探索染色质三维构象的“工具箱”研究进展

林达^{1,2}, 张斯姮^{1,2}, 张智慧^{1,2}, 徐伟泽^{1,2}, 洪萍^{3,4}, 李国亮^{3,4*}, 曹罡^{1,2*}

1. 华中农业大学, 农业微生物学国家重点实验室, 武汉 430070;

2. 华中农业大学动物科学技术学院、动物医学院, 武汉 430070;

3. 华中农业大学, 作物遗传改良国家重点实验室, 武汉 430070;

4. 华中农业大学信息学院, 武汉 430070

* 联系人, E-mail: guoliang.li@mail.hzau.edu.cn; gcao@mail.hzau.edu.cn

收稿日期: 2019-12-20; 接受日期: 2020-02-13; 网络版发表日期: 2020-04-30

国家自然科学基金(批准号: 31771402, 31900432)、博士后创新人才支持计划(批准号: BX20180113)和中国博士后科学基金(批准号: 2018M640710)资助

摘要 三维基因组学是研究基因组三维空间结构与功能的一门学科。基于不同的研究目的和生物学问题, 目前发展出了各种各样的三维基因组学研究方法, 这些方法对推动三维基因组学发展有积极的影响。本文以三维基因组学的提出和发展过程为线索, 对三维基因组学的主要研究方法进行了系统地归类与介绍, 包括基于测序技术(基于邻近连接)的Hi-C, capture Hi-C, single-cell Hi-C, DLO Hi-C, ChIA-PET, Hi-ChIP等; 基于测序技术(不基于邻近连接)的ChIA-Drop, SPRITE, GAM; 以及基于影像技术的seqFISH, CRISPRainbow, SABER amplifies FISH。最后, 本文对染色体构象捕获技术的发展方向进行了简单的展望。

关键词 三维基因组学, 染色体构象, Hi-C, DLO Hi-C, ChIA-PET, FISH

人类基因组包含2万个编码基因以及大量的调控元件。在过去的10年时间里, 研究人员根据细胞类型的不同, 把活跃在不同细胞中的调控元件进行了聚类^[1]。DNA元件百科全书计划(Encyclopedia of DNA Elements, ENCODE)、表观组学线路图计划(Roadmap Epigenome)、人类表观遗传学合作组织(International Human Epigenome Consortium)以及哺乳动物基因组的功能性注释计划(FANTOM projects)已经注释了数千个基因, 以及数百万个与这些基因相关的调控元件^[2~5]。但是, 这些调控元件是如何跨越几千碱基甚至上百万碱基的距离对靶基因发挥调节作用, 还没有清

晰的解释^[6]。

染色质的空间折叠及其在细胞核中的结构对核内基因的表达有深远的影响。例如, 增强子必须要与目的基因发生空间上的互作, 才可以激活目的基因的表达^[7~9], 并且基因的沉默以及基因组DNA的复制都与染色质的折叠密切相关^[10,11]。此外, 全基因组关联分析(genome-wide association studies, GWAS)已经确定了大量与疾病相关的突变位点, 这些突变位点大部分都是位于非编码区的潜在调控元件^[12]。癌细胞中存在大量的基因重排现象, 这都与基因组DNA的三维结构有或多或少的联系^[13,14]。2017年, 研究人员通过对哺乳

引用格式: 林达, 张斯姮, 张智慧, 等. 探索染色质三维构象的“工具箱”研究进展. 中国科学: 生命科学, 2020, 50: 497–505

Lin D, Zhang S H, Zhang Z H, et al. The “toolbox” for chromatin conformation capture (in Chinese). Sci Sin Vitae, 2020, 50: 497–505, doi: 10.1360/SSV-2019-0224

动物早期胚胎发育过程中染色体三维结构重构过程的解析, 发现了染色体三维结构在受精初期呈现出极其松散的状态, 且不具有任何的topologically associated domain(TAD)^[15], A/B compartment^[16]等结构特征. 随着胚胎的发育, 染色体的三维构象及高级结构才慢慢成熟建立^[17]. 随后, 2018年研究人员发现, 在甲型流感病毒(influenza A virus, IAV)感染期间, 转录紊乱导致的RNA聚合酶II过度转录延伸会迅速改变宿主细胞染色质的相互作用. 这些互作的改变通常发生在高转录水平基因的3'末端. 此外, 为了利于转录因子的结合, 转录延伸会将处于转录抑制状态的染色质区域(B compartment)转换为转录激活状态的染色质区域(A compartment)^[18]. 这些研究成果都说明了染色质三维结构与转录调控之间的紧密联系, 也为揭示疾病发生的深层次原因提供了理论基础.

