

文章编号:1000-8551(2020)08-1690-08

# 绵羊 *GlyCAM-1* 基因克隆、核苷酸变异及其组织表达分析

郝志云 王继卿\* 胡江 刘秀 李少斌 罗玉柱\*

(甘肃农业大学动物科学技术学院/甘肃省草食动物生物技术重点实验室/甘肃省牛羊基因改良工程实验室,甘肃 兰州 730070)

**摘要:**糖基化依赖细胞粘附分子(*GlyCAM-1*)是粘液素(mucin)糖蛋白家族的成员,是乳腺细胞合成的一种乳脂球膜蛋白(MFGMPs)的重要组成部分。*GlyCAM-1*在乳腺发育和泌乳中发挥着重要作用,为探究该基因的序列特征、核苷酸序列变异和表达特征,以高泌乳量的小尾寒羊(泌乳高峰期和空怀期)和低泌乳量的甘肃高山细毛羊(泌乳高峰期)的乳腺组织为研究对象,利用克隆测序、RT-qPCR、生物信息学等方法克隆绵羊*GlyCAM-1*基因的编码区,分析*GlyCAM-1*的理化性质和蛋白质结构,并检测*GlyCAM-1*基因的组织表达特性。结果表明,绵羊的*GlyCAM-1*基因CDS区全长465 bp,编码154个氨基酸。测序结果表明,在该基因CDS区检测到7个SNPs,其中2个为同义突变,5个为错义突变。*GlyCAM-1*二级结构主要以 $\alpha$ -螺旋和无规则卷曲为主,延伸链则散布于整个蛋白质结构中。String分析结果表明,*GlyCAM-1*蛋白与CD34、MadCAM-1都作为L-选择素(L-selectin)的配体发挥作用;RT-qPCR结果表明,*GlyCAM-1*基因表达具有组织特异性、品种特异性和时空特异性,乳腺、心脏、肝脏、肺脏、脾脏、肾脏、卵巢和背最长肌8个组织中,*GlyCAM-1*基因只在乳腺和肺脏组织中表达,在其余6个组织中均不表达,其中乳腺中的表达量最高;在泌乳高峰期的乳腺组织中,*GlyCAM-1*在高泌乳量的小尾寒羊中的表达量显著高于甘肃高山细毛羊( $P < 0.05$ );在小尾寒羊的乳腺组织中,*GlyCAM-1*在泌乳高峰期的表达量极显著高于空怀期( $P < 0.01$ )。本研究结果丰富了绵羊基因组数据,为深入研究*GlyCAM-1*基因的泌乳生物学功能及其机理提供了基础数据。

**关键词:***GlyCAM-1*基因;乳腺;绵羊;克隆;组织表达;SNPs

DOI:10.11869/j.issn.100-8551.2020.08.1690

乳腺是哺乳动物唯一重复经历生长、功能分化和退化的器官,在动物的繁育过程中发挥着至关重要的作用<sup>[1]</sup>。乳蛋白是乳汁的重要组成部分,是衡量乳品质的重要指标,主要由酪蛋白、乳清蛋白和乳脂球膜蛋白(milk fat globule membrane proteins, MFGMPs)等组成。MFGMPs仅占乳蛋白总含量的1%~2%,但具有抑制癌细胞、抗病毒感染、增强机体免疫等重要功能<sup>[2-4]</sup>。糖基化依赖细胞粘附分子(glycosylation-dependent cell adhesion molecule 1, *GlyCAM-1*)属于粘液素(mucin)糖蛋白家族的一员,是MFGMPs的重要

组成部分,具有参与细胞识别、信号转导以及介导白细胞粘附和淋巴细胞归巢等生物学功能,同时在免疫应答和炎症反应等过程中发挥重要作用<sup>[5]</sup>。此外,*GlyCAM-1*作为胚胎和子宫内膜特异受体的配体,参与整合素介导的粘膜级联反应<sup>[6]</sup>。在乳腺中,*GlyCAM-1*作为一种受激素调节的MFGMPs重要组成部分参与乳腺发育及乳汁分泌过程<sup>[7-8]</sup>。Suárez Vega等<sup>[9]</sup>通过对2个乳用绵羊品种(*Ovis aries*) Assaf(年产量400 kg)和Churra(年产量117 kg)在泌乳第10、第50、第120、第150天的乳体细胞(milk somatic cells, MSC)进

收稿日期:2018-12-18 接受日期:2019-02-24

基金项目:国家自然科学基金(31860635),甘肃农业大学“伏羲青年英才培育计划项目(Gaufx-02Y02),甘肃省基础研究创新群体项目(18JR3RA190、17JR5RA137)

作者简介:郝志云,男,主要从事动物遗传育种理论与应用。E-mail: haozy2018@163.com

\* 通讯作者:王继卿,男,副研究员,主要从事现代生物技术与动物育种应用。E-mail: wangjq@gsau.edu.cn;

罗玉柱,男,教授,主要从事现代生物技术与动物育种应用。E-mail: luoyz@gsau.edu.cn。同为通讯作者。

行转录组测序发现, *GlyCAM-1* 在整个泌乳时期均高表达( $x_{FRPM} = 26\ 274$ ),但在各时期表达量差异不显著( $P > 0.05$ )。Paten 等<sup>[10]</sup>在研究罗姆尼母羊时发现, *GlyCAM-1* 在产后 15 d 乳腺组织中的表达量高于妊娠后 135 d,表明在乳汁分泌过程中 *GlyCAM-1* 基因的表达量会上升。由此表明, *GlyCAM-1* 在家畜泌乳和乳腺发育中发挥了重要作用,但是关于该基因的序列特征、核苷酸序列变异和表达特征尚不清楚。

