

• 生物技术与方法 •

# 染色质调节元件与不同启动子的相互作用对基因表达调控的影响

王斌<sup>1</sup>, 郭庆<sup>1</sup>, 刘凌云<sup>2</sup>, 李美荃<sup>1</sup>, 戴利利<sup>1</sup>, 陈奇娜<sup>3</sup>, 刘雪莉<sup>3</sup>, 翟书华<sup>1</sup>

1 昆明学院 农学与生命科学学院, 云南 昆明 650224

2 河北省唐山市丰南区第一中学, 河北 唐山 063300

3 解放军联勤保障部队 第 920 医院, 云南 昆明 650031

王斌, 郭庆, 刘凌云, 等. 染色质调节元件与不同启动子的相互作用对基因表达调控的影响. 生物工程学报, 2021, 37(9): 3310-3322.

Wang B, Guo Q, Liu LY, et al. Effect of interactions of chromatin regulatory elements with different promoters on the regulation of gene expression. Chin J Biotech, 2021, 37(9): 3310-3322.

**摘要:** 为探究 DNA 序列元件对不同启动子调节转基因稳定表达的影响, 利用遍在染色质开放元件 (Ubiquitous chromatin opening elements, UCOE) 和基质黏附序列 (Scaffold/matrix-attachment regions, MAR) 分别与含增强子的 *oct4* 基因启动子、含 CpG 岛的 *sox2* 基因启动子和不含调控元件的 *nanog* 基因启动子以及同时包含增强子和 CpG 岛的 CMV 启动子组合构建 pOCT4-MAR、pOCT4-UCOE、pSOX2-MAR、pSOX2-UCOE、pNANOG-MAR、pNANOG-UCOE、pCMV-UCOE、pCMV-MAR 等质粒, 分析这些质粒稳定转染后的表达量和嵌合表达差异。结果发现, UCOE 与含增强子元件的 *oct4* 启动子组合能较稳定高效表达, 而 MAR 与含 CpG 岛的 *sox2* 启动子组合能较稳定高效表达。利用排除位置效应原因的嵌合表达对染色质高级结构调控基因表达的稳定性分析表明: (1) 通常情况下 UCOE 比 MAR 调节的表达载体的表达更高效和更稳定; UCOE 连接含 CpG 岛的启动子形成开放染色质调节的高表达更稳定; (2) MAR 与启动子上 TATA 盒或增强子可能通过染色质环产生高表达, 但相对不稳定。结论: 染色质调节元件 UCOE 和 MAR 与启动子调控元件之间能通过染色质开放状态或染色质环调控基因稳定表达。

**关键词:** 染色质调节序列, 启动子调控元件, 开放染色质, 染色质环

**Received:** November 25, 2020; **Accepted:** March 28, 2021

**Supported by:** National Natural Science Foundation of China (Nos. 31760256, 32060147), Joint Special Programme for Local Colleges and Universities in Yunnan Province, China (No. 202001BA070001-217), Teacher Research Project of Yunnan Education Department, China (No. 2020J0510).

**Corresponding author:** Shuhua Zhai. Tel: +86-871-65098532; E-mail: 937701466@qq.com

国家自然科学基金 (Nos. 31760256, 32060147), 云南省地方高校联合基金 (No. 202001BA070001-217), 云南省教育厅项目 (No. 2020J0510) 资助。

# Effect of interactions of chromatin regulatory elements with different promoters on the regulation of gene expression

Bin Wang<sup>1</sup>, Qing Guo<sup>1</sup>, Lingyun Liu<sup>2</sup>, Meiquan Li<sup>1</sup>, Lili Dai<sup>1</sup>, Qina Chen<sup>3</sup>, Xueli Liu<sup>3</sup>, and Shuhua Zhai<sup>1</sup>

<sup>1</sup> School of Agriculture and Life Sciences, Kunming University, Kunming 650224, Yunnan, China

<sup>2</sup> No. 1 High School, Fengnan District, Tangshan 063300, Hebei, China

<sup>3</sup> 920th Hospital of Joint Logistics Support Force, Kunming 650031, Yunnan, China

**Abstract:** The effect of altering the promoter region of ubiquitous chromatin-opening element (UCOE) and matrix attachment region (MAR) on stable and efficient expression of genes was investigated. Four different promoters were tested, namely, *oct4* containing an enhancer region, *sox2* having a CpG island, *nanog* having no regulatory elements, and CMV containing a CpG island and an enhancer region. Eight reporter plasmids were constructed: pOCT4-UCOE, pOCT4-MAR, pSOX2-UCOE, pSOX2-MAR, pNANOG-UCOE, pNANOG-MAR, pCMV-UCOE, and pCMV-MAR. Stable and efficient expression was observed when UCOE combined with the *oct4* promoter, whereas the *sox2* was the best promoter suited for MAR. Comparison of the stable clones of *oct4*-UCOE and *sox2*-MAR showed that UCOE-regulated expression is more stable and efficient than MAR-regulated expression. When CpG island-containing promoter is linked with UCOE, stable and efficient expression could be observed. These data suggest that an enhancer region in the promoter leads to high, yet unstable expression when combined with UCOE, whereas CpG islands stabilize expression. In conclusion, UCOE and MAR interact with regulatory elements on the promoter by altering the chromatin open state and chromatin loop to regulate gene expression.

**Keywords:** chromatin-modifying sequences, regulatory elements on promoter, opening chromatin, chromatin loop

重组治疗性抗体或蛋白药物 (统称重组药) 是重大疾病如癌症、类风湿关节炎、乙型肝炎和艾滋病等疾病的有效药物,其突出特点是研发和生产成本高,导致药品价格昂贵,因此降低研发和生产成本是重组药生产的迫切需求。作为研发和生产中的主要障碍,重组药细胞株在生产过程中的不稳定现象仍然有许多未解之谜,一般认为表达载体的整合位置<sup>[1]</sup>和表观遗传因素引起的转基因沉默<sup>[2-3]</sup>是重要原因。稳定的表观遗传机制包括组蛋白修饰(低乙酰化)、DNA 甲基化<sup>[4]</sup>等可遗传的表观标记和位置效应<sup>[2]</sup>,同时不稳定的染色质因素如非整合位置原因导致的染色质高级结构变化也可能是影响表达的重要原因,其中包括开放染色质层面的变化及染色质与染色质之间的交互作用。

开放染色质通常是细胞结构上呈解聚状态,位于间期核中心,在细胞周期的早期进行复制。其组蛋白高度乙酰化,相连 DNA 的 CpG 二核苷

酸通常无甲基化。用 DNase I 和其他内切酶进行分析时,发现有两类不同的染色质开放区域,分别是:长度为 200–600 bp 的位点开放染色质,比聚集区染色质敏感 2 倍 (DNase I 超敏位点) 以上;以及长展开放染色质 (大于 600 bp),比聚集区染色质敏感 10 倍以上<sup>[5]</sup>。基因的高效稳定表达受开放染色质的调控,但是如何调控开放染色质进而调控基因的稳定高效表达至今仍然不完全清楚。

