

乳腺生物反应器的表达优化策略

刘晓燕 祝振亮 史广宇 华梓宇 杨晨 张涌 刘军

(西北农林科技大学动物医学院, 杨凌 712100)

摘要: 乳腺生物反应器是指将外源基因导入动物基因组并在动物乳腺中特异性表达, 利用动物乳腺合成、分泌蛋白的功能, 在其乳汁中获得外源蛋白的技术。乳腺生物反应器凭借其高表达、低成本以及合成蛋白质的结构接近天然蛋白质等优势而被视为药用和营养蛋白生产的一次技术革新, 然而由于外源基因随机整合以及重组蛋白表达不稳定等问题极大地限制了其应用。本文结合乳腺生物反应器的发展现状, 从利用基因编辑技术、筛选合适的外源基因整合位点以及改进外源基因调控序列 3 个方面对乳腺生物反应器优化策略进行了综述, 以期为提高乳腺生物反应器生产重组蛋白的表达提供理论借鉴。

关键词: 乳腺生物反应器; 染色质开放性; 定向整合; CRISPR/Cas9; 调控序列; 靶靶位点

DOI:10.13560/j.cnki.biotech.bull.1985.2022-1150

Strategies to Optimize the Expression of Mammary Gland Bioreactor

LIU Xiao-yan ZHU Zhen-liang SHI Guang-yu HUA Zi-yu YANG Chen ZHANG Yong LIU Jun

(College of Veterinary Medicine, Northwest A&F University, Yangling 712100)

Abstract: Mammary gland bioreactor refers to the technology of importing foreign genes into animal genomes which express specifically in animal mammary gland, and using the function of synthesizing and secreting proteins of mammary glands to obtain recombinant proteins in the milk. Mammary gland bioreactor is regarded as a technological innovation for the production of medicinal protein and nutritional protein in the bio-pharmaceutical industry due to its advantages of high expression, low cost and synthesized protein being close to natural proteins in structure. However, due to the random integration of foreign genes and unstable expression of recombinant proteins, its application is largely limited. Based on the current development status of mammary gland bioreactor, this paper reviewed the optimization strategies of mammary gland bioreactor from three perspectives: using gene editing technology, screening suitable foreign gene integration sites and improving the regulatory sequence of exogenous proteins, so as to provide theoretical reference for improving the expression of recombinant protein produced by mammary gland bioreactor

Key words: mammary gland bioreactor; chromatin accessibility; site-specific integration; CRISPR/Cas9; regulatory element; target site

近年来生物制药产业发展迅猛, 市场对生物制药的极大需求表明药用重组蛋白有巨大的市场潜力^[1]。与酵母、细菌、哺乳动物细胞等几种重组蛋白生产系统相比, 乳腺生物反应器更高效、经济且可以生产复杂的具有生物活性的蛋白质, 因而具有更大的优势和前景^[2-4]。羊、牛、猪、兔、鼠等动物都已用于乳腺生物反应器的研究, 并已表达多种

重组蛋白, 如人溶菌酶、人乳铁蛋白、抗凝血酶 III (antithrombin III, Atryn)、凝血因子、人体组织型纤溶酶原激活剂 (tPA) 等 (表 1)。目前美国 GTC 公司利用山羊乳腺表达的 Atryn 和荷兰的 Pharming 公司利用兔乳腺生产的 C1 酯酶抑制剂 (Ruconest) 已在欧盟和美国获得上市批准。

然而目前大多数重组药物蛋白主要由中国仓鼠

收稿日期: 2022-09-30

基金项目: 国家自然科学基金面上项目 (32072805)

作者简介: 刘晓燕, 女, 研究方向: 动物医学; E-mail: 1710969610@nwfau.edu.cn

通讯作者: 刘军, 男, 博士, 教授, 研究方向: 动物胚胎工程; E-mail: liujun2013@nwsuaf.edu.cn

表 1 乳腺生物反应器生产重组蛋白实例

重组蛋白 Recombinant protein	物种 Species	表达含量 Expression content / ($\mu\text{g} \cdot \text{mL}^{-1}$)	时间 / 年 Time/Year	参考文献 References
溶菌酶 Lysozym	奶牛 <i>Bovine</i>	3 100	2016	[5]
人乳铁蛋白 Human lactoferrin	山羊 <i>Capra hircus</i>	3 200	2015	[6]
人血清白蛋白 HSA	奶牛 <i>Bovine</i>	2 500	2016	[7]
抗胰蛋白酶 AAT	绵羊 <i>Ovis aries</i>	650	2000	[8]
重组人纤溶酶原激活物 Recombinant human plasminogen activator	兔 <i>Leporidae</i>	15.2–630	2016	[9]
凝血因子 VIII FVIII	小鼠 <i>Mus musculus</i>	1 000–4 000	2015	[10]
凝血因子 IX FIX	猪 <i>Sus</i>	100–400	2008	[11]
促红细胞生成素 EPO	山羊 <i>Capra hircus</i>	未报道 Not reported	2002	[12]
人 C1 抑制剂 Human C1 inhibitor	兔 <i>Leporidae</i>	未报道 Not reported	2003	[13]
人酸性 α - 葡萄糖苷酶 Human acid alpha-glucosidase	小鼠 <i>Mus musculus</i>	1.5	1996	[14]
粒细胞巨噬细胞集落刺激因子 GM-CSF	小鼠 <i>Mus musculus</i>	200–4 600	1997	[15]
重组人抗凝血酶 Recombinant human antithrombin	山羊 <i>Capra hircus</i>	未报道 Not reported	1998	[16]
人尿激酶 Human urokinase	小鼠 <i>Mus musculus</i>	1 000–2 000	1990	[17]
胰岛素样生长因子 -1 IGF-1	兔 <i>Leporidae</i>	1 000	1992	[18]
铜 / 锌超氧化物歧化酶、细胞外超氧化物歧化酶 CuZn-SOD、EC-SOD	山羊 <i>Capra hircus</i>	100.14 ± 5.09 、 279.10 ± 5.38	2018	[19]
卵泡刺激素 FSH α 、FSH β	山羊 <i>Capra hircus</i>	0.28、0.30	2021	[20]
芳基烷基胺 N- 乙酰转移酶、乙酰 5- 羟色胺甲基 转移酶 AANAT、ASMT	绵羊 <i>Ovis aries</i>	未报道 Not reported	2017	[21]
葡萄糖脑苷脂酶 Glucocerebrosidase	山羊 <i>Capra hircus</i>	111.1 ± 8.1	2015	[22]
人蛋白 C Human protein C	猪 <i>Sus</i>	100–1 000	1994	[23]
抗程序性细胞死亡 1 抗体 PD-1 antibody	小鼠 <i>Mus musculus</i>	80.52 ± 0.82	2020	[24]
丁酰胆碱酯酶 Butyrylcholinesterase	山羊 <i>Capra hircus</i>	1 000–5 000	2008	[25]
生长激素 Growth hormone	兔 <i>Leporidae</i>	10	2012	[26]
胶原 Collagen	小鼠 <i>Mus musculus</i>	200	2000	[27]
抗 CD20 单克隆抗体 Anti-CD20 mAb	小鼠 <i>Mus musculus</i>	17	2008	[28]

Note: HSA : Human serum albumin ; AAT : antitrypsin ; EPO: hemopoietin ; GM-CSF : granulocyte macrophage colony stimulating factor ; IGF-1 : insulin-like growth factor-1 ; CuZn-SOD: copper/zinc superoxide dismutase ; EC-SOD: extracellular superoxide dismutase ; AANAT: aryl alkylamine N-acetyltransferase ; ASMT: acetyl 5-serotonin methyltransferase

卵巢细胞（Chinese hamster ovary cell, CHO 细胞）生产，由于随机整合导致外源基因沉默或转基因动物表型发生改变，同时由于缺少对外源基因的有效调控导致转基因表达水平低等问题，在实际应用中乳腺生物反应器表达重组蛋白的效率未能达到理想水平^[4, 29]，因此，避免外源基因的随机整合并提高外源蛋白表达水平是乳腺生物反应器优化的关键。目前有多种方法可介导外源基因的整合。位点特异性重组酶可介导外源基因定点整合，但存在可逆、高脱靶率等缺陷^[30]。近几年基因编辑技术的出现，特

别是 CRISPR/Cas9 系统的应用，可以将外源基因精准整合到基因组特定位点，很大程度上克服了外源基因随机整合问题。另外，有研究表明，靶位点高开放性的染色质不仅可提高 CRISPR/Cas9 基因编辑的效率^[31]，也有助于增强外源基因的表达^[32]，因此在高开放位点插入外源基因，也许能提高乳腺生物反应器的表达效率。此外，参与转录及转录后调控的顺式作用元件对转基因表达的水平 and 稳定性有重要影响。因此，本文就目前乳腺生物反应器重组蛋白表达效率低的问题提出了 3 条优化策略：

(1) 采用基因编辑技术定点整合外源基因, 如利用 CRISPR/Cas9、TALENs、ZFNs 三大基因编辑技术;(2) 以泌乳期乳腺上皮细胞染色质开放位点作为整合位点提高外源基因表达, 具体措施为: 通过 ATAC-seq 等测序技术并结合位点安全性评价筛选出泌乳期特异性开放、安全的位点作为靶位点;(3) 优化外源基因的调控序列, 具体包括: 改进启动子中转录因子结合基序, 选择合适的增强子, 优化内含子对外源基因的调控, 在表达载体中添加染色质修饰元件, 优化 3'UTR 提高外源基因 mRNA 稳定性等。

