



(12) 发明专利

(10) 授权公告号 CN 111492068 B

(45) 授权公告日 2025.03.21

(21) 申请号 201880081439.8

(22) 申请日 2018.12.18

(65) 同一申请的已公布的文献号
申请公布号 CN 111492068 A

(43) 申请公布日 2020.08.04

(30) 优先权数据
62/607,812 2017.12.19 US

(85) PCT国际申请进入国家阶段日
2020.06.17

(86) PCT国际申请的申请数据
PCT/US2018/066291 2018.12.18

(87) PCT国际申请的公布数据
W02019/126209 EN 2019.06.27

(73) 专利权人 贝克顿迪金森公司

地址 美国新泽西州

(72) 发明人 格雷琴·因博恩·拉姆
艾琳·夏姆

(74) 专利代理机构 北京安信方达知识产权代理
有限公司 11262
专利代理人 刘晓杰 武晶晶

(51) Int. Cl.
C12Q 1/6869 (2006.01)
C12Q 1/6804 (2006.01)

(56) 对比文件
WO 2015031691 A1, 2015.03.05

审查员 常子月

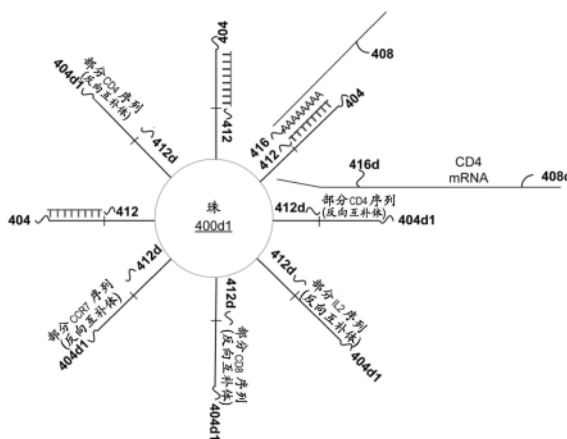
权利要求书14页 说明书62页
序列表3页 附图13页

(54) 发明名称

与寡核苷酸相关联的颗粒

(57) 摘要

本文披露了用于对核酸进行条形码化的方法、组合物和试剂盒。例如,提供了用于进行条形码化的颗粒,其中所述颗粒包含:第一多个寡核苷酸条形码,其各自包含靶结合区;和第二多个寡核苷酸,其各自包含靶特异性靶结合区。



1. 一种合成颗粒,其包含:
与所述合成颗粒偶联的第一多个寡核苷酸条形码;以及
与所述合成颗粒偶联的第二多个寡核苷酸条形码,
其中所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码中的每个包含相同的细胞标记序列和分子标记序列,
其中所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码中的至少10个包含不同的分子标记序列,
其中每个细胞标记序列包含至少6个核苷酸,其中每个分子标记序列包含至少6个核苷酸,
其中所述第一多个寡核苷酸条形码中的每个包含靶结合区,所述靶结合区包含聚(dT)区或随机多聚体,并且
其中所述第二多个寡核苷酸条形码中的每个包含靶特异性靶结合区,所述靶特异性靶结合区包含用于与靶核酸杂交的靶特异性序列,
其中所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码的比率是在从1:10至10:1的范围内。
2. 根据权利要求1所述的合成颗粒,其中所述第一多个寡核苷酸条形码的寡核苷酸条形码的靶结合区包含聚(dT)区。
3. 根据权利要求2所述的合成颗粒,其中所述第二多个寡核苷酸条形码的寡核苷酸条形码包含:所述聚(dT)区;和在所述聚(dT)区的3'的靶特异性靶结合区。
4. 根据权利要求1所述的合成颗粒,其中所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的每个包含样品标记序列。
5. 根据权利要求1所述的合成颗粒,其中所述靶结合区和所述靶特异性靶结合区具有至多50%序列同一性。
6. 根据权利要求1所述的合成颗粒,其中所述第一多个寡核苷酸条形码中的至少两个具有不同长度,和/或所述第二多个寡核苷酸条形码中的至少两个具有不同长度。
7. 根据权利要求1所述的合成颗粒,其中所述第一多个寡核苷酸条形码中的至少两个的序列是相同的,和/或所述第二多个寡核苷酸条形码中的至少两个的序列是相同的。
8. 根据权利要求1所述的合成颗粒,其中所述第一多个寡核苷酸条形码中的至少两个包含相同的序列,和/或所述第二多个寡核苷酸条形码中的至少两个包含相同的序列。
9. 根据权利要求1-8中任一项所述的合成颗粒,其包含第一多个间隔子寡核苷酸,其中所述第一多个间隔子寡核苷酸中的每个不包含所述靶结合区或所述靶特异性靶结合区。
10. 根据权利要求9所述的合成颗粒,其中所述合成颗粒包含第二多个间隔子寡核苷酸,其中所述第二多个间隔子寡核苷酸中的每个不包含所述靶结合区或所述靶特异性靶结合区。
11. 根据权利要求1-8中任一项所述的合成颗粒,其中所述第一多个寡核苷酸条形码中的至少一个比所述第二多个寡核苷酸条形码中的至少一个更短。
12. 根据权利要求1-8中任一项所述的合成颗粒,其中所述第二多个寡核苷酸条形码中的至少一个比所述第一多个寡核苷酸条形码中的至少一个更短。
13. 根据权利要求1-8中任一项所述的合成颗粒,其中所述第一多个寡核苷酸条形码

和/或所述第二多个寡核苷酸条形码中的至少100个包含不同的分子标记序列。

14. 根据权利要求1-8中任一项所述的合成颗粒,其中所述第二多个寡核苷酸条形码中的两个寡核苷酸条形码的两个靶特异性靶结合区包含不同的靶特异性序列。

15. 根据权利要求1-8中任一项所述的合成颗粒,其中所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的每个包含空间标记序列,并且其中所述空间标记序列彼此差异至少一个核苷酸。

16. 根据权利要求1-8中任一项所述的合成颗粒,其中所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的每个包含通用序列,其中所述通用序列是相同的。

17. 根据权利要求1-8中任一项所述的合成颗粒,其中所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少一个固定在所述合成颗粒上。

18. 根据权利要求1-8中任一项所述的合成颗粒,其中所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少一个部分固定在所述合成颗粒上。

19. 根据权利要求1-8中任一项所述的合成颗粒,其中所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少一个包封在所述合成颗粒中。

20. 根据权利要求1-8中任一项所述的合成颗粒,其中所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少一个部分包封在所述合成颗粒中。

21. 根据权利要求1-8中任一项所述的合成颗粒,其中所述合成颗粒是可破坏的。

22. 根据权利要求1-8中任一项所述的合成颗粒,其中所述合成颗粒包含珠。

23. 根据权利要求22所述的合成颗粒,其中所述珠包含磁珠。

24. 根据权利要求1-8和10中任一项所述的合成颗粒,其中所述合成颗粒包含可破坏的水凝胶颗粒。

25. 根据权利要求1-8中任一项所述的合成颗粒,

其中所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的每个包含接头官能团,

其中所述合成颗粒包含固体支持物官能团,并且

其中所述支持物官能团和所述接头官能团彼此相关联。

26. 根据权利要求25所述的合成颗粒,其中所述接头官能团和所述支持物官能团单独地选自下组,该组由以下各项组成:C6、生物素、链霉亲和素、一个或多个伯胺、一个或多个醛、一个或多个酮、及其任何组合。

27. 一种试剂盒,其包含:

权利要求1-8中任一项所述的多个合成颗粒;以及
用于使用所述多个合成颗粒的说明书。

28. 根据权利要求27所述的试剂盒,其中所述多个合成颗粒包含至少94个合成颗粒。

29. 根据权利要求27所述的试剂盒,其包含多个分区,每个分区包含所述多个合成颗粒中的合成颗粒。

30. 根据权利要求29所述的试剂盒,其中所述多个分区包含微孔阵列的微孔。

31. 根据权利要求29所述的试剂盒,其中所述多个分区包含多个液滴。

32. 根据权利要求27所述的试剂盒,其包含缓冲液。

33. 根据权利要求27所述的试剂盒,其包含盒。

34. 根据权利要求27所述的试剂盒,其包含用于逆转录反应的一种或多种试剂。
35. 根据权利要求27所述的试剂盒,其包含用于扩增反应的一种或多种试剂。
36. 根据权利要求27所述的试剂盒,其中每个合成颗粒包含第一多个间隔子寡核苷酸,其中所述第一多个间隔子寡核苷酸中的每个不包含靶所述结合区和所述靶特异性靶结合区。
37. 根据权利要求36所述的试剂盒,其中每个合成颗粒包含第二多个间隔子寡核苷酸,其中所述第二多个间隔子寡核苷酸中的每个不包含所述靶结合区和所述靶特异性靶结合区。
38. 根据权利要求27所述的试剂盒,其中所述第一多个寡核苷酸条形码中的至少一个比所述第二多个寡核苷酸条形码中的至少一个更短。
39. 根据权利要求27所述的试剂盒,其中所述第二多个寡核苷酸条形码中的至少一个比所述第一多个寡核苷酸条形码中的至少一个更短。
40. 根据权利要求27所述的试剂盒,其中所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少100个包含不同的分子标记序列。
41. 根据权利要求27所述的试剂盒,其中所述第二多个寡核苷酸条形码中的两个寡核苷酸条形码的两个靶特异性靶结合区包含不同的靶特异性序列。
42. 根据权利要求27所述的试剂盒,其中所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的每个包含空间标记序列,并且其中空间标记序列彼此差异至少一个核苷酸。
43. 根据权利要求27所述的试剂盒,其中所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的每个包含通用序列,其中通用序列是相同的。
44. 根据权利要求27所述的试剂盒,其中所述多个合成颗粒寡核苷酸中的至少一个固定在所述合成颗粒上。
45. 根据权利要求27所述的试剂盒,其中所述第一多个寡核苷酸条形码中的至少一个部分固定在所述合成颗粒上。
46. 根据权利要求27所述的试剂盒,其中所述第一多个寡核苷酸条形码中的至少一个包封在所述合成颗粒中。
47. 根据权利要求27所述的试剂盒,其中所述第一多个寡核苷酸条形码中的至少一个部分包封在所述合成颗粒中。
48. 根据权利要求27所述的试剂盒,其中所述合成颗粒是可破坏的。
49. 根据权利要求27所述的试剂盒,其中所述合成颗粒包含珠。
50. 根据权利要求49所述的试剂盒,其中所述珠包含磁珠。
51. 根据权利要求27所述的试剂盒,其中所述合成颗粒包含可破坏的水凝胶颗粒。
52. 根据权利要求27所述的试剂盒,
其中所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的每个包含接头官能团,
其中所述合成颗粒包含固体支持物官能团,并且
其中所述支持物官能团和所述接头官能团彼此相关联。
53. 根据权利要求52所述的试剂盒,其中所述接头官能团和所述支持物官能团单独地

选自下组,该组由以下各项组成:C6、生物素、链霉亲和素、一个或多个伯胺、一个或多个醛、一个或多个酮、及其任何组合。

54.一种用于确定样品中的靶核酸的数量的方法,所述方法包括:

使用与权利要求1-8中任一项所述的合成颗粒相关联的第一多个寡核苷酸条形码,对样品中第一多个靶核酸的拷贝进行条形码化,以产生第一多个经条形码化的靶核酸的拷贝;

使用与所述合成颗粒相关联的第二多个寡核苷酸条形码,对所述样品中第二多个靶核酸的拷贝进行条形码化,以产生第二多个经条形码化的靶核酸的拷贝;以及

使用所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码的分子标记序列评估所述第一多个靶核酸和所述第二多个靶核酸中的每个的数量,

所述方法不用于疾病诊断。

55.根据权利要求54所述的方法,其中所述第一多个靶核酸包含细胞的mRNA分子,并且其中所述第二多个靶核酸包含所述细胞的DNA分子。

56.根据权利要求54所述的方法,其中所述第一多个靶核酸包含细胞的mRNA分子,并且其中所述第二多个靶核酸包含感染所述细胞的病毒、细菌和/或真菌的RNA分子。

57.根据权利要求54所述的方法,其中使用与所述合成颗粒相关联的所述第一多个寡核苷酸条形码对所述样品中所述第一多个靶核酸的拷贝进行条形码化包括:将所述第一多个靶核酸与所述合成颗粒上的所述第一多个寡核苷酸条形码杂交,其中使用与所述合成颗粒相关联的所述第二多个寡核苷酸条形码对所述样品中所述第二多个靶核酸的拷贝进行条形码化包括:将所述第二多个靶核酸与所述合成颗粒上的所述第二多个寡核苷酸条形码杂交。

58.根据权利要求54所述的方法,其中对所述第一多个靶核酸的拷贝进行条形码化包括:逆转录所述第一多个靶核酸以产生所述第一多个靶核酸的经条形码化的cDNA,并且其中对所述第二多个靶核酸的拷贝进行条形码化包括:逆转录所述第二多个靶核酸以产生所述第二多个靶核酸的经条形码化的cDNA。

59.根据权利要求58所述的方法,所述方法包括:使用第二链合成,扩增所述第一多个靶核酸的经条形码化的cDNA和/或扩增所述第二多个靶核酸的经条形码化的cDNA,以分别产生第一经双链化的经条形码化的靶核酸和第二经双链化的经条形码化的靶核酸。

60.根据权利要求59所述的方法,所述方法包括:扩增所述第一经双链化的经条形码化的靶核酸和所述第二经双链化的经条形码化的靶核酸,以分别产生第一经标记的靶扩增子和第二经标记的靶扩增子。

61.一种用于捕获样品中的靶核酸的方法,所述方法包括:

使用与权利要求1-8中任一项所述的合成颗粒相关联的第一多个寡核苷酸条形码,对样品中第一多个靶核酸的拷贝进行杂交,以产生第一多个经条形码化的靶核酸的拷贝;

使用与所述合成颗粒相关联的第二多个寡核苷酸条形码,对第二多个靶核酸的拷贝进行杂交,以产生第二多个经条形码化的靶核酸的拷贝,

所述方法不用于疾病诊断。

62.根据权利要求61所述的方法,其中所述第二多个经条形码化的靶核酸的拷贝包含所述第二多个靶核酸中靶核酸的至少50%的拷贝。

63. 根据权利要求61所述的方法,其中所述第二多个经条形码化的靶核酸的拷贝包含所述第二多个靶核酸中靶核酸的至少75%的拷贝。

64. 根据权利要求61所述的方法,其中所述第二多个经条形码化的靶核酸的拷贝包含所述第二多个靶核酸中靶核酸的至少95%的拷贝。

65. 根据权利要求61所述的方法,其中所述第二多个经条形码化的靶核酸的拷贝包含所述第二多个靶核酸中各个靶核酸的至少50%的拷贝。

66. 根据权利要求61所述的方法,其中所述第二多个经条形码化的靶核酸的拷贝包含所述多个靶核酸中各个靶核酸的至少75%的拷贝。

67. 根据权利要求61所述的方法,其中所述第二多个经条形码化的靶核酸的拷贝包含所述第二多个靶核酸中各个靶核酸的至少95%的拷贝。

68. 根据权利要求61所述的方法,其中所述第一多个靶核酸和所述第二多个靶核酸包含mRNA,所述方法包括:

逆转录所述第一多个靶核酸,以产生所述第一多个靶核酸的第一经条形码化的cDNA;并且

逆转录所述第二多个靶核酸,以产生所述第二多个靶核酸的第二经条形码化的cDNA。

69. 根据权利要求68所述的方法,所述方法包括:使用第二链合成,扩增所述第一经条形码化的cDNA和所述第二经条形码化的cDNA,以分别产生第一经条形码化的双链靶和第二经条形码化的双链靶。

70. 根据权利要求69所述的方法,所述方法包括:扩增所述第一经条形码化的靶核酸和所述第二经条形码化的靶核酸,以分别产生第一经标记的靶扩增子和第二经标记的靶扩增子。

71. 根据权利要求54或61所述的方法,其中所述第一多个靶核酸和所述第二多个靶核酸来自单个细胞,所述方法包括将所述合成颗粒与所述单个细胞相关联。

72. 根据权利要求71所述的方法,其中所述单个细胞包括稀有细胞、肿瘤细胞、被病毒多核苷酸感染的细胞、或其任何组合。

73. 根据权利要求71所述的方法,其中所述单个细胞包括来自人的细胞。

74. 根据权利要求71所述的方法,其中所述单个细胞包括来自组织的细胞。

75. 根据权利要求71所述的方法,其中所述单个细胞包括来自肿瘤细胞的细胞。

76. 根据权利要求71所述的方法,所述方法包括:在将所述合成颗粒与所述单个细胞相关联之后,裂解所述单个细胞。

77. 根据权利要求76所述的方法,其中裂解所述单个细胞包括将所述样品加热、使所述样品与洗涤剂接触、改变所述样品的pH、或其任何组合。

78. 根据权利要求71所述的方法,其中所述合成颗粒和所述单个细胞在相同的孔中。

79. 根据权利要求71所述的方法,其中所述合成颗粒和所述单个细胞在相同的液滴中。

80. 根据权利要求54或61所述的方法,其中所述第一多个靶核酸和/或所述第二多个靶核酸包含核糖核酸(RNA)、RNA降解产物、及其任何组合。

81. 根据权利要求54或61所述的方法,其中所述第一多个靶核酸和/或所述第二多个靶核酸包含信使RNA(mRNA)、微小RNA、小干扰RNA(siRNA)、病毒RNA、及其任何组合。

82. 根据权利要求54或61所述的方法,其中所述第一多个靶核酸和/或所述第二多个靶

核酸包含各自含有聚(A)尾的RNA或不含有任何聚(A)尾的RNA。

83. 根据权利要求54或61所述的方法,其中所述合成颗粒包含第一多个间隔子寡核苷酸,其中所述第一多个间隔子寡核苷酸中的每个不包含所述靶结合区和所述靶特异性靶结合区。

84. 根据权利要求83所述的方法,所述合成颗粒包含第二多个间隔子寡核苷酸,其中所述第二多个间隔子寡核苷酸中的每个不包含所述靶结合区和所述靶特异性靶结合区。

85. 根据权利要求54或61所述的方法,其中所述第一多个寡核苷酸条形码中的至少一个比所述第二多个寡核苷酸条形码中的至少一个更短。

86. 根据权利要求54或61所述的方法,其中所述第二多个寡核苷酸条形码中的至少一个比所述第一多个寡核苷酸条形码中的至少一个更短。

87. 根据权利要求54或61所述的方法,其中所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少100个包含不同的分子标记序列。

88. 根据权利要求54或61所述的方法,其中所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的每个包含空间标记序列,并且其中空间标记序列彼此差异至少一个核苷酸。

89. 根据权利要求54或61所述的方法,其中所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的每个包含通用序列,其中通用序列是相同的。

90. 根据权利要求54或61所述的方法,其中所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少一个固定在所述合成颗粒上。

91. 根据权利要求54或61所述的方法,其中所述第一多个寡核苷酸条形码中的至少一个部分固定在所述合成颗粒上。

92. 根据权利要求54或61所述的方法,其中所述第一多个寡核苷酸条形码中的至少一个包封在所述合成颗粒中。

93. 根据权利要求54或61所述的方法,其中所述第一多个寡核苷酸条形码中的至少一个部分包封在所述合成颗粒中。

94. 根据权利要求54或61所述的方法,其中所述合成颗粒是可破坏的。

95. 根据权利要求54或61所述的方法,其中所述合成颗粒包含珠。

96. 根据权利要求95所述的方法,其中所述珠包含磁珠。

97. 根据权利要求54或61所述的方法,其中所述合成颗粒包含可破坏的水凝胶颗粒。

98. 根据权利要求54或61所述的方法,

其中所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的每个包含接头官能团,

其中所述合成颗粒包含固体支持物官能团,并且

其中所述支持物官能团和所述接头官能团彼此相关联。

99. 根据权利要求98所述的方法,其中所述接头官能团和所述支持物官能团单独地选自下组,该组由以下各项组成:C6、生物素、链霉亲和素、一个或多个伯胺、一个或多个醛、一个或多个酮、及其任何组合。

100. 一种合成颗粒,其包含:

与所述合成颗粒偶联的多个寡核苷酸条形码,其中所述多个寡核苷酸条形码中的每个

包含相同的细胞标记序列、分子标记序列和与样品中的靶核酸杂交的靶结合区,并且其中所述多个寡核苷酸条形码中的至少10个包含不同的分子标记序列,其中每个细胞标记序列包含至少6个核苷酸,其中每个分子标记序列包含至少6个核苷酸;以及

与所述合成颗粒偶联的第一多个间隔子寡核苷酸,其中所述第一多个间隔子寡核苷酸中的每个不包含所述靶结合区,其中所述多个寡核苷酸条形码与所述第一多个间隔子寡核苷酸是不同的分子,并且其中与仅与所述多个寡核苷酸条形码偶联的合成颗粒相比,与所述合成颗粒偶联的所述第一多个间隔子寡核苷酸降低所述合成颗粒上的所述多个寡核苷酸条形码的密度,其中所述多个寡核苷酸条形码和所述第一多个间隔子寡核苷酸的比率是在从1:10至10:1的范围内。

101. 根据权利要求100所述的合成颗粒,其中所述第一多个间隔子寡核苷酸中的每个包含细胞标记序列或分子标记序列。

102. 根据权利要求100所述的合成颗粒,其中所述第一多个间隔子寡核苷酸中的至少一个比所述多个寡核苷酸条形码中的至少一个更短。

103. 根据权利要求100所述的合成颗粒,其中所述第一多个间隔子寡核苷酸中的每个包含与所述靶结合区相对应的区,并且其中与所述靶结合区相对应的所述区和所述靶结合区具有不同的序列。

104. 根据权利要求103所述的合成颗粒,其中与所述靶结合区相对应的所述区和所述靶结合区具有至多50%序列同一性。

105. 根据权利要求100所述的合成颗粒,其中所述第一多个间隔子寡核苷酸中的至少两个具有不同长度。

106. 根据权利要求100所述的合成颗粒,其中所述第一多个间隔子寡核苷酸中的至少两个的序列是相同的。

107. 根据权利要求100所述的合成颗粒,其中所述第一多个间隔子寡核苷酸中的至少两个包含相同的序列。

108. 根据权利要求100-107中任一项所述的合成颗粒,其中所述合成颗粒包含第二多个间隔子寡核苷酸,其中所述第二多个间隔子寡核苷酸中的每个不包含所述靶结合区。

109. 根据权利要求108所述的合成颗粒,其中所述第二多个间隔子寡核苷酸中的至少一个比所述第一多个间隔子寡核苷酸中的至少一个更短。

110. 根据权利要求108所述的合成颗粒,其中所述第二多个间隔子寡核苷酸中的至少一个比所述多个寡核苷酸条形码中的至少一个更短。

111. 根据权利要求108所述的合成颗粒,其中所述第二多个间隔子寡核苷酸中的至少两个具有不同长度。

112. 根据权利要求108所述的合成颗粒,其中所述第二多个间隔子寡核苷酸中的至少两个的序列是相同的。

113. 根据权利要求108所述的合成颗粒,其中所述第二多个间隔子寡核苷酸中的至少两个包含相同的序列。

114. 根据权利要求100-107中任一项所述的合成颗粒,其中所述多个寡核苷酸条形码中的至少100个包含不同的分子标记序列。

115. 根据权利要求100-107中任一项所述的合成颗粒,其中所述靶结合区包含基因特

异性序列或寡聚 (dT) 序列。

116. 根据权利要求100-107中任一项所述的合成颗粒,其中所述多个寡核苷酸条形码中的两个寡核苷酸条形码的两个靶结合区包含不同的基因特异性序列。

117. 根据权利要求100-107中任一项所述的合成颗粒,其中所述多个寡核苷酸条形码中的每个包含空间标记序列,并且其中所述多个寡核苷酸条形码的空间标记序列彼此差异至少一个核苷酸。

118. 根据权利要求100-107中任一项所述的合成颗粒,其中所述多个寡核苷酸条形码中的每个包含通用标记序列,其中所述多个寡核苷酸条形码的通用标记序列是相同的。

119. 根据权利要求100-107中任一项所述的合成颗粒,其中所述多个寡核苷酸条形码中的至少一个固定在所述合成颗粒上。

120. 根据权利要求100-107中任一项所述的合成颗粒,其中所述多个寡核苷酸条形码中的至少一个部分固定在所述合成颗粒上。

121. 根据权利要求100-107中任一项所述的合成颗粒,其中所述多个寡核苷酸条形码中的至少一个包封在所述合成颗粒中。

122. 根据权利要求100-107中任一项所述的合成颗粒,其中所述多个寡核苷酸条形码中的至少一个部分包封在所述合成颗粒中。

123. 根据权利要求100-107中任一项所述的合成颗粒,其中所述合成颗粒是可破坏的。

124. 根据权利要求100-107中任一项所述的合成颗粒,其中所述合成颗粒包含珠。

125. 根据权利要求124所述的合成颗粒,其中珠包括琼脂糖珠、磁珠、缀合珠、二氧化硅珠、抗生物素微珠、抗荧光染料微珠、或其任何组合。

126. 根据权利要求124所述的合成颗粒,其中珠包括琼脂糖凝胶珠、链霉亲和素蛋白珠、蛋白A缀合珠、蛋白G缀合珠、蛋白A/G缀合珠、蛋白L缀合珠、寡聚 (dT) 缀合珠、二氧化硅样珠、或其任何组合。

127. 根据权利要求100-107中任一项所述的合成颗粒,其中所述合成颗粒包括选自下组的材料,该组由以下组成:玻璃、水凝胶、顺磁物质、陶瓷、塑料、钛、乳胶、纤维素、尼龙、硅酮、及其任何组合。

128. 根据权利要求100-107中任一项所述的合成颗粒,其中所述合成颗粒包括选自下组的材料,该组由以下组成:聚二甲基硅氧烷 (PDMS)、聚苯乙烯、聚丙烯、琼脂糖、明胶、及其任何组合。

129. 根据权利要求100-107中任一项所述的合成颗粒,其中所述合成颗粒包括选自下组的材料,该组由以下组成:甲基苯乙烯、丙烯酸聚合物、琼脂糖凝胶、及其任何组合。

130. 根据权利要求100-107中任一项所述的合成颗粒,其中所述合成颗粒包含可破坏的水凝胶颗粒。

131. 根据权利要求100-107中任一项所述的合成颗粒,
其中所述多个寡核苷酸条形码中的每个包含接头官能团,
其中所述合成颗粒包含固体支持物官能团,并且
其中所述支持物官能团和所述接头官能团彼此相关联。

132. 根据权利要求131所述的合成颗粒,其中所述接头官能团和所述支持物官能团单独地选自下组,该组由以下各项组成:C6、生物素、链霉亲和素、一个或多个伯胺、一个或多

个醛、一个或多个酮、及其任何组合。

133. 一种试剂盒,其包含:

多个根据权利要求100-107中任一项所述的合成颗粒,;以及
用于使用所述多个寡核苷酸条形码的说明书。

134. 根据权利要求133所述的试剂盒,其中所述多个合成颗粒包含至少94个合成颗粒。

135. 根据权利要求133所述的试剂盒,其包含微孔阵列,其中所述微孔阵列的每个微孔包含所述多个合成颗粒中的合成颗粒。

136. 根据权利要求133所述的试剂盒,其包含缓冲液。

137. 根据权利要求133所述的试剂盒,其包含盒。

138. 根据权利要求133所述的试剂盒,其包含用于逆转录反应的一种或多种试剂。

139. 根据权利要求133所述的试剂盒,其包含用于扩增反应的一种或多种试剂。

140. 根据权利要求133所述的试剂盒,其包含第二多个间隔子寡核苷酸,其中所述第二多个间隔子寡核苷酸中的每个不包含所述靶结合区。

141. 根据权利要求140所述的试剂盒,其中所述第二多个间隔子寡核苷酸中的至少一个比所述第一多个间隔子寡核苷酸中的至少一个更短。

142. 根据权利要求140所述的试剂盒,其中所述第二多个间隔子寡核苷酸中的至少一个比所述多个寡核苷酸条形码中的至少一个更短。

143. 根据权利要求140所述的试剂盒,其中所述第二多个间隔子寡核苷酸中的至少两个具有不同长度。

144. 根据权利要求140所述的试剂盒,其中所述第二多个间隔子寡核苷酸中的至少两个的序列是相同的。

145. 根据权利要求140所述的试剂盒,其中所述第二多个间隔子寡核苷酸中的至少两个包含相同的序列。

146. 根据权利要求133所述的试剂盒,其中所述多个寡核苷酸条形码中的至少100个包含不同的分子标记序列。

147. 根据权利要求133所述的试剂盒,其中所述靶结合区包含基因特异性序列或寡聚(dT)序列。

148. 根据权利要求133所述的试剂盒,其中所述多个寡核苷酸条形码中的两个寡核苷酸条形码的两个靶结合区包含不同的基因特异性序列。

149. 根据权利要求133所述的试剂盒,其中所述多个寡核苷酸条形码中的每个包含空间标记序列,并且其中所述多个寡核苷酸条形码的空间标记序列彼此差异至少一个核苷酸。

150. 根据权利要求133所述的试剂盒,其中所述多个寡核苷酸条形码中的每个包含通用标记序列,其中所述多个寡核苷酸条形码的通用标记序列是相同的。

151. 根据权利要求133所述的试剂盒,其中所述多个寡核苷酸条形码中的至少一个固定在所述合成颗粒上。

152. 根据权利要求133所述的试剂盒,其中所述多个寡核苷酸条形码中的至少一个部分固定在所述合成颗粒上。

153. 根据权利要求133所述的试剂盒,其中所述多个寡核苷酸条形码中的至少一个包

封在所述合成颗粒中。

154. 根据权利要求133所述的试剂盒,其中所述多个寡核苷酸条形码中的至少一个部分包封在所述合成颗粒中。

155. 根据权利要求133所述的试剂盒,其中所述合成颗粒是可破坏的。

156. 根据权利要求133所述的试剂盒,其中所述合成颗粒包含珠。

157. 根据权利要求156所述的试剂盒,其中珠包括琼脂糖珠、磁珠、缀合珠、二氧化硅珠、抗生物素微珠、抗荧光染料微珠、或其任何组合。

158. 根据权利要求156所述的试剂盒,其中珠包括琼脂糖凝胶珠、链霉亲和素蛋白珠、蛋白A缀合珠、蛋白G缀合珠、蛋白A/G缀合珠、蛋白L缀合珠、寡聚(dT)缀合珠、二氧化硅样珠、或其任何组合。

159. 根据权利要求133所述的试剂盒,其中所述合成颗粒包括选自下组的材料,该组由以下组成:玻璃、水凝胶、顺磁物质、陶瓷、塑料、钛、乳胶、纤维素、尼龙、硅酮、及其任何组合。

160. 根据权利要求133所述的试剂盒,其中所述合成颗粒包括选自下组的材料,该组由以下组成:聚二甲基硅氧烷(PDMS)、聚苯乙烯、聚丙烯、琼脂糖、明胶、及其任何组合。

161. 根据权利要求133所述的试剂盒,其中所述合成颗粒包括选自下组的材料,该组由以下组成:甲基苯乙烯、丙烯酸聚合物、琼脂糖凝胶、及其任何组合。

162. 根据权利要求133所述的试剂盒,其中所述合成颗粒包含可破坏的水凝胶颗粒。

163. 根据权利要求133所述的试剂盒

其中所述多个寡核苷酸条形码中的每个包含接头官能团,

其中所述合成颗粒包含固体支持物官能团,并且

其中所述支持物官能团和所述接头官能团彼此相关联。

164. 根据权利要求163所述的试剂盒,其中所述接头官能团和所述支持物官能团单独地选自下组,该组由以下各项组成:C6、生物素、链霉亲和素、一个或多个伯胺、一个或多个醛、一个或多个酮、及其任何组合。

165. 一种用于确定样品中的靶核酸的数量的方法,所述方法包括:

使用与权利要求100-107中任一项所述的合成颗粒相关联的多个寡核苷酸条形码,对样品中多个靶核酸的拷贝进行条形码化,以产生多个经条形码化的靶核酸的拷贝,

并且

使用所述多个寡核苷酸条形码的分子标记序列,对所述样品中所述多个靶核酸中每个的数量进行评估,

所述方法不用于疾病诊断。

166. 根据权利要求165所述的方法,其中使用与所述合成颗粒相关联的所述多个寡核苷酸条形码对所述样品中所述多个靶核酸的拷贝进行条形码化包括:将所述多个靶核酸与所述合成颗粒上的所述多个寡核苷酸条形码杂交。

167. 根据权利要求165所述的方法,其中对所述多个靶核酸的拷贝进行条形码化包括:逆转录所述多个靶核酸,以产生所述多个靶核酸的经条形码化的cDNA。

168. 根据权利要求167所述的方法,所述方法包括:使用第二链合成,扩增所述经条形码化的cDNA,以产生经条形码化的靶核酸,其中所述经条形码化的靶核酸是双链的。

169. 根据权利要求168所述的方法,所述方法包括:扩增所述经条形码化的靶核酸,以产生经标记的靶扩增子。

170. 一种用于捕获样品中的靶核酸的方法,所述方法包括:

使用与权利要求100-107中任一项所述的合成颗粒相关联的多个寡核苷酸条形码,对样品中多个靶核酸的拷贝进行杂交,以产生多个经条形码化的靶核酸的拷贝,所述方法不用于疾病诊断。

171. 根据权利要求170所述的方法,其中所述多个经条形码化的靶核酸的拷贝包含所述样品中所述多个靶核酸中靶核酸的至少50%的拷贝。

172. 根据权利要求170所述的方法,其中所述多个经条形码化的靶核酸的拷贝包含所述样品中所述多个靶核酸中靶核酸的至少75%的拷贝。

173. 根据权利要求170所述的方法,其中所述多个经条形码化的靶核酸的拷贝包含所述样品中所述多个靶核酸中靶核酸的至少95%的拷贝。

174. 根据权利要求170所述的方法,其中所述多个经条形码化的靶核酸的拷贝包含所述样品中所述多个靶核酸中各个靶核酸的至少50%的拷贝。

175. 根据权利要求170所述的方法,其中所述多个经条形码化的靶核酸的拷贝包含所述样品中所述多个靶核酸中各个靶核酸的至少75%的拷贝。

176. 根据权利要求170所述的方法,其中所述多个经条形码化的靶核酸的拷贝包含所述样品中所述多个靶核酸中各个靶核酸的至少95%的拷贝。

177. 根据权利要求170所述的方法,所述方法包括:逆转录所述多个靶核酸,以产生所述多个靶核酸的经条形码化的cDNA。

178. 根据权利要求177所述的方法,所述方法包括:使用第二链合成,扩增所述经条形码化的cDNA,以产生经条形码化的靶核酸,其中所述经条形码化的靶核酸是双链的。

179. 根据权利要求178所述的方法,所述方法包括:扩增所述经条形码化的靶核酸,以产生经标记的靶扩增子。

180. 根据权利要求165或170所述的方法,其中所述样品包含单个细胞,所述方法包括将所述合成颗粒与所述样品中的所述单个细胞相关联。

181. 根据权利要求180所述的方法,其中所述单个细胞包括稀有细胞、肿瘤细胞、被病毒多核苷酸感染的细胞、及其任何组合。

182. 根据权利要求180所述的方法,其中所述单个细胞包括来自人的细胞。

183. 根据权利要求180所述的方法,其中所述单个细胞包括来自组织的细胞。

184. 根据权利要求180所述的方法,其中所述单个细胞包括来自肿瘤细胞的细胞。

185. 根据权利要求180所述的方法,所述方法包括:在将所述合成颗粒与所述单个细胞相关联之后,裂解所述单个细胞。

186. 根据权利要求185所述的方法,其中裂解所述单个细胞包括将所述样品加热、使所述样品与洗涤剂接触、改变所述样品的pH、或其任何组合。

187. 根据权利要求180所述的方法,其中所述合成颗粒和所述单个细胞在相同的孔中。

188. 根据权利要求180所述的方法,其中所述合成颗粒和所述单个细胞在相同的液滴中。

189. 根据权利要求165或170所述的方法,其中所述多个靶核酸包含核糖核酸(RNA)、

RNA降解产物、及其任何组合。

190. 根据权利要求165或170所述的方法,其中所述多个靶核酸包含信使RNA(mRNA)、微小RNA、小干扰RNA(siRNA)、及其任何组合。

191. 根据权利要求165或170所述的方法,其中所述多个靶核酸包含各自含有聚(A)尾的RNA。

192. 根据权利要求165或170所述的方法,其中所述合成颗粒包含第二多个间隔子寡核苷酸,其中所述第二多个间隔子寡核苷酸中的每个不包含所述靶结合区。

193. 根据权利要求192所述的方法,其中所述第二多个间隔子寡核苷酸中的至少一个比所述第一多个间隔子寡核苷酸中的至少一个更短。

194. 根据权利要求192所述的方法,其中所述第二多个间隔子寡核苷酸中的至少一个比所述多个寡核苷酸条形码中的至少一个更短。

195. 根据权利要求192所述的方法,其中所述第二多个间隔子寡核苷酸中的至少两个具有不同长度。

196. 根据权利要求192所述的方法,其中所述第二多个间隔子寡核苷酸中的至少两个的序列是相同的。

197. 根据权利要求192所述的方法,其中所述第二多个间隔子寡核苷酸中的至少两个包含相同的序列。

198. 根据权利要求165或170所述的方法,其中所述多个寡核苷酸条形码中的至少100个包含不同的分子标记序列。

199. 根据权利要求165或170所述的方法,其中所述靶结合区包含基因特异性序列或寡聚(dT)序列。

200. 根据权利要求165或170所述的方法,其中所述多个寡核苷酸条形码中的两个寡核苷酸条形码的两个靶结合区包含不同的基因特异性序列。

201. 根据权利要求165或170所述的方法,其中所述多个寡核苷酸条形码中的每个包含空间标记序列,并且其中所述多个寡核苷酸条形码的空间标记序列彼此差异至少一个核苷酸。

202. 根据权利要求165或170所述的方法,其中所述多个寡核苷酸条形码中的每个包含通用标记序列,其中所述多个寡核苷酸条形码的通用标记序列是相同的。

203. 根据权利要求165或170所述的方法,其中所述多个寡核苷酸条形码中的至少一个固定在所述合成颗粒上。

204. 根据权利要求165或170所述的方法,其中所述多个寡核苷酸条形码中的至少一个部分固定在所述合成颗粒上。

205. 根据权利要求165或170所述的方法,其中所述多个寡核苷酸条形码中的至少一个包封在所述合成颗粒中。

206. 根据权利要求165或170所述的方法,其中所述多个寡核苷酸条形码中的至少一个部分包封在所述合成颗粒中。

207. 根据权利要求165或170所述的方法,其中所述合成颗粒是可破坏的。

208. 根据权利要求165或170所述的方法,其中所述合成颗粒包含珠。

209. 根据权利要求208所述的方法,其中珠包括琼脂糖珠、磁珠、缀合珠、二氧化硅珠、

抗生物素微珠、抗荧光染料微珠、或其任何组合。

210. 根据权利要求208所述的方法,其中珠包括琼脂糖凝胶珠、链霉亲和素蛋白珠、蛋白A缀合珠、蛋白G缀合珠、蛋白A/G缀合珠、蛋白L缀合珠、寡聚(dT)缀合珠、二氧化硅样珠、或其任何组合。

