



# 人溶菌酶密码子优化及其在牛乳腺细胞中高效表达

田媛<sup>1</sup>, 王力<sup>1</sup>, 龙凤<sup>1</sup>, 咎林森<sup>1,2</sup>, 成功<sup>1,2</sup>

(<sup>1</sup>西北农林科技大学动物科技学院, 陕西杨凌 712100; <sup>2</sup>国家肉牛改良中心, 陕西杨凌 712100)

**摘要:**【目的】分析牛乳腺中 7 种主要乳蛋白基因密码子使用偏好性, 筛选高频密码子和低频密码子, 依据其对人溶菌酶基因进行密码子局部优化和全局优化, 通过牛乳腺上皮细胞等多种细胞对优化效果进行分析评价, 为提高重组人溶菌酶表达量, 开发新型、高效、安全的重组人溶菌酶提供理论依据。【方法】利用 CodonW 和 EMBOS 等软件对 7 种主要牛乳蛋白和人溶菌酶基因密码子偏好进行生物信息学分析, 筛选牛乳蛋白基因高频、低频密码子并根据牛乳蛋白基因密码子使用偏好对人溶菌酶基因翻译起始区前 22 位密码子(LYZop22)和全局密码子(LYZop)分别进行优化。构建人溶菌酶-荧光素酶融合表达载体(pGL3-LYZcw/op22/op)和人溶菌酶过表达载体(pcDNA-LYZcw/op22/op), 将上述载体分别转染牛乳腺上皮细胞(BMEC)、牛成纤维细胞(BFFC)和 C127 小鼠乳腺上皮细胞等 3 种细胞, 通过荧光素酶、实时荧光定量 PCR 及 Western-blot 等方法检测密码子优化对溶菌酶表达的影响。【结果】牛乳蛋白基因密码子偏好以 GC 结尾, GC3s 平均含量为  $0.537 \pm 0.062$ , 而人溶菌酶 GC3s 含量为 0.407, 偏好以 AT 结尾; 聚类结果表明, 牛乳中酪蛋白与乳清蛋白类基因在密码子使用偏好性上也存在一定差异。依据筛选获得的牛主要乳蛋白基因 5 个高频密码子(RSCU > 1.5)和 7 个低频密码子(RSCU < 0.5)对人溶菌酶基因密码子进行优化并转染多种细胞进行表达效果分析。荧光素酶分析发现, 相比野生型溶菌酶密码子(LYZcw), 翻译起始区密码子优化类型 LYZop22 分别在 BMEC、BFFC 细胞中提高 1.48 倍( $P < 0.01$ )和 1.30 倍( $P > 0.05$ ); 而全局密码子优化类型 LYZop 分别在 BMEC、BFFC 中提高 2.2 倍( $P < 0.01$ )和 2.44 倍( $P < 0.01$ ), 说明密码子优化能明显提高人溶菌酶在多种细胞中表达量。实时荧光定量 PCR 结果表明, LYZop22 相比 LYZcw 在 BMEC 和 BFFC 细胞中分别提高了 2.08 倍( $P < 0.05$ )和 1.5 倍( $P > 0.05$ ), 而 LYZop 则分别提高了 22 倍( $P < 0.01$ )和 17.8 倍( $P < 0.01$ ), mRNA 表达水平与密码子优化后的 mRNA 二级结构稳定性呈正相关。Western-blot 结果也进一步表明, 密码子优化后的 LYZop22 和 LYZop 能明显提高重组人溶菌酶在牛乳腺上皮细胞中的表达量。上述结果表明, 根据牛乳腺中主要乳蛋白基因密码子使用偏好进行密码子优化, 能显著提高人溶菌酶在牛乳腺上皮细胞及成纤维细胞中的表达量, 且溶菌酶基因全局密码子优化效果优于翻译起始区密码子优化效果。【结论】通过生物信息学分析获得了牛乳蛋白基因使密码子使用偏好及高低频密码子; 依据牛乳蛋白密码子使用偏好性对人溶菌酶密码子优化能显著提高重组人溶菌酶 mRNA 水平和蛋白水平的表达量, 为今后利用生物反应器高效生产重组人溶菌酶奠定基础。

**关键词:** 人溶菌酶; 牛乳蛋白基因; 密码子优化; 荧光素酶检测; 实时荧光定量 PCR; Western-blot

## Codon Optimization of Human Lysozyme and High-Efficiency Expression in Bovine Mammary Cells

TIAN Yuan<sup>1</sup>, WANG Li<sup>1</sup>, LONG Feng<sup>1</sup>, ZAN LinSen<sup>1,2</sup>, CHENG Gong<sup>1,2</sup>

(<sup>1</sup> College of Animal Science and Technology, Northwest A&F University, Yangling 712100, Shaanxi; <sup>2</sup> National Beef Cattle Improvement Center, Yangling 712100, Shaanxi)

收稿日期: 2019-09-18; 接受日期: 2020-05-29

基金项目: 陕西省重点研发计划(2019NY-098)、国家重点研发计划(2018YFD0501700)

联系方式: 田媛, E-mail: 17319512949@163.com。通信作者成功, E-mail: chenggong@nwfau.edu.cn

**Abstract:** 【Objective】The aims of present study were to investigate the codon usage bias of seven milk protein genes in cattle mammary gland, to optimize human lysozyme gene partially and totally based on codon usage bias of milk protein genes, and to evaluate the codon optimized effect of human lysozyme in multiple cells, so as to provide theoretical basis for increasing recombinant human lysozyme expression and developing new high-effect and safe recombinant human lysozyme. 【Method】The codon usage bias and high/low frequency codons of bovine major milk protein genes were confirmed through bioinformatics analysis by CodonW and EMBOSS software. Human lysozyme gene was partially (22 codons in the translation start region, named LYZop22) and totally (named LYZop) optimized according to the codon usage frequency of bovine major milk protein genes. The human lysozyme-luciferase fusion expression vector (pGL3-LYZcw/op22/op) and overexpression vector (pcDNA-LYZcw/op22/op) were constructed and transfected to cattle mammary epithelial cells (BMEC), cattle fibroblasts (BFFC) and mouse mammary epithelial cells for codon optimization effect evaluation by luciferase, real-time qPCR and western blot methods. 【Result】The cattle milk protein genes prefer GC end, which the GC3s content was  $0.537 \pm 0.062$ . However, the human lysozyme gene prefer AT end, and the GC3s content was 0.407. Moreover, differences in codon usage bias between casein protein genes and whey protein genes were found by cluster analysis. Five high frequency codons ( $RSCU > 1.5$ ) and seven low frequency codons ( $RSCU < 0.5$ ) were found in the bovine milk protein genes. The human lysozyme gene was optimized according to the bovine milk protein codon usage bias and evaluated by multiple cells transfection. The luciferase result showed that, the codon optimization could significantly increase the human lysozyme expression, which were 1.48-fold ( $P < 0.01$ ) and 1.30-fold ( $P > 0.05$ ) increased in BMEC and BFFC of LYZop22, respectively. While, 2.2-fold and 2.44-fold increase in the BMEC and BFFC of LYZop, respectively. Real-time qPCR result showed that, the mRNA levels of human lysozyme were increased in BMEC (2.08-fold,  $P < 0.05$ ) and BFFC (1.5-fold,  $P > 0.05$ ) of LYZop22. While the significantly increase in mRNA level of human lysozyme were found in the BMEC (17.8-fold,  $P < 0.01$ ) and BFFC (22-fold,  $P < 0.01$ ) of LYZop, the totally codon optimized type. In addition, there was a positive correlation between mRNA level and mRNA second structure stability. Western-blot result showed that, the expression of human lysozyme was effectively increased both in partially and totally codon optimized human lysozyme gene in the BMEC. Moreover, the effect on human lysozyme expression in totally codon optimization type was better than partially codon optimization type. 【Conclusion】The codon usage bias and high/low frequency codons of bovine major milk protein genes were obtained, and the codon optimization of human lysozyme gene according bovine major milk protein gene codon usage bias could significantly improve the mRNA and protein level of human lysozyme, which would lay the foundation for producing recombinant human lysozyme effectively by bioreactor.

