

文章编号: 2095-2163(2020)03-0348-05

中图分类号: TP391; Q811.4

文献标志码: A

# Chia-Pet 技术与应用研究

许立

(哈尔滨工业大学 计算机科学与技术学院, 哈尔滨 150001)

**摘要:** 特定 DNA 调控元件之间的长距离染色质接触在基因表达调控中起着关键作用,在理解信号网络和细胞状态时,必须对这些三维(3D)染色质结构中的相互作用进行全局表征。利用成对末端标记序列(Chia-Pet)进行染色质相互作用分析是一种将功能染色质结构转化为数百万个短标记序列的方法。自2009年开发以来,在染色质相互作用分析中具有独特的优势,从而为转录调控的研究提供了新的视角。本文介绍了 Chia-Pet 的实验方案和数据分析过程,分析了几种常用工具各自的特点和适用范围,帮助研究人员选用合适的方法以获得更可靠的结果。

**关键词:** 基因表达调控; 三维染色质结构; Chia-Pet

## Research on methods and technology of Chia-Pet

XU Li

(School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China)

**[Abstract]** Long-distance chromatin contact between specific DNA regulatory elements plays a key role in gene expression regulation. Global characterization of the interactions in these three-dimensional (3D) chromatin structures is essential in understanding signal networks and cell states. Chromatin interaction analysis using Paired-End-Tag sequencing (Chia-Pet) is a method for transforming functional chromatin structures into millions of short labeling sequences. Since its development in 2009, it has unique advantages in chromatin interaction analysis, which provides a new perspective for the study of transcriptional regulation. This paper introduces the experimental scheme and data analysis process of Chia-Pet, analyses the characteristics and application scope of several commonly used tools, and helps researchers choose appropriate methods to obtain more reliable results.

**[Key words]** gene expression regulation; three-dimensional chromatin structures; Chia-Pet

## 0 引言

转录调控是真核生物中一个复杂而有序的过程,其中染色质相互作用起着关键作用,从而调节基因表达,并进一步影响其他细胞的活动。许多研究转录因子(tf)与转录调控的相结合的技术已经被开发出来。例如染色质免疫沉淀(chip)微阵列(chip chip)<sup>[1]</sup>、chip pet<sup>[2]</sup>和 chip seq<sup>[3]</sup>,但却无法确定远端tf结合位点的靶基因。另一个挑战是确定这种远端结合位点是否具有功能性,即通过染色体环在物理上接近靶基因启动子,或吸引RNA聚合酶II复合物进行基因转录。因此,鉴定全基因组远端染色质相互作用,将调控元件引导至目标基因,可能为转录调控的研究提供新的视角。染色体构象捕获(3c)<sup>[4]</sup>及其衍生物,4c<sup>[5-6]</sup>和5c<sup>[7]</sup>可以揭示参与转录调控的长程染色质相互作用,但这些技术受到限制,或者是因为其整体性较低,如3c,或者是因其无法在整个基因组中绘制高分辨率的相互作用区域<sup>[8]</sup>。染色质相互作用分析与配对末端标记测序(Chia-Pet)方法就能够符合分析高吞吐量 and 高分辨

率基因组水平上染色质相互作用这些要求。与Hi-C<sup>[9]</sup>相比,Chia-Pet在与功能研究相关的蛋白质相关的更高分辨率上更好,确定TF结合位点和染色质相互作用,为以三维(3D)方式研究长程染色质相互作用奠定了坚实的基础,并提供了更可靠的方式。目前,Chia-Pet已成功应用于人MCF7细胞<sup>[10]</sup>、人癌细胞<sup>[11]</sup>、人T细胞<sup>[12]</sup>、小鼠胚胎干细胞<sup>[13]</sup>、小鼠神经祖细胞<sup>[14]</sup>和小鼠B细胞<sup>[15]</sup>以及其他细胞<sup>[16]</sup>。

