



(12) 发明专利申请

(10) 申请公布号 CN 118103523 A

(43) 申请公布日 2024. 05. 28

(21) 申请号 202280036303.1

(22) 申请日 2022.05.19

(30) 优先权数据

2107192.3 2021.05.19 GB

(85) PCT国际申请进入国家阶段日

2023.11.20

(86) PCT国际申请的申请数据

PCT/GB2022/051268 2022.05.19

(87) PCT国际申请的公布数据

W02022/243691 EN 2022.11.24

(71) 申请人 牛津纳米孔科技公开有限公司

地址 英国牛津郡

申请人 帝国理工学院创新有限公司

(72) 发明人 安德鲁·约翰·赫伦

理查德·亚历山大·古铁雷斯

约书亚·埃德尔

亚历山大·伊万诺夫

卡罗琳·科赫 薛梁

本尼迪克特·雷利-奥唐纳

(74) 专利代理机构 北京市柳沈律师事务所

11105

专利代理师 陈铎

(51) Int.Cl.

C12Q 1/6825 (2006.01)

C12Q 1/6869 (2006.01)

权利要求书2页 说明书40页

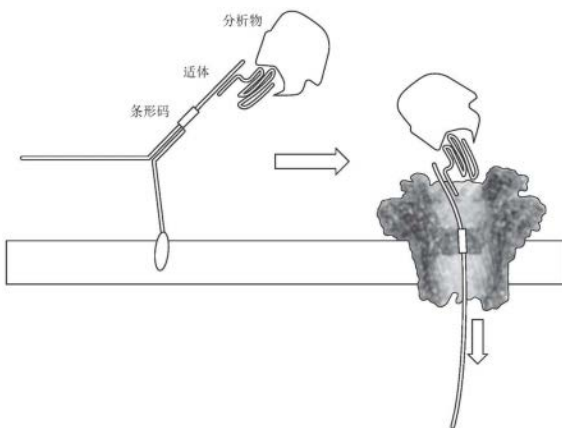
序列表55页 附图40页

(54) 发明名称

方法

(57) 摘要

本发明公开了一种用于检测样品中的多种分子的方法,所述方法包括:(a)使所述样品与载体和纳米孔接触,其中所述载体包括单链前导区、标识符区和对待检测分子具有特异性的分子结合区,并且其中马达蛋白与所述载体结合使得其可以控制所述标识符区在所述纳米孔内的移动;(b)当载体在所述纳米孔内移动时进行一种或多种光学或电学测量,以表征所述标识符区并确定所述分子是否与所述分子结合区结合。



1. 一种用于检测样品中的多种分子的方法,所述方法包括:
 - (a) 使所述样品与载体和纳米孔接触,其中所述载体包括单链前导区、标识符区和对待检测分子具有特异性的分子结合区,并且其中马达蛋白与所述载体结合使得其可以控制所述标识符区在所述纳米孔内的移动;
 - (b) 当载体在所述纳米孔内移动时进行一种或多种光学或电学测量,以表征所述标识符区并确定所述分子是否与所述分子结合区结合。
2. 根据权利要求1所述的方法,其中所述载体进一步包括在所结合的马达蛋白与所述分子结合区之间的间隔子。
3. 根据权利要求2所述的方法,其中所述载体依次包括单链前导区、标识符区、间隔子和分子结合区。
4. 根据权利要求1至3中任一项所述的方法,其中所述分子结合区和/或所述标识符区是多核苷酸。
5. 根据前述权利要求中任一项所述的方法,其中所述分子结合区或其部分是所述标识符区。
6. 根据权利要求4所述的方法,其中所述标识符区是多核苷酸并且包括条形码序列。
7. 根据前述权利要求中任一项所述的方法,其中所述载体包括多于一个标识符区和/或多于一个分子结合区。
8. 根据权利要求7所述的方法,其中所述载体包括多于一个标识符区,并且所述方法用于检测多个样品中的多种分子,所述载体中的一个标识符区对于所述样品是唯一的,并且所述载体中的一个标识符区对于与所述载体结合的所述分子是唯一的。
9. 根据前述权利要求中任一项所述的方法,其中所述分子包括神经递质、蛋白质和/或miRNA。
10. 根据前述权利要求中任一项所述的方法,其中所述分子结合区是适体。
11. 根据前述权利要求中任一项所述的方法,其中所述分子结合区是抗体、抗体片段、纳米抗体或亲和体。
12. 根据前述权利要求中任一项所述的方法,其中所述分子结合区与miRNA互补。
13. 根据前述权利要求中任一项所述的方法,其中所述标识符区是多核苷酸,并且所述方法包括:确定所述标识符区的多核苷酸序列。
14. 根据前述权利要求中任一项所述的方法,其中所述方法用于检测所述分子的存在或不存在。
15. 根据前述权利要求中任一项所述的方法,其中所述方法用于测定所述分子的浓度。
16. 根据前述权利要求中任一项所述的方法,其中所述多种分子是10种或更多种不同分子。
17. 根据前述权利要求中任一项所述的方法,其中所述马达蛋白是解旋酶、聚合酶、核酸酶、转位酶或拓扑异构酶。
18. 根据前述权利要求中任一项所述的方法,其中所述纳米孔是蛋白质孔、固态孔或DNA折纸孔。
19. 一种载体,其包括单链前导区、标识符区和对待检测分子具有特异性的分子结合区,其中马达蛋白在所述单链前导区与所述标识符区之间的位置处与所述载体结合。

20. 根据权利要求19所述的载体,其进一步包括在所述标识符区与所述分子结合区之间的间隔子和/或与所述分子结合区特异性结合的分子。

21. 一种用于多种分子的载体群体,其中所述载体如权利要求19或20所定义,并且所述群体中的不同载体包括不同标识符区和不同分子结合区。

22. 一种用于检测样品中的多种分子的试剂盒,所述试剂盒包括:

(i) 载体群体,其中每个载体包括标识符区和对待检测分子具有特异性的分子结合区,并且所述群体中的不同载体包括不同标识符区和不同分子结合区;

(ii) 包括单链前导区的衔接子;以及

(iii) 马达蛋白。

23. 一种用于检测样品中的多种分子的系统,所述系统包括:

(i) 根据权利要求21所述的载体群体;以及

(iii) 纳米孔。

方法

技术领域

[0001] 本发明涉及使用纳米孔技术检测样品中的分子的多重方法。本发明也涉及用于结合和识别分子的载体、此类载体的群体以及包括此类载体的试剂盒和系统。

背景技术

[0002] 生物传感器是医学诊断的重要部分并且是快速发展行业的一部分。目前用于生物标志物(例如,蛋白质)感测的最先进的检测技术通常与通过ELISA测定的定量光学读数或定性颜色读数相结合,并且受限于低浓度,这掩盖了稀少事件。基于抗体的检测技术通常在范围上受限于多肽靶标,并且在低浓度下灵敏度低。基于质谱(MS)的技术通常需要大量的样品制备,并且也可能在低浓度下灵敏度低,而如果要分析多个靶标,靶向MS技术则需要大样品量。

[0003] 纳米孔技术先前已经用于检测非多核苷酸分子。WO 2013/121201描述了使用包括适体的探针和跨膜孔技术来确定一种或多种分子的存在或不存在的方法,并举例说明了对凝血酶的检测。DNA载体先前已经用于实现对靶标的选择性、无标记检测,但仅限于单一分析物并且不能轻易扩展(参见,例如,Sze等人“Nature comms 8.1 (2017):1-10,以及Cai等人Nature comms 10.1 (2019):1-9)。

[0004] 此外,一些生物标志物群体,例如,miRNA群体具有短寿命并且以低浓度存在,这意味着它们目前在临床上无法获得。

[0005] 因此,需要能够实现同时检测复杂样品(诸如生物体液)中的多种可溶性蛋白质、miRNA和其他分子(诸如生物标志物)的分析方法。此外,也需要此类技术能够检测非常低浓度的靶分子。能够实现该目标的技术有望产生深远的影响,例如,在医疗保健领域用于诊断和监测疾病进展。此类技术也可以在不同领域得到应用,诸如针对污染物或其他杂质的存在研究水样。

发明内容

[0006] 本公开涉及利用纳米孔技术等来检测蛋白质、miRNA和其他生物标志物以及其他类型的分子的方法。该技术具有以高灵敏度和快速读数直接在未处理的样品中进行高度多重检测的潜力。

[0007] 因此,本文提供了:

[0008] -一种用于检测样品中的多种分子的方法,该方法包括:

[0009] (a) 使该样品与载体和纳米孔接触,其中该载体包括单链前导区、标识符区和对待检测分子具有特异性的分子结合区,并且其中马达蛋白与该载体结合使得其可以控

[0010] 制该标识符区在该纳米孔内的移动;

[0011] (b) 当载体在该纳米孔内移动时进行一种或多种光学或电学测量,以表征该标识

[0012] 符区并确定该分子是否与该分子结合区结合;

[0013] -载体,其包括单链前导区、标识符区和对待检测分子具有特异性的分子结合区,

其中马达蛋白在该单链前导区与该标识符区之间的位置处与该载体结合；

[0014] -用于多种分子的载体群体,其中每个载体包括单链前导区、标识符区和对待检测分子具有特异性的分子结合区,其中马达蛋白在该单链前导区与该标识符区之间的位置处与该载体结合,并且该群体中的不同载体包括不同标识符区和不同分子结合区；

[0015] -一种用于检测样品中的多种分子的试剂盒,该试剂盒包括：

[0016] (i) 载体群体,其中每个载体包括标识符区和对待检测分子具有特异性的分子结

[0017] 合区,并且该群体中的不同载体包括不同标识符区和不同分子结合区；

[0018] (ii) 包括单链前导区的衔接子；以及

[0019] (ii) 马达蛋白；以及

[0020] -一种用于检测样品中的多种分子的系统,该系统包括：

[0021] (i) 载体群体,其中每个载体包括单链前导区、标识符区和对待检测分子具有特异性的分子结合区,其中马达蛋白在该单链前导区与该标识符区之间的位置处与该载体结合,并且该群体中的不同载体包括不同标识符区和不同分子结合区；以及

[0022] (iii) 纳米孔。

附图说明

[0023] 图1. 条形码测序和对生物标志物的检测的示意图。总体而言,主要由条形码和具有结合靶向分析物的结合区(例如,cDNA、适体、抗体)组成的定制设计链被束缚在膜上。通过易位通过生物纳米孔(CsgG)来检测靶结合链。

[0024] 图2. (A) 示例性完整载体的示意图,该载体包括:(i) 用于促进穿入纳米孔的前导区；(ii) 具有用于提高捕获率的胆固醇接头的系链；(iii) 提供波形信号用于在施加电压下与该纳米孔偶联时对条形码进行测序的马达蛋白；(iv) 多核苷酸标识符区段(例如,条形码或可能重复的多个条形码)；(v) 连接该条形码的一个或多个间隔子以及(vi) 分子结合区,诸如适体/c-miRNA/抗体,其选择性地靶向例如miRNA、蛋白质或神经递质。(B) 示例性完整载体的示意图,该载体包括衔接子,该衔接子由以下组成:(i) 用于促进穿入纳米孔的前导区；(ii) 具有用于提高捕获率的胆固醇接头的系链；(iii) 提供波形信号用于在施加电压下与该纳米孔偶联时对条形码进行测序的马达蛋白。将衔接子连接至由以下组成的DNA链:(i) 与互补链杂交以具有单个A突出端的衔接子连接部分；(ii) 多核苷酸标识符区段(条形码或可能重复的多个条形码)；(iii) 将条形码连接至(iv) 一个或多个分子结合区的一个或多个间隔子,诸如选择性结合例如靶miRNA、蛋白质或神经递质的适体/c-miRNA/抗体。

[0025] 图3. 载体和用于确定分子在分子结合区上的存在或不存在的方法的图示。酶消化用于去除/消化未与对其具有特异性的分子结合的任何分子结合区。(a) 针对切口酶或核酸内切酶的位点包括在载体中,使得当分子结合区与对其具有特异性的分子结合时,该位点对该切口酶或该核酸内切酶隐藏,但当该分子与该分子结合区结合时,该位点被暴露。在适合于对对其具有特异性的分子的分子结合区的条件下使样品与载体接触之后,可以添加切口酶或核酸内切酶,使得切口被引入其中该分子结合区未与对其具有特异性的分子结合的任何载体中。在这样的消化之后,基于标识符区(条形码序列)产生的信号之后当前信号的存在或不存在,可以将与靶分子结合的载体和与靶分子结合的载体区分开。(b) 在适合于对对其具有特异性的分子的分子结合区的条件下使样品与载体接触之后,可以添加核酸外切

酶,使得该分子结合区被消化。在这样的消化之后,基于标识符区(条形码序列)产生的信号之后当前信号的存在或不存在,可以将与靶分子结合的载体和与靶分子结合的载体区分开。

[0026] 图4.条形码测序和去多重。(a)所有序列均使用碱基识别算法(参见载玻片19-25)进行碱基识别,并与参考序列(即条形码序列)进行比对。最大比对分数用于对条形码进行分类。如果最大比对分数和第二高分太接近,则该事件不会被分类。此外,p值用于对事件进行分类并去除误报分类。(b)使用该方法,所有记录事件的86%被使用和分类,准确性达到99.95%。

