



(12) 发明专利申请

(10) 申请公布号 CN 112673428 A

(43) 申请公布日 2021.04.16

(21) 申请号 202080003506.1

阿塔万·卡鲁纳卡拉恩

(22) 申请日 2020.05.26

阿尔文·伊马迪 莫瑞克·萧

(30) 优先权数据

62/855,690 2019.05.31 US

62/855,616 2019.05.31 US

(74) 专利代理机构 北京安信方达知识产权代理有限公司 11262

代理人 刘晓杰 郑霞

(85) PCT国际申请进入国家阶段日

2020.12.22

(51) Int.Cl.

G11C 13/04 (2006.01)

(86) PCT国际申请的申请数据

PCT/US2020/034515 2020.05.26

(87) PCT国际申请的公布数据

W02020/243075 EN 2020.12.03

(71) 申请人 伊鲁米那股份有限公司

地址 美国加利福尼亚州

(72) 发明人 克雷格·切希拉 阿里·阿加

斯坦利·洪 塔伦·库拉纳

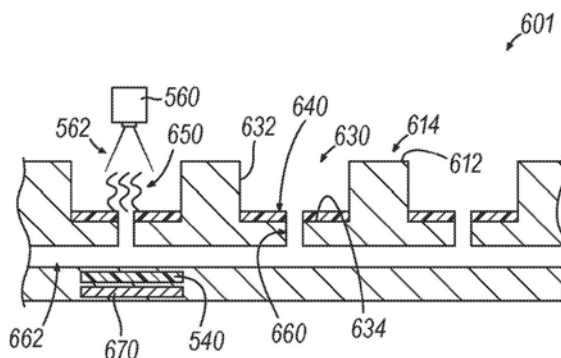
权利要求书9页 说明书44页 附图16页

(54) 发明名称

存储设备、系统和方法

(57) 摘要

一种系统将输入数据作为机器写入的多核苷酸写入存储设备；并且从存储设备读取机器写入的多核苷酸作为输出数据。该存储设备包括流通池，该流通池包括多个存储阱，机器写入的多核苷酸可以存储在该存储阱中。存储设备可以包括对应于存储阱的一组电极，其允许与流通池表面上的阱选择性相互作用。存储设备的操作可以包括接收与存储设备中特定位置相关联的读取请求，创建位于存储设备中特定位置的核苷酸序列的拷贝，将核苷酸序列的拷贝转移到读取位置，并在读取位置读取核苷酸序列的拷贝。



1. 一种用于非易失性存储的系统,其包括:  
处理器和存储器;  
模块接口;  
可移除存储设备,其用于接收在所述模块接口中,所述可移除存储设备包括:  
壳体,其适于与所述模块接口耦合,并将所述可移除存储设备固定在静止位置,以及  
流通池,其位于所述壳体内,所述流通池包括具有从所述流通池的第一表面可进入的  
开口侧的多个阱,其中所述阱适于包含存储机器写入数据的多核苷酸;  
流控设备,其用于向所述可移除存储设备提供一种或更多种试剂;和  
测序设备,其用于分析所述可移除存储设备内的多核苷酸并确定所述多核苷酸内的核  
苷酸;  
其中所述处理器用于执行数据读取过程以:  
确定指示所述多个阱中的阱中的多核苷酸的核苷酸的一组测序数据,以及  
基于所述一组测序数据确定一组机器写入的二进制数据。
2. 根据权利要求1所述的系统,其中,当执行所述数据读取过程时,所述处理器:  
识别所述多个阱中的要从中读取数据的阱,  
操作所述流控设备和所述测序设备对所述阱中的多核苷酸进行合成测序,  
在所述合成测序过程中从所述测序设备中捕获所述一组测序数据,以及  
基于编码方案将所述一组测序数据转换成所述一组二进制数据。
3. 根据权利要求2所述的系统,其中所述编码方案包括一组编码规则,所述一组编码规  
则可由所述处理器用于在具有基数为2的二进制格式和具有大于2的基数的多核苷酸格式  
之间转换数据。
4. 根据权利要求3所述的系统,其中所述多核苷酸格式的基数是基数4。
5. 根据权利要求1所述的系统,其中所述流通池被密封在所述壳体内,所述可移除存储  
设备还包括:  
测序接口,其定位在所述壳体上,并且当所述可移除存储设备处于所述静止位置时,将  
光从所述测序设备传输到所述阱,以及  
流控接口,其定位在所述壳体上,并且当所述可移除存储设备处于所述静止位置时,将  
流体从所述流控设备传输到所述阱。
6. 根据权利要求5所述的系统,其中所述测序接口包括玻璃基底,并且其中所述流控接  
口包括被动歧管。
7. 根据权利要求1所述的系统,还包括在所述可移除存储设备内提供核苷酸和创建多  
核苷酸的合成设备,其中所述处理器还在执行数据写入过程时:  
接收一组输入数据,  
基于所述一组输入数据确定一组输入测序数据,以及  
基于所述一组输入测序数据,操作所述合成设备以在所述多个阱的目标阱中产生输入  
多核苷酸。
8. 根据权利要求7所述的系统,其中所述处理器还:  
存储描述所述多个阱的每个阱的内容的一组索引数据,  
基于所述一组索引数据识别所述多个阱中的要从中读取数据的阱,

基于所述一组索引数据确定所述目标阱,以及  
在所述目标阱中产生所述输入多核苷酸后,更新所述一组索引数据。

9. 根据权利要求8所述的系统,所述可移除存储设备还包括设备存储器,其中所述处理器还将所述可移除存储设备的一组索引数据存储在该设备存储器上。

10. 根据权利要求9所述的系统,其中所述设备存储器用于无线地接收所述一组索引数据。

11. 根据权利要求1所述的系统,所述可移除存储设备还包括位于所述壳体上的电接口,并且当所述可移除存储设备处于所述静止位置时,所述电接口与所述测序设备交换电信号。

12. 根据权利要求11所述的系统,所述可移除存储设备还包括位于所述流通池的第二表面上的集成电路,其中所述电接口用于提供电力和指令以操作所述集成电路,其中所述第二表面与所述第一表面相对。

13. 根据权利要求12所述的系统,其中所述集成电路用于基于从所述测序设备接收的信号选择性地:

将光发射到所述多个阱中的每个阱中,

检测从与所述多个阱的每个阱中的核苷酸相关的标记发出的光的荧光,以及

向所述测序设备提供一组荧光数据,其中所述一组荧光数据可用于确定所述一组测序数据。

14. 根据权利要求1所述的系统,所述系统还包括调节设备,其中所述处理器还:

接收指示所述可移除存储设备已经与所述模块接口耦合的信号,

操作所述流控设备以向所述流通池提供制备流体,其中所述制备流体适于去除所述流通池上的防腐涂层,以使所述流通池能够进行读取和写入操作,以及

操作所述调节设备以使所述流通池达到预定温度。

15. 根据权利要求1所述的系统,所述系统还包括调节设备,其中所述调节设备用于实现所述多个阱中的多核苷酸的保存。

16. 根据权利要求15所述的系统,其中所述调节设备用于冷冻干燥所述流通池以保存所述多个阱中的多核苷酸。

17. 根据权利要求15所述的系统,其中所述调节设备包括所述流控设备的特征,所述特征将防腐流体分配到所述流通池,从而在所述流通池上施加防腐涂层,以保存和保护所述多个阱中的多核苷酸。

18. 根据权利要求15所述的系统,其中所述处理器还:

接收指示将从所述模块接口卸载所述可移除存储设备的信号,以及

响应于指示将从所述模块接口卸载所述可移除存储设备的信号,激活所述调节设备。

19. 根据权利要求1所述的系统,其中所述处理器还:

接收指示所述可移除存储设备已经从所述模块接口卸载的信号,以及

响应于指示所述可移除存储设备已经从所述模块接口卸载的信号,向用户提供所述可移除存储设备被卸载的通知。

20. 一种用于非易失性存储的方法,包括:

将可移除存储设备放置在存储仪器的模块接口中,所述可移除存储设备包括流通池,

其中所述模块接口适于将所述可移除存储设备固定在静止位置；

操作测序设备对所述流通池的多个阱中的阱内的多核苷酸进行合成测序，以确定描述所述多核苷酸的核苷酸的一组测序数据；

操作流控设备，以向所述流通池提供一种或更多种流体类型，用于利用所述测序设备进行合成测序；和

基于所述一组测序数据和编码方案来确定一组二进制数据，其中所述编码方案描述了将信息从二进制格式转换为与所述一组测序数据相关联的非二进制格式的一组规则。

21. 根据权利要求20所述的方法，还包括：

接收一组输入数据；

基于所述一组输入数据和所述编码方案确定一组输入测序数据；和

基于所述一组输入测序数据，操作合成设备以在所述多个阱的目标阱中产生输入多核苷酸。

22. 根据权利要求20所述的方法，还包括接收指示所述可移除存储设备已经与所述模块接口耦合的信号。

23. 根据权利要求22所述的方法，还包括响应于指示所述可移除存储设备已经与所述模块接口耦合的信号，操作所述流控设备以向所述流通池提供制备流体，其中所述制备流体适于移除所述流通池上的防腐涂层，以使所述流通池能够进行读取和写入操作。

24. 根据权利要求22所述的方法，还包括响应于指示所述可移除存储设备已经与所述模块接口耦合的信号来操作调节设备，其中所述调节设备使所述流通池达到读取和写入操作所需的预定温度。

25. 根据权利要求20所述的方法，还包括接收指示所述可移除存储设备将从所述模块接口卸载的信号。

26. 根据权利要求25所述的方法，还包括响应于指示所述可移除存储设备将从所述模块接口卸载的信号，激活调节设备。

27. 根据权利要求26所述的方法，其中所激活的调节设备向所述流通池分配防腐流体，从而在所述流通池上施加防腐涂层，以保存和保护所述多个阱中的多核苷酸。

28. 根据权利要求26所述的方法，其中所激活的调节设备冷冻干燥所述流通池以保存所述多个阱中的多核苷酸。

29. 根据权利要求20所述的方法，还包括：

接收指示所述可移除存储设备已经从所述模块接口卸载的信号；和

响应于指示所述可移除存储设备已经从所述模块接口卸载的信号，向用户提供所述可移除存储设备被卸载的通知。

30. 一种用于非易失性存储的可移除存储设备，包括：

壳体，其适于与存储仪器的模块接口耦合，并将所述可移除存储设备固定在静止位置；

流通池，其密封在所述壳体内，所述流通池包括具有可从所述流通池的第一表面进入的开口侧的多个阱，所述阱适于包含存储机器写入数据的多核苷酸；

测序接口，其定位在所述流通池附近，并且适于当所述可移除存储设备处于静止位置时将光传输到所述阱；和

流控接口，其位于所述壳体上，并且适于当所述可移除存储设备处于所述静止位置时，

将流体从所述存储仪器的流控设备传输到所述阱。

31. 根据权利要求30所述的可移除存储设备,还包括电接口,所述电接口位于所述壳体上,以在所述可移除存储设备处于所述静止位置时与所述存储仪器交换电信号。

32. 根据权利要求31所述的可移除存储设备,还包括位于所述流通池的第二表面上的集成电路,所述电接口用于提供电力和指令以操作集成电路。

33. 根据权利要求32所述的可移除存储设备,其中所述第二表面与所述第一表面相对,并且其中所述集成电路基于从所述存储仪器接收的信号选择性地:

经由所述测序接口将光发射到所述多个阱中的每个阱中,

检测从与所述多个阱的每个阱中的核苷酸相关的标记发出的光的荧光,以及

向所述存储仪器提供一组荧光数据,其中所述一组荧光数据可用于确定一组测序数据,所述一组测序数据描述所述多个阱的每个阱中的多核苷酸的核苷酸。

34. 根据权利要求30所述的可移除存储设备,其中,所述可移除存储设备包括:

电源连接,其用于将所述可移除存储设备与电源耦合,以及

一组传感器,其用于检测所述流通池的一个或更多个特性,并在检测到的特性超过预定阈值时提供指示,其中所述一组传感器包括选自有温度传感器、湿度传感器、光传感器和辐射传感器组成的组的一个或更多个传感器。

35. 根据权利要求30所述的可移除存储设备,还包括至少一个传感器,所述至少一个传感器用于指示一个或更多个环境条件是否落在预定范围之外,所述一个或更多个环境条件包括湿度、温度、光或辐射中的一个或更多个。

36. 根据权利要求30所述的可移除存储设备,所述流通池包括具有穿过所述阱的底部区域形成的多个开口的基底。

37. 根据权利要求36所述的可移除存储设备,所述流通池还包括在穿过所述阱的底部区域形成的开口中的导电材料。

38. 根据权利要求37所述的可移除存储设备,所述导电材料包括氧化铟锡。

39. 根据权利要求37所述的可移除存储设备,还包括在穿过所述阱的底部区域形成的开口中的透明各向异性材料。

40. 根据权利要求30所述的可移除存储设备,所述流通池包括下侧,所述下侧具有一个或更多个焊盘以接触位于所述流通池下方的集成电路芯片。

41. 根据权利要求30所述的可移除存储设备,所述流通池包括玻璃。

42. 根据权利要求30所述的可移除存储设备,还包括薄膜介质,以将多核苷酸转移到所述流通池中的相应测序表面。

43. 根据权利要求42所述的可移除存储设备,其中所述多核苷酸被电泳地转移到所述流通池中的相应测序表面。

44. 根据权利要求42所述的可移除存储设备,其中所述多核苷酸通过印迹转移到所述流通池中的相应测序表面。

45. 一种方法,包括:

接收读取请求,其中所述读取请求与存储设备中的特定位置相关联;

创建位于所述存储设备中的所述特定位置的多核苷酸的拷贝;

将所述多核苷酸的拷贝转移到读取位置;和

在所述读取位置读取所述多核苷酸的拷贝。

46. 根据权利要求45所述的方法, 其中:

所述存储设备包括一组一个或更多个写入位置;

所述特定位置包括在所述一组一个或更多个写入位置中;

所述读取位置不包括在所述一组一个或更多个写入位置中; 和

每当所述存储设备接收到写入请求时, 所述存储设备只将数据写到由所述一组一个或更多个写入位置所包括的位置。

47. 根据权利要求46所述的方法, 其中:

所述存储设备包括流通池;

来自所述一个或更多个写入位置的每个写入位置设置在所述流通池的第一表面上; 和

所述读取位置设置在所述流通池的第二表面上。

48. 根据权利要求47所述的方法, 其中所述第一表面和所述第二表面彼此相对。

49. 根据权利要求47所述的方法, 其中:

所述第一表面是所述流通池的底表面; 和

所述第二表面是所述流通池的顶表面。

50. 根据权利要求46所述的方法, 其中来自所述一组一个或更多个写入位置的每个写入位置具有对应的一组电极, 以使用该写入位置中的核苷酸来编码数据。

51. 根据权利要求50所述的方法, 其中:

所述存储设备包括专用集成电路; 和

所述专用集成电路对于来自所述一组一个或更多个写入位置的每个写入位置, 驱动对应于该位置的一组电极, 从而使用该写入位置中的核苷酸编码数据。

52. 根据权利要求51所述的方法, 其中所述存储设备包括第二集成电路, 其中所述第二集成电路用于从所述读取位置处的核苷酸读取数据。

53. 根据权利要求52所述的方法, 其中所述专用集成电路和所述第二集成电路都是互补金属氧化物半导体 (CMOS) 芯片。

54. 根据权利要求46所述的方法, 其中:

来自所述一组一个或更多个写入位置的每个写入位置是可寻址的; 和

除了所述一组一个或更多个写入位置之外, 所述存储设备不包括任何可寻址位置来以多核苷酸的形式存储数据。

55. 根据权利要求46所述的方法, 其中:

所述存储设备存储一组一个或更多个逻辑数据分组;

所述方法包括创建索引, 所述索引为来自所述一组一个或更多个逻辑数据分组的每个逻辑数据分组指定关联的一组一个或更多个位置; 和

对于来自所述一组一个或更多个逻辑数据分组的每个逻辑数据分组, 来自该逻辑数据分组的关联的一组一个或更多个位置的每个位置是由所述一组一个或更多个写入位置包括的写入位置。

56. 根据权利要求55所述的方法, 其中所述方法包括:

接收从所述一组一个或更多个逻辑数据分组中检索特定逻辑数据分组的请求;

从所述索引中检索所述特定逻辑数据分组的关联的一组一个或更多个位置; 和

基于所述特定逻辑数据分组的关联的一组一个或多个位置来确定所述存储设备中的所述特定位置。

57. 根据权利要求55所述的方法,其中,对于来自所述一组一个或多个逻辑数据分组的每个逻辑数据分组,该逻辑数据分组的关联的一组一个或多个位置包括开始位置和结束位置。

58. 根据权利要求55所述的方法,其中,对于来自所述一组一个或多个逻辑数据分组的每个逻辑数据分组,来自该逻辑数据分组的关联的一组一个或多个位置的每个位置是二维网格中的位置。

59. 根据权利要求55所述的方法,其中所述方法包括:

将所述索引存储在非核苷酸存储器中;和

使用珠上的核苷酸存储所述索引。

60. 根据权利要求45所述的方法,其中将所述多核苷酸的拷贝转移到读取位置包括:

将所述多核苷酸的拷贝固定到所述特定位置的珠;和

将所述珠从所述特定位置转移到所述读取位置。

61. 根据权利要求45所述的方法,其中所述特定位置在珠的表面上。

62. 根据权利要求45所述的方法,其中使用选自以下项组成的组的技术将所述多核苷酸的拷贝转移到所述读取位置:

电泳;

介电泳;和

流体层流流动。

63. 一种系统,包括一个或多个存储指令的非暂时性计算机可读介质,以执行一种方法,所述方法包括:

接收读取请求,其中所述读取请求与存储设备中的特定位置相关联;

创建位于所述存储设备中的所述特定位置的多核苷酸的拷贝;

将所述多核苷酸的拷贝转移到读取位置;和

在所述读取位置读取所述多核苷酸的拷贝。

64. 根据权利要求63所述的系统,其中:

所述存储设备包括一组一个或多个写入位置;

所述特定位置包括在所述一组一个或多个写入位置中;

所述读取位置不包括在所述一组一个或多个写入位置中;和

每当所述存储设备接收到写入请求时,所述存储设备仅将数据写入由所述一组一个或多个写入位置所包括的位置。

65. 根据权利要求64所述的系统,其中:

所述存储设备包括流通池;

来自所述一个或多个写入位置的每个写入位置设置在所述流通池的第一表面上;和

所述读取位置设置在所述流通池的第二表面上。

66. 根据权利要求65所述的系统,其中所述第一表面和所述第二表面彼此相对。

67. 根据权利要求65所述的系统,其中:

所述第一表面是所述流通池的底表面;和

所述第二表面是所述流通池的顶表面。

68. 根据权利要求64所述的系统,其中来自所述一组一个或更多个写入位置的每个写入位置具有对应的一组电极,以使用该写入位置中的核苷酸来编码数据。

69. 根据权利要求68所述的系统,其中:

所述存储设备包括专用集成电路;和

所述专用集成电路对于来自所述一组一个或更多个写入位置的每个写入位置,驱动对应于该位置的一组电极,从而使用该写入位置中的核苷酸编码数据。

70. 根据权利要求69所述的系统,其中所述存储设备包括第二集成电路,其中所述第二集成电路用于从所述读取位置处的核苷酸读取数据。

71. 根据权利要求70所述的系统,其中所述专用集成电路和所述第二集成电路都是CMOS芯片。

72. 根据权利要求64所述的系统,其中:

来自所述一组一个或更多个写入位置的每个写入位置是可寻址的;和

除了所述一组一个或更多个写入位置之外,所述存储设备不包括任何可寻址位置来以多核苷酸的形式存储数据。

73. 根据权利要求64所述的系统,其中:

所述存储设备存储一组一个或更多个逻辑数据分组;

所述方法包括创建索引,所述索引为来自所述一组一个或更多个逻辑数据分组的每个逻辑数据分组指定关联的一组一个或更多个位置;和

对于来自所述一组一个或更多个逻辑数据分组的每个逻辑数据分组,来自该逻辑数据分组的关联的一组一个或更多个位置的每个位置是由所述一组一个或更多个写入位置包括的写入位置。

74. 根据权利要求73所述的系统,其中所述方法包括:

接收从所述一组一个或更多个逻辑数据分组中检索特定逻辑数据分组的请求;

从所述索引中检索所述特定逻辑数据分组的关联的一组一个或更多个位置;和

基于所述特定逻辑数据分组的关联的一组一个或更多个位置来确定所述存储设备中的所述特定位置。

75. 根据权利要求73所述的系统,其中,对于来自所述一组一个或更多个逻辑数据分组的每个逻辑数据分组,该逻辑数据分组的关联的一组一个或更多个位置包括开始位置和结束位置。

76. 根据权利要求73所述的系统,其中,对于来自所述一组一个或更多个逻辑数据分组的每个逻辑数据分组,来自该逻辑数据分组的关联的一组一个或更多个位置的每个位置是二维网格中的位置。

77. 根据权利要求73所述的系统,其中所述方法包括:

将所述索引存储在非核苷酸存储器中;和

使用珠上的核苷酸存储所述索引。

78. 根据权利要求63所述的系统,其中将所述多核苷酸的拷贝转移到读取位置包括:

将所述多核苷酸的拷贝固定到所述特定位置的珠;和

将所述珠从所述特定位置转移到所述读取位置。



79. 根据权利要求63所述的系统,其中所述特定位置在珠的表面上。

80. 根据权利要求63所述的系统,其中使用选自以下项组成的组的技术将所述多核苷酸的拷贝转移到所述读取位置:

电泳;

介电泳;和

流体层流流动。

81. 一个或多个存储用于执行方法的指令的非暂时性计算机可读介质,所述方法包括:

接收读取请求,其中所述读取请求与存储设备中的特定位置相关联;

创建位于所述存储设备中的所述特定位置的多核苷酸的拷贝;

将所述多核苷酸的拷贝转移到读取位置;和

在所述读取位置读取所述多核苷酸的拷贝。

82. 根据权利要求81所述的计算机可读介质,其中:

所述存储设备包括一组一个或多个写入位置;

所述特定位置包括在所述一组一个或多个写入位置中;

所述读取位置不包括在所述一组一个或多个写入位置中;和

每当所述存储设备接收到写入请求时,所述存储设备只将数据写到由所述一组一个或多个写入位置所包括的位置。

83. 根据权利要求82所述的计算机可读介质,其中:

所述存储设备包括流通池;

来自所述一个或多个写入位置的每个写入位置设置在所述流通池的第一表面上;和

所述读取位置设置在所述流通池的第二表面上。

84. 根据权利要求83所述的计算机可读介质,其中所述第一表面和所述第二表面彼此相对。

85. 根据权利要求83所述的计算机可读介质,其中:

所述第一表面是所述流通池的底表面;和

所述第二表面是所述流通池的顶表面。

86. 根据权利要求82所述的计算机可读介质,其中,来自所述一组一个或多个写入位置的每个写入位置具有对应的一组电极,以使用该写入位置中的核苷酸来编码数据。

87. 根据权利要求86所述的计算机可读介质,其中:

所述存储设备包括专用集成电路;和

所述专用集成电路对于来自所述一组一个或多个写入位置的每个写入位置,驱动对应于该位置的一组电极,从而使用该写入位置中的核苷酸编码数据。

88. 根据权利要求87所述的计算机可读介质,其中所述存储设备包括第二集成电路,其中所述第二集成电路用于从所述读取位置处的核苷酸读取数据。

89. 根据权利要求88所述的计算机可读介质,其中所述专用集成电路和所述第二集成电路都是CMOS芯片。

90. 根据权利要求82所述的计算机可读介质,其中:

来自所述一组一个或多个写入位置的每个写入位置是可寻址的;和

除了所述一组一个或更多个写入位置之外,所述存储设备不包括任何可寻址位置来以多核苷酸的形式存储数据。

91. 根据权利要求82所述的计算机可读介质,其中:

所述存储设备存储一组一个或更多个逻辑数据分组;

所述方法包括创建索引,所述索引为来自所述一组一个或更多个逻辑数据分组的每个逻辑数据分组指定关联的一组一个或更多个位置;和

对于来自所述一组一个或更多个逻辑数据分组的每个逻辑数据分组,来自该逻辑数据分组的关联的一组一个或更多个位置的每个位置是由所述一组一个或更多个写入位置包括的写入位置。

92. 根据权利要求91所述的计算机可读介质,其中所述方法包括:

接收从所述一组一个或更多个逻辑数据分组中检索特定逻辑数据分组的请求;

从所述索引中检索所述特定逻辑数据分组的关联的一组一个或更多个位置;和

基于所述特定逻辑数据分组的关联的一组一个或更多个位置来确定所述存储设备中的所述特定位置。

93. 根据权利要求91所述的计算机可读介质,其中,对于来自所述一组一个或更多个逻辑数据分组的每个逻辑数据分组,该逻辑数据分组的关联的一组一个或更多个位置包括开始位置和结束位置。

94. 根据权利要求91所述的计算机可读介质,其中,对于来自所述一组一个或更多个逻辑数据分组的每个逻辑数据分组,来自该逻辑数据分组的关联的一组一个或更多个位置的每个位置是二维网格中的位置。

95. 根据权利要求91所述的计算机可读介质,其中所述方法包括:

将所述索引存储在非核苷酸存储器中;和

使用珠上的核苷酸存储所述索引。

96. 根据权利要求81所述的计算机可读介质,其中将所述多核苷酸的拷贝转移到读取位置包括:

将所述多核苷酸的拷贝固定到所述特定位置的珠;和

将所述珠从所述特定位置转移到所述读取位置。

97. 根据权利要求81所述的计算机可读介质,其中所述特定位置在珠的表面上。

98. 根据权利要求81所述的计算机可读介质,其中使用选自由以下项组成的组的技术将所述多核苷酸的拷贝转移到所述读取位置:

电泳;

介电泳;和

流体层流流动。

## 存储设备、系统和方法

[0001] 相关申请

[0002] 本申请要求于2019年5月31日提交的题为“Non-Volatile Polynucleotide Storage Device, System, and Method”的第62/855,616号美国临时专利申请的优先权,其全部内容通过引用结合于此。本申请还要求于2019年5月31日提交的题为“DNA Storage Device with Separate Reading and Writing Locations”的第62/855,690号美国临时专利申请的优先权,其全部内容通过引用结合于此。

[0003] 背景

[0004] 计算机系统使用各种不同的机制来存储数据,包括磁存储、光存储和固态存储。这种形式的数据存储可能在读写速度、数据保留时间、功耗或数据密度方面存在缺陷。

[0005] 就像天然存在的DNA可以被读取一样,机器写入的DNA也可以被读取。现有的DNA读取技术可以包括基于阵列的循环测序测定(例如合成测序(SBS)),其中通过酶操作的迭代循环对密集阵列的DNA特征(例如模板核酸)进行测序。在每个循环之后,可以捕获图像,并随后与其他图像一起分析,以确定机器写入的DNA特征的序列。在另一生物化学测定中,具有可识别标记(例如,荧光标记)的未知分析物可暴露于已知探针阵列,该探针在该阵列内具有预定地址。观察探针和未知分析物之间发生的化学反应可以帮助识别或揭示分析物的特性。

[0006] 概述

[0007] 本文描述了用于数据的DNA存储的设备、系统和方法,包括用于每阱激活、同时读写高速缓存和用于同时读写能力的多卷管理的特征。

[0008] 一种实现涉及一种用于非易失性存储的系统,包括:处理器和存储器;模块接口;用于接收在模块接口中的可移除存储设备,该可移除存储设备包括:适于与模块接口耦合并将可移除存储设备固定在静止位置的壳体,以及位于壳体内的流通池,该流通池包括多个具有从流通池的第一表面可接近的开口侧的阱,其中阱适于包含存储机器写入数据的多核苷酸;流控设备,用于向可移除存储设备提供一种或更多种试剂;以及测序设备,用于分析可移除存储设备内的多核苷酸并确定核苷酸;其中所述处理器将执行数据读取过程以:确定指示所述多个阱中的阱中的多核苷酸的一组测序数据,并基于该组测序数据确定一组机器写入的二进制数据。

[0009] 上述实现中的任何一种或更多种存在变化,其中当执行数据读取过程时,处理器:识别多个阱中的要从中读取数据的阱,操作流控设备和测序设备以对阱中的多核苷酸执行合成测序,在合成测序期间从测序设备捕获一组测序数据,并基于编码方案将该组测序数据转换成一组二进制数据。

[0010] 上述实现中的任何一种或更多种存在变化,其中编码方案包括一组可由处理器使用的编码规则,以在具有基数为2的二进制格式和具有基数大于2的多核苷酸格式之间转换数据。

[0011] 上述实现中的任何一种或更多种存在变化,其中多核苷酸格式的基数是基数4。

[0012] 上述实现中的任何一种或更多种存在变化,其中流通池密封在壳体内,可移除存

储设备还包括：测序接口，其位于壳体上，当可移除存储设备处于静止位置时，将光从测序设备传输到阱；以及流控接口，其位于壳体上，当可移除存储设备处于静止位置时，将流体从流控设备传输到阱。

[0013] 上述实现中的任何一种或更多种存在变化，其中测序接口包括玻璃基底，并且其中流控接口包括被动歧管。

[0014] 上述实现中的任何一种或更多种存在变化，还包括在可移除存储设备内提供核苷酸和产生多核苷酸的合成设备，其中当执行数据写入过程时，处理器还：接收一组输入数据，基于该组输入数据确定一组输入测序数据，并操作合成设备以基于该组输入测序数据在多个阱的目标阱中产生输入多核苷酸。

[0015] 上述实现中的任何一种或更多种存在变化，其中处理器进一步：存储描述多个阱的每个阱的内容的一组索引数据，基于该组索引数据识别多个阱中的要从中读取数据的阱，基于该组索引数据确定目标阱，并在目标阱中创建输入多核苷酸之后更新该组索引数据。

[0016] 上述实现中的任何一种或更多种存在变化，可移除存储设备还包括设备存储器，其中处理器还将可移除存储设备的一组索引数据存储在该可移除存储设备的设备存储器上。

