

Significant positive correlation exists between the family expansion of DD, CARD, DED, and PYRIN in evolution

To identify the members of the DEATH protein superfamily and to explore the evolutionary patterns of its four subfamilies (DD, CARD, DED, and PYRIN) in species evolution, we first identified 32 DD domain-containing protein-coding genes, 22 CARD protein-coding genes, seven DED genes, and 13 PYRIN genes in the sheep genome (Fig. 1a, Additional files 1, 2, 3, 4: Table S1-S4). Both were named according to their orthologues in humans. Furthermore, we identified members of the DEATH superfamily in 15 other organisms, including seven vertebrates and eight invertebrates. Generally, the quantities of DD, CARD, DED, and PYRIN genes increased with the evolutionary process. To compare the expansion patterns of the four subfamilies, Pearson correlation analyses were conducted by calculating the gene numbers of the DD, CARD, DED, and PYRIN families in different organisms. The statistical results indicated that there were significant positive correlations between the expansion of the DD, CARD, DED, and PYRIN families (Fig. 1b). The Pearson's r value ranged from 0.5629 to 0.958 (mean value = 0.7836). Among the four subfamilies, DD, CARD, and DED had a strong correlation. The Pearson's r values of PYRIN with the other three subfamilies were relatively weaker than those of DD-CARD, DD-DED, and CARD-DED, because there were no PYRIN genes identified in the eight invertebrates that were analysed in this study.

Phylogenetic patterns of the DEATH superfamily

In an attempt to establish appropriate phylogenetic comparisons, we separately analysed the phylogenetic patterns of four subfamilies of the DEATH superfamily. First, 62 DD protein sequences from sheep and four other organisms, including *Gallus gallus* (chicken, which is the representative species of birds), *Danio rerio* (zebrafish, which is the representative species of marine vertebrates), *Octopus bimaculoides* (octopus, which represents marine invertebrates) and *Trichoplusia ni* (moth, which represents terrestrial invertebrates), were aligned by using the ClustalW method, after which N-J phylogenetic trees were built. Similarly, the phylogenetic trees of DD (125 proteins), DED (22 proteins), and PYRIN (23 proteins) were built with similar methods. The results showed that the DD, CARD, DED, and PYRIN proteins had similar phylogenetic patterns,

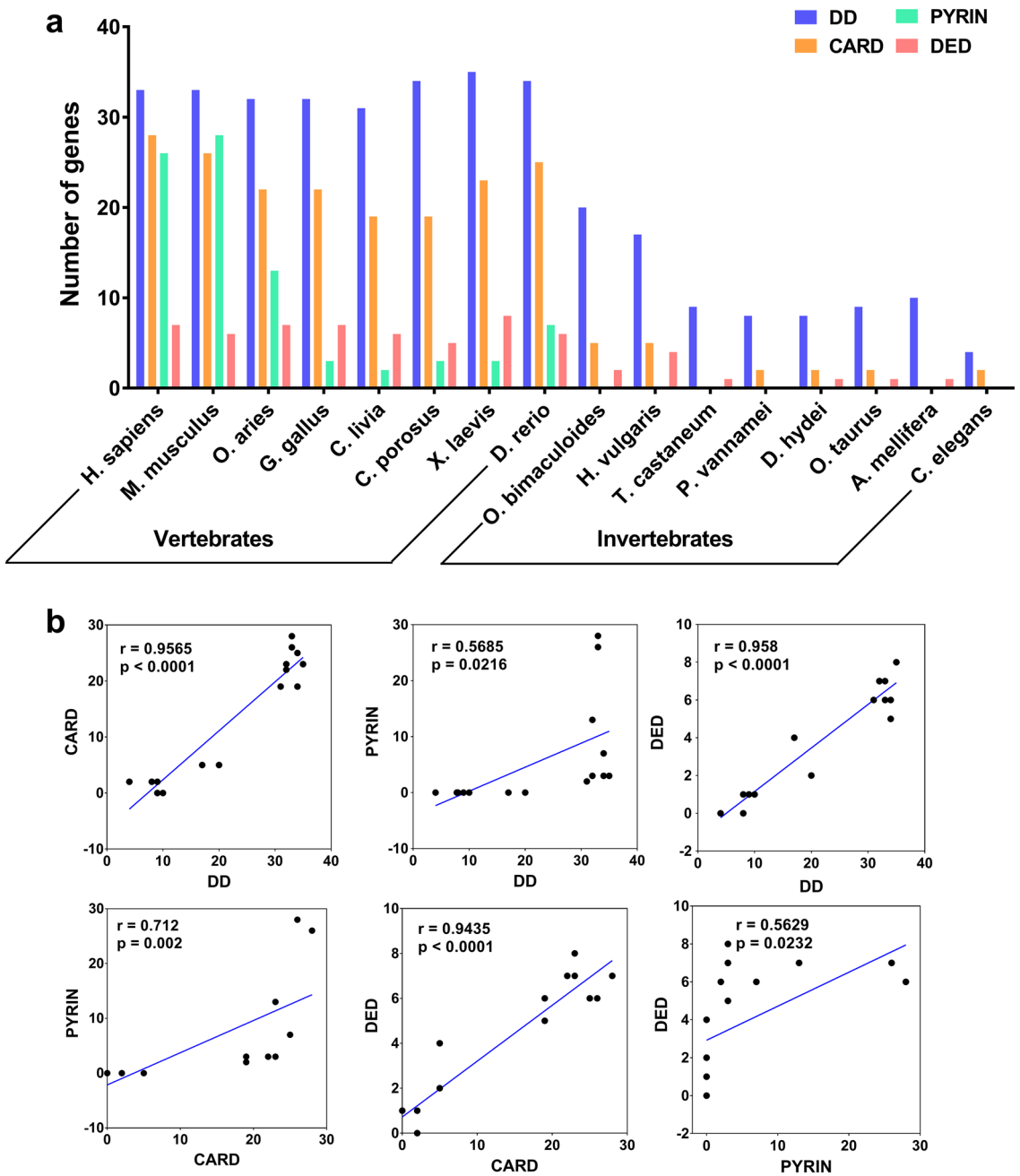


Fig. 1 The numbers of genes from four subfamilies of DEATH (DD, CARD, DED, and PYRIN) were correlated in species evolution. **a** The number of four subfamily members in different species. **b** Pearson's correlation analyses of the number of DD, CARD, DED, and PYRIN genes identified from the genomes of different species

and both of these trees could be categorized into four or five clades. Furthermore, group IV of DD (42 proteins), DED (11 proteins), and PYRIN (9 proteins) was the largest group in the respective tree, and clade V of CARD was its largest group, comprising 31 proteins (Fig. 2). These correlations suggested that these four subfamilies could coevolve because functionally complementary gene divergence and duplication events are usually prone to be retained by natural selection [32].