[0027] 图5.测序、比对和条形码分类。(a)条形码测序的准确性>90%,误报率的几率为0.0001%。(b)混淆矩阵示出对错误条形码分类的偏好非常低。

[0028] 条形码1: TGCTACTCTCCTCATAAGCAGTCCGGTGTATCGAT,

[0029] 条形码2: ATCGCTACGCCTTCGGCTCGTAATCATAGTCGAGT,

[0030] 条形码3: AGCTCAGAGCAGGTCACTCAAGATACGAGCTGCGT,

[0031] 条形码4: GTAAGTCTGCATCAGCGCGCGGCTGTGCGAGGATA,

[0032] 条形码5: CTACGACAGTACGCTAGCAAGGATAGACACTACGA,

[0033] 条形码6: TACTGAACACAAGTTCGTCGTCGAGCAATCACAAT,

[0034] 条形码7: AGTCTACCATTACTTGGATCGGATTAGCCTCACTC,

[0035] 条形码8: TGCACGAGTGCGTGTCAACCGTCCAGATGCTCGTG,

[0036] 条形码9: CTAGTGCGCAGTTGTCTCGGCGGAGTTGAGACTGA,

[0037] 条形码10: GATCATGGTAGTCTTCAAGATCGAGTATGTCTGTC。

[0038] 图6.在复杂的混合物中区分了10种条形码载体。在相同浓度下,在条形码载体检测率中没有观察到偏差。使用的条形码是上面的条形码1-10。

[0039] 图7.停滞分析用于确定靶标是否已经与条形码链结合。(a)如果靶分析物未与条形码链结合,则当前信号不表明停滞(在停留时间、电流振幅中)。(b)结合的分析物(此处给出了互补miRNA的示例)使载体停滞,其产生独特的电流分布。

[0040] 图8.没有miRNA的载体(条形码6)的停滞(6.4%)显著低于添加miRNA时的停滞(62.18%)。

[0041] 图9.对具有浓度依赖性的miRNA多重检测。10种不同条形码使得能够检测和区分10种不同miRNA。捕获率中存在一些观察到的异质性,但在0-5nM miRNA下动态范围保持相似。使用的条形码是上面的条形码1-10。

[0042] miRNA 1: CAGCAGCACACUGUGGUUUGU,

[0043] miRNA 2: AGAGCUUAGCUGAUUGGUGAAC,

[0044] miRNA 3: UAGCUUAUCAGACUGAUGUUG,

[0045] miRNA 4: ACCUGGCAUACAAUGUAGAUUU,

[0046] miRNA 5: UGUAAACAUCCCCGACUGGAAG,

[0047] miRNA 6: UGUAAACAUCCUACACUCUCAGC,

[0048] miRNA 7: AGCUGGUAAAAUGGAACCAAU,

[0049] miRNA 8: GAGCUUUUGGCCCGGGUUAUAC,

[0050] miRNA 9: AACAUUCAUUGCUGUCGGUGGGU,

[0051] miRNA 10:UAGCACCAUCUGAAAUCGGUUA。

[0052] 图10.对凝血酶和15-mer凝血酶结合适体之间的结合的检测。与凝血酶结合之后,波形事件示出更长的停留时间,其中与400nM凝血酶结合之后发现在电流翻转方面停滞显著增加,对应于G-四链体和适体-蛋白质相互作用的解旋。本实验中使用实例1中提供的第一凝血酶载体。

[0053] 图11.凝血酶和15-mer凝血酶结合适体之间的结合的浓度依赖性。通过将凝血酶浓度从0nM增加到400nM来验证结合。随着凝血酶浓度的增加,观察到载体的停滞增加,随着停留时间的延长,出现更多的波形事件。本实验中使用实例1中提供的第一凝血酶载体。

[0054] 图12.使用茎环适体对血清素进行检测。使用实例1中提供的与血清素适体相关联的条形码序列。示出了载体中的血清素和适体的结构。提供了示例性电流迹线,示出了条形码与孔相互作用时产生的信号以及适体与孔相互作用时产生的信号。如表格和图表所示,适体展开引起的平均延迟随着血清素浓度的增加而增加。

[0055] 图13.使用茎环适体对血清素进行检测。(A)不存在血清素和存在40mM血清素的情况下的电流迹线。如电流与停留时间图所示,在存在血清素的情况下停留时间增加。(B)示出了针对0mM、2.5mM、5mM、10mM、20mM和40mM血清素的电流与停留时间图。观察到停滞事件的浓度依赖性增加。

[0056] 图14.使用茎环适体对乙酰胆碱进行检测。示出了使用的条形码和适体序列。提供了有和没有乙酰胆碱的载体的示例性测量。观察到乙酰胆碱浓度与延迟百分比之间的相关性。

[0057] 图15.对没有马达蛋白的分子的检测。(A)使用凝血酶结合适体(TBA)检测凝血酶浓度的增加。阴影区指示TBA的信号。信号处于较低nA的事件指示TBA结合凝血酶。(B)使用血清素结合适体(SBA)检测血清素浓度的增加。阴影区与针对SBA结合血清素的单独SBA观察到的信号相关。(C)对多重样品中的条形码的检测。条形码是基于条形码区的信号幅度来区分的。

[0058] 图16.多重筛选和检测策略实例。载体群体可以用于生成用于筛查和诊断的生物护照。

[0059] 图17.载体链序列-miRNA。

[0060] 图18.载体链序列-蛋白质和神经递质。

[0061] 图19.条形码分类的混淆矩阵。

[0062] 图20.对多个miRNA的检测:A.与对照(0nM)相比,具有10nM miRNA的条形码样品(条形码13)的易位时间的增加。B.(顶部)条形码事件的特征电流迹线以及相关移动标准偏差图。如果移动标准偏差低于某个阈值(0.003),则该事件被分类为延迟。(底部)延迟事件的特征电流迹线以及其相关移动标准偏差图。C.单一条形码(条形码38)滴定曲线($n=5$)。D.同一样品中40个不同条形码随着(各自)miRNA浓度增加的多重滴定曲线($n=5$)。E.箱线图示出了针对在多重实验($n=5$)中添加的10nM miRNA检测到的延迟,叠加了针对添加了10nM miRNA的每个条形码的单个实验的散点图($n=1$,点)。

[0063] 图21.多重实验中未知miRNA浓度的定量。(A)单独绘制针对40个条形码多重实验的滴定曲线。每条曲线均采用Hill拟合函数进行拟合。基于标准曲线确定添加的miRNA的真实浓度(深灰色)和添加的miRNA的预测浓度(浅灰色)。结果示出了预测浓度与实际浓度之

间有非常高的重叠。(B) 残差分析示出了预测值减去实际值 ($n=12$) 之间的差异。

[0064] 图22. 对cTnI的检测。A: cTnI适体序列。B: 事件时间 $\pm 30\text{ng/ml}$ cTnI B总事件时间的比较。C: 到达C3峰的事件时间。D: 从C3峰到结束的事件时间。cTnI的浓度-延迟关系, 在 $t > 95\%$ 的对照事件的情况下, 事件被‘延迟’。E: 总事件时间。F: 到达C3峰的事件时间。G: 从C3峰到结束的事件时间。

具体实施方式

[0065] 分子载体和多重方法

[0066] 本公开涉及检测样品中的多种分子的方法。该方法包括: 使样品与载体接触, 其中该载体包括与对待检测分子具有特异性的分子结合区相关联的标识符区。当载体在马达蛋白的控制下在检测器 (诸如孔) 内移动时, 表征标识符区并确定分子是否与分子结合区结合。因此, 该方法根据相关联的标识符区的表征以及对分子是否与分子结合区结合的确下来识别待检测的分子是存在还是不存在。可以检测样品中的多种分子。如下文更详细地解释, 该方法允许正确地检测和识别样品中的多种分子。

[0067] 如上所述, 用于识别样品中的分子的方法是本领域已知的。本领域已知的一种方法涉及包括适体和尾部的探针 (参见, WO 2013/121201)。提供了具有不同适体和不同尾部的不同探针, 并且根据分析物是否与适体结合, 每个尾部可以对流过孔的电流产生不同影响。以这种方式, 该方法可以检测样品中的多种分析物。与没有分析物与适体结合的探针相比, 当分析物与适体结合时, 通过使尾部通过孔的移动停滞来检测与适体结合的分析物的存在或不存在。

[0068] 然而, 现有技术的方法受限于可以生成的探针的多样性和不同尾部的数量。仅通过改变探针的长度、探针尾部中存在或不存在双链区以及每个载体中不同适体的不同结合亲和力来将不同分析物彼此区分。这些限制限制了样品中可以相互区分的不同分析物的数量。

[0069] 发明人已经设计了在样品内大量不同分析物和/或载体之间进行区分的方法。发明人已经设计了包含位于载体上的马达蛋白的载体, 使得其可以控制孔内分子特异性标识符区的移动。受控的移动允许例如通过测序来准确表征标识符区。当以这种方式控制载体通过孔的移动时, 包含马达蛋白的载体中的被设计用于检测不同分子的标识符区之间的差异可以小得多。因此, 本发明人设计的载体允许对大量不同标识符区进行无标记识别。因此, 可以使用载体来执行用于识别单个样品中的多种分析物的高度多重方法。通过包括针对与不同样品接触的载体的替代标识符区, 本发明的方法也允许同时测量来自多个样品的多种分析物。

[0070] 本文公开的改进方法在检测样品中的非常低水平的分析物方面表现良好 (即具有高灵敏度)。本文公开的方法使得能够使用超稀释样品 (亚皮摩尔水平) 高效检测和筛选例如稀有蛋白质和miRNA分子, 并且可以以高通量方式实现单分子灵敏度。例如, 本文公开的方法可以用于检测以低至约 1pM 至约 1fM 的浓度存在于样品中的分子, 诸如蛋白质或多核苷酸, 诸如miRNA。标准的基于适体的分子检测方法需要更高浓度的分析物以便产生可检测的信号。本文公开的方法使得能够在单分子水平上进行检测, 并且可以用于确定样品中分子的浓度。分子与载体的分子结合区的特异性结合允许增加分子的有效浓度, 例如通过经由

载体上的膜锚定物定位到包括孔的膜。本文公开的载体也提供了允许对未处理的样品(诸如例如,天然临床样品)中的分子进行多重检测的益处。

[0071] 因此,本文提供了一种用于检测样品中的多种分子的方法,该方法包括:

[0072] (a) 使该样品与载体和纳米孔接触,其中该载体包括单链前导区、标识符区和对待检测分子具有特异性的分子结合区,并且其中马达蛋白与该载体结合使得其可以控制该标识符区在该纳米孔内的移动;

[0073] (b) 当载体在纳米孔内移动时进行一种或多种光学或电学测量,以表征该标识符区并确定该分子是否与该分子结合区结合。

[0074] 表征标识符区

[0075] 载体包括一个或多个标识符区。该方法要求载体与检测器相互作用,例如在孔内移动,以表征标识符区。下面讨论可以用来表征标识符区的合适的测量。

[0076] 优选地,标识符区包括或者是多核苷酸。优选地,确定标识符区的多核苷酸序列。可以使用任何合适的技术。本公开特别适合于单分子表征和对低浓度分子的检测。本文更详细地讨论了示例性合适的测序技术。例如,在一些优选的实施例中,测序技术是纳米孔感测方法。纳米孔感测方法在此详细描述。然而,本文公开的方法不限于纳米孔感测。其他单分子测序技术适用于本文公开的方法。

[0077] 在纳米孔链测序中,标识符区在孔内移动。当标识符区在孔内移动时记录的信号允许确定标识符区的序列。标识符区的其他特征(例如,(i)多核苷酸的长度、(ii)多核苷酸的身份、(iii)多核苷酸的二级结构以及(iv)多核苷酸是否被修饰)可以另选地或附加地被确定以表征该标识符区。

[0078] 纳米孔通道中载体分子的存在对通过孔的开放通道离子流有影响。这就是孔通道“分子感测”的本质。可以使用合适的测量技术来测量开放通道离子流的变化,例如,通过电流的变化(例如,WO 2000/28312以及D.Stoddart等人,Proc.Natl.Acad.Sci.,2010,106,7702-7或WO 2009/077734)。如通过电流的减少来测量的离子流减少的程度与孔内或孔附近的障碍物的大小有关。类似的信息可以使用光学方法获得,例如,Huang等人,Nature Nanotechnology 10,986-991(2015)中所公开。因此,感兴趣分子(例如,靶多核苷酸)在孔中或孔附近的结合提供了可检测和可测量的事件,从而形成“生物传感器”的基础。

[0079] 当核酸分子或单个碱基在孔内移动时(例如,当它通过纳米孔的通道时),碱基之间的尺寸差异导致通过通道的离子流的直接相关减少。可以记录离子流的变化。用于记录离子流变化的合适的电测量技术在例如WO 2000/28312以及D.Stoddart等人,Proc.Natl.Acad.Sci.,2010,106,第7702-7页(单通道记录设备);以及例如WO 2009/077734(多通道记录技术)中描述。通过适当的校准,离子流的特征减少可以用于实时识别穿过通道的特定核苷酸和相关联的碱基。在典型的纳米孔核酸测序中,当感兴趣核酸序列的各个核苷酸依次通过纳米孔的通道时,由于通道被核苷酸部分堵塞,开放通道离子流减少。使用上述合适的记录技术来测量离子流的这种减少。离子流的减少可以被校准为通过通道的已知核苷酸的测量离子流的减少,从而产生用于确定哪个核苷酸正在通过通道的手段,并且因此当顺序进行时,产生确定穿过纳米孔的核酸的核苷酸序列的方法。为了准确确定单个核苷酸,通常需要使通过通道的离子流的减少与通过缩颈(或“读数头”)的单个核苷酸的大小直接相关。应当理解,例如,可以对经由相关马达蛋白(诸如聚合酶或解旋酶)的作

用“穿过”孔的完整核酸聚合物执行测序。本文更详细地描述了合适的马达蛋白。

[0080] 确定分子结合

[0081] 该方法确定分子是否与载体的分子结合区结合。在一些实施例中,该方法用于检测与载体特异性结合的分子的存在或不存在。与载体特异性结合的分子的存在或不存在指示一个或多个样品中分子的存在或不存在。

[0082] 该方法也可以用于测定分子的浓度,诸如分子结合载体与游离载体的相对浓度或样品中分子的绝对浓度。相对浓度和绝对浓度可以以任何合适的方式从所获得的测量结果计算。实例和附图中提供了确定相对浓度和绝对浓度的方法的实例。

[0083] 载体的分子结合区与检测器的相互作用用于确定分子是否与至载体的分子结合区结合。

[0084] 当使用纳米孔检测时,分子结合区影响流过孔的电流,取决于该分子结合区是否与分子特异性结合。当分子未结合时,分子结合区以一种方式影响流过孔的电流,并且当分子结合时,该分子结合区以不同方式影响流过孔的电流。这很重要,因为它允许使用该方法确定与分子结合区特异性结合的分子的存在或不存在,并因此确定样品中分子的存在或不存在。