调节开放染色质的 DNA 元件比较受关注,包括行使组织特异的或遍在基因表达模式的遗传调控元件<sup>[6]</sup>,如遍在染色质开放元件 (Ubiquitous chromatin opening elements, UCOE)<sup>[7]</sup>,以及通过物理阻隔形成开放染色质边界的绝缘子 (Insulator) 或障碍元件<sup>[6]</sup>。UCOE 是建立和维持转录专一开放染色质结构功能的遗传控制元件,是 Allen 等发现的 *HNRPA2B1-CBX3* 位点 (A2UCOE)

一段 1.5 kb 的片段<sup>[7]</sup>。UCOE 序列在与某些启动子结合时,通过抑制启动子区 DNA 甲基化克服转基因沉默<sup>[8]</sup>。

染色质上 DNA 环也可以看作染色质之间相互作用的一种方式。MAR (Scaffold/matrix-attachment regions)<sup>[9]</sup>是 DNA 环上与核骨架附着的 DNA 元件<sup>[10]</sup>。MAR 序列与核基质有着较高的亲和性,能将转基因与周围的基因组的负面作用隔绝开(绝缘功能),因此能克服转基因位置效应<sup>[11-12]</sup>。

染色质调节序列 MAR 和 UCOE 通过抑制转基因沉默和克服位置效应使转基因相对稳定地表达。然而在表达载体构建中,发现 MAR 和 UCOE 序列对不同启动子表达量的变化有着不同的影响,有些启动子与染色质调节元件结合对下游基因的改变很小,有些则改变较大,有较好的激活效应;且并非染色质调节序列元件与所有不同启动子结合都能有效地激活基因表达<sup>[13-14]</sup>。

对真核生物的强启动子特征进行分析,发现在转录起始位点 (Transcription start site, TSS) 附近存在许多转录调控元件参与染色质的调节,这些元件是主动调节关联启动子区域染色质重塑的 DNA 序列。如 TATA 盒是募集 RNA 酶复合体 (PIC) 的地方, TATA 盒和启动子区核小体位置对基因的表达水平影响较大<sup>[15]</sup>;而哺乳动物  $\beta$ -珠蛋白 ( $\beta$ -globin) 基因 TATA 盒与核小体迁移有关<sup>[16]</sup>。启动子上的 CpG 岛则是表观沉默研究的焦点, CpG 岛是 DNA 甲基化最密集的地方,也是通过去 DNA 甲基化调节基因转录的关键元件。高甲基化的 CpG 岛会有选择性地募集大量染色质,低甲基化的 CpG 岛区染色质常常形成半胱氨酸甲基化缺失、低水平的组蛋白 H1、高水平的组蛋白乙酰化及无核小体区域的 DNase I 的高敏感性位点<sup>[17]</sup>。启动子上的近端增强子也以不同的方式影响 K4、K9 和 K36 位的组蛋白甲基化。在增强子连接的基因编码区所有 3 个赖氨酸残基都高度三甲基化,而 K9 和 K36 位的二甲基化不受

增强子的影响;在增强子区域只在 K4 和 K9 处出现单甲基化,有别于基因编码区的组蛋白甲基化修饰<sup>[18]</sup>。

干细胞因子 Oct4、Sox2 和 Nanog 的基因是决定干细胞分化和成体细胞去分化的 3 个关键基因,其在成体细胞中几乎不表达,而在干细胞中则高表达,这 3 个基因的体内表达调控明显受环境引起的表观遗传因素的影响<sup>[19-20]</sup>。分析发现, *oct4* 基因启动子上主要包含增强子序列<sup>[21]</sup>, *sox2* 基因启动子上包含 TATA 盒和 CpG 岛<sup>[22-23]</sup>,而 *nanog* 基因启动子上除近端启动子区外,不包括任何顺式调节元件<sup>[23]</sup>。

本研究选择 *oct4* 启动子和 *sox2* 启动子为研究对象,构建 UCOE 和 MAR 与之相连的表达载体,分析 UCOE 和 MAR 分别与 TATA box、CpG 岛和近端增强子之间的相互作用对表达稳定和高效的影响,通过深入理解 UCOE 对启动子区域的 DNA 甲基化抑制机制和 MAR 的染色质环形成机制以及与 3 个启动子调节元件的相互作用,探讨染色质高级结构变化对转基因高效稳定表达的影响。

## 1 材料与方法

### 1.1 表达载体的构建

以商业载体 pEGFP-C1 (购自淼灵质粒平台)为模板进行改造。用 *Ase* I 和 *Age* I 酶同时将 pEGFP-C1 上的 CMV 启动子切除,替换为 *oct4* (Ensembl:ENSG00000204531)、*sox2* (Ensembl:ENSG00000181449) 基因的启动子中包含主要调控元件的部分序列 (0-1 500 bp),获得中间载体 pOCT4、pSOX2。分别在中间载体 pOCT4、pSOX2 和 *oct4* 启动子前端插入 MAR 序列和 UCOE 序列,获得 pOCT4-MAR, pOCT4-UCOE 和 pSOX2-MAR、pSOX2-UCOE;作为对照,以 *nanog* (Ensembl:ENSG00000111704) 和 pEGFP-C1 的 CMV 启动子构建 pNANOG-MAR、pNANOG-UCOE 和 pCMV-MAR、pCMV-UCOE 等质粒 (图 1)。

