

诱导重编程过程中的表观遗传重塑

陈 珺 高亚威 陈嘉瑜 高绍荣*

(同济大学生命科学与技术学院, 上海 200092)

摘要 多能干细胞(pluripotent stem cell, PSC)是一类具有自我更新能力和多向分化潜能的细胞, 具有广泛的临床应用前景, 诱导性多功能干细胞(induced pluripotent stem cell, iPS cell)的获得, 解决了传统方式中的细胞来源和伦理学等问题, 从理论研究和应用上实现了体细胞重编程的重大突破, 也为疾病发生机制研究、药物筛选、个性化药物选择、细胞治疗和再生医学等研究创造了难得的机会, 从而开启了多能干细胞应用的新纪元. iPS 过程中有很多问题尚未得到解决, 尤其是诱导重编程的分子机制方面, 这也是近年来干细胞领域研究的热点, 其中如何实现表观遗传的重编程被认为是亟待解决的核心问题之一. 本文结合我们的研究, 主要介绍诱导重编程领域表观遗传修饰重塑机制的研究进展, 并展望未来研究中大规模信息整合分析的重要性.

关键词 iPS 诱导, 组蛋白修饰, DNA 修饰, 表观遗传重编程
学科分类号 Q71

DOI: 10.16476/j.pibb.2015.0276

体细胞重编程(somatic cell reprogramming)是指采用特定的方法将已经分化的体细胞逆转回到多能性或全能性状态的过程. 无论是先驱性的体细胞核移植、细胞融合技术, 还是近些年新兴的 iPS 诱导技术, 都强有力地证实了已分化细胞状态的可塑性, 并颠覆了过去“分化过程不可逆”的认知. 2012 年, 约翰·戈登和山中伸弥两位科学家由于他们对体细胞重编程领域卓越的贡献而共享了诺贝尔生理医学奖. 其中, 戈登在 1962 年将蛙的小肠上皮细胞核移入去核卵中, 得到了完全健康的蝌蚪, 首次实现了分化细胞的重编程, 开创了克隆动物的先河. 但是, 是否完整的分化细胞, 而非细胞核, 也具有逆分化成为多能细胞的能力依然是未知的. 2006 年, 日本科学家山中伸弥采用了一种简单到不可思议的实验手段实现了体细胞的重编程, 完美地回答了这个问题. 他发现, 只需要在小鼠的分化细胞中过表达一组转录因子 OSKM(Oct4、Sox2、Klf4、c-Myc)就足以实现多能性的重新获得, 并得到一种与胚胎多能干细胞极为类似的细胞, 他将其命名为诱导多能干细胞(induced pluripotent stem cells, iPSCs). iPS 技术, 由于其操作简单、重复度高, 又规避了核移植技术中胚胎获得的伦理性问

题, 在全世界范围内获得了广泛的应用, 并推动了重编程领域在过去近 10 年中的爆炸式发展.

研究体细胞重编程的机制具有十分重大的意义. 一方面, iPS 细胞在临床医学、疾病机制以及药物研发领域都有非常广阔的应用前景, 如何高效地获得高质量、安全的 iPS 细胞正是科学研究的重中之重. 研究重编程的机制则有利于解决这些问题. 另一方面, 可以借用体细胞重编程手段来研究细胞命运决定机制, 帮助我们了解生物体发育程序的奥秘.

对于转录因子诱导重编程机制的研究存在多个方向. 一些研究从影响重编程的因子入手, 通过 RNAi 和小分子的筛选等方法, 寻找替代转录因子或辅助重编程发生的基因与小分子. 也有一些研究从 ES 细胞的核心调控机制以及体细胞与 ES 细胞的简单比较入手, 预测可能对重编程产生影响的多种因子. 在较早的研究时期, 我们实验室发现单独表达 Oct4 可以实现滋养层干细胞的重编程^[1]. 裴端

* 通讯联系人.

Tel: 021-65988653, E-mail: gaoshaorong@tongji.edu.cn

收稿日期: 2015-09-06, 接受日期: 2015-09-23

卿研究组和邓宏魁研究组随后分别建立了通过单独的 Oct4 过表达和小分子掺入等特殊的培养条件, 实现人和鼠体细胞重编程的方法^[2-3]. 邓宏魁研究组在已建立的 Oct4 单因子诱导重编程的基础上, 进一步开发出了完全利用小分子诱导实现成纤维细胞重编程为 iPS 细胞的方法. 这一方法基于原有的转录因子诱导体系, 同样属于诱导重编程的因子替代方面的研究. 尤其目前的证据依然认为化学诱导重编程的过程, 与传统诱导相比具有相当程度的相似性, 如诱导时间长(16~21 天)、效率很低(0.2%)、依赖于激活多能性因子内源表达等^[4]. 小分子诱导重编程的出现为 iPS 细胞在临床领域的应用提供了更为安全的材料.

另一些有关诱导重编程的研究则是利用遗传和分子生物学以及大规模测序等技术, 分析重编程过程的分子机制. 其中, 解析诱导重编程过程中的各种变化, 则是机制研究首先面临的问题. 相对于体内发生的重编程事件, 体细胞在体外被诱导产生 iPS 细胞, 是相对长时间且低效率的过程. 研究发现, 转录因子的种类、表达水平、诱导的初始细胞以及诱导条件等, 都将对 iPS 细胞的诱导效率和诱导时间产生明显的影响. 为避免病毒感染和插入随机性带来的影响, 强力霉素(doxycycline, Dox)控制的表达载体用于 iPS 细胞诱导中. 以此为基础, 研究人员可以通过第一轮重编程或基因定点打靶技术, 将重编程因子整合在 rtTA 转基因小鼠的基因组中, 通过四倍体囊胚互补技术或转基因动物技术, 获得背景单一的可诱导体细胞^[5-8]. 二次重编程系统可以实现多种不同来源初始细胞的重编程诱导. 更为重要的是, 二次诱导重编程过程中没有新的病毒插入, 并保持了相对均一的细胞群体, 因此具有相对稳定、可重复的重编程效率和动力学, 被认为是研究转录因子介导的体细胞重编程机制的重要平台^[9].