**Key words:** human lysozyme; cattle milk protein genes; codon optimization; luciferase analysis; real-time qPCR; Western-blot

## 0 引言

【研究意义】抗生素滥用所引发的细菌耐药性、抗生素残留及食品安全等问题已被视为全球危害公众健康的问题,人们迫切需要开发新型广谱抗菌剂。溶菌酶又称胞壁质酶或N-乙酰胞壁质聚糖水解酶,在唾液、泪液、血清、人乳、牛乳和禽类蛋清中广泛分布,是一种能水解致病菌中黏多糖的碱性酶,具有广谱抗菌、消炎、抗病毒及免疫调节等多种功效,在畜牧、农产品、食品和医疗等行业广泛应用<sup>[1-3]</sup>。人溶菌酶和鸡蛋清溶菌酶氨基酸序列具有59%同源性,但抗菌活性远高于目前使用较广的鸡蛋清溶菌酶,具有更广阔的应用前景<sup>[4]</sup>。然而,通过传统方法从母乳、胎盘、唾液中提取分离或微生物表达系统生产重组人溶菌酶,获得量很少,稳定性低,无法满足研究和潜在市场应用的需求<sup>[5]</sup>。基因工程不断发展为大规模高效生产重组人溶菌酶提供了可能。因此,进一步优化表达

系统,通过生物反应器实现重组人溶菌酶高效表达,将为今后重组人溶菌酶的工业化生产和应用奠定基础。【前人研究进展】2006年,MEGA首次利用山羊乳腺生物反应器生产了重组人溶菌酶<sup>[6]</sup>,随后在牛<sup>[7-9]</sup>、猪<sup>[10]</sup>、山羊<sup>[11]</sup>等动物的研究中均有成功实现了重组人溶菌酶在乳腺中表达的报道。然而,不容忽视的是,异源重组蛋白表达量低是目前生物反应器中亟待解决的难题。研究发现,宿主细胞密码子使用偏好很大程度上影响了异源重组蛋白在宿主细胞的高效表达<sup>[12]</sup>。组成蛋白质的氨基酸有20种,而编码氨基酸的密码子有64种(简并密码子)。不同生物,甚至同种生物不同的蛋白质编码基因,对简并密码子使用频率并不相同,具有一定的偏爱性,其中使用频率高密码子称为最佳密码子,而那些很少被利用的密码子称为稀有或低频密码子。研究发现,少量的DNA密码子的改变,在很大程度上影响了蛋白翻译速率,即使替换单个简并密码子,其蛋白合成的速度会减缓到

其正常速度的 1/10 甚至更慢<sup>[13]</sup>，密码子使用偏好与蛋白翻译效率呈显著正相关<sup>[14]</sup>。通过蛋白质和 mRNA 表达数据库联合分析发现，密码子的选择决定了蛋白的表达量，生物体通过密码子的选择进而调控内源蛋白的表达丰度<sup>[15]</sup>。密码子的选择在生物反应器研究过程中十分重要，外源基因含有宿主细胞基因使用的低频密码子或者偏好密码子不一致情况下，往往会影响外源基因的 mRNA 的稳定性和翻译效率，造成重组蛋白表达量低<sup>[16]</sup>。前人研究结果比较发现，转基因山羊、牛乳腺中重组人乳铁蛋白表达量为 1.5—30 g·L<sup>-1</sup>，而重组人溶菌酶的表达量只有 0.025—0.27 g·L<sup>-1</sup><sup>[17]</sup>，表达量低造成提纯成本高，难以达到乳腺生物反应器工业化生产要求。【本研究切入点】因此，人溶菌酶基因密码子使用偏好性是否与牛主要乳蛋白密码子使用偏好一致？这种差异是否是导致人溶菌酶在乳腺生物反应器中表达量低的主要原因？密码子优化能否进一步提高重组人溶菌酶表达量？这些问题仍待进一步研究。【拟解决的关键问题】本试验通过对牛乳腺主要乳蛋白密码子使用偏好性进行分析，并根据牛乳蛋白高频密码子和低频密码子对人溶菌酶基因密码子进行翻译起始区优化和全局优化设计，通过荧光素酶、实时荧光定量 PCR 及 Western-blot 等方法从 mRNA 和蛋白水平对密码子优化效果进行分析比较，实现人溶菌酶基因的高效表达，为开展重组人溶菌酶的工业化生产及应用奠定基础。

1 材料与方法

试验于 2019 年在西北农林科技大学国家肉牛改良中心实验室完成。

1.1 材料

PGL3-Basic 荧光素酶载体、海肾荧光素酶载体为实验室保存。*E.coli* DH 5a 菌株和 pMD19-T easy 载体购自大连 Takara 公司。牛成纤维细胞通过耳组织块法分离获得，实验室冻存。牛乳腺上皮细胞系购自通派（上海）生物科技有限公司，小鼠乳腺上皮细胞系 C127 购自北京协和细胞库。主要试剂包括：高保真 PCR 酶、限制性内切酶、PrimeScript™ RT reagent Kit with gDNA Eraser 反转录试剂盒和 SYBR Premix EX Taq 定量 PCR 试剂盒购自宝生物工程（大连）有限公司；T4 DNA 连接酶、Dual-Glo 双荧光素酶检测试剂盒购自普洛麦格（北京）生物技术有限公司；Endo-free Plasmid Kit、胶回收试剂盒购自 Omega Biotek 股份有限公司；Lipo3000 转染试剂及 opti-MEM 购自

Invitrogen 公司；Western-blot 细胞裂解液、BCA 蛋白浓度测定试剂盒及 SDS-PAGE 蛋白上样缓冲液购自上海碧云天生物技术有限公司；蛋白酶抑制剂购自罗氏公司。兔源溶菌酶多抗、兔源 GAPDH 单抗和 HRP 标记的羊抗兔二抗分别购自 Santa Cruz（sc-292849），Abcam（ab181603）和生工生物工程（上海）股份有限公司（D110058）。

1.2 牛主要乳蛋白基因和人溶菌酶基因密码子使用偏好分析

选取  $\alpha$ s1、 $\alpha$ s2、 $\beta$ 、 $\kappa$ -酪蛋白， $\beta$ -乳球蛋白， $\alpha$ -乳清白蛋白，乳铁蛋白等 7 种牛乳中主要乳蛋白基因 CDS 序列，用于密码子使用偏好性分析。具体乳蛋白基因名称、GenBank 登录号、CDS 序列长度见表 1。

表 1 7 种牛乳蛋白基因 GenBank 登录号及序列信息  
Table 1 GenBank accession number and sequence information of cattle seven milk protein genes

基因名称 Gene name	GenBank 号 GenBank No.	序列长度 Length of CDS
<i>alpha-s1 casein (CSN1S1)</i>	M33123	645 bp
<i>alpha-s2 casein (CSN1S2)</i>	BC114773	669 bp
<i>beta casein (CSN2)</i>	NM_181008	675 bp
<i>kappa casein (CSN3)</i>	BC102120	573 bp
<i>beta-lactoglobulin (LGB)</i>	NM_173929	537 bp
<i>alpha-lactalbumin (LALBA)</i>	M18780	429 bp
<i>lactotransferrin (LTF)</i>	NM_180998	2127 bp

