



(12)发明专利申请

(10)申请公布号 CN 107810415 A

(43)申请公布日 2018.03.16

(21)申请号 201680034428.5

(74)专利代理机构 北京市柳沈律师事务所

(22)申请日 2016.05.09

11105

代理人 张文辉

(30)优先权数据

62/159,710 2015.05.11 US

(51)Int.Cl.

(85)PCT国际申请进入国家阶段日

2017.12.12

G01N 33/53(2006.01)

(86)PCT国际申请的申请数据

PCT/US2016/031524 2016.05.09

G01N 33/58(2006.01)

(87)PCT国际申请的公布数据

W02016/183029 EN 2016.11.17

G01N 33/68(2006.01)

(71)申请人 亿明达股份有限公司

地址 美国加利福尼亚州

C12Q 1/68(2018.01)

C40B 20/04(2006.01)

C40B 50/16(2006.01)

B01J 19/00(2006.01)

B01L 3/00(2006.01)

(72)发明人 M.赫 M.普雷维特 M.戈利恩斯基

M.W.克林杰 S.佩萨乔维奇

J.M.鲍特尔

权利要求书8页 说明书30页 附图12页

(54)发明名称

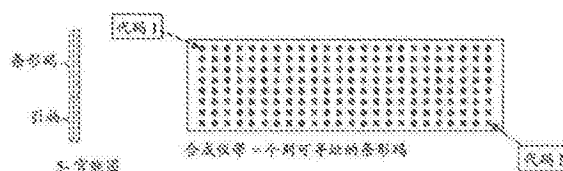
用于发现和分析治疗剂的平台

(57)摘要

表征候选试剂的方法,其包括以下步骤:(a)提供附着于核酸标签的候选试剂文库;(b)使所述文库与固体支持物接触以将所述候选试剂附着于所述固体支持物,由此形成候选试剂阵列;(c)使所述阵列与筛选剂接触,其中所述阵列中的一种或多种候选试剂与所述筛选剂反应;(d)检测阵列以测定阵列中的至少一种候选试剂与筛选剂反应;(e)对核酸标签测序以测定附着于阵列中候选试剂的标签序列;并且(f)基于附着于所述至少一种候选试剂的所述标签序列,鉴定所述阵列中与所述筛选剂反应的所述至少一种候选试剂。

步骤 1: 合成 500,000 个 23 聚核苷酸阵列 + 核苷酸 (-5' 端, 3' 端)

1. 合成阵列 = 500,000 个核苷酸 / 天 10 天内 500,000 个核苷酸
2. 条形码 = 16 核苷酸; "引物" = 10 核苷酸



1. 表征候选试剂的方法,其包括:

(a) 提供候选试剂文库,其中每个候选试剂附着于具有标签序列的核酸标签;

(b) 使所述候选试剂文库与固体支持物接触以将所述候选试剂附着于所述固体支持物,由此形成候选试剂阵列,所述候选试剂阵列包含所述固体支持物上的个别特征,其各自附着于来自所述文库的个别候选试剂;

(c) 使所述候选试剂阵列与筛选剂接触,其中所述阵列中的一种或多种候选试剂与所述筛选剂反应;

(d) 在所述阵列与所述筛选剂接触期间或之后检测所述阵列,从而测定所述阵列中的至少一种候选试剂与所述筛选剂反应;

(e) 对所述阵列上的所述核酸标签测序以测定附着于每种所述候选试剂的标签序列;并且

(f) 基于附着于所述至少一种候选试剂的所述标签序列,鉴定所述阵列中与所述筛选剂反应的所述至少一种候选试剂。

2. 权利要求1的方法,其中所述候选试剂选自下组:蛋白质、核酸、细胞和小分子。

3. 权利要求2的方法,其中所述蛋白质选自下组:抗体、酶、受体、激酶、磷酸酶、聚合酶、蛋白酶、酯酶、组蛋白修饰酶和核激素受体。

4. 权利要求1的方法,其中所述细胞选自下组:从多细胞生物体分离的遗传天然细胞、包含单细胞生物体的遗传天然细胞、基因工程细胞 (genetically engineered cell) 和培养细胞。

5. 权利要求1或4的方法,其中所述细胞是干细胞或免疫细胞。

6. 权利要求2的方法,其中所述小分子选自下组:候选酶抑制剂、候选抗生素、候选抗病毒剂、候选杀虫剂、候选激素、细胞信号传导的候选激活剂、细胞信号传导的候选抑制剂和候选酶活化剂。

7. 权利要求1的方法,其中步骤 (a) 包括组合合成候选试剂文库,其中如下追踪每个候选试剂上进行的组合合成的个别反应:通过将一个或多个核苷酸的独特签名(signature)添加到附着于每个所述候选试剂的核酸标签,由此提供候选试剂文库,其中每个候选试剂附着于独特的核酸标签。

8. 权利要求1的方法,其中所述固体支持物包含核酸引物,并且所述候选试剂经由所述核酸标签与所述核酸引物的杂交而附着于所述固体支持物。

9. 权利要求8的方法,其中所述核酸标签包含通用引物结合序列,所述核酸引物包含通用引物序列,并且所述候选试剂经由所述通用引物结合序列与所述通用引物序列的杂交附着于所述固体支持物。

10. 权利要求1的方法,其中所述筛选剂通过结合所述一种或多种候选试剂或通过阻断所述候选试剂与对所述候选试剂具有亲和力的分析物之间的结合而与所述一种或多种候选试剂反应。

11. 权利要求10的方法,其中检测所述阵列包括检测与所述一种或多种候选试剂结合的所述筛选剂。

12. 权利要求11的方法,其中所述筛选剂是发光的,并且所述检测包括检测所述阵列上的发光。

13. 权利要求1的方法,其中所述筛选剂通过化学修饰所述一种或多种候选试剂与所述一种或多种候选试剂反应。

14. 权利要求13的方法,其中检测所述阵列包括检测所述一个或多个经修饰的候选试剂。

15. 权利要求14的方法,其中所述一种或多种经修饰的候选试剂是发光的,并且所述检测包括检测所述阵列上的发光。

16. 权利要求1的方法,其中所述筛选剂通过产生分析物产物与所述一种或多种候选试剂反应。

17. 权利要求16的方法,其中检测所述阵列包括检测所述分析物产物。

18. 权利要求17的方法,其中所述分析物产物是发光的,并且所述检测包括检测所述阵列上的发光。

19. 权利要求1的方法,其中检测所述阵列包括对于所述阵列上的一个或多个所述个别特征在数个时间点获取信号。

20. 权利要求1的方法,其中所述信号包括光学信号。

21. 权利要求1的方法,其进一步包括扩增所述核酸标签以在所述个别特征处产生所述核酸标签的扩增子。

22. 权利要求21的方法,其中所述核酸标签的测序包括在所述个别特征处测序所述核酸标签的扩增子。

23. 权利要求22的方法,其中所述核酸标签的测序包括检测指示所述序列的光学信号。

24. 权利要求1的方法,其中所述固体支持物位于流动池内。

25. 用于产生蛋白质阵列的方法,其包括:

(a) 提供mRNA分子文库,其中所述文库中的个别mRNA分子包含靶序列和标签序列,

(b) 从所述文库中衍生第一子文库,所述第一子文库包含具有所述标签序列或其互补物的核酸,其中所述核酸附着于固体支持物上的个别特征,

(c) 从所述文库中衍生第二子文库,所述第二子文库包含具有所述靶序列和所述标签序列或其互补物的核酸;

(d) 使所述第二子文库与所述第一子文库接触,从而经由所述标签序列与其互补物的杂交将所述第二子文库的核酸附着于所述固体支持物;并且

(e) 翻译所述固体支持物上的所述靶序列以产生附着于所述个别特征的蛋白质阵列。

26. 权利要求25的方法,其进一步包括下述步骤:对所述固体支持物上的所述标签序列或其互补物测序,从而测定所述固体支持物上的个别特征处所述标签序列或其互补物的位置。

27. 权利要求25或26的方法,其中所述第一子文库的所述核酸进一步包含所述靶序列。

28. 权利要求27的方法,其还包括对所述固体支持物上的所述靶序列测序的步骤。

29. 筛选蛋白质的方法,其包括

(i) 产生根据权利要求25的蛋白质阵列;

(ii) 将所述蛋白质阵列与筛选剂接触,其中所述阵列中的一种或多种蛋白质与所述筛选剂反应;并且

(iii) 在与所述筛选剂接触期间或之后检测所述蛋白质阵列,由此测定所述阵列中的

至少一种蛋白质与所述筛选剂反应。

30. 权利要求29的方法,其进一步包括下述步骤:对所述固体支持物上的所述标签序列或其互补物测序,从而测定所述固体支持物上的个别特征处所述标签序列或其互补物的位置。

31. 权利要求30的方法,其进一步包括基于附着于所述至少一种蛋白质的所述标签序列来鉴定所述阵列中与所述筛选剂反应的所述至少一种蛋白质的步骤。

32. 权利要求29的方法,其中所述筛选剂通过结合所述一种或多种蛋白质或通过阻断所述一种或多种蛋白质与对所述一种或多种蛋白质具有亲和力的分析物之间的结合而与所述一种或多种蛋白质反应。

33. 权利要求32的方法,其中检测所述阵列包括检测与所述一种或多种蛋白质结合的所述筛选剂。

34. 权利要求33的方法,其中所述筛选剂是发光的,并且所述检测包括检测所述阵列上的发光。

35. 权利要求32的方法,其中所述筛选剂通过化学修饰所述一种或多种蛋白质与所述一种或多种蛋白质反应。

36. 权利要求35的方法,其中检测所述阵列包括检测所述一种或多种经修饰的蛋白质。

37. 权利要求36的方法,其中所述一种或多种经修饰的蛋白质是发光的,并且所述检测包括检测所述阵列上的发光。

38. 权利要求32的方法,其中所述筛选剂通过产生分析物产物与所述一种或多种蛋白质反应。

39. 权利要求38的方法,其中检测所述阵列包括检测所述分析物产物。

40. 权利要求39的方法,其中所述分析物产物是发光的,并且所述检测包括检测所述阵列上的发光。

41. 权利要求32的方法,其中所述阵列的检测包括对于所述阵列上的一个或多个个别特征在数个时间点获取信号。

42. 权利要求41的方法,其中所述信号包括光学信号。

43. 权利要求25的方法,其中所述mRNA分子文库包含相同基因的多个变体。

44. 权利要求43的方法,其中所述变体通过随机诱变产生。

45. 权利要求25或43的方法,其中所述mRNA分子文库源自重组生物体。

46. 权利要求25的方法,其中所述蛋白质选自下组:抗体、酶、受体、激酶、磷酸酶、聚合酶、蛋白酶、酯酶、组蛋白修饰酶和核激素受体。

47. 权利要求25的方法,其中通过方法衍生所述第一子文库,所述方法包括使所述文库的mRNA分子与所述固体支持物接触以将所述mRNA分子附着于所述固体支持物。

48. 权利要求47的方法,其中在所述固体支持物上扩增所述mRNA分子以产生所述标签序列的互补物。

49. 权利要求25的方法,其中所述固体支持物包含核酸引物,并且所述mRNA分子经由与所述核酸引物杂交而附着于所述固体支持物。

50. 权利要求49的方法,其中所述mRNA分子包含通用引物结合序列,所述核酸引物包含通用引物序列,并且所述mRNA分子经由所述通用引物结合序列与所述通用引物序列的杂交

而附着于所述固体支持物。

51. 权利要求25的方法,其中通过方法衍生所述第一子文库,所述方法包括逆转录所述个别mRNA分子或包含所述标签序列的所述个别mRNA分子的一部分。

52. 权利要求51的方法,其中在所述固体支持物上扩增逆转录的mRNA分子或其部分以产生所述标签序列的互补物。

53. 权利要求25的方法,其中在流体样品中提供所述mRNA分子文库,并且其中所述第一子文库和所述第二子文库源自所述流体样品的不同级份。

54. 权利要求25的方法,其中所述固体支持物位于流动池内。

55. 权利要求25的方法,其中所述蛋白质共价附着于所述mRNA分子。

56. 权利要求25的方法,其还包括选择性除去附着于所述阵列的一个或多个特征的mRNA分子或蛋白质的步骤。

57. 权利要求56的方法,其中所述选择性除去包括激光介导的键切割,所述键将所述mRNA分子或蛋白质附着于所述特征。

58. 权利要求25的方法,其中所述第一子文库包含具有所述标签序列的互补物的核酸,其中所述第二子文库包含具有所述靶序列和所述标签序列的RNA分子,

其中(d)包括使所述第二子文库与所述第一子文库接触,从而经由所述标签序列与其互补物的杂交将所述第二子文库的mRNA分子附着于所述固体支持物。

59. 权利要求25的方法,其中所述第一子文库包含具有所述标签序列的核酸,

其中所述第二子文库包含具有所述靶序列和所述标签序列的互补物的cDNA分子,

其中(d)包括使所述第二子文库与所述第一子文库接触,从而经由所述标签序列与其互补物的杂交将所述第二子文库的cDNA分子附着于所述固体支持物,并且

其中(e)包括逆转录所述cDNA分子以在所述固体支持物上产生mRNA分子并翻译所述固体支持物上的mRNA分子以产生附着于所述个别特征的蛋白质阵列。

60. 权利要求25的方法,其中用所述固体支持物上的核糖体翻译所述靶序列,并用嘌呤霉素处理所述核糖体以产生附着于所述个别特征的蛋白质阵列。

61. 权利要求25的方法,其中所述固体支持物包含至少 1×10^6 个特征。

62. 权利要求25的方法,其中所述固体支持物上的所述特征的平均间距小于10微米。

63. 权利要求25的方法,其中所述特征包括小于100平方微米的平均面积。

64. 阵列,其包含

(a) mRNA分子文库,其中所述文库中的个别mRNA分子包含靶序列和标签序列,

(b) 包含具有所述标签序列的互补物的核酸的固体支持物,

其中所述核酸附着于固体支持物上的个别特征,

其中所述个别mRNA分子的标签序列与所述固体支持物上的所述个别特征处相应的互补标签序列杂交,并且

其中通过翻译所述mRNA分子衍生的蛋白质附着于相应的mRNA分子。

65. 权利要求64的阵列,其中所述蛋白质经由核糖体附着于相应的mRNA分子。

66. 权利要求65的方法,其中所述蛋白质共价附着于所述核糖体。

67. 权利要求64的阵列,其中所述蛋白质是发光标记的。

68. 权利要求67的阵列,其中发光标记的筛选剂与蛋白质的结合位点特异性结合,从而

所述蛋白质是发光标记的。

69. 权利要求64的阵列,其中所述固体支持物包含至少 1×10^6 个所述特征。

70. 权利要求64的阵列,其中所述固体支持物上的所述特征的平均间距小于10微米。

71. 权利要求64的阵列,其中所述特征包括小于100平方微米的平均面积。

72. 权利要求64的阵列,其中所述mRNA分子文库包含相同基因的多个变体。

73. 权利要求64的阵列,其中所述蛋白质选自下组:抗体、酶、受体、激酶、磷酸酶、聚合酶、蛋白酶、酯酶、组蛋白修饰酶和核激素受体。

74. 权利要求64的阵列,其中所述固体支持物位于流动池内。

75. 筛选细胞的方法,其包括:

(a) 提供多个不同细胞,其中每个所述不同细胞包含具有标签序列的核酸标签;

(b) 使所述不同细胞的混合物与固体支持物接触以形成附着于所述固体支持物的细胞阵列;

(c) 筛选所述固体支持物上的细胞阵列的至少一种光学特征,其中筛选反应包括检测附着于所述固体支持物上的个别细胞;

(d) 对附着于所述固体支持物的核酸标签的所述标签序列测序;并且

(e) 基于所述候选细胞的所述光学特征和所述标签序列,将所述阵列中的至少一个细胞鉴定为候选细胞。

76. 权利要求75的方法,其中所述细胞选自下组:从多细胞生物体分离的天然细胞、包含单细胞生物体的天然细胞、基因工程细胞和培养细胞。

77. 权利要求76的方法,其中所述细胞是干细胞或免疫细胞。

78. 权利要求75的方法,其中步骤(a)包括将所述不同细胞分离到分开的容器中,并且将具有标签序列的核酸添加到每个所述容器,由此提供多个不同细胞,其中每个所述不同细胞位于分开的容器中并且包含具有独特标签序列的核酸标签。

79. 权利要求75的方法,其中具有标签序列的所述核酸附着于珠,并且所述珠结合所述多个不同细胞中的细胞。

80. 权利要求79的方法,其中所述珠包含具有特异性结合亲和力的抗体或所述细胞。

81. 权利要求75的方法,其中通过共价附着于质膜脂质或脂肪酸将具有标签序列的所述核酸附着于所述多个不同细胞中的细胞。

82. 权利要求75的方法,其中经由共价附着于质膜脂质中的蛋白质将具有标签序列的所述核酸附着于所述多个不同细胞中的细胞。

83. 权利要求75的方法,其中所述固体支持物包含核酸引物,并且所述细胞经由所述核酸标签与所述核酸引物的杂交而附着于所述固体支持物。

84. 权利要求83的方法,其中所述核酸标签包含通用引物结合序列,所述核酸引物包含通用引物序列,并且所述候选试剂经由所述通用引物结合序列与所述通用引物序列的杂交而附着于所述固体支持物。

85. 权利要求75的方法,其中所述阵列的检测包括对于所述阵列上的一个或多个所述个别特征在数个时间点获取信号。

86. 权利要求85的方法,其中所述信号包括光学信号。

87. 权利要求75的方法,其还包括扩增所述核酸标签以在所述固体支持物上产生所述

核酸标签的扩增子。

88. 权利要求87的方法,其中所述核酸标签的测序包括对所述核酸标签的扩增子测序。

89. 权利要求88的方法,其中所述核酸标签的测序包括检测指示所述序列的光学信号。

90. 权利要求75的方法,其中所述固体支持物位于流动池内。

91. 权利要求75的方法,其中所述细胞是活细胞。

92. 权利要求90的方法,其还包括从所述固体支持物除去所述至少一个候选细胞。

93. 权利要求92的方法,其还包括在从所述固体支持物除去之后培养所述至少一个候选细胞,由此复制所述至少一个细胞。

94. 权利要求75的方法,其中所述不同细胞包含经遗传修饰的细胞。

95. 权利要求94的方法,其中所述经遗传修饰的细胞包含选自下组的遗传修饰:编码非天然重组蛋白、编码突变体重组蛋白、具有天然存在的蛋白质的缺失、抑制天然存在的蛋白质的表达、增强天然存在的蛋白质的表达、产生非天然分析物和抑制天然分析物的产生。

96. 权利要求75的方法,其中所述细胞阵列的筛选包括用筛选剂处理所述细胞。

97. 权利要求96的方法,其中所述筛选剂与所述至少一个候选细胞结合。

98. 权利要求97的方法,其中所述筛选剂是发光的,并且其中所述筛选反应包括检测所述至少一个候选细胞的发光。

99. 权利要求96的方法,其中所述筛选剂修饰所述至少一个候选细胞。

100. 权利要求96的方法,其中所述筛选剂刺激所述至少一个候选细胞。

101. 权利要求99或100的方法,其中所述筛选剂增加或减少所述至少一个候选细胞的发光,并且其中所述筛选反应包括检测所述至少一个候选细胞的发光。

102. 权利要求75的方法,其进一步包括扩增具有标签序列的核酸标签,以在每个所述不同细胞上附着所述核酸标签的多个拷贝。

103. 权利要求75的方法,其中(d)包括在将所述核酸标签附着于所述固体支持物的条件下从所述固体支持物除去所述个别细胞,然后对附着于所述固体支持物的核酸标签的所述标签序列测序。

104. 权利要求75的方法,其中(d)包括复制所述核酸标签以产生附着于所述固体支持物的核酸标签拷贝,并且然后对附着于所述固体支持物的核酸标签拷贝的所述标签序列测序。

105. 权利要求75的方法,其中(d)包括复制所述核酸标签以产生附着于所述固体支持物的核酸标签拷贝,然后在将所述核酸标签拷贝附着于所述固体支持物的条件下从所述固体支持物除去所述个别细胞,并且然后对附着于固体支持物的核酸标签的所述标签序列测序。

106. 用于生产蛋白质阵列的方法,其包括:

(a) 提供附着于固体支持物的cDNA分子的文库;

(b) 扩增所述固体支持物上的所述cDNA分子以形成簇,其中每个簇包含来自所述文库的特定cDNA分子的多个拷贝;

(c) 在所述簇处转录所述多个拷贝以产生附着于每个所述簇的多个mRNA分子;并且

(d) 翻译所述簇处的所述mRNA分子以产生附着于每个所述簇的多个蛋白质。

107. 权利要求106的方法,其进一步包括对所述固体支持物上的每个所述mRNA分子的

至少一部分测序的步骤。

108. 筛选蛋白质的方法, 其包括

(i) 产生根据权利要求106的蛋白质阵列;

(ii) 使所述固体支持物上的所述蛋白质与筛选剂接触, 其中一种或多种所述蛋白质与所述筛选剂反应; 并且

(iii) 在与所述筛选剂接触期间或之后检测所述固体支持物上的所述蛋白质, 从而测定与所述筛选剂反应的至少一种蛋白质。

109. 权利要求108的方法, 其还包括对所述固体支持物上的每种所述mRNA分子的至少一部分测序的步骤。

110. 权利要求109的方法, 其进一步包括下述步骤: 基于附着于所述至少一种蛋白质的所述mRNA分子的至少一部分的序列来鉴定所述阵列中与所述筛选剂反应的至少一种蛋白质。

111. 权利要求108的方法, 其中所述筛选剂通过与所述一种或多种蛋白质结合或通过阻断所述一种或多种蛋白质与对所述一种或多种蛋白质具有亲和力的分析物之间的结合而与所述一种或多种蛋白质反应。

112. 权利要求111的方法, 其中所述检测包括检测与所述一种或多种蛋白质结合的所述筛选剂。

113. 权利要求112的方法, 其中所述筛选剂是发光的, 并且所述检测包括检测发光。

114. 权利要求108的方法, 其中所述筛选剂通过化学修饰所述一种或多种蛋白质与所述一种或多种蛋白质反应。

115. 权利要求114的方法, 其中所述检测包括检测所述一种或多种经修饰的蛋白质。

116. 权利要求115的方法, 其中所述一种或多种经修饰的蛋白质是发光的, 并且所述检测包括检测发光。

117. 权利要求108的方法, 其中所述筛选剂通过产生分析物产物与所述一种或多种蛋白质反应。

118. 权利要求117的方法, 其中所述检测包括检测所述分析物产物。

119. 权利要求118的方法, 其中所述分析物产物是发光的, 并且所述检测包括检测发光。

120. 权利要求108的方法, 其中所述阵列的检测包括对所述一个或多个所述个别簇在几个时间点时获取信号。

121. 权利要求120的方法, 其中所述信号包括光学信号。

122. 权利要求106的方法, 其中所述mRNA分子文库包含相同基因的多个变体。

123. 权利要求122的方法, 其中通过随机诱变产生所述变体。

124. 权利要求106或122的方法, 其中所述cDNA分子文库源自重组生物体。

125. 权利要求106的方法, 其中所述蛋白质选自下组: 抗体、酶、受体、激酶、磷酸酶、聚合酶、蛋白酶、酯酶、组蛋白修饰酶和核激素受体。

126. 权利要求106的方法, 其中所述固体支持物包含核酸引物, 并且所述cDNA分子经由与所述核酸引物杂交而附着于所述固体支持物。

127. 权利要求126的方法, 其中所述cDNA分子包含通用引物结合序列, 所述核酸引物包

含通用引物序列,并且所述cDNA分子经由所述通用引物结合序列与所述通用引物序列杂交而附着于所述固体支持物。

128. 权利要求106的方法,其中所述固体支持物位于流动池内。

129. 权利要求106的方法,其还包括选择性除去附着于一个或多个簇的蛋白质的步骤。

130. 权利要求129的方法,其中所述选择性除去包括激光介导的键切割,所述键将所述蛋白质分子附着于所述固体支持物。

131. 权利要求106的方法,其中用所述固体支持物上的核糖体翻译所述mRNA分子,并且用嘌呤霉素处理所述核糖体以将所述蛋白质附着于所述mRNA分子。

132. 权利要求106的方法,其中所述固体支持物包含至少 1×10^6 个所述簇。

133. 权利要求106的方法,其中所述固体支持物上的所述簇的平均间距小于10微米。

134. 权利要求106的方法,其中所述簇包含小于100平方微米的平均面积。

135. 阵列,其包含

(a) 固体支持物;

