

# 猪早期胚胎发育中 *SOX2* 基因启动子活性分析

赵剑超, 柴壮, 郭诗萌, 刘忠华

东北农业大学生命科学学院, 哈尔滨 150030

**摘要:** 转录因子 *SOX2* (sex determining region Y-box2) 在早期胚胎发育第一次谱系分化以及内细胞团多能性维持等方面具有重要的作用。但是目前有关 *SOX2* 基因启动子的系统研究较少, 尤其是在猪 (*Sus scrofa*) 中尚无相关报道。为系统分析猪 *SOX2* 基因启动子在早期胚胎中的活性, 本研究通过优化显微注射体系中绿色荧光蛋白表达载体种类和注射时间, 构建了适合猪 *SOX2* 基因启动子活性分析的参照体系和显微注射体系; 对猪 *SOX2* 基因翻译起始位点上游 5000 bp 区域进行转录因子结合位点的预测, 发现该区域存在 4 个转录因子结合位点簇; 针对上述区域设计并构建相应的缺失型 *SOX2* 基因启动子报告载体, 利用建立的显微注射体系将其导入胚胎, 通过 mCherry 荧光强度以及 qRT-PCR 定量分析 *SOX2* 基因启动子中不同转录因子结合位点簇对启动子活性的影响。结果表明, 与全长 *SOX2* 基因启动子相比, *SOX2* 基因翻译起始位点上游 2254~2442 bp 区域缺失后, 猪 4-细胞和 8-细胞胚胎中 *SOX2* 基因启动子活性下降至 17.8%, 该缺失区域中仅包含两个 NF-AT(nuclear factor of activated T cells) 转录因子结合位点。因此, 本研究结果推测猪 *SOX2* 基因启动子中 NF-AT 转录因子结合位点是影响猪早期胚胎 *SOX2* 基因启动子活性的关键位点。本研究为揭示猪早期胚胎发育中 *SOX2* 基因表达调控机制提供了数据支持。

**关键词:** *SOX2* 基因; 启动子活性; 猪

## Analysis of *SOX2* gene promoter activity in porcine early embryonic development

Jianchao Zhao, Zhuang Chai, Shimeng Guo, Zhonghua Liu

College of Life Science, Northeast Agricultural University, Harbin 150030, China

**Abstract:** *SOX2* (sex determining region Y-box2) is one of the critical pluripotent factors that play a crucial role in the first lineage differentiation and maintenance of pluripotency in inner cell mass during early embryonic development. However, there are few researches about the regulation of the *SOX2* promoter, especially in *Sus scrofa*. To analyze the activity of *SOX2* promoter in early porcine embryos, we determined the control system and established the microinjection system for assessing *SOX2* promoter activity by analyzing the embryonic development and the expression of enhanced

收稿日期: 2019-02-21; 修回日期: 2019-06-17

基金项目: 国家自然科学基金面上项目(编号: 31872360)资助[Supported by the National Natural Science Foundation of China (No. 31872360)]

作者简介: 赵剑超, 硕士研究生, 专业方向: 发育生物学。E-mail: arabidopsis@yeah.net

通讯作者: 刘忠华, 教授, 博士生导师, 研究方向: 早期胚胎发育与猪干细胞研究。E-mail: liuzhonghua@neau.edu.cn

DOI: 10.16288/j.ycz.19-045

网络出版时间: 2019/7/5 14:07:50

URI: <http://kns.cnki.net/kcms/detail/11.1913.R.20190705.1407.002.html>

green fluorescence protein (EGFP) after micro-injected different EGFP plasmids at different times after activation of the oocytes. Then, we analyzed the structure of 5000 bp upstream of the *SOX2* translation initiation site and found there were four transcription factor binding site clusters. Next, we designed and constructed promoter-containing plasmids to analyze the function of each cluster. To detect the activity of different promoters, we assessed the mCherry expression in protein levels and mRNA levels by analyzing the mCherry fluorescence intensity and qRT-PCR after injecting plasmids into embryos. These results showed that the activity of the shorted promoter, with the region from 2254 bp to 2442 bp upstream of translation initiation site deleted, decreased to 17.8% at 4-cell and 8-cell stages compared with the full-length promoter. This region included two NF-AT transcription factor binding sites, which indicated that the NF-AT binding site is a key region to regulate the activity of the *SOX2* promoter. The results provide important data for determination the mechanism of porcine *SOX2* regulation.

**Keywords:** *SOX2* gene; promoter activity; pig

*SOX2* 蛋白作为重要的转录因子, 广泛表达于胚胎、原始生殖细胞、神经元等多种组织与细胞中。在胚胎发育过程中, *SOX2* 基因的表达是大脑皮层形成和发育所必需的, 同时也参与指导食道与前胃的形成<sup>[1]</sup>。此外, *SOX2* 基因在多种癌细胞与肿瘤干细胞中广泛表达<sup>[2~4]</sup>。近年来, *SOX2* 转录因子在早期胚胎发育和维持干细胞多能性方面的作用受到广泛关注<sup>[5~7]</sup>。

在哺乳动物早期胚胎发育中, 第一次谱系分化发生在囊胚腔形成时期, 胚胎由全能性的卵裂球分化形成多潜能性的滋养层细胞(trophectoderm, TE)和内细胞团细胞(inner cell mass, ICM)。然而, 关于第一次谱系分化形成滋养层和内细胞团细胞的机制目前尚不清楚。*SOX2* 蛋白作为核心多能性转录因子之一, 在早期胚胎发育中第一次谱系分化以及内细胞团多能性维持等方面具有重要的作用<sup>[8,9]</sup>。2003 年, Avilion 等<sup>[10]</sup>证实小鼠(*Mus musculus*)*Sox2* 基因敲除的囊胚中仍存在 ICM, 但导致其丧失多能性, 同时失去产生胚胎干细胞的能力。2014 年, Wicklow 等<sup>[11]</sup>发现小鼠中 *Sox2* 基因的表达不受 *CDX2* 蛋白调控, 且 *Sox2* 基因在内细胞团中特异性表达并具有重要功能。在猪早期胚胎发育中, *SOX2* 转录因子备受关注。2015 年, Liu 等<sup>[12]</sup>研究表明, 虽然 *SOX2* 基因在猪内细胞团中特异性表达, 但表达模式与小鼠不同: (1)在早期囊胚内细胞团前体细胞中即出现 *SOX2* 基因特异性表达, 但此时 *CDX2* 基因尚未表达; (2)在囊胚期, 滋养层细胞中 *SOX2* 基因表达沉默先

于极化的发生, 而小鼠中由于细胞的极化导致滋养层细胞中 *Sox2* 基因表达沉默。表达模式的差异暗示猪 *SOX2* 基因表达上游调控机制不同于小鼠。此外, 在猪诱导性多能性干细胞(induced pluripotent stem cells, iPSC)的建系研究中, 利用 *SOX2* 蛋白在内的经典 4 因子对体细胞重编程, 虽然可以获得具有一定多能性的 iPS 细胞系<sup>[13~16]</sup>, 但猪 iPSC 同时具有内源基因不激活、外源基因不沉默以及不具嵌合能力等问题<sup>[17]</sup>。同时, 参考小鼠和人胚胎干细胞系建系的经验, 至今仍无法获得猪胚胎干细胞系<sup>[18,19]</sup>。上述研究表明, 猪与小鼠、人等多能性调控网络存在差异, 建立猪多能性相关调控网络机制具有重要的科学意义。

本研究通过检测 *SOX2* 基因启动子介导的报告基因表达水平评价其启动活性, 预测了启动子区域转录因子结合位点, 分析了各区域对 *SOX2* 基因启动子活性的影响, 进而揭示猪早期胚胎发育中 *SOX2* 基因表达调控关键因子, 为进一步研究猪早期胚胎第一次谱系分化以及多能性干细胞建立提供理论基础<sup>[20]</sup>。

