

论著·基础研究

双同源盒诱导小鼠胚胎干细胞向胚外内胚层分化的机制研究

洪磊^{1*}, 郭传亮^{1*}, 蔡勤¹, 李婉睿¹, 曾溢滔¹, 薛燕^{1,2#}, 曾凡一^{1,2#}

1. 上海市儿童医院, 上海交通大学医学附属儿童医院上海医学遗传研究所, 国家卫健委医学胚胎分子生物学重点实验室, 上海市胚胎与生殖工程重点实验室, 上海 200040; 2. 上海交通大学基础医学院组织胚胎学与遗传发育学系, 上海 200025

[摘要] 目的·探索双同源盒 (double homeobox, DUX) 蛋白对小鼠胚胎干细胞 (mouse embryonic stem cell, mESC) 向胚外内胚层 (extraembryonic endoderm, XEN) 分化潜能的影响及可能的作用机制。方法·使用慢病毒体系在 mESC 中构建过表达 DUX 细胞系, 利用流式细胞术检测 DUX 过表达前后 2 细胞样细胞 (2-cell-like cell, 2CLC) 的比例, 并使用实时定量聚合酶链反应 (real-time quantitative reverse transcription polymerase chain reaction, RT-qPCR) 检测 2 细胞期特异性基因, 如内源性 *Dux*、锌指和 SCAN 结构域的蛋白质 4c (zinc finger and SCAN domain containing 4c, *Zscan4c*)、锌指蛋白 352 (zinc finger protein 352, *Zfp352*) 和鼠内源性反转录病毒-聚合酶 (murine endogenous retrovirus-L polymerase, *MERVL-pol*) 的表达。RT-qPCR 检测过表达 DUX 的 mESC 多能性因子 [nanog homeobox (*Nanog*)、kruppel-like transcription factor 4 (*Klf4*)、性别决定区 Y 框蛋白 2 (sex determining region Y-box 2, *Sox2*)、八聚体结合转录因子 4 (octamer-binding transcription factor 4, *Oct4*)] 和自然分化状态下各胚层标志基因 [内胚层 (endodermal): GATA 结合蛋白 4 (GATA binding protein 4, *Gata4*)、*Gata6*、*Sox17*; 外胚层 (ectodermal): 微 III 型 β 微管蛋白 3 (tubulin beta 3 class III, *Tubb3*)、巢蛋白 (*Nestin*); 中胚层 (mesodermal): 心脏和神经嵴衍生物表达转录本 1 (heart and neural crest derivatives expressed 1, *Hand1*)、肌源性分化蛋白 1 (myogenic differentiation 1, *Myod1*)、激酶插入结构域受体 (kinase insert domain protein receptor, *Flkl1*)] 的表达。挖掘公共转录组测序 (RNA sequencing, RNA-seq) 数据, 通过分析胚外内胚层标志基因的表达水平, 明确 DUX 对 mESC 向胚外内胚层分化的影响; 通过对差异基因的功能及通路进行基因本体论 (Gene Ontology, GO) 富集分析、京都基因和基因组数据库 (Kyoto Encyclopedia of Genes and Genomes, KEGG) 富集分析和基因集富集分析 (gene set enrichment analysis, GSEA), 找出 DUX 作用的信号通路; 深入分析已有的染色质免疫共沉淀技术结合二代测序 (chromatin immunoprecipitation sequencing, ChIP-seq) 数据, 探究 DUX 的潜在靶基因。结果·2CLC 比例升高和 2 细胞期标志基因表达上调, 证明已成功构建过表达 DUX 细胞系。分子生物学实验显示过表达 DUX 后可有效维持 mESC 的多能性, 与公共 RNA-seq 数据分析结果一致; 差异基因分析发现, 内胚层基因出现特异性上调; 诱导 mESC 自然分化后, RT-qPCR 检测实验表明 XEN 标志基因 (*Gata4*、*Gata6*、*Sox17*) 的 mRNA 表达出现显著上调 ($P < 0.001$), 而中胚层、外胚层基因没有特异性变化。GSEA 结果提示 DUX 可能激活了视黄醇代谢信号通路, ChIP-seq 数据解析进一步揭示在 DUX 结合的 peaks 中存在已知的视黄酸受体 motif, 可激活下游与 XEN 发育相关的靶基因。结论·DUX 与视黄酸信号通路密切相关, 预示其激活了视黄酸信号通路, 促进 mESC 倾向 XEN 分化。

[关键词] 小鼠胚胎干细胞; 胚外内胚层细胞; 双同源盒; 视黄酸信号通路**[DOI]** 10.3969/j.issn.1674-8115.2024.11.003 **[中图分类号]** R394.1 **[文献标志码]** A

Mechanism of DUX-induced differentiation of mESC into extraembryonic endoderm

HONG Lei^{1*}, GUO Chuanliang^{1*}, CAI Qin¹, LI Wanrui¹, ZENG Yitao¹, XUE Yan^{1,2#}, ZENG Fanyì^{1,2#}

1. Shanghai Institute of Medical Genetics, Shanghai Children's Hospital, Shanghai Jiao Tong University School of Medicine, NHC Key Laboratory of Medical Embryogenesis and Developmental Molecular Biology & Shanghai Key Laboratory of Embryo and Reproduction

[基金项目] 国家重点研发计划 (2019YFA0801402); 国家自然科学基金 (82271890); 上海市临床重点专科项目 (shslczdk05705); 上海市重点学科项目 (2017ZZ02019); 上海市地方高水平大学创新研究团队 (SHSMU-ZDCX20212200); 澳门科学技术发展基金 (FDCT) (0092/2022/A2, 003/2022/ALC)。
[作者简介] 洪磊 (1996—), 男, 硕士生; 电子信箱: honglei_shnu@163.com。郭传亮 (1986—), 男, 博士, 助理研究员; 电子信箱: chuanliangguo@163.com。*为共同第一作者。

[通信作者] 曾凡一, 电子信箱: fzeng@vip.163.com。薛燕, 电子信箱: xueyan@shsmu.edu.cn。#为共同通信作者。

[Funding Information] National Key Research and Development Program of China (2019YFA0801402); National Natural Science Foundation of China (82271890); Shanghai Key Clinical Specialty Project (shslczdk05705); Shanghai Top Priority Key Discipline Project (2017ZZ02019); Innovative Research Team of High-Level Local Universities in Shanghai (SHSMU-ZDCX20212200); Macau Science and Technology Development Fund (FDCT) (0092/2022/A2, 003/2022/ALC)。

[Corresponding Author] ZENG Fanyì, E-mail: fzeng@vip.163.com. XUE Yan, E-mail: xueyan@shsmu.edu.cn. #Co-corresponding authors.

