

DOI: 10.13376/j.cblls/2017129

文章编号: 1004-0374(2017)10-0968-09



江赐忠, 同济大学生命科学与技术学院特聘教授, 博士生导师, 教育部“新世纪优秀人才”, 上海市“曙光学者”、“浦江人才”、“东方学者”获得者。承担多项国家重大科学研究计划、国家自然科学基金委重大研究与面上项目。研究方向为胚胎发育与细胞命运转化中的染色质重塑及其表观遗传调控机制。课题组近期主要工作包括: (1) 果蝇染色质重塑酶 Brahma 在早期胚胎发育中调控核小体定位的机制; (2) 小鼠体细胞重编程中组蛋白修饰变化与 Oct4 动态结合的分子路线图, 及两者互作对全能性获得与维持的作用机制; (3) 人胚胎干细胞向神经外胚层细胞分化中染色质重塑模式及其调控机制。在 *Nucleic Acids Res*、*Cell Reports*、*Cell Death Differ* 等著名学术期刊发表论文 60 余篇。这些发现极大提高人们对细胞命运转化中染色质重塑及其作用机制的理解。



高绍荣, 现任同济大学生命科学与技术学院院长, 国家杰出青年科学基金获得者, 教育部“长江学者”特聘教授, “万人计划”领军人才, 科技部中青年科技创新领军人才, 国家百千万人才工程“有突出贡献中青年专家”, 曾担任“干细胞研究”国家重大科学研究计划专家组成员。高绍荣教授主要利用体细胞核移植与诱导多能干细胞技术从事哺乳动物早期胚胎发育和体细胞重编程分子机制研究, 并在体细胞重编程分子机制和诱导多能干细胞多能性等多项研究中取得一系列重要进展, 已经在包括 *Nature*、*Science*、*Nat Genet*、*Cell Stem Cell* 等国际权威学术期刊发表论文 110 余篇。研究成果曾获得国家自然科学奖二等奖、“周光召基金会杰出青年基础科学奖”、“谈家桢生命科学创新奖”和“药明康德生命化学研究奖”等奖项。

细胞命运转化中的染色质重塑及其表观遗传调控机制

江赐忠*, 高绍荣*

(同济大学生命科学与技术学院, 上海市信号转导与疾病研究重点实验室, 上海 200092)

摘要: 染色质重塑是重要的表观遗传调控机制, 参与调控许多重要生物过程, 但细胞命运转化中染色质变化模式及其调控机制并不清楚。围绕这一问题, 江赐忠和高绍荣实验室研究了小鼠体细胞重编程及人胚胎干细胞 (ESC) 定向分化成神经外胚层细胞 (NEC) 中染色质重塑及其作用机制, 取得一系列前沿性进展: 揭示体细胞重编程中发生精确的核小体重塑, 使之获得与 ESC 没有差别的染色质结构; 绘制体细胞重编程中核心全能性转录因子 Oct4 的动态结合靶点与关键组蛋白修饰变化的分子路线图, 及其两者互作对全能性获得与维持的调控机制; 发现人 ESC 向 NEC 分化中, 紧挨转录起始位点上游的核小体缺失区发生核小体丢失, 激活 NEC 相关基因, 组蛋白乙酰基转移酶 KAT2B 增加 H3K9ac 信号来招募转录因子 Sox2 结合到 NEC 特异靶点激活靶点基因, 促进 NEC 分化。这些成果极大提高人们对细胞命运转化中染色质重塑及其表观遗传调控机制的认识。

收稿日期: 2017-08-19

基金项目: 国家自然科学基金项目(91019017, 91319306, 91519309)

*通信作者: E-mail: czjiang@tongji.edu.cn (江赐忠); gaoshaorong@tongji.edu.cn (高绍荣)

关键词: 胚胎干细胞; iPSC; 神经外胚层细胞; 组蛋白修饰; 核小体; Oct4
中图分类号: Q343.21; Q344.13 文献标志码: A

Chromatin remodeling and its epigenetic regulatory mechanisms in cell fate transition

JIANG Ci-Zhong*, GAO Shao-Rong*

(Shanghai Key Laboratory of Signaling and Disease Research, School of Life Sciences and Technology, Tongji University, Shanghai 200092, China)

Abstract: Chromatin remodeling is an important epigenetic regulatory mechanism, and takes part in controlling many biological processes. However, the pattern and the functions of chromatin remodeling in cell fate transition remain enigmatic. To address this issue, we studied chromatin remodeling in mouse somatic cell reprogramming and the differentiation of human embryonic cells (ESC) into neuroectodermal cells (NEC), respectively, and achieved a series of progress. The results show that accurate nucleosome remodeling takes place and results in a chromatin structure in iPSC highly similar to that in ESC. The core pluripotency factor Oct4 plays pivotal roles in somatic reprogramming. We depicted the molecular roadmap of dynamic Oct4 binding and key histone modification changes in the course of somatic reprogramming, and revealed the functions of their interactions in gain and maintenance of pluripotency. In the process of human ESC differentiating to NEC, we found that nucleosome eviction occurs in the nucleosome depletion regions right upstream of transcription start sites and activates these NEC-related genes. Acetyltransferase KAT2B deposits H3K9ac signal to recruit the transcription factor Sox2 binding to the target sites specific in NEC and activate the target genes, therefore facilitating the differentiation of NEC. These findings greatly improve our understanding of chromatin remodeling in cell fate transition and its associated epigenetic regulatory roles.