## 1 材料与方法

### 1.1 质粒载体构建

本研究所用的质粒包括: pEGFP-C1 (Clontech, 美国), pRL-TK (Promega E2241, 美国), FUGW

(Addgene 14883, 美国)和 pLV-mCherry (Addgene 36084, 美国)。载体构建中片段克隆所用的 DNA 聚合酶为 PrimeSTAR<sup>®</sup> Max DNA Polymerase (TaKaRa R045A, 日本), 反应体系为 25  $\mu$ L, 反应条件中退火温度设定为 58  $^{\circ}$ C, 时间 15 s; DNA 片段连接所用

的试剂盒为 NEBuilder<sup>®</sup>高保真 DNA 组装克隆试剂盒(NEB E2621S, 美国), 引物重复序列长度 25 bp, 20  $\mu$ L 反应体系中共加入 0.2 pmol DNA 片段总量, 反应时间 15 min。PCR 和 qRT-PCR 所用全部引物信息见表 1。

表 1 引物信息

Table 1 Primer sequences used in this study

引物名称	引物序列(5'→3')	用途
UBC pro-p*	GATCTGGCCTCCGCGCCGGGTTTTG GTCTAACAAAAAGCCAAAAACG	克隆 UBC 启动子片段
CMV pro-p*	CGCGTTACATAACTTACGGTAAA AGCTCTGCTTATATAGACCTCCC	克隆 CMV 启动子片段
EGFP-p*	ATGGTGAGCAAGGGCGAGGAGCT CACTTGTACAGCTCGTCCATGC	克隆 EGFP 片段
mCherry-p*	ATGGTGAGCAAGGGCGAGGAGGA CTACTTGTACAGCTCGTCCATGC	克隆 mCherry 片段
pRL line-p	TTCTAGAGCGGCCGCTTCGAGCA GTCGAGCCATGTGAGCAAAAGGC	线性化 pRL-TK 质粒
pEGFP line-p	GGTTTAGTGAACCGTCAGATCCG GAACTCCATATATGGGCTATGAA	线性化 pEGFP-C1 质粒
C1 line-p	CCGGGATCCACCGATCTAGATA TGGGCTATGAACTAATGACCCCG	线性化 pEGFP-C1 质粒
sox2+-p	GTGGGAGATCCTTCACAGTCTA GCTTCTCCGTCTCCGACAAA	克隆猪 SOX2 基因上游序列
sox2-p*	TCTGGAAGTCTTGTGGCATGATC GCGGGCGCTGTGCGCGGGCCCGG	克隆 sox2 pro 5k 片段
sox2 pro[1]-p*	TCTGGAAGTCTTGTGGCATGATC CGCCTTTTCGAAGGAAGTGGGTAAAC	克隆 sox2 pro [1]的线性报告载体
sox2 pro[2]-p*	GCACCACTGAGATCCCTCCTTTA AACAAAGCCCCGTCTAAGTTTCC	克隆 sox2 pro [2]的线性报告载体
sox2 pro[3]-p*	ACCTATCGAGTTTGACAAGAACTG TGGTACTCTTCAAAATAAAGCAGTG	克隆 sox2 pro [3]的线性报告载体
sox2 pro[4]-p*	ACCATTATCCACCAACCCCTTG GCGGGCGCTGTGCGCGGGCCCGG	克隆 sox2 pro [4]的线性报告载体
sox2-q	GCGGAGTGGAACCTTTTGTCCT GGGAAGCGTGACTTATCCTTCT	猪 SOX2 qRT-PCR 扩增引物
18S-q	GATGGGCGGCGGAAAATTG TCCTCAACACCACATGAGCA	猪 18S rRNA qRT-PCR 扩增引物
EGFP-q	AGTGCTTCAGCCGCTACCC AGTTCACCTTGATGCCGTTCTT	EGFP qRT-PCR 扩增引物
mCherry-q	CCCGAGGGCTTCAAGTGGGA CCGTCCGAGGGGAAGTTGGT	mCherry qRT-PCR 扩增引物

\*引物序列中未标出用于无缝克隆的重复序列。

以 FUGW 质粒为模板, 利用引物 UBC pro-p 克隆得到泛素 C (ubiquitin-C, UBC) 启动子片段; 以 pEGFP-C1 质粒为基础载体, 利用引物 pEGFP line-p 克隆得到线性化的 pEGFP-C1 质粒载体; 将 UBC 启动子片段和线性化的 pEGFP-C1 质粒载体连接, 获得 pEGFP-UBC 质粒。pEGFP-C1 和 pEGFP-UBC 质粒特点: 以 pEGFP-C1 质粒为基础骨架的 CMV (cytomegalovirus) 或 UBC 启动子介导的 EGFP 表达载体 (图 1, A 和 B; 表 2)。

以 pEGFP-C1 质粒为模板, 利用引物 CMV pro-p 和 EGFP-p 分别克隆得到 CMV 启动子片段和增强绿色荧光蛋白 (enhanced green fluorescence protein, EGFP) 片段; 以 pRL-TK 质粒为基础载体, 利用引物 pRL line-p 克隆得到线性化的 pRL-TK 质粒载体; 将 CMV 启动子片段、EGFP 片段和线性化的 pRL-TK 质粒载体连接, 获得 pRL-CMV 质粒; 将 UBC 启动子片段、EGFP 片段和线性化的 pRL-TK 质粒载体连接, 获得 pRL-UBC 质粒。其特点为: 以 pRL-TK 质粒为基础骨架的 CMV 或 UBC 启动子介导的 EGFP 表达载体

(图 1, C~E; 表 2)。

以猪基因组为模板, 利用引物 *sox2*+p 进行克隆; 再以克隆产物为模板, 利用引物 *sox2*-p 克隆得到 *SOX2* 翻译起始位点上游 5000 bp 片段, 即本研究中 *SOX2* 基因全长启动子, 命名为 *sox2* pro 5k。以 pLV-mCherry 质粒为模板, 利用引物 mCherry-p 克隆得到 mCherry 片段; 以 pEGFP-C1 质粒为载体骨架, 利用引物 C1 line-p 克隆得到线性化的载体骨架; 将 *sox2* pro 5k 片段、mCherry 片段和线性化的载体骨架连接, 获得质粒 Sp0, 其特点为: 以 pEGFP-C1 质粒为基础骨架的 5000 bp *SOX2* 基因启动子介导的 mCherry 表达载体。

以 Sp0 质粒为模板, 分别以 *sox2* pro[1]-p、*sox2* pro[2]-p、*sox2* pro[3]-p 和 *sox2* pro[4]-p 为引物进行克隆, 将得到的克隆产物直接自连, 分别获得质粒 Sp1、Sp2、Sp3 和 Sp4, 其特点为: 分别缺失相应结合位点簇的 *SOX2* 基因启动子介导的 mCherry 表达载体。