[0017] 上述实现中的任何一种或更多种存在变化，其中设备存储器用于无线地接收该组索引数据。

[0018] 上述实现中的任何一种或更多种存在变化，可移除存储设备还包括位于壳体上的电接口，当可移除存储设备处于静止位置时，该电接口与测序设备交换电信号。

[0019] 上述实现中的任何一种或更多种存在变化，可移除存储设备还包括位于流通池的第二表面上的集成电路，其中电接口用于提供电力和指令以操作集成电路，其中第二表面与第一表面相对。

[0020] 上述实现中的任何一种或更多种存在变化，其中集成电路基于从测序设备接收的信号选择性地：将光发射到多个阱的每个阱中，检测从与多个阱的每个阱中的核苷酸相关联的标记发射的光的荧光，并向测序设备提供一组荧光数据，其中该组荧光数据可用于确定该组测序数据。

[0021] 上述实现中的任何一种或更多种存在变化，该系统还包括调节设备，其中处理器还：接收指示可移除存储设备已经与模块接口耦合的信号，操作流控设备以向流通池提供制备流体，其中制备流体适于移除流通池上的防腐涂层以使流通池能够进行读取和写入操作，并且操作调节设备以使流通池达到预定温度。

[0022] 上述实现中的任何一种或更多种存在变化，该系统还包括调节设备，其中该调节设备用于实现多个阱中的多核苷酸的保存。

[0023] 上述实现中的任何一种或更多种存在变化，其中调节设备冷冻干燥流通池以保存多个阱中的多核苷酸。

[0024] 上述实现中的任何一种或更多种存在变化，其中调节设备包括流控设备的特征，该特征将防腐流体分配到流通池，从而在流通池上施加防腐涂层，以保存和保护多个阱中的多核苷酸。

[0025] 上述实现中的任何一种或更多种存在变化，其中处理器进一步：接收指示可移除

存储设备将从模块接口卸载的信号,并响应于指示可移除存储设备将从模块接口卸载的信号来激活调节设备。

[0026] 上述实现中的任何一种或更多种存在变化,其中处理器还:接收指示可移除存储设备已经从模块接口卸载的信号,并且响应于指示可移除存储设备已经从模块接口卸载的信号,向用户提供可移除存储设备被卸载的通知。

[0027] 另一实现涉及一种用于非易失性存储的方法,包括:将可移除存储设备放置在存储仪器的模块接口中,该可移除存储设备包括流通池,其中该模块接口适于将该可移除存储设备固定在静止位置;操作测序设备对流通池的多个阱中的阱内的多核苷酸进行合成测序,以确定描述多核苷酸的核苷酸的一组测序数据;操作流控设备,以向流通池提供一种或更多种流体类型,用于利用测序设备进行合成测序;以及基于该组测序数据和编码方案来确定一组二进制数据,其中该编码方案描述了将信息从二进制格式转换成与该组测序数据相关联的非二进制格式的一组规则。

[0028] 上述实现中的任何一种或更多种存在变化,进一步包括:接收一组输入数据;基于该组输入数据和编码方案确定一组输入测序数据;以及基于该组输入测序数据操作合成设备以在多个阱的目标阱中产生输入多核苷酸。

[0029] 上述实现中的任何一种或更多种存在变化,还包括接收指示可移除存储设备已经与模块接口耦合的信号。

[0030] 上述实现中的任何一种或更多种存在变化,还包括响应于指示可移除存储设备已经与模块接口耦合的信号,操作流控设备以向流通池提供制备流体,其中制备流体适于移除流通池上的防腐涂层,以使流通池能够进行读取和写入操作。

[0031] 上述实现中的任何一种或更多种存在变化,还包括响应于指示可移除存储设备已经与模块接口耦合的信号来操作调节设备,其中调节设备将流通池带到读取和写入操作所需的预定温度。

[0032] 上述实现中的任何一种或更多种存在变化,还包括接收指示可移除存储设备将从模块接口卸载的信号。

[0033] 上述实现中的任何一种或更多种存在变化,还包括响应于指示可移除存储设备将从模块接口卸载的信号来激活调节设备。

[0034] 上述实现中的任何一种或更多种存在变化,其中激活的调节设备向流通池分配防腐流体,从而在流通池上施加防腐涂层,以保存和保护多个阱中的多核苷酸。

[0035] 上述实现中的任何一种或更多种存在变化,其中激活的调节设备冷冻干燥流通池以保存多个阱中的多核苷酸。

[0036] 上述实现中的任何一种或更多种存在变化,还包括:接收指示可移除存储设备已经从模块接口卸载的信号;以及响应于指示可移除存储设备已经从模块接口卸载的信号,向用户提供可移除存储设备被卸载的通知。

[0037] 另一实现涉及一种用于非易失性存储的可移除存储设备,包括:适于与存储仪器的模块接口耦合并将可移除存储设备固定在静止位置的壳体;密封在所述壳体内部的流通池,所述流通池包括多个具有可从流通池的第一表面进入的开口侧的阱,所述阱适于包含存储机器写入数据的多核苷酸;测序接口,其定位在流通池附近,并且适于当可移除存储设备处于静止位置时将光传输到阱;以及流控接口,该流控接口定位在壳体上,并且适于当可

移除存储设备处于静止位置时将流体从存储仪器的流控设备传输到阱。

[0038] 上述实现中的任何一种或更多种存在变化,还包括位于壳体上的电接口,以在可移除存储设备处于静止位置时与存储仪器交换电信号。

[0039] 上述实现中的任何一种或更多种存在变化,还包括位于流通池的第二表面上的集成电路,电接口用于提供电力和指令以操作集成电路。

[0040] 上述实现中的任何一种或更多种存在变化,其中第二表面与第一表面相对,并且其中集成电路基于从存储仪器接收的信号选择性地:经由测序接口将光发射到多个阱的每个阱中,检测从与多个阱的每个阱中的核苷酸相关联的标记发射的光的荧光,并且向存储仪器提供一组荧光数据,其中该组荧光数据可用于确定一组测序数据,该组测序数据描述了多个阱的每个阱中多核苷酸的核苷酸。

[0041] 上述实现中的任何一种或更多种存在变化,其中可移除存储设备包括:将可移除存储设备与电源耦合的电源连接,以及检测流通池的一个或多个特性并在检测到的特性超过预定阈值时提供指示的一组传感器,其中该组传感器包括选自温度传感器、湿度传感器、光传感器和辐射传感器构成的组的一个或多个传感器。

[0042] 上述实现中的任何一种或更多种存在变化,还包括至少一个传感器,该至少一个传感器指示一个或多个环境条件是否落在预定范围之外,该一个或多个环境条件包括湿度、温度、光或辐射中的一个或多个。

[0043] 上述实现中的任何一种或更多种存在变化,流通池包括具有穿过阱底部区域形成的多个开口的基底。

[0044] 上述实现中的任何一种或更多种存在变化,流通池还包括在穿过阱底部区域形成的开口中的导电材料。

[0045] 上述实现中的任何一种或更多种存在变化,导电材料包括氧化铟锡。

[0046] 上述实现中的任何一种或更多种存在变化,进一步包括在穿过阱的底部区域形成的开口中的透明各向异性材料。

[0047] 上述实现中的任何一种或更多种存在变化,流通池包括具有一个或多个焊盘的下侧,以接触位于流通池下方的集成电路芯片。

[0048] 上述实现中的任何一种或更多种存在变化,流通池包括玻璃。

[0049] 上述实现中的任何一种或更多种存在变化,进一步包括将多核苷酸转移到流通池中相应测序表面的薄膜介质。

[0050] 上述实现中的任何一种或更多种存在变化,其中多核苷酸电泳地转移到流通池中相应的测序表面。

[0051] 上述实现中的任何一种或更多种存在变化,其中多核苷酸通过印迹转移到流通池中相应的测序表面。

[0052] 一种实现涉及一种方法,包括接收读取请求,其中读取请求与存储设备中的特定位置相关联;创建位于存储设备中特定位置的多核苷酸的拷贝;将多核苷酸的拷贝转移到读取位置;以及在读取位置读取多核苷酸的拷贝。

[0053] 上述实现中的任何一种或更多种存在变化,其中存储设备包括一组一个或多个写入位置;特定位置包括在一组一个或多个写入位置中;读取位置不包括在一组一个或多个写入位置中;并且每当存储设备接收到写入请求时,存储设备仅将数据写入由一组

一个或多个写入位置所包括的位置。

[0054] 上述实现中的任何一种或更多种存在变化,其中存储设备包括流通池;来自一个或多个写入位置的每个写入位置设置在流通池的第一表面上;并且读取位置设置在流通池的第二表面上。

[0055] 上述实现中的任何一种或更多种存在变化,其中第一表面和第二表面彼此相对。

[0056] 上述实现中的任何一种或更多种存在变化,其中第一表面是流通池的底表面;第二表面是流通池的顶表面。

[0057] 上述实现中的任何一种或更多种存在变化,其中来自一组一个或多个写入位置的每个写入位置具有对应的一组电极,以使用该写入位置中的核苷酸来编码数据。

[0058] 上述实现中的任何一种或更多种存在变化,其中:存储设备包括专用集成电路;并且专用集成电路对于来自一组一个或多个写入位置的每个写入位置,驱动对应于该位置的一组电极,从而使用该写入位置中的核苷酸编码数据。

[0059] 上述实现中的任何一种或更多种存在变化,其中存储设备包括第二集成电路,其中第二集成电路从读取位置的核苷酸读取数据。

[0060] 上述实现中的任何一种或更多种存在变化,其中专用集成电路和第二集成电路都是互补金属氧化物半导体 (CMOS) 芯片。

[0061] 上述实现中的任何一种或更多种存在变化,其中:来自一组一个或多个写入位置的每个写入位置是可寻址的;并且除了一组一个或多个写入位置之外,存储设备不包括任何可寻址位置来以多核苷酸的形式存储数据。

[0062] 上述实现中的任何一种或更多种存在变化,其中:存储设备存储一组一个或多个逻辑数据分组;该方法包括创建索引,该索引为来自一组一个或多个逻辑数据分组的每个逻辑数据分组指定关联的一组一个或多个位置;并且对于来自一组一个或多个逻辑数据分组的每个逻辑数据分组,来自该逻辑数据分组的关联的一组一个或多个位置的每个位置是由一组一个或多个写入位置包括的写入位置。

[0063] 上述实现中的任何一种或更多种存在变化,其中该方法包括接收从一组一个或多个逻辑数据分组中检索特定逻辑数据分组的请求;从索引中检索特定逻辑数据分组的关联的一组一个或多个位置;以及基于特定逻辑数据分组的关联的一组一个或多个位置来确定存储设备中的特定位置。

[0064] 上述实现中的任何一种或更多种存在变化,其中,对于来自一组一个或多个逻辑数据分组的每个逻辑数据分组,该逻辑数据分组的关联的一组一个或多个位置包括开始位置和结束位置。

[0065] 上述实现中的任何一种或更多种存在变化,其中,对于来自一组一个或多个逻辑数据分组的每个逻辑数据分组,来自该逻辑数据分组的关联的一组一个或多个位置的每个位置是二维网格中的位置。

[0066] 上述实现中的任何一种或更多种存在变化,其中该方法包括将索引存储在非核苷酸存储器中;以及使用珠上的核苷酸存储该索引。

[0067] 上述实现中的任何一种或更多种存在变化,其中将多核苷酸的拷贝转移到读取位置包括将多核苷酸的拷贝固定到特定位置的珠上;以及将珠从特定位置转移到读取位置。

[0068] 上述实现中的任何一种或更多种存在变化,其中特定位置在珠的表面上。

[0069] 上述实现中的任何一种或更多种存在变化,其中使用选自以下项构成的组的技术将多核苷酸的拷贝转移到读取位置:电泳;介电泳;和流体层流流动。

[0070] 另一实现涉及一种系统,该系统包括一个或多个存储指令的非暂时性计算机可读介质,以执行包括接收读取请求的方法,其中该读取请求与存储设备中的特定位置相关联;创建位于存储设备中特定位置的多核苷酸的拷贝;将多核苷酸的拷贝转移到读取位置;以及在读取位置读取多核苷酸的拷贝。

[0071] 上述实现中的任何一种或更多种存在变化,其中存储设备包括一组一个或多个写入位置;特定位置包括在一组一个或多个写入位置中;读取位置不包括在一组一个或多个写入位置中;并且每当存储设备接收到写入请求时,存储设备仅将数据写入由一组一个或多个写入位置所包括的位置。

[0072] 上述实现中的任何一种或更多种存在变化,其中存储设备包括流通池;来自一个或多个写入位置的每个写入位置设置在流通池的第一表面上;并且读取位置设置在流通池的第二表面上。

[0073] 上述实现中的任何一种或更多种存在变化,其中第一表面和第二表面彼此相对。

[0074] 上述实现中的任何一种或更多种存在变化,其中第一表面是流通池的底表面;第二表面是流通池的顶表面。

[0075] 上述实现中的任何一种或更多种存在变化,其中来自一组一个或多个写入位置的每个写入位置具有对应的一组电极,以使用该写入位置中的核苷酸来编码数据。

[0076] 上述实现中的任何一种或更多种存在变化,其中存储设备包括专用集成电路;并且专用集成电路对于来自一组一个或多个写入位置的每个写入位置,驱动对应于该位置的一组电极,从而使用该写入位置中的核苷酸编码数据。

[0077] 上述实现中的任何一种或更多种存在变化,其中存储设备包括第二集成电路,其中第二集成电路从读取位置的核苷酸读取数据。

[0078] 上述实现中的任何一种或更多种存在变化,其中专用集成电路和第二集成电路都是CMOS芯片。

[0079] 上述实现中的任何一种或更多种存在变化,其中来自一组一个或多个写入位置的每个写入位置是可寻址的;并且除了一组一个或多个写入位置之外,存储设备不包括任何可寻址位置来以多核苷酸的形式存储数据。

[0080] 上述实现中的任何一种或更多种存在变化,其中存储设备存储一组一个或多个逻辑数据分组;该方法包括创建索引,该索引为来自一组一个或多个逻辑数据分组的每个逻辑数据分组指定关联的一组一个或多个位置;并且对于来自一组一个或多个逻辑数据分组的每个逻辑数据分组,来自该逻辑数据分组的关联的一组一个或多个位置的每个位置是由一组一个或多个写入位置包括的写入位置。

[0081] 上述实现中的任何一种或更多种存在变化,其中该方法包括:接收从一组一个或多个逻辑数据分组中检索特定逻辑数据分组的请求;从索引中检索特定逻辑数据分组的关联的一组一个或多个位置;以及基于特定逻辑数据分组的关联的一组一个或多个位置来确定存储设备中的特定位置。

[0082] 上述实现中的任何一种或更多种存在变化,其中,对于来自一组一个或多个逻辑数据分组的每个逻辑数据分组,该逻辑数据分组的关联的一组一个或多个位置包括开



始位置和结束位置。

[0083] 上述实现中的任何一种或更多种存在变化,其中,对于来自一组一个或更多个逻辑数据分组的每个逻辑数据分组,来自该逻辑数据分组的关联的一组一个或更多个位置的每个位置是二维网格中的位置。

[0084] 上述实现中的任何一种或更多种存在变化,其中该方法包括:将索引存储在非核苷酸存储器中;以及使用珠上的核苷酸存储该索引。

[0085] 上述实现中的任何一种或更多种存在变化,其中将多核苷酸的拷贝转移到读取位置包括将多核苷酸的拷贝固定到特定位置的珠上;以及将珠从特定位置转移到读取位置。

[0086] 上述实现中的任何一种或更多种存在变化,其中特定位置在珠的表面上。

[0087] 上述实现中的任何一种或更多种存在变化,其中使用选自以下项构成的组的技术将多核苷酸的拷贝转移到读取位置:电泳;介电泳;和流体层流流动。

[0088] 一种实现涉及存储用于执行方法的指令的一个或更多个非暂时性计算机可读介质,该方法包括:接收读取请求,其中该读取请求与存储设备中的特定位置相关联;创建位于存储设备中特定位置的多核苷酸的拷贝;将多核苷酸的拷贝转移到读取位置;以及在读取位置读取多核苷酸的拷贝。

[0089] 上述实现中的任何一种或更多种存在变化,其中:存储设备包括一组一个或更多个写入位置;特定位置包括在一组一个或更多个写入位置中;读取位置不包括在一组一个或更多个写入位置中;并且每当存储设备接收到写入请求时,存储设备仅将数据写入由一组一个或更多个写入位置所包括的位置。

[0090] 上述实现中的任何一种或更多种存在变化,其中:存储设备包括流通池;来自一个或更多个写入位置的每个写入位置设置在流通池的第一表面上;并且读取位置设置在流通池的第二表面上。

[0091] 上述实现中的任何一种或更多种存在变化,其中第一表面和第二表面彼此相对。

[0092] 上述实现中的任何一种或更多种存在变化,其中:第一表面是流通池的底表面;第二表面是流通池的顶表面。

[0093] 上述实现中的任何一种或更多种存在变化,其中来自一组一个或更多个写入位置的每个写入位置具有对应的一组电极,以使用该写入位置中的核苷酸来编码数据。

[0094] 上述实现中的任何一种或更多种存在变化,其中:存储设备包括专用集成电路;并且专用集成电路对于来自一组一个或更多个写入位置的每个写入位置,驱动对应于该位置的一组电极,从而使用该写入位置中的核苷酸编码数据。

[0095] 上述实现中的任何一种或更多种存在变化,其中存储设备包括第二集成电路,其中第二集成电路从读取位置的核苷酸读取数据。

[0096] 上述实现中的任何一种或更多种存在变化,其中专用集成电路和第二集成电路都是CMOS芯片。

[0097] 上述实现中的任何一种或更多种存在变化,其中:来自一组一个或更多个写入位置的每个写入位置是可寻址的;并且除了一组一个或更多个写入位置之外,存储设备不包括任何可寻址位置来以多核苷酸的形式存储数据。

[0098] 上述实现中的任何一种或更多种存在变化,其中:存储设备存储一组一个或更多个逻辑数据分组;该方法包括创建索引,该索引为来自一组一个或更多个逻辑数据分组的

每个逻辑数据分组指定关联的一组一个或多个位置；并且对于来自一组一个或多个逻辑数据分组的每个逻辑数据分组，来自该逻辑数据分组的关联的一组一个或多个位置的每个位置是由一组一个或多个写入位置包括的写入位置。

[0099] 上述实现中的任何一种或更多种存在变化，其中该方法包括：接收从一组一个或多个逻辑数据分组中检索特定逻辑数据分组的请求；从索引中检索特定逻辑数据分组的关联的一组一个或多个位置；以及基于特定逻辑数据分组的关联的一组一个或多个位置来确定存储设备中的特定位置。

[0100] 上述实现中的任何一种或更多种存在变化，其中，对于来自一组一个或多个逻辑数据分组的每个逻辑数据分组，该逻辑数据分组的关联的一组一个或多个位置包括开始位置和结束位置。

[0101] 上述实现中的任何一种或更多种存在变化，其中，对于来自一组一个或多个逻辑数据分组的每个逻辑数据分组，来自该逻辑数据分组的关联的一组一个或多个位置的每个位置是二维网格中的位置。

[0102] 上述实现中的任何一种或更多种存在变化，其中该方法包括：将索引存储在非核苷酸存储器中；以及使用珠上的核苷酸存储该索引。

[0103] 上述实现中的任何一种或更多种存在变化，其中将多核苷酸的拷贝转移到读取位置包括：将多核苷酸的拷贝固定到特定位置的珠上；以及将珠从特定位置转移到读取位置。

[0104] 上述实现中的任何一种或更多种存在变化，其中特定位置在珠的表面上。

[0105] 上述实现中的任何一种或更多种存在变化，其中使用选自以下项构成的组的技术将多核苷酸的拷贝转移到读取位置：电泳；介电泳；和流体层流流动。

[0106] 应当理解，前述概念和下面更详细讨论的附加概念的所有组合（只要这些概念不相互矛盾）都被认为是这里公开的发明主题的一部分，并且实现了这里描述的益处/优点。特别是，出现在本公开的结尾处的所要求保护的主题的所有组合被设想为本文公开的创造性主题的一部分。

[0107] 附图简述

[0108] 在附图和下面的描述中阐述了一个或多个实现方案的细节。根据说明书、附图和权利要求，其他特征、方面和优点将变得明显，其中：

[0109] 图1描绘了可用于进行生化过程的系统的示例的示意框图；

[0110] 图2描绘了可与图1的系统一起使用的可消耗盒的示例的示意性截面框图；

[0111] 图3描绘了可用于图1的系统的流通池的示例的透视图；

[0112] 图4描绘了图3的流通池的通道的放大透视图；

[0113] 图5描绘了可以结合到图4的通道中的阱的示例的示意性截面图；

[0114] 图6描绘了用于读取多核苷酸的过程的示例的流程图；

[0115] 图7描绘了可以结合到图4的通道中的阱的另一个例子的示意性截面图；

[0116] 图8描绘了用于写入多核苷酸的过程的示例的流程图；

[0117] 图9描绘了电极组件的示例的俯视图；

[0118] 图10描绘了可结合到图4的通道中的阱的另一示例的示意性截面图；

[0119] 图11描述了多核苷酸存储系统的一个例子的示意图；

[0120] 图12A描绘了说明可与图11的多核苷酸存储系统一起使用的存储设备的示例的示

意图；

[0121] 图12B描绘了说明可与图11的多核苷酸存储系统一起使用的存储设备的替代实例的示意图；

[0122] 图12C描绘了说明可与图11的多核苷酸存储系统一起使用的存储设备的又一替代实例的示意图；

[0123] 图13描绘了可被执行以向图15的存储设备提供读取和写入操作的缓存的过程的流程图；

[0124] 图14描绘了可被执行以提供与存储设备的冗余数据写入和读取的过程的流程图；

[0125] 图15描绘了可被执行以提供与存储设备的高速数据写入和读取的过程的流程图；

[0126] 图16A描绘了可以用图11的多核苷酸存储系统执行以安装图12A-12C中任一个的存储设备的过程的流程图；

[0127] 图16B描绘了可以用图11的多核苷酸存储系统执行以卸载图12A-12C中任一个的存储设备的过程的流程图；

[0128] 图17A描绘了说明膜复制过程的第一方面的示意图；

[0129] 图17B描绘了说明图17A的膜复制过程的第二方面的示意图；

[0130] 图18描绘了可结合到图4的通道中的阱的另一示例的示意性截面图；和

[0131] 图19描绘了珠和流通池之间的关系的示例的示意框图。

[0132] 将认识到,附图中的一些或所有附图是用于说明目的的示意性图示。为了说明一个或更多个实现方案的目的而提供附图,明确地理解这些附图将不用于限制权利要求的范围或含义。

[0133] 详细描述

[0134] 在一些方面,本文公开了用于DNA存储设备的方法和系统,该DNA存储设备可以是可移除和便携式的,并且可以用作大规模和小规模存档目的的DNA硬盘驱动器设备。在一些进一步的方面,本文公开了用于减轻包含机器写入的DNA的DNA存储设备中的数据错误风险的方法和系统。机器写入的DNA可以替代传统形式的数据存储(如磁存储、光存储和固态存储)。机器写入的DNA可以提供更快的读写速度、更长的数据保留时间、更低的功耗和更高的数据密度。虽然这里描述的例子指的是“DNA存储系统”或“DNA存储设备”,但应该理解这只是多核苷酸存储的一个例子。本文的教导可以容易地应用于利用不一定是DNA形式的多核苷酸的存储系统和设备。因此,本发明不限于使用DNA作为本文所述的唯一一种用于存储的多核苷酸。此外,多核苷酸只是如本文所述可用于存储的生物材料的一个例子。

[0135] 数字信息如何存储在DNA中的例子在2015年9月17日公布的题为“High-Capacity of Storage of Digital Information in DNA”的第2015/0261664号美国出版物中公开,其全部内容通过引用结合于此。例如,可以使用来自编码理论的方法来增强来自DNA片段的编码消息的可恢复性,包括禁止已知与现有高通量技术中的较高错误率相关联的DNA均聚物(即,多于一个相同碱基的运行)。此外,类似于奇偶校验位的检错部分可以被集成到代码中的索引信息中。更复杂的方案,包括但不限于纠错码,实际上,信息学中目前采用的基本上任何形式的数字数据安全(例如,基于RAID的方案),可以在DNA存储方案的未来发展中实施。信息的DNA编码可以用软件计算。组成每个计算机文件的字节可以由没有均聚物的DNA序列通过编码方案来表示,以产生编码文件,该编码文件用形成DNA序列的五个或六个碱基

来替换每个字节。

[0136] 编码方案中使用的代码可以被构造成允许接近运行长度受限信道的最佳信息容量的直接编码(例如,没有重复的核苷酸),尽管也可以使用其他编码方案。所得的电子DNA序列可能太长,不容易通过标准寡核苷酸合成产生,并且可能被分成长度为100个碱基的重叠片段,其中75个碱基重叠。为了降低系统合成误差引入任何特定碱基运行的风险,片段中的交替片段可以被转换成它们的反向互补,这意味着每个碱基可以被“写”四次,每个方向两次。然后,每个片段可以用索引信息加上简单的错误检测信息来扩充,该索引信息允许确定该段起源的计算机文件及其在该计算机文件中的位置。这种索引信息也可以被编码为非重复的DNA核苷酸,并附加到DNA片段的信息存储碱基上。将DNA片段分成100个碱基的长度和75个碱基的重叠纯粹是任意的和说明性的,应当理解,可以使用其他长度和重叠,而不是限制性的。

[0137] 可以使用DNA片段的其他编码方案,例如提供增强的纠错性能。可以增加索引信息量,以便允许对更多或更大的文件进行编码。为了避免DNA片段中的系统模式,编码方案的一个扩展可以是添加信息中的改变。一种方法是对DNA片段中的信息进行“混洗”,如果知道混洗的模式,就可以检索到信息。不同模式的混洗可用于不同的DNA片段。另一种方法是在每个DNA片段的信息中加入一定程度的随机性。为此,可以通过使用一系列随机数字和包含编码在DNA片段中的信息的数字的模加来使用一系列随机数字。如果知道所用的一系列随机数字,可以在解码过程中通过模减法检索信息。不同系列的随机数字可用于不同的DNA片段。每个串的数据编码部分可包含香农信息,每DNA碱基5.07位,这接近于运行长度限制为1的基数4通道每DNA碱基5.05位的理论最佳值。索引实现可以允许 $3^{14} = 4782969$ 个唯一的数据位置。将用于指定文件和文件内位置的索引三进制数位(因此也是基数)的数量增加两个,即16个,得到 $3^{16} = 43046721$ 个唯一位置,超过了嵌套引物分子存储器(NPMM)方案的实际最大值16.8M。

[0138] DNA片段设计可以在三个不同的运行中合成(DNA片段随机分配给运行),以创建每个DNA片段设计大约 $1.2 \times 10^7$ 个拷贝。可以使用亚磷酰胺化学,并且可以使用原位微阵列合成平台中的喷墨印刷和流通池反应器技术。无水室中的喷墨印刷可以允许将非常少量的亚磷酰胺输送到2D平面表面上的受限耦合区域,导致平行添加数十万个碱基。随后的氧化和脱三苯甲基可以在流通池反应器中进行。一旦完成DNA合成,寡核苷酸就可以从表面切割下来并脱保护。

[0139] 然后将衔接子添加到DNA片段中,以便能够制作DNA片段的多个拷贝。没有衔接子的DNA片段可能需要额外的化学过程,通过在DNA片段的末端添加额外的基团来“启动”合成多个拷贝的化学过程。寡核苷酸可以使用聚合酶链反应(PCR)方法和成对末端PCR引物进行扩增,然后进行珠纯化和定量。然后寡核苷酸可以被测序以产生104个碱基的读数。然后,数字信息解码可以通过从两端测序每个寡核苷酸的中间碱基,并快速计算全长寡核苷酸和去除与设计不一致的序列读数来进行。序列读取可以使用计算机软件进行解码,该软件可以完全逆转编码过程。奇偶校验三进制数位指示错误或可以明确解码或分配给重建的计算机文件的序列读取可以被丢弃。每个解码文件中的位置可以在多个不同的测序DNA寡核苷酸中检测到,简单多数表决可用于解决由DNA合成或测序错误引起的任何差异。

[0140] 虽然在机器写入的DNA的上下文中提供了几个例子,但是可以设想,这里描述的原

理可以应用于其他种类的机器写入的生物材料。

[0141] 如本文所用,术语“机器写入的DNA”应理解为包括一条或多条多核苷酸链,所述多核苷酸链由机器产生,或以其他方式由机器修饰,以存储数据或其他信息。本文多核苷酸的一个例子是DNA。要注意的是,尽管在本公开全文中使用了在被读取或写入的DNA的上下文中的术语“DNA”,但是该术语仅用作多核苷酸的代表性例子,并且可以包含多核苷酸的概念。“机器”,如这里所使用的“机器写入的”,可以包括专门设计用于写入DNA的仪器或系统,如这里更详细描述。该系统可以是非生物的或生物的。在一个例子中,生物系统可以包括或者是聚合酶。例如,聚合酶可以是末端脱氧核苷酸转移酶(TdT)。在生物系统中,该过程可以另外由机器硬件(例如,处理器)或算法控制。“机器写入的DNA”可以包括任何具有一个或多个由机器写入的碱基序列的多核苷酸。虽然机器写入的DNA在本文中被用作例子,但是其它多核苷酸链可以替代本文所述的机器写入的DNA。“机器写入的DNA”可以包括天然碱基和天然碱基的修饰,包括但不限于甲基化或其他化学标签修饰的碱基;一种类似于DNA的人工合成聚合物,如肽核酸(PNA);或者吗啉代DNA。“机器写入的DNA”还可以包括由至少一条来源于自然界(例如,从天然存在的生物体中提取)的碱基链形成的DNA链或其它多核苷酸,机器写入的碱基链以平行方式或端对端方式固定于其上。在其他实现中,“机器写入的DNA”可以由生物系统(例如酶)写入,代替或补充在此描述的非生物系统(例如电极机器)对DNA进行写入。换句话说,“机器写入的DNA”可以是机器直接写入的;或者通过由算法和/或机器控制的酶(例如聚合酶)写入的。

