



(12) 发明专利申请

(10) 申请公布号 CN 105392898 A

(43) 申请公布日 2016. 03. 09

(21) 申请号 201480037597. 5

(51) Int. Cl.

(22) 申请日 2014. 04. 30

C12Q 1/68(2006. 01)

(30) 优先权数据

61/817, 651 2013. 04. 30 US

61/971, 974 2014. 03. 28 US

(85) PCT国际申请进入国家阶段日

2015. 12. 30

(86) PCT国际申请的申请数据

PCT/US2014/036258 2014. 04. 30

(87) PCT国际申请的公布数据

W02014/182528 EN 2014. 11. 13

(71) 申请人 加州理工学院

地址 美国加利福尼亚州

(72) 发明人 龙·蔡 埃里克·吕贝克

铁木尔·契彦塔耶夫

阿梅特·科斯坤 庭-芳·何

长·赫·孙 希尔·沙哈

(74) 专利代理机构 北京安信方达知识产权代理

有限公司 11262

代理人 刘小立 郑霞

权利要求书7页 说明书50页 附图26页

(54) 发明名称

通过顺序杂交编条形码的分子多重标记

(57) 摘要

本发明, 除其他以外, 提供了用于检测和 / 或定量细胞、组织、器官或生物体中的核酸的技术。在一些实施方案中, 本发明通过顺序编条形码提供了用于大量靶的高通量谱分析的方法, 所述靶诸如转录物和 / 或 DNA 基因座。



1. 一种组合物,所述组合物包含多种可检测地标记的寡核苷酸,所述多种可检测地标记的寡核苷酸的每种靶向转录物或 DNA 基因座并且被可检测部分标记,因此所述组合物至少包含:

(i) 第一寡核苷酸,所述第一寡核苷酸靶向第一转录物或 DNA 基因座并被第一可检测部分标记;和

(ii) 第二寡核苷酸,所述第二寡核苷酸靶向第二转录物或 DNA 基因座并被第二可检测部分标记。

2. 一种试剂盒,所述试剂盒包含多种可检测地标记的寡核苷酸,所述多种可检测地标记的寡核苷酸的每种靶向转录物或 DNA 基因座并被可检测部分标记,因此所述试剂盒至少包含:

(i) 第一寡核苷酸,所述第一寡核苷酸靶向第一转录物或 DNA 基因座并被第一可检测部分标记;

(ii) 第二寡核苷酸,所述第二寡核苷酸靶向第二转录物或 DNA 基因座并被第二可检测部分标记;

(iii) 第三寡核苷酸,所述第三寡核苷酸任选地在序列上与所述第一寡核苷酸相同,所述第三寡核苷酸靶向所述第一转录物或 DNA 基因座并被所述第一可检测部分、所述第二可检测部分或第三可检测部分标记;和

(iv) 第四寡核苷酸,所述第四寡核苷酸任选地在序列上与所述第二寡核苷酸相同,所述第四寡核苷酸靶向所述第二转录物或 DNA 基因座并被所述第一可检测部分、所述第二可检测部分、所述第三可检测部分或第四可检测部分标记,

其中所述第三寡核苷酸被与所述第一寡核苷酸不同的可检测部分标记或所述第四寡核苷酸被与所述第二寡核苷酸不同的可检测部分标记,或两者。

3. 如前述权利要求中任一项所述的组合物或试剂盒,其中每种寡核苷酸被荧光团标记。

4. 如前述权利要求中任一项所述的组合物或试剂盒,其中所述第一寡核苷酸和所述第二寡核苷酸靶向不同的转录物或 DNA 基因座。

5. 如前述权利要求中任一项所述的组合物或试剂盒,其中所述第一寡核苷酸和所述第二寡核苷酸靶向相同的转录物或 DNA 基因座。

6. 如权利要求 5 所述的组合物或试剂盒,其中多种可检测地标记的寡核苷酸包含靶向相同的转录物或 DNA 基因座的两种或更多种可检测地标记的寡核苷酸。

7. 如权利要求 5 所述的组合物或试剂盒,其中所述第一寡核苷酸和所述第二寡核苷酸具有不同的寡核苷酸序列。

8. 如前述权利要求中任一项所述的组合物或试剂盒,其中所述第一可检测部分和所述第二可检测部分是不同的。

9. 如前述权利要求中任一项所述的组合物或试剂盒,其中所述第一可检测部分和所述第二可检测部分是相同的。

10. 如前述权利要求中任一项所述的组合物或试剂盒,其中所述第一寡核苷酸和所述第二寡核苷酸具有小于 68% 的序列同一性。

11. 如前述权利要求中任一项所述的组合物或试剂盒,其中每种寡核苷酸与任何其他

核苷酸具有小于 50% 的序列同一性。

12. 如前述权利要求中任一项所述的组合物或试剂盒, 其中每种寡核苷酸的长度小于 29 个碱基对。

13. 如前述权利要求中任一项所述的组合物或试剂盒, 其中所述多种可检测地标记的寡核苷酸靶向至少 10 种不同的转录物和 / 或 DNA 基因座。

14. 如前述权利要求中任一项所述的组合物或试剂盒, 包含至少两种可检测部分。

15. 如前述权利要求中任一项所述的组合物, 所述组合物还包含:

(iii) 第三寡核苷酸, 所述第三寡核苷酸任选地在序列上与所述第一寡核苷酸相同, 所述第三寡核苷酸靶向所述第一转录物或 DNA 基因座; 和

(iv) 第四寡核苷酸, 所述第四寡核苷酸任选地在序列上与所述第二寡核苷酸相同, 所述第四寡核苷酸靶向所述第二转录物或 DNA 基因座,

其中所述第三寡核苷酸被与所述第一寡核苷酸不同的可检测部分标记, 或所述第四寡核苷酸被与所述第二寡核苷酸不同的可检测部分标记, 或两者。

16. 如前述权利要求中任一项所述的组合物或试剂盒, 其中所述第三寡核苷酸在序列上与所述第一寡核苷酸相同。

17. 如前述权利要求中任一项所述的组合物或试剂盒, 其中所述第三寡核苷酸被与所述第一寡核苷酸不同的可检测部分标记。

18. 如前述权利要求中任一项所述的组合物或试剂盒, 其中所述第四寡核苷酸在序列上与所述第二寡核苷酸相同。

19. 如前述权利要求中任一项所述的组合物或试剂盒, 其中所述第四寡核苷酸被与所述第二寡核苷酸不同的可检测部分标记。

20. 一种方法, 所述方法包括以下步骤:

(a) 进行第一接触步骤, 所述第一接触步骤包括使包含多种核酸的细胞与第一多种可检测地标记的寡核苷酸接触, 所述第一多种可检测地标记的寡核苷酸的每种靶向核酸并被可检测部分标记, 因此组合物至少包含:

(i) 第一寡核苷酸, 所述第一寡核苷酸靶向第一核酸并被第一可检测部分标记; 和

(ii) 第二寡核苷酸, 所述第二寡核苷酸靶向第二核酸并被第二可检测部分标记;

(b) 在所述第一接触步骤之后对所述细胞成像, 以便检测所述第一多种的寡核苷酸与其靶的相互作用;

(c) 进行第二接触步骤, 所述第二接触步骤包括使所述细胞与第二多种可检测地标记的寡核苷酸接触, 所述第二多种包括靶向被所述第一多种靶向的重叠核酸的寡核苷酸, 因此所述第二多种至少包含:

(i) 第三寡核苷酸, 所述第三寡核苷酸任选地在序列上与所述第一寡核苷酸相同, 所述第三寡核苷酸靶向所述第一核酸; 和

(ii) 第四寡核苷酸, 所述第四寡核苷酸任选地在序列上与所述第二寡核苷酸相同, 所述第三寡核苷酸靶向所述第二核苷酸,

其中所述第二多种与所述第一多种的不同在于, 存在于所述第二多种中的寡核苷酸的至少一种被与所述第一多种中的靶向相同的核酸的对应寡核苷酸不同的可检测部分标记, 因此在所述第二多种中:

(iii) 所述第三寡核苷酸被所述第一可检测部分、所述第二可检测部分或第三可检测部分标记；并且

(iv) 所述第四寡核苷酸被所述第一可检测部分、所述第二可检测部分、所述第三可检测部分或第四可检测部分标记，

其中所述第三寡核苷酸被与所述第一寡核苷酸所用的不同的可检测部分标记，或所述第四寡核苷酸被与所述第二寡核苷酸所用的不同的可检测部分标记，或两者；

(d) 在所述第二接触步骤之后对所述细胞成像，因此以便检测所述第二多种的寡核苷酸与其靶的相互作用；以及

(e) 任选地重复所述接触步骤和成像步骤，每次用新的多种可检测地标记的寡核苷酸，所述新的多种可检测地标记的寡核苷酸包含靶向被所述第一多种和所述第二多种靶向的重叠核酸的寡核苷酸，其中由于标记靶向相同核酸的寡核苷酸的可检测部分中的至少一种差异，所利用的每个多种与所利用的每个其他的多种不同。

21. 一种方法，所述方法包括以下步骤：

(a) 进行第一接触步骤，所述第一接触步骤包括使包含多种转录物和 DNA 基因座的细胞与第一多种可检测地标记的寡核苷酸接触，所述第一多种可检测地标记的寡核苷酸的每种靶向转录物或 DNA 基因座并被可检测部分标记，因此组合物至少包含：

(i) 第一寡核苷酸，所述第一寡核苷酸靶向第一转录物或 DNA 基因座并用第一可检测部分标记；和

(ii) 第二寡核苷酸，所述第二寡核苷酸靶向第二转录物或 DNA 基因座并用第二可检测部分标记；

(b) 在所述第一接触步骤之后对所述细胞成像，以便检测所述第一多种的寡核苷酸与其靶的杂交；

(c) 进行第二接触步骤，所述第二接触步骤包括使所述细胞与第二多种可检测地标记的寡核苷酸接触，该第二多种包含靶向被所述第一多种靶向的重叠转录物和 / 或 DNA 基因座的寡核苷酸，因此所述第二多种至少包含：

(i) 第三寡核苷酸，所述第三寡核苷酸任选地在序列上与所述第一寡核苷酸相同，所述第三寡核苷酸靶向所述第一转录物或 DNA 基因座；和

(ii) 第四寡核苷酸，所述第四寡核苷酸任选地在序列上与所述第二寡核苷酸相同，所述第四寡核苷酸靶向所述第二转录物或 DNA 基因座，

其中所述第二多种与所述第一多种的不同在于，存在于所述第二多种中的寡核苷酸的至少一种被与所述第一多种中靶向相同的转录物或 DNA 基因座的对应寡核苷酸不同的可检测部分标记，因此在所述第二多种中：

(iii) 所述第三寡核苷酸被所述第一可检测部分、所述第二可检测部分或第三可检测部分标记；并且

(iv) 所述第四寡核苷酸被所述第一可检测部分、所述第二可检测部分、所述第三可检测部分或第四可检测部分标记，

其中所述第三寡核苷酸被与所述第一寡核苷酸所用的不同的可检测部分标记，或所述第四寡核苷酸用与所述第二寡核苷酸所用的不同的可检测部分标记，或两者；

(d) 在所述第二接触步骤之后对所述细胞成像，以便检测所述第二多种的寡核苷酸与

其靶的杂交被检测出；以及

(e) 任选地重复所述接触步骤和成像步骤,每次用新的多种可检测地标记的寡核苷酸,所述新的多种可检测地标记的寡核苷酸包含靶向所述第一多种和所述第二多种靶向的重叠转录物或 DNA 基因座的寡核苷酸,其中由于标记靶向相同转录物或 DNA 基因座的寡核苷酸的可检测部分中的至少一种差异,所利用的每个多种与所利用的每个其他的多种不同。

22. 如权利要求 20 或 21 所述的方法,其中可检测地标记的寡核苷酸的每个多种包含 F 种可检测部分,其中 F 至少是 2。

23. 如权利要求 20-22 中任一项所述的方法,所述方法包括 N 个接触步骤,其中 N 至少是 2。

24. 如权利要求 20-23 中任一项所述的方法,其中可检测地标记的核苷酸的每个多种靶向 (F)^N个转录物和 / 或 DNA 基因座。

25. 如权利要求 20-23 中任一项所述的方法,其中可检测地标记的寡核苷酸的每个多种靶向少于 (F)^N个转录物和 / 或 DNA 基因座。

26. 如权利要求 20-25 中任一项所述的方法,其中可检测地标记的寡核苷酸的每个多种靶向相同的转录物和 / 或 DNA 基因座。

27. 如权利要求 20-26 中任一项所述的方法,所述方法任选地在成像步骤之后包括去除多种可检测地标记的寡核苷酸的步骤。

28. 如权利要求 20-27 中任一项所述的方法,所述方法在每个成像步骤之后包括去除多种可检测地标记的寡核苷酸的步骤。

29. 如权利要求 28 所述的方法,其中所述去除的步骤包括使所述多种可检测地标记的寡核苷酸与消化可检测地标记的寡核苷酸的酶接触。

30. 如权利要求 29 所述的方法,其中所述去除的步骤包括使所述多种可检测地标记的寡核苷酸与 DNA 酶接触。

31. 如权利要求 29 所述的方法,其中所述去除的步骤包括使所述多种可检测地标记的寡核苷酸与 RNA 酶接触。

32. 如权利要求 20-31 中任一项所述的方法,其中所述去除的步骤包括光漂白。

33. 如权利要求 20-32 中任一项所述的方法,其中每个多种包含靶向相同的转录物或 DNA 基因座的两种或更多种可检测地标记的寡核苷酸。

34. 如权利要求 33 所述的方法,其中靶向相同的转录物或 DNA 基因座的所有可检测地标记的寡核苷酸被提供相同的颜色的荧光团标记。

35. 如权利要求 33 所述的方法,其中靶向相同的转录物或 DNA 基因座的所有可检测地标记的寡核苷酸被相同的荧光团标记。

36. 如权利要求 20-35 中任一项所述的方法,其中可检测地标记的寡核苷酸的每个多种与另一个不同。

37. 如权利要求 20-36 中任一项所述的方法,其中所述第三寡核苷酸被与所述第一寡核苷酸不同的可检测部分标记,并且所述第四寡核苷酸被与所述第二寡核苷酸相同的可检测部分标记;或所述第三寡核苷酸被与所述第一寡核苷酸相同的可检测部分标记,并且所述第四寡核苷酸被与所述第二寡核苷酸不同的可检测部分标记。

38. 如权利要求 20-36 中任一项所述的方法,其中所述第三寡核苷酸被与所述第一寡

核苷酸不同的可检测部分标记,并且所述第四寡核苷酸被与所述第二寡核苷酸不同的可检测部分标记。

39. 如权利要求 20-38 中任一项所述的方法,其中成像步骤包括在接触步骤之后对所述细胞成像,因此多种可检测地标记的寡核苷酸与其靶的杂交被定量。

40. 如权利要求 20-39 中任一项所述的方法,所述方法还包括 HCR、光片照明显微术、CLARITY 或其组合。

41. 如前述权利要求中任一项所述的组合物、试剂盒或方法,其中可检测地标记的寡核苷酸的每个多种的靶是预先确定的。

42. 如前述权利要求中任一项所述的组合物、试剂盒或方法,其中每种可检测地标记的寡核苷酸的量是预先确定的。

43. 如前述权利要求中任一项所述的组合物、试剂盒或方法,其中可检测地标记的寡核苷酸与中间寡核苷酸杂交,所述中间寡核苷酸与靶杂交。

44. 如权利要求 43 所述的组合物、试剂盒或方法,其中中间寡核苷酸包含与所述靶的序列互补的序列和与可检测地标记的寡核苷酸互补的突出端序列。

45. 如前述权利要求中任一项所述的组合物、试剂盒或方法,其中存在两种或更多种中间寡核苷酸与靶杂交。

46. 如权利要求 45 所述的组合物,其中与靶杂交的每种中间寡核苷酸包含相同的突出端序列。

47. 如前述权利要求中任一项所述的组合物、试剂盒或方法,其中可检测地标记的寡核苷酸包含 5' - 磷酸化。

48. 如前述权利要求中任一项所述的组合物、试剂盒或方法,其中可检测地标记的寡核苷酸包含一个或更多个尿嘧啶。

49. 如权利要求 43-46 中任一项所述的组合物、试剂盒或方法,其中可检测地标记的寡核苷酸在与中间寡核苷酸杂交后,具有凹陷的 3' - 末端。

50. 如权利要求 20-49 中任一项所述的方法,所述方法还包括以下步骤:

(f) 进行接触步骤,所述接触步骤包括使包含多种核酸的细胞与多种中间寡核苷酸接触,所述多种中间寡核苷酸的每种:

(i) 靶向核酸并且任选地被可检测部分标记;并且

(ii) 在与所述靶杂交后包含突出端序列;以及

(g) 任选地对所述细胞成像,以便检测所述中间寡核苷酸与其靶之间的相互作用。

51. 如权利要求 50 所述的方法,其中所述中间寡核苷酸不被可检测部分标记。

52. 如权利要求 50 所述的方法,其中不存在步骤 (g)。

53. 如权利要求 50 所述的方法,其中步骤 (f) 和任选地步骤 (g) 在步骤 (a) 之前进行。

54. 如权利要求 20-53 中任一项所述的方法,其中去除步骤不破坏中间寡核苷酸。

55. 如权利要求 50-54 中任一项所述的方法,其中所述核酸是转录物或 DNA 基因座。

56. 如权利要求 20-55 中任一项所述的方法,其中所述细胞是人类细胞。

57. 如权利要求 20-56 中任一项所述的方法,其中所述细胞在组织或器官内。

58. 如前述权利要求中任一项所述的组合物、试剂盒或方法,其中至少一种可检测地标记的寡核苷酸被 HCR 标记。

59. 如前述权利要求中任一项所述的组合物、试剂盒或方法,其中至少一种可检测地标记的寡核苷酸与中间寡核苷酸杂交,所述中间寡核苷酸与转录物或 DNA 基因座杂交。

60. 如前述权利要求中任一项所述的方法,所述方法包括去除步骤,所述去除步骤包括使用核酸外切酶 III、Lambda 核酸外切酶或尿嘧啶特异性切除试剂。

61. 一种光片照明显微镜,其中样品被从其侧部照射。

62. 如权利要求 61 所述的显微镜,其中所述样品被从其侧部照射,并且光片平行于样品台。

63. 如权利要求 61 或 62 所述的显微镜,其中所述光片垂直于检测物体。

64. 如权利要求 61-63 中任一项所述的显微镜,其中所述样品被沿着 z- 轴扫描。

65. 一种用于制备靶核酸的方法,所述靶核酸包含第一序列,所述方法包括以下步骤:

1) 提供包含所述第一序列的第一核酸,其中所述第一序列在两末端的侧翼是切口核酸内切酶位点;

2) 扩增所述第一核酸或所述第一核酸的一部分以提供第二核酸,所述第二核酸包含所述第一序列和侧翼切口核酸内切酶位点;以及

3) 使所述第二核酸与对应于所述侧翼切口核酸内切酶位点的一种或更多种切口核酸内切酶接触。

66. 如权利要求 65 所述的方法,其中两个侧翼切口核酸内切酶位点是相同的。

67. 如权利要求 66 所述的方法,其中使用对应于相同的切口核酸内切酶位点的一种切口核酸内切酶。

68. 如权利要求 65 所述的方法,其中两个侧翼切口核酸内切酶位点是不同的。

69. 如权利要求 68 所述的方法,其中使用两种切口核酸内切酶,所述两种切口核酸内切酶的每种独立地对应于切口核酸内切酶位点。

70. 如权利要求 65-69 中任一项所述的方法,所述方法包括变性步骤,其中双链核酸被转变成单链核酸。

71. 一种用于制备靶核酸的方法,所述靶核酸包含第一序列,所述方法包括以下步骤:

1) 提供包含所述第一序列或其互补序列的第一核酸,其中所述第一序列或其互补序列的侧翼为至少一个限制性位点;

2) 扩增所述第一核酸或所述第一核酸的一部分,以提供第二核酸,所述第二核酸包含所述第一序列和所述至少一个侧翼限制性位点;以及

3) 使所述第二核酸与对应于所述至少一个侧翼限制性位点的限制性酶接触,以提供包含凹陷末端的第三核酸;

4) 使所述第三核酸与核酸酶接触,以使如果所述互补序列存在,选择性地消化包含所述互补序列的链,同时保留包含所述第一序列的链。

72. 如权利要求 71 所述的方法,其中所述第一序列或其互补序列在两端的侧翼独立地为限制性位点。

73. 如权利要求 71 或 72 所述的方法,其中所述第二核酸与对应于所述至少一个侧翼限制性位点的每个的限制性酶接触。

74. 如权利要求 71-73 中任一项所述的方法,其中包含所述互补序列的所述链在限制性消化后具有 3' 凹陷末端。

75. 如权利要求 71-74 中任一项所述的方法,其中包含所述第一序列的所述链在限制性消化后具有 5' 凹陷末端。

76. 如权利要求 71-75 中任一项所述的方法,其中所述核酸酶是核酸外切酶 III。

77. 如权利要求 65-76 中任一项所述的方法,其中所述靶核酸是 DNA。

78. 如权利要求 65-77 中任一项所述的方法,其中所述靶核酸具有与所述第一序列相同的序列。

79. 如权利要求 65-78 中任一项所述的方法,其中在随后的接触步骤之前,所述第二核酸和 / 或所述第三核酸任选地且独立地被修饰。

80. 如权利要求 65-79 中任一项所述的方法,其中所述靶核酸是包含第一序列和第二序列的中间寡核苷酸,所述第一序列与靶杂交,所述第二序列与被 HCR 标记的可检测地标记的寡核苷酸杂交。

81. 如权利要求 65-79 中任一项所述的方法,其中所述靶核酸是桥接探针。

通过顺序杂交编条形码的分子多重标记

[0001] 对相关申请的交叉引用

[0002] 本申请要求 2013 年 4 月 30 日提交的美国临时申请序列号 61/817,651 和 2014 年 3 月 28 日提交的美国临时申请序列号 61/971,974 的优先权,通过引用各自以其整体并入本文。

[0003] 关于联邦政府赞助的研究或开发的声明

[0004] 依据由美国国家卫生研究院授予的拨款号 R01HD075605,美国政府对此发明具有一定的权利。

[0005] 发明背景

[0006] 细胞的转录谱分析 (profiling) 是许多目的必需的。能够在单细胞中分辨多种 mRNA 的显微成像可以提供关于转录物丰度和定位的有价值的信息,对理解细胞鉴定的分子基础和开发疾病治疗是重要的。因此,对通过例如显微成像对细胞中转录物的谱分析新的和改进的方法存在需要。

[0007] 发明概述

[0008] 本发明提供与对细胞中的,尤其是单细胞中的转录物或 DNA 基因座谱分析的现有技术相关的挑战或缺点的某些洞见。此外,本发明提供用于有效地实现此类谱分析包括单细胞的此类谱分析的新技术。提供的技术是广泛有用的,包括例如用于分离的细胞、组织中的细胞、器官中的细胞和 / 或生物体中的细胞的谱分析。

[0009] 例如,本发明提供以下洞见:现有技术诸如单细胞 RNA-seq 或 qPCR,需要将单细胞分离并且将其放入多孔格式中,这会是成本高昂的、劳动密集的并且倾向产生伪结果 (prone to artifacts) 的多重步骤的方法。此外,本发明认识到,首先利用酶促反应将 mRNA 转换成 DNA 模板的现有原位测序技术可以是非常低效的 (例如在 mRNA 向 DNA 转化过程中),因此,通常,只有一小部分的 RNA 被转化和检测。本发明提供了这样的具体洞见:关于这种低效率 (对于 RT 估计在 1% 和对于 PLA 在 10%) 的一种主要缺点是它可以在基因表达测量中引入显著的噪音和 (ad) 偏倚。本发明还认识到,利用单分子荧光原位杂交 (smFISH) 的现有光谱 mRNA 编条形码技术需要不同的荧光基团用于规模放大,并且可能在可生成的条形码的数量上被限制。smFISH 还需要在杂交期间将探针分裂成编条形码的亚组。因为 smFISH 通常使用两种或更多种颜色用于靶,其在图像中产生高密度的对象,这可增加数据分析的复杂性。

[0010] 除其他方面,本发明提供了用于对例如转录物和 / 或 DNA 基因座谱分析的新技术,该新技术克服了与本发明之前的方法有关的问题的一种或多种或所有。在一些实施方案中,本发明提供用于通过允许不同靶的多路复用的顺序条形码方案在细胞中检测多个靶 (例如转录物或 DNA 基因座) 的方法。

[0011] 在一些实施方案中,本发明提供了方法,所述方法包括以下步骤:

[0012] (a) 进行第一接触步骤,所述第一接触步骤包括使包含多种核酸的细胞与第一多种可检测地标记的寡核苷酸接触,所述第一多种可检测地标记的寡核苷酸的每一种靶向核酸并且被可检测部分标记,因此组合物至少包含:

[0013] (i) 第一寡核苷酸, 所述第一寡核苷酸靶向第一核酸并且被第一可检测部分标记; 和

[0014] (ii) 第二寡核苷酸, 所述第二寡核苷酸靶向第二核酸并且被第二可检测部分标记;

[0015] (b) 在第一接触步骤之后对细胞成像, 以便检测所述第一多种的寡核苷酸与其靶的相互作用;

[0016] (c) 进行第二接触步骤, 所述第二接触步骤包括使所述细胞与第二多种可检测地标记的寡核苷酸接触, 所述第二多种包括靶向被所述第一多种靶向的重叠核酸的寡核苷酸, 因此所述第二多种至少包含:

[0017] (i) 第三寡核苷酸, 所述第三寡核苷酸任选地在序列上与所述第一寡核苷酸相同, 所述第三寡核苷酸靶向所述第一核酸; 和

[0018] (ii) 第四寡核苷酸, 所述第四寡核苷酸任选地在序列上与第二寡核苷酸相同, 所述第四寡核苷酸靶向所述第二核酸,

[0019] 其中所述第二多种与所述第一多种的不同在于, 存在于所述第二多种中的寡核苷酸的至少一种被与所述第一多种中靶向相同核酸的对应的寡核苷酸不同的可检测部分标记, 因此在所述第二多种中:

[0020] (iii) 所述第三寡核苷酸被所述第一可检测部分、所述第二可检测部分或第三可检测部分标记; 并且

[0021] (iv) 第四寡核苷酸被所述第一可检测部分、所述第二可检测部分、所述第三可检测部分或第四可检测部分标记,

[0022] 其中所述第三寡核苷酸被与所述第一寡核苷酸所用的不同的可检测部分标记, 或所述第四寡核苷酸被与所述第二寡核苷酸所用的不同的可检测部分标记, 或两者;

[0023] (d) 在所述第二接触步骤之后对细胞成像, 以便检测所述第二多种的寡核苷酸与它们的靶的相互作用; 以及

[0024] (e) 任选地所述重复接触和成像步骤, 每次用新的多种可检测地标记的寡核苷酸, 所述新的多种可检测地标记的寡核苷酸包含靶向被所述第一多种和第二多种靶向的重叠核酸的寡核苷酸, 其中由于标记靶向相同核酸的寡核苷酸的可检测部分中的至少一种差异, 所利用的每个多种与所利用的每个其他的多种不同。

[0025] 在一些实施方案中, 本发明 (例如, 如在图 1 中呈现的), 提供了方法, 所述方法包括以下步骤:

[0026] (a) 进行第一接触步骤, 所述第一接触步骤包括使包含多种转录物和 DNA 基因座的细胞与第一多种可检测地标记的寡核苷酸接触, 所述第一多种可检测地标记的寡核苷酸的每一种靶向转录物或 DNA 基因座并且被可检测部分标记, 因此组合物至少包含:

[0027] (i) 第一寡核苷酸, 所述第一寡核苷酸靶向第一转录物或 DNA 基因座并且被第一可检测部分标记; 和

[0028] (ii) 第二寡核苷酸, 所述第二寡核苷酸靶向第二转录物或 DNA 基因座并且被第二可检测部分标记;

[0029] (b) 在第一接触步骤之后对细胞成像, 以便检测所述第一多种的寡核苷酸与它们的靶的识别;

[0030] (c) 进行第二接触步骤,所述第二接触步骤包括使所述细胞与第二多种可检测地标记的寡核苷酸接触,所述第二多种包括靶向被第一多种靶向的重叠转录物和 / 或 DNA 基因座的寡核苷酸,因此所述第二多种至少包含:

[0031] (i) 第三寡核苷酸,所述第三寡核苷酸任选地在序列上与第一寡核苷酸相同,所述第三寡核苷酸靶向所述第一转录物或 DNA 基因座;和

[0032] (ii) 第四寡核苷酸,所述第四寡核苷酸任选地在序列上与第二寡核苷酸相同,所述第四寡核苷酸靶向所述第二转录物或 DNA 基因座,其中所述第二多种与第一多种的不同在于,存在于第二多种中的寡核苷酸的至少一种被与所述第一多种中靶向相同的转录物或 DNA 基因座的对应的寡核苷酸不同的可检测部分标记,因此在所述第二多种中:

[0033] (iii) 所述第三寡核苷酸被所述第一可检测部分、所述第二可检测部分或第三可检测部分标记;并且

[0034] (iv) 所述第四寡核苷酸被所述第一可检测部分、所述第二可检测部分、所述第三可检测部分或第四可检测部分标记,

[0035] 其中所述第三寡核苷酸被与所述第一寡核苷酸所用的不同的可检测部分标记,或第四寡核苷酸被与第二寡核苷酸所用的不同的可检测部分标记,或两者;

[0036] (d) 在所述第二接触步骤之后对细胞成像,以便检测检测所述第二多种的寡核苷酸与它们的靶的识别;以及

[0037] (e) 任选地重复所述接触和成像步骤,每一次用新的多种可检测地标记的寡核苷酸,所述新的多种可检测地标记的寡核苷酸包含靶向被所述第一多种和第二多种靶向的重叠转录物或 DNA 基因座的寡核苷酸,其中由于标记靶向相同转录物或 DNA 基因座的寡核苷酸的可检测部分中的至少一种差异,所利用的每个多种与所利用的每个其他的多种不同。

[0038] 在一些实施方案中,被可检测地标记的寡核苷酸靶向的核酸是转录物和 / 或 DNA 基因座,或包含转录物和 / 或 DNA 基因座。在一些实施方案中,被可检测地标记的寡核苷酸靶向的核酸是转录物,或包括转录物。在一些实施方案中,被可检测地标记的寡核苷酸靶向的核酸是转录物。在一些实施方案中,被可检测地标记的寡核苷酸靶向的核酸是 DNA 基因座,或包括 DNA 基因座。在一些实施方案中,被可检测地标记的寡核苷酸靶向的核酸是 DNA 基因座。在一些实施方案中,在接触步骤中使用的可检测地标记的寡核苷酸的每个多种靶向相同的转录物和 / 或 DNA 基因座。

[0039] 在一些实施方案中,在接触步骤中使用的多种可检测地标记的寡核苷酸称为一组可检测地标记的寡核苷酸。在一些实施方案中,一组可检测地标记的寡核苷酸的靶称为一组靶。在一些实施方案中,组中的靶是或包括转录物。在一些实施方案中,组中的靶是转录物。在一些实施方案中,组中的每一种靶是或包括转录物。在一些实施方案中,组中的每一种靶是转录物。在一些实施方案中,组中的靶是或包括 DNA 基因座。在一些实施方案中,组中的靶是 DNA 基因座。在一些实施方案中,组中的每一种靶是或包括 DNA 基因座。在一些实施方案中,组中的每一种靶是 DNA 基因座。

[0040] 在一些实施方案中,提供的方法任选地包括在成像步骤之后去除多种可检测地标记的寡核苷酸的步骤。在一些实施方案中,提供的方法包括在每个成像步骤之后去除多种可检测地标记的寡核苷酸的步骤。在一些实施方案中,去除的步骤包括使多种可检测地标记的寡核苷酸与消化可检测地标记的寡核苷酸的酶接触。在一些实施方案中,去除步骤包

括使多种可检测地标记的寡核苷酸与 DNA 酶接触。在一些实施方案中,去除步骤包括使多种可检测地标记的寡核苷酸与 RNA 酶接触。在一些实施方案中,去除步骤包括光漂白。

[0041] 在一些实施方案中,每一组包括靶向相同的转录物和 / 或 DNA 基因座的两种或更多种可检测地标记的寡核苷酸。在一些实施方案中,靶向相同的转录物和 / 或 DNA 基因座的组中的两种或更多种可检测地标记的寡核苷酸产生相同的可检测的信号。在一些实施方案中,靶向相同的转录物和 / 或 DNA 基因座的组中的所有可检测地标记的寡核苷酸产生相同的可检测的信号。在一些实施方案中,其中可检测地标记的寡核苷酸被荧光团标记,可检测的信号是某种颜色。在一些实施方案中,靶向相同的转录物和 / 或 DNA 基因座的组中的所有可检测地标记的寡核苷酸被提供相同的可检测的颜色的荧光团标记。

[0042] 在一些实施方案中,靶向相同的转录物和 / 或 DNA 基因座的组中的两种或更多种可检测地标记的寡核苷酸具有相同的可检测的标记。在一些实施方案中,靶向相同的转录物和 / 或 DNA 基因座的组中的所有可检测地标记的寡核苷酸具有相同的可检测的标记。在一些实施方案中,靶向相同的转录物和 / 或 DNA 基因座的所有可检测地标记的寡核苷酸具有相同的荧光团。

[0043] 在一些实施方案中,本发明提供可用于进行所提供的方法的组合物。

[0044] 在一些实施方案中,本发明提供组合物,所述组合物包含多种可检测地标记的寡核苷酸,所述可检测地标记的寡核苷酸的每一种靶向核酸并且被可检测的部分标记,因此所述组合物至少包含:

[0045] (i) 第一寡核苷酸,所述第一寡核苷酸靶向第一核酸并且被第一可检测部分标记;和

[0046] (ii) 第二寡核苷酸,所述第二寡核苷酸靶向第二核酸并且被第二可检测部分标记。

[0047] 在一些实施方案中,本发明提供了试剂盒,该试剂盒包含多种可检测地标记的寡核苷酸,所述可检测地标记的寡核苷酸的每一种靶向核酸并且被可检测部分标记,因此所述试剂盒至少包含:

[0048] (i) 第一寡核苷酸,所述第一寡核苷酸靶向第一核酸并且被第一可检测部分标记;

[0049] (ii) 第二寡核苷酸,所述第二寡核苷酸靶向第二核酸并且被第二可检测部分标记。

[0050] (iii) 第三寡核苷酸,所述第三寡核苷酸任选地在序列上与所述第一寡核苷酸相同,所述第三寡核苷酸靶向所述第一核酸并且被所述第一可检测部分、所述第二可检测部分或第三可检测部分标记;和

[0051] (iv) 第四寡核苷酸,所述第四寡核苷酸任选地在序列上与所述第二寡核苷酸相同,所述第四寡核苷酸靶向核酸并且被所述第一可检测部分、所述第二可检测部分、所述第三可检测部分或第四可检测部分标记,

[0052] 其中所述第三寡核苷酸被与所述第一寡核苷酸不同的可检测部分标记,或所述第四寡核苷酸被与所述第二寡核苷酸不同的可检测部分标记,或两者。

[0053] 在一些实施方案中,可检测部分是荧光团或包括荧光团。

[0054] 在一些实施方案中,多种可检测地标记的寡核苷酸靶向两种或更多种核酸

(“靶”)。在一些实施方案中,靶是或包括转录物。在一些实施方案中,靶是转录物。在一些实施方案中,靶是 RNA。在一些实施方案中,靶是 mRNA。在一些实施方案中,靶是 tRNA。在一些实施方案中,靶是 rRNA。在一些实施方案中,靶是非编码 RNA。在一些实施方案中,靶是或包括 DNA 基因座。在一些实施方案中,转录物是 DNA 基因座。在一些实施方案中,靶是转录物的基因座。在一些实施方案中, DNA 序列的不同的转录物,诸如基因的剪接变体,构成不同的靶,其中所述变体的一种或更多种可以独立地被靶向和检测或定量。在一些实施方案中,本发明提供检测个体剪接变体的方法、组合物或试剂盒。在一些实施方案中,本发明提供用于检测单核苷酸多态性 (SNP) 的方法、组合物或试剂盒。

[0055] 在一些实施方案中,提供的方法定量靶,例如,转录物或 DNA 基因座。

[0056] 在一些实施方案中,靶向相同的靶的寡核苷酸具有相同的序列组,即,当其在不同的步骤被应用时,寡核苷酸之间的差异在部分间,而不是序列间。

[0057] 定义

[0058] 动物:如本文使用的,术语“动物”是指动物界的任何成员。在一些实施方案中,“动物”指在任何发育阶段的人类。在一些实施方案中,“动物”指在任何发育阶段的非人类动物。在某些实施方案中,非人类动物是哺乳动物(例如,啮齿动物、小鼠、大鼠、兔、猴、狗、猫、绵羊、牛、灵长类和/或猪)。在一些实施方案中,动物包括,但不局限于:哺乳动物、鸟类、爬行类、两栖类、鱼类和/或蠕虫。在一些实施方案中,动物可以是转基因动物、遗传工程化的动物和/或无性系。

[0059] 大约:如本文使用的,关于数字的术语“大约”或“约”通常被采用以包括落入数字的任一方向(多于或少于)的 5%、10%、15%或 20%的范围内的数字,除非另有说明或另外从上下文文明显(除非这样的数字将小于可能值的 0%或超过 100%)。在一些实施方案中,关于剂量的术语“约”的使用意指 $\pm 5\text{mg/kg/天}$ 。

[0060] 同源性:“同源性”或“同一性”或“相似性”指的是两个核酸分子之间的序列相似性。同源性和同一性可以通过比较为了比较目的而可对齐的每个序列中的位置而被各自确定。在相比较的序列中相当的位置被相同碱基占据,则分子在那个位置上是相同的;当相当的位点被相同的或相似的核酸残基(例如,在空间和/或电子上的性质相似)占据,则分子可以被称为在那个位置上同源(相似)。表示为同源性/相似性或同一性的百分比是指比较的序列共享的位置上相同的或相似的核酸的数目的函数。“无关的”或“非同源”的序列与本文描述的序列共享少于 40%的同一性、少于 35%的同一性、少于 30%的同一性或少于 25%的同一性。在两个序列的比较中,残基(氨基酸或核酸)的不存在或额外的残基的存在也减少同一性和同源性/相似性。