构建完成的质粒去内毒素提取后-20℃保存。

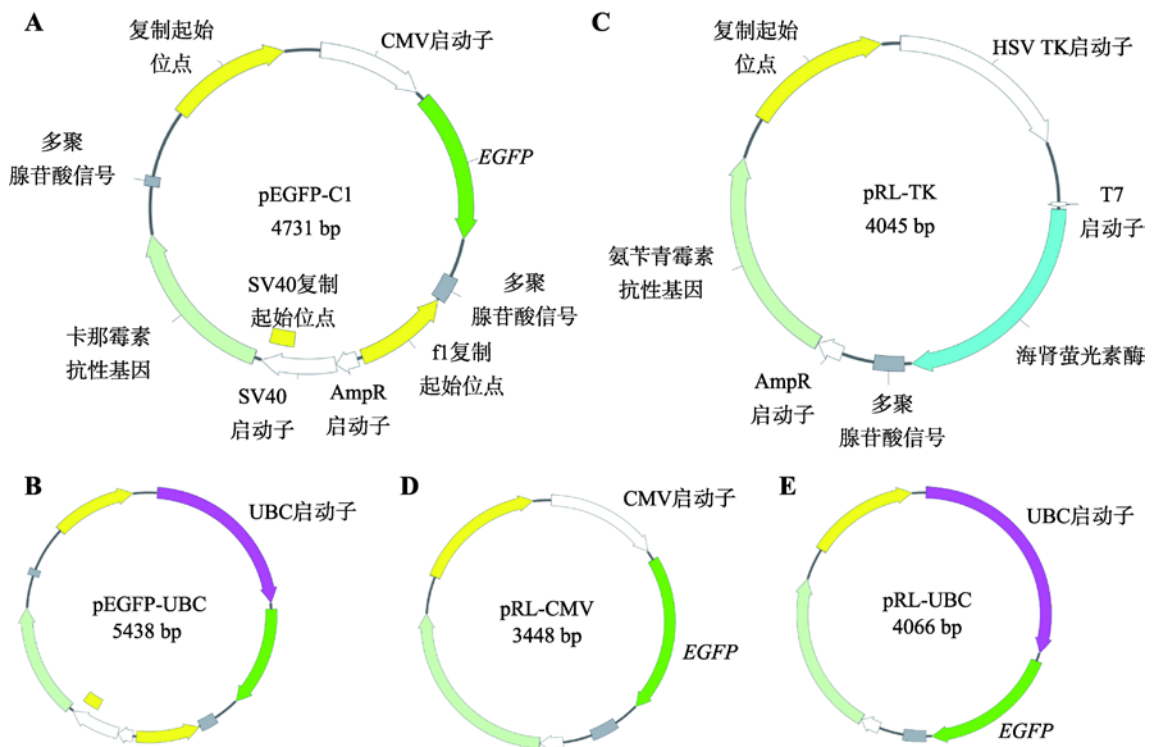


图 1 报告质粒图谱

Fig. 1 Map of reporter plasmids

A, B: 报告质粒 pEGFP-C1 和 pEGFP-UBC 构建图谱; C~E: pRL-TK 质粒和 pRL-CMV、pRL-UBC 报告质粒构建图谱。

表 2 报告质粒信息对比

Table 2 Comparison of reporter plasmid information

名称	大小(bp)	质粒骨架	启动子	报告基因
pEGFP-C1	4731	pEGF-C1	CMV	EGFP
pEGFP-UBC	5348	pEGF-C1	UBC	EGFP
pRL-CMV	3448	pRL-TK	CMV	EGFP
pRL-UBC	4066	pRL-TK	UBC	EGFP

## 1.2 胚胎培养与显微注射

从屠宰场收集猪卵巢, 在 37 ℃ 生理盐水中抽取直径 2~8 mm 卵泡。将获得的卵泡液静置, 弃上清, 并用 HEPES 缓冲液清洗稀释后, 在 39 ℃ 恒温工作台上借助显微镜挑出有 3 层以上卵丘细胞包裹的卵母细胞, 50 枚一组培养在 500 μL 预平衡的成熟培养液(MAT)中, 条件为 39 ℃、5% CO<sub>2</sub> 的饱和湿度环境。避光静置培养 42 h 后, 将卵母细胞置于 0.5% 透明质酸酶中震荡去除卵丘细胞, 在 39 ℃ 恒温工作台上借助显微镜挑出具有第一极体的 MII 期卵母细胞。将获得的 MII 期卵母细胞置于细胞融合液(FM)中排成一行, 附加两次 120 V/mm、30 μs 的矩形脉冲以完成孤雌激活, 之后 50 枚一组培养在 500 μL 预平衡的胚胎培养液(PZM-3), 条件为 39 ℃、5% CO<sub>2</sub> 的饱和湿度环境中。

显微操作系统为倒置显微镜(Nikon TS-100, 日本)、显微注射系统(Nikon IM-9A, 日本)及 Femtojet Express 定量注射系统(Eppendorf, 德国)。取 100 μL 胚胎操作液(MAN)于平皿中央, 并用石蜡油覆盖后放置于显微操作仪 39 ℃ 恒温工作台上, 向 MAN 液滴中加入卵母细胞准备注射。将 5 μL 质粒溶液通过微量上样枪头加至注射针内进行注射操作。注射系统相关参数为: 注射压力(Pi)100 hPa、注射时间(Ti)0.5 s 和补偿压力(Pc)40 hPa。显微镜下对猪卵母细胞进行胞质注射, 注射量约为 30 pL/枚。

## 1.3 启动子信息获取与预测

在 NCBI 数据库中获取猪 *SOX2* 基因上游序列。选择翻译起始位点上游 5000 bp 序列, 以翻译起始位点第一个腺嘌呤脱氧核糖核苷酸为+1, 上游为负进行定位标识。利用 LASAGNA-Search 2.0 启动子预测数据库([http://biogrid-lasagna.engr.uconn.edu/lasagna\\_](http://biogrid-lasagna.engr.uconn.edu/lasagna_)

search/)对 5000 bp 序列进行转录因子结合位点预测。Matrix-Derived Models 参数选择 TRANSFAC Matrix 模型, 将预测结果中  $P < 1 \times 10^{-5}$  的结果列为候选位点(表 3)。

表 3 sox2 pro 5k 片段中预测的转录因子结合位点

Table 3 The predict transcription factor binding sites in sox2 pro 5k sequence

序号	名称	位置	备注
1	NF-AT	-470~-459	位于 d1 区域
2	NF-Y	-481~-467	
3	NF-Y	-482~-469	
4	GC box	-546~-533	
5	SREBP-1	-678~-668	
6	NF-AT	-1336~-1325	位于 d2 区域
7	C/EBP	-1337~-1325	
8	Ik-3	-1472~-1460	
9	STAT5B	-1473~-1461	
10	STAT5A	-1473~-1461	
11	STAT5A	-1474~-1461	位于 d3 区域
12	AP-2rep	-1496~-1490	
13	STAT5A	-1503~-1496	
14	Lmo2 complex	-1601~-1591	
15	AREB6	-1601~-1590	
16	NF-kappaB	-1654~-1641	位于 d4 区域
17	SOX-9	-1762~-1749	
18	YY1	-1864~-1848	
19	NF-AT	-2271~-2260	
20	NF-AT	-2432~-2421	
21	Tal-1beta:ITF-2	-3109~-3094	位于 d4 区域
22	Tal-1alpha:E47	-3109~-3094	
23	Tal-1alpha:E47	-3109~-3094	
24	TATA	-3194~-3185	
25	Freac-7	-3415~-3400	
26	BSAP	-3512~-3485	
27	IRF-1	-3747~-3735	
28	c-Ets-1(p54)	-3854~-3845	
29	Lhx3	-4075~-4066	
30	FOXO1	-4235~-4227	
31	CHX-10	-4262~-4252	
32	CHX-10	-4263~-4253	
33	MRF-2	-4288~-4278	
34	E4BP4	-4494~-4483	

## 1.4 启动子活性分析

利用荧光报告基因的平均荧光强度与其 mRNA 相对含量评估相应启动子活性。对注射后 72 h 胚胎通过 CCD 数字成像系统采集荧光镜下图像信息, 利用软件 Adobe Photoshop CC 2015 分析 mCherry 和 EGFP 平均荧光强度。各组分别收集 EGFP 阳性 4-细胞和 8-细胞胚胎共 100 枚, 利用酸性操作液(1 mL MAN 加 20  $\mu$ L 浓盐酸)去除透明带后置于 RNase-free 1.5 mL 离心管, 液氮冷冻后-80 保存。对收集的胚胎样本利用 Purelink RNA mini kit (Ambion 12183018A, 美国)提取总 RNA, 利用 High Capacity cDNA Reverse Transcription Kit (Applied Biosystems 4368814, 美国)进行总 RNA 反转录, 获得 cDNA, 利用 SYBR premix Ex Taq™ (TaKaRa RR420, 日本)和引物 mCherry-q、EGFP-q、sox2-q、18S-q 分别对 *mCherry*、*EGFP*、*SOX2* 和猪 18S rRNA 进行 qRT-PCR 扩增。数据处理时, 以猪 18S rRNA 基因作为内参校正 *SOX2* 基因检测结果 ( $\Delta$ SOX2), 以 *EGFP* 基因校正 *mCherry* 基因检测结果 ( $\Delta$ mCherry), 再用  $\Delta$ SOX2 对  $\Delta$ mCherry 进行二次校正后作为最终 qRT-PCR 的检测结果, 以此评价各启动子活性。具体 qRT-PCR 数据处理公式为: 相对表达量 =  $2^{-[(\text{实验组} \Delta C_{t_{mCherry}} - \Delta C_{t_{SOX2}}) - (\text{对照组} \Delta C_{t_{mCherry}} - \Delta C_{t_{SOX2}})]}$ , 其中  $\Delta C_{t_{mCherry}} = C_{t_{mCherry}} - C_{t_{EGFP}}$ ,  $\Delta C_{t_{SOX2}} = C_{t_{SOX2}} - C_{t_{18S rRNA}}$ ,  $C_t$  为 qRT-PCR 中的阈值循环。