(b) 附着于所述固体支持物的不同cDNA分子的文库,其中每个不同的cDNA分子附着于所述固体支持物上的个别特征,并且其中每个特征包括特定cDNA分子的多个拷贝;

(c) 附着于所述cDNA分子的mRNA分子,其中每个所述cDNA分子与相应的附着的mRNA分子互补;和

(d) 附着于所述mRNA分子的蛋白质分子,其中每个所述蛋白质分子由相应的附着的mRNA分子编码。

136. 权利要求135的阵列,其中所述蛋白质经由核糖体附着于相应的mRNA分子。

137. 权利要求136的方法,其中所述蛋白质共价附着于所述核糖体。

138. 权利要求135的阵列,其中所述mRNA分子经由RNA聚合酶附着于相应的cDNA分子。

139. 权利要求138的方法,其中所述mRNA分子共价附着于所述cDNA分子。

140. 权利要求135的阵列,其中所述蛋白质是发光标记的。

141. 权利要求140的阵列,其中发光标记的筛选剂与蛋白质的结合位点特异性结合,从而所述蛋白质是发光标记的。

142. 权利要求135的阵列,其中所述固体支持物包含至少 1×10^6 个所述特征。

143. 权利要求135的阵列,其中所述固体支持物上的所述特征的平均间距小于10微米。

144. 权利要求135的阵列,其中所述特征包括小于100平方微米的平均面积。

145. 权利要求135的阵列,其中所述cDNA分子文库包含相同基因的多个变体。

146. 权利要求135的阵列,其中所述蛋白质选自下组:抗体、酶、受体、激酶、磷酸酶、聚合酶、蛋白酶、酯酶、组蛋白修饰酶和核激素受体。

用于发现和分析治疗剂的平台

[0001] 发明背景

[0002] 本公开一般涉及具有医学、农业和工业用途的试剂(例如小分子、蛋白质或细胞)的发现,并且更具体地涉及用于筛选用于此类用途的候选试剂的平台。

[0003] 目前,高通量筛选药物发现利用多个步骤和平台来对小分子的大文库筛选显示针对特定靶标的功效和效力的“前导物(lead)”或“命中(hit)”,以及针对一组分子靶标的有利毒性概况。

[0004] 在该过程的第一步中,针对预定靶标以高通量的方式筛选小分子文库。这些文库通常是数十万种化合物的数目级。目前的高通量方法可以是自动化的。因此,随着机器人技术和自动化技术的发展,高成本的仪器可以在低微升范围内使用大量的输入试剂来筛选100,000个分子/天。测定法可以是同质的或异质的,前者相对简单且价格便宜,而后者更灵敏但更为复杂、耗时且昂贵。

[0005] 尽管最初的小分子筛选测定法有效鉴定“命中”或“前导物”,但这些测定法通常只是第一步。这些前导物或命中通常在遗传毒性、药理学毒性/细胞毒性和生物体细胞毒性方面进行进一步筛选,使用针对明确定义的靶标的序型分析(profiling)测定法进行以消除可以具有不利临床效应的候选物。这些测定法中的每种使用采用独立工作流的多种平台和成套测定法独立进行,所述工作流进行策划和数据挖掘,以便对命中或前导物的效用产生共识。

[0006] 蛋白质进化方法构成了另一种类型的筛选。这些方法涵盖了大量的排列。虽然这些方法的通量、速度和低成本已经趋向于使其成为快速蛋白质进化的主要选择,但是方法通常是复杂的并且带来挑战。例如,基于乳液的筛选提出合并/混合液滴(用于多重筛选)的困难、破坏乳剂(回收感兴趣组分)的困难、源自由于各种液滴尺寸引起的浓度变化的复杂化、以及液滴的交叉污染。

[0007] 与基于小分子或基于蛋白质的疗法相比,基于细胞的疗法具有潜在有利的治疗能力。在主要优点间:细胞可以感测外部信号,移动到身体内的特定位置,并且整合多种刺激,并以复杂的行为(如释放特定的效应分子)作出反应。完全实现基于细胞的治疗剂的潜力将受益于精确工程化改造治疗性细胞,使其“行为”可以在时间和空间上进行控制。

[0008] 细胞工程中的当前工作流通常包括以下步骤:(i)设计负责感测、整合和响应刺激物的细胞内信号传导回路;(ii)复杂的遗传工程,以整合负责介导细胞中那些功能的基因;和(iii)筛选方法以鉴定文库中所有细胞中携带基因互补物的那些基因,其更好地进行期望的功能。由于期望的细胞行为的复杂性,目前的筛选方法并不理想。例如:基于荧光激活细胞分选(FACS)的方法是高通量,但仅查看细胞行为的快照(snapshot)。另一方面,荧光显微术可以详细追踪细胞行为的动力学,但是为低通量。

[0009] 因此,需要对小分子、蛋白质、细胞和其它试剂筛选有益特性的平台和方法。本公开解决了这种需要并提供了其它优点。

[0010] 发明概述

[0011] 本公开提供了表征候选试剂的方法。方法可以包括以下步骤:(a)提供候选试剂文

库,其中每个候选试剂附着于具有标签序列的核酸标签;(b)使所述候选试剂文库与固体支持物接触以将所述候选试剂附着于所述固体支持物,由此形成候选试剂阵列,所述候选试剂阵列包含所述固体支持物上的个别特征,其各自附着于来自所述文库的个别候选试剂;(c)使所述候选试剂阵列与筛选剂接触,其中所述阵列中的一种或多种候选试剂与所述筛选剂反应;(d)在所述阵列与所述筛选剂接触期间或之后检测所述阵列,从而测定所述阵列中的至少一种候选试剂与所述筛选剂反应;(e)对所述阵列上的所述核酸标签测序以测定附着于每种所述候选试剂的标签序列;并且(f)基于附着于所述至少一种候选试剂的所述标签序列,鉴定所述阵列中与所述筛选剂反应的所述至少一种候选试剂。

[0012] 本公开还提供了用于生产蛋白质阵列的方法。方法可以包括以下步骤:(a)提供附着于固体支持物的cDNA分子的文库;(b)扩增所述固体支持物上的所述cDNA分子以形成簇,其中每个簇包含来自所述文库的特定cDNA分子的多个拷贝;(c)在所述簇处转录所述多个拷贝以产生附着于每个所述簇的多个mRNA分子;并且(d)翻译所述簇处的所述mRNA分子以产生附着于每个所述簇的多个蛋白质。

[0013] 本公开还提供了用于生产蛋白质阵列的方法,所述方法包括以下步骤:(a)提供mRNA分子文库,其中所述文库中的个别mRNA分子包含靶序列和标签序列,(b)从所述文库中衍生第一子文库,所述第一子文库包含具有所述标签序列或其互补物的核酸,其中所述核酸附着于固体支持物上的个别特征,(c)从所述文库中衍生第二子文库,所述第二子文库包含具有所述靶序列和所述标签序列或其互补物的核酸;(d)使所述第二子文库与所述第一子文库接触,从而经由所述标签序列及其互补物的杂交将所述第二子文库的核酸附着于所述固体支持物;并且(e)翻译所述固体支持物上的所述靶序列以产生附着于所述个别特征的蛋白质阵列。

[0014] 本公开还提供了筛选细胞的方法。方法可以包括以下步骤:(a)提供多个不同细胞,其中每个所述不同细胞包含具有标签序列的核酸标签;(b)使所述不同细胞的混合物与固体支持物接触以形成附着于所述固体支持物的细胞阵列;(c)筛选所述固体支持物上的细胞阵列的至少一种光学特征,其中筛选反应包括检测附着于所述固体支持物上的个别细胞;(d)对附着于所述固体支持物的核酸标签的所述标签序列测序;并且(e)基于所述候选细胞的所述光学特征和所述标签序列,将所述阵列中的至少一个细胞鉴定为候选细胞。

[0015] 本公开提供了阵列,其包含(a)固体支持物;(b)附着于所述固体支持物上的不同cDNA分子的文库,其中每个不同的cDNA分子附着于所述固体支持物上的个别特征,并且其中每个特征包括特定cDNA分子的多个拷贝;(c)附着于所述cDNA分子的mRNA分子,其中每个所述cDNA分子与相应的附着的mRNA分子互补;和(d)附着于所述mRNA分子的蛋白质分子,其中每个所述蛋白质分子由相应的附着的mRNA分子编码。

[0016] 本公开还提供了阵列,该阵列包括以下阵列,其包含(a)mRNA分子文库,其中所述文库中的个别mRNA分子包含靶序列和标签序列,(b)包含具有所述标签序列的互补物的核酸的固体支持物,其中所述核酸附着于固体支持物上的个别特征,其中所述个别mRNA分子的标签序列与所述固体支持物上的所述个别特征处相应的互补标签序列杂交,并且其中通过翻译所述mRNA分子衍生的蛋白质附着于相应的mRNA分子上。

[0017] 附图简述

[0018] 图1A至图1F显示了候选试剂的高通量筛选过程中的步骤。

[0019] 图2显示了用于高通量合成和筛选候选蛋白质试剂的方法。

[0020] 图3A和图3B显示了候选蛋白质试剂的高通量标签化(tagging)、合成和筛选的过程。

[0021] 图4显示了筛选剂暴露于附着于流动池的细胞和从各种细胞对筛选剂的应答预期的荧光对时间的图。

[0022] 图5显示了分选到个别微孔中,然后用微孔特异性标签标签化的细胞。

[0023] 图6显示了使用编码珠解码固体支持物上细胞位置的。

[0024] 图7显示了使用核酸标签解码固体支持物上的细胞位置。

[0025] 图8显示了核酸标签从细胞转移到固体支持物上的位点和固体支持物上核酸标签的测序。

[0026] 图9显示了通过将一对核酸附着于细胞膜中的脂肪酸来标签化的细胞。

[0027] 图10显示了流动池表面上核酸标签化的细胞的捕捉。

[0028] 图11显示了经由核酸标签的限制性内切核酸酶切割从流动池表面分离细胞。

[0029] 图12显示了基于获得细胞的容器中解码标签的位置鉴定流动池表面上的细胞。

[0030] 发明详述

[0031] 本公开提供用于高通量筛选候选试剂,例如用于治疗用途的候选物的装置和方法。本文阐述的具体实施方案有益地使用核酸测序技术和装置。使用如本文阐述的测序技术和装置的优点是空间排列大量不同的候选试剂的能力,使得可以区分与每种候选试剂相关的核酸并且使得可以在固体支持物上个别检测每种候选试剂对特定刺激物的应答。

[0032] 在具体的实施方案中,在文库中提供了多种不同的候选试剂,其中每种候选试剂与独特的核酸标签附着(或以其它方式相关联)。候选试剂可以是多种物质中的任一种,包括但不限于核酸、蛋白质、细胞和小分子。这些标签化的候选试剂可以附着于固体支持物,使得个别成员形成空间上分开的候选试剂阵列。然后,可以通过暴露于筛选剂(或其它刺激物)筛选候选试剂阵列,并且可以在固体支持物上空间分辨的位置处检测候选试剂的反应。一个或多个“命中”的位置可以基于阵列上的预期的、期望的或独特的信号的空间分辨检测来鉴定。还可以在阵列上进行测序反应(在筛选步骤之前或之后),以便就其在阵列中的位置而言鉴定每个标签。可以通过将“命中”的位置与该位置处的标签的身份相关联来鉴定候选试剂。

[0033] 本文提出的方法或装置中使用的一些候选试剂是基于核酸的,包括例如蛋白质和细胞。在使用蛋白质的几个实施方案中,可以测定编码蛋白质的DNA或RNA分子的序列以区分一种蛋白质与另一种蛋白质。此外,可以基于已知的遗传密码从该RNA序列推断个别蛋白质的氨基酸序列。然而,在一些实施方案中,蛋白质和编码蛋白质的核酸都不需要测序。相反,每种蛋白质可以与标签一起附着(或以其它方式相关联),所述标签已经与蛋白质序列先验相关。因此,对标签测序可以足以区分一种蛋白质与另一种。类似地,基于细胞的候选试剂含有可以测序以鉴定群体中的个别细胞的核酸。此外,可以评估核酸序列以测定个别细胞的有用特征。此外,使用或先验分配的标签可以允许区分细胞特征,而不需要对细胞的其它内容物测序。

[0034] 虽然本文中就筛选候选试剂的治疗功能而言例示了本公开的方法和装置,但是应当理解,可以筛选其它功能或结构特征。例如,所述方法和装置可用于筛选毒性、农业用途

(例如杀虫剂、生长因子、激素等)、工业用途(例如催化剂、染料、塑料等)、营养(例如香料、防腐剂等)、环境清理等。通常,这些方法和装置可以用于筛选生物学或非生物学功能。

[0035] 本文阐述的方法和装置提供了在高通量筛选期间进行定量测量的能力的优点。例如,商品化测序平台(例如由Illumina, Inc. (San Diego, CA)商业化的测序平台)包括具有用于量化荧光信号的相对宽的动态范围的精密光学器件。另一个优点是跟踪候选试剂阵列中多个位置处的筛选反应的时间动力学的能力。相比之下,在许多传统筛选中使用的流体分类方法仅提供通过的候选试剂的快照,从而将测量限制于单一时间点。本文阐述的方法和装置提供了与流体分选技术相比同等或在一些情况下更好的高通量,但是具有筛选结果的基于时间的测量的额外益处。本文阐述的装置和方法的另一个优点是检测筛选结果并且以允许标签与筛选结果的空间相关以从筛选鉴定“命中”的方式来检测标签的能力。

[0036] 除非另有规定,本文使用的术语将被理解为具有相关技术中的普通含义。下面列出本文使用的几个术语及其含义。

[0037] 如本文所用,当提及核酸使用时,术语“扩增子”是指复制核酸的产物,其中产物具有与核酸的核苷酸的至少一部分相同或互补的核苷酸序列。可以通过使用核酸或其扩增子作为模板的多种扩增方法中的任一种来产生扩增子,所述扩增方法包括例如聚合酶延伸、聚合酶链式反应(PCR)、滚环扩增子(RCA)、多重置换扩增(MDA)、连接延伸或连接链式反应。扩增子可以是具有特定核苷酸序列的单拷贝(例如PCR产物)或核苷酸序列的多个拷贝(例如RCA的多联体产物)。靶核酸的第一扩增子通常是互补拷贝。随后的扩增子是在产生第一扩增子后自靶核酸或自扩增子创建的拷贝。随后的扩增子可以具有与靶核酸基本上互补或与靶核酸基本上相同的序列。

[0038] 如本文所用,术语“阵列”是指可以根据相对位置相互区分的特征或位点的群体。在阵列的不同位点处的不同分子或其它物质可以根据阵列中位点的位置来彼此区分。阵列的个别位点可以包括特定类型的一种或多种分子(或其它物质)。例如,位点可以包括具有特定序列的单一核酸分子,或者位点可以包括具有相同序列(和/或其互补物)的几个核酸分子。阵列的位置可以是位于同一基底上的不同特征。示例性的特征包括但不限于基底中的孔、基底中或基底上的珠(或其它颗粒)、从基底的凸起、基底上的脊或基底中的通道。阵列的位点可以是各自承载不同分子(或其它物质)的不同基底。可以根据与基底相关联的固体支持物上的基底位置或者根据液体或凝胶中的基底位置来鉴定附着于不同基底的不同物质。分开的基底位于固体支持物上的示例性阵列包括但不限于在孔中具有珠的那些阵列。

[0039] 如本文所用,术语“附着/附接”是指两个物体彼此连接、紧固、粘附、连接或结合的状态。例如,分析物如核酸可通过共价键或非共价键附着于诸如凝胶或固体支持物的材料。共价键以原子之间共享电子对为特征。非共价键是不涉及共享电子对的化学键,并且可以包括例如氢键、离子键、范德华力、亲水性相互作用和疏水性相互作用。在一些实施方案中,附着可以经由捕捉剂发生。捕捉剂可以包括例如抗体、受体、核酸、配体、凝集素、碳水化合物、亲合素、生物素或其类似物。

[0040] 如本文所用,术语“候选试剂”旨在表示怀疑具有特定结构或功能的物质。示例性的物质包括但不限于分子细胞和亚细胞组分。分子可以任选地为生物活性分子,如蛋白质、氨基酸、核酸(例如DNA或RNA)、核苷酸、多糖、糖类、代谢物、维生素、酶辅因子等。其它候选

试剂包括大环、环状肽、融合分子(例如核酸-蛋白质融合体)或展示的构建体(例如噬菌体上的肽)。可以怀疑候选试剂具有的示例性功能包括但不限于另一种试剂的活化另一种试剂的抑制、另一种试剂的化学修饰、另一种试剂的降解、另一种试剂的合成,其中另一种试剂可以任选地是上文作为候选试剂示例的任何一种或多种物质。候选试剂的结构可以是用于上述物质或本领域已知的其它物质的任何已知的或可疑的结构。

[0041] 如本文所用,术语“不同”在提及核酸使用时意指核酸具有彼此不同的核苷酸序列。两种或更多种核酸可以具有沿其整个长度不同的核苷酸序列。或者,两种或更多种核酸可以具有沿着其长度的实质部分不同的核苷酸序列。例如,两种或更多种核酸可以具有两种或更多种分子不同的靶核苷酸序列部分,而还具有两种或更多种分子上相同的通用序列部分。该术语可以类似地应用于基于氨基酸序列差异可区分为彼此不同的蛋白质。

[0042] 如本文所用,术语“每个”当提及物质集合使用时旨在鉴定集合中的个别物质,但不一定是指集合中的每个物质。若明确的公开或上下文清楚相反规定,则可以存在例外。

[0043] 如本文所用,术语“延伸”在提及核酸使用时意图指对核酸添加至少一个核苷酸或寡核苷酸。在具体的实施方案中,例如可以经由聚合酶催化(例如DNA聚合酶、RNA聚合酶或逆转录酶)将一个或多个核苷酸添加到核酸的3'末端。可以使用化学或酶促方法将一个或多个核苷酸添加到核酸的3'或5'末端。例如,可以经由化学或酶促(例如连接酶催化)方法将一个或多个寡核苷酸添加到核酸的3'或5'端。核酸可以以模板引导的方式延伸,由此延伸产物与模板核酸互补,所述模板核酸与延伸的核酸杂交。

[0044] 如本文所用,术语“特征”是指阵列中特定分子或细胞种类的位置。特征可以仅含有单一分子(或细胞),或者它可以含有同一种类的几个分子(或细胞)的群体。在一些实施方案中,在附着分子或细胞之前,特征存在于固体支持物上。在其它实施方案中,该特征通过将分子或细胞附着于固体支持物而创建。阵列的特征通常是离散的。离散特征可以是连续的或者它们可以彼此之间有间隔。特征的大小和/或特征之间的间隔可以变化,使得阵列可以是高密度、中等密度或较低密度。高密度阵列的特征为具有以小于约15 μm 分开的位点。中等密度阵列具有以约15至30 μm 分开的位点,而低密度阵列具有以大于30 μm 分开的位点。本文中有用的阵列可以具有例如以小于100 μm 、50 μm 、10 μm 、5 μm 、1 μm 或0.5 μm 分开的位点。本公开的装置或方法可用于足以区分上述密度或密度范围的位点的分辨率检测阵列。

[0045] 如本文所用,术语“流动池”旨在表示具有可进行反应的室、用于将试剂递送至室的入口以及用于从室除去试剂的出口的容器。在一些实施方案中,室配置用于检测室中发生的反应。例如,室可以包括一个或多个透明表面,允许室中阵列、光学标记的分子等的光学检测。示例性的流动池包括但不限于用于核酸测序装置中的那些,例如由Illumina, Inc. (San Diego, CA)商业化的Genome Analyzer[®]、MiSeq[®]、NextSeq[®]或HiSeq[®]平台;或由Life Technologies (Carlsbad, CA)商业化的SOLiD[™]或Ion Torrent[™]测序平台的流动池。示例性的流动池及其制造和使用方法也记载于例如WO 2014/142841 A1;美国专利申请公开No. 2010/0111768 A1和美国专利No. 8,951,781,每篇通过引用并入本文。

[0046] 如本文所用,术语“文库”旨在表示包括几个不同物质的集合。集合中的物质可以在结构和/或功能上有所不同。例如,集合可以包括具有不同核苷酸序列的核酸,或者集合可以包括具有不同一级结构(即氨基酸序列)、二级结构、三级结构或四级结构的蛋白质。但是,应当理解的是,文库中的物质可以有一些冗余。例如,特定核酸或蛋白质的多个拷贝可

以存在于文库中,该文库仍然包含大量不同的核酸或蛋白质。可以在文库中的物质的示例性类型包括本文中就候选试剂或筛选剂而言阐述的物质。

[0047] 如本文所用,术语“发光”是指发出冷体辐射。该术语旨在区分作为由于热而从材料发射的辐射的白炽。通常,当能量源将原子的电子从其最低能量基态置换为较高能量激发态;然后电子以辐射的形式返回能量,因此它可以回落到基态时会产生发光。一种特别有用的发光物质类型是当激发辐射提供能量时发出冷体辐射的物质。此类物质称为“荧光”或“光致发光”。可以认为荧光或光致发光为在某个波长处由物质的辐射发射,其是在另一种波长处照射物质的结果。

[0048] 如本文所用,术语“核酸”和“核苷酸”旨在与其在本领域中的使用一致并且包括天然存在的种类或其功能类似物。核酸的特别有用的功能类似物能够以序列特异性方式与核酸杂交或能够用作特定核苷酸序列复制的模板。天然存在的核酸通常具有含磷酸二酯键的主链。类似物结构可以具有交替的主链连接,包括本领域已知的多种连接之任一种。天然存在的核酸通常具有脱氧核糖(例如在脱氧核糖核酸(DNA)中发现)或核糖(例如在核糖核酸(RNA)中发现)。核酸可以含有具有本领域已知的这些糖部分的多种类似物中的任一种的核苷酸。核酸可以包括天然或非天然核苷酸。就这点而言,天然脱氧核糖核酸可以具有一种或多种选自下组的碱基:腺嘌呤、胸腺嘧啶、胞嘧啶或鸟嘌呤,并且核糖核酸可以具有一种或多种选自下组的碱基:尿嘧啶、腺嘌呤、胞嘧啶或鸟嘌呤。可以包括在核酸或核苷酸中的有用的非天然碱基是本领域已知的。