[0085] 载体中的分子特异性标识符区(即,仅存在于也包括对给定的感兴趣分子具有特异性的分子结合区的载体中的标识符区)产生的信号用于识别与该载体结合或未结合的分子并且因此识别在样品中存在或不存在的分子。

[0086] 分子结合区对流过孔的电流的影响取决于分子是否与分子结合区结合,可以基于载体或该载体的分子结合区在孔内移动所花费的时间来测量。例如,当分子结合区是适体时,该适体的二级和三级结构可以可检测地减慢或暂时阻止载体在孔内的前进。当适体与其同源分子结合时,载体可能必须克服更高的能量障碍以在孔内前进,并且该适体穿过孔的前进速度可能比该适体未与分子结合时可检测地慢(诸如延长的相互作用)。

[0087] 在一些实施例中,分子结合区可以是与靶分子(诸如miRNA)互补的多核苷酸。在此类实施例中,当分子结合区未与分子结合时,载体可以在孔内以“正常”速度前进。当分子结合区与其靶分子(诸如miRNA)结合时,分子结合区现在的双链部分可能影响载体在孔内的前进,使得前进可检测地减慢或停滞。以这种方式,可以确定与载体的分子结合区特异性结合的分子的存在或不存在。

[0088] 在一些实施例中,分子结合区可以是抗体、抗体片段、纳米抗体或亲和体。虽然这种分子结合区的存在可能阻止整个载体通过孔的移动,但是可以确定分子的存在或不存在。不希望受理论的束缚,载体的这种分子结合区充当孔的“泄漏塞”。纳米孔测量对系统中的微小变化极其敏感,从而能够在各个核苷酸碱基之间进行区分,并且因此可以根据分子是否与分子结合区结合来检测塞的“泄漏”的差异。

[0089] 可以执行对照实验以确定载体与分子特异性结合时与分子未结合时相比,分子结合区对流过孔的电流和/或该载体的前进的影响。然后可以将对测试样品执行本文所述的方法的结果与源自此类对照实验的结果进行比较,以确定特定分子是存在还是不存在于测试样品中。这在W0 2013/121201中更详细地描述。

[0090] 可以使用替代方法来确定分子是否与分子结合区结合。例如,可以使用基于消化的方法,其依赖于分子结合时与分子未结合时相比的分子结合区的消化的差异。通常,与载

体的分子结合区结合的分子可以保护分子结合区免遭消化。

[0091] 例如,当分子结合区是多核苷酸时,可以使用位点特异性核酸内切酶,诸如限制酶或单链切口酶。当分子未结合时,多核苷酸分子结合区被消化,并且当载体在孔内移动时检测到其不存在。当分子结合时,多核苷酸分子结合区不被消化或简单地被“切开”,导致载体在孔内的前进减慢或停滞,如上所述,并且任选地对分子结合区的表征。

[0092] 消化方法也可以应用于其中分子结合区是多核苷酸的情况,诸如抗体、抗体片段、纳米抗体或亲和体。分子与分子结合区的结合可以导致对蛋白酶靶位点的保护,并且因此分子结合区不被消化。在不存在分子的情况下,分子结合区可以被消化。信号的差异可以用于确定载体上分子的存在或不存在。

[0093] 在一些实施例中,本文所述的间隔子可以位于载体的分子结合区的前导序列侧上。间隔子导致孔内载体的前进减慢或停滞,从而提供测量信号中分子结合区的位置的指标。此外,当分子结合时,间隔子的存在也允许放大载体在孔内的前进减慢或停滞,从而提供更清晰的信号来确定分子在载体上的存在或不存在。

[0094] 该方法也允许测定分子的浓度,如实例中更详细地描述。在一些实施例中,确定样品中分子结合载体与游离载体相比的相对浓度。这对于确定样品之间的分子水平的相对变化可能有用。在一些实施例中,可以测定样品内分子的绝对浓度。这可以使用标准曲线作为参考来执行,如实例中所举例说明。

[0095] 该方法优选是允许同时检测多种分子的多重方法。例如,该方法可以用于检测2种或更多种,诸如5种或更多种、10种或更多种、50种或更多种、100种或更多种(例如,200种至500种)、500种或更多种(例如,600种至1000种)或至少1000种(例如,1000种至10000种)不同分子。

[0096] 马达蛋白

[0097] 如本领域技术人员将理解的,任何合适的马达蛋白可以用于本文提供的方法和产品中。马达蛋白可以是能够与多核苷酸结合并控制其相对于检测器(诸如纳米孔)例如通过孔的移動的任何蛋白。在一些实施例中,多于一种马达蛋白与载体结合。

[0098] 在一些实施例中,马达蛋白是或源自多核苷酸处理酶。多核苷酸处理酶是能够与多核苷酸的至少一种特性相互作用并对其进行修饰的多肽。酶可以通过将多核苷酸定向或移动至特定位置来修饰该多核苷酸。

[0099] 在一些实施例中,马达蛋白源自酶分类(EC)组3.1.11、3.1.13、3.1.14、3.1.15、3.1.16、3.1.21、3.1.22、3.1.25、3.1.26、3.1.27、3.1.30和3.1.31中的任一者的成员。

[0100] 通常,马达蛋白是解旋酶、聚合酶、核酸外切酶、拓扑异构酶或它们的变体。

[0101] 当载体处于溶液中时,马达蛋白通常停滞在载体上。在一些实施例中,载体上的马达蛋白被修饰以防止该马达蛋白从该载体脱离(除了通过使间隔子的端脱离)。可以以任何合适的方式调整马达蛋白。例如,可以将马达蛋白加载到载体上并且然后进行修饰以便防止其从间隔子脱离。或者,可以修饰马达蛋白以防止其在加载到载体上之前从该载体脱离。对马达蛋白进行修饰以便防止其从载体脱离可以使用本领域已知的方法来实现,诸如WO 2014/013260中讨论的那些,其通过引用的方式整体并入本文,并且特别参考描述对马达蛋白(诸如解旋酶)进行修饰以便防止其与多核苷酸链脱离的段落。

[0102] 例如,马达蛋白可以具有多核苷酸非结合开口;例如,当马达蛋白从链脱离时多核

核苷酸链可以通过的空腔、裂缝或空隙。在一些实施例中,多核苷酸非结合开口是当马达蛋白从间隔子脱离时间间隔子可以穿过的开口。在一些实施例中,针对给定的马达蛋白的多核苷酸非结合开口可以通过参考其结构来确定,例如,通过参考其X射线晶体结构。X射线晶体结构可以在存在和/或不存在多核苷酸底物的情况下获得。在一些实施例中,多核苷酸非结合开口在给定的马达蛋白中的位置可以使用本领域已知的标准包通过分子建模来推断或确认。在一些实施例中,多核苷酸非结合开口可以通过一个或多个部分(例如,马达蛋白的一个或多个结构域)的移动来瞬时产生。

[0103] 可以通过关闭多核苷酸非结合开口来修饰马达蛋白。因此,关闭多核苷酸非结合开口可以防止马达蛋白从间隔子脱离。例如,可以通过共价关闭多核苷酸非结合开口来修饰马达蛋白。在一些实施例中,用于以这种方式寻址的优选马达蛋白是解旋酶。

[0104] 在一个实施例中,马达蛋白是核酸外切酶。合适的酶包括但不限于来自大肠杆菌的核酸外切酶I (SEQ ID NO:1)、来自大肠杆菌的核酸外切酶III (SEQ ID NO:2)、来自嗜热链球菌的RecJ (SEQ ID NO:3)和噬菌体 λ 外切核酸酶 (SEQ ID NO:4)、TatD外切核酸酶以及它们的变体。包含SEQ ID NO:3所示序列或其变体的三个亚基相互作用形成三聚体核酸外切酶。

[0105] 在一个实施例中,马达蛋白是聚合酶。聚合酶可以是PyroPhage® 3173DNA聚合酶(其可从Lucigen® Corporation商购)、SD聚合酶(可从Bioron®商购)、来自NEB的Klenow或其变体。在一个实施例中,酶是Phi29 DNA聚合酶 (SEQ ID NO:5)或其变体。可以用于本公开的Phi29聚合酶的修饰版本公开于美国专利号5,576,204中。

[0106] 在一个实施例中,马达蛋白是拓扑异构酶。在一个实施例中,拓扑异构酶是部分分类(EC)组5.99.1.2和5.99.1.3中的任一者的成员。拓扑异构酶可以是逆转录酶,其是能够催化从RNA模板形成cDNA的酶。它们可从例如New England Biolabs®和Invitrogen®商购。

[0107] 在一个实施例中,马达蛋白是解旋酶。根据本文提供的方法可以使用任何合适的解旋酶。例如,根据本公开内容使用的该马达蛋白或每种马达蛋白可以独立地选自Hel308解旋酶、RecD解旋酶、TraI解旋酶、TrwC解旋酶、XPD解旋酶和Dda解旋酶或它们的变体。单体解旋酶可以包括附接在一起的几个结构域。例如,TraI解旋酶和TraI亚组解旋酶可以含有两个RecD解旋酶结构域、一个松弛酶结构域和一个C末端结构域。这些结构域通常形成单体解旋酶,该解旋酶能够在不形成寡聚体的情况下发挥作用。合适的解旋酶的具体实例包括Hel308、NS3、Dda、UvrD、Rep、PcrA、Pif1和TraI。这些解旋酶通常作用于单链DNA。可以沿着双链DNA的两条链移动的解旋酶的实例包括FtfK和六聚酶复合物或多亚基复合物,诸如RecBCD。Hel308解旋酶在公开(诸如WO 2013/057495)中描述,其全部内容通过引用的方式并入。RecD解旋酶在公开(诸如WO 2013/098562)中描述,其全部内容通过引用的方式并入。XPD解旋酶在公开(诸如WO 2013/098561)中描述,其全部内容通过引用的方式并入。Dda解旋酶在公开(诸如WO 2015/055981和WO 2016/055777)中描述,其全部内容通过引用的方式并入。

[0108] 在一个实施例中,解旋酶包括SEQ ID NO:6所示的序列(Trwc Cba)或其变体、SEQ ID NO:7所示的序列(Hel308 Mbu)或其变体或SEQ ID NO:8所示的序列(Dda)或其变体。变体可以以本文讨论的任何方式不同于天然序列。SEQ ID NO:8的示例性变体包括E94C/A360C。SEQ ID NO:8的另一示例性变体包括E94C/A360C,并且然后包括(Δ M1)G1G2(即,删

除M1,并且然后添加G1和G2)。

[0109] 在一些实施例中,马达蛋白(例如解旋酶)可以在至少两种活动操作模式(当马达蛋白具备促进移动的所有必要组分,例如燃料和辅因子(诸如本文讨论的ATP和Mg²⁺时))和一种非活动操作模式(当马达蛋白不具备促进移动的必要时成分时)下控制多核苷酸的移动。

[0110] 当具备促进移动的所有必要组分时(即在活动模式下),马达蛋白(例如解旋酶)沿着多核苷酸以5'至3'或3'至5'方向移动(取决于马达蛋白)。在其中马达蛋白用于控制多核苷酸链相对于纳米孔的移动的实施例中,马达蛋白可以用于将多核苷酸移离(例如移出)孔(例如,逆着施加的场)或移动多核苷酸朝向(例如,进入)孔(例如,利用施加的场)。例如,当多核苷酸的马达蛋白朝向其移动的端被孔捕获时,该马达蛋白逆着由施加的电势产生的场的方向工作,并且将穿入的多核苷酸拉出孔(例如,进入顺式室)。然而,当马达蛋白远离其移动的端被捕获在孔中时,该马达蛋白以施加的电势产生的场的方向工作,并将穿入的多核苷酸推入孔中(例如,进入反式室)。

[0111] 当马达蛋白(例如,解旋酶)不具备促进移动所需的组分时(即,在非活动模式下),它可以与多核苷酸结合,并且在其相对于纳米孔移动时充当减慢多核苷酸的移动的制动器,例如,通过由施加的电势产生的场拉入孔中。在非活动模式下,多核苷酸的哪一端被捕获并不重要,是施加的场决定了多核苷酸相对于孔的移动,并且马达蛋白充当制动器。当在非活动模式下时,马达蛋白对多核苷酸的移动控制可以以多种方式描述,包括棘轮、滑动和制动。

[0112] 在活动模式下,马达蛋白通常消耗燃料分子。燃料通常是游离核苷酸或游离核苷酸类似物。游离核苷酸可以是但不限于以下中的一种或多种:单磷酸腺苷(AMP)、二磷酸腺苷(ADP)、三磷酸腺苷(ATP)、单磷酸鸟苷(GMP)、二磷酸鸟苷(GDP)、三磷酸鸟苷(GTP)、单磷酸胸苷(TMP)、二磷酸胸苷(TDP)、三磷酸胸苷(TTP)、单磷酸尿苷(UMP)、二磷酸尿苷(UDP)、三磷酸尿苷(UTP)、单磷酸胞苷(CMP)、二磷酸胞苷(CDP)、三磷酸胞苷(CTP)、单磷酸环腺苷(cAMP)、单磷酸环鸟苷(cGMP)、单磷酸脱氧腺苷(dAMP)、二磷酸脱氧腺苷(dADP)、三磷酸脱氧腺苷(dATP)、单磷酸脱氧鸟苷(dGMP)、二磷酸脱氧鸟苷(dGDP)、三磷酸脱氧鸟苷(dGTP)、单磷酸脱氧胸苷(dTMP)、二磷酸脱氧胸苷(dTDP)、三磷酸脱氧胸苷(dTTP)、单磷酸脱氧尿苷(dUMP)、二磷酸脱氧尿苷(dUDP)、三磷酸脱氧尿苷(dUTP)、单磷酸脱氧胞苷(dCMP)、二磷酸脱氧胞苷(dCDP)和三磷酸脱氧胞苷(dCTP)。游离核苷酸通常选自AMP、TMP、GMP、CMP、UMP、dAMP、dTMP、dGMP或dCMP。游离核苷酸通常是三磷酸腺苷(ATP)。