为了系统研究绵羊 *GlyCAM-1* 基因的核苷酸序列特征及其变异、组织表达特征,本研究以高泌乳量的小尾寒羊(泌乳高峰期和空怀期)和低泌乳量的甘肃高山细毛羊(泌乳高峰期)为对象,应用基因克隆、生物信息学、测序、实时荧光定量 PCR (real time-quantitative PCR, RT-qPCR) 等技术,对该基因进行了 CDS 区域克隆、氨基酸序列特征分析,研究了该基因 CDS 区的核苷酸序列变异,检测了乳腺、心脏、肝脏、脾脏等 8 个组织中该基因的 mRNA 表达特征,以期为研究 *GlyCAM-1* 基因的泌乳生物学功能提供基础数据。

## 1 材料与方法

### 1.1 试验材料

试验样品采自甘肃省天祝县松山镇金子河绵羊繁育场,选择处于泌乳高峰期(产后第 3 周)的高泌乳量小尾寒羊和低泌乳量甘肃高山细毛羊各 3 只,另外选择处于空怀期的小尾寒羊 3 只,合计 9 只试验羊。要求所有试验羊均为 3 岁、第 4 胎、在同一饲养条件下饲养管理。屠宰后立即采取乳腺、心脏、肝脏、肺脏、脾脏、肾脏、卵巢和背最长肌各 0.5 g 左右。乳腺组织样用焦碳酸二乙酯(diethyl pyrocarbonate, DEPC)水冲洗

除去残留的乳汁之后,装入 2 mL 冻存管后立即置于液氮中保存,带回实验室后于  $-80^{\circ}\text{C}$  保存备用。

### 1.2 试验方法

按照 Trizol Reagent (Invitrogen, CA, USA) 说明书要求,提取组织中的总 RNA。提取后的 RNA 用 1.5% 琼脂糖凝胶电泳检测其完整性,用 Nanoprop 2000 型超微量分光光度计(塞默飞,美国)检测  $\text{OD}_{260}/\text{OD}_{280}$  及其浓度,质量合格的 RNA 保存于  $-80^{\circ}\text{C}$  冰箱中备用。根据天根反转录 FastKing RT kit(with gDNase)说明书要求,对 RNA 进行反转录,得到的 cDNA 按照其浓度稀释成  $100\ \text{ng}\cdot\mu\text{L}^{-1}$  溶液,  $-20^{\circ}\text{C}$  冰箱保存备用。

### 1.3 *GLYCAL-1* 基因引物设计及克隆

依据 GenBank 中预测的绵羊 *GlyCAM-1* 基因的 mRNA 序列(XM\_012149360.2),利用 Primer 5.0 设计 P1 和 P2 引物,其中 P1 引物用于 *GlyCAM-1* 基因 CDS 区的克隆,P2 引物用于 RT-qPCR 分析,  $\beta$ -actin 基因为内参引物,引物具体信息见表 1。以小尾寒羊乳腺组织获得的 cDNA 为模板,采用  $20\ \mu\text{L}$  反应体系,包括 cDNA  $0.8\ \mu\text{L}$ (约  $100\ \text{ng}\cdot\mu\text{L}^{-1}$ ),  $5\ \text{U}\cdot\mu\text{L}^{-1}$  Taq 预混酶  $10\ \mu\text{L}$ ,  $0.25\ \mu\text{mol}\cdot\text{L}^{-1}$  上、下游引物各  $0.8\ \mu\text{L}$ , ddH<sub>2</sub>O 补足至  $20\ \mu\text{L}$ 。PCR 反应条件:  $94^{\circ}\text{C}$  预变性 5 min;  $94^{\circ}\text{C}$  变性 30 s,  $58^{\circ}\text{C}$  退火 30 s,  $72^{\circ}\text{C}$  延伸 30 s,共 35 个循环;  $72^{\circ}\text{C}$  终延伸 7 min;  $4^{\circ}\text{C}$  保存。PCR 产物用 1.5% 琼脂糖电泳检测。选择目的条带进行切胶回收并连接至克隆载体 pMD19-T,经转化、PCR 鉴定后挑取阳性单克隆穿刺培养,送至上海生工生物科技有限公司进行测序。以泌乳高峰期小尾寒羊(3 个)和甘肃高山细毛羊(3 个)乳腺组织的 cDNA 为模板,扩增 P1 引物,将其扩增片段直接测序鉴定其核苷酸序列变异特征。

表 1 绵羊 *GlyCAM-1* 基因的扩增引物信息

Table 1 Primer information for sheep *GlyCAM-1* gene

引物名称 Primer	引物序列(5'-3') Primer sequence(5'-3')	扩增长度 Product length/bp	退火温度 T <sub>m</sub> /°C
P1	F: ATGAAATTCCTCTGCGTCTTG R: CTACGGCTTCACCACTTCAGAG	465	58
P2	F: GCCTCTGCTCAGTTCATC R: TCACATTTGGCTCTTTGG	106	60
$\beta$ -actin	F: AGCCTTCCTTCCTGGGCATGGA R: GCACAGCACCGTGTGGCGTAGA	113	60