Key words: embryonic stem cells; iPSC; neuroectodermal cells; histone modification; nucleosome; Oct4

多能干细胞是一类具有自我更新与多向分化潜能的细胞。多能性是指具备分化成各种胚层, 甚至发育为整个个体的能力。自我更新是指无限复制生长而不改变细胞特征的能力。多能干细胞主要可分为胚胎干细胞与诱导性多能干细胞 (iPSC)。由于多能干细胞是研究发育与疾病的很好体外模型, 以及在转化医学中有巨大的潜在应用价值, 干细胞的研究受到极大重视并已经成为热点。多能性的获得与维持、定向分化效率和安全性是干细胞研究与应用的三个核心问题。尽管这些年对干细胞的研究取得许多重大成果, 但对这三个核心问题的分子机制还没有完全了解清楚。

1 干细胞的获得

胚胎干细胞 (ESC) 是最原始的多能干细胞, 来源于囊胚的内细胞团。1981年, 生物学家首次从小鼠囊胚中分离并成功建立小鼠 ESC 系^[1-2]。而直到相隔 17 年后的 1998 年, 科学家才从人囊胚中成功分离得到人的 ESC^[3]。这些所分离并成功体外培养

的细胞可在一定条件下分化成各种不同的组织细胞, 极大促进发育学中复杂事件的分子机制研究, 以及细胞替代疗法和器官移植的发展。然而, 从早期胚胎中分离干细胞有很高的技术要求, 存在免疫排斥与伦理上的问题, 这极大地限制 ESC 在基础研究与临床上的应用。因此, 需要建立非来源于胚胎的, 没有免疫排斥的多能干细胞。

2006 年, 日本 Yamanaka 实验室将外源的四种转录因子 (Oct4、Sox2、c-Myc 和 Klf4) 导入小鼠成纤维细胞 (MEF), 成功逆转得到多能干细胞^[4]。随后 2007 年, 日本和美国研究人员又各自独立把人体皮肤细胞改造成 iPSC^[5-6]。然而, 这种通过慢病毒介导过表达外源转录因子重编程体细胞获得的 iPSC 存在安全隐患。2013 年, 北京大学邓宏魁课题组利用一组化学小分子成功重编程体细胞, 获得在基因表达谱、表观状态、分化潜能与生殖传递上和 ESC 相似的 iPSC^[7]。由于 iPSC 可用患者自身体细胞诱导获得而不具有免疫排斥性, 且 iPSC 不用到胚胎而不存在伦理问题, 因此, iPSC 的诞生极大

地促进了干细胞的基础研究与临床应用。

2 干细胞多能性与染色质调控

细胞命运转化,即干细胞分化成组织细胞或体细胞重编程为 iPSC。这一过程细胞状态发生巨大变化,由复杂的调控网络完成。而该过程中精密的转录调控是在一个染色质开放的环境下实现的。事实上,人们已经逐渐认识到在多能性干细胞中染色质是开放的,而在分化的细胞中染色质是关闭的^[8]。电子显微镜技术首次证实 ESC 与分化细胞中染色质结构的差别,即分化细胞中广泛存在异染色质而在未分化细胞中则很少^[9]。

组蛋白修饰是重要的染色质属性,可以标记区分出染色质活跃区与抑制区。ESC 中染色质全局上开放,并富含活跃型组蛋白修饰 H3K4me3 与组蛋白 H3 和 H4 的乙酰化^[10-11]。相反,ESC 中含有低水平的异染色质关联的组蛋白修饰^[12]。与此结论相一致的发现是,异染色质标记 H3K9 甲基化可以阻碍体细胞重编程^[13]。因此,染色质状态在干细胞多能性获得与维持中有重要调控作用。理解染色质重塑对干细胞多能性的分子调控机制是非常重要的,将促进多能性干细胞在临床安全高效的应用。

3 体细胞重编程中染色质的精确重塑

经体细胞重编程获得的 iPSC 和 ESC 一样具有全能性与自我更新的能力,多能性标记基因的表达相似,而且均能发育成具有繁殖能力的子代。Guenther 等^[14]研究进一步表明,人 ESC 与 iPSC 在基因表达谱、DNA 甲基化及关键组蛋白修饰模式方面高度相似。然而,在多能性获得与维持中有重要调控作用的染色质结构在 iPSC 与 ESC 中有多大相似并不清楚。

为解决这一问题,我们利用小鼠次级体细胞重编程技术,从第一代 all-iPSC 小鼠的脂肪前体细胞与造血前体细胞(均为中胚层细胞)、胃内皮细胞(内胚层细胞)和表皮细胞(外胚层细胞)经诱导过表达外源 4 个转录因子(Oct4、Sox2、c-Myc 和 Klf4)重编程细胞得到次级 iPSC,分别建成 iPSC 系 32、16-6、T2 与 S8,以小鼠 ESC 系 R1 作为 ESC 代表。用 RNA-seq 技术测定这 5 种细胞系的基因表达谱。核小体是真核生物染色质的基本结构单元。核小体在基因组上的定位直接影响染色质结构。因此,我们用 MNase-seq 技术测定这 5 种细胞系的全基因组高分辨率核小体定位图谱。