[0113] 针对马达蛋白的辅因子是允许马达蛋白发挥作用的因子。辅因子优选是二价金属阳离子。二价金属阳离子优选是Mg²⁺、Mn²⁺、Ca²⁺或Co²⁺。辅因子最优选是Mg²⁺。

[0114] 在本文所述的方法中,马达蛋白与载体结合,使得其可以控制标识符区在检测器(诸如跨膜孔)内的移动。

[0115] 载体在检测器内的移动可以通过任何合适的方式来控制。在一些实施例中,构建体的移动由物理或化学力(电势)驱动。在一些实施例中,物理力由电势(例如,电压)或温度梯度等提供。

[0116] 在一些实施例中,检测器是纳米孔,并且当跨纳米孔施加电势时,构建体相对于该纳米孔移动。多核苷酸带负电荷,并且因此跨纳米孔施加电势将导致多核苷酸在施加的电势的影响下相对于纳米孔移动。例如,如果相对于纳米孔的顺式侧向纳米孔的反式侧施加

正电压电势,则这将诱导带负电荷的分析物从纳米孔的顺式侧移动到纳米孔的反式侧。类似地,如果相对于纳米孔的顺式侧向纳米孔的反式侧施加正电压电势,则这将阻碍带负电荷的分析物从纳米孔的反式侧移动到纳米孔的顺式侧。如果相对于纳米孔的顺式侧向纳米孔的反式侧施加负电压电势,则将发生相反的情况。本文更详细地描述了施加适当电压的装置和方法。在一些实施例中,化学力由浓度(例如pH)梯度提供。

[0117] 样品

[0118] 样品可以是任何合适的样品。样品可以是生物样品。本文所述的任何方法可以在从任何生物体或微生物获得或提取的样品上体外执行。生物体或微生物通常是古生菌、原核生物或真核生物,并且通常属于五个界中的一个界:植物界、动物界、真菌界、原核生物界和原生界。在一些实施例中,本文所述的各个方面的方法可以在从任何病毒获得或提取的样品上体外执行。

[0119] 样品优选是流体样品。样品可以是复合生物流体。样品通常包含体液。体液可以获自人或动物。人或动物可能患有、怀疑患有疾病或处于患病风险中。样品可以是尿液、淋巴液、唾液、粘液、精液、脑脊液或羊水、全血、血浆或血清。通常,样品源自人类,但或者,它可以来自另一哺乳动物(诸如来自商业养殖的动物(诸如马、牛、羊或猪))或者可以另选地是宠物,诸如猫或狗。

[0120] 或者,植物来源的样品通常获自商业作物,诸如谷类、豆类、水果或蔬菜,例如,小麦、大麦、燕麦、油菜籽、玉米、大豆、稻米、香蕉、苹果、西红柿、马铃薯、葡萄、烟草、豆类、扁豆、甘蔗、可可、棉花、茶或咖啡。

[0121] 样品可以是非生物样品。非生物样品优选是流体样品。非生物样品的实例包括手术液、水(诸如饮用水、海水或河水)以及用于实验室测试的试剂。

[0122] 可以在测定之前对样品进行处理,例如通过离心或通过穿过过滤掉不需要的分子或细胞(诸如红细胞)的膜。取样后可以立即进行测量。样品通常也可以在测定前进行储存,优选低于-70°C。

[0123] 在一些实施例中,样品可以包含基因组DNA。基因组DNA可以被片段化,或者本文所述的任何方法还可以包括片段化基因组DNA。DNA可以通过任何合适的方法片段化。例如,片段化DNA的方法是本领域已知的。此类方法可以使用转座酶,诸如MuA转座酶或市售的G-管。

[0124] 本文公开的方法可以用于使用单一测定来检测来自一个或多个样品的一种或多种分子并确定在哪个样品中检测到该分子。当使用包括与特定分子结合区相关联的标识符区和与特定样品相关联的标识符区的载体时,可以实现这一点。一个或多个样品可以是2个或更多个样品,诸如至少3、4、5、10、20、50或100个样品。样品可以例如取自不同患者、患者体内不同类型的组织或者在不同时间点取得。不同时间点可以以秒、分钟、天、月或年来间隔。样品可以包括一个或多个对照样品。

[0125] 可以在没有或最少样品制备的情况下询问样品,或者可以在用于该方法之前对样品进行处理,例如以除去杂质或浓缩待检测的分子类型。使用未处理或最低限度处理的样品的能力导致从样品收集到分析的快速周转。

[0126] 分子

[0127] 本文所述的载体包括对待检测分子具有特异性的分子结合区。本文公开的方法用于检测多种分子。本文提供的术语“分子”可以与术语“分析物”互换使用。

[0128] 分子可以是能够被分子结合区特异性结合的任何分子。例如,分子可以是金属离子、无机盐、聚合物、氨基酸、肽、多肽、蛋白质、核苷酸、寡核苷酸、多核苷酸、染料、漂白剂、药物、诊断剂、消遣性药物、爆炸物和/或环境污染物。分子可以是生物标志物。该方法可以包括:检测两种或更多种相同类型的分子,诸如两种或更多种蛋白质、两种或更多种核苷酸或者两种或更多种药物。该方法可以包括:检测两种或更多种不同类型的分子,诸如一种或多种蛋白质、一种或多种核苷酸和一种或多种药物。

[0129] 分子可以从细胞分泌。或者,分子可以存在于细胞内部,使得在可以执行该方法之前必须从细胞中提取分子。

[0130] 在一种实施例中,分子选自氨基酸、肽、多肽、蛋白质、核苷酸、寡核苷酸和/或多核苷酸。

[0131] 在一种实施例中,分子选自氨基酸、肽、多肽和/或蛋白质。氨基酸、肽、多肽或蛋白质可以是天然存在的或非天然存在的。多肽或蛋白质可以在其中包括合成的或经修饰的氨基酸。对氨基酸的许多不同类型的修饰是本领域已知的。下面参考跨膜孔讨论合适的氨基酸及其修饰。出于本公开的目的,应当理解,可以通过本领域可用的任何方法来修饰分子。

[0132] 蛋白质可以是酶、抗体、激素、生物标志物、生长因子或生长调节蛋白,诸如细胞因子。细胞因子可以选自白细胞介素(诸如IL-1、IL-1、IL-2、IL-4、IL-5、IL-6、IL-10、IL-12和IL-13)、干扰素(诸如IFN- γ)以及其他细胞因子(诸如TNF- α)。蛋白质可以是细菌蛋白质、真菌蛋白质、病毒蛋白质或源于寄生虫的蛋白质。

[0133] 在一个实施例中,分子选自核苷酸、寡核苷酸和/或多核苷酸。核苷酸通常含有核碱基、糖和至少一个磷酸基团。核碱基通常是杂环的。核碱基包括但不限于嘌呤和嘧啶,并且更具体地为腺嘌呤、鸟嘌呤、胸腺嘧啶、尿嘧啶和胞嘧啶。糖通常是戊糖。核苷酸糖包括但不限于核糖和脱氧核糖。核苷酸通常是核糖核苷酸或脱氧核糖核苷酸。核苷酸通常含有单磷酸盐、二磷酸盐或三磷酸盐。磷酸盐可以附接在核苷酸的5'或3'侧。

[0134] 核苷酸包括但不限于单磷酸腺苷(AMP)、二磷酸腺苷(ADP)、三磷酸腺苷(ATP)、单磷酸鸟苷(GMP)、二磷酸鸟苷(GDP)、三磷酸鸟苷(GTP)、单磷酸胸苷(TMP)、二磷酸胸苷(TDP)、三磷酸胸苷(TTP)、单磷酸尿苷(UMP)、二磷酸尿苷(UDP)、三磷酸尿苷(UTP)、单磷酸胞苷(CMP)、二磷酸胞苷(CDP)、三磷酸胞苷(CTP)、5-甲基胞苷单磷酸、5-甲基胞苷二磷酸、5-甲基胞苷三磷酸、5-羟甲基胞苷单磷酸、5-羟甲基胞苷二磷酸、5-羟甲基胞苷三磷酸、单磷酸环腺苷(cAMP)、单磷酸环鸟苷(cGMP)、单磷酸脱氧腺苷(dAMP)、二磷酸脱氧腺苷(dADP)、三磷酸脱氧腺苷(dATP)、单磷酸脱氧鸟苷(dGMP)、二磷酸脱氧鸟苷(dGDP)、三磷酸脱氧鸟苷(dGTP)、单磷酸脱氧胸苷(dTMP)、二磷酸脱氧胸苷(dTDP)、三磷酸脱氧胸苷(dTTP)、单磷酸脱氧尿苷(dUMP)、二磷酸脱氧尿苷(dUDP)、三磷酸脱氧尿苷(dUTP)、单磷酸脱氧胞苷(dCMP)、二磷酸脱氧胞苷(dCDP)和三磷酸脱氧胞苷(dCTP)、5-甲基-2'-脱氧胞苷单磷酸、5-甲基-2'-脱氧胞苷二磷酸、5-甲基-2'-脱氧胞苷三磷酸、5-羟甲基-2'-脱氧胞苷单磷酸、5-羟甲基-2'-脱氧胞苷二磷酸以及5-羟甲基-2'-脱氧胞苷三磷酸。核苷酸优选选自AMP、TMP、GMP、UMP、dAMP、dTMP、dGMP或dCMP。核苷酸可以是无碱基的(即缺乏核碱基)。核苷酸可以含有附加的修饰。特别地,合适的修饰核苷酸包括但不限于:2'氨基嘧啶(诸如2'-氨基胞苷和2'-氨基尿苷)、2'-羟基嘌呤(诸如2'-氟嘧啶(诸如2'-氟胞苷和2'氟尿苷)、羟基嘧啶(诸如5'- α -对硼酸尿苷)、2'-邻甲基核苷酸(诸如2'-邻甲基腺苷、2'-邻甲基鸟苷、

2'-邻甲基胞苷和2'-邻甲基尿苷)、4'-硫代嘧啶(诸如4'-硫代尿苷和4'-硫代胞苷),并且核苷酸具有核碱基修饰(诸如5-戊炔基-2'-脱氧尿苷、5-(3-氨基丙基)-尿苷和1,6-二氨基己基-N-5-氨基甲酰基甲基尿苷)。

[0135] 寡核苷酸是短核苷酸聚合物,其通常具有50或更少个核苷酸,诸如40或更少、30或更少、20或更少、10或更少或5或更少个核苷酸。寡核苷酸可以包括上文讨论的任何核苷酸,包括无碱基和经修饰的核苷酸。

[0136] 多核苷酸可以是单链的或双链的。多核苷酸的至少一部分可以是双链的。多核苷酸可以是核酸,诸如脱氧核糖核酸(DNA)或核糖核酸(RNA)。多核苷酸可以包括与DNA的一条链杂交的RNA的一条链。多核苷酸可以是本领域已知的任何合成核酸,诸如肽核酸(PNA)、甘油核酸(GNA)、苏糖核酸(TNA)、锁核酸(LNA)或具有核苷酸侧链的其他合成聚合物。多核苷酸可以包括上文讨论的任何核苷酸,包括经修饰的核苷酸。

[0137] 多核苷酸可以为任何长度。例如,多核苷酸在长度上可以是至少10、至少50、至少100、至少150、至少200、至少250、至少300、至少400或至少500个核苷酸或核苷酸对。多核苷酸在长度上可以是1000个或更多个核苷酸或核苷酸对、5000个或更多个核苷酸或核苷酸对或者100000个或更多个核苷酸或核苷酸对。

[0138] 在一种实施例中,分子是微小RNA(miRNA)。MiRNA是单链RNA多核苷酸分子,其在基因表达的转录后调控中发挥作用。

[0139] 这些分子可能与特定表型或与特定类型的细胞相关联。例如,这些分子可以指示细菌细胞。这些分子可以指示病毒、真菌或寄生虫。这些分子可能是一组特定的消遣性药物(诸如SAMHSA 5组测试)、爆炸物或环境污染物。

[0140] 在一个实施例中,分子是可以用于诊断或预测疾病或病症的生物标志物。生物标志物可以是上述任何分子,诸如蛋白质或多核苷酸。合适的生物标志物组是本领域已知的,例如,如描述于Edwards, A.V.G.等人(2008) *Mol. Cell. Proteomics* 7, 第1824-1837页; Jacquet, S.等人(2009) *Mol. Cell. Proteomics* 8, 第2687-2699页; Anderson N.L.等人(2010) *Clin. Chem.* 56, 177-185中。疾病或病症优选是癌症、冠心病、心血管疾病或败血症。

[0141] 在一个实施例中,分子是神经递质。神经递质是通过突触在细胞之间传递信号的分子。神经递质的实例包括乙酰胆碱、多巴胺、肾上腺素、去甲肾上腺素、核苷酸(诸如ATP)、氨基酸(诸如谷氨酸、天冬氨酸和 δ -氨基丁酸)以及脑啡肽。

[0142] 前导序列

[0143] 本公开的载体包括单链前导序列。前导序列通常包括聚合物,诸如多核苷酸(例如DNA或RNA)、经修饰的多核苷酸(诸如无碱基DNA)、PNA、LNA、聚乙二醇(PEG)或多肽。在一些实施例中,前导序列包括DNA的单链,诸如聚dT区段。前导序列可以为任何长度,但在长度上通常为10至150个核苷酸,诸如在长度为20至120、30至100、40至80或50至70个核苷酸。

[0144] 标识符区

[0145] 本公开的载体包括标识符区。载体可以包括多于一个标识符区,诸如2个或更多个、3个、4个、5个或更多个(诸如例如约10个)标识符区。在一些实施例中,载体上的不同标识符区可以与不同分子结合区相关联,使得当标识符区在检测器内通过时,可以确定一个或多个相关联的分子结合区的身份。因此,载体可以包括一系列标识符区和分子结合区。载体被布置成使得标识符区在检测器内的移动由与载体结合的马达蛋白控制。

