

DOI:10.12113/202012013

牦牛 *NLRP* 基因家族全基因组鉴定及转录组分析

鲍 麒^{1,2}, 张晓兰^{1,2}, 黄 纯^{1,2}, 梁春年^{1,2}, 李吉叶³, 褚 敏^{1,2}, 阎 萍^{1,2*}

(1.中国农业科学院 兰州畜牧与兽药研究所, 兰州 730050; 2.甘肃省牦牛繁育工程重点实验室, 兰州 730050; 3.青海大通种牛场, 西宁 810000)

摘 要: *NLRP* 基因家族在哺乳动物基因组中分布广泛, 在免疫应答和生殖相关过程中发挥重要作用。本研究根据已报道的人类 *NLRP* 基因, 利用 PFAM、CDD、UniProt 等数据库, 采用 ClustalW、MEGA X、ExPASy、MEME、GSDS 和 STRING 等软件对牦牛 *NLRP* 家族成员进行全基因组鉴定和分析, 并且基于高通量测序数据分析 *NLRP* 基因在牦牛不同组织中的表达情况。研究共鉴定出 8 个不同的 *NLRP* 基因家族成员, 包括 *NLRP2*、*NLRP3*、*NLRP5*、*NLRP6*、*NLRP8*、*NLRP9*、*NLRP12*、*NLRP14*, 其中 *NLRP9* 有 2 个可变剪切。牦牛 *NLRP* 蛋白均有多个磷酸化位点, GRAVY 值均小于 0, 属于亲水性蛋白; 不稳定系数均大于 40, 均为不稳定蛋白, 无跨膜结构。系统进化树显示同一家族成员氨基酸序列相似性高, 在 6 个物种中都聚在一起, 免疫相关基因与生殖发育相关基因在进化中已经分开。蛋白互作网络显示牦牛 *NLRP* 蛋白与其他物种已经研究的功能相似, 不仅与参与免疫相关的蛋白互作, 也与生殖标记蛋白互作。如 *NLRP3* 与多种参与炎症反应的蛋白 PYCARD、MLKL、GSDMD 互作, *NLRP5* 也与生殖细胞标记蛋白 GDF9、BMP15、ZAR1 等互作。转录组分析结果显示, 牦牛 *NLRP* 基因在 6 种组织中的表达较为微弱或者不表达, 具有组织特异性。牦牛 *NLRP* 家族在结构及理化性质上与其他物种同源基因具有相同的特点, 部分试验和预测证实其同样参与了免疫和生殖相关过程。本研究结论不仅对牦牛 *NLRP* 基因家族进行了系统分析, 同时为进一步研究牦牛先天免疫和生殖相关过程提供理论依据。

关键词: *NLRP*; 牦牛; 生物信息; 基因家族; 全基因组鉴定

中图分类号: S828 **文献标志码:** A **文章编号:** 1672-5565(2022)02-113-11

Genome-wide identification and transcriptome analysis of *NLRP* gene family of yak

BAO Qi^{1,2}, ZHANG Xiaolan^{1,2}, HUANG Chun^{1,2}, LIANG Chunnian^{1,2}, LI Jiye³, CHU Min^{1,2}, YAN Ping^{1,2*}

(1.Lanzhou Institute of Husbandry and Pharmaceutical Sciences, Chinese Academy of Agricultural Sciences, Lanzhou 730050, China; 2.Key Laboratory of Yak Breeding Engineering, Lanzhou 730050, China; 3.Datong Cattle Farm in Qinghai Province, Xining 810000, China)

Abstract: *NLRP* genes are widely distributed in mammalian genomes and play an important role in immune response and reproduction-related processes. According to the reported human *NLRP* genes, based on PFAM, CDD, UniProt, and other databases, the ClustalW, MEGA X, ExPASy, MEME, GSDS, and STRING software systems were used to identify and analyze the *NLRP* family members of yak, and based on high-throughput sequencing data, the *NLRP* gene expression in different tissues of yak was analyzed. A total of eight different *NLRP* gene family members were identified, including *NLRP2*, *NLRP3*, *NLRP5*, *NLRP6*, *NLRP8*, *NLRP9*, *NLRP12*, and *NLRP14*, among which *NLRP9* had two variable splicings. The *NLRP* proteins had multiple phosphorylation sites, and the GRAVY values were all less than 0, which were hydrophilic proteins; the instability coefficients were all greater than 40, and they were all unstable proteins with no transmembrane structure. The phylogenetic tree showed that the amino acid sequences of members of the same family were highly similar, and they were clustered in six species. The immune-related genes and the genes related to reproductive development had been separated in evolution. The protein interaction network showed that members of the yak *NLRP* family had similar functions with those studied in

收稿日期: 2020-12-31; 修回日期: 2021-06-26.

基金项目: 中国农业科学院创新工程项目 (No.CAAS-ASTIP-2014-LIHPS-01); 科技援青合作专项 (No.2020-QY-212); 现代农业 (肉牛牦牛) 产业技术体系建设专项资金 (No.CARS-37); 甘肃省科技计划资助 (No.20JR5RA580).

作者简介: 鲍麒, 男, 硕士研究生, 研究方向: 动物基因组学. E-mail: baoqin@163.com.

*** 通信作者:** 阎萍, 女, 研究员, 研究方向: 动物遗传育种与繁殖. E-mail: pingyanlz@163.com.

other species. They not only interacted with immune-related proteins but also with reproductive marker proteins. For example, NLRP3 interacted with a variety of proteins (PYCARD, MLKL, and GSDMD) involved in inflammation, and NLRP5 also interacts with germ cell marker proteins such as GDF9, BMP15, and ZAR1. The results of transcriptome analysis showed that the expression of yak *NLRP* gene in six kinds of tissues was weak or not expressed, which was tissue-specific. The yak *NLRP* family had the same characteristics as homologous genes of other species in terms of structure and physical and chemical properties. Some experiments and predictions have confirmed that it is also involved in immune and reproduction-related processes. The study not only conducted a systematic analysis of the yak *NLRP* gene family, but also provided theoretical basis for further research on the innate immunity and reproduction-related processes of yak.

Keywords: NLRP; Yak; Bioinformatics; Gene family; Genome-wide identification

NLRP 蛋白家族最早在人上被发现,共有 14 个家族成员,可以分为免疫相关蛋白和生殖相关蛋白。免疫相关蛋白主要在先天免疫系统发挥功能,包括 NLRP1、NLRP3、NLRP6、NLRP10 和 NLRP12,生殖相关蛋白主要在哺乳动物生殖系统中起作用,包括 NLRP2、NLRP4、NLRP5、NLRP7、NLRP8、NLRP9、NLRP11、NLRP13 和 NLRP14^[1]。NLRP 蛋白的基本结构主要由 3 个保守的结构域组成,第一个是氨基端的 PYD(Pyrin domain),该结构域包含 92 个氨基酸,主要功能是介导蛋白之间的相互作用,第二个是中间的 NACHT 结构域,该结构域是以 4 个相近的含核苷水解酶域的蛋白合起来命名的,包括神经元凋亡抑制蛋白(Neuronal apoptosis inhibitor protein, NAIP)、II 型主要组织相容性复合体转录激活子(Class II major histocompatibility complex trans-activator, CII TA)、来源于柄孢霉的非亲和位点蛋白(HET-E)、哺乳动物端粒酶相关蛋白(Telomerase associated protein 1, TP1)。最后是羧基端的 LRRs(Leucine-rich repeats)结构域^[2],该结构域含有 9-12 个亮氨酸重复,主要功能是介导自身调节、蛋白间的相互作用和病原相关分子模式的传感。