无论是转录水平,还是蛋白质水平,多能性标记基因 Oct4、Sox2、Klf4、c-Myc、Nanog、Lin28a 在这 5 种全能性细胞中都处于高水平,没有差别。相反,组织分化基因 SSEA4 在这 5 种细胞中都没有表达。进一步基因表达谱相关性分析显示,基因表达谱在这 5 种细胞两两之间有高度相关性(spearman $R \geq 0.95$)。这表明体细胞重编程获得的 iPSC 与 ESC 在基因表达谱上没有差别。

我们用 10 kb 的窗口扫描整个基因组,计算每个窗口的核小体信号,从而得到全基因组核小体信号分布。比较分析结果显示,全基因组核小体信号分布在这 5 种细胞两两之间高度相似(spearman $R \geq 0.9$)。而且,这 5 种多能干细胞与前人发表的小鼠 ESC 全基因组核小体信号分布高度相似(spearman $R \geq 0.8$),但与前人发表的小鼠 MEF 全基因组核小体信号分布存在显著差别(spearman $R \leq 0.4$)。

核小体在转录起始位点(TSS)附近呈现经典的 -1、NDR (nucleosome depletion region,核小体缺失区)、+1、+2、+3 等分布模式,对基因转录调控有着重要作用^[15]。因此,我们着重比较这 5 种多能性细胞 TSS 附近核小体的分布。结果显示,这 5 种多能性细胞 TSS 附近的这种核小体分布模式没有差别,而且对基因转录调控作用也一致,即表达基因 TSS 附近呈现上述经典核小体分布模式,TSS 落在 NDR 内;而沉默基因 TSS 正好被一个核小体占据,TSS 附近没有其他核小体占据。

由各种多能性转录因子构成的调控网络对多能干细胞状态的维持起重要作用。因此,我们进一步分析核小体定位与这些转录因子结合靶点的空间关系。结果表明,核心转录因子 Oct4、Sox2、Nanog 等结合到两个核小体间的连接区。由于连接区是开放的染色质区,便于这些转录因子结合到靶点,从而维持细胞多能性。另一类与自我更新相关的转录因子(如 Essrb、Zfx、Tcfcp2l1 等)结合在核小体上,它们可能通过核小体旋转设置使转录因子结合到靶点。有意思的是,转录因子 E2F1 与 c-Myc 倾向结合在核小体两端边上。Orkin 和 Hochedlinger^[16]报道,c-Myc 不在含有 Oct4、Sox2、Nanog 等多能性转录因子的核心多能性调控网络中。所以,c-Myc 与 Oct4、Sox2、Nanog 的结合靶点和核小体有不同的空间拓补关系就不奇怪了。而绝缘蛋白 CTCF 则以核小体大小为间隔周期性地结合到连接区。这种结合模式有助于 CTCF 作为绝缘子对抑制与激活染色质区域进行划界。这 4 类转录因子结合靶点与核小

体的空间关系在这 5 种多能性细胞中都是一样的。

综上所述, 体细胞重编程中发生精确的核小体重塑, 使得核小体定位模式、对基因活性的调控关系及与不同转录因子形成各自独特的空间关系, 在不同胚层来源诱导得到的 iPSC 与 ESC 中没有区别, 从而达到精确调控基因转录表达, 获得与维持多能性 (图 1) 的目的。该结果发表在 *BMC Biology* 杂志上^[17]。

4 体细胞重编程中组蛋白修饰与 Oct4 的分子调控路线

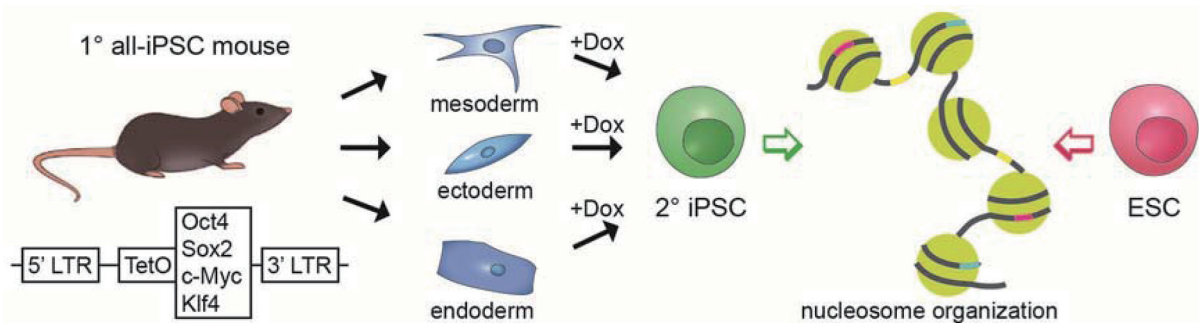
核心多能性转录因子 Oct4 在体细胞重编程中起重要转录调控功能。然而, Oct4 在体细胞重编程中对 DNA 的动态结合与调控作用并不清楚。表观修饰的变化在体细胞重编程中也起重要作用。Polo 等^[18]报道, 小鼠 iPSC 的四因子诱导过程中, 共价区域在第一波基因表达变化后逐渐建立, 而 DNA 甲基化变化则发生在第二波基因表达变化之后。但是, 在体细胞重编程中, Oct4 动态结合与染色质结构变化的相互作用及其在多能性的获得与维持中的调控作用并不清楚。