1 外源基因定点整合解决位置效应

目前可通过两种方法将外源基因插入基因组中, 一种方法是利用病毒载体或原核显微注射线性化基因片段随机整合外源基因, 但由于大多数基因组随机整合位点是转录抑制的, 可能导致组蛋白去乙酰化和启动子甲基化而引起基因沉默^[30, 33-34]。另一种方法是利用细胞的自我修复机制定点整合外源基因, 需要在基因组特定位点引入双链断裂 (DSB), 随后通过同源重组介导的双链 DNA 修复 (HDR) 介导转基因插入^[35]。后者是目前最常用的转基因整合的方法。DSB 的产生至关重要, 而基因编辑技术尤其是 CRISPR/Cas9 技术的进步使其变得简单高效, 只需要根据整合位点的核苷酸序列设计 sgRNA, 构建 CRISPR/Cas9 表达载体, 与同源重组修复模板共同转染细胞, 经过一系列的筛选鉴定即可将外源基因精准整合到靶位点^[36]。通过这种方法将外源基因插入合适的位点, 可有效解决位置效应, 降低转基因表达的异质性。

2 打靶染色质高开放性位点提高转基因表达水平

2.1 常用基因打靶位点

随着基因编辑技术的发展, 外源基因定点整合不再是难题, 但打靶位点的匮乏不能满足该技术的应用。如今已证实转基因整合的位置不同, 会导致多达 1 000 倍的转录差异, 因此打靶位点的选择十分重要^[37]。目前 *Rosa26* (*the reverse orientation splice acceptor β -geo 26 locus*), *H11* (*Hipp11*), *Col1a1* (*$\alpha 1(1)$ procollagen*), *AAVS1* (*the adeno-associated virus site 1*)、 β -乳球蛋白 (*BLG*) 基因和 β -酪蛋白 (β -casein) 基因

等位点已应用于家畜基因打靶, 其中 BLG 和 β -casein 是制备乳腺生物反应器最常用的几个打靶位点。对小鼠 β -casein 基因敲除实验证明该基因的缺失对小鼠的生存没有影响^[38], 宋绍征等^[39]已证实山羊 BLG 基因位点打靶获得的重构胚具有类似干细胞的全能发育性, 并且早期胚胎的发育也未受到影响。 β -酪蛋白和 β -乳球蛋白是反刍动物乳汁中含量较高的蛋白, 在该基因位点打靶, 可利用其内源性调控序列调控转基因表达^[40]。此外, 利用表达牛乳中主要过敏源的乳蛋白基因位点整合外源基因, 一方面可以改善牛乳的营养成分, 另一方面还能降低牛乳的致敏性^[39, 41]。但需要注意的是打靶乳蛋白基因座可能会影响乳腺组织的正常泌乳, 而且根据已有研究, 在常用的几个打靶位点整合外源基因依旧存在表达量低、表达不稳定的问题, 因此还需寻找更合适的打靶位点 (表 2)。

2.2 打靶位点染色质开放性对转基因表达的影响

将转基因整合到已知具有高转录活性的位置可以增加其表达, 而位点的转录活性与染色质开放性密切相关, 有研究者指出整合位点的开放性是影响转基因表达水平的主导因素^[50]。染色质开放性是指核大分子能够物理接近染色质 DNA 的程度, 真核生物的 DNA 包裹组蛋白形成核小体, 在细胞核中被紧密压缩, 而基因的表达依赖于调控因子与 DNA 序列结合所发挥的功能, 但核小体的存在阻碍调控因子对 DNA 序列的接近, 从而使基因表达受到影响。开放的染色质通常指缺乏核小体的区域, 目标基因及其上下游调控基因位点的染色质开放性变化可能会导致转录因子等调控因子的结合发生改变, 进而影响转录效率, 因此基因的高水平表达受开放染色质的调控, 转基因的表达与染色质区域的开放性有关。已有研究证明了在更加开放的染色质区域进行基因编辑能显著提高转基因的表达效率^[51-55]。

染色质开放性受 DNA 甲基化、组蛋白修饰以及染色质重塑等表观遗传修饰影响, 组蛋白乙酰化通过中和电荷、破坏氨基酸间的静电相互作用降低染色质的紧密状态^[56]; DNA 甲基化与染色质去乙酰化有关^[56]; 染色质重塑复合物可被分为 4 个亚类: ISWI、CHD、SWI/SNF 和 INO80, 即使它们发挥着

表 2 常用家畜打靶位点

打靶位点 Target site	物种 Species	转基因 Transgenes	基因编辑技术 Gene editing technology	时间 / 年 Time/Year	参考文献 Referencet
Rosa26	绵羊 <i>Ovis aries</i>	turboGFP	CRISPR/Cas9	2016	[42]
	猪 <i>Sus</i>	tdTomato	TALENs	2014	[43]
	牛 <i>Bovine</i>	自然抗性相关巨噬细胞蛋白 1 NRAMP1	CRISPR/Cas9	2021	[44]
β -casein	牛 <i>Bovine</i>	人成纤维细胞生长因 2 Human fibroblast growth factor 2	CRISPR/Cas9	2015	[45]
	绵羊 <i>Ovis aries</i>	芳基烷基胺 N- 乙酰转移酶、乙酰 5- 羟色胺 甲基转移酶 AANAT、ASMT	CRISPR/Cas9	2017	[21]
	奶牛 <i>Bovine</i>	溶葡萄球菌酶 Lysostaphin	ZFNickases	2013	[46]
BLG	奶牛 <i>Bovine</i>	人血清白蛋白 HSA	TALENs	2016	[7]
	山羊 <i>Capra hircus</i>	人 α - 乳清蛋白 Human α -whey protein	Homologous recombination	2019	[41]
	山羊 <i>Capra hircus</i>	人乳铁蛋白 Human lactoferrin	CRISPR/Cas9	2017	[47]
	山羊 <i>Capra hircus</i>	人乳铁蛋白 Human lactoferrin	CRISPR/Cas9	2020	[39]
Col1a1	绵羊 <i>Ovis aries</i>	抗胰蛋白酶 AAT	Homologous recombination	2000	[8]
AAVS1	山羊 <i>Capra hircus</i>	转化生长因子 β 1 TGF- β 1	TALENs	2014	[48]
H11	猪 <i>Sus</i>	绿色荧光蛋白 GFP	CRISPR/Cas9	2015	[49]

Note : NRAMP1 : natural resistance-associated macrophage protein 1 ; TGF- β 1 : transforming growth factor- β 1 ; GFP : green fluorescent protein

不同的功能（核小体组装、染色质访问和核小体编辑），但都具有使 DNA 易位和打破组蛋白与 DNA 接触的机制以实现核小体重新定位、滑动、排出或编辑的功能^[57]。其中 SWI/SNF 与染色质开放性最为相关，它使核小体沿着 DNA 滑动而排出核小体组分或完整的核小体，从而促进染色质开放^[57]。因此染色质表观遗传修饰特征可以反映染色质开放性。相关研究表明抑制转基因表达的位点常常具有不开放染色质结构的表观遗传特性，如缺乏 H3K4 甲基化和 H3 乙酰化，富集 DNA 甲基化等，而在具有象征开放染色质区域表观遗传特征的区域进行基因编辑能显著提高转基因表达效率^[51-55]。最近，对猪基因组顺式调控元件的研究验证了富含 H3K4me3 和 H3K27ac 标志的活性启动子有更高的基因转录激活水平和表达量^[58]，而富含 H3K4me3 和 H3K27ac 表观遗传修饰的活性启动子和增强子常出现在开放染色质区域^[59-62]。

随着泌乳进程的发展或激素分泌的变化，由于 DNA 甲基化、组蛋白修饰以及染色质重塑等表观遗传机制的调控，泌乳相关基因的开放程度也会随之改变，泌乳相关基因的染色质开放性处于一个动态

变化的过程，当开放程度的改变发生在启动子或增强子等区域时，则会影响基因的表达。并且，相比于奶期，某些基因在泌乳期会特异性高开放并高水平表达^[3, 63-65]。因此，筛选泌乳期高开放的基因区域并对其进行打靶可以促进转基因在泌乳期的表达。值得一提的是，随着 DNase-seq、ATAC-seq、CHIP-seq 等染色质开放位点测序技术的发展，获得特定时期的染色质开放位点已不再是难题，尤其是 ATAC-seq 凭借其简单、省时、灵敏度高且不需要大量细胞的优势而被广泛应用^[66]。

2.3 打靶位点的安全性评价

对于乳腺生物反应器重组蛋白的生产而言，筛选外源基因整合位点既要保证转基因稳定高水平表达，也要降低打靶对宿主表型造成不良影响的风险，即保证靶位点的安全性。外源基因与宿主细胞基因组之间的相互作用，一方面可能影响外源基因本身的表达，导致其表达衰减或完全沉默，另一方面还可能对内源性基因产生影响，最终导致宿主细胞的克隆扩增或恶性转化^[67]，理想的靶位点必须避免上述情况的发生。“基因安全港”（genomic safe harbors, GSHs）是指基因组中能容纳外源基因序列，使之稳