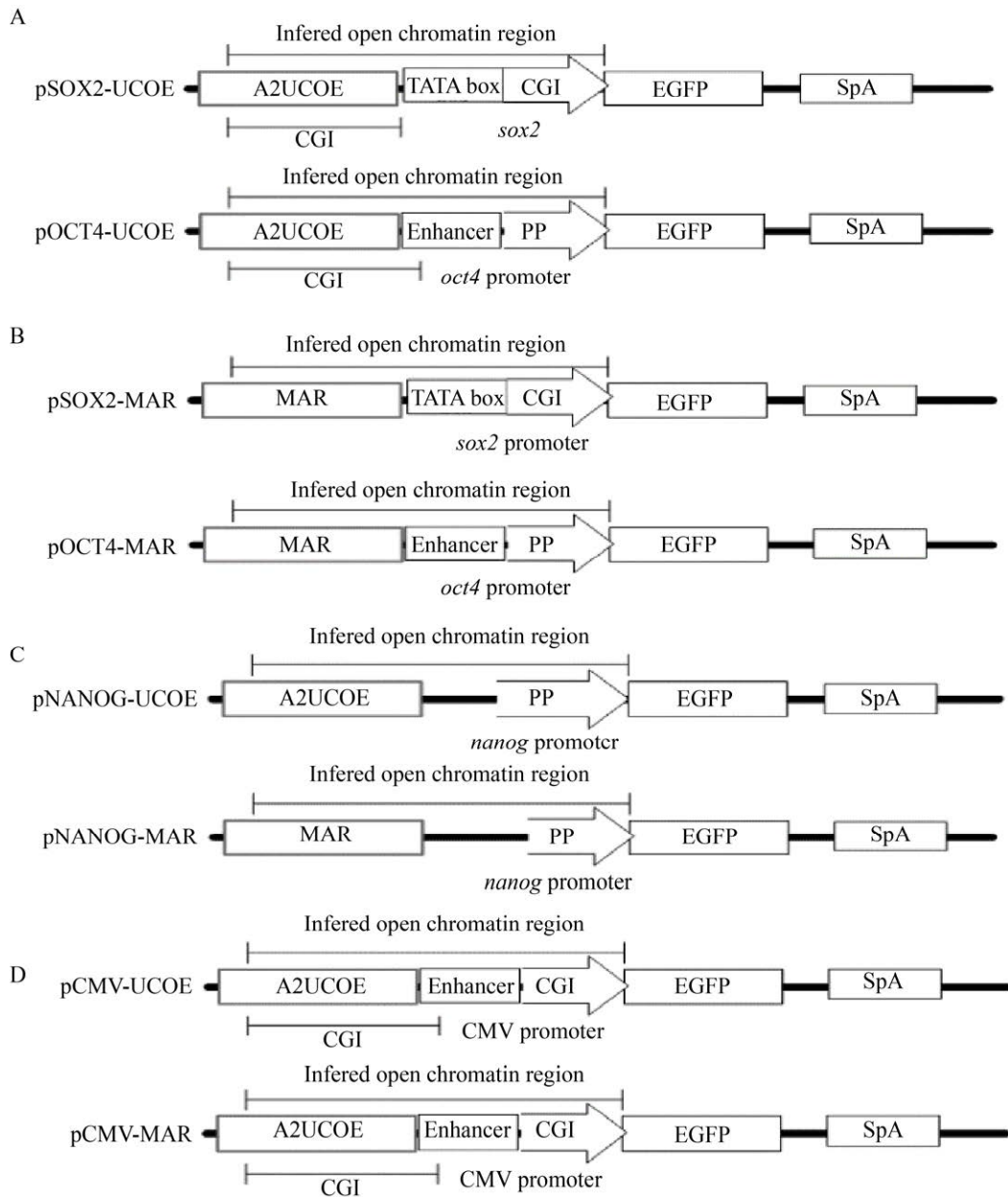


图1 UCOE和MAR分别与4种启动子 $oct4$ 、 $sox2$ 、 $nanog$ 和CMV结合的结构

Fig. 1 Schematic diagram of plasmids containing UCOE and MAR combined four different promoters:  $oct4$ ,  $sox2$ ,  $nanog$ , and CMV, respectively. CGI indicates CpG island, PP indicates proximal promoter, SpA indicates SV40 polyA.

## 1.2 转染及G418筛选

用电击转化仪 (BTX ECM 830) 转化, 电击杯中加入  $1 \times 10^6$  个细胞, 280 V、20 ms 电击一次。转染后培养 24 h, 待细胞密度增至 50%–70% 汇合时加 150  $\mu\text{g}/\text{mL}$  G418 进行筛选培养, 直至大量细

胞死亡后, 用 75  $\mu\text{g}/\text{mL}$  G418 继续培养至细胞克隆出现。

## 1.3 定量检测标记蛋白EGFP表达量

在所有这些改造的载体中, 通过分析其表达框上的增强型绿色荧光蛋白 (EGFP) 基因表达

来观察载体表达的效率。

### 1.3.1 荧光强度分析

利用荧光显微镜 (LEICA DMI 3000B), 统一拍照参数: 10 倍焦距; 70%绿光、蓝光和白光强度, 每孔从不同角度随机拍摄 4 张照片, 并用 ImageJ 软件进行荧光强度分析。

### 1.3.2 qRT-PCR 分析

提取细胞总 RNA, 用 Trizol 试剂 (Invitrogen) 进行 RNA 提取并用逆转录试剂盒 (Abm 公司, 中国江苏, Cat#G492) 逆转录为 cDNA。定量 PCR 反应 (Abm 公司, 中国江苏, Cat#MasterMix-S) 体系为 20  $\mu$ L, 内参基因用 *gapdh*, 引物均分别用 *egfp* 和 *gapdh* 基因的通用检测引物。将配制好的样品和内参在定量 PCR 仪 (EppendorfQX200) 上各做 3 个重复。使用  $2^{-\Delta\Delta C_t}$  方法分析基因表达相对变化。

### 1.3.3 ELISA 分析

用蛋白质含量为 1–10  $\mu$ g/mL 的小鼠 GFP 单抗 (Proteintech) 液。设置标准曲线对照 10 孔, 空白对照、阴性对照及阳性对照。4  $^{\circ}$ C 过夜进行包被。加 0.1 mL 待检样品, 37  $^{\circ}$ C 孵育 1 h, 洗涤后加酶标抗体 (Proteintech), 在各反应孔中加入新鲜稀释的酶标抗体 0.1 mL, 37  $^{\circ}$ C 孵育 0.5–1.0 h, 洗涤后加底物液显色反应, 在酶标仪上 (美国 Molecular Devices: SpectraMax M2) 用 450 nm 波长紫外光检测各孔  $OD_{450}$  值。

### 1.4 稳定表达和不同时间点荧光强度分析

所有转染上述质粒的 CHO 细胞, 经 G418 筛选 7–8 d 后形成稳定转染的细胞, 并在持续添加 75  $\mu$ g/mL G418 保持筛选压力。将稳转细胞按 1 000 个/孔铺到 6 孔板中, 等细胞克隆长出来后, 进行荧光强度分析。从转接到 6 孔板中开始计算, 每隔 5 d 全面观察检测一次荧光强度。分别计算 10 d、20 d 和 30 d 的荧光强度。

### 1.5 嵌合表达分析

分别对 pOCT4-MAR、pOCT4-UCOE、pSOX2-MAR、pSOX2-UCOE 等 4 个载体和 pNANOG-

MAR、pNANOG-UCOE 和 pCMV-MAR、pCMV-UCOE 等对照质粒的 24 个稳转细胞单克隆, 即每个细胞克隆 3 个重复, 用 Image J 软件中 Minimum、Maxentropy、Mean 等 3 种算法进行荧光强度分析。三种算法分别计算同一个样本中最强表达的荧光细胞、大于平均荧光强度的所有细胞和大于最小荧光强度的所有细胞, 计算 3 种不同细胞的荧光强度的标准误差, 来衡量每个样本中 *egfp* 基因嵌合表达多样性和差异程度, 标准误差越大, 则表明样本中嵌合表达程度越大, 每个稳转细胞样品之间的嵌合程度用 Origin2019b 进行 *t*-test 成对检验统计分析。

## 2 结果与分析

### 2.1 MAR 与 *sox2* 启动子或 UCOE 与 *oct4* 启动子调控高效表达

为了验证染色质调节元件 UCOE 和 MAR 与启动子上的顺式作用元件如 CpG 岛、增强子和 TATA 盒之间存在共同作用, 对构建的 pOCT4-MAR、pOCT4-UCOE、pSOX2-MAR、pSOX2-UCOE 等质粒载体进行稳定转染, 并对这些启动子启动的 *egfp* 基因通过 EGFP 基因荧光强度分析、qRT-PCR 分析和 ELISA 分析 (图 2), 结果显示, 上述 CHO 细胞中稳定转染 25 d 后, UCOE 调节的启动子, 其表达普遍获得提高, 其中 pOCT4-UCOE 中的 *egfp* 表达增强最为显著; MAR 调节的启动子对基因表达的影响普遍不显著。但例外的是, pSOX2-MAR 显著增强, 而 pSOX2-UCOE 则没有明显增强 (表 1, 图 2)。因为 OCT4 启动子由近端增强子和核心启动子组成, 因此推测 UCOE 与增强子组合可能增强基因高表达; 而 SOX2 启动子由 CpG 岛和 TATA 盒组成, 因此推测 MAR 与 CpG 岛或 TATA 盒之间的相互作用可能增强基因表达。进一步通过对照质粒分析发现, 用 MAR 调节含增强子和 CpG 岛的 CMV 启动子并不增强表达, 因此推测在 *sox2* 基因启动子上, MAR 可能更多与 TATA 盒相互作用或者

