



# (12) 发明专利申请

(10) 申请公布号 CN 115485391 A

(43) 申请公布日 2022. 12. 16

(21) 申请号 202180029166.4

(22) 申请日 2021.02.17

(30) 优先权数据

62/977,512 2020.02.17 US

(85) PCT国际申请进入国家阶段日

2022.10.17

(86) PCT国际申请的申请数据

PCT/US2021/018391 2021.02.17

(87) PCT国际申请的公布数据

WO2021/167986 EN 2021.08.26

(71) 申请人 10X基因组学有限公司

地址 美国加利福尼亚州

申请人 P·萨伦

(72) 发明人 P·萨伦 P·斯托尔

(74) 专利代理机构 北京植众德本知识产权代理

有限公司 16083

专利代理师 张彦彦

(51) Int.Cl.

C12Q 1/6841(2006.01)

权利要求书7页 说明书53页

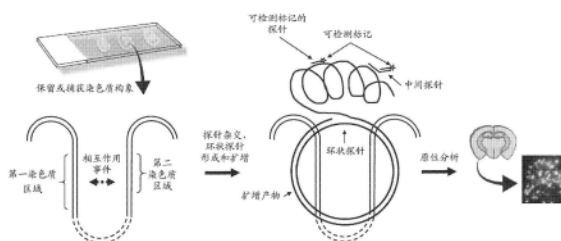
序列表2页 附图14页

## (54) 发明名称

染色质相互作用的原位分析

## (57) 摘要

在本文所述的一些实施方案中,是原位执行的用于分析细胞中的或样品诸如非均质化组织样品的细胞中的染色质相互作用事件的方法。所述方法可以包括生物样品中跨细胞群的染色质相互作用事件的空间分析。所述方法还可以包括获得生物样品、将探针与参与染色质相互作用事件的靶核酸序列杂交并产生包含靶核酸序列的全部或部分的核酸序列、扩增如此产生的所述核酸序列并原位检测所述扩增的核酸序列。



1. 一种用于分析染色质相互作用的方法,其包括:
  - (a) 使样品与第一探针和第二探针同时接触或以任意顺序依次接触,其中:

所述样品包含彼此接近的第一染色质区域和第二染色质区域,其中所述接近与细胞中所述第一和第二染色质区域之间的染色质相互作用事件相关,

所述第一探针与所述第一染色质区域中的第一核酸链杂交,并且所述第二探针与所述第二染色质区域中的第二核酸链杂交,并且

所述第一和第二探针彼此桥接或通过一种或多种桥接探针桥接;和
  - (b) 连接所述第一和/或第二探针的末端或所述一种或多种桥接探针的末端以形成环状探针,其中所述环状探针包含所述第一核酸链或其互补链的序列和/或所述第二核酸链或其互补链的序列,

其中检测所述环状探针的扩增产物,从而分析所述染色质相互作用事件。
2. 如权利要求1所述的方法,其中所述第一和第二染色质区域在相同的分子例如所述细胞中的相同染色体中。
3. 如权利要求2所述的方法,其中所述第一和第二染色质区域在核酸序列中不重叠,并且不通过所述细胞中的共同磷酸二酯键直接连接。
4. 如权利要求3所述的方法,其中所述第一和第二染色质区域之间的基因组距离是至少0.5kb、1kb、2kb、5kb、10kb、15kb、20kb、25kb、30kb、35kb、40kb、45kb、50kb、100kb、150kb、或200kb、或在两者之间的任何范围内、或大于200kb。
5. 如权利要求1所述的方法,其中所述第一和第二染色质区域在不同的分子例如所述细胞中的不同染色体中。
6. 如权利要求1至5中任一项所述的方法,其中所述第一和/或第二染色质区域包含染色体核酸,例如异染色质或常染色质中的染色体DNA。
7. 如权利要求1至6中任一项所述的方法,其中所述第一和/或第二染色质区域在核小体区域或无核小体区域中,例如,在诸如细胞核或线粒体的细胞区室中。
8. 如权利要求1至7中任一项所述的方法,其中所述第一和/或第二染色质区域在30nm纤维、间期染色体、中期染色体、端粒或着丝粒中。
9. 如权利要求1至8中任一项所述的方法,其中所述第一和/或第二染色质区域包含调控元件。
10. 如权利要求1至9中任一项所述的方法,其中所述第一和/或第二染色质区域包含启动子或其元件(例如,核心启动子元件或近端启动子元件)和/或长程调控元件(例如,增强子、沉默子、绝缘子或基因座控制区(LCR))。
11. 如权利要求10所述的方法,其中所述第一和第二探针分别与启动子序列和增强子序列杂交,或反之亦然。
12. 如权利要求1至11中任一项所述的方法,其中所述染色质相互作用事件由一种或多种染色质相关因子介导。
13. 如权利要求12所述的方法,其中所述一种或多种染色质相关因子包括多核苷酸(例如,DNA或RNA)、多肽(例如,蛋白质或肽)、碳水化合物、脂质、小分子和/或其缀合物、衍生物、代谢物或类似物。
14. 如权利要求12或13所述的方法,其中所述一种或多种染色质相关因子包括转录因

子、激活因子、抑制因子、染色质重塑因子、聚合酶、复制因子、DNA修复因子、组蛋白(例如,包含翻译后修饰的组蛋白)、组蛋白修饰酶、DNA修饰酶(例如,DNA甲基化酶)和/或其辅因子或复合物。

15. 如权利要求1至14中任一项所述的方法,其中所述染色质相互作用事件在所述接触步骤之前或期间发生。

16. 如权利要求1至15中任一项所述的方法,其中所述染色质相互作用事件将所述第一和第二染色质区域一起保持在染色质构象中。

17. 如权利要求16所述的方法,其还包括保留或捕获所述染色质构象。

18. 如权利要求1至17中任一项所述的方法,其还包括在所述接触步骤之前用交联剂处理所述样品。

19. 如权利要求18所述的方法,其中所述第一和/或第二染色质区域中的核酸通过所述交联剂彼此交联和/或与一种或多种染色质相关因子交联。

20. 如权利要求18或19所述的方法,其中所述交联剂选自由以下组成的组:甲醛、UV辐射、戊二醛、双(亚胺酯)、双(琥珀酰亚胺酯)、二异氰酸酯、二酰氯、1-乙基-3-(3-二甲氨基丙基)碳二亚胺、丝裂霉素C、氮芥、美法仑、1,3-丁二烯二环氧化物、顺式二胺二氯铂(II)、环磷酰胺、双琥珀酰亚胺戊二酸酯和丙酸二硫代双琥珀酰亚胺酯。

21. 如权利要求1至20中任一项所述的方法,其中所述第一核酸链是有义链或反义链,且/或其中所述第二核酸链是有义链或反义链。

22. 如权利要求1至21中任一项所述的方法,其中所述第一和/或第二探针是染色质接近探针,任选地其中所述第一染色质区域中的所述第一核酸链的互补链和/或所述第二染色质区域中的所述第二核酸链的互补链不与探针例如染色质接近探针杂交。

23. 如权利要求1至21中任一项所述的方法,其中所述第一染色质区域中的所述第一核酸链的互补链与第一染色质接近探针杂交,且/或其中所述第二染色质区域中的所述第二核酸链的互补链与第二染色质接近探针杂交。

24. 如权利要求22或23所述的方法,其中所述染色质接近探针(例如,所述第一和/或第二染色质接近探针)包含一种或多种天然核酸和/或一种或多种合成核酸类似物。

25. 如权利要求24所述的方法,其中所述一种或多种合成核酸类似物是异种核酸(XNA),任选地选自由以下组成的组:1,5-脱水己糖醇核酸(HNA)、环己烯核酸(CeNA)、苏糖核酸(TNA)、乙二醇核酸(GNA)、锁核酸(LNA)、肽核酸(PNA)和氟阿拉伯核酸(FANA)、及其组合。

26. 如权利要求22至25中任一项所述的方法,其中所述染色质接近探针(例如,所述第一和/或第二染色质接近探针)包含肽核酸(PNA)。

27. 如权利要求26所述的方法,其中所述第一和/或第二染色质接近探针是PNA探针,并且所述方法还包括使所述样品与所述第一和第二染色质接近探针同时接触或以任意顺序依次接触的步骤。

28. 如权利要求23至27中任一项所述的方法,其中在使所述样品与所述第一和/或第二探针接触之前使所述样品与所述第一和/或第二染色质接近探针接触。

29. 如权利要求23至28中任一项所述的方法,其中所述第一或第二核酸链和所述对应互补链之间的所述杂交的解链温度( $T_m$ )比所述第一或第二染色质接近探针和所述对应互

补链之间的所述杂交的 $T_m$ 低例如约5°C、约10°C、约15°C、约20°C、约25°C、约30°C、约35°C、约40°C、约45°C、约50°C、或大于50°C。

30. 如权利要求22至29中任一项所述的方法,其中在允许探针接近所述第一和/或第二染色质区域中的单链序列的条件下使所述样品与所述染色质接近探针(例如,所述第一和/或第二染色质接近探针)接触。

31. 如权利要求30所述的方法,其中使所述样品在至少100nM、200nM、500nM、1 $\mu$ M、2 $\mu$ M或大于2 $\mu$ M的探针浓度下与所述染色质接近探针接触。

32. 如权利要求30或31所述的方法,其中使所述样品在至少40°C、45°C、50°C或大于50°C的温度下与所述染色质接近探针接触。

33. 如权利要求30至32中任一项所述的方法,其中将所述样品与所述染色质接近探针一起孵育至少30分钟、1小时、2小时、5小时或超过5小时。

34. 如权利要求30至33中任一项所述的方法,其还包括例如使用一次或多次洗涤诸如严格洗涤去除不特异性杂交的所述染色质接近探针的分子。

35. 如权利要求30至34中任一项所述的方法,其还包括例如使用一次或多次洗涤诸如严格洗涤去除不特异性杂交的所述第一和/或第二探针和/或所述一种或多种桥接探针的分子。

36. 如权利要求1至35中任一项所述的方法,其中所述第一和/或第二探针分别包含3'突出端和5'突出端,所述3'突出端和5'突出端位于与所述第一和第二核酸链杂交的序列的侧翼。

37. 如权利要求36所述的方法,其中所述第一探针的所述3'突出端与所述第二探针的所述5'突出端至少部分互补,且/或所述第一探针的所述5'突出端与所述第二探针的所述3'突出端至少部分互补。

38. 如权利要求37所述的方法,其还包括使用所述第二探针作为模板来延伸所述第一探针的3'端(例如,使用引物延伸和/或连接反应),以及将所述第一探针的所述延伸的3'端与所述第一探针的所述任选延伸的5'端连接(例如,使用连接反应),从而环化所述第一探针以形成所述环状探针。

39. 如权利要求37或38所述的方法,其还包括使用所述第一探针作为模板来延伸所述第二探针的3'端(例如,使用引物延伸和/或连接反应),以及将所述第二探针的所述延伸的3'端与所述第二探针的所述任选延伸的5'端连接(例如,使用连接反应),从而环化所述第二探针以形成所述环状探针。

40. 如权利要求36所述的方法,其中所述第一探针的所述3'突出端和所述第二探针的所述5'突出端与第一桥接探针序列至少部分互补,并且所述第一探针的所述5'突出端和所述第二探针的所述3'突出端与第二桥接探针序列至少部分互补,任选地其中所述第一和第二桥接探针序列在相同的桥接探针中或在不同的桥接探针中。

41. 如权利要求40所述的方法,其还包括使用所述一种或多种桥接探针作为模板来延伸所述第一探针的3'端(例如,使用引物延伸和/或连接反应),以及将所述第一探针的所述延伸的3'端与所述第二探针的所述任选延伸的5'端连接(例如,使用连接反应)。

42. 如权利要求41所述的方法,其还包括使用所述一种或多种桥接探针作为模板来延伸所述第二探针的3'端(例如,使用引物延伸和/或连接反应),以及将所述第二探针的所述

延伸的3'端与所述第一探针的所述任选延伸的5'端连接(例如,使用连接反应),从而连接所述第一和第二探针以形成所述环状探针。

43.如权利要求40至42中任一项所述的方法,其中所述第一和/或第二桥接探针序列位于与所述第一染色质区域中的所述第一核酸链的互补链杂交的探针(例如,染色质接近探针)中和/或与所述第二染色质区域中的所述第二核酸链的互补链杂交的探针(例如,染色质接近探针)中。

44.如权利要求36所述的方法,其中所述第一探针的所述3'突出端和所述第二探针的所述5'突出端与第一桥接探针至少部分互补,并且所述第一探针的所述5'突出端和所述第二探针的所述3'突出端与第二桥接探针至少部分互补。

45.如权利要求44所述的方法,其还包括使用所述第一探针作为模板来延伸所述第一桥接探针的3'端(例如,使用引物延伸和/或连接反应),以及将所述第一桥接探针的所述延伸的3'端与所述第二桥接探针的所述任选延伸的5'端连接(例如,使用连接反应)。

46.如权利要求45所述的方法,其还包括使用所述第二探针作为模板来延伸所述第二桥接探针的3'端(例如,使用引物延伸和/或连接反应),以及将所述第二桥接探针的所述延伸的3'端与所述第一桥接探针的所述任选延伸的5'端连接(例如,使用连接反应),从而连接所述第一和第二桥接探针以形成所述环状探针。

47.如权利要求1至46中任一项所述的方法,其中所述环状探针在所述样品中原位形成。

48.如权利要求1至47中任一项所述的方法,其中所述第一探针、所述第二探针、所述桥接探针、所述环状探针和/或所述染色质接近探针包含条形码序列。

49.如权利要求48所述的方法,其中所述环状探针包含一种或多种条形码序列,所述条形码序列对应于例如所述第一染色质区域和/或所述第二染色质区域中的感兴趣的核酸序列。

50.如权利要求49所述的方法,其中所述扩增产物包含所述一种或多种条形码序列或其互补序列、所述第一核酸链或其互补链的序列、以及所述第二核酸链或其互补链的序列的多个拷贝。

51.如权利要求1至50中任一项所述的方法,其中所述扩增产物在所述样品中原位形成。

52.如权利要求1至51中任一项所述的方法,其中所述扩增产物使用所述环状探针的滚环扩增(RCA)形成,任选地其中所述RCA可以是线性RCA、分支RCA、树突状RCA或其任何组合,并且任选地其中所述扩增产物使用Phi29聚合酶形成,其中所述扩增在约20°C至约60°C、任选约30°C至约40°C的温度下执行。

53.如权利要求1至52中任一项所述的方法,其中所述扩增产物的所述检测包括用荧光团、同位素、质量标签或其组合来标记所述扩增产物。

54.如权利要求1至53中任一项所述的方法,其中所述扩增产物的所述检测包括将一种或多种探针与所述扩增产物直接或间接杂交,任选地其中荧光标记的探针直接与所述扩增产物杂交,或者任选地其中荧光标记的探针直接与中间探针杂交,所述中间探针与所述扩增产物杂交。

55.如权利要求1至54中任一项所述的方法,其中所述扩增产物在所述样品中原位检

测。

56. 如权利要求1至55中任一项所述的方法,其中所述扩增产物的所述检测包括例如使用荧光显微镜检查对所述样品进行成像。

57. 如权利要求1至56中任一项所述的方法,其中所述扩增产物的所述检测包括例如通过对所述扩增产物的全部或部分进行测序和/或与所述扩增产物原位杂交(例如,连续荧光原位杂交)来确定所述扩增产物的序列。

58. 如权利要求57所述的方法,其中所述测序包括边杂交边测序、边连接边测序、边合成边测序和/或边结合边测序。

59. 如权利要求1至58中任一项所述的方法,其中所述样品是组织样品,例如非均质组织样品,诸如厚度任选地在约1 $\mu\text{m}$ 和约50 $\mu\text{m}$ 之间的组织切片,例如厚度在约5 $\mu\text{m}$ 和约35 $\mu\text{m}$ 之间的组织切片。

60. 如权利要求1至59中任一项所述的方法,其中所述样品是处理过的或清除过的样品。

61. 如权利要求1至60中任一项所述的方法,其中所述样品是固定的或不固定的,例如福尔马林固定的石蜡包埋(FFPE)样品、冷冻组织样品或新鲜组织样品。

62. 如权利要求1至61中任一项所述的方法,其中所述样品在诸如载玻片的基底上。

63. 如权利要求1至62中任一项所述的方法,其中所述样品包埋在基质例如用于从所述样品中附着生物分子和/或其产物的官能化水凝胶中。

64. 如权利要求1至63中任一项所述的方法,其中分析所述样品以例如平行或依次地检测多个染色质相互作用事件。

65. 如权利要求64所述的方法,其中第一环状探针与第一染色质相互作用事件相关,并且第二环状探针与第二染色质相互作用事件相关,并且所述第一和第二环状探针各自包含对应于所述第一和/或第二探针或所述染色质相互作用事件的条形码序列或其互补序列。

66. 如权利要求1至65中任一项所述的方法,其中所述染色质相互作用事件涉及除了所述第一和第二染色质区域之外的一个或多个染色质区域,并且所述环状探针还包含所述一个或多个其他染色质区域中的核酸链的序列或其互补序列。

67. 一种用于分析染色质相互作用的方法,其包括:

(a) 使样品与第一染色质接近探针和第二染色质接近探针同时接触或以任意顺序依次接触,其中:

所述样品包含彼此接近的第一染色质区域和第二染色质区域,其中所述接近与细胞中所述第一和第二染色质区域之间的染色质相互作用事件相关,并且

所述第一染色质接近探针与所述第一染色质区域中的第一核酸链的互补链杂交,并且所述第二染色质接近探针与所述第二染色质区域中的第二核酸链的互补链杂交,

从而允许另外的探针接近所述第一和第二核酸链(例如,通过在所述第一和第二核酸链中为另外的探针提供单链结合位点);

(b) 使所述样品与第一检测探针和第二检测探针同时接触或以任意顺序依次接触,其中:

所述第一检测探针与所述第一核酸链杂交,并且所述第二检测探针与所述第二核酸链杂交;

(c) 使所述样品与桥接所述第一检测探针的3'端和所述第二检测探针的5'端的第一桥接探针和桥接所述第二检测探针的3'端和所述第一检测探针的5'端的第二桥接探针接触;

(d) 使用所述第一和第二检测探针作为模板来环化所述第一和第二桥接探针以形成环状探针,其中所述环状探针包含所述第一核酸链的序列和所述第二核酸链的序列;

(e) 在所述样品中原位生成所述环状探针的滚环扩增(RCA)产物;和

(f) 检测所述RCA产物,从而分析所述染色质相互作用事件。

68. 如权利要求67所述的方法,其中所述染色质相互作用事件将所述第一和第二染色质区域一起保持在染色质构象中,并且所述方法还包括在(a)之前用交联剂保留或捕获所述染色质构象。

69. 如权利要求67或68所述的方法,其中所述第一和第二染色质接近探针是肽核酸(PNA)探针。

70. 如权利要求67至69中任一项所述的方法,其中所述第一和第二检测探针不分别从所述第一和第二染色质区域置换所述第一和第二染色质接近探针。

71. 一种试剂盒,其包括能够与第一染色质区域中的第一核酸链杂交的第一探针,以及能够与第二染色质区域中的第二核酸链杂交的第二探针,其中:

已知所述第一染色质区域和所述第二染色质区域参与或怀疑参与细胞中的染色质相互作用事件,

在与其中保留或捕获与所述染色质相互作用事件相关的染色质构象的样品中的所述第一和第二核酸链杂交时,所述第一和第二探针彼此桥接或通过一种或多种桥接探针桥接,并且

在连接所述第一和/或第二探针的末端或所述一种或多种桥接探针的末端时,形成环状探针,并且所述环状探针包含所述第一核酸链或其互补链的序列和/或所述第二核酸链或其互补链的序列。

72. 如权利要求71所述的试剂盒,其还包括所述一种或多种桥接探针。

73. 如权利要求71或72所述用于分析多个染色质相互作用事件的试剂盒,其包括多组所述第一探针、所述第二探针和任选的所述一种或多种桥接探针,其中每组用于分析不同的染色质相互作用事件。

74. 一种试剂盒,其包括:

(i) 第一染色质接近探针和第二染色质接近探针,其中:

所述第一染色质接近探针能够与第一染色质区域中的第一核酸链的互补链杂交,

所述第二染色质接近探针能够与第二染色质区域中的第二核酸链的互补链杂交,并且

已知所述第一染色质区域和所述第二染色质区域参与或怀疑参与细胞中的染色质相互作用事件;

(ii) 第一检测探针和第二检测探针,其中:

所述第一检测探针能够与所述第一核酸链杂交,并且所述第二检测探针能够与所述第二核酸链杂交;和

(iii) 能够桥接所述第一检测探针的3'端和所述第二检测探针的5'端的第一桥接探针,和能够桥接所述第二检测探针的3'端和所述第一检测探针的5'端的第二桥接探针,

在与其中保留或捕获与所述染色质相互作用事件相关的染色质构象的样品中的所述

第一和第二核酸链杂交时,所述第一和第二检测探针通过所述第一和第二桥接探针桥接,并且

在使用所述第一和第二检测探针作为模板来连接所述第一和第二桥接探针的末端时,形成环状探针,并且所述环状探针包含所述第一核酸链的序列和所述第二核酸链的序列。

75.如权利要求74所述的试剂盒,其中所述第一和第二染色质接近探针是肽核酸(PNA)探针。

76.如权利要求74或75所述用于分析多个染色质相互作用事件的试剂盒,其包括多组所述第一和第二染色质接近探针、所述第一和第二检测探针、和/或所述第一和第二桥接探针,其中每组用于分析不同的染色质相互作用事件。

77.如权利要求71至76中任一项所述的试剂盒,其还包括酶并且任选地包括聚合酶(例如,Phi29聚合酶)和/或连接酶。



## 染色质相互作用的原位分析

[0001] 相关申请的交叉引用

[0002] 本申请要求2020年2月17日提交的题为“METHOD FOR SPATIAL ANALYSIS OF CHROMATIN INTERACTION EVENTS PERFORMED IN SITU”的美国临时专利申请号62/977,512的优先权,所述申请出于所有目的以引用方式整体并入本文。

[0003] 关于ASCII文本文件的序列表的提交

[0004] 关于ASCII文本文件的以下提交的内容以引用方式整体并入本文:计算机可读形式(CRF)的序列表(文件名:202412006440SEQLIST.TXT,记录日期:2021年2月12日,大小:2KB)。

### 技术领域

[0005] 本公开在一些方面涉及用于分析染色质相互作用事件的方法,所述事件涉及细胞或系统内的核酸序列。

### 背景技术

[0006] 脱氧核糖核酸(DNA)通常被视为线性分子,很少关注三维组织。但是染色体不是刚性的,并且虽然两个基因座之间的线性距离可能很大,但折叠时,空间距离可能很小。例如,虽然以三维染色质结构组织的染色体DNA的区域可能被许多兆碱基分开,但它们也可以在三维空间中直接相邻。

[0007] 调控元件(例如,基因增强子、沉默子和绝缘子元件)是短片段,其可以含有激活或抑制基因的转录因子的一个或多个结合位点。调控元件经常远离它们的靶基因,并且尽管它们可以通过特定因子的结合来识别,但经常不清楚它们与哪些基因相互作用。然而,在基因组的空间组织中,它们接近于它们的靶基因。

[0008] 了解核酸如何相互作用,或许更重要的是,这种相互作用或缺乏这种相互作用如何调控细胞过程,是一个相对较新的探索领域。例如,了解染色体折叠和其中的模式可以深入了解染色质结构、基因活性和细胞功能状态之间的复杂关系。在癌基因和其他疾病相关基因的情况下,长程遗传调节因子的鉴定对于鉴定导致疾病状态的基因组变异和导致疾病状态的过程将非常有用。

[0009] 检查调控元件(诸如启动子和增强子)如何在整个基因组中发挥作用以调控功能(诸如转录)的染色质相互作用研究是已知的。但是,迄今为止,尚未跨组织切片使用空间分辨率原位检查染色质相互作用事件。这种策略将揭示基因组中跨靶调控元件相互作用的、组织样品中跨细胞群的协同活性。因此,需要改进的方法来原位分析靶调控区域在组织样品中跨细胞群的三维染色质结构中的相互作用,并且本公开解决了这个需求和其他需求。

### 发明内容

[0010] 在一些方面,本文提供了用于分析染色质相互作用的方法,其包括:(a)使样品与第一探针和第二探针同时接触,或者以任意顺序依次接触,其中:样品包含彼此接近的第一

染色质区域和第二染色质区域,其中接近与细胞中第一和第二染色质区域之间的染色质相互作用事件相关,第一探针与第一染色质区域中的第一核酸链杂交并且第二探针与第二染色质区域中的第二核酸链杂交,并且第一和第二探针彼此桥接或通过一种或多种桥接探针桥接;和(b)连接第一和/或第二探针的末端或一种或多种桥接探针的末端以形成环状探针,其中环状探针包含第一核酸链或其互补链的序列和/或第二核酸链或其互补链的序列,其中检测环状探针的扩增产物,从而分析染色质相互作用事件。

[0011] 在一些实施方案中,第一和第二染色质区域可以在相同的分子例如细胞中的相同染色体中。在一些实施方案中,第一和第二染色质区域可以在核酸序列中不重叠并且可以不通过细胞中的共同磷酸二酯键直接连接。在一些实施方案中,第一和第二染色质区域之间的基因组距离可以是至少0.5kb、1kb、2kb、5kb、10kb、15kb、20kb、25kb、30kb、35kb、40kb、45kb、50kb、100kb、150kb、或200kb、或在两者之间的任何范围内、或大于200kb。

[0012] 在一些实施方案中,第一和第二染色质区域可以在不同的分子例如细胞中的不同染色体中。

[0013] 在任何前述实施方案中,第一和/或第二染色质区域可以包含染色体核酸,例如异染色质或常染色质中的染色体DNA。

[0014] 在任何前述实施方案中,第一和/或第二染色质区域可以在核小体区域或无核小体区域中,例如在诸如细胞核或线粒体的细胞区室中。

[0015] 在任何前述实施方案中,第一和/或第二染色质区域可以在30nm纤维、间期染色体、中期染色体、端粒或着丝粒中。

[0016] 在任何前述实施方案中,第一和/或第二染色质区域可以包含调控元件。

[0017] 在任何前述实施方案中,第一和/或第二染色质区域可以包含启动子或其元件(例如,核心启动子元件或近端启动子元件)和/或长程调控元件(例如,增强子、沉默子、绝缘子或基因座控制区(LCR))。在任何前述实施方案中,第一和第二探针可以分别与启动子序列和增强子序列杂交,或反之亦然。

[0018] 在任何前述实施方案中,染色质相互作用事件可以由一种或多种染色质相关因子介导。在任何前述实施方案中,一种或多种染色质相关因子可以包括多核苷酸(例如,DNA或RNA)、多肽(例如,蛋白质或肽)、碳水化合物、脂质、小分子和/或其缀合物、衍生物、代谢物或类似物。在任何前述实施方案中,一种或多种染色质相关因子可以包括转录因子、激活因子、抑制因子、染色质重塑因子、聚合酶、复制因子、DNA修复因子、组蛋白(例如,包含翻译后修饰的组蛋白)、组蛋白修饰酶、DNA修饰酶(例如,DNA甲基化酶)和/或其辅因子或复合物。

[0019] 在任何前述实施方案中,染色质相互作用事件可以在接触步骤之前或期间发生。在任何前述实施方案中,染色质相互作用事件可以将第一和第二染色质区域一起保持在染色质构象中。在任何前述实施方案中,方法还可以包括保留或捕获染色质构象。

[0020] 在任何前述实施方案中,方法还可以包括在接触步骤之前用交联剂处理样品。在任何前述实施方案中,第一和/或第二染色质区域中的核酸可以通过交联剂彼此交联和/或与一种或多种染色质相关因子交联。在任何前述实施方案中,交联剂可以包括甲醛、UV辐射、戊二醛、双(亚胺酯)、双(琥珀酰亚胺酯)、二异氰酸酯、二酰氯、1-乙基-3-(3-二甲基氨基丙基)碳二亚胺、丝裂霉素C(mitomycin C)、氮芥、美法仑(melphalan)、1,3-丁二烯双环氧化物、顺式二胺二氯铂(II)、环磷酰胺、双琥珀酰亚胺戊二酸酯或丙酸二硫代双琥珀酰亚

胺酯、及其任何组合。

[0021] 在任何前述实施方案中,第一核酸链可以是有义链或反义链,且/或其中第二核酸链可以是有义链或反义链。例如,第一和第二核酸链可以都是有义链,或者第一和第二核酸链可以都是反义链。在其他实例中,第一核酸链是有义链并且第二核酸链是反义链,或反之亦然。

[0022] 在任何前述实施方案中,第一和/或第二探针可以是染色质接近探针(chromatin accessing probe),例如开放探针(诸如PNA探针),其打开染色质中的DNA双链体以提供另一个探针的单链结合位点。在任何前述实施方案中,第一染色质区域中的第一核酸链的互补链和/或第二染色质区域中的第二核酸链的互补链可以与一种或多种另外的探针杂交,诸如具有“一手柄”(例如,不与染色质DNA杂交的3'和/或5'突出端)的检测探针和/或另外的染色质接近探针。在任何前述实施方案中,第一染色质区域中的第一核酸链的互补链和/或第二染色质区域中的第二核酸链的互补链可以不与另外的探针(例如,另外的染色质接近探针)杂交。

[0023] 在任何前述实施方案中,第一染色质区域中的第一核酸链的互补链可以与第一染色质接近探针杂交,且/或其中第二染色质区域中的第二核酸链的互补链可以与第二染色质接近探针杂交。

[0024] 在任何前述实施方案中,染色质接近探针(例如,分别靶向第一和第二染色质区域的第一和/或第二染色质接近探针)可以包含一种或多种天然核酸残基和/或一种或多种合成核酸类似物的一种或多种残基。在任何前述实施方案中,一种或多种合成核酸类似物可以包括异种核酸(XNA)并且可以任选地包括1,5-脱水己糖醇核酸(HNA)、环己烯核酸(CeNA)、苏糖核酸(TNA)、乙二醇核酸(GNA)、锁核酸(LNA)、肽核酸(PNA)或氟阿拉伯核酸(fluoro arabino nucleic acid,FANA)、或其任何组合。

[0025] 在任何前述实施方案中,染色质接近探针(例如,第一和/或第二染色质接近探针)可以包含一种或多种肽核酸(PNA)残基。在任何前述实施方案中,第一和/或第二染色质接近探针可以是PNA探针,并且方法还可以包括使样品与第一和第二染色质接近探针同时接触或以任意顺序依次接触的步骤。在任何前述实施方案中,样品可以在使样品与第一和/或第二探针接触之前与第一和/或第二染色质接近探针接触。

[0026] 在任何前述实施方案中,第一或第二核酸链和对应互补链之间的杂交的解链温度( $T_m$ )可以比第一或第二染色质接近探针和对应互补链之间的杂交的 $T_m$ 低例如约5°C、约10°C、约15°C、约20°C、约25°C、约30°C、约35°C、约40°C、约45°C、约50°C、或大于50°C。在一些实施方案中, $T_m$ 的这种差异允许染色质接近探针分离染色质中的双链体的链。染色质接近探针可以与染色质中的核酸链稳定杂交,从而使未缠绕的双链体的另一条链成为单链并且可供一种或多种另外的探针(诸如具有5'和/或3'突出端的检测探针)接近。

[0027] 在任何前述实施方案中,样品可以在允许探针接近第一和/或第二染色质区域中的单链序列的条件下与染色质接近探针(例如,第一和/或第二染色质接近探针)接触。探针接近可以包括通过染色质接近探针和/或另外的探针(诸如检测探针和/或另外的染色质接近探针)的接近。在此类条件下,染色质接近探针可以不包括XNA(例如,PNA)残基并且可以由天然核酸残基组成。在任何前述实施方案中,可以在至少100nM、200nM、500nM、1 $\mu$ M、2 $\mu$ M或大于2 $\mu$ M的染色质接近探针(例如,PNA探针或天然核酸残基的探针)浓度下使样品与染色质