定且充分表达,并对受体细胞和生物体没有负面影响的基因打靶位点。最初研究者对“基因安全港”提出五个衡量标准:1)该位点到任何基因5'端的距离大于50 kb;2)该位点到癌相关基因的距离大于300 kb;3)该位点到任何microRNA的距离大于300 kb;4)该位点位于基因转录单元外;5)该位点位于超保守区域外^[68]。提出这5个标准主要考虑了两个方面,一是与已知在癌症中起作用的基因或microRNA基因保持必要距离,二是避免内源性编码基因和保守基因元件的破坏^[67]。随着对基因结构、功能和调控等方面的更多研究,“基因安全港”的标准也得到进一步发展。基因安全港的概念为筛选乳腺生物反应器安全打靶位点提供了参考。

一些调控性非编码RNA已被证明在各种细胞和其生理过程中发挥重要作用,因此Papapetrou等^[67]提出判定打靶位点的安全性还应考虑对调控性非编码RNA的影响。最近Pellenz等^[69]增加了筛选安全打靶位点的标准,即靶位点应具有低转录活性、不在拷贝数可变区域、位于开放染色质区域且在基因组中只有一个拷贝。另外,染色质拓扑关联结构域(topologically associating domains, TADs)可作为基因安全港的标记。TAD间有边界隔离,在每一个TAD内的DNA序列与域内其他DNA序列的相互作用频率显著高于与域外的DNA序列的相互作用频率^[70]。Hilliard等^[71]指出TAD边界与基因表达显著关联,保守边界之间的稳定TAD区域表现出对基因表达下调的抵抗力,因此保守的TAD也许可以作为安全港区域的标记,帮助我们筛选打靶位点。近几年有人指出DNA酶I超敏位点(DNaseI)可能富含DNA转录因子结合位点和转录调控元件^[72],所以在筛选安全打靶位点时排除DNaseI超敏区域或许可以避免对内源基因转录调控的不良影响^[73]。

目前多数研究者认为与基因内打靶位点相比,打靶基因间区域的安全性更高。尽管在基因间区域包含许多功能性元件,如转录因子结合位点、增强子样序列和转录终止子等,但与基因内区域相比,基因间区域的破坏对细胞功能的影响更小,尤其是非模式生物的大多数基因功能注释不全,打靶基因区域有更多的不确定性。然而,基因间区域经常处于沉默状态,在此区域插入的外源基因能否有效表

达仍有待商榷^[50,68]。需要注意的是,打靶位点很难保证外源基因持久稳定地表达,转基因可能在细胞分化过程中被沉默,如被我们普遍长久当作安全位点的AAVS1位点在内胚层和中胚层分化时可能发生甲基化,从而导致基因表达沉默^[74]。

3 调控序列对外源蛋白表达效率的影响

虽然靶位点对转基因定点整合及表达效率有极为重要的作用,但外源蛋白表达效率是多种因素共同作用的结果,定点整合外源基因到染色质开放区域不能保证其高表达效率,因此仅通过选择合适的靶位点并不能完全解决乳腺生物反应器的问题,还需要考虑基因表达的转录及转录后调控。表达载体的调控元件通常包括启动子、增强子、内含子、绝缘子以及染色质修饰元件如基质附着区域(matrix attachment region, MAR)、普遍存在的染色质开放元件(ubiquitous chromatin opening elements, UCOE)、稳定抗阻遏元件(stabilizing anti-repressor element, STAR)和5'/3'非编码区域(untranslated region, UTR)等。启动子具有结合并正确定位转录起始复合物的作用,为启动转录所必需的调控序列;增强子包含基因调控所需的特定转录因子的结合位点,具有刺激和调节基因转录活动的作用;非编码区域影响转基因表达水平和稳定性。这里将主要介绍启动子、增强子、内含子、3'UTR和染色质修饰元件(包括绝缘子、UCOE、MAR和STAR)对重组蛋白表达效率的影响。表3列出了部分在试验中已被证实对转基因表达有积极作用的调控序列。

3.1 启动子

哺乳动物细胞的转录过程较为复杂,受顺式调控元件和反式作用因子的调节。其中,启动子在整合和处理与转录起始相关的信号方面起着关键作用,较弱的启动子会增加基因间转录的干扰性,而强启动子的基因表达量通常会更高,因此启动子的选择也影响着基因表达的效率^[50]。在乳腺细胞产生乳蛋白的过程中,参与转录激活的转录因子具有保守性,所以一种动物的启动子可以在另一种动物的乳腺细胞中有效调控基因的转录^[35]。山羊 β -酪蛋白(*goat β -casein*)、牛 α -s1酪蛋白(*bovine α SI-casein*)、鼠乳清酸性蛋白(*Mouse WAP*)和绵羊 β -乳球蛋白(*ovine*

BLG) 等基因的启动子已成功用于乳腺生物反应器生产重组蛋白, 其中 β -酪蛋白基因启动子作为最有效的启动子之一, 广泛用于小鼠、山羊和奶牛乳腺生物反应器^[92], 但这些启动子仍旧不能满足工业化生产的需求。启动子是转录激活的第一步, 同时也决定着转录的组织特异性, 因此在不改变启动子泌乳期特异性表达的基础上对其进行改进十分重要。

参与乳蛋白表达的关键转录因子都不是乳腺所特有的, 但由于特定的组合而导致乳蛋白基因启动子在泌乳期间特异性激活转录, 所以有必要阐明乳蛋白在生理状态下的转录激活过程。乳腺分泌乳汁的过程受到激素的调控, 而启动子和增强子含有多种与激素相关的转录因子结合元件, 如激活性转录因子 STAT5 和 C/EBP 以及抑制性转录因子 YY1 等。以往的研究已经阐述了催乳激素激活 β -酪蛋白表达的信号通路: 当缺乏催乳激素的刺激时, YY-1 与 β -casein 基因的远端启动子结合, 并通过影响染色质的表观遗传修饰, 促进组蛋白 H3 的去乙酰化和二甲基化以抑制染色质的开放, 从而抑制 β -酪蛋白基因的表达; 而经催乳激素刺激后, 激活 JAK2-STAT5 信号通路, 进而使 STAT5 蛋白被磷酸化, 磷酸化的 STAT5a 和 STAT5b 形成同源或异源二聚体^[93]。二聚体 STAT5、GR 和 C/EBP 等迅速与 β -casein 基因近端启动子和远端增强子结合, 招募共激活因子 p300 促进组蛋白 H3 乙酰化, 从而使染色质结构开放, 促进 β -casein 基因的表达^[94]。在这些乳蛋白基因转录相关的转录因子中, STAT5 起着核心作用^[95]。研究者发现, 敲除或者突变 STAT5 基因序列会导致乳蛋白基因的表达显著降低^[95], 而 STAT5 的过度表达使乳蛋白基因的表达显著升高^[93]。另外, 介于 YY1 的转录抑制作用, 近端启动子区 YY1 负调控元件的缺失可能对提高催乳素诱导的近端乳蛋白表达至关重要^[96]。

在以上研究的基础上, 改进转录因子结合基序, 增强激活性转录因子与启动子的结合, 而不改变启动子乳腺特异性启动转录的特性, 将有助于增强外源蛋白在乳腺中特异性表达。

3.2 增强子

即使核心启动子足以启动组装 RNA 聚合酶 II

的机制, 但缺乏增强子调控的转录水平通常较弱^[62]。利用增强子能显著提高乳腺中转基因的表达式, Cheng 等^[77]将 CMV 增强子和乳腺特异性启动子杂合以调控小鼠中重组人乳铁蛋白的表达, 并使目标蛋白增加了约 10 万倍, 且目标蛋白并没有在乳腺以外的组织中表达。之后他们使用山羊 β -酪蛋白/CMV 嵌合启动子驱动目标蛋白在山羊和兔乳腺中的特异性表达也获得了较理想的结果^[9, 19]。在转录过程中, 先锋转录因子先与增强子结合, 随后募集其他转录因子、辅调节因子、染色质重塑因子和修饰因子以及 RNA 聚合酶 II 等进一步激活转录^[97]。增强子的功能主要由其包含的密集的转录因子结合位点反映, 一方面增强子中聚集的多个转录因子结合位点使转录因子协同作用从而逐渐克服增强子形成核小体的倾向, 进而实现基因调控, 另一方面因为增强子表达多个不同转录因子的结合位点, 使得一个增强子可以在多个细胞类型中调节同源基因^[97], 这为我们利用不同来源的增强子调节转基因表达提供了便利。

增强子通常混合在其他基因座中, 想要充分利用增强子的转录激活效应就要注重其激活作用的特异性, 多种因素决定了增强子-启动子相互作用的特异性, 如启动子的染色质可及性、核心启动子内的序列、结构蛋白在启动子近端的结合以及周围的染色质环境等^[98]。哺乳动物含有的增强子数目远多于启动子, 一个启动子可能与多个增强子相互作用, 在同一基因附近具有类似活性的多个增强子相互补充, 并且功能冗余增强子能在其他增强子损失时发挥补偿、缓冲的作用, 从而赋予表型稳健性^[99-100]。Panigrahi 等^[97]发现同一基因附近的多个增强子中最远端的增强子直接与启动子接触, 并将增强子链中的其他增强子带到目标启动子处, 因此最远端增强子在靶基因调节中起着关键作用。另外, 这些相似的增强子在基因表达方面可能起着协同调节的作用, 例如小鼠 *Wap* 基因在孕期的表达上调了近 1 000 倍, 虽然远端增强子的作用最强, 但这是 3 个增强子的联合作用导致的结果^[101]。

