(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 118202097 A
(43)申请公布日 2024.06.14

(21)申请号 202280029417.3
(22)申请日 2022.03.04
(30)优先权数据
63/156,858 2021.03.04 US
(85)PCT国际申请进入国家阶段日
2023.10.19
(86)PCT国际申请的申请数据
PCT/US2022/019045 2022.03.04
(87)PCT国际申请的公布数据
WO2022/187719 EN 2022.09.09
(71)申请人 俄勒冈健康与科学大学
地址 美国俄勒冈州
(72)发明人 C·M·丹巴赫尔
(74)专利代理机构 北京律盟知识产权代理有限责任公司 11287
专利代理师 刘媛媛
(51)Int.Cl.
C40B 70/00(2006.01)
C12N 15/10(2006.01)
B82Y 5/00(2006.01)
B01J 19/00(2006.01)

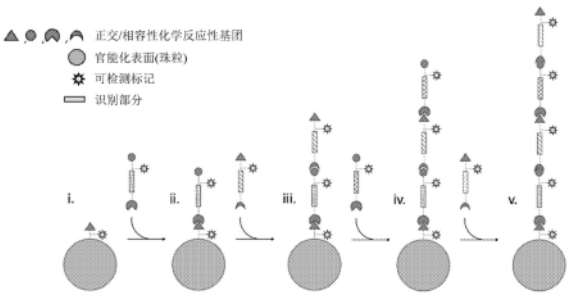
权利要求书4页 说明书50页
序列表（电子公布） 附图33页

(54)发明名称

分子条形码以及相关方法和系统

(57)摘要

本公开涉及分子条形码,其包括由一或多个标识体构成的集合。公开了使用所述分子条形码如将材料分类的方法。本文公开了制备所述分子条形码的方法和用于使用所述分子条形码的系统。



CN 118202097 A

1.一种分子条形码,其包括由一或多个标识体构成的集合,所述由标识体构成的集合中的每个标识体都包括由一或多个可检测标记构成的集合,所述一或多个可检测标记操作性地连接到支架部分,所述支架部分包括:

识别部分,所述识别部分与所述由标识体构成的集合中的至少一个标识体的识别部分正交;以及

切割位点,所述切割位点操作性地连接到所述由一或多个可检测标记构成的集合,所述一或多个可检测标记以三维布置的形式彼此级联以编码和形成所述分子条形码。

2.根据权利要求1所述的分子条形码,由此应用与所述识别部分具有正交反应性的切割剂在至少一个标识体的所述切割位点处切割所述分子条形码,以根据所述标识体的正交性使所述一或多个可检测标记从三维布置中解离,并且由此诱导可检测信号响应以解码所述分子条形码。

3.根据权利要求1或权利要求2所述的分子条形码,其中所述支架部分包括多肽、氨基酸、环肽、核酸或其组合。

4.根据权利要求1至3中任一权利要求所述的分子条形码,其中所述支架部分包括双链DNA(dsDNA),所述识别部分包括核酸酶识别部分,并且所述一或多个可检测标记包括荧光团。

5.根据权利要求1至4中任一权利要求所述的分子条形码,其中所述支架部分包括双链DNA,所述识别部分包括蛋白酶识别部分,并且所述一或多个可检测标记包括荧光团。

6.根据权利要求1至4中任一权利要求所述的分子条形码,其中所述支架部分包括双链DNA,所述识别部分包括化学切割识别部分,并且所述一或多个可检测标记包括荧光团。

7.根据权利要求1至4中任一权利要求所述的分子条形码,其中所述支架部分包括氨基酸,所述识别部分包括蛋白酶识别部分,并且所述一或多个可检测标记包括荧光团。

8.根据权利要求1至4中任一权利要求所述的分子条形码,其中所述支架部分包括氨基酸,所述识别部分包括化学切割识别部分,并且所述一或多个可检测标记包括荧光团。

9.根据权利要求1至8中任一权利要求所述的分子条形码,其中所述识别部分包括所述切割位点。

10.根据权利要求1至9中任一权利要求所述的分子条形码,其进一步包括位于所述由标识体构成的集合中的每个标识体之间的标识体接头。

11.根据权利要求10所述的分子条形码,其中所述标识体接头选自以下群组:点击化学部分、天然化学连接接头和酶介导的连接接头。

12.根据权利要求1至11中任一权利要求所述的分子条形码,其中所述由标识体构成的集合中的至少一个标识体的所述识别部分包括化学接头、肽或核酸。

13.根据权利要求1至12中任一权利要求所述的分子条形码,其中所述由标识体构成的集合中的所述可检测标记中的至少一或多个可检测标记包括荧光团。

14.根据权利要求1至13中任一权利要求所述的分子条形码,其中所述由标识体构成的集合中的每个识别部分都包括至少一个选自以下的部分:蛋白酶识别部分、核酸内切酶识别部分、能通过亲和试剂识别的表位、核酸探针识别部分、经修饰的肽侧链和非天然肽侧链。

15.根据权利要求1至14中任一权利要求所述的分子条形码,其中所述分子条形码附接

到珠粒。

16.根据权利要求15所述的分子条形码,其中所述珠粒进一步附接到测试剂。

17.根据权利要求15所述的分子条形码,其中所述珠粒进一步附接到化学结构单元。

18.根据权利要求1至17中任一权利要求所述的分子条形码,其进一步包括与所述分子条形码的所述三维布置具有互信息的材料。

19.一种将材料分类的方法,所述方法包括:

向环境中引入由一或多个根据权利要求1至18中任一权利要求所述的分子条形码构成的集合,其中所述集合中的至少一个分子条形码与材料具有互信息;

将正交反应性切割剂选择性地应用于所述由分子条形码构成的集合以在与所述切割剂具有反应性的所述切割位点处切割所述分子条形码,从而根据所述标识体的正交性使标识体的所述可检测标记从所述三维布置中解离,并且由此诱导可检测信号响应以解码所述分子条形码。

20.一种筛选方法,其包括:

将多个测试构建体引入到环境中,其中每个测试构建体都包括与根据权利要求1至18中任一权利要求所述的分子条形码操作性地偶联的测试剂,其中所述分子条形码的所述三维布置与所述测试剂具有互信息;

针对由一或多个靶标构成的集合筛选所述多个测试构建体;

检测所述测试构建体中的一或多个测试构建体的活性;

将正交反应性切割剂选择性地应用于所述多个测试构建体以在与所述切割剂具有反应性的所述切割位点处切割偶联的分子条形码,从而根据所述标识体的正交性使标识体的所述可检测标记从所述三维布置中解离,并且由此诱导可检测信号响应以解码所述偶联的条形码,并由此标识具有针对所述由靶标构成的集合的活性的测试剂。

21.根据权利要求19或权利要求20所述的方法,其中所述可检测信号响应是在不进行富集的情况下诱导的。

22.根据权利要求20或权利要求21所述的方法,其中将正交反应性切割剂选择性地应用于所述多个测试构建体包括依序地应用所述切割试剂。

23.根据权利要求22所述的方法,其中依序地应用所述切割试剂包括迭代地应用单一切割剂或依序地应用两种或更多种切割剂。

24.根据权利要求19至23中任一权利要求所述的方法,其中所述可检测信号响应包括检测又一个或多个解离的可检测标记的存在或不存在或调节信号频率或信号幅度。

25.一种用于产生分子条形码的方法,所述方法包括:

提供由一或多个标识体构成的集合,所述由标识体构成的集合中的每个标识体都包括:由一或多个可检测标记构成的集合,所述一或多个可检测标记操作性地连接到支架部分,所述支架部分包括识别部分和切割位点,所述识别部分与所述由标识体构成的集合中的至少一个标识体的识别部分正交,所述切割位点操作性地连接到由一或多个可检测标记构成的集合;

从所述由标识体构成的集合中选择第一标识体和第二标识体,并且将所述第一标识体与所述第二标识体以三维布置的方式级联,以编码和形成分子条形码;

(任选地)从所述由标识体构成的集合中选择一个标识体,并且将所述标识体与所述分

子条形码级联,以修改所述三维布置并且进一步编码和形成所述分子条形码;

(任选地)重复一次至n次所述从所述由标识体构成的集合中选择一个标识体并且将所述标识体与所述分子条形码级联的步骤。

26.根据权利要求25所述的方法,其进一步包括将所述第一标识体操作性地附接到材料。

27.根据权利要求26所述的方法,其中所述材料为表面。

28.根据权利要求27所述的方法,其中所述表面包括微粒或珠粒的外表面或内表面。

29.根据权利要求27或权利要求28所述的方法,其包括将所述第一标识体通过高亲和力结合蛋白操作性地附接到所述表面。

30.根据权利要求25所述的方法,其进一步包括将所述第一标识体附接到三臂接头。

31.根据权利要求25所述的方法,其中提供由一或多个标识体构成的集合进一步包括将所述由标识体构成的集合中的每个标识体配置成与在前的标识体级联并且促进标识体以线性三维布置的方式级联,并且由此依序地形成和编码所述分子条形码。

32.根据权利要求31所述的方法,其中将所述由标识体构成的集合配置成与在前的标识体特异性地结合包括通过连接到所述在前的标识体、从所述在前的标识体延伸或合成到所述在前的标识体上进行级联。

33.根据权利要求25至33中任一权利要求所述的方法,其进一步包括:

提供由一或多个化学结构单元构成的集合,所述由化学结构单元构成的集合中的每个化学结构单元都被配置成与在前的化学结构单元级联;

从所述由化学结构单元构成的集合中选择第一化学结构单元,并且将所述第一化学结构单元与所述第一标识体级联;

(任选地)从所述由被配置成与所述在前的化学结构单元级联的化学结构单元构成的集合中选择一个化学结构单元,并且将所述化学结构单元与所述在前的化学结构单元级联,并由此形成一系列化学结构单元;以及

(任选地)重复一次至N次所述从所述由被配置成与所述在前的化学结构单元级联的化学结构单元构成的集合中选择一个化学结构单元并且将所述化学结构单元与所述在前的化学结构单元级联的步骤。

34.根据权利要求33所述的方法,其中将所述系列的化学结构单元级联包括使用非核酸相容性反应。

35.一种用于编码和解码分子条形码的系统,所述系统包括:

条形码编码系统,所述条形码编码系统被配置成向环境中引入由一或多个经编码的根据权利要求1至8中任一权利要求所述的分子条形码构成的集合,其中所述集合中的至少一个分子条形码与材料具有互信息;

切割系统,所述切割系统被配置成将正交反应性切割剂选择性地应用于环境以在标识体的具有与所述切割剂具有反应性的识别部分的切割位点处切割所述由分子条形码构成的集合,从而根据所述标识体的正交性使所述可检测标记从所述由分子条形码构成的集合中的分子条形码的三维布置中解离,并且由此从一个中诱导可检测信号响应以解码所述由分子条形码构成的集合;

(任选地)照明系统,所述照明系统被配置成向所述由分子条形码构成的集合选择性地

传递光以诱导所述可检测信号响应;

检测系统,所述检测系统用于从所述由分子条形码构成的集合中检测所述可检测信号响应;以及

处理器,所述处理器与所述照明系统和所述光检测系统耦接,并且被配置成促进解码所述由分子条形码构成的集合。

分子条形码以及相关方法和系统

[0001] 版权声明

[0002] ©2022俄勒冈健康与科学大学。本专利文件的公开内容的一部分含有受版权保护的材料。版权所有人不反对任何人对本专利文件或专利公开内容进行复制再现,因为其出现在专利与商标局的专利文档和记录中,但无论如何在其它方面保留所有版权。37CFR§1.71(d)。

技术领域

[0003] 本公开涉及生物技术。更具体地,本公开涉及分子条形码以及相关方法和系统。

附图说明

[0004] 图1为用于非DNA条形码的构建步骤的图形表示。

[0005] 图2为用于与下一代测序(NGS)条形码相关的条形码的构建步骤的图形表示。

[0006] 图3为非核酸条形码的图形表示和条形码去卷积期间的潜在成像结果。

[0007] 图4为用于与NGS条形码相关的未标记的条形码的构建步骤的图形表示。这些可以是任选地标记的。

[0008] 图5为用于在不使用DNA的情况下编码化学文库的构建步骤的图形表示。

[0009] 图6为用于与NGS条形码相关的可切割条形码的构建步骤的图形表示。

[0010] 图7为用于标识非DNA分子条形码的解码步骤的图形表示。

[0011] 图8为用于标识与NGS条形码相关的条形码的解码步骤的图形表示。

[0012] 图9为具有核酸内切酶识别位点的双链DNA NGS相容性条形码和条形码去卷积期间的潜在成像结果的图形表示。

[0013] 图10为具有正交蛋白酶识别位点的非核酸NGS相容性条形码和条形码去卷积期间的潜在成像结果的图形表示。

[0014] 图11为视觉NGS条形码的去卷积步骤的图形表示。

[0015] 图12为与化学构建块连接的非DNA条形码的构建的图形表示。

[0016] 图13为非DNA条形码编码的化学文库和条形码去卷积期间的潜在成像结果的图形表示。

[0017] 图14为示范性双标记的双链DNA标识体的图形表示。

[0018] 图15为连接到可检测标记的支架剂的图形表示。

[0019] 图16为具有双标记的环肽构建块的非核酸条形码以及条形码可以如何解码的图形表示。

[0020] 图17为具有可以为用于附接条形码以供单细胞分析的表面的封盖珠粒的微型孔的图形表示。

[0021] 图18为由五个单标记FND标识体构成的集合的文本表示,所述单标记FND标识体具有配置有正交粘性端的支架部分、正交识别部分和切割位点以及经修饰的核苷酸。

[0022] 图19A表示用通过与未标记的dsDNA连接进行附接的第一标记的dsDNA标识体段

(Id1/HindIII-AF750)标记的链霉亲和素珠粒,所述未标记的dsDNA通过生物素部分附接到珠粒。

[0023] 图19B表示用第二标记的dsDNA标识体段(Id2/SpeI-AF647)标记的链霉亲和素珠粒。

[0024] 图19C表示用第三标记的dsDNA标识体段(Id3/XhoI-ATTO550)标记的链霉亲和素珠粒。

[0025] 图19D表示用第四标记的dsDNA标识体段(Id4/NotI-ATTO488)标记的链霉亲和素珠粒。

[0026] 图20A表示在四个荧光通道中成像的未经切割的珠粒:647nm(向上斜条纹)、550nm(向下斜条纹)、488nm(横条纹)和750nm(虚线),并且在右侧绘制了针对每种荧光团获得的平均强度值。

[0027] 图20B表示暴露于NotI酶之后珠粒的平均强度。

[0028] 图20C表示暴露于XhoI酶之后珠粒的平均强度。

[0029] 图20D表示暴露于SpeI酶之后珠粒的平均强度。

[0030] 图21A呈现了条形图,所述条形图表示针对单标记dsDNA标识体链获得的OCS结果,其中循环的顺序是颠倒的以供分析。

[0031] 图21B呈现了条形图,所述条形图表示针对混合标记dsDNA标识体链获得的OCS结果,其中循环的顺序是颠倒的以供分析。

[0032] 图22A表示六个携带相同的用于由经OCS实验的3个循环跟踪的FOV中的单独珠粒生成数据的单标记标识体链序列的I珠粒中的第一珠粒。

[0033] 图22B表示六个携带相同的用于由经OCS实验的3个循环跟踪的FOV中的单独珠粒生成数据的单标记标识体链序列的I珠粒中的第二珠粒。

[0034] 图22C表示携带相同的用于由经OCS实验的3个循环跟踪的FOV中的单独珠粒生成数据的单标记标识体链序列的第三珠粒。

[0035] 图22D表示携带相同的用于由经OCS实验的3个循环跟踪的FOV中的单独珠粒生成数据的单标记标识体链序列的第四珠粒。

[0036] 图22E表示携带相同的用于由经OCS实验的3个循环跟踪的FOV中的单独珠粒生成数据的单标记标识体链序列的第五珠粒。

[0037] 图22F表示携带相同的用于由经OCS实验的3个循环跟踪的FOV中的单独珠粒生成数据的单标记标识体链序列的第六珠粒。

[0038] 图23A用图形表示了跨OCS实验用序列

[0039] Id1/HindIII-AF647--Id2/SpeI-AF750--Id3/XhoI-ATTO488--Id4/NotI-ATTO488标记的珠粒的珠粒文库。

[0040] 图23B用图形表示了跨OCS实验用序列

[0041] Id1/HindIII-ATTO488--Id2/SpeI-ATTO550--Id3/XhoI-AF647--Id4/NotI-ATTO488标记的珠粒的珠粒文库。

[0042] 图23C用图形表示了跨OCS实验用序列

[0043] Id1/HindIII-ATTO550--Id2/SpeI-AF750--Id3/XhoI-ATTO488--Id4/NotI-AF750标记的珠粒的珠粒文库。

[0044] 图24A表示含有两个AF-750标记的AF-750标记的发夹寡核苷酸的SpeI酶切割。

[0045] 图24B呈现了SpeI酶切割之前和之后来自AF-647和AF-750标记两者的信号的图。

[0046] 图25A表示ATTO550标记从链霉亲和素珠粒开始进行的连接以及通过荧光成像得到的证实。

[0047] 图25B表示含有通过连接正交地附接到珠粒上的三个受体寡核苷酸中的第二受体寡核苷酸的AF-647标记的第二标记的DNA发夹以及通过荧光成像得到的证实。

[0048] 图25C表示含有通过连接正交地附接到珠粒上的三个受体寡核苷酸中的最后一个受体寡核苷酸的ATTO-488标记的第三标记的DNA发夹以及通过荧光成像得到的证实。

[0049] 图26A表示在SpeI酶切割之前和之后具有三种不同环标识体Id1/SpeI-ATTO-550、Id2/XhoI-AF-647和Id3/NotI-ATTO-488的珠粒。

[0050] 图26B呈现了条形图,所述条形图展示了ATTO-550标记的高效SpeI切割。

[0051] 图27A表示含有第一标记的ssDNA的链霉亲和素珠粒的第一编码循环(ATTO-488标记),所述第一标记的ssDNA与第一未标记的ssDNA杂交,所述第一未标记的ssDNA通过生物素部分附接到珠粒。

[0052] 图27B表示链霉亲和素珠粒的第二编码循环(AF-647标记)。

[0053] 图28A表示利用在含有三个不同的杂交的标识体Id1/NotI-ATTO-488、Id2/SpeI-AF-647和Id3/XhoI-ATTO-550的珠粒上的第一酶(NotI)进行的第一切割。

[0054] 图28B表示利用在含有三种不同的杂交的标识体的珠粒上的第二酶(SpeI)进行的第二切割。

[0055] 图28C提供了高效NotI和SpeI酶切割的荧光成像证实的图。

具体实施方式

[0056] 在第一实施例中,本文所述的分子条形码包括由一或多个标识体构成的集合,所述集合中的每个标识体都包含一或多个可检测标记,所述一或多个可检测标记操作性地连接到支架部分。所述支架部分包括:识别部分和切割位点,所述识别部分与所述由标识体构成的集合中的至少一个标识体的所述识别部分正交,所述切割位点操作性地连接到所述一或多个可检测标记,所述一或多个可检测标记以三维布置的方式彼此级联以编码和形成所述分子条形码。

[0057] 可能存在其中分子条形码包含不可切割标识体(图2)的情况。在一些实施例中,不可切割标识体符号可以被包含在分子条形码中,以便于在可切割标识体符号的所述可检测标记与所述分子条形码解离后检测其信号响应。在一些实施例中,不可切割标识体符号可以作为首先编码的标识体符号而操作性地连接到固相材料。在一些实施例中,不可切割标识体包括缺乏化学反应性识别部分或切割位点的支架部分。

[0058] 短语“与...操作性地偶联”和“与...偶联”是指两个或更多个实体之间的任何形式的相互作用,包含静电、酶、共价、离子或其它化学相互作用。两个实体可以彼此相互作用,即使其彼此不直接接触。例如,两个实体可以通过中间实体而彼此相互作用。

[0059] 如本文所用,蛋白酶切割位点序列和蛋白酶识别序列是正交反应性蛋白酶在此结合的肽序列。如本文所用,蛋白酶切割位点序列进一步包含其肽序列内的切割位点。

[0060] 将正交反应性切割剂应用于所述分子条形码在标识体的具有与所述切割剂具有

反应性的识别部分的所述切割位点处切割所述分子条形码,以根据所述标识体的正交性使所述标识体的所述可检测标记从所述三维布置中解离,并且由此诱导可检测信号响应以解码所述分子条形码。

[0061] 所述分子条形码的所述支架部分可以包括多肽、氨基酸、环肽、核酸或其组合。所述支架部分可以包括聚乙二醇(PEG)n。“n”可以等于1-12,更优选地等于1-6,最优选地等于2-5。标记的环肽的实例示出于图15和图16中。例如,可检测标记可以为荧光团或化学保护的量子点。使可检测标记远离彼此间隔开可以防止淬灭和/或光谱移位。切割位点可以为用于分别通过小分子、蛋白酶、核酸内切酶或RNA酶H切割的化学接头、肽、ssDNA、dsDNA或RNA。

[0062] 在一些实施例中,支架部分包括N末端赖氨酸残基,所述N末端赖氨酸残基可用化学基团(例如,反式环辛烯(TCO),使用NHS-PEG4-TCO)修饰,所述化学基团适于与含有C末端甲基四嗪基团的其它标识体的支架部分级联。标识体的C末端修饰可以使用可商购获得的含有马来酰亚胺反应性部分的异双功能交联剂以附接到所安装的半胱氨酸巯基(甲基四嗪-PEG4-马来酰亚胺)来实现。在一些实施例中,支架部分可以通过其C末端半胱氨酸残基而被修饰成含有DBCO基团(使用DBCO-PEG4-马来酰亚胺),而携带N末端裂解物的标识体的支架部分可以通过N末端赖氨酸残基而被修饰成含有叠氮化物基团。在一些实施例中,末端标识体可以被包含在分子条形码中以充当“封盖”标识体,并且因此被修饰成在其C末端处不含有化学反应性点击部分。

[0063] 在一些实施例中,标识体支架部分包括通过其N末端和C末端与含有化学反应性基团的接头试剂共价连接的肽片段,所述化学反应性基团实现了使用分裂-池化方法对标识体符号的正交化学级联。在一些实施例中,标识体的识别部分为蛋白酶切割位点序列。在一些实施例中,标识体的识别部分为蛋白酶识别序列。在一些实施例中,标识体支架部分包括被修饰成含有N末端赖氨酸残基、C末端半胱氨酸残基和携带叠氮化物基团的内部非天然氨基酸的多肽片段,并且可以从商业供应商(例如,位于加利福尼亚州费利蒙市的34801Campus Dr.(34801Campus Dr.Fremont,CA 94555)的阿纳斯派克股份有限公司(Anaspec Inc.))订购。

[0064] 在一些实施例中,支架部分包括用于通过天然化学连接促进携带N末端半胱氨酸的标识体的级联的N末端半胱氨酸残基。在一些实施例中,支架部分包括用于促进携带酶识别标签的标识体的级联的酶识别标签。

[0065] 在一些实施例中,点击化学可以用于将标识体级联,因为点击反应是正交的,可以具有相比于天然化学连接期望的反应动力学,可能需要较少的标识体来形成分子条形码,并且由此降低成本。

[0066] 在一些实施例中,支架部分包括用于促进标识体通过化学连接的级联的环肽、间隔接头或其它间隔分子。在此,环肽、间隔接头或间隔分子起到将标识体单体单元彼此间隔开环肽、间隔接头或间隔分子的分子距离。

[0067] 在一优选实施例中,支架部分操作性地连接到由一或多个可检测标记构成的集合。在包括操作性地连接到两个或更多个可检测标记(即,双标记)的支架部分的实施例中,支架部分可以被配置成将标识体的可检测标记彼此间隔开。例如,将可检测标记相对于彼此间隔开可以增强标识体的组合标记性能。

[0068] 如图16所示,支架部分包括环肽,并且可以相容地附接(或插入)在两个标识体之

间。在图16所示的配置中,每个标识体的朝向将是不相关的(标识体可以通过其N末端或C末端的附接进行级联),因为可检测标记可以在标识体之间附接到支架部分,并且蛋白水解位点的朝向将不影响蛋白酶活性。

[0069] 在一些实施例中,支架部分包括具有5'端和3'端的单链DNA(ssDNA)分子,所述ssDNA分子用可检测标记通过经修饰的内部碱基进行荧光标记,并且与互补的ssDNA预杂交以形成dsDNA双链体。所形成的dsDNA双链体可以编码单个核酸内切酶限制性序列(或位点类型)的一或多个拷贝。本领域的普通技术人员,借助于本公开将理解,可检测标记从所形成且经编码的分子条形码中的高效解离可以通过将核酸内切酶限制性序列沿着支架部分选择性地间隔开来促进。在一些实施例中,含有核酸酶识别序列的dsDNA双链体的一条链包括用于附接预先指定的荧光团的内部氨基修饰。

[0070] 在一优选实施例中,可检测标记因为其亮度(即,量子产率)和多路复用能力而包括一或多个Q点。在一些实施例中,可检测标记包括具有不同发射波长的荧光团,所述荧光团可以共价附接到标识体的支架部分。在一些实施例中,所形成且经编码的分子条形码的可检测信号响应的组合可以包括或可以不包括连续发射光谱,因为形成分子条形码的标识体符号的三维布置可以变化。

[0071] 在一些实施例中,如果Q点是通过分裂-池化并行使用的,则可以形成并编码分子条形码以包含约一百万(10^6)个标识体类别变异。用组合Q点标记对每个标识体进行的双标记可以将这一数字扩展至约十亿(10^{10})个标识体类别变异。在一些实施例中,荧光团掺杂的纳米颗粒的使用可以提供进一步的标识体类别变异。

[0072] 由标识体构成的集合中的至少一个标识体的识别部分可以包括化学接头、肽或核酸。切割位点可以位于识别部分内。具体地,由标识体构成的集合中的每个识别部分都可以包括至少一个选自以下的部分:蛋白酶识别部分、核酸内切酶识别部分、能通过亲和试剂识别的表位、核酸探针识别部分、经修饰的肽侧链和非天然肽侧链。

[0073] 具有两个可检测标记和相同的识别位点中的两个识别位点的示例性标识体示出于图14中。

[0074] 本文公开的分子条形码可以通过将一个标识体与另一个标识体通过特异性相容化学附接来构建,所述特异性相容化学如点击化学部分、天然化学连接接头或酶介导的连接接头,例如通过T4 DNA连接酶的DNA连接。分子条形码可进一步包括位于由标识体构成的集合中的每个标识体之间的标识体接头。标识体接头可以选自以下群组:点击化学部分、天然化学连接接头和酶介导的连接接头。

[0075] 在一些实施例中,标识体可以按预先限定的顺序依序添加,以经连续轮次的分裂和池化来形成条形码。每个轮次的标识体添加将使标识体能够通过正交化学连接进行级联。

[0076] 在一些实施例中,标识体通过连接进行的级联可以在相同的反应缓冲液(1X磷酸盐缓冲盐水(PBS)或1X T4 DNA连接酶缓冲液)中使用要级联的摩尔过量的传入标识体来进行。

[0077] 在一些实施例中,安装在支架部分内的N末端赖氨酸残基可以用化学基团(例如,反式环辛烯(TCO),使用NHS-PEG4-TCO)修饰,所述化学基团适于与含有C末端甲基四嗪基团的其它标识体附接。标识体的C末端修饰可以使用可商购获得的含有马来酰亚胺反应性部

分的异双功能交联剂以附接到所安装的半胱氨酸巯基(甲基四嗪-PEG4-马来酰亚胺)来实现。在一些实施例中,一些标识体可以通过其C末端半胱氨酸残基而被修饰成含有DBCO基团(使用DBCO-PEG4-马来酰亚胺),而一些标识体可以通过其N末端赖氨酸残基而被修饰成含有叠氮化物基团。在一些实施例中,末端标识体被包含在所形成的分子条形码中并且充当"封盖"标识体,并且因此被修饰成在其C末端处不含有化学反应性点击部分。

[0078] 与下一代测序条形码连接的分子条形码可以具有包括双链DNA的支架部分和包括蛋白质、DNA、RNA或可化学切割部分的识别部分。双链DNA可以具有供捕获的3'重叠,如Poly(T),即唯一的分子缀合,和/或5'PCR手柄(图9和图10)。

[0079] 识别部分可以包括构成条形码的每个标识体类别的单个核酸内切酶切割位点类型(图5)。这些分子条形码的切割剂可以是蛋白酶(图8、图10)、核酸酶(图7、图9)或化学切割剂。本领域已知许多核酸内切酶相对于彼此具有正交反应性。例如,NotI、XhoI、SpeI和HindIII对分别含有GCGGCCGC、CTCGAG、ACTAGT和AAGCTT的dsDNA序列具有特异性。因此,在一些实施例中,NotI不应高效地特异性切割XhoI、SpeI或HindIII所识别的位点。对于所列出的核酸酶中的每种核酸酶也是如此,因为在存在由其它酶所识别的底物的情况,每种核酸酶都应能够仅切割(主要地)其相应底物。

[0080] 分子条形码可以具有包括双链DNA的支架部分、包括核酸酶识别部分的识别部分以及包括荧光团的一或多个可检测标记。可替代地,分子条形码可以具有包括双链DNA的支架部分、包括蛋白酶识别部分的识别部分以及包括荧光团的一或多个可检测标记。进一步可替代地,第一实施例的分子条形码可以具有包括双链DNA的支架部分、包括化学切割识别部分的识别部分以及包括荧光团的一或多个可检测标记。

[0081] 非DNA分子条形码可以具有支架部分,包括氨基酸、肽或蛋白质。识别部分可以由化学可切割部分制成,并且可以包含多于一个可检测标记,例如,两个或更多个不同标记的组合。这些条形码的切割剂可以为蛋白酶或化学切割剂,所述蛋白酶或化学切割剂可以依序添加。可替代地,分子条形码可以具有包括氨基酸的支架部分、包括化学切割识别部分的识别部分以及包括荧光团的一或多个可检测标记。进一步可替代地,分子条形码可以具有包括氨基酸的支架部分、包括蛋白酶识别部分的识别部分以及包括荧光团的一或多个可检测标记。

[0082] 识别部分可以包括化学接头、肽或核酸。含有上面的识别基序的肽的标识体属于不同的类别;一种类别是由其所识别的蛋白酶来定义的。

[0083] 可检测标记中的一或多个可检测标记可以包括荧光团。

[0084] 本公开的益处是,分子条形码可以与和分子条形码的三维布置具有互信息的材料缔合。因此,所述材料可以通过解码分子条形码来进行分类。例如,可以将由一或多个分子条形码构成的集合引入到环境(例如,化学或生化系统)中,其中所述集合中的至少一个分子条形码与也存在于所述环境中的材料具有互信息。用于由分子条形码构成的集合的正交反应性切割剂可以被选择性地应用于在对切割剂具有反应性的切割位点处切割分子条形码。所述切割根据标识体的正交性使标识体的可检测标记从三维布置中解离。可检测标记的解离诱导可检测信号响应。所述信号响应可以用于解码由一或多个分子条形码构成的集合(即,确定所述特定标识体的身份和顺序)。解码由一或多个分子条形码构成的集合揭示了与材料共有的互信息(例如,材料的身份和/或浓度)。这种材料也可以与和用于编码其身

份的分子条形码相同的珠粒连接。

[0085] 如本文所用,“以编码”是指将信息、数据或分类指令转化为经转换的格式。

[0086] 如本文所用,“以解码”是指反转编码过程以从经转化的格式中提取信息。

[0087] 如本文所用,“互信息”是指可通过观察第二变量获得的关于第一变量的量信息。

[0088] 如本文所用,“正交”是指多组分系统中的在特定反应条件集合下与特定试剂具有化学反应性的组分,而多组分系统中的至少一种其它组分与所述试剂具有有限的反应性或无反应性,即使多组分系统中的所有组分都存在于相同的环境中也是如此。

[0089] 如本文所用,“正交”是指多组分系统中的在特定反应条件集合下与特定试剂具有化学反应性的组分,而所述系统中的至少一或多种组分与所述试剂不具有反应性,即使所述系统中的所有组分都存在于相同的环境中也是如此。同样,“正交反应性”是指具有正交反应性的材料。

[0090] 如本文所用,“标识体”是给予单个“标识分子”的名称。如本文所公开的,单个标识体分子包括由一或多个可检测标记构成的集合,所述一或多个可检测标记在其并入并编码分子条形码之前操作性地连接到支架部分,所述支架部分包括识别部分和切割位点。

[0091] 如本文所用,“标识体类别”是给予由具有本质上相同的配置的一或多个标识体构成的集合的名称。因此,一种类别中的标识体应基本上相同地对刺激作出反应。

[0092] 如本文所用,“标识体符号”是已并入到分子条形码中或已与材料物理地缔合的标识体类别的有形实例。

[0093] 如本文所用,“识别部分(recognition moiety)”和“识别部分(recognition portion)”可互换使用,并且是指对化学或酶切割剂具有反应性的化学部分。

[0094] 如本文所用,“切割位点”是两个解离的分子之间的切割点。在一些实施例中,切割位点位于识别部分内。在一些实施例中,切割位点位于识别部分之外或远离识别部分。

[0095] 如本文所用,“符号”是起到如事实、质量、数据等信息或其它形式的信息的视觉或有形表示的作用的物体。

[0096] 例如,材料可以包括与分子条形码(统一被称为“测试构建体”)操作性地偶联的测试剂,其中分子条形码的三维布置与测试剂具有互信息。在本实例中,解码条形码表明测试剂是现在还是过去存在于具有分子条形码的环境中。因此,利用条形码的测试构建体可以用于筛选。

[0097] 例如,筛选方法可以包含将多个测试构建体引入到环境中,其中每个测试构建体都包括与唯一的分子条形码操作性地偶联的测试剂,其中分子条形码的三维布置与测试剂具有互信息。针对由一或多个靶标构成的集合筛选多个测试构建体。确定测试构建体中的一或多个测试构建体的活性。将正交反应性切割剂选择性地应用于多个测试构建体,以在对切割剂具有反应性的切割位点处切割偶联的分子条形码(即,主要在仅具有对切割剂具有反应性的识别部分的标识体的切割位点处切割分子条形码)。这仅根据标识体的正交性而使标识体的可检测标记从三维布置中解离。因此,可以使用可检测信号响应以解码条形码(即,在本实例中,标识条形码标识体序列),并且由此标识针对由靶标构成的集合具有活性的对应测试剂。

[0098] 对一些分子条形码中的可检测标记的识别和切割可能需要按从最外面的段朝珠粒的次序进行(例如,图7),否则条形码信息可能丢失。另外,对一些分子条形码中的可检测

标记的识别和切割可能需要按已知次序进行,以保留条形码信息(例如,图8、图11)。识别部分可以由蛋白质或可化学切割部分构成。任选地,识别部分含有多于一个可检测标记,例如,两个或更多个不同标记的组合。

[0099] 有利的是,可以在不进行富集的情况下诱导可检测信号响应。可检测信号响应可以是信号强度,如荧光强度的提高或降低。

[0100] 将正交反应性切割剂选择性地应用于多个测试构建体可以涉及依序地应用切割试剂(例如,迭代地应用单一切割剂或单独地或依序地应用两种或更多种切割剂)。切割剂可以是不同的蛋白酶和/或不同的化学切割剂。

[0101] 用于产生分子条形码的方法包含提供由一或多个标识体构成的集合,所述由标识体构成的集合中的每个标识体都包括由一或多个可检测标记构成的集合,所述一或多个可检测标记操作性地连接到支架部分,所述支架部分包括识别部分和切割位点,所述识别部分与所述由标识体构成的集合中的至少一个标识体的所述识别部分正交,所述切割位点操作性地连接到由一或多个可检测标记构成的集合。标识体可以以组合方式进行标记。例如,10个不同的量子点可以针对7个标识体条形码中的每个标识体用相同的蛋白酶位点编码,这可能导致编码1000万个不同的条形码或珠粒。所述方法包含从所述由标识体构成的集合中选择第一标识体和第二标识体,并且将所述第一标识体与所述第二标识体以三维布置的方式级联,以编码和形成分子条形码。所述方法可以继续从所述由标识体构成的集合中选择一个标识体,并且将所述标识体与所述分子条形码级联,以修改所述三维布置并且进一步编码和形成所述分子条形码。此最后的步骤可以重复一次至n次。

[0102] 标识体链可以通过顺序分裂-池化方法构建,并且在每个轮次中添加不同类别。在一些实施例中,可以并行使用各种标识体类别以通过分裂-池化连接来编码和形成分子条形码,以在不依赖于基于DNA测序的解码的情况下创建分子条形码的组合文库。在一些实施例中,可以在存在同时形成和编码的其它分子条形码的情况下产生单独分子条形码。此类其它分子条形码可以包含附接到珠粒上的位点而非标识体条形码附接位点上的下一代测序条形码。

[0103] 在一些实施例中,标识体可以用于通过分裂-池化珠粒来构建基于标识体的分子条形码的组合链。本质上,每个轮次的添加将唯一的标识体符号添加到每个组合链中。本质上,每个所产生的分子条形码将含有根据并入到分子条形码中的标识体符号的三维布置组织的可检测标记的组合。

[0104] 标识体可以通过天然化学连接、点击化学连接、酶促地通过肽连接酶、分选酶或SFP合成酶附接。在每次切割循环期间,仅标识条形码内的标识体的一种类别。

[0105] 第一标识体可以与材料,如材料的表面,如珠粒的表面操作性地偶联。第一标识体可以通过使用普通化学而附接到珠粒。第一标识体可以为寡聚体,并且可以通过其5'端附接到珠粒(图6)。珠粒可以与测试剂操作性地偶联。珠粒可以为封盖珠粒(图17)。第一标识体可以通过高亲和力结合蛋白或三臂接头与材料表面操作性地偶联。

[0106] 提供由一或多个标识体构成的集合可以包含将所述由标识体构成的集合中的每个标识体配置成与在前的标识体级联并且促进标识体以线性三维布置的方式级联,并且由此依序地形成和编码所述分子条形码。将所述由标识体构成的集合配置成与在前的标识体特异性地结合可以包含通过连接到所述在前的标识体、从所述在前的标识体延伸或合成到

所述在前的标识体上进行级联。例如,包括双链DNA的标识体可以具有用于连接的悬垂端。

[0107] 所述方法可以包含提供由一或多个化学构建块构成的集合,所述由化学构建块构成的集合中的每个化学构建块被配置成与在前的化学构建块级联;从所述由化学构建块构成的集合中选择第一化学构建块,并且将所述第一化学构建块与所述第一标识体级联。换言之,在化学构建块附接之后,构建块被视觉条形码段编码(图12和图13)。所述方法可以继续从所述由被配置成与所述在前的化学构建块级联的化学构建块构成的集合中选择一个化学构建块,并且将所述化学构建块与所述在前的化学构建块级联,并由此形成一系列化学构建块。此最后的步骤可以重复一次至n次。

[0108] 有益的是,将所述系列的化学构建块级联可以包含使用非核酸相容性反应。因此,反应可以用于构建主要基于核酸的条形码通常无法获得的分子条形码。

[0109] 化学文库构建后,每个珠粒可以展示唯一的化合物,所述唯一的化合物的身份被编码在组合视觉条形码中。然后可以在使用各种已知的选择方案时选择化合物的文库。在选择之后,可能不需要在非生产性化合物之上富集所选化合物。这种条形码类型可以实现在不使用DNA的情况下构建和编码高度多样化的化学文库。文库成员可以如前所述使用例如条形码段的正交切割以大规模并行方式进行区分。

[0110] 在一些实施例中,分子条形码文库可以形成于珠粒上,并且珠粒可以固定在表面上并且在与实验溶液接触之前和之后成像。首先可以使用配置有适当激发波长和发射滤光片的标准荧光显微镜对所固定的珠粒进行成像,以记录给定视场或所关注区中构成每个条形码的视觉可检测标记的组合。

[0111] 代替标准荧光显微镜,倒置荧光显微镜、如内置的具有适当配置的荧光扫描仪的任何放大设备、可以对荧光珠粒进行成像的设备都可以工作。自动成像仪可用于各种应用。例如,可以使用分子动力学公司扫描仪(Molecular Dynamics scanner)或耐细隆公司扫描仪(Nexcelom scanner),用于使用标准4荧光团的细胞阵列应用。为了检测Q点,需要特殊的配置,例如能够区分所有11个可能的Q点的滤光片转轮。

[0112] 在特定实施例中,在一定数量的荧光团附接后,在所有构建的标识体上可能需要一种颜色恒定的颜色。例如,如果存在11个可区分的Q点,则可以使用10个Q点来构建条形码,并且第11个Q点可以用作要作为所有标识体上的常量使用的标准颜色。这可能是有益的,因为此荧光团可以作为定量标准,在每个图像中任何给定链中的其它荧光团都可以针对所述定量标准进行比较。这可以使用户看到其它信号如何相对于所述恒定信号进行比较,并且更准确地测量信号的分数。本领域的普通技术人员,借助于本公开将理解,还可以使用其它类型的内部控制。

[0113] 本文还公开了一种用于编码和解码分子条形码的系统。此类系统包含条形码编码系统,所述条形码编码系统被配置成向环境中引入由一或多个经编码的分子条形码构成的集合,其中所述由经编码的分子条形码构成的集合中的至少一个分子条形码与材料具有互信息。所述系统进一步包含切割系统,所述切割系统被配置成将正交反应性切割剂选择性地应用于环境以在标识体的切割位点处切割由分子条形码构成的集合,在所述由分子条形码构成的集合中,存在与切割剂具有反应性的识别部分,以根据标识体的正交性使可检测标记从分子条形码集合中的分子条形码的三维布置中解离,由此从一个中诱导可检测信号响应以解码由分子条形码构成的集合。所述系统进一步包含检测系统,所述检测系统用于

从由分子条形码构成的集合中检测可检测信号响应,并且可以进一步包含照明系统,所述照明系统被配置成向由分子条形码构成的集合选择性地传递光以诱导可检测信号响应。

[0114] 流动池

[0115] 使用预先制造的塑料盖与从微流控芯片工作室公司(Microfluidic Chip Shop)购买的粘合剂构建流动池—所述流动池被切割,以使得能够在安装在载玻片上时产生16个单独泳道。流动池泳道各自的体积为大约15-20μl。使用聚-L-赖氨酸涂覆的玻璃载玻片以构建流动池;以这种方式,可以使用赖氨酸侧链上的游离伯胺以用NHS-LC-生物素进行修饰(在500μM浓度下在NHS缀合缓冲液,在室温下,持续至少1小时)。生物素化后,用200μl 2X杂交缓冲液冲洗流动池泳道,以使其准备好用于引入寡核苷酸涂覆的链霉亲和素珠粒,以用于固定和下游解码实验。