[0061] 在一些实施方案中,术语“同源性”描述基于数学的序列相似性的比较,其被用于鉴定具有相似功能或基序的基因。本文描述的核酸序列可以被用作“查询序列”来针对公共数据库进行搜索,例如,来鉴定其他家族成员、相关的序列或同源物。在一些实施方案中,这种搜索可以使用 Altschul, 等人 (1990) J. Mol. Biol. 215:403-10 的 NBLAST 和 XBLAST 程序(版本 2.0)来进行。在一些实施方案中,可以用 NBLAST 程序,得分 = 100, 字长 = 12 进行 BLAST 核苷酸搜索,来获得与本发明的核酸分子同源的核苷酸序列。在一些实施方案中,为了获得用于比较目的的缺口的比对,可以利用描述于 Altschul 等人, (1997) Nucleic Acids Res. 25(17):3389-3402 中的 Gapped BLAST。当利用 BLAST 和 Gapped BLAST 程序时,可以

使用各程序（例如，XBLAST 和 BLAST）的默认参数（参见 www.ncbi.nlm.nih.gov）。

[0062] 同一性：如本文使用的，“同一性”意指当序列被比对至最大序列匹配时，即，考虑到缺口和插入，在两个或更多个序列中对应位置上的相同的核苷酸残基的百分比。可以通过已知的方法容易地计算同一性，包括但不限于在以下描述的那些（Computational Molecular Biology, Lesk, A. M., 编著, Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D. W., 编著, Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part I, Griffin, A. M., 和 Griffin, H. G., 编著, Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; 和 Sequence Analysis Primer, Gribskov, M. 和 Devereux, J., 编著, M Stockton Press, New York, 1991; 和 Carillo, H. 和 Lipman, D., SIAM J. Applied Math., 48:1073(1988)。确定同一性的方法被设计来给出被测试的序列之间的最大匹配。此外，确定同一性的方法被编入公众可得的计算机程序。确定两个序列之间的同一性的计算机程序方法包括，但不限于，GCG 程序包（Devereux, J. 等人, Nucleic Acids Research 12(1):387(1984))、BLASTP、BLASTN 和 FASTA(Altschul, S. F. 等人, J. Molec. Biol. 215:403-410(1990) 和 Altschul 等人. Nuc. Acids Res. 25:3389-3402(1997))。BLAST X 程序是从 NCBI 和其他来源公众可得的 (BLAST Manual, Altschul, S. 等人, NCBI NLM NIH Bethesda, Md. 20894; Altschul, S. 等人, J. Mol. Biol. 215:403-410(1990))。还可以使用公知的 Smith Waterman 算法来确定同一性。

[0063] 体外：如本文使用的，术语“体外”是指发生在人工环境中的事件，例如，在试管或反应容器中，在细胞培养物中，等等，而不是在生物体内（例如，动物、植物和 / 或微生物）。

[0064] 体内：如本文使用的，术语“体内”是指发生在生物体内（例如，动物、植物和 / 或微生物）的事件。

[0065] 寡核苷酸：术语“寡核苷酸”是指核苷酸单体的聚合物或寡聚物，含有核酸碱基、被修饰的核酸碱基、糖类、被修饰的糖类、磷酸键或被修饰的键的任何组合。

[0066] 本发明的寡核苷酸可以是各种长度的。在具体的实施方案中，寡核苷酸长度可以在从约 2 个核苷酸至约 200 个核苷酸的范围内。在多种相关的实施方案中，单链的、双链的和三链的寡核苷酸长度可以在从约 4 个核苷酸至约 10 个核苷酸、从约 10 个核苷酸至约 50 个核苷酸、从约 20 个核苷酸至约 50 个核苷酸、从约 15 个核苷酸至约 30 个核苷酸、从约 20 个核苷酸至约 30 个核苷酸的范围内。在一些实施方案中，寡核苷酸长度在从约 9 个核苷酸至约 39 个核苷酸的范围内。在一些实施方案中，寡核苷酸长度是至少 4 个核苷酸。在一些实施方案中，寡核苷酸长度是至少 5 个核苷酸。在一些实施方案中，寡核苷酸长度是至少 6 个核苷酸。在一些实施方案中，寡核苷酸长度是至少 7 个核苷酸。在一些实施方案中，寡核苷酸长度是至少 8 个核苷酸。在一些实施方案中，寡核苷酸长度是至少 9 个核苷酸。在一些实施方案中，寡核苷酸长度是至少 10 个核苷酸。在一些实施方案中，寡核苷酸长度是至少 11 个核苷酸。在一些实施方案中，寡核苷酸长度是至少 12 个核苷酸。在一些实施方案中，寡核苷酸长度是至少 15 个核苷酸。在一些实施方案中，寡核苷酸长度是至少 20 个核苷酸。在一些实施方案中，寡核苷酸长度是至少 25 个核苷酸。在一些实施方案中，寡核苷酸长度是至少 30 个核苷酸。在一些实施方案中，寡核苷酸是长度为至少 18 个核苷酸的互补链的双链体。在一些实施方案中，寡核苷酸是长度为至少 21 个核苷酸的互补链的双链体。

[0067] 预先确定的：预先确定的意味着是故意选择，例如，与随机地发生或实现相对。可能包含如下的某些个体寡核苷酸的组合物不是“预先确定的”组合物：它们是通过不能被控制来有意地产生特定寡核苷酸的过程偶然产生的个体寡核苷酸。在一些实施方案中，预先确定的组合物是可以被有意地重现的组合物（例如，通过重复受控的过程）。

[0068] 样品：如本文使用的，术语“样品”是指从如本文描述的感兴趣的来源获得或衍生的生物样品。在一些实施方案中，感兴趣的来源包括生物体，诸如动物或人类。在一些实施方案中，生物样品包括生物组织或生物流体。在一些实施方案中，生物样品是或包括骨髓；血液；血细胞；腹水；组织或细针穿刺活检样品；包含细胞的体液；游离核酸；痰；唾液；尿液；脑脊液、腹腔积液；胸腔积液；粪便；淋巴；妇科液；皮肤拭样；阴道拭样；口腔拭样；鼻拭样；洗涤或灌洗液诸如导管灌洗液或支气管肺泡灌洗液；抽吸物；刮涂物（scraping）；骨髓样本；组织活检样本；手术样本；粪便、其他体液、分泌物和 / 或排泄物；和 / 或其所来自的细胞等。在一些实施方案中，生物样品是或包括从个体获得的细胞。在一些实施方案中，样品是通过任何适当手段从感兴趣的来源直接获取的“原始样品”。例如，在一些实施方案中，通过选自由活检（例如，细针抽吸或组织活检）、手术、体液（例如，血液、淋巴、粪便等）的收集等组成的组的方法获取原始生物样品。在一些实施方案中，如从上下文中将明确的，术语“样品”是指通过处理（例如，通过去除原始样品的一种或更多种组分和 / 或通过添加一种或更多种试剂至）原始样品获得的制备物。例如，使用半透膜过滤。这样的“处理的样品”可包括，例如从样品提取的或通过使原始样品经历诸如 mRNA 的扩增或反转录、某些成分的分和 / 或纯化等的技术获得的核酸或蛋白质。

[0069] 受试者：如本文使用的，术语“受试者”或“测试受试者”是指为了例如实验、诊断、预防和 / 或治疗目的，依照本发明所提供的化合物或组合物被施用至的任何生物体。典型的受试者包括动物（例如，哺乳动物诸如小鼠、大鼠、兔、非人类灵长类动物、和人类；昆虫；蠕虫等）和植物。在一些实施方案中，受试者可能患有疾病（disease）、紊乱（disorder）和 / 或状况（condition）或对疾病、紊乱和 / 或状况易感。

[0070] 基本上：如本文使用的，术语“基本上”是指表现感兴趣的特征或性质的总的或接近总的范围或程度的定性条件。生物领域中普通技术人员将理解的是，生物和化学现象很少（若存在）进行完全和 / 或进行至完备或达到或避免绝对的结果。因此术语“基本上”在本文被用以捕获许多生物和 / 或化学现象中固有的完整性的潜在缺乏。

[0071] 患有：“患有”疾病、紊乱和 / 或状况的个体已被诊断为具有和 / 或展示疾病、紊乱和 / 或状况的一种或更多种症状。

[0072] 易感：对疾病、紊乱和 / 或状况“易感”的个体，是与一般公众的成员相比具有发展疾病、紊乱、和 / 或状况的较高风险的个体。在一些实施方案中，对疾病、紊乱和 / 或状况易感的个体可以未被诊断患有疾病、紊乱和 / 或状况。在一些实施方案中，对疾病、紊乱和 / 或状况易感的个体可显示疾病、紊乱和 / 或状况的症状。在一些实施方案中，对疾病、紊乱和 / 或状况易感的个体可以未表现出疾病、紊乱和 / 或状况的症状。在一些实施方案中，对疾病、紊乱和 / 或状况易感的个体将会发展疾病、紊乱和 / 或状况。在一些实施方案中，对疾病、紊乱和 / 或状况易感的个体将不会发展疾病、紊乱和 / 或状况。

[0073] 治疗：如本文使用的，术语“治疗（treat）”、“治疗（treatment）”或“治疗（treating）”是指任何用于将疾病、紊乱和 / 或状况的一种或更多种症状或特征部分地或

完全地减轻、改善、缓解、抑制、预防、延迟发作、减少严重程度和 / 或减少发生的方法。治疗可以对未表现出疾病、紊乱和或状况迹象的受试者施用。在一些实施方案中,为了例如减少发展与疾病、紊乱、和 / 或状况相关的病理学的风险的目的,治疗可以对仅表现出疾病、紊乱和或状况的早期迹象的受试者施用。

[0074] 野生型:如本文使用的,术语“野生型”具有其领域内理解的含义,指的是具有如同在自然界中发现的处于“正常”(与突变、患病、改变等相对)状态或环境中的结构和 / 或活性的实体。本领域普通技术人员将领会的是,野生型的基因和多肽通常以多种不同的形式(例如,等位基因)存在。

[0075] 附图简述

[0076] 图 1. 由本公开内容提供的方法被展示在图 1 中。

[0077] 图 2. 所提供的方法的示例性顺序编条形码。(a) 顺序编条形码的示意图。在每一轮杂交中,多种探针(例如,24种)被杂交在每一种转录物上,成像并且随后通过 DNA 酶 I 处理来剥离。可以在不同轮的杂交中使用相同的探针序列,但是探针被偶联至不同的荧光团上。(b) 来自多酵母细胞上三轮杂交的复合四色 FISH 数据。十二个基因被 2 轮杂交编码,第三轮杂交利用与杂交 1 相同的探针进行。加框的区域在每个图像的右下角被放大。显示匹配的点并且提取条形码。非共定位的点,并非意图被理论限制,可能是由于细胞中探针的非特异性结合以及错误杂交。每个条形码的数目被定量来提供单细胞中对应的转录物的丰度。(c) 示例性条形码。mRNA 1:黄-蓝-黄;mRNA 2:绿-紫-绿;mRNA 3:紫-蓝-紫;和 mRNA 4:蓝-紫-蓝。

[0078] 图 3. 顺序杂交和编条形码的示意图。(a) 顺序杂交和编条形码的示意图。(b) 细胞的 FISH 图像的示意图。在每一轮的杂交中,检测的是相同的点,但是与转录物相缔合的染料变化了。mRNA 的身份被编码在杂交的染料的时间层序(temporal sequence)中。

[0079] 图 4. DNA 酶 I 有效去除结合到 mRNA 的 smFISH 探针。DNA 酶 I 有效去除结合到 mRNA 的 smFISH 探针。在抗漂白缓冲液中 4 小时的 DNA 酶 I 处理之前和之后,将点成像。处理之后的强度比率的平均值、中位数和 STD 是 11.5%、8.3%和 11%。在 DNA 酶 I 处理之后和之前绘制每一个点的点强度的比率。 $n = 1084$ 个点。

[0080] 图 5. DNA 酶 I 处理之后光漂白去除残余强度。DNA 酶 I 处理之后光漂白去除残余强度。4 小时的 DNA 酶 I 处理之后通过 10 秒钟的激发来漂白点。漂白之后的强度比率的平均值、中位数和 STD 是 0.03%、0.01%和 0.049%。在 DNA 酶 I 处理之后和之前绘制每一个点的点强度的比率。 $n = 1286$ 个点。

[0081] 图 6. 经过多轮再杂交,mRNA 是稳定的。经过多轮再杂交,mRNA 是稳定的。经过 6 次杂交绘制 smRISH 点的强度分布。两次杂交被重复 3 次以产生 6 次总的杂交。通过与下一个相同的杂交中的点的共定位来鉴定点。对于每一个箱形图,计数的点的数目是在 191 和 1337 之间。

[0082] 图 7. 每个细胞中,从前两轮的杂交中鉴定的条形码在随后一轮的杂交中重新出现的比例。每个细胞中,从前两轮的杂交中鉴定的条形码在随后一轮的杂交中重新出现的比例。经过所有三次杂交通过共定位鉴定条形码。77.9±5.6%的条形码重新出现。 $n = 37$ 个细胞。

[0083] 图 8. 在杂交 1 和杂交 3 中 FISH 点之间的逐点置换。在杂交 1 和杂交 3 中 FISH

点之间的逐点置换。在杂交 1 和杂交 3 中的 Cy5 图像中的 FISH 点被提取,用 2D 高斯 (2D Gaussians) 拟合。逐点的置换被示于 3D 直方图中。标准差是 105.8nm,表明,mRNA 可以在两轮杂交之间定位至 100nm。n = 1199 个点。

[0084] 图 9. 在相同的探针组的重复杂交之间鉴定的条形码 (杂交 1 和杂交 3)。在相同的探针组的重复杂交之间鉴定的条形码 (杂交 1 和杂交 3)。通过杂交之间的共定位鉴定条形码。每一列与一个个体细胞对应。每一行与在杂交 1 和杂交 3 之间鉴定的特定的条形码对应。加粗的行名称对应于重复的颜色条形码,其应该在杂交 1 和杂交 3 之间共定位。未加粗的行名称对应于假阳性条形码。例如,对 (Alexa 532, Alexa 532) 检测到了大量的条形码,表明在 Alexa 532 通道中点的共定位, n = 37 个细胞。A1532 = Alexa 532。A1594 = Alexa 594。A1647 = Alexa 647。

[0085] 图 10. 来自条形码提取的单细胞 mRNA 水平。来自条形码提取的单细胞 mRNA 水平。通过杂交 1 和杂交 2 之间的共定位鉴定条形码。每一列与个体细胞对应。n = 37 个细胞。A1532 = Alexa 532。A1594 = Alexa 594。A1647 = Alexa 647。从顶部到底部 :YLR194c、CMK2、GYP7、PMC1、NPT1、SOK2、UIP3、RCN2、DOA1、HSP30、PUN1 和 YPS1。

[0086] 图 11. 在小鼠胚胎干细胞 (mESC) 中 Nanog Alexa 647 探针的 DNA 酶 I 剥离。在小鼠胚胎干细胞 (mESC) 中 Nanog Alexa 647 探针的 DNA 酶 I 剥离。在 mESC 中靶向 Nanog 的四十八个探针被杂交。通过以 3 单位 / μ L 浓度的 DNA 酶 I 孵育 30 分钟将探针剥离。

[0087] 图 12. 在小鼠胚胎干细胞 (mESC) 中 Nanog mRNA 的再杂交。在小鼠胚胎干细胞 (mESC) 中 Nanog mRNA 的再杂交。通过以 3 单位 / μ L 浓度的 DNA 酶 I 孵育 30 分钟将探针剥离。Nanog Alexa 647 探针被再杂交持续 12 小时并且将其成像。图像是产生自每 1.5 μ m 拍照一次的 11 个图像的 z 向堆叠的 2D 最大投射。

[0088] 图 13. 在皮质中的 β -肌动蛋白 (红色) 的 HCR 检测和在 100 μ m 冠状切片中用逆行示踪物 (荧光金,绿) 的可视化。完整的冠状切片在 10X 和 60X (放大的插图) 两种情况下成像。在 60X 的图像中,个体红点对应于单个 β -肌动蛋白 mRNA 分子。 β -肌动蛋白的表达可以通过计数荧光焦点来定量,同时检测被逆行示踪物 (绿色) 标记的神经元的末梢亚群。

[0089] 图 14. 用 HCR 标记的可检测标记的寡核苷酸与用荧光团直接标记的 smFISH 探针检测 20 μ m 脑切片中的 RNA 的单分子方面一样特异。HCR 探针 (左) 和 smFISH 探针 (右) 同时靶向 β -肌动蛋白并且共定位。注意 HCR 的改善的 S/N。

[0090] 图 15. 用 HCR 标记的可检测标记的寡核苷酸在 20 μ m 脑切片中良好地再杂交。HCR 点在 *hyb1* 和 *hyb2* 中共定位,在两个杂交之间 DNA 酶处理。这显示 HCR 能与 seqFISH 方案完全整合。

[0091] 图 16. CLARITY 与尼氏 :Thy-1-eYFP 小鼠大脑的 1mm 厚的冠状切片 (前囟 AP, 2.3-1.3mm) 被清晰化并用 NeuroTrace 荧光尼氏染剂 (1:100 稀释,48 小时,RT) 染色。左,运动皮质的 3d 冠状渲染。右,第五层运动皮质的 100 μ m-厚的截面。箭头指示锥体神经元的顶端树突 (红-尼氏荧光,绿-eYFP)。

[0092] 图 17. 一种示例性的光片显微镜的示意图。

[0093] 图 18. SPIM 检测在 100 μ m CLARITY 脑切片中的单个 mRNA。切片被跨越 100 μ m 扫描。图像随后被配准和固定 (stitch) 至三维重建。图像中的衍射极限点与被 HCR 检测的

单个 β -肌动蛋白 mRNA 对应。比例尺是 $15\ \mu\text{m}$ 。

[0094] 图 19. Allen Reference Atlas 的连续冠状水平上的皮层躯体感觉运动 (cortical somatic sensorimotor) 子网络的连接图。其显示, 躯体感觉运动区的四个主要组成部分 (SSp、SSs、MOp 和 MOs) 的每个, 被分为 4 个功能域。这些功能相关的域与所有其他的广泛互联并且形成四个主要皮层躯体感觉运动子网络: 口面咽 (orofaciopharyngeal) (orf, 蓝色)、上肢 (ul, 绿色)、下肢和躯干 (ll/tr, 红色) 和须 (bfd. cm&w, 黄色)。数字说明切片相对前囟的位置 (mm)。所提供的方法可以表征在不同的子网络内这些不同的域的每一个中的投射神经元的连通性和分子身份。

[0095] 图 20. 用于自动地检测和绘制逆行标记的神经元和基因条形码信息的示例性信息学工作流程。A. 用 CTb 标记 (粉红色) 和尼氏染色 (蓝色) 的原始图像。加框的区域显示 CTb 标记的神经元的特写镜头。B. 多通道原始图像被转换成灰度图用于分割。C. 通过分割管道运行个体示踪剂通道图像, 所述分割管道离散地分离组织背景与标记的细胞。白点是标记的胞体 (somata) 的重构。D. 重新集成的多画面 tiff 与用于配准的 ARA 中冠状切片相关。E. 使用所提供的开发的配准软件, 多重画面 tiff 被扭曲以将组织的轮廓和细胞结构描述两者与其对应的 ARA 水平对齐。F. 通过分割过程提取的细胞被空间上标准化并且可以与在 ARA 内的层-特异性的或亚-核-特异性的 ROI (sub-nucleus-specific) 相关联。G. 分割的和配准的标记重建, 以及其伴随的 seqFISH 数据, 在 iConnectome FISH 查看器上公众可得。分析选项卡提供了关于注射部位、示踪剂类型、由 ROI 标记的细胞的数量信息, 这可以进一步消除层特异性的细胞计数和细胞的基因表达的歧义。

[0096] 图 21. 利用核酸外切酶 III (Exo III) 的杂交链式反应 (HCR) 再杂交。(a) 桥接链和 HCR 聚合物的 exoIII 消化的图示。ExoIII 以 3' 到 5' 方向消化桥接链和 HCR 聚合物为 dNMP, 留下中间探针链结合至靶 (例如, mRNA) 的。可以随后将新的桥接链杂交至具有不同的起始序列的靶结合探针, 所述起始序列起始不同发夹组与不同的荧光染料的聚合。(b) 阐释 (a) 中显示的示意图在 T3T 小鼠成纤维细胞系中使用抗 β -肌动蛋白 (Actb) 转录物的探针的应用的原始数据。

[0097] 图 22. 使用 Lambda 核酸外切酶 (λ -exo) 的杂交链式反应 (HCR) 再杂交。(a) 桥接链的 λ -exo 消化的图示。 λ -exo 以 5' 至 3' 的方向选择性消化 5' 磷酸化的桥接链, 从结合至靶 (例如, mRNA) 的中间探针释放 HCR 聚合物。释放的聚合物被洗涤缓冲液洗掉。可以随后将新的桥接链杂交至具有不同的起始序列的靶结合探针, 所述起始序列起始不同发夹组与不同的荧光染料的聚合。(b) 阐释 (a) 中显示的示意图在 T3T 小鼠成纤维细胞系中使用抗 β -肌动蛋白 (Actb) 转录物的探针的应用的原始数据。

[0098] 图 23. 使用尿嘧啶-特异性的切除试剂 (USER) 的杂交链式反应 (HCR) 再杂交。(a) 桥接链的 USER 消化的图示。USER 选择性消化在桥接链中的脱氧尿苷核苷酸, 导致桥接链变得片段化。片段随后从中间探针链上解离, 从结合至靶 (例如, mRNA) 的探针释放 HCR 聚合物。释放的聚合物被洗涤缓冲液洗掉。可以随后将新的桥接链杂交至具有不同的起始序列的靶结合探针, 所述起始序列起始不同发夹组与不同的荧光染料的聚合。(b) 阐释 (a) 中显示的示意图在 T3T 小鼠成纤维细胞系中使用抗 β -肌动蛋白 (Actb) 转录物的探针的应用的原始数据。

[0099] 图 24. 使用互补的寡核苷酸 (cTOE) 的可检测标记的寡核苷酸的示例性去除。

[0100] 图 25. 示例性寡核苷酸制备。原始寡核苷酸（如在本图中示范的，探针）文库包含若干个探针亚文库。每个亚文库具有特定组的引物，其可以被用于利用 PCR 扩增亚文库。当期望的亚文库被扩增，将产物与切口酶一起孵育。酶在探针链上的识别位点切割磷酸二酯键。将所得产物变性，并且使其在变性凝胶上运行，允许所需的探针序列被释放。可以随后从凝胶切掉探针条带并且将其提取。所提取的产物可用于杂交。

[0101] 图 26. 示例性寡核苷酸制备。产物是凝胶上的第三个条带。该文库具有与其相关的许多不同的引物，一个引物组用于每个亚组的探针。示例性引物是随机的 20 个核苷酸序列，具有 45-55% 的 GC 含量和大约 55°C 的 T_m。切口核酸内切酶位点是 GTCTCNN；对应的缺口核酸内切酶是 Nt. BsmAI。产物探针是与具有 45-55% 的 GC 含量的感兴趣的 mRNA 互补的 20 聚体 DNA 序列。

[0102] 某些实施方案的详细描述

[0103] 除了其他以外，本发明提供了用于细胞中的核酸（例如，转录物和 / 或 DNA 基因座）谱分析的新的方法、组合物和 / 或试剂盒。

[0104] 在一些实施方案中，本发明提供了用于细胞中的核酸（例如，转录物和 / 或 DNA 基因座）谱分析的方法。在一些实施方案中，所提供的方法对单一细胞中的多个靶进行谱分析。所提供的方法可以，除了其他以外，通过顺序编条形码用有限数目的可检测标记对大量靶（转录物、DNA 基因座或其组合）谱分析。

[0105] 图 1 示出了根据本发明的方法。如所示出的，本发明提供了这样的方法：其中用标记的探针的多轮杂交（接触步骤）对细胞中存在的核酸（例如，mRNA）谱分析。特别地，如图 1 中所示出的，提供了与细胞中的核酸靶杂交的探针的组，其中在单个组内标记探针（即，与不同靶杂交的可检测地标记的寡核苷酸）并且，此外，在不同的组中至少一个探针被不同地标记。

[0106] 在一些实施方案中，本发明（例如，如在图 1 中所展示的），提供了包括以下步骤的方法：

[0107] (a) 进行第一接触步骤，所述第一接触步骤包括使包含多种转录物和 DNA 基因座的细胞与第一多种可检测地标记的寡核苷酸接触，所述第一多种可检测地标记的寡核苷酸的每种靶向转录物或 DNA 基因座并且被可检测部分标记，因此组合物至少包含：

[0108] (i) 第一寡核苷酸，所述第一寡核苷酸靶向第一转录物或 DNA 基因座并且被第一可检测部分标记；和

[0109] (ii) 第二寡核苷酸，所述第二寡核苷酸靶向第二转录物或 DNA 基因座并且被第二可检测部分标记；

[0110] (b) 在第一接触步骤之后对细胞成像，检测所述第一多种的寡核苷酸与其靶的杂交；

[0111] (c) 进行第二接触步骤，所述第二接触步骤包括使所述细胞与第二多种可检测地标记的寡核苷酸接触，所述第二多种包含靶向被所述第一多种靶向的重叠转录物和 / 或 DNA 基因座的寡核苷酸，因此所述第二多种至少包含：

[0112] (i) 第三寡核苷酸，所述第三寡核苷酸任选地在序列上与所述第一寡核苷酸相同，所述第三寡核苷酸任靶向所述第一转录物或 DNA 基因座；和

[0113] (ii) 第四寡核苷酸，所述第四寡核苷酸任选地在序列上与所述第二寡核苷酸相

同,所述第四寡核苷酸靶向所述第二转录物或 DNA 基因座,其中所述第二多种与第一多种的不同在于存在于所述第二多种中的寡核苷酸的至少一种被与所述第一多种中靶向相同转录物或 DNA 基因座的对应寡核苷酸不同的可检测部分标记,因此在所述第二多种中:

[0114] (iii) 所述第三寡核苷酸被所述第一可检测部分、所述第二可检测部分或第三可检测部分标记;并且

[0115] (iv) 所述第四寡核苷酸被所述第一可检测部分、所述第二可检测部分、所述第三可检测部分或第四可检测部分标记,其中所述第三寡核苷酸被与第一寡核苷酸所用的不同的可检测部分标记,或第四寡核苷酸被与第二寡核苷酸所用的不同的可检测部分标记,或两者;

[0116] (d) 在所述第二接触步骤之后对细胞成像,以便检测所述第二多种的寡核苷酸与其靶的杂交;以及

[0117] (e) 任选地重复所述接触和成像步骤,每次用新的多种可检测地标记的寡核苷酸,所述新的多种可检测地标记的寡核苷酸包含靶向被所述第一多种和第二多种靶向的重叠转录物或 DNA 基因座的寡核苷酸,其中由于标记靶向相同转录物或 DNA 基因座的核苷酸的可检测部分上的至少一种差异,所利用的每个多种与所利用的每个其他的多种不同。

[0118] 如本文使用的,可检测地标记的寡核苷酸被可检测部分标记。在一些实施方案中,可检测地标记的寡核苷酸包含一个可检测部分。在一些实施方案中,可检测地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,可检测地标记的寡核苷酸具有一个可检测部分。在一些实施方案中,可检测地标记的寡核苷酸具有两个或更多个可检测部分。

[0119] 在一些实施方案中,可检测部分是荧光团或包含荧光团。标记有荧光团、的示例性可检测地标记的寡核苷酸包括但不限于用于荧光原位杂交 (FISH) 的探针。如本领域普通技术人员所广泛已知并实践的,FISH 被用于,除了其他以外,检测和定位特定 DNA 序列或 RNA 靶的存在或不存在。用于设计和制备经标记的可检测地标记的寡核苷酸的方法在本领域是广泛地已知的,包括但不限于在美国专利申请公布 US20120142014 中所描述的那些方法。但是,由于诸如荧光团可得性的限制,FISH 仅能够被用于对给定的实验中的有限数目的靶谱分析。通过对多种不同的靶顺序编条形码,本发明所提供的方法能够对大量靶、多达 F^N 个,谱分析,其中 F 是可检测部分(在 FISH 的情况中,荧光团)的类型的数目,并且 N 是接触步骤(在 FISH 的情况中,杂交)的数目。例如,当 F 是 4 且 N 是 8 时,几乎整个转录组 ($4^8 = 65,536$) 可被谱分析。在一些实施方案中, F 至少是 2。在一些实施方案中, F 是 3。在一些实施方案中, F 是 4。在一些实施方案中, F 是 5。在一些实施方案中, F 是 6。在一些实施方案中, F 是 7。在一些实施方案中, F 是 8。在一些实施方案中, F 是 9。在一些实施方案中, F 是 10。在一些实施方案中, F 是 11。在一些实施方案中, F 是 12。在一些实施方案中, F 是 13。在一些实施方案中, F 是 14。在一些实施方案中, F 是 15。在一些实施方案中, F 大于 15。在一些实施方案中, N 是 2。在一些实施方案中, N 大于 2。在一些实施方案中, N 是 3。在一些实施方案中, N 大于 3。在一些实施方案中, N 是 4。在一些实施方案中, N 大于 4。在一些实施方案中, N 是 5。在一些实施方案中, N 大于 5。在一些实施方案中, N 是 6。在一些实施方案中, N 大于 6。在一些实施方案中, N 是 7。在一些实施方案中, N 大于 7。在一些实施方案中, N 是 8。在一些实施方案中, N 大于 8。在一些实施方案中, N

是 9。在一些实施方案中, N 大于 9。在一些实施方案中, N 是 10。在一些实施方案中, N 大于 10。在一些实施方案中, 多种可检测地标记的寡核苷酸靶向至少 100 个靶。

[0120] 在接触步骤中, 可以在可检测地标记的寡核苷酸与其靶结合之前、同时或之后对其标记。在一些实施方案中, 可检测地标记的寡核苷酸, 诸如标记荧光团的寡核苷酸, 在其与其靶结合之前被标记。在一些实施方案中, 可检测地标记的寡核苷酸, 在其与其靶结合的同时被标记。在一些实施方案中, 可检测地标记的寡核苷酸, 在其与其靶结合之后被标记。在一些实施方案中, 可检测地标记的寡核苷酸在杂交之后通过用杂交链式反应 (HCR) 正交扩增而被标记 (Choi, HM., Nat Biotechnol. 201011 月 ;28(11):1208-12)。在一些实施方案中, 可检测地标记的寡核苷酸包含部分, 例如核酸序列, 能够在成像步骤中提供信号的该一个或更多部分可以直接或间接地与寡核苷酸连接。

[0121] 在一些实施方案中, 可以将相同类型的标签附连到针对不同靶的不同探针。在一些实施方案中, 在接触步骤中使用的多种可检测地标记的寡核苷酸中, 针对相同靶的探针具有相同的标签 (一组可检测地标记的寡核苷酸)。在数轮接触和成像后, 每个靶具有其自身独特的标签组合 (顺序编条形码), 因此可以获得靶的信息, 例如, 数量信息和 / 或空间信息。例如, 当荧光团被用于标记可检测地标记的寡核苷酸时, N 个步骤后, 靶会具有顺序条形码 $F_1F_2\cdots F_N$, 其中 F_n 是在第 n 次成像中用于靶的荧光团的颜色。由于它们的条形码的不同 (例如, 红红蓝红与红红红蓝比较), 可将一个靶与另一个区分开。

[0122] 在一些实施方案中, 本发明的标签是或包含一种或更多种荧光染料, 包括但不限于荧光素、罗丹明、Alexa Fluor 类、Dylight Fluor 类、ATTO 类染料或其任何类似物或衍生物。

[0123] 在一些实施方案中, 本发明的标签包括但不限于荧光素和荧光素的化学衍生物; 曙红; 羧基荧光素; 异硫氰酸荧光素 (FITC); 荧光素酰胺 (Fluorescein amidite) (FAM); 赤藓红; 玫瑰红; 从细菌铜绿假单胞菌 (*Pseudomonas aeruginosa*) 分泌的荧光素; 亚甲基蓝; 激光染料、罗丹明类染料 (例如, 罗丹明、罗丹明 6G、罗丹明 B、罗丹明 123、金胺 0、磺酰罗丹明 101、磺酰罗丹明 B 和德克萨斯红)。

[0124] 在一些实施方案中, 本发明的标签包括但不限于 ATTO 类染料; 吖啶类染料 (例如, 吖啶橙、吖啶黄); Alexa Fluor; 7-氨基放线菌素 D; 8-苯胺基萘-1-磺酸盐; 金胺-罗丹明染料; 苯并葱酮; 5, 12-双(苯乙炔基)并四苯; 9, 10-双(苯乙炔基)葱; Blacklight paint; 脑虹 (Brainbow); 钙黄绿素; 羧基荧光素; 羧基二乙酸荧光素琥珀酰亚胺酯; 羧基荧光素琥珀酰亚胺酯; 1-氯-9, 10-双(苯乙炔基)葱; 2-氯-9, 10-双(苯乙炔基)葱; 2-氯-9, 10-二苯基葱; 香豆素; 花菁类染料 (例如, 花菁诸如 Cy3 和 Cy5、DiOC6、SYBR Green I); DAPI、Dark quencher、Dylight Fluor、Fluo-4、FluoProbes; 荧光酮类染料 (例如, 钙黄绿素、羧基荧光素、羧基二乙酸荧光素琥珀酰亚胺酯、羧基荧光素琥珀酰亚胺酯、曙红、曙红 B、曙红 Y、赤藓红、荧光素、异硫氰酸荧光素、荧光素酰胺、印度黄、汞溴红); Fluoro-Jade 染料; Fura-2; Fura-2-乙酰氧基甲酯; 绿色荧光蛋白、Hoechst 染料、印度黄、Indo-1、荧光黄、萤光素、部花菁、光增白剂、噁嗪类染料 (例如, 甲酚紫、尼罗蓝、尼罗红); 花; 菲啶类染料 (溴化乙锭和碘化丙啶); 焰红染料、藻胆素、藻红蛋白、藻红素、8-羟基-1, 3, 6-三磺酸茚 (Pyranine)、罗丹明、罗丹明 123、罗丹明 6G、RiboGreen、RoGFP、红荧烯、SYBR Green I、(E)-二苯乙烯、(Z)-二苯乙烯、磺酰罗丹明 101、磺酰罗丹明 B、Synapto-pHluorin、四苯

基丁二烯、四钠三(红菲绕啉二磺酸)钌(II)(Tetrasodium tris(bathophenanthroline disulfonate)ruthenium(II))、德克萨斯红、TSQ、伞形花内酯或黄色荧光蛋白。

[0125] 在一些实施方案中,本发明的标签包括但不限于荧光染料的 Alexa Fluor 家族(Molecular Probes, Oregon)。在荧光显微术和细胞生物学中, Alexa Fluor 类染料被广泛地用作细胞和组织标记。 Alexa Fluor 系列的激发和发射光谱覆盖可见光谱并且延伸至红外光谱。大致地根据该家族的个体成员的最大激发(以 nm)将其编号。某些 Alexa Fluor 类染料通过香豆素、罗丹明、咕吨(诸如荧光素)和花菁类染料的磺化合成。在一些实施方案中,磺化使 Alexa Fluor 类染料带负电荷并且具亲水性。在一些实施方案中, Alexa Fluor 类染料比具有相当的激发和发射的普通染料(例如,荧光素、罗丹明)更稳定、更明亮并且对 pH 更不敏感,并且在某种程度上来讲是较新的花菁系列。示例性 Alexa Fluor 类染料包括但不限于 Alexa-350、Alexa-405、Alexa-430、Alexa-488、Alexa-500、Alexa-514、Alexa-532、Alexa-546、Alexa-555、Alexa-568、Alexa-594、Alexa-610、Alexa-633、Alexa-647、Alexa-660、Alexa-680、Alexa-700 或 Alexa-750。

[0126] 在一些实施方案中,本发明的标记包括荧光类染料的 DyLight Fluor 家族的一种或更多种(Dyomics and Thermo Fisher Scientific)。示例性 Alexa Fluor 家族染料包括但不限于 DyLight-350、DyLight-405、DyLight-488、DyLight-549、DyLight-594、DyLight-633、DyLight-649、DyLight-680、DyLight-750 或 DyLight-800。

[0127] 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。

[0128] 在一些实施方案中,本领域技术人员理解在特定的循环中特定探针的标记的选择可以基于多种因素来确定,所述多种因素包括,例如,大小、所产生的信号的类型、附连至或并入探针的方式、细胞组分的性质包括细胞组分在细胞内的位置、细胞的性质、被分析的相互作用的类型等等。

[0129] 例如,在一些实施方案中,探针用 Cy3 或 Cy5 标记, Cy3 或 Cy5 被合成为携带 N-羧基琥珀酰亚胺酯(NHS-酯)反应性基团。由于 NHS-酯容易地与脂肪胺基团反应,所以核苷酸可以用氨基烷基修饰。这可以通过在合成反应期间并入氨基烷基修饰的核苷酸进行。在一些实施方案中,每 60 个碱基使用一个标签以避免猝灭作用。

[0130] 可检测地标记的寡核苷酸可以与靶杂交,所述靶例如转录物或 DNA 基因座。在一些实施方案中,靶是或包含转录物。在一些实施方案中,靶是转录物。在一些实施方案中,转录物是 RNA。在一些实施方案中,转录物是 mRNA。在一些实施方案中,转录物是 tRNA。在一些实施方案中,转录物是 rRNA。在一些实施方案中,转录物是 snRNA。在一些实施方案中, RNA 是非编码 RNA。示例性非编码 RNA 类型在本领域是广泛已知的,包括但不限于长非编码 RNA(lncRNA)、微小 RNA(miRNA)、短干扰 RNA(siRNA)、piwi 相互作用 RNA(piRNA)、核仁小 RNA(snoRNA) 和其他短 RNA。在一些实施方案中, RNA 是 lncRNA。在一些实施方案中, RNA 是 miRNA。在一些实施方案中, RNA 是 piRNA。在一些实施方案中, RNA 是 snoRNA。

