



(12) 发明专利

(10) 授权公告号 CN 110168346 B

(45) 授权公告日 2023. 03. 03

(21) 申请号 201780082227.7

C12Q 1/6841 (2018.01)

(22) 申请日 2017.11.08

(56) 对比文件

(65) 同一申请的已公布的文献号  
申请公布号 CN 110168346 A

US 2016304952 A1, 2016.10.20

US 2016304952 A1, 2016.10.20

CN 104350372 A, 2015.02.11

(43) 申请公布日 2019.08.23

CN 105556309 A, 2016.05.04

(30) 优先权数据

US 2016305856 A1, 2016.10.20

62/419,033 2016.11.08 US

CN 105531377 A, 2016.04.27

62/569,127 2017.10.06 US

CN 102171234 A, 2011.08.31

(85) PCT国际申请进入国家阶段日  
2019.07.04

CN 1570140 A, 2005.01.26

WO 0123614 A1, 2001.04.05

US 2002058248 A1, 2002.05.16

(86) PCT国际申请的申请数据  
PCT/US2017/060558 2017.11.08

Jeffrey R. Moffitt et al..High-throughput single-cell gene-expression profiling with multiplexed error-robust fluorescence in situ hybridization.

(87) PCT国际申请的公布数据  
W02018/089438 EN 2018.05.17

《Proceedings of the National Academy of Sciences》.2016,第113卷(第39期),第11046-11051页.

(73) 专利权人 哈佛学院院长及董事  
地址 美国马萨诸塞

Nikolay T et al..smiFISH and FISH-quant - a flexible single RNA detection approach with super-resolution capability.《Nucleic Acids Research》.2016,第44卷(第2期),e165. (续)

(72) 发明人 庄小威 J·R·莫菲特 王桂萍

审查员 韩雅琪

(74) 专利代理机构 中国贸促会专利商标事务所  
有限公司 11038

专利代理师 张小勇

(51) Int.Cl.

G01N 1/36 (2006.01)

C12Q 1/6806 (2018.01)

C12Q 1/6834 (2018.01)

权利要求书4页 说明书30页 附图13页

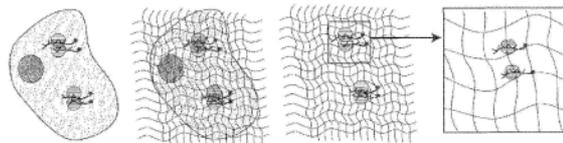
(54) 发明名称

使用MERFISH、扩展显微术和相关技术进行多重成像

例如mRNA,例如通过将核酸探针结合到样品上。其他方面一般性涉及用于此类方法的组合物或装置、用于此类方法的试剂盒等。

(57) 摘要

本发明一般性涉及显微术和涉及用于成像或测定例如细胞内的核酸或其他所需靶标的系统和方法。在某些方面,样品包含在可膨胀材料内,可膨胀材料以某种方式膨胀和成像。材料的膨胀改善了后续图像的有效分辨率。例如,这可以与其他超分辨率技术(例如STORM)和/或与诸如MERFISH的技术组合以用于测定样品内的核酸



CN 110168346 B

[接上页]

(56) 对比文件

Chen, F et al. Nanoscale imaging of RNA with expansion microscopy.《Nat Methods》.2016,第13卷(第8期),全文.

Chen F et al..Expansion microscopy.《Science》.2015,第347卷(第6221期),第543-第548页.

1. 一种方法,其包括:  
将细胞包埋到可膨胀材料中;  
将来自细胞的核酸固定至可膨胀材料,其包括将核酸暴露于多种锚定探针,所述多种锚定探针在每种核酸的序列的单个区域处将核酸固定至可膨胀材料,其中每种锚定探针包含与来自细胞的核酸的至少一部分基本上互补的核酸序列和与可膨胀材料反应的部分;  
使可膨胀材料膨胀;  
将可膨胀材料暴露于多种核酸探针;和  
测定核酸探针与固定的核酸的结合。
2. 根据权利要求1所述的方法,包括在位于每种核酸的5'末端的单个区域处将核酸固定至可膨胀材料。
3. 根据权利要求1所述的方法,包括在位于每种核酸的3'末端的单个区域处将核酸固定至可膨胀材料。
4. 根据权利要求1所述的方法,包括在位于每种核酸的中心部分的单个区域处将核酸固定至可膨胀材料。
5. 根据权利要求1-4中任一项所述的方法,其中对每种核酸使用一种独特的锚定探针。
6. 根据权利要求1-4中任一项所述的方法,其中对每种核酸使用多种独特的锚定探针。
7. 根据权利要求1-4中任一项所述的方法,其中对于每种核酸使用至少5种独特的锚定探针。
8. 根据权利要求1-4中任一项所述的方法,其中对于每种核酸使用至少10种独特的锚定探针。
9. 根据权利要求1-4中任一项所述的方法,其中多种锚定探针中的至少一些锚定探针包含亚磷酰胺修饰的寡核苷酸。
10. 根据权利要求1-4中任一项所述的方法,其中多种锚定探针中的至少一些锚定探针包含多聚dT序列。
11. 根据权利要求1所述的方法,其中将细胞包埋到可膨胀材料中包括用可膨胀材料的前体包围细胞,并使可膨胀材料的前体形成围绕细胞的可膨胀材料。
12. 根据权利要求11所述的方法,其中可膨胀材料的前体包括单体。
13. 根据权利要求11或12中任一项所述的方法,其中使前体形成可膨胀材料包括使前体聚合。
14. 根据权利要求11或12中任一项所述的方法,其中使前体形成可膨胀材料包括将前体暴露于交联剂。
15. 根据权利要求11或12中任一项所述的方法,其中前体包括丙烯酰胺。
16. 根据权利要求11或12中任一项所述的方法,其中前体包括丙烯酸酯。
17. 根据权利要求1-4、11或12中任一项所述的方法,其中可膨胀材料包括凝胶。
18. 根据权利要求1-4、11或12中任一项所述的方法,其中可膨胀材料包括聚丙烯酰胺。
19. 根据权利要求1-4、11或12中任一项所述的方法,其中可膨胀材料包括甲基丙烯酸酯。
20. 根据权利要求1-4、11或12中任一项所述的方法,包括使可膨胀材料各向同性地膨胀。

21. 根据权利要求1-4、11或12中任一项所述的方法,还包括固定细胞。
22. 根据权利要求1-4、11或12中任一项所述的方法,还包括固定细胞,其中固定细胞包括将细胞暴露于多聚甲醛。
23. 根据权利要求1-4、11或12中任一项所述的方法,其中来自细胞的核酸包含RNA。
24. 根据权利要求1-4、11或12中任一项所述的方法,其中来自细胞的核酸包含mRNA。
25. 根据权利要求1-4、11或12中任一项所述的方法,其中来自细胞的核酸包含DNA。
26. 根据权利要求1-4、11或12中任一项所述的方法,其中使可膨胀材料膨胀包括将可膨胀材料暴露于包含水的溶液。
27. 根据权利要求1-4、11或12中任一项所述的方法,其中使可膨胀材料膨胀包括将可膨胀材料暴露于相对于可膨胀材料低渗的溶液。
28. 根据权利要求1-4、11或12中任一项所述的方法,还包括,对于多种核酸探针中至少一些探针的每一种,测定核酸探针在样品内的结合。
29. 根据权利要求1-4、11或12中任一项所述的方法,还包括基于多种核酸探针的结合产生代码字;并且对于至少一些代码字,将代码字与有效代码字匹配,其中如果没有找到匹配,则对代码字应用错误纠正以形成有效代码字。
30. 根据权利要求1-4、11或12中任一项所述的方法,包括将细胞暴露于至少5种不同的核酸探针。
31. 根据权利要求1-4、11或12中任一项所述的方法,包括将细胞同时暴露于多种核酸探针。
32. 根据权利要求1-4、11或12中任一项所述的方法,包括将细胞顺序地暴露于多种核酸探针。
33. 根据权利要求1-4、11或12中任一项所述的方法,其中多种核酸探针包含具有不同序列的核酸探针的组合性组合。
34. 根据权利要求1所述的方法,其中多种核酸探针中的至少一些包含含有靶标序列的第一部分和含有一种或多种读出序列的第二部分。
35. 根据权利要求34所述的方法,其中多种核酸探针包含可区分的核酸探针,其由取自一种或多种读出序列的一种或多种读出序列的组合性组合形成。
36. 根据权利要求35所述的方法,包括将细胞暴露于包含第一信号传导实体的第一二级探针,第一二级探针能够结合核酸探针的一些读出序列,以及通过测定第一信号传导实体来测定核酸探针的结合。
37. 根据权利要求36所述的方法,还包括将细胞暴露于包含第二信号传导实体的第二二级探针,第二二级探针能够结合核酸探针的一些读出序列,以及通过测定第二信号传导实体来测定核酸探针的结合。
38. 根据权利要求1-4、11、12或34-37中任一项所述的方法,其中多种核酸探针中的至少一些包含DNA。
39. 根据权利要求1-4、11、12或34-37中任一项所述的方法,其中多种核酸探针中的至少一些包含RNA。
40. 根据权利要求1-4、11、12或34-37中任一项所述的方法,其中多种核酸探针具有10至300个核苷酸的平均长度。

41. 根据权利要求1-4、11、12或34-37中任一项所述的方法,其中多种核酸探针中的至少一些被配置为结合来自细胞的核酸。

42. 根据权利要求1-4、11、12或34-37中任一项所述的方法,其中核酸探针与细胞内靶标的至少一些结合是特异性结合。

43. 根据权利要求1-4、11、12或34-37中任一项所述的方法,其中多种核酸探针中的至少一些被配置为结合RNA。

44. 根据权利要求1-4、11、12或34-37中任一项所述的方法,其中多种核酸探针中的至少一些被配置为结合DNA。

45. 根据权利要求1-4、11、12或34-37中任一项所述的方法,包括在可膨胀材料膨胀后,以优于300nm的分辨率测定样品中核酸探针的结合。

46. 根据权利要求1-4、11、12或34-37中任一项所述的方法,包括在可膨胀材料膨胀后,以优于100nm的分辨率测定样品中核酸探针的结合。

47. 根据权利要求1-4、11、12或34-37中任一项所述的方法,包括在可膨胀材料膨胀之前,以优于30nm的有效分辨率测定样品中核酸探针的结合。

48. 根据权利要求1-4、11、12或34-37中任一项所述的方法,包括在可膨胀材料膨胀之前,以优于10nm的有效分辨率测定样品中核酸探针的结合。

49. 根据权利要求1-4、11、12或34-37中任一项所述的方法,包括通过对至少一部分可膨胀材料成像来测定核酸探针的结合。

50. 根据权利要求1-4、11、12或34-37中任一项所述的方法,包括使用光学成像技术测定核酸探针的结合。

51. 根据权利要求1-4、11、12或34-37中任一项所述的方法,包括使用荧光成像技术测定核酸探针的结合。

52. 根据权利要求1-4、11、12或34-37中任一项所述的方法,包括使用超分辨率荧光成像技术测定核酸探针的结合。

53. 根据权利要求1-4、11、12或34-37中任一项所述的方法,还包括通过测定核酸探针的结合来确定细胞的基因型。

54. 一种方法,其包括:

将多种靶标固定至可膨胀材料,其包括将靶标暴露于多种锚定探针,所述多种锚定探针在每种靶核酸的单个区域处将核酸固定至可膨胀材料,其中多种锚定探针中的每一种锚定探针包含与靶标的至少一部分基本上互补的核酸序列和与可膨胀材料反应的部分,并且其中固定至可膨胀材料的多种靶标中至少50%的靶标固定在单个区域处;和

使可膨胀材料膨胀。

55. 一种方法,其包括:

将多种核酸固定至可膨胀材料,其包括将多种核酸暴露于多种锚定探针,所述多种锚定探针在每种靶核酸的单个区域处将核酸固定至可膨胀材料,其中多种锚定探针中的每一种锚定探针包含与多种核酸的至少一部分基本上互补的核酸序列和与可膨胀材料反应的部分,并且其中固定至可膨胀材料的多种核酸中至少50%的核酸固定在单个区域处;

将可膨胀材料暴露于多种核酸探针;

使可膨胀材料膨胀;和

测定核酸探针与固定的核酸的结合。

56. 一种物品,其包含:

可膨胀材料,其包含包埋的细胞和经多种锚定探针固定至可膨胀材料的多种核酸,每一种锚定探针包含与来自多种核酸的至少一部分基本上互补的核酸序列和与可膨胀材料反应的部分,其中固定至可膨胀材料的多种核酸中至少50%的核酸固定在单个区域处。

57. 一种物品,其包含:

聚合物,其包含膨胀的细胞和经多种锚定探针固定至聚合物的多种核酸,每一种锚定探针包含与多种核酸的至少一部分基本上互补的核酸序列和与可膨胀材料反应的部分,其中细胞在聚合物内膨胀至其正常大小的至少5倍,并且其中固定至可膨胀材料的多种核酸中至少50%的核酸固定在单个区域处。

## 使用MERFISH、扩展显微术和相关技术进行多重成像

[0001] 相关申请

[0002] 本申请要求由Zhuang等人于2016年11月8日提交的题为“Matrix Imprinting and Clearing”的美国临时专利申请序列号62/419,033和Zhuang等人于2017年10月6日提交的题为“Multiplexed Imaging Using MERFISH, Expansion Microscopy, and Related Technologies”的美国临时专利申请序列号62/569,127的权益。这些中的每一篇都通过引用整体并入本文。

[0003] 政府资助

[0004] 本发明是在美国国立卫生研究院授予的基金号OD022125下由美国政府支持完成的。美国政府拥有本发明的某些权利。

[0005] 领域

[0006] 本发明一般性涉及显微术和涉及用于成像或测定例如细胞内的核酸或其他所需靶标的系统和方法。

[0007] 背景

[0008] 基于图像的单细胞转录组学(其中通过成像鉴定、计数和原位定位RNA种类)天然地保留RNA的天然空间背景。多重错误-稳健荧光原位杂交(MERFISH)是用于单细胞转录组成像的方法,并且已经证明在单细胞中分析数百至数千个RNA。在MERFISH中,通过组合标记方法鉴定RNA,所述组合标记方法使用错误稳健的二进制条形码和随后连续轮的单分子荧光原位杂交(smFISH)以读出这些条形码来编码RNA种类。RNA鉴定的准确性依赖于个体RNA分子的空间分离信号,这限制了可以测量的RNA的密度并使得大量高丰度RNA的多重成像具有挑战性。因此,需要改进成像技术。

[0009] 概述

[0010] 本发明一般性涉及显微术和涉及用于成像或测定例如细胞内的核酸或其它所需靶标的系统和方法。在一些情况下,本发明的主题涉及相关的产品、特定问题的替代解决方案和/或一个或多个系统和/或物品的多种不同用途。

[0011] 在一个方面,本发明一般性涉及包含可膨胀材料的制品。该材料可包括包埋的细胞和固定至可膨胀材料的多个核酸。在一些情况下,固定至可膨胀材料的多个核酸中的至少50%的核酸固定在单一点(single points)上。

[0012] 根据另一方面,本发明一般性涉及包含扩增细胞和固定在聚合物上的多个核酸的聚合物。在一些实施方案中,细胞在聚合物内膨胀至其正常尺寸的至少5倍。在某些情况下,固定至可膨胀材料的多个核酸中至少50%固定在单一点上。

[0013] 另一方面,本发明一般性涉及一种方法。在一组实施方案中,该方法包括将细胞包埋到可膨胀材料中,将来自细胞的核酸固定至可膨胀材料,使可膨胀材料膨胀,将可膨胀材料暴露于多个核酸探针,以及测定核酸探针与固定的核酸的结合。

[0014] 在另一组实施方案中,该方法包括将多个靶标固定至可膨胀材料,并使可膨胀材料膨胀。在一些情况下,固定至可膨胀材料的多个靶标中至少50%的靶标固定在单一点上。

[0015] 在另一组实施方案中,该方法包括将多个核酸固定至可膨胀材料,将可膨胀材料暴露于多个核酸探针,使可膨胀材料膨胀,以及测定核酸探针与固定的核酸的结合。在某些实施方案中,固定至可膨胀材料的多个核酸中至少50%的核酸固定在单一的点上。

[0016] 另一方面,本发明涵盖制备本文所述的一个或多个实施方案(例如MERFISH和扩展显微术)的方法。在另一方面,本发明涵盖使用本文所述的一个或多个实施方案(例如MERFISH和扩展显微术)的方法。

[0017] 当结合附图考虑本发明的各种非限制性实施方案的以下详细描述时,本发明的其他优点和新颖特征将变得显而易见。

[0018] 附图简要说明

[0019] 将参考附图通过示例的方式描述本发明的非限制性实施方案,附图是示意性的并且不旨在按比例绘制。在附图中,所说明的每个相同或几乎相同的组分通常由单一数字表示。为了清楚起见,并非每个组分都标记在每个图中,也没有示出本发明的每个实施方案的每个组分,其中说明对于本领域普通技术人员来说理解本发明不是必需的。在附图中:

[0020] 图1A-1E示出了未膨胀样品中RNA的MERFISH测量;

[0021] 图2A-2H示出了在本发明的一个实施方案中膨胀样品中RNA的MERFISH测量;

[0022] 图3A-3H示出了在本发明的又一个实施方案中与MERFISH组合的膨胀样品中的成像;和

[0023] 图4A-4C示意性地示出了在一个或多个锚点处的靶标的附着。

[0024] 详述

[0025] 本发明一般性涉及显微术和涉及用于成像或测定例如细胞内的核酸或其他所需靶标的系统和方法。在某些方面,样品包含在可膨胀材料内,可膨胀材料以某种方式膨胀和成像。材料的膨胀改善了后续图像的有效分辨率。例如,这可以与其他超分辨率技术(例如STORM)和/或与诸如MERFISH的技术组合以用于测定样品内的核酸例如mRNA,例如通过将核酸探针结合到样品上。其他方面一般性涉及用于此类方法的组合物或装置、用于此类方法的试剂盒等。

[0026] 在某些方面,样品包埋或包含在可膨胀材料例如聚电解质凝胶中。在一些情况下,样品可以是例如细胞或其他生物结构或非生物结构。在可膨胀材料膨胀时,样品在物理上而不是在光学上有效放大。样品的组分通过材料的膨胀彼此分开,但通常保持相当的几何形状或拓扑结构,特别是当材料的膨胀基本上各向同性地发生时。因此,最初彼此接近的组分现在更加分开,并且能够更容易地检测,例如使用成像或其他技术。

[0027] 膨胀技术可以与用于测定样品中的核酸例如mRNA的技术组合。在一些情况下,例如,在可膨胀材料膨胀之前,样品内的核酸附着于(例如,共价地)可膨胀材料。然后可以例如使用(例如重复地)暴露于核酸探针(其可以用于测定核酸在材料内的分布)来检测核酸。例如,在诸如MERFISH(多重错误-稳健荧光原位杂交)的技术中,将样品暴露于不同轮的核酸探针,并且能够使用荧光或其他技术测定核酸的结合。在某些情况下,可以使用相对少量的标记来鉴定相对大量的不同靶标,例如,通过使用各种组合方法。在一些实施方案中,还可以使用错误检查和/或错误纠正代码来增强鉴定。参见,例如,Zhuang等人的题为“Systems and Methods for Determining Nucleic Acids”的美国专利申请序列号15/329,683,其于2017年8月3日公开为美国专利申请公开号2017/0220733,其全部内容通过引