近年来对 NLRP 基因的研究主要集中在炎症小体上。NLRP 家族中的炎症体包含 NLRP1 和 NLRP3 炎症小体。炎症小体是由胞浆内模式识别受体参与组装的多蛋白复合物。炎症小体活化后可促进炎症因子的成熟和释放,从而参与先天免疫过程。炎症小体募集天冬氨酸蛋白水解酶-1(Caspase-1),引起 caspase-1 蛋白水解 pro-IL-1b 和 pro-IL-18,产生成熟的活性细胞因子^[3]。NLRP1 炎症体由感受器 NLRP1、结合蛋白 ASE 和效应蛋白 Caspase-1 三个部分组成,能够促进 IL-18, IL-1b 及 IL-33 的成熟,在炎症发生时起作用^[3]。在 NLRP 蛋白中, NLRP10 是唯一一个缺乏 LRR 结构域的家族成员。有研究表明,该蛋白是炎症和细胞凋亡的多功能负调节剂,能激活促炎性半胱天冬酶、半胱天冬酶-1

和半胱天冬酶-5^[4]。另外,与生殖相关的 *NLRP* 基因也备受关注。小鼠中 *Nlrp2* 是卵母细胞质量的关键调节因子,敲除掉该基因的雌性小鼠的繁殖率会随着年龄的增长而显著下降^[5]。*Nlrp4* 和 *Nlrp9* 在小鼠上具有特异性复制的特点,在小鼠中 *Nlrp4* 有 7 个拷贝基因,*Nlrp9* 有 3 个拷贝基因^[6]。*Nlrp4g* 参与了小鼠早期胚胎的发育,在卵母细胞中特异性表达^[7]。*Nlrp9* 的复制基因可能没有具体的功能,有研究发现同时敲除 *Nlrp9a* 和 *Nlrp9e* 或在合子中沉默 *Nlrp9b* 对小鼠早期胚胎发育没有明显的影响,表明在小鼠早期胚胎发育过程中可能不需要 *Nlrp9* 的参与^[8]。啮齿类动物中没有 *Nlrp7*,被认为是在物种进化过程中由 *Nlrp2* 复制而形成的,二者基因位点相近,由 *Nlrp2* 执行 *Nlrp7* 的功能^[9]。*NLRP7* 不仅在免疫活动中发挥作用,还是一种母源印记基因,相关研究表明 *NLRP7* 印记状态的异常会导致人胚胎发育的异常^[10],从而造成葡萄胎或者流产现象的发生。

牦牛是生活在海拔 2 500-6 000 m 的青藏高原及其邻近地区的一种独特的家畜,自古以来,牦牛是青藏高原地区牧民必不可缺的家畜,为高原地区的牧民提供了吃、住、行等基本生存保证^[11]。牦牛能适应高海拔地区的严寒、缺氧、缺草等恶劣自然条件,是哺乳动物中研究高海拔适应性的理想物种^[12]。与大部分一年一胎的牛科动物相比,牦牛性成熟时间晚^[13],生殖性能差,每两年生育一胎或每三年生育两胎。其次,牦牛体外受精的囊胚发育率低,影响牦牛生产。近年来,多个 *NLRP* 基因在牦牛生殖相关的研究中被报道,如陈虹^[14]在牦牛卵巢蛋白质分析中鉴定到 *NLRP5* 在卵母细胞和受精卵中高表达;杨芳^[15]通过对牦牛和犏牛睾丸的比较转录组分析发现 *NLRP14* 的上调可能会造成原始生殖细胞或未分化精原细胞的积累,进而加剧生殖细胞凋亡,表明该家族基因在牦牛上也发挥重要作用。目前,*NLRP* 基因家族在人、小鼠基因组上被完整鉴

定,但在牦牛上尚无报道。本研究对牦牛 *NLRP* 基因家族进行系统性分析,可为进一步研究牦牛 *NLRP* 基因家族的功能奠定基础。

1 材料与方法

1.1 牦牛 *NLRP* 基因家族的鉴定

从 UniProt 数据库^[16]下载已知人类 *NLRP* 蛋白序列(见表 1),使用 BLSAT+软件^[17]构建本地数据

库,同源比对牦牛蛋白序列。牦牛参考基因组、蛋白序列和注释文件来源于 NCBI (<https://www.ncbi.nlm.nih.gov/genome/?term=wild+yak>)。同时利用 HMMER 软件对人 *NLRP* 蛋白序列建立 HMM 模型,在牦牛全基因组蛋白序列中搜索同源序列,筛选阈值设置为 $E\text{-value}<1\times10^{-5}$,去除可变剪切后,同时满足两种鉴定方法的序列被认为是牦牛 *NLRP* 基因家族成员,基因名字以基因组注释文件给定的名称命名。

表 1 人类 *NLRP* 蛋白信息
Table 1 Human *NLRP* proteins

缩写	蛋白名字	UniProt 数据库登录号
NLRP1	NACHT, LRR, and PYD domains-containing protein 1	Q9C000
NLRP2	NACHT, LRR, and PYD domains-containing protein 2	Q9NX02
NLRP3	NACHT, LRR, and PYD domains-containing protein 3	Q96P20
NLRP4	NACHT, LRR, and PYD domains-containing protein 4	Q96MN2
NLRP5	NACHT, LRR, and PYD domains-containing protein 5	P59047
NLRP6	NACHT, LRR and PYD domains-containing protein 6	P59044
NLRP7	NACHT, LRR, and PYD domains-containing protein 7	Q8WX94
NLRP8	NACHT, LRR, and PYD domains-containing protein 8	Q86W28
NLRP9	NACHT, LRR, and PYD domains-containing protein 9	Q7RTR0
NLRP10	NACHT, LRR, and PYD domains-containing protein 10	Q86W26
NLRP11	NACHT, LRR and PYD domains-containing protein 11	P59045
NLRP12	NACHT, LRR and PYD domains-containing protein 12	P59046
NLRP13	NACHT, LRR, and PYD domains-containing protein 13	Q86W25
NLRP14	NACHT, LRR, and PYD domains-containing protein 14	Q86W24

1.2 牦牛 *NLRP* 基因家族系统进化分析

采用 ClustalW^[18]对鉴定出的牦牛 *NLRP* 蛋白进行序列比对,利用 MEGA-X 软件^[19]对牦牛、小鼠、大鼠、狗、牛、人 *NLRP* 蛋白序列进行多重序列比对,采用最大似然法构建系统进化树。最后使用 Evolview 软件(<https://evolgenius.info//evolview-v2/#login>)对进化树进行美化处理。

1.3 牦牛 *NLRP* 蛋白理化性质分析

利用在线生物信息学软件 Expasy Prototparam^[20]、SignalP 5.0 Server^[21]、TMHMM Server v.2.0(<http://www.cbs.dtu.dk/services/TMHMM/>)、Wolf PSORT^[22]分别预测 *NLRP* 蛋白长度、分子量、等电点、不稳定系数、信号肽、跨膜结构和蛋白亚细胞定位情况。采用在线工具 GSDS^[23]对牦牛 *NLRP* 基因结构进行可视化分析。

1.4 基因结构和保守结构域分析

使用 NCBI 在线工具 CD-search^[24]和 PFAM 数据库^[25]对鉴定出的牦牛 *NLRP* 蛋白进行保守结构域鉴定。利用 MEME^[26]分析 *NLRP* 保守基序(默认

参数),并用 TBtools^[27]进行展示。运用 NetPhos 3.1 Server(<http://www.cbs.dtu.dk/services/NetPhos/>)分析 *NLRP* 蛋白的磷酸化位点,并且统计各个位点的个数。采用 SOPMA^[28]预测牦牛 *NLRP* 蛋白的二级结构。使用 STRING 数据库^[29]查询牦牛 *NLRP* 的蛋白互作网络。