[0146] 在一些实施例中,载体上多于一个标识符区的存在可以用于区分来自不同样品的载体。在一些实施例中,包括多于一个标识符区的载体可以用于检测多个样品中的多种分子的方法,并且因此载体可以包括例如对样品唯一的一个标识符区和对与该载体结合的分子唯一的一个标识符区。载体的标识符区被定位成使得当该载体接触跨膜孔时,与该载体结合的马达蛋白控制标识符区在跨膜孔内的移动。

[0147] 标识符区的目的是充当与载体结合的一个或多个分子的身份的唯一信号。另一标识符区可以充当载体来源的唯一信号,例如以识别该载体已经接触的样品。在一些实施例中,标识符区是多核苷酸或包括多核苷酸序列。核苷酸可以是以下讨论的任何核苷酸。标识符多核苷酸在长度上可以是2至300个核苷酸,诸如2至200、2至100、2至75、2至50、2至40、2至30、2至25、4至100、4至75、4至50、4至40、4至30、4至25、4至20、4至15、4至10、6至100、6至75、6至50、6至40、6至30、6至25、6至20、6至15、6至10、8至100、8至75、8至50、8至40、8至30、8至25、8至20或8至15个核苷酸。

[0148] 在一些实施例中,分子结合区或其部分是标识符区。例如,标识符区可以与分子结合区重叠。在此类实施例中,分子结合区和标识符区优选是多核苷酸。当分子结合区是多核苷酸,诸如适体或与靶多核苷酸杂交的多核苷酸时,分子结合区的多核苷酸序列对于结合的分子是唯一的,并且因此起到识别该分子的作用。在一些实施例中,标识符区和分子结合区不重叠。

[0149] 在一些实施例中,标识符区包括条形码序列。多核苷酸条形码是本领域众所周知的(Kozarewa, I. 等人, (2011), *Methods Mol. Biol.* 733, 第279-298页)。条形码是多核苷酸的特定序列,其以特定且已知的方式影响流过孔的电流。条形码序列在长度上通常为2或更多个核苷酸,诸如4或更多、8或更多或者12或更多个核苷酸。在一些实施例中,条形码序列在长度上为2至50个核苷酸,诸如2至45、2至40、2至35、2至30、2至25、4至50、4至45、4至40、4至35、4至30、4至25、4至20、4至15、4至10、6至50、6至45、6至40、6至35、6至30、6至25、6至20、6至15、6至10、8至50、8至45、8至40、8至35、8至30、8至25、8至20或8至15个核苷酸。在一些实施例中,条形码序列在长度上为10至50、10至45、10至40、10至35、10至30、10至25或10至20个核苷酸。在一些实施例中,条形码序列在长度上为15至50、15至45、15至40、15至35、15至30或15至25个核苷酸。在一些实施例中,条形码序列在长度上为20至50、20至45、20至40、20至35、20至30个核苷酸。在一些实施例中,条形码序列在长度上为25至50、25至45、25至40或25至35个核苷酸。在一些实施例中,条形码序列在长度上约为10、15、20、25、30、35、40、45或50个核苷酸。

[0150] 条形码的长度越长,可以使用的独特组合的数量就越多。为了提高条形码测序的准确度,可以在载体中重复条形码以使得条形码能够被校对。因此,标识符区可以包括2个或更多个(诸如3至10个,例如3个或更多个、4个或更多个、5个或更多个或6个或更多个)条形码拷贝。

[0151] 条形编码允许强调多重检测。例如,4个碱基的条形码将生成($4^4=256$)个独特的配置,其最多将允许256个蛋白质或miRNA靶标。碱基数量可以增加,生成例如 4^8 、 4^{12} 个独特组合。例如,8个碱基的序列可以用于生成65,536个独特的条形码。这通常是感测和诊断方面的巨大进步,其中通常在任何时间只能选择性地探测一个或少数分子,诸如多达约5或约10个分子。

[0152] 在一些实施例中,标识符区可以包括间隔子或一系列间隔子,如本文所述。一系列间隔子可以包括2个或更多个(例如,3个或更多个、4个或更多个、5个或更多个或6个或更多个、7个或更多个、8个或更多个、9个或更多个或者10个或更多个)间隔子。一系列间隔子可以包括20个或更多个、50个或更多个或100个或更多个间隔子。一系列间隔子可以包括2至1000个间隔子,诸如2至100、2至50、2至20或2至10个间隔子。一系列间隔子中的间隔子可以相同或不同。当标识符区在孔内移动时,可以测量间隔子的特征信号。不同载体分子中的间隔子的类型和数量可以基于所测量的信号来区分。一个或多个间隔子可以是本文所述的任何间隔子,例如,iSp9和iSp18间隔子。

[0153] 分子结合区

[0154] 本公开的载体包括对待检测分子具有特异性的分子结合区。在一些实施例中,载体可以包括多于一个分子结合区。载体可以包括一个或多个相同类型的分子结合区和/或可以包括两个或更多个不同分子结合区。两个或更多个分子结合区可以是3个或更多个(诸如4个、5个、6个或更多个(诸如例如约10个))分子结合区。载体中的不同分子结合区通常与不同分子特异性结合。这允许使用单个载体检测多个分子(分析物)。

[0155] 在载体中,标识符区可以与每个分子结合区相关联。通常,标识符区将被定位成使得其在其相关联的分子结合区与检测器相互作用之前穿过检测器(诸如孔)。

[0156] 在载体中,一个标识符区可以与一个或多个分子结合区(诸如2个或更多个、3个或更多个(例如,4至10个)分子结合区)相关联。在这种情况下,2个或更多个分子结合区通常与同一分子结合。2个或更多个分子结合区可以对不同分子具有特异性。2个或更多个分子结合区可以被间隔子或一系列间隔子(诸如本文定义的那些)分开。间隔子可以将分子结合区与相关联的标识符区分开。

[0157] 可以使用任何分子结合区,只要它与感兴趣分子特异性结合,使得当载体与孔接触时,可以确定与分子结合区结合的分子的存在或不存在。例如,分子结合区可以是:适体;互补DNA序列;肽或蛋白质,诸如抗体、抗体片段、纳米抗体或亲和体;点击化学反应基团;生物素或链霉亲和素;或其他。

[0158] 在一个实施例中,分子结合区是适体。适体是与一个或多个分子结合的小分子。合适的适体和产生适体的方法是本领域已知的并且例如在WO 2013/121201中进行了描述,该专利通过引用的方式并入本文。适体可以使用SELEX (Stoltenburg, R. 等人, (2007), *Biomolecular Engineering* 24, 第381-403页; Tuerk, C. 等人, *Science* 249, 第505-510页; Bock, L.C. 等人, (1992), *Nature* 355, 第564-566页) 或NON-SELEX (Berezovski, M. 等人 (2006), *Journal of the American Chemical Society* 128, 第1410-1411页) 产生。适体可以是肽适体或寡核苷酸适体。在一个实施例中,适体是肽适体。肽适体可以包括任何氨基酸。氨基酸可以是以下讨论的任何氨基酸。在一个实施例中,适体是寡核苷酸适体。寡核苷酸适体可以包括任何核苷酸。核苷酸可以是以上讨论的任何核苷酸。适体可以为任何长度。适体在长度上通常为至少15个氨基酸或核苷酸,诸如在长度上为约15至约50、约20至约40或约25至约30个氨基酸或核苷酸。

[0159] 在一个实施例中,分子结合区是多核苷酸。在一个实施例中,多核苷酸是适体。在一个实施例中,多核苷酸包括与待检测的多核苷酸分子(靶多核苷酸)互补的序列。靶多核苷酸可以为任何长度。例如,靶多核苷酸在长度上可以是至少10、至少50、至少100、至少

150、至少200、至少250、至少300、至少400或至少500个核苷酸或核苷酸对。靶多核苷酸可以是寡核苷酸。寡核苷酸是短核苷酸聚合物,其通常具有50或更少个核苷酸,诸如40或更少、30或更少、20或更少、10或更少或5或更少个核苷酸。分子结合区优选与miRNA互补。miRNA是在转录后基因调控中起作用的短非编码RNA,并且在长度上通常为21至23个核苷酸,但在长度上可以为18至30个核苷酸,诸如在长度上为20至25个核苷酸。

[0160] 在待检测的分子是多核苷酸的情况下,分子结合区可以包括与靶多核苷酸的补体具有90%或更高(诸如至少97%、98%或99%)同一性的序列。在这样的分子结合区中,补体中的一个或多个(例如,2个、3个、4个或5个)核苷酸可以被可以与靶多肽中的相应核苷酸碱基配对的非规范核苷酸替代。优选地,分子结合区包括靶多核苷酸的补体。

[0161] 在一个实施例中,分子结合区是抗体、抗体片段、纳米抗体或亲和体。如本文使用的术语“抗体”可以涉及整个抗体(包括两条重链和两条轻链)以及其抗原结合片段。抗体可以包括但不限于多克隆、单克隆、嵌合、dAb(结构域抗体)、单链、Fab、Fab'和F(ab')₂片段以及scFv。纳米抗体是单结构域抗体,诸如V_H片段或V_{NAR}片段。亲和体是包括三螺旋支架结构域的抗体模拟物,在三个螺旋中的两个螺旋上具有氨基酸取代,从而允许氨基酸序列和潜在抗原结合的巨大多样性。亲合体在Frejd、Fredrik Y.和Kyu-Tae Kim. (Experimental & molecular medicine 49.3 (2017): e306-e306) 和 Löfblom, John等人 (FEBS letters 584.12 (2010): 2670-2680) 中讨论。合适的抗体、抗体片段、纳米抗体和亲和体是本领域已知的或者可以通过标准方法制备。

[0162] 附接多肽的方法是本领域众所周知的。例如,可以使用利用分选酶介导的反应对蛋白质进行位点特异性C末端、N末端或内环标记,如在Guimaraes等人Nature protocols 8.9 (2013): 1787, Theile等人Nature protocols 8.9 (2013): 1800以及Koussa等人Methods 67.2 (2014): 134-141中所述,所有这些均通过引用的方式并入本文。技术人员可以利用合适的技术将蛋白质(诸如抗体)掺入载体中。

[0163] 对待检测分子具有特异性的分子结合区能够以比其与无关分子结合时更大的亲和力与其预期靶分子(其旨在检测的分子)结合。当待检测分子是蛋白质时,无关分子可以是无关对照蛋白质,诸如牛血清白蛋白。当待检测的分子是多核苷酸时,无关分子可以是乱序对照多核苷酸(例如,具有与预期靶分子相同数量和类型的核苷酸的随机多核苷酸序列)。待检测分子优选以是对照至少10、至少50、至少100、至少500或至少1000倍大的亲和力与分子结合区结合。亲和力可以通过本领域已知的方法确定。例如,亲和力可以通过ELISA测定、生物层干涉测量法、表面等离子体共振、动力学方法或平衡/溶液方法来确定。技术人员将认识到哪些分子特异性结合分子结合区。

[0164] 例如,与预期靶miRNA分子具有相似序列的miRNA多核苷酸或者与共享密切相关结构域的蛋白质可能会发生一些交叉反应。优选地,分子结合区以更大的亲和力与其靶分子结合,例如,是相关分子(诸如相关多核苷酸(例如,具有相似序列的miRNA多核苷酸)或相关蛋白质(诸如同源物))至少10、至少50、至少100、至少500或至少1000倍大的亲和力。

[0165] 间隔子

[0166] 在本文提供的方法的一些实施例中,载体包括间隔子。间隔子优选定位在所结合的马达蛋白和分子结合区之间。当马达蛋白与间隔子相互作用时,载体在孔内的移动停滞(或者换句话说,减慢或延迟)。间隔子更优选紧邻分子结合区定位。当载体在检测器内移动

时,在分子结合区与孔相互作用之前,马达蛋白的移动被间隔子停滞。当马达蛋白停滞在间隔子处时,与没有间隔子的类似载体相比,当分子与分子结合区结合时,可能会产生放大的光学信号或电信号。当标识符区与分子结合区分开时,间隔子优选位于标识符区和分子结合区之间的载体中。

[0167] 当载体移动通过检测器时,例如当载体相对于纳米孔移动时,在马达蛋白遇到间隔子时会产生独特的电信号或光学信号。例如,定位在所结合的马达蛋白与分子结合区之间的间隔子可以充当独特的信号,以允许清楚地识别分子结合区与检测器相互作用时产生的信号,例如,位于载体在纳米孔内移动时产生的信号/迹线/曲线中。因此,间隔子可以用作标记物来定位分子结合区在检测器内移动时产生的信号,从而有助于确定与分子结合区特异性结合的分子的存在或不存在。

[0168] 间隔子可以提供阻碍马达蛋白移动的能量屏障。例如,间隔子可以通过减少马达蛋白对多核苷酸的牵引力来使马达蛋白停滞。这可以例如通过使用无碱基间隔子(即其中碱基从载体中的一个或多个核苷酸中去除的间隔子)来实现。

[0169] 间隔子可以物理地阻止马达蛋白的移动,例如通过引入大的化学基团来物理地阻碍马达蛋白的移动。间隔子可以是多核苷酸的双链区。

[0170] 间隔子可以包括线性分子,诸如聚合物。通常,线性间隔子具有与靶多核苷酸不同结构。例如,如果靶多核苷酸是DNA,则该间隔子或每个间隔子通常不包括DNA。具体地,如果靶多核苷酸是脱氧核糖核酸(DNA)或核糖核酸(RNA),则该间隔子或每个间隔子优选包括肽核酸(PNA)、甘油核酸(GNA)、苏糖核酸(TNA)、锁核酸(LNA)或具有核苷酸侧链的合成聚合物。在一些实施例中,间隔子可以包括一个或多个硝基吡啶、一个或多个肌苷、一个或多个嘧啶、一个或多个2-氨基嘌呤、一个或多个2-6-二氨基嘌呤、一个或多个5-溴-脱氧尿苷、一个或多个反向胸苷(反向dT)、一个或多个反向双脱氧胸苷(ddT)、一个或多个双脱氧胞苷(ddC)、一个或多个5-甲基胞苷、一个或多个5-羟甲基胞苷、一个或多个2'-邻甲基RNA碱基、一个或多个异脱氧胞苷(Iso-dCs)、一个或多个异脱氧鸟苷(Iso-dGs)、一个或多个C3($\text{OC}_3\text{H}_6\text{OPO}_3$)基团、一个或多个光可裂解(PC) [$\text{OC}_3\text{H}_6-\text{C}(\text{O})\text{NHCH}_2-\text{C}_6\text{H}_3\text{NO}_2-\text{CH}(\text{CH}_3)\text{OPO}_3$]基团、一个或多个己二醇基团、一个或多个间隔子9(iSp9) [$(\text{OCH}_2\text{CH}_2)_3\text{OPO}_3$]基团、或一个或多个间隔子18(iSp18) [$(\text{OCH}_2\text{CH}_2)_6\text{OPO}_3$]基团;或一个或多个硫醇连接。间隔子可以包括这些基团的任何组合。这些基团中的许多基团可从IDT®(Integrated DNA Technologies®)商购。例如,C3、iSp9和iSp18间隔子均可从IDT®获得。间隔子可以包括任何数量的上述基团作为间隔子单元。