[0142] “机器写入的DNA”可以包括已经用已知技术从原始形式(例如,照片、文本文档等)转化为二进制编码序列的转换而来的数据,然后用已知技术将该二进制编码序列转化为DNA碱基序列,然后用机器以一条或多条DNA链或其它多核苷酸的形式产生该DNA碱基序列。或者,可以生成“机器写入的DNA”,以索引或以其他方式跟踪预先存在的DNA,存储来自任何其他来源的数据或信息,并用于任何合适的目的,而不一定需要将原始数据转换成二进制代码的中间步骤。

[0143] 如下文更详细描述,机器写入的DNA可以被写入反应位点和/或从反应位点读取。如本文所用,术语“反应位点”是可以发生至少一个指定反应的局部区域。反应位点可以包括反应结构或基底的支撑表面,在该支撑表面上可以固定物质。例如,反应位点可以是一个离散的空间区域,其中写入了一组离散的多核苷酸链或其他多核苷酸。反应位点可以允许与相邻反应位点的反应分离的化学反应。提供DNA机器写入的设备可以包括带有具有写入特征(例如,电极)和/或读取特征的阱的流通池。在一些情况下,反应位点可以包括反应结构的表面(其可以位于流通池的通道中),该表面上已经具有反应组分,例如其上的多核苷酸集落。在一些流通池中,集落中的多核苷酸具有相同的序列,例如,单链或双链模板的克隆拷贝。然而,在一些流通池中,反应位点可以仅包含单个多核苷酸分子,例如单链或双链形式。

[0144] 多个反应位点可以沿着流通池的反应结构随机分布,或者可以以预定的方式排列(例如,在矩阵中并排排列,例如在微阵列中)。反应位点还可以包括反应室、凹部或阱,其至少部分地限定了被配置成分隔指定反应的空间区域或体积。如本文所用,术语“反应室”或“反应凹部”包括支撑结构的限定空间区域(其通常与流动通道流体地联接)。反应凹部可以至少部分地与周围环境或其他空间区域分离。例如,多个反应凹部可以通过共享的壁彼此分开。作为更具体的例子,反应凹部可以是纳米阱,其包括由检测表面的内表面限定的凹

痕、凹坑、阱、凹槽、空腔或凹陷,并且具有开口或孔径(即,侧面开放),使得纳米阱可以与流动通道流体地联接。

[0145] 多个反应位点可以沿着流通池的反应结构随机分布,或者可以以预定的方式排列(例如,在矩阵中并排排列,例如在微阵列中)。反应位点还可以包括反应室、凹部或阱,其至少部分地限定了被配置成分隔指定反应的空间区域或体积。如本文所用,术语“反应室”或“反应凹部”包括支撑结构的限定空间区域(其通常与流动通道流体地联接)。反应凹部可以至少部分地与周围环境或其他空间区域分离。例如,多个反应凹部可以通过共享的壁彼此分开。作为更具体的例子,反应凹部可以是纳米阱,其包括由检测表面的内表面限定的凹痕、凹坑、阱、凹槽、空腔或凹陷,并且具有开口或孔径(即,侧面开放),使得纳米阱可以与流动通道流体地联接。

[0146] 为了读取机器写入的DNA,可以定义反应位点的一个或更多个离散的可检测区域。这种可检测区域可以是可成像区域、电检测区域或基于读取过程中存在的核苷酸类型在性质上具有可测量的变化(或性质上没有变化)的其他类型的区域。

[0147] 如本文所用,术语“像素”是指离散的可成像区域。每个可成像区域可以包括存在多核苷酸的隔室或离散空间区域。在一些情况下,像素可以包括两个或更多个反应位点(例如,两个或更多个反应室、两个或更多个反应凹部、两个或更多个阱等。在一些其他情况下,像素可以仅包括一个反应位点。使用相应的检测设备,例如图像传感器或其他光检测设备来检测每个像素。光检测设备可以使用集成电路制造工艺来制造,例如用于制造电荷耦合设备电路(CCD)或互补金属氧化物半导体(CMOS)设备或电路的工艺。光检测设备因此可以包括例如一种或更多种半导体材料,并且可以采取例如CMOS光检测设备(例如CMOS图像传感器)或CCD图像传感器(另一种类型的图像传感器)的形式。CMOS图像传感器可以包括光传感器阵列(例如光电二极管)。在一个实现中,单个图像传感器可以与物镜一起使用,以在成像事件期间捕获几个“像素”。在一些其他实现中,每个分立的光电二极管或光传感器可以捕获相应的像素。在一些实现中,一个或更多个检测设备的光传感器(例如,光电二极管)可以与相应的反应位点相关联。与反应位点相关联的光传感器可以检测来自相关反应位点的光发射。在一些实现中,当在相关的反应位点发生指定的反应时,可以通过至少一个光导来进行光发射的检测。在一些实现中,多个光传感器(例如,光检测或摄像设备的几个像素)可以与单个反应位点相关联。在一些实现中,单个光传感器(例如,单个像素)可以与单个反应位点或一组反应位点相关联。

[0148] 如本文所用,术语“合成”应理解为包括由机器产生DNA以存储数据或其他信息的过程。因此,机器写入的DNA可以构成合成的DNA。如本文所用,术语“可消耗的盒”、“试剂盒”、“可移除的盒”和/或“盒”是指构成盒或盒系统的组件的相同盒和/或部件的组合。本文所述的盒可以独立于具有反应位点的元件,例如具有多个阱的流通池。在一些情况下,流通池可以可移除地插入到盒中,然后盒被插入到仪器中。在一些其他实现中,流通池可以在没有盒的情况下可移除地插入到仪器中。如本文所使用的,术语“生化分析”可以包括生物分析或化学分析中的至少一种。

[0149] 术语“基于”应理解为是指某事物至少部分地由其被指示为“基于”的事物所决定。为了表明某物必须完全由其他东西决定,它被描述为完全基于它完全由什么决定。

[0150] 术语“非核苷酸存储器”应理解为是指能够以不同于核苷酸的形式存储数据或指

令的对象、设备或设备组合,所述数据或指令可由设备检索和/或处理。“非核苷酸存储器”的例子包括固态存储器、磁存储器、硬盘驱动器、光驱和前述的组合(例如,磁光存储元件)。

[0151] 术语“DNA存储设备”应理解为是指被配置为以多核苷酸序列(如机器写入的DNA)的形式存储数据或指令的对象、设备或设备组合。“DNA存储设备”的例子包括具有如本文所述的可寻址阱的流通池、包含多个这种流通池的系统、以及存储已经从合成核苷酸序列的表面切割下来的核苷酸序列的管或其它容器。如本文所用,术语“核苷酸序列”或“多核苷酸序列”应理解为包括多核苷酸分子,以及分子的基础序列,这取决于上下文。多核苷酸序列可以包含(或编码)指示某些物理特征的信息。

[0152] 本文所述的实现可用于执行用于可消耗盒制备和/或生物化学分析和/或机器写入的DNA合成的指定反应。

#### [0153] I. 系统概述

[0154] 图1是被配置为进行生化分析和/或合成的系统100的示意图。系统100可以包括基础仪器102,该基础仪器102被配置为接收和可分离地接合可移除盒200和/或具有一个或更多个反应位点的部件。基础仪器102和可移除盒200可以被配置成相互作用以将生物材料运输到系统100内的不同位置和/或进行包括生物材料的指定反应,以便为后续分析(例如,通过合成生物材料)制备生物材料,并且可选地,检测生物材料的一个或更多个事件。在一些实现中,基础仪器102可以被配置成检测直接在可移除盒200上的生物材料的一个或更多个事件。该事件可以指示与生物材料一起的指定反应。可移除盒200可以根据这里描述的任何盒来构造。

[0155] 尽管下面参考了如图1所示的基础仪器102和可移除盒200,但是应当理解,基础仪器102和可移除盒200仅示出了系统100的一种实现,并且存在其他实现。例如,基础仪器102和可移除盒200包括各种部件和特征,它们共同执行用于制备生物材料和/或分析生物材料的若干操作。此外,尽管本文所述的可移除盒200包括具有反应位点的元件,例如具有多个阱的流通池,但是其他盒可以独立于具有反应位点的元件,并且具有反应位点的元件可以单独插入到基础仪器102中。也就是说,在一些情况下,流通池可以可移除地插入到可移除盒200中,然后可移除盒200被插入到基础仪器102中。在一些其他实现中,流通池可以可移除地直接插入到基础仪器102中,而没有可移除的盒200。在另外的实现中,流通池可以集成到插入到基础仪器102中的可移除盒200中。

[0156] 在示出的实现中,基础仪器102和可移除盒200中的每一个能够执行某些功能。然而,应理解,基础仪器102和可移除盒200可以执行不同的功能和/或可以共有这样的功能。例如,基础仪器102被显示为包括检测组件110(例如,成像设备),其被配置为检测可移除盒200处的指定反应。在替代实现中,可移除盒200可以包括检测组件,并且可以通信地耦合到基础仪器102的一个或更多个部件。作为另一个示例,基础仪器102是不提供、接纳液体或与可移除盒200交换液体的“干”仪器(“dry” instrument)。也就是说,如图所示,可移除盒200包括可消耗试剂部分210和流通池接收部分220。可消耗试剂部分210可以包含在生化分析和/或合成期间使用的试剂。流通池接收部分220可包括光学透明区域或其他可检测区域,用于检测组件110对流通池接收部分220内发生的一个或更多个事件进行检测。在替代实现中,基础仪器102可以向可移除盒200提供例如试剂或其他液体,这些试剂或其他液体随后被可移除盒200消耗(例如,用于指定的反应或合成过程)。

[0157] 如本文所用,生物材料可包括一种或更多种生物或化学物质,例如核苷、核苷酸、核酸、多核苷酸、寡核苷酸、蛋白质、酶、肽、寡肽、多肽、抗体、抗原、配体、受体、多糖、碳水化合物、多磷酸盐、纳米孔、细胞器、脂质层、细胞、组织、生物体和/或生物活性化合物,例如上述物种的类似物或模拟物。在一些情况下,生物材料可以包括全血、淋巴液、血清、血浆、汗液、泪液、唾液、痰液、脑脊液、羊水、精液、阴道分泌物、浆液、滑液、心包液、腹膜液、胸膜液、渗出液、渗出物、囊液、胆汁、尿液、胃液、肠液、粪便样本、含有单个或更多细胞的液体、含有细胞器的液体、流化组织、流化生物体、包括病毒病原体的病毒、含有多细胞生物体的液体、生物拭子和生物洗液。在一些情况下,生物材料可以包括一组合成序列,包括但不限于机器写入的DNA,其可以是固定的(例如,附着在盒中的特定阱中)或不固定的(例如,储存在试管中)。

[0158] 在一些实现中,生物材料可包括添加的材料,例如水、去离子水、盐水溶液、酸性溶液、碱性溶液、洗涤剂溶液和/或pH缓冲液。添加的材料还可以包括将在指定的测定方案期间使用以进行生化反应的试剂。例如,添加的液体可以包括与生物材料进行多个聚合酶链反应(PCR)循环的材料。在其他方面,添加的材料可以是生物材料的载体,例如细胞培养基或其他缓冲的和/或pH调节的和/或等渗的载体,其可以允许或保持生物材料的生物功能。

[0159] 然而,应该理解的是,被分析的生物材料可以是与装载到系统100中或由系统100产生的生物材料不同的形式或状态。例如,装载到系统100中的生物材料可以包括全血或唾液或细胞群,其随后被处理(例如,通过分离或扩增程序)以提供制备的核酸。然后通过系统100分析(例如,通过PCR定量地或通过SBS测序)制备的核酸。因此,当在描述诸如PCR的第一操作的同时使用术语“生物材料”并且在描述随后的第二操作,例如测序时再次使用时,应当理解,第二操作中的生物材料可以相对于在第一操作之前或在第一操作期间的生物材料被改变。例如,可以对扩增子核酸进行测序(例如SBS),所述扩增子核酸由在先前扩增(例如PCR)中扩增的模板核酸产生。在这种情况下,扩增子是模板的拷贝,并且扩增子与模板的量相比以更高的量存在。

[0160] 在一些实现中,系统100可以基于用户提供的物质(例如,全血或唾液或细胞群)自动制备用于生化分析的样本。然而,在其它实现中,系统100可分析生物材料,其部分地或预先被制备用于由用户分析。例如,用户可以提供包含已经从全血中分离和/或扩增的核酸的溶液;或者可以提供其中RNA或DNA序列部分或全部暴露用于处理的病毒样品。

[0161] 如本文中所使用的,“指定反应”包括所讨论的分析物的在化学、电学、物理或光学性质(或特性)中的至少一个方面的改变。在特定的实现中,指定的反应是缔合型结合事件(例如,将所讨论的分析物与荧光标记的生物分子掺合)。指定的反应可以是解离型结合事件(例如,从所讨论的分析物释放荧光标记的生物分子)。指定的反应可以是化学转化、化学变化或化学相互作用。指定的反应也可以是电特性方面的变化。例如,指定的反应可以是溶液中离子浓度的变化。一些反应包括但不限于化学反应,例如,还原、氧化、加成、消除、重排、酯化、酰胺化、醚化、环化或取代;其中第一化学物质结合第二化学物质的结合相互作用;其中两种或更多种化学物质彼此分离的解离反应;荧光;发光;生物发光;化学发光;以及生物反应,例如,核酸复制、核酸扩增、核酸杂交、核酸连接、磷酸化、酶催化、受体结合,或配体结合。指定的反应也可以是质子的添加或移除,例如,作为周围溶液或环境的pH变化而可检测到的。附加的指定的反应可以是检测穿过膜(例如,天然或合成双层膜)的离子流。例



如,当离子流过膜时,电流被中断,并且可以检测到中断。也可以使用带电标签的现场感测;热感测和其他合适的分析感测技术也是如此。

[0162] 在特定的实现中,指定的反应包括将荧光标记的分子掺合到分析物中。分析物可以是寡核苷酸,并且荧光标记的分子可以是核苷酸。当激发光指向具有所标记的核苷酸的寡核苷酸时,可以检测该指定的反应,并且荧光团发射可检测的荧光信号。在可选的实现中,检测到的荧光是化学发光和/或生物发光的结果。指定的反应还可以例如通过使供体荧光团接近受体荧光团而增加荧光共振能量转移 (fluorescence (或 Förster) resonance energy transfer, FRET), 通过分离供体荧光团和受体荧光团来减少 FRET, 通过从荧光团分离猝灭剂来增加荧光或通过共同定位猝灭剂和荧光团来增加荧光。

[0163] 如本文中所使用的,“反应成分”包括可用于获得指定反应的任何物质。例如,反应成分包括试剂、例如酶的催化剂、用于反应的反应物、样品、反应产物、其它生物分子、盐、金属辅因子、螯合剂以及缓冲溶液(例如,氢化缓冲液)。反应成分可以单独地在溶液中或结合在一种或更多种混合物中被输送到流控网络中的不同位置。例如,反应成分可以被输送到其中生物材料被固定的反应室。反应成分可以与生物材料直接地或间接地相互作用。在一些实现中,可移除盒200预装载有一种或更多种参与执行指定测定方案的反应成分。预装载可以在由使用者(例如,在消费者的设施处)接收到盒200之前在一个位置(例如,制造设施)处发生。例如,一种或更多种反应成分或试剂可以预装载到可消耗试剂部分210中。在一些实现中,可移除盒200也可以在流通池接收部分220中预装流通池。

[0164] 在一些实现中,基础仪器102可以构造成每个环节与一个可移除盒200相互作用。在该环节之后,该可移除盒200可以由另一个可移除盒200替换。在一些实现中,基础仪器102可以构造成每个环节与一个以上的可移除盒200相互作用。如本文所使用的,术语“环节”包括执行样品制备和/或生化分析方案中的至少一个。样品制备可包括合成生物材料;和/或分离、隔离、修饰和/或扩增生物材料的一种或更多种成分,使得制备的生物材料适于分析。在一些实现中,环节可以包括其中进行多个受控反应的连续活动,直到(a)已经进行指定数量的反应,(b)已经检测指定数量的事件,(c)已经经过指定时间段的系统时间,(d)信噪比已下降到指定阈值;(e)已经识别目标成分;(f)已经检测到系统失效或故障;和/或(g)已经耗尽用于进行反应的一种或更多种资源。可选地,环节可以包括将系统活动暂停一段时间(例如,分钟、小时、天、周)并且稍后完成该环节,直到(a)-(g)中的至少一个发生。

[0165] 测定方案可以包括用于进行指定的反应、检测指定的反应和/或分析指定的反应的一系列操作。共同地,可移除盒200和基础仪器102可以包括用于执行不同操作的部件。测定方案的操作可以包括流控操作、热控制操作、检测操作和/或机械操作。

[0166] 流控操作包括控制穿过系统100的流体(例如,液体或气体)的流动,系统100可以由基础仪器102和/或可移除盒200致动。在一个例子中,流体是液体形式。例如,流控操作可以包括控制泵以诱导生物材料或反应成分流入反应室中。

[0167] 热控制操作可以包括控制系统100的指定部分的温度,例如可移除盒200的一个或更多个部分的温度。举例来说,热控制操作可以包括升高或降低在其包括生物材料的流体储存在其中的聚合酶链反应(PCR)区域的温度。

[0168] 检测操作可以包括控制检测器的致动或监测检测器的活动以检测生物材料的预定性质、特性或特征。作为一个示例,检测操作可以包括捕获包括生物材料的指定区域的图

像以检测来自指定区域的荧光发射。检测操作可以包括控制光源以照射生物材料或控制检测器以观察生物材料。

[0169] 机械操作可以包括控制指定部件的移动或位置。例如,机械操作可以包括控制马达以移动基础仪器102中的可操作地接合可移除盒200中的可移动阀的阀控制部件。在一些例子中,不同操作的组合可以同时发生。例如,当泵控制穿过反应室的流体流时,检测器可以捕获反应室的图像。在一些例子中,指向不同生物材料的不同操作可以同时发生。例如,第一生物材料可以经历扩增(例如,PCR),同时第二生物材料可以经历检测。

[0170] 类似或相同的流控元件(例如,通道、端口、储器等)可以被不同地标记,以更容易地区分这些流控元件。例如,端口可以被称为储器端口、供应端口、网络端口、馈送端口等。应理解,被不同地标记的两个或更多个流控元件(例如,储器通道、样本通道、流动通道、桥接通道)不要求流控元件在结构上不同。另外,可以修改权利要求以添加这样的标记以更容易地区分权利要求中的这样的流控元件。

[0171] 本文所使用的“液体”是相对不可压缩的并且具有流动和符合保持物质的容器或通道的形状的能力的物质。液体可以是水基的,并且包括极性分子,极性分子呈现将液体保持在一起的表面张力。液体还可以包括例如在油基或非水物质中的非极性分子。应当理解,在本申请中提及的液体可以包括包含两种或更多种液体的组合的液体。例如,单独的试剂溶液可以随后组合以进行指定的反应。

[0172] 一个或更多个实现可以包括将生物材料(例如,模板核酸)保留在分析生物材料的指定位置处。如本文中所使用的,当关于生物材料使用时,术语“保留”包括将生物材料附着到表面或将生物材料限制在指定空间内。如本文中所使用的,当关于生物材料使用时,术语“固定”包括将生物材料附着到固体载体中的表面或固体载体上的表面。固定可以包括将生物材料以分子水平附接到该表面。例如,可以使用包括非共价相互作用(例如,静电力、范德华力以及疏水界面的脱水)和共价结合技术的吸附技术将生物材料固定到基底表面,其中功能团或接头促进将生物样品附接到该表面。将生物材料固定到基底的表面可以基于基底表面、携带生物材料的液体介质的特性以及生物材料自身的特性。在一些例子中,基底表面可以被功能化(例如,化学地或物理地改变)以促进将生物材料固定到基底表面。该基底表面可以首先被改变以具有结合至表面的官能团。然后,官能团可以结合到生物材料以将生物材料固定在其上。在某些情况下,生物材料可以通过凝胶固定在表面上。

[0173] 在一些实现中,核酸可以固定化到表面并使用桥式扩增进行扩增。另一种在表面上扩增核酸的有用方法是滚环扩增(RCA),例如,使用下面更详细阐述的方法。在一些实现中,核酸可以附接至表面并使用一个或更多个引物对被扩增。例如,一种引物可以在溶液中,并且另一种引物可以固定在表面上(例如,5'-附接)。举例来说,核酸分子可以与表面上的引物中的一个杂交,随后延伸所固定的引物以产生核酸的第一拷贝。溶液中的引物然后与核酸的第一拷贝杂交,核酸的第一拷贝可以使用核酸的第一拷贝作为模板进行延伸。任选地,在产生核酸的第一拷贝之后,原始核酸分子可以与表面上的第二固定化的引物杂交,并且可以同时或在溶液中的引物延伸之后延伸。在任何实现中,使用溶液中的固定引物和引物的重复延伸轮次(例如扩增)可用于提供核酸的多个拷贝。在一些实现中,生物材料可以与构造成在生物材料的扩增(例如,PCR)期间使用的反应成分一起被限定在预定空间内。

[0174] 本文阐述的一个或更多个实现可以构造成执行是扩增(或PCR)方案或包括扩增

(例如PCR)方案的测定方案。在扩增方案期间,可以改变储器或通道内生物材料的温度,以便扩增目标序列或生物材料(例如,生物材料的DNA)。举例来说,生物材料可以经历(1)大约95°C的预加热阶段大约75秒;(2)大约95°C的变性阶段大约15秒;(3)大约59°C的退火-延伸阶段(annealing-extension stage)大约45秒;以及(4)大约72°C的温度保持阶段大约60秒。实现可以执行多个扩增循环。应注意,上述循环仅描述了一个特定的实现,并且可选实现可以包括对扩增方案的修改。

[0175] 本文阐述的方法和系统可以使用具有多种密度中任何密度的特征的阵列,包括例如至少约10个特征/cm<sup>2</sup>、约100个特征/cm<sup>2</sup>、约500个特征/cm<sup>2</sup>、约1000个特征/cm<sup>2</sup>、约5000个特征/cm<sup>2</sup>、约10000个特征/cm<sup>2</sup>、约50000个特征/cm<sup>2</sup>、约100000个特征/cm<sup>2</sup>、约1000000个特征/cm<sup>2</sup>、约5000000个特征/cm<sup>2</sup>,或更高。本文阐述的方法和装置可以包括具有至少足以以这些密度中的一个或更多个来解析各个特征的分辨率的检测部件或设备。

[0176] 基础仪器102可以包括用户界面130,用户界面130构造成接收用于进行指定的测定方案的使用者输入和/或构造成使用者传送关于测定的信息。用户界面130可以与基础仪器102合并在一起。例如,用户界面130可以包括触摸屏,触摸屏附接到基础仪器102的外壳并且构造成识别来自使用者的触摸和相对于显示在触摸屏上的信息的触摸的位置。可选地,用户界面130可以相对于基础仪器102远离地定位。

[0177] II. 盒

[0178] 可移除盒200被构造成在盒室140处可分离地接合或可移除地联接到基础仪器102。如本文所用,当术语“可分离地接合”或“可移除地联接”(或类似术语)用于描述可移除盒200和基础仪器102之间的关系时。该术语旨在表示可移除盒200和基础仪器102之间的连接是可分离的,而不会破坏基础仪器102。因此,可移除盒200可以以电的方式可分离地接合到基础仪器102,使得基础仪器102的电触头不被破坏。可移除盒200可以以机械方式可分离地接合到基础仪器102,使得保持可移除盒200的基础仪器102的特征(例如盒室140)不会被破坏。该可移除盒200可以以流体的方式可分离地接合到基础仪器102,使得基础仪器102的端口不被破坏。例如,如果仅需要对部件进行简单的调整(例如,重新对齐)或者需要简单的更换(例如更换喷嘴),则不认为基础仪器102被“破坏”。当部件可彼此分离而无需过度的努力或在分离部件上花费大量时间时,部件(例如,可移除盒200和基础仪器102)可以是容易地可分离的。在一些实现中,可移除盒200和基础仪器102可以是容易地可分离的,而不破坏可移除盒200或基础仪器102。

[0179] 在一些实现中,可移除盒200可以在与基础仪器102在一起的时间段期间被永久地改变或部分地损坏。例如,容纳液体的容器可以包括箔盖,该箔盖被刺穿以允许液体流过系统100。在这种实现中,箔盖可能被损坏,使得损坏的容器将被另一个容器替换。在特定的实现中,可移除盒200是一次性盒,使得可移除盒200可以在单次使用之后被替换以及任选地被处理掉。类似地,可移除盒200的流通池可以是单独一次性的,使得流通池可以在单次使用后被替换并可选地被处理掉。

[0180] 在其它的实现中,可移除盒200可以在与基础仪器102接合时用于一个以上的环节,和/或可以从基础仪器102移除,重新装载试剂,并重新接合到基础仪器102以进行额外的指定反应。因此,在一些情况下,可移除盒200可被整修,使得同一个可移除盒200可以与不同的消耗品(例如,反应成分和生物材料)一起使用。在从位于客户设施处的基础仪器102

移除盒200之后,可以在制造设施处进行整修。

[0181] 盒室140可以包括狭槽、安装件、连接器接口和/或任何其他特征,以接收可移除盒200或其一部分,从而与基础仪器102相互作用。

[0182] 可移除盒200可以包括流控网络,该流控网络可以保持并引导流体(例如,液体或气体)通过其中。流控网络可以包括多个互连的流控元件,该多个互连的流控元件能够储存流体和/或允许流体流过其中。流控元件的非限制性示例包括通道、通道的端口、腔、储存设备、存储设备的储器、反应室、废物储器、检测室,用于反应和检测的多用途室等。例如,可消耗试剂部分210可以包括一个或更多个储存试剂的试剂阱或试剂室,并且可以是流控网络的一部分或耦合到流控网络。该流控元件可以以指定的方式流体地联接至彼此,使得系统100能够执行样品制备和/或分析。

[0183] 如本文所使用的,术语“流体地联接”(或类似术语)是指连接在一起的两个空间区域,使得液体或气体可以在该两个空间区域之间被引导。在一些例子中,流体联接允许流体在该两个空间区域之间被来回引导。在另一些例子中,流体联接是单向的,使得在该两个空间区域之间仅存在一个流动方向。例如,测定储器可以与通道流体地联接,使得液体可以从测定储器输送到通道中。然而,在一些实现中,可能不能将通道中的流体引导回至测定储器中。在特定的实现中,流控网络可以构造成接收生物材料并引导生物材料穿过样品制备和/或样品分析。流控网络可以将生物材料其它反应成分引导到废物储器。

[0184] 图2描绘了可消耗盒300的实现。可消耗盒可以是组合的可移除盒的一部分,例如图1的可移除盒200的可消耗试剂部分210;或者可以是单独的试剂盒。可消耗盒300可以包括外壳302和顶部304。外壳302可以包括非导电聚合物或其他材料,并且被形成为制造一个或更多个试剂室310、320、330。试剂室310、320、330的尺寸可以变化,以适应存储在其中的不同体积的试剂。例如,第一室310可以大于第二室320,第二室320可以大于第三室330。第一室310的尺寸适于容纳更大体积的特定试剂,例如缓冲试剂。第二室320的尺寸可以设置成容纳比第一室310更小体积的试剂,例如容纳裂解试剂的试剂室。第三室330的尺寸可以设置成容纳比第一室310和第二室320更小体积的试剂,例如容纳全功能含核苷酸试剂的试剂室。

[0185] 在图示的实现中,外壳302具有多个外壳壁或侧面350,在其中形成室310、320、330。在图示的实现中,外壳302形成至少基本上整体或单片的结构。在替代实现中,外壳302可以由一个或更多个子部件构成,这些子部件被组合以形成外壳302,例如用于室310、320和330的独立形成的隔室。

[0186] 一旦试剂被提供到相应的室310、320、330中,外壳302可以由顶部304密封。顶部304可以包括导电或非导电材料。例如,顶部304可以是铝箔密封件,其粘接到外壳302的顶面,以将试剂密封在其各自的室310、320、330内。在其他实现中,顶部304可以是塑料密封件,该塑料密封件粘接到外壳302的顶面,以将试剂密封在它们各自的室310、320、330内。

[0187] 在一些实现中,外壳302还可以包括标识器390。标识器390可以是射频识别(RFID)应答器、条形码、识别芯片和/或其他标识符。在一些实现中,标识器390可以嵌入外壳302中或者附着到外表面。标识器390可以包括用于可消耗盒300的唯一标识符的数据和/或用于可消耗盒300的类型的的数据。如本文所述,标识器390的数据可由基础仪器102或被配置用于加热可消耗盒300的单独设备读取。

[0188] 在一些实现中,可消耗盒300可以包括其他部件,例如阀、泵、流体线路、端口等。在一些实现中,可消耗盒300可以包含在另一个外壳内。

[0189] III. 系统控制器

[0190] 基础仪器102还可以包括系统控制器120,该系统控制器120被配置为控制可移除盒200和/或检测组件110中的至少一个的操作。系统控制器120可以利用专用硬件电路、板、DSP、处理器等的任何组合来实现。可选地,系统控制器120可以利用具有单个处理器或多个处理器的现成的PC来实现,其中功能操作分布在处理器之间。作为另一选择,可以利用混合配置来实现系统控制器120,在该混合配置中利用专用硬件来执行某些模块功能,而利用现成的PC等来执行其余的模块功能。

[0191] 系统控制器120可以包括多个电路模块,多个电路模块构造成控制基础仪器102和/或可移除盒200的某些部件的操作。这里的术语“模块”可以指被配置成执行特定任务的硬件设备。例如,电路模块可以包括流量控制模块,该流量控制模块被配置成控制通过可移除盒200的流控网络的流体流量。流量控制模块可以可操作地连接到阀致动器和/或系统泵。流量控制模块可以选择性地启动阀致动器和/或系统泵,以引导流体流过一个或多个路径和/或阻止流体流过一个或多个路径。

[0192] 系统控制器120还可以包括热控制模块。热控制模块可以控制热循环器或其他热部件,以提供和/或移除来自可移除盒200的样品制备区域和/或可移除盒200的任何其他区域的热能。在一个特定的例子中,热循环器可以根据PCR方案增加和/或降低生物材料经历的温度。

[0193] 系统控制器120还可以包括检测模块,其构造成控制检测组件110以获得关于生物材料的数据。如果检测组件110是可移除盒200的一部分,则检测模块可以通过直接有线连接或者通过接触阵列来控制检测组件110的操作。检测模块可以控制检测组件110在预定时间或预定时间段获取数据。举例来说,当生物材料具有附着于其上的荧光团时,检测模块可以控制检测组件110捕获可移除盒的流通池接收部分220的反应室的图像。在一些实现中,可以获得多个图像。

