

牦牛 α -乳白蛋白基因单核苷酸多态性分析

崔艳华¹, 曹 喻¹, 曲晓军², 李海梅¹, 董爱军¹, 马 莺¹

(1. 哈尔滨工业大学 食品科学与工程学院, 150090 哈尔滨; 2. 黑龙江省科学院微生物研究所, 150010 哈尔滨)

摘 要: 为明确牦牛乳中 α -乳白蛋白遗传多样性, 以 α -乳白蛋白基因(*LAA*)的 I、II、III 和 IV 外显子为研究对象, 采用聚合酶链反应-单链构象多态性(PCR-SSCP)方法对麦洼牦牛的 *LAA* 进行分析, 旨在从分子水平揭示 *LAA* 在牦牛乳中遗传多样性分布. 牛属不同来源的 α -乳白蛋白基因的编码序列被用于构建系统发育树. 结果表明, 牦牛乳中 *LAA* I 外显子区域以杂合体形式存在, 并发现了一个新变异体 JN084189. 不同来源 *LAA* 的系统发育树分析表明, *LAA* 在进化上存在一定的物种特异性.

关键词: 牦牛乳; α -乳白蛋白; 遗传多样性; 聚合酶链反应-单链构象多态性

中图分类号: Q52

文献标志码: A

文章编号: 0367-6234(2012)06-0092-05

A single nucleotide polymorphism and sequence analysis of α -lactalbumin gene (*LAA*) in Chinese *Bos grunniens* (yak)

CUI Yan-hua¹, CAO Yu¹, QU Xiao-jun², LI Hai-mei¹, DONG Ai-jun¹, MA Ying¹

(1. School of Food Science and Engineering, Harbin Institute of Technology, 150090 Harbin, China;

2. Institute of Microbiology, Heilongjiang Science Academy, 150010 Harbin, China)

Abstract: A polymorphism of α -lactalbumin gene *LAA* exon I, II, III, and IV has been identified by PCR-single strand conformation polymorphism protocol to reveal the distribution of diversity of α -lactalbumin in Maiwa yak at the molecular level. The sequences corresponding to the coding region of α -lactalbumin gene were used for phylogenetic analysis to establish the phylogenetic relationship of *LAA* variants in the *Bos* genus. Experimental results indicated that most yak milks were with hybrid genotype in exon I of α -lactalbumin. A new variant had been identified in Maiwa yak breed by sequencing analysis and entered into GenBank with accession number JN084189. The phylogenetic analysis presented α -lactalbumin evolution consistent with species at a degree.

Key words: yak milk; α -lactalbumin; polymorphism; PCR-SSCP

牛乳蛋白中主要包含两大类蛋白, 即酪蛋白和乳清蛋白. 其中乳清蛋白占总蛋白的 12% ~ 14%, 包括 α -乳白蛋白(α -LA)、 β -乳球蛋白(β -LG)、血清白蛋白(SA)、免疫球蛋白(Ig)和乳铁蛋白(LF)^[1-2]. 由于氨基酸的置换或缺失、磷酸化位点不同、糖基化差异等造成了牛乳蛋白存在着遗传变异体. 研究表明, 某些乳蛋白遗传变异体与产乳量、乳组成、干酪产量等重要性状密切相关, 因此, 乳蛋白多样性研究引起了广泛关注^[1-3].