随着技术的进步和二次诱导平台的建立, 研究人员分析了 iPS 细胞产生的进程, 包括细胞形态、细胞内基因表达谱改变、细胞的表观遗传变化等. 人们发现体细胞的诱导重编程, 将经历一系列被严格调控的中间状态^[9-11]. 其中裴端卿研究组与国外研究人员一起验证了成纤维细胞到 iPS 细胞重编程过程中的间充质向上皮转化的转变过程(mesenchymal-to-epithelial transition, MET), 这一过程对于细胞的形态变化、细胞增殖速度改变、SSEA1 蛋白表达以及克隆样的细胞聚集都具有重

要作用^[10, 12-14]. 更多的研究证明, 整个重编程过程中, 存在一系列可供标识的细胞内信号, 如体细胞特异的 Thy1 的表达下调、SSEA1 以及 Oct4 和 Nanog 的相继表达等^[5].

虽然诱导体系各有不同, 诱导的具体进程也有所差异, 无法改变的是诱导过程的核心和本质是实现体细胞的表观遗传重编程. 随着机制研究的深入, 对于表观遗传重编程机制的解答, 被认为是诱导重编程研究中的核心问题. Jaenisch 研究组通过在二次诱导平台下的 pre-B 细胞单细胞诱导, 发现重编程的低效率并不是由于重编程因子转入的差别引发的^[15]. 他们的发现揭示了转录因子诱导的重编程过程中的随机性, 同时暗示, 体细胞特异的表观遗传修饰, 阻碍和限制了重编程的发生. Hanna 研究组最近的研究也证明, 解除诱导重编程中 Mbd3 介导的“刹车”效应, 将极大地提高重编程的效率, 甚至改变诱导重编程的动态模式^[16].

表观遗传变化是指 DNA 序列不发生变化的情况下基因表达发生的可遗传改变. 对于从受精卵到个体的发育过程来说, 体细胞的遗传物质是相同的, 但体细胞的各种特性千差万别, 所以有观点认为表观遗传学可以涵盖从遗传物质开始到个体成型的所有调控过程^[17]. 在分子水平上, 染色质上的表观遗传修饰主要包括: DNA 的甲基化、组蛋白修饰、核小体重构和染色质高级结构. 这些表观遗传修饰独有地存在于各个细胞系中, 并且相互关联, 与细胞内的转录因子等共同维持细胞内基因的转录、应答、分化、去分化等一系列的变化. 从本质上来说, 无论是发育过程中从全能受精卵开始分化成为特化的组织细胞, 还是体细胞被重编程为多能干细胞都是表观遗传上的变化. 体细胞重编程的核心正是表观遗传的重塑. 诱导重编程中的表观遗传修饰重塑, 主要包括两个重大事件: 打破原有体细胞的表观遗传修饰稳态和建立 ES 细胞特异的表观遗传修饰稳态. 在多项研究中, 发现很多组蛋白修饰相关和 DNA 修饰相关的基因能够对 iPS 诱导产生影响.

中国在干细胞以及重编程领域资金投入巨大, 科技部 973、863 计划以及国家自然科学基金、国家发改委产业化计划等都从不同的层面上给予干细胞研究多个重大重点项目的资金支持, 也取得了丰硕的成果. 下面, 将从组蛋白修饰与染色体高级构象、DNA 修饰以及 X 染色体激活等方面, 结合我们实验室取得的研究成果, 介绍体细胞诱导重编程

中表观遗传机制的研究进展。

1 诱导重编程中的组蛋白变化

核小体是染色质的基本组成单位, 由 4 种组蛋白 H2A、H2B、H3、H4 形成的八聚体和缠绕在上面的 DNA 组成。核心组蛋白的尾端上存在着甲基化、乙酰化、磷酸化、泛素化、类泛素化、ADP 核糖基化等多种翻译后修饰的位点, 这些修饰可以通过影响 DNA/核小体结合构象或者调节蛋白质识别来参与基因转录等过程, 被称为“组蛋白编码”(histone code)^[18](表 1)。对于重编程过程中组蛋白修饰动态变化的研究, 科学家们主要采用两种不同的策略: a. 过表达或敲低组蛋白修饰酶, 观察其对重编程的影响; b. 利用高通量测序手段, 对不同组蛋白修饰进行染色质免疫共沉淀(ChIP-seq), 分析全基因组水平上表观修饰的分布和变化情况。

基因的启动子区主要存在激活性的 H3K4me3 和抑制性的 H3K27me3 这两种组蛋白甲基化修饰。2011 年, Koche 等^[19]利用 CFSE 染色方法, 于分离二次 iPS 诱导的早期(72 h 内), 分离了不同分裂次数的细胞, 并使用高通量测序手段, 综合分析了前 3 次细胞周期中组蛋白 H3K4 和 H3K27 甲基化、DNA 甲基化以及转录组的变化。他们发现, 在重编程早期, 基因组上广泛出现了 H3K4me2 的获得, 并认为这些修饰的获得与后续过程中 ES 细胞活化区域 H3K4me3 的获得存在一定关系。这些修饰往往集中于体细胞中外源转录因子可以结合且未被 DNA 甲基化高度修饰的区域, 而 H3K4me2 在早期的获得并不能够直接改变基因的转录水平。相反, H3K27me3 在重编程初期并未检测到显著的变化, 说明这种抑制性修饰可能在更晚的时候才发生变化^[19]。2012 年, Konrad 研究组在基于流式分选细胞的重编程平台中, 观察到了与 mRNA 变化一致的两波 H3K4me3/H3K27me3 变化, 分别发生在重编程的早期和晚期。但是, 对胚胎干细胞中特有的既具有 H3K4me3 又具有 H3K27me3 信号的双价启动子(bivalent promoter)的分析显示, 这类基因的数目在重编程过程中缓慢持续上升, 说明双价性(bivalency)是逐步建立的^[20]。在重编程过程中, 与 H3K4me3 信号的丢失和 H3K27me3 信号的同时或随后增强一致发生的是体细胞特异基因的沉默。与之形成对比, 重编程过程中多能基因上 H3K4me3 信号的获得是不同步的, 这暗示了重编程过程中内源的多能基因激活是一个随机的过程^[20]。另一个有