### 1.4 *GLYCAM-1* 基因生物信息学分析

基于克隆获得的绵羊 *GlyCAM-1* 基因序列,利用

DNASar 7.1 软件查找该基因的开放阅读框(open reading frame, ORF),并利用 NCBI 中的 BLAST 进行验

证;用 ExPasy 站点上的 ProtParam 软件分析绵羊 GlyCAM-1 的氨基酸特征;利用 ExPaSy 站点上的 SOPMA 分析其蛋白结构;用 NetPhos 3.1 预测氨基酸序列中潜在的磷酸化位点<sup>[11]</sup>;利用 GO 和 KEGG 分析其参与的生物学功能和代谢通路,利用 String<sup>[12]</sup> 在线分析 GlyCAM-1 的蛋白网络互作关系,用 MEGA 5.0 构建系统进化树。

### 1.5 GlyCAM-1 基因表达量分析

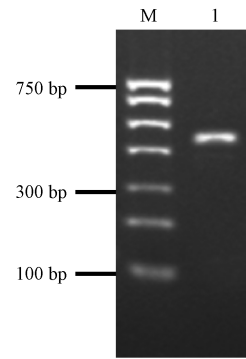
采用相对定量 SYBR Green I (TaKaRa, 大连) 染料法检测 GlyCAM-1 基因的组织表达量。采用 20  $\mu\text{L}$  体系,包括 cDNA 模板 2  $\mu\text{L}$  (<100 ng), 0.25  $\mu\text{mol}\cdot\text{L}^{-1}$  上下游引物各 0.4  $\mu\text{L}$ 、10  $\mu\text{mol}\cdot\text{L}^{-1}$  SYBR qPCR Master Mix 10  $\mu\text{L}$ 、ROX II 0.4  $\mu\text{L}$  和 RNase-free ddH<sub>2</sub>O 6.8  $\mu\text{L}$ 。RT-qPCR 反应条件:95 $^{\circ}\text{C}$  预变性 5 min;95 $^{\circ}\text{C}$  变性 15 s,60 $^{\circ}\text{C}$  退火 34 s,72 $^{\circ}\text{C}$  延伸 30 s,40 个循环。溶解反应:95 $^{\circ}\text{C}$  变性 15 s;60 $^{\circ}\text{C}$  采集荧光信号 60 s。每个反应设 3 个平行对照。待反应结束后依据结果报告,分析其荧光定量溶解曲线,判断是否存在引物二聚体,以及 PCR 产物的准确性,收集数据,计算表达量<sup>[13]</sup>。反应在 Applied Biosystems QuantStudio<sup>®</sup> 6 Flex 仪 (Thermo Fisher, USA) 中进行。

## 2 结果与分析

### 2.1 GlyCAM-1 基因的克隆

以小尾寒羊的乳腺 cDNA 为模板,用 RT-PCR 进行扩增,PCR 产物用 1.5% 琼脂糖凝胶电泳检测。结果表明,目的片段条带 (465 bp) 跟预期结果一致 (图 1)。测序后获得该基因的完整 CDS 区序列,大小为 465 bp,编码 154 个氨基酸。序列已提交至 GenBank,登录号为 MH917952.1。

通过 BLAST 分析发现, MH917952.1 与山羊的 GlyCAM-1 具有较高的序列同源性 (98.7%), 与黄牛 GlyCAM-1 之间的同源性为 88.96%。利用 GenBank



Note: M: DNA Marker; 1: P1.

图 1 绵羊 GlyCAM-1 基因 PCR 电泳结果

Fig.1 Electrophoresis results of GlyCAM-1 gene of sheep

中山羊 (*Capra hircus*)、野牦牛 (*Bos mutus*)、黄牛 (*Bos taurus*)、藏羚羊 (*Pantholops hodgsonii*)、大鼠 (*Rattus norvegicus*)、小鼠 (*Mus musculus*) 和人类 (*Homo sapiens*) GlyCAM-1 的氨基酸序列,结合本研究中获得 MH917952.1 构建系统进化树。由图 2 可知, MH917952.1 首先与山羊的 GlyCAM-1 序列聚到一支,然后与藏羚羊、野牦牛和黄牛的序列聚到一起,说明获得的 MH917952.1 为绵羊的 GlyCAM-1 序列。

### 2.2 GlyCAM-1 蛋白理化性质

绵羊 GlyCAM-1 蛋白的原子组成为 C<sub>749</sub>H<sub>1234</sub>N<sub>200</sub>O<sub>241</sub>S<sub>4</sub>, 分子质量为 17 025.47, 等电点为 5.27。在编码的 154 个氨基酸中,根据 R 侧链的电荷数和极性,可以分为两类:1) 极性氨基酸,占总氨基酸数目的 57.15%,包括 20 个带正电荷的极性氨基酸 (R + K + H) (12.99%), 22 个带负电荷的极性氨基酸 (D + E) (14.29%) 和 46 个不带电荷的极性氨基酸 (G、S、T、C、Y、N、Q) (29.87%); 2) 66 个非极性氨基酸 (A、V、I、L、F、W、M、P), 占总氨基酸数目的 42.85%。GlyCAM-1 的消光系数在 280 nm 时为 1 490, 不稳定系数为 49.83, 说明该蛋白是一个不稳定蛋白。GlyCAM-1 在哺乳动物网织红细胞内的半衰期为 30