目前很多对增强子的研究集中于超级增强子, 超级增强子是被相关转录因子、RNA 聚合酶 II 及其相关的辅因子和染色质调节因子占据的一大簇增强子, 驱动定义细胞特性的基因的表达^[102]。相比

常规增强子, 超级增强子对转录的影响更大、更有利于促进转基因的表达^[103]。Zeng 等^[104]便发现在乳腺中, *Wap* 基因超级增强子和 *Tbgr4* 基因启动子的组合使 *Tbgr4* 基因的表达增加了 5 倍。不过由于超级增强子具有启动子特异性, 它与自身靶启动子组合上调基因表达的能力远优于与其他启动子的组合^[104], 因此在通过超级增强子提高转基因表达时还需要将其与自身靶启动子结合以更好发挥超级增强子转录激活的潜力。

需要注意的是, 增强子不仅具有转录激活的功能, 随着增强子上组装的蛋白质不同, 增强子也可能抑制转录^[97]; 最近研究也证实了位于内含子中的某些增强子在控制组织特异性表达中起核心作用^[105]; 启动子和增强子的功能也有重叠, 某些增强子具有驱动局部转录起始的启动子功能, 同时, 某些启动子又具有增强子的增强转录功能^[106-107], 例如 Zhu 等^[108]发现随着活性启动子数量的增加, 其附近的增强子的数量会减少, 这可能是由于活性启动子具有普遍的增强子样功能, 这也使得对启动子和增强子的定义变得模糊。

3.3 内含子

在哺乳动物、植物和真菌等真核生物体中, 都发现了内含子提高基因表达的能力, 加入内含子已被公认为是提高转基因表达水平的一种途径。目前已有通过添加内含子提高 CHO 细胞重组蛋白产量的报道^[79, 109], 在乳腺生物反应器中, β -casein 基因的第二个内含子也常被添加到载体中以提高外源蛋白的表达^[78]。内含子增加基因表达的这种现象被称为“内含子介导的增强”(intron-mediated enhancement, IME)^[110], 但是这种现象背后的分子机制仍不完全清楚。Bieberstein 等^[111]发现哺乳动物细胞中内含子与染色质修饰有关, 内含子可促进启动子 H3K4me3 和 H3K9ac, 从而有助于将通用转录因子招募到启动子, 提高基因表达水平。真核生物的转录通常启动子在正义方向和上游反义方向启动, 在正义方向 mRNA 的转录持续到聚合酶到达基因的 3' 端, 而反义方向的转录很快会终止, 这种启动子的方向性被认为与内含子有关, 其主要依赖于内含子导致的基因环状结构的形成, 该环状结构有助于基因 5' 端附

近招募终止因子以终止启动子的上游反义转录, 从而赋予启动子方向性, 有研究表明该环状结构缺陷的突变体会丧失转录的方向性^[112]。不过内含子促进基因环状结构形成的机制及基因环状结构增强转录的机制仍不清楚。此外, 内含子促进转录还可能与其促进 mRNA 输出和稳定 mRNA 结构等机制有关。有人指出 mRNA 被剪接后, 剪接体沉积于外显子与外显子连接处形成外显子连接复合体(exon junction complex, EJC), 有助于 mRNA 向胞质输出^[113], 而一些动植物的全基因组分析表明 mRNA 稳定性与基因中内含子的数量呈正相关^[114]。研究表明一个转基因必须至少包含一个内含子, 由于许多基因的第一内含子包含转录因子结合位点, 将其添加到载体中有利于转基因表达^[115]。

同一内含子插入同一基因的不同位置或是在同一基因的不同位置插入不同内含子都对 IME 有重要影响^[114]。并不是所有内含子都有增强转录的能力, 有些内含子甚至抑制基因的表达^[116]。通常认为内含子必需位于转录序列内, 距离转录起始位点必须小于 1 kb 才能有效提高基因表达^[117], 且距离启动子更近的内含子有更强的增强潜力^[118]。但最近 Dwyer 等^[119]将内含子插入酵母基因不同位置时发现随着远离启动子的内含子对终止子的接近, 内含子重新对基因表达产生积极影响。该现象可能与终止子的近端内含子影响转录起始前复合物(pre-initiation-complex, PIC)的组装和启动子方向性有关^[119]。总之, 由于内含子介导转基因表达效率提高的复杂性, 还需要进一步探究内含子类型、位置与基因、细胞类型以及生理条件的关系, 以获得最有利于外源基因在乳腺生物反应器中表达的内含子组合。

3.4 染色质修饰元件

由于宿主基因和引入的转基因之间的关系复杂, 即使将外源基因稳定整合到靶位点, 在转基因附近染色质结构和内源性调控元件的影响下也有可能引发位置效应, 导致转基因沉默。在载体上添加染色质修饰元件有助于将打靶位点的染色质环境维持在有利于转录的状态, 从而提高转基因表达水平和稳定性。研究表明, MAR、绝缘子、UCOE 和 STAR

等染色质修饰元件能有效提高转基因表达效率。根据染色质修饰元件功能可将其大致分为两类：一类可以阻止异染色质标记扩散到常染色质区域，形成染色体区域之间的边界，如 MAR 和绝缘子；一类可以主动使染色质进入允许转录的状态，如 UCOE 和 STAR^[120]。

绝缘子是在相邻染色质区域之间形成功能边界的 DNA 序列，不同于传统的顺式调控元件，绝缘子本身并不表现出固有的转录增强或抑制活性，而是通过保护其他顺式调控元件不受周围染色质的干扰而发挥作用。根据绝缘子功能可以将其分为两类：一类绝缘子可以干扰某个基因增强子与另一个基因启动子之间的相互作用，从而保护后者的转录活性，这类绝缘子称为增强子阻断绝缘子（enhancer-blocking insulators）；另一类绝缘子具有屏障功能，可以防止异染色质扩散到邻近区域，称为屏障绝缘子（barrier insulators）^[121]。最初人们认为绝缘子是功能上独立的元件，它们的活动与周围染色质无关且不会与其他绝缘子进行任何形式的交互作用，不过之后的试验表明，不仅相邻近的绝缘子之间会有相互作用，而且远距离的绝缘子间也可以进行物理接触^[122]。目前一些研究已经证实了绝缘子在避免位置效应和增强转基因表达方面的作用，但基于不同的试验条件，绝缘子呈现出了不同的效果^[123]。当前在乳腺生物反应器中应用最广泛的是鸡 β 珠蛋白绝缘子（chicken β -globin insulator）^[79-80]，但对其在乳腺生物反应器中的效果还需要进一步探索。

MAR 是染色质中的一段与核基质特异结合的序列，其与核基质结合使染色质形成独立的环状结构，通过形成不同的调节域来隔绝周围染色质产生的影响以增加基因的表达^[124]。研究已表明在载体中添加 MAR 能改善转基因的表达，且通过添加 MAR 优化表达载体已被商业化^[125]。不过，在使用不同的 MAR、载体或细胞系统时结果可能不同甚至相互矛盾，因此推断 MAR 功能或活性受到载体成分的影响^[126-127]。

UCOE 是位于普遍表达的管家基因启动子内的无甲基化 CpG 岛，可以通过抑制启动子 DNA 甲基化而避免转基因沉默^[128]。研究者已证实 UCOE 在瞬时和稳定转染中皆可以提高转基因表达水平^[129]，

且 UCOE 与 MAR 类似，都具有启动子特异性^[125]。Wang 等^[130]发现 MAR、UCOE 与启动子上调控元件相互作用调节染色质的改变，并且 UCOE 与含增强子的启动子结合更有利于基因表达，而 MAR 与含 CpG 岛和 TATA 盒的启动子结合更利于基因表达，其可能机制是 MAR 更有利于转录起始区含 CpG 岛的启动子上形成核小体占位的合理分布；而 UCOE 则有利于带增强子的启动子上核小体占位的“合理”分布。另外，与含 MAR 的载体相比，通常情况下含 UCOE 的表达载体具有更高效和更稳定的基因表达^[130]。目前 UCOE 也已被用于构建生产重组蛋白的商业化表达载体^[125]。有研究者对比 4 种染色质修饰元件后发现，A2UCOE 相比 MAR X-S29、STAR40 和 cHS4 这几种常用染色质修饰元件有更强的增强转基因表达量和稳定性的能力^[90]。

STAR 可以抵消染色质相关阻遏物的副作用，使染色质进入转录允许状态，从而提高重组蛋白的表达水平和稳定性。Kwaks 等^[91]首次报道了 STAR40 对 CHO 细胞外源蛋白碱性磷酸酶表达量和表达稳定性的增强效果，但在 Saunders 等^[90]的研究中却发现 STAR40 仅具有有限的活性，不过他们验证了 STAR7 和 STAR67 在 CHO 细胞中增强外源蛋白表达的功效，特别是将其组合使用时，二者显示出了更强的效果。目前对 STAR 的报道还较少，STAR 也还未被用于工业生产^[125]。