表 1 利用 MAR、UCOE 调节含增强子、CpG 岛、TATA 盒的启动子的载体及在 CHO 细胞中的表达  
 Table 1 The expression profiles in CHO cell with plasmids containing MAR or UCOE and different promoters. The promoter may contain enhancer, CpG island, TATA box, or their combinations

Vectors	Target promoter	Elements in promoters	Chromatin regulators	Effect of expression
pOCT4	<i>oct4</i> promoter	Enhancer	None	Control
pOCT4-MAR	<i>oct4</i> promoter	Enhancer	MAR	Insignificant
pOCT4-UCOE	<i>oct4</i> promoter	Enhancer	UCOE	Improved
pSOX2	<i>sox2</i> promoter	TATA box+CpG island	None	Control
pSOX2-MAR	<i>sox2</i> promoter	TATA box+CpG island	MAR	Improved
pSOX2-UCOE	<i>sox2</i> promoter	TATA box+CpG island	UCOE	Insignificant
pNANOG	<i>nanog</i> promoter	None	None	Control
pNANOG-UCOE	<i>nanog</i> promoter	None	UCOE	Improved
pNANOG-MAR	<i>nanog</i> promoter	None	MAR	Insignificant
pEGFP-C1	CMV early promoter	Enhancer+CpG island	None	Control
pCMV-UCOE	CMV early promoter	Enhancer+CpG island	UCOE	Improved
pCMV-MAR	CMV early promoter	Enhancer+CpG island	MAR	Insignificant

是两个元件之间同时起作用增强基因表达。而 UCOE 调节 CMV 启动子的表达增强并不特别显著,推测 UCOE 与 CpG 岛共同存在启动子区时可能对表达并无益处。

## 2.2 稳定转染后表达量变化分析

根据 2.1 中载体表达模式,为了探讨 UCOE 和 MAR 是否与含 TATA 盒、CpG 岛和增强子等启动子元件共同作用的表达规律,对转染了 pOCT4-MAR、pOCT4-UCOE、pSOX2-MAR、pSOX2-UCOE、pNANOG-MAR、pNANOG-UCOE 质粒的 CHO-k1 细胞荧光强度进行了连续的观察分析,并在转染后 10 d、20 d 和 30 d 三个时间点进行荧光拍照,以观察每个载体表达的稳定性(图 3),用 ImageJ 软件中 Mean 算法和 Maximum 算法对最强的荧光细胞强度和平均细胞荧光强度进行分析(图 4)。从平均细胞荧光强度来看,所有质粒表达的 EGFP 荧光强度 10–30 d 基本呈逐渐下降的趋势;但是对最强表达的荧光细胞分析表明,pOCT4-UCOE 载体能稳定持续表达达到 20 d 后表达量才开始下降;pSOX2-MAR 载体则在 10 d 稳转后表达量不断下降,但在 20 d 后表

达量开始上升,到 30 d 时表达量又恢复到与第 10 天持平的水平,变化幅度大且没有规律性;而 pNANOG-UCOE 载体则由稳转第 10 天时较低的表达量开始不断增强到 20 d 达到最高表达量,之后才开始下降。这提示 pOCT4-UCOE、pSOX2-MAR、pNANOG-UCOE 等 3 个载体在第 10 天时都不是相对高表达的载体,但在 20 d 后相对表达量都较高,表明 3 个载体表达量更稳定。

## 2.3 染色质调节元件与启动子相互作用调节的高效表达与稳定性关联分析

为了进一步理解 UCOE 和 MAR 分别与含 TATA 盒、CpG 岛和增强子等启动子元件的相互作用构成的开放染色质状态的稳定程度,对稳转后的单个细胞克隆的生长速度、表达强度差异及细胞分裂能力进行分析。以单个细胞克隆为研究对象的分析发现随机整合了同一个质粒的不同单细胞克隆生长速率不同(图 5A 左),由于不同单克隆质粒整合位置不同,其开放染色质状态也可能不同,提示外源启动子的开放染色质状态可能影响细胞生产速率。同时质粒整合于同一单克隆中不同细胞荧光强度也不同(图 5A 中),提示染色质调