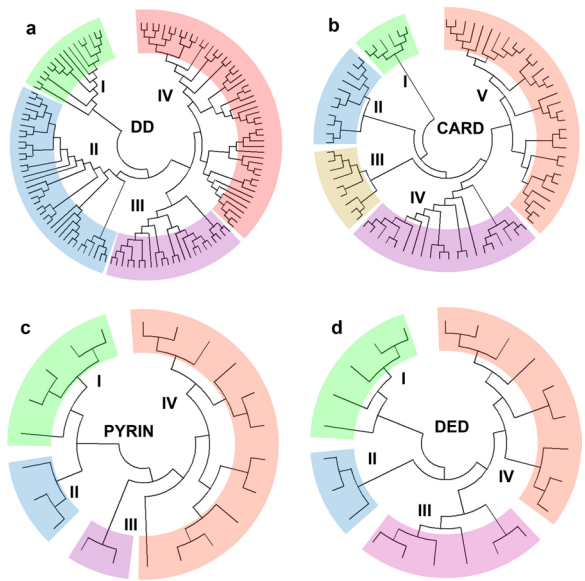


Fig. 2 Phylogenetic classification of four subfamilies of DEATH proteins of different species. These phylogenetic trees were constructed through the neighbor-joining method with 1,000 bootstrap replicates based on DD, CARD, DED, and PYRIN protein sequences from five organisms (three vertebrates: *Ovis aries*, *Gallus gallus*, and *Danio rerio*; and two invertebrates: *Octopus bimaculoides* and *Trichoplusia ni*). These different colors are only used to indicate that one phylogenetic tree can be divided into several major clades, and the same colors do not suggest that these clades in different trees have common characteristics

Collinearity analysis of DEATH superfamily genes from common milk-yielding animals

Sheep, goats, cattle, and horses are common milk-yielding animals. We surveyed the collinearity relationships among the orthologous DEATH superfamily genes from sheep (*Ovis aries*), goat (*Capra hircus*), cattle (*Bos taurus*) and horse (*Equus caballus*) to investigate the potential clues of evolutionary events between these livestock. The results indicated that 25 of 32 (78%) DD genes in sheep had collinearity relationships with 22 goat genes, 24 cattle genes, and 25 horse genes. Moreover, thirteen out of twenty-two (59%) CARD genes in sheep had collinearity relationships with 13 goat genes, 13 cattle genes, and 13 horse genes. Three out of seven (43%) sheep DED genes had collinearity relationships with three goat genes, two cattle genes, and one horse gene. In addition, eleven out of thirteen (85%) sheep PYRIN genes had collinearity relationships with 11 goat genes, 13 cattle genes and 12 horse genes (Fig. 3). It was remarkable that the total number of collinearity relations of the DEATH superfamily genes between sheep-goat were similar to sheep-cattle and sheep-horse. A reasonable explanation for this result is that both of them belong to the superorder *Laurasiatheria*.

Furthermore, the evolutionary constraints of orthologous pairs of DEATH superfamily genes were determined according to nonsynonymous substitutions (Ka) and synonymous substitutions (Ks). We calculated the Ks and Ka values of DD, CARD, DED, and PYRIN gene pairs and determined the selection types in the evolution of DEATH superfamily members through Ka/Ks ratios (Fig. 4a). The Ka/Ks ratios of all of the DD, CARD, and DED genes and most PYRIN genes were lower than 1, which indicated that negative selection drove the nucleotide substitutions of these genes. There was no significant difference between the Ka and Ks values of the pairs of DD, CARD, DED, and PYRIN genes; however, the Ka/Ks ratios of DD orthologous pairs were significantly lower than those of CARD, DED, and PYRIN (Fig. 4b). Among the PYRIN gene pairs, the average Ka/Ks values of IFI203 in six organism pairs (0.8413) were significantly higher than the mean value of all PYRIN genes (0.3613), which indicated that these three genes underwent less negative selection than the average. In particular, the Ka/Ks ratios of *IFI203* within *Bovidae* animals were significantly higher than average, with the Ka/Ks values of *IFI203* from sheep-cattle being 1.2025 and goat-cattle being 0.9924, which suggested that the *IFI203* gene was undergoing positive selection during the evolution and expansion of the *Bovidae* lineage.

The DEATH superfamily expanded mainly in the late evolutionary stage

In this study, we identified the orthologues of DD, CARD, DED, and PYRIN proteins in yeast (single-celled eukaryote), Arabidopsis (plant), Drosophila (invertebrate), zebrafish (marine vertebrate) and sheep (terrestrial mammal). The origin of these organisms represented different time points in species evolution, we were able to determine the family expansion of these four subfamilies through the existence of orthologues. The members of the sheep DEATH family with orthologs in all of the other four organisms were labelled '####', and one (3.13%) DD protein belonged to this group, which indicated that it originated from the emergence of eukaryotes. No DEATH superfamily members in sheep have orthologues both in Arabidopsis, Drosophila and zebrafish (###*). In addition, ten (31.25%) DD proteins and three CARD proteins had orthologues in Drosophila and zebrafish (##**), thus suggesting that these members appeared before the appearance of vertebrates. A total of 15 DDs, 15 CARDs, 7 DEDs, and 8 PYRINs in sheep had orthologues only in zebrafish (#***), thus indicating that these members appeared before the separation between terrestrial and aquatic vertebrates. Furthermore, six DDs, 4 CARDs, and 5 PYRINs have no orthologues in the other four organisms (****), thus suggesting that they emerged with the

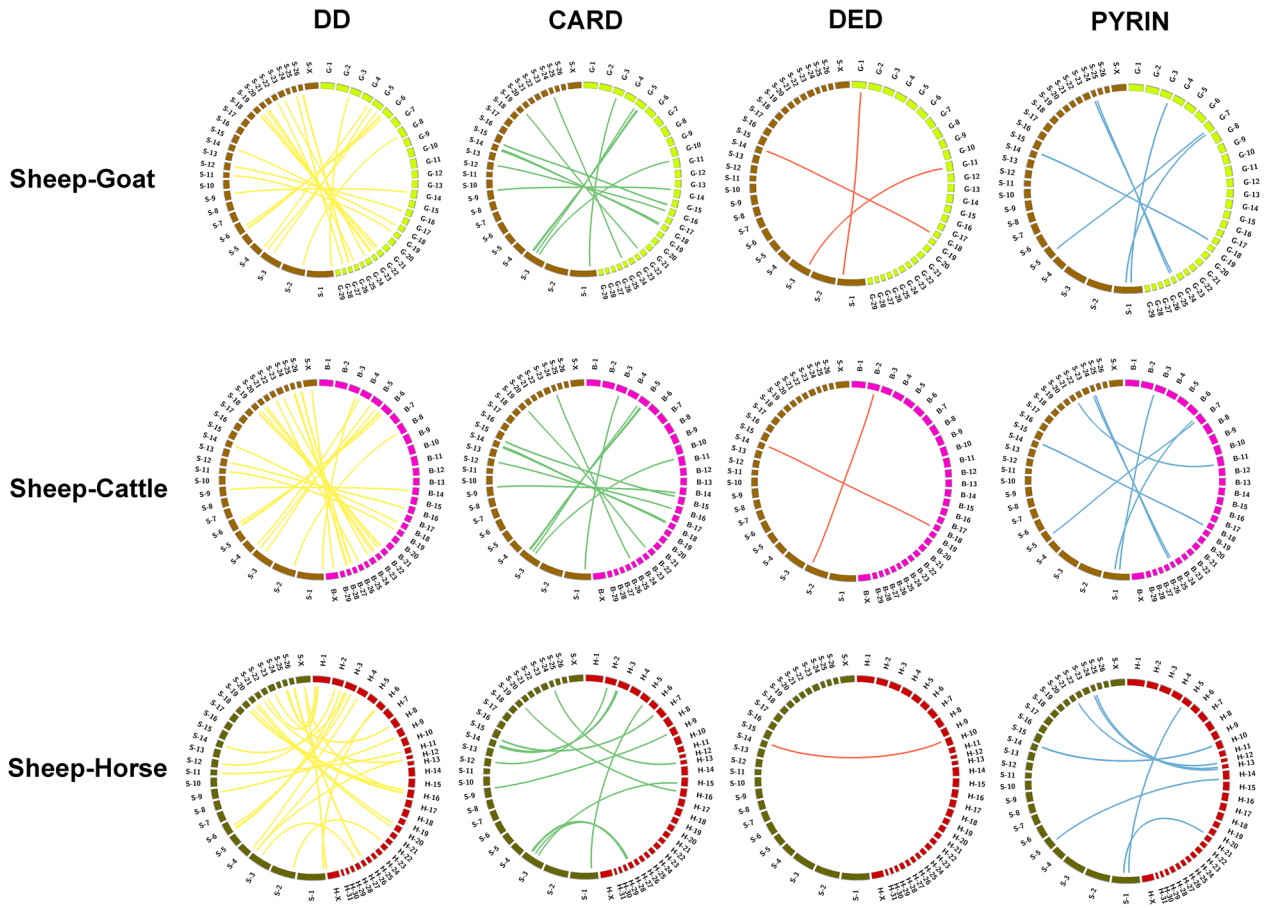


Fig. 3 Collinearity relationships of the DEATH superfamily genes from common milk-yielding animals. The collinearity relationships of the DD, CARD, DED, and PYRIN genes from sheep-goat, sheep-cattle, and sheep-horse. Highlight lines link these collinear genes between different species. These coloured segments in circles represent autosomes and X chromosomes from sheep (S), cattle (B), and horse (H)

lineage expansion of terrestrial vertebrates (Fig. 5a). All of these results showed that the DEATH superfamily originated from one DD protein-coding gene in a single-cellular eukaryote and expanded in the late evolutionary stage (Fig. 5b).