Engineering, Shanghai 200040, China; 2. Department of Histoembryology, Genetics & Development, Shanghai Jiao Tong University College of Basic Medical Sciences, Shanghai 200025, China

[Abstract] **Objective** To explore the effect of double homeobox (DUX) protein on the differentiation potential of mouse embryonic stem cells (mESCs) into extraembryonic endoderm (XEN) and the possible mechanism of its action. **Methods** Overexpression of DUX cell lines in mESCs was achieved by using a lentiviral system. The proportion of 2-cell-like cells (2CLCs) before and after DUX overexpression was detected by flow cytometry, and the expression of 2-cell stage-specific genes, *Dux*, zinc finger and SCAN domain containing 4c (*Zscan4c*), zinc finger protein 352 (*Zfp352*) and murine endogenous retrovirus-L polymerase (*MERVL-pol*), were detected by real-time quantitative reverse transcription polymerase chain reaction (RT-qPCR). RT-qPCR assay was used to detect the expression of pluripotency factors, nanog homeobox (*Nanog*), kruppel-like transcription factor 4 (*Klf4*), sex determining region Y-box 2 (*Sox2*), and octamer-binding transcription factor 4 (*Oct4*), in pluripotent state, as well as the expression of signature genes for different germ layers in the differentiated state [endodermal: GATA binding protein 4 (*Gata4*), GATA binding protein 6 (*Gata6*), and sex determining region Y-box 17 (*Sox17*); ectodermal: *Nestin* and tubulin beta 3 class III (*Tubb3*); mesodermal: heart and neural crest derivatives expressed 1 (*Hand1*), myogenic differentiation 1 (*Myod1*), and kinase insert domain protein receptor (*Flkl1*)]. Public RNA sequencing (RNA-seq) data were mined to further clarify the effect of DUX on the differentiation of mESCs into extraembryonic endoderm. Functional and pathway enrichment analyses of differentially expressed genes were performed using Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), and gene set enrichment analysis (GSEA) to identify the signaling pathways regulated by DUX. Additionally, an in-depth analysis of existing chromatin immunoprecipitation sequencing (ChIP-seq) data was conducted to explore the potential target genes of DUX. **Results** Molecular biology experiments showed that overexpression of DUX could effectively maintain the pluripotency of mESCs, which was consistent with the analysis of public RNA-seq data. Differential gene analysis revealed that endodermal genes were specifically upregulated. After differentiation assay of mESCs, RT-qPCR assay experiments showed that mRNA expression of the XEN marker genes (*Gata4*, *Gata6*, *Sox17*) was significantly upregulated ($P < 0.001$). In contrast, there was no specific change in mesodermal and ectodermal genes. GSEA enrichment analysis indicated that DUX might activate the retinoid metabolism signaling pathway, and the analysis of the ChIP-seq data further revealed the presence of a large number of known retinoic acid receptor motif in DUX-bound peaks, which could activate downstream target genes related to the development of the XEN. **Conclusion** DUX has a strong correlation with the retinoic acid signaling pathway and it is predicted to activate the retinoic acid signaling pathway, which could promote the tendency of mESCs toward XEN differentiation.

[Key words] mouse embryonic stem cell (mESC); extraembryonic endoderm (XEN); double homeobox (DUX); retinoic acid signaling pathway

细胞的全能性是指单个细胞产生完整胚胎及其胚胎外成分的能力。而在小鼠中, 这种能力仅限于受精卵和2细胞的发育阶段^[1], 伴随着发育的进行, 分裂的细胞发育潜能逐渐受限, 并在囊胚阶段分化成不同的细胞系而丧失全能性。在小鼠胚胎早期发育过程中, 全能卵裂球分化为内细胞团 (inner cell mass, ICM) 和滋养外胚层 (trophectoderm, TE), 内细胞团形成囊胚的原始内胚层 (primitive endoderm, PrE) 和上胚层 (epiblast, EPI)^[2]。

来源于囊胚内细胞团体外培养的小鼠胚胎干细胞 (mouse embryonic stem cell, mESC), 具有分化成几乎所有细胞类型的能力, 常称为多能干细胞^[3-4]; 与全能干细胞相比, 其不能分化形成胚外组织。有研究从mESC转化或直接从八细胞期胚泡中体外培养获得发育潜能扩展或扩大的多能干细胞 (extended or expanded pluripotent stem cell, EPSC)^[5-7]。这2种细胞类型都显示出多能性得到了扩展的分子和功能特征, 如全能性相关标记基因的表达, 以及通过嵌合对EPI、PE和TE细胞谱系的贡献。然而EPSC倾向于

PrE谱系细胞分化, 这导致使用EPSC构建的“扩展多潜能干细胞类囊胚 (EPS-blastoids)”存在严重的植入后缺陷^[8-9]。

体外培养的mESC保留了极少量的细胞亚群 (0.1%~1%), 这些亚群表达鼠内源性反转录病毒 (murine endogenous retrovirus-L, MERVL) 和金属硫蛋白-2 (metallothionein-2, MT-2) 转座子以及其他全能性基因 [如锌指和SCAN结构域的蛋白质4 (zinc finger and SCAN domain containing 4, *Zscan4*) 家族], 表现出2细胞期胚胎的多个特征, 如2细胞胚胎特异性转录物的表达^[10]、组蛋白流动性的增加^[11]以及内细胞团和滋养外胚层发育潜能的促进^[10], 因而这一少量的细胞亚群也被称为2细胞样细胞 (2-cell-like cell, 2CLC)。由于2细胞胚胎难以大量获取, 而体外培养的mESC中可相对容易地分离富集得到2CLC, 使其成为了解全能性和合子基因组激活的良好模型^[12]。但是许多研究表明一些体外建立的2CLC表达谱更接近于囊胚期细胞^[13-14], 这说明可以使用这些细胞系去构建类囊胚, 模拟TE、EPI与PrE

的发育过程。值得注意的是,缺失2CLC的mESC表现出向中胚层和外胚层分化的倾向^[10]。然而,仍不清楚的是这些2CLC是否存在与EPSC类似的分化倾向性,即倾向于胚外内胚层分化。

双同源盒(double homeobox, DUX)家族蛋白是mESC进入2C样状态的关键驱动因素。DUX(小鼠)或DUX4(人类)可能是胚胎基因组激活(zygotic genome activation, ZGA)的关键因子,诱导人类和小鼠中最早的ZGA基因波^[15-17]。在mESC中,诱导DUX表达可在24 h内将MERVL阳性细胞的效率提高到10%~74%^[16]。此外,DUX诱导的2C-like状态是短暂的、可逆的,类似于自发形成的2CLC^[18]。围绕DUX上下游调控机制已发现很多早期胚胎发育过程中的关键转录分子或表观遗传学修饰因子^[19-23],而关于DUX如何通过诱导2CLC比例增加导致mESC在分化过程中命运转变的分子机制,迄今未见报道。本文重点聚焦于此,探索DUX对mESC分化潜能的影响,并分析潜在的作用机制。

1 材料与方法

1.1 细胞株和质粒

本研究使用mESC细胞株CGR8.8由斯坦福大学的Dr. Chen教授赠予;HEK293T细胞购自中国科学院上海生命科学研究院细胞库。

用于2CLC示踪的载体由北京协和医学院黄粤教授赠予;用于DUX过表达的载体为pCW57-MCS1-P2A-MCS2-mDux;用于慢病毒包装的辅助载体为psPAX2、pMD2.G。

1.2 主要试剂与仪器

MEM培养基、胎牛血清、 β -巯基乙醇、胰蛋白酶(Gibco, 美国),青-链霉素双抗、L-谷氨酰胺、非必需氨基酸(上海源培生物科技股份有限公司),LIF(Millipore, 美国),PD0325901、CHIR99021(上海陶术生物科技股份有限公司),明胶(Sigma, 德国),Lipofectamine 3000、Trizol(Invitrogen, 美国),潮霉素、嘌呤霉素、阿霉素、氨苄霉素、酵母粉、胰蛋白胨(上海碧云天生物有限公司),聚合酶链反应(polymerase chain reaction, PCR)高保真酶、SYBR Green Fast qPCR Mix、cDNA反转录试剂盒