为了系统评价Chia-Pet的方法,本文将详细探讨该方法的实验方案,与此同时,为方便后续研究分析,很多分析Chia-Pet数据的计算方法被提出,本文对这些计算方法进行了较为全面的研究与论述。

## 1 Chia-Pet 实验方案介绍

对端测序的结果存储在2个fastq文件中,可以使用Chia-Pet工具<sup>[17]</sup>或其他方法<sup>[18]</sup>进行处理。通常,Chia-Pet数据处理有7个步骤(见图1),分别是:连接子过滤;Pet映射;冗余去除;自连和互连Pet分类;结合位点分析自连Pet;用互连Pet进行染色质相互作用分析;染色质相互作用数据的可视化。

作者简介: 许立(1994-),男,硕士研究生,主要研究方向:机器学习、数据挖掘、生物信息学。

收稿日期: 2019-05-20

在第一步中,连接体将与参考半连接体核苷酸序列对齐。除标签序列外,有 2 种半连接体,分别命名为 A 和 B,而且具有相同的核苷酸。因此,根据连接体的组成将 PET 分为 2 类:相同的连接体(AA 或 BB)和不同的连接体(AB 或 BA)。然后将连接体从原始测序片段中排除,并保留剩余的 DNA 片段以供进一步分析。在连接体过滤后,使用 BWA<sup>[19]</sup>、Bowtie<sup>[20]</sup>、Batmis<sup>[21]</sup>或其他绘图工具将短 DNA 序列与参考基因组对齐。使用 samtools<sup>[22]</sup>和 bedtols<sup>[23]</sup>过滤掉冗余和低质量的映射序列。自连 PET 是指从两端循环的单个 DNA 片段的测序片段,并在同一染色体上的短距离内映射到基因组。互连 PET 是指来自不同 DNA 片段的测序片段,通常 2 个标签位于不同染色体中或长距离位于同一染色体中。虽然使用自连 PET 来确定基因组上的蛋白质结合位点,但是互连 PET 可以通过聚类来预测染色质相互作用。在此基础上还须确保 2 个结合位点之间的交互集群确实存在或者是偶然发生的。Li 等人<sup>[17]</sup>使用基于超几何分布的 Fisher 精确检验来量化相互作用频率。Paulsen 等人<sup>[18]</sup>提出了一种基于非中心超几何分布的新统计模型,该模型将基因组距离依赖关系考虑在内进行 p 值估计。最后,构建 Chia-Pet 浏览器来报告数据并可可视化结合位点以及交互集群。

通过数据处理获得的相互作用需通过湿实验室进行验证。短基因组距离中 DNA 元件间的相互作用可以通过 3C 实验验证。对于远距离相互作用中的 DNA 片段(位于不同染色体或同一染色体中的两个锚点,距离超过 100 万碱基对),可以使用显微镜技术,如 DNA 荧光原位杂交(DNA-FISH)<sup>[24]</sup>直接观察相互作用锚的位置和核中的相对空间距离。

## 2 Chia-Pet 数据分析方法

### 2.1 Chia-Pet Tool 介绍

Chia-Pet 工具(CPT)<sup>[17]</sup>是第一个解决与 Chia-Pet 数据相关的独特问题的软件,并建立了有效的数据处理工作流程。CPT 的研发者采用随机抽样和统计方法来滤除溶液中而非体内随机结扎产生的噪声。算法提出在染色质片段群中,每个染色质片段具有以随机方式与任何其他片段连接的相等机会,并且不同结合位点之间的相互作用彼此独立的假设。在该随机模型下,互连 PET 的数量遵循超几何分布。定义如下:

$$P(n_{ij} | n, n_i, n_j) = \frac{\binom{n_i}{n_{ij}} \binom{2n - n_i}{n_j - n_{ij}}}{\binom{2n}{n_i}} \quad (1)$$