需要注意的是，对 MAR、UCOE 和 STAR 等染色质修饰元件影响外源蛋白表达的研究大多基于 CHO 细胞，而染色质修饰元件对转基因的影响基本都具有特异性，外源蛋白表达系统、细胞类型和载体构建等因素都会对其效果产生影响，因此还需要进一步探索其在乳腺生物反应器中的功能。

3.5 优化 3'UTR 提高 mRNA 稳定性

基因表达受到转录调控以及转录后调控的影响，从基因结构的角度考虑，5'UTR 主要影响转录激活以及翻译激活，3'UTR 则主要调控 mRNA 稳定性。在反刍动物中，乳蛋白主要由酪蛋白组成，其中 α s2 酪蛋白与 β 酪蛋白的 mRNA 丰度相差不大，但最终蛋白质积累量却相差几倍，关键就在于 mRNA 稳定性的差异。

表 3 外源蛋白调控序列

Table 3 Regulatory sequences of exogenous proteins

类型 Type	名称 Name	转基因 Transgene	表达系统 Expression system	参考文献 Reference
启动子 Promoter	β -lactoglobulin promoter	溶菌酶 Lysozyme	山羊乳腺生物反应器 Goat mammary gland bioreactor	[75]
	α S1-casein promoter	抗程序性细胞死亡 1 抗体 PD-1 antibody	小鼠乳腺生物反应器 Mouse mammary gland bioreactor	[24]
	CMV promoter	凝血因子Ⅷ FVIII	小鼠乳腺生物反应器 Mouse mammary gland bioreactor	[10]
	CMV enhancer	人乳铁蛋白 Human lactoferrin	山羊乳腺生物反应器 Goat mammary gland bioreactor	[76]
增强子 Enhancer	CMV enhancer	人乳铁蛋白 Human lactoferrin	小鼠乳腺生物反应器 Mouse mammary gland bioreactor	[77]
	β -casein intron	溶菌酶 Lysozyme	小鼠乳腺生物反应器 Mouse mammary gland bioreactor	[78]
内含子 Intron	SV40 intron	促红细胞生成素 EPO	中国仓鼠卵巢细胞 CHO cell	[79]
	β -lactoglobulin insulator	溶菌酶 Lysozyme	猪乳腺生物反应器 Porcine mammary gland bioreactor	[80]
绝缘子 Insulator	eHS4 insulator	人血清白蛋白 HSA	牛乳腺生物反应器 Cattle mammary gland bioreactor	[81]
	tDNA insulator	单克隆抗体 Monoclonal antibody	中国仓鼠卵巢细胞 CHO cell	[82]
MAR	MAR X-29	增强型绿色荧光蛋白 eGFP	中国仓鼠卵巢细胞 CHO cell	[83]
	Chicken lysozyme MAR	免疫球蛋白 G IgG	中国仓鼠卵巢细胞 CHO cell	[84]
	IFN- β MAR	促红细胞生成素、肝细胞生长因子 EPO, HGF	中国仓鼠卵巢细胞 CHO cell	[85]
	β -globin MAR	可溶性 TGF- β II 型受体 sTbetaRII	中国仓鼠卵巢细胞 CHO cell	[86]
	TOP1 MAR	增强型绿色荧光蛋白 eGFP	中国仓鼠卵巢细胞 CHO cell	[87]
	DHFR intron MAR	增强型绿色荧光蛋白 eGFP	中国仓鼠卵巢细胞 CHO cell	[88]
	MAR6	增强型绿色荧光蛋白 eGFP	中国仓鼠卵巢细胞 CHO cell	[88]
	UCOE	绿色荧光蛋白 GFP	中国仓鼠卵巢细胞 CHO cell	[89]
STAR	STAR7	单克隆抗体 Monoclonal antibody	中国仓鼠卵巢细胞 CHO cell	[90]
	STAR67	单克隆抗体 Monoclonal antibody	中国仓鼠卵巢细胞 CHO cell	[90]
	STAR40	分泌碱性磷酸酶 SEAP	中国仓鼠卵巢细胞 CHO cell	[91]

Note : HGF : hepatocyte growth factor; sTbetaRII : soluble TGF-beta type II receptor; SEAP : secreted alkaline phosphatase

3'UTRs 通过调节 mRNA 稳定性和翻译来决定蛋白质水平, 主要由富含 AU 的元件 (AU-rich elements, ARE) 和 miRNAs 介导^[131]。ARE 元件广泛存在于哺乳动物基因的 3'UTRs 中, 根据与之结合的 RNA 结合蛋白 (RNA binding protein, RBP) 的不同, 起到稳定 mRNA 或诱导 mRNA 衰变的作用。 β 酪蛋白的 ARE 通过结合 HuR 蛋白而稳定 mRNA 从而促进 CSN2 mRNA 的高表达^[132], 因此可在转基因的 3'UTRs 中添加 ARE 元件以提高乳腺中外源基因 mRNA 的稳定性。除了 ARE 元件外, 还有多种 RNA 基序可用于增强 mRNA 的稳定性^[133-135], 但在乳腺中是否具有相同作用还有待探究。

4 总结与展望

在众多可用于生物制药的表达系统中, 乳腺生物反应器以其低成本、高效益的特性脱颖而出, 然而外源基因随机整合以及外源蛋白表达不稳定等问题长期限制了其发展。近几年基因编辑技术的出现, 使乳腺生物反应器迎来新机遇。通过基因编辑技术将外源基因定向整合到合适的靶位点, 可以有效避免位置效应的发生并使外源基因稳定表达, 同时通过优化调控序列构建高效表达载体又可以进一步提高外源蛋白表达水平。

除以上提到的问题外, 乳腺生物反应器还存在其他一些问题待克服, 如体细胞核移植效率低, 有人畜共患病风险, 重组蛋白的提纯难度较大等。目前乳腺生物反应器生产重组蛋白的技术依旧不够成熟, 如何优化仍然充满挑战, 但乳腺生物反应器投入药用蛋白规模化生产、成为药用蛋白生产的主流是生物制药业发展的必然趋势。随着对基因表达调控研究的深入、分子生物学技术的进步, 以及对乳腺生物反应器的不断优化, 研发出更有效、更安全、更经济的乳腺生物反应器将成为可能, 届时将满足全球对治疗性蛋白日益增长的需求, 从而减少疾病的发生, 以维护人类及其他生物的健康和福祉。

参考文献

- [1] O' Flaherty R, Bergin A, Flampouri E, et al. Mammalian cell culture for production of recombinant proteins: a review of the critical steps in their biomanufacturing [J]. *Biotechnol Adv*, 2020, 43: 107552.
- [2] Wang YL, Zhao SH, Bai L, et al. Expression systems and species used for transgenic animal bioreactors [J]. *Biomed Res Int*, 2013, 2013: 580463.
- [3] Qian X, Zhao FQ. Current major advances in the regulation of milk protein gene expression [J]. *Crit Rev Eukaryot Gene Expr*, 2014, 24 (4): 357-378.
- [4] Clark AJ. The mammary gland as a bioreactor: expression, processing, and production of recombinant proteins [J]. *J Mammary Gland Biol Neoplasia*, 1998, 3 (3): 337-350.
- [5] Lu D, Liu S, Ding FR, et al. Large-scale production of functional human lysozyme from marker-free transgenic cloned cows [J]. *Sci Rep*, 2016, 6: 22947.
- [6] Cui CC, Song YJ, Liu J, et al. Gene targeting by TALEN-induced homologous recombination in goats directs production of β -lactoglobulin-free, high-human lactoferrin milk [J]. *Sci Rep*, 2015, 5: 10482.
- [7] Luo Y, Wang YS, Liu J, et al. Generation of TALE nickase-mediated gene-targeted cows expressing human serum albumin in mammary glands [J]. *Sci Rep*, 2016, 6: 20657.
- [8] McCreath KJ, Howcroft J, Campbell KH, et al. Production of gene-targeted sheep by nuclear transfer from cultured somatic cells [J]. *Nature*, 2000, 405 (6790): 1066-1069.
- [9] Song SZ, Ge X, Cheng YB, et al. High-level expression of a novel recombinant human plasminogen activator (rhPA) in the milk of transgenic rabbits and its thrombolytic bioactivity in vitro [J]. *Mol Biol Rep*, 2016, 43 (8): 775-783.
- [10] Wang Q, Hao S, Ma L, et al. Comparison of human coagulation factor VIII expression directed by *Cytomegalovirus* and mammary gland-specific promoters in HC11 cells and transgenic mice [J]. *Blood Coagul Fibrinolysis*, 2015, 26 (7): 755-761.
- [11] Gil GC, Velander WH, van Cott KE. Analysis of the N-glycans of recombinant human Factor IX purified from transgenic pig milk [J]. *Glycobiology*, 2008, 18 (7): 526-539.
- [12] 成勇, 王玉阁, 罗金平等. 由成年转基因山羊体细胞而来的克隆山羊 [J]. *生物工程学报*, 2002, 18 (1): 79-83.
Cheng Y, Wang YG, Luo JP, et al. Cloned goats produced from the somatic cells of an adult transgenic goat [J]. *Chinese Journal of Biotechnology*, 2002, 18 (1): 79-83.
- [13] Koles K, van Berkel PHC, Pieper FR, et al. N- and O-glycans of recombinant human C1 inhibitor expressed in the milk of transgenic rabbits [J]. *Glycobiology*, 2004, 14 (1): 51-64.