Expression patterns of DEATH superfamily members in the mammary gland of sheep in late pregnancy and lactation

It is known that the expression levels of many receptor genes, catalytic genes, and signal transducers are altered during the transition between late pregnancy and lactation [33, 34]. To examine the functional roles of the DEATH superfamily genes during this physiological switch, the expression levels of DEATH superfamily members were analysed in the mammary tissues of sheep. Herein, we analysed the RNA-seq data of tissue samples from the mammary glands of 15 sheep. These sheep were divided into three groups according to their nutritional levels: ad libitum nutrition (HM), maintenance nutrition (MM), and submaintenance nutrition (LM). First, we analysed the expression patterns of DD, CARD, DED, and PYRIN members during the late pregnancy stage (LP) and the lactation period (L) under different nutritional conditions (Fig. 6a–h). The results showed that nutrition levels could affect the expression level of almost all members of the DEATH superfamily. More than 50% of DEATH family members (DD: 53.6%, CARD: 55%, DED: 83.3%, and PYRIN: 78.8%) were upregulated in lactation at high nutritional levels compared to maintenance nutrition (Fig. 6i). Approximately half of the DEATH genes (DD: 46.4%, CARD: 55.0%, DED: 66.7%, and PYRIN:55.6%) were negatively regulated in late pregnancy at a high nutritional level (Fig. 6k). By comparing the number of genes that were down- or upregulated in the LP and L periods under high and low nutritional levels, we found that the expression changes of the four subfamily genes were positively correlated between the DD, CARD, DED, and PYRIN subfamilies (Fig. 6j, l).

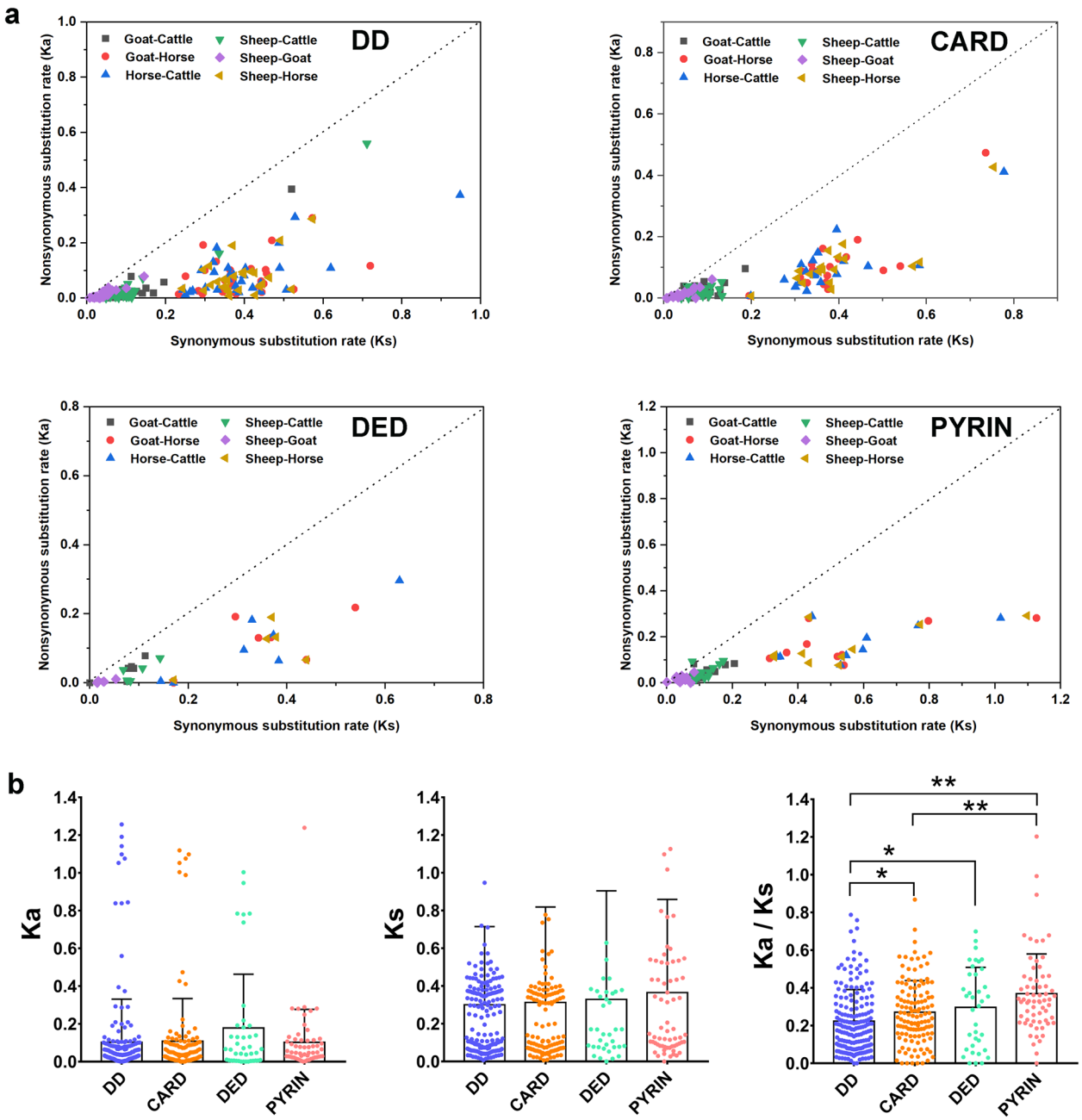


Fig. 4 Evolutionary constraints and selection pressures on DEATH superfamily members. **a** The Ka and Ks values of DD, CARD, DED, and PYRIN gene pairs, and these gene pairs from sheep-goat, sheep-cattle, sheep-horse, goat-horse, goat-cattle, and hose-cattle are marked with different colours. The dashed line in squares indicates Ka/Ks = 1. **b** Comparisons of the Ka, Ks, and Ka/Ks values of different gene pairs. * $P < 0.05$, ** $P < 0.01$

To explore the potential roles of DEATH members in the lactation period, we further compared the expression levels of DEATH genes in LP and Lactation. Compared to the LP period, half or more DEATH genes (DD: 71.4%, CARD: 65.0%, DED: 50.0%, and PYRIN: 88.9%) were upregulated in the lactation period, and these upregulations were independent of nutritional conditions (Fig. 7). This result implied that the DEATH superfamily may play crucial roles in lactation in the sheep mammary gland.