α -LA 占牛乳中乳清蛋白的 5% (质量分数), 由 *LAA* 基因编码^[4]. α -LA 基因 (*LAA*) 全长 2 kb, 包括 4 个外显子和 3 个内含子, 定位于牛基因组的第 5 号染色体上^[5]. 研究表明, 在牛乳中, α -LA 不仅作为牛乳蛋白中的一部分存在, 还对乳糖的生成及牛乳的分泌起调控作用^[6-7]. 同时, α -LA 具有抗菌活性, 并激活粘液代谢, 具有保护胃肠道免受病原菌侵袭的作用^[7]. 现已在不同来源的牛乳中发现了 3 种 α -LA 的变异体, 即变异体 A、B 和 C^[3]. 目前, 对于牦牛 α -LA 研究主要集中在蛋白质水平, 分子水平研究鲜有报道, 遗传多样性分子水平研究未见报道^[8]. 本研究以 *LAA* 的 I、II、III 和 IV 外显子为研究对象, 采用 PCR-SSCP 方法分析了牦牛乳 *LAA* 的遗传多样性. 并

收稿日期: 2011-06-05.

基金项目: 国家自然科学基金资助项目(30871953;31071571).

作者简介: 崔艳华(1978—), 女, 博士, 副教授;

马 莺(1961—), 女, 教授, 博士生导师.

通信作者: 马 莺, maying@hit.edu.cn.

采用生物信息学方法对现有不同来源 α -LA 变异体进行分析,建立系统发育树,旨在揭示其多样性的分布及进化。

1 实验

1.1 牦牛乳

34份牦牛乳取自四川红原麦洼牦牛。取样后加入叠氮化钠(NaN_3 , 0.4 g/L)防止蛋白沉淀,存于 $-20\text{ }^\circ\text{C}$ 。

1.2 试剂

蔗糖、Tris 碱、硼酸、盐酸、氯化镁、Triton-X-100、氯化钠、乙二醇四乙酸二钠盐、异丙醇、无水乙醇、丙烯酰胺、甲叉双丙烯酰胺、N,N,N',N'-四甲基乙二胺、过硫酸铵等均为分析纯,购自北京试剂公司。蛋白酶K购自TAKARA公司。KOD-Plus-DNA

聚合酶、Target cloneTM-plus kit 购自东洋纺(上海)生物科技有限公司。引物由上海生工合成。

1.3 主要仪器

高速冷冻离心机(Biofuge Stratos, Heraeus)、电泳仪(BioRad)、通用突变检测系统(DCodeTM, BioRad)、凝胶成像系统(BioRad)、PCR仪(TAKARA公司)。

1.4 牦牛乳 DNA 的提取

采用常规离心法提取牦牛乳中体细胞。以乳体细胞为材料,提取牦牛基因组DNA,方法见文献[9]。

1.5 引物以及 PCR-SSCP

根据牛 α -LA 基因(NCBI 序列号 X06366)的I、II、III和IV外显子区域设计引物LAA-E1F和LAA-E1R、LAA-E2F和LAA-E2R、LAA-E3F和LAA-E3R、LAA-E4F和LAA-E4R,分别用于扩增 α -LA的I、II、III和IV外显子及内含子部分区域(见表1)。

表1 本研究所用引物

引物名称	引物序列 (5'-3')	对应位置 [#]	外显子位置 [#]	长度/bp	$t_m/^\circ\text{C}$	PCR 反应 退火温度/ $^\circ\text{C}$	扩增片段 大小/bp
LAA-E1F	TGGATGTAAGGCTTGATG	664-680	Exon1	18	50.3	50.5	309
LAA-E1R	AAGAGGATGAAGAGAATGG	955-972	766-898	19	50.2		
LAA-E2F	AGATTGGTTGGAGAGCCT	1122-1138	Exon 2	18	53.6	52.0	392
LAA-E2R	GTATTATCAAGTCCCACAGT	1495-1513	1220-1378	20	51.6		
LAA-E3F	CAGAAGGCAACAGGCATA	1737-1753	Exon3	18	53.9	52.0	369
LAA-E3R	GGACTGAAGTAAGTGAAGC	2088-2105	1852-1927	19	52.1		
LAA-E4F	CTACATCCTAAGGCACGC	2203-2219	Exon4	18	54.2	53.0	316
LAA-E4R	GACAGAAGCAGCAAAGAC	2502-2518	2432-2489	18	53.1		