然而,CPT 未能解决基于 3C 的技术中最常见的偏差来源,即由线性基因组距离引起的非特异性相互作用,即随机聚合物效应。CPT 使用超几何分布来评估相互作用的显着性。该模型假设任何 2 个基因组基因座同样可能由 PET 连接,而无论基因组距离如何。因此,除了检测真正的循环相互作用之外,该方法可能检测到由于基因组接近而非基因组循环所带来的非随机交互频率的大量假阳性。实际上,最近的一项研究表明,CPT 确定的相互作用表明与高分辨率 Hi-C 数据集的一致性较差。此外,CPT 非常难以安装,因为需要专门特定的操作系统进行配置,包括复杂的编程语言和环境,主要有 C、perl、python、R、MySQL、Apache Web 服务器和 PHP。

### 2.2 ChiaSig 介绍

Paulsen 等人描述了一种处理 ChiaPET 数据的新方法,方法中具体考虑了零模型中的基因组距离。该方法名为 ChiaSig<sup>[18]</sup>,使用非中心超几何来识别统计上显著的相互作用。该算法在 CPT 的基础上改进,并考虑结合位点 *i* 和 *j* 间相互作用的 Chia-Pet 数量,用 NCHG 分布代替超几何分布,定义如下:

$$P(n_{ij} | n, n_i, n_j, w_{ij}) = \frac{\binom{n_i}{n_{ij}} \binom{2n - n_i}{n_j - n_{ij}} w_{ij}^{n_{ij}}}{\sum_{n'_{ij}} \binom{n_i}{n'_{ij}} \binom{2n - n_i}{n_j - n'_{ij}} w_{ij}^{n'_{ij}}} \quad (2)$$

分析 Chia-Pet 数据的基本步骤见图 1。沿着基因组序列紧密相连的区域有更高的机会形成随机接触,因此有必要将基于不同序列距离的预期交互频率也纳入模型中。假设  $\lambda_{ij}$  为结合位点区域 *i* 和 *j* 之间的预期相互作用频率,根据 *i* 和 *j* 间的基因组距离来估计。研究时可参考如下数学公式:

$$w_{ij} = \frac{\lambda_{ij}(2\lambda - \lambda_i - \lambda_j + \lambda_{ij})}{(\lambda_i - \lambda_{ij})(\lambda_j - \lambda_{ij})} \quad (3)$$

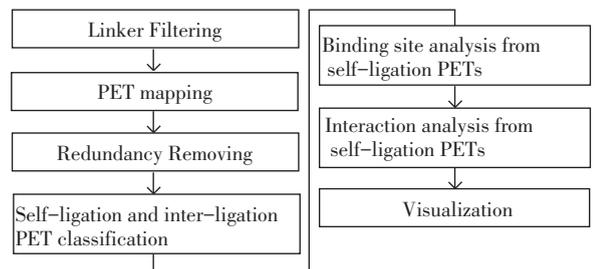


图 1 分析 Chia-Pet 数据的基本步骤

Fig. 1 Fundamental steps to analyze Chia-Pet data

正如预期的那样,这种方法识别的交互要比 CPT 少得多。虽然这种方法能够产生准确的交互,

但是软件只执行 Chia-Pet 数据分析、交互评分中的最后一步。因此,用户必须编写自己的软件来查找和删除链接器序列、对齐宠物、删除重复项、调用峰值、将宠物分组到交互中并确定宠物距离的下限。因此,该软件仅对具有重要编程技能的研究人员有用。已经描述了其他软件包,但这些软件包或者不公开,或者与 CPT 和 Chiasig 有类似的限制。

### 2.3 Mango 介绍

Mango<sup>[25]</sup> 将基因组位点间相互作用的可能性作为距离和峰深的函数进行建模,并使用该模型为相互作用分配统计置信度。值得注意的是,Mango 用一种简单而健壮的贝叶斯方法取代了计算上昂贵的距离匹配重布线方法。