为解决这个问题, 我们对小鼠次级体细胞重编程中的 9 个时间点 (MEF, 第 1、3、5、7、11、15、18 天, 成熟 iPSC) 的细胞, 测定转录组、关键组蛋白修饰以及外源的和总的 Oct4 的 DNA 结合位点, 再用生物信息学方法分析体细胞重编程中基因表达变化、Oct4 对 DNA 的动态结合和组蛋白修饰变化, 及其在多能性获得与维持中的作用。结果显示, 外源 Oct4 与总 Oct4 在基因组上的结合情况, 从 MEF 到重编程的第 7 天相关性都很高, 但从第 11 天起相关性开始显著下降。这暗示很可能内源 Oct4 第 7 天后开始激活转录表达, 并慢慢取代外源 Oct4, 导

致外源 Oct4 在总 Oct4 中的比重下降。定量 PCR 证实从第 11 天起内源 Oct4 开始转录, 并随着重编程进行快速上升。Western blotting 也证实从第 11 天起内源 Oct4 蛋白开始出现。对 DNA 上总 Oct4 结合信号的聚类分析得到 7 类: 有 4 类是从第 1 天开始就结合, 其中一类第 7 天后就脱落了, 两类则到晚期 (第 15 天) 后才脱落, 另一类则自始至终都保持结合, 这一类显著富集在启动子区; 有 2 类则分别从中晚期才开始并保持结合到成熟 iPSC; 最后一小类只在第 7~15 天期间有结合。有意思的是, Oct4 一旦结合到靶点就很少在中间脱落, 反之亦然。此外, Oct4 结合随着重编程趋向于向 TSS 靠拢。

Oct4 靶点序列分析发现, 所有靶点都富集核心多能性转录因子 (如 Sox2、Nanog、Smad) 的基序 (motif), 这说明核心全能性调控网络在重编程中的重要性。此外, 重编程中自始至终都结合的这组 Oct4 靶点序列则富集 Elf 家族等多种转录因子基序。有意思的是, 这些转录因子 (如 Yy1 与 Nfya) 的转录水平也在重编程中自始至终维持高位; 而在重编程中晚期才开始结合的那两组 Oct4 靶点序列则主要富集与多能性相关转录因子 (如 Esrrb、Tcfap2c、Zfp423、Nr5a2) 的基序, 有意思的是, 这些转录因子的转录水平也从中晚期开始上升。这说明 Oct4 结合精确激活相关基因转录。

那影响 Oct4 结合的因素又有什么? 利用已发表的 MEF 与多能性干细胞染色质结构数据分析发现, 重编程第 1 天 Oct4 在靶点的结合信号强弱与这些靶点在 MEF 时的核小体占有成反比。反之, 那些重编程一开始就结合 Oct4 的靶点上的 FAIRE 信号 (代表染色质开放程度) 第 1 天就很强, 而中晚期才结合 Oct4 的靶点上第 1 天没有 FAIRE 信号。这些发现表明, 重编程中染色质的开放是 Oct4 结



初级all-iPSC小鼠的不同胚层细胞加Dox诱导后获得次级iPSC。该重编程中发生精确的核小体重塑, 使次级iPSC获得与ESC没有差别的核小体分布。DNA链上的红、黄、蓝短线代表与核小体有特定空间关系的与多能性相关的不同类型转录因子基序。

图1 小鼠体细胞重编程中的精确核小体重塑模型图

合到靶点的一个前提。与此一致的是,重编程中 Oct4 靶点上的结合信号与激活性组蛋白修饰 H3K4me3 和 H3K27ac 信号呈正相关,而与抑制性组蛋白修饰 H3K27me3 没有相关性。此外,重编程一开始就结合,但中期脱落的这组 Oct4 靶点上的 DNA 甲基化水平在重编程第 1 天时显著低于成熟的 iPSC。相反,重编程中晚期才开始结合但保持到成熟 iPSC 的这组 Oct4 靶点上的 DNA 甲基化水平在重编程第 1 天时显著高于成熟的 iPSC。这表明 Oct4 靶点在 Oct4 结合前有去 DNA 甲基化的过程。总的来说,在体细胞重编程中,Oct4 靶点发生表观信号变化帮助 Oct4 结合到靶点上,促进多能性获得与维持。