[0131] 在一些实施方案中,靶是或包含 DNA 基因座。在一些实施方案中,当靶是 DNA 基因

座时,可检测地标记的寡核苷酸任选地包含一个或更多个 RNA 核苷酸或 RNA 区段。可检测地标记的寡核苷酸包含在成像之后可以被选择性地去除而不降解 DNA 靶的 RNA 序列,例如,通过 RNA 特异性酶消化去除。特异性降解 RNA 而不降解 DNA 的示例性酶包括但不局限于多种 RNA 酶,诸如 RNA 酶 A 和 RNA 酶 H。

[0132] 在一些实施方案中,可检测地标记的寡核苷酸与其靶例如转录物或 DNA 基因座直接地杂交。在一些实施方案中,可检测地标记的寡核苷酸通过与一个或更多个中间物结合或杂交来与其靶特异性相互作用(特异性识别其靶),所述一个或更多个中间物(例如寡核苷酸)与靶结合、杂交或以其他方式特异性地连接。在一些实施方案中,中间寡核苷酸以突出端与其靶杂交,因此具有互补序列的第二寡核苷酸(“桥接寡核苷酸”或“桥接探针”)可以与所述中间寡核苷酸结合。在一些实施方案中,中间物靶向核酸并且任选地被可检测部分标记,并且在与靶杂交之后包含突出端序列。在一些实施方案中,中间物包含与靶杂交的序列、突出端序列、和任选地可检测部分。在一些实施方案中,中间物包含与靶杂交的序列和突出端序列。在一些实施方案中,中间物不具有可检测部分。在一些实施方案中,第二寡核苷酸是可检测地标记的寡核苷酸。在一些实施方案中,第二可检测地标记的寡核苷酸被染料标记。在一些实施方案中,可检测地标记的寡核苷酸被 HCR 聚合物标记。在一些实施方案中,贯穿多次接触、去除和/或成像步骤,中间寡核苷酸保持与靶结合;在接触和成像步骤中通过可检测标签的组合来提供顺序条形码,所述可检测标签通过桥接探针与中间寡核苷酸连接。例如,当可检测地标记的寡核苷酸被用作桥接探针时,通过可检测地标记的寡核苷酸提供条形码,所述可检测地标记的寡核苷酸通过它们的突出端序列与中间寡核苷酸杂交。在成像步骤之后,桥接寡核苷酸被任选地如本文描述的去除。在一些实施方案中,对于一个靶采用 1 个中间寡核苷酸。在一些实施方案中,对于一个靶采用 2 个或更多个中间寡核苷酸。在一些实施方案中,对于一个靶采用 3 个或更多个中间寡核苷酸。在一些实施方案中,对于一个靶采用 4 个或更多个中间寡核苷酸。在一些实施方案中,对于一个靶采用 5 个或更多个中间寡核苷酸。在一些实施方案中,对于一个靶采用 6 个或更多个中间寡核苷酸。在一些实施方案中,对于一个靶采用 7 个或更多个中间寡核苷酸。在一些实施方案中,对于一个靶采用 8 个或更多个中间寡核苷酸。在一些实施方案中,对于一个靶采用 9 个或更多个中间寡核苷酸。在一些实施方案中,对于一个靶采用 10 个或更多个中间寡核苷酸。在一些实施方案中,对于一个靶采用 11 个或更多个中间寡核苷酸。在一些实施方案中,对于一个靶采用 12 个或更多个中间寡核苷酸。在一些实施方案中,对于一个靶采用 13 个或更多个中间寡核苷酸。在一些实施方案中,对于一个靶采用 13 个或更多个中间寡核苷酸。在一些实施方案中,对于一个靶采用 15 个或更多个中间寡核苷酸。在一些实施方案中,对于一个靶采用 16 个或更多个中间寡核苷酸。在一些实施方案中,对于一个靶采用 17 个或更多个中间寡核苷酸。在一些实施方案中,对于一个靶采用 18 个或更多个中间寡核苷酸。在一些实施方案中,对于一个靶采用 19 个或更多个中间寡核苷酸。在一些实施方案中,对于一个靶采用 20 个或更多个中间寡核苷酸。在一些实施方案中,对于一个靶采用 21 个或更多个中间寡核苷酸。在一些实施方案中,对于一个靶采用 22 个或更多个中间寡核苷酸。在一些实施方案中,对于一个靶采用 23 个或更多个中间寡核苷酸。在一些实施方案中,对于一个靶采用 24 个或更多个中间寡核苷酸。在一些实施方案中,对于一个靶采用 25 个或更多个中间寡核苷酸。在一些实施方案中,对于一个靶采用 30 个或更多个中间寡核苷酸。

在一些实施方案中,对于一个靶采用 40 个或更多个中间寡核苷酸。在一些实施方案中,对于一个靶采用 50 个或更多个中间寡核苷酸。

[0133] 在一些实施方案中,每种中间寡核苷酸与靶的一个不同序列杂交。在一些实施方案中,靶的每种中间寡核苷酸包含相同的突出端序列。在一些实施方案中,针对靶的每种可检测地标记的寡核苷酸包含与靶的所有中间寡核苷酸共有的相同突出端序列互补的相同的序列。在一些实施方案中,中间寡核苷酸包含与靶互补的序列和与可检测地标记的寡核苷酸互补的序列。

[0134] 在一些实施方案中,所提供的方法还包括以下步骤:

[0135] (f) 进行接触步骤,所述接触步骤包括使包含多种核酸的细胞与多种中间寡核苷酸接触,所述多种中间寡核苷酸的每种:

[0136] (i) 靶向核酸并任选地被可检测部分标记;并且

[0137] (ii) 在与靶杂交后包含突出端序列;以及

[0138] (g) 任选地对细胞成像,以便检测中间寡核苷酸与其靶之间的相互作用。

[0139] 在一些实施方案中,步骤 (f) 和任选步骤 (g) 在步骤 (a) 之前进行。在一些实施方案中,步骤 (f) 是被进行的步骤 (a)。在一些实施方案中,去除步骤保留中间寡核苷酸。

[0140] 在一些实施方案中,所提供的技术被用于对由于剪接变异、RNA 编辑、寡核苷酸修饰或其组合形成的不同转录物谱分析。在一些实施方案中,靶是 RNA 剪接变体。在一些实施方案中,所提供的技术对基因的一种或更多种剪接变体例如基因的一种或更多种剪接变体的位置和数量谱分析。在一些实施方案中,所提供的方法或组合物对不同的剪接变体谱分析。在一些实施方案中,包含一种或更多种变体的外显子通过顺序杂交和编条形码而被靶向并编条形码。在一些实施方案中,剪接变体包含由剪接得到的一个或更多个可区分的序列,并且此类序列被靶向。在一些实施方案中,通过靶向外显子和 / 或可区分的序列,所提供的技术能够对一种或更多种特定的剪接变体或 mRNA 的整个剪接库 (splicing repertoire) 谱分析。如本领域所广泛已知的,mRNA 剪接对许多生物过程和疾病是重要的,例如,神经系统疾病,如自闭症或唐氏综合征。负责细胞间粘附和突触发生的分子被剪接并且已知在大脑中它们的缺损产生错误传送 (miswiring) 并导致疾病。

[0141] 在一些实施方案中,可检测地标记的寡核苷酸靶向由序列编辑、化学修饰和 / 或其组合导致的序列修饰。在一些实施方案中,任选地在转化过程后,修饰的核酸靶与未修饰的靶相比与一种或更多种不同的互补序列杂交,并且利用选择性与修饰的核酸杂交的一种或更多种寡核苷酸而被谱分析。在一些实施方案中,靶是通过 RNA 编辑的 RNA (Brennicke, A., A. Marchfelder, 等人 . (1999). "RNA editing". FEMS Microbiol Rev 23 (3) :297-316)。在一些实施方案种,所提供的技术对通过 RNA 编辑形成的不同 RNA 变体谱分析。在一些实施方案种,所述提供的技术对修饰的寡核苷酸谱分析。在一些实施方案中,所提供的技术对甲基化的 RNA 谱分析 (Song CX, Yi C, He C. Mapping recently identified nucleotide variants in the genome and transcriptome. Nat Biotechnol. 2012 11 月; 30 (1 1) :1107-16)。在某些实施方案中,所提供的技术对甲基化的 DNA 谱分析。在一些实施方案中,靶是单核苷酸多态性 (SNP)。

[0142] 在一些实施方案中,通过对靶谱分析,所提供的技术,除了其他以外,提供在一些情况中在单一细胞、组织、器官或生物体中的靶的数量和 / 或位置的信息。在一些实施方案

中,转录物的谱分析可被用来定性地和 / 或定量地界定细胞、组织、器官或生物体内基因表达的时空模式。

[0143] 在一些实施方案中,一组中的每种可检测地标记的寡核苷酸具有不同的靶,例如,转录物或 DNA 基因座。在一些实施方案中,一组中的两种或更多种可检测地标记的寡核苷酸具有相同的靶。在一些实施方案中,两种或更多种可检测地标记的寡核苷酸靶向相同的转录物。在一些实施方案中,两种或更多种可检测地标记的寡核苷酸靶向相同的 DNA 基因座。在一些实施方案中,约 2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、40、50、60、70、80、90 或 100 种可检测地标记的寡核苷酸靶向相同的靶。在一些实施方案中,两种或更多种可检测地标记的寡核苷酸靶向相同的靶。在一些实施方案中,5 种或更多种可检测地标记的寡核苷酸靶向相同的靶。在一些实施方案中,10 种或更多种可检测地标记的寡核苷酸靶向相同的靶。在一些实施方案中,15 种或更多种可检测地标记的寡核苷酸靶向相同的靶。在一些实施方案中,20 种或更多种可检测地标记的寡核苷酸靶向相同的靶。在一些实施方案中,25 种或更多种可检测地标记的寡核苷酸靶向相同的靶。在一些实施方案中,30 种或更多种可检测地标记的寡核苷酸靶向相同的靶。在一些实施方案中,35 种或更多种可检测地标记的寡核苷酸靶向相同的靶。在一些实施方案中,40 种或更多种可检测地标记的寡核苷酸靶向相同的靶。在一些实施方案中,45 种或更多种可检测地标记的寡核苷酸靶向相同的靶。在一些实施方案中,50 种或更多种可检测地标记的寡核苷酸靶向相同的靶。在一些实施方案中,60 种或更多种可检测地标记的寡核苷酸靶向相同的靶。在一些实施方案中,70 种或更多种可检测地标记的寡核苷酸靶向相同的靶。在一些实施方案中,80 种或更多种可检测地标记的寡核苷酸靶向相同的靶。在一些实施方案中,90 种或更多种可检测地标记的寡核苷酸靶向相同的靶。在一些实施方案中,100 种或更多种可检测地标记的寡核苷酸靶向相同的靶。在一些实施方案中,约 1-10 种可检测地标记的寡核苷酸靶向相同的靶。在一些实施方案中,约 5-15 种可检测地标记的寡核苷酸靶向相同的靶。在一些实施方案中,约 10-20 种可检测地标记的寡核苷酸靶向相同的靶。在一些实施方案中,约 15-25 种可检测地标记的寡核苷酸靶向相同的靶。在一些实施方案中,约 20-30 种可检测地标记的寡核苷酸靶向相同的靶。在一些实施方案中,约 25-35 种可检测地标记的寡核苷酸靶向相同的靶。在一些实施方案中,约 30-40 种可检测地标记的寡核苷酸靶向相同的靶。在一些实施方案中,约 35-45 种可检测地标记的寡核苷酸靶向相同的靶。在一些实施方案中,约 40-50 种可检测地标记的寡核苷酸靶向相同的靶。在一些实施方案中,约 45-55 种可检测地标记的寡核苷酸靶向相同的靶。在一些实施方案中,约 50-70 种可检测地标记的寡核苷酸靶向相同的靶。在一些实施方案中,约 60-80 种可检测地标记的寡核苷酸靶向相同的靶。在一些实施方案中,约 70-90 种可检测地标记的寡核苷酸靶向相同的靶。在一些实施方案中,约 80-100 种可检测地标记的寡核苷酸靶向相同的靶。

[0144] 在一些实施方案中,除了其他以外,对相同的靶使用多种可检测地标记的寡核苷酸增加信号强度。在一些实施方案中,靶向相同靶的组中的每种可检测地标记的寡核苷酸与靶的不同部分相互作用。

[0145] 在一些实施方案中,在一组中针对靶的所有可检测地标记的寡核苷酸具有相同的可检测部分。在一些实施方案中,所有可检测地标记的寡核苷酸以相同的方式被标记。在一些实施方案中,针对靶的所有可检测地标记的寡核苷酸具有相同的荧光团。

[0146] 在一些实施方案中,针对靶的可检测地标记的寡核苷酸定位在靶的被靶向区域内。被靶向区域可以具有多种长度。在一些实施方案中,被靶向区域的长度为约 20bp。在一些实施方案中,被靶向区域的长度为约 30bp。在一些实施方案中,被靶向区域的长度为约 40bp。在一些实施方案中,被靶向区域的长度为约 50bp。在一些实施方案中,被靶向区域的长度为约 60bp。在一些实施方案中,被靶向区域的长度为约 80bp。在一些实施方案中,被靶向区域的长度为约 100bp。在一些实施方案中,被靶向区域的长度为约 150bp。在一些实施方案中,被靶向区域的长度为约 200bp。在一些实施方案中,被靶向区域的长度为约 250bp。在一些实施方案中,被靶向区域的长度为约 300bp。在一些实施方案中,被靶向区域的长度为约 350bp。在一些实施方案中,被靶向区域的长度为约 400bp。在一些实施方案中,被靶向区域的长度为约 450bp。在一些实施方案中,被靶向区域的长度为约 500bp。在一些实施方案中,被靶向区域的长度为约 600bp。在一些实施方案中,被靶向区域的长度为约 700bp。在一些实施方案中,被靶向区域的长度为约 800bp。在一些实施方案中,被靶向区域的长度为约 900bp。在一些实施方案中,被靶向区域的长度为约 1,000bp。在一些实施方案中,针对靶的可检测地标记的寡核苷酸被定位为在靶上邻近彼此。

[0147] 如本领域普通技术人员所理解的,不同的技术可以被用于成像步骤。示例性方法包括但不局限于落射荧光显微术、共聚焦显微术、不同类型的超分辨率显微术 (PALM/STORM、SSIM/GSD/STED) 和光片照明显微术 (SPIM 等)。

[0148] 示例性超分辨率技术包括但不局限于 I⁵M 和 4Pi-显微术、受激发射损耗显微术 (STEDM)、基态损耗显微术 (GSDM)、空间结构照明显微术 (SSIM)、光激活定位显微术 (PALM)、可逆饱和光线性荧光跃迁 (RESOLFT)、全内反射荧光显微镜 (TIRFM)、荧光 -PALM (FPALM)、随机光学重建显微术 (STORM)、一纳米精度的荧光成像 (FIONA) 及其组合。例如:Chi, 2009 "Super-resolution microscopy: breaking the limits, Nature Methods 6(1):15-18; Blow 2008, "New ways to see a smaller world," Nature 456:825-828; Hell, 等人, 2007, "Far-Field Optical Nanoscopy," Science 316:1153; R. Heintzmann 和 G. Ficz, 2006, "Breaking the resolution limit in light microscopy," Briefings in Functional Genomics and Proteomics 5(4):289-301; Garini 等人, 2005, "From micro to nano: recent advances in high-resolution microscopy," Current Opinion in Biotechnology 16:3-12; 和 Bewersdorf 等人, 2006, "Comparison of I⁵M and 4Pi-microscopy," 222(2):105-117; 和 Wells, 2004, "Man the Nanoscopes," JCB 164(3):337-340。

[0149] 在一些实施方案中,使用电子显微镜 (EM)。

[0150] 在一些实施方案中,成像步骤检测靶。在一些实施方案中,成像步骤定位靶。在一些实施方案中,成像步骤提供靶的三维空间信息。在一些实施方案中,成像步骤定量靶。通过使用多次接触和成像步骤,所提供的方法能够以出乎意料的高通量对大量的靶提供空间信息和 / 或数量信息。例如,当使用 F 种可检测的不同类型的标签时,在 N 次接触和成像步骤之后可以获得规模多达 F^N 个靶的空间信息和 / 或数量信息。

[0151] 在一些实施方案中,所提供的方法包括在接触和 / 或成像步骤之前或之后的另外的步骤。在一些实施方案中,所提供的方法在每个成像步骤之后包括去除多种可检测地标记的寡核苷酸的步骤。在一些实施方案中,去除步骤包括降解可检测地标记的寡核苷酸。在

一些实施方案中,去除步骤不显著地降解靶,使得如果期望,能靶可用于下一轮接触和/或成像步骤。在一些实施方案中,去除步骤包括使多种可检测地标记的寡核苷酸与消化可检测地标记的寡核苷酸的酶接触。在一些实施方案中,去除步骤包括使可检测地标记的寡核苷酸与 DNA 酶或 RNA 酶接触。例如,在一些实施方案中,可检测地标记的寡核苷酸包含 DNA 序列并且 DNA 酶被用于其的降解;在一些其他的实施方案中,可检测地标记的寡核苷酸包含 RNA 序列并且 RNA 酶被用于其的降解。在一些实施方案中,去除步骤包括降解可检测部分。在一些实施方案中,去除步骤包括光漂白。

[0152] 在一些实施方案中,一组可检测地标记的寡核苷酸的靶还是另一组的靶。在一些实施方案中,一组可检测地标记的寡核苷酸的靶与另一组的靶重叠。在一些实施方案中,重叠多于 10%。在一些实施方案中,重叠多于 20%。在一些实施方案中,重叠多于 30%。在一些实施方案中,重叠多于 40%。在一些实施方案中,重叠多于 50%。在一些实施方案中,重叠多于 60%。在一些实施方案中,重叠多于 70%。在一些实施方案中,重叠多于 80%。在一些实施方案中,重叠多于 90%。在一些实施方案中,重叠多于 91%。在一些实施方案中,重叠多于 92%。在一些实施方案中,重叠多于 93%。在一些实施方案中,重叠多于 94%。在一些实施方案中,重叠多于 95%。在一些实施方案中,重叠多于 96%。在一些实施方案中,重叠多于 97%。在一些实施方案中,重叠多于 98%。在一些实施方案中,重叠多于 99%。在一些实施方案中,重叠多于 99.5%。在一些实施方案中,重叠多于 99.6%。在一些实施方案中,重叠多于 99.7%。在一些实施方案中,重叠多于 99.8%。在一些实施方案中,重叠多于 99.9%。在一些实施方案中,重叠是 100%。在一些实施方案中,一组可检测地标记的寡核苷酸的靶与另一组的靶相同。在一些实施方案中,每组可检测地标记的寡核苷酸靶向相同的靶。

[0153] 在一些实施方案中,在第二接触步骤中靶向第一转录物或 DNA 基因座(第一靶)的第三可检测地标记的寡核苷酸任选地与靶向第一转录物或 DNA 基因座的第一可检测地标记的寡核苷酸具有相同的序列。在一些实施方案中,序列是相同的。在一些实施方案中,序列是不同的。相似地,在一些实施方案中,在第二接触步骤中靶向第二转录物或 DNA 基因座(第一靶)的第四可检测地标记的寡核苷酸任选地与靶向第一转录物或 DNA 基因座的第二可检测地标记的寡核苷酸具有相同的序列。在一些实施方案中,序列是相同的。在一些实施方案中,序列是不同的。

[0154] 在一些实施方案中,第二多种与第一多种的不同在于存在于第二多种中的寡核苷酸的至少一种被与第一多种中的靶向相同的转录物或 DNA 基因座的对应寡核苷酸不同的可检测部分标记。在一些实施方案中,可检测地标记的寡核苷酸的每个多种与另一多种的不同在于存在于一个多种的至少一种寡核苷酸被与在另一多种中的靶向相同的转录物或 DNA 基因座的对应寡核苷酸不同的可检测部分标记。

[0155] 在一些实施方案中,可检测地标记的寡核苷酸具有 [S]-[L] 结构,其中 [S] 是寡核苷酸序列, [L] 是可检测部分或可检测部分的组合。在一些实施方案中, [L] 包含多个单位的可检测标记的,例如荧光团,所述多个单位的每个独立地与寡核苷酸序列例如 [S] 的一个或更多个核苷部分缔合。在一些实施方案中,附连至相同的可检测地标记的寡核苷酸的每种可检测标记提供相同的可检测信号。在一些实施方案中,附连至相同寡核苷酸序列的所有可检测标记是相同的。

[0156] 在一些实施方案中,在两组或更多组可检测地标记的寡核苷酸之中靶向相同的靶的寡核苷酸具有相同的序列组,即,在可检测地标记的寡核苷酸之间的差异,如果存在,在可检测部分内,而不是序列内。例如,在一组可检测地标记的寡核苷酸中,靶向第一靶的可检测地标记的寡核苷酸都具有相同的可检测部分,或检测部分的组合 $[L]_1$:

[0157] $[S]_1-[L]_1$ 、 $[S]_2-[L]_1$ 、 \cdots 、 $[S]_n-[L]_1$, 其中 n 是针对靶的可检测地标记的寡核苷酸的数目,例如,3-50 的整数;

[0158] 在另一组可检测地标记的寡核苷酸中,其中靶向相同的靶的寡核苷酸被不同地标记,靶向相同靶的寡核苷酸具有相同的寡核苷酸序列组 ($[S]_1$ 、 $[S]_2$ 、 \cdots 、 $[S]_n$),还有一个不同的 $[L]_2$:

[0159] $[S]_1-[L]_2$ 、 $[S]_2-[L]_2$ 、 \cdots 、 $[S]_n-[L]_2$, 其中 $[L]_1$ 被与 $[L]_2$ 不同地检测。

[0160] 为了示例本发明的某些实施方案,以下提供了两步骤、两标记、4 靶 ($F^N = 2^2 = 4$) 过程,其中靶向每组中的相同靶的所有可检测地标记的寡核苷酸独立地具有相同的可检测部分:

[0161] 步骤 1. 使靶与第一多种 (P1) 可检测地标记的寡核苷酸接触:

[0162] 靶 T1 : $[S]_{P1 T1 1}[L]_1$ 、 $[S]_{P1 T1 2}[L]_1$ 、 $[S]_{P1 T1 3}[L]_1$ 、 \cdots 、 $[S]_{P1 T1 P1T1}[L]_1$, 其中 P1T1 是第一多种中靶向 T1 的可检测地标记的寡核苷酸的数目,并且 $[L]_1$ 是第一可检测标记;

[0163] 靶 T2 : $[S]_{P1 T2 1}[L]_1$ 、 $[S]_{P1 T2 2}[L]_1$ 、 $[S]_{P1 T2 3}[L]_1$ 、 \cdots 、 $[S]_{P1 T2 P1T2}[L]_1$, 其中 P1T2 是第一多种中靶向 T2 的可检测地标记的寡核苷酸的数目;

[0164] 靶 T3 : $[S]_{P1 T3 1}[L]_2$ 、 $[S]_{P1 T3 2}[L]_2$ 、 $[S]_{P1 T3 3}[L]_2$ 、 \cdots 、 $[S]_{P1 T3 P1T3}[L]_2$, 其中 P1T3 是第一多种中靶向 T3 的可检测地标记的寡核苷酸的数目,并且 $[L]_2$ 是与 $[L]_1$ 不同的可检测的标记;

[0165] 靶 T4 : $[S]_{P1 T4 1}[L]_2$ 、 $[S]_{P1 T4 2}[L]_2$ 、 $[S]_{P1 T4 3}[L]_2$ 、 \cdots 、 $[S]_{P1 T4 P1T4}[L]_2$, 其中 P1T4 是第一多种中靶向 T4 的可检测地标记的寡核苷酸的数目。

[0166] 步骤 2 : 成像;

[0167] 步骤 3 : 从靶去除 P1;

[0168] 步骤 4 : 使靶与第二多种 (P2) 可检测地标记的寡核苷酸接触:

[0169] 靶 T1 : $[S]_{P2 T1 1}[L]_1$ 、 $[S]_{P2 T1 2}[L]_1$ 、 $[S]_{P2 T1 3}[L]_1$ 、 \cdots 、 $[S]_{P2 T1 P2T1}[L]_1$, 其中 P2T1 是第二多种中靶向 T1 的可检测地标记的寡核苷酸的数目;

[0170] 靶 T2 : $[S]_{P2 T2 1}[L]_2$ 、 $[S]_{P2 T2 2}[L]_2$ 、 $[S]_{P2 T2 3}[L]_2$ 、 \cdots 、 $[S]_{P2 T2 P2T2}[L]_2$, 其中 P2T2 是第二多种中靶向 T2 的可检测地标记的寡核苷酸的数目;

[0171] 靶 T3 : $[S]_{P2 T3 1}[L]_1$ 、 $[S]_{P2 T3 2}[L]_1$ 、 $[S]_{P2 T3 3}[L]_1$ 、 \cdots 、 $[S]_{P2 T3 P2T3}[L]_1$, 其中 P2T3 是第二多种中靶向 T3 的可检测地标记的寡核苷酸的数目;

[0172] 靶 T4 : $[S]_{P2 T4 1}[L]_2$ 、 $[S]_{P2 T4 2}[L]_2$ 、 $[S]_{P2 T4 3}[L]_2$ 、 \cdots 、 $[S]_{P2 T4 P2T4}[L]_2$, 其中 P2T4 是第二多种中靶向 T4 的可检测地标记的寡核苷酸的数目。

[0173] 步骤 5 : 成像。

[0174] 在两次成像步骤之后,每种靶具有其自身独特的顺序条形码:

[0175] T1 : $[L]_1[L]_1$;

[0176] T2 : $[L]_1[L]_2$;

[0177] T3 : $[L]_2[L]_1$; 并且

[0178] $T4 : [L]_2 [L]_2$ 。

[0179] 在一些实施方案中,另外的条形码,也可以使用 $T1--$ 、 $T2--$ 、 $--T1$ 、 $--T2$,其中 $--$ 指示对于该步骤没有信号。

[0180] 在以上示例性过程中, $P1T1$ 、 $P1T2$ 、 $P1T3$ 、 $P1T4$ 、 $P2T1$ 、 $P2T2$ 、 $P2T3$ 和 $P2T4$ 的每个独立地是自然数(大于0的整数)。在一些实施方案中, $P1T1 = P2T1$ 。在一些实施方案中, $P1T2 = P2T2$ 。在一些实施方案中, $P1T3 = P2T3$ 。在一些实施方案中, $P1T4 = P2T4$ 。在一些实施方案中,一种可检测地标记的寡核苷酸被用于靶。在一些实施方案中,两种或更多种可检测地标记的寡核苷酸被用于靶。

[0181] 在一些实施方案中,在每个多种中靶向相同靶的可检测地标记的寡核苷酸具有相同的序列组。例如,对于在以上实例中的靶 $T1$, $[S]_{P1 T1 1}$ 至 $[S]_{P1 T1 P1T1}$ 的每个独立地具有与 $[S]_{P2 T1 1}$ 至 $[S]_{P2 T1 P2T1}$ 之一相同的序列,并且 $[S]_{P2 T1 1}$ 至 $[S]_{P2 T1 P2T1}$ 的每个独立地具有与 $[S]_{P1 T1 1}$ 至 $[S]_{P1 T1 P1T1}$ 之一相同的序列。在一些实施方案中,在每个多种中靶向相同靶的可检测地标记的寡核苷酸具有不同的序列组。

[0182] 在一些实施方案中,所提供的方法在成像步骤之后任选地包括去除多种可检测地标记的寡核苷酸的步骤。在一些实施方案中,所提供的方法在成像步骤之后包括去除步骤。在一些实施方案中,所提供的方法在除了最后的成像步骤之外的每个成像步骤之后包括去除步骤。在一些实施方案中,所提供的方法在每个成像步骤之后包括去除步骤。

[0183] 在所提供的方法中的去除步骤可以达到多种目的中的一种或更多种。在一些实施方案中,去除步骤将多种可检测地标记的寡核苷酸从靶去除以使得靶可被用于与另一种可检测地标记的寡核苷酸相互作用。在一些实施方案中,去除步骤去除多种可检测地标记的寡核苷酸以使得一个多种可检测地标记的寡核苷酸的可检测部分不干扰另一个多种可检测地标记的寡核苷酸与靶的结合的检测。在一些实施方案中,去除步骤去除至少 80% 的可检测地标记的寡核苷酸。在一些实施方案中,去除步骤去除至少 85% 的可检测地标记的寡核苷酸。在一些实施方案中,去除步骤去除至少 90% 的可检测地标记的寡核苷酸。在一些实施方案中,去除步骤去除至少 91% 的可检测地标记的寡核苷酸。在一些实施方案中,去除步骤去除至少 92% 的可检测地标记的寡核苷酸。在一些实施方案中,去除步骤去除至少 93% 的可检测地标记的寡核苷酸。在一些实施方案中,去除步骤去除至少 94% 的可检测地标记的寡核苷酸。在一些实施方案中,去除步骤去除至少 95% 的可检测地标记的寡核苷酸。在一些实施方案中,去除步骤去除至少 96% 的可检测地标记的寡核苷酸。在一些实施方案中,去除步骤去除至少 97% 的可检测地标记的寡核苷酸。在一些实施方案中,去除步骤去除至少 98% 的可检测地标记的寡核苷酸。在一些实施方案中,去除步骤去除至少 99% 的可检测地标记的寡核苷酸。在一些实施方案中,去除步骤去除至少 99.1% 的可检测地标记的寡核苷酸。在一些实施方案中,去除步骤去除至少 99.2% 的可检测地标记的寡核苷酸。在一些实施方案中,去除步骤去除至少 99.3% 的可检测地标记的寡核苷酸。在一些实施方案中,去除步骤去除至少 99.4% 的可检测地标记的寡核苷酸。在一些实施方案中,去除步骤去除至少 99.5% 的可检测地标记的寡核苷酸。在一些实施方案中,去除步骤去除至少 80% 的可检测信号。在一些实施方案中,去除步骤去除至少 85% 的可检测信号。在一些实施方案中,去除步骤去除至少 90% 的可检测信号。在一些实施方案中,去除步骤去除至少 91% 的可检测信号。在一些实施方案中,去除步骤去除至少 92% 的可检测信号。在一些实施方案

中,去除步骤去除至少 93%的可检测信号。在一些实施方案中,去除步骤去除至少 94%的可检测信号。在一些实施方案中,去除步骤去除至少 95%的可检测信号。在一些实施方案中,去除步骤去除至少 96%的可检测信号。在一些实施方案中,去除步骤去除至少 97%的可检测信号。在一些实施方案中,去除步骤去除至少 98%的可检测信号。在一些实施方案中,去除步骤去除至少 99%的可检测信号。在一些实施方案中,去除步骤去除至少 99.5%的可检测信号。在一些实施方案中,去除步骤去除 100%的可检测信号。在一些实施方案中,在去除步骤之后没有信号可被成像步骤检测出。

[0184] 去除步骤任选地保留靶(例如,转录物或 DNA 基因座)用于进一步使用,例如,通过另外的接触和/或成像步骤进一步检测或定量。在一些实施方案中,去除步骤保留至少 80%的靶。保留的靶的百分比可以例如通过比较在去除步骤之前和之后采集的数据而被测量出,任选地利用相同的接触和成像方案进行。在一些实施方案中,去除步骤保留至少 85%的靶。在一些实施方案中,去除步骤保留至少 90%的靶。在一些实施方案中,去除步骤保留至少 91%的靶。在一些实施方案中,去除步骤保留至少 92%的靶。在一些实施方案中,去除步骤保留至少 93%的靶。在一些实施方案中,去除步骤保留至少 94%的靶。在一些实施方案中,去除步骤保留至少 95%的靶。在一些实施方案中,去除步骤保留至少 96%的靶。在一些实施方案中,去除步骤保留至少 97%的靶。在一些实施方案中,去除步骤保留至少 98%的靶。在一些实施方案中,去除步骤保留至少 99%的靶。

[0185] 用于去除可检测地标记的寡核苷酸的方法在领域是被广泛已知的。在一些实施方案中,去除步骤包括降解可检测地标记的寡核苷酸。在一些实施方案中,通过酶消化去除可检测地标记的寡核苷酸。在一些实施方案中,去除步骤包括使多种可检测地标记的寡核苷酸与消化可检测地标记的寡核苷酸的酶接触。

[0186] 合适的酶是本领域中被广泛使用的。例如,根据可检测地标记的寡核苷酸和/或靶的类型,可以使用 DNA 酶或 RNA 酶。在一些实施方案中,包含用于检测/定量 RNA 靶的 DNA 序列的可检测地标记的寡核苷酸通过 DNA 酶例如 DNA 酶 I 来消化。在一些实施方案中,包含用于检测/定量 DNA 靶的 RNA 序列的可检测地标记的寡核苷酸通过 RNA 酶来消化。在一些实施方案中,可检测地标记的 RNA 寡核苷酸被用来靶向 DNA 基因座。

[0187] 在一些实施方案中,可检测地标记的寡核苷酸通过与一个或多个中间物结合或杂交来与其靶相互作用,所述一个或多个中间物诸如与靶结合、杂交或以其他方式连接的寡核苷酸。在一些实施方案中,可检测地标记的寡核苷酸通过与中间寡核苷酸杂交来与靶相互作用,所述中间寡核苷酸与靶杂交,其中中间寡核苷酸包含与靶互补的序列和与可检测地标记的寡核苷酸互补的序列(突出端)。在一些实施方案中,去除步骤将可检测地标记的寡核苷酸去除,任选地使中间寡核苷酸保持完整。在一些实施方案中,去除步骤将可检测地标记的寡核苷酸去除并且使中间寡核苷酸保持完整。在一些实施方案中,从化学或酶学角度,可检测地标记的寡核苷酸与中间物不同,以使得可检测地标记的寡核苷酸可被选择性地去除。

[0188] 在一些实施方案中,中间 DNA 寡核苷酸被用来以突出端(例如,20nt)与 DNA 基因座杂交,以使得包含 RNA 序列的桥接寡核苷酸可与互补序列(例如, RNA 桥接探针)结合。RNA 桥接探针可以用染料或 HCR 聚合物(还可以是 DNA)直接地标记。成像之后,可使用 RNA 酶来消化掉 RNA 桥接探针,同时使 DNA 探针保持完整地杂交在 DNA 基因座上。这样的方法

提供多种优点。例如,随后的接触步骤仅包括 RNA 桥接探针以突出端与 DNA 寡核苷酸杂交,并且避免使双链 DNA 解链并与 DNA 寡核苷酸杂交,这是个困难的过程。另外,可以使靶向相同基因的所有 DNA 寡核苷酸(例如,20 个至 40 个)的突出端相同,以使得每轮杂交每个基因仅需要一种类型的 RNA 桥接探针。为了针对不同的杂交(接触步骤)切换颜色,技术人员可以用不同的标记或不同的 HCR 聚合物改变 RNA 桥接探针。还可以使用这样的 DNA 桥接探针:可以例如用桥或 HCR 发夹上的诸如 EcoRI 的特异性酶限制位点来特异性地去除。将细胞与合适的核酸酶一起孵育可以消化掉所有的可检测部分而不影响 DNA 基因座和/或其上杂交的探针。

[0189] 在一些实施方案中,可检测地标记的寡核苷酸包含 5' - 磷酸化并且可以被 Lambda 核酸外切酶降解,而中间寡核苷酸未被 5' - 磷酸化并且不能被 Lambda 核酸外切酶降解。

[0190] 在一些实施方案中,可检测地标记的寡核苷酸包含尿嘧啶。在一些实施方案中,可检测地标记的寡核苷酸包含尿嘧啶,并且可以被 USER™酶(New England BioLabs, Ipswich, Massachusetts, MA, US) 降解,而中间寡核苷酸不包含尿嘧啶并且不能被 USER™酶降解。

[0191] 在一些实施方案中,与中间寡核苷酸的突出端杂交的寡核苷酸在与突出端杂交时具有凹陷的 3' - 末端。具有凹陷的 3' - 末端的可检测地标记的寡核苷酸当与中间寡核苷酸杂交时可以被核酸外切酶 III 选择性地消化。由于核酸外切酶 III 对不具有凹陷的 3' - 末端或其 3' - 末端是 RNA-DNA 双链体的中间寡核苷酸的活性弱得多,所以它们可以保持完整。

[0192] 在一些实施方案中,当涉及到酶时,在产生最佳结果的温度进行去除步骤。在一些实施方案中,在约 37°C 进行去除步骤。在一些实施方案中,在室温进行去除步骤。在一些实施方案中,用 Lambda 核酸外切酶的消化在约 37°C 进行。在一些实施方案中,用 USER™酶的消化在约 37°C 进行。在一些实施方案中,用 USER™酶的消化在室温进行。在一些实施方案中,用核酸外切酶 III 的消化在约 37°C 进行。在一些实施方案中,用核酸外切酶 III 的消化在室温进行。

[0193] 在一些实施方案中,中间寡核苷酸和用于可检测地标记的寡核苷酸结合的突出端序列的使用提供了多种优点。在一些实施方案中,突出端序列和可检测地标记的寡核苷酸之间的杂交的历程比中间寡核苷酸与靶之间的杂交的历程快。在一些实施方案中,针对靶的所有中间寡核苷酸包含相同的突出端序列,并且针对靶的所有可检测地标记的寡核苷酸包含用于与相同的突出端序列结合的相同的互补序列。在一些实施方案中,一组可检测地标记的寡核苷酸和一组中间寡核苷酸之间的杂交比一组中间寡核苷酸和一组靶之间的杂交快多达约 20-40 倍。在一些实施方案中,可检测地标记的寡核苷酸和中间寡核苷酸之间的杂交可以在 30 分钟内完成,与之相比,在一些情况中,对于中间寡核苷酸和靶之间的杂交长达约 12 小时。