[0194] 任选地,系统控制器120可以包括分析模块,分析模块构造成分析数据以向系统100的使用者提供至少部分结果。例如,分析模块可以分析由检测组件110提供的成像数据。该分析可以包括识别生物材料的核酸序列。

[0195] 以上描述的系统控制器120和/或电路模块可以包括一个或多个基于逻辑的设备,包括一个或多个微控制器、处理器、精简指令集计算机(RISC)、专用集成电路(ASIC)、现场可编程门阵列(FPGA)、逻辑电路以及能够执行本文所描述的功能的任何其它电路。在一个实现中,系统控制器120和/或电路模块执行存储在计算机或机器可读介质中的一组指令,以便执行一个或多个测定方案和其他操作。该组指令可以信息源或物理存储元件的形式存储在基础仪器102和/或可移除盒200内。由系统100执行的方案可以用于执行,例如,机器写入DNA或以其他方式合成DNA(例如,将二进制数据转换成DNA序列,然后合成DNA链或代表二进制数据的其他多核苷酸)、DNA或RNA的定量分析、蛋白质分析、DNA测序(例如,合成测序(SBS))、样品制备和/或用于测序的片段库的制备。

[0196] 指令集可以包括指令系统100执行诸如本文所描述的各种实现的方法和过程的特定操作的各种命令。指令集可以是软件程序的形式。如本文所使用的,术语“软件”和“固件”

是可互换的,并且包括储存在包括RAM存储器、ROM存储器、EPROM存储器、EEPROM存储器以及非易失性RAM (NVRAM) 存储器的存储器中以由计算机执行的任何计算机程序。上面的存储器类型仅仅是示例,并且因此并不限制可用于储存计算机程序的存储器的类型。

[0197] 软件可以是各种形式,例如系统软件或应用软件。此外,软件可以是单独程序的集合或者更大程序或程序模块的一部分内的程序模块的形式。软件还可以包括面向对象编程形式的模块化编程。在获得检测数据之后,检测数据可以由系统100自动处理,响应于使用者输入而被处理,或响应于由另一处理机器做出的请求(例如,通过通信链路的远程请求)而被处理。

[0198] 系统控制器120可以经由可以是硬连线的或无线的通信链路连接到系统100的其它部件或子系统。系统控制器120还可以通信地连接到场外系统或服务器。系统控制器120可以接收来自用户界面130的使用者输入或命令。用户界面130可以包括键盘、鼠标、触摸屏面板和/或语音识别系统等。

[0199] 系统控制器120用于提供处理能力,例如储存、译解和/或执行软件指令,以及控制系统100的整体操作。系统控制器120可以被配置并且编程以控制各种部件的数据方面和/或功率方面。尽管系统控制器120在图1中被表示为单个结构,但是应理解,系统控制器120可以包括在不同位置遍布系统100分布的多个单独的部件(例如,处理器)。在一些实现中,一个或更多个部件可以与基础仪器102集成,并且一个或更多个部件可以相对于基础仪器102远离地定位。

#### [0200] IV. 流通池

[0201] 图3-4描绘了可用于系统100的流通池400的例子。该示例的流通池包括限定多个细长流动通道410的主体,该流动通道410凹入主体402的上表面404下方。流动通道410通常彼此平行,并且基本上沿着主体402的整个长度延伸。虽然示出了五个流动通道410,但是流通池400可以包括任何其他合适数量的流动通道410,包括多于或少于五个流动通道410。该示例的流通池400还包括一组入口端口420和一组出口端口422,每个端口420、422与相应的流动通道410相关联。因此,每个入口端口420可用于将流体(例如,试剂等)传送到相应的通道410;而每个出口端口422可用于传送来自相应流动通道410的流体。

[0202] 在一些形式中,流通池400直接集成到可移除盒200的流通池接收部分220中。在一些其它形式中,流通池400可移除地与可移除盒200的流通池接收部分220连接。在流通池400直接集成到流通池接收部分220中或者可移除地与流通池接收部分220耦合的形式中,流通池400的流动通道410可以经由入口端口420从可消耗试剂部分210接收流体,入口端口420可以与存储在可消耗试剂部分210中的试剂流体地联接。当然,流动通道410可以经由端口420、422与各种其他流体源或储器等耦合。作为另一个说明性变型,一些形式的可消耗盒300可以被配置成可移除地接收或以其他方式集成流通池400。在这种形式中,流通池400的流动通道410可以通过入口端口420接收来自试剂室310、320、330的流体。根据这里的教导,流通池400可以结合到系统100中的其他合适的方式对于本领域技术人员来说是明显的。

[0203] 图4更详细地示出了流通池400的流动通道410。如图所示,流动通道410包括形成在流动通道410的基面412中的多个阱430。如下文将更详细描述,每个阱430被配置为包含DNA链或其他多核苷酸,例如机器写入的多核苷酸。在一些形式中,每个阱430具有圆柱形构造,具有大致圆形的横截面轮廓。在一些其他形式中,每个阱430具有多边形(例如,六边

形、八边形等) 截面轮廓。可选地, 阱430可以具有任何其他合适的构造。还应当理解, 阱430可以以任何合适的图案布置, 包括但不限于网格图案。

[0204] 图5示出了流通池500内的通道的一部分, 该流通池500是流通池400的变型的一个例子。换句话说, 图5中描绘的通道是流通池400的流动通道410的变型。该流通池500可操作来读取固定到流通池500中的阱530的底面534的多核苷酸链550。仅作为示例, 多核苷酸链550被固定的底面534可以包括用叠氮基封端的共嵌段聚合物。仅作为进一步的例子, 这种聚合物可以包括根据于2015年4月21日发布的题为“Polymer Coatings”的美国专利第9, 012, 022号的至少一些教导提供的聚(N-(5-叠氮基乙酰胺基戊基) 丙烯酰胺-共-丙烯酰胺) (PAZAM) 涂层, 该专利的全部内容通过引用结合于此。这种聚合物可以掺入本文所述的各种流通池中的任何一种。

[0205] 在本示例中, 阱530由流通池500的基面512提供的间隙空间514分隔开。每个阱530具有侧壁532和底面534。该示例中的流通池500可操作以在每个阱530下方提供图像传感器540。在一些形式中, 每个阱530具有至少一个对应的图像传感器540, 图像传感器540相对于阱530固定就位。每个图像传感器540可以包括CMOS图像传感器、CCD图像传感器或任何其他合适类型的图像传感器。仅作为示例, 每个阱530可以具有一个相关联的图像传感器540或多个相关联的图像传感器540。作为另一种变型, 单个图像传感器540可以与两个或更多个阱530相关联。在一些形式中, 一个或更多个图像传感器540相对于阱530移动, 使得单个图像传感器540或单组图像传感器540可以相对于阱530移动。作为又一变型, 流通池500可以相对于单个图像传感器540或单组图像传感器540移动, 它们可以至少基本上固定在原位。

[0206] 每个图像传感器540可以直接结合到流通池500中。可选地, 每个图像传感器540可以直接结合到诸如可移除盒200的盒中, 流通池500集成到盒中或者以其他方式与盒耦合。作为又一说明性变型, 每个图像传感器540可以直接结合到基础仪器102中(例如, 作为上述检测组件110的一部分)。无论图像传感器540位于何处, 图像传感器540都可以集成到包括其他部件(例如, 控制电路等)的印刷电路中。在一个或更多个图像传感器540不直接结合到流通池500中的形式中, 流通池500可以包括光学透射特征(例如, 窗口等)允许一个或更多个图像传感器540捕获由与多核苷酸链550相关联的一个或更多个荧光团发射的荧光, 多核苷酸链550固定到流通池500中的阱530的底面534, 如下文更详细描述。还应该理解, 各种光学元件(例如, 透镜、光波导等)可以置于阱530的底面534和相应的图像传感器540之间。

[0207] 同样如图5所示, 光源560可操作以将光562投射到阱530中。在一些形式中, 每个阱530具有至少一个对应的光源560, 光源560相对于阱530固定在原位。仅作为示例, 每个阱530可以具有一个相关联的光源560或多个相关联的光源560。作为另一种变型, 单个光源560可以与两个或更多个阱530相关联。在一些其他形式中, 一个或更多个光源560相对于阱530移动, 使得单个光源560或单组光源560可以相对于阱530移动。作为又一变型, 流通池500可以相对于单个光源560或单组光源560移动, 这些光源可以基本上固定在原位。仅作为示例, 每个光源560可以包括一个或更多个激光器。在另一个例子中, 光源560可以包括一个或更多个二极管。

[0208] 每个光源560可以直接结合到流通池500中。可选地, 每个光源560可以直接结合到诸如可移除盒200的盒中, 流通池500集成到盒中或者以其他方式与盒耦合。作为又一说明性变型, 每个光源560可以直接结合到基础仪器102中(例如, 作为上述检测组件110的一部

分)。在一个或多个光源560不直接结合到流通池500中的形式中,流通池500可以包括光学透射特征(例如,窗口等)允许阱530接收由一个或多个光源560发射的光,从而使得光能够到达固定到阱530的底面534的多核苷酸链550。还应该理解,各种光学元件(例如,透镜、光波导等)可以置于阱530和相应的光源560之间。

[0209] 如本文别处所述以及如图6的框590所示,DNA读取过程可以从在目标阱530中进行测序反应开始(例如,根据于2016年9月27日发布的题为“Methods and Compositions for Nucleic Acid Sequencing”的美国专利第9,453,258号的至少一些教导,其全部内容通过引用结合于此)。接下来,如图6的框592所示,光源560在目标阱530上被激活,从而照亮目标阱530。投射光562可以导致与多核苷酸链550相关的荧光团发荧光。因此,如图6的框594所示,相应的图像传感器540可以检测从与多核苷酸链550相关联的一个或多个荧光团发射的荧光。基础仪器102的系统控制器120可以驱动光源560发光。基础仪器102的系统控制器120还可以处理从图像传感器540获得的图像数据,该图像数据表示来自阱530中的多核苷酸链550的荧光发射分布。使用来自图像传感器540的图像数据,并且如图6的框596所示,系统控制器120可以确定每个多核苷酸链550中的碱基序列。仅作为示例,该方法和装备可用于绘制基因组或以其他方式确定与天然存在的生物体相关的生物信息,其中DNA链或其他多核苷酸从天然存在的生物体获得或以其他方式基于天然存在的生物体。或者,可以利用上述过程和装备来获得存储在机器写入的DNA中的数据,这将在下面更详细地描述。

[0210] 仅作为进一步的例子,当执行图6所示的上述程序时,时空测序反应可以利用一种或更多种化学和成像事件或步骤来区分在测序反应期间掺入生长的核酸链中的多种分析物(例如,四种核苷酸);或者,在具有四种不同核苷酸的混合物中可以检测到少于四种不同的颜色,同时仍然导致四种不同核苷酸的确定(例如,在测序反应中)。一对核苷酸类型可以在同一波长检测,但基于该对中的一个成员相比于另一个的强度中的差异区分,或基于对该对的一个成员的改变(例如,经由化学修饰、光化学修饰或物理修饰)导致表现信号相比于该对的另一个成员检测的信号出现或消失来区分。

[0211] V. 机器写入生物材料

[0212] 在一些实现中,诸如图1中所示的系统100的系统100可以被配置成合成生物材料(例如,多核苷酸,诸如DNA)以编码数据,该数据随后可以通过诸如上述那些的测定的执行来检索。在一些实现中,这种类型的编码可以通过给核苷酸碱基赋值来执行(例如,二进制值,例如0或1,三进制值,例如0、1或2,等等),将待编码的数据转换成相关值的串(例如,使用ASCII编码方案将文本消息转换成二进制串),然后创建由核苷酸组成的一个或多个多核苷酸,所述核苷酸在对应于通过转换数据获得的串的序列中具有碱基。

[0213] 在一些实现中,可以使用具有如图7所示配置的阱阵列630的流通池400的形式来进行这种多核苷酸的产生。图7示出了流通池600内的通道的一部分,其是流通池400的变型的例子。换句话说,图7中描绘的通道是流通池400的流动通道410的变型。在该示例中,每个阱630凹入流通池600的基面612之下。因此,阱630通过间隙空间614彼此隔开。仅作为示例,阱630可以沿着流通池600的基面612以网格或任何其他合适的图案布置。该示例的每个阱630包括侧壁632和底面634。该示例的每个阱630还包括位于阱630的底面634上的相应电极组件640。在一些形式中,每个电极组件640仅包括单个电极元件。在一些其他形式中,每个电极组件640包括多个电极元件或段。术语“电极”和“电极组件”在这里应该理解为是可互



换的。

[0214] 基础仪器102可操作以独立地激活电极组件640,使得一个或更多个电极组件640可处于激活状态,而一个或更多个其他电极组件640不处于激活状态。在一些形式中,使用CMOS设备或其他设备来控制电极组件640。这种CMOS设备可以直接集成到流通池600中,可以集成到流通池600所包含的盒(例如,盒200)中,或者直接集成到基础仪器102中。如图7所示,每个电极组件640沿着底面634的整个宽度延伸,终止于相应阱630的侧壁632。在其他形式中,每个电极组件640可以仅沿着底面634的一部分延伸。例如,一些形式的电极组件640可以相对于侧壁632在内部终止。虽然每个电极组件540在图5中被示意性地描绘为单个元件,但是应当理解,每个电极组件540实际上可以由多个离散的电极形成,而不仅仅由一个单个电极组成。

[0215] 如图7所示,通过激活相关阱630的电极组件640来电化学产生酸,可以在单个阱630中产生特定的多核苷酸链650,该酸可以将阱630中的多核苷酸链650的端基脱保护。仅举例来说,多核苷酸链650可以使用一端具有化学物质如硅烷化学物质而另一端具有DNA合成相容化学物质(例如,酶结合的短寡核苷酸)的接头化学连接到阱630底部的表面。

[0216] 为了便于试剂交换(例如,解封闭剂的传输),在该示例中,每个电极组件640和每个阱630的底面634可以包括至少一个开口660。开口660可以与在阱630下方、底面634下方延伸的流动通道662流体地联接。为了提供穿过电极组件640的这种开口660,电极组件640可以是环形的,可以放置在象限中,可以放置在阱630的周边或侧壁632上,或者可以以其他合适的方式放置或成形,以避免干扰试剂交换和/或光的通过(例如,可以在涉及荧光发射检测的测序过程中使用)。在其他实现中,试剂可以在没有开口660的情况下提供到流通池600的流动通道中。应当理解,开口660可以是可选的,并且在一些形式中可以省略。类似地,流动通道662可以是可选的,并且在一些形式中可以省略。

[0217] 图9示出了电极组件640可以采取的形式的例子。在该示例中,电极组件640包括四个分立的电极段642、644、646、648,它们一起限定了环形形状。电极段642、644、646、648因此被配置为环的离散但相邻的象限。每个电极段642、644、646、648可以被配置成提供与特定核苷酸唯一相关的预定电荷。例如,电极段642可以被配置成提供与腺嘌呤唯一相关的电荷;电极段644可以被配置成提供与胞嘧啶唯一相关的电荷;电极段646可以被配置成提供与鸟嘌呤唯一相关的电荷;并且电极段648可以被配置成提供与胸腺嘧啶唯一相关的电荷。当这四种核苷酸的混物流经阱630上方的流动通道时,电极段642、644、646、648的激活可导致来自该流动通道的相应核苷酸粘附到链650上。因此,当电极段642被激活时,它可以实现腺嘌呤向链650的写入;当电极段644被激活时,它可以实现胞嘧啶向链650的写入;当电极段646被激活时,它可以实现鸟嘌呤向链650的写入;并且当电极段648被激活时,它可以实现胸腺嘧啶向链650的写入。这种写入可以通过激活的电极段642、644、646、648将酶的抑制剂与与激活的电极段642、644、646、648相关联的像素杂交来提供。虽然电极段642、644、646、648在图9中显示为形成环形,但是应当理解,电极段642、644、646、648可以形成任何其他合适的形状。在其他实现中,单个电极可用于电极组件640,并且电荷可被调节以掺入要写入DNA链或其他多核苷酸的各种核苷酸。

[0218] 作为另一个例子,电极组件640可以被激活以提供局部的(例如,位于电极组件640设置于其中的阱630内)电化学产生的pH变化;和/或局部地电化学产生部分(例如,还原或

氧化试剂),以从核苷酸中去除嵌段。作为另一种变体,不同的核苷酸可以具有不同的嵌段;并且这些嵌段可以基于传送到阱630的光(例如,从光源560投射的光562)的波长被光切割。作为另一种变体,不同的核苷酸可以具有不同的嵌段;并且这些嵌段可以基于某些其他条件被切割。例如,可以基于还原条件加上高局部pH或低局部pH的组合来移除四个嵌段中的一个;基于氧化条件加上高局部pH或低局部pH的组合,可以除去四个嵌段中的另一个;基于光和高局部pH的组合,可以去除四个嵌段中的另一个;以及基于光和低局部pH的组合,可以去除四个嵌段中的另一个。因此,可以同时掺入四个核苷酸,但是响应于四组不同的条件发生选择性去嵌段。

[0219] 电极组件640进一步在电极段642、644、646、648的排列中心限定开口660。如上所述,该开口660可以为流动通道662和阱630之间的流体连通提供路径,从而允许流经流动通道662的试剂等到达阱630。同样如上所述,一些变型可以省略流动通道662并以某种其他方式(例如,通过被动扩散等)提供试剂等到阱630的连通。不管流体是否通过开口660连通,开口660可以在读取周期期间提供通过阱630底部的光传输路径,如本文所述。在一些形式中,开口660可以是可选的,因此可以省略。在省略开口660的形式中,流体可以经由位于阱630上方或相对于阱630定位的一个或更多个流动通道连通到阱630。此外,在读取周期期间,可能不需要开口660来提供穿过阱630底部的光传输路径。例如,如下文关于流通池601所述,电极组件640可以包括光学透明材料(例如,光学透明导电膜(TCF)等),并且流通池600本身可以包括光学透明材料(例如,玻璃),使得电极组件640和形成流通池600的材料可以允许与机器写入的多核苷酸链650相关联的一个或更多个荧光团发射的荧光到达阱630下方的图像传感器540。

[0220] 图8显示了可在流通池600中用于机器写入多核苷酸或其他核苷酸序列的过程的例子。在该过程开始时,如图8的第一框690所示,核苷酸可以通过阱630流入流通池600。如图8中的下一个框692所示,然后可以激活电极组件640,以将第一核苷酸写入目标阱630底部的引物。如图8的下一个框694所示,终止子随后可从目标阱630中的刚刚写入的第一个核苷酸被切割掉。鉴于本文的教导,终止子可以从第一个核苷酸上切割下来的各种合适的方式对于本领域技术人员来说是明显的。一旦终止子被从第一个核苷酸上切下,如图8的下一个框696所示,电极组件640可以被激活以将第二个核苷酸写入第一个核苷酸。虽然图8中未示出,但是可以将终止子从第二个核苷酸上切下,然后将第三个核苷酸写入第二个核苷酸,以此类推,直到已经写入了所需的核苷酸序列。

[0221] 在一些实现中,通过诸如DNA的生物材料的合成进行的数据编码可以以其他方式进行。例如,在一些实现中,流通池600可能完全没有电极组件640。例如,去嵌段试剂可以通过开口660从流动通道662选择性地连通到阱630。这可以消除电极组件640选择性激活核苷酸的需要。作为另一个例子,阱630的阵列可以暴露于包含可以用于编码数据的所有核苷酸碱基的溶液中,然后通过使用来自空间光调制器(SLM)的光为单个阱630选择性地激活单个核苷酸。作为另一个例子,在一些实现中,可以为单个碱基分配组合值(例如,腺嘌呤可以用于编码二联体00,鸟嘌呤可以用于编码二联体01,胞嘧啶可以用于编码二联体10,胸腺嘧啶可以用于编码二联体11),以增加所产生的多核苷酸的存储密度。根据本公开,其他示例也是可能的,并且对于本领域技术人员来说将是明显的。因此,上述合成生物材料如DNA来编码数据的描述应该理解为仅仅是说明性的;不应被视为限制。

[0222] VI. 读取机器写入的生物材料

[0223] 在流通池600的一个或多个阱630中的多核苷酸链650已经被机器写入之后,多核苷酸链650可以随后被读取以提取机器写入的多核苷酸链650中存储的任何数据或其他信息。这种读取过程可以使用诸如图5所示和上述的布置来执行。换句话说,一个或多个光源560可用于照射与机器写入的多核苷酸链650相关联的一个或多个荧光团;并且一个或多个图像传感器540可以用于检测由与机器写入的多核苷酸链650相关联的被照亮的的一个或多个荧光团发射的荧光。由与机器写入的多核苷酸链650相关联的被照射的一个或多个荧光团发射的光的荧光分布可以被处理以确定机器写入的多核苷酸链650中的碱基序列。机器写入的多核苷酸链650中的这种确定的碱基序列可以被处理以确定存储在机器写入的多核苷酸链650中的数据或其他信息。

[0224] 在一些形式中,机器写入的多核苷酸链650在储存期间保留在包含阱630的流通池600中。当需要读取机器写入的多核苷酸链650时,流通池600可以允许直接从流通池读取机器写入的多核苷酸链650。仅作为示例,包含阱630的流通池600可以被容纳在包含光源560和/或图像传感器540的盒(例如,盒200)或基础仪器102中,使得机器写入的多核苷酸链650被直接从阱630读取。

[0225] 作为另一个说明性例子,包含阱630的流通池可以直接结合光源560或图像传感器540中的一个或两个。图10示出了流通池601的示例,其包括具有电极组件640的阱630、一个或多个图像传感器540和控制电路670。类似于图5中描绘的流通池500,该示例的流通池601可操作来接收从光源560投射的光562。该投射光562可以导致与机器写入的多核苷酸链650相关联的一个或多个荧光团发荧光;并且相应的图像传感器540可以捕获从与机器写入的多核苷酸链650相关联的一个或多个荧光团发出的荧光。

[0226] 如上在流通池500的情况下所述,流通池601的每个阱650可以包括其自己的图像传感器540和/或其自己的光源560;或者这些部件可以如上所述以其他方式配置和布置。在本示例中,从与机器写入的多核苷酸链650相关联的一个或多个荧光团发射的荧光可以经由开口660到达图像传感器540。此外,或者替代地,电极组件640可以包括光学透明材料(例如,光学透明导电膜(TCF)等),并且流通池601本身可以包括光学透明材料(例如,玻璃),使得电极组件640和形成流通池601的材料可以允许从与机器写入的多核苷酸链650相关联的一个或多个荧光团发射的荧光到达图像传感器540。此外,各种光学元件(例如,透镜、光波导等)可以置于阱650和相应的图像传感器之间,以确保图像传感器540仅接收从与所需阱630的机器写入的多核苷酸链650相关的一个或多个荧光团发出的荧光。

[0227] 在本示例中,控制电路670直接集成到流通池601中。仅作为示例,控制电路670可以包括CMOS芯片和/或其他印刷电路配置/部件。控制电路670可以与图像传感器540、电极组件640和/或光源560连通。在这种情况下,“连通”意味着控制电路670与图像传感器540、电极组件640和/或光源560电连通。例如,控制电路670可操作来接收和处理来自图像传感器540的信号,该信号代表由图像传感器540拾取的图像。本文中的“连通”还可以包括控制电路670向图像传感器540、电极组件640和/或光源560提供电能。

[0228] 在一些形式中,每个图像传感器540具有相应的控制电路670。在一些其他形式中,控制电路670与流通池601中的几个(如果不是全部的话)图像传感器耦合。鉴于本文的教导,可用于实现这一点的各种合适的部件和配置对于本领域技术人员来说将是明显的。还

应当理解,除了集成到流通池601中之外或代替集成到流通池601中,控制电路670可以整体或部分集成到盒(例如,可移除盒200)中和/或基础仪器102中。

[0229] 作为又一个说明性例子,不管是使用像图7的流通池600那样的只写流通池还是像图10的流通池601那样的读写流通池,机器写入的多核苷酸链650可以在合成后从阱630转移。这可能发生在合成完成后不久,就在机器写入的多核苷酸链650被读取之前,或在任何其他合适的时间。在这样的形式中,机器写入的多核苷酸链650可以被转移到只读流通池,如图5所示的流通池500;然后在只读流通池500中被读取。或者,可以使用任何其他合适的设备或过程。

[0230] 在一些实现中,读取通过生物材料合成编码的数据可以通过确定储存感兴趣的合成链650的阱630,然后使用诸如前述的技术(例如,合成测序)对那些链650进行测序来实现。在一些实现中,为了便于读取存储在核苷酸序列中的数据,当存储数据时,可以用显示合成编码该数据的链650所位于的阱630的信息来更新索引。例如,当被配置成合成能够存储高达256位数据的串650的系统100的实现被用于存储一兆位(1,048,576位)文件时,系统控制器120可以执行如下步骤:1)将文件分成4,096个256位段;2)识别流通池600、601中当前未用于存储数据的4096个阱630的序列;3)将4,096段写入4,096个阱430、530;4)更新索引以指示从第一个识别的阱630开始并在最后一个识别的阱630结束的序列正被用于存储该文件。随后,当做出读取文件的请求时,可以使用索引来识别包含相关链650的阱630,可以对来自那些阱630的链650进行测序,并且可以将序列组合并转换成适当的编码格式(例如,二进制),然后可以返回该组合并转换的数据作为对读取请求的响应。

[0231] 在一些实现中,可以以其他方式读取先前通过生物材料合成编码的数据。例如,在一些实现中,如果要写入对应于4,096个阱630的文件,而不是识别4,096个顺序阱630以将其写入,则控制器可以识别4,096个阱630,然后在那些阱630没有形成连续序列的情况下,用对应于该文件的多个位置更新索引。作为另一个例子,在一些实现中,系统控制器120可以将阱630分组在一起(例如,分成128个阱630的分组),而不是识别单个阱630,从而减少与存储位置数据相关联的开销(即,通过将寻址要求从每个阱630一个地址减少到每组阱630一个地址)。作为另一个例子,在存储反映其中已经合成了DNA链或其他多核苷酸的阱630的位置的数据的实现中,该数据可以以各种方式存储,例如序列标识符(例如,阱1、阱2、阱3等)或坐标(例如,阵列中阱的位置的X和Y坐标)。

[0232] 作为另一个例子,在一些实现中,不是从其中合成链650的阱630读取链650,而是可以从其他位置读取链650。例如,链650可被合成以包括地址,然后从阱630切割并存储在管中以供以后检索,在此期间,所包括的地址信息可用于识别对应于特定文件的链650。作为另一个说明性的例子,可以使用聚合酶从表面复制链650,然后洗脱并储存在试管中。或者,可以使用与DNA链或其它多核苷酸杂交的生物素化寡核苷酸,并在分配在阱630中的链霉亲和素珠上捕获延伸产物,将链650复制到珠上。根据本公开,其他示例也是可能的,并且对于本领域技术人员来说将是明显的。因此,上述检索通过合成生物材料编码的数据的描述应该理解为仅仅是说明性的;不应被视为限制。

[0233] 在此描述的实现可以利用用于流通池表面的聚合物涂层,例如在于2015年4月21日发布的题为“Polymer Coatings”的第9,012,022号美国专利中描述的,其全部内容通过引用结合于此。本文所述的实现可以利用一种或更多种具有可检测标记和可切割接头的标

记核苷酸,例如于2008年8月19日发布的题为“Labelled Nucleotide Strands”的第7,414,116号美国专利中所述的那些,其全部内容通过引用结合于此。例如,本文所述的实现可利用可切割的接头,该接头可通过与水溶性磷或水溶性含过渡金属的催化剂接触而被切割,该催化剂具有作为可检测标记的荧光团。本文所述的实现可以使用双通道检测方法检测多核苷酸的核苷酸,例如于2016年9月27日发布的题为“Methods and Compositions for Nucleic Acid Sequencing”的第9,453,258号美国专利中所述的方法,其全部内容通过引用结合于此。例如,本文所述的实现可以利用基于荧光的SBS方法,该方法具有在第一通道中检测到的第一核苷酸类型(例如,dATP具有当被第一激发波长激发时在第一通道中检测到的标记),在第二通道中检测到的第二核苷酸类型(例如,dCTP具有当被第二激发波长激发时在第二通道中检测到的标记),在第一和第二通道中都检测到的第三核苷酸类型(例如,当被第一和/或第二激发波长激发时,dTTP具有至少一个在两个通道中被检测到的标记),和第四种核苷酸类型,其缺少在任一通道中未被检测到或最低限度被检测到的标记(例如,无标记的dGTP)。这里描述的盒和/或流通池的实现可以根据以下描述的一个或多个教导来构造:发布于2014年12月9日的题为“Biosensors for Biological or Chemical Analysis and Systems and Methods for Same”的第8,906,320号美国专利,其全部内容通过引用结合于此;题为“Gel Patterned Surfaces”,于2016年12月6日发布的第9,512,422号美国专利,其全部内容通过引用结合于此;题为“Biosensors for Biological or Chemical Analysis and Methods of Manufacturing the Same”,于2019年4月9日发布的第10,254,225号美国专利,其全部内容通过引用结合于此;和/或题为“Cartridge Assembly”于2018年5月3日公布的第2018/0117587号美国公开,其全部内容通过引用结合于此。

[0234] VII. 用于多核苷酸硬盘驱动器设备的系统和方法

[0235] 随着向多核苷酸存储设备写入数据(例如合成)和从多核苷酸存储设备读取数据(例如测序)期间所使用的装置和材料的成本、速度和可用性的提高,在各种环境中可能会达到这样一个点,其中基于DNA的存储是基于常规方法(例如磁性、电子)的存储的可行的(如果不是优选的)补充或替代。例如,在数据中心规模的长期归档场景中,基于DNA的存储可以提供许多优势,如稳定性、数据密度和持久性。这些优势可以抵消基于DNA的存储所需的系统和仪器的相对成本和复杂性,因为几个这样的DNA存储系统可以服务于单个数据中心。