211. 根据权利要求165或170所述的方法,其中所述合成颗粒包括选自下组的材料,该组由以下组成:玻璃、水凝胶、顺磁物质、陶瓷、塑料、钛、乳胶、纤维素、尼龙、硅酮、及其任何组合。

212. 根据权利要求165或170中任一项所述的合成颗粒,其中所述合成颗粒包括选自下组的材料,该组由以下组成:聚二甲基硅氧烷(PDMS)、聚苯乙烯、聚丙烯、琼脂糖、明胶、及其任何组合。

213. 根据权利要求165或170中任一项所述的合成颗粒,其中所述合成颗粒包括选自下组的材料,该组由以下组成:甲基苯乙烯、丙烯酸聚合物、琼脂糖凝胶、及其任何组合。

214. 根据权利要求165或170所述的方法,其中所述合成颗粒包含可破坏的水凝胶颗粒。

215. 根据权利要求165或170所述的方法,
其中所述多个寡核苷酸条形码中的每个包含接头官能团,
其中所述合成颗粒包含固体支持物官能团,并且
其中所述支持物官能团和所述接头官能团彼此相关联。

216. 根据权利要求215所述的方法,其中所述接头官能团和所述支持物官能团单独地选自下组,该组由以下各项组成:C6、生物素、链霉亲和素、一个或多个伯胺、一个或多个醛、一个或多个酮、及其任何组合。

217. 根据权利要求1-8中任一项所述的合成颗粒,其中所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码的比率是在从1:3至2:1的范围内。

218. 根据权利要求1-8中任一项所述的合成颗粒,其中所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码的比率是至多2:1。

219. 根据权利要求27所述的试剂盒,其中所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码的比率是在从1:3至2:1的范围内。

220. 根据权利要求27所述的试剂盒,其中所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码的比率是至多2:1。

221. 根据权利要求54所述的方法,其中所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码的比率是在从1:3至2:1的范围内。

222. 根据权利要求54所述的方法,其中所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码的比率是至多2:1。

223. 根据权利要求61所述的方法,其中所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码的比率是在从1:3至2:1的范围内。

224. 根据权利要求61所述的方法,其中所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码的比率是至多2:1。

225. 根据权利要求100-107中任一项所述的合成颗粒,其中所述多个寡核苷酸条形码和所述第一多个间隔子寡核苷酸的比率是在从1:3至2:1的范围内。

226. 根据权利要求100-107中任一项所述的合成颗粒, 其中所述多个寡核苷酸条形码和所述第一多个间隔子寡核苷酸的比率是至少1:3。

227. 根据权利要求133所述的试剂盒, 其中所述多个寡核苷酸条形码和所述第一多个间隔子寡核苷酸的比率是在从1:3至2:1的范围内。

228. 根据权利要求133所述的试剂盒, 其中所述多个寡核苷酸条形码和所述第一多个间隔子寡核苷酸的比率是至少1:3。

229. 根据权利要求165所述的方法, 其中所述多个寡核苷酸条形码和所述第一多个间隔子寡核苷酸的比率是在从1:3至2:1的范围内。

230. 根据权利要求165所述的方法, 其中所述多个寡核苷酸条形码和所述第一多个间隔子寡核苷酸的比率是至少1:3。

231. 根据权利要求170所述的方法, 其中所述多个寡核苷酸条形码和所述第一多个间隔子寡核苷酸的比率是在从1:3至2:1的范围内。

232. 根据权利要求170所述的方法, 其中所述多个寡核苷酸条形码和所述第一多个间隔子寡核苷酸的比率是至少1:3。

与寡核苷酸相关联的颗粒

[0001] 相关申请的交叉引用

[0002] 本申请根据35 U.S.C. §119(e) 要求于2017年12月19日提交的美国临时专利号62/607,812的优先权,其内容通过引用以其全文并入本文。

[0003] 序列表的引用

[0004] 本申请是连同电子格式的序列表一起提交的。序列表被提供为题为Sequence_Listing_BDCRI_030W0.txt的文件,创建于2018年12月18日,大小是2.1千字节。将电子格式的序列表的信息通过引用以其全文并入本文。

[0005] 发明背景

技术领域

[0006] 本公开总体上涉及靶的分子条形码化的领域,并更具体地涉及靶捕获。

背景技术

[0007] 诸如使用与核酸条形码偶联的珠的条形码化的方法和技术对于细胞分析有用,特别是使用例如逆转录、聚合酶链式反应(PCR)扩增、和下一代测序(NGS)以解密基因表达谱来确定细胞的状态。然而,可以改善这些方法和技术的敏感性。

发明内容

[0008] 本文的披露内容包括合成颗粒的实施例。在一些实施例中,所述合成颗粒包含:第一多个寡核苷酸条形码;和第二多个寡核苷酸条形码,其中所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码中的每个包含相同的细胞标记序列并且包含分子标记序列,其中所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码中的至少10个包含不同的分子标记序列,其中所述第一多个寡核苷酸条形码中的每个包含靶结合区,并且其中所述第二多个寡核苷酸条形码中的每个包含靶特异性靶结合区。

[0009] 在一些实施例中,所述第一多个寡核苷酸条形码的寡核苷酸条形码的靶结合区包含靶特异性靶结合区。在一些实施例中,所述第一多个寡核苷酸条形码的寡核苷酸条形码的靶结合区包含聚(dT)区。所述第二多个寡核苷酸条形码的寡核苷酸条形码可以包含聚(dT)区。

[0010] 在一些实施例中,所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码的比率是在从1:10至10:1的范围内。所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的每个可以包含样品标记序列。所述第二多个寡核苷酸条形码中的至少一个比所述第一多个寡核苷酸条形码中的至少一个更短。所述靶结合区和所述靶特异性靶结合区可以具有至多50%序列同一性。所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少两个可以具有不同长度。所述第二多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少两个的序列可以是相同的。所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少两个可以包含相同的序列。

[0011] 在一些实施例中,合成颗粒包含第一多个间隔物寡核苷酸,其中所述第一多个间隔子寡核苷酸中的每个不包含靶结合区和靶特异性靶结合区。合成颗粒可以包含第二多个间隔子寡核苷酸,其中所述第二多个间隔子寡核苷酸中的每个不包含靶结合区和靶特异性靶结合区。

[0012] 在一些实施例中,所述第一多个间隔子寡核苷酸中的至少一个比所述第二多个间隔子寡核苷酸中的至少一个更短。所述第二多个间隔子寡核苷酸中的至少一个可以比所述第一多个间隔子寡核苷酸中的至少一个更短。所述第一多个间隔子寡核苷酸和/或所述第二多个间隔子寡核苷酸中的至少两个可以具有不同长度。所述第一多个间隔子寡核苷酸和/或所述第二多个间隔子寡核苷酸中的至少两个的序列可以是相同的。所述第一多个间隔子寡核苷酸和/或所述第二多个间隔子寡核苷酸中的至少两个可以包含相同的序列。

[0013] 在一些实施例中,所述第一多个寡核苷酸条形码中的至少一个比第二多个寡核苷酸条形码中的至少一个更短。所述第二多个寡核苷酸条形码中的至少一个可以比所述第一多个寡核苷酸条形码中的至少一个更短。所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少两个可以具有不同长度。所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少两个的序列可以是相同的。所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少两个可以包含相同的序列。

[0014] 在一些实施例中,所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少100个包含不同的分子标记序列。靶结合区可以包含靶特异性序列、寡聚(dT)序列、随机多聚体、或其任何组合。所述第二多个寡核苷酸条形码中的两个寡核苷酸条形码的两个靶特异性靶结合区可以包含不同的靶特异性序列。每个细胞标记序列可以包含至少6个核苷酸。每个分子标记序列可以包含至少6个核苷酸。所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的每个可以包含空间标记序列,并且其中所述空间标记序列彼此之间相差至少一个核苷酸。所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的每个可以包含通用序列,其中所述通用序列是相同的。

[0015] 在一些实施例中,所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少一个固定在合成颗粒上。所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少一个可以部分固定在合成颗粒上。所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少一个可以包封在合成颗粒中。所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少一个可以部分包封在合成颗粒中。合成颗粒可以是可破坏的。合成颗粒可以包含珠。珠可以包含磁珠。合成颗粒可以包含可破坏的水凝胶颗粒,例如可破坏的水凝胶珠。

[0016] 在一些实施例中,所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的每个可以包含接头官能团,合成颗粒可以包含固体支持物官能团,并且所述支持物官能团和所述接头官能团可以彼此相关联。所述接头官能团和所述支持物官能团可以单独地选自下组,该组由以下各项组成:C6、生物素、链霉亲和素、一个或多个伯胺、一个或多个醛、一个或多个酮、一个或多个硫化物及其任何组合。

[0017] 本文的披露内容包括试剂盒或组合物的实施例。在一些实施例中,所述试剂盒或组合物包含:多个合成颗粒,其中每个合成颗粒包含:第一多个寡核苷酸条形码;和第二多个寡核苷酸条形码,其中所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码中

的每个包含相同的细胞标记序列并且包含分子标记序列,其中所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码中的至少10个包含不同的分子标记序列,其中所述第一多个寡核苷酸条形码中的每个包含靶结合区,并且其中所述第二多个寡核苷酸条形码中的每个包含靶特异性靶结合区;以及任选地,用于使用所述多个合成颗粒的说明书。在一些实施例中,所述试剂盒包含缓冲液、用于逆转录反应的一种或多种试剂、用于扩增反应的一种或多种试剂或其组合。

[0018] 在一些实施例中,所述多个合成颗粒包含至少94个合成颗粒。所述试剂盒可以包含多个分区,每个分区包含所述多个合成颗粒中的合成颗粒。所述多个分区可以包含微孔阵列的微孔。所述多个分区可以包含多个液滴。

[0019] 在一些实施例中,所述第一多个寡核苷酸条形码的寡核苷酸条形码的靶结合区包含靶特异性靶结合区。所述第一多个寡核苷酸条形码的寡核苷酸条形码的靶结合区可以包含聚(dT)区。所述第二多个寡核苷酸条形码的寡核苷酸条形码可以包含聚(dT)区。

[0020] 在一些实施例中,所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码的比率是在从1:10至10:1的范围内。所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的每个可以包含样品标记序列。所述第二多个寡核苷酸条形码中的至少一个可以比所述第一多个寡核苷酸条形码中的至少一个更短。所述靶结合区和所述序列特异性靶结合区可以具有至多50%序列同一性。所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少两个可以具有不同长度。所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少两个的序列可以是相同的。所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少两个可以包含相同的序列。

[0021] 在一些实施例中,每个合成颗粒包含第一多个间隔物寡核苷酸,其中所述第一多个间隔子寡核苷酸中的每个不包含靶结合区和靶特异性靶结合区。每个合成颗粒可以包含第二多个间隔子寡核苷酸,其中所述第二多个间隔子寡核苷酸中的每个不包含靶结合区和靶特异性靶结合区。

[0022] 在一些实施例中,所述第一多个间隔子寡核苷酸中的至少一个比所述第二多个间隔子寡核苷酸中的至少一个更短。所述第二多个间隔子寡核苷酸中的至少一个可以比所述第一多个间隔子寡核苷酸中的至少一个更短。所述第一多个间隔子寡核苷酸和/或所述第二多个间隔子寡核苷酸中的至少两个可以具有不同长度。所述第一多个间隔子寡核苷酸和/或所述第二多个间隔子寡核苷酸中的至少两个的序列可以是相同的。所述第一多个间隔子寡核苷酸和/或所述第二多个间隔子寡核苷酸中的至少两个可以包含相同的序列。

[0023] 在一些实施例中,所述第一多个寡核苷酸条形码中的至少一个比第二多个寡核苷酸条形码中的至少一个更短。所述第二多个寡核苷酸条形码中的至少一个可以比所述第一多个寡核苷酸条形码中的至少一个更短。所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少两个可以具有不同长度。所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少两个的序列可以是相同的。所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少两个可以包含相同的序列。

[0024] 在一些实施例中,所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少100个包含不同的分子标记序列。靶结合区可以包含靶特异性序列、寡聚(dT)序列、随机多聚体、或其任何组合。所述第二多个寡核苷酸条形码中的两个寡核苷酸条形码

的两个靶特异性靶结合区可以包含不同的靶特异性序列。每个细胞标记序列可以包含至少6个核苷酸。每个分子标记序列可以包含至少6个核苷酸。所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的每个可以包含空间标记序列,并且其中空间标记序列彼此之间相差至少一个核苷酸。所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的每个可以包含通用序列,其中通用序列是相同的。

[0025] 在一些实施例中,所述多个合成颗粒寡核苷酸中的至少一个固定在合成颗粒上。所述第一多个寡核苷酸条形码中的至少一个可以部分固定在合成颗粒上。所述第一多个寡核苷酸条形码中的至少一个可以包封在合成颗粒中。所述第一多个寡核苷酸条形码中的至少一个可以部分包封在合成颗粒中。合成颗粒可以是可破坏的。合成颗粒可以包含珠。珠可以包含磁珠。合成颗粒可以包含可破坏的水凝胶颗粒,例如水凝胶珠。

[0026] 在一些实施例中,所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的每个包含接头官能团,所述合成颗粒包含固体支持物官能团,并且所述支持物官能团和所述接头官能团彼此相关联。所述接头官能团和所述支持物官能团可以单独地选自下组,该组由以下各项组成:C6、生物素、链霉亲和素、一个或多个伯胺、一个或多个醛、一个或多个酮、一个或多个硫化物及其任何组合。

[0027] 本文的披露内容包括如下方法的实施例,所述方法用于确定样品中靶的数量。在一些实施例中,所述方法包括:使用与合成颗粒相关联的第一多个寡核苷酸条形码,对第一多个靶的拷贝进行条形码化,以产生第一多个经条形码化的靶的拷贝;使用与所述合成颗粒相关联的第二多个寡核苷酸条形码,对样品中第二多个靶的拷贝进行条形码化,以产生第二多个经条形码化的靶的拷贝,其中所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码中的每个包含相同的细胞标记序列并且包含分子标记序列,其中所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码中的至少10个包含不同的分子标记序列,其中所述第一多个寡核苷酸条形码中的每个包含靶结合区,并且其中所述第二多个寡核苷酸条形码中的每个包含靶特异性靶结合区;并且使用所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码的分子标记序列对所述第一多个靶和/或所述第二多个靶中的每个的数量进行评估。

[0028] 在一些实施例中,所述第一多个寡核苷酸条形码的寡核苷酸条形码的靶结合区包含靶特异性靶结合区。所述第一多个寡核苷酸条形码的寡核苷酸条形码的靶结合区可以包含聚(dT)区。所述第二多个寡核苷酸条形码的寡核苷酸条形码可以包含聚(dT)区。

[0029] 在一些实施例中,第一多个靶包含细胞的mRNA分子,并且其中第二多个靶包含细胞的DNA分子。所述第一多个靶可以包含细胞的mRNA分子,并且所述第二多个靶可以包含感染细胞的病毒、细菌和/或真菌的RNA分子。

[0030] 在一些实施例中,使用与合成颗粒相关联的第一多个寡核苷酸条形码对样品中第一多个靶的拷贝进行条形码化包括:将所述第一多个靶与所述合成颗粒上的所述第一多个寡核苷酸条形码杂交,其中使用与所述合成颗粒相关联的所述第二多个寡核苷酸条形码对所述样品中所述第二多个靶的拷贝进行条形码化包括:将所述第二多个靶与所述合成颗粒上的所述第二多个寡核苷酸条形码杂交。对所述第一多个靶的拷贝进行条形码化可以包括:逆转录所述第一多个靶以产生所述第一多个靶的经条形码化的cDNA,并且其中对所述第二多个靶的拷贝进行条形码化包括:逆转录所述第二多个靶以产生所述第二多个靶的经

条形码化的cDNA。所述方法可以包括使用第二链合成,扩增所述第一多个靶的经条形码化的cDNA和/或扩增所述第二多个靶的经条形码化的cDNA,以分别产生第一经双链化的经条形码化的靶和第二经双链化的经条形码化的靶。所述方法可以包括扩增第一经双链化的经条形码化的靶和第二经双链化的经条形码化的靶,以分别产生第一经标记的靶扩增子和第二经标记的靶扩增子。

[0031] 本文的披露内容包括如下方法的实施例,所述方法用于捕获样品中的靶。在一些实施例中,所述方法包括:使用与合成颗粒相关联的第一多个寡核苷酸条形码,对第一多个靶的拷贝进行杂交,以产生第一多个经条形码化的靶的拷贝;使用与合成颗粒相关联的第二多个寡核苷酸条形码,对第二多个靶的拷贝进行杂交,以产生第二多个经条形码化的靶的拷贝,其中所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码中的每个包含相同的细胞标记序列并且包含分子标记序列,其中所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码中的至少10个包含不同的分子标记序列,其中所述第一多个寡核苷酸条形码中的每个包含靶结合区,并且其中所述第二多个寡核苷酸条形码中的每个包含靶特异性靶结合区。

[0032] 在一些实施例中,所述第二多个经条形码化的靶的拷贝包含所述多个靶中靶的至少50%的拷贝。所述第二多个经条形码化的靶的拷贝可以包含所述第二多个靶中靶的至少75%的拷贝。所述第二多个经条形码化的靶的拷贝可以包含所述第二多个靶中靶的至少95%的拷贝。所述第二多个经条形码化的靶的拷贝可以包含所述第二多个靶中每个靶的至少50%的拷贝。所述第二多个经条形码化的靶的拷贝可以包含所述第二多个靶中每个靶的至少75%的拷贝。所述第二多个经条形码化的靶的拷贝可以包含所述第二多个靶中每个靶的至少95%的拷贝。

[0033] 在一些实施例中,所述第一多个靶和所述第二多个靶包含mRNA,所述方法包括:逆转录所述第一多个靶,以产生所述第一多个靶的第一经条形码化的cDNA;并且逆转录所述第二多个靶,以产生所述第二多个靶的第二经条形码化的cDNA。所述方法可以包括使用第二链合成,扩增所述第一条形码化的cDNA和所述第二经条形码化的cDNA,以分别产生第一经双链化的经条形码化的靶和第二经双链化的经条形码化的靶。所述方法可以包括扩增所述第一经条形码化的靶和所述第二经条形码化的靶,以分别产生第一经标记的靶扩增子和第二经标记的靶扩增子。

[0034] 在一些实施例中,所述第一多个靶和所述第二多个靶来自单个细胞,所述方法包括将合成颗粒与单个细胞相关联。所述单个细胞可以包含稀有细胞、肿瘤细胞、来自人的细胞、来自组织的细胞、来自肿瘤细胞、被病毒多核苷酸感染的细胞、或其任何组合。

[0035] 在一些实施例中,所述方法包括在将所述合成颗粒与所述单个细胞相关联之后,裂解所述单个细胞。裂解所述单个细胞可以包括将样品加热、使样品与洗涤剂接触、改变样品的pH、或其任何组合。所述合成颗粒和所述单个细胞可以在相同的孔中。所述合成颗粒和所述单个细胞可以在相同的液滴中。

[0036] 在一些实施例中,所述第一多个靶和/或所述第二多个靶包括核糖核酸(RNA)、信使RNA(mRNA)、微小RNA、小干扰RNA(siRNA)、RNA降解产物、各自含有聚(A)尾的RNA、不含有任何聚(A)尾的RNA、病毒RNA、及其任何组合。

[0037] 在一些实施例中,所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码

的比率是在从1:10至10:1的范围内。所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的每个可以包含样品标记序列。所述第二多个寡核苷酸条形码中的至少一个可以比所述第一多个寡核苷酸条形码中的至少一个更短。所述靶结合区和所述靶特异性靶结合区可以具有至多50%序列同一性。所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少两个可以具有不同长度。所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少两个的序列是相同的。所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少两个可以包含相同的序列。

[0038] 在一些实施例中,合成颗粒包含第一多个间隔物寡核苷酸,其中所述第一多个间隔子寡核苷酸中的每个不包含靶结合区和靶特异性靶结合区。合成颗粒可以包含第二多个间隔子寡核苷酸,其中所述第二多个间隔子寡核苷酸中的每个不包含靶结合区和靶特异性靶结合区。

[0039] 在一些实施例中,所述第一多个间隔子寡核苷酸中的至少一个比所述第二多个间隔子寡核苷酸中的至少一个更短。所述第二多个间隔子寡核苷酸中的至少一个可以比所述第一多个间隔子寡核苷酸中的至少一个更短。所述第一多个间隔子寡核苷酸和/或所述第二多个间隔子寡核苷酸中的至少两个可以具有不同长度。所述第一多个间隔子寡核苷酸和/或所述第二多个间隔子寡核苷酸中的至少两个的序列可以是相同的。所述第一多个间隔子寡核苷酸和/或所述第二多个间隔子寡核苷酸中的至少两个可以包含相同的序列。

[0040] 在一些实施例中,所述第一多个寡核苷酸条形码中的至少一个比第二多个寡核苷酸条形码中的至少一个更短。所述第二多个寡核苷酸条形码中的至少一个可以比所述第一多个寡核苷酸条形码中的至少一个更短。所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少两个可以具有不同长度。所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少两个的序列可以是相同的。所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少两个可以包含相同的序列。

[0041] 在一些实施例中,所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少100个包含不同的分子标记序列。靶结合区可以包含靶特异性序列、寡聚(dT)序列、随机多聚体、或其任何组合。所述第二多个寡核苷酸条形码中的两个寡核苷酸条形码的两个靶结合区可以包含靶特异性序列。每个细胞标记序列可以包含至少6个核苷酸。每个分子标记序列可以包含至少6个核苷酸。所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的每个可以包含空间标记序列,并且其中空间标记序列彼此之间相差至少一个核苷酸。所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的每个可以包含通用序列,其中通用序列是相同的。

[0042] 在一些实施例中,所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的至少一个固定在合成颗粒上。所述第一多个寡核苷酸条形码中的至少一个可以部分固定在合成颗粒上。所述第一多个寡核苷酸条形码中的至少一个可以包封在合成颗粒中。所述第一多个寡核苷酸条形码中的至少一个可以部分包封在合成颗粒中。合成颗粒可以是可破坏的。合成颗粒可以包含珠。珠可以包含磁珠。合成颗粒可以包含可破坏的水凝胶颗粒,例如水凝胶珠。

[0043] 在一些实施例中,所述第一多个寡核苷酸条形码和/或所述第二多个寡核苷酸条形码中的每个包含接头官能团,所述合成颗粒包含固体支持物官能团,并且所述支持物官

能团和所述接头官能团彼此相关联。所述接头官能团和所述支持物官能团可以单独地选自下组,该组由以下各项组成:C6、生物素、链霉亲和素、一个或多个伯胺、一个或多个醛、一个或多个酮、及其任何组合。

[0044] 本文的披露内容包括合成颗粒的实施例。在一些实施例中,所述合成颗粒包含:多个寡核苷酸条形码,其中所述多个寡核苷酸条形码中的每个包含相同的细胞标记序列、分子标记序列和靶结合区,并且其中所述多个寡核苷酸条形码中的至少10个包含不同的分子标记序列;和第一多个间隔子寡核苷酸,其中所述第一多个间隔子寡核苷酸中的每个不包含靶结合区。所述多个寡核苷酸条形码和所述第一多个间隔子寡核苷酸的比率可以在从1:10至10:1的范围内。所述第一多个间隔子寡核苷酸中的每个可以包含细胞标记序列或分子标记序列。

[0045] 在一些实施例中,所述第一多个间隔子寡核苷酸中的至少一个可以比所述多个寡核苷酸条形码中的至少一个更短。所述第一多个间隔子寡核苷酸中的每个可以包含与所述靶结合区相对应的区,并且其中与所述靶结合区相对应的所述区和所述靶结合区具有不同的序列。与所述靶结合区相对应的所述区和所述靶结合区可以具有至多50%序列同一性。所述第一多个间隔子寡核苷酸中的至少两个可以具有不同长度。所述第一多个间隔子寡核苷酸中的至少两个的序列可以是相同的。所述第一多个间隔子寡核苷酸中的至少两个可以包含相同的序列。

[0046] 在一些实施例中,合成颗粒包含第二多个间隔子寡核苷酸,其中所述第二多个间隔子寡核苷酸中的每个不包含所述靶结合区。所述第二多个间隔子寡核苷酸中的至少一个可以比所述第一多个间隔子寡核苷酸中的至少一个更短。所述第二多个间隔子寡核苷酸中的至少一个可以比所述多个寡核苷酸条形码中的至少一个更短。所述第二多个间隔子寡核苷酸中的至少两个可以具有不同长度。所述第二多个间隔子寡核苷酸中的至少两个的序列可以是相同的。所述第二多个间隔子寡核苷酸中的至少两个包含相同的序列。

[0047] 在一些实施例中,所述多个寡核苷酸条形码中的至少100个包含不同的分子标记序列。所述靶结合区可以包含基因特异性序列、寡聚(dT)序列、随机多聚体、或其任何组合。所述多个寡核苷酸条形码中的两个寡核苷酸条形码的两个靶结合区可以包含不同的基因特异性序列。所述多个寡核苷酸条形码的每个细胞标记序列可以包含至少6个核苷酸。所述多个寡核苷酸条形码的每个分子标记序列可以包含至少6个核苷酸。所述多个寡核苷酸条形码中的每个可以包含空间标记序列,并且其中所述多个寡核苷酸条形码的空间标记序列彼此差异至少一个核苷酸。所述多个寡核苷酸条形码中的每个可以包含通用标记序列,其中所述多个寡核苷酸条形码的通用标记序列是相同的。

[0048] 在一些实施例中,所述多个寡核苷酸条形码中的至少一个固定在、部分固定在合成颗粒上,包封在、部分包封在合成颗粒中,或其任何组合。合成颗粒可以是可破坏的。合成颗粒可以包含珠。珠可以包括琼脂糖凝胶珠、链霉亲和素蛋白珠、琼脂糖珠、磁珠、缀合珠、蛋白A缀合珠、蛋白G缀合珠、蛋白A/G缀合珠、蛋白L缀合珠、寡(dT)缀合珠、二氧化硅珠、二氧化硅样珠、抗生物素微珠、抗荧光染料微珠、或其任何组合。合成颗粒可以包含选自下组的材料,该组由以下组成:聚二甲基硅氧烷(PDMS)、聚苯乙烯、玻璃、聚丙烯、琼脂糖、明胶、水凝胶、顺磁物质、陶瓷、塑料、玻璃、甲基苯乙烯、丙烯酸聚合物、钛、乳胶、琼脂糖凝胶、纤维素、尼龙、硅酮、及其任何组合。合成颗粒可以包含可破坏的水凝胶颗粒。所述多个寡核苷

酸条形码中的每个可以包含接头官能团,合成颗粒可以包含固体支持物官能团,和/或所述支持物官能团和所述接头官能团可以彼此相关联。所述接头官能团和所述支持物官能团可以单独地选自下组,该组由以下各项组成:C6、生物素、链霉亲和素、一个或多个伯胺、一个或多个醛、一个或多个酮、及其任何组合。在一些实施例中,所述多个合成颗粒包含至少94个合成颗粒。试剂盒可以包含微孔阵列,其中所述微孔阵列中的每个微孔包含所述多个合成颗粒中的合成颗粒。试剂盒可以包含缓冲液、盒、用于逆转录反应的一种或多种试剂、用于扩增反应的一种或多种试剂或其组合。所述多个寡核苷酸条形码和所述第一多个间隔子寡核苷酸的比率可以在从1:10至10:1的范围内。所述第一多个间隔子寡核苷酸中的每个可以包含细胞标记序列或分子标记序列。

[0049] 本文提供了包括试剂盒的实施例。在一些实施例中,所述试剂盒包含:多个合成颗粒,其中每个合成颗粒包含:多个寡核苷酸条形码,其中所述多个寡核苷酸条形码中的每个包含相同的细胞标记序列、分子标记序列和靶结合区,并且其中所述多个寡核苷酸条形码中的至少10包含不同的分子标记序列;和第一多个间隔子寡核苷酸,其中所述第一多个间隔子寡核苷酸中的每个不包含靶结合区。所述试剂盒可以包含:用于使用所述多个寡核苷酸条形码的说明书。

[0050] 在一些实施例中,所述第一多个间隔子寡核苷酸中的至少一个比所述多个寡核苷酸条形码中的至少一个更短。所述第一多个间隔子寡核苷酸中的每一个可以包含与所述靶结合区相对应的区。与所述靶结合区相对应的所述区和所述靶结合区可以具有不同的序列。与所述靶结合区相对应的所述区和所述靶结合区可以具有至多50%序列同一性。所述第一多个间隔子寡核苷酸中的至少两个可以具有不同长度。所述第一多个间隔子寡核苷酸中的至少两个的序列可以是相同的。所述第一多个间隔子寡核苷酸中的至少两个可以包含相同的序列。

[0051] 在一些实施例中,合成颗粒包含第二多个间隔子寡核苷酸,其中所述第二多个间隔子寡核苷酸中的每个不包含所述靶结合区。所述第二多个间隔子寡核苷酸中的至少一个可以比所述第一多个间隔子寡核苷酸中的至少一个更短。所述第二多个间隔子寡核苷酸中的至少一个可以比所述多个寡核苷酸条形码中的至少一个更短。所述第二多个间隔子寡核苷酸中的至少两个可以具有不同长度。所述第二多个间隔子寡核苷酸中的至少两个的序列可以是相同的。所述第二多个间隔子寡核苷酸中的至少两个可以包含相同的序列。

[0052] 在一些实施例中,所述多个寡核苷酸条形码中的至少100个包含不同的分子标记序列。所述靶结合区可以包含基因特异性序列、寡聚(dT)序列、随机多聚体、或其任何组合。所述多个寡核苷酸条形码中的两个寡核苷酸条形码的两个靶结合区可以包含不同的基因特异性序列。所述多个寡核苷酸条形码的每个细胞标记序列可以包含至少6个核苷酸。所述多个寡核苷酸条形码的每个分子标记序列可以包含至少6个核苷酸。所述多个寡核苷酸条形码中的每个可以包含空间标记序列,并且其中所述多个寡核苷酸条形码的空间标记序列彼此差异至少一个核苷酸。所述多个寡核苷酸条形码中的每个可以包含通用标记序列,其中所述多个寡核苷酸条形码的通用标记序列是相同的。

[0053] 在一些实施例中,所述多个寡核苷酸条形码中的至少一个固定在、部分固定在合成颗粒上,包封在、部分包封在合成颗粒中,或其任何组合。合成颗粒可以是可破坏的。合成颗粒可以包含珠。珠可以包括琼脂糖凝胶珠、链霉亲和素蛋白珠、琼脂糖珠、磁珠、缀合珠、

蛋白A缀合珠、蛋白G缀合珠、蛋白A/G缀合珠、蛋白L缀合珠、寡(dT)缀合珠、二氧化硅珠、二氧化硅样珠、抗生物素微珠、抗荧光染料微珠、或其任何组合。合成颗粒可以包含选自下组的材料,该组由以下组成:聚二甲基硅氧烷(PDMS)、聚苯乙烯、玻璃、聚丙烯、琼脂糖、明胶、水凝胶、顺磁物质、陶瓷、塑料、玻璃、甲基苯乙烯、丙烯酸聚合物、钛、乳胶、琼脂糖凝胶、纤维素、尼龙、硅酮、及其任何组合。合成颗粒可以包含可破坏的水凝胶颗粒。所述多个寡核苷酸条形码中的每个可以包含接头官能团,合成颗粒可以包含固体支持物官能团,和/或所述支持物官能团和所述接头官能团可以彼此相关联。所述接头官能团和所述支持物官能团可以单独地选自下组,该组由以下各项组成:C6、生物素、链霉亲和素、一个或多个伯胺、一个或多个醛、一个或多个酮、及其任何组合。

[0054] 本文披露的实施例包括如下方法,所述方法用于确定样品中靶的数量。在一些实施例中,所述方法包括:使用与合成颗粒相关联的多个寡核苷酸条形码,对样品中多个靶的拷贝进行条形码化,以产生多个经条形码化的靶的拷贝,其中所述多个寡核苷酸条形码中的每个包含相同的细胞标记序列、分子标记序列和靶结合区,并且其中所述多个寡核苷酸条形码中的至少10个包含不同的分子标记序列,其中所述合成颗粒包含第一多个间隔子寡核苷酸,并且其中所述第一多个间隔子寡核苷酸中的每个不包含靶结合区;并且使用所述多个寡核苷酸条形码的分子标记序列,对所述样品中所述多个靶中每个的数量进行评估。在一些实施例中,使用与所述合成颗粒相关联的所述多个寡核苷酸条形码对样品中多个靶的拷贝进行条形码化可以包括:将所述多个靶与所述合成颗粒上的所述多个寡核苷酸条形码杂交。对所述多个靶的拷贝进行条形码化可以包括:逆转录所述多个靶,以产生所述多个靶的经条形码化的cDNA。所述方法可以包括:使用第二链合成,扩增所述经条形码化的cDNA,以产生经条形码化的靶,其中所述经条形码化的靶是双链的。所述方法可以包括:扩增所述经条形码化的靶,以产生经标记的靶扩增子。

[0055] 本文披露了如下方法的实施例,所述方法用于捕获样品中的靶。在一些实施例中,所述方法包括:使用与合成颗粒相关联的多个寡核苷酸条形码,对样品中多个靶的拷贝杂交,以产生多个经条形码化的靶的拷贝,其中所述多个寡核苷酸条形码中的每个包含相同的细胞标记序列、分子标记序列和靶结合区,并且其中所述多个寡核苷酸条形码中的至少10个包含不同的分子标记序列,其中所述合成颗粒包含第一多个间隔子寡核苷酸,并且其中所述第一多个间隔子寡核苷酸中的每个不包含靶结合区。在一些实施例中,所述多个经条形码化的靶的拷贝可以包含所述样品中所述多个靶中靶的拷贝的至少50%、75%、95%、或其任何组合。所述多个经条形码化的靶的拷贝可以包含所述样品中所述多个靶中每个靶的拷贝的至少50%、75%、95%、或其任何组合。所述方法可以包括:逆转录所述多个靶,以产生所述多个靶的经条形码化的cDNA。所述方法可以包括:使用第二链合成,扩增所述经条形码化的cDNA,以产生经条形码化的靶,其中所述经条形码化的靶是双链的。所述方法可以包括:包括扩增所述经条形码化的靶,以产生经标记的靶扩增子。

[0056] 在一些实施例中,所述样品包含单个细胞,所述方法包括将所述合成颗粒与所述样品中的所述单个细胞相关联。单个细胞可以包含稀有细胞、肿瘤细胞、来自人的细胞、来自组织的细胞、来自肿瘤的细胞、被病毒多核苷酸感染的细胞、及其任何组合。所述方法可以包括:在将所述合成颗粒与所述单个细胞相关联之后,裂解所述单个细胞。所述方法可以包括:裂解所述单个细胞包括将所述样品加热、使所述样品与洗涤剂接触、改变所述样品的

pH、或其任何组合。所述合成颗粒和所述单个细胞可以在相同的孔中。所述合成颗粒和所述单个细胞可以在相同的液滴中。所述多个靶可以包含核糖核酸(RNA)、信使RNA(mRNA)、微小RNA、小干扰RNA(siRNA)、RNA降解产物、各自含有聚(A)尾的RNA、及其任何组合。

[0057] 在一些实施例中,所述多个寡核苷酸条形码和所述第一多个间隔子寡核苷酸的比率是在从1:10至10:1范围。所述第一多个间隔子寡核苷酸中的每个可以包含细胞标记序列或分子标记序列。所述第一多个间隔子寡核苷酸中的至少一个可以比所述多个寡核苷酸条形码中的至少一个更短。所述第一多个间隔子寡核苷酸中的每个可以包含与所述靶结合区相对应的区,并且其中与所述靶结合区相对应的所述区和所述靶结合区具有不同的序列。与所述靶结合区相对应的所述区和所述靶结合区可以具有至多50%序列同一性。所述第一多个间隔子寡核苷酸中的至少两个可以具有不同长度。所述第一多个间隔子寡核苷酸中的至少两个的序列可以是相同的。所述第一多个间隔子寡核苷酸中的至少两个可以包含相同的序列。

[0058] 在一些实施例中,所述合成颗粒可以包含第二多个间隔子寡核苷酸,其中所述第二多个间隔子寡核苷酸中的每个不包含靶结合区。所述第二多个间隔子寡核苷酸中的至少一个可以比所述第一多个间隔子寡核苷酸中的至少一个更短。所述第二多个间隔子寡核苷酸中的至少一个可以比所述多个寡核苷酸条形码中的至少一个更短。所述第二多个间隔子寡核苷酸中的至少两个可以具有不同长度。所述第二多个间隔子寡核苷酸中的至少两个的序列可以是相同的。所述第二多个间隔子寡核苷酸中的至少两个可以包含相同的序列。

[0059] 在一些实施例中,所述多个寡核苷酸条形码中的至少100个包含不同的分子标记序列。所述靶结合区可以包含基因特异性序列、寡聚(dT)序列、随机多聚体、或其任何组合。所述多个寡核苷酸条形码中的两个寡核苷酸条形码的两个靶结合区可以包含不同的基因特异性序列。所述多个寡核苷酸条形码的每个细胞标记序列可以包含至少6个核苷酸。所述多个寡核苷酸条形码的每个分子标记序列可以包含至少6个核苷酸。所述多个寡核苷酸条形码中的每个可以包含空间标记序列,并且其中所述多个寡核苷酸条形码的空间标记序列彼此差异至少一个核苷酸。所述多个寡核苷酸条形码中的每个可以包含通用标记序列,其中所述多个寡核苷酸条形码的通用标记序列是相同的。

[0060] 在一些实施例中,所述多个寡核苷酸条形码中的至少一个固定在、部分固定在合成颗粒上,包封在、部分包封在合成颗粒中,或其任何组合。合成颗粒可以是可破坏的。合成颗粒可以包含珠。珠可以包括琼脂糖凝胶珠、链霉亲和素蛋白珠、琼脂糖珠、磁珠、缀合珠、蛋白A缀合珠、蛋白G缀合珠、蛋白A/G缀合珠、蛋白L缀合珠、寡(dT)缀合珠、二氧化硅珠、二氧化硅样珠、抗生物素微珠、抗荧光染料微珠、或其任何组合。合成颗粒可以包含选自下组的材料,该组由以下组成:聚二甲基硅氧烷(PDMS)、聚苯乙烯、玻璃、聚丙烯、琼脂糖、明胶、水凝胶、顺磁物质、陶瓷、塑料、玻璃、甲基苯乙烯、丙烯酸聚合物、钛、乳胶、琼脂糖凝胶、纤维素、尼龙、硅酮、及其任何组合。合成颗粒可以包含可破坏的水凝胶颗粒。

[0061] 在一些实施例中,所述多个寡核苷酸条形码中的每个包含接头官能团,所述合成颗粒包含固体支持物官能团,和/或所述支持物官能团和所述接头官能团彼此相关联。所述接头官能团和所述支持物官能团可以单独地选自下组,该组由以下各项组成:C6、生物素、链霉亲和素、一个或多个伯胺、一个或多个醛、一个或多个酮、及其任何组合。

