

动物早期胚胎发育中染色质结构的继承和重编程

柯玉文^{1,2}, 刘江^{1,2}

1. 中国科学院北京基因组研究所, 中国科学院基因组科学与信息重点实验室, 北京 100101

2. 中国科学院大学, 北京 100049

摘要: 染色质开放性和染色质三维高级结构在基因表达和调控中发挥着非常重要的作用, 广泛参与分化、发育、肿瘤发生等细胞生理过程, 是表观遗传研究的热点领域之一。动物胚胎发育起始于终端分化的卵子受精形成全能性的受精卵。在精卵结合的过程中, 染色质开放性和染色质三维高级结构发生了剧烈的变化, 经历继承、重编程、重新建立的过程, 并指导调控受精卵分化发育最终成为多细胞、多器官组织的新生命个体。本文介绍了近年来研究染色质开放性和染色质三维高级结构的实验分析技术手段, 染色质结构在动物早期胚胎发育过程中的变化规律及其在早期胚胎发育中的作用, 染色质结构与其他表观遗传信息(甲基化、组蛋白修饰等)关系方面的重要研究进展和存在的科学问题, 以期表观遗传调控早期胚胎发育的研究提供参考。

关键词: 表观遗传; 胚胎发育; 染色质开放性; 染色质三维高级结构

The inheritance and reprogramming of chromatin structure in early animal embryos

Yuwen Ke^{1,2}, Jiang Liu^{1,2}

1. CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100101, China

2. University of Chinese Academy of Sciences, Beijing 100049, China

Abstract: The chromatin accessibility and 3D high-order structure play an important role in gene expression and regulation, and are involved in numerous cellular processes such as differentiation, development and tumorigenesis. It is currently one of the hottest research fields in epigenetics. The animal embryonic development initiates from the terminally differentiated mature eggs, which are fertilized by sperms to establish totipotent zygotes. During this complex process, the epigenome encompassing chromatin accessibility and high-order structures changes dramatically, undergoes inheritance, reprogramming and re-establishment, so that the zygote could eventually develop into a new multicellular and multi-tissue organism. In this review, we summarize the research methods and technology for chromatin accessibility and 3D high-order structures, the dynamics of chromatin structures during the animal embryonic development and how to

收稿日期: 2018-07-04; 修回日期: 2018-10-08

基金项目: 国家自然科学基金项目(编号: 31425015)资助[Supported by the National Natural Science Foundation of China (No. 31425015)]

作者简介: 柯玉文, 博士, 副研究员, 研究方向: 表观遗传。E-mail: keyw@big.ac.cn

通讯作者: 刘江, 博士, 研究员, 研究方向: 表观遗传。E-mail: liuj@big.ac.cn

DOI: 10.16288/j.ycz.18-189

网络出版时间: 2018/10/20 14:54:02

URI: <http://kns.cnki.net/kcms/detail/11.1913.R.20181020.1453.006.html>

regulate early embryogenesis, the relationship between chromatin structure and other epigenetic information such as methylation and histone modification. We hope that this review may shed light on the research of epigenome regulating embryo development.

Keywords: epigenome; embryo development; chromatin accessibility; chromatin 3D high-order structure

多细胞生物个体里的绝大部分细胞都携带同样的遗传物质 DNA, 然而不同类型细胞的形态、功能却千差万别。这种异质性是由各种细胞不同的基因表达模式造成的, 不同的基因表达模式由表观遗传所调控从而决定细胞命运。表观遗传是指由非 DNA 序列改变引起的、可遗传的基因表达调控方式, 主要包括 DNA 修饰、组蛋白修饰(histone modification)、非编码 RNA (no-coding RNA)、染色质开放性(chromatin accessibility)及染色质三维高级结构(3D chromatin organization)等^[1~3]。表观遗传广泛参与分化、发育、肿瘤发生等细胞生理过程。

两性生殖动物生命起始于卵子和精子的结合, 经过细胞分裂、基因正确有序激活表达等过程形成胚胎, 并进一步分化出不同的细胞类型和器官发育成新的个体。终端分化的精子和卵子的表观遗传状态有着巨大的差别, 融合后形成全能性的受精卵(zygote), 来自父源和母源的基因组在表观遗传状态逐步趋向一致, 最终会重编程为相同的状态。这种重编程对胚胎的早期发育起着至关重要的作用, 一旦出现差错会导致胚胎致死或者疾病、畸形^[4, 5]。

由于早期胚胎材料的稀缺及研究手段有限, 长期以来人们对早期胚胎表观遗传的继承和重编程规律不甚了解。近几年来基因组测序技术高速发展, 以及各种微量高效 DNA 分析新技术不断涌现, 例如甲基化测序(whole-genome bisulfite sequencing methylC-seq or BS-seq)^[6~8]、染色质免疫共沉淀测序(chromatin immunoprecipitation sequencing, ChIP-seq)^[9~12]、转座酶可接近性核染色质区域测序(assay for transposase accessible chromatin, ATAC-seq)^[13~16]、DNA 核酸酶超敏感位点测序(DNase I hypersensitive sites sequencing, DNase-seq)^[17~20]、高通量染色质构象捕获测序(high input chromosome conformation capture, Hi-C)^[21~25]等。这些技术使得研究早期胚胎表观遗传动态变化成为可能。鉴于脊椎动物早期胚胎的甲基化、组蛋

白修饰的重编程规律和调控作用已有多篇优秀文章研究讨论^[3, 26~28], 本文重点介绍近年来动物早期胚胎发育中染色质开放性及染色质三维高级结构方面重编程规律的研究进展。

1 染色质开放性重编程

在真核生物中, 顺式调控元件(*cis*-regulatory elements, 包括启动子 promoter、增强子 enhancer、沉默子 silencer、隔离子 insulator 等)与细胞类型特异的转录因子(transcription factor)共同指导了细胞的命运决定和个体的发育。转录因子通过 motif 识别或其他蛋白质介导结合到特定的 DNA 片段上, 发挥调控基因表达的重要作用。基因组 DNA 某个特定区域上的核小体包装不紧密甚至缺乏核小体, DNA 暴露出来, 可被核酸酶 (DNase) 识别并特异酶切, 称之为染色质开放性或可接近性区域^[29~31]。诸多研究表明转录因子往往结合在开放的染色质片段上。因此定位染色质的开放区域能够鉴定细胞或组织特异的调控元件和重要的转录因子, 并可研究染色质的开放状态改变与基因转录的关系。目前 DNase-seq 或 ATAC-seq 方法是定位染色质开放区域的两种常用方法。DNase-seq 利用核酸酶 特异酶切无核小体包装的裸露 DNA, 对酶切下来的 DNA 小片段建库测序, 得到核酸酶 超敏感位点 DHS (DNase hypersensitive sites)^[18, 19]; 而 ATAC-seq 同样是利用转座酶 Transposase 不能进入核小体连接致密的地方, 但能进入松散的区域并切割下暴露的 DNA 并同时连接上特异性的接头 adapters, 连接上 adapters 的 DNA 片段被分离出来用于二代测序^[16]。因此, DNase-seq 和 ATAC-seq 得到的都是全基因组尺度上处于开放状态的染色质区域。在分化的细胞或组织中, DNase-seq 和 ATAC-seq 得到的结果基本一致。