[0049] 如本文所用,术语“间距”在提及阵列的特征使用时旨在表示相邻特征的中心到中心间隔。特征模式可以按平均间距来表征。该模式可以是有序的,使得平均间距周围的变化系数较小,或者该模式可以是随机的,在这种情况下,变化系数可以相对较大。在任一种情况下,平均间距可以是例如至少约10nm、0.1 μ m、0.5 μ m、1 μ m、5 μ m、10 μ m、100 μ m或更大。或者/另外,平均间距可以是例如至多约100 μ m、10 μ m、5 μ m、1 μ m、0.5 μ m或0.1 μ m或更小。当然,特征的特定模式的平均间距可以在选自上述范围的较低值之一和上限值之一之间。

[0050] 如本文所用,术语“蛋白质”或“氨基酸”旨在与其在本领域中的使用一致并且包括天然存在的种类或其功能类似物。天然存在的蛋白质通常具有包含肽键的主链。类似物结构可以具有交替主链连接,包括本领域中已知的多种主链连接之一。天然存在的蛋白质通常具有选自以下的天然氨基酸:精氨酸、组氨酸、赖氨酸、天冬氨酸、谷氨酸、丝氨酸、苏氨酸、天冬酰胺、谷氨酰胺、半胱氨酸、甘氨酸、脯氨酸、丙氨酸、缬氨酸、异亮氨酸、亮氨酸、甲硫氨酸、苯丙氨酸、酪氨酸、色氨酸及其天然存在的修饰。一些天然存在的修饰包括磷酸化(例如丝氨酸、苏氨酸、组氨酸、天冬氨酸和谷氨酸的磷酸化)、异戊二烯化(prenylation, isoprenylation)、酰化、烷基化、糖基化、生物素化、泛素化等。蛋白质可以包括其上具有非天然部分的天然氨基酸。蛋白质可以包括天然或非天然氨基酸。

[0051] 如本文所用,当提及第一试剂和第二试剂使用时,术语“反应”意指修饰试剂之一或两者的化学结构、在两种试剂之间创建一个或多个共价键、允许试剂之一催化对另一种试剂的化学结构的修饰、或者特异性结合两种试剂(例如通过非共价相互作用)的行为。示例性的反应包括但不限于化学反应如还原、氧化、加成、消除、重排、酯化、酰胺化、醚化、环化或取代;第一试剂以特定的亲和力与第二试剂结合的结合相互作用;两种或更多种试剂彼此分离的解离反应;荧光;发光;化学发光;和生物反应,如核酸复制、核酸扩增、核酸杂

交、核酸连接、磷酸化、酶促催化、受体结合或配体结合。

[0052] 如本文所用,术语“重组”旨在指非天然存在的遗传构建体。一个实例是组合来自超过一种起源的遗传物质的产物。示例性的重组分子包括但不限于DNA、RNA和蛋白质。可以衍生重组分子的起源包括例如来自不同生物体的相似遗传元件、来自相同生物体的不同遗传元件、来自不同生物体的不同遗传元件、合成遗传元件或合成和天然遗传元件的组合。

[0053] 如本文所用,术语“筛选剂”旨在表示具有与第二候选试剂相比对第一候选试剂具有选择性的结构或功能的物质。在许多实施方案中,筛选剂的结构或功能在其用于本公开的方法或组合物之前是已知的。示例性的物质包括但不限于分子、细胞和亚细胞组分。分子可以任选地为生物活性分子,如蛋白质、氨基酸、核酸(例如DNA或RNA)、核苷酸、多糖、糖类、代谢物、维生素、酶辅因子等。示例性的选择性功能包括但不限于另一种试剂的活化、另一种试剂的抑制、另一种试剂的化学修饰、另一种物质的降解、另一种物质的合成;其中所述另一种试剂可以是以上例举为候选试剂的任何一种或多种物质。筛选剂的结构可以是上述物质或本领域已知的其它物质的任何已知的或可疑的结构。

[0054] 如本文所用,与第二事物相比“选择性”操纵第一事物(或第一件事物的“选择性”操纵)意指与对第二事物的影响相比,操纵对第一事物具有更大的影响。操纵不需要对第二事物有任何影响。操纵对第一件事物具有的影响可以比对第二事物的影响大至少1%、10%、50%、90%或99%。操纵对第一件事物具有的影响可以比对第二事物的影响高至少2倍、5倍、10倍、100倍、 1×10^3 倍、 1×10^4 倍或 1×10^6 倍。操作可以包括例如修饰、接触、处理、改变、切割(例如化学键)、光化学切割(例如化学键)、形成(例如化学键)、光化学形成(例如化学键)、共价修饰、非共价修饰、破坏、光消融、除去、合成、聚合、光聚合、扩增(例如核酸)、复制(例如核酸)、延伸(例如核酸),连接(例如核酸)或本文阐述或本领域中以其它方式已知的其它操作。

[0055] 如本文所用,术语“小分子”旨在表示具有小于约1000道尔顿的分子量的化合物。在具体的实施方案中,小分子是非聚合的。然而,在其它实施方案中,小分子可以是二聚体或三聚体。还应当理解,小分子可以是能够掺入聚合物中的单体。特别有用的小分子是有机化合物。有用的小分子可以具有小于900、800、600、400、200或100道尔顿的分子量。

[0056] 如本文所用,术语“固体支持物”是指不溶于含水液体的刚性基底。基底可以是无孔的或多孔的。基底可以任选地能够吸收液体(例如由于孔隙率),但是通常具有足够的刚性,使得基底在吸收液体时基本不膨胀,并且当通过干燥除去液体时基本上不收缩。无孔固体支持物通常对液体或气体是不可渗透的。示例性固体支持物包括但不限于玻璃和改性或官能化玻璃、塑料(例如丙烯酸类树脂(acrylics)、聚苯乙烯和苯乙烯和其它材料的共聚物、聚丙烯、聚乙烯、聚丁烯、聚氨酯、特氟隆™、环状烯烃、聚酰亚胺等)、尼龙、陶瓷、树脂、Zeonor、二氧化硅或基于二氧化硅的材料,包括硅和改性硅、碳、金属、无机玻璃、光纤束和聚合物。一些实施方案的特别有用的固体支持物是流动池的组分或位于流动池装置内。示例性的流动池在本文中进一步详细阐述。

[0057] 如本文所用,术语“子文库”旨在表示具有包括来自文库的物质或来自文库的物质的拷贝的呈现(representation)的集合。通过与文库中的呈现相比,子文库中的呈现可以是完整的或部分的。可以通过从文库中分开至少一些物质或者通过制备文库中至少部分物质的拷贝来衍生子文库。

[0058] 如本文所用,术语“标签序列”旨在表示核酸中的一系列核苷酸,其可用于鉴定或表征与核酸附着(或缔合)的试剂。标签序列可以是天然存在的序列或在获得核酸的生物体中不天然存在的序列。在具体的实施方案中,与生物样品一起使用的一个或多个标签序列不天然存在于生物样品的基因组、转录物组或其它核酸中。例如,标签序列可以与特定生物样品中的核酸序列具有小于80%、70%、60%、50%或40%的序列同一性。

[0059] 如本文所用,术语“通用序列”是指即使分子也具有彼此不同的序列区域,与两个或更多个核酸分子共同的一系列核苷酸。存在于分子集合的不同成员中的通用序列可以允许使用与通用序列互补的通用捕捉核酸群体捕捉多个不同的核酸。类似地,存在于分子集合的不同成员中的通用序列可以允许使用与通用序列互补的通用引物群体来复制或扩增多种不同的核酸。因此,通用捕捉核酸或通用引物包括可以与通用序列特异性杂交的序列。

[0060] 考虑到上述定义,可以理解以下阐述的和权利要求中记载的实施方案。

[0061] 本公开提供了表征候选试剂的方法。该方法可以包括以下步骤:(a) 提供候选试剂文库,其中每个候选试剂附着于具有标签序列的核酸标签;(b) 使所述候选试剂文库与固体支持物接触以将所述候选试剂附着于所述固体支持物,由此形成候选试剂阵列,所述候选试剂阵列包含所述固体支持物上的个别特征,其各自附着于来自所述文库的个别候选试剂;(c) 使所述候选试剂阵列与筛选剂接触,其中所述阵列中的一种或多种候选试剂与所述筛选剂反应;(d) 在所述阵列与所述筛选剂接触期间或之后检测所述阵列,从而测定所述阵列中的至少一种候选试剂与所述筛选剂反应;(e) 对所述阵列上的所述核酸标签测序以测定附着于每种所述候选试剂的标签序列;并且(f) 基于附着于所述至少一种候选试剂的所述标签序列,鉴定所述阵列中与所述筛选剂反应的所述至少一种候选试剂。

[0062] 多种候选试剂中的任一种可以在本文阐述的方法中使用。在一些实施方案中,候选试剂选自下组:蛋白质、核酸、细胞或小分子。应该理解,在一些实施方案中,使用的候选试剂将排除一种或多种类型的物质,如蛋白质、核酸、细胞或小分子。贯穿本公开阐述了可以纳入或从方法或组合物中排除的这些类型的物质的实例。

[0063] 有用的候选试剂可以是具有或怀疑具有生物活性的分子。示例性的生物活性包括但不限于治疗活性、毒性、激素活性、生物分子或细胞的活化、生物分子或细胞的抑制、抗生素活性、抗病毒活性、杀虫剂活性、对生物体或其器官的影响,如精神药理学效应或免疫学效应,等等。特别有用类型的候选试剂是酶抑制剂或酶激活剂,包括例如靶向本文所述的酶的那些。细胞信号传导的候选激活剂或细胞信号传导的候选抑制剂也是有用的。然而,在一些实施方案中,候选试剂不需要具有或也不需要怀疑具有生物学活性。在某些情况下,候选试剂会具有或怀疑具有非生物活性。可以为非生物的活性的实例包括但不限于工业催化、食品防腐、石油加工、聚合物合成等。

[0064] 候选试剂可以是聚合的或非聚合的。特别有用的聚合物包括但不限于蛋白质、核酸、多糖、蛋白质核酸(PNA)和塑料。有用的非聚合物分子包括例如脂质、氨基酸、核苷酸、酶辅因子、代谢物、单糖和其它小分子。

[0065] 可用作候选试剂的蛋白质的实例包括但不限于抗体;酶,如氧化还原酶、转移酶、水解酶、裂合酶、异构酶、连接酶、激酶、磷酸酶、聚合酶、蛋白酶、核酸酶,纤维素酶、木质酶、淀粉酶、脂肪酶、甘露聚糖酶、淀粉酶、葡聚糖酶、木瓜蛋白酶、肾素、组蛋白修饰酶和酯酶;或受体如G偶联受体、细胞表面受体、免疫受体、感觉受体和核激素受体。

[0066] 另一种特别有用类型的候选试剂是细胞,包括例如来自生物体的细胞,包括但不限于哺乳动物,如啮齿类、小鼠、大鼠、兔、豚鼠、有蹄动物、马、绵羊、猪、山羊、牛、猫、狗、灵长类(即人或非人灵长类);植物,如拟南芥(*Arabidopsis thaliana*)、玉米、高粱、燕麦、小麦、稻、芥花(*canola*)或大豆;藻类,如莱茵衣藻(*Chlamydomonas reinhardtii*);线虫,如秀丽隐杆线虫(*Caenorhabditis elegans*);昆虫,如黑腹果蝇(*Drosophila melanogaster*)、蚊、果蝇、蜜蜂或蜘蛛;鱼,如斑马鱼;爬行类;两栖类,如青蛙或非洲爪蟾(*Xenopus laevis*);盘基网柄菌(*Dictyostelium discoideum*);真菌如卡氏肺囊虫(*pneumocystis carinii*)、*Takifugu rubripes*,酵母,酿酒酵母(*Saccharomyces cerevisiae*)或粟酒裂殖酵母(*Schizosaccharomyces pombe*);或恶性疟原虫。其它生物体包括原核生物,如细菌,大肠杆菌、葡萄球菌(*staphylococci*)或肺炎支原体(*mycoplasma pneumoniae*);古菌(*archae*);病毒,如丙型肝炎病毒、埃博拉病毒或人免疫缺陷病毒;或类病毒。细胞可以源自上述生物体的均质培养物或群体,或者备选来自几种不同生物体的集合,例如在混合培养物、群落或生态系统中。

[0067] 本文使用的候选试剂可以是天然存在的,例如从天然群体中收获。例如,候选试剂可以是多细胞生物体或群落中分离的遗传天然细胞。类似地,可以使用从一种或多种遗传天然细胞或生物体获得的蛋白质、核酸或其它生物分子。或者,候选试剂可以是合成的,或者它们可以是天然存在的试剂的工程化变体。例如,基因工程化细胞、蛋白质或核酸可用于本文中阐述的方法中。用作候选试剂的细胞或细胞组分可以源自单细胞生物体或多细胞生物体。在具体的实施方案中,源自这些细胞类型中一种或多种的干细胞、免疫细胞或生物组分可以用作本文的候选试剂。

[0068] 在一些实施方案中,可以与候选试剂的合成平行合成核酸标签。例如,可以使用组合化学方法来合成候选试剂文库,并且可以添加特定核苷酸(或核苷酸序列)以指示将许多不同化学部分中的哪个添加到文库的每个成员中。因此,附着于特定候选试剂的标签中的核苷酸序列提供了特定候选试剂的合成史,任选地可以用于测定该特定候选试剂的化学结构的信息。用于附着和组合合成方法的化学物质描述于例如美国专利号5,565,324; 5,573,905或6,060,596,每篇通过引用并入本文。

[0069] 本文阐述的方法的具体实施方案可以包括组合合成候选试剂文库的步骤,其中如下追踪对每种候选试剂进行的组合合成的个别反应:通过将个或多个核苷酸的独特签名添加到附着于每种候选试剂的核酸标签,从而提供候选试剂文库,其中每种候选试剂附着于独特的核酸标签。

[0070] 核酸标签可以使用本领域已知的适合于与核酸和特定类型的所使用的候选试剂一起使用的多种化学物质中的任一种附着于候选试剂。在具体的实施方案中,将核酸标签共价附着于候选试剂。在候选试剂是核酸或由核酸编码的情况下,连续核酸可以包括标签序列和候选试剂序列。也可以使用化学方法将核酸标签共价附着于充当或编码候选试剂的另一种核酸。适合于将核酸标签附着于候选试剂(无论候选试剂是核酸还是其它种类)的化学物质包括例如N-羟基琥珀酰亚胺酯(NHS酯)、亚氨酸酯、肼、碳二亚胺、马来酰亚胺、卤代乙酰基、吡啶基二硫化物、双吡丙啶(*diazirines*)、点击化学物质(*clickchemistry*) (参见例如美国专利号6,737,236;7,427,678;7,375,234;7,763,736;或8,129,542,其各自通过引用并入本文)或巯基。其它有用的化学物质包括已经用于将核酸附着于珠或其它固体支

持物上的那些物质,如美国专利号7,259,258或7,504,499中阐述,每篇通过引用并入本文。

[0071] 标签序列可以是多种长度中的任一种。较长的序列通常可以适应群体的更大数目和种类的标签。通常,多个中的所有探针将具有相同的长度标签(尽管具有不同的序列),但是对于不同的探针也可以使用不同的长度标签。标签序列的长度可以是至少2、4、6、8、10、12、15、20或更多个核苷酸。或者/另外,标签序列的长度可以是至多20、15、12、10、8、6、4或更少的核苷酸。可以使用的标签序列的实例在例如美国专利公开号2014/0342921A1和美国专利号8,460,865中阐述,其各自通过引用并入本文。

[0072] 本公开的方法可以包括使候选试剂文库与固体支持物接触以将候选试剂附着于固体支持物的步骤。因此,可以在支持体上形成候选试剂的阵列,该阵列包括各自附着于来自文库的个别候选试剂的个别特征。

[0073] 可以使用多种固体支持物中的任一种。特别有用的固体支持物是用于核酸阵列的固体支持物。实例包括玻璃、改性玻璃、功能化玻璃、无机玻璃、微球体(例如惰性和/或磁性颗粒)、塑料、多糖、尼龙、硝酸纤维素、陶瓷、树脂、二氧化硅、基于二氧化硅的材料、碳、金属、光纤或光纤束、聚合物和多孔(例如微量滴定)板。示例性的塑料包括丙烯酸类树脂、聚苯乙烯、聚乙烯和其它材料的共聚物、聚丙烯、聚乙烯、聚丁烯,聚氨酯和Teflon™。示例性的基于二氧化硅的材料包括硅和各种形式的改性硅。

[0074] 在具体的实施方案中,固体支持物可以在容器诸如孔、管、通道、比色皿、培养皿、瓶内或者是所述容器的一部分。特别有用的容器是流动池,例如如WO 2014/142841A1;美国专利申请公开No.2010/0111768A1和美国专利No.8,951,781或Bentley et al.,Nature 456:53-59(2008)中描述,每篇通过引用并入本文。示例性流动池是从Illumina, Inc. (San Diego, CA) 商购获得的那些,其与测序平台(如Genome Analyzer[®]、MiSeq[®]、NextSeq[®]或HiSeq[®]平台)一起使用。另一个特别有用的容器是多孔板或微量滴定板中的孔。

[0075] 任选地,固体支持物可以包括凝胶涂层材料。经由凝胶将核酸附着于固体支持物以从Illumina Inc. (San Diego, CA) 商购或记载于美国专利申请公开号2011/0059865A1、2014/0079923A1、或2015/0005447A1;或PCT公开号WO 2008/093098(其各自通过引用并入本文)的流动池例示。可用于本文中阐述的方法和装置中的示例性凝胶包括但不限于具有胶体结构,如琼脂糖;聚合物网状结构,如明胶;或交联的聚合物结构,如聚丙烯酰胺, SFA(参见例如美国专利申请公开号2011/0059865A1,其通过引用并入本文)或PAZAM(参见例如美国申请公开号2014/0079923 A1或2015/0005447 A1,每篇通过引用并入本文)的凝胶。可以使用凝胶直接(例如经由凝胶与候选试剂之间的共价键)或经由凝胶附着的核酸与已经附着于候选试剂的互补核酸的杂交来附接候选试剂。

[0076] 在一些实施方案中,固体支持物可以配置为可以附着核酸和/或候选试剂的特征阵列。在具体的实施方案中,每个特征将容纳不超过一个候选试剂,或者将以其它方式配置为包含候选试剂的特定混合物的单一种类。特征可以以多种期望格式中的任一种存在。例如,特征可以是孔、凹坑、通道、脊、凸起区域、栓(pegs)、柱等。在一些实施方案中,特征可以包含珠。然而,在具体的实施方案中,特征不需要包含珠或颗粒。示例性的特征包括存在于用于由454LifeSciences(Roche的子公司,Basel Switzerland)或Ion Torrent(Life Technologies的子公司,Carlsbad California)出售的商业测序平台的基底中的孔。具有

孔的其它基底包括例如在美国专利号6,266,459;6,355,431;6,770,441;6,859,570;6,210,891;6,258,568;6,274,320;美国专利公开号2009/0026082 A1;2009/0127589 A1;2010/0137143 A1;2010/0282617 A1或PCT公开号WO 00/63437中描述的蚀刻纤维光学器件和其它基底,每篇通过引用并入本文。在一些实施方案中,基底的孔可以包括如美国专利申请公开号2014/0243224 A1中阐述的凝胶材料(具有或不具有珠),其通过引用并入本文。

[0077] 固体支持物上的特征可以是非金属表面,如玻璃、塑料或上文例举的其它材料上的金属特征。具有金属特征的示例性固体支持物及其制造方法在美国专利号8,895,249或美国专利申请公开号2014/0243224 A1中提供,每篇通过引用并入本文。

[0078] 特征可以在固体支持物上以点或斑的网格出现。特征可以以重复模式或以不规则的非重复模式定位。特别有用的重复模式是六边形模式、直线模式、网格模式、具有反射对称性的模式、具有旋转对称性的模式等。不对称的模式也可以是有用的。间距在最近的相邻特征的不同对之间可以是相同的,或者间距在最近的相邻特征的不同对之间可以变化。

[0079] 高密度阵列的特征为具有小于约15 μm 的平均间距(对于相邻特征)。中密等阵列具有约15至30 μm 的平均间距,而低密度阵列具有大于30 μm 的平均间距。可用于本发明的阵列可以具有小于100 μm 、50 μm 、10 μm 、5 μm 、1 μm 或0.5 μm 的平均间距。上文或本文其它地方阐述的平均间距值和范围旨在适用于有序阵列或随机阵列。

[0080] 在具体的实施方案中,固体支持物上的特征可以各自具有大于约100 nm^2 、250 nm^2 、500 nm^2 、1 μm^2 、2.5 μm^2 、5 μm^2 、10 μm^2 、100 μm^2 、或500 μm^2 或更大的面积。或者/另外,特征可以各自具有小于约1 mm^2 、500 μm^2 、100 μm^2 、25 μm^2 、10 μm^2 、5 μm^2 、1 μm^2 、500 nm^2 、或100 nm^2 或更小的面积。上述范围可以描述从上方观察或成像时固体支持物上的珠或其它颗粒的表观面积。

[0081] 在具体的实施方案中,固体支持物可以包括珠或其它颗粒的集合。具有位于表面上的珠的阵列的实例包括其中珠位于孔中的那些阵列,诸如BeadChipTM array(Illumina Inc., San Diego CA)、用于来自454LifeSciences(Roche的子公司,Basel Switzerland)的测序平台中的基底或用于来自IonTorrent(Life Technologies的子公司,Carlsbad California)的测序平台的基底。具有位于表面上的珠的其它固体支持物描述于美国专利号6,266,459;6,355,431;6,770,441;6,859,570;6,210,891;6,258,568;或6,274,320;美国专利申请公开号2009/0026082A1;2009/0127589A1;2010/0137143A1;或2010/0282617A1或PCT公开号WO 00/63437,每篇通过引用并入本文。几篇上述参考文献描述了在固体支持物之中或之上加载珠前将核酸附着于珠的方法。因此,珠的集合可以包括各自具有附着独特核酸的不同珠。然而,应当理解的是,可以使珠包括通用引物,然后将珠粒加载到阵列上,从而形成用于本文阐述的方法的通用阵列。候选试剂可以在固体支持物上加载珠之前或之后附着于珠。如先前阐述,通常用于珠阵列的固体支持物可以在没有珠的情况下使用。例如,可以将核酸(如探针或引物)或候选试剂直接附着于孔或孔中的凝胶材料。因此,上述参考文献例示了材料、组合物或装置,其可以为了在本文中阐述的方法和组合物中使用而修改。

[0082] 因此,本文阐述的方法中使用的固体支持物可以包括珠阵列,其中不同的候选试剂或不同的核酸附着于阵列中的不同珠。在该实施方案中,每个珠可以附着于不同的候选试剂或核酸,并且珠可以随机分布在固体支持物上,以有效地将不同的核酸附着于固体支持物。任选地,固体支持物可以包括具有容纳不超过单一珠或单一候选试剂的尺寸的孔。在