[0236] 对于这种情况,提供一种可以与DNA存储系统耦合或安装用于写入和读取的盒或模块,然后从DNA存储系统卸载下来并重新定位到长期存档区域进行存储,可以提供许多优点。例如,提供专门构建的模块允许在模块中包含额外的功能,以改善读取和写入过程,以及改善在模块中存储数据的机器写入的DNA的长期稳定性。作为另一个例子,提供单个模块,从用户的角度来看,该模块似乎是耐用的封闭系统,在外观和功能上类似于许多传统的便携式存储驱动器和设备,可以提高用户对DNA存储设备的可接受性和可用性。例如,一些用户可能会错误地认为,与传统的存储方法相比,机器写入的DNA是脆弱的或不可靠的,特别是当DNA存储设备看起来过于复杂或流通池表面和存储阱向外可见时。

[0237] 上述优点的真实例子可以在磁盘存储驱动器的情况下看到,磁盘存储驱动器包含以七千RPM或更高速度旋转的磁盘,而读/写头以仅几纳米的“飞行高度”漂浮在盘上方。使

用这种设备,盘片上薄如指纹甚至烟雾颗粒的障碍物可能导致与读取头的灾难性碰撞以及存储数据的完全或部分丢失。然而,由于这种复杂性被保护在专门开发的机箱内,并且通常用户看不到,磁盘硬盘驱动器仍然是存储的流行选择,并且被认为提供了健壮、可靠的存储,尽管它们很复杂。

#### [0238] A. 示例性DNA存储系统

[0239] 当在此描述时,可操作来读取编码为多核苷酸(例如DNA)的数字数据、将数字数据编码和写入DNA或两者的系统可被称为用于DNA存储的系统,或“DNA存储系统”。应当理解,这样的系统可以包括各种部件和设备,这些部件和设备可以被组装成单件装备(例如,可以被组装并可通信地耦合在壳体内),或者可以是单独的装备,这些装备可以被连接、布置或两者兼有,以便提供所描述的特征。

[0240] 作为示例,图11示出了DNA存储系统1300的示例的示意图。DNA存储系统1300包括一组仪器1301和存储设备或模块1320,它们可以与DNA存储系统1300集成在一起,或者在一些实现中,通过诸如模块接收器1321的模块接口与DNA存储系统1300可移除地耦合。该组仪器1301可以对应于上述基础仪器102。该组仪器1301可以组装在单个装备内,或者可以是一个或多个分开的装备,它们被布置、连接或两者兼有,以便提供所描述的功能。该组仪器1301包括存储控制器1302,该存储控制器1302可以包括被配置为存储和执行指令以操作该组仪器1301的一个或多个处理器和存储器。该组仪器1301还包括测序设备1304、合成设备1306、流控设备1308和电接口1310。

[0241] 在一些实现中,存储控制器1302、测序设备1304、合成设备1306、流控设备1308和电接口1310可以是其间具有一个或多个流控、电或机械接口的独立设备。在其他实现中,存储控制器1302、测序设备1304、合成设备1306、流控设备1308和电接口1310可以集成到单个设备中,其中测序设备1304、合成设备1306、流控设备1308和电接口1310中的每一个形成其子部件。

[0242] 测序设备1304可操作以读取在存储设备1320的一个或多个阱中编码和存储为多核苷酸(例如DNA)的数据,并且可包括诸如成像设备、光学传感器、照明设备(例如LED、其他照明器)和可用于检测储存在阱中的DNA的特性的其他设备的特征(例如上述与SBS相关的过程和设备,其中与单个核苷酸相关的荧光标记或标签可由光学传感器检测)。

[0243] 合成设备1306可操作来在存储设备1320的一个或多个阱内合成具有特定核苷酸排列的多核苷酸。合成设备1306包括或与单个核苷酸或其他生物材料的储存器和可操作以将输入生物材料传送到一个或多个阱的输入递送设备耦合。在一些实现中,这可以包括一组位于阱附近的电极,其可操作以将特定的核苷酸吸引到特定的阱,同时输入递送设备通过入口端口420向流动通道410提供核苷酸载液或核苷酸写入试剂。在一些实现中,这可以包括核苷酸注射头,其可以定位在期望的阱附近,使得一个或多个核苷酸可以以期望的顺序释放。

[0244] 流控设备1308可包括本文所述的与流控相关的任何设备或特征,例如流控网络、泵、阀和可操作以在期望的体积和压力下向一个或多个流动通道410或一个或多个流动通道410上的特定位置提供期望的流体类型的其他部件。在一些实现中,流控设备1308可以包括电润湿特征,该电润湿特征可操作以将期望体积的流体精确地引导至期望位置,而不是用流体淹没流动通道410。在一些实现中,流控设备1308可用于为DNA存储系统1300的

一个或更多个部件提供热控制或调节。作为一个例子,在由流控设备1308传输的流体需要在与存储设备1320一起使用之前被加热的情况下,流体可以被引导通过存储控制器1302的部分(例如,靠近处理器、存储器和散热器),以便将热量从那些部件吸入流体。当被转移的流体在使用前需要冷却时,流体可以通过DNA存储设备1300的冷却系统。可包括在DNA存储系统1300中的其他类似特征包括可被定位成加热或冷却流体、或加热或冷却存储设备1320本身的热控制设备,并且可包括吹气温度控制设备、固态热电板和其他设备。

[0245] 流控设备1308提供的流体可以包括用测序设备1304和合成设备1306执行的各种过程中产生和使用的流体试剂,并且还可以包括非功能性流体,例如用于冲洗和清洁DNA存储系统1300的一个或更多个部件的蒸馏水。测序设备1304使用的试剂可以不同于合成设备1306使用的试剂,并且每个设备本身可以在合成和测序的不同部分使用一种或更多种不同的试剂。当在本文中使用时,在测序操作期间可以提供的任何变化的试剂可以统称为核苷酸读取试剂,而在合成操作期间可以提供的任何变化的试剂可以统称为核苷酸写入试剂。

[0246] 在一些实现中,例如在存储设备1320包括需要电力或数据传输的特征的情况下,电接口1310可以包括有线导电连接,可以包括能够与存储设备1320交换电力、数据或两者的无线收发器设备(例如,RFID、NFC、蓝牙、光发射器、感应充电设备),其中存储设备1320包括需要电力或数据传输的特征。

[0247] 在存储设备1320是可移除的情况下,DNA存储系统1300还可以包括一个或更多个特征,以在使用期间将存储设备1320相对于一组仪器1301耦合和静态定位。这些特征可以包括模块接收器1321,其可以是存储设备1320可以安置在其中的槽,以及引导特征(例如,轨道)和锁定特征,以高度精确和固定地定位存储设备1320,使得一组仪器1301中的一个或更多个可重复地和自动地定位,以与其相应的接口相互作用。在其他实现中,存储设备1320可以包括与该组仪器1301的部件接口的流通池1322。

[0248] 应该理解的是,DNA存储系统1300仅仅是一个例子,根据本公开,许多变化是可能的,并且对于本领域技术人员来说是明显的。作为示例,一组仪器1301和存储设备1320可以具有比所示更少的部件或更多的部件。作为另一个例子,存储设备1320的一些实现可以包括该组仪器1301的部件,例如其中测序设备1304的多个电极被集成在存储设备1320自身的内部流通池之上或之内。在这种情况下,测序设备1304与一组仪器1301配对的部分可以包括导电开关网络,该网络允许电信号产生并传输到安装在流通池内或流通池上的期望电极,这将在下面更详细地描述。

[0249] B. 示例DNA硬盘设备

[0250] 图12A-12B各自示出了可用作具有DNA存储系统(例如DNA存储系统1300)的DNA硬盘驱动模块的DNA存储设备或模块700、702(例如存储设备1320)的例子。图12A示出了包括壳体706的DNA存储设备700,一个或更多个安装导向件708定位在壳体706上,并适于帮助将存储设备700与DNA存储系统1300耦合,以相对于一组仪器1301精确定位和固定存储设备700。安装导向件708可以包括轨道、主动锁定键机构、被动锁定键机构和闩锁或锁定机构中的一个或更多个,该轨道帮助滑动受控的方向和距离,该主动锁定键机构在耦合期间由接收器部分(例如,模块接收器1321)的相应部分移动或被移动,以仅当以正确的取向插入时才允许完全耦合,该被动锁定键机构在耦合期间配合或接收DNA存储系统1300的静态部分,以确保正确的取向,该闩锁或锁定机构在相对于DNA存储系统1300实现正确的定向和定位

时固定存储设备700。

[0251] 壳体706可以包括内部710,当存储设备700没有与DNA存储系统1300耦合时,内部710可以被密封或屏蔽以防止周围环境,或者可以一直被密封以防止周围环境。在一些实现中,壳体706可以被密封以防止内部710中的部件容易接近或可见,但是没有被密封以防止空气、湿度或周围环境的其他方面。在一些实现中,壳体706可以包括门或盖,当存储设备700与DNA存储系统1300耦合以允许完全耦合时,该门或盖滑动、旋转或以其他方式移动。在一些实现中,内部710内的设备或部件本身可以相对于内部710的环境和壳体706外部的周围环境单独密封,无论壳体706本身是否以任何方式密封。密封和屏蔽特征可包括,例如,部件的气密密封和耦合、减少环境温度对内部710的影响的绝缘材料、减少运动、振动或其他物理冲击对内部710内的部件的影响的减震材料、防止外部辐射源影响内部710的辐射屏蔽材料以及其他特征。

[0252] 存储设备700包括一个或更多个流通池1322,流通池1322本身包括多个存储阱1348。流通池1322及其相应的存储阱1348可包括一个或更多个关于图3中描述的流通池400、图5中描述的流通池500、图6中描述的流通池600、图8中描述的流通池601或本文描述的其他流通池描述的特征。

[0253] 存储设备700包括多个接口,这些接口可以部分在内部710内,并且部分被壳体706暴露(例如,在外壳的表面上可见,或者当盖子、门或其他设备在耦合期间移位时暴露)。所示的接口包括测序接口1324、合成接口1326、流控接口1328和模块电子接口1330。

[0254] 在一些实现中,流通池1322、测序接口1324、合成接口1326、流控接口1328和一组模块电子设备1330可以是其间具有一个或更多个流控、电或机械接口的独立设备。在其他实现中,流通池1322、测序接口1324、合成接口1326、流控接口1328和一组模块电子设备1330可以集成到单个设备中,其中流通池1322、测序接口1324、合成接口1326、流控接口1328和一组模块电子设备1330中的每一个形成其子部件。

[0255] 当与DNA存储系统1300耦合时,测序设备1304通过测序接口1324与流通池1322相互作用。测序接口1324可以包括玻璃盖、玻璃基底或被配置成允许测序设备1304与流通池1322和存储阱1348相互作用的其他接口表面。在测序设备1304包括可用于检测阱内的带标签或被标记的核苷酸或其他物质的光学传感器和光源的示例中,测序接口1324可包括透明玻璃盖,所述玻璃盖覆盖流通池1322(例如,如在流动通道410的上下文中所描述的),并防止在其中传输的流体泄漏,同时还在每个方向上将光传输至存储阱1348和从存储阱1348传输光。在一些实现中,测序接口1324可以包括一个或更多个波导,以选择性地照亮流通池1322的一个或更多个部分。

[0256] 合成接口1326被配置成允许合成设备1306与流通池1322的一个或更多个阱相互作用,因此将根据特定的合成设备1306而变化。在其他实现中,合成设备1306可以在没有阱的流通池1322的特定表面上合成DNA核苷酸。在一些实现中,合成接口可以包括导电层或耦合,该导电层或耦合从电极接收电特性并将它们传导至流通池1322的阱附近的区域。在一些实现中,合成接口1326可以包括或者可以与流控接口1328的一些或全部相互作用,例如在合成设备1306在合成期间提供核苷酸载液的情况下。在一些实现中,合成接口1326可以包括密封流通池1322的多孔膜,同时当由核苷酸注射头或端口在期望的位置注射时,允许核苷酸进入流通池1322或特定的存储阱1348。在一些实现中,合成接口1326可以由柔性材



料形成,可以包括多个小阀,或者可以包括被配置为在核苷酸注射头或端口提供核苷酸后自密封的其他特征。

[0257] 流控接口1328将基于流控设备1308的特定实现而变化,并且可以包括本文所述的流控网络的一个或更多个特征,例如入口端口420、出口端口422和其他部件。流控接口1328可以是被动的,例如歧管网络,其向一个或更多个存储阱1348提供相等的供应流体流,或者一组入口,其可以向特定组的存储阱1348(例如,流动通道410的单个通道的阱)提供流体。在一些实现中,流控接口1328还可以包括一个或更多个主动特征或与一个或更多个主动特征一起操作,例如流控致动1360,其可以包括阀、分流器、泵或其他特征,以提供对流体施加到流通池1322表面的更多控制。阀可以是电子控制的,或者可以作为来自流控设备1308的输入流体的特定压力的结果而被激活。泵或电润湿表面可以使用由电接口1310提供的电力和控制信号进行电子控制,并且可以用于向特定的一组存储阱1348提供流体,而不是向整个流通池1322提供流体。

[0258] 存储设备700还可以包括一个或更多个由来自外部源的电力或控制信号操作的部件,或者需要与外部源交换数据的部件。这些部件可以与存储设备700的模块电子接口1330耦合,当存储设备700与DNA存储系统1300耦合时,模块电子接口1330本身与DNA存储系统1300的电子接口1310耦合。存储设备700的部件可以通过模块电子接口1330从DNA存储系统1300接收电力,与存储控制器1302交换数据,或者与和存储控制器1302通信耦合的其他设备交换数据。

[0259] 存储设备700还可以包括高速缓存存储器1334。高速缓存存储器可以包括非易失性电子存储器,其可以接收和存储来自其他设备的数据,例如存储控制器1302、与存储控制器1302通信耦合的另一设备或者存储设备700本身的另一设备。高速缓存存储器1334可以接收和存储索引数据,该索引数据描述存储阱1348的状态和内容、存储在存储阱1348中的特定文件或数据的位置、可与DNA存储系统1300上的相应信息一起使用的编码或解密信息,以从存储阱1348读取数据。在涉及存储设备700的读取和写入操作期间,存储控制器1302可以提供和修改这样的信息。高速缓存存储器1334也可用于存储被写入存储阱1348的数据部分,如将在下面更详细描述,以便在一些情况下允许同时读取和写入数据到存储设备700。

[0260] 存储设备700还可以包括一组模块传感器1336。模块传感器可以包括例如不同位置(例如,在内部710的壁上,在流通池1322的表面上)的温度或湿度传感器、振动或冲击传感器、辐射传感器、光传感器和其他环境传感器。该组模块传感器1336可以基于感测到的条件产生信号和数据,这些信号和数据可以存储在高速缓存存储器1334上,当与DNA存储系统1300耦合时提供给存储控制器1302;或者提供给DNA存储系统1300或存储设备700的另一个设备或部件。

[0261] 存储设备700还可以包括电源1338,例如电池、太阳能电池、感应充电接收器、定向能量接收器或可以代替模块电子接口1330或除模块电子接口1330之外产生或接收功率的其他设备。在一些实现中,电源1338可以为存储设备700的一些或所有部件供电,即使当没有耦合到DNA存储系统1300时。作为一个例子,当存储设备700从DNA存储系统700卸载时,例如在长期存储或运输到长期存储期间,电源1338的电池可以通过模块电子接口1330充电,以便为模块传感器1336供电。作为另一个例子,感应充电接收器可以在存储时给存储设备

700供电,其中存储设备700被放置用于存档存储的存储架或货架系统包括与感应充电接收器对准的感应充电表面。

[0262] 存储设备700还可以包括状态指示器1340,其可以包括灯指示器、音频指示器、显示屏或可操作来向用户提供信息的其他设备。状态指示器1340可以由用户交互以查看由模块传感器1336或电源1338提供的信息,或者可以由这样的设备激活以提供与存储设备700相关的警告。例如,当模块传感器1336检测到可能损坏存储在存储设备700中的机器写入的DNA的温度、湿度、运动、辐射或其他特征时,状态指示器1340可以被激活以点亮红色闪烁发光二极管(LED),从扬声器发出警告声音或提供另一个警告指示。

[0263] 存储设备700还可以包括无线通信设备1342,无线通信设备1342可以包括Wi-Fi收发器、蓝牙或NFC收发器、RFID发射器或能够无线传输数据、接收数据或两者的另一设备。在一些实现中,无线通信设备1342可以基于来自模块传感器1336的数据(例如,诸如经由低能量蓝牙收发器向存档存储架附近的基站提供的温度或辐射警报警告,其随后可以生成各种类型的警报)生成与这种传输范围内的其他设备的电子通信。在一些实现中,无线通信设备1342的RFID可以存储存储设备700的唯一标识符,当存储设备700被插入时,该标识符可以被提供给DNA存储设备1300。在一些实现中,这种RFID可以接收和存储来自模块传感器1336的信息,该信息可以由放置在存储设备700附近的手持RFID读取器读取,以便无线地接收存储设备700最近生成的温度数据或其他传感器获取的数据。在一些实现中,无线通信设备1342的Wi-Fi收发器可以连接到本地可用网络,并提供对高速缓存存储器1334的内容的网络访问,以帮助远程确定存储在存储设备的流通池1322上的特定存储设备或特定文件或数据的位置和内容。

[0264] 应当理解,存储设备700可以以各种方式实现,并且不是图12A所示的所有特征都是必需的或者将出现在每个存储设备中。作为一个例子,一些旨在提供非常简单、低成本存储的存储设备可以仅包括几个简单的接口(例如,提供对流通池1322的光学访问的玻璃盖)、流通池和包括一个或更多个安装导向件以帮助将模块定向和插入DNA存储系统1300的壳体。甚至高端存储设备也可能不包括存储设备700的所有特征,因此应当理解,所示的特征仅是示例,并且基于存储设备700的存储设备的变化实现将包括这些特征的变化组合或省略。

[0265] 此外,一些存储设备可以包括图12A中未明确示出的附加特征。作为示例,存储设备的一些实现,例如存储设备700,可以包括流体储器,该流体储器包含用于写入数据、读取数据或两者的一种或更多种试剂,使得存储设备700在运输期间与试剂配对并携带试剂。在这样的实现中,流控设备1308可以包括泵或经由流控网络提供加压液体或气体,以将所需的试剂从这样的内部储器移动到流通池1322,代替或除了提供试剂本身之外。

[0266] 作为示例,图12B示出了存储设备702的替代示例,其可与诸如DNA存储系统1300的DNA存储系统一起使用。存储设备702包括图12A的存储设备700的许多特征,具有在存储设备700的上下文中描述的类似功能。该示例的存储设备702还包括一组流通池工具1350,其可以包括一个或更多个可操作来帮助合成、测序和影响流通池1322的其他操作的特征。用于存储设备702的一组流通池工具1350包括流控致动设备1360、测序致动设备1362和合成致动设备1364。该组流通池工具可由集成电路1352操作,如下文将更详细描述,或者可通过经由模块电子接口1330接收的控制信号操作,或者两者兼有。

[0267] 流控致动设备1360可包括主动流控特征,其可操作以控制各种流体向存储阱1348或流通池1322的其他区域的输送。作为示例,流控致动设备1360可以包括电润湿表面,该电润湿表面被配置成接收电信号并产生相应的电特性,以便将受控量的流体输送到电润湿表面上的特定位置。流控致动设备1360还可包括电致动阀、泵、喷嘴和主动流控网络的其他特征,其可操作以向流通池1322的表面提供受控的流体输送。

[0268] 测序致动设备1362可以包括类似于流控致动设备1360的主动流控特征,并且可以是单独的部件,或者可以在不同的实现中与流控致动设备1360共享。测序致动设备1362还可包括主动光学特征,其可用于在可能在测序期间发生的照明和成像过程期间限制附近存储阱1348之间的光学串扰,或者可用于提供照明或其他特征。作为示例,测序致动设备1362可以包括电致变色材料和部件,其可操作来改变存储阱1348的光传输特性,以允许测序,或者降低对附近阱的串扰的可能性。作为另一个例子,测序致动设备1362可以包括多个照明器(例如,LED),每个照明器对应于并可操作来照明一个或更多个存储阱1348。

[0269] 合成致动设备1364可以包括在合成期间可用的主动特征,例如与流控致动设备1360的那些相似或共享的主动流控特征;并且还可以包括可操作来促进或抑制合成活性的主动特征。作为示例,合成致动设备1364可以包括在一对一、多对一、一对多或多对多的基础上对应于存储阱1348的多个阱电极。如结合图6-8和电极组件640所述,可以单独或共同操作阱电极,以在存储阱1348附近产生电特性,该电特性可以促进或抑制与存储阱1348内机器写入的DNA的合成和测序相关的某些事件(例如,在SBS的情况下,其中测序过程可以包括一个或更多个类似于合成期间执行的子过程)。

[0270] 合成致动设备1364的阱电极可以实施为电极组件640,可以安装在储存阱1348的侧壁上,可以安装在储存阱1348周边的流通池1322的表面上,或者可以安装在其他地方。阱电极可以安装在流通池1322的顶侧(例如,具有阱开口的表面)或存储阱1348内,向每个电极供电的电路从该电极延伸到流通池1322的与顶侧相对的下侧。

[0271] 合成致动设备1364的阱电极可以与存储阱1348一一对应或多对一对应,使得一个或更多个阱电极可以被激活以在单个存储阱1348处产生期望的电特性。在具有不同对应关系的实现中,可以激活单个阱电极以在几个存储阱1348处产生期望的电特性(例如,使得单个阱电极的激活可以实现第一阱中的合成(框1224)和第二阱中的合成(框1228),或者第一阱和第二阱的测序,其中这些阱彼此接近并且在空间上与单个阱电极相关联)。

[0272] 存储设备702还包括集成电路1352,其可以包括例如互补金属氧化物半导体(CMOS)。集成电路1352可以被配置成经由模块电子接口1330从存储控制器1302或从另一设备接收控制信号,并且选择性地激活一个或更多个流通池工具1350。以这种方式,集成电路1352可以操作流控致动设备1360(例如,通过激活电润湿表面来传输流体)、测序致动设备1362(例如,通过激活电致变色特征来减少串扰)或合成致动设备1364。

[0273] 作为示例,集成电路1352可以操作合成致动设备1364的阱电极,以在一个或更多个存储阱1348处产生期望的电特性。电特性可以包括电流或电压的存在或不存在,以及改变电流或电压的幅度、频率、顺序或其他特性。集成电路1352可与流通池1322的下侧静态耦合并组装在壳体706内,并且可另外包括照明器、光电二极管和/或类似地可由存储控制器1302或另一设备操作的其他设备,以帮助存储阱1348内的测序和合成,如关于电极组件640所述。

[0274] 当耦合到流通池1322时,集成电路1352的多个电连接将与合成致动设备1364的每个电极的电路对准并导电耦合,或者类似地与一组流通池工具1350的其他特征耦合。在一些实现中,集成电路1352和阱电极之间的电连接可以包括透明导电材料,例如铟锡氧化物材料,以有助于耐用性和导电性,同时不妨碍来自流通池1322下侧的存储阱1348的照明和光学成像。在一些实现中,透明导电材料以及导电材料穿过的阱的结构可以是各向异性的,以便进一步控制或限制光传输到阱中。在一些实现中,阱电极和集成电路1352之间的导电耦合可以包括定位成彼此接触并传导电信号的多个导电焊盘。

[0275] 图12C示出了可与诸如DNA存储系统1300的DNA存储系统一起使用的存储设备704的替代示例。存储设备704包括图12B的存储设备702的几个特征。然而,存储设备704不包括集成电路1352,而是包括集成电路槽1354。集成电路槽1354位于流通池1322的下方并靠近流通池1322;并且被成形并适于接收DNA存储系统1300的外部集成电路1356,并且当存储设备704与DNA存储系统1300耦合时,将该集成电路放置成与流通池1322接触。这可以包括外部集成电路1356被静态地定位在DNA存储系统1300的模块接收器1321内或附近,使得当存储设备704被插入时,它被水平地插入到集成电路槽1354中;或者可以包括可机械定位在接收器中的外部集成电路1356,使得它首先水平插入集成电路槽1354中以实现与流通池1322对准,然后垂直定位成与流通池1322接触。这也可包括外部集成电路1356,其可由DNA存储系统1300电子定位,并可通过电子操作该组仪器1301的相应部件从集成电路槽1354自由插入和移除。

[0276] 集成电路槽1354和相应的外部集成电路1356可以被成形并适于提供外部集成电路1356到流通池1322的可重复的、精确的、临时的耦合。这种精确耦合允许外部集成电路1356的多个电触点与该组流通池工具1350的电触点对准并导电耦合,以允许这些特征的控制和操作。作为示例,在测序致动设备1364包括多个阱电极的情况下,集成电路1356可以与合成致动设备1364的每个阱电极对准并耦合。这种对准还可以对准并允许外部集成电路1356的任何照明或光学感测特征与相应的存储阱1348对准。存储设备704对于某些应用可能是有利的,因为它允许外部集成电路1356或CMOS(其可能与相对较高的复杂性和成本相关联)被实现为一组仪器1301的一部分并与多个存储设备一起使用,而不是要求每个存储设备包括其自己的集成电路。

[0277] 上述存储设备的一些实现可以包括比上述更少的部件和特征或者附加的部件和特征。作为示例,图15A和15B示出了在表面906和膜904之间复制DNA的过程的图示。参考图17A的图示900,膜904可以是薄膜介质,其可以从存储设备(例如存储设备700)插入和移除。薄膜904可以包含在壳体或盒内,以帮助插入存储设备700,并在运输过程中保护薄膜904。薄膜904可包含对应于存储阱1348的多个阱或指定区域。表面906可以是流通池1322的表面,存储阱1348可从该表面进入。

[0278] 在所示的例子中,表面906包括多个引物908(例如寡核苷酸),它们可以一对一地位于每个存储阱1348内。薄膜904包括多个机器写入的多核苷酸910,其位于薄膜904的阱内或薄膜的对应于储存阱1348的指定区域。图示900示出了薄膜904位于表面906的近似上方,使得引物908可以与机器写入的多核苷酸910相互作用。在这种相互作用期间,引物908由机器写入的多核苷酸910触发,导致薄膜904的阱或区域中的每个机器写入的多核苷酸910复制到表面906的相应阱。

[0279] 参考图17B的图示902,当薄膜904随后被移除时,机器写入的DNA的复制品保持存储在表面906的阱内。可以看出,通过插入薄膜904,机器写入的DNA可以复制到表面906的一个或更多个阱中。这对于使用DNA存储系统1300或类似设备在第一位置制备机器写入的DNA并将其写入薄膜904,然后将机器写入的DNA从薄膜904复制到第二位置的存储设备700的存储阱1348可能是有用的。薄膜904和表面906之间的引物的输送可以通过一组仪器1301、一组流通池工具1350或两者来执行。作为示例,合成致动设备1364的电极可被操作以将复制的机器写入的DNA吸引或排斥到目的地表面,或者来自流控致动设备1360的受控流体流可将复制的机器写入的DNA传输到目的地表面。

[0280] 上述过程也可用于将机器写入的DNA从存储设备700复制到薄膜904,然后薄膜904可被运输和测序或在别处读取。例如,薄膜904可以用对应于表面906的阱的多个引物908制备,而表面906的阱包含多个机器写入的多核苷酸910。薄膜904可以被插入到存储设备700中,并且定位在表面906附近,使得机器写入的多核苷酸910从表面906复制到薄膜904上。然后,薄膜904可以被传送到DNA存储系统1300或类似的设备并由其读取。上述每一项也可以通过在适当位置(例如,在表面906的阱内,从阱内的链或在薄膜904上,从薄膜904上存在的链)产生DNA的复制链,然后将复制链电泳地转移和结合到它们相应的目的地来进行。这种转移也可以通过物理接触将复制表面印迹到目标表面上,或者通过使用流控设备1308或另一设备来进行。

#### [0281] C. 管理DNA硬盘驱动模块的方法

[0282] 如关于高速缓存存储器1334所描述的,使用高速缓存策略来管理与存储设备(例如存储设备700等)的数据写入和读取可能是有利的。作为一个例子,在数据被写入存储设备700并且在写入流通池1322完成之前用户开始从DNA存储系统1300的一组仪器1301卸载卷的情况下,未写入的数据可以被存储在电子存储器上,并且将与存储设备700一起移动,直到它可以被写入流通池1322的这种时间。作为另一个例子,在为存储设备700维护阱内容、文件位置或其他类似数据的索引的情况下,在与存储设备700一起移动的电子存储器上维护该索引,而不是将该索引存储在云存储中或者存储在DNA存储系统1300的一组仪器1301的永久存储卷上,或者除此之外,可能是有用的。以这种方式,如果存储设备700与不同的DNA存储系统1300的一组仪器1301一起运输和安装,阱索引立即可用,而不需要从云存储卷或从先前耦合的DNA存储系统1300的一组仪器1301获得信息。

[0283] 作为高速缓存策略的另一示例,图13描绘了当诸如高速缓存存储器1334的电子存储器可用时,可以被执行以向存储设备700提供读取和写入操作的高速缓存的过程1120的流程图。当用DNA存储系统1300将数据写入存储设备700时,可以将附加数据写入高速缓存存储器1334。这可以包括将文件索引写入(框1122)到高速缓存存储器1334,其可以描述多个阱的内容和特定文件或数据的位置,无论是存储在多个阱中、存储在高速缓存存储器1334上还是两者都有。这种信息可以用于允许以后访问和检索所请求的数据。单个文件或数据束的校验和数据也可以被写入(框1124)高速缓存存储器1334,并且可以与文件索引相关联。

[0284] 在高速缓存存储器1334上存储文件和数据的索引、校验和列表或两者,与在流通池1322上存储这样的数据并要求对其进行测序或者在云存储或服务器上存储这样的数据并要求在驱动器内容可访问之前对其进行访问和检索相比,可以在将来实现更快的数据读

取和写入。除了将其存储在其他位置(例如,在流通池1322本身上或在网络可访问卷上)之外,将这种数据存储在高速缓存存储器1334上还提供了这种数据的冗余存储的额外优点,因为文件表和索引的丢失可能导致存储在流通池1322上的数据的完全丢失,或者基于逐阱检查重建文件索引的时间和资源成本大大增加。

