



(12) 发明专利

(10) 授权公告号 CN 105658813 B

(45) 授权公告日 2021.01.05

(21) 申请号 201480049002.8

(72) 发明人 P. 弗拉泽 C. 奥斯本

(22) 申请日 2014.09.04

S. 舍恩费尔德

(65) 同一申请的已公布的文献号

(74) 专利代理机构 中国专利代理(香港)有限公司 72001

申请公布号 CN 105658813 A

代理人 李唐 彭昶

(43) 申请公布日 2016.06.08

(51) Int.CI.

(30) 优先权数据

C12Q 1/6806 (2018.01)

1315777.1 2013.09.05 GB

C12Q 1/6883 (2018.01)

(85) PCT国际申请进入国家阶段日

C12Q 1/6876 (2018.01)

2016.03.04

(56) 对比文件

(86) PCT国际申请的申请数据

CN 103180459 A, 2013.06.26

PCT/GB2014/052664 2014.09.04

WO 2012/108864 A1, 2012.08.16

(87) PCT国际申请的公布数据

WO 2012/159025 A2, 2012.11.22

W02015/033134 EN 2015.03.12

CN 101688237 A, 2010.03.31

(73) 专利权人 巴布拉哈姆研究院

审查员 蒋红云

地址 英国剑桥郡

权利要求书1页 说明书33页

序列表5页 附图3页

(54) 发明名称

包括选择和富集步骤的染色体构象捕获方法

(57) 摘要

本发明涉及通过使用分离的核酸分子鉴定与靶核酸区段相互作用的核酸区段的方法,以及用于所述方法的试剂盒。本发明还涉及鉴定一种或多种指示特定疾病状态的相互作用核酸区段的方法。

1. 用于鉴定与靶核酸区段子集相互作用的核酸区段的方法,所述方法包括以下步骤:
  - (a) 交联包含靶核酸区段子集的核酸组合物;
  - (b) 通过限制性消化片段化交联的核酸组合物;
  - (c) 用生物素标记片段的末端;
  - (d) 连接片段化的核酸区段以产生连接的片段;
  - (e) 逆转交联;
  - (f) 将所述片段剪切,随后末端修复;
  - (g) 用链霉亲和素将片段下拉;
  - (h) 衔接头连接,随后进行PCR;
  - (i) 通过添加结合靶核酸区段子集的分离的核酸分子进行启动子捕获,其中所述分离的核酸分子用结合对的第一半标记;
  - (j) 通过使用所述结合对的第二半分离连接的片段,其包含结合到分离的核酸分子的靶核酸区段的子集;
  - (k) 使用PCR进行扩增;和
  - (l) 测序以鉴定与靶核酸区段子集相互作用的核酸区段。
2. 权利要求1的方法,其中靶核酸区段子集选自启动子、沉默基因、增强子或绝缘子。
3. 权利要求1或权利要求2的方法,其中所述分离的核酸分子从细菌人工染色体、F黏粒或粘粒获得。
4. 权利要求1或权利要求2的方法,其中所述分离的核酸分子是RNA。
5. 权利要求1或权利要求2的方法,其中所述结合对的第一半包含生物素并且所述结合对的第二半包含链霉亲和素。
6. 权利要求1或权利要求2的方法,其中所述核酸组合物源自哺乳动物细胞核。
7. 权利要求1或权利要求2的方法,其中所述核酸组合物源自人细胞核。
8. 权利要求1或权利要求2的方法,其中所述核酸组合物源自非人细胞核。
9. 权利要求1或权利要求2的方法,其中所述核酸组合物源自小鼠细胞核或植物细胞核。
10. 能够进行权利要求1-9任一项的方法的缓冲液和试剂在制备用于鉴定一种或多种指示特定疾病状态的相互作用核酸区段的方法的试剂盒中的用途,所述方法包括:
  - a) 对从具有特定疾病状态的个体获得的核酸组合物执行权利要求1-9任一项的方法;
  - b) 定量核酸区段和靶核酸区段之间的相互作用频率;
  - c) 比较来自具有所述疾病状态的个体的核酸组合物中相互作用频率与来自健康主体的正常对照核组合物中相互作用频率,使得核酸组合物中相互作用频率的差异指示特定疾病状态。
11. 权利要求10的用途,其中所述疾病状态选自癌症、自身免疫性疾病、发育性疾病或遗传障碍。
12. 试剂盒,其用于鉴定与靶核酸区段子集相互作用的核酸区段,其包含能够进行权利要求1-9任一项的方法的缓冲液和试剂。

## 包括选择和富集步骤的染色体构象捕获方法

### 发明领域

[0001] 本发明涉及用于鉴定与靶核酸区段相互作用的核酸区段的方法,特别是用于鉴定与靶核酸区段的子集相互作用的核酸区段的方法,以及用于进行该方法的试剂盒。本发明还涉及鉴定一种或多种指示特定疾病状态的相互作用核酸区段的方法。

### [0002] 发明背景

[0003] 调控元件在生物体的遗传控制中发挥关键作用,并已被证明促成健康和疾病(例如癌症和自身免疫性疾病)。这些调控元件的鉴定也可以在分子生物学中具有潜在的应用,例如在表达系统和遗传修饰技术中。然而,虽然数以千计的调控元件已经在小鼠和人的基因组中作图,但是确定它们调节哪些目标基因代表一个重大的挑战。已经证明调控DNA序列(例如,增强子)可以跨过(bridge)相当的基因组距离以与它们的靶基因相互作用。基因及其调控区之间的染色体接触可以被捕获,但目前缺失在全基因组规模将基因连接到其调控序列的方法。

[0004] 开发以鉴定基因组基因座之间相互作用的最早的方法之一是染色体构象捕获(3C)技术(Dekker 等 *Science* (2002) 295:1306–1311)。这涉及通过如下产生3C文库:交联核组合物,以便空间上紧密邻近的基因组基因座被连接;通过消化除去交联之间的插入DNA(intervening DNA)环;并且连接和逆转交联相互作用区域以生成3C文库。所述文库然后可以用于计算已知序列间相互作用频率。但是此方法需要先前的相互作用的知识,以便检测目标相互作用区域。

[0005] 从那之后,该技术已经进一步开发以尝试和克服3C法的限制。例如4C法使用反向PCR以鉴定其中特异性相互作用是未知的诱饵序列的相互作用(Simonis 等 *Nat. Genet.* (2006) 38:1341–1347;和Zhao 等 *Nat. Genet.* (2006) 38:1348–1354)。然而,该方法对时间和资源两者是昂贵的,设计用来检测与仅一个诱饵基因组基因座的相互作用,并且仅是半定量的。

[0006] Hi-C方法使用连接标志物以分离细胞中所有的连接的相互作用序列(参见W0 2010/036323和Lieberman-Aiden 等, *Science* (2009) 326:289–93)。虽然这提供了关于在特定时间点上核组合物内发生的所有相互作用的信息,但是每个细胞有约800,000个可连接的片段(例如,如果具有6个碱基对的识别位点的限制性内切酶,如Hind III或Bg I II已被用于Hi-C文库的生成),这使得文库过于复杂且阻止特定个别区域之间的相互作用的问询。

[0007] 因此,需要提供克服了现有方法局限性的用于鉴定核酸相互作用的改进的方法。

### [0008] 发明概述

[0009] 根据本发明的第一个方面,提供了用于鉴定与靶核酸区段子集相互作用的核酸区段的方法,所述方法包括以下步骤:

[0010] (a)获得包含靶核酸区段子集的核酸组合物;

[0011] (b)交联核酸组合物;

[0012] (c)通过超声处理片段化交联的核组合物;

[0013] (d)连接从步骤(c)得到的片段化的核酸区段以产生连接的片段;

[0014] (e) 添加结合靶核酸区段子集的分离的核酸分子,其中所述分离的核酸分子用结合对的第一半标记;

[0015] (f) 通过使用所述结合对的第二半分离连接的片段,其包含结合到分离的核酸分子的靶核酸区段的子集;和

[0016] (g) 测序来自步骤(f)的分离的连接的片段以鉴定与靶核酸区段子集相互作用的核酸区段。

[0017] 根据本发明的进一步的方面,提供了鉴定一种或多种指示特定疾病状态的相互作用核酸区段的方法,其包括:

[0018] a) 对从具有特定疾病状态的个体获得的核酸组合物执行本文定义的方法;

[0019] b) 定量核酸区段和靶核酸区段之间的相互作用频率;

[0020] c) 比较来自具有所述疾病状态的个体的核酸组合物中相互作用频率与来自健康主体的正常对照核酸组合物中相互作用频率,使得核酸组合物中相互作用频率的差异指示特定疾病状态。

[0021] 根据本发明进一步的方面,提供了用于鉴定与靶核酸区段子集相互作用的核酸区段的试剂盒,其包含能够进行本文所限定的方法的缓冲液和试剂。

[0022] 附图简述

[0023] 图1:本文描述的发明的实施方案的示意图概述。

[0024] 图2:使用本文所述方法获得的结果。图2(A)展示了MYC基因周围的基因组区域和MYC基因在CD34+细胞中的染色体相互作用概况。图2(B)显示了CD34+细胞核的DNA FISH的实例。图2(C)显示使用3C方法对MYC启动子和在所述基因下游1.8Mb处区域的相互作用的验证。

[0025] 发明详述

[0026] 根据本发明的第一个方面,提供了用于鉴定与靶核酸区段子集相互作用的核酸区段的方法,所述方法包括以下步骤:

[0027] (a) 获得包含靶核酸区段子集的核酸组合物;

[0028] (b) 交联核酸组合物;

[0029] (c) 通过超声处理片段化交联的核酸组合物;

[0030] (d) 连接从步骤(c)得到的片段化的核酸区段以产生连接的片段;

[0031] (e) 添加结合靶核酸区段子集的分离的核酸分子,其中所述分离的核酸分子用结合对的第一半标记;

[0032] (f) 通过使用所述结合对的第二半分离连接的片段,其包含结合到分离的核酸分子的靶核酸区段的子集;和

[0033] (g) 测序来自步骤(f)的分离的连接的片段以鉴定与靶核酸区段子集相互作用的核酸区段。

[0034] 根据可以提及的本发明的第二方面,提供了用于鉴定与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:

[0035] (a) 获得包含靶核酸区段的核酸组合物;

[0036] (b) 交联核酸组合物;

[0037] (c) 片段化交联的核组合物;

- [0038] (d) 连接从步骤(c)得到的片段化的核酸区段以产生连接的片段；  
[0039] (e) 添加配置以结合靶核酸区段的分离的核酸分子，其中所述分离的核酸分子用结合对的第一半标记；  
[0040] (f) 通过使用所述结合对的第二半分离连接的片段，其包含结合到分离的核酸分子的靶核酸区段；和  
[0041] (g) 测序来自步骤(f)的分离的连接的片段以鉴定与靶核酸区段相互作用的核酸区段。

[0042] 本发明的方法提供了通过使用分离的核酸分子分离靶核酸区段用于鉴定相互作用序列的手段，特别是分离靶核酸区段子集的分离的核酸分子。这有助于在非常复杂的文库内将数据集中在特定的相互作用，并可以用于组织信息成各种子集，其取决于所用的分离的核酸分子的类型（例如启动子，以鉴定启动子相互作用）。然后可以获得在特定子集内的染色体相互作用的详细信息。

[0043] 尽管以前的技术，诸如4C，允许一个或几个启动子的全基因组相互作用的捕获，但本文描述的方法可以在一个单一实验中捕获超过22,000个启动子以及它们的相互作用基因组基因座。此外，本方法产生了显著更定量的读出。

[0044] 全基因组关联研究(GWAS)已经鉴定了数千个与疾病有联系的单核苷酸多态性(SNP)。然而，许多这些SNP位置距离基因很远，使得预测它们对其作用的基因非常具有挑战性。因此，本方法提供了鉴定相互作用核酸区段的一种方式，即使它们在基因组内位置远离彼此。

[0045] 如本文所用，提及“核酸区段”，相当于提及“核酸序列”，并且指的是核苷酸（即，例如，腺嘌呤(A)，胸腺嘧啶(T)，胞嘧啶(C)，鸟嘌呤(G)，和/或尿嘧啶(U)）的任何聚合物。这种聚合物可以产生或不可以产生功能性基因组片段或基因。核酸序列的组合最终可包括一个染色体。包含脱氧核糖核苷的核酸序列被称为脱氧核糖核酸(DNA)。包含核糖核苷的核酸序列被称为核糖核酸(RNA)。RNA可进一步表征为几个类型，诸如蛋白质编码RNA、信使RNA(mRNA)、转移RNA(tRNA)、长的非编码RNA(lnRNA)、长基因间非编码RNA(lincRNA)、反义RNA(asRNA)、微RNA(miRNA)、短干扰RNA(siRNA)、小核(snRNA)和小核仁RNA(snoRNA)。

[0046] “单核苷酸多态性”或“SNP”是基因组内的单核苷酸变化（即A,C,G或T），其在生物物种的成员之间或成对染色体之间不同。

[0047] 应当理解，“靶核酸区段”指的是用户熟知的目标序列。只分离含有靶核酸区段的连接片段有助于集中数据以鉴定与特定目标基因的特异性相互作用。

[0048] 本文提及的术语“相互作用”，是指两个元件之间的关联(association)，例如在本方法中，核酸区段和靶核酸区段之间的基因组相互作用。相互作用通常导致一个相互作用元件对另一个具有影响，例如，沉默或激活其结合的元件。相互作用可以在线性基因组序列上位置靠近在一起或相距甚远的两个核酸区段之间发生。

[0049] 本文提及的术语“核酸组合物”，是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被组织成染色体，其中蛋白（即，例如，组蛋白）可以与具有调控功能的染色体结合。在一个实施方案中，核酸组合物包含核组合物。这样的核组合物可以通常包括核基因组组构或染色质。

[0050] 如本文所用，提及“交联”，是指两种化合物之间任何稳定的化学结合，使得它们可

以作为一个单元被进一步加工。这种稳定性可以基于共价和/或非共价键合(例如,离子的)。例如,核酸和/或蛋白质可以是通过化学剂(即,例如,固定剂)、热、压力、pH变化或辐射而交联,使得它们在常规实验室程序(即,例如,提取、洗涤、离心等)过程中保持其空间关系。如本文所用交联等价于术语“固定”,其适用于固定任何和所有的细胞过程的任何方法或过程。因此,交联/固定的细胞,精确地保持核酸组合物内的组分之间在固定时的空间关系。许多化学品能够提供固定,包括但不限于,甲醛、福尔马林、或戊二醛。

[0051] 提及如本文所用术语“片段”,指的是比其来源的序列短的任何核酸序列。片段可以是任何大小,范围从几兆碱基和/或千碱基到只有几个核苷酸长。片段合适地长度大于5个核苷酸碱基,例如长度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,例如长度1、5、10、20、25、50、75、100、200、300、400或500个核苷酸千碱基。本发明的第二方面中,方法诸如限制性内切酶消化、超声处理、酸孵育、碱孵育、微流化等,都可以使用以片段化核酸组合物。

[0052] 本文提及的术语“连接的”,是指通常包含磷酸二酯键的两个核酸区段的任何连接。通常通过催化酶(即,例如,连接酶诸如T4 DNA连接酶)的存在在辅因子试剂和能量来源(即,例如,三磷酸腺苷(ATP))的存在下促进连接。在本文描述的方法中,已被交联的两个核酸区段的片段连接在一起以产生一个单一的连接的片段。