附图说明

[0062] 图1说明了非限制性示例性条形码(例如随机条形码)。

[0063] 图2显示了进行条形码化和数字计数(例如进行随机条形码化和数字计数)的非限制性示例性工作流程。

[0064] 图3是显示用于从多个靶产生经(例如经条形码化的靶随机经条形码化的靶)的索引文库的非限制性示例性过程的示意图。

[0065] 图4A是与多个寡核苷酸条形码相关联的非限制性示例性合成颗粒。图4B是与多个寡核苷酸条形码和多个间隔子寡核苷酸相关联的非限制性示例性合成颗粒。图4C1-4C2是与多个寡核苷酸条形码和多个间隔子寡核苷酸相关联的其他非限制性示例性合成颗粒。图4D1-4D2是与具有靶特异性靶结合区的多个寡核苷酸条形码相关联的非限制性示例性合成颗粒。图4D1-4D2中展示的合成颗粒可以包含图4B和4C1-4C2中说明的所述多个间隔子寡核苷酸中的一个或多个(未示出)。

[0066] 图5A-5C显示如下非限制性示例性比较:具有能够与mRNA分子的聚(A)尾杂交的聚(dT)区的寡核苷酸条形码(图5A);具有能够与一个或多个目的靶结合的靶特异性靶结合区的寡核苷酸条形码(图5B);具有聚(dT)区和靶特异性靶结合区的寡核苷酸条形码(图5C)。

[0067] 图6是柱图,显示使用珠,在RNA捕获中的非限制性示例性改善,每个珠包含具有能够捕获RNA分子的寡聚(dT)尾和多个间隔子寡核苷酸的多个寡核苷酸条形码。

[0068] 图7是显示使用具有TCR特异性寡核苷酸条形码的珠,T细胞受体(TCR)捕获效率的非限制性示例性比较的柱图。

具体实施方式

[0069] 在以下详细说明中参考了形成本文的一部分的附图。在附图中,除非上下文另有指示,否则相似的符号通常标识相似的部件。在具体实施例、附图和权利要求中描述的说明性实施例并不意味着是限制性的。在不脱离本文提出的主题的精神或范围的情况下,可以利用其他实施例,并且可以做出其他改变。容易理解的是,如本文一般描述的以及图中说明的本披露的方面能以各种不同的配置来布置、替换、组合、分离和设计,所有这些都都在本文中明确考虑并且构成本披露内容的一部分。

[0070] 来自GenBank的所有专利、公开的专利申请、其他出版物、和序列,以及本文提及的其他数据库关于相关技术通过引用以其整体并入。

[0071] 确定核酸或靶(例如信使核糖核苷酸(mRNA)分子)的数量对于确定例如在不同发育阶段或在不同环境条件下在细胞中表达的基因是临床上重要的。然而,确定核酸分子(例如,mRNA分子)的绝对数目也是非常具有挑战性的,尤其是当分子数目非常小时。确定样品中分子的绝对数量的一种方法是数字聚合酶链式反应(PCR)。理想地,PCR在每个循环中产生分子的相同拷贝。然而,PCR可具有缺点使得每个分子复制具有随机概率,且此概率根据PCR循环和基因序列而变化,这导致扩增偏差和不准确的基因表达测量。

[0072] 可以将具有独特分子标记(ML,也称为分子指数(MI))的条形码(例如随机条形码)用于计数分子数目。对每个细胞标记独特的具有分子标记的条形码可用于计数每个细胞中的分子数目。用于进行条形码化的非限制性示例性测定包括PreciseTM测定(细胞研究公司(Cellular Research, Inc.) (帕洛阿尔托,加利福尼亚州))、ResolveTM测定(细胞研究公司

(帕洛阿尔托,加利福尼亚州))或Rhapsody™测定(贝克顿迪金森公司(Becton,Dickinson and Company)(富兰克林湖,新泽西))。然而,这些方法和技术可能引入误差,如果不校正的话,可能导致估计过高的细胞计数。

[0073] Rhapsody™测定法可利用具有大量的(例如6561至65536个)条形码(例如随机条形码)的非耗尽性池、聚(T)寡核苷酸上的独特分子标记,以在RT步骤期间与样品中的所有聚(A)-mRNA杂交。除了分子标记,可以使用条形码的细胞标记来鉴定微孔板的每个孔中的每个单细胞。条形码可包括通用PCR引发位点。在RT期间,靶基因分子与条形码随机地反应。每个靶分子可以与条形码(例如随机条形码)杂交,从而产生经条形码化的互补核糖核苷酸(cDNA)分子(例如经随机条形码化的cDNA分子)。在标记后,可将来自微孔板微孔的经条形码化的cDNA分子合并到单个管中用于PCR扩增和测序。可以分析原始测序数据来得到具有独特分子标记的条形码的数目。

[0074] 本文的披露内容包括合成颗粒的实施例。在一些实施例中,所述合成颗粒包含:第一多个寡核苷酸条形码;和第二多个寡核苷酸条形码,其中所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码中的每个包含相同的细胞标记序列并且包含分子标记序列,其中所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码中的至少10个包含不同的分子标记序列,其中所述第一多个寡核苷酸条形码中的每个包含靶结合区,并且其中所述第二多个寡核苷酸条形码中的每个包含靶特异性靶结合区。在一些实施例中,所述第一多个寡核苷酸条形码的寡核苷酸条形码的靶结合区包含靶特异性靶结合区。在一些实施例中,所述第一多个寡核苷酸条形码的寡核苷酸条形码的靶结合区包含聚(dT)区。所述第二多个寡核苷酸条形码的寡核苷酸条形码可以包含聚(dT)区。

[0075] 本文的披露内容包括试剂盒或组合物的实施例。在一些实施例中,所述试剂盒或组合物包含:多个合成颗粒,其中每个合成颗粒包含:第一多个寡核苷酸条形码;和第二多个寡核苷酸条形码,其中所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码中的每个包含相同的细胞标记序列并且包含分子标记序列,其中所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码中的至少10个包含不同的分子标记序列,其中所述第一多个寡核苷酸条形码中的每个包含靶结合区,并且其中所述第二多个寡核苷酸条形码中的每个包含靶特异性靶结合区;以及任选地,用于使用所述多个合成颗粒的说明书。在一些实施例中,所述试剂盒包含缓冲液、用于逆转录反应的一种或多种试剂、用于扩增反应的一种或多种试剂或其组合。在一些实施例中,所述多个合成颗粒包含至少94个合成颗粒。所述试剂盒可以包含多个分区,每个分区包含所述多个合成颗粒中的合成颗粒。所述多个分区可以包含微孔阵列的微孔。所述多个分区可以包含多个液滴。在一些实施例中,所述第一多个寡核苷酸条形码的寡核苷酸条形码的靶结合区包含靶特异性靶结合区。所述第一多个寡核苷酸条形码的寡核苷酸条形码的靶结合区可以包含聚(dT)区。所述第二多个寡核苷酸条形码的寡核苷酸条形码可以包含聚(dT)区。

[0076] 本文的披露内容包括如下方法的实施例,所述方法用于确定样品中靶的数量。在一些实施例中,所述方法包括:使用与合成颗粒相关联的第一多个寡核苷酸条形码,对第一多个靶的拷贝进行条形码化,以产生第一多个经条形码化的靶的拷贝;使用与所述合成颗粒相关联的第二多个寡核苷酸条形码,对样品中第二多个靶的拷贝进行条形码化,以产生第二多个经条形码化的靶的拷贝,其中所述第一多个寡核苷酸条形码和所述第二多个寡核

苷酸条形码中的每个包含相同的细胞标记序列并且包含分子标记序列,其中所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码中的至少10个包含不同的分子标记序列,其中所述第一多个寡核苷酸条形码中的每个包含靶结合区,并且其中所述第二多个寡核苷酸条形码中的每个包含靶特异性靶结合区;并且使用所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码的分子标记序列对所述第一多个靶和/或所述第二多个靶中的每个的数量进行评估。在一些实施例中,所述第一多个寡核苷酸条形码的寡核苷酸条形码的靶结合区包含靶特异性靶结合区。所述第一多个寡核苷酸条形码的寡核苷酸条形码的靶结合区可以包含聚(dT)区。所述第二多个寡核苷酸条形码的寡核苷酸条形码可以包含聚(dT)区。在一些实施例中,第一多个靶包含细胞的mRNA分子,并且其中第二多个靶包含细胞的DNA分子。所述第一多个靶可以包含细胞的mRNA分子,并且所述第二多个靶可以包含感染细胞的病毒、细菌和/或真菌的RNA分子。

[0077] 本文披露了与多个寡核苷酸相关联的颗粒的实施例。在一些实施例中,所述颗粒是合成颗粒。在一些实施例中,所述合成颗粒包含:多个寡核苷酸条形码,其中所述多个寡核苷酸条形码中的每个包含相同的细胞标记、分子标记和靶结合区,并且其中所述多个寡核苷酸条形码中的至少10个包含不同的分子标记序列;和第一多个间隔子寡核苷酸,其中所述第一多个间隔子寡核苷酸中的每个不包含靶结合区。所述多个寡核苷酸条形码和所述第一多个间隔子寡核苷酸的比率可以在从1:10至10:1的范围内。所述第一多个间隔子寡核苷酸中的每个可以包含细胞标记或分子标记。

[0078] 本文还提供了包括试剂盒的实施例,所述试剂盒包含本披露的一个或多个合成颗粒。在一些实施例中,所述试剂盒包含:多个合成颗粒,其中每个合成颗粒包含:多个寡核苷酸条形码,其中所述多个寡核苷酸条形码中的每个包含相同的细胞标记、分子标记和靶结合区,并且其中所述多个寡核苷酸条形码中的至少10个包含不同的分子标记序列;和第一多个间隔子寡核苷酸,其中所述第一多个间隔子寡核苷酸中的每个不包含靶结合区。所述试剂盒可以包含:用于使用所述多个寡核苷酸条形码的说明书。

[0079] 本文披露的实施例包括如下方法,所述方法用于确定样品中靶的数量。在一些实施例中,所述方法包括:使用与合成颗粒相关联的多个寡核苷酸条形码,对样品中多个靶的拷贝进行条形码化,以产生多个经条形码化的靶的拷贝,其中所述多个寡核苷酸条形码中的每个包含相同的细胞标记、分子标记和靶结合区,并且其中所述多个寡核苷酸条形码中的至少10个包含不同的分子标记序列,其中所述合成颗粒包含第一多个间隔子寡核苷酸,并且其中所述第一多个间隔子寡核苷酸中的每个不包含靶结合区;并且使用所述多个寡核苷酸条形码的分子标记,对所述样品中所述多个靶中每个的数量进行评估。

[0080] 本文披露了如下方法的实施例,所述方法用于捕获样品中的靶。在一些实施例中,所述方法包括:使用与合成颗粒相关联的多个寡核苷酸条形码,对样品中多个靶的拷贝杂交,以产生多个经条形码化的靶的拷贝,其中所述多个寡核苷酸条形码中的每个包含相同的细胞标记、分子标记和靶结合区,并且其中所述多个寡核苷酸条形码中的至少10个包含不同的分子标记序列,其中所述合成颗粒包含第一多个间隔子寡核苷酸,并且其中所述第一多个间隔子寡核苷酸中的每个不包含靶结合区。

[0081] 一些定义

[0082] 除非另外定义,本文所用的技术术语和科学术语具有与本披露所属领域的普通技

术人员通常所理解的含义。参见,例如Singleton等人.,*Dictionary of Microbiology and Molecular Biology* 2nd ed. [微生物学和分子生物学词典,第2版],J.Wiley&Sons[约翰威利父子公司](纽约,纽约州1994);Sambrook等人.,*Molecular Cloning, A Laboratory Manual* [分子克隆,实验室手册],Cold Spring Harbor Press[冷泉港出版社](冷泉港,纽约1989)。出于本披露的目的,以下术语定义如下。

[0083] 如本文所用的,术语“衔接子”可以意指促进相关联的核酸的扩增或测序的序列。相关联的核酸可包括靶核酸。相关联的核酸可包含空间标记、靶标记、样品标记、索引标记、条形码、随机条形码、或分子标记中的一个或多个。衔接子可以是线性的。衔接子可以是预腺苷酸化的衔接子。衔接子可以是双链或单链的。一个或多个衔接子可以位于核酸的5' 或 3' 端。当衔接子在5' 和3' 端包括已知序列时,已知序列可以是相同或不同的序列。位于多核苷酸的5' 和/或3' 端的衔接子能够与固定在表面上一个或多个寡核苷酸杂交。在一些实施例中,衔接子可以包括通用序列。通用序列可以是两个或更多个核酸分子共有的核苷酸序列的区域。两个或更多个核酸分子可具有不同序列的区域。因此,例如,5' 衔接子可以包括相同和/或通用核酸序列,并且3' 衔接子可以包括相同和/或通用序列。可存在于多个核酸分子的不同成员中的通用序列可允许使用与通用序列互补的单个通用引物复制或扩增多个不同序列。相似地,可以存在于核酸分子的集合中的不同成员中的至少一个、两个(例如,一对)或更多个通用序列可以允许使用与通用序列互补的至少一个、两个(例如,一对)或更多个单个通用引物复制或扩增多个不同序列。因此,通用引物包括可与此类通用序列杂交的序列。可以修饰携带靶核酸序列的分子以将通用衔接子(例如,非靶核酸序列)附接至不同靶核酸序列的一端或两端。与靶核酸附接的一个或多个通用引物可以提供通用引物杂交的位点。与靶核酸附接的一个或多个通用引物可以彼此相同或不同。

[0084] 如本文使用的,术语“关联”或“与……相关联”可意指两个或更多个种类可以被鉴定为在某个时间点处共定位。关联可意指两个或更多个种类在或曾经在相似的容器内。关联可以是信息学关联,其中例如关于两个或更多个种类的数字信息被存储并且可以用于确定所述种类中的一个或多个在某个时间点处共定位。关联可以是直接或间接的物理关联。在一些实施例中,两个或更多个相关联的种类彼此之间或与共同的固体或半固体表面是“拴系的”、“附接的”或“固定的”。关联可以指用于将标记附接到固体或半固体支持物(例如合成颗粒或珠)上的共价或非共价方式。关联可以是靶与标记之间的共价键。

[0085] 如本文使用的,术语“互补性”可以指两个核苷酸之间精确配对的能力。例如,如果核酸的在给定位置的核苷酸能够与另一个核酸的核苷酸以氢键结合,则两个核酸被认为在所述位置处是彼此互补的。两单链核酸分子之间的互补性可以是“部分的”,其中所述核苷酸中仅一些结合,或者当所述单链分子之间存在完全互补性时,这种互补性可以是完全的。如果第一核苷酸序列与第二核苷酸序列互补,则可以认为第一核苷酸序列是第二序列的“互补体”。如果第一核苷酸序列互补于和第二序列相反的序列(即,核苷酸顺序相反),则可以认为第一核苷酸序列是第二序列的“反向互补体”。如本文使用的,术语“互补体”、“互补”和“反向互补体”可以互换使用。从本披露可以理解,如果一个分子可以与另一个分子杂交,则其可以是杂交的分子的互补体。

[0086] 如本文使用的,术语“数字计数”可以指用于估计样品中靶分子数量的方法。数字计数可以包括确定已经与样品中的靶相关联的独特标记的数量的步骤。这种随机方法将计

数分子的问题从相同分子的定位和鉴定之一转化为有关检测到一组预定义标记的一系列是/否数字问题。

[0087] 如本文使用的,术语(一个或多个)“标记”可以指与样品中的靶相关联的核酸代码。标记可以是例如核酸标记。标记可以是完全或部分可扩增的标记。标记可以是完全或部分可测序的标记。标记可以是可鉴定为有区别的天然核酸的一部分。标记可以是已知的序列。标记可以包括核酸序列的接点,例如天然和非天然序列的接点。如本文使用的,术语“标记”可以与术语“索引”、“标签”或“标记-标签”互换使用。标记可以传达信息。例如,在各种实施例中,可以使用标记来确定样品的身份、样品的来源、细胞的身份和/或靶。

[0088] 如本文使用的,术语“非耗尽性储库(non-depleting reservoir)”可以指由许多不同标记组成的随机条形码池。非耗尽性储库可以包括大量不同的随机条形码,使得当非耗尽性储库与靶池相关联时,每个靶可能与独特的随机条形码相关联。每个经标记的靶分子的独特性可以通过随机选择的统计来确定,并且取决于与多样的标记相比在集合中相同的靶分子的拷贝数。所得的经标记的靶分子集合的大小可以通过条形码化处理的随机性质来确定,然后对检测到的随机条形码的数目的分析允许计算原始集合或样品中存在的靶分子的数目。当存在的靶分子的拷贝数与独特的随机条形码的数目的比率低时,经标记的靶分子是高度独特的(即,用给定的标记来标记多于一个靶分子的概率非常低)。

[0089] 如本文所用的,术语“核酸”是指多核苷酸序列、或其片段。核酸可以包括核苷酸。核酸对于细胞可以是外源的或内源的。核酸可以存在于无细胞环境中。核酸可以是基因或其片段。核酸可以是DNA。核酸可以是RNA。核酸可以包括一种或多种类似物(例如改变的骨架、糖或核碱基)。类似物的一些非限制性实例包括:5-溴尿嘧啶、肽核酸、外来核酸、吗啉代、锁核酸、二醇核酸、苏糖核酸、二脱氧核苷酸、虫草菌素、7-脱氮-GTP、荧光团(例如,罗丹明或与糖连接的荧光黄素)、含有核苷酸的硫醇、生物素连接的核苷酸、荧光基类似物、CpG岛、甲基-7-鸟苷、甲基化的核苷酸、肌苷、硫代尿苷、假尿苷、二氢尿苷、癸苷、以及怀俄苷。“核酸”、“多核苷酸”、“靶多核苷酸”和“靶核酸”可以互换使用。

[0090] 核酸可以包括一种或多种修饰(例如,碱基修饰、骨架修饰),以为核酸提供新的或增强的特征(例如,改进的稳定性)。核酸可以包括核酸亲和标签。核苷可以是碱基-糖组合。核苷的碱基部分可以是杂环碱基。此类杂环碱基的两个最常见的类别是嘌呤和嘧啶。核苷酸可以是进一步包括与核苷的糖部分共价连接的磷酸基团的核苷。对于包括呋喃戊糖的那些核苷,磷酸基团可以连接到糖的2'、3'或5'羟基部分。在形成核酸中,磷酸基团可以将相邻的核苷彼此共价连接以形成线性高分子化合物。转而成此线性高分子化合物的各自端可以进一步接合而形成环状化合物;然而,线性化合物通常是合适的。此外,线性化合物可以具有内部核苷酸碱基互补性,并且因此可以按产生完全或部分双链化合物的方式折叠。在核酸中,所述磷酸基团通常可以被称为形成核酸的核苷间骨架。连键或骨架可以是3'到5'磷酸二酯键。

[0091] 核酸可以包括修饰的骨架和/或修饰的核苷间键。修饰的骨架可以包括在骨架中保留磷原子和在骨架中不具有磷原子的那些。其中含有磷原子的合适修饰的核酸骨架可以含有例如硫代磷酸酯;手性硫代磷酸酯;二硫代磷酸酯;磷酸三酯;氨基烷基磷酸三酯;甲基磷酸酯和其他烷基磷酸酯,如3'-亚烷基磷酸酯、5'-亚烷基磷酸酯;手性磷酸酯;亚磷酸酯;包括3'-氨基磷酰胺酯和氨基烷基磷酰胺酯的磷酰胺酯;磷二酰胺酯;硫代羰基磷酰胺酯;

硫代羰基烷基膦酸酯;硫代羰基烷基磷酸三酯;硒代磷酸酯;以及具有正常3'-5'键的硼烷磷酸酯,2'-5'连接的类似物和具有反向极性的那些,其中一个或多个核苷酸间键是3'至3'、5'至5'或2'至2'键。

[0092] 核酸可以包括由短链烷基或环烷基核苷间键、混合杂原子、和烷基或环烷基核苷间键或者一个或多个短链杂原子的或杂环的核苷间键形成的多核苷酸骨架。这些可包括具有以下结构的那些:吗啉代键(从核苷的糖部分部分地形成);硅氧烷骨架;硫化物、亚砷和砷骨架;甲酰乙酰基和硫代甲酰乙酰基骨架;亚甲基甲酰乙酰基和硫代甲酰乙酰基骨架;核糖乙酰基骨架;含烯的骨架;氨基磺酸盐骨架;亚甲亚氨基和亚甲胍基骨架;磺酸酯和磺酰胺骨架;酰胺骨架;和具有混合的N、O、S和CH₂组分部分的其他骨架。

[0093] 核酸可以包括核酸模拟物。术语“模拟物”可以旨在包括其中只有呋喃糖环或呋喃糖环和核苷酸间键两者被非呋喃糖基团替代的多核苷酸,仅替代呋喃糖环可以称为糖替代物。可以保持杂环碱基部分或修饰的杂环碱基部分以便与适当的靶核酸杂交。一种这样的核酸可以是肽核酸(PNA)。在PNA中,多核苷酸的糖骨架可以被含酰胺的骨架(特别是氨基乙基甘氨酸骨架)替代。核苷酸可以被保持并且直接或间接地结合至骨架的酰胺部分的氮杂氮原子上。PNA化合物中的骨架可以包括两个或更多个连接的氨基乙基甘氨酸单元,其给予PNA含酰胺的骨架。杂环碱基部分可以直接或间接地结合到骨架的酰胺部分的氮杂氮原子上。

[0094] 核酸可以包括吗啉代骨架结构。例如,核酸可以包括代替核糖环的6元吗啉代环。在这些实施例的一些中,磷二酰胺酯或其他非磷酸二酯核苷间键可替代磷酸二酯键。

[0095] 核酸可以包括具有附接到吗啉代环上的杂环碱基的连接的吗啉代单元(即吗啉代核酸)。连接基团可以连接吗啉代核酸中的吗啉代单体单元。非离子型基于吗啉代的寡聚化合物可以与细胞蛋白具有较少的不希望的作用。基于吗啉代的多核苷酸可以是核酸的非离子模拟物。吗啉代类别中的多种化合物可以使用不同的连接基团连接。另一类多核苷酸模拟物可称为环己烯基核酸(CeNA)。通常存在于核酸分子中的呋喃糖环可以被环己烯基环替代。可以制备CeNA DMT保护的亚磷酰胺单体,并用于使用亚磷酰胺化学的寡聚化合物合成。将CeNA单体并入核酸链可以增加DNA/RNA杂交体的稳定性。CeNA寡聚腺苷酸可以与具有和天然复合物相似的稳定性的核酸互补体形成复合物。另外的修饰可以包括锁核酸(LNA),其中2'-羟基基团连接到糖环的4'碳原子,从而形成2'-C,4'-C-氧亚甲基键,由此形成双环糖部分。所述键可以是桥连2'氧原子和4'碳原子的基团亚甲基(-CH₂-),其中n是1或2。LNA和LNA类似物可以表现出与互补性核酸非常高的双链体热稳定性(T_m=+3°C至+10°C)、对3'-外切核苷酸降解的稳定性以及良好的溶解度特性。

[0096] 核酸还可以包括核碱基(通常简称为“碱基”)修饰或取代。如本文使用的,“未修饰的”或“天然的”核碱基可以包括嘌呤碱基(例如腺嘌呤(A)和鸟嘌呤(G))、以及嘧啶碱基(例如胸腺嘧啶(T)、胞嘧啶(C)和尿嘧啶(U))。经修饰的核碱基可以包括其他合成以及天然的核碱基,如5-甲基胞嘧啶(5-me-C),5-羟甲基胞嘧啶,黄嘌呤,次黄嘌呤,2-氨基腺嘌呤,腺嘌呤和鸟嘌呤的6-甲基以及其他烷基衍生物,腺嘌呤和鸟嘌呤的2-丙基以及其他烷基衍生物,2-硫尿嘧啶,2-硫胸腺嘧啶以及2-硫胞嘧啶,5-卤代尿嘧啶以及胞嘧啶,5-丙炔基(-C≡C-CH₃)尿嘧啶及胞嘧啶以及嘧啶碱基的其他炔基衍生物,6-偶氮基尿嘧啶,胞嘧啶以及胸腺嘧啶,5-尿嘧啶(假尿嘧啶),4-硫尿嘧啶,8-卤基、8-氨基、8-巯基、8-硫烷基、8-羟基以

及其他8-取代的腺嘌呤和鸟嘌呤,5-卤基特别是5-溴、5-三氟甲基以及其他5-取代的尿嘧啶和胞嘧啶、7-甲基鸟嘌呤和7-甲基腺嘌呤、2-F-腺嘌呤、2-氨基腺嘌呤、8-氮杂鸟嘌呤和8-氮杂腺嘌呤、7-脱氮杂鸟嘌呤和7-脱氮杂腺嘌呤、以及3-脱氮杂鸟嘌呤和3-脱氮杂腺嘌呤。经修饰的核碱基可以包括三环嘧啶如吩噻嗪胞苷(1H-嘧啶并(5,4-b)(1,4)苯并噻嗪-2(3H)-酮)、吩噻嗪胞苷(1H-嘧啶并(5,4-b)(1,4)苯并噻嗪-2(3H)-酮),G-夹(c1amp)如取代的吩噻嗪胞苷(例如9-(2-氨基乙氧基)-H-嘧啶并(5,4-(b)(1,4)苯并噻嗪-2(3H)-酮)、吩噻嗪胞苷(1H-嘧啶并(5,4-b)(1,4)苯并噻嗪-2(3H)-酮),G-夹(c1amp)如取代的吩噻嗪胞苷(例如9-(2-氨基乙氧基)-H-嘧啶并(5,4-(b)(1,4)苯并噻嗪-2(3H)-酮)、吡啶胞苷(2H-嘧啶并(4,5-b)吡啶-2-酮)、吡啶并吡啶胞苷(H吡啶并(3',2':4,5)吡咯并[2,3-d]嘧啶-2-酮)。

[0097] 如本文使用的,术语“样品”可以指包括靶的组合物。用于通过披露的方法、装置、和系统进行分析的合适样品包括细胞、组织、器官、或生物体。

[0098] 如本文使用的,术语“采样装置”或“装置”可以指可以取一部分样品和/或将所述部分放置在基底上的装置。采样装置可以指例如荧光激活细胞分选(FACS)机、细胞分选机、活检针、活检装置、组织切片装置、微流体装置、叶栅和/或超薄切片机。

[0099] 如本文使用的,术语“固体支持物”可以指可以附接多个随机条形码的离散固体或半固体表面。固体支持物可以包括任何类型的实心的、多孔的或空心的球体、球、承座、圆柱体或其他类似配置,其由塑料、陶瓷、金属或高分子材料(例如,水凝胶)构成,其上可以固定核酸(例如,共价地或非共价地)。固体支持物可以包括可以是球形的(例如,微球)或具有非球形或不规则形状的离散颗粒,所述形状是诸如立方体、长方形、锥形、圆柱形、圆锥形、椭圆形或圆盘形等。以阵列间隔开的多个固体支持物可以不包括基底。固体支持物可以是颗粒,例如合成颗粒或珠。

[0100] 固体支持物可以指“基底”。基底可以是一种固体支持物。基底可以指可以在其上进行本披露的方法的连续的固体或半固体表面。例如,基底可以指阵列、盒、芯片、装置和载玻片。

[0101] 如本文使用的,术语“空间标记”可以指其可以在空间中与位置相关联的标记。

[0102] 如本文使用的,术语“随机条形码”可以指包含标记的多核苷酸序列。随机条形码可以是可用于随机条形码化的多核苷酸序列。随机条形码可用于对样品中的靶定量。随机条形码可用于控制标记与靶相关联后可能发生的错误。例如,随机条形码可用于评估扩增或测序错误。与靶相关联的随机条形码可以称为随机条形码-靶或随机条形码-标签-靶。

[0103] 如本文所用,术语“基因特异性随机条形码”可以指包括标记和基因特异性的靶结合区的多核苷酸序列。随机条形码可以是可用于随机条形码化的多核苷酸序列。随机条形码可用于对样品中的靶定量。随机条形码可用于控制标记与靶相关联后可能发生的错误。例如,随机条形码可用于评估扩增或测序错误。与靶相关联的随机条形码可以称为随机条形码-靶或随机条形码-标签-靶。

[0104] 如本文使用的,术语“随机条形码化”可以指核酸的随机标记(例如,条形码化)。随机条形码化可以利用递归泊松策略来关联并对与靶相关联的标记进行定量。如本文所用的,术语“随机条形码化”可以与“基因特异性随机条形码化”互换地使用。

[0105] 如本文使用的,术语“靶”可以指可与随机条形码相关联的组合物。用于通过披露

的方法、装置和系统进行分析的示例性合适的靶包括寡核苷酸、DNA、RNA、mRNA、微小RNA、tRNA等。靶可以是单链的或双链的。在一些实施例中,靶可以是蛋白质。在一些实施例中,靶是脂质。

[0106] 如本文所用的,术语“逆转录酶”可以指具有逆转录酶活性(即,催化从RNA模板合成DNA)的一组酶。通常,这样的酶包括但不限于逆转录病毒逆转录酶、逆转录转座子逆转录酶、逆转录质粒逆转录酶、逆转录子逆转录酶、细菌逆转录酶、II型内含子衍生的逆转录酶,及其突变体、变体或衍生物。非逆转录病毒逆转录酶包括非LTR逆转录转座子逆转录酶、逆转录质粒逆转录酶、逆转录子逆转录酶和II型内含子逆转录酶。II型内含子逆转录酶的实例包括乳酸乳球菌LI.LtrB内含子逆转录酶、细长嗜热聚球藻(*Thermosynechococcus elongatus*)TeI4c内含子逆转录酶或嗜热脂肪土芽孢杆菌GsI-IIIC内含子逆转录酶。其他类别的逆转录酶可以包括许多类型的非逆转录病毒逆转录酶(即,逆转录子、II型内含子、以及多样性产生型逆转录元件等等)。

[0107] 本文中公开了用于鉴定信号细胞标记的系统和方法。在一些实施例中,所述方法包括:(a)使用多个随机条形码使细胞样品中的多个靶条随机条形码化以创造多个经随机条形码化的靶,其中该多个随机条形码中的每个包含细胞标记和分子标记;(b)获得所述多个经随机条形码化的靶的测序数据;(c)确定与所述多个随机条形码的细胞标记中的每个相关联的具有不同序列的分子标记的数目;(d)基于与细胞标记中的每个相关联的具有不同序列的分子标记的数目,确定所述多个随机条形码的细胞标记中的每个的等级;(e)基于(c)中确定的与细胞标记中的每个相关联的具有不同序列的分子标记的数目和(d)中确定的细胞标记中的每个的等级,生成累积和图;(f)生成该累积和图的二阶导数图;(g)确定所述累积和图的二阶导数图的最小值,其中所述二阶导数图的最小值对应于细胞标记阈值;和(h)基于(c)中确定的与细胞标记中的每个相关联的具有不同序列的分子标记的数目和(g)中确定的细胞标记阈值,将细胞标记中的每个鉴定为信号细胞标记或噪声细胞标记。

[0108] 条形码

[0109] 条形码化(诸如随机条形码化)已经在如下文献中进行了描述:例如Fu等人.,*Proc Natl Acad Sci U.S.A.* [美国国家科学院院刊],2011年5月31日,108(22):9026-31;美国专利申请公开号US 2011/0160078;Fan等人.,*Science* [科学],2015年2月6日,347(6222):1258367;美国专利申请公开号US2015/0299784;和PCT申请公开号W02015/031691;将这些文献中的每个的内容(包括任何支持性或补充性信息或材料)通过引用以其全文并入本文中。在一些实施例中,本文披露的条形码可以是随机条形码,所述随机条形码可以是可用于对靶进行随机标记(例如,条形码,标签)的多核苷酸序列。如果随机条形码的不同的条形码序列的数量与待标记的任何靶的出现的数量的比率可以是、或约1:1、2:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1、10:1、11:1、12:1、13:1、14:1、15:1、16:1、17:1、18:1、19:1、20:1、30:1、40:1、50:1、60:1、70:1、80:1、90:1、100:1、或在这些值的任何两个之间的数字或范围,则条形码可以称为随机条形码。靶可以是,例如包括具有相同或几乎相同序列的mRNA分子的mRNA种类。如果随机条形码的不同的条形码序列的数量与待标记的任何靶的出现的数量的比率是至少、或至多1:1、2:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1、10:1、11:1、12:1、13:1、14:1、15:1、16:1、17:1、18:1、19:1、20:1、30:1、40:1、50:1、60:1、70:1、80:1、90:1、或100:1,则条形码可以称为随机条形码。随机条形码的条形码序列可以称为分子标记。

[0110] 条形码(例如随机条形码)可以包括一个或多个标记。示例性标记可以包括通用标记、细胞标记、条形码序列(例如,分子标记)、样品标记、板标记、空间标记、和/或前空间(pre-spatial)标记。图1展示了具有空间标记的示例性条形码104。条形码104可包括可将条形码与固体支持物108连接的5' 胺。条形码可以包括通用标记、维度标记、空间标记、细胞标记、和/或分子标记。条形码中不同标记(包括但不限于通用标记、维度标记、空间标记、细胞标记、和分子标记)的顺序可以改变。例如,如图1中显示,通用标记可以是5' -末端标记,且分子标记可以是3' -末端标记。空间标记、维度标记、和细胞标记可以处于任何顺序。在一些实施例中,通用标记、空间标记、维度标记、细胞标记、和分子标记是处于任何顺序的。条形码可以包括靶结合区。靶结合区可以与样品中的靶(例如,靶核酸、RNA、mRNA、DNA)相互作用。例如,靶结合区可以包括可以与mRNA的聚(A)尾相互作用的寡聚(dT)序列。在一些情况下,条形码的标记(例如,通用标记、维度标记、空间标记、细胞标记和条形码序列)可以由1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20个或更多个核苷酸分开。

[0111] 标记(例如细胞标记)可包括一组独特的定义长度的核酸子序列,例如每个七个核苷酸(相当于一些汉明错误校正代码中使用的比特数目),其可以设计为提供错误校正能力。可以设计包括七个核苷酸序列的错误校正子序列组,使得所述组中的序列的任何成对组合展现出定义的“遗传距离”(或错配碱基数),例如一组纠错子序列能被设计为展现三个核苷酸的遗传距离。在这种情况下,对于经标记的靶核酸分子的序列数据组中的错误校正序列的审查(在下面更全面地描述)能允许检测或校正扩增或测序错误。在一些实施例中,用于产生错误校正代码的核酸子序列的长度可以变化,例如,它们可以是、或约1、2、3、4、5、6、7、8、9、10、15、20、30、31、40、50个、或在这些值的任何两个之间的数字或范围的核苷酸长度。在一些实施例中,其他长度的核酸子序列可以用来产生错误校正代码。

[0112] 条形码可以包括靶结合区。靶结合区可以与样品中的靶相互作用。所述靶可以是、或包括核糖核酸(RNA)、信使RNA(mRNA)、微小RNA、小干扰RNA(siRNA)、RNA降解产物、各自含有聚(A)尾的RNA、或其任何组合。在一些实施例中,多个靶可以包括脱氧核糖核酸(DNA)。

[0113] 在一些实施例中,靶结合区可以包括可以与mRNA的聚(A)尾相互作用的寡聚(dT)序列。条形码的一个或多个标记(例如,通用标记、维度标记、空间标记、细胞标记、和条形码序列(例如,分子标记))可以通过一个或多个间隔子与条形码的剩余标记的另一个或两个分开。间隔子的长度可以是例如1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20个或更多个核苷酸。在一些实施例中,条形码的标记中没有标记被间隔物分开。

[0114] 通用标记

[0115] 条形码可以包括一个或多个通用标记。在一些实施例中,对于条形码组中的所有条形码(附接到给定的固体支持物上的),一个或多个通用标记可以是相同的。在一些实施例中,对于所有附接到多个颗粒(例如,合成颗粒,如珠)上的条形码,一个或多个通用标记可以是相同的。在一些实施例中,通用标记可以包括能够与测序引物杂交的核酸序列。测序引物可以用于对包括通用标记的条形码进行测序。测序引物(例如,通用测序引物)可以包括与高通量测序平台相联系的测序引物。在一些实施例中,通用标记可以包括能够与PCR引物杂交的核酸序列。能够与测序或PCR引物杂交的通用标记的核酸序列可以被称为引物结合位点。通用标记可以包括可用于引发条形码转录的序列。通用标记可以包括可用于延伸条形码或条

形码内的区域的序列。通用标记的长度可以是、或约1、2、3、4、5、10、15、20、25、30、35、40、45、50个核苷酸、或在这些值的任何两个之间的数字或范围的核苷酸。例如,通用标记可以包括至少约10个核苷酸。通用标记的长度可以是至少、或至多1、2、3、4、5、10、15、20、25、30、35、40、45、50、100、200、或300个核苷酸。在一些实施例中,可切割接头或修饰的核苷酸可以是通用标记序列的一部分,以使条形码能够从支持物上被切割下来。

[0116] 维度标记

[0117] 条形码可以包括一个或多个维度标记。在一些实施例中,维度标记可以包括提供关于标记(例如,随机标记)发生的维度的信息的核酸序列。例如,维度标记可以提供关于对靶进行随机条形码化的时间的信息。维度标记可以与样品中条形码化(例如,随机条形码化)的时间相关联。维度标记可以在标记的时间处被激活。不同的维度标记可以在不同的时间被激活。该维度标记提供关于靶、靶组和/或样品被随机条形码化的顺序的信息。例如,在细胞周期的G0期可以对细胞群进行随机条形码化。在细胞周期的G1期,可以用条形码(例如,随机条形码)对所述细胞再次进行脉冲处理。在细胞周期的S期,可以用条形码对所述细胞再次进行脉冲处理,等等。每个脉冲(例如,细胞周期的每个阶段)处的条形码可以包含不同的维度标记。以这种方式,所述维度标记提供关于哪些靶在细胞周期的哪个时期被标记的信息。维度标记可以探询许多不同的生物阶段。示例性的生物学时间可以包括但不限于细胞周期、转录(例如,转录起始)和转录物降解。在另一个实例中,样品(例如,细胞、细胞群)可以在用药物和/或疗法治疗之前和/或之后随机标记。不同靶的拷贝数的变化可以指示样品对药物和/或疗法的反应。

[0118] 维度标记可以是可激活的。可以在特定时间点激活可激活的维度标记。可激活的标记可以被例如组成性地激活(例如,不关闭)。所述可激活的维度标记可以被例如可逆地激活(例如,所述可激活的维度标记可以打开和关闭)。所述维度标记可以被例如可逆地激活至少1、2、3、4、5、6、7、8、9、10次或更多次。所述维度标记可以被可逆地激活例如至少1、2、3、4、5、6、7、8、9、10次或更多次。在一些实施例中,可以用荧光、光、化学事件(例如,切割,另一种分子的连接,修饰的添加(例如,聚乙二醇化、sumo化、乙酰化、甲基化、去乙酰化、去甲基化)、光化学事件(例如,光锁定)以及引入非天然的核苷酸将所述维度标记激活。