在哺乳动物早期胚胎发育过程中, 染色质开放

性经历了剧烈的变化, 并且染色质的开放决定了受精卵全基因组转录激活(zygotic genome activation, ZGA)事件。Lu 等^[32]用少量细胞 DNase-seq 方法生成小鼠(*Mus musculus*)受精卵至桑椹胚(morula)时期的染色质开放性图谱发现: 受精卵时期开放区域较少, 仅有 844 个; 随着胚胎发育进行, DHS 数目逐步增加, 尤其是 8 细胞时期增长显著, 达到 12 000 多个; 并且发现早期建立的 DHS 会一直维持至后面的发育阶段。各个时期 DHS 的分布情况显示: 受精卵时期 DHS 有 87.2% 位于启动子区, 8 细胞胚胎时期 DHS 将近一半位于远端调控区(distal region), 该结果说明基因表达随着发育进行而趋向受到多种调控元件的特异性调控。通过分析精子和卵子的 DHS 信号发现: 两者的染色质开放程度比较低, 但卵子比精子更开放一些。手动分离受精卵期间的雌雄原核分别分析父源和母源的染色质开放状态发现: 在受精后 7.5 h 的 PN3 阶段受精卵中, 父源基因组的开放状态已经与母源基本一致。而在后续的胚胎发育时期, 两个亲本的基因组开放状态高度相似, 仅仅发现一些区域的开放状态呈现亲本偏好形式, 其中部分是印记基因(imprinting gene)。这结果说明受精后, 父源和母源的染色质开放性便迅速重编程为相似状态。

同样使用 DNase-seq 方法研究人早期胚胎, 结果发现 DHS 数量也是随着胚胎发育逐渐增加, 总的趋势和小鼠高度相似, 说明人和小鼠早期胚胎的染色质开放性建立模式比较保守^[33]。小鼠 ZGA 发生在 2 细胞胚胎时期而人发生在 8 细胞胚胎时期^[34], 有意思的是 DHS 与 ZGA 密切相关。结合 DHS 和 ZGA 激活的基因分析发现: 小鼠 2 细胞胚胎时期的 DHS 富集于启动子区域并且将近 32% 的 ZGA 激活基因的启动子具有 DHS; 但人 4 细胞时期胚胎才仅有 8% ZGA 基因的启动子具有 DHS, 而 8 细胞阶段达到了 81%^[33]。这一结果说明小鼠和人启动子区 DHS 的建立与 ZGA 密切相关。

在大部分分化组织和细胞中, 使用 DNase-seq 和 ATAC-seq 得到的染色质开放性图谱基本一致^[16], 但在精子和早期胚胎中却有较大的差别。使用 ATAC-seq 方法研究小鼠精子时发现, 精子里大部分启动子和增强子区处于开放状态, 并与胚胎干细胞或其他成体分化细胞高度相似^[35], 但精子的 DHS 信

号不强^[32]。另外与 DNase-seq 结果不同, ATAC-seq 信号在小鼠和人早期胚胎尤其是 ZGA 前的染色质上广泛存在^[36, 37]。在小鼠早 2 细胞(early 2-cell) ZGA 前 ATAC-seq 信号会发现它们富集于重复元件(repeat elements)上, 并且这些信号里较强的一部分位于 MERV1 上或附近, 与此一致的是 RNA 分析结果显示 MERV1 在早 2 细胞 ZGA 前活跃表达。总体来说, 小鼠早 2 细胞胚胎染色质的 ATAC-seq 信号暗示其可能处于一种更松散的状态。而在 2 细胞至内细胞团(inner cell mass, ICM)时期, 早期胚胎与胚胎干细胞相似, 大部分 ATAC-seq 信号位于启动子区和远端元件区。另一方面, 2 细胞时期在转录终止位点下游(transcription end site, TES)有非同寻常的强烈 ATAC-seq 信号, 这种富集在后面时期减弱, 在体细胞里已经非常微弱。与 ATAC-seq 小鼠早期胚胎结论一致, 人在 ZGA 前的 1~4 细胞时期染色质上也广泛存在 ATAC peaks 信号。这些 1~4 细胞时期的 ATAC peaks 很多集中在 CpG 含量较高的启动子区, 并且大部分在 8 细胞时期维持, 说明启动子区的提前开放与基因的未来激活相关; 同时很多远端非启动子区也有很强的 ATAC peaks 信号, 并且这些区域富集于转录因子的结合位点, 然而这些 ATAC peaks 信号随着 ZGA 反而大量消失, 随后胚胎细胞在很多新的调控元件位置建立起开放染色质区域。在 ZGA 后胚胎时期, ATAC-seq 得到的染色质开放性图谱与 DNase-seq 比较类似, 例如在人胚胎中, 2 细胞时期 ATAC peaks 有 22 977 个, 但 DHS 仅有 729 个, 8 细胞时期两者的差异就小, 分别为 40 426 个和 39 813 个^[33, 37]。这说明早期胚胎 ZGA 前的染色质结构非常独特, 不同的检测方法得到的信号不同, 哪种方法得到的信号更接近真实的染色质开放状态有待进一步的验证。

转录因子在早期胚胎染色质开放性建立及启动 ZGA 发挥着至关重要的作用。OCT4 是一种干细胞多能性维持的重要转录因子, 在胚胎发育、胚层分化中起着重要的调控作用, 但它在小鼠和人中的作用规律不完全相同^[38]。分析早期胚胎 OCT4 基因上的 DHS 位点发现, 在人中 DHSs 位于 OCT4 基因内区域, 而在小鼠中位于启动子区域^[33]。通过转录因子结合片段富集分析(TF-binding motif enrichment