[0116] 分子条形码去卷积

[0117] 将20μl用分子条形码涂覆的0.2-1mg/ml链霉亲和素珠粒在1X杂交缓冲液中引入到流动池泳道中。允许珠粒沉降至少1小时以促进表面附接。然后将流动池泳道用200μl1X杂交缓冲液冲洗,以去除未结合的珠粒,然后用100μl 1X Cutsmart缓冲液冲洗。所有标识体解码实验都是使用含有50-500U单个限制性酶的切割溶液进行的。在一些实施例中,使用约5U/ul浓度的单一限制性酶。这些溶液是如下基本上去除甘油的:通过首先将酶储备液2倍稀释于1X Cutsmart缓冲液(NEB)中,并且通过7K MWCO Zeba柱(可从赛默飞世尔公司(ThermoFisher,Inc.)商购获得)缓冲液交换到1X Cutsmart缓冲液中。在引入第一切割剂之前,获取未切割的1μm用各种分子条形码涂覆的MyOne T1珠粒的图像。在每个解码循环期间,在热块上,将切割溶液引入到流动池的泳道中,并且在37℃下温育15-30分钟。在每个解码循环之后,获取珠粒的图像,并且将其编译用于下游分析。

[0118] 收集链霉亲和素珠粒图像

[0119] 在具有HCX PL FLUOTAR L 40x(NA-0.6)CORR PH2物镜、具有卢门科公司(Lumencor)光谱X光引擎(395、440、470、550、640、748)和徕卡公司(Leica)DFC9000 GTC相机的徕卡公司Thunder系统上对链霉亲和素珠粒进行成像。对于图像获取,利用以下滤光片设置(Quad立方体-激发:375-407、462-496、542-566、622-654,DC:415、500、572、660,发射:420-450、506-532、578-610、666-724;以及Y7立方体,激发:672-748,DC:760,发射:765-855)。另外的DFT5快速滤光片转轮位于立方体的下游,并且包含以下LP滤光片:440%、510%、590%、700%和100%。

[0120] 对与链霉亲和素珠粒缀合的分子条形码的循环性解码的图像分析

[0121] 将图像加载到Volocity¬Æ数据库中,并且按以下方式进行分析。首先,将单独图像按其获取的颠倒顺序制成时间系列。然后对图像进行运动校正,并且裁切到从场的中间开始的区域,以使不均匀照明最小化。对于每个图像,在10个珠粒和一个背景区域周围绘制相同大小的ROI,用于计算背景减法。从循环中的下一个图像中减去针对每种标记获得的值,以清楚地标识每次切割循环期间释放的荧光团标记。例如,在对三个图像(未切割的、通过RE1切割的和通过RE2切割的)进行的OCS实验中,从在倒数第二个图像(通过RE1切割之后)中获取的数据中减去在最后一个图像(通过RE2切割之后)中获取的数据。从在第一个图像(未切割)中获取的数据中减去在倒数第二个图像(通过RE1切割之后)中获取的数据。这使得能够准确地鉴定每次切割循环期间的信号损失。

[0122] 链霉亲和素珠粒示出在图19A中,其含有通过与未标记的dsDNA连接进行附接的第一标记的dsDNA标识体段(Id1/HindIII-AF750),所述未标记的dsDNA通过生物素部分附接到珠粒。dsDNA标识体段,也显示为位于珠粒右侧,含有AF750标记(AF-750)(向下斜条纹),所述AF-750使用NHS-AF750通过经内部氨基修饰的碱基附接到DNA双链体的寡核苷酸之一。附接的标记的标识体还含有酶可接近的限制性核酸内切酶位点(HindIII),所述HindIII定位在附接的标记与珠粒之间。Id1/HindIII-AF750寡核苷酸双链体通过与未标记的寡核苷酸(Id0)的连接反应附接到珠粒,所述未标记的寡核苷酸不含有可识别的限制性核酸内切酶位点(不可切割),但含有用于附接到链霉亲和素珠粒的5'生物素修饰。在溶液中进行dsDNA标识体的连接,洗涤珠粒,并且将珠粒的样品固定在流动池中的经生物素修饰的表面上,以进行成像。通过四个荧光发射通道488nm、550nm、647nm和750nm中的荧光成像证实Id1/HindIII-AF750与珠粒的附接,其中仅在750nm波长中观察到强发射信号。证实Id1/HindIII-AF750 dsDNA标识体与Id0寡核苷酸双链体的高效连接的图像以条形图形式呈现,所述条形图是使用从以下成像的所有4个发射波长以及许多携带相同寡核苷酸的珠粒(平均强度值)获取的数据绘制的:488nm(横条纹)、550nm(向上斜条纹)、647nm(虚线)和750nm(向下斜条纹)。图中的条的高度对应于在成像的所有四个通道中观察到的发射的相对量值,所述相对量值以荧光强度单位(FU)表示。

[0123] 链霉亲和素珠粒示出在图19B中,含有第二标记的dsDNA标识体段(Id2/SpeI-AF647)。第二dsDNA标识体段,也显示为位于珠粒右侧,含有AF647标记(AF-647)(虚线),所述AF-647使用NHS-AF647通过经内部氨基修饰的碱基附接到DNA双链体的寡核苷酸之一。链中的第二附接的标记的标识体还含有酶可接近的限制性核酸内切酶位点(SpeI),所述SpeI定位在附接的标记与珠粒之间。Id2/SpeI-AF647寡核苷酸双链体通过与第一标记的dsDNA标识体段(Id1/HindIII-AF750)双链体的连接反应附接到珠粒。在溶液中进行dsDNA标识体的连接,洗涤珠粒,并且将珠粒的样品固定在流动池中的经生物素修饰的表面上,以进行成像。珠粒上Id1/HindIII-AF750--Id2/SpeI-AF647连接产物的成功形成通过四个荧光发射通道488nm、550nm、647nm和750nm中的荧光成像得到证实,其中现在750nm和647nm发射波长两者观察到强信号,但550nm或488nm波长未观察到强信号。证实Id2/SpeI-AF647 dsDNA标识体与Id1/HindIII-AF750dsDNA标识体的高效连接的图像以条形图形式呈现,所述条形图是使用从以下成像的所有4个发射波长以及许多携带相同寡核苷酸的珠粒(平均强度值)获取的数据绘制的:488nm(横条纹)、550nm(向上斜条纹)、647nm(虚线)和750nm(向下斜条纹)。图中的条的高度对应于在成像的所有四个通道中观察到的发射的相对量值,所述相对量值以荧光强度单位(FU)表示。

[0124] 链霉亲和素珠粒示出在图19C中,含有第三标记的dsDNA标识体段(Id3/XhoI-ATTO550)。第三dsDNA标识体段,也显示为位于珠粒右侧,含有ATTO550标记(ATTO-550)(向上斜条纹),所述ATTO-550使用NHS-ATTO550通过经内部氨基修饰的碱基附接到DNA双链体的寡核苷酸之一。链中的第三附接的标记的标识体还含有酶可接近的限制性核酸内切酶位点(XhoI),所述XhoI定位在附接的标记与珠粒之间。Id3/XhoI-ATTO550寡核苷酸双链体通过与第二标记的dsDNA标识体段(Id2/XhoI-AF647)双链体的连接反应附接到珠粒。在溶液中进行dsDNA标识体的连接,洗涤珠粒,并且将珠粒的样品固定在流动池中的经生物素修饰的表面上,以进行成像。珠粒上Id1/HindIII-AF750--Id2/SpeI-AF647--Id3/XhoI-ATTO550

连接产物的成功形成通过四个荧光发射通道488nm、550nm、647nm和750nm中的荧光成像得到证实,其中现在750nm、647nm和550nm发射波长观察到强信号,但488nm波长未观察到强信号。证实Id3/XhoI-ATTO550 dsDNA标识体与生长的标识体链的高效连接的图像以条形图形式呈现,所述条形图是使用从以下成像的所有4个发射波长以及许多携带相同寡核苷酸的珠粒(平均强度值)获取的数据绘制的:488nm(横条纹)、550nm(向上斜条纹)、647nm(虚线)和750nm(向下斜条纹)。图中的条的高度对应于在成像的所有四个通道中观察到的发射的相对量值,所述相对量值以荧光强度单位(FU)表示。

[0125] 链霉亲和素珠粒示出在图19D中,含有第四标记的dsDNA标识体段(Id4/NotI-ATTO488)。第四dsDNA标识体段,也显示为位于珠粒右侧,含有ATTO488标记(ATTO-488)(横条纹),所述ATTO-488使用NHS-ATTO488通过经内部氨基修饰的碱基附接到DNA双链体的寡核苷酸之一。链中的第四附接的标记的标识体还含有酶可接近的限制性核酸内切酶位点(NotI),所述NotI定位在附接的标记与珠粒之间。Id4/NotI-ATTO488寡核苷酸双链体通过与第三标记的dsDNA标识体段(Id3/XhoI-AF647)双链体的连接反应附接到珠粒。在溶液中进行dsDNA标识体的连接,洗涤珠粒,并且将珠粒的样品固定在流动池中的经生物素修饰的表面上,以进行成像。珠粒上Id1/HindIII-AF750--Id2/SpeI-AF647--Id3/XhoI-ATTO550--Id4/NotI-ATTO488连接产物的成功形成通过四个荧光发射通道488nm、550nm、647nm和750nm中的荧光成像得到证实,其中现在750nm、647nm、550nm和488nm发射波长观察到强信号。证实Id4/NotI-ATTO488 dsDNA标识体与生长的标识体链的高效连接的图像以条形图形式呈现,所述条形图是使用从以下成像的所有4个发射波长以及许多携带相同寡核苷酸的珠粒(平均强度值)获取的数据绘制的:488nm(横条纹)、550nm(向上斜条纹)、647nm(虚线)和750nm(向下斜条纹)。图中的条的高度对应于在成像的所有四个通道中观察到的发射的相对量值,所述相对量值以荧光强度单位(FU)表示。

[0126] 解码标记的dsDNA标识体链;3次切割循环:

[0127] 用四种不同荧光团标记的单一组合对珠粒进行编码,其中每个唯一的标记附接到dsDNA标识体链内的不同标识体段。通过先前所述的四个轮次的连接形成dsDNA标识体链,如图19所示。对于本实验,链序列为:

[0128] Id1/HindIII-ATTO550--Id2/SpeI-AF750--Id3/XhoI-AF647--Id4/NotI-ATTO488。因此,预期的是携带此类dsDNA标识体链序列的珠粒将在正交切割的第一循环中在仅暴露于NotI酶之后丢失488nm通道中的荧光信号;然后在正交切割的第二循环中在仅暴露于XhoI酶之后丢失647nm通道中的荧光信号;并且在正交切割的第三循环中在仅暴露于SpeI酶之后丢失750nm通道中的荧光信号。在此,将珠粒固定在流动池的表面上,并且使用这些酶按照前述循环顺序进行正交切割测序(OCS)的三个循环,其中在每个循环之前和之后获取图像。首先,在四个荧光通道中对未经切割的珠粒(图20A)进行成像:647nm(向上斜条纹)、550nm(向下斜条纹)、488nm(横条纹)和750nm(虚线),并且在右侧绘制了针对每种荧光团获得的平均强度值。在获取的未经切割的珠粒的图像中观察到所有四种荧光团的信号(右边的条形图)。然后将流动池置于37℃热块上,并且暴露于含有1XCutSmart缓冲液和5U/ul浓度的NotI酶的溶液中持续15分钟。然后用1X CutSmart缓冲液冲洗流动池泳道,并且将流动池返回到显微镜用于成像。从同一流动池泳道的同一位置获取暴露于NotI酶之后的珠粒的图像(图20B),所述流动池泳道被成像用于对未经切割的珠粒进行定量(尽管在每

次切割循环期间将流动池从显微镜中取出并置于热块上以进行温育,但将其返回到显微镜,并且在实验期间在获取的所有图像中都保持此FOV)。珠粒在暴露于NotI后在488nm通道中显示出显著较低的信号,如图20B中所示(右侧的平均强度图);NotI暴露之前,来自珠粒的488nm信号发射的平均强度为大约9,300FU,并且暴露于NotI后,此信号下降至大约2,000FU。重要的是,针对所有其它荧光标记观察到的信号强度保持基本上不变。为了启动正交切割的第二循环,将流动池返回到37℃热块上,并且暴露于含有1X CutSmart缓冲液和5U/ul浓度的XhoI酶的溶液中持续15分钟。然后用1X CutSmart缓冲液冲洗流动池泳道,并且将流动池返回到显微镜用于成像。暴露于XhoI酶之后的珠粒的图像(图20C)显示出来自AF647标记的信号发射显著减少。来自XhoI暴露之前的珠粒的647nm信号发射的平均强度为大约8,000FU,并且暴露于XhoI后,此信号下降至大约1,000FU。重要的是,针对两个未经切割的荧光标记观察到的信号强度保持基本上不变。为了启动正交切割的第三循环,将流动池返回到37℃热块上,并且暴露于含有1X CutSmart缓冲液和5U/ul浓度的SpeI酶的溶液中持续15分钟。然后用1X CutSmart缓冲液冲洗流动池泳道,并且将流动池返回到显微镜用于成像。暴露于SpeI酶之后的珠粒的图像显示出来自AF750标记的信号发射显著减少(图20D)。来自SpeI暴露之前的珠粒的750nm信号发射的平均强度为大约5,500FU,并且暴露于SpeI后,此信号下降至大约400FU。重要的是,针对最终的未经切割的荧光标记(ATTO550)观察到的信号强度保持基本上不变。链序列:

[0129] Id1/HindIII-ATTO550--Id2/SpeI-AF750--Id3/XhoI-AF647--Id4/NotI-ATTO488暴露于NotI,然后暴露于XhoI,之后暴露于SpeI产生了预期的解码结果。

[0130] 解码不同标记的dsDNA标识体链:单个段和混合段

[0131] 构建单标记链,由此将Id1/HindIII、Id2/SpeI、Id3/XhoI和Id4/NotI的段各自用先前列出的四种荧光团(ATTO488、ATTO550、AF647和AF750)差异地进行标记,以制成总共16个不同的标记的dsDNA标识体段。选择每种类型(Id1/HindIII、Id2/SpeI、Id3/XhoI和Id4/NotI)的一个差异标记的段以编码预先限定的链序列。在此使用的单标记dsDNA标识体链序列为:Id1/HindIII-ATTO550--Id2/SpeI-AF750--Id3/XhoI-AF647--Id4/NotI-ATTO488。通过先前描述的分裂-池化连接策略构建含有预先限定的单标记链序列的链霉亲和素珠粒,并且将其固定在流动池的一个泳道中的经生物素修饰的表面上以进行成像。在与含有限制性酶的溶液接触之前,在所有四个荧光通道中对单个FOV中的许多珠粒进行成像。然后,将珠粒进行如前所述的OCS的三个循环(循环1的NotI、循环2的XhoI和SpeI),之后在每个循环后进行成像。对于成像系列中的每次采集,对单个FOV中的许多珠粒的数据求平均,并且在OCS实验的所有循环中对同一FOV进行成像。虽然Id1可通过HindIII切割,并且用ATTO488标记,但在此不切割此Id段,因为来自单个荧光团标记的信号很容易可视化,不需要切割此段。将通过在每次切割循环之前和之后对所有四种标记进行成像获得的强度数据输入到表格中,其中循环是按时间顺序列出的。从原始数据中可以清楚地看出,哪种荧光团是在OCS的每个循环期间被切割的。然而,为了更清楚地限定附接到链中的每个Id段的标记的顺序,将数据绘制为从珠粒“读取”获得的序列。针对单标记dsDNA标识体链获得的OCS结果展示在图21A中示出的条形图中,其中循环的顺序颠倒以供分析。以此方式,在本实验中的正交切割(通过SpeI切割)的最后循环之后提取的图像成为系列中的“第一”图像,XhoI切割之后获取的图像成为所述系列中的“第二”图像,NotI切割之后获取的图像成为所述系列中的“第

三”图像,并且获取的未经切割的珠粒的图像成为所述系列中的“最后”图像。虽然是按时间颠倒的顺序,但从“前一个”循环中减去所有四种荧光团的强度数据。然后针对单标记dsDNA标识体链解码成像系列中的每个循环绘制所有四种荧光团的所得数据,如图21A所示。在此实例中,珠粒是在SpeI切割之后进行成像的,并且珠粒上的唯一剩余的荧光团ATTO550(向下斜条纹)观察到强信号,所述ATTO550附接到Id1。从在XhoI切割之后获取的图像中减去针对SpeI切割之后的图像中的所有四种荧光团观察到的强度数据,以得到在SpeI切割循环期间被切割的清晰的单一标记AF750(浅色虚线),所述AF750附接到Id2。同样,从在NotI切割之后获取的图像中减去XhoI切割后在珠粒上观察到的所有四个通道的信号,以便观察在XhoI切割时去除的单一标记AF647(深色虚线),所述AF647附接到Id3。最后,从获得自未经切割的珠粒的获取的图像中的荧光团强度中减去在NotI切割之后在图像中在珠粒上观察到的所有强度,以便观察NotI切割时去除的单一标记ATTO488(横条纹),所述ATTO488附接到Id4。

[0132] 为了增加使用dsDNA标识体链编码的文库的理论多样性,构建了经混合的片段链,由此在段连接之前,将两种常见类型的差异标记的dsDNA标识体以1:1摩尔比以10种不同的预先限定的可视觉辨别的组合的形式同等地混合(例如,Id2/SpeI-ATTO488与Id2/SpeI-AF647混合)。10种荧光团组合可以使用4种不同的标记变得在视觉上可区分:488/488、488/550、488/647、488/750、550/550、550-/647、550/750、647/647、647-/750和750/750。选择每种类型的差异标记的段的一个组合(Id1/HindIII、Id2/SpeI、Id3/XhoI和Id4/NotI),以编码预先限定的链序列,其中在链中的每个Id段位置处具有混合标记。在此使用的混合标记的dsDNA标识体链序列为:Id1/HindIII-ATTO550/AF750--Id2/SpeI-AF647/ATTO488--Id3/XhoI-ATTO550/AF647--Id4/NotI-AF750/ATTO488。通过先前描述的分裂-池化连接策略构建含有预先限定的混合标记链序列的链霉亲和素珠粒,并且将其固定在流动池的一个泳道中的经生物素修饰的表面上以进行成像。在与含有限制性酶的溶液接触之前,在所有四个荧光通道中对单个FOV中的许多珠粒进行成像。然后,将珠粒进行如前所述的OCS的三个循环(循环1的NotI、循环2的XhoI和SpeI),之后在每个循环后进行成像。对于成像系列中的每次采集,对单个FOV中的许多珠粒的数据求平均,并且跨OCS系列的所有循环对同一FOV进行成像。虽然Id1可通过HindIII切割,并且用作用AF750和ATTO488两者标记的1:1混合物,但在此不切割此Id段,因为来自仅两种剩余的荧光团标记的信号很容易可视化(在SpeI切割之后),不需要切割此段。将通过在每次切割循环之前和之后对所有四种标记进行成像获得的强度数据输入到表格中,其中循环是按时间顺序列出的。初始地从原始数据中不可以清楚地看出哪种荧光团是在OCS的每个循环期间被切割的。为了更清楚地限定附接到链中的每个Id段的标记的顺序,将数据绘制为从珠粒“读取”获得的序列,如利用单标记链数据进行的。针对混合标记dsDNA标识体链获得的OCS结果展示在图21B中示出的条形图中,其中循环的顺序颠倒,以用于如针对单标记链数据进行的分析。以此方式,在本实验中的正交切割(通过SpeI切割)的最后循环之后提取的图像成为系列中的“第一”图像,XhoI切割之后获取的图像成为所述系列中的“第二”图像,NotI切割之后获取的图像成为所述系列中的“第三”图像,并且获取的未经切割的珠粒的图像成为所述系列中的“最后”图像。虽然是按时间颠倒的顺序,但从“前一个”循环中减去所有四种荧光团的强度数据。然后针对混合标记dsDNA标识体链解码成像系列中的每个循环绘制所有四种荧光团的所得数据,如图3B所示。在此

实例中,珠粒是在SpeI切割之后进行成像的,并且针对珠粒上的仅两种剩余的荧光团ATTO550和AF750(分别地,向下斜条纹和浅色虚线条)观察到强信号,所述ATTO550和AF750为预先限定的附接到Id1的标记的组合。从在XhoI切割之后获取的图像中减去针对SpeI切割之后的图像中的所有四种荧光团观察到的强度数据,以得到在SpeI切割循环期间被切割的两种清晰解析的标记ATTO488和AF647(分别地,横条纹和深色虚线条),所述ATTO488和AF647为预先限定的附接到Id2的标记的组合。同样,从在NotI切割之后获取的图像中减去XhoI切割之后在珠粒上观察到的所有四个通道的信号,以便观察在XhoI切割之后去除的标记(不幸的是,在此实验中成像误差使附接到Id3的荧光团不可区分)。最后,从获得自未经切割的珠粒的获取的图像中的荧光团强度中减去在NotI切割之后在图像中在珠粒上观察到的所有强度,以便标识在NotI切割时去除的两种标记ATTO488和AF750(分别地,横条纹和浅色虚线条),所述ATTO488和AF750为附接到Id4的标记的组合。

[0133] 解码dsDNA标识体链:经3个循环跟踪单独珠粒(4个图像)

[0134] 为了编码和解码高多样性的珠粒文库,需要准确跟踪来自跨OCS解码实验的所有循环的位于多个荧光通道中的单独珠粒的荧光强度。图22A至22F示出了经OCS实验的3个循环跟踪的六个单独的携带相同的用于由FOV中的单独珠粒生成数据的单标记标识体链序列(Id1/HindIII-ATTO550--Id2/SpeI-AF750--Id3/XhoI-AF647--Id4/NotI-ATTO488;图21A)的珠粒。将珠粒固定在如前所述的流动池中,并且将流动池转移到热块,以在其如上所述的循环期间利用每种切割剂进行温育。按照对如上针对图21的产生描述的颠倒时间顺序的循环的分析绘制了在OCS的每个循环之前和之后在对所有4种荧光团进行成像之后获得的数据。本实验中跟踪的所有单独珠粒(共10个;在图21中示出6个)展示了经OCS的三个循环去除荧光团的预期序列。首先,通过NotI从Id4段中切割ATTO488(横条纹),然后通过XhoI从Id3中去除AF647(深色虚线),然后通过SpeI从Id2段中去除AF750(浅色虚线),并且Id1段作为SpeI切割后珠粒上的唯一标记而被保留,所述Id1段含有ATTO550(向下斜条纹)。基于这些数据,实验和分析配置已经准备好用于跨OCS的至少3个循环以大规模并行方式同时跟踪许多单独珠粒。

[0135] 标记的dsDNA标识体链结构:解码256种标记的珠粒的文库

[0136] 为了展示对标记的携带dsDNA标识体链的珠粒文库的组合编码,将四种不同的可区分标记,即ATTO488(横条纹)、ATTO550(向下斜条纹)、AF647(暗虚线)和AF750(浅色虚线)附接到四个dsDNA标识体链段(Id1-4),以在每个dsDNA标识体段位置处创建4个差异标记的选项。经通过连接进行的4个轮次的分裂和池化,构建了荧光团标记的dsDNA标识体链的256种不同序列(组合)的文库。简而言之,使含有Id0连接受体dsDNA双链体(此双链体不含有视觉标记,但确实含有用于附接到珠粒的5'生物素修饰,以及适于与Id1段连接的5'悬垂部)的珠粒分裂到4个孔中,以通过如在方法部分中描述的连接附接第一标记的Id段(Id1/HindIII-ATTO488、Id1/HindIII-ATTO550、Id1/HindIII-AF647或Id1/HindIII-AF750)。将连接淬灭,然后将珠粒洗涤并池化以进行混合,并且然后使珠粒分裂到4个孔中,以通过连接附接第二标记的Id段(Id2/SpeI-ATTO488、Id2/SpeI-ATTO550、Id2/SpeI-AF647或Id2/SpeI-AF750)。将连接淬灭,然后将珠粒洗涤并池化以进行混合,并且然后使珠粒分裂到4个孔中,以通过连接附接第三标记的Id段(Id3/XhoI-ATTO488、Id3/XhoI-ATTO550、Id3/XhoI-AF647或Id3/XhoI-AF750)。将连接淬灭,然后将珠粒洗涤并池化以进行混合,并且然后使珠

粒分裂到4个孔中,以通过连接附接第四标记的Id段(Id4/NotI-ATTO488、Id4/NotI-ATTO550、Id4/NotI-AF647或Id4/NotI-AF750)。在淬灭最后的连接反应后,将珠粒洗涤、池化并固定在流动池的单个泳道中的经生物素修饰的表面上。使流动池泳道中的珠粒经受OCS工作流中的3个解码循环。从贯穿整个实验跟踪的单独珠粒中获得的OCS数据示出在图23中。

[0137] 在此,如前所述,在限制性酶暴露之前对珠粒进行成像,然后在暴露于每种单独的酶(每个正交切割事件)之后进行成像。在切割循环1中,使含有5 U/ul NotI酶的溶液流动到流动池泳道中,并且在设置为37℃的热块上温育15分钟。在切割循环1之后,使流动池返回到显微镜,并且在暴露于NotI酶之后对用于对未经切割的珠粒行成像的同一FOV进行成像。这实现了对跨所述系列的前两个图像的相同的单独珠粒进行跟踪。在切割循环2中暴露于XhoI酶,并且在切割循环3中暴露于SpeI酶的情况下重复此过程。获得了附接到跨本文示出的OCS实验跟踪的三个珠粒的所有四种荧光团标记的强度值,利用如上所述的“前一个”循环减法按颠倒时间顺序分析图像,并且绘制标记去除的逐循环可视化的值。在此实验中跨所有切割循环跟踪的三个珠粒示出在图23A、23B和23C中。图23中示出的三个珠粒中的每个珠粒都含有唯一的dsDNA标识体链序列,所述dsDNA标识体链序列易于解析:a)珠粒5序列:

[0138] Id1/HindIII-AF647--Id2/SpeI-AF750--Id3/XhoI-ATTO488--Id4/NotI-ATTO488;b)珠粒6序列:Id1/HindIII-ATTO488--Id2/SpeI-ATTO550--Id3/XhoI-AF647--Id4/NotI-ATTO488;以及c)珠粒2序列:

[0139] Id1/HindIII-ATTO550--Id2/SpeI-AF750--Id3/XhoI-ATTO488--Id4/NotI-AF750)。下面针对所跟踪的3个珠粒示出了展示在所述系列中的每个成像步骤获得的数据,即,通过每种酶释放的每个Id段(Id)的说明的条形图。成像循环由竖直虚线分开。

[0140] 标记的DNA发夹结构

[0141] 通过将预杂交的dsDNA双链体(含有位于双链体的一条寡核苷酸链的5’端上的生物素修饰和附接到双链体的另一条链的可检测标记(AF-647,a和b中的横条纹))和含有两个AF-750标记的发夹寡核苷酸(a和b中的浅色虚线)连接来构建标识体。所产生的连接的发夹含有三个可检测标记,即一个用于附接到链霉亲和素珠粒的5’生物素修饰和位于AF-647标记侧面的两个完全形成且正交的限制性位点。两个限制性位点中的第一限制性位点(RE1位点,对通过SpeI进行的切割具有特异性)定位在附接在发夹的茎内的AF-647标记与附接到发夹的环的AF-750标记之间。这种标识体结构使得能够构建短的标识体链,以用通过可正交切割的接头附接的荧光团的许多不同组合装饰珠粒。这种发夹结构确保了在切割之前限制性位点保持为双链区,因为在此互补链是通过发夹结构的环融合的。连接后,将此寡核苷酸附接到链霉亲和素珠粒,并且将珠粒固定在流动池中的经生物素修饰的表面上,以在暴露于OCS的一个循环之前和之后进行成像(图24A)。从单个视场中的许多珠粒中获得值,并且将平均强度值以百分比绘制。在酶暴露之前,珠粒表现出来自AF-647和AF-750标记两者的清晰可测量信号(图24B之前),并且因此在没有发生切割的情况下,这些值是切割后的图像所预期的信号的100%。在暴露于SpeI酶时,双AF-750标记的发夹“帽”从发夹标识体中释放,如在切割后提取的图像中观察到的(之后的图24B)。在切割后提取的图像中,珠粒含有的AF-647信号比在切割之前获取的图像中观察到的AF-647信号略大100%,但含有的AF-

750信号比暴露于SpeI的AF-750信号小20％。虽然发夹的第二标记可通过XhoI切割,但此荧光团作为珠粒上的最高剩余信号可被清楚地区分,并且因此它其切割不是其标识所必须的。另外,先前已经展示了对保留在珠粒上的结构的成功切割(在此)。

[0142] 通过逐步连接编码标识体环结构

[0143] 链霉亲和素珠粒示出在图25A中,其含有通过与第一未标记的DNA发夹连接进行附接的第一标记的DNA发夹,所述第一未标记的DNA发夹通过生物素部分附接到珠粒。第一标记的DNA发夹在与附接到珠粒的第一连接受体DNA发夹寡核苷酸发生模板化连接时产生标记的标识体段,以作为含有位于dsDNA区内的经编码的限制性位点的ssDNA环。每种标记的DNA发夹的连接被设计成仅对一个受体发夹寡核苷酸具有特异性,从而使得能够每次对一种标记的发夹寡核苷酸进行正交逐步连接。这允许在构建含有由DNA环构成的标识体的珠粒的文库时采取分裂和池化方法。为了展示珠粒上DNA环标识体的逐步构建,在三个轮次的编码中使用了三个标记的DNA发夹,以模拟分裂和池化工作流。含有ATTO550标记(向下斜条纹)的第一发夹附接到珠粒上的三个受体寡核苷酸之一。新形成的标记的DNA环标识体还含有酶可接近的限制性核酸内切酶位点(SpeI),所述SpeI定位在附接的标记与珠粒之间。通过三个荧光发射通道488nm、550nm和647nm中的荧光成像证实了ATTO550标记与珠粒的连接,其中仅在550nm波长中观察到强发射信号。证实高效连接的图像以条形图形式呈现,所述条形图是使用从以下成像的所有3个发射波长以及许多携带相同寡核苷酸的珠粒(平均强度值)获取的数据绘制的:488nm(横条纹)、550nm(向下斜条纹)、647nm(虚线)。图中的条高度对应于在成像的所有三个通道中观察到的发射的相对量值,所述相对量值以荧光强度单位(FU)表示。图25B示出了含有通过连接正交地附接到珠粒上的三个受体寡核苷酸中的第二受体寡核苷酸的AF-647标记(虚线)的第二标记的DNA发夹。新形成的标记的DNA环标识体还含有酶可接近的限制性核酸内切酶位点(XhoI),所述XhoI定位在附接的标记与珠粒之间。通过三个荧光发射通道488nm、550nm和647nm中的荧光成像证实了AF-647标记与珠粒的连接,其中现在ATTO-550标记和AF-647标记两者观察到强发射信号。证实高效连接的图像以条形图形式呈现,所述条形图是使用从以下成像的所有3个发射波长以及许多携带相同寡核苷酸的珠粒(平均强度值)获取的数据绘制的:488nm(横条纹)、550nm(向下斜条纹)、647nm(虚线)。图中的条高度对应于在成像的所有三个通道中观察到的发射的相对量值,所述相对量值以荧光强度单位(FU)表示。图25C示出了含有通过连接正交地附接到珠粒上的三个受体寡核苷酸中的最后一个(第三)受体寡核苷酸的ATTO-488标记(横条纹)的第三标记的DNA发夹。新形成的标记的DNA环标识体还含有酶可接近的限制性核酸内切酶位点(NotI),所述NotI定位在附接的标记与珠粒之间。通过三个荧光发射通道488nm、550nm和647nm中的荧光成像证实了ATTO-488标记与珠粒的连接,其中现在所有三种标记(ATTO-550、AF-647和ATTO-488)都观察到强发射信号。证实高效连接的图像以条形图形式呈现,所述条形图是使用从以下成像的所有3个发射波长以及许多携带相同寡核苷酸的珠粒(平均强度值)获取的数据绘制的:488nm(横条纹)、550nm(向下斜条纹)、647nm(虚线)。图中的条高度对应于在成像的所有三个通道中观察到的发射的相对量值,所述相对量值以荧光强度单位(FU)表示。

[0144] 解码标识体环结构

[0145] 为了证实在此产生的标记的标识体环结构在OCS工作流中是功能性的,使经受三

个轮次的差异标记的DNA环标识体构建的珠粒暴露于正交切割的单个循环。将在(图25)中构建的相同珠粒固定在流动池中的经生物素修饰的表面上,并且在此使用。这些珠粒含有三种不同的环标识体,即Id1/SpeI-ATTO-550、Id2/XhoI-AF-647和Id3/NotI-ATTO-488(图26A)。在暴露于SpeI酶之前和之后对珠粒进行成像。在暴露于SpeI后,通过三个荧光发射通道488nm、550nm和647nm中的荧光成像证实了ATTO-550标记从珠粒上的高效切割,其中观察到相对于两种其它标记获得的信号,ATTO-550标记的信号显著损失。证实高效SpeI切割的图像以条形图形式呈现(图26B),所述条形图是使用从以下成像的所有3个发射波长以及许多携带相同寡核苷酸的珠粒(平均强度值)获取的数据绘制的:488nm(横条纹)、550nm(向下斜条纹)、647nm(虚线)。图中的条高度对应于在成像的所有三个通道中观察到的发射的相对量值,所述相对量值以荧光强度单位(FU)表示。

[0146] 通过逐步杂交编码标识体

[0147] 链霉亲和素珠粒示出在图27A中,其含有与第一未标记的ssDNA杂交的第一标记的ssDNA,所述第一未标记的ssDNA通过生物素部分附接到珠粒。第一标记的DNA在与附接到珠粒的三个正交互补链的模板化杂交时产生标记的标识体段,作为含有位于dsDNA区内的经编码的限制性位点的dsDNA双链体。所需的限制性酶识别序列的一半编码在每条互补链中。每个标记的ssDNA的杂交被设计成仅对珠粒上的一个互补寡核苷酸具有特异性,从而使得能够一次对一个标记的ssDNA寡核苷酸进行正交逐步杂交。这允许在构建含有由标记的杂交的dsDNA构成的标识体的珠粒的文库时采取分裂和池化方法。为了展示珠粒上杂交的标识体的逐步构建,在三个轮次的编码中使用了三条标记的ssDNA链,以模拟分裂和池化工作流。使含有ATTO-488标记(横条纹)的第一ssDNA寡核苷酸与珠粒上的三个互补寡核苷酸之一杂交。新形成的标记的dsDNA杂交的标识体还含有酶可接近的限制性核酸内切酶位点(NotI),所述NotI定位在附接的标记与珠粒之间。通过三个荧光发射通道488nm、550nm和647nm中的荧光成像证实了ATTO-488标记与珠粒的连接,其中仅在488nm波长中观察到强发射信号。证实高效杂交的图像以条形图形式呈现,所述条形图是使用从以下成像的所有3个发射波长以及许多携带相同寡核苷酸的珠粒(平均强度值)获取的数据绘制的:488nm(横条纹)、550nm(向下斜条纹)、647nm(虚线)。图中的条的高度对应于在成像的所有三个通道中观察到的发射的相对量值,所述相对量值以荧光强度单位(FU)表示。图27B示出了含有与珠粒上的三个互补寡核苷酸中的第二互补寡核苷酸正交地杂交的AF-647标记(虚线)的第二标记的ssDNA寡核苷酸。新形成的标记的dsDNA杂交的标识体还含有酶可接近的限制性核酸内切酶位点(SpeI),所述SpeI定位在附接的标记与珠粒之间。通过三个荧光发射通道488nm、550nm和647nm中的荧光成像证实了AF-647标记与珠粒的杂交,其中现在ATTO-488标记和AF-647标记两者观察到强发射信号。证实高效杂交的图像以条形图形式呈现,所述条形图是使用从以下成像的所有3个发射波长以及许多携带相同寡核苷酸的珠粒(平均强度值)获取的数据绘制的:488nm(横条纹)、550nm(向下斜条纹)、647nm(虚线)。图中的条的高度对应于在成像的所有三个通道中观察到的发射的相对量值,所述相对量值以荧光强度单位(FU)表示。图27C示出了含有与珠粒上的三个互补寡核苷酸中的最后一个(第三)互补寡核苷酸正交地杂交的ATTO-550标记(向下斜条纹)的第三标记的ssDNA寡核苷酸。新形成的dsDNA标识体还含有酶可接近的限制性核酸内切酶位点(XhoI),所述XhoI定位在附接的标记与珠粒之间。通过三个荧光发射通道488nm、550nm和647nm中的荧光成像证实了ATTO-550

标记与珠粒的连接,其中现在所有三种标记(ATTO-550、AF-647和ATTO-488)都观察到强发射信号。证实高效杂交的图像以条形图形式呈现,所述条形图是使用从以下成像的所有3个发射波长以及许多携带相同寡核苷酸的珠粒(平均强度值)获取的数据绘制的:488nm(横条纹)、550nm(向下斜条纹)、647nm(虚线)。图中的条的高度对应于在成像的所有三个通道中观察到的发射的相对量值,所述相对量值以荧光强度单位(FU)表示。

[0148] 解码杂交的标识体

[0149] 为了证实在此通过杂交产生的标记的标识体结构在OCS工作流中是功能性的,使经受三个轮次的通过杂交进行的差异标记的ssDNA标识体构建的珠粒暴露于正交切割的两个循环。将在(图27)中构建的相同珠粒固定在流动池中的经生物素修饰的表面上,并且在此使用。这些珠粒含有三种不同的杂交的标识体,即Id1/NotI-ATTO-488、Id2/SpeI-AF-647和Id3/XhoI-ATTO-550。在以已知顺序逐循环暴露于两种酶(首先暴露于NotI,然后暴露于SpeI)之前和之后如先前所述(图28A和28B)对珠粒进行成像。在暴露于NotI后,通过三个荧光发射通道488nm、550nm和647nm中的荧光成像证实了ATTO-488标记从珠粒上的高效切割,其中观察到相对于两种其它标记获得的信号,ATTO-488标记的信号显著损失。证实高效NotI切割的图像以条形图形式呈现(图28C),所述条形图是使用从以下成像的所有3个发射波长以及许多携带相同寡核苷酸的珠粒(平均强度值)获取的数据绘制的:488nm(横条纹)、550nm(向下斜条纹)、647nm(虚线)。图中的条的高度对应于在成像的所有三个通道中观察到的发射的相对量值,所述相对量值以荧光强度单位(FU)表示。然后使珠粒暴露于通过SpeI酶进行的正交切割的第二循环,并且再次进行成像。在暴露于SpeI后,通过三个荧光发射通道488nm、550nm和647nm中的荧光成像证实了AF-647标记从珠粒上的高效切割,其中观察到相对于SpeI切割之前的珠粒并且相对于珠粒(ATTO-550)上的剩余的信号,AF-647标记的信号显著损失。证实高效SpeI切割的图像以条形图形式呈现(图10C),所述条形图是使用从以下成像的所有3个发射波长以及许多携带相同寡核苷酸的珠粒(平均强度值)获取的数据绘制的:488nm(横条纹)、550nm(向下斜条纹)、647nm(虚线)。图中的条的高度对应于在成像的所有三个通道中观察到的发射的相对量值,所述相对量值以荧光强度单位(FU)表示。

[0150] 实例

[0151] 以下实施例进一步描述和展示了所公开的分子条形码的实施例的使用。实例仅出于说明的目的而给出,并且不应被解释为限制分子条形码的使用。在不脱离本公开的精神和范围的情况下,这些实例的许多变化是可能的。

[0152] 实例1基于荧光团/蛋白酶位点/非DNA(FPND)标识体类别的分子条形码

[0153] 实例1说明了荧光团/蛋白酶位点/非DNA(FPND)标识体在分子条形码中用于原位标记指定用于视觉解码的材料的用途(图1-3、15、16)。通常,FPDN标识体包括基于荧光团的可检测标记,所述基于荧光团的可检测标记操作性地连接到基于多肽的支架部分,其中切割位点是正交反应性蛋白酶切割位点,并且识别部分是蛋白酶切割位点序列或正交反应性蛋白酶的蛋白酶识别序列。

[0154] 为了证实基于FPND标识体的分子条形码的编码和解码能力,可以进行四个循环的正交蛋白酶切割实验(实验1)。在此,可以并行使用各种FPND标识体类别以通过分裂-池化连接来编码和形成分子条形码,以在不依赖于基于DNA测序的解码的情况下创建分子条形码组合文库。

[0155] 实验1中的每个分子条形码包括源自一或多种FPDN标识体类别的五个标识体符号(Id_{1-5})构成的集合。每个Id_1符号源自不可切割的FPDN标识体类别,并且Id_2、Id_3、Id_4和Id_5符号中的每一个源自可切割的FPDN标识体类别。每个可切割的FPDN标识体被配置成与至少一种蛋白酶正交地发生反应,所述至少一种蛋白酶选自烟草蚀刻病毒蛋白酶(TEVp)、烟草叶脉斑点病毒蛋白酶(TVMVp)、芜菁花叶病毒蛋白酶(TUMVp)和向日葵轻度花叶病毒蛋白酶(SuMMVp)。例如,本领域已知TEVp在人类蛋白质组中不具有已知的脱靶底物,并且可以用作正交反应性切割剂。并且本领域已知的是TEVp、TVMPp1、TUMVp2和SuMMVp已在正交蛋白酶体系中实施。因此,TEVp不应在TVMVp、TUMVp或SuMMV所识别的切割位点处高效地切割分子条形码。换言之,每种蛋白酶都应能够在存在其它蛋白酶所识别的底物的情况下在具有包括相应的蛋白酶的切割位点序列的识别部分的标识体的切割位点处切割分子条形码。也可以使用因子Xa。因子Xa在精氨酸残基之后在其优选的切割位点中切割Ile-(Glu或Asp)-Gly-Arg。其将不切割脯氨酸或精氨酸之后的位点。