接近探针接触,这可以通过染色质接近探针促进染色质中双链体的打开以为一种或多种另外的探针提供单链结合位点。在任何前述实施方案中,可以在至少40°C、45°C、50°C、或大于50°C的温度下使样品与染色质接近探针(例如,PNA探针或天然核酸残基的探针)接触,这可以促进染色质中双链体的打开以为染色质接近探针和/或另一种探针提供单链结合位点。在任何前述实施方案中,样品可以与染色质接近探针一起孵育至少30分钟、1小时、2小时、5小时或超过5小时,并且孵育可以促进染色质中双链体的打开以为染色质接近探针和/或另一种探针提供单链结合位点。

[0028] 染色质接近探针可以是检测探针(例如,在与染色质中的单链区域结合后具有5'和/或3'突出端的探针)。在一些实施方案中,染色质接近探针本身不是检测探针,并且在与染色质中的单链区域结合后可以不包含突出端,或者仅包含一个突出端(3'或5'),或者包含3'和5'突出端,其中3'和/或5'突出端可以促进环状探针的形成,但不形成环状探针的一部分。

[0029] 在任何前述实施方案中,方法还可以包括例如使用一次或多次洗涤诸如严格洗涤去除不特异性杂交的染色质接近探针的分子。

[0030] 在任何前述实施方案中,方法还可以包括例如使用一次或多次洗涤诸如严格洗涤去除不特异性杂交的第一和/或第二探针和/或一种或多种桥接探针的分子。

[0031] 在任何前述实施方案中,第一和/或第二探针可以分别包含3'突出端和5'突出端,所述3'突出端和5'突出端位于与第一和第二核酸链杂交的序列的侧翼。

[0032] 在任何前述实施方案中,第一探针的3'突出端可以与第二探针的5'突出端至少部分互补,且/或第一探针的5'突出端可以与第二探针的3'突出端至少部分互补。

[0033] 在任何前述实施方案中,方法还可以包括使用第二探针作为模板来延伸第一探针的3'端(例如,使用引物延伸和/或连接反应),以及将第一探针的延伸的3'端与第一探针的任选延伸的5'端连接(例如,使用连接反应),从而环化第一探针以形成环状探针。

[0034] 在任何前述实施方案中,方法还可以包括使用第一探针作为模板来延伸第二探针的3'端(例如,使用引物延伸和/或连接反应),以及将第二探针的延伸的3'端与第二探针的任选延伸的5'端连接(例如,使用连接反应),从而将第二探针环化以形成环状探针。

[0035] 在任何前述实施方案中,第一探针的3'突出端和第二探针的5'突出端可以与第一桥接探针序列至少部分互补,并且第一探针的5'突出端和第二探针的3'突出端可以与第二桥接探针序列至少部分互补,任选地其中第一和第二桥接探针序列可以在相同的桥接探针中或在不同的桥接探针中。

[0036] 在任何前述实施方案中,方法还可以包括使用桥接探针作为模板来延伸第一探针的3'端(例如,使用引物延伸和/或连接反应),以及将第一探针的延伸的3'端与第二探针的任选延伸的5'端连接(例如,使用连接反应)。

[0037] 在任何前述实施方案中,方法还可以包括使用桥接探针作为模板来延伸第二探针的3'端(例如,使用引物延伸和/或连接反应),以及将第二探针的延伸的3'端与第一探针的任选延伸的5'端连接(例如,使用连接反应),从而连接第一和第二探针以形成环状探针。

[0038] 在任何前述实施方案中,第一和/或第二桥接探针序列可以位于与第一染色质区域中的第一核酸链的互补链杂交的探针(例如,染色质接近探针)中和/或与第二染色质区域中的第二核酸链的互补链杂交的探针(例如,染色质接近探针)中。

[0039] 在任何前述实施方案中,第一探针的3'突出端和第二探针的5'突出端可以与第一桥接探针至少部分互补,并且第一探针的5'突出端和第二探针的3'突出端可以与第二桥接探针至少部分互补。

[0040] 在任何前述实施方案中,方法还可以包括使用第一探针作为模板来延伸第一桥接探针的3'端(例如,使用引物延伸和/或连接反应),以及将第一桥接探针的延伸的3'端与第二桥接探针的任选延伸的5'端连接(例如,使用连接反应)。

[0041] 在任何前述实施方案中,方法还可以包括使用第二探针作为模板来延伸第二桥接探针的3'端(例如,使用引物延伸和/或连接反应),以及将第二桥接探针的延伸的3'端与第一桥接探针的任选延伸的5'端连接(例如,使用连接反应),从而连接第一和第二桥接探针以形成环状探针。

[0042] 在任何前述实施方案中,环状探针可以在样品中原位形成。

[0043] 在任何前述实施方案中,第一探针、第二探针、桥接探针、环状探针和/或染色质接近探针可以包含条形码序列。在任何前述实施方案中,环状探针可以包含一种或多种条形码序列,所述条形码序列对应于例如第一染色质区域和/或在第二染色质区域中的感兴趣的核酸序列。在任何前述实施方案中,扩增产物可以包含一种或多种条形码序列或其互补序列、第一核酸链或其互补链的序列、以及第二核酸链或其互补链的序列的多个拷贝。

[0044] 在任何前述实施方案中,扩增产物可以在样品中原位形成。

[0045] 在任何前述实施方案中,扩增产物可以使用环状探针的滚环扩增(RCA)形成,任选地其中RCA可以是线性RCA、分支RCA、树突状RCA或其任何组合,并且任选地其中扩增产物可以使用Phi29聚合酶形成,其中扩增可以在约20°C至约60°C、任选约30°C至约40°C的温度下执行。

[0046] 在任何前述实施方案中,扩增产物的检测可以包括用荧光团、同位素、质量标签或其组合来标记扩增产物。

[0047] 在任何前述实施方案中,扩增产物的检测可以包括将一种或多种探针与扩增产物直接或间接杂交,任选地其中荧光标记的探针可以直接与扩增产物杂交,或者任选地其中荧光标记的探针可以直接与中间探针杂交,所述中间探针可以与扩增产物杂交。

[0048] 在任何前述实施方案中,扩增产物可以在样品中原位检测。

[0049] 在任何前述实施方案中,扩增产物的检测可以包括例如使用荧光显微镜检查对样品进行成像。

[0050] 在任何前述实施方案中,扩增产物的检测可以包括例如通过对扩增产物的全部或部分进行测序和/或与扩增产物原位杂交(例如,连续荧光原位杂交)确定扩增产物的序列。

[0051] 在任何前述实施方案中,测序可以包括边杂交边测序、边连接边测序、边合成边测序和/或边结合边测序。

[0052] 在任何前述实施方案中,样品可以是组织样品,例如非均质组织样品,诸如厚度任选地在约1 $\mu$ m和约50 $\mu$ m之间(例如约5 $\mu$ m和约35 $\mu$ m之间)的组织切片。

[0053] 在任何前述实施方案中,样品可以是处理过的或清除过的样品。

[0054] 在任何前述实施方案中,样品可以是固定的或不固定的,例如福尔马林固定的石蜡包埋(FFPE)样品、冷冻组织样品或新鲜组织样品。

[0055] 在任何前述实施方案中,样品可以在基底(诸如载玻片)上。

[0056] 在任何前述实施方案中,样品可以包埋在基质例如用于从样品中附着生物分子和/或其产物的官能化水凝胶中。

[0057] 在任何前述实施方案中,可以分析样品以例如平行或依次地检测多个染色质相互作用事件。

[0058] 在任何前述实施方案中,第一环状探针可以与第一染色质相互作用事件相关,并且第二环状探针可以与第二染色质相互作用事件相关,并且第一和第二环状探针各自可以包含可以对应于第一和/或第二探针或染色质相互作用事件的条形码序列或其互补序列。

[0059] 在任何前述实施方案中,染色质相互作用事件可以涉及除了第一和第二染色质区域之外的一个或多个染色质区域,并且环状探针还可以包含一个或多个其他染色质区域中的核酸链的序列或其互补序列。

[0060] 在一些方面,本文提供了一种用于分析染色质相互作用的方法,其包括:(a)使样品与第一染色质接近探针和第二染色质接近探针同时接触或以任意顺序依次接触,其中:样品包含彼此接近的第一染色质区域和第二染色质区域,其中接近与细胞中的第一和第二染色质区域之间的染色质相互作用事件相关,并且第一染色质接近探针与第一染色质区域中的第一核酸链的互补链杂交,并且第二染色质接近探针与第二染色质区域中的第二核酸链的互补链杂交,从而允许另外的探针接近第一和第二核酸链(例如,通过在第一和第二核酸链中为另外的探针提供单链结合位点);(b)使样品与第一探针和第二探针同时接触或以任意顺序依次接触,其中:第一探针与第一核酸链杂交,并且第二探针与第二核酸链杂交;(c)使样品与桥接第一探针的3'端和第二探针的5'端的第一桥接探针和桥接第二探针的3'端和第一探针的5'端的第二桥接探针接触;(d)使用第一和第二探针作为模板来环化第一和第二桥接探针以形成环状探针,其中环状探针包含第一核酸链的序列和第二核酸链的序列;(e)在样品中原位生成环状探针的滚环扩增(RCA)产物;和(f)检测RCA产物,从而分析染色质相互作用事件。

[0061] 在任何前述实施方案中,染色质相互作用事件可以将第一和第二染色质区域一起保持在染色质构象中,并且方法还可以包括在(a)之前用交联剂保留或捕获染色质构象。在任何前述实施方案中,第一和第二染色质接近探针可以是肽核酸(PNA)探针。在任何前述实施方案中,第一和第二探针可以不分别从第一和第二染色质区域置换第一和第二染色质接近探针。在一些方面,使样品在与第一和第二探针接触之前与第一和第二染色质接近探针接触。染色质接近探针可以与互补链稳定结合,使得染色质中的第一和第二核酸链成为单链,从而为第一和第二探针提供结合位点。在一些方面,第一和第二探针与第一和第二核酸链结合,而不是与第一和第二染色质接近探针结合。在一些实施方案中,染色质接近探针结合互补链的亲合力大于染色质中的核酸链结合其互补链的亲合力,由此染色质接近探针打开染色质中的双链体。在一些实施方案中,染色质接近探针结合互补链的亲合力大于检测探针(例如,第一和第二探针)对染色质接近探针的亲合力,使得染色质接近探针以及第一和第二探针与染色质中的链结合,而不是形成染色质接近探针/检测探针双链体。

[0062] 在任何前述实施方案中,方法可以用于原位染色质相互作用事件的空间分析。在一些实施方案中,染色质相互作用事件跨越生物样品中的细胞群。

[0063] 在一些方面,本文提供了一种试剂盒,其包含:能够与第一染色质区域中的第一核酸链杂交的第一探针,和能够与第二染色质区域中的第二核酸链杂交的第二探针,其中:已

知第一染色质区域和第二染色质区域参与或怀疑参与细胞中的染色质相互作用事件；在与其中与染色质相互作用事件相关的染色质构象被保留或捕获的样品中的第一和第二核酸链杂交时，第一和第二探针彼此桥接或通过一种或多种桥接探针桥接；并且在连接第一和/或第二探针的末端或一种或多种桥接探针的末端时，形成环状探针并且所述环状探针包含第一核酸链或其互补链的序列和/或第二核酸链或其互补链的序列。在一些实施方案中，试剂盒还包含一种或多种桥接探针。在一些方面，本文提供了一种用于分析多个染色质相互作用事件的试剂盒，其包含多组第一探针、第二探针和任选的一种或多种桥接探针，其中每组用于分析不同的染色质相互作用事件。

[0064] 在一些方面，本文提供了一种试剂盒，其包括：(i) 第一染色质接近探针和第二染色质接近探针，其中：第一染色质接近探针能够与第一染色质区域中的第一核酸链的互补链杂交；第二染色质接近探针能够与第二染色质区域中的第二核酸链的互补链杂交；并且已知第一染色质区域和第二染色质区域参与或怀疑参与细胞中的染色质相互作用事件；(ii) 第一检测探针和第二检测探针，其中：第一检测探针能够与第一核酸链杂交，并且第二检测探针能够与第二核酸链杂交；和(iii) 能够桥接第一检测探针的3'端和第二检测探针的5'端的第一桥接探针，以及能够桥接第二检测探针的3'端和第一检测探针的5'端的第二桥接探针；在与其中与染色质相互作用事件相关的染色质构象被保留或捕获的样品中的第一和第二核酸链杂交时，第一和第二检测探针通过第一和第二桥接探针桥接；并且在使用第一和第二检测探针为模板连接第一和第二桥接探针的末端时，形成环状探针，并且所述环状探针包含第一核酸链的序列和第二核酸链的序列。在任何前述实施方案中，第一和第二染色质接近探针可以是肽核酸(PNA)探针。在一些方面，本文提供了一种用于分析多个染色质相互作用事件的试剂盒，其包括多组第一和第二染色质接近探针、第一和第二检测探针、和/或第一和第二桥接探针，其中每组用于分析不同的染色质相互作用事件。

[0065] 在任何前述实施方案中，试剂盒还可以包括酶，所述酶任选地包括聚合酶(例如，Phi29聚合酶)和/或连接酶。

[0066] 在一些实施方案中，方法包括：a) 获得生物样品；b) 将包含至少两种单一寡核苷酸序列的第一组探针(例如，一种或多种染色质接近探针)与至少两种一级靶核酸序列的有义链或反义链(例如，染色质中的第一或第二核酸链的互补链)杂交，其中所述至少两种一级靶核酸序列(例如，第一和第二染色质区域)参与染色质相互作用事件；c) 将包含至少两种单一寡核苷酸序列的第二组探针(例如，用于第一和第二染色质区域的第一和第二探针，诸如具有突出端的检测探针)与所述至少两种一级靶核酸序列的有义链或反义链(例如，第一或第二核酸链)杂交；d) 将第三组探针(例如，一种或多种桥接探针或环化探针)与所述第二组探针杂交，其中所述第二探针由于与参与染色质相互作用事件的至少两种一级靶核酸序列杂交而在空间上彼此接近，并且其中所述第三探针(例如，一种或多种桥接探针或环化探针)桥接所述第二探针的末端；e) 使用聚合酶延伸第三组探针(例如，一种或多种桥接探针或环化探针)以形成环化的核酸，所述环化的核酸包含参与染色质相互作用事件的至少两种一级靶核酸序列的至少一部分；f) 扩增所述环状核酸以生成包含可原位检测的信号产生标签的扩增子群；和g) 通过所述信号产生标签原位检测所述带标签的扩增子群。

[0067] 在一些实施方案中，方法可以是多重的并且包括以下的步骤：淬灭从所述信号产生标签生成的信号；和重复步骤b) - g)，其中所述第一组和第二组探针各自包含至少两种单

一寡核苷酸序列,所述单一寡核苷酸序列与参与染色质相互作用事件的至少两种二级靶核酸序列的有义链或反义链杂交。在一些实施方案中,方法包括猝灭来自所述信号产生标签的信号。在一些实施方案中,方法包括连续多轮次重复步骤b)-g),其中每个连续轮次包含具有与用于其他轮次的第一和第二组探针不同的寡核苷酸序列的第一和第二组探针,并且其中第一和第二组探针与不同于其他轮次中的靶核酸序列的靶核酸序列杂交。

[0068] 在一个实施方案中,使用交联剂将所述生物样品内的染色质与染色质相关因子化学交联。在一些实施方案中,将染色质DNA与染色质相关因子交联。在一个实施方案中,交联剂可以选自包括以下的组:甲醛、UV辐射、戊二醛、双(亚胺酯)、双(琥珀酰亚胺酯)、二异氰酸酯、二酰氯、和1-乙基-3-(3-二甲基氨基丙基)碳二亚胺。在替代实施方案中,染色质不与染色质相关因子交联。

[0069] 在一个实施方案中,一级靶核酸序列包含基因组DNA。在一些实施方案中,一级靶核酸序列包含靶启动子DNA序列和靶增强子DNA序列。

[0070] 在一些实施方案中,所述第一组探针之一(例如,用于第一染色质区域的染色质接近探针)和所述第二组探针之一(例如,用于第一染色质区域的第一探针,诸如检测探针)与靶启动子DNA序列杂交,并且其中所述第一组探针之一(例如,用于第二染色质区域的染色质接近探针)和所述第二组探针之一(例如,用于第二染色质区域的第二探针,诸如检测探针)与靶增强子DNA序列杂交。例如,第一染色质区域包含靶启动子序列或其部分,而第二染色质区域包含靶增强子序列或其部分,或反之亦然。在一个实施方案中,第一组探针(例如,染色质接近探针)包括肽核酸探针或PNA。在另一个实施方案中,第二组探针包含含有3'和5'突出端的检测探针。在另一个实施方案中,第三组探针(例如,一种或多种桥接探针或环化探针)包含与所述第二组探针的所述3'和5'突出端互补的序列并且与所述第二组探针杂交以形成所述桥。

[0071] 在一些实施方案中,所述扩增步骤通过滚环扩增进行并且对所述扩增子进行化学标记。在一个实施方案中,将扩增子用荧光标记的探针进行化学标记。在另一个实施方案中,检测步骤包括荧光原位杂交。

[0072] 在一些实施方案中,方法涉及制备包含生物样品的载体基底。在一个实施方案中,载体基底是其上具有所述生物样品的载玻片。在一个实施方案中,将生物样品通过固定剂固定。在替代实施方案中,不对生物样品进行固定。

[0073] 在一些实施方案中,染色质相关因子可以包括蛋白质、DNA、RNA、小分子和代谢物。在一个实施方案中,染色质相关因子是转录因子。

[0074] 在一些实施方案中,方法还包括通过原位测序对所述一个或多个扩增子群进行测序的步骤。

[0075] 在另一个实施方案中,用于原位空间分析染色质相互作用事件的方法包括以下步骤:在合适的载体基底上制备组织样品;通过交联剂将所述组织样品内的染色质与染色质相关因子交联;将包含至少两种单一寡核苷酸序列的第一寡核苷酸组(例如,与染色质区域中的核酸链杂交的一种或多种探针,例如检测探针)与至少两种靶核酸序列的第一链杂交,其中所述至少两种靶核酸序列参与染色质相互作用事件;将第二寡核苷酸组(例如,一种或多种桥接探针或环化探针)与所述第一寡核苷酸组杂交,其中所述第一寡核苷酸由于与参与染色质相互作用事件的至少两种靶核酸序列杂交而彼此接近,并且其中所述第二寡核苷酸



酸桥接所述第一寡核苷酸的末端；延伸第二寡核苷酸以形成环状寡核苷酸，所述环状寡核苷酸包含参与染色质相互作用事件的至少两种靶核酸序列的至少一部分；扩增所述环状寡核苷酸以生成可原位检测的扩增子群；以及原位检测所述扩增子。

[0076] 在又一个实施方案中，用于原位空间分析染色质相互作用事件的方法包括以下步骤：a) 从至少一个受试者获得生物样品；b) 制备包含所述生物样品的载体基底；c) 使用交联剂将所述生物样品内的染色质与染色质相关因子交联；d) 将多个探针与参与染色质相互作用事件的多个一级靶核酸序列原位杂交，其中所述多个探针至少包含与环化探针杂交的检测探针；e) 使用聚合酶延伸环化探针以形成环化的核酸，所述环化的核酸包含参与染色质相互作用事件的一级靶核酸序列的至少一部分；f) 扩增所述环状核酸以生成一个或多个扩增子群，每个所述扩增子群包含可原位检测的信号产生标签；和g) 通过所述信号产生标签原位检测所述带标签的扩增子群。

[0077] 在一些实施方案中，信号产生标签对于一个或多个扩增子群是唯一的并且能够区分一个群与另一个群。

## 附图说明

[0078] 通过参考以下附图公开了本发明的代表性实施方案。

[0079] 图1示出了用于在生物组织样品中原位分析第一和第二染色质区域之间的染色质相互作用事件的示例性方法。虚线指示第一和第二染色质区域可以在相同的分子中或在不同分子中（例如，不同的染色体）。

[0080] 图2A示出了一种或多种探针与染色质区域中的核酸链杂交的示例性实施方案，其中第一和第二探针是染色质接近探针。第一染色质区域中的第一核酸链的互补链和/或第二染色质区域中的第二核酸链的互补链可以但不需要结合染色质接近探针。使互补链成为单链以为另外的探针（诸如具有突出端、可以连接以形成环状探针的检测探针）提供结合位点。

[0081] 图2B示出了一种或多种探针与染色质区域中的核酸链杂交的示例性实施方案，其中第一探针是与第一染色质区域杂交的染色质接近探针/检测探针，并且第二探针是与第二染色质区域中的第二核酸链杂交的检测探针。提供单独的染色质接近探针以打开第二染色质区域中的双链体，使得第二探针可以结合。

[0082] 图3A示出了一种或多种探针与染色质区域中的核酸链杂交的示例性实施方案，其中提供染色质接近探针以打开第一和第二染色质区域中的双链体，使得第一和第二探针（具有突出端的检测探针）可以结合。

[0083] 图3B示出了一种或多种探针与染色质区域中的核酸链杂交的示例性实施方案。多个染色质接近探针可以用于打开染色质中的双链体。在第一染色质区域中，与互补链上的相邻序列杂交的两种染色质接近探针使第一核酸链成为单链。第一探针的靶结合序列可能与两种染色质接近探针不互补，以便减少探针之间的双链体形成。在染色质接近探针/互补染色质链杂交时，染色质接近探针可以与互补链交联（例如，在第二染色质区域中）以减少探针双链体的形成并有利于检测探针/染色质链双链体的形成。

[0084] 图4示出了其中第一和第二探针彼此桥接，以及连接第一探针的末端或第二探针的末端以形成环状探针的示例性实施方案。可以使用一种探针作为模板来延伸另一种探

针,例如,使用5'至3'引物延伸(例如,使用DNA聚合酶)之后连接(例如,使用连接酶),和/或使用探针与模板的杂交和杂交探针的连接。

[0085] 图5A示出了其中第一和第二探针通过桥接探针桥接的示例性实施方案。延伸桥接探针的末端,然后使用第一或第二探针作为模板来连接,从而形成环状探针,所述环状探针包含桥接探针的序列、与第一探针互补的序列(例如,染色质中的第一核酸链中的序列)、和与第二探针互补的序列(例如,染色质中的第二核酸链中的序列)。由于第一和第二探针不形成环状探针的一部分,因此可以阻断第一和/或第二探针的5'和/或3'端以进行连接。第一和/或第二探针的5'和/或3'端可以任选地包含一种或多种不与桥接探针杂交的另外的序列。

[0086] 图5B示出了其中第一和第二探针通过桥接探针桥接的示例性实施方案。桥接第一和第二探针末端的桥接探针序列可以通过接头或间隔物连接,所述接头或间隔物可以包含核酸或非核酸部分。由于桥接探针不形成环状探针的一部分,因此可以阻断桥接探针的5'和/或3'端以进行连接。桥接探针的5'和/或3'端可以任选地包含一种或多种不与第一或第二探针杂交的另外的序列。第一和第二探针的末端可以任选地延伸,然后使用桥接探针作为模板来连接,从而形成环状探针,所述环状探针包含第一探针的序列(其与染色质中的第一核酸链互补)和第二探针的序列(其与染色质中的第二核酸链互补)。

[0087] 图6A示出了示例性实施方案,其中第一和/或第二桥接探针序列在与第一染色质区域中的第一核酸链的互补链杂交的探针(例如,染色质接近探针)中,或者在与第二染色质区域中的第二核酸链的互补链杂交的探针中。第一和第二探针的末端可以任选地延伸,然后使用染色质接近探针的突出端序列作为模板来连接,从而形成环状探针,所述环状探针包含第一探针的序列(其与染色质中的第一核酸链互补)和第二探针的序列(其与染色质中的第二核酸链互补)。染色质接近探针的突出端序列可以是5'或3'端阻断的或包含不与第一或第二探针杂交的另外的序列。在一些情况下,在环状探针形成后,染色质接近探针的突出端序列可以用作环状探针的滚环扩增(RCA)引物。可以提供单独的RCA引物。

[0088] 图6B示出了示例性实施方案,其中一个桥接探针序列在染色质接近探针中,并且另一个桥接探针序列在与染色质接近探针分开提供的桥接探针中。

[0089] 图7示出了组织样品中跨细胞群的染色质相互作用事件的原位空间分析的实例。染色质中由转录因子(TF)介导的启动子区域和相互作用区域(例如,增强子)之间的染色质相互作用作为非限制性实例在图7-10中示出。

[0090] 图8示出了示例性实施方案,其中PNA开放探针和检测探针与参与染色质相互作用事件的靶序列杂交。

[0091] 图9示出了方法,其包括杂交环化探针以桥接检测探针末端、以及使用检测探针作为模板来延伸环化探针(任一检测探针可以在5'至3'或3'至5'方向延伸)的示例性步骤。延伸的环化探针的末端可以连接(例如,连接)以形成环状探针。

[0092] 图10示出了包括滚环扩增和荧光信号检测的示例性步骤的方法。示例了可检测标记的RCA扩增子。然而,扩增子本身可以不被可检测标记,并且可以被可检测标记的探针(未示出)检测。或者,扩增子可以被中间探针(其本身可以不被可检测标记;未示出)识别,并且中间探针可以被可检测标记的探针检测。

## 具体实施方式

[0093] 本申请中提及的所有出版物,包括专利文献、科学文章和数据库,出于所有目的都以引用方式整体并入,所达到的程度如同每个单独的出版物都以引用方式单独并入一样。如果本文中列出的定义与以引用的方式并入本文的专利、申请、公布的申请以及其它出版物中提出的定义相反或以其它方式不一致,则本文提出的定义优先于以引用的方式并入本文的定义。

[0094] 本文所用的小节标题仅出于组织性目的并且不解释为限制所描述的主题。

[0095] I. 样品和分析物

[0096] A. 样品和样品处理

[0097] 本文公开的样品可以是或来源于其中期望分析染色质相互作用的任何生物样品。本文公开的方法和组合物可以用于分析生物样品,所述生物样品可以使用多种技术(包括但不限于活检、外科手术和激光捕获显微镜检查(LCM))中的任一种从受试者获得,并且通常包括来自受试者的细胞和/或其他生物材料。生物样品也可以从非哺乳动物生物体(例如,植物、昆虫、蛛形纲动物、线虫、真菌或两栖动物)获得。生物样品也可以从真核生物获得,诸如组织样品、患者来源的类器官(PDO)或患者来源的异种移植物(PDX)。来自生物体的生物样品可以包含一种或多种其他生物体或来自其中的组分。例如,除了哺乳动物细胞和非细胞组织组分之外,哺乳动物组织切片可以包含朊病毒、类病毒、病毒、细菌、真菌或来自其他生物体的组分。可以从中获得生物样品的受试者可以是健康或无症状的个体、患有或怀疑患有疾病或疾病倾向的个体(例如,患有疾病(诸如癌症)的患者)、和/或需要治疗或怀疑需要治疗的个体。

[0098] 生物样品可以包括任何数量的大分子,例如,细胞大分子和细胞器(例如,线粒体和细胞核)。生物样品可以作为组织样品获得,诸如组织切片、活检、核心活检、针抽吸物或细针抽吸物。样品可以是流体样品,例如血液样品、尿液样品或唾液样品,并且可以在将细胞或细胞组分放置到基底上之后分析细胞和其中的细胞组分(诸如染色质)。样品可以是皮肤样品、结肠样品、面颊拭子、组织学样品、组织病理学样品、血浆或血清样品、肿瘤样品、活细胞、培养细胞、临床样品,诸如例如全血或血液来源的产物、血细胞、或培养的组织或细胞,包括细胞悬浮液。在一些实施方案中,生物样品可以包括沉积在表面上的细胞,并且细胞可以从身体样品(例如血液、血浆、血清、尿液、唾液、粘膜排泄物、痰、粪便和泪液)中分离。

[0099] 生物样品可以来源于本文提到的受试者或生物体的同质培养物或群体,或者另选地来自例如群落或生态系统中的若干种不同生物体的集合。

[0100] 生物样品可以包括一种或多种患病细胞。患病细胞可以具有改变的代谢性质、基因表达、蛋白质表达和/或形态学特征。疾病的实例包括炎症病症、代谢病症、神经系统病症和癌症。癌细胞可以来源于实体瘤、血液恶性肿瘤、细胞系,或作为循环肿瘤细胞获得。生物样品还可以包括胎儿细胞和免疫细胞。

[0101] 生物样品可以包括3D基质中的分析物(例如,蛋白质、RNA和/或DNA)。在一些实施方案中,来源于分析物(例如,蛋白质、RNA和/或DNA)或与分析物相关的扩增子(例如,滚环扩增产物)可以嵌入3D基质中。在一些实施方案中,3D基质可以包含化学和/或酶促连接(例如,通过交联)的天然分子和/或合成分子的网络。在一些实施方案中,3D基质可以包含合成

聚合物。在一些实施方案中,3D基质包含水凝胶。

[0102] 在一些实施方案中,本文的基底可以是不溶于水性液体并且允许将生物样品、分析物、特征和/或试剂(例如,探针)放置在支持物上的任何支持物。在一些实施方案中,生物样品可以附着至基底。取决于样品的性质和分析方法中的后续步骤,生物样品的附着可以是不可逆的或可逆的。在某些实施方案中,可以通过向基底施加合适的聚合物涂层并使样品与聚合物涂层接触而将样品可逆地附着至基底。然后可以例如使用至少部分溶解聚合物涂层的有机溶剂将样品从基底分离。水凝胶是适用于此目的的聚合物的实例。

[0103] 在一些实施方案中,基底可以用一种或多种物质涂覆或官能化以促进样品与基底的附着。可以用于涂覆或官能化基底的合适物质包括但不限于凝集素、聚赖氨酸、抗体和多糖。

[0104] 可以执行多个步骤来制备或处理用于测定和/或测定期间的生物样品。除非另有说明,否则下文所述的制备或处理步骤一般可以任何方式和任何顺序组合,以适当地制备或处理特定样品以用于和/或分析。

[0105] (i) 组织切片

[0106] 生物样品可以从受试者中收获(例如,经由手术活检、全受试者切片)或在生长基底或培养皿上作为细胞群在体外生长,并且制备为组织薄片或组织切片用于分析。生长的样品可以足够薄,无需进一步处理步骤即可用于分析。或者,可以使用机械切割设备(诸如振动刀片切片机)将生长的样品和经由活检或切片获得的样品制备为薄组织切片。作为另一个替代方案,在一些实施方案中,可以通过向合适的基底材料施加生物样品的触印来制备薄组织切片。

[0107] 组织切片的厚度可以是细胞的最大横截面尺寸的一部分(例如,小于0.9、0.8、0.7、0.6、0.5、0.4、0.3、0.2或0.1)。然而,也可以使用厚度大于最大横截面细胞尺寸的组织切片。例如,可以使用低温恒温器切片,其可以是例如10 $\mu$ m-20 $\mu$ m厚。

[0108] 更一般地来讲,组织切片的厚度通常取决于用于制备切片的方法和组织的物理特性,并且因此可以制备并使用具有多种不同厚度的切片。例如,组织切片的厚度可以是至少0.1 $\mu$ m、0.2 $\mu$ m、0.3 $\mu$ m、0.4 $\mu$ m、0.5 $\mu$ m、0.7 $\mu$ m、1.0 $\mu$ m、1.5 $\mu$ m、2 $\mu$ m、3 $\mu$ m、4 $\mu$ m、5 $\mu$ m、6 $\mu$ m、7 $\mu$ m、8 $\mu$ m、9 $\mu$ m、10 $\mu$ m、12 $\mu$ m、13 $\mu$ m、14 $\mu$ m、15 $\mu$ m、20 $\mu$ m、30 $\mu$ m、40 $\mu$ m、或50 $\mu$ m。如果期望或方便,也可以使用更厚的切片,例如至少70 $\mu$ m、80 $\mu$ m、90 $\mu$ m或100 $\mu$ m或更厚。通常,组织切片的厚度在1 $\mu$ m-100 $\mu$ m、1 $\mu$ m-50 $\mu$ m、1 $\mu$ m-30 $\mu$ m、1 $\mu$ m-25 $\mu$ m、1 $\mu$ m-20 $\mu$ m、1 $\mu$ m-15 $\mu$ m、1 $\mu$ m-10 $\mu$ m、2 $\mu$ m-8 $\mu$ m、3 $\mu$ m-7 $\mu$ m、或4 $\mu$ m-6 $\mu$ m之间,但如上文提到的,也可以分析厚度大于或小于这些范围的切片。