analysis)发现 OCT4 在人 ZGA 期(8 细胞期)高度富集在 DHS 上,相比之下,OCT4 在小鼠 ZGA(2 细胞期)期是不富集在 DHSs 上或者富集很少,这个结果暗示 OCT4 可能在人 ZGA 中具有重要作用^[33]。敲低 OCT4 (siOCT4)实验证明了这一点:OCT4 表达下降后,人 8 细胞期近 25%的 ZGA 基因表达随之下调,其中 85%下调基因的启动子区在 8 细胞期已建立 DHS;但小鼠 4 细胞期只有 1.4%的 ZGA 基因表达随之下调,说明 OCT4 及其密切相关的 DHS 在人 ZGA 事件中发挥了重要的作用而对小鼠 ZGA 影响较小^[33]。另外,研究还发现在小鼠中 OCT4 表达下降后,部分 8 细胞时期获得的 DHS 消失或减弱,并且这些 DHS 较多富集于远端调控区,暗示 OCT4 在小鼠胚胎中参与基因的远端调控作用^[32]。这些结果说明 OCT4 在小鼠和人 ZGA 事件和胚胎发育中发挥不同的调控作用。同样通过富集分析发现:小鼠 2 细胞启动子 DHS 富集于 Nfya 结合片段。Nfya 是 Nfy 复合物的一个亚基,该复合物能结合到基因组上并促进染色质开放^[39]。在小鼠胚胎中敲低 Nfya 后,约 28.4%的 2 细胞 DHS 减弱并且减弱的 DHS 富集于启动子区,同时发现 15.1%的 ZGA 基因表达下调,说明 Nfya 在小鼠中参与部分启动子 DHS 的形成及 ZGA 的激活^[32]。总之,转录因子广泛参与哺乳动物早期胚胎 DHS 的建立,调控着 ZGA 事件和胚胎发育。还有哪些转录因子在其中发挥作用有待于进一步研究。

胚胎发育早期的染色质开放性与重复元件表达活跃密切相关。小鼠和人基因组包含着大量的重复元件,重复元件在物种进化中具有非常重要的作用,部分重复元件在早期胚胎中转录活跃但在终端分化的细胞、组织中不表达^[40-42]。与此对应的是无论是 ATAC-seq 还是 DNase-seq 信号在早期胚胎中均发现有富集于重复元件上^[33, 36, 37]。例如在小鼠早 2 细胞时期,富集于重复元件特别是 MERVL 上;小鼠 2 细胞时期基因启动子和远端调控区的染色质开放信号特异地集中在基因组的重复序列上,尤其是 SINE、ERV 最为富集,并与此时期重复序列表达较活跃相关联。在人中部分转座元件例如 HERV-K、HERV-H、SVA、AluS,在早期胚胎中高表达而在分化组织中没有表达。分析人早期胚胎 DNase-seq 信号发现相

较于其他分化组织,这些转座子区域 DHS 在早期胚胎更加富集,说明转座子活化与染色质开放性相关联^[33]。同样在人 8 细胞时期,远端调控区的 ATAC-seq 信号特异地集中在 Alu、SINE、LTR,而 ERVK, SVA 和 ERV1 同时富集于 8 细胞时期和 ICM^[37]。这些结果说明早期胚胎重复元件的转录活跃与其区域上染色质的开放性密切相关。

在哺乳动物早期胚胎发育过程中,DNA 甲基化会经历一个重编程的过程:精子和卵子中的 DNA 甲基化程度都很高并且精子的甲基化水平显著高于卵子;受精后父源和母源基因组上甲基化均被大规模擦除,到植入前的囊胚阶段,胚胎的 DNA 甲基化水平降到最低点;当胚胎着床以后,通过从头甲基化(*de novo* methylation)的方式,胚胎的 DNA 甲基化水平渐渐升高^[43-45]。分析 DHSs 与 DNA 甲基化的相关性发现,开放区域富集于低甲基化区域,具有 DHS 信号的区域的甲基化水平比没有 DHS 信号区域的低^[33];无论 DNase-seq 还是 ATAC-seq 都发现人早期胚胎中高开放性的启动子一般都是高 CpG 密度,而低开放性的启动子一般都是低 CpG 密度^[33, 37];同时发现高 CpG 的启动子区域趋向于在早期发育阶段建立 DHSs,而低 CpG 的启动子区域趋向则相反^[33]。ZGA 前检测到的 ATAC-seq 信号很大程度上与卵子的部分甲基化区域(partially methylated domains, PMDs)重合,例如人 2 细胞胚胎有 73%的 ATAC peaks 与 PMD 重合^[37]。另外通过小鼠实验发现,雌原核的远端 ATAC peaks 与卵子 PMD 区域及受精卵母源非经典 H3K3me3 宽峰有极大程度的重合,而雄原核则没有。ZGA 后,PMD 区域上的 ATAC 开放信号和非经典 H3K3me3 均消失。父源基因组上 ZGA 前建立的 H3K3me3 也与父源的 ATAC 开放信号相重叠。这些结果说明染色质开放性与 DNA 甲基化及组蛋白修饰有一定程度上的相关性。

