



(12) 发明专利申请

(10) 申请公布号 CN 114787378 A

(43) 申请公布日 2022. 07. 22

(21) 申请号 202080069743.8

(74) 专利代理机构 北京市中咨律师事务所
11247

(22) 申请日 2020.10.05

专利代理师 胡志君 黄革生

(30) 优先权数据

1914325.4 2019.10.04 GB

(51) Int.Cl.

C12Q 1/6806 (2018.01)

(85) PCT国际申请进入国家阶段日

C12Q 1/6869 (2018.01)

2022.04.01

C12Q 1/6809 (2018.01)

(86) PCT国际申请的申请数据

C12Q 1/6813 (2018.01)

PCT/GB2020/052448 2020.10.05

(87) PCT国际申请的公布数据

W02021/064430 EN 2021.04.08

(71) 申请人 巴布拉罕姆研究所

地址 英国剑桥郡

(72) 发明人 V·马利舍瓦 S·舍恩菲尔德

M·斯皮瓦科夫 长野卓

P·弗雷泽

权利要求书2页 说明书22页

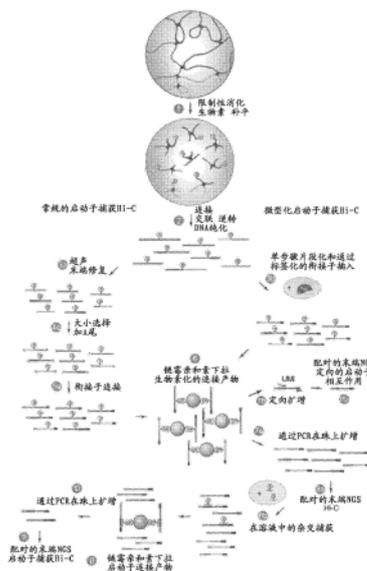
序列表4页 附图2页

(54) 发明名称

新方法

(57) 摘要

本发明涉及一种用于鉴定与一个靶核酸区段或多个靶核酸区段相互作用的核酸区段的方法以及实施该方法的试剂盒。本发明还涉及鉴定一个或多个指示特定疾病的相互作用性核酸区段的方法。



1. 一种用于鉴定与一个靶核酸区段或多个靶核酸区段相互作用的核酸区段的方法,所述方法包括以下步骤:

- (a) 获得包含该一个靶核酸区段或多个靶核酸区段的核酸组合物;
- (b) 交联该核酸组合物;
- (c) 使用核酸内切酶使该交联的核酸组合物片段化;
- (d) 用一个或多个包含共价连接的生物素部分的核苷酸填充该片段化的交联核酸区段的末端;
- (e) 连接从步骤(d)获得的该片段化核酸区段以产生已连接片段;
- (f) 使用重组酶对该已连接片段实施单步骤片段化和寡核苷酸插入;
- (g) 富集包含步骤(d)的生物素部分的片段;
- (h) 富集包含该一个靶核酸区段或多个靶核酸区段的片段;
- (i) 对步骤(h)中获得的该富集的片段测序,以鉴定与该一个靶核酸区段或多个靶核酸区段相互作用的核酸区段。

2. 根据权利要求1所述的方法,其中步骤(h)包括进行定向扩增以富集包含该一个靶核酸区段或多个靶核酸区段的片段。

3. 根据权利要求1所述的方法,其中步骤(h)包括:

(i) 添加与该一个靶核酸区段或多个靶核酸区段结合的分离子核酸分子,其中用结合对子的第一半部分标记所述分离子核酸分子;并且

(ii) 通过使用所述结合对子的第二半部分,分离出含有与分离子核酸分子结合的该一个靶核酸区段或多个靶核酸区段的片段。

4. 根据权利要求1至3中任一项所述的方法,其中通过标签化进行步骤(f)。

5. 根据权利要求1至4中任一项所述的方法,其中步骤(e)利用胞核内连接。

6. 根据权利要求1至5中任一项所述的方法,其中重组酶是逆转录病毒整合酶,如突变型转座酶,如超活性Tn5转座酶。

7. 根据权利要求1至6中任一项所述的方法,其中重组酶包含测序用配对末端衔接子序列或其片段。

8. 根据权利要求1至7中任一项所述的方法,其中所述寡核苷酸和/或衔接子序列包含条形码序列。

9. 根据权利要求7或权利要求8所述的方法,其中所述寡核苷酸和/或衔接子序列选自: SEQ ID NO:1、SEQ ID NO:2和/或SEQ ID NO:3,或是使后续文库制备和测序成为可能的寡核苷酸序列。

10. 根据权利要求1至9中任一项所述的方法,其中在步骤(h)添加分离子核酸分子是在借助衔接子序列的互补性防止已连接片段与其他已连接片段结合的序列如阻断物序列存在下进行的。

11. 根据权利要求1至10中任一项所述的方法,其中所述一个靶核酸区段或多个靶核酸区段选自:启动子、沉默子、增强子或绝缘子。

12. 根据权利要求1至11中任一项所述的方法,其中从细菌人工染色体(BAC)、F粘粒或粘粒获得分离子核酸分子。

13. 根据权利要求1或3至12中任一项所述的方法,其中分离子核酸分子是RNA。

14. 根据权利要求3至13中任一项所述的方法,其中结合对子的第一半部分包含生物素并且结合对子的第二半部分包含链霉亲和素。

15. 根据权利要求1至14中任一项所述的方法,其中步骤(c)所使用的限制性酶是HindIII或Dpn II。

16. 根据权利要求3至15中任一项所述的方法,其额外地包括在步骤(g)扩增包含生物素部分的富集片段。

17. 根据权利要求1至16中任一项所述的方法,其额外地包括在步骤(i)之前扩增分离的已连接片段。

18. 根据权利要求1、2或4至17中任一项所述的方法,其中通过PCR进行定向扩增或扩增。

19. 根据权利要求1至18中任一项所述的方法,其中所述核酸组合物衍生自哺乳动物细胞核,如人细胞核。

20. 根据权利要求1至19中任一项所述的方法,其中所述核酸组合物衍生自非人细胞核,如小鼠细胞核或植物细胞核。

21. 根据权利要求1至20中任一项所述的方法,其中所述核酸组合物衍生自10000、50000、20万、50万个或1百万个细胞。

22. 鉴定一个或多个指示特定疾病状态的相互作用性核酸区段的方法,包括:

(a) 对从患有特定疾病的个体获得的核酸组合物实施权利要求1至21中任一项所述的方法;

(b) 对核酸区段与一个靶核酸区段或多个靶核酸区段之间相互作用的频率定量;和

(c) 将来自具有所述疾病状态的个体的该核酸组合物中的相互作用频率与来自健康受试者的正常对照核组合物中的相互作用频率进行比较,从而核酸组合物中相互作用频率的差异指示特定疾病。

23. 根据权利要求22所述的方法,其中所述疾病状态选自:癌症、自身免疫性疾病、发育性疾病或遗传病。

24. 用于鉴定与一个靶核酸区段或多个靶核酸区段相互作用的核酸区段的试剂盒,包含能够实施根据权利要求1至21中任一项所述方法的缓冲剂和试剂。

25. 根据权利要求24所述的试剂盒,其中所述重组酶是逆转录病毒整合酶或转座酶,如突变型转座酶,如超活性Tn5转座酶。

新方法

技术领域

[0001] 本发明涉及一种用于鉴定与一个靶核酸区段或多个靶核酸区段相互作用的核酸区段的方法以及实施该方法的试剂盒。本发明还涉及鉴定一个或多个指示特定疾病的相互作用性核酸区段的方法。

背景技术

[0002] 调节元件在生物的遗传控制中发挥核心作用并且已经显示其对健康和疾病(例如在癌症和自体免疫疾病中)有贡献。已经展示这类调节元件(例如,增强子)可以距它们的靶基因在相当大的基因组距离(按线性标度)存在。已经开发了捕获这些调节元件及其靶基因的方法并且这些方法广泛用来研究影响关于基因表达和表型建立的调节性场景动力学(regulatory landscape dynamics),以及用来研究疾病形成中基因修饰的作用。然而,采用低细胞数量作业时,确定这些调节元件调节哪些靶基因提出了重大挑战。

[0003] 为鉴定基因组座位之间相互作用所开发的首类方法之一是染色体构象捕获(3C)技术(Dekker等人,Science (2002) 295:1306-1311)。这种方法涉及通过以下方式创建3C文库:交联核组合物,从而空间紧密邻近的基因组座位变得连接;通过消化过程移除交联之间的间插性DNA环;并且连接并逆转相互作用性区域的交联,以生成3C文库。该文库随后可以用来检测/鉴定已知序列之间相互作用的频率。然而,这种方法要求事先知晓相互作用,以便检测相互作用性目的区域。从那时起,已经进一步发展该技术以克服伴随3C方法的局限性。

[0004] Hi-C是不需要关于目标相互作用组(interactome of interest)的任何现有知识的全基因组方法。这种方法使用接界标志物(junction markers)来分离细胞中全部连接的相互作用性序列(参见W0 2010/036323和Lieberman-Aiden等人,2009)。尽管这提供了核组合物中特定时间点处存在的全部相互作用的信息,但所得到的文库是极端复杂的,这妨碍以鉴定特定元件(如启动子和增强子)之间明显相互作用所要求的分辨率分析它们。为了克服这种限制,已经开发了捕获Hi-C技术,其涉及富集Hi-C文库用于染色体相互作用的捕获步骤,所述文库包含至少在一个末端处的目的区域(参见W0 2015/033134、Dryden等人,2014和Schoenfelder等人,2015)。W0 2015/033134公开了一种通过使用分离用核酸分子,鉴定与该一个靶核酸区段或多个靶核酸区段相互作用的核酸区段的方法和试剂盒。然而,这种方法要求始于庞大数目的细胞(3-4千万个细胞),这在采用罕见细胞类型、来自生物早期发育阶段的细胞或患者/活检样品作业时不可能。

[0005] 因此需要提供一种克服现有可用方法学局限性的鉴定核酸相互作用的改进方法。

[0006] 发明概述

[0007] 根据本发明的第一方面,提供一种用于鉴定与一个靶核酸区段或多个靶核酸区段相互作用的核酸区段的方法,所述方法包括以下步骤:

[0008] (a) 获得包含该一个靶核酸区段或多个靶核酸区段的核酸组合物;

[0009] (b) 交联该核酸组合物;