[0156] 在实验1中,FPND标识体支架部分包括通过其N末端和C末端与含有化学反应性基团的接头试剂共价连接的肽片段,并且可检测标记包括一或多个Q点。向含有蛋白水解位点的肽片段中的携带经修饰的叠氮化物的氨基酸中的Id_1、Id_2、Id_3、Id_4和Id_5分别添加Q点荧光团。在文库中的任何所形成的FPDN标识体条形码(包括Id_1、Id_2、Id_3、Id_4和Id_5符号)上展示的荧光团的组合可以具有或可以不具有连续发射光谱,因为分子条形码内构成文库的FPND标识体符号的三维布置将会变化。

[0157] 可以对充当切割剂的蛋白酶进行检测限、特异性和精度实验(一式三份),以建立与基于FPND标识体的分子条形码的蛋白酶正交反应性的再现性。

[0158] A.FPND标识体级联

[0159] 如先前所述,FPND标识体可以含有附接在其N末端和C末端处的正交点击化学修饰。在实验1中,每种Id_1、Id_3和Id_5标识体类别都包括位于其N末端处的TCO基团,而每种Id_2和Id_4标识体类别都包括位于其C末端处的甲基四嗪基团。Id_1和Id_3标识体类别的C末端包括DBCO基团,并且Id_2和Id_4标识体类别包括位于其N末端处的叠氮化物基团。

[0160] 第一标识体附接(添加Id_1以附接到经甲基四嗪修饰的珠粒)为在37℃下温育的45分钟反应。在所有反应期间,珠粒都保持为1mg/ml,并且对于所有随后添加,都使用10μM浓度的标识体单位。在1X PBS中进行三个洗涤步骤之后,使珠粒经受更高浓度(100uM)的游离TCO,以确保所有甲基四嗪位点都饱和。在1X PBS中进行三个洗涤步骤后,珠粒准备好添加第二标识体单位(Id_2)。对Id_2的附接,以及向FPND分子条形码的每次随后标识体添加(即,Id_3、Id_4和Id_5)重复相同的反应条件。由于所列出的点击反应的反应动力学,在每次标识体添加之后,通过添加适当的点击反应基团来"追赶"反应,以确保在添加下一个标识体之前,用于标识体附接的所有反应位点是饱和的。

[0161] B.每个FPND标识体文库类别的制备

[0162] Id_{1-5}标识体类别的识别部分分别包括以下肽序列:Id_1不可切割并且不含有切割位点;Id_2包括KGGSGGGSACVYHQSGGAzGGSC(SEQ ID NO:1),所述序列含有TUMV切割位点;Id_3包括KGGSGGGSEEIHLQSGGGAzGGSC(SEQ ID NO:2),所述序列含有SuMMV切割位点;Id_4包括KGGSGGGSETVRFQGGGAzGGSC(SEQ ID NO:3),所述序列含有TVMV切割位点;并且Id_5包括KGGSGGGSENLYFQSGGAzGGSC(SEQ ID NO:4),所述序列含有TEV切割位点。

[0163] 总体上,可以通过首先使用用于可视化的经DBCO修饰的荧光团来修饰肽(通过内部地安装在所有肽的C末端附近的携带叠氮化物的残基)来产生FPND标识体。为此,40μM由每种肽构成的样品可以在200uM下与指定的经DBCO修饰的荧光团混合,并且允许其在1X PBS中在37℃下反应至少3小时;此反应也可以在室温下保持过夜。然后可以使用HPLC、透析或脱盐柱对经荧光团修饰的肽进行纯化,以去除过量的(和任何未经反应的)DBCO-荧光团试剂,并将肽交换到NHS相容性反应缓冲液中(100mM磷酸钠缓冲液,pH 8.5,补充有80mM KCl)。标记的肽也可以使用本领域已知的各种方法沉淀,并且重悬于NHS相容性反应缓冲液中。然后可以使标记的肽在NHS相容性反应缓冲液中为20μM浓度,并且与100μM其指定NHS试剂(对于Id_1、Id_3和Id_5肽片段,为NHS-PEG4-TCO;对于Id_2和Id_4,为NHS-PEG4-叠氮化物)混合,以实现在室温下过夜缀合。经N末端(单)点击修饰的标记的肽可以使用HPLC、透析或脱盐柱进行纯化,以去除过量的(和任何未孔反应的)含NHS的试剂,并且将肽交换到1X PBS中。标记的且经单修饰的肽也可以使用本领域已知的各种方法沉淀,并且重悬于1X PBS缓冲液中。然后可以使标记的且经单修饰的肽在1X PBS中为20μM浓度,并且与100μM其指定马来酰亚胺试剂(对于Id_1、Id_3和Id_5肽片段,为马来酰亚胺-PEG4-DBCO;对于Id_2和Id_4,为马来酰亚胺-PEG_4-甲基四嗪)混合,以实现在室温下过夜缀合。标记的且经双修饰的肽可以使用HPLC、透析或脱盐柱进行纯化,以去除过量的(和任何未孔反应的)含马来酰亚胺的试剂,并将肽交换到1X PBS中。标记的且经双独修饰的肽也可以使用本领域已知的各种方法沉淀,并且重悬于1X PBS缓冲液中。FPND标识体然后可以立即使用,储存在4℃下持续几天,或储存在-20℃下,以长期储存。

[0164] C.实验1-基于FPND标识体的分子条形码验证

[0165] 可以如下完成对基于FPND标识体的分子条形码性能的验证:使用通过切割实验进行的四循环去卷积/解码,以证实标识体形成和编码基于FPND标识体的分子条形码的级联,以及其在通过正交反应性切割解码时的可检测信号响应。

[0166] 通常,为了检测可检测信号响应,进行了一系列成像和切割步骤以解码存在于组合文库的成员(即,固定到珠粒的分子条形码)上的基于FPND标识体的分子条形码。此过程可以循环重复,由此可以在每次切割循环期间引入下一种正交蛋白酶,然后进行成像步骤。

[0167] 在实验1中,在第一切割步骤中引入TEV蛋白酶,以降低仅附接到TEVp反应性标识体(即,Id_5)的荧光团的观察到的信号强度,然后进行第一成像步骤,以在视觉上标识从并入在包括组合文库的分子条形码中的Id_5符号中切割的可检测标记。切割循环按以下顺序引入所列蛋白酶:循环1期间TEVp;循环2期间TVMVp;循环3期间SuMMVp;并且循环4期间TUMVp。

[0168] Id_1是添加到链的第一标识体,并且因此直接附接到固体载体(珠粒);在此Id_1不包含切割位点,并且其识别部分可以为[不可切割的接头]。如以上所列出的,Id_1含有N末端TCO修饰,所述TCO修饰可以用于附接到固体载体(珠粒);固体载体的表面被修饰成含有用于附接Id_1的相容性甲基四嗪部分。Id_2、Id_3、Id_4和Id_5的识别部分和切割位点为分别地TUMVp、SuMMVp、TVMVp和TEV蛋白酶的识别序列和切割位点。

[0169] 如实验1中所用,Id_{1-5}标识体形成于如上所述的珠粒上。然后,在用于解码分子条形码的第一切割步骤中,使珠粒暴露于含有2个单位的TEVp的溶液中,以在30℃下在蛋白酶反应缓冲液(pH 7.5的50mM Tris-HCl、0.5mM EDTA、1mM DTT)中进行30-60分钟温育。温育

后,在第一成像步骤中对珠粒进行成像,并且检测响应于TEVp的荧光团。在此,可检测信号响应包括在第一切割步骤期间从分子条形码中解离的FPND标识体符号的可检测标记的发射波长强度的降低。在洗涤步骤之后,在第二切割步骤中,使珠粒暴露于含有2个单位的TVMVp的溶液,以在30℃下在蛋白酶反应缓冲液中进行30-60分钟温育。温育后,在第二成像步骤中,对珠粒进行成像,并且检测响应于TVMVp的荧光团。在此,可检测信号响应包括在第二切割步骤期间从分子条形码中解离的FPND标识体符号的可检测标记的发射波长强度的降低。在另外的洗涤步骤之后,在第三切割步骤中,使珠粒暴露于含有2个单位的SuMMVp的溶液,以在30℃下在蛋白酶反应缓冲液中进行30-60分钟温育。温育后,在第三成像步骤中,对珠粒进行成像,并且检测响应于SuMMVp的荧光团。在此,可检测信号响应包括在第三切割步骤期间从分子条形码中解离的FPND标识体符号的可检测标记的发射波长强度的降低。最后,在第四切割步骤中,使珠粒暴露于含有2个单位的TUMVp的溶液,以在30℃下在蛋白酶反应缓冲液中进行30-60分钟温育。温育后,在第四成像步骤中,对珠粒进行成像,并且检测响应于TUMVp的荧光团。在此,可检测信号响应包括在切割步骤期间从分子条形码中解离的FPND标识体符号的可检测标记的发射波长强度的降低。在去除最终的可切割标识体标记之后,不可切割标记将保持是从珠粒发射的最强信号,从而允许其进行标识并潜在用于定量。在分裂-池化和级联期间形成的每个条形码的荧光团的序列是在视觉去卷积期间确定的,并且可以与其所附接的珠粒相关。

[0170] 实例2基于荧光团/核酸酶位点(FNS)标识体的分子条形码

[0171] 实例2说明了基于荧光团/核酸酶位点(FNS)标识体的分子条形码用于原位标记指定用于视觉解码的材料的用途(图5、7、9)。

[0172] 可以使用由一或多种FNS标识体类别构成的集合以通过分裂-池化连接来编码和形成视觉分子条形码,以创建不需要基于DNA测序的解码的条形码的组合文库。如实例2中所示,可以在同一实验中并且在存在同时构建的许多不同的条形码链的情况下产生单个条形码。本领域已知许多核酸内切酶相对于彼此具有正交反应性。例如,NotI、XhoI、SpeI和HindIII对分别含有GCGGCCGC、CTCGAG、ACTAGT和AAGCTT的dsDNA序列具有特异性。因此,NotI不应高效地特异性切割XhoI、SpeI或HindIII所识别的位点。对于所列出的核酸酶中的每种核酸酶也是如此,因为在存在由其它酶所识别的底物的情况,每种核酸酶都应能够仅切割(主要地)其相应底物。此方面提供了基于标识体的分子条形码标识期间的正交性。所展示的标识体条形码文库的每个成员含有五种类别(Id_{1-5}),所述五种类别由四个可切割Q荧光团/核酸酶位点标识体符号(Id_2、Id_3、Id_4和Id_5)以及一个不可切割标识体符号(Id_1)组成。

[0173] 实验2为四循环正交核酸酶切割实验,其是通过对反应性荧光团进行视觉成像证实的,被进行以证实基于FNS标识体的分子条形码的编码和解码。实验2包括第一成像步骤,所述第一成像步骤用于记录由构成存在于文库的任何成员(珠粒)上的基于FNS标识体的分子条形码的可检测标记所传递的可检测信号。在第一切割步骤中引入NotI,以降低仅附接到NotI核酸酶反应性标识体的荧光团的观察到的信号强度,从而允许视觉表示附接到Id_5的标记。此过程可以循环重复,由此可以在每次切割步骤期间引入下一种正交核酸酶,然后进行成像步骤。

[0174] 在第一成像步骤之后,实验2的每个循环包括切割步骤和成像步骤。核酸酶切割剂

按以下顺序应用于分子条形码:循环1期间NotI;循环2期间XhoI;循环3期间SpeI;并且循环4期间HindIII。每个循环之后的成像将使得能够对每个构成基于FNS标识体的分子条形码的视觉标记添加的顺序进行去卷积。可以对充当切割剂的核酸酶进行检测限、特异性和精度实验(一式三份),以建立与基于FNS标识体的分子条形码的核酸酶正交反应性的再现性。

[0175] 在一些实施例中,FNS标识体支架部分包含具有5'端和3'端的ssDNA,所述ssDNA用可检测标记通过经修饰的内部碱基进行荧光标记,并且与互补的ssDNA预杂交以形成dsDNA双链体。所形成的dsDNA双链体可以编码单个核酸内切酶限制性序列(或位点类型)的一或多个拷贝。本领域的普通技术人员,借助于本公开将理解,可检测标记从所形成且经编码的分子条形码中的高效解离可以通过将核酸内切酶限制性序列沿着支架部分选择性地间隔开来促进。在一些实施例中,含有核酸酶识别序列的dsDNA双链体的一条链包括用于附接预先指定的荧光团的内部氨基修饰。

[0176] 如实验2中所用,每种FNS标识体类别成员的支架部分都包括相同的限制性核酸内切酶位点类型,但接收不同且可区分的荧光团。然后可以使用氨基反应性异双功能交联试剂NHS-PEG_4-TCO,以化学修饰dsDNA上的所安装的伯胺,以含有点击相容性TCO基团。经甲基四嗪修饰的荧光团然后可以与标识体dsDNA双链体缀合。

[0177] 如实验2中所用,FNS标识体的支架部分包括未配对的3'端,所述未配对的3'端适于与其它FNS标识体类别连接,以便于通过分裂-池化构建标识体的组合链。向正在生长的FNS标识体分子条形码的每个轮次的添加将添加用于正在形成的分子条形码的唯一的标识体符号。

[0178] 如实验2中所用,Id_1为所添加的第一标识体,并且因此直接附接到固体载体(即,珠粒)。Id_1不包含切割位点,并且其识别部分为不可切割接头。Id_1含有用于附接到固体载体的经5'氨基修饰的碱基。如实验2中所用,Id_2、Id_3、Id_4和Id_5的识别部分和切割位点为分别地HindIII、SpeI、XhoI和NotI核酸内切酶的识别序列和切割位点。

[0179] 如实验2中所用,Id_{1-5}类别的可检测标记包括一或多个量子点。每种FNS标识体类别包括具有不同发射波长的量子点,所述量子点共价附接到携带特定核酸酶位点类型的每个类别特异性dsDNA的不同反应混合物中的FNS标识体(或板的孔)。如实验1中所用,通过位于dsDNA序列的一或两条链内的经内部修饰的碱基分别向Id_1、Id_2、Id_3、Id_4和Id_5添加量子点可检测标记。在文库中的基于经编码的FNS标识体的分子条形码(本文示出的Id_1、Id_2、Id_3、Id_4和Id_5)上展示的荧光团的组合可以具有或可以不具有连续发射光谱,因为包括构成文库的分子条形码的标识体符号的三维构型是通过如本文所述的正交核酸酶的循环确定的。

[0180] A.FNS标识体级联

[0181] 如实验2中所用,FNS标识体识别部分包括形成dsDNA双链体的互补ssDNA链。FNS标识体切割位点包括包含在dsDNA双链体中的特定限制性核酸内切酶位点,并是通过杂交形成的,由此两条链共有显著互补性,但在其3'端处保持不配对。每种FNS标识体类别的未配对的3'端被设计成与相邻的类别成员上的受体3'端互补。

[0182] 在一些实施例中,FNS标识体可以按预先限定的顺序依序添加,以经连续轮次的分裂和池化来形成条形码。每个轮次的添加都将使得能够通过杂交的标识体的ssDNA片段的5'和3'端的酶连接来进行标识体dsDNA的级联。每个标识体的连接可以在1X连接酶缓冲液(pH 7.5的50mM Tris-HCl、10mM MgCl2、1mM ATP、10mM DTT)中,在存在500-1,000U T4 DNA

连接酶和20-80U多核苷酸激酶的情况下,经37℃下的25分钟温育进行。图示的荧光团/核酸酶位点标识体条形码含有5个通过酶连接进行连接的类别成员。

[0183] B.FNS标识体的制备

[0184] NotI、XhoI、SpeI和HindIII为对分别含有GCGGCCGC、CTCGAG、ACTAGT和AAGCTT的dsDNA序列具有特异性的核酸内切酶。如实验2中所用,Id_1被配置成不可切割,并且ID_{2-5}的识别部分分别包括:Id_2包括

[0185] ATTTATTTAAGCTTATTA/iAmMC6T/TATTT(SEQ ID NO:5),所述序列含有HindIII切割位点;Id_3包括ATTTATTTAACTAGTATTA/iAmMC6T/TATTT(SEQ ID NO:6),所述序列含有SpeI切割位点;Id_4包括ATTTATTTACTCGAGATTA/iAmMC6T/TATTT(SEQ ID NO:7),所述序列含有XhoI切割位点;并且Id_5包括

[0186] ATTTATTTAGCGGCCGCATTA/iAmMC6T/TATTT(SEQ ID NO:8),所述序列含有NotI切割位点。Id_1的不可切割的dsDNA可以由任何不能被任何用于识别构成基于FSN标识体的分子条形码的其它类别成员的核酸酶识别的DNA序列组成。

[0187] 如本文所用,“iAmMC6T”是指可从集成DNA技术公司(Integrated DNA Technologies,Inc.,美国爱荷华市52241克拉尔维尔的商业公园1710(1710 Commercial Park,Coralville,Iowa 52241,USA))商购获得的Int氨基修饰剂C6 dT。在一些实施例中,经修饰的核苷酸(例如,iAmMC6T)被配置成离识别部分足够远,以便在经修饰的核苷酸被标记时不干扰切割剂进入识别部分。在一些实施例中,经修饰的核苷酸被配置成离识别部分至少6bp。在一些实施例中,内部地放置在环状dsDNA内的经修饰的核苷酸被放置成离彼此足够远,以避免基于FRET的淬灭。在一些实施例中,内部地放置在环状dsDNA内的经修饰的核苷酸被放置成离彼此足够远,以避免酶识别位点并促进酶可及性。

[0188] 如实验2中所用,通过首先将一条互补链粘接来产生每种FNS标识体类别,所述互补链与上面列出的类别特异性ssDNA链(Id_{2-5})一起产生可行的dsDNA限制性位点。这是以1:1摩尔比在NHS相容性粘接缓冲液(100mM磷酸钠缓冲液,pH值8.5,补充有80mM KCl)中在热块上在95℃下持续5分钟完成的,然后将热块移除,并且允许在工作台上经相当于约2小时的时间段冷却至室温。

[0189] 在一些实施例中,携带单个核酸内切酶位点类型的标识体类别特异性dsDNA可以使用NHS-PEG_4-反式环辛烯进行修饰(通过以上序列中示出的内部氨基修饰),使得其含有用于下游缀合的相容化学。为此,将100μM标记的dsDNA与500μM NHS-PEG_4-反式环辛烯(TCO)在NHS相容性粘接缓冲液(100mM磷酸钠缓冲液,pH 8.5,补充有80mM KCl)中混合,并且允许在室温下反应过夜。然后可以通过脱盐以去除过量的(和任何未经反应的)NHS-PEG_4-TCO试剂来对经TCO修饰的dsDNA双链体进行纯化。许多经甲基四嗪修饰的荧光团都是可商购获得的,并且可以用于以指定方式标记每个类别特异性标识体。为此,每个类别特异性标识体可以分裂到不同的孔中,并且每种类别的100μM经TCO修饰的dsDNA可以与在NHS相容性粘接缓冲液中以200μM使用的不同的经甲基四嗪修饰的荧光团缀合。此反应可以在室温下进行30分钟。然后可以将缀合的FNS标识体缓冲液交换到储存缓冲液(50mM Tris-HCl,pH 7.5,补充有100mM NaCl)中,并且可以立即使用,或可以储存在4℃下持续几天或储存在-20℃下,以便长期储存。

[0190] C.基于FNS标识体的分子条形码验证

[0191] 对基于FNS标识体的分子条形码的标识可以使用通过切割实验进行的四循环去卷积/解码完成,以证实Id_{1-5}的可检测信号响应,并证实基于FNS标识体的分子条形码的相应依序级联和编码。

[0192] 在一些实施例中,分子条形码文库可以形成于珠粒上,并且珠粒可以固定在表面上并且在与实验溶液接触之前和之后成像。首先可以使用配置有适当激发波长和发射滤光片的标准荧光显微镜对所固定的珠粒进行成像,以记录给定视场或所关注区中由构成每个分子条形码的可检测标记传递的可检测信号响应的组合。所列出的核酸内切酶的所有"高保真"版本都可以在相同的缓冲液(可从美国马萨诸塞州01969罗利纽伯里波特转盘428的新英格兰生物实验室公司(New England Biolabs,Inc.at 428Newburyport Turnpike,Rowley,MA01969,U.S.A.)商购获得的1X CUTSMART缓冲液)中以100%效率起作用。

[0193] 如实验2中所用,在第一切割步骤中,将珠粒暴露在含有5个单位NotI-HF核酸内切酶的溶液中,以在37℃下在CUTSMART缓冲液(50mM乙酸钾、20mM三乙酸盐、10mM乙酸镁、100ug/ml BSA;缓冲液pH在25℃下为7.9)中进行5-10分钟温育。温育后,可以在第二成像步骤中对珠粒进行成像,以检测响应NotI核酸内切酶活性的荧光团可检测标记所传递的可检测信号响应。在此,可检测信号响应包括在第一切割步骤期间从分子条形码中解离的FNS标识体符号的可检测标记的发射波长强度的降低。任选地,可以在将珠粒暴露于第二切割步骤之前进行洗涤步骤。

[0194] 第二切割步骤包括将珠粒暴露于含有5个单位的XhoI-HF的溶液中,以在37℃下在CUTSMART缓冲液中进行5-10分钟温育。温育后,在第二成像步骤中对珠粒进行成像,并且检测响应于XhoI核酸内切酶活性的荧光团可检测标记所传递的可检测信号响应。在此,可检测信号响应包括在第二切割步骤期间从分子条形码中解离的FNS标识体符号的可检测标记的发射波长强度的降低。任选地,可以在将珠粒暴露于第三切割步骤之前进行洗涤步骤。

[0195] 第三切割步骤包括将珠粒暴露于含有5个单位的SpeI-HF的溶液中,以37℃下在CUTSMART缓冲液中进行5-10分钟温育。温育后,在第二成像步骤中对珠粒进行成像,并且检测响应于SpeI核酸内切酶活性的荧光团可检测标记所传递的可检测信号响应。在此,可检测信号响应包括在第三切割步骤期间从分子条形码中解离的FNS标识体符号的可检测标记的发射波长强度的降低。任选地,可以在将珠粒暴露于第四切割步骤之前进行洗涤步骤。

[0196] 最后,第四切割步骤包括将珠粒暴露于含有5个单位的HindIII的溶液中,在37℃下在CUTSMART缓冲液中进行5-10分钟温育。温育后,在第四成像步骤中对珠粒进行成像,并且检测响应于HindIII核酸内切酶活性的荧光团可检测标记所传递的可检测信号响应。在此,可检测信号响应包括在第四切割步骤期间从分子条形码中解离的HindIII FNS标识体符号的可检测标记的波长相关的发射波长的信号强度降低。

[0197] 任选地,在每个循环的切割步骤期间可以引入多于一个限制性核酸内切酶,但核酸内切酶应在这样的条件下引入,即由此不同的核酸内切酶具有明显不同的动力学切割速率,并且在切割期间应获取多个图像。

[0198] 在去除最终的可切割标识体标记之后,不可切割的可检测标记将传递从珠粒发射的最强可检测信号,从而允许对其进行标识。在视觉反去卷积期间确定的每个FSN标识体条形码的荧光团的序列可以归因于以这种方式附接到其上的每个珠粒。

[0199] 实例3基于荧光团/蛋白酶位点/dsDNA(FPD)标识体的分子条形码

[0200] 实例3说明了基于荧光团/蛋白酶位点/dsDNA(FPD)标识体的分子条形码用于原位标记被设计成用于视觉解码以及在NGS反应环境中解码的材料的用途(图6、8、10、11)。可以使用由FPD标识体类别中的一或多种FPD标识体类别构成的集合以通过分裂-池化连接来编码和形成NGS相容性视觉分子条形码,以创建不需要基于DNA测序的解码的条形码的组合文库。如实例3中所示,可以在同一实验中并且在存在同时形成和编码的许多不同的分子条形码链的情况下产生单个分子条形码。如实例3中所用,组合文库中的每个分子条形码都含有五种FPD标识体类别(Id_{1-5}),所述五种FPD标识体类别包括四个可切割FPD标识体符号(Id_2、Id_3、Id_4和Id_5)以及一个不可切割FPD标识体符号(Id_1)。

[0201] 实验3为四循环正交蛋白酶切割实验,其是通过对反应性荧光团进行视觉成像证实的,被进行以验证基于FPD标识体的分子条形码的编码和解码。实验3包括第一成像步骤,所述第一成像步骤用于记录由存在于文库的任何成员(珠粒)上的基于FNS标识体的分子条形码的可检测标记所传递的可检测信号。在第一循环中引入TEV蛋白酶,以降低并入到分子条形码中的TEV蛋白酶反应性FPD标识体符号的荧光团可检测标记的信号强度,从而允许对从Id_5切割的标记进行视觉标识。此过程是循环重复的,由此在每次切割循环期间引入下一种正交蛋白酶,然后进行成像步骤。

[0202] 如实验3中所用,在切割步骤期间按以下顺序引入蛋白酶切割剂:循环1期间TEVp;循环2期间TVMVp;循环3期间SuMMVp;并且循环4期间TUMVp。每个循环之后的成像将实现并入到基于FPD标识体的分子条形码中的FPD标识体符号的可检测标记的三维布置去卷积。可以对充当切割剂的蛋白酶进行检测限、特异性和精度实验(一式三份),以建立与基于荧光团/蛋白酶位点/dsDNA标识体的分子条形码的蛋白酶正交反应性的再现性。

[0203] A.FPD标识体支架部分

[0204] 在一些实施例中,FPD标识体支架部分包括多肽片段,所述多肽片段使用氨基反应性异双功能交联试剂(NHS-PEG_4-甲基四嗪)通过单链DNA(ssDNA)中的经内部氨基修饰的碱基与具有3'和5'端的ssDNA片段共价连接。含有肽识别序列的多肽片段可以被修饰为含有N末端赖氨酸残基,并且随后可以利用化学基团进行修饰,所述化学基团适于使用不同的氨基反应性异双功能交联试剂(NHS-PEG_4-反式环辛烯)附接到经修饰的ssDNA。在将肽片段与经修饰的ssDNA寡核苷酸缀合之后,特异性且互补的ssDNA寡核苷酸(具有与附接到肽片段但处于反向互补朝向的ssDNA寡核苷酸相同的信息)可以与支架部分的ssDNA杂交,以产生具有未配对的3'端的dsDNA物种,所述未配对的3'端将适于与其它FPD标识体类别连接,以通过分裂-池化构建标识体的组合链。向基于FPD标识体的分子条形码的每个轮次的添加将添加用于形成分子条形码的唯一的标识体符号。附接到肽片段(C末端)的相对端是如下所述的可检测标记(荧光团)。附接到每个FPD标识体的可检测标记的发射波长(观察到的颜色)与所附接的dsDNA的序列相关,由此将可以从视觉条形码获得的信息与可以在DNA测序之后从所形成的NGS条形码获得的信息偶联。

[0205] 如实验3中所用,Id_1是添加到链的第一标识体,并且因此直接附接到固体载体(珠粒);在此Id_1不包含切割位点,并且其识别部分可以为不可切割的接头。Id_1包括用于附接到固体载体的经5'氨基修饰的碱基,并且可以编码用于检索和下游NGS文库制备的5'常数区。Id_5包括用于从细胞裂解物或生物样品中捕获大分子的3'捕获序列。Id_2、Id_3、Id_4和Id_5的识别部分和切割位点为分别地TUMVp、SuMMVp、TVMVp和TEV蛋白酶的识别序列和切割位点。

[0206] B.FPD标识体可检测标记

[0207] 如实验3中所用,选择包括多个可区分的可检测标记的量子点荧光团可检测标记,用于对实例3中的FPD标识体进行组合标记。在不同的反应混合物(或板的孔)中对于每种类别特异性肽,传递可检测信号的不同发射波长的可检测标记可以共价附接到每个FPD标识体类别。

[0208] 在一些实施例中,可检测标记可以在其相应支架部分的相对端(C末端)处添加到Id_1、Id_2、Id_3、Id_4和Id_5,标识体的dsDNA部分与所述相对端缀合。组合文库中的基于任何所形成的FPD标识体的条形码(本文示出的Id_1、Id_2、Id_3、Id_4和Id_5)上展示的荧光团的组合可以传递或可以不传递具有连续发射光谱的可检测信号,因为并入到构成组合文库的分子条形码中的FPD标识体符号的可检测标记的三维布置是通过如上所述的正交蛋白酶的循环确定的。

[0209] C.FPD标识体级联

[0210] 每个ssDNA片段的序列被设计成记录附接的可检测标记的发射波长。通过杂交形成构成每个FPD标识体支架部分的互补的ssDNA链,由此两条链共有显著互补性,但在其3'端保持不配对。每种FPD标识体类别的未配对的3'端被设计成与并入到分子条形码中的相邻FPD标识体符号上的受体3'端互补。FPD标识体可以按预先限定的顺序依序添加,以经连续轮次的分裂和池化形成分子条形码。每个轮次的FPD标识体添加都将使得能够通过杂交的标识体的ssDNA片段的5'和3'端的酶连接来进行FPD标识体dsDNA的级联。每个标识体的连接可以在1X连接酶缓冲液(pH 7.5的50mM Tris-HCl、10mM MgCl2、1mM ATP、10mM DTT)中,在存在500-1,000U T4 DNA连接酶和20-80U多核苷酸激酶的情况下,经37℃下的25分钟温育进行。最终"封盖"FPD标识体符号的dsDNA可以被设计成编码FPD标识体的可检测标记的波长以及被设计成用于从细胞裂解物中捕获大分子的3'捕获区。唯一的分子标识体(UMI)也可以作为随机化的或半随机化的碱基的连续延伸段而被包含在封盖FPD标识体中,或可以添加到NGS条形码中,其中每种标识体类别位于随机化的或半随机化的碱基的较小延伸段中。

[0211] D.FPD标识体的制备

[0212] 被设计成包括单个蛋白酶识别位点的标识体类别特异性寡肽可以从商业供应商处购买。

[0213] 如实验3中所用,Id_1被配置成不可切割,并且Id_{2-5}的识别部分包括肽序列,所述肽序列分别包含:Id_2,其包括KGGSGGGSACVYHQSGGSGGSC(SEQ ID NO:9),所述序列含有TUMV切割位点;Id_3,其包括KGGSGGGSEEIHLQSGGSGGSC(SEQ ID NO:10),所述序列含有SuMMV切割位点;Id_4,其包括KGGSGGGSETVRFQGGGSGGSC(SEQ ID NO:11),所述序列含有TVMV切割位点;以及Id_5,其包括KGGSGGGSENLYFQSGGSGGSC(SEQ ID NO:12),所述序列含有TEV切割位点。

[0214] 如实验3中所用,每种FPD标识体类别都是如下产生的:通过首先使用NHS-PEG_4-反式环辛烯对肽进行修饰(通过安装在在肽的N末端处的赖氨酸残基),使得其含有用于下游缀合的相容化学。为此,将100μM肽与500μM NHS-PEG_4-反式环辛烯(TCO)在NHS肽反应缓冲液(100mM磷酸钠,pH 8.5,补充有80mM KCl和70mM NaCl)中混合,并且允许在室温下反应过夜。使用HPLC以去除过量的(和任何未经反应的)NHS-PEG_4-TCO试剂来对经TCO修饰的肽进行纯化。在NHS相容性粘接缓冲液(100mM磷酸钠缓冲液,pH 8.5,补充有80mM KCl)中,通过

与500μM NHS-PEG4-甲基四嗪在室温下过夜温育,对含有内部氨基修饰并携带5'磷酸基团的100μM寡核苷酸进行修饰。然后使用7K MWCO Zeba脱盐柱以去除过量的和任何未经反应的NHS-PEG4-甲基四嗪试剂将经甲基四嗪修饰的寡核苷酸缓冲液交换到NHS相容性粘接缓冲液中两次。通过各自在25μM下以1:1比率混合,将每种FPD标识体类别的经TCO修饰的肽与不同的经四嗪修饰的ssDNA寡核苷酸(其各自含有对应于与标识体缀合的可检测标记的不同核苷酸序列)缀合。允许此反应在室温下进行30分钟。通过与缀合物以1:1摩尔比在NHS相容性粘接缓冲液(100mM磷酸钠缓冲液,pH 8.5,补充有80mM KCl)中在热块上在95℃下温育5分钟将互补的ssDNA寡核苷酸粘接以形成dsDNA。然后将热块移除,并且允许在工作台上经相当于约2小时的时间段冷却至室温。Id_1类别成员接收互补的寡核苷酸链,所述互补的寡核苷酸链含有用于化学修饰和随后附接到固体载体的5'氨基修饰。所有互补的ssDNA寡核苷酸都含有用于能够在条形码形成期间与相邻的标识体的有效连接的5'磷酸修饰。然后用经马来酰亚胺修饰的荧光团通过其C末端处的安装的半胱氨酸残基,通过将每种荧光团添加到每种FPD标识体类别(每种FPD标识体类别接收一个指定荧光团,所述指定荧光团被记录在所附接的dsDNA内)来对粘接的标识体进行标记。然后将FPD标识体缓冲液交换到储存缓冲液(50mM Tris-HCl,pH 7.5,补充有100mM NaCl)中,并且可以立即使用,或可以储存在4℃下持续几天或储存在-20℃下,以便长期储存。

[0215] E.基于FPD标识体的分子条形码验证

[0216] 对FPD标识体分子条形码的标识可以使用通过切割实验进行的四循环去卷积/解码完成,以证实可检测信号响应,并证实FPD标识体分子条形码的相应形成和编码。在一些实施例中,基于FDP标识体的分子条形码文库可以形成于珠粒上,并且珠粒可以固定在表面上并且在与实验溶液接触之前和之后成像。首先将使用配置有适当激发波长和发射滤光片的标准荧光显微镜对所固定的珠粒进行成像,以记录给定视场或所关注区中构成每个条形码的视觉可检测标记的组合。

[0217] 如实验3中所用,第一切割步骤包括在第一切割步骤中使珠粒暴露于含有2个单位的TEVp的溶液中,以在30℃下在蛋白酶反应缓冲液(pH 7.5的50mM Tris-HCl、0.5mM EDTA、1mM DTT)中进行30-60分钟温育。温育后,在第一成像步骤中对珠粒进行成像,并且检测传递对TEVp的可检测信号响应的荧光团可检测标记。在此,可检测信号响应包括在第一切割步骤期间从分子条形码中解离的FPD标识体符号的可检测标记的发射波长强度的降低。任选地,可以在第二切割步骤之前进行洗涤步骤。

[0218] 如实验3中所用,第二切割步骤包括将珠粒暴露于含有2个单位的TVMVp的溶液中,以在30℃下在蛋白酶反应缓冲液中进行30-60分钟温育。温育后,在第二成像步骤中,对珠粒进行成像,并且检测响应于TVMVp的荧光团可检测标记。在此,可检测信号响应包括在第二切割步骤期间从分子条形码中解离的FPD标识体符号的可检测标记的发射波长强度的降低。任选地,可以在第三切割步骤之前进行洗涤步骤。

[0219] 如实验3中所用,第三切割步骤包括将珠粒暴露于含有2个单位的SuMMVp的溶液中,以在30℃下在蛋白酶反应缓冲液中进行30-60分钟温育。温育后,在第二成像步骤中,对珠粒进行成像,并且检测传递对SuMMVp的可检测信号响应的荧光团可检测标记。在此,可检测信号响应包括在第三切割步骤期间从分子条形码中解离的FPD标识体符号的可检测标记的发射波长强度的降低。任选地,可以在第三切割步骤之前进行洗涤步骤。

[0220] 如实验3中所用,第四切割步骤包括将珠粒暴露于含有2个单位的TUMVp的溶液中,以在30℃下在蛋白酶反应缓冲液中进行30-60分钟温育。温育后,在第四成像步骤中,对珠粒进行成像,并且检测传递对TUMVp的可检测信号响应的荧光团可检测标记。在此,可检测信号响应包括在第四切割步骤期间从分子条形码中解离的FPD标识体符号的可检测标记的发射波长强度的降低。

[0221] 在通过从可切割标识体中切割最终的可检测标记去除之后,不可切割的标记将保持为从珠粒传递的最强可检测信号,从而允许对其进行标识。在视觉去卷积期间确定的每个分子条形码的荧光团的序列呈现在NGS条形码的序列中,所述序列可以在NGS实验后检索并且与珠粒相关。

[0222] F.实例3说明

[0223] 在一些实施例中,每一种FPD标识体类别的识别部分都包括实质上相同的蛋白酶识别位点。

[0224] 在一些实施例中,标识体类别大小基于多少唯一的可检测标记是可用的。在Q点的情况下,此数字为约10,而在标准荧光团的情况下,此数字为约四至约六。双标记使类别大小为100,因为每个类别成员都可以用两个Q点来标记。这也可以使用标准荧光团的组合来完成。

[0225] 在一些实施例中,通过首先通过肽的N末端和位于DNA序列中间的内部核苷酸修饰将不同的ssDNA链与同一肽(其具有相同的蛋白酶识别位点,并且在所有情况下相同的缀合化学)缀合来构建由一或多个FPD标识体构成的集合的结构部分。这可以在不同的管或孔(例如孔1-10)中完成。然后(在FPD标识体的结构部分仍分隔在1-10个不同的管/孔中的情况下),将结构部分洗涤或以其它方式纯化,并且将仅一种Q点类型附接到结构部分中的每个结构部分,并且由此构成FPD标识体类别(通过肽的C末端)。FPD标识体类别构建可以在不同的反应中进行,以产生另外的FPD标识体类别,以包含在由FPD标识体构成的集合中。这被设计成使得知道哪个ssDNA与哪个Q点搭配。然后(在FPD标识体类别仍分隔在1-10个不同的管/孔中的情况下),将FPD标识体类别洗涤或以其它方式纯化,并且将互补的ssDNA与由FPD标识体构成的集合中的每个标识体杂交(互补的ssDNA被配置成与由FPD标识体构成的集合中的标识体的结构部分的ssDNA互补)以形成具有3'未配对的端部的dsDNA。本领域的普通技术人员将理解,互补的ssDNA将不与错误的FPD标识体类别连接,由此向用户提供对如何配置和构造FPD标识体类别的某种程度的控制。在一些实施例中,用户可以将FPD标识体类别配置成具有反应特异性,在所述反应特异性的情况下,例如,类别B将基本上仅与类别A和C连接;类别3将基本上仅与类别B和D相连,以此类推。在一些实施例中,FPD标识体的构建包括首先将肽与Q点缀合,之后添加预杂交的dsDNA段以完成结构部分并形成FPD标识体。

[0226] 在一些实施例中,FPD标识体构建过程可以对每种FPD标识体类别(每个含有不同蛋白酶识别位点的肽)进行重复,并且FPD标识体类别然后可以被布置成用于在管中或在一个板或多个板中储存,这取决于每种类别的大小的类别的数量。例如,如实例3中所示,五种包括十个标识体的FPD标识体类别各自可以置于板的50个不同的孔中。因此,在制备包括珠粒和由一或多个分子条形码构成的集合的组合文库时,珠粒将具有在10个不同的孔中附接到其的第一FPD标识体类别。然后可以将珠粒洗涤、池化并随机地分裂到接下来的10个孔中,以添加第二FPD标识体类别,从而提供对文库的组合多样性的覆盖(现在为跨2种类别呈

现的100个不同成员的所有可能组合)。然后重复以上过程;可以将珠粒洗涤、池化并随机地分裂到接下来的10个孔中,以添加第三FPD标识体类别,从而产生文库的跨3种类别呈现的1,000个成员。

[0227] 在一些实施例中,重复FPD标识体构建过程,直至添加了总共5种FPD标识体类别——从而构成1M成员组合文库。重要的是,实例3中的最终"封盖"FPD标识体类别包括用于从反应混合物(例如细胞裂解物)中捕获大分子的3'捕获区。此3'捕获区可以是用于捕获聚腺苷酸化的RNA的poly(T)标签,或用于捕获在测定中用作报告基因的核苷酸标签的大分子的特定标签。

[0228] 在解码包括组合文库的所形成且经编码的分子条形码方面,在第一切割步骤之前,从并入到分子条形码中的每个标识体符号传递的可检测信号的发射光谱将组合出现,其中出现多个重叠光谱(类似于彩虹)。因此,在第一切割步骤之前,并入到分子条形码中的FPD标识体符号的三维布置是不可解码的。然而,已知FPD标识体类别被添加到组合文库中的顺序,三维布置允许通过切割步骤一次一个蛋白酶并且以已知的顺序循环蛋白酶来解码分子条形码,因为每种切割剂蛋白酶将基本上仅切割源自FPD标识体类别的FPD标识体符号的可检测标记,所述可检测标记包括与切割剂具有正交反应性的识别部分。

[0229] 换言之,分子条形码解码包括检测在暴露于对应蛋白酶期间哪些可检测信号响应(在本实施例中颜色)被传递。在一些实施例中,当在暴露于第一蛋白酶之后成像时,彩虹颜色的珠粒将丢失整个颜色,或者对于给定颜色的强度将显著降低,然后将洗涤并添加下一种蛋白酶,并且在进行成像时,第二种颜色将丢失或强度将显著降低。

[0230] 实例4基于荧光团/核酸酶位点/dsDNA(FND)标识体的分子条形码

[0231] 实例4说明了基于荧光团/核酸酶位点/dsDNA(FND)标识体的分子条形码用于原位标记被设计成用于视觉解码以及在下一代测序(NGS)反应环境中解码的材料的用途(图5、7、8、9、11)。可以使用一系列FND标识体类别以通过分裂-池化连接来编码和形成NGS相容性视觉分子条形码,以创建不需要基于DNA测序的解码的条形码的组合文库。如实例4中所示,可以在同一实验中并且在存在同时形成和编码的许多不同的分子条形码的情况下产生单个分子条形码。

[0232] 本领域的普通技术人员将理解,许多核酸内切酶相对于彼此具有正交反应性。例如,NotI、XhoI、SpeI和HindIII对分别含有GCGGCCGC、CTCGAG、ACTAGT和AAGCTT的dsDNA序列具有特异性。由此,NotI不应高效地特异性切割XhoI、SpeI或HindIII所识别的位点。换言之,每种核酸内切酶都应能够在存在其它核酸内切酶所识别的底物的情况下在具有包括相应的核酸内切酶切割位点序列的识别部分的标识体的切割位点处切割分子条形码,由此促进在基于FND标识体的分子条形码解码期间核酸内切酶切割剂的正交反应性。