The 3D structures of the DD, CARD, DED, and PYRIN domains are highly evolutionarily conserved

To compare the structural characteristics of the DD, CARD, DED and PYRIN domains of different organisms,

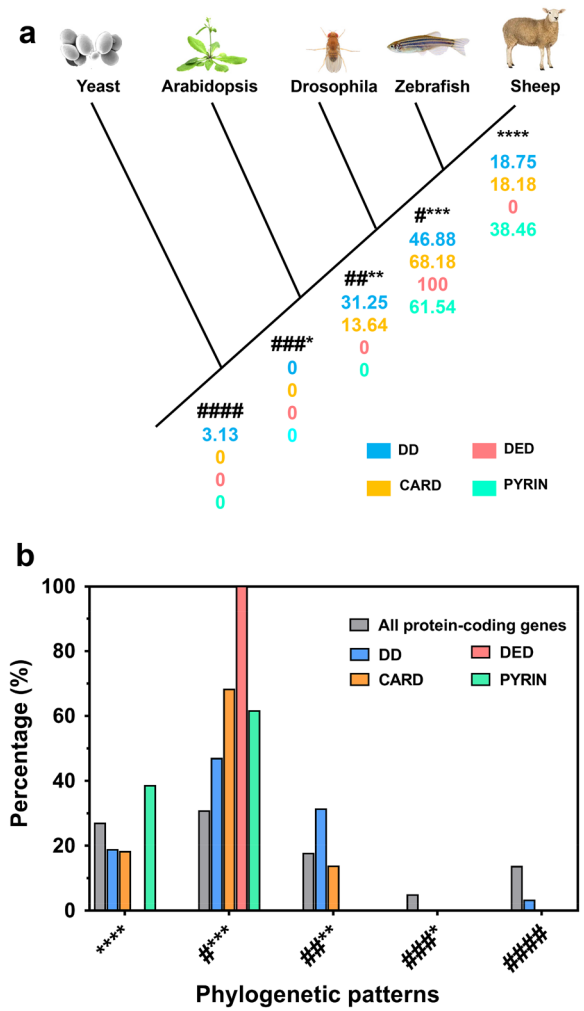


Fig. 5 The DEATH superfamily underwent more expansion at the late evolutionary stage. **a** The family phylogenetic patterns of the DD, CARD, DED, and PYRIN genes were analysed. The symbols "****", "#***", "##**", "###*", and "####" represent different phylogenetic patterns. The numbers labelled below the symbols separately represent the percentages of the DD, CARD, DED, and PYRIN genes. **b** The ratios of four subfamily genes and all sheep protein-coding genes in different phylogenetic patterns

we downloaded the 3D structures of these domains from PDB and AlphaFold database (Additional file 5: Table S5). All of these domains were composed of 5–7 α-helixes (Fig. 8a-d). Structural alignments showed that these helixes had similar spatial conformations in the domains of DD, CARD, DED and PYRIN, and the root mean square deviation (RMSD) between different domains from different species suggested that DD (average value = 3.02), CARD (average value = 3.62), DED (average value = 2.63), and PYRIN (average value = 1.72) were structurally conserved in biological evolution (Fig. 8 e, f). Surface conservation analyses showed that the DD, CARD, DED, and PYRIN domains contained many evolutionarily conserved residues on the surface (Fig. 9a). These conserved residues were randomly distributed in these domains (Fig. 9b). Statistical results indicated that the CARD domain contained more conserved surface residues than the other three domains (Fig. 9c).

Discussion

Sheep is an important livestock species, and sheep milk is a nutrient-rich food for humans. The characterization and identification of lactation phenotype-associated genes are crucial for the molecular breeding of sheep. Currently, genome-wide association study (GWAS)-based studies are the dominant approach to identify the molecular genetic basis related to the economic phenotype of livestock [35]. However, SNP-based GWASs rely on the individual information and cannot detect most genetic information or rare mutations with complex characteristics [36, 37]. Many previous reports have indicated that some members of the DEATH superfamily are involved in the regulation of lactation in mammals [38–40]. In this study, we employed a hypothesis-driven approach that focused on the relationship between the DEATH protein superfamily and the lactation traits of sheep. The genome-wide analysis of one particular gene family can provide us with a systematic understanding of this family and evaluate its functional roles in traits of interested. Herein, 74 DEATH superfamily members were identified in the sheep genome, and its four subfamilies were strongly correlated in evolution. The immune-related gene *IFI203*, which is a member of the PYRIN subfamily, was undergoing strong positive selection in sheep. Furthermore, these DEATH genes exhibited similar expression patterns under different nutritional conditions in the mammary gland in the late pregnancy and lactation period, and different subfamilies were positively correlated in expression levels. Additionally, most DEATH genes were upregulated during the lactation period in sheep, thus suggesting their important role in lactation. These data can help to understand the evolutionary history of the DEATH domain superfamily and provide new genetic markers for sheep breeds.

Coevolution represents the process of reciprocal evolution changes that occur between pairs or groups of species as they interact with each other [41]. At the molecular level, some functionally related proteins usually have coevolutionary linkages. The coevolutionary relationship between different protein families indicates protein interactions and functional relationships [42]. Moreover, the analyses of coevolved protein families prompted the prediction of protein interactions through the characteristics of coevolution. Coevolution analyses of protein families will be important in demonstrating