[0053] 本文提及“分离的核酸分子”是指由结合到靶核酸区段的核酸形成的分子。例如,分离的核酸分子可含有靶核酸区段的互补序列,其然后将形成与靶核酸区段的核苷酸碱基的相互作用(即,形成碱基对(bp))。应当理解,分离的核酸分子,例如生物素化的RNA,并不需要含有靶核酸区段的整个互补序列以形成互补相互作用并将其从核酸组合物分离。分离的核酸分子可以是至少10个核苷酸碱基长,例如,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个核苷酸碱基长。

[0054] 本文提及“结合对”是指特异性识别彼此以形成连接的至少两个部分(即,第一半和第二半)。合适的结合对包括,例如,生物素和抗生物素蛋白或生物素和抗生物素蛋白的衍生物,诸如链霉亲和素和中性亲和素(neutravidin)。

[0055] 本文提及“标记”或“标记的”指的是通过连接标志物区分靶标的过程,其中所述标志物包括对配体具有独特亲和力的特定部分(即,亲和标签)。例如,标记可用于选择性地纯化分离的核酸序列(即,例如,通过亲和层析)。这种标记可包括但不限于,生物素标记、组氨酸标记(即,6His)或FLAG标记。

[0056] 在一个实施方案中,靶核酸分子,即,靶核酸分子的子集,选自启动子、增强子、沉默基因或绝缘子。在进一步的实施方案中,所述靶核酸分子是启动子。在进一步的替代实施方案中,所述靶核酸分子是绝缘子。

[0057] 本文提及的术语“启动子”是指利于可操作连接的编码区域的转录起始的核酸序列。启动子有时称作“转录起始区”。调控元件通常与启动子相互作用,以激活或抑制转录。

[0058] 本发明人已使用本发明的方法,以鉴定数千个启动子相互作用,且每个启动子发生十到二十个相互作用。本文所描述的方法已经鉴定了一些相互作用是细胞特异性的,或者与不同的疾病状态相关联。也已经鉴定了相互作用的核酸区段之间大范围的分隔距离—大部分相互作用在100千碱基内,但有些可以扩展到2兆碱基和更多。有趣的是,该方法也被用于显示活性和非活性基因两者都形成相互作用。

[0059] 被鉴定为与启动子相互作用的核酸区段是对于正确的遗传控制所需的调控元件的候选。他们的破坏可能会改变转录输出，并导致疾病，因此，连接这些元件与它们的靶基因可以为新的治疗提供潜在的新的药物靶标。

[0060] 鉴定哪些调控元件与启动子相互作用对理解遗传相互作用是至关重要的。本方法还提供了在特定时间点观察核酸组合物内相互作用的快速方法(snapshot)，因此可以设想，该方法可以在一系列的时间点或发育状态或实验条件下进行以建立细胞的核酸组合物内相互作用变化的画面。

[0061] 应该理解的是，在一个实施方案中，靶核酸区段与包含调控元件的核酸区段相互作用。在进一步的实施方案中，调控元件包括增强子、沉默基因或绝缘子。

[0062] 如本文所用术语“调控基因”，指的是编码蛋白质的任何核酸序列，其中所述蛋白质结合相同或不同的核酸序列，从而调节转录速率或以其他方式影响相同或不同核酸序列的表达水平。如本文所用的术语“调控元件”，是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，各种调控元件可以包括，但不限于，增强子、激活子、阻遏物、绝缘子、启动子或沉默基因。

[0063] 在一个实施方案中，所述靶核酸分子是通过染色质免疫沉淀(ChIP)测序鉴定的基因组位点。ChIP测序实验通过交联核酸组合物内的蛋白质-DNA复合物分析蛋白质-DNA相互作用。然后(通过免疫沉淀)分离蛋白质-DNA复合物，随后测序蛋白质结合的基因组区域。

[0064] 应当设想，在一些实施方案中，核酸区段位于与靶核酸区段相同的染色体上。可选地，核酸区段位于与靶核酸区段不同的染色体上。

[0065] 该方法可以用于鉴定长距离相互作用、短距离相互作用或紧密近邻的相互作用。如本文所用的术语“长距离相互作用”，指的是检测线性基因组序列内相距很远的相互作用核酸区段。这种类型的相互作用可以鉴定两个基因组区域，其例如，位于同一染色体的不同臂，或位于不同的染色体上。如本文所用的术语“短距离相互作用”，指的是检测位于基因组内彼此相对接近的相互作用核酸区段。如本文所用的术语“紧密邻近的相互作用”，指的是检测线性基因组内非常靠近彼此和例如相同基因的一部分的相互作用核酸区段。

[0066] 本发明人已证明SNP更经常位于一个相互作用核酸区段而不是偶然(by chance)所预期的，因此本发明的方法可用于鉴定哪些SNP相互作用并因此有可能调控特定基因。

[0067] 在一个实施方案中，分离的核酸分子是从细菌人工染色体(BAC), F黏粒或粘粒获得。在进一步的实施方案中，分离的核酸分子从细菌人工染色体(BAC)中获得。

[0068] 在一个实施方案中，分离的核酸分子为DNA、cDNA或RNA。在进一步的实施方案中，分离的核酸分子是RNA。

[0069] 分离核酸分子可以在一个合适的方法使用，诸如溶液杂交选择(参见W02009/099602)。在该方法中，产生了一组“诱饵”序列，以形成可被用于从样品(即‘池’)分离靶核酸的子集的杂交混合物。

[0070] 在一个实施方案中，结合对的第一半包含生物素并且结合对的第二半包括链霉亲和素。

[0071] 在一个实施方案中，片段化(即，本发明第二方面的步骤(c))使用限制性内切酶或超声处理进行。在一个进一步的实施方案中，本发明第二方面的步骤(c)使用限制性内切酶进行。为避免疑问，应当理解，本发明第一方面的步骤(c)使用超声处理来进行。

[0072] 如本文所用的术语“限制性内切酶”，指的是在特定的碱基对序列切割核酸的任何蛋白。裂解可导致平的或粘性末端，这取决于所选限制性内切酶的类型。限制性内切酶的实例包括，但不限于，EcoRI、EcoRII、BamHI、HindIII、DpnII、BgIII、NcoI、TaqI、NotI、HinfI、Sau3A、PvuII、SmaI、HaeIII、HgaI、AluI、EcoRV、KpnI、PstI、SacI、SalI、ScaI、SpeI、SphI、StuI、XbaI。在一个实施方案中，限制性内切酶是HindIII。

[0073] 在一个实施方案中，该方法还包括步骤(d)期间并入连接标志物到连接片段。

[0074] 如本文所用的术语“连接标志物”，是指任何化合物或化学部分，其能够被并入核酸内，并且可以提供选择性纯化的基础。例如，连接标志物可以包括，但不限于，标记的核苷酸接头(例如生物素)，标记和/或修饰的核苷酸，切口平移，引物接头，或加标签的接头。

[0075] 连接标志物允许连接的片段在分离步骤(f)之前被纯化，因此，这确保只有连接的序列被分离的核酸分子结合，而不是非连接(即，无相互作用)片段。

[0076] 在一个实施方案中，连接标志物包括标记的核苷酸接头(即，例如，生物素)。在进一步的实施方案中，连接标志物包括生物素。在一个实施方案中，连接标志物包括修饰的核苷酸。在一个实施方案中，连接标志物包括引物接头。

[0077] 在一个实施方案中，在不加入连接标志物的情况下，核酸片段在步骤(d)中被阻止彼此连接，即连接标志物将包含接头序列。

[0078] 在一个实施方案中，所述方法还包括在步骤(e)之前逆转交联(reversing the cross-linking)。应该理解的是，存在若干在本领域中已知的方式来逆转交联，并且其将取决于最初形成交联的方式。例如，交联可通过使交联的核酸组合物经历高热而逆转，诸如高于50°C、55°C、60°C、65°C、70°C、75°C、80°C、85°C或更高。此外，交联的核酸组合物可能需要经受高热超过1个小时，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。在一个实施方案中，在步骤(e)之前逆转交联包括在蛋白酶K的存在下在65°C孵育该交联的核酸组合物至少8小时(即过夜)。

[0079] 在一个实施方案中，所述方法还包括在步骤(e)之前纯化所述核酸组合物以除去不包含连接标志物的任何片段。

[0080] 本文提及的“纯化”，可以指核酸组合物已经历处理(即，例如分级)以除去各种其它组分，并且该组合物基本上保留其表达的生物学活性。当使用术语“基本上纯化的”时，此指示将是指其中核酸形成组合物的主要组分的组合物，诸如构成约50%、约60%、约70%、约80%、约90%、约95%或更多的组合物(即，例如，重量/重量和/或重量/体积)。

[0081] 在一个实施方案中，该方法还包括在步骤(g)之前将配对的末端衔接头序列连接到分离的靶连接的片段的末端。

[0082] 如本文所用的术语“配对的末端衔接头”，指的是允许自动高通量测序以从两端同时读取的任何引物对集合。例如，与这些衔接头相容的这样的高通量测序装置包括但不限于Solexa(Illumina)、454系统、和/或ABI SOLiD。例如，该方法可以包括使用与聚A尾结合的通用引物。

[0083] 在一个实施方案中，所述方法还包括在步骤(g)之前扩增分离的目标连接片段。在进一步的实施方案中，扩增通过聚合酶链式反应(PCR)进行。

[0084] 在一个实施方案中，核酸组合物源自哺乳动物细胞核。在进一步的实施方案中，哺乳动物细胞核可以是人细胞核。本领域中可获得许多细胞用于在本文描述的方法中，例

如GM12878(人类淋巴母细胞系)或CD34+(人离体造血祖细胞)。

[0085] 应当理解,本文所描述的方法在一系列生物中发现实用性,不只是人类。例如,也可以使用本方法,以鉴定植物和动物中的基因组相互作用。

[0086] 因此,在替代实施方案中,核酸组合物源自非人细胞核。在一个实施方案中,所述非人细胞选自(但不限于)植物、酵母、小鼠、牛、猪、马、狗、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。

[0087] 根据本发明的进一步的方面,提供了鉴定一种或多种指示特定疾病状态的相互作用核酸区段的方法,其包括:

[0088] a)对从具有特定疾病状态的个体获得的核酸组合物执行本文定义的方法;

[0089] b)定量核酸区段和靶核酸区段之间的相互作用频率;

[0090] c)比较来自具有所述疾病状态的个体的核酸组合物中相互作用频率与来自健康主体的正常对照核酸组合物中相互作用频率,使得核酸组合物中相互作用频率的差异指示特定疾病状态。

[0091] 如本文所用提及“相互作用的频率”或“相互作用频率”,是指核酸组合物(即样品)内发生的特异性相互作用的次数。在一些情况下,在核酸的组合物中相比来自健康主体的正常对照核酸组合物更低的相互作用频率,指示特定疾病状态(即,因为核酸区段相互作用更不频繁)。可选地,在核酸的组合物中相比来自健康主体的正常对照核酸组合物更高的相互作用频率,指示特定疾病状态(即,因为核酸区段相互作用更频繁)。在一些情况下,差异将表示为至少0.5倍差异,诸如1倍、1.5倍、2倍、2.5倍、3倍、4倍、5倍、7倍或10倍差异。

[0092] 在本发明的一个方面中,相互作用频率可以被用于确定两个不同的核酸区段的空间接近性。随着相互作用频率的增加,这两个基因组区域在三维核空间中是物理上彼此接近的可能性增大。相反,随着相互作用频率的降低,这两个基因组区域在三维核空间中是物理上彼此接近的可能性降低。

[0093] 可以通过任何合适的方法进行定量,以计算来自患者的核酸组合物中或核酸组合物样品或其稀释物的纯化物或提取物中的相互作用频率。例如,高通量测序的结果也可以使得能够检查特定的相互作用频率。在本发明的方法中,定量可以通过测量一个或多个样品中的靶核酸区段或连接产物的浓度来进行。核酸组合物可以从生物样品中的细胞获得,所述样品可以包括脑脊髓液(CSF)、全血、血清、血浆、或来自其的提取物或纯化物,或其稀释物。在一个实施方案中,生物样品可以是脑脊髓液(CSF)、全血、血清或血浆。生物样品还包括组织匀浆、组织切片,以及来自活的主体的活检样本,或死后取样(taken post-mortem)。可以通常的方式制备(例如适当时稀释或浓缩)和储存样品。

[0094] 在一个实施方案中,所述疾病状态选自:癌症、自身免疫性疾病、发育性疾病、遗传障碍、糖尿病、心血管疾病、肾脏疾病、肺疾病、肝脏疾病、神经系统疾病、病毒感染或细菌感染。在进一步的实施方案中,所述疾病状态是癌症或自身免疫性疾病。在又进一步的实施方案中,所述疾病状态是癌症,例如乳腺癌、肠癌、膀胱癌、骨癌、脑癌、子宫颈癌、结肠癌、子宫内膜癌、食道癌、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。

[0095] 本文提及的“自身免疫性疾病”包括从靶向人自己身体的免疫应答产生的状况,例如急性播散性脑脊髓炎(ADEM)、强直性脊柱炎、白塞氏病、乳糜泻、克罗恩病、1型糖尿病、格

雷夫斯病、格林-巴利综合征(GBS)、牛皮癣、类风湿性关节炎、风湿热、斯耶格伦综合征、溃疡性结肠炎和血管炎。

[0096] 本文提及的“发育性疾病”包括下列状况，通常源自童年，诸如学习障碍、交流障碍、自闭症、注意力缺陷多动障碍(ADHD)和发育性协调障碍。

[0097] 本文提及的“遗传障碍”包括源自基因组中一个或多个异常的状况，诸如安格尔曼综合征、卡纳万病、腓骨肌萎缩症、色盲、猫叫综合征、囊性纤维化病、唐氏综合征、杜氏肌营养不良症、血色病、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、普瑞德-威利综合征、镰状细胞病、泰伊-萨克斯二氏病和特纳综合征。

[0098] 根据本发明进一步的方面，提供了用于鉴定与靶核酸区段子集相互作用的核酸区段的试剂盒，其包含能够进行本文所限定的方法的缓冲液和试剂。

[0099] 根据可以提及的本发明进一步的方面，提供了用于鉴定与靶核酸区段相互作用的核酸区段的试剂盒，其包含能够进行本文所限定的方法的缓冲液和试剂。

[0100] 所述试剂盒可以包含用于执行该方法的一个或多个物品和/或试剂。例如，用于本文描述的方法的寡核苷酸探针和/或扩增引物对可以以分离的形式提供，并且可以是试剂盒的一部分，例如在合适的容器诸如小瓶中，其中的内容物被保护免于外部环境。该试剂盒可包含用于根据本文中所述方法的方案使用的说明书。其中核酸意在用于在PCR中使用的试剂盒可以包含对于该反应所需的一种或多种其它试剂，诸如聚合酶、核苷酸、缓冲溶液等。