## 2 结果与分析

### 2.1 猪胚胎显微注射体系的建立

#### 2.1.1 报告载体的筛选

对猪胚胎显微注射体系中报告载体的类型进行筛选, 以获得高效表达外源报告基因的载体。选择两种载体骨架以及两种广谱型启动子共计 4 种载体作为备选报告载体, 分别是 pEGFP-C1、pEGFP-UBC、pRL-CMV 和 pRL-UBC (图 1)。孤雌激活后 4 h 用上述 4 种质粒对胚胎进行显微注射, 注射质粒浓度为 20 ng/ $\mu$ L。观察 EGFP 蛋白表达情况以及胚胎发育情况(图 2, A 和 B)。结果表明, 注射 pEGFP-C1 质粒

的胚胎 EGFP 阳性率为 39.33% ( $\pm 1.76\%$ ), 显著高于其他组(图 2, C 和 D)。由于早期胚胎发育会受外源 DNA 浓度的影响, 因此又分别选择 20 ng/ $\mu$ L 与 100 ng/ $\mu$ L 浓度的 pEGFP-C1 进行显微注射。经观察发现, 当注射浓度提高至 100 ng/ $\mu$ L 时, 注射胚胎发育情况与 20 ng/ $\mu$ L 浓度组无显著差异(图 2B), 但 EGFP 表达效果更好(图 2, C 和 D)。因此, 在后续研究中选择 pEGFP-C1 质粒为内参和基本载体骨架, 质粒总浓度为 100 ng/ $\mu$ L。

#### 2.1.2 注射时间选择

为了确定猪胚胎显微注射体系中最佳注射时间, 分别针对孤雌前卵母细胞(-0 h)与孤雌激活 0 h、2 h、4 h、6 h、8 h、10 h 以及 11 h 后胚胎显微注射 pEGFP-C1 质粒, 观察各组胚胎发育与 EGFP 蛋白表达情况(图 3)。结果表明, 不同注射时间对于胚胎发育与 EGFP 蛋白表达影响较大。其中-0 h 组对胚胎发育影响最小, 但 EGFP 蛋白表达水平最低, 而 2 h、4 h 与 6 h 组对胚胎发育影响较小, 同时 EGFP 蛋白高水平表达(图 3, A 和 B)。综合胚胎发育情况及 EGFP 蛋白表达水平, 选择孤雌后 2~6 h 作为最佳注射时间。此外, 观察发现胚胎在 4-细胞期开始表达 EGFP 蛋白, 在 8-细胞期 EGFP 蛋白表达阳性胚胎比例达到最高。因此, 在后续 *SOX2* 基因启动子活性检测研究中选择孤雌后 72 h 收集 4-细胞和 8-细胞胚胎进行检测(图 3 C)。

### 2.2 启动子结构预测分析与报告载体构建

本研究对猪 *SOX2* 基因翻译起始位点上游 5000 bp 序列进行转录因子结合位点预测分析。结果发现, 该区域中转录因子结合位点成簇分布于 4 个区域, 以与翻译起始位点的距离分别命名 d1~d4(图 4 A, 表 3)。针对 4 个结合位点簇分别设计并构建相应缺失的 *SOX2* 基因启动子报告载体, 将构建的启动子片段分别命名为 sox2 pro [1~4], 构建的质粒分别命名为 Sp1~4 (图 4, B 和 C)。同时针对猪 *SOX2* 基因翻译起始位点上游 5000 bp 序列, 与小鼠进行相似性比对分析, 结果显示, 序列相似点的排布近似斜率为 1 的直线, 相似性为 71% (图 4 D)。因此, 小鼠的相关研究结果在猪中具有一定参考价值。

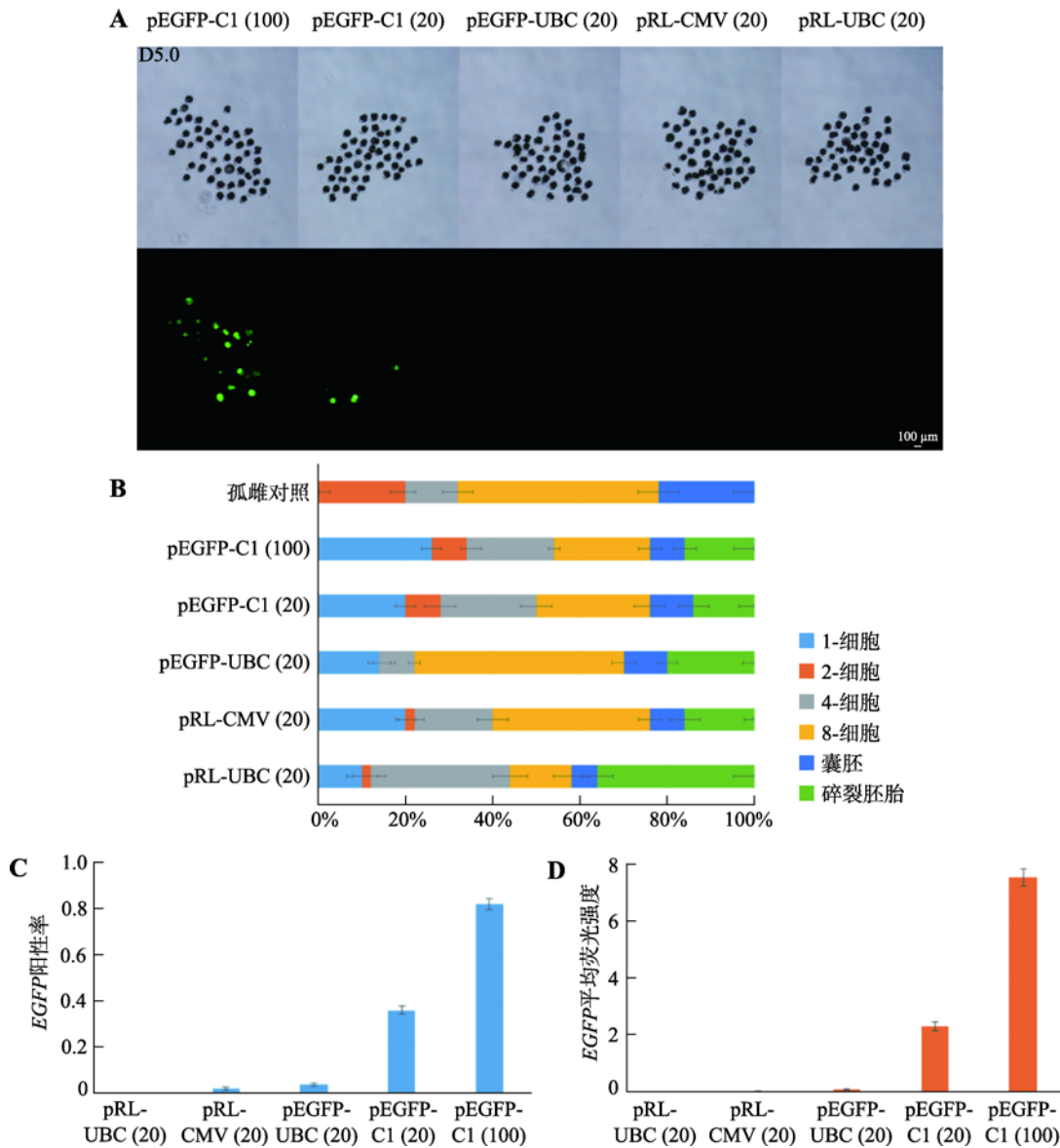


图 2 4 种质粒对胚胎荧光标记的效果

Fig. 2 The fluorescent labeling effects of four plasmids on embryos

A: 用 4 种质粒分别对孤雌激活 4 h 后的胚胎进行注射, 培养 5 d 后荧光镜下观察 EGFP 基因表达情况; B: 以孤雌胚胎为对照, 各注射组培养 5 d 内胚胎各阶段发育情况统计; C: 各注射组在培养 5 d 后胚胎 EGFP 基因表达阳性率; D: 各注射组在培养 5 d 后胚胎 EGFP 平均荧光强度。括号中 20 和 100 分别代表质粒浓度为 20 ng/μL 和 100 ng/μL。