综上所述,小鼠和人在早期胚胎发育过程中的染色质开放性总体建立模式非常相似,都经历了逐步建立的过程,并且与 ZGA 密切相关。

2 染色质三维高级结构重编程

在高等真核生物细胞里,例如人线性非组

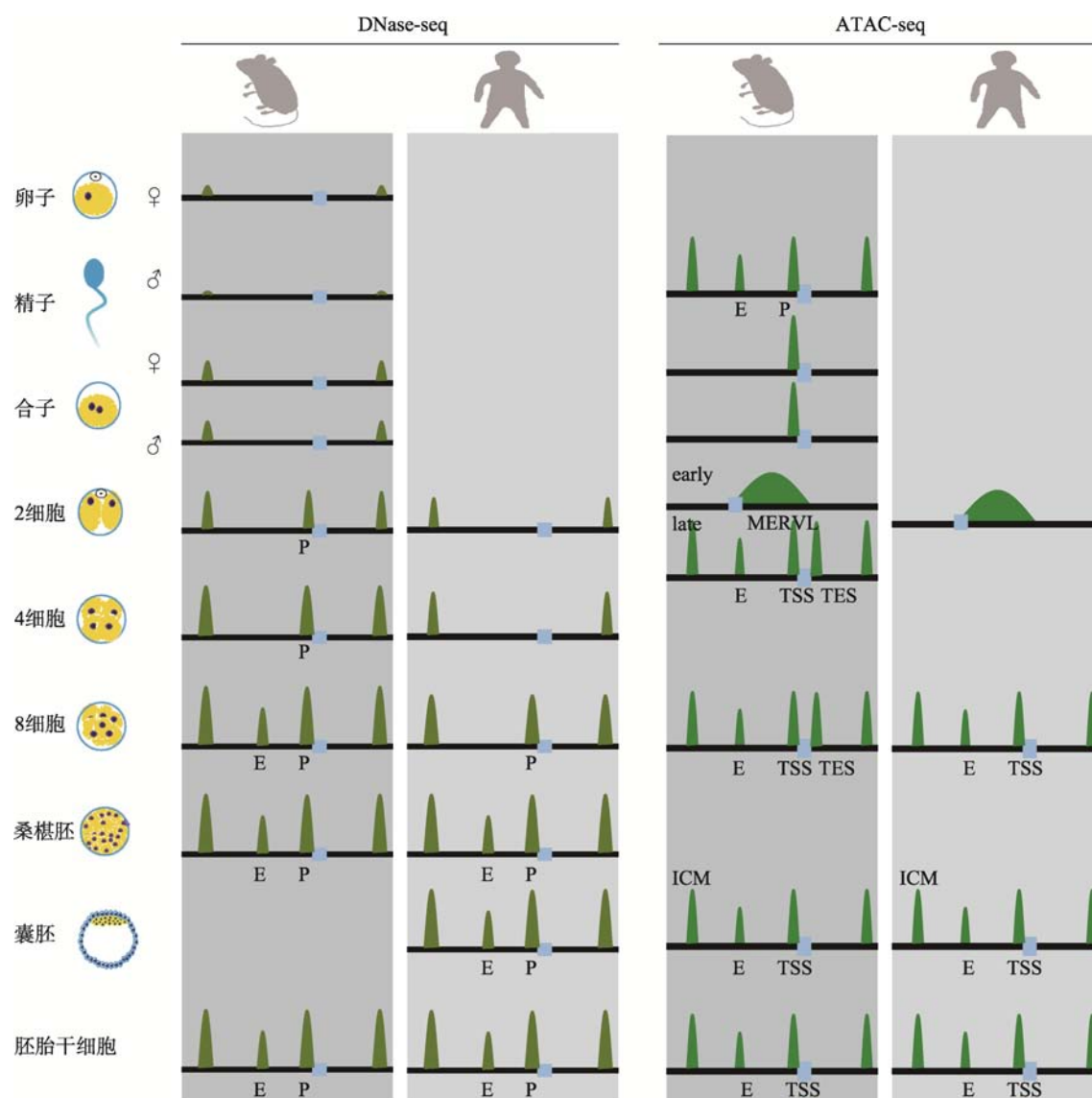


图 1 染色质开放性在小鼠和人早期胚胎发育的重编程

Fig. 1 The reprogramming of chromatin accessibility during early embryo development in the mouse and human

左侧部分显示 DNase-seq 方法检测到小鼠和人早期胚胎中的染色质开放状态模式, DNase-seq 信号在 ZGA 前基本检测不到; 右侧部分显示 ATAC-seq 方法检测到小鼠和人早期胚胎中的染色质开放状态模式, ATAC-seq 信号在 ZGA 前的染色质上广泛存在。在 ZGA 后, DNase-seq 与 ATAC-seq 得到的染色质开放性图谱比较类似。P 代表启动子(promoter); E 代表增强子(enhancer); TSS 代表转录起始位点(transcription start site); TES 代表转录终止位点(transcription end site)。

装 DNA 长度可达 2 m, 而细胞核直径仅有 10 μm 。为适应有限的细胞核空间, 在组蛋白及其他蛋白介导下, 染色质经过非常复杂有序的折叠形成多层次立体空间结构, 同时正确地按照时空顺序执行功能以确保生物机体正常生活及生长发育。染色质三维高级结构是表观遗传非常重要的部分, 在基因组调控中发挥着非常重要的作用, 如基因表达、DNA 复

制、分裂期基因组从母代传递到子代、维持基因组稳定性等^[46~48]。目前一般认为染色质按不同长度尺度分层次组装: 染色质疆域(chromosome territories), 即不同的染色质在细胞核内折叠、占据不同的特定的区域; 活性或惰性区室(A/B-compartment)结构, 与基因表达活性密切相关, 活性的开放的 A-compartment 的基因往往表达水平较高, 而惰性的封闭的

B-compartment 基因表达水平较低; 拓扑相关区域 (topologically associating domains, TADs) 结构, 染色质上一段连续的区域, 此区域上的所有位点 (loci) 互相之间相互作用频率高, 但与区域外的位点相互作用极少, TADs 普遍存在, 目前普遍认为 TAD 结构是基本的结构和功能单元; 亚 TAD 结构及环状 (loop) 结构, 具有组织特异性, 亚 TAD 结构与功能元件间的特殊 loop 相互作用直接相关, loop 相互作用很普遍, 常发生在活跃基因启动子、活跃增强子及 CTCF 结合位点上。TAD 和 compartment 是染色质三维高级结构的基础组成部分, 存在于所有的体细胞、组织和体外培养细胞株中。TAD 结构在种属间相当保守^[49], 胚胎干细胞 (ESC) 分化或诱导多能干细胞 (iPSC) 转化过程中, TAD 结构基本不变, 仅有小部分 compartment 发生活性 (A)/惰性 (B) 状态改变^[50, 51]。在细胞有丝分裂期, 基本检测不到染色体的 TAD 和 compartment 结构^[52]。传统上主要借助于显微镜研究染色质三维高级结构, 但显微镜在分辨率及高通量产出方面有很大的局限性。2002 年以来, 发展起一类基于分子生化的染色质构象捕获技术 (3C、4C、5C、Hi-C、ChIA-PET 等), 这些方法利用的是近距离连接反应原理, 即低浓度下分子内的连接反应优先发生。通过检测基因组上任意一对位点的相对相互作用频率 (relative frequency of interaction), 而描绘出染色质的精细分子构象。它们超越了显微镜的分辨率局限, 是一种侧面反映染色质构象的方法^[53]。其中高通量全基因组染色质构象捕获技术 Hi-C 以能检测全基因组的相互作用图谱及其高通量特性, 得到较为广泛的应用。通过改进 Hi-C 方法适用于少量细胞, 能帮助研究人员了解动物早期胚胎发育中的染色质三维高级结构动态变化规律。

小鼠成熟精子中染色质上大部分组蛋白被精蛋白取代, 仅保留约 1% 的组蛋白。小鼠成熟卵子停滞于第二次减数分裂中期 (M_{II}), 受精后才完成分裂排出第二极体。显微镜形态学观察结果显示: 小鼠配子和体细胞的染色质结构具有显著性差异, 精子、卵子染色质形态高度浓缩紧致, 而体细胞间期的染色质比较松散。然而通过 Hi-C 测序方法得到的小鼠精子高分辨率染色质三维构象图谱结果却出乎意料, 结果显示: 小鼠精子中存在典型的 TAD 和 comp-

artment 结构, 无论是 TAD 数目、TAD 结构模式还是 compartment 模式与小鼠体细胞的高度相似^[35, 54, 55]。对精子染色质结构进行进一步的分析发现: 精子同一条染色质内普遍存在着超远程 (大于 4 Mb) 的相互作用, 这种超远程的相互作用往往以 TAD 为基本单元, 发生在 TAD-TAD 间; 同时还发现精子不同染色质间的相互作用频率相对其他胚胎时期也高得多^[54], 这些结果非常符合精子染色质高度紧缩的生理特点。小鼠未成熟卵细胞 (non-surrounded nucleolus, NSN) 的染色质结构与体细胞的很相似, 具有典型的 loop、TAD、compartment 结构^[56]; 成熟为卵子 (surrounded-nucleolus, SN) 过程中, 逐渐关闭转录, 染色质与核膜连接消失, loop、TAD、compartment 强度逐渐降低; 最终成熟卵子染色质呈现出一种均一性结构, 并且缺乏 TADs 和 compartment 结构^[54, 55], 卵子的这种均一性结构与细胞分裂期染色质结构高度相似, 这非常符合成熟卵子处于 M 期的生理特点。