此类构造中,可以由于源自孔中的珠配合的力而将珠附着于孔。也可以使用附着化学物质或粘合剂将珠保持在孔中。

[0083] 固体支持物可以包含多个核酸或候选试剂或可以通过本文阐述的方法制备以附着多个核酸或候选试剂。例如,固体支持物可以包括至少 $10,100,1 \times 10^3,1 \times 10^4,1 \times 10^5,1 \times 10^6,1 \times 10^7,1 \times 10^8,1 \times 10^9$ 或更多不同的核酸或候选试剂。或者/另外,固体支持物可以包括至多 $1 \times 10^9,1 \times 10^8,1 \times 10^7,1 \times 10^6,1 \times 10^5,1 \times 10^4,1 \times 10^3,100,10$ 或更少的不同核酸或候选试剂。应当理解的是,例如,当已经扩增候选试剂的核酸组分以形成簇时,每个不同的核酸或候选试剂可以以多个拷贝存在。因此,上述范围可以描述固体支持物上不同候选试剂或核酸簇的数目。还应当理解的是,上述范围可以描述不同标签或本文中阐述为对于特定核酸或候选试剂独特的其它序列元件的数目。或者或另外,范围可以描述使用本文阐述的方法在固体支持物上创建的延伸的核酸或修饰的候选试剂的数目。

[0084] 在使固体支持物与核酸或候选试剂接触前,特征可以存在于固体支持物上。例如,在核酸或候选试剂经由与引物杂交而附着于支持物的实施方案中,引物可以在特征处附着,而特征外部的间质区域基本上缺少任何引物。可以使用美国专利号8,895,249、美国专利号8,778,849、或美国专利号申请公开号2014/0243224A1(每篇通过提及并入本文)中阐述的方法将核酸或候选试剂在固体支持物上的预形成特征处捕捉,并且任选地在固体支持物上扩增。

[0085] 在一些实施方案中,在将核酸标签和/或候选试剂附着于固体支持物期间或之后形成特征。例如,固体支持物可以具有引物的草坪(lawn)或者可以一其它方式缺乏特征。在这种情况下,可以通过在固体支持物上附着核酸或候选试剂来形成特征。任选地,可以在固体支持物上扩增捕捉的核酸,使得所得到的簇成为特征。虽然上文将附着例示为在引物和另一个核酸的互补部分之间的捕捉,但应理解的是,除引物之外的捕捉部分可以在预形成的特征处或作为草坪存在。其它示例性的捕捉部分包括但不限于能够与核酸或候选试剂反应以创建共价键的化学部分或能够与核酸或候选试剂上的配体非共价结合的受体。

[0086] 在具体的实施方案中,将核酸引物附着于固体支持物。具有附着的核酸标签的候选试剂文库可以经由核酸标签与附着的核酸引物的杂交而附着于固体支持物。例如,核酸标签可以包括通用引物结合序列,核酸引物可以包括通用引物序列,并且候选试剂可以经由通用引物结合序列与通用引物序列的杂交而附着于固体支持物。作为通用引物的备选,固体支持物可以包括与特定标签序列杂交的靶物特异性引物。

[0087] 每个特征的单一候选试剂阵列可以通过将候选试剂附着于固体支持物而形成。因此,固体支持物上的一个或多个特征可以各自包括单一候选试剂(例如单分子、单细胞或其它单一物质)。在一些实施方案中,特征可以配置为适应不超过特定类型的单一候选试剂。然而,无论特征是否能容纳超过一种候选试剂,特征仍可以包括不超过单一候选试剂、不超过单一核酸标签,或不超过单一候选试剂和单一核酸酸标签。或者,个别特征可以包括多个候选试剂和/或核酸标签。例如,个别特征可以包括具有彼此相同序列的核酸分子全体和/或蛋白质全体。在具体的实施方案中,全体可以通过从单一核酸模板扩增以产生扩增子来产生,例如作为附着于每个特征的簇。

[0088] 本文阐述的方法可以使用多种扩增技术中的任一种。可以使用的示例性技术包括但不限于聚合酶链式反应(PCR)、滚环扩增(RCA)、多重置换扩增(MDA)或随机引物扩增

(RPA)。在一些实施方案中,扩增可以在溶液中进行。优选地,在本公开的方法中使用的扩增技术将在固相上进行。具体地,可以在扩增技术中延伸附着于固体支持物并与核酸标签杂交的一种或多种引物种类(例如,用于存在于核酸标签中的一个或多个通用引物结合位点的通用引物)。以固相PCR实施方案为例,可以将用于扩增的一个或两个引物附着于固体支持物上(例如经由凝胶)。利用两种附着于固体支持物的引物的形式通常称为桥式扩增,因为双链扩增子在已经被复制的模板序列侧翼的两个固体支持物附着的引物之间形成桥样结构。可以用于桥式扩增的示例性的试剂和条件描述于例如美国专利号5,641,658、7,115,400、或8,895,249;或美国专利公开号2002/0055100A1、2004/0096853A1、2004/0002090A1、2007/0128624A1或2008/0009420A1,其各自通过引用并入本文。也可以用附着于固体支持物的扩增引物之一和溶液中的第二引物进行固相PCR扩增。使用固体支持物附着的引物和可溶性引物的组合的示例性形式是乳剂PCR中使用的形式,如记载于例如Dressman et al., Proc. Natl. Acad. Sci. USA 100:8817-8822 (2003)、WO 05/010145、或美国专利申请公开号2005/0130173A1或2005/0064460A1,每篇通过引用并入本文。乳液PCR例示了该形式,并且应当理解的是,出于本文阐述的方法的目的,乳液的使用是任选的,并且事实上对于几个实施方案,不使用乳液。

[0089] 可以为了在本公开的方法中使用而修改RCA技术。可以在RCA反应中使用的示例性组分和RCA产生扩增子的原理描述于例如Lizardi et al., Nat. Genet. 19:225-232 (1998)和美国专利申请公开号2007/0099208A1,每篇通过引用并入本文。用于RCA的引物可以在溶液或附着于固体支持物。引物可以是本文所述的一种或多种通用引物。

[0090] 可以为了在本公开的方法中使用而修改MDA技术。MDA的一些基本原理和有用条件描述于例如Dean et al., Proc. Natl. Acad. Sci. USA 99:5261-66 (2002); Lage et al., Genome Research 13:294-307 (2003); Walker et al., Molecular Methods for Virus Detection, Academic Press, Inc., 1995; Walker et al., Nucl. Acids Res. 20:1691-96 (1992); US 5,455,166; US 5,130,238; 和US6,214,587,每篇通过引用并入本文。用于MDA的引物可以在溶液或附着于扩增位点处附着于固体支持物。同样,引物可以是本文所述的一种或多种通用引物。

[0091] 在具体的实施方案中,可以使用上面例举的扩增技术的组合。例如,RCA和MDA可以组合使用,其中RCA用于在溶液中产生多联体扩增子(例如使用溶液相引物)。然后,可以使用扩增子作为模板,用于使用附着于固体支持物的引物(例如通用引物)进行MDA。在此实例中,将组合的RCA和MDA步骤后产生的扩增子附着于固体支持物。

[0092] 本公开的方法可以包括使候选试剂阵列与筛选剂或刺激物接触的步骤。阵列中的一种或多种候选试剂可以任选地与筛选剂反应或响应刺激物。因此,可以将一种或多种候选试剂分类为命中。

[0093] 在一些实施方案中,筛选剂将通过结合一种或多种候选试剂或通过阻断候选试剂与对候选试剂具有亲和力的分析物之间的结合而与阵列中的一种或多种候选试剂反应。可以检测阵列以鉴定筛选剂被结合的特征。例如,筛选剂可产生可检测的信号,例如由于筛选剂上存在标记物。

[0094] 在一些实施方案中,筛选剂将修饰候选试剂驻留的阵列上的一个或多个特征。例如,筛选剂可以通过化学修饰候选试剂、筛选剂或二者与候选试剂反应。可以检测阵列以鉴

定已经进行修饰的特征。特别地,修饰可以产生例如由于将标记物附着于候选试剂驻留的特征而可检测的信号。示例性修饰包括将标记物部分添加到候选试剂驻留的特征;将亲和部分添加到候选试剂驻留的特征,其中亲和部分结合标记物;在候选试剂驻留的特征处除去荧光猝灭剂,其中由于除去猝灭剂而产生发光信号;将Forster共振能量转移(FRET)供体或受体部分添加到候选试剂驻留的特征,其中添加改变从特征发射的表观发光波长;从候选试剂驻留的特征中除去标记物部分;从候选试剂驻留的特征中除去亲和部分,其中亲和部分结合标记物;从候选试剂驻留的特征中除去Forster共振能量转移(FRET)供体或受体部分,其中除去改变从特征发射的表观发光波长等。

[0095] 在一些实施方案中,筛选剂将与阵列中的一种或多种候选试剂反应以产生分析物。可以检测阵列以鉴定已经产生分析物的特征。例如,分析物可以产生可检测的信号。可以产生的示例性分析物包括发光标记物、用于可检测受体的配体、由酶用于产生可检测产物的底物等。

[0096] 在一些实施方案中,筛选剂不需要与候选试剂文库直接接触。相反,捕捉剂可以与筛选剂接触以产生捕捉产物,然后可以使捕捉产物与候选试剂的集合接触。在另一个实施方案中,可以使捕捉剂与候选试剂的集合接触以产生捕捉产物,然后可以使捕捉产物与筛选剂接触。如通过使用捕捉产物的实施方案例示,筛选剂不需要与候选试剂进行直接接触以使筛选剂可用于指示候选试剂的期望的或可疑的结构或功能。

[0097] 本文中阐述的方法可以进一步包括以下步骤:在使阵列与筛选剂接触期间或之后检测候选试剂阵列,由此测定阵列中的至少一种候选试剂与筛选剂反应。附着有阵列的固体支持物可以处于本文阐述的多种状态中的任一种。例如,固体支持物可以包括候选试剂以及对其附着的核酸标签。或者,固体支持物可以不包括核酸标签,而是处于从候选试剂中除去核酸标签后的状态。在检测步骤期间,核酸(例如标签和/或基于核酸的候选试剂)可以是单分子形式或全体形式。在更进一步的实施方案中,固体支持物可以不包括候选试剂,而是处于除去候选试剂之后的状态。因此,检测可以在本文阐述的方法中的多个点中的任一处发生。

[0098] 可以在本文阐述的筛选步骤中检测多种信号中的任一种,包括例如光学信号,诸如辐射的吸光度、发光发射、发光寿命、发光极化等;Rayleigh和/或Mie散射;磁性能;电性能;电荷;质量;放射性等。可以在本文阐述的方法中检测的示例性标记物包括但不限于荧光团、发光体(luminophore)、生色团、纳米粒子(例如金、银、碳纳米管)、重原子、放射性同位素、质量标记物、电荷标记物、自旋标记物、受体、配体等。

[0099] 具体的实施方案使用成像技术。可以使用本领域已知的检测装置获得图像。实例包括配置为光、明场、暗场、相衬、荧光、反射、干涉或共焦成像的显微镜。在具体的实施方案中,可以使用荧光显微镜(例如共焦荧光显微镜)来检测例如凭借荧光标记物而为荧光的筛选剂(或其它分析物)。还可以使用具有用于荧光检测的光学器件的核酸测序装置(诸如由Illumina, Inc. (San Diego, CA)商业化的Genome Analyzer[®]、MiSeq[®]、NextSeq[®]或HiSeq[®]平台装置;或由Life Technologies (Carlsbad, CA)商业化的SOLiD[™]测序平台)来使荧光样本成像。可以使用的其它成像光学器件包括Bentley et al., Nature 456:53-59 (2008), PCT公开号WO 91/06678, WO 04/018497或WO 07/123744;美国专利号7,057,026, 7,

329,492,7,211,414,7,315,019或7,405,281,和美国专利申请公开号2008/0108082中描述的检测装置中找到的,每篇通过引用并入本文。

[0100] 可以以期望的分辨率获得固体支持物的图像,例如以区分固体支持物上的阵列中的候选试剂。因此,分辨率可以足以区分以至少0.5 μ m、1 μ m、5 μ m、10 μ m、50 μ m、100 μ m、500 μ m、1mm或更大分开的阵列的特征。或者/另外,分辨率可设置为区分以至多1mm、500 μ m、100 μ m、50 μ m、10 μ m、5 μ m、1 μ m、0.5 μ m或更小分开的阵列的特征。

[0101] 本公开的方法可以包括对核酸标签测序以测定附着于固体支持物(例如阵列)上的候选试剂的标签序列的步骤。在许多实施方案中,候选试剂随机位于阵列中,并且测序反应提供信息以定位每个不同的候选试剂。可以在附着有候选试剂及其核酸标签的固体支持物上进行测序。因此,核酸标签不需要从固体支持物上取出以测定其序列。在一些情况下,在对核酸标签进行测序反应之前,可以任选地将候选试剂从固体支持物中取出。

[0102] 在一些实施方案中,候选试剂是核酸或附着于编码候选试剂的核酸。在此类实施方案中,可以使用测序技术来测定候选试剂的序列。可以在附着有候选试剂的固体支持物上进行测序。可以将候选试剂及其核酸标签一起测序(例如在连续测序读取中)或分别测序(例如,可以在标签的测序和候选试剂的测序之间进行非测序步骤)。

[0103] 本公开的方法可以采用候选试剂阵列的基于时间的检测或动力学测量。因此,阵列的检测可以包括在数个时间点时获取阵列上的一个或多个个别特征的信号。可以在基于时间或动力学的分析中检测本文阐述的多种信号和标签中的任一种。

[0104] 本公开的方法可以包括对核酸标签测序以测定附着于阵列中的候选试剂的标签序列的步骤。在一些实施方案中,候选试剂是核酸,候选试剂由与其附着的核酸编码(例如,附着于mRNA或cDNA的蛋白质候选试剂),或者候选试剂含有核酸(例如含有核酸的细胞)。测序技术,如合成测序(SBS)技术,是一种特别有用的方法。

[0105] SBS可以如下进行。为了启动第一SBS循环,可以使一种或多种标记的核苷酸、DNA聚合酶、SBS引物等与固体支持物上的一个或多个特征(例如,将核酸附着于固体支持物的特征)接触。可以检测SBS引物延伸使标记核苷酸被掺入的那些特征。任选地,核苷酸可以包括可逆终止部分,该可逆终止部分一旦已经将核苷酸添加到SBS引物便终止进一步的引物延伸。例如,可以将具有可逆终止剂部分的核苷酸类似物添加到引物,使得随后的延伸不能发生,直到递送去封闭剂以除去该部分。因此,对于使用可逆终止的实施方案,可以将去封剂递送到固体支持物(在检测发生之前或之后)。可以在各个递送步骤之间进行清洗。然后,可以将循环重复n次以将引物延伸n个核苷酸,由此检测长度n的序列。可以为了与本公开的组合物、装置或方法一起使用而容易改编的示例性SBS程序、流体系统和检测平台记载于例如Bentley et al.,*Nature* 456:53-59 (2008),PCT公开号W0 91/06678,W0 04/018497或W007/123744;美国专利号7,057,026,7,329,492,7,211,414,7,315,019或7,405,281,和美国专利申请公开号2008/0108082,每篇通过引用并入本文。

[0106] 可以使用其它使用循环反应的测序程序,如焦磷酸测序。焦磷酸测序检测特定的核苷酸被掺入新生核酸链中时无机焦磷酸(PPi)的释放(Ronaghi,et al.,*Analytical Biochemistry* 242(1),84-9 (1996);Ronaghi,*Genome Res.*11(1),3-11 (2001);Ronaghi et al.*Science* 281(5375),363 (1998);或美国专利号6,210,891,6,258,568或6,274,320,其各自通过引用并入本文)。在焦磷酸测序中,释放的PPi可以通过ATP硫酸化酶立即转化为三

磷酸腺苷(ATP)来检测,并且可通过萤光素酶产生的光子检测产生的ATP水平。因此,可以通过发光检测系统来监测测序反应。用于基于荧光的检测系统的激发辐射源对于焦磷酸测序程序不是必需的。可用于将焦磷酸测序法应用于本公开的装置,组合物或方法的有用的流体系统,检测器和程序记载于例如PCT专利申请公开号W02012/058096,美国专利申请公开号2005/0191698A1,或美国专利号7,595,883或7,244,559,每篇通过引用并入本文。

[0107] 连接测序反应也是有用的,包括例如Shendure et al.*Science*309:1728-1732 (2005);或美国专利号5,599,675或5,750,341中描述的连接测序反应,每篇通过引用并入本文。一些实施方案可以包括杂交测序程序,如记载于例如Bains et al.,*Journal of Theoretical Biology* 135 (3),303-7 (1988);Drmanac et al.,*Nature Biotechnology* 16,54-58 (1998);Fodor et al.,*Science* 251 (4995),767-773 (1995);或PCT专利申请公开号W0 1989/10977,每篇通过引用并入本文。在连接测序和杂交测序两者中,存在于阵列位点处的核酸经历寡核苷酸递送和检测的重复循环。本文或本文引用的参考文献中阐述的组合物、装置或方法可以容易地适用于连接测序或杂交测序程序。典型地,寡核苷酸是荧光标记的,并且可以使用荧光检测器来检测,所述荧光检测器类似于本文或本文引用的参考文献中就SBS程序而言描述的那些。

[0108] 一些测序实施方案可以利用涉及DNA聚合酶活性的实时监测的方法。例如,可以通过带有荧光团的聚合酶和 γ -磷酸盐标记的核苷酸之间的荧光共振能量转移(FRET)相互作用或用零式波导(zeromode waveguides,ZMW)来检测核苷酸掺入。用于基于FRET的测序的技术和试剂描述于例如Levene et al.*Science* 299,682-686 (2003);Lundquist et al.*Opt.Lett.*33,1026-1028 (2008);Korlach et al.*Proc.Natl.Acad.Sci.USA* 105,1176-1181 (2008),每篇均通过引用并入本文。

[0109] 一些测序实施方案包括检测在将核苷酸掺入延伸产物中时释放的质子。例如,基于释放质子的检测的测序可以使用从Ion Torrent (Guilford,CT,Thermo Fisher子公司)商购的电检测器和相关技术或美国专利申请公布号2009/0026082 A1;2009/0127589 A1;2010/0137143 A1;或US 2010/0282617 A1中描述的测序方法和系统,每篇通过引用并入本文。

[0110] 本公开的方法可以进一步包括从固体支持物中取出一种或多种候选试剂的步骤。在具体的实施方案中,可以在对特征处的核酸标签测序之前,从阵列的特征中取出候选试剂。当候选试剂干扰测序技术时或当期望保护候选试剂免于测序技术中使用的试剂时,这可以是有利的。在一些实施方案中,可以从阵列中选择性除去已经在筛选中鉴定为命中的一种或多种候选试剂。

[0111] 在一些实施方案中,候选试剂可以通过切割将它附着于固体支持物的系链来取出。例如,候选试剂可以经由与其附着的核酸标签栓系到固体支持物。在通过与互补核酸杂交将标签核酸附着于支持物的情况下,可以通过使杂交复合物变性来实现取出。或者,核酸标签可以包含位于对候选试剂的附着点和标签序列之间的切割位点。如此,候选试剂可以通过靶向切割位点的切割反应与标签序列分开。在一些方法中,可以在扩增或修饰步骤期间将切割位点引入到核酸中。例如,可以在延伸步骤期间将切割位点引入延伸的引物中。

[0112] 示例性的切割位点包括但不限于对导致键断裂的化学、酶学或物理过程易感的部分。例如,位置可以由内切核酸酶识别的核苷酸序列。合适的内切核酸酶及其识别序列在

本领域中是众所周知的,并且在许多情况下甚至可以是商品化的(例如来自New England Biolabs,Beverly MA;ThermoFisher,Waltham,MA或Sigma Aldrich,St.Louis MO)。特别有用的内切核酸酶将在核酸链中远离其在核酸中的结合位点的3'的位点处断裂键,其实例包括I型或II型限制性内切核酸酶。

[0113] 光不稳定部分提供有用的切割位点。示例性的光不稳定部分包括但不限于(1-(4,5-二甲氧基-2-硝基苯基)乙基)酯(即DMNPE)和(1-(2-硝基苯基)乙基)酯(即NPE)。见Meth.Enzymol.291:307-347(1998),其通过引入并入本文。其它光不稳定部分及其合成和使用的方法描述于WO 91/06678或美国专利申请公开号20100092957 A1,每篇通过引用并入本文。光不稳定部分可以有利地用于从个别特征中靶向切割候选试剂。可以使用空间分辨辐射(例如来自激光、聚焦光源和/或空间滤器/掩模)来解决经由光不稳定部分栓系特定候选试剂的个别特征。因此,可以选择性除去特定的候选试剂(与同一阵列上的其它特征的候选试剂相比)。

[0114] 在一些实施方案中,切割位点是无碱基位点或具有易于被除去以创建无碱基位点的碱基的核苷酸。易于被除去以形成无碱基位点的核苷酸的实例包括尿嘧啶和8-氧代-鸟嘌呤。在一个具体的实施方案中,使用购自New England Biolabs的USER™试剂。可以用于切割核酸的切割位点和方法的其它实例例如在美国专利号7,960,120中阐述,其通过引用并入本文。

[0115] 本公开进一步提供了用于生产蛋白质阵列的方法。该方法可以包括以下步骤:(a)提供附着于固体支持物的cDNA分子的文库;(b)扩增固体支持物上的cDNA分子以形成簇,其中每个簇包括来自文库的特定cDNA分子的多个拷贝;(c)在所述簇处转录所述多个拷贝以产生附着于每个所述簇的多个mRNA分子;和(d)翻译簇中的mRNA分子以产生附着于每个簇的多个蛋白质。

[0116] 本公开还提供了用于产生蛋白质阵列的方法,其包括以下步骤:(a)提供mRNA分子文库,其中所述文库中的个别mRNA分子包含靶序列和标签序列,(b)从所述文库中衍生第一子文库,所述第一子文库包含具有所述标签序列或其互补物的核酸,其中所述核酸附着于固体支持物上的个别特征,(c)从所述文库中衍生第二子文库,所述第二子文库包含具有所述靶序列和所述标签序列或其互补物的核酸;(d)使所述第二子文库与所述第一子文库接触,从而经由所述标签序列及其互补物的杂交将所述第二子文库的核酸附着于所述固体支持物;并且(e)翻译所述固体支持物上的所述靶序列以产生附着于所述个别特征的蛋白质阵列。

[0117] 基因组DNA(gDNA)、信使RNA(mRNA)或拷贝DNA(cDNA)分子的文库可以源自表达一种或多种感兴趣蛋白质的一种或多种生物体。例如,文库可以从怀疑表达期望的蛋白质或具有期望活性的蛋白质的生物体获得。可以使用本领域已知的任何生物体,例如本文阐述的那些。生物体可以是工业、治疗、诊断或预后兴趣的。例如,可以从人生物体中获得文库并且出于预后或诊断目的进行筛选。或者,可以怀疑生物体表达产生治疗效果或具有工业应用的基因。在一些实施方案中,文库是从非人生物体获得的,或者在其它情况下是非人起源的。

[0118] 在一些应用中,gDNA、mRNA或cDNA分子文库包括在重组生物体文库中表达的单一基因的变体。可以使用已知的蛋白质工程技术产生此类变体,包括但不限于Sambrook et

al., *Molecular Cloning: A Laboratory Manual*, Third Ed., Cold Spring Harbor Laboratory, New York (2001) 和于 Ausubel et al., *Current Protocols in Molecular Biology*, John Wiley and Sons, Baltimore, Md. (1999) 中阐述的, 其各自通过引用并入本文。例如, 可以通过随机诱变基因的全部或部分来产生变体。也可以使用半随机诱变技术。