注:[#]为与X06366序列对应位置。

以提取的牦牛基因组DNA为模板,分别以LAA-E1F和LAA-E1R、LAA-E2F和LAA-E2R、LAA-E3F和LAA-E3R、LAA-E4F和LAA-E4R为引物扩增 α -LA的I、II、III和IV外显子区域以及部分内含子区域。

PCR反应程序为 $94\text{ }^\circ\text{C}$ 2 min 预变性,然后以 $94\text{ }^\circ\text{C}$ 15 s、退火30 s、 $68\text{ }^\circ\text{C}$ 25 s为1个循环,共进行30个循环,之后 $68\text{ }^\circ\text{C}$ 延伸7 min, $4\text{ }^\circ\text{C}$ 保存。PCR反应体系为:10×PCR buffer 2 μL , dNTPs (2.5 mmol/L) 2 μL , MgCl_2 (25 mmol/L) 0.8 μL ,上下游引物(10 $\mu\text{mol/L}$)各1 μL , KOD-Plus-DNA聚合酶2 U, DNA 10~50 ng,用双蒸水补足体积至20 μL 。取5 μL PCR产物,以1.5%琼脂糖凝胶进行电泳检测产物纯度。

10 μL 的PCR产物用于PCR-SSCP分析。加入等体积的2×SSCP Gel Loading Dye(DCodeTM),样品在 $95\text{ }^\circ\text{C}$ 反应10 min后迅速冰浴。点样于20 cm×20 cm、8%聚丙烯酰胺凝胶(丙烯酰胺与双丙烯酰胺的质量比为37.5:1)、 $15\text{ }^\circ\text{C}$ 、150 V下

电泳7 h。

1.6 克隆与测序

采用Target cloneTM-plus kit将PCR产物克隆到pTA2载体上。将筛选获得阳性克隆进行测序,正反向同时测序。获得的序列信息提交至GenBank,序列号为JN084188(YAK1)、JN084189(YAK2)。

1.7 牛属 α -乳白蛋白序列分析

LAA基因核酸序列来自NCBI数据库(www.ncbi.nlm.nih.gov),其中包括家牛LAA基因核酸序列(*Bos taurus*, *Bos indicus*) AB052163、AB052164、AB052165、AB052166、AB052167、AF249896、X06366、EU200932、EU422984;两个牦牛(*Bos grunniens*)LAA基因核酸序列 AF194372和GU562877;6个水牛(*Bubalus bubalus*)LAA基因核酸序列 EF408824、AY726609、AY726610、AY726611、AY726612、DQ785796。

α -LA的I、II、III和IV外显子区域被用于构建系统进化分析。采用CLUSTAL X软件进行多

重序列比对^[10], TREECON 软件计算遗传距离, 并建立系统发育树^[11].

2 结果与分析

2.1 牦牛 α -LA 遗传多样性分析

2.1.1 PCR 扩增

以牦牛乳基因组 DNA 为模板, 分别以 LAA-E1F 和 LAA-E1R、LAA-E2F 和 LAA-E2R、LAA-E3F 和 LAA-E3R、LAA-E4F 和 LAA-E4R 为引物扩增 α -LA 的 I、II、III 和 IV 外显子以及部分内含子区域, 片段大小约分别为 300、400、400 和 300 bp(图 1), 与预期相符. 为便于阐述, 将上述片段命名为 A、B、C、D 片段.