在精子和卵子结合形成受精卵的过程中, 父源染色质上的精蛋白替换回组蛋白, 母源染色质解压缩。利用电子光谱成像研究小鼠早期胚胎发现: 受精卵时期染色质呈现普遍消散 (dispersed) 状态, 在核膜附近没有明显的聚集; 2 细胞时期开始出现离散凝集区块, 在核膜附近聚集开始增多; 在 8 细胞时期和多能外胚层细胞里, 染色质呈现纤维网状结构, 与胚胎干细胞的极度相似^[57]。小鼠早期胚胎高分辨率染色质三维构象图谱与此成像表型相一致, 三维构象图谱分析结果显示^[54, 55]: 受精后, 染色质的三维结构迅速呈现为一种极其松散的状态; 受精卵和 2 细胞时期胚胎的 TAD 和 compartment 结构不清晰; 清晰的 TAD 和 compartment 结构随着胚胎发育的进行逐步建立起来; 8 细胞时期胚胎具有了比较经典的 TAD 和 compartment 结构。值得注意的是, 分析受精后处于细胞间期的受精卵和 2 细胞时期胚胎发现: 这两个时期胚胎细胞里的染色质竟然没有明显的 TAD 和 compartment 结构, 这是一种非常独特的结构。进一步分析发现, 这种结构与其他处于细胞间期的体细胞或体外培养细胞株的染色质结构都不同, 与有丝分裂期染色体结构也显著不同。仔细对比分析受精卵时期和胚胎发育后面时期的染色质相互作用图谱发现: 在后面时期的典型 TAD 的位置上,

受精卵在该位置有较高频率的相互作用(相对于背景而言),可以称之为 TAD 雏形或者类 TAD 结构,同样受精卵时期存在着 compartment 结构雏形^[55, 56]。随着胚胎发育的进行, TAD 雏形内部的相互作用频率增高,与 TAD 外部的相互作用减少,边界隔绝(insulation of boundary)效应增强, compartment 强度增加,直至 8 细胞时期胚胎形成经典的 TAD 和 compartment 结构。

精子和卵子的染色质三维高级结构差异显著,受精后父源和母源的染色质结构会发生怎样的变化呢?利用品系间的单核苷酸多态性(single nucleotide polymorphisms, SNPs)分别对来源于父源和母源的基因组进行研究发现,两者染色质结构建立模式高度相似:同样在受精卵和 2 细胞时期胚胎中没有清晰的 TAD 和 compartment 结构,随着发育的进行逐渐地建立起来经典的 TAD 和 compartment 结构^[54, 55]。在 PN3、PN5 阶段的受精卵中,发现父源染色质的远距离相互作用比母源的少一些,意味着受精卵中父源染色质在组蛋白替换后迅速呈现更为松散的状态。而在 compartment 结构上,受精卵中父源和母源的染色质也有细微差别,父源染色质 compartment 强度要高一些,这种差别一直持续到 8 细胞^[56, 58, 59]。总的来说,受精后父源和母源染色质迅速呈现相似而又有微小差别结构,发育至胚胎后期已基本一致,仅在一些等位特异(allele-specific)区域例如印记基因^[54]具有结构微小差异。

小鼠 ZGA 发生在晚 2 细胞期(late 2-cell), TAD 和 compartment 结构伴随着 ZGA 的发生而逐步建立。在果蝇(*Drosophila melanogaster*)胚胎发育中也同样发现 TAD 和 compartment 结构伴随着 ZGA 的发生而逐步建立。果蝇受精后仅细胞核进行分裂,并没有细胞质的分裂,当细胞核分裂至第 13 个周期(nuclear cycle, nc13)时才形成细胞膜。在 nc14 发生 ZGA,同时果蝇胚胎里染色质的 TAD 和 compartment 结构开始逐步建立^[60]。使用药物 α -amanitin 处理小鼠受精卵,抑制 ZGA 的发生,发现染色质仍然能够建立 TAD 和 compartment 结构,说明染色质三维高级结构的建立不依赖于 ZGA^[54, 55]。同样,果蝇早期胚胎染色质三维高级结构的建立不依赖于 ZGA^[60]。染色质三维高级结构的逐步建立伴随着 ZGA 发生

但又同时不依赖于 ZGA 这一规律特点是否在其他物种中也存在还有待于进一步验证,各物种早期胚胎染色质三维高级结构的建立机制还知之甚少,需要更多的研究探索。另一方面通过药物 aphidicolin 抑制小鼠受精卵 DNA 复制,发现不能建立 TAD 和 compartment 结构,说明染色质三维高级结构的建立依赖于基因组的复制^[54]。小鼠受精卵复制分裂至 2 细胞时期而果蝇复制分裂至 nc14 开始建立染色质三维高级结构,这其中蕴含着怎样的调节机制还有待进一步探索。

胚胎干细胞的分化过程和体细胞重编程为 iPSCs 过程都会发生部分 compartment 状态转换,及伴随基因表达的协同变化。与此现象相一致,在胚胎发育过程中同样有部分 compartment 发生了 A/B 状态的转换。使用基因本体(Gene Ontology, GO)方法分析时期特异(stage-specific) compartment 上的基因发现:这些基因富集于生殖系统发育或胚胎发育过程,这说明染色质三维高级结构参与调控生殖发育及胚胎发育的进行^[54]。分析染色质三维高级结构与其他表观遗传信息发现有着密切的关联^[54]:例如非甲基化 CpG 相对比较富集于 A-compartment; A-compartment 上的启动子(promoter)有 60% 处于非甲基化状态,而 B-compartment 上的启动子仅有 20% 处于非甲基化状态。总之,非甲基化启动子往往调控基因更好的表达,位于 A-compartment 的基因一般表达水平更高。早期胚胎发育会发生全基因组去甲基化,分析发现 A-compartment 上的 CpG 甲基化的下降水平比 B-compartment 上的 CpG 甲基化的下降水平要高,暗示 A-compartment 更利于 DNA 甲基转移酶接近执行去甲基化功能^[54, 58]。通过整合小鼠早期胚胎染色质开放性数据发现,各时期 DHSs 富集于 A-compartment,同时某个特定胚胎时期新获得的 DHS 位点也是富集于 A-compartment。总而言之, A-compartment 甲基化程度更低,开放性更好,转录更活跃。