意思的发现是, H3K27 甲基转移酶 PRC2 复合物与 H3K27me3/2 特异性的去甲基酶 Utx1 敲低都会降低 iPS 诱导效率, 这说明 H3K27me3 的建立与去除对重编程都十分关键^[21-22]。

在基因的增强子区域, 主要富集着 H3K4me1 以及激活性的 H3K27ac 信号。Koche 等^[19]的研究发现, 与基因启动子区相比, 增强子区域在重编程初期的变化更为显著, 这之前观察到的不同种类细胞中增强子的差异性高于启动子的现象是一致的^[23]。研究表明, 在细胞分化过程中, 增强子的建立和消除与 H3K4 甲基转移酶 MLL4 和 H3K4 去甲基酶 LSD1 的功能有关^[24-25]。与分化过程相类似, 在重编程过程中, 不同基因的增强子也在沉默(inactive)、停滞(poised)和激活(active)等不同的状态中转换, 因此, 可以推断 MLL4 与 LSD1 也可能在这个过程中起着重要作用。组蛋白乙酰化可以中和赖氨酸上的正电荷, 使染色质结构松散, 因此通常被认为是激活性的组蛋白标志。H3K27ac 是基因增强子活化的独特标志。我们实验室未发表的数据也表明, 在重编程过程中多能基因的增强子上 H3K27ac 的建立与细胞多能性的逐渐获得是高度一致的。此外, 在诱导培养基内加入组蛋白去乙酰化酶抑制剂丙戊酸(valproic acid, VPA)或者丁酸盐(butyrate)也都能够提高重编程效率^[26-30]。

组蛋白修饰 H3K9me3 是异染色质的标志, 其在 ES 细胞中 H3K9me3 标记区域仅占~4%, 而在体细胞中可达~12%^[31]。研究表明, H3K9me3 是完全重编程的阻碍之一^[32-33]。2013 年, 我国裴端卿研究组发表于 *Nature Genetics* 的工作发现, 在未完全重编程的 pre-iPS 细胞中, 尽管 Oct4 的表达水平很高, 但是却不能结合到许多胚胎干细胞特异的结合位点上, 实现多能性网络的建立。进一步的研究中他们发现 H3K9me3 这一抑制性的表观遗传修饰可能是 Oct4 在这些位点上结合的阻碍。他们使用 siRNA 降低 H3K9 甲基转移酶 Setdb1 的表达, 结果可以高效地将 pre-iPS 细胞转变为 iPS 细胞, 这证明 H3K9me3 的确是重编程晚期的重要阻碍^[32]。

此外, 裴端卿研究组也发现, 能够极大地提高重编程效率的小分子维生素 C 可以与 H3K36me2/3 去甲基酶 Jhdmla/1b 协同作用, 调控 H3K36 甲基化水平, 促进细胞增殖, 抑制细胞衰老, 从而促进体细胞重编程^[34-35]。

其他已知的在重编程过程中十分重要的组蛋白修饰还包括 H3K79 甲基化和 H2AK119 泛素化

等^[21, 36]. 抑制 H3K79 甲基转移酶 Dot1l 的表达量, 或者加入小分子 EPZ004777 来抑制 Dot1l 酶活性, 都可以加速重编程. 这一功能被证明主要是通过促进重编程早期细胞的间充质向上皮转化过程(MET)来实现^[21]. 而 PRC1 复合物组分 Bmi1 被证明可以取代 SKM, 仅与 Oct4 一起实现成纤维细胞的 iPS 诱导^[36].

除了常规的 H2A、H2B、H3、H4 组蛋白外, 细胞内还存在着多种组蛋白变体, 这些变体嵌入核小体形成了结构和功能各异的核小体变体, 在基因组的特定位置或者在特定时期结合. 越来越多的研究表明, 组蛋白变体在体细胞重编程中也起到了非常重要的作用^[37]. 日本理化所研究者在利用哺乳动物的卵母细胞寻找参与诱导全能胚胎干细胞的分子的实验中发现, H2A/H2B 的组蛋白突变体 TH2A/TH2B 可以将 iPSC 细胞的诱导效率提高将近 20 倍. 在重编程过程中 TH2A/TH2B 主要富集在 X 染色体上, 它们的过表达能够加大染色质的开放程度, 与重编程因子协同作用, 促进体细胞重编程^[38]. 另一个 H2A 变体 macroH2A 则被认为是重编程过程中的主要屏障之一^[39]. 近期, 两个不同研究组的独立工作都表明, 在重编程过程中 H2A.X 的重定位对于体细胞的完全重编程十分重要, H2A.X 的定位信息可以帮助我们判断 iPSC 的质量^[40-42]. 在过去核移植、胚胎干细胞以及生殖干细胞的研究中已经证实, 多种组蛋白变体 H2A.Z、H3.3、H3.2、CenH3、macroH2A 以及它们对应的组蛋白伴侣等都对于干性维持、基因调控有着复杂的影响^[37], 因此, 我们有理由相信, 组蛋白变体在体细胞重编程过程中的差异性分布和动态变化对于多能性的获得是十分重要的.