(南京诺唯赞生物科技股份有限公司),感受态、质粒提取试剂盒(天根生化科技有限公司),焦碳酸二乙酯(diethyl pyrocarbonate, DEPC)、无水乙醇、异丙醇、三氯甲烷(上海大合化学制品有限公司)。PCR仪(Roche, 瑞士),实时荧光定量PCR仪、Nanodrop 2000、细胞培养箱(ThermoFisher, 美国),显微镜(Olympus, 日本)。

1.3 实验方法

1.3.1 细胞培养 HEK293T培养基体系为高糖DMEM培养基添加10%的胎牛血清及1%青-链霉素双抗。mESC培养基体系为Knockout DMEM培养基添加15%的胚胎干细胞专用血清、1%青-链霉素双抗、1% L-谷氨酰胺、1%非必需氨基酸、0.55 mmol/L β -巯基乙醇和LIF(10^4 U/mL)。撤去LIF,将mESC悬浮培养在细菌培养皿中,形成拟胚体。继续培养48 h,每日换液并进行后续实验。细胞均置于5% CO₂的37 °C培养箱中进行培养。

1.3.2 建立稳定MERVL::tdTomato报告ESC细胞系 使用Lipofectamine 3000将含有MERVL::tdTomato的载体转染至mESC中,转染48 h后将细胞传代至10 cm直径培养皿中,并更换含有150 mg/mL潮霉素的培养液筛选7 d,待单个细胞长大时挑选单克隆。

1.3.3 建立DUX过表达的mESC细胞系 首先借助于HEK293T细胞进行慢病毒载体的包装,将HEK293T细胞传代至6孔板中,传代比例为第2日细胞密度达到70%~80%。慢病毒包装质粒psPAX2、包膜质粒pMD2.G和目的基因质粒按3:1:4比例混合加入至opti-MEM培养基中,并加入P3000试剂混匀;Lipofectamine 3000与另一份opti-MEM培养基混匀,将两者溶液混匀后室温静置15 min。将混合溶液缓慢滴加到培养液中,轻轻摇匀,细胞正常培养24 h后换新鲜培养液。收集转染48 h的培养上清液,3000 \times g、4 °C离心5 min,去除细胞沉淀,使用0.45 μ m的滤器过滤液体。含有慢病毒的滤液可直接用于感染或者冻存于-80 °C。mESC在感染前1 d铺板,感染当日密度为20%~30%。感染时去除旧培养液,换成一半培养液和一半慢病毒滤液混合并加入终浓度8 μ g/mL的凝聚胺,于37 °C、5% CO₂的培养箱中继续培养,24 h后换成正常新鲜培养液。感染48 h后,将细胞传代至10 cm直径培养皿中,并更换含有1 μ g/mL嘌呤

霉素的培养液筛选 7 d, 待单个细胞长大时挑选单克隆。将挑出的单克隆一分为二, 一份加入含有 2 μg/mL 四环素 (doxycycline) (记为 plusDox), 另一份加正常 mESC 培养液 (记为 noDox), 培养 24 h 后荧光显微镜观察, 保存有红色荧光的克隆进行后续的实验。

1.3.4 细胞总 RNA 提取、反转录和实时定量聚合酶链反应 (real-time quantitative reverse transcription polymerase chain reaction, RT-qPCR) 将细胞从培养箱中取出, 弃上清液, PBS 洗涤 1 遍, 0.25% 胰酶消化 3 min, 使用含血清培养液终止消化, 800×g 离心 3 min, 弃上清液。PBS 重悬清洗 1 遍, 800×g 离心 3 min, 弃上清液, 然后向 EP 管中加入 1 mL 的 Trizol 充分吹打混匀, 放在室温静置 10 min。向每份 EP 管中加入 200 μL 三氯甲烷, 涡旋剧烈震荡 15 s, 室温静置 3 min。10 000×g, 4 °C 离心 15 min。将上层澄清液转移至新的 EP 管中。向澄清液中加入等体积的异丙醇 (一般为 500 μL), 上下颠倒混匀, 室温静置 10 min。10 000×g, 4 °C 离心 10 min。吸上清液, 加入 1 mL 的 75% 乙醇混匀重悬。5 000×g, 4 °C 离心

5 min。重复上一步。弃上清液, 将 EP 管打开放在超净台里, 室温鼓风吹干, 一般 5~10 min, 注意风干时间不要过长。加入适量的 DEPC 水溶解 RNA 沉淀, 可以冰上或 4 °C 溶解 1 h 左右, 短暂低速离心, Nanodrop 2000 测浓度后进行接下来的反转录实验, 剩下的 RNA 冻在 -80 °C。

取 1 μg RNA 加 4 μL 4×gDNA wiper Mix 并用 RNase-free ddH₂O 调整体积至 16 μL, 42 °C 孵育 2 min。再加入 4 μL 5×HiScript II qRT SuperMix II, 50 °C, 15 min; 然后 85 °C, 5 s。

用 ddH₂O 将上述反转录后的 cDNA 溶液体积调整至 100 μL。反应体系为 2 μL cDNA、0.5 μL qPCR-F (10 μmol/L)、0.5 μL qPCR-R (10 μmol/L)、5 μL 2×ChamQ Universal SYBR qPCR Master Mix 和 7 μL ddH₂O。反应条件如下: 95 °C 初始变性 3 min, 然后 95 °C 5 s、60 °C 30 s、72 °C 15 s, 40 个循环。以 3-磷酸-甘油醛脱氢酶 (glyceraldehyde-3-phosphate dehydrogenase, *Gapdh*) 作为参考基因 (表 1); 采用 2^{-ΔΔC_t} 方法计算相对表达量。

表 1 用于 RT-qPCR 的引物序列

Tab 1 Primer sequences used for RT-qPCR

Gene	Forward primer (5'→3')	Reverse primer (5'→3')
<i>Gapdh</i>	AGGTCGGTGTGAACGGATTG	TGTAGACCATGTAGTTGAGGTCA
<i>Dux</i>	ACTTCTAGCCCCAGCGACTC	CCATGCTGCCAGGATTTCTA
<i>Zscan4c</i>	GTCCTGACAGAGGCCTGCC	GAGATGTCTGAAGAGGCAAT
<i>Zfp352</i>	AAGTCCCACATCTGAAGAAACAC	GGGTATGAGGATTCACCCACA
<i>MERVL-pol</i>	ATCTCCTGGCACCTGGTATG	AGAAGAAGGCATTTGCCAGA
<i>Nanog</i>	TCTTCCTGGTCCCCACAGTTT	GCAAGAATAGTTCTCGGGATGAA
<i>Sox2</i>	GCGGAGTGAAACTTTTGTC	CGGGAAGCGTGACTTATCCTT
<i>Klf4</i>	GTGCCCCGACTAACCGTTG	GTCGTTGAACTCCTCGGTCT
<i>Oct4</i>	GGCTTCAGACTTCGCCTCC	AACCTGAGGTCCACAGTATGC
<i>Gata4</i>	CCCTACCCAGCCTACATGG	ACATATCGAGATTGGGGTGTCT
<i>Gata6</i>	GAGCTGGTGCTACCAAGAGG	TGCAAAAGCCCATCTCTTCT
<i>Sox17</i>	GAATTTCCATCTCCACCTCCG	CGATTGGCACCTTTCACCC
<i>Nestin</i>	CTGCAGGCCACTGAAAAGTT	GACCCTGCTTCTCCTGCTC
<i>Tubb3</i>	CGGCAACTATGTAGGGGACT	CCAGCACCACTCTGACCAA
<i>Hand1</i>	AAGGATGCACAAGCAGGTGAC	TTTAATCCTCTTCTCGCCGGG
<i>Myod1</i>	CCACTCCGGACATAGACTTG	AAAAGCGCAGGTCTGGTGAG
<i>Flk1</i>	CCTGGTCAAACAGCTCATCA	AAGCGTCTGCCTCAATCACT