1.5 牦牛 *NLRP* 基因家族的表达分析

牦牛不同组织的转录组数据下载自 NCBI 数据库(PRJNA362606)^[30]。首先,利用 fastq-dump 软件(<https://ncbi.github.io/sra-tools/fastq-dump.html>)将下载得到的 SRA 文件转换为 fastq 格式文件,对于上述所有过滤后的 RNA-Seq 数据,使用 fastp^[31]对下载的原始数据需要进行质控和预处理,得到的 clean data 用于后续分析。在 Hisat2(2.1.0 版)中使用 Hisat2-build 命令^[32],默认参数下构建基因组索引。随后利用 featureCounts 计数,得到每个基因在各个样本中的 raw counts。在这项研究中,我们使用 FPKM 作为一个标准化的转录丰度(Fragments Per Kilobase of exon model per Million mapped fragments),

利用 R 脚本根据各个样本的 raw counts 计算出基因的表达量构建出基因表达矩阵,共 18 个转录组样本。最后提取 *NLRP* 基因的表达数据,利用 R 软件对其进行分析、聚类 and 可视化。

2 结果分析

2.1 牦牛 NLRP 基因家族成员的鉴定及蛋白理化特性分析

从牦牛基因组中共鉴定出 8 个 *NLRP* 基因家族成员(见表 2),按照注释文件中的信息命名。蛋白基本理化性质分析结果显示:牦牛 *NLRP* 蛋白分子量介于 113 117.63–124 957.17 Da 之间,蛋白长度介于 996–1 098 之间,等电点介于 5.52–8.27 之间,消

光系数介于 70 340–130 560 之间,不稳定系数介于 44.74–57.41 之间,脂肪族指数介于 88.57–98.02 之间。8 个 *NLRP* 蛋白均为不稳定、无信号肽、无跨膜结构域的疏水蛋白。亚细胞定位结果显示,不同的 *NLRP* 蛋白定位有所差异,*NLRP3*、*NLRP6* 和 *NLRP12* 均定位于细胞质膜,*NLRP5* 和 *NLRP14* 定位于细胞质基质或者细胞核,*NLRP8* 定位于细胞核中,*NLRP2* 和 *NLRP9* 定位于细胞质基质中。所有的 *NLRP* 蛋白都有丝氨酸、苏氨酸、酪氨酸磷酸化位点(见表 3),磷酸化位点数依次为丝氨酸>苏氨酸>酪氨酸,每个家族成员的磷酸化位点个数介于 86~102 个之间。其中,*NLRP14* 含有的磷酸化位点最多(102 个),*NLRP2* 最少(83 个)。

表 2 牦牛 NLRP 蛋白基本信息和编码蛋白理化性质分析结果

Table 2 Basic information of NLRP proteins and physiological and chemical properties of their encoded proteins										
基因名字	等电点	分子量 /Da	长度 /aa	消光系数	不稳定系数	脂肪族指数	亲水性	信号肽	跨膜结构域	亚细胞定位
NLRP14	5.52	124 957.17	1 094	127 455	51.26	95.61	0.122	无	0	细胞质基质/细胞核
NLRP3	6.50	118 309.48	1 032	127 095	46.89	93.72	0.233	无	0	细胞质膜
NLRP5	6.58	121 122.96	1 098	114 130	44.74	96.07	0.080	无	0	细胞质基质/细胞核
NLRP9	5.58	114 151.39	996	130 560	46.75	93.29	0.070	无	0	细胞质基质
NLRP12	7.78	117 992.56	1 054	109 300	47.23	98.02	0.131	无	0	细胞质膜
NLRP6	6.62	113 117.63	1 025	70 340	55.22	88.57	0.289	无	0	细胞质膜
NLRP8	8.27	120 358.60	1 066	113 755	57.41	96.94	0.103	无	0	细胞核
NLRP2	6.16	114 237.22	1 012	105 080	52.32	94.74	0.159	无	0	细胞质基质

表 3 牦牛 NLRP 蛋白的磷酸化位点预测

Table 3 Phosphorylation site prediction of NLRPs of yak			
蛋白名称	丝氨酸	苏氨酸	酪氨酸
NLRP14	67	29	6
NLRP2	54	26	3
NLRP3	55	24	9
NLRP5	54	30	3
NLRP6	51	35	4
NLRP8	63	24	4
NLRP9	57	26	4
NLRP12	54	29	7

2.2 NLRP 蛋白氨基酸序列比对和系统进化树构建

通过 ClustalW 对 8 个 *NLRP* 蛋白进行多重序列比对,结果显示 *NLRP* 蛋白的氨基酸序列相似度低(见图 1)。两两比对分数在 17.1%到 38.4%不等。其中 *NLRP3* 和 *NLRP12* 的比对分数最高(38.4%),*NLRP2* 和 *NLRP5* 的比对分数最低(17.1%)。序列中部分位点 Y、R、E、V、F 在所有蛋白序列中完全一致。

利用 MEGA X 对 8 个牦牛、8 个黄牛、14 个人、

10 个狗、12 个大鼠和 19 个小鼠 *NLRP* 蛋白氨基酸序列进行多重序列比对和构建系统进化树(见图 2)。从进化树的结构可以看出不同家族成员在进化上已经发生了分离,同一家族成员氨基酸序列相似度高,在 6 个物种中都聚在一起。免疫相关和生殖相关的 *NLRP* 蛋白已分开(Group I 和 Group II),生殖相关的 *NLRP* 蛋白之间相似度要高于免疫相关蛋白。通过与已验证的人、小鼠等物种的 *NLRP* 蛋白序列比对,进一步确认注释文件的准确性。

2.3 基因结构和蛋白保守结构域分析

基因结构分析结果显示(见图 3),*NLRP* 基因内含子数量介于 8–12 个之间,外显子数目介于 9–13 个之间。5′-UTR 和 3′-UTR 的长短差异较大,3′-UTR 长度要大于 5′-UTR。*NLRP8* 和 *NLRP2* 没有 5′-UTR 和 3′-UTR 结构,*NLRP9*、*NLRP5*、*NLRP12* 和 *NLRP6* 没有 5′-UTR 结构。从基因结构预测的结果上看,免疫相关蛋白和生殖相关蛋白没有显著区别。

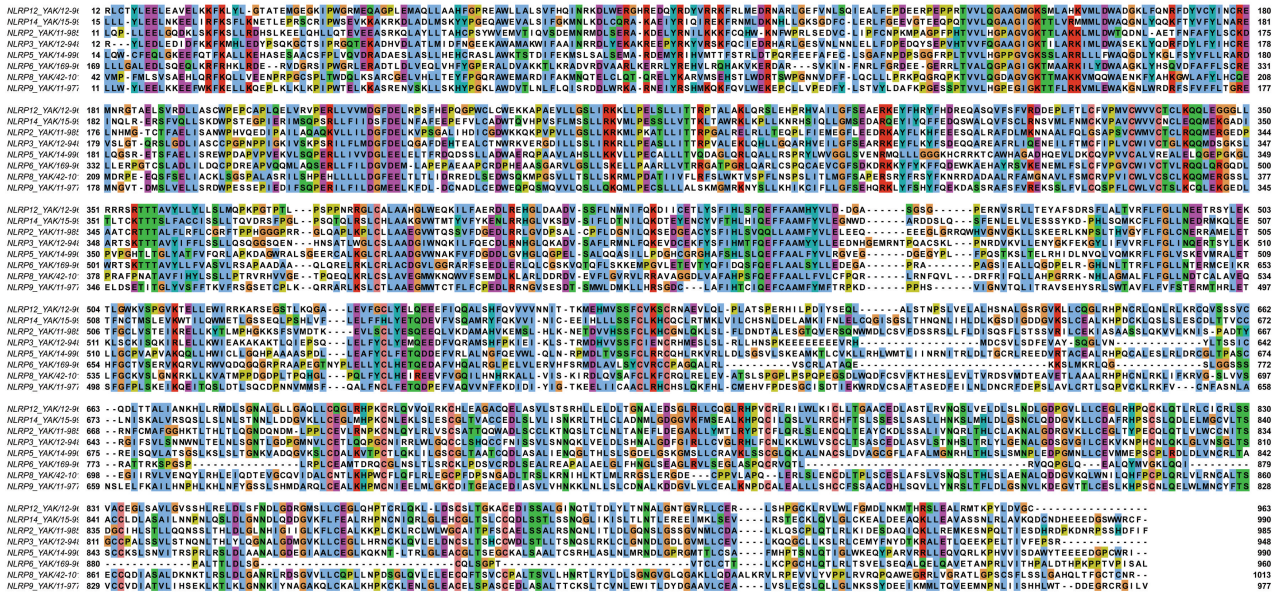


图1 牦牛 *NLRP* 蛋白氨基酸序列比对情况
Fig.1 Amino acid sequence alignment of *NLRP* proteins in yak