综上所述,早期胚胎染色质具有独特的结构,伴随着 ZGA 发生逐步建立起经典的高级结构。这种建立模式在其他种属生物是否一样,以及何种机制指导胚胎细胞建立正确的染色质结构,都是亟待解决的问题。

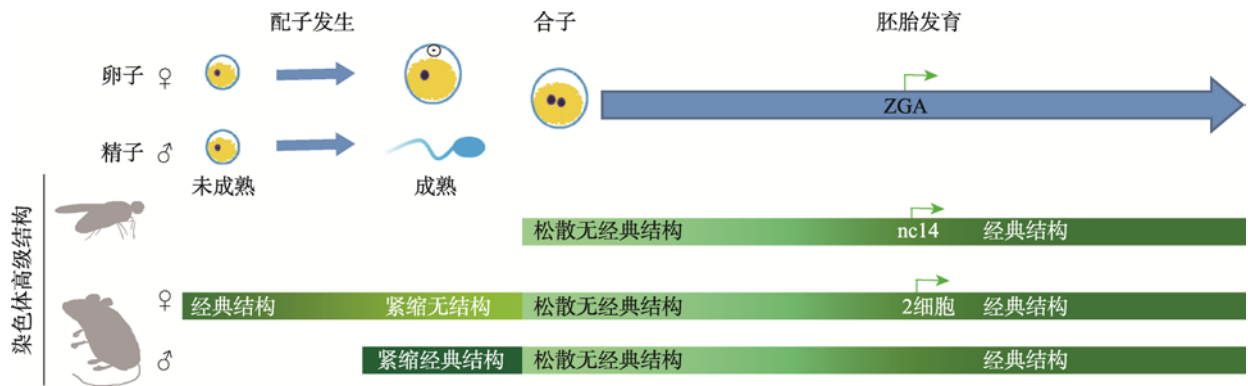


图 2 染色质三维高级结构在早期胚胎的重编程

Fig. 2 The reprogramming of chromatin high-order structures during embryogenesis

果蝇和小鼠早期胚胎发育过程中,伴随着 ZGA 发生逐步建立起经典的染色质三维高级结构。ZGA 代表受精卵基因组转录激活。nc14 代表细胞核分裂第 14 个周期。

3 结语与展望

动物早期胚胎发育是一个非常特异的过程,终端分化的配子融合形成全能性的胚胎,最终分化形成各种细胞、组织和器官,除了 DNA 序列,表观遗传信息、RNA、蛋白质、信号通路网络、代谢甚至细胞形态都发生了剧烈的变化。微量 DNA 分析技术的不断发展使得早期胚胎的表观遗传图谱初步浮现,然而还有太多的谜题等待解答:表观遗传信息有些能遗传到下一代而有些不能,是什么在其中发挥了选择作用?早期胚胎中在没有模板的情况下,某些表观遗传信息是如何建立起新的模式?表观遗传信息之间更深层次的关联以及如何协调指导早期胚胎发育等等,这些都是仍然有待阐明的热点问题。另一方面,近年来蓬勃发展的单细胞转录组测序技术显示即使同类型细胞也存在细胞-细胞之间的异质性,这也暗示着细胞-细胞之间表观遗传调控的异质性。单细胞转录组技术显示高度对称相似的早期胚胎 4 细胞时期胚胎已呈现细胞-细胞之间表达谱的异质性^[61]。单细胞转录组和表观遗传信息整合分析将帮助我们研究何种因素影响细胞个体基因表达模式,每个细胞将要分化形成的细胞类型,哪些关键基因和表观遗传信息驱动发育阶段前进,帮助我们全面了解绘制早期胚胎发育的表观遗传图谱及细胞命运决定的调控机理。因此单细胞表观遗传测序分析技术将是未来研究热点之一。早期胚胎表观遗传

的研究有助于人们全面认识人早期胚胎发育机制,为解决临床生殖医学的各种问题提供线索,从而提高人口质量。

参考文献(References):

- [1] Wang KC, Chang HY. Epigenomics: technologies and applications. *Circ Res*, 2018, 122(9): 1191–1199. [DOI]
- [2] Rivera CM, Ren B. Mapping human epigenomes. *Cell*, 2013, 155(1): 39–55. [DOI]
- [3] Xu Q, Xie W. Epigenome in early mammalian development: inheritance, reprogramming and establishment. *Trends Cell Biol*, 2018, 28(3): 237–253. [DOI]
- [4] Dean W, Lucifero D, Santos F. DNA methylation in mammalian development and disease. *Birth Defects Res C Embryo Today*, 2005, 75(2): 98–111. [DOI]
- [5] Marcho C, Cui W, Mager J. Epigenetic dynamics during preimplantation development. *Reproduction*, 2015, 150(3): R109–120. [DOI]
- [6] Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM, Edsall L, Antosiewicz-Bourget J, Stewart R, Ruotti V, Millar AH, Thomson JA, Ren B, Ecker JR. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*, 2009, 462(7271): 315–322. [DOI]
- [7] Lister R, Mukamel EA, Nery JR, Urich M, Puddifoot CA, Johnson ND, Lucero J, Huang Y, Dwork AJ, Schultz MD, Yu M, Tonti-Filippini J, Heyn H, Hu S, Wu JC, Rao A, Esteller M, He C, Haghghi FG, Sejnowski TJ, Behrens MM, Ecker JR. Global epigenomic reconfiguration during mammalian brain development. *Science*, 2013, 341(6146):