利用 CodonW1.42 和 EMBOSS (<http://imed.med.ucm.es/EMBOSS/>) 中 CUSP、CHiPs 等工具对牛主要乳蛋白基因密码子及人溶菌酶密码子使用特性进行分析。分析的特性参数主要包括：A3s，G3s，C3s，同义密码子在第 3 位上相应碱基的出现频率（T3s）、基因的 G+C 含量（GC）、密码子第 3 位的 G+C 含量（GC3s）、密码子适应性指数（CAI）、有效密码子数（ENc）等。同义密码子的相对使用度（RSCU）是对同义密码子的使用偏好性评估，该值等于同义密码子的实际观测值与同义密码子平均使用期望值的比值。如果密码子使用无偏好性，则 RSCU 值为 1；如该密码子比其他同义密码子使用更频繁，则其 RSCU 值大于 1，反之亦然。将 7 个牛乳蛋白基因和人溶菌酶基因各自作为一个对象，以密码子的 RSCU 值作为变量，对应每个编码氨基酸密码子利用 TBtools 0.665 绘制热图并进行聚类分析。

1.3 溶菌酶密码子优化及基因克隆

根据牛主要乳蛋白基因密码子使用偏好性, 对人溶菌酶基因局部(翻译起始区前 22 位密码子)和全局优化设计。利用 mfold 在线软件对优化后的人溶菌酶基因序列进行 RNA 二级结构分析 (<http://unafold.rna.albany.edu/>)。

采用 Trizol 法提取人 Hela 细胞总 RNA, 通过凝胶电泳和核酸定量检测 RNA 质量和浓度。根据 Takara 反转录试剂盒操作说明对 RNA 进行反转录获得 cDNA。以 cDNA 为模板, 利用 P-pGL3-cw 引物(表 2)克隆获得人溶菌酶编码区序列(野生型溶菌

酶序列, *LYZcw*)。反应条件为: 98 ℃, 30 s 模板变性; 98℃, 5 s; 57℃, 15 s; 72℃, 15 s, 30 个循环。72℃, 5 min 延伸。以 *LYZcw* cDNA 为模板, 利用 P-pGL3-op22 引物(表 2)克隆获得翻译起始区密码子优化序列(*LYZop22*), PCR 反应条件同上。对上述获得的 *LYZcw*、*LYZop22* PCR 产物经过胶回收纯化、序列末端加 A 反应、T-A 克隆并 DH5α 转化后进行测序验证(*LYZcw-T*、*LYZop22-T*)。全局密码子优化的溶菌酶序列(*LYZop*), 经生工生物工程(上海)股份有限公司全基因合成并克隆到 T 载体中(*LYZop-T*)。

表 2 *LYZcw*、*LYZop22*、*LYZop* 序列克隆引物信息及基因合成  
Table 2 Primers for the amplication of *LYZcw*, *LYZop22* and *LYZop*

引物名称 Primer name	基因名称 Gene name	引物序列 Primer sequence	产物长度 Product length	酶切位点 Restriction site
P-pGL3-cw	<i>LYZcw</i>	F: CCATGGCCATGAAGGCTCTCATTG R: CCATGGTTACTCCACATCACTCCACAACCTTGAAC	459bp	<i>NcoI</i> / <i>Nco</i>
P-pGL3-op22	<i>LYZop22</i>	F: CCATGGACCATGAAGGCCCTGATCGTGCTGGGAC TGGTGCTGCTGTCTGTGACCGTGCAGGGCAAGGT GTTTGAGAGGTGTGAGTTGGCCAGAACTCTG R: CCATGGTCACTCCACAACCTTGAAC	459bp	<i>NcoI</i> / <i>NcoI</i>
—	<i>LYZop</i>	全基因合成 Whole gene synthesis	459 bp	<i>NcoI</i> / <i>NcoI</i>
P-pcDNA-cw	<i>LYZcw</i>	F: AAGCTTCCATGAAGGCTCTCATTG R: CTCGAGCTCCACAACCTTGAAC	456 bp	<i>HindIII</i> / <i>XhoI</i>
P-pcDNA-op22	<i>LYZop22</i>	F: CCATGGACCATGAAGGCCCTGATCG R: CTCGAGCTCCACAACCTTGAAC	456bp	<i>HindIII</i> / <i>XhoI</i>
P-pcDNA-op	<i>LYZop</i>	F: CCATGGACCATGAAGGCCCTGATCG R: CTCGAG CACTCCACATCCCTGCACAT	456 bp	<i>HindIII</i> / <i>XhoI</i>

1.4 pGL3-LYZcw/op22/op 溶菌酶-荧光素酶融合表达构建

以 Invitrogen 公司 pGL3-control 载体为骨架载体, 构建人溶菌酶基因与萤火虫荧光素酶基因融合表达载体。通过 *NcoI* 内切酶分别酶切 *LYZcw-T*、*LYZop22-T*、*LYZop-T* 和 pGL3-control 载体。1%琼脂糖凝胶电泳回收 *LYZcw*、*LYZop22*、*LYZop* 溶菌酶基因序列和线性化的 pGL3-control 载体。通过 T4 DNA 连接酶进行连接, 分别构建 pGL3-LYZcw/op22/op 荧光素酶融合表达载体。测序筛选鉴定 *LYZcw/op22/op* 序列以正向插入并与萤火虫荧光素酶基因正确融合的质粒用于下一步研究。

1.5 细胞转染及荧光素酶检测密码子优化效果

分别将实验室冻存的牛乳腺上皮细胞系(BMEC)、牛成纤维细胞(BFFC)及 C127 小鼠乳腺上皮细胞

系复苏, 用完全培养基(DMEM 高糖+10% FBS +1% PS)重悬, 接种到 T75 培养瓶中, 置于 37 ℃, 5% CO<sub>2</sub> 培养箱中培养。待细胞汇合度达到 90%时进行传代接种于 24 孔培养板中(2×10<sup>5</sup>/孔)。接种后的第二天, 按照每孔 0.8 μg 质粒(PGL-LYZcw/LYZop/LYZop22 : pRL-tk = 10 : 1) : 2 μLlipo3000 比例转染上述细胞, 每组 4 个复孔, 转染 48 h 后通过荧光素酶活性检测比较人溶菌酶密码子优化效果。

1.6 pcDNA3.1-LYZcw/op22/op 溶菌酶过表达载体构建

分别以 pGL3-LYZcw/op22/op 质粒为模板, 利用 P-pcDNA-cw、P-pcDNA-op22、P-pcDNA-op 引物(表 2)克隆 *LYZcw*、*LYZop22*、*LYZop* 序列(PCR 反应条件参照 1.3)并连接到 T 载体。使用 *HindIII*、*XhoI* 内

切酶分别酶切 pcDNA3.1 (+)、LYZcw-T、LYZop22-T vector、LYZop- T, 连接构建 pcDNA-LYZcw/op22/op 人溶菌酶过表达载体, 测序做进一步鉴定。

1.7 细胞转染及实时荧光定量 PCR 检测密码子优化效果

将培养的牛成纤维细胞 (BFFC)、牛乳腺上皮细胞 (BMEC) 及 C127 细胞传代接种于 12 孔培养板中 ( $2 \times 10^5$ /孔)。接种后的第二天, 按照 lipo3000 转染试剂说明书, 每孔 1.6  $\mu$ g 质粒 (pcDNA 空载、pcDNA-LYZcw/op22/op): 4  $\mu$ L lipo3000 比例转染上述细胞, 每组 4 个复孔。24 h 后按照 1.3 中方法提取细胞总 RNA 并反转录获得 cDNA, 用于下一步定量 PCR 检测。

对不同密码子优化方式对重组人溶菌酶 mRNA 水平表达量进行定量分析。反应条件为: 预变性 95  $^{\circ}$ C, 30 s; 95  $^{\circ}$ C, 5 s; 60  $^{\circ}$ C, 34 s, 40 个循环。循环结束后进行熔解曲线分析, 确保扩增条带单一。