由于使用方便和准确性的提高,Mango 将通过对 Chia-Pet 数据集的分析,大幅提升揭示三维染色质结构特征和功能的能力。同时也纠正了非特定的相互作用,可以作为一个基因组接近和峰深的函数。本次研究证明,与 CPT(现有的 Chia-Pet 分析管道)和 Chiasig(为 Chia-Pet 交互提供统计置信度估计的软件包)相比,Mango 表现出更高的准确性。将 Mango 应用于多个 Chia-Pet 数据集,可以独立复制与 NAT 相关的发现。三维染色质环的结构,包括对具有内向基序的 CTCF 结合位点的强富集。

除了提高准确性之外,Mango 的可用性也颇受青睐。Mango 被设计成所有的研究人员都可以使用。Mango 很容易安装,只需一个命令就可以完成从 fastq 到交互的所有步骤。

### 2.4 MICC 介绍

MICC<sup>[26]</sup>,一种易于使用的 R 包,用于处理 Chia-

Pet 数据。MICC 旨在以高灵敏度检测染色质相互作用,同时将错误发现率(FDR)控制在合理水平。MICC 的输入是源自 Chia-Pet 数据的原始 PET 簇。MICC 的最终输出包括:将 PET 簇描述为真实相互作用簇的后验概率列表和相应的 FDR。在不同数据集的相同 FDR 上,MICC 总能检测到比 Chia-Pet 工具和 Chiasig 更多的相互作用。此外,MICC 检测到的相互作用在生物学重复之间也更加一致。

### 2.5 各种工具之间的性能比较

Mango 仅依赖 4 个广泛使用且易于安装的软件包。相比之下,CPT 需要具体的操作系统配置,主要有复杂的编程语言和环境阵列,包括 C、Perl、Python、R、Mysql、Apache Web Server 和 Php,并附带 7 页的安装指南。Chiasig 可以轻松安装,但只执行分析 Chia PET 数据所需的单个步骤。因此,用户要编写自己的代码来执行大多数处理步骤,包括连接解析、PET 映射、冗余去除、峰值调用和距离过滤,详见表 1。

MICC,从 Chia-Pet 的数据中检测显著染色质相互作用。与 Chia-Pet 工具相比,MICC 使用较低深度的测序库恢复了较高深度测序库中检测到的交互作用的显著比例。同时,还为宠物集群提供了更一致的排序,从而可以提高实验复制之间的再现性。通过与 5C 数据的比较,分析后发现 MICC 能比 Chiasig 更有效地检测相互作用。此外,MICC 检测到的低 PET 计数的相互作用与 5C 数据有很大的重叠,这表明 MICC 寻找弱相互作用是可行的。这些特性使 MICC 优于其他现有的工具,特别是在以较少的排序深度处理 ChiaPET 数据时。

表 1 Chia-Pet 软件工具包功能的比较

Tab. 1 Comparison of Chia-Pet software packages

Software	连接过滤	PET 映射	冗余去除	峰值调用	自连区分	聚类	统计分析
CPT	√	√	√	√	√	√	√
Chiasig							√
Mango	√	√	√	√	√	√	√
MICC						√	√

## 3 Chia-Pet 技术应用

### 3.1 研究 DNA 片段之间的相互作用

Chip Seq 用于分析 DNA 和蛋白质之间的相互作用,而 Chia-Pet 则从根本上研究 DNA 片段之间的相互作用。Fullwood 等人<sup>[27]</sup> 使用 Chia-Pet 技术构建了由人乳腺癌细胞系 MCF7 的雌激素受体  $\alpha$ (ER- $\alpha$ ) 结合的染色质相互作用网络,发现长程

ER- $\alpha$ 结合位点主要位于启动子区域。Handoko 等人<sup>[28]</sup> 发现 CTCF 介导的小鼠胚胎多能干细胞相互作用。Chia-Pet 揭示的 5 个不同的染色质结构域为染色体结构组织提供了新的 CTCF 功能模型,并将增强子与基因转录调控的启动子连接起来。