2.1.2 PCR-SSCP 分析与测序分析

以 LAA-E1F 和 LAA-E1R 引物, 扩增获得的 PCR 产物 A 片段(664-972, X06366 相应位置)包含了 α -乳白蛋白基因的外显子 I 区域(766-898)、信号肽区域(766-822)、内含子 I 部分区域(899-972). 在 34 个牦牛乳样品中, PCR-SSCP 分析均为杂合型(见图 2(a)). 选 5 个样品与 pTA2 载体连接, 分别选取阳性克隆 10 个进行正反向测序. 结果发现, 在内含子 I 区域无差异, 而信号肽区域(800)、外显子 I 区域(829)处存在差异(见图 3). 变异体 1 序列与目前已知牛属 α -乳白蛋白外显子 I 区域均存在差异, 为 1 个新变异体, 已提交 GenBank, 获得序列号 JN084189. 变异体 2 与来自大多数其他来源的 α -乳白蛋白的外显子 I 区域 100% 同源.

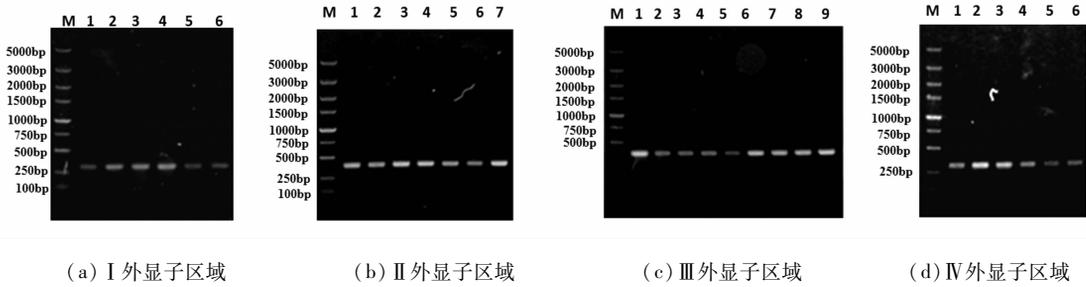


图 1 牦牛 α -乳白蛋白基因的 I、II、III 和 IV 外显子区域 PCR 分析

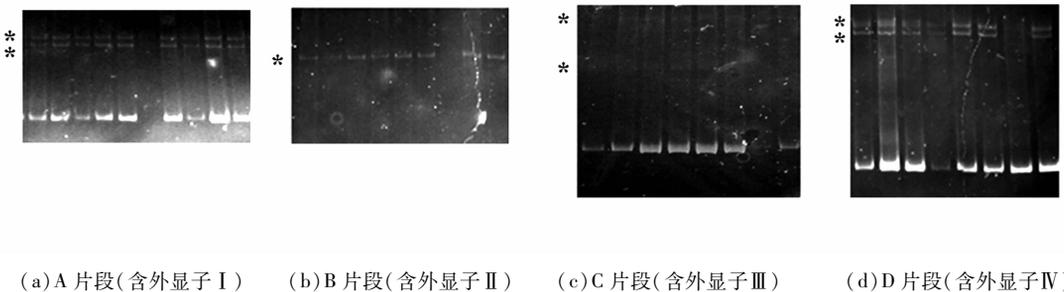


图 2 牦牛 α -乳白蛋白基因的 PCR-SSCP 分析

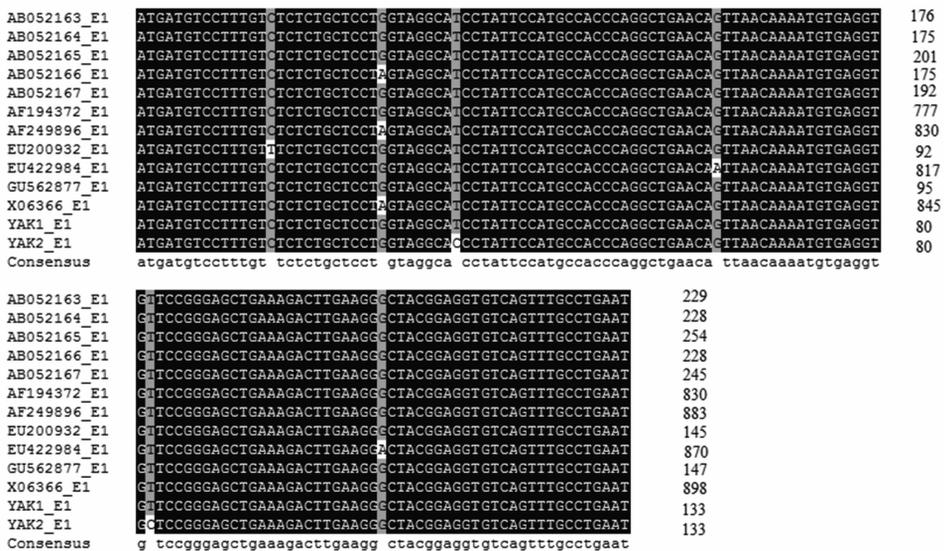


图 3 牛属乳白蛋白基因的外显子 I 序列比对图

以 LAA-E2F 和 LAA-E2R 引物,扩增获得的 PCR 产物 B 片段(1122 - 1513, X06366 相应位置)包含了 α -乳白蛋白基因的外显子 II 区域(1220 - 1378)和内含子 II 部分区域(1379 - 1513). 在 34 个牦牛乳样品中,PCR-SSCP 分析无差异(见图 2(b)). 测序结果发现,在内含子 II 部分区域和外显子 II 区域均无差异. 且外显子 II 与 GU562877 的外显子 II 的同源性为 100%.