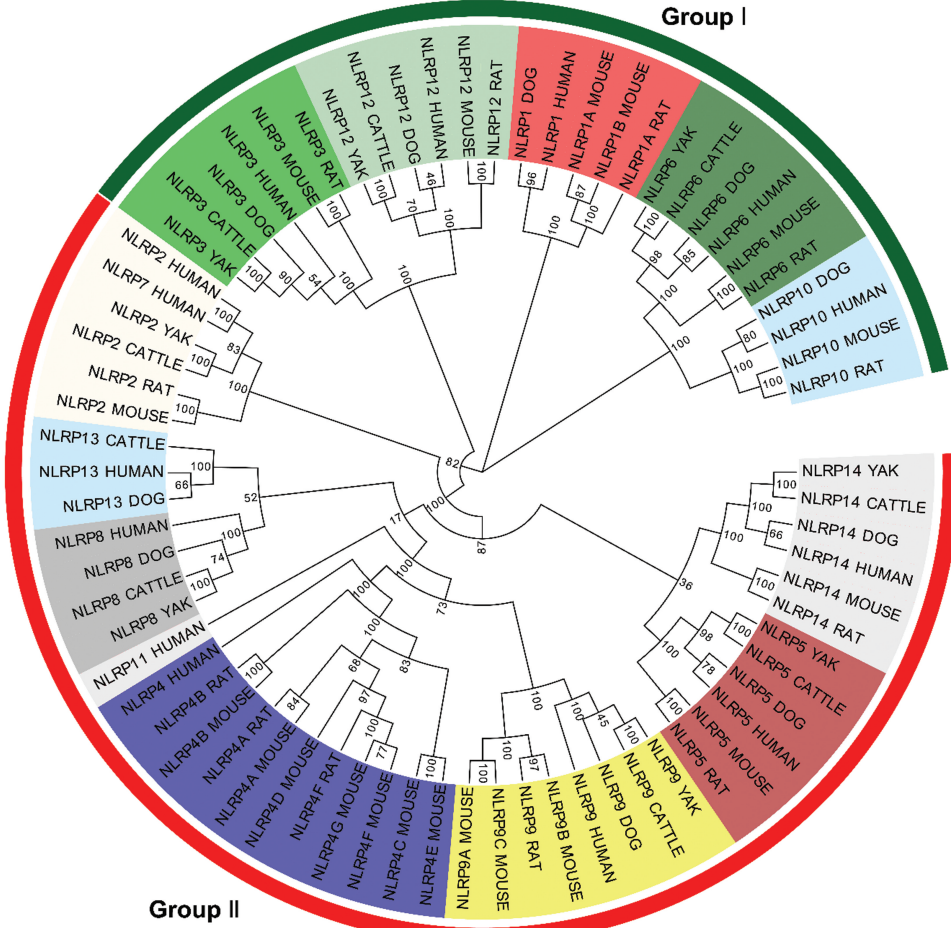


图2 牦牛 *NLRP* 蛋白的系统进化分析
Fig.2 Phylogenetic analysis of *NLRP* proteins in yak

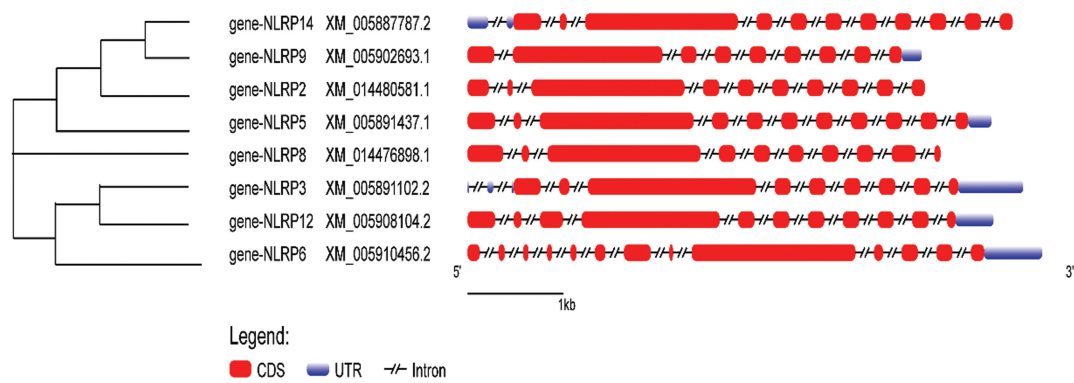


图 3 牦牛 *NLRP* 家族成员基因结构

Fig.3 Gene structure of *NLRP* family members in yak

MEME 预测结果显示牦牛 NLRP 蛋白存在 10 种保守基序 (motif 1-10) (见图 4), 长度分别为 34、49、49、48、21、48、50、41、39、29 bp。NLRP5、NLRP12 和 NLRP14 没有 motif 4。NLRP14 的 motif 数量最多 (16 个), NLRP6 数量最少 (10 个)。10 个 motif 中, C、D、L、N、G、D 位点的一致性较好 (见图 5)。牦牛 NLRP 蛋白与其他物种相同, 都具有该家族代表性保守结构域: NACHT、PYRIN 和 LRR (见图 6)。

2.4 牦牛 *NLRP* 家族蛋白结构和蛋白互作网络分析

对牦牛 NLRP 蛋白的氨基酸序列进行二级结构预测, 结果 (见表 4), 牦牛 NLRP 蛋白二级结构主要元件为 α -螺旋, 其次依次为无规则卷曲、延伸链和 β -转角。为了解蛋白之间的功能联系和在生物系统中的工作原理, 在蛋白互作网络数据库中查询了 8 个 NLRP 蛋白的互作网络 (见图 7)。NLRP2 和 NLRP14 的调控网络相似, 都与 PYCARD、1L1B、

1L18、NEK7 和 MAVS 互作。NLRP3、NLRP6 和 NLRP12 也与 PYCARD 互作。NLRP5、NLRP8 和 NLRP9 具有独立的调控网络。NLRP5 与 BMP15、ZAR1、RAD16 等蛋白存在相互作用。NLRP8 和 MEFV、TAF8、TAF3 等蛋白互作。NLRP9 与 TBK1、UBC、TTC12 等蛋白互作。

表 4 NLRP 蛋白二级结构

Table 4 Secondary structure of NLRP proteins in yak

蛋白名字	α -螺旋	β -转角	延伸链	无规则卷曲
NLRP2	58.60	3.46	7.21	30.73
NLRP3	58.43	3.59	6.30	31.69
NLRP5	56.56	3.10	6.01	34.34
NLRP6	54.63	4.29	7.51	33.56
NLRP8	58.07	3.85	5.82	32.27
NLRP9	60.74	3.31	6.83	29.12
NLRP12	59.49	3.89	5.79	30.83
NLRP14	58.68	4.11	6.76	30.44