[0109] 也可以从单个生物样品中获得多个切片。例如,可以通过使用切片刀片对活检样品执行连续切片来从手术活检样品中获得多个组织切片。以这种方式可以保留连续切片之间的空间信息,并且可以对切片进行连续分析以获得关于生物样品的三维信息。

[0110] (ii) 冷冻

[0111] 在一些实施方案中,生物样品(例如,如上所述的组织切片)可以通过在适合维持或保持组织结构的完整性(例如,物理特性)的温度下深度冷冻来制备。可以使用任何数量的合适的方法将冷冻的组织样品切片(例如,切成薄片)到基底表面上。例如,可以使用设定在适合维持组织样品的结构完整性和样品中的核酸的化学性质的温度下的冷冻切片机(例如,低温恒温器)来制备组织样品。这种温度可以例如低于-15 $^{\circ}$ C、低于-20 $^{\circ}$ C或低于-25 $^{\circ}$ C。

[0112] (iii) 固定和后固定

[0113] 在一些实施方案中,生物样品可以使用福尔马林固定和石蜡包埋 (FFPE) 来制备,福尔马林固定和石蜡包埋是已建立的方法。在一些实施方案中,可以使用福尔马林固定和石蜡包埋来制备细胞悬浮液和其他非组织样品。在固定样品并包埋在石蜡或树脂块中之后,可以如上所述对样品进行切片。在分析之前,可以通过在适当的溶剂 (例如,二甲苯) 中孵育组织切片然后冲洗 (例如,99.5%乙醇2分钟,96%乙醇2分钟,和70%乙醇2分钟) 来从组织切片中去除石蜡包埋材料 (例如,去石蜡化)。

[0114] 作为上述福尔马林固定的替代方案,可以在分析之前将生物样品固定在多种其他固定剂中的任一种中以保持样品的生物结构。例如,样品可以经由浸入乙醇、甲醇、丙酮、多聚甲醛 (PFA) -Triton及其组合中来固定。

[0115] 在一些实施方案中,丙酮固定用于新鲜冷冻样品,其可以包括但不限于皮质组织、小鼠嗅球、人脑肿瘤、人死后脑和乳腺癌样品。当执行丙酮固定时,可以不执行预透化步骤 (如下所述)。或者,丙酮固定可以与透化步骤结合执行。

[0116] 在一些实施方案中,本文提供的方法包括一个或多个固定后 (post-fixing) (也称为固定后 (postfixation)) 步骤。在一些实施方案中,在使样品与本文公开的多核苷酸 (例如,一种或多种探针,诸如环状或锁式探针) 接触之后,执行一个或多个固定后步骤。在一些实施方案中,当在样品中形成包含探针和靶标的杂交复合物之后,执行一个或多个固定后步骤。在一些实施方案中,在本文公开的连接反应 (诸如环化锁式探针的连接) 之前,执行一个或多个固定后步骤。

[0117] 在一些实施方案中,在使样品与用于非核酸分析物 (诸如蛋白质分析物) 的结合剂或标记剂 (例如,抗体或其抗原结合片段) 接触之后,执行一个或多个固定后步骤。标记剂可以包含核酸分子 (例如,报告寡核苷酸),其包含对应于标记剂的序列并因此对应于 (例如,唯一地鉴定) 分析物。在一些实施方案中,标记剂可以包含含有一种或多种条形码序列的报告寡核苷酸。

[0118] 可以使用本文公开的任何合适的固定试剂 (例如,DEPC-PBS中的3% (重量/体积) 多聚甲醛) 来执行固定后步骤。

[0119] (iv) 包埋

[0120] 作为上述石蜡包埋的替代方案,可以在切片和其他处理步骤之前将生物样品包埋在多种其他包埋材料中的任一种中以为样品提供结构基底。在一些情况下,可以例如在分析从样品获得的组织切片之前去除包埋材料。合适的包埋材料包括但不限于蜡、树脂 (例如,甲基丙烯酸酯树脂)、环氧树脂和琼脂。

[0121] 在一些实施方案中,生物样品可以包埋在基质 (例如,水凝胶基质) 中。在一些方面,包埋材料可以向样品施加一次或多次。以这种方式包埋样品通常涉及使生物样品与水凝胶接触,使得生物样品被水凝胶包围。例如,可以通过使样品与合适的聚合物材料接触并激活聚合物材料以形成水凝胶来包埋样品。在一些实施方案中,形成水凝胶,使得水凝胶内化在生物样品内。

[0122] 在一些实施方案中,生物样品经由形成水凝胶的聚合物材料的交联而固定在水凝胶中。交联可以化学和/或光化学执行,或者另选地通过本领域已知的任何其他水凝胶形成方法执行。

[0123] 水凝胶基质的组成和对生物样品的应用通常取决于生物样品的性质和制备(例如,切片、非切片、固定类型)。作为一个实例,在生物样品是组织切片的情况下,水凝胶基质可以包括单体溶液和过硫酸铵(APS)引发剂/四甲基乙二胺(TEMED)促进剂溶液。作为另一个实例,在生物样品由细胞(例如,培养的细胞或从组织样品解离的细胞)组成的情况下,可以将细胞与单体溶液和APS/TEMED溶液一起孵育。对于细胞,水凝胶基质凝胶在隔室中形成,包括但不限于用于培养、维持或运输细胞的装置。例如,水凝胶基质可以使用添加到隔室中达到约0.1 $\mu$ m至约2mm范围内的深度的单体溶液加APS/TEMED形成。

[0124] 生物样品的水凝胶包埋的另外的方法和方面在例如Chen等人,Science 347(6221):543-548,2015中有所描述,所述文献的全部内容以引用方式并入本文。

[0125] (v) 染色

[0126] 为了促进可视化,可以使用多种染色剂和染色技术对生物样品进行染色。在一些实施方案中,例如,可以使用任何数量的染色剂对样品进行染色,所述染色剂包括但不限于吖啶橙、俾斯麦棕(Bismarck brown)、胭脂红、考马斯蓝、甲酚紫、DAPI、曙红、溴化乙锭、酸性品红、苏木精、赫斯特染色剂(Hoechst stains)、碘、甲基绿、亚甲基蓝、中性红、尼罗蓝、尼罗红、四氧化钼、碘化丙啶、罗丹明或番红。

[0127] 可以使用苏木精和曙红(H&E)染色技术、使用巴氏染色技术(Papanicolaou staining technique)、马森三色染色技术(Masson's trichrome staining technique)、银染色技术、苏丹染色技术(Sudan staining technique)和/或使用高碘酸-希夫(Periodic Acid Schiff,PAS)染色技术对样品进行染色。PAS染色通常在福尔马林或丙酮固定后执行。在一些实施方案中,可以使用罗氏染色剂(Romanowsky stain),包括瑞氏染色剂(Wright's stain)、哲纳尔氏染色剂(Jenner's stain)、坎-格伦瓦德染色剂(Can-Grunwald stain)、利什曼染色剂(Leishman stain)和吉姆萨染色剂(Giemsa stain),对样品进行染色。

[0128] 在一些实施方案中,可以对生物样品进行脱色。使生物样品脱色或褪色的方法是本领域已知的,并且通常取决于向样品施加的染色剂的性质。例如,在一些实施方案中,经由抗体偶联向样品施加一种或多种免疫荧光染色剂。可以使用诸如经由用还原剂处理和去污剂洗涤来切割二硫键、离液盐处理、用抗原修复溶液处理和用酸性甘氨酸缓冲液处理的技术来去除此类染色剂。用于多重染色和脱色的方法在例如Bolognesi等人,J.Histochem.Cytochem.2017;65(8):431-444、Lin等人,Nat Commun.2015;6:8390、Pirici等人,J.Histochem.Cytochem.2009;57:567-75和Glass等人,J.Histochem.Cytochem.2009;57:899-905中有所描述,所述文献各自的全部内容以引用方式并入本文。

[0129] (vi) 等距膨胀

[0130] 在一些实施方案中,嵌入在基质(例如,水凝胶)中的生物样品可以等距膨胀。可以使用的等距膨胀方法包括水合作用,这是膨胀显微镜检查中的准备步骤,如Chen等人,Science 347(6221):543-548,2015中所述。

[0131] 等距膨胀可以通过将生物样品的一种或多种组分锚定到凝胶上,然后形成凝胶、蛋白水解和溶胀来执行。在一些实施方案中,样品中的分析物、分析物的产物和/或与样品中的分析物相关的探针可以锚定到基质(例如,水凝胶)上。生物样品的等距膨胀可以在将

生物样品固定在基底上之前发生,或者在将生物样品固定在基底上之后发生。在一些实施方案中,等距膨胀的生物样品可以在使基底与本文公开的探针接触之前从基底去除。

[0132] 一般来讲,用于执行生物样品的等距膨胀的步骤可以取决于样品的特性(例如,组织切片的厚度、固定、交联)和/或感兴趣的分析物的特性(例如,将RNA、DNA和蛋白质锚定到凝胶上的不同条件)。

[0133] 在一些实施方案中,将生物样品中的蛋白质锚定到可溶胀的凝胶(诸如聚电解质凝胶)上。抗体可以在锚定至可溶胀凝胶之前、之后或与锚定至可溶胀凝胶结合来针对蛋白质。生物样品中的DNA和/或RNA也可以经由合适的接头锚定到可溶胀凝胶上。此类接头的实例包括但不限于6-((丙烯酰基)氨基)己酸(Acryloyl-X SE)(可从ThermoFisher,Waltham,MA获得)、Label-IT胺(可从MirusBio,Madison,WI获得)和Label X(在例如Chen等人,Nat.Methods 13:679-684,2016中有所描述,所述文献的全部内容以引用方式并入本文)。

[0134] 样品的等距膨胀可以提高样品的后续分析的空间分辨率。可以通过将等距膨胀的样品与未等距膨胀的样品进行比较来确定空间谱分析中增加的分辨率。

[0135] 在一些实施方案中,生物样品等距膨胀至其非膨胀尺寸的至少2倍、2.1倍、2.2倍、2.3倍、2.4倍、2.5倍、2.6倍、2.7倍、2.8倍、2.9倍、3倍、3.1倍、3.2倍、3.3倍、3.4倍、3.5倍、3.6倍、3.7倍、3.8倍、3.9倍、4倍、4.1倍、4.2倍、4.3倍、4.4倍、4.5倍、4.6倍、4.7倍、4.8倍、或4.9倍的尺寸。在一些实施方案中,样品等距膨胀至其未膨胀尺寸的至少2倍且小于20倍。

[0136] (vii) 交联和去交联

[0137] 在一些实施方案中,生物样品在本文公开的测定步骤之前、期间或之后可逆或不可逆地交联。交联剂包括促进一个分子与另一个分子的连接的化学剂或甚至光。交联剂可以是蛋白质-核酸交联剂、核酸-核酸交联剂和/或蛋白质-蛋白质交联剂。此类剂的实例是本领域已知的。在一些实施方案中,交联剂是可逆的交联剂。在一些实施方案中,交联剂是不可逆的交联剂。

[0138] 在一些实例中,将染色质交联以保持与染色质DNA复合的染色质相关因子以用于分析本文公开的染色质相互作用事件。在一些实施方案中,使待分析的样品与蛋白质-核酸交联剂、核酸-核酸交联剂、蛋白质-蛋白质交联剂或其任意组合接触。在一些实施方案中,与染色质DNA相互作用的蛋白质和/或核酸与染色质DNA交联,使得交联的蛋白质和/或核酸的分析可以与它们结合的染色质DNA的分析组合执行(例如,平行或依次地)。在一些实施方案中,可以辨别染色质相关因子和染色质DNA之间的一级、二级和三级相互作用。在一些实例中,交联剂是可逆的交联剂,使得交联的分子可以容易地分离。在一些实施方案中,交联剂是不可逆的交联剂,使得交联的分子不可以容易地分离。在一些实例中,交联剂是光,诸如UV光。在一些实例中,交联剂是光激活的。这些交联剂包括甲醛、戊二酸二琥珀酰亚胺酯、UV-254、补骨脂素及其衍生物,诸如氨基甲基三氧沙林、戊二醛、乙二醇双[琥珀酰亚胺基琥珀酸酯]和本领域技术人员已知的其他化合物,包括可从万维网上获得的Thermo Scientific Pierce Cross-linking Technical Handbook,Thermo Scientific(2009)中描述的那些。

[0139] 在一些方面,分析物、多核苷酸和/或分析物的扩增产物(例如,扩增子)或与其结合的探针可以锚定到聚合物基质上。例如,聚合物基质可以是水凝胶。在一些实施方案中,可以修饰多核苷酸探针和/或其扩增产物(例如,扩增子)中的一种或多种以含有官能团,所

述官能团可用作锚定位点以将多核苷酸探针和/或扩增产物与聚合物基质连接。

[0140] 在一些实施方案中,生物样品经由形成水凝胶的聚合物材料的交联而固定在水凝胶中。交联可以化学和/或光化学执行,或者另选地通过本领域已知的任何其他水凝胶形成方法执行。水凝胶可以包括包含网络的大分子聚合物凝胶。在网络中,一些聚合物链可以任选地交联,尽管交联并不总是发生。

[0141] 在一些实施方案中,水凝胶可以包括水凝胶亚基,诸如但不限于丙烯酰胺、双丙烯酰胺、聚丙烯酰胺及其衍生物、聚(乙二醇)及其衍生物(例如,PEG-丙烯酸酯(PEG-DA)、PEG-RGD)、明胶-甲基丙烯酸(GelMA)、甲基丙烯酸化透明质酸(MeHA)、聚脂肪族聚氨酯、聚醚聚氨酯、聚酯聚氨酯、聚乙烯共聚物、聚酰胺、聚乙烯醇、聚丙二醇、聚环丁烷氧化物、聚乙烯吡咯烷酮、聚丙烯酰胺、聚(丙烯酸羟乙酯)和聚(甲基丙烯酸羟乙酯)、胶原蛋白、透明质酸、壳聚糖、右旋糖酐、琼脂糖、明胶、藻酸盐、蛋白质聚合物、甲基纤维素等,以及它们的组合。

[0142] 在一些实施方案中,水凝胶包括杂合材料,例如,水凝胶材料包括合成和天然聚合物两者的元素。合适的水凝胶的实例在例如例如美国专利号6,391,937、9,512,422和9,889,422、以及美国专利申请公开号2017/0253918、2018/0052081和2010/0055733中有所描述,所述专利各自的全部内容以引用方式并入本文。

[0143] 在一些实施方案中,水凝胶可以形成基底。在一些实施方案中,基底包括水凝胶和一种或多种第二材料。在一些实施方案中,将水凝胶置于一种或多种第二材料的顶部。例如,水凝胶可以预先形成,并且然后与一种或多种第二材料一起放置在顶部、底部或任何其他配置中。在一些实施方案中,水凝胶形成在基底形成期间接触一种或多种第二材料之后发生。水凝胶的形成也可以在位于基底上的结构(例如,孔、脊、突起和/或标记)内发生。

[0144] 在一些实施方案中,基底上的水凝胶形成在向样品提供探针之前、同时或之后发生。例如,可以在已经含有探针的基底上执行水凝胶形成。

[0145] 在一些实施方案中,水凝胶形成在生物样品内发生。在一些实施方案中,将生物样品(例如,组织切片)包埋在水凝胶中。在一些实施方案中,将水凝胶亚基输注到生物样品中,并且通过外部或内部刺激引发水凝胶的聚合。

[0146] 在其中在生物样品内形成水凝胶的实施方案中,可以使用官能化化学。在一些实施方案中,官能化化学包括水凝胶-组织化学(HTC)。适用于HTC的任何水凝胶组织骨架(例如,合成的或天然的)都可以于锚定生物大分子和调节官能化。使用HTC骨架变体的方法的非限制性实例包括CLARITY、PACT、ExM、SWITCH和ePACT。在一些实施方案中,生物样品内的水凝胶形成是永久性的。例如,生物大分子可以永久地粘附至水凝胶,从而允许多轮询问。在一些实施方案中,生物样品内的水凝胶形成是可逆的。

[0147] 在一些实施方案中,在聚合之前、同时和/或之后将另外的试剂添加到水凝胶亚基中。例如,另外的试剂可以包括但不限于用于扩增核酸并将条形码连接至扩增片段的寡核苷酸(例如,探针)、片段DNA的核酸内切酶、DNA的片段化缓冲液(fragmentation buffer)、DNA聚合酶、dNTP。可以使用其他酶,包括但不限于RNA聚合酶、转座酶、连接酶、蛋白酶K和DNA酶。另外的试剂还可以包括逆转录酶,包括具有末端转移酶活性的酶、引物和转换寡核苷酸。在一些实施方案中,在聚合之前、同时和/或之后将光标记添加到水凝胶亚基中。

[0148] 在一些实施方案中,在聚合之前、同时和/或之后将HTC试剂添加到水凝胶中。在一些实施方案中,在聚合之前、同时和/或之后将细胞标记剂添加到水凝胶中。在一些实施方



案中,在聚合之前、同时和/或之后将细胞穿透剂添加到水凝胶中。

[0149] 可以使用任何合适的方法清除包埋在生物样品中的水凝胶。例如,可以使用电泳组织清除方法从水凝胶包埋样品中去除生物大分子。在一些实施方案中,在清除水凝胶之前或之后将水凝胶包埋样品储存在介质(例如,封固介质、甲基纤维素或其他半固体介质)中。

[0150] 在一些实施方案中,本文公开的方法包括使可逆交联的生物样品去交联。去交联不必是完全的。在一些实施方案中,仅使可逆交联的生物样品中的一部分交联分子去交联并允许迁移。

[0151] (viii) 组织透化和处理

[0152] 在一些实施方案中,可以将生物样品透化以促进分析物转移出样品,和/或促进物质(诸如探针)转移到样品中。如果样品未充分透化,则从样品中捕获的分析物的量可能太低而无法实现充分分析。相反,如果组织样品的渗透性太强,则可能会丢失组织样品内的分析物的相对空间关系。因此,期望在透化组织样品的程度足够获得良好的信号强度,同时仍维持样品中的分析物分布的空间分辨率之间的平衡。

[0153] 一般来讲,可以通过将样品暴露于一种或多种透化剂来使生物样品透化。用于此目的的合适的剂包括但不限于有机溶剂(例如,丙酮、乙醇和甲醇)、交联剂(例如,多聚甲醛)、去污剂(例如,皂苷、Triton X-100<sup>TM</sup>或Tween-20<sup>TM</sup>)和酶(例如,胰蛋白酶、蛋白酶)。在一些实施方案中,生物样品可以与细胞透化剂一起孵育以促进样品的透化。用于样品透化的另外的方法在例如Jamur等人,Method Mol.Biol.588:63-66,2010中有所描述,所述文献的全部内容以引用方式并入本文。用于样品透化的任何合适的方法通常可以与本文所述的样品结合使用。

[0154] 在一些实施方案中,可以通过将一种或多种裂解试剂添加到样品中来使生物样品透化。合适的裂解剂的实例包括但不限于生物活性试剂,诸如用于裂解不同细胞类型(例如,革兰氏阳性或阴性细菌、植物、酵母、哺乳动物)的裂解酶,诸如溶菌酶、无色肽酶、溶葡萄球菌素(lysostaphin)、labiase、细胞裂解酶(kitalase)、溶壁酶和多种其他可商购获得的裂解酶。

[0155] 其他裂解剂可以另外地或另选地添加到生物样品中以促进透化。例如,可以使用基于表面活性剂的裂解溶液来裂解样品细胞。裂解溶液可以包括离子表面活性剂,诸如例如十二烷基肌氨酸钠和十二烷基硫酸钠(SDS)。更一般地来讲,化学裂解剂可以包括但不限于有机溶剂、螯合剂、去污剂、表面活性剂和离液剂。

[0156] 在一些实施方案中,可以通过非化学透化方法使生物样品透化。非化学透化方法是本领域已知的。例如,可以使用的非化学透化方法包括但不限于物理裂解技术(诸如电穿孔)、机械透化方法(例如,使用均质器和研磨球以机械方式破坏样品组织结构的珠击打)、声学透化(例如,声波降解法)和热裂解技术(诸如加热以诱导样品的热透化)。

[0157] 在分析样品之前,可以将另外的试剂添加到生物样品中以执行各种功能。在一些实施方案中,可以将DNA酶和RNA酶灭活剂或抑制剂(诸如蛋白酶K)和/或螯合剂(诸如EDTA)添加到样品中。例如,本文公开的方法可以包括用于增加核酸对结合的可及性的步骤,例如,打开细胞中的DNA以供探针杂交的变性步骤。例如,可以使用蛋白酶K处理来释放与蛋白质结合的DNA。

[0158] B. 分析物

[0159] 本文公开的方法和组合物可以用于检测和分析参与染色质相互作用的多种不同分析物。可以直接或间接检测靶标或分析物。

[0160] 分析物可以来源于特定类型的细胞和/或特定亚细胞区域,例如来自细胞核或来自线粒体。可以使用特异性靶向某些细胞区室和细胞器的透化剂来选择性地从细胞中释放分析物以用于分析,和/或允许一种或多种试剂(例如,用于分析物检测的探针)接近细胞或细胞区室或细胞器中的分析物。

[0161] 分析物可以包括可能参与染色质相互作用的任何生物分子、大分子或化学化合物,包括蛋白质或肽、脂质或核酸分子,或包括有机或无机分子的小分子。分析物可以是为其开发特异性结合配偶体(例如,亲和结合配偶体)的任何物质或实体。这种特异性结合配偶体可以是核酸探针(用于核酸分析物)并且可以直接导致生成RCA模板(例如,锁式或其他可环化探针)。或者,例如在使用或生成可以是RCA模板的环状核酸分子的测定中,特异性结合配偶体可以与核酸偶联,所述核酸可以使用RCA策略检测。

[0162] 特别感兴趣的分析物可以包括核酸分子,诸如DNA(例如,基因组DNA、线粒体DNA、质体DNA、病毒DNA等)和RNA(例如,mRNA、微RNA、rRNA、snRNA、病毒RNA等),以及合成和/或修饰的核酸分子(例如,包括核酸结构域,所述核酸结构域包含合成或修饰的核苷酸(诸如LNA、PNA、吗啉代等)或由其组成)、蛋白质分子(诸如肽、多肽、蛋白质或朊病毒或包括蛋白质或多肽组分等或其片段的任何分子)、或脂质或碳水化合物分子、或包含脂质或碳水化合物组分的任何分子。分析物可以是单个分子或含有两个或更多个分子亚基的复合物,例如,包括但不限于蛋白质-DNA复合物,其可以彼此共价结合或不彼此共价结合,并且可以相同或不同。因此,除了细胞或微生物之外,这种复合分析物也可以是蛋白质复合物或蛋白质相互作用。这种复合物或相互作用因此可以是同源多聚体或异源多聚体。分子的聚集体(例如,蛋白质)也可以是靶标分析物,例如相同蛋白质或不同蛋白质的聚集体。分析物也可以是蛋白质或肽与核酸分子(诸如DNA或RNA)之间的复合物,例如蛋白质与核酸(例如,染色质相关因子、调节因子(诸如转录因子)和DNA或RNA)之间的相互作用。

[0163] 本文公开的方法和组合物可以用于分析任何数量的分析物。例如,所分析的分析物的数量可以是存在于样品区域中或基底的单个特征内的至少约2种、至少约3种、至少约4种、至少约5种、至少约6种、至少约7种、至少约8种、至少约9种、至少约10种、至少约11种、至少约12种、至少约13种、至少约14种、至少约15种、至少约20种、至少约25种、至少约30种、至少约40种、至少约50种、至少约100种、至少约1,000种、至少约10,000种、至少约100,000或更多种不同的分析物。

[0164] 在本文所述的任何实施方案中,分析物包含靶序列。在一些实施方案中,靶序列对于样品可以是内源的,例如,与细胞的染色质缔合或在细胞的染色质中。生物样品可以包含一种或多种感兴趣的分析物,例如,各自参与相同分子或不同分子中的两个或更多个染色质区域的多个染色质相互作用事件。提供了用于执行多重测定以分析单个生物样品中的细胞中的两种或更多种不同染色质相互作用事件的方法。

[0165] II. 染色质相互作用的分析

[0166] 在真核基因组中,染色体DNA将自身缠绕在组蛋白周围(即,一核小体”),从而形成被称为染色质的复合物。染色质的紧密或松散包装有助于控制基因表达。紧密包装的染色

质(一封闭染色质”)通常不允许基因表达,而更松散包装、可接近的染色质区域(一开放染色质”)与基因产物的活跃转录相关。用于探测全基因组DNA可及性的方法已被证明在鉴定多种细胞类型中的调控元件和量化导致基因表达的激活或抑制的变化方面极为有效。示例性的基因组结构映射技术在US 10,526,639 B2中有所描述,所述专利以引用方式整体并入。

[0167] 当RNA聚合酶以及一般转录因子与启动子区域结合以形成起始前复合物(封闭复合物)时,真核生物中的基因转录开始。复合物从封闭状态转换为开放状态导致两条DNA链分离,并且将模板链呈递至RNA聚合酶的活性位点。启动子是通常紧邻转录起始位点(TSS)的区域,其结合并定位起始前复合物。TATA盒结合蛋白(TBP)与高度保守的TATA盒结合以引发转录复合物组装。真核基因还含有被称为顺式作用控制元件的其他调节序列,其与转录激活因子或抑制因子结合以增加或减少从核心启动子的转录。这些调控元件包括增强子、沉默子和绝缘子。在许多情况下,调控元件可以位于距其靶基因至多若干兆碱基的距离处。证据显示,基因表达的长程控制可以通过基因和这些调控元件之间的直接物理相互作用来介导。转录因子是一组参与转录起始和调控的蛋白质,其与启动子的特定序列元件结合并且将RNA聚合酶募集到转录起始位点。RNA聚合酶II的这些因子的实例包括TFIID、TFIIA、TFIIB、TFIIF、TFIIE、和TFIIH。Orphanides, G., Lagrange, T., Reinberg, D. (1996). The general transcription factors of RNA polymerase II, *Genes&Development*, 10 (21): 2657-83. doi:10.1101/gad.10.21.2657. 增强子是相对较短的DNA区域,其可以被转录因子结合以影响基因转录。人类基因组中有数十万个增强子,并且它们存在于原核生物和真核生物中。Pennacchio, L.A, Bickmore, W., Dean, A., Nobrega, M.A., Bejerano, G. (2013). Enhancers: five essential questions, *Nature Reviews. Genetics*, 14 (4): 288-95. doi: 10.1038/nrg3458; Kulaeva O.I., Nizovtseva, E.V., Polikanov, Y.S., Ulianov, S.V., Studitsky, V.M. (2012). Distant activation of transcription: mechanisms of enhancer action, *Molecular and Cellular Biology*, 32 (24): 4892-7. doi:10.1128/MCB.01127-12. 沉默子是与被称为抑制因子的转录调节因子结合的DNA序列,其阻止RNA聚合酶将DNA序列转录成RNA。最常见的位置位于靶标基因的上游。Maston, G., Sarah, E., Michael, G., (2006). Transcriptional regulatory elements in the Human Genome, *Annual Review of Genomics and Human Genetics*, 7: 29-59. doi:10.1146/annurev.genom.7.080505.115623.