用并入本文。

[0028] 在一些情况下,可以在膨胀之前相对于可膨胀材料固定一种或多种靶标,例如核酸如mRNA。在某些实施方案中,靶标在单个点(a single point)处附着到可膨胀材料。这在一些实施方案中是重要的,因为在两个或更多个点处附着的靶标可以与材料一起膨胀,这可能使得以高分辨率分辨靶标更加困难。例如,参考图4A,靶标10和11显示在可膨胀材料20内。例如,靶标10和11可以是核酸。在图4B中,靶标10和11在两个点处(例如,它们的端部)附着,并且在通过可膨胀材料膨胀时,靶标10和11与膨胀材料的其余部分一起膨胀(例如,“拉伸”),因为它们的锚定点(显示为圆点(dots))随着膨胀材料一起膨胀。相反,在图4C中,靶标10和11仅在一个点处(显示为圆点(a dot))附着,并且在可膨胀材料膨胀时,靶标不会膨胀或拉伸。因此,在图4C中,靶标10和11现在更加分开,并且可以更容易地测定,而在图4B中,靶标10和11可能仍然是物理上重叠的,并且因此尽管有膨胀过程但仍可能产生不可区分的信号。因此,在本发明的某些实施方案中,至少一些靶标例如核酸在单个点处附着至可膨胀材料。在一些实施方案中,分子的行为可以是以与图4C中所述相同的方式,即使其通过多个附着物附着于凝胶,例如,如果附着物沿着分子的长度彼此足够接近。

[0029] 另外,可以从样品中“清除”可能有助于背景的样品的组分,例如蛋白质、脂质和/或其他非靶标,以改善在某些情况下的测定。然而,例如通过将核酸(或其他靶标)附着到可膨胀材料上,可以防止核酸或其他所需靶标被清除,如上所述。以这种方式,可膨胀材料可以与诸如MERFISH的技术组合以促进mRNA、其他核酸或其他靶标的检测。这种技术的非限制性实例可以在例如题为“Matrix Imprinting and Clearing”的美国专利申请公开号62/419,033中找到,其全部内容通过引用并入本文。

[0030] 可以使用多种技术来测定结合,包括光学技术,例如荧光显微术。在一些情况下,空间位置可以以超分辨率确定,或者以比光的波长或衍射极限更好的分辨率确定(尽管在其他实施方案中,不需要超分辨率)。例如,可以使用诸如STORM(随机光学重建显微术)的技术。参见,例如,Zhuang等人于2010年11月23日授权的题为“Sub-Diffraction Limit Image Resolution and Other Imaging Techniques”的美国专利号7,838,302,其全部内容通过引用并入本文。

[0031] 以上讨论是本发明的一个实施方案的非限制性实例,其能够用于测定样品中的核酸或其他靶标。然而,其他实施方案也是可能的。因此,更一般地,本发明的各个方面涉及用于成像或测定例如在细胞或其他样品中的核酸或其他所需靶标的各种系统和方法。

[0032] 如上所述,在一个方面,样品包埋或包含在可膨胀材料内。样品可以是任何合适的样品,并且在一些实施方案中可以是生物样品。在一些情况下,样品含有DNA和/或RNA,例如可以在样品中测定的DNA和/或RNA。(在其他实施方案中,可以测定样品中的其他靶标。)在一些情况下,样品可以包括细胞,例如哺乳动物细胞(包括人细胞)或其他类型的细胞。在某些情况下,样本可以含有病毒。另外,在一些情况下,样品可以是例如来自活组织检查、人工生长或培养等的组织样品。

[0033] 在某些实施方案中,可膨胀材料是能够膨胀的材料,例如,当暴露于水或另一种合适的液体时。例如,该材料可以表现出至少1.1、至少1.2至少1.3、至少1.5、至少2、至少3、至少4、至少5、至少7、至少10或至少15等的尺寸相对变化,和/或小于15、小于10、小于7、小于5、小于4、小于3、小于2、小于1.5、小于1.3或小于1.2的尺寸相对变化(即,2的尺寸相对变化

意味着样品在线性维度上加倍),或者与这些相反(即,2的尺寸相对变化意味着样品在线性维度上减半)。

[0034] 在一些实施方案中,可膨胀材料可以是在膨胀过程中不会显著变形的材料(例如,可膨胀材料可以在所有3个维度上基本上均匀地或各向同性地膨胀),尽管在一些情况下,可膨胀材料可表现出一些变形或非各向同性膨胀。例如,可膨胀材料可以相对于正交尺寸在一个维度上相对于较短的线性膨胀在线性维度上膨胀小于150%、小于130%、小于125%、小于120%、小于115%、小于110%或小于105%。

[0035] 在某些情况下,可膨胀材料是聚合物。合适的聚合物的非限制性实例包括聚电解质和琼脂糖。在某些情况下,聚合物是凝胶或水凝胶。各种聚合物能够用于各种实施方案中,包括但不限于丙烯酸、丙烯酰胺、乙二醇二丙烯酸酯、乙二醇二甲基丙烯酸酯、聚(乙二醇二甲基丙烯酸酯)、聚(N-异丙基丙烯酰胺)、甲基纤维素、(环氧乙烷)-(环氧丙烷)-(环氧乙烷)三元共聚物、海藻酸钠、聚(乙烯醇)、藻酸盐、壳聚糖、阿拉伯树胶、明胶、琼脂糖等。在一些情况下,可以选择相对光学透明的聚合物。在一些情况下,可膨胀材料可由单体或低聚物形成,例如,包含一种或多种取代或未取代的甲基丙烯酸酯、丙烯酸酯、丙烯酰胺、甲基丙烯酰胺、乙烯醇、乙烯胺、烯丙胺、烯丙醇,包括其二乙烯基交联剂(例如N,N-亚烷基双丙烯酰胺如N,N-亚甲基双丙烯酰胺)等。在一些情况下,可存在聚合引发剂和/或交联剂。例如,前体可包括一种或多种交联剂,其可用于在其形成时交联聚合的可膨胀材料,例如在聚合过程中。

[0036] 在一些情况下,可以通过将材料暴露于水、包含水的溶液或其他合适的介质(例如,液体介质)来促进可膨胀材料的膨胀。不希望受任何理论束缚,据信流入材料的水可促进可膨胀材料的膨胀。在某些实施方案中,溶液相对于可膨胀材料可以是低渗的,这可以促进水输送到可膨胀材料中,例如,由于张力的差异。在其他实施方案中可以使用其他使聚合物膨胀的方式,例如通过改变pH、外部电/磁场的变化、温度的变化、对光的响应等。

[0037] 在某些情况下,在膨胀已发生后,可以限制材料的膨胀。例如,可以稳定材料,使得其他缓冲液的交换不会引起尺寸变化。这种稳定化可以例如通过可膨胀材料的化学改性或将可膨胀材料包埋到不可膨胀的第二材料中而发生。

[0038] 作为非限制性实例,可以将样品暴露于一种或多种单体或其他前体,其能够反应形成可膨胀材料。在一些情况下,前体可以在反应之前渗透、扩散或以其他方式运输通过样品。例如,在一些情况下,前体可以存在于液体(例如,水、盐水或其他水性介质)中,其可以以某种方式渗透通过样品。

[0039] 在一些情况下,样品可以包埋到相对大的聚合物或凝胶中,然后其能够在一些情况下进行切面(section)或切片(slice)以产生用于分析的较小部分,例如,使用本领域普通技术人员通常可获得的各种切片技术。例如,组织或器官可以固定在合适的聚合物或凝胶中。

[0040] 在某些实施方案中,凝胶的膨胀可以解开物理重叠的分子。在非限制性实例中,两个RNA分子可在样品中物理重叠。通过将RNA分子在一个位置上锚定,凝胶的膨胀可以分开这些分子,使得它们不再物理重叠,只要每个RNA的附着点在样品内充分分开即可。在一些实施方案中,例如,如果RNA分子沿其长度在多个位置锚定到凝胶上,则凝胶的膨胀可能不会物理地分开这两个分子。相反,凝胶的膨胀可能在附着点之间拉伸这些分子,使得重

叠的这些分子的两个区域在膨胀后仍然重叠。在这种情况下,这些分子可能不能在物理上(例如通过光学技术)区分。在一些情况下,确定性的(例如,靶向RNA的特定位置)相对于随机的(例如,靶向RNA上的大量不同位置,其中随机选择特定靶标位置)锚定方法可能在RNA将在凝胶的多个位置上锚定的可能性上不同,因此在重叠时不能物理分离。RNA中3'或5'末端或确定位置的特异性靶向可能在某些情况下导致确定性锚定,其不会延伸RNA。此外,应该理解的是,尽管在该实施方案中使用了RNA,但是这些考虑也适用于样品中的任何其他生物分子,例如,DNA或本文描述的其他靶标。

[0041] 在一个方面,可以对膨胀的样品进行成像或研究以测定例如在包含在可膨胀材料中的细胞、组织或其他样品中的核酸或其他所需靶标。用于测定核酸或其他所需靶标的技术包括但不限于MERFISH、smFISH或诸如Zhuang等人于2017年1月27日提交的题为“Systems and Methods for Determining Nucleic Acids”的美国专利申请序列号15/329,683(其于2017年8月3日公布为美国专利申请公开号2017/0220733)或Zhuang等人于2017年1月27日提交的题为“Probe Library Construction”的美国专利申请序列号15/329,651(其于2017年7月27日公布为美国专利申请公开号2017/0212986)中公开的技术,这些文献各自通过引用整体并入本文。另外,在一些情况下,所需的靶标可以固定在可膨胀材料(例如聚合物或凝胶)内,而其他组分被“清除”,例如通过降解和/或物理去除,例如,如在Zhuang等人于2016年11月8日提交的题为“Matrix Imprinting and Clearing”的美国临时专利申请序列号62/419,033(通过引用整体并入本文)中讨论的。

[0042] 如果需要测定核酸,则核酸可以是例如细胞(或其他样品)中存在的DNA、RNA或其他核酸。核酸可以是细胞内源的,或添加到细胞中。例如,核酸可以是病毒的核酸,或人工产生的核酸。在一些情况下,待测定的核酸可以由细胞表达。在一些实施方案中,核酸是RNA。RNA可以是编码和/或非编码RNA。可研究的细胞内RNA的非限制性实例包括mRNA、siRNA、rRNA、miRNA、tRNA、lncRNA、snoRNA、snRNA、exRNA、piRNA等。

[0043] 在一些情况下,可研究细胞内的大部分核酸。例如,在一些情况下,可以测定细胞中存在的足够的RNA以产生细胞的部分或完整转录组。在一些情况下,测定细胞内的至少4种mRNA,并且在一些情况下,可测定细胞内的至少3、至少4、至少7、至少8、至少12、至少14、至少15、至少16、至少22、至少30、至少31、至少32、至少50、至少63、至少64、至少72、至少75、至少100、至少127、至少128、至少140、至少255、至少256、至少500、至少1,000、至少1,500、至少2,000、至少2,500、至少3,000、至少4,000、至少5,000、至少7,500、至少10,000、至少12,000、至少15,000、至少20,000、至少25,000、至少30,000、至少40,000、至少50,000、至少75,000或至少100,000种类型的mRNA。

[0044] 在一些情况下,可以测定细胞的转录组。应当理解,转录组通常包括细胞中产生的所有RNA分子,而不仅仅是mRNA。因此,例如,转录组还可以包括rRNA、tRNA、siRNA等。在一些实施方案中,可以测定细胞的至少5%、至少10%、至少15%、至少20%、至少25%、至少30%、至少40%、至少50%、至少60%、至少70%、至少80%、至少90%或100%的转录组。

[0045] 细胞或其他样品中的一种或多种核酸的测定可以是定性的和/或定量的。另外,该测定也可以是空间的,例如,可以在二维或三维中确定细胞或其他样品中核酸的位置。在一些实施方案中,可以测定细胞(或其他样品)内核酸的位置、数量和/或浓度。

[0046] 在一些情况下,可以测定细胞的大部分基因组。测定的基因组区段可以是连续的

或散布在基因组上。例如,在一些情况下,测定细胞中至少4个基因组区段,并且在一些情况下,可以测定细胞中至少3、至少4、至少7、至少8、至少12、至少14、至少15、至少16、至少22、至少30、至少31、至少32、至少50、至少63、至少64、至少72、至少75、至少100、至少127、至少128、至少140、至少255、至少256、至少500、至少1,000、至少1,500、至少2,000、至少2,500、至少3,000、至少4,000、至少5,000、至少7,500、至少10,000、至少12,000、至少15,000、至少20,000、至少25,000、至少30,000、至少40,000、至少50,000、至少75,000或至少100,000个基因组区段。

[0047] 在一些情况下,可以测定细胞的整个基因组。应当理解,基因组一般涵盖细胞中产生的所有DNA分子,而不仅仅是染色体DNA。因此,例如,在某些情况下,基因组还可以包括线粒体DNA、叶绿体DNA、质粒DNA等。在一些实施方案中,可以测定细胞的至少约5%、至少约10%、至少约15%、至少约20%、至少约25%、至少约30%、至少约40%、至少约50%、至少约60%、至少约70%、至少约80%、至少约90%或100%的基因组。

[0048] 然而,如所讨论的,应当理解,在本发明的其他实施方案中,可以测定或固定例如除了核酸以外和/或代替核酸的其他靶标。例如,在本发明的一些实施方案中,待测定或固定的靶标可包括蛋白质(例如,抗体、酶、结构蛋白)、脂质、碳水化合物、病毒等。在一个实施方案中,细胞组分例如蛋白质能够通过与其结合(例如抗体/免疫球蛋白(一抗、二抗、纳米抗体、抗体片段、IgG、IgM、IgA、IgD、IgE等))来进行检测,所述抗体与锚定在聚合物或凝胶上的寡核苷酸探针缀合。然后可以去除这些组分,留下寡核苷酸探针通过其他核酸探针的杂交来检测,其与细胞核酸检测相似或相同。在另一个实施方案中,即使从凝胶或聚合物中去除原始组分,也可以在同一样品中同时检测多种不同的细胞种类。例如,如上所述,可以通过核酸探针的杂交检测RNA分子,同时通过抗体-寡核苷酸缀合物检测蛋白质。

[0049] 如上所述,样品可以部分或完全固定或包埋聚合物或凝胶中。例如,如上所述,样品可以包埋可膨胀材料中。在一组实施方案中,可以在聚合过程中使用锚定探针。锚定探针可包括例如在聚合过程中和/或之后能够与可膨胀材料聚合的锚定部分和能够例如化学和/或物理地固定靶标的靶向部分。例如,在聚丙烯酰胺的情况下,锚定探针可以包括能够聚合并掺入聚合物中的亚磷酰胺部分。作为另一个实例,锚定探针可以含有与靶标(其是核酸,例如RNA(例如mRNA)或DNA)互补的核酸序列作为靶向部分。靶向部分可以是靶标特异性的,和/或可以随机地与样品内的不同靶标相关联(例如,由于非特异性结合)。其他部分也可以存在于锚定探针内。

[0050] 例如,为了与靶核酸缔合,锚定探针可以包含与靶核酸的至少一部分基本上互补的核酸序列。例如,核酸可以与该核酸的至少5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个或更多个核苷酸互补。在一些情况下,互补性可能是精确的(Watson-Crick互补性),或者可能存在1、2个或更多个错配。

[0051] 因此,锚定探针可以在一些实施方案中包含能够与核酸分子相互作用并与其结合的部分,和/或需要固定的其他分子,例如蛋白质或脂质、其他所需靶标等。固定的可以是共价或非共价。例如,为了固定靶核酸,锚定探针可以包含含有亚磷酰胺部分(例如,在5'末端、3'末端、内部碱基处等)的核酸,以及能够识别靶核酸的部分。

[0052] 在一些情况下,例如在转录组分析的情况下,锚定探针能够配置成固定mRNA。例如,在一组实施方案中,锚定探针可例如顺序地含有用于结合mRNA的多聚腺苷酸尾的多个

胸腺嘧啶核苷酸。因此,例如,锚定探针能够具有锚定探针内的至少5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个或更多个连续的胸腺嘧啶核苷酸(例如,多聚dT部分)。在一些情况下,至少一些胸腺嘧啶核苷酸可以是“锁”胸腺嘧啶核苷酸。这些可包含这些胸腺嘧啶核苷酸的至少20%、至少30%、至少40%、至少50%、至少60%、至少70%或至少80%。在某些实施方案中,锁核苷酸和非锁核苷酸可以交替。这种锁胸腺嘧啶核苷酸可用于例如稳定mRNA的多聚腺苷酸尾部与锚定探针的杂交。

[0053] 在另一组实施方案中,锚定探针可包含与mRNA(或另一种靶核酸)基本上互补的序列,如上所述。该序列可以与靶核酸的全部或仅一部分(例如末端部分(例如,朝向5'末端或3'末端),或末端部分之间的中间部分)基本上互补。例如,可以使用具有与DNA或RNA靶基本上互补的部分的锚定探针来固定核酸。锚定探针和核酸之间可以有例如5个或更多个、6个或更多个、7个或更多个、8个或更多个、9个或更多个、10个或更多个、12个或更多个、13个或更多个、14个或更多个、15个或更多个、20个或更多个、25个或更多个、30个或更多个、35个或更多个、40个或更多个、45个或更多个或50个或更多个互补核苷酸。

[0054] 在另一组实施方案中,可以通过上述方法将多个锚定探针靶向至感兴趣的每个RNA。例如,可以将2个或更多个、3个或更多个、4个或更多个、5个或更多个、10个或更多个、15个或更多个、或20个或更多个不同的核酸锚定探针靶向至给定的RNA。这些探针可以靶向该RNA的小区域,使得如果多个探针结合并将RNA锚定到凝胶上,则大部分RNA在膨胀期间保持未拉伸。如果一次靶向多种RNA种类,则应用于样品的独特锚定探针的数量可以是例如5个或更多个、10个或更多个、15个或更多个、20个或更多个、100个或更多个、200个或更多个、1,000个或更多个或10,000个或更多个。

[0055] 其他方法可用于锚定核酸或其中需要固定的其他分子。在一组实施方案中,可以通过共价键合来固定核酸,例如DNA或RNA。例如,在一组实施方案中,可以使用烷基化试剂,其与RNA或DNA共价结合并含有能够在聚合时掺入聚电解质中的第二化学部分。在另一组实施方案中,RNA分子中的末端核糖可以使用高碘酸钠(或另一种氧化剂)氧化以产生醛,其可以与丙烯酰胺或其他聚合物或凝胶交联。在其他实施方案中,可以使用能够修饰碱基的化学试剂,例如醛,例如多聚甲醛或戊二醛,烷化剂或含琥珀酰亚胺基的基团;修饰末端磷酸盐的化学试剂,如碳酰亚胺,如EDC(1-乙基-3-(3-二甲基氨基丙基)碳二亚胺);修饰内部糖类的化学试剂,例如对-马来酰亚胺基-苯基异氰酸酯;或修饰末端糖的化学试剂,如高碘酸钠。在一些情况下,这些化学试剂可以带有第二化学部分,其随后能够直接交联至凝胶或聚合物,和/或可以用能够直接交联至凝胶或聚合物的化合物进一步修饰。

[0056] 在另一组实施方案中,核酸可以在聚合物或凝胶内物理缠结,例如由于它们的长度,并因此不能从它们在凝胶内的原始位置扩散。

[0057] 在其他实施方案中,类似的锚定探针可用于将其他组分固定到聚合物或凝胶上。例如,在一组实施方案中,可以修饰能够特异性结合合适靶标(例如,另一种蛋白质、脂质、碳水化合物、病毒等)的抗体,以包括能够变成掺入聚合物或凝胶中的亚磷酰胺部分。