[0119] 在本文阐述的方法中表达的基因(例如 mRNA、cDNA 或蛋白质)可以选自下组: 抗体、酶、受体、激酶、磷酸酶、聚合酶、蛋白酶、酯酶、核激素受体、组蛋白修饰酶和本文中例示的其它酶。基因可以是重组的, 例如在与基因的天然来源不同的生物体中表达, 或者具有对于基因而言非天然的修饰。在具体的实施方案中, 可以从通常用于克隆或表达的细菌、酵母、昆虫、哺乳动物或其它细胞系统分离 gDNA、mRNA、cDNA 或蛋白质文库。

[0120] 可以构建不同核酸(例如 gDNA、mRNA 或 cDNA)的文库, 使得文库的每个个别成员包含蛋白质编码序列和标签序列。可以随机或系统性将标签序列分配给每个蛋白质编码序列(即, 在已知标签与已知变体先验关联的情况下)。作为实例, 当合成也编码蛋白质的核酸构建体时, 可以通过将随机核苷酸掺入标签区域中的一个或多个位置中随机分配标签。在另一个实例中, 可以将不同核酸标签的群体与不同蛋白质编码核酸的群体随机连接, 使得平均而言, 每种蛋白质编码核酸将与独特的标签序列连接。作为系统性标签化的实例, 可以以下述方式将特定标签合成到核酸构建体中, 使得它与构建体中的蛋白质编码序列中的突变或变体先验关联。

[0121] 本文阐述的方法可以包括从文库中衍生第一和第二子文库的步骤。在一些实施方案中, 通过从文库中取出级份或通过物理分开文库来衍生子文库。一般地, 文库将是包含每个候选试剂的多个拷贝的流体。可以使用多种候选试剂中的任一种, 如 gDNA、mRNA、cDNA、大环、环状肽、融合分子(例如核酸-蛋白质融合体)、展示的构建体(例如噬菌体上的肽)或蛋白质等。因此, 通过分级或分开流体衍生的子文库预计将具有大致相同的内容。在一些情况下, 可以在核酸文库(例如 DNA 或 RNA 文库)上进行扩增步骤以创建多个拷贝。

[0122] 可以使用本文其它地方阐述的方法将 gDNA、mRNA、cDNA 或蛋白质文库附着于固体支持物。在一个具体的实施方案中, 使用 mRNA 分子文库, 其中每个 mRNA 分子包括编码候选试剂的序列以及标签序列。通过包括使文库的 mRNA 分子与固体支持物接触以将 mRNA 分子附着于固体支持物的方法衍生第一子文库。例如, 固体支持物可以具有固体支持物附着的核酸引物, 并且 mRNA 分子可以经由与核酸引物杂交而附着于固体支持物。作为更具体的实例, mRNA 分子可以包括通用引物结合序列, 核酸引物包括通用引物序列, 并且 mRNA 分子可以经由通用引物结合序列与通用引物引物序列的杂交而附着于固体支持物。然后, 可以将附着的 mRNA 分子在固体支持物上复制或扩增以产生标签序列的互补物。然后可以将与第一子文库从相同的文库衍生的第二子文库与第一子文库的扩增子(或拷贝)接触, 从而经由标签序列及其互补物的杂交将第二子文库的核酸附着于固体支持物。以这种方式, 固体支持物已经经修饰而包括可用于筛选反应或翻译以产生由第二子文库的 mRNA 分子表达的蛋白质文库的 mRNA 分子文库。

[0123] 在一些实施方案中, 可以在溶液中复制或扩增第一核酸或第二核酸子文库之一。因此, 扩增或复制的子文库将包括可以与另一个子文库中的标签序列杂交的互补标签序列。可以将子文库之任一附着于固体支持物, 并且可以将另一个子文库杂交以创建具有核酸分子文库的固体支持物, 所述核酸分子文库可用于筛选反应、产生 mRNA 的转录反应和/或

产生蛋白质文库的翻译反应。

[0124] 对于利用mRNA文库的实施方案,可以通过包括逆转录文库中的一些或全部个别mRNA分子以产生cDNA的方法衍生第一子文库。逆转录可以发生在溶液中或固体支持物上。然后可以在溶液中或在固体支持物上扩增得到的cDNA分子。因此,可以产生具有文库中标签序列的互补物的cDNA分子。在一些情况下,cDNA分子可以包含不超过mRNA分子的一部分,例如,表达来自多结构域蛋白的特定结构域或结构域子集的部分。

[0125] 本发明的方法可以包括将cDNA转录成mRNA的步骤。可以使用已知的混合物,如 Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Third Ed., Cold Spring Harbor Laboratory, New York (2001) 和于 Ausubel et al., *Current Protocols in Molecular Biology*, John Wiley and Sons, Baltimore, Md. (1999) (每篇通过引用并入本文) 中阐述的混合物,或者商品化的混合物实施转录。在具体的实施方案中,对支持物附着的cDNA分子进行转录。转录反应的mRNA产物可以附着在产生它们的特征处。例如,可以使用本文实施例11中阐述的技术来实现附着。将蛋白质附着于编码它们的DNA分子也是有用的(例如通过直接附着或乳胶液滴中的共定位)。蛋白质也可以附着于编码它们的DNA分子附近(例如核酸可编程蛋白质阵列(NAPPA))。

[0126] 本发明的方法可以包括将mRNA翻译成蛋白质的步骤。可以使用已知的混合物,如 Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Third Ed., Cold Spring Harbor Laboratory, New York (2001) 和于 Ausubel et al., *Current Protocols in Molecular Biology*, John Wiley and Sons, Baltimore, Md. (1999) (每篇通过引用并入本文) 中阐述的混合物,或者商品化的混合物实施翻译。在具体的实施方案中,对支持物附着的mRNA分子进行翻译。翻译反应的蛋白质产物可以附着在产生它们的特征处。例如,可以使用本文实施例11中阐述的技术来实现附着。

[0127] 本文阐述的用于生产蛋白质阵列的方法可以用于筛选蛋白质的方法中。该方法可以包括以下步骤:(i) 生产蛋白质阵列;(ii) 使所述蛋白质阵列与筛选剂接触,其中所述阵列中的一种或多种蛋白质与所述筛选剂(或刺激物)反应;和(iii) 在与筛选剂(或刺激物)接触期间或之后检测蛋白质阵列,由此测定阵列中的至少一种蛋白质与筛选剂(或刺激物)反应。

[0128] 在具体的实施方案中,筛选剂通过与一种或多种蛋白质结合或通过阻断蛋白质与对蛋白质具有亲和力的分析物之间的结合而与一种或多种蛋白质反应。阵列的检测可以包括检测与一种或多种蛋白质结合的筛选剂。例如,筛选剂可以是发光的,并且检测可以通过检测阵列上的发光来进行。

[0129] 在一些实施方案中,筛选剂通过化学修饰一种或多种蛋白质与一种或多种蛋白质反应。在这种情况下,阵列的检测可以通过检测一种或多种经修饰的蛋白质来实现。例如,一种或多种经修饰的蛋白质可以是发光的,并且可以通过检测阵列上的发光来实现检测。

[0130] 在一些实施方案中,筛选剂通过产生分析物产物与一种或多种蛋白质反应。在这种情况下,阵列的检测包括检测分析物产物。例如,分析物产物可以是发光的,并且检测可以通过检测阵列上的发光来进行。

[0131] 如本文中先前阐述,阵列的检测可以包括在几个时间点获取阵列上的一个或多个个别特征的信号。

[0132] 可以使用本文中阐述或本领域已知的方法对作为gDNA、cDNA或mRNA分子的一部分的标签序列测序。例如,可以在进行、进行了或将进行筛选的固体支持物上进行测序反应。因此,本文阐述的方法可以包括下述步骤:对固体支持物上的标签序列或其互补物测序,由此测定固体支持物上个别特征处的标签序列或其互补物位置。此类方法可以进一步包括基于附着于至少一种蛋白质的标签序列鉴定阵列中与筛选剂反应的至少一种蛋白质的步骤。

[0133] 本文阐述的方法可以进一步包括对固体支持物上的候选gDNA、cDNA或mRNA测序的步骤。因此,可以在固体支持物上进行一个或多个测序反应以测定标签序列的身份和/或候选序列的身份,从而测定固体支持物上的个别特征处期望的候选试剂的位置。

[0134] 在一些情况下,gDNA、cDNA或mRNA可以比测序方法的平均读段长度更长。在此类情况下,可以对候选核酸进行几次测序运行,每次使用不同的引物。例如,第一测序运行可以使用与本文或本文引用的参考文献中阐述的测序方法和平台中通常采用的通用引发序列杂交的引物。随后的测序运行可以使用第二测序引物进行,该第二测序引物在第一测序引物已经杂交的位置的下游杂交。可以根据测序运行的预期读段长度来分隔引物,使得可以获得模板核酸的重叠或邻接读段。可以对于足够数目的测序引物重复几次测序运行以覆盖候选核酸的期望区域。

[0135] 当候选核酸长于测序方法的平均读段长度时的另一项选择是使用盒式方法。在盒式方法中,在候选核酸中突变或修饰核酸的小区或域。盒可以选择为足够短以便在单次测序运行中读取。只要在盒上游的候选核酸的区域在文库的所有成员中是一致的,它可以有效地充当通用引发位点,并且测序引物可以与该区域杂交以在该盒处进行定位测序。

[0136] 在一些实施方案中,候选gDNA、cDNA或mRNA不需要测序。相反,标签序列可以已经与已知候选核酸序列先验相关,使得对标签测序将足以鉴定候选核酸。在特定特征处存在的标签的先验知识可以与筛选步骤中从特征观察到的信号相关联,以鉴定候选核酸、其编码的蛋白质或蛋白质的性质。

[0137] 在一些实施方案中,候选核酸(例如gDNA、cDNA或mRNA)或蛋白质可通过切割将它附着于固体支持物的系链来取出。例如,此类候选试剂可以通过标签序列栓系于固体支持物。切割位点可以位于与候选试剂的附着点和标签序列之间。如此,候选试剂可以通过靶向切割位点的切割反应与标签序列分开。示例性的切割位点包括但不限于本文其它地方所述的或本领域中已知的那些切割位点。

[0138] 因此,本公开的方法可以包括选择性除去附着于阵列的一个或多个特征的gDNA、cDNA、mRNA或蛋白质分子的步骤。选择性除去可以包括光介导的光不稳定性键的切割,例如使用本文其它地方阐述的试剂和方法进行,所述光不稳定性键将分子附着于特征。

[0139] 本公开提供了阵列,其包括(a)固体支持物;(b)附着于所述固体支持物的不同cDNA分子的文库,其中每个不同的cDNA分子附着于所述固体支持物上的个别特征,并且其中每个特征包括特定cDNA分子的多个拷贝;(c)附着于所述cDNA分子的mRNA分子,其中每个所述cDNA分子与相应的附着的mRNA分子互补;和(d)附着于所述mRNA分子的蛋白质分子,其中每个所述蛋白质分子由相应的附着的mRNA分子编码。

[0140] 本公开进一步提供了阵列,其包括(a)mRNA分子文库,其中所述文库中的个别mRNA分子包含靶序列和标签序列,(b)包含具有所述标签序列的互补物的核酸的固体支持物,其中所述核酸附着于固体支持物上的个别特征,其中所述个别mRNA分子的标签序列与所述固

体支持物上的所述个别特征处相应的互补标签序列杂交,并且其中通过翻译所述mRNA分子衍生的蛋白质附着于相应的mRNA分子上。

[0141] 本公开的阵列可以包括从本文阐述的方法的一个或多个步骤产生的组件。例如,根据本文阐述的方法,核酸(如gDNA、cDNA或mRNA种类)可以附着于固体支持物。类似地,根据本文阐述的方法,蛋白质可以附着于固体支持物。在具体的实施方案中,如实施例11和111中所阐述,cDNA可以附着于固体支持物,编码的mRNA可以附着于cDNA,并且编码的蛋白质可以附着于mRNA。如那些实施例中所阐述,经由cDNA的附着是任选的,并且取而代之,mRNA可以附着于固体支持物(例如经由与互补标签的杂交或通过其它附着方法),并且编码的蛋白质可以附着于mRNA。

[0142] 任选地,mRNA分子可以通过与RNA聚合酶形成的复合物而附着于其编码cDNA分子。附着可以通过复合的cDNA、mRNA和RNA聚合酶之间的共价交联来介导。类似地,蛋白质分子可以经由与核糖体形成的复合物附着于其编码mRNA上。附着可以通过复合的mRNA、蛋白质和核糖体之间的共价交联来介导。

[0143] 本文阐述的筛选步骤中的多种筛选剂、标记物或产物之任一种可以存在于本公开的阵列中。例如,蛋白质定位的特征可以经发光标记的,任选地经由发光分子(如筛选剂)与蛋白质的特异性非共价结合或经由共价附着于蛋白质的发光部分。不是所有的功能都必然是标记的。相反,可以选择性标记包含具有期望的或选择的活性的蛋白质的特征亚组,而不标记其它特征。例如,根据文库的内容和筛选的性质,可以标记少于50%、25%、10%、5%或1%的特征。

[0144] 本文阐述的多种候选试剂和/或核酸标签中的任一种可存在于本公开的阵列中。阵列上的不同种类的数目、含有阵列上的种类的特征的密度或附接在特定特征处的每个种类的数目可以在本文中就制备和使用阵列的方法所阐述的范围内。

[0145] 本公开还提供了筛选细胞的方法。该方法可以包括以下步骤:(a)提供多个不同细胞,其中每个所述不同细胞包含具有标签序列的核酸标签;(b)使所述不同细胞的混合物与固体支持物接触以形成附着于所述固体支持物的细胞阵列;(c)筛选所述固体支持物上的细胞阵列的至少一种光学特征,其中筛选反应包括检测附着于所述固体支持物上的个别细胞;(d)对附着于所述固体支持物的核酸标签的所述标签序列测序;并且(e)基于所述候选细胞的所述光学特征和所述标签序列,将所述阵列中的至少一个细胞鉴定为候选细胞。

[0146] 在具体的实施方案中,使用细胞作为本文阐述的方法中的候选试剂。细胞可以是多细胞生物体分离的天然细胞、包含单细胞生物体的天然细胞、来自多细胞生物体的基因工程细胞或基因工程单细胞生物体。用于本文中阐述的方法或组合物中的细胞可以从天然来源获得,或者它们可以从离体培养物中获得。

[0147] 可以体外复制、培养或扩增的细胞是特别有用的。例如,在筛选步骤之前或之后制备候选细胞的一个或多个拷贝可以是有用的。因此,在筛选(或者细胞的克隆)中鉴定为命中的细胞可以从预形成的原种(preformed stock)或者从筛选中使用的固体支持物中分离。例如,可以进一步使用或操纵分离的细胞以更充分地表征细胞或在治疗程序中使用细胞。

[0148] 对于采用遗传修饰细胞的实施方案,遗传修饰可以是本领域已知的任何种类,包括例如导致表达非天然重组蛋白、表达突变重组蛋白、缺失全部或部分天然存在的蛋白质

的编码序列、抑制天然存在的蛋白质的表达、增强天然存在的蛋白质的表达、产生非天然分析物或抑制天然分析物的产生的那些。例如,候选细胞文库中的一种或多种基因的编码序列可以包含点突变、缺失(例如除去整个蛋白质编码序列、蛋白质的结构域或其它部分)或插入(例如嵌合体)。

[0149] 文库中的候选细胞可以包括核酸标签。在实施例1V中阐述和/或在下面阐述用于对细胞进行标签化的几种示例性方法。

[0150] 在一些实施方案中,核酸标签可以共价附着于细胞表面。通常,采用方法仅用单一核酸标签序列来对文库中的每个细胞进行标签化(尽管每个细胞中或每个细胞上可以存在多个核酸分子,每个核酸分子具有相同标签序列的拷贝)。例如,可以物理分离细胞,使得每个细胞可以单独地与核酸标签分子反应。用于物理分离细胞的方法包括例如将每个细胞分离到单独的容器、微板上的孔、阵列上的特征、珠、液滴致动器装置(droplet actuator device)中的流体微滴、乳剂中的流体微滴、或囊泡。

[0151] 用于创建可以递送核酸标签的液滴的特别有用的方法包括例如由RainDance Technologies(Billerica,MA)商业化的或在美国专利号9,017,623或8,857,462中描述的,每篇通过引用并入本文。用于创建和添加标签到液滴的其它方法由10X Genomics商业化或美国专利申请公开号2014/0155295A1;2014/0206554A1;2014/0227684A1或2014/0378322A1中描述,每篇通过引用并入本文。可以修改这些方法,使得将细胞(或本文阐述的其它候选试剂)加载到个别液滴中,并且加载的液滴与含有核酸标签的流体相互作用,使得个别液滴以单一标签核酸种类和单个候选试剂种类结束。当然,核酸种类的多个拷贝或候选试剂种类的多个拷贝可以存在于个别液滴中。在一些实施方案中,核酸标签附着于递送至个别液滴的珠。例如,可以使用本文实施例1V中阐述的附着化学物质将核酸标签附着于细胞表面。

[0152] 可用于分离细胞(或其它候选试剂)以例如对细胞进行标签化的特别有用的微滴操作装置是微滴致动器,如记载于例如美国专利号8,637,242,美国专利号6,911,132,题目为“Apparatus for Manipulating Droplets by Electrowetting-Based Techniques,”2005年6月28日公告;Pamula et al.,美国专利公开号20060194331,题目为“Apparatuses and Methods for Manipulating Droplets on a Printed Circuit Board,”2006年8月31日公布;Pollack et al.,国际专利公开号W0/2007/120241,题目为“Droplet-Based Biochemistry,”2007年10月25日公布;Shenderov,美国专利号6,773,566,题目为“Electrostatic Actuators for Microfluidics and Methods for Using Same,”2004年8月10日公告;Shenderov,美国专利号6,565,727,题目为“Actuators for Microfluidics Without Moving Parts,”2003年5月20日公告;Kim et al.,美国专利公开号20030205632,题目为“Electrowetting-driven Micropumping,”2003年11月6日公布;Kim et al.,美国专利公开号20060164490,题目为“Method and Apparatus for Promoting the Complete Transfer of Liquid Drops from a Nozzle,”2006年7月27日公布;Kim et al.,美国专利公开号20070023292,题目为“Small Object Moving on Printed Circuit Board,”2007年2月1日公布;Shah et al.,美国专利公开号20090283407,题目为“Method for Using Magnetic Particles in Droplet Microfluidics,”2009年11月19日公布;Kim et al.,美国专利公开号20100096266,题目为“Method and Apparatus for Real-time Feedback

Control of Electrical Manipulation of Droplets on Chip,”2010年4月22日公布; Velev,美国专利号7,547,380,题目为“Droplet Transportation Devices and Methods Having a Fluid Surface,”2009年6月16日公告;Sterling et al.,美国专利号7,163,612,题目为“Method,Apparatus and Article for Microfluidic Control via Electrowetting,for Chemical,Biochemical and Biological Assays and the Like,”2007年1月16日公告;Becker et al.,美国专利号7,641,779,题目为“Method and Apparatus for Programmable Fluidic Processing,”2010年1月5日公告;Becker et al.,美国专利号6,977,033,题目为“Method and Apparatus for Programmable Fluidic Processing,”2005年12月20日公告;Decre et al.,美国专利号7,328,979,题目为“System for Manipulation of a Body of Fluid,”2008年2月12日公告;Yamakawa et al.,美国专利公开号20060039823,题目为“Chemical Analysis Apparatus,”2006年2月23日公布;Wu,美国专利公开号20110048951,题目为“Digital Microfluidics Based Apparatus for Heat-exchanging Chemical Processes,”2011年3月3日公布;Fouillet et al.,美国专利公开号20090192044,题目为“Electrode Addressing Method,”2009年7月30日公布;Fouillet et al.,美国专利号7,052,244,题目为“Device for Displacement of Small Liquid Volumes Along a Micro-catenary Line by Electrostatic Forces,”2006年5月30日公告;Marchand et al.,美国专利公开号20080124252,题目为“Droplet Microreactor,”2008年5月29日公布;Adachi et al.,美国专利公开号20090321262,题目为“Liquid Transfer Device,”2009年12月31日公布;Roux et al.,美国专利公开号20050179746,题目为“Device for Controlling the Displacement of a Drop Between Two or Several Solid Substrates,”2005年8月18日公布;以及Dhindsa et al.,“Virtual Electrowetting Channels:Electronic Liquid Transport with Continuous Channel Functionality,”Lab Chip,10:832-836 (2010),每篇通过引用并入本文。

[0153] 在具体的实施方案中,可以遗传修饰文库中的细胞以包含核酸标签。例如,群体中的个别细胞可以携带编码标签序列的质粒,或者可以修饰个别细胞的基因组以包含标签序列。在一些情况下,在核酸构建体中编码标签序列,所述核酸构建体还包括将在本文阐述的方法中筛选的遗传变体。

[0154] 可以构建不同细胞的文库,使得每个个别细胞获得随机分配的标签序列,或者可以修饰每个细胞以包括已知的标签序列。例如,随机标签化可以通过在合成核酸构建体时将随机核苷酸掺入标签区域中的一个或多个位置中来进行,所述核酸构建体也将编码待筛选的遗传变体。在另一个实例中,不同核酸标签的群体可以与不同变体编码核酸的群体随机连接,使得平均而言,每个变体核酸将与独特的标签序列连接。然后可以将构建体添加到文库中的细胞。类似地,不同核酸标签的群体可以共价附着于不同细胞的群体,使得平均每个细胞附着于独特的标签序列。

[0155] 在其它实施方案中,可以构建不同细胞的文库,使得每个个别细胞获取与已知细胞先验相关的已知标签。作为先验标签化的实例,可以以下述方式将特定标签合成到核酸构建体中,使得它与该构建体中的突变或变体相关。在一些实施方案中,细胞在物理上彼此分开,并且使具有已知标签序列的核酸与细胞接触以形成单一类型的标签与单一细胞的共价附着。

[0156] 在一个具体的实施方案中,核酸标签可以附着于珠,并且珠可以结合细胞。任选地,珠可以附着于具有特异性结合亲和力的抗体或细胞。然而,可以使用其它附着形式来将珠附着于细胞,包括例如本文在将核酸附着于候选试剂或固体支持物的背景中例示的。

[0157] 在一些实施方案中,核酸标签可以通过将核酸标签共价附着于细胞的质膜脂质或脂肪酸而附着于细胞。或者,核酸标签可以通过与细胞的质膜脂质中的蛋白质共价附着而附着于细胞。作为共价附着的备选,核酸标签可以包括与细胞表面上的配体结合的受体,或者核酸标签可以包括与细胞表面上的受体结合的配体。在下面的实施例IV中阐述示例性的附着方法。

[0158] 可以使用本文阐述的一种或多种方法将细胞文库附着于阵列,用于将其它类型的候选试剂附着于固体支持物。例如,固体支持物可以包括核酸引物,并且细胞可以经由核酸标签与核酸引物的杂交而附着于固体支持物。在一些情况下,核酸标签可以包括通用引物结合序列,核酸引物可以包括通用引物序列,并且候选试剂可以经由通用引物结合序列与通用引物序列的杂交而附着于固体支持物。

[0159] 本公开的方法可以包括延伸与核酸(如核酸标签)杂交的固体支持物附着的引物的步骤。所得到的延伸引物将包括标签序列和来自核酸的其它序列(尽管是互补形式)。因此,延伸的引物是来自候选试剂的核酸的空间标签化形式。应该理解的是,存在于核酸中的除了标签序列之外的序列元件也可以包含在延伸的引物中。此类元件包括例如引物结合位点、切割位点、其它标签序列(例如样品鉴定标签)、捕捉序列、用于核酸结合蛋白或核酸酶的识别位点等。