[0285] 当存储设备700与DNA存储系统1300耦合时,系统可以从用户或者从与DNA存储系统1300通信的系统和设备接收(框1126)读取操作,可以接收(框1136)写入操作,或者两者都接收。如已经讨论的,在一些情况下,可能不可能允许同时读取和写入数据到流通池1322的分离阱。这可能是由于DNA存储系统1300的限制或存储设备700的限制。作为示例,存储设备700的一些实现可以被认为具有两种互斥模式:读取模式和写入模式。由于不同的流控、阱调节和每次操作所需的其他设备,同时读取和写入可能不可靠或不可能。此外,从读取模式切换到写入模式可能需要花费时间和浪费试剂。虽然这样的实现可能特别受益于所公开的高速缓存策略,但是应该理解,即使支持同时读取和写入的实现也可能遇到它们可能受益于高速缓存策略的各种场景。

[0286] 在接收到读取操作的情况下(框1126),DNA存储系统1300可以检查文件索引,以定位所请求的数据并确定其当前是否存储在高速缓存存储器1334上(框1128)。在不同的情况下,所请求的数据可以在高速缓存存储器1334上可用。作为示例,在数据被写入存储设备700并且随后不久被请求的情况下,它仍然可以被存储在高速缓存存储器1334中。作为另一个例子,在数据最近基于请求从流通池1322读取的情况下,它可以存储在高速缓存存储器1334中,直到被重写。作为另一个例子,DNA存储系统1300可以被配置为在可能的情况下将存储在流通池1322上的某些数据标记为也保存在高速缓存存储器1334中。这种配置可以由用户提供的手动配置,或者可以基于DNA存储系统1300的自动确定,该自动确定本身基于对这种数据的读取请求的频率。

[0287] 在所请求的数据可从高速缓存存储器1334获得的情况下,DNA存储系统1300将从高速缓存存储器1334读取数据(框1130)以服务该请求,这可以允许存储设备700保持在写模式,同时允许数据被读取(例如,从高速缓存存储器1334而不是从流通池1322)。在所请求的数据没有存储在高速缓存存储器1334中的情况下,当这样的功能可用时(例如,当存储设备700处于读取模式或者当读取操作以其他方式可用时),DNA存储系统可以从流通池1322读取数据(框1132)。在存储设备700处于写入模式并主动向流通池1322写入数据的情况下,由于在模式之间切换的时间和试剂成本,优先切换回读取模式可能不是有利的。然而,在排队待写入的数据具有可存储在高速缓存存储器1334上的大小的情况下,切换到读取模式并允许从阱中读取所请求的数据(框1132)同时输入数据存储在高速缓存存储器1334上可能是有利的。在这种情况下,已经请求写入和读取数据的用户或其他系统或设备察觉到这些动作正在同时执行,因为输出数据正在从流通池1322读取,而输入数据正在被写入高速缓存存储器1334。

[0288] 在每次读取操作之后,可以更新高速缓存存储器1334上的文件索引或另一个数据集(框1134)以反映最近请求的数据的读取频率,这样的数据集对于将来确定由于使用频率(例如,可以每单个数据请求的数据)或使用模式(例如,每个星期五请求的数据可以在星期四以低优先级被循环到高速缓存存储器1334中,使得它在一段时间内完成,在这段时间内可能有减少的数据读写请求)而应被循环到高速缓存存储器1334中的数据有用。

[0289] 继续参考图13,其中接收到写入操作(框1136),DNA存储系统1300可以确定(框1138)存储设备700当前是否处于读取模式。在存储设备700当前处于读取模式的情况下,可以将与写入操作相关联的输入数据写入(框1142)到高速缓存存储器1334,并且标记为在可用时写入到流通池1322。在存储设备已经处于写入模式的情况下,输入数据可以被写入(框1140)到流通池1322的一个或更多个阱。

[0290] 所公开的高速缓存方法还可能受到高速缓存策略的影响,该高速缓存策略优先于其他考虑因素而停留在当前模式(例如,读取模式或写模式),使得所有排队的读操作可以在切换到写模式之前被执行,而不管请求的到达顺序,并且甚至可以在最后一个读取请求完成之后的短暂时间段内保持读取模式,以便允许其他读取请求在切换模式之前到达并被服务。除了减少在给定时间段内执行的模式切换的总数之外,这种策略还降低了特定于读取的试剂与特定于写入的试剂交叉污染的风险,否则,由于模式切换更频繁,这种交叉污染会更频繁地发生。

[0291] 当从流通池1322写入和读取数据时,存储设备700还可以受益于数据克隆和数据分段策略。作为示例,图14描绘了过程1220的流程图,过程1220可被执行以提供与诸如存储设备700的存储设备的冗余数据写入和读取操作。当接收到提供应当写入流通池的阱的输入数据的输入请求(框1222)时,DNA存储系统1300可以根据其能力以并行或紧密顺序合成(框1224)和将输入数据存储在第一阱中,并且合成(框1228)和将输入数据存储在第二阱中。虽然图14描述了第一阱和第二阱,但是应当理解,数据可以有利地跨三个或更多阱执行。第一阱和第二阱的冗余可用于错误检查,例如通过对第一阱和/或第二阱的多核苷酸中的一个或两个进行测序,以确定在读取和/或写入过程中是否发生定相或预定相。假设在合成中没有错误,机器写入的多核苷酸的相同拷贝对应于第一个阱和第二个阱中的输入数据。根据第一阱和第二阱的性质,这提供了各种优点。作为示例,参考存储设备700,第一阱可以位于流通池1322中的第一个中,而第二阱可以位于流通池1322中的第二个中。在这种情况下,写入的数据被冗余地存储在两个独立的流通池中,这对于数据完整性和最小化数据丢失的风险是理想的。

[0292] 即使使用了另一个存储设备,例如存储设备1320,并且第一阱和第二阱各自位于流通池1322中,跨同一流通池中各个阱克隆数据也有一些优点。除了提供数据冗余以降低由于阱故障或一个阱内的DNA意外降解而导致数据丢失的风险之外,克隆写入的数据还可以提供额外的错误检查能力,无论两个阱位于何处。作为一个例子,在输入数据被写入如图14所示的分开阱的情况下,然后通过测序(框1226)和从第一阱读取数据,以及测序(框1230)和从第二阱读取数据,从两个分开的阱读回输入数据。从每个阱中读取的输出的比较(框1232)或输出的校验和将指示写入一个池的多核苷酸是错误合成的还是随后由于某种原因降解了。

[0293] 在一些实现中,第二阱中多核苷酸的合成可以基于第一阱中多核苷酸的合成来完成。也就是说,可以对写入第一阱中的多核苷酸进行克隆扩增,并且可以将一个或更多个克隆的多核苷酸储存在第二阱中和/或流体储存室中。可以对第一阱中克隆扩增的多核苷酸进行测序,以确定第一阱中核苷酸的序列。可以将第一阱中的核苷酸序列与指示的写入多核苷酸进行比较,以确定在写入过程中是否发生了任何定相或预定相错误。如果发生错误,存储在第二阱和/或流体存储室中的一个或更多个克隆的多核苷酸可能由于损坏而被丢

弃,并且写入过程可能再次发生。如果没有发生错误,则储存在第二阱和/或流体储存室中的一种或更多种克隆的多核苷酸可以被克隆并储存在第一阱和/或一个或多个其它阱中,以提供本文所述的两种或更多种相同的多核苷酸。

[0294] 图15描绘了过程1240的流程图,过程1240可被执行以提供与存储设备的高速数据写入和读取。图15的过程可以使用存储设备700来执行。当接收到输入请求时(框1242),DNA存储系统1300可以将与输入请求相关联的数据分段(框1244)成多个部分(例如,两个或更多)。系统可以将第一部分合成并写入(框1246)到第一阱中,并且并行地将第二部分合成并写入(框1248)到第二阱中,而不是将整个输入写入到单个阱中。在第一阱和第二阱可以以这种方式分别合成而不影响另一个阱的写入速度的情况下,结果是相对于将整个输入写入单个阱,输入数据可以以大约两倍的速度完全写入DNA储存器。类似地,当接收到输出请求时(框1250),DNA存储系统1300可以测序(框1252)和从第二阱读取第二部分,并且并行地,测序(框1254)和从第一阱读取第一部分。然后,这两个部分可以被重组,并且可以提供完整的输出(框1256)。同样在这种情况下,相对于从单阱中读取数据,可以以更高的速度从DNA储存器中读取输出数据。

[0295] D. 安装和卸载DNA硬盘驱动器的方法

[0296] 如已经讨论的,所公开的DNA硬盘驱动器的一些实现可以暂时与DNA存储系统1300耦合,并且可以从其移除;并且可以用于在DNA存储系统1300和其他系统、设备或存储位置之间的便携性。在诸如图12A的存储设备700的DNA硬盘驱动器由于任何原因可以与DNA存储系统1300耦合和分离的这种情况和其他情况下,DNA存储系统1300、存储设备700或两者可以被配置为在与DNA存储系统1300安装(例如,耦合)期间和从DNA存储系统1300卸载(例如,分离)期间执行某些过程。

[0297] 作为示例,图14A和14B示出了当安装并随后卸载存储设备(例如存储设备700)时,可以用本文公开的系统和存储设备(例如DNA存储系统1300)执行的过程。图16A示出了可被执行以安装存储设备并使其可用于读取和写入数据的过程800。存储设备700可以被物理地安装(框802)到DNA存储系统,例如通过将存储设备700插入模块接收器1321槽或隔间直到完全就位,这可以由来自一个或多个安装引导件708的物理反馈或由DNA存储系统1300的用户界面产生的听觉响应、视觉响应或其他指示来指示。DNA存储系统1300然后可以识别(框804)并确定存储设备700的状态信息,其可以包括例如识别类型或型号、基于标识符唯一地识别特定存储设备700、或者识别与存储设备700相关联的阱索引或文件索引信息。

[0298] 存储设备700的识别(框804)可以以不同的方式执行,包括通过DNA存储系统1300的光学扫描仪对准并识别壳体706上的光学代码,无线收发器从RFID存储器或其他无线通信设备1342接收信息,或者存储控制器1302从高速缓存存储器1334接收信息。可以通过经由有线或无线连接从高速缓存存储器1342接收这样的信息来执行阱状态、文件状态和其他信息的确定(框804)。还可以使用存储设备700的唯一标识来执行确定(框804),以查询存储这种信息的另一设备或部件,例如存储控制器1302可以访问本地可用存储驱动器或远程可用存储驱动器来确定这种情况。

[0299] 然后,DNA存储系统1300可以逻辑地安装(框806)存储设备700,这可以包括重新配置其自己的软件,例如操作系统、文件系统或软件应用,以指示存储卷700准备好用于作为DNA存储系统1300的存储卷来读取和写入数据。



[0300] 在接收到具有一组输入数据的写入请求的情况下(框808),DNA存储系统1300可以调节和准备(框810)存储设备700,用于将数据写入一个或多个存储阱1348,然后在对应于接收到的输入数据的那些阱内合成(框812)机器写入的DNA。如已经描述的,这可以包括激活一组仪器1301中的一个或多个,以向存储设备700提供试剂流体、热输入、光输入和其他输入,以便在阱内为机器写入的多核苷酸创建期望的核苷酸序列,其中期望的核苷酸序列通过基于编码方案将该组输入数据编码成DNA格式来确定。期望的核苷酸序列还可以包括附加到该组输入数据的其他信息,其可以包括散列或校验和信息、指示特定文件或数据集的结束或开始的标记或其他信息。所需的核苷酸序列可以在阱中作为机器写入的单链DNA和机器写入的双链DNA(每条单链相互镜像),或者作为两者合成。

[0301] 写入操作完成后,DNA存储系统1300可以提供(框814)操作确认,并更新一个或多个状态表以反映操作的完成。确认信息可以包括例如操作的时间和日期、输入数据的校验和或哈希值、受操作影响的存储阱1348、存储设备700的标识以及其他信息,其中的一些或全部可以被添加到阱状态表、文件索引表或其他表中。

[0302] 在接收到指示期望输出数据的读取请求(框816)的情况下,DNA存储系统1300可以调节和准备(框818)存储设备用于读取数据,如已经描述的,然后对来自一个或多个存储阱1348的所请求的输出数据进行测序(框820)。可以接收读取请求(框816)作为文件或数据集的描述,其在存储阱1348内的位置可以通过使用阱状态表、文件索引表或其他类似目录来确定。针对读取操作和数据测序(框820)对存储设备700进行调节(框818)可以包括向受影响的存储阱1348提供试剂输入、热输入、光输入和其他输入,并捕获输出。作为一个例子,输出可以包括来自与存储阱1348内的机器写入的DNA成对的光学标记的核苷酸的发射光,其可以从指示机器写入的多核苷酸的核苷酸的序列转换成匹配所请求的输出数据的数字数据。在完成读取操作之后,DNA存储系统1300可以提供(框814)对操作的确认,并更新一个或多个状态表。读取操作的确认可以包括诸如存储阱1348读取的标识、合成前校验和、合成后校验和、作为编码核苷酸序列或解码成数字形式的输出数据以及其他信息的信息。

[0303] 在安装时,存储设备700可以执行多个读取和写入操作,直到接收到卸载请求(框832)。卸载请求可以是与DNA存储系统1300的手动交互的结果,例如通过按下按钮来指示存储设备700需要被移除,或者是由于存储设备700处于或接近其存储容量而自动移除。当接收到时,DNA存储系统1300可以确定(框834)存储设备700的状态,例如它当前是否正在被写入或读取。在一些实现中,在当前操作正在被执行的情况下,DNA存储系统1300可以延迟进一步的处理,直到读取或写入操作完成。

[0304] 在一些实现中,DNA存储系统1300可以响应于任何卸载请求或者响应于高优先级或立即卸载请求,在读取或写入操作完成之前终止读取或写入操作。在这种情况下,DNA存储系统1300可以通过配置其操作系统、文件系统或软件应用来反映存储设备700不再可用于对与存储设备700交互的设备或软件应用的读取或写入操作,来立即逻辑地卸载(框836)该模块。然后,DNA存储系统1300可以确定存储设备700当前是否处于写入过程中(框838),如果是,则可以将输入数据的未写入部分写入诸如高速缓存存储器1334的位置。以这种方式,在文件或数据集仅被部分写入存储阱1348的情况下,未写入的部分可以被存储(框840)到高速缓存存储器1334,因为将文件的两个部分物理地分组在同一存储设备700上可能是有利的。

[0305] 在存储(框840)输入的未写入部分之后,或者当存储设备700未被使用或者涉及数据读取过程时,DNA存储系统1300可以采取任何必要的动作来关闭(框842)当前或最近涉及读取或写入操作的任何阱。在最近的合成操作后,这可能包括提供安全储存当前形式的DNA或保护DNA免于降解所需的任何封闭核苷酸序列或试剂。然后,DNA存储系统1300可以使用流控设备1308从存储设备700中排空(框844)任何读或写试剂或其他流体,并且还可以向存储设备700提供(框846)存储流体。存储流体可以包括防腐剂或其他被配置为保护和保存存储阱1348内的机器写入的DNA的流体,并且可以例如填充各个阱或者为每个阱提供薄的保护涂层。防腐流体可以包括例如抗氧化剂、干燥剂、聚合物涂层或其他物质或物质的组合,例如携带抗氧化剂的液体聚合物涂层。在存储设备700的未来安装期间,保护涂层可以在存储设备700的后续写入或读取之前被冲洗和移除。

[0306] DNA存储系统1300还可以执行其他过程来调节(框848)用于存储的模块,这可以包括降低存储设备700的温度(例如,降低绝缘壳体706内的流通池1322的温度)和降低内部710的湿度,以允许存储设备700稳定地运输到温度和湿度受控的存储区域。作为调节(框848)过程的一个例子,DNA存储系统1300可以通过提供热输入和压力输入来冷冻干燥流通池1322,以升华流通池1322内的液体,然后以气体形式排空液体。根据安全冷冻干燥流通池1322所需的时间量,冷冻干燥的调节(框848)可以由DNA存储系统1300的子系统或独立设备来执行,使得存储设备700可以从DNA存储系统1300的主模块接收器1321移除,并放置在备用接收器中,该备用接收器不能读取或写入机器写入的数据到存储阱1348,但是能够与存储设备700接口以提供调节(框848)。这样的设备可以靠近DNA存储系统1300,或者可以位于存储设备700将被存档的环境控制的存储区域中。

[0307] DNA存储系统1300还可以校准(框850)一个或更多个模块传感器1336(如果存在)。这可以包括,在调节(框848)存储设备700之后,基于调节时的温度、湿度、光、运动、辐射或其他可检测特性的已知值,配置和校准模块传感器1336以提供对这些特性的精确感测(框848)(例如,流通池1322的温度和湿度将基于这些特性的配置调节值而已知,而当存储设备700与DNA存储系统1300或子系统耦合时,运动、光和辐射可以被假设为接近零)。传感器的校准(框850)还可以包括校准、配置或启用与存储设备700将被存储的房间相关联的一个或更多个传感器,或者与存储设备700将被存储在其中的特定机架、架子或存储壳体相关联的一个或更多个传感器。作为另一个例子,一个或更多个存储设备可以被光学菊花链连接,并且光时域反射计可以被用来共同监控存储设备的各方面。

[0308] DNA存储系统1300还可以更新(框852)存储在高速缓存存储器1334或另一存储器上的一个或更多个状态表,以反映卸载的时间和日期、卸载时执行的过程(例如,指示存储设备700是否被提供防腐流体或涂层(框846)以及类型、存储设备700被调节的状态(框848)、存储设备700是否包含存储在高速缓存存储器(框840)上的任何未写入的输入数据,和其他在稍后安装存储设备700时可能有用的信息。

[0309] DNA存储系统1300还可以禁用或移除(框854)一个或更多个模块锁,该模块锁将存储设备700物理固定在与耦合的DNA存储系统1300的模块接收器1321或槽内。这些可以包括电子操作的或释放的闩锁,该闩锁抓住安装引导件708或壳体706的其他部分,并且在存储设备700被物理安装后将存储设备700固定在原位(框802)。一旦存储设备700可以被安全地移除而没有数据丢失或装备损坏,DNA存储系统1300可以通过经由DNA存储系统1300的用

户界面提供听觉、视觉或其他输出来指示(框856)存储设备700的卸载完成。

[0310] 如同其他示例,应该理解,图14A和14B所示的过程是示例,并且各种安装和卸载过程可以包括比所示更少的过程或未示出的附加过程,根据本公开,这种变化对于本领域技术人员来说是明显的。

[0311] VIII. 专用的读取和写入位置的分离

[0312] 在一些实现中,以核苷酸序列形式编码的数据读取的进一步变化也是可能的。例如,在一些实现中,当要读取数据时,不是对用于编码数据的链进行测序,而是可以首先拷贝那些链(例如,通过PCR),并且可以通过对拷贝进行测序来读取数据,从而降低测序可能将错误引入到原件中的风险。图18提供了流通池1400内的通道的一部分的不按比例图示,其是流通池400的变型的示例。换句话说,图18中描绘的通道是流通池400的流动通道410的变型。在这个例子中,用于读取如图5所示的核苷酸序列的阱530包括在通道的顶面1412中,而用于写入如图6所示的核苷酸序列的阱630包括在基面612中。为了支持这一点,在一些实现中,流通池1400的每个表面可以具有专用控制电路,例如用于基面612上的写入阱的第一专用CMOS电路,以及用于顶面1412上的读取阱的第二专用CMOS电路。

[0313] 在该示例中,除了先前在图5和图6的上下文中描述的用于写入和读取机器写入的多核苷酸的部件之外,流通池1400还包括多个珠1410,其可用于将机器写入的多核苷酸650的拷贝从基面612中的阱630转移到顶面1412中的阱530。在一些实现中,当要使用包括具有图18的配置的流通池1400的DNA存储设备来执行写入命令时,写入表面(例如,基面612)上的一个或更多个阱630可以被识别为未使用,并且编码要写入的数据的多核苷酸可以以先前在图6的上下文中描述的方式在那些阱630中合成。另外,在一些实现中,当在如图18所示配置的流通池1400上合成机器写入的多核苷酸时,可以合成那些链,以不仅包括作为写入请求的主题的数据,还包括可以匹配先前已经固定到珠1410表面的序列的结合序列。

[0314] 在一些实现中,当接收到对已经存储在如图18所示配置的流通池1400中的数据的读取请求时,一个或更多个珠1410可以被移动到已经写入链650的阱630。在一些实现中,这可以通过使用阱630中的电极640来产生可以吸引珠1410的磁场(例如,在珠是顺磁性的实现中)或吸引珠1410的电场(例如,在珠具有电荷并且可以被电泳装置吸引的实现中)来实现。在其他实现中,例如在阱630中的电极640不够强以产生必需的场的情况下,可以使用单独的电极,其可以将珠吸引到或靠近已经写入链650的阱630。一旦珠1410已经移动,链650可以被拷贝,并且拷贝上的结合序列可以粘附到先前已经固定到珠1410表面的匹配序列上。然后,用于将珠1410移动到链被转移到的位置的电极可以被去激活,并且珠1410可以被转移到拷贝的链可以被读取的位置。例如,在一些实现中,可以激活相对表面(例如,顶表面1412)上的对应阱530中的电极1440,以将珠1410转移到该对应阱530。或者,在一些实现中,不对应于单个阱的磁体或电极(例如,移动到流通池的顶表面1412的物理附近的磁体)可以将珠1410移动到可以读取拷贝的链的位置,而不是必须将它们移动到对应于链最初被写入的阱的特定读取阱。然后,可以如先前在图5的上下文中所描述的那样对拷贝进行测序,从而允许读取数据,而不会增加可能与对原始链进行测序相关联的数据损坏的风险。

[0315] 在一些实现中,珠和阱之间的其他关系也是可能的。例如,在一些实现中,机器写入的多核苷酸650可以直接写在珠本身上,而不是使用珠1410来运输最初被写入池1400的基面612中的阱630中的序列。例如,在一些实现中,当数据将被编码到机器写入的多核苷酸

650中时,具有直接固定到其表面的引物的磁珠1410可以被转移到适当的阱中,并且机器写入的多核苷酸可以由这些引物而不是(或除此之外)阱表面上的引物构建。在这种类型的一些实现中,当合成机器写入的多核苷酸时,它们可以被合成以不仅包括被写入DNA存储设备的数据,还包括识别合成该多核苷酸的阱的预定序列,从而即使该序列可能从未被物理固定到可寻址的阱,也能够保持数据地址信息。

[0316] 在一些实现中,如上所述的寻址数据也可以(或者可替换地)用于支持通过珠多路复用来自多个写入阱的数据,所述珠可以将机器写入的链从那些阱转移到流通池的另一表面中的读取阱。例如,在一些实现中,可以提供具有匹配 $n$ 个不同序列的寡核苷酸固定到其表面的磁珠1410(例如, $x$ 个寡核苷酸匹配序列1, $x$ 个寡核苷酸匹配序列2, $\dots$ , $x$ 个寡核苷酸匹配序列 $n$ )。然后,这些序列可被视为通道标识符,珠在流通池的写表面和读表面之间起到多路复用 $n$ 通道连接的作用。在一些实现中,这可以使用设置有对应于流通池的写表面中的阱的独特序列的条形码的珠来实施。图19示出了简化的(例如,通过描绘长度为4的条形码序列,而不是在实践中优选使用的更长的序列)。在如图19所示配置的实现中,当数据要被编码时,它可以被写入机器写入的多核苷酸中,该多核苷酸不仅可以包括数据,还可以包括珠1410表面上对应于它被写入的阱的序列之一。例如,当数据将被写入第一阱631时,它可以被写入机器写入的多核苷酸中,该多核苷酸不仅可以包括数据,还可以包括对应于第一阱的序列,即AATC。类似地,当数据将被写入第二阱632、第三阱633或第四阱634时,它可以被写入机器写入的多核苷酸中,该多核苷酸不仅包括数据,还包括对应于它被写入的阱的序列,即GCCA(对于第二阱632)、TTAG(对于第三阱633)或CGGT(对于第四阱634)。随后,当要读取数据时,可以拷贝来自第一至第四阱631、632、633、634的链,并且该拷贝结合到珠表面上相应的第一至第四组结合位点1411、1412、1413、1414中的寡核苷酸。然后,具有结合的机器写入的多核苷酸的珠1410可以被运送到流通池的读取表面,并且它们结合的寡核苷酸可以用于确定它们最初被写入的阱631、632、633、634(即,以序列AATC开始的链可以被识别为源自第一阱631,以序列GCCA开始的链可以被识别为源自第二阱632,等等)。

[0317] 应该理解,在图19的上下文中对以上提供的复用的描述的变化也是可能的,并且可以包括在一些实现中。例如,在一些实现中,不是每个阱631、632、633、634具有唯一匹配珠1410表面上的一组结合位点的识别序列,而是在用于识别阱631、632、633、634的序列和珠1410表面上的结合位点之间可能存在多对一的关系。例如,如果每个阱631、632、633、634具有 $m$ 个碱基地址序列,则该序列中的前 $p$ 个碱基可以指示珠表面上的哪组结合位点/哪条通道可以用于运输该阱的链,剩余的碱基用于将通道内的阱彼此区分。此外,在一些实现中,珠表面上的条形码设置可潜在地用于在转移过程中替代流通池上的空间条形码设置(例如,当机器写入的多核苷酸被拷贝到珠上时,它们可与用于将它们结合到珠表面的序列一起被拷贝,而不是与用于识别单个阱的完整序列一起被拷贝)。

[0318] 作为潜在变化的另一个例子,在一些情况下,当需要将机器写入的多核苷酸的拷贝从流通池中的写入位置运输到读取位置时,链的末端可以被脱保护,并且可以向其中添加可以结合珠的特定部分的序列,而不是将特定的阱预先连接到珠表面的特定部分。例如,对于如图19所示的珠1410的珠,在这种类型的实现中,如果要运输来自四个阱的链,那么可以结合到四个结合位点1411、1412、1413、1414的序列可以被添加到它们,使得这些链可以在单个珠1410的表面上同时运输,而不管它们最初被写入的阱。

[0319] 还应该理解,在存在珠的实现中,它们的使用的变化可以扩展到复用的不同方法之外。例如,在一些实现中,除了使用珠来传输和/或存储机器写入的多核苷酸中编码的数据之外,或者作为其替代,珠可以用于存储管理信息,例如指示各种逻辑数据分组(例如,文件)存储在哪里的索引。在一些实现中,这种类型的管理信息存储在珠上,它可以存储在机器写入的多核苷酸中,该多核苷酸以特定的独特序列开始,该序列将这些链与编码文件或类似数据的其他链区分开来。这些独特的序列也可用于捕获对这种类型的管理信息进行编码的机器写入的链,以便当要检索数据时,可捕获这些链并对其进行测序,以确定要检索的数据位于何处。在一些实现中,管理信息也可以或替代地以其他方式存储在珠上,例如颜色编码或一些其他形式的光学编码(例如全息条形码)。此外,在管理信息存储在珠上的一些实现中,它也可以存储在非核苷酸存储器中,在珠上的管理信息丢失或损坏的情况下,它可以提供冗余。类似地,在一些实现中,珠可以仅暂时包含在流通池中,并且可以在数据被写入它们之后从流通池中移出,例如将它们存储在不同的存储模块(例如,如前所述的管)中,或者处置它们(例如,在已经写入或绑定到它们的数据被读取并且不再需要珠之后)。

[0320] 其他变化也是可能的。例如,在一些实现中,使用通过复制编码信息的机器写入的多核苷酸来读取信息的技术,然后从拷贝而不是原始拷贝读取信息,可以使用除珠之外的技术,例如电泳、双向电泳、流体层流流动或其他传输技术,从原始链被写入的位置移动拷贝。应该注意的是,这些技术也可以用于在珠存在的位置运输珠。作为另一个例子,在流通池中有专用读取和写入位置的一些实现中,每个写入位置可以有一个专用读取位置(例如,每个写入阱有一个读取阱,反之亦然)。在其他实现中,可能存在不同的关系,例如可能存在多个写入位置(例如,不同的阱)和仅一个读取位置(例如,没有任何阱的流通池的表面)。作为另一个例子,在一些实现中,读取和写入位置可以以彼此不同的位置关系设置。例如,在一些实现中,专用读取位置和写入位置可以在流通池通道的同一表面上(例如,在流通池通道的基面上)。此外,在一些实现中,机器写入的多核苷酸可以在整个部分中从写入位置移动到读取位置,而不是逐阱移动(例如,珠可以移动到包括要读取的数据的区域,并且该区域中所有链的拷贝可以结合到珠并被运输,而不是基于逐阱结合和/或运输链)。

[0321] IX. 杂项

[0322] 包括专利、专利申请和文章在内的所有参考文献都通过引用整体明确地并入本文。

[0323] 提供前述描述是为了使本领域技术人员能够实践这里描述的各种配置。虽然已经参照各种附图和配置具体描述了本主题技术,但是应当理解,这些仅仅是为了说明的目的,并且不应当被视为限制本主题技术的范围。

[0324] 如本文所使用的,以单数形式叙述并且以单词“一(a)”或“一个(an)”开始的元件或步骤应当被理解为不排除多个所述元件或步骤,除非明确陈述了这种排除。此外,提及“一个实现”并不旨在解释为排除也包含所陈述的特征的另外的实现的存在。另外,除非明确地相反指出,否则“包括(comprising)”或“具有(having)”具有特定特性的一个元件或多个元件的实现可以包括另外元件,而不管它们是否具有该特性。

[0325] 在整个本说明书中使用的术语“基本上”和“大约”用于描述并考虑到例如由于处理中的变化而引起的小波动。例如,它们可以指小于或等于 $\pm 5\%$ 、例如小于或等于 $\pm 2\%$ 、例如小于或等于 $\pm 1\%$ 、例如小于或等于 $\pm 0.5\%$ 、例如小于或等于 $\pm 0.2\%$ 、例如小于或等

于 $\pm 0.1\%$ 、例如小于或等于 $\pm 0.05\%$ 。

[0326] 实现主题技术可能有许多其他方式。在不脱离本主题技术的范围的情况下,这里描述的各种功能和元件可以不同于那些示出的功能和元件来划分。对这些实现的各种修改对于本领域技术人员来说是明显的,并且这里定义的一般原理可以应用于其他实现。因此,在不脱离本主题技术的范围的情况下,本领域普通技术人员可以对本主题技术进行许多改变和修改。例如,可以使用不同数量的给定模块或单元,可以使用不同类型或类型的给定模块或单元,可以添加给定模块或单元,或者可以省略给定模块或单元。