1.8 Western-blot 检测密码子优化效果

将培养的牛乳腺细胞 (BMEC) 经胰酶消化后接种于 6 孔培养板中 ( $4 \times 10^5$  / 孔)。接种后的第二天, 按照 lipo3000 转染试剂说明书, 每孔 4.0  $\mu$ g 质粒 (pcDNA 空载、pcDNA-LYZcw/op22/op): 10  $\mu$ L lipo3000 比例转染上述细胞, 每组 4 个复孔。转染 48 h 后, 收集蛋白用于 Western-blot 检测。蛋白电泳采用分离胶浓度为 12%, 蛋白上样量 40  $\mu$ g。电泳条件为浓缩胶 80 V, 分离胶 120 V, 待溴酚蓝跑出胶后终止电泳。蛋白转膜按照伯乐半干转膜仪操作说明, 200

mA, 1h 进行转膜。转膜结束后, PVDF 膜经 5%脱脂奶粉封闭——一抗 4  $^{\circ}$ C 孵育过夜 (人溶菌酶、GAPDH 1 : 1 000 稀释) ——TBST 洗涤 (3 次, 每次 10 min) ——二抗室温孵育 2 h (1 : 3 000 稀释) ——TBST 洗涤 (3 次, 每次 10 min) ——显影——ChemiDoc XRS<sup>+</sup> 仪器成像。

1.9 数据分析

采用  $2^{-\Delta\Delta C_t}$  方法对 real-time qPCR 分析计算目的基因相对表达量。GraphPad Prism 6.0 进行 One-way ANOVA 分析和作图。数据均以平均数  $\pm$  标准误 (Mean $\pm$ SEM) 表示。

2 结果

2.1 牛乳蛋白及人溶菌酶基因密码子组成分析

通过 CodonW 软件对牛乳蛋白及人溶菌酶基因密码子组成进行分析, 结果表明, 牛乳蛋白基因密码子 GC 含量平均为  $0.485 \pm 0.026$ , 第三位碱基 GC3s 含量平均为  $0.537 \pm 0.062$ 。人溶菌酶基因密码子 GC 含量为 0.473, 第三位碱基 GC3s 含量为 0.407。上述结果表明, 人溶菌酶基因整体 GC 含量与牛乳蛋白基因 GC 含量相近, 但第三位碱基 GC 含量低于牛乳蛋白基因 GC 含量。牛乳蛋白基因密码子偏好以 GC 结尾, 而人溶菌酶基因密码子偏好以 AT 结尾 (表 3)。

ENc 值范围为 20 (每个氨基酸只使用一个密码子) 到 61 (各个密码子被均衡使用), 其值越低, 偏好性越强。牛乳蛋白及人溶菌酶 ENc 值分别为 49.814 和 50.27, 表明牛乳蛋白和人溶菌酶基因密码子使用相

表 3 牛乳蛋白及人溶菌酶基因密码子组成分析  
Table 3 Codon composition analysis of bovine milk proteins and human lysozyme gene

Bos taurus	A3s	T3s	C3s	G3s	GC3s	GC	CAI	ENc
$\alpha$ s1 酪蛋白 <i>alpha-s1 casein</i>	0.311	0.461	0.290	0.258	0.408	0.445	0.227	52.300
$\alpha$ s2 酪蛋白 <i>alpha-s2 casein</i>	0.253	0.453	0.340	0.335	0.484	0.417	0.317	49.970
$\beta$ 酪蛋白 <i>beta casein</i>	0.201	0.421	0.342	0.298	0.505	0.512	0.234	48.750
$\kappa$ -酪蛋白 <i>kappa casein</i>	0.399	0.412	0.243	0.194	0.341	0.430	0.240	58.250
乳球蛋白 <i>lactoglobulin</i>	0.079	0.147	0.574	0.519	0.825	0.586	0.349	36.270
$\alpha$ 乳清白蛋白 <i>alpha-lactalbumin</i>	0.283	0.349	0.395	0.337	0.526	0.441	0.281	55.510
乳铁蛋白 <i>lactotransferrin</i>	0.170	0.242	0.410	0.449	0.671	0.561	0.262	47.650
Mean $\pm$ SEM	0.242 $\pm$ 0.039	0.355 $\pm$ 0.045	0.371 $\pm$ 0.040	0.341 $\pm$ 0.042	0.537 $\pm$ 0.062	0.485 $\pm$ 0.026	0.273 $\pm$ 0.017	49.814 $\pm$ 2.668
人溶菌酶 <i>human lysozyme</i>	0.3143	0.4032	0.2742	0.2323	0.407	0.473	0.193	50.270

对较平均。ENC vs GC3s 线性回归分析结果表明, 7 种乳蛋白基因有效密码子数 ENc 值与第三位碱基 GC3s 呈显著负相关 ( $R = -0.8968$ ,  $P < 0.01$ ) (图 1), 表明密码子使用偏好性越强 (ENc 值越小) 的基因, 其 GC3s 值越高, 主要偏好以 GC 碱基结尾的密码子。因此, 牛乳蛋白基因密码子的使用偏好性受 GC3s 影响。

2.2 牛乳蛋白基因密码子使用偏好性分析

对编码氨基酸且为简并密码子的 7 种牛乳蛋白基

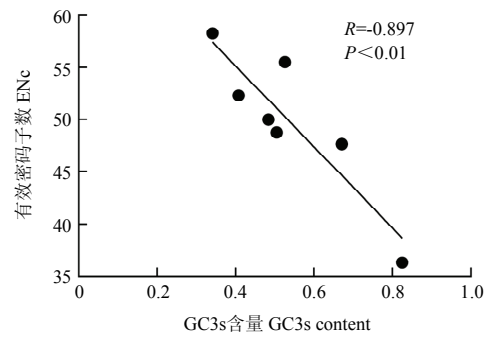
因使用的共 59 个密码子 (不含 Met、Trp 及 3 个终止密码子) 进行同义密码子的相对使用度 (RSCU) 分析。结果表明, 牛乳蛋白基因对 CTG (2.71)、AGG (2.32)、GCC (1.73)、GTG (1.68)、ATC (1.64) 等 5 个密码子具有明显的偏好性 ( $RSCU > 1.5$ ), 为高频密码子。而 TTA (0.12)、ATA (0.18)、TCG (0.19)、CTA (0.25)、CCG (0.26)、GCG (0.45)、ACG (0.48) 等 7 个密码子在牛乳蛋白基因中使用较小 ( $RSCU < 0.5$ ), 为低频密码子 (表 4)。

表 4 牛乳蛋白基因密码子使用偏好性分析

Table 4 Codon preference analysis of bovine milk protein genes

密码子	氨基酸	频率/1000	个数	同义密码子	密码子	氨基酸	频率/1000	个数	同义密码子
Codon	Amino acid	Frequency/1000	Number	相对使用度 RSCU	Codon	Amino acid	Frequency/1000	Number	相对使用度 RSCU
GCA	A	13.793	26	0.73	AAT	N	19.629	37	0.96
<b>GCC</b>	A	32.891	62	1.73	CCA	P	17.507	33	1.06
<u>GCG</u>	A	8.488	16	0.45	CCC	P	19.629	37	1.19
GCT	A	20.69	39	1.09	<u>CCG</u>	P	4.244	8	0.26
TGC	C	14.854	28	0.98	CCT	P	24.403	46	1.48
TGT	C	15.385	29	1.02	CAA	Q	16.976	32	0.57
GAC	D	22.281	42	1.09	CAG	Q	42.44	80	1.43
GAT	D	18.568	35	0.91	AGA	R	5.836	11	1.06
GAA	E	33.952	64	0.90	<u>AGG</u>	R	12.732	24	2.32
GAG	E	41.91	79	1.10	CGA	R	2.122	4	0.39
TTC	F	21.22	40	1.13	CGC	R	3.183	6	0.58
TTT	F	16.446	31	0.87	CGG	R	6.897	13	1.26
GGA	G	11.671	22	1.07	CGT	R	2.122	4	0.39
GGC	G	15.385	29	1.41	AGC	S	14.324	27	1.30
GGG	G	9.019	17	0.83	AGT	S	13.263	25	1.20
GGT	G	7.427	14	0.68	TCA	S	7.427	14	0.67
CAC	H	11.141	21	1.35	TCC	S	14.854	28	1.34
CAT	H	5.305	10	0.65	<u>TCG</u>	S	2.122	4	0.19
<u>ATA</u>	I	2.653	5	0.18	TCT	S	14.324	27	1.30
<b>ATC</b>	I	24.403	46	1.64	ACA	T	11.141	21	0.84
ATT	I	17.507	33	1.18	ACC	T	18.037	34	1.36
AAA	K	33.952	64	0.88	<u>ACG</u>	T	6.366	12	0.48
AAG	K	42.971	81	1.12	ACT	T	17.507	33	1.32
<u>CTA</u>	L	4.244	8	0.25	GTA	V	6.897	13	0.41
CTC	L	19.629	37	1.14	GTC	V	15.915	30	0.95
<b>CTG</b>	L	46.684	88	2.71	<b>GTG</b>	V	28.117	53	1.68
CTT	L	18.037	34	1.05	GTT	V	15.915	30	0.95
<u>TTA</u>	L	2.122	4	0.12	TAC	Y	15.385	29	0.91
TTG	L	12.732	24	0.74	TAT	Y	18.568	35	1.09
AAC	N	21.22	40	1.04					