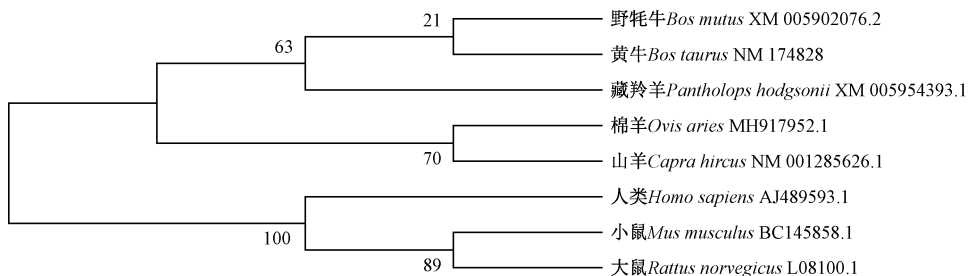


图 2 GlyCAM-1 氨基酸序列的同源性分析

Fig.2 Homology analysis of the amino acid sequence of GlyCAM-1

h, 脂肪系数为 95.06, 总平均亲水性为 -0.442, 说明该蛋白是一个可溶性蛋白。

绵羊 *GlyCAM-1* 蛋白的二级结构由 76 个  $\alpha$ -螺旋 (helix) (49.35%)、73 个无规则卷曲 (coil) (47.40%) 和 5 个延伸链 (extended strand) (3.25%) 组成。磷酸化位点检测结果表明, 该蛋白功能区域富含丝氨酸 (16 个) 和苏氨酸 (10 个) 磷酸化位点。GO 功能注释结果表明, *GlyCAM-1* 基因主要注释到蛋白质结合、离子结合及细胞粘附等生物学过程中。KEGG 分析结果表明, *GlyCAM-1* 基因注释到细胞粘附分子代谢通路中。String 分析结果表明, *GlyCAM-1* 同催乳素 (prolactin, PRL) 蛋白、粘膜血管贴壁细胞黏附分子 1 (mucosal vascular addressing cell adhesion molecule 1, MadCAM-1)、CD34 和 L-选择素 (L-selectin) 具有直接的联系 (图 3)。

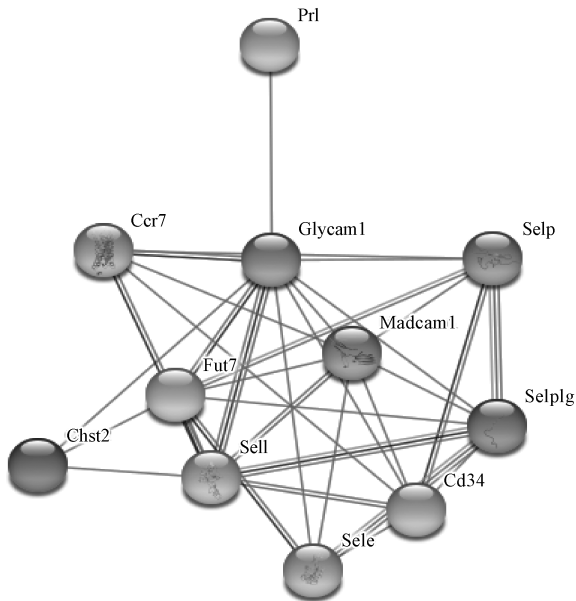


图 3 绵羊 *GlyCAM-1* 蛋白网络互作图

Fig.3 The protein interaction network of sheep *GlyCAM-1*

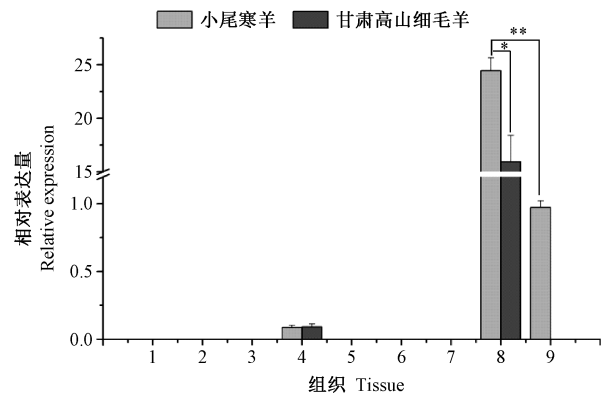
### 2.3 *GlyCAM-1* 基因核苷酸序列变异特征分析

测序结果表明, 在 *GlyCAM-1* 基因 CDS 区检测到 7 个 SNPs, 其中 2 个为同义突变: c. 39 C>T (p. A) 和 c. 258 G>A (p. L); 5 个为错义突变, 分别为 c. 41 C>T (p. A14V)、c. 43 A>G (p. T15A)、c. 279 A>T (p. Q93H)、c. 292 A>T (p. T98S) 和 c. 371 A>G (p. H124R)。

### 2.4 *GlyCAM-1* 基因的表达特征分析

利用 RT-qPCR 法检测处于泌乳高峰期的小尾寒羊和甘肃高山细毛羊中 *GlyCAM-1* 基因在心脏、肝脏、脾脏、肺脏、肾脏、背最长肌、卵巢和乳腺组织中的表达

量, 同时检测该基因在空怀期小尾寒羊乳腺组织中的表达量。由图 4 可知, *GlyCAM-1* 基因的表达具有明显的组织特异性。在检测的 8 个组织中, *GlyCAM-1* 基因在乳腺中的表达量最高, 在肺脏组织中微弱表达, 在其他 6 个组织中均不表达。同时 *GlyCAM-1* 的表达也呈现出品种特异性, 在泌乳高峰期的乳腺组织中, *GlyCAM-1* 基因在小尾寒羊中的表达量是甘肃高山细毛羊的 1.6 倍 ( $P < 0.05$ )。此外, 不同泌乳时期, *GlyCAM-1* 的表达量也不同。小尾寒羊中 *GlyCAM-1* 基因在泌乳高峰期中的表达量是空怀期表达量的 25 倍 ( $P < 0.01$ )。