2 诱导重编程过程中染色体高级结构的变化

多能干细胞具有非常特殊而松散的染色质高级结构. 与已分化细胞相比, ES 细胞中的异染色体集落(heterochromatin foci)更大更分散, 染色质结构蛋白(chromatin architectural proteins)与 DNA 的结合更松散而高度动态, 同时全基因组的转录都更为活跃^[31, 43]. 因此, 重编程过程中, 体细胞必须要发生染色质高级结构的变化. 此外, 研究发现, 染色质折叠互作可以形成大的调控功能域, 称为拓扑结构域(topological domains), 可以实现远端调控子对基因转录活性的调控^[44-45], 这些发现都暗示了重编

程过程中染色体高级结构的变化可能直接驱动体细胞表观组和转录组的转变, 参与诱导多能性的建立. 近年来, 基于染色质构象捕获(chromosome conformation capture, 3C)的长距离染色体互作检测技术得到了长足的发展, 在重编程机制研究领域也得到了应用. 4C-seq (circular chromosome conformation capture sequencing)是一种基因组水平的 3C 衍生技术, 它主要检测某一选定位点与基因组上其他位置之前相互接触的频率, 经过复杂的生物信息学分析, 能够给予我们细胞内 DNA 空间构成的信息. 利用 4C-seq 和转录因子 ChIP-seq 数据, de Wit 等^[46]证实了多能干细胞中基因组围绕多能因子 Oct4 和 Nanog 形成了特殊的细胞核 3D 结构, 具有多能因子结合位点的增强子通过多能因子高效地结合在一起, 以促进多能性的维持. 2013 年, Lu 研究组和 Konrad 研究组分别在 *Cell Stem Cell* 上发表文章, 他们利用 4C 测序和荧光原位杂交技术检测重编程过程中 Oct4 和 Nanog 位点的全基因组染色质互作图谱, 实验的结果都证明了多能位点的顺式互作网络的重置是多能基因激活的先决条件^[47-48]. 其中, Lu 研究组的文章揭示了重编程因子 Klf4 在 Oct4 位点的高级染色质结构建立过程中的重要作用^[48], 而 Konrad 研究组的文章则强调了 Mediator 和 cohesin 蛋白在染色质成环及 3D 结构形成中的功能^[47].

3 诱导重编程中的 DNA 修饰变化

不仅是组蛋白修饰, DNA 的修饰变化也被认为是表观遗传重编程中的重要事件(表 1). 传统研究的 DNA 修饰主要是甲基化修饰. DNA 的甲基化由 DNA 甲基转移酶(DNA methyltransferase, Dnmt)催化产生, 主要产物为 5-甲基胞嘧啶(5mC)^[17]. 哺乳动物的 DNA 甲基化主要发生在 CpG 二核苷酸. 作为一种重要的表观遗传修饰, DNA 甲基化可以导致长效性基因沉默、广泛参与基因表达的调控、组蛋白修饰的建立等过程, 并对基因组稳定性维持具有重要作用^[17, 49-51].

在体细胞中, 很多多能性相关的基因和调控区域都被 H3K9me3 和 DNA 的甲基化覆盖, 以维持基因沉默的稳定性, 而传统研究认为体细胞特异基因的沉默也可能由甲基化导致^[9]. 前面提到的组蛋白修饰变化和转录因子结合的研究也都表明, DNA 甲基化和 H3K9me3 是重编程过程中必须打破的壁垒. Jaenisch 研究组在 pre-iPS 为平台的研究中

发现, DNA 去甲基化参与了核心转录因子 Oct4、Nanog、Utf1 等基因在诱导后期的激活^[12].

然而直到最近几年, DNA 去甲基化的可能机制才开始随着 Tet 蛋白家族和 DNA 羟甲基化研究的发展而逐渐清晰起来. Tet 蛋白家族包括 Tet1、2、3 三个成员, 可以将甲基化胞嘧啶(5mC)氧化为羟甲基化胞嘧啶(5hmC)的形式, 从而进一步通过继续氧化为 5caC、5fC 并为 TDG 识别切除, 或者被 AID 加氨, 生成 5hmU 并被 SMUG1 切除. 切除后的 DNA 最终经由 DNA 结合修复(BER)信号通路, 被修补为无修饰的胞嘧啶实现主动去甲基化. 除了主动去甲基化, 5hmC 也可以通过复制稀释这一类被动的方式实现其修饰的去除^[52-53]. 在这一研究领域中, 华人科学家以及我国本土的研究团队发挥了不可替代的重要作用. 华人科学家张毅与中国科学院徐国良研究组, 分别独立证明了 5hmC 继续氧化产物的存在及其主动去除的机制, 并在探讨 5hmC 在胚胎干细胞与早期发育受精过程中的作用方面取得了重要研究成果^[52-54].

同时多个研究团队发现, 5hmC 在某些细胞中具有相对稳定地存在, 并可能参与了转录的调控, 因此 5hmC 也被认为是表观修饰的一种. 目前的研究认为, Tet 家族蛋白催化 5mC 到 5hmC 的转化, 参与了体内受精过程和 PGC 发育等过程的去甲基化^[54-56]. 不仅在体内, 对体外的多种重编程过程, 包括体细胞核移植、细胞融合以及现在讨论的诱导重编程中, Tet 蛋白介导的 5hmC, 对于 DNA 甲基化的去除、多能性基因的激活和组蛋白修饰的重建都十分重要^[52, 57-61].