[0058] 另外,应当理解,在各种实施方案中,可以以任何合适的顺序进行样品在可膨胀材料内的包埋和核酸(或其他所需靶标)的固定。例如,可以在包埋样品之前、期间或之后进行固定。在一些情况下,在凝胶或聚合物形成之前或期间,可以对靶标进行化学修饰或反应以与凝胶或聚合物交联。

[0059] 如上所述,在本发明的一些实施方案中,锚定探针在单个点处将靶标(例如核酸)固定至可膨胀材料。在一些情况下,固定至可膨胀材料的至少20%、至少30%、至少40%、至少50%、至少60%、至少70%、至少80%、至少90%或至少95%(在数量上)的靶标固定在单个点上。通过仅使用单个点,可以相对于可膨胀材料固定相对大量的靶标,而不必影响靶标的功能。例如,多个附着点可能潜在地破坏靶标的结构,例如,由于在可膨胀材料的膨胀期间靶标的“拉伸”(参见例如图4A-4C),或者由于锚定导致的靶标的替换或扭曲(例如,如果锚定发生在用于结合或结构稳定性的位点处或附近,等)。然而,应该理解的是,在一些实施方案中,靶标可以在多于一个点处附着于可膨胀材料。

[0060] 可以使用多种技术在单个点处附着靶标例如核酸(例如RNA或DNA)。例如,可以使用与核酸的特定区域互补的锚定探针(例如,基于序列互补性)。在一些情况下,锚定探针可以个别地特异于每种核酸(例如,具有与该核酸序列互补的独特核酸序列)。因此,能够将诸如DNA或RNA的核酸固定至可膨胀材料,例如,在由锚定探针内的核酸序列控制的单个点或区域处。然而,在其他实施方案中,锚定探针可以包括与核酸的共同特征(例如mRNA的多聚腺苷酸尾)互补的核酸,使得相同的锚定探针能够固定几种不同的核酸(例如,样品中的不同mRNA序列)。

[0061] 另外,在一些实施方案中,可以使用多于一种类型的锚定探针,例如,靶向靶标的相同或不同部分。例如,在一些实施方案中,可以将一个或多个锚定探针应用于样品,使得即使不是每个靶标都被每个锚定探针锚定,也将相当大百分比的靶标固定至可膨胀材料。在一些情况下,例如,至少20%、至少30%、至少40%、至少50%、至少60%、至少70%、至少80%或至少90%的靶标例如通过一个或多个锚定探针固定至可膨胀材料。

[0062] 除了互补核酸序列之外,可以使用其他方法来锚定核酸。例如,在一组实施方案中,RNA分子的3' OH末端可以例如化学地靶向以用于固定至可膨胀材料。例如,3' OH末端可以通过用诸如高碘酸盐(例如高碘酸钠或高碘酸钾)的试剂氧化而转化成醛。然后醛能够与肼、胺或其他部分反应,这些部分能够使锚定探针能够掺入可膨胀材料中。例如,肼或胺能够与化学部分(例如丙烯酰胺、甲基丙烯酰胺或能够掺入可膨胀材料如聚丙烯酰胺中的其它化学物质)连接。上面已经描述了附着技术的其它实例。因此,在一些实施方案中,RNA分子可仅在它们的3' OH末端处锚定。

[0063] 作为另一个非限制性实例,3' OH可以是RNA的酶促延伸可以添加锚定探针的位点。例如,多聚腺苷酸聚合酶可用于将修饰的A核苷酸添加到RNA分子的3' 末端。如上所述,这些A核苷酸可用作多聚T探针的结合位点。作为另一个实例,可以化学修饰A核苷酸,使得它们含有可用于锚定RNA的合适化学部分。例如,含有点击化学试剂叠氮化物的腺苷核苷酸能够商业购买。然后,这可以与通过NHS酯与甲基丙烯酸酯基团连接的DBCO基团(二苯并环辛炔)反应,然后能够将其掺入可膨胀材料中,从而通过3' OH末端将RNA固定至可膨胀材料。

[0064] 作为又一个实例,在一些实施方案中,RNA分子的5' 末端可以例如化学地靶向用于固定至可膨胀材料。例如,碳二亚胺基团能够用于与未加帽的RNA分子上存在的5' 磷酸反应。可以使用脱帽酶如Dcp2(mRNA-脱帽酶2)使加帽的RNA分子脱帽。例如,诸如EDC(1-乙基-3-(3-二甲基-氨基丙基)碳二亚胺)的交联剂可以与未加帽的RNA反应以特异性标记RNA的5' 末端。然后将这些分子附着到锚定部分如丙烯酰胺或甲基丙烯酸酯上,例如通过与含有伯胺的甲基丙烯酸酯分子反应。然后能够将这些部分掺入可膨胀材料中,从而通过5' 末

端将RNA固定至可膨胀材料。

[0065] 在各种实施方案中,其他分子(例如蛋白质)也可锚定至可膨胀基质。例如,蛋白质可以用含有锚定部分的试剂标记。多种方法能够用于蛋白质的化学修饰,包括通过醛交联至胺,通过二硫键或马来酰亚胺与半胱氨酸交联等。在一些实施方案中蛋白质也可以通过与以上述方式标记的抗体的相互作用锚定至可膨胀的凝胶。

[0066] 在将核酸或其他合适的分子固定到聚合物或凝胶上之后,可以“清除”样品中的其他组分。这种清除可以包括非所需的靶标的组分的去除和/或非所需的靶标的组分的降解(例如,到较小的组分、非荧光的组分等)。在一些情况下,可以清除样品中至少50%、至少60%、至少70%、至少80%或至少90%的非所需的组分。在某些实施方案中也能够进行多个清除步骤,例如,以去除各种非所需的组分。如所讨论的,据信去除这些组分可降低分析期间的背景(例如,通过降低背景和/或脱靶结合),而所需组分(例如核酸)能被固定并因此不被清除。

[0067] 例如,可以使用酶、变性剂、螯合剂、化学试剂等从样品中清除蛋白质,其可以将蛋白质分解成更小的组分和/或氨基酸。这些较小的组分可以更容易地被物理去除,和/或可以足够小或惰性,使得它们不会显著影响背景。类似地,可以使用表面活性剂等从样品中清除脂质。在某些情况下,使用这些中的一种或多种,例如同时或顺序使用。合适的酶的非限制性实例包括蛋白酶,例如蛋白酶K、胰酶或肽酶,或消化酶,例如胰蛋白酶、胃蛋白酶或胰凝乳蛋白酶。合适的变性剂的非限制性实例包括盐酸胍、丙酮、乙酸、尿素或高氯酸锂。能够使蛋白质变性的化学试剂的非限制性实例包括溶剂例如苯酚、氯仿、异氰酸胍、尿素、甲酰胺等。表面活性剂的非限制性实例包括Triton X-100(聚乙二醇对-(1,1,3,3-四甲基丁基)-苯基醚)、SDS(十二烷基硫酸钠)、Igepal CA-630或泊洛沙姆。螯合剂的非限制性实例包括乙二胺四乙酸(EDTA)、柠檬酸盐或聚天冬氨酸。在一些实施方案中,可将诸如这些的化合物施加于样品以清除蛋白质、脂质和/或其他组分。例如,可以将缓冲溶液(例如,含有Tris或三(羟甲基)氨基甲烷)施加于样品,然后去除。

[0068] 可用于去除DNA的DNA酶的非限制性实例包括DNA酶I、dsDNA酶、多种限制酶等。清除RNA的技术的非限制性实例包括RNA酶,例如RNA酶A、RNA酶T或RNA酶H或化学试剂,例如通过碱水解(例如,通过将pH增加至大于10)。去除糖或细胞外基质的系统的非限制性实例包括酶,例如几丁质酶、肝素酶或其他糖基化酶。去除脂质的系统的非限制性实例包括酶如脂肪酶,化学试剂如醇(如甲醇或乙醇),或洗涤剂如Triton X-100或十二烷基硫酸钠。其中许多是商业上容易获得的。以这种方式,可以去除样品的背景,这可以促进核酸探针或其他所需靶标的分析,例如使用荧光显微术或本文所讨论的其他技术。如上所述,在各种实施方案中,可以固定各种靶标(例如,核酸、某些蛋白质、脂质、病毒等),同时可以使用合适的试剂或酶清除其他非靶标。作为非限制性实例,如果固定蛋白质(例如抗体),则可以使用RNA酶、DNA酶、去除脂质、糖等的系统。

[0069] 在一些情况下,所需的靶标是核酸。在一组实施方案中,作为说明性的非限制性实例,可以通过同时和/或顺序地将样品暴露于一种或多种类型的核酸探针来研究样品。例如,在一组实施方案中,核酸探针可以包括smFISH或MERFISH探针,例如国际专利申请公开号W0 2016/018960或W0 2016/018963中讨论的那些,这些文献各自通过引用整体并入本文。然而,应当理解,以下内容仅作为示例,并且在其他实施方案中,所需的靶标可以是例如

蛋白质、脂质、病毒等。

[0070] 核酸探针可包含核酸(或能与核酸杂交的实体,例如,特异性地杂交),例如DNA、RNA、LNA(锁核酸)、PNA(肽核酸)或其组合。在一些情况下,核酸探针内还可以存在另外的组分,例如,如下文所讨论。可以使用任何合适的方法将核酸探针引入细胞或其他样品中。

[0071] 例如,在一些实施方案中,在引入核酸探针之前固定细胞或其他样品,例如,以保留核酸在样品内的位置。用于固定细胞和组织的技术是本领域普通技术人员已知的。作为非限制性实例,可以使用化学品如甲醛、多聚甲醛、戊二醛、乙醇、甲醇、丙酮、乙酸等固定细胞。在一个实施方案中,可以使用Hepes-谷氨酸缓冲液介导的有机溶剂(HOPE)固定细胞。在一些实施方案中,在形成之前和/或在形成可膨胀材料之后,可以固定细胞。在一些情况下,细胞在可膨胀材料膨胀之前固定。

[0072] 可以使用任何合适的方法将核酸探针引入细胞(或其他样品)中。在一些情况下,细胞可以被充分透化,使得核酸探针可以通过使含有核酸探针的流体在细胞周围流动而引入细胞中。在一些情况下,细胞可以充分透化作为固定过程的一部分;在其他实施方案中,细胞可以通过暴露于某些化学物质如乙醇、甲醇、Triton X-100等而透化。另外,在一些实施方案中,诸如电穿孔或显微注射的技术可用于将核酸探针引入细胞或其他样品中。

[0073] 本发明的某些方面通常涉及被引入细胞(或其他样品)的核酸探针。取决于应用,探针可以包含可以与核酸杂交(通常通过Watson-Crick碱基配对)的多种实体中的任何一种,例如DNA、RNA、LNA、PNA等。核酸探针通常含有能够与靶核酸的至少一部分结合的靶序列,在某些情况下是特异性地结合。当引入细胞或其他样品时,核酸探针可能能够结合特定的靶核酸(例如,mRNA或本文所述的其他核酸)。在一些情况下,核酸探针可以使用信号传导实体(例如,如下文所讨论)和/或通过使用能够结合核酸探针(即,一级核酸探针)的二级核酸探针来测定。下面详细讨论这种核酸探针的测定。

[0074] 在一些情况下,可以例如同时将多于一种类型的(一级)核酸探针施加于样品。例如,可以存在至少2、至少5、至少10、至少25、至少50、至少75、至少100、至少300、至少1,000、至少3,000、至少10,000、至少30,000、至少50,000、至少100,000、至少250,000、至少500,000或至少1,000,000个可区分的核酸探针,其例如同时或顺序施加于样品。

[0075] 靶序列可以位于核酸探针(或一级核酸探针或编码核酸探针)内的任何位置。靶序列可含有与靶核酸的一部分基本上互补的区域。在一些情况下,所述部分可以为至少50%、至少60%、至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或100%互补。在一些情况下,靶序列的长度可以是至少5、至少10、至少15、至少20、至少25、至少30、至少35、至少40、至少50、至少60、至少65、至少75、至少100、至少125、至少150、至少175、至少200、至少250、至少300、至少350、至少400或至少450个核苷酸。在某些情况下,靶序列的长度可以不超过500、不超过450、不超过400、不超过350、不超过300、不超过250、不超过200、不超过175、不超过150、不超过125、不超过100、不超过75、不超过60、不超过65、不超过60、不超过55、不超过50、不超过45、不超过40、不超过35、不超过30、不超过20或不超过10个核苷酸。任何这些的组合也是可能的,例如,靶序列可具有10至30个核苷酸、20至40个核苷酸、5至50个核苷酸、10至200个核苷酸或25至35个核苷酸、10至300个核苷酸等的长度。通常,基于Watson-Crick核苷酸碱基配对确定互补性。

[0076] (一级)核酸探针的靶序列可以参考怀疑存在于细胞或其他样品中的靶核酸来确定。例如,可以使用蛋白质序列通过测定表达形成蛋白质的核酸来确定蛋白质的靶核酸。在一些情况下,仅使用编码蛋白质的核酸的一部分,例如具有如上所述的长度。另外,在一些情况下,可以使用可以用于识别特定靶标的多于一个靶序列。例如,可以顺序和/或同时使用多个探针,其能够与同一靶标的不同区域结合或杂交。杂交通常是指退火过程,通过该过程,互补的单链核酸通过Watson-Crick核苷酸碱基配对(例如,氢键合、鸟嘌呤-胞嘧啶和腺嘌呤-胸腺嘧啶)缔合以形成双链核酸。

[0077] 在一些实施方案中,核酸探针,例如一级核酸探针,还可以包含一个或多个“读出”序列。但是,应该理解,读出序列不是在所有情况下都是必需的。在一些实施方案中,核酸探针可包含1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16或更多个,20或更多个,32或更多个,40或更多个,50或更多个,64或更多个,75或更多个,100或更多个,128或更多个读出序列。读出序列可位于核酸探针内的任何位置。如果存在多于一个读出序列,则读出序列可以彼此相邻定位,和/或与其他序列是散布的。

[0078] 如果存在,读出序列可以具有任何长度。如果使用多于一个读出序列,则读出序列可以独立地具有相同或不同的长度。例如,读出序列的长度可以是至少5、至少10、至少15、至少20、至少25、至少30、至少35、至少40、至少50、至少60、至少65、至少75、至少100、至少125、至少150、至少175、至少200、至少250、至少300、至少350、至少400或至少450个核苷酸。在某些情况下,读出顺序的长度可以不超过500、不超过450、不超过400、不超过350、不超过300、不超过250、不超过200、不超过175、不超过150、不超过125、不超过100、不超过75、不超过60、不超过65、不超过60、不超过55、不超过50、不超过45、不超过40、不超过35、不超过30、不超过20或不超过10个核苷酸。任何这些的组合也是可能的,例如,读出序列可具有10至30个核苷酸、20至40个核苷酸、5至50个核苷酸、10至200个核苷酸、或25至35个核苷酸、10至300个核苷酸等的长度。

[0079] 在一些实施方案中,读出序列可以是任意的或随机的。在某些情况下,选择读出序列以减少或最小化与细胞或其他样品的其他组分的同源性,例如,使得读出序列本身不与怀疑在细胞或其他样品内的其他核酸结合或杂交。在一些情况下,同源性可小于10%、小于8%、小于7%、小于6%、小于5%、小于4%、小于3%、小于2%或小于1%。在某些情况下,可能存在少于20个碱基对、少于18个碱基对、少于15个碱基对、少于14个碱基对、少于13个碱基对、少于12个碱基对、少于11个碱基对或少于10个碱基对的同源性。在某些情况下,碱基对是连续的(sequential)。

[0080] 在一组实施方案中,核酸探针群可含有一定数量的读出序列,在某些情况下,其可少于核酸探针的靶标数。本领域普通技术人员将意识到,如果存在一个信号传导实体和n个读出序列,则通常可以独特地鉴定 $2^n - 1$ 个不同的核酸靶标。但是,并非所有可能的组合都需要使用。例如,核酸探针群可靶向12种不同的核酸序列,但含有不超过8种读出序列。作为另一个例子,核酸群可以靶向140种不同的核酸种类,但含有不超过16种读出序列。可以通过使用每个探针内的读出序列的不同组合来分别鉴定不同的核酸序列靶标。例如,每个探针可含有1、2、3、4、5、6、7、8、9、10、11、12、13、14、15个等或更多个读出序列。在一些情况下,核酸探针群可各自含有相同数量的读出序列,但在其他情况下,在各种探针上可存在不同数量的读出序列。

[0081] 作为非限制性实例,第一核酸探针可含有第一靶序列、第一读出序列和第二读出序列,而不同的第二核酸探针可含有第二靶序列、相同的第一读出序列和代替第二读出序列的第三读出序列。由此可以通过确定存在的或与给定探针或位置相关的各种读出序列来区分这样的探针,如本文所讨论的。

[0082] 另外,在某些实施方案中,核酸探针(及其在编码探针上的相应的互补位点)可以仅使用4种碱基中的仅2种或仅3种碱基制备,例如在探针中省略所有“G”或省略所有“C”。缺乏“G”或“C”的序列在某些实施方案中可形成非常少的二级结构,并且能够有助于更均匀、更快速的杂交。

[0083] 在一些实施方案中,核酸探针可含有信号传导实体。应该理解,信号传导实体不是在所有情况下都是必需的;例如,在一些实施方案中,可以使用二级核酸探针测定核酸探针,如下面另外详细讨论的。能够使用的信号传导实体的示例也在下面更详细地讨论。

[0084] 其他组分也可以存在于核酸探针内。例如,在一组实施方案中,可以存在一个或多个引物序列,例如,以允许探针的酶促扩增。本领域普通技术人员将知道适合于诸如扩增(例如,使用PCR或其他合适技术)的应用的引物序列。许多这样的引物序列可商购获得。可存在于一级核酸探针内的序列的其他实例包括但不限于启动子序列、操纵子、鉴定序列、无义序列等。

[0085] 通常,引物是单链或部分双链的核酸(例如DNA),其用作核酸合成的起始点,允许聚合酶如核酸聚合酶延伸引物并复制互补链。引物(例如,被设计为)与靶核酸互补并与靶核酸杂交。在一些实施方案中,引物是合成引物。在一些实施方案中,引物是非天然存在的引物。引物通常具有10至50个核苷酸的长度。例如,引物可具有10至40、10至30、10至20、25至50、15至40、15至30、20至50、20至40或20至30个核苷酸的长度。在一些实施方案中,引物具有18至24个核苷酸的长度。

[0086] 另外,核酸探针的组分可以以任何合适的顺序排列。例如,在一个实施方案中,组分可以在核酸探针中排列为:引物-读出序列-靶向序列-读出序列-反向引物。该结构中的“读出序列”可以各自包含任何数量(包括0)的读出序列,只要探针中存在至少一个读出序列即可。非限制性实例结构包括引物靶向序列-读出序列-反向引物,引物-读出序列-靶向序列-反向引物,靶向序列-引物-靶向序列-读出序列-反向引物,靶向序列-引物-读出序列-靶向序列-反向引物,引物-靶向序列-读出序列-靶向序列-反向引物,靶向序列-引物-读出序列-反向引物,靶向序列-读出序列-引物,读出序列-靶向序列-引物,读出序列-引物靶向序列-反向引物等。此外,反向引物在一些实施方案中是任选的,包括在所有上述实例中。