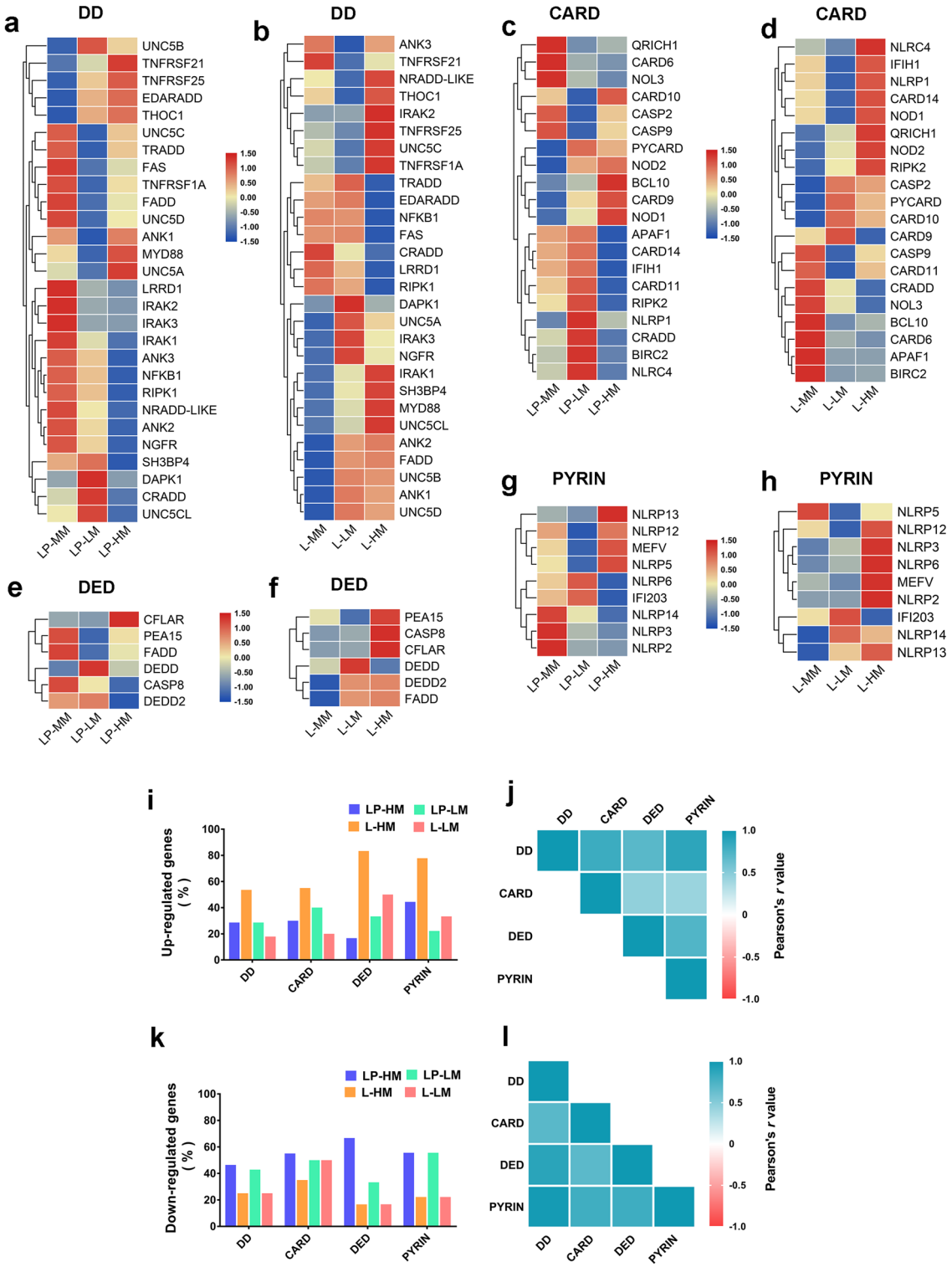


Fig. 6 Expression profiles of sheep DEATH genes under different nutritional conditions during late pregnancy and lactation periods. **a**–**h** Expression profiles of the DD, CARD, DED, and PYRIN genes in the late pregnancy (LP) and lactation (L) periods under high maintenance nutrition (HM), moderate maintenance nutrition (MM), and lower maintenance nutrition (LM). **i** The percentages of upregulated DD, CARD, DED, and PYRIN genes under different nutrition levels. **j** Pearson correlation analyses of the upregulated DD, CARD, DED, and PYRIN genes under different nutrition levels. **k** The percentages of downregulated DD, CARD, DED, and PYRIN genes under different nutritional levels. **l** Pearson's correlation analyses of downregulated DD, CARD, DED, and PYRIN genes under different nutritional levels

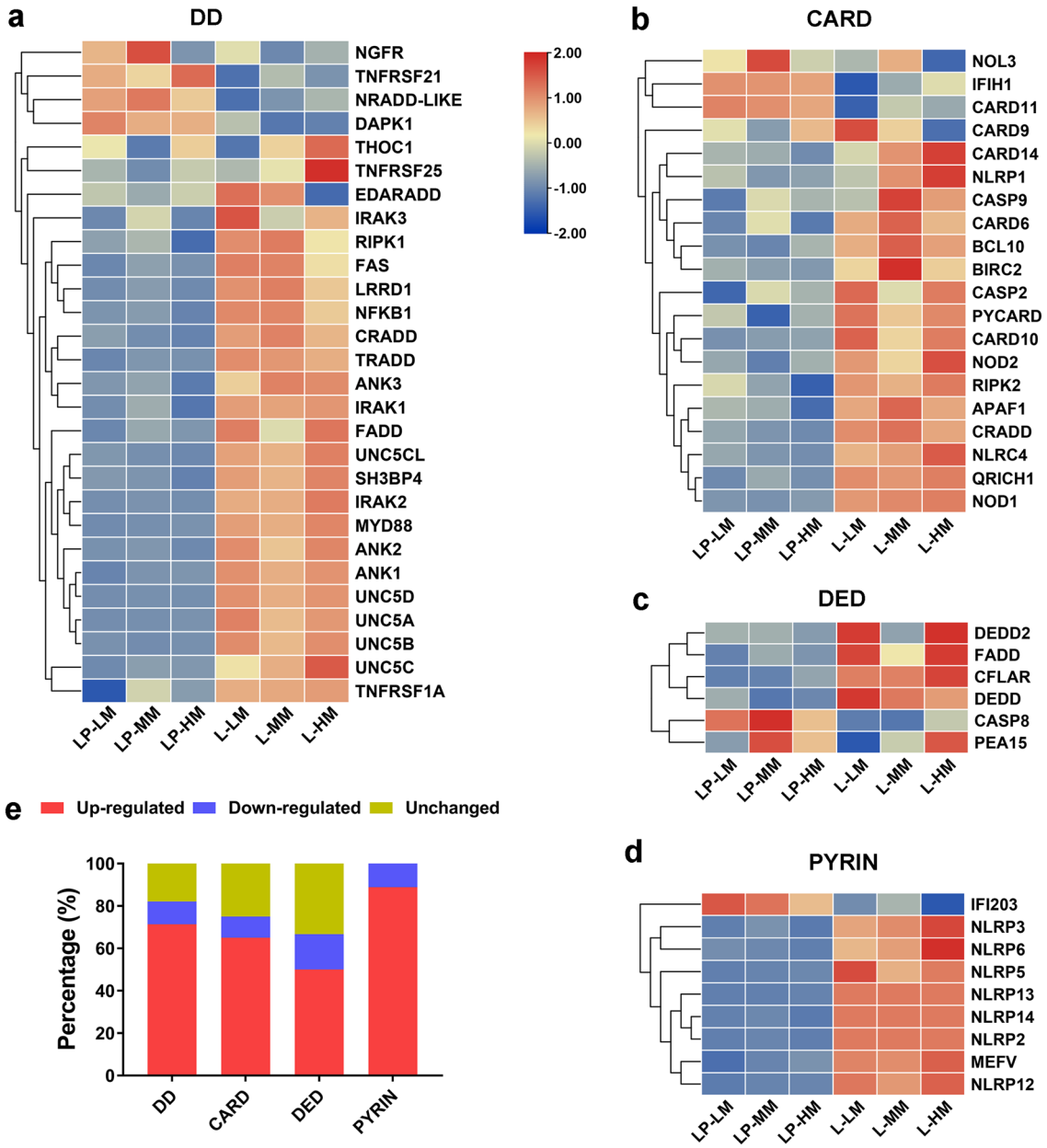


Fig. 7 Comparisons of the expressional characters of sheep DEATH genes during the late pregnancy and lactation periods. **a**–**d** Comparisons of the expression of DD, CARD, DED, and PYRIN genes in the late pregnancy (LP) and lactation (L) periods under HM, MM, and LM conditions. **e** The percentages of down- and upregulated DD, CARD, DED, and PYRIN genes in **a**-**d**

the molecular mechanisms underlying the similarity observed in phylogenetic trees of related proteins and in distinguishing direct molecular interactions from functional constraints. The MAGO and Y14 proteins are the core components of the exon junction complex, and both proteins play important roles in RNA metabolism and organism development. A previous report suggested that MAYO and Y14 are coevolved protein families in evolution [43]. Proliferating cell nuclear antigen (PCNA) coevolved with its binding partner across the phylogeny of fungi, which contributes to speciation in the fungal lineage [44]. In this study, the identification of DD, CARD, DED, and PYRIN subfamily members in vertebrates and invertebrates, as well as the phylogenetic analyses of the four subfamily members suggested that these four subfamilies coevolved in evolution. This coevolution contributes to the consistency in quantity and function of four subfamilies and further guarantees the normal operation