- [14] Bijvoet AG, Kroos MA, Pieper FR, et al. Expression of cDNA-encoded human acid alpha-glucosidase in milk of transgenic mice [J] . *Biochim Biophys Acta*, 1996, 1308 (2) : 93-96.
- [15] Uusi-Oukari M, Hyttinen JM, Korhonen VP, et al. Bovine alpha s1-casein gene sequences direct high level expression of human granulocyte-macrophage colony-stimulating factor in the milk of transgenic mice [J] . *Transgenic Res*, 1997, 6 (1) : 75-84.
- [16] Edmunds T, van Patten SM, Pollock J, et al. Transgenically produced human antithrombin: structural and functional comparison to human plasma-derived antithrombin [J] . *Blood*, 1998, 91 (12) : 4561-4571.
- [17] Meade H, Gates L, Lacy E, et al. Bovine AlphaS1-casein gene sequences direct high level expression of active human urokinase in mouse milk [J] . *Nature Biotechnology*, 1990, 8 (5) : 443-446.
- [18] Brem G, Hartl P, Besenfelder U, et al. Expression of synthetic cDNA sequences encoding human insulin-like growth factor-1 (IGF-1) in the mammary gland of transgenic rabbits [J] . *Gene*, 1994, 149 (2) : 351-355.
- [19] Lu R, Zhang T, Wu DJ, et al. Production of functional human CuZn-SOD and EC-SOD in bitransgenic cloned goat milk [J] . *Transgenic Res*, 2018, 27 (4) : 343-354.
- [20] Hua RM, Liu JX, Li Y, et al. Novel functional recombinant human follicle-stimulating hormone acquired from goat milk [J] . *J Agric Food Chem*, 2021, 69 (9) : 2793-2804.
- [21] Ma T, Tao J, Yang M, et al. An AANAT/ASMT transgenic animal model constructed with CRISPR/Cas9 system serving as the mammary gland bioreactor to produce melatonin-enriched milk in sheep [J] . *J Pineal Res*, 2017, 63 (1) : 2017 Aug;63 (1) .
- [22] Tavares KC, Dias AC, Lazzarotto CR, et al. Transient expression of functional glucocerebrosidase for treatment of Gaucher' s disease in the goat mammary gland [J] . *Mol Biotechnol*, 2016, 58 (1) : 47-55.
- [23] Morcöl T, Akers RM, Johnson JL, et al. The porcine mammary gland as a bioreactor for complex proteins [J] . *Ann N Y Acad Sci*, 1994, 721: 218-233.
- [24] Gong GH, Zhang W, Xie LP, et al. Expression of a recombinant anti-programed cell death 1 antibody in the mammary gland of transgenic mice [J] . *Prep Biochem Biotechnol*, 2021, 51 (2) : 183-190.
- [25] Baldassarre H, Hockley DK, Olaniyan B, et al. Milk composition studies in transgenic goats expressing recombinant human butyrylcholinesterase in the mammary gland [J] . *Transgenic Res*, 2008, 17 (5) : 863-872.
- [26] Lipinski D, Zeyland J, Szalata M, et al. Expression of human growth hormone in the milk of transgenic rabbits with transgene mapped to the telomere region of chromosome 7q [J] . *J Appl Genet*, 2012, 53 (4) : 435-442.
- [27] Bulleid NJ, John DC, Kadler KE. Recombinant expression systems for the production of collagen [J] . *Biochem Soc Trans*, 2000, 28 (4) : 350-353.
- [28] Tang B, Yu SS, Zheng M, et al. High level expression of a functional human/mouse chimeric anti-CD20 monoclonal antibody in milk of transgenic mice [J] . *Transgenic Res*, 2008, 17 (4) : 727-732.
- [29] Yu HQ, Wang XB, Zhu L, et al. Establishment of a rapid and scalable gene expression system in livestock by site-specific integration [J] . *Gene*, 2013, 515 (2) : 367-371.
- [30] Ahmadi M, Damavandi N, Akbari Eidgahi MR, et al. Utilization of site-specific recombination in biopharmaceutical production [J] . *Iran Biomed J*, 2016, 20 (2) : 68-76.
- [31] Jensen KT, Fløe L, Petersen TS, et al. Chromatin accessibility and guide sequence secondary structure affect CRISPR-Cas9 gene editing efficiency [J] . *FEBS Lett*, 2017, 591 (13) : 1892-1901.
- [32] Barnes LM, Bentley CM, Dickson AJ. Stability of protein production from recombinant mammalian cells [J] . *Biotechnol Bioeng*, 2003, 81 (6) : 631-639.
- [33] Houdebine LM. Production of pharmaceutical proteins by transgenic animals [J] . *Comp Immunol Microbiol Infect Dis*, 2009, 32 (2) : 107-121.
- [34] Wang TY, Guo X. Expression vector cassette engineering for recombinant therapeutic production in mammalian cell systems [J] . *Appl Microbiol Biotechnol*, 2020, 104 (13) : 5673-5688.
- [35] Shepelev MV, Kalinichenko SV, Deykin AV, et al. Production of recombinant proteins in the milk of transgenic animals: current state and prospects [J] . *Acta Naturae*, 2018, 10 (3) : 40-47.
- [36] Ran FA, Hsu PD, Wright J, et al. Genome engineering using the CRISPR-Cas9 system [J] . *Nat Protoc*, 2013, 8 (11) : 2281-2308.
- [37] Pavani G, Amendola M. Targeted gene delivery: where to land [J] . *Front Genome Ed*, 2021, 2: 609650.
- [38] Kumar S, Clarke AR, Hooper ML, et al. Milk composition and lactation of beta-casein-deficient mice [J] . *PNAS*, 1994, 91 (13) : 6138-6142.

- [39] 宋绍征, 张婷, 潘生强, 等. CRISPR/Cas9 系统介导的人乳铁蛋白基因在山羊 β -乳球蛋白基因座定点敲入 [J]. 中国农业大学学报, 2020, 25 (7): 111-119.
- Song SZ, Zhang T, Pan SQ, et al. hLF gene knock-in at the BLG locus of goat by CRISPR/Cas9 system [J]. J China Agric Univ, 2020, 25 (7): 111-119.
- [40] 冀艳华, 徐乔璐, 刘军, 等. TALEN 介导猪瘟病毒结构蛋白基因 E0 敲入山羊乳腺上皮细胞 β -乳球蛋白基因座 [J]. 中国兽医学报, 2017, 37 (7): 1206-1211.
- Ji YH, Xu QL, Liu J, et al. CSFV E0 gene knocked-in β -lactoglobulin locus in goat mammary epithelial cells by TALEN [J]. Chin J Vet Sci, 2017, 37 (7): 1206-1211.
- [41] An LY, Yang L, Huang YJ, et al. Generating goat mammary gland bioreactors for producing recombinant proteins by gene targeting [J]. Methods Mol Biol, 2019, 1874: 391-401.
- [42] Wu MM, Wei CH, Lian ZX, et al. Rosa26-targeted sheep gene knock-in via CRISPR-Cas9 system [J]. Sci Rep, 2016, 6: 24360.
- [43] Li XP, Yang Y, Bu L, et al. Rosa26-targeted swine models for stable gene over-expression and Cre-mediated lineage tracing [J]. Cell Res, 2014, 24 (4): 501-504.
- [44] Yuan MK, Zhang JC, Gao YP, et al. HMEJ-based safe-harbor genome editing enables efficient generation of cattle with increased resistance to tuberculosis [J]. J Biol Chem, 2021, 296: 100497.
- [45] Jeong YH, Kim YJ, Kim EY, et al. Knock-in fibroblasts and transgenic blastocysts for expression of human FGF2 in the bovine β -casein gene locus using CRISPR/Cas9 nuclease-mediated homologous recombination [J]. Zygote, 2016, 24 (3): 442-456.
- [46] Liu X, Wang YS, Guo WJ, et al. Zinc-finger nickase-mediated insertion of the lysostaphin gene into the beta-casein locus in cloned cows [J]. Nat Commun, 2013, 4: 2565.
- [47] 周文君, 郭日红, 邓明田, 等. RS-1 提高 CRISPR-Cas9 系统介导的人乳铁蛋白基因敲入效率 [J]. 生物工程学报, 2017, 33 (8): 1224-1234.
- Zhou WJ, Guo RH, Deng MT, et al. RS-1 enhanced the efficiency of CRISPR-Cas9 mediated knock-in of human lactoferrin [J]. Chin J Biotechnol, 2017, 33 (8): 1224-1234.
- [48] Hu SW, Hall J, Wang Z, et al. TALEN-mediated targeted insertion of transforming growth factor-beta 1 (TGF-beta 1) gene into the goat AAVS1 locus [J]. Transgenic Research, 2014, 23 (1): 203-204.
- [49] Ruan JX, Li HG, Xu K, et al. Highly efficient CRISPR/Cas9-mediated transgene knockin at the H11 locus in pigs [J]. Sci Rep, 2015, 5: 14253.
- [50] Brady JR, Tan MC, Whittaker CA, et al. Identifying improved sites for heterologous gene integration using ATAC-seq [J]. ACS Synth Biol, 2020, 9 (9): 2515-2524.
- [51] Iler N, Goodwin AJ, McInerney J, et al. Targeted remodeling of human beta-globin promoter chromatin structure produces increased expression and decreased silencing [J]. Blood Cells Mol Dis, 1999, 25 (1): 47-60.
- [52] Nemeth MJ, Bodine DM, Garrett LJ, et al. An erythroid-specific chromatin opening element reorganizes beta-globin promoter chromatin structure and augments gene expression [J]. Blood Cells Mol Dis, 2001, 27 (4): 767-780.
- [53] Eyquem J, Poirot L, Galetto R, et al. Characterization of three loci for homologous gene targeting and transgene expression [J]. Biotechnol Bioeng, 2013, 110 (8): 2225-2235.
- [54] Yan CH, Boyd DD. Histone H3 acetylation and H3 K4 methylation define distinct chromatin regions permissive for transgene expression [J]. Mol Cell Biol, 2006, 26 (17): 6357-6371.
- [55] Veith N, Ziehr H, MacLeod RAF, et al. Mechanisms underlying epigenetic and transcriptional heterogeneity in Chinese hamster ovary (CHO) cell lines [J]. BMC Biotechnol, 2016, 16: 6.
- [56] Singh K, Erdman RA, Swanson KM, et al. Epigenetic regulation of milk production in dairy cows [J]. J Mammary Gland Biol Neoplasia, 2010, 15 (1): 101-112.
- [57] Clapier CR, Iwasa J, Cairns BR, et al. Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes [J]. Nat Rev Mol Cell Biol, 2017, 18 (7): 407-422.
- [58] Zhao YX, Hou Y, Xu YY, et al. A compendium and comparative epigenomics analysis of *Cis*-regulatory elements in the pig genome [J]. Nat Commun, 2021, 12 (1): 2217.
- [59] Reilly SK, Yin J, Ayoub AE, et al. Evolutionary genomics. Evolutionary changes in promoter and enhancer activity during human corticogenesis [J]. Science, 2015, 347 (6226): 1155-1159.
- [60] Chen KF, Chen Z, Wu DY, et al. Broad H3K4me3 is associated with increased transcription elongation and enhancer activity at tumor-suppressor genes [J]. Nat Genet, 2015, 47 (10): 1149-1157.
- [61] Bartosovic M, Kabbe M, Castelo-Branco G. Single-cell CUT&Tag profiles histone modifications and transcription factors in complex