[0087] 在将核酸探针引入细胞或其他样品后,可以通过测定信号传导实体(如果存在)直接测定核酸探针,和/或可以通过使用一种或多种二级核酸探针测定核酸探针。根据本发明的某些方面。如上所述,在某些情况下,测定可以是空间的,例如,在二维或三维中。另外,在一些情况下,测定可以是定量的,例如,可以测定一级核酸探针(和靶核酸)的量或浓度。另外,取决于应用,二级探针可包含能够与核酸杂交的多种实体中的任何一种,例如DNA、RNA、LNA和/或PNA等。下面更详细地讨论信号传导实体。

[0088] 二级核酸探针可含有能够与一级核酸探针的读出序列结合或杂交的识别序列。在一些情况下,结合是特异性的,或者结合可以使得识别序列优选地仅与存在的读出序列中一个结合或杂交。二级核酸探针还可含有一个或多个信号传导实体。如果使用多于一种二

级核酸探针,则信号传导实体可以相同或不同。

[0089] 识别序列可以具有任何长度,并且多个识别序列可以具有相同或不同的长度。如果使用多于一个识别序列,则识别序列可以独立地具有相同或不同的长度。例如,识别序列的长度可以是至少5、至少10、至少15、至少20、至少25、至少30、至少35、至少40或至少50个核苷酸。在一些情况下,识别序列的长度可以不超过75、不超过60、不超过65、不超过60、不超过55、不超过50、不超过45、不超过40、不超过35、不超过30、不超过20或不超过10个核苷酸。任何这些的组合也是可能的,例如,识别序列可以具有10至30、20至40、或25至35个核苷酸等的长度。在一个实施方案中,识别序列具有与读出序列相同的长度。此外,在一些情况下,识别序列可以与一级核酸探针的读出序列至少50%、至少60%、至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或至少100%互补。

[0090] 如上所述,在一些情况下,二级核酸探针可包含一个或多个信号传导实体。下面更详细地讨论信号传导实体的实例。

[0091] 如所讨论的,在本发明的某些方面,使用含有各种“读出序列”的核酸探针。例如,一级核酸探针群可含有某些“读出序列”,其能结合某些二级核酸探针,并且使用二级核酸探针(例如,其包含信号传导实体)在样品内测定一级核酸探针的位置。如上所述,在一些情况下,可以以各种组合将读出序列群组合以产生不同的核酸探针,例如,使得相对少量的读出序列可用于产生相对大量的不同核酸探针。

[0092] 因此,在一些情况下,一级核酸探针(或其他核酸探针)群可各自含有一定数量的读出序列,其中一些在不同的一级核酸探针之间共有,使得一级核酸探针的总群体可包含一定数量的读出序列。核酸探针群可具有任何合适数量的读出序列。例如,一级核酸探针可以具有1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20个等的读出序列。在一些实施方案中,也可以有多于20个。另外,在一些情况下,核酸探针群总共可具有1个或更多个、2个或更多个、3个或更多个、4个或更多个、5个或更多个、6个或更多个、7个或更多个、8个或更多个、9个或更多个、10个或更多个、11个或更多个、12个或更多个、13个或更多个、14个或更多个、15个或更多个、16个或更多个、20个或更多个、24个或更多个、32个或更多个、40个或更多个、50个或更多个、60个或更多个、64个或更多个、100个或更多个、128个或更多个等的可能存在的读出序列,尽管一些或所有探针可各自包含多于一个读出序列,如本文所讨论的。另外,在一些实施方案中,核酸探针群可以具有不超过100、不超过80、不超过64、不超过60、不超过50、不超过40、不超过32、不超过24、不超过20、不超过16、不超过15、不超过14、不超过13、不超过12、不超过11、不超过10、不超过9、不超过8、不超过7、不超过6、不超过5、不超过4、不超过3或不超过两个存在的读出序列。任何这些的组合也是可能的,例如,核酸探针群可以总共包含10至15个读出序列。

[0093] 作为从相对少量的读出序列组合产生相对大量核酸探针的方法的非限制性实例,在6种不同类型的核酸探针的群体中,每种核酸探针包含一个或多个读出序列,群体内的读出序列的总数可以不大于4。应当理解,尽管在该实例中使用了4个读出序列以便于解释,但是在其他实施方案中,可以实现更大数量的核酸探针,例如根据应用,使用5、8、10、16、32个等或更多个读出序列,或本文所述的任何其他合适数量的读出序列。如果每个一级核酸探针含有两个不同的读出序列,则通过使用4个这样的读出序列(A、B、C和D),可以单独鉴定多

达6个探针。应当注意,在该实例中,核酸探针上的读出序列的排序不是必需的,即“AB”和“BA”可以被视为同义的(尽管在其他实施方案中,读出序列的排序可以是必不可少,“AB”和“BA”可能不一定是同义的)。类似地,如果在一级核酸探针群体中使用5个读出序列(A、B、C、D和E),则可以单独鉴定多达10个探针。例如,本领域普通技术人员将理解,对于在每个探针上具有n个读出序列的群体中的k个读出序列,可以产生多达 $\binom{n}{k}$ 个不同的探针,假设读出序列的排序不是必需的;因为并非所有探针都需要具有相同数量的读出序列,并且不需要在每个实施方案中使用所有读出序列组合,所以在某些实施方案中也可以使用多于或少于这个数量的不同探针。另外,还应该理解,在一些实施方案中,每个探针上的读出序列的数量不必相同。例如,一些探针可以含有2个读出序列,而其他探针可以含有3个读出序列。

[0094] 在一些方面,样品中核酸探针的读出序列和/或结合模式可用于定义错误检测和/或错误校正代码,例如,以减少或防止核酸的错误鉴定或误差。因此,例如,如果指示了结合(例如,如使用信号传导实体确定的),则可以用“1”标识该位置;相反,如果没有指示结合,则可以用“0”标识位置(或者在某些情况下,反之亦然)。然后能够使用例如使用不同核酸探针的多轮结合测定来产生例如用于该空间位置的“代码字”。在一些实施方案中,可以对代码字进行错误检测和/或校正。例如,可以组织代码字,使得如果对于给定的一组读出序列或核酸探针的结合模式没有找到匹配,则可以将匹配识别为错误,并且可选地,可以对序列应用错误校正以确定核酸探针的正确靶标。在一些情况下,代码字可以具有较少的“字母”或位置,即由代码字编码的核酸的总数,例如,其中每个代码字编码不同的核酸。

[0095] 这种错误检测和/或错误校正代码可以采用各种形式。先前已在诸如电信行业的其他环境中开发了各种这样的代码,例如Golay代码或汉明代码。在一组实施方案中,分配核酸探针的读出序列或结合模式,使得并非每种可能的组合都被分配。

[0096] 例如,如果4个读出序列是可能的并且一级核酸探针含有2个读出序列,那么可以鉴定多达6个一级核酸探针;但是,使用的一级核酸探针的数量可能少于6。类似地,对于在每个一级核酸探针上具有n个读出序列的群体中的k个读出序列,可以产生 $\binom{n}{k}$ 个不同的探针,但是使用的一级核酸探针的数量可以是多于或少于 $\binom{n}{k}$ 的任何数量。另外,这些可以被随机分配,或以特定方式分配以增加检测和/或纠正错误的能力。

[0097] 作为另一个例子,如果使用多轮核酸探针,则可以任意选择轮数。如果在每一轮中,每个靶标可以给出两种可能的结果,例如被检测到或未被检测到,则n轮探针可能有多达 $2^n$ 个不同的靶标,但实际使用的核酸靶标的数量可以是小于 $2^n$ 的任何数量。例如,如果在每一轮中,每个靶标可以给出多于两种可能的结果,例如在不同的颜色通道中检测到,则n轮探针可能有超过 $2^n$ (例如 $3^n$ 、 $4^n$ .....)个不同的靶标。在一些情况下,实际使用的核酸靶标的数量可以是小于该数量的任何数量。另外,这些可以被随机分配,或以特定方式分配以增加检测和/或纠正错误的能力。

[0098] 例如,在一组实施方案中,可以在代码空间内分配代码字或核酸探针,使得该分配由汉明距离分开,汉明距离测量给定模式中导致核酸探针被误解为不同的有效核酸探针的不正确“读出”的数量。在某些情况下,汉明距离可以是至少2,至少3,至少4,至少5,至少6等。另外,在一组实施方案中,分配可以形成为汉明代码,例如汉明(7,4)代码,汉明(15,11)

代码,汉明(31,26)代码,汉明(63,57)代码,汉明(127,120)代码等。在另一组实施方案中,分配可以形成SECDED代码,例如SECDED(8,4)代码,SECDED(16,4)代码,SCEDED(16,11)代码,SCEDED(22,16)代码,SCEDED(39,32)代码,SCEDED(72,64)代码等。在另一组实施方案中,分配可以形成扩展的二进制Golay代码、完美的二进制Golay代码或三进制Golay代码。在另一组实施方案中,分配可以表示从上述任何代码中获取的可能值的子集。

[0099] 例如,可以通过仅使用包含固定数量的“1”位(例如4)的二进制字来形成具有相同的SECDED代码的错误校正特性的代码,以对靶标进行编码。在另一组实施方案中,为了解决不对称读出错误,分配可以表示从上述代码中获取的可能值的子集。例如,在某些情况下,其中“1”位的数量对于所有使用的二进制字可以是固定的代码可以在“0”位被测量为“1”或“1”位被测量为“0”处的速率不同时消除具有不同数量的“1”的字的偏倚测量。

[0100] 因此,在一些实施方案中,一旦确定了代码字(例如,如本文所讨论的),就可以将代码字与已知的核酸代码字进行比较。如果发现匹配,则可以鉴定或确定核酸靶标。如果没有找到匹配,则可以鉴定代码字的读出中的错误。在一些情况下,还可以应用错误校正来确定正确的代码字,从而导致核酸靶标的正确身份。在一些情况下,可以选择代码字,使得在假设仅存在一个错误的情况下仅可能有一个可能的正确代码字,因此,核酸靶标的仅一个正确身份是可能的。在某些情况下,这也可以推广到更大的代码字间距或汉明距离;例如,可以选择代码字,使得如果存在两个、三个或四个错误(或者在一些情况下更多),则只有一个可能的正确代码字可用,并且因此只有核酸靶标的一个正确身份是可能的。

[0101] 错误校正代码可以是二进制错误校正代码,或者它可以基于其他编号系统,例如三进制或四进制错误校正代码。例如,在一组实施方案中,可以使用多于一种类型的信号传导实体并将其分配至错误校正代码内的不同数字。因此,作为非限制性示例,第一信号传导实体(或在某些情况下,多于一个信号传导实体)可以被分配为“1”,并且第二信号传导实体(或在某些情况下,可以是多于一个信号传导实体)可以被分配为“2”(“0”表示不存在信号传导实体),并且分配的代码字用于定义三进制错误校正代码。类似地,第三信号传导实体可以另外被分配为“3”以产生四进制错误校正代码等。

[0102] 如上所讨论的,在某些方面,确定信号传导实体,例如,以确定核酸探针和/或产生代码字。在一些情况下,可以使用各种技术,例如在空间上确定样品内的信号传导实体。在一些实施方案中,信号传导实体可以是荧光的,并且用于确定样品内荧光的技术,例如荧光显微术或共聚焦显微术,可以用于在空间上识别细胞内信号传导实体的位置。在某些情况下,样品内的实体的位置可以在两个或甚至三个维度中确定。另外,在某些实施方案中,多于一个信号传导实体可以一次(例如,具有不同颜色或发射的信号传导实体)和/或顺序地确定。

[0103] 另外,在某些实施方案中,可以测定鉴定的核酸靶标的置信水平。例如,可以使用精确匹配的数量与具有一个或多个一位(one-bit)错误的匹配的数量的比率来测定置信水平。在某些情况下,可以仅使用具有大于特定值的置信度比率的匹配。例如,在某些实施方案中,仅当匹配的置信度比率为大于约0.01、大于约0.03、大于约0.05、大于约0.1、大于约0.3、大于约0.5、大于约1、大于约3、大于约5、大于约10、大于约30、大于约50、大于约100、大于约300、大于约500、大于约1000或任何其他合适的值时,才可接受匹配。另外,在某些实施方案中,仅当所鉴定的核酸靶标的置信度比率比内标或假阳性对照大约0.01、约0.03、约

0.05、约0.1、约0.3、约0.5、约1、约3、约5、约10、约30、约50、约100、约300、约500、约1000或任何其它合适的值时,才可接受匹配。

[0104] 在一些实施方案中,可以以相对高的分辨率确定实体的空间位置(并因此,实体可以与之关联的核酸探针)。例如,可以以优于约100微米,优于约30微米,优于约10微米,优于约3微米,优于约1微米,优于约800nm,优于约800微米,优于约600nm,优于约500nm,优于约400nm,优于约300nm,优于约200nm,优于约100nm,优于约90nm,优于约80nm,优于约70nm,优于约70nm,优于约60nm,优于约50nm,优于约40nm,优于约30nm,优于约20nm,或优于约10nm等的空间分辨率确定位置。

[0105] 有多种技术能够光学地确定或成像实体或靶标的空间位置,例如,使用荧光显微术,使用放射性,通过与合适的发色团缀合等。例如,可以在本发明的各种实施方案中使用的各种常规显微术技术包括但不限于落射荧光显微术、全内反射显微术、高倾斜薄照(HILO)显微术、光片显微术、扫描共聚焦显微术、扫描线共聚焦显微术、旋转盘共聚焦显微术或其他可比较的传统显微技术。

[0106] 在一些实施方案中,可以使用用于标记核酸例如DNA或RNA的原位杂交(ISH)技术,例如,其中核酸探针可以与样品中的核酸杂交。这些可以例如以细胞级或单分子级分辨率进行。在一些情况下,ISH探针可以包括RNA、DNA、PNA、LNA、其他合成核苷酸等,和/或这些中任何一种的组合。杂交探针的存在可以例如如下测量:使用放射性标记的核酸探针的放射性,使用例如生物素标记的核酸探针进行的免疫组织化学,使用例如可以结合酶例如辣根过氧化物酶的探针和方法例如酪胺信号放大的酶促发色团或荧光团产生,使用直接用荧光团标记的核酸探针的荧光成像,或二级核酸探针与这些一级探针的杂交,其中通过任何上述方法检测二级探针。

[0107] 在一些情况下,空间位置可以以超分辨率确定,或者以比光的波长或衍射极限更好的分辨率确定(尽管在其他实施方案中,不需要超分辨率)。非限制性实例包括STORM(随机光学重建显微术),STED(受激发射耗尽显微术),NSOM(近场扫描光学显微术),4Pi显微术,SIM(结构照明显微术),SMI(空间调制照明)显微术,RESOLFT(可逆饱和光学线性荧光过渡显微术),GSD(基态消光显微术),SSIM(饱和结构照明显微术),SPDM(光谱精密距离显微术),光活化定位显微术(PALM),荧光光活化定位显微术(FPALM),LIMON(3D光学显微术纳米尺寸显微术),超分辨率光学波动成像(SOFI)等。参见,例如,2010年11月23日授权的Zhuang等人的题为“Sub-Diffraction Limit Image Resolution and Other Imaging Techniques”的美国专利号7,838,302;2013年10月22日授权的Zhuang等人的题为“Sub-diffraction Limit Image Resolution in Three Dimensions”的美国专利号8,564,792;或者2013年6月20日公开的Zhuang等人的题为“High Resolution Dual-Objective Microscopy”的国际专利申请公开号W02013/090360,其各自通过引用整体并入本文。

[0108] 在一个实施方案中,可以通过单高斯模式激光线照射样品。在一些实施方案中,通过使这些激光线穿过通过压电或其他机械装置振动的多模光纤,可以使照明轮廓变平。在一些实施方案中,可以通过使单模高斯光束通过各种折射光束整形器(例如,piShaper或一系列堆叠的Powell透镜)来平坦化照明轮廓。在又一组实施方案中,高斯光束可以通过各种不同的漫射元件,例如磨砂玻璃或工程化漫射器,其在某些情况下可以高速旋转以去除残留的激光散斑。在又一个实施方案中,激光照射可以通过一系列小透镜阵列以产生近似于

平坦照明场的照明重叠图像。

[0109] 在一些实施方案中,可以确定实体的空间位置的质心。例如,可以使用本领域普通技术人员已知的图像分析算法在图像或一系列图像内确定信号传导实体的质心。在一些情况下,可以选择算法以确定样品中的非重叠单个发射器和/或部分重叠的单个发射器。合适技术的非限制性示例包括最大似然算法、最小二乘算法、贝叶斯算法、压缩感测算法等。在某些情况下也可以使用这些技术的组合。

[0110] 另外,在某些情况下可以使信号传导实体失活。例如,在一些实施方案中,可以将含有信号传导实体的第一二级核酸探针应用于可以识别第一读出序列的样品,然后可以在将第二二级核酸探针应用于样品之前使第一二级核酸探针失活。如果使用多个信号传导实体,则可以使用相同或不同的技术来使信号传导实体失活,并且可以例如顺序地或同时地使多个信号传导实体中的一些或全部失活。

[0111] 失活可以通过去除信号传导实体(例如,来自样品,或来自核酸探针等)和/或通过以某种方式化学改变信号传导实体(例如通过光漂白信号传导实体,漂白或化学改变信号传导实体的结构,例如通过还原等)来引起。例如,在一组实施方案中,荧光信号传导实体可以通过化学或光学技术失活,例如氧化、光漂白、化学漂白、严格洗涤或酶消化或通过暴露于酶的反应、使信号传导实体与其他组分(例如,探针)解离、信号传导实体的化学反应(例如,与能够改变信号传导实体的结构的反应物)等。例如,漂白可以通过暴露于氧气、还原剂而发生,或者信号传导实体可以从核酸探针化学裂解并通过流体流动洗掉。

[0112] 在一些实施方案中,各种核酸探针(包括一级和/或二级核酸探针)可包括一种或多种信号传导实体。如果使用多于一种核酸探针,则信号传导实体可各自相同或不同。在某些实施方案中,信号传导实体是能够发光的任何实体。例如,在一个实施方案中,信号传导实体是荧光的。在其他实施方案中,信号传导实体可以是磷光的、放射性的、吸收性的等。在一些情况下,信号传导实体是可以以相对高的分辨率(例如,在比可见光的波长或衍射极限更好的分辨率下)在样品内测定的任何实体。信号传导实体可以是例如染料、小分子、肽或蛋白质等。在某些情况下,信号传导实体可以是单个分子。如果使用多个第二核酸探针,则核酸探针可包含相同或不同的信号传导实体。

[0113] 信号传导实体的非限制性实例包括荧光实体(荧光团)或磷光实体,例如,花青染料(例如,Cy2,Cy3,Cy3B,Cy5,Cy5.5,Cy7等),Alexa Fluor染料,Atto染料,光切换染料,光活化染料,荧光染料,金属纳米粒子,半导体纳米粒子或“量子点”,荧光蛋白如GFP(绿色荧光蛋白),或光活化荧光蛋白,如PAGFP,PSCFP,PSCFP2,Dendra,Dendra2,EosFP,tdEos,mEos2,mEos3,PAmCherry,PAtagRFP,mMaple,mMaple2和mMaple3。其他合适的信号传导实体是本领域普通技术人员已知的。参见,例如,美国专利号7,838,302或美国专利申请序列号61/979,436,其各自通过引用整体并入本文。在某些情况下,可以使用光谱上不同的荧光染料。

[0114] 在一组实施方案中,信号传导实体可以通过可以被裂解以释放信号传导实体的键与寡核苷酸序列连接。在一组实施方案中,荧光团可以通过可裂解键(例如光可裂解键)与寡核苷酸缀合。光可裂解键的非限制性实例包括但不限于1-(2-硝基苄基)乙基,2-硝基苄基,生物素亚磷酰胺,丙烯酰亚磷酰胺,二乙基氨基香豆素,1-(4,5-二甲氧基-2-硝基苄基)乙基,环十二烷基(二甲氧基-2-硝基苄基)乙基,4-氨基甲基-3-硝基苄基,(4-硝基-3-(1-