方框表示牛乳蛋白基因中高频密码子, 下划线表示牛乳蛋白基因中低频密码子  
Box shows the high frequency codons of milk protein genes, underline shows the low frequency codons



圆点表示不同乳蛋白基因 GC3s 值对应的 ENc 值  
The dot represents the ENc value corresponding to the GC3s value of different milk protein genes

图 1 牛乳蛋白基因 ENc vs GC3s 相关性分析  
Fig. 1 Correlation analysis of ENc vs GC3s of bovine milk protein genes

表 5 人溶菌酶基因密码子优化情况  
Table 5 Codon Optimization of human lysozyme gene

LYZcw		LYZop22		LYZop	
1	ATG AAG GCT CTC ATT GTT	1	ATG AAG <u>GCC CTG ATC GTG</u>	1	ATG AAG <u>GCC CTG ATC GTG</u>
19	CTG GGG CTT GTC CTC CTT	19	CTG <u>GGA CTG GTG CTG CTG</u>	19	CTG <u>GGA CTG GTG CTG CTG</u>
37	TCT GTT ACG GTC CAG GGC	37	TCT <u>GTG ACC GTG</u> CAG GGC	37	TCT <u>GTG ACC GTG</u> CAG GGC
55	AAG GTC TTT GAA AGG TGT	55	AAG <u>GTG</u> TTT <u>GAG</u> AGG TGT	55	AAG <u>GTG</u> TTT <u>GAG</u> AGG <u>TGC</u>
73	GAG TTG GCC AGA ACT CTG	73	GAG TTG GCC AGA ACT CTG	73	GAG <u>CTG</u> GCC <u>AGG</u> ACT CTG
91	AAA AGA TTG GGA ATG GAT	91	AAA AGA TTG GGA ATG GAT	91	<u>AAG AGG CTG</u> GGA ATG GAT
109	GGC TAC AGG GGA ATC AGC	109	GGC TAC AGG GGA ATC AGC	109	<u>GGC TAT</u> AGG GGA ATC AGC
127	CTA GCA AAC TGG ATG TGT	127	CTA GCA AAC TGG ATG TGT	127	<u>CTG GCC</u> AAC TGG ATG TGT
145	TTG GCC AAA TGG GAG AGT	145	TTG GCC AAA TGG GAG AGT	145	<u>CTG GCC AAG TGG GAG TCT</u>
163	GGT TAC AAC ACA CGA GCT	163	GGT TAC AAC ACA CGA GCT	163	<u>GGC TAT</u> AAC <u>ACC AGG GCC</u>
181	ACA AAC TAC AAT GCT GGA	181	ACA AAC TAC AAT GCT GGA	181	<u>ACC</u> AAC <u>TAT AAC GCC</u> GGA
199	GAC AGA AGC ACT GAT TAT	199	GAC AGA AGC ACT GAT TAT	199	GAC <u>AGG</u> AGC <u>ACC GAC</u> TAT
217	GGG ATA TTT CAG ATC AAT	217	GGG ATA TTT CAG ATC AAT	217	<u>GGC ATC</u> TTT CAG ATC <u>AAC</u>
235	AGC CGC TAC TGG TGT AAT	235	AGC CGC TAC TGG TGT AAT	235	<u>TCT AGG TAT</u> TGG TGT <u>AAC</u>
253	GAT GGC AAA ACC CCA GGA	253	GAT GGC AAA ACC CCA GGA	253	<u>GAC</u> GGC <u>AAG</u> ACC <u>CCT GGC</u>
271	GCA GTT AAT GCC TGT CAT	271	GCA GTT AAT GCC TGT CAT	271	<u>GCC GTG</u> AAT GCC TGT <u>CAC</u>
289	TTA TCC TGC AGT GCT TTG	289	TTA TCC TGC AGT GCT TTG	289	<u>CTG TCT TGT AGC GCC CTG</u>
307	CTG CAA GAT AAC ATC GCT	307	CTG CAA GAT AAC ATC GCT	307	CTG <u>CAG GAC</u> AAC ATC <u>GCC</u>
325	GAT GCT GTA GCT TGT GCA	325	GAT GCT GTA GCT TGT GCA	325	GAT GCT <u>GTG GCC TGC GCC</u>
343	AAG AGG GTT GTC CGT GAT	343	AAG AGG GTT GTC CGT GAT	343	AAG AGG <u>GTG GTG AGG</u> GAT
361	CCA CAA GGC ATT AGA GCA	361	CCA CAA GGC ATT AGA GCA	361	<u>CCT CAG GGA ATC AGG GCT</u>
379	TGG GTG GCA TGG AGA AAT	379	TGG GTG GCA TGG AGA AAT	379	TGG GTG <u>GCC</u> TGG <u>AGG</u> AAT
397	CGT TGT CAA AAC AGA GAT	397	CGT TGT CAA AAC AGA GAT	397	<u>AGG TGC CAG</u> AAC <u>AGG GAC</u>
415	GTC CGT CAG TAT GTT CAA	415	GTC CGT CAG TAT GTT CAA	415	<u>GTG AGG</u> CAG TAT <u>GTG CAG</u>
433	GGT TGT GGA GTG TAA	433	GGT TGT GGA GTG TAA	433	<u>GGA</u> TGT GGA GTG TAA

下划线表示优化后的密码子 Underline shows the codons which have been optimized

2.3 基因密码子偏好性的聚类分析

以密码子的 RSCU 值作为变量, 对具有简并性的 59 个编码氨基酸的密码子进行聚类分析。依据 RSCU 值可以将 7 种牛乳蛋白较好的分成为两大类: 酪蛋白类 (CSN2、CSN1S2、CSN3 和 CSN1S1) 和乳清蛋白类 (LTF、LGB 和 LALBA), 而人溶菌酶基因密码子使用偏好与牛乳酪蛋白类更接近 (图 2)。结果表明, 即使在同一乳腺组织内, 编码牛乳酪蛋白类基因密码子和乳清蛋白类基因密码子在使用偏好性上可能也存在一定差异。

2.4 人溶菌酶基因密码子优化及分析

根据表 3 结果中牛乳蛋白基因高频密码子及低频密码子使用情况, 对人溶菌酶密码子进行了优化。其中 LYZop 对溶菌酶全局密码子进行了优化, 而 LYZop22 只对翻译起始区前 22 位密码子进行了优化 (表 5)。



## 2.5 人溶菌酶-荧光素融合载体构建及溶菌酶表达分析

电泳和测序结果表明, 成功克隆到了人溶菌酶野生型LYZcw、翻译起始区密码子优化的LYZop22 CDS区序列及全基因合成的全局密码子优化的LYZop CDS区序列, 且与预期序列信息一致(图4)。NcoI酶切和测序结果表明, 成功构建了pGL3-LYZcw/op22/op 3个人溶菌酶-荧光素酶融合表达载体, 且LYZcw、LYZop22、LYZop编码区与萤火虫荧光素酶基因正确融合, 没有移码突变(图5)。