[0168] 在一个方面,本文公开了分析生物样品的方法,其包括:a)获得生物样品;b)将包含至少两种单一寡核苷酸序列的第一组探针(例如,与染色质区域中的核酸链杂交的一种或多种探针)与至少两种一级靶核酸序列的有义链或反义链(例如,第一或第二核酸链)杂交,其中所述至少两种一级靶核酸序列参与染色质相互作用事件;c)将包含至少两种单一寡核苷酸序列的第二组探针(例如,与染色质区域中的核酸链杂交的一种或多种探针)与前述至少两种一级靶核酸序列的有义链或反义链(例如,第一或第二核酸链)杂交;d)将第三组探针(例如,一种或多种桥接探针或环化探针)与前述第二组探针杂交,其中所述第二探针(例如,与染色质区域中的核酸链杂交的一种或多种探针)由于与参与染色质相互作用事件的至少两种一级靶核酸序列杂交而在空间上彼此接近,并且其中所述第三探针(例如,一种或多种桥接探针或环化探针)桥接所述第二探针的末端;e)使用聚合酶延伸第三组探针

(例如,一种或多种桥接探针或环化探针)以形成环化的核酸,所述环化的核酸包含参与染色质相互作用事件的至少两种一级靶核酸序列的至少一部分;f)扩增所述环状核酸以生成包含可原位检测的信号产生标签的扩增子群;和g)通过所述信号产生标签原位检测所述带标签的扩增子群。

[0169] 在一些实施方案中,本文所述的方法在图1中举例说明。样品可以是固定的或不固定的,例如福尔马林固定的石蜡包埋 (FFPE) 样品、冷冻组织样品或新鲜组织样品。如图1所示,样品可以是基底(例如,载玻片上的组织)上的样品。在I.A部分中概述了适用于本文提供的方法的样品和样品制备技术。在一些实施方案中,样品的亚细胞结构和染色质构象在样品制备期间保持不变(例如,通过交联)。在一些实施方案中,通过所提供的方法来分析涉及第一和第二染色质区域的染色质相互作用事件。在一些实施方案中,生成扩增产物(例如,通过生成环状探针的滚环扩增(RCA)产物)并使用检测到的标记探针进行检测。在一些实施方案中,原位分析样品以例如平行或依次地检测多个染色质相互作用事件。

[0170] 如本文所用,术语“样品”一般是指受试者的生物样品。生物样品可以包含任何数量的大分子,例如细胞大分子。样品可以是细胞样品。样品可以是细胞系或细胞培养样品。样品可以包括一种或多种细胞。样品可以包括一种或多种微生物。生物样品可以是核酸样品或蛋白质样品。生物样品也可以是碳水化合物样品或脂质样品。生物样品可以来源于另一个样品。样品可以是组织样品,诸如活检、核心活检、针抽吸物或细针抽吸物。样品可以是流体样品,诸如血液样品、尿液样品或唾液样品。样品可以是皮肤样品。样品可以是面颊拭子。样品可以是血浆或血清样品。样品可以是无细胞(cell-free)或无细胞(cell free)样品。无细胞样品可以包括细胞外多核苷酸。细胞外多核苷酸可以从身体样品中分离,所述身体样品可以选自由以下组成的组:血液、血浆、血清、尿液、唾液、粘膜排泄物、痰、粪便和泪液。

[0171] 如本文所用,术语“基因组”一般是指来自受试者的基因组信息,其可以是例如受试者的遗传信息的至少一部分或全部。基因组可以在DNA或RNA中编码。基因组可以包含编码区(例如,编码蛋白质的编码区)以及非编码区。基因组可以包括生物体中所有染色体的序列。例如,人基因组通常总共有46条染色体。所有这些的序列一起可以构成人基因组。

[0172] 如本文所用,术语“染色质相互作用事件”一般是指不同染色质位点之间的长程相互作用,从而形成染色质的高级构象以维持染色质结构或促进基因表达。

[0173] 如本文所用,术语“转录调控区”或调控区一般是指基因组DNA内位于基因上游或下游的区域,例如10kb-1Mb、50kb-500kb、100kb-200kb,包括启动子、增强子。反式作用因子(例如,转录因子)结合的位点区域。

[0174] 如本文所用,术语“结合基序”一般是指存在于基因组DNA上并且可以被反式作用因子(诸如转录因子)结合以调控靶基因(例如,调控本发明中的效应基因的表达)的元件。

[0175] 如本文所用,术语“染色质相互作用频率”在本文中也称为“Hi-C相互作用频率”或Hi-C接触频率,一般是指在执行染色质构象分析时存在于染色质相互作用中的不同区域。相互作用的信号表示为两端落在特定区域内的Hi-C数据中的读段的数量。

[0176] 如本文所用,术语“染色质开放区域序列”一般是指由于核小体结合等而暴露在染色质中并且可以被反式作用因子(诸如转录因子)结合的DNA序列。

[0177] 如本文所用,术语“染色体构象捕获”或“染色质构象捕获技术”一般是指用于分

析细胞中的染色质的空间组织的所有技术。此类技术能够通过染色质的不同空间位置之间的关系建立染色质三维结构信息,所述技术还包括高通量测序染色质构象捕获技术。此类相互作用可以来自生物学功能,包括但不限于启动子-增强子相互作用。合适的染色体构象捕获技术过程的实例包括但不限于3C和Hi-C。这些过程已在以下出版物中进行了描述: Dekker, J. 等人 (2002). Capturing chromosome conformation, *Science*, 295, 1306-1311; Belton, J. 等人 (2012). Hi-C: A comprehensive technique to capture the conformation of genomes, *Methods*, 58, 3. 使用转座酶剪切和标记染色质DNA的另外的方法在US 2016/0208323 A1中有所描述,所述文献以引用方式整体并入。

[0178] 如本文所用,术语“高通量染色质构象捕获技术”一般是指将高通量测序技术与生物信息学分析方法相组合,以高效分析全基因组范围内的整个染色质DNA的空间位置并实现高分辨率。一种用于染色质三维结构和染色质相互作用信息的方法。技术至少包括Hi-C、基于Hi-C的改进技术原位Hi-C,以及基于原位Hi-C方法进一步引入桥接接头后获得的BL-Hi-C。另外,ChIA-PET方法也是一种高通量的染色质构象捕获技术。

[0179] 在一些实施方案中,可用技术可以用于确认通过本文所述的方法生成的数据。例如,如本文所用,术语“使用高通量测序的转座酶可接近染色质的测定”或“ATAC-seq”一般是指用于研究分子生物学中的染色质可及性的技术。它由ATAC实验和高通量测序两部分组成。ATAC-seq方法使用人工转座子探测DNA可及性,所述方法将特定序列插入染色质的可接近区域。因为转座酶只能将序列插入到不被转录因子和/或核小体结合的染色质可及区域中,所以测序读段可以用于推断染色质可及性增加的区域。因此,基因组中某些基因座序列的富集表明所述区域中没有核小体并且处于松散暴露状态,在所述松散暴露状态中核机构(诸如DNA结合蛋白)可以进入,从而提供关于染色质区段的转录活性状态的信息。

[0180] 在一些实施方案中,将对参与染色质相互作用事件的靶核酸序列进行测序。在一个实施方案中,将对包含参与染色质相互作用事件的全部或部分核酸序列的扩增产物(“扩增子”)进行测序。如本文所用,术语“测序”一般是指用于确定一种或多种多核苷酸中的核苷酸碱基序列的方法和技术。多核苷酸可以是例如核酸分子,诸如脱氧核糖核酸(DNA)或核糖核酸(RNA),包括其变体或衍生物(例如,单链DNA)。可以通过目前可用的各种系统执行测序,所述系统诸如但不限于 Illumina®、Pacific Biosciences (PacBio®)、Oxford Nanopore®或Life Technologies (Ion Torrent®)的测序系统。或者或另外,可以使用核酸扩增、聚合酶链式反应(PCR)(例如,数字PCR、定量PCR或实时PCR)或恒温扩增执行测序。此类系统可以提供对应于受试者(例如,人)的遗传信息的多种原始遗传数据,如通过系统由受试者提供的样品生成。在一些实例中,此类系统提供测序读段(在本文中也称为“读段”)。读段可以包括对应于已测序的核酸分子序列的一串核苷酸碱基。在一些情况下,本文提供的系统和方法可以与蛋白质组信息一起使用。本领域已知的其他测序方法,诸如使用可逆终止子通过延伸的测序、原位测序、荧光原位测序(FISSEQ)、焦磷酸测序、大规模平行签名测序(MPSS)等适用于本发明的方法中。可逆终止方法使用与可逆终止和可去除荧光相结合的逐步边合成边测序生物化学。

[0181] 原位测序技术在Ke, R. 等人, (2013). In situ sequencing for RNA analysis in preserved tissue and cells, *Nat. Methods*, 10, 857-860中有所描述。

[0182] FISSEQ是一种通过以下步骤延伸DNA的方法：将单一类型的荧光标记的三磷酸核苷酸添加到反应中；洗去未掺入的核苷酸，通过测量荧光来检测核苷酸的掺入，以及重复所述循环。在每个循环中，将来自先前循环的荧光漂白或数字减去，或者从核苷酸上切割荧光团并洗去。FISSEQ在例如Lee等人(2014).*Science*.343,1360-3中有所描述。

[0183] 焦磷酸测序是其中在每个核苷酸掺入事件期间(即,当将核苷酸添加到生长的多核苷酸序列中时)释放焦磷酸盐(PPi)的方法。在DNA聚合酶催化的反应中释放的PPi在可以目视检测到偶联反应中被ATP硫酸化酶和荧光素酶检测到。添加的核苷酸被核苷酸降解酶连续降解。在第一种添加的核苷酸被降解后,可以添加下一种核苷酸。随着这个程序的重复,可以推断出更长的模板序列。焦磷酸测序在Ronaghi等人(1998)*Science* 281,363中进一步描述。

[0184] MPSS同时利用基于连接的DNA测序。将包含所有可能的突出端的标记衔接子的混合物与四个核苷酸的靶序列退火。在衔接子成功连接后检测到标记。然后使用限制酶切割DNA模板以暴露接下来的四个碱基。MPSS在Brenner等人(2000)*Nat. Biotech.*18,630中进一步描述。

[0185] 在一个实施方案中,可以使用可检测标记或信号产生标签来对生物样品标记或加标签。通常,标记或标签将化学(例如,共价键合、氢键合或离子键合)结合样品或其组分。可检测标记可以对特定靶标(例如,生物标志物或分子类别)具有选择性,这可以用抗体或其他靶标特异性结合剂来实现。可检测标记可以包含可见组分,典型的如染料或荧光分子(例如,荧光团);然而,还可以考虑标记所使用的任何信号传导方式。例如,荧光标记的生物样品是通过如下技术进行标记以辅助显微分析(例如,荧光显微镜检查)的生物样品,所述技术诸如但不限于的免疫荧光、免疫组织化学或免疫细胞化学染色。合适的荧光显微镜检查技术的实例在Sanderson,M.J.等人(2014).*Fluorescence Microscopy*,Cold Spring Harbor Protocols,2014(10),doi:10.1101/pdb.top071795中有所描述。因此,可检测标记可以与生物样品或其目标组分化学连接。可检测标记可以是抗体和/或荧光染料,其中抗体和/或荧光染料还包含将样品与组合物、水凝胶或其他可溶胀材料连接或交联的物理、生物或化学锚或部分。可检测标记可以与核酸衔接子连接。标记的样品还可以包括多于一种标记。例如,每种标记可以具有特定的或可区分的荧光性质,例如可区分的激发和发射波长。此外,每种标记可以具有不同的靶特异性结合剂,所述结合剂对样品中的特定和可区分靶标或样品的组分具有选择性。

[0186] 如本文所用,术语“寡核苷酸”一般是指由两个或更多个脱氧核糖核苷酸或核糖核苷酸组成的分子。确切的尺寸将取决于许多因素,这些因素继而取决于寡核苷酸的最终功能或用途。寡核苷酸可以以任何方式产生,包括化学合成、DNA复制、逆转录或其组合。

[0187] 术语“引物”是指当置于引发引物延伸的条件下时可以充当合成起始点的寡核苷酸。寡核苷酸“引物”可以天然存在(如在纯化的限制性消化物中)或可以合成产生。将引物选择为在其3'端上具有与模板的特定序列链“基本上”互补的区域。引物必须充分互补才能与模板链杂交,从而发生引物延长。引物序列不必反映模板的确切序列。例如,非互补核苷酸片段可以与引物的5'端连接,而剩余的引物序列与所述链基本上互补。非互补碱基或更长的序列可以散布在引物中,前提条件是引物序列与模板序列有足够的互补性以进行杂交并且从而形成模板引物复合物以用于合成引物的延伸产物。

[0188] 在真核基因组中,染色体DNA将自身缠绕在组蛋白周围(即,“核小体”),从而形成被称为染色质的复合物。染色质的紧密或松散包装有助于控制基因表达。紧密包装的染色质(“封闭染色质”)通常不允许基因表达,而更松散包装、可接近的染色质区域(“开放染色质”)与基因产物的活跃转录相关。用于探测全基因组DNA可及性的方法已被证明在鉴定多种细胞类型中的调控元件和量化导致基因表达的激活或抑制的变化方面极为有效。

[0189] 本发明可用于诊断应用。例如,在一些实施方案中,本文所述的方法用于确定感兴趣基因的染色体构象状态和鉴定/表征其参与的顺式和反式染色质相互作用。某些疾病状态的特征是异常(例如,增加或减少)基因表达,其在一些发育状态下或在一些细胞类型中与染色质相互作用的不同模式相关。因此,方法通过检测染色质相互作用事件的这些模式可用于诊断此类疾病状态。在一些实施方案中,方法用于比较处于激活或失活状态的基因与测试样品的相互作用谱,以便确定基因的激活状态。

[0190] 在仍其他实施方案中,本文所述的方法可以用于检测表达改变的变体(例如,多态性)基因。例如,在一些基因中,某些单核苷酸多态性(SNP)的存在与疾病状态或改变的基因功能(例如,药物代谢)相关。在一些实施方案中,所公开的方法用于比较已知的SNP与测试样品的相互作用谱,以便确定变体基因的染色质相互作用,这在一些情况下可能与基因在一些发育阶段或一些细胞类型中的激活状态相关。

[0191] 在又其他的实施方案中,本文所述的方法可以用于检测指示基因组重排(包括但不限于易位、缺失、融合和倒位)的染色质相互作用模式。在一些实施方案中,本发明用于比较已知的基因重排与测试样品的相互作用谱,以便确定变体基因的染色质相互作用,这在一些情况下可能与基因在一些发育阶段或一些细胞类型中的激活状态相关。在一些实施方案中,待用作对照的相互作用谱是使用例如本发明的方法实验性地生成的。

[0192] 在另外的实施方案中,可以利用诊断签名。在一些实施方案中,诊断签名给出关于特定疾病的诊断倾向和预后的信息。例如,在一些实施方案中,诊断特征预测未来的基因组重排或检测与特定疾病状态或预后相关的染色体构象特征(例如,成环或反式相互作用)。在一些实施方案中,待用作对照的诊断签名(例如,指示给定疾病状态或预后)是使用例如本发明的方法实验性地生成的。

[0193] 在又其他的实施方案中,本发明的方法可用于研究应用。此类应用包括但不限于发育和分化中的基因调控的研究、疾病中的基因调控研究、药物代谢中的基因调控研究以及变体基因的调控的研究。在一些实施方案中,研究应用利用来自人受试者的样品。在其他实施方案中,研究应用利用来自非人动物(例如,非人哺乳动物)的测试样品。在一些实施方案中,非人动物是转基因动物。

[0194] 生物样品可以是包含交联的或能够交联的DNA的任何物理实体。然而,在一些实施方案中,方法不使用交联。样品可以是或可以来源于生物材料。样品可以是或可以来源于一种或多种细胞、一种或多种细胞核或一种或多种组织样品。实体可以是或可以来源于存在DNA(诸如染色质)的任何实体。样品可以是或可以来源于一种或多种分离的细胞或一种或多种分离的组织样品或一种或多种分离的细胞核。样品可以是或可以来源于活细胞和/或死细胞和/或核裂解物和/或分离的染色质。样品可以来源于待测试未来将罹患疾病的可能性的受试者。样品可以来源于有活力或无活力的患者材料。

[0195] 细胞可以来源于细胞培养物或从来自活生物体或死亡生物体(即,死后)、或来自

完整实验生物体(例如,完整的黑腹果蝇(*D.melanogaster*)胚胎或秀丽隐杆线虫(*C.elegans*)胚胎)、或来自微生物的混合物的特定组织离体分析。可以选择用于分析的细胞(例如通过在细胞周期的特定阶段同步细胞,或通过分选细胞,例如通过荧光激活细胞分选来捕获表达特定标志物的细胞类型(例如,使用对在感兴趣的细胞类型或细胞阶段中唯一表达的蛋白质具有特异性的抗体),或通过原位杂交进行检测,例如使用检测在感兴趣的细胞类型中特异性表达的特定例如mRNA或其他RNA的核酸探针或使用显示出特定基因的表达或特定阶段的特性的荧光标志物(诸如GFP)。例如,在Pitx3转录因子的启动子控制下的GFP转基因可以用于标记表达多巴胺的神经元(Maxwell S.等人(2005).Pitx3 regulates tyrosine hydroxylase expression in the substantia nigra and identifies a subgroup of mesencephalic dopaminergic progenitor neurons during mouse development.Dev.Biol.282(2):467-479)。细胞可以用剂预处理,例如,以测试药物对基因座的共分离或定位的影响,或者在生物体的生命周期期间研究,以了解发育、衰老和退化。

[0196] 任选地,根据组织的物种和类型来制备单细胞悬浮液,例如,可以制备哺乳动物实体组织的单细胞悬浮液。可以通过也与3C技术兼容的任何程序来进行单细胞悬浮液的制备。与保留交联的染色质接触的染色质样品生产兼容的若干单细胞制备的详细描述可见于例如Hagege,H.等人(2007).Quantitative analysis of chromosome conformation capture assays(3C-qPCR),Nature Protocols,2:1722。在多细胞生物体的情况下,单细胞悬浮液的制备可以从组织解剖开始,之后用胶原酶处理,或者对于软组织(例如,小鼠胸腺或胎肝),之后使组织通过细胞过滤网(例如,40微米网孔),或者在细胞在体外培养物或微生物培养物中生长的情况下,通过在适合细胞类型的力下离心培养物,之后以适当强度重悬以产生细胞损伤或死亡最少的单细胞悬浮液。使用已发表的方案或其后的开发,也可以应用于死后样品(Mitchell,A.C.等人(2014).The genome in three dimensions:a new frontier in human brain disease.Biol.Psychiatry,75:961)。使用植物材料(Grab,S.等人(2013).Characterization of chromosomal architecture in Arabidopsis by chromosome conformation capture.Genome Biology,14:R129)和昆虫组织(Ghavi-Helm,Y.等人(2014).Enhancer loops appear stable during development and are associated with paused polymerase,Nature,512:96),生产含有反映染色质接触的交联DNA片段的染色质的方法也是可能的。

[0197] 在一些实施方案中,细胞核、细胞、组织或整个生物体可以用交联剂(例如,化学交联剂)处理。在一些实施方案中,在将探针与生物样品杂交之前,可以用交联剂处理生物样品。交联剂诱导蛋白质彼此之间以及核酸(DNA和/或RNA)与蛋白质之间的键联。本发明的方法与也与当前基于3C的方法兼容的交联条件兼容。在一个实施方案中,交联剂是甲醛。例如,甲醛可以在缓冲溶液(例如,PBS pH 7.0-8.0的缓冲溶液)中以0.5%-4%或约1%-2%(全部重量/重量)的浓度使用,或直接将浓缩的交联剂溶液直接添加到细胞培养基中5-120分钟,或在一个实施方案中,约10-20分钟来使用。根据本发明,还可以使用除甲醛之外的交联剂,诸如UV光、丝裂霉素C、氮芥、美法仑、1,3-丁二烯二环氧化物、顺式二胺二氯铂(II)和环磷酸胺、双琥珀酰亚胺戊二酸酯、丙酸二硫代双琥珀酰亚胺酯以及戊二醛。

[0198] 在另一个实施方案中,没有交联步骤的方法也是可能的。例如,可以使用玻璃化细



胞的低温球磨(Oeffinger, M. 等人(2007). *Comprehensive analysis of diverse ribonucleoprotein complexes*, *Nature Methods*, 4:951-6; Hakhverdyan 等人(2015). *Rapid, optimized interatomic screening*, *Nature Methods*, 12:553)。在替代实施方案中, 诸如CUT&RUN的方案是合适的, 它们已成功用于谱分析不溶性染色质和检测无交联的长程3D接触。Skene, P. J., Henikoff, J. G., Henikoff, S. (2018). *Targeted in Situ Genome-Wide Profiling With High Efficiency for Low Cell Number*, *Nature Protocols*, 13(5), 1006-1019, doi:10.1038/nprot.2018.015。

[0199] 在一些实施方案中, 待测定的样品可以放置在基底或阵列上, 其是分子(诸如生物大分子(诸如肽或核酸分子))或生物样品(诸如组织切片)在基底上或基底内的可寻址位置中的排列。显微镜载玻片上的组织切片(诸如石蜡固定、未固定的冷冻切片, 包括FFPE切片)中的组织。

[0200] 在阵列内, 每个排列的样品都是可寻址的, 因为它的位置可以在阵列的至少两个维度内可靠且一致地确定。阵列上的特征应用位置可以采用不同的形状。例如, 阵列可以是规则的(诸如排列成均匀的行和列)或不规则的。因此, 在有序阵列中, 每个样品的位置在其被应用于阵列时分配给所述样品, 并且可以提供键以将每个位置与适当的靶标或特征位置相关联。通常, 有序阵列以对称网格图案排列, 但样品可以以其他图案排列(诸如呈放射状分布的线、螺旋线或有序簇)。可寻址阵列可以是计算机可读的, 因为可以对计算机进行编程以将阵列上的地址与关于该位置的样品的信息(诸如杂交或结合数据, 包括例如信号强度)相关联。在计算机可读格式的一些实例中, 阵列中的单个特征有规则地排列, 例如呈Cartesian网格图案, 所述单个特征可以通过计算机与位置信息相关联。

[0201] 组织阵列, 也称为组织微阵列或TMA, 包括单个显微镜载玻片上的多个正常和/或疾病组织(诸如具有或不具有相关正常相邻组织的癌组织)的切片。例如, 组织微阵列允许在单个实验中分析大量肿瘤上的一种或多种标志物的表达。

[0202] 在一些实施方案中, 染色质是基因组染色质。在示例性实施方案中, 染色质是真核细胞的基因组染色质。真核细胞的基因组染色质的主要功能可以是将DNA包装成更小的体积以适合细胞、增强DNA以允许有丝分裂、防止DNA损伤、并控制基因表达和DNA复制。染色质的结构取决于若干因素。整体结构取决于细胞周期的阶段: 在间期期间, 染色质在结构上是松散的, 以允许转录和复制DNA的RNA和DNA聚合酶进入。间期期间的染色质的局部结构取决于存在于DNA上的基因: 编码活跃转录的基因的DNA包装更松散, 并且发现与RNA聚合酶缔合(称为常染色质), 而发现编码非活性基因的DNA与结构蛋白缔合并且包装更紧密(异染色质)。染色质中结构蛋白的表观遗传化学修饰也改变了局部染色质结构, 在于通过甲基化和乙酰化对组蛋白进行特定的化学修饰。随着细胞准备分裂, 即进入有丝分裂或减数分裂, 染色质包装得更紧密, 以促进后期期间的染色体分离。

[0203] 如上所述, 真核细胞的基因组染色质可以包含组装成更高级结构或功能区域的DNA序列和多个DNA结合蛋白以及某些RNA序列。如本文所用, “染色质的结构或功能特征”是指特征在于或编码诸如调控功能(诸如启动子、终止子、翻译起始、增强子等)的功能或结构特征(诸如异染色质、常染色质、核小体、端粒或着丝粒)的染色质特征。核酸序列的物理特征可以包括功能作用, 并且反之亦然。如下所述, 本发明的染色质可以是染色质片段, 并且因此可以包括染色质的物理或功能特征的片段, 或者不包括物理或功能特征或已知的物

理或功能特征。

[0204] 基因组真核染色质的主要蛋白质组分是将DNA压缩成核小体的组蛋白。核小体包含八组组蛋白,其周围缠绕着一段长约150bp至约250bp的双链DNA序列。组蛋白H2A、H2B、H3和H4是核小体的一部分,而组蛋白H1可以将相邻的核小体连接在一起形成更高级的结构。组蛋白经受翻译后修饰,所述翻译后修饰可能影响组蛋白调控染色质功能的功能。此类修饰可以包括甲基化、瓜氨酸化、乙酰化、磷酸化、SUMO化、泛素化和ADP-核糖基化。

[0205] 许多其他多肽和蛋白质复合物与核小体和组蛋白相互作用以调控染色质功能。如本文所用的一多肽复合物”旨在描述组装在一起以形成统一缔合的因子的蛋白质和多肽。多肽复合物的成员可以经由非共价键或共价键彼此相互作用。通常,多肽复合物的成员将协同作用以实现与核酸序列结合或与已经与染色质中的核酸序列缔合或结合的多肽和蛋白质结合。染色质缔合多肽复合物可以包含多种蛋白质和/或多肽,它们各自用于与可以与复合物永久缔合或瞬时缔合的其他多肽相互作用,这取决于细胞条件和细胞周期内的位置。因此,响应于不同的生理条件或作为细胞周期的因素,特定的多肽复合物的组成成员在不同发育阶段可能不同。例如,在动物中,具有已知染色质重塑活性的多肽复合物包括Polycomb群基因沉默复合物,以及Trithorax群基因激活复合物。

[0206] 此外,与本发明的染色质缔合的蛋白质可以是在细胞中正常表达的蛋白质或可以是在细胞中表达的外源异源蛋白质。在一些实施方案中,与本发明的染色质缔合的蛋白质是在细胞中正常表达的蛋白质。在其他实施方案中,与本发明的染色质缔合的蛋白质是在细胞中正常表达的蛋白质。

[0207] 在一些实施方案中,本文所述的方法分析感兴趣区域中的核苷酸序列或靶核苷酸序列与其他序列之间的相互作用。靶核苷酸序列可以是一个(或多个)染色体内的感兴趣的基因组区域,并且可以包括调控元件。在一些实施方案中,调控元件选自包括启动子、增强子、沉默子或绝缘子的组。靶核苷酸序列可以包含感兴趣的特定基因座。基因座是基因或DNA序列的特定位置或染色体上的位置。感兴趣的基因组区域可以包含特定基因座(诸如基因序列)以及一个或两个侧翼区域。

[0208] 一其他核苷酸序列”(即感兴趣区域内的核苷酸序列与其相互作用的核苷酸序列)本身可以位于感兴趣的区域中,或者它们可以来自其他区域,诸如来自不同染色体的相同染色体的其他部分。当基因的调控已经变化或当基因丢失时,如果发生疾病,与此类区域的相互作用可能发生变化。

[0209] 如先前所述,本发明的方法可以利用描述于Dekker等人,(2002) Science 295:1306中的染色质构象捕获(3C)或描述于Belton,J.等人(2012).Hi-C:A comprehensive technique to capture the conformation of genomes,Methods,58(3)中的Hi-C.3C样模板可以使用已知方法制备,诸如由Splinter等人,(2004)Methods Enzymol.375,493-507描述的方法。简而言之,在一个实施方案中,使用交联剂(诸如甲醛)固定样品(诸如细胞、组织或细胞核)。然后执行初级限制性酶消化,使得DNA在交联细胞核的背景中被消化。然后在低DNA浓度下执行分子内连接,这有利于交联DNA片段之间的连接(即,分子内连接)而不是非交联DNA片段之间的连接(即,分子间或随机连接)。接下来,逆转交联,并且可以纯化DNA。产生的3C模板含有连接的限制性片段,因为它们最初靠近核空间。

[0210] 在一个实施方案中,可以使用可检测标记来对生物样品标记或加标签。如本文所

用的术语“标记”是指可以用于提供可检测信号并且可以与核酸或蛋白质连接的任何原子或分子。标记可提供可通过荧光、放射性、比色法、重量法、X射线衍射或吸收、磁性、酶活性等检测到的信号。

[0211] 如本文所用,术语“杂交(hybridize/hybridization)”是指互补序列通过碱基配对相互作用与靶核酸(待检测的序列)退火(Marmur and Lane, Proc. Natl. Acad. Sci. USA 46:453 [1960]和Doty等人, Proc. Natl. Acad. Sci. USA 46:461 [1960])。术语“退火的”和“杂交的”自始至终可互换使用,并且旨在涵盖寡核苷酸和靶核酸之间的任何特异性和可再现的相互作用,包括仅具有部分互补性的区域的结合和利用非规范相互作用以获得稳定性和/或特异性的结合相互作用。能够选择性杂交的核苷酸序列将一般与寡核苷酸探针长度上的对应互补核苷酸序列至少75%、85%、90%、95%或98%同源。选择性由杂交过程中的盐和温度条件确定。“特异性杂交”是指分子在严格条件下(例如,65°C和 $0.1 \times \text{SSC}$  { $1 \times \text{SSC} = 0.15\text{M NaCl}, 0.015\text{M}$ 柠檬酸钠pH 7.0})仅与核苷酸序列结合、双链体化或杂交。严格条件是寡核苷酸探针将与其靶序列杂交但不与其他序列杂交的条件。严格条件是序列依赖性的并且在不同的环境中是不同的。较长的序列在较高温度下特异性杂交。一般来讲,非常严格的条件选择为比特定序列在限定的离子强度和pH下的热解链点( $T_m$ )低约5°C。杂交温度是低于解链温度( $T_m$ )的温度,并且杂交温度越接近 $T_m$ ,杂交越严格,这意味着错配的DNA序列将不彼此杂交。在一些实施方案中,寡核苷酸序列超过基因组DNA以确保有效的(和可量化的)杂交。通常,严格条件包括在pH 7.0至8.3下至少约0.01至1.0M Na离子浓度(或其他盐)的盐浓度。严格条件还可以通过添加去稳定剂(诸如甲酰胺或四烷基铵盐)来实现。核酸双链体的稳定性通过解链温度或 $-T_m$ 来测量。特定条件下核酸双链体的 $T_m$ 是平均一半碱基对已经解离的温度。

[0212] 如本文所用的术语“探针”是指如下寡核苷酸,其由于探针中的至少一个序列与另一个核酸中的序列的互补性或其他可再现的吸引相互作用的方式而与另一个核酸中的序列形成双链体结构或其他复合物。合适地,寡核苷酸探针的长度将至少为15、20、25、30或40个核苷酸。

[0213] 参考附图,图1示出了用于在生物组织样品中原位分析第一和第二染色质区域之间的染色质相互作用事件的示例性方法。如图1所示,可以首先处理生物样品(诸如组织样品)以保留或捕获与染色质相互作用事件(诸如第一和第二染色质区域之间的染色质相互作用事件)相关的染色质构象。然后可以使样品与第一探针和第二探针同时接触或以任意顺序依次接触。由于染色质构象的保留或捕获,第一染色质区域和第二染色质区域保持彼此接近,即使在染色质相互作用事件(和/或参与相互作用的染色质相关因子)不再存在之后也是如此。在一些实施方案中,第一探针与第一染色质区域中的第一条核酸链杂交,并且第二探针与第二染色质区域中的第二条核酸链杂交,并且第一和第二探针彼此桥接或通过一种或多种桥接探针桥接。在一些实施方案中,方法包括连接第一和/或第二探针的末端或一种或多种桥接探针的末端以形成环状探针,其中环状探针包含第一核酸链或其互补链的序列和/或第二核酸链或其互补链的序列,其中检测环状探针的扩增产物,从而分析染色质相互作用事件,例如使用原位分析。

[0214] 在一些实施方案中,方法可以包括分别连接(例如,使用连接反应)用于第一和第二染色质区域的第一和第二探针。在一些实施方案中,方法可以包括经由两个桥接探针来

连接(例如,使用连接反应)第一探针和第二探针。在一些实施方案中,方法可以包括连接(例如,使用连接反应)两个桥接探针以形成环状探针。在一些方面,可以设计第一和/或第二探针(例如,检测探针)、桥接探针和染色质接近探针的合适组合以形成环状探针。

[0215] 在一些实施方案中,使样品与作为染色质接近探针的探针接触。在一些实施方案中,第一和/或第二探针是染色质接近探针,例如,第一和/或第二染色质接近探针。在一些实施方案中,第一染色质区域中的第一核酸链的互补链和/或第二染色质区域中的第二核酸链的互补链不与探针(例如,染色质接近探针)杂交。在一些实施方案中,第一染色质区域中的第一核酸链的互补链与第一染色质接近探针杂交。在一些实施方案中,第二染色质区域中的第二核酸链的互补链与第二染色质接近探针杂交。在一些实施方案中,使样品同时与第一和第二染色质接近探针接触。在一些实施方案中,使样品依次与第一和第二染色质接近探针接触。在一些实施方案中,使样品在与第二染色质接近探针接触之前与第一染色质接近探针接触。在一些实施方案中,使样品在与第一染色质接近探针接触之前与第二染色质接近探针接触。

[0216] 图2A示出了一种或多种探针与染色质区域中的核酸链杂交的示例性实施方案,其中第一和第二探针是染色质接近探针。第一染色质区域中的第一核酸链的互补链和/或第二染色质区域中的第二核酸链的互补链可以但不需要结合染色质接近探针。使互补链成为单链以为另外的探针(诸如具有突出端、可以连接以形成环状探针的检测探针)提供结合位点。

[0217] 图2B示出了一种或多种探针与染色质区域中的核酸链杂交的示例性实施方案,其中第一探针是与第一染色质区域杂交的染色质接近探针/检测探针,并且第二探针是与第二染色质区域中的第二核酸链杂交的检测探针。提供单独的染色质接近探针以打开第二染色质区域中的双链体,使得第二探针可以结合。

[0218] 图3A示出了一种或多种探针与染色质区域中的核酸链杂交的示例性实施方案,其中提供第一和第二染色质接近探针以打开第一和第二染色质区域中的双链体,使得第一和第二探针(具有突出端的检测探针)可以结合。

[0219] 图3B示出了一种或多种探针与染色质区域中的核酸链杂交的示例性实施方案。第一染色质接近探针可以包括用于打开染色质中的双链体的多个染色质接近探针。在第一染色质区域中,与互补链上的相邻序列杂交的两种染色质接近探针使第一核酸链成为单链。第一探针的靶结合序列可能与两种染色质接近探针不互补,以便减少探针之间的双链体形成。在染色质接近探针/互补染色质链杂交时,第二染色质接近探针可以与互补链交联(例如,在第二染色质区域中)以减少探针双链体的形成并有利于检测探针/染色质链双链体的形成。

[0220] 在一些实施方案中,第一染色质接近探针包含肽核酸(PNA)。在一些实施方案中,第二染色质接近探针包含肽核酸(PNA)。在一些实施方案中,第一和第二染色质接近探针均包含肽核酸(PNA)。在一些实施方案中,第一和/或第二染色质接近探针是PNA探针。在一些实施方案中,使样品与作为PNA探针的第一和第二染色质接近探针(例如,第一染色质接近探针是PNA探针和/或第二染色质接近探针是PNA探针)同时接触或以任意顺序依次接触。

[0221] 在一些实施方案中,第一染色质接近探针包含一种或多种天然核酸和/或一种或多种修饰的或合成的核酸类似物,诸如肽核酸(PNA)。在一些实施方案中,第二染色质接近

探针包含一种或多种天然核酸和/或一种或多种合成核酸类似物。在一些实施方案中,第一和第二染色质接近探针均包含一种或多种天然核酸、一种或多种修饰的核酸和/或一种或多种合成核酸类似物。

[0222] 在一些实施方案中,一种或多种合成核酸类似物是异种核酸(XNA),任选地选自自由以下组成的组:1,5-脱水己糖醇核酸(HNA)、环己烯核酸(CeNA)、苏糖核酸(TNA)、乙二醇核酸(GNA)、锁核酸(LNA)、肽核酸(PNA)和氟阿拉伯核酸(FANA)、及其任何组合。