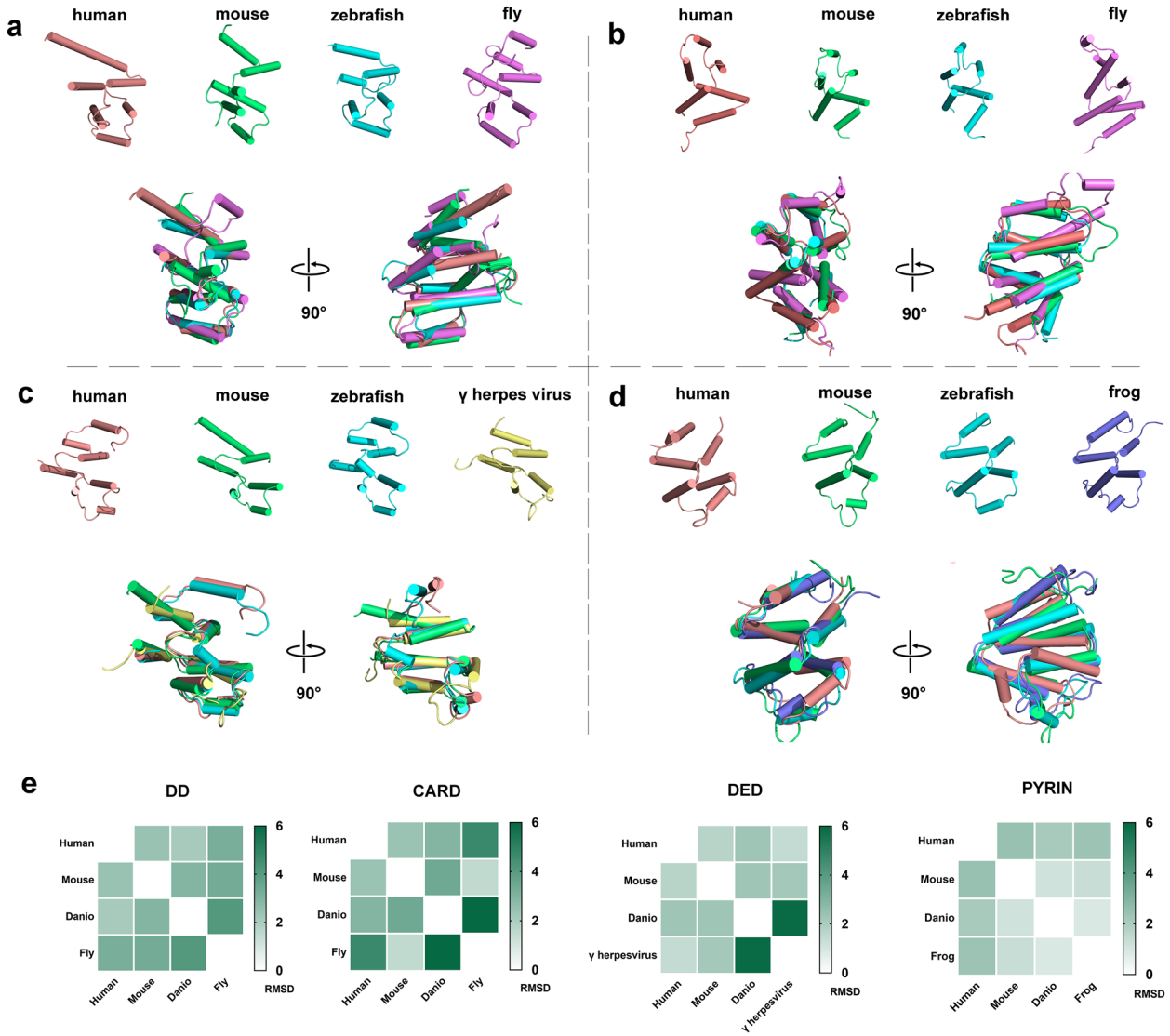


Fig. 8 The 3D structures of the DD, CARD, DED, and PYRIN domains were highly evolutionarily conserved. **a**–**d** 3D structures of the DD, CARD, DED, and PYRIN domains from human, mouse, zebrafish, and frog/fly/γ herpes virus and their structural alignments. The cylinder elements in each structure represented α-helices. **e** RMSD matrix of the DD, CARD, DED, and PYRIN domains of different organisms

of death signal transduction in eukaryotes, such as the apoptosis signal.

Epithelial apoptosis is one of the most important cellular events in the development of the mammary gland. Most secretory epithelial cells in the lactating mammary undergo dramatic apoptosis during mammary gland involution [45]. Many members of the DEATH superfamily play essential roles in many apoptotic events of mammary cells. For example, The Fas-associated death domain protein (FADD) is an important mediator of apoptosis, and the dominant negative FADD transgenic mice exhibited increased mammary secretory alveolar cell apoptosis and impaired milk production performance [38]. In addition, nucleotide oligomerization domain (NOD)-like receptors 1 (NOD1) is a CARD domain-containing protein, and CASP8 is a DED domain-containing protein; a recent study indicated that a high-concentrate diet can stimulate mammary cell apoptosis in dairy cows via the NOD1/CASP8 pathway [46]. In this study, we found that most members of the DEATH superfamily were upregulated during the lactation period in the mammary gland of sheep. A reasonable explanation for this scenario was that most DEATH family members were associated with apoptotic events, which were highly activated during the lactation period in the mammary

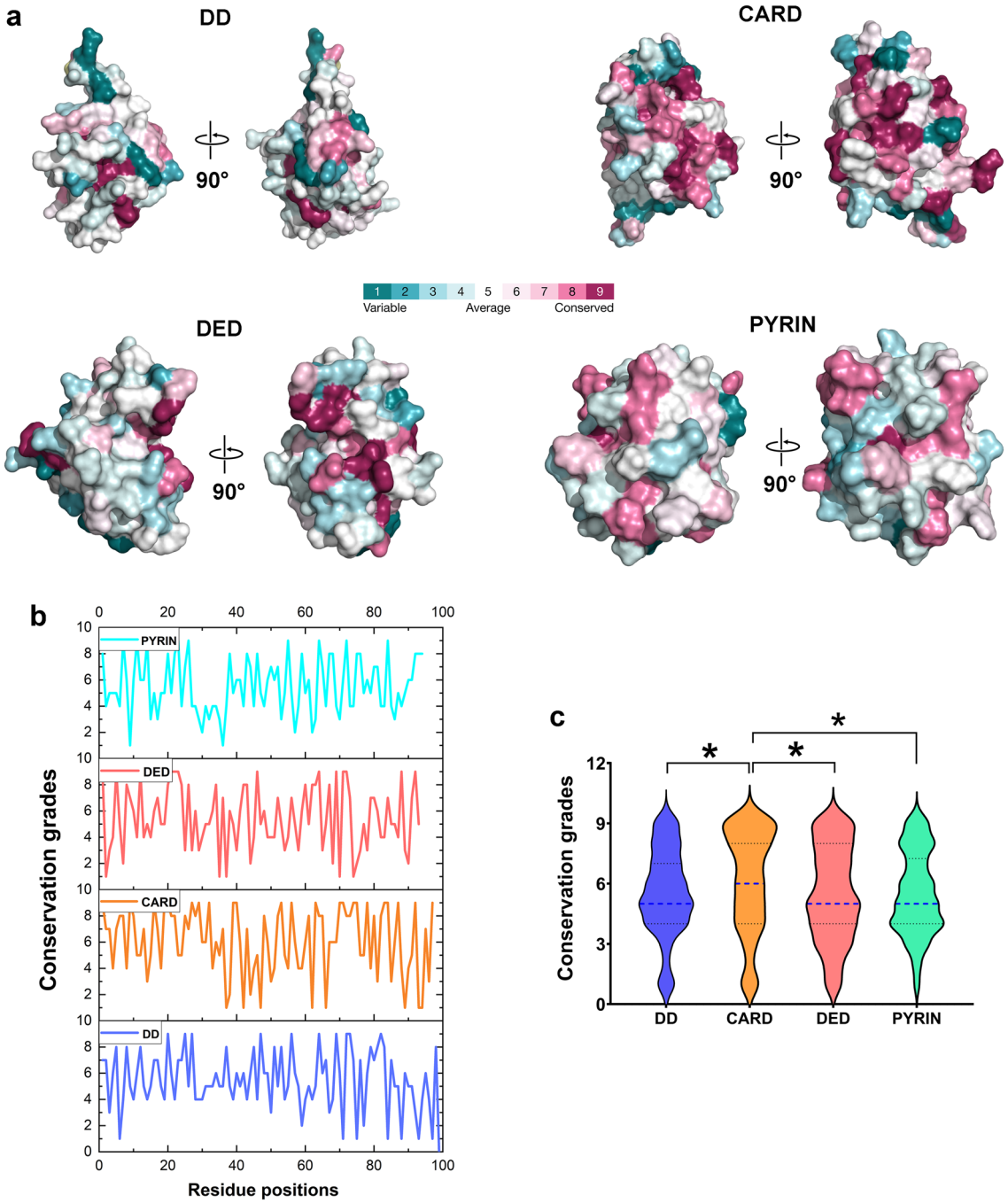


Fig. 9 Sequence conservation analyses of the DD, CARD, DED, and PYRIN domains. **a** Conserved residues on the surface of each domain. Cyan indicates variable residues, and purplish red indicates conserved residues. **b** Conservation profiles of all of the residues in each domain. Residues in each domain were scored from one to nine based on sequence alignments that were achieved in the Uniref-90 blast database. **c** Statistical analyses of the residue conservation grades of four domains

gland [45]. In addition, nutritional conditions could affect the expression of DEATH family members. Interestingly, we found the expression levels of DD, CARD, DED, and PYRIN subfamily members were correlated at different nutritional levels. Previous reports suggest that nutrition could influence mammary epithelial cell apoptosis through the oxidative stress imposed by feed and the tissue's ability to prevent damage caused by reactive oxygen species (ROS) from feed [9, 47]. We speculate that these apoptotic events induced by

oxidative stress furtherly regulate the expression of apoptosis-related Death superfamily members, finally leading to the correlated expression patterns that regulated by nutrition levels.