- [0010] (c) 使用核酸内切酶使该交联的核酸组合物片段化；
- [0011] (d) 用一个或多个包含共价连接的生物素部分的核苷酸填充该片段化的交联核酸区段的末端；
- [0012] (e) 连接从步骤(d)获得的该片段化核酸区段以产生已连接片段；
- [0013] (f) 使用重组酶对该已连接片段实施单步骤片段化和寡核苷酸插入；
- [0014] (g) 富集包含步骤(d)的生物素部分的片段；
- [0015] (h) 富集包含该一个靶核酸区段或多个靶核酸区段的片段；
- [0016] (i) 对步骤(h)中获得的该富集的片段测序,以鉴定与该一个靶核酸区段或多个靶核酸区段相互作用的核酸区段。
- [0017] 根据本发明的又一个方面,提供一种鉴定一个或多个指示特定疾病状态的相互作用性核酸区段的方法,所述方法包括:
- [0018] (a) 对从患有特定疾病的个体获得的核酸组合物实施如本文所述的方法；
- [0019] (b) 对核酸区段与一个靶核酸区段或多个靶核酸区段之间相互作用的频率定量；
- 和
- [0020] (c) 将来自具有所述疾病状态的个体的该核酸组合物中的相互作用频率与来自健康受试者的正常对照核组合物中的相互作用频率进行比较,从而核酸组合物中相互作用频率的差异指示特定疾病。
- [0021] 根据本发明的再一个方面,提供一种用于鉴定与该一个靶核酸区段或多个靶核酸区段相互作用的核酸区段的试剂盒,包含能够实施本文所述方法的缓冲液和试剂。
- [0022] 附图简述
- [0023] 图1:与本文展示的微型化启动子捕获Hi-C方案相比的常规方案的对比性示意概览。缩略语:Tn5-重组酶/转座酶,B-生物素,NGS-下一代测序,UMI-唯一分子标识符(unique molecular identifier)。
- [0024] 图2:使用如本文所述的方法获得的结果。试样-人CD4+T细胞为50000个细胞,和1百万个细胞起始材料(分别是常规Hi-C方案中所用起始材料的1/600和1/30)中使用ChiCAGO时调出的显著的相互作用。
- [0025] 发明详述
- [0026] 根据本发明的第一方面,提供一种用于鉴定与一个靶核酸区段或多个靶核酸区段相互作用的核酸区段的方法,所述方法包括以下步骤:
- [0027] (a) 获得包含该一个靶核酸区段或多个靶核酸区段的核酸组合物；
- [0028] (b) 交联该核酸组合物；
- [0029] (c) 使用核酸内切酶使该交联的核酸组合物片段化；
- [0030] (d) 用一个或多个包含共价连接的生物素部分的核苷酸填充该片段化的交联核酸区段的末端；
- [0031] (e) 连接从步骤(d)获得的该片段化核酸区段以产生已连接片段；
- [0032] (f) 使用转座酶对该已连接片段实施单步骤片段化和寡核苷酸插入；
- [0033] (g) 富集包含步骤(d)的生物素部分的片段；
- [0034] (h) 富集包含该一个靶核酸区段或多个靶核酸区段的片段；
- [0035] (i) 对步骤(h)中获得的该富集的片段测序,以鉴定与该一个靶核酸区段或多个靶

核酸区段相互作用的核酸区段。

[0036] 本发明的方法提供一种通过使用分离一个靶核酸区段或多个靶核酸区段的定向扩增或分离用核酸分子来鉴定相互作用性序列和核酸区段的手段。这类方法具有使得数据聚焦于极复杂文库中特定相互作用上的优点。另外,包括定向扩增或添加与一个靶核酸区段或多个靶核酸区段结合的分​​离用核酸分子的方法也可以用来根据为选择所用的试剂的类型或所用的分离用核酸分子的类型(例如鉴定启动子相互作用的启动子),将信息组织成多个子集。随后可以获得关于目标靶特定群组中染色体相互作用的详细信息。

[0037] 本发明的方法进一步提供使用重组酶的单步骤片段化和寡核苷酸插入。这种单步骤片段化和寡核苷酸插入具有提供以下方法的优点,所述方法具有显著更少的总体步骤和减少的核酸组合物操纵。例如,在本发明方法的具体实施方案中,可以从获得核酸组合物(如本文所述的步骤(a))至连接片段化的核酸区段(如本文所述的步骤(e))以及还从进行单步骤片段化和寡核苷酸插入(如本文所述的步骤(f))至富集包含生物素的片段(如本文所述的步骤(g)),使用单管。不包括单步骤片段化和寡核苷酸插入的方法需要借助物理或酶促手段(例如超声处理或限制性酶消化)进行分别的片段化、文库片段末端修复、在文库片段3'末端添加dATP、大小选择、连接寡核苷酸序列及从未连接的寡核苷酸纯化片段(如在WO 2015/033134中公开的那些方法)。因此,将可以理解,本发明提供用于鉴定与该一个靶核酸区段或多个靶核酸区段相互作用的核酸区段的方法,所述方法更简单,包括更少步骤并且可以包含较短的完成时间框。因此,本发明的方法比常规方案更快并且降低产生文库的总费用。还将理解,这类优点可以导致核酸组合物损耗减少。这种减少的核酸组合物损耗允许起始材料的量(例如,从中获得核酸组合物的细胞的数目)减少,或可用于后续分析的所得核酸组合物增加。

[0038] 另外,尽管先前技术如4C允许捕获一个或一些启动子的全基因组相互作用,但本文所述的方法可以在单一实验中捕获超过22000个启动子及它们的相互作用性基因组座位。另外,本发明方法产生显著更多的定量性读段。

[0039] 全基因组关联研究(GWAS)已经鉴定数千个与疾病连锁的单核苷酸多态性(SNP)。然而,这些SNP中许多SNP距基因很远距离存在,这使预测它们作用于哪些基因非常有挑战性。因此,本发明方法提供了鉴定相互作用性核酸区段的方式,即使它们在基因组中彼此远距离存在。

[0040] 如本文所用对“核酸区段”的提及等同于对“核酸序列”的提及,并且指核苷酸(即例如,腺嘌呤(A)、胸苷(T)、胞嘧啶(C)、鸟苷(G)和/或尿嘧啶(U))的任何聚合物。这种聚合物可以产生或不产生有功能的基因组片段或基因。核酸序列组合最终可以包含染色体。包含脱氧核糖核苷的核酸序列称作脱氧核糖核酸(DNA)。包含核糖核苷的核酸序列称作核糖核酸(RNA)。RNA可以进一步表征成数个类型,如编码蛋白质的RNA、信使RNA(mRNA)、转运RNA(tRNA)、非编码性长RNA(lncRNA)、基因间非编码性长RNA(lincRNA)、反义RNA(asRNA)、微RNA(miRNA)、短干扰性RNA(siRNA)、核小RNA(snRNA)和核仁小RNA(snoRNA)。

[0041] “单核苷酸多态性”或“SNP”是基因组中在生物学物种成员之间或配对染色体之间有差异的单核苷酸变异(即A、C、G或T)。

[0042] 应当理解,术语“靶核酸区段”指用户已知的一个目标序列或多个目标序列。仅分离含有该靶核酸区段的已连接片段有助聚拢数据以鉴定与特定基因或目的基因区段的特

异性相互作用。备选地,进行定向扩增以富集包含该一个靶核酸区段或多个靶核酸区段的片段有助于聚拢数据以通过增加组合物中包含该一个靶核酸区段或多个靶核酸区段的片段的比例,鉴定与特定基因或目的基因区段的特异性相互作用。

[0043] 本文中对术语“相互作用”或“相互作用性”的提及,指两个元件之间的关联,例如在本发明的方法中,指核酸区段与靶核酸区段之间的基因组相互作用。相互作用可以引起一个相互作用性元件影响另一元件,例如沉默或激活与之结合的元件。相互作用可以发生在线性基因组序列上接近或远隔的两个核酸区段之间。因此,在一个实施方案中,与该一个靶核酸区段或多个靶核酸区段相互作用的该一个核酸区段或多个核酸区段在线性基因组序列上与所述一个靶核酸区段或多个靶核酸区段紧邻,例如在相同染色体上彼此相对地接近。在又一个实施方案中,与该一个靶核酸区段或多个靶核酸区段相互作用的该一个核酸区段或多个核酸区段在线性基因组序列上与所述靶核酸区段远隔,例如,存在于不同的染色体上;或者,如果在相同的染色体上,则进一步远离。

[0044] 本文中对术语“核酸组合物”的提及,指包含核酸和蛋白质的任何组合物。核酸组合物中的核酸可以组织成染色体,其中蛋白质(即例如,组蛋白)可以与具有调节功能的染色体结合。在一个实施方案中,核酸组合物包含核组合物。这种核组合物一般可以包含基因组组织结构或染色质。

[0045] 如本文所用对“交联性”或“交联”的提及,指两种化合物之间的任何稳定化学结合,从而它们可以作为一个单元进一步加工。这种稳定性可以基于共价键合和/或非共价键合(例如离子性键合)。例如,核酸和/或蛋白质可以通过化学物质(即例如,固定剂)、热、压力、pH变化或辐射交联,从而它们在常规实验室程序(即例如,提取,洗涤,离心等)期间维持它们的空间关系。如本文所用的交联等同于术语“固定”或“固定法”,其适用于使任何和全部细胞过程固定的任何方法或过程。交联/固定的细胞因此准确地在固定时维持核酸组合物中各组分之间的空间关系。许多化学物质能够提供固定,所述化学物质包括但不限于甲醛、福尔马林或戊二醛。

[0046] 如本文所用对术语“片段”的提及,指比衍生该片段的序列更短的任何核酸序列。片段可以具有任何大小,范围从数兆碱基和/或数千碱基至仅几个核苷酸长度。片段适当地在长度上大于5个核苷酸碱基,例如长10、15、20、25、30、40、50、100、250、500、750、1000、2000、5000或10000个核苷酸碱基。片段可以甚至更长,例如长1、5、10、20、25、50、75、100、200、300、400或500千个核苷酸碱基。诸如限制性酶消化、超声处理、酸孵育、碱孵育、微流化等方法均可以用来使核酸组合物片段化。

[0047] 在一些实施方案中,使用核酸内切酶进行片段化(即步骤(c))。合适的核酸内切酶实例包括但不限于序列特异性核酸内切酶,如限制性酶,和非序列特异性核酸内切酶,如MNase或DNA酶。

[0048] 因此,在一个实施方案中,核酸内切酶是序列特异性核酸内切酶,如限制性酶。如本文所用的术语“限制性酶”指在特定碱基对序列处切割核酸的任何蛋白质。取决于所选限制性酶的类型,切割可以产生平末端或粘末端。限制性酶的实例包括但不限于Eco RI、Eco RII、Bam HI、Hind III、Dpn II、Bgl II、Nco I、Taq I、Not I、Hinf I、Sau 3A、Pvu II、Sma I、Hae III、Hga I、Alu I、Eco RV、Kpn I、Pst I、Sac I、Sal I、Sca I、Spe I、Sph I、Stu I、Xba I。在又一个实施方案中,使用限制性酶进行片段化(即步骤(c))。在一个实施方案中,

限制性酶是Hind III。在又一个实施方案中,限制性酶是Dpn II。

[0049] 在一个备选实施方案中,核酸内切酶是非序列特异性核酸内切酶。如本文所用的术语“非序列特异性核酸内切酶”指切割核酸并且不限于所述核酸之序列的任何蛋白质,例如它们可以在其中不结合蛋白质(例如核小体和/或转录因子)的任何区域切割核酸。非序列特异性核酸内切酶的实例是本领域已知的并且包括但不限于DNA酶、RNA酶和MNase。MNase是衍生自细菌金黄色葡萄球菌(*Staphylococcus aureus*)的非特异性核酸内切-外切酶,其结合并且切割染色质上未结合蛋白质的DNA区域-与组蛋白或其他结合染色质的蛋白质结合的DNA依旧未遭消化。在又一个实施方案中,使用非序列特异性核酸内切酶进行片段化(即步骤(c))。

[0050] 在另一实施方案中,使用超声处理进行片段化(即步骤(c))。

[0051] 本文中术语“填充片段末端或核酸区段末端”的提及,指在片段化后添加核苷酸至交联的核酸组合物或区段的3'末端。这种填充包括添加dATP、dCTP、dGTP和/或dTTP核苷酸至核酸组合物或区段的3'末端。为了允许富集已经连接并且因此含有连接接界的核酸片段或区段,如本文所述的一个或多个用于填充的核苷酸可以包含共价连接的生物素部分。因此,在一个实施方案中,填充该片段化的交联核酸区段的末端包括添加生物素部分至交联的核酸片段的末端。在又一个实施方案中,填充该片段化的交联核酸区段的末端包括用“接界标志物”“标记”交联的核酸片段的末端。这种“标记”末端或添加生物素部分至交联的核酸片段的末端允许后续选择或富集已经根据步骤(e)和如本文所述的方法连接的核酸片段和/或区段。