[0233] 如实例4所用,包括珠粒和由一或多个基于FND标识体的分子条形码构成的集合的组合文库包含五个FND标识体类别(Id_{1-5}),其中Id_1被配置为不可切割的标识体类别,并且Id_{2-4}被配置为可切割标识体类别。实验4为四循环正交核酸酶切割实验,其是通过对反应性荧光团进行视觉成像证实的,被进行以验证在珠粒上对基于FND标识体的分子条形码的编码和解码。

[0234] 实验4包括四个循环,每个循环包括一或多个切割步骤以及之后的一或多个成像步骤。在切割步骤期间,正交切割剂核酸酶可以一次一个应用于组合文库的珠粒,以促进珠

粒上的所形成且经编码的分子条形码的正交切割。可以进行初步成像步骤以检测从并入到存在于珠粒中的任何珠粒上的分子条形码中的FND标识体符号的可检测标记传递的可检测信号响应。

[0235] 如实验4中所展示的,在第一切割步骤期间的第一循环中引入NotI,以降低通过与NotI核酸酶具有反应性的FND标识体符号(在这种情况下为Id_5符号)的可检测标记传递的可检测信号响应的强度,从而允许解码或视觉标识通过切割从Id_5符号解离的可检测标记。如实验4中所展示的,在切割步骤的一或多个切割步骤期间按以下顺序引入正交反应性切割剂核酸酶:循环1期间NotI;循环2期间XhoI;循环3期间SpeI;并且循环4期间HindIII。在成像步骤中的一或多个成像步骤期间对从解离的可检测标记传递的可检测信号响应的检测有助于解码并入到分子条形码中的FND标识体符号的三维布置。可以对充当切割剂的核酸酶进行检测限、特异性和精度实验(一式三份),以建立与基于FND标识体的分子条形码的核酸酶正交反应性的再现性。

[0236] A.FND标识体支架组成

[0237] 在一些实施例中,FND标识体支架部分包括具有单个核酸内切酶限制序列的一或多个拷贝的预杂交的双链DNA(dsDNA)。具有核酸内切酶识别序列的预杂交的双链dsDNA的一条链承载5'氨基修饰。然后可以使用氨基反应性异双功能交联试剂NHS-PEG_4-TCO,以化学修饰预杂交的dsDNA上的伯胺,以含有点击相容性TCO基团。预杂交的dsDNA通过ssDNA中的经内部氨基修饰的碱基与具有3'和5'端的单链DNA(ssDNA)片段共价连接,所述经内部氨基修饰的碱基可以使用氨基反应性异双功能交联试剂(NHS-PEG_4-甲基四嗪)转化为点击相容性甲基四嗪基团。在将与杂交的dsDNA片段与经修饰的ssDNA寡核苷酸缀合之后,特异性且互补的ssDNA寡核苷酸(具有与附接到dsDNA片段但处于反向互补朝向的ssDNA寡核苷酸相同的信息)并且可以与标识体的ssDNA杂交,以产生具有未配对的3'端的dsDNA物种,所述未配对的3'端将适于与其它FND标识体类别连接,以促进通过分裂-池化构建标识体的组合链。

[0238] 在一些实施例中,将包括至少一个可检测标记(在本例中为荧光团)的标记的ssDNA链粘接到含有特定核酸内切酶位点的所附接的ssDNA链。在一些实施例中,标记的ssDNA链在其相对的5'和3'端被双标记以包含四个可检测标记。在一些实施例中,双标记的ssDNA包括四个不同的可检测标记,以提供16种FND标识体类别,每个可检测标记具有唯一的可检测信号响应。在一些实施例中,可检测标记以相对于彼此间隔开的距离操作性地连接到支架部分,以减少可检测标记之间的荧光猝灭或其它非特异性反应。

[0239] 如图14所示,标记的ssDNA链在其相对的5'端和3'端上被双标记,并且相关联的dsDNA双链体的序列含有同一类型的两个限制性位点。每个FND标识体的可检测标记的发射波长(观察到的颜色)与所附接的dsDNA(能够连接)的序列相关,由此将可以从分子条形码获得的信息与可以在DNA测序之后从形成的NGS条形码获得的信息偶联。为避免交叉反应,构成所形成的NGS条形码的所设计的序列不应含有可以通过所使用的任何限制性酶切割的序列。Id_1为所添加的第一标识体,并且因此直接附接到固体载体(珠粒)。在此,Id_1不包含切割位点,并且其识别部分为不可切割接头。Id_1包括用于附接到固体载体的经5'氨基修饰的碱基,并且可以编码用于检索和下游NGS文库制备的5'常数区。Id_5的结构部分含有从细胞裂解物或生物样品中捕获大分子的3'捕获序列。Id_2、Id_3、Id_4和Id_5的识别部分和切割位点

为分别地HindIII、SpeI、XhoI和NotI核酸内切酶的识别序列和切割位点。

[0240] B.FND标识体可检测标记

[0241] 在本实例中,选择包括多个可区分标记的可商购获得的荧光团以对FND标识体进行组合标记。在一优选实施例中,FND标识体的可检测标记包括一或多个Q点。在一些实施例中,不同发射波长的荧光团可以共价附接到携带特定核酸酶位点类型的每个类别特异性dsDNA的不同反应混合物(或板的孔)中的每种标识体类别。由五个FND标识体类别类型构成的单个分子条形码在实例4中展示。可以将荧光团在粘接到附接到构成标识体条形码的NGS部分的有连接能力的dsDNA段的链的ssDNA链的任一端(示出了3'端)处添加到Id_1、Id_2、Id_3、Id_4和Id_5。在组合文库(本文示出的Id_1、Id_2、Id_3、Id_4和Id_5)中的任何所形成的单个基于FND标识体的分子条形码上显示的荧光团的组合可以具有或可以不具有连续发射光谱,因为并入到分子条形码中的FND标识体符号的三维布置是通过循环使用如本文所述的正交反应性核酸酶的切割步骤确定的。

[0242] C.FND标识体级联

[0243] 每个ssDNA片段的序列被设计成记录发射波长,即从并入到分子条形码中的FND标识体符号的解离的可检测标记传递的可检测信号响应。通过杂交形成每个FND标识体的互补的ssDNA链,由此两条链共有显著互补性,但在其3'端保持不配对。在一些实施例中,级联的FND标识体符号的未配对的3'端被配置成与相邻FND标识体符号的受体3'端互补。FND标识体可以按预先限定的顺序依序添加,以经连续轮次的分裂和池化形成分子条形码。每个轮次的FND标识体类别添加都将使得能够通过杂交的标识体的ssDNA片段的5'和3'端的酶连接来进行标识体dsDNA的级联。每个标识体的连接可以在1X连接酶缓冲液(pH 7.5的50mM Tris-HCl、10mM MgCl2、1mM ATP、10mM DTT)中,在存在500-1,000U T4 DNA连接酶和20-80U多核苷酸激酶的情况下,经37℃下的25分钟温育进行。在一些实施例中,最终的"封盖"FND标识体类别包括dsDNA(构成条形码的NGS部分),以编码附接到标识体多肽蛋白酶识别部分的标记的波长,以及被设计成从细胞裂解物中捕获大分子的3'捕获区。唯一的分子标识体(UMI)也可以作为随机化的或半随机化的碱基的连续段延伸而被包含在封盖FND标识体类别内,或可以添加到NGS条形码中,其中每种标识体类别位于随机化的或半随机化的碱基的较小延伸段中。为了避免交叉反应性,构成所形成的NGS条形码的所配置的序列,包含构成UMI的那些序列,基本上不含有可以通过所使用的任何限制性酶切割的序列。在一些实施例中,可以使用所配置的分子条形码的计算过滤以避免包含用于通过切割实验解码的限制性位点之一。FND标识体条形码含有5个通过酶连接进行连接的类别成员。

[0244] D.FND标识体类别制备

[0245] NotI、XhoI、SpeI和HindIII对分别含有GCGGCCGC、CTCGAG、ACTAGT和AAGCTT的dsDNA序列具有特异性。构成含有一或多种限制性核酸内切酶位点类型的标记的dsDNA的每个FND标识体类别特异性ssDNA可以从商业供应商处购买。如实验4中所用,FND标识体类别的识别部分和切割位点包括核苷酸序列,所述核苷酸序列分别被配置为:(Id_1不可切割);Id_2包括/5AmMC6/ATTTATTTAAGCTTATTATTATTT(SEQ ID NO:13),所述序列含有HindIII切割位点;Id_3包括

[0246] /5AmMC6/ATTTATTTAACTAGTATTATTATTT(SEQ ID NO:14),所述序列含有SpeI切割位点;Id_4包括/5AmMC6/ATTTATTTACTCGAGATTATTATTT(SEQ ID NO:15),所述序列含有XhoI

切割位点;并且Id_5包括/5AmMC6/ATTTATTTAGCGGCCGCATTATTATTT(SEQ ID NO:16),所述序列含有NotI切割位点。

[0247] 在一些实施例中,氨基修饰仅在构成附接到珠粒的FND标识体支架部分的寡核苷酸之一的5'端上。因此,所有其它dsDNA FND标识体支架部分都含有内部氨基修饰,使得端部可用于连接侧接的FND标识体支架部分。例如,在一些实施例中,氨基修饰位于所有杂交FND标识体的5'端上,因为对于每个杂交双链体,一条链的5'端被生物素化,并且双链体中的另一条链含有通过经NHS修饰的荧光团附接的标记。在一些实施例中,氨基修饰包含在由环状或环形基于dsDNA的标识体构成的集合中的所有标识体的支架部分的内部位置内,使得自由端将可用于连接。

[0248] 如本文所用,"5AmMC6"是指氨基修饰剂C6。技术人员将理解,氨基修饰剂用于将伯氨基引入到寡核苷酸中。例如,氨基修饰剂可以与NHS酯或异硫氰酸酯荧光可检测标记结合使用。技术人员将理解,氨基修饰剂C6可以在寡核苷酸合成期间掺入,并且可以用于在六碳间隔子的末端处用伯氨基标记寡核苷酸的5'端。如本文所用,"N-羟基琥珀酰亚胺"(NHS)是指具有式$(CH_2CO)_2NOH$的有机化合物。技术人员将理解,与蛋白质可以在游离氨基上通过酯介导的衍生化非选择性地进行的方式类似,N-羟基琥珀酰亚胺酯或"NHS酯"可以用于位点特异性地修饰安装在含有所设计的位置处的此类基团的合成的寡核苷酸内的伯胺基(参见例如,南达(Nanda)等人,酶学方法(Methods Enzymol.),536:87-94(2014))。例如,技术人员将理解,NHS酯可以用于标记伯胺(R-NH2)或蛋白质、经胺修饰的核苷酸和其它含胺的分子。因此,如本文所用,"NHS荧光团"是指与NHS缀合的任何荧光团。

[0249] 在一些实施例中,通过首先将一条互补的标记的链粘接来产生FND标识体类别,所述互补的标记的链与上面列出的类别特异性ssDNA链(Id_{2-5})一起产生可行的dsDNA限制性位点。这可以以1:1摩尔比在NHS相容性粘接缓冲液(100mM磷酸钠缓冲液,pH值8.5,补充有80mM KCl)中在热块上在95℃下经5分钟完成,然后将热块移除,并且允许在工作台上经相当于约2小时的时间段冷却至室温。携带单个核酸内切酶位点类型的类别特异性dsDNA可以使用NHS-PEG_4-反式环辛烯进行修饰(通过以上序列中示出的5'氨基修饰),使得其含有用于下游缀合的相容化学。为此,100μM标记的dsDNA可以与500μM NHS-PEG_4-反式环辛烯(TCO)在NHS相容性粘接缓冲液(100mM磷酸钠缓冲液,pH 8.5,补充有80mM KCl)中混合,并且在室温下反应过夜。可以通过脱盐以去除过量的(和任何未经反应的)NHS-PEG_4-TCO试剂来对经TCO修饰的dsDNA双链体进行纯化。通过以1:1摩尔比在NHS相容性粘接缓冲液(100mM磷酸钠缓冲液,pH值8.5,补充有80mM KCl)中在热块上在95℃下经5分钟进行温育来将含有内部氨基修饰的100μM寡核苷酸首先粘接到互补的ssDNA中以形成编码条形码的NGS部分的dsDNA,然后将热块移除,并且在工作台上经相当于约2小时的时间段冷却至室温。

[0250] 如实验4中所用,Id_1类别NGS条形码部分接收互补的寡核苷酸链,所述互补的寡核苷酸链含有用于化学修饰和随后附接到固体载体的5'氨基修饰。所有互补的ssDNA寡核苷酸都包括用于能够在条形码形成期间与相邻的标识体的有效连接的5'磷酸修饰。在NHS相容性粘接缓冲液中通过与500μM NHS-PEG4-甲基四嗪在室温下过夜温育对也携带5'磷酸基的经内部氨基修饰的寡核苷酸进行修饰。然后使用7K MWCO Zeba脱盐柱以去除过量的和任何未经反应的NHS-PEG4-甲基四嗪试剂来将经甲基四嗪修饰的寡核苷酸缓冲液交换到NHS相容性粘接缓冲液中两次。通过各自在25μM下以1:1比率混合,将每种类别的经TCO修饰的

dsDNA与不同的经甲基四嗪修饰的ssDNA寡核苷酸(其各自含有对应于已经与标识体缀合的可检测标记的不同核苷酸序列)缀合。可以允许此反应在室温下进行30分钟。然后可以将缀合的FND标识体缓冲液交换到储存缓冲液(50mM Tris-HCl,pH 7.5,补充有100mM NaCl)中,并且可以立即使用,可以储存在4℃下持续几天或储存在-20℃下以便长期储存。

[0251] E.基于FND标识体的分子条形码验证

[0252] 对基于FND标识体的分子条形码的标识可以使用通过切割实验进行的四循环去卷积/解码来完成,以证实从通过切割从分子条形码解离以解码其经编码的三维布置的可检测标记传递的可检测信号响应。在一些实施例中,分子条形码组合文库可以形成于珠粒上,并且珠粒可以固定在表面上并且在与实验溶液接触之前和之后成像。首先将使用配置有适当激发波长和发射滤光片的标准荧光显微镜对所固定的珠粒进行成像,以记录给定视场或所关注区中构成每个条形码的视觉可检测标记的组合。通常,本文列出的核酸内切酶的"高保真"版本都可以在相同的缓冲液(可从美国马萨诸塞州01969罗利纽伯里波特转盘428的新英格兰生物实验室公司商购获得的1X CUTSMART缓冲液)中以100%效率起作用。

[0253] 如实验4中所用,在第一切割步骤中,将珠粒暴露于含有5个单位NotI-HF核酸内切酶的溶液中,以在37℃下在CUTSMART缓冲液(50mM乙酸钾、20mM三乙酸盐、10mM乙酸镁、100ug/ml BSA;缓冲液pH在25℃下为7.9)中进行5-10分钟温育。温育后,在第一成像步骤中,对珠粒进行成像,并且检测响应于NotI核酸内切酶活性的可检测标记。在此,可检测信号响应包括在第一切割步骤期间从分子条形码中解离的FND标识体符号的可检测标记的发射波长强度的降低。任选地,可以在第二切割步骤之前进行洗涤步骤。

[0254] 如实验4中所用,第二切割步骤包括将珠粒暴露于含有5个单位的XhoI-HF的溶液,以在37℃下在CUTSMART缓冲液中进行5-10分钟温育。温育后,在第二成像步骤中,对珠粒进行成像,并且检测响应于XhoI核酸内切酶活性的荧光团可检测标记。在此,可检测信号响应包括在第二切割步骤期间从分子条形码中解离的FND标识体符号的可检测标记的发射波长强度的降低。任选地,可以在第三切割步骤之前进行洗涤步骤。

[0255] 如实验4中所用,第三切割步骤包括将珠粒暴露于含有5个单位的SpeI-HF的溶液中,以37℃下在CUTSMART缓冲液中进行5-10分钟温育。温育后,在第三成像步骤中,对珠粒进行成像,并且检测响应于SpeI核酸内切酶活性的荧光团。在此,可检测信号响应包括在第三切割步骤期间从分子条形码中解离的FND标识体符号的可检测标记的发射波长强度的降低。任选地,可以在第四切割步骤之前进行洗涤步骤。

[0256] 如实验4中所用,第四切割步骤包括将珠粒暴露于含有5个单位的HindIII的溶液中,以在37℃下在CUTSMART缓冲液中进行5-10分钟温育。温育后,在第四成像步骤中,对珠粒进行成像,并且检测响应于HindIII核酸内切酶活性的荧光团。在此,可检测信号响应包括在第四切割步骤期间从分子条形码中解离的FND标识体符号的可检测标记的发射波长强度的降低。

[0257] 在一些实施例中,在通过切割进行的每个循环的解码中可以引入多于一个限制性核酸内切酶,但这些核酸内切酶必须在这样的条件下引入,即由此不同的核酸内切酶具有明显不同的动力学切割速率,并且在切割期间应获取多个图像。

[0258] 在去除最终的可切割标识体标记之后,不可切割标记将保持是从珠粒发射的最强信号,从而允许对其进行检测和标识。在视觉去卷积期间确定的每个条形码的荧光团的序

列呈现在NGS条形码的序列中,所述序列可以在NGS实验后检索并且与珠粒相关。

[0259] 实例5通过对单标记和双标记FND标识体进行分裂-池化连接产生的分子条形码

[0260] 如实例5所公开的,产生单标记和双标记荧光团/核酸酶位点/dsDNA(FND)标识体类别,以通过分裂-池化连接编码和形成分子条形码,以创建分子条形码组合文库。进行通过对反应性荧光团可检测标记进行视觉成像证实的多循环正交核酸酶切割实验以验证对基于FND标识体的分子条形码的编码和解码。将已知的切割剂以已知顺序应用于核酸酶切割实验,以促进单标记和双标记FND标识体类别的去卷积。然后按颠倒时间顺序对图像进行分析,以观察伪"信号增益"对实时信号损失。

[0261] 从由五个FND标识体类别(FND Id_{1-5})构成的集合中产生基于单标记和双标记FND标识体的分子条形码组合文库。由五个经NHS修饰的荧光团构成的集合与如本文所述的经氨基修饰的碱基直接缀合。单标记FND标识体类别包括在切割时与单一标记相关的识别部分,而双标记FND标识体类别包括在切割时与两种标记相关的识别部分。例如,利用未混合体积的AF-750标记的基于HindIII的FND标识体构建的组合文库在通过HindIII切割时将损失750nm光谱,而用50/50比率的分别地AF-750标记的基于HindIII的FND标识体和ATTO-550基于HindIII的FND标识体构建的组合文库将在切割时损失550nm和750nm光谱,从而使可以在一个切割循环期间收集的互信息有效翻倍。

[0262] 图18为由五个FND标识体(FND ID_{1-5})构成的集合的文本表示,所述五个FND标识体具有配置有正交粘性端的支架部分、正交识别部分和切割位点以及经修饰的核苷酸。如图18所示,FND Id_{1-5}的每个支架部分都包括未标记的单链寡核苷酸,所述未标记的单链寡核苷酸被配置成与互补的单链标记的寡核苷酸粘接,以形成具有未配对的单链悬垂部分或"粘性端"的双链体寡核苷酸。在一些实施例中,未配对的粘性端可以包括1个至5个核苷酸。在一些实施例中,未配对的粘性端可以包括多于5个核苷酸。观察到,随着其核苷酸的数量增加,更多正交粘性端构型成为可能。在一些实施例中,所述连接为模板化连接。如本文所用,"模板连接"或"模板化连接"统称为特定于正确互补粘性端的连接反应。技术人员将理解,将可能使用少到仅一个碱基悬垂部,因为这是在本领域已知的许多现有试剂盒中通常连接到DNA上用于测序的NGS衔接子所使用的构型。在一些实施例中,增加的正交性可用于实现多个标识体类别的同时连接。因此,在一些实施例中,构成由标识体类别构成的集合的所有标识体类别可以在单个反应中连接在一起,因为所述集合中的每个粘性端对都被配置成彼此正交。例如,具有粘性端的被配置成粘接到Id_2标识体的粘性端并保持与Id_3标识体的粘性端正交的Id_1标识体允许将Id_1标识体仅与Id_2标识体选择性地连接。技术人员将理解,粘性端连接的过程促进具有互补粘性端的双链体寡核苷酸对的选择性连接。

[0263] A.FND标识体支架组成

[0264] 产生FND Id_{1-5}以进行两个正交核酸酶切割实验,从而从单标记和双标记FND标识体类别中产生组合文库。FND Id_1的支架部分包括未标记的SEQ ID NO:17,所述SEQ ID NO:17被配置成粘接到标记的SEQ ID NO:18。对应于SEQ ID NO:18的位置1至5的核苷酸包括粘性端,所述粘性端被配置成粘接到SEQ ID NO:19的粘性端退。位置1和25分别通过5'磷酸化进行修饰并且被修饰为3'氨基修饰剂C6 dT。

[0265] 在一些实施例中,支架部分被配置成具有没有肽或核苷酸序列的识别部分,所述识别部分与已知的切割剂反应以使其不可切割。在一些实施例中,不可切割的标识体可用

作最后可视化的标识体,因为一些切割剂随着空间位阻增加而变得不那么有效。例如,观察到在直接从固相,如链霉亲和素珠粒中切割dsDNA时,HindIII的效率降低。观察到,在某些情况下,在切割步骤后观察到单一标记时,不需要进行最后的切割步骤。这类似于具有标记的不可切割的最终标识体符号。例如,在本文所述的FND ID_{1-5}的情况下,在SpeI切割后观察到单一标记时,则不需要切割形成分子条形码的残余部分的HindIII标识体符号。因此,在一些实施例中,最后可视化的标识体可以被配置成是可切割的,因为在通过成像对其进行标识时,将不需要将其切割。

[0266] FND Id_2的支架部分包括未标记的SEQ ID NO:19,所述SEQ ID NO:19被配置成粘接到标记的SEQ ID NO:20。对应于SEQ ID NO:19的位置1至5的核苷酸包括粘性端,所述粘性端被配置成粘接到SEQ ID NO:18的粘性端。SEQ ID NO:19的位置13至18包括HindIII识别部分和切割位点。SEQ ID NO:19的位置1通过5'磷酸化进行修饰。对应于SEQ ID NO:20的位置1至5的核苷酸包括粘性端,所述粘性端被配置成粘接到SEQ ID NO:21的粘性端。SEQ ID NO:20的位置16-21包括HindIII识别部分和切割位点。SEQ ID NO:20的位置1和9分别通过5'磷酸化进行修饰并且被修饰为Int氨基修饰剂C6 dT。

[0267] FND Id_3的支架部分包括未标记的SEQ ID NO:21,所述SEQ ID NO:21被配置成粘接到标记的SEQ ID NO:22。对应于SEQ ID NO:21的位置1至5的核苷酸包括粘性端,所述粘性端被配置成粘接到SEQ ID NO:20的粘性端。SEQ ID NO:21的位置13至18包括SpeI识别部分和切割位点。SEQ ID NO:21的位置1通过5'磷酸化进行修饰。对应于SEQ ID NO:22的位置1至5的核苷酸包括粘性端,所述粘性端被配置成粘接到SEQ ID NO:23的粘性端。SEQ ID NO:22的位置16-21包括SpeI识别部分和切割位点。SEQ ID NO:22的位置1和9分别通过5'磷酸化进行修饰并且被修饰为Int氨基修饰剂C6 dT。

[0268] FND Id_4的支架部分包括未标记的SEQ ID NO:23,所述SEQ ID NO:23被配置成粘接到标记的SEQ ID NO:24。对应于SEQ ID NO:23的位置1至5的核苷酸包括粘性端,所述粘性端被配置成粘接到SEQ ID NO:22的粘性端。SEQ ID NO:21的位置13至18包括XhoI识别部分和切割位点。SEQ ID NO:21的位置1通过5'磷酸化进行修饰。对应于SEQ ID NO:24的位置1至5的核苷酸包括粘性端,所述粘性端被配置成粘接到SEQ ID NO:25的粘性端。SEQ ID NO:22的位置16至21包括XhoI识别部分和切割位点。SEQ ID NO:24的位置1和9分别通过5'磷酸化进行修饰并且被修饰为Int氨基修饰剂C6 dT。

[0269] FND Id_5的支架部分包括未标记的SEQ ID NO:25,所述SEQ ID NO:25被配置成粘接到标记的SEQ ID NO:26。对应于SEQ ID NO:25的位置1至5的核苷酸包括粘性端,所述粘性端被配置成粘接到SEQ ID NO:24的粘性端。SEQ ID NO:25的位置11至18包括NotI识别部分和切割位点。SEQ ID NO:25的位置1通过5'磷酸化进行修饰。对应于SEQ ID NO:26的位置11至18的核苷酸包括NotI识别部分和切割位点。SEQ ID NO:26的位置1通过5AmMC6T进行修饰。

[0270] 如本文所用,"3AmMC6T"是指3'氨基修饰剂C6 dT,这是一种可从集成DNA技术公司(IDT)(美国爱荷华市52241克拉尔维尔的商业公园1710)商购获得的经修饰的核苷酸。如本文所用,"5'生物素-TEG"或"5生物素Teg"统称为附接到15个原子的混合极性三乙基乙二醇间隔子的可作为可以由IDT公司(IDT)在合成期间安装的修饰的生物素分子。技术人员将理解,5'生物素-TEG可以并入在寡核苷酸的5'或3'端。如本文所用,"5Phos"是指5'磷酸化,例

如寡核苷酸的在其5'端处的磷酸化。技术人员将理解,如果使用寡核苷酸作为DNA连接酶的底物,则需要5'磷酸化。如本文所用,"5AmMC6T"是指5'氨基修饰剂C6 dT,这是一种可从集成DNA技术公司(美国爱荷华市52241克拉尔维尔的商业公园1710)商购获得的经修饰的核苷酸。

[0271] B.单标记和双标记FND标识体可检测标记组成

[0272] 对于单标记正交核酸酶切割实验,FND Id_{1-5}的可检测标记被相应地配置:FND Id_1是未标记的;FND Id_2包括单个AF-750;FND Id_3包括单个AF-647;FND Id_4包括单个ATTO-550;并且FND Id_5包括单个ATTO-488。如本文所公开的,FND Id_{1-5}的每个支架部分是通过NHS酯介导的衍生化标记的。使用AF-750以标记SEQ ID NO:20的位于位置9处的3'Int氨基修饰剂C6 dT。使用AF-647以标记SEQ ID NO:22的位于位置9处的Int氨基修饰剂C6 dT。使用ATTO-555以标记SEQ ID NO:24的位于位置9处的Int氨基修饰剂C6 dT。使用ATTO-488以标记SEQ ID NO:26的位于位置1处的5'氨基修饰剂C6 dT。

[0273] 对于双标记正交核酸酶切割实验,FND Id_{1-5}的可检测标记被相应地配置:FND Id_1是未标记的;FND Id_2包括大约50％ AF-750和大约50％ ATTO-550;FND Id_3包括大约50％ AF-647和大约ATTO-488;FND Id_4包括大约50％ AF-647和大约50％ ATTO-550;并且FND Id_5包括大约50％ ATTO-488和大约50％ AF-750。如本文所公开的,FND Id_{1-5}的每个支架部分是通过NHS酯介导的衍生化标记的。使用AF-750和ATTO-550以标记SEQ ID NO:20的位于位置9处的3'Int氨基修饰剂C6 dT。使用AF-647和ATTO-488以标记SEQ ID NO:22的位于位置9处的Int氨基修饰剂C6 dT。使用AF-647和ATTO-555以标记SEQ ID NO:24的位于位置9处的Int氨基修饰剂C6 dT。使用ATTO-488和AF-750以标记SEQ ID NO:26的位于位置1处的5'氨基修饰剂C6 dT。

[0274] 如本文所用,"ATTO-488"为最大吸收为501nm并且最大荧光为523nm的ATTO荧光染料。技术人员将理解,ATTO-488在480nm至515nm范的围内被更高效地激发。如本文所用,"ATTO-550"为最大吸收为554nm并且最大荧光为576nm的ATTO荧光染料。技术人员将理解,ATTO-550在540nm至565nm范的围内被更高效地激发。ATTO-488和ATTO-550可从ATTO-Tec有限公司(ATTO-Tec GmbH,锡根57074的Martinshardt 7(Martinshardt 7,57074Siegen);info@att-tec.com;产品编号:AD 488,和产品编号:AD550)。

[0275] 如本文所用,"AF-405"为Alexa Fluor 405染料,这是一种在401nm处具有激发峰并且在421nm处具有发射峰的发射蓝色的合成荧光团。如本文所用,"AF-647"为Alexa Fluor 647染料,这是一种具有适于594nm或633nm激光线的激发的远红外荧光染料。如本文所用,"AF-750"为Alexa Fluor 750染料,这是一种具有适于633nm激光线的激发或染料泵浦的激发的明亮的近红外荧光染料。AF-405、AF-647和AF-750可从赛默飞世尔科技公司(美国马萨诸塞州02451沃尔瑟姆第三大街168(168Third Avenue,Waltham,MA 02451,USA);AF-405的目录号为A30000;AF-647的目录号为A20006;AF-750的目录号为A20011)商购获得。

[0276] 在一些实施例中,AF-405、AF-647、AF-750、ATTO-488和ATTO-550可互换使用,以产生不同类别的标识体。技术人员将理解,可以使用任何NHS缀合的荧光团(NHS-荧光团)以通过NHS酯介导的衍生化来标记任何游离氨基。NHS缀合的荧光团可从本文提供的供应商处购买。

[0277] C.单标记和双标记FND标识体可检测标记构建

[0278] 首先,将FND Id_{1-5}(SEQ ID NO:18、SEQ ID NO:20、SEQ ID NO:22、SEQ ID NO:24、SEQ ID NO:26)的标记的寡核苷酸上的游离氨基(NH_2)用荧光团使用20μl 200μM寡核苷酸标记,并且重悬于pH为8.5的100mM NaPO4(磷酸钠缓冲液)中,并且然后与2μl 100mM NHS-荧光团(例如,重悬于无水二甲基甲酰胺(DMF)中的NHS-ATTO-488、NHS-ATTO-550、NHS-AF-647或NHS-AF-750)在过夜反应中在室温下混合。

[0279] 然后用80μl 2X杂交缓冲液(pH 7.5的20mM Tris、1M NaCl、1mM EDTA)将标记的寡核苷酸稀释至100μl体积,以淬灭任何剩余的未经反应的NHS基团。将100μl 40μM标记的寡核苷酸添加到如本文所述的其对应的未标记的寡核苷酸中并且与所述未标记的寡核苷酸混合(即,SEQ ID NO:17与SEQ ID NO:18混合;SEQ ID NO:19与SEQ ID NO:20混合;SEQ ID NO:21与SEQ ID NO:22混合;SEQ ID NO:23与SEQ ID NO:24混合;SEQ ID NO:25与SEQ ID NO:26混合),用于随后的粘接。

[0280] C.FND标识体支架部分粘接

[0281] 然后通过以下进行粘接:将所有FND Id_{1-5}支架部分寡核苷酸双链体暴露于热块中90℃下持续大约5分钟,并且然后将热块放置在实验室工作台上,直至其达到室温为止(大约1.5小时)。

[0282] 将10μl生物素化的FND Id_1支架部分寡核苷酸双链体在90μl杂交缓冲液中10倍稀释至200nM,并且将100μl每种差异标记的双链体与100μl 0.2mg/ml Dynabeads MyOne链霉亲和素T1珠粒(链霉亲和素珠粒)(可从美国马萨诸塞州02451沃尔瑟姆第三大街168的赛默飞世尔科技公司商购获得;目录号65601)混合,并且然后预洗涤3X并且通过上下移液快速重悬于杂交缓冲液中(在结合前),以提供1:1(体积:体积)的均匀涂层。在热块上在37℃下经30分钟偶尔轻敲管进行结合。然后将缀合的链霉亲和素珠粒拉动在磁铁上以去除未结合的支架部分寡核苷酸双链体。去除溶液,并且如下洗涤缀合的链霉亲和素珠粒:在300μl补充有0.01% Tween-20(可从美国密苏里州68178圣路易斯市的西格玛奥德里奇公司邮政信箱14508(Sigma-Aldrich,Inx.,PO Box 14508,St.Louis,MO 68178,USA)商购获得;Cas号:9005-64-5)的2X杂交缓冲液中洗涤3X;在100μl 1X T4 DNA连接酶缓冲液(可从美国马萨诸塞州01938-2723伊普斯威奇乡村公路240的新英格兰生物实验室公司(NEB)(New England Biolabs,Inc.(NEB),240 County Road,Ipswich,MA 01938-2723,USA)商购获得;info@neb.com)洗涤2X;并且然后重悬于1X T4 DNA连接酶缓冲液中。然后在水浴超声器中对缀合的链霉亲和素珠粒进行超声处理持续3分钟,以打破珠粒簇,并且储存在4℃下或冰上,用于下游连接反应。

[0283] 连接步骤是在室温或至多37℃的温度下,在存在T4多核苷酸激酶(PNK)(可从NEB公司商购获得;(5μl/250μl连接反应);目录号:M0201S或Mo201L)和T4 DNA连接酶(可从NEB公司商购获得;(5μl/250μl连接反应);目录号:M0202S、M0202T、M0202L或M0202M)两者的情况下在NEB公司提供的1X T4 DNA连接酶缓冲液中进行的。使用0.1mg/ml用本文所述的FND Id_1支架部分寡核苷酸双链体涂覆的链霉亲和素珠粒进行珠粒上连接(在NEB公司提供的100μl 1X连接缓冲液中洗涤2X),以附接FND Id_2支架部分寡核苷酸双链体。在相同条件下进行FND Id_{3-5}支架部分寡核苷酸双链体的所有随后连接。连接之间的洗涤为:用200μl补充有0.01% Tween-20的2X杂交缓冲液进行1次洗涤,然后用200μl 1X杂交缓冲液进行2次洗

涤,然后用NEB公司提供的100μl 1X连接缓冲液进行2次洗涤。

[0284] 在一些实施例中,可以在溶液中在与链霉亲和素珠粒连接之前连接FND标识体。例如,在通过FND Id_1与链霉亲和素珠粒连接之前,将由FND Id_2构成的集合与由FND Id_3构成的集合预先连接。这是通过以下进行的:将5μl FND Id_2支架部分寡核苷酸双链体与5μl FND Id_3支架部分寡核苷酸双链体(20μM储备液浓度)预混合,或将5μl FND Id_4支架部分寡核苷酸双链体与例如5μl FND Id_5支架部分寡核苷酸双链体、与25μl 10X连接酶缓冲液(NEB公司)、205μl H_2O以及5μl PNK(NEB公司)和5μl T4 DNA连接酶(NEB公司)预混合。这些经预混合的标识体连接是在介于室温与37℃之间的温度下进行的。然后将如上制备的5μl链霉亲和素珠粒(用Id1寡核苷酸双链体涂覆的0.2mg珠粒,所述珠粒用200μl补充有0.01%Tween-20的2X杂交缓冲液洗涤3X,然后用200μl 1X杂交缓冲液洗涤两次,然后用NEB公司提供的100μl 1X连接缓冲液洗涤另外两次,并且重悬于50μl连接缓冲液中)移液到新管的底部,以添加各种250μl连接反应。

[0285] 连接是在存在T4多核苷酸激酶(PNK)和T4 DNA连接酶的情况下进行的。在每个轮次的连接之间,使用0.5-1mM EDTA进行广泛洗涤,以淬灭任何剩余的连接酶活性,使用高盐(500mM-1M NaCl)以去除来自dsDNA双链体的连接酶,并且使用洗涤剂(0.01%Tween-20)以帮助从双链体中进行连接酶去除并且防止珠粒结块;所述珠粒结块可能导致在随后的连接期间珠粒涂层不均匀,因此为了进一步防止这种情况,在成像之前,并且在暴露于任何酶催化的反应之前,对珠粒进行超声处理持续至少3分钟。进一步的细节,请参加方法部分。

[0286] 用于编码的所有连接完成之后,将链霉亲和素珠粒用200μl补充有0.01% Tween-20的2X杂交缓冲液洗涤3X,用200μl 1X杂交缓冲液洗涤2X,重悬于20μl 1X杂交缓冲液中,超声处理3分钟,并且上样到含有经生物素修饰的表面的流动池中用于固定化和随后的切割实验。

[0287] D.结果

[0288] 编码基于单标记和双标记FND标识体的分子条形码

[0289] 如实例5中所示,通过连接差异标记的预杂交的寡核苷酸双链体构建基于FND标识体的分子条形码。使用组合连接策略,由5个段(即,每个分子条形码5个标识体符号)构建具有4个不同的可区分标记的分子条形码。如本文所公开的,将由FND标识体构成的集合通过对集合中的支架部分寡核苷酸双链体的寡核苷酸链之一进行的5'生物素修饰与链霉亲和素珠粒结合。(技术人员将理解,5'生物素修饰可以应用于寡核苷酸双链体的任一条链。)在此,与链霉亲和素珠粒结合的FND Id_1的支架部分寡核苷酸双链体是未标记的,并且产生了适于与FND Id_2的支架部分寡核苷酸双链体连接的5'悬垂部(磷酸化的)。

[0290] 如实例5中所公开的,FND Id_2的支架部分用AF-750标记,并且包括酶可接近的HindIII切割位点和适于与FND Id_3的支架部分寡核苷酸双链体连接的5'悬垂部(磷酸化的)两者。将FND Id_2与FND Id_1连接后,将这些链霉亲和素珠粒的样品固定在流动池中用于成像。在4个荧光通道中提取与含有FND Id_1和FND Id_2的分子条形码结合的链霉亲和素珠粒的图像,并且如图19中所示绘制强度值。图19A表示用通过与未标记的dsDNA连接进行附接的第一标记的dsDNA标识体段(Id1/HindIII-AF750)标记的链霉亲和素珠粒,所述未标记的dsDNA通过生物素部分附接到珠粒。图19B表示用第二标记的dsDNA标识体段(Id2/SpeI-AF647)标记的链霉亲和素珠粒。图19C表示用第三标记的dsDNA标识体段(Id3/XhoI-

ATTO550)标记的链霉亲和素珠粒。图19D表示用第四标记的dsDNA标识体段(Id4/NotI-ATTO488)标记的链霉亲和素珠粒。图20A表示在四个荧光通道中成像的未经切割的珠粒:647nm(向上斜条纹)、550nm(向下斜条纹)、488nm(横条纹)和750nm(虚线),并且在右侧绘制了针对每种荧光团获得的平均强度值。

[0291] 在含有FND Id_1/FND Id_2连接产物的珠粒上观察到清晰的AF-750标记检测(图19A),这表明连接反应是高效的,而在任何其它荧光通道中几乎没有观察到信号。与包括连接的(即,经编码的)FND Id_1和FND Id_2标识体符号的分子条形码结合的链霉亲和素珠粒然后进行另一轮次的连接,由此FND Id_3与正在生长的分子条形码连接(即,级联)。如实例5中所公开的,FND Id_3的支架部分用AF-647标记,并且包括酶可接近的SpeI切割位点和适于与FND Id_4的支架部分寡核苷酸双链体连接的5'悬垂部(磷酸化的)两者。将FND Id_3连接后,将珠粒的样品固定在流动池中用于成像。在4个荧光通道中提取与包括连接的(即,经编码的)FND Id_1/FND Id_2/FND Id_3/标识体符号的分子条形码结合的链霉亲和素珠粒的图像,并且绘制强度值(图19)。在与包括连接的FND Id_1/FND Id_2/FND Id_3/标识体符号的分子条形码结合的链霉亲和素珠粒上检测到AF-750和AF-647(图19B),这表明FND Id_3的连接成功。在其它两个荧光通道(488nm或550nm发射)中,在对与连接的FND Id_1/FND Id_2/FND Id_3/标识体符号结合的链霉亲和素珠粒进行成像时,未观察到信号。与包括连接的(即,经编码的)FND Id_1/FND Id_2/FND Id_3/标识体符号的分子条形码结合的链霉亲和素珠粒然后进行另一轮次的连接,由此FND Id_4与正在生长的分子条形码连接(即,级联)。如实例5中所公开的,FND Id_4的支架部分用ATTO-550标记,并且包括酶可接近的XhoI限制性位点和适于与FND Id_5的支架部分寡核苷酸双链体连接的5'悬垂部(磷酸化的)两者。将FND Id_4连接后,将链霉亲和素珠粒的样品固定在流动池中用于成像。在4个荧光通道中提取与包括连接的(即,经编码的)FND Id_1/FND Id_2/FND Id_3/FND Id_4/标识体符号的分子条形码结合的链霉亲和素珠粒的图像,并且绘制强度值(图19)。在与包括FND Id_1/FND Id_2/FND Id_3/FND Id_4/标识体符号的分子条形码结合的链霉亲和素珠粒上清晰地观察到AF-750、AF-647和ATTO-550(图19C),这表明FND Id_4的连接成功。在其余荧光通道(488nm发射)中,在对与包括FND Id_1/FND Id_2/FND Id_3/FND Id_4/标识体符号的分子条形码结合的链霉亲和素珠粒进行成像时,未观察到信号。与包括FND Id_1/FND Id_2/FND Id_3/FND Id_4/标识体符号的分子条形码结合的链霉亲和素珠粒然后进行另一轮次的连接,由此FND Id_5与正在生长的分子条形码连接(即,级联)。如实例5中所公开的,FND Id_5的支架部分用ATTO-488标记,并且包括酶可接近的NotI限制性位点(即,切割位点)。将FND Id_5连接后,将链霉亲和素珠粒的样品固定在流动池中用于成像。在4个荧光通道中提取与包括FND Id_1/FND Id_2/FND Id_3/FND Id_4/FND Id_5/标识体符号的分子条形码结合的链霉亲和素珠粒的图像,并且绘制强度值(图19)。在与包括FND Id_1/FND Id_2/FND Id_3/FND Id_4/FND Id_5/标识体符号的分子条形码结合的链霉亲和素珠粒上观察到所有四种标记(AF-750、AF-647、ATTO-550和ATTO-488)的信号(图19D),这表明FND Id_5的连接成功。这些结果表明,可以使用分裂和池化方法构建标识体的差异标记的链(即,分子条形码),以产生标记的珠粒的组合文库。图1中示出的连接的标识体链以及含有其它各种标记组合的其它链是以同样的方式制备的(在珠粒上产生),并且将珠粒固定在流动池的泳道中用于随后的解码实验。