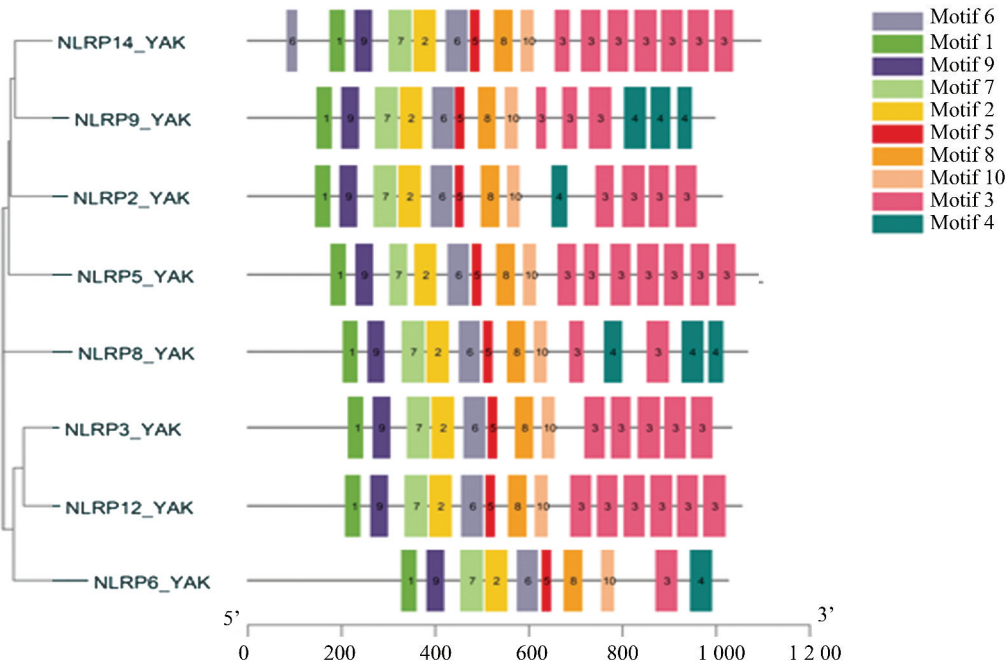


图 4 牦牛 *NLRP* 蛋白保守基序分布

Fig.4 Conserved motifs of *NLRP* proteins in yak



图 5 牦牛 *NLRP* 蛋白典型保守基序序列图

Fig.5 Typical conserved motif sequences of *NLRP* proteins in yak

注:每个字母代表一种氨基酸,字母的大小代表该位点氨基酸的个数,与氨基酸一致性有关,字母越大个数越少,说明该位点的一致性越好,保守性也越好,反之说明该位点的一致性较差,保守性低。

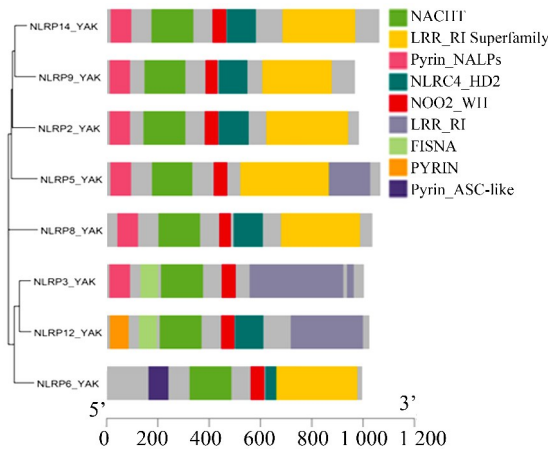


图 6 牦牛 *NLRP* 蛋白保守结构域分布

2.5 *NLRP* 基因在牦牛不同组织中的表达

为了研究 *NLRP* 基因家族在牦牛不同组织中的表达情况,本研究从公共数据库 NCBI 下载了牦牛转录组测序数据^[30],获得了 8 个牦牛 *NLRP* 基因在 6 种组织中的表达模式。结果显示,牦牛 *NLRP* 基因在 6 种组织中的表达较为微弱或者不表达(见图 8)。*NLRP2* 只在肺脏和脾脏中检测到表达。*NLRP3*、*NLRP5* 在 6 种组织中均有表达,二者均在脾脏中的表达量最高。*NLRP6* 在肾脏中的表达量要高于脾脏和肺脏。*NLRP8* 在脾脏和肝脏中检测到表达。*NLRP9* 仅在脾脏中检测到。*NLRP12* 在心脏中不表达,在其他组织中表达。*NLRP14* 只在肺和肾脏中检测到。

Fig.6 Conserved domains of *NLRP* proteins in yak

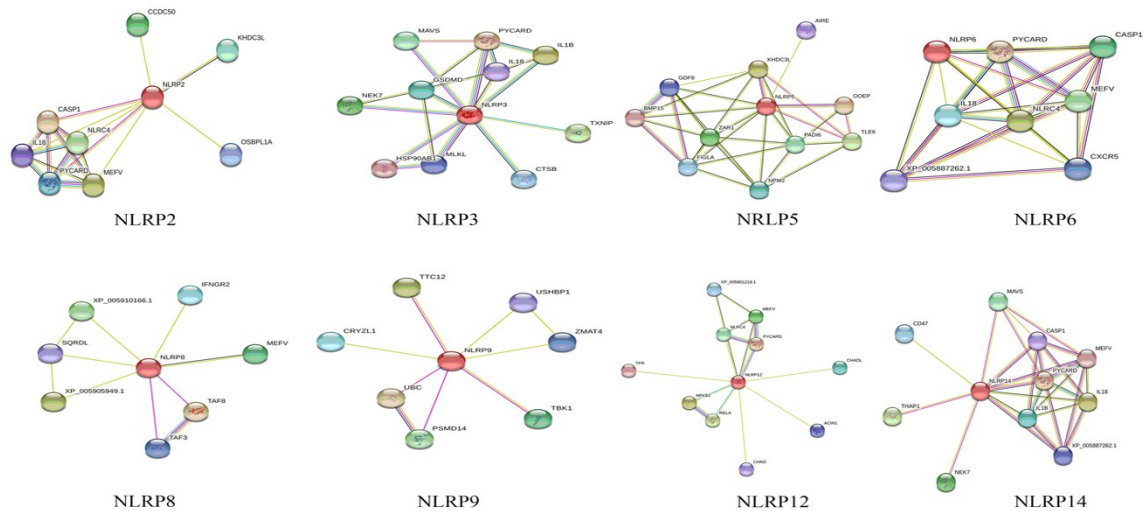


图 7 牦牛 *NLRP* 家族成员蛋白互作网络

Fig.7 Protein-protein interaction network of *NLRP* family members in yak

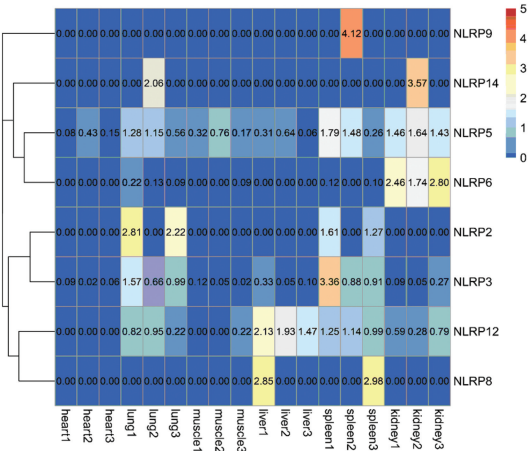


图 8 *NLRP* 基因在牦牛不同组织的表达模式
Fig.8 Expression patterns of *NLRP* genes in different tissues of yak