启动子是重要的基因转录调控元件。我们发现 Oct4 在启动子区的结合信号变化能准确地把重编程过程中分为起始、成熟和稳定三个阶段。而且,启动子区 Oct4 结合变化与组蛋白修饰变化协同调控基因转录。例如,重编程中 Oct4 一直保持结合的启动子上的 H3K4me3 也一直维持高水平,但 H3K27me3 维持低水平或无;相应地,这些基因(主要为持家基因)转录水平远高于平均水平。有意思的是,重编程中中晚期 Oct4 开始脱落的启动子上的 H3K4me3 也在中晚期下降或丢失,而 H3K27me3 则在中晚期上升。相反,重编程中中晚期 Oct4 才开始结合的启动子上的 H3K4me3 也在中晚期上升,而 H3K27me3 则在中晚期下降或丢失。这说明启动子区染色质重塑协同 Oct4 结合到靶点一道调控基因转录。

增强子是另一类远离 TSS 的转录调控元件。我们用组蛋白修饰 H3K4me1 定义出增强子,再根据增强子上的 H3K27ac 信号得到增强子状态:关闭(H3K4me1-),激活(H3K4me1+/H3K27ac+),停滞(H3K4me1+/H3K27ac-)。结果表明,重编程中 99.2% 的增强子状态都发生变化。MEF 激活的增强子在重编程开始到第 7 天期间大部分变为非激活的。而 iPSC 激活的增强子大部分是从第 15 天开始变为激活的。有意思的,增强子在由激活变为非激活后就维持非激活状态。类似地,增强子在由非激活变为激活后就维持激活状态。一直维持停滞状态的增强子主要富集发育与分化功能,而一直维持激活状态的增强子主要富集基本生物过程的持家基因功能。有意思的是,在重编程中晚期才激活的增强子主要富集干性获得功能。这说明增强子上的染色质状态变化与体细胞重编程中多能性的获得与维持密

切相关。

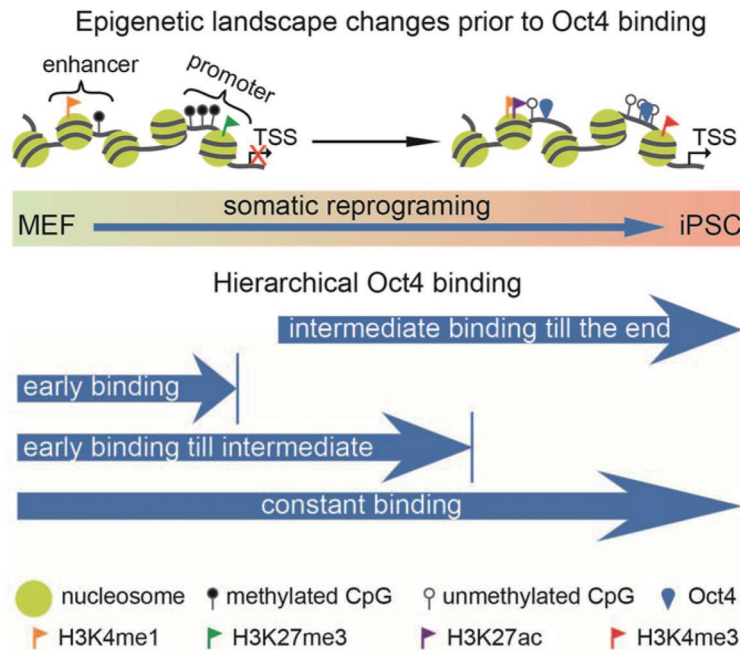
类似地,我们检测重编程中增强子上组蛋白修饰的变化与 Oct4 结合变化的关系,发现激活性组蛋白修饰 H3K4me1 和 H3K27ac 与 Oct4 结合都有很高的正相关,而抑制性 H3K27me3 和主要分布在启动子区的激活性 H3K4m3 都与 Oct4 结合不相关。我们进一步对 ESC 特异的超级增强子分析,发现 H3K27ac 信号逐渐增强的模式与 Oct4 信号增强模式完全一致^[19]。这说明很可能随着 H3K27ac 信号增强,相应染色质变开放,让增强子激活,招募上调转录因子结合,从而激活基因,促进多能性获得与维持。

多能性的建立是通过多能调控网络实现的。为此,我们收集了构成多能性调控网络的基因。根据它们在重编程中的表达谱,分为一直激活的、早期、中期和晚期激活的 4 组。有意思的是,每组基因的表达模式与对应增强子上 Oct4 结合模式相一致,即一直激活的基因的增强子上一直有 Oct4 结合。晚期激活的基因的增强子也是到晚期 Oct4 才结合上。这表明 Oct4 层级式地结合到不同多能性调控网络相关基因的增强子上,然后激活靶基因,使细胞获得多能性。

综上所述,体细胞重编程中,Oct4 连续地结合到靶点,很少结合上去又掉下来,反之亦然。染色质重塑与 DNA 甲基化变化使染色质处于“允许”状态,促进 Oct4 层级式地结合到靶点,在不同阶段激活不同功能的基因,使细胞获得多能性。该研究结果绘制了体细胞重编程中组蛋白修饰变化与 Oct4 结合的动态变化分子路线图,并揭示两者互作在多能性获得与维持中的作用机制(图 2)。该结果发表在 *Cell Reports* 杂志上^[20]。我们的结果发表后,加州大学的一个研究组在 *Cell* 发表论文利用类似的方法研究了 4 个转录因子的结合情况^[21],结果与我们的结论相似。