[0292] 解码基于单标记和双标记FND标识体的分子条形码

[0293] 在一些实施例中,三维布置是线性的,具有一个或两个开口端。在一些实施例中,三维布置是环形的,不具有开口端。在一些实施例中,三维布置是发夹形式,具有环状端和开口端。技术人员将理解,两个或更多个dsDNA寡核苷酸的粘性端连接通常将形成段的线性链,每个段在一起并且在至少一个端部处连接,以形成链。

[0294] 图20B表示暴露于NotI酶之后珠粒的平均强度。图20C表示暴露于XhoI酶之后珠粒的平均强度。图20D表示暴露于SpeI酶之后珠粒的平均强度。图20A-D示出了对通过由一或多个FND标识体构成的集合的级联形成和编码在三维布置中的分子条形码的解码。如图20A-D所示,分子条形码是由连接的FND Id_{1-5}标识体符号和链霉亲和素珠粒的线性分段的链形成和编码的。如图20A-D所示,FND Id_1标识体符号是未标记的;FND Id_2标识体符号是用ATTO-550标记的并且编码酶可接近的HindIII识别部分;FND Id_3标识体符号是用AF-647标记的并且编码酶可接近的SpeI识别部分;FND Id_4标识体符号是用AF-750标记的并且编码酶可接近的XhoI识别部分;并且FND Id_5标识体符号是用ATTO-488标记的并且编码酶可接近的NotI限制性位点。为了进行解码,将携带所描述的标识体链(和标记组合)的链霉亲和素珠粒固定在流动池的表面上,并且在流动池的单个泳道内进行成像。在此解码实验中,任何给定视场(FOV)中的所有珠粒都携带相同的标识体链。在暴露于任何限制性酶之前,对流动池泳道中的链霉亲和素珠粒进行成像,由此观察到所有四种可检测标记的信号(图20A)。为了展示每个循环对一种标记的受控去除,首先使流动池泳道中的链霉亲和素珠粒与含有200U NotI酶的溶液(通过脱盐基本上去除了甘油)在新英格兰生物实验室公司提供的1X Cutsmart缓冲液(目录号B7204)中接触,并且在温度受控的热块室(脱离显微镜)中在37℃下温育30分钟。

[0295] 用NotI酶温育之后,用至少100μl 1X Cutsmart缓冲液冲洗流动池泳道,并且将流动池放回在显微镜台上。在所有4个荧光通道中获取同一流动池泳道的图像(来自不同FOV,因为本实验不需要跟踪单独珠粒)。成像时,从此通过正交切割进行的解码的第一循环中观察到NotI可切割的标记(ATTO-488)被清晰地去除(图20B)。在所有其它通道(550nm、647nm和750nm)中观察到的信号保持为高,这表明通过NotI进行的切割是特异性的,并且与标识体链内的其它经编码的限制性位点正交。然后从显微镜中去除流动池,并且使珠粒与含有400U XhoI酶的溶液(通过脱盐基本上去除了甘油)在新英格兰生物实验室公司(NEB)提供的1X Cutsmart缓冲液中接触,并且在温度受控的热块室(脱离显微镜)中在37℃下温育30分钟。用XhoI酶温育之后,用至少100μl 1X Cutsmart缓冲液冲洗流动池泳道,并且将流动池放回在显微镜台上。然后在所有4个荧光通道中获取同一流动池泳道的图像。成像时,从此通过正交切割进行的解码的第二循环中观察到XhoI可切割的标记(AF-750)被清晰地去除(图20C)。在这些图像中,链霉亲和素珠粒已经去除了FND Id_5和FND Id_4缔合的标记,这表明与NotI一样,XhoI酶也能够仅响应其经编码的限制性位点来进行特异性且正交的切割。然后从显微镜中去除流动池,并且使珠粒与含有300U SpeI酶的溶液(通过脱盐基本上去除了甘油)在新英格兰生物实验室公司(NEB)提供的1X Cutsmart缓冲液中接触,并且在温度受控的热块室(脱离显微镜)中在37℃下温育30分钟。用SpeI酶温育之后,用至少100μl 1X Cutsmart缓冲液冲洗流动池泳道,并且将流动池放回在显微镜台上。然后在所有4个荧光通道中获取同一流动池泳道的图像。成像时,观察到经受此通过正交切割进行的解码的第三循环的链霉亲和素珠粒去除了操作性地连接到FND Id_3的SpeI切割的AF-647(图20d),并且

仅观察到单个剩余标记(ATTO-550)的强信号。保留在珠粒上的最终标记可在图像中清晰地区分,并且因此不需要随后切割HindIII可切割标记(FND Id_2)。因此,标识体链的差异标记的组合可以通过分裂和池化连接策略进行编码,并且随后可以通过正交切割反应的循环以大量平行方式解码。

[0296] 解码具有单和双(经混合的)可检测标记的FND标识体符号

[0297] 如图21B所示,分子条形码包括一个未标记的FND标识体符号(FND Id_0)和四个双标记的FND标识体符号(FND Id_1-550/750、FND Id_2-488/647、FND Id_3-750/488和FND Id_4-488)。

[0298] 图21A呈现了条形图,所述条形图表示针对单标记dsDNA标识体链获得的OCS结果,其中循环的顺序是颠倒的以供分析。

[0299] 图21B呈现了条形图,所述条形图表示针对混合标记dsDNA标识体链获得的OCS结果,其中循环的顺序是颠倒的以供分析。

[0300] 图21A和21B分别是条形图,其示出了对包含具有单和双(经混合的)可检测标记的差异标记的FND标识体符号的分子条形码的解码。如图21A所示,对包括四个单标记FND标识体符号(使用FND标识体类别编码:FND Id_1-550、FND Id_2-750、FND Id_3-647和FND Id_4-488)的分子条形码进行了构建并且然后解码,以证明基于单标记FND标识体的分子条形码的组合文库理论容量。例如,为了构建由包括双标记的标识体符号的分子条形码构成的集合,将FND Id_2-488与FND Id_2-550以等摩尔比混合,以产生每个双标记标识体总共10种可视觉辨别的可检测标记的组合(而不是每个单标记标识体4种不同可能性)。混合用于构建双标记的标识体链的每个双标记FND标识体的可视觉辨别的组合为488/488、488/550、488/647、488/750、550/550、550-/647、550/750、647/647、647-/750和750/750。对于这些实验,如先前所述在链霉亲和素珠粒上构建(即,通过多个轮次的连接以及其间的洗涤)预先确定的链构型(即,预先确定的单标记或链中的每个标识体位置的处的标记的混合物)。然而,通过所有包括链的标识体段在单个连接反应内的溶液中连接构建,然后随后捕获到链霉亲和素珠粒上的单标记链在性能方面未显示出差异(数据未示出)。针对携带单标记和双标记标识体链的链霉亲和素珠粒,在每个循环时在所有四个荧光通道中获取在每个视场中跨多于100个珠粒得到的平均珠粒强度。获取未切割的链霉亲和素珠粒的图像,随后用NotI切割链霉亲和素珠粒(循环1),随后用XhoI切割链霉亲和素珠粒(循环2),并且随后用SpeI切割链霉亲和素珠粒(循环3)。然后按颠倒的时间顺序对图像进行分析。例如,将在循环3(通过SpeI切割)之后获取的图像作为系列中的第一图像进行分析;将在循环2(通过XhoI切割)之后获取的图像作为系列中的第二图像进行分析;将循环1之后获取的图像(通过NotI切割)作为系列中的第三图像进行分析;并且将未切割的珠粒作为系列中的第四图像进行分析。在处于颠倒的时间顺序时,从“下一个”图像中减去在“前一个”图像中获得的所有四个荧光通道的强度值,以便逐循环确定与每个标识体段缔合的标记。例如,从针对所有四个荧光通道从XhoI切割之后获取的图像中获得的强度值中减去针对所有四个荧光通道从SpeI切割之后获取的图像中获得的强度值,并且对于系列中的每个循环都进行此分析。分析之后,在每个循环时针对单标记和双标记的链绘制在所有四个荧光通道中获得的强度值。图21A示出了从携带单标记标识体链的珠粒中释放的标记的序列或顺序。图21B示出了从携带双标记的链的链霉亲和素珠粒中释放的顺序。在单标记实验中,跨所有切割循环释放的标记的顺序

(即,序列)易于解析;在图中仅对标记的FND标识体符号编号(Id_0是未标记的并附接到珠粒)。在双标记的实验中,成像错误阻止了对Id_3的识别,但附接到链中的侧接的标识体段的标记易于解析。在单标记实验中解码的dsDNA标识体链序列为Id_1-550、Id_2-750、Id_3-647和Id_4-488。在双标记实验中解码的dsDNA标识体链序列为Id_1-550/750、Id_2-488/647,Id_3未确定,以及Id_4-750/488。这些结果表明,荧光标记的珠粒的组合文库可以使用这种方法进行编码和解码。例如,此类组合文库的理论多样性(基于示出的实验开发)对于单标记链将为256,并且对于双标记链将为10,000。以其最大理论多样性解码此类文库将需要跟踪单独链霉亲和素珠粒。

[0301] 解码基于FND标识体的分子条形码:跟踪4个图像内的9个单独珠粒(每个标识体单个标记,3个切割循环)

[0302] 如本文先前所公开的,相较于通过每个单独标识体的连接以及其间的洗涤逐步构建的基于FND标识体的分子条形码,包括FND Id_1-550、FND Id_2-750、FND Id_3-647和FND Id_4-488符号的分子条形码(通过在单个连接反应内对所有标识体进行的溶液中连接构建,并且然后捕获到珠粒上)在性能方面未显示出明显差异(从各个解码实验中进行的定性和定量评估)。因此,在溶液中预先连接针对单个珠粒跟踪实验(如图22所示)构建的基于单标记FND标识体的分子条形码,之后用于涂覆珠粒。

[0303] 为了证明跨所有切割循环对单独珠粒的跟踪,将与包括标记的FND标识体符号的单个预先确定的序列的分子条形码缀合的链霉亲和素珠粒固定在流动池中,并在暴露于每种切割剂(循环1中NotI,循环2中XhoI,并且循环3中SpeI)之前和之后进行成像。在所有循环的成像之后,选择9个单独珠粒用于跟踪,以颠倒的时间顺序对图像进行分析,如上所述减去来自“前一”循环的值,并且针对每个循环的所有9个珠粒绘制所产生的强度值。

[0304] 图22A表示六个携带相同的用于由经OCS实验的3个循环跟踪的FOV中的单独珠粒生成数据的单标记标识体链序列的I珠粒中的第一珠粒。图22B表示六个携带相同的用于由经OCS实验的3个循环跟踪的FOV中的单独珠粒生成数据的单标记标识体链序列的I珠粒中的第二珠粒。图22C表示携带相同的用于由经OCS实验的3个循环跟踪的FOV中的单独珠粒生成数据的单标记标识体链序列的第三珠粒。图22D表示携带相同的用于由经OCS实验的3个循环跟踪的FOV中的单独珠粒生成数据的单标记标识体链序列的第四珠粒。图22E表示携带相同的用于由经OCS实验的3个循环跟踪的FOV中的单独珠粒生成数据的单标记标识体链序列的第五珠粒。图22F表示携带相同的用于由经OCS实验的3个循环跟踪的FOV中的单独珠粒生成数据的单标记标识体链序列的第六珠粒。

[0305] 图22示出了从9个单独跟踪的链霉亲和素珠粒中获得的FND标识体“正交切割测序”(OCS)数据。本实验中跟踪的所有链霉亲和素珠粒在暴露于预先确定顺序的正交切割剂时显示出预期(正确)顺序的荧光团释放。这些结果表明,与包括许多不同的唯一标记的FND标识体符号序列的基于FND标识体的分子条形码缀合的链霉亲和素珠粒可以按高度并行的方式解析。

[0306] 编码和解码256个成员组合珠粒文库:包括不同颜色代码的FND标识体分子条形码

[0307] 为了证明并行对许多链霉亲和素珠粒进行编码,将单标记FND标识体以随机方式并入到含有差异标记组合的链中。组合分子条形码构建是通过多个轮次的FND标识体连接步骤之间分裂和池化链霉亲和素珠粒来进行的。为了产生总多样性为256的荧光链霉亲和

素珠粒的文库,将4种不同的标记放置在四种FND标识体类别(FND(256)Id_{1-4})中的每种FND标识体类别上,以产生256种可能的标记组合。FND(256)Id_0是未标记的,并且首先附接到链霉亲和素珠粒,以充当用于连接FND(256)Id_1的受体(先前在本文中在构建单标记链时进行描述)。FND(256)Id_{1-3}被配置成分别具有HindII、SpeI、XhoI识别部分。更具体地,首先用所有4种荧光团(ATTO-488、ATTO-550、AF-647和AF-750)通过实例5中公开的标记方法对FND(256)Id_{1-4}进行标记。将每个标记的寡核苷酸双链体保持在标准96孔板的单独孔中。以此方式,在分裂和池化的每个轮次期间,每个FND标识体可使用4种不同标记的选项。在分子条形码构建的第一轮次中,将0.2mg/ml涂覆有FND(256)Id_0的珠粒分裂到4个孔中,并且暴露于含有1X T4 DNA连接酶缓冲液的补充有T4多核苷酸激酶(PNK)、T4 DNA连接酶和四种差异标记的FND(256)Id_1(即,FND(256)Id_1-488、FND(256)Id_1-550、FND(256)Id_1-647或FND(256)Id_1-750)之一的300nM浓度的溶液中。温育(大约15分钟至20分钟)以促进FND(256)Id_1的高效连接之后,将链霉亲和素珠粒(磁性)在相同的孔中(其在所述孔中连接)利用用于淬灭连接反应并且洗掉未连接的FND(256)Id_1的溶液洗涤几次。然后将与四种差异标记的FND(256)Id_1缀合的链霉亲和素珠粒池化到同一管中以均匀混合。然后将链霉亲和素珠粒分裂到4个孔中,并且暴露于含有4种可用于FND(256)Id_2的差异标记的选项中的每种选项的连接溶液。重复此分裂和池化程序,以将FND(256)Id_{1-4}的所有四种标识体类别连接,以产生256个不同珠粒的文库。使用本文先前描述的OCS程序,对来自文库的单独珠粒进行跟踪,按如上所述的颠倒的时间顺序对图像进行分析(利用“前一”循环减法),并且针对每个循环绘制从在所有四个通道中进行的成像获得的荧光值。

[0308] 图23示出了来自256个成员文库的3个单独链霉亲和素珠粒的OCS测序读段。图23中示出的荧光团从3个珠粒释放的顺序在本实验中易于解析。与双标记实验一起,这些数据表明,使用本文所述的方法可以对10,000个珠粒的文库进行编码和解码。在链中添加更多标识体的情况下,或者在使用额外的可单独检测的标记(例如量子点)的情况下,使用OCS可以对大得多的文库进行编码和解码。

[0309] 实例6编码和解码基于环标识体的分子条形码

[0310] A.环标识体组成和构建

[0311] 如实例6中公开的,三种标识体类别(环Id_{1-3})被构建成具有包括连接的dsDNA环的支架部分(本文中被称为“环标识体”或“环Id”),并且用于构建基于环Id的分子条形码的集合。环Id_{1-3}中的每个环的支架部分被配置成包括一或多个限制性位点和用于附接到链霉亲和素珠粒(通过NHS-LC-生物素;可从美国密苏里州68178圣路易斯市的西格玛奥德里奇公司邮政信箱14508商购获得;CAS编号:35013-72-0)和NHS-荧光团(可从美国马萨诸塞州02451沃尔瑟姆第三大街168的赛默飞世尔科技公司商购获得;目录号65601)的经氨基修饰的碱基。所有寡核苷酸-链霉亲和素缀合如本文先前所述进行。

[0312] 环Id_1的支架部分包括Hp_双_NH2_XhoI寡核苷酸(SEQ ID NO:27)和Hp_iNH2_XhoI_珠粒寡核苷酸(SEQ ID NO:28)。对应于SEQ ID NO:27的位置1的核苷酸通过5’磷酸化进行修饰。SEQ ID NO:27的位置14和20为经Int氨基修饰剂C6 dT修饰的核苷酸。SEQ ID NO:27的位置34至37包括粘性端,所述粘性端被配置成粘接到SEQ ID NO:28的粘性端。SEQ ID NO:27的位置13至22包括茎环特征。对应于SEQ ID NO:28的位置1的核苷酸通过5’磷酸化进行修饰。SEQ ID NO:28的位置17为经Int氨基修饰剂C6 dT修饰的核苷酸。SEQ ID NO:

28的位置34至37包括粘性端,所述粘性端被配置成粘接到SEQ ID NO:27的粘性端。SEQ ID NO:27的位置13至21包括茎环特征。

[0313] 环Id_2的支架部分包括Hp_双_NH2_SpeI寡核苷酸(SEQ ID NO:29)和Hp_iNH2_SpeI_珠粒寡核苷酸(SEQ ID NO:30)。对应于SEQ ID NO:29的位置1的核苷酸通过5'磷酸化进行修饰。SEQ ID NO:29的位置14和20为经Int氨基修饰剂C6 dT修饰的核苷酸。SEQ ID NO:29的位置34至37包括粘性端,所述粘性端被配置成粘接到SEQ ID NO:30的粘性端。SEQ ID NO:29的位置13至21包括茎环特征。对应于SEQ ID NO:30的位置1的核苷酸通过5'磷酸化进行修饰。SEQ ID NO:30的位置17为经Int氨基修饰剂C6 dT修饰的核苷酸。SEQ ID NO:30的位置34至37包括粘性端,所述粘性端被配置成粘接到SEQ ID NO:29的粘性端。SEQ ID NO:30的位置13至21包括茎环特征。

[0314] 环Id_3的支架部分包括Hp_双_NH2_NotI寡核苷酸(SEQ ID NO:31)和Hp_iNH2_NotI_珠粒寡核苷酸(SEQ ID NO:32)。对应于SEQ ID NO:31的位置1的核苷酸通过5'磷酸化进行修饰。SEQ ID NO:31的位置16和22为经Int氨基修饰剂C6 dT修饰的核苷酸。SEQ ID NO:31的位置38至41包括粘性端,所述粘性端被配置成粘接到SEQ ID NO:32的粘性端。SEQ ID NO:31的位置13至21包括茎环特征。对应于SEQ ID NO:32的位置1的核苷酸通过5'磷酸化进行修饰。SEQ ID NO:32的位置17为经Int氨基修饰剂C6 dT修饰的核苷酸。SEQ ID NO:32的位置34至37包括粘性端,所述粘性端被配置成粘接到SEQ ID NO:31的粘性端。SEQ ID NO:32的位置13至21包括茎环特征。

[0315] 将由Hp_iNH2_XhoI_珠粒寡核苷酸构成的集合、由Hp_iNH2_SpeI_珠粒寡核苷酸构成的集合和由Hp_iNH2_NotI_珠粒寡核苷酸构成的集合(统称为"HP_iNH2寡核苷酸")合成以准备好与链霉亲和素珠粒的集合缀合。将Hp_iNH2寡核苷酸与胺反应性生物素(NHS-LC-生物素)组合,以使生物素与Hp_iNH2寡核苷酸在其游离胺处(即,对应于所有三种HP_iNH2寡核苷酸的位置17的核苷酸处)缀合。将每个由生物素化的HP_iNH2寡核苷酸构成的集合重悬至200nM的最终浓度。

[0316] 将由Hp_双_NH2_Xhol寡核苷酸构成的集合、由Hp_双_NH2_SpeI寡核苷酸构成的集合和由Hp_双_NH2_NotI寡核苷酸构成的集合(统称为"Hp_双_NH2寡核苷酸")合成,并且用荧光团标记,以准备好与其相应的Hp_双_NH2寡核苷酸连接。将胺反应性ATTO-550(NHS-ATTO-550)与由Hp_双_NH2_Xhol寡核苷酸构成的集合组合以在对应于位置14和20的核苷酸处对其进行双标记。将胺反应性ATTO-647(NHS-ATTO-647)与由Hp_双_NH2_SpeI寡核苷酸构成的集合组合以在对应于位置14和20的核苷酸处对其进行双标记。将胺反应性ATTO-488(NHS-ATTO-488)与由Hp_双_NH2_NotI寡核苷酸构成的集合组合以在对应于位置16和22的核苷酸处对其进行双标记。将每个由标记的Hp_双_NH2寡核苷酸构成的集合重悬至200nM的最终浓度。

[0317] 为了形成和编码由基于环Id的分子条形码构成的集合,首先通过本文先前所述的方法将所有三个由生物素化的HP_iNH2寡核苷酸构成的集合与由链霉亲和素珠粒构成的集合结合。将由链霉亲和素珠粒构成的集合洗涤以去除未结合的寡核苷酸(用200μl补充有0.01% Tween-20的2X杂交缓冲液洗涤3X,然后用200μl 1X杂交缓冲液洗涤两次),并且然后用NEB公司提供的100μl 1X连接缓冲液洗涤两次或更多次,重悬于50μl连接缓冲液中,并且超声处理3分钟以在连接反应之前破坏珠粒的团块。

[0318] 连接反应如本文先前所述如下进行:使用0.2mg/ml用来自所有三个由HP_iNH2寡核苷酸构成的集合的成员涂覆的链霉亲和素珠粒,并且在单个连接轮次中引入每个由Hp_双_NH2寡核苷酸构成的集合。换言之,将由链霉亲和素珠粒构成的集合:首先在第一连接循环中通过与珠粒缀合的Hp_iNH2_XhoI_珠粒寡核苷酸的粘性端连接与由Hp_双_NH2_Xhol寡核苷酸构成的集合连接;其次在第二连接循环中通过与珠粒缀合的Hp_iNH2_SpeI_珠粒寡核苷酸的粘性端连接与由Hp_双_NH2_SpeI寡核苷酸构成的集合连接;并且然后第三,在第三连接循环中通过Hp_iNH2_NotI_珠粒寡核苷酸与由Hp_双_NH2_NotI寡核苷酸构成的集合连接,以完成环Id_{1-3}的构建并形成由基于环Id的分子条形码构成的集合。因此,每个基于环Id的分子条形码的三维布置包括由缀合的经编码的环Id_{1-3}符号和单个链霉亲和素珠粒构成的集合。

[0319] B.解码基于环Id的分子条形码

[0320] 编码的三个连接轮次完成之后,将链霉亲和素珠粒用200μl补充有0.01% Tween-20的2X杂交缓冲液洗涤3X,用200μl 1X杂交缓冲液洗涤2X,重悬于20μl 1X杂交缓冲液中,超声处理3分钟,并且上样到含有经生物素修饰的表面的流动池中用于固定化和随后的切割实验。

[0321] 在一些实施例中,可以使用OCS相容性链霉亲和素珠粒文库以对其组分与OCS代码协调的其它分子进行共编码。例如,核酸内的条形码捕获用于NGS工作流的寡核苷酸(如本文所述,或作为附接到同一珠粒的单独分子),并且这可以通过协调连接来完成。在一些实施例中,可以承受合成化学反应的标识体构型可用于使得能够对化学文库进行编码。OCS相容性文库可以由包含聚合物而非DNA的支架部分构建,以解决这个问题(如本文先前所述)。技术人员将理解,当DNA的线性5'和3'端暴露时,其在很大程度上容易降解,并且因此,在环形化的情况下,受更多保护以免于由于暴露于一些化学反应而造成降解。

[0322] 为了进一步探索不同标识体三维布置在OCS工作流中的功效,如前所述,在珠粒上逐步构建含有编码不同限制性酶识别位点的dsDNA的长区的环形ssDNA环的标记的组合(图25)。将三种不同的ssDNA发夹寡核苷酸HP_iNH2寡核苷酸附接到链亲和素珠粒,所述链霉亲和素珠粒各自含有适于仅与三种不同标记的发夹寡核苷酸之一连接的5'悬垂部。在如上所述的连接溶液中每次引入一个标记的发夹,经3个构建轮次(编码)。在环标识体构建的每个轮次之后,提取固定在流动池中的珠粒在相关荧光通道中的图像(以获取488nm、550nm和647nm发射的数据)。从给定FOV中的许多链霉亲和素珠粒的平均值中获得强度值,并且在原始图像旁绘制针对每个标记绘制强度值。从成像中显而易见的是,在与每个标记的发夹逐步连接之后,仅在预期通道中获得珠粒的荧光强度。

[0323] 此过程可以在如前所述的分裂和池化程序中重复,其中连接步骤通过每个池化步骤之前的洗涤步骤分开。这些结果显示,标识体可以构建为DNA环,并且这些结构可以以组合的形式编码。

[0324] 为了测试OCS工作流中的环结构,将含有三个形成的环标识体的珠粒固定在流动池中,并暴露于含有300U SpeI酶的溶液中(图8A)。在通过SpeI切割之前和之后提取图像,并且针对三个相关的荧光通道绘制从珠粒获得的总强度值(图8B)。事实上,连接后,环形产物包括适于通过SpeI切割的完全形成的酶可接近的dsDNA限制性位点。综上所述,这些结果显示,标识体可以被构建为DNA的环,并且其可以以组合的形式编码以制备OCS相容性文库。

[0325] 实例7编码和解码基于发夹标识体的分子条形码

[0326] 为了探索不同标识体三维结构在OCS工作流中的功效,借助于通过本文先前公开的方法使标记的且生物素化的dsDNA受体与双标记的发夹连接来构建包括两个由酶可接近的SpeI切割位点分开的差异标记的荧光ssDNA发夹标识体(图24A)。在暴露于含300U SpeI的1X CutSmart缓冲液(NEB公司)之前和之后(图24B),在两个荧光通道中从为涂覆的FSH标识体的链霉亲和素珠粒(如前所述固定在流动池中)中获取图像。获取的图像显示,在37℃下暴露于300U SpeI酶5分钟之后,AF-750被清楚地去除,而珠粒上留下来自AF-647标记的强烈信号。对来自每个FOV中的许多珠粒的值进行求和并求平均(此处不跟踪单独珠粒),然后将荧光强度值校正为百分比值,以便视觉表示SpeI切割后来自珠粒的总信号损失%(750nm发射)(图24C)。在此实验中,切割后AF-750信号的损失为约82%。这些结果表明,此处设计的发夹结构可以充当两个段的标识体,从而允许不同发夹的组合含有限制性位点和可视觉辨别的标记的不同组合,以用于构建和产生(编码)OCS相容性文库。

[0327] 如实例7中公开的,三种标识体类别(发夹Id_{1-3})被构建成具有包括茎双链体部分和环部分的支架部分(本文中被称为“发夹标识体”),并且用于构建基于发夹Id的分子条形码的集合。发夹Id_{1-3}中的每一种的茎双链体部分包括第一和第二寡核苷酸,所述第一和第二寡核苷酸被配置成杂交并形成具有可用于连接的自由粘性端的dsDNA双链体。在形成双链体时,第一和第二寡核苷酸包括第一和第二限制性核酸内切酶识别部分(在本文分别被称为“RE1位点”和“RE2位点”)。发夹Id_{1-3}中的每一种的第一寡核苷酸被配置成具有可用于使双链体生物素化的经5AmMC6修饰的核苷酸。发夹Id_{1-3}中的每一种的第二寡核苷酸被配置成具有可用于通过NHS荧光团进行荧光标记的经内部iAmMC6T修饰的核苷酸。如实例7中所用,发夹Id_{1-3}的每个环部分都包括被配置成与步骤-环结构形成dsDNA双链体的ssDNA Hp_双_NH2寡核苷酸(SEQ ID NO:39)。对应于SEQ ID NO:39的位置1的核苷酸通过5’磷酸化进行修饰。SEQ ID NO:39的位置14和20为经Int氨基修饰剂C6 dT修饰的寡核苷酸。位置13至21包括茎环结构。每个环部分被配置成具有茎-环结构内可用于通过NHS荧光团进行荧光标记的位置处的一或多个经iAmMC6T修饰的核苷酸。发夹Id_{1-3}中的每个发夹的支架部分被配置成包括一或多个识别部分和用于附接到链霉亲和素珠粒(通过NHS-LC-生物素;可从美国密苏里州68178圣路易斯市的西格玛奥德里奇公司邮政信箱14508商购获得;CAS编号:35013-72-0)和NHS-荧光团(可从美国马萨诸塞州02451沃尔瑟姆第三大街168的赛默飞世尔科技公司商购获得;目录号65601)的经氨基修饰的碱基。所有寡核苷酸-链霉亲和素缀合如本文先前所述进行。

[0328] 发夹Id_1的茎双链体部分的第一寡核苷酸包括Ab_茎1_SpeI_XhoI寡核苷酸(SEQ ID NO:33)。对应于SEQ ID NO:33的位置1的核苷酸为经氨基修饰剂C6修饰的核苷酸。SEQ ID NO:33的位置17至22和31至36分别包括SpeI识别部分和XhoI识别部分。SEQ ID NO:33的位置41至44包括粘性端,所述粘性端被配置成粘接到SEQ ID NO:39的粘性端。发夹Id_1的茎双链体部分的第二寡核苷酸包括Ab_茎1_comp寡核苷酸(SEQ ID NO:34)。对应于SEQ ID NO:34的位置1的核苷酸通过5’磷酸化进行修饰。SEQ ID NO:34的位置14为经Int氨基修饰剂C6 dT修饰的核苷酸。SEQ ID NO:28的位置5至10和18至23分别包括XhoI识别部分和SpeI识别部分。

[0329] 发夹Id_2的茎双链体部分的第一寡核苷酸包括Ab_茎2_HindIII_SpeI寡核苷酸

(SEQ ID NO:35)。对应于SEQ ID NO:35的位置1的核苷酸通过氨基修饰剂C6进行修饰。SEQ ID NO:35的位置16至21和30至35分别包括HindIII识别部分和SpeI识别部分。SEQ ID NO:35的位置40至43包括粘性端,所述粘性端被配置成粘接到SEQ ID NO:39的粘性端。发夹Id_2的茎双链体部分的第二寡核苷酸包括Ab_茎2_comp寡核苷酸(SEQ ID NO:36)。对应于SEQ ID NO:36的位置1的核苷酸通过5'磷酸化进行修饰。SEQ ID NO:36的位置14为经Int氨基修饰剂C6 dT修饰的核苷酸。SEQ ID NO:36的位置5至10和18至23分别包括HindIII识别部分和SpeI识别部分。

[0330] 发夹Id_3的茎双链体部分的第一寡核苷酸包括Ab_茎3_EcoRI_HindIII寡核苷酸(SEQ ID NO:37)。对应于SEQ ID NO:37的位置1的核苷酸是通过氨基修饰剂C6进行修饰的。SEQ ID NO:37的位置16至21和30至35分别包括EcoRI识别部分和HindIII识别部分。SEQ ID NO:33的位置40至43包括粘性端,所述粘性端被配置成粘接到SEQ ID NO:39的粘性端。发夹Id_3的茎双链体部分的第二寡核苷酸包括Ab_茎3_comp寡核苷酸(SEQ ID NO:38)。对应于SEQ ID NO:38的位置1的核苷酸通过5'磷酸化进行修饰。SEQ ID NO:34的位置14为经Int氨基修饰剂C6 dT修饰的核苷酸。SEQ ID NO:28的位置5至10和18至23分别包括HindIII识别部分和EcoRI识别部分。

[0331] 将由Ab_茎1_SpeI_XhoI寡核苷酸构成的集合、由Ab-茎1_SpeI_XhoI寡核苷酸构成的集合、由Ab_茎2_HindIII_SpeI构成的集合、由Ab-茎2_comp寡核苷酸构成的集合、由Ab_茎3_EcoRI_HindIII寡核苷酸构成的集合和由Ab_茎3_comp寡核苷酸构成的集合(本文统称为"茎寡核苷酸")合成,以准备好用于双链体形成和与由链霉亲和素珠粒构成的集合。

[0332] 如实例7中所用,从IDT公司订购呈经纯化的寡核苷酸(HPLC)或标准经脱盐的寡核苷酸形式的所有寡核苷酸。首先对含有5'氨基修饰的标准经脱盐的寡核苷酸进行脱盐或沉淀,以去除从合成遗留的任何过量游离胺。将茎寡核苷酸的5'氨基(200uM)利用用于附接到MyOne T1链霉亲和素珠粒的适当化学通过将重悬于无水DMF(可从密苏里州68178圣路易斯市的西格玛奥德里奇公司邮政信箱14508商购获得;Cas编号:68302-57-8)中的NHS-LC-生物素添加至大约2.0mM的最终浓度进行修饰,并且允许在室温下在NHS缀合缓冲液(NCB公司:100mM NaPO4,pH 8.5)中反应过夜。然后将此反应使用7K MWCO Zeba脱盐柱(可从美国马萨诸塞州02451沃尔瑟姆第三大街168的赛默飞世尔科技公司商购获得;目录号89891)两次缓冲液交换到水中,以去除所有过量的未经反应的生物素。

[0333] 用各种NHS-荧光团标记经内部氨基修饰的寡核苷酸(即,Ab_茎1_comp、Ab_茎2_comp、Ab_茎2_comp、HP_双_NH2)如本文先前所述进行(使1mM NHS-荧光团试剂与200μM寡核苷酸在黑暗中在室温下在NHS缀合缓冲液中反应过夜,以防止荧光团光漂白)。第二天通过用2X杂交缓冲液(20mM pH 7.5的Tris-HCl/1M NaCl/1mM EDTA)10倍稀释至20uM的浓度来将这些反应淬灭。

[0334] 如图<tbd>所示,Ab_茎1_SpeI_XhoI被生物素化,Ab_茎1_comp用NHS-AF-647标记,并且将两种寡核苷酸在1X杂交缓冲液中以1:1.2摩尔比(分别地;10μM生物素化的寡核苷酸:12μM标记的寡核苷酸)通过以下粘接:在90℃下加热5分钟,并且在热块中经约30分钟缓慢冷却至RT。这种粘接的生物素化的且647标记的寡核苷酸双链体在溶液中与HP_双_NH2寡核苷酸连接,所述HP_双_NH2寡核苷酸是用如上所述的单一荧光团类型(如实例中所示的NHS-AF750;赛默公司)预标记的。通过将200nM的粘接的双链体在1X T4 DNA连接酶缓冲液

(NEB公司)中与最终浓度为400nM的AF750标记的HP_双_NH2寡核苷酸、与5μl T4 PNK(NEB公司)和5μl T4 DNA连接酶(NEB公司)在250μl连接反应中在37℃下混合30分钟进行溶液中连接。在连接反应中与生物素化的寡核苷酸结合之前,将连接直接捕获到如上所述预先洗涤(用2X杂交缓冲液进行2次洗涤)的MyOne T1链霉亲和素珠粒上。通过将0.2mg/ml的250μl经洗涤的珠粒在2X杂交缓冲液中与250μl连接反应混合启动结合,并且允许结合37℃下进行30分钟。然后将珠粒用200μl补充有0.01％ Tween-20的2X杂交缓冲液洗涤3X,用200μl 1X杂交缓冲液洗涤2X,重悬于20μl 1X杂交缓冲液中,超声处理3分钟,并且上样到含有经生物素修饰的表面的流动池中用于固定化和随后的切割实验。

[0335] 实例9编码和解码由ssDNA杂交(Hyb)标识体产生的分子条形码

[0336] 如实例9中所公开的,由三种标识体类别(Hyb Id_{1-3})形成和编码由三个基于Hyb Id的分子条形码构成的集合,所述三种标识体类别中的每种标识体类别包括具有第一杂交寡核苷酸和第二杂交寡核苷酸的支架部分,所述杂交核苷酸被配置成彼此杂交并形成具有一或多个正交切割位点的生物素化的荧光标记的dsDNA双链体(统称为“杂交标识体”或“Hyb Id”),所述一或多个正交切割位点用于构建由基于Hyb Id的分子条形码构成的集合。Hyb Id_{1-3}的支架部分分别包括XhoI识别部分、SpeI识别部分和NotI识别部分。

[0337] Hyb Id_1的支架部分包括Hyb_5NH2_XhoI_珠粒第一杂交寡核苷酸(SEQ ID NO:40)(在本文还被称为“Hyb_5NH2_XhoI_珠粒寡核苷酸”)和Hyb_5NH2_XhoI_comp第二杂交寡核苷酸(SEQ ID NO:41)(在本文还被称为“Hyb_5NH2_XhoI_comp寡核苷酸”)。对应于SEQ ID NO:40的位置1的核苷酸为氨基修饰剂C6修饰的。SEQ ID NO:40的位置22至27包括XhoI识别部分(在实例9中还被称为“RE3位点”)。对应于SEQ ID NO:41的位置1的核苷酸为氨基修饰剂C6修饰的。SEQ ID NO:41的位置8至13包括XhoI识别部分。

[0338] Hyb Id_2的支架部分包括Hyb_5NH2_SpeI_珠粒第一杂交寡核苷酸(SEQ ID NO:42)(在本文还被称为“Hyb_SpeI_XhoI_珠粒寡核苷酸”)和Hyb_5NH2_SpeI_comp第二杂交寡核苷酸(SEQ ID NO:43)(在本文还被称为“Hyb_5NH2_SpeI_comp寡核苷酸”)。对应于SEQ ID NO:42的位置1的核苷酸为氨基修饰剂C6修饰的。SEQ ID NO:42的位置21至26包括SpeI识别部分(在实例9中还被称为“RE2位点”)。对应于SEQ ID NO:43的位置1的核苷酸为氨基修饰剂C6修饰的。

[0339] Hyb Id_3的支架部分包括Hyb_5NH2_NotI_珠粒第一杂交寡核苷酸(SEQ ID NO:44)(在本文还被称为“Hyb_5NH2_XhoI_珠粒寡核苷酸”)和Hyb_5NH2_XhoI_comp第二杂交寡核苷酸(SEQ ID NO:45)(在本文还被称为“Hyb_5NH2_XhoI_comp寡核苷酸”)。对应于SEQ ID NO:44的位置1的核苷酸为氨基修饰剂C6修饰的。SEQ ID NO:44的位置22至27包括NotI识别部分(在实例9中还被称为“RE3位点”)。对应于SEQ ID NO:41的位置1的核苷酸为氨基修饰剂C6修饰的。

[0340] 将胺反应性ATTO-550(NHS-ATTO-550)与由Hyb_5NH2_XhoI_comp寡核苷酸构成的集合组合以在其5’游离胺(即,对应于SEQ ID NO:41的位置1的核苷酸)处对其进行标记。将胺反应性AF-647(NHS-AF-647)与由Hyb_5NH2_SpeI_comp寡核苷酸构成的集合组合以在其5’游离胺(即,对应于SEQ ID NO:43的位置1的核苷酸)处对其进行标记。将胺反应性ATTO-488(NHS-ATTO-488)与由Hyb_5NH2_NotI_comp寡核苷酸构成的集合组合以在其5’游离胺(即,对应于SEQ ID NO:45的位置1的核苷酸)处对其进行标记。

[0341] 将所标记的由Hyb_5NH2_XhoI_comp寡核苷酸、Hyb_5NH2_SpeI_comp寡核苷酸和Hyb_5NH2_NotI_comp寡核苷酸(在实例9中统称为“comp寡核苷酸”)构成的集合

[0342] 如本文已经描述的进行生物素化,并且将所有三种寡核苷酸添加到MyOne T1链霉亲和素珠粒中用于涂覆。用如上所述的各种NHS-荧光团试剂标记comp寡核苷酸。为了对用生物素化的寡核苷酸涂覆的珠粒进行组合编码,在分裂和池化循环期间一次引入一个标记的杂交寡核苷酸。这些标记的杂交寡核苷酸在每个编码循环期间以200nM的浓度添加在0.5X杂交缓冲液中。标记的链与珠粒在37℃下温育30分钟,其中偶尔轻敲管以促进捕获。在每个编码循环之间,将珠粒用200μl补充有0.01％ Tween-20的2X杂交缓冲液洗涤3X,用200μl 1X

[0343] 杂交缓冲液洗涤2X,重悬于20μl 1X杂交缓冲液中,并且超声处理3分钟。通过杂交编码标识体的三个轮次完成之后,将珠粒用200μl补充有0.01％ Tween-20的2X杂交缓冲液洗涤3X,用200μl 1X杂交缓冲液洗涤2X,重悬于20μl 1X杂交缓冲液中,超声处理3分钟,并且上样到含有经生物素修饰的表面的流动池中用于固定化和随后的切割实验。

[0344] 将胺反应性ATTO-550(NHS-ATTO-550)与由Hyb_5NH2_XhoI_comp寡核苷酸构成的集合组合以在其5’游离胺(即,对应于SEQ ID NO:41的位置1的核苷酸)处对其进行标记。将胺反应性AF-647(NHS-AF-647)与由Hyb_5NH2_SpeI_comp寡核苷酸构成的集合组合以在其5’游离胺(即,对应于SEQ ID NO:43的位置1的核苷酸)处对其进行标记。将胺反应性ATTO-488(NHS-ATTO-488)与由Hyb_5NH2_NotI_comp寡核苷酸构成的集合组合以在其5’游离胺(即,对应于SEQ ID NO:45的位置1的核苷酸)处对其进行标记。实例9结束。

[0345] 成像系统动态范围

[0346] 在一些实施例中,成像方法可能受到成像系统的动态范围的限制(参见维斯里德(Weissleder)等人,IEEE量子电子学选题杂志(IEEE J.Sel.Top Quantum Electron);1月-2月;25(1):6801507(2019))。例如,在构建本文所述的文库时,一些链霉亲和素珠粒将具有同一荧光团的多个拷贝,并且发现使任何荧光通道中的信号饱和的链霉亲和素珠粒可能难以解析。因此,在一些实施例中,高动态范围成像的使用将改进可以使用组合物和本文所公开的方法构建的文库多样性。例如,获取每个实验数据点的三个图像:一个在低暴露下提取,一个在中等暴露下提取,并且一个在较高暴露下提取。这三个图像在数学上反向拼接在一起,以创建具有非常高动态范围的连续“图像”。以此方式,含有荧光团的非常少的拷贝的链霉亲和素珠粒可以在与含有所述荧光团的许多拷贝的链霉亲和素珠粒相同的视场(实验)中进行成像。具有荧光团的非常少的拷贝的链霉亲和素珠粒需要较高暴露以使其值上升到可以对其进行准确定量的范围,并且具有荧光团的许多拷贝的珠粒需要较低暴露以使其值下降到低于饱和,并且进入可以对其进行准确定量的范围。因此,在一些实施例中,使用高动态范围成像允许链霉亲和素珠粒在同一实验中有较大偏差,以增加用本文所公开的组合物和方法构建的文库的多样性。