[0194] 在一些实施方案中,在去除步骤中使用链置换来去除可检测地标记的寡核苷酸。在一些实施方案中,在去除步骤中利用加热来使可检测地标记的寡核苷酸解离。

[0195] 在一些实施方案中,去除步骤包括光漂白。在一些实施方案中,光漂白破坏可检测地标记的寡核苷酸的染料,诸如荧光团。

[0196] 在一些实施方案中,第一组和第二组可检测地标记的寡核苷酸靶向每个靶的不同序列,并且在第一成像步骤之后的去除步骤是任选的。例如,一种策略是用不同的 DNA 探针(可检测地标记的 DNA 寡核苷酸)来靶向相同的 RNA,以使得第一多种探针可以靶向 RNA 上

的一组序列,并且第二多种探针靶向相同 RNA 上的不同的序列组。在第一杂交(接触)时,使用第一多种探针。它们随后可以被成像并且任选地被 DNA 酶或破坏寡核苷酸或染料的其他方法光漂白或消化。第二组探针可以被杂交并且被成像,而不受第一组探针的干扰。

[0197] 在一些实施方案中,所提供的方法任选地包括 HCR、光片照明显微术、CLARITY 或其组合。在一些实施方案中,所提供的方法允许对组织、器官或生物体中的靶直接谱分析。在一些实施方案中,器官是大脑。在一些实施方案中,所提供的方法允许对完整的大脑或组织中的转录物直接成像。在一些实施方案中,所提供的方法还包括 HCR。在一些实施方案中,所提供的方法还包括光片照明显微术。在一些实施方案中,所提供的方法还包括 CLARITY。

[0198] 所提供的方法提供许多超越本发明之前的方法的优点。例如,在一些实施方案中,所提供的方法在合理的成本下提供了高通量。在一些实施方案中,所提供的方法提供了靶的直接探测,而不转化靶或扩增靶。在一些实施方案中,所提供的方法使得能够快速规模放大而不需要大量的可检测标记。在一些实施方案中,所提供的方法可以对相同的靶应用多种标签并且因此增加信号强度。在一些实施方案中,所提供的方法提供了所述优点的组合。

[0199] 在一些实施方案中,本发明提供了包含多种可检测地标记的寡核苷酸组合物,所述组合物例如在所提供的方法中使用。示例性组合物包括但不限于在本文的示例性方法实施方案中描述的那些组合物。

[0200] 在一些实施方案中,本发明提供了包含多种可检测地标记的寡核苷酸的组合物,所述可检测地标记的寡核苷酸的每种靶向核酸并且被可检测部分标记,因此所述组合物至少包含:

[0201] (i) 第一寡核苷酸,所述第一寡核苷酸靶向第一核酸并且被第一可检测部分标记;以及

[0202] (ii) 第二寡核苷酸,所述第二寡核苷酸靶向第二核酸并且被第二可检测部分标记。

[0203] 在一些实施方案中,本发明提供了包含多种可检测地标记的寡核苷酸的组合物,所述可检测地标记的寡核苷酸的每种靶向转录物或 DNA 基因座并且被可检测部分标记,因此所述组合物至少包含:

[0204] (i) 第一寡核苷酸,所述第一寡核苷酸靶向第一转录物或 DNA 基因座并且被第一可检测部分标记;以及

[0205] (ii) 第二寡核苷酸,所述第二寡核苷酸靶向第二转录物或 DNA 基因座并且被第二可检测部分标记。

[0206] 在一些实施方案中,本发明提供了包含多种可检测地标记的寡核苷酸的试剂盒,所述可检测地标记的寡核苷酸的每种靶向转录物或 DNA 基因座并且被可检测部分标记,因此所述试剂盒至少包含:

[0207] (i) 第一寡核苷酸,所述第一寡核苷酸靶向第一转录物或 DNA 基因座并且被第一可检测部分标记;

[0208] (ii) 第二寡核苷酸,所述第二寡核苷酸靶向第二转录物或 DNA 基因座并且被第二可检测部分标记。

[0209] (iii) 第三寡核苷酸,所述第三寡核苷酸任选地在序列上与第一寡核苷酸相同,所述第三寡核苷酸靶向第一转录物或 DNA 基因座并且被第一可检测部分、第二可检测部分或

第三可检测部分标记;以及

[0210] (iv) 第四寡核苷酸,所述第四寡核苷酸任选地在序列上与第二寡核苷酸相同,所述第四寡核苷酸靶向第二转录物或 DNA 基因座,并且被第一可检测部分、第二可检测部分、第三可检测部分或第四可检测部分标记,其中第三寡核苷酸被与第一寡核苷酸不同的可检测部分标记,或第四寡核苷酸用与第二寡核苷酸不同的可检测部分标记,或两者。

[0211] 在一些实施方案中,组合物中的靶向相同的靶(转录物或 DNA 基因座)的可检测地标记的寡核苷酸被提供相同的可检测信号或在成像步骤中无法区分的可检测信号的部分标记。在一些实施方案中,组合物中的靶向相同的靶的可检测地标记的寡核苷酸被相同的可检测部分标记。

[0212] 在一些实施方案中,可检测部分是或包含荧光团。在一些实施方案中,可检测部分是荧光团。示范性荧光团是领域内被广泛已知和使用的,例如但不限于荧光素、罗丹明、Alexa Fluor 类、Dylight Fluor 类、ATTO 类染料或其任何类似物或衍生物。

[0213] 在一些实施方案中,第一可检测地标记的寡核苷酸和第二可检测地标记的寡核苷酸靶向不同的靶。在一些实施方案中,第一可检测地标记的寡核苷酸和第二可检测地标记的寡核苷酸靶向相同的靶。在一些实施方案中,组合物或试剂盒中的可检测地标记的寡核苷酸靶向两种或更多种靶,例如,转录物和/或 DNA 基因座。在一些实施方案中,组合物或试剂盒中的可检测地标记的寡核苷酸靶向两种或更多种转录物。在一些实施方案中,组合物或试剂盒中的可检测地标记的寡核苷酸靶向两种或更多种 DNA 基因座。在一些实施方案中,组合物或试剂盒中的可检测地标记的寡核苷酸靶向至少 4 种靶。在一些实施方案中,组合物或试剂盒中的可检测地标记的寡核苷酸靶向至少 9 种靶。在一些实施方案中,组合物或试剂盒中的可检测地标记的寡核苷酸靶向至少 16 种靶。在一些实施方案中,组合物或试剂盒中的可检测地标记的寡核苷酸靶向至少 25 种靶。在一些实施方案中,组合物或试剂盒中的可检测地标记的寡核苷酸靶向至少 36 种靶。在一些实施方案中,组合物或试剂盒中的可检测地标记的寡核苷酸靶向至少 50 种靶。在一些实施方案中,组合物或试剂盒中的可检测地标记的寡核苷酸靶向至少 100 种靶。在一些实施方案中,组合物或试剂盒中的可检测地标记的寡核苷酸靶向至少 200 种靶。在一些实施方案中,组合物或试剂盒中的可检测地标记的寡核苷酸靶向至少 500 种靶。在一些实施方案中,组合物或试剂盒中的可检测地标记的寡核苷酸靶向至少 1,000 种靶。在一些实施方案中,组合物或试剂盒中的可检测地标记的寡核苷酸靶向至少 5,000 种靶。在一些实施方案中,组合物或试剂盒中的可检测地标记的寡核苷酸靶向至少 10,000 种靶。在一些实施方案中,组合物或试剂盒中的可检测地标记的寡核苷酸靶向至少 50,000 种靶。在一些实施方案中,组合物或试剂盒中的可检测地标记的寡核苷酸靶向至少 100,000 种靶。在一些实施方案中,组合物或试剂盒中的可检测地标记的寡核苷酸靶向至少 1,000,000 种靶。

[0214] 在一些实施方案中,第一寡核苷酸和第二寡核苷酸具有不同的寡核苷酸序列。在一些实施方案中,第一可检测部分和第二可检测部分是不同的。在一些实施方案中,第一可检测部分和第二可检测部分是相同的。

[0215] 在一些实施方案中,第一寡核苷酸和第二寡核苷酸共有少于 5% 的序列同一性。在一些实施方案中,第一寡核苷酸和第二寡核苷酸共有少于 10% 的序列同一性。在一些实施方案中,第一寡核苷酸和第二寡核苷酸共有少于 20% 的序列同一性。在一些实施方案中,

第一寡核苷酸和第二寡核苷酸共有少于 30% 的序列同一性。在一些实施方案中,第一寡核苷酸和第二寡核苷酸共有少于 40% 的序列同一性。在一些实施方案中,第一寡核苷酸和第二寡核苷酸共有少于 50% 的序列同一性。在一些实施方案中,第一寡核苷酸和第二寡核苷酸共有少于 60% 的序列同一性。在一些实施方案中,第一寡核苷酸和第二寡核苷酸共有少于 65% 的序列同一性。在一些实施方案中,第一寡核苷酸和第二寡核苷酸共有少于 68% 的序列同一性。在一些实施方案中,第一寡核苷酸和第二寡核苷酸共有少于 70% 的序列同一性。在一些实施方案中,第一寡核苷酸和第二寡核苷酸共有少于 80% 的序列同一性。在一些实施方案中,第一寡核苷酸和第二寡核苷酸共有少于 90% 的序列同一性。

[0216] 在一些实施方案中,每种寡核苷酸与任何其他寡核苷酸共有少于 5% 的序列同一性。在一些实施方案中,每种寡核苷酸与其他任何寡核苷酸共有少于 10% 的序列同一性。在一些实施方案中,每种寡核苷酸与任何其他寡核苷酸共有少于 20% 的序列同一性。在一些实施方案中,每种寡核苷酸与任何其他寡核苷酸共有少于 30% 的序列同一性。在一些实施方案中,每种寡核苷酸与任何其他寡核苷酸共有少于 40% 的序列同一性。在一些实施方案中,每种寡核苷酸与任何其他寡核苷酸共有少于 50% 的序列同一性。在一些实施方案中,每种寡核苷酸与任何其他寡核苷酸共有少于 55% 的序列同一性。在一些实施方案中,每种寡核苷酸与任何其他寡核苷酸共有少于 60% 的序列同一性。在一些实施方案中,每种寡核苷酸与任何其他寡核苷酸共有少于 65% 的序列同一性。在一些实施方案中,每种寡核苷酸与任何其他寡核苷酸共有少于 68% 的序列同一性。在一些实施方案中,每种寡核苷酸与任何其他寡核苷酸共有少于 70% 的序列同一性。在一些实施方案中,每种寡核苷酸与任何其他寡核苷酸共有少于 80% 的序列同一性。在一些实施方案中,每种寡核苷酸与任何其他寡核苷酸共有少于 90% 的序列同一性。

[0217] 在一些实施方案中,组合物或试剂盒包含靶向相同的靶的两种或更多种可检测地标记的寡核苷酸。在一些实施方案中,5、10、20、30、40、50 种或更多种可检测地标记的寡核苷酸靶向相同的靶。

[0218] 可检测地标记的寡核苷酸可以为多种合适的长度。在一些实施方案中,可检测地标记的寡核苷酸的长度为 15 个碱基对。在一些实施方案中,可检测地标记的寡核苷酸的长度为 16 个碱基对。在一些实施方案中,可检测地标记的寡核苷酸的长度为 17 个碱基对。在一些实施方案中,可检测地标记的寡核苷酸的长度为 18 个碱基对。在一些实施方案中,可检测地标记的寡核苷酸的长度为 19 个碱基对。在一些实施方案中,可检测地标记的寡核苷酸的长度为 20 个碱基对。在一些实施方案中,可检测地标记的寡核苷酸的长度为 21 个碱基对。在一些实施方案中,可检测地标记的寡核苷酸的长度为 22 个碱基对。在一些实施方案中,可检测地标记的寡核苷酸的长度为 23 个碱基对。在一些实施方案中,可检测地标记的寡核苷酸的长度为 24 个碱基对。在一些实施方案中,可检测地标记的寡核苷酸的长度为 25 个碱基对。在一些实施方案中,可检测地标记的寡核苷酸的长度为 26 个碱基对。在一些实施方案中,可检测地标记的寡核苷酸的长度为 27 个碱基对。在一些实施方案中,可检测地标记的寡核苷酸的长度为 28 个碱基对。在一些实施方案中,可检测地标记的寡核苷酸的长度为 29 个碱基对。在一些实施方案中,可检测地标记的寡核苷酸的长度为 30 个碱基对。在一些实施方案中,可检测地标记的寡核苷酸的长度为至少 15 个碱基对。在一些实施方案中,可检测地标记的寡核苷酸的长度为至少 16 个碱基对。在一些实施方案中,可检

[0221] 在一些实施方案中,组合物还包含:

[0222] (iii) 第三寡核苷酸,所述第三寡核苷酸任选地在序列上与第一寡核苷酸相同,所述第三寡核苷酸靶向第一转录物或 DNA 基因座;以及

[0223] (iv) 第四寡核苷酸,所述第四寡核苷酸任选地在序列上与第二寡核苷酸相同,所述第四寡核苷酸靶向第二转录物或 DNA 基因座,其中第三寡核苷酸被与第一寡核苷酸不同的可检测部分标记,或第四寡核苷酸被与第二寡核苷酸不同的可检测部分标记,或两者。

[0224] 在一些实施方案中,第三寡核苷酸在序列上与第一寡核苷酸相同。在一些实施方案中,第三寡核苷酸包含与第一寡核苷酸重叠的序列。在一些实施方案中,第三寡核苷酸与第一寡核苷酸具有少于 50% 的序列同一性。在一些实施方案中,第三寡核苷酸与第一寡核苷酸具有少于 40% 的序列同一性。在一些实施方案中,第三寡核苷酸与第一寡核苷酸具有少于 30% 的序列同一性。在一些实施方案中,第三寡核苷酸与第一寡核苷酸具有少于 20% 的序列同一性。在一些实施方案中,第三寡核苷酸与第一寡核苷酸具有少于 10% 的序列同一性。在一些实施方案中,第三寡核苷酸与第一寡核苷酸具有少于 5% 的序列同一性。

[0225] 在一些实施方案中,第四寡核苷酸在序列上与第二寡核苷酸相同。在一些实施方案中,第四寡核苷酸包含与第二寡核苷酸重叠的序列。在一些实施方案中,第四寡核苷酸与第二寡核苷酸具有少于 50% 的序列同一性。在一些实施方案中,第四寡核苷酸与第二寡核苷酸具有少于 40% 的序列同一性。在一些实施方案中,第四寡核苷酸与第二寡核苷酸具有少于 30% 的序列同一性。在一些实施方案中,第四寡核苷酸与第二寡核苷酸具有少于 20% 的序列同一性。在一些实施方案中,第四寡核苷酸与第二寡核苷酸具有少于 10% 的序列同一性。在一些实施方案中,第四寡核苷酸与第二寡核苷酸具有少于 5% 的序列同一性。

[0226] 在一些实施方案中,第三寡核苷酸被与第一寡核苷酸不同的可检测部分标记。在一些实施方案中,第四寡核苷酸被与第二寡核苷酸不同的可检测部分标记。

[0227] 在一些实施方案中,多种、组合物、试剂盒或方法中的可检测地标记的寡核苷酸的量是预先确定的。在一些实施方案中,多种、组合物、试剂盒或方法中的 5% 的可检测地标记的寡核苷酸的量是预先确定的。在一些实施方案中,多种、组合物、试剂盒或方法中的 10% 的可检测地标记的寡核苷酸的量是预先确定的。在一些实施方案中,多种、组合物、试剂盒或方法中的 20% 的可检测地标记的寡核苷酸的量是预先确定的。在一些实施方案中,多种、组合物、试剂盒或方法中的 30% 的可检测地标记的寡核苷酸的量是预先确定的。在一些实施方案中,多种、组合物、试剂盒或方法中的 40% 的可检测地标记的寡核苷酸的量是预先确定的。在一些实施方案中,多种、组合物、试剂盒或方法中的 50% 的可检测地标记的寡核苷酸的量是预先确定的。在一些实施方案中,多种、组合物、试剂盒或方法中的 60% 的可检测地标记的寡核苷酸的量是预先确定的。在一些实施方案中,多种、组合物、试剂盒或方法中的 70% 的可检测地标记的寡核苷酸的量是预先确定的。在一些实施方案中,多种、组合物、试剂盒或方法中的 80% 的可检测地标记的寡核苷酸的量是预先确定的。在一些实施方案中,多种、组合物、试剂盒或方法中的 90% 的可检测地标记的寡核苷酸的量是预先确定的。

[0228] 在一些实施方案中,多种、组合物、试剂盒或方法中的至少 5 种可检测地标记的寡核苷酸的量是预先确定的。在一些实施方案中,多种、组合物、试剂盒或方法中的至少 10 种可检测地标记的寡核苷酸的量是预先确定的。在一些实施方案中,多种、组合物、试剂盒或方法中的至少 20 种可检测地标记的寡核苷酸的量是预先确定的。在一些实施方案中,多

种、组合物、试剂盒或方法中的至少 30 种可检测地标记的寡核苷酸的量是预先确定的。在一些实施方案中,多种、组合物、试剂盒或方法中的至少 40 种可检测地标记的寡核苷酸的量是预先确定的。在一些实施方案中,多种、组合物、试剂盒或方法中的至少 50 种可检测地标记的寡核苷酸的量是预先确定的。在一些实施方案中,多种、组合物、试剂盒或方法中的至少 60 种可检测地标记的寡核苷酸的量是预先确定的。在一些实施方案中,多种、组合物、试剂盒或方法中的至少 70 种可检测地标记的寡核苷酸的量是预先确定的。在一些实施方案中,多种、组合物、试剂盒或方法中的至少 80 种可检测地标记的寡核苷酸的量是预先确定的。在一些实施方案中,多种、组合物、试剂盒或方法中的至少 90 种可检测地标记的寡核苷酸的量是预先确定的。在一些实施方案中,多种、组合物、试剂盒或方法中的至少每种可检测地标记的寡核苷酸的量是预先确定的。

[0229] 在一些实施方案中,对于一种靶提供两种或更多种可检测地标记的寡核苷酸。在一些实施方案中,针对一种靶的所有可检测地标记的寡核苷酸的总量是预先确定的。在一些实施方案中,针对一种靶的所有可检测地标记的寡核苷酸的总量是预先确定的,其中针对所述靶的每种可检测地标记的寡核苷酸的量是独立地且任选地预先确定的。在一些实施方案中,针对多种靶的每种的所有可检测地标记的寡核苷酸的总量是预先确定的。在一些实施方案中,多种靶具有至少两种靶。在一些实施方案中,多种靶具有至少五种靶。在一些实施方案中,多种靶具有至少 10 种靶。在一些实施方案中,多种靶具有至少 50 种靶。在一些实施方案中,多种靶具有至少 100 种靶。在一些实施方案中,多种靶具有至少 500 种靶。在一些实施方案中,多种靶具有至少 1,000 种靶。

[0230] 在一些实施方案中,多种、组合物、试剂盒或方法的靶是预先确定的。在一些实施方案中,多种、组合物、试剂盒或方法的至少 10 种靶是预先确定的。在一些实施方案中,多种、组合物、试剂盒或方法的至少 50 种靶是预先确定的。在一些实施方案中,多种、组合物、试剂盒或方法的至少 100 种靶是预先确定的。在一些实施方案中,多种、组合物、试剂盒或方法的至少 1,000 种靶是预先确定的。在一些实施方案中,多种、组合物、试剂盒或方法的多达 F^N 种靶是预先确定的,其中 F 是多种中的可检测部分的数目,且 N 是成像步骤的数目。

[0231] 用于合成可检测地标记的寡核苷酸的方法是在领域中被广泛已知和实践的,例如,参见 Lubeck, E. & Cai, L. *Nat. Methods* 9, 743-48 (2012)。寡核苷酸也是从多个供应商商业上可得的。在一些实施方案中,本发明提供了用于制备可检测地标记的寡核苷酸的方法。在一些实施方案中,本发明提供了用于制备中间寡核苷酸的方法。在一些实施方案中,本发明提供了用于制备桥接寡核苷酸的方法。

[0232] 在一些实施方案中,本发明提供了用于制备具有第一序列的靶核酸的方法,该方法包括以下步骤:

[0233] 1) 提供包含第一序列的第一核酸,其中第一序列在两末端的侧翼为切口核酸内切酶位点;

[0234] 2) 扩增第一核酸或第一核酸的一部分,以提供第二核酸,所述第二核酸包含第一序列和侧翼切口核酸内切酶位点;以及

[0235] 3) 使第二核酸与对应于侧翼切口核酸内切酶位点的一种或更多种切口核酸内切酶接触。

[0236] 在一些实施方案中,具有第一序列的靶核酸是单链的。在一些实施方案中,扩增步

骤包括聚合酶链式反应 (PCR)。在一些实施方案中,所提供的方法还包括变性的步骤,其中双链的第二核酸变性并且双链变成单链的。在一些实施方案中,所提供的方法还包括分离具有第一序列的核酸。在一些实施方案中,在与切口核酸内切酶接触之前,第二核酸任选地被修饰。在一些实施方案中,所提供的方法还包括标记具有第一序列的核酸。

[0237] 在一些实施方案中,两个侧翼核酸内切酶位点是相同的。在一些实施方案中,使用对应于相同的切口核酸内切酶位点的一种切口核酸内切酶。在一些实施方案中,两个侧翼核酸内切酶位点是不同的。在一些实施方案中,使用两种切口核酸内切酶,所述两种切口核酸内切酶的每种独立地对应于一个切口内切核酸位点。

[0238] 在一些实施方案中,所提供的技术的寡核苷酸从寡核苷酸池 (pool) 产生。在一些实施方案中,此类池是商业上可得的。在一些实施方案中初始 DNA 寡核苷酸池由被组织成亚组的多达 12,000 种或更多种单链序列组成。每种序列被设计以使得期望的序列 (例如,探针序列) 的侧翼为切口核酸内切酶位点以及正向引物序列和反向引物序列。正向引物序列和反向引物序列对所期望的序列所属的亚组特异。引物对被用于利用聚合酶链式反应 (PCR) 扩增亚组。PCR 反应的产物被分离并通过切口核酸内切酶消化。与切口酶一起孵育的时间基于所使用的酶的量 and 回收的 DNA 的量而变化。在一些实施方案中,在约 1 小时内约 10 单位的酶消化约 1 μ g 的 DNA。然后在缓冲液例如 2x 加载缓冲液 (96% 甲酰胺 /20mM EDTA) 和水中纯化和重构样品,以制备最终的加载缓冲液 (48% 甲酰胺 /10mM EDTA) 并且变性,例如,通过加热至 95°C 以使 DNA 完全变性。纯化变性的 DNA 并且分离所期望的产物。在一些实施方案中,纯化和 / 或分离包括电泳。在图 25 中图示了示例性过程。

[0239] 在一些实施方案中,本发明提供了用于制备具有第一序列的靶核酸的方法,所述方法包括以下步骤:

[0240] 1) 提供包含第一序列或其互补序列的第一核酸,其中第一序列或其互补序列的侧翼为至少一个限制性位点;

[0241] 2) 扩增第一核酸或第一核酸的一部分,以提供第二核酸,所述第二核酸包含第一序列和至少一个侧翼限制性位点;以及

[0242] 3) 使第二核酸与对应于至少一个侧翼限制性位点的限制性酶接触,以提供包含凹陷末端的第三核酸;

[0243] 4) 使第三核酸与核酸酶接触,以选择性地消化包含互补序列 (如果存在) 的链,同时保留包含第一序列的链。

[0244] 在一些实施方案中,第一序列或其互补序列在每个末端侧翼独立地为限制性位点。

[0245] 在一些实施方案中,本发明提供了用于制备具有第一序列的靶核酸的方法,所述方法包括以下步骤:

[0246] 1) 提供包含第一序列或其互补序列的第一核酸,其中第一序列或其互补序列在两端侧翼均为限制性位点;

[0247] 2) 扩增第一核酸或第一核酸的一部分,以提供第二核酸,所述第二核酸包含第一序列和侧翼限制性位点;以及

[0248] 3) 使第二核酸与对应于侧翼限制性位点的限制性酶接触,以提供包含凹陷末端的第三核酸;

[0249] 4) 使第三核酸与核酸酶接触,以选择性地消化包含互补序列(如果存在)的链,同时保留包含第一序列的链。

[0250] 在一些实施方案中,具有第一序列的靶核酸是单链的。在一些实施方案中,扩增步骤包括PCR。在一些实施方案中,所提供的方法还包括分离具有第一序列的核酸。在一些实施方案中,在与限制性酶接触之前,第二核酸任选地被修饰。在一些实施方案中,在与核酸酶接触之前,第三核酸任选地被修饰。在一些实施方案中,核酸酶是核酸外切酶 III,其优先降解带有 3' - 凹陷末端的链,并且可以保留带有 5' - 凹陷末端的链。在一些实施方案中,限制性酶产生 5' - 凹陷末端。在一些实施方案中,限制性酶产生 3' - 凹陷末端。在一些实施方案中,在限制性消化之后,互补序列具有 3' 凹陷末端。在一些实施方案中,在限制性消化之后包含互补序列的链具有 3' 凹陷末端,并且在限制性消化之后包含第一序列的链具有 5' 凹陷末端。在一些实施方案中,所提供的方法还包括标记具有第一序列的核酸。

[0251] 在一些实施方案中,单链的寡核苷酸,例如用于 seqFISH 的探针或中间寡核苷酸,可利用核酸酶消化诸如 *exoIII* 核酸酶消化来产生。可以使用扩增(例如,PCR)产物上的两个限制性位点而非两个切口位点位于探针和/或适体序列的侧翼。在一些实施方案中,一个限制性位点留下 3' 凹陷末端同时另一个留下 5' 凹陷末端。例如, *EcoRI* 和 *BamHI* 留下 5' 凹陷末端,同时 *BmtI* 和 *PacI* 留下 3' 凹陷末端。此类限制性酶是领域中广泛已知和使用的。核酸外切酶 III 优先降解 3' 凹陷末端,并保留带有 5' 凹陷末端的链。这提供了利用 PCR 和限制性核酸酶从寡核苷酸池产生单链探针的另一种机制。

[0252] 在一些实施方案中,所提供的靶核酸是 DNA。在一些实施方案中,靶核酸具有与第一序列相同的序列。在一些实施方案中,靶核酸是中间寡核苷酸,包含与靶杂交的第一序列和与第二寡核苷酸杂交的第二序列,所述靶例如转录物或 DNA 基因座,所述第二寡核苷酸例如可检测地标记的寡核苷酸。在一些实施方案中,靶核酸是中间寡核苷酸,包含与靶杂交的第一序列和与由 HCR 标记的可检测地标记的寡核苷酸杂交的第二序列。在一些实施方案中,靶核酸是桥接探针。

[0253] 在一些实施方案中,所提供的方法被用于诊断疾病,其中所述疾病与转录物或 DNA 基因座的异常数目有关。在一些实施方案中,所提供的方法被用来选择治疗的受试者。在一些实施方案中,所提供的方法被用来监控治疗方案。在一些实施方案中,所提供的方法中的细胞来自受试者。在一些实施方案中,所提供的方法中的细胞是哺乳动物细胞。在一些实施方案中,所提供的方法中的细胞是人类细胞。在一些实施方案中,所提供的方法中的细胞来自受试者。在一些实施方案中,所提供的方法中的细胞来自动物。在一些实施方案中,所提供的方法中的细胞来自人类受试者。在一些实施方案中,所提供的方法中的细胞从人类受试者分离。在一些实施方案中,所提供的的方法中的细胞来自病变组织或对疾病易感的组织。由于能够同时检测并且定量大量的靶,所提供的方法对诊断、治疗监控和患者分层提供了显著的优势。

[0254] 在一些实施方案中,所提供的技术任选地包括对蛋白、神经活动和/或结构布置谱分析。在一些实施方案中,所提供的方法包括对同一样品中的蛋白质谱分析。在一些实施方案中,所提供的方法包括对同一样品中的神经活动谱分析。在一些实施方案中,所提供的方法包括对结构布置谱分析。

[0255] 如同本文所示例的,所提供的技术适用于很多种样品。例如,HCR-seqFISH 在脑切

片中有效并且 SPIM 可以粗略地检测 CLARITY 脑切片中的单一 mRNA。在一些实施方案中,所提供的技术可用于对神经退行性疾病的小鼠模型或人类大脑中的靶谱分析。在本发明之前没有其他的技术能够达到同样质量和数量的数据。

实施例

[0256] 上述内容已经描述了本发明的某些非限制性实施方案。相应地,应理解本文描述的本发明的实施方案仅仅例证了本发明的原理的应用。本文详细描述的例子性实施方案并非意图限制权利要求的范围。

[0257] 通过顺序杂交和编条形码对核酸原位谱分析

[0258] 如在本文的非限制性实例中所描述的,通过所提供的方法,通过连续数论的接触、成像和去除步骤对细胞中的核酸例如 mRNA 谱分析(图 2(a) 和图 3)。由于在细胞中转录物被固定,在多次杂交期间对应的荧光斑点保持在原位,并且可以被对齐以读取荧光团序列。此顺序条形码被设计以独特地鉴定 mRNA。

[0259] 在每轮杂交期间,每种转录物被一组可检测地标记的寡核苷酸(在该情况中,FISH 探针被单一类型的荧光团标记)靶向。对样品成像并且随后用 DNA 酶 I 对其处理以去除 FISH 探针。在随后的轮次中,使 mRNA 与具有相同的寡核苷酸序列但现在被不同的染料标记的 FISH 探针杂交。可用的条形码的数目规模为 F^N , 其中 F 是荧光团的数目且 N 是杂交轮次数。例如,用 4 种染料,8 轮杂交可以覆盖几乎整个转录组 ($4^8 = 65,536$)。

[0260] 作为例证,用 4 种染料和 2 轮杂交对单一酵母细胞中的 12 个基因编条形码 ($4^2 = 16$, 4 个条形码被忽略;每次杂交进行三个循环)。细胞被固定在玻璃表面上。DNA 探针被杂交、成像并且随后通过 DNA 酶 I 处理去除 ($88.5 \pm 11.0\%$ (SE) 的效率,图 4)。剩余的信号被光漂白(图 5)。甚至在 6 次杂交之后,观察到 mRNA 为原始强度的 $70.9 \pm 21.8\%$ (SE) (图 6)。观察到在前两次杂交中共定位的点的 $77.9 \pm 5.6\%$ (SE) 也与第三次杂交共定位(图 7 和图 8)。通过计数细胞中对应的条形码的出现来定量 mRNA 丰度(图 9 和图 10, $n = 37$ 个细胞)。显示在哺乳动物细胞中 mRNA 可以被有效地剥离并再杂交(图 11 和图 12)。如本文证实的,所提供的方法具有超越本发明之前的已知方法的许多优点。例如,所提供的方法快速地放大规模;即使使用两种染料,编码能力在原则上是无限的 (2^N)。在每个接触步骤期间,可使用增加信号的亮度的针对探针的所有可用的可检测地标记的寡核苷酸(在此实施例中,FISH 探针)。所提供的方法的读数也是稳健的并且使能够对天然样品充分 Z-叠加(Z-stacks)。所提供的方法可以利用可检测地标记的寡核苷酸的高杂交效率的优势,所述可检测地标记的寡核苷酸诸如 FISH 探针 (>95% 的 mRNA 被检测出;Lubeck, E. & Cai, L. Nat. Methods 9, 743-48 (2012))。申请人注意到还可以设计可检测地标记的寡核苷酸例如 FISH 探针,以分辨单一细胞中的大量剪接异构体、SNP 以及染色体基因座 (Levesque, M. J. & Raj, A. Nat Meth 10, 246-248 (2013))。与超分辨率方法组合 (Lubeck, E. & Cai, L. Nat. Methods 9, 743-48 (2012)), 所提供的方法使复合样品诸如大脑中的大量靶(例如转录组)能够被以单细胞分辨率直接地成像。

[0261] 方法和程序

[0262] 样品制备:使 MDN1-GFP 酵母细胞在补充有 50mM CaCl_2 的 YPD 中生长,至 OD 为 0.3。将细胞在 1% 甲醛 5% 醋酸中固定 5 分钟,在缓冲液 B 中冲洗 3X 并在 30°C 下去壁

(spheroplasted) 持续 1 小时。将细胞在 70% EtOH 中在 -20℃ 下贮存长达两周。

[0263] 通过以下准备盖玻片:用 1M NaOH 和 100% EtOH 的交替溶液超声 3X, 随后在丙酮中超声最终一轮。在丙酮中制备 2% 的 (3-氨基丙基) 三乙氧基硅烷 (Sigma 440140) 溶液并且将经清洁的盖玻片立即浸入该溶液中持续两分钟。冲洗胺改性的盖玻片并将其在室温下在超纯水中贮存。

[0264] 在 23℃ 下用 0.5U/μL 的 DNA 酶 I 溶液 (Roche04716728001) 预处理被固定的酵母细胞, 持续 30 分钟。处理之后, 通过在两片胺改性的盖玻片之间物理地按压酵母的稀释溶液使酵母细胞粘附至被包被的盖玻片。随后小心将盖玻片分开 (pealed apart) 并且将其立即浸入 1% 甲醛溶液中持续 2.5 分钟。固定之后将盖玻片干燥并且通过使包被粘附剂的流动细胞粘附至盖玻片来构建流动胞 (GraceBio Labs SA84-0.5-SecureSeal)。将 FluoSphere365nm 荧光珠添加至盖玻片来测量多次杂交的漂移 (Life F8805)。将用封口膜覆盖的流动细胞在 4℃ 下贮存。

[0265] 可检测地标记的寡核苷酸的制备: 根据 Lubeck, E. & Cai, L. Nat. Methods9, 743-48 (2012) 中的方法制备探针。对于每种靶, 使用 24 种探针。将针对每组基因的所有 24 种探针与所使用的四种染料 Alexa 532、594、Cy5 和 Cy7 之一偶联。

[0266] 杂交: 使流动细胞在 10% 硫酸葡聚糖 (Sigma D8906)、10% 甲酰胺和 2X SSC 的杂交缓冲液中在 2nM/ 探针的浓度下杂交过夜。杂交之后, 将样品在添加至室温样品前预热至 37℃ 持续 10 分钟的 30% 甲酰胺、0.1% Triton-X100 缓冲液中洗涤。用 2X SSC 对样品洗涤数次以去除扩散的探针。

[0267] 成像: 将样品浸入抗漂白缓冲液 (Swoboda, M. ACS Nano6, 6364-69 (2012)): 20mM Tris-HCl、50mM NaCl、0.8% 葡萄糖、饱和 Trolox (Sigma: 53188-07-1)、 OD_{405nm} 为 0.05 的吡喃糖氧化酶 (Sigma P4234) 和 1/1000 稀释的过氧化氢酶 (Sigma: 9001-05-2)。

[0268] 探针置换: 成像之后, 在 DNA 酶 I 缓冲液 (Roche) 中洗涤细胞并且允许细胞在 0.5U/μL DNA 酶 I (Roche) 中静置 4 小时。为抑制 DNA 酶, 用 30% 甲酰胺、0.1% Triton-X100、2X SSC 对细胞洗涤两 X。DNA 酶处理之后在抗漂白缓冲液中对细胞再成像一次, 以确定 DNA 酶 I 探针剥离比率。为了去除剩余的探针信号, 在所有成像通道中将样品用 10 秒的激发漂白并且用标准激发时间再一次成像以记录残留信号。

[0269] 再杂交: 根据前面概述的条件使样品在显微镜上再杂交。在显微镜上杂交期间用封口膜覆盖样品以防止蒸发。

[0270] 对同一样品进行至少六轮杂交。每轮杂交在显微镜下过夜发生, 并在日间进行 DNA 酶处理和成像。在对应于此所应用的迭代杂交方案中, 利用两轮杂交对 mRNA 编条形码。随后重复条形码方案, 以使得使用相同的探针进行 hyb1 和 hyb3, 而用另一组探针进行 hyb2 和 hyb4。Hby1 和 hyb3 之间的共定位给出了对于所检测的转录物的校准, 而 hyb1 和 hyb2 产生编条形码的数据。

[0271] 数据分析: 用 ImageJ、Python 和 Matlab 进行数据分析。由于样本在实验期间漂移, 利用从每个成像位置的 DAPI 通道确定的基于交叉关联的配准方法使原始图像对齐。然后将漂移校正传播到对应于相同位置的其他 4 个颜色通道。然后对图像去卷积来降低邻近的 FISH 点之间的重叠。个体通道中的点重叠很少被观察到, 但是当图像被叠加时, 在不同通道中的点会以其点扩散函数 (PSF) 重叠。基于迭代 Lucy-Richardson 算法 (Lucy,

L. B. The Astronomical Journal. 79, 745 (1974) 和 Richardson, W. H. J. Opt. Soc. Am. 62, 55-59 (1972)) 来处理原始数据。显微镜的 PSF 通过对在显微镜的 DAPI 通道中测量的珠图像 (~ 200nm 直径) 求平均值来评估。利用该测量的点扩散函数连同 Lucy-Richardson 算法, 在计算该过程超过 20 次迭代之后, 我们对在 FISH 图像中的荧光发射分布进行最大似然性评价。该去卷积方法的输出提供了分辨的 FISH 数据且增加了条形码分配的精确度。

[0272] 利用局部最大值函数来鉴定对应于图像中的 FISH 信号的点。低于阈值的点被丢弃, 不用于进一步分析。通过优化与相同的探针组杂交的 hyb1 和 hyb3 图像之间的共定位来确定阈值的值。将每个 PSF 的最大强度像素用作该 mRNA 分子的位置的代表。

[0273] 从对应于 hyb1 和 hyb2 中的 mRNA 的点自动提取条形码。算法计算了在 hyb1 中鉴定的每个点和在 hyb2 中鉴定的所有点之间的成对距离。对于 hyb1 中的每个点, 在 hyb2 中的最近的邻点被鉴定出。如果该距离是 0 像素或 1 像素并且 hyb2 中的点的最近的邻点也是 hyb1 中的原始点, 则条形码对被确认。对称的最近邻点的要求减少了条形码的假分配。为减少 cy7 的假阳性, 要求在 hyb1 中检测到的点 cy7 在 hyb3 中以 cy7 重复出现。