- tissues [J]. *Nat Biotechnol*, 2021, 39 (7): 825-835.
- [62] Shlyueva D, Stampfel G, Stark A. Transcriptional enhancers: from properties to genome-wide predictions [J]. *Nat Rev Genet*, 2014, 15 (4): 272-286.
- [63] Rijnkels M, Freeman-Zadrowski C, Hernandez J, et al. Epigenetic modifications unlock the milk protein gene loci during mouse mammary gland development and differentiation [J]. *PLoS One*, 2013, 8 (1): e53270.
- [64] Singh K, Molenaar AJ, Swanson KM, et al. Epigenetics: a possible role in acute and transgenerational regulation of dairy cow milk production [J]. *Animal*, 2012, 6 (3): 375-381.
- [65] Zhang XY, Zhang SH, Ma L, et al. Reduced representation bisulfite sequencing (RRBS) of dairy goat mammary glands reveals DNA methylation profiles of integrated genome-wide and critical milk-related genes [J]. *Oncotarget*, 2017, 8 (70): 115326-115344.
- [66] Yan F, Powell DR, Curtis DJ, et al. From reads to insight: a hitchhiker's guide to ATAC-seq data analysis [J]. *Genome Biol*, 2020, 21 (1): 22.
- [67] Papapetrou EP, Schambach A. Gene insertion into genomic safe harbors for human gene therapy [J]. *Mol Ther*, 2016, 24 (4): 678-684.
- [68] Sadelain M, Papapetrou EP, Bushman FD. Safe harbours for the integration of new DNA in the human genome [J]. *Nat Rev Cancer*, 2011, 12 (1): 51-58.
- [69] Pellenz S, Phelps M, Tang WL, et al. New human chromosomal sites with "safe harbor" potential for targeted transgene insertion [J]. *Hum Gene Ther*, 2019, 30 (7): 814-828.
- [70] Beagan JA, Phillips-Cremins JE. On the existence and functionality of topologically associating domains [J]. *Nat Genet*, 2020, 52 (1): 8-16.
- [71] Hilliard W, Lee KH. Systematic identification of safe harbor regions in the CHO genome through a comprehensive epigenome analysis [J]. *Biotechnol Bioeng*, 2021, 118 (2): 659-675.
- [72] Meuleman W, Muratov A, Rynes E, et al. Index and biological spectrum of human DNase I hypersensitive sites [J]. *Nature*, 2020, 584 (7820): 244-251.
- [73] Autio MI, Motakis E, Perrin A, et al. Computationally defined and *in vitro* validated putative genomic safe harbour loci for transgene expression in human cells [J]. *bioRxiv*, 2022, DOI:10.1101/2021.12.07.471422.
- [74] Bhagwan JR, Collins E, Mosqueira D, et al. Variable expression and silencing of CRISPR-Cas9 targeted transgenes identifies the AAVS1 locus as not an entirely safe harbour [J]. *F1000Research*, 2019, 8: 1911.
- [75] Yu HQ, Chen JQ, Liu SG, et al. Large-scale production of functional human lysozyme in transgenic cloned goats [J]. *J Biotechnol*, 2013, 168 (4): 676-683.
- [76] Zhang T, Yuan YG, Lu R, et al. The goat β -casein/CMV chimeric promoter drives the expression of hLF in transgenic goats produced by cell transgene microinjection [J]. *Int J Mol Med*, 2019, 44 (6): 2057-2064.
- [77] Cheng Y, An LY, Yuan YG, et al. Hybrid expression cassettes consisting of a milk protein promoter and a cytomegalovirus enhancer significantly increase mammary-specific expression of human lactoferrin in transgenic mice [J]. *Mol Reprod Dev*, 2012, 79 (8): 573-585.
- [78] Li GC, Shi WQ, Chen G, et al. Construction and *in vivo* evaluation of a mammary gland-specific expression vector for human lysozyme [J]. *Plasmid*, 2014, 76: 47-53.
- [79] Xu DH, Wang XY, Jia YL, et al. SV40 intron, a potent strong intron element that effectively increases transgene expression in transfected Chinese hamster ovary cells [J]. *J Cell Mol Med*, 2018, 22 (4): 2231-2239.
- [80] Lu D, Li QY, Wu ZB, et al. High-level recombinant human lysozyme expressed in milk of transgenic pigs can inhibit the growth of *Escherichia coli* in the duodenum and influence intestinal morphology of sucking pigs [J]. *PLoS One*, 2014, 9 (2): e89130.
- [81] Luo Y, Liu J, Liu QQ, et al. Chicken hypersensitive site-4 insulator increases human serum albumin expression in bovine mammary epithelial cells modified with phiC31 integrase [J]. *Biotechnol Lett*, 2013, 35 (4): 529-537.
- [82] Naderi F, Hashemi M, Bayat H, et al. The augmenting effects of the tDNA insulator on stable expression of monoclonal antibody in Chinese hamster ovary cells [J]. *Monoclon Antib Immunodiagn Immunother*, 2018, 37 (5): 200-206.
- [83] Zhang JH, Zhang JH, Wang XY, et al. Distance effect characteristic of the matrix attachment region increases recombinant protein expression in Chinese hamster ovary cells [J]. *Biotechnol Lett*, 2020, 42 (2): 187-196.
- [84] Girod PA, Zahn-Zabal M, Mermod N. Use of the chicken lysozyme 5' matrix attachment region to generate high producer CHO cell lines