[0119] 在一些实施例中,所述维度标记对于附接到给定固体支持物(例如合成颗粒,如珠)上的所有条形码(例如,随机条形码)可以是相同的,但对于不同的固体支持物(例如合成颗粒)是不同的。在一些实施例中,相同固体支持物上的至少60%、70%、80%、85%、90%、95%、97%、99%或100%的条形码可以包含相同的维度标记。在一些实施例中,相同固体支持物上的至少60%的条形码可以包括相同的维度标记。在一些实施例中,相同固体支持物上的至少95%的条形码可以包括相同的维度标记。

[0120] 多个固体支持物(例如,合成颗粒)可以表现多达 10^6 个或更多个独特维度标记序列。维度标记的长度可以是、或约1、2、3、4、5、10、15、20、25、30、35、40、45、50个核苷酸、或在这些值的任何两个之间的数字或范围的核苷酸。维度标记的长度可以是至少、或至多1、2、3、4、5、10、15、20、25、30、35、40、45、50、100、200、或300个核苷酸。维度标记可以包括在约5至约200个之间的核苷酸。维度标记可以包括在约10至约150个之间的核苷酸。维度标记可以包括长度在约20至约125个之间的核苷酸。

[0121] 空间标记

[0122] 条形码可以包括一个或多个空间标记。在一些实施例中,空间标记可以包括提供与条形码相关联的靶分子的空间取向的信息的核酸序列。空间标记可以与样品中的坐标相关联。所述坐标可以是固定的坐标。例如可以参考基底固定坐标。空间标记可以参考二维或三维网格。可以参考界标固定坐标。在空间中界标是可被鉴定的。界标可以是可被成像的结构。界标可以是生物学结构,例如解剖学界标。界标可以是细胞界标,例如细胞器。界标可以是非天然界标,诸如具有可鉴定标识(诸如色码、条形码、磁性、荧光、放射性或独特尺寸或形状)的结构。空间标记可以与物理分区(例如,孔、容器、微球、管、微胶囊或液滴)相关联。在一些实施例中,将多个空间标记一起用于编码在空间中的一个或多个位置。

[0123] 所述空间标记对于附接到给定固体支持物(例如,合成颗粒,如珠)上的所有条形码可以是相同的,但对于不同的固体支持物(例如,合成颗粒)是不同的。在一些实施例中,包括相同空间标记的相同固体支持物上的条形码的百分比可以是、或约60%、70%、80%、85%、90%、95%、97%、99%、100%、或在这些值的任何两个之间的数字或范围。在一些实施例中,包括相同空间标记的相同固体支持物上的条形码的百分比可以是至少、或至多60%、70%、80%、85%、90%、95%、97%、99%、或100%。在一些实施例中,相同固体支持物上的至少60%的条形码可以包括相同的空间标记。在一些实施例中,相同固体支持物上的至少95%的条形码可以包括相同的空间标记。

[0124] 多个固体支持物(例如,合成颗粒,如珠)可以表现多达 10^6 个或更多个独特空间标记序列。空间标记的长度可以是、或约1、2、3、4、5、10、15、20、25、30、35、40、45、50个核苷酸、或在这些值的任何两个之间的数字或范围的核苷酸。空间标记的长度可以是至少或至多1、2、3、4、5、10、15、20、25、30、35、40、45、50、100、200、或300个核苷酸。空间标记可以包括在约5至约200个之间的核苷酸。空间标记可以包括在约10至约150个之间的核苷酸。空间标记可以包括长度在约20至约125个之间的核苷酸。

[0125] 细胞标记

[0126] 条形码可以包括一个或多个细胞标记。在一些实施例中,细胞标记可以包括提供用于确定哪个靶核酸来自哪个细胞的信息的核酸序列。在一些实施例中,所述细胞标记对于所有附接到给定固体支持物(例如,合成颗粒,如珠)上的条形码是相同的,但对于不同的固体支持物(例如,合成颗粒)是不同的。在一些实施例中,包括相同细胞标记的相同固体支持物上的条形码的百分比可以是、或约60%、70%、80%、85%、90%、95%、97%、99%、100%、或在这些值的任何两个之间的数字或范围。在一些实施例中,包括相同细胞标记的相同固体支持物上的条形码的百分比可以是、或约60%、70%、80%、85%、90%、95%、97%、99%、或100%。例如,相同固体支持物上的至少60%的条形码可以包括相同的细胞标记。作为另一个实例,相同固体支持物上的至少95%的条形码可以包括相同的细胞标记。

[0127] 多个固体支持物(例如,合成颗粒)可以表现多达 10^6 个或更多个独特细胞标记序列。细胞标记的长度可以是、或约1、2、3、4、5、10、15、20、25、30、35、40、45、50个核苷酸、或在这些值的任何两个之间的数字或范围的核苷酸。细胞标记的长度可以是至少、或至多1、2、3、4、5、10、15、20、25、30、35、40、45、50、100、200、或300个核苷酸。例如,细胞标记可以包括约5至约200个之间的核苷酸。作为另一个实例,细胞标记可以包括约10至约150个之间的核苷酸。还作为另一个实例,细胞标记可以包括长度在约20至约125个之间的核苷酸。

[0128] 条形码序列

[0129] 条形码可以包括一个或多个条形码序列。在一些实施例中,条形码序列可以包括为与条形码杂交的特定类型的靶核酸种类提供鉴定信息的核酸序列。条形码序列可以包括如下核酸序列,所述核酸序列为与条形码(例如,靶结合区)杂交的靶核酸种类的特定出现提供计数器(例如,提供粗略近似)。

[0130] 在一些实施例中,将条形码序列的多元组附接到给定固体支持物(例如,合成颗粒,如珠)。在一些实施例中,可以有、或约有 10^2 、 10^3 、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、 10^9 个、或在这些值的任何两个之间的数字或范围的独特分子标记序列。例如,多个条形码可以包括具有不同序列的约6561个条形码序列。作为另一个实例,多个条形码可以包括具有不同序列的约65536个条形码序列。在一些实施例中,可以有至少、或至多 10^2 、 10^3 、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 或 10^9 个独特条形码序列。可以将所述独特分子标记序列附接到给定固体支持物(例如,合成颗粒)上。

[0131] 条形码的长度可以是、或是约1、2、3、4、5、10、15、20、25、30、35、40、45、50个、或在这些值的任何两个之间的数字或范围的核苷酸。条形码的长度可以是至少、或至多1、2、3、4、5、10、15、20、25、30、35、40、45、50、100、200、或300个核苷酸。

[0132] 分子标记

[0133] 条形码可以包括一个或多个分子标记。分子标记可以包括条形码序列。在一些实施例中,分子标记可以包括为与随机条形码杂交的特定类型的靶核酸种类提供鉴定信息的核酸序列。分子标记可以包括如下核酸序列,该核酸序列为与随机条形码(例如,靶结合区)杂交的靶核酸种类的特定出现提供计数器。

[0134] 在一些实施例中,将分子标记的多元组附接到给定固体支持物(例如,合成颗粒,如珠)。在一些实施例中,可以有、或约有 10^2 、 10^3 、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、 10^9 个、或许多或一定范围的独特分子标记序列。例如,多个随机条形码可包括具有不同序列的约6561个分子标记。作为另一个实例,多个随机条形码可包括具有不同序列的约65536个分子标记。在一些实施例中,可以有至少、或至多 10^2 、 10^3 、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、或 10^9 个独特分子标记序列。可以将具有独特分子标记序列的随机条形码附接到给定固体支持物(例如,合成颗粒)。

[0135] 对于使用多个随机条形码的随机条形码化,不同分子标记序列的数量与任何靶的出现的数量的比率可以是、或约1:1、2:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1、10:1、11:1、12:1、13:1、14:1、15:1、16:1、17:1、18:1、19:1、20:1、30:1、40:1、50:1、60:1、70:1、80:1、90:1、100:1、或在这些值的任何两个之间的数字或范围。靶可以是包括具有相同或几乎相同序列的mRNA分子的mRNA种类。在一些实施例中,不同分子标记序列的数目与任何靶的出现次数的比率是至少、或至多1:1、2:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1、10:1、11:1、12:1、13:1、14:1、15:1、16:1、17:1、18:1、19:1、20:1、30:1、40:1、50:1、60:1、70:1、80:1、90:1、或100:1。

[0136] 分子标记的长度可以是或是约1、2、3、4、5、10、15、20、25、30、35、40、45、50个、或在这些值的任何两个之间的数字或范围的核苷酸。分子标记的长度可以是至少、或至多1、2、3、4、5、10、15、20、25、30、35、40、45、50、100、200、或300个核苷酸。

[0137] 靶结合区

[0138] 条形码可以包含一个或多个靶结合区,诸如捕获探针。在一些实施例中,靶结合区可以与目的靶杂交。在一些实施例中,靶结合区可以包含与靶(例如,靶核酸,如待分析的细

胞核酸) 特异性杂交的核酸序列。例如, 靶结合区可以与靶核酸在特定的基因序列上杂交。在一些实施例中, 靶结合区可以包含可附接(例如, 杂交) 至特定靶核酸的特定位置的核酸序列。在一些实施例中, 靶结合区可以包含能够与限制性酶位点突出端(例如, EcoRI 粘性末端突出端) 进行特异性杂交的核酸序列。然后条形码可以连接到包括与限制性位点突出端互补的序列的任何核酸分子。

[0139] 在一些实施例中, 靶结合区可以包含非特异性靶核酸序列。非特异性靶核酸序列可以指独立于靶核酸的特定序列可与多个靶核酸结合的序列。例如, 靶结合区可以包含与 mRNA 分子上的聚(A) 尾杂交的随机多聚体序列或寡聚(dT) 序列。随机多聚体序列可以是例如随机二聚体、三聚体、四聚体、五聚体、六聚体、七聚体、八聚体、九聚体、十聚体或任何长度的更高多聚体序列。在一些实施例中, 对于附接到给定合成颗粒(例如, 珠) 的所有条形码, 所述靶结合区是相同的。在一些实施例中, 对于附接到给定合成颗粒上的多个条形码, 靶结合区可以包含两个或更多个不同的靶结合序列。靶结合区的长度可以是或约5、10、15、20、25、30、35、40、45、50个核苷酸、或在这些值的任何两个之间的数字或范围的核苷酸。靶结合区的长度可以是至多约5、10、15、20、25、30、35、40、45、50个或更多个核苷酸。

[0140] 在一些实施例中, 靶结合区可以包括寡聚(dT), 所述寡聚(dT) 可以与包括聚腺苷酸化末端的mRNA杂交。靶结合区可以是基因特异性的。例如, 可以将靶结合区配置为与靶的特定区域杂交。靶结合区的长度可以是或是约1、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个核苷酸、或在这些值的任何两个之间的数字或范围的核苷酸。靶结合区的长度可以是至少、或至多1、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、或30个核苷酸。靶结合区的长度可以是约5至30个核苷酸。当条形码包括基因特异性靶结合区时, 所述条形码可以称为基因特异性条形码。

[0141] 在一些实施例中, 条形码不包括靶结合区。在一些实施例中, 条形码包括与具有如下序列的靶结合区相对应的区, 所述序列对样品的一个或多个细胞中的一些、大多数、基本上所有或所有mRNA分子具有低结合亲和力(例如, 不结合)。例如, 条形码可以包含与具有不与目的mRNA分子结合的序列的靶结合区相对应的区。如果第一条条形码的靶结合区包括能够与mRNA分子的聚(A) 尾杂交的寡聚(dT) 序列, 第二条条形码的相对应的区可以包括例如, 与聚(dT) 序列不相似或基本相似的序列。如果第一条条形码的靶结合区包含具有能够与特定基因序列特异性杂交的序列的靶结合区, 第二条条形码的相对应的区可以包括例如与靶结合区不相似或基本上相似的序列。

[0142] 定向特性

[0143] 条形码可以包括一种或多种可用于定向(例如, 比对) 条形码的定向特性。条形码可以包括用于等电聚焦的部分。不同的条形码可以包含不同的等电聚焦点。当将这些条形码被引入样品中时, 所述样品可以经历等电聚焦, 以便于将所述条形码定位成已知的方式。以这种方式, 所述定向特性可以用于开发样品中条形码的已知的映射。示例性定向特性可以包括电泳迁移率(例如, 基于条形码的尺寸)、等电点、自旋、电导率和/或自组装。例如, 条形码具有自组装的定向特性, 当激活时可以自组装成特定定向(例如, 核酸纳米结构)。

[0144] 亲和特性

[0145] 条形码可以包括一种或多种亲和特性。例如, 空间标记可以包括亲和特性。亲和特

性可以包括化学和/或生物部分,所述特性可以促进所述条形码与另一种实体(例如,细胞受体)的结合。例如,亲和特性可以包括抗体,例如,对于样品上的具体部分(例如,受体)特异性的抗体。在一些实施例中,抗体可以将条形码引导到特定细胞类型或分子上。在特定细胞类型或分子处的和/或附近的靶可以被随机标记。在一些实施例中,除了空间标记的核苷酸序列,亲和特性可以提供空间信息,因为所述抗体可以将所述条形码引导至特定位置。抗体可以是治疗性抗体,例如单克隆抗体或多克隆抗体。抗体可以是人源化的或嵌合的。抗体可以是裸抗体或融合抗体。

[0146] 抗体可以是全长(即,天然存在的或通过正常免疫球蛋白基因片段重组过程形成的)免疫球蛋白分子(例如,IgG抗体)或免疫球蛋白分子的免疫活性(即,特异性结合)部分(像抗体片段)。

[0147] 抗体片段可以是例如抗体的一部分,诸如F(ab')₂、Fab'、Fab、Fv、sFv等。在一些实施例中,抗体片段可以与由全长抗体识别的相同的抗原结合。抗体片段可以包括由抗体的可变区组成的分离的片段,诸如由重链和轻链的可变区组成的“Fv”片段和其中轻链和重链可变区通过肽接头连接的重组单链多肽分子(“scFv蛋白”)。示例性抗体可以包括但不限于癌细胞抗体、病毒抗体、结合至细胞表面受体(CD8、CD34、CD45)的抗体、和治疗性抗体。

[0148] 通用衔接子引物

[0149] 条形码可以包括一个或多个通用衔接子引物。例如,基因特异性条形码(诸如基因特异性随机条形码)可以包括通用衔接子引物。通用衔接子引物可以指在所有条形码上通用的核苷酸序列。通用衔接子引物可以用于构建基因特异性条形码。通用衔接子引物的长度可以是、或是约1、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个核苷酸、或在这些值的任何两个之间的数字或范围的核苷酸。通用衔接子引物的长度可以是至少、或至多1、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、或30个核苷酸。通用衔接子引物的长度可以是5至30个核苷酸。

[0150] 接头

[0151] 当条形码包括多于一个类型的标记(例如,多于一个细胞标记或多于一个条形码序列,如一个分子标记)时,这些标记可以穿插着接头标记序列。接头标记序列的长度可以是至少约5、10、15、20、25、30、35、40、45、50个或更多个核苷酸。接头标记序列的长度可以是至多约5、10、15、20、25、30、35、40、45、50个或更多个核苷酸。在一些情况下,接头标记序列的长度是12个核苷酸。可以将接头标记序列用于促进条形码的合成。该接头标记可以包括错误校正(例如,汉明)代码。

[0152] 固体支持物

[0153] 在一些实施例中,本文披露的条形码(诸如随机条形码)可以与固体支持物相关联。例如,固体支持物可以是颗粒或合成颗粒。在一些实施例中,固体支持物上的多个条形码(例如,第一多个条形码)的一些或所有条形码序列(如,随机条形码(例如,第一条形码序列)的分子标记)具有至少一个核苷酸的差异。相同固体支持物上的条形码的细胞标记可以是相同的。不同的固体支持物上的条形码的细胞标记可以具有至少一个核苷酸的差异。例如,第一固体支持物上的第一多个条形码的第一细胞标记可以具有相同的序列,且第二固体支持物上的第二多个条形码的第二细胞标记可以具有相同的序列。第一固体支持物上的

第一多个条形码的第一细胞标记和第二固体支持物上的第二多个条形码的第二细胞标记可以具有至少一个核苷酸的差异。细胞标记可以是例如约5至20个核苷酸长。条形码序列可以是例如约5至20个核苷酸长。合成颗粒可以是例如珠。

[0154] 合成颗粒可以是例如硅胶珠、可控孔径玻璃珠、磁珠、Dynabead、交联葡聚糖/琼脂糖凝胶珠、珠状纤维素、聚苯乙烯珠、或其任何组合。合成颗粒可以包含诸如聚二甲基硅氧烷(PDMS)、聚苯乙烯、玻璃、聚丙烯、琼脂糖、明胶、水凝胶、顺磁物质、陶瓷、塑料、玻璃、甲基苯乙烯、丙烯酸聚合物、钛、乳胶、琼脂糖凝胶、纤维素、尼龙、硅酮、或其任何组合的材料。

[0155] 在一些实施例中,合成颗粒可以是高分子颗粒(例如,可变形的颗粒或凝胶颗粒),其用条形码或随机条形码官能化(诸如来自10X基因组公司(10X Genomics)(旧金山,加利福尼亚)的凝胶珠)。在一些实现方式中,凝胶颗粒可以包含一种或多种基于聚合物的凝胶。例如,可以通过将一种或多种聚合物前体包封进液滴来产生凝胶颗粒。在将聚合物前体暴露于促进剂(例如,四甲基乙二胺(TEMED))后,可以产生凝胶颗粒。

[0156] 在一些实施例中,颗粒可以是可降解的。例如,高分子颗粒或珠可以例如在所期望的条件下溶解、融化或降解。所期望的条件可以包括环境条件。所期望的条件可导致高分子颗粒以受控方式溶解、融化或降解。由于化学刺激、物理刺激、生物刺激、热刺激、磁刺激、电刺激、光刺激或其任何组合,凝胶颗粒可以溶解、融化或降解。

[0157] 试剂(诸如寡核苷酸条形码)例如可以偶联/固定到合成颗粒的内表面(寡核苷酸条形码和/或用于产生寡核苷酸条形码的材料的扩散的可接近内部)和/或凝胶颗粒的外表面或本文描述的任何其他微胶囊。颗粒可以是例如凝胶珠。相关联(例如,偶联/固定)可以经由任何形式的化学键(例如,共价键、离子键)或物理现象(例如,范德华力、偶极-偶极相互作用等)。在一些实施例中,试剂与颗粒或本文描述的任何其他固体支持物(例如,微胶囊)的相关联(例如,偶联/固定)可以是可逆的,例如经由不稳定部分(例如,经由化学交联剂,包括本文描述的化学交联剂)。在施加刺激后,不稳定部分可以被切割并释放固定化的试剂。在一些实施例中,不稳定部分是二硫键。例如,在经由二硫键将寡核苷酸条形码固定到凝胶颗粒上的情况下,将二硫键暴露于还原剂可以切割二硫键并从颗粒释放寡核苷酸条形码。不稳定部分可以作为凝胶颗粒、珠或微胶囊的一部分、作为将试剂与凝胶珠或微胶囊连接的化学接头的一部分、和/或作为试剂的一部分包括在内。在一些实施例中,多个条形码的至少一个条形码可固定在颗粒上、部分固定在颗粒上、包封在颗粒中、部分包封在颗粒中、或其任何组合。

[0158] 在一些实施例中,颗粒(例如,凝胶珠)可以包含多种不同的聚合物,包括但不限于:聚合物、热敏聚合物、光敏聚合物、磁性聚合物、pH敏感聚合物、盐敏感聚合物、化学敏感聚合物、聚电解质、多糖、肽、蛋白质和/或塑料。聚合物可包括但不限于以下材料:如聚(N-异丙基丙烯酰胺)(PNIPAAm)、聚(苯乙烯磺酸酯)(PSS)、聚(烯丙基胺)(PAAm)、聚(丙烯酸)(PAA)、聚(乙烯亚胺)(PEI)、聚(双烯丙基二甲基-氯化铵)(PDADMAC)、聚(吡咯)(PPy)、聚(乙基吡咯烷酮)(PVPON)、聚(乙基吡啶)(PVP)、聚(甲基丙烯酸甲酯)(PMAA)、聚(甲基丙烯酸甲酯)(PMMA)、聚苯乙烯(PS)、聚(四氢呋喃)(PTHF)、聚(邻苯二甲醛)(PTHF)、聚(己基紫精)(PHV)、聚(L-赖氨酸)(PLL)、聚(L-精氨酸)(PARG)、聚(乳酸-聚羟基乙酸)(PLGA)。

[0159] 许多化学刺激可用于触发颗粒的破坏、溶解、或降解。这些化学改变的实例可包括但不限于pH介导的颗粒壁改变、经由交联键的化学裂解使颗粒壁分解、颗粒壁的触发解聚、

和颗粒壁转换反应。批量改变也可用于触发颗粒的破坏。

[0160] 通过各种刺激对微胶囊或颗粒的批量或物理变化在设计胶囊以释放试剂方面也提供了许多优点。在宏观尺度上发生批量或物理变化,其中颗粒破裂是由刺激引起的机械-物理力的结果。这些过程可包括但不限于压力引起的破裂、颗粒壁熔化、或颗粒壁的孔隙率的改变。

[0161] 生物刺激也可用于触发颗粒的破坏、溶解、或降解。通常,生物触发剂类似于化学触发剂,但是许多实例使用生物分子、或生命系统中常见的分子,如酶、肽、糖、脂肪酸、核酸等。例如,颗粒可包含具有肽交联的聚合物,所述肽交联通过特定蛋白酶对切割敏感。更具体地,一个实例可包括含有GFLGK肽交联的微胶囊。在添加生物触发物(如蛋白酶组织蛋白酶B)后,壳孔的肽交联被切割且颗粒的内容物被释放。在其他情况下,蛋白酶可以是热激活的。在另一个实例中,颗粒包括含有纤维素的壳壁。水解酶壳聚糖的添加用作纤维素键裂解、壳壁解聚、和内部内容物释放的生物触发剂。

[0162] 还可以在施加热刺激后诱导颗粒释放其内容物。温度的变化可导致颗粒的各种变化。热量的变化可能导致颗粒熔化,使得颗粒壁崩解。在其他情况下,热量可能增加颗粒内部组分的内部压力,使得颗粒破裂或爆炸。在仍其他情况下,热量可以使颗粒变成收缩的脱水状态。热量还可以作用于颗粒壁内的热敏聚合物,从而引起颗粒的破坏。

[0163] 将磁性纳米颗粒包括在微胶囊的颗粒壁中可以允许颗粒的触发破裂以及将颗粒引导成阵列。本披露的装置可以包含用于任一目的的磁性颗粒。在一个实例中,将 Fe_3O_4 纳米颗粒并入含聚电解质的颗粒中在振荡磁场刺激的存在下触发破裂。

[0164] 由于电刺激的结果,颗粒也可能被破坏、溶解、或降解。与先前部分中描述的磁性颗粒相似,电敏颗粒可以允许颗粒的触发破裂以及其他功能,如电场中的对准、电导率或氧化还原反应。在一个实例中,含有电敏材料的颗粒在电场中排列,从而可以控制内部试剂的释放。在其他实例中,电场可以在颗粒壁本身内引起氧化还原反应,这可以增加孔隙率。

[0165] 也可用光刺激来破坏颗粒。许多光触发是可能的,并可以包括使用各种分子(如能够吸收特定波长范围的光子的纳米颗粒和发色团)的系统。例如,金属氧化物涂层可用作胶囊触发剂。涂覆有 SiO_2 的聚电解质胶囊的UV照射可导致颗粒壁的崩解。在又一个实例中,可以将可光切换材料(如偶氮苯基团)并入颗粒壁中。在施加UV或可见光后,如这些的化学物质在吸收光子后经历可逆的顺式-反式异构化。在此方面,光子切换的并入导致颗粒壁在施加光触发剂后可崩解或变得更多孔。

[0166] 例如,在图2中展示的条形码化(例如,随机条形码化)的非限制性实例中,在框208处将细胞(诸如单细胞)引入微孔阵列的多个微孔之后,在框212处可以将颗粒引入微孔阵列的多个微孔。每个微孔可以包含颗粒。颗粒可以包含多个条形码。条形码可以包含附接到颗粒的5' 胺区。条形码可以包括通用标记、条形码序列(例如,分子标记)、靶结合区、或其任何组合。

[0167] 本文披露的条形码可以与固体支持物(例如,颗粒或珠)相关联(例如,附接)。与固体支持物关联的条形码每个可包括选自下组的条形码序列,该组由以下组成:具有独特序列的至少100或1000个条形码序列。在一些实施例中,与固体支持物关联的不同条形码可包含不同序列的条形码序列。在一些实施例中,与固体支持物相关联的条形码的百分比包括相同的细胞标记。例如,所述百分比可以是、或是约60%、70%、80%、85%、90%、95%、

97%、99%、100%、或在这些值的任何两个之间的数字或范围。作为另一个实例,该百分比可以是至少、或至多60%、70%、80%、85%、90%、95%、97%、99%、或100%。在一些实施例中,与固体支持物相关联的条形码可以具有相同的细胞标记。与不同固体支持物相关联的条形码可以具有选自下组的不同的细胞标记,该组由以下组成:具有独特序列的至少100或1000个细胞标记。

[0168] 本文披露的条形码可以与固体支持物(例如,颗粒,如珠)相关联(例如,附接)。在一些实施例中,可以用包括与多个条形码关联的多个合成的颗粒的固体支持物对样品中的多个靶进行随机地条形码化。在一些实施例中,固体支持物可包括与多个条形码相关联的多个合成的颗粒。不同固体支持物上的多个条形码的空间标记可以具有至少一个核苷酸的差异。固体支持物例如可以包括处于二维或三维的多个条形码。合成的颗粒可以是珠。颗粒可以是硅胶珠、可控孔径玻璃珠、磁珠、Dynabead、交联葡聚糖/琼脂糖珠、珠状纤维素、聚苯乙烯珠、或其任何组合。固体支持物可包括聚合物、基质、水凝胶、针阵列装置、抗体、或其任何组合。在一些实施例中,固体支持物可以自由浮动。在一些实施例中,固体支持物可嵌入半固体或固体阵列中。条形码可以不与固体支持物相关联。条形码可以是单独的核苷酸。条形码可与基底相关联。

[0169] 如本文使用的,术语“拴系的”、“附接的”和“固定的”可互换使用,并且可以指用于将条形码附接到固体支持物上的共价或非共价方式。可以将多种不同的固体支持物中的任何一种用作固体支持物,以用于附接预先合成的条形码或用于条形码的原位固相合成。

[0170] 在一些实施例中,固体支持物是颗粒,例如珠。颗粒可以包含一种或多种类型的实心的、多孔的或空心的球体、球、承座、圆柱体或其他相似配置,其上可以固定核酸(例如,共价地或非共价地)。颗粒可以例如由塑料、陶瓷、金属、聚合物材料、或其任何组合构成。颗粒可以是、或包括球形的(例如,微球)或具有非球形或不规则形状的离散颗粒,所述形状是诸如立方形、长方形、锥形、圆柱形、圆锥形、椭圆形或圆盘形等。在一些实施例中,颗粒的形状可以是非球形的。

[0171] 所述颗粒可以包含多种材料,包括但不限于顺磁性材料(例如,镁、钼、锂和钽)、超顺磁性材料(例如,铁氧体(Fe_3O_4 ;磁铁矿)纳米颗粒)、铁磁材料(例如,铁、镍、钴,其一些合金,以及一些稀土金属化合物)、陶瓷、塑料、玻璃、聚苯乙烯、二氧化硅、甲基苯乙烯、丙烯酸聚合物、钛、胶乳、交联琼脂糖、琼脂糖、水凝胶、聚合物、纤维素、尼龙、或其任何组合。

[0172] 在一些实施例中,颗粒(例如,标记所附接的颗粒)是水凝胶珠。在一些实施例中,所述颗粒包含水凝胶。

[0173] 本文披露的一些实施例包含一个或多个颗粒(例如珠)。颗粒的每个可包含多个寡核苷酸(例如,条形码)。多个寡核苷酸的每个可包含条形码序列(例如,分子标记)、细胞标记、和靶结合区(例如,寡聚(dT)序列、基因特异性序列、随机多聚体、或其组合)。所述多个寡核苷酸的每个的细胞标记序列可以是相同的。不同颗粒上的寡核苷酸的细胞标记序列可以是不同的,使得可以鉴定不同颗粒上的寡核苷酸。在不同实现方式中,不同细胞标记序列的数量可以是不同的。在一些实施例中,细胞标记序列的数量可以是、或是约10、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、10000、20000、30000、40000、50000、60000、70000、80000、90000、100000、 10^6 、 10^7 、 10^8 、 10^9 、在这些值的任何两个之间的数字或范围、或更多。在一些实施例中,细胞标记序列的数目可

以是至少、或至多10、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、10000、20000、30000、40000、50000、60000、70000、80000、90000、100000、 10^6 、 10^7 、 10^8 、或 10^9 。在一些实施例中,多个颗粒中不超过1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、或更多个包括具有相同细胞序列的寡核苷酸。在一些实施例中,包括具有相同细胞序列的寡核苷酸的多个颗粒可以是至多0.1%、0.2%、0.3%、0.4%、0.5%、0.6%、0.7%、0.8%、0.9%、1%、2%、3%、4%、5%、6%、7%、8%、9%、10%或更多。在一些实施例中,多个颗粒中没有颗粒具有相同的细胞标记序列。

[0174] 在每个颗粒上的多个寡核苷酸可以包含不同的条形码序列(例如,分子标记)。在一些实施例中,条形码序列的数量可以是、或约10、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、10000、20000、30000、40000、50000、60000、70000、80000、90000、100000、 10^6 、 10^7 、 10^8 、 10^9 、或在这些值的任何两个之间的数字或范围。在一些实施例中,条形码序列的数目可以是至少、或至多10、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、10000、20000、30000、40000、50000、60000、70000、80000、90000、100000、 10^6 、 10^7 、 10^8 、或 10^9 。例如,多个寡核苷酸中的至少100个包含不同的条形码序列。作为另一个实例,在单个颗粒中,多个寡核苷酸中的至少100、500、1000、5000、10000、15000、20000、50000个、这些值的任何两个之间的数字或范围、或更多个包含不同的条形码序列。一些实施例提供了包含条形码的多个颗粒。在一些实施例中,待标记的靶和不同条形码序列的出现(或拷贝或数量)的比率可以是至少1:1、1:2、1:3、1:4、1:5、1:6、1:7、1:8、1:9、1:10、1:11、1:12、1:13、1:14、1:15、1:16、1:17、1:18、1:19、1:20、1:30、1:40、1:50、1:60、1:70、1:80、1:90、或更高。在一些实施例中,多个寡核苷酸的每个进一步包括样品标记、通用标记、或两者。颗粒例如可以是纳米颗粒或微颗粒。

[0175] 颗粒的尺寸可以改变。例如,颗粒的直径范围可以从0.1微米至50微米。在一些实施例中,颗粒的直径可以是、或约是0.1、0.5、1、2、3、4、5、6、7、8、9、10、20、30、40、50微米、或在这些值的任何两个之间的数字或范围。

[0176] 颗粒的直径可以与基底的孔的直径相关。在一些实施例中,相比孔的直径,颗粒的直径可以更长或更短、或约更长或更短10%、20%、30%、40%、50%、60%、70%、80%、90%、100%、或在这些值的任何两个之间的数字或范围。颗粒的直径可以与细胞(例如,由基底的孔截留的单个细胞)的直径有关。在一些实施例中,相比孔的直径,颗粒的直径可以是更长或更短至少或至多10%、20%、30%、40%、50%、60%、70%、80%、90%或100%。颗粒的直径可以与细胞(例如,由基底的孔截留的单个细胞)的直径有关。在一些实施例中,相比细胞的直径,颗粒的直径可以更长或更短、或约更长或更短10%、20%、30%、40%、50%、60%、70%、80%、90%、100%、150%、200%、250%、300%、或在这些值的任何两个之间的数字或范围。在一些实施例中,相比细胞的直径,颗粒的直径可以是更长或更短至少或至多10%、20%、30%、40%、50%、60%、70%、80%、90%、100%、150%、200%、250%、或300%。

[0177] 颗粒可以附接到和/或包埋在基底中。可以将颗粒附接到和/或嵌入凝胶、水凝胶、聚合物和/或基质中。使用存在于颗粒中可以充当位置地址的条形码上的空间标记,可以鉴定在基质(例如,凝胶、基质、支架,或聚合物)中颗粒的空间位置。

[0178] 颗粒的实例可以包含但不限于链霉亲和素珠、琼脂糖珠、磁珠、Dynabeads®、MACS®微珠、缀合抗体的珠(例如,抗免疫球蛋白微珠)、缀合A蛋白的珠、缀合G蛋白的珠、缀合A/G蛋白的珠、缀合L蛋白的珠、缀合寡聚(dT)的珠、二氧化硅珠、二氧化硅样珠、抗生物素微珠、抗荧光染料微珠以及BcMag™羧基封端的磁珠。

[0179] 颗粒可以关联有(例如浸渍有)量子点或荧光染料,以使其在一个荧光光通道或多个光通道中是荧光的。颗粒可以关联有氧化铁或氧化铬,以使其具有顺磁性或铁磁性。颗粒是可被鉴定的。例如,使用照相机可以将颗粒成像。颗粒可以具有与所述颗粒相关联的可检测的代码。例如,颗粒可以包含条形码。颗粒的尺寸可以变化,例如由于在有机或无机溶液中的溶胀。颗粒可以是疏水的。颗粒可以是亲水的。颗粒可以是生物相容的。

[0180] 可以使固体支持物(例如,颗粒)可视化。固体支持物可以包括可视化标签(例如,荧光染料)。可以用标识(例如,数字)将固体支持物(例如,颗粒)蚀刻。通过对颗粒成像可以将所述标识可视化。

[0181] 固体支持物可以包括不溶性、半溶性或不溶性材料。当固体支持物包括接头、支架、结构单元或附接至其上的其他反应性部分时,它可以被称为“官能化的”,而当固体支持物缺少附接至其上的这样一个反应性部分时,它可以被称为“非官能化的”。固体支持物可以在溶液中不受约束地利用,如以微量滴定孔形式;以流通形式,如在柱中;或在试纸条(dipstick)中。

[0182] 固体支持物可以包括膜、纸、塑料、涂覆的表面、平表面、玻璃、载玻片、芯片、或其任何组合。固体支持物可以采用树脂、凝胶、微球或其他几何构型的形式。固体支持物可以包含二氧化硅芯片;微颗粒;纳米颗粒;平板;阵列;毛细管;平支持物,如玻璃纤维过滤器,玻璃表面,金属表面(钢、金、银、铝、硅以及铜),玻璃支持物,塑料支持物,硅支持物,芯片,过滤器,膜,微孔板,载玻片;塑料材料包括多孔板或膜(例如,由聚乙烯、聚丙烯、聚酰胺、聚偏二氟乙烯形成);和/或晶片;梳状物(comb);针或针头(例如,适于组合合成或分析的针阵列);或平表面(如晶片(例如,硅晶片)、带有具有或不具有滤底的凹陷的晶片的凹陷或纳升孔的阵列中的颗粒(例如合成颗粒,如珠)。

[0183] 所述固体支持物可以包括聚合物基质(例如,凝胶、水凝胶)。该聚合物基质可能能够渗透细胞内间隙(例如,细胞器周围)。该聚合物基质可能能够贯穿循环系统进行泵送。

[0184] 固体支持物可以是生物分子。例如,固体支持物可以是或可以包含核酸、蛋白质、抗体、组蛋白、细胞区室、脂质、碳水化合物等。作为生物分子的固体支持物可被扩增、翻译、转录、降解和/或修饰(例如,聚乙二醇化、sumo化、乙酰化、甲基化)。除了附接至生物分子的空间标记之外,作为生物分子的固体支持物可以提供空间和时间信息。例如,生物分子可以在未修饰时包括第一构象,但是在修饰时可以改变为第二构象。这些不同的构象可以将本披露的条形码(例如,随机条形码)暴露给靶。例如,生物分子可以包括由于生物分子的折叠而不可接近的条形码。在修饰生物分子(例如,乙酰化)时,这些生物分子可以改变构象以暴露这些条形码。修饰的时间设置可以为本披露的条形码化的方法提供另一时间维度。

[0185] 在一些实施例中,包括本披露的条形码试剂的生物分子可以位于细胞的细胞质中。激活时,该生物分子可以移动到细胞核,于此可以进行条形码化。以这种方式,所述生物分子的修饰可以编码由条形码鉴别的靶的另外的空间-时间信息。

[0186] 基底和微孔阵列

[0187] 如本文所用,基底可以指一种固体支持物。基底可以指可包括本披露的条形码和随机条形码的固体支持物。例如,基底可以包括多个微孔。例如,基底可以是包括两个或更多个微孔的孔阵列。在一些实施例中,微孔可以包括具有确定体积的小反应室。在一些实施例中,微孔可以截留一个或多个细胞。在一些实施例中,微孔只能截留一个细胞。在一些实施例中,微孔可以截留一个或多个固体支持物。在一些实施例中,微孔只能截留一个固体支持物。在一些实施例中,微孔截留单个细胞和单个固体支持物(例如,颗粒)。微孔可以包括本披露的组合条形码试剂。

[0188] 在固体支持物上的条形码的合成

[0189] 可以在固体支持物(例如,颗粒,如合成颗粒或珠)上合成条形码(例如,随机条形码)。通过多种固定技术(涉及在固体支持物和条形码上的官能团对)中的任一种,可以将预先合成的条形码(例如,包含可以与固体支持物连接的5'胺)附接到固体支持物(例如,颗粒)上。条形码可以包含官能团。固体支持物(例如,颗粒)可以包含官能团。所述条形码官能团和所述固体支持物官能团可以包含例如,生物素、链霉亲和素、一个或多个伯胺、一个或多个羧基、一个或多个羟基、一种或多种醛、一个或多个酮及其任何组合。可以通过将条形码上的5'氨基基团偶联(例如,使用1-乙基-3-(3-二甲基氨基丙基)碳二亚胺)至所述官能化的固体支持物的羧基基团来将条形码(例如,随机条形码)拴系于固体支持物。可以通过进行多次漂洗步骤从所述反应混合物中除去残留的未偶联的条形码。在一些实施例中,所述条形码和固体支持物通过接头分子(例如,短的官能化的烃分子或聚环氧乙烷分子)使用类似的附接化学而间接附接。所述接头可以是可切割的接头,例如酸不稳定接头或可光切割接头。

[0190] 可以使用多种固相寡核苷酸合成技术(例如,磷酸二酯合成、磷酸三酯合成、亚磷酸三酯合成和亚磷酰胺合成)中的任一种在固体支持物(例如,颗粒)上合成条形码(例如,随机条形码)。单一核苷酸能以逐步的方式与生长中的拴系寡核苷酸偶联。若干寡核苷酸的短的预合成序列(或嵌段)可以与生长中的拴系条形码偶联。