[0171] 在一些实施例中,间隔子可以包括一种或多种导致马达蛋白停滞的化学基团。在一些实施例中,合适的化学基团是一种或多种侧基化学基团。一种或多种化学基团可以附接到载体中的一种或多种核碱基。一种或多种化学基团可以附接到载体的主链。可以存在任何数量的合适的化学基团,诸如2、3、4、5、6、7、8、9、10、11、12或更多个。合适的基团包括但不限于荧光团、链霉亲和素和/或生物素、胆固醇、亚甲蓝、二硝基苯酚(DNP)、地高辛和/或抗地高辛基团和二苄基环辛炔基团。在一些实施例中,间隔子可以包括淋巴细胞。在一些实施例中,间隔子可以包括聚合物,其为多肽或聚乙二醇(PEG)。

[0172] 在一些实施例中,间隔子可以包括一种或多种无碱基核苷酸(即,缺乏核碱基的核苷酸),诸如2、3、4、5、6、7、8、9、10、11、12或更多种脱碱基核苷酸。核碱基可以被无碱基核苷

酸中的-H(idSp)或-OH取代。可以通过从一个或多个相邻核苷酸去除核碱基来将无碱基间隔子插入靶多核苷酸中。例如,可以修饰多核苷酸以包括3-甲基腺嘌呤、7-甲基鸟嘌呤、1,N6-乙烯腺嘌呤肌苷或次黄嘌呤,并且可以使用人烷基腺嘌呤DNA糖基化酶(hAAG)从这些核苷酸中去除核碱基。或者,可以修饰多核苷酸以包括尿嘧啶和用尿嘧啶-DNA糖基化酶(UDG)去除的核碱基。在一个实施例中,一个或多个间隔子不包括任何无碱基核苷酸。

[0173] 一个或多个间隔子可以存在于载体的其他地方。间隔子可以包括任何合适数量的间隔子。例如,载体可以包括两个或更多个、3个或更多个或5个或更多个间隔子,诸如1至约20个间隔子,例如1至约10个间隔子。

[0174] 停滞区

[0175] 在一些实施例中,载体包括停滞区。当载体处于溶液中时(即在与孔接触并在孔内移动之前),载体的停滞区为马达蛋白提供定位在载体上的位置。停滞区通常存在于前导区和标识符区之间。这使得前导区能够与检测器相互作用,诸如穿入孔中。它也将马达蛋白定位在载体上,使得它准备好控制在载体与检测器相互作用时标识符区通过检测器(例如孔)的移动。

[0176] 在一些实施例中,停滞区是间隔子,如本文和WO 2020/234612中所述。载体还可以包括封闭部分,其防止马达蛋白从间隔子移开。

[0177] 封闭部分通常是防止马达蛋白在与马达蛋白天然加工多核苷酸的方向相反的方向上移动的部分。例如,如果马达蛋白在5'至3'的方向上天然加工多核苷酸链,则合适的封闭部分可以是防止马达蛋白在3'至5'的方向上移动的部分。类似地,如果马达蛋白在3'至5'的方向上天然加工多核苷酸链,则合适的封闭部分可以是防止马达蛋白在5'至3'的方向上移动的部分。

[0178] 封闭部分通常与载体结合,以防止马达蛋白从间隔子移开。防止马达蛋白从间隔子移开可以通过提供立体块以物理方式阻止马达蛋白的移动来实现。防止马达蛋白从间隔子移开可以通过使用化学封闭部分来实现,马达蛋白不能移动越过或通过该化学封闭部分。在一些实施例中,封闭部分包括本文讨论的间隔子基团中的一种或多种间隔子基团。在其他实施例中,封闭部分可以包括多核苷酸链。

[0179] 载体也可以包括连接至停滞区或间隔子的加载位点。加载位点是用于将马达蛋白加载到多核苷酸衔接子上的位点。合适的加载位点更详细地描述于WO 2020/234612中。

[0180] 将马达蛋白加载到多核苷酸上、使马达蛋白停滞在溶液中的多核苷酸上的方法、合适的间隔子和合适的封闭部分更详细地描述于WO 2020/234612(其通过引用的方式并入本文)和WO 2014/135838(其通过引用的方式并入本文)中。

[0181] 锚定物

[0182] 在一些实施例中,载体包括附接到载体的膜锚定物或跨膜孔锚定物。锚定物可以共价或非共价附接到载体。例如,锚定物可以附接到与载体的多核苷酸区杂交的寡核苷酸。载体的与锚定物寡核苷酸杂交的多核苷酸区就其不阻止分子与分子结合区的特异性结合而言与分子结合区不同。

[0183] 在一些实施例中,锚定物有助于根据本文公开的方法表征靶多核苷酸。例如,在包括使载体与跨膜孔接触的方法中,膜锚定物或跨膜孔锚定物可以促进所选载体在跨膜孔周围的定位。术语锚定物和系链在本文中可互换使用。

[0184] 锚定物可以是插入膜中的多肽锚定物和/或疏水锚定物。在一个实施例中,疏水锚定物是脂质、脂肪酸、甾醇、碳纳米管、多肽、蛋白质或氨基酸,例如胆固醇、棕榈酸酯或生育酚。锚定物可以包括硫醇、生物素或表面活性剂。

[0185] 一方面,锚定物可以是生物素(用于与链霉亲和素结合)、直链淀粉(用于与麦芽糖结合蛋白或融合蛋白结合)、Ni-NTA(用于与多聚组氨酸或多聚组氨酸标记蛋白结合)或肽(诸如抗原)。

[0186] 在一个实施例中,锚定物包括一个接头或2个、3个、4个或更多个接头。优选的接头包括但不限于聚合物,诸如多核苷酸、聚乙二醇(PEG)、多糖和多肽。这些接头可以是直链、支链或环状的。例如,接头可以是环状多核苷酸。衔接子可以与环状多核苷酸接头上的互补序列杂交。一种或多种锚定物或一种或多种接头可以包括可以被切割或分解的组分,诸如限制位点或光不稳定基团。接头可以用马来酰亚胺基团功能化以附接到蛋白质中的半胱氨酸残基。合适的接头描述于WO 2010/086602中。

[0187] 在一个实施例中,锚定物是胆固醇或脂肪酰基链。例如,可以使用长度为6至30个碳原子的任何脂肪酰基链,诸如十六烷酸。合适的锚定物和将锚定物附接到衔接子的方法的实例公开于WO 2012/164270和WO 2015/150786中。可以使用相同的方法将锚定物附接到载体。

[0188] 检测器

[0189] 任何合适的检测器都可以用于本文所述的方法中。检测器可以是在测序方法中有用的任何检测器。例如,纳米孔测序或单分子实时测序(例如通过合成测序)技术。优选地,本文使用的方法中的检测器是纳米孔。任何合适的纳米孔都可以用在本文所述的方法中。在一个实施例中,纳米孔是跨膜孔。

[0190] 跨膜孔是在某种程度上穿过膜的结构。它允许由施加的电势驱动的水合离子流过膜或在膜内流动。跨膜孔通常穿过整个膜,使得水合离子可以从膜的一侧流到膜的另一侧。然而,跨膜孔不必穿过膜。它的一端可以是封闭的。例如,孔可以是膜中的孔、间隙、通道、沟槽或狭缝,水合离子可以沿着其流动或流入其中。

[0191] 任何跨膜孔均可以用在本文提供的方法中。孔可以是生物的或人造的。合适的孔包括但不限于蛋白质孔、多核苷酸孔和固态孔。孔可以是DNA折纸孔(Langecker等人, Science, 2012; 338: 932-936)。合适的DNA折纸孔公开于WO 2013/083983、WO 2018/011603和WO 2020/025974中。

[0192] 在一个实施例中,纳米孔是跨膜蛋白孔。跨膜蛋白孔是允许水合离子(诸如多核苷酸)从膜的一侧流到该膜的另一侧的多肽或多肽的集合。在本文提供的方法中,跨膜蛋白孔能够形成允许由施加的电势驱动的水合离子从膜的一侧流到另一侧的孔。跨膜蛋白孔优选允许多核苷酸从膜(诸如三嵌段共聚物膜)的一侧流到另一侧。跨膜蛋白孔允许多核苷酸移动通过孔。

[0193] 在一个实施例中,纳米孔是跨膜蛋白孔,其是单体或寡聚物。孔优选地由数个重复亚基组成,诸如至少6、至少7、至少8、至少9、至少10、至少11、至少12、至少13、至少14、至少15或至少16个亚基。孔优选是六聚体、七聚体、八聚体或九聚体孔。孔可以是同源低聚物或异源低聚物。

[0194] 在一个实施例中,跨膜蛋白孔包括离子可以流过的桶或通道。孔的亚基通常围绕

中心轴并为跨膜 β -桶或通道或跨膜 α -螺旋束或通道贡献链。

[0195] 通常,跨膜蛋白孔的桶或通道包括促进与分析物诸如靶多核苷酸(如本文所述)相互作用的氨基酸。这些氨基酸优选位于桶或通道的缩颈附近。跨膜蛋白孔通常包括一种或多种带正电荷的氨基酸,诸如精氨酸、赖氨酸或组氨酸或芳香族氨基酸,诸如酪氨酸或色氨酸。这些氨基酸通常促进孔与核苷酸、多核苷酸或核酸之间的相互作用。

[0196] 在一个实施例中,纳米孔是源自 β -桶孔或 α -螺旋束孔的跨膜蛋白孔。 β -桶孔包括由 β -链形成的桶或通道。合适的 β -桶孔包括但不限于 β -毒素(诸如 α -溶血素、炭疽毒素和杀白细胞素)和细菌的外膜蛋白/孔蛋白(诸如耻垢分枝杆菌孔蛋白(Msp),例如MspA、MspB、MspC或MspD、CsgG、外膜孔蛋白F(OmpF)、外膜孔蛋白G(OmpG)、外膜磷脂酶A和奈瑟菌自转运脂蛋白(NalP))以及其他孔,诸如溶素。 α -螺旋束孔包括由 α -螺旋形成的桶或通道。合适的 α -螺旋束孔包括但不限于内膜蛋白和 α 外膜蛋白,诸如WZA和ClyA毒素。

[0197] 在一个实施例中,纳米孔是源自或基于Msp、 α -溶血素(α -HL)、胞溶素、CsgG、ClyA、Sp1或溶血蛋白茶香毒素C(FraC)的跨膜孔。

[0198] 在一个实施例中,纳米孔是源自CsgG的跨膜蛋白孔,例如来自大肠杆菌链K-12子链MC4100的CsgG。这样的孔是低聚的并且通常包括7、8、9或10个源自CsgG的单体。孔可以是源自包括相同单体的CsgG的同源寡聚孔。或者,孔可以是源自CsgG的异源低聚孔,其包括至少一个不同于其他单体的单体。源自CsgG的合适孔的实例公开于例如WO 2016/034591、WO 2017/149316、WO 2017/149317、WO 2017/149318、WO 2018/211241和WO 2019/002893中。

[0199] 在一个实施例中,纳米孔是源自溶菌素的跨膜孔。源自溶菌素的合适孔的实例公开于WO 2013/153359中。

[0200] 在一个实施例中,纳米孔是源自或基于 α -溶血素(α -HL)的跨膜孔。野生型 α -溶血素孔由7个相同的单体或亚基组成(即,它是七聚体)。 α -溶血素孔可以是 α -溶血素-NN或其变体。变体优选地在位置E111和K147处包括N残基。

[0201] 在一个实施例中,纳米孔是源自Msp(例如,来自MspA)的跨膜蛋白孔。源自MspA的合适孔的实例公开于WO 2012/107778中。

[0202] 在一个实施例中,纳米孔是源自或基于ClyA的跨膜孔。源自ClyA的合适孔的实例公开于WO 2014/153625中。

[0203] 在一个实施例中,检测器是纳米移液管。纳米移液管通常具有约10nm的直径,诸如约10nm至约12、约15、约18或约20nm。纳米移液管可以由石英毛细管、玻璃和/或碳制成,诸如涂覆有碳层的玻璃。合适的纳米移液管在本领域是已知的。

[0204] 膜

[0205] 在包括使用纳米孔的实施例中,纳米孔通常存在于膜中,例如纳米孔穿过膜和/或提供穿过膜的通道。可以使用任何合适的膜并且合适的膜是本领域已知的。

[0206] 膜优选是两亲层。两亲层是由两亲分子(诸如磷脂)形成的层,其具有亲水性和亲脂性。两亲分子可以是合成的或天然存在的。非天然存在的两亲物和形成单层的两亲物是本领域已知的并且包括例如嵌段共聚物(Gonzalez-Perez等人,Langmuir,2009,25,10447-10450)。

[0207] 嵌段共聚物可以是二嵌段(由两个单体亚基组成),但也可以由两个以上的单体亚基构建以形成表现为两亲物的更复杂的排列。共聚物可以是三嵌段、四嵌段或五嵌段共聚

物。膜优选是三嵌段共聚物膜。

[0208] 膜可以例如是国际申请号W02014/064443或W02014/064444中公开的膜中的一者。

[0209] 两亲分子可以被化学修饰或功能化以促进锚定物的偶联。两亲层可以是单层或双层。两亲层通常是平面的。两亲层可以是弯曲的。两亲层可以被支撑。

[0210] 两亲膜通常是天然可移动的,本质上充当二维流体,脂质扩散速率约为 $10^{-8} \text{ cm s}^{-1}$ 。这意味着孔和锚定载体通常可以在两亲膜内移动。