[0160] 引物的延伸可以使用本文例示的或本领域已知的用于核酸扩增或核酸测序的方法进行。在具体的实施方案中,可以通过经由聚合酶催化(例如DNA聚合酶、RNA聚合酶或逆转录酶)将一个或多个核苷酸添加至引物的3'末端。可以使用化学或酶促方法将一个或多个核苷酸添加到引物的3'或5'端。例如,可以通过化学或酶促(例如连接酶催化)方法将一个或多个寡核苷酸添加至引物的3'或5'端。引物可以以模板指导的方式延伸,由此延伸产物同与引物杂交的模板核酸互补。在一些实施方案中,使用RNA模板通过逆转录酶延伸DNA引物,由此产生cDNA。因此,在本文阐述的方法中制备的延伸探针可以是逆转录的DNA分子。用于延伸核酸的示例性方法阐述于美国专利申请公开号US 2005/0037393A1或美国专利号8,288,103或8,486,625,每篇通过引用并入本文。

[0161] 可以通过延伸复制与引物杂交的核酸的全部或部分。例如,延伸的探针可以包括从核酸复制的至少1、2、5、10、25、50、100、200、500、1000或更多个核苷酸。可以控制延伸产物的长度,例如,在延伸反应中使用可逆终止的核苷酸并进行有限数目的延伸循环。可以如SBS技术所例示的那样运行循环,并且使用经标记的核苷酸是不必要的。因此,在本文阐述的方法中产生的延伸引物可以包含不超过1000、500、200、100、50、25、10、5、2或1个从核酸复制的核苷酸。当然,延伸的探针可以是上文阐述的范围之内或之外的任何长度。

[0162] 可以使用本文中就筛选其它候选试剂而言阐述的方法来筛选附着于阵列的细胞文库。在具体的实施方案中,筛选细胞阵列可以包括用筛选剂处理细胞的步骤。因此,筛选剂可以与阵列上的至少一个候选细胞结合。任选地,筛选剂是发光的,并且将通过检测至少一个候选细胞的发光来进行筛选反应。

[0163] 在一些实施方案中,递送到阵列的筛选剂将修饰阵列中的至少一个候选细胞。例

如,筛选剂可以刺激阵列上的至少一个候选细胞。或者,筛选剂可以抑制或甚至杀死阵列上的至少一个候选细胞

[0164] 与细胞阵列接触的筛选剂可以增加或减少至少一个候选细胞的发光,并且筛选反应可以包括检测至少一个候选细胞的发光的步骤。

[0165] 如本文其它地方所述,检测步骤可以包括动力学或基于时间的测量。例如,检测细胞阵列可以包括在几个时间点获取阵列上的一个或多个个别特征的信号的步骤。

[0166] 可以在本文中阐述的方法中对附着于细胞或存在于细胞中的核酸标签测序。可以在细胞上对附着于细胞的核酸测序。或者,可以经由例如引物延伸方法(如实施例IV中例示的那些引物延伸方法)将核酸转移到固体支持物上,并且可以在固体支持物上对含有标签序列(或其互补物)的延伸引物的区域测序。本文其它地方阐述的用于对固体支持物连接的核酸测序的方法可以用于附着于细胞或与衍生核酸标签的细胞附近的特征的核酸标签。

[0167] 在一些情况下,可以期望从固体支持物除去细胞,同时使核酸标签(或标签的拷贝)附着于固体支持物。然后可以在没有细胞的情况下对标签序列测序。

[0168] 对于核酸标签存在于细胞中的实施方案,可以在阵列的表面上使细胞裂解,导致细胞内容物的释放和核酸标签的局部捕捉。核酸标签可以包括与固体支持物上的引物互补的通用序列区。因此,可以以适合扩增和/或测序的形式在固体支持物上捕捉核酸标签。可以任选地扩增固体支持物上捕捉的核酸。可以使用本文阐述的方法在固体支持物上对固体支持物上捕捉的核酸或由其产生的扩增子测序。

[0169] 本公开的方法可以包括从阵列取出一个或多个细胞的步骤。例如,可以选择性除去筛选步骤中鉴定为命中的一个或多个细胞。任选地,可以使用保留细胞存活力的条件。如此,可以培养、复制或扩增取出的细胞。可以使用本文别处对其它候选试剂阐述的技术和试剂取出细胞。例如,使用光不稳定性接头和空间过滤光束对于将特定的细胞与阵列中的其它细胞分离是特别有用的。

[0170] 在一些实施方案中,在筛选步骤之后,不需要从阵列中取出活细胞。相反,标签序列可以已经与已知细胞先验关联,使得对标签测序将足以鉴定细胞,而不需要进一步分离或表征细胞。存在于特定特征处的标签的先验知识可以与筛选步骤中从特征观察到的信号相关联以鉴定细胞。

[0171] 以下实施例旨在例示而不是限制本发明。

[0172] 实施例I

[0173] 小分子筛选

[0174] 新一代测序平台,如由Illumina(San Diego,CA)商业化的测序平台,为构建用于药物发现的集成高通量筛选平台提供了基础。通过用实现对靶分子附着的官能团修饰核酸,可以利用测序流动池作为基底来构建用于高通量化合物筛选的阵列。可以在改变附着靶物的情况下使用相同的平台进行多种筛选测定法。在图1A到图1F中图示并在下面进行了用于在测序平台上筛选小分子的示例性方法。

[0175] 图1A显示第一步,其中在基于带的合成仪器(tape-based synthesis instrument)上合成50,000个不同的核酸分子,每个核酸分子长29个核苷酸。带包括分别合成个别核酸种类的个别可寻址位点。核酸分子在5'末端还包含官能团(FG),例如巯基、胺或N-羟基琥珀酰亚胺基团。核酸的5'区域包括编码通用引物的10个核苷酸(对于所有的50,

000个核酸都是相同的),并且3'区域编码50,000个不同的标签序列之一(也称为“编码”序列)。示例性的基于带的DNA合成仪记载于美国专利申请公开文本号2011/0178285 A1,其通过引用并入本文。

[0176] 图1B中显示了第二步,其中将单模板杂合核酸与合成带的每个孔中的核酸连接以创建通用突出端引物。如图所示,在不同核酸的通用引物序列与单模板杂合物上的突出端的杂交之后发生连接。任选地,可以在连接事件之前从带中除去不同的核酸(例如纯化或修饰核酸)。其它任选的构造包括将官能团附着于单模板杂合物而非不同的核酸。或者/另外,单模板杂合物可以具有表面附着部分,如生物素、叠氮化物或炔烃基团。

[0177] 如图1C所示,进行第三步以机器人方式将特定候选试剂(例如“化合物”)添加到合成器带上的每个个别孔。每种候选试剂包括与不同核酸上的官能团反应的反应基团。因此,每个候选试剂被附着于合成仪带上的特定核酸标签,从而创建编码的候选试剂。

[0178] 如图1D所示,进行步骤4以将编码的候选试剂固定在流动池表面上的可区分的位点上。通过核酸的单模板部分与流动池表面上的互补通用引物的杂交可发生固定化。可以通过PAZAM凝胶或其它水凝胶将引物附着于流动池表面,如例如记载于美国专利申请公开号2011/0059865A1,或美国专利号9,012,022,每篇通过引用并入本文。

[0179] 在步骤5中,如图1E所示,使筛选剂(也称为“靶分子”)与位于流动池表面上的固定化的编码候选试剂接触。筛选剂包括可以使用诸如Illumina测序仪的光学装置检测的荧光标记物。可以例如实时获得图像以测定在每个可区分的位点处的筛选剂的结合动力学。在图1E的实例中,附着于代码N的候选试剂是“命中”,如由期望的结合动力学概貌测定。

[0180] 如图1F所示,代码N的序列可以根据在流动池上进行的测序方案来测定。代码N的序列可以与合成仪带上合成它的位置相关联,并且继而可以基于哪个候选试剂被递送到合成仪带上合成代码N的位置的知识来测定“命中”的身份。

[0181] 实施例II

[0182] 蛋白质筛选

[0183] 固相扩增方法允许核酸的高度多重呈现。特别有用的固相扩增方法是桥式扩增(也称为簇形成),其可以如美国专利号5,641,658;美国专利申请公开号2002/0055100;美国专利号7,115,400;美国专利申请公开号2004/0096853;美国专利申请公开号2004/0002090;美国专利申请公开号2007/0128624;和美国专利申请公开号2008/0009420(每篇通过引用并入本文)所述的那样进行。

[0184] 该实施例描述了使用由固相扩增得到的核酸簇在流动池上形成蛋白质阵列。该技术的一些关键优点是:(a)它提供了大量的含蛋白质的特征,这些特征的尺寸相对较小,但在表面上可以光学区分,从而可以筛选大量不同的蛋白质的活性;和(b)特征包含编码蛋白质的核酸,核酸适于测序,由此可以鉴定流动池的每个特征处的蛋白质。

[0185] 如图2所示,构建cDNA文库以具有P5和P7引物结合位点以及转录和翻译位点的衔接头。将文库附着于流动池,并使用桥式扩增在流动池的表面上扩增文库成员以形成簇。对于描述P5和P7引物结合序列、用衔接头制备核酸文库的方法、以及制备簇的方法,参见例如上文就桥式扩增技术而言引用的参考文献,Bentley et al.,Nature 456:53-59(2008),美国专利申请公开号2011/0059865A1和美国专利号7,741,463,每篇通过引用并入本文。

[0186] 使用RNA聚合酶将得到的cDNA簇转录成mRNA。将mRNA维持于转录它的cDNA簇附近,

通过 (a) 没有在cDNA构建体的末端的转录终止位点,使得RNA聚合酶停止,并且mRNA转录物保持与cDNA结合, (b) 使用已知转录停止位点 (例如来自大肠杆菌的Trp操纵子), 其会类似停止RNA聚合酶, (c) 使用停止转录的化学方法 (例如双环霉素); 或 (d) 在cDNA簇中具有与转录的RNA杂交的互补序列。

[0187] 然后,将mRNA (其定位于转录它的cDNA簇) 翻译以生成蛋白质。这将通过mRNA中的起始密码子 (其可以与cDNA文库构建体的5' 末端处的衔接头中的转录起始序列一起引入) 来促进。此外,mRNA在3' 末端不具有终止密码子。如图2所示,缺少终止密码子导致mRNA末端的核糖体停滞,使得蛋白质经由核糖体保持附着于mRNA (mRNA继而通过停顿的RNA聚合酶附着于cDNA簇)。或者,核糖体可以通过化学手段 (例如氯霉素/嘌呤霉素) 停止。可以使用容易获得的兔网状细胞裂解物、细菌S30或小麦胚芽提取物系统进行偶联的转录和翻译。

[0188] 在备选的方法中,可以在溶液中生成蛋白质,然后附着于流动池或其它表面。虽然此方法不包括表面上的扩增,但是当基于溶液的技术更容易获得或适用于正在评估的蛋白质时可以是有益的。用筛选剂或其它刺激物筛选得到的蛋白质阵列以鉴定命中。对簇中的cDNA测序以鉴定蛋白质命中。

[0189] 实施例III

[0190] 使用标签的蛋白质筛选

[0191] 新一代测序平台,如由Illumina (San Diego,CA) 商业化的测序平台为建立用于蛋白质进化的整合高通量筛选平台提供了基础。图3A和图3B中图示了一个示例性方法。

[0192] 作为第一步,创建具有对感兴趣的靶蛋白质和随机标签 (也称为“代码”) 的随机或半随机突变的载体文库。构建载体以表达包含蛋白质的编码序列和标签的mRNA。不同随机标签的数目超过文库中蛋白质变体的数目,使得每个变体mRNA具有包含独特标签的高概率。例如,10个核苷酸的随机化序列将提供 10^6 个不同的标签,而20个核苷酸的随机化序列将提供 10^{12} 个不同的标签。

[0193] 作为突变的随机引入和随机标签的使用的替代,第一步可以以进行预定义的突变并且与已知标签相关联的方式进行。

[0194] 在第二步中,培养已经用载体转化的细菌文库以从相应载体产生mRNA。裂解文库以从文库 (即mRNA文库) 中的各种载体中释放mRNA转录物的混合物。

[0195] 在第三步中,将mRNA文库拆分以创建两个子文库。该文库包括每个成员的多个拷贝,使得拆分文库导致两个子文库中存在的相同成员。再次,将独特的标签序列附着于每个蛋白质变体序列。

[0196] 在第四步中,如图3B所示,将第一子文库附着于流动池,在流动池上扩增,并对所得扩增子的编码序列和相应标签测序。如此,定位附着每个成员的位点。然后,将第二子文库的成员与流动池接触以使来自第二子文库的成员经由标签序列的互补性与来自第一子文库的相同成员产生的扩增子杂交。如此,可以基于测序结果和标签与其互补物之间预期的杂交来推断第二子文库的每个成员所在的位点。然后翻译已经在流动池上捕捉的第二子文库的成员以在流动池上表达蛋白质靶标。每种蛋白质都以其保持与其翻译的RNA附着的方式表达。例如,可以使用实施例II中阐述的技术来实现附着。然后用筛选剂或其它刺激物筛选蛋白质,允许以高通量方式鉴定期望的功能。结果是蛋白质活性信号,其定位于流动池上与编码蛋白质的RNA序列相关的位点。

[0197] 该方法可以允许每个流动池泳道具有3000万个蛋白质或更高的高通量。该筛选过程避免了与基于乳液的筛选相关的挑战。例如,可以使用IIIllumina测序仪器在流动池上的每个特征处获得定量且基于动力学的数据,以允许比使用传统的基于乳剂的技术可获得的标准更精确且更具有选择性的筛选标准。

[0198] 实施例IV

[0199] 细胞筛选

[0200] 该实施例描述了筛选细胞的方法,其提供荧光激活细胞分选(FACS)的高通量优势以及实时跟踪动态过程的能力,这通常是与荧光显微术相关的优点,同时增加下一代测序鉴定群体中的个别细胞的优点。因此,本文阐述的方法允许显示期望行为的细胞的回收和表征。具体地,本实施例使来自IIIllumina Inc. (San Diego,CA)的下一代测序平台适合于通过荧光以超高通量监测个别细胞响应特定刺激的表型行为(例如通过荧光报道分子的表达来测量),随后对每个个别细胞表面上显示的标签进行测序,用于随后鉴定和回收。该平台可为细胞工程的高通量筛选提供标准方法。

[0201] 细胞具有与小分子和蛋白质不同的,并且在许多情况下相对于小分子和蛋白质的能力改善的潜在的治疗能力。细胞可以主动地移动到身体中的特定位置,它们可以感知多种外部刺激物,并且它们可以整合多种信息源,以精确的输出来作出应答。另外,基于细胞的应答可以具有复杂的动力学模式,其可以通过预先编程的功能或通过添加外部因素来控制。目前,正在探索基于细胞的治疗来治疗极其多种疾病,从使用工程化微生物治疗克罗恩氏病到使用患者衍生的工程化免疫细胞治疗癌症。重要的是,微生物或人细胞可以经工程化改造而成为治疗上有用的。

[0202] 用于细胞工程化改造的工作流

[0203] 细胞工程(在微生物或人细胞中)是一个考虑许多变量的复杂过程。因此,不同于尝试单一设计,有利的是设计然后可以选择或筛选的候选细胞的文库。具体而言,细胞工程可以包括以下步骤:(1)遗传回路的设计,其可以包括用于外部信号的一个或多个传感器、信息整合网络和信号依赖性响应(通常涉及基因表达的响应);(2)在候选细胞文库中修饰上文提及的组分的一个或多个遗传靶标(例如使用CRISPR-Cas9);和(3)高通量筛选以鉴定在许多不同的候选细胞中赋予期望的细胞行为的命中。

[0204] 强力的和高通量的筛选技术可以为细胞工程提供重大的优势,因为需要测试的细胞数目较大(通常尝试数千种不同的设计),并且由于工程化过程中引入的修改将显现大量且变化的细胞行为。期望的筛选也将能够遵循复杂的时间动力学,同时实现非常高的通量。本文阐述的筛选提供了这些优点。

[0205] 细胞筛选方法有以下两个阶段:

[0206] 阶段1:表型监测。如图4所示,将具有合适的荧光报道物的活工程化细胞文库加载到用于测序平台(例如IIIllumina MiSeq[®]、NextSeq[®]、HiSeq[®]或Genome Analyzer[®]平台)的流动池中。个别细胞将随机分布并附着于流动池表面(使用特定的附着化学物质,取决于待分析的特定细胞)。以限定的时间间隔记录每个个别细胞的荧光,如目前在测序期间对IIIllumina流动池所做的那样扫描流动池表面。注意,个别细胞具有的平均大小可以与目前测序簇相当(或更大),从而使测序平台的分辨率非常适合于筛选。另外,细胞可以携带数千个荧光报道分子,因此信号强度将会与目前通过测序平台检测到的信号强度相似(或更

大)。在特定的时间,将细胞暴露于筛选剂(或其它刺激物),并且细胞的应答(例如,如报道基因表达的变化,或任何其它类型的细胞报道物的荧光强度的变化-如细胞质中的钙浓度、通过亚细胞定位变化进行荧光团淬灭等反应)。通过监测荧光报告物随时间的变化,将有可能测定响应多种多样的刺激物(或刺激物的组合)的表型变化的时间动力学。

[0207] 阶段2:通过解码个别标签进行单细胞鉴定。对群体中的每个细胞分配标签。下文描述了使细胞进行标签化的三种不同技术。在第一步中,将标签附着于个别细胞上。不依赖于标签类型(例如珠、膜附着的核酸等),首先将细胞分选到个别容器(例如微孔)中,使得一个细胞位于每个容器中。然后为每个容器中的细胞提供独特的标签,如图5所示。

[0208] 2.1用珠的单细胞条形码编码(Barcoding)。用1至10个珠对群体中的每个细胞(例如对于T细胞为约10 μ m直径)进行标签化。珠可以具有约1-2 μ m的直径。有用的珠包括例如在Illumina的BeadArray技术中使用的或在美国专利号7,622,294或8,741,630中描述的,每篇通过引用并入本文。通过先前将个别细胞分离到个别容器中,并且对每个容器添加共享相同的独特标签(即每个容器和每个标签之间存在已知的对应)的珠,用编码单一标签的珠对每个细胞进行标签化。细胞标签化可以以多种方式实现,例如通过将珠与识别天然存在于靶细胞表面上的特定表位的抗体共价连接进行。注意,如此虽然每个珠标签化的细胞将随机分布在流动池的表面上,但每个细胞将与其自身的珠在空间上共同定位(图6)。因此,在阶段1之后,通过解码流动池上的珠并将鉴定的标签分配给与相应的珠最接近的细胞来揭示每个细胞的身份。可以使用合成测序技术(参见例如Bentley et al., Nature 456:53-59 (2008), PCT公开号WO 91/06678, WO 04/018497或WO 07/123744;美国专利号7,057,026, 7,329,492, 7,211,414, 7,315,019或7,405,281,以及美国专利申请公开号2008/0108082A1,每篇通过引用并入本文)或美国专利号8,460,865或Gunderson et al., Genome Research 14:870-877 (2004)中阐述的解码技术(每篇通过引用并入本文)来解码珠上的标签。

[0209] 2.2使用序列标签的单细胞鉴定。珠标签化的替代或补充是使用每个细胞表面上直接显示的唯一序列标识符。为此,在细胞加载到流动池之前(例如在阶段1之前),群体中的每个细胞将用表面展示的个别单链DNA标签进行修饰。核酸标签附着化学物质将取决于特定的细胞类型。例如,如由Sledz et al., J Am Chem Soc 134:765-8 (2012) (其通过引用并入本文)所报道,人细胞可以用与质膜脂质共价连接的单链DNA分子标签化。作为另一个实例,酵母细胞可以用与Aga1-Aga2细胞壁复合物上展示的HaLoTag蛋白共价连接的单链DNA分子标签化。

[0210] 每个个别标签至少由三个区域组成。区域1连接到附着部分(例如脂质或蛋白质,取决于细胞类型)。区域2由限定数目的碱基组成,所述数据取决于需要的复杂性(plexity)(使用所有4种核苷酸,可能产生长度为“n”的序列的4ⁿ个组合;例如,4个碱基长的序列将涵盖用于鉴定256个不同个别细胞的标签,5个碱基长的序列将导致1024个不同的标签等)。因此,凭借少达9个碱基的标签序列,就有可能个别鉴定超过250,000个不同的细胞。区域3由与测序引物互补的DNA序列组成。注意每个细胞将用编码单一标签的核酸进行标签化,这通过预先分离容器中的单个细胞,并对每个容器添加共享相同独特标签序列的核酸(如前所述,每个容器与每个标签之间有已知的对应)。

[0211] 在分开的步骤(已经获得表型测量之前或之后进行)中,将例如以下面描述的两种

方式之一来测定每个个别细胞的身份。

[0212] 2.2.1附着于每个细胞的特定标签的原位测序。每个细胞将显示特定标签的多个拷贝。因此,容易在测序仪上检测信号(见图7)。此外,鉴定每个标签的读段长度相对较短(例如,9个碱基足以鉴定>250,000个标签),完全在从商品化测序仪典型的数百个核苷酸的读取长度内。

[0213] 原位标签测序可以以多种方式实现。一种可能性是直接在活细胞表面上测序。另一种可能性是在测序之前固定细胞(如果从活细胞泄漏的核苷酸干扰测序反应)。第三种可能性是从每个细胞分离和捕捉核酸标签。为此,用包含与所有细胞附着的核酸共有的区域互补的区域的捕捉探针(例如测序引物)修饰流动池表面,允许细胞初始附着于表面,并促进测序前捕捉分离的核酸。任选地,可以在核酸标签的表面捕捉之后除去细胞。

[0214] 2.2.2.附着于流动池的标签拷贝的原位测序。如图8所示,将细胞表面展示的核酸标签与流动池表面附着的引物(具有部分互补序列)杂交。细胞表面核酸标签通过用特异性限制酶消化而与细胞分离,然后将细胞从流动池中洗出。然后将流动池表面附着的核酸标签延伸,有效地创建初始细胞标签的拷贝,然后对其测序。在许多情况下,不需要桥式扩增拷贝,因为细胞可能留下由数千个核酸组成的直径约1 μ m的“斑点”。然而,如果想要的话,可以进行桥式扩增(在这种情况下,可以对核酸标签添加合适的引物结合位点,并且可以将第二引物类型附着于流动池表面)。

[0215] 在某些情况下,如图9所示,可以使用膜锚定核酸将标签附着于个别细胞。膜锚定核酸将包括接头区、限制性位点(RE)、测序引物结合位点(SBS3)、标签序列(条形码)和捕捉序列。然后如Sieden et al., J Am Chem Soc 134:765-8 (2012) (其通过引用并入本文)中所述,或如图10中所示,通过具有与捕捉序列互补的序列的表面附着的核酸表面捕捉细胞。

[0216] 在筛选步骤中记录指示细胞表型行为的荧光信号(例如在特定的培养基条件下随时间段测量来自报道物的荧光)之后,可以将核酸标签在流动池表面上复制,然后可以将细胞冲走。表面上的拷贝有效地保留了有关哪些细胞位于流动池表面上的哪个位置的空间信息,如图11所示。

[0217] 通过记录表面上每个细胞的空间位置,可以将标签身份分配给每个记录的荧光时间系列,从而有效地将表型响应与细胞身份相关联。然后可以使用这些身份来鉴定单个容器中的个别克隆(其中最初对细胞进行标签化,如图12中所示),或者通过用与选择的细胞表面上存在的标签互补的核酸下拉来从细胞混合物回收选择的细胞。

[0218] 贯穿本申请,引用了各种出版物、专利申请或专利。在此通过引用将这些出版物的全部公开并入本申请中,以更完全地描述本发明所属的技术发展水平。

[0219] 术语“包括”在本文中旨在是开放式的,不仅包括叙述的要素,还包括任何另外的要素。

[0220] 虽然已经参照上面提供的实施方案描述了本发明,但是应当理解,在不脱离本发明的情况下可以进行各种修改。因此,本发明仅由权利要求书限制。

步骤 1: 合成 500,000 个 29 聚体寡聚物 + 附着基团 (-SH、胺; NHS).