[0052] 接界标志物允许在富集步骤(h)之前纯化已连接片段,因此确保仅富集连接的序列,而不富集未连接的(即非相互作用性)片段。

[0053] 在某些实施方案中,接界标志物包含标记的核苷酸接头(即包含共价连接的生物素部分的核苷酸)。在又一个实施方案中,接界标志物包含生物素。在一个实施方案中,接界标志物可以包含修饰的核苷酸。在一个实施方案中,接界标志物可以包含寡核苷酸接头序列。

[0054] 本文中对术语“连接的”或“连接”的提及,指两个核酸区段的任何键接,其通常包含磷酸二酯键。该键接通常在辅因子试剂和能量源(即例如,三磷酸腺苷(ATP))存在下受催化酶(即例如,连接酶如T4 DNA连接酶)的存在促进。在本文所述的方法中,为了产生单个已连接片段,将两个已经交联的核酸区段的片段连接在一起。

[0055] 在一个实施方案中,连接片段化核酸区段以产生已连接片段(即步骤(e))利用了胞核内连接。因此,在某些实施方案中,通过胞核内连接进行片段化核酸区段的连接。这种胞核内连接(in-nucleus ligation)具有可以使用小体积试剂的优点,这导致核酸组合物损耗减少,并且因此还可以允许起始材料的量减少。例如,可以减少从中获得核酸组合物的细胞的数目,或可以增加可用于后续分析的所得核酸组合物。

[0056] 本文中对“单步骤片段化和寡核苷酸插入”的提及,指在单步骤中使已连接片段片段化和插入寡核苷酸序列。这类方法利用与寡核苷酸序列结合并且将这些序列插到片段上的重组酶。这个过程又称作“标签化(tagmentation)”。因此,在一个实施方案中,单步骤片段化和寡核苷酸插入包括标签化。

[0057] 对于上文提到的那些优点,单步骤片段化和寡核苷酸连接的优点进一步包括不需

要从未连接片段移除已经并入核酸组合物中的任何结合对子元件(如生物素),不需要对已连接片段进行大小选择,借助重组酶的酶促片段化消除了末端修复的必要,因为未进行超声处理,并且不必进行A-尾添加。另外,插入寡核苷酸和/或衔接子(adapter)序列(可以包括条形码序列和/或唯一分子标识符)与片段化同时进行。这类条形码序列或唯一分子标识符可以允许在后续分析和加工中鉴定特定核酸组合物并且允许多个核酸组合物在后续步骤中合并,同时保留鉴定和分析各个核酸组合物的能力。因此,在一个实施方案中,寡核苷酸序列是“衔接子序列”,其允许后续文库制备和含有衔接子的核酸片段测序或使之成为可能。在又一个实施方案中,该衔接子包含条形码序列和/或唯一分子标识符。

[0058] 在又一个实施方案中,单步骤片段化和寡核苷酸插入包括向已连接片段插入条形码序列。在一个实施方案中,配对的末端衔接子序列包含条形码序列和/或唯一分子标识符。

[0059] 如本文展示的利用单步骤片段化和寡核苷酸连接(例如标签化)的本发明方法的再一个优点是与先前的公开方案相比,获得显著富集的包含该一个靶核酸区段或多个靶核酸区段的片段文库。例如,与根据先前已知或常规Hi-C方案产生的文库相比,可以生成在至少5倍和20倍之间或在至少5倍和80倍之间的富集值。在一个实施方案中,与根据常规Hi-C方案生成的文库相比,可以根据本文所述的方法生成富集至少5倍、至少10倍、至少15倍或至少20倍的文库。在又一个实施方案中,可以根据本文所述的方法生成富集至少10倍、至少11倍、至少12倍、至少13倍、至少14倍或至少15倍的文库。在又一个实施方案中,可以根据本文所述的方法生成富集至少50倍、至少55倍或至少60倍的文库。将可以理解,与根据常规Hi-C方案生成的文库相比,实施如本文所述的方法时获得的任何富集值可以取决于用于使交联的核酸组合物片段化的核酸内切酶的身份。例如,当使用限制性酶Hind III时,可以获得高达20倍的富集值。备选地,当使用限制性酶Dpn II时,可以获得高达80倍的富集值。

[0060] 如本文所用的术语“配对末端衔接子”指允许自动化高通量测序从两个末端读取的任何引物对组。例如,与这些衔接子兼容的这类高通量测序装置包括但不限于Solexa (Illumina)、454System和/或ABI SOLiD。例如,该方法可以包括使用与聚腺苷酸尾连接的通用引物。

[0061] 将理解,本发明方法中适用的重组酶包括能够移除(或切割)序列并且将其插入寡核苷酸或核酸片段中的任何酶。这类重组酶的实例包括逆转录病毒整合酶和转座酶如MuA、Tn5、Tn7和Tc1/水手型转座酶。因此,在本发明方法的一个实施方案中,重组酶是逆转录病毒整合酶。在又一个实施方案中,重组酶是转座酶,如Tn5转座酶。对于重组酶,整合酶或转座酶在本文展示的方法中有活性,可以将酶突变以克服这类酶天然存在的低活性水平。因此,在又一个实施方案中,重组酶是突变型转座酶,如超活性转座酶。这种超活性转座酶可以是突变型Tn5转座酶。在一个实施方案中,重组酶是Tn5转座酶,如超活性Tn5转座酶。

[0062] Tn5转座酶是重组酶蛋白的RNA酶超家族成员,该成员包括逆转录病毒整合酶并且通过所谓“切割和黏贴”机理,催化核酸的一部分(称作转座子)移动至基因组的另一个部分或另一个基因组。重组酶如转座酶和转座子元件可见于某些细菌中,并且参与获得抗生素耐药性。转座酶通常无活性并且在该蛋白质的活性部位或他处突变可以导致超活性酶产生。本领域已知产生Tn5转座酶的方法(Picelli等人,(2014) Genome Research 24:2033-2040)。然而,当纯化Tn5转座酶时,可以通过使用寡核苷酸序列,如衔接子序列,进一步改造

这些方法。

[0063] 纯化重组酶(例如Tn5转座酶)时所用的寡核苷酸序列(如衔接子序列)随着该酶并入并且随后由所述重组酶插入核酸片段或区段中。这类序列可以具有序列多样性并且包含额外元件,所述元件使得进一步加工插入了这类序列的核酸片段或区段成为可能。例如,用已纯化重组酶并入的寡核苷酸可以包含测序用衔接子序列和/或条形码序列。然而,将理解,全部这类寡核苷酸均包含允许用该酶并入的转座子序列或元件。转座子序列或元件的实例包括Tn5转座酶相容性镶嵌式末端(Mosaic End,ME)序列和空间上与重组酶和/或转座酶的结合袋兼容的序列

[0064] 因此,根据一个实施方案,该方法的重组酶包含镶嵌式末端双链(MEDS)寡核苷酸,所述寡核苷酸包含配对末端衔接子序列的半部分。在又一个实施方案中,重组酶包含测序用配对末端衔接子序列。在进一步的其他实施例中,转座酶可以包含寡核苷酸,所述寡核苷酸包含测序用配对末端衔接子序列,额外地包含条形码序列。在其他实施方案中,寡核苷酸序列选自:如本文所述的SEQ ID NO:1、SEQ ID NO:2和/或SEQ ID NO:3。在一个备选实施方案中,寡核苷酸序列包含使后续文库制备和测序成为可能的任何序列。将理解这类序列使扩增和分离核酸区段以及结合所述核酸区段以便通过高通量或下一代测序法分析序列成为可能。下一代测序平台的实例包括:Roche 454(即Roche 454GS FLX)、Applied Biosystems的SOLiD系统(即SOLiDv4)、Illumina的GAIIx、HiSeq 2000和MiSeq测序仪、Life Technologies的基于Ion Torrent半导体的测序仪、Pacific Biosciences的PacBio RS和Oxford Nanopore的MinION。

[0065] 本文中对“进行富集”或“富集”的提及,指对核酸区段的任何分离或目的核酸区段或靶核酸区段的比例相对于核酸组合物中的其他核酸区段增加。将可以理解,这类称谓包括术语“进行分离”、“分离”、“分开”、“移除”、“纯化”等。例如,富集或分离目的核酸区段或靶核酸区段可以包括正向方法,如“分出”目的核酸区段或靶核酸区段,或可以包括负向方法,如排除不作为目的的核酸区段或排除不包含靶核酸区段的核酸区段。备选地,富集或分离可以包括选择性或定向扩增目的核酸区段或靶核酸区段。这类选择性或定向扩增目的核酸区段或靶核酸区段将增加这类区段在核酸组合物中的比例(即富集所述区段)。

[0066] 在一个实施方案中,所述富集步骤(h)包括步骤:进行定向扩增以富集包含该一个靶核酸区段或多个靶核酸区段的片段。

[0067] 在一个备选实施方案中,所述富集步骤(h)包括以下步骤:

[0068] (i) 添加与靶核酸区段结合的分离用核酸分子,其中用结合对子的第一半部分标记所述分离用核酸分子;并且

[0069] (ii) 通过使用所述结合对子的第二半部分,分离出含有与分离用核酸分子结合的靶核酸区段的片段,

[0070] 以富集包含该一个靶核酸区段或多个靶核酸区段的片段。在某些实施方案中,上述步骤(i)和(ii)可以依次进行,即,进行步骤(i),而随后进行步骤(ii)。在其他实施方案中,上述步骤(i)和(ii)可以同步进行。

[0071] 因此,本发明方法的富集步骤(h)包括富集包含特定靶区段或序列的核酸片段或目的区段或靶核酸区段。

[0072] 本文中对“定向扩增”的提及,指使用优先扩增特定目的核酸区段或靶核酸区段的

方法扩增。这种定向扩增可以利用与靶核酸区段中存在的靶核酸区段或序列(例如启动子或沉默子序列)互补的特定引物序列。因此,在一个实施方案中,引物序列与启动子序列互补。在另一个实施方案中,引物序列与包含SNP的序列互补。在本文展示的方法中使用的引物序列可以包含参与后续加工或分析已扩增核酸区段的额外元件。例如,引物序列可以包含如本文所述的衔接子序列用于测序或可用于鉴定该一个核酸区段或一组核酸区段(例如衍生自特定样品的那些区段)的唯一分子标识符。备选地或额外地,定向扩增可以利用可能有利于扩增靶核酸区段或包含该靶核酸区段的片段的特定条件。将可以理解,可以通过本领域已知的任何方法,如聚合酶链反应(PCR)进行扩增。进一步可以理解,如本文所述的定向扩增可以包括扩增溶液中或用于富集的支持物部分(如珠)上的核酸区段。也可以在扩增之前进行引物序列延伸,从而支持物部分上的扩增可以额外地包括在扩增所述延伸的序列和核酸区段之前延伸引物序列的步骤。

[0073] 本文中对“分离用核酸分子”的提及,指由下述核酸形成的分子,所述核酸配置成与靶核酸区段结合。例如,分离用核酸分子可以含有针对靶核酸区段的互补序列,所述互补序列随后将与靶核酸区段的核苷酸碱基形成相互作用(即,形成碱基对(bp))。应当理解,分离用核酸分子,例如生物素酰化的RNA,不需要含有针对靶核酸区段的完整互补序列以形成互补性相互作用和从核酸组合物分离它。分离用核酸分子可以长至少10个核苷酸碱基,例如长至少20、30、40、50、60、70、80、90、100、130、150、170、200、300、400、500、750、1000、2000、3000、4000或5000个核苷酸碱基。