[0211] 膜可以是脂质双层。脂质双层可以是任何脂质双层。合适的脂质双层包括但不限于平面脂质双层、支撑双层或脂质体。脂质双层优选是平面脂质双层。合适的脂质双层公开于W0 2008/102121、W0 2009/077734和W0 2006/100484中。

[0212] 脂质双层可以由干燥的脂质形成,如W0 2009/077734中所述。脂质双层可以跨开口形成,如W02009/077734中所述。

[0213] 可以使用形成脂质双层的任何脂质组合物。选择脂质组合物,使得形成具有所需特性(诸如表面电荷、支撑膜蛋白的能力、堆积密度或机械特性)的脂质双层。脂质组合物可以包括一种或多种不同脂质。例如,脂质组合物可以含有多达100种脂质。脂质组合物优选含有1至10种脂质。脂质组合物可以包括天然存在的脂质和/或人造脂质。

[0214] 膜可以包括固态层。固态层可以由有机和无机材料形成,包括但不限于微电子材料、绝缘材料(诸如 Si_3N_4 、 Al_2O_3 和 SiO)、有机和无机聚合物(诸如聚酰胺、塑料(诸如Teflon®)或弹性体(诸如双组分加成固化硅橡胶))以及玻璃。固态层可以由石墨烯形成。合适的石墨烯层公开于W0 2009/035647中。如果膜包括固态层,则孔通常存在于固态层内含有的两亲膜或层中,例如,固态层内的洞、孔、间隙、通道、沟槽或狭缝内。技术人员可以制备合适的固态/两亲混合系统。合适的系统公开于W0 2009/020682和W0 2012/005857。可以使用上面讨论的任何两亲膜或层。

[0215] 本文公开的方法可以使用(i)包括孔的人造两亲层、(ii)包括孔的分离的天然存在的脂质双层来执行。人造两亲层通常是人造三嵌段共聚物层。除了孔之外,层还可以包括其他跨膜和/或膜内蛋白质以及其他分子。下面讨论合适的装置和条件。本公开的方法通常在体外执行。

[0216] 表征

[0217] 本公开的方法包括:表征载体的标识符区,并确定分子是否与分子结合区结合,如本文更详细地描述。

[0218] 可以使用任何合适的检测器系统来进行表征并确定分子是否与分子结合区结合。表征并确定分子是否与分子结合区结合可以例如使用适合于研究膜/孔系统(其中孔被插入膜中)的任何装置来执行。该方法可以使用适合于跨膜孔感测的任何装置来执行。例如,装置可以包括包括水溶液的室和将该室分成两个部分的屏障。屏障可以具有孔,在该孔中形成含有跨膜孔的膜。本文描述了跨膜孔。

[0219] 表征方法可以使用W0 2008/102120、W0 2010/122293或W0 2000/028312中描述的装置来执行。

[0220] 表征方法可以涉及当载体在检测器(例如,纳米孔)内移动时进行一种或多种光学或电学测量。电学测量可以通常通过测量电流来测量流过纳米孔的离子电流。可能的电学测量包括:电流测量、阻抗测量、隧道效应或电子隧道效应测量(Ivanov AP等人,Nano

Lett. 2011年1月12日; 11(1): 279-85) 以及FET测量 (国际申请W0 2005/124888), 例如, 电压FET测量。在一些实施例中, 信号可以是穿过固态纳米孔的电子隧道效应或穿过固态纳米孔的电压FET测量。

[0221] 或者, 可以光学测量通过纳米孔的离子流, 诸如Heron等人: J. Am. Chem. Soc. 第9卷, 131, 第5期, 2009年。使用纳米孔进行光学聚合物测序的方法描述于W0 2016/009180中。

[0222] 装置也可以包括能够施加电势并测量穿过膜和孔的电信号的电路。表征方法可以使用膜片钳或电压钳来执行。表征方法优选涉及电压钳的使用。

[0223] 表征方法可以在基于硅的孔阵列上执行, 其中每个阵列包括128个、256个、512个、1024个、2000个、3000个、4000个、6000个、10000个、12000个、15000个或更多个孔。

[0224] 表征方法可以涉及测量流过孔的电流。方法通常通过在膜和孔上施加电压来执行。使用的电压通常为+2V至-2V, 通常为-400mV至+400mV。所使用的电压优选在具有选自-400mV、-300mV、-200mV、-150mV、-100mV、-50mV、-20mV和0mV的下限和独立地选自+10mV、+20mV、+50mV、+100mV、+150mV、+200mV、+300mV和+400mV的上限的范围内。所使用的电压更优选在100mV至240mV的范围内并且最优选在120mV至220mV的范围内。通过使用增加的施加电势可以增加孔对不同核苷酸的区分。

[0225] 表征方法通常在任何电荷载体的存在下执行, 诸如金属盐 (例如, 碱金属盐)、卤化物盐 (例如, 氯化物盐 (诸如碱金属氯化物盐))。电荷载体可以包括离子液体或有机盐, 例如, 四甲基氯化铵、三甲基苯基氯化铵、苯基三甲基氯化铵或1-乙基-3-甲基氯化咪唑鎓。在上面讨论的示例性装置中, 盐存在于室中的水溶液中。通常使用氯化钾 (KCl)、氯化钠 (NaCl) 或氯化铯 (CsCl)。优选KCl。盐可以是碱土金属盐, 诸如氯化钙 (CaCl₂)。盐浓度可以处于饱和状态。盐浓度可以是3M或更低, 并且通常为0.1至2.5M、0.3至1.9M、0.5至1.8M、0.7至1.7M、0.9至1.6M或1M至1.4M。浓度优选为150mM至1M。表征方法优选使用至少0.3M (诸如至少0.4M、至少0.5M、至少0.6M、至少0.8M、至少1.0M、至少1.5M、至少2.0M、至少2.5M或至少3.0M) 的盐浓度来执行。高盐浓度提供高信噪比并允许在正常电流波动的背景下识别指示结合/无结合的电流。

[0226] 表征方法通常在缓冲液的存在下执行。在上面讨论的示例性装置中, 缓冲液存在于室中的水溶液中。可以使用任何合适的缓冲液。通常, 缓冲液是HEPES。另一合适的缓冲液是Tris-HCl缓冲液。该方法通常在4.0至12.0、4.5至10.0、5.0至9.0、5.5至8.8、6.0至8.7或7.0至8.8或7.5至8.5的pH下执行。所使用的pH优选为约7.5。

[0227] 表征方法可以在0°C至100°C、15°C至95°C、16至90°C、17°C至85°C、18°C至80°C、19°C至70°C或20°C至60°C执行。表征方法通常在室温下执行。表征方法任选地在支持酶功能的温度下进行, 诸如约37°C。

[0228] 载体、载体群体、试剂盒和系统

[0229] 本文也提供载体、载体群体、试剂盒和系统。

[0230] 本公开的载体包括单链前导区、标识符区和对待检测分子具有特异性的分子结合区, 其中马达蛋白在该单链前导区与多核苷酸标识符区之间的位置处与该载体结合。如上文所述的前导序列、标识符区、分子结合区和马达蛋白可以应用于所讨论的载体、载体群体、试剂盒和系统的任何实施例中。载体还可以包括上面讨论的任何附加特征。

[0231] 也提供了本文所述的用于多种分子的载体群体。群体中的不同载体可以包括不同

标识符区和不同分子结合区。例如,群体中的每个载体可以包括唯一的标识符区和唯一的分子结合区。群体中每个载体可以有多个拷贝。换句话说,与载体的分子结合区相关联的标识符区可以不同于与群体中的不同分子结合的每个其他分子结合区相关联的标识符区。在一些实施例中,载体的标识符区不同于群体中的与不同分子结合的其他载体的标识符区。

[0232] 术语“相关联”是指标识符区和一个或多个分子结合区存在于同一载体上,使得标识符区可以用于唯一地识别同一载体上的一个或多个分子结合区。通常,标识符区位于载体中,使得其在分子结合区与检测器相互作用之前在马达蛋白的控制下与检测器相互作用。标识符区可以紧邻分子结合区或者可以通过接头分开。接头在长度上通常为2至约50个碱基,诸如3至约20个碱基,优选约5至约10个碱基。接头可以包括本文所述的任何核苷酸。也可以将一个或多个间隔子定位在标识符区与其相关联的分子结合区之间。间隔子可以在间隔子的一侧或两侧上存在有或不存在接头。

[0233] 也提供了用于检测样品中的多种分子的试剂盒,其包括(i)如本文所述的载体群体以及(ii)马达蛋白。

[0234] 还提供了用于检测样品中的多种分子的系统,其包括(i)如本文所述的载体群体、(ii)马达蛋白以及(iii)跨膜孔。

[0235] 试剂盒或系统可以包括多于一个如本文所定义的载体群体。除了与分子结合区相关联的标识符区之外,群体中的每个载体还可以包括该群体中的所有载体所共有的并且不存在于试剂盒或系统中的任何其他群体的载体中的另外的标识符区,即,该另外的标识符区对于该群体的载体来说是唯一的。因此,试剂盒或系统可以包括两个或更多个载体群体,诸如3个、4个、5个或更多个(例如,10个或更多个、20个或更多个、或者50个或更多个)载体群体,其中每个群体中的载体包括另外的载体,该另外的标识符区对于该群体的载体来说是唯一的。这样的试剂盒或系统可以用于同时分析多个样品。待使用相同检测器检测的每个样品中存在的分子和另外的标识符区可以用于确定给定分子在哪个或哪些样品中存在或不存在。

[0236] 定义

[0237] 当提及单数名词时使用不定冠词或定冠词时,例如,“一(a)”或“一个(an)”、“该(the)”,这包括该名词的复数,除非另有特别说明。当说明书和权利要求书中使用术语“包括”时,其不排除其他要素或步骤。此外,说明书和权利要求书中的术语第一、第二、第三等用于区分相似要素,而不一定用于描述顺序或时间次序。应当理解,如此使用的术语在适当情况下是可互换的,并且本文所描述的本发明的实施例能够以不同于本文描述或说明的其它顺序操作。提供以下术语或定义仅为了帮助理解本发明。除非特别定义,否则本文所使用的所有术语具有与本发明的领域的技术人员所理解的含义相同的含义。针对本领域的定义和术语,从业者特别要参考Sambrook等人,Molecular Cloning:A Laboratory Manual,第4版,Cold Spring Harbor Press,Plainsview,New York(2012);以及Ausubel等人,Current Protocols in Molecular Biology(附录114),John Wiley&Sons,New York(2016)。本文提供的定义不应被解释为具有小于本领域普通技术人员理解的范围。

[0238] 当提及如量、持续时间等可测量的值时,如本文所使用的“约”意味着涵盖与指定值的 $\pm 20\%$ 或 $\pm 10\%$,更优选地 $\pm 5\%$,甚至更优选地 $\pm 1\%$,以及仍更优选 $\pm 0.1\%$ 的偏差,因为此类偏差适合于执行所公开的方法。

[0239] 如本文所使用的“核苷酸序列”、“DNA序列”或“一个或多个核酸分子”是指任何长度的核苷酸的聚合形式,核糖核苷酸或脱氧核糖核苷酸。该术语仅指分子的一级结构。因此,该术语包括双链和单链DNA和RNA。如本文所使用的术语“核酸”是单链或双链共价连接的核苷酸序列,其中每个核苷酸上的3'和5'端通过磷酸二酯键连接。多核苷酸可以由脱氧核糖核苷酸碱基或核糖核苷酸碱基组成。核酸可以在体外合成制造或从天然来源分离。核酸还可以包括经修饰的DNA或RNA(例如,已经甲基化的DNA或RNA)或已经进行翻译后修饰的RNA(例如,用7-甲基鸟苷进行5'-加帽、3'-加工(诸如切割以及聚腺苷酸化已经剪接))。核酸还可以包括合成核酸(XNA),诸如己糖醇核酸(HNA)、环己烯核酸(CeNA)、苏糖核酸(TNA)、甘油核酸(GNA)、锁核酸(LNA)和肽核酸(PNA)。核酸(本文也称为“多核苷酸”)的大小对于双链多核苷酸通常表示为碱基对(bp)的数量或者在单链多核苷酸的情况下表示为核苷酸的数量(nt)。一千bp或nt等于千碱基(kb)。在长度上小于约40个核苷酸的多核苷酸通常被称为“寡核苷酸”并且可以包括用于诸如经由聚合酶链式反应(PCR)操作DNA的引物。

[0240] 在本公开内容的上下文中,术语“氨基酸”以其最广泛的含义使用,并且意在包括含有胺(NH₂)和羧基(COOH)官能团以及对每种氨基酸具有特异性的侧链(例如,R基团)的有机化合物。在一些实施例中,氨基酸是指天然存在的L α -氨基酸或残基。本文使用天然存在的氨基酸的常用一字母和三字母缩写:A=Ala;C=Cys;D=Asp;E=Glu;F=Phe;G=Gly;H=His;I=Ile;K=Lys;L=Leu;M=Met;N=Asn;P=Pro;Q=Gln;R=Arg;S=Ser;T=Thr;V=Val;W=Trp;以及Y=Tyr(Lehninger, A.L., (1975) Biochemistry, 第2版, 第71-92页, Worth Publishers, New York)。通用术语“氨基酸”进一步包括D-氨基酸、反向翻转氨基酸以及经化学修饰的氨基酸(诸如氨基酸类似物)、通常不掺入蛋白质中的天然存在的氨基酸(诸如正亮氨酸)以及具有本领域已知的特性作为氨基酸(诸如 β -氨基酸)特征的化学合成氨基酸。例如,苯丙氨酸或脯氨酸的类似物或模拟物(其允许与天然苯丙氨酸或脯氨酸相同的肽化合物构象限制)包括在氨基酸的定义内。此类类似物和模拟物在本文中称为相应氨基酸的“功能等同物”。Roberts和Vellaccio列出了氨基酸的其他实例,The Peptides: Analysis, Synthesis, Biology, Gross和Meiehofer编辑,第5卷,第341页,Academic Press, Inc., N.Y. 1983,其通过引用的方式并入本文。