为了解决不同的生物学问题, 现阶段发展出各式各样的染色体构象捕获技术, 大部分技术都是基于邻近连接来固定染色体互作信息, 包括chromosome conformation capture(3C)^[19], high-throughput chromosome conformation capture(Hi-C)^[16], chromatin interaction analysis by paired-end tag sequencing(ChIA-PET)^[20]等. 这样做的优势是可以简单迅速地通过酶连反应将远距离染色体的互作信息固定下来. 但是邻近连接对染色体间互作的捕获能力有限, 为了获取更丰富的基因组互作信息, 近期研究人员开发了几种不依赖于邻近连接的染色体构象捕获技术, 包括基于“拆分池”的SPRITE技术^[21], 基于10x Genomics^[22]测序平台的ChIA-Drop技术^[23]以及基于冰冻切片的genome architecture mapping(GAM)技术^[24]. 利用这些技术, 研究人员描绘了由核体(nuclear bodies)介导的染色体间互作模型^[21], 并且观测到了由RNA聚合酶II转录延伸所导致的染色体三维结构的改变^[23]. 以上提出的所有高通量染色体构象捕获技术, 都依托二代或三代测序平台来读取基因组互作信息. 为了直接肉眼观测到染色质的空间结构, 研究人员将CRISPR(clustered regularly interspaced short palindromic repeats)技术应用在基因组标记中, 开发出可对活细胞基因组多位点同时标记的CRISPRainbow技术^[25], 该技术能够同时标记多达7个不同位置的基因组位点, 每种位点对应一种不同的颜色, 允许人们实时追踪染色体构象的改变. 此外, 庄小威课题组^[26,27]将超分辨显微镜技术与原位杂交技术

相结合, 成功观测到了染色质的具体互作情况.

针对目前已经开发的三维基因组学技术, 本文进行了整理和归类, 为相关学者和研究人员提供参考(表1). 下文进行相关的介绍.

1 测序技术(基于邻近连接)

1.1 染色体构象捕获技术

针对不同的研究目的, 三维基因组的研究方法有很多. 最早的染色体构象捕获技术(3C), 是在2002年由Dekker等人^[19]提出来的. 该方法首先会对细胞核进行甲醛固定, 甲醛的交联作用可以将细胞核内的互作DNA与蛋白交联在一起, 把全基因组范围内的基因组互作信息都固定下来; 再通过限制性内切酶酶切、扩大反应体积邻近连接, 将线性距离很远、但空间距离很近的两个调控区域酶连在一起; 最后通过荧光定量PCR, 验证感兴趣的基因组互作是否真实发生. 3C技术的建立, 在DNA FISH(fluorescence *in situ* hybridization)^[39]之外, 首次提供了一种对染色体空间构象研究的方法, 开创了基因组学研究的新领域.

1.2 Hi-C技术

由于3C技术提出的年代较早, 高通量测序技术还没有普及, 所以3C有个很大的缺陷就是通量太低, 只能对一对基因组互作位点进行验证. 随着高通量测序的发展, Dekker课题组^[16]将高通量测序技术与3C相结合, 于2009年发表了Hi-C技术. Hi-C技术将3C文库构建过程进行了改进: 当染色质酶切结束后, 用生物素修饰的核苷酸对DNA黏性末端进行标记, 然后再进行邻近连接; 连接结束后, 用超声的方法对回收的连接产物进行打断, 然后再用链霉亲和素回收互作DNA结点; 最后将回收的互作DNA构建高通量测序文库进行高通量测序.

通过对Hi-C数据的分析, 研究人员发现, 基因组按照互作模式的不同可以划分为两类不同的区间, 也叫做A/B compartment^[40]. A compartment一般处于常染色质区, 这个区域基因转录活性较高, 基因调控元件大量结合, 是一个较松散的染色质区域. 而B compartment处于异染色质区域, 该区域基因的转录活性较低, 基因组紧密包裹, 调控元件往往很难在该区域结合.

如果进一步对A/B compartment进行划分, 基因组

表 1 主要染色体构象捕获技术比较及其技术特点

Table 1 Comparison of major chromosome conformation capture techniques and their technical characteristics