[0327] 带下划线和/或斜体的标题和副标题仅用于方便,并不限制主题技术,也不涉及对主题技术描述的解释。本领域普通技术人员已知或以后将会知道的贯穿本公开描述的各种实现的元件的所有结构和功能等同物通过引用明确地结合于此,并且旨在被本主题技术所包含。此外,无论在上面的描述中是否明确叙述了这种公开,本文中公开的任何内容都不旨在献给公众。

[0328] 应该认识到,前述概念的所有组合和下面更详细讨论的额外概念(假设这样概念不相互不一致)被设想为本文公开的创造性主题的部分。特别是,出现在本公开的结尾处的所要求保护的主题的所有组合被设想为本文公开的创造性主题的一部分。

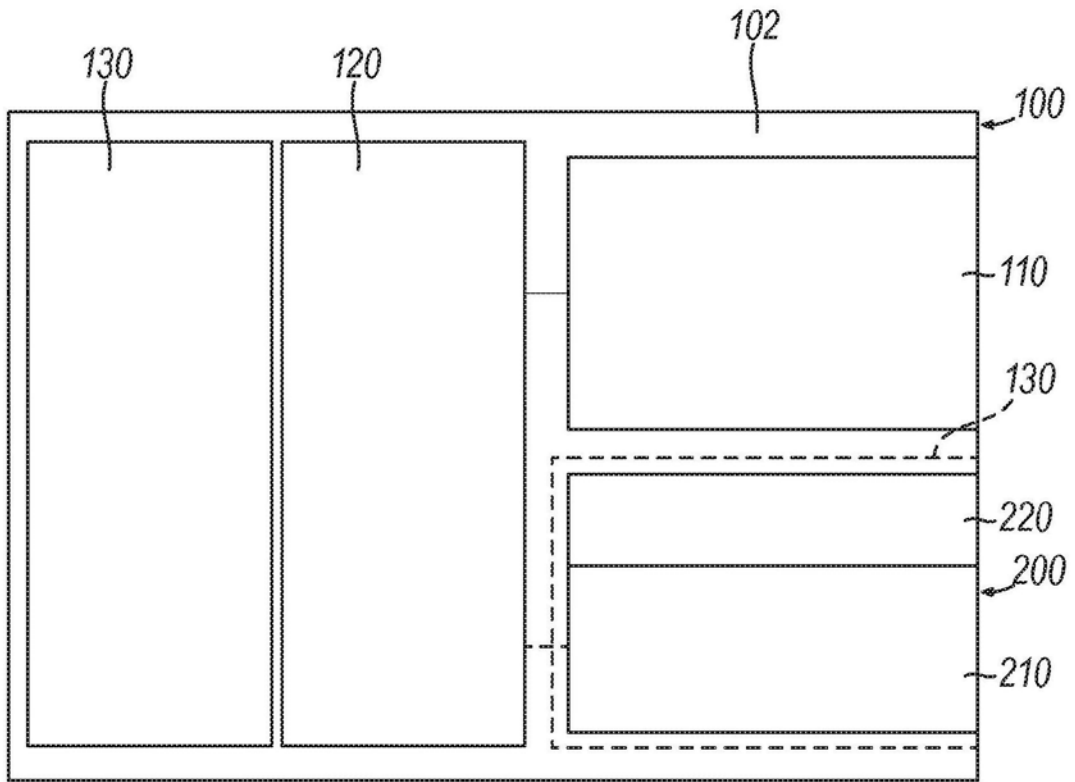


图1

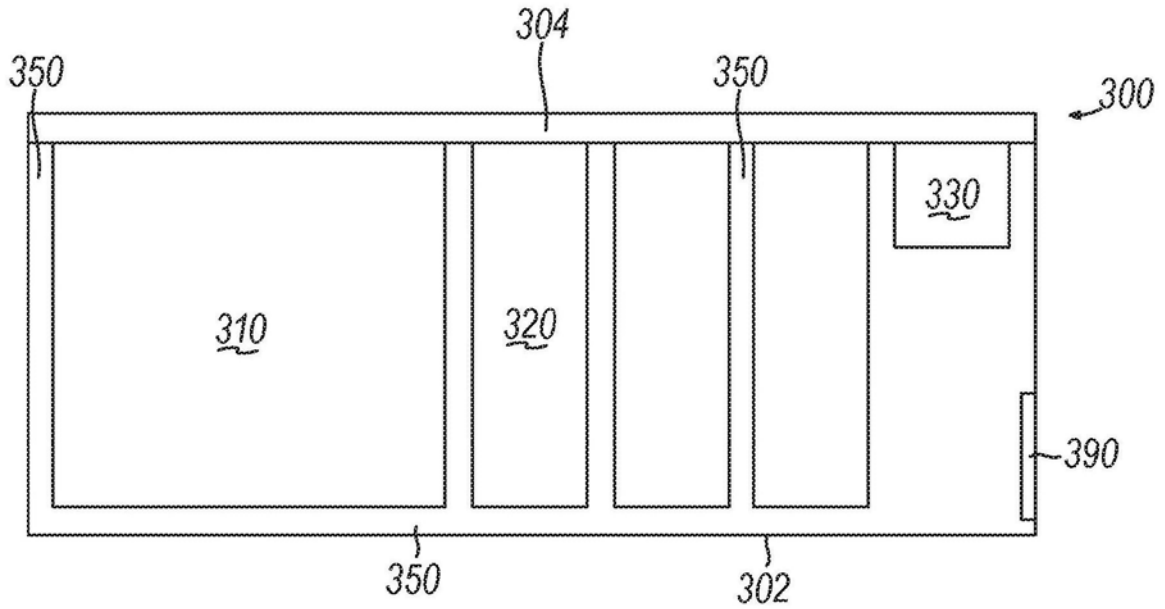


图2

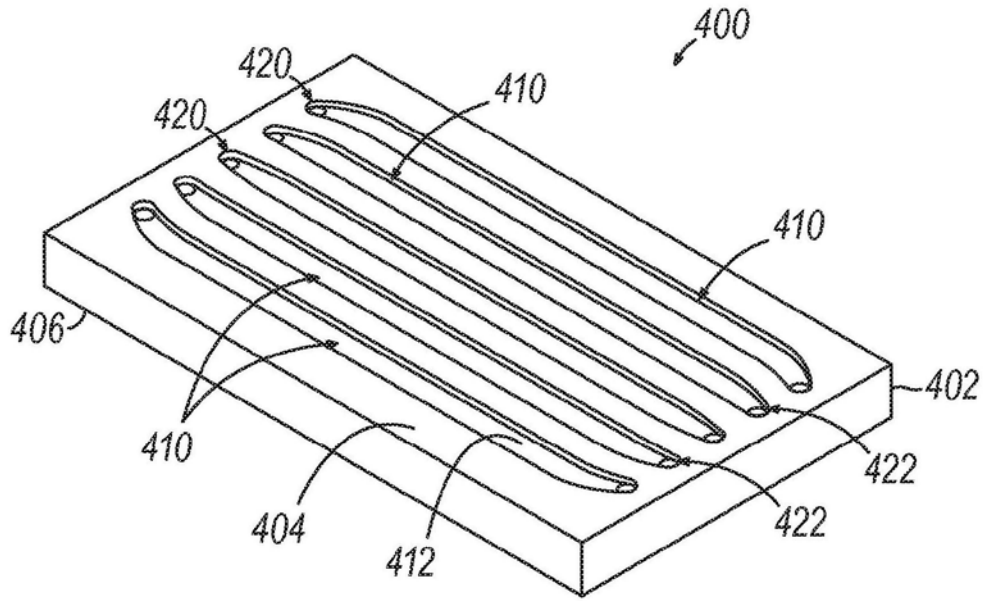


图3

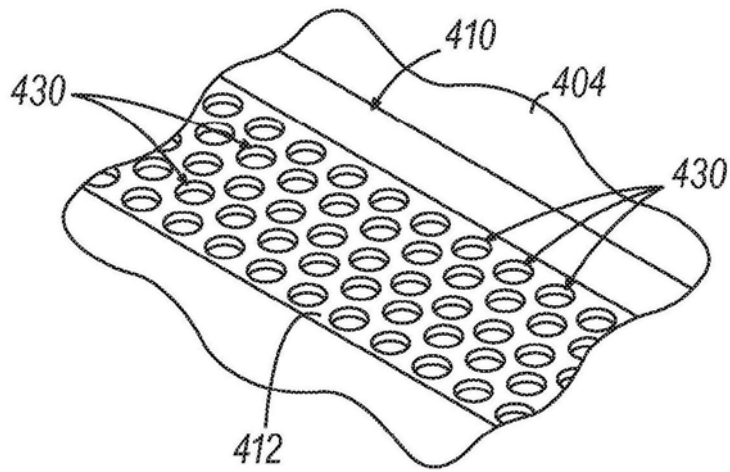


图4



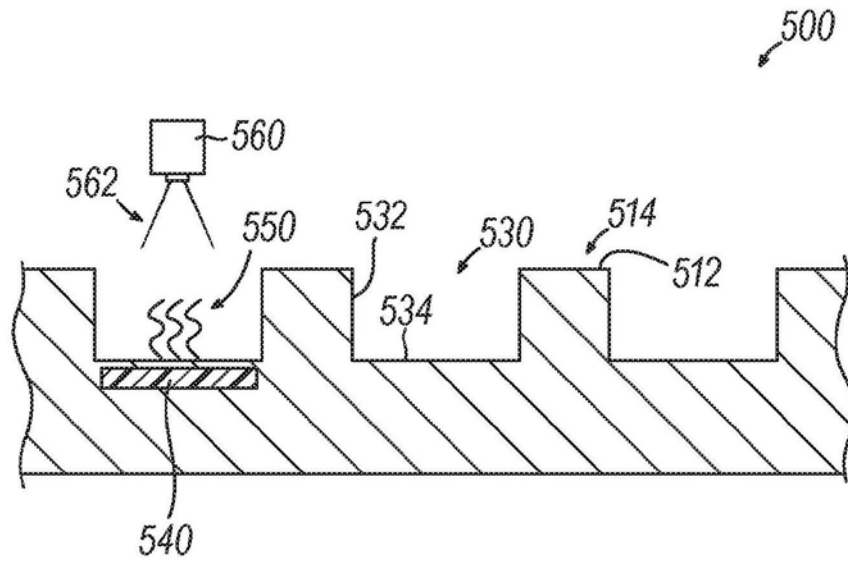


图5

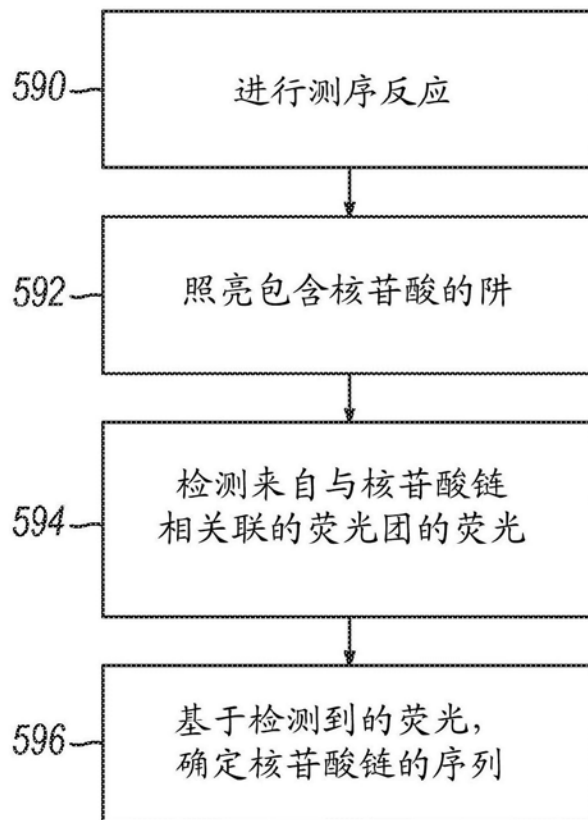


图6

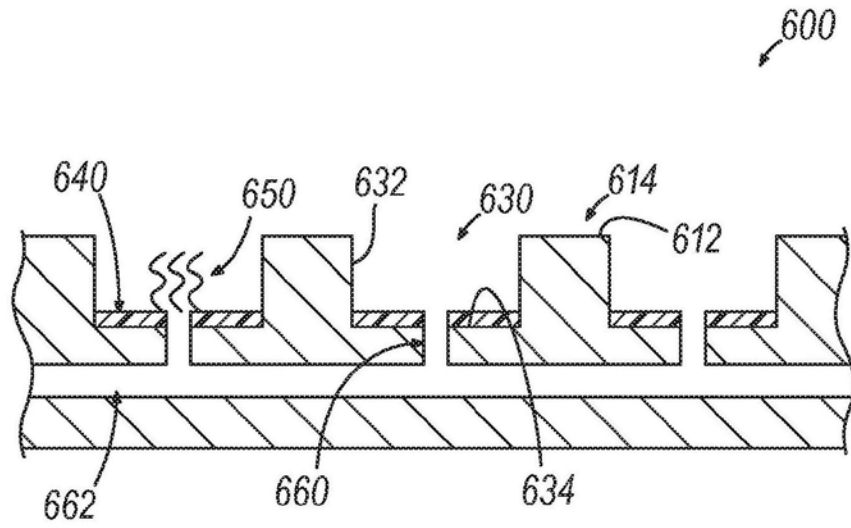


图7

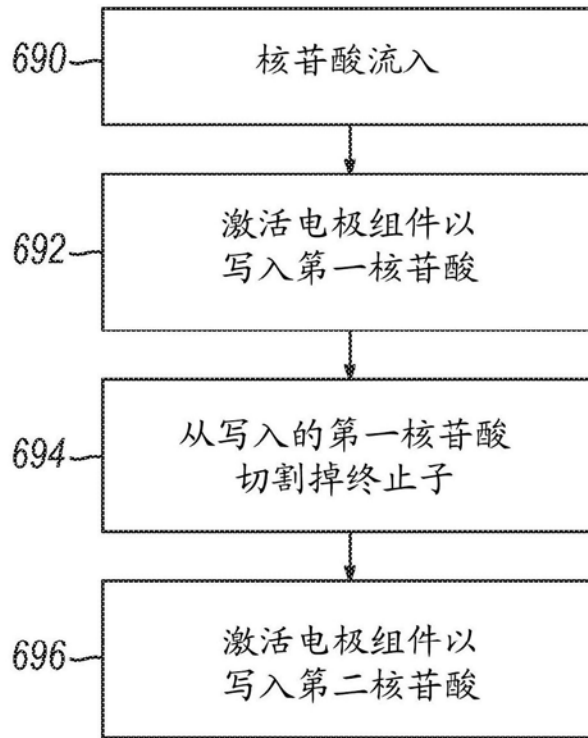


图8

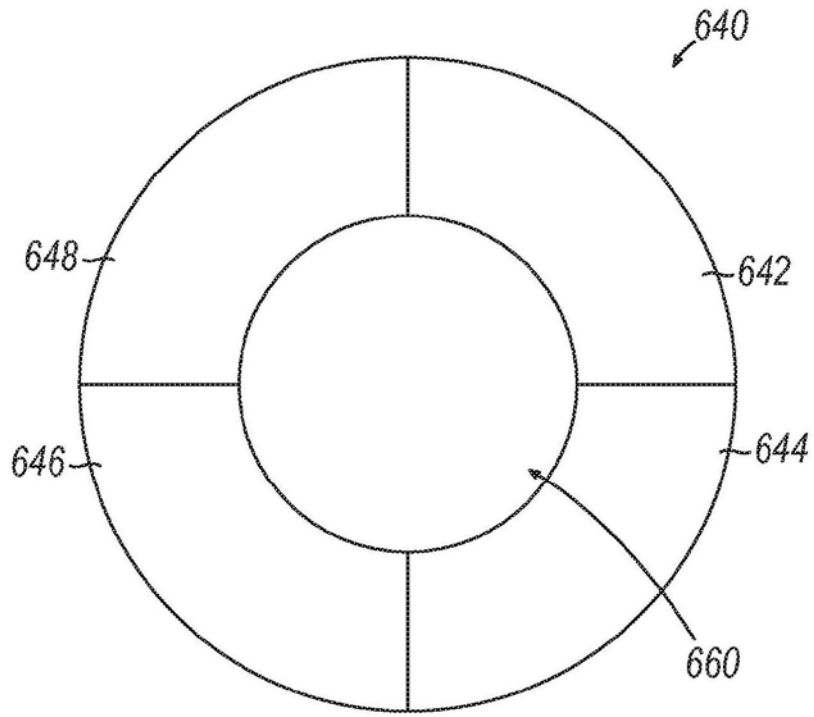