[0074] 在一个实施方案中,在65°C和72°C之间添加与该一个靶核酸区段或多个靶核酸区段结合的分用核酸分子。在一个具体实施方案中,在65°C添加分离用核酸分子。因此,在又一个实施方案中,以上富集步骤(h)的步骤(i)在65°C和72°C之间(如在65°C)进行。在另一个实施方案中,在68°C和72°C之间使用所述结合对子的第二半部分分离出含有与分离用核酸分子结合的靶核酸区段的片段。在一个具体实施方案中,在68°C使用结合对子的第二半部分分离片段。因此,因此,在再一个实施方案中,富集步骤(h)的步骤(ii)在68°C和72°C之间(如在68°C)进行。

[0075] 在一个实施方案中,在阻断用或阻断物序列存在下添加分离用核酸分子。这类阻断物序列防止包含衔接子序列的已连接片段通过衔接子序列中的任何互补性与包含衔接子序列的其他已连接片段结合。因此,这类阻断物序列防止不包含该一个靶核酸区段或多个靶核酸区段的片段与的确包含该一个靶核酸区段或多个靶核酸区段的片段结合。在某些实施方案中,将阻断物序列在添加分离用核酸分子之前添加至已连接片段。在备选实施方案中,将阻断物序列同步地或连同分离用核酸分子一起添加至已连接片段。因此将理解,在一个实施方案中,阻断物序列包含与连接至片段的衔接子序列兼容的任何序列,如与特定衔接子序列互补的序列。在又一个实施方案中,阻断物序列包含与包含配对末端衔接子序列的半部分的MEDS寡核苷酸兼容(如与之互补)的任何序列。在一些实施方案中,阻断物序列选自如本文所述的SEQ ID NO:10、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:15、SEQ ID NO:16和/或SEQ ID NO:17。

[0076] 额外地,本发明方法的富集步骤(h)可以根据本领域已知的方法及使用本领域已知的试剂进行。例如,在富集步骤(h)包括如本文所述的与靶核酸区段结合的分用核酸分子情况下,该方法或该方法的步骤可以在二价阳离子盐浓度高(如在100mM与600mM之间)的

缓冲液存在下进行。盐可以按2.5:1和60:1之间的摩尔比存在。也可以存在体积排斥剂/增稠剂,例如按0.002%与0.1%之间的浓度。额外地或备选地,所述方法或该方法的步骤可以包括在缓冲液存在下孵育核酸组合物。孵育可以持续8小时或更短时间,任选地在两个不同温度孵育,其中两个不同温度在2次和100次之间循环。US 9,587,268中描述了这类缓冲液和方法的实例。因此将理解,在根据本文公开的某些实施方案进行富集步骤(h)时,包含分离用核酸分子的富集比使用常规试剂和反应方法时更迅速。

[0077] 本文中“结合对子”的提及,指彼此特异性识别以形成接合的至少两个部分(即第一半部分和第二半部分)。合适的结合对子例如包括生物素与亲和素或生物素与亲和素的衍生物如链霉亲和素与中性亲和素。

[0078] 本文中“标记的”或“标记”的提及,指通过连接标志物区分靶的过程,其中标志物包含对配体具有独特亲和力的特定部分(即亲和性标签)。例如,标记物可以起到选择性纯化分离用核酸序列的作用(即例如,借助亲和色谱)。这种标记物可以包括,但不限于生物素标记物、组氨酸标记物(即6His)或FLAG标记物。

[0079] 在一个实施方案中,分离用核酸分子包含生物素。在又一个实施方案中,将分离用核酸分子用生物素标记。在又一个实施方案中,将分离用核酸分子用组氨酸标记物或FLAG标记物标记。因此,根据某些实施方案,结合对子可以包含标记物(如组氨酸或FLAG标记物)和抗体。

[0080] 在一个实施方案中,靶核酸区段选自启动子、增强子、沉默子或绝缘子。在又一个实施方案中,靶核酸区段是启动子。在进一步备选的实施方案中,靶核酸区段是绝缘子。

[0081] 本文中对术语“一个启动子”和“多个启动子”的提及,指促进有效连接的编码区的转录起始的核酸序列。启动子有时称作“转录起始区”。调节元件经常与启动子相互作用以激活或抑制转录。

[0082] 本发明人已经使用本发明的方法鉴定了数千个启动子相互作用,每个启动子发生十个至二十个相互作用。本文所述的方法已经鉴定一些相互作用有细胞特异性或与不同的疾病状态相关。也已经鉴定到相互作用性核酸区段之间广泛的间隔距离-大部分相互作用在100千碱基范围内,但是可能延伸至2兆碱基及更远。有趣地,该方法也已经用来显示有活性基因和无活性基因均形成相互作用。

[0083] 经鉴定与启动子相互作用的核酸区段是为恰当遗传控制所需的调节元件候选物。它们的破坏可能改变转录性输出并促成疾病,因此将这些元件与它们的靶基因联系可能为新疗法提供潜在的新药靶。

[0084] 鉴定哪些调节元件与启动子相互作用对理解基因相互作用至关重要。本发明方法还提供对特定时间点处核酸组合物中的相互作用的快照浏览(snapshot look),因此构思该方法可以历经一系列的时间点或发育状态或实验条件进行,以构建细胞的核酸组合物中相互作用变化的画面。

[0085] 应当理解在一个实施方案中,靶核酸区段与包含调节元件的核酸区段相互作用。在又一个实施方案中,调节元件包括增强子、沉默子或绝缘子。

[0086] 如本文所用的术语“调节基因”指任何编码蛋白质的核酸序列,其中所述蛋白质与相同或不同的核酸序列结合,因而调节相同或不同核酸序列的转录速率或另外影响其表达水平。如本文所用的术语“调节元件”指影响另一个基因组元件的活性状态的任何核酸序

列。例如,各种调节元件可以包括但不限于增强子、激活子、阻遏子、绝缘子、启动子或沉默子。

[0087] 在一个实施方案中,靶核酸分子是通过染色质免疫沉淀(ChIP)测序所鉴定的基因组部位。ChIP测序实验通过使核酸组合物中的蛋白质-DNA复合体交联,分析蛋白质-DNA相互作用。随后将蛋白质-DNA复合体(通过免疫沉淀法)分离,之后对与蛋白质结合的基因组区域测序。

[0088] 在一些实施方案中,将构思核酸区段位于与靶核酸区段相同的染色体上。备选地,核酸区段位于与靶核酸区段不同的染色体上。

[0089] 该方法可以用来鉴定长程相互作用、短程相互作用或紧密相邻相互作用。如本文所用的术语“长程相互作用”指检测线性基因组序列中远隔的相互作用性核酸区段。这种类型的相互作用可以鉴定例如位于相同染色体的不同臂上或位于不同染色体上的两个基因组区域。如本文所用的术语“短程相互作用”指检测基因组中彼此定位相对接近的相互作用性核酸区段。如本文所用的术语“紧密相邻相互作用”指检测在线性基因组和例如相同基因的部分中彼此非常接近的相互作用性核酸区段。

[0090] 本发明人已经显示SNP更经常地位于相互作用性核酸区段中,多于预期偶然存在,因此本发明的方法可以用来鉴定哪些SNP与特定基因相互作用并且因此可能调节这些基因。

[0091] 因此,将可以从本文展示的公开内容理解,本发明方法可以用来鉴定核酸组合物中的任何核酸相互作用,尤其DNA-DNA相互作用。

[0092] 在一个实施方案中,从细菌人工染色体(BAC)、F粘粒或粘粒获得分离用核酸分子。在又一个实施方案中,从细菌人工染色体(BAC)获得分离用核酸分子。

[0093] 在一个实施方案中,分离用核酸分子是DNA、cDNA或RNA。在又一个实施方案中,分离用核酸分子是RNA。

[0094] 分离用核酸分子可以用于合适的方法中,如溶液杂交选择法(参见WO 2009/099602)。在这种方法中,生成一套‘诱饵’序列以形成可以用来从样品(即‘池塘’)分离靶核酸子群的杂交混合物。

[0095] 在一个实施方案中,结合对子的第一半部分包含生物素并且结合对子的第二半部分包含链霉亲和素。

[0096] 在一个实施方案中,该方法额外地包括在步骤(f)之前使交联逆转。应当理解本领域已知使交联逆转的数个方式,并且它将取决于最初形成交联的方式。例如,可以通过使交联的核酸组合物经历高热,如超过50°C、55°C、60°C、65°C、70°C、75°C、80°C、85°C或更高,使交联逆转。另外,可能需要使交联的核酸组合物经历高热长于1小时,例如,至少5小时、6小时、7小时、8小时、9小时、10小时或12小时或更长时间。在一个实施方案中,步骤(f)之前使交联逆转包括将交联的核酸组合物在蛋白酶K存在下于65°C孵育至少8小时(即过夜)。

[0097] 在一个实施方案中,该方法额外地包括在步骤(f)之前纯化核酸组合物以移除不含有接界标志物的任何片段。

[0098] 本文中对“纯化”的提及可以指这样的核酸组合物,其已经受到处理(即例如,分级)以移除多种其他组分并且所述组合物基本上保留其已表现的生物学活性。在使用术语“基本上纯化”的情况下,这种指称将指这样的组合物,其中所述核酸构成组合物的主要组

分,如构成组合物的约50%、约60%、约70%、约80%、约90%、约95%或更多(即例如,重量/重量(w/w)、体积/体积(v/v)和/或重量/体积(w/v))。

[0099] 在一个实施方案中,该方法额外地包括在步骤(i)之前扩增分离的靶连接片段。在又一个实施方案中,通过聚合酶链反应(PCR)进行扩增。

[0100] 在一个实施方案中,核酸组合物衍生自哺乳动物细胞核。在又一个实施方案中,哺乳动物细胞核可以是人细胞核。许多的人细胞在本领域可用于本文所述的方法中,例如GM12878(类淋巴母细胞人细胞系)或CD34+(人离体造血祖细胞)。

[0101] 将可以理解,本文所述的方法用于一系列生物中,不仅仅是人类中。例如,本发明的方法也可以用来鉴定植物和动物中的基因组相互作用。

[0102] 因此,在备选实施方案中,核酸组合物衍生自非人细胞核。在一个实施方案中,非人细胞选自包括但不限于:植物、酵母、小鼠、奶牛、猪、马、犬、猫、山羊或绵羊。在一个实施方案中,非人细胞核是小鼠细胞核或植物细胞核。

[0103] 将从如本文中提到的本发明优点理解,本发明方法在本文提到的步骤期间减少核酸组合物损耗。这种减少的核酸组合物损耗可以允许起始材料的量(例如,从中获得核酸组合物的细胞的数目)减少。因此,在一个实施方案中,核酸组合物可以衍生自数目比之前的启动子捕获或构象捕获技术更少的细胞。在又一个实施方案中,核酸组合物衍生自1百万个或更少的细胞、0.5百万个或更少的细胞、0.2百万个或更少的细胞、50000个或更少的细胞或10000个或更少的细胞。在又一个实施方案中,核酸组合物衍生自1百万个细胞、0.5百万个细胞、0.2百万个细胞、50000个细胞或10000个细胞。在某些实施方案中,核酸组合物衍生自1百万个细胞,50000个细胞或10000个细胞。

[0104] 在一个实施方案中,如本文所述的方法包括以下步骤:

[0105] (i) 使包含一个靶核酸区段或多个靶核酸区段的核酸组合物交联;

[0106] (ii) 使交联的核酸组合物片段化;

[0107] (iii) 用生物素标记片段的末端;

[0108] (iv) 连接片段化的核酸区段以产生已连接片段;

[0109] (v) 使交联逆转;

[0110] (vi) 使用转座酶对该已连接片段实施单步骤片段化和衔接子插入;

[0111] (vii) 用链霉亲和素减少已连接片段;