- [J]. *Biotechnol Bioeng*, 2005, 91 (1): 1-11.
- [85] Kim JD, Yoon Y, Hwang HY, et al. Efficient selection of stable Chinese hamster ovary (CHO) cell lines for expression of recombinant proteins by using human interferon beta SAR element [J]. *Biotechnol Prog*, 2005, 21 (3): 933-937.
- [86] Kim JM, Kim JS, Park DH, et al. Improved recombinant gene expression in CHO cells using matrix attachment regions [J]. *J Biotechnol*, 2004, 107 (2): 95-105.
- [87] Jia YL, Guo X, Wang XC, et al. Human genome-derived TOP1 matrix attachment region enhances transgene expression in the transfected CHO cells [J]. *Biotechnol Lett*, 2019, 41 (6/7): 701-709.
- [88] Tian ZW, Xu DH, Wang TY, et al. Identification of a potent MAR element from the human genome and assessment of its activity in stably transfected CHO cells [J]. *J Cell Mol Med*, 2018, 22 (2): 1095-1102.
- [89] Benton T, Chen T, McEntee M, et al. The use of UCOE vectors in combination with a preadapted serum free, suspension cell line allows for rapid production of large quantities of protein [J]. *Cytotechnology*, 2002, 38 (1-3): 43-46.
- [90] Saunders F, Sweeney B, Antoniou MN, et al. Chromatin function modifying elements in an industrial antibody production platform—comparison of UCOE, MAR, STAR and cHS4 elements [J]. *PLoS One*, 2015, 10 (4): e0120096.
- [91] Kwaks THJ, Barnett P, Hemrika W, et al. Identification of anti-repressor elements that confer high and stable protein production in mammalian cells [J]. *Nat Biotechnol*, 2003, 21 (5): 553-558.
- [92] Maksimenko OG, Deykin AV, Khodarovich YM, et al. Use of transgenic animals in biotechnology: prospects and problems [J]. *Acta Naturae*, 2013, 5 (1): 33-46.
- [93] Li S, Huang SH, Qiao SY, et al. Cloning and functional characterization of STAT5a and STAT5b genes in buffalo mammary epithelial cells [J]. *Anim Biotechnol*, 2020, 31 (1): 59-66.
- [94] Qian X, Zhao FQ. Regulatory roles of Oct proteins in the mammary gland [J]. *Biochim Biophys Acta*, 2016, 1859 (6): 812-819.
- [95] Song N, Luo J, Huang L, et al. Mutation of signal transducer and activator of transcription 5 (STAT5) binding sites decreases milk allergen α S1-casein content in goat mammary epithelial cells [J]. *Foods*, 2022, 11 (3): 346.
- [96] Kung MH, Lee YJ, Hsu JT, et al. A functional study of proximal goat β -casein promoter and intron 1 in immortalized goat mammary epithelial cells [J]. *J Dairy Sci*, 2015, 98 (6): 3859-3875.
- [97] Panigrahi A, O'Malley BW. Mechanisms of enhancer action: the known and the unknown [J]. *Genome Biol*, 2021, 22 (1): 108.
- [98] Galouzis CC, Furlong EEM. Regulating specificity in enhancer-promoter communication [J]. *Curr Opin Cell Biol*, 2022, 75: 102065.
- [99] Osterwalder M, Barozzi I, Tissières V, et al. Enhancer redundancy provides phenotypic robustness in mammalian development [J]. *Nature*, 2018, 554 (7691): 239-243.
- [100] Song W, Sharan R, Ovcharenko I. The first enhancer in an enhancer chain safeguards subsequent enhancer-promoter contacts.
- [101] Shin HY, Willi M, HyunYoo K, et al. Hierarchy within the mammary STAT5-driven wap super-enhancer [J]. *Nat Genet*, 2016, 48 (8): 904-911.
- [102] Hnisz D, Abraham BJ, Lee TI, et al. Super-enhancers in the control of cell identity and disease [J]. *Cell*, 2013, 155 (4): 934-947.
- [103] Lee Z, Raabe M, Hu WS. Epigenomic features revealed by ATAC-seq impact transgene expression in CHO cells [J]. *Biotechnol Bioeng*, 2021, 118 (5): 1851-1861.
- [104] Zeng XK, Lee HK, Wang CC, et al. The interdependence of mammary-specific super-enhancers and their native promoters facilitates gene activation during pregnancy [J]. *Exp Mol Med*, 2020, 52 (4): 682-690.
- [105] Borsari B, Villegas-Mirón P, Pérez-Lluch S, et al. Enhancers with tissue-specific activity are enriched in intronic regions [J]. *Genome Res*, 2021, 31 (8): 1325-1336.
- [106] Andersson R, Sandelin A. Determinants of enhancer and promoter activities of regulatory elements [J]. *Nat Rev Genet*, 2020, 21 (2): 71-87.
- [107] Corrales M, Rosado A, Cortini R, et al. Clustering of *Drosophila* housekeeping promoters facilitates their expression [J]. *Genome Res*, 2017, 27 (7): 1153-1161.
- [108] Zhu I, Song W, Ovcharenko I, et al. A model of active transcription hubs that unifies the roles of active promoters and enhancers [J]. *Nucleic Acids Res*, 2021, 49 (8): 4493-4505.
- [109] Dong WH, Li CP, Yang Y, et al. Increasing transgenic expression in recombinant Chinese hamster ovary cells using introns in different directions [J]. *Sheng Wu Gong Cheng Xue Bao*, 2019, 35 (6): 1071-1078.

- [110] Chorev M, Carmel L. The function of introns [J] . Front Genet, 2012, 3: 55.
- [111] Bieberstein NI, Carrillo Oesterreich F, Straube K, et al. First exon length controls active chromatin signatures and transcription [J] . Cell Rep, 2012, 2 (1) : 62-68.
- [112] Agarwal N, Ansari A. Enhancement of transcription by a splicing-competent intron is dependent on promoter directionality [J] . PLoS Genet, 2016, 12 (5) : e1006047.
- [113] Palazzo AF, Mahadevan K, Tarnawsky SP. ALREX-elements and introns: two identity elements that promote mRNA nuclear export [J] . Wiley Interdiscip Rev RNA, 2013, 4 (5) : 523-533.
- [114] Shaul O. How introns enhance gene expression [J] . Int J Biochem Cell Biol, 2017, 91 (Pt B) : 145-155.
- [115] Houdebine LM. Design of vectors for optimizing transgene expression [M] //Transgenic Animal Technology. Amsterdam: Elsevier, 2014: 489-511.
- [116] Qu GS, Piazza CL, Smith D, et al. Group II intron inhibits conjugative relaxase expression in bacteria by mRNA targeting [J] . eLife, 2018, 7: e34268.
- [117] Rose AB. Introns as gene regulators: a brick on the accelerator [J] . Front Genet, 2019, 9: 672.
- [118] Gallegos JE, Rose AB. The enduring mystery of intron-mediated enhancement [J] . Plant Sci, 2015, 237: 8-15.
- [119] Dwyer K, Agarwal N, Gega A, et al. Proximity to the promoter and Terminator regions regulates the transcription enhancement potential of an intron [J] . Front Mol Biosci, 2021, 8: 712639.
- [120] Romanova N, Noll T. Engineered and natural promoters and chromatin-modifying elements for recombinant protein expression in CHO cells [J] . Biotechnol J, 2018, 13 (3) : e1700232.
- [121] Reddi PP, Urekar CJ, Abhyankar MM, et al. Role of an insulator in testis-specific gene transcription [J] . Ann N Y Acad Sci, 2007, 1120: 95-103.
- [122] Chetverina D, Aoki T, Erokhin M, et al. Making connections: insulators organize eukaryotic chromosomes into independent cis-regulatory networks [J] . Bioessays, 2014, 36 (2) : 163-172.
- [123] Lu XB, Guo YH, Huang W. Characterization of the cHS4 insulator in mouse embryonic stem cells [J] . FEBS Open Bio, 2020, 10 (4) : 644-656.
- [124] Kisseljova NP, Dmitriev P, Katargin A, et al. DNA polymorphism and epigenetic marks modulate the affinity of a scaffold/matrix attachment region to the nuclear matrix [J] . Eur J Hum Genet, 2014, 22 (9) : 1117-1123.
- [125] Guo X, Wang C, Wang TY. Chromatin-modifying elements for recombinant protein production in mammalian cell systems [J] . Crit Rev Biotechnol, 2020, 40 (7) : 1035-1043.
- [126] Harraghy N, Calabrese D, Fisch I, et al. Epigenetic regulatory elements: recent advances in understanding their mode of action and use for recombinant protein production in mammalian cells [J] . Biotechnol J, 2015, 10 (7) : 967-978.
- [127] Zhao CP, Guo X, Chen SJ, et al. Matrix attachment region combinations increase transgene expression in transfected Chinese hamster ovary cells [J] . Sci Rep, 2017, 7: 42805.
- [128] Nematpour F, Mahboudi F, Vaziri B, et al. Evaluating the expression profile and stability of different UCOE containing vector combinations in mAb-producing CHO cells [J] . BMC Biotechnol, 2017, 17 (1) : 18.
- [129] Rocha-Pizaña MDR, Ascencio-Favela G, Soto-García BM, et al. Evaluation of changes in promoters, use of UCOES and chain order to improve the antibody production in CHO cells [J] . Protein Expr Purif, 2017, 132: 108-115.
- [130] Wang B, Guo Q, Liu LY, et al. Effect of interactions of chromatin regulatory elements with different promoters on the regulation of gene expression [J] . Chinese journal of biotechnology, 2021, 37 (9) : 3310-3322.
- [131] Mayr C. What are 3' UTRs doing? [J] . Cold Spring Harb Perspect Biol, 2019, 11 (10) : a034728.
- [132] Kim JJ, Yu J, Bag J, et al. Translation attenuation via 3' terminal codon usage in bovine *csn1s2* is responsible for the difference in α s2- and β -casein profile in milk [J] . RNA Biol, 2015, 12 (3) : 354-367.
- [133] Goodarzi H, Najafabadi HS, Oikonomou P, et al. Systematic discovery of structural elements governing stability of mammalian messenger RNAs [J] . Nature, 2012, 485 (7397) : 264-268.
- [134] Cohen-Zontag O, Baez C, Lim LQJ, et al. A secretion-enhancing cis regulatory targeting element (SECRete) involved in mRNA localization and protein synthesis [J] . PLoS Genet, 2019, 15 (7) : e1008248.
- [135] He Z, Song D, van Zalen S, et al. Structural determinants of human ζ -globin mRNA stability [J] . J Hematol Oncol, 2014, 7: 35.