[0347] 对于本领域的技术人员将显而易见的是,在不偏离本公开的基本原理的情况下,可以对上述实施例的细节进行许多改变。

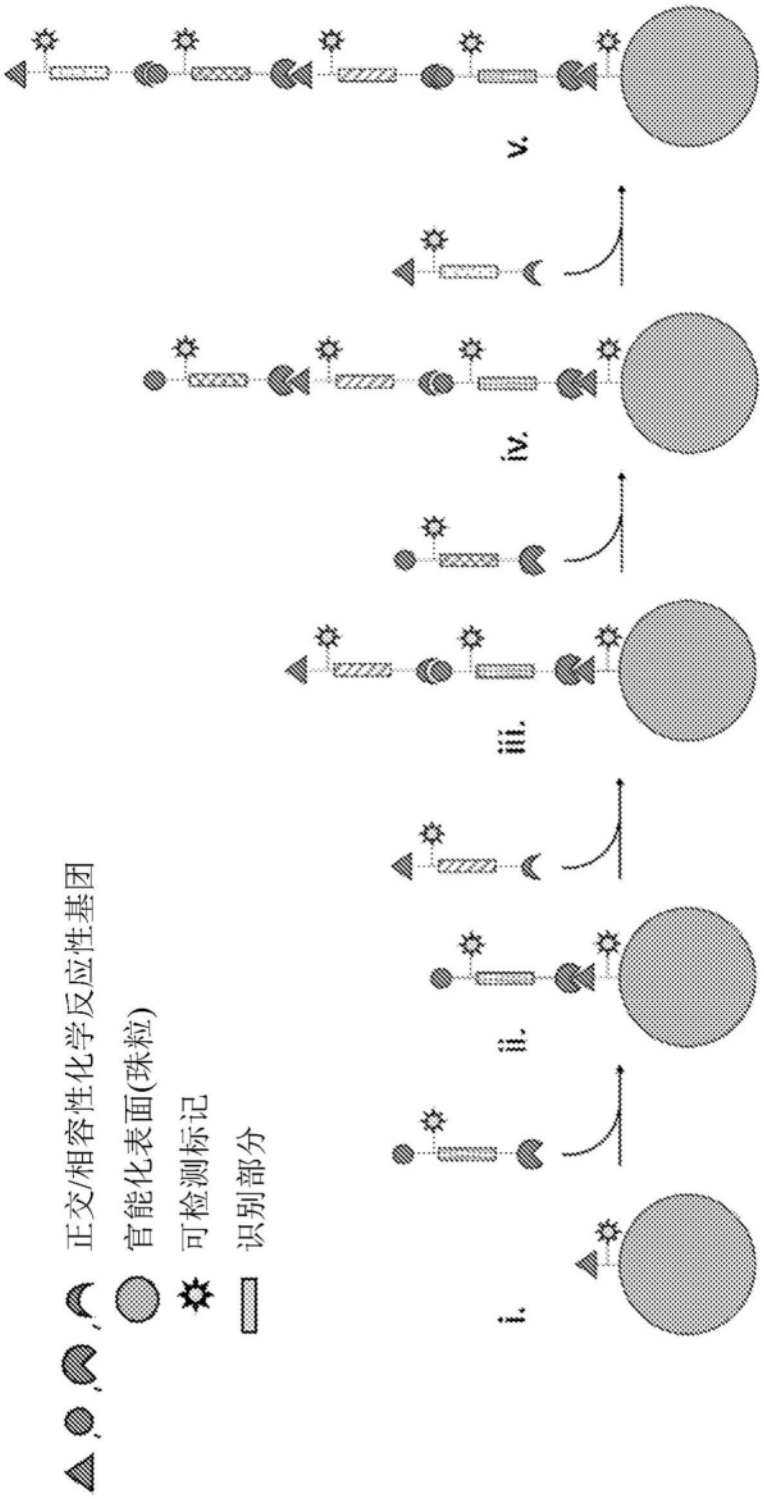


图1

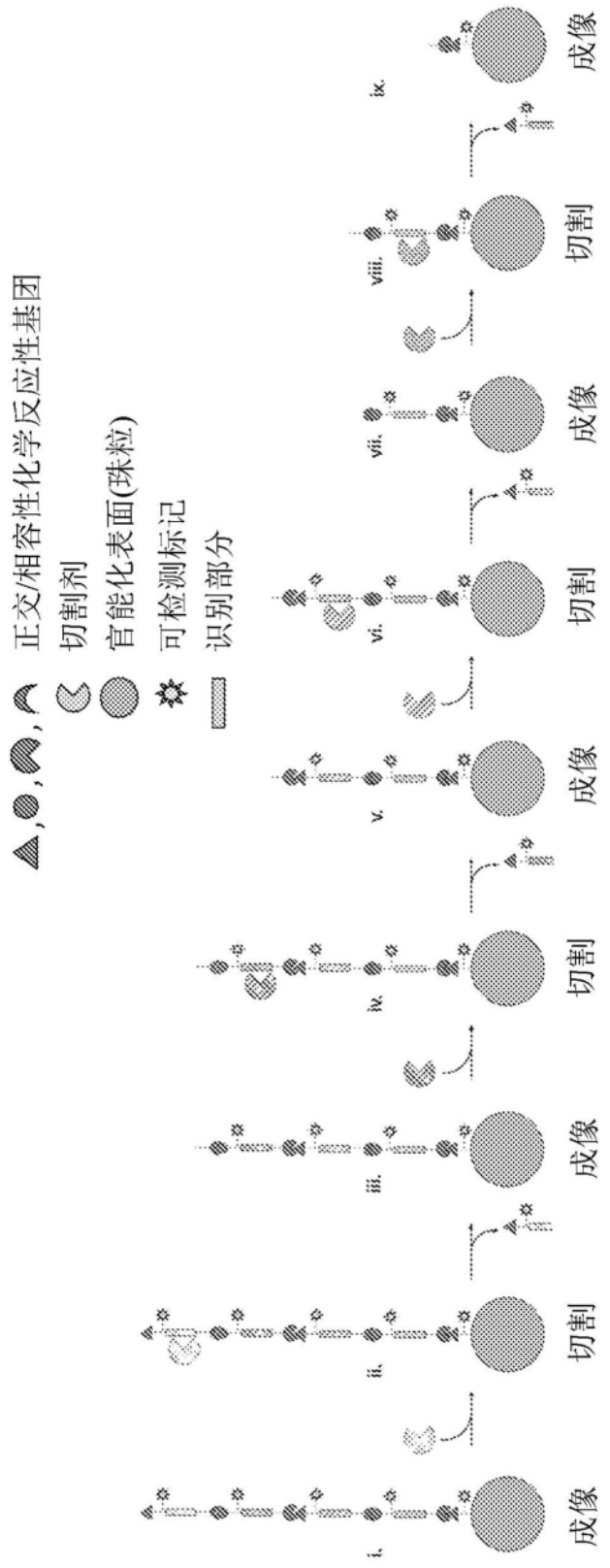

正交/相容性化学反应性基团
切割剂
官能化表面(珠粒)
可检测标记
识别部分
i.
ii.
iii.
iv.
v.
vi.
vii.
viii.
ix.
成像
切割
成像
切割
成像
切割
成像
切割
成像

图2

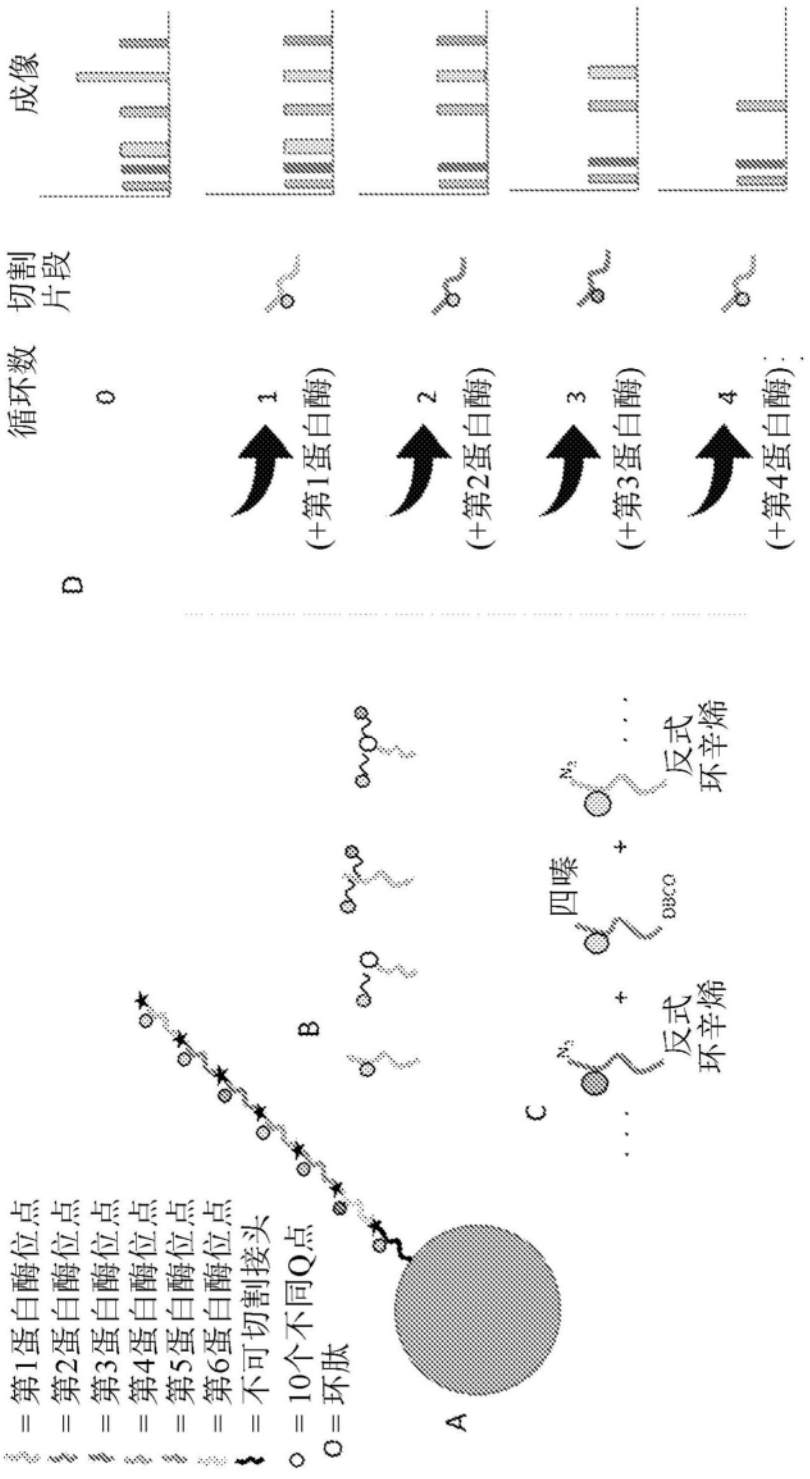

= 第1蛋白酶位点
= 第2蛋白酶位点
= 第3蛋白酶位点
= 第4蛋白酶位点
= 第5蛋白酶位点
= 第6蛋白酶位点
= 不可切割接头
= 10个不同Q点
= 环肽
A
B
C
四嗪
DBCO
反式
环辛烯
D
循环数
切割
片段
成像
0
1
(+第1蛋白酶)
2
(+第2蛋白酶)
3
(+第3蛋白酶)
4
(+第4蛋白酶)

图3

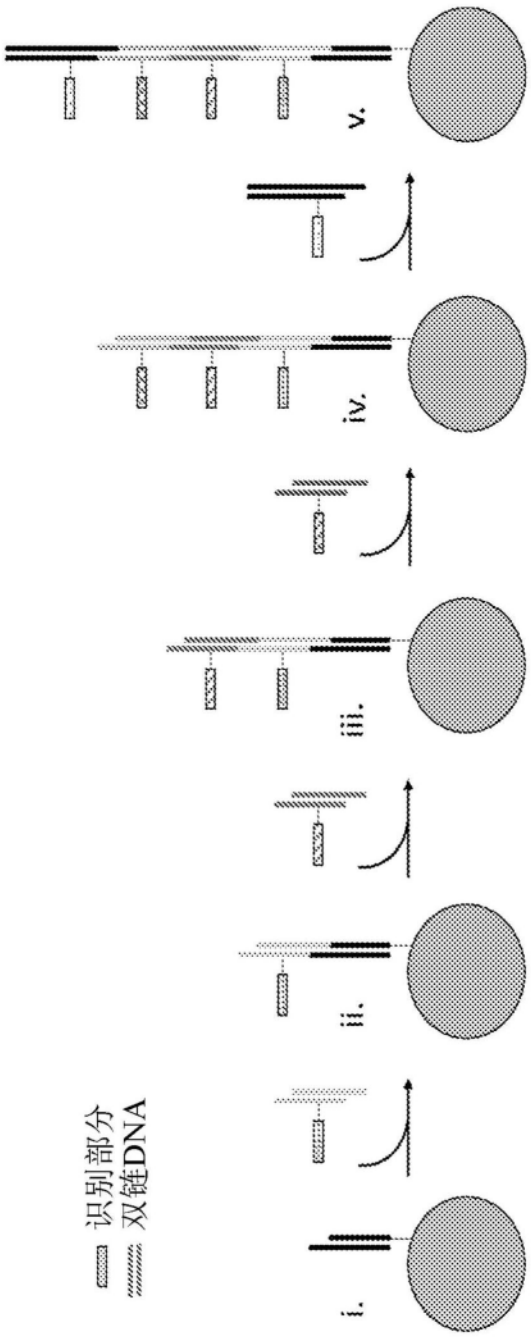


图4

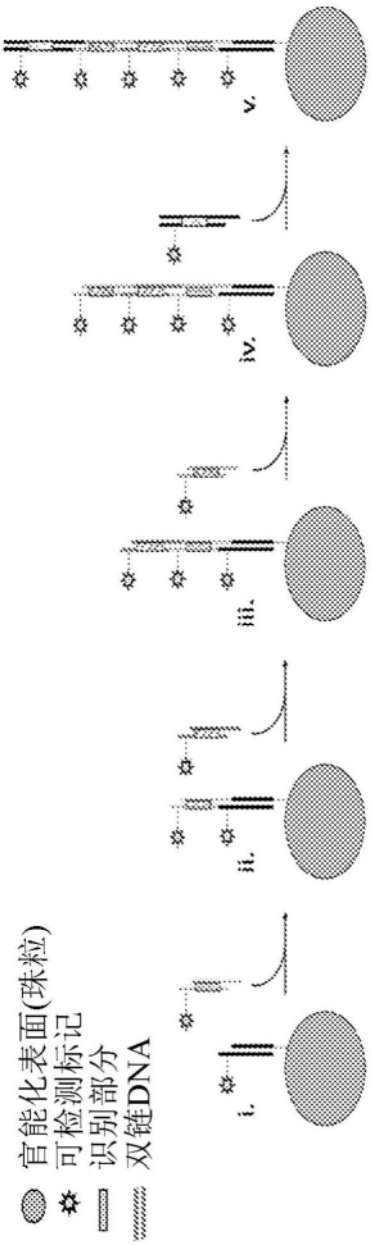


图5

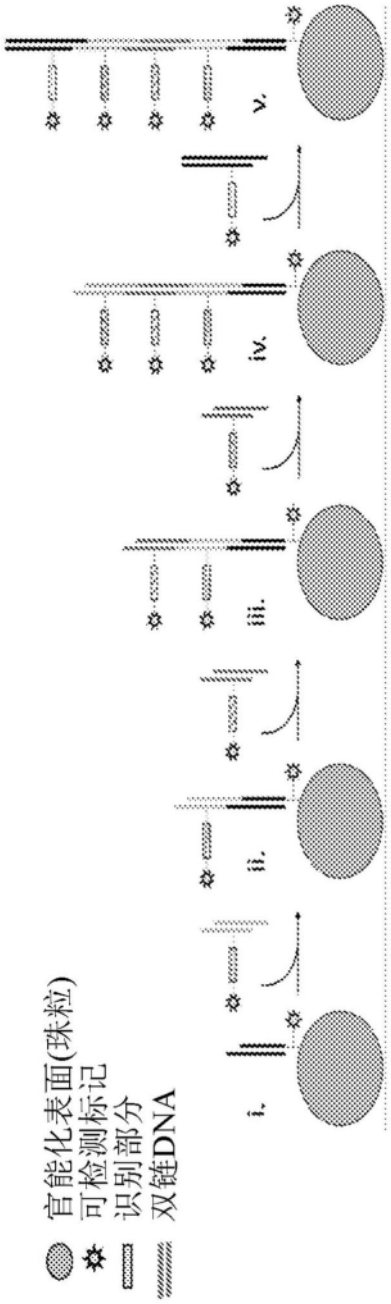


图6

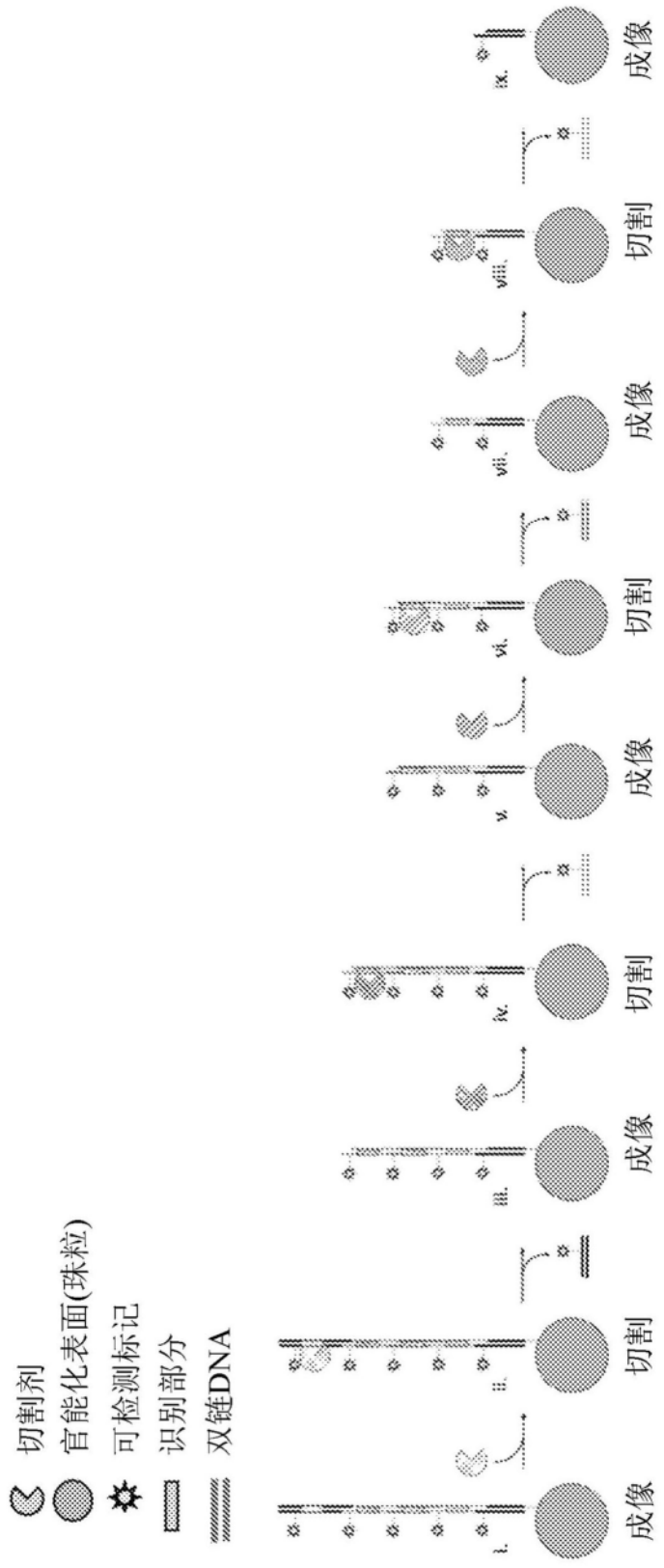


图7

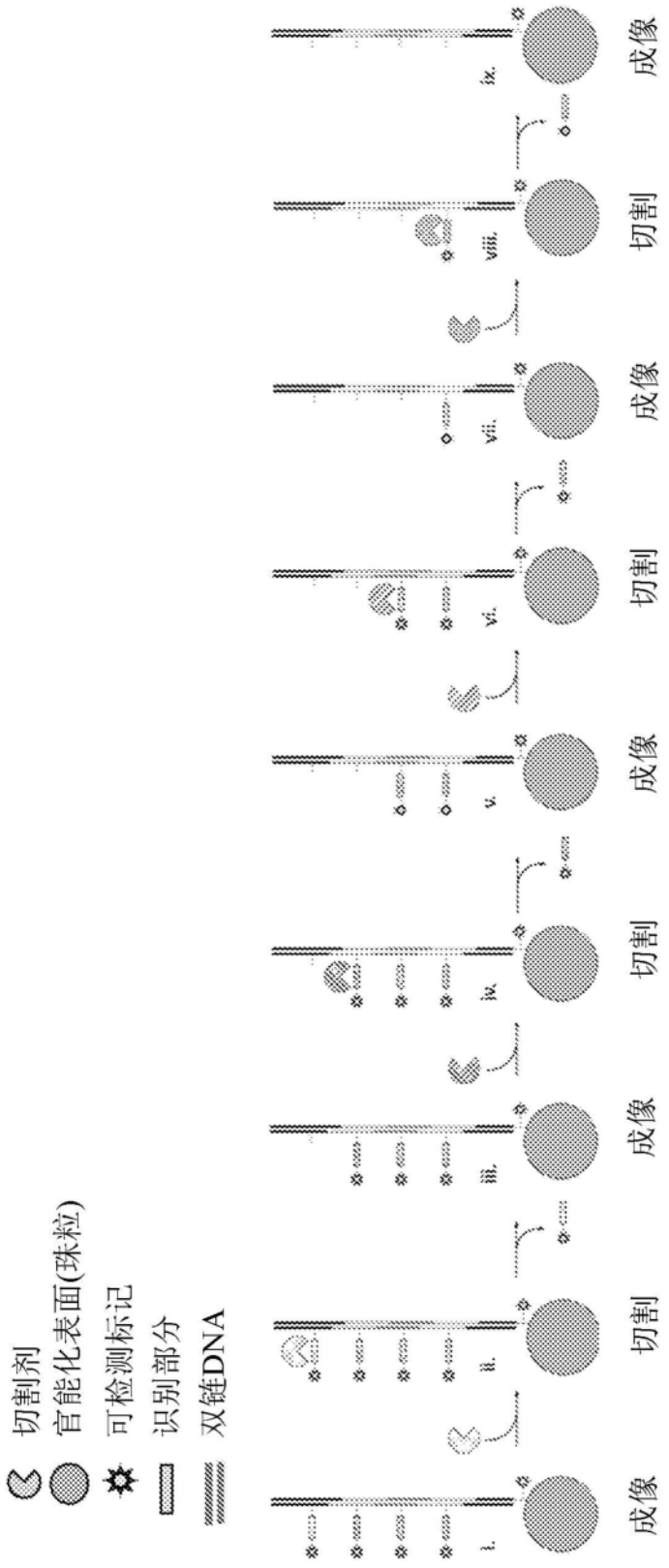

切割剂
官能化表面(珠粒)
可检测标记
识别部分
双链DNA
i.
ii.
iii.
iv.
v.
vi.
vii.
viii.
ix.
成像
切割
成像
切割
成像
切割
成像
切割
成像

图8

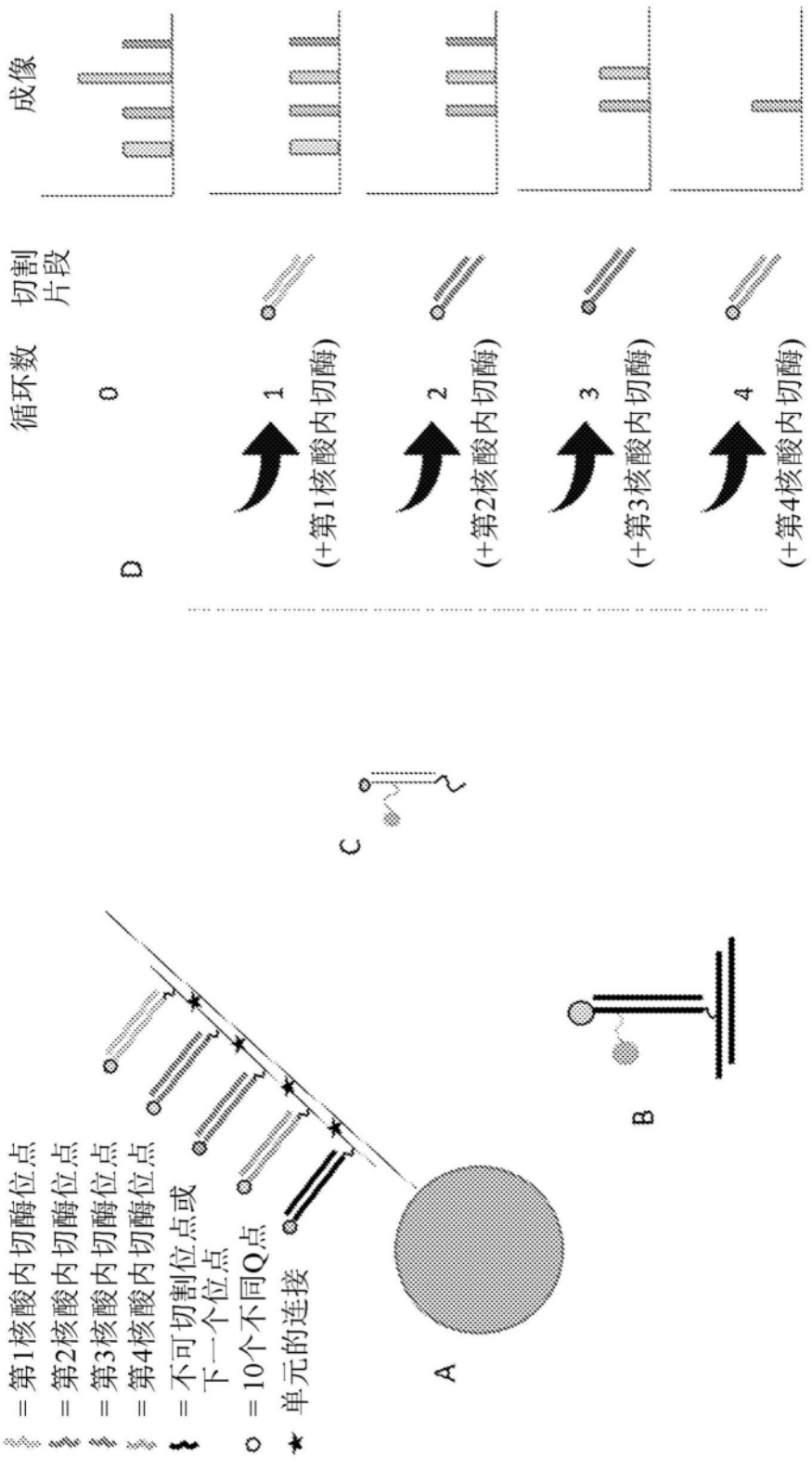

= 第1核酸内切酶位点
= 第2核酸内切酶位点
= 第3核酸内切酶位点
= 第4核酸内切酶位点
= 不可切割位点或下一个位点
= 10个不同Q点
单元的连接
A
B
C
D
循环数
切割片段
成像
0
1
(+第1核酸内切酶)
2
(+第2核酸内切酶)
3
(+第3核酸内切酶)
4
(+第4核酸内切酶)

图9

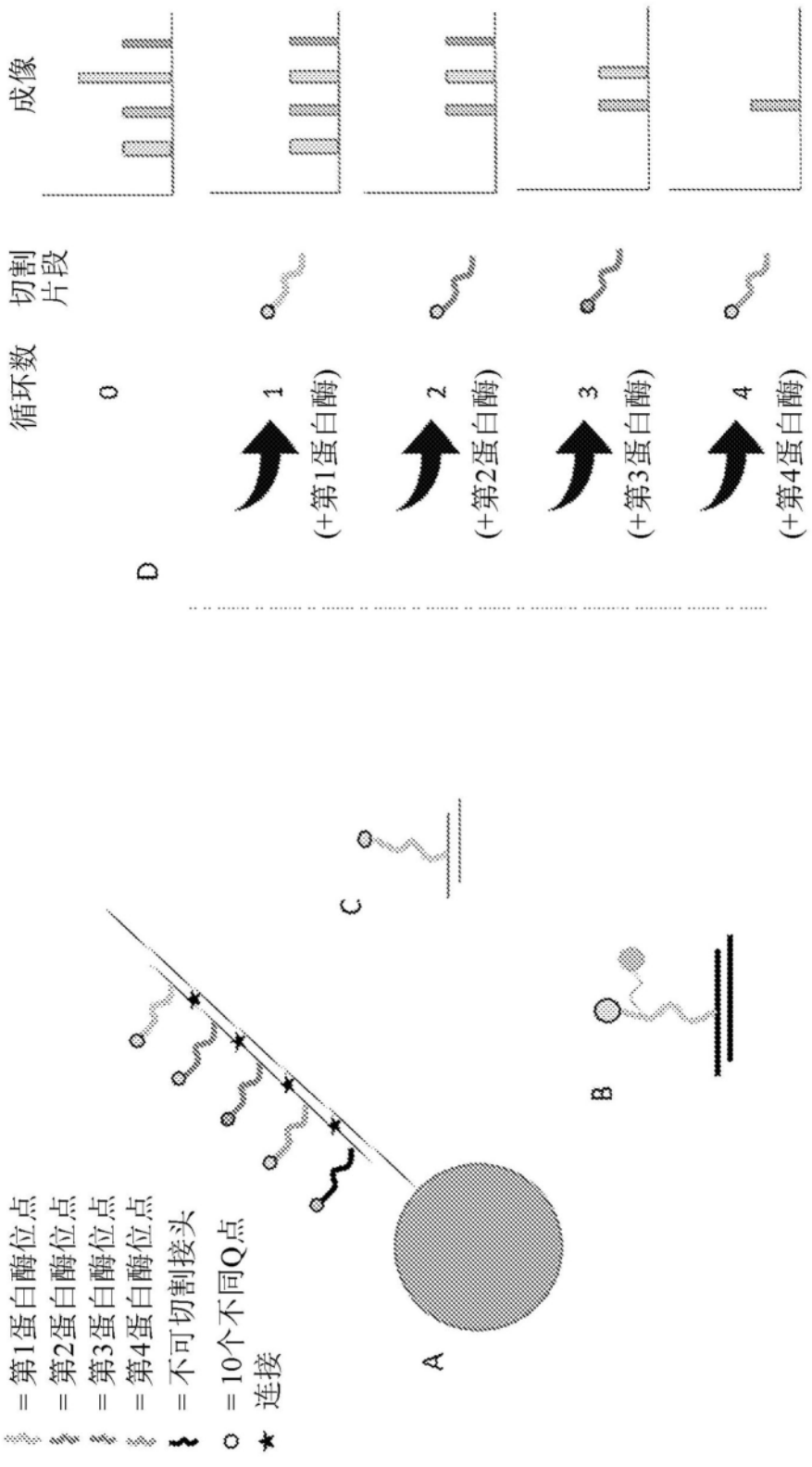

= 第1蛋白酶位点
= 第2蛋白酶位点
= 第3蛋白酶位点
= 第4蛋白酶位点
= 不可切割接头
= 10个不同Q点
连接
A
B
C
D
循环数
切割片段
成像
0
1
(+第1蛋白酶)
2
(+第2蛋白酶)
3
(+第3蛋白酶)
4
(+第4蛋白酶)

图10

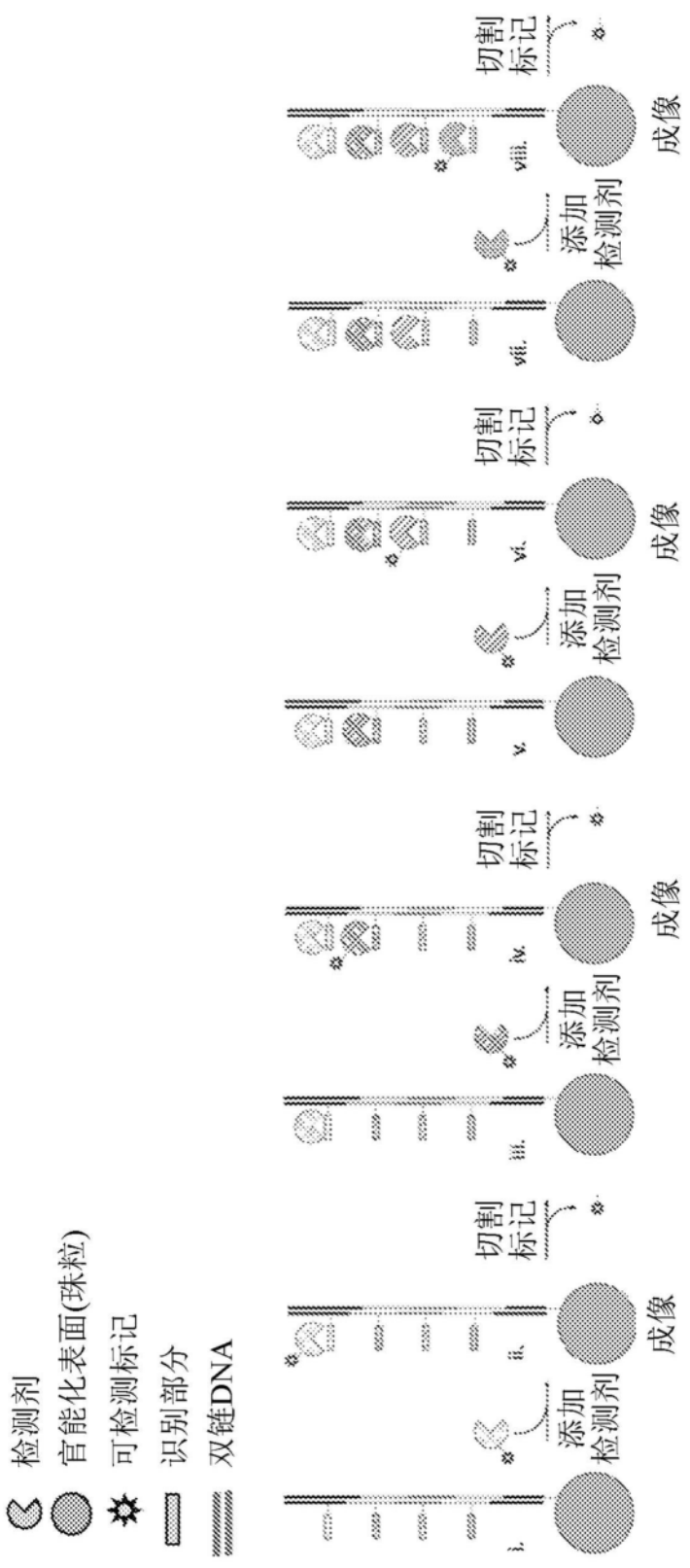

检测剂
官能化表面(珠粒)
可检测标记
识别部分
双链DNA
i.
添加检测剂
切割标记
成像
ii.
添加检测剂
切割标记
成像
iii.
添加检测剂
切割标记
成像
iv.
添加检测剂
切割标记
成像
v.
添加检测剂
切割标记
成像
vi.
添加检测剂
切割标记
成像
vii.
添加检测剂
切割标记
成像
viii.

图11

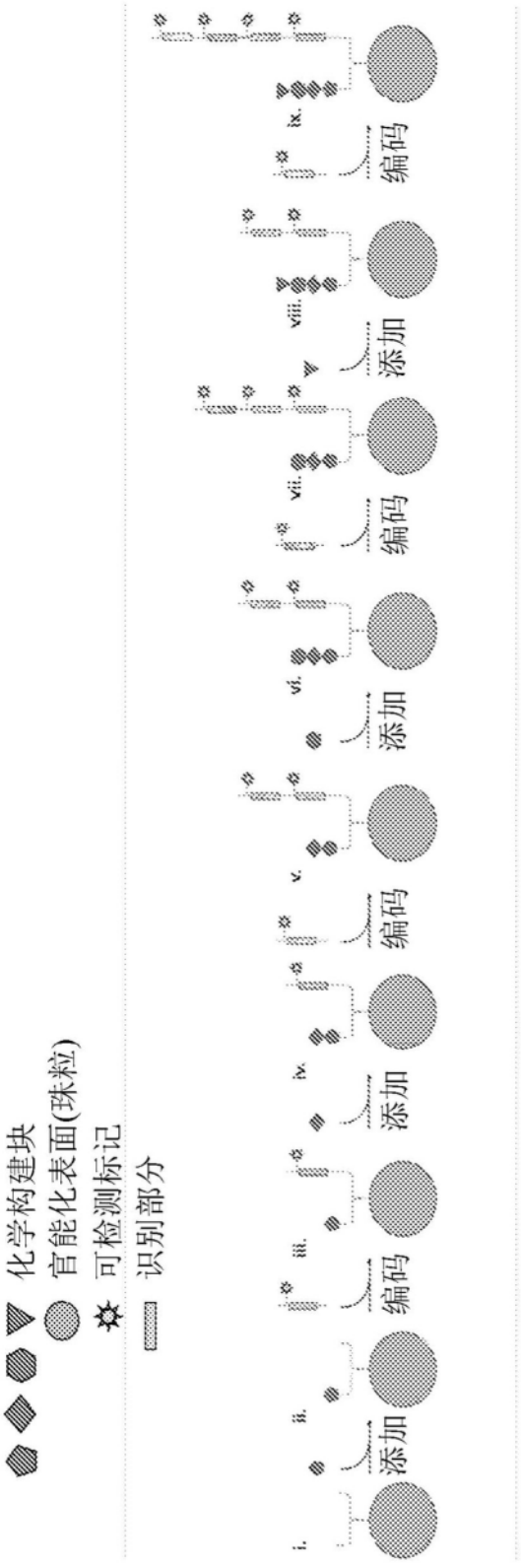


图12

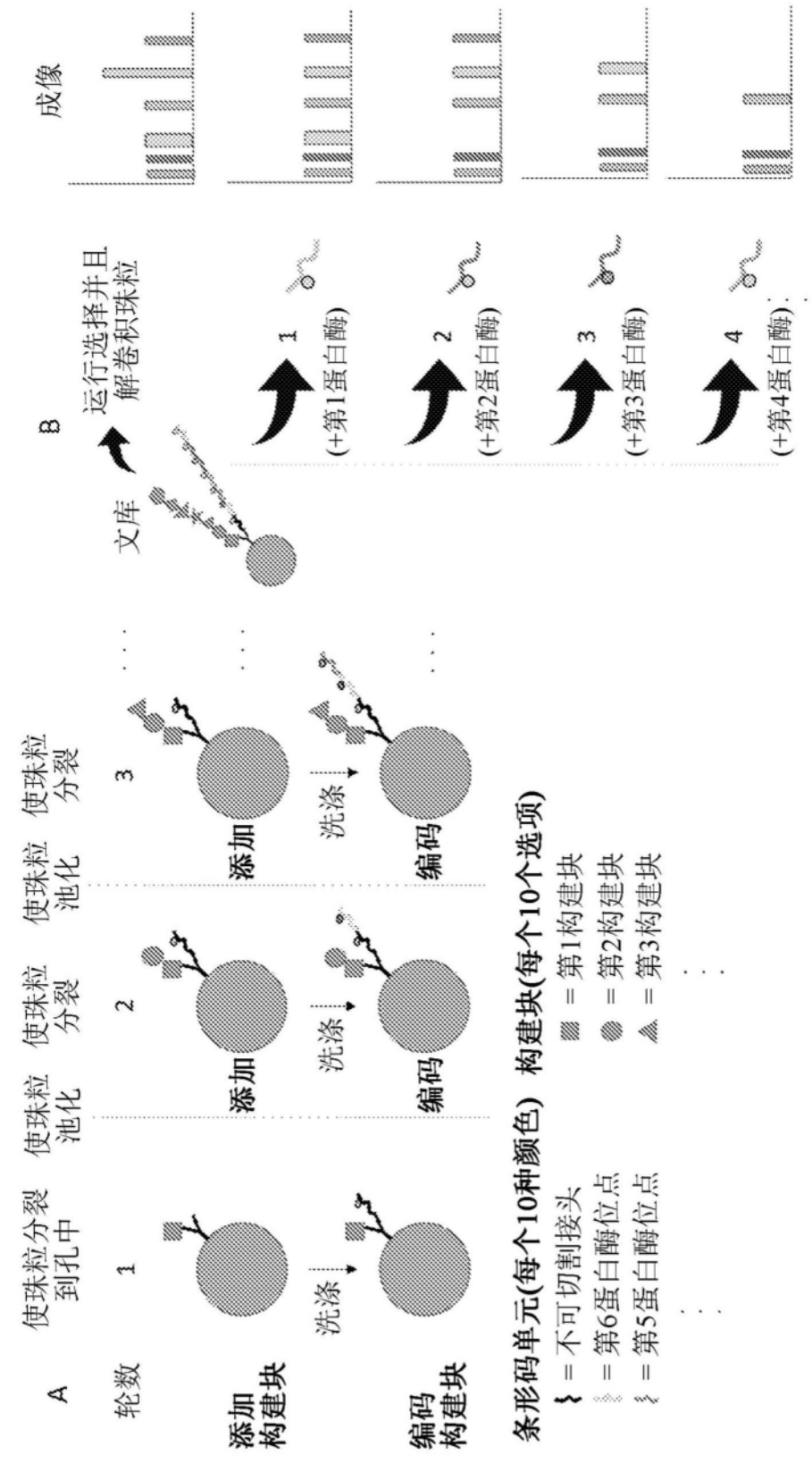

A
使珠粒分裂到孔中
使珠粒池化
使珠粒分裂
使珠粒池化
使珠粒分裂
轮数
1
2
3
添加构建块
编码构建块
洗涤
添加
编码
洗涤
添加
编码
条形码单元(每个10种颜色)
构建块(每个10个选项)
= 不可切割接头
= 第6蛋白酶位点
= 第5蛋白酶位点
= 第1构建块
= 第2构建块
= 第3构建块
B
运行选择并且解卷积珠粒
文库
成像
1
(+第1蛋白酶)
2
(+第2蛋白酶)
3
(+第3蛋白酶)
4
(+第4蛋白酶)