氯羰基氧基乙基) 苯基) 甲基-S-乙酰硫代酸酯, (4-硝基-3-(1-叔羰基氧基乙基) 苯基) 甲基-3-(2-吡啶基二硫代丙酸) 酯, 3-(4,4'-二甲氧基三苯甲基)-1-(2-硝基苯基)-丙烷-1,3-二醇-[2-氰基乙基]- (N,N-二异丙基)]-亚磷酰胺, 1-[2-硝基-5-(6-三氟乙酰基己酰氨基甲基) 苯基]-乙基-[2-氰基-乙基- (N,N-二异丙基)]-亚磷酰胺, 1-[2-硝基-5-(6-(4,4'-二甲氧基三苯甲基氧基) 丁酰氨基甲基) 苯基]-乙基-[2-氰基乙基- (N,N-二异丙基)]-亚磷酰胺, 1-[2-硝基-5-(6-(N-(4,4'-二甲氧基三苯甲基)))-生物素酰胺基己酰胺基-甲基) 苯基]-乙基-[2-氰基乙基- (N,N-二异丙基)]-亚磷酰胺, 或类似的连接基。在另一组实施方案中, 荧光团可以通过二硫键与寡核苷酸缀合。二硫键可以被各种还原剂裂解, 例如但不限于二硫苏糖醇, 二硫赤藓糖醇,  $\beta$ -巯基乙醇, 硼氢化钠, 硫氧还蛋白, 谷氧还蛋白, 胰蛋白酶原, 胍, 二异丁基氢化铝, 草酸, 甲酸, 抗坏血酸, 亚磷酸, 氯化锡, 谷胱甘肽, 硫代乙醇酸, 2,3-二巯基丙醇, 2-巯基乙胺, 2-氨基乙醇, 三(2-羧乙基) 膦, 双(2-巯基乙基) 砷, N,N'-二甲基-N,N'-双(巯基乙酰基) 胍, 3-巯基丙酸盐, 二甲基甲酰胺, 硫代丙基-琼脂糖, 三正丁基膦, 半胱氨酸, 硫酸铁, 亚硫酸钠, 亚磷酸盐, 次磷酸盐, 硫代磷酸盐等, 和/或其任意的组合。在另一个实施方案中, 荧光团可以通过一种或多种硫代磷酸酯修饰的核苷酸(其中硫修饰取代桥接和/或非桥接氧) 与寡核苷酸缀合。在某些实施方案中, 通过添加化合物(例如但不限于碘乙醇, 在乙醇中混合的碘, 硝酸银或氯化汞), 荧光团可以从寡核苷酸上裂解下来。在又一组实施方案中, 信号传导实体可以通过还原或氧化进行化学失活。例如, 在一个实施方案中, 可以使用硼氢化钠将发色团如Cy5或Cy7还原至稳定的非荧光状态。在另一组实施方案中, 荧光团可以通过偶氮键与寡核苷酸缀合, 偶氮键可以用2-[(2-N-芳基氨基) 苯基偶氮]吡啶裂解。在另一组实施方案中, 荧光团可以通过合适的核酸区段与寡核苷酸缀合, 所述核酸区段可以在适当暴露于DNA酶(例如外切脱氧核糖核酸酶或内切脱氧核糖核酸酶) 时裂解。实例包括但不限于脱氧核糖核酸酶I或脱氧核糖核酸酶II。在一组实施方案中, 裂解可以通过限制性内切核酸酶发生。可能合适的限制性内切核酸酶的非限制性实例包括BamHI, BsrI, NotI, XmaI, PspAI, DpnI, MboI, MnlI, Eco57I, Ksp632I, DraIII, AhaII, SmaI, MluI, HpaI, ApaI, BclI, BstEII, TaqI, EcoRI, SacI, HindII, HaeII, DraII, Tsp509I, Sau3AI, PacI等。已经详细研究了超过3000种限制酶, 其中超过600种可商购获得。在另一组实施方案中, 荧光团可以与生物素缀合, 并且寡核苷酸与抗生物素蛋白或链霉抗生物素蛋白缀合。生物素与抗生物素蛋白或链霉抗生物素蛋白之间的相互作用允许荧光团与寡核苷酸缀合, 同时充分暴露于过量添加的游离生物素可以“竞争胜过”键联, 从而导致裂解发生。另外, 在另一组实施方案中, 可以使用相应的“立足点探针”去除探针, “立足点探针”包含与探针相同的序列, 以及与编码探针同源的额外数量的碱基(例如, 1-20个额外的碱基, 例如, 5个额外的碱基)。这些探针可以通过链置换相互作用去除标记的读出探针。

[0115] 如本文所用, 术语“光”通常是指具有任何合适波长(或等效地, 频率) 的电磁辐射。例如, 在一些实施方案中, 光可以包括光学或视觉范围内的波长(例如, 具有约400nm至约700nm的波长, 即“可见光”), 红外波长(例如, 具有约300微米至700纳米的波长), 紫外波长(例如, 具有约400纳米至约10纳米的波长) 等。在某些情况下, 如下面详细讨论的, 可以使用多于一个实体, 即化学上不同或独特(例如在结构上) 的实体。然而, 在其他情况下, 实体可以在化学上相同或至少基本上在化学上相同。

[0116] 本发明的另一方面涉及一种计算机实现的方法。例如, 可以提供能够自动和/或重

复执行本文描述的任何方法的计算机和/或自动系统。如本文所使用的，“自动”设备是指能够在没有人的指导的情况下操作的设备，即，自动设备可以在任何人已完成采取任何动作以促进功能（例如通过输入指令到计算机中以启动过程）之后的一段时间内执行功能。通常，自动设备可以在此时间点之后执行重复功能。在某些情况下，处理步骤也可以记录在机器可读介质上。

[0117] 例如，在一些情况下，计算机可用于控制样品的成像，例如，使用荧光显微术、STORM或其他超分辨率技术，例如本文所述的那些。在一些情况下，计算机还可以控制操作诸如图像分析中的漂移校正、物理对准、杂交和簇对齐，簇解码（例如，荧光簇解码），错误检测或校正（例如，如本文所讨论的），降噪，从背景特征（例如图像中的噪声或碎片）识别前景特征等。作为示例，计算机可用于控制样品内的信号传导实体的激活和/或激发，和/或信号传导实体的图像的获取。在一组实施方案中，可以使用具有各种波长和/或强度的光来激发样品，并且可以使用计算机将用于激发样品的光的波长序列与获取的包含信号传导实体的样品的图像相关联。例如，计算机可以将具有各种波长和/或强度的光应用于样品，以在每个感兴趣区域中产生不同平均数量的信号传导实体（例如，每个位置一个激活的实体，每个位置两个激活的实体等）。在一些情况下，如上所述，该信息可用于构建图像和/或确定信号传导实体的位置，在一些情况下以高分辨率确定。

[0118] 以下文献通过引用整体并入本文：由Zhuang等人于2007年8月7日提交的题为“Sub-Diffraction Image Resolution and Other Imaging Techniques”的国际专利申请号PCT/US2007/017618，于2008年7月31日公布为WO 2008/091296；Zhuang等人的于2010年8月17日授权的题为“Sub-Diffraction Image Resolution and Other Imaging Techniques”的美国专利号7,776,613；Zhuang等人的于2010年11月23日授权的题为“Sub-Diffraction Image Resolution and Other Imaging Techniques”的美国专利号7,838,302；Zhuang等人于2015年7月29日提交的题为“Probe Library Construction”的国际专利申请号PCT/US2015/042559，于2016年2月4日公布为WO 2016/018963；Zhuang等人于2015年7月29日提交的题为“Systems and Methods for Determining Nucleic Acids”的国际专利申请号PCT/US2015/042556，于2016年2月4日公布为WO 2016/018960；Zhuang等人于2017年1月27日提交的题为“Systems and Methods for Determining Nucleic Acids”的美国专利申请序列号15/329,683，于2017年8月3日公布为美国专利申请公开号2017/0220733；Zhuang等人于2017年1月27日提交的题为“Probe Library Construction”的美国专利申请序列号15/329,651，于2017年7月27日公布为美国专利申请公开号2017/0212986；Zhuang等人于2016年8月31日提交的题为“Sub-Diffraction Image Resolution and Other Imaging Techniques”的美国专利申请序列号15/252,307，于2016年12月22日公布为美国专利申请公开号2016/0370295；Zhuang等人于2016年11月8日提交的题为“Matrix Imprinting and Clearing”的美国临时专利申请序列号62/419,033；Zhuang等人于2017年5月26日提交的题为“Systems and Methods for High-Throughput Image-Based Screening”的美国临时专利申请序列号62/511,920；和Zhuang等人于2017年10月6日提交的题为“Multiplexed Imaging using MERFISH, Expansion Microscopy, and Related Technologies”的美国临时专利申请序列号62/569,127。

[0119] 以下实施例旨在举例说明本发明的某些实施方案，但不是示例本发明的全部范

围。

#### [0120] 实施例1

[0121] 以下一些实施例示出了基于本发明某些实施方案的膨胀MERFISH,其利用扩展显微术显著增加MERFISH可测量的RNA的总密度。在这些实施例中,RNA锚定在聚合物凝胶上,该聚合物凝胶体积膨胀12倍以从物理上降低RNA密度。如此,在由丰富的RNA组成的MERFISH文库(其总密度比先前测量的RNA文库高>10倍)中证明了RNA的准确鉴定。

[0122] 另外,以下实施例中的一些示出了基于某些实施方案将免疫荧光与膨胀MERFISH组合,其利用亚磷酰胺修饰的寡核苷酸缀合的抗体揭示蛋白质定位和表达以及RNA的测量。在这些实施例中,样品在蛋白水解之前用亚磷酰胺修饰的寡核苷酸-缀合的抗体染色,从而提供与可膨胀凝胶的交联,然后进行消化。将连接的寡核苷酸染色以显示抗体的原始位置以及RNA的MERFISH读出。

[0123] 单细胞转录组学的基于原位成像的方法不仅可以确定待测定的个体细胞的表达谱,还可以确定待定位的个体RNA分子的空间位置。这些方法提供了强有力的手段来映射细胞内RNA的空间组织和组织中转录不同的细胞。目前可用的基于图像的单细胞RNA分析方法依赖于多重荧光原位杂交(FISH)或原位测序。特别地,多重错误稳健FISH(MERFISH),一种大规模多重形式的单分子FISH(smFISH),允许在转录组规模上进行RNA成像。作为一种在细胞内对个体RNA分子进行成像的有效方法,smFISH可提供单一细胞中的RNA的精确拷贝数和空间组织。MERFISH通过用寡核苷酸探针组合标记RNA使smFISH测量多重化,所述寡核苷酸探针含有错误稳健条形码并通过连续(sequential)轮的smFISH成像测量这些条形码。使用这种方法,已经证明使用错误检测/校正条形码方案同时成像个体细胞中成百上千至成千上万种RNA种类。参见,例如,题为“Systems and Methods for Determining Nucleic Acids”的美国专利申请公开号2017-0220733,名称为“确定核酸的系统和方法”,和题为“Probe Library Construction”的美国专利申请公开号2017-0212986,其各自通过引用整体并入本文。

[0124] 还可以增加MERFISH在单日时长的测量中对数以万计细胞的测量通量。此外,通过将细胞RNA锚定在聚合物基质上并去除产生荧光背景的其他细胞成分,已开发了提高信号背景比的样品清除方法,这种清除方法允许对组织切片进行高质量的MERFISH测量。参见,例如,题为“Matrix Imprinting and Clearing”的美国专利申请公开号62/419,033;和题为“Systems and methods for high-throughput image-based screening”的美国专利申请公开号62/511,920,各自通过引用整体并入本文。

[0125] 为了准确鉴定RNA分子,MERFISH以及其他多重FISH或基于原位测序的RNA分析方法需要来自个体RNA的非重叠荧光信号。然而,由于衍射极限,当分子彼此足够接近时,它们的荧光信号会重叠,从而限制了能够同时成像的RNA的密度。例如,利用超分辨率成像方法(通过光学手段或通过样品膨胀)能够克服该问题。这些实施例示出了样品膨胀和扩展显微术(ExM),以显著提高分辨相近分子的能力。

[0126] 在该方法中,将所需信号与可膨胀聚丙烯酰胺凝胶缀合,然后通过改变缓冲液的离子强度使凝胶物理膨胀,从而分开以其他方式会产生重叠荧光信号的分子。这些实施例证实了组合MERFISH和扩展显微术的方法。通过亚磷酰胺修饰的多聚dT寡核苷酸将mRNA锚定至可膨胀的聚合物凝胶,并在培养的人骨肉瘤细胞(U-2 OS)中成像高丰度RNA文库,其含

有约130个RNA种类。在没有凝胶膨胀的情况下,这些RNA没有很好地得到分辨,因此以约15-20%的相对低的效率得到检测。相反,在膨胀的样品中,个体RNA分子变得很好地分辨,导致其检测效率的显著提高。与smFISH和批量测序结果的比较表明,膨胀样品中的这些RNA以高准确度和接近100%的效率得到检测。这些实施例还证明了在这些膨胀样品中同时进行MERFISH RNA成像和蛋白质免疫荧光成像的能力。

#### [0127] 实施例2

[0128] RNA密度对MERFISH测量检测效率的影响。为了示出RNA密度对多重smFISH测量的影响,使用先前公布的16位修改的汉明距离4 (MHD4) 二进制码(其允许错误检测和校正)测量高丰度129-RNA文库。该代码包括140个独特的代码字,其中约129个用于编码RNA。约11个用作不对应于任何RNA的空白对照。在129个靶向RNA中,106个丰度范围为每个细胞40-250个拷贝,其余23个跨越每个细胞1-1000个拷贝的丰度范围,以量化不同丰度的性能。该文库中RNA的总丰度是先前使用MHD4代码测量的检测效率为约80-90%的130-RNA文库的14倍。

[0129] 在MERFISH测量中,用两组探针标记RNA。在第一步中,每个细胞RNA与称为“编码探针”的寡核苷酸探针复合组杂交,其包含结合细胞RNA的靶向序列和测定这些RNA的条形码的读出序列。在第二步中,通过一系列smFISH测量检测读出序列和从而检测条形码,所述一系列smFISH测量中的每轮测量使用与一个或多个读出序列互补的一个或多个读出探针。使用基于基质印迹的清除方法在U-2 OS细胞中进行MERFISH测量。参见题为“Matrix Imprinting and Clearing”的美国专利申请公开号62/419,033,其通过引用并入本文。简而言之,将细胞固定、透化、用针对129个RNA种类的编码探针以及靶向聚腺苷酸化(多聚腺苷酸)RNA的亚磷酰胺修饰的多聚dT探针标记。将细胞包埋在聚合物凝胶中,通过蛋白酶K消化和洗涤剂提取来去除细胞蛋白质和脂质,同时通过多聚dT探针将多聚腺苷酸RNA锚定在凝胶上。清除后,进行八轮双色smFISH测量,每轮使用两个读出探针,以读出RNA上的16位条形码,在连续轮的smFISH成像之间将化学裂解用于去除与读出探针连接的荧光团。因为细胞的高度大于单个光学切片的厚度,所以样品用多个z切片成像以确保检测到细胞内>90%的RNA分子。

[0130] 与先前对低丰度RNA文库的MERFISH测量不同(由于与该丰度高14倍的RNA文库相关的高分子密度),在每轮成像中大部分smFISH信号在空间上重叠(图1A、B)。结果,只有一小部分RNA分子是可解码的(图1C)。与大量RNA-seq的比较显示,通过MERFISH检测到的这些RNA的每细胞平均拷贝数与通过RNA-seq测量的RNA丰度相关,其中 $\log_{10}$ 值之间的Pearson相关系数为0.6(图1D)。为了确定检测效率,将MERFISH结果与该文库中12个基因的smFISH测量结果进行比较,还使用基于基质印迹的清除方法进行。该比较显示通过MERFISH测定的每个细胞的拷贝数仅为通过smFISH测定的那些的21% $\pm$ 4%(平均值 $\pm$ SEM)或16%(中值)(图1E)。

[0131] 图1显示了未膨胀的U-2 OS样品中高丰度RNA文库的MERFISH测量。使用如上所述的基于基质印迹的清除方法清除样品。图1A显示了用129个RNA的编码探针染色的未膨胀的U-2 OS样品中MERFISH测量的图像,并在z-扫描的一个焦平面用Cy5标记的读出探针可视化。实线标记细胞的边缘,虚线标记细胞核的DAPI染色区域。图1B显示了对于(图1A)中加框的子区域的所有8轮双色smFISH成像的高通滤波荧光图像。不同的阴影分别代表Cy5通道(绿色)和Alexa 750(红色)通道。图1C显示了基于它们的经测量的二进制条形码的图1A中

所有解码的RNA的定位。插图显示了图1B中所示区域的所有解码的RNA的定位。显示了所有光学切片上的解码RNA。实线标记细胞的边缘,虚线标记细胞核的DAPI染色区域。图1D显示了通过MERFISH测定的129个RNA种类的每个细胞的平均RNA拷贝数相对于通过RNA-seq测定的丰度。 $\log_{10}$ 值之间的Pearson相关系数( $\rho_{10}$ )为0.6。图1E显示了129个RNA中的12个的通过MERFISH测定的每个细胞的平均RNA拷贝数相对于通过smFISH测量的平均RNA拷贝数。对每个数据集进行5微米深度的z扫描,以确保至少90%的RNA分子包括在成像深度中。 $\log_{10}$ 值之间的Pearson相关系数( $\rho_{10}$ )为0.75。由MERFISH测定的拷贝数值与smFISH测定的拷贝数值的平均比率为21% $\pm$ 4%(SEM,n=12个RNA种类),中值为16%。

### [0132] 实施例3

[0133] 使用扩展显微术进行高RNA密度MERFISH测量。据信降低的MERFISH检测效率可能部分归因于重叠的单分子信号。如本实施例中所示,通过超分辨率成像方法(通过光学手段或通过样品膨胀)能够克服该问题。本实施例示出了在单个位置处锚定的分子的样品膨胀和扩展显微术(ExM)以增加分子之间的距离并显著提高分辨相近的分子的能力。

[0134] 在本实施例中,将细胞固定并透化,并用如上所述的编码探针和多聚dT锚定探针标记。然后将细胞包埋在可膨胀的聚合物凝胶中。然后,用蛋白酶K消化细胞以除去蛋白质并使凝胶的机械性质均质化。然后将凝胶在低盐缓冲液中膨胀,并最后再次包埋在不可膨胀的凝胶中以使其稳定在膨胀状态(图2A)。RNA在其多聚腺苷酸尾部锚定至凝胶,这在个体RNA分子和凝胶之间产生单点的接触,并避免在膨胀过程中拉伸mRNA。这应有助于分开相近的分子。在将样品包埋到凝胶中之前对MERFISH编码探针进行染色,因为发现至少在一些实施方案中,MERFISH探针在样品包埋两次后不能可靠地良好穿透。使用低盐缓冲液(0.5 $\times$ 盐水-柠檬酸钠(SSC))代替水进行透析以在膨胀期间保持探针的特异性结合,这导致较低的膨胀因子。凝胶体积的测量显示在0.5 $\times$ 盐水-柠檬酸钠(SSC)中的透析产生12倍的体积膨胀。注意到,只要膨胀足以分开相邻的RNA分子,较低的膨胀因子在某些应用中具有允许更快成像的优点。