5 人 ESC 向神经外胚层细胞(NEC)分化中的染色质重塑及表观调控机制

染色质重塑提供转录调控的表观基础,对许多生物过程,如细胞分化、胚胎发育等有重要作用。多能干细胞与分化细胞之间的全基因组染色质结构比较发现,核小体占有差别富集在与多能性和分化关联的位点^[22]。人多能干细胞与供体成纤维细胞存在许多 H3K4me3 与 H2K27me3 分布不同的区域^[14]。NEC 能分化成各类神经细胞,是中枢神经系统发育



小鼠体细胞重编程中表观修饰发生变化, 促进Oct4层级式地结合到靶点激活靶基因, 使细胞获得与维持全能性。

图2 小鼠体细胞重编程中Oct4动态结合与组蛋白修饰变化的分子路线图

的重要资源^[23]。但人ESC分化成NEC过程中染色质结构发生怎样变化, 如何调控神经分化重要转录因子与基因转录并不清楚。为此, 我们利用人ESC向NEC体外分化系统, 制备这两类细胞的基因表达谱、核小体图谱、核心组蛋白修饰图谱以及与NEC分化关联的转录因子Sox2的DNA结合图谱, 进行表观基因组学分析, 揭示染色质结构在ESC向NEC分化中的变化模式及表观调控机制。

ESC与NEC的核小体定位图谱比较发现, NEC中全局上核小体占有升高, 核小体模糊性(fuzziness)下降。这表明分化后, NEC中的核小体更致密, 定位更稳定。结合表达谱数据, 我们发现NDR(核小体缺失区)中的核小体占有与基因表达水平是呈反比的。有意思的是, ESC中在紧挨TSS上游的经典NDR上有一部分基因被核小体占有。这组基因NDR上的核小体在NEC中丢失而呈现开放状态, 表达水平也显著上升。这组基因包括许多NEC分化中起重要作用的转录因子, 如Pax6、Sox9、Lhx2、Meis2等。进一步分析发现这些丢失的核小体与H2A.Z核小体富集在一起。这表明H2A.Z可能促进了NDR上核小体的解聚。

比较分化前后启动子上的组蛋白修饰发现主要变化是, 只含H3K9ac修饰的启动子数量上升, 这主要是因为H3K4me3+/H3K9ac+启动子数量的减

少; H3K9ac+/H3K27me3+的启动子数量也上升, 这主要是因为H3K4me3+/H3K9ac+/H3K27me3+启动子数量的减少; H3K4me3+/H3K9ac+/H3K27me3+的启动子数量下降。有意思的是, 上述NDR发生核小体丢失的那组基因的TSS附近H3K9ac信号显著上升, H3K27me3下降, 而H3K4me3不变, 因此H3K36me3水平与RNA Pol II结合显著上升, 促进基因转录。

为了鉴定哪种组蛋白修饰在NEC分化中起关键作用, 我们收集多能性、NEC、其他谱系3组特异基因集合。结果发现, 多能性特异基因的抑制主要是通过启动子上H3K4me3的下降与H3K27me3的上升; 而NEC特异基因的激活主要是通过H3K27me3的下降与H3K9ac的上升; 其他谱系特异基因表达水平变化不大, 只有H3K4me3发生轻微下降。这暗示NEC分化中H3K9ac与H3K27me3可能起重要作用。

为了验证我们的猜想, 我们构建了组蛋白修饰指数 $HMI = \log_2(FC_{K9ac}) - \log_2(FC_{K27me3})$, FC指某种组修饰水平NEC对ESC的比值。HMI的生物学意义是分化后激活性H3K9ac与抑制性H3K27me3的叠加变化。对 $HMI \leq -1$ 的基因, 即启动子组蛋白修饰变化为净抑制性水平达到2倍或以上的基因, 它们的表达水平在NEC中显著下降。相反, 对

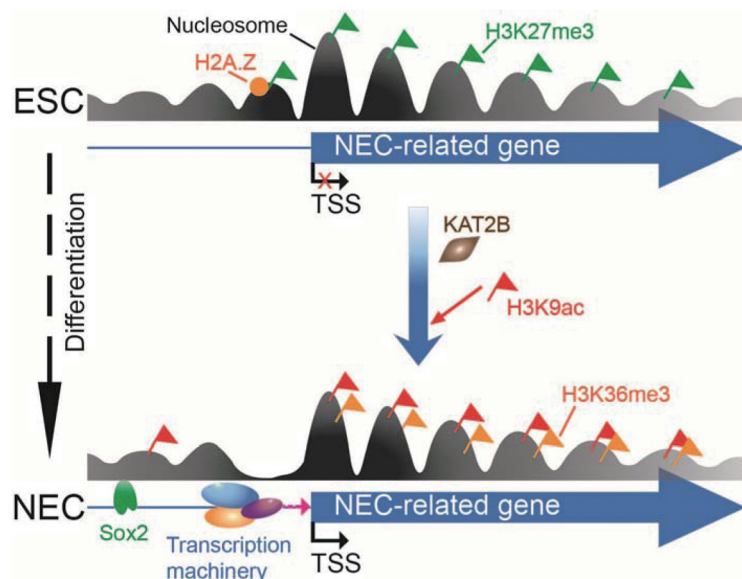
HMI ≥ 1 的基因, 即启动子组蛋白修饰变化为净激活水平达到 2 倍或以上的基因, 它们的表达水平在 NEC 中显著上升。而对于 $-1 \leq \text{HMI} \leq 1$ 基因, 它们的表达水平没有显著变化。接着, 结合转录组数据, 分别计算 H3K4Me3、H3K9ac、H3K27me3 与 H3K36me3 的受试者工作特征曲线 (receiver operating characteristic curve, 简称 ROC 曲线), 发现 H3K9ac 与 H3K27me3 能很好地预测基因表达水平变化。接着, 鉴定分化中 H3K9ac 或 H3K27me3 发生显著变化的区域, 发现这些区域主要富集在启动子区, 而且富含与 NEC 分化中起重要作用的转录因子的基序 (motif), 如 Pax6、Sox2、Sox9、Six6、Nr2f2 等。这证实 H3K9ac 与 H3K27me3 在 NEC 分化中确实起重要调控作用。