In this study, the Ka/Ks analyses showed that negative selection was the main selection pressure acting on most members of the DEATH superfamily. Unusually, we found that the PYRIN domain-containing protein IFI203 was undergoing significant positive selection in sheep. IFI203 belongs to the interferon-inducible protein family and functions as a regulator of lipid homeostasis and triacylglycerol accumulation [48]. Lipids are one of the most important components of milk, and sheep milk contains more fat than goat and cow milk, comprising approximately 7 g total fat in 100 g milk [49]. Today, an increasing number of sheep are raised throughout the world for sheep milk and wool. We speculated that IFI203 is under positive selection due to its crucial role in lipid metabolism in sheep.

The tertiary structures of evolutionarily conserved proteins are often more similar than the primary structures; therefore, analyzing the 3D structure of protein domains is necessary to study protein conservation. Ordinarily, protein domains retain its own molecular functions independent of the rest components of protein [50]. Various protein domains that play an important role in biological activities were structurally conserved in evolution, such as the Toll/IL-1 receptor (TIR) domain, DEP domain, and Wnt domain [2, 25, 26]. It is known that proteins can exert molecular functions in proper conformational states, and conformational divergence usually leads to functional discrepancies [51]. Apoptosis is an essential and conserved biological process that occurs in almost all multicellular organisms. These DD, CARD, DED, and PYRIN domain-containing proteins are crucial components in apoptosis. Herein, we found that these DD, CARD, DED, and PYRIN domains were conformationally conserved in evolution. The structural conservation of these domains guaranteed the molecular functional consistency of the DEATH superfamily members in cellular apoptotic events, and the expansion of the DEATH superfamily further contributed to the mechanical complexity and component diversity of apoptosis in mammals.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40538-023-00406-7.

Additional file 1: Table S1. Characterization of DD domain-containing protein coding genes in sheep.

Additional file 2: Table S2. Characterization of CARD domain-containing protein coding genes in sheep.

Additional file 3: Table S3. Characterization of DED domain-containing protein coding genes in sheep.

Additional file 4: Table S4. Characterization of PYRIN domain-containing protein coding genes in sheep.

Additional file 5: Table S5. Three-dimensional structures of the DD, CARD, DED, and PYRIN domains here used.

Acknowledgements
Not applicable.

Author contributions
Z.G.: Conceptualization, Investigation, Writing and original draft preparation. S.H.: Investigation, Writing and original draft preparation. Y.H.: Investigation. G.G.: Conceptualization, Supervision, Validation. J.Z.: Conceptualization, Data Curation, Supervision. All authors read and approved the final manuscript.

Funding
This work was supported by the National Natural Science Foundation of China (No: 32200974, 82200420), Natural Science Foundation of Shaanxi Province (No: 2020JQ-715), the Scientific Research Foundation of Tangdu Hospital (No: 2021ZTXM-026).

Availability of data and materials
All data generated or analysed during this study are included in this published article and its supplementary information files.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interest.

Author details
[1]School of Food and Biological Engineering, Shaanxi University of Science and Technology, Xi'an 710021, China. [2]Department of Cardiology, The Second Affiliated Hospital of Air Force Medical University, Xi'an 710038, China.

Received: 21 February 2023 Accepted: 30 April 2023
Published online: 10 May 2023

References

1. Pulina G, et al. Effects of nutrition on the contents of fat, protein, somatic cells, aromatic compounds, and undesirable substances in sheep milk. Anim Feed Sci Technol. 2006;131(3):255–91. https://doi.org/10.1016/j.anifeedsci.2006.05.023.
2. Gai Z, et al. Whole genome-wide analysis of DEP family members in sheep (Ovis aries) reveals their potential roles in regulating lactation. Chem Biol Technol Agric. 2022;9(1):68. https://doi.org/10.1186/s40538-022-00336-w.
3. Ni Y, et al. Three lactation-related hormones: regulation of hypothalamus-pituitary axis and function on lactation. Mol Cell Endocrinol. 2021;520:111084. https://doi.org/10.1016/j.mce.2020.111084.
4. Klier CM, et al. The role of estrogen and progesterone in depression after birth. J Psychiatr Res. 2007;41(3–4):273–9. https://doi.org/10.1016/j.jpsychires.2006.09.002.
5. Scandolo DG, et al. Efficacy of estradiol or GnRH in combination with progesterone intravaginal devices to control the follicular wave dynamics and resulting fertility in lactating dairy cows. Anim Reprod Sci. 2020;223:106646. https://doi.org/10.1016/j.anireprosci.2020.106646.