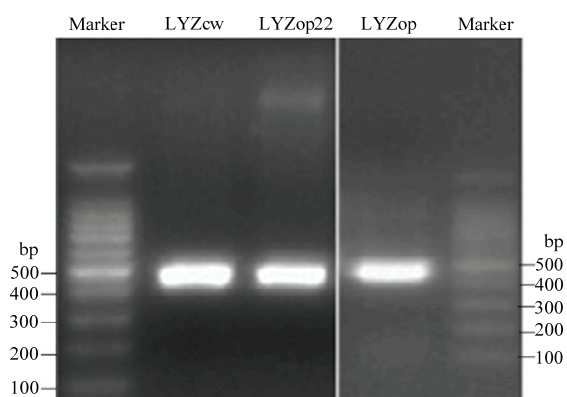


图4 LYZcw、LYZop22和LYZop电泳检测

Fig. 4 LYZcw, LYZop22 and LYZop electrophoresis detection

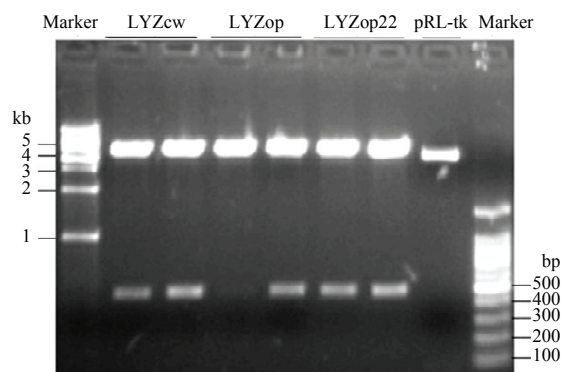


图5 pGL3-LYZcw/op22/op载体NcoI酶切鉴定

Fig. 5 pGL3-LYZcw/op22/op vector identified by NcoI enzyme digestion

通过lipo3000将pGL-LYZcw+pRL-tk、pGL-LYZop+pRL-tk、pGL-LYZop22+pRL-tk分别转染BMEC、BFFC、C127细胞, 48 h后检测荧光素酶表达量。荧光素酶检测结果表明, 相比LYZcw, 全局密码子优化的LYZop在BMEC和对照BFFC、C127细胞中分别

提高了2.20倍( $P<0.01$ )、2.44倍( $P<0.01$ )和1.99倍( $P<0.01$ )。翻译起始区密码子优化的LYZop22在BMEC细胞中相比野生型提高了1.48倍( $P<0.01$ ), 而在对照BFFC、C127细胞中差异不显著( $P>0.05$ )(图6)。

## 2.6 mRNA、蛋白水平检测人溶菌酶基因密码子优化效果

构建pcDNA-LYZcw/op22/op过表达载体, 并分别转染BMEC、BFFC和C127细胞, 检测密码子优化对人溶菌酶mRNA表达量的影响。实时荧光定量PCR结果表明, LYZop22相比LYZcw在BMEC和BFFC细胞中分别提高了2.08倍( $P<0.05$ )和1.5倍( $P>0.05$ )。而LYZop相比野生型LYZcw在上述两个细胞中分别提高了17.8倍( $P<0.01$ )和22倍( $P<0.01$ )(图7)。上述结果表明, 根据牛乳腺主要蛋白进行密码子优化后的溶菌酶能显著提高其在牛乳腺上皮细胞和成纤维细胞转录水平表达量。

进一步对pcDNA-LYZcw/op22/op过表达载体转染的牛BMEC细胞进行人溶菌酶蛋白表达检测。相比野生型LYZcw, 密码子优化的LYZop、LYZop22明显提高了人溶菌酶在牛BMEC细胞中的表达量, 且全局密码子优化的LYZop溶菌酶表达量高于翻译起始区密码子优化的LYZop22(图8-A)。上述结果表明, 密码子优化能明显提高人溶菌酶在牛乳腺上皮细胞中的表达。然而, 经全局密码子优化的人溶菌酶基因在牛乳腺上皮细胞中蛋白表达水平提高倍数低于mRNA水平提高倍数(图8-B)。

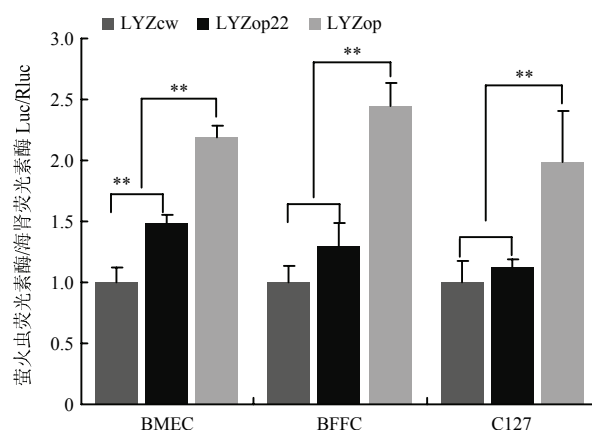


图6 荧光素酶检测人溶菌酶基因密码子优化效果

Fig. 6 The effect of codon optimization analysis of human lysozyme gene by luciferase

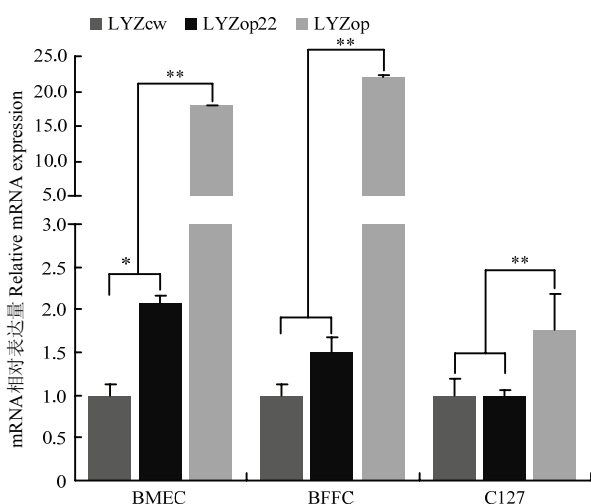


图 7 mRNA 水平检测人溶菌酶基因密码子优化效果

Fig. 7 The effect of codon optimization analysis on human mRNA level of human lysozyme gene

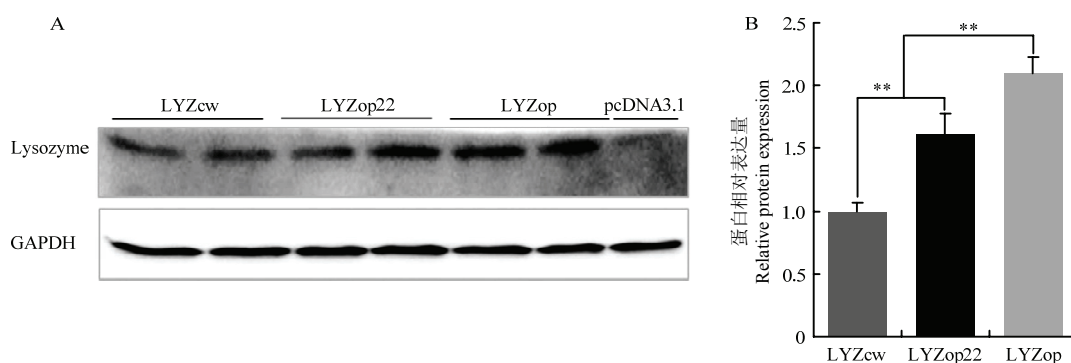


图 8 蛋白水平检测密码子优化对人溶菌酶表达量的影响

Fig. 8 The effect of codon optimization analysis on protein level of human lysozyme