[0112] (viii) 进行包含该一个靶核酸区段或多个靶核酸区段的片段的定向扩增;并且

[0113] (ix) 测序以鉴定与该一个靶核酸区段或多个靶核酸区段相互作用的核酸区段。

[0114] 在另一个实施方案中,如本文所述的方法包括以下步骤:

[0115] (i) 使包含一个靶核酸区段或多个靶核酸区段的核酸组合物交联;

[0116] (ii) 使交联的核酸组合物片段化;

[0117] (iii) 用生物素标记片段的末端;

[0118] (iv) 连接片段化的核酸区段以产生已连接片段;

[0119] (v) 使交联逆转;

[0120] (vi) 使用转座酶对该已连接片段实施单步骤片段化和衔接子插入;

[0121] (vii) 用链霉亲和素减少片段并且使用PCR扩增;

[0122] (viii) 通过添加与该一个靶核酸区段或多个靶核酸区段结合的分离子核酸分子

捕获启动子,其中用结合对子的第一半部分标记所述分离用核酸分子;(ix)通过使用所述结合对子的第二半部分,分离出含有与分离用核酸分子结合的该一个靶核酸区段或多个靶核酸区段的已连接片段;

[0123] (x) 使用PCR扩增;并且

[0124] (xi) 测序以鉴定与该一个靶核酸区段或多个靶核酸区段相互作用的核酸区段。

[0125] 根据本发明的又一个方面,提供一种鉴定一个或多个指示特定疾病状态的相互作用性核酸区段的方法,所述方法包括:

[0126] a) 对从具有特定疾病状态的个体获得的核酸组合物进行本文所述的方法;

[0127] b) 对核酸区与一个靶核酸区段或多个靶核酸区段之间相互作用的频率定量;

[0128] c) 将来自具有所述疾病状态的个体的该核酸组合物中的相互作用频率与来自健康受试者的正常对照核酸组合物中的相互作用频率进行比较,从而核酸组合物中相互作用频率的差异指示特定疾病。

[0129] 如本文所用对“相互作用的频率”或“相互作用频率”的提及,指特定相互作用在核酸组合物(即样品)中出现的次数。在一些情况下,与来自健康受试者的正常对照核酸组合物比较,核酸组合物中较低的相互作用频率指示特定疾病状态(即原因在于核酸区段更不频繁地相互作用)。备选地,与来自健康受试者的正常对照核酸组合物比较,核酸组合物中较高的相互作用频率指示特定疾病状态(即原因在于核酸区段更频繁地相互作用)。在一些情况下,差异将由至少0.5倍差异,如1倍、1.5倍、2倍、2.5倍、3倍、4倍、5倍、7倍或10倍差异代表。

[0130] 在本发明的一个方面,相互作用的频率可以用来确定两个不同核酸区段的空间邻近状态。随着相互作用频率升高,两个基因组区域在3D核空间中彼此物理临近的概率升高。反过来,随着相互作用频率下降,两个基因组区域在3D核空间中彼此物理临近的概率下降。

[0131] 可以通过适合计算来自患者或纯化或提取核酸组合物样品或其稀释物的核酸组合物中相互作用频率的任何方法进行定量。例如,高通量测序结果还可以使得检查特定相互作用的频率成为可能。在本发明的方法中,可以通过测量一个样品或多个样品中靶核酸区段或连接产物的浓度进行定量。核酸组合物可以从生物样品中的细胞获得,所述生物样品可以包括脑脊液(CSF)、全血、血清、血浆或来自它们的提取物或纯化物,或其稀释物。在一个实施方案中,生物样品可以是脑脊液(CSF)、全血、血清或血浆。生物样品还包括来自活受试者的或死后取得的组织匀浆液、组织切片和活检标本。可以制备样品,例如根据需要稀释或浓缩,并且以常规方式保存。

[0132] 在一个实施方案中,疾病状态选自:癌症、自身免疫疾病、发育性疾病、遗传病、糖尿病、心血管疾病、肾脏疾病、肺部疾病、肝脏疾病、神经系统疾病、病毒性感染或细菌性感染。在又一个实施方案中,疾病状态是癌症或自身免疫疾病。在又一个实施方案中,疾病状态是癌症,例如乳腺癌、肠癌、膀胱癌、骨癌、脑癌、颈癌、结肠癌、子宫内膜癌、食管癌、肾癌、肝癌、肺癌、卵巢癌、胰腺癌、前列腺癌、皮肤癌、胃癌、睾丸癌、甲状腺癌或子宫癌、白血病、淋巴瘤、骨髓瘤或黑素瘤。

[0133] 本文中对“自身免疫疾病”的提及包括源自靶向人自身身体的免疫应答的疾病,例如急性播散性脑脊髓炎(ADEM)、强直性脊柱炎、Behçet病、小肠吸收不良(Celiac disease)、克罗恩氏病、1型糖尿病、Graves病、Guillain-Barré综合征(GBS)、银屑病、类风

湿性关节炎、风湿热、Sjögren 综合征、溃疡性结肠炎和血管炎。

[0134] 本文中“发育疾病”的提及包括多种疾病,通常源自儿童期,如学习失能、沟通障碍、自闭症、注意力缺陷多动障碍(ADHD)和发育性共济失调。

[0135] 本文中“遗传病”的提及包括源自基因组中一个或多个异常的疾病,如Angelman综合征、卡纳万病、Charcot-Marie-Tooth病、色盲、Cri du chat综合征、囊性纤维化、唐氏综合征、杜兴氏肌营养不良、血色素沉着症、血友病、先天性睾丸发育不全综合征、神经纤维瘤病、苯丙酮尿症、多囊性肾病、Prader-Willi综合征、镰刀形红细胞疾病、Tay-Sachs病和Turner综合征。

[0136] 根据本发明的又一个方面,提供一种用于鉴定与一个靶核酸区段或多个靶核酸区段相互作用的核酸区段的试剂盒,所述试剂盒包含能够实施本文所述方法的缓冲液和试剂。

[0137] 试剂盒可以包括一种或多种实施该方法的物品和/或试剂。例如,用于本文所述方法中的寡核苷酸探针、扩增引物对和/或重组酶结合的寡核苷酸可以以分离的形式提供并且可以是试剂盒的部分,例如在其中保护内容物免于外部环境的合适容器如小瓶中。试剂盒可以包括根据本文所述方法的方案使用的说明书。其中核酸意图用于PCR中的试剂盒可以包括反应所需的一种或多种其他试剂,如聚合酶、核苷酸、缓冲溶液等。

[0138] 在一个实施方案中,试剂盒包含重组酶。在又一个实施方案中,包含于如本文所述的试剂盒中的重组酶是转座酶,如超活性突变型转座酶,例如超活性突变型Tn5转座酶。

[0139] 根据本发明的再一个方面,提供如本文所述的重组酶能够进行单步骤片段化和衔接子插入。因此,本文还提供能够进行标签化的重组酶。

[0140] 在一个实施方案中,本文提供的重组酶是超活性突变型转座酶。在又一个实施方案中,转座酶是超活性突变型Tn5转座酶。在又一个实施方案中,转座酶包含配对末端衔接子序列。

[0141] 应当理解除了前述那些之外,还可以在本文所述的实施例中见到试剂盒中待纳入的缓冲液和试剂的类型实例。

[0142] 以下研究和方案阐述本文所述方法的实施方案:

实施例

[0143] 缩略语:

[0144] BB 结合缓冲液

[0145] BSA 牛血清白蛋白

[0146] dd 双脱氧

[0147] EDTA 乙二胺四乙酸

[0148] HB Agilent杂交缓冲液(HBI、HBII、HBIII和HBIV)

[0149] NaCl 氯化钠

[0150] NTB 无吐温的缓冲液

[0151] PBS 磷酸盐缓冲盐水

[0152] PCR 聚合酶链反应

[0153] PE 配对末端

- [0154] rpm 转/分钟
- [0155] SDS 十二烷基硫酸钠
- [0156] SPRI珠 固相可逆固定化珠
- [0157] TB 吐温缓冲液
- [0158] Tn5 转座酶
- [0159] Tris-HCl 三(羟甲基)氨基甲烷盐酸盐
- [0160] WB 洗涤缓冲液
- [0161] 细胞固定
- [0162] 1. 对于单次实验, 须将最小50000个细胞在室温以甲醛终浓度2%固定10分钟。
- [0163] • 用0.125M终浓度的甘氨酸猝灭。
- [0164] • 在1500rpm (400xg) 于4°C离心5分钟。
- [0165] • 丢弃上清液, 在100 μ l冷的1xPBS中小心重悬沉淀物。在1500rpm (400xg) 于4°C离心5分钟。
- [0166] • 丢弃上清液并且在液氮中速冻或直接推进到下一个步骤。
- [0167] 细胞透化和限制性酶切消化
- [0168] 2. 在100 μ l冰冷裂解缓冲液中重悬来自步骤1的经固定细胞沉淀物。将管在冰上孵育30分钟。
- [0169] 3. 将管以约600g于4°C离心5分钟。
- [0170] 4. 移除上清液, 留下约20 μ l含胞核沉淀物的溶液。
- [0171] 5. 将沉淀物在100 μ l 1.2x NE缓冲液3 (如果使用Dpn II) 或NE缓冲液2 (如果使用Hind III) 中洗涤两次。
- [0172] 6. 移除上清液, 留下约20 μ l。添加334 μ l 1.2x NE缓冲液3 (或如果用Hind III作业, 添加NE缓冲液2)。
- [0173] 7. 添加12 μ l 10%SDS (终浓度0.3%, w/v); 以950转/分钟于37°C在热混合仪上振摇1小时。
- [0174] 8. 添加80 μ l 10%Triton (终浓度1.8%, v/v); 以950转/分钟于37°C在热混合仪上振摇1小时。
- [0175] 9. 添加30 μ l Dpn II (50U/ μ l) (或15 μ l Hind III-100U/ μ l和15 μ l H₂O) 并且按950转/分钟于37°C在热混合仪上振摇12-16小时。
- [0176] 生物素标记和Hi-C连接
- [0177] 10. 将消化混合物短暂离心。
- [0178] 11. 添加4.5 μ l dCTP、dTTP和dGTP (10mM混合物)、37.5 μ l生物素-dATP和10 μ l Klenow (5U/ μ l)。在37°C孵育45分钟, 每30秒以700转/分钟振摇10秒。
- [0179] 12. 以600g于4°C离心6分钟。
- [0180] 13. 移除上清液, 留下包含沉淀物的约50 μ l。
- [0181] 14. 添加835 μ l H₂O/100 μ l T4 DNA连接酶缓冲液/5 μ l BSA (20mg/ml)/10 μ l T4 DNA连接酶 (Invitrogen)。
- [0182] 15. 在16°C孵育最短4小时 (或长达12小时)。
- [0183] 纯化Hi-C DNA