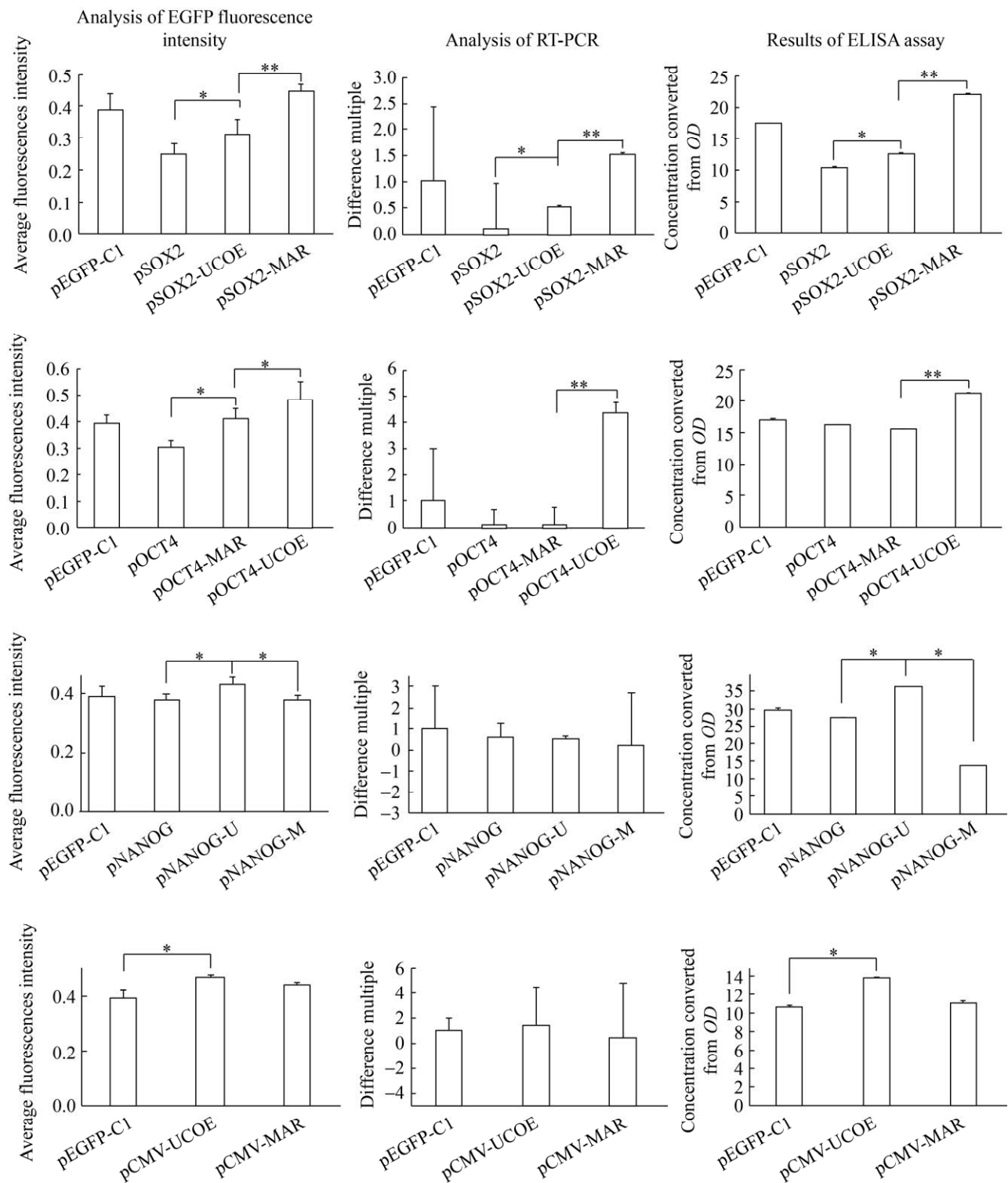


图2 稳定转染 25 d 以后, 由 TATA 盒和 CpG 岛组成的 *sox2* 启动子 (第一排) 和由含增强子的 *oct4* 启动子 (第二排) 被染色质元件 UCOE 和 MAR 调控后, 获得的增强表达模式相反

Fig. 2 After 25 days of stable transfection, *sox2* promoter (first row) and *oct4* promoter (second row) were regulated by chromatin elements UCOE and MAR, respectively, and showed different enhanced expression patterns.

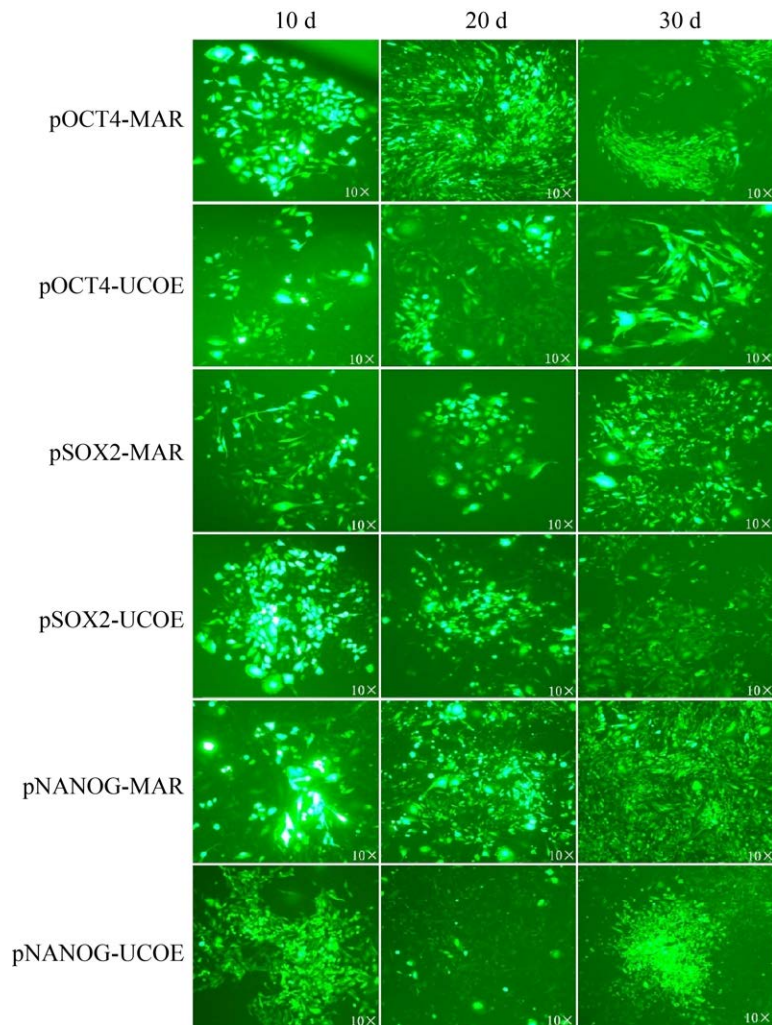


图3 依据 UCOE 和 MAR 与包含不同调节元件启动子的相互作用构建的表达载体分别在 10 d、20 d、30 d 的表达分析

Fig. 3 Gene expression profiles in CHO-K1 cells at time points of 10 d, 20 d and 30 d. These plasmids contain UCOE or MAR and promoters with different regulatory elements.

节元件与启动子相互作用形成的开放染色质在整合位置相同的情况下，仍然还受其他因素的影响。另外同一单克隆中存在荧光强度特别高的细胞，其细胞分裂能力显著下降（图 5A 右），提示某些与细胞分裂相关联的位置可能存在延展的开放染色质。

以相同整合位置的单克隆为研究对象，对其表现出较大的表达差异定义为单克隆嵌合表达。通过分析每一个质粒的单克隆嵌合表达程度，作为比较各个质粒之间的嵌合表达差异的依据，以评价整合位置相对固定的前提下对不同质粒的开

放染色质结构变化规律。分别对 pOCT4-MAR、pOCT4-UCOE；pSOX2-MAR、pSOX2-UCOE；pNANOG-MAR、pNANOG-UCOE 和 pCMV-MAR、pCMV-UCOE 等 8 个质粒的 24 个单克隆，即每个质粒取 3 个不同时间点的稳转单克隆为分析对象，每个克隆重复进行 Minimum、Maxentropy、Mean 等 3 种算法的荧光强度分析，3 种算法分别取不同荧光强度的细胞进行计算，以每一个单克隆中 3 种算法获得的荧光强度之间的标准误差表示细胞嵌合表达程度，标准误差越高表示其嵌合表达程度越高。



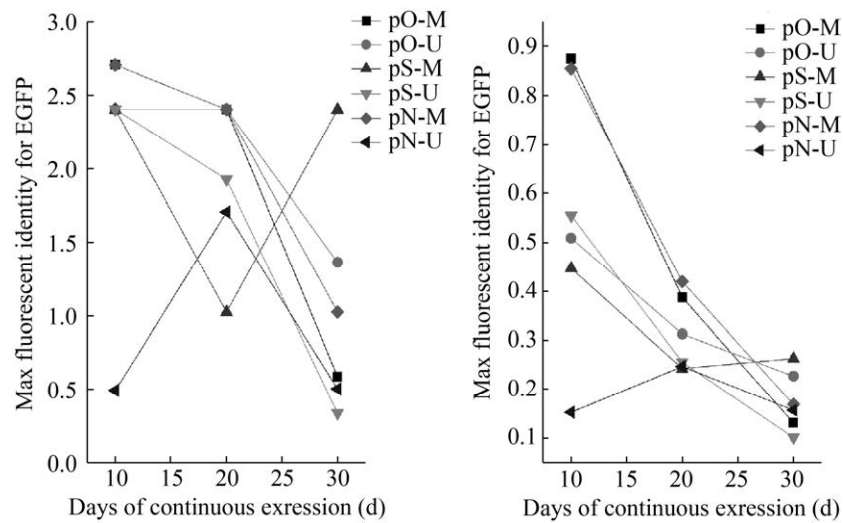


图4 依据UCOE和MAR与带不同调节元件的启动子发生相互作用所构建的表达载体在CHO细胞中稳转单克隆的嵌合表达分析

Fig. 4 Chimeric expression analysis of monoclones that stably transfected plasmids behalf of interaction between UCOE/MAR with different promoters.