使用偏好性均存在差异,正是这种差异可能进一步调节着蛋白的表达量<sup>[18]</sup>。本研究中根据牛乳中 7 种主要乳蛋白基因密码子偏好性进行聚类分析发现,牛乳中酪蛋白类(CSN2、CSN1S2、CSN3 和 CSN1S1)和乳清蛋白类(LTF、LGB 和 LALBA)可以明显的分为两类。表明,同一组织内牛乳酪蛋白和乳清蛋白基因密码子使用偏好存在一定的差异,而这种密码子使用偏好性可能在一定程度上影响着牛乳酪蛋白和乳清蛋白表达量。密码子使用偏好性分析发现,牛乳蛋白基因中发现 5 个高频密码子和 7 个低频密码子,且牛乳蛋白基因密码子第三位碱基偏好以 GC 结尾,而人溶菌酶偏好以 AT 结尾。而较高的 GC 含量,特别是

### 3 讨论

人溶菌酶作为非特异性免疫因子,对金黄色葡萄球菌、链球菌、大肠杆菌等均具有广谱的抗菌作用,在体内发挥着杀菌消炎、免疫调节、改善胃肠道菌群的作用。同时,在畜牧、农产品、食品、医疗和化妆品轻工等行业也具有广泛应用前景。利用现代基因工程和乳腺生物反应器生产重组人溶菌酶克服了传统提取方法蛋白稳定性差、成本高的不足,但如何进一步提高重组蛋白的表达量并用于下一步工业化生产仍是乳腺生物反应器首要考虑的问题。

近年来研究发现,不同生物密码子的选择不仅对蛋白高水平表达发挥作用,同时对微量的调控蛋白表达也起到调节作用<sup>[13]</sup>,这一发现对于提高外源重组蛋白的表达和生产具有重要的指导意义。研究发现,同一物种,不同组织甚至同一组织内不同蛋白其密码子

简并碱基 GC 含量在哺乳动物细胞中与 mRNA 丰度和蛋白表达量呈正相关<sup>[19-20]</sup>。表明,人溶菌酶密码子使用偏好性不利于在牛乳中高效表达,这也可能是造成牛乳腺生物反应器中重组人溶菌酶表达量低的主要原因<sup>[17, 21]</sup>。

密码子使用偏好性在调节 mRNA 稳定性、蛋白翻译起始、延伸等方面发挥了重要的作用<sup>[22-23]</sup>。本研究中局部和全局密码子优化后的人溶菌酶 mRNA 二级结构自由能 MFE 分别由 -138.83 kcal/mol 降低到了 -141.02 kcal/mol 和 -178.12 kcal/mol。前人研究发现, mRNA 折叠强度与 mRNA 丰度和蛋白表达量呈正相关<sup>[24-25]</sup>。本研究发现,随着密码子优化后人溶菌酶

mRNA 二级结构稳定性的增加,其 mRNA 表达量也相应提高。在牛 BMEC 等 3 种细胞中,全局密码子优化的人溶菌酶 mRNA 水平显著高于局部密码子优化的人溶菌酶,这一结果也与优化后 mRNA 稳定性呈正相关。表明,密码子优化可能通过增加 mRNA 二级结构稳定性和 mRNA 转录效率进而提高了人溶菌酶 mRNA 表达水平。

研究发现,mRNA 二级结构 5'端的茎环结构有利于蛋白翻译起始和延伸,进而提高蛋白表达量<sup>[26-27]</sup>。密码子优化后的溶菌酶尽管整体 mRNA MFE 明显降低,但进一步对 5'端翻译起始后 100 bp mRNA 碱基进行二级结构分析发现,密码子优化并没有明显改变人溶菌酶基因前 100bp 二级结构自由能,反而在翻译起始区引入了更多的茎环结构,而 5'端茎环结构的引入,有利于提高蛋白翻译起始效率<sup>[28]</sup>。酿酒酵母中研究发现,蛋白表达丰度与 mRNA 折叠强度存在正相关,而有效的 mRNA 翻译起始效率和较高的 mRNA 二级结构缩短了核糖体之间的距离,使 mRNA 的翻译效率提高,进而提高了蛋白的表达丰度<sup>[25, 29]</sup>。本研究中也发现,翻译区密码子优化后牛 BMEC 和 BFFC 细胞中人溶菌酶基因 mRNA 水平和蛋白水平相比野生型都有一定程度的提高。全局密码子优化的人溶菌酶基因,在 mRNA 翻译起始区引入了更多的茎环结构,同时, mRNA 整体折叠也更稳定。因此,密码子优化后的人溶菌酶表达量在 mRNA 水平和蛋白水平相比野生型均得到明显的提高。

密码子优化是提高外源重组蛋白表达的重要途径之一。尽管目前已开发许多软件可根据宿主密码子使用偏好对目的基因密码子进行优化和设计,然而关于密码子优化的“金标准”仍未统一,密码子优化的效果也不尽如人意,归根结底是由于目前关于密码子优化是如何影响外源蛋白表达仍存在很大的争议。密码子优化在 mRNA 转录效率<sup>[30]</sup>、稳定性<sup>[31]</sup>,蛋白翻译起始<sup>[32]</sup>、延伸效率、蛋白折叠及蛋白稳定性<sup>[33]</sup>等多个环节均发挥了调节作用。ZHOU 等研究发现,密码子优化提高了目的基因蛋白和 RNA 表达水平,但主要是体现在 mRNA 转录水平的提高<sup>[30]</sup>。PRESNYAK 等<sup>[31]</sup>认为,密码子优化是决定 mRNA 稳定性的主要因素。本研究中发现,经全局密码子优化的人溶菌酶基因 mRNA 二级结构稳定性得到明显提高,其在牛乳腺上皮细胞中转录水平提高了 17.8 倍,然而其蛋白水平提高倍数却低于 mRNA 水

平提高倍数。表明,密码子优化大幅提高了重组人溶菌酶 mRNA 的表达水平,但 mRNA 表达水平与蛋白表达水平并不一定呈强相关<sup>[34]</sup>,而 mRNA 转录后的修饰可能进一步影响了蛋白的表达。因此,通过密码子优化提高 mRNA 表达水平的同时如何进一步提高蛋白表达水平及其内在分子机制仍有待于进一步研究。

## 4 结论

本研究通过生物信息学分析了牛乳腺主要乳蛋白基因密码子使用偏好,获得了牛乳蛋白基因密码子使用偏好及使用的高低频密码子;依据牛乳蛋白基因密码子使用偏好性对人溶菌酶全局密码子优化能显著提高重组人溶菌酶 mRNA 水平和蛋白水平的表达量,为今后利用生物反应器高效生产重组人溶菌酶奠定基础。

## References

- [1] 李永富,孙震,史锋,张颖,马金生,蒋立凤. 溶菌酶对猪肉的保鲜作用. 上海农业学报, 2009, 25(04): 61-63.  
LI Y F, SUN Z, SHI F, ZHANG Y, MA J S, JIANG L F. The effect of lysozyme on keeping chilled pork fresh. *Acta Agriculturae Shanghai*, 2009, 25(04): 61-63. (in Chinese)
- [2] 朱军英,崔立,沈彦萍,施彬. 溶菌酶对仔猪腹泻的预防作用. 中国畜牧杂志, 2009, 45(03): 40-42.  
ZHU J Y, CUI L, SHEN Y P, SHI B. The preventive effect of lysozyme on diarrhea in piglets. *Chinese Journal of Animal Science*, 2009, 45(03): 40-42. (in Chinese)
- [3] RAGLAND S A, CRISS A K. From bacterial killing to immune modulation: Recent insights into the functions of lysozyme. *PLoS Pathogens*, 2017, 13(9): e1006512.
- [4] THELLIN O, ZORZI D, MELIN P, DELVENNE P, HEINEN E, ZORZI W, EL MOUALIJ B, QUATRESOOZ P. Hen egg white lysozyme is less active against *Gardnerella vaginalis* biofilm than human lysozyme. *Microbiology Research International*, 2018, 6(1): 1-6.
- [5] ERCAN D, DEMIRCI A. Recent advances for the production and recovery methods of lysozyme. *Critical Reviews in Biotechnology*, 2016, 36(6): 1078-1088.
- [6] MAGA E A, SHOEMAKER C F, ROWE J D, BONDURANT R H, ANDERSON G B, MURRAY J D. Production and processing of milk from transgenic goats expressing human lysozyme in the mammary gland. *Journal of Dairy Science*, 2006, 89(2): 518-524.