[0274] 在该非限制性实施例中, 由于 DNA 酶 I 的低成本和快速活性, 申请人用 DNA 酶 I 去除探针。申请人注意到将探针从 mRNA 去除并使其完整的任何方法均可被用于所提供的编条形码方法, 例如但不限于, 链置换 (Duose, D. Y. Nucleic Acids Research, 40, 3289-3298 (2012)) 和高温洗涤或甲酰胺洗涤。申请人注意到 DNA 酶 I 不要求从标准 smFISH 探针重新设计探针并且不因严苛的洗涤干扰样品。

[0275] 在一些实施方案中, 观察到来自 dsDNA 的 DAPI 信号在几秒钟内快速消失, 而 smFISH 探针需要质的较长的时间段 (数十分钟) 来被降解。不意图受理论限制, 相对于 dsDNA 裂解速率, DNA 酶 I 的探针去除效率是低的。在短时间内去除过程仍然被观察到。

[0276] 在某些实验中, 在 DNA 酶 I 处理后, 11.5% 的荧光信号保持在 mRNA 上。通过漂白, 剩余的信号几乎降低至零。申请人注意到在光漂白之前可达成更完整的信号和 / 或探针去除, 以使得更多的 mRNA 可被用于之后轮次的杂交。申请人注意到光漂白对于编条形码并非必需的, 但在一些实施方案中, 通过去除可能在进一步的编条形码轮次中给出假阳性的残留信号, 光漂白确实简化了过程。与 mRNA 的结合的 11.5% 的残留探针中的一些可抑制进一步的杂交轮次。申请人注意到由于所展示的 5 次杂交的数据在杂交效率上显示出轻微下降, 残留探针不显著抑制进一步的杂交轮。

[0277] 对脑组织中的核酸谱分析

[0278] 对完整脑切片中的细胞的转录谱分析对于理解细胞身份的分子基础是重要的。然而, 本发明之前, 在天然神经网络的解剖学背景下对单一细胞中的转录物丰度和定位定量地谱分析在技术上是困难的。皮层躯体感觉子网络被用作实例来说明示例性提供的技术的可行性和实用性, 例如, 利用 FISH 原位测序 (seqFISH) 和连接组学 (connectomics) 对不同功能区域内的不同的神经元群体中的多种基因谱分析, 所述不同功能区域诸如在初级躯体感觉皮层区 (SSp)、初级躯体运动皮层区 (MOp)、次级躯体运动皮层区 (MOs) 和辅助躯体感觉皮层区 (SSs) 中的那些区域。

[0279] 如本文广泛描述的, 在一些实施方案中, 本发明提供了经由通过 FISH 原位“测序”对单一细胞中的基因表达谱分析的技术, 例如如由图 1 和图 2 所图示。为了检测个体 mRNA, 单分子荧光原位杂交 (smFISH), 和与 mRNA 序列互补的 20mer 寡核苷酸探针

一起被使用 (Fan, Y. Braut, SA, Lin, Q., Singer, RH, Skoultchi, AI. Determination of transgenic loci by expression FISH. *Genomics*. 2001 年 10 月 2 日; 71(1):66-9; Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S. Stochastic mRNA synthesis in mammalian cells. *PLoS Biol*. 2006 年 10 月; 4(10):e309)。通过将 24 种此类用荧光团标记的探针放置在 mRNA 上, 细胞中的单一转录物变得可容易地原位检测。显示出通过 smFISH, 几乎所有的可被检测出的 mRNA 被观察到 (Lubeck, E. L. Cai. Single cell systems biology by super-resolution imaging and combinatorial labeling. *Nature Methods* 9, 743-48 (2012))。所提供的方法是高度定量的且保留了组织样本内的空间信息, 而不物理地分离单一细胞或使用匀浆。

[0280] 在一些实施方案中, 为了区分不同的 mRNA 种类, 利用顺序的杂交轮用可检测地标记的寡核苷酸诸如 FISH 探针对 mRNA 编条形码。在一轮杂交期间, 每种转录物被用单一类型的荧光团标记的一组多种探针例如 24 种 FISH 探针靶向。对样品成像并且通过酶促消化去除 FISH 探针。然后在随后的轮次中使 mRNA 与相同的 FISH 探针 (但在一些情况下, 现在被不同的染料标记) 杂交。由于转录物被固定在细胞中, 所以在多轮杂交期间, 对应于单一 mRNA 的荧光点保持在原位, 并且可以被对齐以读取颜色序列。因此每种 mRNA 种类被分配独特的条形码。给定的细胞中的每种转录物的数目可以通过计数对应的条形码的数目来确定。在图 1、2 或 3 中图示了示例性方法。

[0281] 所提供的技术可以利用 FISH 的高杂交效率 (> 95% 的 mRNA 被检测出) 的优势。在一些实施方案中, 对于鉴定转录物, 碱基对分辨率是不需要的, 尽管如果期望, 碱基对分辨率是可达到的。所提供的方法的可用的条形码的数目规模为 F^N , 其中 F 是不同的荧光团的数目且 N 是杂交轮次的数目。用 5 种不同的染料和 3 轮杂交, 可对 125 种独特的核酸谱分析。通过 6 轮杂交 ($5^6 = 15,625$) 几乎整个转录组可以被覆盖, 例如利用分辨细胞中的所有转录物的超分辨率显微术。在一些实施方案中, 使用常规显微术来检测较少但仍然大量的靶, 诸如 100 种基因多路复用 (multiplex), 所述常规显微术诸如常规的落射荧光显微术, 其实施起来简单且稳健。

[0282] 在多个杂交循环中, 探针可被剥离并与相同的 mRNA 再杂交 (图 2)。许多商业上可得的荧光团效果稳健, 诸如对于编条形码给出至少 7 种颜色的 AlexaFluor 488、532、594、647、700、750 和 790。即便在 6 轮杂交结束时, 探针可以以原始强度的 $70.9 \pm 21.8\%$ 与剥离的 mRNA 再杂交 (图 6)。作为例证, 用 4 种染料和 2 轮杂交, 被编条形码的 12 种基因被在单一酵母细胞中编条形码 ($4^2 = 16$, 图 3, C)。

[0283] 由于 12 种基因的多路复用图像只占据每个荧光通道的 5% 的光学空间, 因此在细胞中存在足够的光学空间进行多种例如 100 种基因的多路复用。虽然图 3 中的所有 4 个荧光通道的复合图像看起来稠密, 但是每个荧光通道中的点是分散地分布的。可以用 2 维度高斯曲线对每个点拟合来提取其图心位置并进一步降低与其他荧光通道中的点的重叠。显示在不同的杂交轮次之间相同点重新对齐至 100nm (图 8)。

[0284] 在一些实施方案中, 100 种基因多路复用可以用 3 轮杂交快速地进行。在一些实施方案中, 每个杂交循环包括约 4 小时的杂交、约 1 小时的成像和约 1 小时的 DNA 酶处理和洗涤, 每次的时间长度可以任选且独立地变化。在一些实施方案中, 3 轮杂交花费约 18 小时。在一些实施方案中, 限速步骤是成像时间, 而不是杂交时间, 因为一个脑切片可被成像同时

在同一显微镜上的另一切片杂交。在一些实施方案中,单个显微镜可以 8 个切片同时地多路复用并且在 18 小时内三个杂交循环结束时对所有 8 个切片获得 100 种基因数据。

[0285] 在一些实施方案中,在显微镜上含有 10^6 个细胞的 $10\text{mm} \times 5\text{mm} \times 10 \mu\text{m}$ 的脑切片可以在 35 分钟内被成像并分析。在一些实施方案中,用 $20\times$ 气体目镜 (airobjective) 和 $13\text{mm} \times 13\text{mm}$ 摄像头芯片,显微镜上的单个视野 (FOV) 为 $0.5\text{mm} \times 0.5\text{mm} \times 2 \mu\text{m}$ 。在一些实施方案中,每个 FOV 以 100 毫秒曝光并读取。在一些实施方案中,以 XYZ 并在不同的颜色通道中扫描样品引起在快照之间 200 毫秒的时间延迟。在一些实施方案中,整个大脑切片可以在 2000 秒或 35 分钟内被成像。对于 100 种基因多重化所需的 3 轮杂交,总的成像时间是 105 分钟。在一些实施方案中,在一台显微镜上整个小鼠大脑可以在 30 天内被成像。当使用多路显微镜时,时间帧可被进一步缩短。在一些实施方案中,所提供的方法可以用 500 个切片对整个小鼠大脑成像,花费少于 \$25,000。

[0286] 与本发明之前已知的其他方法相比,所提供的技术提供了多种优势。除了其他以外,所提供的技术是定量的、保留空间信息并便宜地扩大至整个组织、器官和 / 或生物体规模。

[0287] 与本发明之前的单细胞 RNA-seq 比较。不像单细胞 RNA-seq 或 qPCR,需要将单细胞分离并放入 96 孔的格式中,所提供的方法诸如 seqFISH 可以用自动显微术以少的额外花费在细胞的天然解剖学背景中扫描大量细胞。用微流体装置来达到相同水平的通量将会是经济上不可行的且劳动密集的。在一些实施方案中,所提供的技术的主要花费是探针合成的初始花费,所述初始花费被以下事实抵消:探针被合成后,它们可以被用于很多,例如 1000 到 10,000 次或甚至更多的反应。

[0288] 所提供的方法诸如 seqFISH 基于单分子 FISH 并且可以以绝对定量测量低拷贝数的转录物。用这种方法获得的数据是高度地定量的且使能够进行高质量的统计分析。相比而言,目前的单细胞 qPCR 和 RNA-seq 实验因自反转录 (RT) 误差和 PCR 扩增误差的偏倚而在定量能力方面有。

[0289] 与本发明之前的其他原位测序方法比较。smFISH 方法的一个主要优点是被靶向的几乎所有 mRNA 可被观察到。经确定结合在 mRNA 上的每种 FISH 探针的效率是 50 % -60 % (Lubeck, E. & Cai, L. Nat. Methods 9, 743-48 (2012); Levesque, M. J. & Raj, A. Nat Meth 10, 246-248 (2013))。靶向多种,例如每 mRNA 24 种至 48 种探针保证了至少 10 种探针在几乎每个 mRNA 上杂交,在非特异性背景下提供良好信号。用 FISH 探针直接地探测 mRNA 是高度地特异的并且确保了所有的转录物被检测到。

[0290] 相反,在通过测序反应检测 DNA 模板之前,许多其他原位测序方法利用酶促反应将 mRNA 转化成 DNA 模板,而不是直接靶向 mRNA。然而,mRNA 至 DNA 的转化过程是高度地无效率的,并且只有小比例的 RNA 被转化并被检测出。低效率 (对于反转录 (RT) 被评估为 1% 并且和对于锁式连接 (padlock ligation) (PLA) 被评估为 10%) 的一个示例性主要缺点是其可在基因表达测量中引入显著的噪音和偏倚。

[0291] 假定典型的细胞大小为 ($10-20 \mu\text{m}^3$),在该细胞中存在约 25,000 个衍射限制的点。在一些实施方案中,对于单细胞中的转录物检测这是可用的固定物 (realestate)。在 seqFISH 中,所选择的一组基因,诸如转录因子 (TF) 和细胞粘附分子 (CAM),可以高精度度被成像和定量。如果选择了具有每基因 100 个转录物的平均拷贝数的靶基因,则可以进行

高度地定量的 100-200 个基因的谱分析实验。相反,对于很多其他原位测序方法,固定物的大部分被用来对核糖体 RNA 以及管家基因测序;感兴趣的基因诸如对神经细胞身份特异的那些基因是严重不足且难以检测的。

[0292] 在一些实施方案中,所提供的方法使用杂交链式反应(HCR)(Choi, 等人, Programmable in situ amplification for multiplexed imaging of mRNA expression *Nature Biotechnol*, 28, 1208-1212, (2010)) 来放大 FISH 信号,所述 FISH 信号允许用 20X 气体目镜得到大 FOV 成像,但同时保留了 smFISH 的高检测效率。

[0293] 与本发明之前的对多路复用 RNA 编超分辨率编条形码的方法比较。在一些实施方案中,与本发明之前的通过 smFISH 对 mRNA 谱编条形码(Femino 等人, Visualization of single RNA transcripts in situ. *Science*. 1998 年 4 月 24 日; 280(5363):585-90; Kosman 等人, Multiplex detection of RNA expression in *Drosophila* embryos. *Science*. 2004 年 8 月 6 日; 305(5685):846; Levsky 等人, Single-cell gene expression profiling. *Science*. 2002 年 8 月 2 日; 297(5582):836-40; Lubeck 等人, Single cell systems biology by super-resolution imaging and combinatorial labeling. *Nature Methods* 9, 743-48(2012); 和 Levesque 等人, *Nat Meth* 10, 246-248(2013)) 相比,所提供的方法具有很多优点,其中针对特定 mRNA 的探针被分成被不同染料标记的亚组。除了其他以外,所提供的技术不需要很多不同的荧光团来放大规模;即便用两种染料,编码容量是巨大的,并且可以使用重复的条形码(例如,红色-红色-红色)。比较而言,本发明之前的对 RNA 谱编条形码在可产生的条形码的数目上受限制(~ 30)。在提供的方法中,在每轮杂交期间,所有可检测地标记的寡核苷酸诸如针对转录物的 FISH 探针可以被立即使用,而不是将探针划分成亚组。除了其他以外,由于每个 mRNA 上的信号较强,所提供的技术提供了条形码读取的改善的稳健性。与本发明之前的方法相比,由于每个 mRNA 可以具有较少的颜色,在一些实施方案中,在每轮杂交期间具有单种颜色,而不是在本发明之前谱条形编码方案中的至少三种颜色,所以图像中目标的密度较低。在超分辨率显微术是必要的之前,如果期望,则较低的图像密度可极大地简化数据分析并且允许更多的基因被多路复用。申请人注意到在所提供的实施方案中,某些谱编条形码的方法、探针、和/或超分辨率显微术可被使用并且可以是有用的实施方案。在一些实施方案中,为了用所提供的技术诸如 seqFISH 对转录组谱分析,超分辨率显微术被用于分辨细胞中的数百万的转录物。

[0294] 除了转录物的谱分析,所提供的技术可以分辨对于相同的 mRNA 分子的多种可选剪接选择和 RNA 编辑。由于测序读段通常太短以致无法与相同转录物内的外显子关联,可选剪接异构体难以通过测序方法探测。所提供的方法诸如 seqFISH 允许大脑切片中的个体细胞内的全部剪接异构体直接可视化。类似地,检测单核苷酸多态性(SNP)的 smFISH 方法可以适用于 seqFISH,以对神经元或其他细胞类型中的被编辑的转录物成像。

[0295] 在一些实施方案中,所提供的技术提供了用于以下的高效和有成本效益的流水线(pipelines):通过顺序 FISH(seqFISH)、和整体 seqFISH(integrate seqFISH) 和连接组学的基因原位谱分析,以对皮层中的躯体运动神经元谱分析,来鉴定对应于细胞身份的组合物分子标志物。

[0296] 大脑内基因表达映射的原位定量

[0297] 光片照明显微术被应用于对 CLARITY 清晰化的大脑切片直接成像。在一些实施方

案中,每台机器 1 个月内映射一个小鼠大脑。在一些实施方案中,用 4-5 台机器在一周内映射一个小鼠大脑。

[0298] 放大:FISH 信号的放大允许用 20X 低 NA 物镜实现对大脑切片大 FOV 成像。在一些实施方案中,所提供的方法使用被杂交链式反应 (HCR) 标记的可检测地标记的寡核苷酸 (Choi 等人,2010) 来增加信背比 (signal-to-background) 和 / 或保留 FISH 方法的特异性和多路复用能力。通过这种方法,与 mRNA 靶互补的核酸探针触发了链式反应,在所述链式反应中亚稳态的荧光标记的核酸发夹自组装成栓荧光放大聚合物 (tethered fluorescent amplification polymers)。利用携带光谱不同的荧光团的正交 HCR 放大器,可对所有通道同时进行原位放大。

[0299] 在一些实施方案中,带有 HCR (HCR 探针) 的可检测地标记的寡核苷酸包含与靶 mRNA 互补的 20-nt 区域,加上 40-nt 的 HCR 引发物。探针的杂交在 10% 甲酰胺的严格度下进行,随后在允许条件下进行扩增步骤。可以对如发夹浓度的条件进行优化以达到最佳结果;申请人指出,在某些情况中,较高浓度的发夹增加反应速率。每个 HCR 探针可以被扩增至衍射极限斑点。在一些实施方案中,在衍射极限点的规模内,FISH 信号被放大约 10-20 倍。点亮度可以被进一步增强,同时通过例如在每个探针内并入多个 HCR 引发物和 / 或用多个荧光团标记每个 HCR 扩增发夹来保持衍射极限点规模。

[0300] 在大脑切片中,从 mRNA 直接地观察到 HCR 放大的信号。当靶向相同的 mRNA 时,HCR 探针以 90% 的比率与 smFISH 点共定位,但是要亮 10-20 倍 (图 14)。这允许 HCR 探针超越大脑的自发荧光被容易地检测到 (图 13)。高定位比率证明了 HCR 是与 smFISH 一样特异的并且大部分转录物被检测到。

[0301] HCR 探针容易地被剥离并再杂交,并且可以与本文描述的 seqFISH 方案充分地整合。图 15 显示了在大脑切片中在两轮不同的杂交中 HCR 靶向的相同基因。两次杂交之间的良好的共定位显示了 HCR-seqFISH 稳健地对大脑中的 mRNA 编条形码。

[0302] HCR 方案与 smFISH 杂交花费相同的时间范围并且不增加测定的循环时间。HCR 中的初始杂交步骤在时间上与 smFISH 相似,而第二扩增步骤在 30 分钟与 1 小时内发生。使 RNA 探针与转录物杂交和任选地使用可选的发夹类型来放大信号的可选方法能够进一步减少循环时间。在一些实施方案中,HCR 消除了对购买用胺标记的寡核苷酸探针的需求。除了其他以外,HCR 能够潜在地降低约一半的试剂花费,例如,至每大脑 \$10,000。

[0303] 自动化。硬件和软件两者的自动化可被应用于有效地扩大规模,例如,映射 100 种基因和 / 或减少人类劳力和 / 或错误。在一些实施方案中,技术的关键部分被整合以产生用于组织和 / 或器官成像,诸如大脑切片的成像的流水线和 / 或优化工作流程。除了其他以外,自动化射流 (fluidics)、图像采集和 / 或整合分析可以独立地且任选地与快速杂交循环时间和成像时间组合。

[0304] 硬件。在一些实施方案中,自动化系统需要最小的来自用户的干预并且在用户已设定实验后可自动地执行图像采集。在一些实施方案中,每个测序仪 (sequencer) 由进行成像的自动化落射荧光显微镜和进行连续杂交的自动化射流系统组成或包括进行成像的自动化落射荧光显微镜和进行连续杂交的自动化射流系统。在一些实施方案中,压缩空气被用来将试剂推进带有固定在底部盖玻片上的细胞和组织的 1 cm×1 cm 孔中。不意图被理论限制,申请人指出在一些实施方案中,由于杂交缓冲液的高粘度,压缩空气驱动系统消

除了死体积并且还可被精确控制以递送限定体积的试剂。在一些实施方案中,单独的真空线被用于净化室。在一些实施方案中,所提供的方案的工作流程与在本发明的时期本领域所熟知的现有 DNA 测序仪相似。

[0305] 在一些实施方案中,在每个杂交循环期间,机器自动地使样品与探针杂交、用缓冲液洗涤以去除过量的探针和 / 或扫描大脑切片用于成像。在一些实施方案中,其中在去除步骤中使用 DNA 酶,成像之后 DNA 酶流入以去除探针。在彻底的洗涤之后,随后可进行另一轮杂交。在杂交时间期间,显微镜移动至镜台上的不同位置以对已经杂交和洗涤的另一个大脑切片成像。以如此方式,大部分时间相机在采集图像,同时镜台上的其他样品在杂交。

[0306] 软件。在一些实施方案中,使用软件,例如来自自动化控制过程和分析数据。在一些实施方案中,将在 Micromanager (由 National Institute of Health 支持的免费软件) 中编写代码,来控制显微镜以及射流元件。在一些实施方案中,通过 Micromanager 来控制阀门、镜台、光源、相机和 / 或显微镜。

[0307] 在一些实施方案中,压缩传感被用于密集的图像 (Zhu 等人, Faster STORM using compressed sensing. Nat. Methods. 2012, 9(7):721-3) 并且去卷积方法被用于将密集的簇中的点区分开。在一些实施方案中,在图像分析上的改进提高了所提供方法例如 seqFISH 的多路复用能力 (例如,超出 100 种基因多路复用的约 4-5 倍)。在一些实施方案中,以相似的方式从 Illumina GAI 测序仪改进至 HiSeq 仪器来提高效率,其中利用图像处理方法来分析测序芯片上的密集堆积的簇增加了通量。在一些实施方案中,数据采集和分析被整合在用户友好的包内。

[0308] 在一些实施方案中,所提供的技术提供了用于数据分析的软件包。在一些实施方案中,所提供的技术提供了用于在 Python 和 Matlab 中分析数据的软件包。所提供技术的图像可以是多种尺寸的,并且如果期望能任选地被优化。在一些实施方案中,每个 FOV 为 14 位深的 6 兆像素,对应于每个图像 1.5MB 的数据。在一些实施方案中,每个运行产生约 100GB 的数据。在一些实施方案中,所提供的技术提供了用于数据处理的 / 或缓解数据阻塞 (log jam) 的方法。在一些实施方案中,数据阻塞通过以下缓解:从每个图像分割出点,用 2 维高斯分布对其拟合,并记录拟合点 (fits) 的中心位置。在一些实施方案中,所提供的技术通过丢弃原始图像并保存处理的数据来节省计算机空间。

[0309] 用 CLARITY 清晰化的大脑切片的光片照明显微术。在一些实施方案中,所提供的技术提供了用于对组织、器官和 / 或有机体成像的方法。在一些实施方案中,所提供的技术提供了用于测量厚的组织或器官的方法。在一些实施方案中,厚的组织或器官具有约 100 μm 或大于 100 μm 的厚度。在一些实施方案中,所提供的技术保留了超出了单一细胞内的远程投射和形态。在一些实施方案中,光片照明显微术被用来测量厚的组织或器官。在一些实施方案中,组织、器官和 / 或生物体通过 CLARITY 被清晰化 (Chung 等人, Structural and molecular interrogation of intact biological systems, Nature, 2013, doi:10.1038/nature12107)。

[0310] 在一些实施方案中,为了对较好的保留远程投射和形态的较厚的大脑切片 (>100 μm) 成像,光片照明显微术 (也被称为选择性平面照明显微术 (SPIM)) 被应用在 CLARITY 清晰化的大脑组织上。在一些实施方案中,CLARITY 方法使大脑变的透明,用于神经元的组分及其分子身份的可视化和鉴定。在一些实施方案中,通过去除光散射脂质,

CLARITY 使大脑组织变成光学上透明且高分子 - 可渗透的, 所述光散射脂质被多孔水凝胶置换以保留大脑组织的形态, 以使得研究可以被进行而无需将大脑薄切, 这使感兴趣的神经元及其远程突触连接能够可视化。不意图被理论限制, 申请人注意到与本发明之前的以前在培养物或薄的切片中进行的 FISH 相比, 所提供的技术可以使用较厚的组织并且允许个体神经元或 3D 神经网络转录组的更精确的重构。图 16 图示了一种成功的、经验证的基于 Clarity 的方案, 来制备与 FISH 染色相容的光学上清晰的厚切片: (1) 将 2mL Eppendorf 管中的 100 微米的冠状大脑切片在 1.5mL 的 4% 丙烯酰胺、2% 甲醛、0.25% 热引发剂、1×PBS 中在 4 度下孵育过夜; (2) 将氮气鼓穿水凝胶溶液, 持续 10 秒; (3) 将脱气样品在 42 度下孵育 2 小时, 以使聚合; (4) 将样品在 PBS 中洗涤 3 次并且在 10% SDS、1×PBS 中在 37 度下孵育 4 小时, 以使清晰化; 以及 (5) 将样品在 PBS 中洗涤 3 次并且即用于 seqFISH。

[0311] 在一些实施方案中, 所提供的技术提供了用于使背景最小化或者避免虚化背景的方法。在一些实施方案中, 所提供的技术利用使背景最小化或避免虚化背景的成像技术。在一些实施方案中, SPIM 被用于具有较高虚化背景的较厚的切片。在一些实施方案中, 虽然共聚焦显微镜能够拒绝这样背景, 但是它扫描缓慢并且在对较下层成像时光漂白样品的较上层。在一些实施方案中, 在 SPIM 中只有被成像的层被激发光照射。在一些实施方案中, 在用于所提供的技术的 SPIM 设置中, 彼此垂直放置的两个物镜被悬于样品之上约 55°。在一些实施方案中, 光片利用使光线的一个光轴线聚焦在约 10 μm 高和约 200 μm 的有效宽或 FOV 内的柱面透镜来产生。

[0312] 在一些实施方案中, 本发明为所提供的方法提供了显微镜设置。在一些实施方案中, 本发明提供了光片显微镜, 其中样品从侧部被照射。在一些实施方案中, 光片与样品台平行。在一些实施方案中, 光片与检测物镜垂直。在图 17 中图示了光片照射显微镜的示例性设置。通过改装两个反射镜和柱面透镜, 产生光片的平面并且其从侧面照射样品, 并且垂直于检测物镜 (中)。底部反射镜与柱面透镜连接并且被直接安装在物镜的相同底座上。对于该配置, 物镜与照射片同步地移动, 允许沿着 z- 轴 (右, 上部) 扫描样品。右 (下部) 图还显示了样品被置于杂交室内, 并且由下面的空气物镜成像。

[0313] 如在图 18 中所图示的, 用 CLARITY 清晰化并与针对 β-肌动蛋白的 HCR 探针杂交的 100 μm 大脑切片获得 SPIM 图像。采用具有 0.5 μm 间隔的 200 个光学切片来产生重构。用 20× 水浸物镜来观察清晰的 HCR 信号。β-肌动蛋白 mRNA 被高表达, 对细胞体中的大量的点做出了解释。但是, 在轴突中也观察到了清晰的衍射极限点。

[0314] 在一些实施方案中, 可改变针对 CLARITY 清晰化的大脑的和 SPIM 显微术的 HCR-seqFISH 方案。在 4-5 小时内 100 μm 的切片被有效杂交, 指示可检测地标记的寡核苷酸探针可以在 100 μm 厚但清晰的切片中容易地扩散。另外, DNA 酶也容易地扩散以从切片剥离 HCR 信号 (图 15)。在一些实施方案中, 所提供的技术提供了可检测地标记的寡核苷酸, 诸如 FISH 和 HCR 探针, 它们比本领域普通技术人员常规扩散进入 1mm 厚的冠状切片的抗体小, 并且提供了针对整个组织或器官对靶的谱分析, 例如, 对 CLARITY 清晰化的整个大脑进行 seqFISH。

[0315] 在一些实施方案中, 所提供的技术提供了显微术的几何形状。在一些实施方案中, 所提供的技术提供了 SPIM 的可选几何形状, 以使得厚的大脑切片 (>100 μm) 和潜在的整个 CLARITY 清晰化的大脑可以被安装在具有长工作距物镜的落射荧光显微镜上。在一些实施

方案中,光片垂直于成像轴产生,并跨越 200-300 μm 以 10 μm 宽划分以一定倾角安装在显微镜上的样品。在一些实施方案中,所提供的几何形状允许指导开发的流室和自动化设计转移。在一些实施方案中,大脑切片中的基准标志物被用来记录连续的切片。在一些实施方案中,在划分之前,纳米级标尺被注射入大脑,允许不同分区间的良好记录。

[0316] 速度。在一些实施方案中,成像速度限制了最终通量。在一些实施方案中,所提供的 HCR 扩增提供了超过足够成像的光子数目,并且可以使用不那么昂贵的相机来对样品成像。在一些实施方案中,来自收集物镜 (collection objective) 的光可以用成像平面二色和滤器分成多个例如 6 个不同的光束 (例如,5 个荧光和 1 个 DAPI)。这极大地增加了原位“测序仪”的通量,使得在单台显微镜上在 1 周内可完成整个大脑。

[0317] 靶基因选择。在一些实施方案中,本发明提供了用于选择和成像一组靶的技术,所述一组靶诸如一组转录物和 / 或 DNA 基因座 (例如,如所示例的 100 个靶的组)。在一些实施方案中,来自 Allen brain atlas (ABA) 的原位数据库选择靶基因。多重标准可以被用于选择感兴趣的基因,例如,皮层区中的可能代表细胞身份的那些基因。从重叠的标准计算选择最佳的基因的组是熟知的 (2. Alon, N; Moshkovitz, Dana; Safra, Shmuel (2006), "Algorithmic construction of sets for k-restrictions", ACM Trans. Algorithms (ACM) 2 (2): 153-177, ISSN 1549-6325; 8. Cormen, TH.; Leiserson, Charles E.; Rivest, Ronald L.; Stein, Clifford (2001), Introduction to Algorithms, Cambridge, Mass.: MIT Press and McGraw-Hill, 第 1033-1038 页, ISBN0-262-03293-7; 12. Feige, U (1998), "A threshold of $\ln n$ for approximating set cover", Journal of the ACM (ACM) 45 (4): 634-652, ISSN 0004-5411)。在一些实施方案中,集合覆盖启发法 (set-cover-heuristics, Pe'er, 2002) 被用于选择以下基因: 1. 已知限定细胞亚型的基因; 2. 在 ABA 中表现出“盐和胡椒 (salt and pepper)”的表达模式的基因; 3. 属于诸如转录因子、离子通道、GPCR 和神经妥乐平的基因家族的基因; 和 4. 被从皮层样品的 RNA seq 实验剔除的基因。例如, SLC1A3 标识胶质细胞而 SLC6A1 标识抑制性神经元, 并且 SLC17A7 标识兴奋性神经元。在一些实施方案中,具有异质性表达模式的基因诸如 PVALB、SST 和 CALB 标识出抑制性神经元的亚组。以下显示了 100 种基因的示例性组:

[0318]

基因名称	表达谱
SLC6A1	所用抑制性的 (I)
SLC17A7	所有兴奋性的 (E)
SLC1A3	神经胶质
PVALB;	亚组 I
SST;	亚组 I
CALB2;	亚组 I

[0319]

LER5;	同形皮质
TNNC1;	同形皮质
MYL4;	同形皮质
SATB2;	同形皮质
CCL27a;	同形皮质
BOC;	初级运动 LI
DACT2;	初级运动 LI
LHX1;	初级运动 LI
PVRL3;	初级运动 LI
SLC44a3;	初级运动 L2/L3
KLK5;	初级运动 L2/L3
TNNC1	初级运动 L2/L3
WNT6	初级运动 L2/L3
ZMAT4	初级运动 L5
STARD8	初级运动 L5
TCF21	初级运动 L5
MYL4	初级运动 L5
KRT80	初级运动 L6a
OLFR19	初级运动 L6a
TBC1d30	初级运动 L6a
OLF16	初级运动 L6b
EAR6	初级运动 L6b
CHIT1	初级运动 L6b
SLN	次级运动 LI
ADAMTS8	次级运动 LI
EPYC	次级运动 LI
KCNV1	次级运动 LI
pcdh7	次级运动 L2/L3
GLT8d2	次级运动 L2/L3
HKDC1	次级运动 L2/L3
SRPX	次级运动 L3
ZFP458	次级运动 L3
SLC30a8	次级运动 L3
GK5	次级运动 L5
TEX28	次级运动 L5
MS4a10	次级运动 L5
K RT16	次级运动 L6a

[0320]

KRT42	次级运动 L6a
DOC2a	次级运动 L6a
KR T33b	次级运动 L6b
YBX	次级运动 L6b
PNPLA5	次级运动 L6b
TMEM215	初级躯体感觉 LI
SDC1	初级躯体感觉 LI
PREX1	初级躯体感觉 LI
DIEXF	初级躯体感觉 LI
DHRS7c	初级躯体感觉 L2/L3
DDIT41	初级躯体感觉 L2/L3
TDG	初级躯体感觉 L2/L3
EPSTI1	初级躯体感觉 L2/L3
RORb	初级躯体感觉 L4
GSC2	初级躯体感觉 L4
KRT10	初级躯体感觉 L4
GCA	初级躯体感觉 L4
DCBLD2	初级躯体感觉 L5
ABCD2	初级躯体感觉 L5
GTDC1	初级躯体感觉 L5
IL17RA	初级躯体感觉 L6a
TBR1	初级躯体感觉 L6a
PPID	初级躯体感觉 L6a
IGHM	初级躯体感觉 L6b
MMGT1	初级躯体感觉 L6b
CPLX3	初级躯体感觉 L6b
ART2b	次级躯体感觉 LI
GNB4	次级躯体感觉 LI
B3GAT2	次级躯体感觉 LI
PDC	次级躯体感觉 L2/L3
ADIG	次级躯体感觉 L2/L3
FPR1	次级躯体感觉 L2/L3
INHBC	次级躯体感觉 L4
RUFY4	次级躯体感觉 L4
HGFAC	次级躯体感觉 L4
EFCAB4b	次级躯体感觉 L5
SSTR2	次级躯体感觉 L5

[0321]

ZFP395	次级躯体感觉 L5
CCDC36	次级躯体感觉 L6a
ST14	次级躯体感觉 L6a
MYL12b	次级躯体感觉 L6b
RSP02	次级躯体感觉 L6b
NDNF	L1 (I)
RASGRF2	L2/3 (I)
CUX2	L2/3/4
RORB	L4
SCNN1A	L4
ETV1	L5
FEZF2	L5
BCL6	L5
TRIB2	L5a
FOXP2	L6
TLE4	L6/L6b
CTGF	L6b
CYLD	L2/3
CMTM3	L2/3
ANKRD6	L2/3

[0322] seqFISH 与蛋白质检测、细胞器标志物和活性测量集成。在一些实施方案中,所提供的技术例如 seqFISH 允许以单细胞分辨率对同一样品中的 RNA 以及蛋白质、神经活动和结构布置原位多路复用分析。在另外的一轮杂交中可使特定靶的抗体与样品杂交。在一些实施方案中,所提供的方法任选地包括免疫染色的步骤。在一些实施方案中,在多种抗体被用于顺序的杂交轮中检测许多蛋白质靶 (Schubert W 等人, Analyzing proteome topology and function by automated multidimensional fluorescence microscopy. Nat Biotechnol (2006) 24(10):1270-1278)。申请人注意到在细胞中存在超出 mRNA 多达约 100-1000 倍或更多倍的较高的蛋白丰度。被靶向的蛋白可以标识诸如线粒体、ER、转运小泡、细胞骨架的细胞器以及突触连接。例如,可以使用 MAP2 抗体来标识出细胞边界,以帮助划分轴突和树突。

[0323] 在通过所提供的方法 (例如, seqFISH) 对转录谱分析之前,可以在落射荧光和光片照明显微镜上对大脑切片的实况观察成像。在大脑切片中钙 (Nakai J, Ohkura M, Imoto K (February 2001). "A high signal-to-noise Ca(2+) probe composed of a single green fluorescent protein". Nat. Biotechnol. 19(2):137-41; Akerboom 等人, "Optimization of a GCaMP calcium indicator for neural activity imaging." J Neurosci. 2012 年 10 月 3 日; 32(40):13819-40; Stosiek 等人, "In vivo two-photon calcium imaging of neuronal networks." Proceedings of the National Academy of Sciences 100(12):7319) 和电压传感器 (Cohen, 等人, "Optical Measurement of Membrane Potential" in Reviews of Physiology. "Biochemistry and Pharmacology,

第 83 卷, 第 35-88 页, 1978(6 月); Mutoh 等人, Genetically Engineered Fluorescent Voltage Reporters ACS Chem Neurosci. 2012 年 8 月 15 日; 3(8):585-592; Peterka 等人, Imaging voltage in neurons. Neuron. 2011 年 1 月 3 日; 69(1):9-21) 可被成像。SPIM 允许对大脑切片中的这些传感器有效且快速地成像。将大脑切片固定在显微镜上并且所提供的方案诸如 seqFISH 方案可以用自动化的射流来进行。在一些实施方案中,除了实况测量之外,检测活动依赖性即早基因 (IEG) 的 mRNA 作为在神经元中的整体神经活动的衡量。例如,在神经元中容易地检测到 CamKII 和 cFos 具有异质性表达水平;根据丰度可将它们分别并入基因的组中,例如,示例性 100 种基因多路复用,或在另外的循环中进行 FISH。

[0324] 集成连接组学和 seqFISH 以鉴定不同的躯体感觉运动神经网络内不同的神经元的分子身份。

[0325] 利用所提供的技术系统地表征躯体感觉运动神经网络内不同的神经元群体的分子身份。在一些实施方案中,相同的功能子网络内的神经元群体可共用共同的标志物基因的组,但也具有在细胞水平上定义身份的其他基因的异质性表达。在一些实施方案中,不同子网络中的细胞在其表达方式上更为不同。其每个控制基础类型的感觉运动功能的示例性皮层-皮层躯体子网络为:(1) 口面咽 (orofaciopharyngeal), 针对饮食;(2) 上肢 (upper limbs), 针对伸出和抓取;(3) 下肢 (lower limbs), 针对运动;和 (4) 须 (whisker), 针对有节奏的须运动。在一些实施方案中,所提供的技术提供了新颖且严密的方法,用于表征具有不同神经网络的皮层神经元的分子身份,并且为了解遗传回路提供了非常宝贵的信息,所述遗传回路为哺乳动物大脑的布线图的基础。