	技术分类	解析精度	技术特点
测序技术 (基于邻近连接)	3C ^[18]	群体细胞	有选择性地验证基因组两个位点是否存在空间上的互作, 但通量较低, 只能进行一对一的位置互作验证.
	Hi-C ^[15]	群体细胞	通量高, 可捕获基因组全局的染色质空间构象和互作信息, 但随机连接噪音较高.
	<i>In situ</i> Hi-C ^[28]	群体细胞	在Hi-C的基础上, 采用细胞核内连接的方式, 降低了数据的噪音, 但需要大量细胞才可进行文库构建.
	Capture Hi-C ^[29,30]	群体细胞	采用探针捕获的方式, 只捕获感兴趣基因组位点的互作信息, 需设计大量的捕获探针, 捕获效率取决于捕获位点的选择.
	Single cell Hi-C ^[31]	单细胞水平	前期步骤与 <i>in situ</i> Hi-C基本类似, 邻近连接后, 会用玻璃毛细管分选单细胞核来构建Hi-C文库, 可以解析单细胞染色质的空间构象, 但分辨率较低.
	Single-nucleus Hi-C ^[32]	单细胞水平	引入了单细胞基因组扩增技术, 将邻近连接的DNA产物进行放大, 不需要生物素标记, 能解析单细胞染色质空间构象, 但单细胞扩增的偏好性可能会导致结果不准确.
	Dip-C ^[33]	单细胞水平	采用了优化后的单细胞基因组扩增技术, 偏好性更小, 不需要生物素标记, 能解析单细胞染色质空间构象.
	sciHi-C ^[34]	单细胞水平	采用拆分池技术, 既能捕获群体细胞的染色质空间构象, 又可以解析群体细胞内单个细胞的染色质空间结构, 但需要大量细胞才可进行文库构建.
	DLO Hi-C ^[35]	群体细胞	实验过程不需要生物素标记, 并且引入了不同的DNA接头, 对随机连接的噪音进行评估, 由于MmeI酶切长度的限制, 适用于人、小鼠等基因组碱基均衡的物种, 不适用于玉米、高粱等基因组重复序列高的物种.
	ChIA-PET ^[36]	群体细胞	加入了抗体富集的步骤, 能捕获特定转录因子介导的基因组互作信息, 但需要大量细胞才可进行文库构建.
测序技术 (不基于邻近连接)	HiChIP ^[37]	群体细胞	结合了 <i>in situ</i> Hi-C和转座酶Tn5介导的文库构建的技术优势, 大大降低了起始细胞量, 同时提高了DNA互作片段的捕获效率.
	GAM ^[23]	单细胞水平	不依赖甲醛交联, 不需要邻近连接, 起始细胞量少, 分辨率大小取决于冰冻切片的厚度, 现阶段分辨率较低.
	SPRITE ^[20]	群体细胞	采用拆分池技术, 通过给每个蛋白-DNA分子加上不同的标签, 捕获单分子内的染色质互作信息, 现阶段需要大量细胞才可进行文库构建.
影像技术	ChIA-Drop ^[22]	群体细胞	利用10x Genomics平台, 将每个蛋白-DNA分子形成一个油包水的液滴, 并加上唯一的标签, 最后捕获单分子内的染色质互作信息, 现阶段需要大量细胞才可进行文库构建.
	CRISPRainbow ^[24]	单细胞水平	活细胞成像, 能够实时追踪特定定位点的染色体构象的变化情况, 不同位点用不同的颜色来进行区分, 由于荧光颜色数目的限制, 目前能够实时观测的染色体位点较少.
	DNA seqFISH ^[25-27]	单细胞水平	采用原位杂交的方式, 通过多轮DNA杂交, 解析染色质某一区域的精细折叠情况, 目前只能解析局部区域的染色质精细结构.
	SABER amplifies FISH ^[38]	单细胞水平	利用引物交换反应(primer-exchange reaction, PER), 将传统FISH杂交探针的信号进行放大, 增强荧光信号强度, 该方法可用于DNA与RNA的检测.

可以被划分为更小的、一般以Mb大小为单位的拓扑结构域——TAD^[11]. TAD是基因组折叠的单元, 在一个TAD结构内部, 大量的转录调控元件折叠在一起, 互作紧密, 而在不同的TAD之间, 互作较少.

由于测序深度的限制, 以前的Hi-C数据分辨率较低, 难以捕获近距离的启动子和增强子互作信息, 只能在Mb或者几十kb的分辨率下对A/B compartment以

及TAD进行分析. 2014年发表的原位Hi-C(*in situ* Hi-C)^[28]对传统的Hi-C方法进行了优化, 采用4碱基酶(MboI)以及核内邻近连接的方法, 在有足够测序深度的情况下, 将Hi-C的分辨率提高到1 kb. 通过原位Hi-C, 研究人员阐明了CTCF蛋白是如何介导染色体折叠(loop)形成的, 并在GM12878细胞系中, 对9448个loop进行了注释. 高分辨率的Hi-C数据可以将具体的启动

子与增强子通过loop连接起来, 精确地寻找某个启动子区域的调控元件。此外, 研究人员利用基因组特异性杂交探针与Hi-C文库进行杂交, 只对感兴趣的基因组调控区域及其互作区域进行捕获, 选择性地分析癌症风险基因、易感基因等特定区域的基因组互作情况, 这就是现在应用广泛的capture Hi-C^[29,30]。