在描述人类 T 细胞中增强子-启动子相互作用后,Chepelev 等人<sup>[29]</sup> 提出增强子以细胞特异性的方

式增加其靶基因的表达,相互作用的启动子是共存的。此外,细胞核中的染色体在多个层次上被组织起来发挥作用,除 CTCF 外,还有许多因素可能参与 T 细胞的这一过程。在未来的研究中,需要对详细的机制进行探讨。

He 等人<sup>[30]</sup>根据 Chip Seq 获得的 ER- $\alpha$  结合峰计算 DNA 环化的可能性,继而预测 ER- $\alpha$  介导的染色质相互作用。这是第一个使用 Chip Seq 预测染色质相互作用的工作,为 Chia-Pet 提供了补充。

### 3.2 构建染色质相互作用网络

与许多细胞网络一样,染色质交互网络<sup>[31]</sup>具有无标度和模块化拓扑结构,多数节点仅参与一个或两个交互,而一些节点与不成比例的大量节点连接。染色质相互作用网络被组织成“社区”,社区内的基因以协调方式执行相关功能并对外部刺激做出反应,意味着这些社区可能在数百万年进化过程中被塑造。

在未来的研究中,不仅可以将该方法应用于其他特定类型的基因,还可以将相互依赖的网络结合起来,因为细胞活动一起发生并且相互联系。此外,染色质相互作用网络可能奠定 3D 或甚至 4D 基因组波的基础,从静态转变为动态<sup>[31]</sup>。

### 3.3 染色质相互作用的功能研究

目前,已有多种方法用于研究 Chia-Pet 鉴定的染色质相互作用的功能,即:荧光素酶报告基因测定<sup>[11]</sup>、目的蛋白的表达水平敲定实验<sup>[11]</sup>、来自转基因实验的增强子测定法鉴定的调控元件<sup>[15]</sup>、基因组编辑方法(如锌指核酸酶基因组编辑,TALENs 和 CRISPR / Cas9)干扰染色质相互作用<sup>[16]</sup>。

### 3.4 染色质三维结构的重建

染色质的精确三维结构提供了更好的生物学功能景观。到目前为止,远距离相互作用的数据适合于重建三维基因组结构。2 个 3c 衍生物,即 hi-c<sup>[10]</sup> 和 Chia-Pet<sup>[9]</sup>,实际上反映了整个基因组的结构。Hi-C 技术可以捕获所有的交互,但是分辨率很低。Chia-Pet 技术大大提高了分辨率,但只能识别已知蛋白质介导的相互作用。因此,Chia-Pet 数据可用于进行更为密集的建模。

对染色质的三维结构进行建模主要有 2 种方法<sup>[32]</sup>。一种是物理模型,如用于解释实验结果的珠子串模型;另一种是用于重建结构的非线性优化模型。其中,物理模型方法中必须考虑许多物理性质。重建结构的非线性优化模型的第一步是将染色质相互作用频率转换为空间距离,基于此将空间距离转

换为三维结构。由于缺乏直接参数来评估在全基因组范围内建立的三维结构,电子显微镜的发展将在促进染色质三维结构的研究中发挥重要作用。染色质相互作用的可视化与功能测定结合是一种重要的方式,可以让人们对基因组结构有更直观的印象,并全面了解基因组的功能。

## 4 结束语

本文介绍了 Chia-Pet 的实验方案和数据分析过程,分析了几种常用工具的特点和适用范围,有助于研究中选用合适的方法以获得更可靠的结果。现已成功地应用于转录调控分析的许多研究中,并已鉴定出不同的染色质相互作用模型。尽管如此,在 Chia-Pet 协议和分析管道方面仍有亟待改进之处,使协议更加简洁和易于执行,数据分析过程更加自动化和可定制。