以 LAA-E3F 和 LAA-E3R 引物,扩增获得的 PCR 产物 C 片段(1737 - 2105, X06366 相应位置)包含了 α -乳白蛋白基因的外显子 III 区域(1852 - 1927)和内含子 II 部分区域(1928 - 2105). 在 34 个牦牛乳样品中,PCR-SSCP 分析均为杂合型(见图 2(c)). 测序分析表明,内含子 II 区域含有 1 个碱基差异(即 1985 位置为 A 或者 G, X06366 相应位置),而外显子 III 区域无差异.

以 LAA-E4F 和 LAA-E4R 引物,扩增获得的 PCR 产物 D 片段(2203 - 2518, X06366 相应位置)包含了 α -乳白蛋白基因的外显子 IV 区域(2432 - 2489)和内含子 III 部分区域(2203 -

2431). 在 34 个牦牛乳样品中,PCR-SSCP 分析均为杂合型(见图 2(d)). 测序结果发现在内含子 III 中存在 2 个碱基差异(即 2238 位置为 A 或者 G, 2276 位置为 G 或者 A, X06366 相应位置),而外显子 IV 与 AF194372、GU562877 序列完全相同.

2.2 不同来源的 α -LA 遗传多样性和进化分析

来自牛属和水牛属,包括 *Bos taurus*、*Bos indicus*、*Bos grunniens* 和 *Bubalus bubalus* 的 α -LA 外显子 I、II、III 和 IV 用于构建系统发育树,进行聚类分析. 以来自山羊 (*Capra hircus*) 的 α -LA 基因 (LAA) 相应序列 (EF564266) 作为参比. 当前,已命名 3 个 α -LA 变异体,即 A、B、C. B 变异体最常见,在瘤牛、欧洲牛中均有发现. 在爪哇野牛 (*Bos javanicus*) 中发现的 C 变异体,与 B 变异体的差异在于某处谷氨酸 (Glu) 被谷氨酰胺 (Gln) 替代^[12]. 由于 C 变异体未有具体序列信息,未用于聚类分析.

研究表明,可划分为 4 个分支,其中包括 1 个山羊的 α -LA 基因 (即 IV 组) 和 3 个主要序列分支 (I - III 组)(见图 4).