统计分析结果表明,UCOE或MAR与含CpG岛的 $sox2$ 启动子和CMV启动子共同作用时启动的嵌合表达程度差异不显著;而与不含CpG岛、仅含增强子的 $oct4$ 启动子和不带任何调节元件的 $nanog$ 启动子共同作用时,则差异显著(图5B-C),提示启动子上CpG岛可能具有稳定UCOE或MAR调节的开放染色质的作用。另外,比较UCOE和MAR对嵌合表达的影响发现,除只含增强子的 $oct4$ 启动子中pOCT4-UCOE嵌合表达程度显著高于pOCT4-MAR外,其余3个启动子中UCOE调节的单克隆嵌合表达程度都低于带MAR序列的质粒,提示UCOE具有降低嵌合表达的作用;而增强子在嵌合表达中发挥不同于其他调节元件的作用。同时对两组嵌合表达差异显著的质粒进行比较时发现,带增强子的 $oct4$ 启动子和不含任何调节元件的 $nanog$ 启动子表现出完全相反的表达模式,即对 $oct4$ 启动子来说,pOCT4-UCOE嵌合表达程度显著高于pOCT4-MAR;对 $nanog$ 启动子来说,pNANOG-UCOE的嵌合表达程度显著低于pNANOG-MAR(图5B-C),如果UCOE具有降低嵌合表达的作用,则pOCT4-

UCOE的高嵌合表达进一步说明增强子可能是引起高嵌合表达的关键调节元件。

对比8个质粒的单克隆嵌合表达差异和基因高效表达关联分析发现,对 $oct4$ 启动子来说,UCOE调节的基因高效表达导致高嵌合,而对 $sox2$ 启动子被MAR调节的高效表达与高嵌合的关联不显著;对 $nanog$ 启动子(无调控元件)和CMV(CpG岛+增强子)启动子分析发现,UCOE调节两个启动子高效表达都呈低嵌合表达(图5D-E;表2)。以上结果表明,染色质调节序列UCOE调节的高效表达较为稳定,只有 $oct4$ 启动子是例外。但相较于 $oct4$ 启动子上由UCOE+增强子调节的高表达高嵌合,MAR+TATA box+CpG岛调节高表达和低嵌合,表现出更好的稳定性。同时CMV启动子上UCOE+增强子+CpG岛组合则也调节高表达和低嵌合,由UCOE+ $nanog$ 同样调控高表达和低嵌合,说明UCOE普遍调节高表达和低嵌合。而UCOE+增强子调节的高表达例外地出现高嵌合,提示增强子调节的高表达可能不稳定,同时UCOE+增强子+CpG岛组合则调节高表达和低嵌合,说明CpG岛的存在可能抑制了增强

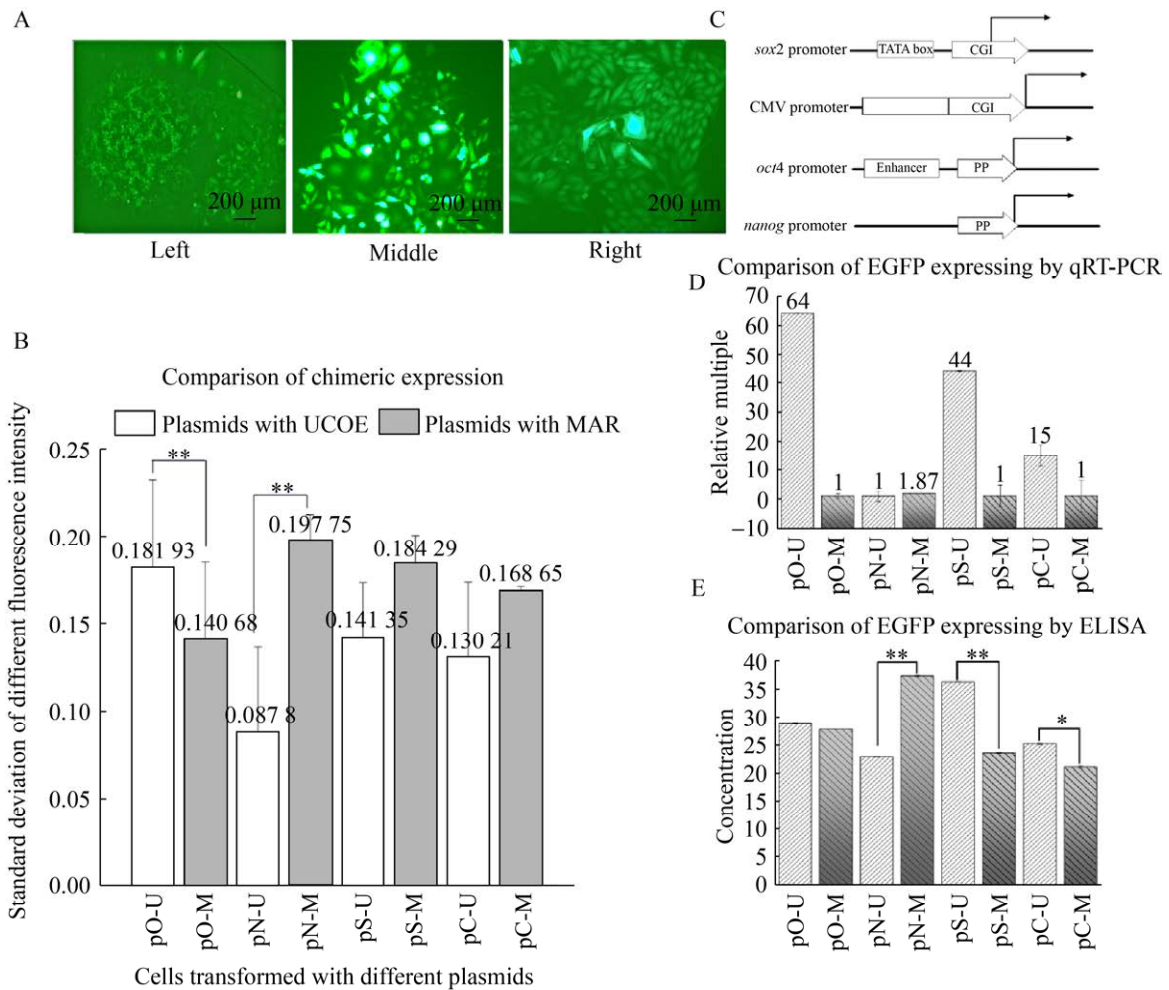


图 5 质粒稳定转染后单克隆细胞中出现的嵌合表达现象

Fig. 5 Chimeric expression occurred in monoclonal cells transfected with the plasmids. (A) Left: different monoclonal cells showed different growth rates; Middle: variable fluorescence intensity appeared in different cells in the same monoclonal cell; Right: cells showed different dividing ability in the same monoclonal cell. (B) Analysis of chimeric expression in cells transfected with different plasmids. (C) Structures of four different promoters. (D–E) Analysis of *egfp* expression driven by different promoters in different plasmids.