1.3 单细胞Hi-C染色体构象捕获技术

传统的Hi-C技术是通过捕获群体细胞的平均染色质来解析一群细胞的染色质三维构象。2013年, 为了解析不同单细胞间染色质构象的差异, Nagano等人^[31]发表了单细胞Hi-C技术(single-cell Hi-C), 该技术以Hi-C技术为基础, 以单个细胞核为研究对象, 能够分辨单细胞染色质的构象模型, 并阐明染色质的相互作用以及基因组功能的受控机制。不同组织类型的单细胞染色质三维结构的解析是现阶段表观遗传学的研究热点^[41]。但单细胞Hi-C技术基于传统Hi-C技术, 需要生物素对染色质末端进行分子标记, 且缺乏体外线性扩增的步骤, 微量互作DNA的回收难度大, 技术门槛很高, 大部分实验室都难以进行单细胞Hi-C实验。后期慢慢发展出一些基于单细胞扩增的单细胞Hi-C技术, 例如single-nucleus Hi-C^[32]和Dip-C^[33]。此外, 研究人员还将原位Hi-C与“拆分池”技术^[42]相结合, 发明了sciHi-C(single-cell combinatorial indexed Hi-C)^[34]。该技术用带有标签的DNA接头将每个细胞进行“打码”标记, 然后进行混合、拆分, 再经过测序接头的第二轮“打码”标记, 将每个细胞都带上唯一的标记, 最后混合建库、测序, 这样就可以既看到“树木”, 又看到“森林”地获取群体细胞和单个细胞的染色质互作信息。

1.4 DLO Hi-C

DLO Hi-C(digestion-ligation-only Hi-C)^[35]最大的特点是采用同时酶切酶连的方式, 将含有MmeI的DNA接头连接在染色体内切酶切口末端上, 然后进行邻近酶连, 最后用MmeI内切酶酶切消化, 回收固定大小的互作DNA片段进行测序。DLO Hi-C最显著的特征是引入了含有MmeI的DNA接头, 通过凝胶电泳、切胶回收固定大小的互作DNA片段就可捕获染色体的互作信息, 因此取消了繁琐生物素标记、链霉亲和素磁珠回收等步骤, 大大地降低了背景数据噪音, 获取

了高质量的染色质互作数据, 并可在测序前通过A、B两种接头的连接情况, 预判文库中随机连接噪音的大小。此外, DLO Hi-C使用EGS(ethylene glycol bis(succinimidyl succinate))和甲醛水溶液对细胞进行双交联固定, 可以捕获非直接接触的DNA-蛋白弱互作, 同时避免后期实验过程中解交联的发生。与传统的Hi-C技术相比, DLO Hi-C提高了测序数据利用率、缩短了文库构建时间、降低了文库构建成本。

1.5 ChIA-PET

Hi-C捕获到的是全基因组范围内的互作信息, 所有的调控元件都包含在其中。如果想对某个转录因子所介导的基因组互作进行分析, Hi-C的测序深度往往难以达到。2009年, 也就是Hi-C建立的同一年, 研究人员平行发表了另一种染色质构象捕获技术ChIA-PET^[36]。ChIA-PET技术将ChIP^[43-46]技术与染色质构象捕获技术结合到一起, 在染色质构象捕获过程中, 加入了抗体富集的步骤, 特异性地捕获感兴趣的转录因子或者其他调控元件介导的基因组互作信息^[47]。利用ChIA-PET技术, 研究人员全面地描绘了CCCTC结合因子(CCCTC-binding factor, CTCF)与RNA聚合酶II(RNAPII)共同介导的染色质折叠以及这些特异性的染色质折叠对转录活性的影响^[8,48]。

1.6 HiChIP

由于ChIA-PET的初始细胞量较大, 一般需要 1×10^7 个细胞才可进行抗体富集及后期的文库构建, 为了进一步提高捕获效率, 研究人员将*in situ* Hi-C和转座酶Tn5介导的文库构建方法相结合, 开发出了Hi-ChIP(*in situ* Hi-C followed by chromatin immunoprecipitation)^[49]技术。HiChIP的前期步骤与*in situ* Hi-C^[28]类似, 在细胞核内对染色质进行限制性内切酶酶切、生物素末端标记以及邻近连接。当核内邻近连接结束后, 研究人员对细胞核进行裂解、超声以及抗体富集, 并对感兴趣的转录因子介导的染色质互作信息进行捕获。得到DNA-蛋白复合物后, 进行解交联, 再用链霉亲和素磁珠回收互作DNA结点, 最后在链霉亲和素磁珠上直接进行转座酶Tn5介导的文库构建。HiChIP结合了*in situ* Hi-C和转座酶Tn5介导的文库构建的技术优势, 大大地降低了起始细胞量, 同时提高了DNA互作片段的捕获效率。

1.7 其他基于邻近连接的染色体构象捕获技术

除了上述代表性的染色体构象捕获技术外, 国内外开发了大量的高效的染色体构象捕获技术, 包括利用DNA酶进行染色质片段化的DNase Hi-C^[50]和*in situ* DNase Hi-C^[37]; 利用桥式接头(bridge linkers)进行邻近连接的BL Hi-C^[51]和BAT Hi-C^[52]; 利用苯酚氯仿对开放区域染色质互作片段进行分离的OCEAN-C^[53]; 利用完整的高通量测序接头, 直接与互作DNA片段进行连接, 不通过PCR扩增直接测序的SAFE Hi-C^[54]. 这些优秀的染色体构象捕获技术在一定程度上推动着三维基因组学的发展.