[0101] 应该理解的是，包含在试剂盒中的各类缓冲液和试剂的实例可以见于本文所述实施例。

[0102] 下述研究和方案阐释了本文所述方法的实施方案：

[0103] 缩写：

[0104] BAC 细菌人工染色体

[0105] BB 结合缓冲液

[0106] BSA 牛血清白蛋白

[0107] dd 双脱氧

[0108] DMEM Dulbecco氏改良Eagle培养基

[0109] EDTA 乙二胺四乙酸

[0110] FBS 胎牛血清

[0111] NaCl 氯化钠

[0112] NTB 无吐温缓冲液

[0113] PBS 磷酸盐缓冲盐水

[0114] PCR 聚合酶链式反应

[0115] PE 配对末端

[0116] rpm 每分钟转数

[0117] SCRiBL 与诱饵基因座相互作用的区域的序列捕获

[0118] SDS 十二烷基硫酸钠

[0119] SPRI珠 固相可逆固定化珠

[0120] TB 吐温缓冲液

- [0121] Tris-HCl 三(羟甲基)氨基甲烷盐酸盐
- [0122] WB 洗涤缓冲液。
- [0123] 实施例1:启动子捕获Hi-C方案
- [0124] 第一天:细胞的固定,HindIII消化
- [0125] 1. 用 $3 \times 10^7 - 4 \times 10^7$  个细胞开始。用室温的DMEM/10%FBS制成37ml。
- [0126] 2. 加入甲醛至2%的终浓度,并在室温下固定准确的10分钟,同时在摇动装置上混合。
- [0127] 3. 通过加入6ml冷的1M甘氨酸(最终0.125M)淬灭反应。
- [0128] 4. 在室温下孵育5分钟,随后在冰上15分钟。
- [0129] 5. 在4℃以1500rpm(400x g)离心10分钟。
- [0130] 6. 弃上清,在冷的1x PBS中小心重悬沉淀,添加冷1x PBS到50ml的最终体积。
- [0131] 7. 在4℃以1500rpm(400x g)离心10分钟,然后弃上清。细胞可以在液氮中快速冷冻,并在-80℃储存细胞。
- [0132] 8. 在50 ml冰冷的裂解缓冲液(10mM Tris-HCl pH8,10 mM NaCl,0.2% Igepal CA-630中重悬细胞,并添加一片蛋白酶抑制剂混合物(Roche完全的,无EDTA,11873580001)。在临用前新鲜制备。
- [0133] 9. 任选地,但对某些细胞类型建议:在30分钟的孵育期间,进行杜恩斯(Dounce)匀浆细胞( $2 \times 10$ 次冲击,中间休息5分钟,同时细胞被保持在冰上)(参见下文步骤10)。
- [0134] 10. 冰上孵育30分钟,并偶尔混合。
- [0135] 11. 在1800rpm下于4℃离心染色质5分钟,弃上清。
- [0136] 12. 在细胞沉淀的顶部小心铺层大约500μL的1.25x NE缓冲液2,但是不重悬。此步骤是除去任何剩余的裂解缓冲液。弃上清。
- [0137] 13. 每5至6 Mio细胞,在250μl的1.25x NE缓冲液2中重悬沉淀。
- [0138] 例如:当用30 Mio细胞开始时,在1.5 ml 1.25x NE缓冲液2中重悬沉淀。
- [0139] 14. 分成250μL的等分试样,分别对应5到6 Mio细胞(开始)。
- [0140] 例如:当用30 Mio细胞开始时,分成6等分试样,每个250μl。
- [0141] 15. 向每个等分试样中,添加108μl 1.25x NE缓冲液2(最终体积358μl)。
- [0142] 16. 每管添加11μl 10% SDS,小心混合,并在37℃孵育60分钟,以950rpm进行旋转。
- [0143] 此步骤除去不直接交联到DNA的蛋白。
- [0144] 17. 每管通过加入75μl 10%的Triton X-100淬灭SDS,并在37℃孵育60分钟,以950rpm进行旋转。
- [0145] 18. 每管通过加入1500单位的HindIII (NEB R0104T) 消化染色质并在37℃孵育过夜同时旋转(950rpm)。
- [0146] 第2天:DNA末端的生物素化,连接
- [0147] 注意:通过制备连接缓冲液混合物(参见第2天,步骤5)开始这一天,但只有临使用前添加10倍连接缓冲液。
- [0148] 1. 将管置于冰上。用Hi-C样品如下所述进行生物素化步骤。
- [0149] 2. 为了填补限制性片段突出端并用生物素标记DNA末端,添加6μl 10x NEB2,2μL

H<sub>2</sub>O, 1.5μl 10 mM的dCTP, 1.5μl 10mM的dGTP, 1.5μl 10 mM的dTTP, 37.5μl 0.4 mM的生物素-14-dATP (Life Technologies 19524-016), 以及10μl 5U/μl 克列诺酶(Klenow) (DNA聚合酶I大片段, NEB M0210L)。小心混合并在37°C孵育60分钟。不旋转,但小心上下吸打(大约每10分钟)混合,避免气泡。

[0150] 3. 将管置于冰上。向所有管中添加86μl 10%的SDS(最终浓度1.47%)以使酶失活。

[0151] 4. 以950rpm在65°C孵育酶正好30分钟并之后立即将管置于冰上。

[0152] 5. 在孵育期间(第2天,步骤4),每个样品准备1个15 ml Falcon管,每个管具有7.61 ml连接混合物(745μl 10%的Triton X-100,820μl 10×连接缓冲液(NEB B0202S),82μl 10 mg/ml BSA(NEB B9001S)和5.965 ml水)。将每个消化的染色质混合物(第2天,步骤4)转移到对应的15ml管。

[0153] 6. 在37°C孵育60分钟,并偶尔混合。

[0154] 7. 向每个管中加入50μl 1U/μl T4 DNA 连接酶(Invitrogen 15224-025)。通过颠倒所述管混合并在16°C孵育4小时。

[0155] 8. 在室温孵育另外30分钟。

[0156] 9. 通过每管中加入60μl的10 mg/ml的蛋白酶K(Roche 03115879001)和在65°C孵育所述管过夜来逆转交联并降解蛋白。

[0157] 第3天:DNA纯化第I部分

[0158] 1. 在过夜温育后,每管加入额外的60μl 10 mg/ml的蛋白酶K,并继续在65°C再孵育2小时。

[0159] 2. 冷却反应混合物至室温。添加12.5μl 10mg/ml的RNase A (Roche 10109142001),并于37°C孵育60分钟。

[0160] 3. 通过进行苯酚提取纯化在这些管中的DNA。转移反应混合物至50ml falcon管,并添加8 ml pH为8.0的苯酚(Sigma P4557)。涡旋1分钟,和在3500rpm下离心管10分钟,然后小心地将尽可能多的水相转移到新的50 ml falcon管。

[0161] 4. “反提取”:将2.5 ml 1xTE(10 mM Tris-HCl pH8,1mM EDTA)转移到包含来自第3天步骤3的较低相的管中。涡旋1分钟,和在3500rpm下离心管10分钟,然后小心地将尽可能多的上面水相转移到来自第3天步骤3的对应管中的上清液中。

[0162] 5. 按照第3天步骤3中所述步骤,使用10 ml苯酚 pH8.0:氯仿(Sigma P3803)重复提取。

[0163] 6. 为了使DNA沉淀,加入1000μL 3M的乙酸钠pH5.2,和25 ml冰冷的100%乙醇至每个管。在-20°C孵育过夜。

[0164] 第4天:DNA纯化,第II部分

[0165] 1. 在4°C以3500rpm离心30分钟。

[0166] 2. 弃上清,并在400μl 1xTE中重悬每个沉淀。在此步骤该DNA可以储存在-20°C,如果必要。

[0167] 3. 通过进行2次苯酚:氯仿提取进行另一轮纯化。加入450μl苯酚pH8.0:氯仿(1:1)并涡旋1分钟。将管以14,000rpm离心10分钟,并将水相转移到新管中。在第二次提取后,通过加入0.1x体积的3M乙酸钠pH5.2和2.5x体积100%乙醇沉淀DNA,并在-20°C下孵育过

夜。

[0168] 第5天:DNA纯化,第III部分

[0169] 1. 离心下沉淀的DNA(在4°C以14,000rpm进行30分钟)后,用70%乙醇洗涤各DNA沉淀3次,并在25μl 1×TE中重新悬浮每一DNA沉淀。

[0170] 2. 合并所有Hi-C管的内容物。

[0171] 3. 使用Quant-iT PicoGreen (Life Technologies P7589) 测定法确定DNA的产量。制备1:200,1:500和1:1000稀释液用于每个Hi-C文库。

[0172] 第6+7天:Hi-C连接效率和质量控制

[0173] 1. 为了检查文库的质量和数量,在0.8%琼脂糖凝胶上运行来自Hi-C文库的2μl 和6μl 1:10稀释的等分试样。文库应当运行为超过10 kb的非常紧的条带。

[0174] 2. Hi-C标记和Hi-C连接效率通过PCR消化测定法来验证。HindIII位点(AAGCTT)的成功填补和连接产生了用于限制性内切酶NheI的位点(GCTAGC)。

[0175] 设置五个(相同)25μl PCR反应以扩增短距离连接产物(由两个相邻的限制性片段形成),例如弗雷泽实验室 *Calr* 对照。使用200 ng的每个Hi-C文库作为模板。然后合并、纯化(Qiagen PCR纯化试剂盒)PCR产物,并确定浓度(NanoDrop)。随后,分成四个样品(未消化,用HindIII消化,用NheI消化,和用HindIII和NheI两者消化)。每个消化目标是500至600 ng PCR产物。在1.5%凝胶上运行消化的样品后,3C和Hi-C连接事件的相对数量可以通过量化切割和未切割的带的强度进行估计。

[0176] 3. 检查Hi-C文库中已知的长距离相互作用,诸如在小鼠红系细胞中*Hbb*和*Eraf*之间。

[0177] 第8天:从非连接的DNA末端除去生物素

[0178] 1. 在非连接的DNA末端的生物素-14-dATP用T4 DNA聚合酶的核酸外切酶活性除去。对于每个反应(每个Hi-C文库设置总共至少8个反应,对应总共40μg),混合5μg的Hi-C文库与0.5μl 10 mg/ml BSA,5μl 10×NE缓冲液2,2μl 2.5mM dATP,和5μl T4 DNA聚合酶(NEB M0203L),以50μl的总体积。在20°C下孵育4小时。

[0179] 2. 通过加入2ml 0.5 M EDTA pH8.0到每个管来停止反应。

[0180] 3. 从第8天步骤1合并各自2个反应物(DNA总量现在每个样品大约10μg)。

[0181] 为了纯化DNA,如第4天步骤3中所述进行苯酚:氯仿提取,随后乙醇沉淀。在-20°C 孵育几个小时至过夜。

[0182] 4. 在4°C以14,000rpm离心样品30分钟。弃上清,并用70%乙醇洗涤一次。弃上清,风干DNA沉淀,每个样品再悬浮于130μl H<sub>2</sub>O(在此阶段各样品应包含约10mg DNA)。

[0183] 第9天:DNA剪切和末端修复

[0184] 1. 使用Covaris E220进行超声处理。设置DNA剪切以获得具有约400bp峰的片段:

[0185] 占空比10%

[0186] 峰入射功率(W):140

[0187] 每一次发射的循环(Cycles per burst):200

[0188] 时间:55秒。

[0189] 2. 超声处理后,将各样品的整个体积(130μl)转移到新鲜的Eppendorf管中。加入以下:

- [0190] 18 $\mu$ l 10X连接缓冲液(NEB B0202S)
- [0191] 18 $\mu$ l 2.5mM的dNTP混合物
- [0192] 6.5 $\mu$ l T4 DNA聚合酶(NEB M0203L)
- [0193] 6.5 $\mu$ l T4 DNA多核苷酸激酶(NEB M0201L)
- [0194] 1.3 $\mu$ l 克列诺酶(NEB M0210L)
- [0195] 在室温下孵育30分钟。
- [0196] 3. 每个样品分成两个(每个含有约5 $\mu$ g的DNA,其是用于以下使用的柱的最大容量)。
- [0197] 根据制造商的指导用MinElute柱(Qiagen 28004)纯化DNA。用15 $\mu$ l TLE(10 mM Tris pH8.0,0.1mM EDTA)洗脱每个柱,并用另外15 $\mu$ L TLE重复洗脱。
- [0198] 第10天:dATP的添加,SPRI珠大小选择
- [0199] 1. 向来自第9天步骤3的剪切和末端修复的DNA(30 $\mu$ L)加入下列:
- [0200] 5 $\mu$ l NE缓冲液 2 10x
- [0201] 11.5 $\mu$ l dATP 1mM
- [0202] 3.5 $\mu$ l 克列诺外切酶(Klenow exo-) (NEB M0212L)
- [0203] 在37°C孵育30分钟。
- [0204] 2. 在65°C孵育20分钟以使酶失活,之后立即置于冰上。
- [0205] 3. 通过双面SPRI珠大小选择法(0.6x随后是0.9x)对200和650碱基对之间的片段进行大小选择。通过涡旋充分混合SPRI珠(Beckman Coulter Ampure XP珠 A63881),保持在室温下至少30分钟。
- [0206] a)合并来自第10天步骤1的两个A尾样品(目前总体积100 $\mu$ l)到新的管(A)。
- [0207] b)对于每个样品,用180 $\mu$ l SPRI珠溶液制备一个管(B)。
- [0208] c)从管(B)中取出60 $\mu$ l SPRI珠溶液加入具有100 $\mu$ l DNA溶液的管(A)(0.6x)。彻底混合,在室温下孵育10分钟,并放置在磁力分离器上并回收含所需大小范围内的DNA的未结合的上清液到新的管(管C)。丢弃含有珠的管(A)。
- [0209] d)浓缩SPRI珠:将管(B)(包含120 $\mu$ l 溶液中的SPRI珠)放置在磁铁上,并去除所有上清液仅剩30 $\mu$ l(即弃约90 $\mu$ l上清液)。从磁铁上取下管(B)并在剩余的30 $\mu$ l体积中重悬珠。
- [0210] e)从管(B)将30 $\mu$ l浓缩的珠(见上文)加入到管(C)(0.9x SPRI珠,即溶液中DNA相对SPRI珠的比例为1:0.9)。混合均匀,在室温下孵育至少10分钟,放置在磁铁上,并且弃上清。用新鲜制备的70%的乙醇洗两次。丢弃管(B)。
- [0211] 4. 在管(C)中在50 $\mu$ l TLE(Tris低-EDTA;10mM Tris,0.1mM EDTA)中重悬结合珠的DNA,在室温下孵育5分钟,放置在磁铁上并转移上清液(包含您的大小选择的DNA)到一个新的管(D)中。丢弃含珠的管(C)。
- [0212] 5. 合并所有的Hi-C文库样品。
- [0213] 第11天:生物素-链霉亲和素下拉(pulldown)和衔接头连接
- [0214] 1. 使用Quant-iT PicoGreen (Life Technologies P7589)测定法确定DNA的产量。制备1:20和1:50稀释液用于您的样品。从40 $\mu$ g的起始材料的预期产量大约是10 $\mu$ g。
- [0215] 为了避免过载珠(参见第11天,步骤4),建议每150 $\mu$ l链霉亲和素珠使用不大于最大2.5 $\mu$ g的DNA。

[0216] 2. 通过退火PE衔接头引物1和2生成PE衔接头:

[0217] 5'-P-GATCGGAAGAGCGGTTAGCAGGAATGCCGAG-3' (SEQ ID NO: 1)  
5'-ACACTCTTCCCTACACGACGCTTCCGATCT-3' (SEQ ID NO: 2)。

[0218] 混合15μl每种PE衔接头引物与70μl H<sub>2</sub>O(总体积100μl)。在PCR仪中,运行以下程序:95℃5分钟,然后以每分钟1℃降低温度,直至达到4℃。

[0219] 制备10至20μl的PE衔接头等分试样(15μM),储存在-20℃,并在临用前解冻等分试样。

[0220] 3. 制备缓冲液:

[0221] a) TB (吐温缓冲液): 5mM Tris, 0.5mM EDTA, 1M NaCl, 0.05% 吐温

[0222] 每个样品制备1600μl

[0223] b) NTB (无吐温缓冲液): 5mM Tris, 0.5mM EDTA, 1M NaCl

[0224] 每个样品制备600μl

[0225] c) NTB 2x: 10mM Tris, 1mM EDTA, 2M NaCl

[0226] 每个样品制备300μl

[0227] d) 连接缓冲液 (10x 连接缓冲液 NEB B0202S): 1x

[0228] 每个样品制备150μl

[0229] e) NE缓冲液2: 1x

[0230] 每个样品制备300μl。

[0231] 4. 转移150μl Dynabeads MyOne 链霉亲和素C1珠(Life Technologies 650.01)悬浮液到低结合1.5ml Eppendorf管中。

[0232] 每2μg至2.5μg的DNA设置一个反应,如第11天步骤1确定的。

[0233] 5. 用400μl TB洗涤珠两次。洗涤步骤总是如下:

[0234] a)将样品置于磁力分离器上并回收珠

[0235] b)弃上清

[0236] c)加入新的缓冲液,充分混合,并将样品转移到新的管

[0237] d)在室温下旋转样品3分钟

[0238] e) 参见a)。

[0239] 6. 在300μl NTB 2x中重悬珠。用TLE将Hi-C DNA(参见第11天步骤1)体积补足到300μl。

[0240] 在Hi-C文库DNA的量超过为2.5μg的情况下(参见第11天,步骤1和步骤4),制备足夠数量的Hi-C样品,每个在300μl TLE的总体积中包含最多2.5μg的DNA。

[0241] 组合珠与Hi-C DNA(总体积600μl)。在室温下缓慢旋转30分钟。

[0242] 7. 将样品置于磁力分离器上,回收具有结合的Hi-C DNA的珠,弃去上清液,并用400μl NTB洗涤一次,然后用200μl 1×连接缓冲液再洗涤一次。

[0243] 8. 在50μl 1x 连接缓冲液中重悬并转移到新的管。加入4μl 退火的15μM PE衔接头(参见第11天,步骤2)和4μl T4 DNA连接酶(NEB M0202S)。在室温下缓慢旋转2小时。

[0244] 9. 将样品置于磁力分离器上,回收Hi-C结合珠,并用每次400 μl TB洗涤两次。

[0245] 10. 用200μl 1xNTB洗涤,然后用200μl 1xNE缓冲液2洗涤。

[0246] 11. 用60μl的1X NE缓冲液2洗涤,然后在40μl 1X NE缓冲液2中重悬珠并转移到

新管。如果每个Hi-C文库执行超过一个链霉亲和素-生物素下拉(即,如果Hi-C文库的起始量超过2.5μgDNA,参见第11天,步骤1和4),则合并所有反应物。储存在4℃。

[0247] 第12天:测试PCR以确定用于Hi-C文库扩增的条件

[0248] 1. 为了确定用于Hi-C文库的最佳PCR循环数,设置测试PCR具有6、9、12和15个扩增循环。设置4个反应,每个具有:

- [0249] 2.5μl 珠上的Hi-C文库DNA
- [0250] 5μl 缓冲液 5x (Phusion NEB F531)
- [0251] 0.7μl dATP 10mM
- [0252] 0.7μl dCTP 10mM
- [0253] 0.7μl dGTP 10mM
- [0254] 0.7μl dTTP 10mM
- [0255] 0.075μl PE PCR 引物 1.0 (100μM)
- [0256] 0.075μl PE PCR 引物 2.0 (100μM)
- [0257] 0.3μl Phusion 聚合酶 (NEB F531)
- [0258] 14.25μl H<sub>2</sub>O

PE PCR 引物 1.0:

AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTTTCCGATCT

[0259] (SEQ ID NO: 3)

PE PCR 引物 2.0:

CAAGCAGAAGACGGCATACGAGATCGGTCTGGCATTCTGCTGAACCGCTTTCCGA

[0260] TCT (SEQ ID NO: 4)。

[0261] 2. 用下列条件运行4个不同的PCR进行n个循环:

- [0262] 1个循环: 98℃ 30 秒
- [0263] 65℃ 30 秒
- [0264] 72℃ 30 秒
- [0265] n-2个循环: 98℃ 10 秒
- [0266] 65℃ 30 秒
- [0267] 72℃ 30 秒
- [0268] 1个循环: 98℃ 10 秒
- [0269] 65℃ 30 秒
- [0270] 72℃ 7 分钟。

[0271] 3. 通过在1.5%琼脂糖凝胶上运行整个反应物(25μl)检查扩增的DNA的量。通常,在300到600bp范围内的成片条带(smear)应该刚好在9(在某些情况下,6)个扩增循环时约可见,并随着PCR循环数目增加而强度增加。

[0272] 第13天:Hi-C文库的最终PCR扩增

[0273] 1. 设置 x PCR反应(HiC文库剩余体积除以2.5),每个具有25μl,如第12天步骤1中所述,用一个PCR条件(即循环数)。PCR循环的数量将取决于Hi-C文库的可用量和测试PCR(第12天),但通常6-7个循环应当足以获得用于捕获Hi-C所需的约500ng Hi-C文库(参见下

文)。

[0274] 2. 合并来自第13天步骤1的所有个别PCR反应物。置于磁力分离器上并转移上清液到新的1.5ml低结合Eppendorf管中。在初始量的1X NE缓冲液2中重悬链霉亲和素珠(参见第11天,步骤11),并保留作为备份。确定含有扩增的Hi-C文库的上清液的体积,并通过加入1.8x体积的SPRI珠(Beckman Coulter Ampure XP珠A63881),按照制造商的指导进行纯化。重悬在100μl TLE中。

[0275] 3. 通过向来自第13天步骤2的100μl Hi-C文库加入180μl SPRI珠,重复SPRI珠纯化(如第13天步骤2中所述)。重悬在终体积20μl TLE中。

[0276] 4. 通过Bioanalyzer (Agilent) 检查Hi-C文库的质量和数量。保留50 ng至100 ng的Hi-C文库用于下一代测序。

[0277] 第14天:Hi-C文库与生物素-RNA的杂交

[0278] 1. 制备Hi-C文库(池):转移等同于500 ng的Hi-C文库的体积至1.5 ml Eppendorf管中,并用SpeedVac浓缩。保留50 ng至100 ng的Hi-C文库用于下一代测序。蒸发所有液体后,在3.4μl的H<sub>2</sub>O中重悬Hi-C DNA沉淀。加入下列组分:

[0279] 2.5μl 定制的封闭液(blocker) 1 (Agilent Technologies)

[0280] 2.5μl 定制的封闭液 2 (Agilent Technologies)

[0281] 0.6μl 定制的封闭液 3 (Agilent Technologies)

[0282] 充分重悬,转移到PCR条(strip)中(Agilent 410022),用PCR条管盖封闭(Agilent 光学盖 8x 条 401425)并保持在冰上。

[0283] 2. 通过混合下列组分制备杂交缓冲液(每个捕获Hi-C样品49μl):

[0284] 25μl SureSelect 杂交溶液 1 (Agilent Technologies)

[0285] 1μl SureSelect 杂交溶液 2 (Agilent Technologies)

[0286] 10μl SureSelect 杂交溶液 3 (Agilent Technologies)

[0287] 13μl SureSelect 杂交溶液 4 (Agilent Technologies)

[0288] 充分混合,加热到65°C持续5分钟,转移到PCR条中(Agilent 410022),用PCR条管盖封闭(Agilent 光学盖 8x 条 401425)并保持在室温。

[0289] 3. 制备生物素-RNA(诱饵):转移5μl生物素化的RNA(100ng/μl,定制的,Agilent Technologies)到1.5ml低结合Eppendorf管中。

[0290] 使用不含核酸酶的水,以制备1:4稀释的SureSelect RNase Block(Agilent Technologies),并向生物素化的RNA加入2μl RNase Block稀释液。充分混合,转移到PCR条中(Agilent 410022),用PCR条管盖封闭(Agilent 光学盖 8x 条 401425)并保持在冰上。

[0291] 4. 杂交反应:设置PCR仪器(PTC-200, MJ Research)到下列程序:

[0292] 1.95°C持续5分钟

[0293] 2.保持65°C。

[0294] PCR仪器盖必须加热。在整个过程中,动作迅速,并尽量使PCR仪器盖打开尽可能短的时间。样品的蒸发将导致未达最佳的杂交条件。

[0295] 将具有池Hi-C文库的PCR条(DNA;第14天步骤1)转移到PCR仪器中,在下面标记黑色的位置,并开始PCR程序。

[0296]

A									
B									
C									
D									DNA
E									
F									
G									
H									

[0297] 刚超过5分钟后(一旦温度已达到65℃),将具有杂交缓冲液的PCR条(来自第14天,步骤2)转移到在PCR仪器中,在下面标记灰色的位置。孵育5分钟。

[0298]

A									
B									Hyb
C									
D									DNA
E									
F									
G									
H									

[0299] 5分钟后(从PCR程序开始起10分钟),将具有生物素化RNA诱饵的PCR条(参见第14天,步骤3)转移到在PCR仪器中,在下面标记交叉影线的位置。孵育2分钟。

[0300]

A									
B									Hyb
C									
D									DNA
E									
F									RNA
G									
H									

[0301] 2分钟后,从含有杂交缓冲液的PCR条和含有生物素化RNA诱饵的PCR条上取下盖子。吸取13μl杂交缓冲液到7μl RNA诱饵中(从灰色进入交叉影线)。丢弃含有杂交缓冲液的PCR条。立即进行下一步。

[0302]

A										
B										Hyb
C										
D										DNA
E										
F										RNA
G										
H										

[0303] 从含有池Hi-C文库(DNA)的PCR条上取下盖子。吸取9μl的Hi-C文库(由于蒸发可以略少)到20μl的具有杂交缓冲液的RNA诱饵中(从黑色进入交叉影线)。丢弃含有Hi-C文库的空的PCR条。

[0304]

A										
B										
C										
D										DNA
E										
F										RNA
G										
H										

[0305] 用PCR条管盖(Agilent 光学盖 8x 条 401425)立即封闭剩余的PCR条(现在含Hi-C文库/杂交缓冲液/RNA诱饵),并在65°C孵育24小时。

[0306]

A										
B										
C										
D										
E										
F										RNA
G										
H										

[0307] 第15天:链霉亲和素-生物素下拉和洗涤

[0308] 1. 制备缓冲液:

[0309] 室温的结合缓冲液(BB, Agilent Technologies)

[0310] 室温的洗涤缓冲液I(WB I, Agilent Technologies)

- [0311] 在65°C的洗涤缓冲液II(WB II, Agilent Technologies)
- [0312] 室温的NEB2 1x (NEB B7002S)。
- [0313] 2. 洗涤磁珠:在每个捕获Hi-C样品加入60μl到1.5 ml低结合Eppendorf管之前,充分混合Dynabeads MyOne 链霉亲和素 T1 (Life Technologies 65601)。如下洗涤珠(对于所有后续洗涤步骤相同的程序):
- [0314] a) 加入200μl BB
- [0315] b) 涡旋混合(在低至中间设定)持续5秒。
- [0316] c) 将管置于Dyna1 磁力分离器上 (Life Technologies)
- [0317] d) 回收珠,弃上清。
- [0318] 重复步骤a)至d)总共3次洗涤。
- [0319] 3. 生物素-链霉亲和素下拉:链霉亲和素珠在新的低结合Eppendorf管中200μl的BB中的时候,打开PCR仪器的盖子(在PCR仪器正在运行时),吸取整个杂交反应物到含有链霉亲和素珠的管中。
- [0320] 在旋转轮上在室温下孵育30分钟。
- [0321] 4. 洗涤,30分钟后,将样品置于磁力分离器上,弃上清,在500μl WB I中重悬珠,并转移到新的管中。在室温下孵育15分钟。每2-3分钟涡旋每次5秒。
- [0322] 在磁力分离器上分离珠和缓冲液,并除去上清液。在500μl WB II(预热到65°C)中重悬并转移到新的管。
- [0323] 在65°C孵育10分钟,并且每2至3分钟涡旋(在低到中间设定)5秒。重复在WB II中洗涤总共3次,均在65°C。
- [0324] 5. 除去上清液后,在200μl 1xNEB2中重悬并立即转移到新管中。直接放回到磁力分离器上并除去上清液。
- [0325] 6. 重悬在30μl 1x NEB2中并转移到新的管。“抓(catch)”在珠上的RNA/DNA混合物杂交物现在已准备好用于PCR扩增。
- [0326] 第16天:确定用于捕获Hi-C文库扩增的PCR条件
- [0327] 1. 为了确定用于Hi-C文库的最佳PCR循环数,设置测试PCR具有6、7和9个扩增循环。设置3个反应,每个具有:
- [0328] 2.5μl 珠上的捕获Hi-C文库DNA
- [0329] 5μl 缓冲液 5x (Phusion NEB F531)
- [0330] 0.7μl dATP 10mM
- [0331] 0.7μl dCTP 10mM
- [0332] 0.7μl dGTP 10mM
- [0333] 0.7μl dTTP 10mM
- [0334] 0.075μl PE PCR 引物 1.0 (100μM)
- [0335] 0.075μl PE PCR 引物 2.0 (100μM)
- [0336] 0.3μl Phusion 聚合酶 (NEB F531)
- [0337] 14.25μl H<sub>2</sub>O

## PE PCR 引物 1.0:

AATGATAACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTCCGATCT

[0338] (SEQ ID NO: 3)

## PE PCR 引物 2.0:

CAAGCAGAAGACGGCATACGAGATCGGTCTGGCATTCTGCTGAACCGCTCTCCGA

[0339] TCT (SEQ ID NO: 4)。

[0340] 2. 用下列条件运行3个不同的PCR进行n个循环:

[0341] 1个循环: 98°C 30 秒

[0342] 65°C 30 秒

[0343] 72°C 30 秒

[0344] n-2个循环: 98°C 10 秒

[0345] 65°C 30 秒

[0346] 72°C 30 秒

[0347] 1个循环: 98°C 10 秒

[0348] 65°C 30 秒

[0349] 72°C 7 分钟。

[0350] 3. 通过在1.5%琼脂糖凝胶上运行整个反应物(25μl)检查扩增的DNA的量。通常,在300-600bp范围内的成片条带应当刚好在6-7个扩增循环时约可见。

[0351] 第17天:捕获Hi-C文库的最终PCR扩增

[0352] 1. 设置 x PCR反应(捕获HiC文库剩余体积除以2.5),每个具有25μl,如第16天步骤1中所述,用一个PCR条件(即,选择的PCR循环数)。通常,4个PCR循环应当足以获得适合用于Illumina测序的捕获Hi-C文库的量。

[0353] 2. 合并来自第17天步骤1的所有个别PCR反应物。置于磁力分离器上并转移上清液到新的1.5ml低结合Eppendorf管中。在初始量的1X NE缓冲液2中重悬链霉亲和素珠(参见第15天,步骤6),并保留作为备份。确定捕获Hi-C文库的体积,并通过加入1.8倍体积的SPRI珠(Beckman Coulter Ampure XP珠A63881),按照制造商的指导进行纯化。重悬在100μl TLE中。

[0354] 3. 通过向来自第17天步骤2的100μl Hi-C文库加入180μl SPRI珠,重复SPRI珠纯化(如第17天步骤2中所述)。重悬在终体积20μl TLE中。