1. 合成仪 = 50,000 个寡聚物 / 天 10 天内 500,000 个寡聚物
2. 条形码 = 19 聚体; “引物” = 10 聚体

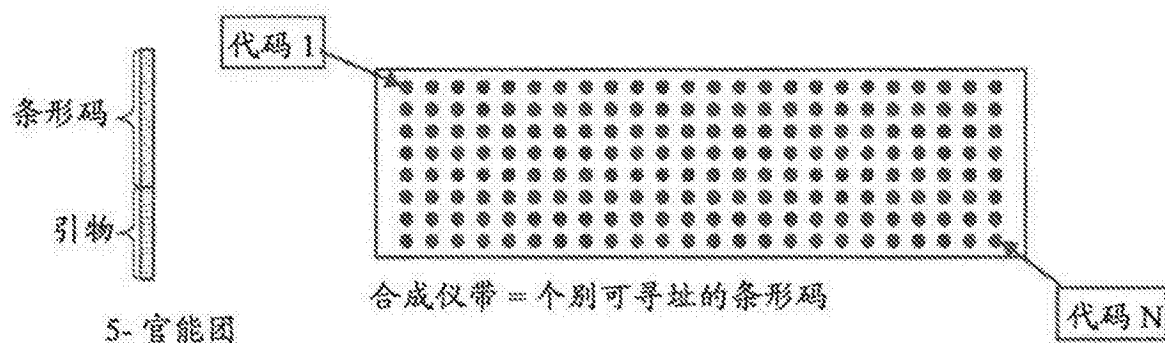


图1A

步骤 2: 在个别孔中连接突出端引物与单模板序列

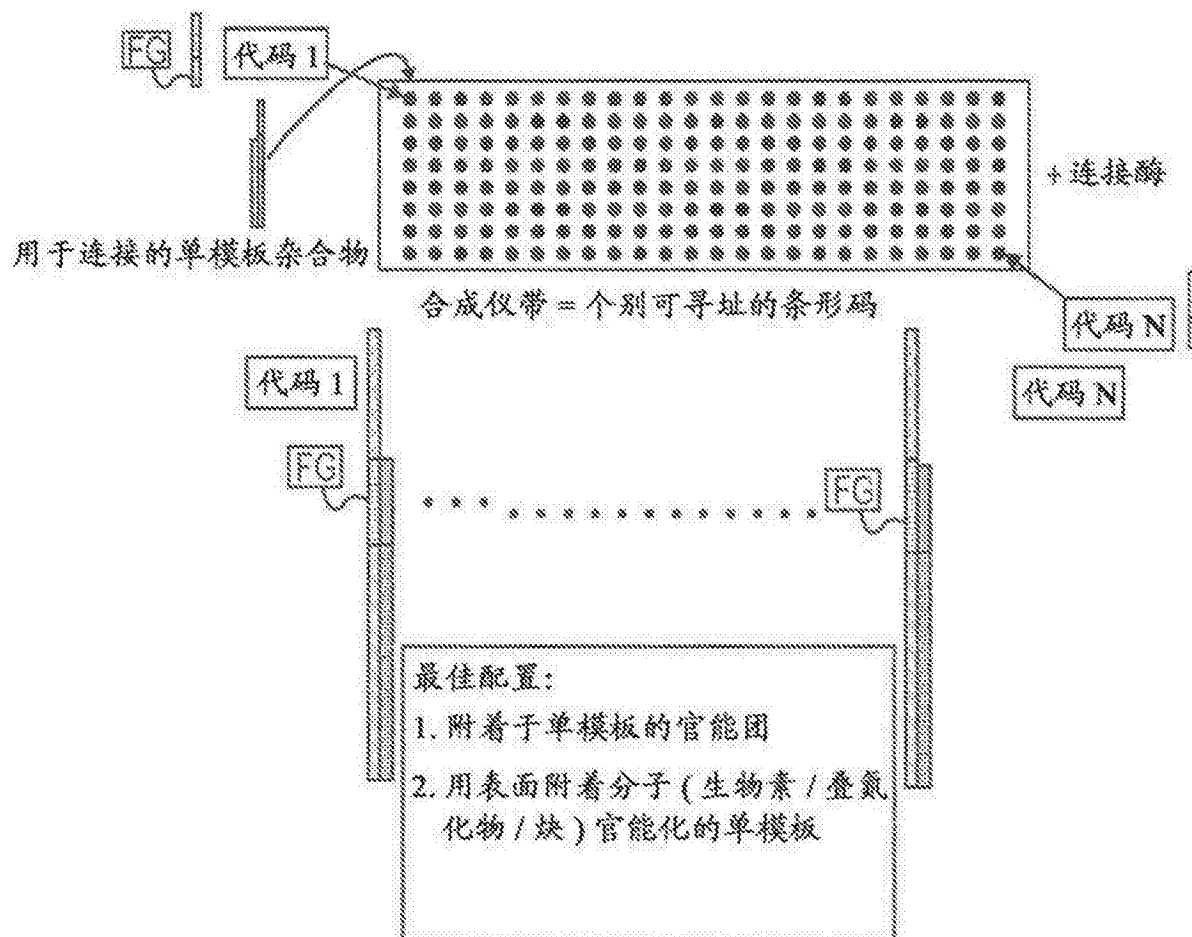


图1B

步骤 3: 机器人添加具有反应性基团 (RG) 的化合物至具有官能化代码的个别孔

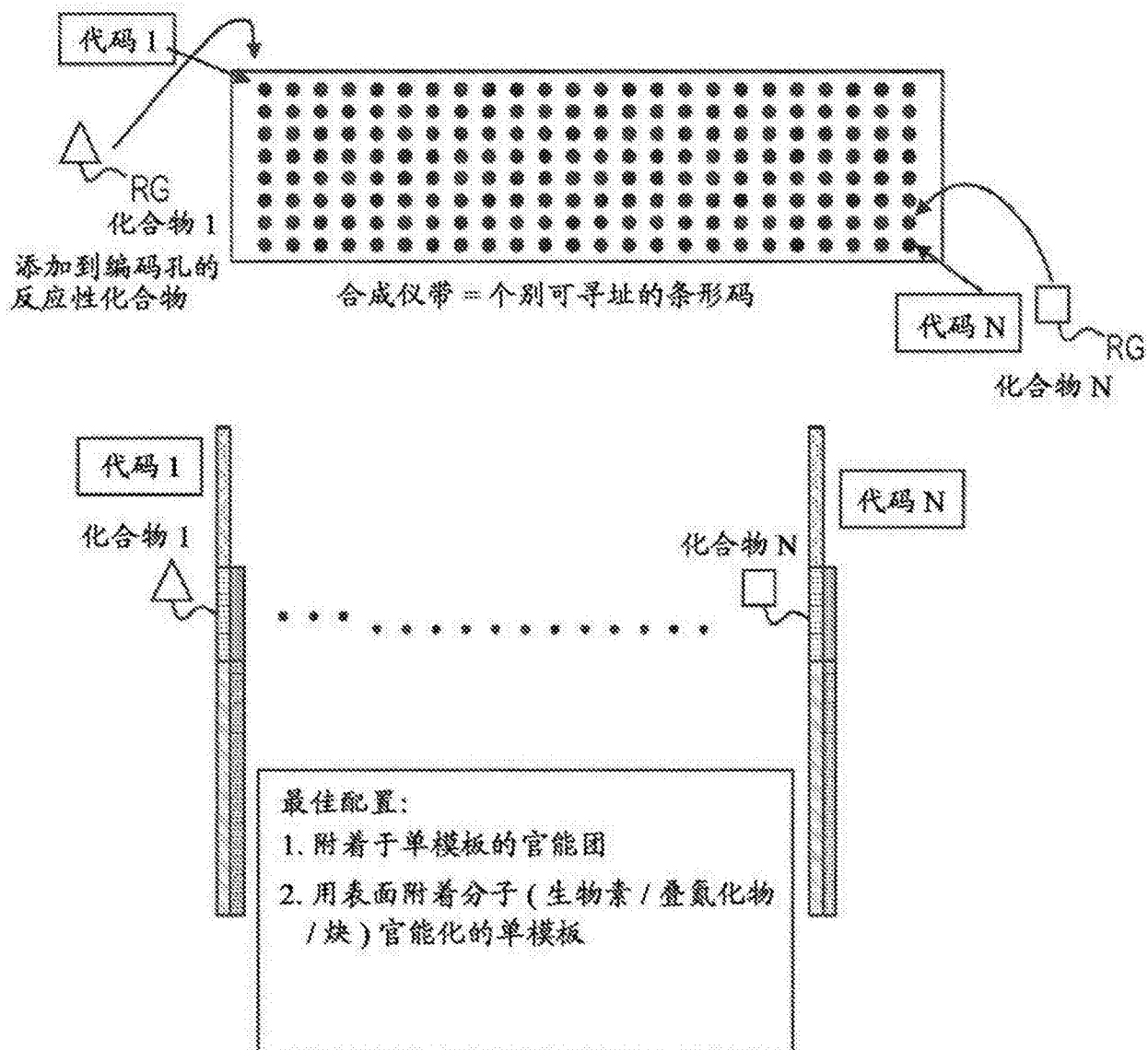


图1C

步骤 4: 经由杂交或表面功能团将化合物固定化至单分子表面
(或 Pazam 流动池表面)

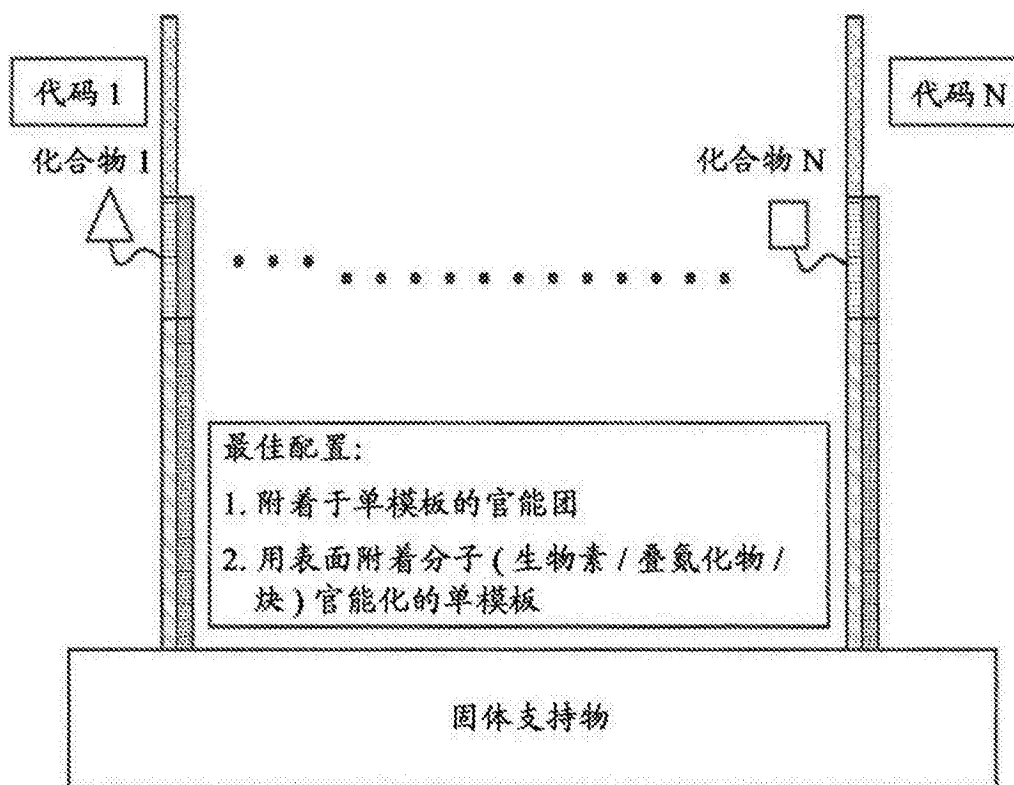


图1D

步骤 5: 在靶分子中流动并且记录对化合物的结合动力学的实时图像

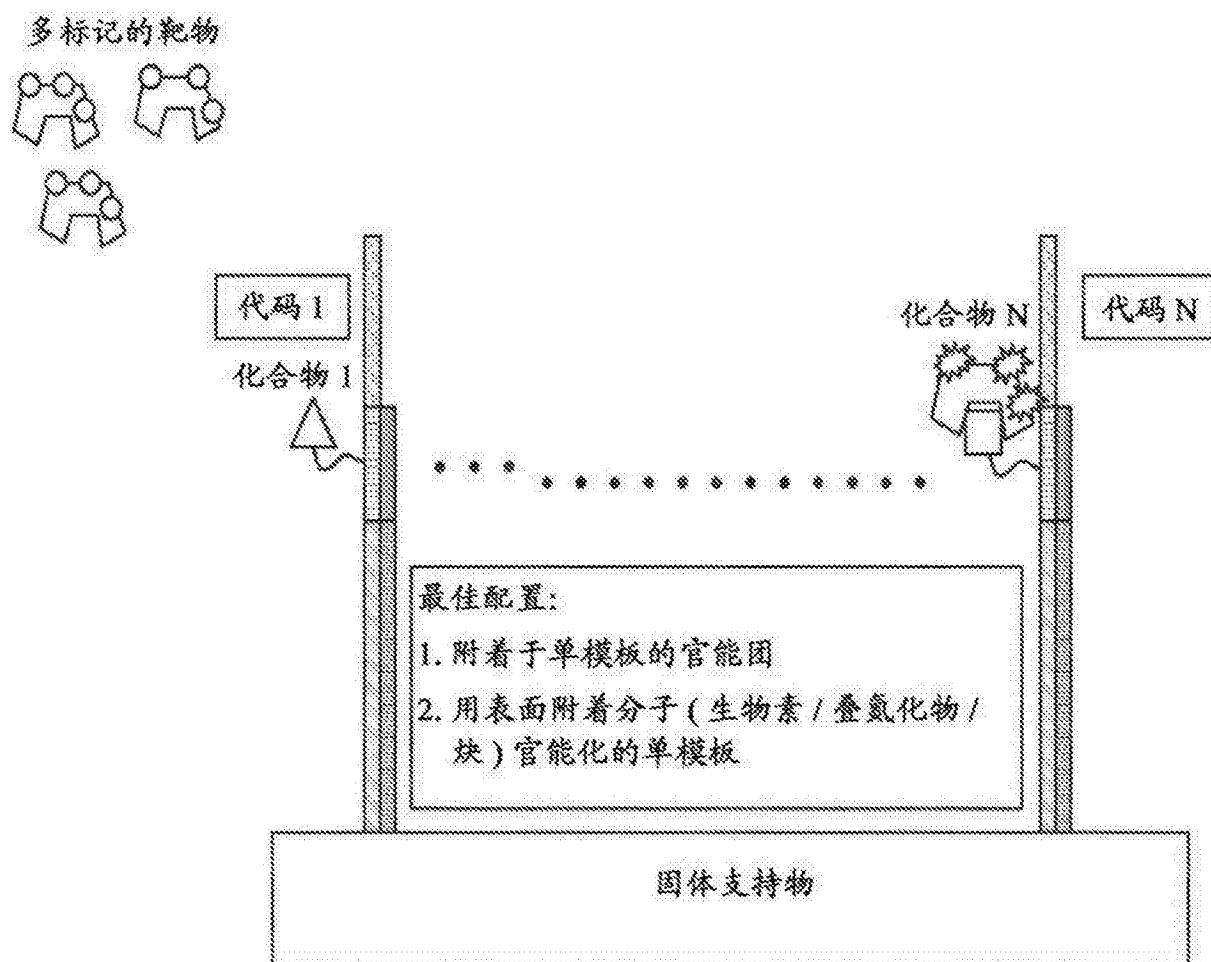


图1E

步骤 6: 对编码分子测序以鉴定“前导物 (leads)”或“命中”。

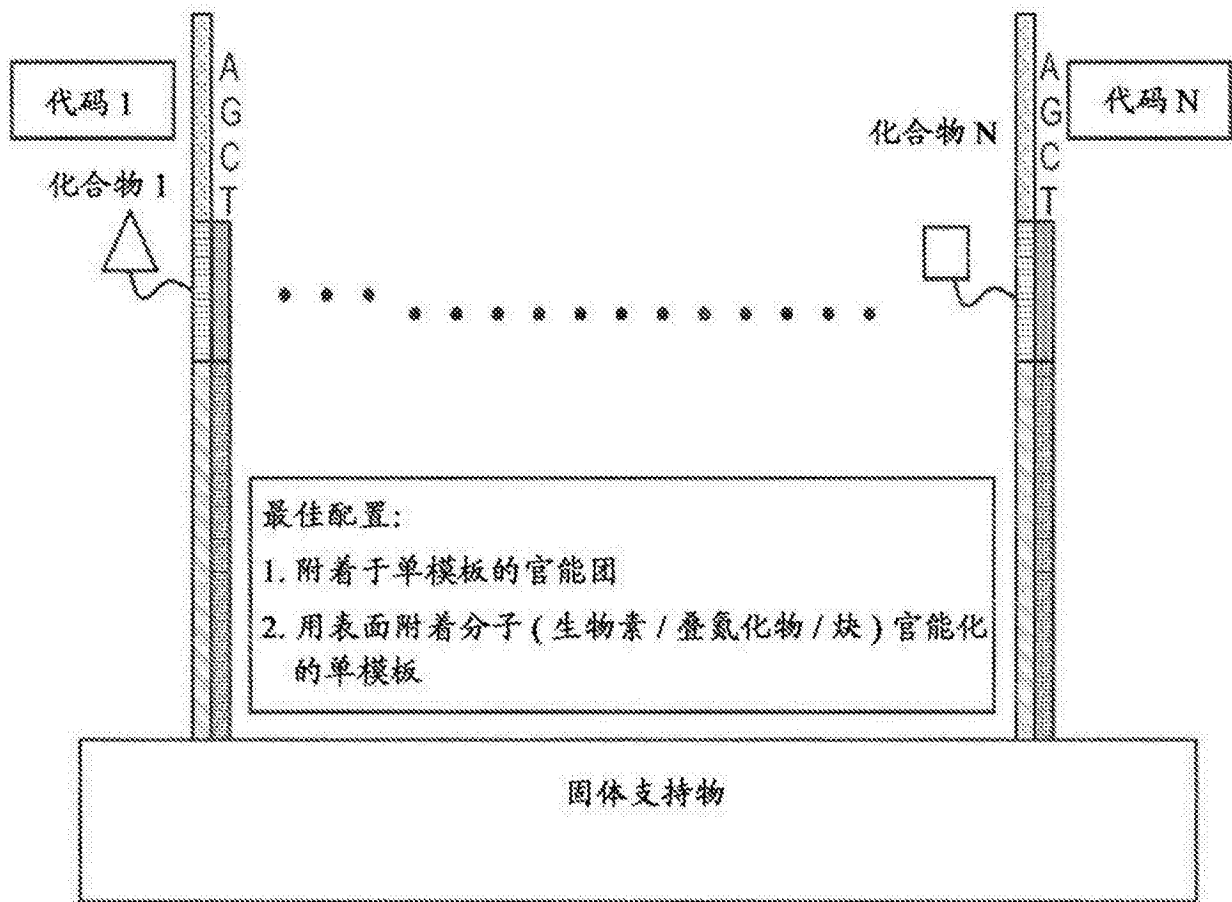


图1F

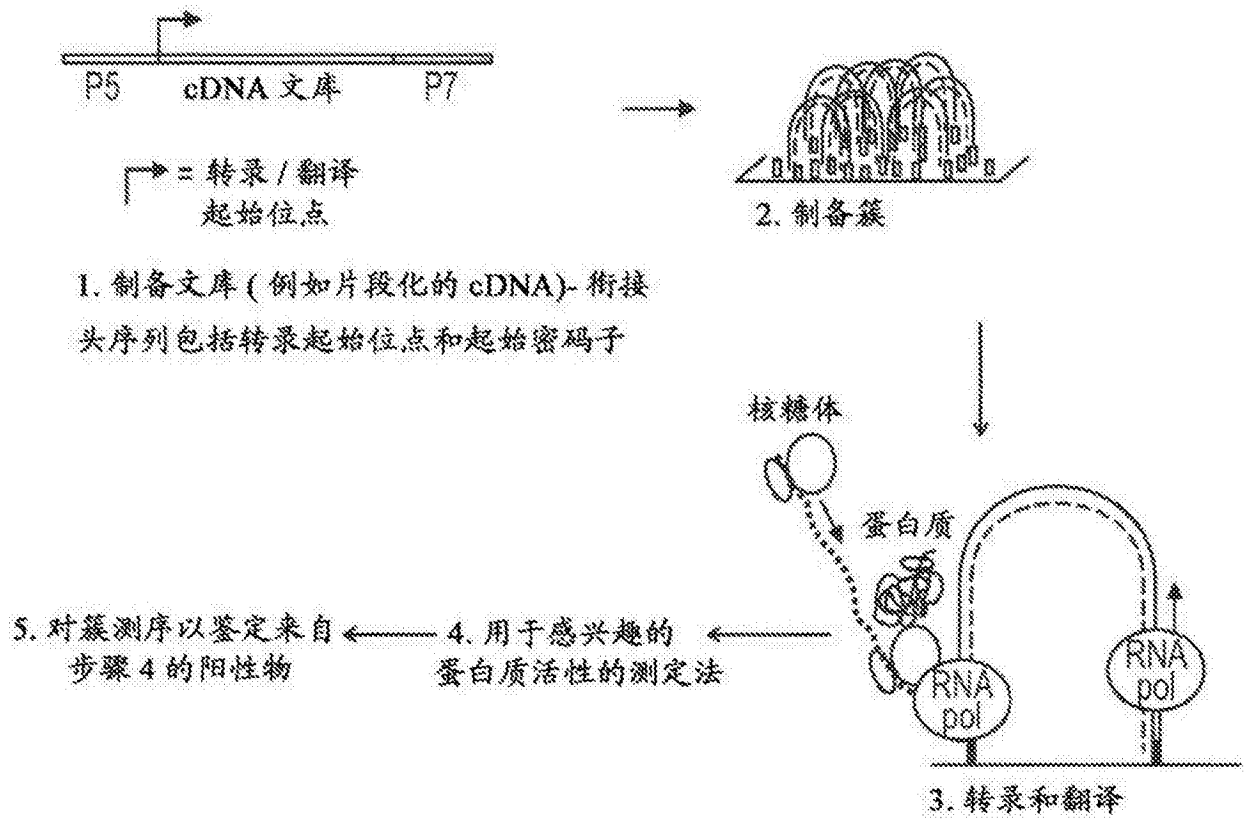
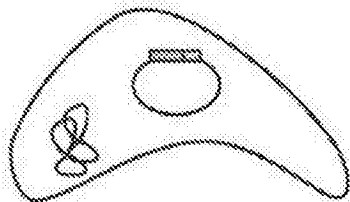