### 2.3 启动子活性检测

将质粒 Sp0、Sp1、Sp2、Sp3、Sp4 与内参质粒 pEGFP-C1 以 4 : 1 的比例混合, 终浓度为 100 ng/μL。对猪胚胎进行注射并培养 72 h, 以检测 mCherry 以及 EGFP 蛋白的表达情况(图 5 A)。结果显示, 与 Sp0 注射组相比, Sp1 注射组胚胎无法检测到 mCherry 蛋白信号; Sp3 与 Sp4 注射组 mCherry 蛋白信号显

著下降; 而 Sp2 注射组胚胎 mCherry 蛋白水平与 Sp0 注射组无显著差异(图 5 B)。进一步利用 qRT-PCR 对各组胚胎的 mCherry 基因相对表达量进行定量分析。结果显示, 相对于 Sp0 注射组胚胎, 各组胚胎 mCherry 基因相对表达量均有不同程度的下降, 但与荧光检测结果不同的是, Sp3 注射组胚胎 mCherry 基因的 mRNA 表达水平最低, 仅为 Sp0 注射组的  $0.18(\pm 0.0034)$  (图 5 C)。据此推断, Sp3 质粒中



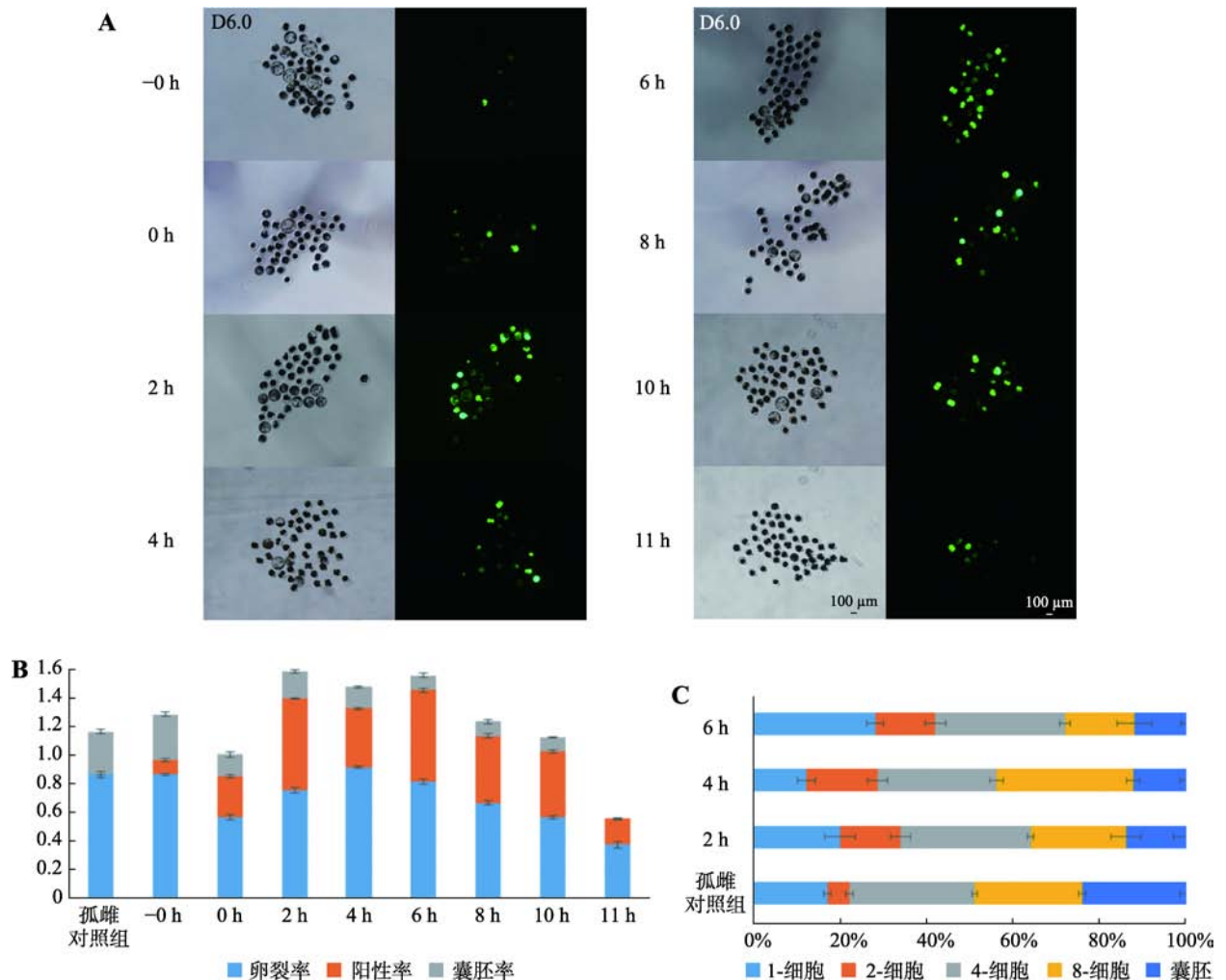


图 3 孤雌激活前后不同注射时间的荧光标记效果

**Fig. 3 The fluorescent labeling effects of different injection times of parthenogenetic activation**

A: 分别在孤雌前(-0 h)和孤雌后 0 h、2 h、4 h、6 h、8 h、10 h、11 h 注射 pEGFP-C1 质粒, 培养 6 d 后观察 *EGFP* 基因表达情况; B: 各注射组胚胎 *EGFP* 基因表达阳性率及发育率统计; C: 2 h、4 h 和 6 h 注射组在培养 6 d 时间内胚胎各阶段发育情况统计。

缺失的转录因子结合位点簇 d3 可能为 *SOX2* 基因启动子活性关键区域。根据转录因子结合位点预测结果, d3 中仅包含两个 NF-AT 转录因子结合位点, 推测 NF-AT 转录因子家族是调节猪早期胚胎中 *SOX2* 基因表达的关键转录因子。

### 3 讨论

*SOX2* 蛋白是哺乳动物早期胚胎发育中关键转录因子之一, 对于 *SOX2* 基因上游调控网络的研究可以揭示早期胚胎发育中细胞命运决定、胚胎干细胞多能性维持等相关分子机制<sup>[21]</sup>。本研究通过对猪

早期胚胎中 *SOX2* 基因启动子活性分析, 发现 NF-AT 转录因子结合位点是 4-细胞与 8-细胞胚胎中调控 *SOX2* 基因表达的关键顺式调控元件。NF-AT 是一个钙离子敏感的转录因子家族, 在内皮细胞、骨骼肌细胞、平滑肌细胞和神经元等多种细胞中广泛表达<sup>[22]</sup>。但是 NF-AT 转录因子在早期胚胎发育中的作用功能尚不明确, 同时 NF-AT 蛋白调控 *SOX2* 基因表达的相关研究尚未报道。因此, 研究 NF-AT 转录因子对于猪早期胚胎发育的调控功能具有重要意义。

对 Sp1 注射组的报告基因检测时发现, 其蛋白水平与 mRNA 水平并不一致。在小鼠中, *Sox2* 基因启动子无“TATA box”核心启动元件, 而上游约