[0355] 4. 在测序前,通过Bioanalyzer (Agilent) 和定量PCR检查捕获Hi-C文库的质量和数量。

[0356] 实施例2:与诱饵基因座(SCRiBL)Hi-C方案相互作用的区域的序列捕获

[0357] 本研究描述了下列SCRiBL Hi-C的三个变体,其中的两个(下面的b和c)允许样品加条形码和多路复用(multiplexing):

[0358] a) 常规SCRiBL Hi-C(未加条形码的)

[0359] b) 捕获前加条形码的SCRiBL Hi-C

[0360] c) 捕获后加条形码的SCRiBL Hi-C。

[0361] 按照实施例1中方案从第1到10天,然后继续进行下列步骤:

[0362] a) 常规SCRiBL Hi-C:

[0363] 按照实施例1中方案从第11到13天,然后继续进行下面的步骤14。

[0364] b) 捕获前加条形码的SCRiBL Hi-C:

[0365] 第11天:生物素-链霉亲和素下拉(pulldown)和衔接头连接

[0366] 1. 使用Quant-iT PicoGreen (Life Technologies P7589) 测定法确定DNA的产量。制备样品的1:20和1:50稀释液。从40μg的起始材料的预期产量大约是10μg。

[0367] 为了避免过载珠(参见第11天,步骤4),建议每150μl链霉亲和素珠使用不大于最大2.5μg的DNA。

[0368] 2. 通过使TruSeq通用衔接头引物和TruSeq索引的衔接头引物退火生成TruSeq衔接头:

5'-P-

[0369] GATCGGAAGAGCACACGTCTGAACTCCAGTCACNNNNNNATCTGTATGCCGTCTTCT  
GCTTG-3' (SEQ ID NO: 5)

[0370] (应当注意“NNNNNN”是条形码)

5'-

[0371] AATGATACGGCGACCACCGAGATCTACACTTTCCCTACACGACGCTTCCGATCT-  
3' (SEQ ID NO: 3)

[0372] 混合15μl每种TruSeq衔接头引物与70μl H<sub>2</sub>O(总体积100μl)。在PCR仪中,运行以下程序:95°C 5分钟,然后以每分钟1°C降低温度,直至达到4°C。

[0373] 制备10至20μl的TruSeq衔接头等分试样(15μM),储存在-20°C,并在临用前解冻等分试样。

[0374] 3. 制备缓冲液:

[0375] a) TB (吐温缓冲液): 5mM Tris, 0.5mM EDTA, 1M NaCl, 0.05% 吐温

[0376] 每个样品制备1600μl

[0377] b) NTB (无吐温缓冲液): 5mM Tris, 0.5mM EDTA, 1M NaCl

[0378] 每个样品制备600μl

[0379] c) NTB 2x: 10mM Tris, 1mM EDTA, 2M NaCl

[0380] 每个样品制备300μl

[0381] d) 连接缓冲液 (10x 连接缓冲液 NEB B0202S): 1x

[0382] 每个样品制备150μl

[0383] e) NE缓冲液2: 1x

[0384] 每个样品制备300μl。

[0385] 4. 转移150μl Dynabeads MyOne 链霉亲和素C1珠(Life Technologies 650.01)悬浮液到低结合1.5ml Eppendorf管中。

[0386] 每2μg至2.5μg的DNA设置一个反应,如第11天步骤1确定的。

[0387] 5. 用400μl TB洗涤珠两次。洗涤步骤总是如下:

[0388] a) 将样品置于磁力分离器上并回收珠

[0389] b) 弃上清

- [0390] c) 加入新的缓冲液,充分混合,并将样品转移到新的管
- [0391] d) 在室温下旋转样品3分钟
- [0392] e) 参见a)。
- [0393] 6. 在300 $\mu$ l NTB 2x中重悬珠。用TLE将Hi-C DNA(参见第11天步骤1)体积补足到300 $\mu$ l。
- [0394] 在Hi-C文库DNA的量超过为2.5 $\mu$ g的情况下(参见第11天,步骤1和步骤4),制备足够数量的Hi-C样品,每个在300 $\mu$ l TLE的总体积中包含最多2.5 $\mu$ g的DNA。
- [0395] 组合珠与Hi-C DNA(总体积600 $\mu$ l)。在室温下缓慢旋转30分钟。
- [0396] 7. 将样品置于磁力分离器上,回收具有结合的Hi-C DNA的珠,弃去上清液,并用400 $\mu$ l NTB洗涤一次,然后用200 $\mu$ l 1 $\times$ 连接缓冲液再洗涤一次。
- [0397] 8. 重悬在50 $\mu$ l 1 $\times$ 连接缓冲液中并转移到新的管。加入4 $\mu$ l 退火的15 $\mu$ M TruSeq衔接头(参见第11天,步骤2)和4 $\mu$ l T4 DNA连接酶(NEB M0202S)。在室温下缓慢旋转2小时。
- [0398] 9. 将样品置于磁力分离器上,回收Hi-C结合珠,并每次用400  $\mu$ l TB洗涤两次。
- [0399] 10. 用200 $\mu$ l 1xNTB洗涤,然后用200 $\mu$ l 1xNE缓冲液2洗涤。
- [0400] 11. 用60 $\mu$ l的1X NE缓冲液2洗涤,然后在40 $\mu$ l 1X NE缓冲液2中重悬珠并转移到新管。如果每个Hi-C文库执行超过一个链霉亲和素-生物素下拉(即,如果Hi-C文库的起始量超过2.5 $\mu$ gDNA,参见第11天,步骤1和4),则合并所有反应物。储存在4℃。

**[0401] 第12天:测试PCR以确定用于Hi-C文库扩增的条件**

- [0402] 1. 为了确定用于Hi-C文库的最佳PCR循环数,设置测试PCR具有6、9、12和15个扩增循环。设置4个反应,每个具有:

- |        |               |                                 |
|--------|---------------|---------------------------------|
| [0403] | 2.5 $\mu$ l   | 珠上的Hi-C文库DNA                    |
| [0404] | 5 $\mu$ l     | 缓冲液 5x (Phusion NEB F531)       |
| [0405] | 0.7 $\mu$ l   | dATP 10mM                       |
| [0406] | 0.7 $\mu$ l   | dCTP 10mM                       |
| [0407] | 0.7 $\mu$ l   | dGTP 10mM                       |
| [0408] | 0.7 $\mu$ l   | dTTP 10mM                       |
| [0409] | 0.075 $\mu$ l | TruSeq PCR 引物 1.0 (100 $\mu$ M) |
| [0410] | 0.075 $\mu$ l | TruSeq PCR 引物 2.0 (100 $\mu$ M) |
| [0411] | 0.3 $\mu$ l   | Phusion 聚合酶(NEB F531)           |
| [0412] | 14.25 $\mu$ l | H <sub>2</sub> O                |

TruSeq PCR 引物 1.0:

AATGATAACGGCGACCACCGAGATCTACACTCTTCCCTACACGA

- [0413] (SEQ ID NO: 6)

TruSeq PCR 引物 2.0:

- [0414] CAAGCAGAAGACGGCATACGAGAT (SEQ ID NO: 7)。

- [0415] 2. 用下列条件运行4个不同的PCR进行n个循环:

[0416] 1个循环: 98℃ 30 秒

[0417] 62℃ 30 秒

- [0418] 72℃ 30 秒
- [0419] n-2个循环: 98℃ 10 秒
- [0420] 62℃ 30 秒
- [0421] 72℃ 30 秒
- [0422] 1个循环: 98℃ 10 秒
- [0423] 62℃ 30 秒
- [0424] 72℃ 7 分钟。

[0425] 3. 通过在1.5%琼脂糖凝胶上运行整个反应物(25μl)检查扩增的DNA的量。通常，在300到600bp范围内的成片条带应该刚好在9(在某些情况下,6)个扩增循环时约可见，并随着PCR循环数目增加而强度增加。

[0426] 第13天:Hi-C文库的最终PCR扩增

[0427] 1. 设置 x PCR反应(HiC文库剩余体积除以2.5), 每个具有25μl, 如第12天步骤1中所述, 用一个PCR条件(即循环数)。PCR循环的数量将取决于Hi-C文库的可用量和测试PCR(第12天), 但通常6-7个循环应当足以获得用于捕获Hi-C所需的约500ng Hi-C文库(参见下文)。

[0428] 2. 合并来自第13天步骤1的所有个别PCR反应物。置于磁力分离器上并转移上清液到新的1.5ml低结合Eppendorf管中。在初始量的1X NE缓冲液2中重悬链霉亲和素珠(参见第11天, 步骤11), 并保留作为备份。确定含有扩增的Hi-C文库的上清液的体积, 并通过加入1.8x体积的SPRI珠(Beckman Coulter Ampure XP珠A63881), 按照制造商的指导进行纯化。重悬在100μl TLE中。

[0429] 3. 通过向来自第13天步骤2的100μl Hi-C文库加入180μl SPRI珠, 重复SPRI珠纯化(如第13天步骤2中所述)。重悬在终体积20μl TLE中。

[0430] 4. 通过Bioanalyzer (Agilent) 检查Hi-C文库的质量和数量。保留50 ng至100 ng的Hi-C文库用于下一代测序。

[0431] 前进到第14天

[0432] c) 捕获后加条形码的SCRiBL Hi-C:

[0433] 第11天:生物素-链霉亲和素下拉(pulldown)和衔接头连接

[0434] 1. 使用Quant-iT PicoGreen (Life Technologies P7589) 测定法确定DNA的产量。制备样品的1:20和1:50稀释液。从40μg的起始材料的预期产量大约是10μg。

[0435] 为了避免过载珠(参见第11天, 步骤4), 建议每150μl链霉亲和素珠使用不大于最大2.5μg的DNA。

[0436] 2. 通过退火SCRiBL衔接头引物1和2生成SCRiBL衔接头:

5'-P-GATCGGAAGAGCACACGTCTGAACTCCAGTCAC-3' (SEQ ID NO: 8)

[0437] 5'-ACACTCTTCCCTACACGACGCTTCCGATCT-3' (SEQ ID NO: 2) 。

[0438] 混合15μl每种SCRiBL衔接头引物与70μl H<sub>2</sub>O(总体积100μl)。在PCR仪中, 运行以下程序: 95℃ 5分钟, 然后以每分钟1℃降低温度, 直至达到4℃。

[0439] 制备10至20μl的SCRiBL衔接头等分试样(15μM), 储存在-20℃, 并在临用前解冻等分试样。

- [0440] 3. 制备缓冲液：
- [0441] a) TB (吐温缓冲液)：5mM Tris, 0.5mM EDTA, 1M NaCl, 0.05% 吐温
- [0442] 每个样品制备1600μl
- [0443] b) NTB (无吐温缓冲液)：5mM Tris, 0.5mM EDTA, 1M NaCl
- [0444] 每个样品制备600μl
- [0445] c) NTB 2x: 10mM Tris, 1mM EDTA, 2M NaCl
- [0446] 每个样品制备300μl
- [0447] d) 连接缓冲液 (10x 连接缓冲液 NEB B0202S)：1x
- [0448] 每个样品制备150μl
- [0449] e) NE缓冲液2: 1x
- [0450] 每个样品制备300μl。
- [0451] 4. 转移150μl Dynabeads MyOne 链霉亲和素C1珠 (Life Technologies 650.01) 悬浮液到低结合1.5ml Eppendorf管中。
- [0452] 每2μg至2.5μg的DNA设置一个反应,如第11天步骤1确定的。
- [0453] 5. 用400μl TB洗涤珠两次。洗涤步骤总是如下：
- [0454] a) 将样品置于磁力分离器上并回收珠
- [0455] b) 弃上清
- [0456] c) 加入新的缓冲液,充分混合,并将样品转移到新的管
- [0457] d) 在室温下旋转样品3分钟
- [0458] e) 参见a)。
- [0459] 6. 在300μl NTB 2x中重悬珠。用TLE将Hi-C DNA(参见第11天步骤1)体积补足到300μl。
- [0460] 在Hi-C文库DNA的量超过为2.5μg的情况下(参见第11天,步骤1和步骤4),制备足够数量的Hi-C样品,每个在300μl TLE的总体积中包含最多2.5μg的DNA。
- [0461] 组合珠与Hi-C DNA(总体积600μl)。在室温下缓慢旋转30分钟。
- [0462] 7. 将样品置于磁力分离器上,回收具有结合的Hi-C DNA的珠,弃去上清液,并用400μl NTB洗涤一次,然后用200μl 1×连接缓冲液再洗涤一次。
- [0463] 8. 重悬在50μl 1x 连接缓冲液中并转移到新的管。加入4μl 退火的15μM SCRiBL 衔接头(参见第11天,步骤2)和4μl T4 DNA连接酶 (NEB M0202S)。在室温下缓慢旋转2小时。
- [0464] 9. 将样品置于磁力分离器上,回收Hi-C结合珠,并每次用400 μl TB洗涤两次。
- [0465] 10. 用200μl 1xNTB洗涤,然后用200μl 1xNE缓冲液2洗涤。
- [0466] 11. 用60μl的1X NE缓冲液2洗涤,然后在40μl 1X NE缓冲液2中重悬珠并转移到新管。如果每个Hi-C文库执行超过一个链霉亲和素-生物素下拉(即,如果Hi-C文库的起始量超过2.5μgDNA,参见第11天,步骤1和4),则合并所有反应物。储存在4℃。
- [0467] 第12天:测试PCR以确定用于Hi-C文库扩增的条件
- [0468] 1. 为了确定用于Hi-C文库的最佳PCR循环数,设置测试PCR具有6、9、12和15个扩增循环。设置4个反应,每个具有:
- [0469] 2.5μl 珠上的Hi-C文库DNA
- [0470] 5μl 缓冲液 5x (Phusion NEB F531)

[0471]	0.7μl	dATP 10mM
[0472]	0.7μl	dCTP 10mM
[0473]	0.7μl	dGTP 10mM
[0474]	0.7μl	dTTP 10mM
[0475]	0.075μl	SCRiBL 捕获前PCR 引物 1.0 (100μM)
[0476]	0.075μl	SCRiBL 捕获前PCR 引物 2.0 (100μM)
[0477]	0.3μl	Phusion 聚合酶 (NEB F531)
[0478]	14.25μl	H <sub>2</sub> O

**SCRiBL 预捕获 PCR 引物 1.0:**

[0479] **ACACTCTTCCCTACACGACGCTCTCCGATCT** (SEQ ID NO: 2)

**SCRiBL 预捕获 PCR 引物 2.0:**

[0480] **GTGACTGGAGTTCAGACGTGTGCTCTCCGATC** (SEQ ID NO: 9)。

[0481] 2. 用下列条件运行4个不同的PCR进行n个循环:

[0482] 1个循环: 98°C 30 秒

[0483] 65°C 30 秒

[0484] 72°C 30 秒

[0485] n-2个循环: 98°C 10 秒

[0486] 65°C 30 秒

[0487] 72°C 30 秒

[0488] 1个循环: 98°C 10 秒

[0489] 65°C 30 秒

[0490] 72°C 7 分钟。

[0491] 3. 通过在1.5%琼脂糖凝胶上运行整个反应物(25μl)检查扩增的DNA的量。通常，在300到600bp范围内的成片条带应该刚好在9(在某些情况下,6)个扩增循环时约可见，并随着PCR循环数目增加而强度增加。