我们在实验中发现, 多能性转录因子 Sox2 在 NEC 分化中也持续高表达。那 Sox2 是如何在这两种截然不同的细胞中发挥功能的? ESC 与 NEC 中全基因组 ChIP-seq (染色质免疫沉淀结合测序) 实验测定 Sox2 在这两种细胞中的结合位置, 比较发现 Sox2 在 ESC 与 NEC 中只有很少的共同结合位点, 大部分都是特异性结合位点。同时, 在 ESC 特异的结合位点含有大量全能性转录因子的基序, 如 Oct4、Nanog、Tcf-Nanog 等。而在 NEC 特异的结合位点含有大量 NEC 分化相关转录因子的基序, 如 Pax6、Sox9、Nkx2.1 等。这表明 Sox2 协同不同

转录因子调控不同靶基因, 在 ESC 中维持全能性, 而在 NEC 中促进向神经分化。

接着, 我们研究染色质重塑对 Sox2 结合靶点的调控作用。结果发现, ESC 与 NEC 中的 Sox2 靶点上的核小体占有没有差别, 但 H3K9ac 信号显著上升, H3K27me3 信号显著下降。同时, Sox2 靶点显著富集在 H3K9ac 或 H3K27me3 发生显著变化的区域。这表明 H3K9ac 与 H3K27me3 对 Sox2 在靶点的结合有重要调控作用。巧合的是, 转录组数据发现组蛋白乙酰基转移酶 KAT2B 在分化后表达发生显著变化。因此, 我们推断 KAT2B 在 NEC 分化中起重要作用。为此, 我们在 ESC 中敲降 Kat2b 再进行分化, 发现 NEC 中特异的 Sox2 靶点上的 H3K9ac 信号与 Sox2 结合水平显著下降, 甚至消失, 另一个 NEC 中关键的转录因子 Pax6 结合水平也显著下降, 甚至消失, 因此分化不到 NEC。这说明 KAT2B 通过增加 Sox2 靶点的 H3K9ac 信号招募 Sox2 结合, 从而激活基因促进 NEC 分化。

综上所述, 人 ESC 向 NEC 分化过程中, 全局上核小体占有升高, 定位更稳定。紧挨 TSS 上游的 NDR 发生核小体丢失, 激活 NEC 分化中起重要作用的基因。启动子上 H3K9ac 升高与 H3K27me3 下降协同激活 NEC 关联基因转录。组蛋白乙酰基转移酶 KAT2B 通过增加 H3K9ac 信号招募 Sox2 与 Pax6 结合到 NEC 特异性靶点, 激活相应靶基因,



H2A.Z促进启动子区NDR的核小体逐出, 让转录机器结合上; 伴随H3K27me3信号的减弱或丢失, 乙酰基转移酶KAT2B增加H3K9ac信号, 招募Sox2结合到靶点, 两者共同激活NEC关联基因, 从而促进ESC向NEC分化。

图3 人ESC向NEC分化中染色质重塑及其调控作用机制图

促进 NEC 分化 (图 3)。该结果发表在 *Cell Death and Differentiation* 杂志上^[24]。

6 结语

以“细胞编程与重编程的表观遗传机制”重大研究计划为契机, 我们实验室重点研究细胞命运转化中染色质重塑及其表观遗传调控作用机制, 取得一系列前沿性进展: 揭示了小鼠体细胞重编程中的精确核小体重塑, 获得与 ESC 没有区别的染色质结构, 使之有一样的转录程序, 获得与维持多能性^[17]; 绘制了体细胞重编程中 Oct4 动态结合靶点与关键组蛋白修饰的分子路线图, 及阐述两者互作在多能性获得与维持中的作用机制^[20]; 揭示人 ESC 向 NEC 分化中染色质重塑模式, 及组蛋白乙酰基转移酶 KAT2B 通过增加 H3K9ac 信号招募转录因子 Sox2 结合到 NEC 特异靶点, 激活相应靶基因, 促进 NEC 分化^[24]。这一系列自主创新研究成果提高我们对细胞命运转化中染色质重塑及其表观调控功能机制的认识, 为从染色质结构上研究干细胞多能性状态及分化效率提供新方法。此外, 课题所获得的高通量数据是研究细胞命运转化中表观遗传调控机制的宝贵数据。