[0326] 利用一系列神经通路,可以生成数字皮层连接图 (atlas) 来展示束路追踪研究的原始图像。可以用图表重构通路来生成皮层-皮层连接图,以助于大规模数据的分析。基于皮层内连接,可以建立四个不同的皮层-皮层躯体子网络,其每个控制基础类型的感觉运动功能。这些子网络的每个包含初级躯体感觉 (SSp) 皮层区、初级躯体运动 (MOp) 皮层区、次级躯体运动 (MOs) 皮层区和辅助躯体运动 (SSs) 皮层区中的 4-5 个不同的功能域 (domain), 根据这些皮层区与对应于特定躯体分区的其他躯体感觉运动区的连接强度这些皮层区被进一步细分。在一些实施方案中,口面咽子网络包括五个主要的节点:(1) SSp 口和鼻域 (SSp-m/n); (2) MOp 口面域 (Mop-orf); (3) MOs 喙背外侧 (rostradorsolateral) 域 (MOs-rdl); (4) SSp 桶状区前外侧 (barrel field anterolateral) 域 (SSp-bfd. al); 和 (5) SSs 喙和尾腹侧 (caudoventral) 域 (SSs-r&cv)。在一些实施方案中,上肢子网络的四个主要节点包括 (1) SSp 上肢 (SSp-ul); (2) Mop-ul; (3) 喙背侧 (rostradorsal) MOs (MOs-rd); 和 (4) 尾背侧 (caudodorsal) SSs (SSs-CD)。在一些实施方案中,下肢/躯干子网络包括 SSp 下肢/躯干区域 (SSp-ll/tr)、MOp-ll/tr 和喙背侧中部 (rostradorsomedial) MOs (MOs-rdm) (图 10B-D)。在一些实施方案中,须子网络包括尾中部 (caudomedial) SSp-bfd (SSp-bfd. cm), 对应于须初级运动皮层 (vM1) 的 MOp-w 和尾背侧 SSs (SSs-cd; 图 19)。示例性数据由 Mouse Connectome Project (www.mouseconnectome.org) 描述。

[0327] 为了确定这些躯体感觉运动子网络的每个中的不同神经元群体的分子身份,使用多荧光逆行示踪剂来标记神经元,并且所提供的技术诸如 seqFISH 可被应用于以单细胞分辨率确定逆行地标记的群体的基因表达谱。为了标记神经元群体,将多种(例如,五种)逆行示踪剂注射入同一动物中的每个感觉运动子网络的主要节点之一的五个主要

靶（示踪剂信息如下）。例如，在一个动物中，回路示踪剂被注射入口面咽子网络的主要的皮层节点中的两个（SSp-bfd. al 和 SSs-r&cv）和皮层下节点中的三个（尾壳腹外侧（caudoputamenventrolateral）域，CP-v1；丘脑腹后内侧核，VPM；和腹侧三叉神经脊束核，SPV）。这同时地背景（back）标记口面咽子网络的所有其他节点中的五种不同的神经元。在该实施例中，标记的神经元在 SSSp-m/n 域中且在 Mop-oro 域中。

[0328] 另一方面，示踪剂可被注射入四种不同的 SSp 躯体分区（即 SSp-m/n、SSp-ul、SSp-ll/tr 和 SSp-bfd. cm），其每种属于不同的躯体子网络。这同时地标记与不同的子网络相关的皮层区中的不同神经元群体。在这种情况下，例如，背景标记的神经元可在与每个子网络相关的 MOp 域中观察到，所述 MOp 域即 Mop-orf、Mop-ul、Mop-ll/tr 和 Mop-w。被应用于四个躯体感觉运动子网络的每一个的所有主要节点和皮层下靶的注射策略标记了每个子网络的不同神经元群体。

[0329] 注射示踪剂后（例如，注射示踪剂后一周），处死动物，并收获其大脑并以 20 μ m 或 100 μ m 的厚度冠状地切片，用于被背景标记的神经元的 seqFISH 分析。利用例如 Allen Brain Atlas 项目的在线数字基因表达数据库（www.Brian-Map.org）（Lein 等人，2007Genome-wide atlas of gene expression in the adult mouse brain. Nature, 11；445(7124):168-76）预选基因用于谱分析，所述基因诸如在躯体感觉运动皮层区（SSp、Mop、MOs、SSs）中富集表达的示例性的约 100 种基因。

[0330] 注射策略和注射后处理。使用三百只 4 周龄的 C57B1/6 雄性小鼠。在一只动物中，如以上描述的将五种荧光逆行示踪剂注射入同一躯体感觉运动子网络内的不同节点，或不同躯体子网络的不同节点中（图 19）。示踪剂是荧光金（FG，黄色）、与 488 或 647（CTb-488[绿色]、CTb-647[粉红色]）缀合的霍乱毒素 b、Red retrobeads（RR，红色）、和与 655（WGA-Qdot655，白色）缀合的麦胚凝集素。由于 Qdot655 具有与 CTb647 不同的激发波长，所以 Qdot655 可以被捕捉入不同的通道中并用独特的色度伪上色。经由立体定向手术注射（离子透入地或以压力注射）示踪剂。手术和灌注的细节被描述于例如 Hintiryan 等人，Comprehensive connectivity of the mouse main olfactory bulb: analysis and online digital atlas. Front Neuroanat. 2012 年 8 月 7 日；6:30. eCollection 2012。在一些实施方案中，在准确相同的坐标中用所使用的相同的示踪剂注射两只成对小鼠。动物中的一个被用来证实标记的细胞和注射位点的位置，而另一只经受所提供的例如 seqFISH 方法。在示踪剂运输后灌注动物，并且大脑被冠状地切成 50 μ m 厚度的片并且以四个系列收集。四个系列的切片中的一个用荧光尼氏染色液（NeuroTrace Blue[NTB]）复染色、安装在载玻片上、并利用 Olympus VS120 虚拟显微术系统成像。在一些实施方案中，尼氏染色提供了用于使背景标记的细胞的精确解剖位置可视化的细胞结构背景。通过信息流水线对这些图像处理，以使得每个个体图像忠实地配准到它的 Allen Reference Atlas (ARA) 的相应水平。该尼氏染色连同所提供的信息工具一起使能够自动且精确的计数每个 ROI（在该情况中，躯体感觉运动区的不同域）中的每种示踪剂标记的神经元群体的大约数目。分布方式被自动地标绘到对应的图水平以生成其连接图。

[0331] 将成对的小鼠同时处死，并且将其大脑以 20 μ m 或 100 μ m 的厚度切片用于 seqFISH 分析。将这些切片首先在 20 \times （或 10 \times ）物镜下成像以展现用不同的示踪剂背景标记的神经元。对于示例性的 100 种基因的组，包含躯体感觉运动区的贯穿所有的冠状水

平的大脑切片被用于进行 seqFISH。这种方法展现了每个示踪剂标记的神经元中的基因表达。所有图像首先被分析每种个体示踪剂标记的神经元的基因表达谱。每个切片被配准回其配对的大脑的最接近的匹配切片,以使得来自配对的大脑的近似相同冠状水平处的切片可以在连接体(connectome)视窗中被并排展示。如此,在不同神经元群体的最接近匹配的解剖学背景中,其分子谱被展示。在一些实施方案中,在每个逆行标记的神经元群体中,基因表达谱被关联。

[0332] 结果。在一些实施方案中,在不同的躯体感觉运动区内不同的逆行地标记的神经元群体展示出不同的转录组谱;甚至用不同的示踪剂标记的相同域(例如,SSp-m)中的神经元群体展示出与具有不同连接谱的其邻近神经元不同的基因表达谱。在一些实施方案中,在相同子网络内的不同躯体感觉运动节点(例如,SSp-m、MOp-orf、SSp-bfd.a1和SSs-r)内的不同神经元群体共用共同的网络特异的基因,而不同的神经网络内(例如,口面咽和下肢/躯干子网络)的神经元群体展示出非常不同的转录组谱。在一些实施方案中,对于不同皮层区和不同层中的那些神经元,鉴定区域(例如,SSp或MOp)或层(不同层)特异基因。

[0333] 如示例的,所提供的技术提供了荧光束示踪与 seqFISH 技术的独特组合,来以亚神经元分辨率和忠实的解剖学背景表征不同的躯体感觉运动网络内基于连接的神经元群体(细胞类型)的分子身份。参见,例如,图3和图9,为示例性结果。在一些实施方案中,所提供的技术包括平行地测量其他参数(即抗体和细胞器染色,以及 IEG 表达水平),并且可被应用于整个新皮层或大脑。

[0334] 用于在线分析并展示数据的高通量流水线和信息学工具。在一些实施方案中,所提供的技术提供了用于通过例如可公开访问的数据库,诸如 www.MouseConnectome.org) 在线分析并展示数据高通量流水线和信息学工具,。在一些实施方案中,所提供的技术提供了与 Mouse Connectome Project 的集成, Mouse Connectome Project 的广泛研究范围和多种荧光成像的使用使其在连接组学群体(community)之中成为有价值的工具并且非常适于小鼠大脑中的远程连接研究。例如,其提供了在线可视化工具,所述可视化工具允许用户将多荧光标记的途径在自身的细胞结构背景之上并且在对应的 ARA 图谱水平上可视化。在一些实施方案中,为了使 seqFISH 信息与其对应的逆行地标记的体细胞密质部分忠实地关联,被标记的细胞体被从组织背景并从同一切片的图像离散地分割出,但在不同的轮次中取得(例如,首先针对成像逆行示踪剂,然后针对 seqFISH 中不同的 mRNA),并且通过它们相对于切片和/或组织中的解剖学地标上的固定的参考点的坐标来在空间上索引化。在一些实施方案中,为了将数据与在 ARA 中限定的立体定位坐标关联,本发明提供了极大地增加配准精确度(即,将每个扫描的显微图像扭曲至 ARA 的对应水平的形状)和图像分割的新颖的配准流水线,该图像分割自动且精确地枚举给定的 ROI(例如,SSp-m, MOp-11)中的荧光标记的神经元的。在一些实施方案中,所提供的技术共同地允许大脑内的标记和来自多种示踪剂的 seqFISH 数据,并且跨越多个大脑,被整理成可视化和注释目的单个解剖学框架(single anatomical framework)。

[0335] 在一些实施方案中,在对应的 Allen Reference Atlas(Dong, H. W. (2008). The Allen Reference Atlas: A Digital Color Brain Atlas of C57BL/6J Male Mouse, John Wiley&Sons) 的图水平上配准图像。由配准过程产生的变形矩阵被应用到原始分辨率的图

像上,来获得高分辨率的扭曲图像。在配准和配准精化后,NeuroTrace® 荧光尼氏染色被转化成明场图像。接下来,每个图像的每个通道被调整亮度和对比度,以使工具中的标记可见度和质量最大化,所述工具例如 iConnectome。在修正(即,偏斜度、角度)和 JPEG2000 文件格式转换后,可以将图像公布至 iConnectome 查看器 (www.mouseconnectome.org)。

[0336] FISH 可视化工具。在一些实施方案中,产生的所有连接数据通过 MCP 信息学流水线处理并且通过新的 iConnectome FISH 查看器 (www.mouseconnectome.org) 在线展示。不同于展示两个顺行地 (PHAL 和 BDA) 标记和两个逆行地 (FG 和 CTb) 标记的 iConnectome 查看器,iConnectome FISH 可以展示多达五种带有逆行荧光染料的不同神经元群体。如上提到的,可将每组注射剂给予一对小鼠。一只可根据常规 MCP 流水线处理并且在 iConnectome FISH 查看器中被展示,以在多荧光标记的神经元群体自身尼氏染色的细胞结构背景内和它们对应的 ARA 水平内呈现多荧光标记的神经元群体。这些可以为跨越整个大脑的每个荧光标记的神经元群体提供精确的解剖学信息。在 seqFISH 之后,来自其配对伴侣的大脑切片被配准到其配对伴侣所配准至的最接近的 ARA 水平并且可以在不同的窗口中被并排显示。在神经元中表达的基因列表在侧部面板上被列出。点击基因之后,表达该基因的荧光标记的神经元可以变亮,来指示其表达的位置。这提供了在连接背景和解剖学背景的全局背景内显示神经元群体的分子身份的实用的方式。

[0337] 在一些实施方案中,开发了对应的数据库,所述数据库允许用户对这些数据进行分析并且将神经连接与其分子身份关联。该信息学工具基于如下数据库建立:所述数据库利用基因编条形码来存储与每个逆行标记的神经元群体相关的信息(例如细胞数目、解剖学位置)。该数据库可以帮助用户鉴定在相同神经网络或不同神经网络内的神经元的对应的基因条形码。

[0338] 映射整个大脑。在一些实施方案中,所提供的技术在单个神经元水平上具有足够的灵敏度、选择性、自动化和/或时空分辨率,用于对整个大脑高通量分析逆行标记的神经元中的基因表达。

[0339] 去除步骤的另外的示例性方法

[0340] 在一些实施方案中,本发明提供了用于从靶去除可检测地标记的寡核苷酸的多种方法。在一些实施方案中,核酸外切酶 III (ExoIII) 被用来去除可检测地标记的寡核苷酸。图 21 图示了使用 Exo III 的 HCR 再杂交的示例性方法。在图 21 中,Exo III 消化桥接链和 HCR 聚合物,保持中间寡核苷酸完整,用于与新的桥接链杂交。在图 21 (b) 中展示了使用靶向 T3T 小鼠成纤维细胞中的 β -肌动蛋白 (Actb) 转录物的可检测地标记的寡核苷酸的示例性数据。左边的图像显示了使用 Alexa 488 染料的 Actb 转录物的初始杂交和扩增。中间的图像显示了室温下在外切酶 III 中孵育 1 小时后 Alexa488 通道中的信号完全丢失。右边的图像显示了仅添加新的桥接链以及用 Alexa 647 染料标记的相应发夹之后,Actb 转录物的再扩增。调整图像的对比度以图示方法的某些特征。

[0341] 在一些实施方案中,Lambda 核酸外切酶 (λ -exo) 被用来去除可检测地标记的寡核苷酸。图 22 图示了使用 λ -exo 的 HCR 再杂交的示例性方法。在图 22 中, λ -exo 消化 5' 磷酸化的桥接链并且从与靶(例如,mRNA)结合的中间寡核苷酸释放 HCR 聚合物,并且保持中间寡核苷酸的完整用于在洗涤去释放的聚合物之后与新的桥接链杂交。在图 22 (b) 中展示了使用靶向 T3T 小鼠成纤维细胞中的 β -肌动蛋白 (Actb) 转录物的可检测地标记的

寡核苷酸的示例性数据。左边的图像显示了使用 Alexa 488 染料的 Actb 转录物的初始杂交和扩增。中间的图像显示了在 37°C 下在 λ -exo 中孵育 1 小时后 Alexa 488 通道中的信号丢失。右边图像显示了在用洗涤缓冲液洗涤并仅添加新的桥接链连同用 Alexa 647 染料标记的对应的发夹之后, Actb 转录物的再扩增。调整图像的对比度以图示方法的某些特征。

[0342] 在一些实施方案中, 尿嘧啶特异性切除试剂 (USER) 被用来去除可检测地测标记的寡核苷酸。图 23 图示了使用 USER 的 HCR 再杂交的示例性方法。在图 23 中, USER 在桥接链中的脱氧尿苷核苷酸处消化并且从与靶 (例如, mRNA) 结合的中间寡核苷酸释放 HCR 聚合物, 并且保持中间核苷酸的完整用于在洗涤去片段和释放的聚合物之后与新的桥接链杂交。在图 23(b) 中展示了使用靶向 T3T 小鼠成纤维细胞中的 β -肌动蛋白 (ACTb) 的可检测地标记的寡核苷酸的示例性数据。左边的图像显示了使用 Alexa 488 染料的 Actb 转录物的初始杂交和扩增。中间的图像显示了在 37°C 下在 USER 中孵育 1 小时后 Alexa 488 通道中的信号丢失。右边图像显示了在用洗涤缓冲液洗涤并仅添加新的桥接链连同用 Alexa 647 染料标记的对应的发夹之后, Actb 转录物的再扩增。调整图像的对比度以图示方法的某些特征。

[0343] 在一些实施方案中, 通过使用互补寡核苷酸 (cTOE) 置换来去除可检测地标记的寡核苷酸。在一些实施方案中, 置换包括使用葡聚糖或其衍生物、盐和 / 或有机溶剂。在一些实施方案中, 置换包括使用葡聚糖或其衍生物。在一些实施方案中, 置换包括使用硫酸葡聚糖。在一些实施方案中, 置换包括使用盐。在一些实施方案中, 盐是 $MgCl_2$ 。在一些实施方案中, 置换包括使用有机溶剂。在一些实施方案中, 有机溶剂是甲酰胺。可以单独地或以组合优化多种因素, 例如但不限于 cTOE 浓度、孵育时间、缓冲液的组成以及有机溶剂的类型和 / 或浓度。图 24 显示了使用 cTOE 置换 smFISH 探针的示例性数据。显示了待置换的 smFISH 探针 (Alexa 647) 和共定位的 smFISH 探针 (Alexa 532) 之间的荧光强度的平均比率。进行多种处理, 其中比较了 cTOE 的浓度、杂交缓冲液组成和置换时间。所有的置换探针条件导致比对照显著更多的置换, 在所述对照中细胞被置于 10% DS 中并且不添加 cTOE。不意图被理论限制, 申请人注意到, 除了其他以外, 增加 cTOE 的浓度、增加 cTOE 探针杂交的时间量、将缓冲液调整为 10mM $MgCl_2$ 或 10% 甲酰胺全都导致置换增加。cTOE 以 2.5 μM 在 10% 硫酸葡聚糖 (DS) 中持续 2 小时导致 Alexa 647 smFISH 信号的最小残留但比基线信号较小的增加, 所述基线信号通过杂交 Alexa 594 (A594) 而不是 Alexa 647 且不添加 cTOE 来确定。

[0344] 寡核苷酸制备的另加的实施例

[0345] 通过 PCR 扩增一组序列 (图 25)。将产物分离, 例如, 在 -20°C 下用 5 体积沉淀缓冲液 (30:1 EtOH:1M NaAc) 沉淀至少 10 分钟。将沉淀混合物离心 10 分钟。丢弃上清液并且在切口酶缓冲液中用合适单位的酶重构寡核苷酸团块, 所述合适单位的酶基于在 1 小时中约 10 单位的酶消化约 1 μg DNA。孵育时间结束后, 将样品再次沉淀并在 2X 加载缓冲液 (96% 甲酰胺 / 20mM EDTA) 和水中重构以制备最终的加载缓冲液 (48% 甲酰胺 / 10mM EDTA)。将样品加热到 95°C 以使 DNA 完全变性。然后将变性的 DNA 加载进变性聚丙烯酰胺凝胶 (8M 尿素 10-12% 丙烯酰胺) 中。使凝胶在 250V 下运行 1 小时, 或按期望优化。电泳后使用 1x sybr gold 对凝胶染色 15 分钟且然后可视化。合适的条带被切出、破碎、并在 DI 水中孵育 2 小时。孵育后, 对样品再次沉淀并且然后使用真空柱纯化。用 30 μL 无 RNA 酶水洗脱柱以

得到最终产物,如在图 26 中所示的。

[0346] 在一些实施方案中,所提供的方法使用限制酶位点而不是切口核酸内切酶位点。与图 25 中的扩增步骤相似,通过 PCR 扩增 5'-端侧翼为 BamHI 位点且 3'-端侧翼为 AatII 位点的一组序列。将 PCR 产物在 -20°C 下用 5 体积的沉淀缓冲液 (30:1EtOH:1M NaOAc) 沉淀至少 10 分钟并分离,随后用 BamHI 和 AatII 消化。将产物再次纯化并经受 *exo III* 消化。去除消化的核酸提供产物寡核苷酸。

[0347] 等同物

[0348] 已经描述了本发明的一些例证性实施方案,对于本领域技术人员应该是明显的是上述仅仅是例证性的而非限制性的,仅以实例的方式呈现。大量的修改和其他例证性的实施方案在领域普通技术人员的理解范围内并且被构思为落在本发明的范围内。特别地,虽然本文展示的许多实例涉及方法行为或系统元素的特定组合,应理解这些行为和这些元素可以其他方式组合来完成相同的目标。仅结合一个实施方案讨论的行动、元素和特征不意图排斥其他实施方案中的类似角色。此外,对于在以下权利要求中引用的一个或更多个方法加功能限制,方法不意图被局限于在本文中公开的用于执行所述的功能的方法,而是意图在范围上覆盖现在已知或以后开发的用于执行所述的功能的任何方法。

[0349] 在权利要求中使用序数术语如“第一”、“第二”、“第三”等,来修饰权利要求元件,其自身并不意味着任何优先性、优越性、或一种权利要求相对另一个的顺序或方法的行为被执行的时间顺序,而仅被用作标记以区分具有某个名字的一个权利要求元件和具有相同的名字(但使用了序数术语)的另一个元件,以区分权利要求元件。相似地,在权利要求中,使用的 a)、b) 等,或 i)、ii) 等其自身并不意味着权利要求中任何优先性、优越性或步骤的次序。相似地,在说明中使用的这些术语其自身并非具有任何所需的优先性、优越性或次序。

[0350] 认为上述书面说明足以使本领域技术人员能够实践本发明。本发明不被所提供的实施例的范围限制,因为实施例意图作为本发明的一方面的单个例证并且其他功能上等同的实施方案在本发明的范围内。除了本文显示和描述的那些之外,根据前述说明本发明的多种改变对于本领域技术人员来说将变得明显,并且落入所附权利要求的范围内。本发明的优点和目的不必要被本发明的每个实施方案包括。

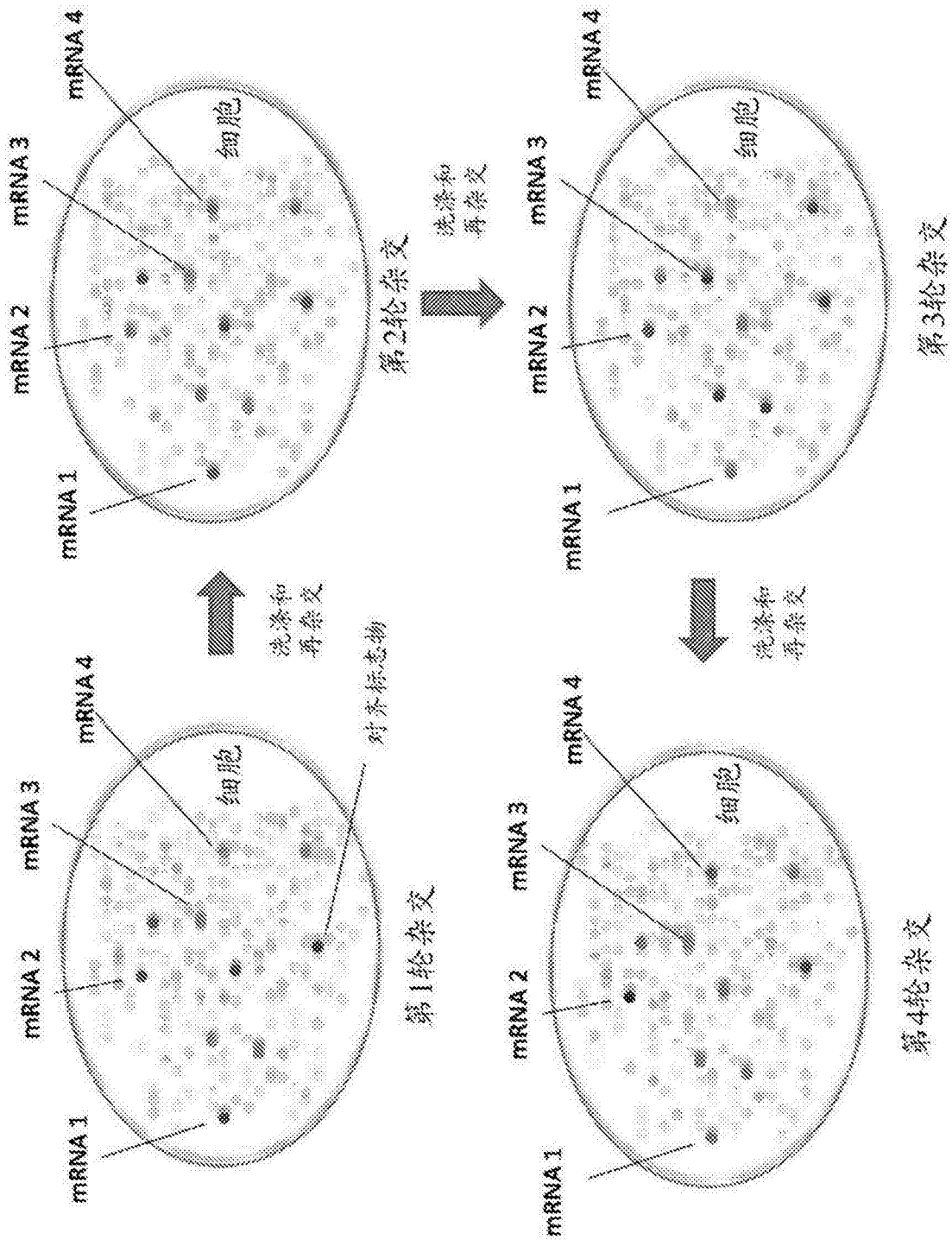
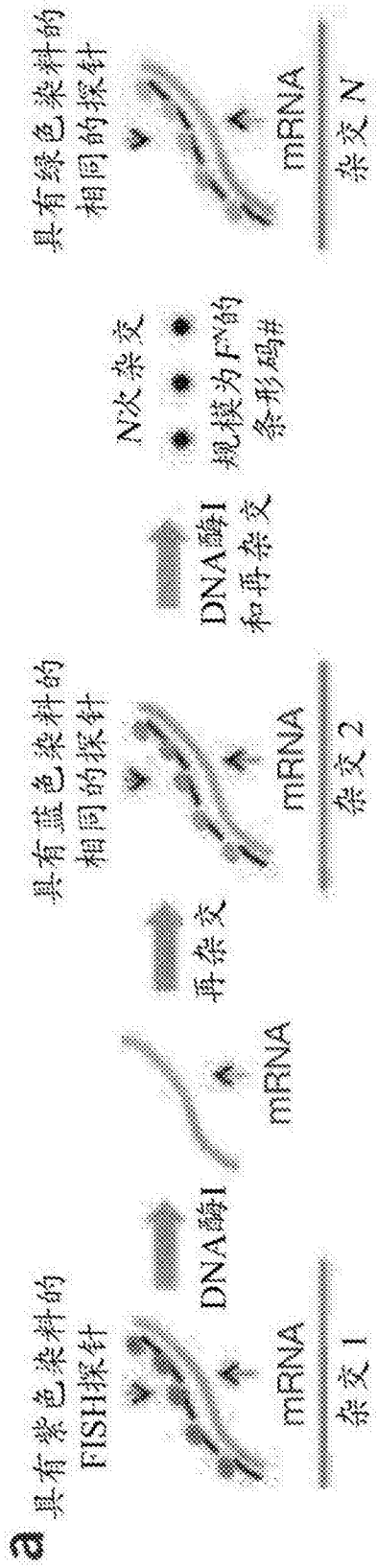


图 1



复合四色FISH图像

b

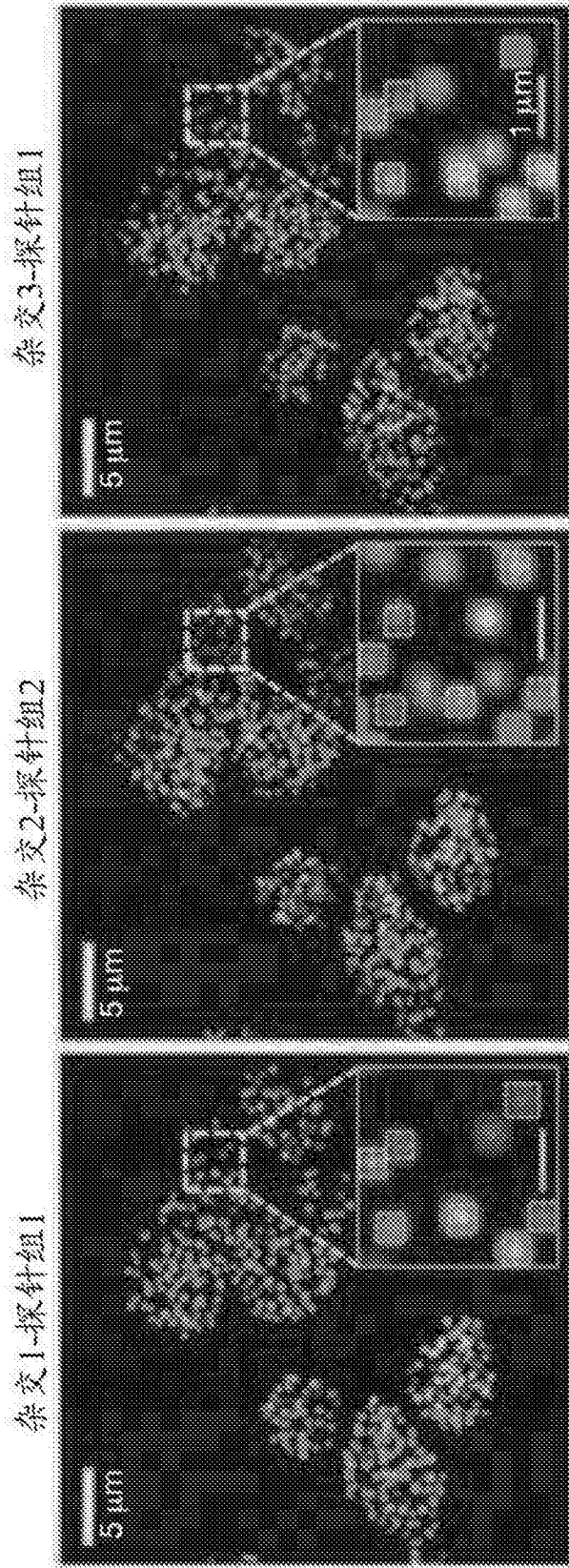


图 2

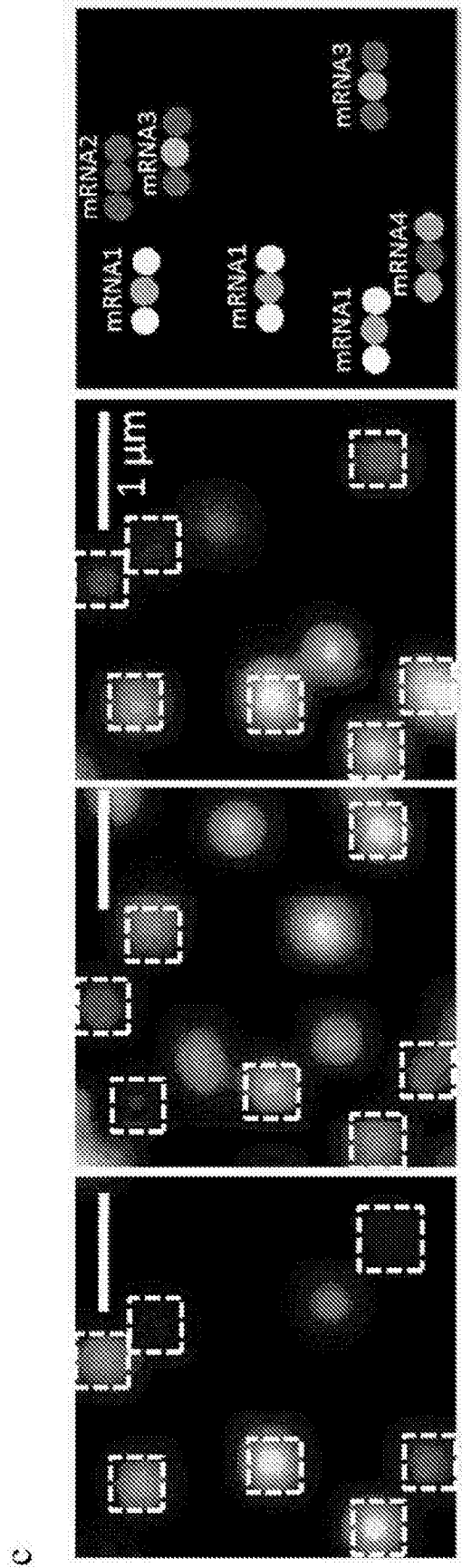


图2(续)