表 2 8 个质粒的单克隆嵌合表达与基因表达的关系

Table 2 The relationship between chimeric expression and gene expression in cells transfected with 8 plasmids

Plasmids	Regulatory elements in promoters	Chimeric expression	Expression level	Stability of opening chromatin
pOCT4-UCOE	UCOE+enhancer	High	High	Unstable
pOCT4-MAR	MAR+enhancer	Low	Low	Stable
pSOX2-UCOE	CpG island+TATA box+UCOE	Insignificant	Low	Relatively stable
pSOX2-MAR	CpG island+TATA box+MAR	Insignificant	High	Relatively unstable
pNANOG-UCOE	UCOE	Low	High	Stable
pNANOG-MAR	MAR	High	Low	Unstable
pCMV-UCOE	UCOE+enhancer+CpG island	Low	High	Stable
pCMV-MAR	MAR+enhancer+CpG island	High	Low	Unstable

子带来的不稳定。综合低表达高嵌合的情况判断(表2),通常情况下UCOE调节的高效表达比MAR更稳定;但当MAR与含TATA盒的*sox2*启动子结合时也能高效且相对稳定地表达,由于上述分析提示的CpG岛在稳定表达中发挥关键作用,而*sox2*启动子上TATA盒只与MAR结合时在稳定开放染色质中起作用。

### 3 讨论

#### 3.1 MAR或UCOE序列通过与启动子调控元件共同作用调控稳定表达

转录水平的调控是基因表达调控的首要环节,MAR和UCOE对基因表达的调控主要针对转录水平的调控。真核生物转录由起始、延伸和终止等环节组成,转录的起始由RNA聚合酶II复合体的形成和起始开始,经典理论阐明了启动子上的调控元件在辅助RNA聚合酶II募集、形成中的作用,RNA聚合酶II的形成取决于染色质的解聚、拓扑结构的改变等环节,因此染色质水平的改变,是转录调控的重要组成部分。但是在更深入的染色质调控研究上,仍有许多待探索的空间,如开放染色质的精细调控方面仍然未有定论。MAR和UCOE序列通过不同的机制形成开放染色质调控基因表达,目前已成为哺乳动物细胞稳定表达不可或缺的重要元件。

本研究表明,MAR与含CpG岛和TATA盒的*sox2*基因启动子结合后,其载体pSOX2-MAR在稳定转染后的表达稳定性显著强于UCOE与*sox2*启动子的组合。这可能是由于正常情况下*sox2*启动子TSS位置上含有CpG岛,启动子上的CpG岛不易被甲基化因而在起始位点形成稳定的核小体缺失区域<sup>[24]</sup>,是基因稳定转录的基础,而UCOE的作用是通过阻止启动子上CpG位点甲基化<sup>[8]</sup>,因此在*sox2*启动子上加入UCOE在功能上重复且不会发挥叠加作用;而MAR通过形成染色质环<sup>[25]</sup>可改善*sox2*启动子上有利于转录的核小体占位。UCOE和MAR的不同干预中,MAR

除了稳定含TATA盒的*sox2*启动子的pSOX2-MAR高表达外,对包括含增强子的*oct4*启动子、不含调节序列的*nanog*和同时含CpG及增强子的CMV启动子均低于UCOE的调节效果,提示TATA盒可能与MAR有相互作用。TATA盒促进形成RNA聚合酶复合体,其中中间蛋白(Mediator)可参与染色质环的形成影响染色质重塑<sup>[22,26]</sup>,反过来MAR形成染色质环可能通过中间蛋白与TATA盒发生相互作用。

而另一方面,UCOE与含增强子的*oct4*启动子结合,其载体pOCT4-UCOE表达稳定性优于MAR与*oct4*启动子组合。这可能是由于*oct4*启动子主要包含近端增强子和Sp1结合区的近端启动子,因此UCOE的甲基化抵抗作用能够对去除启动子上CG的甲基化从而减少核小体的占位,进而增强表达;而MAR对启动子近端增强子区H3K9上乙酰化的阻止,可能是妨碍其发挥转录增强效应<sup>[27]</sup>的原因。

该结果意味着MAR或UCOE序列对启动子上染色质的调节受到启动子上的调控元件影响。具体来说,MAR更有利于TSS区含CpG岛的启动子上形成核小体占位的合理分布,甚至是核小体缺失;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。即启动子上调控元件与MAR或UCOE相互作用调节染色质的改变。

#### 3.2 排除位置效应的表达不稳定可能归因于染色质高级结构的动态变化

转基因表达中因为位置效应导致的嵌合表达是主要障碍,位置效应本质是染色质构象问题。通过分析整合位置一致的单克隆中细胞嵌合表达,则反映的是相同染色质构象背景下的动态变化,往往是外源载体稳定性的本质。

对pOCT4-MAR、pOCT4-UCOE、pSOX2-MAR、pSOX2-UCOE及其对照pNANOG-MAR、pNANOG-UCOE和pCMV-MAR、pCMV-UCOE等几个载体中的单克隆细胞嵌合表达表明,

UCOE 与启动子组合的载体普遍稳定, 而 MAR 与启动子的组合, 无论是低表达还是高表达都相对不稳定。对高表达的 pOCT4-UCOE、pSOX2-MAR 和 pNANOG-UCOE 在稳定转染后随时间点发生的表达变化来看, 高效表达的载体都呈现较大波动, 尤其 pSOX2-MAR 的表达量变化尤其波动较大且呈现无规律性。MAR 通过形成三维结构染色质环增强基因高表达<sup>[9]</sup>, pSOX2-MAR 可能通过 TATA 盒和 MAR 共同形成染色质环, 染色质环可能相对不稳定。真核生物中染色质三维结构的形成是相对瞬时的状态, 但这种高级结构的改变可能促成启动子上形成稳定的转录复合体, 以及通过改变核小体位置形成二维构象的开放染色质, 并最终导致组蛋白或 DNA 的修饰等稳定的表观标记的形成, 从而对基因转录产生稳定的影响<sup>[28]</sup>。UCOE 通过与 *oct4* 启动子和 *nanog* 启动子形成的高效表达载体表现出的相对稳定, 可能是 UCOE 介导的非甲基化 GC 更有利于稳定的核小体缺失即开放染色质的形成, 含 CpG 岛的载体也通过去甲基化 CG 形成更稳定的开放染色质。

### 3.3 染色质高级结构对表达不稳定的影响

MAR 和 TATA 盒与染色质 DNA 环的形成相关, 同时通过比较发现染色质 DNA 环形成的高表达不稳定。而另一方面含增强子的启动子表达也不稳定。最近大量的研究表明, 活跃的增强子 (Active enhancer) 可以像启动子一样独立地启动自身转录 RNA, 命名为 eRNA<sup>[28]</sup>。eRNA 通过结合中间蛋白、黏连蛋白和装载因子 Nipbl (Nipped-B-like protein) 形成黏连蛋白相关的增强子-启动子染色质 DNA 环, 将增强子置于与靶基因启动子接近的三维位置<sup>[26,28]</sup>, 从而在三维结构层面激活基因转录、起始和延伸。通过 MAR 与 TATA 盒可能涉及的染色质 DNA 环和同样涉及染色质 DNA 环的高级结构对表达的不稳定现象, 表明染色质 DNA 环对表达的影响相对 DNA 去甲基化带来的开放染色质更稳定。