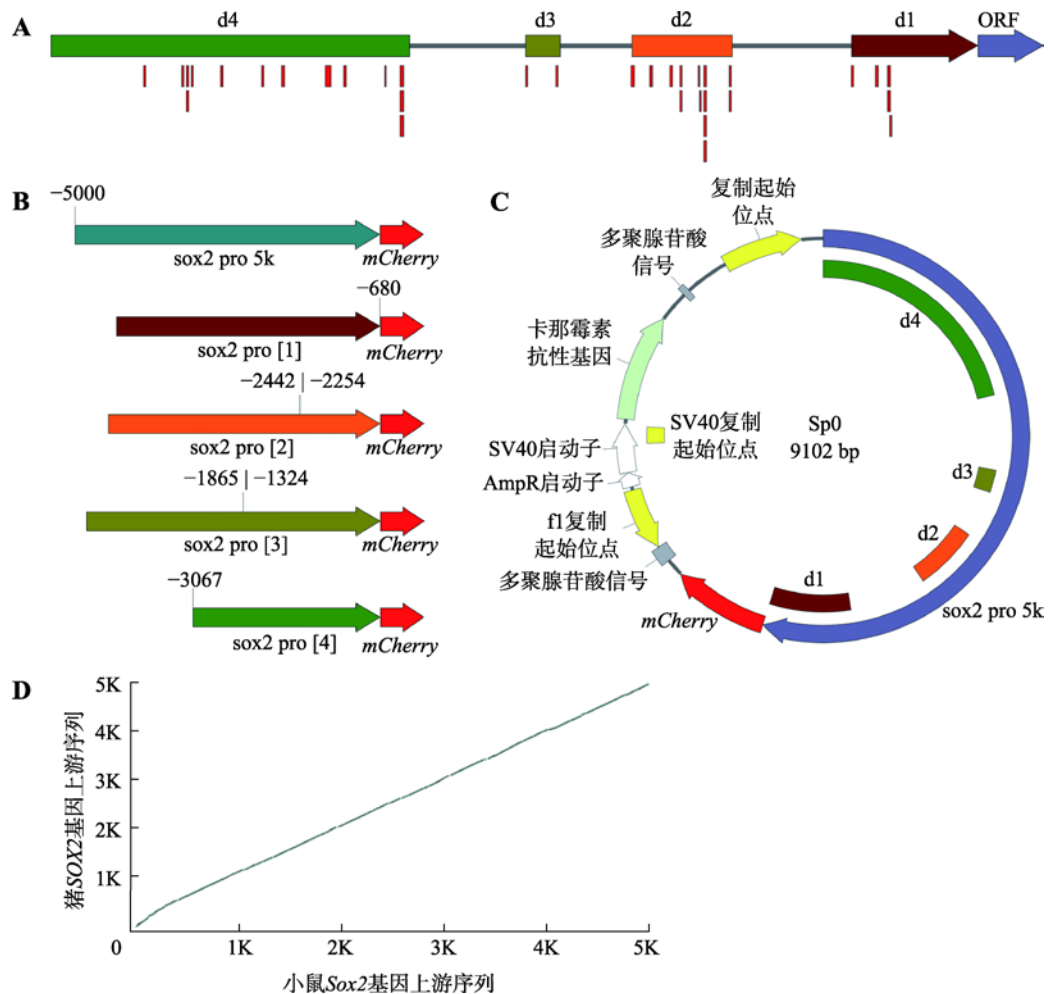


图 4 猪 *SOX2* 翻译起始位点上游 5000 bp 序列结构分析

Fig. 4 Sequence analysis of 5000 bp upstream of the porcine *SOX2* translation initiation site

A 猪 *SOX2* 基因翻译起始位点上游 5000 bp 序列预测的转录因子结合位点(红色条)以及 4 个转录因子结合位点成簇分布的区域 d1~d4; B: 猪 *SOX2* 基因翻译起始位点上游 5000 bp 全长启动子、4 个转录因子结合位点成簇分布的区域 d1~d4 和构建的各区域缺失型启动子 *sox2 pro* [1~4]; C: *SOX2* 基因启动子报告载体图谱; D: 猪 *SOX2* 翻译起始位点上游 5000 bp 序列与小鼠对应序列相似性比较。

60 bp 处存在倒转的“CCAAT box”元件<sup>[23,24]</sup>, 该元件作为 *Sox2* 基因启动子的核心启动元件, 转录因子 NF-Y (nuclear factor-Y) 可识别、结合“CCAAT box”并募集 RNA 聚合酶复合体, 启动 *Sox2* 基因的转录<sup>[25~27]</sup>。而 Sp1 质粒中 *sox2 pro* [1] 启动子片段缺失区域 d1 上存在预测的 NF-Y 结合位点, 且与翻译起始位点的距离同小鼠类似, 因此推测 d1 中包含核心启动元件以及转录起始位点。然而本研究结果显示, Sp1 注射组中可以检测到报告基因的 mRNA 表达。据此推断, 在 *SOX2* 基因启动子的其他区域中存在第二转录起始位点, 由于 *sox2 pro* [1] 中并未破坏第二转录起始位点, 导致该组中检测到报告基因的

mRNA 表达, 但由于移码突变产生的无义蛋白不具有荧光特性, 因此无法检测到报告基因蛋白。预测分析结果显示 d2 区域包括 C/EBP (CCAAT-enhancer-binding protein) 结合位点, 而 C/EBP 转录因子家族具有募集 RNA 聚合酶复合体、启动转录的能力<sup>[28~30]</sup>, 因此推测第二转录起始位点位于 d2 中。在后续研究中, 本课题组将针对这两个潜在的转录起始位点进行分析, 明确猪 *SOX2* 基因转录起始位点特性。

根据上述结果以及相关分析, 可以确定调控 *SOX2* 基因启动子的潜在关键因子包括 NF-AT<sup>[31]</sup>、NF-Y<sup>[32]</sup>、C/EBP<sup>[33]</sup> 和 AP-2 rep (Krueppel-like factor

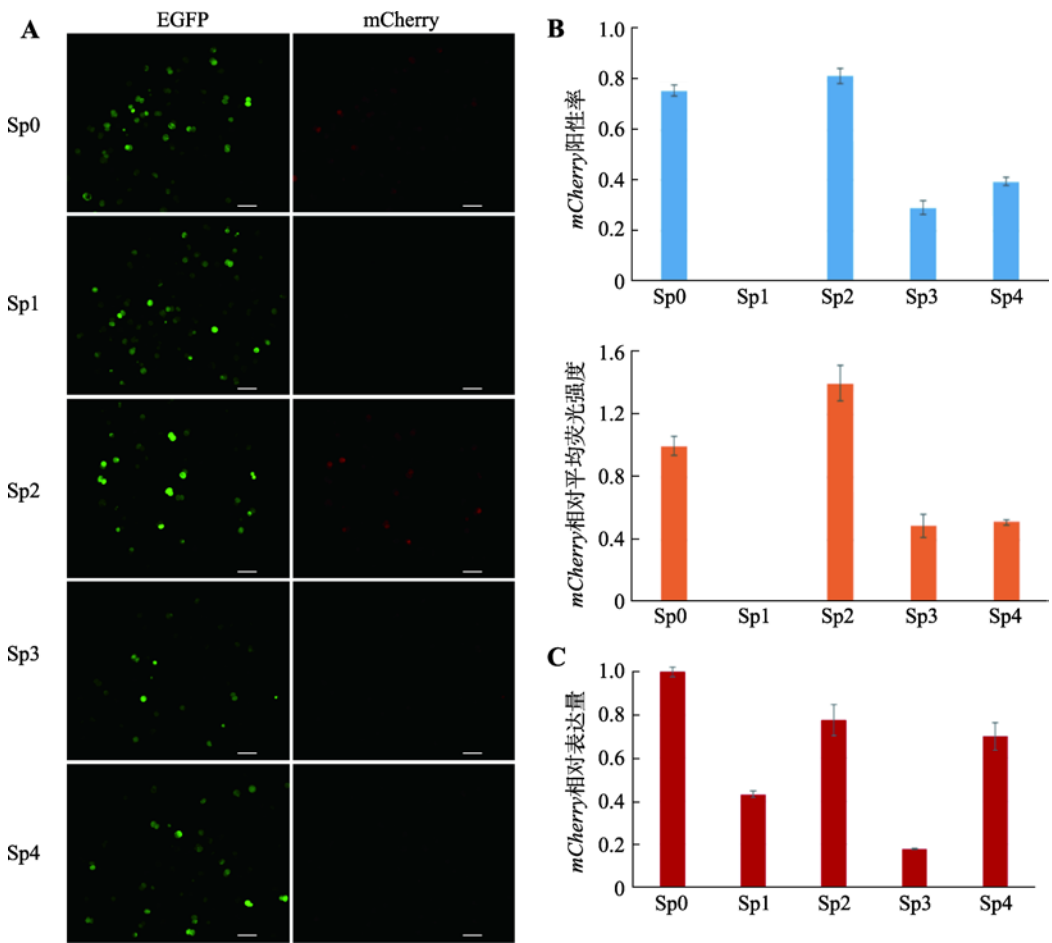


图 5 各启动子片段的活性分析

Fig. 5 Analysis of each promoter fragment activity

A: 孤雌胚胎分别注射 Sp0~4 各组质粒并培养 72 h 后在荧光镜下观察 EGFP 和 mCherry 基因表达情况 (Bar=100  $\mu$ m); B: 孤雌胚胎注射并培养 72 h 后各注射组 mCherry 基因表达统计; C: 孤雌胚胎注射并培养 72 h 后 qRT-PCR 检测 mCherry 相对表达量。