- [0184] 16. 将管以600g于4℃离心6分钟。
- [0185] 17. 移除800μl上清液,管中留下200μl。
- [0186] 18. 添加15μl蛋白酶K (10mg/ml)。在65℃孵育4小时(任选)。
- [0187] 19. 添加15μl蛋白酶K (10mg/ml)。在65℃孵育过夜。
- [0188] 20. 遵循生产商的说明,用1x体积的SPRI珠(Beckman Coulter Ampure XP珠 A63881)纯化。切勿使珠过干,因为这可能减少长DNA片段的回收。在无核酸酶的水中孵育10分钟。
- [0189] 标签化
- [0190] 21. 如下设立若干标签化反应(根据收集的DNA的总量):
- [0191] • Xμl DNA
- [0192] • 4μl标签化缓冲剂(5X)
- [0193] • Yμl Tn5
- [0194] • 16-X-Y无核酸酶的水
- [0195] 在55℃无混合情况下孵育7分钟。
- [0196] 以约400bp的DNA片段分布为目标。
- [0197] 作为指南:如果约50ng DNA作业,使用0.5-1μl 12.3uM Tn5。如果用100-300ng DNA作业,使用1μl 24.6uM Tn5。
- [0198] 为了结果更好——滴定Tn5的量以得到正确的片段分布。
- [0199] 22. 在TapeStation或Bioanalyzer上检查DNA片段分布:
- [0200] • 利用1μl标签化混合物,添加3μl H₂O和1μl 0.2%SDS。在55℃孵育7分钟。
- [0201] • 通过添加5μl 0.2%SDS并且在55℃孵育7分钟除去转座酶。
- [0202] • 从这种混合使用2μl加载于TapeStation或Bioanalyzer上。
- [0203] 如果分布正确——添加1μl无核酸酶的水至来自步骤21的初始标签化混合物并且通过添加5μl 0.2%SDS并且在55℃孵育7分钟除去Tn5。
- [0204] 23. 将25μl这种标签化混合物与来自步骤22的剩余3μl混合物合并。
- [0205] 减少Hi-C连接产物
- [0206] 24. 为了减少连接事件,每份样品使用25μl链霉亲和素MyOne C1 Dyna珠。为了准备珠子,将它们用400μl TB缓冲液洗涤两次(每次洗涤转动3分钟)。在50μl 2xNTB缓冲液中重悬25μl珠。
- [0207] 25. 将珠(来自之前的步骤)与22μl TLE和28μl标签化混合物(来自步骤24)混合在一起。在室温孵育45分钟,缓慢转动。
- [0208] 26. 将珠用100μl 1xNTB洗涤4次,随后用50μl TLE洗涤两次。将珠在25μl无核酸酶的水中重悬。
- [0209] 文库制备
- [0210] 27. 如下进行5反应:
- [0211] • 5μl来自之前的混合物
- [0212] • 29.5μl H₂O
- [0213] • 10μl KAPA HiFi缓冲液(5x)
- [0214] • 1.5μl dNTPs (10mM)

- [0215] • 1 μ l KAPA HiFi DNA聚合酶
- [0216] • 3 μ l i7/i5引物(10 μ M)混合物
- [0217] PCR条件:
- [0218] • 在72 $^{\circ}$ C 3分钟
- [0219] • 4-7个以下循环 {10秒95 $^{\circ}$ C ;30秒55 $^{\circ}$ C ;30秒72 $^{\circ}$ C }
- [0220] • 在72 $^{\circ}$ C 5分钟
- [0221] 28. 将反应合并并用Ampure SPRI珠(1x比率)纯化。通过TapeStation/Bioanalyzer和Qubit检查捕获Hi-C文库的质量和数量。
- [0222] 采用生物素-RNA捕获杂交Hi-C文库-方法1
- [0223] 配制三个PCR管条:“DNA”、“杂交”和“RNA”。
- [0224] 29a. 制备Hi-C文库:将体积等同于300ng与1 μ g之间、尤其500ng的Hi-C文库转移入1.5ml微量离心管并且使用SpeedVac(45 $^{\circ}$ C,约15分钟)干燥。在4 μ l无核酸酶的水中重悬Hi-C DNA沉淀物。
- [0225] 30a. 配制阻断混合物。每份样品:
- [0226] • 阻断物#1-2.5 μ l (Agilent Technologies)
- [0227] • 阻断物#2-2.5 μ l (Agilent Technologies)
- [0228] • 定制阻断物-1 μ l
- [0229] 31a. 将来自之前的步骤的阻断混合物与来自步骤29的DNA文库混合。将10 μ l DNA文库转移入相应PCR管条的孔中。保持在冰上。
- [0230] 32a. 配制杂交混合物。使其保持在室温。
- [0231] • HBI-25 μ l
- [0232] • HBII-1 μ l
- [0233] • HBIII-10 μ l
- [0234] • HBIV-13 μ l
- [0235] 彻底混合;如果沉淀物已物形成,则在65 $^{\circ}$ C加热5分钟。将每次捕获30 μ l等分至“杂交”PCR管条(Agilent 410022)中的每个孔,用PCR排管盖(Agilent光学盖8x管条401425)盖上并且保持在室温。
- [0236] 33a. 配制RNA酶阻断溶液1:4(例如3 μ l RNA酶阻断物+9 μ l水)。
- [0237] 34a. 制备生物素-RNA。每次捕获:将5 μ l定制诱饵(或2 μ l定制诱饵+3 μ l无核酸酶的水,如果捕获体系尺寸<3Mb)与2 μ l RNA酶阻断稀释物混合。转移7 μ l这些混合物至“RNA”PCR管条。保持在冰上。
- [0238] 35a. 杂交反应:设定PCR热循环为以下程序:95 $^{\circ}$ C持续5分钟,65 $^{\circ}$ C- ∞
- [0239] 须加热PCR仪盖。在整个流程期间,迅速作业并且尽力保持PCR仪盖打开时间尽可能最短。样品蒸发将导致次优杂交条件。
- [0240] 36a. 将含Hi-C文库的“DNA”PCR管条转移至PCR仪,处于以下黑色标记的位置,并且启动PCR程序。将DNA在95 $^{\circ}$ C孵育5分钟。

[0241]

A											
B											
C											
D											DNA
E											
F											
G											
H											

[0242] 37a. 一旦温度已经达到65℃, 则将含杂交缓冲液的“杂交”PCR管条转移至PCR仪, 处于以下灰色标记的位置。在65℃孵育5分钟。

[0243]

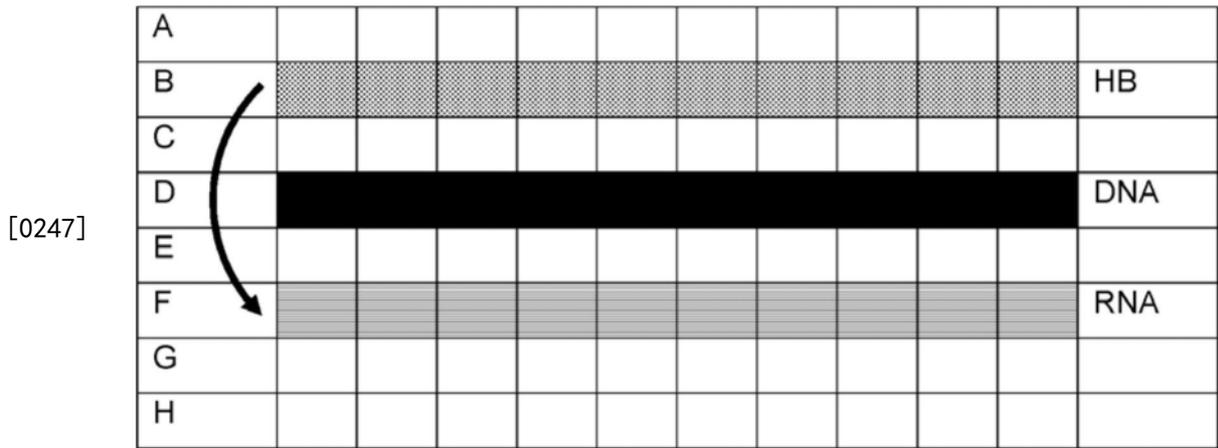
A											
B											HB
C											
D											DNA
E											
F											
G											
H											

[0244] 38a. 将含生物素酰化RNA诱饵的“DNA”PCR管条转移至PCR仪, 处于以下交叉阴影线标记的位置。孵育2分钟。

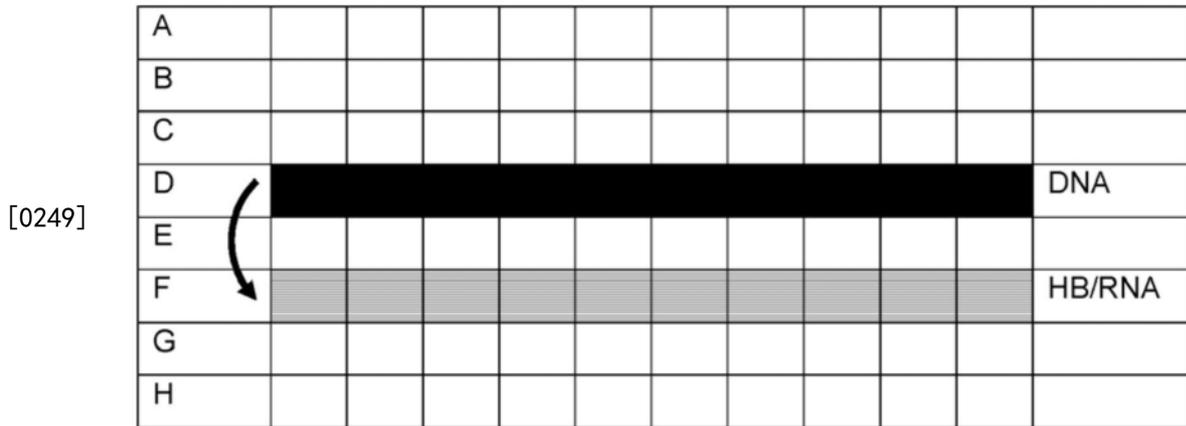
[0245]

A											
B											HB
C											
D											DNA
E											
F											RNA
G											
H											

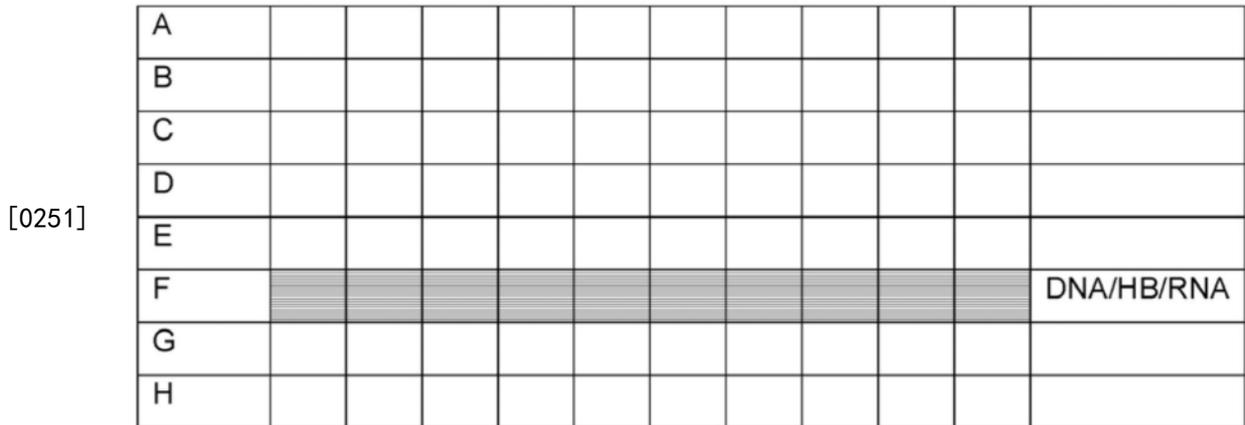
[0246] 39a. 打开“杂交”和“RNA”管条。将13μl 杂交缓冲液吸取至7μl RNA诱饵 (从灰色至交叉阴影线处) 中。丢弃盛有杂交缓冲液的PCR管条。立即推进到下一个步骤。



[0247] 40a. 从装有Hi-C文库的“DNA”PCR管条揭下盖子。将10μl Hi-C文库吸取至含杂交缓冲液的20μl RNA诱饵(从黑色至加交叉阴影线处)中。核实DNA PCR管条中无任何残留并丢弃它。