注: 1~8 依次表示心脏、肝脏、脾脏、肺脏、肾脏、卵巢、背最长肌、泌乳高峰期小尾寒羊和甘肃高山细毛羊的乳腺组织; 9 为空怀期小尾寒羊的乳腺组织; \* 和 \*\* 表示不同乳腺组织间差异显著 ( $P < 0.05$ ) 和极显著 ( $P < 0.01$ )。

Note: 1~8: Heart, liver, spleen, lung, kidney, ovary, longissimus dorsi and mammary gland at the peak of lactation, respectively. 9: The mammary gland tissue of non-lactating period in the small-tailed Han sheep. \* and \*\* indicates the significant difference at 0.05 level and 0.01 level between different gland tissue.

图 4 *GlyCAM-1* 基因在小尾寒羊和甘肃高山细毛羊不同组织中的表达量

Fig.4 Expression of *GlyCAM-1* gene in different tissues of Small Tail Han sheep and Gansu alpine merino sheep

## 3 讨论

### 3.1 MFGMPs 的生物学功能

MFGMPs 是乳中的一种重要蛋白质, 主要以乳脂肪球膜的形式存在, 在乳腺分泌细胞内形成并分泌到乳汁中。尽管 MFGMPs 只占牛奶总蛋白质含量的 1%~2%, 但具有多样性的生物学功能, 在幼崽的细胞生长过程及防御机能<sup>[3, 14]</sup>、动物的免疫和发育等过程发挥了重要作用。此外, Haramizu 等<sup>[15]</sup> 研究发现, MFGMPs 还可以提高小鼠的耐力和肌肉运动能力。在

新生儿的辅食中添加 MFGMPs,可减少腹泻的发生<sup>[2]</sup>。MFGMPs 的数量和含量存在物种差异性,如奶牛<sup>[16]</sup>、奶山羊<sup>[17]</sup>和绵羊<sup>[4]</sup>乳中的 MFGMPs 的数量和含量各不相同。粘液素是粘膜物理屏障的组成部分,可以作为物理屏障防御病原微生物对呼吸道、胃肠道和子宫内膜等部位的感染。作为 MFGMPs 的重要组成部分,GlyCAM-1 同样在免疫保护和细胞粘附方面具有重要的作用<sup>[18]</sup>。研究发现,GlyCAM-1 的表达量与乳蛋白含量呈正相关。Ji 等<sup>[19]</sup>利用 iTRAQ(isobaric tags for relative and absolute quantitation)技术鉴定了牦牛和奶牛的乳脂球膜蛋白,发现在高乳蛋白率(5.16%)的牦牛乳汁中,GlyCAM-1 高度富集,并且是低乳蛋白率(2.79%)奶牛的 10.15 倍。由于绵羊的乳蛋白含量(5.15%)显著高于山羊(4.10%)和奶牛<sup>[20-21]</sup>,推测绵羊乳中 GlyCAM-1 表达量可能也高于山羊和奶牛,但还需进一步证实。

### 3.2 GlyCAM-1 序列特征及生物信息学分析

本研究利用 RT-PCR 方法首次克隆了绵羊 GlyCAM-1 基因的编码区序列。序列比对表明,绵羊 GlyCAM-1 的氨基酸序列与山羊具有较高的同源性(98.7%),与黄牛之间的同源性为 88.96%。通过直接测序技术在 GlyCAM-1 基因中检测到 7 个 SNPs,表明该基因具有丰富的多态性。李晶等<sup>[22]</sup>研究发现,GlyCAM-1 的核苷酸变异与奶牛乳房炎有相关性,GlyCAM-1 第 3 外显子 2081 处 A>T 的变异降低了乳房炎的感染率。GO 功能注释表明 GlyCAM-1 基因主要参与蛋白结合、离子结合等生物学过程,KEGG 结果表明 GlyCAM-1 参与细胞分子粘附代谢通路。细胞粘附分子是在细胞表面表达的糖蛋白,且在包括止血、免疫应答、炎症、胚胎发生等多种生物过程中起关键作用。在淋巴器官中,GlyCAM-1 作为 L-选择蛋白的配体参与了淋巴细胞和内皮细胞之间的相互作用<sup>[23]</sup>,使细胞与细胞间、细胞与基质间,或细胞-基质-细胞间发生粘附,从而参与细胞的识别、活化、信号转导和细胞的增殖与分化。因此,GlyCAM-1 可以通过该途径调控乳腺细胞的增殖、分化和凋亡<sup>[24]</sup>。

蛋白互作研究可以在分子水平揭示蛋白质功能以及包括生长、发育、代谢、分化和凋亡在内的细胞活动规律<sup>[24]</sup>。通过蛋白-蛋白互作网络分析发现,GlyCAM-1 与 PRL、L-Selectin、MadCAM-1 和 CD34 蛋白具有直接的作用。GlyCAM-1、MadCAM-1 和 CD34 都是 L-选择素的配体,具有粘蛋白样多肽结构和硫酸化的 O-连接聚糖<sup>[25]</sup>,它们通过粘附参与细胞的识别、活化、信号转导增殖及凋亡等过程。蛋白质半衰期与其稳定