Note: *Zscan4c*—zinc finger and SCAN domain containing 4c; *Zfp352*—zinc finger protein 352; *MERVL-pol*—murine endogenous retrovirus-L polymerase; *Nanog*—nanog homeobox; *Sox2*—sex determining region Y-box 2; *Klf4*—kruppel-like transcription factor 4; *Oct4*—octamer-binding transcription factor 4; *Gata4*—GATA binding protein 4; *Gata6*—GATA binding protein 6; *Sox17*—sex determining region Y-box 17; *Tubb3*—tubulin beta 3 class III; *Hand1*—heart and neural crest derivatives expressed 1; *Myod1*—myogenic differentiation 1; *Flk1*—kinase insert domain protein receptor.

1.3.5 RNA-seq 数据分析 从 EMBL-EBI 数据库分别下载视黄酸处理 (PRJNA677836^[24]) 和 DUX 过表达 (PRJNA338980^[16]) mESC 的 RNA-seq 数据, 样本信息如表 2。

表 2 用于 RNA-seq 数据分析的样本信息

Tab 2 Sample information for RNA-seq data analysis

Group	Sample	Description
DUX	Unsorted-NoDox-rep1/2	Untreated with tetracycline and unsorted 2CLCs
	Unsorted-plusDox-rep1/2	Treated with tetracycline and unsorted 2CLCs
RA	2iL-ESCs-rep1/2/3	Control
	RA-ESCs-rep1/2/3	Treated with RA

为了分析基因表达, 本文使用 trim-galore (v0.5.0) 对原始数据进行过滤, 得到高质量的 clean data。利用 hisat2 (v2.1.0) 软件将质量控制后的序列比对到参考基因组 (mm10), 并采用 stringtie2 (v2.1.7) 软件分析对已知的基因和转录本进行表达定量。通过 DESeq2 分析差异表达基因, 筛选标准为 $P<0.05$ 且差异倍数 (fold change, FC) 对数的绝对值 ($|\log_2FC|$) >1 。使用 R (v4.3.1) 包 clusterProfile

(v4.8.2) 对筛选出的差异基因进行基因本体论 (Gene Ontology, GO) 富集分析、京都基因和基因组数据库 (Kyoto Encyclopedia of Genes and Genomes, KEGG) 富集分析和基因集富集分析 (gene set enrichment analysis, GSEA)。

1.3.6 ChIP-seq 数据分析 从 EMBL-EBI 数据库下载 HA-tagged-Dux ChIP-seq 数据 (PRJNA377313^[16]), 样本信息如表 3。

表 3 用于 ChIP-seq 数据分析的样本信息

Tab 3 Sample information for ChIP-seq data analysis

Sample	Description
Unsorted-plusDox-Input	Input. Treated with tetracycline and unsorted 2CLCs
Unsorted-plusDox-HAChIP-12hr-rep1/2	Addition of tetracycline-treated 12hr, unsorted 2CLCs
Unsorted-plusDox-HAChIP-18hr-rep1/2	Addition of tetracycline-treated 18hr, unsorted 2CLCs

使用 bowtie2 (v2.2.5) 将 ChIP-seq reads 比对到参考基因组 (mm10) 上, 丢弃线粒体 DNA 或未分配对应序列的读取。对于单一的比对, 使用 SAMtools (v1.9) 过滤重复读取和低映射质量读取。使用 MACS2 (v2.1.2) 调取 ChIP-seq 富集峰, 用 IDR (v2.0.3) 合并共同峰, 用于后续分析。使用 R 包 ChIPseeker (v1.39.0) 对富集的峰进行注释。使用 MEME 套件 v5.0.5 中的 analysis of motif enrichment (AME) 软件进行转录因子结合位点富集分析, 从 JASPAR 下载用于扫描的结合基序矩阵。使用 deepTools (v3.5.5) 和 RPKM (reads per kilo base per million mapped reads) 归一化方法生成 bigwig 文件。在 IGV (Integrative Genomics Viewer) 中创建并展示图表。

1.4 统计学方法

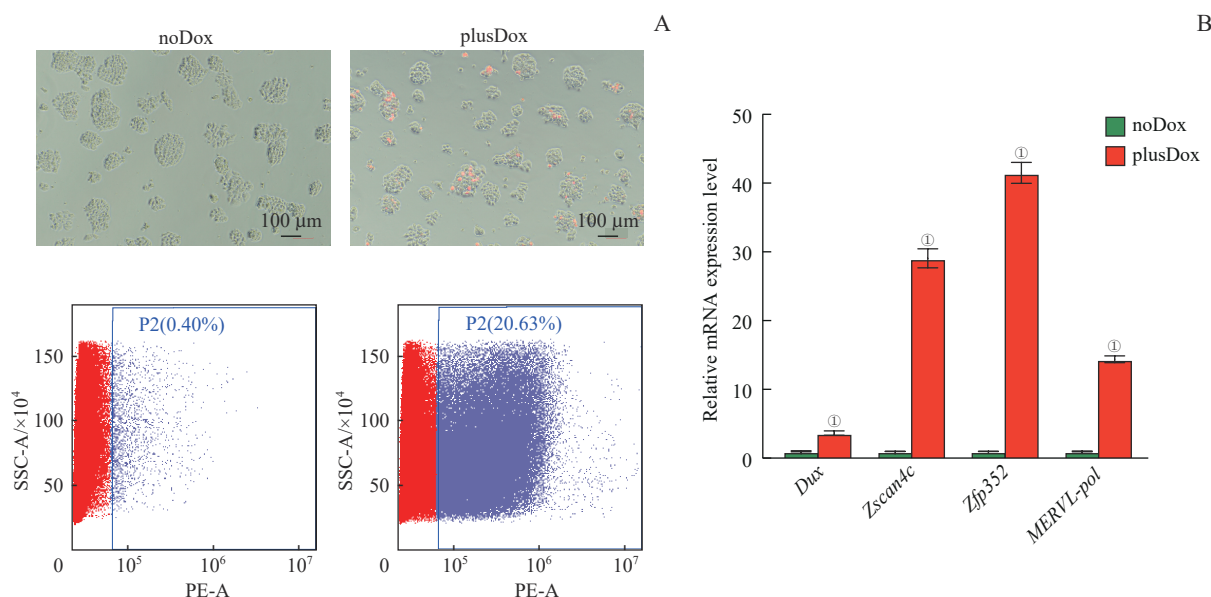
本研究中的实验数据均来自 3 次独立重复操作,

使用 GraphPad Prism 9 软件进行统计分析和图表绘制。定量数据以 $\bar{x}\pm s$ 表示, 并采用 Student's *t*-test 检验进行两两比较。 $P<0.05$ 表示差异具有统计学意义。

2 结果

2.1 建立诱导 DUX 过表达细胞系

首先在 mESC 中转染 MERV L:: tdTomato 质粒, 构建 2CLC 指示性细胞系。在此细胞系基础上, 使用慢病毒感染过表达 DUX。在加入四环素培养 24 h 后, 流式细胞计数检测结果显示红色荧光细胞比例显著提高, 2CLC 占比从 0.40% 升至 20.63% (图 1A)。RT-qPCR 检测 2 细胞期特异性基因 *MERV L-pol*、*Zscan4c*、*Zfp352* 和内源 *Dux* 的 mRNA 表达, 结果均显著升高 (图 1B)。这些数据表明已成功构建 DUX 诱导过表达细胞系。



Note: A. Microscopy pictures and flow cytometry showing the percentage of 2CLCs. B. RT-qPCR validation of selected 2C marker genes and MERVL-polymerase gene (*MERVL-pol*) in noDox (control) and plusDox (DUX overexpression) ES cells. ^① $P < 0.001$, compared with the control group.