[0249] 立即以未用过的PCR排管盖封闭剩余的“RNA”PCR管条(现在盛装Hi-C文库/杂交缓冲液/RNA诱饵)并且在65℃孵育24小时。



[0251] 链霉亲和素-生物素减少和洗涤-待随以上方法1一起使用

[0252] 41a. 配制缓冲液:

[0253] 在室温配制结合缓冲液(BB,Agilent Technologies)

[0254] 在室温配制洗涤缓冲液I(WB I,Agilent Technologies)

[0255] 在65℃和72℃之间,尤其在65℃配制洗涤缓冲液II(WB II Agilent

Technologies)

[0257] 在室温配制NEB2 1x (NEB B7002S)。

[0258] 42a. 洗涤磁珠:

[0259] 彻底混合Dyna珠MyOne链霉亲和素T1 (Life Technologies 65601), 之后向1.5ml 低结合性微量离心管添加60 μ l/捕获Hi-C样品。如下洗涤珠子(全部后续洗涤步骤的程序相同):

[0260] • 添加200 μ l BB

[0261] • 在涡旋混合器(设定为低至中等)上混合5秒。

[0262] • 将管置于Dyna1磁分离器(Life Technologies)上

[0263] • 回收珠子, 丢弃上清液

[0264] 重复步骤a)至d), 总计洗涤3次。

[0265] 43a. 生物素-链霉亲和素减少:

[0266] 在Dyna珠MyOne链霉亲和素T1珠处于未用过的低结合微量离心管中的200 μ l BB时, 打开PCR仪的盖子(同时PCR仪正在运行)并且将整个杂交反应吸取至盛装有链霉亲和素珠的管中。在室温于转轮上孵育30分钟。

[0267] 44a. 洗涤:

[0268] 在30分钟后, 将样品置于磁分离器上, 丢弃上清液。

[0269] 使珠重悬于500 μ l WB I中, 并且转移至新管。在室温孵育15分钟。每隔2至3分钟涡旋混合5秒。

[0270] 在磁分离器上分离珠子和缓冲液并且移除上清液。重悬于500 μ l WB II(预温至65 $^{\circ}$ C和72 $^{\circ}$ C之间, 尤其65 $^{\circ}$ C)中并且转移至新管。在65 $^{\circ}$ C和72 $^{\circ}$ C之间, 尤其在65 $^{\circ}$ C孵育10分钟, 并且每2至3分钟涡旋混合(设定为低至中等)5秒。重复总计3次WB II中洗涤, 均在65 $^{\circ}$ C和72 $^{\circ}$ C之间, 尤其在65 $^{\circ}$ C进行。

[0271] 重悬于200 μ l Neb2 1X中。直接置于磁体上。移除上清液并且重悬于30 μ l Neb2 1X中。

[0272] 珠上的RNA/DNA混合物杂交体‘catch’现在为PCR扩增(步骤45)做好准备。

[0273] 采用生物素-RNA捕获杂交Hi-C文库-方法2(使用含高浓度二价阳离子盐的缓冲液-本文中称作“快速杂交”)

[0274] 如本文所述, 根据使用这种方法的实施方案, 准备时间可以大幅度减少(例如, 减少至约2小时45分钟)。

[0275] 29b. 在室温预温快速杂交缓冲液直至解冻并且保持在室温直至准备使用

[0276] 30b. 配制阻断混合物:

[0277] • 2.5 μ l 1mg/ml来自衍生出核酸组合物的相同物种的Cot-1 DNA, 如人Cot-1 DNA

[0278] • 2.5 μ l 10mg/ml鲑精DNA

[0279] • 1 μ l定制阻断物

[0280] 31b. 如下在室温设置阻断反应:

[0281] • 添加6 μ l上文配制的阻断混合物至11 μ l制备的DNA样品(大约100ng-1 μ g, 例如500ng)

[0282] • 上下抽吸以混合。短暂离心。

- [0283] 32b. 如下文显示对热循环仪编程。启动程序并且立即敲击暂停按钮。这将在添加阻断混合物至预制备的基因组DNA片段文库同时, 加热盖子。
- [0284] • 变性-95°C 5分钟
- [0285] • 阻断-65°C 10分钟
- [0286] • 杂交-50个循环-65°C 1分钟; 37°C 3秒
- [0287] • 保存-保持65°C
- [0288] 33b. 将样品置入热循环仪并且恢复程序以进行变性并阻断。
- [0289] 34b. 在热循环仪上孵育样品的同时, 在冰上配制捕获诱饵混合物。
- [0290] 35b. 稀释SureSelect RNase Block以捕获(1份RNase Block:3份水):
- [0291] • 混合1 μ l RNase Block (Agilent Technologies Inc.)
- [0292] • 3 μ l 水
- [0293] 36b. 配制杂交混合物:
- [0294] • 2 μ l 稀释的SureSelect RNase Block
- [0295] • 5 μ l SureSelect定制诱饵(或2 μ l SureSelect定制诱饵+3 μ l 无核酸酶的水, 如果捕获系统<3Mb)
- [0296] • 6 μ l 室温5X快速杂交缓冲液
- [0297] 37b. 当热循环在65°C 达到第一杂交循环时, 敲击暂停按钮。热循环仪现在维持在65°C。打开热循环仪盖子并且将13 μ l 杂交混合物吸取至每个相应的阻断反应中。通过缓慢上下抽吸8至10次充分混合。杂交反应现在为30 μ l。
- [0298] 38b. 用盖密封各孔, 关上热循环仪盖子, 并且敲击运行按钮以恢复程序运行循环杂交条件, 同时启动加热盖。
- [0299] 39b. 准备磁珠 (Dyna珠MyOne链霉亲和素T1, Invitrogen)
- [0300] • 在涡旋混合器上剧烈重悬Dyna1 (Invitrogen) 磁珠
- [0301] • 对每份杂交样品, 使用60 μ l Dyna珠T1磁珠
- [0302] • 洗涤珠子:
- [0303] (a) 添加200 μ l SureSelect结合缓冲液 (Agilent Technologies Inc.)
- [0304] (b) 通过上下抽吸10次混合珠子
- [0305] (c) 将管置于磁力座上
- [0306] (d) 等待2-5分钟并且丢弃上清液
- [0307] (e) 重复步骤(a) 至步骤(d), 总计洗涤3次
- [0308] (f) 使珠重悬于20 μ l SureSelect结合缓冲液中
- [0309] 40b. 使用链霉亲和素珠捕获杂交的DNA
- [0310] • 在孵育后, 从热循环仪取出样品并且在室温短暂离心以收集液体
- [0311] • 添加每份样本的整个杂交混合物至洗涤过和备妥的相应Dyna1 MyOne T1链霉亲和素珠溶液并且颠倒管条-管/平板以混合3至5次
- [0312] • 在转床或摇床上室温振摇30分钟来孵育杂交-捕获/珠溶液
- [0313] • 通过等分出1500 μ l/样品, 在68°C 和72°C 之间, 尤其在68°C 预温洗涤缓冲液#2
- [0314] • 30分钟后短暂离心杂交-捕获/珠溶液
- [0315] 41b. 洗涤珠子:

- [0316] (a) 在磁分离器上分离珠子和缓冲液并且移除上清液
- [0317] (b) 通过上下抽吸8-10次使珠重悬于500 μ l洗涤缓冲液#1中,随后在23 $^{\circ}$ C搁置10分钟。在磁分离器上分离珠子和缓冲液并且移除上清液
- [0318] (c) 重复步骤(a)至(b)
- [0319] (d) 在磁力座上分离珠子和缓冲液1分钟并且移除上清液
- [0320] (e) 添加500 μ l预温的洗涤缓冲液#2。缓慢上下抽吸10次以重悬珠子。当上下抽吸洗涤缓冲液时,使缓冲液直接分散沉淀的珠子以更快重悬它们
- [0321] (f) 在68 $^{\circ}$ C和72 $^{\circ}$ C之间,尤其在68 $^{\circ}$ C孵育样品10分钟
- [0322] (g) 重复步骤(d)至步骤(f),总计洗涤3次
- [0323] (h) 在磁力座上分离珠子和缓冲液。确保已经移除整个洗涤缓冲液#2
- [0324] (i) 重悬于50 μ l无核酸酶的水中,在磁力座上分离珠子并且移除上清液
- [0325] (j) 使珠重悬于23 μ l无核酸酶的水中,并且推进到PCR。
- [0326] 推进到捕获Hi-C文库的PCR扩增(步骤45)。
- [0327] 捕获Hi-C文库的PCR扩增
- [0328] 45. 以如下5个扩增循环设立PCR:
- [0329] • 5 μ l来自之前的混合物
- [0330] • 29.5 μ l mQ(水)
- [0331] • 10 μ l KAPA HiFi缓冲液(5x)
- [0332] • 1.5 μ l dNTPs(10mM)
- [0333] • 1 μ l KAPA HiFi DNA聚合酶
- [0334] • 3 μ l引物(10 μ M)混合物(P5-FCA-R和FCA-P7F)
- [0335] PCR条件:
- [0336] • 在95 $^{\circ}$ C3分钟
- [0337] • 5个以下循环{20秒95 $^{\circ}$ C;30秒55 $^{\circ}$ C;30秒72 $^{\circ}$ C}
- [0338] • 在72 $^{\circ}$ C3分钟
- [0339] 46. 合并来自以上步骤的全部各PCR反应。置于磁分离器上并且将上清液转移至未用过的1.5ml低结合性(lobind)微量离心管中。遵循生产商的说明,用1X体积的SPRI珠(Beckman Coulter Ampure XP珠A63881)纯化。重悬于终体积20 μ l TLE或无核酸酶的水中。
- [0340] 通过TapeStation/Bioanalyzer和KAPA qPCR检查捕获Hi-C文库的质量和数量。
- [0341] Tn5转座酶衔接子序列
- [0342] 用于Tn5转座酶上装配的序列:
- [0343]
- | | | |
|----------|---|-------------|
| Tn5MErev | 5' - [phos]CTGTCTCTTATACACATCT-3' | SEQ ID NO:1 |
| FC-A | 5' -GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-3' | SEQ ID NO:2 |
| FC-B | 5' -TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-3' | SEQ ID NO:3 |
- [0344] 用于预捕获PCR的引物

[0345]	Dual-i7-rcN701	CAAGCAGAAGACGGCATAACGAGATTAAGGCGAGTC TCGTGGGCTCGG	SEQ ID NO: 4
	Dual-i7-rcN702	CAAGCAGAAGACGGCATAACGAGATCGTACTAGGTC TCGTGGGCTCGG	SEQ ID NO: 5
	Dual-i7-rcN705	CAAGCAGAAGACGGCATAACGAGATGGACTCCTGTC TCGTGGGCTCGG	SEQ ID NO: 6
	Dual-i7-rcN706	CAAGCAGAAGACGGCATAACGAGATTAGGCATGGTC TCGTGGGCTCGG	SEQ ID NO: 7
	Dual-i5-S503	AATGATACGGCGACCACCGAGATCTACACTATCCTC TTCGTCGGCAGCGTC	SEQ ID NO: 8
	Dual-i5-S504	AATGATACGGCGACCACCGAGATCTACACAGAGTA GATCGTCGGCAGCGTC	SEQ ID NO: 9

[0346] “rc”表示条形码序列为反向互补。

[0347] 阻断物序列

[0348]	i5Rdd	TCGTGGCAGCGTCAGATGTGTATAAGAGA/3ddC/	SEQ ID NO:10
	ddi7F	GTCTCGTGGGCTCGGAGATGTGTATAAGAGA/3ddC/	SEQ ID NO:11
	i5F	CTGTCTTTATACACATCTGACGCTGCCGACGA	SEQ ID NO:12
	P5-FCA-F	GTGTAGATCTCGGTGGTCGCCGTATCATT	SEQ ID NO:13
	P5-FCA-R	AATGATACGGCGACCACCGAGATCTACAC	SEQ ID NO:14
	i7R	CTGTCTTTATACACATCTCCGAGCCCACGAGAC	SEQ ID NO:15
	FCA-P7F	CAAGCAGAAGACGGCATAACGAGAT	SEQ ID NO:16
	FCA-P7R	ATCTCGTATGCCGTCTTCTGCTTG	SEQ ID NO:17