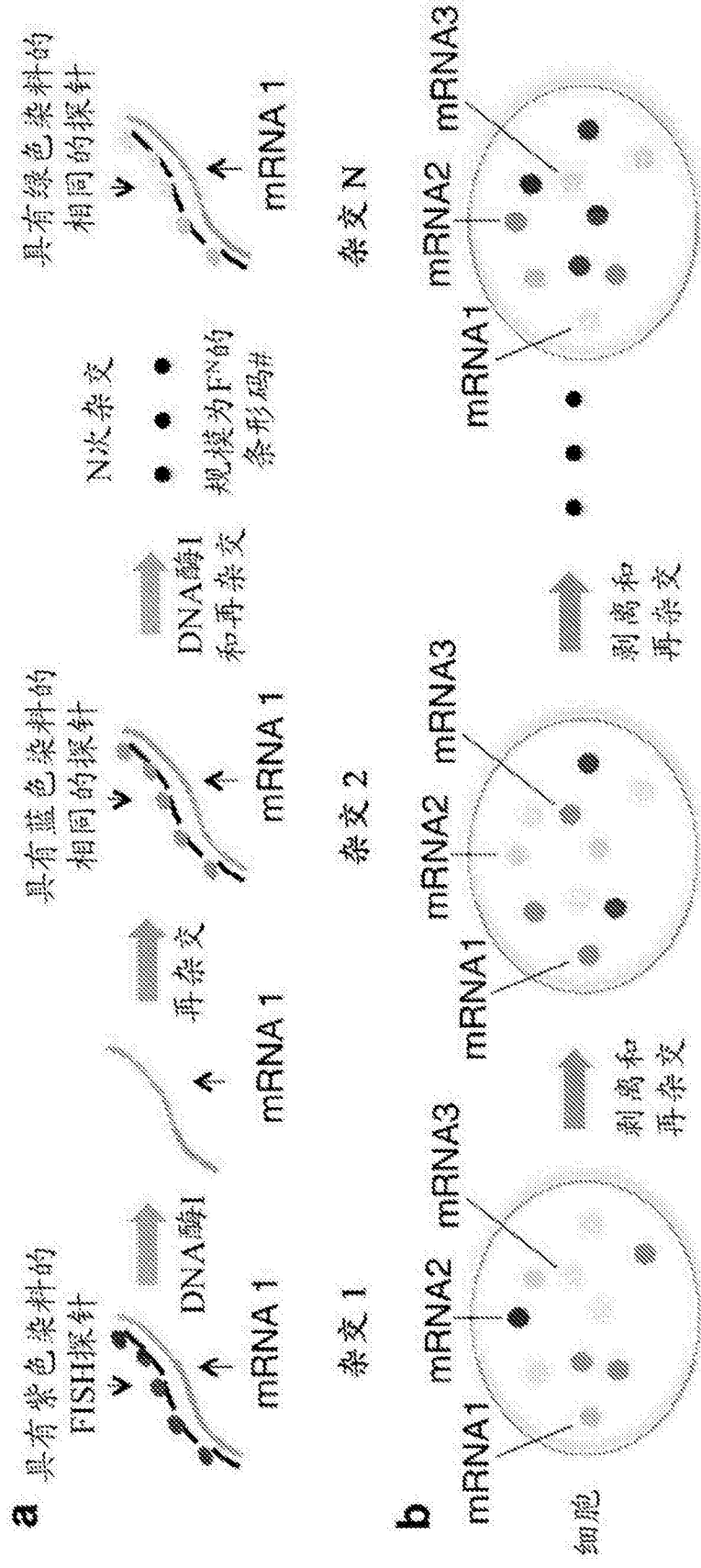


图 3

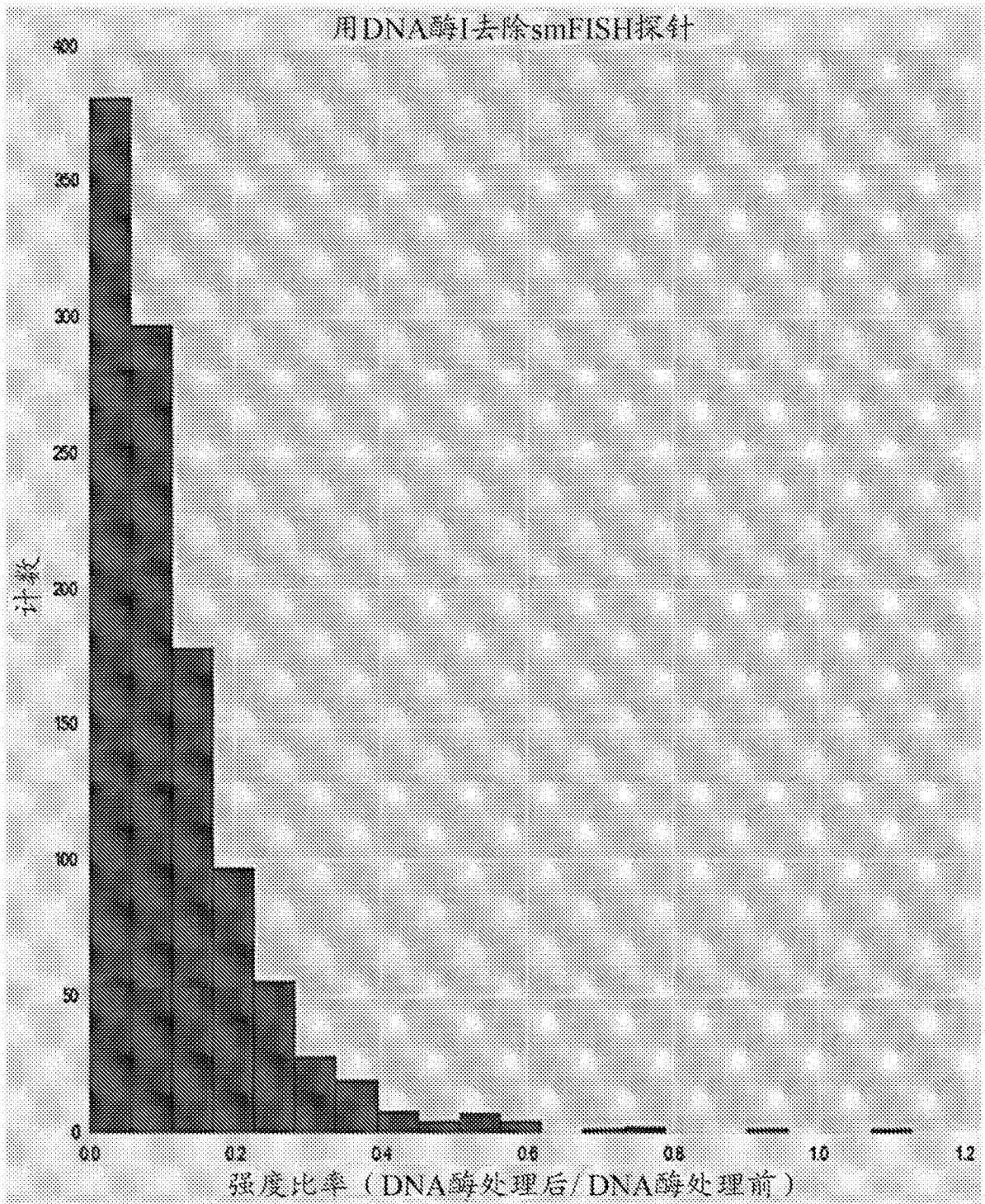


图 4

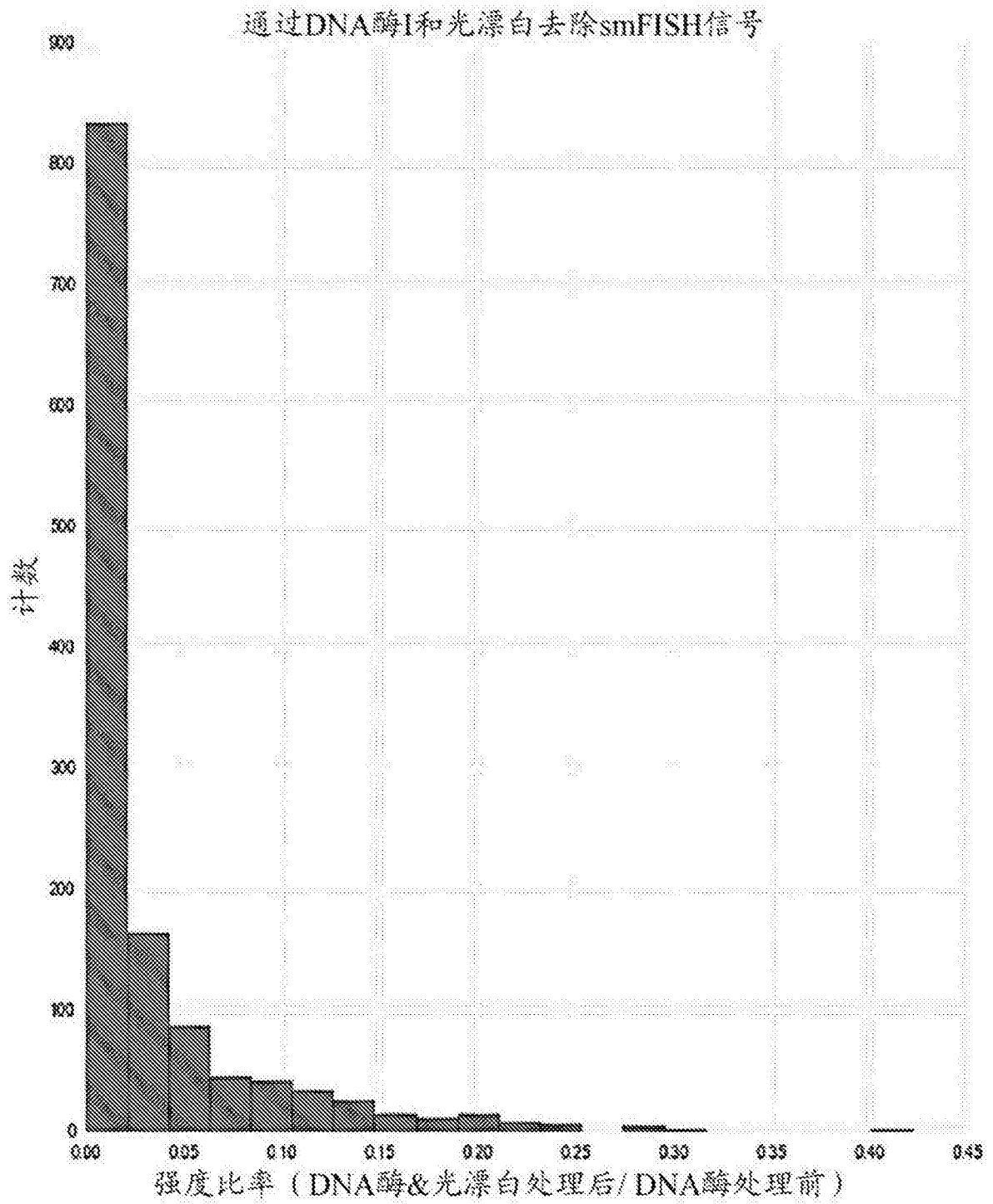


图 5

强度 v. 杂交

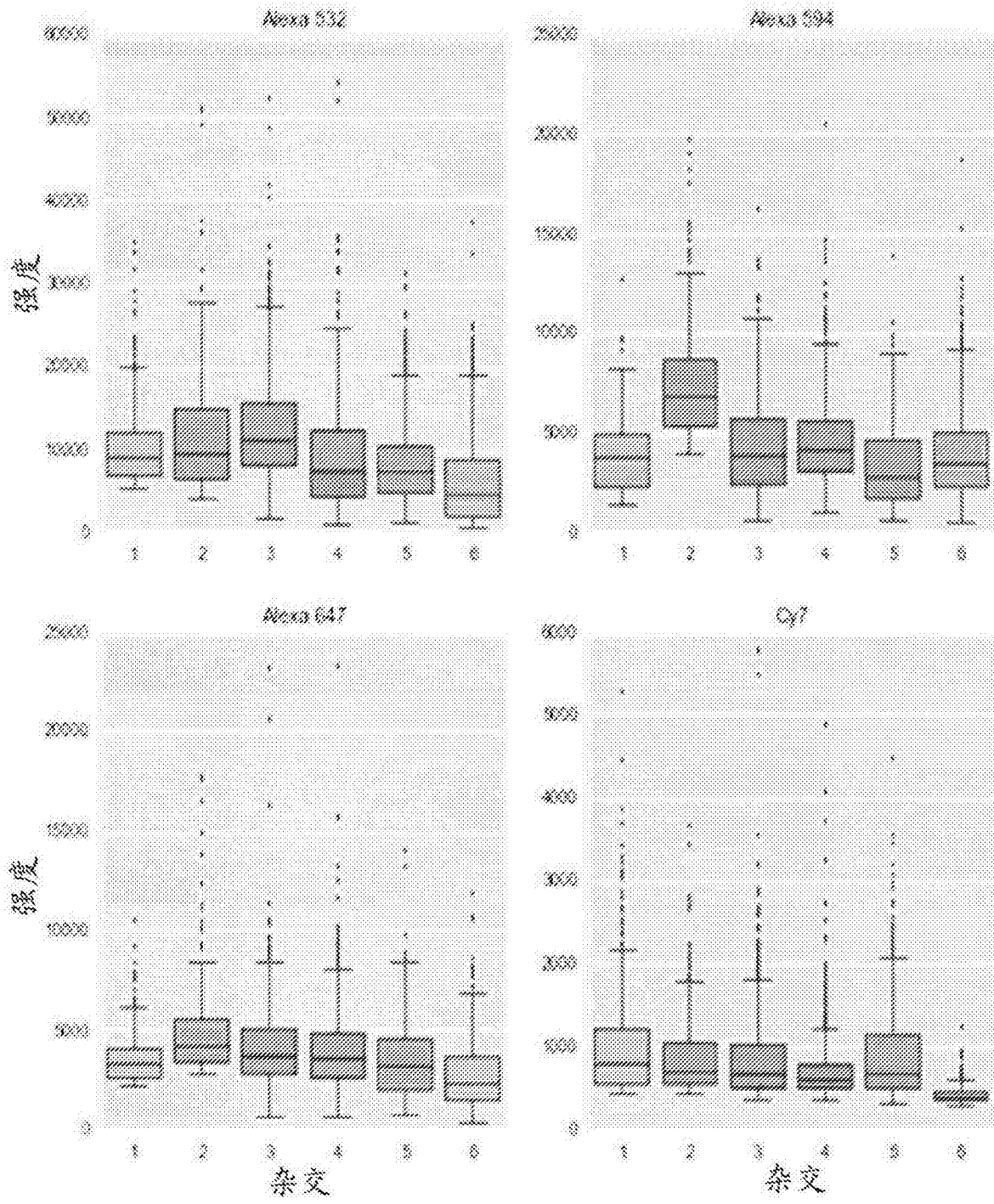


图 6

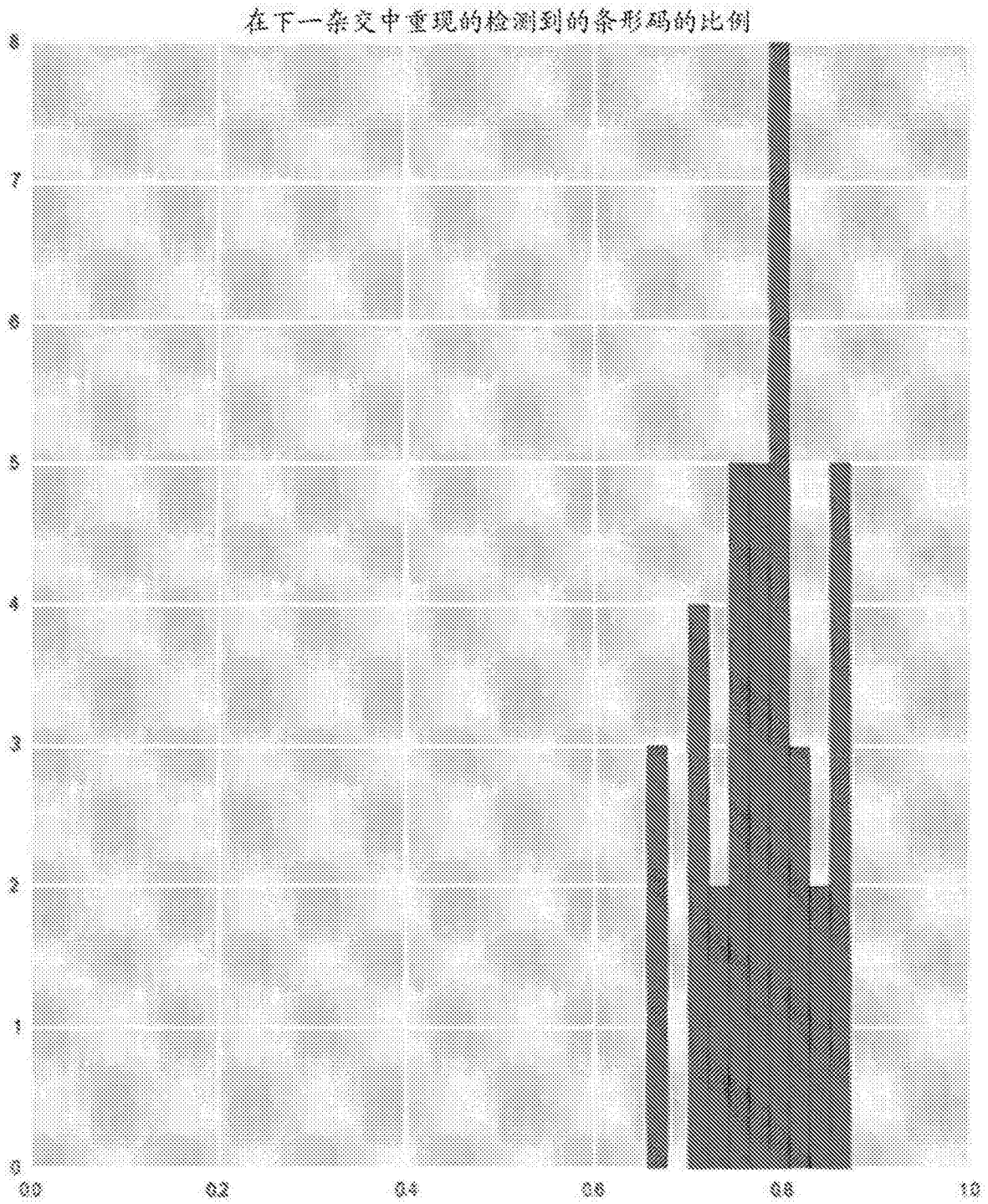


图 7

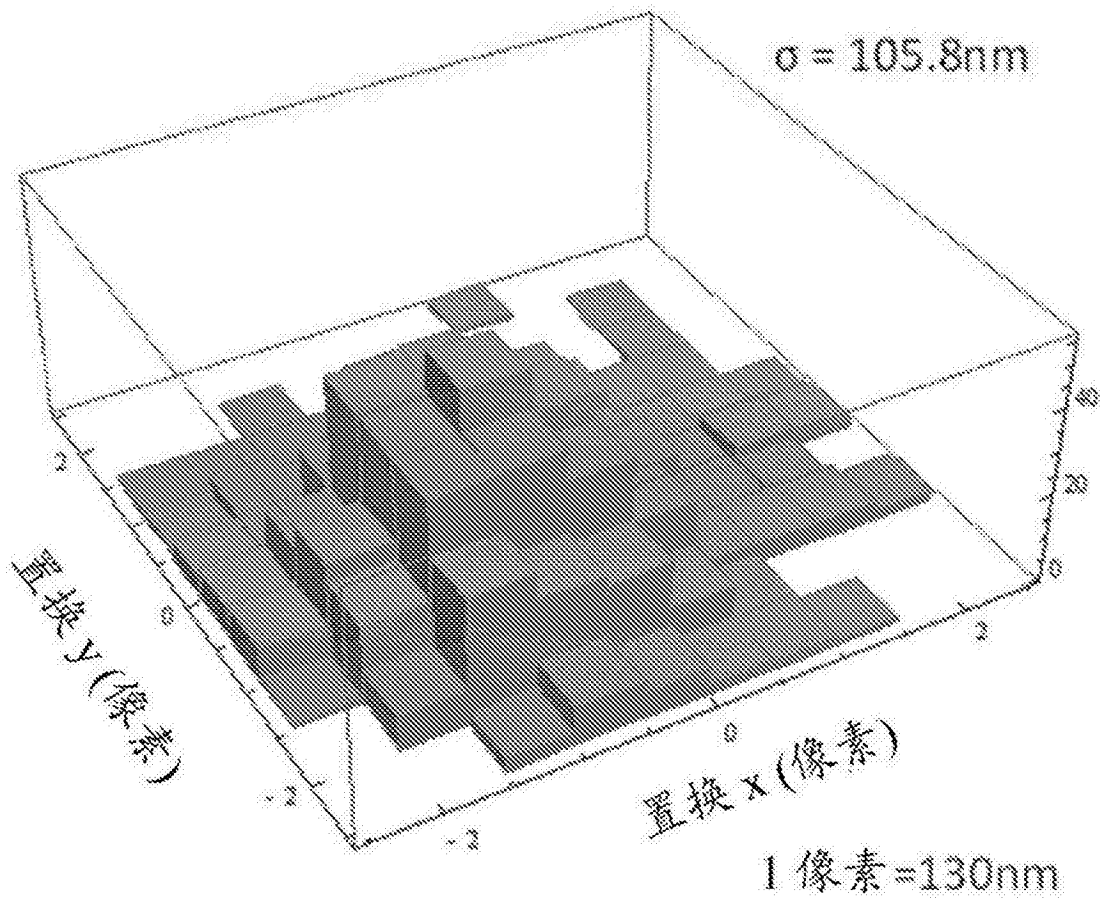


图 8

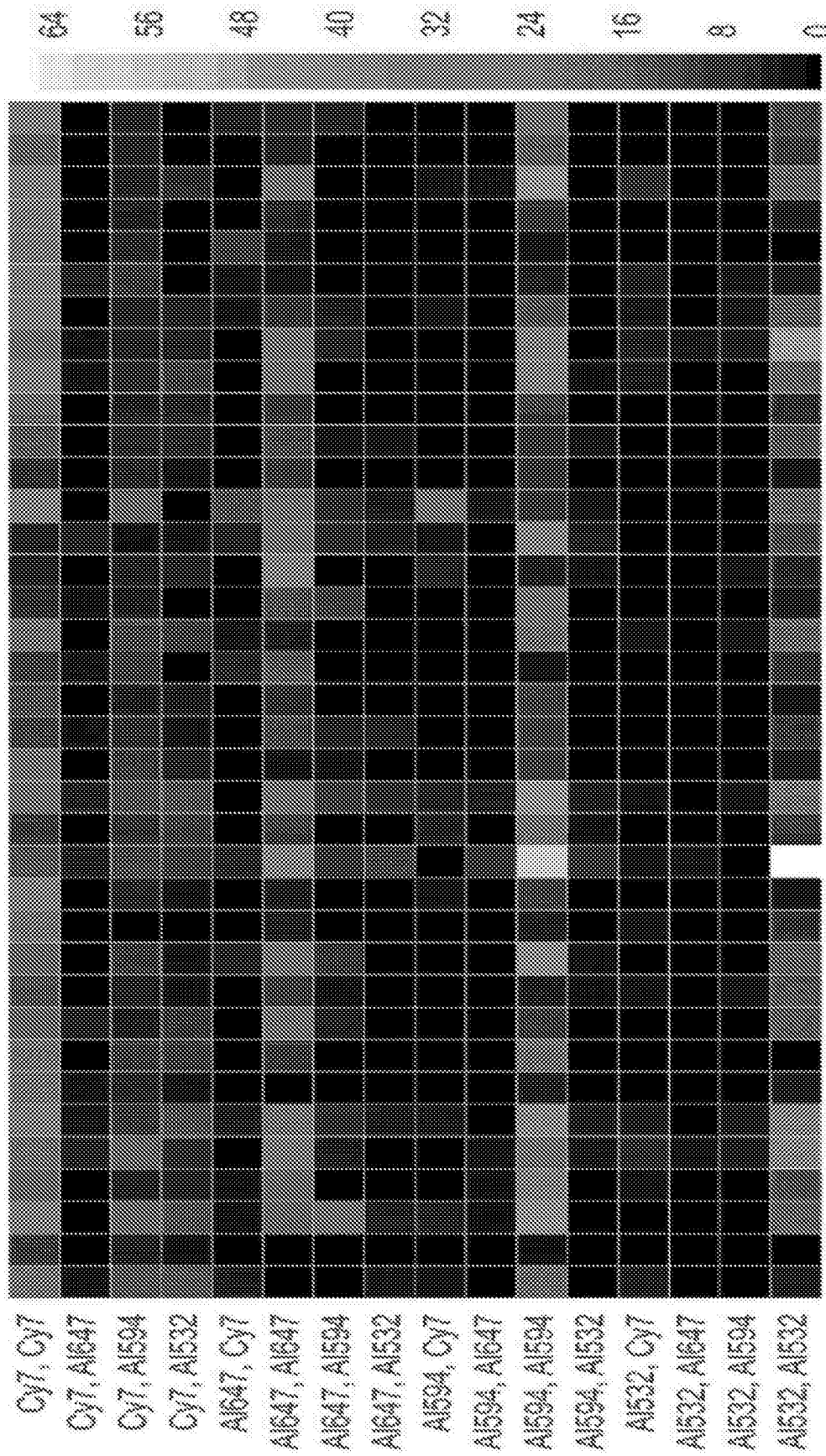


图 9

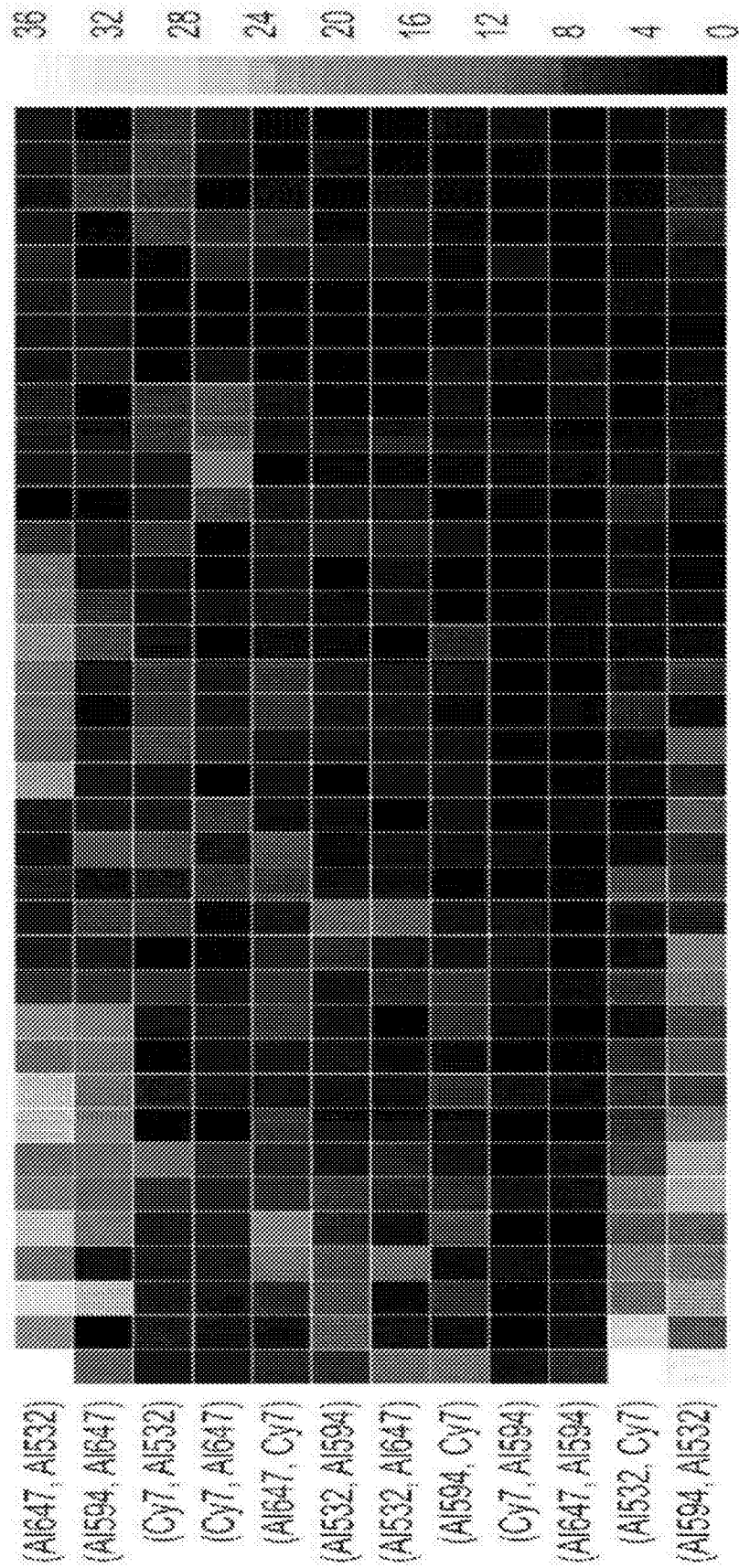


图 10

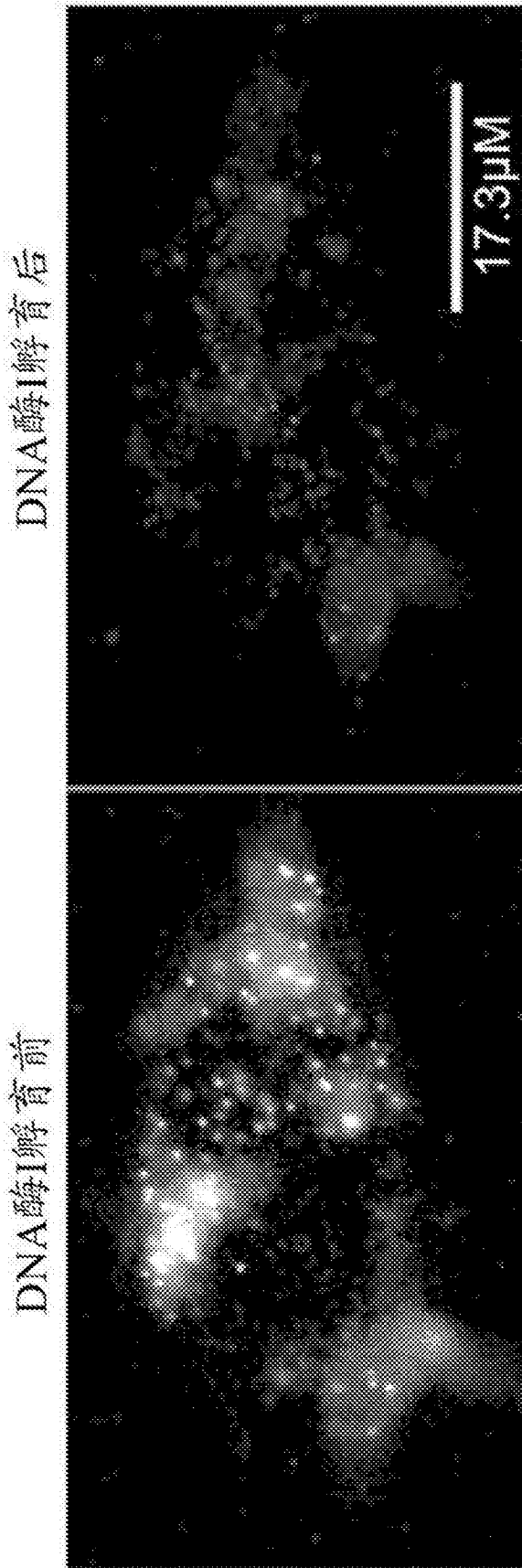


图 11

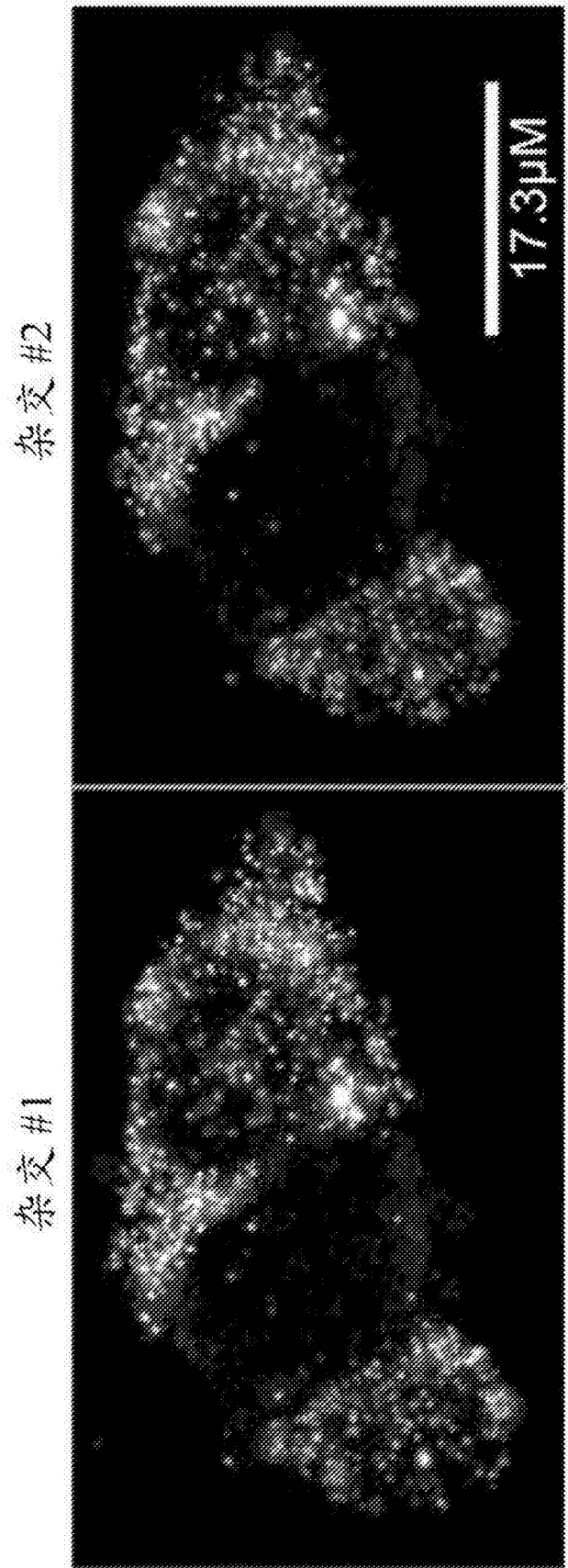


图 12

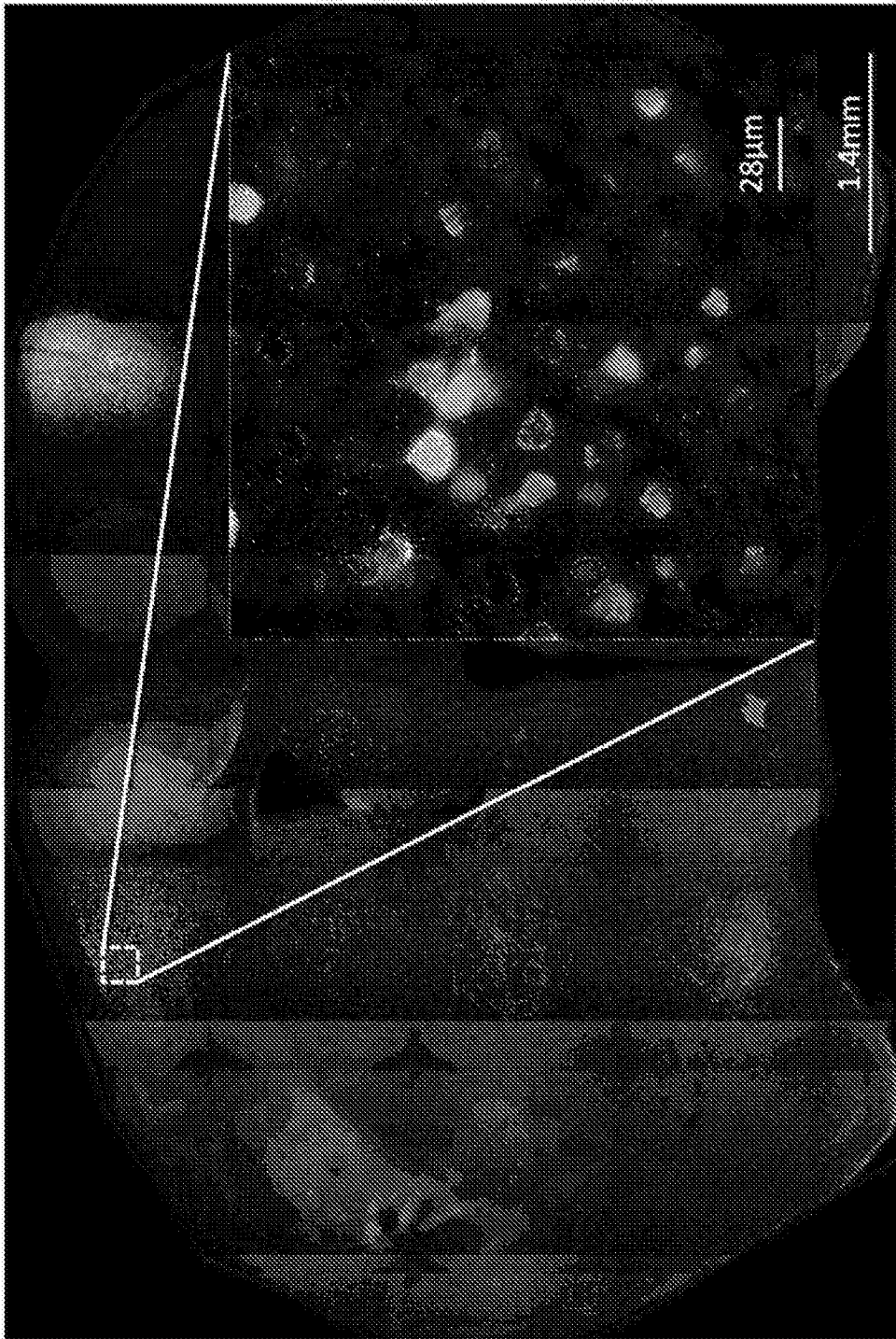


图 13

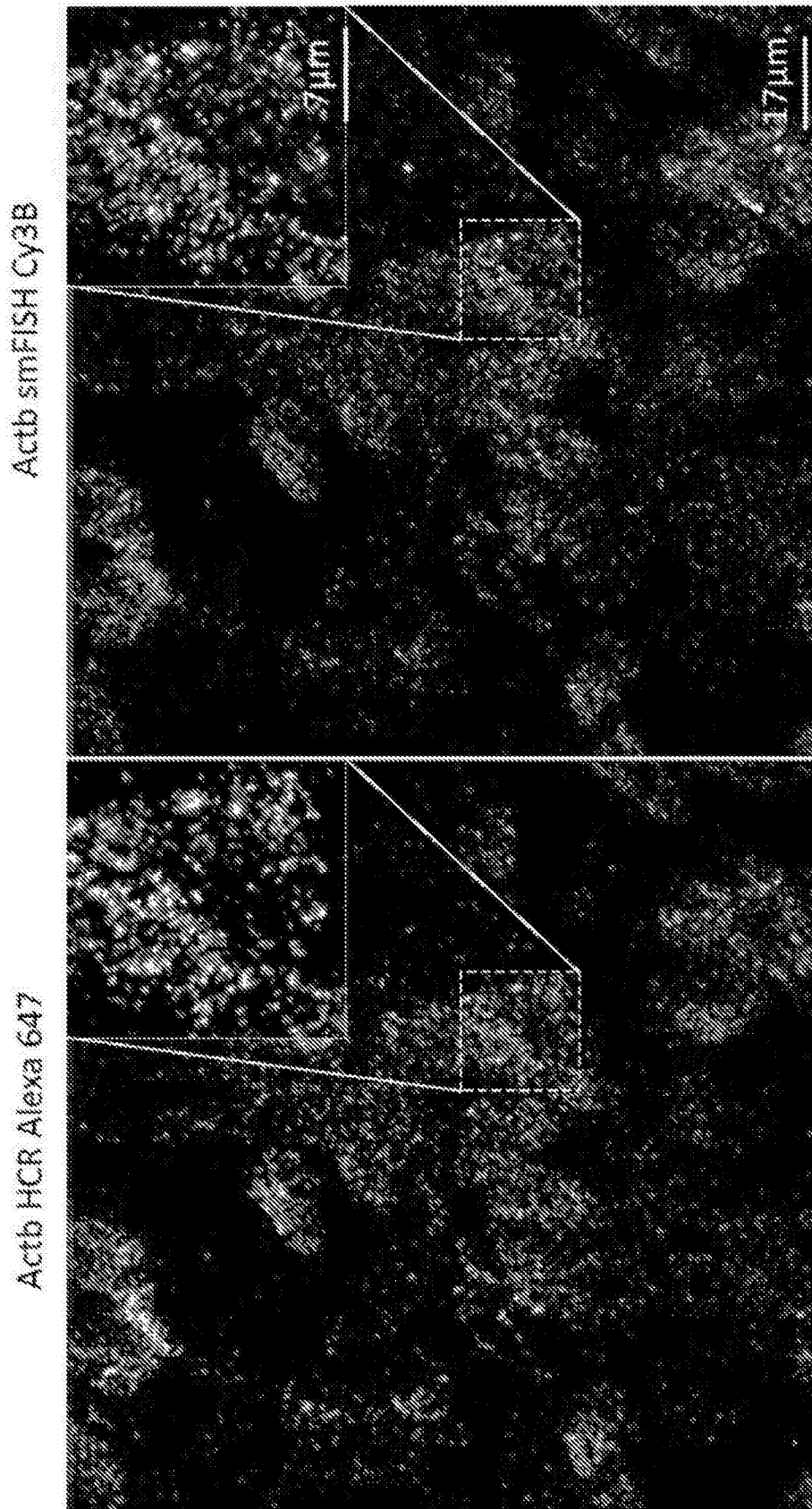