12) 4 种转录因子<sup>[34]</sup>。Cao 等<sup>[35]</sup>对猪早期胚胎发育表达模式进行分析,发现上述 4 种转录因子在早期胚胎发育中都有表达。其中, NF-Y 和 C/EBP 转录因子可能与猪 *SOX2* 基因的转录起始相关; NF-AT 转录因子可以增强 *SOX2* 基因启动子的启动活性;而 AP-2 rep 转录因子或对 NF-AT 转录因子的功能产生影响<sup>[36,37]</sup>。因此推断这 4 类转录因子是潜在调节 *SOX2* 基因表达的关键因子。

双荧光素酶检测系统<sup>[38]</sup>虽然是启动子活性检测的常规技术<sup>[39]</sup>,然而受到早期胚胎样品收集困难以及细胞数低的制约,本研究无法采用双荧光素酶检测系统进行 *SOX2* 基因启动子活性评价。因此选择报告基因表达水平检测的方式评价启动子活性<sup>[40]</sup>。通过 qRT-PCR 检测报告基因 mRNA 的相对表达水平,

避免蛋白降解周期对结果的干扰;利用内参质粒的表达水平检测,排除显微注射体系的系统误差;利用内源 *SOX2* 基因的表达水平检测,排除各组间因胚胎状态差异导致的误差,通过计算获得报告基因相对表达水平可以最高程度地体现相应启动子的转录活性。同时,本研究对胚胎显微注射条件进行探索,最终确定孤雌激活 2~6 h 进行显微注射可以在保证报告基因表达的前提下最小限度地影响胚胎发育。本研究技术体系可以更广泛地应用于猪早期胚胎中启动子活性的研究,筛选分析猪早期胚胎中相关基因调控元件,揭示猪早期胚胎发育特点。

综上所述,本研究利用报告基因表达检测的方法,以 pEGFP-C1 质粒为内参和启动子报告载体骨架,对孤雌后 2~4 h 胚胎注射,并在 72 h 后检测报

告基因表达水平, 评估启动子活性。分析猪 *SOX2* 基因翻译起始位点上游 5000 bp, 发现 4 个转录因子结合位点成簇分布的区域, 类似于小鼠 *Sox2* 基因表达调控中超级增强子(super-enhancer)结构<sup>[41,42]</sup>。通过分析这 4 个区域对猪 *SOX2* 基因启动子活性的影响, 发现存在 NF-TA 转录因子结合位点簇的关键区域, 该区域的缺失导致 *SOX2* 基因启动子活性降低。本研究结果为猪 *SOX2* 基因启动子结构以及上游调控机制的研究提供了数据支持。

## 参考文献(References):

- [1] Sarkar A, Hochedlinger K. The sox family of transcription factors: versatile regulators of stem and progenitor cell fate. *Cell Stem Cell*, 2013, 12(1): 15–30. [DOI]
- [2] Liang S, Furuhashi M, Nakane R, Nakazawa S, Goudarzi H, Hamada J, Iizasa H. Isolation and characterization of human breast cancer cells with sox2 promoter activity. *Biochem Bioph Res Co*, 2013, 437(2): 205–211. [DOI]
- [3] Zhu F, Qian W, Zhang H, Liang Y, Wu M, Zhang Y, Zhang X, Gao Q, Li Y. Sox2 is a marker for stem-like tumor cells in bladder cancer. *Stem Cell Rep*, 2017, 9(2): 429–437. [DOI]
- [4] Stolzenburg S, Rots MG, Beltran AS, Rivenbark AG, Yuan X, Qian H, Strahl BD, Blancafort P. Targeted silencing of the oncogenic transcription factor sox2 in breast cancer. *Nucleic Acids Res*, 2012, 40(14): 6725–6740. [DOI]
- [5] Wang Z, Oron E, Nelson B, Razis S, Ivanova N. Distinct lineage specification roles for nanog, oct4, and sox2 in human embryonic stem cells. *Cell Stem Cell*, 2012, 10(4): 440–454. [DOI]
- [6] Arnold K, Sarkar A, Yram MA, Polo JM, Bronson R, Sengupta S, Seandel M, Geijsen N, Hochedlinger K. Sox2(+) adult stem and progenitor cells are important for tissue regeneration and survival of mice. *Cell Stem Cell*, 2011, 9(4): 317–329. [DOI]
- [7] Hu Y H, Yao JH. Progress on pluripotency factors in zebrafish. *Hereditas (Beijing)*, 2012, 34(9): 1097–1107. 胡雨, 姚纪花. 斑马鱼多能性因子的研究进展. *遗传*, 2012, 34(9): 1097–1107. [DOI]
- [8] Rizzino A, Wuebben EL. Sox2/oct4: a delicately balanced partnership in pluripotent stem cells and embryogenesis. *Biochim Biophys Acta*, 2016, 1859(6): 780–791. [DOI]
- [9] White MD, Angiolini JF, Alvarez YD, Kaur G, Zhao ZW, Mocskos E, Bruno L, Bissiere S, Levi V, Plachta N. Long-lived binding of sox2 to DNA predicts cell fate in the four-cell mouse embryo. *Cell*, 2016, 165(1): 75–87. [DOI]
- [10] Avilion AA, Nicolis SK, Pevny LH, Perez L, Vivian N, Lovell-Badge R. Multipotent cell lineages in early mouse development depend on sox2 function. *Genes Dev*, 2003, 17(1): 126–140. [DOI]
- [11] Wicklow E, Blij S, Frum T, Hirate Y, Lang RA, Sasaki H, Ralston A. Hippo pathway members restrict sox2 to the inner cell mass where it promotes icm fates in the mouse blastocyst. *PLoS Genet*, 2014, 10(10): e1004618. [DOI]
- [12] Liu S, Bou G, Sun R, Guo S, Xue B, Wei R, Cooney AJ, Liu Z. Sox2 is the faithful marker for pluripotency in pig: evidence from embryonic studies. *Dev Dynam*, 2015, 244(4): 619–627. [DOI]
- [13] Wu Z, Chen J, Ren J, Bao L, Liao J, Cui C, Rao L, Li H, Gu Y, Dai H, Zhu H, Teng X, Cheng L, Xiao L. Generation of pig induced pluripotent stem cells with a drug-inducible system. *J Mol Cell Biol*, 2009, 1(1): 46–54. [DOI]
- [14] Gu Q, Hao J, Hai T, Wang J, Jia Y, Kong Q, Wang J, Feng C, Xue B, Xie B, Liu S, Li J, He Y, Sun J, Liu L, Wang L, Liu Z, Zhou Q. Efficient generation of mouse escs-like pig induced pluripotent stem cells. *Protein Cell*, 2014, 5(5): 338–342. [DOI]
- [15] Esteban MA, Xu J, Yang J, Peng M, Qin D, Li W, Jiang Z, Chen J, Deng K, Zhong M, Cai J, Lai L, Pei D. Generation of induced pluripotent stem cell lines from tibetan miniature pig. *J Biol Chem*, 2009, 284(26): 17634–17640. [DOI]
- [16] West FD, Uhl EW, Liu Y, Stowe H, Lu Y, Yu P, Gallegos-Cardenas A, Pratt SL, Stice SL. Brief report: chimeric pigs produced from induced pluripotent stem cells demonstrate germline transmission and no evidence of tumor formation in young pigs. *Stem Cells*, 2011, 29(10): 1640–1643. [DOI]
- [17] Fan N, Chen J, Shang Z, Dou H, Ji G, Zou Q, Wu L, He L, Wang F, Liu K, Liu N, Han J, Zhou Q, Pan D, Yang D, Zhao B, Ouyang Z, Liu Z, Zhao Y, Lin L, Zhong C, Wang Q, Wang S, Xu Y, Luan J, Liang Y, Yang Z, Li J, Lu C, Vajta G, Li Z, Ouyang H, Wang H, Wang Y, Yang Y, Liu Z, Wei H, Luan Z, Esteban MA, Deng H, Yang H, Pei D, Li N, Pei G, Liu L, Du Y, Xiao L, Lai L. Piglets cloned from induced pluripotent stem cells. *Cell Res*, 2013, 23(1): 162–166. [DOI]
- [18] Brevini T, Pennarossa G, Maffei S, Gandolfi F. Pluripotency network in porcine embryos and derived cell lines. *Reprod Domest Anim*, 2012, 47(Suppl. 4): 86–91. [DOI]
- [19] Roberts RM, Yuan Y, Genovese N, Ezashi T. Livestock models for exploiting the promise of pluripotent stem cells. *ILAR J*, 2015, 56(1): 74–82. [DOI]
- [20] Goolam M, Scialdone A, Graham SJL, Macaulay IC, Jedrusik A, Hupalowska A, Voet T, Marioni JC, Zernicka-Goetz M. Heterogeneity in oct4 and sox2 targets biases cell fate in 4-cell mouse embryos. *Cell*, 2016, 165(1): 61–74. [DOI]