1237905. [DOI]
- [8] Xie W, Barr CL, Kim A, Yue F, Lee AY, Eubanks J, Dempster EL, Ren B. Base-resolution analyses of sequence and parent-of-origin dependent DNA methylation in the mouse genome. *Cell*, 2012, 148(4): 816–831. [DOI]
- [9] Rotem A, Ram O, Shores N, Sperling RA, Goren A, Weitz DA, Bernstein BE. Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. *Nat Biotechnol*, 2015, 33(11): 1165–1172. [DOI]
- [10] Brind'amour J, Liu S, Hudson M, Chen C, Karimi MM, Lorincz MC. An ultra-low-input native ChIP-seq protocol for genome-wide profiling of rare cell populations. *Nat Commun*, 2015, 6: 6033. [DOI]
- [11] Shankaranarayanan P, Mendoza-Parra MA, Walia M, Wang L, Li N, Trindade LM, Gronemeyer H. Single-tube linear DNA amplification (LinDA) for robust ChIP-seq. *Nat Methods*, 2011, 8(7): 565–567. [DOI]
- [12] Lara-Astiaso D, Weiner A, Lorenzo-Vivas E, Zaretzky I, Jaitin DA, David E, Keren-Shaul H, Mildner A, Winter D, Jung S, Friedman N, Amit I. Immunogenetics chromatin state dynamics during blood formation. *Science*, 2014, 345(6199): 943–949. [DOI]
- [13] Buenrostro JD, Wu B, Litzenburger UM, Ruff D, Gonzales ML, Snyder MP, Chang HY, Greenleaf WJ. Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature*, 2015, 523(7561): 486–490. [DOI]
- [14] Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. ATAC-seq: a method for assaying chromatin accessibility genome-wide. *Curr Protoc Mol Biol*, 2015, 109: 21.29.1–9. [DOI]
- [15] Cusanovich DA, Daza R, Adey A, Pliner HA, Christiansen L, Gunderson KL, Steemers FJ, Trapnell C, Shendure J. Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science*, 2015, 348(6237): 910–914. [DOI]
- [16] Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods*, 2013, 10(12): 1213–1218. [DOI]
- [17] Jin W, Tang Q, Wan M, Cui K, Zhang Y, Ren G, Ni B, Sklar J, Przytycka TM, Childs R, Levens D, Zhao K. Genome-wide detection of DNase I hypersensitive sites in single cells and FFPE tissue samples. *Nature*, 2015, 528(7580): 142–146. [DOI]
- [18] Song L, Crawford GE. DNase-seq: a high-resolution technique for mapping active gene regulatory elements across the genome from mammalian cells. *Cold Spring Harb Protoc*, 2010, 2010(2): pdb.prot5384. [DOI]
- [19] John S, Sabo PJ, Canfield TK, Lee K, Vong S, Weaver M, Wang H, Vierstra J, Reynolds AP, Thurman RE, Stamatoyannopoulos JA. Genome-scale mapping of DNase I hypersensitivity. *Curr Protoc Mol Biol*, 2013, Chapter 27: Unit 21.27. [DOI]
- [20] John S, Sabo PJ, Thurman RE, Sung MH, Biddie SC, Johnson TA, Hager GL, Stamatoyannopoulos JA. Chromatin accessibility pre-determines glucocorticoid receptor binding patterns. *Nat Genet*, 2011, 43(3): 264–268. [DOI]
- [21] Lieberman-Aiden E, Van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, 2009, 326(5950): 289–293. [DOI]
- [22] Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, Aiden EL. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*, 2014, 159(7): 1665–1680. [DOI]
- [23] Nagano T, Lubling Y, Stevens TJ, Schoenfelder S, Yaffe E, Dean W, Laue ED, Tanay A, Fraser P. Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature*, 2013, 502(7469): 59–64. [DOI]
- [24] Stevens TJ, Lando D, Basu S, Atkinson LP, Cao Y, Lee SF, Leeb M, Wohlfahrt KJ, Boucher W, O'shaughnessy-Kirwan A, Cramard J, Faure AJ, Ralser M, Blanco E, Morey L, Sanso M, Palayret MGS, Lehner B, Di Croce L, Wutz A, Hendrich B, Klennerman D, Laue ED. 3D structures of individual mammalian genomes studied by single-cell Hi-C. *Nature*, 2017, 544(7648): 59–64. [DOI]
- [25] Ramani V, Deng X, Qiu R, Gunderson KL, Steemers FJ, Disteché CM, Noble WS, Duan Z, Shendure J. Massively multiplex single-cell Hi-C. *Nat Methods*, 2017, 14(3): 263–266. [DOI]
- [26] Ci W, Liu J. Programming and inheritance of parental DNA methylomes in vertebrates. *Physiology (Bethesda)*, 2015, 30(1): 63–68. [DOI]
- [27] Smith ZD, Meissner A. DNA methylation: roles in mammalian development. *Nat Rev Genet*, 2013, 14(3): 204–220. [DOI]
- [28] Deng D. DNA methylation and demethylation: current status and future perspective. *Hereditas (Beijing)*, 2014,