[0191] 条形码(例如,随机条形码)可以通过散布逐步或嵌段偶联反应用一轮或多轮分离-聚池合成进行合成,其中合成颗粒的总池被分成多个单独的较小的池,然后其各自经受不同的偶联反应,随后重组和混合单独的池,以使生长中的条形码序列跨整个颗粒池随机化。分离-聚池合成是组合合成过程的实例,其中使用最少数量的化学偶联步骤合成最多数量的化学化合物。由此产生的化合物文库的潜在多样性由可用于每个偶联步骤的唯一性结构单元(例如核苷酸)的数目以及用于产生所述文库的偶联步骤的数目决定。例如,在每个步骤包括10轮偶联使用4种不同核苷酸的分离-聚池合成将产生 $4^{10} = 1,048,576$ 个唯一性核苷酸序列。在一些实施例中,分离-聚池合成可以使用酶促方法(例如聚合酶延伸或连接反应)而不是化学偶联来进行。例如,在每轮分离-聚池聚合酶延伸反应中,在给定的池中拴系于颗粒的条形码的3'末端可以与一组半随机引物的5'末端杂交,例如具有 $5' - (M)_k - (X)_i - (N)_j - 3'$ 结构的引物,其中 $(X)_i$ 是*i*个核苷酸长的随机核苷酸序列(所述引物组包括 $(X)_i$ 的所有可能的组合), $(N)_j$ 是特定核苷酸(或*j*个核苷酸的系列),并且 $(M)_k$ 是特定核苷酸(或*k*个核苷酸的系列),其中将不同的脱氧核糖核苷酸三磷酸(dNTP)添加至每个池并通过聚合酶并入拴系寡核苷酸中。

[0192] 条形码化的方法

[0193] 本披露提供了如下方法,所述方法用于评估样品中不同靶的数量。在一些实施例中,对多个靶进行条形码化包括将多个条形码与多个靶杂交以创建经条形码化的靶(例如,经随机条形码化的靶)。对多个靶进行条形码化可以包括产生经条形码化的靶的索引文库。产生经条形码化的靶的索引文库可以用包括多个条形码(例如,随机条形码)的固体支持物进行。

[0194] 使样品和条形码接触

[0195] 本披露提供了用于使样品(例如,细胞)与本披露的基底接触的方法。可以使包括例如细胞、器官或组织薄片的样品与条形码(例如,随机条形码)接触。例如,通过重力流可以使所述细胞接触,其中可以使所述细胞沉淀并且产生单层细胞。所述样品可以是组织薄片。可以将薄片置于基底上。该样品可以是一维的(例如,形成平面)。可以将所述样品(例如,细胞)涂布于基底上,例如,通过在基底上生长/培养所述细胞。

[0196] 当条形码靠近靶时,靶可以与条形码进行杂交。条形码可以按不可耗尽的比率接触,使得每个不同的靶可以与本披露的不同条形码相关联。为了确保靶与条形码之间的有效关联,可以将靶与条形码交联。

[0197] 细胞裂解

[0198] 在细胞和条形码的分布之后,可以裂解细胞以释放靶分子。细胞裂解可以通过多种手段中的任何一种来完成,例如通过化学或生化手段,通过渗透冲击,或通过热裂解、机械裂解或光学裂解。可以通过添加包括洗涤剂(例如SDS、十二烷基硫酸锂、Triton X-100、Tween-20或NP-40)的细胞裂解缓冲液、有机溶剂(例如甲醇或丙酮)或消化酶(例如蛋白酶K、胃蛋白酶或胰蛋白酶)或其任何组合来裂解细胞。为了增加靶和条形码的关联,可通过例如降低裂解物的温度和/或增加裂解物的粘度来改变靶分子的扩散速率。

[0199] 在一些实施例中,可以使用滤纸将样品裂解。可以用在滤纸上部的裂解缓冲液浸渍该滤纸。在压力下可以将滤纸应用于样品,该压力可以促进样品的裂解,以及样品的靶与基底的杂交。

[0200] 在一些实施例中,裂解可以通过机械裂解、热裂解、光学裂解、和/或化学裂解来进行。化学裂解可以包括使用消化酶类,如蛋白酶K、胃蛋白酶、以及胰蛋白酶。可以通过将裂解缓冲液添加到基底中进行裂解。裂解缓冲液可以包括Tris HCl。裂解缓冲液可以包含至少约、或至多约0.01、0.05、0.1、0.5或1M或更多或更少的Tris HCl。裂解缓冲液可以包含约0.01、0.05、0.1、0.5或1M Tris HCl。裂解缓冲液的pH可以至少约1、2、3、4、5、6、7、8、9、10、或更高。裂解缓冲液的pH可以是至多约1、2、3、4、5、6、7、8、9、10、或更高。在一些实施例中,该裂解缓冲液的pH是约7.5。该裂解缓冲液可以包括盐(例如,LiCl)。在裂解缓冲液中盐的浓度可以是至少约0.1M、0.5M、或1M、或更高。在裂解缓冲液中盐的浓度可以是至多约0.1M、0.5M、或1M、或更高。在一些实施例中,在裂解缓冲液中盐的浓度是约0.5M。裂解缓冲液可以包括洗涤剂(例如,SDS、十二烷基硫酸锂、曲通X、tween、NP-40)。在裂解缓冲液中洗涤剂的浓度可以是至少约0.0001%、0.0005%、0.001%、0.005%、0.01%、0.05%、0.1%、0.5%、1%、2%、3%、4%、5%、6%、或7%、或更高。在裂解缓冲液中洗涤剂的浓度可以是至多约0.0001%、0.0005%、0.001%、0.005%、0.01%、0.05%、0.1%、0.5%、1%、2%、3%、4%、5%、6%、或7%、或更高。在一些实施例中,在裂解缓冲液中洗涤剂的浓度是约1%十二烷基硫酸锂。该裂解方法中所用时间可以依赖于所用洗涤剂的量。在一些实施例中,所用洗涤剂

越多,裂解所需时间越短。裂解缓冲液可以包括螯合剂(例如,EDTA、EGTA)。在裂解缓冲液中螯合剂的浓度可以是至少约1、5、10、15、20、25、或30mM或更高。在裂解缓冲液中螯合剂的浓度可以是至多约1、5、10、15、20、25、或30mM或更高。在一些实施例中,在裂解缓冲液中的螯合剂的浓度是约10mM。裂解缓冲液可以包括还原剂(例如, β -巯基乙醇、DTT)。在裂解缓冲液中还原剂的浓度可以是至少约1、5、10、15、或20mM或更高。在裂解缓冲液中还原剂的浓度可以是至多约1、5、10、15、或20mM或更高。在一些实施例中,在裂解缓冲液中还原剂的浓度是约5mM。在一些实施例中,裂解缓冲液可以包括约0.1M Tris HCl、约pH 7.5、约0.5M LiCl、约1%十二烷基硫酸锂、约10mM EDTA,以及约5mM DTT。

[0201] 可以在约4°C、10°C、15°C、20°C、25°C、或30°C的温度进行裂解。裂解可以进行约1分钟、5分钟、10分钟、15分钟、20分钟或更多分钟。裂解的细胞可以包含至少约100000、200000、300000、400000、500000、600000、700000个或更多个靶核酸分子。裂解的细胞可以包含至多约100000、200000、300000、400000、500000、600000、700000个或更多个靶核酸分子。

[0202] 将条形码附接到靶核酸分子

[0203] 在细胞裂解和核酸分子从释放之后,核酸分子可以随机地与共定位的固体支持物的条形码相关联。关联可以包括将条形码的靶识别区与靶核酸分子的互补部分的杂交(例如,条形码的寡聚(dT)可与靶的聚(A)尾相互作用)。可以选择用于杂交的测定条件(例如缓冲液pH、离子强度、温度等)以促进形成特定的稳定的杂交体。在一些实施例中,可以将从裂解的细胞释放出的核酸分子与基底上的多个探针(例如,与基底上的探针杂交)相关联。当所述探针包括寡聚(dT)时,可以将mRNA分子与探针杂交,并且进行逆转录。可以将寡核苷酸的寡聚(dT)部分充当用于cDNA分子的第一链合成的引物。例如,图2中(在框216上)展示的条形码化的非限制性实例中,mRNA分子可以与颗粒上的条形码杂交。例如,单链的核苷酸片段可以与条形码的靶结合区杂交。

[0204] 附接可以进一步包括将条形码的靶识别区与靶核酸分子的一部分连接。例如,靶结合区可以包含可能能够与限制性位点突出端(例如,EcoRI粘性末端突出端)进行特异性杂交的核酸序列。测定程序还可以包括用限制性酶(例如EcoRI)处理靶核酸以产生限制性位点突出端。然后条形码可以连接到包括与限制性位点突出端互补的序列的任何核酸分子。连接酶(例如,T4 DNA连接酶)可用于连接两个片段。

[0205] 例如,在图2(在框220处)中说明的条形码化的非限制性实例中,随后可以将来自多个细胞(或多个样品)的经标记的靶(例如,靶-条形码分子)例如池化至管中。经标记的靶可以通过例如回收条形码和/或附接靶-条形码分子的颗粒来池化。

[0206] 可以通过使用磁性颗粒和外部施加的磁场来实现附接的靶-条形码分子的基于固体支持物的集合的检索。一旦所述靶-条形码分子已经池化,所有进一步的处理可以在单个反应容器中进行。进一步的处理可以包括例如逆转录反应、扩增反应、切割反应、解离反应和/或核酸延伸反应。进一步的处理反应可以在微孔内进行,即,不首先池化来自多个细胞的经标记的靶核酸分子。

[0207] 逆转录

[0208] 本披露提供了使用逆转录来产生靶-条形码缀合物的方法(在图2的框224中)。靶-条形码缀合物可以包括条形码以及靶核酸(即,条形码化的cDNA分子,如随机条形码化的

cDNA分子)的全部或部分的互补性序列。关联的RNA分子的逆转录可以通过添加逆转录引物连同逆转录酶一起而发生。逆转录引物可以是寡聚dT引物、随机六核苷酸引物或靶特异性寡核苷酸引物。寡聚(dT)引物的长度可以是、或可以是约12至18个核苷酸,并与哺乳动物mRNA的3'端的内源性聚(A)尾结合。随机六核苷酸引物可在多个互补位点处结合至mRNA。靶特异性寡核苷酸引物通常选择性地引发目的mRNA。

[0209] 在一些实施例中,标记的RNA分子的逆转录可通过添加逆转录引物而进行。在一些实施例中,所述逆转录引物是寡聚(dT)引物、随机六核苷酸引物或靶特异性寡核苷酸引物。通常,寡聚(dT)引物的长度为12-18个核苷酸,并结合至在哺乳动物mRNA的3'端的内源性聚(A)尾。随机六核苷酸引物可在多个互补位点处结合至mRNA。靶特异性寡核苷酸引物通常选择性地引发目的mRNA。

[0210] 逆转录可以重复地发生以产生多个经标记的cDNA分子。本文披露的方法可以包括进行至少约1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20次逆转录反应。所述方法可以包括进行至少约25、30、35、40、45、50、55、60、65、70、75、80、85、90、95或100次逆转录反应。

[0211] 扩增

[0212] 可以进行一个或多个核酸扩增反应(例如,在图2的框228中)以产生经标记的靶核酸分子的多个拷贝。扩增能以多路方式进行,其中多个靶核酸序列同时进行扩增。扩增反应可用于向核酸分子添加测序衔接子。扩增反应可以包括扩增样品标记(如果存在)的至少一部分。扩增反应可以包括扩增细胞标记和/或条形码序列(例如,分子标记)的至少一部分。扩增反应可以包括扩增样品标签、细胞标记、空间标记、条形码(例如,分子标记)、靶核酸或其组合的至少一部分。扩增反应可包括扩增多个核酸的0.5%、1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、97%、100%、或在这些值的任何两个之间的数字或范围。所述方法可以进一步包括进行一个或多个cDNA合成反应以产生包括样品标记、细胞标记、空间标记和/或条形码序列(例如,分子标记)的靶-条形码分子的一个或多个cDNA拷贝。

[0213] 在一些实施例中,可以使用聚合酶链式反应(PCR)进行扩增。如本文使用的,PCR可以指用于通过DNA的互补链的同时引物延伸使特异性DNA序列体外扩增的反应。如本文使用的,PCR可以包括所述反应的派生形式,包括但不限于RT-PCR、实时PCR、巢式PCR、定量PCR、多重PCR、数字PCR、和组装PCR。

[0214] 经标记的核酸的扩增可以包括基于非PCR的方法。非基于PCR的方法的实例包括但不限于多重置换扩增(MDA)、转录介导的扩增(TMA)、基于核酸序列的扩增(NASBA)、链置换扩增(SDA)、实时SDA、滚环扩增或环到环扩增(circle-to-circle amplification)。其他非基于PCR的扩增方法包括DNA依赖性RNA聚合酶驱动的RNA转录扩增或RNA指导的DNA合成和转录的多个循环以扩增DNA或RNA靶、连接酶链式反应(LCR)、和Q β 复制酶(Q β)方法、回文探针的使用、链置换扩增、使用限制性内切核酸酶的寡核苷酸驱动的扩增、使引物与核酸序列杂交并且将所得双链体在延伸反应和扩增之前切割的扩增方法、使用缺乏5'外切核酸酶活性的核酸聚合酶的链置换扩增、滚环扩增和分支延伸扩增(RAM)。在一些实施例中,扩增不产生环化转录物。

[0215] 在一些实施例中,本文披露的方法进一步包括对经标记的核酸(例如,经标记的

RNA、经标记的DNA、经标记的cDNA)进行聚合酶链式反应,以产生经标记的扩增子(例如,经随机标记的扩增子)。经标记的扩增子可以是双链分子。双链分子可包括双链RNA分子、双链DNA分子或者与DNA分子杂交的RNA分子。双链分子的一条或两条链可以包括样品标记、空间标记、细胞标记、和/或条形码序列(例如,分子标记)。经标记的扩增子可以是单链分子。单链分子可包括DNA、RNA或其组合。本披露的核酸可以包括合成的或改变的核酸。

[0216] 扩增可以包括使用一个或多个非天然核苷酸。非天然核苷酸可以包括光不稳定或可触发的核苷酸。非天然核苷酸的实例可以包括但不限于肽核酸(PNA)、吗啉代和锁核酸(LNA)、以及二醇核酸(GNA)与苏糖核酸(TNA)。可以将非天然核苷酸添加至扩增反应的一个或多个循环中。添加非天然核苷酸也可以用于鉴别扩增反应中特定循环或时间点的产物。

[0217] 进行一个或多个扩增反应可以包括使用一个或多个引物。一个或多个引物可以包括例如1、2、3、4、5、6、7、8、9、10、11、12、13、14、或15个或更多个核苷酸。一个或多个引物可以包括至少1、2、3、4、5、6、7、8、9、10、11、12、13、14、或15个或更多个核苷酸。一个或多个引物可以包括少于12-15个核苷酸。一个或多个引物可以退火至多个经标记的靶(例如,随机地经标记的靶)的至少一部分。一个或多个引物可以退火至多个经标记的靶的3'端或5'端。一个或多个引物可以退火至多个经标记的靶的内部区域。内部区可以是该多个经标记的靶的3'端的至少约50、100、150、200、220、230、240、250、260、270、280、290、300、310、320、330、340、350、360、370、380、390、400、410、420、430、440、450、460、470、480、490、500、510、520、530、540、550、560、570、580、590、600、650、700、750、800、850、900或1000个核苷酸。一个或多个引物可以包括一组固定的引物。一个或多个引物可以包括至少一个或多个定制引物。一个或多个引物可以包括至少一个或多个对照引物。一个或多个引物可以包括至少一个或多个基因特异性引物。

[0218] 一个或多个引物可以包括通用引物。通用引物可以退火至通用引物结合位点。一个或多个定制引物可以退火至第一样品标记、第二样品标记、空间标记、细胞标记、条形码序列(例如,分子标记)、靶、或其任何组合。一个或多个引物可以包括通用引物和定制引物。定制引物可以设计用于扩增一个或多个靶。靶可以包括一个或多个样品中总核酸的子集。靶可以包括一个或多个样品中总的经标记的靶的子集。一个或多个引物可以包括至少96个或更多个定制引物。一个或多个引物可以包括至少960个或更多个定制引物。一个或多个引物可以包括至少9600个或更多个定制引物。一个或多个定制引物可以退火至两个或更多个不同的经标记的核酸。两个或更多个不同的经标记的核酸可以对应于一个或多个基因。

[0219] 可以在本披露的方法中使用任何扩增方案。例如,在一个方案中,第一轮PCR可以使用基因特异性引物和针对通用亿明达(Illumina)测序引物1序列的引物来扩增附接到颗粒(例如,珠)上的分子。第二轮PCR可以使用侧翼于亿明达测序引物2序列的巢式基因特异性引物和针对通用亿明达测序引物1序列的引物扩增第一PCR产物。第三轮PCR添加P5和P7以及样品索引,以便使PCR产物进入亿明达测序文库。使用150bp x 2测序的测序可以揭示读段1上的细胞标记和条形码序列(例如,分子标记)、读段2上的基因、以及索引1读段上的样品索引。

[0220] 在一些实施例中,使用化学切割可以将核酸从基底中去除。例如,可以将存在于核酸中的化学基团或经修饰的碱基用于促进将其从固体支持物中去除。例如,酶可以用于从基底中去除核酸。例如,通过限制性内切核酸酶消化,可以将核酸从基底中去除。例如,使用

尿嘧啶-d-糖基化酶 (UDG) 处理含有dUTP或ddUTP的核酸可以从基底中去除核酸。例如,可以使用用于核苷酸切除(例如,碱基切除修复酶(例如,脱嘌呤/脱嘧啶(AP)核酸内切酶))的酶将核酸从基底中去除。在一些实施例中,可以使用可光解(photocleavable)基团以及光将核酸从基底中去除。在一些实施例中,可以使用可切割接头从基底中去除核酸。例如,可切割接头可以包括以下中的至少一种:生物素/亲和素、生物素/链霉亲和素蛋白、生物素/中性链亲和素、Ig蛋白A、光不稳定性接头、酸或碱不稳定性接头基团、或适配体。

[0221] 当探针是基因特异性时,可以将所述分子与探针杂交,并且进行逆转录和/或扩增。在一些实施例中,在核酸已经合成(例如,逆转录)之后,可以将其扩增。扩增能以多重方式进行,其中多个靶核酸序列同时进行扩增。扩增可以将测序衔接子添加至核酸。

[0222] 在一些实施例中,例如,用桥接扩增可以将扩增在基底上进行。cDNA可以是同聚物尾部,使用基底上的寡聚(dT)探针,以产生用于桥接扩增的相容端。在桥接扩增中,与模板核酸的3'端互补的引物可以是共价附接到固体颗粒的每对引物的第一引物。当含有模板核酸的样品与颗粒接触并进行单个热循环时,可以将模板分子退火至第一引物,并且第一引物通过添加核苷酸而向前延伸以形成双链体分子,所述双链体分子由模板分子和与模板互补的新形成的DNA链构成。在下一循环的加热步骤中,双链体分子可以变性,从颗粒释放模板分子,并通过第一引物将互补性DNA链附接至颗粒。在随后的退火和延伸步骤的退火阶段中,互补链可以与第二引物杂交,所述第二引物在从第一引物去除的位置处与互补链的片段互补。所述杂交可导致互补链在通过共价键固定到第一引物的第一和第二引物之间形成桥接,并通过杂交形成第二引物。在延伸阶段,通过在相同的反应混合物中添加核苷酸,第二引物可以按相反方向延伸,从而将桥转化为双链桥。然后开始下一个循环,并且所述双链桥可以变性以产生两个单链核酸分子,每个单链核酸分子的一端分别经第一和第二引物附接至颗粒表面,其中每个单链核酸分子的另一端是未附接的。在所述第二个循环的退火和延伸步骤中,每条链可以与先前未使用的另外的互补引物杂交在相同的颗粒上,以形成新的单链桥。将现在杂交的两个先前未使用的引物延伸从而将两个新桥转换成双链桥。

[0223] 扩增反应可以包括扩增多个核酸的至少1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、97%或100%。

[0224] 对经标记的核酸的扩增可以包括基于PCR的方法或非基于PCR的方法。对经标记的核酸的扩增可以包括对经标记的核酸的索引式扩增。对经标记的核酸的扩增可以包括对经标记的核酸的线性扩增。扩增可以通过聚合酶链式反应(PCR)来进行。PCR可指用于通过DNA的互补链的同时引物延伸使特异性DNA序列体外扩增的反应。PCR可涵盖所述反应的派生形式,包括但不限于,RT-PCR、实时PCR、巢式PCR、定量PCR、多重PCR、数字PCR、抑制PCR、半抑制PCR以及装配PCR。

[0225] 在一些实施例中,所述经标记的核酸的扩增包括非基于PCR的方法。非基于PCR的方法的实例包括但不限于多重置换扩增(MDA)、转录介导的扩增(TMA)、基于核酸序列的扩增(NASBA)、链置换扩增(SDA)、实时SDA、滚环扩增或环到环扩增(circle-to-circle amplification)。其他非基于PCR的扩增方法包括DNA依赖性RNA聚合酶驱动的RNA转录扩增或RNA指导的DNA合成和转录的多个循环以扩增DNA或RNA靶、连接酶链式反应(LCR)、 Φ 复制酶(Φ)、回文探针的使用、链置换扩增、使用限制性内切核酸酶的寡核苷酸驱动的扩增、使

引物与核酸序列杂交并且将所得双链体在延伸反应和扩增之前切割的扩增方法、使用缺乏5'外切核酸酶活性的核酸聚合酶的链置换扩增、滚环扩增和分支延伸扩增(RAM)。

[0226] 在一些实施例中,本文披露的所述方法进一步包括对扩增的扩增子(例如,靶)进行巢式聚合酶链式反应。扩增子可以是双链分子。双链分子可包括双链RNA分子、双链DNA分子或者与DNA分子杂交的RNA分子。双链分子的一条或两条链可包括样品标签或分子鉴定物标记。可替代地,所述扩增子可以是单链分子。单链分子可包括DNA、RNA或其组合。本发明的核酸可以包括合成的或改变的核酸。

[0227] 在一些实施例中,所述方法包括反复扩增经标记的核酸以产生多个扩增子。本文披露的方法可包括进行至少约1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20次扩增反应。可替代地,所述方法包括进行至少约25、30、35、40、45、50、55、60、65、70、75、80、85、90、95或100次扩增反应。

[0228] 扩增可进一步包括将一个或多个对照核酸添加至一个或多个包括多个核酸的样品中。扩增可进一步包括将一个或多个对照核酸添加至多个核酸中。对照核酸可以包括对照标记。

[0229] 扩增可以包括使用一个或多个非天然核苷酸。非天然核苷酸可以包括光不稳定和/或可触发的核苷酸。非天然核苷酸的实例包括但不限于肽核酸(PNA)、吗啉代和锁核酸(LNA)以及二醇核酸(GNA)与苏糖核酸(TNA)。可以将非天然核苷酸添加至扩增反应的一个或多个循环中。添加非天然核苷酸也可以用于鉴别扩增反应中特定循环或时间点的产物。

[0230] 进行一个或多个扩增反应可以包括使用一个或多个引物。一个或多个引物可以包括一个或多个寡核苷酸。一个或多个寡核苷酸可以包括至少约7至9个核苷酸。一个或多个寡核苷酸可以包括少于12至15个核苷酸。一个或多个引物可以退火至多个经标记的核酸的至少一部分。一个或多个引物可以退火至多个经标记的核酸的3'端和/或5'端。一个或多个引物可以退火至多个经标记的核酸的内部区。内部区可以是所述多个经标记的核酸的3'端的至少约50、100、150、200、220、230、240、250、260、270、280、290、300、310、320、330、340、350、360、370、380、390、400、410、420、430、440、450、460、470、480、490、500、510、520、530、540、550、560、570、580、590、600、650、700、750、800、850、900或1000个核苷酸。一个或多个引物可以包括一组固定的引物。一个或多个引物可以包括至少一个或多个定制引物。一个或多个引物可以包括至少一个或多个对照引物。一个或多个引物可以包括至少一个或多个管家基因引物。一个或多个引物可以包括通用引物。通用引物可以退火至通用引物结合位点。一个或多个定制引物可以退火至第一样品标签、第二样品标签、分子鉴定物标记、核酸或它们的产物。一个或多个引物可以包括通用引物和定制引物。定制引物可以被设计成扩增一个或多个靶核酸。靶核酸可以包括一个或多个样品中总核酸的子集。在一些实施例中,所述引物是附接至本披露的阵列的探针。

[0231] 在一些实施例中,条形码化(例如,随机地条形码化)在样品中的多个靶进一步包括产生条形码化的片段的索引文库。不同的条形码的条形码序列(例如,不同的随机条形码的分子标记)可以彼此不同。产生经条形码化的靶(例如,经随机条形码化的靶)的索引文库包括从样品中的多个靶产生多个索引多核苷酸。例如,对于包括第一索引靶和第二索引靶的经条形码化的靶的索引文库,第一索引多核苷酸的标记区与第二索引多核苷酸的标记区可以具有、具有约、具有至少、或具有至多1、2、3、4、5、6、7、8、9、10、20、30、40、50个核苷酸的

差异、或在这些值的任何两个之间的数字或范围的核苷酸差异。在一些实施例中,产生经条形码化的靶的索引文库包括使多个靶(例如mRNA分子)与包括聚(T)区和标记区的多个寡核苷酸接触;以及使用逆转录酶进行第一链合成以产生单链标记的cDNA分子(每个包括cDNA区和标记区),其中多个靶包含不同序列的至少两个mRNA分子,且多个寡核苷酸包含不同序列的至少两个寡核苷酸。产生经条形码化的靶的索引文库可进一步包括扩增单链标记的cDNA分子以产生双链标记的cDNA分子;以及在双链标记的cDNA分子上进行巢式PCR以产生标记的扩增子。在一些实施例中,所述方法可包括产生衔接子标记的扩增子。

[0232] 随机条形码化可以使用核酸条形码或标签以标记单个核酸(例如,DNA或RNA)分子。在一些实施例中,其涉及将DNA条形码或标签添加至cDNA分子,因为它们是从mRNA产生的。可以进行巢式PCR以最小化PCR扩增偏差。可以使用例如下一代测序(NGS)添加衔接子用于测序。例如在图2的框232处,可以使用测序结果以确定靶的一个或多个拷贝的细胞标记、条形码序列(例如,分子标记)、和核苷酸片段的序列。

[0233] 图3是显示产生经条形码化的靶(例如,经随机条形码化的靶)例如mRNA的索引文库的非限制性示例性过程的示意图。如步骤1显示,逆转录过程可以编码具有独特条形码序列(例如,分子标记)、细胞标记和通用PCR位点的每个mRNA分子。例如,通过将一组条形码(例如随机条形码)310杂交(例如随机杂交)到RNA分子302的聚(A)尾区308,可以将RNA分子302逆转录以产生经标记的cDNA分子304(包括cDNA区306)。每个条形码310可包括靶结合区,例如聚(dT)区312、条形码序列或分子标记314、和通用PCR区316。

[0234] 在一些实施例中,细胞标记可以包括3至20个核苷酸。在一些实施例中,条形码序列(例如,分子标记)可包括3至20个核苷酸。在一些实施例中,多个随机条形码的每个进一步包括通用标记和细胞标记的一个或多个,其中通用标记对于固体支持物上的多个随机条形码是相同的且细胞标记对于固体支持物上的多个随机条形码是相同的。在一些实施例中,通用标记可以包括3至20个核苷酸。在一些实施例中,细胞标记包括3至20个核苷酸。

[0235] 在一些实施例中,标记区314可包括条形码序列或分子标记318和细胞标记320。在一些实施例中,标记区314可包括通用标记、维度标记、和细胞标记的一个或多个。条形码序列或分子标记318的长度可以是、可以是约、可以是至少、或可以是至多1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100个核苷酸、或在这些值的任何之间的数字或范围的核苷酸。细胞标记320的长度可以是、可以是约、可以是至少、或可以是至多1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100个核苷酸、或在这些值的任何之间的数字或范围的核苷酸。通用标记的长度可以是、可以是约、可以是至少、或可以是至多1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100个核苷酸、或在这些值的任何之间的数字或范围的核苷酸。对于固体支持物上的多个随机条形码,通用标记可以是相同的,且对于固体支持物上的多个随机条形码,细胞标记是相同的。维度标记的长度可以是、可以是约、可以是至少、或可以是至多1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100个核苷酸、或在这些值的任何之间的数字或范围的核苷酸。

[0236] 在一些实施例中,标记区314可包括、包括约、包括至少、或包括至多1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000个不同标记、或在这些值的任何之间的数字或范围的不同标记,如条形码序列或分子标记318和细胞标记320。每个标记的长度可以是、可以是约、可以是至少、或可以是至多1、2、3、4、5、6、

7、8、9、10、20、30、40、50、60、70、80、90、100个核苷酸、或在这些值的任何之间的数字或范围的核苷酸。一组条形码或随机条形码310可以含有、含有约、含有至少、或可以是至多10、20、40、50、70、80、90、 10^2 、 10^3 、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、 10^9 、 10^{10} 、 10^{11} 、 10^{12} 、 10^{13} 、 10^{14} 、 10^{15} 、 10^{20} 个条形码或随机条形码310、或在这些值的任何之间的数字或范围的条形码或随机条形码310。并且条形码或随机条形码310的组可以例如，各自含有独特标记区314。经标记的cDNA分子304可以进行纯化以去除过量条形码或随机条形码310。纯化可以包括Ampure珠纯化。

[0237] 如步骤2所示，来自逆转录过程的产物在步骤1中可以池化至1管中，且用第1PCR引物池和第1通用PCR引物进行PCR扩增。因为独特标记区314，池化是可能的。特别地，可以将标记的cDNA分子304扩增以产生巢式PCR标记的扩增子322。扩增可以包括多重PCR扩增。扩增可以包括在单一反应体积中用96种多重引物进行的多重PCR扩增。在一些实施例中，在单一反应体积中，多重PCR扩增可以利用、利用约、利用至少、或利用至多10、20、40、50、70、80、90、 10^2 、 10^3 、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、 10^9 、 10^{10} 、 10^{11} 、 10^{12} 、 10^{13} 、 10^{14} 、 10^{15} 、 10^{20} 个多重引物、或在这些值的任何之间的数字或范围的多重引物。扩增可包括靶向特异性基因的定制引物326A-C和通用引物328的第1PCR引物池324。定制引物326可以与经标记的cDNA分子304的cDNA部分306'内的区域杂交。通用引物328可以与经标记的cDNA分子304的通用PCR区域316杂交。

[0238] 如图3的步骤3中显示，来自步骤2中的PCR扩增的产物可以用巢式PCR引物池和第2通用PCR引物进行扩增。巢式PCR可以最小化PCR扩增偏差。例如，巢式PCR标记的扩增子322可通过巢式PCR进行进一步扩增。巢式PCR可以包括在单个反应体积中用巢式PCR引物332a-c的巢式PCR引物池330和第2通用PCR引物328'的多重PCR。巢式PCR引物池328可含有、含有约、含有至少、或含有至多1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000个不同巢式PCR引物330、或在这些值的任何之间的数字或范围的不同巢式PCR引物330。巢式PCR引物332可含有衔接子334，并与经标记的扩增子322的cDNA部分306'内的区域杂交。通用引物328'可含有衔接子336，并与经标记的扩增子322的通用PCR区域316杂交。因此，步骤3产生衔接子标记的扩增子338。在一些实施例中，巢式PCR引物332和第2通用PCR引物328'可以不含有衔接子334和336。相反，衔接子334和336可以连接到巢式PCR的产物以产生衔接子标记的扩增子338。

[0239] 如步骤4中显示，可以使用文库扩增引物将来自步骤3的PCR产物进行PCR扩增用于测序。特别地，可以将衔接子334和336用于对衔接子标记的扩增子338执行一个或多个另外的测定。衔接子334和336可以与引物340和342杂交。一个或多个引物340和342可以是PCR扩增引物。一个或多个引物340和342可以是测序引物。一个或多个衔接子334和336可以用于衔接子标记的扩增子338的进一步扩增。一个或多个衔接子334和336可以用于对衔接子标记的扩增子338进行测序。引物342可含有板索引344，使得使用同一组条形码或随机条形码310产生的扩增子可以使用下一代测序(NGS)在一轮测序反应中测序。

[0240] 具有间隔子寡核苷酸的合成颗粒

[0241] 本文的披露内容包括颗粒(例如，合成颗粒)，其中所述颗粒包含：多个寡核苷酸条形码(例如，捕获寡核苷酸条形码)，其中所述多个寡核苷酸条形码中的每个包含相同的细胞标记、分子标记和靶结合区，并且其中所述多个寡核苷酸条形码中的至少10个包含不同的分子标记序列)；和第一多个间隔子寡核苷酸(或具有靶特异性靶结合区的寡核苷酸)，其

中所述第一多个间隔子寡核苷酸中的每个不包含所述寡核苷酸条形码(例如,捕获寡核苷酸条形码)中存在的靶结合区(即,靶结合区的核酸序列)。例如,所述间隔子寡核苷酸不包含与所述寡核苷酸条形码(例如,捕获寡核苷酸条形码)中存在的靶结合区具有超过95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、或25%序列同一性的核酸序列。在一些实施例中,所述间隔子寡核苷酸包含与所述寡核苷酸条形码(例如,捕获寡核苷酸条形码)中存在的靶结合区具有不超过95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、或25%序列同一性的核酸序列。在一些实施例中,所述间隔子寡核苷酸不包含长度大于例如10-30个核苷酸长的区,所述区与所述寡核苷酸条形码(例如,捕获寡核苷酸条形码)中存在的靶结合区具有超过95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、或25%序列同一性。在一些实施例中,所述间隔子寡核苷酸在期望的核酸杂交条件下不与靶结合区特异性杂交的序列杂交。在一些实施例中,所述间隔子寡核苷酸在期望的核酸杂交条件下不与靶结合区特异性杂交的序列特异性结合。

[0242] 寡核苷酸条形码的密度

[0243] 如本文所述,用于进行条形码化(例如,随机条形码化)的颗粒(例如,合成颗粒)上的寡核苷酸条形码(例如,随机条形码)的密度可以影响靶捕获靶的效率(例如,捕获mRNA靶的效率)。图4A是与多个寡核苷酸条形码相关联的非限制性示例性合成颗粒。例如,合成颗粒400a可以与寡核苷酸条形码404(展示了具有16个寡核苷酸的条形码404)相关联(例如,偶联或缀合)。寡核苷酸条形码404还在本文中被称作捕获寡核苷酸。如果合成颗粒400a上的寡核苷酸条形码404的密度较高,相比如果合成颗粒上寡核苷酸条形码的密度较低,则靶408捕获的效率可能更低(例如,参见图4B和4C)。不受任何具体理论的束缚,如果合成颗粒400a上的寡核苷酸条形码404的密度太高,则在所述寡核苷酸条形码404之间可能存在太多的空间位阻,使得如果密度较低,则所述合成颗粒400a捕获更少的靶分子408。例如,如果寡核苷酸条形码404的靶结合区412(展示的聚(dT)尾)的密度过高,则所述寡核苷酸条形码404中可能存在空间位阻(或靶结合区中的空间位阻),使得一些靶分子408(所展示的靶分子408的聚(A)尾416)不与所述寡核苷酸条形码404杂交。因此,在一些实施例中,例如通过使用间隔子寡核苷酸与颗粒相关联,降低颗粒中捕获寡核苷酸条形码的密度可能是有利的。如本文所述,间隔子寡核苷酸的形式、类型、长度、密度和位置可以发生变化。与颗粒相关联的间隔子寡核苷酸的数目可以大于、小于、等于、或基本上等于与相同颗粒相关联的捕获寡核苷酸的数目。尽管所展示的合成颗粒400a与具有16个寡核苷酸的条形码404相关联,但是所述寡核苷酸条形码404的数量仅用于说明,而不旨在进行限制。

[0244] 图4B是与多个寡核苷酸条形码和多个间隔子寡核苷酸相关联的非限制性示例性合成颗粒。如所展示的,合成颗粒400b可以包括间隔子寡核苷酸420b(所展示的六个间隔子寡核苷酸420b)和寡核苷酸条形码404。寡核苷酸条形码404可以包括能够捕获(例如,杂交)靶分子408的靶结合区412。间隔子寡核苷酸420b可以包括间隔子区424b,其与靶结合区412相对应,不能捕获(例如,杂交)靶分子408。在合成颗粒400b上包括间隔子寡核苷酸420b可以改善靶408的捕获(例如,mRNA捕获)和随后对捕获的靶分子408的定量。例如,合成颗粒400b可以与寡核苷酸条形码404(例如,各自具有聚(dT)区的条形码,所述条形码可以经由聚(dA)尾416结合到mRNA分子或种类的拷贝或出现)和间隔子寡核苷酸420b两者偶联,这可

以改善mRNA捕获和基因表达定量。尽管所展示的合成颗粒400b与具有10个寡核苷酸条形码的404和六个间隔子寡核苷酸420b相关联,但是寡核苷酸条形码404和间隔子寡核苷酸的数量仅用于说明,而不旨在进行限制。

[0245] 相比寡核苷酸条形码,与合成颗粒相关联的间隔子寡核苷酸可以更短或更长。例如,相比寡核苷酸条形码,与合成颗粒相关联的所有间隔子寡核苷酸可以更短(例如,参见图4C1)。作为另一个实例,相比寡核苷酸条形码,与合成颗粒相关联的所有间隔子寡核苷酸可以更长(例如,参见图4C2)。例如,与合成颗粒相关联的一些间隔子寡核苷酸可以比寡核苷酸条形码更长,并且与合成颗粒相关联的一些间隔子寡核苷酸可以比寡核苷酸条形码更短。

[0246] 图4C1是与多个寡核苷酸条形码和多个间隔子寡核苷酸相关联的非限制性示例性合成颗粒,每个间隔子寡核苷酸相比寡核苷酸条形码更短。相比寡核苷酸条形码404,在合成颗粒400c1上的间隔子寡核苷酸420c1可以更短。可以首先产生与具有相同序列的寡核苷酸条形码404(或具有相同序列的靶结合区412)相关联的合成颗粒400c1。可以将酶(例如核酸外切酶(5'或3')或核酸内切酶)用于去除寡核苷酸条形码404中的一些的靶结合区域412。在合成颗粒上所得的更短的寡核苷酸条形码在本文中被称为间隔子寡核苷酸420c。

[0247] 在一些实施例中,可以通过将图4C1中展示的寡核苷酸条形码404和间隔子寡核苷酸420c1两者与合成颗粒400c1相关联(例如,偶联),降低寡核苷酸条形码404的密度,并随后使用核酸外切酶或核酸内切酶降低密度。