图1 2CLC的鉴定

Fig 1 Identification of 2CLCs

2.2 DUX促进mESC向胚外内胚层分化

值得注意的是, 2CLC对嵌合体的胚胎(着床后的外胚层衍生)和胚胎外(原始内胚层和滋养外胚层衍生)组织都有贡献^[10]。相对于对照组, 过表达DUX形成2CLC后, RT-qPCR检测多能性因子的表达情况, 结果显示多能性因子表达没有明显变化(图2A)。对DUX过表达RNA-seq数据分析发现, 除了2C特异性基因明显高表达外(图2B), 内胚层标志基因也高表达(图2C)。在ES细胞培养阶段短暂诱导DUX表达12 h, 随后进行拟胚体培养, 第4日RT-qPCR检测三胚层标志因子的表达; 相对于对照组, DUX的激活导致胚外内胚层基因表达明显上调, 外胚层、中胚层基因没有显著变化(图2D)。这些结果表明DUX除了促进2CLC形成外, 还具有促进mESC向胚外内胚层分化的作用。

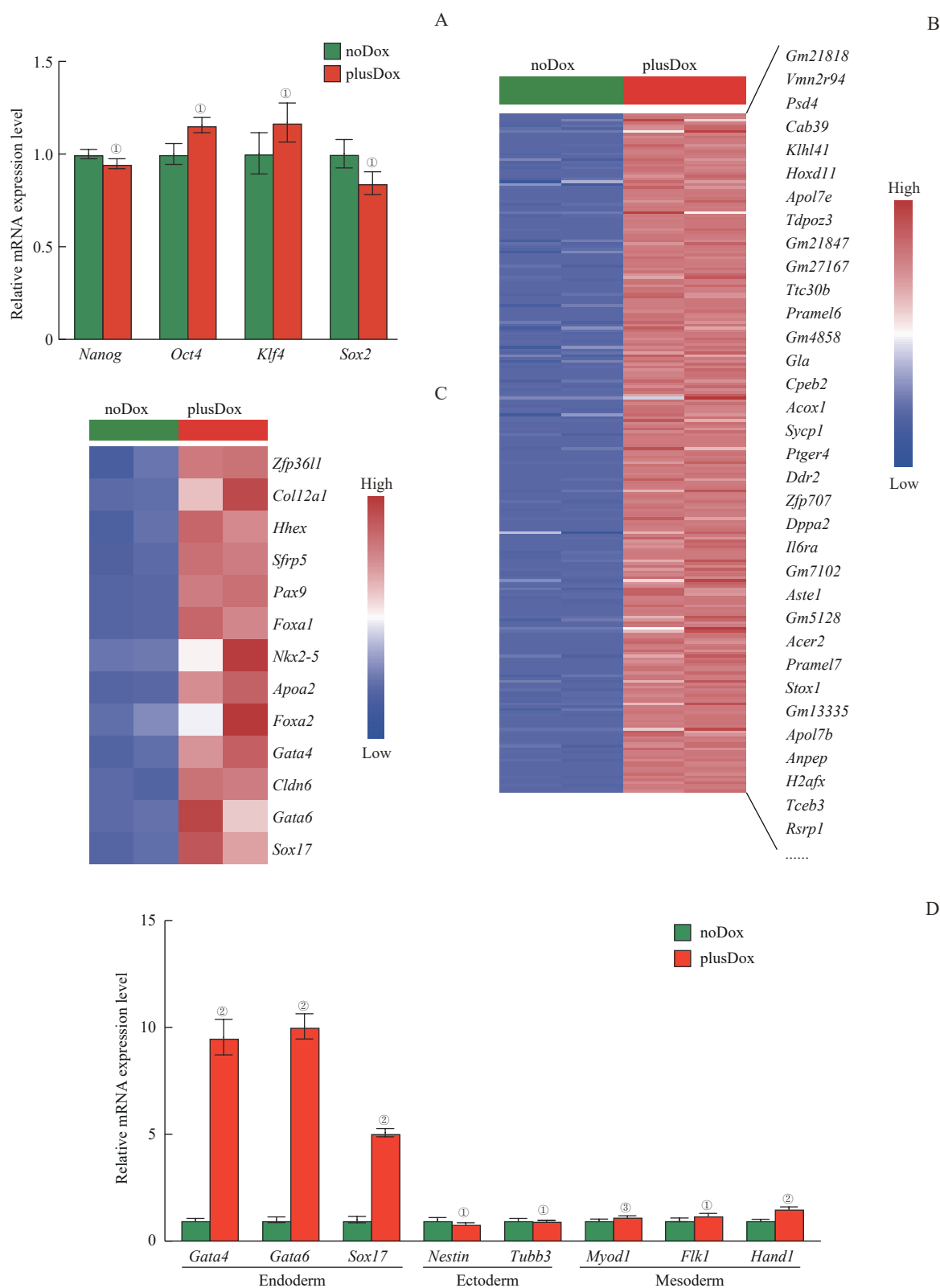
2.3 DUX激活视黄酸信号通路

为了探究DUX促进mESC向胚外胚层分化的作用机制, 通过重新分析文献RNA-seq数据, GO分析结果显示DUX促进的上调基因主要富集在模式规范过程、区域化等(图3A), GSEA结果显示DUX会激活视黄醇代谢通路(图3B), 激活视黄酸

信号通路相关的表达(图3C)。视黄酸是视黄醇代谢的产物, 同样能促进2CLC的形成。对已发表的视黄酸处理的mESC测序数据进行比对分析, 发现DUX上调的富集通路与视黄酸促进的上调差异基因非常相似(图3D)。以上数据表明DUX可能激活视黄酸信号通路, 通过视黄酸途径影响细胞状态。