2 测序技术(不基于邻近连接)

染色体间互作大部分都是核体(nuclear bodies)介导发生的^[55], 而核体的分子量和空间体积很大, 直径可达0.5~2 μm ^[56], 互作的DNA距离很远, 传统的邻近连接很难将染色体间互作信息捕获下来. 近几年发表了几种不依赖于邻近连接的染色体构象捕获技术, 分别是genome architecture mapping(GAM)^[24], split-pool recognition of interactions by tag extension(SPRITE)^[21]以及ChIA-Drop^[23], 本文分别对这几种技术进行简要介绍.

2.1 GAM

GAM^[24]将冰冻切片技术引入到染色体构象捕获中, 将细胞核切成厚度为120 nm的薄片, 然后利用激光显微切割技术和全基因组扩增技术, 将每一切片薄层内部的基因组互作DNA构建成为高通量测序文库, 最后测序、分析统计每一层的基因组互作情况, 获取全基因组互作图谱. GAM技术不需要对细胞进行交联固定, 不需要邻近连接, 只需要几百个细胞即可获取基因组互作信息, 其利用物理切片的方法研究基因组三维结构, 具有很强创新性.

2.2 SPRITE

SPRITE^[21]则利用超声和酶解的方式, 将染色质打断成150~1000 bp的蛋白-DNA复合物, 然后利用“拆分池”的方式, 将蛋白-DNA复合物的末端连接上带有特异识别序列的DNA标签接头, 然后进行混合、拆分、

加第二轮DNA标签接头, 如此反复, 这样可以把每个单分子蛋白-DNA复合物的互作DNA末端都带上相同且区别于其他蛋白-DNA复合物分子的标签, 最后通过建库测序, 获取每个蛋白-DNA复合物分子的DNA互作信息. SPRITE很好地解决了邻近连接的技术缺陷, 绘制出了更为精细的染色体间的互作图谱.

2.3 ChIA-Drop

2019年发表的ChIA-Drop^[23]将ChIA-PET和微流控技术(如10x Genomics平台)结合在一起, 其先将染色质片段化成6000 bp左右, 然后用抗体对片段化的染色质进行特异性富集; 结束后, 将染色质单分子利用10x Genomics平台配套试剂盒以及特异性的标签引物混合制备成“油包水”的小液滴; 对每个小液滴进行单分子扩增以及标记后, 利用高通量测序平台对其进行测序; 最后获取每个蛋白-DNA复合物内部包含的DNA互作信息^[23].

3 影像技术

3.1 CRISPRainbow

为了直接肉眼观测到染色质的空间结构, 2016年, 来自麻省大学医学院的科学家开发出了基于CRISPR/Cas9^[57-59]的染色体实时示踪技术: CRISPRainbow^[25]. 该技术将guide RNA(gRNA)进行了改造, 让其能与不同颜色的荧光蛋白特异地结合在一起. 不同颜色的排列组合, 可以对不同的基因组位点进行区分, 允许人们实时追踪染色体构象的改变. CRISPRainbow能够同时确定多达6种不同的DNA位点, 每种位点对应一种不同的颜色. 在活细胞中采用CRISPRainbow能够追踪基因组动态的拓扑运动, 具有重要的生物学意义.

3.2 seqFISH

同样基于显微成像技术, 庄小威课题组^[26,27]将超分辨显微技术与RNA原位杂交测序技术(sequential FISH, seqFISH)^[60-62]相结合, 开发出针对于基因组DNA的seqFISH技术, 并成功观测到染色体的精细结构. 定位特定区域在基因组上的互作情况一般分为两步: 第一步是杂交阶段, 该阶段将第一轮杂交探针与染色体进行杂交, 每个杂交位点的范围为30 kb, 其中均匀分布300条杂交探针, 每条杂交探针有一个25 bp

的显色探针杂交区; 第二步是显色阶段, 这一步会用带有荧光的针对于第一个杂交位点的显色探针, 与杂交探针的25 bp的显色区进行杂交, 然后显微成像、光淬灭, 再用带有荧光标记的探针对第二个杂交位点的显色探针与杂交探针的25 bp的显色区进行杂交, 然后显微成像、光淬灭, 如此反复, 获取某一特定区域的基因组互作情况^[27]。基于原位杂交与显微成像的染色体构象捕获技术, 不仅可以分析基因组特定位点的空间互作情况, 还可以观测这一区域在细胞核中的位置分布, 应用前景广阔。