[0135] 值得注意的是,在这些实验的膨胀样品中,个体RNA分子变得很好地分辨(图2B、C)并且被成功解码(图2D)。这些RNA的通过MERFISH检测的每个细胞的平均拷贝数与通过RNA-seq测量的RNA丰度相关,其中 $\log_{10}$ 值之间的Pearson相关系数为0.83(图2E)。

[0136] 为了量化膨胀对解码性能的改善,在膨胀和未膨胀样品之间比较每个细胞的每个RNA种类的平均MERFISH计数。发现在膨胀样品中检测到的129个RNA种类的每个细胞的拷贝数与未膨胀样品中的拷贝数强相关,其中 $\log_{10}$ 值之间的Pearson相关系数高至0.88(图2F)。然而,对于这些129个RNA种类,在膨胀样品中检测到的每个细胞的拷贝数是未膨胀样品中检测到的拷贝数的7.5 $\pm$ 0.3倍(平均值 $\pm$ SEM)或6.6倍(中值),这表明对于该高丰度文库,与未膨胀样品相比,膨胀样品具有显著更高的检测效率。每个细胞的拷贝数结果在实验的重复之间是高度可重复的(图2G)。与smFISH测量的比较显示,通过MERFISH测定的每个细胞的拷贝数是通过smFISH测定的拷贝数的105 $\pm$ 9%(平均值 $\pm$ SEM)或111%(中值)(图2H),这表明膨胀后的检测效率接近100%。

[0137] 图2显示了膨胀的U-2 OS细胞中高丰度RNA文库的MERFISH测量。图2A显示了膨胀MERFISH的基本实现的示意图。用亚磷酰胺修饰的多聚dT锚定探针和MERFISH编码探针共同染色靶标RNA,然后通过多聚dT探针锚定到可膨胀的聚合物凝胶上。然后,将样品在具有蛋

白酶K和SDS的消化缓冲液中孵育,分别用于去除蛋白质和脂质,以使凝胶的机械性质均质化。最后,然后通过低盐缓冲液中透析使凝胶膨胀,并再次将膨胀的凝胶包埋在聚丙烯酰胺凝胶中,使其稳定在膨胀状态。图2B显示了在z-扫描的一个焦平面处用129种RNA的编码探针染色、包埋、清除、膨胀、再包埋和用Cy5标记的读出探针可视化的U-2 OS样品中MERFISH测量的图像。实线标记细胞的边缘,虚线标记细胞核的DAPI染色区域。图2C显示了对于图2B中加框的子区域的所有8轮双色smFISH成像的高通滤波荧光图像。阴影表示Cy5通道(绿色)和Alexa 750通道(红色)。图2D显示了基于它们的经测量的二进制条形码着色的图2B中所有解码的RNA的定位。插图显示了来自图2C中所示区域的所有解码的RNA的定位。显示了所有光学切片上的解码RNA。实线标记细胞的边缘,虚线标记细胞核的DAPI染色区域。图2E显示了通过MERFISH测定的129个RNA种类的每个细胞的平均RNA拷贝数相对于通过RNA-seq测定的丰度。 $\log_{10}$ 值之间的Pearson相关系数( $\rho_{10}$ )为0.83。图2F显示了通过膨胀MERFISH测定的129个RNA种类的每个细胞的平均RNA拷贝数相对于未膨胀样品中MERFISH测定的平均RNA拷贝数。对膨胀MERFISH测量进行12微米深度的z扫描,对未膨胀的样品进行5微米深度的z扫描,以确保至少90%的RNA分子包括在成像深度中。 $\log_{10}$ 值之间的Pearson相关系数( $\rho_{10}$ )是0.88。对于这129个RNA种类在膨胀样品中检测到的每个细胞的拷贝数是在未膨胀样品中检测到的拷贝数的7.5+/-0.3倍(平均值+/-SEM)或6.6倍(中值)。图2G显示了在一个膨胀样品中相对于在第二个膨胀样品中检测到的129个RNA种类的每个细胞的平均RNA拷贝数。 $\log_{10}$ 值之间的Pearson相关系数( $\rho_{10}$ )是0.94。图2H显示了129个RNA中的12个的通过膨胀MERFISH测定的每个细胞的平均RNA拷贝数相对于通过经清除和未膨胀样品的smFISH测量的平均RNA拷贝数。如图1所示对未膨胀但经清除的样品进行smFISH测量。 $\log_{10}$ 值之间的Pearson相关系数( $\rho_{10}$ )为0.96。MERFISH和smFISH结果之间的平均比率为105%+/-9%(SEM,n=12个RNA种类),中值比率为111%。

#### [0138] 实施例4

[0139] 将免疫荧光成像整合到膨胀样品的MERFISH实验中。细胞是由不同结构和区室组成的高度复杂的系统,免疫荧光是一种用于可视化特定亚细胞结构和区室的强大技术。因此,本实施例测试了膨胀样品中是否可以组合免疫荧光和MERFISH成像。已经使用寡核苷酸缀合的抗体和具有甲基丙烯酰基的互补探针以将信号掺入凝胶中在免疫染色的样品中证明了膨胀显微术。然后这些寡核苷酸标记蛋白质靶标的位置,并且可以通过荧光团标记的互补寡核苷酸的杂交来检测,所述互补寡核苷酸被整合到MERFISH读出测量中。在实验过程中,在细胞与MERFISH编码探针和RNA锚定探针杂交后,用寡核苷酸标记的(二级)抗体对蛋白质靶标进行免疫染色,然后将标记的样品包埋在可膨胀的凝胶中。将亚磷酰胺修饰添加到抗体上的寡核苷酸中,使得它可以在包埋步骤期间掺入聚合物凝胶中。如上所述,在消化、凝胶膨胀和第二次包埋在不可膨胀的凝胶中后,进行MERFISH读出程序以首先检测RNA编码探针中的读出序列,然后进行另外一轮FISH检测以读出代表蛋白质靶标的寡核苷酸序列。

[0140] 对于本实施例,钙粘蛋白在培养的U-2 OS细胞中免疫染色,其用上述相同的高丰度MERFISH文库染色。观察到细胞周围钙粘蛋白的特异性染色(图3A)以及相同细胞中清楚分辨的smFISH斑点(图3B-D)。与免疫荧光的组合不影响MERFISH成像质量(图3E和3F)。观察到MERFISH结果和RNA-seq结果的相同的高相关性, $\log_{10}$ 值之间的Pearson相关系数为0.83

(图3E)。在免疫染色样品中检测到的每个细胞的每个RNA种类的平均拷贝数与未经免疫染色的检测样品中的那些强烈相关,  $\log_{10}$  值之间的Pearson相关系数为0.95(图3F)。平均而言, 免疫染色和未免疫染色的样品之间每个细胞的每个RNA种类的拷贝数比率为99% $\pm$ 3%(SEM), 这表明当与免疫荧光组合时MERFISH的性能的削弱可忽略不计。膨胀的和MERFISH标记的样品中钙粘蛋白的染色也看起来类似于对照U-2 OS样品中的钙粘蛋白染色, 该对照在固定后直接免疫染色而没有RNA的MERFISH成像(图3G)。为了量化抗体染色是否受MERFISH RNA标记、凝胶包埋、消化和膨胀程序的影响, 在膨胀MERFISH样品与对照样品中比较细胞周围的每单位长度的钙粘蛋白强度(标准化至膨胀前的实际长度), 其具有可比较的结果(图3H)。

[0141] 图3显示了将免疫荧光成像整合到膨胀样品的MERFISH测量中。图3A显示了用MERFISH编码探针、钙粘蛋白第一抗体、寡核苷酸缀合的第二抗体和Cy5-缀合的互补探针染色的膨胀的U-2 OS样品的图像。图3B显示了图3A中的加框子区域的MERFISH测量的smFISH图像, 其在z扫描的一个焦平面处用Cy5标记的读出探针可视化。图3C显示了针对图3B中的加框子区域的所有8轮双色smFISH成像的高通滤波荧光图像。阴影表示Cy5通道(绿色)和Alexa 750通道(红色)。图3D显示了基于它们的经测量的二进制条形码着色的图3B中所有解码的RNA的定位。插图显示了图3C中所示区域的所有解码的RNA的定位。显示了所有光学切片上的解码RNA。图3E显示了通过MERFISH测定的129个RNA种类的每个细胞的平均RNA拷贝数相对于通过RNA-seq测定的丰度。 $\log_{10}$  值之间的Pearson相关系数( $\rho_{10}$ )为0.83。图3F显示了免疫染色样品中的通过膨胀MERFISH测定的129个RNA种类的每个细胞的平均RNA拷贝数相对于仅MERFISH的样品中的平均RNA拷贝数。对两者进行12微米深度的z扫描。 $\log_{10}$  值之间的Pearson相关系数( $\rho_{10}$ )为0.95。在两个实验中测定的RNA拷贝数之间的平均比率为99% $\pm$ 3%(SEM, n=129个RNA种类), 中值比率为94%。图3G显示了在固定和透化后立即用钙粘蛋白第一抗体、寡核苷酸缀合的第二抗体和Cy5-缀合的互补探针染色的未膨胀的培养的U-2 OS细胞的图像。图3H显示了膨胀的MERFISH和对照样品中每单位长度的钙粘蛋白强度(标准化至膨胀前的实际长度)。在去除背景后, 针对z扫描的堆叠组合强度。

#### [0142] 实施例5

[0143] 这些实施例已经证明了将MERFISH和扩展显微术组合起来测量高丰度RNA文库的方法。已显示, 当总RNA密度较高时, 来自相近RNA分子的信号之间的重叠降低了MERFISH检测效率, 并且样品膨胀能够克服这种重叠问题并大大提高检测效率, 对于这些高丰度的RNA文库恢复MERFISH的近100%的检测效率。

[0144] 每个哺乳动物细胞都有数以万计的不同RNA种类。MERFISH成像能够用于测定单细胞中的成千上万的RNA种类。为了增加能够在单细胞中同时成像的RNA种类的数量同时保持相似数量的成像轮, 每轮成像的RNA的密度可以增加并且可能成为能够成像的RNA数量的重要限制因素。膨胀MERFISH方法能够促进能够在单细胞中测量的RNA种类数量的显著增加。

[0145] 在这些实施例中, 在这些膨胀样品中证明了MERFISH与免疫荧光的组合, 其能够提供关于转录组分析的细胞背景的重要信息。在具有荧光团缀合的抗体的常规免疫荧光中, 光谱不同的荧光团的数量限制了能够同时研究的靶标的数量。寡核苷酸缀合的抗体的使用潜在地允许使用仅一个或几个光谱不同的荧光团在连续轮的杂交和成像中可视化许多蛋

白质。因此, MERFISH和免疫荧光成像的组合可以允许在单细胞中同时进行组合的蛋白质组学和转录组学测量。

#### [0146] 实施例6

[0147] 本实施例描述了在一些上述实施例中使用的各种材料和方法。

[0148] 编码探针的设计。MERFISH编码探针使用16位汉明权重-4汉明距离-4代码设计, 具有140种可能的条形码。在这种编码方案中, 所使用的所有条形码之间分开的汉明距离至少为4, 因此必须不正确地读出至少4位才能将一个有效条形码改变为另一个。使用4的恒定汉明权重(即每个条形码中的“1”位的数量)来避免由于“1”到“0”和“0”到“1”误差的差分率引起的潜在测量偏倚。编码探针组含有92个编码探针/RNA。每个编码探针包含使用pipeline设计的30-nt靶区域, 两侧是从分配给每个RNA的四个中随机选择的两个20-nt读出序列, 一个是在5'末端的20-nt引发区, 另一个是在3'末端的使用T7启动子的反向互补物的20-nt引发区。在读出序列和靶区域之间添加额外的腺苷核苷酸间隔区, 以防止读出序列中的末端鸟嘌呤三联体, 并防止靶区域与来自相邻序列的G组合形成G四联体。5'末端的引发区在末端具有胸腺嘧啶, 其置于5'末端的引发区和编码区的连接处。这被设计成在探针构建的逆转录步骤中使用的正向引物上掺入尿嘧啶, 其能够通过尿嘧啶特异性切除试剂(USER)酶裂解。可裂解的引物设计与使用T7启动子的反向互补物作为第二引物一起允许构建长度为72nt的编码探针。探针长度的减少可以促进探针的渗透并减少非特异性结合。

[0149] 为了在我们的文库中包含一些丰富但短的RNA, 不是设计靶向RNA的不同区域的探针, 而是允许探针与另一个探针共有至多20nt。这种修改有助于将可为每个基因设计的探针数量增加2倍。从RNA的所有潜在靶区域中随机选择92个靶区域/RNA。通过允许相邻编码探针之间至多20-nt的重叠, 该设计允许短至1, 200nt的RNA被92个具有30-nt靶向序列的编码探针靶向。由于给定的细胞RNA通常由92个编码探针中少于三分之一的探针结合, 因此预期具有重叠靶向区域的编码探针基本上不会相互干扰, 但会部分地补偿由于局部靶标RNA上难以接近的局部区域(例如二级结构)导致的结合减少或合成期间探针的丢失。

[0150] 构建编码探针。编码探针组由复杂的寡核苷酸库构建。简而言之, 寡核苷酸库(CustomArray)通过有限循环PCR扩增以制备体外转录模板, 通过体外转录将这些模板转化为RNA(New England Biolabs), 并通过逆转录将RNA转化为DNA(Maxima RT H, Thermo Fisher Scientific)。然后用1:30(体积比)稀释的USER酶(New England Biolabs)在37°C下孵育24小时以消化探针, 以在5'末端的引发区和靶区域之间的尿嘧啶位点处裂解掉引发区。之后, 通过用于除去RNA的碱水解和柱纯化(Zymo Research)纯化DNA。将最终探针重悬于无RNA酶的水中并在-20°C下储存。

[0151] 寡核苷酸与第二抗体缀合。含有所需读出探针的寡核苷酸通过NHS-酯和无铜点击化学的组合与第二抗体缀合。首先, 使用NHS-酯化学法, 用无铜点击交联剂标记第二抗体。具体地, 根据制造商的说明, 使用基于旋转柱的透析膜(Amicon, 100kDa截留分子量)从未缀合的驴抗兔第二抗体(Thermoscientific)中除去叠氮化物防腐剂。将NHS-酯、PEG5、DBCO交联剂(Kerafast)在无水DMSO(Thermoscientific)中稀释至10微摩尔浓度。然后将2微升该交联剂与100微升的2mg/mL抗体在1×磷酸盐缓冲盐水(PBS)中混合。将该反应室温孵育1小时, 然后使用如上所述的Amicon柱通过第二轮纯化终止。通过样品在280nm(抗体)和309nm(DBCO)的相对吸收度测定每个抗体的DBCO交联剂的平均数。平均而言, 上述交联剂的量产

生每个抗体约7个DBCO交联剂。

[0152] 含有所需序列以及允许与聚合物凝胶交联的5'-亚磷酰胺和允许与DBCO标记的抗体交联的3'-叠氮化物的寡核苷酸探针从IDT订购并在1×PBS中悬浮至100微摩尔。然后将20微升合适的寡核苷酸加入到100微升DBCO标记的抗体中,终浓度为约2mg/mL。将该反应在4℃孵育至少12小时。标记的抗体未进一步纯化,因为预期不与抗体缀合的残留寡核苷酸易于从样品中洗去。

[0153] 细胞培养和固定。用含有10% (体积比) FBS (Thermo Fisher Scientific) 的Eagle最小必需培养基(ATCC)培养U-2 OS细胞(ATCC)。将细胞铺板在40mm直径的no.1.5的盖玻片(Bioprotech)上,每个盖玻片350,000个细胞,并在培养皿中在37℃和5%CO<sub>2</sub>孵育48小时。将细胞固定、透化并用编码探针染色。简而言之,将细胞在1×PBS中的4% (体积比) 多聚甲醛(Electron Microscopy Sciences)中室温固定15分钟,用1×PBS洗涤三次,用1×PBS中的0.5% (体积比) Triton X-100 (Sigma) 室温透化10分钟,并用1×PBS洗涤一次。

[0154] 编码探针染色用于MERFISH测量。将透化的细胞在编码洗涤缓冲液中孵育5分钟,所述洗涤缓冲液包含2×盐水-柠檬酸钠(SSC) (Ambion) 和30% (体积比) 甲酰胺(Ambion)。然后将编码杂交缓冲液中的30微升的约300微摩尔编码探针和3.3微摩尔锚定探针(20-nt序列的交替dT和胸苷锁定核酸(dT+),其具有读出序列的20-nt反向互补序列和5'-亚磷酰胺修饰(Integrated DNA Technologies))加入到Parafilm(Bemis)的表面,并用含有细胞的盖玻片覆盖。然后将样品在杂交炉内的潮湿室中在37℃孵育40小时。编码杂交缓冲液由补充有0.1% (重量体积比) 酵母tRNA (Life Technologies)、1% (体积比) 鼠RNA酶抑制剂(New England Biolabs) 和10% (重量体积比) 葡聚糖硫酸盐(Sigma) 的编码洗涤缓冲液组成。然后用编码洗涤缓冲液洗涤细胞,并在47℃孵育30分钟;该洗涤步骤重复一次。

[0155] 免疫染色。在用编码和锚定探针染色后,将样品用4% (体积比) 多聚甲醛(Electron Microscopy Sciences) 室温后固定10分钟。然后将样品在含有补充有3% (体积比) RNasin核糖核酸酶抑制剂(Promega)、6% (体积比) 鼠RNA酶抑制剂(New England Biolabs) 和1mg/ml酵母tRNA (Life Technologies) 的2×SSC (Ambion) 中的4% (重量体积比) UltraPure BSA (Thermo Fisher Scientific) 的封闭缓冲液中室温封闭30分钟。将样品与第一抗体(抗泛钙黏着蛋白,Abcam) 在封闭缓冲液中以2微克/ml的浓度室温孵育1小时,并用2×SSC洗涤三次,每次10分钟。然后将样品与寡核苷酸标记的第二抗体在封闭缓冲液中以3.75微克/ml的浓度室温孵育1小时,然后用2×SSC洗涤三次,每次10分钟。用4% (体积比) 多聚甲醛(Electron Microscopy Sciences) 再次固定样品10分钟。

[0156] 未膨胀样品的包埋、消化和清除。首先将盖玻片上的染色样品与脱气的聚丙烯酰胺溶液一起孵育5分钟,该溶液含有4% (体积比) 的19:1丙烯酰胺/双丙烯酰胺(BioRad)、60mM Tris·HCl pH 8 (ThermoFisher)、0.3M NaCl (ThermoFisher)、0.2% (体积比) 四甲基乙二胺(TEMED) 和1:25,000稀释的0.1微米直径的羧酸盐修饰的橙色荧光珠(Life Technologies)。珠子用作基准标记,用于在多轮smFISH成像中拍摄的图像的对齐。

[0157] 为了浇铸薄的聚丙烯酰胺薄膜,对于每个样品,将50微升的最终溶液添加至已用1mL GelSlick (Lonza) 预处理5分钟(以便不粘至聚丙烯酰胺凝胶) 的玻璃板(TED Pella) 的表面。将如上所述处理的盖玻片上的样品吸出并轻轻倒置在该50微升液滴上,以在盖玻片和玻璃板之间形成薄层溶液。然后使该溶液在填充氮气的自制室中室温聚合2小时。然后轻