图13

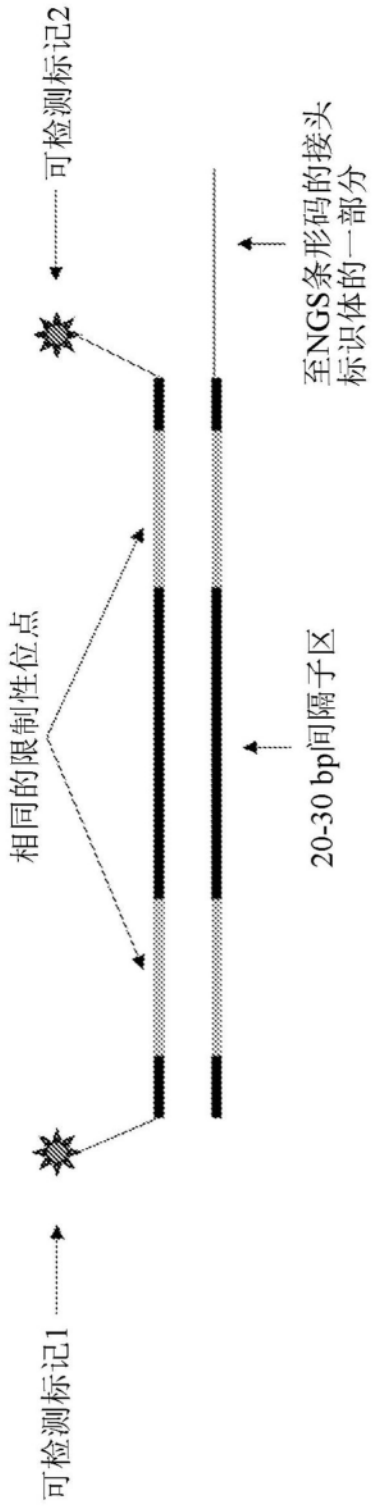


图14

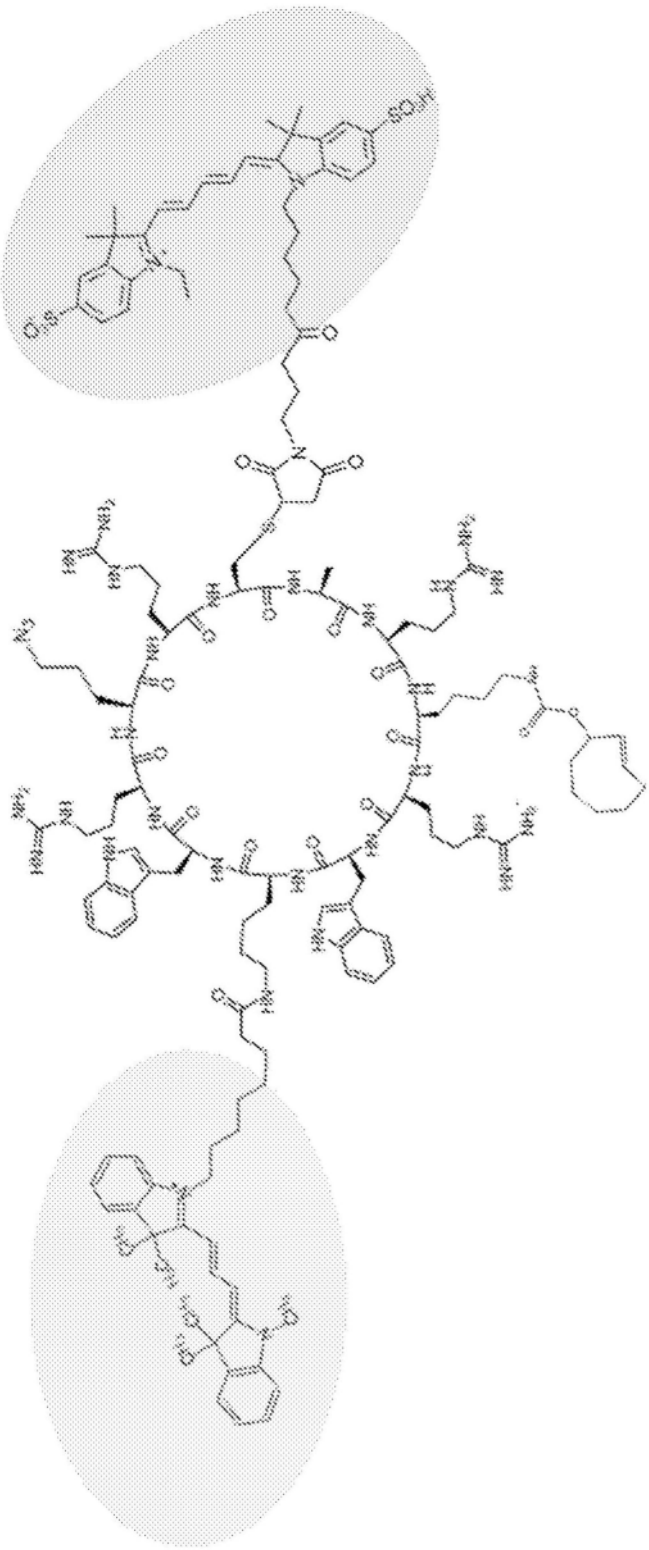


图15

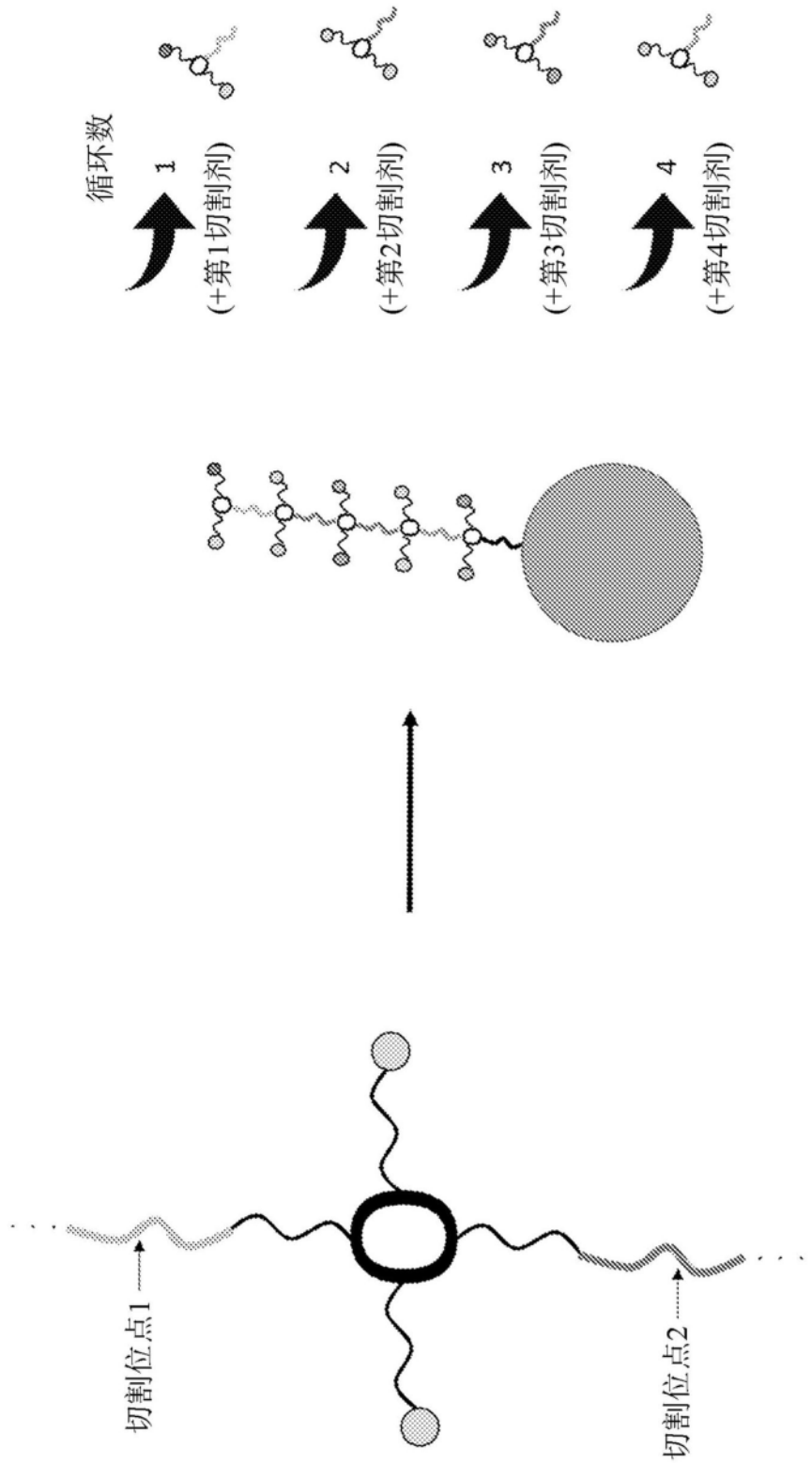


图16

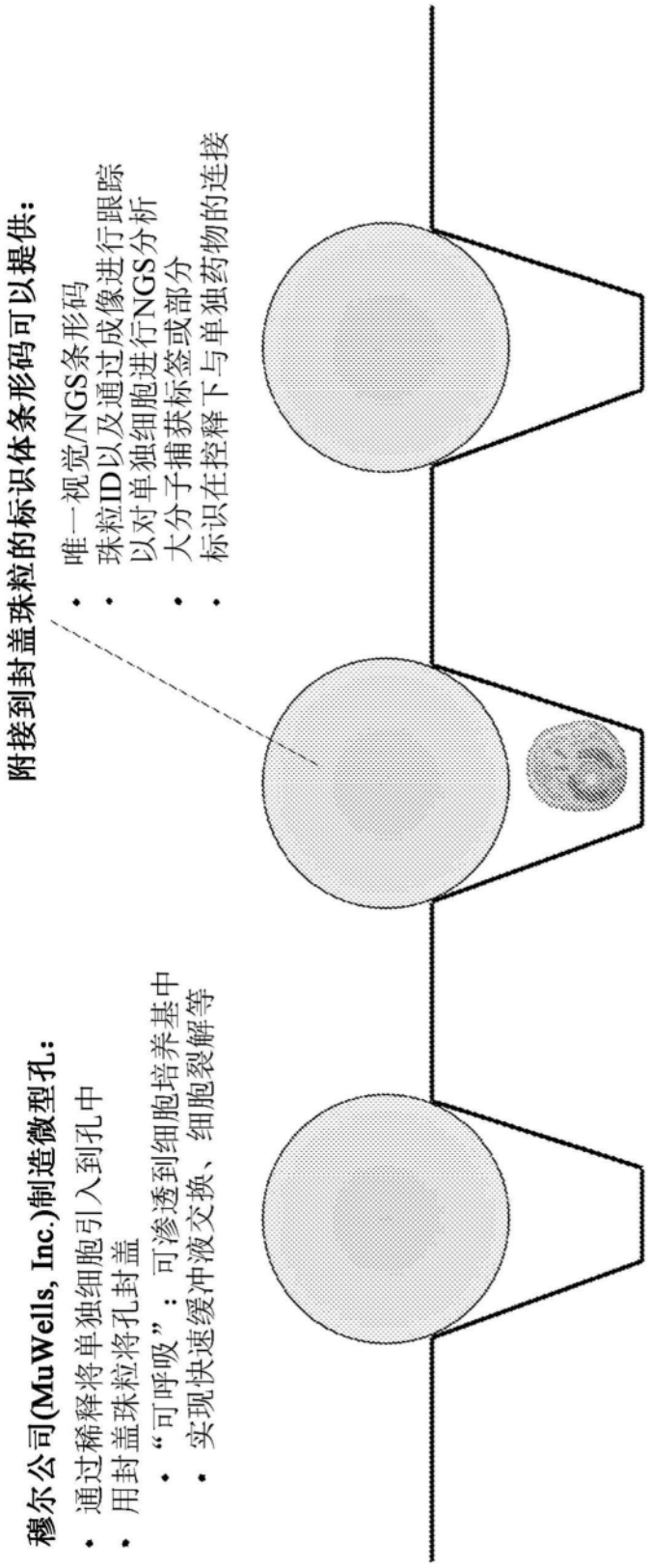

穆尔公司(MuWells, Inc.)制造微型孔:
• 通过稀释将单独细胞引入到孔中
• 用封盖珠粒将孔封盖
• “可呼吸”：可渗透到细胞培养基中
• 实现快速缓冲液交换、细胞裂解等
附接到封盖珠粒的标识体条形码可以提供:
• 唯一视觉/NGS条形码
• 珠粒ID以及通过成像进行跟踪以对单独细胞进行NGS分析
• 大分子捕获标签或部分
• 标识在控释下与单独药物的连接
数量：微型孔可以呈96孔板形式
96孔板的每个孔约22,000个微型孔
在95%置信度下对于单独细胞约10%负载
每个96孔板约200,000个单独细胞

图17

FND Id_1-生物素未标记的

/5BiotinTEG/ATTATTTTATTAGCTAGTACGCTACGACGTC

FND Id_1互补的

/5Phos/[TGCAG]GACGTCGTAGCGTACTAGC/3AmMC6T/

FND Id_2未标记的HindIII

/5Phos/[CTGCA]CGACTGC{AAGCTT}GATACTAGCA

FND Id_2互补的

/5Phos/[ACACT]TGC/iAmMC6T/AGTATC{AAGCTT}GCAGTCG

FND Id_3未标记的SpeI

/5Phos/[AGTGT]CGACTGCACTAGTGATACTAGCA

FND Id_3互补的

/5Phos/[TATGC]TGC/iAmMC6T/AGTATC{ACTAGT}GCAGTCG

FND Id_4未标记的XhoI

/5Phos/[GCATA]CGACTGC{CTCGAG}GATACTAGCA

FND Id_4互补的

/5Phos/[ATGGA]TGC/iAmMC6T/AGTATC{CTCGAG}GCAGTCG

FND Id_5未标记的NotI

/5Phos/[TCCAT]CGACT{GCGGCCGC}GATACTAGCA

FND Id_5互补的

/5AmMC6T/GCTAGTATC{GCGGCCGC}AGTCG

- 注意：
  - “[]”括号格式分组表示形成粘性端的[核苷酸]和相关[连接剂]
  - “{}”括号格式分组表示形成限制性位点的[核苷酸]和相关[切割剂]

图18

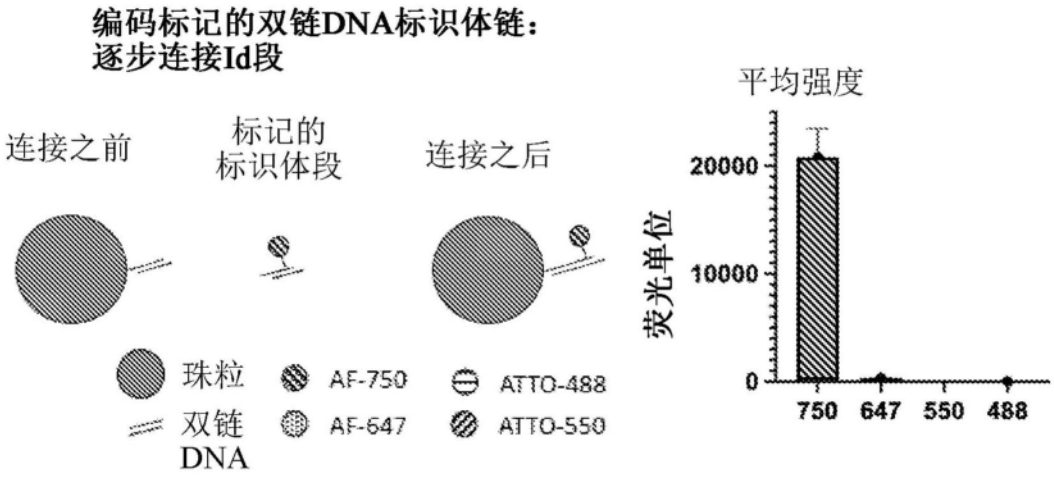


图19A

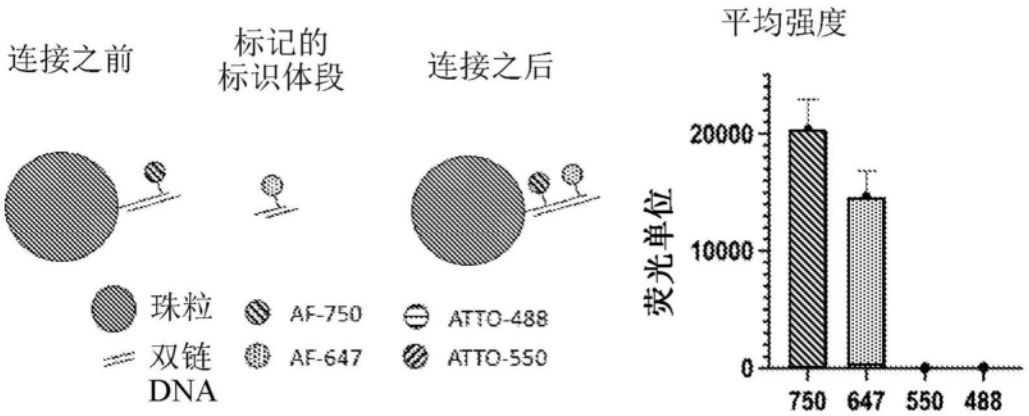


图19B

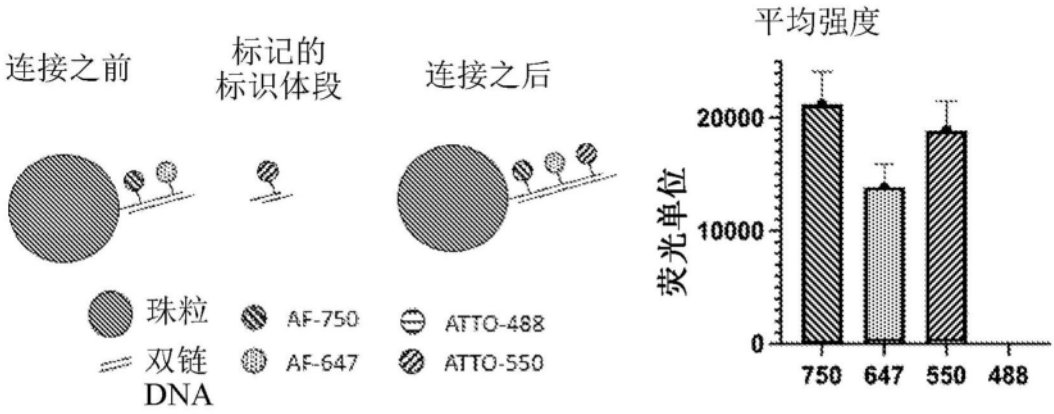


图19C

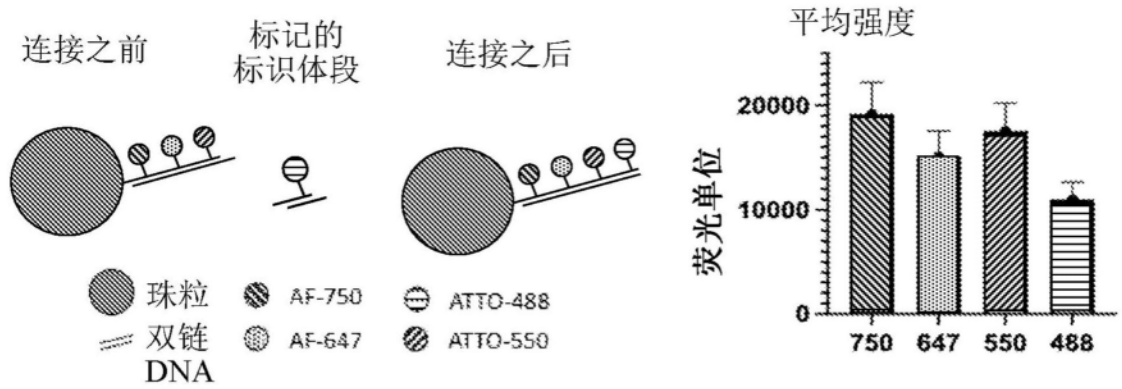


图19D

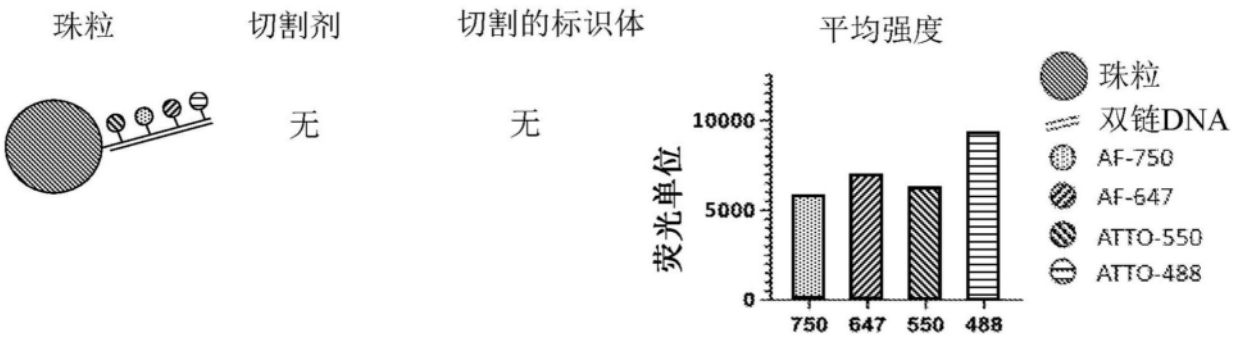


图20A

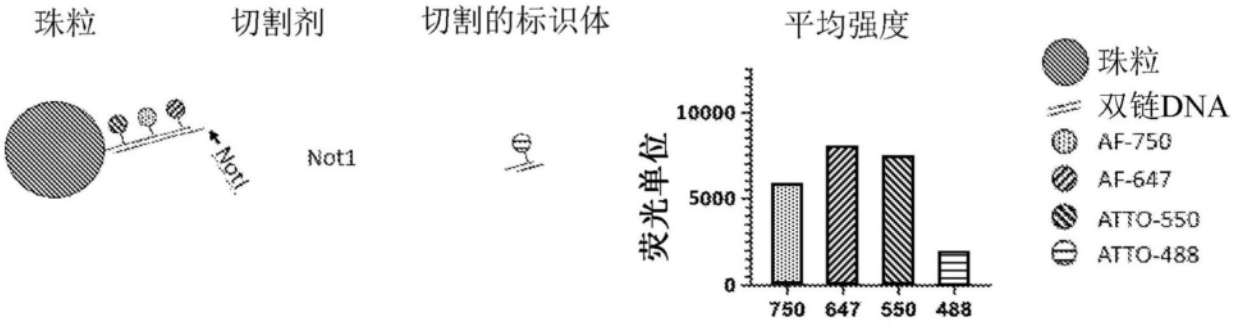


图20B

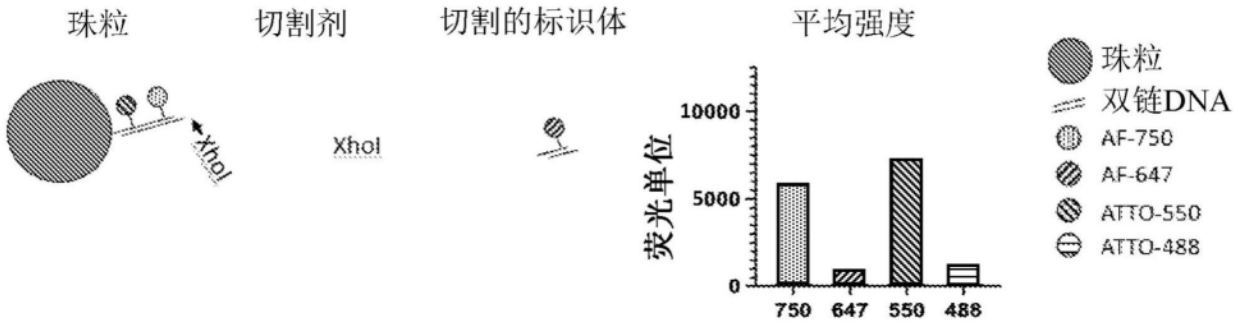


图20C

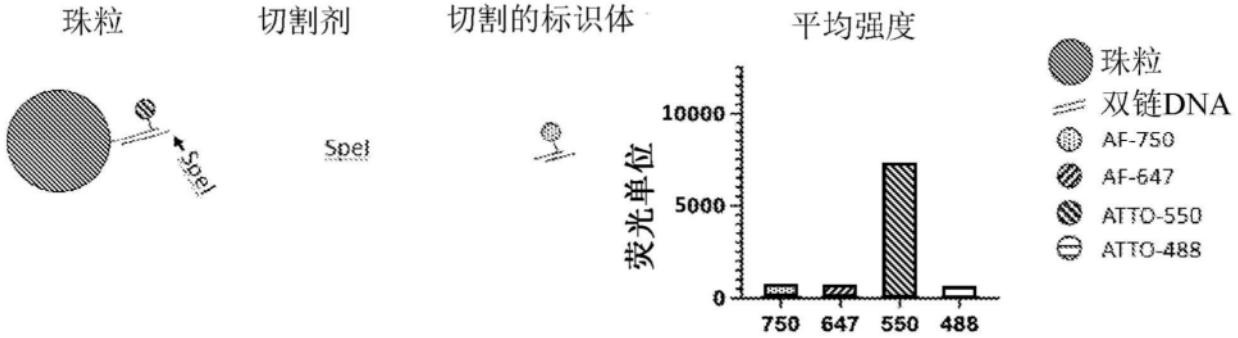


图20D

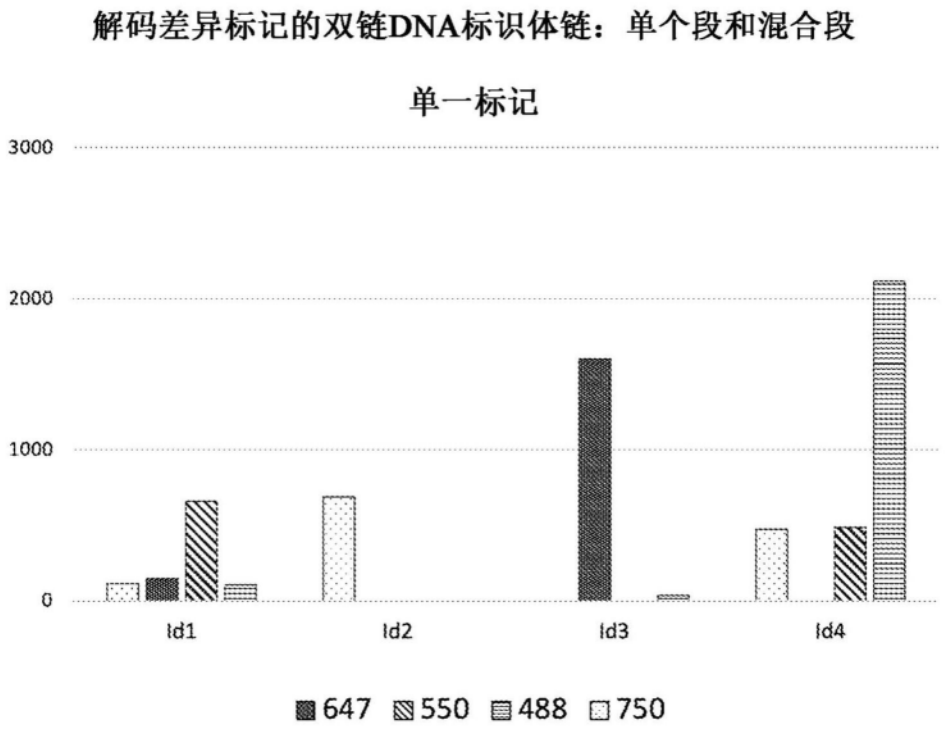


图21A

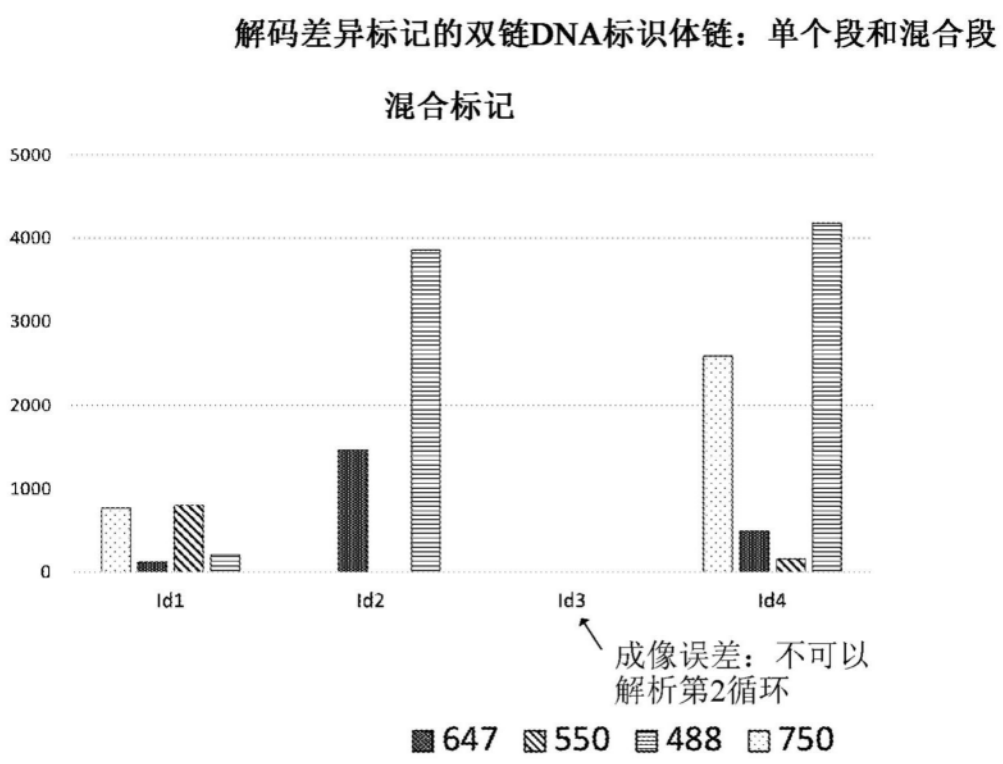

解码差异标记的双链DNA标识体链：单个段和混合段
混合标记
5000
4000
3000
2000
1000
0
Id1
Id2
Id3
Id4
成像误差：不可以
解析第2循环
647
550
488
750

图21B

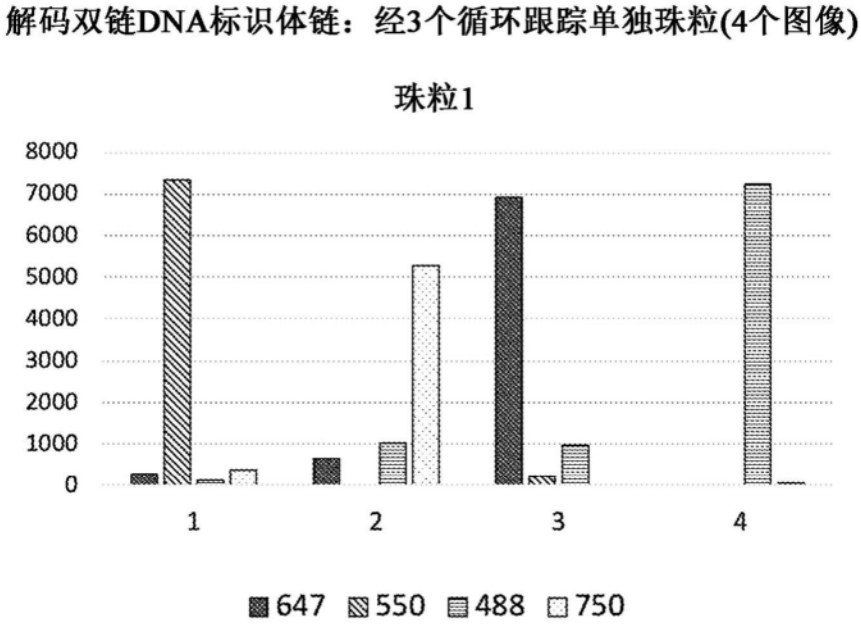

解码双链DNA标识体链：经3个循环跟踪单独珠粒(4个图像)
珠粒1
8000
7000
6000
5000
4000
3000
2000
1000
0
1
2
3
4
647
550
488
750

图22A

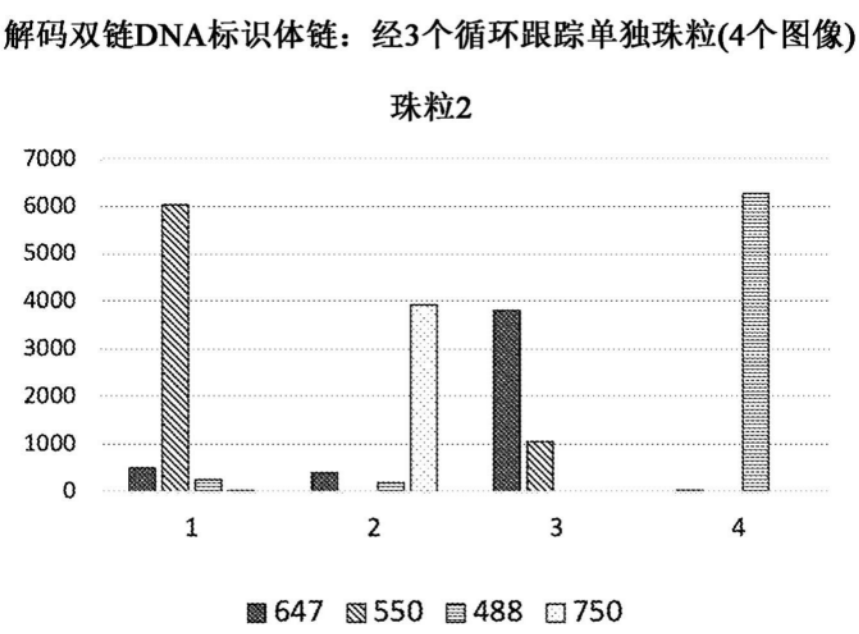

解码双链DNA标识体链：经3个循环跟踪单独珠粒(4个图像)
珠粒2
7000
6000
5000
4000
3000
2000
1000
0
1
2
3
4
647
550
488
750

图22B

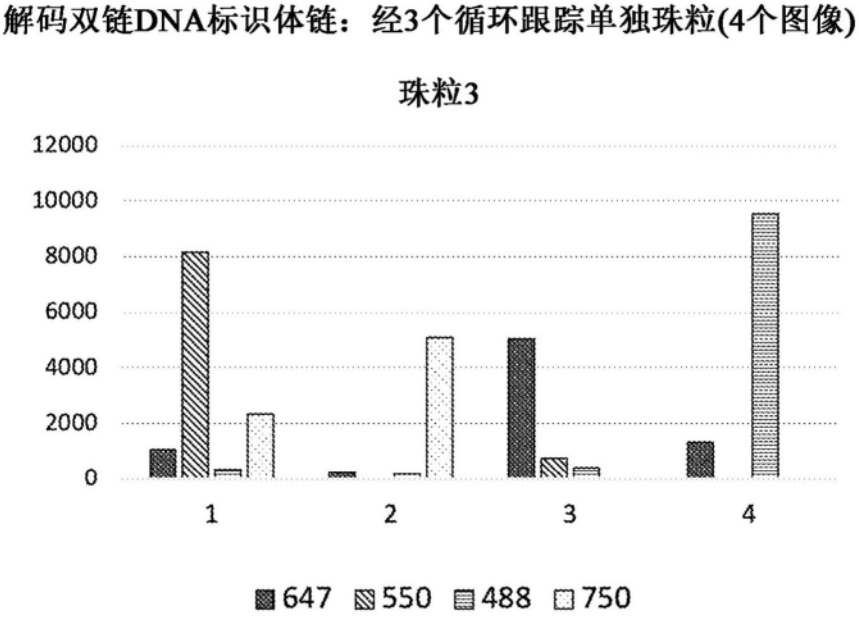

解码双链DNA标识体链：经3个循环跟踪单独珠粒(4个图像)
珠粒3
12000
10000
8000
6000
4000
2000
0
1
2
3
4
647
550
488
750

图22C

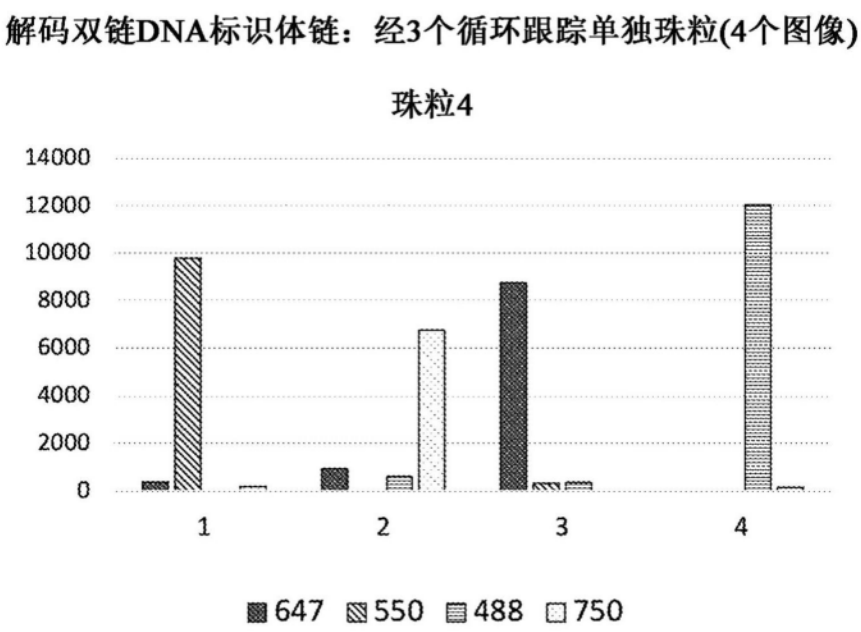

解码双链DNA标识体链：经3个循环跟踪单独珠粒(4个图像)
珠粒4
14000
12000
10000
8000
6000
4000
2000
0
1
2
3
4
647
550
488
750

图22D

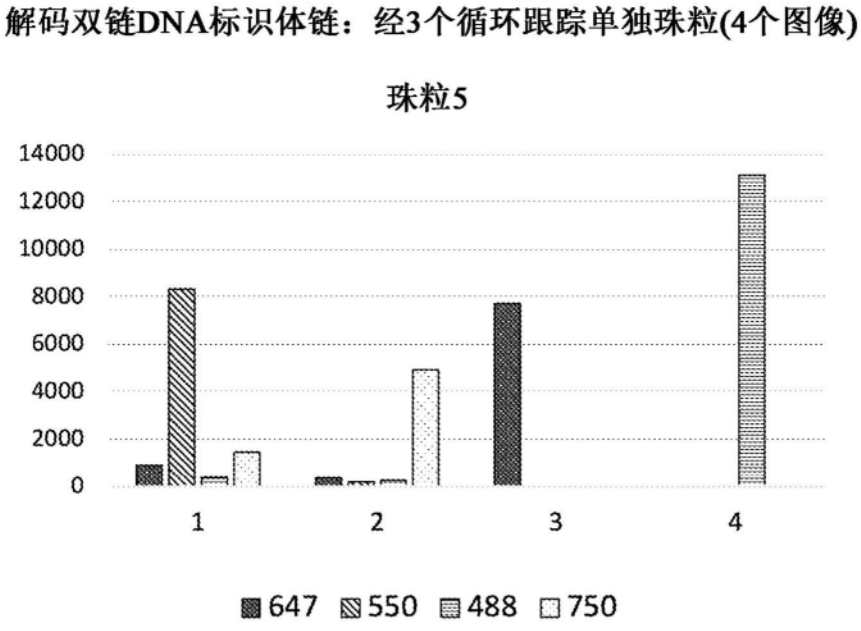

解码双链DNA标识体链：经3个循环跟踪单独珠粒(4个图像)
珠粒5
14000
12000
10000
8000
6000
4000
2000
0
1
2
3
4
647
550
488
750

图22E

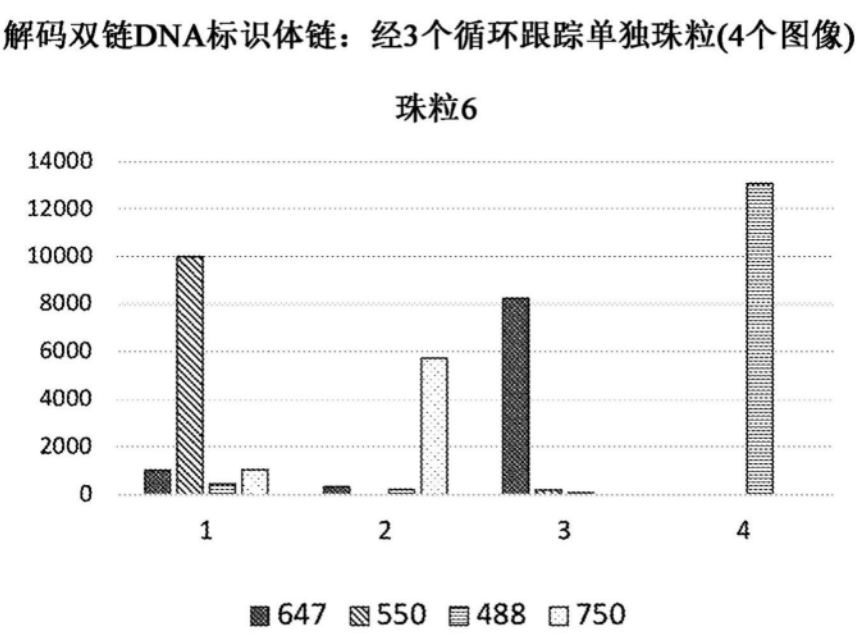

解码双链DNA标识体链：经3个循环跟踪单独珠粒(4个图像)
珠粒6
14000
12000
10000
8000
6000
4000
2000
0
1
2
3
4
647
550
488
750