6. Wang Y, et al. Bta-miR-34b controls milk fat biosynthesis via the Akt/mTOR signaling pathway by targeting RAI14 in bovine mammary epithelial cells. J Anim Sci Biotechnol. 2021;12(1):83. https://doi.org/10.1186/s40104-021-00598-8.
7. King K, et al. Dynamics of lipid droplets in the endometrium and fatty acids and oxylipins in the uterine lumen, blood, and milk of lactating cows during diestrus. J Dairy Sci. 2021;104(3):3676–92. https://doi.org/10.3168/jds.2020-19196.
8. Gai Z, et al. Downregulation of CASTOR1 inhibits heat-stress-induced apoptosis and promotes casein and lipid synthesis in mammary epithelial cells. J Agric Food Chem. 2022;70(17):5386–95. https://doi.org/10.1021/acs.jafc.2c00877.
9. Stefanon B, et al. Mammary apoptosis and lactation persistency in dairy animals. J Dairy Res. 2002;69(1):37–52. https://doi.org/10.1017/s0022029901005246.
10. Hall AE, et al. The cytoskeleton adaptor protein ankyrin-1 is upregulated by p53 following DNA damage and alters cell migration. Cell Death Dis. 2016;7:e2184. https://doi.org/10.1038/cddis.2016.91.
11. Ray P, et al. Crocetin exploits p53-induced death domain (PIDD) and FAS-associated death domain (FADD) proteins to induce apoptosis in colorectal cancer. Sci Rep. 2016;6:32979. https://doi.org/10.1038/srep32979.
12. Thompson R, et al. An Inhibitor of PIDDosome Formation. Mol Cell. 2015;58(5):767–79. https://doi.org/10.1016/j.molcel.2015.03.034.
13. Williams ME, et al. UNC5A promotes neuronal apoptosis during spinal cord development independent of netrin-1. Nat Neurosci. 2006;9(8):996–8. https://doi.org/10.1038/nn1736.
14. Freitas C, Larrivee B, Eichmann A. Netrins and UNC5 receptors in angiogenesis. Angiogenesis. 2008;11(1):23–9. https://doi.org/10.1007/s10456-008-9096-2.
15. So T, Ishii N. The TNF-TNFR family of Co-signal molecules. Adv Exp Med Biol. 2019;1189:53–84. https://doi.org/10.1007/978-981-32-9717-3_3.
16. Fu Q, et al. miR-29a up-regulation in AR42J cells contributes to apoptosis via targeting TNFRSF1A gene. World J Gastroenterol. 2016;22(20):4881–90. https://doi.org/10.3748/wjg.v22.i20.4881.
17. Kumar S. Caspase 2 in apoptosis, the DNA damage response and tumour suppression: enigma no more? Nat Rev Cancer. 2009;9(12):897–903. https://doi.org/10.1038/nrc2745.
18. Wurstle ML, Laussmann MA, Rehm M. The central role of initiator caspase-9 in apoptosis signal transduction and the regulation of its activation and activity on the apoptosome. Exp Cell Res. 2012;318(11):1213–20. https://doi.org/10.1016/j.yexcr.2012.02.013.
19. Fritsch M, et al. Caspase-8 is the molecular switch for apoptosis, necroptosis and pyroptosis. Nature. 2019;575(7784):683–7. https://doi.org/10.1038/s41586-019-1770-6.
20. Newton K, et al. Cleavage of RIPK1 by caspase-8 is crucial for limiting apoptosis and necroptosis. Nature. 2019;574(7778):428–31. https://doi.org/10.1038/s41586-019-1548-x.
21. Li C, et al. Caspase-8 mutations in head and neck cancer confer resistance to death receptor-mediated apoptosis and enhance migration, invasion, and tumor growth. Mol Oncol. 2014;8(7):1220–30. https://doi.org/10.1016/j.molonc.2014.03.018.
22. Shimada K, et al. Oxidized mitochondrial DNA activates the NLRP3 inflammasome during apoptosis. Immunity. 2012;36(3):401–14. https://doi.org/10.1016/j.immuni.2012.01.009.
23. Wang J, et al. NLRP3 inflammasome mediates angiotensin II-induced islet beta cell apoptosis. Acta Biochim Biophys Sin (Shanghai). 2019;51(5):501–8. https://doi.org/10.1093/abbs/gmz032.
24. Rossi MN, et al. NLRP2 regulates proinflammatory and antiapoptotic responses in proximal tubular epithelial cells. Front Cell Dev Biol. 2019;7:252. https://doi.org/10.3389/fcell.2019.00252.
25. Gai Z, et al. Whole genome level analysis of the Wnt and DIX gene families in mice and their coordination relationship in regulating cardiac hypertrophy. Front Genet. 2021;12:608936. https://doi.org/10.3389/fgene.2021.608936.
26. Zhao J, et al. Genome-wide identification and characterization of Toll-like receptors (TLRs) in housefly (Musca domestica) and their roles in the insecticide resistance. Int J Biol Macromol. 2020;150:141–51. https://doi.org/10.1016/j.ijbiomac.2020.02.061.
27. Wang Y, et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 2012;40(7):e49. https://doi.org/10.1093/nar/gkr1293.
28. Wang D, et al. KaKs_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. Genomics Proteomics Bioinformatics. 2010;8(1):77–80. https://doi.org/10.1016/s1672-0229(10)60008-3.
29. Zou XD, et al. Genome-wide analysis of WD40 protein family in human. Sci Rep. 2016;6:39262. https://doi.org/10.1038/srep39262.
30. Persson E, Sonnhammer ELL. InParanoiDB 9: ortholog groups for protein domains and full-length proteins. J Mol Biol. 2023. https://doi.org/10.1016/j.jmb.2023.168001.
31. Jumper J, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596(7873):583–9. https://doi.org/10.1038/s41586-021-03819-2.
32. Fryxell KJ. The coevolution of gene family trees. Trends Genet. 1996;12(9):364–9. https://doi.org/10.1016/s0168-9525(96)80020-5.
33. Sosa-Castillo E, Rodriguez-Cruz M, Molto-Puigmarti C. Genomics of lactation: role of nutrigenomics and nutrigenetics in the fatty acid composition of human milk. Br J Nutr. 2017;118(3):161–8. https://doi.org/10.1017/S0007114517001854.
34. Gessner DK, et al. Expression of target genes of nuclear factor E2-related factor 2 in the liver of dairy cows in the transition period and at different stages of lactation. J Dairy Sci. 2013;96(2):1038–43. https://doi.org/10.3168/jds.2012-5967.
35. Alqudah AM, et al. GWAS: fast-forwarding gene identification and characterization in temperate Cereals: lessons from Barley - A review. J Adv Res. 2020;22:119–35. https://doi.org/10.1016/j.jare.2019.10.013.
36. Tam V, et al. Benefits and limitations of genome-wide association studies. Nat Rev Genet. 2019;20(8):467–84. https://doi.org/10.1038/s41576-019-0127-1.
37. Li YR, Keating BJ. Trans-ethnic genome-wide association studies: advantages and challenges of mapping in diverse populations. Genome Med. 2014;6(10):91. https://doi.org/10.1186/s13073-014-0091-5.
38. Shackleton M, et al. Impaired lactation in mice expressing dominant-negative FADD in mammary epithelium. Dev Dyn. 2009;238(4):1010–6. https://doi.org/10.1002/dvdy.21917.
39. Adams HA, et al. Identification of a nonsense mutation in APAF1 that is likely causal for a decrease in reproductive efficiency in Holstein dairy cattle. J Dairy Sci. 2016;99(8):6693–701. https://doi.org/10.3168/jds.2015-10517.
40. Tan X, et al. Down-regulation of NOD1 in neutrophils of periparturient dairy cows. Vet Immunol Immunopathol. 2012;150(1–2):133–9. https://doi.org/10.1016/j.vetimm.2012.09.004.
41. Groussin M, Mazel F, Alm EJ. Co-evolution and Co-speciation of Host-Gut bacteria systems. Cell Host Microbe. 2020;28(1):12–22. https://doi.org/10.1016/j.chom.2020.06.013.
42. Ochoa D, Pazos F. Practical aspects of protein co-evolution. Front Cell Dev Biol. 2014;2:14. https://doi.org/10.3389/fcell.2014.00014.
43. Gong P, Zhao M, He C. Slow co-evolution of the MAGO and Y14 protein families is required for the maintenance of their obligate heterodimerization mode. PLoS One. 2014;9(1):e84842. https://doi.org/10.1371/journal.pone.0084842.
44. Zamir L, et al. Tight coevolution of proliferating cell nuclear antigen (PCNA)-partner interaction networks in fungi leads to interspecies network incompatibility. Proc Natl Acad Sci U S A. 2012;109(7):E406–14. https://doi.org/10.1073/pnas.1108633109.
45. Jena MK, et al. Molecular mechanism of mammary gland involution: an update. Dev Biol. 2019;445(2):145–55. https://doi.org/10.1016/j.ydbio.2018.11.002.
46. Ul Aabdin Z, et al. High-Concentrate feeding to dairy cows induces apoptosis via the NOD1/Caspase-8 pathway in mammary epithelial cells. Genes (Basel). 2020. https://doi.org/10.3390/genes11010107.
47. Dai W, et al. Organelles coordinate milk production and secretion during lactation: insights into mammary pathologies. Prog Lipid Res. 2022;86:101159. https://doi.org/10.1016/j.plipres.2022.101159.
48. Norheim F, et al. Genetic regulation of liver lipids in a mouse model of insulin resistance and hepatic steatosis. Mol Syst Biol. 2021;17(1):e9684. https://doi.org/10.15252/msb.20209684.
49. Rasika DMD, et al. Probiotics and prebiotics in non-bovine milk. Adv Food Nutr Res. 2020;94:339–84. https://doi.org/10.1016/bs.afnr.2020.06.008.

50. Korasick D, Jez JM. Protein domains: structure, function, and methods. Encyclopedia Cell Biol. 2016. https://doi.org/10.1016/B978-0-12-394447-4.10011-2.
51. Sakamaki K, et al. Conservation of structure and function in vertebrate c-FLIP proteins despite rapid evolutionary change. Biochem Biophys Rep. 2015;3:175–89. https://doi.org/10.1016/j.bbrep.2015.08.005.

Publisher's Note