- [21] Yang N, Wang Y, Hui L, Li X, Jiang X. Silencing sox2 expression by rna interference inhibits proliferation, invasion and metastasis, and induces apoptosis through map4k4/jnk signaling pathway in human laryngeal cancer tu212 cells. *J Histochem Cytochem*, 2015, 63(9): 721–733. [DOI]
- [22] Rinne A, Banach K, Blatter LA. Regulation of nuclear factor of activated t cells (nfat) in vascular endothelial cells. *J Mol Cell Cardiol*, 2009, 47(3): 400–410. [DOI]
- [23] Wiebe MS, Wilder PJ, Kelly D, Rizzino A. Isolation, characterization, and differential expression of the murine sox-2 promoter. *Gene*, 2000, 246(1–2): 383–393. [DOI]
- [24] Lis M, Walther D. The orientation of transcription factor binding site motifs in gene promoter regions: does it matter? *BMC Genomics*, 2016, 17: 185. [DOI]
- [25] Sikorska M, Sandhu JK, Deb-Rinker P, Jezierski A, Leblanc J, Charlebois C, Ribocco-Lutkiewicz M, Bani-Yaghoub M, Walker PR. Epigenetic modifications of sox2 enhancers, srr1 and srr2, correlate with in vitro neural differentiation. *J Neurosci Res*, 2008, 86(8): 1680–1693. [DOI]
- [26] Mojsin M, Topalovic V, Marjanovic Vicentic J, Stevanovic M. Transcription factor nf-y inhibits cell growth and decreases sox2 expression in human embryonal carcinoma cell line nt2/d1. *Biochemistry (Mosc)*, 2015, 80(2): 202–207. [DOI]
- [27] Wiebe MS, Wilder PJ, Kelly D, Rizzino A. Isolation, characterization, and differential expression of the murine sox-2 promoter. *Gene*, 2000, 246(1–2): 383–393. [DOI]
- [28] Di Stefano B, Sardina JL, van Oevelen C, Collombet S, Kallin EM, Vicent GP, Lu J, Thieffry D, Beato M, Graf T. C/EBP $\alpha$  poises b cells for rapid reprogramming into induced pluripotent stem cells. *Nature*, 2014, 506(7487): 235–239. [DOI]
- [29] Bueno C, Sardina JL, Di Stefano B, Romero-Moya D, Muñoz-López A, Ariza L, Chillón MC, Balanzategui A, Castaño J, Herreros A, Fraga MF, Fernández A, Granada I, Quintana-Bustamante O, Segovia JC, Nishimura K, Ohtaka M, Nakanishi M, Graf T, Menendez P. Reprogramming human b cells into induced pluripotent stem cells and its enhancement by c/EBP $\alpha$ . *Leukemia*, 2016, 30(3): 674–682. [DOI]
- [30] Borrelli S, Fanoni D, Dolfini D, Alotto D, Ravo M, Grober OM, Weisz A, Castagnoli C, Berti E, Vigano MA, Mantovani R. C/EBP $\delta$  gene targets in human keratinocytes. *PLoS One*, 2010, 5(11): e13789. [DOI]
- [31] Kel A, Kel-Margoulis O, Babenko V, Wingender E. Recognition of nfatp/ap-1 composite elements within genes induced upon the activation of immune cells. *J Mol Biol*, 1999, 288(3): 353–376. [DOI]
- [32] Mantovani R. A survey of 178 nf-y binding ccaat boxes. *Nucleic Acids Res*, 1998, 26(5): 1135–1143. [DOI]
- [33] Grange T, Roux J, Rigaud G, Pictet R. Cell-type specific activity of two glucocorticoid responsive units of rat tyrosine aminotransferase gene is associated with multiple binding sites for c/ebp and a novel liver-specific nuclear factor. *Nucleic Acids Res*, 1991, 19(1): 131–139. [DOI]
- [34] Roth C, Schuierer M, Günther K, Buettner R. Genomic structure and DNA binding properties of the human zinc finger transcriptional repressor ap-2rep (klf12). *Genomics*, 2000, 63(3): 384–390. [DOI]
- [35] Cao S, Han J, Wu J, Li Q, Liu S, Zhang W, Pei Y, Ruan X, Liu Z, Wang X, Lim B, Li N. Specific gene-regulation networks during the pre-implantation development of the pig embryo as revealed by deep sequencing. *BMC Genomics*, 2014, 15: 4. [DOI]
- [36] Hilger-Eversheim K, Moser M, Schorle H, Buettner R. Regulatory roles of ap-2 transcription factors in vertebrate development, apoptosis and cell-cycle control. *Gene*, 2000, 260(1–2): 1–12. [DOI]
- [37] Macián F, López-Rodríguez C, Rao A. Partners in transcription: Nfat and ap-1. *Oncogene*, 2001, 20(19): 2476–2489. [DOI]
- [38] Ning Ding, Yuan Gao, Wang N, Li H. Functional analysis of the chicken ppar $\gamma$  gene flanking region and c/EBP $\alpha$  mediated gene regulation. *Comp Biochem Phys B*, 2011, 158(4): 297–303. [DOI]
- [39] Cheng M, Zhang WJ, Xing TY, Yan XH, Li YM, Li H, Wang N. Functional analysis of the upstream regulatory region of chicken mir-17-92 cluster. *Hereditas (Beijing)*, 2016, 38(8): 724–735.
- 程敏, 张文建, 邢天宇, 闫晓红, 李玉茂, 李辉, 王宁. 鸡 miR-17-92 基因簇上游调控区功能分析. *遗传*, 2016, 38(8): 724–735. [DOI]
- [40] Yu L, Domann FE. Rapid and direct quantitative rt-pcr method to measure promoter activity. *Biotechnol Prog*, 2006, 22(5): 1461–1463. [DOI]
- [41] Zhang S, Magnusson G. Kilham polyomavirus: activation of gene expression and DNA replication in mouse fibroblast cells by an enhancer substitution. *J Virol*, 2001, 75(21): 10015–10023. [DOI]
- [42] Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, Rahl PB, Lee TI, Young RA. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell*, 2013, 153(2): 307–319. [DOI]