[0223] 在一些实施方案中,第一或第二核酸链和对应互补链之间的杂交的解链温度( $T_m$ )比第一或第二染色质接近探针和对应互补链之间的杂交的 $T_m$ 低例如约5°C、约10°C、约15°C、约20°C、约25°C、约30°C、约35°C、约40°C、约45°C、约50°C、或大于50°C。

[0224] 在一些实施方案中,在允许探针接近第一和/或第二染色质区域中的单链序列的条件下使样品与染色质接近探针(例如,第一和/或第二染色质接近探针)接触。在一些实施方案中,允许探针接近第一和/或第二染色质区域中的单链序列的条件包括高浓度的染色质接近探针。在一些实施方案中,使样品在至少约100nM、200nM、500nM、1 $\mu$ M、2 $\mu$ M或大于约2 $\mu$ M的探针浓度下与染色质接近探针接触。在一些实施方案中,使样品在约100nM和约5 $\mu$ M之间(诸如约100nM至约1 $\mu$ M、约500nM至约2 $\mu$ M、约1 $\mu$ M至约3 $\mu$ M、和约2 $\mu$ M至约5 $\mu$ M中任一个之间)的探针浓度下与染色质接近探针接触。

[0225] 在一些实施方案中,使样品在至少约40°C、45°C、50°C或大于约50°C的温度下与染色质接近探针接触。在一些实施方案中,使样品在约40°C和约100°C之间(诸如约40°C至约45°C、约45°C至约50°C、约50°C至约55°C、约55°C至约60°C的温度、或约60°C至约100°C中任一个之间)的温度下与染色质接近探针接触。在一些实施方案中,将样品与染色质接近探针一起孵育至少30分钟、1小时、2小时、5小时或超过5小时。

[0226] 在一些实施方案中,去除染色质接近探针的分子。在一些实施方案中,方法包括例如使用一次或多次洗涤诸如严格洗涤去除不特异性杂交的染色质接近探针的分子。

[0227] 在一些实施方案中,连接第一和/或第二探针的末端(任选地其中第一和/或第二探针是染色质接近探针)以形成环状探针。在一些实施方案中,可以直接连接第一和/或第二探针的末端以形成环状探针。在一些实施方案中,第一和/或第二探针彼此桥接。在一些实施方案中,第一和/或第二探针分别包含3'突出端和5'突出端,所述3'突出端和5'突出端位于与第一和第二核酸链杂交的序列的侧翼。图4图解说明了第一探针的3'突出端可以与第二探针的5'突出端至少部分互补,且/或第一探针的5'突出端可以与第二探针的3'突出端至少部分互补。在一些实施方案中,使用第二探针作为模板来延伸第一探针的3'端(例如,使用引物延伸和/或连接反应),以及将第一探针的延伸的3'端与第一探针的任选延伸的5'端连接(例如,使用连接反应),从而环化第一探针以形成环状探针,其中环状探针包含第一核酸链或其互补链的序列和/或第二核酸链或其互补链的序列(图4)。

[0228] 在一些实施方案中,第一和第二探针(任选地其中第一和第二探针是染色质接近探针)通过一种或多种桥接探针桥接(图5A-5B)。在一些实施方案中,连接一种或多种桥接探针的末端以形成环状探针,其中环状探针包含第一核酸链或其互补链的序列和/或第二核酸链或其互补链的序列。在图5A中,第一探针的3'突出端和第二探针的5'突出端与第一桥接探针序列至少部分互补,并且第一探针的5'突出端和第二探针的3'突出端与第二桥接探针序列至少部分互补。在一些实施方案中,第一和第二桥接探针序列在相同的桥接探针

中或在不同的桥接探针中。任选地,桥接探针可以包含接头/间隔物序列(图5B;桥接探针内的虚线)。在一些实施方案中,第一探针的3'突出端和第二探针的5'突出端可以与第一桥接探针至少部分互补,并且第一探针的5'突出端和第二探针的3'突出端可以与第二桥接探针至少部分互补。在一些实施方案中,第一和/或第二探针包含不与第一和/或第二桥接探针互补的另外的3'突出端。在一些实施方案中,第一和/或第二探针不包含不与第一和/或第二桥接探针互补的另外的3'突出端。

[0229] 在一些实施方案中,可以使用第一探针作为模板来延伸第一桥接探针的3'端(例如,使用引物延伸和/或连接反应)。在一些实施方案中,方法包括将第一桥接探针的延伸的3'端与第二桥接探针的任选延伸的5'端连接(例如,使用连接反应)。在一些实施方案中,方法包括使用第二探针作为模板来延伸第二桥接探针的3'端(例如,使用引物延伸和/或连接反应),以及将第二桥接探针的延伸的3'端与第一桥接探针的任选延伸的5'端连接(例如,使用连接反应),从而连接第一和第二桥接探针以形成环状探针。

[0230] 图6A-6B示出了方法,其中第一和/或第二桥接探针序列位于与第一染色质区域中的第一核酸链的互补链杂交的探针(例如,染色质接近探针)中和/或与第二染色质区域中的第二核酸链的互补链杂交的探针(例如,染色质接近探针)中。在一些实施方案中,第一和第二桥接探针序列位于与第一染色质区域中的第一核酸链的互补链杂交的探针(例如,染色质接近探针)中和/或与第二染色质区域中的第二核酸链的互补链杂交的探针(例如,染色质接近探针)中(例如,图6A)。在一些实施方案中,第一桥接探针序列位于与第一染色质区域中的第一核酸链的互补链杂交的探针(例如,染色质接近探针)中和/或与第二染色质区域中的第二核酸链的互补链杂交的探针(例如,染色质接近探针)中。在一些实施方案中,第二桥接探针序列不位于与第一染色质区域中的第一核酸链的互补链杂交的探针(例如,染色质接近探针)中和/或与第二染色质区域中的第二核酸链的互补链杂交的探针(例如,染色质接近探针)中(例如,图6B)。

[0231] 在一些实施方案中,第一探针的3'端通过一个桥接探针与第二探针的5'端连接,并且第一探针的5'端与第二探针的3'端彼此连接(例如,经由与第一染色质区域中的第一核酸链的互补链杂交的探针(例如,染色质接近探针)和/或与第二染色质区域中的第二核酸链的互补链杂交的探针(例如,染色质接近探针)中的第一桥接探针序列)。在一些实施方案中,第一探针的5'端通过一个桥接探针与第二探针的3'端连接,并且第一探针的3'端与第二探针的5'端彼此连接(例如,经由与第一染色质区域中的第一核酸链的互补链杂交的探针(例如,染色质接近探针)和/或与第二染色质区域中的第二核酸链的互补链杂交的探针(例如,染色质接近探针)中的第一桥接探针序列)。

[0232] 在一些实施方案中,环状探针在样品中原位形成。

[0233] 在一些实施方案中,探针寡核苷酸之间的间隙可以在连接之前首先使用例如Mu聚合酶、DNA聚合酶、RNA聚合酶、逆转录酶、VENT聚合酶、Taq聚合酶和/或其任何组合、衍生物和变体(例如,工程化突变体)进行填充。在一些实施方案中,测定还可以包括连接产物的扩增。

[0234] 在一些实施方案中,例如使用一次或多次洗涤诸如严格洗涤去除不特异性杂交的第一和/或第二探针和/或一种或多种桥接探针的分子。在一些实施方案中,例如使用一次或多次洗涤诸如严格洗涤去除不特异性杂交的第一和/或第二探针和/或一种或多种桥接

探针的分子。

[0235] 在一些实施方案中,第一探针、第二探针、桥接探针、环状探针和/或染色质接近探针包含条形码序列。在一些实施方案中,环状探针包含一种或多种条形码序列,所述条形码序列对应于例如第一染色质区域和/或第二染色质区域中的感兴趣的核酸序列。

[0236] 在一些实施方案中,第一探针、第二探针、桥接探针、环状探针和/或染色质接近探针的条形码序列可以是相同的或不同的。这些核酸条形码可以用于例如通过样品、生物体等来标记片段化DNA,例如使得可以同时分析多个样品同时保留关于样品来源的信息。一般来讲,条形码可以包括可以用于鉴定一种或多种特定核酸的一种或多种核苷酸序列。条形码可以是人工序列,或者可以是天然存在的序列。条形码可以包含至少约1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20或更多个连续核苷酸。在一些实施方案中,条形码包含至少约10、20、30、40、50、60、70、80、90、100或更多个连续核苷酸。在一些实施方案中,包含条形码的核酸群体中的至少一部分条形码是不同的。在一些实施方案中,至少约10%、20%、30%、40%、50%、60%、70%、80%、90%、95%、99%的条形码是不同的。在更多此类实施方案中,所有的条形码都是不同的。包含条形码的核酸群体中的不同条形码的多样性可以随机生成或非随机生成。在一些实施方案中,转座子序列包含至少一个条形码。在一些实施方案中,转座子序列包含含有第一条条形码序列和第二条条形码序列的条形码。在一些此类实施方案中,可以鉴定第一条条形码序列或指定其与第二条条形码序列配对。例如,使用包括多个已知彼此配对的第一和第二条条形码序列的参考表,可以知道已知的第一条形码序列与已知的第二条条形码序列配对。在另一个实例中,第一条条形码序列可以包含与第二条条形码序列相同的序列。在另一个实例中,第一条条形码序列可以包含第二条条形码序列的反向互补序列。在一些实施方案中,第一条条形码序列和第二条条形码序列是不同的(一双码(bi-code)”)。应当理解,在一些实施方案中,大量可用的条形码允许每种带条形码的核酸分子包含唯一身份标识。模板核酸混合物中的每种分子的唯一身份标识可以用于若干种应用,以例如在单倍型测序、亲本等位基因鉴别、宏基因组测序和基因组的样品测序中鉴定具有多个染色体、基因组、细胞、细胞类型、细胞疾病状态和物种的样品中的单个核酸分子。

[0237] 在一些实施方案中,本文提供的方法可以是多重的(例如,分析样品以检测多个染色质相互作用事件)。在一些实施方案中,多个染色质相互作用事件的检测平行发生。在一些实施方案中,多个染色质相互作用事件的检测依次发生。在一些实施方案中,第一环状探针与第一染色质相互作用事件相关,并且第二环状探针与第二染色质相互作用事件相关,并且第一和第二环状探针各自包含对应于第一和/或第二探针或染色质相互作用事件的条形码序列或其互补序列。

[0238] 在一些实施方案中,染色质相互作用事件涉及除了第一和第二染色质区域之外的一个或多个染色质区域,并且环状探针还包含一个或多个其他染色质区域中的核酸链的序列或其互补序列。

[0239] 参考图7-10中描绘的实施方案,在步骤1和2中,用生物样品(例如,组织样品)制备载玻片并使载玻片经受甲醛交联剂以将染色质与相关因子化学交联。在图8中,引入了包含特异性靶向相互作用的感兴趣基因组区域(例如,在空间上彼此紧密接近的相互作用的启动子和增强子核酸序列)的寡核苷酸序列的肽核酸(PNA)探针,并将其与基因组序列杂交。另外,如图8所示,引入了包含特异性靶向相同的相互作用的感兴趣基因组区域的寡核苷酸

序列的检测探针。在一个实施方案中,检测探针与双链基因组序列的一条链杂交,而PNA探针与双链基因组序列的另一条链杂交。检测探针还可以包含3'和5'突出端或延伸。在图9中,引入了环化探针并使其与检测探针的3'和5'突出端杂交,因此在与增强子区域杂交的检测探针和与参与染色质相互作用事件的启动子区域杂交的检测探针之间形成桥。如图9所示,使用DNA聚合酶延伸环化探针以形成包含感兴趣的靶相互作用序列的全部或部分的统一环状核酸分子。

[0240] 在图10中,经由滚环扩增(RCA)原位扩增统一环状核酸,并进行荧光团标记以用于通过荧光原位杂交(FISH)进行的检测。利用锁式探针和滚环扩增的合适策略在Qian,X.等人(2018). *A spatial atlas of inhibitory cell types in mouse hippocampus*, bioRxiv doi:<https://doi.org/10.1101/431957>中有所描述。在图10中,示例了可检测标记的RCA扩增子。然而,扩增子本身可以不被可检测标记,并且可以被可检测标记的探针(未示出)检测。或者,扩增子可以被中间探针(其本身可以不被可检测标记;未示出)识别,并且中间探针可以被可检测标记的探针检测。如本文所用,可检测标记的寡核苷酸(例如,检测探针)用可检测部分进行标记。在一些实施方案中,可检测标记的寡核苷酸包含一个可检测部分。在一些实施方案中,可检测标记的寡核苷酸包含两个或更多个可检测部分。在一些实施方案中,可检测标记的寡核苷酸具有一个可检测部分。在一些实施方案中,可检测标记的寡核苷酸具有两个或更多个可检测部分。用于结合和鉴定靶核酸的探针和方法已在例如US2003/0013091、US2007/0166708、US2010/0015607、US2010/0261026、US2010/0262374、US2010/0112710、US2010/0047924和US2014/0371088中有所描述,所述专利各自以引用方式整体并入本文。在一些实施方案中,可检测部分是或包含纳米材料。在一些实施方案中,可检测部分是或包含纳米颗粒。在一些实施方案中,可检测部分是或包含量子点。在一些实施方案中,可检测部分是量子点。在一些实施方案中,可检测部分包含量子点。在一些实施方案中,可检测部分是或包含金纳米颗粒。在一些实施方案中,可检测部分是金纳米颗粒。在一些实施方案中,可检测部分包含金纳米颗粒。本领域技术人员理解,在一些实施方案中,用于特定探针的标记的选择可以基于多种因素来确定,包括例如尺寸、生成的信号类型、连接或掺入探针中的方式、细胞成分的性质(包括它们在细胞内的位置)、细胞的性质、被分析的相互作用的类型等。例如,在一些实施方案中,将探针用Cy3或Cy5标记,所述Cy3或Cy5已经被合成以携带N-羟基琥珀酰亚胺酯(NHS酯)反应基团。由于NHS酯容易与脂肪胺基团反应,因此可以用氨基烷基修饰核苷酸。这可以通过在合成反应期间掺入经氨基烷基修饰的核苷酸来完成。在一些实施方案中,每60个碱基使用一个标记以避免淬灭效应。

[0241] 在一些实施方案中,可以使用一组或两组探针,例如,不包括检测探针或PNA探针。另外,可以使用替代探针设计,例如锁式探针,只要最终扩增子包含参与染色质相互作用事件的靶序列(例如,启动子/增强子)的全部或部分即可。

[0242] 在一些实施方案中,本文描述的方法可以是多重的。例如,在检测和成像技术之后,可以对样品(例如,基底上的生物或组织样品)进行化学处理(例如,漂白)并从第一个标记淬灭信号,并且对参与染色质相互作用事件的其他靶核酸序列重复上述程序。所述程序可以重复多轮,直到被分子降解或其他抑制性问题阻止—每一轮利用可区分的标签,从而跨细胞群创建多个染色质相互作用事件的分层阵列。可商购获得的合适技术是可用的,例如MACSima<sup>TM</sup> Imaging Platform或Akoya Biosciences Phenoptics Multispectral



Imaging platform.

[0243] 在一些实施方案中,可以通过原位测序技术或其他合适的技术对滚环扩增生成的扩增产物进行原位测序。

[0244] 在一些实施方案中,扩增产物包含一种或多种条形码序列或其互补序列、第一核酸链或其互补链的序列、以及第二核酸链或其互补链的序列的多个拷贝。在一些实施方案中,扩增产物在样品中原位形成。在一些实施方案中,扩增产物使用环状探针的滚环扩增(RCA)形成,任选地其中RCA是线性RCA、分支RCA、树突状RCA或其任何组合。在一些实施方案中,使用Phi29聚合酶形成扩增产物。在一些实施方案中,扩增在约20°C和约60°C之间(任选地在约30°C和约40°C之间)的温度下执行。

[0245] 在一些实施方案中,扩增通过执行滚环扩增(RCA)来实现。在其他实施方案中,添加与环状探针或环化的探针杂交的引物并照此用于扩增。在一些实施方案中,RCA包括线性RCA、分支RCA、树枝状RCA或其任何组合。

[0246] 在一些实施方案中,扩增在约20°C和约60°C之间的温度下执行。在一些实施方案中,扩增在约30°C和约40°C之间的温度下执行。在一些方面,扩增步骤(诸如滚环扩增(RCA))在处于或约25°C和处于或约50°C之间(诸如处于或约25°C、27°C、29°C、31°C、33°C、35°C、37°C、39°C、41°C、43°C、45°C、47°C或49°C)的温度下执行。

[0247] 在一些实施方案中,在适当的dNTP前体和其他辅因子的存在下添加DNA聚合酶后,将引物延长以产生环状模板的多个拷贝。这种扩增步骤可以利用恒温扩增或非恒温扩增。在一些实施方案中,在杂交复合物形成和扩增探针缔合后,对杂交复合物进行滚环扩增以产生含有cDNA的多个拷贝的cDNA纳米球(即,扩增子)。用于滚环扩增(RCA)的技术是本领域已知的,诸如线性RCA、分支RCA、树突状RCA或其任何组合。(参见例如,Baner等人,Nucleic Acids Research,26:5073-5078,1998;Lizardi等人,Nature Genetics 19:226,1998;Mohsen等人,Acc Chem Res.2016November 15;49(11):2540-2550;Schweitzer等人Proc.Natl Acad.Sci.USA 97:101 13-1 19,2000;Faruqi等人,BMC Genomics 2:4,2000;Nallur等人,Nucl.Acids Res.29:e118,2001;Dean等人Genome Res.11:1095-1099,2001;Schweitzer等人,Nature Biotech.20:359-365,2002;美国专利号6,054,274、6,291,187、6,323,009、6,344,329和6,368,801)。用于RCA的示例性聚合酶包括DNA聚合酶,诸如phi29( $\phi$ 29)聚合酶、Klenow片段、嗜热脂肪芽孢杆菌(*Bacillus stearothermophilus*)DNA聚合酶(BST)、T4 DNA聚合酶、T7 DNA聚合酶或DNA聚合酶I。在一些方面,可以采用已被工程改造或突变以具有期望特性的DNA聚合酶。在一些实施方案中,聚合酶是phi29 DNA聚合酶。

[0248] 在一些方面,在扩增步骤期间,可以将修饰的核苷酸添加到反应中以将修饰的核苷酸掺入扩增产物(例如,纳米球)中。示例性的修饰的核苷酸包括胺修饰的核苷酸。在方法的一些方面,例如,对于将生成的扩增产物(例如,纳米球)与支架、细胞结构和/或其他扩增产物(例如,其他纳米球)锚定或交联。在一些方面,扩增产物包含修饰的核苷酸,诸如胺修饰的核苷酸。在一些实施方案中,胺修饰的核苷酸包含丙烯酸N-羟基琥珀酰亚胺部分修饰。其他胺修饰的核苷酸的实例包括但不限于5-氨基烯丙基-dUTP部分修饰、5-炔丙基氨基-dCTP部分修饰、N6-6-氨基己基-dATP部分修饰或7-脱氮-7-炔丙基氨基-dATP部分修饰。

[0249] 在一些方面,多核苷酸和/或扩增产物(例如,扩增子)可以锚定至聚合物基质。例如,聚合物基质可以是水凝胶。在一些实施方案中,可以修饰多核苷酸探针中的一种或多种

以含有官能团,所述官能团可用作锚定位点以将多核苷酸探针和/或扩增产物与聚合物基质连接。可以根据所提供的实施方案采用的示例性修饰和聚合物基质包括例如描述于WO 2014/163886、WO 2017/079406、US 2016/0024555、US 2018/0251833和US 2017/0219465中的那些。在一些实例中,支架还含有修饰或官能团,所述修饰或官能团可以与探针组或扩增产物的修饰或官能团反应或掺入其中。在一些实例中,支架可以包含寡核苷酸、聚合物或化学基团,以提供基质和/或支撑结构。

[0250] 扩增产物一般可以固定在基质内被扩增的核酸的位置,从而创建扩增子的局部集落。扩增产物可以通过空间因子固定在基质内。扩增产物也可以通过共价或非共价键合固定在基质内。以这种方式,可以认为扩增产物附着在基质上。通过固定在基质上(诸如通过共价键合或交联),维持了原始扩增子的尺寸和空间关系。通过固定在基质上(诸如通过共价键合或交联),扩增产物在机械应力下对移动和拆散具有抗性。

[0251] 在一些方面,扩增产物与周围基质共聚和/或共价连接,从而保留它们的空间关系和其固有的任何信息。例如,如果扩增产物是由包埋在基质中的细胞内的DNA或RNA生成的那些,则也可以将扩增产物官能化以形成与基质的共价连接,从而保留它们在细胞内的空间信息,从而提供亚细胞定位分布模式。在一些实施方案中,所提供的方法涉及在水凝胶亚基的存在下包埋一种或多种多核苷酸探针组和/或扩增产物以形成一种或多种水凝胶包埋的扩增产物。在一些实施方案中,所描述的水凝胶-组织化学包括将核酸与原位合成的水凝胶共价连接以用于组织清除、酶扩散和多循环测序,而现有的水凝胶-组织化学方法则不能。在一些实施方案中,为了使扩增产物能够包埋在组织-水凝胶环境中,将胺修饰的核苷酸包括在扩增步骤(例如,RCA)中、使用丙烯酸N-羟基琥珀酰亚胺酯用丙烯酰胺部分官能化、并与丙烯酰胺单体共聚以形成水凝胶。

[0252] 在一些实施方案中,可以通过顺序荧光杂交执行序列分析(例如,边杂交边测序)。顺序荧光杂交可以涉及包含寡核苷酸和可检测标记的检测探针的顺序杂交。

[0253] 在一些实施方案中,可以通过边合成边测序(SBS)来执行测序。在一些实施方案中,测序引物与一个或多个条形码处或附近的序列互补。在此类实施方案中,边合成边测序可以包括逆转录和/或扩增,以便生成模板序列,引物序列可以从所述模板序列结合。示例性SBS方法包括描述于例如但不限于US 2007/0166705、US 2006/0188901、US 7,057,026、US 2006/0240439、US 2006/0281109、US20110059865、US 2005/0100900、US2008/0280773、US20090118128、US 2012/0270305、US 2013/0260372和US 2013/0079232中的那些。

[0254] 在一些实施方案中,可以使用通过连接进行的单分子测序来执行测序。此类技术利用DNA连接酶来掺入寡核苷酸并鉴定这种寡核苷酸的掺入。寡核苷酸通常具有和与寡核苷酸杂交的序列中的特定核苷酸的身份相关的不同标记。例如,涉及边连接边测序的方面和特征在例如Shendure等人Science(2005),309:1728-1732和US 5,599,675;US 5,750,341;US 6,969,488;US 6,172,218;和US 6,306,597中有所描述。

[0255] 在一些实施方案中,扩增产物的检测包括用荧光团、同位素、质量标签或其组合来标记扩增产物。在一些实施方案中,扩增产物的检测包括将一种或多种探针与扩增产物(例如,可检测标记的探针)直接或间接杂交,任选地其中荧光标记的探针直接与扩增产物杂交,或者任选地其中荧光标记的探针直接与中间探针杂交,所述中间探针与扩增产物杂交。在一些实施方案中,

[0256] 扩增产物在样品中原位检测。

[0257] 在一些实施方案中,条形码被可检测标记的检测寡核苷酸(诸如荧光标记的寡核苷酸)靶向。在一些实施方案中,使用一种或多种解码方案来解码用于序列确定的信号(诸如荧光)。在本文的任何实施方案中,条形码可以使用任何合适的方法或技术进行分析(例如,检测或测序),所述合适的方法或技术包括本文所述的那些,诸如靶标的RNA顺序探测(RNA SPOT)、顺序荧光原位杂交(seqFISH)、单分子荧光原位杂交(smFISH)、多重抗误荧光原位杂交(MERFISH)、基于杂交的原位测序(HybISS)、原位测序、靶向原位测序、荧光原位测序(FISSEQ)或空间分辨的转录扩增子读出映射(STARmap)。在一些实施方案中,本文提供的方法包括通过顺序杂交和用多个标记的探针检测来分析条形码。示例性解码方案在例如Eng等人,“Transcriptome-scale Super-Resolved Imaging in Tissues by RNA SeqFISH+,”*Nature* 568(7751):235-239(2019);Chen等人,“Spatially resolved,highly multiplexed RNA profiling in single cells,”*Science*;348(6233):aaa6090(2015);Gyllborg等人,*Nucleic Acids Res*(2020)48(19):e112;US 10,457,980B2;US 2016/0369329 A1;WO 2018/026873 A1;和US 2017/0220733 A1中有所描述,所述文献和专利全部以引用方式整体并入本文。在一些实施方案中,这些测定能够同时实现信号放大、组合解码和纠错方案。

[0258] 在一些实施方案中,核酸杂交可以用于测序。这些方法利用与条形码序列的至少一部分互补的标记的核酸解码器探针。可以使用具有可区分标记的许多不同探针的集合来执行多重解码。核酸杂交测序的非限制性实例在例如US 8,460,865和Gunderson等人,*Genome Research* 14:870-877(2004)中有所描述,所述专利和文献各自的全部内容以引用方式并入本文。在任何前述实施方案中,本文提供的方法可以包括通过顺序杂交和用多个标记的探针(例如,检测寡核苷酸)检测来分析条形码。

[0259] 在一些实施方案中,可以在测序期间使用DNA聚合酶活性的实时监测。例如,可以通过荧光共振能量转移(FRET)检测核苷酸掺入,如例如Levene等人,*Science*(2003),299,682-686、Lundquist等人,*Opt.Lett.*(2008),33,1026-1028和term“perfectly”等人,*Proc.Natl.Acad.Sci.USA*(2008),105,1176-1181中所述。

[0260] 在一些方面,可以在室温下进行分析和/或序列确定,以在低背景噪声和错误减少的情况下最好地保留组织形态。在一些实施方案中,分析和/或序列确定包括在测序进行时消除错误累积。

[0261] 在一些实施方案中,分析和/或序列确定涉及洗涤以去除未结合的多核苷酸,其后揭示荧光产物用于成像。

[0262] 在一些方面,检测(包括成像)使用多种不同类型的显微镜检查中的任一种进行,例如共聚焦显微镜检查、双光子显微镜检查、明场显微镜检查、完整组织膨胀显微镜检查和/或CLARITY<sup>TM</sup>优化的光片显微镜检查(COLM)。

[0263] 在一些实施方案中,荧光显微镜检查用于检测探针的检测和成像。在一些方面,荧光显微镜是使用荧光和磷光代替反射和吸收或者除反射和吸收之外使用荧光和磷光来研究有机或无机物质的性质的光学显微镜。在荧光显微镜检查中,用激发样品中的荧光的波长的光照射样品。然后通过显微镜物镜对通常波长比照射更长的荧光进行成像。在这种技术中可以使用两种滤光片;确保照射接近单色且处于正确的波长的照明(或激发)滤光片、

以及确保没有激发光源到达检测器的第二发射(或抑止)滤光片。或者,这些功能均可以通过单个二向色滤光片来实现。“荧光显微镜”包括使用荧光生成图像的任何显微镜,无论是更简单的设置(如落射荧光显微镜)还是更复杂的设计(诸如共聚焦显微镜),其使用光学切片来获得更好的荧光图像分辨率。

[0264] 在一些实施方案中,共聚焦显微镜检查用于检测探针的检测和成像。共聚焦显微镜使用在探测器前面的光学共轭平面中的点照射和针孔来消除失焦信号。由于仅可以检测到非常接近焦平面的荧光产生的光,因此图像的分辨率(特别是在样品深度方向上)比宽场显微镜要好得多。然而,由于来自样品荧光的大部分光在针孔处被阻挡,这种增加的分辨率是以信号强度降低为代价的——因此通常需要长时间曝光。由于一次仅照射样品中的一个点,因此2D或3D成像需要在样本中的规则光栅(即,平行扫描线的矩形图案)上进行扫描。焦平面的可实现厚度主要由所用光的波长除以物镜的数值孔径限定,但也由样本的光学性质限定。可能的薄光学切片使这些类型的显微镜特别擅长样品的3D成像和表面谱分析。CLARITY™优化的光片显微镜(COLM)为大型澄清样品的快速3D成像提供了替代显微镜检查。COLM询问大型免疫染色组织、允许采集速度提高、并产生更高质量的生成数据。

[0265] 可以采用的其他类型的显微镜检查包括明场显微镜检查、倾斜照明显微镜检查、暗场显微镜检查、相差、微分干涉差(DIC)显微镜检查、干涉反射显微镜检查(也称为反射干涉差或RIC)、单平面照明显微镜检查(SPIM)、超分辨率显微镜检查、激光显微镜检查、电子显微镜检查(EM)、透射电子显微镜检查(TEM)、扫描电子显微镜检查(SEM)、反射电子显微镜检查(REM)、扫描透射电子显微镜检查(STEM)和低电压电子显微镜检查(LVEM)、扫描探针显微镜检查(SPM)、原子力显微镜检查(ATM)、弹道电子发射显微镜检查(BEEM)、化学力显微镜检查(CFM)、导电原子力显微镜检查(C-AFM)、电化学扫描隧道显微镜(ECSTM)、静电力显微镜检查(EFM)、流体力显微镜(FluidFM)、力调制显微镜检查(FMM)、面向特征的扫描探针显微镜检查(FOSPM)、开尔文探针力显微镜检查(KPFM)、磁力显微镜检查(MFM)、磁共振力显微镜检查(MRFM)、近场扫描光学显微镜检查(NSOM)(或SNOM,扫描近场光学显微镜检查SNOM)、压电响应力显微镜检查(PFM)、PSTM,光子扫描隧道显微镜检查(PSTM)、PTMS,光热显微光谱/显微镜检查(PTMS)、SCM,扫描电容显微镜检查(SCM)、SECM,扫描电化学显微镜检查(SECM)、SGM,扫描门显微镜检查(SGM)、SHPM,扫描霍尔探针显微镜检查(SHPM)、SICM,扫描离子电导显微镜检查(SICM)、SPSM,自旋极化扫描隧道显微镜检查(SPSM)、SSRM,扫描扩散电阻显微镜检查(SSRM)、SthM,扫描热显微镜检查(SthM)、STM,扫描隧道显微镜检查(STM)、STP,扫描隧道电位法(STP)、SVM,扫描电压显微镜检查(SVM)、以及同步加速器X射线扫描隧道显微镜检查(SXSTM)和完整组织膨胀显微镜检查(exM)。

[0266] 本文所述的探针可用于分析与靶核酸或靶分析物相关的延伸产物(例如,RCA产物)的序列。在一些实施方案中,待分析的序列是核酸条形码序列。例如,核酸条形码序列可以对应于生物样品中的分析物或其部分(例如,核酸序列)或对应于用于分析物或其部分的标记剂。

[0267] 在一些实施方案中,检测还包括对包含探针和靶核酸或靶分析物(例如,或其产物或衍生物)的复合物进行成像,例如通过光学成像。例如,一种或多种探针可以是包含一种或多种核酸条形码序列的条形码探针,所述核酸条形码序列可以直接或间接地被可检测标记的检测探针(例如,荧光标记的检测探针)结合。可以分析可检测信号或一系列信号(诸如

包含空间模式和/或时间模式的荧光)以揭示样品中的一种或多种分析物的存在/不存在、分布、位置、量、水平、表达或活性。在一些实施方案中,在组织样品中原位分析(例如,通过成像)一种或多种分析物(例如,靶序列),而不需要迁移出组织样品的细胞。在一些实施方案中,在组织样品中原位分析(例如,通过成像)一种或多种蛋白质分析物,不需要迁移出组织样品(例如,迁移到基底上)。在一些实施方案中,探针包含分析物结合部分(例如,抗体),并且核酸条形码序列在原位分析期间不被切割。