## 参考文献

- [1] REN B, ROBERT F, WYRICK J J, et al. Genome-wide location and function of DNA binding proteins [J]. *Science*, 2000, 290 (5500): 2306.
- [2] WEI C L, WU Q, VEGA V B, et al. A global map of p53 transcription-factor binding sites in the human genome [J]. *Cell*, 2006, 124(1): 207.
- [3] JOHNSON D S, MORTAZAVI A, MYERS R M, et al. Genome-wide mapping of in vivo protein-DNA interactions [J]. *Science*, 2007, 316(5830): 1497.
- [4] DEKKER J, RIPPE K, DEKKER M, et al. Capturing chromosome conformation [J]. *science*, 2002, 295(5558): 1306.
- [5] ZHAO Z, TAVOOSIDANA G, SJÖLINDER M, et al. Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions [J]. *Nature genetics*, 2006, 38(11): 1341.
- [6] SIMONIS M, KLOUS P, SPLINTER E, et al. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C) [J]. *Nature genetics*, 2006, 38(11): 1348.
- [7] DOSTIE J, RICHMOND T A, ARNAOUT R A, et al. Chromosome conformation capture carbon copy (5C): A massively parallel solution for mapping interactions between genomic elements [J]. *Genome research*, 2006, 16(10): 1299.
- [8] de WIT E, De LAAT W. A decade of 3C technologies: Insights into nuclear organization [J]. *Genes & development*, 2012, 26 (1): 11.
- [9] LIEBERMAN-AIDEN E, Van BERKUM N L, WILLIAMS L, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome [J]. *Science*, 2009, 326 (5950): 289.
- [10] LI G, RUAN X, AUERBACH R K, et al. Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation [J]. *Cell*, 2012, 148(1-2): 84.
- [11] CHEPELEV I, WEI G, WANGSA D, et al. Characterization of genome-wide enhancer-promoter interactions reveals co-expression of interacting genes and modes of higher order chromatin organization [J]. *Cell research*, 2012, 22(3): 490.

- [12] HANDOKO L, XU H, LI G, et al. CTCF-mediated functional chromatin interactome in pluripotent cells [J]. *Nature genetics*, 2011, 43(7): 630.
- [13] DOWEN J M, FAN Z P, HNISZ D, et al. Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes [J]. *Cell*, 2014, 159(2): 374.
- [14] KIEFFER-KWON K R, TANG Z, MATHE E, et al. Interactome maps of mouse gene regulatory domains reveal basic principles of transcriptional regulation [J]. *Cell*, 2013, 155(7): 1507.
- [15] PAPANTONIS A, KOHRO T, BABOO S, et al. TNF $\alpha$  signals through specialized factories where responsive coding and miRNA genes are transcribed [J]. *The EMBO Journal*, 2012, 31(23): 4404.
- [16] DEMARE L E, LENG J, COTNEY J, et al. The genomic landscape of cohesin-associated chromatin interactions [J]. *Genome research*, 2013, 23(8): 1224.
- [17] LI G, FULLWOOD M J, XU H, et al. ChIA-PET tool for comprehensive chromatin interaction analysis with paired-end tag sequencing [J]. *Genome biology*, 2010, 11(2): R22.
- [18] PAULSEN J, RØDLAND E A, HOLDEN L, et al. A statistical model of ChIA-PET data for accurate detection of chromatin 3D interactions [J]. *Nucleic acids research*, 2014, 42(18): e143.
- [19] LI H, DURBIN R. Fast and accurate short read alignment with Burrows - Wheeler transform [J]. *Bioinformatics*, 2009, 25(14): 1754.
- [20] LANGMEAD B, TRAPNELL C, POP M, et al. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome [J]. *Genome biology*, 2009, 10(3): R25.
- [21] TENNAKOON C, PURBOJATI R W, SUNG W K. BatMis: A fast algorithm for k-mismatch mapping [J]. *Bioinformatics*, 2012, 28(16): 2122.
- [22] LI H, HANDSAKER B, WYSOKER A, et al. The sequence alignment/map format and SAMtools [J]. *Bioinformatics*, 2009, 25(16): 2078.
- [23] QUINLAN A R, HALL I M. BEDTools: A flexible suite of utilities for comparing genomic features [J]. *Bioinformatics*, 2010, 26(6): 841.
- [24] LANGER-SAFER P R, LEVINE M, WARD D C. Immunological method for mapping genes on Drosophila polytene chromosomes [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 1982, 79: 4381.
- [25] PHANSTIEL D H, BOLE A P, HEI DARI N, et al. Mango: A bias-correcting ChIA-PET analysis pipeline [J]. *Bioinformatics*, 2015, 31(19): 3092.
- [26] HE C, ZHANG M Q, WANG X. MICC: An R package for identifying chromatin interactions from ChIA - PET data [J]. *Bioinformatics*, 2015, 31(23): 3832.
- [27] FULLWOOD M J, LIU M H, PAN Y F, et al. An oestrogen-receptor -  $\alpha$  - bound human chromatin interactome [J]. *Nature*, 2009, 462(7269): 58.
- [28] HANDOKO L, XU H, LI G, et al. CTCF-mediated functional chromatin interactome in pluripotent cells [J]. *Nature genetics*, 2011, 43(7): 630.
- [29] CHEPELEV I, WEI G, WANGSA D, et al. Characterization of genome-wide enhancer-promoter interactions reveals co-expression of interacting genes and modes of higher order chromatin organization [J]. *Cell research*, 2012, 22(3): 490.
- [30] HE C, WANG X, ZHANG M Q. Nucleosome eviction and multiple co-factor binding predict estrogen-receptor- $\alpha$ -associated long-range interactions [J]. *Nucleic acids research*, 2014, 42(11): 6935.
- [31] SANDHU K S, LI G, POH H M, et al. Large-scale functional organization of long-range chromatin interaction networks [J]. *Cell reports*, 2012, 2(5): 1207.
- [32] PENG C, LI G L, ZHANG H Y, et al. Reconstruction of three-dimensional structures of chromatin and its biological implications [J]. *Scientia Sinica Vitae*, 2014, 44(8): 794.