性有关。贾浩等<sup>[26]</sup>发现,蛋白质的半衰期长,其稳定性高。本研究结果表明,GlyCAM-1 基因编码的蛋白具有较长的半衰期,但属于不稳定蛋白,这可能与该蛋白富含丝氨酸和苏氨酸磷酸化位点有关,磷酸化使得该蛋白的稳定性降低<sup>[27-28]</sup>。

### 3.3 GlyCAM-1 基因组织表达特征分析

本研究中,GlyCAM-1 基因的表达具有明显的组织特异性、品种特异性和时空特异性。在 8 个组织中,GlyCAM-1 基因在泌乳期的乳腺组织中高表达,在空怀期乳腺组织和肺脏组织中也有少量表达,其他组织中均不表达,这与前人研究结果相一致。研究表明,GLYCAM-1 基因在泌乳期奶牛和小鼠的乳腺组织中高度表达,同时在肺脏组织、子宫中也表达<sup>[29-31]</sup>。GlyCAM-1 在乳腺上皮细胞合成后分泌到乳汁中<sup>[7]</sup>,其表达量受到激素和外源性刺激的调控。例如催乳素(prolactin,PRL)能诱导 GlyCAM-1 的表达,而孕酮则抑制了该基因的表达<sup>[32]</sup>。断奶 10 d 后,乳腺中的 GlyCAM-1 表达量降至极低的水平,表明乳腺中 GlyCAM-1 的调节与其他乳蛋白质相似,都需要通过哺乳幼崽等外源性刺激才能继续表达<sup>[33]</sup>。GlyCAM-1 基因在子宫和胚胎着床阶段的表达也进行了较多的研究。绵羊子宫中 GlyCAM-1 mRNA 和蛋白质的水平与处于发情周期的子宫内膜上皮中孕酮受体表达的变化趋势非常相似,这表明黄体酮可能是子宫中 GlyCAM-1 的潜在调节因子<sup>[31]</sup>。本试验中,GlyCAM-1 基因在高泌乳量小尾寒羊乳腺组织中的表达量显著高于低泌乳量甘肃高山细毛羊乳腺组织,说明该基因能促进绵羊的泌乳,这与 Romney 绵羊乳腺组织中的研究结果相似<sup>[10]</sup>。在产奶量有显著差异的 Assaf(400 kg)和 Churra(117 kg)绵羊之间,GlyCAM-1 表达量并无显著差异( $P>0.05$ )<sup>[9]</sup>,这可能与研究中不同的 RNA 来源有关,本研究分析的是乳腺组织,而 Suárez Vega 等<sup>[9]</sup>采用乳汁中的 MSC 检测其各时期的基因表达谱。Cánovas 等<sup>[34]</sup>和 Wickramasinghe 等<sup>[35]</sup>已证实奶牛乳腺和 MSC 之间基因的表达模式有较大的差异。

本试验结果表明,在空怀期小尾寒羊乳腺组织中,GlyCAM-1 基因的表达量很低,但在乳汁合成过程中,该基因的表达量又上升。在乳腺的发育及乳汁合成中,该基因的表达量变化趋势与激素分泌有关,GlyCAM-1 基因受到 PRL 的诱导才能表达。PRL 在发情期开始分泌,导致该基因在泌乳高峰期有较高的表达量,而在空怀期表达量很低<sup>[36]</sup>。同时,GlyCAM-1 在不同乳腺发育时期的表达量差异可能还与腺泡发育有关<sup>[32]</sup>。与空怀期相比,泌乳高峰期的乳腺中含有大量的腺泡,这也导

致了该基因在泌乳高峰期有较高的表达量。

## 4 结论

本研究首次克隆获得了小尾寒羊 *GlyCAM-1* 基因的 CDS 全长序列,该基因编码 154 个氨基酸。绵羊的 *GlyCAM-1* 氨基酸序列与山羊和牛的 *GlyCAM-1* 具有较高的序列同源性。*GlyCAM-1* 蛋白的二级结构主要以  $\alpha$ -螺旋和无规则卷曲为主。蛋白互作分析结果表明,*GlyCAM-1* 蛋白与 CD34、MadCAM-1 都作为 L-选择素的配体发挥作用。此外,*GlyCAM-1* 基因的 CDS 区发现了 7 个 SNPs(包括 2 个同义突变和 5 个非同义突变)。*GlyCAM-1* 基因在绵羊上的表达具有明显的组织特异性、品种特异性和时空特异性。本研究结果丰富了绵羊基因组数据,为深入研究 *GlyCAM-1* 基因的泌乳生物学功能及其机理提供了基础数据。