轻地分开盖玻片和玻璃板,并用具有 $2\times$ SSC (Ambion) 中的2% (重量体积比) SDS (Thermo Fisher Scientific)、0.5% (体积比) Triton X-100的消化缓冲液洗涤PA膜一次。消化缓冲液使用2% (重量体积比) SDS以促进脂质去除。洗涤后,用补充有1% (体积比) 蛋白酶K (New England Biolabs) 的消化缓冲液覆盖凝胶。将样品在该缓冲液中在潮湿的 $37^{\circ}\text{C}$ 培养箱中消化 $>12$ 小时,然后用 $2\times$ SSC洗涤三次。MERFISH测量立即进行或将样品在 $4^{\circ}\text{C}$ 下储存在补充有0.1% (体积比) 鼠RNA酶抑制剂 (New England Biolabs) 的 $2\times$ SSC中不超过48小时。

[0158] 未膨胀样品的盖玻片的硅烷化。通过在室温下浸入37% (体积比) HCl和甲醇的1:1混合物中洗涤40mm直径的No.1.5盖玻片 (Bioprotechs) 30分钟。然后将盖玻片在去离子水中漂洗三次,并在70% (体积比) 乙醇中漂洗一次。盖玻片用氮气吹干,然后在室温下浸入0.1% (体积比) 三乙胺 (Millipore) 和0.2% (体积比) 烯丙基三氯硅烷 (Sigma) 的氯仿溶液中30分钟。盖玻片各自用氯仿和乙醇洗涤一次,然后用氮气吹干。然后在使用前将硅烷化的盖玻片在室温下在干燥室中储存过夜,以使硅烷层脱水。

[0159] 膨胀样品的包埋、消化和清除。制备含有2M NaCl (Thermo Fisher Scientific)、7.7% (重量比) 丙烯酸钠 (Sigma)、4% (体积比) 的19:1丙烯酰胺/双丙烯酰胺 (BioRad) 和60mM Tris HCl pH8 (Thermo Fisher Scientific) 的单体溶液,在 $-20^{\circ}\text{C}$ 下等分冷冻,并在使用前解冻。将单体溶液脱气并在使用前冷却至 $4^{\circ}\text{C}$ 。将终浓度为0.2% (体积比) 的TEMED和1:5,000稀释的0.1微米直径的羧酸酯改性的橙色荧光珠 (Life Technologies) 加入到溶液中。珠子用作基准标记,用于在多轮smFISH成像中拍摄的图像的对齐。将染色的样品在溶液中室温孵育5分钟。然后将溶液保持在冰上并进一步补充终浓度为0.2% (重量体积比) 的过硫酸铵 (Sigma)。

[0160] 为了浇铸薄的聚合物膜,对于每个样品,将50微升的最终溶液添加至已用1mL GelSlick (Lonza) 预处理5分钟 (以便不粘至凝胶) 的玻璃板 (TED Pella) 的表面。将如上所述处理的盖玻片上的样品吸出并轻轻倒置在该50微升液滴上,以在盖玻片和玻璃板之间形成薄层溶液。然后使该溶液在填充氮气的自制室中室温聚合2小时。然后轻轻分开盖玻片和玻璃板,如上所述用具有 $2\times$ SSC (Ambion) 中的2% (重量体积比) SDS (Thermo Fisher Scientific)、0.5% (体积比) Triton X-100的消化缓冲液洗涤凝胶一次。洗涤后,使用剃刀刀片将凝胶薄膜小心地修剪成所需的大小,并用补充有1% (体积比) 蛋白酶K (New England Biolabs) 的消化缓冲液覆盖。将样品在该缓冲液中在潮湿的 $37^{\circ}\text{C}$ 培养箱中消化 $>12$ 小时,然后用 $2\times$ SSC (Ambion) 洗涤三次。在消化过程中凝胶会膨胀约1.5倍。

[0161] 膨胀和再包埋。消化后,将样品在补充有0.2% (体积比) 蛋白酶K (New England Biolabs) 的 $0.5\times$ SSC缓冲液中室温膨胀。加入蛋白酶K以将样品保持在无RNA酶的环境中。每30分钟更换缓冲液直至样品不再膨胀 (通常约2小时)。将膨胀的凝胶重新包埋到聚丙烯酰胺凝胶中以稳定凝胶用于进行连续轮的读出探针杂交和成像。简言之,将样品在由4% (体积比) 的19:1丙烯酰胺/双丙烯酰胺 (BioRad) 和30mM NaCl (Thermo Fisher Scientific)、6mM Tris HCl pH 8 (Thermo Fisher Scientific) 和0.2% (体积比) 的TEMED (Sigma) 组成的再包埋溶液中在振荡器上室温孵育20分钟。然后将再包埋溶液保持在冰上并进一步补充过硫酸铵 (Sigma),终浓度为0.2% (重量体积比)。将凝胶置于结合硅烷处理的盖玻片上 (见下文),用溶液冲洗并用KimWipes (Kimtech) 快速干燥。将带有凝胶的盖玻片放入自制的氮气室中,盖上玻璃板 (TED Pella) 并室温聚合1小时。

[0162] 盖玻片的结合硅烷处理。将直径40毫米的no.1.5盖玻片(Bioptechs)在1M氢氧化钾中超声处理30分钟,用去离子水洗涤三次并再次在70%(体积比)乙醇中超声处理30分钟。然后将盖玻片在室温下浸入5%(体积比)冰醋酸和99%(体积比)乙醇中的0.38%(体积比)结合硅烷(GE Healthcare)中1小时。用70%(体积比)乙醇快速洗涤三次后,将盖玻片放入60℃的烘箱中直至完全干燥。盖玻片用干燥剂能够在密封容器中储存长达一个月。

[0163] 连续轮的读出探针染色和成像。为了帮助选择正确的成像焦平面,将包埋的样品在盘中与第一对双色读出探针在杂交缓冲液中室温杂交30分钟(膨胀样品)或10分钟(未膨胀样品),杂交缓冲液由无核酸酶的水中的2×SSC(Ambion)、5%(体积比)碳酸亚乙酯(Sigma)、0.1%(体积比)鼠RNA酶抑制剂和3nM适当的读出探针组成,在由无核酸酶水中的2×SSC(Ambion)和10%(体积比)碳酸亚乙酯(Sigma)组成的洗涤缓冲液中洗涤20分钟(膨胀样品)或7分钟(未膨胀的样品)。然后用2×SSC洗涤样品一次,用2×SSC(Ambion)中10微克/微升的DAPI染色10分钟,并在2×SSC(Ambion)中洗涤3次,每次5分钟。

[0164] 将样品安装到流动室中,并且通过包含三个计算机控制的八通阀(Hamilton)和计算机控制的蠕动泵(Gilson)的自制流体系统控制通过该室进行缓冲液交换以进行随后的成像和杂交,其中对于膨胀样品使用延长的流动和孵育时间(除了前面部分所述的杂交和洗涤时间)。三(2-羧乙基)膦盐酸盐(TCEP)孵育时间增加到30分钟,并且缓冲液交换时间增加到7分钟以允许扩散以在凝胶内部达到平衡。

[0165] 为了准确地比较不同样品之间每个细胞的计数,应捕获细胞中超过90%的靶标RNA。为了量化不同z平面上的RNA分布,在跨越细胞的离散成像平面上以0.5微米的步长进行光学切片,并且将smFISH信号的分布量化为z位置的函数。发现>90%的RNA位于未膨胀样品中的前5微米深度体积中,并且位于来自盖玻片表面的膨胀样品中的前12微米深度体积中。因此,扫描5微米深度的体积用于未膨胀的样品,并且扫描12微米深度的体积用于步长为1微米的膨胀样品。z方向的扫描由Nano-F200纳米定位仪(Mad City Labs)控制。

[0166] 在高通量成像平台上进行连续MERFISH成像和信号去除。简而言之,在杂交后,使用2,048×2,048像素的科学互补金属氧化物半导体(sCMOS)相机结合高数值孔径(NA=1.3)和高放大倍数(60×)硅油物镜以223×223微米的FOV面积对样品成像。在成像约100-400个FOV后,通过将样品在由2×SSC和50mM TCEP(Sigma)组成的还原性裂解缓冲液中孵育来熄灭读出探针的荧光。将杂交、成像和化学裂解过程重复八次,其中405nm DAPI通道与第一轮读出成像一起成像。

[0167] 图像配准(Image Registration)和解码。如下进行跨成像轮的相同FOV的图像的配准以及RNA条形码的解码。简而言之,基于基准珠的位置校正来自不同成像轮的每个位置处的z-堆叠的横向偏移。将经校正的z-堆叠进行高通滤波以去除背景,解卷积以收紧RNA斑点,然后进行低通滤波以连接在图像之间的位置中略有不同的RNA质心。为了校正颜色通道之间的亮度差异,首先通过均衡其强度直方图来对图像进行标准化,并通过迭代过程进一步细化,以消除不同位之间的荧光强度的显著变化。针对每个焦平面处的每个FOV中的每个像素观察到的16个标准化强度值的集合(对应于16个读出探针成像)表示16维空间中的矢量。将像素矢量标准化并与16位MHD4代码中的140个条形码中的每一个进行比较。如果矢量和条形码之间的欧几里德距离小于由单-位(single-bit)误差的距离定义的给定阈值,则将像素分配给条形码。使用具有最大邻域连通性的3-D连接阵列将相邻像素组合成单个推

定的RNA。使用基于倒置条形码密度的分水岭算法计算细胞边界,其中DAPI染色区域作为初始种子。

[0168] 计算拆分到由哈佛大学的FAS Division of Science, Research Computing Group支持的Odyssey机群和包含两个10核Intel Xeon E5-2680 2.8GHz CPU和256GB RAM的桌面服务器之间。

[0169] 单分子FISH。每个smFISH探针包含30-nt靶区域和20-nt定制设计的读出序列。为每个基因设计了48个探针。透化后,将细胞在编码洗涤缓冲液中孵育5分钟,所述洗涤缓冲液包含2×盐水-柠檬酸钠(SSC)(Ambion)和30%(体积比)甲酰胺(Ambion)。然后将30微升的编码杂交缓冲液中的2微摩尔smFISH探针和3.3微摩尔锚定探针添加至Parafilm(Bemis)表面,并用含细胞的盖玻片覆盖。然后将样品在杂交炉内的潮湿室中在37°C下孵育24小时。编码杂交缓冲液由补充有0.1%(重量体积比)酵母tRNA(Life Technologies)、1%(体积比)鼠RNA酶抑制剂(New England Biolabs)和10%(重量体积比)葡聚糖硫酸盐(Sigma)的编码洗涤缓冲液组成。然后用编码洗涤缓冲液洗涤细胞,并在47°C孵育30分钟;该洗涤步骤重复一次。然后使用如上对未膨胀的样品的包埋、消化和清除所述的基于基质印迹的清除方法包埋和清除细胞。将包埋的样品在培养皿中与杂交缓冲液中的Cy5 20-nt读出探针杂交,洗涤并用DAPI染色,如连续轮的读出探针染色和成像部分所述。收集图像,以1微米的步长扫描5微米深度的体积,并收集每个smFISH探针组100个FOV。

[0170] 虽然已在本文中描述和说明了本发明的几个实施方案,但本领域普通技术人员将容易地设想用于进行所述功能和/或获得本文所述的结果和/或一个或多个有利方面的各种其它方法和/或结构,并且这样的变化和/或修改的每一种被认为在本发明的范围内。更通常地,本领域技术人员将容易地理解,本文所述的所有参数、尺度、材料和构型意为示例性的并且实际参数、尺度、材料和/或构型将取决于对其使用本发明的教导的具体应用。本领域技术人员将认可或能够确定本文所述的特定的本发明的实施方案的许多等同物(通过使用不超过常规实验)。因此,应理解,前述实施方案仅通过举例的方式提出,并且在所附权利要求和其等同物的范围内,本发明可按与明确描述和要求保护的方式不同的方式来实施。本发明涉及本文所述的各种个体特性、系统、物品、材料、试剂盒和/或方法。另外,如果这样的特性、系统、物品、材料、试剂盒和/或方法不是相互矛盾的,则两个或更多个这样的特性、系统、物品、材料、试剂盒和/或方法的任意组合包括在本发明的范围内。

[0171] 在本说明书和通过引用并入的文献包括冲突和/或不一致的公开内容的情况下,以本说明书为准。如果通过引用并入的两个或更多个文献包括相互冲突和/或不一致的公开内容,则以具有较晚生效日期的文献为准。

[0172] 如本文中所述定义和使用的,所有定义应当被理解为优先于词典定义、通过引用并入的文献中的定义和/或所定义的术语的一般含义。

[0173] 除非明确地指示与之相反,否则不定冠词“一种/一个(a)”和“一种/一个(an)”,如本文中在说明书和权利要求中所用,应当被理解为意指“至少一种/一个”。

[0174] 短语“和/或”,如在本文中在说明书和权利要求中所用,应当被理解为意指所连接的元素的“任一个或两者”,即元素在一些情况下结合地存在以及在其它情况下分离地存在。利用“和/或”列出的多个元素应当以相同的方式来解释,即所连接的元素的“一个或多个”。除通过“和/或”从句明确确定的元素外,还可任选地存在其它元素,无论与明确确定的

那些元素相关还是无关。因此,作为非限定性实例,对“A和/或B”的提及,当与开放性措辞诸如“包含/包括”结合使用时,在一个实施方案中可仅指A(任选地包括除B外的元素);在另一个实施方案中可仅指B(任选地包括除A外的元素);在另一个实施方案中,可指A和B(任选地包括其它元素);依此类推。

[0175] 如本文中在说明书和权利要求中所用,“或”应当被理解为具有与如上定义的“和/或”相同的含义。例如,当在列表中分开各项时,“或”或“和/或”应当被理解为是包含性的,即包含至少一个,但也包括许多个元素或一系列元素中的多于一个,以及任选地,包括另外未列出的项。只有明确地指明相反的术语,例如“……中的仅一个”或“……中的恰好一个”或当用于权利要求中时的“由……组成”将指包含多个元素或一系列元素中的正好一个元素。一般地,如本文中所述的术语“或”应当仅在冠有排他性的术语诸如“任一”、“……之一”、“……中的仅一个”或“……中的恰好一个”时被解释为表示排他性选择(即“一个或另一个但非两个”)。

[0176] 如本文中在说明书和权利要求中所用,关于一系列一个或多个元素的短语“至少一个”应当被理解为意指选自一系列元素中的任何一个或多个元素的至少一个元素,但不一定包括元素列表内明确列出的每一个元素的至少一个并且不排除元素列表中的元素的任何组合。该定义还允许可任选地存在除短语“至少一个”所指的元素列表内明确确定的元素外的元素,无论与那些明确确定的元素相关还是不相关。因此,作为非限定性实例,“A和B的至少一个”(或,等同地,“A或B的至少一个”或,等同地“A和/或B的至少一个”)可在一个实施方案中指至少一个(任选地包括不止一个)A而无B存在(且任选地包括除B外的元素);在另一个实施方案中指至少一个(任选地包括不止一个)B而无A存在(且任选地包括除A外的元素);在另一个实施方案中指至少一个(任选地包括不止一个)A和至少一个(任选地包括不止一个)B(且任选地包括其它元素);依此类推。

[0177] 当在本文中使用时,词语“约”来表示数字时,应该理解,本发明的又一个实施方案包括未通过词语“约”的存在而修饰的该数字。

[0178] 还应当理解,除非明确地指明与之相反,否则在包括不止一个步骤或行为的本文请求保护的任何方法中,所述方法的步骤或行为的顺序不必须地限定于其中叙述所述方法的步骤或行为的顺序。

[0179] 在权利要求中,以及在上述说明书中,所有过渡短语诸如“包含”、“包括”、“携带”、“具有”、“含有”、“拥有”、“牵涉”、“持有”、“由……组成”等被理解为开放性的,即意指包括但不限于。只有过渡短语“由……组成”和“基本上由……组成”应当分别是封闭性的或半封闭性的过渡短语,如美国专利局专利审查程序手册第2111.03节中所示的。