注:基因表达热图中的每个小方块代表每个基因,基因的表达量由小方块的颜色深度来表示。行表示每个基因在不同样本中的表达,列表示每个样本中所有基因的表达。表达量值按照行进行标准化,并基于平均连接算法进行聚类。

3 讨 论

随着基因组学的发展,动物基因组的测定以及注释信息的完善,使得我们能够通过生物信息学的方法分析基因的相关信息,从序列的角度直接解析基因的各种信息,对目标基因可能存在的生物学功能进行预测^[33]。本研究共鉴定出牦牛 *NLRP* 基因家族的 8 个成员,其中免疫相关的有 *NLRP3*、*NLRP6* 和 *NLRP12*,生殖相关的有 *NLRP2*、*NLRP5*、*NLRP8*、*NLRP9* 和 *NLRP14*。

NLRP 基因在很多物种中被鉴定出来,但是在物种之间存在很大差异。例如,在灵长类上已鉴定出 14 个 *NLRP* 基因^[34],在黄牛上已鉴定出 8 个 *NLRP* 基因^[35],在小鼠上已鉴定出 20 个 *NLRP* 基因^[36],在狗上已鉴定出 11 个 *NLRP* 基因^[35]。小鼠基因组上缺失了 *NLRP7*、*NLRP8*、*NLRP11* 和 *NLRP13*,但在 *NLRP1*、*NLRP4* 和 *NLRP9* 上发生了复制,其发生复制和平行扩张的都是与生殖相关的基因。与小鼠不同的是,牦牛 *NLRP* 家族中免疫相关基因缺失 *NLRP1* 和 *NLRP10*,生殖相关基因缺失 *NLRP4*、*NLRP7*、*NLRP11* 和 *NLRP13*,同时也没有鉴定出平行扩张的基因。物种间的比对结果显示, *NLRP* 家族部分基因在进化过程中是保守的,如 *NLRP3*(牦牛与人有 84% 相同,牦牛与小鼠有 80% 相同,人与小鼠有 82% 相同,牦牛与黄牛有 99% 相同);而部分生殖相关基因 *NLRP5*(牦牛与人有 71% 相同,牦牛与小鼠有 63% 相同,人与小鼠有 50% 相

同,牦牛与黄牛有 99% 相同)与 *NLRP14*(牦牛与人有 83% 相同,牦牛与小鼠有 70% 相同,人与小鼠有 62% 相同,牦牛与黄牛有 99% 相同)序列在物种间的差异较大,这可能是生殖相关基因进化更为快速导致的。牦牛 *NLRP* 蛋白除了都含有 *NLRP* 蛋白家族独有的 *PYRIN*、*NACHT* 和 *LRR* 保守结构域,还预测出 *NOD2_WH* 和 *FISNA* 结构域(见图 6)。*NOD2_WH* 这种有翼的螺旋结构在 *NOD2* 蛋白中被发现^[37],但它的分子功能尚不清楚。*FISNA* 结构域在鱼类和其他脊椎动物中存在,该域经常与 *NACHT* 域相关联^[38],具体功能尚无报道。

一些蛋白在基因表达调控、生长发育、信号因子传递及细胞周期调控等生命过程中存在相互作用,这种不同蛋白之间的相互作用构成了蛋白互作网络。结果显示,牦牛 *NLRP2*、*NLRP3*、*NLRP6*、*NLRP12*、*NLRP14* 与均与参与炎症反应的蛋白 *PYCARD* 存在相互作用。*PYCARD* 是凋亡相关斑点样蛋白,能够被 *NLRP3* 募集,并且与天冬氨酸蛋白水解酶-1(*caspase-1*)形成多蛋白聚合的炎症小体,它的作用是激活 *caspase-1*,通过促进炎症因子 *IL-1 β* 和 *IL-18* 的成熟和分泌,从而促进细胞凋亡,调节固有免疫应答^[39]。*NLRP2*、*NLRP6*、*NLRP14* 均与 *MLKL* 互作。*MLKL* 是混合谱系激酶结构域样蛋白,相关研究表明在没有凋亡蛋白抑制剂(*IAPs*)和半胱氨酸的天冬氨酸蛋白水解酶-8(*caspase-8*)情况下,*MLKL* 对于 *TLR* 诱导的 *NLRP3* 激活有重要作用^[40]。*NLRP5* 与 *TLE6*、*BMP15*、*FIGLA*、*NPM2*、*GDF9* 等参与生殖过程的蛋白互作,*BMP15* 蛋白由卵母细胞衍生生长因子基因 *BMP15* 编码,在绵羊中发现该基因的自然突变可以以剂量敏感的方式导致排卵率和不育表型的增加^[41]。另外,绵羊 *GDF9* 基因突变以类似于 *BMP15* 失活突变的方式导致排卵率增加和不孕,表明 *GDF9* 对绵羊正常卵泡发育至关重要^[42]。

转录组分析结果显示牦牛 *NLRP* 基因家族成员在 6 个组织中的表达存在差异。免疫相关基因 *NLRP3* 在脾脏中表达较高,可能与脾脏作为最大的免疫器官有关。*NLRP5* 在 6 个组织中表达,在脾脏中的表达量要高于其他组织(见图 8)。相关研究表明 *NLRP3* 在猪上的同源物在早期胚胎发育中起作用,受精卵中 *NLRP5* 的表达被抑制会导致胚胎的早期发育受到抑制^[43]。在合子基因组激活前,*NLRP5* 的转录物在母猪植入前胚胎中显著下调直到盒子基因组激活后检测不到任何表达,表明 *NLRP5* 基因是母猪早期胚胎发生所必需的,提示该基因可能在合子基因组激活中发挥重要作用。蛋白互作网络揭示

NLRP5 也与生殖细胞标记蛋白 *GDF9*、*BMP15*、*ZAR1* 等存在相互作用(见图 7)。*NLRP8* 在牦牛脾脏和肝脏中检测到微弱表达,该基因在啮齿类动物基因组中缺失,在其他物种中其研究尚无报道。*NLRP14* 只在肺脏和肾脏中检测到,相关研究表明其在卵子发生和早期胚胎发育中发挥重要作用,该基因的沉默会导致胚胎发育阻滞^[45]。

从蛋白互作网络以及基因表达谱的结果来看,部分 *NLRP* 基因及蛋白的功能很难完全划分为免疫和生殖两部分,它们的功能是复杂的。从文献报道的结果来看,*NLRP* 基因既可以在免疫方面发挥作用,又能够参与动物的生殖过程。*NLRP7* 是与动物生殖相关的基因,但是也有研究发现该基因在先天免疫系统也起作用,编码的蛋白同样能够招募 ASC 促进 caspase-1 激活和促进 IL-1 调节因子和 IL-18 的成熟^[46]。*NLRP9* 在此前研究中被认为是生殖相关基因,在 小鼠卵巢中表达,参与合子基因组的激活^[47]。此外,玉香等^[48] 在对 *NLRP9* 在仔猪组织中的表达、定位及在猪小肠上皮细胞(IPEC-J2)细胞中的功能研究中发现 *NLRP9* 参与了肠粘膜免疫过程,其表达主要定位于小肠上皮、睾丸和卵巢中。

4 结 论

通过同源比对和保守结构域搜索,在牦牛基因组中鉴定出牦牛 8 个 *NLRP* 家族成员,并且通过系统进化分析确定牦牛 *NLRP* 基因家族的分类和进化关系,并且通过生物信息学对牦牛 *NLRP* 基因家族的基因结构、理化性质、蛋白保守结构域和二级结构进行了预测分析。通过 STRING 数据库查询蛋白互作关系,进一步了解牦牛 *NLRP* 蛋白与其他蛋白之间的功能联系,为确定基因功能提供参考。最后利用转录组测序数据,初步展示了 *NLRP* 基因在牦牛不同组织中的表达情况,上述分析将为后续牦牛 *NLRP* 基因功能研究提供理论基础。