2.4 DUX占据视黄酸受体结合基序

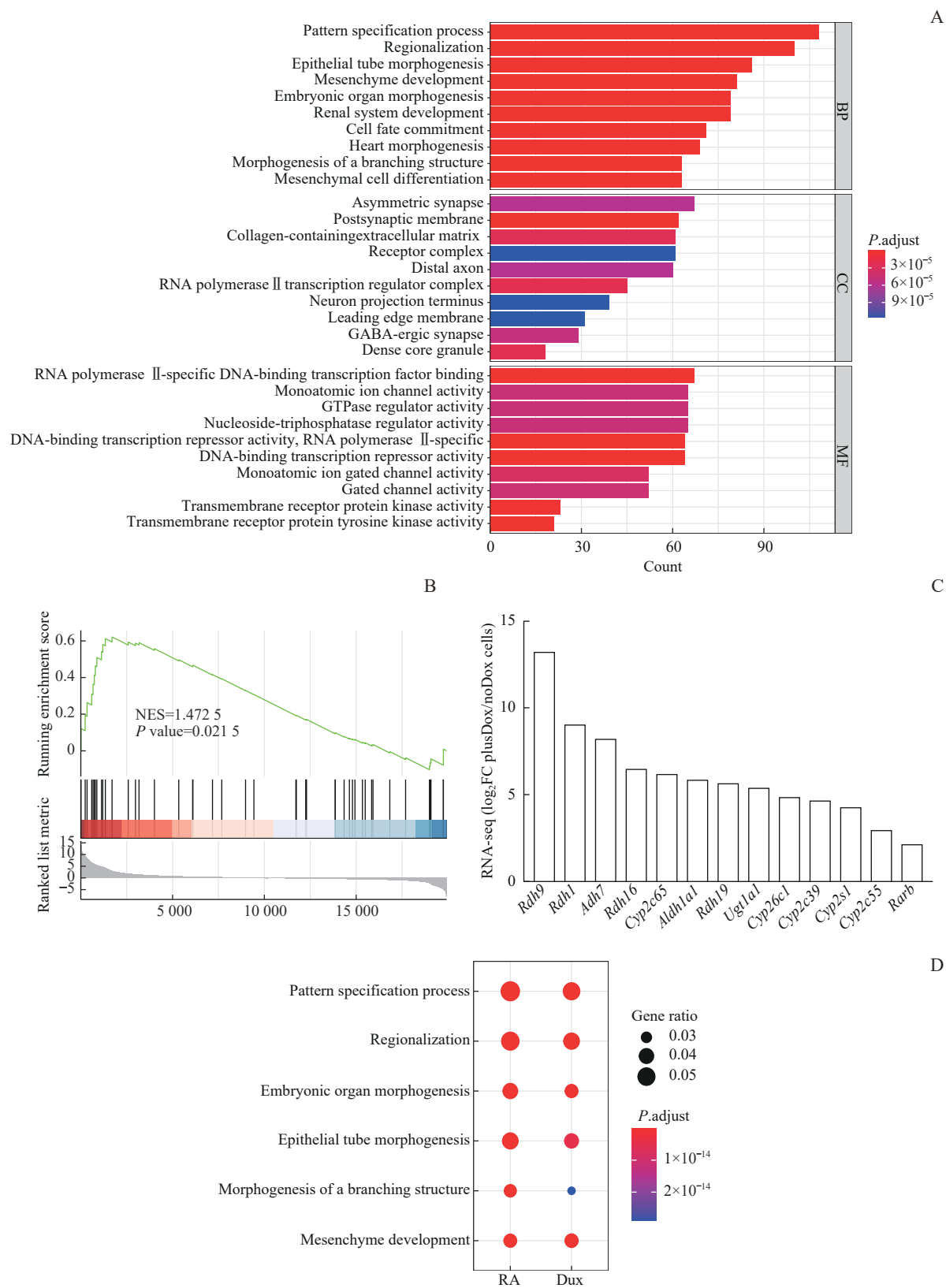
为了进一步探索DUX调控视黄酸通路的潜在分子机制, 通过已发表的DUX ChIP-seq数据, 使用AME软件, 分析视黄酸 γ 受体(retinoic acid receptor gamma, RARG)与DUX占据位点的共定位情况。结果显示: 两者交集很多是与发育相关的基因(图4A), 比如2C特异性基因*Zscan4*家族成员基因; 另外DUX与RARG共定位在*Gata4*基因的远端增强子上(图4B)。进一步KEGG分析发现, 该交集基因会富集在丝裂原活化蛋白激酶(mitogen-activated protein kinase, MAPK)信号通路(图4C)。这些分析结果表明DUX与视黄酸受体存在一些相同的DNA结合位点, 它们共同的靶基因构成了调节ESC分化潜能的网络。



Note: A. RT-qPCR validation of selected pluripotent marker genes in noDox (control) and plusDox (DUX overexpression) ES cells. B/C. Heatmap representation of expression levels of genes involved in 2CLCs (B) or endoderm (C) specification in noDox (control) and plusDox (DUX overexpression) cells. D. RT-qPCR was employed to assess the RNA expression levels of marker factors for endoderm, mesoderm and ectoderm on the fourth day of embryoid body differentiation in noDox (control) and plusDox (DUX overexpression) cells. ^① $P>0.05$, ^② $P<0.001$, ^③ $P=0.0013$, compared with the control group.

图2 DUX诱导的2CLC的发育潜能的检测

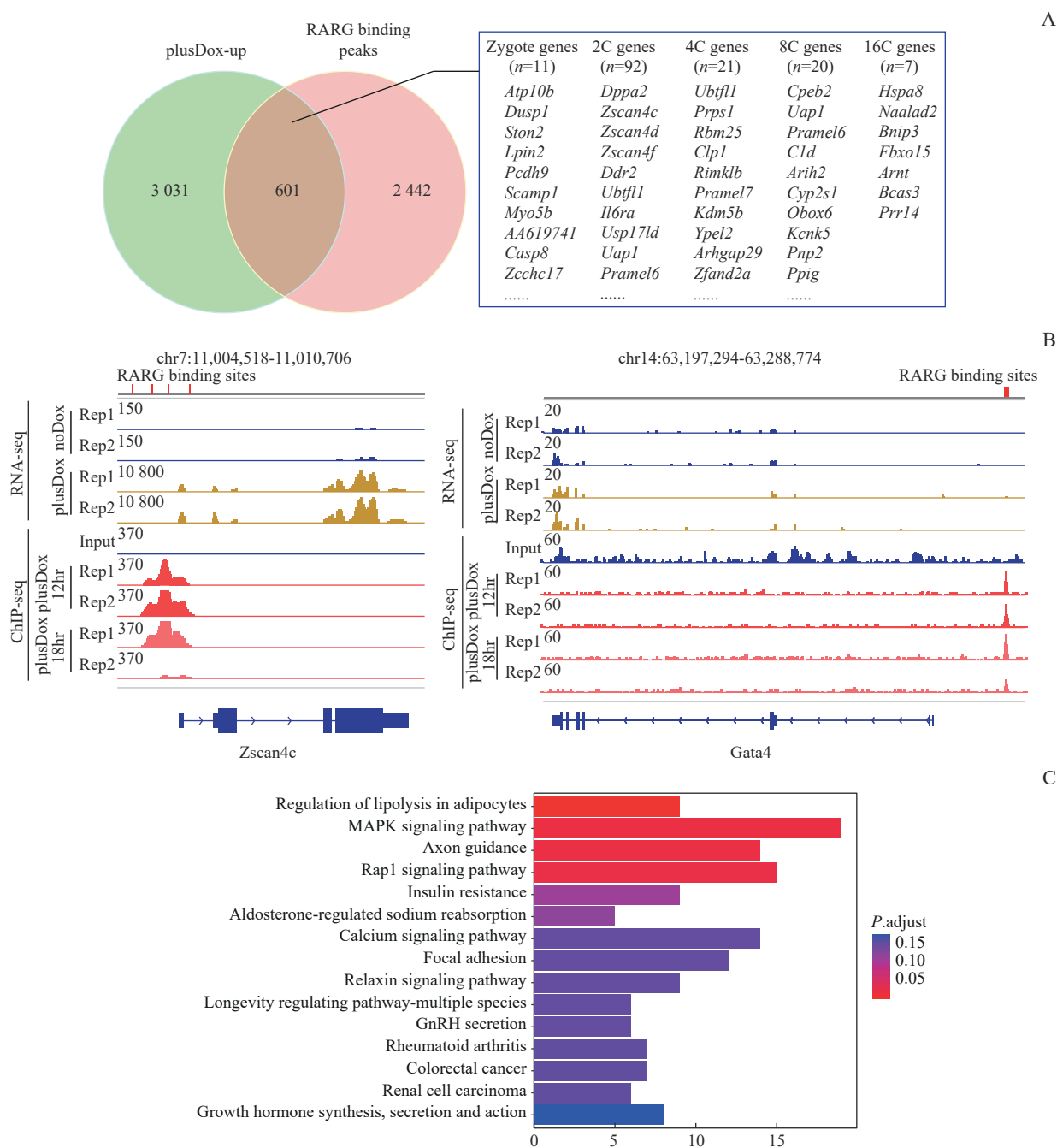
Fig 2 Detection of developmental potential of DUX-induced 2CLCs