3.3 SABER amplifies FISH

在seqFISH中, 一般一条杂交探针只结合一条显色探针, 为了检测一个位点的荧光信号, 往往需要数百条杂交探针进行杂交与显色^[27], 缺少杂交信号放大的步骤。2019年发表的SABER amplifies FISH^[63], 将交换反应信号扩增(signal amplification by exchange reaction, SABER)技术引入FISH中, 通过引物交换反应(primer-exchange reaction, PER)^[63], 将普通的FISH探针加上长单链DNA串联体, 在串联体上可杂交多个互补荧光成像单链以放大FISH信号。利用SABER进行信号扩增放大与传统的rolling-circle amplification(RCA)^[64]和hybridization chain reaction(HCR)^[38,65,66]扩增放大技术相比, 有以下5个优势: (i) 信号扩增可控性高, 效率高; (ii)

能够对多条杂交探针信号同时进行扩增放大; (iii) 扩增后的探针具有很高的杂交效率; (iv) 扩增后探针的杂交与成像步骤简单; (v) 扩增成本较低。利用SABER amplifies FISH, 研究人员成功解析了人类X染色体的空间结构。

4 展望

由于组织样本的异质性高, 很难获取高纯度的组织细胞样本, 而Hi-C实验又要求较大的起始细胞量, 所以现阶段进行染色体构象捕获实验的样本, 大部分还是细胞系。虽然现阶段开发了一些基于微量细胞和单细胞的染色体构象捕获技术, 但由于其实验操作难度大, 只有极少数实验室在进行单细胞和微量细胞Hi-C的实验, 后期需开发一些简单、可靠、重复性好的针对于微量细胞的染色体构象捕获技术, 来满足对微量组织样本的实验需求。此外, 由于技术原因, 以前的染色体构象捕获技术对染色体间的互作捕获能力非常有限, 导致主要研究都集中在染色体内互作, 对染色体间互作情况知之甚少。后期需要开发更多的不基于邻近连接的、简单且易操作的染色体构象捕获技术, 来观测不同细胞周期、不同发育阶段、不同生理病理条件、不同药物处理后的细胞染色体间互作情况, 为探索细胞转录调控和表观遗传研究奠定基础。

参考文献

- 1 Dekker J, Belmont A S, Guttman M, et al. The 4D nucleome project. *Nature*, 2017, 549: 219–226
- 2 Kundaje A, Meuleman W, Ernst J, et al. Integrative analysis of 111 reference human epigenomes. *Nature*, 2015, 518: 317–330
- 3 Stunnenberg H G, Abignani S, Adams D, et al. The International Human Epigenome Consortium: A blueprint for scientific collaboration and discovery. *Cell*, 2016, 167: 1145–1149
- 4 Carninci P, Kasukawa T, Katayama S, et al. The transcriptional landscape of the mammalian genome. *Science*, 2005, 309: 1559–1563
- 5 ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature*, 2012, 489: 57–74
- 6 Li G L, Ruan Y J, Gu R S, et al. Emergence of 3D genomics (in Chinese). *Chin Sci Bull*, 2014, 59: 1165–1172 [李国亮, 阮一骏, 谷瑞升, 等. 起航三维基因组学研究. 科学通报, 2014, 59: 1165–1172]
- 7 Tolhuis B, Palstra R J, Splinter E, et al. Looping and interaction between hypersensitive sites in the active β -globin locus. *Mol Cell*, 2002, 10: 1453–1465
- 8 Tang Z, Luo O J, Li X, et al. CTCF-mediated human 3D genome architecture reveals chromatin topology for transcription. *Cell*, 2015, 163: 1611–1627
- 9 Deng W, Rupin J W, Krivega I, et al. Reactivation of developmentally silenced globin genes by forced chromatin looping. *Cell*, 2014, 158: 849–860
- 10 Guelen L, Pagie L, Brasset E, et al. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature*,