图2

1. 创建具有 mRNA 转录物 + 索引的细菌文库
(索引: $10\text{bp} = 10^6$ 个代码; $20\text{bp} = 10^{12}$ 个代码)



2. 裂解细菌 - 收集 mRNA 转录物



3. 1/2 样品 → 测序确定突变的索引

$$\text{突变 } Y_1 = \text{代码 } X_1$$

$$\text{突变 } Y_2 = \text{代码 } X_2$$

$$\text{突变 } Y_N = \text{代码 } X_N$$

4. 1/2 样品 → 在流动池上翻译 / 测定法

图3A

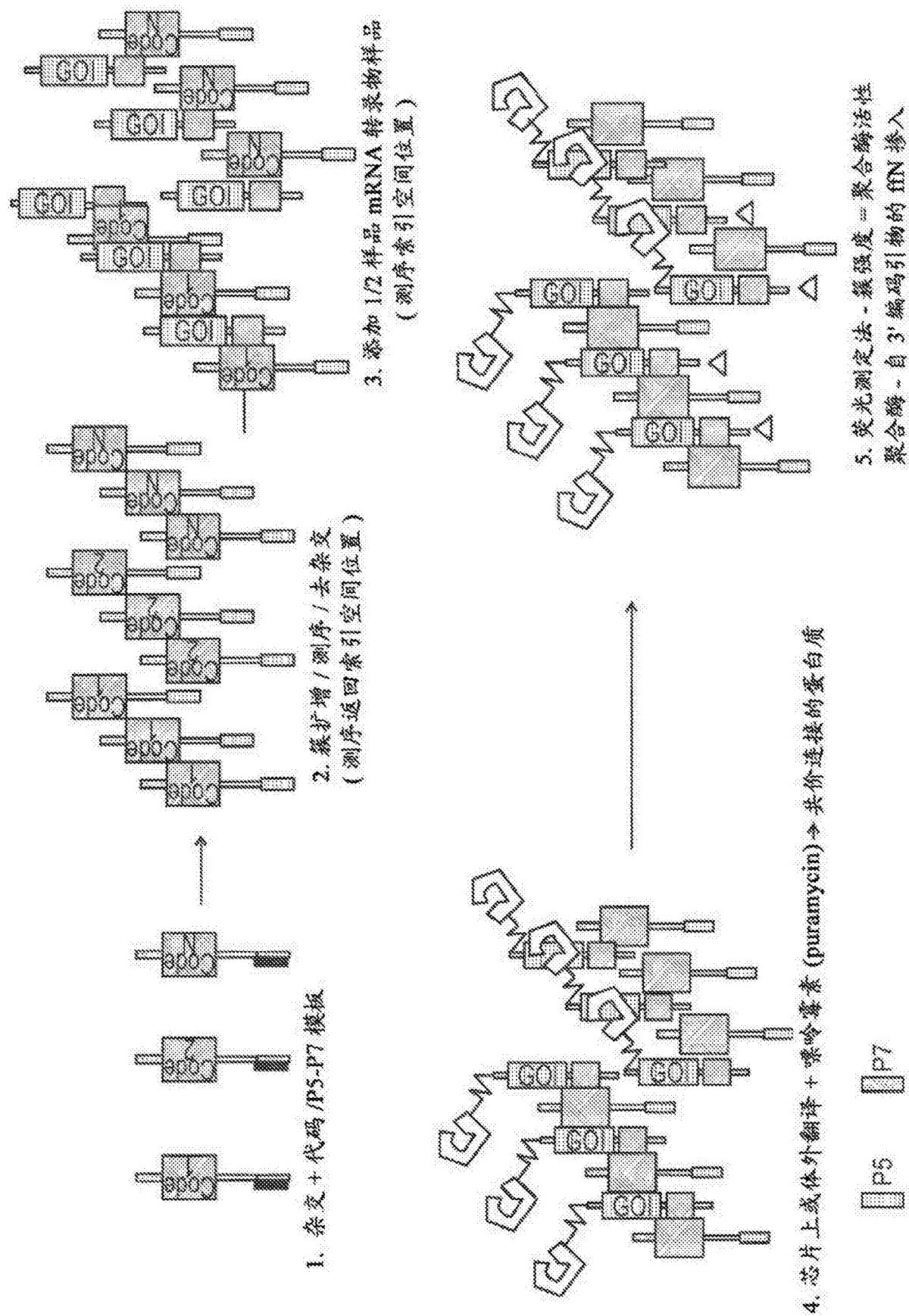


图3B

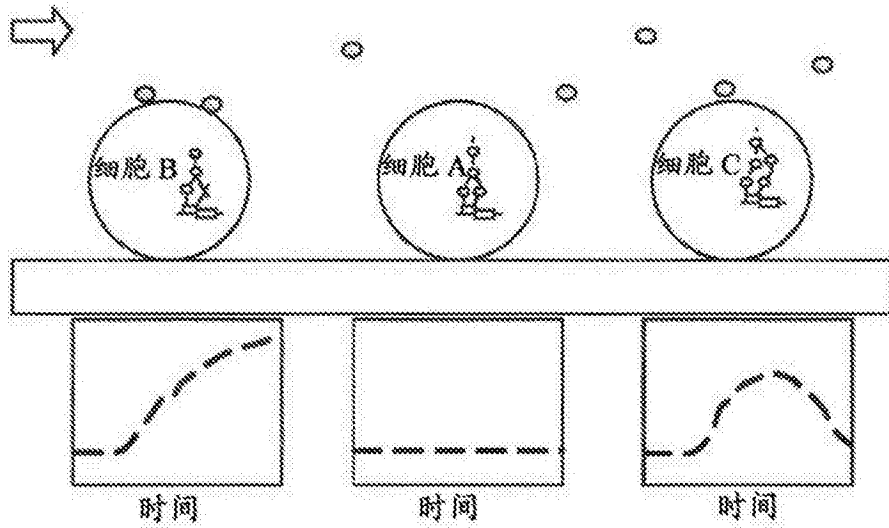


图4

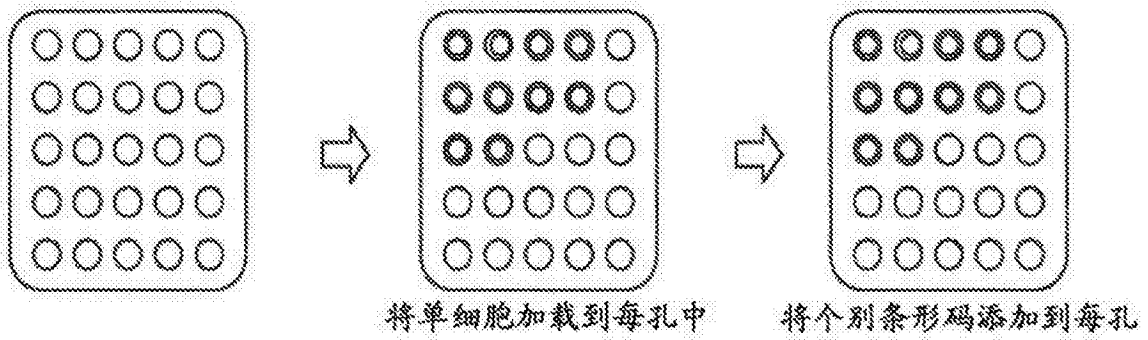


图5

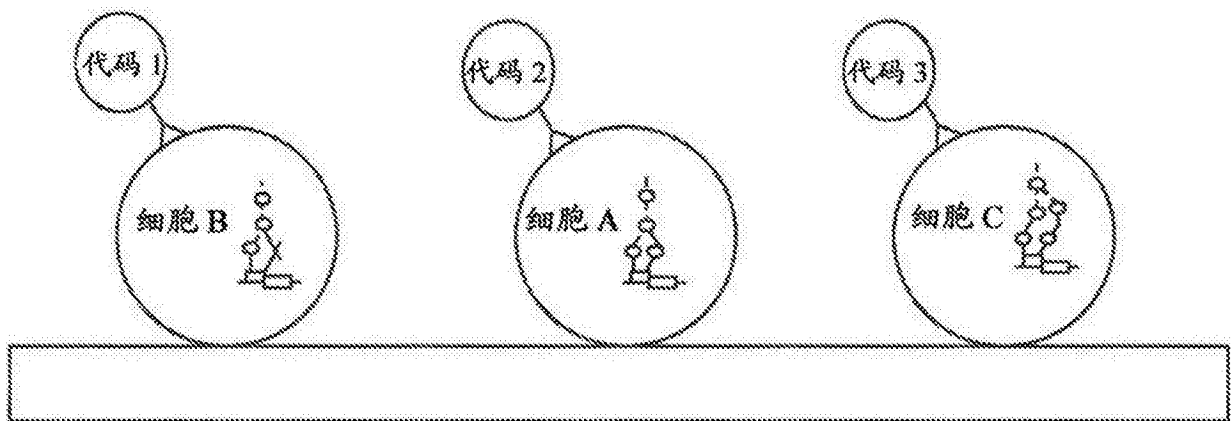


图6

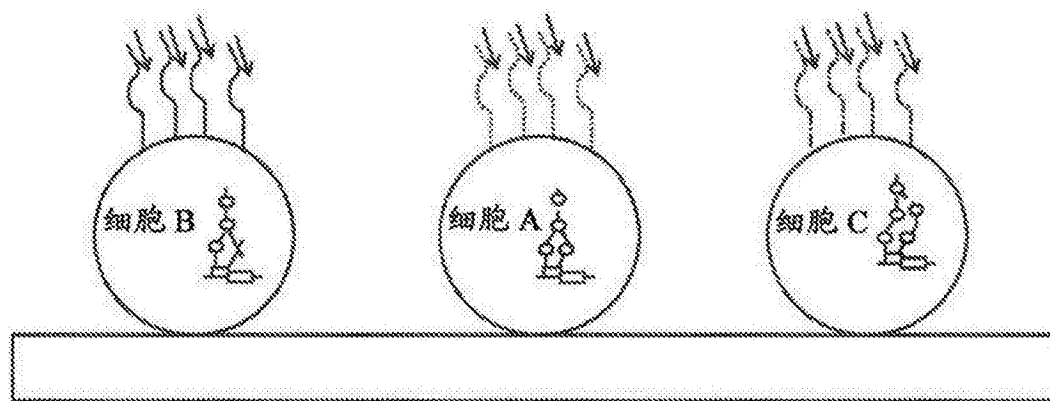


图7

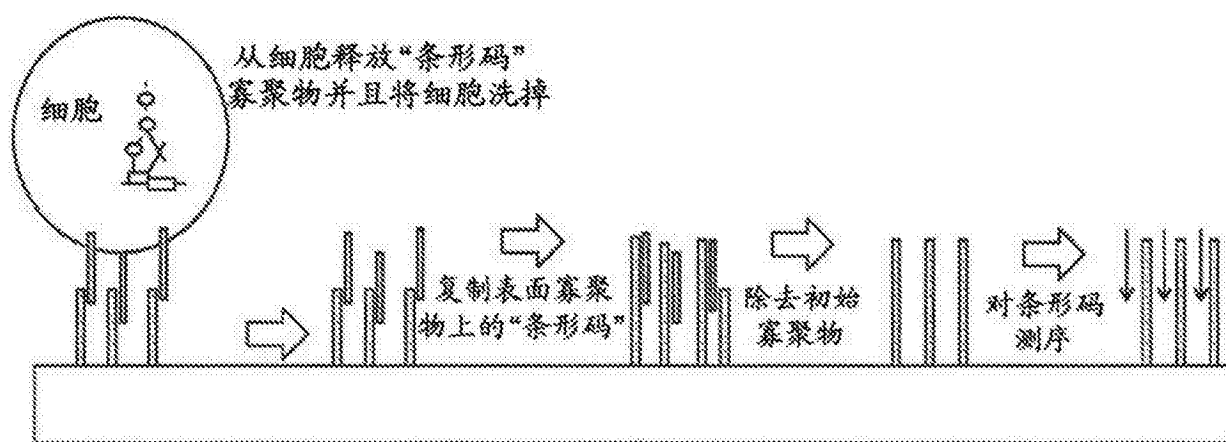


图8

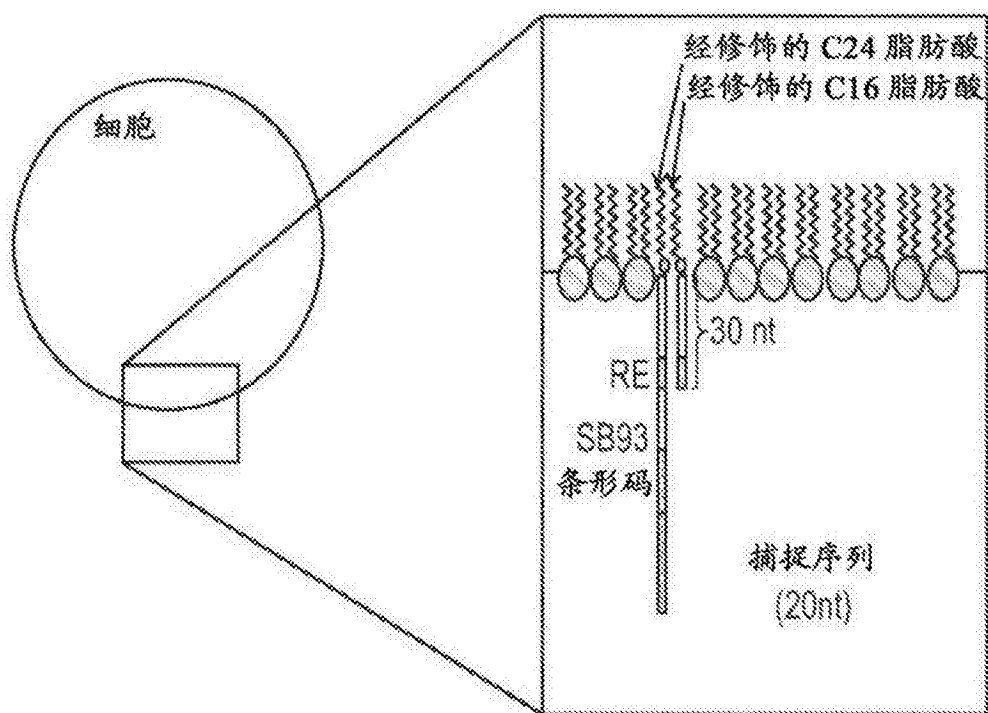


图9

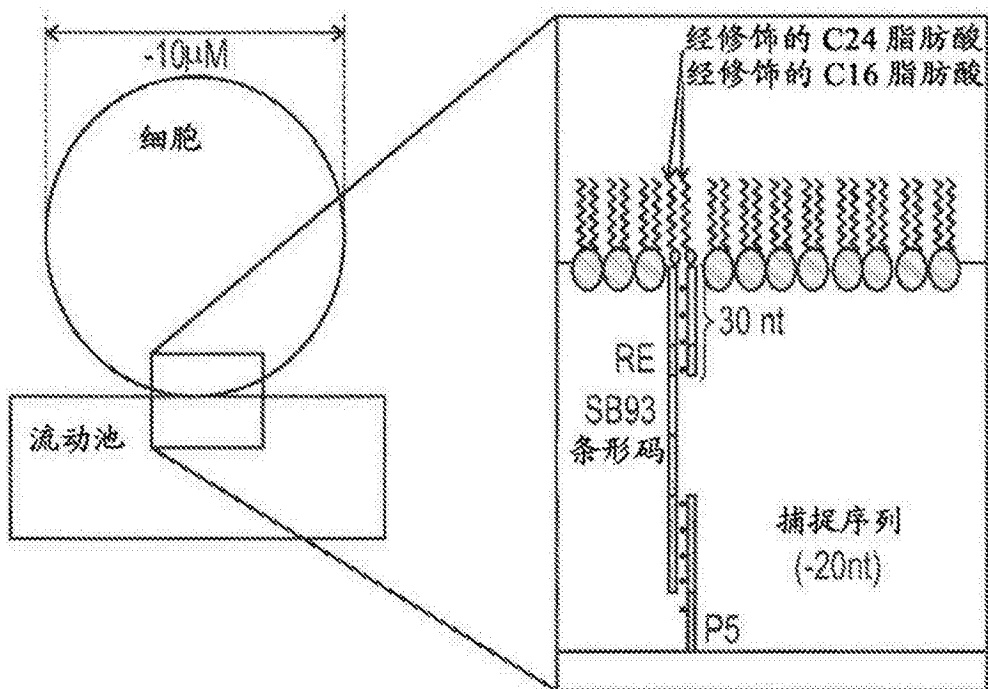


图10

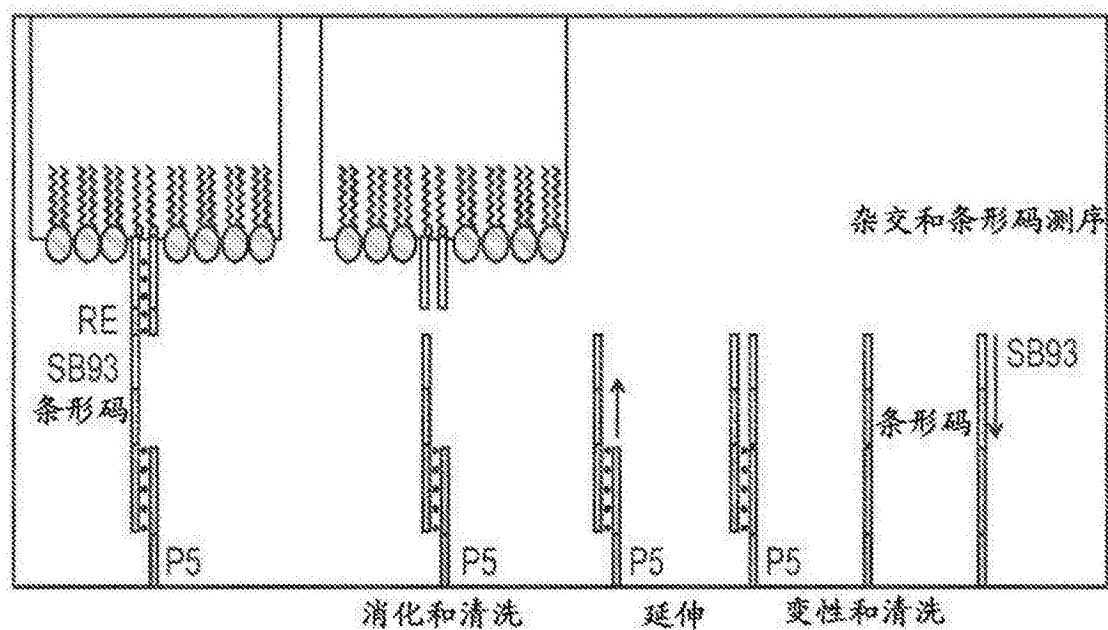


图11

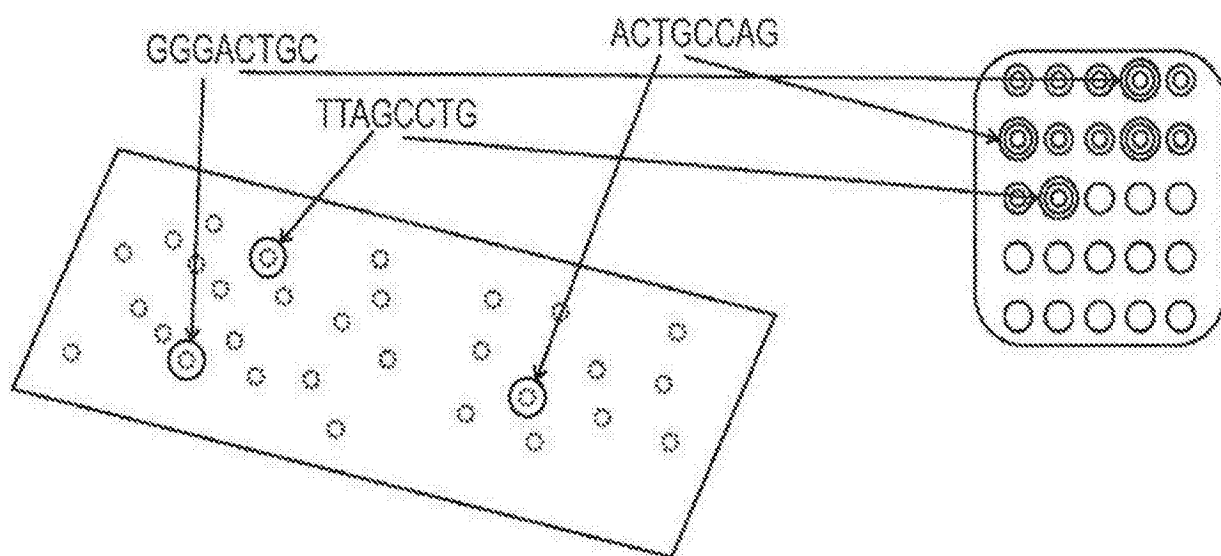


图12

专利名称(译)	用于发现和分析治疗剂的平台		
公开(公告)号	CN107810415A	公开(公告)日	2018-03-16
申请号	CN201680034428.5	申请日	2016-05-09
[标]申请(专利权)人(译)	亿明达股份有限公司		
申请(专利权)人(译)	亿明达股份有限公司		
当前申请(专利权)人(译)	亿明达股份有限公司		
[标]发明人	M 赫 M 普雷维特 M 戈利恩斯基 M W 克林杰 S 佩萨乔维奇 J M 鲍特尔		
发明人	M.赫 M.普雷维特 M.戈利恩斯基 M.W.克林杰 S.佩萨乔维奇 J.M.鲍特尔		
IPC分类号	G01N33/53 G01N33/58 G01N33/68 C12Q1/68 C40B20/04 C40B50/16 B01J19/00 B01L3/00		
CPC分类号	B01J19/0046 B01J2219/005 B01J2219/00626 B01J2219/00743 B01L3/5085 C12Q1/6869 C40B20/04 C40B50/16 G01N33/53 G01N33/58 G01N33/68 B01J2219/00317 B01J2219/00459 B01J2219/00527 B01J2219/00547 B01J2219/00572 B01L3/54 B01L2300/021 C12N15/1062 C12N15/1065 C12N15/1068 C12N15/1075 C12Q2563/131 C12Q2563/179 C12Q2565/514 C12Q2565/515		
代理人(译)	张文辉		
优先权	62/159710 2015-05-11 US		
外部链接	Espacenet SIPO		

摘要(译)

表征候选试剂的方法，其包括以下步骤：(a)提供附着于核酸标签的候选试剂文库；(b)使所述文库与固体支持物接触以将所述候选试剂附着于所述固体支持物，由此形成候选试剂阵列；(c)使所述阵列与筛选剂接触，其中所述阵列中的一种或多种候选试剂与所述筛选剂反应；(d)检测阵列以测定阵列中的至少一种候选试剂与筛选剂反应；(e)对核酸标签测序以测定附着于阵列中候选试剂的标签序列；并且(f)基于附着于所述至少一种候选试剂的所述标签序列，鉴定所述阵列中与所述筛选剂反应的所述至少一种候选试剂。

步骤1: 合成 500,000 个 29 聚体寡聚物 + 附着基团 (-SH, 胺; NHS).

1. 合成仪 = 50,000 个寡聚物 / 天 10 天内 500,000 个寡聚物
2. 条形码 = 19 聚体; “引物” = 10 聚体