[0248] 图4C2是与多个寡核苷酸条形码和多个间隔子寡核苷酸相关联的非限制性示例性合成颗粒,每个间隔子寡核苷酸相比寡核苷酸条形码更长。相比寡核苷酸条形码404,在合成颗粒400c2上的间隔子寡核苷酸420c2可以更长。例如,可以首先产生与具有相同序列的寡核苷酸条形码404(或具有相同序列的靶结合区412)相关联的合成颗粒400c2。可以将诸如聚合酶或连接酶的酶用于将一个或多个序列通过延伸或连接添加至寡核苷酸条形码404中的一些,以产生比寡核苷酸条形码404更长的间隔子寡核苷酸420c2。可以添加相同序列或不同序列以产生合成颗粒400c2的间隔区寡核苷酸420c2。在一些实施例中,为产生间隔子寡核苷酸420c2而添加的不同序列的数量可以是、或约1、2、3、4、5、6、7、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、或在这些值的任何两个之间的数字或范围。在一些实施例中,为产生间隔子寡核苷酸420c2而添加的不同序列的数量可以是至少、或至多1、2、3、4、5、6、7、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、或1000。合成颗粒400c2的间隔子寡核苷酸420c2可以具有相同的序列或不同的序列。在一些实施例中,所述间隔子寡核苷酸420c2可以具有、或具有约1、2、3、4、5、6、7、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000个或在这些值的任何两个之间的数字或范围的序列。在一些实施例中,间隔子寡核苷酸420c2可以具有至少、或具有至多1、2、3、4、5、6、7、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、或1000个不同的序列。合成颗粒400c2的间隔子寡核苷酸420c2可以具有相同的长度或不同长度。在一些实施例中,所述间隔子寡核苷酸420c2可以具有、或具有约1、2、3、4、5、6、7、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000个或在这些值的任何两个之间的数字或范围的不同长度。在一些实施例中,间隔子寡核苷酸420c2可以具有至少、或具有至多1、2、3、4、5、6、7、9、10、20、

30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、或1000个不同长度。在合成颗粒上所得的更长的寡核苷酸条形码在本文中被称为间隔子寡核苷酸420c1。

[0249] 在一些实施例中,可以通过将图4C2中展示的寡核苷酸条形码404和间隔子寡核苷酸420c2两者与合成颗粒400c2相关联(例如,偶联),降低寡核苷酸条形码404的密度。

[0250] 寡核苷酸条形码的密度可以通过调节用于将寡核苷酸条形码与合成颗粒相关联的相关联条件来改变(例如,减少交联剂的量,在偶联过程中减少寡核苷酸条形码的量)。通过改变相关联条件来准确地控制相关联(例如,偶联)可能需要大量测试。在一些实施例中,可以首先通过调节相关联条件来降低合成颗粒上的寡核苷酸条形码的密度,并使用核酸外切酶或核酸内切酶进一步降低。在一些实施例中,可通过在调节以产生具有较低密度的寡核苷酸条形码的合成颗粒的条件下使寡核苷酸条形码和间隔子寡核苷酸与合成颗粒相关联来降低寡核苷酸条形码的密度,并使用核酸外切酶或核酸内切酶进一步降低。

[0251] 在一些实施例中,间隔子寡核苷酸(例如,在合成颗粒上相关联的间隔子寡核苷酸)可以具有相同长度或不同长度。在一些实施例中,相比一个或多个寡核苷酸条形码,一个或多个间隔子寡核苷酸可以更短。例如,间隔子寡核苷酸和寡核苷酸条形码可以是相同的,除了所述寡核苷酸条形码包含靶结合区,并且所述间隔子寡核苷酸不包含靶结合区。在一些实施例中,一个或多个间隔子寡核苷酸可以具有与一个或多个寡核苷酸条形码相同的长度。例如,在一些实施例中(除了寡核苷酸条形码包含靶结合区,并且间隔子寡核苷酸包含与靶结合区具有相同长度但具有不同核苷酸序列的区),间隔子寡核苷酸和寡核苷酸条形码可以是相同的。

[0252] 具有靶特异性靶结合区的寡核苷酸条形码

[0253] 本文的披露内容包括合成颗粒的实施例。在一些实施例中,所述合成颗粒包含:第一多个寡核苷酸条形码;和第二多个寡核苷酸条形码,其中所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码中的每个包含相同的细胞标记序列并且包含分子标记序列,其中所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码中的至少10个包含不同的分子标记序列,其中所述第一多个寡核苷酸条形码中的每个包含靶结合区,并且其中所述第二多个寡核苷酸条形码中的每个包含靶特异性靶结合区。在一些实施例中,所述第一多个寡核苷酸条形码的寡核苷酸条形码的靶结合区包含靶特异性靶结合区。在一些实施例中,所述第一多个寡核苷酸条形码的寡核苷酸条形码的靶结合区包含聚(dT)区。所述第二多个寡核苷酸条形码的寡核苷酸条形码可以包含聚(dT)区。

[0254] 本文的披露内容包括如下方法的实施例,所述方法用于确定样品中靶的数量。在一些实施例中,所述方法包括:使用与合成颗粒相关联的第一多个寡核苷酸条形码,对第一多个靶的拷贝进行条形码化,以产生第一多个经条形码化的靶的拷贝;使用与所述合成颗粒相关联的第二多个寡核苷酸条形码,对样品中第二多个靶的拷贝进行条形码化,以产生第二多个经条形码化的靶的拷贝,其中所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码中的每个包含相同的细胞标记序列并且包含分子标记序列,其中所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码中的至少10个包含不同的分子标记序列,其中所述第一多个寡核苷酸条形码中的每个包含靶结合区,并且其中所述第二多个寡核苷酸条形码中的每个包含靶特异性靶结合区;并且使用所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码的分子标记序列对所述第一多个靶和/或所述第二多个靶中的每

个的数量进行评估。在一些实施例中,所述第一多个寡核苷酸条形码的寡核苷酸条形码的靶结合区包含靶特异性靶结合区。所述第一多个寡核苷酸条形码的寡核苷酸条形码的靶结合区可以包含聚(dT)区。所述第二多个寡核苷酸条形码的寡核苷酸条形码可以包含聚(dT)区。在一些实施例中,第一多个靶包含细胞的mRNA分子,并且其中第二多个靶包含细胞的DNA分子。所述第一多个靶可以包含细胞的mRNA分子,并且所述第二多个靶可以包含感染细胞的病毒、细菌和/或真菌的RNA分子。

[0255] 图4D1-4D2是与具有靶特异性靶结合区的多个寡核苷酸条形码相关联的非限制性示例性合成颗粒。合成颗粒上的寡核苷酸条形码的靶结合区可以包含、或可以是靶特异性靶结合区和/或非靶特异性靶结合区(例如,聚(dT)区和随机多聚体)。在合成颗粒(例如,与间隔子寡核苷酸相关联的合成颗粒)上的寡核苷酸条形码的靶结合区可以包括、或是聚(dT)区404、随机多聚体、和/或靶特异性序列412d(例如,能够与一个或多个目的靶杂交的序列)。例如,与合成颗粒400d1相关联的寡核苷酸条形码404d1可以包括能够特异性结合许多目的靶(例如,四个目的靶,如CD4、IL2、CD8和CR7)的靶结合区412d1和具有聚(dT)区412的寡核苷酸条形码404(图4D1)。

[0256] 在一些实施例中,所述目的靶可以包括与细胞组分结合试剂(例如,蛋白质结合试剂,如抗体或蛋白质结合适配子)缀合的寡核苷酸,其中所述寡核苷酸包含针对与其缀合的细胞组分结合试剂的独特标识。已经在美国专利申请公开号US 2018/0088112和美国专利申请公开号US 2018/0346970(这些文献中每个的内容通过引用以其全文并入本文中)中描述了与寡核苷酸(如条形码化的抗体)相关联的细胞组分结合试剂和它们的用途(如对细胞和样品的样品索引、鉴定或追踪以及蛋白质表达谱分析)。

[0257] 样品索引、鉴定或追踪. 在一些实施例中,进行样品索引、鉴定或追踪可以包括:使来自多个样品中的每个的一个或多个细胞与多个样品索引组合物的样品索引组合物接触,其中所述一个或多个细胞中的每个包含一个或多个细胞组分靶(例如,蛋白质靶),其中所述多个样品索引组合物中的至少一个样品索引组合物包含两个或更多个细胞组分结合试剂(例如抗原结合试剂或抗体),其中所述两个或更多个细胞组分结合试剂中的每个与样品索引寡核苷酸相关联,其中所述两个或更多个细胞组分结合试剂中的至少一个能够与所述一个或多个细胞组分靶中的至少一个特异性结合,其中所述样品索引寡核苷酸包含样品索引序列,并且其中所述多个样品索引组合物中的至少两个样品索引组合物的样品索引序列包含不同的序列;使用多个条形码(例如,与合成颗粒相关联的寡核苷酸条形码)对样品索引寡核苷酸进行条形码化,以创建多个经条形码化的样品索引寡核苷酸;获取所述多个经条形码化的样品索引寡核苷酸的测序数据;并且基于所述多个经条形码化的样品索引寡核苷酸中的至少一个经条形码化的样品索引寡核苷酸的样品索引序列,鉴定所述一个或多个细胞中的至少一个细胞的样品来源。在一些实施例中,样品索引序列的长度是25-60个核苷酸(例如长度是45个核苷酸)、长度是约128个核苷酸,或长度是至少128个核苷酸。多个样品索引组合物中的至少10个样品索引组合物的样品索引序列可以包含不同的序列。多个样品索引组合物中的至少10个样品索引组合物的样品索引序列可以包含不同的序列。多个样品索引组合物中的至少10个样品索引组合物的样品索引序列包含不同的序列。在一些实施例中,样品索引寡核苷酸与多个样品的细胞的基因组序列不同源。样品索引寡核苷酸可以包含分子标记序列;条形码的靶结合区可以结合(例如,杂交)到(例如,聚(A)区或给聚(A)区)

的结合序列;或其组合。样品索引寡核苷酸可以包含与多个条形码中至少一个条形码的捕获序列(例如,靶结合区)互补的序列。条形码的靶结合区可以包括捕获序列。例如靶结合区可以包含聚(dT)区。例如与条形码的捕获序列互补的样品索引寡核苷酸的序列可以包含聚(dA)尾。在一些实施例中,样品索引寡核苷酸可以包含分子标记。

[0258] 表达谱. 在一些实施例中,进行蛋白质表达谱分析(或细胞组分靶的谱分析)可以包括:提供包含多个蛋白质靶(或其他细胞组分靶,例如碳水化合物、脂质等)的样品;提供多个组合物,所述组合物各自包含与寡核苷酸缀合的蛋白质结合试剂(或细胞组分结合试剂),其中所述寡核苷酸包含针对与其缀合的蛋白质结合试剂的独特标识,并且所述蛋白质结合试剂(或细胞组分结合试剂)能够特异性结合所述多个蛋白质靶(或细胞组分靶)中的至少一个;使所述多个组合物与所述样品接触,用于与所述多个蛋白质靶(或细胞组分靶)特异性结合;去除未结合的组合物;提供多个寡核苷酸探针(例如,与合成颗粒相关联的寡核苷酸条形码),其中所述多个寡核苷酸探针中的每个包含靶结合区和条形码序列(例如,分子标记序列),其中所述条形码序列来自独特条形码序列的多元组;使所述多个寡核苷酸探针与所述多个种组合物的寡核苷酸接触;延伸与所述寡核苷酸杂交的寡核苷酸探针以产生多个经标记的核酸,其中所述经标记的核酸中的每个包含独特标识和条形码序列;并且确定每个独特标识的独特条形码序列的数量,由此确定所述样品中每个蛋白质靶的量。在一些实施例中,独特标识包括长度为25-45个核苷酸的核苷酸序列。在一些实施例中,独特标识选自独特标识的多元组。在一些实施例中,独特标识的多元组包括至少100种不同的独特标识。在一些实施例中,独特标识的多元组包括至少1,000种不同的独特标识。在一些实施例中,独特标识的多元组包括至少10,000种不同的独特标识。在一些实施例中,多个组合物包括多个抗体、多个适配子、或其组合。在一些实施例中,寡核苷酸通过接头与蛋白质结合试剂缀合。在一些实施例中,所述接头包括化学基团。在一些实施例中,所述寡核苷酸包括接头。在一些实施例中,所述化学基团可逆地附接至蛋白质结合试剂。在一些实施例中,所述化学基团选自下组,该组由以下组成:UV可光解基团、链霉亲和素蛋白、生物素、胺、及其任何组合。在一些实施例中,样品包括单个细胞。在一些实施例中,多个蛋白质靶在单个细胞的表面表达。在一些实施例中,去除未结合的组合物包括用洗涤缓冲液洗涤单个细胞。在一些实施例中,这些方法包括裂解单个细胞。在一些实施例中,这些方法包括从蛋白质结合试剂中分离寡核苷酸。在一些实施例中,通过UV光解、化学处理(二硫苏糖醇)、加热、酶处理、或其任何组合将寡核苷酸从蛋白质结合试剂中分离。在一些实施例中,寡核苷酸探针中的每一个包括细胞标记、通用引物的结合位点、或其任何组合。在一些实施例中,靶结合区包括聚(dT)。在一些实施例中,多个寡核苷酸探针被固定在固体支持物上。在一些实施例中,固体支持物是珠。在一些实施例中,这些方法进一步包括扩增多个经标记的核酸以产生多个扩增子。在一些实施例中,扩增包括对条形码序列的至少一部分和独特标识的至少一部分进行PCR扩增。在一些实施例中,独特条形码序列的多元组包括至少100个独特条形码序列。在一些实施例中,独特条形码序列的多元组包括至少1,000个独特条形码序列。在一些实施例中,独特条形码序列的多元组包括至少10,000个独特条形码序列。多个组合物可以进一步包括不与寡核苷酸缀合的第二蛋白质结合试剂。蛋白质结合试剂和第二蛋白质结合试剂可以是相同的。在一些实施例中,这些方法进一步包括对多个扩增子进行测序。在一些实施例中,测序包括对条形码序列的至少一部分和独特标识的至少一部分进行测序。可

以将所述方法用于对样品中的多个蛋白质靶和多个核酸靶分子进行同时定量分析。

[0259] 在一些实施例中,与细胞组分结合试剂缀合的寡核苷酸可以具有聚(dA)区416,所述聚(dA)区416可以与寡核苷酸条形码404的聚(dT)区412杂交。在一些实施例中,寡核苷酸条形码404d1可以具有靶特异性靶结合区412d,所述靶结合区412d可以特异性地杂交和捕获与细胞组分结合试剂相关联的寡核苷酸。与靶标特异性靶结合区412结合或杂交的细胞组分结合试剂相关联的寡核苷酸区可以在寡核苷酸的3'端(与mRNA的聚(dA)区相对应)、在寡核苷酸的5'端、或在寡核苷酸的内部。在一些实施例中,寡核苷酸在被具有靶特异性靶结合区412d的寡核苷酸条形码404d1捕获之前,从细胞组分结合试剂解离。

[0260] 如图4D1所展示的,寡核苷酸条形码404的聚(dT)区可以捕获(例如杂交)具有聚(A)尾416的靶(如mRNA分子)。寡核苷酸条形码404的聚(dT)区412可以捕获具有聚(A)区或尾416的靶。寡核苷酸条形码404d1的靶特异性靶结合区412d可以捕获(例如,杂交)目的靶。例如,靶特异性靶结合区412d可以包括或可以是部分CD4序列,所述序列可以与CD4 mRNA分子的非聚(A)区416d(或具有非A碱基的CD4 mRNA的区416d)杂交并捕获CD4 mRNA分子。作为另一个实例,靶标特异性靶结合区412d可以包括或者可以是用于捕获不具有聚(A)尾的靶(例如病毒mRNA)的序列。在一些实施例中,合成颗粒400d1可用于同时确定被病毒感染的细胞的基因型和/或表型以及病毒的基因型和/或表型。例如,合成颗粒400d1可以包括具有用于捕获被感染细胞的mRNA的聚(dT)区412的寡核苷酸条形码404和具有用于捕获病毒RNA(例如,病毒mRNA)的靶特异性靶结合区域412d的寡核苷酸条形码404d1。在一些实施例中,可以通过将寡核苷酸条形码404与聚(dT)区412以及将寡核苷酸条形码404d1与靶特异性靶结合区412d相关联(例如,偶联)至合成颗粒400d1,来产生珠400d1。

[0261] 参考图4D2,与合成颗粒400d2相关联的寡核苷酸条形码404d2可以包括能够特异性结合许多目的靶的靶结合区412d1和聚(dT)区404(图4D2)。如图4D2所展示,合成颗粒400d2还可以包括具有非靶特异性靶结合区(例如,聚(dT)区400)的寡核苷酸条形码404。如图4D2所展示,寡核苷酸条形码404的聚(dT)区412可以捕获(例如杂交)具有聚(A)尾416的靶(如mRNA分子)。寡核苷酸条形码404的聚(dT)区412可以捕获具有聚(A)区或尾416的靶。寡核苷酸条形码404d2的靶特异性靶结合区412d可以捕获(例如,杂交)目的靶。如图4D2所展示,在一些实施例中,具有靶特异性靶结合区412d的寡核苷酸条形码404d2可以包括聚(dT)区412。在一些实施例中,合成颗粒400d2可用于同时确定被病毒感染的细胞的基因型和/或表型以及病毒的基因型和/或表型。例如,合成颗粒400d2可以包括具有用于捕获被感染细胞的mRNA的聚(dT)区412的寡核苷酸条形码404和具有用于捕获病毒RNA(例如,病毒mRNA)的靶特异性靶结合区域412d的寡核苷酸条形码404d2。

[0262] 在一些实施例中,可以通过将寡核苷酸条形码404与聚(dT)区412以及将寡核苷酸条形码404d2与靶特异性靶结合区412d和聚(dT)区412相关联(例如,偶联)至合成颗粒400d,来产生珠400d2。在一些实施例中,可以首先产生与具有相同序列(或聚(dT)区412)的寡核苷酸条形码404相关联的合成颗粒400d2。可以将诸如聚合酶或连接酶的酶用于通过延伸或连接将靶特异性靶结合区412d添加至寡核苷酸条形码404中的一些,以产生具有靶特异性靶结合区412d和聚(dT)区412的寡核苷酸条形码404d2。例如,具有聚(dA)序列的单链DNA模板(例如,在模板的3'端)和目的靶的部分序列(例如,在模板的5'端)可以与寡核苷酸条形码404的聚(dT)区412杂交,这可以通过聚合酶延伸以产生具有靶特异性靶结合区412d

和聚(dT)区412的寡核苷酸条形码404d2。作为另一个实例,可以使用部分双链体DNA模板,产生具有靶特异性靶结合区412和聚(dT)区412的寡核苷酸条形码404d2。所述部分双链体DNA模板的一条链可以包括聚(dA)序列(例如,在链的5'端)和目的靶的部分序列(例如,在链的3'端)。部分双链体DNA模板的其他链可以包括靶特异性靶结合区。可以通过使一条链上的目的靶的部分序列与另一条链上的靶特异性靶结合区杂交来形成部分双链体DNA模板。经由寡核苷酸条形码404的聚(dT)区412和部分双链体DNA模板的聚(dA)序列的杂交,可以将部分双链体DNA模板与寡核苷酸条形码404杂交。杂交后,可以使用连接酶将部分双链DNA模板与具有聚(dT)区412的寡核苷酸条形码404连接。可以通过使连接至寡核苷酸条形码404的部分双链体DNA模板变性来去除具有聚(dA)序列的模板链,以产生具有靶特异性靶结合区412d和聚(dT)区412的寡核苷酸条形码404。合成颗粒400d1、400d2可以包括间隔子寡核苷酸(未示出)。

[0263] 图5A-5C显示以下各项的非限制性示例性对比:具有能够与mRNA分子的聚(A)尾杂交的聚(dT)区412的寡核苷酸条形码404(图5A);具有能够与一个或多个目的基因结合的靶特异性靶结合区412的寡核苷酸条形码404d1;和具有聚(dT)区412和靶特异性靶结合区412d的寡核苷酸条形码404d2。寡核苷酸条形码404、404d1、404d2可以与单一合成颗粒相关联。在一些实施例中,代替聚(dT)寡核苷酸(用于经由聚(A)尾捕获mRNA)与颗粒(如合成颗粒或珠)的偶联(例如,共价偶联),具有各种靶特异性序列412(例如,基因特异性)的寡核苷酸条形码可以与颗粒偶联。这些寡核苷酸可以与具有聚(dT)区404的寡核苷酸具有相似的结构。寡核苷酸条形码404d1、404d2可以具有不同的序列,用于捕获RNA(包括不具有聚(A)尾的RNA分子)。例如,取决于偶联的寡核苷酸404d1、404d2的靶,颗粒可以允许捕获RNA的特定子集。由于单个细胞具有约300,000个mRNA分子,即使这些分子的仅一个子集最终将被靶向基因组扩增,需要偶联到每个颗粒上以捕获所有mRNA分子的具有聚(dT)区的寡核苷酸条形码的数量也很大(例如,大于300,000)。使用靶特异性寡核苷酸条形码404d1、404d2,可以产生更具选择性的一组cDNA分子,因为仅捕获了用于随后扩增和测序的mRNA分子。此外,用于随后扩增和测序的分子(例如,针对每个基因或针对所有目的基因)的数量可能远低于300,000个拷贝,因此需要将靶向所述基因或所有目的基因的更少的寡核苷酸条形码404d1、404d2与颗粒偶联。所述颗粒可以与寡核苷酸条形码404d1、404d2偶联,所述寡核苷酸条形码能够特异性捕获预期由一种或多种细胞类型(例如,CD4+辅助T细胞)表达的基因的mRNA分子。因此,仅这些基因的mRNA分子可以被捕获并进一步分析(例如表达定量)。

[0264] 第一多个寡核苷酸(例如,间隔子寡核苷酸和靶特异性寡核苷酸条形码)

[0265] 本文披露了合成颗粒的实施例。在一些实施例中,合成颗粒(例如,参考4B和4C1-4C2的合成颗粒400b、400c1、400c2)包括:多个寡核苷酸条形码(例如,在图4B和4C1-4C2中的寡核苷酸条形码404,和在图4D1-4D2中的寡核苷酸条形码404d1、404d2)。多个寡核苷酸条形码404、404d1、404d中的每个可以包含相同的细胞标记、分子标记和靶结合区(例如,在图4B、4C1-4C2和4D1-4D2中的靶结合区412、412d1、412d2)。多个寡核苷酸条形码中的至少、或至多10、100、1000、10000、100000、或更多个可以包含不同的分子标记序列。合成颗粒可以包括第一多个间隔子寡核苷酸(例如,参考图4B和4C1-4C2所描述的间隔子寡核苷酸420b、420c1、420c2)。

[0266] 在一些实施例中,第一多个间隔子寡核苷酸(或具有靶特异性靶结合区的寡核苷酸

酸条形码,在本文中被称为靶特异性寡核苷酸条形码)的数量可以是或可以约是0、1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、 10^9 、或在这些值的任何两个之间的数字或范围。在一些实施例中,第一多个间隔子寡核苷酸(或具有靶特异性靶结合区的寡核苷酸条形码)的数量可以是至少或至多0、1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 或 10^9 。

[0267] 在不同的实现方式中,多个寡核苷酸条形码和第一多个间隔子寡核苷酸(或具有靶特异性靶结合区的寡核苷酸条形码)的比率可以是不同的,范围从 $1:10^9$ 至 $10^9:1$ 。如本文所使用的,两个所述多个寡核苷酸的比率是指第一多个寡核苷酸中寡核苷酸的数量和第二多个寡核苷酸中寡核苷酸的数量之间的比率。在一些实施例中,多个寡核苷酸条形码和第一多个间隔子寡核苷酸(或具有靶特异性靶结合区的寡核苷酸条形码)的比率可以是或可以约是1:1、1:2、1:3、1:4、1:5、1:6、1:7、1:8、1:9、1:10、1:20、1:30、1:40、1:50、1:60、1:70、1:80、1:90、1:100、1:200、1:300、1:400、1:500、1:600、1:700、1:800、1:900、1:1000、1:2000、1:3000、1:4000、1:5000、1:6000、1:7000、1:8000、1:9000、 $1:10^4$ 、 $1:10^5$ 、 $1:10^6$ 、 $1:10^7$ 、 $1:10^8$ 、 $1:10^9$ 、或在这些值的任何两个之间的数字或范围。在一些实施例中,多个寡核苷酸条形码和第一多个间隔子寡核苷酸的比率可以是至少或至多1:1、1:2、1:3、1:4、1:5、1:6、1:7、1:8、1:9、1:10、1:20、1:30、1:40、1:50、1:60、1:70、1:80、1:90、1:100、1:200、1:300、1:400、1:500、1:600、1:700、1:800、1:900、1:1000、1:2000、1:3000、1:4000、1:5000、1:6000、1:7000、1:8000、1:9000、 $1:10^4$ 、 $1:10^5$ 、 $1:10^6$ 、 $1:10^7$ 、 $1:10^8$ 或 $1:10^9$ 。在一些实施例中,多个寡核苷酸条形码和第一多个间隔子寡核苷酸(或具有靶特异性靶结合区的寡核苷酸条形码)的比率可以是或可以约是1:1、2:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1、10:1、20:1、30:1、40:1、50:1、60:1、70:1、80:1、90:1、100:1、200:1、300:1、400:1、500:1、600:1、700:1、800:1、900:1、1000:1、2000:1、3000:1、4000:1、5000:1、6000:1、7000:1、8000:1、9000:1、 $10^4:1$ 、 $10^5:1$ 、 $10^6:1$ 、 $10^7:1$ 、 $10^8:1$ 、 $10^9:1$ 、或在这些值的任何两个之间的数字或范围。在一些实施例中,多个寡核苷酸条形码和第一多个间隔子寡核苷酸(或具有靶特异性靶结合区的寡核苷酸条形码)的比率可以是至少或至多1:1、2:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1、10:1、20:1、30:1、40:1、50:1、60:1、70:1、80:1、90:1、100:1、200:1、300:1、400:1、500:1、600:1、700:1、800:1、900:1、1000:1、2000:1、3000:1、4000:1、5000:1、6000:1、7000:1、8000:1、9000:1、 $10^4:1$ 、 $10^5:1$ 、 $10^6:1$ 、 $10^7:1$ 、 $10^8:1$ 、或 $10^9:1$ 。

[0268] 相比寡核苷酸条形码,间隔子寡核苷酸(或具有靶特异性靶结合区的寡核苷酸条形码)可以更短、具有相同的长度或更长。在不同的实现方式中,相比每个寡核苷酸条形码(或至少一个寡核苷酸条形码)更短、具有相同的长度或更长的间隔子寡核苷酸(或具有靶特异性靶结合区的寡核苷酸条形码)的数量可以是不同的。在一些实施例中,相比每个寡核苷酸条形码(或至少一个寡核苷酸条形码)更短、具有相同的长度或更长的间隔子寡核苷酸(或具有靶特异性靶结合区的寡核苷酸条形码)的数量可以是或可以约1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、 10^9 、或在这些值的任何两个之间的数字或范围。在一些实施例中,相比每个寡核苷酸条形码(或至少一个寡核苷酸条形

码)更短、具有相同的长度或更长的间隔子寡核苷酸(或具有靶特异性靶结合区的寡核苷酸条形码)的数量可以是至少或至多1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、或 10^9 。例如,相比多个寡核苷酸条形码中的至少一个,所述第一多个间隔子寡核苷酸中的至少一个可以更短。

[0269] 在一些实施例中,相比每个(或至少一个)间隔子寡核苷酸(或具有靶特异性靶结合区的寡核苷酸),更短、具有相同长度或更长的寡核苷酸条形码的数量可以是或可以约1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、 10^9 、或在这些值的任何两个之间的数字或范围。在一些实施例中,相比每个(或至少一个)间隔子寡核苷酸(或具有靶特异性靶结合区的寡核苷酸条形码),更短、具有相同长度或更长的寡核苷酸条形码的数量可以是至少或至多1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、或 10^9 。

[0270] 所述第一多个间隔子寡核苷酸中的每个可以不包含靶结合区。例如,图4B中展示的间隔子寡核苷酸420b可以包括与靶结合区相对应、不能用于与目的靶杂交的间隔子区424b。作为另一个实例,相比缺少靶结合区412的寡核苷酸条形码404,在图4C1中展示的间隔子寡核苷酸420c1更短。作为另一个实例,相比寡核苷酸条形码404,图4C2中展示的间隔子寡核苷酸420c1更长。所述第一多个间隔子寡核苷酸(或具有靶特异性靶结合区的寡核苷酸条形码)中的每个可以包含细胞标记或分子标记。所述第一多个间隔子寡核苷酸中的每个可以包含与靶结合区相对应的区(本文中被称为间隔子区,例如在图4B中的间隔子区424b)。

[0271] 与靶结合区相对应的区424b和靶结合区(或两个靶结合区)可以具有不同的序列。间隔子区424b和靶结合区(或两个靶结合区)可以是不同的。在不同的实现方式中,在间隔子区424b和靶结合区(或在两个靶结合区(如聚(dT)区412和靶特异性靶结合区412d)之间)之间的序列同一性可以是不同的(在1%至99%的范围内)。在一些实施例中,在间隔子区424b和靶结合区412、412d(或两个靶结合区)之间的序列同一性可以是或可以约是1%、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%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、或在这些值的任何两个之间的数字或范围。在一些实施例中,在间隔子区424b和靶结合区412、412d(或在两个靶结合区之间)之间的序列同一性可以是至少、或至多1%、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%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、

60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、或99%。例如,与靶结合区和靶结合区412、412d(或两个靶结合区)相对应的区424b可以具有至多50%序列同一性。在一些实施例中,间隔子区424b和靶结合区412、412d(或两个靶结合区)可以彼此差异或约差异1、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、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99、100、或在这些值的任何两个之间的数字或范围。在一些实施例中,间隔子区424b和靶结合区412、412d(或两个靶结合区)可以彼此差异至少、或至多1、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、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99、或100。

[0272] 在不同实现方式中,具有不同长度(例如在图4C1-4C2中,间隔子寡核苷酸420c1、420c2)的第一多个间隔子寡核苷酸(或具有靶特异性靶结合区的寡核苷酸条形码)的数量可以是不同的。在一些实施例中,具有不同长度的第一多个间隔子寡核苷酸(或具有靶特异性靶结合区的寡核苷酸条形码)的数量可以是或可以约是0、1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、 10^9 、或在这些值的任何两个之间的数字或范围。在一些实施例中,具有不同长度的第一多个间隔子寡核苷酸(或具有靶特异性靶结合区的寡核苷酸条形码)的数量可以是至少或至多0、1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、或 10^9 。例如,第一多个间隔子寡核苷酸(或具有靶特异性靶结合区的寡核苷酸条形码)中的至少两个可以具有不同的长度。在一些实施例中,具有不同长度的第一多个间隔子寡核苷酸(或具有靶特异性靶结合区的寡核苷酸条形码)的百分比可以是或可以约是1%、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%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、100%、或在这些值的任何两个之间的数字或范围。在一些实施例中,具有不同长度的第一多个间隔子寡核苷酸(或具有靶特异性靶结合区的寡核苷酸条形码)的百分比可以是至少或至多1%、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%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、

49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%。

[0273] 在不同的实现方式中,具有相同序列的第一多个间隔子寡核苷酸(或具有靶特异性靶结合区的寡核苷酸条形码)的数量可以是不同的。在一些实施例中,具有相同序列的第一多个间隔子寡核苷酸(或具有靶特异性靶结合区的寡核苷酸条形码)的数量可以是或可以约是0、1、2 3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、 10^9 、或在这些值的任何两个之间的数字或范围。在一些实施例中,具有相同序列的第一多个间隔子寡核苷酸(或具有靶特异性靶结合区的寡核苷酸条形码)的数量可以是至少、至多0、1、2 3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、或 10^9 。例如,所述第一多个间隔子寡核苷酸(或具有靶特异性靶结合区的寡核苷酸条形码)中的至少两个的序列可以是相同的。在不同的实现方式中,具有相同子序列的第一多个间隔子寡核苷酸(或具有靶特异性靶结合区的寡核苷酸条形码)的数量可以是不同的。在一些实施例中,具有相同子序列的第一多个间隔子寡核苷酸(或具有靶特异性靶结合区的寡核苷酸条形码)的数量可以是或可以约是0、1、2 3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、 10^9 、或在这些值的任何两个之间的数字或范围。在一些实施例中,具有相同子序列的第一多个间隔子寡核苷酸(或具有靶特异性靶结合区的寡核苷酸条形码)的数量可以是至少、至多0、1、2 3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、或 10^9 。例如,第一多个间隔子寡核苷酸(或具有靶特异性靶结合区的寡核苷酸条形码)中的至少两个可以包含相同的序列。

[0274] 第二多个寡核苷酸(例如,间隔子寡核苷酸和靶特异性寡核苷酸条形码)

[0275] 在一些实施例中,合成颗粒包含第二多个间隔子寡核苷酸(或具有靶特异性靶结合区的寡核苷酸条形码)。所述第二多个间隔子寡核苷酸中的每个可以不包含靶结合区。所述第一多个间隔子寡核苷酸中的一个或多个和所述第二多个间隔子寡核苷酸中的一个或多个可以具有相同的长度和不同的核苷酸序列。所述第一多个间隔子寡核苷酸中的一个或多个和所述第二多个间隔子寡核苷酸中的一个或多个可以具有不同长度和相似或不同的核苷酸序列。所述第一多个靶特异性寡核苷酸中的一个或多个和所述第二多个靶特异性寡核苷酸中的一个或多个可以具有相同的长度和不同的核苷酸序列。所述第一多个靶特异性寡核苷酸中的一个或多个和所述第二多个靶特异性寡核苷酸中的一个或多个可以具有不同长度和相似或不同的核苷酸序列。

[0276] 在一些实施例中,第二多个间隔子寡核苷酸(或具有靶特异性靶结合区的寡核苷酸条形码)的数量可以是或可以约是0、1、2 3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、 10^9 、或在这些值的任何两个之间的数字或范围。在一些实

施例中,第二多个间隔子寡核苷酸(或具有靶特异性靶结合区的寡核苷酸条形码)的数量可以是至少或至多0、1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、或 10^9 。

[0277] 第二多个间隔子寡核苷酸的第二间隔子寡核苷酸(或第二多个靶特异性寡核苷酸条形码的第二靶特异性寡核苷酸条形码)可以具有不同长度。相比第一多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)中的至少一个或每个,所述数量的第二多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)可以更短、可以具有相同长度,或可以更长。在一些实施例中,相比第一多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)中的至少一个或每个更短、具有相同长度、或更长的第二多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)的数量可以是或可以约是0、1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、 10^9 、或在这些值的任何两个之间的数字或范围。在一些实施例中,相比第一多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)中的至少一个或每个,更短、具有相同长度、或更长的第二多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)的数量可以是至少或至多0、1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、或 10^9 。例如,相比第一多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)中的至少一个,所述第二多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)中的至少一个可以更短。在一些实施例中,相比第一多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)中的至少一个或每个更短、具有相同长度、或更长的第二多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)的百分比可以是或可以约是1%、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%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、100%、或在这些值的任何两个之间的数字或范围。在一些实施例中,相比第一多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)中的至少一个或每个更短、具有相同长度、或更长的第二多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)的百分比可以是或可以约是1%、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%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%。

[0278] 在一些实施例中,相比第二多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)的

至少一个或每个更短、具有相同长度、或更长的第二多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)的数量可以是或可以约是0、1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、 10^9 、或在这些值的任何两个之间的数字或范围。在一些实施例中,相比第二多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)中的至少一个或每个更短、具有相同长度、或更长的第一多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)的数量可以是至少或至多0、1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、或 10^9 。例如,相比多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)中的至少一个,所述第二多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)中的至少一个可以更短。在一些实施例中,相比第二多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)中的至少一个或每个更短、具有相同长度、或更长的第一多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)的百分比可以是或可以约是1%、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%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、100%、或在这些值的任何两个之间的数字或范围。在一些实施例中,相比第二多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)中的至少一个或每个更短、具有相同长度、或更长的第一多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)的百分比可以是或可以约是1%、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%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、或100%。

[0279] 在不同的实现方式中,具有不同长度的第二多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)的数量可以是不同的。在一些实施例中,具有不同长度的第二多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)的数量可以是或可以约是0、1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、 10^9 、或在这些值的任何两个之间的数字或范围。在一些实施例中,具有不同长度的第二多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)的数量可以是至少或至多0、1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、或 10^9 。例如,所述第二多个间隔子寡核苷酸中的至少两个可以具有不同长度。在一些实施例中,具有不同长度的第二多个间隔子寡核苷酸(或靶特异

性寡核苷酸条形码)的百分比可以是或可以约是1%、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%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、100%、或在这些值的任何两个之间的数字或范围。在一些实施例中,具有不同长度的第二多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)的百分比可以是至少或至多1%、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%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%。

[0280] 在不同的实现方式中,具有相同序列的第二多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)的数量可以是不同的。在一些实施例中,具有相同序列的第二多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)的数量可以是或可以约是0、1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、 10^9 、或在这些值的任何两个之间的数字或范围。在一些实施例中,具有相同序列的第二多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)的数量可以是至少、至多0、1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、或 10^9 。例如,所述第二多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)中的至少两个的序列可以是相同的。在不同的实现方式中,具有相同子序列的第二多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)的数量可以是不同的。在一些实施例中,具有相同子序列的第二多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)的数量可以是或可以约是0、1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、 10^9 、或在这些值的任何两个之间的数字或范围。在一些实施例中,具有相同子序列的第二多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)的数量可以是至少、至多0、1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、或 10^9 。例如,第二多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码)中的至少两个可以包含相同的序列。

[0281] 靶结合区

[0282] 如图4B和4C1-4C2中所展示的,合成颗粒400b、400c1、400c2可以包括间隔子寡核苷酸420b、420c1、420c2和寡核苷酸条形码404。寡核苷酸条形码404可以包括能够捕获(例

如,杂交)靶分子408的靶结合区412。间隔子寡核苷酸420b可以包括间隔子区424b,其与靶结合区412相对应,不能捕获(例如,杂交)靶分子408(或靶分子的子集)。所述间隔子寡核苷酸420c1、420c2可以包括或缺少与靶结合区412相对应的区。在合成颗粒400b、400c1、400c2上包括间隔子寡核苷酸420b、420c1、420c2可以改善靶408的捕获(例如,mRNA捕获)和随后对捕获的靶分子408的定量。例如,合成颗粒400b、400c1、400c2可以与寡核苷酸条形码404(例如,各自具有聚(dT)区的条形码,所述条形码可以结合到mRNA分子或种类的拷贝或出现)和间隔子寡核苷酸420b、420c1、420c2两者偶联,这可以改善mRNA捕获和基因表达定量。在一些实施例中,靶结合区可以与目的靶杂交。在一些实施例中,靶结合区可以包含与靶(例如,靶核酸,如待分析的细胞核酸)特异性杂交的核酸序列。例如,靶结合区可以与靶核酸在特定的基因序列上杂交。在一些实施例中,靶结合区可以包含可附接(例如,杂交)至特定靶核酸的特定位置的核酸序列。在一些实施例中,靶结合区可以包含能够与限制性酶位点突出端(例如,EcoRI粘性末端突出端)进行特异性杂交的核酸序列。然后条形码可以连接到包括与限制性位点突出端互补的序列的任何核酸分子。