参考文献(References)

[1] 吴比. 小鼠 *Nlrp14* 基因敲除模型的建立[D]. 哈尔滨: 东北农业大学,2013.
WU Bi. Establishing *Nlrp14* knock-out mouse model[D]. Harbin: Northeast Agricultural University,2013.
[2] CAO J Y, ZHOU L T, LI Z L, et al. Dopamine D1 receptor agonist A68930 attenuates acute kidney injury by inhibiting *NLRP3* inflammasome activation[J]. Journal of Pharmacological Sciences,2020,143(3): 226–233. DOI: 10.1016/j.

jphs.2020.04.005.
[3] HARRIS P A, DURAISWAMI C, FISHER D T, et al. High throughput screening identifies ATP-competitive inhibitors of the *NLRP1* inflammasome [J]. Bioorganic & Medicinal Chemistry Letters, 2015, 25 (14): 2739–2743. DOI: 10.1016/j.bmcl.2015.05.032.
[4] HA H J, KIM D S, HAHN Y. A 2.7-kb deletion in the human *NLRP10* gene exon 2 occurred after the human-chimpanzee divergence [J]. Biochemical Genetics, 2009, 47 (9/10): 665–670. DOI: 10.1007/s10528-009-9262-2.
[5] KUCHMIY A A, D’HONT J, HOCHEPIED T, et al. *NLRP2* controls age-associated maternal fertility [J]. Journal of Experimental Medicine, 2016, 213 (13): 2851–2860. DOI: 10.1084/jem.20160900.
[6] 陈静, 刘芳, 霍建超, 等. *NLRP* 家族中生殖相关基因的研究进展 [J]. 畜牧与兽医, 2017, 49 (6): 200–203. DOI: CNKI;SUN;XMY.S.0.2017-06-04.
CHEN Jing, LIU Fang, HUO Jianchao, et al. Research progress of *NLRP* reproduction related genes [J]. Animal Husbandry & Veterinary Medicine, 2017, 49 (6): 200–203. DOI: CNKI;SUN;XMY.S.0.2017-06-04.
[7] 彭辉. *Nlrp* 基因家族的表达及 *Nlrp2* 和 *Nlrp4g* 在 小鼠早期胚胎发育过程中的功能 [D]. 杨凌: 西北农林科技大学, 2012.
PENG Hui. Expression analysis of *Nlrp* gene family and function of *Nlrp2* and *Nlrp4g* in early embryonic development in the mouse [D]. Yangling: Northwest A & F University, 2012.
[8] 林秀娇. *Nlrp9b* 在 小鼠早期胚胎发育中的表达、定位及功能 [D]. 福州: 福建农林大学, 2015.
LIN Xiujiao. The expression, localization and function of *Nlrp8b* gene in the early mouse embryo [D]. Fuzhou: Fujian Agriculture and Forestry University, 2015.
[9] SLIM R, WALLACE E P. *NLRP7* and the genetics of hydatidiform moles: Recent advances and new challenges [J]. Frontiers in Immunology, 2013, 4: 242. DOI: 10.3389/fimmu.2013.00242.
[10] 李广栋, 崔炜, 田秀芝, 等. *NLRP7* 基因在母源印记中的研究进展 [J]. 中国畜牧兽医, 2018, 45 (4): 1002–1008. DOI: 10.16431/j.cnki.1671-7236.2018.04.021.
LI Guangdong, CUI Wei, TIAN Xiuzhi, et al. Research progress on *NLRP7* gene in maternal imprinting [J]. China Animal Husbandry & Veterinary Medicine, 2018, 45 (4): 1002–1008. DOI: 10.16431/j.cnki.1671-7236.2018.04.021.
[11] 胡泉军. 牦牛基因组数据库建设 [D]. 兰州: 兰州大学, 2014.
HU Qunjun. Database construction of yak genome [D]. Lanzhou: Lanzhou University, 2014.
[12] 石仙, 熊显荣, 兰道亮, 等. 牦牛 *CCND2* 基因的克隆及其在不同发情时期卵巢中的表达 [J]. 中国农业科学,

- 2017,50(13):2604–2613. DOI: 10.3864/j.issn.0578–1752.2017.13.018.
- SHI Xian, XIONG Xianrong, LAN Daoliang, et al. Cloning and expression analysis of CCND2 gene in yak ovaries during different periods of estrus[J]. *Scientia Agricultura Sinica*,2017,50(13):2604–2613. DOI: 10.3864/j.issn.0578–1752.2017.13.018.
- [13] 闫振龙. 牦牛睾丸促性腺激素受体和性激素受体增龄性变化与免疫调控研究[D]. 兰州: 甘肃农业大学, 2014.
- YAN Zhenlong. The immine regulation study of gonadotropin receptor and hormone receptor in yak testis (*Bos grunniens*) at different ages [D]. Lanzhou: Gansu Agricultural University, 2014.
- [14] 陈虹. 牦牛卵巢 Label-free 蛋白组学分析及相关蛋白在生殖系统的表达[D]. 兰州: 甘肃农业大学, 2018:2604–2613.
- CHEN Hong. Analysis of label-free quantification proteomics in yak's ovary and expression of related proteins in the reproductive system[D]. Lanzhou: Gansu Agricultural University, 2018.
- [15] 杨芳. 基于 RNA-Seq 对牦牛和犏牛睾丸转录组的比较分析[D]. 成都: 四川农业大学, 2017.
- YANG Fang. Comparative analysis of testis transcriptome for yak and cattleyak based on RNA-seq [D]. Chengdu: Sichuan Agricultural University, 2017.
- [16] CONSORTIUM U P. UniProt: a hub for protein information [J]. *Nucleic Acids Research*, 2015(D1): 204–212. DOI: 10.1093/nar/gku989.
- [17] CAMACHO C, COULOURIS G, AVAGYAN V, et al. BLAST+: architecture and applications[J]. *BMC Bioinformatics*, 2009, 10(1): 421. DOI: 10.1186/1471–2105–10–421.
- [18] LARKIN M A, BLACKSHIELDS G, BROWN N P, et al. ClustalW and ClustalX version 2[J]. *Bioinformatics*, 2007, 23(21): 2947–2948. DOI: 10.1093/bioinformatics/btm404.
- [19] SUDHIR K, GLEN S, LI M, et al. MEGA X: Molecular evolutionary genetics analysis across computing platforms [J]. *Molecular Biology & Evolution*, 2018(6): 6. DOI: 10.1093/molbev/msy096.
- [20] GASTEIGER E, GATTIKER A, HOOGLAND C, et al. ExPASy: the proteomics server for in-depth protein knowledge and analysis [J]. *Nucleic Acids Research*, 2003, 31(13): 3784–3788. DOI: 10.1093/nar/gkg563.
- [21] SignalP 5.0 improves signal peptide predictions using deep neural networks[J]. *Nature Biotechnology*, 2019, 37: 420–423. DOI: 10.1038/s41587–019–0036–z.
- [22] PAUL H, KEUN-JOON P, TAKESHI O, et al. WoLF PSORT: protein localization predictor [J]. *Nucleic Acids Research*, 2007, 35(Web Server issue): 585–587. DOI: 10.1093/nar/gkm259.
- [23] HU B, JIN J, GUO A Y, et al. GSDS 2.0: an upgraded gene feature visualization server [J]. *Bioinformatics*, 2014, 31(8): 1296. DOI: 10.1093/bioinformatics/btu817.
- [24] MARCHLER-BAUER A, BRYANT S H. CD-Search: protein domain annotations on the fly [J]. *Nucleic Acids Research*, 2004, 32(S2): W327–W331. DOI: 10.1093/nar/gkh454.
- [25] BATEMAN A, COIN L, DURBIN R, et al. The Pfam protein families database [J]. *Nucleic Acids Research*, 2004, 32(S1): D138–D141. DOI: 10.1093/nar/gkh121.
- [26] BAILEY T L, MIKAEL B, BUSKE F A, et al. MEME Suite: tools for motif discovery and searching [J]. *Nucleic Acids Research*, 2009, 37(Web Server issue): W202–W208. DOI: 10.1093/nar/gkp335.
- [27] CHEN C, CHEN H, ZHANG Y, et al. TBtools – an integrative toolkit developed for interactive analyses of big biological data [J]. *Molecular Plant*, 2020, 13(8): 1194–1202. DOI: 10.1016/j.molp.2020.06.009.
- [28] Geourjon C, Deleage G. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments [J]. *Computer Applications in the Biosciences*, 1995, 11(6): 681–684. DOI: 10.1093/bioinformatics/11.6.681.
- [29] SZKLARCZYK D, MORRIS J H, COOK H, et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible [J]. *Nucleic acids research*, 2016: D362–D368. DOI: 10.1093/nar/gkw937.
- [30] TANG Q, GU Y, ZHOU X, et al. Comparative transcriptomics of five high-altitude vertebrates and their low-altitude relatives [J]. *GigaScience*, 6(12): 1–9. DOI: 10.1093/gigascience/gix105.
- [31] CHEN S, ZHOU Y, CHEN Y, et al. fastp: an ultra-fast all-in-one FASTQ preprocessor [J]. *Bioinformatics*, 2018, 34(17): i884–i890. DOI: 10.1093/bioinformatics/bty560.
- [32] WEN G. A simple process of RNA-sequence analyses by Hisat2, Htseq and DESeq2 [C]// *Proceedings of the 2017 International Conference on Biomedical Engineering and Bioinformatics*. New York: ACM, 2017. DOI: 10.1145/3143344.3143354.
- [33] 罗荣松, 冯郁, 宋明民, 等. 小鼠 Nlrp 基因家族生物信息学分析 [J]. *基因组学与应用生物学*, 2018, 37(3): 1168–1176. DOI: 10.13417/j.gab.037.001168.
- LUO Rongsong, FENG Yu, SONG Mingmin, et al. Bioinformatics analysis of Nlrp gene family in rat [J]. *Genomics and Applied Biology*, 2018, 37(3): 1168–1176. DOI: 10.13417/j.gab.037.001168.
- [34] TSCHOPP J, MARTINON F, BURNS K. NALPs: A novel protein family involved in inflammation [J]. *Nature*