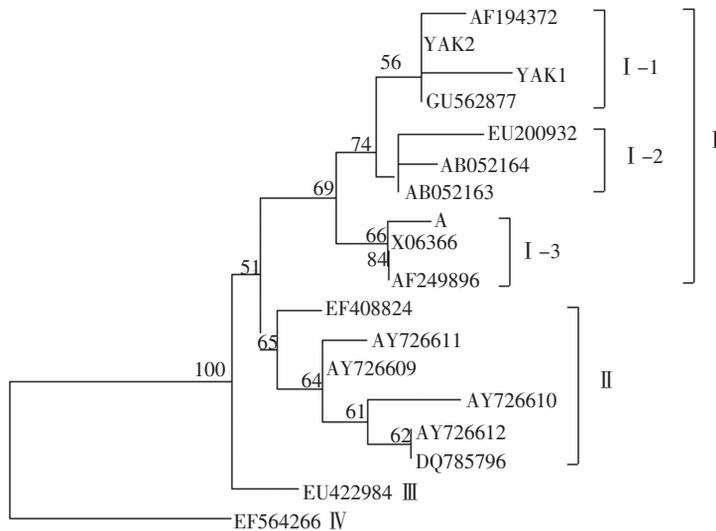


图 4 不同来源 LAA 编码序列的聚类图

第 I 组,均来自牛属,可分为 3 个亚组. I - 1 组包括变异体 AF194372、GU562877、YAK1 和 YAK2,均来自牦牛. 本研究获得 YAK2 序列与 GU562877 同源性为 100%,与 AF194372 相差 1 个碱基(见表 2). 而本研究发现的 YAK1 为一个新的变异体,与 GU562877 相差两个碱基,与 AF194372 相差 3 个碱基. I - 2 组包括 AB052164、AB05163、EU200932,其中 AB052164、AB05163 相差 1 个碱基,而 EU200932 与 AB052164 相差两个碱基,与 AB052164 相差 3 个碱基. I - 3 组包括 A 变异体、X06366、

AF249896,其中 X06366 和 AF249896 序列相同,为 B 变异体. 当前发现大多数牛属 α -LA 均为 B 变异体,与 A 变异体仅仅相差 1 个碱基. 因此,推测以上变异体直接由 α -LA * B 进化而来.

第 II 组包括 EF408824、AY726609、AY726610、AY726611、AY726612、DQ785796,均来自水牛 (*Bubalus bubalus*). 第 III 组 EU422984 独立构成 1 组. EU422984 来自瘤牛 (*Bos indicus*),序列较为独特. 聚类分析表明, α -LA 基因呈现了明显的物种差异. 牦牛 α -LA 基因与来自家牛的 α -LA 基因存在显著的差异.

表 2 牛属 α -乳白蛋白基因变异体分布

变异体	变异位置 [#]														来源,分布	
	780	792	800	829	848	852	873	1231	1233	1335	1354	1377	2471	2485		
B (X06366)	C Val	A Leu	T Ile	G Gln	T Phe	G Arg	G Gly	G Ala	G Ala	C Asp	A Asn	C Asp	G Gln	A Lys	所有品种, 最为普遍 (较普遍)所有 <i>B. indicus</i> 品种, 某些 <i>B. taurus</i>	
A						A His										
YAK1		G Leu	C Thr		C Leu			A Thr	G Thr	T Asp						<i>B. grunniens</i>
YAK2		G Leu						A Thr	G Thr	T Asp						<i>B. grunniens</i>
AF194372		G Leu						A Thr	G Thr	T Asp	G Asp					<i>B. grunniens</i>
GU562877		G Leu						A Thr	G Thr	T Asp						<i>B. grunniens</i>
AB052166		G Leu						A Thr	G Thr							<i>B. taurus</i>
AB052164		G Leu						A Thr	G Thr				A Gln			<i>B. taurus</i>
EU200932	T Val	G Leu						A Thr	G Thr					G Arg		<i>B. indicus</i>
EU422984		G Leu		A Gln			A Asp	A Thr	A Thr			T Asp				<i>B. indicus</i>

#—第 1 行为核酸序列位置;第 2 行为氨基酸序列位置. 氨基酸序列包括信号肽部分,即为前体蛋白. 加粗为差异位点.