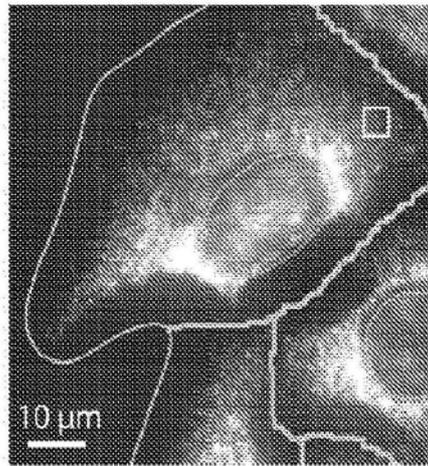


图1A

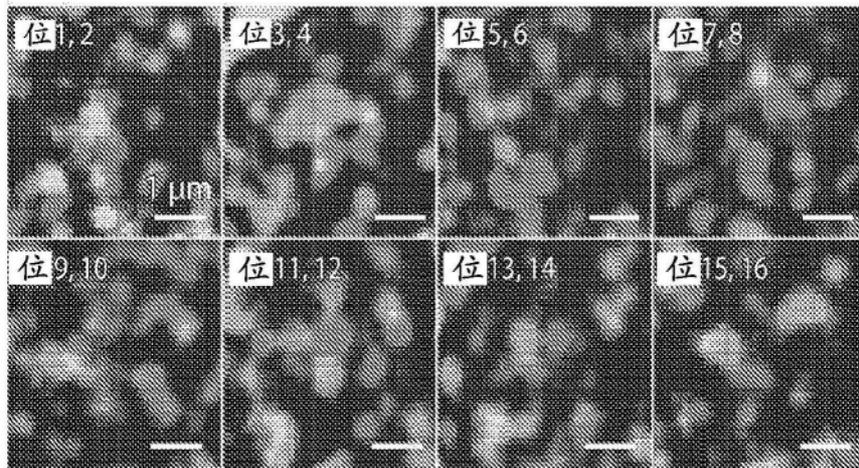


图1B

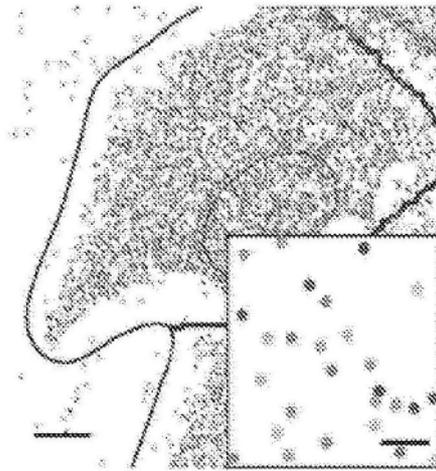


图1C

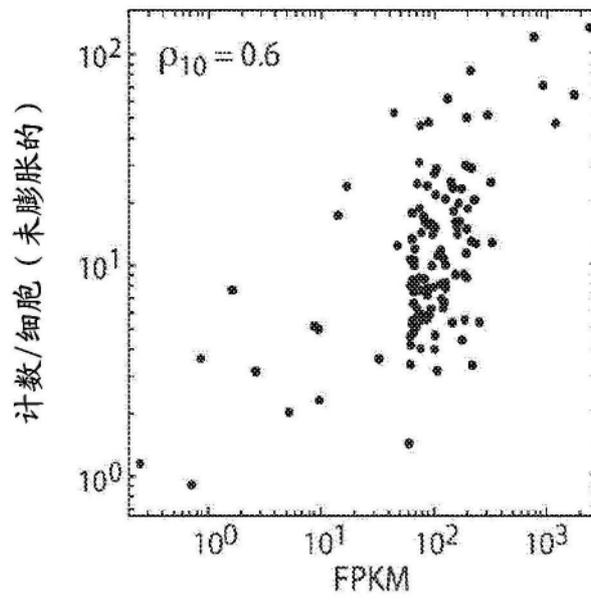


图1D

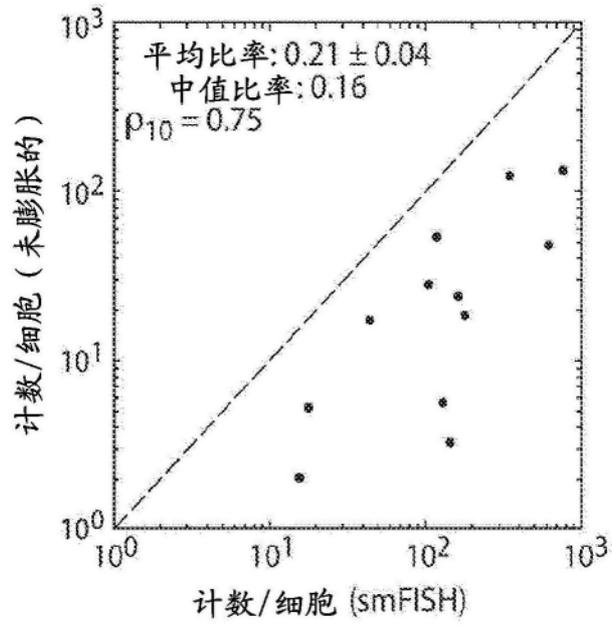


图1E

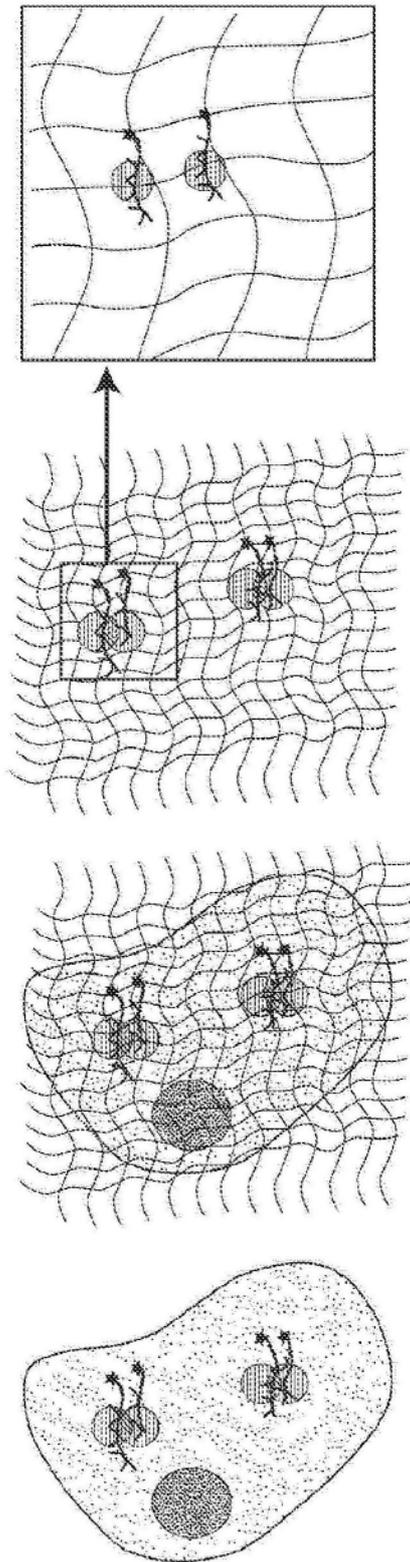


图2A

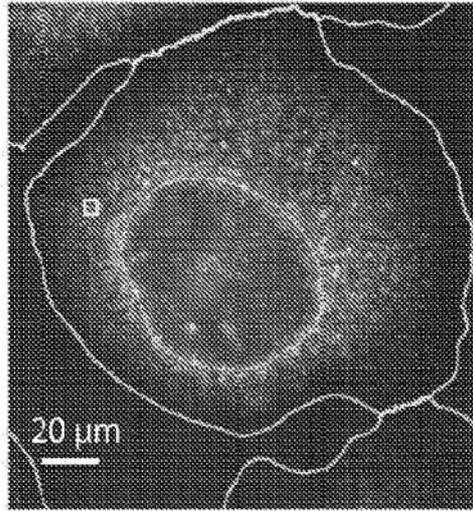


图2B

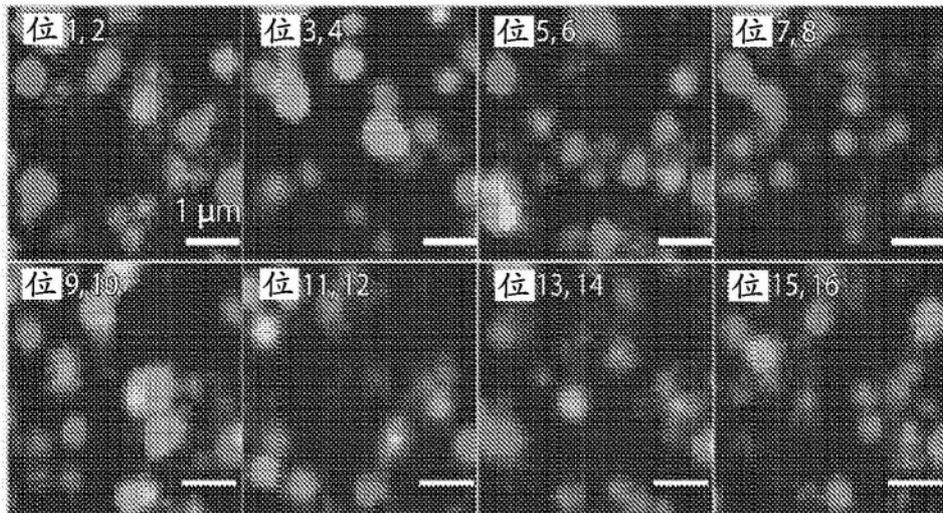


图2C

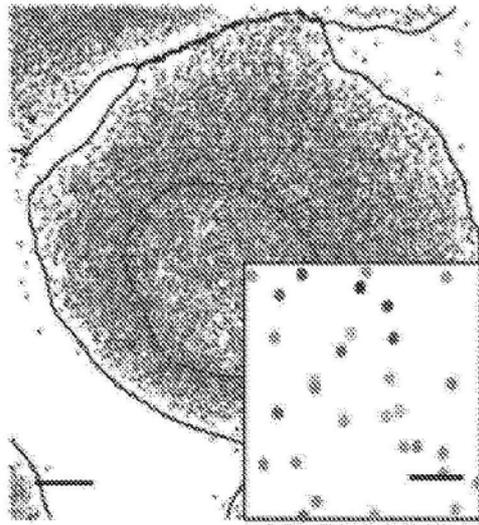


图2D

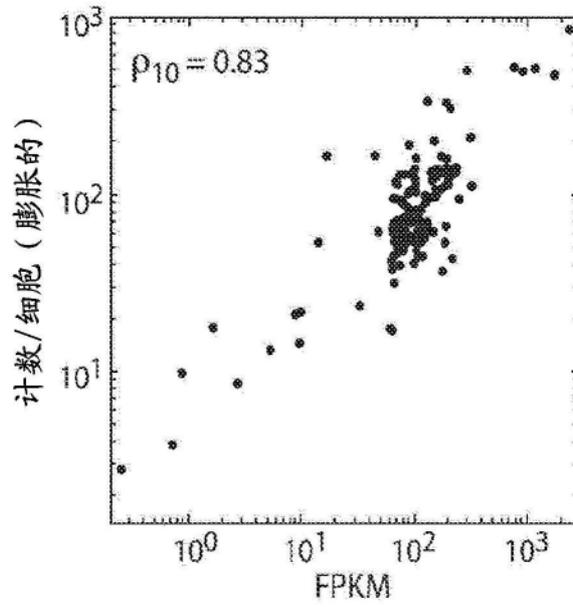


图2E

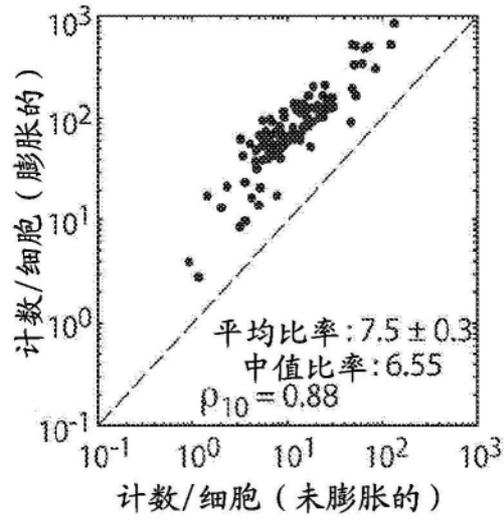


图2F

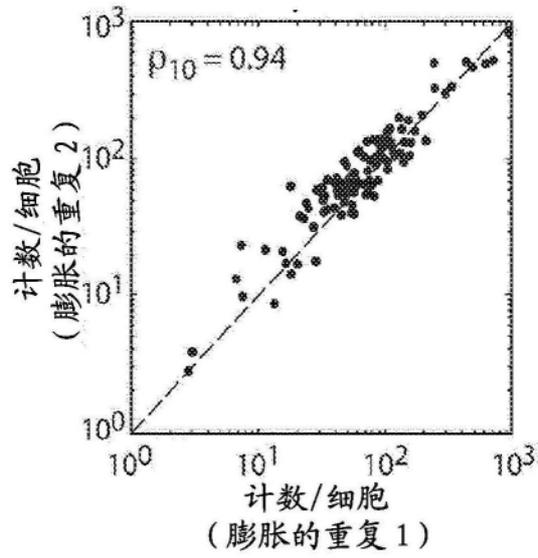


图2G

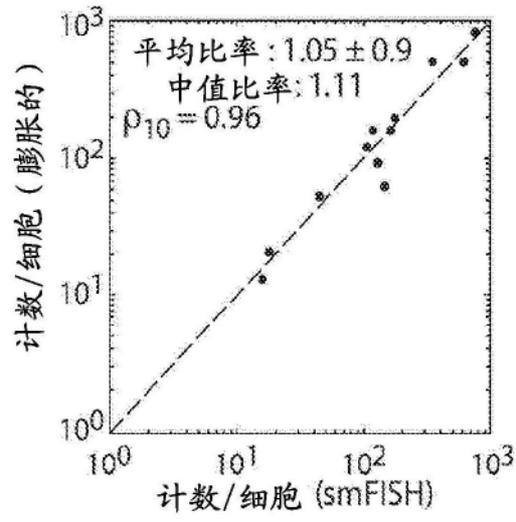


图2H

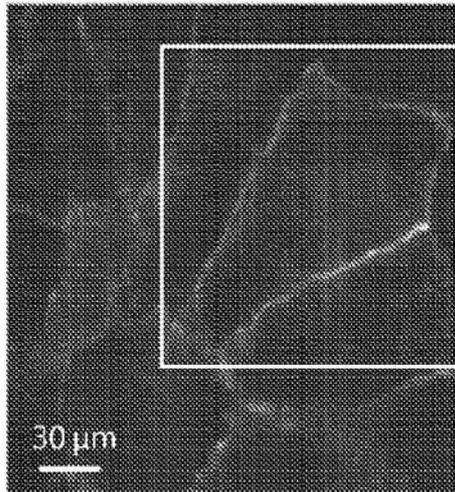


图3A

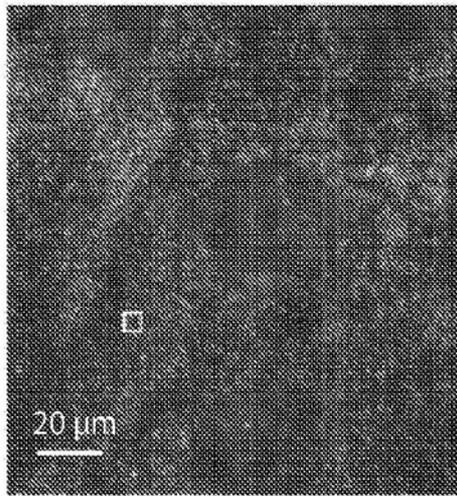


图3B

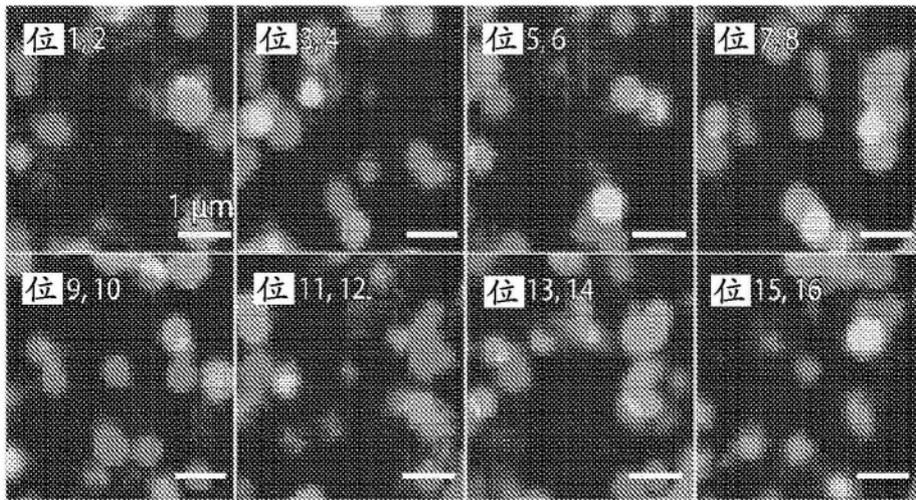


图3C

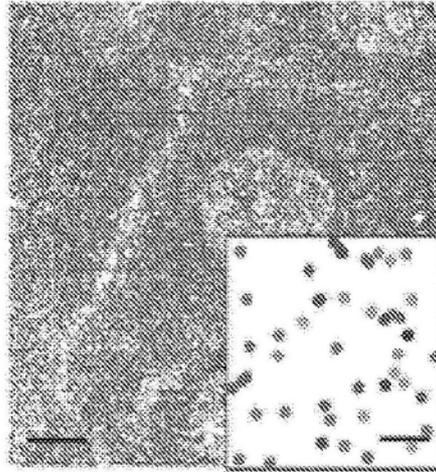


图3D

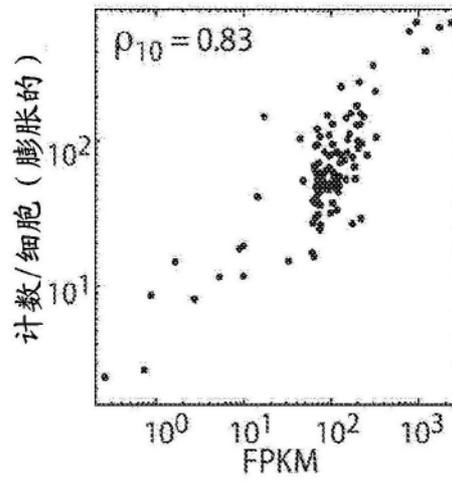


图3E

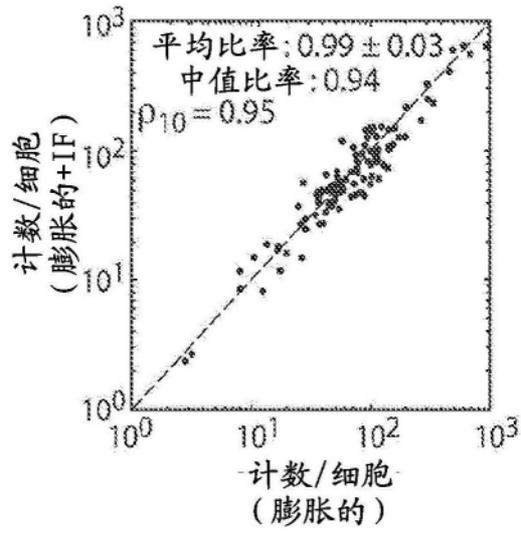


图3F

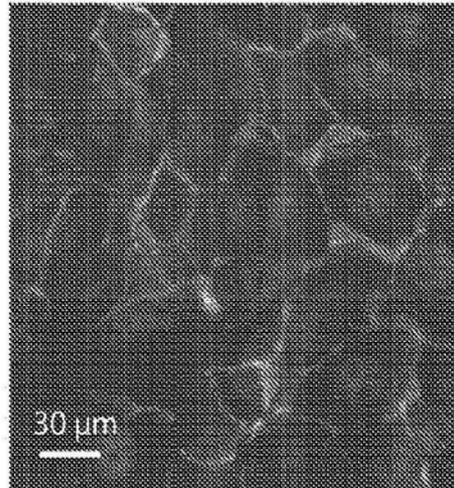


图3G

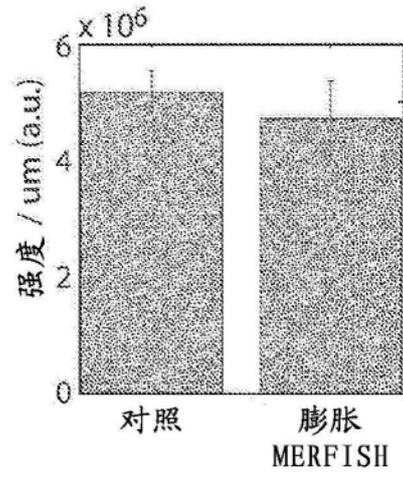


图3H

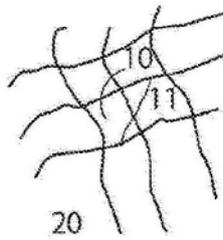


图4A

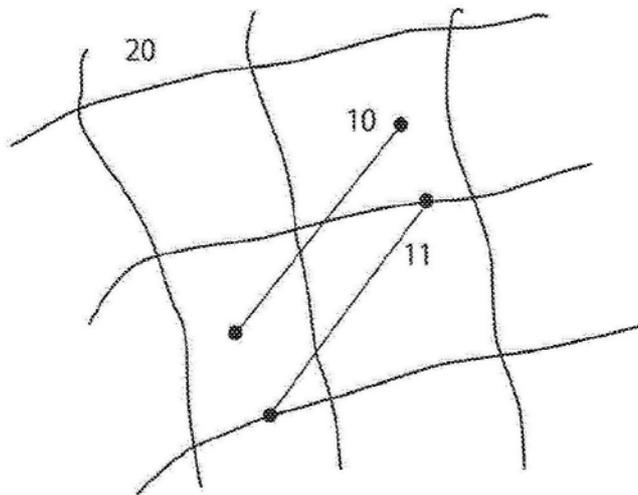


图4B

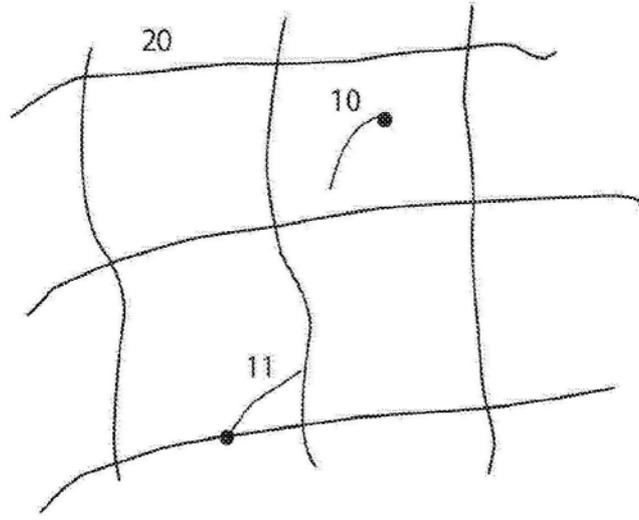


图4C