(上接第347页)

较高的关键词有“management(管理)”、“corporate governance(公司治理)”、“internal audit(内部审计)”,其次是“audit(审计)”、“performance(绩效)”、“innovation(创新)”。这些都是当下内部审计研究领域的热点。

### 3 结束语

综上所述,在2006~2018年间,内部审计、内部控制和公司治理是内部审计研究领域的重点,且已初步形成一定的研究规模,但相较于目前变化迅捷的市场环境来说还很薄弱,仍需要后续不断的探索研究与实践检验。

内部审计的完善依赖于客观准确的评价活动和高效的内部控制,因此,未来内部审计的研究必将与评价活动和内部控制进行紧密结合。随着人工智能与未来内部审计结合日趋成熟,届时内部审计的方

法、技术和工具等方面均将涌现更多的研究成果。

### 参考文献

- [1] 王光远, 瞿曲. 公司治理中的内部审计—受托责任视角的内部治理机制观 [J]. *审计研究*, 2006(2): 29.
- [2] SHIFFRIN R M, BRNER K. Mapping knowledge domains [J]. *Proceedings of the National Academy of the USA*, 2004, 101(1): 5183.
- [3] 刘则渊, 王贤文. 生态经济学研究前沿及其演进的可视化分析 [J]. *西南林学院学报*, 2008, 28(4): 3.
- [4] 丁涛. 企业内部审计与内部控制关系探讨 [J]. *财会通讯*, 2016(16): 122.
- [5] 杨雄胜. 内部控制理论研究新视野 [J]. *会计研究*, 2005(7): 49.
- [6] 吴水澎, 陈汉文, 邵贤弟. 企业内部控制理论的发展与启示 [J]. *会计研究*, 2000(5): 2.
- [7] 李连华. 公司治理结构与内部控制的链接与互动 [J]. *会计研究*, 2005(2): 64.
- [8] 房宏君. 国内科技人才研究的来源分析和热点分析 [J]. *人力资源管理*, 2011(12): 120.