[0283] 在一些实施例中,靶结合区可以包含非特异性靶核酸序列(本文中还被称为非靶特异性靶结合区)。非特异性靶核酸序列可以指独立于靶核酸的特定序列可与多个靶核酸结合的序列。例如,靶结合区可以包含随机多聚体序列,或与mRNA分子上的聚(A)尾杂交的寡(dT)序列。随机多聚体序列可以是例如随机二聚体、三聚体、四聚体、五聚体、六聚体、七聚体、八聚体、九聚体、十聚体或任何长度的更高多聚体序列。在一些实施例中,对于附接到给定合成颗粒(例如,珠)的所有条形码,所述靶结合区是相同的。在一些实施例中,对于附接到给定合成颗粒上的多个条形码,靶结合区可以包含两个或更多个不同的靶结合序列。靶结合区的长度可以是或约5、10、15、20、25、30、35、40、45、50个核苷酸、或在这些值的任何两个之间的数字或范围的核苷酸。靶结合区的长度可以是至多约5、10、15、20、25、30、35、40、45、50个或更多个核苷酸。

[0284] 在一些实施例中,靶结合区可以包含寡聚(dT),所述寡聚(dT)可以与包含聚腺苷酸化(聚(A))末端的mRNA杂交。靶结合区可以是靶特异性的(例如,基因特异的)。例如,可以将靶结合区配置为与靶的特定区域杂交。靶结合区的长度可以是或是约1、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个核苷酸、或在这些值的任何两个之间的数字或范围的核苷酸。靶结合区的长度可以是至少、或至多1、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、或30个核苷酸。靶结合区的长度可以是约5至30个核苷酸。当条形码包括基因特异性靶结合区时,所述条形码可以称为基因特异性条形码。

[0285] 所述靶结合区可以包含基因特异性序列、寡聚(dT)序列、随机多聚体、或其任何组合。在不同的实现方式中,多个寡核苷酸条形码(具有不同基因特异性序列)的寡核苷酸条形码的靶结合区的数量可以是不同的。在一些实施例中,多个寡核苷酸条形码(具有不同基因特异性序列)的寡核苷酸条形码的靶结合区的数量可以是或可以约是0、1、2 3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、 10^9 、或在这些值的任何两个之间的数字或范围。在一些实施例中,多个寡核苷酸条形码(具有不同基因特异性序列)的寡核苷酸条形码的靶结合区的数量可以是至少或至多0、1、2 3、4、5、6、7、8、9、10、20、30、

40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、或 10^9 。在一些实施例中,多个寡核苷酸条形码(具有不同基因特异性序列)的寡核苷酸条形码的靶结合区的百分比可以是或可以约是1%、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%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、100%、或在这些值的任何两个之间的数字或范围。在一些实施例中,多个寡核苷酸条形码(具有不同基因特异性序列)的寡核苷酸条形码的靶结合区的百分比可以是至少或至多1%、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%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%。例如,多个寡核苷酸条形码中的两个寡核苷酸条形码的两个靶结合区可以包含不同的基因特异性序列。

[0286] 在一些实施例中,多个寡核苷酸条形码中的至少或至多100、1000、10000、100000、或更多个包含不同的分子标记序列。多个寡核苷酸条形码中的每个细胞标记可以包含至少、或至多1、2、3、4、5、6、7、8、9、10或更多个核苷酸。多个寡核苷酸条形码中的每个分子标记可以包含至少、或至多1、2、3、4、5、6、7、8、9、10或更多个核苷酸。多个寡核苷酸条形码中的每个可以包含空间标记,并且其中所述多个寡核苷酸条形码的空间标记彼此差异至少一个核苷酸。多个寡核苷酸条形码中的每个可以包含通用标记,其中所述多个寡核苷酸条形码的通用标记是相同的。

[0287] 在一些实施例中,所述多个寡核苷酸条形码中的至少一个固定在、部分固定在合成颗粒上,包封在、部分包封在合成颗粒中,或其任何组合。合成颗粒可以是可破坏的。合成颗粒可以包含颗粒。颗粒可以包括琼脂糖凝胶珠、链霉亲和素蛋白珠、琼脂糖珠、磁珠、缀合珠、蛋白A缀合珠、蛋白G缀合珠、蛋白A/G缀合珠、蛋白L缀合珠、寡(dT)缀合珠、二氧化硅珠、二氧化硅样珠、抗生物素微珠、抗荧光染料微珠、或其任何组合。合成颗粒可以包含选自下组的材料,该组由以下组成:聚二甲基硅氧烷(PDMS)、聚苯乙烯、玻璃、聚丙烯、琼脂糖、明胶、水凝胶、顺磁物质、陶瓷、塑料、玻璃、甲基苯乙烯、丙烯酸聚合物、钛、乳胶、琼脂糖凝胶、纤维素、尼龙、硅酮、及其任何组合。合成颗粒可以包含可破坏的水凝胶颗粒。所述多个寡核苷酸条形码中的每个可以包含接头官能团,合成颗粒可以包含固体支持物官能团,和/或所述支持物官能团和所述接头官能团可以彼此相关联。所述接头官能团和所述支持物官能团可以单独地选自下组,该组由以下各项组成:C6、生物素、链霉亲和素、一个或多个伯胺、一个或多个醛、一个或多个酮、及其任何组合。在一些实施例中,所述多个合成颗粒包含至少

94个合成颗粒。试剂盒可以包含微孔阵列,其中所述微孔阵列中的每个微孔包含所述多个合成颗粒中的合成颗粒。试剂盒可以包含缓冲液、盒、用于逆转录反应的一种或多种试剂、用于扩增反应的一种或多种试剂或其组合。

[0288] 本文的披露内容包括试剂盒或组合物的实施例。在一些实施例中,所述试剂盒或组合物包含:多个合成颗粒,其中每个合成颗粒包含:第一多个寡核苷酸条形码;和第二多个寡核苷酸条形码,其中所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码中的每个包含相同的细胞标记序列并且包含分子标记序列,其中所述第一多个寡核苷酸条形码和所述第二多个寡核苷酸条形码中的至少10个包含不同的分子标记序列,其中所述第一多个寡核苷酸条形码中的每个包含靶结合区,并且其中所述第二多个寡核苷酸条形码中的每个包含靶特异性靶结合区;以及任选地,用于使用所述多个合成颗粒的说明书。在一些实施例中,所述试剂盒包含缓冲液、用于逆转录反应的一种或多种试剂、用于扩增反应的一种或多种试剂或其组合。在一些实施例中,所述多个合成颗粒包含至少94个合成颗粒。所述试剂盒可以包含多个分区,每个分区包含所述多个合成颗粒中的合成颗粒。所述多个分区可以包含微孔阵列的微孔。所述多个分区可以包含多个液滴。在一些实施例中,所述第一多个寡核苷酸条形码的寡核苷酸条形码的靶结合区包含靶特异性靶结合区。所述第一多个寡核苷酸条形码的寡核苷酸条形码的靶结合区可以包含聚(dT)区。所述第二多个寡核苷酸条形码的寡核苷酸条形码可以包含聚(dT)区。

[0289] 本文提供的披露包括试剂盒的实施例,所述试剂盒包含本披露的一个或多个合成颗粒。在一些实施例中,所述试剂盒包含:多个合成颗粒,其中每个合成颗粒包含:多个寡核苷酸条形码,其中所述多个寡核苷酸条形码中的每个包含相同的细胞标记、分子标记和靶结合区,并且其中所述多个寡核苷酸条形码中的至少10个包含不同的分子标记序列;和第一多个间隔子寡核苷酸(或靶特异性寡核苷酸条形码),其中所述第一多个间隔子寡核苷酸中的每个不包含靶结合区。所述试剂盒可以包含:用于使用所述多个寡核苷酸条形码的说明书。

[0290] 使用具有间隔子寡核苷酸的合成颗粒进行条形码化

[0291] 本文披露的实施例包括如下方法,所述方法用于确定样品中靶的数量(参考图2和3进行描述)。在一些实施例中,所述方法包括:使用与合成颗粒相关联的多个寡核苷酸条形码,对样品中多个靶的拷贝进行条形码化,以产生多个经条形码化的靶的拷贝,其中所述多个寡核苷酸条形码中的每个包含相同的细胞标记、分子标记和靶结合区,并且其中所述多个寡核苷酸条形码中的至少10、100、1000、10000或更多个包含不同的分子标记序列,其中所述合成颗粒包含第一多个间隔子寡核苷酸,并且其中所述第一多个间隔子寡核苷酸中的每个不包含靶结合区;并且使用所述多个寡核苷酸条形码的分子标记,对所述样品中所述多个靶中每个的数量进行评估。在一些实施例中,使用与所述合成颗粒相关联的所述多个寡核苷酸条形码对样品中多个靶的拷贝进行条形码化可以包括:将所述多个靶与所述合成颗粒上的所述多个寡核苷酸条形码杂交。对所述多个靶的拷贝进行条形码化可以包括:逆转录所述多个靶,以产生所述多个靶的经条形码化的cDNA。所述方法可以包括:使用第二链合成,扩增所述经条形码化的cDNA,以产生经条形码化的靶,其中所述经条形码化的靶是双链的。所述方法可以包括:扩增所述经条形码化的靶,以产生经标记的靶扩增子。

[0292] 本文披露了如下方法的实施例,所述方法用于捕获样品中的靶。在一些实施例中,

所述方法包括：使用与合成颗粒相关联的多个寡核苷酸条形码，对样品中多个靶的拷贝进行杂交，以产生多个经条形码化的靶的拷贝，其中所述多个寡核苷酸条形码中的每个包含相同的细胞标记、分子标记和靶结合区，并且其中所述多个寡核苷酸条形码中的至少10、100、1000、10000或更多个包含不同的分子标记序列，其中所述合成颗粒包含第一多个间隔子寡核苷酸，并且其中所述第一多个间隔子寡核苷酸中的每个不包含靶结合区。所述方法可以包括：逆转录所述多个靶，以产生所述多个靶的经条形码化的cDNA。所述方法可以包括：使用第二链合成，扩增所述经条形码化的cDNA，以产生经条形码化的靶，其中所述经条形码化的靶是双链的。所述方法可以包括：包括扩增所述经条形码化的靶，以产生经标记的靶扩增子。

[0293] 多个经条形码化的靶的拷贝可以包含样品中多个靶中靶的拷贝的百分比、许多靶（例如2、3、4、5、6、7、8、9、10个或更多个靶），或样品中的每个靶。这种百分比是捕获效率的量度或指示。在一些实施例中，为产生多个经条形码化的靶的拷贝，与寡核苷酸条形码杂交的样品中多个靶中的靶、两个靶、三个靶或每个靶的拷贝的百分比可以是或可以约是1%、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%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、100%、或在这些值的任何两个之间的数字或范围。在一些实施例中，为产生多个经条形码化的靶的拷贝，与寡核苷酸条形码杂交的样品中多个靶中的靶、两个靶、三个靶或每个靶的拷贝的百分比可以是至少、或至多1%、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%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、或100%。

[0294] 在一些实施例中，所述样品包含单个细胞，所述方法包括将所述合成颗粒与所述样品中的所述单个细胞相关联。单个细胞可以包含稀有细胞、肿瘤细胞、来自人的细胞、来自组织的细胞、来自肿瘤的细胞、被病毒多核苷酸感染的细胞、及其任何组合。所述方法可以包括：在将所述合成颗粒与所述单个细胞相关联之后，裂解所述单个细胞。所述方法可以包括：裂解所述单个细胞包括将所述样品加热、使所述样品与洗涤剂接触、改变所述样品的pH、或其任何组合。所述合成颗粒和所述单个细胞可以在相同的孔中。所述合成颗粒和所述单个细胞可以在相同的液滴中。所述多个靶可以包含核糖核酸（RNA）、信使RNA（mRNA）、微小RNA、小干扰RNA（siRNA）、RNA降解产物、各自含有聚（A）尾的RNA、及其任何组合。

[0295] 靶

[0296] 靶可以是或包括DNA靶、RNA靶标、或其组合。在一些实施例中，靶是具有聚（A）尾的

mRNA靶。例如,靶可以是编码重免疫球蛋白(例如,IgA、IgD、IgE1、IgE2、IgG和IgM)和轻免疫球蛋白(例如,IgK和IgL)的mRNA。例如,具有带有对重免疫球蛋白和轻免疫球蛋白特异的靶结合区的寡核苷酸条形码的珠可用于确定在免疫细胞(例如像T细胞,B细胞)中重免疫球蛋白和轻免疫球蛋白的配对。此类珠可以用于确定抗体的重免疫球蛋白和轻免疫球蛋白的配对,从而确定抗体库。在一些实施例中,所述靶是没有聚(A)尾的RNA靶(在本文中也称为非聚腺苷酸化的靶),例如非编码RNA、组蛋白mRNA和病毒RNA(例如,呼肠孤病毒mRNA)。例如,靶可以来自两种或更多种生物,例如细胞的RNA以及细菌、真菌或病毒的RNA。细胞可以被细菌、真菌和/或病毒感染。例如,具有带有聚(dT)区(用于靶结合)的寡核苷酸条形码和具有针对病毒mRNA的靶结合区的寡核苷酸条形码的珠可用于确定在感染期间的细胞表型和病毒基因型。在一些实施例中,所述靶是DNA靶(例如,不具有聚(A)尾的基因组DNA片段)。

[0297] 在一些实施例中,靶(例如,基因的mRNA转录物)包括CD4、FOXO1、CD45RO、MYC、IL1R2、PRF1、GZMK、LGALS1、IL17F、IL23R、LYNX1、PRDM1、SELL、SMAD4、ICOS、IKZF5、RORC、AHRR、CTLA4、ITGB7、ENTPD1、CCR8、TSHR、TGFB2、IL12A、IL7R、HLA-DMA、CCR5、TIAF1、BCL6、BHLHE40、CXCR4、和CD307c。靶可以包括CD3D、GSTP1、TCF7、CD3E、RNB6、RB1、MYB、CD3G、KRT8、CDH1、ERBB3、ERBB2、TCTN1、ESR1、CDKN1A、和TFF3。靶可以包括ABCB1、ABCG2、ADAM23、AKT1、APC、AR、ATM、BAD、BCL2、BIRC5、BRCA1、BRCA2、C4A、CA12、CCNA1、CCND1、CCND2、CCNE1、CDH1、CDH13、CDK2、CD326、CDKN1A、CDKN1C、CDKN2A、CSF1、CST6、CTNBN1、CTSD、EGF、EGFR、EMAP-2、ERBB2、ERBB3、ESR1、ESR2、FOXA1、GATA3、GLI1、GPI、GRB7、GSTP1、HIC1、HPRT1、ID1、IGF1、IGF1R、IGFBP3、IL6、JUN、KRT18、KRT19、KRT5、KRT8、LAMP1、MAPK1、MAPK3、MAPK8、MGMT、MKI67、MLH1、MMP2、MMP9、MUC1、MYB、MYC、NME1、NOTCH1、NR3C1、PGR、PLAU、PRDM2、PSMB2、PSMB4、PTEN、PTGS2、PYCARD、RAB7A、RARA、RARB、RASSF1、RB1、REEP5、RNB6、SERPINE1、SFN、SFRP1、SLC39A6、SLIT2、SNAI2、SRC、TBC1D9、TCTN1、TFF3、TGFB1、THBS1、TP73、TWIST1、VEGFA、XBP1、CD3E、CD3G、CD3G、TCF7、ALCAM、CD25、ITGA6、THY1、PROM1、和CXCR4。靶可以包括BRCA1、BRCA2、TP53、PTEN、MSH2、MLH1、MSH6、PMS2、EPCAM、APC、RB1、MEN1、RET、和VHL。

[0298] 所述靶可能与以下有关:血液和淋巴疾病;癌症;消化系统;耳、鼻和喉;眼疾病;女性特有的疾病;男性特有的疾病;腺体和激素;心脏和血管;免疫系统疾病;男性特有的疾病;肌肉和骨骼;新生儿疾病;神经系统;营养和代谢性疾病;呼吸疾病;和/或皮肤和结缔组织。

[0299] 样品和目的细胞

[0300] 在一些实施例中,样品可以是或包括临床样品、生物样品、环境样品或其组合。例如,样品可包括来自受试者的一种或多种生物流体、组织和细胞。在一些实施例中,样品可以包括血液、尿液、脑脊液、胸膜液、羊水、精液、唾液、骨髓、活检样品、或其组合。在一些实施例中,目的细胞可包括干细胞、癌细胞、血细胞,外周血单核细胞、循环肿瘤细胞(CTC)、乳腺癌细胞、处于所需细胞周期阶段的细胞、或其组合。

[0301] 如本文所用,术语“目的细胞”是指正在研究的细胞。样品中目的细胞的非限制性实例包括表达恶性表型的细胞;肿瘤细胞,例如已经从肿瘤脱落到血液或其他体液或骨髓中的肿瘤细胞;良性肿瘤细胞;癌细胞;外周血中的癌细胞;甲状腺癌细胞;乳腺癌细胞;循环肿瘤细胞(“CTC”);白血病细胞;癌症干细胞;来自不同细胞周期阶段(G0/G1、S、G2)的单个细胞;带有X和Y染色体的精子;干细胞;胎儿或成人干细胞;多能干细胞;地中海贫血患者

者中的有核红细胞(“NRBC”);胎儿细胞,如母体外周血中的胎儿细胞;母体循环中的胎儿有核红细胞(“FNRBC”);以及由CD71、CD8、CD34或CD133表征的细胞。样品可以包含来自混合癌细胞样品中的细胞。

[0302] 目的细胞的其他实例包括但不限于以下细胞:循环内皮细胞;被病毒感染的细胞,例如被HIV感染的细胞,被目的基因转染的细胞;和存在于遭受自身免疫或自身反应性障碍的受试者的外周血中的T细胞或B细胞的异常亚型;活化的淋巴细胞;抗原呈递细胞,例如单核细胞和树突状细胞;被致病生物或寄生生物(例如,支原体)感染的细胞,含有细胞内寄生虫的细胞;以及在稀释流体(如尿液)中的细胞或微生物。

[0303] 细胞系的非限制性实例包括:Jurkat细胞(一种T白血病细胞系);SKBR3(一种腺癌衍生的乳腺癌细胞系,以Her2/neu的过度表达而熟知);T47D(一种导管癌衍生的乳腺癌细胞系,显示出由低至中等的Her2/neu表达);和HeLa。

[0304] 在一些实施例中,多个靶可以包括在一个或多个样品中。样品可包括一个或多个细胞、或者来自一个或多个细胞的核酸。样品可以是单细胞或来自单细胞的核酸。一个或多个细胞可以是一种或多种细胞类型。一种或多种细胞类型中的至少一种可以是脑细胞、心脏细胞、癌细胞、循环肿瘤细胞、器官细胞、上皮细胞、转移性细胞、良性细胞、原代细胞、循环细胞或其任何组合。

[0305] 用于在本披露的方法中使用的样品可以包括一个或多个细胞。样品可以指一个或多个细胞。在一些实施例中,多个细胞可包括一种或多种细胞类型。一种或多种细胞类型中的至少一种可以是脑细胞、心脏细胞、癌细胞、循环肿瘤细胞、器官细胞、上皮细胞、转移性细胞、良性细胞、原代细胞、循环细胞或其任何组合。在一些实施例中,该细胞是从癌组织切除的癌细胞,例如乳腺癌、肺癌、结肠癌、前列腺癌、卵巢癌、胰腺癌、脑癌、黑色素瘤和非黑色素瘤皮肤癌等。在一些实施例中,该细胞源自癌症,但是从体液收集(例如,循环肿瘤细胞)。癌症的非限制性实例可以包括腺瘤、腺癌、鳞状细胞癌、基底细胞癌、小细胞癌、大细胞未分化癌、软骨肉瘤、以及纤维肉瘤。样品可包括组织、单层细胞、固定的细胞、组织切片、或其任何组合。样品可包括生物学样品、临床样品、环境样品、生物学流体、来自受试者的组织或细胞。样品可以获得自人、哺乳动物、狗、大鼠、小鼠、鱼、蝇、蠕虫、植物、真菌、细菌、病毒、脊椎动物、或无脊椎动物。

[0306] 在一些实施例中,该细胞是已经被病毒感染并含有病毒寡核苷酸的细胞。在一些实施例中,病毒感染可以由如单链(+链或“有义”)DNA病毒(例如,细小病毒)、或双链RNA病毒(例如,呼吸道肠道病毒)的病毒引起。本文披露的珠可用于确定经感染的细胞的细胞表型和感染细胞的病毒的基因型。例如,具有带有聚(dT)区(用于靶结合)的寡核苷酸条形码和具有针对病毒mRNA的靶结合区的寡核苷酸条形码的珠可用于同时确定感染期间的细胞表型和/或基因型和病毒基因型和/或表型。在一些实施例中,细胞是已经被细菌(或真菌)感染并含有细菌寡核苷酸(或真菌寡核苷酸)的细胞。本文披露的珠可用于同时确定表型被感染细胞的基因型和/或表型以及感染生物的基因型和/或表型。在一些实施例中,所述细胞是细菌。这些可以包括革兰氏阳性细菌或革兰氏阴性细菌。在一些实施例中,所述细胞是真菌。在一些实施例中,所述细胞是原生动物或其他寄生虫。在一些实施例中,所述细胞是寄生虫的宿主细胞。本文披露的珠可用于同时确定宿主细胞的表型和/或基因型以及寄生虫的基因型和/或表型。

[0307] 如本文使用的,术语“细胞”可以指一个或多个细胞。在一些实施例中,细胞是正常细胞,例如,处于不同发育阶段的人细胞,或来自不同器官或组织类型的人细胞。在一些实施例中,该细胞是非人类细胞,例如其他类型的哺乳动物细胞(例如小鼠、大鼠、猪、狗、牛或马)。在一些实施例中,所述细胞是其他类型的动物或植物细胞。在其他实施例中,细胞可以是任何原核或真核细胞。

[0308] 在一些实施例中,在将细胞与珠关联之前对所述细胞进行分选。例如,所述细胞可以通过荧光活化细胞分选或磁活化细胞分选进行分选,或更普遍地通过流式细胞术进行分选。可以按尺寸过滤细胞。在一些实施例中,滞留物包含有待与珠相关联的细胞。在一些实施例中,流过物(flow through)包含有待与珠相关联的细胞。

[0309] 样品可以指多个细胞。样品可以指单层细胞。样品可以指薄片(例如,组织薄片)。样品可以指细胞的固体或半固体集合,其可以放置在阵列上的一个维度上。

[0310] 测序

[0311] 在一些实施例中,估算不同的经条形码化的靶(例如经随机条形码化的靶)的数目可以包括确定经标记的靶、空间标记、分子标记、样品标记、细胞标记、或其任何产物(例如经标记的扩增子、或经标记的cDNA分子)的序列。扩增的靶可以经受测序。确定经条形码化的靶(例如经随机条形码化的靶)或其任何产物的序列可以包括进行测序反应以确定样品标记、空间标记、细胞标记、分子标记的至少一部分的序列,经标记的靶(例如经随机标记的靶)、其互补体、其反向互补体的至少一部分的序列,或其任何组合的序列。

[0312] 经条形码化的靶或随机地经条形码化的靶(例如,经扩增的核酸、经标记的核酸、经标记的核酸的cDNA拷贝等)的序列的确定可以使用如下各种测序方法来进行,所述方法包括但不限于:杂交测序(SBH)、连接测序(SBL)、定量增量荧光核苷酸添加测序(QIFNAS)、逐步连接和切割、荧光共振能量转移(FRET)、分子信标、TaqMan报告探针消化、焦磷酸测序、原位荧光测序(FISSEQ)、FISSEQ珠、摆动测序、多重测序、聚合菌落(POLONY)测序;纳米网格滚环测序(ROLONY)、等位基因特异性寡聚连接测定(例如,寡聚连接测定(OLA)、使用连接的线性探针和滚环扩增(RCA)读数的单一模板分子OLA、连接的挂锁探针、使用连接的圆形挂锁探针和滚环扩增(RCA)读数的或单一模板分子OLA)等。

[0313] 在一些实施例中,确定经条形码化的靶(例如经随机条形码化的靶)或其任何产物的序列包括配对端测序、纳米孔测序、高通量测序、鸟枪法测序、染料终止剂测序、多重引物DNA测序、引物步移、桑格双脱氧测序法、马克西姆-吉尔伯特(Maxim-Gilbert)测序、焦磷酸测序、真正的单分子测序或其任何组合。可替代地,可以通过电子显微镜分析法或化学敏感场效应晶体管(chemFET)阵列来确定经条形码化的靶或其任何产物的序列。

[0314] 可以使用高通量测序方法,如使用平台(如Roche 454、Illumina Solexa、ABI-SOLiD、ION Torrent、Complete Genomics、Pacific Bioscience、Helicos、或Polonator平台)的循环阵列测序。在一些实施例中,测序可以包括MiSeq测序。在一些实施例中,测序可以包括HiSeq测序。

[0315] 经标记的靶(例如经随机标记的靶)可以包括代表来自生物体基因组基因的约0.01%至生物体基因组基因的约100%的核酸。例如,可以使用包括多个多聚体的靶标互补区域,通过从所述样品中捕获含有互补序列的基因,对约0.01%的生物体基因组基因至约100%的生物体基因组基因进行测序。在一些实施例中,经条形码化的靶包括代表从约

0.01%的生物体转录组转录物至约100%的生物体转录组转录物的核酸。例如,可以使用包括聚(T)尾的靶标互补区域,通过从所述样品中捕获mRNA,对约0.501%的生物体转录组转录物至约100%的生物体转录组转录物进行测序。

[0316] 确定多个条形码(例如随机条形码)的空间标记和分子标记的序列可包括对多个条形码的0.00001%、0.0001%、0.001%、0.01%、0.1%、1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、20%、30%、40%、50%、60%、70%、80%、90%、99%、100%、或在这些值的任何两个之间的数字或范围进行测序。确定多个条形码的标记(例如样品标记、空间标记、和分子标记)的序列可包括对多个条形码中的1、10、20、30、40、50、60、70、80、90、100、 10^3 、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、 10^9 、 10^{10} 、 10^{11} 、 10^{12} 、 10^{13} 、 10^{14} 、 10^{15} 、 10^{16} 、 10^{17} 、 10^{18} 、 10^{19} 、 10^{20} 个、或在这些值的任何两个之间的数字或范围进行测序。对多个条形码中的一些或所有进行测序可包括产生具有、具有约、具有至少、或具有至多10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、10000个、或在这些值的任何两个之间的数字或范围内的核苷酸或碱基的读段长度的序列。

[0317] 测序可包括对经条形码化的靶的至少或至少约10、20、30、40、50、60、70、80、90、100个或更多个核苷酸或碱基对进行测序。例如,测序可包括通过对多个经条形码化的靶进行聚合酶链式反应(PCR)扩增,产生测序数据,其中序列具有50、75、或100个、或更多个核苷酸的读段长度。测序可包括对经条形码化的靶的至少或至少约200、300、400、500、600、700、800、900、1,000个或更多个核苷酸或碱基对进行测序。测序可包括对经条形码化的靶的至少或至少约1500、2000、3000、4000、5000、6000、7000、8000、9000、或10000个或更多个核苷酸或碱基对进行测序。

[0318] 测序可以包含至少约200、300、400、500、600、700、800、900、1,000个或更多个测序读段/运行。在一些实施例中,测序包括对至少或至少约1500、2000、3000、4000、5000、6000、7000、8000、9000、或10000个或更多个测序读段/运行进行测序。测序可以包括小于或等于约1,600,000,000个测序读段/运行。测序可以包括小于或等于约200,000,000个读段/运行。

[0319] 实例

[0320] 以上讨论的实施例的一些方面在以下实例中进一步详细披露,其不是旨在以任何方式限制本披露的范围。

[0321] 实例1

[0322] 使用具有间隔子寡核苷酸的珠改善RNA捕获

[0323] 该实施显示,与不具有间隔子寡核苷酸的珠相比,具有间隔子寡核苷酸的珠可以用更高的效率捕获RNA分子。

[0324] 对于具有间隔子寡核苷酸的珠,使用批量RNA的定量聚合酶链式反应(qPCR)实验显示,与不具有间隔子寡核苷酸的珠相比,提高了捕获效率(图6)。图6是柱图,显示使用珠,在RNA捕获中的非限制性示例性改善,每个珠包含具有能够捕获RNA分子的寡聚(dT)尾和多个间隔子寡核苷酸的多个寡核苷酸条形码。将相同量的RNA与10000个珠杂交(具有或不具有间隔子寡核苷酸)。从捕获的RNA分子生成cDNA后,用扩增GAPDH的引物进行qPCR。与不具有间隔子寡核苷酸的珠相比,具有间隔子寡核苷酸的珠低约0.3-0.5Cqs(定量循环,来自扩增的荧光超过背景荧光的循环)。这意味着,相对于仅有dT的珠捕获的GAPDH,间隔子珠捕获

的GAPDH的量约为125% - 138%。

[0325] 总之,这些数据表明可通过使用间隔子寡核苷酸降低寡核苷酸条形码的密度来提高RNA捕获效率。

[0326] 实例2

[0327] 靶特异性RNA捕获

[0328] 该实例显示,与具有用于RNA捕获的聚(dT)区的寡核苷酸的珠相比,具有靶特异性寡核苷酸的珠可以用更高的效率捕获目的RNA分子。

[0329] 靶特异性寡核苷酸包括具有序列5' - ACAAGTCTGTCTG - 3' (SEQ ID NO:4) 和5' - CACACCCAAAAGGCCACACTGGT - 3' (SEQ ID NO:5) 的靶结合区,所述序列对TCR α 和TCR β 的恒定区具有特异性。对于具有靶特异性寡核苷酸(对TCR α 和TCR β 的恒定区具有特异性)的珠,用批量RNA的定量聚合酶链式反应(qPCR)实验显示与具有聚(dT)序列的寡核苷酸的珠相比,提高了捕获效率(图143)。图143是显示使用具有TCR特异性寡核苷酸条形码的珠,T细胞受体(TCR)捕获效率的非限制性示例性比较的柱图。将相同量的RNA杂交到10000个珠上(具有带有聚(dT)靶结合区的寡核苷酸的珠、或具有带有聚(dT)靶结合区以及对TCR α 和TCR β 的恒定区特异性的靶结合区的寡核苷酸的珠)。从捕获的RNA分子生成cDNA后,使用靶向TCR α 可变区和恒定区以及TCR β 可变区和恒定区的引物进行qPCR,以扩增TCR α 和TCR β 。相对于被具有带有用于非特异性靶结合的寡聚(dT)区的寡核苷酸条形码的珠捕获的TCR α ,被具有靶特异性寡核苷酸条形码的珠捕获的TCR α 分子的量为约260%。相对于被具有带有用于非特异性靶结合的寡聚(dT)区的寡核苷酸条形码的珠捕获的TCR β ,被具有靶特异性寡核苷酸条形码的珠捕获的TCR β 分子的量为约425%。

[0330] 总之,这些数据表明使用具有靶特异性靶结合区的寡核苷酸条形码可以提高特异性RNA捕获效率。

[0331] 在至少一些先前描述的实施例中,在一个实施例中使用的一个或多个元素可以互换地用于另一个实施例中,除非这种替换在技术上不可行。本领域技术人员将理解,在不脱离所要求保护的的主题的范围的情况下,可以对上述方法和结构进行各种其他的省略、添加和修改。所有此类修改和改变都旨在落在由所附权利要求书限定的主题的范围之内。

[0332] 关于本文中基本上任何复数和/或单数术语,在对于背景和/或应用适当的情况下,本领域技术人员可以从复数转换为单数和/或从单数转换为复数。为了清楚起见,可以在本文明确阐述各种单数/复数排列。如本说明书和所附权利要求书中使用的,除非上下文另有明确指示,否则单数形式“一个/一种(a/an)”和“所述(the)”包括复数的提及物。除非另外说明,在本文中对“或”的任何提及旨在涵盖“和/或”。

[0333] 本领域技术人员将理解,一般来说,本文使用的术语,尤其是所附权利要求书(例如,所附权利要求书的主体)中的术语,通常旨在作为“开放性的”术语(例如,术语“包括(including)”应解释为“包括但不限于(including but not limited to)”,术语“具有(having)”应解释为“具有至少(having at least)”,术语“包括(includes)”应解释为“包括但不限于(includes but is not limited to)”等)。本领域技术人员将进一步理解,如果预期到所介绍的权利要求陈述的特定数目,这样的预期将明确地陈述于权利要求中,并且在不存在这种陈述的情况下没有这种意图存在。例如,作为对理解的帮助,以下所附权利要求书可以包含介绍性短语“至少一个”和“一个或多个”的使用,以介绍权利要求陈述。然

而,此类短语的使用不应解读为意味着由不定冠词“一个”或“一种”介绍权利要求陈述会将任何包含这种介绍的权利要求陈述的具体权利要求限制到包含仅一个这种陈述的实施例中,甚至当相同的权利要求包括介绍性短语“一个或多个”或“至少一个”以及不定冠词如“一个”或“一种”时也是如此(例如,“一个”和/或“一种”应解释为意指“至少一个”或“一个或多个”);这对于使用定冠词来介绍权利要求陈述同样适用。此外,即使明确地陈述了介绍的权利要求陈述的特定数目,本领域技术人员将认识到,这种陈述应解释为意指至少所陈述的数字(例如,仅陈述“两个陈述”而没有其他修饰词意指至少两个陈述、或两个或更多个陈述)。此外,在使用类似于“A、B和C等中的至少一个”的惯例的那些情况下,通常这种句法结构是在本领域技术人员将理解该惯例的意义上预期(例如,“具有A、B和C中的至少一个的系统”将包括但不限于仅具有A,仅具有B,仅具有C,A和B一起,A和C一起,B和C一起,和/或A、B、和C一起等的系统)。在使用类似于“A、B或C等中的至少一个”的惯例的那些情况下,通常这种句法结构是在本领域技术人员将理解该惯例的意义上预期(例如,“具有A、B或C中的至少一个的系统”将包括但不限于仅具有A,仅具有B,仅具有C,A和B一起,A和C一起,B和C一起,和/或A、B、和C一起等的系统)。本领域技术人员将进一步理解,实际上,无论在说明书、权利要求书还是在附图中,呈现两个或更多个替代术语的任何分离性词语和/或短语应被理解为考虑到包括术语之一、任一术语或两个术语的可能性。

[0334] 此外,当本披露的特征或方面以马库什组(Markush group)描述时,本领域技术人员将意识到本披露还由此以马库什组的任何单独的成员或成员子组描述。

[0335] 如本领域技术人员将理解的,出于任何和所有目的,如在提供书面描述方面,本文披露的所有范围还包括任何和所有可能的它的子范围和子范围组合。任何列出的范围都可以很容易地被识别为充分描述并使相同的范围能被分解为至少相等的一半,三分之一,四分之一,五分之一,十分之一等。作为非限制性示例,这里讨论的每个范围可以容易地分解为下三分之一,中三分之一和上三分之一等。如本领域技术人员还将理解的,所有语言,如“多至”、“至少”、“大于”、“小于”等包括所陈述的数字,并且指代可以随后分解为如上讨论的子范围的范围。最后,如本领域技术人员将理解的,范围包括每个单独的成员。因此,例如,具有1-3个物品的组是指具有1、2或3个物品的组。类似地,具有1-5个物品的组指代具有1、2、3、4或5个物品的组,等等。

[0336] 尽管本文已经披露了各种方面和实施例,但其他方面和实施例对本领域技术人员将是明显的。本文披露的各种方面和实施例用于说明的目的而并不意于限制由以下权利要求所指出的真实范围和精神。

- [0001] 序列表
- [0002] <110> 赛卢拉研究公司(Cellular Research, Inc.)
- [0003] 格雷琴·因博恩·拉姆 (Lam, Gretchen Yinbon)
- [0004] 艾琳·夏姆 (Shum, Eleen)
- [0005] <120> 与寡核苷酸相关联的颗粒
- [0006] <130> BDCRI.030W0
- [0007] <150> 62/607,812
- [0008] <151> 2017-12-19
- [0009] <160> 5
- [0010] <170> PatentIn 3.5版
- [0011] <210> 1
- [0012] <211> 103
- [0013] <212> DNA
- [0014] <213> 人工序列
- [0015] <220>
- [0016] <223> 合成寡核苷酸
- [0017] <220>
- [0018] <221> n
- [0019] <222> (78) .. (85)
- [0020] <223> n = a、t、c或g
- [0021] <220>
- [0022] <221> 尚未归类的特征
- [0023] <222> (78) .. (85)
- [0024] <223> n是a、c、g或t
- [0025] <400> 1
- [0026] tctctctcta cacgacgctc ttccgatctg atggatgact ggcctgcgac tacacgaggt 60
- [0027] agcggtgact tccaaggnnn nnnnnttttt tttttttttt ttt 103
- [0028] <210> 2
- [0029] <211> 103
- [0030] <212> DNA
- [0031] <213> 人工序列
- [0032] <220>
- [0033] <223> 合成寡核苷酸
- [0034] <220>
- [0035] <221> n
- [0036] <222> (78) .. (103)
- [0037] <223> n = a、t、c或g
- [0038] <220>

[0039]	<221>	尚未归类的特征	
[0040]	<222>	(78) .. (103)	
[0041]	<223>	n是a、c、g或t	
[0042]	<400>	2	
[0043]		tctctctcta cacgacgctc ttccgatctg atggatgact ggcctgacgac tacacgaggt	60
[0044]		agcgggtgact tccaaggnnn nnnnnnnnnn nnnnnnnnnn nnn	103
[0045]	<210>	3	
[0046]	<211>	121	
[0047]	<212>	DNA	
[0048]	<213>	人工序列	
[0049]	<220>		
[0050]	<223>	合成寡核苷酸	
[0051]	<220>		
[0052]	<221>	n	
[0053]	<222>	(75) .. (85)	
[0054]	<223>	n = a、t、c或g	
[0055]	<220>		
[0056]	<221>	尚未归类的特征	
[0057]	<222>	(78) .. (85)	
[0058]	<223>	n是a、c、g或t	
[0059]	<220>		
[0060]	<221>	n	
[0061]	<222>	(104) .. (121)	
[0062]	<223>	n = a、t、c或g	
[0063]	<220>		
[0064]	<221>	尚未归类的特征	
[0065]	<222>	(104) .. (121)	
[0066]	<223>	n是a、c、g或t	
[0067]	<400>	3	
[0068]		tctctctcta cacgacgctc ttccgatctg atggatgact ggcctgacgac tacacgaggt	60
[0069]		agcgggtgact tccaaggnnn nnnnnttttt tttttttttt tttnnnnnnn nnnnnnnnnn	120
[0070]		n	121
[0071]	<210>	4	
[0072]	<211>	13	
[0073]	<212>	DNA	
[0074]	<213>	人工序列	
[0075]	<220>		
[0076]	<223>	合成寡核苷酸	
[0077]	<400>	4	

[0078]	acaagtctgt ctg	13
[0079]	<210> 5	
[0080]	<211> 23	
[0081]	<212> DNA	
[0082]	<213> 人工序列	
[0083]	<220>	
[0084]	<223> 合成寡核苷酸	
[0085]	<400> 5	
[0086]	cacacccaaa aggccacact ggt	23