图22F

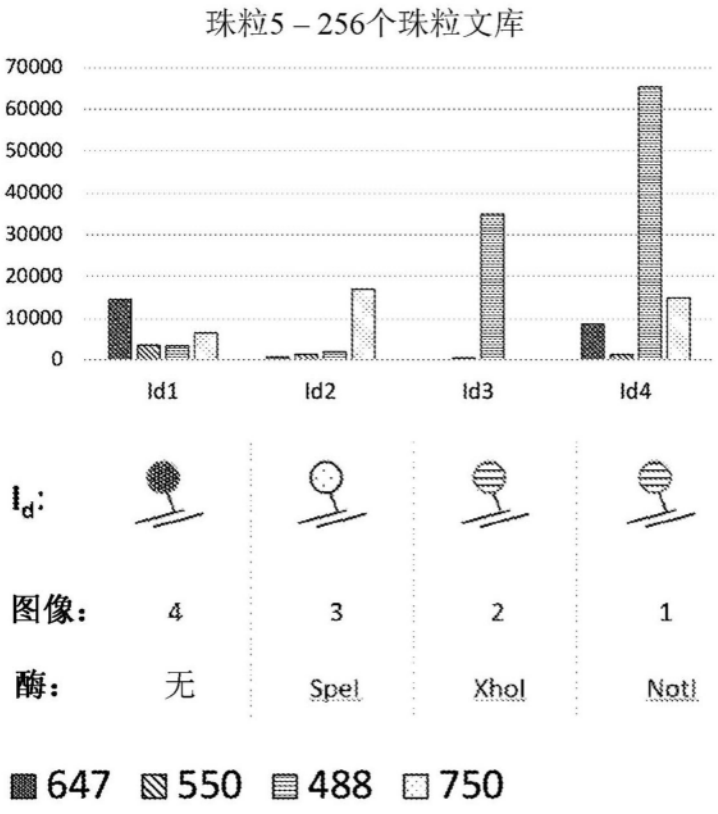

珠粒5 – 256个珠粒文库
70000
60000
50000
40000
30000
20000
10000
0
Id1
Id2
Id3
Id4
Id:
图像:
4
3
2
1
酶:
无
SpeI
XhoI
NotI
647
550
488
750

图23A

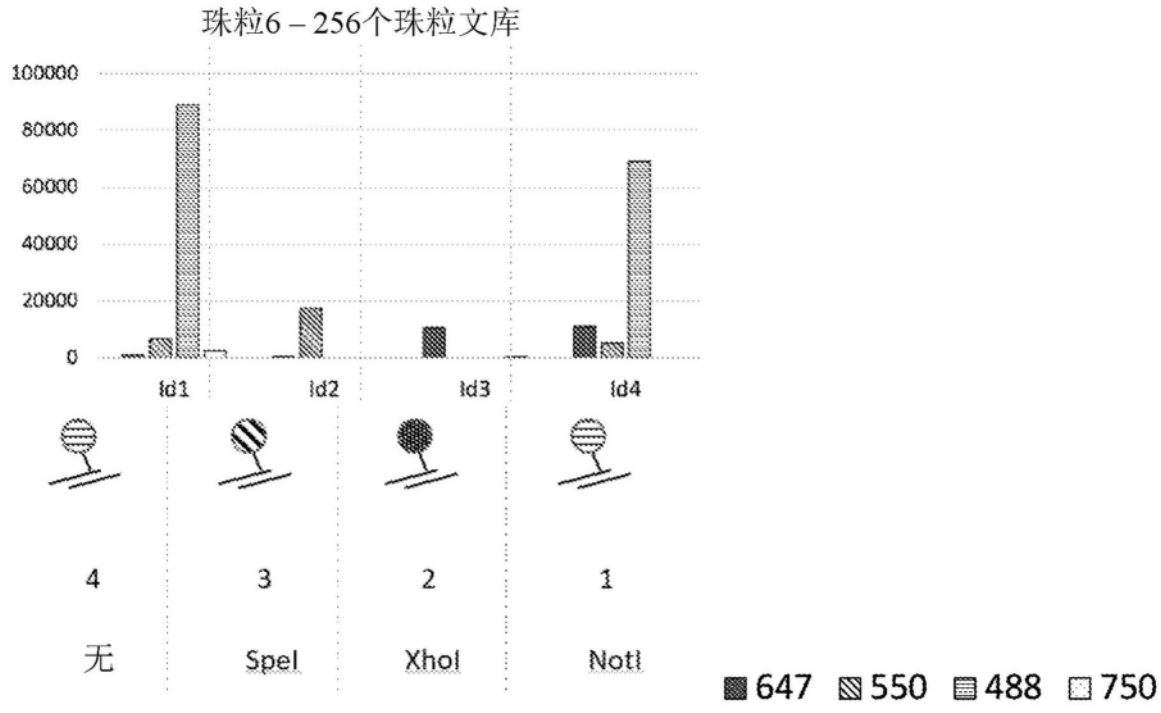

珠粒6 – 256个珠粒文库
100000
80000
60000
40000
20000
0
Id1
Id2
Id3
Id4
4
3
2
1
无
SpeI
XhoI
NotI
647
550
488
750

图23B

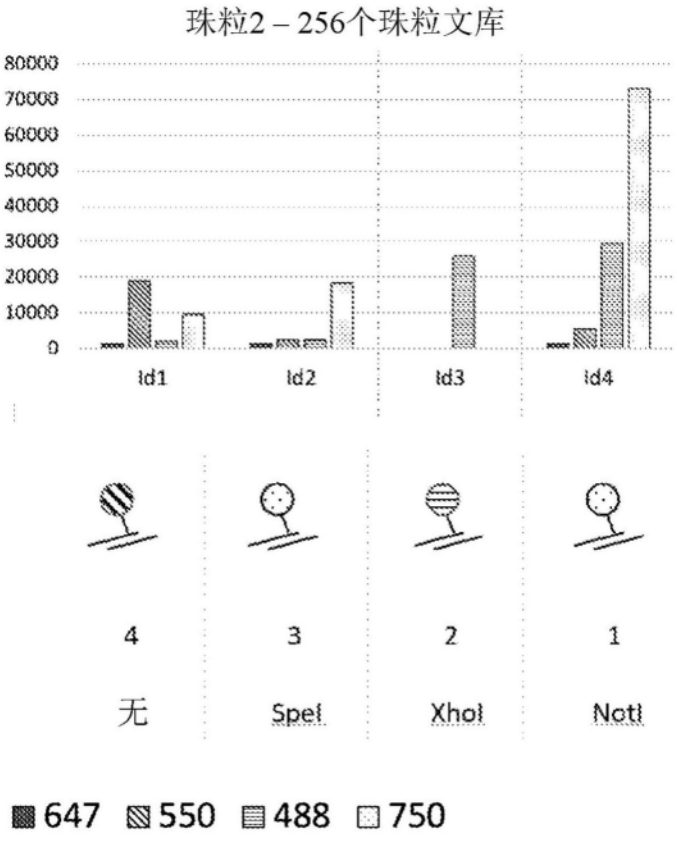

珠粒2 – 256个珠粒文库
80000
70000
60000
50000
40000
30000
20000
10000
0
Id1
Id2
Id3
Id4
4
3
2
1
无
SpeI
XhoI
NotI
647
550
488
750

图23C

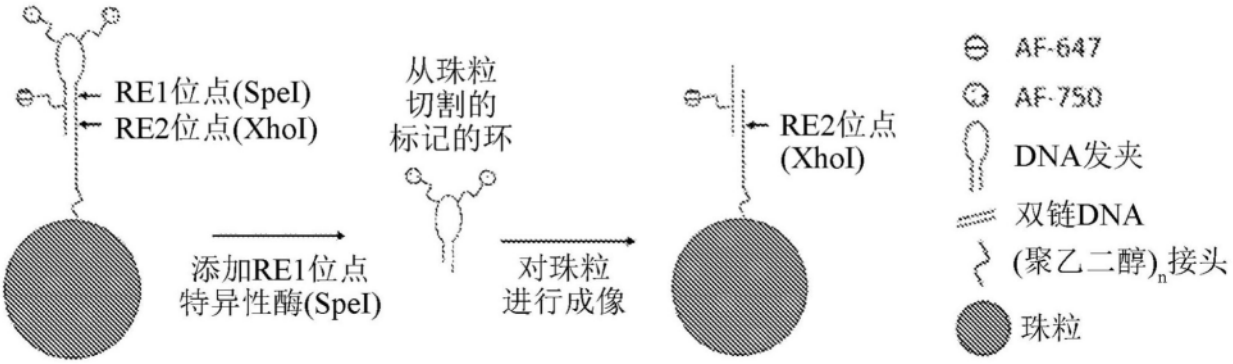

RE1位点(SpeI)
RE2位点(XhoI)
添加RE1位点
特异性酶(SpeI)
从珠粒
切割的
标记的环
对珠粒
进行成像
RE2位点
(XhoI)
AF-647
AF-750
DNA发夹
双链DNA
(聚乙二醇)n接头
珠粒

图24A

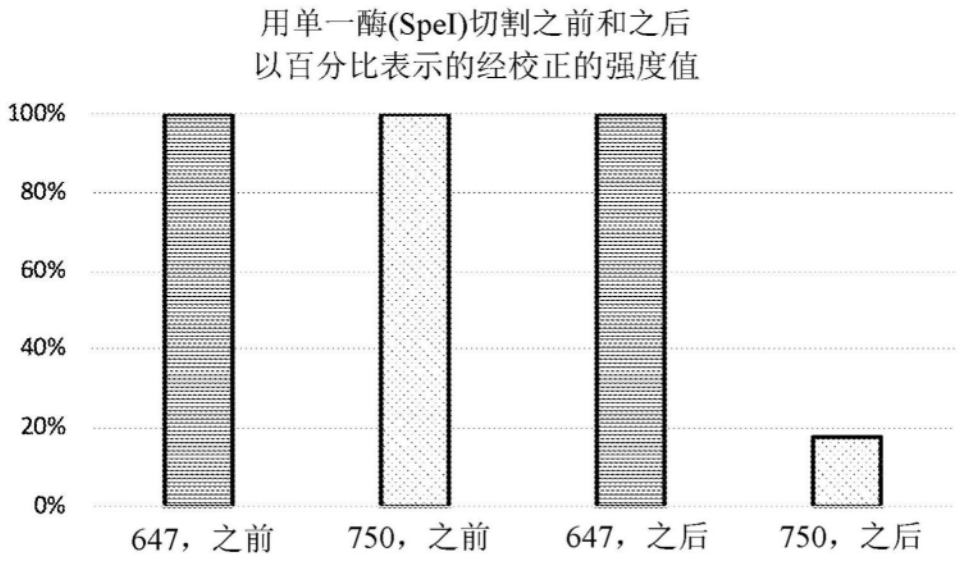

用单一酶(SpeI)切割之前和之后
以百分比表示的经校正的强度值
100%
80%
60%
40%
20%
0%
647，之前
750，之前
647，之后
750，之后

图24B

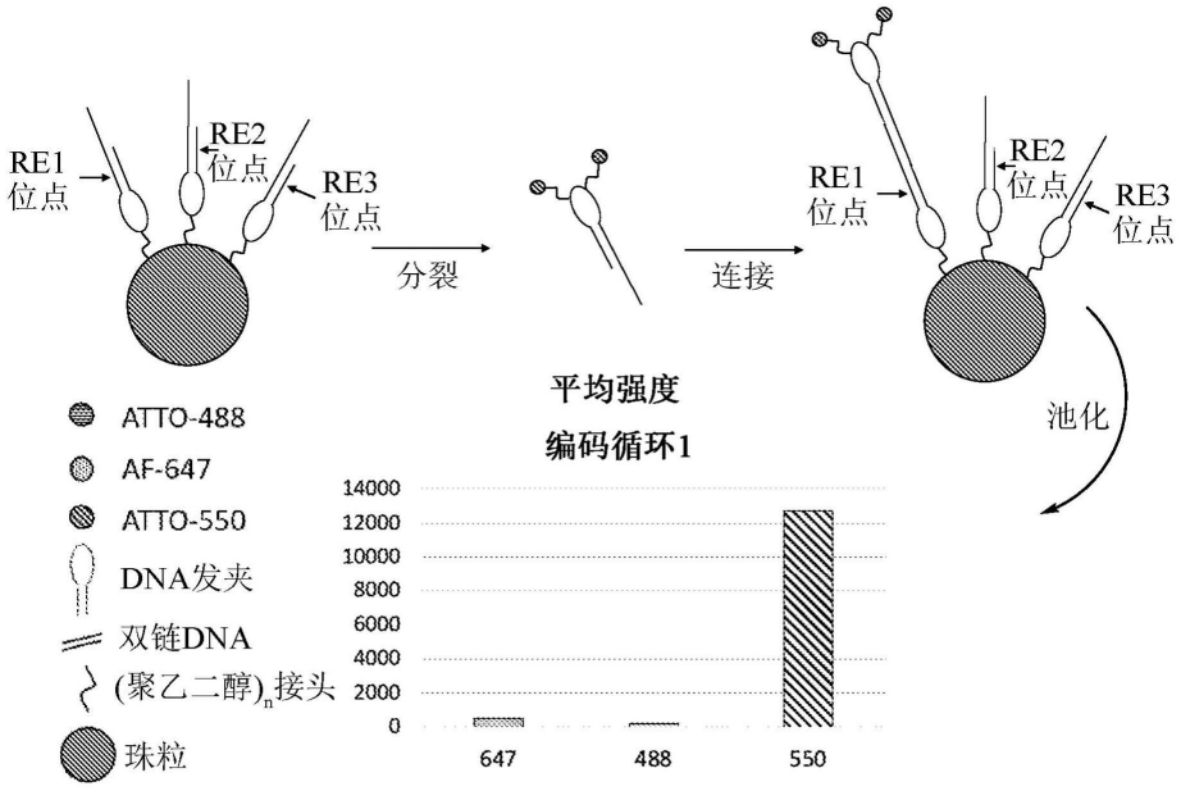

RE1
位点
RE2
位点
RE3
位点
分裂
连接
池化
ATTO-488
AF-647
ATTO-550
DNA发夹
双链DNA
(聚乙二醇)n接头
珠粒
平均强度
编码循环1
14000
12000
10000
8000
6000
4000
2000
0
647
488
550

图25A

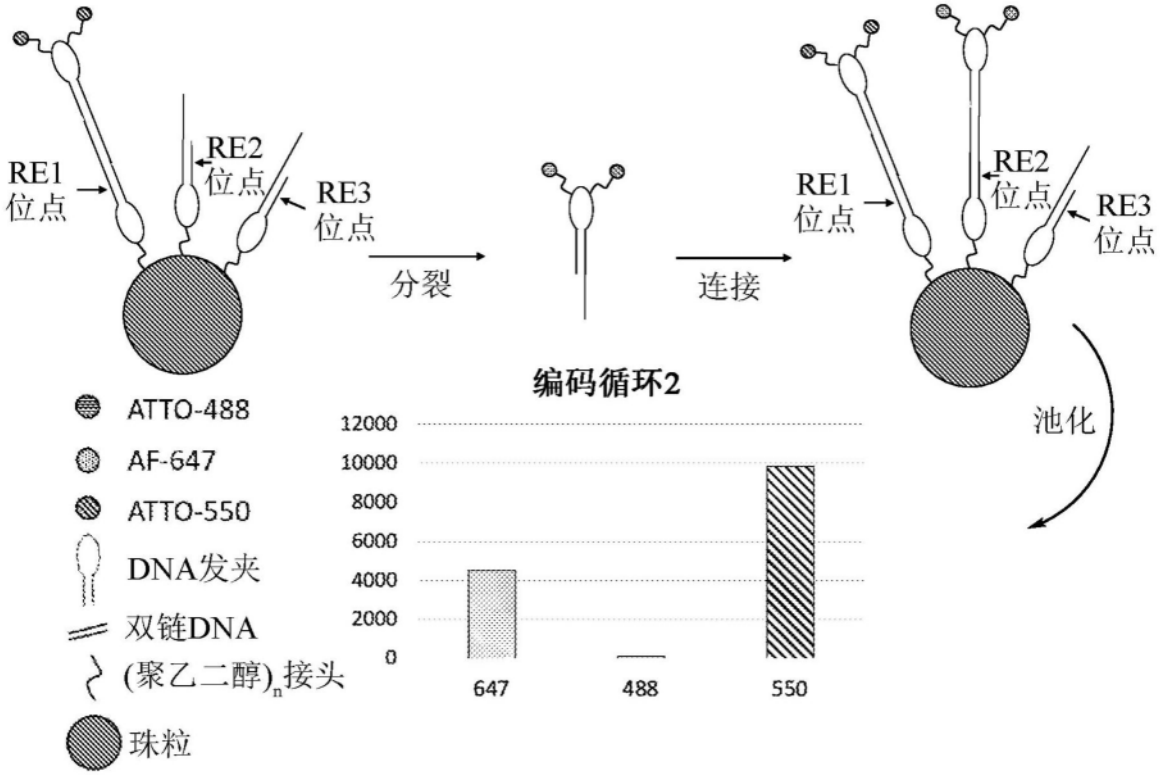

RE1
位点
RE2
位点
RE3
位点
分裂
连接
池化
ATTO-488
AF-647
ATTO-550
DNA发夹
双链DNA
(聚乙二醇)n接头
珠粒
编码循环2
12000
10000
8000
6000
4000
2000
0
647
488
550

图25B

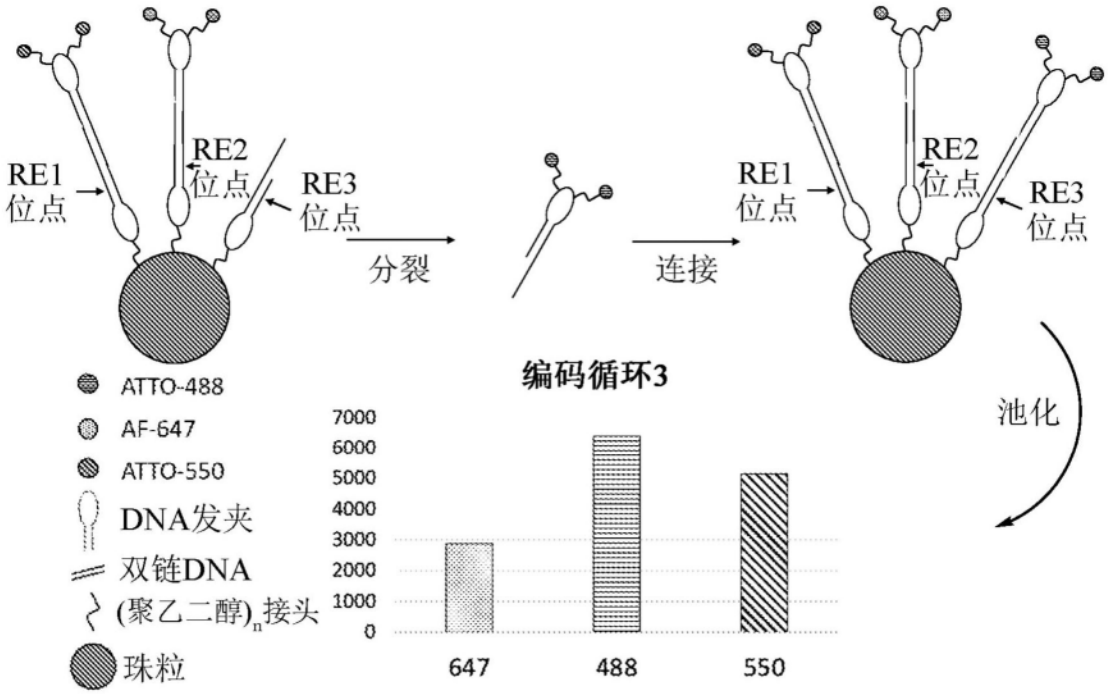

RE1
位点
RE2
位点
RE3
位点
分裂
连接
编码循环3
池化
ATTO-488
AF-647
ATTO-550
DNA发夹
双链DNA
(聚乙二醇)n接头
珠粒
7000
6000
5000
4000
3000
2000
1000
0
647
488
550

图25C

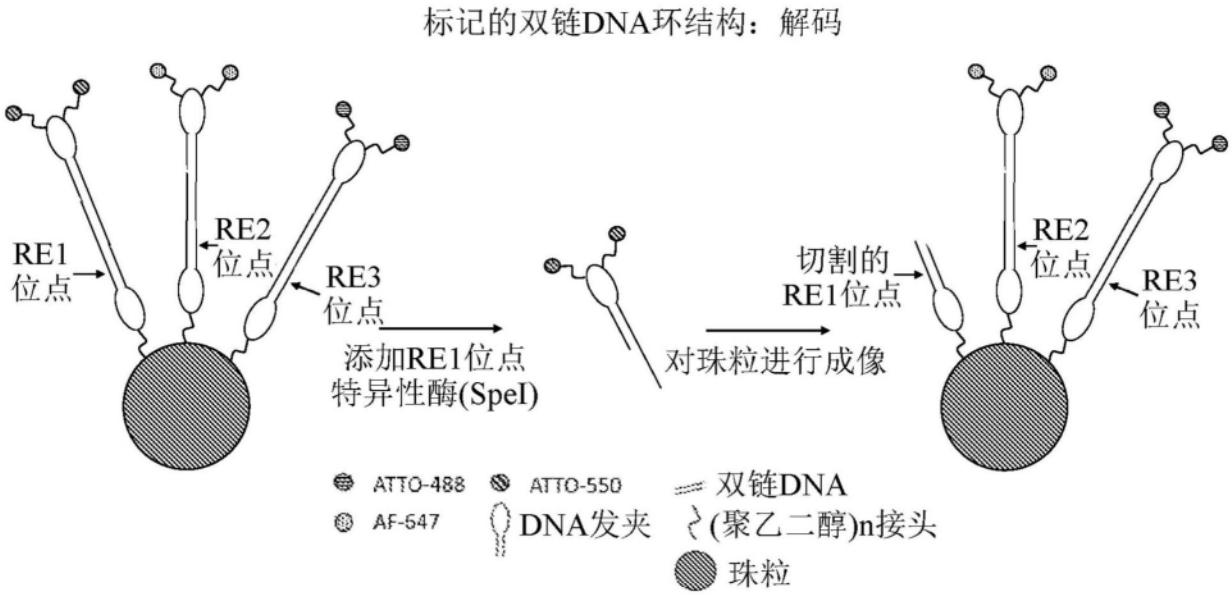

标记的双链DNA环结构：解码
RE1
位点
RE2
位点
RE3
位点
添加RE1位点
特异性酶(SpeI)
对珠粒进行成像
切割的
RE1位点
ATTO-488
ATTO-550
双链DNA
AF-647
DNA发夹
(聚乙二醇)n接头
珠粒

图26A

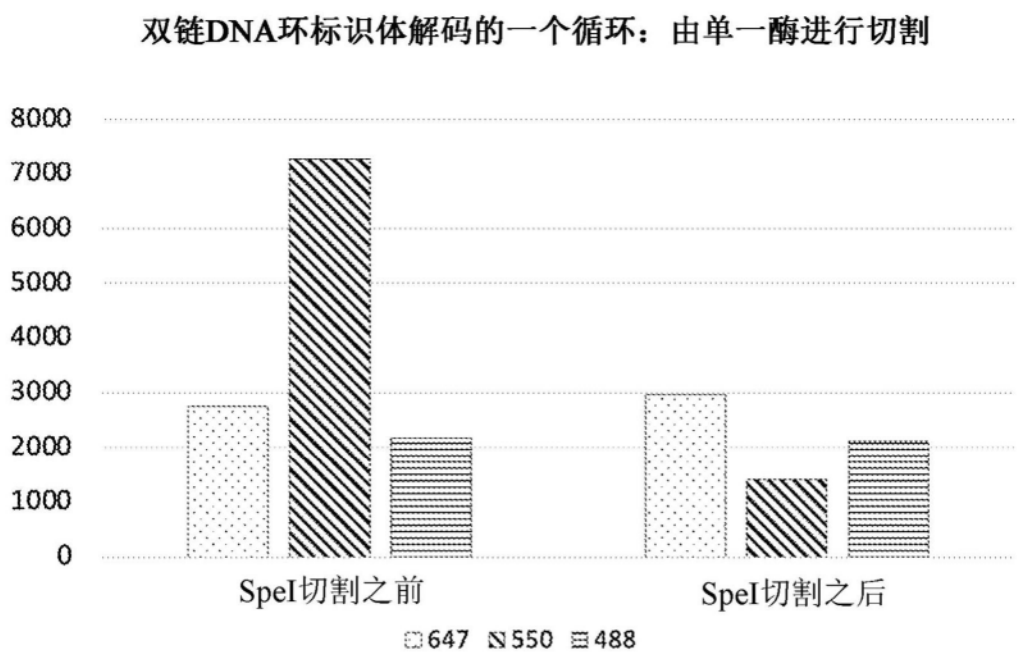

双链DNA环标识体解码的一个循环：由单一酶进行切割
8000
7000
6000
5000
4000
3000
2000
1000
0
SpeI切割之前
SpeI切割之后
647 550 488

图26B

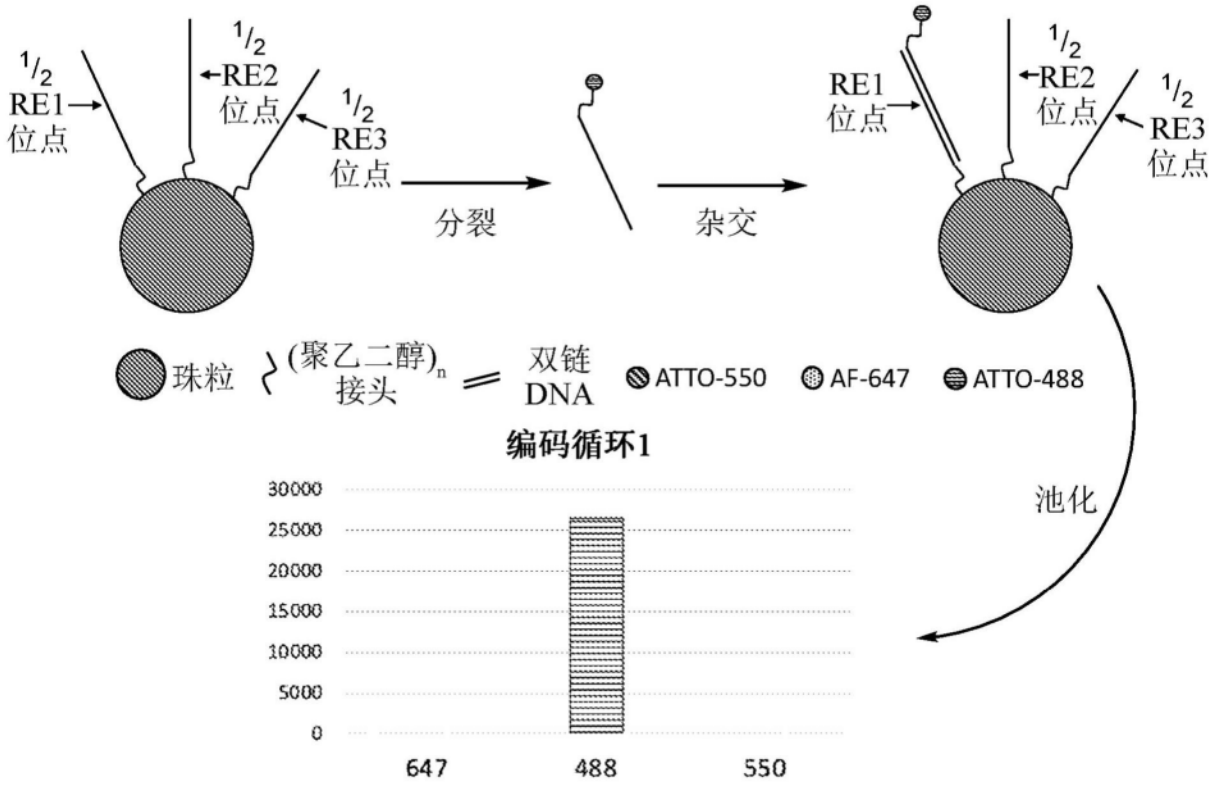

1/2 RE1 位点
1/2 RE2 位点
1/2 RE3 位点
分裂
杂交
RE1 位点
1/2 RE2 位点
1/2 RE3 位点
珠粒
(聚乙二醇)n 接头
双链 DNA
ATTO-550
AF-647
ATTO-488
池化
编码循环1
30000
25000
20000
15000
10000
5000
0
647
488
550

图27A

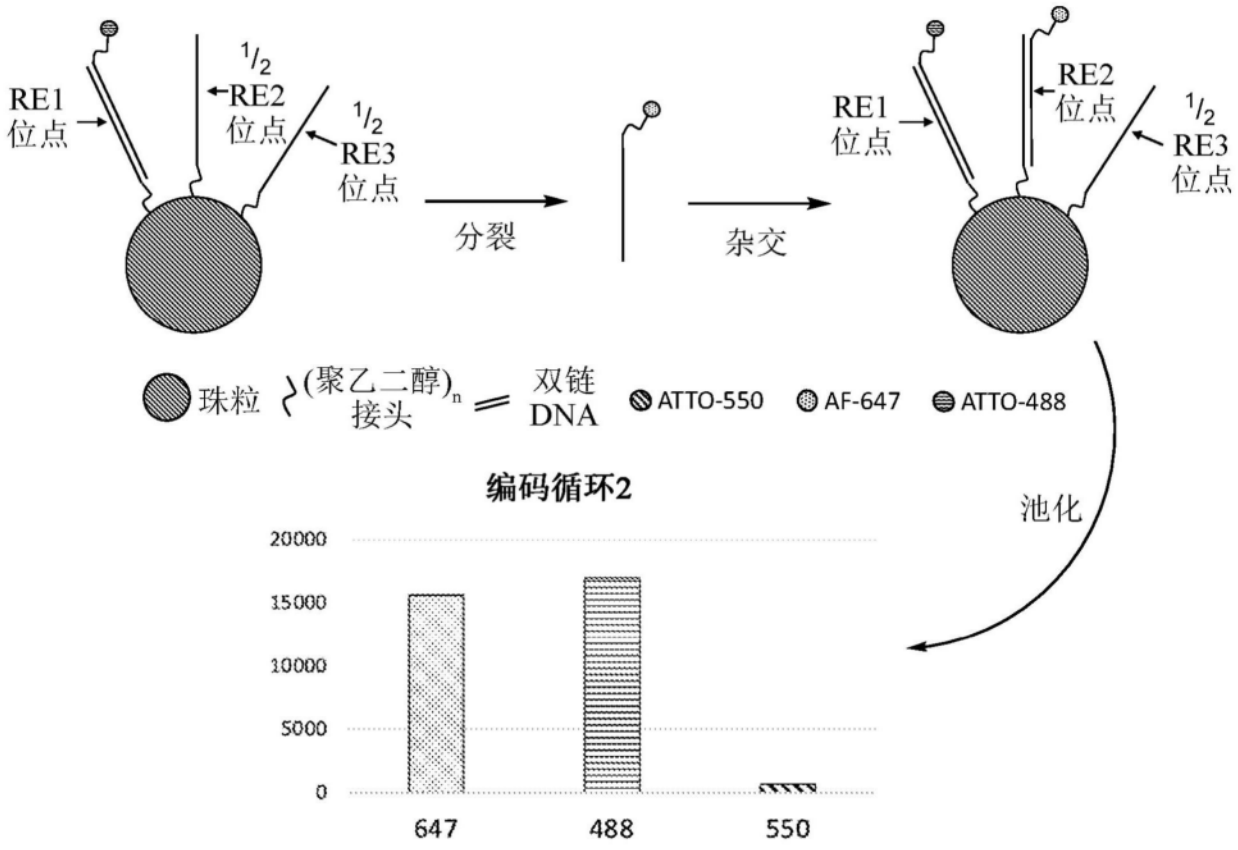

RE1
位点
1/2
RE2
位点
1/2
RE3
位点
分裂
杂交
RE1
位点
RE2
位点
1/2
RE3
位点
珠粒
(聚乙二醇)n
接头
双链
DNA
ATTO-550
AF-647
ATTO-488
池化
编码循环2
20000
15000
10000
5000
0
647
488
550

图27B

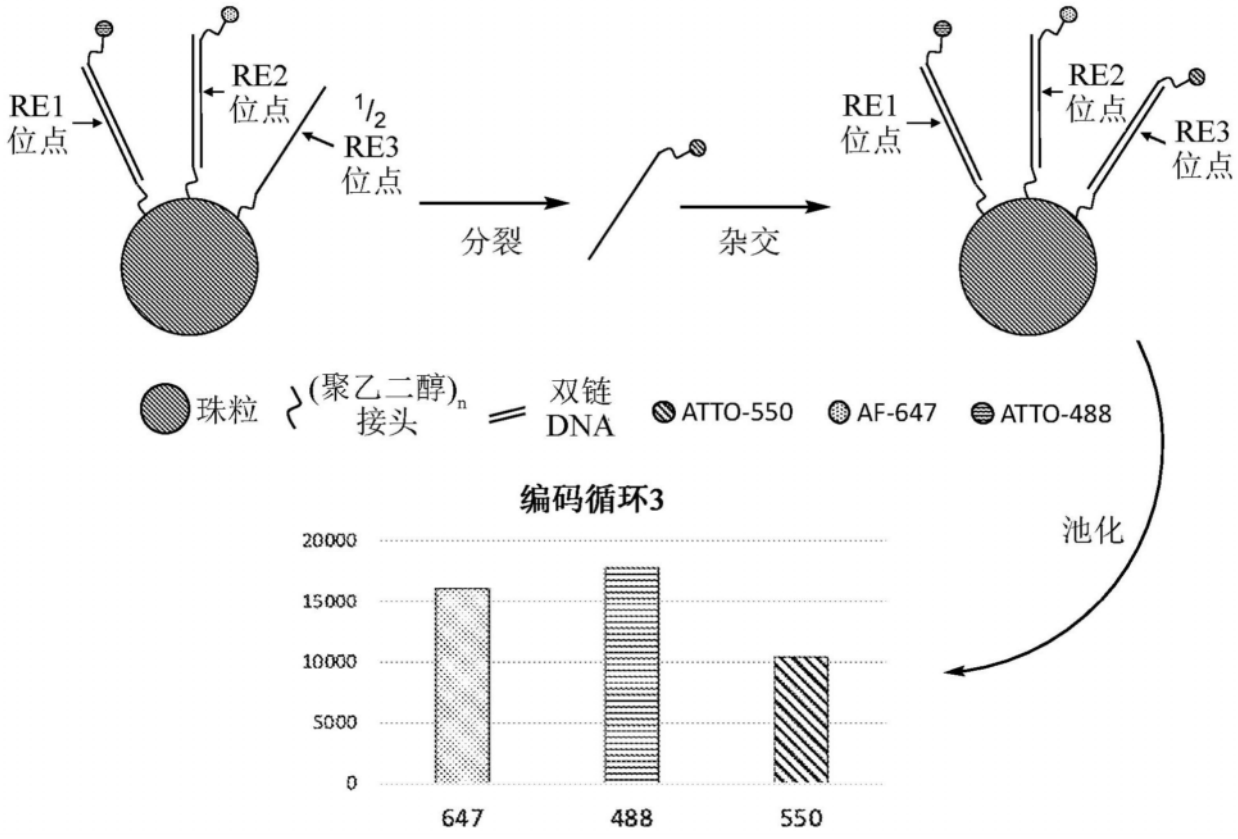

RE1
位点
RE2
位点
1/2
RE3
位点
分裂
杂交
RE1
位点
RE2
位点
RE3
位点
珠粒
(聚乙二醇)n
接头
双链
DNA
ATTO-550
AF-647
ATTO-488
池化
编码循环3
20000
15000
10000
5000
0
647
488
550

图27C

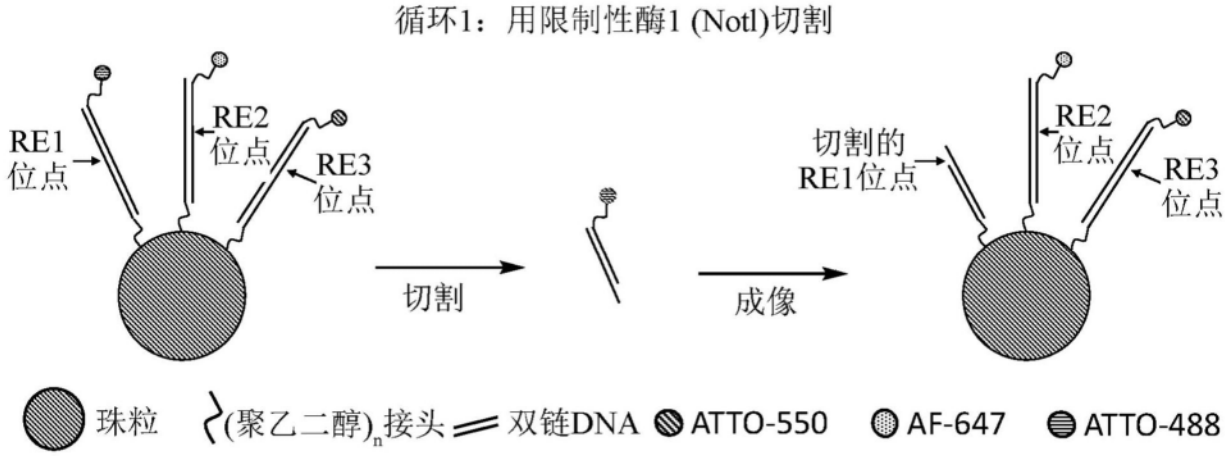

循环1：用限制性酶1 (NotI)切割
RE1
位点
RE2
位点
RE3
位点
切割
切割的
RE1位点
RE2
位点
RE3
位点
成像
珠粒
(聚乙二醇)n接头
双链DNA
ATTO-550
AF-647
ATTO-488

图28A

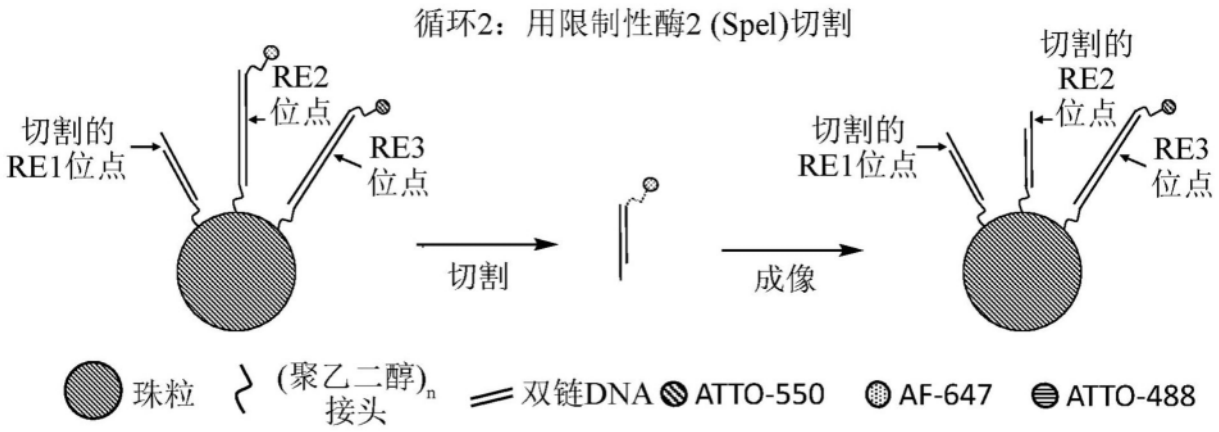

循环2：用限制性酶2 (SpeI)切割
切割的
RE1位点
RE2
位点
RE3
位点
切割
成像
切割的
RE1位点
切割的
RE2
位点
RE3
位点
珠粒
(聚乙二醇)n
接头
双链DNA
ATTO-550
AF-647
ATTO-488

图28B

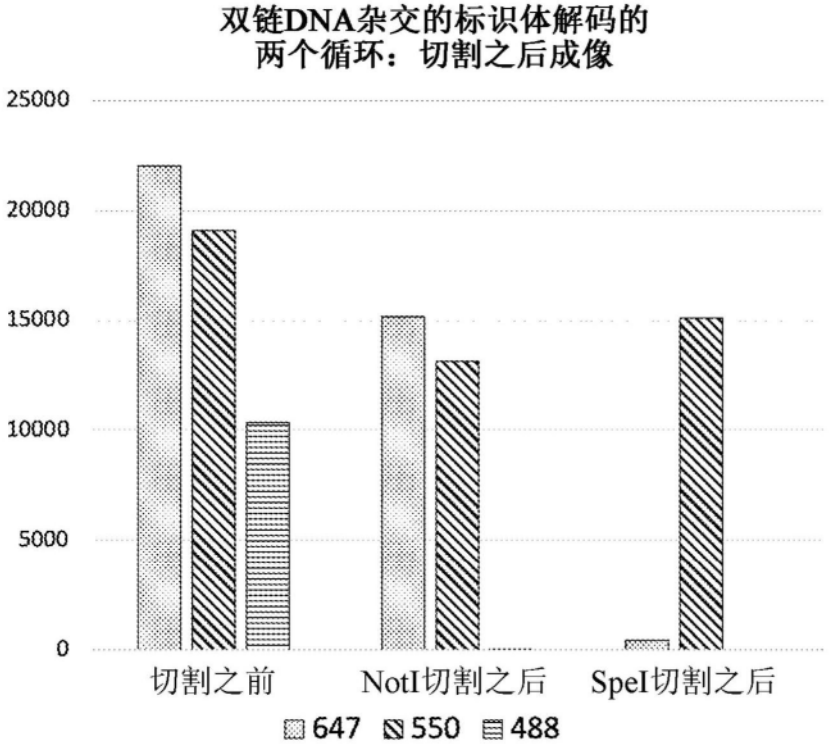

双链DNA杂交的标识体解码的
两个循环：切割之后成像
25000
20000
15000
10000
5000
0
切割之前
NotI切割之后
SpeI切割之后
647
550
488

图28C