图9

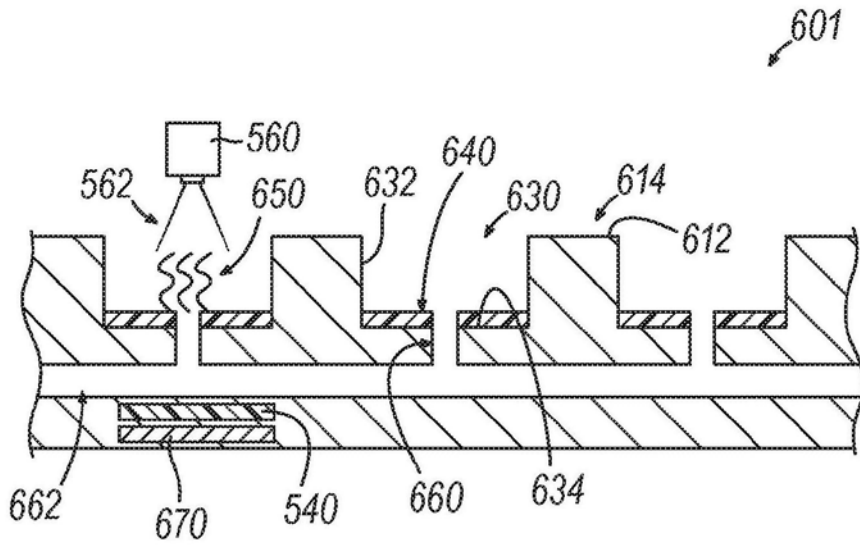


图10

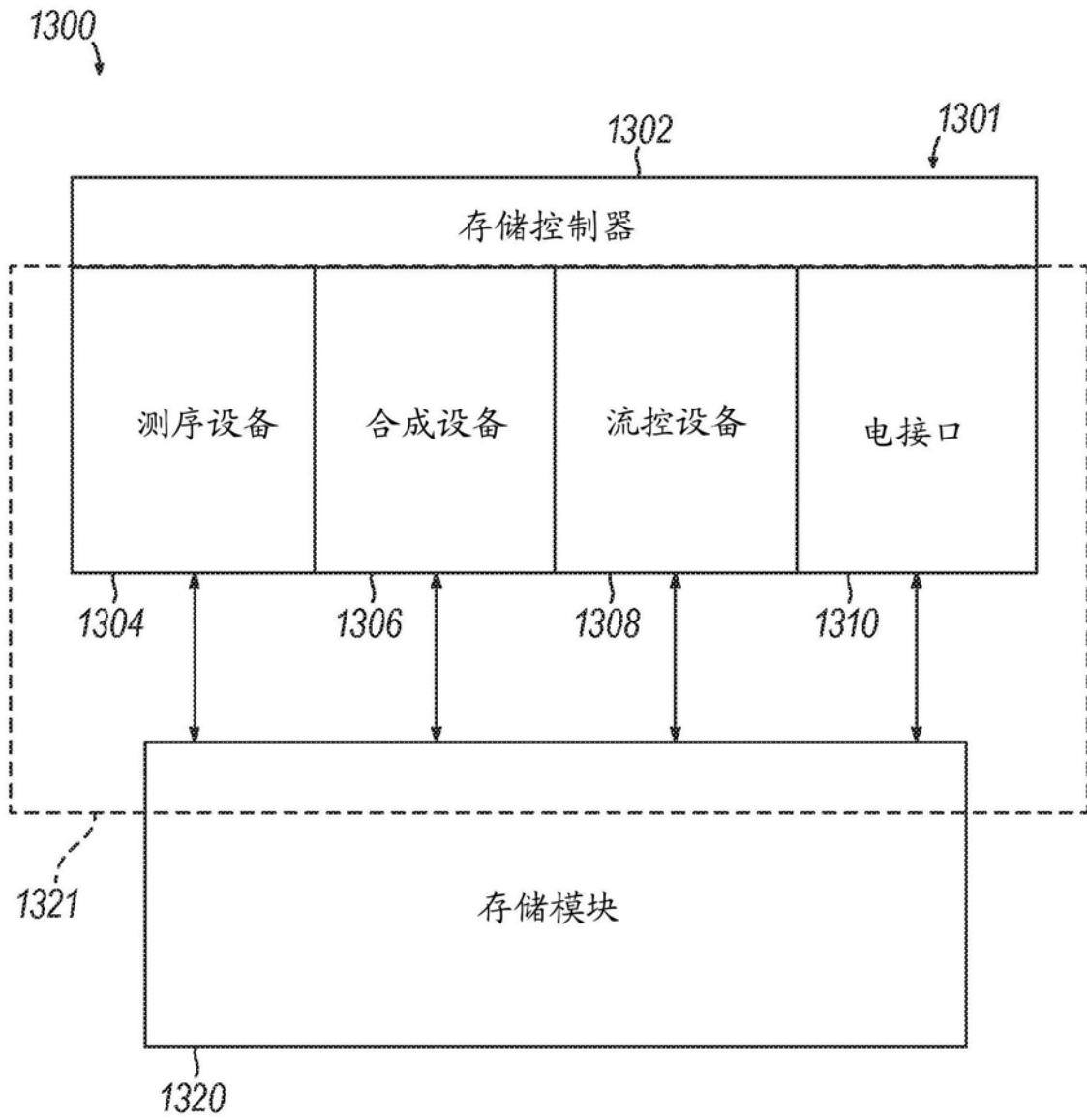


图11

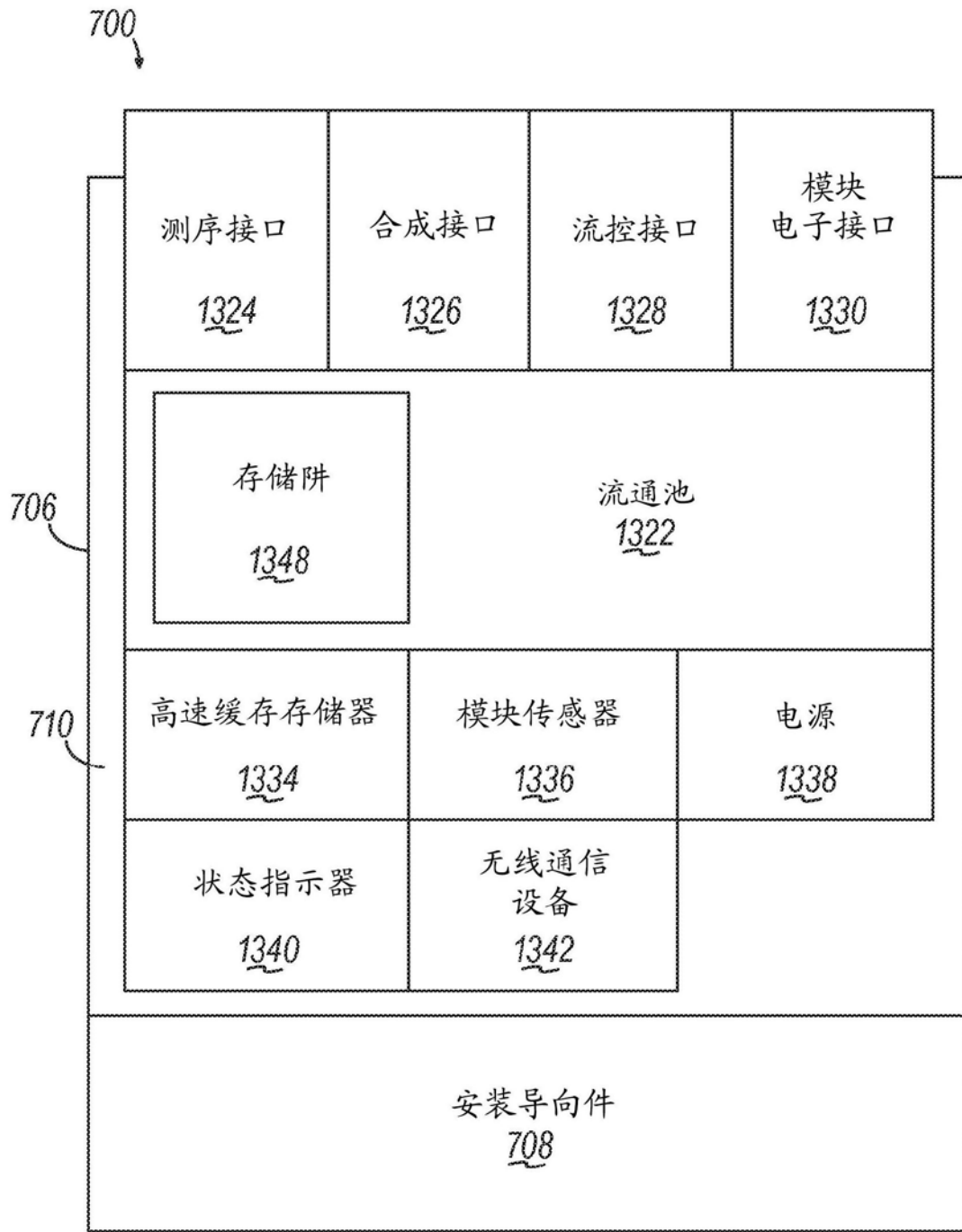


图12A

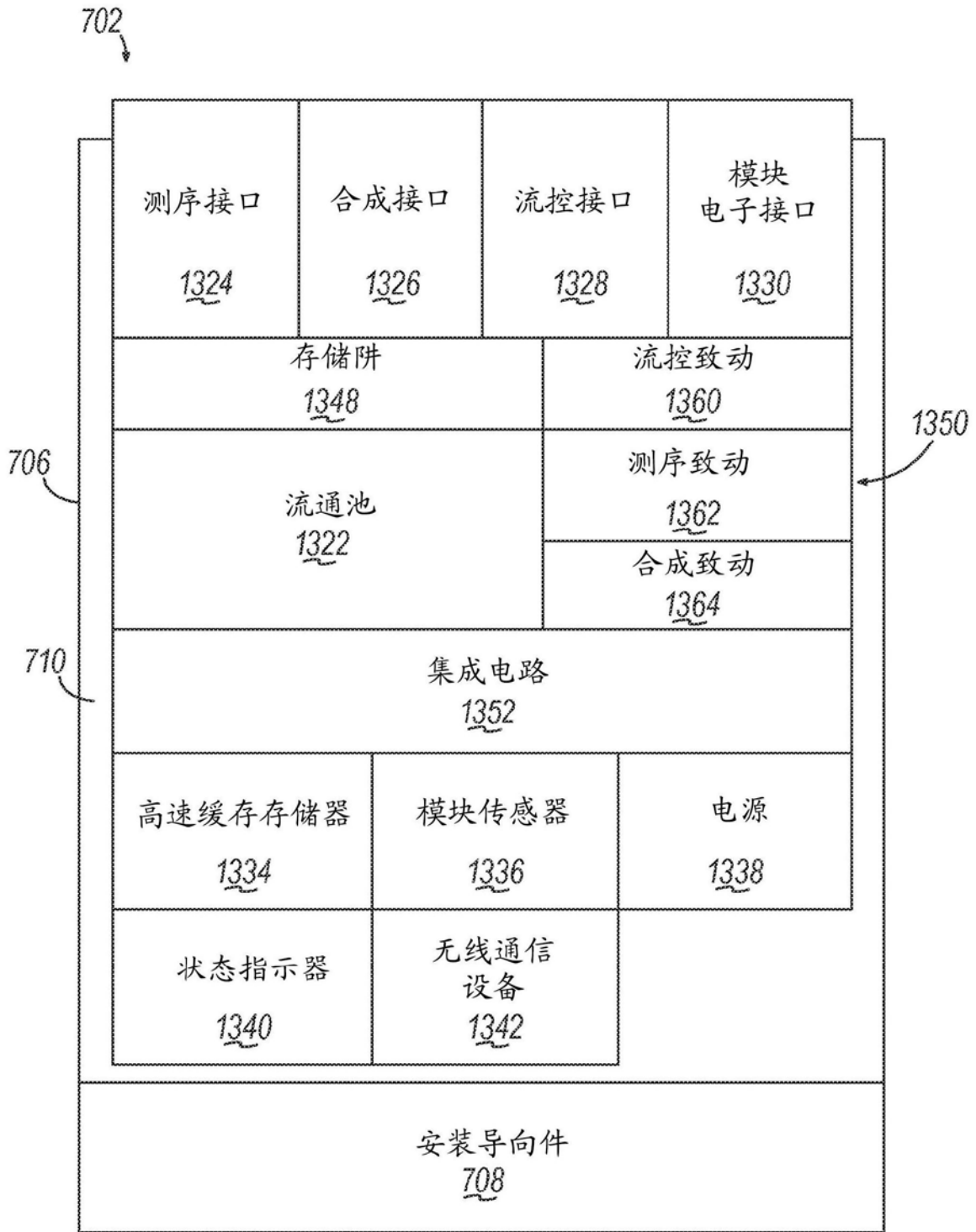


图12B

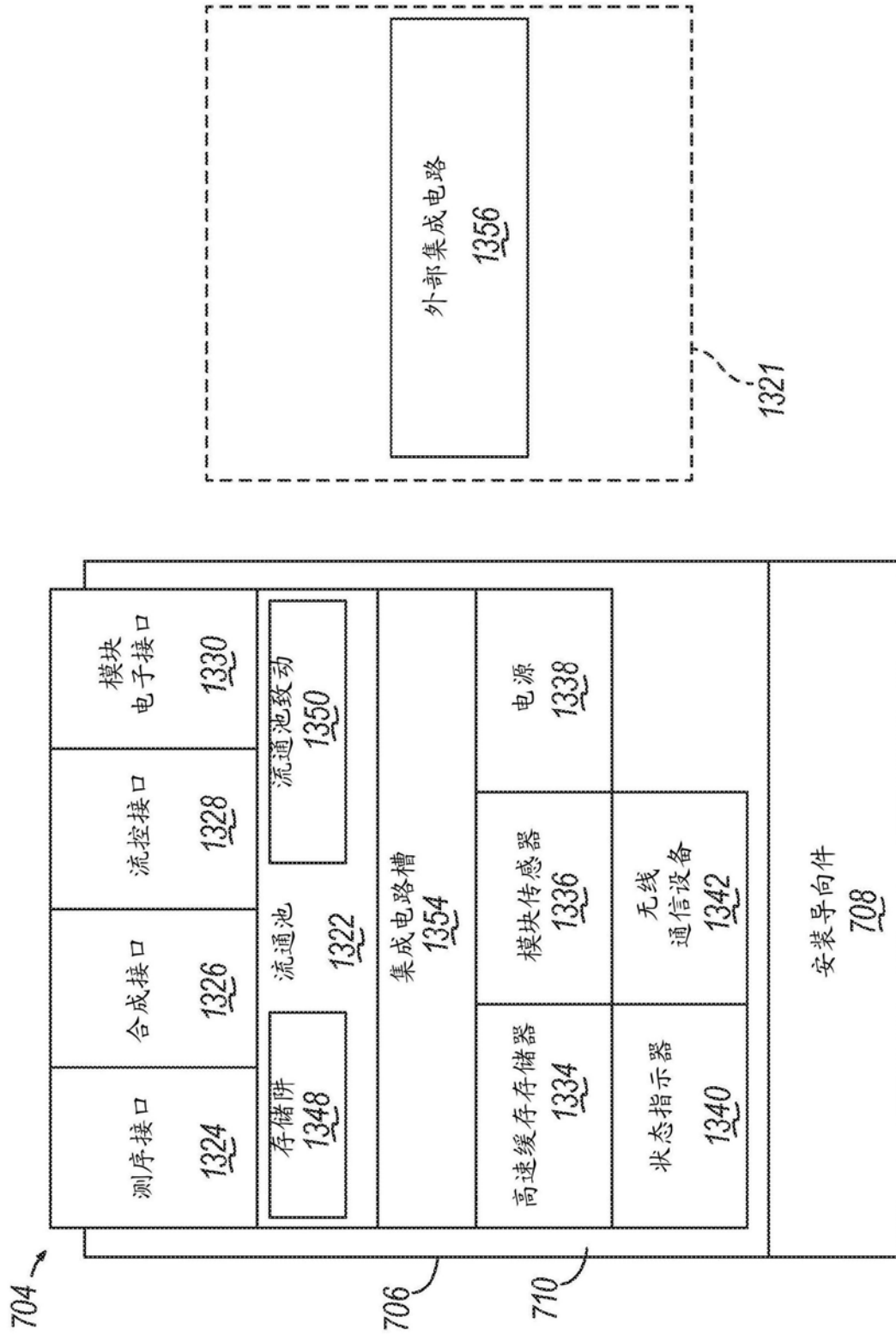


图12C

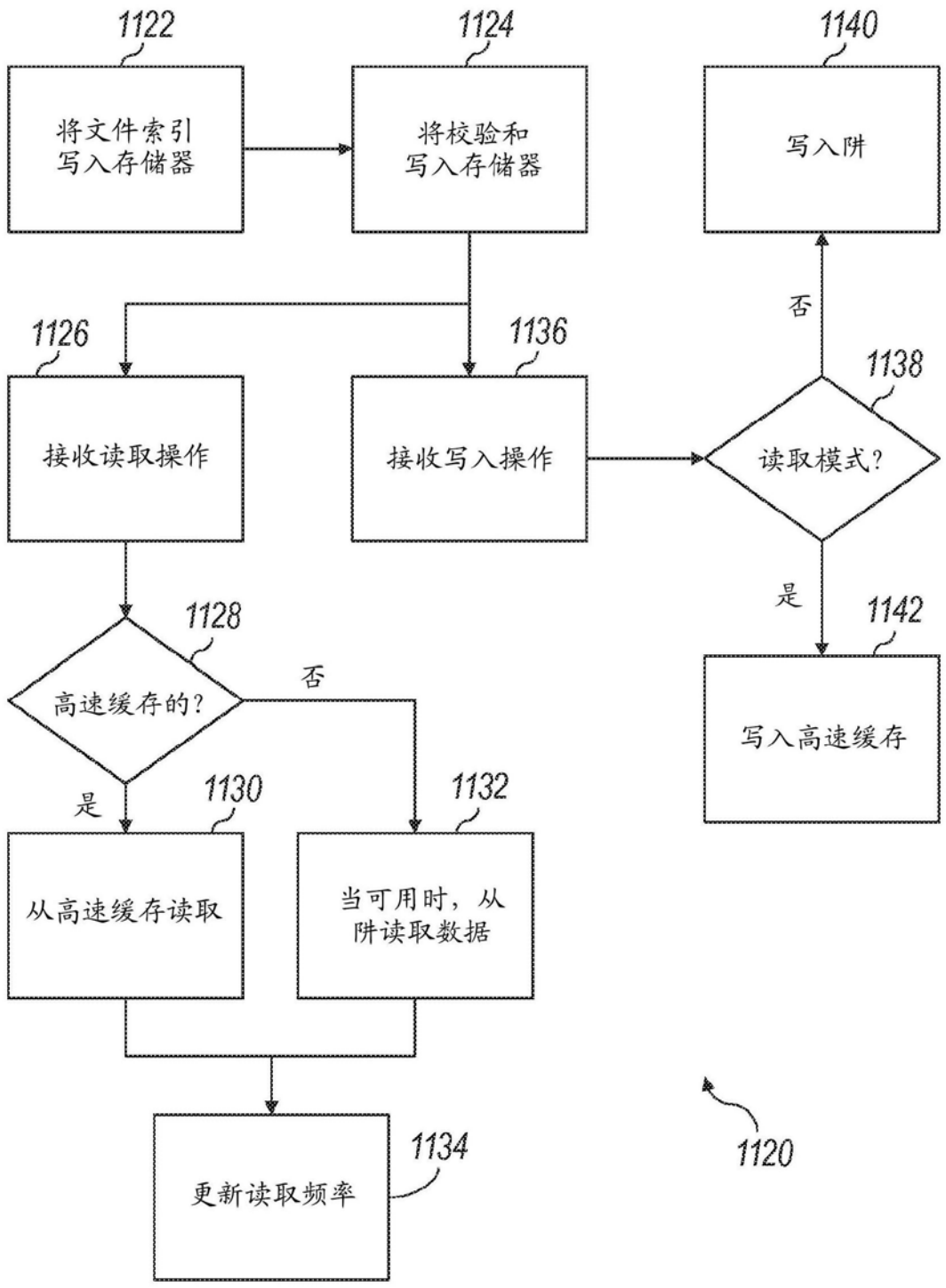


图13



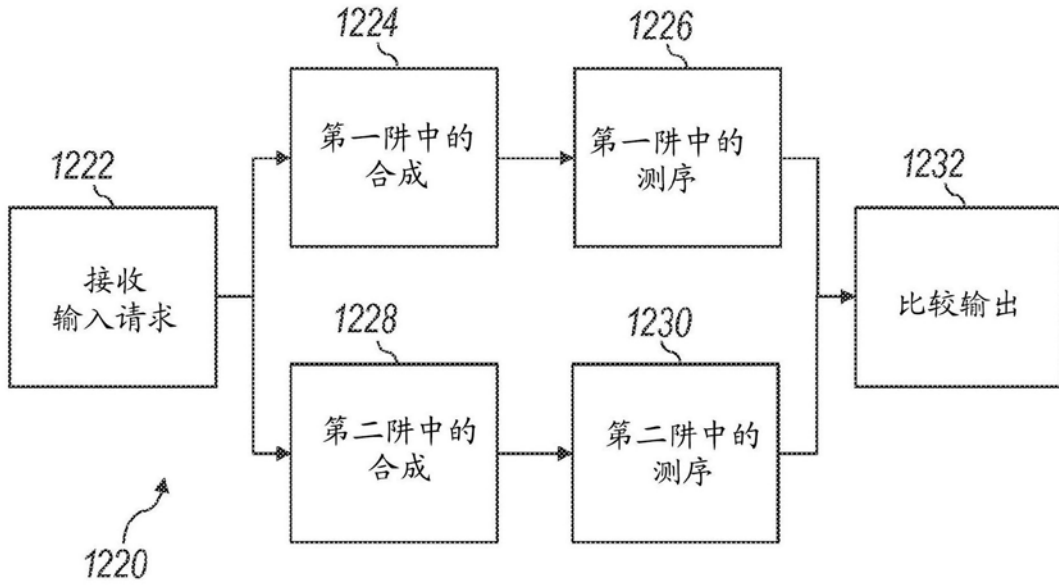


图14

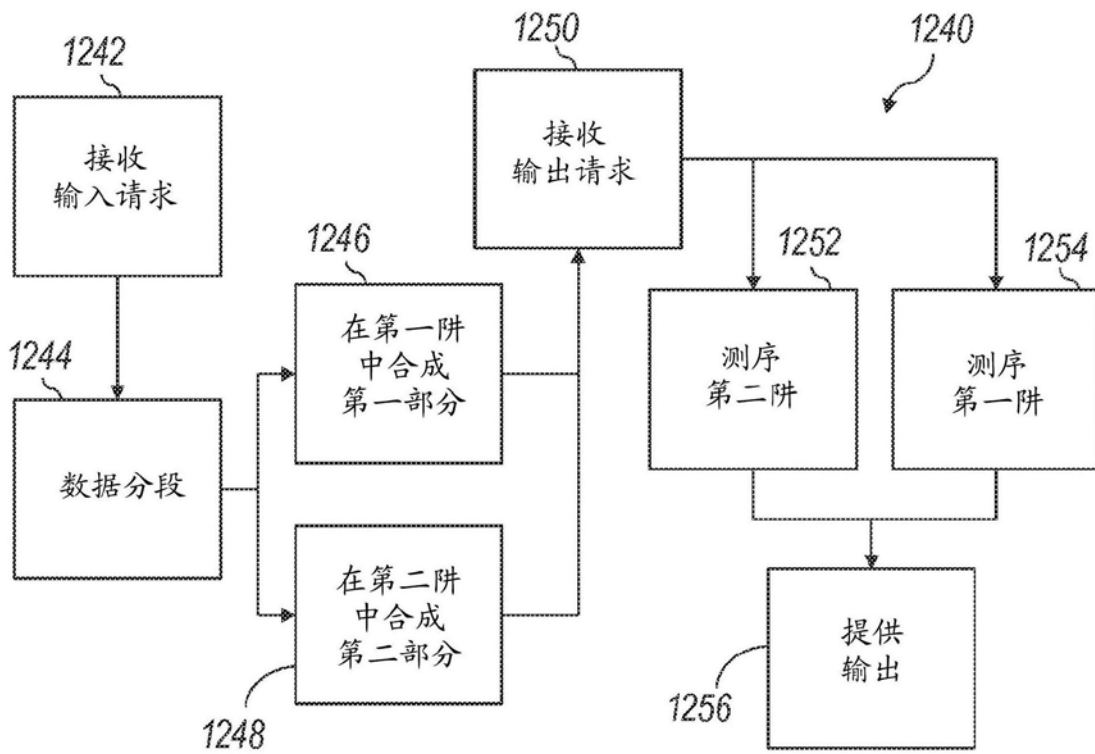


图15

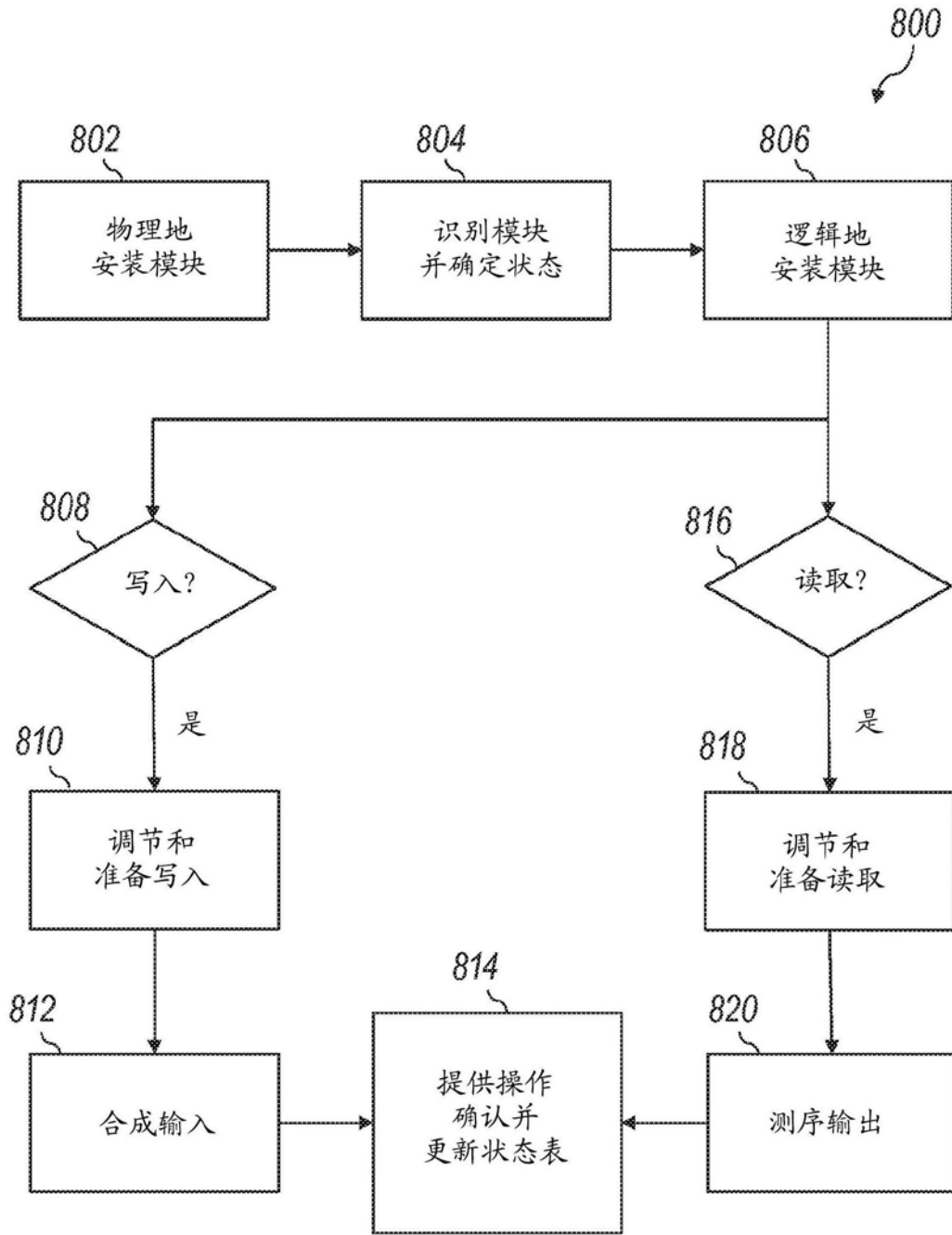


图16A

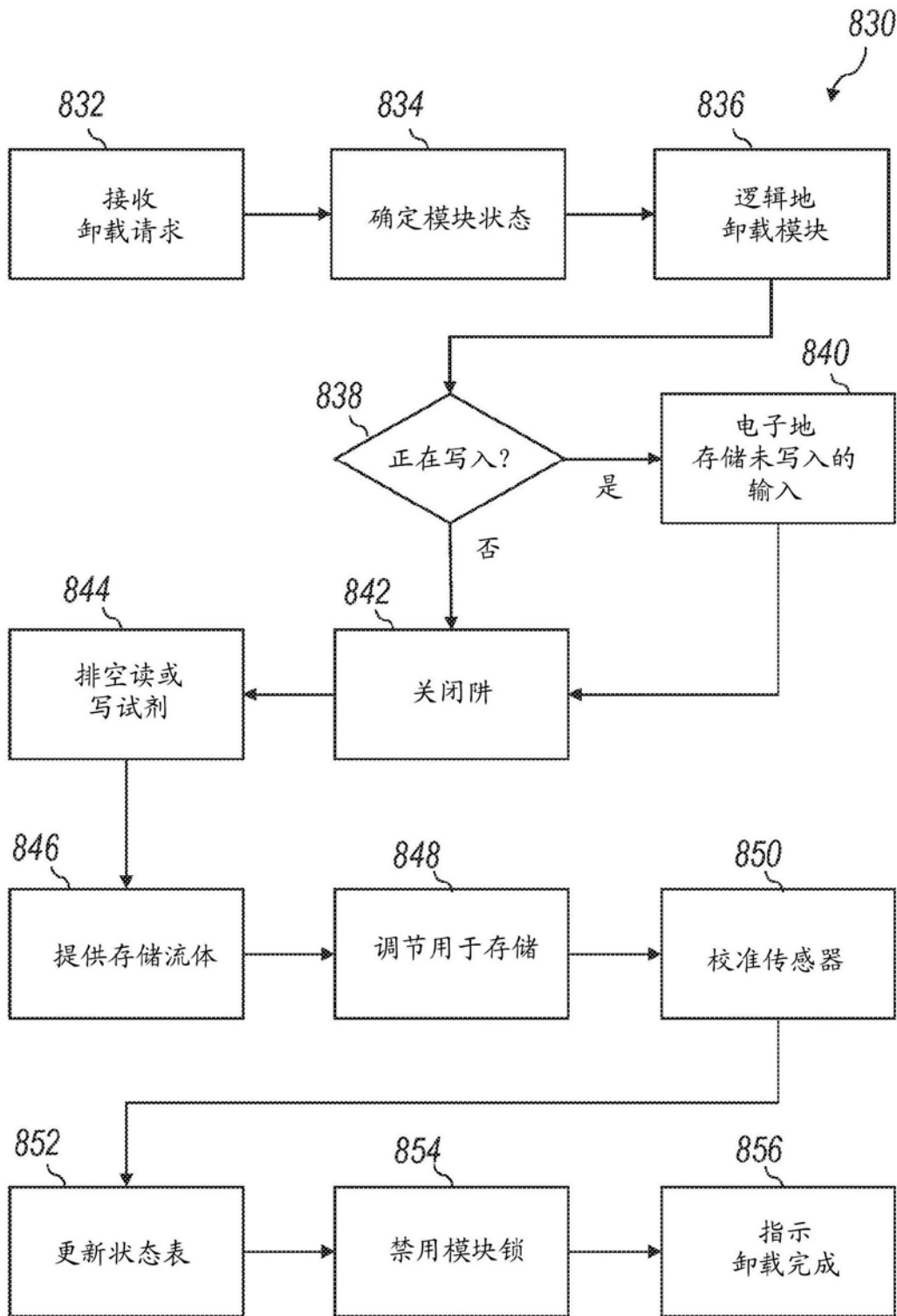


图16B

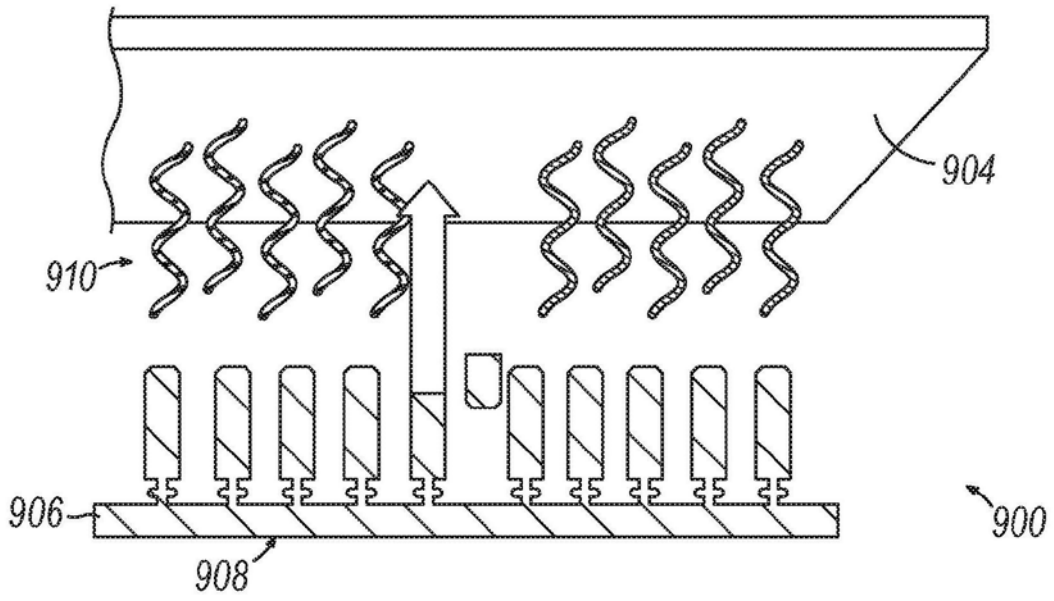


图17A

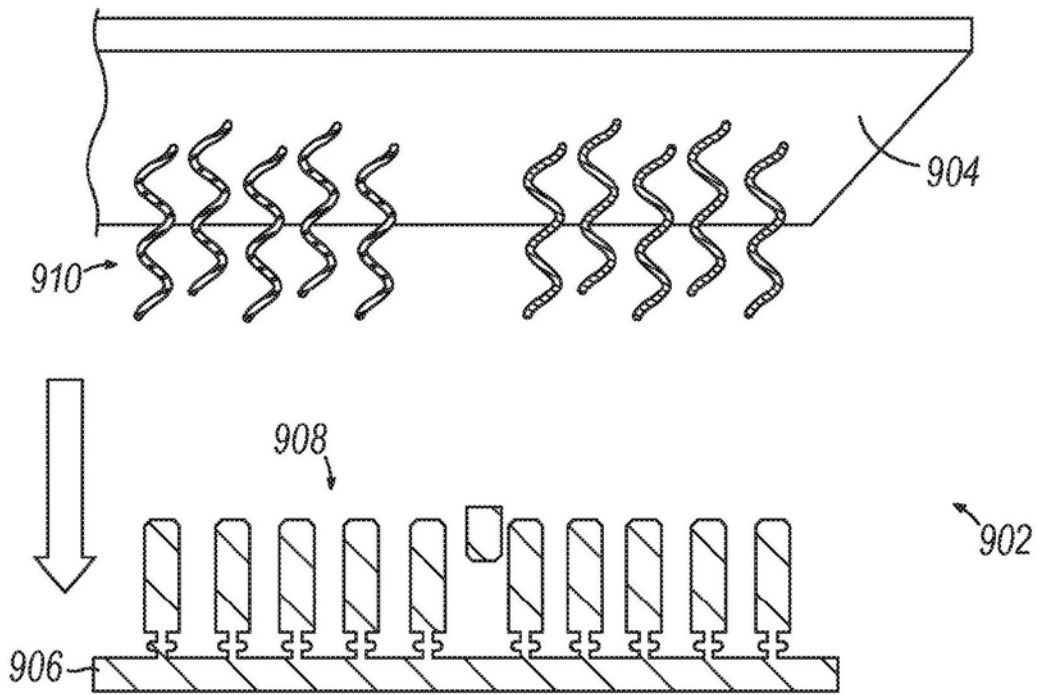


图17B

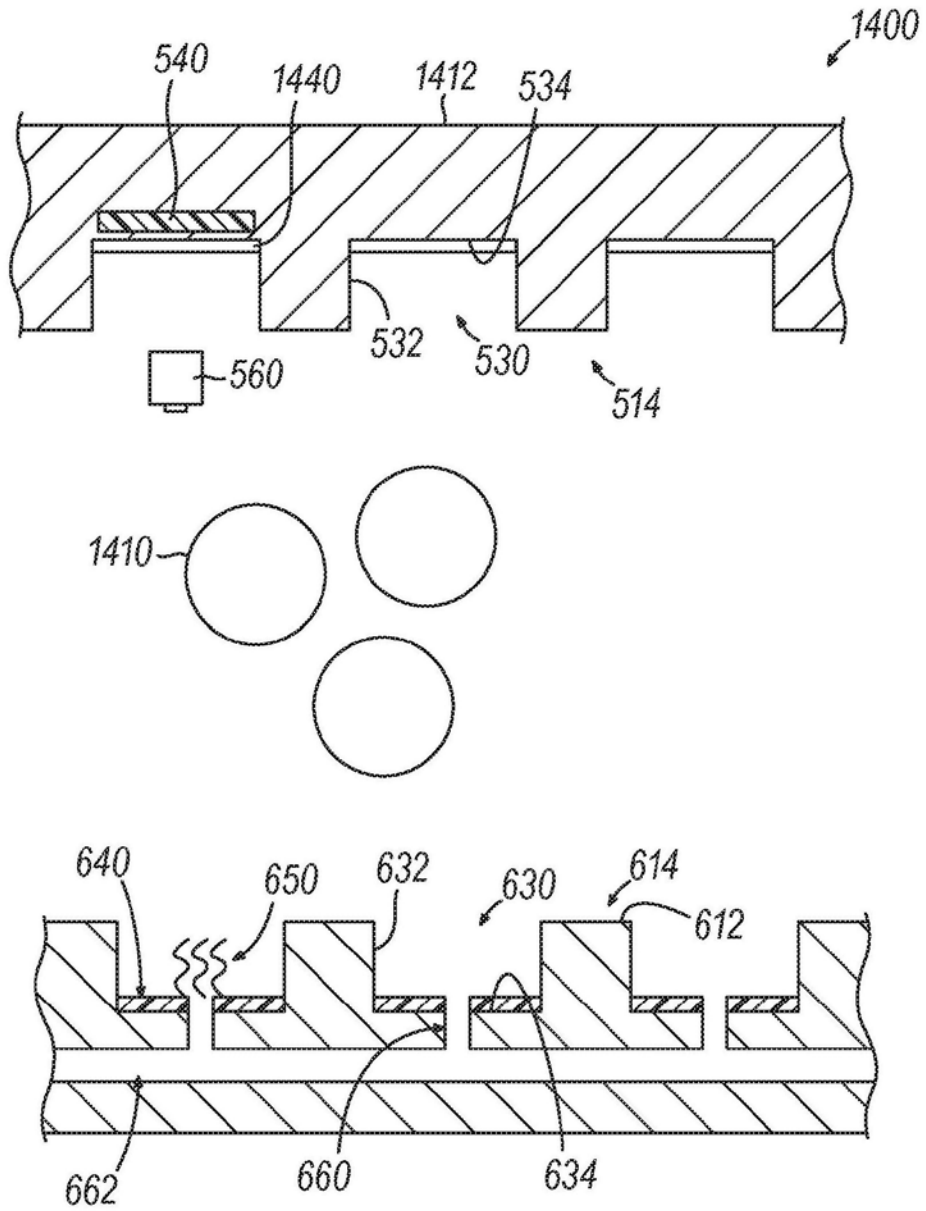


图18

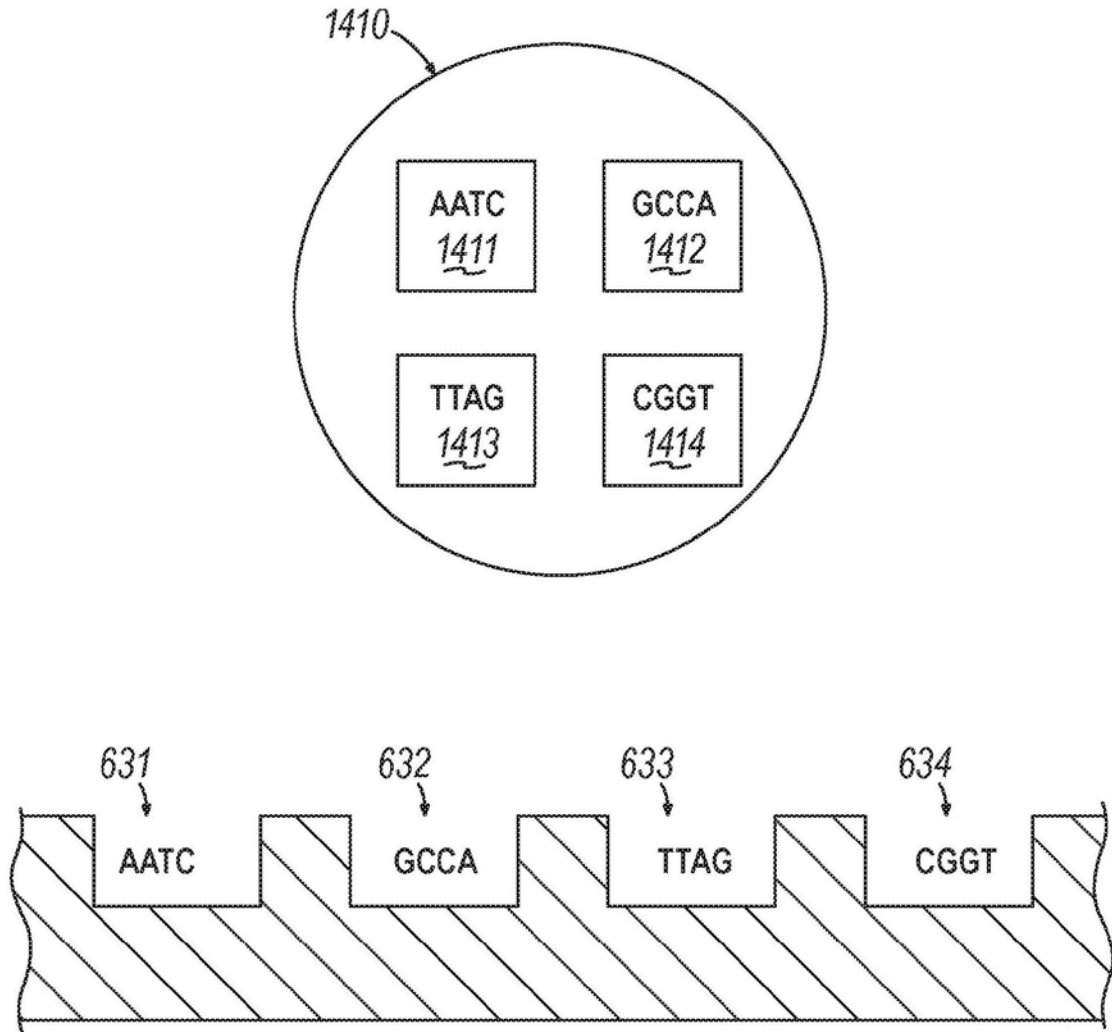


图19