3 结 论

1) 牦牛乳中 α -LA 的 I、III 和 IV 外显子区域的 PCR 片段呈现了相同的带型,均具有两个单链片段,表明多数以杂合体形式存在. 而 II 外显子区域的 PCR 片段则为一个单链片段表明以纯合体形式存在.

2) 对 PCR 产物进行测序分析发现了一个新的变异体 JN084189,该变异体与来自牦牛两个 AF194372、GU562877 在 I 外显子区域存在两个碱基差异,最终导致在蛋白水平上存在两个氨基酸差异,一处位于信号肽中. 新变异体的发现为研究 α -LA 的进化提供了参考.

3) 来自不同来源,包括 *Bos taurus*、*Bos indicus* 和 *Bos grunniens* 的 α -LA 基因序列系统发育分析发现 α -LA 在进化上存在一定的物种特异性. α -LA 具有多种生理功能,本研究将为 α -LA 变异体与生理功能之间的相互关系研究奠定基础.

参考文献:

[1] HECK J, SCHENNINK A, VALENBERG van H, *et al.* Effects of milk protein variants on the protein composition of bovine milk [J]. *J Dairy Science*, 2009, 92 (3): 1192 - 1202.

[2] 崔艳华,马莺,何胜华,等. 牛乳酪蛋白遗传多态性研究 [J]. *中国乳品工业*, 2010, 38(5): 32 - 37.

[3] FARRELL H M, JIMENEZ-FLORES J R, BLECK G T, *et al.* Nomenclature of the proteins of cows' milk-sixth revision [J]. *J Dairy Science*, 2004, 87: 1641 - 1674.

[4] MARTIN P, SZYMANOWSKA M, ZWIERZCHOWSKI

L, *et al.* The impact of genetic polymorphisms on the protein composition of ruminants milks [J]. *Reprod Nutr Dev*, 2002, 42: 433 - 459.

[5] HAYES H, PETIT E. Mapping of the β -lactoglobulin gene and of an immunoglobulin M heavy chain-like sequence in homologous cattle, sheep, and goat chromosomes [J]. *Mamma Genome*, 1993, 4: 207 - 210.

[6] STINNAKRE M G, VILOTTE J L, SOULIER S, *et al.* Creation and phenotypic analysis of α -lactalbumin-deficient mice [J]. *Proc Natl Acad Sci USA*, 1994, 91: 6544 - 6548.

[7] USHIDA Y, SHIMOKAWA Y, TOIDA T, *et al.* Bovine α -lactalbumin stimulates mucus metabolism in gastric mucosa [J]. *J Dairy Science*, 2007, 90: 541 - 546.

[8] BAI W L, YIN R H, ZHENG Y C, *et al.* Cloning and molecular characterization of a yak α -lactalbumin cDNA from mammary tissue [J]. *Livest Science*, 2010, 129: 122 - 128.

[9] 崔艳华,马莺,曲晓军,等. 牦牛乳中 DNA 提取方法的建立与优化 [J]. *哈尔滨工业大学学报*, 2011, 43(2): 1892 - 1895.

[10] THOMPSON J D, GIBSON T J, PLEWNIAC F, *et al.* The clustalx windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools [J]. *Nucl Acids Res*, 1997, 24: 4876 - 4882.

[11] PEER van der Y, WACHTER de R. TREECON for windows: a software package for construction and drawing of evolutionary trees for the microsoft windows environment [J]. *Comput Appl Biosci*, 1994, 10: 569 - 570.

[12] BELL K, HOPPER K E, MCKENZIE H A. Bovine α -lactalbumin C and α_{s1} -, β -, and κ -caseins of Bali (Banteng) cattle, *Bos (Bibos) javanicus* [J]. *Aust J Biol Science*, 1981, 34: 149 - 159.

(编辑 刘 彤)