## 参考文献:

- [ 1 ] Akers R M. A 100-Year review; Mammary development and lactation [J]. *Journal of Dairy Science*, 2017, 100(12):10332-10352
- [ 2 ] Nelly Z, Anne Staudt K, Gitte G, Graciela R, Henry G, Norma V. Efficacy of an MFGM-enriched complementary food in diarrhea, anemia, and micronutrient status in infants [J]. *Journal of Pediatric Gastroenterology and Nutrition*, 2015, 53(5):561-568
- [ 3 ] Riccio P. The proteins of the milk fat globule membrane in the balance [J]. *Trends in Food Science and Technology*, 2004, 15(9):458-461
- [ 4 ] Pisanu S, Ghisaura S, Pagnozzi D, Biosa G, Tanca A, Roggio T, Uzzau S. The sheep milk fat globule membrane proteome [J]. *Journal of Proteomics*, 2011, 74(3):350-358
- [ 5 ] Le P F, Cassy S, Hayes H, Martin P. Structure and expression of goat *GLYCAM1* gene; Lactogenic-dependent expression in ruminant mammary gland and interspecies conservation of the proximal promoter [J]. *Gene*, 2003, 313(1):83-89
- [ 6 ] Lasky L A, Singer M S, Dowbenko D, Imai Y, Henzel W J, Grimely C. An endothelial ligand for L-selectin is a novel mucin-like molecule [J]. *Cell*, 1992, 69(6):927-937
- [ 7 ] Dowbenko D, Kikuta A, Fennie C, Lasky L A. Glycosylation-dependent cell adhesion molecule 1 (*GlyCAM 1*) mucin is expressed by lactating mammary gland epithelial cells and is present in milk [J]. *Journal of Clinical Investigation*, 1993, 92(2):952-960
- [ 8 ] Hwang S T, Singer M S, Yednock T A, Yednock T A, Rosen S D. *GlyCAM-1*, a physiologic ligand for L-selectin, activates beta 2 integrins on naive peripheral lymphocytes [J]. *Journal of Experimental Medicine*, 1996, 184(4):1343-1348
- [ 9 ] SuárezVega A, Gutiérrez Gil B, Klopp C, Arranz J. Characterization and comparative analysis of the milk transcriptome in two dairy sheep breeds using RNA sequencing [J]. *Scientific Report*, 2015, 5:18399
- [ 10 ] Paten A M, Duncan E J, Pain S J, Pain S, Paterson S W, Kenyon P R. Functional development of the adult ovine mammary gland insights from gene expression profiling [J]. *BMC Genomics*, 2015, 16:748-761
- [ 11 ] Nikolaj B, Steen G, Soren B. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites [J]. *Journal of Molecular Biology*, 1999, 294(5):1351-1362
- [ 12 ] Lim D, Kim N K, Park H S, Park H S, Lee H S, Cho Y M, Kim Tae. Identification of candidate genes related to bovine marbling using protein-protein interaction networks [J]. *International Journal of Biological Sciences*, 2011, 7(7):992-1002
- [ 13 ] Li T Y, Diao H Y, Zhao L, Liu N, Yan Y Y, Tian X, Sun W. Identification of suitable reference genes for real-time quantitative PCR analysis of hydrogen peroxide-treated human umbilical vein endothelial cells [J]. *BMC Molecular Biology*, 2017, 18(1):1-8
- [ 14 ] Cavaletto M, Giuffrida M G, Conti A. Milk fat globule membrane components-a proteomic approach [J]. *Advances in Experimental Medicine and Biology*, 2008, 606(1):129-141
- [ 15 ] Haramizu S, Ota N, Otsuka A, Hashizume K, Sugita S, Hase T, Murase T, Shimotoyodome A. Dietary milk fat globule membrane improves endurance capacity in mice [J]. *Physiology Regulatory Integrative Comparative Physiology*, 2014, 307(8):1009-1017
- [ 16 ] Zou X, Guo Z, Jin Q, Huang J H, Xu X B, Wang X G. Composition and microstructure of colostrum and mature bovine milk fat globule membrane [J]. *Food Chemistry*, 2015, 185:362-370
- [ 17 ] Lu J, Liu L, Pang X, Pang X Y, Jia Z H, Zhao L L, Lv J P. Comparative proteomics of milk fat globule membrane in goat colostrum and mature milk [J]. *Food Chemistry*, 2016, 209:10-16
- [ 18 ] Van B J, Dekker J, Buller H A, Einerhand A W. Mucin gene structure and expression; Protection vs. adhesion [J]. *The American Journal of Physiology*, 1995, 269(51):613-627
- [ 19 ] Ji X X, Li X S, Ma Y, Li D. Differences in proteomic profiles of milk fat globule membrane in yak and cow milk [J]. *Food Chemistry*, 2017, 221:1822-1827
- [ 20 ] Leitner G, Lavon Y, Matzrafi Z, Benun O, Bezman D, Merin U. Somatic cell counts, chemical composition and coagulation properties of goat and sheep bulk tank milk [J]. *International Dairy Journal*, 2016, 58:9-13
- [ 21 ] 张容昶, 胡江. 牦牛生产技术[M]. 北京: 金盾出版社, 2002:185-188
- [ 22 ] 李晶, 王洪梅, 王长法, 柳楠, 李秋玲, 侯明海, 李建斌, 仲跻峰. 中国荷斯坦牛 *GlyCAM-1* 基因遗传多态性与乳房炎关联性分析[J]. *中国农业科技导报*, 2008, 10(6):68-73
- [ 23 ] Rosen S D, Bertozzi C R. The selectins and their ligands [J]. *Current Opinion in Cell Biology*, 1994, 6(5):663-673
- [ 24 ] 刘畅. 整联蛋白 alpha6、beta4 介导层黏连蛋白作用调控乳腺上皮细胞[D]. 哈尔滨: 东北农业大学, 2012
- [ 25 ] Uchimura K, Rosen S D. Sulfated L-selectin ligands as a therapeutic target in chronic inflammation [J]. *Trends in Immunology*, 2006, 27(12):559-565
- [ 26 ] 贾浩, 张小白, 宋晓峰. 人类胞内蛋白半衰期与其亚细胞定位的相关性研究[J]. *计算机与应用化学*, 2011, 28(4):411-414