干细胞质量评估是其走向临床应用不容忽视的一步, 也是具有挑战性的一步, 是未来干细胞领域的一个研究难点与热点。传统的分子细胞生物学手段耗时、复杂、不稳定, 而且昂贵, 目前没有可靠标准。累积的遗传突变是衡量 iPSC 质量的一个指标。我们研究发现 Tet 诱导的体细胞重编程会造成突变, 这些突变累积会降低 all-iPSC 小鼠的存活能力。连续重编程六代后的 iPSC 质量大大降低, 产生的 all-iPSC 小鼠完全不能存活^[25]。前述本课题组发现细胞命运转化中的精确染色质重塑, 这表明或许能通过鉴定到起“开关”作用的关键基因组位点或区域的染色质状态来作为评估干细胞质量的标准, 这需要将来更多的探索证明。

[参 考 文 献]

- [1] Evans MJ, Kaufman MH. Establishment in culture of pluripotential cells from mouse embryos. *Nature*, 1981, 292: 154-6
- [2] Martin GR. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci USA*, 1981, 78: 7634-8
- [3] Thomson JA, Itskovitz-Eldor J, Shapiro SS, et al. Embryonic stem cell lines derived from human blastocysts. *Science*, 1998, 282: 1145-7
- [4] Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*, 2006, 126: 663-76
- [5] Takahashi K, Tanabe K, Ohnuki M, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*, 2007, 131: 861-72
- [6] Yu J, Vodyanik MA, Smuga-Otto K, et al. Induced pluripotent stem cell lines derived from human somatic cells. *Science*, 2007, 318: 1917-20
- [7] Hou P, Li Y, Zhang X, et al. Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds. *Science*, 2013, 341: 651-4
- [8] Gaspar-Maia A, Alajem A, Meshorer E, et al. Open chromatin in pluripotency and reprogramming. *Nat Rev Mol Cell Biol*, 2011, 12: 36-47
- [9] Park SH, Park SH, Kook MC, et al. Ultrastructure of human embryonic stem cells and spontaneous and retinoic acid-induced differentiating cells. *Ultrastruct Pathol*, 2004, 28: 229-38
- [10] Azuara V, Perry P, Sauer S, et al. Chromatin signatures of pluripotent cell lines. *Nat Cell Biol*, 2006, 8: 532-8
- [11] Krejci J, Uhlirova R, Galiova G, et al. Genome-wide reduction in H3K9 acetylation during human embryonic stem cell differentiation. *J Cell Physiol*, 2009, 219: 677-87
- [12] Meshorer E, Yellajoshula D, George E, et al. Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Dev Cell*, 2006, 10: 105-16
- [13] Chen J, Liu H, Liu J, et al. H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs. *Nat Genet*, 2013, 45: 34-42
- [14] Guenther MG, Frampton GM, Soldner F, et al. Chromatin structure and gene expression programs of human embryonic and induced pluripotent stem cells. *Cell Stem Cell*, 2010, 7: 249-57
- [15] Jiang C, Pugh BF. Nucleosome positioning and gene regulation: advances through genomics. *Nat Rev Genet*, 2009, 10: 161-72
- [16] Orkin SH, Hochedlinger K. Chromatin connections to pluripotency and cellular reprogramming. *Cell*, 2011, 145: 835-50
- [17] Tao Y, Zheng W, Jiang Y, et al. Nucleosome organizations in induced pluripotent stem cells reprogrammed from somatic cells belonging to three different germ layers. *BMC Biol*, 2014, 12: 109
- [18] Polo JM, Anderssen E, Walsh RM, et al. A molecular roadmap of reprogramming somatic cells into iPS cells. *Cell*, 2012, 151: 1617-32
- [19] Whyte WA, Orlando DA, Hnisz D, et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell*, 2013, 153: 307-19
- [20] Chen J, Chen X, Li M, et al. Hierarchical Oct4 binding in concert with primed epigenetic rearrangements during somatic cell reprogramming. *Cell Rep*, 2016, 14: 1540-54
- [21] Chronis C, Fiziev P, Papp B, et al. Cooperative binding of transcription factors orchestrates reprogramming. *Cell*,

- 2017, 168: 442-59, e420
- [22] West JA, Cook A, Alver BH, et al. Nucleosomal occupancy changes locally over key regulatory regions during cell differentiation and reprogramming. *Nat Commun*, 2014, 5: 4719
- [23] Pankratz MT, Li XJ, Lavaute TM, et al. Directed neural differentiation of human embryonic stem cells via an obligated primitive anterior stage. *Stem Cells*, 2007, 25: 1511-20
- [24] Du Y, Liu Z, Cao X, et al. Nucleosome eviction along with H3K9ac deposition enhances Sox2 binding during human neuroectodermal commitment. *Cell Death Differ*, 2017, 24: 1121-31
- [25] Gao S, Zheng C, Chang G, et al. Unique features of mutations revealed by sequentially reprogrammed induced pluripotent stem cells. *Nat Commun*, 2015, 6: 6318