哥伦比亚大学的 Doege 等^[61]首先报道了 Tet2 的缺失会影响重编程过程中 Nanog 和 Esrrb 基因上 5hmC 的获得. Piccolo 等^[59]利用细胞融合介导的重编程平台, 证明 Tet1 和 Tet2 分别介导了重编程过程中发生在印记基因和多能性基因上的去甲基化. José 与王建龙研究组合作, 发现 Tet1 可以与 Nanog 相互作用. 他们首先在神经干细胞中过表达 OKM 获得重编程的中间状态, 进一步发现具备氧化活性的 Tet1 可以辅助外源 Nanog 更为高效地将这种中间状态的细胞转变为 iPS 细胞^[57]. 这表明 Tet1 对于 iPS 诱导过程中多能性网络的建立可能十分重要^[62]. 我们实验室^[58]在同一时期发表的工作证明, Tet1 催化的 DNA 羟甲基化反应是推动重编程发生的重要动力, 可以直接参与重编程过程中 Oct4 转录调控区域的去甲基化和转录激活. Tet1

更可以取代 Oct4 实现诱导重编程, 在这个 Tet1 介导的重编程体系中, 5hmC 作为一个重要的表观遗传修饰, 参与了细胞多能性的获取和整体表观体系的重建. 裴端卿研究组与徐国良研究组合作, 发现维生素 C 可以提高 Tet1 的催化活性, 并影响体细胞重编程过程. 在较高浓度维生素 C 的情况下, Tet1 会显著抑制体细胞重编程过程早期的 MET 转化过程, 从而阻碍体细胞重编程; 而在维生素 C 浓度低于生理浓度 2 个数量级的情况下, Tet1 的抑制作用被隐藏起来, 转而促进重编程^[60]. 徐国良研究组在另一项研究中发现, 在缺少 Tet 蛋白的情况下, 以成纤维细胞为起点的诱导重编程中期的 MET 过程将不能发生, 而 Tet 和 TDG 在 MET 过程中的关键作用在于调控重编程早期 microRNA 的去甲基化和表达^[60]. 这一结果支持了我们的观点: DNA 羟甲基化和它所介导的 DNA 去甲基化应该不仅发生于重编程晚期, 而其影响也并不局限于多能性相关区域的 DNA 去甲基化.

4 诱导重编程对于 X 染色体活性状态的影响

在人和鼠等胎盘哺乳动物中, 雄性染色体为 XY, 雌性为 XX. 在雌性动物的体细胞内, 2 条 X 染色体只有 1 条染色体具有活性, 这样使得雌雄动物之间虽然 X 染色体的数量不一样, 但 X 染色体上的基因产物的剂量是平衡的, 这个现象被称为剂量补偿效应. 研究认为, 在小鼠、人、兔、牛等哺乳动物的内细胞团细胞形成早期, 2 条 X 染色体都处于活化状态, 而在随后的分化过程中, X 染色体失活(X chromosome inactivation, XCI)机制被启动. 在这一过程中, 非编码 RNA Xist 首先表达, 然后结合于 X 染色体上, 引发 X 染色体相关基因的表达沉默, 并诱导抑制性组蛋白标记的建立. 这也使得探讨诱导重编程对于 X 染色体状态的影响, 成为诱导重编程机制探讨中不可回避的重要问题^[11].

对于小鼠而言, 胚胎干细胞和诱导多能性干细胞中都具有 2 条具有活性的 X 染色体(Xa, Xa), 在分化过程中, 随着 Xist 表达发生 X 染色体的随机失活(Xi, Xa). Kathrin 研究组发现 iPS 细胞内同样含有 2 条活化的 X 染色体, 因此在诱导重编程中, 失去活性的 X 染色体需要重新激活, 这一过程与 X 染色体上抑制性染色体标记(主要是 H3K27me3)去除以及 Xist 基因沉默伴随发生^[64]. 小鼠外胚层干细胞(epiblast stem cells, EpiSCs)转化为

胚胎干细胞的过程中, 已经发生失活的 X 染色体也需要重新激活^[65]. 因此 X 染色体的活性状态被认为是小鼠体系中细胞发育阶段状态的重要指标^[61]. 在我们实验室最新完成的工作中, 发现 Xist 基因的表达在诱导重编程的早期是不可缺少的, 而在 pre-iPS 到最终建立全能性的阶段, Xist 表达被沉默, X 染色体被重新激活^[66].

对于人的 iPS 细胞和胚胎干细胞的研究发现, X 染色体的活性状态比较复杂, 存在三类不同的 X 染色体状态. 其中两类(I 类和 II 类)分别类似于鼠 ES 细胞^[66]和 EpiSC 细胞, 但是这两类细胞在经过长期培养后, 容易出现 Xist 表达缺陷, 从而引发 XCI 受损状态的出现($Xi^{w/oXIST}Xa$, $XeXa$), 这类细胞被证明了存在体内分化能力受损并且癌基因表达增高等缺陷, 因而不适宜临床应用^[67-69]. 更重要的

是, 目前受损的 XCI, 无法通过更改培养条件或加入小分子等方法实现逆转. 因此在对雌性个体, 尤其是女性病人体的诱导重编程和 iPS 细胞维持中, X 染色体的活性状态非常重要. 我们实验室刚刚发表的研究证明, X 染色体状态的不稳定, 保守地存在于较早代数的兔胚胎干细胞, 并与人的胚胎干细胞相似, 在传代中出现 XCI 受损. 更为重要的一点在于, 我们发现, 通过 OSKM 因子介导的重编程以及改善培养条件, 可以修复原本 Xist 基因表达调控的缺陷, 获得双活性状态的 X 染色体, 并修复原本受损的体内分化能力, 以及上调癌基因. 而且受损的 XCI 也可以通过这一方式获得修复, 这为解决女性 iPS 细胞中 X 染色体的不稳定性, 提供了非常宝贵的参考^[70].