[0268] 在一些情况下,对捕获的一幅或多幅图像执行分析,并且分析可以包括处理图像和/或量化观察到的信号。例如,分析可以包括处理一种或多种细胞类型、一种或多种类型的生物标志物、生物标志物的数量或水平和/或在样品的特定区域中检测到的细胞的数量或水平的信息。在一些实施方案中,分析包括检测序列,例如存在于样品中的条形码。在一些实施方案中,分析包括斑点的定量(例如,如果检测到扩增产物)。在一些情况下,分析包括确定是否存在与来自特定组的一种或多种生物标志物相关的特定细胞和/或信号。在一些实施方案中,可以将获得的信息与阳性和阴性对照或与特征的阈值进行比较,以确定样品是否表现出某种特征或表型。在一些情况下,信息可以包括来自细胞、区域的信号,且/或包括来自多种可检测标记的读数。在一些情况下,分析还包括展示来自分析或检测步骤的信息。在一些实施方案中,可以使用软件来自动化数据的处理、分析和/或展示。

#### [0269] III. 组合物和试剂盒

[0270] 在一些实施方案中,本文提供了一种组合物,其包含如第II部分所述的多种探针,诸如:一种或多种第一探针和第二探针(例如,染色质接近探针、检测探针),其中第一探针在第一染色质区域中与第一核酸链杂交,并且第二探针在参与染色质相互作用事件的第二染色质区域中与第二核酸链杂交;和一种或多种桥接探针,其中第一和/或第二探针的末端或一种或多种桥接探针的末端形成环状探针,其中环状探针包含第一核酸链或其互补链的序列和/或第二核酸链或其互补链的序列。

[0271] 本文还提供了试剂盒,例如包括如第II部分所述的探针或探针文库(任选地包含可检测部分,例如信号产生标签)。

[0272] 试剂盒的各种组分可以存在于单独的容器中,或者某些相容的组分可以预先组合到单个容器中。在一些实施方案中,试剂盒还含有用于使用试剂盒的组分来实践所提供的方法的说明书。

[0273] 在一些实施方案中,试剂盒可以含有执行所提供的方法的一个或多个步骤所需的试剂和/或消耗品。在一些实施方案中,试剂盒含有用于固定、包埋和/或透化生物样品的试剂。在一些实施方案中,试剂盒含有试剂,诸如用于连接和/或扩增的酶和缓冲剂,诸如连接酶和/或聚合酶。在一些方面,试剂盒还可以包括本文所述试剂中的任一种,例如洗涤缓冲液和连接缓冲液。在一些实施方案中,试剂盒含有用于检测和/或测序的试剂,诸如条形码检测探针或可检测标记。在一些实施方案中,试剂盒任选地含有其他组分,例如用于另外的测定的核酸引物、酶和试剂、缓冲液、核苷酸、修饰的核苷酸、试剂。

#### [0274] IV. 术语

[0275] 在整个本公开中使用特定术语来解释所述设备、系统、方法和组合物的各个方面。

[0276] 已经描述了本公开的一些例示性实施方案,本领域的技术人员应当清楚,上文仅仅是例示性并且不是限制性的,仅以实例的方式呈现。许多修改和其他例示性实施方案在

本领域的普通技术人员的范围内并且被考虑落入本公开的范围。特别地,尽管本文呈现的许多实例涉及方法动作或系统元素的特定组合,但应当理解,那些动作和那些元素可以以其他方式组合以实现相同的目标。

[0277] 一如本文所用,术语“\_\_\_\_\_”“一般是指……”或类似的句子,不旨在以明示或暗示的方式将该术语的含义限制在其简单或普通的含义之外,并且基于在本专利的任何部分中所作的任何陈述(权利要求的语言除外),这种术语不应被解释为在范围上受到限制。在某种程度上,在本专利末尾的权利要求中引用的任何术语在本专利中以与单一含义一致的方式提及,这样做只是为了清楚起见,以免混淆读者,并且不旨在通过暗示或其他方式将这种权利要求术语限制为该单一含义。最后,除非权利要求元素是通过引用词语“一意指”和不引用任何结构的功能来定义的,否则不旨在基于35U.S.C. §112第6段的应用来解释任何权利要求元素的范围。

[0278] 除非上下文另外明确指出,否则如本文所用,单数形式“一个(种)”和“一所述”包括多个指示物。例如,“一个(种)”表示“至少一个(种)”或“一个或多个(一种或多种)”。还应当理解,除非在本专利中使用句子来明确定义术语

[0279] 如本文所用的术语“一约”或“一大约”是指本技术领域的技术人员容易知道的相应值的通常误差范围。本文中对“一约”一个值或参数的提及包括(并且描述)针对所述值或参数本身的实施方案。例如,根据相关领域中的实践,“一约”可以意指在1个或超过1个标准差以内。或者,“一约”可以意指给定值的至多20%、至多10%、至多5%或至多1%的范围。

[0280] 在整个本公开中,要求保护的主题的各个方面以范围格式呈现。应当理解的是,范围格式的描述仅仅是为了方便以及简洁,并且不应当解释为对要求保护的主体范围的不灵活的限制。因此,范围的描述应当被认为是具有明确公开的所有可能的子范围以及该范围内的单独数值。例如,在提供数值范围的情况下,应当理解,在该范围的上限和下限之间的每个中间值以及在该规定范围内的任何其他规定值或中间值都涵盖在要求保护的主体中。这些较小范围的上下限可独立地包括在较小范围中,并且也涵盖在要求保护的主体中,受规定范围内任何具体排除的限制的限制。在范围包括限值中的一个或两个的情况下,排除那些包括的限值中的一个或两个的范围也包括在要求保护的主体中。无论范围的宽度如何,这都适用。

[0281] 在权利要求中使用顺序术语(诸如“第一”、“第二”、“第三”等)来修饰权利要求元素本身并不意味着一个权利要求元素超过另一个的任何优先、居先或次序、或执行方法动作的时序,而是仅用作区分具有某一名称的一个权利要求元素与(如果不使用顺序术语,那么将)具有相同名称的另一元素以区分所述权利要求元素的标记。类似地,使用a)、b)等或i)、ii)等本身并不意味着权利要求中的步骤的任何优先、居先或次序。类似地,在说明书中使用这些术语本身并不意味着任何需要的优先、居先或次序。

[0282] (i) 条形码

[0283] “一条形码”是传达或能够传达信息(例如,关于样品、珠粒和/或捕获探针中的分析物的信息)的标记或标识符。条形码可以是分析物的一部分,或独立于分析物。条形码可以与分析物连接。特定条形码相对于其他条形码可以是唯一的。

[0284] 条形码可以有多种不同的格式。例如,条形码可以包括多核苷酸条形码、随机核酸和/或氨基酸序列、以及合成核酸和/或氨基酸序列。条形码可以以可逆或不可逆的方式与

分析物或另一个部分或结构连接。可以在样品测序之前或期间将条形码添加到例如脱氧核糖核酸 (DNA) 或核糖核酸 (RNA) 样品的片段中。条形码可以允许单个测序读段的鉴定和/或定量 (例如, 条形码可以是或可以包括唯一分子标识符或—UMI”)。

[0285] 条形码可以例如以单细胞分辨率空间分辨存在于生物样品中的分子组分 (例如, 条形码可以是或可以包括—空间条形码”)。在一些实施方案中, 条形码包括 UMI 和空间条形码两者。在一些实施方案中, 条形码包括一起充当单个条形码的两个或更多个子条形码。例如, 多核苷酸条形码可以包括由一个或多个非条形码序列隔开的两个或多个多核苷酸序列 (例如, 子条形码)。

[0286] (ii) 核酸和核苷酸

[0287] 术语—核酸”和—核苷酸”旨在与其在本领域中的用途一致并且包括天然存在的种类或其功能类似物。核酸的特别有用的功能类似物能够以序列特异性方式与核酸杂交 (例如, 能够与两种核酸杂交, 使得可以在两种杂交的核酸之间发生连接) 或能够用作复制特定核苷酸序列的模板。天然存在的核酸通常具有含有磷酸二酯键的骨架。类似物结构可以具有替代的骨架键联, 包括多种本领域已知的那些中的任一种。天然存在的核酸通常具有脱氧核糖 (例如, 存在于脱氧核糖核酸 (DNA) 中) 或核糖 (例如, 存在于核糖核酸 (RNA) 中)。

[0288] 核酸可以含有具有本领域已知的这些糖部分的多种类似物中的任一种的核苷酸。核酸可以包括天然或非天然核苷酸。在这方面, 天然脱氧核糖核酸可以具有一种或多种选自腺嘌呤 (A)、胸腺嘧啶 (T)、胞嘧啶 (C) 或鸟嘌呤 (G) 组成的组的碱基, 并且核糖核酸可以具有一种或多种选自尿嘧啶 (U)、腺嘌呤 (A)、胞嘧啶 (C) 或鸟嘌呤 (G) 组成的组的碱基。可以包括在核酸或核苷酸中的可用的非天然碱基是本领域已知的。

[0289] 可以用于修饰其结构上任何位置的核苷酸的修饰的碱基部分的实例包括但不限于: 5-氟尿嘧啶、5-溴尿嘧啶、5-氯尿嘧啶、5-碘尿嘧啶、次黄嘌呤、黄嘌呤、乙酰胞嘧啶、5-(羧基羟甲基) 尿嘧啶、5-羧基甲基氨基甲基-2-硫代尿苷、5-羧基甲基氨基甲基尿嘧啶、二氢尿嘧啶、 $\beta$ -D-半乳糖基嘧啶 ( $\beta$ -D-galactosylqueosine)、肌苷、N<sup>6</sup>-异戊烯基腺嘌呤、1-甲基鸟嘌呤、1-甲基肌苷、2,2-二甲基鸟嘌呤、2-甲基腺嘌呤、2-甲基鸟嘌呤、3-甲基胞嘧啶、5-甲基胞嘧啶、N<sup>6</sup>-腺嘌呤、7-甲基鸟嘌呤、5-甲基氨基甲基尿嘧啶、甲氧基氨基甲基-2-硫脲嘧啶、 $\beta$ -D-甘露糖基嘧啶 ( $\beta$ -D-mannosylqueosine)、5'-甲氧基羧基甲基尿嘧啶、5-甲氧基尿嘧啶、2-甲硫基-N<sup>6</sup>-异戊烯基腺嘌呤、尿嘧啶-5-氧乙酸、假尿嘧啶、嘧啶 (queosine)、2-硫胞嘧啶、5-甲基-2-硫尿嘧啶、2-硫尿嘧啶、4-硫尿嘧啶、5-甲基尿嘧啶、尿嘧啶-5-氧乙酸甲酯、尿嘧啶-S-氧乙酸、5-甲基-2-硫尿嘧啶、3-(3-氨基-3-N-2-羧基丙基) 尿嘧啶、2,6-二氨基嘌呤和生物素化类似物等。

[0290] 可以用于修饰其结构上任何位置的核苷酸的修饰的糖部分的实例包括但不限于阿拉伯糖、2-氟阿拉伯糖、木糖和己糖, 或磷酸骨架的修饰组分, 诸如硫代磷酸酯、二硫代磷酸酯、硫代磷酸酰胺酯、氨基磷酸酯、二氨基磷酸酯、甲基磷酸酯、烷基磷酸三酯或甲缩醛 (formacetal) 或其类似物。

[0291] (iii) 探针和靶标

[0292] —探针”或—靶标”在用于提及核酸或核酸序列时, 旨在作为方法或组合物上下文中的核酸或序列的语义标识符, 并且不限制核酸或序列的结构或功能超出明确指出的范

围。

[0293] (iv) 寡核苷酸和多核苷酸

[0294] 术语“寡核苷酸”和“多核苷酸”可互换使用,以指代长度为约2个至约500个核苷酸的单链核苷酸多聚体。寡核苷酸可以是合成的,通过酶促(例如,经由聚合)或使用“一分池”方法来制备。寡核苷酸可以包括核糖核苷酸单体(即,可以是寡核糖核苷酸)和/或脱氧核糖核苷酸单体(即,寡脱氧核糖核苷酸)。在一些实例中,寡核苷酸可以包括寡核苷酸中的脱氧核糖核苷酸单体和核糖核苷酸单体的组合(例如,脱氧核糖核苷酸单体和核糖核苷酸单体的随机或有序组合)。寡核苷酸的长度可以是例如4个至10个、10个至20个、21个至30个、31个至40个、41个至50个、51个至60个、61个至70个、71个至80个、80个至100个、100个至150个、150个至200个、200个至250个、250个至300个、300个至350个、350个至400个、或400个至500个核苷酸。寡核苷酸可以包括一个或多个与多聚体结构连接(例如,共价或非共价)的功能部分。例如,寡核苷酸可以包括一种或多种可检测标记(例如,放射性同位素或荧光团)。

[0295] (v) 杂交(Hybridizing/Hybridize)和退火(Annealing/Anneal)

[0296] 术语“杂交(hybridizing/hybridize)”和“退火(annealing/anneal)”在本公开中可互换使用,并且指代两个不同分子内的基本上互补或互补的核酸序列的配对。配对可以通过任何过程实现,其中核酸序列通过碱基配对与基本上或完全互补的序列连接以形成杂交复合物。出于杂交的目的,如果两个核酸序列的至少60%(例如,至少70%、至少80%或至少90%)的单个碱基彼此互补,则两个核酸序列“基本上互补”。

[0297] (vi) 引物

[0298] “引物”是具有3'端的单链核酸序列,其可以用作核酸延伸反应中的核酸聚合酶的底物。RNA引物由RNA核苷酸形成,并且用于RNA合成,而DNA引物由DNA核苷酸形成,并且用于DNA合成。引物还可以包括RNA核苷酸和DNA核苷酸两者(例如,以随机或设计的模式)。引物还可以包括本文所述的可以具有另外功能的其他天然或合成核苷酸。在一些实例中,DNA引物可以用于引发RNA合成,并且反之亦然(例如,RNA引物可以用于引发DNA合成)。引物的长度可以不同。例如,引物可以是约6个碱基至约120个碱基。例如,引物可以包括至多约25个碱基。在一些情况下,引物可以指引物结合序列。

[0299] (vii) 引物延伸

[0300] “引物延伸”是指其中两个核酸序列通过它们各自的末端互补核酸序列的重叠而连接(例如,杂交)的任何方法。这种连接之后可以是使用其他核酸序列作为延伸模板的一个或两个末端的核酸延伸(例如,酶促延伸)。酶促延伸可以通过酶进行,所述酶包括但不限于聚合酶和/或逆转录酶。

[0301] (viii) 连接

[0302] 在一些实施方案中,本文提供了能够进行DNA模板连接(诸如从cDNA分子)的探针或探针组。参见例如美国专利8,551,710,其以引用方式在此整体并入。在一些实施方案中,本文提供了能够进行RNA模板连接的探针或探针组。参见例如美国专利公开2020/0224244,其以引用方式在此整体并入。在一些实施方案中,探针组是SNAIL探针组。参见例如美国专利公开20190055594,其以引用方式在此整体并入。

[0303] 在一些实施方案中,本文的连接是例如通过酶促方式(例如,连接酶)连接两个(或



更多)彼此接近的核酸序列的接近连接。在一些实施方案中,接近连接可以包括“间隙填充”步骤,其涉及基于模板核酸分子的核酸序列,跨越两个感兴趣的核酸分子之间的距离,通过聚合酶掺入一种或多种核酸(参见例如美国专利号7,264,929,其以引用方式整体并入本文)。多种不同的方法可以用于接近连接核酸分子,包括(但不限于)“粘性末端”和“平末端”连接。此外,可以使用单链连接对单链核酸分子执行接近连接。粘性末端接近连接涉及连接事件本身之前两个待连接的核酸分子之间的互补单链序列的杂交。平末端接近连接一般不包括来自每个核酸分子的互补区域的杂交,因为两个核酸分子在连接位点都缺乏单链突出端。

[0304] 在一些实施方案中,本文提供的是多重接近连接测定。参见例如美国专利公开20140194311,其以引用方式在此整体并入。在一些实施方案中,本文提供了能够进行接近连接的探针或探针组。在一些实施方案中,环状探针可以与靶核酸间接杂交。在一些实施方案中,由能够进行接近连接的探针组(例如,接近连接原位杂交(PLISH)探针组)形成环状构建体。参见例如美国专利公开2020/0224243,其以引用方式在此整体并入。在一些实施方案中,探针组可以包含两种或更多种探针寡核苷酸,每种寡核苷酸包含彼此互补的区域。对于使用接近的寡核苷酸来检测分析物,参见例如Söderberg等人,Methods.(2008),45(3):227-32和美国专利申请公开号2002/0051986,所述文献和专利的全部内容以引用方式并入本文。

[0305] 在一些实施方案中,连接涉及化学连接。在一些实施方案中,连接涉及模板依赖性连接。在一些实施方案中,连接涉及不依赖模板的连接。在一些实施方案中,连接涉及酶促连接。

[0306] 在一些实施方案中,酶促连接涉及使用连接酶。在一些方面,本文使用的连接酶包含通常用于将多核苷酸连接在一起或用于连接单个多核苷酸的末端的酶。可以使用RNA连接酶、DNA连接酶或另一种连接酶将两个核苷酸序列连接在一起。连接酶包括ATP依赖性双链多核苷酸连接酶、NAD<sup>-i</sup>依赖性双链DNA或RNA连接酶和单链多核苷酸连接酶,例如描述于EC 6.5.1.1(ATP依赖性连接酶)、EC 6.5.1.2(NAD<sup>+</sup>依赖性连接酶)、EC 6.5.1.3(RNA连接酶)中的连接酶中的任一种。连接酶的具体实例包括细菌连接酶,诸如大肠杆菌(*E. coli*) DNA连接酶、Tth DNA连接酶、嗜热球菌属种(*Thermococcus* sp.) (菌株9<sup>°N</sup>) DNA连接酶(9<sup>°N</sup> DNA连接酶,New England Biolabs)、Taq DNA连接酶、Ampligase<sup>™</sup>(Epicentre Biotechnologies),以及噬菌体连接酶,诸如T3 DNA连接酶、T4 DNA连接酶和T7 DNA连接酶及其突变。在一些实施方案中,连接酶是T4 RNA连接酶。在一些实施方案中,连接酶是夹板R连接酶(splintR ligase)。在一些实施方案中,连接酶是单链DNA连接酶。在一些实施方案中,连接酶是T4 DNA连接酶。在一些实施方案中,连接酶是具有DNA-夹板DNA连接酶活性的连接酶。在一些实施方案中,连接酶是具有RNA-夹板DNA连接酶活性的连接酶。

[0307] 在一些实施方案中,本文的连接是直接连接。在一些实施方案中,本文的连接是间接连接。“直接连接”意指多核苷酸的末端彼此紧邻杂交,以形成导致它们彼此连接(分子内连接)的连接酶的底物。或者,“间接”意指多核苷酸的末端彼此不相邻地杂交,即被一个或多个中间的核苷酸或“间隙”隔开。在一些实施方案中,所述末端不直接彼此连接,而是经由一种或多种中间核苷酸(所谓“间隙”或“间隙填充”(寡)核苷酸)的中间性或通过延伸探针的3'端以“填充”对应于所述中间核苷酸的“间隙”(分子间连接)来发生。在一些

情况下,多核苷酸的杂交末端之间的一种或多种核苷酸的间隙可以被一种或多种与夹板、锁式探针或靶核酸互补的一间隙”(寡)核苷酸一填充”。间隙可以是1个至60个核苷酸的间隙或1个至40个核苷酸的间隙或3个至40个核苷酸的间隙。在具体的实施方案中,间隙可以是约1个、2个、3个、4个、5个、6个、7个、8个、9个或10个或更多个核苷酸的间隙、所示值之间的任何整数(或整数范围)的核苷酸的间隙。在一些实施方案中,所述末端区域之间的间隙可以被间隙寡核苷酸填充或通过延伸多核苷酸的3'端来填充。在一些情况下,连接涉及将探针的末端与至少一个间隙(寡)核苷酸连接,使得间隙(寡)核苷酸掺入到所得多核苷酸中。在一些实施方案中,本文的连接之前是间隙填充。在其他实施方案中,本文的连接不需要间隙填充。

[0308] 在一些实施方案中,多核苷酸连接产生解链温度高于未连接的多核苷酸的解链温度的多核苷酸。因此,在一些方面,在包括扩增和检测的后续步骤之前,连接使含有连接的多核苷酸的杂交复合物稳定。

[0309] 在一些方面,使用高保真连接酶,诸如热稳定的DNA连接酶(例如,Taq DNA连接酶)。热稳定的DNA连接酶在升高的温度下具有活性,从而允许通过在接近DNA链的解链温度( $T_m$ )的温度下孵育连接来进一步鉴别。与退火的完全碱基配对的底物相比,这选择性地降低了退火的错配底物的浓度(预期具有错配周围略低的 $T_m$ )。因此,可以通过组合连接酶活性位点的内在选择性和平衡条件来实现高保真连接,以降低退火的错配dsDNA的发生率。

[0310] (ix) 核酸延伸

[0311] “核酸延伸”一般涉及将一种或多种核酸(例如A、G、C、T、U、其核苷酸类似物或衍生物)以模板依赖性方式掺入到分子中(诸如但不限于核酸序列),使得通过酶(诸如聚合酶或逆转录酶)掺入连续的核酸,从而产生新合成的核酸分子。例如,与互补核酸序列杂交的引物可以用于通过将互补核酸序列用作核酸合成的模板来合成新的核酸分子。类似地,与poly(dT)序列(例如,捕获结构域)杂交的mRNA转录物的3'聚腺苷酸化尾可以用作对应cDNA分子的单链合成的模板。

[0312] (x) PCR扩增

[0313] “PCR扩增”是指使用聚合酶链式反应(PCR)生成遗传物质包括DNA和RNA序列的拷贝。用于实施PCR的合适试剂和条件在例如美国专利4,683,202、4,683,195、4,800,159、4,965,188和5,512,462中有所描述,所述专利各自的全部内容以引用方式并入本文。在典型的PCR扩增中,反应混合物包括待扩增的遗传物质、酶、用于引物延伸反应的一种或多种引物、和用于反应的试剂。寡核苷酸引物的长度足以在退火条件下提供与互补遗传物质的杂交。引物的长度一般取决于扩增结构域的长度,但通常将为至少4个碱基、至少5个碱基、至少6个碱基、至少8个碱基、至少9个碱基、至少10个碱基对(bp)、至少11bp、至少12bp、至少13bp、至少14bp、至少15bp、至少16bp、至少17bp、至少18bp、至少19bp、至少20bp、至少25bp、至少30bp、至少35bp,并且可以长达40bp或更长,其中引物长度一般在18bp至50bp的范围内。遗传物质可以与单个引物或两个引物(正向和反向引物)的组接触,这取决于是否期望遗传物质的引物延伸、线性或指数扩增。

[0314] 在一些实施方案中,PCR扩增过程使用DNA聚合酶。DNA聚合酶活性可以由一种或多种不同的DNA聚合酶提供。在某些实施方案中,DNA聚合酶来自细菌,例如,DNA聚合酶是细菌DNA聚合酶。例如,DNA聚合酶可以来自埃希氏菌属(*Escherichia*)、芽孢杆菌属(*Bacillus*)、

嗜热菌属(Thermophilus)或火球菌(Pyrococcus)属的细菌。

[0315] 可以使用的DNA聚合酶的合适实例包括但不限于:大肠杆菌DNA聚合酶I、Bsu DNA聚合酶、Bst DNA聚合酶、Taq DNA聚合酶、VENT<sup>TM</sup> DNA聚合酶、DEEPVENT<sup>TM</sup> DNA聚合酶、LongAmp<sup>®</sup> Taq DNA聚合酶、LongAmp<sup>®</sup> Hot Start Taq DNA聚合酶、Crimson LongAmp<sup>®</sup> Taq DNA聚合酶、Crimson Taq DNA聚合酶、OneTaq<sup>®</sup> DNA聚合酶、OneTaq<sup>®</sup> Quick-Load<sup>®</sup> DNA聚合酶、Hemo KlenTaq<sup>®</sup> DNA聚合酶、REDTaq<sup>®</sup> DNA聚合酶、Phusion<sup>®</sup> DNA聚合酶、Phusion<sup>®</sup>高保真DNA聚合酶、Platinum Pfx DNA聚合酶、AccuPrime Pfx DNA聚合酶、Phi29 DNA聚合酶、Klenow片段、Pwo DNA聚合酶、Pfu DNA聚合酶、T4 DNA聚合酶和T7 DNA聚合酶。

[0316] 术语“DNA聚合酶”不仅包括天然存在的酶,还包括其所有修饰的衍生物,还包括天然存在的DNA聚合酶的衍生物。例如,在一些实施方案中,可以修饰DNA聚合酶以去除5'-3'核酸外切酶活性。可以使用的DNA聚合酶的序列修饰衍生物或突变体包括但不限于保留至少一些功能(例如,野生型序列的DNA聚合酶活性)的突变体。在不同反应条件(例如,温度、模板浓度、引物浓度等)下,突变可以影响酶的活性谱,例如提高或降低聚合速率。突变或序列修饰还可以影响酶的核酸外切酶活性和/或热稳定性。

[0317] 在一些实施方案中,PCR扩增可以包括反应,诸如但不限于链置换扩增反应、滚环扩增反应、连接酶链反应、转录介导的扩增反应、恒温扩增反应和/或环介导的扩增反应。

[0318] 在一些实施方案中,PCR扩增使用与靶DNA片段的3'标签互补的单一引物。在一些实施方案中,PCR扩增使用第一和第二引物,其中第一引物的至少3'端部分与靶核酸片段的3'标签的至少一部分互补,并且其中第二引物的至少3'端部分表现出靶核酸片段的5'标签的至少一部分的序列。在一些实施方案中,第一引物的5'端部分与靶核酸片段的3'标签不互补,并且第二引物的5'端部分不表现出靶核酸片段的5'标签的至少一部分的序列。在一些实施方案中,第一引物包括第一通用序列且/或第二引物包括第二通用序列。

[0319] 在一些实施方案中(例如,当PCR扩增捕获的DNA时),可以使用DNA连接酶将PCR扩增产物与另外的序列连接。DNA连接酶活性可以由一种或多种不同的DNA连接酶提供。在一些实施方案中,DNA连接酶来自细菌,例如,DNA连接酶是细菌DNA连接酶。在一些实施方案中,DNA连接酶来自病毒(例如,噬菌体)。例如,DNA连接酶可以是T4 DNA连接酶。适用于连接步骤的其他酶包括但不限于Tth DNA连接酶、Taq DNA连接酶、嗜热菌属种(菌株9oN)DNA连接酶(9oNTM DNA连接酶,可以从New England Biolabs,Ipswich,MA获得)和Ampligase<sup>TM</sup>(可以从Epicentre Biotechnologies, Madison, WI获得)。还可以使用其衍生物(例如,序列修饰的衍生物)和/或突变体。

[0320] 在一些实施方案中,遗传物质通过逆转录聚合酶链式反应(RT-PCR)扩增。期望的逆转录酶活性可以由一种或多种不同的逆转录酶提供,其合适的实例包括但不限于:M-MLV、MuLV、AMV、HIV、ArrayScript<sup>TM</sup>、MultiScribe<sup>TM</sup>、ThermoScript<sup>TM</sup>和SuperScript<sup>®</sup>I、II、III和IV酶。“逆转录酶”不仅包括天然存在的酶,还包括其所有此类修饰的衍生物,还包括天然存在的逆转录酶的衍生物。

[0321] 另外,可以使用M-MLV、MuLV、AMV和HIV逆转录酶的序列修饰的衍生物或突变体来执行逆转录,所述序列修饰的衍生物或突变体包括保留至少一些功能(例如,野生型序列的

逆转录酶活性)的突变体。逆转录酶可以作为组合物的一部分提供,所述组合物包括其他组分,例如,增强或改善逆转录酶活性的稳定组分,诸如RNA酶抑制剂、DNA依赖性DNA合成的抑制剂,例如放线菌素D(actinomycin D)。逆转录酶的许多序列修饰的衍生物或突变体(例如,M-MLV)以及包括未修饰和修饰的酶的组合物是可商购获得的,例如ArrayScript™、MultiScribe™、ThermoScript™和SuperScript® I、II、III和IV酶。

[0322] 某些逆转录酶(例如,禽成髓细胞瘤病毒(Avian Myeloblastosis Virus,AMV)逆转录酶和莫洛尼鼠白血病病毒(Moloney Murine Leukemia Virus,M-MuLV,MMLV)逆转录酶)可以使用RNA(cDNA合成)和单链DNA(ssDNA)两者作为模板来合成互补DNA链。因此,在一些实施方案中,逆转录反应可以使用能够将RNA和ssDNA两者用作延伸反应的模板的酶(逆转录酶),例如AMV或MMLV逆转录酶。

[0323] 在一些实施方案中,使用本领域众所周知的技术(诸如但不限于—TAQMAN™”或“SYBR®”)或在毛细管(—LightCycler®毛细管”)上通过实时PCR进行RNA和/或DNA的定量(也称为定量PCR或qPCR)。在一些实施方案中,遗传物质的定量通过光学吸光度和实时PCR确定。在一些实施方案中,遗传物质的定量通过数字PCR确定。在一些实施方案中,可以将分析的基因与对应于表达(mRNA)和数量(DNA)的参考核酸提取物(DNA和RNA)进行比较,以便比较靶核酸的表达水平。

[0324] (xi) 标记、可检测标记和光学标记

[0325] 术语“可检测标记”、“光学标记”和“标记”在本文中可互换地用于指与待检测的分子(例如,用于原位测定的探针、捕获探针或分析物)缔合(例如,缀合)的直接或间接可检测部分。可检测标记本身可以是直接可检测的(例如,放射性同位素标记或荧光标记),或者在酶标记的情况下,可以是间接可检测的,例如通过催化直接可检测的底物化合物或组合物的化学变化。可检测标记可以适用于小规模检测和/或适用于高通量筛选。因此,合适的可检测标记包括但不限于放射性同位素、荧光团、化学发光化合物、生物发光化合物和染料。

[0326] 可以定性地检测(例如,通过光学或光谱)可检测标记,或可以将其量化。定性检测一般包括其中确认可检测标记的存在(existence/presence)的检测方法,而可量化检测一般包括具有可量化(例如,可数字报告)值(诸如强度、持续时间、极化和/或其他性质)的检测方法。在一些实施方案中,可检测标记结合特征,或结合与特征缔合的捕获探针。例如,可检测标记的特征可以包括与珠粒连接的荧光、比色或化学发光标记(参见例如Rajeswari等人,J.Microbiol Methods 139:22-28,2017和Forcucci等人,J.Biomed Opt.10:105010,2015,所述文献各自的全部内容以引用方式整体并入本文)。

[0327] 在一些实施方案中,可以将多种可检测标记与待检测的特征、捕获探针或组合物连接。例如,可检测标记可以在核酸聚合或扩增期间掺入(例如,Cy5®标记的核苷酸,诸如Cy5®-dCTP)。可以使用任何合适的可检测标记。在一些实施方案中,可检测标记是荧光团。例如,荧光团可以来自包括以下的组:7-AAD(7-氨基放线菌素D)、吖啶橙(+DNA)、吖啶橙(+RNA)、Alexa Fluor® 350、Alexa Fluor® 430、Alexa Fluor® 488、Alexa Fluor® 532、Alexa Fluor® 546、Alexa Fluor® 555、Alexa Fluor® 568、Alexa Fluor® 594、Alexa Fluor® 633、