- 36(5): 403–410.
邓大军. DNA 甲基化和去甲基化的研究现状及思考. 遗传, 2014, 36(5): 403–410. [DOI]
- [29] Boyle AP, Song L, Lee BK, London D, Keefe D, Birney E, Iyer VR, Crawford GE, Furey TS. High-resolution genome-wide in vivo footprinting of diverse transcription factors in human cells. *Genome Res*, 2011, 21(3): 456–464. [DOI]
- [30] Henikoff S. Nucleosome destabilization in the epigenetic regulation of gene expression. *Nat Rev Genet*, 2008, 9(1): 15–26. [DOI]
- [31] Tsompana M, Buck MJ. Chromatin accessibility: a window into the genome. *Epig Chrom*, 2014, 7(1): 33. [DOI]
- [32] Lu F, Liu Y, Inoue A, Suzuki T, Zhao K, Zhang Y. Establishing chromatin regulatory landscape during mouse preimplantation development. *Cell*, 2016, 165(6): 1375–1388. [DOI]
- [33] Gao L, Wu K, Liu Z, Yao X, Yuan S, Tao W, Yi L, Yu G, Hou Z, Fan D, Tian Y, Liu J, Chen ZJ, Liu J. Chromatin accessibility landscape in human early embryos and its association with evolution. *Cell*, 2018, 173(1): 248–259.e15. [DOI]
- [34] Niakan KK, Han J, Pedersen RA, Simon C, Pera RA. Human pre-implantation embryo development. *Development*, 2012, 139(5): 829–841. [DOI]
- [35] Jung YH, Sauria ME, Lyu X, Cheema MS, Ausio J, Taylor J, Corces VG. Chromatin states in mouse sperm correlate with embryonic and adult regulatory landscapes. *Cell Rep*, 2017, 18(6): 1366–1382. [DOI]
- [36] Wu J, Huang B, Chen H, Yin Q, Liu Y, Xiang Y, Zhang B, Liu B, Wang Q, Xia W, Li W, Li Y, Ma J, Peng X, Zheng H, Ming J, Zhang W, Zhang J, Tian G, Xu F, Chang Z, Na J, Yang X, Xie W. The landscape of accessible chromatin in mammalian preimplantation embryos. *Nature*, 2016, 534(7609): 652–657. [DOI]
- [37] Wu J, Xu J, Liu B, Yao G, Wang P, Lin Z, Huang B, Wang X, Li T, Shi S, Zhang N, Duan F, Ming J, Zhang X, Niu W, Song W, Jin H, Guo Y, Dai S, Hu L, Fang L, Wang Q, Li Y, Li W, Na J, Xie W, Sun Y. Chromatin analysis in human early development reveals epigenetic transition during ZGA. *Nature*, 2018, 557(7704): 256–260. [DOI]
- [38] Fogarty NME, McCarthy A, Snijders KE, Powell BE, Kubikova N, Blakeley P, Lea R, Elder K, Wamaitha SE, Kim D, Maciulyte V, Kleinjung J, Kim JS, Wells D, Vallier L, Bertero A, Turner JMA, Niakan KK. Genome editing reveals a role for OCT4 in human embryogenesis. *Nature*, 2017, 550(7674): 67–73. [DOI]
- [39] Nardini M, Gnesutta N, Donati G, Gatta R, Forni C, Fossati A, Vornrhein C, Moras D, Romier C, Bolognesi M, Mantovani R. Sequence-specific transcription factor NF-Y displays histone-like DNA binding and H2B-like ubiquitination. *Cell*, 2013, 152(1–2): 132–143. [DOI]
- [40] Cordaux R, Batzer MA. The impact of retrotransposons on human genome evolution. *Nat Rev Genet*, 2009, 10(10): 691–703. [DOI]
- [41] Grow EJ, Flynn RA, Chavez SL, Bayless NL, Wossidlo M, Wesche DJ, Martin L, Ware CB, Blish CA, Chang HY, Pera RA, Wysocka J. Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells. *Nature*, 2015, 522(7555): 221–225. [DOI]
- [42] Peaston AE, Evsikov AV, Graber JH, De Vries WN, Holbrook AE, Solter D, Knowles BB. Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Dev Cell*, 2004, 7(4): 597–606. [DOI]
- [43] Wang L, Zhang J, Duan J, Gao X, Zhu W, Lu X, Yang L, Zhang J, Li G, Ci W, Li W, Zhou Q, Aluru N, Tang F, He C, Huang X, and Liu J. Programming and inheritance of parental DNA methylomes in mammals. *Cell*, 2014, 157(4): 979–991. [DOI]
- [44] Li G, Yu Y, Fan Y, Li C, Xu X, Duan J, Li R, Kang X, Ma X, Chen X, Ke Y, Yan J, Lian Y, Liu P, Zhao Y, Zhao H, Chen Y, Sun X, Liu J, Qiao J, Liu J. Genome wide abnormal DNA methylome of human blastocyst in assisted reproductive technology. *J Genet Genomics*, 2017, 44(10): 475–481. [DOI]
- [45] Guo H, Zhu P, Yan L, Li R, Hu B, Lian Y, Yan J, Ren X, Lin S, Li J, Jin X, Shi X, Liu P, Wang X, Wang W, Wei Y, Li X, Guo F, Wu X, Fan X, Yong J, Wen L, Xie SX, Tang F, Qiao J. The DNA methylation landscape of human early embryos. *Nature*, 2014, 511(7511): 606–610. [DOI]
- [46] Furlan-Magaril M, Varnai C, Nagano T, Fraser P. 3D genome architecture from populations to single cells. *Curr Opin Genet Dev*, 2015, 31: 36–41. [DOI]
- [47] Fraser J, Williamson I, Bickmore WA, Dostie J. An overview of genome organization and how we got there: from FISH to Hi-C. *Microbiol Mol Biol Rev*, 2015, 79(3): 347–372. [DOI]
- [48] Dekker J, Mirny L. The 3D genome as moderator of chromosomal communication. *Cell*, 2016, 164(6): 1110–1121. [DOI]
- [49] Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. Topological domains in mammalian genomes identified by analysis of chromatin interactions.

- Nature*, 2012, 485(7398): 376–380. [DOI]
- [50] Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, Lee AY, Ye Z, Kim A, Rajagopal N, Xie W, Diao Y, Liang J, Zhao H, Lobanenko VV, Ecker JR, Thomson JA, Ren B. Chromatin architecture reorganization during stem cell differentiation. *Nature*, 2015, 518(7539): 331–336. [DOI]
- [51] Krijger PH, Di Stefano B, De Wit E, Limone F, Van Oevelen C, De Laat W, Graf T. Cell-of-origin-specific 3D genome structure acquired during somatic cell reprogramming. *Cell Stem Cell*, 2016, 18(5): 597–610. [DOI]
- [52] Naumova N, Imakaev M, Fudenberg G, Zhan Y, Lajoie BR, Mirny LA, Dekker J. Organization of the mitotic chromosome. *Science*, 2013, 342(6161): 948–953. [DOI]
- [53] Dekker J, Misteli T. Long-range chromatin interactions. *Cold Spring Harb Perspect Biol*, 2015, 7(10): a019356. [DOI]
- [54] Ke Y, Xu Y, Chen X, Feng S, Liu Z, Sun Y, Yao X, Li F, Zhu W, Gao L, Chen H, Du Z, Xie W, Xu X, Huang X, Liu J. 3D chromatin structures of mature gametes and structural reprogramming during mammalian embryogenesis. *Cell*, 2017, 170(2): 367–381.e20. [DOI]
- [55] Du Z, Zheng H, Huang B, Ma R, Wu J, Zhang X, He J, Xiang Y, Wang Q, Li Y, Ma J, Zhang X, Zhang K, Wang Y, Zhang MQ, Gao J, Dixon JR, Wang X, Zeng J, Xie W. Allelic reprogramming of 3D chromatin architecture during early mammalian development. *Nature*, 2017, 547(7662): 232–235. [DOI]
- [56] Flyamer IM, Gassler J, Imakaev M, Brandao HB, Ulianov SV, Abdennur N, Razin SV, Mirny LA, Tachibana-Konwalski K. Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. *Nature*, 2017, 544(7648): 110–114. [DOI]
- [57] Ahmed K, Dehghani H, Rugg-Gunn P, Fussner E, Rossant J, Bazett-Jones DP. Global chromatin architecture reflects pluripotency and lineage commitment in the early mouse embryo. *PLoS One*, 2010, 5(5): e10531. [DOI]
- [58] Zhang Y, Xiang Y, Yin Q, Du Z, Peng X, Wang Q, Fidalgo M, Xia W, Li Y, Zhao ZA, Zhang W, Ma J, Xu F, Wang J, Li L, Xie W. Dynamic epigenomic landscapes during early lineage specification in mouse embryos. *Nat Genet*, 2017, 50(1): 96–105. [DOI]
- [59] Gassler J, Brandão HB, Imakaev M, Flyamer IM, Ladstätter S, Bickmore WA, Peters JM, Mirny LA, Tachibana K. A mechanism of cohesin-dependent loop extrusion organizes zygotic genome architecture. *EMBO J*, 2017, 36(24): 3600–3618. [DOI]
- [60] Hug CB, Grimaldi AG, Kruse K, Vaquerizas JM. Chromatin architecture emerges during zygotic genome activation independent of transcription. *Cell*, 2017, 169(2): 216–228.e19. [DOI]
- [61] 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]

(责任编辑: 高绍荣)