Table 1 Key epigenetic modifications involved in somatic cell reprogramming

表 1 重编程过程中几种重要的表观遗传修饰

表观遗传修饰	参与重编程的调控基因 *	相关的小分子	对重编程的影响	参考文献
H3K4me3	TrxG		动态变化对重编程十分重要, 与多能基因激活密切相关	[12, 20]
H3K4me1	MLL3/4; LSD1	Parnate	甲基化促进重编程	[19]
H3K27me3	PRC2; UTX1		动态变化对重编程十分重要, 与谱系特异性基因沉默密切相关	[12, 20-22]
H3K9me3	SETDB1、G9A	去甲基化酶抑制剂 BIX-01294	去甲基化促进重编程	[71-72]
H3K27ac	P300、CBP; HDAC	去乙酰化酶抑制剂 VPA、SAHA、TSA、Butyrate	乙酰化促进重编程	[26-29]
H3K79me2/3	DOT1L	DOT1L 的抑制剂 EPZ004777	去甲基化促进重编程	[21]
TH2A、TH2B macroH2A H2A.X			促进重编程 抑制重编程 重定位对 iPSC 细胞质量很重要	[38] [39] [41-42]
DNA 甲基化	DNMT3A、DNMT3B、DNMT1	DNA 甲基转移酶抑制剂 RG108; 5-AZA	去甲基化促进重编程	[12, 27, 71]
DNA 羟甲基化	TET1、TET2、TET3		羟甲基化及去甲基化促进重编程	[57-58]

* 灰色表格内为催化产生对应修饰的蛋白, 白色表格内为催化去除该修饰的蛋白.

5 小结与展望

诱导重编程的出现, 不仅为干细胞研究以及再生医学等领域提供了良好的发展契机, 也启发和推动了体细胞直接转分化方面的探索研究. 目前, 包括神经干细胞、神经细胞、肝细胞、心肌细胞等一系列细胞, 都被发现可以通过 OSKM 诱导或者其他转录因子, 乃至小分子诱导的方法, 从已分化的成体细胞转化得到. 但是对于转分化方面的机制研究, 尤其是核心的表观遗传修饰重塑机制的研究, 目前还基本处于空白. 而在这方面, 我们希望能借鉴诱导重编程机制研究中的经验与方法, 通过机制研究, 实现对转分化方法的改善和新类型转分化的方法设计.

随着对于诱导重编程过程了解的加深, 以及对表观遗传修饰重塑机制研究的不断深入, 越发意识到传统的实验方法和研究思路并不能很好地解答未知问题. 而要想更为深刻地了解诱导重编程甚至细胞转分化等过程的分子机制和解决这些过程中出现的复杂问题, 必须使用更为系统的研究手段进行更为广泛的信息收集, 并采用系统生物学的观点和研究方法思考问题, 构建模型, 通过严格对照试验进行论证. 而新技术的研发, 尤其是微流控技术以及转录组和表观修饰组的微量测定技术等, 势必将在该领域中发挥越来越重要的作用, 并将推动相关技术在临床方面的应用.

参 考 文 献

- [1] Wu T, Wang H, He J, *et al.* Reprogramming of trophoblast stem cells into pluripotent stem cells by Oct4. *Stem cells*, 2011, **29**(5): 755–763
- [2] Li Y, Zhang Q, Yin X, *et al.* Generation of iPSCs from mouse fibroblasts with a single gene, Oct4, and small molecules. *Cell Res*, 2011, **21**(1): 196–204
- [3] Zhu S, Li W, Zhou H, *et al.* Reprogramming of human primary somatic cells by OCT4 and chemical compounds. *Cell Stem Cell*, 2010, **7**(6): 651–655
- [4] Hou P, Li Y, Zhang X, *et al.* Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds. *Science*, 2013, **341**(6146): 651–654
- [5] Brambrink T, Foreman R, Welstead G G, *et al.* Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. *Cell Stem Cell*, 2008, **2**(2): 151–159
- [6] Carey B W, Markoulaki S, Hanna J H, *et al.* Reprogramming factor stoichiometry influences the epigenetic state and biological properties of induced pluripotent stem cells. *Cell Stem Cell*, 2011, **9**(6): 588–598
- [7] Stadtfeld M, Apostolou E, Akutsu H, *et al.* Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. *Nature*, 2010, **465**(7295): 175–181
- [8] Kang L, Wang J, Zhang Y, *et al.* iPSCs can support full-term development of tetraploid blastocyst-complemented embryos. *Cell Stem Cell*, 2009, **5**(2): 135–138
- [9] Plath K, Lowry W E. Progress in understanding reprogramming to the induced pluripotent state. *Nat Rev Genet*, 2011, **12**(4): 253–265
- [10] Samavarchi-Tehrani P, Golipour A, David L, *et al.* Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. *Cell Stem Cell*, 2010, **7**(1): 64–77
- [11] Papp B, Plath K. Epigenetics of reprogramming to induced pluripotency. *Cell*, 2013, **152**(6): 1324–1343
- [12] Mikkelsen T S, Hanna J, Zhang X, *et al.* Dissecting direct reprogramming through integrative genomic analysis. *Nature*, 2008, **454**(7200): 49–55
- [13] Li R, Liang J, Ni S, *et al.* A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. *Cell Stem Cell*, 2010, **7**(1): 51–63
- [14] Smith Z D, Nachman I, Regev A, *et al.* Dynamic single-cell imaging of direct reprogramming reveals an early specifying event. *Nat Biotechnol*, 2010, **28**(5): 521–526
- [15] Hanna J, Saha K, Pando B, *et al.* Direct cell reprogramming is a stochastic process amenable to acceleration. *Nature*, 2009, **462**(7273): 595–601
- [16] Rais Y, Zviran A, Geula S, *et al.* Deterministic direct reprogramming of somatic cells to pluripotency. *Nature*, 2013, **502**(7469): 65–70
- [17] Allis C D, Jenuwein T, Reinberg D. *Epigenetics* [M]. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press, 2007
- [18] Strahl B D, Allis C D. The language of covalent histone modifications. *Nature*, 2000, **403**(6765): 41–45
- [19] Koche R, Smith Z, Adli M, *et al.* Reprogramming factor expression initiates widespread targeted chromatin remodeling. *Cell Stem Cell*, 2011, **8**(1): 96–105
- [20] Polo J, Anderssen E, Walsh R, *et al.* A molecular roadmap of reprogramming somatic cells into iPSCs. *Cell*, 2012, **151**(7): 1617–1632
- [21] Onder T T, Kara N, Cherry A, *et al.* Chromatin-modifying enzymes as modulators of reprogramming. *Nature*, 2012, **483**(7391): 598–602
- [22] Mansour A A, Gafni O, Weinberger L, *et al.* The H3K27 demethylase Utx regulates somatic and germ cell epigenetic reprogramming. *Nature*, 2012, **488**(7411): 409–413
- [23] Heintzman N D, Hon G C, Hawkins R D, *et al.* Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature*, 2009, **459**(7243): 108–112
- [24] Lee J E, Wang C, Xu S, *et al.* H3K4 mono- and di-methyltransferase MLL4 is required for enhancer activation during cell differentiation. *eLife*, 2013, **2**: e01503
- [25] Whyte W A, Bilodeau S, Orlando D A, *et al.* Enhancer decommissioning by LSD1 during embryonic stem cell