[0492] 第13天:Hi-C文库的最终PCR扩增

[0493] 1. 设置 x PCR反应(HiC文库剩余体积除以2.5),每个具有25μl,如第12天步骤1中所述,用一个PCR条件(即循环数)。PCR循环的数量将取决于Hi-C文库的可用量和测试PCR(第12天),但通常6-7个循环应当足以获得用于捕获Hi-C所需的约500ngHi-C文库(参见下文)。

[0494] 2. 合并来自第13天步骤1的所有个别PCR反应物。置于磁力分离器上并转移上清液到新的1.5ml低结合Eppendorf管中。在初始量的1X NE缓冲液2中重悬链霉亲和素珠(参见第11天,步骤11),并保留作为备份。确定含有扩增的Hi-C文库的上清液的体积,并通过加入1.8x体积的SPRI珠(Beckman Coulter Ampure XP珠A63881),按照制造商的指导进行纯化。重悬在100μl TLE中。

[0495] 3. 通过向来自第13天步骤2的100μlHi-C文库加入180μlSPRI珠,重复SPRI珠纯化(如第13天步骤2中所述)。重悬在终体积20μl TLE中。

[0496] 4. 通过Bioanalyzer (Agilent) 检查Hi-C文库的质量和数量。保留50 ng至100

ng的Hi-C文库用于下一代测序。

[0497] 前进到第14天

[0498] 用于靶标富集的生物素化RNA的生成(用于常规SCRiBL Hi-C、捕获前加条形码的SCRiBL Hi-C和捕获后加条形码的SCRiBL Hi-C)

[0499] 第14天:细菌人工染色体(BACs)的消化

[0500] 1. 选择并定购(order)覆盖目标基因组区域的细菌人工染色体(BACs)。使用Macherey Nagel NucleoBond BAC 100试剂盒(REF 740579)按照制造商的指导制备BAC DNA。

[0501] 2. 以200至250 $\mu$ l的总体积,用在NEB缓冲液2(1×最终)中的400U的HindIII在37℃过夜消化等摩尔量的BAC DNA(总共20至25 $\mu$ g)。

[0502] 注意:所用的限制性内切酶必须与实施例1 Hi-C文库生成的第1至13天中相同。

[0503] 第15天:T7启动子衔接头与消化的BAC DNA的连接

[0504] 1. 进行苯酚氯仿提取,随后是氯仿提取。用0.1体积的3M 乙酸钠和2.5体积的100%乙醇在-20℃沉淀最少2小时(或在-20℃过夜)。

[0505] 2. 在4℃以14,000rpm离心20分钟。用800 $\mu$ l 70%乙醇洗涤,再离心,去除乙醇并允许沉淀风干。

[0506] 3. 在25 $\mu$ l 10mM Tris, pH7.5中重悬。通过Nanodrop确定浓度。产量应当为15-20 $\mu$ g DNA。

[0507] 4. 为了制备T7启动子衔接头,混合“Luo T7 HindIII”和“Luo T7 HindIII rc”引物两者各20 $\mu$ l(100 $\mu$ M)与60 $\mu$ l的寡核苷酸(oligo)退火缓冲液(10 mM Tris, pH8.0; 50mM NaCl; 1mM EDTA),使得最终衔接头浓度为20 $\mu$ M。

[0508] 使用HPLC纯化引物。引物序列可以见于这部分方案的末尾。

[0509] 5. 在PCR仪器中退火衔接头:95℃持续5分钟,然后以每分钟1℃的速率降低温度,直至达到4℃。退火的衔接头应当以等分试样冷冻。在实验之前在冰上融化并立即使用。

[0510] 6. 加入15倍摩尔过量的衔接头(6.2 $\mu$ l假设用于20 $\mu$ g的BAC DNA,约4kb的片段)到在1×T4 DNA连接酶缓冲液(NEB B0202S)中的DNA,所述1×T4 DNA连接酶缓冲液是150 $\mu$ L的最终体积中的12 $\mu$ l(400U/ $\mu$ l)T4 DNA连接酶(NEB M0202S)。在16℃孵育过夜。

[0511] 第16天:消化的T7启动子-DNA的超声处理

[0512] 1. 通过在65℃加热10分钟灭活T4 DNA连接酶,在冰上冷却。

[0513] 2. DNA(连接到T7启动子的BAC HindIII片段)现在将是150 $\mu$ L。加入120 $\mu$ l H<sub>2</sub>O以得到270 $\mu$ l。移去10 $\mu$ l以在凝胶上运行(超声处理前样品),并将其余体积分成各130 $\mu$ l(对于通过Covaris E220超声处理所需的体积)两个样品。

[0514] 3. 通过隔垫(septum)吸取进入加金属盖的covaris裸露DNA管(metal-lidded covaris naked DNA tubes)中,避免产生可能干扰超声处理的气泡。

[0515] 4. 使用下列Covaris E220条件以生成具有以200碱基对的峰的片段:

[0516] 10%占空比

[0517] 175W峰入射功率

[0518] 每次发射(brust) 200个循环

[0519] 180秒的处理时间

[0520] 7 °C。

[0521] 5. 超声处理后,将样品从covaris管转移到新的Eppendorf管中。

[0522] 6. 任选:在2%琼脂糖凝胶上运行超声处理前等分试样和超声处理后等分试样,以评价片段化。

[0523] 第17天:末端修复和SPRI珠大小选择

[0524] 1. 回收的超声处理的DNA的体积应当是每个样品130μl。添加17μl 2.5mM 混合的dNTPs, 9.5μl 10X NEB连接酶缓冲液(NEB B0202S),使得终浓度为1x, 6μl T4 DNA 聚合酶(NEB M0203L), 6μl 多核苷酸激酶(NEB M0201L), 1.5μl 克列诺酶(NEB M0210L),达到每个反应170μl的终体积。在室温下孵育30分钟。

[0525] 2. 分开两个170μl的反应物的每一个在两个Eppendorf管(每个包含85μl)中(以避免在后续的步骤中过量加载Qiagen纯化柱)。使用Qiagen PCR纯化试剂盒(Qiagen 28104)纯化在所有四个管中的DNA,每个用50μl洗脱,并且然后在最终洗脱后组合各自两个样品(对于两个样品最后的总体积应当是总共100μl)。

[0526] 3. 执行双侧SPRI珠(Agencourt AMPure XP珠A63881)大小选择(0.7x,随后是1.0x)以分离180至300bp大小范围内的BAC-T7启动子的DNA。超声处理和修复的DNA的回收体积现在应该是100μl(管A)。通过涡旋充分混合SPRI珠,并转移180μl的SPRI珠到新的Eppendorf管中(管B)。将此180μl SPRI珠的70μL加入到管(A)中的100μl DNA中(0.7X),彻底混合,在室温下孵育10分钟,并放置在磁力分离器上并回收含所需大小范围内的DNA的未结合的上清液到新的管(管C)。丢弃含有珠的管(A)。

[0527] 4. 此步骤是为了浓缩珠。将管(B)(包含110μl的SPRI珠)放置在磁铁上,并去除所有上清液仅剩30μl(即弃约80μl上清液)。从磁铁上取下管(B)并在剩余的30μl体积中重悬珠。

[0528] 5. 从管(B)将30μl浓缩的珠加入到管(C)(1.0x SPRI珠,即溶液中DNA与溶液中SPRI珠的比例现在为1:1)。充分混合,在室温下孵育10分钟,放置在磁铁上,并且弃上清。丢弃管(B)。

[0529] 6. 在管(C)中在25μl TLE(Tris低-EDTA;10mM Tris,0.1mM EDTA)中重悬结合珠的DNA,在室温下孵育5分钟,放置在磁铁并转移上清液(包含您的大小选择的DNA)到一个新的管(D)中。丢弃含珠的管(C)。

[0530] 7. 确定大小选择的DNA的浓度。

[0531] 第18天:用生物素-UTP体外转录

[0532] 1. 使用T7 MegaScript试剂盒(Ambion; AM1333)和生物素RNA标记混合物(Roche; 目录号11 685 597 910),或者(更便宜的选项)和生物素-UTP(Roche; 目录号11 388 908 910),并与标准rNTPs混合。设置下列反应:

[0533] 2μl 10X 缓冲液

[0534] 5.5μl DNA 模板(通常100–500 ng)

[0535] 5μl 生物素-UTP(Roche 11 388 908 910 原液是 10 mM)

[0536] 1μl 未标记的 rUTP(100 mM 原液)

[0537] 1.5μl rATP(100 mM 原液)

[0538] 1.5μl rCTP(100 mM 原液)

[0539] 1.5 $\mu$ l rGTP (100 mM 原液)

[0540] 2 $\mu$ l 酶混合物

[0541] 注意:这导致对于每种核苷酸7.5mM的最终核苷酸浓度,并且标记与未标记的UTP的比例是1:2。

[0542] 2. 反应物在37°C孵育过夜(12至16小时)。

[0543] 第19天:生物素化RNA的纯化

[0544] 1. 为了除去模板DNA,用1 $\mu$ l Turbo DNase (Life Technologies AM2238) 在37°C处理15分钟。

[0545] 2. 使用Ambion MEGAclear试剂盒(AM1908M)按照制造商的指导纯化RNA。在50 $\mu$ l的终体积中洗脱。制备5–10 $\mu$ l等分试样并尽快冷冻。预期的产量将通常为20 $\mu$ g–50 $\mu$ g RNA,如通过Nanodrop确定。这个量足以用于40–100个单独的SCRiBL Hi-C反应(每个实验使用500ng生物素化的RNA诱饵)。

[0546] 3. 通过在2%琼脂糖凝胶上(使用Ambion MegaScript试剂盒提供的甲酰胺上样染料)运行1至2 $\mu$ gRNA检查获得的RNA的大小和完整性。

[0547] 衔接头引物(HindIII 5'-突出端加下划线,当使用其它限制性内切酶时相应修饰):

Luo T7 HindIII:

5'-TCTAGTCGACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGA-3'

[0548] (SEQ ID NO: 10)

LuoT7 HindIII rc:

5'-[Phos]AGCTTCGCCCTATAGTGAGTCGTATTACAATTCACTGGCCGTCGACTAGA[Spc

[0549] C3]-3' (SEQ ID NO: 11)。

[0550] Hi-C文库的溶液杂交捕获

[0551] 第20天:Hi-C文库与生物素-RNA靶诱饵的杂交

[0552] 1. 制备Hi-C文库(池):转移等同于500 ng的Hi-C文库(实施例1的第13天步骤4)的体积至1.5ml低结合 Eppendorf管中,并用SpeedVac浓缩。

[0553] 保留50 ng至100 ng的Hi-C文库用于下一代测序。

[0554] 所有的液体蒸发后,在3.5 $\mu$ l H<sub>2</sub>O中重悬Hi-C DNA沉淀并转移到新的1.5ml低结合管(每单个SCRiBL反应)。加入下列组分:

[0555] 2.5 $\mu$ l cot1 DNA (小鼠 cot-1 DNA Invitrogen 18440-016, 1mg/ml;总共2.5  $\mu$ g)

[0556] 1 $\mu$ l 鲑鱼精 DNA (经剪切的, Ambion AM9680 10 mg/ml 原液, 制备 1:4 稀释液 (2.5 mg/ml),使得添加的最终量是2.5 $\mu$ g)。

[0557] 添加下列阻断寡核苷酸(blocking oligos):

[0558] a) 常规SCRiBL Hi-C:

[0559] 对于使用标准PE衔接头和PE PCR 1.0和PE PCR 2.0用于扩增生成的Hi-C文库,使用下列两种阻断寡核苷酸:

[0560] 1.5 $\mu$ l PE 阻断寡核苷酸 1 (200  $\mu$ M 原液;与PE PCR 1.0序列相同):

- [0561] AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTCCGATCT  
(SEQ ID NO: 3)
- [0562] 1.5μl PE 阻断寡核苷酸 2 (200 μM 原液;与PE PCR 2.0序列相同):  
CAAGCAGAAGACGGCATACGAGATCGGTCTCGCATTCTGCTGAACCGCTCTCCGA  
TCT (SEQ ID NO: 4)。
- [0564] b) 捕获前加条形码的SCRiBL Hi-C:
- [0565] 对于使用Illumina TruSeq衔接头和用TruSeq PCR 引物 1.0 和TruSeq PCR 引物 2.0 PCR扩增生成的Hi-C文库,使用下列两种阻断寡核苷酸:
- [0566] 1.5μl TruSeq 通用阻断寡核苷酸(200 μM 原液;与TruSeq通用衔接头序列相同):  
AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTCCGATCT  
(SEQ ID NO: 3)
- [0568] 1.5μl TruSeq 阻断寡核苷酸索引(index) x (200 μM 原液;TruSeq衔接头索引x的反向互补物):  
CAAGCAGAAGACGGCATACGAGATNNNNNNNTGACTGGAGTTCAGACGTGTGCTCTT  
CCGATC (SEQ ID NO: 12)
- [0570] 注意:使用对应用于生成Hi-C文库的Truseq衔接头索引的Truseq阻断寡核苷酸索引。例如,如果您已经使用Truseq衔接头索引6用于Hi-C文库生成:  
GATCGGAAGAGCACACGTCTGAACTCCAGTCACGCCAATATCTCGTATGCCGTCTCT  
[0571] GCTTG (SEQ ID NO: 13)
- [0572] 在这种情况下,使用Truseq阻断寡核苷酸索引6(Truseq衔接头索引6的反向互补物)作为阻断寡核苷酸:
- [0573] CAAGCAGAAGACGGCATACGAGATATTGGCGTGACTGGAGTTCAGACGTGTGCTCTC  
CGATC (SEQ ID NO: 14)。
- [0574] c) 捕获后加条形码的SCRiBL Hi-C:
- [0575] 对于使用SCRiBL衔接头和用SCRiBL捕获前PCR 引物 1.0 和SCRiBL捕获前PCR 引物 2.0 PCR扩增生成的Hi-C文库,使用下列四种阻断寡核苷酸:
- [0576] 0.75μl SCRiBL 阻断寡核苷酸 1 (200 μM 原液):  
[0577] GTGACTGGAGTTCAGACGTGTGCTCTCCGATCdd\* (SEQ ID NO: 15)
- [0578] 0.75μl SCRiBL 阻断寡核苷酸 2 (200 μM 原液):  
[0579] AGATCGGAAGAGCGTCGTAGGGAAAGAGTGT (SEQ ID NO: 16)
- [0580] 0.75μl SCRiBL 阻断寡核苷酸 3 (200 μM 原液):  
[0581] AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC (SEQ ID NO: 17)
- [0582] 0.75μl SCRiBL 阻断寡核苷酸 4 (200 μM 原液):  
[0583] ACACCTTTCCCTACACGACGCTCTCCGATCdd\* (SEQ ID NO: 18)