Note: A. GO analysis of upregulated genes in plusDox ES cells. B. GSEA analysis showing collective changes in the retinol metabolism signaling pathway gene set. C. Expression levels (\log_2FC) (FC, fold change) of selected RA-pathway-related genes in plusDox and noDox cells (ESCs) based on RNA-seq data. D. Comparison of enrichment results from multiple gene lists.

图3 DUX过表达ESC的功能富集分析

Fig 3 Functional enrichment analysis of DUX overexpressed ESCs



Note: A. Intersection of DUX-upregulated genes with RARG and DUX co-localized genes and possible mechanism of regulating mESCs. B. IGV view of RNA-seq and ChIP-seq results. C. KEGG analysis of the intersecting genes.

图4 DUX调控ESC分化潜能的可能机制

Fig 4 Potential mechanism of DUX regulation on ESCs differentiation

3 讨论

DUX是一些2C特异性基因和内源性反转录病毒元件的驱动因素,然而尚不清楚DUX在2C样细胞退出全能性进而向下分化的倾向及潜在机制。本研究首先通过已发表文献的RNA-seq数据分析,发现在维持mESC自我更新培养条件下,DUX过表达的mESC会

高表达一些PrE的标志因子,随着mESC向拟胚体自然分化,过表达DUX会促使拟胚体向胚外内胚层分化。

GSEA结果显示DUX激活了视黄醇代谢通路。视黄醇代谢产物视黄酸是核受体超家族的成员^[25],控制着广泛的发育过程。最近有研究报道低剂量的视黄酸和视黄酸受体 γ 协同作用会促进mESC进入类2C

状态^[26]。本研究发现视黄酸途径促进的上调差异基因GO富集分析的功能区与DUX激活的高度相似,两者有相似的转录水平表达谱。这表明DUX可能通过视黄酸信号通路调节mESC状态。进一步对ChIP-seq数据进行生物信息学分析,证明已知的视黄酸受体motif存在于DUX富集的peaks中,两者能够结合在一些靶基因的启动子或增强子区域,调控这些基因的表达。有趣的是,RARG与DUX共同占据的基因,进一步与DUX激活的基因取交集,这些交集基因会富集在MAPK信号通路上,而持续激活的细胞外信号调节激酶1/2(extracellular regulated protein kinases 1/2, ERK1/2)磷酸化会导致mESC的分化^[27]。此外,视黄酸信号通路也可直接激活MAPK/ERK通路^[28],且视黄酸能够通过诱导GATA4等促分化因子的转录,促进内胚层来源的细胞分化^[29]。本研究发现DUX与视黄酸有相似的结合基序,进而调控相同的靶基因,过表达DUX激活了视黄酸信号通路;如果抑制视黄酸受体,可能会抑制DUX与视黄酸下游相同的靶基因的转录,这些基因既有调节重编程相关的功能,也有促进分化的能力。

此外,本研究将为DUX异常激活导致的疾病治疗提供新思路。人类同源蛋白DUX4在成体肌肉细胞中的异常上调与面肩胛骨肌营养不良症(facioscapulohumeral muscular dystrophy, FSHD)相关^[30]。尽管DUX与DUX4的同源结构域和结合序列存在差异,但是它们都会激活2C基因的表达,其中可能存在的机制为两者通过结合并激活不同的内源性反转录转座子,导致临近基因的强烈表达^[15]。在FSHD中DUX4的不当表达会引发细胞凋亡^[31],而在DUX诱导的2CLC中也观察到严重的DNA损伤^[32],

这种细胞毒性可能源自2C时期特异性基因的异常激活,比如PRAME(preferentially expressed antigen in melanoma)家族基因^[33]等。而在胚胎发育卵裂过程中,DUX的短暂激活并没有导致细胞死亡,这可能得益于ZSCAN4等相关因子的作用。已知ZSCAN4参与了ESC端粒和基因组稳定性的维持,以及保护2C胚胎免受DNA损伤^[34-36]。因此,DUX增强了Zscan4的转录激活,会在一定程度上弥补其对细胞造成的伤害。在视黄酸处理的2CLC形成体系中,PRAME和ZSCAN4家族基因都会被激活^[37-38]。鉴于DUX与视黄酸受体可能存在的交互影响,抑制视黄酸途径的某些关键分子可能在FSHD中抵消DUX4的作用,该领域的研究有望为此类疾病的靶向治疗奠定基础。

利益冲突声明/Conflict of Interests

所有作者声明不存在利益冲突。

All authors disclose no relevant conflict of interests.

作者贡献/Authors' Contributions

洪磊主要完成数据分析处理与文章撰写的工作;洪磊、蔡勤和李婉睿完成实验操作;郭传亮、薛燕、曾溢滔和曾凡一参与文章的写作与修改。所有作者均阅读并同意最终稿件的提交。

The work of data analysis and the manuscript writing were completed by HONG Lei. The experimental operation was completed by HONG Lei, CAI Qin, and LI Wanrui. The manuscript was drafted and revised by GUO Chuanliang, XUE Yan, ZENG Yitao, and ZENG Fanyi. All authors have read the final manuscript and approved the submission.

致谢 感谢贵安科创产业发展有限公司-贵安超级计算中心在提供数据分析平台方面的支持与帮助。

• Received: 2024-04-18

• Accepted: 2024-05-13

• Published online: 2024-11-28

参 · 考 · 文 · 献

- [1] SOLTER D. From teratocarcinomas to embryonic stem cells and beyond: a history of embryonic stem cell research[J]. Nat Rev Genet, 2006, 7(4): 319-327.
- [2] COCKBURN K, ROSSANT J. Making the blastocyst: lessons from the mouse[J]. J Clin Invest, 2010, 120(4): 995-1003.
- [3] EVANS M J, KAUFMAN M H. Establishment in culture of pluripotential cells from mouse embryos[J]. Nature, 1981, 292(5819): 154-156.
- [4] MARTIN G R. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells[J]. Proc Natl Acad Sci U S A, 1981, 78(12): 7634-7638.
- [5] YANG J, RYAN D J, WANG W, et al. Establishment of mouse expanded potential stem cells[J]. Nature, 2017, 550(7676): 393-397.
- [6] YANG Y, LIU B, XU J, et al. Derivation of pluripotent stem cells with *in vivo* embryonic and extraembryonic potency[J]. Cell, 2017, 169(2): 243-257. e25.
- [7] GAO X F, NOWAK-IMIALEK M, CHEN X, et al. Establishment of porcine and human expanded potential stem cells[J]. Nat Cell Biol, 2019, 21(6): 687-699.
- [8] LI R H, ZHONG C Q, YU Y, et al. Generation of blastocyst-like structures from mouse embryonic and adult cell cultures[J]. Cell, 2019, 179(3): 687-702. e18.
- [9] LIU K S, XU X C, BAI D D, et al. Bilineage embryo-like structure from EPS cells can produce live mice with tetraploid trophectoderm[J]. Protein Cell, 2023, 14(4): 262-278.
- [10] MACFARLAN T S, GIFFORD W D, DRISCOLL S, et al.