图 14

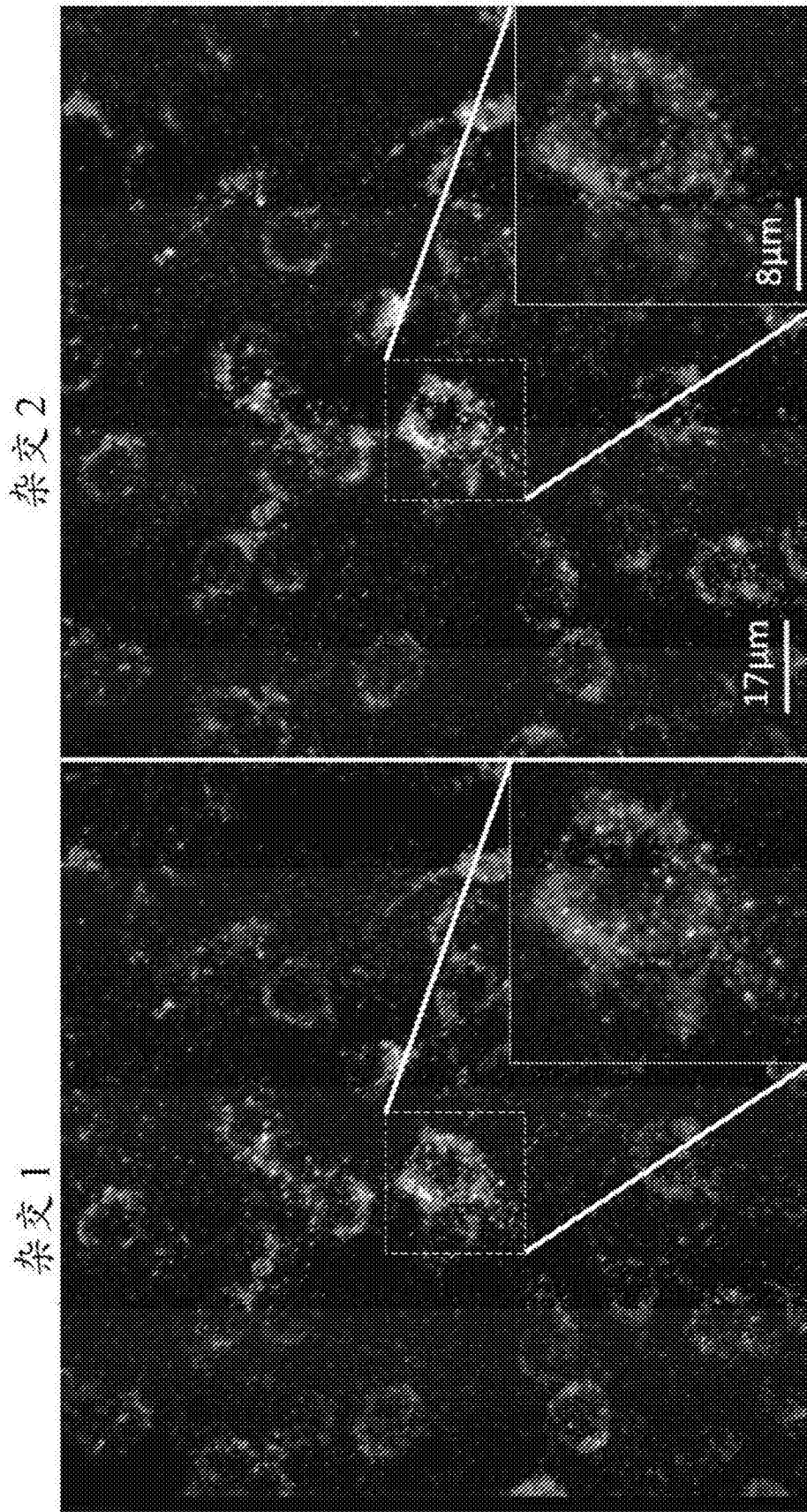


图 15

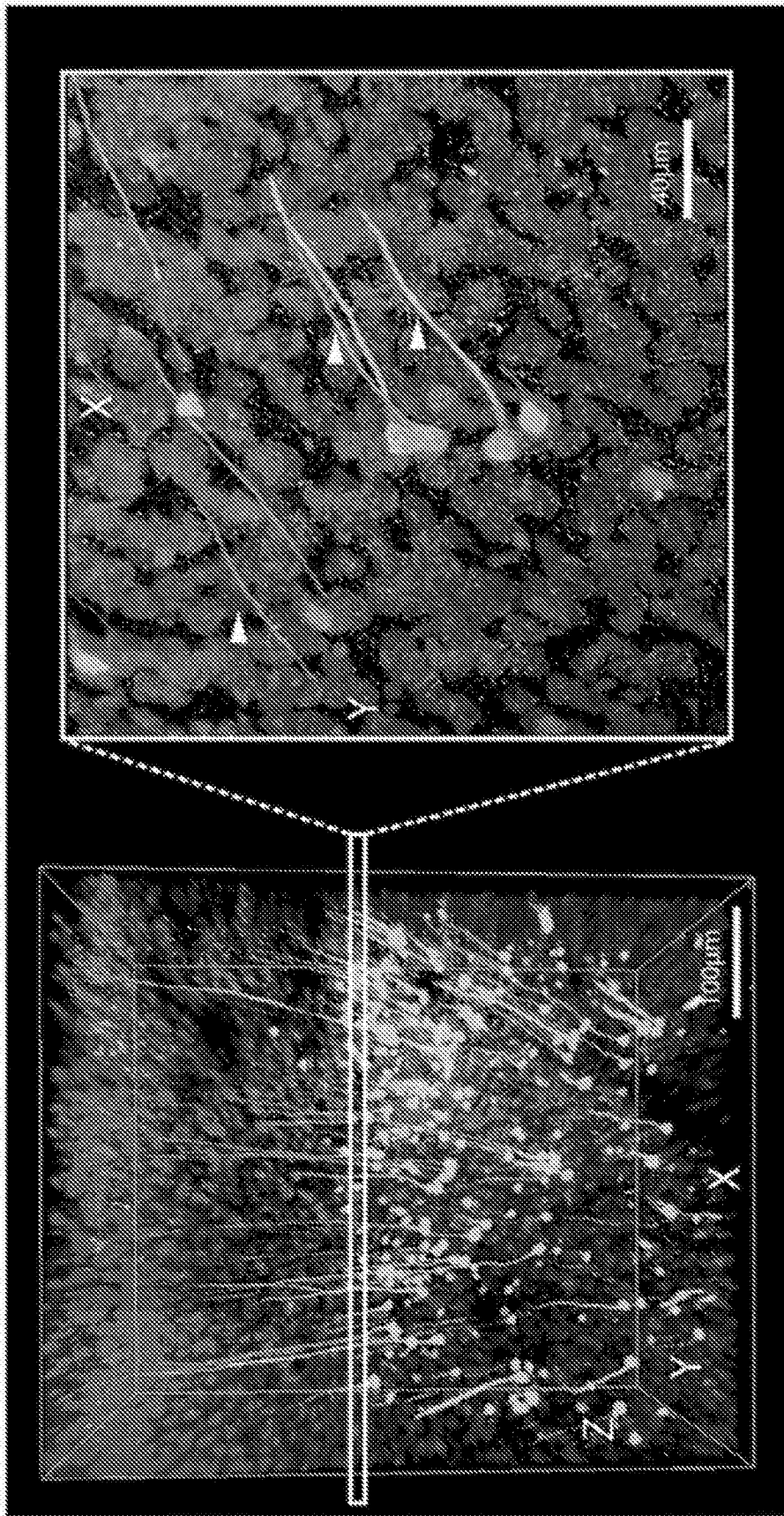


图 16

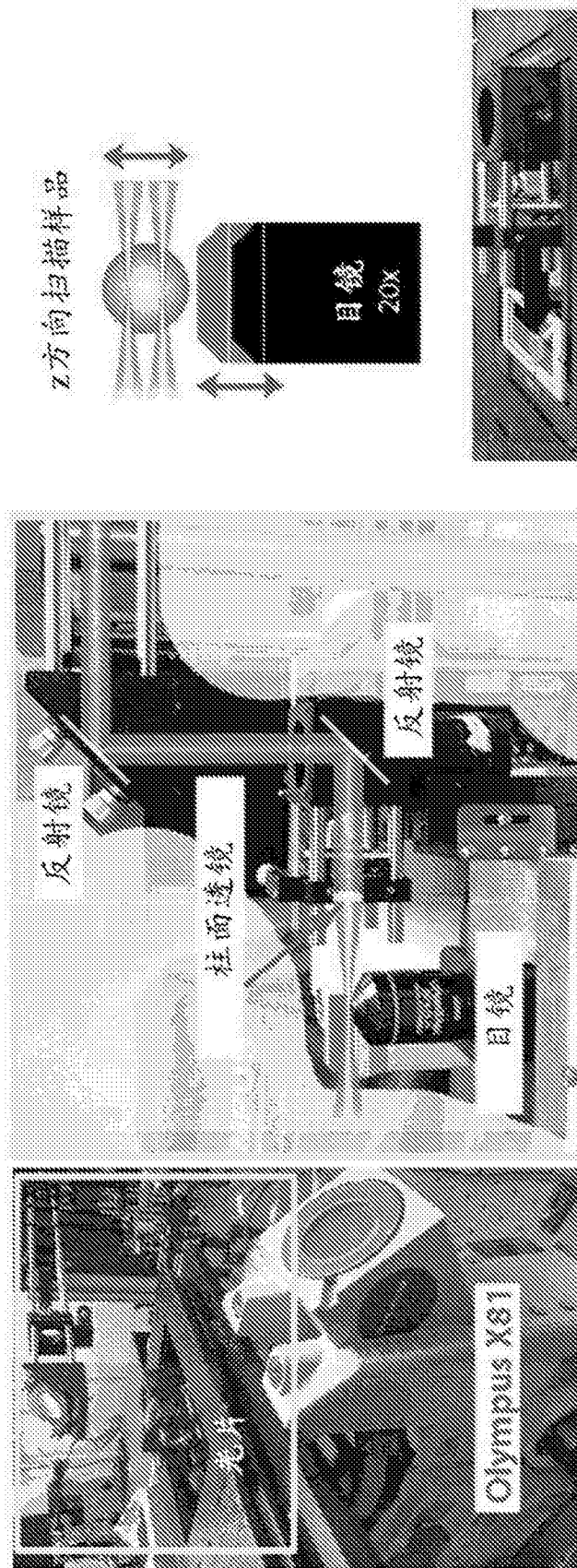


图 17

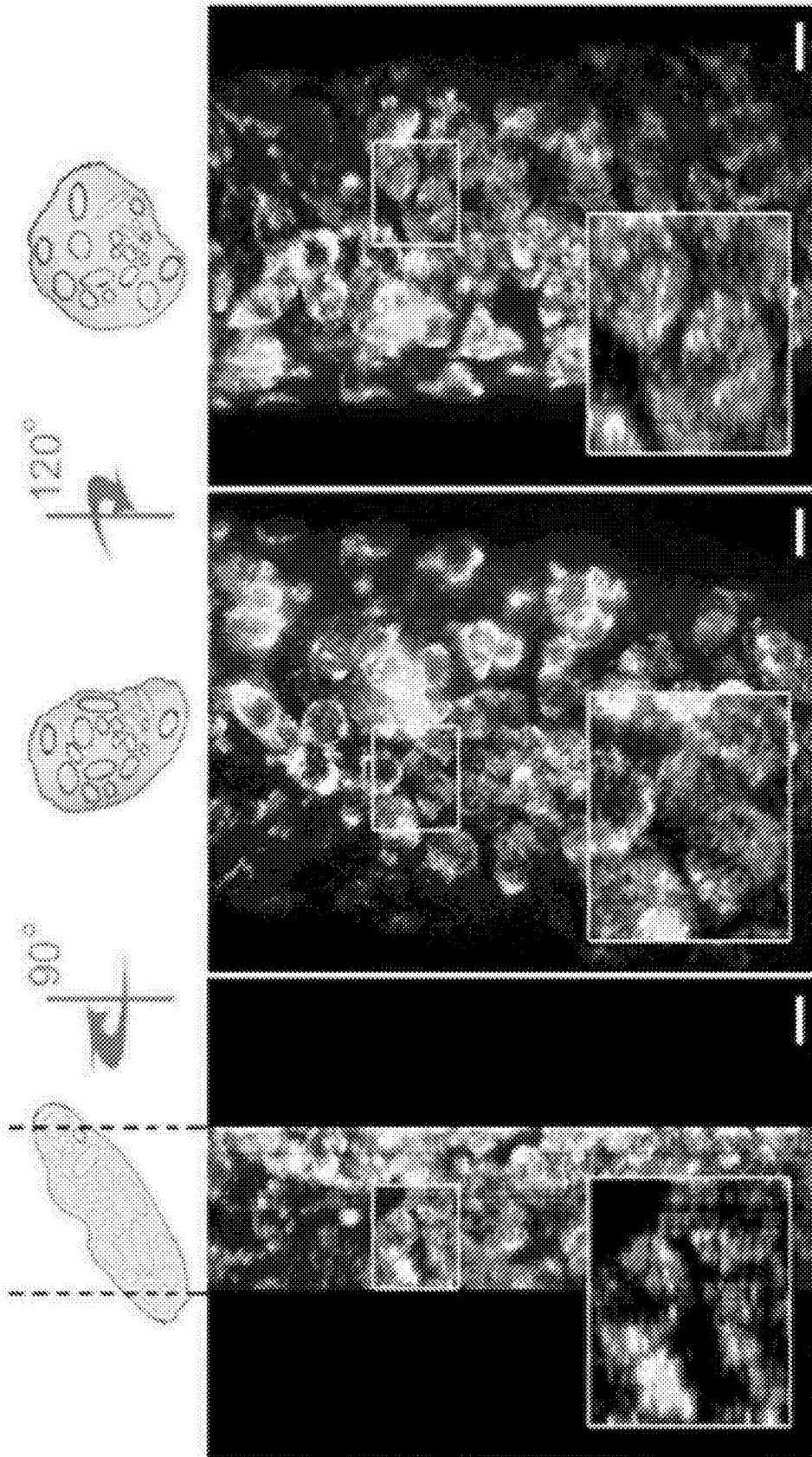


图 18

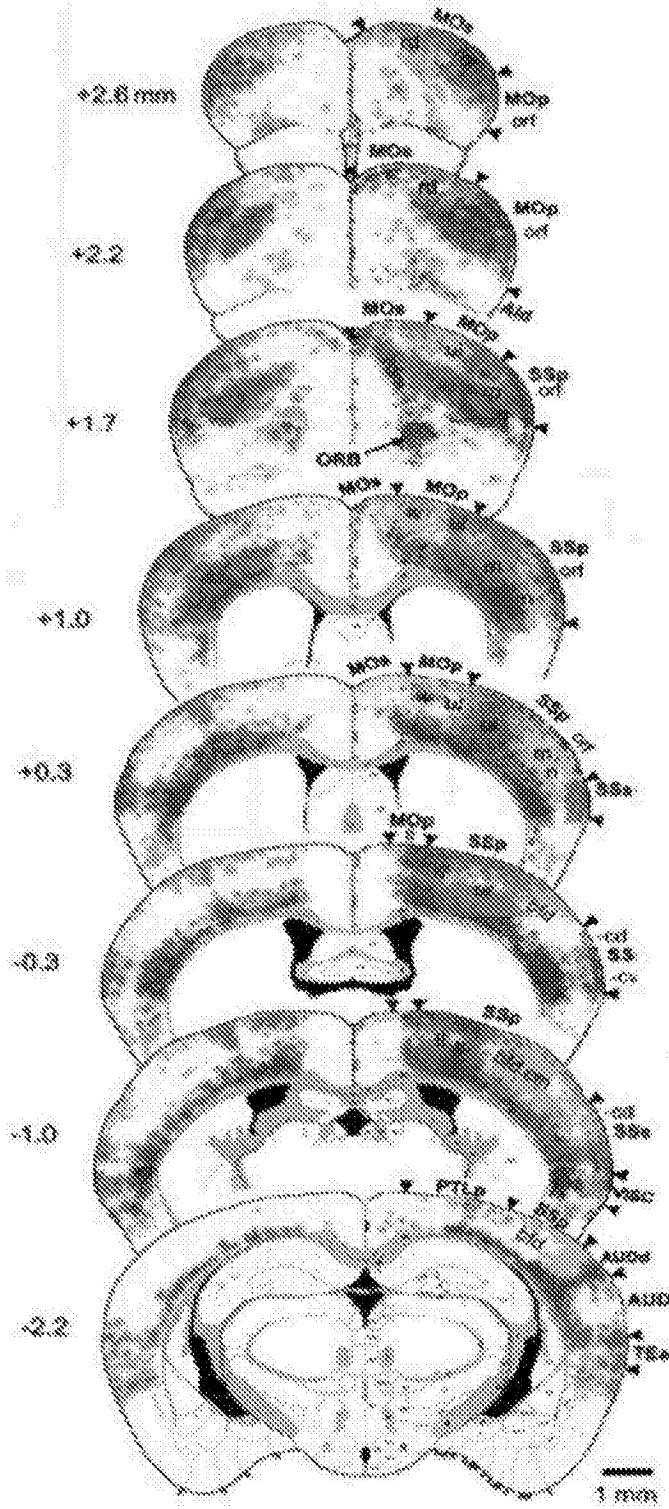


图 19

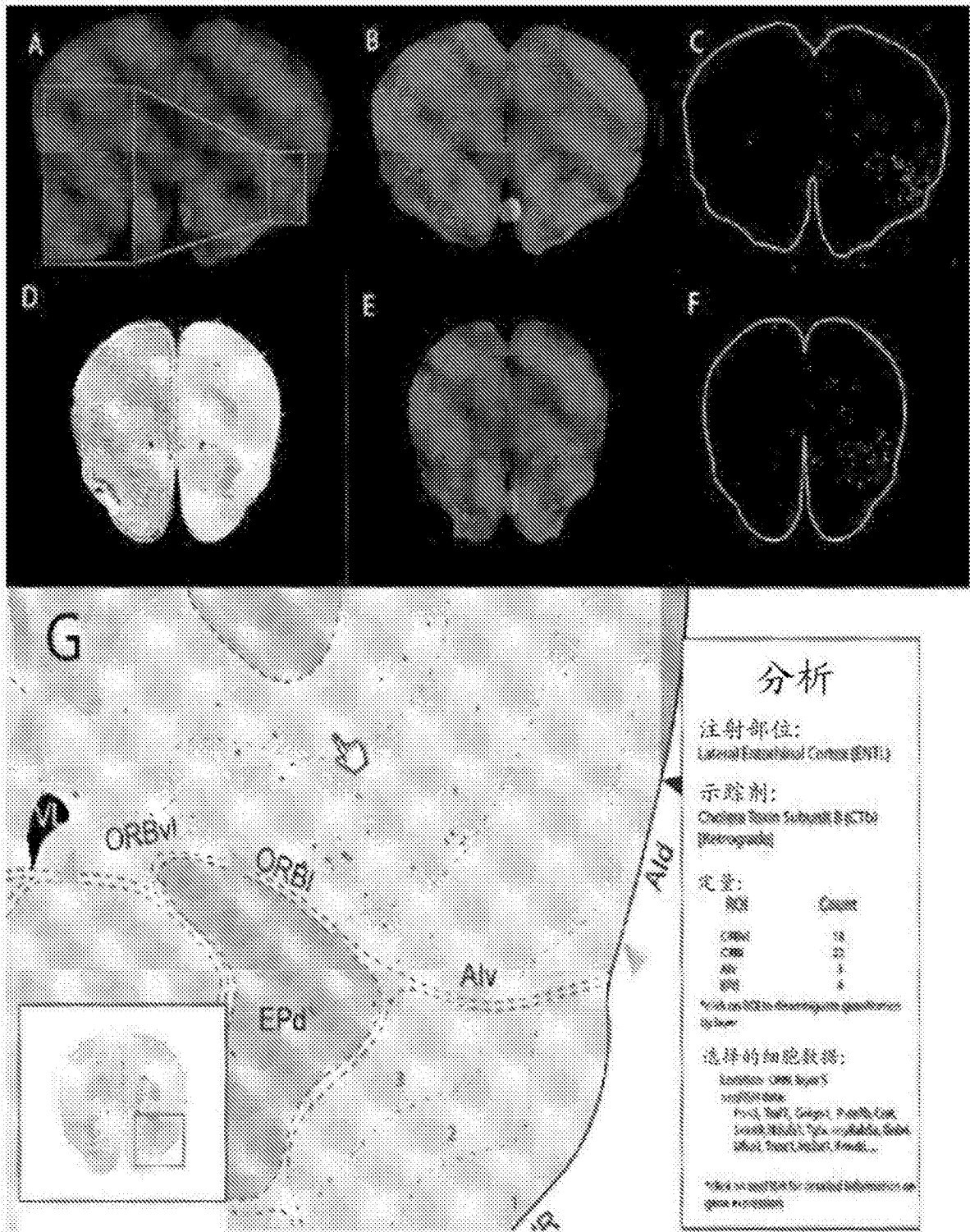


图 20

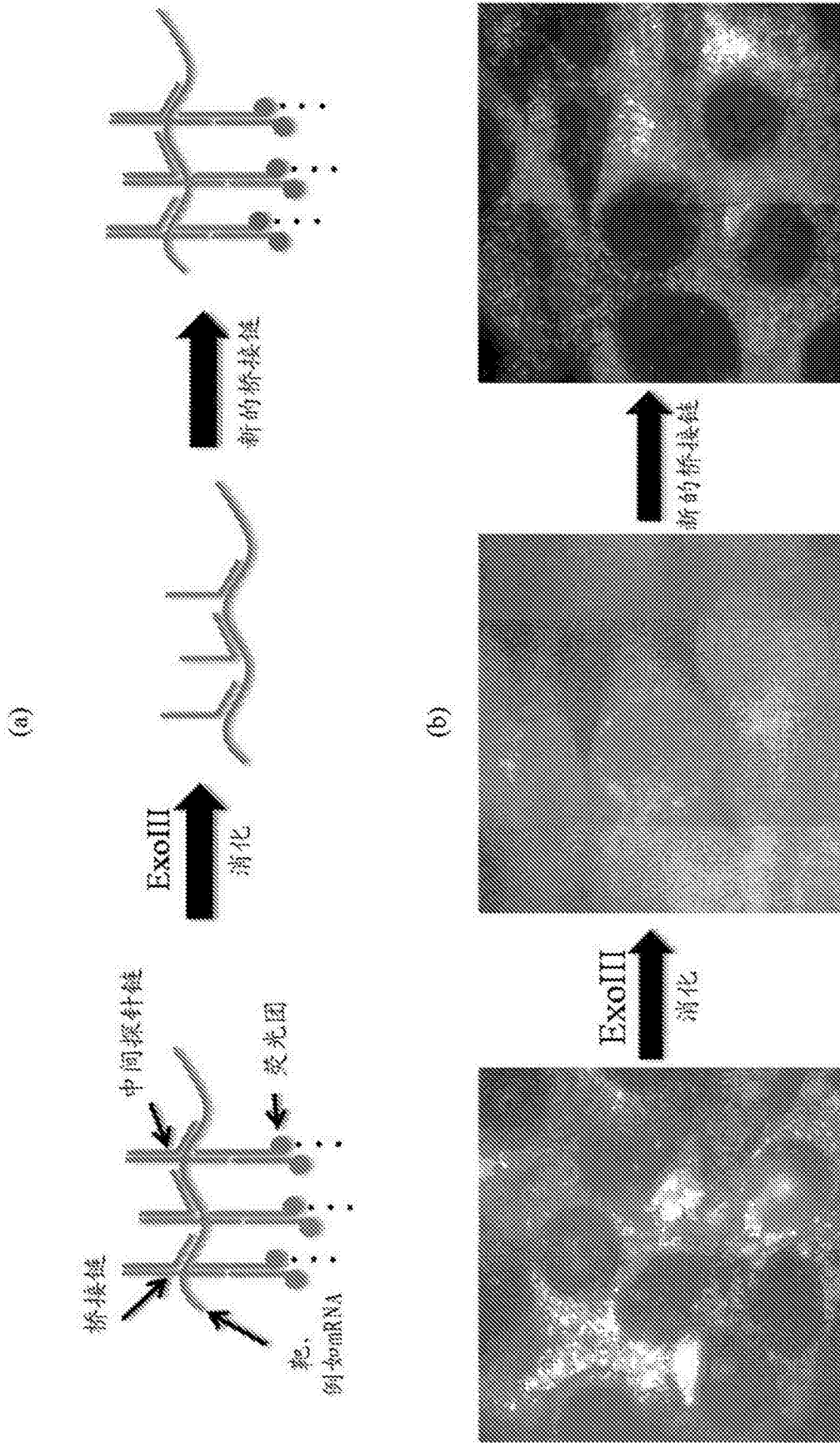


图 21

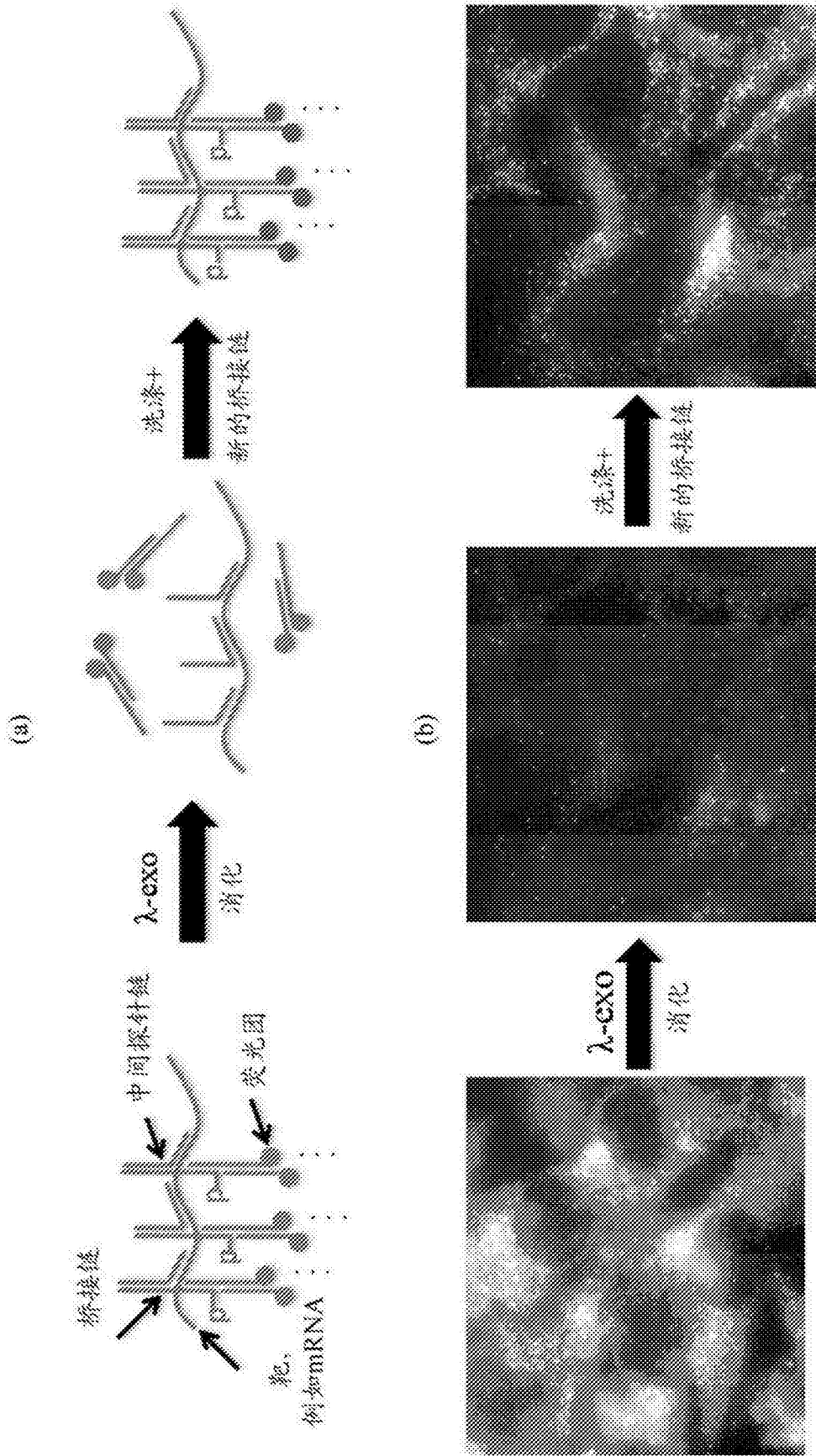


图 22

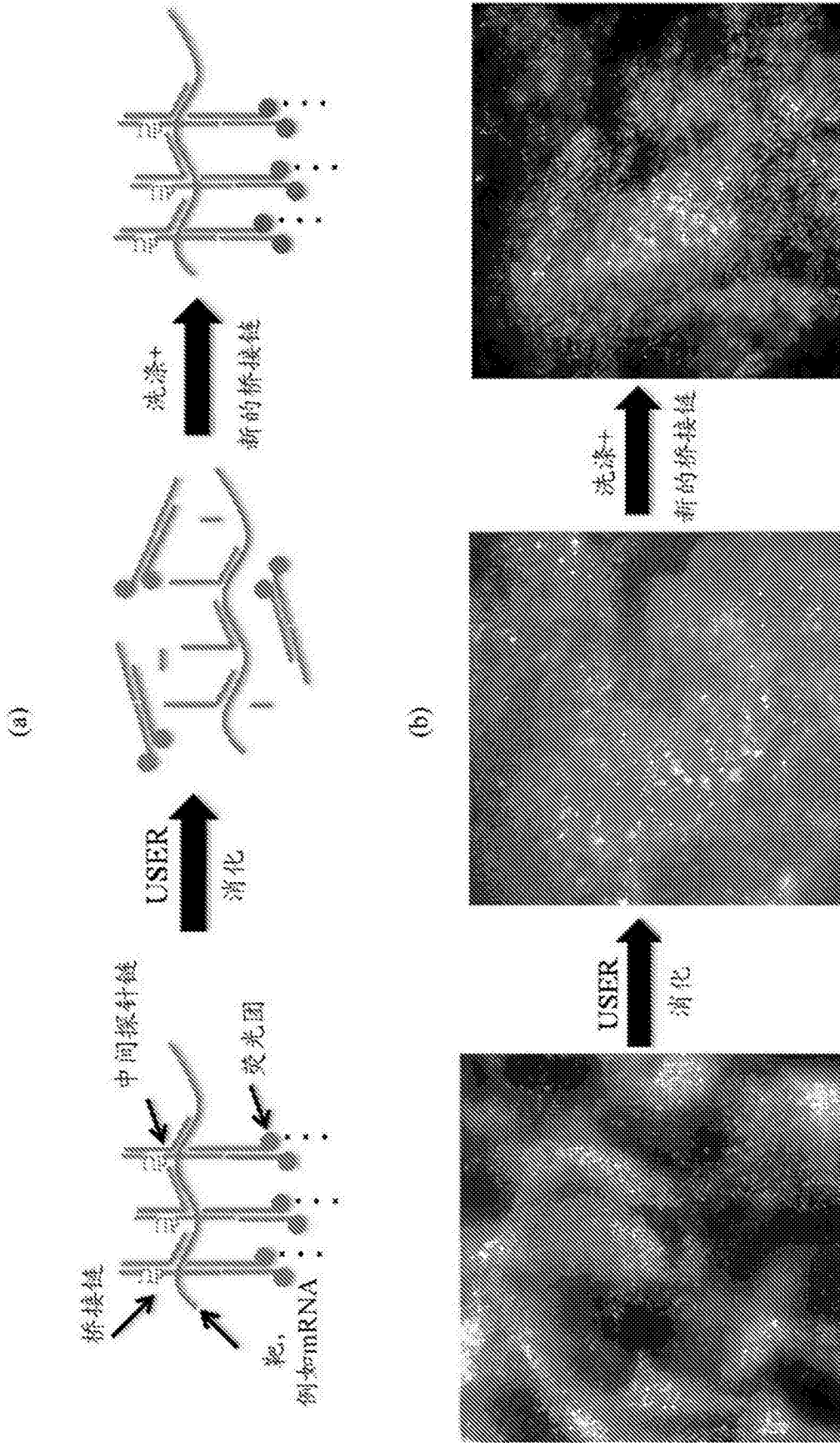


图 23

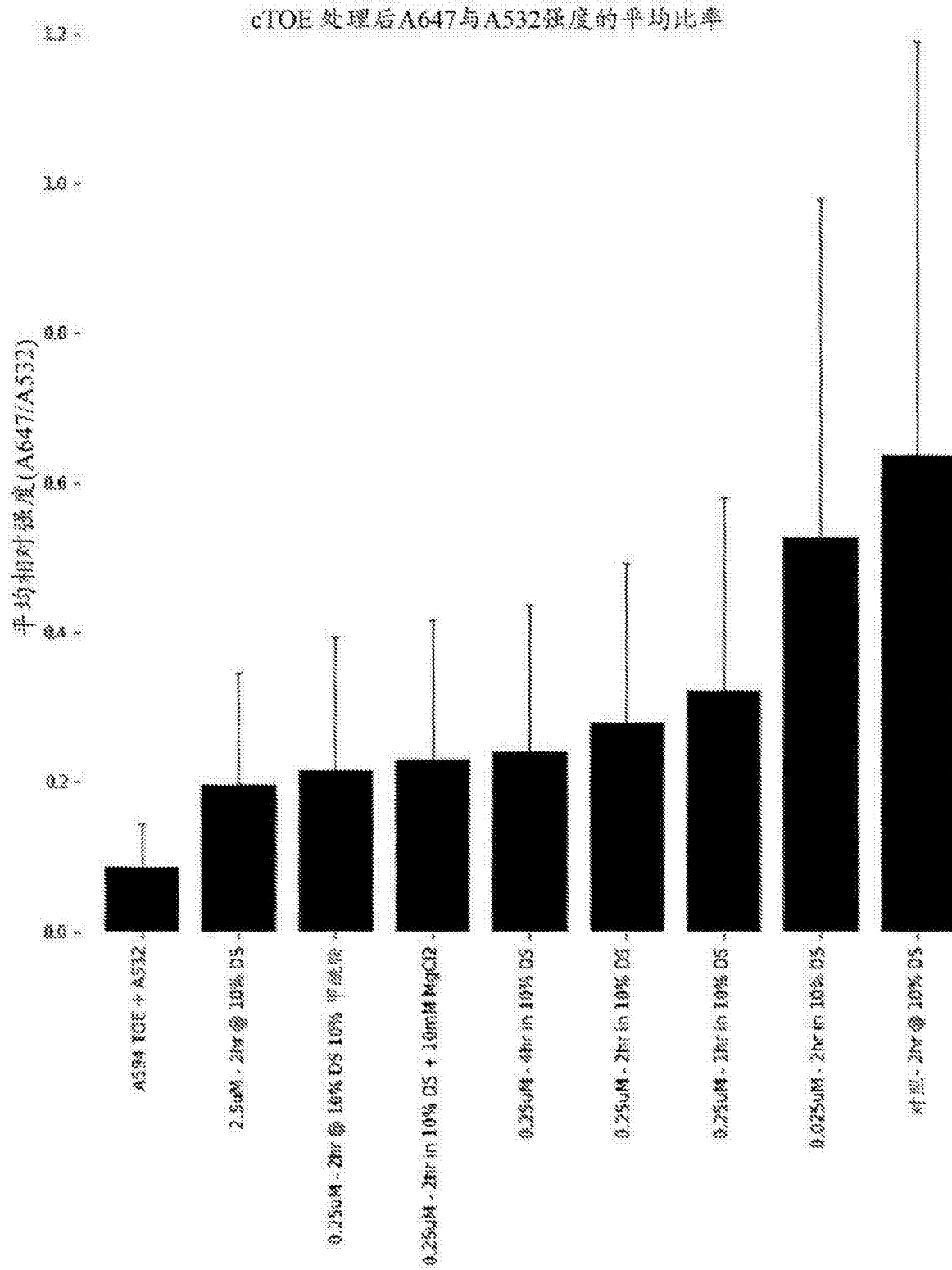


图 24

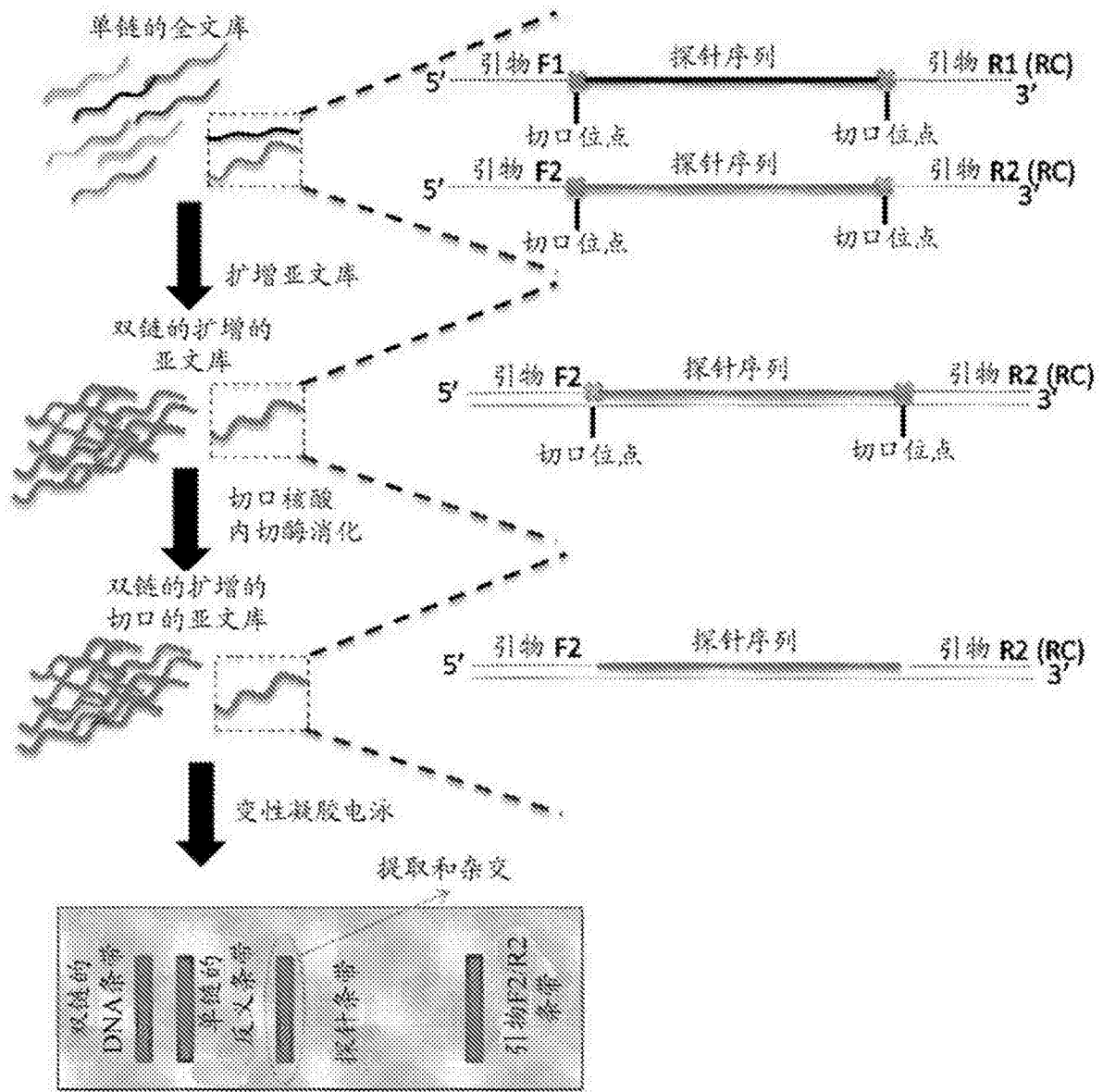


图 25



图 26