Alexa Fluor® 647、Alexa Fluor® 660、Alexa Fluor® 680、Alexa Fluor® 700、Alexa Fluor® 750、别藻蓝蛋白 (Allophycocyanin, APC)、AMCA/AMCA-X、7-氨基放线菌素D (7-AAD)、7-氨基-4-甲基香豆素、6-氨基喹啉、苯胺蓝、ANS、APC-Cy7、ATTO-TAG™ CBQCA、ATTO-TAG™ FQ、Auramine O-Feulgen、BCECF (高pH)、BFP (蓝色荧光蛋白)、BFP/GFP FRET、BOBO™-1/BO-PRO™-1、BOBO™-3/BO-PRO™-3、BODIPY® FL、BODIPY® TMR、BODIPY® TR-X、BODIPY® 530/550、BODIPY® 558/568、BODIPY® 564/570、BODIPY® 581/591、BODIPY® 630/650-X、BODIPY® 650-665-X、BTC、钙黄绿素、钙黄绿素蓝、Calcium Crimson™、Calcium Green-1™、Calcium Orange™、Calcofluor® White、5-羧基荧光素 (5-FAM)、5-羧基萘荧光素、6-羧基罗丹明6G、5-羧基四甲基罗丹明 (5-TAMRA)、羧基-X-罗丹明 (5-ROX)、Cascade Blue®、Cascade Yellow™、CCF2 (GeneBLAzer™)、CFP (青色荧光蛋白)、CFP/YFP FRET、色霉素A3 (Chromomycin A3)、C1-NERF (低pH)、CPM、6-CR 6G、CTC甲腈、Cy2®、Cy3®、Cy3.5®、Cy5®、Cy5.5®、Cy7®、Cychrome (PE-Cy5)、丹磺酰胺 (Dansylamine)、丹磺酰尸胺 (Dansyl cadaverine)、丹磺酰氯 (Dansylchloride)、DAPI、Dapoxyl、DCFH、DHR、DiA (4-Di-16-ASP)、DiD (DiIc18 (5))、DIDS、DiI (DiIc18 (3))、DiO (DiOc18 (3))、DiR (DiIc18 (7))、Di-4 ANEPPS、Di-8 ANEPPS、DM-NERF (4.5-6.5pH)、DsRed (红色荧光蛋白)、EBFP、ECFP、EGFP、ELF® -97乙醇、曙红、赤藓红、溴化乙锭、乙锭同源二聚体-1 (EthD-1)、氯化铊 (III)、5-FAM (5-羧基荧光素)、快蓝、荧光素-dT亚磷酰胺、FITC、Fluo-3、Fluo-4、FluorX®、Fluoro-Gold™ (高pH)、Fluoro-Gold™ (低pH)、Fluoro-Jade、FM® 1-43、Fura-2 (高钙)、Fura-2/BCECF、Fura Red™ (高钙)、Fura Red™/Fluo-3、GeneBLAzer™ (CCF2)、红移GFP (rsGFP)、GFP野生型、GFP/BFP FRET、GFP/DsRed FRET、Hoechst 33342&33258、7-羟基-4-甲基香豆素 (pH 9)、1,5IAEDANS、Indo-1 (高钙)、Indo-1 (低钙)、吖啶二羧花青 (Indodicarbocyanine)、吖啶三羧花青 (Indotricarbocyanine)、JC-1、6-JOE、JOJO™-1/JO-PRO™-1、LDS 751 (+DNA)、LDS 751 (+RNA)、LOLO™-1/LO-PRO™-1、荧光黄、LysoSensor™蓝 (pH 5)、LysoSensor™绿 (pH 5)、LysoSensor™黄/蓝 (pH 4.2)、LysoTracker®绿、LysoTracker®红、LysoTracker®黄、Mag-Fura-2、Mag-Indo-1、Magnesium Green™、Marina Blue®、4-甲基伞形酮、光神霉素 (Mithramycin)、MitoTracker®绿、MitoTracker®橙、MitoTracker®红、NBD (胺)、尼罗红、Oregon Green® 488、Oregon Green® 500、Oregon Green® 514、Pacific Blue、PBF1、PE (R-藻红蛋白)、PE-Cy5、PE-Cy7、PE-德州红 (PE-Texas Red)、PerCP (多甲藻素叶绿素蛋白质 (Peridinin chlorophyll protein))、PerCP-Cy5.5 (TruRed)、PharRed (APC-Cy7)、C-藻青蛋白、R-藻青蛋白、R-藻红蛋白 (PE)、PI (碘化丙啶)、PKH26、PKH67、POPO™-1/PO-PRO™-1、POPO™-3/PO-PRO™-3、碘化丙啶 (PI)、PyMPO、苝、派洛宁Y (Pyronin Y)、Quantam Red (PE-Cy5)、氮芥喹吡因 (Quinacrine Mustard)、R670 (PE-Cy5)、红613 (PE-德州红)、红色荧光蛋白 (DsRed)、试卤灵、RH414、Rhod-2、罗丹明B、罗丹明绿™、罗丹明红™、罗丹明鬼笔环肽、罗丹明110、罗丹明123、5-ROX (羧基-X-罗丹明)、S65A、S65C、S65L、S65T、SBFI、SITS、

SNAFL®-1 (高pH)、SNAFL®-2、SNARF®-1 (高pH)、SNARF®-1 (低pH)、Sodium Green™、SpectrumAqua®、SpectrumGreen® #1、SpectrumGreen® #2、SpectrumOrange®、SpectrumRed®、SYTO® 11、SYTO® 13、SYTO® 17、SYTO® 45、SYTOX® 蓝、SYTOX® 绿、SYTOX® 橙、5-TAMRA (5-羧基四甲基罗丹明)、四甲基罗丹明 (TRITC)、德州红®/德州红®-X、德州红®-X (NHS Ester)、硫杂二羰花青 (Thiadicarbocyanine)、噻唑橙、TOTO®-1/TO-PRO®-1、TOTO®-3/TO-PRO®-3、TO-PRO®-5、Tri-color (PE-Cy5)、TRITC (四甲基罗丹明)、TruRed (PerCP-Cy5.5)、WW 781、X-罗丹明 (XRITC)、Y66F、Y66H、Y66W、YFP (黄色荧光蛋白)、YOYO®-1 / YO-PRO®-1、YOYO®-3/YO-PRO®-3、6-FAM (荧光素)、6-FAM (NHS 酯)、6-FAM (叠氮化物)、HEX、TAMRA (NHS 酯)、Yakima Yellow、MAX、TET、TEX615、ATTO 488、ATTO 532、ATTO 550、ATTO 565、ATTO Rho101、ATTO 590、ATTO 633、ATTO 647N、TYE 563、TYE 665、TYE 705、5 'IRDye® 700、5 'IRDye® 800、5 'IRDye® 800CW (NHS 酯)、WellRED D4染料、WellRED D3染料、WellRED D2染料、Lightcycler® 640 (NHS 酯) 和 Dy 750 (NHS 酯)。

[0328] 如上所述, 在一些实施方案中, 可检测标记是或包括发光或化学发光部分。常见的发光/化学发光部分包括但不限于过氧化物酶, 诸如辣根过氧化物酶 (HRP)、大豆过氧化物酶 (SP)、碱性磷酸酶和萤光素酶。给定适当的底物 (例如, 氧化剂加化学发光化合物), 这些蛋白质部分可以催化化学发光反应。已知许多化合物家族在多种条件下提供化学发光。化学发光化合物家族的非限制性实例包括2,3-二氢-1,4-酞嗪二酮鲁米诺、5-氨基-6,7,8-三甲氧基-和二甲氨基[ca]苯类似物。这些化合物可以在碱性过氧化氢或次氯酸钙和碱的存在下发光。化学发光化合物家族的其他实例包括例如2,4,5-三苯基咪唑、对二甲氨基和-甲氧基取代基、草酸盐 (诸如草酰活性酯)、对硝基苯基、N-烷基吡啶酯、萤光素、光泽精或吡啶酯。在一些实施方案中, 可检测标记是或包括基于金属或基于质量的标记。例如, 小簇金属离子、金属或半导体可以充当质量代码。在一些实例中, 金属可以选自元素周期表的第3-15族, 例如Y、La、Ag、Au、Pt、Ni、Pd、Rh、Ir、Co、Cu、Bi或其组合。

[0329] (xii) 受试者

[0330] 如本文所用, 术语“受试者”一般是指动物, 诸如哺乳动物 (例如, 人) 或禽类 (例如, 鸟), 或其他生物体, 诸如植物。例如, 受试者可以是脊椎动物、哺乳动物、啮齿动物 (例如, 小鼠)、灵长类动物、猿猴或人。动物可以包括但不限于农场动物、体育竞技动物和宠物。受试者可以是健康或无症状的个体、患有或怀疑患有疾病 (例如, 癌症) 或疾病倾向的个体、和/或需要治疗或怀疑需要治疗的个体。受试者可以是患者。受试者可以是微生物 (microorganism/microbe) (例如, 细菌、真菌、古生菌、病毒)。

[0331] V. 示例性实施方案

[0332] 所提供的实施方案包括:

[0333] 实施例1. 一种用于原位空间分析染色质相互作用事件的方法, 所述方法包括:

[0334] a) 获得生物样品;

[0335] b) 将包含至少两种单一寡核苷酸序列的第一组探针与至少两种一级靶核酸序列的有义或反义链杂交,其中所述至少两种一级靶核酸序列参与染色质相互作用事件;

[0336] c) 将包含至少两种单一寡核苷酸序列的第二组探针与所述至少两种一级靶核酸序列的有义或反义链杂交;

[0337] d) 将第三组探针与所述第二组探针杂交,其中所述第二组探针由于与参与染色质相互作用事件的至少两种一级靶核酸序列杂交而在空间上彼此接近,并且其中所述第三组探针桥接所述第二组探针的末端;

[0338] e) 延伸第三组探针以形成环化的核酸,所述环化的核酸包含参与染色质相互作用事件的至少两种一级靶核酸序列的至少一部分;

[0339] f) 扩增所述环状核酸以生成包含可原位检测的信号产生标签的扩增子群;和

[0340] g) 通过所述信号产生标签原位检测所述带标签的扩增子群。

[0341] 实施方案2.如实施方案1所述的方法,其还包括制备包含所述生物样品的载体基底的步骤。

[0342] 实施方案3.如实施方案2所述的方法,其中所述载体基底是其上具有生物样品的载玻片。

[0343] 实施方案4.如实施方案1所述的方法,其中将所述生物样品通过固定剂固定。

[0344] 实施方案5.如实施方案4所述的方法,其中不对所述生物样品进行固定。

[0345] 实施方案6.如实施方案1至5中任一项所述的方法,其还包括使用交联剂将所述生物样品内的染色质DNA与染色质相关因子交联的步骤。

[0346] 实施方案7.如实施方案6所述的方法,其中所述交联剂选自包括以下的组:甲醛、UV辐射、戊二醛、双(亚胺酯)、双(琥珀酰亚胺酯)、二异氰酸酯、二酰氯、1-乙基-3-(3-二甲基氨基丙基)碳二亚胺、丝裂霉素C、氮芥、美法仑、1,3-丁二烯二环氧化物、顺式二胺二氯铂(II)、环磷酰胺、双琥珀酰亚胺戊二酸酯和丙酸二硫代双琥珀酰亚胺酯。

[0347] 实施方案8.如实施方案1至7中任一项所述的方法,其中所述至少两种一级靶核酸序列包含调控元件。

[0348] 实施方案9.如实施方案8所述的方法,其中所述调控元件选自包括启动子、增强子、沉默子或绝缘子的组。

[0349] 实施方案10.如实施方案9所述的方法,其中所述第一组探针之一和所述第二组探针之一与靶启动子DNA序列杂交,并且其中所述第一组探针之一和所述第二组探针之一与靶增强子DNA序列杂交。

[0350] 实施方案11.如实施方案1至10中任一项所述的方法,其中所述第一组探针包括肽核酸探针。

[0351] 实施方案12.如实施方案1至11中任一项所述的方法,其中所述第二组探针包括包含3'和5'突出端的检测探针。

[0352] 实施方案13.如实施方案1至12中任一项所述的方法,其中所述第三组探针包含与所述第二组探针的所述3'和5'突出端互补的序列并且与所述第二组探针杂交以形成所述桥。

[0353] 实施方案14.如实施方案1至13中任一项所述的方法,其中步骤e)包括使用DNA聚合酶来延伸第三组探针。

[0354] 实施方案15.如实施方案1至14中任一项所述的方法,其中所述扩增步骤通过滚环扩增进行。

[0355] 实施方案16.如实施方案1至14中任一项所述的方法,其中所述扩增步骤还包括化学标记所述扩增子。

[0356] 实施方案17.如实施方案16所述的方法,其中所述扩增步骤还包括用荧光标记的探针化学标记所述扩增子。

[0357] 实施方案18.如实施方案1至17中任一项所述的方法,其中所述检测步骤包括荧光显微镜检查。

[0358] 实施方案19.如实施方案1至17中任一项所述的方法,其中所述检测步骤包括荧光原位杂交。

[0359] 实施方案20.如实施方案1至19中任一项所述的方法,其中所述染色质相关因子包括蛋白质、DNA、RNA、小分子和代谢物。

[0360] 实施方案21.如实施方案20所述的方法,其中所述染色质相关因子包括转录因子。

[0361] 实施方案22.如实施方案1至21中任一项所述的方法,其还包括以下的步骤:淬灭从所述信号产生标签生成的信号;和重复步骤b)-g),其中所述第一组和第二组探针各自包含至少两种单一寡核苷酸序列,所述单一寡核苷酸序列与参与染色质相互作用事件的至少两种二级靶核酸序列的有义链或反义链杂交。

[0362] 实施方案23.如实施方案1至21中任一项所述的方法,其还包括以下的步骤:淬灭来自所述信号产生标签的信号;和连续多轮次重复步骤b)-g),其中每个连续轮次包含具有与用于其他轮次的第一和第二组探针不同的寡核苷酸序列的第一和第二组探针,并且其中第一和第二组探针与不同于其他轮次中的靶核酸序列的靶核酸序列杂交。

[0363] 实施方案24.一种用于原位空间分析染色质相互作用事件的方法,所述方法包括以下步骤:获得生物样品;将包含至少两种单一寡核苷酸序列的第一寡核苷酸组与至少两种靶核酸序列的第一链杂交,其中所述至少两种靶核酸序列参与染色质相互作用事件;将第二寡核苷酸组与所述第一寡核苷酸组杂交,其中所述第一寡核苷酸由于与参与染色质相互作用事件的至少两种靶核酸序列杂交而彼此接近,并且其中所述第二寡核苷酸桥接所述第一寡核苷酸的末端;延伸第二寡核苷酸以形成环状寡核苷酸,所述环状寡核苷酸包含参与染色质相互作用事件的至少两种靶核酸序列的至少一部分;扩增所述环状寡核苷酸以生成可原位检测的扩增子群;以及原位检测所述扩增子。

[0364] 实施方案25.如实施方案24所述的方法,其还包括制备包含所述生物样品的载体基底的步骤。

[0365] 实施方案26.如实施方案25所述的方法,其中所述载体基底是其上具有生物样品的载玻片。

[0366] 实施方案27.如实施方案24至26中任一项所述的方法,其中将所述生物样品通过固定剂固定。

[0367] 实施方案28.如实施方案24至26中任一项所述的方法,其中不对所述生物样品进行固定。

[0368] 实施方案29.如实施方案24所述的方法,通过交联剂将组织样品内的染色质与染色质相关因子交联。



[0369] 实施方案30.如实施方案29所述的方法,其中所述交联剂选自包括以下的组:甲醛、UV辐射、戊二醛、双(亚胺酯)、双(琥珀酰亚胺酯)、二异氰酸酯、二酰氯、1-乙基-3-(3-二甲基氨基丙基)碳二亚胺、丝裂霉素C、氮芥、美法仑、1,3-丁二烯二环氧化物、顺式二胺二氯化铂(II)、环磷酰胺、双琥珀酰亚胺戊二酸酯和丙酸二硫代双琥珀酰亚胺酯。

[0370] 实施方案31.如实施方案24至30中任一项所述的方法,其中所述至少两种一级靶核酸序列包含调控元件。

[0371] 实施方案32.如实施方案31所述的方法,其中所述调控元件选自包括启动子、增强子、沉默子或绝缘子的组。

[0372] 实施方案33.如实施方案32所述的方法,其中所述第一寡核苷酸之一与靶启动子DNA序列杂交,并且所述第一寡核苷酸之一与靶增强子DNA序列杂交。

[0373] 实施方案34.如实施方案24至33中任一项所述的方法,其中所述第一寡核苷酸包括包含3'和5'突出端的检测探针。

[0374] 实施方案35.如实施方案24至34中任一项所述的方法,其中所述第二寡核苷酸包含与所述第一寡核苷酸的所述3'和5'突出端互补的序列并且与所述第一寡核苷酸杂交以形成所述桥。

[0375] 实施方案36.如实施方案24至35中任一项所述的方法,其中所述扩增步骤通过滚环扩增进行。

[0376] 实施方案37.如实施方案24至36中任一项所述的方法,其中所述扩增步骤还包括化学标记所述扩增子。

[0377] 实施方案38.如实施方案24至37中任一项所述的方法,其中所述扩增步骤还包括用荧光标记的探针化学标记所述扩增子。

[0378] 实施方案39.如实施方案24至38中任一项所述的方法,其中所述检测步骤包括荧光显微镜检查。

[0379] 实施方案40.如实施方案24至39中任一项所述的方法,其中所述检测步骤包括荧光原位杂交。

[0380] 实施方案41.如实施方案24至40中任一项所述的方法,其中所述染色质相关因子包括蛋白质、DNA、RNA、小分子和代谢物。

[0381] 实施方案42.如实施方案41所述的方法,其中所述染色质相关因子包括转录因子。

[0382] 实施方案43.如实施方案24至42中任一项所述的方法,其中DNA聚合酶用于延伸第二寡核苷酸。

[0383] 实施方案44.一种用于原位空间分析染色质相互作用事件的方法,所述方法包括:

[0384] a) 获得生物样品。

[0385] b) 将多种探针与参与染色质相互作用事件的多种一级靶核酸序列原位杂交,其中所述多种探针至少包括与环化探针杂交的检测探针;

[0386] c) 任选地将包含至少两种单一寡核苷酸序列的第二组探针与所述至少两种一级靶核酸序列的有义或反义链杂交;

[0387] d) 延伸环化探针以形成环化的核酸,所述环化的核酸包含参与染色质相互作用事件的一级靶核酸序列的至少一部分;

[0388] e) 扩增所述环状核酸以生成一个或多个扩增子群,每个所述扩增子群包含可原位

检测的信号产生标签;和

[0389] f) 通过所述信号产生标签原位检测所述带标签的扩增子群。

[0390] 实施方案45. 如实施方案44所述的方法, 其中所述信号产生标签对于一个或多个扩增子群是唯一的并且能够区分一个群与另一个群。

[0391] 实施方案46. 如实施方案44或45所述的方法, 其还包括通过原位测序对所述一个或多个扩增子群进行测序的步骤。

[0392] 实施方案47. 如实施方案44至46中任一项所述的方法, 其还包括制备包含所述生物样品的载体基底的步骤。

[0393] 实施方案48. 如实施方案47所述的方法, 其中所述载体基底是其上具有生物样品的载玻片。

[0394] 实施方案49. 如实施方案44至48中任一项所述的方法, 其中将所述生物样品通过固定剂固定。

[0395] 实施方案50. 如实施方案44至48中任一项所述的方法, 其中不对所述生物样品进行固定。

[0396] 实施方案51. 如实施方案44至50中任一项所述的方法, 其还包括使用交联剂将所述生物样品内的染色质与染色质相关因子交联的步骤;

[0397] 实施方案52. 如实施方案51所述的方法, 其中所述交联剂选自包括以下的组: 甲醛、UV辐射、戊二醛、双(亚胺酯)、双(琥珀酰亚胺酯)、二异氰酸酯、二酰氯、1-乙基-3-(3-二甲氨基丙基)碳二亚胺、丝裂霉素C、氮芥、美法仑、1,3-丁二烯二环氧化物、顺式二胺二氯化铂(II)、环磷酰胺、双琥珀酰亚胺戊二酸酯和丙酸二硫代双琥珀酰亚胺酯。

[0398] 实施方案53. 如实施方案44至52中任一项所述的方法, 其中所述一级靶核酸序列包含调控元件。

[0399] 实施方案54. 如实施方案44至53中任一项所述的方法, 其中所述调控元件选自包括启动子、增强子、沉默子或绝缘子的组。

[0400] 实施方案55. 如实施方案44至54中任一项所述的方法, 其中所述一级靶核酸序列包含基因组DNA。

[0401] 实施方案56. 如实施方案44至55中任一项所述的方法, 其中所述扩增步骤通过滚环扩增进行。

[0402] 实施方案57. 如实施方案44至56中任一项所述的方法, 其中所述扩增步骤还包括化学标记所述扩增子。

[0403] 实施方案58. 如实施方案44至57中任一项所述的方法, 其中所述扩增步骤还包括用荧光标记的探针化学标记所述扩增子。

[0404] 实施方案59. 如实施方案44至58中任一项所述的方法, 其中所述检测步骤包括荧光显微镜检查。

[0405] 实施方案60. 如实施方案44至59中任一项所述的方法, 其中所述检测步骤包括荧光原位杂交。

[0406] 实施方案61. 如实施方案44至60中任一项所述的方法, 其中所述染色质相关因子选自包括蛋白质、DNA、RNA、小分子和代谢物的组。

[0407] 实施方案62. 如实施方案61所述的方法, 其中所述染色质相关因子包括转录因子。

[0408] 实施方案63.如实施方案44所述的方法,其中DNA聚合酶用于延伸环化探针。

[0409] 实施例

[0410] 包括以下实施例仅出于说明的目的而非旨在限制本公开的范围。

[0411] 实施例1:在小鼠皮质中原位映射染色质相互作用

[0412] 细胞核内的DNA的三维构象在调控若干基因组区域(包括启动子和增强子)的表达中发挥重要作用。此实施例展示了用于在小鼠皮质中原位确定染色质相互作用的空间背景的方法。特别地,此实施例展示了,通过使用修改版本的Hi-C(Belton等人,2012;Lieberman-Aiden等人,2009)(一种帮助研究这些相互作用的工具)来获得如先前所述和要求保护的空間背景,可以映射基因组的拓扑结构。

[0413] Synj1是已经在帕金森病(Parkinson's disease)中观察到异常表达水平的基因。Hmox2是参与铁代谢的基因。先前已经确定,在帕金森病期间,脑铁代谢经历功能障碍(Jiang,Wang,Rogers,&Xie,2017)。通过Hi-C实验还显示,Synj1和Hmox2是基因组区域,它们不仅共表达,还在小鼠皮质内定期彼此相互作用(Babaei等人,2015)。上述测定用于观察小鼠皮质组织内两个基因的相互作用模式的空间分布,以增加对参与帕金森病的途径和机制的理解。

基因	转录物	探针序列	SEQ ID NO.
Synj1	NM_001045515.1	ACCAAAGTGTGCAAAGTCC	1
Hmox2	NM_010442.2	GAACCCAGTCTATGCCCCAC	2

[0415] 类似地,Dyrk1a产生磷酸化淀粉样前体蛋白(APP)和tau的激酶。已发现Dyrk1a在阿尔茨海默病(Alzheimer's)患者的死后大脑中上调(Velazquez等人,2019)。Sidt1是编码跨膜家族蛋白的基因,并且被认为在神经退行性疾病(诸如阿尔茨海默病和帕金森病)中发挥潜在作用(Hwang等人,2010)。Dyrk1a和Sidt1是共表达并且在基因组的三维构象中彼此物理相互作用的两个基因。

基因	转录物	探针序列	SEQ ID NO.
Dyrk1a	NM_007890.2	AGCCTGAATCGAGCAGAACC	3
Sidt1	NM_198034.3	CCCAGCATGTAGGAGGCAAA	4

[0417] 实施例2:在人细胞系中原位映射染色质相互作用

[0418] 细胞核内的DNA的三维构象在调控若干基因组区域(包括启动子和增强子)的表达中发挥重要作用。此实施例展示了用于在人主动脉平滑肌细胞(HASMC)系中映射染色质相互作用的方法。

[0419] CD52是在正常以及恶性免疫细胞中都表达的基因,并且被认为参与T细胞的迁移和激活(Toh,Kyaw,Tipping,&Bobik,2013)。尽管据称它在免疫系统中很重要,但它在其他细胞中的生物学作用和机制仍然未知。通过先前的Hi-C实验,已经观察到CD52启动子的表达受到其与被称为增强子的基因组区域的相互作用的影响。已经在人细胞系(诸如HASMC)中观察到CD52与其增强子(chr1:26615350-26617509)的相互作用,并且所述相互作用是有条件地表达的。使用上述测定有助于观察不同细胞类型中相互作用的分布以及哪些空间影响决定了这种表达。

	基因组区域	转录物	探针序列	SEQ ID NO.
[0420]	启动子: CD52_26644410	NM_001 803	GTCTCAGCCTTAGCCCTGTG	5
	相互作用增强子: chr1: 26615350-26617509	-	GCCTCTTCTTAGACGACCTG TGACGAACCC	6

[0421] 类似地,上述测定用于将这种相互作用与在整个组织中普遍存在且高度保守的另一种相互作用进行比较(De Bie等人,2006)。COMMD6基因产生NF- $\kappa$ -B(负责控制DNA的转录、细胞因子产生和细胞存活)-抑制蛋白,并且研究它们参与的途径可以引起鉴定若干种代谢障碍的新候选基因。

	基因组区域	转录物	探针序列	SEQ ID NO.
[0422]	启动子: COMMD6_76123575	NM_001 287394	TCGTCTTTCCA ACTCTGCGT	7
[0423]	相互作用增强子: chr13: 76119844 - 76120980	-	CACCCCCTGAAGTTATGGG GCGAC	8

[0424] 引用:

[0425] 1.Ali et al.(2014).Rolling circle amplification:A versatile tool for chemical biology,materials science and medicine.Chemical Society Reviews.Royal Society of Chemistry.<https://doi.org/10.1039/c3cs60439j>

[0426] 2.Babaei et al.(2015).Hi-C Chromatin Interaction Networks Predict Co-expression in the Mouse Cortex.PLOS Computational Biology,11(5), e1004221.<https://doi.org/10.1371/journal.pcbi.1004221>

[0427] 3.Belton et al.(2012).Hi-C:A comprehensive technique to capture the conformation of genomes.Methods,58(3),268-276.<https://doi.org/10.1016/j.ymeth.2012.05.001>

[0428] 4.Bienko et al.(2013).A versatile genome-scale PCR-based pipeline for high-definition DNA FISH.Nature Methods,10(2),122-124.<https://doi.org/10.1038/nmeth.2306>

[0429] 5.Broude et al.(1999).PNA Openers as a Tool for Direct Quantification of Specific Targets in Duplex DNA,Journal of Biomolecular Structure and Dynamics,17(2),237-244.<https://doi.org/10.1080/07391102.1999.10508356>

[0430] 6.De Bie et al.(2006).Characterization of COMMD protein-protein interactions in NF- $\kappa$ B signalling.Biochemical Journal,398(1),63-71.<https://doi.org/10.1042/BJ20051664>

[0431] 7.Hwang et al.(2010).Glycoproteomics in neurodegenerative diseases.Mass Spectrometry Reviews,29(1),79-125.<https://doi.org/10.1002/mas.20221>

[0432] 8.Jiang et al.(2017,May 1).Brain Iron Metabolism Dysfunction in Parkinson's Disease.Molecular Neurobiology.Humana Press Inc.<https://doi.org/10.1007/s12035-016-9879-1>

[0433] 9.Ke et al.(2013).In situ sequencing for RNA analysis in preserved tissue and cells.Nature Methods,10(9),857-860.<https://doi.org/10.1038/nmeth.2563>

[0434] 10.Klaesson et al.(2018).Improved efficiency of in situ protein analysis by proximity ligation using UnFold probes.Scientific Reports,8(1),5400.<https://doi.org/10.1038/s41598-018-23582-1>

[0435] 11.Kuhn,H.(2002).Rolling-circle amplification under topological constraints.Nucleic Acids Research,30(2),574-580.<https://doi.org/10.1093/nar/30.2.574>

[0436] 12.Lieberman-Aiden et al.(2009).Comprehensive mapping of long-range interactions reveals folding principles of the human genome.Science,326(5950),289-293.

[0437] <https://doi.org/10.1126/science.1181369>

[0438] 13.Sahlen et al.(2015).Genome-wide mapping of promoter-anchored interactions with close to single enhancer resolution.Genome Biology,16(1).<https://doi.org/10.1186/s13059-015-0727-9>

[0439] 14.Stougaard et al.(2011).Strategies for highly sensitive biomarker detection by Rolling Circle Amplification of signals from nucleic acid composed sensors.Integrative Biology,3(10),982.<https://doi.org/10.1039/c1ib00049g>

[0440] 15.Toh et al.(2013).Immune regulation by CD52-expressing CD4 T cells.Cellular and Molecular Immunology,10(5),379-382.<https://doi.org/10.1038/cmi.2013.35>

[0441] 16.Velazquez et al.(2019).Chronic Dyrk1 Inhibition Delays the Onset of AD-Like Pathology in 3xTg-AD Mice.Molecular Neurobiology,56(12),8364-8375.<https://doi.org/10.1007/s12035-019-01684-9>

[0442] 17.Yaroslavsky et al.(2013).Fluorescence imaging of single-copy DNA sequences within the human genome using PNA-directed padlock probe assembly, Chemistry&Biology,20(3),445-453.<https://doi.org/10.1016/j.chembiol.2013.02.012>

[0443] 本发明在范围上不旨在限制于具体公开的实施方案,提供这些实施方案例如来说明本发明的各个方面。根据本文的描述和教导,对所述组合物和方法的各种修改将变得显而易见。在不脱离本公开的真实范围和精神的情况下,可以实践这些变化,并且这些变化旨在落入本公开的范围。

## 序列表

<110> 10x基因组学有限公司(10X GENOMICS, INC.)