- [7] LIU X, WANG Y, TIAN Y, YU Y, GAO M, HU G, SU F, PAN S, LUO Y, GUO Z, QUAN F, ZHANG Y. Generation of mastitis resistance in cows by targeting human lysozyme gene to beta-casein locus using zinc-finger nucleases. *Proceedings of the Royal Society B Biological Sciences*, 2014, 281(1780): 20133368.
- [8] 熊显荣, 李文哲, 王丽君, 王勇胜, 苏建民, 华松, 张涌. 体细胞核移植技术生产转人溶菌酶基因(hLYZ)牛胚胎的研究. *农业生物技术学报*, 2011, 19(02): 294-301.
- XIONG X R, LI W Z, WANG L J, WANG Y S, SU J M, HUA S, ZHANG Y. Human lysozyme gene (hLYZ) transgenic bovine embryos produced by somatic cell nuclear transfer. *Journal of Agricultural Biotechnology*, 2011, 19(02): 294-301. (in Chinese)
- [9] YANG B, WANG J, TANG B, LIU Y, GUO C, YANG P, YU T, LI R, ZHAO J, ZHANG L, DAI Y, LI N. Characterization of bioactive recombinant human lysozyme expressed in milk of cloned transgenic cattle. *PLoS One*, 2011, 6(3): e17593.
- [10] LU D, LIU S, SHANG S, WU F, WEN X, LI Z, LI Y, HU X, ZHAO Y, LI Q, LI N. Production of transgenic-cloned pigs expressing large quantities of recombinant human lysozyme in milk. *PLoS One*, 2015, 10(5): e0123551.
- [11] YU H, CHEN J, LIU S, ZHANG A, XU X, WANG X, LU P, CHENG G. Large-scale production of functional human lysozyme in transgenic cloned goats. *Journal of Biotechnology*, 2013, 168(4): 676-683.
- [12] MAURO V P, CHAPPELL S A. Considerations in the use of codon optimization for recombinant protein expression. *Methods in Molecular Biology*, 2018, 1850: 275-288.
- [13] LI G W, OH E, WEISSMAN J S. The anti-Shine-Dalgarno sequence drives translational pausing and codon choice in bacteria. *Nature*, 2012, 484(7395): 538-541.
- [14] HANSON G, COLLIER J. Codon optimality, bias and usage in translation and mRNA decay. *Nature Reviews Molecular Cell Biology*, 2018, 19(1): 20-30.
- [15] CARNEIRO R L, REQUIAO R D, ROSSETTO S, DOMITROVIC T, PALHANO F L. Codon stabilization coefficient as a metric to gain insights into mRNA stability and codon bias and their relationships with translation. *Nucleic Acids Research*, 2019, 47(5): 2216-2228.
- [16] GUSTAFSSON C, GOVINDARAJAN S, MINSHULL J. Codon bias and heterologous protein expression. *Trends in Biotechnology*, 2004, 22(7): 346-353.
- [17] COOPER C A, MAGA E A, MURRAY J D. Production of human lactoferrin and lysozyme in the milk of transgenic dairy animals: past, present, and future. *Transgenic Research*, 2015, 24(4): 605-614.
- [18] PLOTKIN J B, ROBINS H, LEVINE A J. Tissue-specific codon usage and the expression of human genes. *Proceedings of the National Academy of Sciences of the United States of America*, 2004, 101(34): 12588-12591.
- [19] NEWMAN Z R, YOUNG J M, INGOLIA N T, BARTON G M. Differences in codon bias and GC content contribute to the balanced expression of TLR7 and TLR9. *Proceedings of the National Academy of Sciences*, 2016, 113(10): e1362-e1371.
- [20] KUDLA G, LIPINSKI L, CAFFIN F, HELWAK A, ZYLICZ M. High guanine and cytosine content increases mRNA levels in mammalian cells. *PLoS Biology*, 2006, 4(6): e180.
- [21] PLOTKIN J B, KUDLA G. Synonymous but not the same: the causes and consequences of codon bias. *Nature Reviews Genetics*, 2011, 12(1): 32-42.
- [22] BOEL G, LETSO R, NEELY H, PRICE W N, WONG K H, SU M, LUFF J, VALECHA M, EVERETT J K, ACTON T B, XIAO R, MONTELLONE G T, AALBERTS D P, HUNT J F. Codon influence on protein expression in *E. coli* correlates with mRNA levels. *Nature*, 2016, 529(7586): 358-363.
- [23] HANSON G, COLLIER J. Codon optimality, bias and usage in translation and mRNA decay. *Nature Reviews Molecular Cell Biology*, 2018, 19(1): 20-30.
- [24] WANG X, LI P, GUTENKUNST R N. Systematic effects of mRNA secondary structure on gene expression and molecular function in budding yeast. *BioRxiv*, 2017: 138792.
- [25] ZUR H, TULLER T. Strong association between mRNA folding strength and protein abundance in *S. cerevisiae*. *Embo Reports*, 2012, 13(3): 272-277.
- [26] JAGODNIK J, CHIARUTTINI C, GUILLIER M. Stem-loop structures within mRNA coding sequences activate translation initiation and mediate control by small regulatory RNAs. *Molecular Cell*, 2017, 68(1): 158-170.e3.
- [27] TULLER T, ZUR H. Multiple roles of the coding sequence 5' end in gene expression regulation. *Nucleic Acids Research*, 2015, 43(1): 13-28.
- [28] PAULUS M, HASLBECK M, WATZELE M. RNA stem-loop enhanced expression of previously non-expressible genes. *Nucleic*

- Acids Research*, 2004, 32(9): e78.
- [29] MAO Y, LIU H, LIU Y, TAO S. Deciphering the rules by which dynamics of mRNA secondary structure affect translation efficiency in *Saccharomyces cerevisiae*. *Nucleic Acids Research*, 2014, 42(8): 4813-4822.
- [30] ZHOU Z, DANG Y, ZHOU M, LI L, YU C, FU J, CHEN S, LIU Y. Codon usage is an important determinant of gene expression levels largely through its effects on transcription. *Proceedings of the National Academy of Sciences*, 2016, 113(41): e6117.
- [31] PRESNYAK V, ALHUSAINI N, CHEN Y H, MARTIN S, MORRIS N, KLINE N, OLSON S, WEINBERG D, BAKER K E, GRAVELEY B R, COLLIER J. Codon optimality is a major determinant of mRNA stability. *Cell*, 2015, 160(6): 1111-1124.
- [32] MITARAI N, SNEPPEN K, PEDERSEN S. Ribosome collisions and translation efficiency: optimization by codon usage and mRNA destabilization. *Journal of Molecular Biology*, 2008, 382(1): 236-245.
- [33] YU C H, DANG Y, ZHOU Z, WU C, ZHAO F, SACHS M S, LIU Y. Codon usage influences the local rate of translation elongation to regulate co-translational protein folding. *Molecular Cell*, 2015, 59(5): 744-754.
- [34] KOUSSOUNADIS A, LANGDON S P, UM I H, HARRISON D J, SMITH V A. Relationship between differentially expressed mRNA and mRNA-protein correlations in a xenograft model system. *Scientific Reports*, 2015, 5: 10775.

(责任编辑 林鉴非)