2008, 453: 948–951

- 11 Pope B D, Ryba T, Dileep V, et al. Topologically associating domains are stable units of replication-timing regulation. *Nature*, 2014, 515: 402–405
- 12 Maurano M T, Humbert R, Rynes E, et al. Systematic localization of common disease-associated variation in regulatory DNA. *Science*, 2012, 337: 1190–1195
- 13 Roix J J, McQueen P G, Munson P J, et al. Spatial proximity of translocation-prone gene loci in human lymphomas. *Nat Genet*, 2003, 34: 287–291
- 14 Zhang Y, McCord R P, Ho Y J, et al. Spatial organization of the mouse genome and its role in recurrent chromosomal translocations. *Cell*, 2012, 148: 908–921
- 15 Nora E P, Lajoie B R, Schulz E G, et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature*, 2012, 485: 381–385
- 16 Lieberman-Aiden E, van Berkum N L, Williams L, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, 2009, 326: 289–293
- 17 Du Z, Zheng H, Huang B, et al. Allelic reprogramming of 3D chromatin architecture during early mammalian development. *Nature*, 2017, 547: 232–235
- 18 Heinz S, Texari L, Hayes M G B, et al. Transcription elongation can affect genome 3D structure. *Cell*, 2018, 174: 1522–1536.e22
- 19 Dekker J, Rippe K, Dekker M, et al. Capturing chromosome conformation. *Science*, 2002, 295: 1306–1311
- 20 Zhang J, Poh H M, Peh S Q, et al. ChIA-PET analysis of transcriptional chromatin interactions. *Methods*, 2012, 58: 289–299
- 21 Quinodoz S A, Ollikainen N, Tabak B, et al. Higher-order inter-chromosomal hubs shape 3D genome organization in the nucleus. *Cell*, 2018, 174: 744–757.e24
- 22 Kitzman J O. Haplotypes drop by drop. *Nat Biotechnol*, 2016, 34: 296–298
- 23 Zheng M, Tian S Z, Capurso D, et al. Multiplex chromatin interactions with single-molecule precision. *Nature*, 2019, 566: 558–562
- 24 Beagrie R A, Scialdone A, Schueler M, et al. Complex multi-enhancer contacts captured by genome architecture mapping. *Nature*, 2017, 543: 519–524
- 25 Ma H, Tu L C, Naseri A, et al. Multiplexed labeling of genomic loci with dCas9 and engineered sgRNAs using CRISPRainbow. *Nat Biotechnol*, 2016, 34: 528–530
- 26 Boettiger A N, Bintu B, Moffitt J R, et al. Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature*, 2016, 529: 418–422
- 27 Bintu B, Mateo L J, Su J H, et al. Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. *Science*, 2018, 362: eaau1783
- 28 Rao S S P, Huntley M H, Durand N C, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*, 2014, 159: 1665–1680
- 29 Jäger R, Migliorini G, Henrion M, et al. Capture Hi-C identifies the chromatin interactome of colorectal cancer risk loci. *Nat Commun*, 2015, 6: 6178
- 30 Mifsud B, Tavares-Cadete F, Young A N, et al. Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. *Nat Genet*, 2015, 47: 598–606
- 31 Nagano T, Lubling Y, Stevens T J, et al. Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature*, 2013, 502: 59–64
- 32 Flyamer I M, Gassler J, Imakaev M, et al. Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. *Nature*, 2017, 544: 110–114
- 33 Tan L, Xing D, Chang C H, et al. Three-dimensional genome structures of single diploid human cells. *Science*, 2018, 361: 924–928
- 34 Ramani V, Deng X, Qiu R, et al. Massively multiplex single-cell Hi-C. *Nat Methods*, 2017, 14: 263–266
- 35 Lin D, Hong P, Zhang S, et al. Digestion-ligation-only Hi-C is an efficient and cost-effective method for chromosome conformation capture. *Nat Genet*, 2018, 50: 754–763
- 36 Fullwood M J, Liu M H, Pan Y F, et al. An oestrogen-receptor- α -bound human chromatin interactome. *Nature*, 2009, 462: 58–64
- 37 Ramani V, Cusanovich D A, Hause R J, et al. Mapping 3D genome architecture through *in situ* DNase Hi-C. *Nat Protoc*, 2016, 11: 2104–2121
- 38 Dirks R M, Pierce N A. Triggered amplification by hybridization chain reaction. *Proc Natl Acad Sci USA*, 2004, 101: 15275–15278
- 39 Langer-Safer P R, Levine M, Ward D C. Immunological method for mapping genes on *Drosophila* polytene chromosomes. *Proc Natl Acad Sci*