- differentiation. *Nature*, 2012, **482**(7384): 221–225
- [26] Huangfu D, Osafune K, Maehr R, *et al.* Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. *Nat Biotechnol*, 2008, **26**(11): 1269–1275
- [27] Huangfu D, Maehr R, Guo W, *et al.* Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nat Biotechnol*, 2008, **26**(7): 795–797
- [28] Mali P, Chou B K, Yen J, *et al.* Butyrate greatly enhances derivation of human induced pluripotent stem cells by promoting epigenetic remodeling and the expression of pluripotency-associated genes. *Stem cells*, 2010, **28**(4): 713–720
- [29] Zhang Z H, Wu W S. Sodium butyrate promotes generation of human induced pluripotent stem cells through induction of the miR302/367 cluster. *Stem Cells And Development*, 2013, **22**(16): 2268–2277
- [30] Liang G Y, Taranova O, Xia K, *et al.* Butyrate promotes induced pluripotent stem cell generation. *J Biol Chem*, 2010, **285** (33): 25516–25521
- [31] Gaspar-Maia A, Alajem A, Meshorer E, *et al.* Open chromatin in pluripotency and reprogramming. *Nature Reviews Molecular Cell Biology*, 2011, **12**(1): 36–47
- [32] Chen J K, Liu H, Liu J, *et al.* H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs. *Nat Genet*, 2013, **45**(1): 34–42
- [33] Soufi A, Donahue G, Zaret K. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell*, 2012, **151**(5): 994–1004
- [34] Esteban M A, Wang T, Qin B, *et al.* Vitamin C enhances the generation of mouse and human induced pluripotent stem cells. *Cell Stem Cell*, 2010, **6**(1): 71–79
- [35] Wang T, Chen K, Zeng X, *et al.* The histone demethylases Jhdm1a/1b enhance somatic cell reprogramming in a vitamin-C-dependent manner. *Cell Stem Cell*, 2011, **9**(6): 575–587
- [36] Moon J H, Heo J S, Kim J S, *et al.* Reprogramming fibroblasts into induced pluripotent stem cells with Bmi1. *Cell Research*, 2011, **21**(9): 1305–1315
- [37] Skene P J, Henikoff S. Histone variants in pluripotency and disease. *Development*, 2013, **140**(12): 2513–2524
- [38] Shinagawa T, Takagi T, Tsukamoto D, *et al.* Histone variants enriched in oocytes enhance reprogramming to induced pluripotent stem cells. *Cell Stem Cell*, 2014, **14**(2): 217–227
- [39] Gaspar-Maia A, Qadeer Z A, Hasson D, *et al.* MacroH2A histone variants act as a barrier upon reprogramming towards pluripotency (vol 4, 1565, 2013). *Nature Communications*, 2013, **4**: 1565
- [40] Gao Y, Gao S. Quality control: H2A.X links to better iPSCs. *Cell Stem Cell*, 2014, **15**(3): 259–260
- [41] Buganim Y, Markoulaki S, Van Wietmarschen N, *et al.* The developmental potential of iPSCs is greatly influenced by reprogramming factor selection. *Cell Stem Cell*, 2014, **15**(3): 295–309
- [42] Wu T, Liu Y, Wen D, *et al.* Histone variant H2A.X deposition pattern serves as a functional epigenetic mark for distinguishing the developmental potentials of iPSCs. *Cell Stem Cell*, 2014, **15**(3): 281–294
- [43] Meshorer E, Yellajoshula D, George E, *et al.* Hyperdynamic plasticity in pluripotent embryonic of chromatin proteins stem cells. *Developmental Cell*, 2006, **10**(1): 105–116
- [44] Dixon J R, Selvaraj S, Yue F, *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, 2012, **485**(7398): 376–380
- [45] Chakalova L, Fraser P. Organization of transcription. *Cold Spring Harbor Perspectives In Biology*, 2010, **2**(9): a000729
- [46] De Wit E, Bouwman B A M, Zhu Y, *et al.* The pluripotent genome in three dimensions is shaped around pluripotency factors. *Nature*, 2013, **501**(7466): 227–231
- [47] Apostolou E, Ferrari F, Walsh R M, *et al.* Genome-wide chromatin interactions of the Nanog locus in pluripotency, differentiation, and reprogramming. *Cell Stem Cell*, 2013, **12**(6): 699–712
- [48] Wei Z, Gao F, Kim S, *et al.* Klf4 organizes long-range chromosomal interactions with the oct4 locus in reprogramming and pluripotency. *Cell Stem Cell*, 2013, **13**(1): 36–47
- [49] De Carvalho D D, You J S, Jones P A. DNA methylation and cellular reprogramming. *Trends Cell Biol*, 2010, **20**(10): 609–617
- [50] Cedar H, Bergman Y. Programming of DNA methylation patterns. *Annu Rev Biochem*, 2012, **81**: 97–117
- [51] Cedar H, Bergman Y. Linking DNA methylation and histone modification: patterns and paradigms. *Nat Rev Genet*, 2009, **10**(5): 295–304
- [52] Pastor W A, Aravind L, Rao A. TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. *Nat Rev Mol Cell Biol*, 2013, **14**(6): 341–356
- [53] Wu H, Zhang Y. Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation. *Genes Dev*, 2011, **25**(23): 2436–2452
- [54] Gu T P, Guo F, Yang H, *et al.* The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature*, 2011, **477**(7366): 606–610
- [55] Hackett J A, Sengupta R, Zyllicz J J, *et al.* Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. *Science*, 2013, **339**(6118): 448–452
- [56] Vincent J J, Huang Y, Chen P Y, *et al.* Stage-specific roles for tet1 and tet2 in DNA demethylation in primordial germ cells. *Cell Stem Cell*, 2013, **12**(4): 470–478
- [57] Costa Y, Ding J, Theunissen T W, *et al.* NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. *Nature*, 2013, **495**(7441): 370–374
- [58] Gao Y, Chen J, Li K, *et al.* Replacement of Oct4 by Tet1 during iPSC induction reveals an important role of DNA methylation and hydroxymethylation in reprogramming. *Cell Stem Cell*, 2013, **12**(4): 453–469
- [59] Piccolo F M, Bagci H, Brown K E, *et al.* Different roles for Tet1 and Tet2 proteins in reprogramming-mediated erasure of imprints induced by EGC fusion. *Mol Cell*, 2013, **49**(6): 1023–1033
- [60] Hu X, Zhang L, Mao S Q, *et al.* Tet and TDG mediate DNA demethylation essential for mesenchymal-to-epithelial transition in