<120> 染色质相互作用的原位分析

<130> 20241-20064.40

<140> 尚未分配

<141> 2021-02-17

<150> US 62/977,512

<151> 2020-02-17

<160> 8

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 20

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> 合成构建体

<400> 1

accaaagtgt gcaaactgcc 20

<210> 2

<211> 20

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> 合成构建体

<400> 2

gaacccagtc tatgccccac 20

<210> 3

<211> 20

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> 合成构建体

<400> 3

agcctgaatc gagcagaacc 20

<210> 4

<211> 20

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>  
<223> 合成构建体  
<400> 4  
cccagcatgt aggaggcaaa 20  
<210> 5  
<211> 20  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<220>  
<223> 合成构建体  
<400> 5  
gtctcagcct tagccctgtg 20  
<210> 6  
<211> 30  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<220>  
<223> 合成构建体  
<400> 6  
gcctcttctt agacgacctg tgacgaacct 30  
<210> 7  
<211> 20  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<220>  
<223> 合成构建体  
<400> 7  
tcgtctttcc aactctgcgt 20  
<210> 8  
<211> 24  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<220>  
<223> 合成构建体  
<400> 8  
caccctga agttatgggg cgac 24

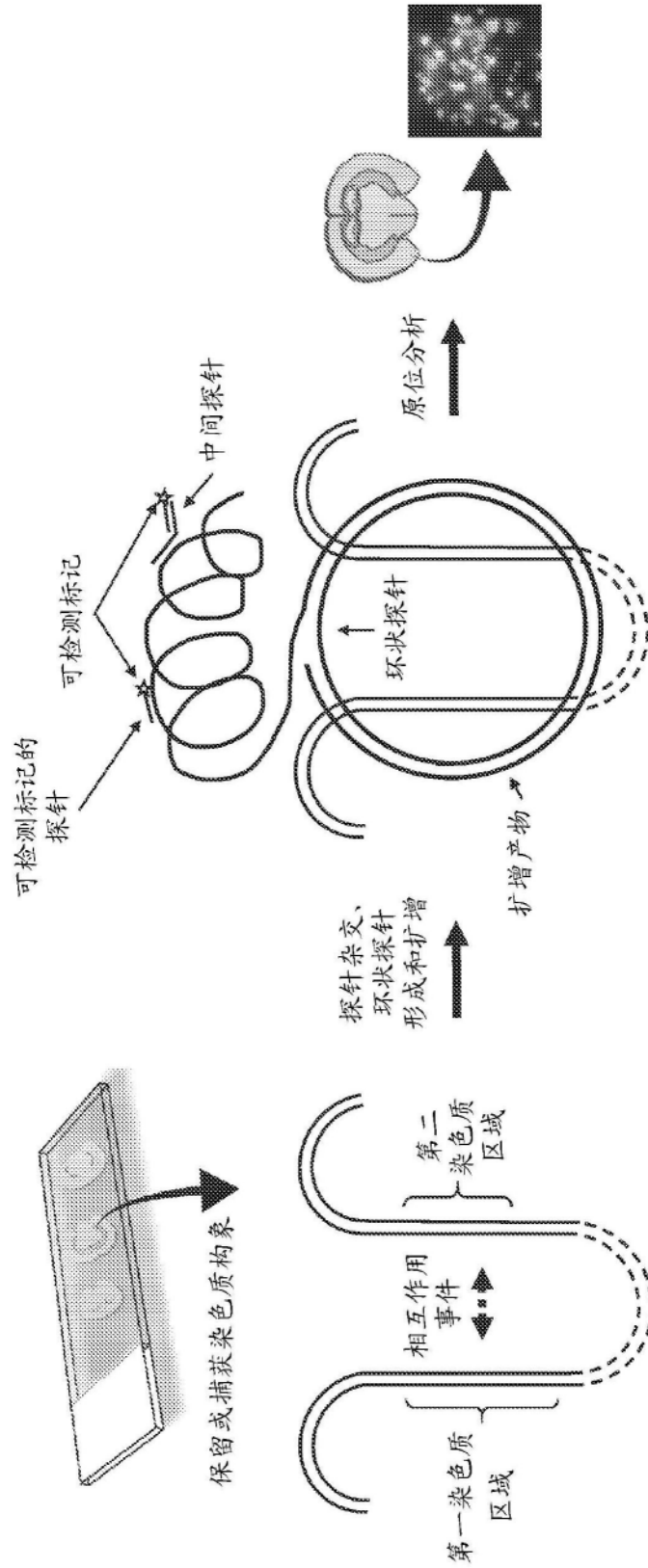


图1



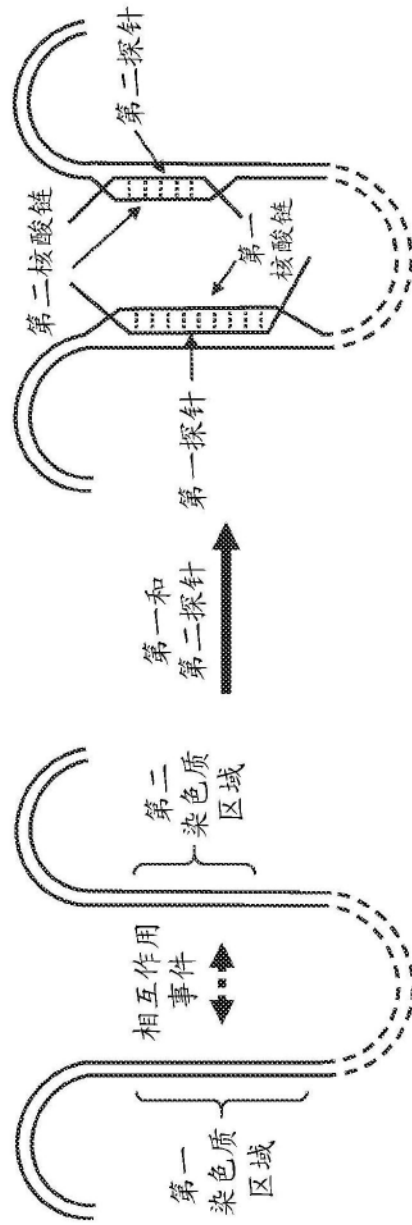


图2A

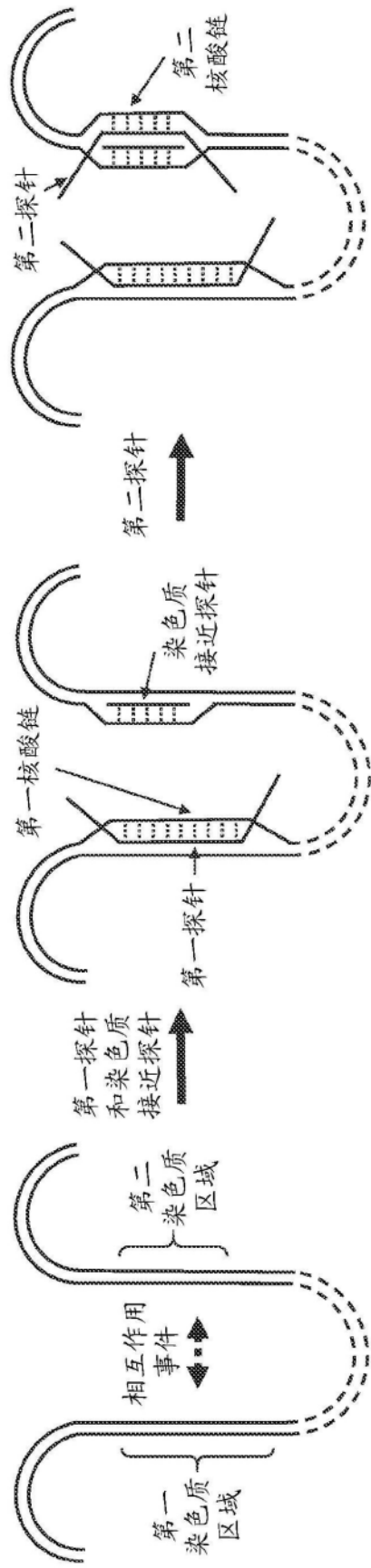


图2B

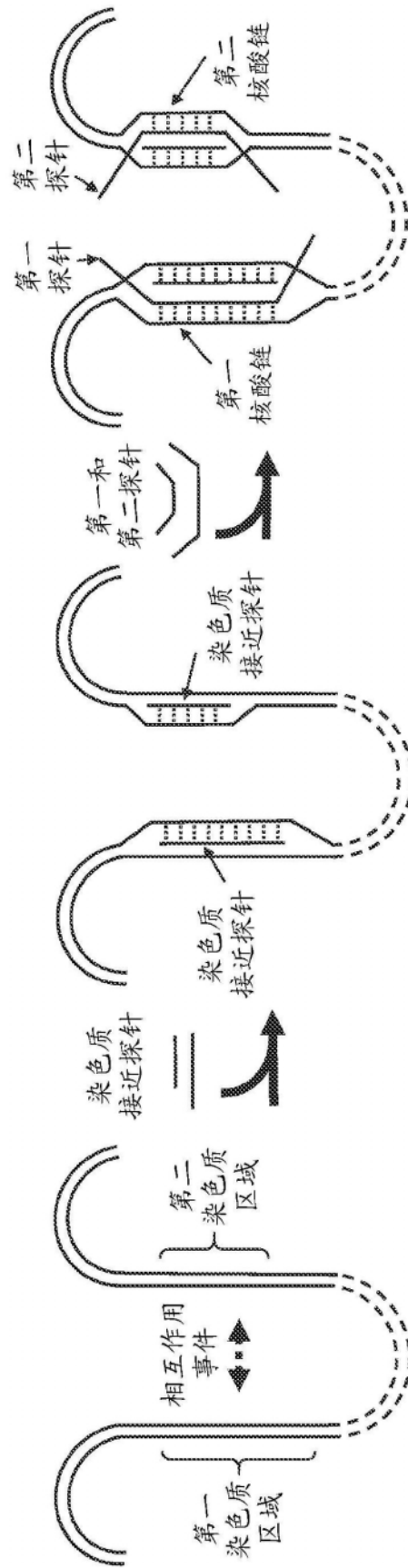


图3A

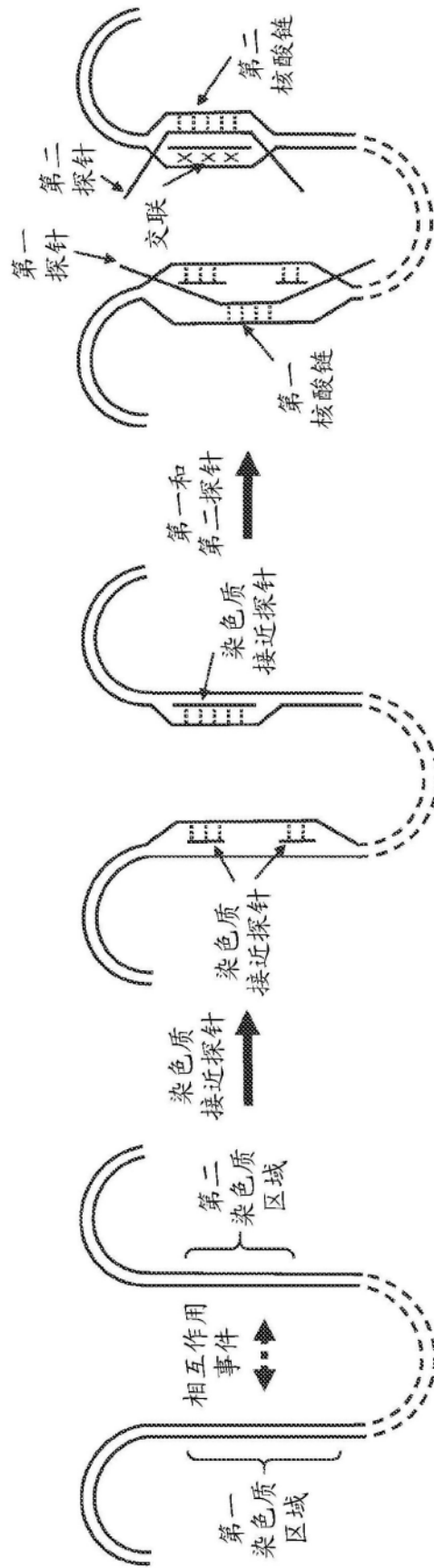


图3B

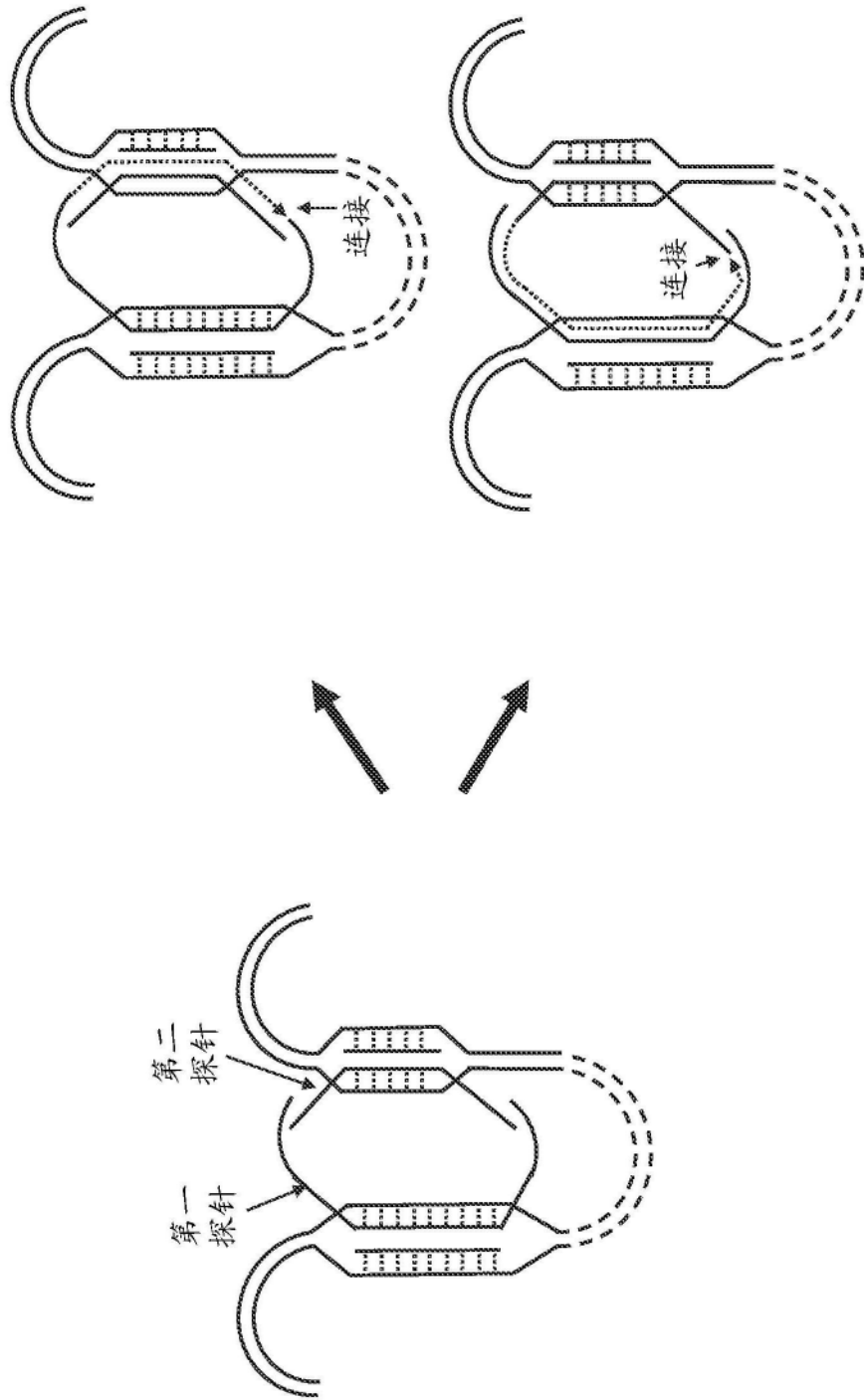


图4

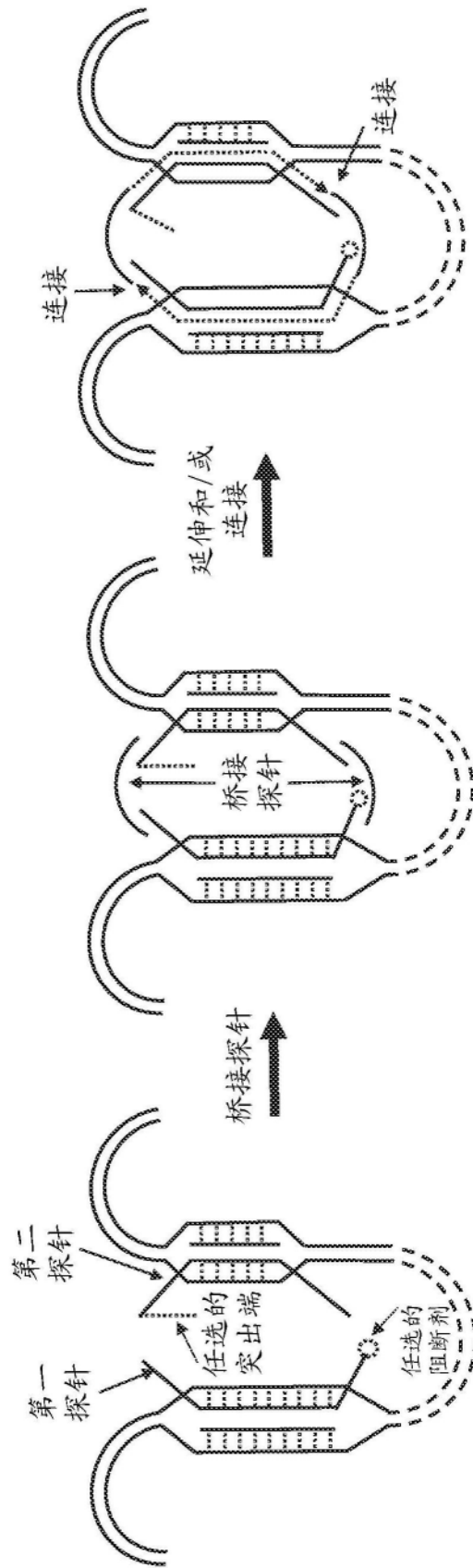


图5A

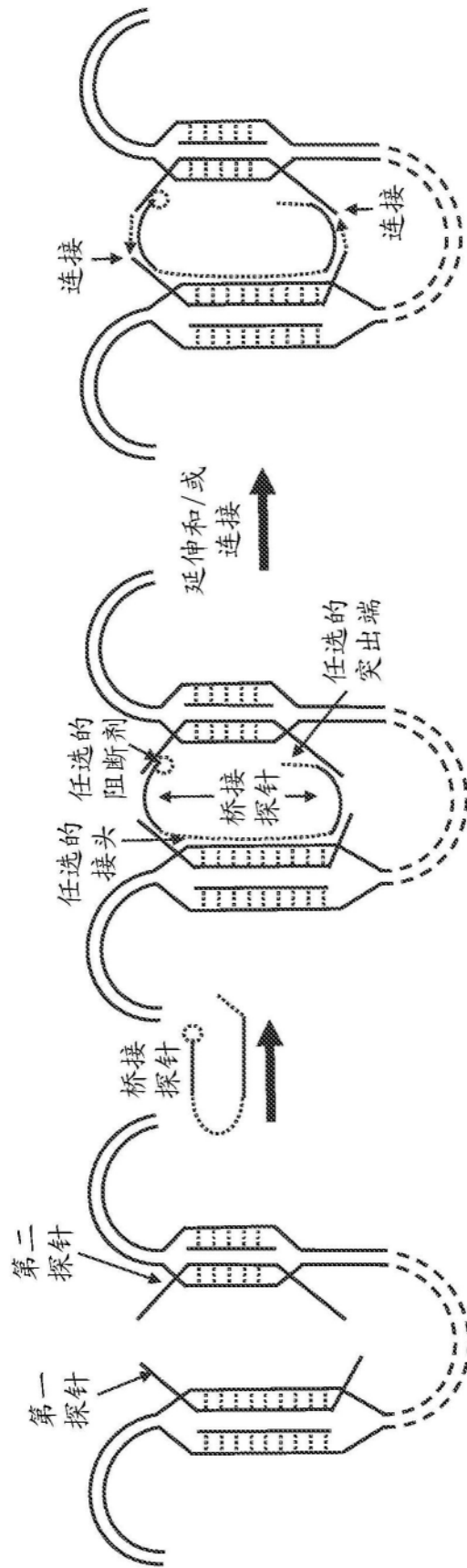


图5B

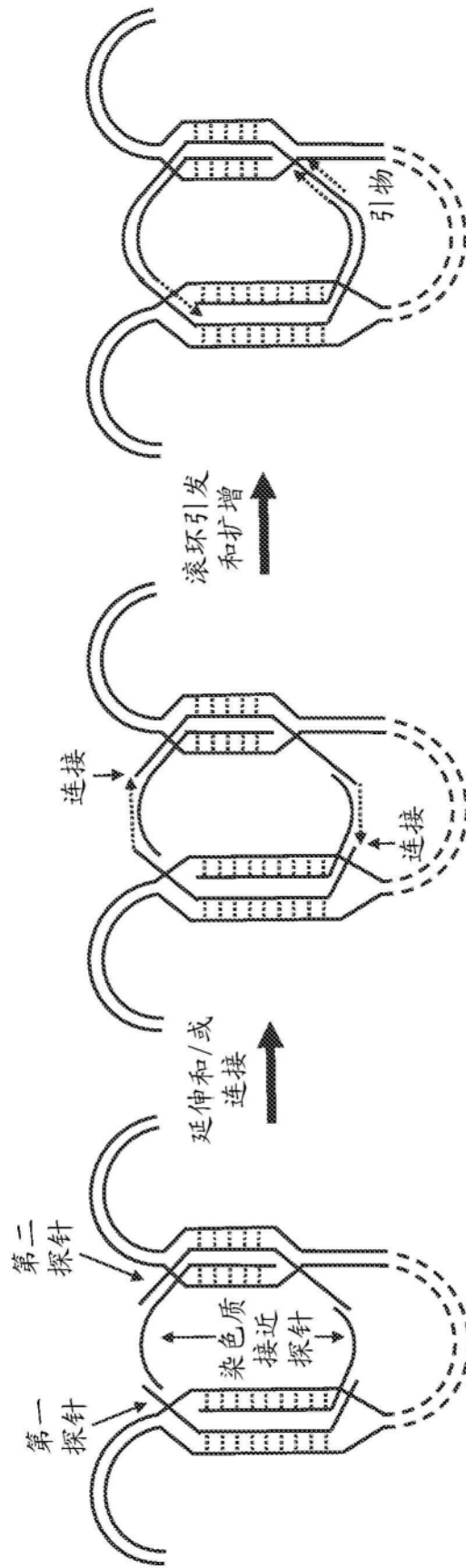


图6A



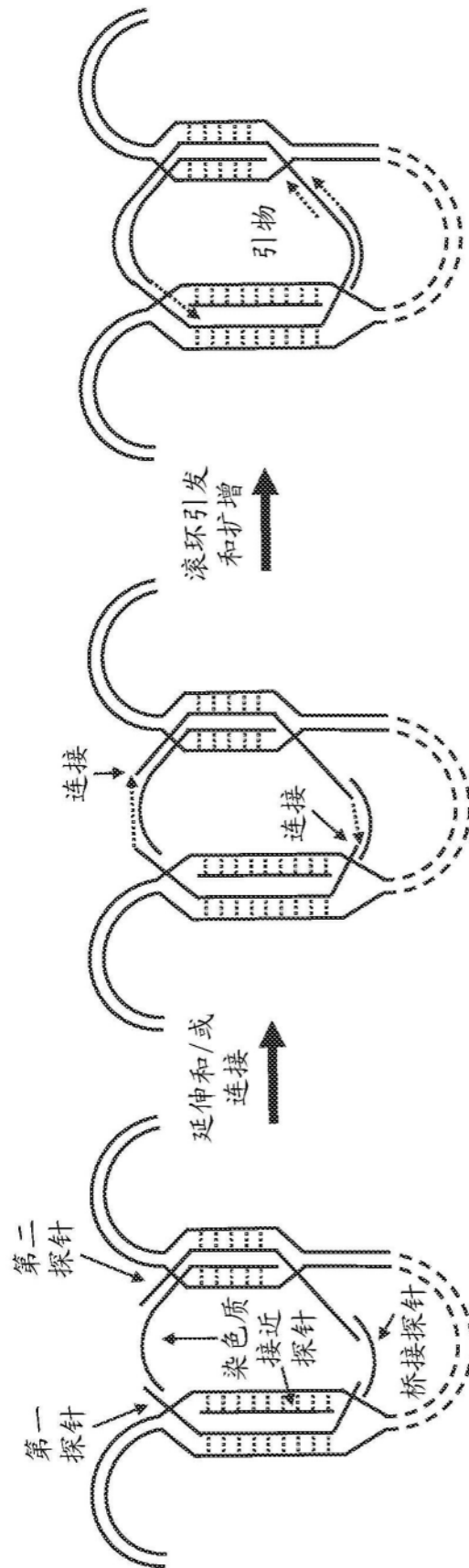


图6B

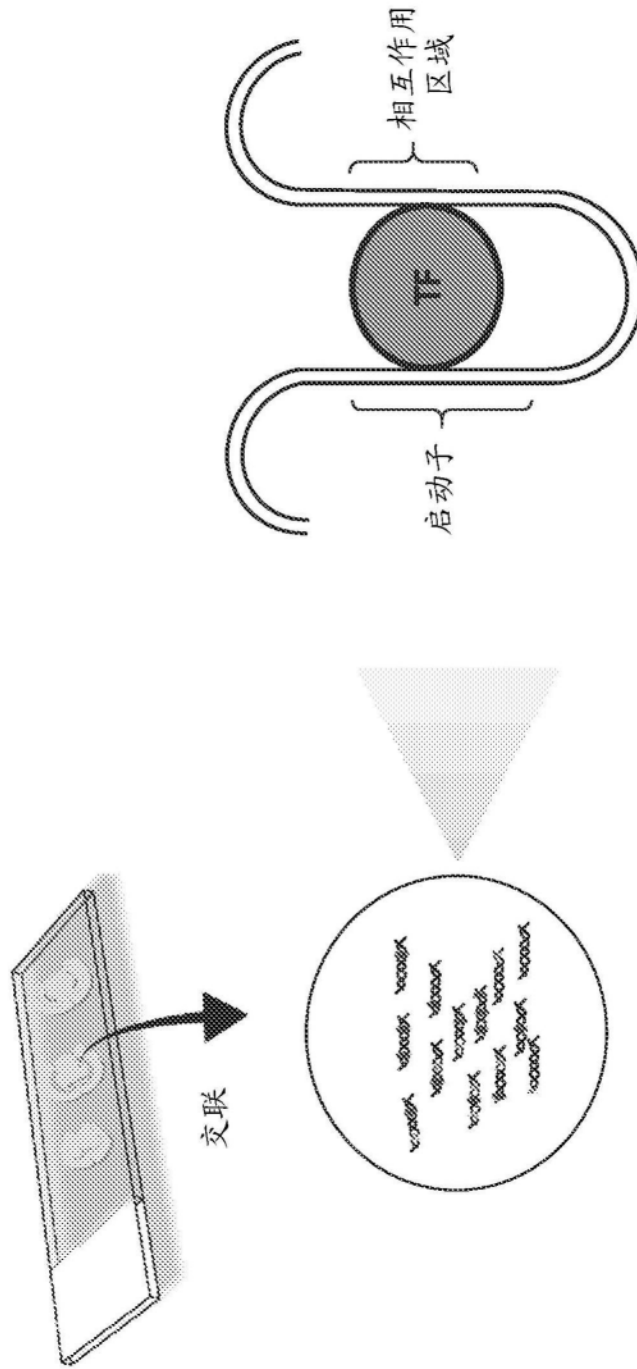


图7

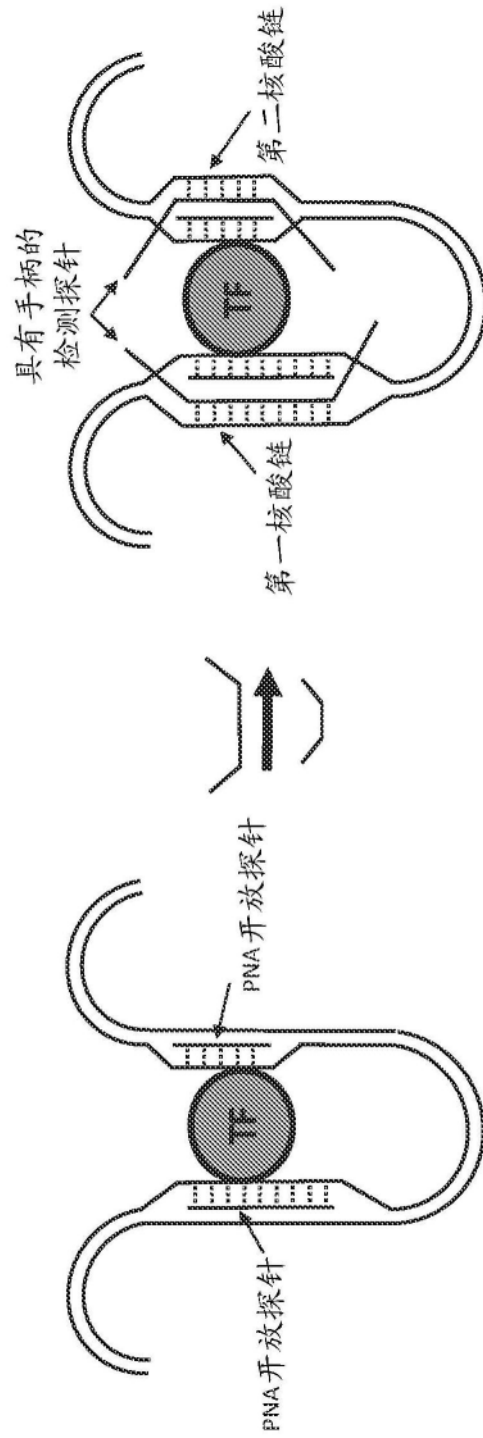
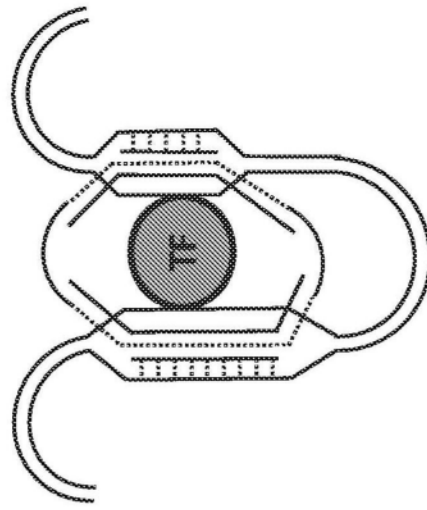
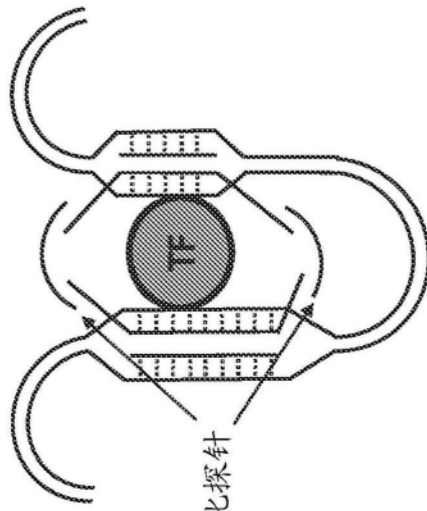


图8



环化探针的  
延伸/环化 ↑



环化探针

图9

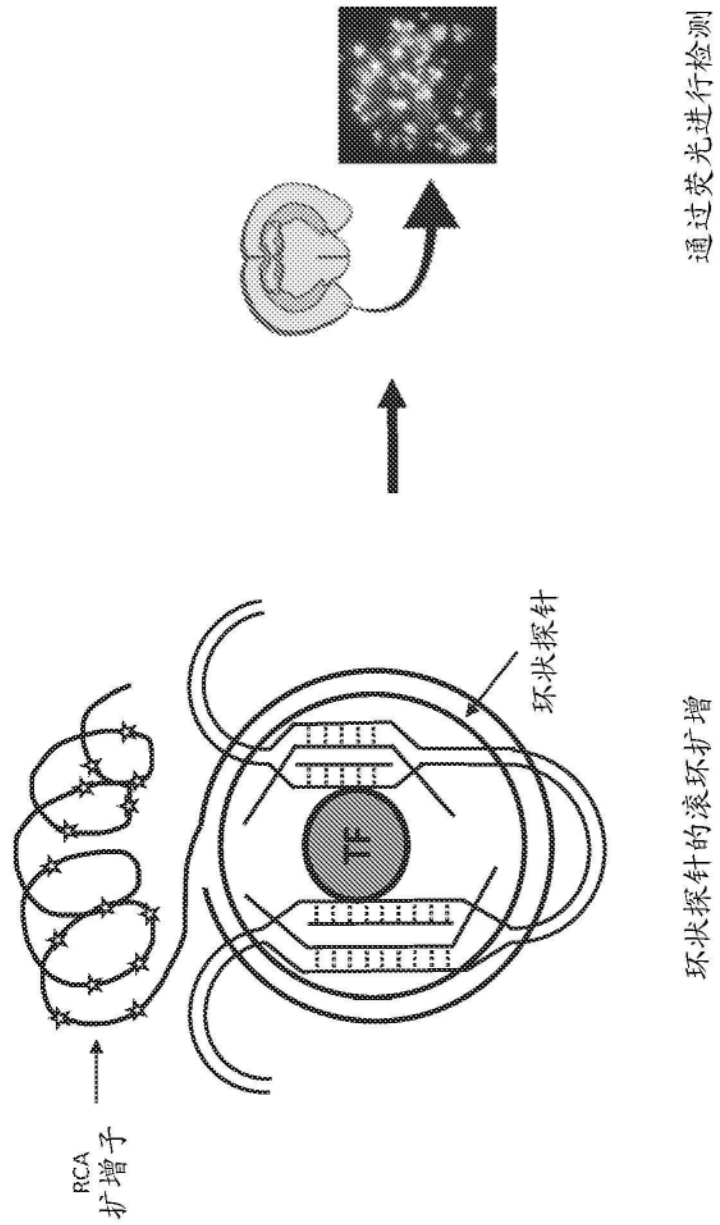


图10