- [0584] \* SCRiBL 阻断寡核苷酸 1 和 4 包含 3'双脱氧修饰
- [0585] 充分重悬,转移到PCR条中(Agilent 410022),用PCR条管盖封闭(Agilent 光学盖 8x 条 401425)并保持在冰上。
- [0586] 2. 制备2.23x杂交缓冲液(每个SCRiBL Hi-C样品30μl);
- [0587] 11.15x SSPE (原液 20x;Gibco 15591-043)
- [0588] 11.15x Denhardt's (原液 50x;Invitrogen 750018)
- [0589] 11.15mM EDTA (原液 500mM;Gibco 15575-038)
- [0590] 0.223% SDS (原液 10%;Promega V6551)
- [0591] 充分混合,加热到65°C持续5分钟,转移到PCR条中(Agilent 410022),用PCR条管盖封闭(Agilent 光学盖 8x 条 401425)并保持在室温。
- [0592] 3. 制备生物素-RNA(诱饵):转移500ng的生物素化的RNA(第19天,步骤2)到1.5 ml低结合的Eppendorf管中,并用H<sub>2</sub>O补足5.5μl的体积。添加1.5μl SUPERase-In(Ambion AM2694,20单位/μl)。转移到PCR条中(Agilent 410022),用PCR条管盖封闭(Agilent 光学盖 8x 条 401425)并保持在冰上(体积现在应当是7μl)。
- [0593] 4. 杂交反应:设置PCR仪器(PTC-200, MJ Research)到下列程序(全程必须加热PCR仪器盖):
- [0594] 1.95°C持续5分钟
- [0595] 2.保持65°C
- [0596] 按照与实施例1的第14天步骤4相同的指导用于转移PCR条。
- [0597] 第21天:链霉亲和素-生物素下拉和洗涤
- [0598] 1. 制备缓冲液:
- [0599] a) 结合缓冲液 (BB) :
- [0600] 1M NaCl
- [0601] 10mM Tris-HCl pH 7.5
- [0602] 1mM EDTA
- [0603] 每个SCRiBL样品制备800μl并保持在室温
- [0604] b) 洗涤缓冲液 I (WB I) :
- [0605] 1xSSC (SSC 20x 原液: Invitrogen 15557-044)
- [0606] 0.1% SDS
- [0607] 每个SCRiBL样品制备500μl并保持在室温
- [0608] c) 洗涤缓冲液 II (WB II) :
- [0609] 0.1xSSC (SSC 20x 原液: Invitrogen 15557-044)
- [0610] 0.1% SDS
- [0611] 每个SCRiBL样品制备1500μl,加热到65°C并保持在此温度
- [0612] d) NEB2 1x (10x 原液;NEB B7002S)
- [0613] 每个SCRiBL样品制备230μl并保持在室温。
- [0614] 2. 洗涤磁珠:在每个SCRiBL Hi-C样品加入60μl到1.5 ml低结合Eppendorf管之前,充分混合Dynabeads MyOne 链霉亲和素 T1 (Life Technologies 65601)。如下洗涤珠(对于所有后续洗涤步骤相同的程序):

- [0615] a) 加入200 $\mu$ l BB
- [0616] b) 涡旋混合(在低至中间设定)持续5秒。
- [0617] c) 将管置于Dyna1 磁力分离器上 (Life Technologies)
- [0618] d) 回收珠,弃上清。
- [0619] 重复步骤a)至d)总共3次洗涤。
- [0620] 3. 生物素-链霉亲和素下拉:链霉亲和素珠在新的低结合Eppendorf管中200 $\mu$ l的BB中的时候,打开PCR仪器的盖子(同时PCR仪器仍然在运行),吸取整个杂交反应物到含有链霉亲和素珠的管中并充分混合。
- [0621] 在该点检查杂交反应的体积。低于20 $\mu$ l的体积意味着过度蒸发可能已导致未达最佳的杂交条件。
- [0622] 在旋转轮上在室温下孵育30分钟。
- [0623] 4. 洗涤,30分钟后,将样品置于磁力分离器上,弃上清,在500 $\mu$ l WB I中重悬珠,并转移到新的管中。在室温下孵育15分钟。每2-3分钟涡旋每次5秒。
- [0624] 在磁力分离器上分离珠和缓冲液,并除去上清液。在500 $\mu$ l WB II(预热到65°C)中重悬并转移到新的管。
- [0625] 在65°C孵育10分钟,并且每2至3分钟涡旋(在低到中间设定)5秒。重复在WB II中总共3次洗涤,均在65°C。
- [0626] 5. 除去上清液后,在200 $\mu$ l 1xNEB2中重悬并立即转移到新管中。直接放回到磁力分离器上并除去上清液。
- [0627] 6. 重悬在30 $\mu$ l 1x NEB2中并转移到新的管。“抓”在珠上的RNA/DNA混合物杂交物现在已准备好用于PCR扩增。
- [0628] 第22天:确定用于SCRiBL Hi-C文库扩增的PCR条件
- [0629] a) 常规SCRiBL Hi-C:
- [0630] 1. 为了确定用于SCRiBL Hi-C文库扩增的最佳PCR循环数,设置测试PCR具有7、9和12个扩增循环。设置3个反应,每个具有:
- [0631] 2.5 $\mu$ l 珠上的SCRiBL Hi-C文库DNA
- [0632] 5 $\mu$ l 缓冲液 5x (Phusion NEB F531)
- [0633] 0.7 $\mu$ l dATP 10mM
- [0634] 0.7 $\mu$ l dCTP 10mM
- [0635] 0.7 $\mu$ l dGTP 10mM
- [0636] 0.7 $\mu$ l dTTP 10mM
- [0637] 0.075 $\mu$ l PE PCR 引物 1.0 (100 $\mu$ M)
- [0638] 0.075 $\mu$ l PE PCR 引物 2.0 (100 $\mu$ M)
- [0639] 0.3 $\mu$ l Phusion 聚合酶 (NEB F531)
- [0640] 14.25 $\mu$ l H<sub>2</sub>O

PE PCR 引物 1.0:

AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTCCGATCT

[0641] (SEQ ID NO: 3)

PE PCR 引物 2.0:

CAAGCAGAAGACGGCATACGAGATCGGTCTCGCATTGCTGAACCGCTCTCCGA

[0642] TCT (SEQ ID NO: 4)。

[0643] 2. 用下列条件运行3个不同的PCR进行n个循环:

[0644] 1个循环: 98°C 30 秒

[0645] 65°C 30 秒

[0646] 72°C 30 秒

[0647] n-2个循环: 98°C 10 秒

[0648] 65°C 30 秒

[0649] 72°C 30 秒

[0650] 1个循环: 98°C 10 秒

[0651] 65°C 30 秒

[0652] 72°C 7 分钟。

[0653] 3. 通过在1.5%琼脂糖凝胶上运行整个反应物(25μl)检查扩增的DNA的量。通常，在300–600bp范围内的成片条带应当刚好在9个扩增循环时约可见。

[0654] 第22天: 确定用于SCRiBL Hi-C文库扩增的PCR条件

[0655] b) 捕获前加条形码的SCRiBL Hi-C

[0656] 1. 为了确定用于SCRiBL Hi-C文库扩增的最佳PCR循环数，设置测试PCR具有7、9和12个扩增循环。设置3个反应，每个具有:

[0657] 2.5μl 珠上的SCRiBL Hi-C文库DNA

[0658] 5μl 缓冲液 5x (Phusion NEB F531)

[0659] 0.7μl dATP 10mM

[0660] 0.7μl dCTP 10mM

[0661] 0.7μl dGTP 10mM

[0662] 0.7μl dTTP 10mM

[0663] 0.075μl TruSeq PCR 引物 1.0 (100μM)

[0664] 0.075μl TruSeq PCR 引物 2.0 (100μM)

[0665] 0.3μl Phusion 聚合酶 (NEB F531)

[0666] 14.25μl H<sub>2</sub>O。

[0667] 2. 用下列条件运行3个不同的PCR进行n个循环:

[0668] 1个循环: 98°C 30 秒

[0669] 62°C 30 秒

[0670] 72°C 30 秒

[0671] n-2个循环: 98°C 10 秒

[0672] 62°C 30 秒

- [0673] 72℃ 30 秒
- [0674] 1个循环: 98℃ 10 秒
- [0675] 62℃ 30 秒
- [0676] 72℃ 7 分钟。

[0677] 3. 通过在1.5%琼脂糖凝胶上运行整个反应物(25μl)检查扩增的DNA的量。通常，在300–600bp范围内的成片条带应当刚好在9个扩增循环时约可见。

[0678] 第22天:确定用于SCRiBL Hi-C文库扩增的PCR条件

- c) 捕获后加条形码的SCRiBL Hi-C

[0680] 1. 为了确定用于SCRiBL Hi-C文库扩增的最佳PCR循环数，设置测试PCR具有7、9和12个扩增循环。设置3个反应，每个具有：

- [0681] 2.5μl 珠上的SCRiBL Hi-C文库DNA
- [0682] 5μl 缓冲液 5x (Phusion NEB F531)
- [0683] 0.7μl dATP 10mM
- [0684] 0.7μl dCTP 10mM
- [0685] 0.7μl dGTP 10mM
- [0686] 0.7μl dTTP 10mM
- [0687] 0.075μl SCRiBL 捕获后PCR 引物 1.0 (100μM)
- [0688] 0.075μl SCRiBL 捕获后PCR 引物 2.0 (100μM)
- [0689] 0.3μl Phusion 聚合酶 (NEB F531)
- [0690] 14.25μl H<sub>2</sub>O

SCRiBL 捕获后 PCR 引物 1.0

AATGATAACGGCGACCACCGAGATCTACACTTTCCCTACACGACGCTTCCGATCT

(SEQ ID NO: 3)

[0691] SCRiBL 捕获后 PCR 引物 2.0:

CAAGCAGAAGACGGCATACGAGATNNNNNNNTGACTGGAGTTCAGACGTGTGCTCTTC

CGATC (SEQ ID NO: 12)

[0692] (请注意“NNNNNN”是条形码)。

[0693] 2. 用下列条件运行3个不同的PCR进行n个循环：

- [0694] 1个循环: 98℃ 30 秒
- [0695] 65℃ 30 秒
- [0696] 72℃ 30 秒
- [0697] n-2个循环: 98℃ 10 秒
- [0698] 65℃ 30 秒
- [0699] 72℃ 30 秒
- [0700] 1个循环: 98℃ 10 秒
- [0701] 65℃ 30 秒
- [0702] 72℃ 7 分钟。

[0703] 3. 通过在1.5%琼脂糖凝胶上运行整个反应物(25μl)检查扩增的DNA的量。通常，在300–600bp范围内的成片条带应当刚好在9个扩增循环时约可见。

[0704] 第23天:SCRiBL Hi-C文库的最终PCR扩增

[0705] 1. 设置 x PCR反应(x 等于SCRiB HiC文库剩余体积除以2.5),每个具有25 $\mu$ l,如上述(参见第22天步骤1),用一个PCR条件(即,选择的PCR循环数)。通常,6个PCR循环应当足以获得适合用于Illumina测序的SCRiBL Hi-C文库的量。

[0706] 2. 合并来自第23天步骤1的所有个别PCR反应。置于磁力分离器上并转移上清液到新的1.5ml低结合Eppendorf管中。在初始量的1X NE缓冲液2中重悬链霉亲和素珠(参见第21天,步骤6),并保留作为备份。确定含有SCRiBL Hi-C文库的上清液的体积,并通过加入1.8x体积的SPRI珠(Beckman Coulter Ampure XP珠A63881),按照制造商的指导进行纯化。重悬在100 $\mu$ l TLE中。

[0707] 3. 通过向来自第23天步骤2的100 $\mu$ lHi-C文库加入180 $\mu$ lSPRI珠,重复SPRI珠纯化(如第23天步骤2中所述)。重悬在终体积20 $\mu$ l TLE中。

[0708] 4. 在测序前,通过Bioanalyzer (Agilent) 和定量PCR检查SCRiBL Hi-C文库的质量和数量。

[0709] 实施例3:MYC基因启动子相互作用的研究

[0710] 本文所述的方法允许对于数以千计的启动子同时捕获染色体相互作用,这随后可以单独分析。作为一个实例,本文显示MYC基因周围的基因组相互作用。结果可参照如本文所示图2进行解释。

[0711] 图2(A)显示MYC基因周围的基因组区域。图2(A)显示与CD34+造血祖细胞中的MYC启动子相互作用的区域位置,如通过本发明的方法所鉴定。星号表示位于如下所述确证的基因的下游1.8 Mb的相互作用区域的位置。

[0712] 图2(B)显示CD34+细胞核的DNA FISH的实例。在这些实验中,对于MYC基因座中的区域(如图2(A)中条所示位置)特异性的荧光标记的DNA探针杂交到固定的CD34+细胞核中的基因组。对935个基因座测量了这些探针之间的相对间隔距离。距离的定量显示为盒须图(box and whiskers graph)。柱1显示了MYC启动子和远端相互作用区域之间的间隔距离。其间隔距离比MYC启动子和在本发明方法中未显示相互作用的插入区域、或插入区域和远端相互作用区域显著更近,分别如柱2和3所示。P-值 < 0.001。

[0713] 图2(C)显示使用已知的3C方法对MYC启动子和在所述基因下游1.8Mb处区域之间的相互作用的验证。MYC启动子、远端相互作用元件和插入区域的特异性引物被用于在来自CD34 +细胞和GM12878细胞的三个独立制备的3C相互作用文库中扩增,其中在本发明的方法中没有检测到相互作用。来自有限个循环的PCR的扩增产物在琼脂糖凝胶上运行。L, DNA梯状标志物;Co, 对照扩增产物;GM, 从三个GM12878 3C文库的扩增;CD, 从三个CD34+ 3C文库的扩增;图2(C)中的图显示了对于GM12878细胞(左)和CD34 +细胞(右),相对对照扩增产物标准化的,与插入区域(黑色柱)和相互作用区域(灰色柱)相互作用的MYC启动子的凝胶的定量。

[0714] 在图2(B)和2(C)中呈现的数据因此用作图2(A)中所示发现的验证。总之,这些结果表明,位于MYC基因的下游1.8Mb的区域与MYC启动子以比在线性基因组序列上接近得多的区域显著更高的频率相互作用,如由本发明的方法所预测的。

[0001]	序列表	
[0002]	<110> Babraham Institute	
[0003]	<120> 包括选择和富集步骤的染色体构象捕获方法	
[0004]	<130> BAB-C-P1604PCT	
[0005]	<160> 18	
[0006]	<170> PatentIn 版本 3.5	
[0007]	<210> 1	
[0008]	<211> 32	
[0009]	<212> DNA	
[0010]	<213> 人工的	
[0011]	<220>	
[0012]	<223> 合成的寡核苷酸	
[0013]	<400> 1	
[0014]	gatcggaaga gcgggttcagc aggaatgccg ag	32
[0015]	<210> 2	
[0016]	<211> 33	
[0017]	<212> DNA	
[0018]	<213> 人工的	
[0019]	<220>	
[0020]	<223> 合成的寡核苷酸	
[0021]	<400> 2	
[0022]	acactctttc cctacacgac gctttccga tct	33
[0023]	<210> 3	
[0024]	<211> 58	
[0025]	<212> DNA	
[0026]	<213> 人工的	
[0027]	<220>	
[0028]	<223> 合成的寡核苷酸	
[0029]	<400> 3	
[0030]	aatgatacgg cgaccaccga gatctacact ctccctac acgacgctt tccgatct	58
[0031]	<210> 4	
[0032]	<211> 58	
[0033]	<212> DNA	
[0034]	<213> 人工的	
[0035]	<220>	
[0036]	<223> 合成的寡核苷酸	
[0037]	<400> 4	
[0038]	caaggcatacg agatcggtct cggcattcct gctgaaccgc tcttccga	58