- somatic cell reprogramming. *Cell Stem Cell*, 2014, **14**(4): 512–522
- [61] Doege C A, Inoue K, Yamashita T, *et al.* Early-stage epigenetic modification during somatic cell reprogramming by Parp1 and Tet2. *Nature*, 2012, **488**(7413): 652–655
- [62] Jackson S A, Sridharan R. The nexus of Tet1 and the pluripotency network. *Cell Stem Cell*, 2013, **12**(4): 387–388
- [63] Chen J, Guo L, Zhang L, *et al.* Vitamin C modulates TET1 function during somatic cell reprogramming. *Nat Genet*, 2013, **45** (12): 1504–1509
- [64] Maherali N, Sridharan R, Xie W, *et al.* Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. *Cell Stem Cell*, 2007, **1**(1): 55–70
- [65] Nichols J, Smith A. Naive and primed pluripotent states. *Cell Stem Cell*, 2009, **4**(6): 487–492
- [66] Chen Q, Gao S, He W, *et al.* Xist repression shows time-dependent effects on the reprogramming of female somatic cells to induced pluripotent stem cells. *Stem cells*, 2014, **32**(10): 2642–2656
- [67] Mekhoubad S, Bock C, De Boer A S, *et al.* Erosion of dosage compensation impacts human iPSC disease modeling. *Cell Stem Cell*, 2012, **10**(5): 595–609
- [68] Tomoda K, Takahashi K, Leung K, *et al.* Derivation conditions impact X-inactivation status in female human induced pluripotent stem cells. *Cell Stem Cell*, 2012, **11**(1): 91–99
- [69] Anguera M C, Sadreyev R, Zhang Z, *et al.* Molecular signatures of human induced pluripotent stem cells highlight sex differences and cancer genes. *Cell Stem Cell*, 2012, **11**(1): 75–90
- [70] Jiang Y, Kou Z, Wu T, *et al.* Xist deficiency and disorders of X-inactivation in rabbit embryonic stem cells can be rescued by transcription-factor-mediated conversion. *Stem Cells Dev*, 2014, **23**(19): 2283–2296
- [71] Shi Y, Despons C, Do J T, *et al.* Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. *Cell Stem Cell*, 2008, **3**(5): 568–574
- [72] Shi Y, Do J T, Despons C, *et al.* A combined chemical and genetic approach for the generation of induced pluripotent stem cells. *Cell Stem Cell*, 2008, **2**(6): 525–528

Remodeling of Epigenetic Landscape During Induced Somatic Cell Reprogramming

CHEN Jun, GAO Ya-Wei, CHEN Jia-Yu, GAO Shao-Rong*

(School of Life Science and Technology, Tongji University, Shanghai 200092, China)

Abstract Pluripotent stem cells are capable of differentiation into different cell types, and therefore have great potential in clinical applications. The achievement of transcription-factor-induced pluripotent stem cells (iPS cell) avoids ethical issues and offers an invaluable source of patient-specific pluripotent stem cells for disease modeling, drug screening, toxicology tests, personalized medicine selection, cell therapy and regenerative medicine. As a milestone, iPS technology initiates a new era of pluripotent stem cells. However, even until now many problems especially the key molecular mechanism of epigenetic reprogramming during this process, are little understood. Uncovering the molecular mechanism of this unique platform would shed light on improving the reprogramming efficiency and iPSC quality, ultimately advancing their therapeutic applications. In this paper we will briefly introduce the significant progress in this field and attach importance to combining high-throughput sequencing technique and systematic biology in future research.

Key words iPS induction, histone modification, DNA modification, epigenetic reprogramming

DOI: 10.16476/j.pibb.2015.0276

*Corresponding author.

Tel: 86-21-65988653, E-mail: gaoshaorong@tongji.edu.cn

Received: September 6, 2015 Accepted: September 23, 2015