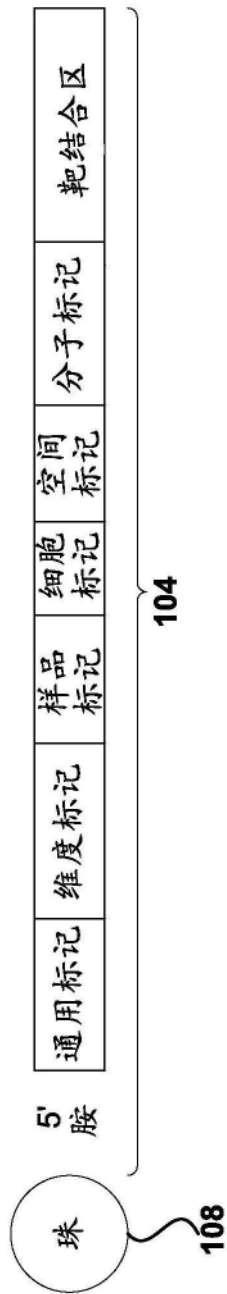


图1

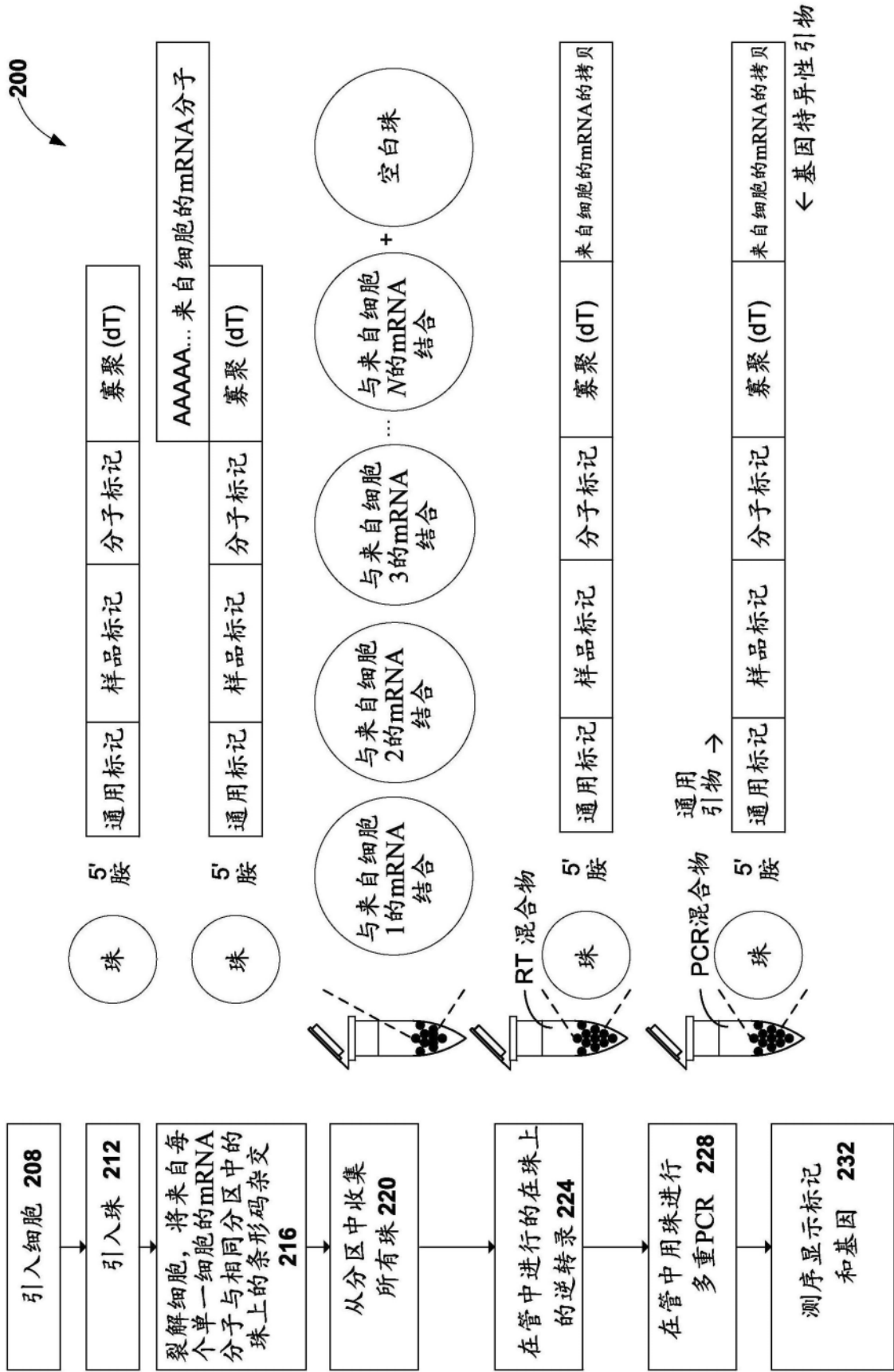


图2

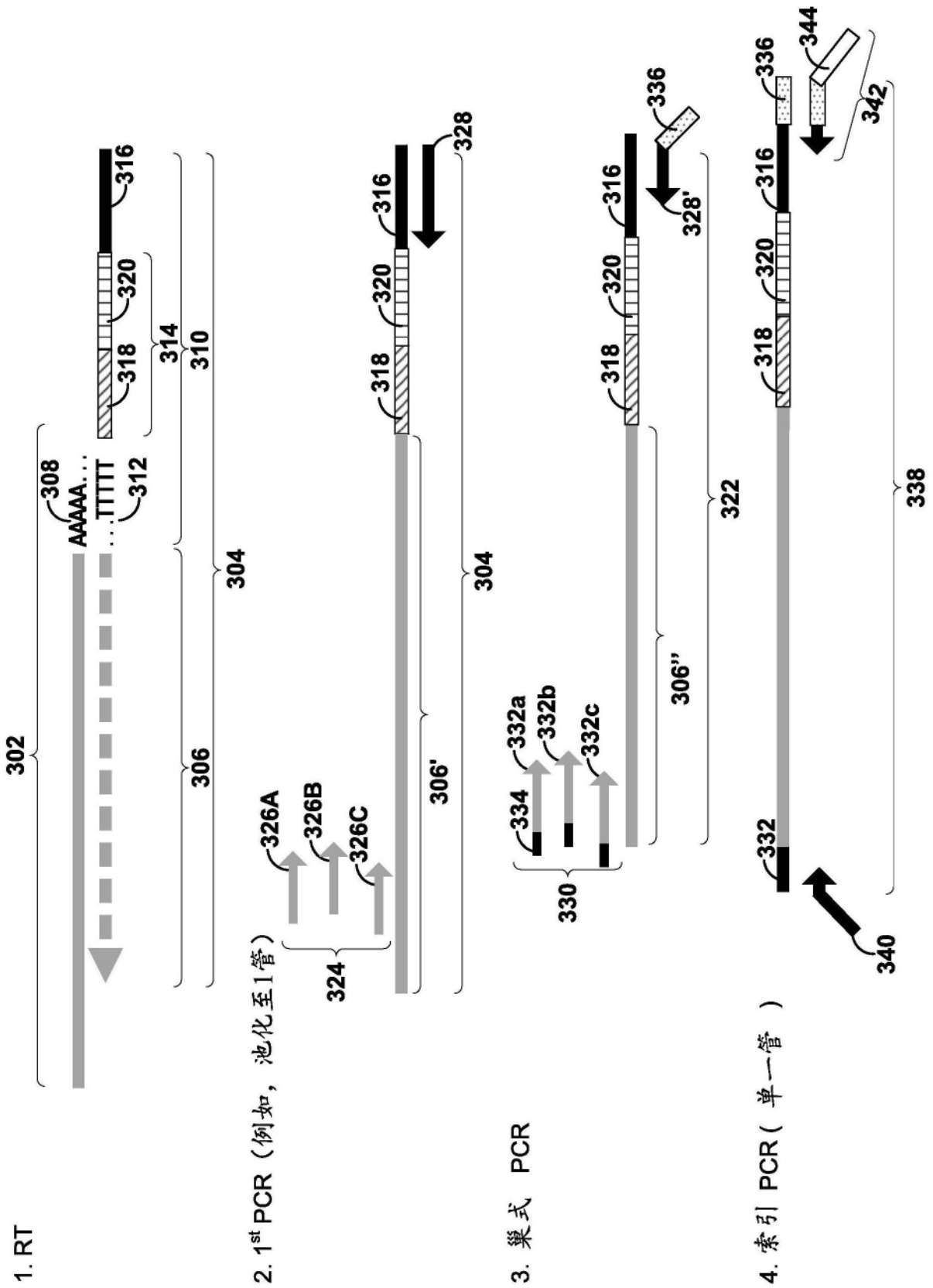


图3

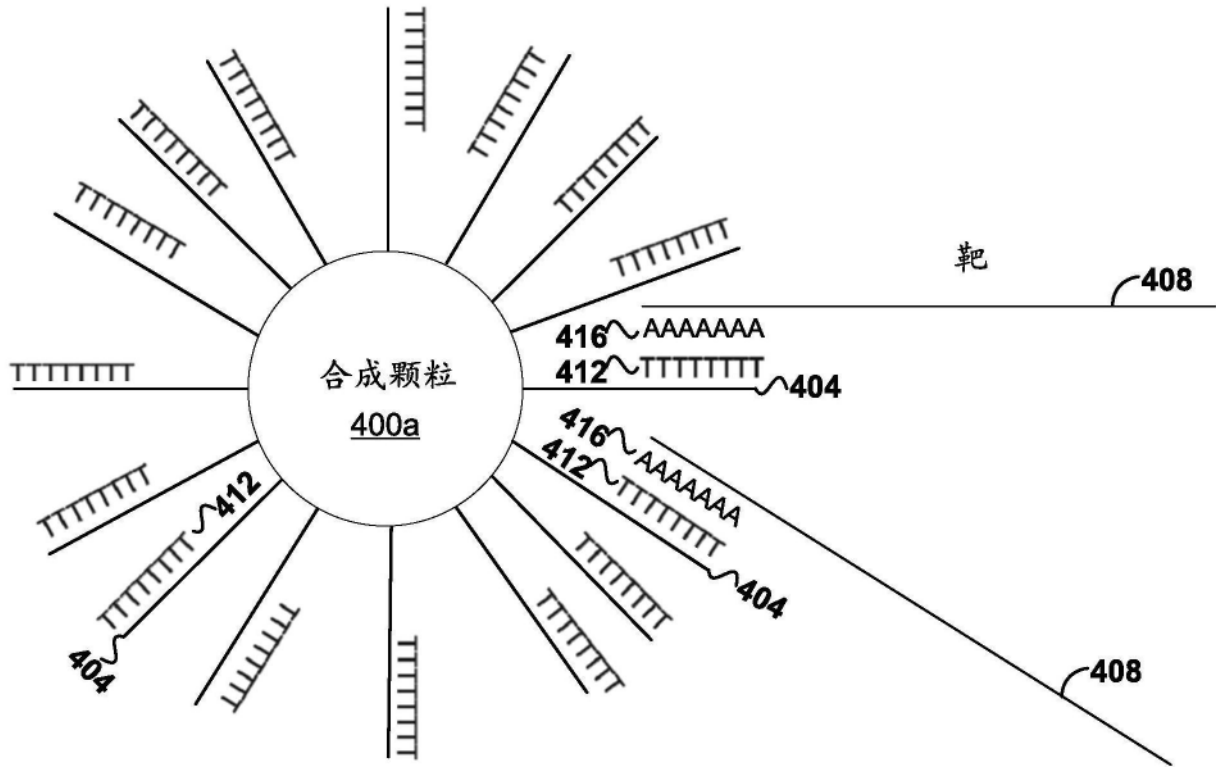


图4A

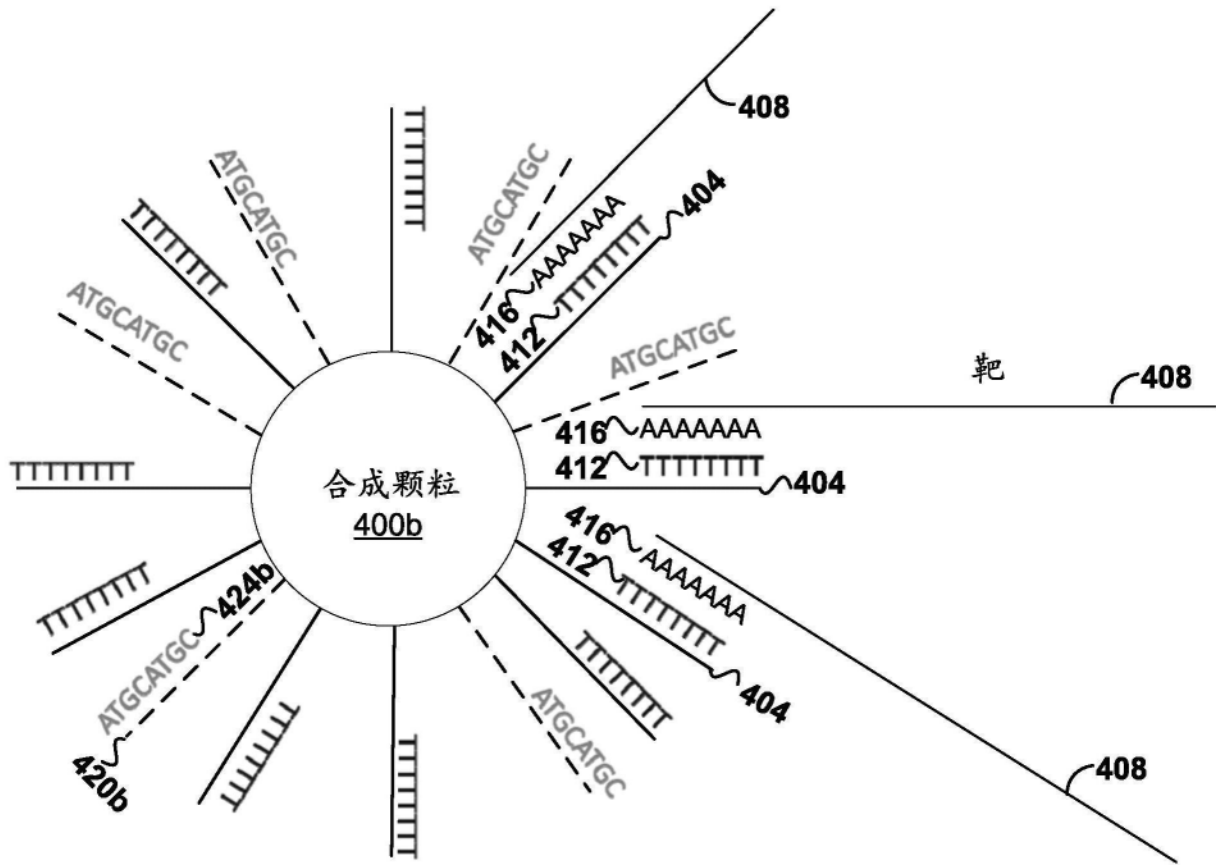


图4B

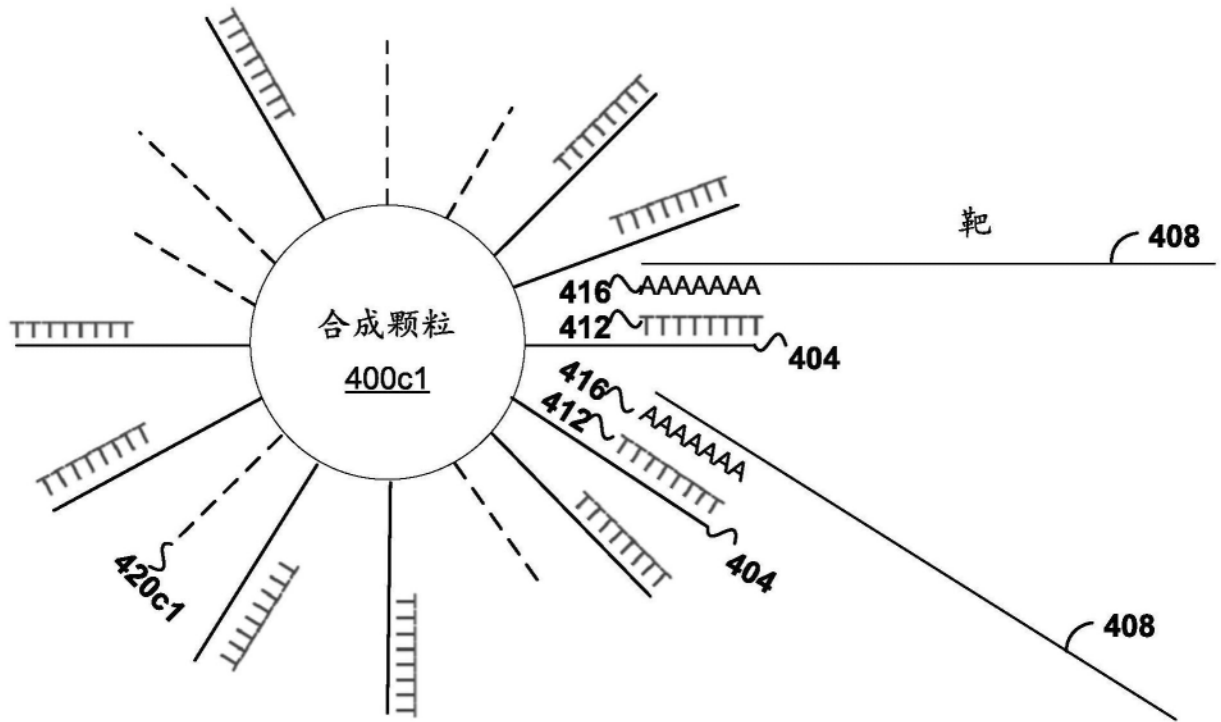


图4C1

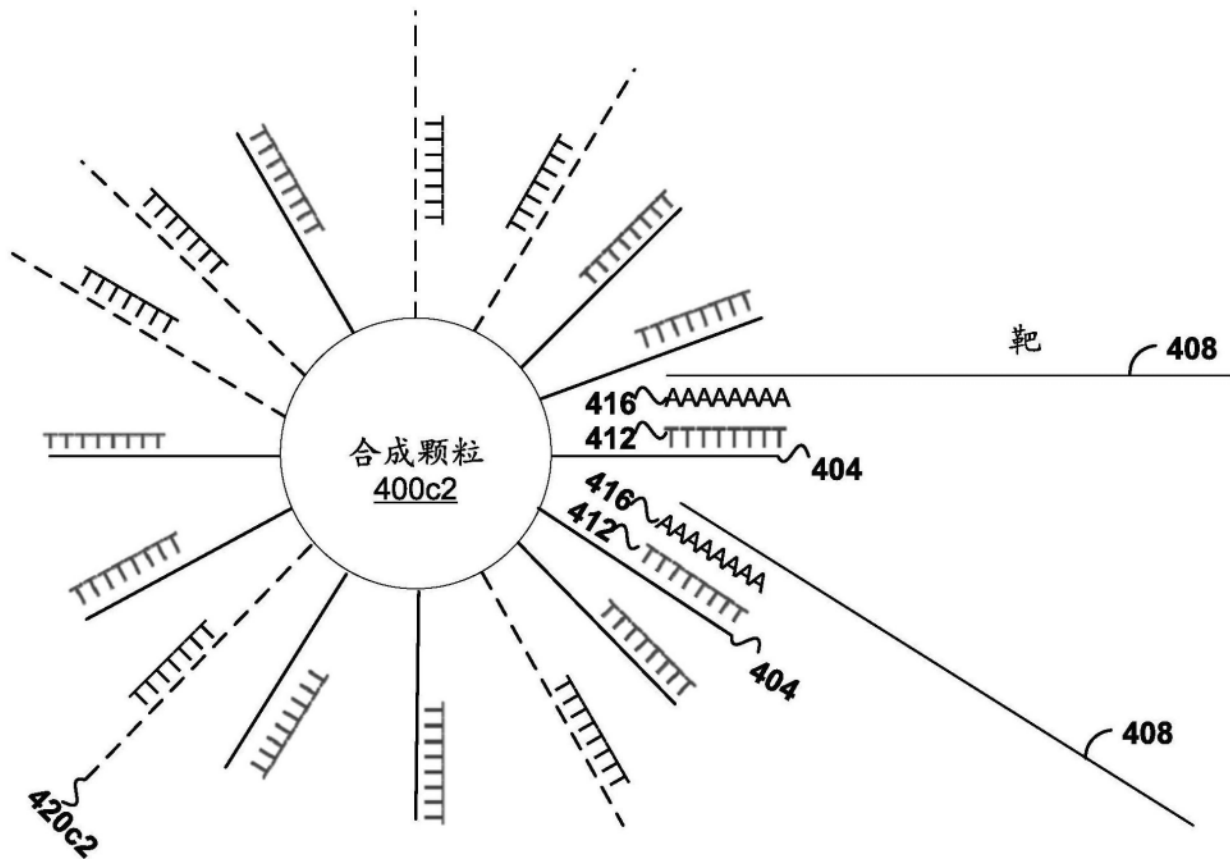


图4C2

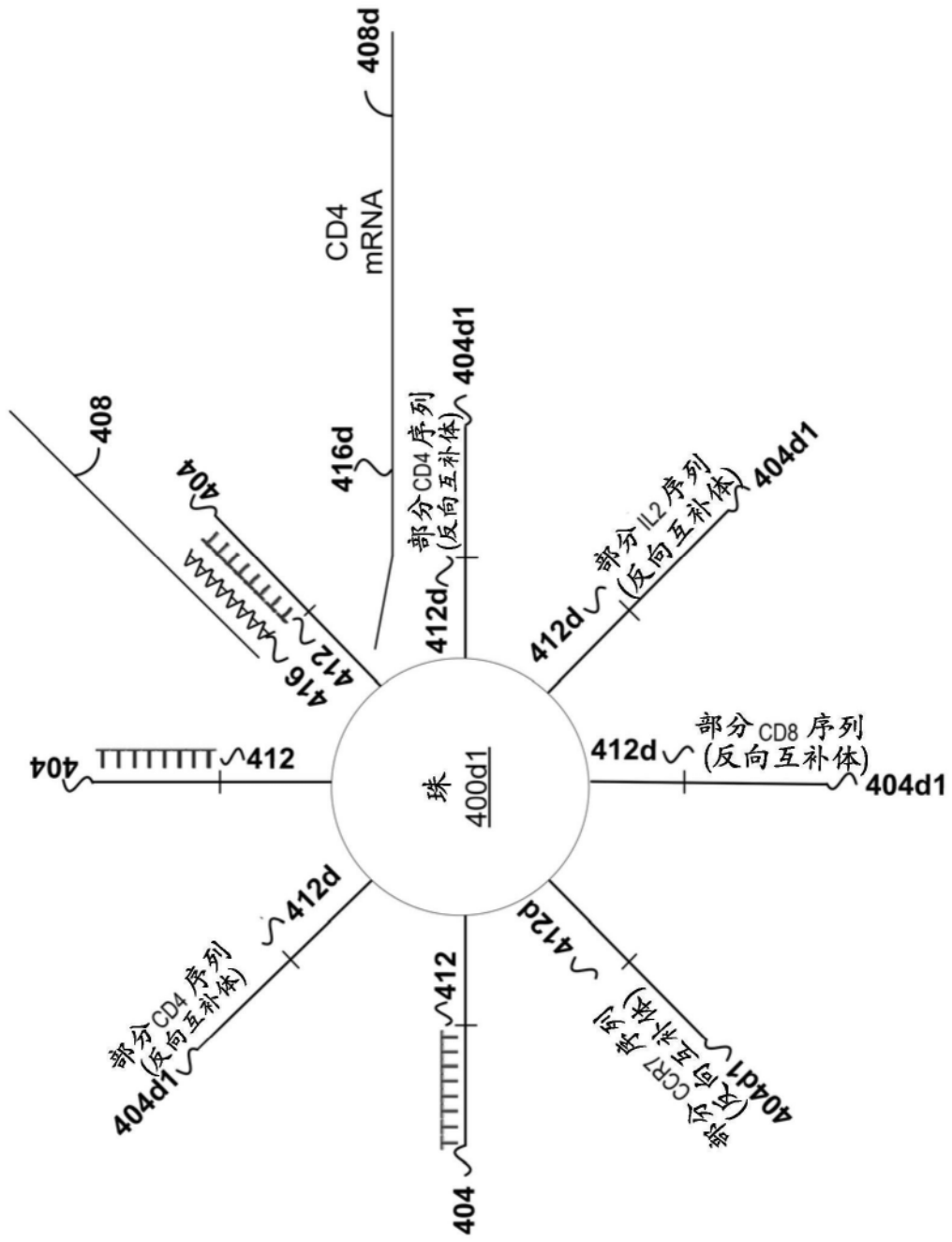


图4D1

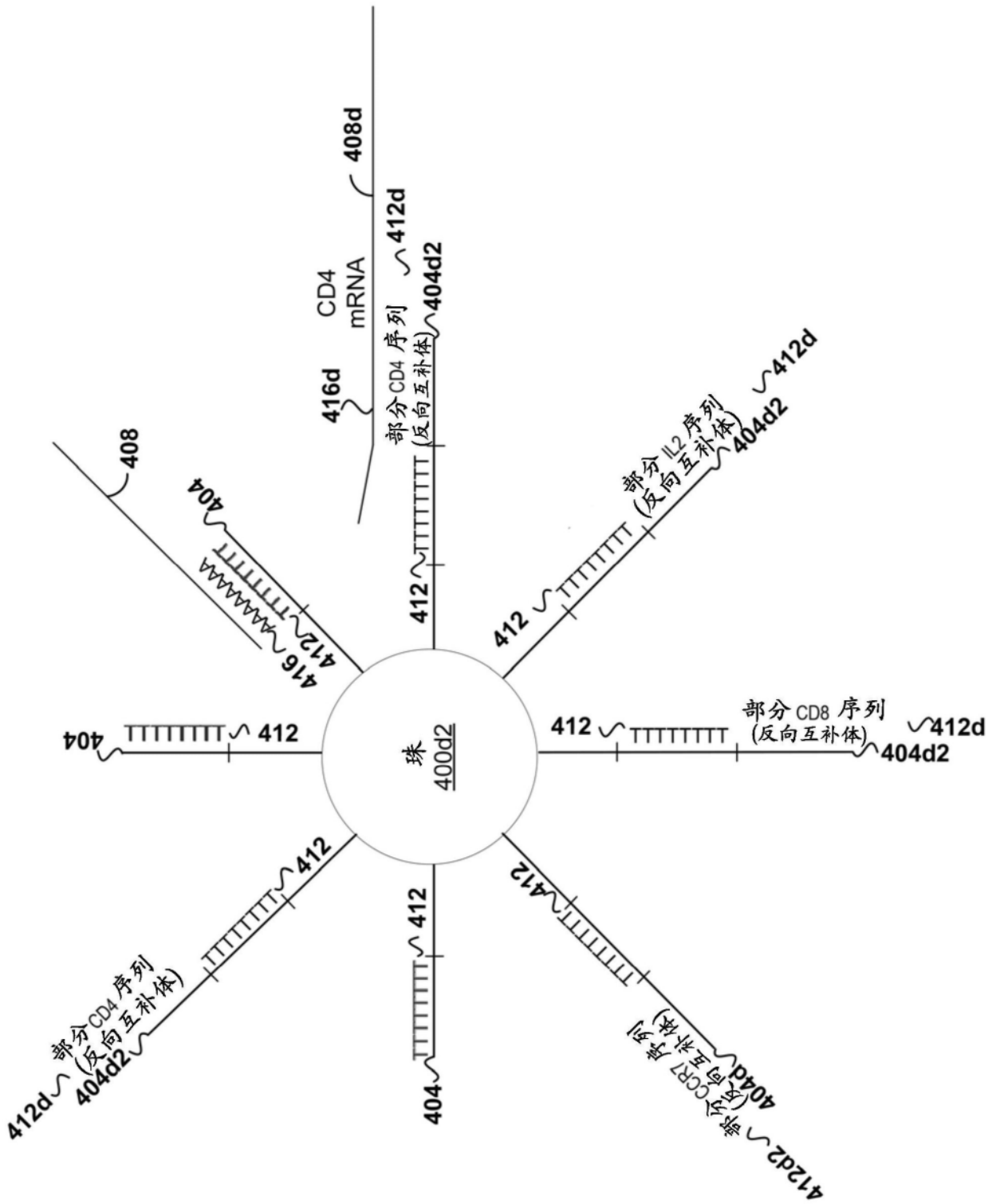


图4D2

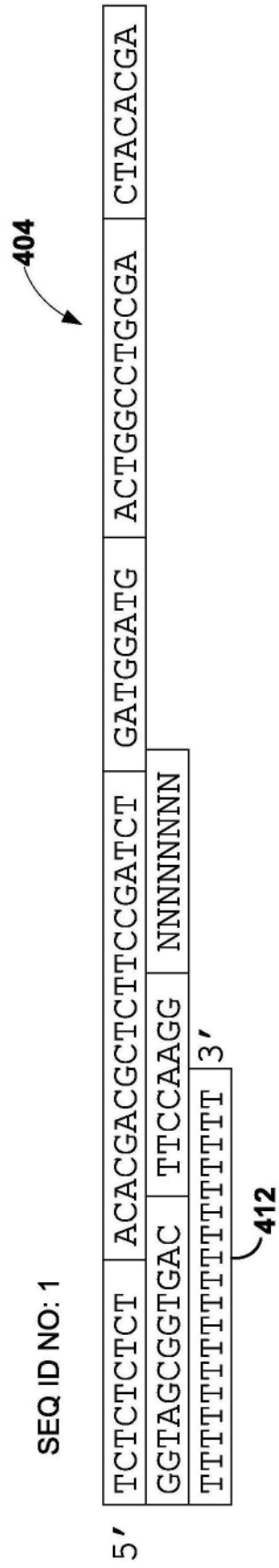


图5A

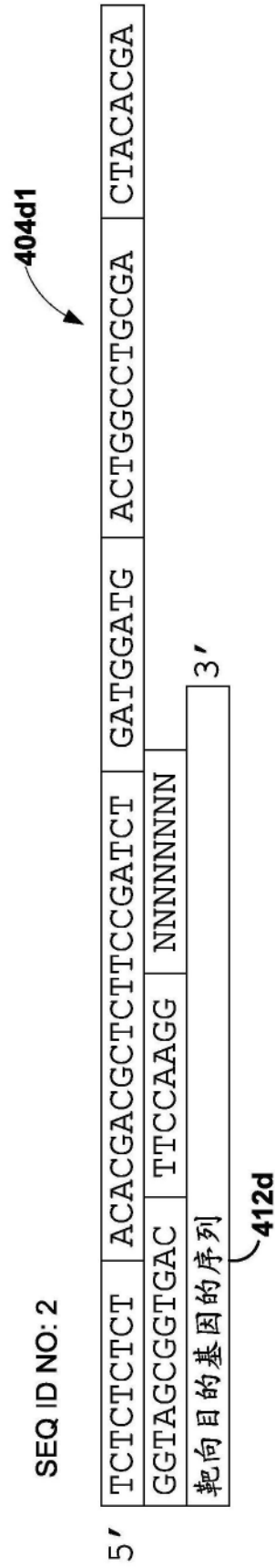
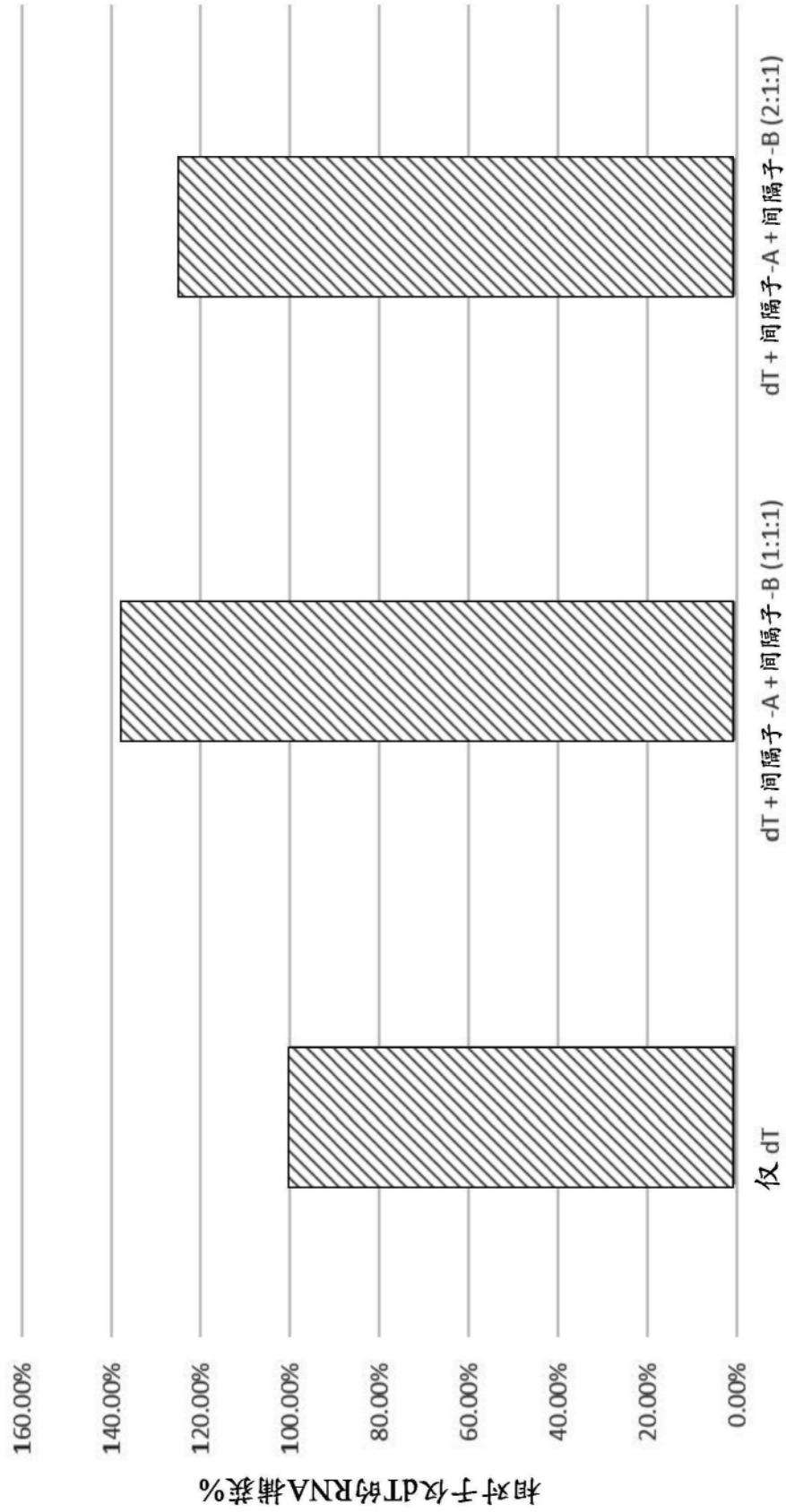


图5B



图5C

针对具有或不具有间隔子的珠进行的RNA捕获效率的比较



珠上的寡聚组合物

图6

针对具dT相比于dT+TCR α特异性+TCR β特异性珠进行的TCR捕获效率的比较

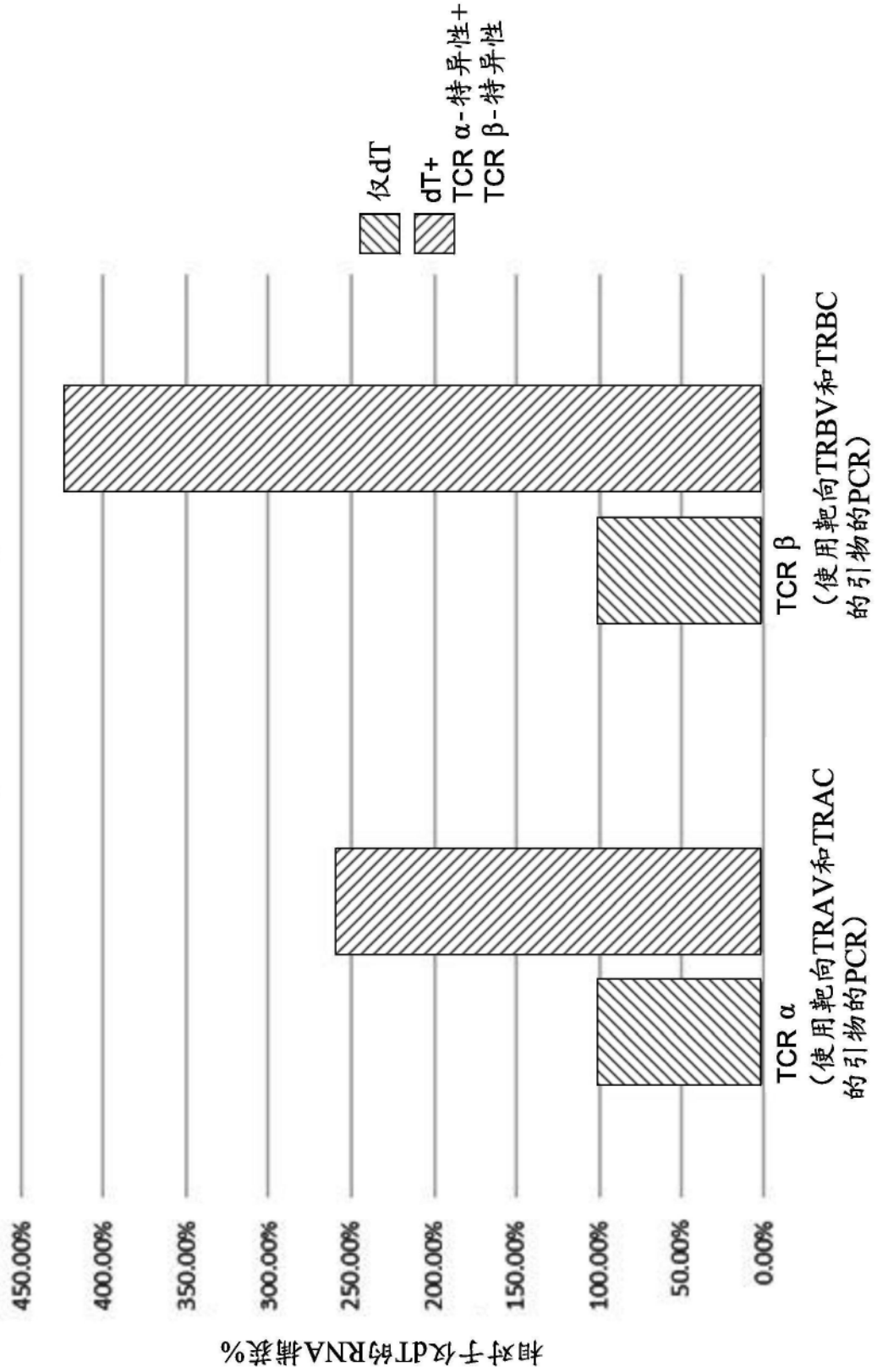


图7