[0039]	<210>	5		
[0040]	<211>	63		
[0041]	<212>	DNA		
[0042]	<213>	人工的		
[0043]	<220>			
[0044]	<223>	合成的寡核苷酸		
[0045]	<220>			
[0046]	<221>	尚未归类的特征		
[0047]	<222>	(34) .. (39)		
[0048]	<223>	n是a、c、g、或t		
[0049]	<400>	5		
[0050]	gatcggaaga	gcacacgtct	gaactccagt cacnnnnnna tctcgatgc cgtttctgc	60
[0051]	ttg			63
[0052]	<210>	6		
[0053]	<211>	44		
[0054]	<212>	DNA		
[0055]	<213>	人工的		
[0056]	<220>			
[0057]	<223>	合成的寡核苷酸		
[0058]	<400>	6		
[0059]	aatgatacgg	cgaccaccga	gatctacact cttccctac acga	44
[0060]	<210>	7		
[0061]	<211>	24		
[0062]	<212>	DNA		
[0063]	<213>	人工的		
[0064]	<220>			
[0065]	<223>	合成的寡核苷酸		
[0066]	<400>	7		
[0067]	caagcagaag	acggcatacg	agat	24
[0068]	<210>	8		
[0069]	<211>	33		
[0070]	<212>	DNA		
[0071]	<213>	人工的		
[0072]	<220>			
[0073]	<223>	合成的寡核苷酸		
[0074]	<400>	8		
[0075]	gatcggaaga	gcacacgtct	gaactccagt cac	33
[0076]	<210>	9		
[0077]	<211>	33		

[0078]	<212>	DNA	
[0079]	<213>	人工的	
[0080]	<220>		
[0081]	<223>	合成的寡核苷酸	
[0082]	<400>	9	
[0083]	gtgactggag ttcagacgtg tgctttccg atc		33
[0084]	<210>	10	
[0085]	<211>	46	
[0086]	<212>	DNA	
[0087]	<213>	人工的	
[0088]	<220>		
[0089]	<223>	合成的寡核苷酸	
[0090]	<400>	10	
[0091]	tcttagtcgac ggccaggtaaa ttgttaatcg actcaactata gggcga		46
[0092]	<210>	11	
[0093]	<211>	50	
[0094]	<212>	DNA	
[0095]	<213>	人工的	
[0096]	<220>		
[0097]	<223>	合成的寡核苷酸	
[0098]	<400>	11	
[0099]	agtttcgccc tatagttagt cgtattacaa ttcactggcc gtcgactaga		50
[0100]	<210>	12	
[0101]	<211>	63	
[0102]	<212>	DNA	
[0103]	<213>	人工的	
[0104]	<220>		
[0105]	<223>	合成的寡核苷酸	
[0106]	<220>		
[0107]	<221>	尚未归类的特征	
[0108]	<222>	(25) .. (30)	
[0109]	<223>	n是a、c、g、或t	
[0110]	<400>	12	
[0111]	caaggcatacg agatnnnnn gtgactggag ttcagacgtg tgctttccg		60
[0112]	atc		63
[0113]	<210>	13	
[0114]	<211>	63	
[0115]	<212>	DNA	
[0116]	<213>	人工的	

[0117]	<220>	
[0118]	<223>	合成的寡核苷酸
[0119]	<400>	13
[0120]	gatcggaaga	gcacacgtct gaactccagt cacgccaata tctcgatgc cgtttctgc 60
[0121]	ttg	63
[0122]	<210>	14
[0123]	<211>	63
[0124]	<212>	DNA
[0125]	<213>	人工的
[0126]	<220>	
[0127]	<223>	合成的寡核苷酸
[0128]	<400>	14
[0129]	caaggagaag	acggcatacg agatattggc gtgactggag ttcagacgtg tgctttccg 60
[0130]	atc	63
[0131]	<210>	15
[0132]	<211>	33
[0133]	<212>	DNA
[0134]	<213>	人工的
[0135]	<220>	
[0136]	<223>	合成的寡核苷酸
[0137]	<400>	15
[0138]	gtgactggag	ttcagacgtg tgctttccg atc 33
[0139]	<210>	16
[0140]	<211>	33
[0141]	<212>	DNA
[0142]	<213>	人工的
[0143]	<220>	
[0144]	<223>	合成的寡核苷酸
[0145]	<400>	16
[0146]	agatcggaag	agcgtcgtgt agggaaagag tgt 33
[0147]	<210>	17
[0148]	<211>	34
[0149]	<212>	DNA
[0150]	<213>	人工的
[0151]	<220>	
[0152]	<223>	合成的寡核苷酸
[0153]	<400>	17
[0154]	agatcggaag	agcacacgtc tgaactccag tcac 34
[0155]	<210>	18

[0156]	<211>	32
[0157]	<212>	DNA
[0158]	<213>	人工的
[0159]	<220>	
[0160]	<223>	合成的寡核苷酸
[0161]	<400>	18
[0162]	acacttttc cctacacgac gctttccga tc	32

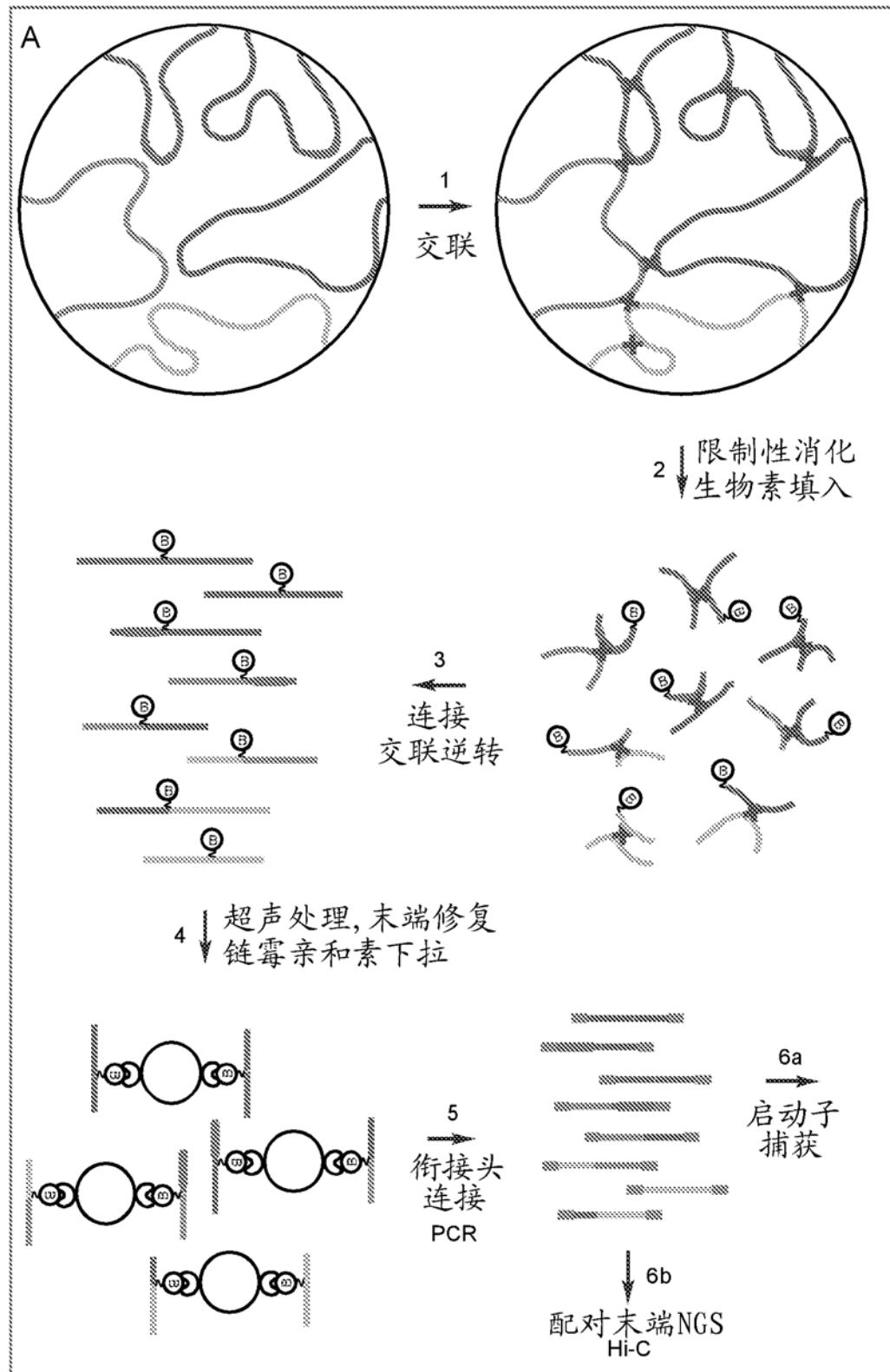


图 1

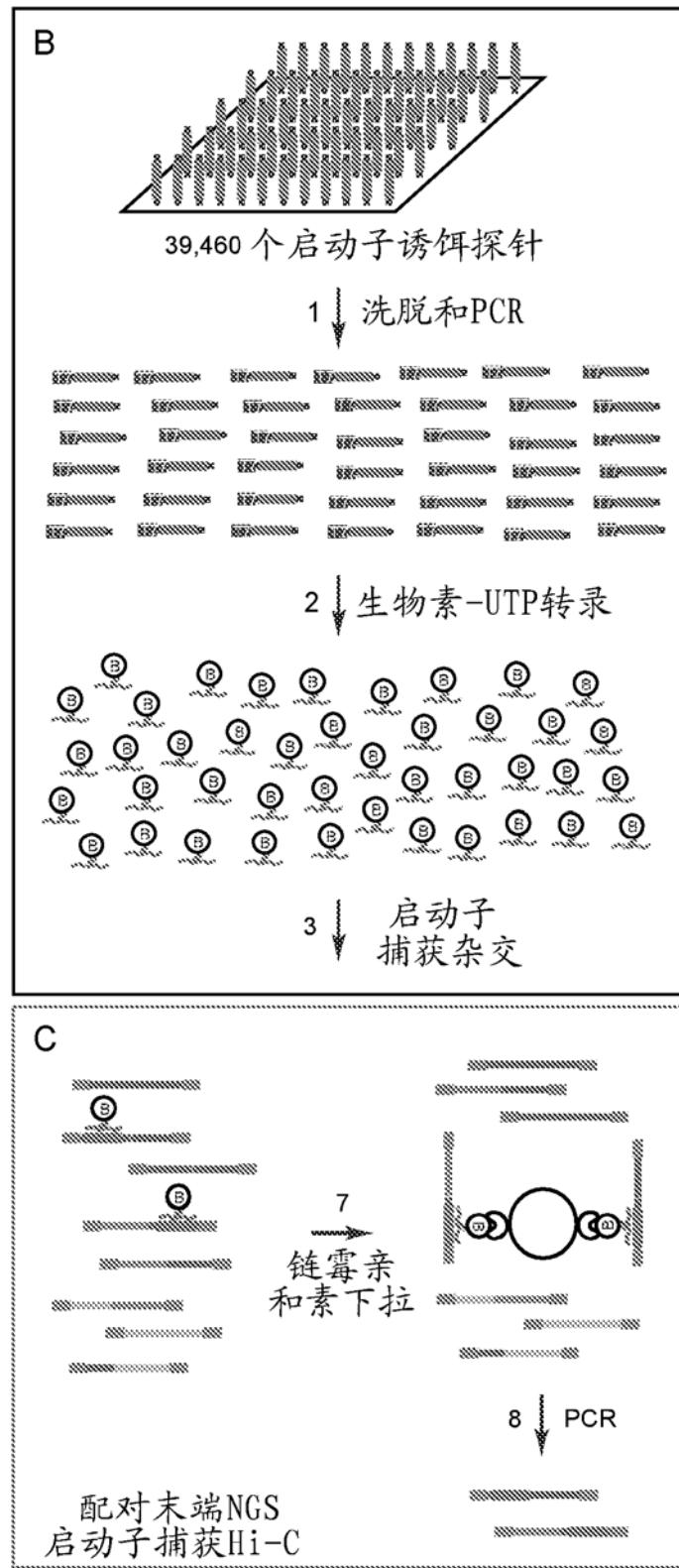


图 1(续)

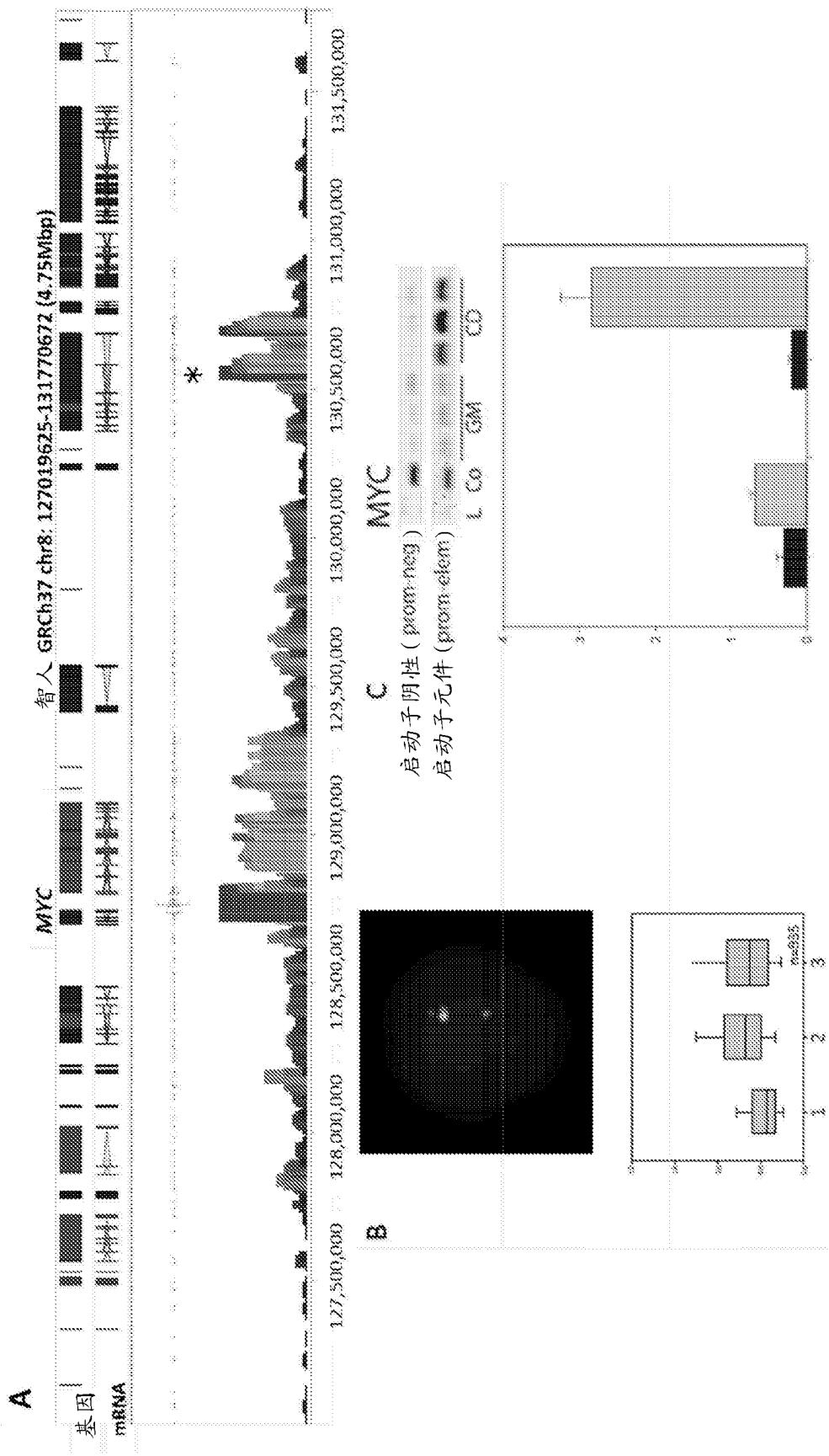


图 2