- [27] 李蒙, 李铮, 李欣, 杜曼婷, 宋璇, 张德权. 磷酸化水平对肌红蛋白稳定性的影响[J]. 中国农业科学, 2017, 50(22):4382-4388
- [28] 卢曾奎, 马友记, 赵兴绪, 张勇, 张利平. 绵羊 *STMN1* 基因的克隆、序列分析及其组织表达[J]. 核农学报, 2018, 32(10):1931-1938
- [29] Groenen M A, Dijkhof R J, van der Poel J J. Characterization of a *GlyCAM1 like* gene (glycosylation dependent cell adhesion molecule 1) which is highly and specifically expressed in the lactating bovine mammary gland [J]. *Gene*, 1995, 158(2):189-195
- [30] Nishimura T, Kohmoto K. Regulation of glycosylation-dependent cell adhesion molecule 1 (*GlyCAM-1*) gene in the mouse mammary gland differs from that of casein genes [J]. *Comparative Biochemistry and Physiology Part B*, 2001, 129(1):149-156
- [31] Spencer T E, Bartol F F, Bazer F W, Johnson G A, Joyce M M. Identification and characterization of glycosylation-dependent cell adhesion molecule 1-like protein expression in the ovine uterus [J]. *Biology of Reproduction*, 1999, 60(2):241-250
- [32] Hou Z, Bailey J P, Vomachka A J, Masada M, Lockefleer J, Horseman N D. Glycosylation-dependent cell adhesion molecule 1 (*GlyCAM 1*) is induced by prolactin and suppressed by progesterone in mammary epithelium [J]. *Endocrinology*, 2000, 141(11):4278-4283
- [33] Hobbs A A, Richards D A, Kessler D J, Rosen J M. Complex hormonal regulation of rat casein gene expression [J]. *Journal of Biological Chemistry*, 1982, 257(7):3598-3605
- [34] Cunos A, Rincon G, Bevilacqua C, Islas-Trejo A, Brenaut P, Hovey R C, Boutinaud M, Morgenthaler C, VanKlompberg M K, Martin P, Medrano J F. Comparison of five different RNA sources to examine the lactating bovine mammary gland transcriptome using RNA-Sequencing [J]. *Scientific Reports*, 2014, 4: 5297
- [35] Wickramasinghe S, Rincon G, Islas-Trejo A, Medrano J F. Transcriptional profiling of bovine milk using RNA sequencing [J]. *BMC Genomics*, 2012, 13:13-45
- [36] 李庆章. 奶山羊乳腺发育与泌乳生物学[M]. 北京: 科学出版社, 2011:102-103

# Gene Cloning, Nucleotide Variation and Tissue Expression Analysis of *GlyCAM-1* in Sheep (*Ovis aries*)

HAO Zhiyun WANG Jiqing\* HU Jiang LIU Xiu LI Shaobin LUO Yuzhu\*

(College of Animal Science and Technology/Gansu Key Laboratory of Herbivorous Animal Biotechnology/

Gansu Engineering Lab of Genetic Improvement in Ruminants, Gansu Agricultural University, Lanzhou, Gansu 730070)

**Abstract:** The glycosylation-dependent cell adhesion molecule (*GlyCAM-1*) is a member of the mucin glycoprotein family and is an important component of milk fat globule membrane protein (MFGMPs) synthesized by mammary epithelial cells. *GlyCAM-1* plays an important role in mammary gland development and lactation. To explore the sequence characteristics, nucleotide sequence variation and expression characteristics of *GlyCAM-1* gene, the mammary gland tissue of high-lactating small-tailed Han sheep (lacking peak and empty period) and low lactation Gansu alpine fine-wool sheep (the peak of lactation) were used to study. The coding region of sheep *GlyCAM-1* gene was cloned by cloning sequencing. The physicochemical properties and structure of *GlyCAM-1* protein were analyzed by bioinformatics and the tissue expression characteristics of the *GlyCAM-1* gene were detected by RT-qPCR. The results showed that the CDS of *GlyCAM-1* gene in sheep was 465 bp in length and encoded 154 amino acids. The sequencing results showed that 7 SNPs were identified in the CDS region of the gene, of which 2 were synonymous mutations and 5 were missense mutations. The secondary structure of *GlyCAM-1* is predominantly composed of  $\alpha$ -helix and random coil, and the extended chain is dispersed throughout the protein structure. Results of String analysis indicated that *GlyCAM-1* protein interacts with CD34 and MadCAM-1 as ligands of L-selectin. RT-qPCR results indicate that the expression of *GlyCAM-1* gene is tissue-specific, breed-specific and space-time specific. Among the eight tissues of the mammary gland, heart, liver, lung, spleen, kidney, ovary and longissimus dorsi. *GlyCAM-1* gene was only expressed in mammary gland and lung tissues, but not in the other six tissues. and the expression level in the mammary gland is higher. In the mammary gland at the peak of lactation, the expression of *GlyCAM-1* level in the high-lactating small-tailed Han sheep was significantly higher than that in the Gansu alpine fine-wool sheep ( $P < 0.05$ ). In the mammary gland of Small Tail Han sheep, the expression at the peak of lactation was significantly higher than that in the empty period ( $P < 0.01$ ). The results of this study enriched the sheep genome data and provided basic data for further study on the lactation biological function and mechanism of *GlyCAM-1* gene.

**Keywords:** *GlyCAM-1* gene, mammary gland, sheep, clone, tissue expression, SNPs