- USA, 1982, 79: 4381–4385
- 40 Fortin J P, Hansen K D. Reconstructing A/B compartments as revealed by Hi-C using long-range correlations in epigenetic data. *Genome Biol*, 2015, 16: 180
- 41 Stevens T J, Lando D, Basu S, et al. 3D structures of individual mammalian genomes studied by single-cell Hi-C. *Nature*, 2017, 544: 59–64
- 42 Cusanovich D A, Daza R, Adey A, et al. Multiplex single-cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science*, 2015, 348: 910–914
- 43 Johnson D S, Mortazavi A, Myers R M, et al. Genome-wide mapping of *in vivo* protein-DNA interactions. *Science*, 2007, 316: 1497–1502
- 44 Barski A, Cuddapah S, Cui K, et al. High-resolution profiling of histone methylations in the human genome. *Cell*, 2007, 129: 823–837
- 45 Robertson G, Hirst M, Bainbridge M, et al. Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat Methods*, 2007, 4: 651–657
- 46 Skene P J, Henikoff S. A simple method for generating high-resolution maps of genome-wide protein binding. *eLife*, 2015, 4: e09225
- 47 Rusk N. When chia pets meet Hi-C. *Nat Methods*, 2009, 6: 863
- 48 Handoko L, Xu H, Li G, et al. CTCF-mediated functional chromatin interactome in pluripotent cells. *Nat Genet*, 2011, 43: 630–638
- 49 Mumbach M R, Rubin A J, Flynn R A, et al. HiChIP: Efficient and sensitive analysis of protein-directed genome architecture. *Nat Methods*, 2016, 13: 919–922
- 50 Ma W, Ay F, Lee C, et al. Fine-scale chromatin interaction maps reveal the *cis*-regulatory landscape of human lincRNA genes. *Nat Methods*, 2015, 12: 71–78
- 51 Liang Z, Li G, Wang Z, et al. BL-Hi-C is an efficient and sensitive approach for capturing structural and regulatory chromatin interactions. *Nat Commun*, 2017, 8: 1622
- 52 Huang J, Jiang Y, Zheng H, et al. BAT Hi-C maps global chromatin interactions in an efficient and economical way. *Methods*, 2020, 170: 38–47
- 53 Li T, Jia L, Cao Y, et al. OCEAN-C: Mapping hubs of open chromatin interactions across the genome reveals gene regulatory networks. *Genome Biol*, 2018, 19: 54
- 54 Niu L, Shen W, Huang Y, et al. Amplification-free library preparation with SAFE Hi-C uses ligation products for deep sequencing to improve traditional Hi-C analysis. *Commun Biol*, 2019, 2: 1
- 55 Hu Y, Plutz M, Belmont A S. *Hsp70* gene association with nuclear speckles is *Hsp70* promoter specific. *J Cell Biol*, 2010, 191: 711–719
- 56 Pederson T. The nucleolus. *Cold Spring Harb Perspect Biol*, 2011, 3: a000638
- 57 Hsu P D, Lander E S, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell*, 2014, 157: 1262–1278
- 58 Zhang F, Wen Y, Guo X. CRISPR/Cas9 for genome editing: Progress, implications and challenges. *Hum Mol Genet*, 2014, 23: R40–R46
- 59 Ran F A, Hsu P D, Wright J, et al. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc*, 2013, 8: 2281–2308
- 60 Shah S, Takei Y, Zhou W, et al. Dynamics and spatial genomics of the nascent transcriptome by intron seqFISH. *Cell*, 2018, 174: 363–376.e16
- 61 Shah S, Lubeck E, Zhou W, et al. seqFISH accurately detects transcripts in single cells and reveals robust spatial organization in the hippocampus. *Neuron*, 2017, 94: 752–758.e1
- 62 Takei Y, Shah S, Harvey S, et al. Multiplexed dynamic imaging of genomic loci by combined CRISPR imaging and DNA sequential FISH. *Biophys J*, 2017, 112: 1773–1776
- 63 Kishi J Y, Lapan S W, Beliveau B J, et al. SABER amplifies FISH: Enhanced multiplexed imaging of RNA and DNA in cells and tissues. *Nat Methods*, 2019, 16: 533–544
- 64 Lizardi P M, Huang X, Zhu Z, et al. Mutation detection and single-molecule counting using isothermal rolling-circle amplification. *Nat Genet*, 1998, 19: 225–232
- 65 Choi H M T, Chang J Y, Trinh L A, et al. Programmable *in situ* amplification for multiplexed imaging of mRNA expression. *Nat Biotechnol*, 2010, 28: 1208–1212
- 66 Choi H M T, Beck V A, Pierce N A. Next-generation *in situ* hybridization chain reaction: Higher gain, lower cost, greater durability. *ACS Nano*, 2014, 8: 4284–4294

The “toolbox” for chromatin conformation capture

LIN Da^{1,2}, ZHANG SiHeng^{1,2}, ZHANG ZhiHui^{1,2}, XU WeiZe^{1,2}, HONG Ping^{3,4}, LI GuoLiang^{3,4} & CAO Gang^{1,2}

1 State Key Laboratory of Agricultural Microbiology, Huazhong Agricultural University, Wuhan 430070, China;

2 College of Veterinary Medicine, Huazhong Agricultural University, Wuhan 430070, China;

3 National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan 430070, China;

4 College of Informatics, Huazhong Agricultural University, Wuhan 430070, China

Studies on the 3D genomics could facilitate the understanding of relationship between genome structure and functions. Based on different research purposes, a variety of chromatin three-dimensional conformation capture methods have been developed, greatly promoting the progress of three-dimensional genomics research. This review aims to systematically introduce and summarize various 3D genomics research methods. The methods are generally divided into three categories. The first category is based on proximity ligation and sequencing, including 3C, Hi-C, *in situ* Hi-C, capture Hi-C, single-cell Hi-C, DLO Hi-C, and ChIA-PET. The other category is not based on proximity ligation, such as ChIA-Drop, SPRITE, and GAM. The last category is based on microscopy, including seqFISH, CRISPRainbow, and SABER amplifies FISH. Finally, the perspective of conformation capture technologies is discussed.

3D genomics, chromatin conformation, Hi-C, DLO Hi-C, ChIA-PET, FISH

doi: [10.1360/SSV-2019-0224](https://doi.org/10.1360/SSV-2019-0224)