- Embryonic stem cell potency fluctuates with endogenous retrovirus activity[J]. *Nature*, 2012, 487(7405): 57-63.
- [11] BOŠKOVIĆ A, EID A, PONTABRY J, et al. Higher chromatin mobility supports totipotency and precedes pluripotency *in vivo*[J]. *Genes Dev*, 2014, 28(10): 1042-1047.
- [12] BAKER C L, PERA M F. Capturing totipotent stem cells[J]. *Cell Stem Cell*, 2018, 22(1): 25-34.
- [13] HU Y Y, YANG Y Y, TAN P C, et al. Induction of mouse totipotent stem cells by a defined chemical cocktail[J]. *Nature*, 2023, 617(7962): 792-797.
- [14] SHEN H, YANG M, LI S Y, et al. Mouse totipotent stem cells captured and maintained through spliceosomal repression[J]. *Cell*, 2021, 184(11): 2843-2859. e20.
- [15] WHIDDON J L, LANGFORD A T, WONG C J, et al. Conservation and innovation in the *DUX4*-family gene network[J]. *Nat Genet*, 2017, 49(6): 935-940.
- [16] HENDRICKSON P G, DORÁIS J A, GROW E J, et al. Conserved roles of mouse *DUX* and human *DUX4* in activating cleavage-stage genes and *MERVL/HERVL* retrotransposons[J]. *Nat Genet*, 2017, 49(6): 925-934.
- [17] DE IACO A, PLANET E, COLUCCIO A, et al. *DUX*-family transcription factors regulate zygotic genome activation in placental mammals[J]. *Nat Genet*, 2017, 49(6): 941-945.
- [18] FU X D, WU X J, DJEKIDEL M N, et al. *Myc* and *Dnmt1* impede the pluripotent to totipotent state transition in embryonic stem cells[J]. *Nat Cell Biol*, 2019, 21(7): 835-844.
- [19] HU Z H, TAN D E K, CHIA G, et al. Maternal factor *NELFA* drives a 2C-like state in mouse embryonic stem cells[J]. *Nat Cell Biol*, 2020, 22(2): 175-186.
- [20] YANG F, HUANG X, ZANG R G, et al. *DUX*-miR-344-ZMYM2-mediated activation of *MERVL* LTRs induces a totipotent 2C-like state[J]. *Cell Stem Cell*, 2020, 26(2): 234-250. e7.
- [21] YANG G, ZHANG L F, LIU W Q, et al. *Dux*-mediated corrections of aberrant H3K9ac during 2-cell genome activation optimize efficiency of somatic cell nuclear transfer[J]. *Cell Stem Cell*, 2021, 28(1): 150-163. e5.
- [22] XU R M, LI S, WU Q, et al. Stage-specific H3K9me3 occupancy ensures retrotransposon silencing in human pre-implantation embryos[J]. *Cell Stem Cell*, 2022, 29(7): 1051-1066. e8.
- [23] ZUO F F, JIANG J Y, FU H P, et al. A *TRIM66/DAX1/Dux* axis suppresses the totipotent 2-cell-like state in murine embryonic stem cells[J]. *Cell Stem Cell*, 2022, 29(6): 948-961. e6.
- [24] WANG Y Q, NA Q, LI X H, et al. Retinoic acid induces *NELFA*-mediated 2C-like state of mouse embryonic stem cells associates with epigenetic modifications and metabolic processes in chemically defined media[J]. *Cell Prolif*, 2021, 54(6): e13049.
- [25] GIGUERE V, ONG E S, SEGUI P, et al. Identification of a receptor for the morphogen retinoic acid[J]. *Nature*, 1987, 330(6149): 624-629.
- [26] ITURBIDE A, RUIZ TEJADA SEGURA M L, NOLL C, et al. Retinoic acid signaling is critical during the totipotency window in early mammalian development[J]. *Nat Struct Mol Biol*, 2021, 28(6): 521-532.
- [27] HAMILTON W B, BRICKMAN J M. Erk signaling suppresses embryonic stem cell self-renewal to specify endoderm[J]. *Cell Rep*, 2014, 9(6): 2056-2070.
- [28] YEN A, ROBERSON M S, VARVAYANIS S, et al. Retinoic acid induced mitogen-activated protein (MAP)/extracellular signal-regulated kinase (ERK) kinase-dependent MAP kinase activation needed to elicit HL-60 cell differentiation and growth arrest[J]. *Cancer Res*, 1998, 58(14): 3163-3172.
- [29] ARCECI R J, KING A A, SIMON M C, et al. Mouse *GATA-4*: a retinoic acid-inducible GATA-binding transcription factor expressed in endodermally derived tissues and heart[J]. *Mol Cell Biol*, 1993, 13(4): 2235-2246.
- [30] MITSUHASHI H, ISHIMARU S, HOMMA S, et al. Functional domains of the FSHD-associated *DUX4* protein[J]. *Biol Open*, 2018, 7(4): bio033977.
- [31] GENG L N, YAO Z Z, SNIDER L, et al. *DUX4* activates germline genes, retroelements, and immune mediators: implications for facioscapulohumeral dystrophy[J]. *Dev Cell*, 2012, 22(1): 38-51.
- [32] OLBRICH T, VEGA-SENDINO M, TILLO D, et al. CTCF is a barrier for 2C-like reprogramming[J]. *Nat Commun*, 2021, 12(1): 4856.
- [33] LEZCANO C, JUNGBLUTH A A, NEHAL K S, et al. *PRAE* expression in melanocytic tumors[J]. *Am J Surg Pathol*, 2018, 42(11): 1456-1465.
- [34] MARKIEWICZ-POTOCZNY M, LOBANOVA A, LOEB A M, et al. TRF2-mediated telomere protection is dispensable in pluripotent stem cells[J]. *Nature*, 2021, 589(7840): 110-115.
- [35] SRINIVASAN R, NADY N, ARORA N, et al. *Zscan4* binds nucleosomal microsatellite DNA and protects mouse two-cell embryos from DNA damage[J]. *Sci Adv*, 2020, 6(12): eaaz9115.
- [36] ZALZMAN M, FALCO G, SHAROVA L V, et al. *Zscan4* regulates telomere elongation and genomic stability in ES cells[J]. *Nature*, 2010, 464(7290): 858-863.
- [37] TAGLIAFERRI D, MAZZONE P, NOVIELLO T M R, et al. Retinoic acid induces embryonic stem cells (ESCs) transition to 2 cell-like state through a coordinated expression of *Dux* and *Duxbl1*[J]. *Front Cell Dev Biol*, 2019, 7: 385.
- [38] NAPOLITANO G, TAGLIAFERRI D, FUSCO S, et al. A novel member of *Prae* family, *Gm12794c*, counteracts retinoic acid differentiation through the methyltransferase activity of *PRC2*[J]. *Cell Death Differ*, 2020, 27(1): 345-362.

[本文编辑] 张慧俊