Reviews Molecular Cell Biology, 2003, 4 (2) : 95 – 104. DOI: 10.1038/nrm1019.

[35] 彭辉, 张涌. NLRP 蛋白家族的研究进展[J]. 中国畜牧杂志, 2014, 50 (3) : 90–93. DOI: 10.3969/j.issn.0258–7033.2014.03.022.

PENG Hui, ZHANG Yong. Research progress of NLRP protein family [J]. China Academic Journal, 2014, 50 (3) : 90–93. DOI: 10.3969/j.issn.0258–7033.2014.03.022.

[36] 彭辉, 张文昌, 肖天放, 等. Nlrp 基因家族部分免疫相关基因在小鼠上的表达[J]. 福建农林大学学报(自然科学版), 2013, 42 (4) : 398–402. DOI: 10.3969/j.issn.1671–5470.2013.04.013.

PENG Hui, ZHANG Wenchang, XIAO Tianfang, et al. Expression of immunization-related Nlrp genes in mouse[J]. Journal of Fujian Agriculture and Forestry University (Natural Science Edition), 2013, 42 (4) : 398–402. DOI: 10.3969/j.issn.1671–5470.2013.04.013.

[37] MAEKAWA S, OHTO U, SHIBATA T, et al. Crystal structure of NOD2 and its implications in human disease[J]. Nature Communications, 2016, 7: 11813. DOI: 10.1038/ncomms11813.

[38] STEIN C, CACCAMO M, LAIRD G, et al. Conservation and divergence of gene families encoding components of innate immune response systems in zebrafish[J]. Genome Biology, 2007, 8 (11) : R251. DOI: 10.1186/gb–2007–8–11–r251.

[39] 李艳蝶, 卢美萍. NLRP3 炎症小体与儿童自身炎症性疾病研究进展[J]. 浙江大学学报(医学版), 2017, 46 (4) : 449–453. DOI: 10.3785/j.issn.1008–9292.2017.08.17.

LI Yandie, LU Meiping. Progress on the study of NLRP3 inflammasome in autoinflammatory disease of children[J]. Journal of Zhejiang University (Medical Sciences), 2017, 46 (4) : 449–453. DOI: 10.3785/j.issn.1008–9292.2017.08.17.

[40] LAWLOR K E, KHAN N, MILDENHALL A, et al. RIPK3 promotes cell death and NLRP3 inflammasome activation in the absence of MLKL[J]. Nature Communications, 2015, 6 (1) : 1–19. DOI: 10.1038/ncomms7282.

[41] GALLOWAY S M, MCNATTY K P, CAMBRIDGE L M, et al. Mutations in an oocyte-derived growth factor gene (BMP15) cause increased ovulation rate and infertility in a dosage-sensitive manner[J]. Nature Genetics, 2000, 25 (3) : 279–283. DOI: 10.1038/77033.

[42] HANRAHAN J P. Mutations in the genes for oocyte-derived growth factors GDF9 and BMP15 are associated with both increased ovulation rate and sterility in Cambridge and Belclare sheep (Ovis aries) [J]. Biology of Reproduction, 2004, 70 (4) : 900–909. DOI: 10.1095/biolreprod.103.023093.

[43] PENG Hui, LIU Fang, LI Wenhao, et al. Knockdown of NLRP5 arrests early embryogenesis in sows [J]. Animal Reproduction Science, 2015, 163: 151–156. DOI: 10.1016/j.anireprosci.2015.11.004.

[44] MICHIHARU H, KIRKMAN N J, MAYO K E, et al. The mouse germ-cell-specific leucine-rich repeat protein NALP14: a member of the NACHT nucleoside triphosphatase family[J]. Biology of Reproduction, 2005 (4) : 879. DOI: 10.1095/biolreprod.104.033753.

[45] XIN T, PASCAL G, MONGET P. Evolution and functional divergence of NLRP genes in mammalian reproductive systems[J]. BMC Evolutionary Biology, 2009, 9 (1) : 202. DOI: 10.1186/1471–2148–9–202.

[46] KHARE S, DORFLEUTNER A, BRYAN N B, et al. An NLRP7-containing inflammasome mediates recognition of microbial lipopeptides in human macrophages[J]. Immunity, 2012, 36 (3) : 464–476. DOI: 10.1016/j.immuni.2012.02.001.

[47] 刘芳, 江宁, 肖琼, 等. Nlrp9b 在小鼠不同组织中的表达和定位[J]. 中国兽医杂志, 2015, 51 (12) : 30–32+50. DOI: CNKI;SUN:ZSYZ.0.2015–12–009.

LIU Fang, JIANG Ning, XIAO Qiong, et al. Expression and localization of NLRP9b gene in mouse different tissues [J]. Chinese Journal of Veterinary Medicine, 2015, 51 (12) : 30–32+50. DOI: CNKI;SUN:ZSYZ.0.2015–12–009.

[48] 玉香. NLRP9 在仔猪组织中的表达、定位及在 IPEC-J2 细胞中的功能研究[D]. 福州: 福建农林大学, 2018.

YU Xiang. Expression and localization of *NLRP9* in piglet tissues and its function in IPEC–K2 cells [D]. Fuzhou: Fujian Agriculture and Forestry University, 2018.