[0349] 缓冲溶液

[0350] 5X快速杂交缓冲液

[0351] 1540mM MgCl₂·6H₂O, 0.0417%w/w HPMC, 100mM Tris (pH 8.0) 和H₂O。

[0352] 洗涤缓冲液#1

[0353] (“低严格性缓冲液”—高盐浓度和低温,以移除非特异性结合的探针) 2XSSC, 0.1% SDS和H₂O。

[0354] 洗涤缓冲液#2

[0355] (“高严格性缓冲液”—低盐浓度和高温,以移除低亲合性杂交探针) 0.1XSSC, 0.1% SDS和H₂O。

序列表

	<110> 巴布拉罕姆研究所	
	<120> 新方法	
	<130> BAB-C-P2543PCT	
	<150> GB1914325.4	
	<151> 2019-10-04	
	<160> 17	
	<170> PatentIn 版本 3.5	
	<210> 1	
	<211> 19	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<220>	
	<221> misc_feature	
	<222> (1).. (1)	
	<223> 5' -[phos]	
	<400> 1	
	ctgtctctta tacacatct	19
	<210> 2	
[0001]	<211> 34	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 2	
	gtctcgtggg ctcggagatg tgtataagag acag	34
	<210> 3	
	<211> 33	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 3	
	tcgtcggcag cgtcagatgt gtataagaga cag	33
	<210> 4	
	<211> 47	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 4	
	caagcagaag acggcatacg agattaaggc gagtctcgtg ggctcgg	47

	<210> 5	
	<211> 47	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 5	
	caagcagaag acggcatacg agatcgtact aggtctcgtg ggctcgg	47
	<210> 6	
	<211> 47	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 6	
	caagcagaag acggcatacg agatggactc ctgtctcgtg ggctcgg	47
	<210> 7	
	<211> 47	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
[0002]	<400> 7	
	caagcagaag acggcatacg agattaggca tggctctcgtg ggctcgg	47
	<210> 8	
	<211> 51	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 8	
	aatgatacgg cgaccaccga gatctacact atcctcttcg tcggcagcgt c	51
	<210> 9	
	<211> 51	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 9	
	aatgatacgg cgaccaccga gatctacaca gtagtagatcg tcggcagcgt c	51
	<210> 10	
	<211> 30	
	<212> DNA	
	<213> 人工序列	

	<220>		
	<223>	合成的	
	<220>		
	<221>	misc_feature	
	<222>	(30)..(30)	
	<223>	/3ddC/	
	<400>	10	
		tcgtcggcag cgtcagatgt gtataagaga	30
	<210>	11	
	<211>	31	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成的	
	<220>		
	<221>	misc_feature	
	<222>	(31)..(31)	
	<223>	/3ddC/	
	<400>	11	
		gtctcgtggg ctcgagatg tgtataagag a	31
	<210>	12	
	<211>	33	
[0003]	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成的	
	<400>	12	
		ctgtctctta tacacatctg acgctgccga cga	33
	<210>	13	
	<211>	29	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成的	
	<400>	13	
		gtgtagatct cggtggtcgc cgtatcatt	29
	<210>	14	
	<211>	29	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成的	
	<400>	14	
		aatgatacgg cgaccaccga gatctacac	29

	<210> 15	
	<211> 34	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 15	
	ctgtctctta tacacatctc cgagcccacg agac	34
	<210> 16	
	<211> 24	
	<212> DNA	
[0004]	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 16	
	caagcagaag acggcatacg agat	24
	<210> 17	
	<211> 24	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成的	
	<400> 17	
	atctcgtatg ccgtcttctg cttg	24

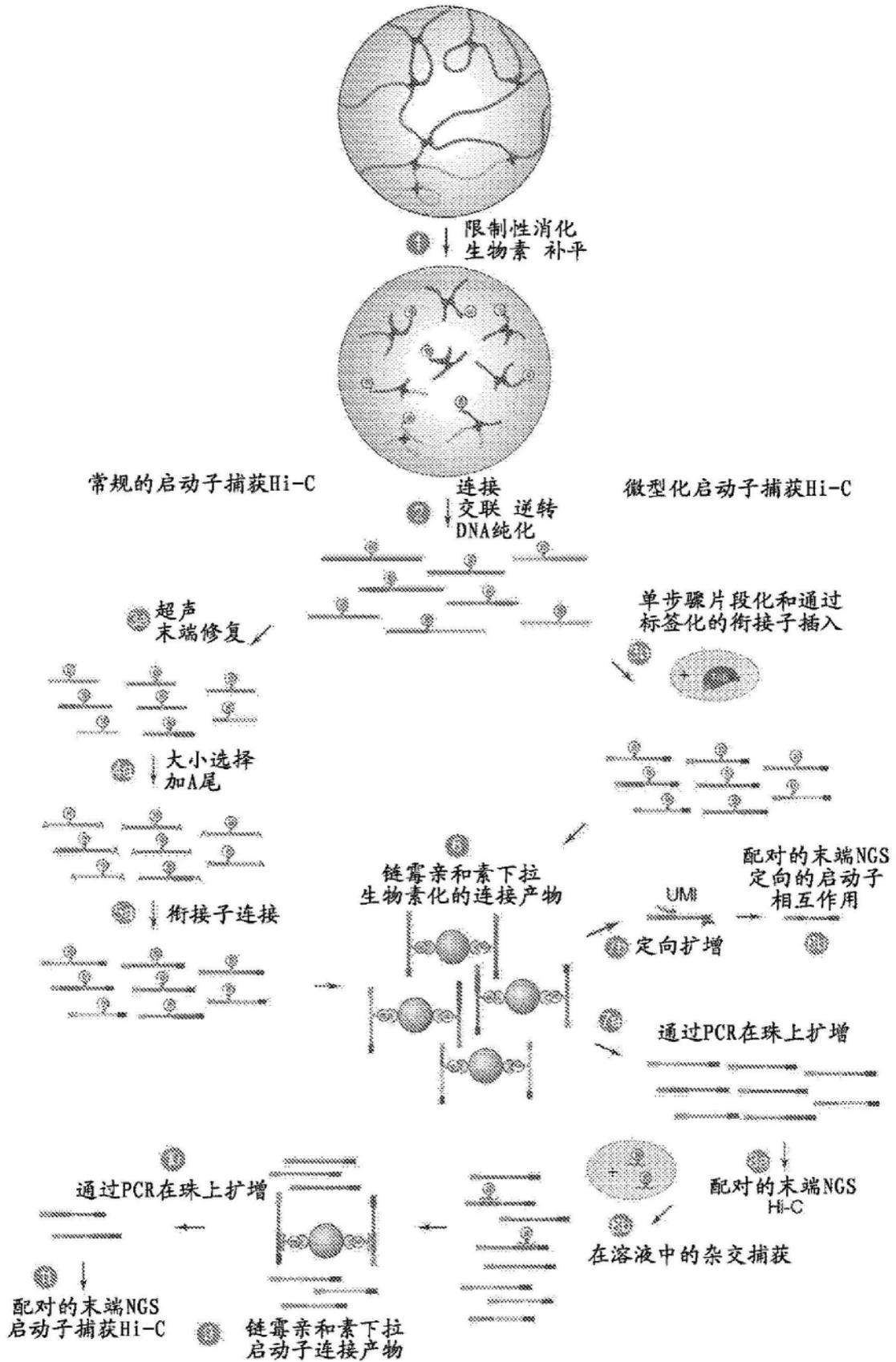


图1

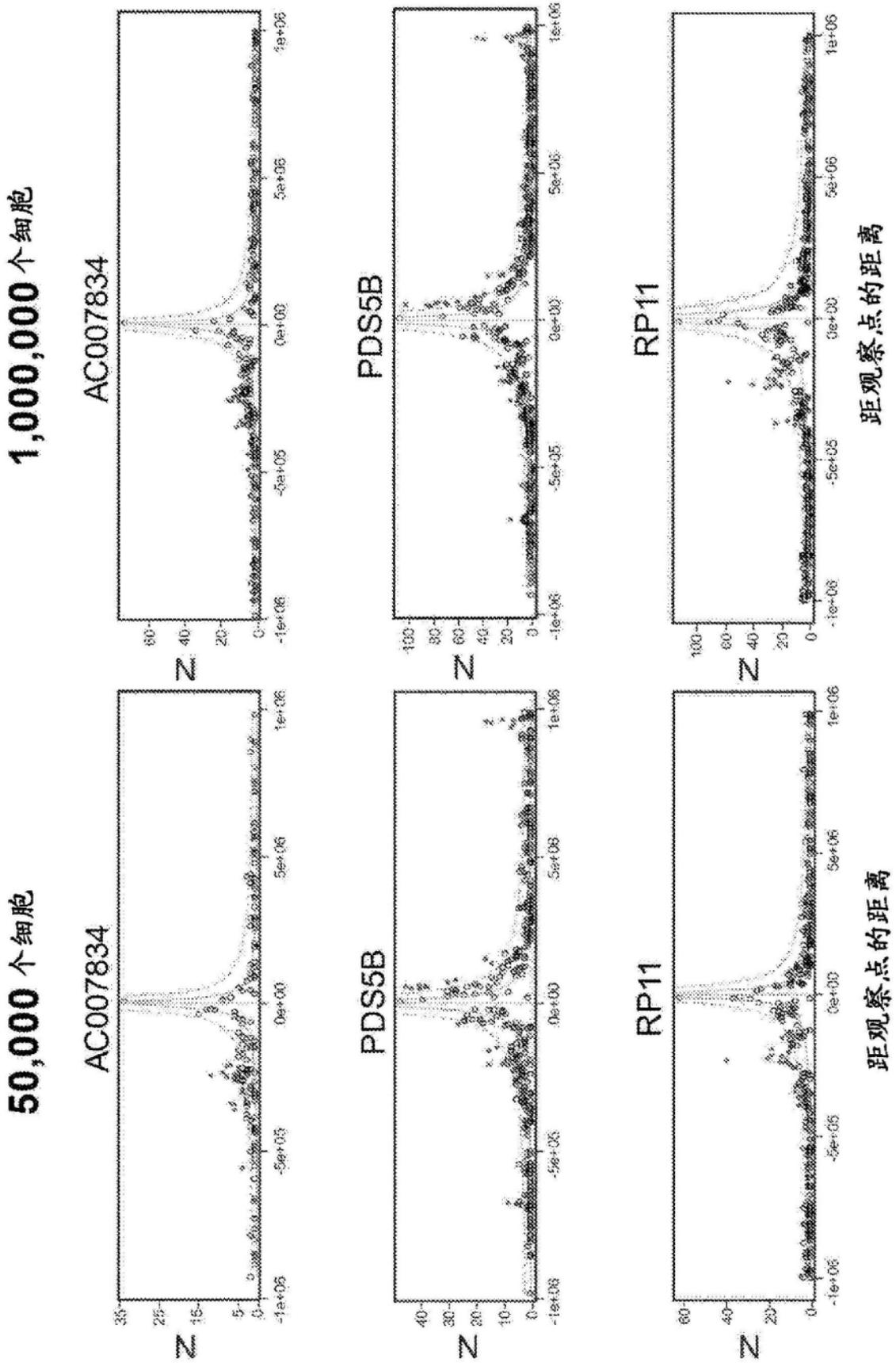


图2