通过上述的分析和讨论, 可以初步归纳染色质高级结构对表达不稳定的影响为: (1) TATA 盒和增强子可能受染色质 DNA 环的调控, 因此容易与 MAR 组合形成高表达, 同时染色质 DNA 环形成的高表达相对不稳定。(2) UCOE 通过抑制启动子区 CG 的甲基化尤其 CpG 岛的去甲基化形成的开放染色质可能更有利于高效表达的稳定。

但是由于以上结论的部分证据仅来自数据的推导, 以上结论仍需进一步获得直观和直接的实验证据进行进一步确认。

**致谢:** 感谢中国医学科学院医学生物学研究所胡云章研究员、孙静副研究员和施建东副研究员在该项目早期给予的帮助和支持。感谢中美瑞康核酸技术 (南通) 研究院有限公司李龙承教授在该项目研究中给予的鼓励和支持。

## REFERENCES

- [1] Lee JS, Kallehauge TB, Pedersen LE, et al. Site-specific integration in CHO cells mediated by CRISPR/Cas9 and homology-directed DNA repair pathway. *Sci Rep*, 2015, 5: 8572.
- [2] Veith N, Ziehr H, Macleod RAF, et al. Mechanisms underlying epigenetic and transcriptional heterogeneity in Chinese hamster ovary (CHO) cell lines. *BMC Biotechnol*, 2016, 16: 6.
- [3] Kim M, O'Callaghan PM, Droms KA, et al. A mechanistic understanding of production instability in CHO cell lines expressing recombinant monoclonal antibodies. *Biotechnol Bioeng*, 2011, 108(10): 2434-2446.
- [4] Osterlehner A, Simmeth S, Göpfert U. Promoter methylation and transgene copy numbers predict unstable protein production in recombinant Chinese hamster ovary cell lines. *Biotechnol Bioeng*, 2011, 108(11): 2670-2681.
- [5] Higgs DR. Do LCRs open chromatin domains? *Cell*, 1998, 95(3): 299-302.
- [6] Antoniou M, Harland L, Mustoe T, et al. Transgenes encompassing dual-promoter CpG islands from the human *TBP* and *HNRPA2B1* loci are resistant to heterochromatin-mediated silencing. *Genomics*, 2003, 82(3): 269-279.

- [7] Allen ML, Antoniou M. Correlation of DNA methylation with histone modifications across the HNRPA2B1-CBX3 ubiquitously-acting chromatin open element (UCOE). *Epigenetics*, 2007, 2(4): 227-236.
- [8] Zhang F, Frost AR, Blundell MP, et al. A ubiquitous chromatin opening element (UCOE) confers resistance to DNA methylation-mediated silencing of lentiviral vectors. *Mol Ther*, 2010, 18(9): 1640-1649.
- [9] Loc PV, Strätling WH. The matrix attachment regions of the chicken lysozyme gene co-map with the boundaries of the chromatin domain. *EMBO J*, 1988, 7(3): 655-664.
- [10] Heng HHQ, Goetze S, Ye CJ, et al. Chromatin loops are selectively anchored using scaffold/matrix-attachment regions. *J Cell Sci*, 2004, 117(Pt 7): 999-1008.
- [11] Goetze S, Baer A, Winkelmann S, et al. Performance of genomic bordering elements at predefined genomic loci. *Mol Cell Biol*, 2005, 25(6): 2260-2272.
- [12] Bode J, Benham C, Knopp A, et al. Transcriptional augmentation: modulation of gene expression by scaffold/matrix-attached regions (S/MAR elements). *Crit Rev Eukaryot Gene Expr*, 2000, 10(1): 73-90.
- [13] 王斌, 刘凌云, 范志伟, 等. MAR 序列调控弱启动子构建高效表达载体的方法: 中国, 110863008B. 2020-10-16.  
Wang B, Liu LY, Fan ZW, et al. Method for constructing efficient expression vector by regulating and controlling weak promoter through MAR sequence: CN, 110863008B. 2020-10-16 (in Chinese).
- [14] 王斌, 刘凌云, 殷旭东, 等. 高 CG 片段与低 CG 启动子联用构建高效表达载体的方法: 中国, 110804626A. 2020-02-18.  
Wang B, Liu LY, Yin XD, et al. Method for constructing efficient expression vector through combination of high CG (cholyglycine) fragment and low CG promoter: CN, 110804626A. 2020-02-18 (in Chinese).
- [15] Hornung G, Bar-Ziv R, Rosin D, et al. Noise-mean relationship in mutated promoters. *Genome Res*, 2012, 22(12): 2409-2417.
- [16] Gui CY, Dean A. A major role for the TATA box in recruitment of chromatin modifying complexes to a globin gene promoter. *Proc Natl Acad Sci USA*, 2003, 100(12): 7009-7014.
- [17] Lin JC, Jeong S, Liang GN, et al. Role of nucleosomal occupancy in the epigenetic silencing of the *MLH1* CpG island. *Cancer Cell*, 2007, 12(5): 432-444.
- [18] Rentoft M, Kim K, Cho Y, et al. Enhancer requirement for histone methylation linked with gene activation. *FEBS J*, 2008, 275(23): 5994-6001.
- [19] Santourlidis S, Wernet P, Ghanjati F, et al. Unrestricted somatic stem cells (USSC) from human umbilical cord blood display uncommitted epigenetic signatures of the major stem cell pluripotency genes. *Stem Cell Res*, 2011, 6(1): 60-69.
- [20] Bem J, Grabowska I. The alchemy—epigenetic regulation of pluripotency. *Postepy Biochem*, 2013, 59(2): 144-156.
- [21] Kellner S, Kikyo N. Transcriptional regulation of the *oct4* gene, a master gene for pluripotency. *Histol Histopathol*, 2010, 25(3): 405-412.
- [22] Wiebe MS, Wilder PJ, Kelly D, et al. Isolation, characterization, and differential expression of the murine Sox-2 promoter. *Gene*, 2000, 246(1/2): 383-393.
- [23] Wang B, Sun J, Shi JD, et al. Small-activating RNA can change nucleosome positioning in human fibroblasts. *J Biomol Screen*, 2016, 21(6): 634-642.
- [24] Collings CK, Anderson JN. Links between DNA methylation and nucleosome occupancy in the human genome. *Epigenet Chromatin*, 2017, 10: 18.
- [25] Razin SV, Iarovaia OV, Sjakste N, et al. Chromatin domains and regulation of transcription. *J Mol Biol*, 2007, 369(3): 597-607.
- [26] Barrero MJ, Malik S. The RNA polymerase II transcriptional machinery and its epigenetic context//Kundu T, Ed. *Epigenetics: Development and Disease*. Dordrecht: Springer, 2013, 61: 237-259.
- [27] 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. *Eur J Hum Genet*, 2014, 22(9): 1117-1123.
- [28] Dekker J. Gene regulation in the third dimension. *Science*, 2008, 319(5871): 1793-1794.

(本文责编 郝丽芳)