[0241] 术语“多肽”和“肽”在本文中可互换使用,指氨基酸残基的聚合物及其变体和合成类似物。因此,这些术语适用于其中一个或多个氨基酸残基是合成的非天然存在的氨基酸(诸如相应的天然存在的氨基酸的化学类似物)的氨基酸聚合物,以及天然存在的氨基酸聚合物。多肽也可以经历成熟或翻译后修饰过程,其可以包括但不限于:糖基化、蛋白水解切割、脂化、信号肽切割、前肽切割、磷酸化等。肽可以使用重组技术来制备,例如通过重组或合成多核苷酸的表达。重组产生的肽通常基本上不含培养基,例如,培养基占蛋白质制剂体积的小于约20%,更优选小于约10%并且最优选小于约5%。

[0242] 术语“蛋白质”用于描述具有二级或三级结构的折叠多肽。蛋白质可以由单一多肽组成或可以包括组装形成多聚体的多个多肽。多聚体可以是同寡聚体或杂寡聚体。蛋白质可以是天然存在的或野生型蛋白质或者经修饰的或非天然存在的蛋白质。例如,蛋白质与野生型蛋白质的不同之处在于添加、取代或缺失一个或多个氨基酸。

[0243] 蛋白质的“变体”涵盖肽、寡肽、多肽、蛋白质和酶,其相对于所讨论的未修饰或野生型蛋白质具有氨基酸取代、缺失和/或插入,并且具有与其所源自的未经修饰的蛋白质相

似的生物学和功能活性。如本文所使用的术语“氨基酸同一性”是指序列在比较窗口内逐个氨基酸基础上相同的程度。因此,“序列同一性百分比”通过以下来计算:在比较窗口内比较两个经过最佳比对的序列,测定相同的氨基酸残基(例如,Ala、Pro、Ser、Thr、Gly、Val、Leu、Ile、Phe、Tyr、Trp、Lys、Arg、His、Asp、Glu、Asn、Gln、Cys和Met)出现在这两个序列中的位置处的数量以产生匹配位置的数量;用匹配位置的数量除以比较窗口中的位置的总数(即,窗口大小);以及将结果乘以100以产生序列同一性百分比。

[0244] 对于本发明的所有方面和实施例,“变体”与相应野生型蛋白质的氨基酸序列具有至少50%、60%、70%、80%、90%、95%或99%的完整序列同一性。序列同一性也可以与全长多核苷酸或多肽的片段或部分相同。因此,序列可能与全长参考序列仅具有50%的总体序列同一性,但特定区、结构域或亚基的序列可以与参考序列共享80%、90%或高达99%的序列同一性。

[0245] 术语“野生型”是指从天然存在的来源分离的基因或基因产物。野生型基因是在群体中最常见的基因,并且因此被任意设计为该基因的“正常”或“野生型”形式。相反,术语“经修饰的”、“突变体”或“变体”是指与野生型基因或基因产物相比时表现出序列修饰(例如,取代、截短或插入)、翻译后修饰和/或功能特性(例如,改变的特性)的基因或基因产物。值得注意的是,可以分离天然存在的突变体;通过与野生型基因或基因产物相比它们具有改变的特性这一事实来识别这些基因。引入或取代天然存在的氨基酸的方法是本领域众所周知的。例如,可以通过在编码突变单体的多核苷酸中的相关位置处用精氨酸密码子(CGT)替代甲硫氨酸密码子(ATG)来用精氨酸(R)取代甲硫氨酸(M)。引入或取代非天然存在的氨基酸的方法也是本领域众所周知的。例如,可以通过将合成的氨酰基-tRNA包括在用于表达突变单体的IVTT系统中来引入非天然存在的氨基酸。或者,它们可以通过在大肠杆菌中表达突变体单体来引入,该大肠杆菌在那些特定氨基酸的合成(即非天然存在)类似物的存在下对特定氨基酸而言是营养缺陷型的。如果突变体单体是使用部分肽合成产生的,则它们也可以通过裸连接产生。保守取代用具有相似化学结构、相似化学性质或相似侧链体积的其他氨基酸替代氨基酸。引入的氨基酸可以与它们取代的氨基酸具有相似的极性、亲水性、疏水性、碱性、酸性、中性或电荷。或者,保守取代可以引入另一芳香族或脂肪族氨基酸来代替预先存在的芳香族或脂肪族氨基酸。保守氨基酸改变是本领域众所周知的并且可以根据如下表1中定义的20种主要氨基酸的特性来选择。当氨基酸具有相似的极性时,这也可以参考表2中氨基酸侧链的亲水等级来确定。

[0246] 表1-氨基酸的化学性质

[0247]	Ala	脂肪族、疏水性、中性	Met	疏水性、中性
	Cys	极性、疏水性、中性	Asn	极性、亲水性、中性
	Asp	极性、亲水性、带电荷(-)	Pro	疏水性、中性
	Glu	极性、亲水性、带电荷(-)	Gln	极性、亲水性、中性
	Phe	芳香族、疏水性、中性	Arg	极性、亲水性、带电荷(+)
	Gly	脂肪族、中性	Ser	极性、亲水性、中性
	His	芳香族、极性、亲水性、带电荷(+)	Thr	极性、亲水性、中性
	Ile	脂肪族、疏水性、中性	Val	脂肪族、疏水性、中性
	Lys	极性、亲水性、带电荷(+)	Trp	芳香族、疏水性、中性
	Leu	脂肪族、疏水性、中性	Tyr	芳香族、极性、疏水性

[0248] 表2-亲水等级

侧链		亲水性
[0249]	Ile	4.5
	Val	4.2
	Leu	3.8
	Phe	2.8
	Cys	2.5
	Met	1.9
	Ala	1.8
	Gly	-0.4
	Thr	-0.7
	Ser	-0.8
[0250]	Trp	-0.9
	Tyr	-1.3
	Pro	-1.6
	His	-3.2
	Glu	-3.5
	Gln	-3.5
	Asp	-3.5
	Asn	-3.5
	Lys	-3.9
	Arg	-4.5

[0251] 如本文更详细描述,突变体或经修饰的蛋白质、单体或肽可以以任何方式并在任何位点进行化学修饰。突变体或经修饰的单体或肽优选通过将分子附接到一个或多个半胱氨酸(半胱氨酸连接)、将分子附接到一个或多个赖氨酸、将分子附接到一个或多个非天然氨基酸、表位的酶修饰或末端的修饰来进行化学修饰。用于执行此类修饰的合适的方法是本领域众所周知的。经修饰的蛋白质、单体或肽的突变体可以通过附接任何分子来进行化学修饰。例如,经修饰的蛋白质、单体或肽的突变体可以通过附接染料或荧光团来进行化学修饰。

[0252] 本发明是关于特定实施例并参考某些附图来进行描述的,但本发明不限于此而仅受权利要求书的限制。权利要求中的任何附图标记不应被解释为限制范围。当然,应当理解,根据本发明的任何特定实施例,不必实现所有方面或优点。因此,例如,本领域技术人员应当认识到,本发明可以以实现或优化这里所教导的一个优点或一组优点而不必实现这里所教导或建议的其他优点的方式来实施或执行。

[0253] 当结合附图阅读时,通过参考以下详细描述可以最好地理解本发明的组织和操作方法及其特征和优点。本发明的方面和优点将从下文中所描述的实施例中显而易见且参考下文中所描述的一个或多个实施例予以阐明。在整个说明书中对“一个实施例”或“实施例”的引用意味着结合该实施例描述的特定特征、结构或特性包括在本发明的至少一个实施例中。因此,在本说明书中各处出现的短语“在一个实施例中”或“在实施例中”不一定都指同一实施例,而是可以指同一实施例。类似地,应当理解,在本发明的示例性实施例的描述中,出于简化公开内容和帮助理解一个或多个各种发明方面的目的,本发明的各种特征有时被组合在单个实施例、附图或其描述中。然而,本公开的方法不应被解释为反映所要求保护的发明要求比每个权利要求中明确列举的特征需要更多特征的意图。相反,如所附权利要求所反映,创造性方面在于少于单个前述公开的实施例的所有特征。

[0254] 应当理解的是,除非上下文另有说明,本公开的“实施例”可以具体地组合在一起。所有公开的实施例的具体组合(除非上下文另有暗示)是所要求保护的发明的进一步公开的实施例。

[0255] 另外,如在本说明书和所附权利要求书中所使用的,除非内容另外明确指出,否则单数形式“一个/种(a/an)”和“所述(the)”包括复数指代物。因此,例如,提及“多核苷酸”包括两个或更多个多核苷酸,提及“马达蛋白”包括两个或更多个这样的蛋白质,提及“解旋酶”包括两个或更多个解旋酶,提及“单体”是指对于两个或更多个单体,提及“孔”包括两个或更多个孔等。

[0256] 本文引用的所有出版物、专利和专利申请,无论上文还是下文,均通过引用的方式整体并入本文。

[0257] 应当理解,虽然本文已经讨论了根据本发明的方法的具体实施例、具体构造以及材料和/或分子,但是在不脱离本发明的范围和精神的情况下,可以进行形式和细节上的各种改变或修改。提供以下实例是为了更好地说明特定实施例,并且它们不应被视为限制本申请。本申请仅由以下权利要求限制。

[0258] 实例

[0259] 在下面的实例中,载体以两部分合成,并将各部分连接在一起。也设想了载体可以整体合成为单个单元而不需要连接。

[0260] 实例1

[0261] 此实例描述了迄今为止设计的一些载体。提供的序列不包括前导区。前导区可以通过使用连接C链(例如,CCCAGCGGAAGTAGGA)添加到所提供的序列,其也包括与包括马达蛋白的前导区和停滞区的多核苷酸的3'端互补的区域,如图2A所示。

[0262] 所提供的载体序列可以与包括蛋白质(诸如凝血酶和SARS-CoV-2刺突蛋白)、神经递质(诸如血清素和多巴胺)和miRNA的分子结合。每个载体的标识符区不同于每个其他载体的标识符区。可以预见,不同载体上的与相同分子结合的不同分子结合区可以使用相同的条形码。

[0263] 序列包括5'连接链、标识符序列或条形码(下划线)、任选的包括iSpC3或iSp18间隔子的间隔区以及分子结合区(斜体)。本实例中的分子结合区是适体或miRNA的互补序列。

[0264] 凝血酶

[0265] /5Phos/CCTAGTTCCGCTGGGGAGCTAGCATAACGTGTCTAACACTGCACAGATGAT/iSpC3/iSpC3/GGTTGGTGTGGTTGG

[0266] /5Phos/CCTAGTTCCGCTGGGATATAGGCATAAAGTAAAATCGTACCAACTCAATC/iSpC3/iSpC3/GGTTGGTGTGGTTGG

[0267] /5Phos/CCTAGTTCCGCTGGGCTCGGCGTTGTGTGTCAAATGGCGTAGATCTGGAT/iSpC3/iSpC3/GGTTGGTGTGGTTGG

[0268] /5Phos/CCTAGTTCCGCTGGGCACAGCCCCATGTAACCCAT/iSp18/iSp18/AGTCCGTGGTAGGGCAGGTTGGGGTGACT

[0269] 血清素

[0270] /5Phos/CCTAGTTCCGCTGGGACTAGATAAAAGGAAGGGAGCACAGTAACGTCGTT/iSpC3/iSpC3/CGACTGGTAGGCAGATAGGGGAAGCTGATTGATG

[0271] 多巴胺

[0272] /5Phos/CCTAGTTCCGCTGGGTCTCCCGTATCCGTGGCTAAACGCCTTCAATCTTA/iSpC3/iSpC3/GGATATTGCGCGATTCCGGTCGGCAGCTTAGGAAGTGCGGTGTC

[0273] SARS-CoV-2, S蛋白质

[0274] /5Phos/CCTAGTTCCGCTGGGACCTTTTCTAGGATGGAACATTCTA/iSpC3/iSpC3/CACGCATAACGTCTTGCGGGGCGGCGGGTTGAGAGGATGTCGGGTGGTTATGCGTG

[0275] miRNA(条形码11-20)

[0276] 衔接子_条形码11_c-has-miR-497-5p

[0277] /5Phos/CCTAGTTCCGCTGGGTGCTACTCTTCCTCATAAGCAGTCCGGTGTATCGAT/iSpC3/iSpC3/ACAAACCACAGTGTGCTGCTG

[0278] 衔接子_条形码12_c-has-miR-27b-5p

[0279] /5Phos/CCTAGTTCCGCTGGGATCGCTACGCCTTCGGCTCGTAATCATAGTCGAGT/iSpC3/iSpC3/GTTCACCAATCAGCTAAGCTCT

[0280] 衔接子_条形码13_c-has-miR-21-5p

[0281] /5Phos/CCTAGTTCCGCTGGGAGCTCAGAGCAGGTCACCTCAAGATACGAGCTGCGT/iSpC3/iSpC3/TCAACATCAGTCTGATAAGCTA

- [0282] 衔接子_条形码14_c-has-miR-221-5p
/5Phos/CCTAGTTCCGCTGGGGTAAAGTCTGCATCAGCGCGCGGCTGTGCGAGGATA/iSp
- [0283] pC3/iSpC3/AAATCTACATTGTATGCCAGGT
- [0284] 衔接子_条形码15_c-has-miR-30d-5p
/5Phos/CCTAGTTCCGCTGGGGTACGACAGTACGCTAGCAAGGATAGACACTACGA/iS
- [0285] pC3/iSpC3/CTTCCAGTCGGGGATGTTTACA
- [0286] 衔接子_条形码16_c-has-miR-30c-5p
/5Phos/CCTAGTTCCGCTGGGTACTGAACACAAGTTCGTCGTCGAGCAATCACAAT/iSp
- [0287] C3/iSpC3/GCTGAGAGTGTAGGATGTTTACA
- [0288] 衔接子_条形码17_c-has-miR-133a-5p
/5Phos/CCTAGTTCCGCTGGGAGTCTACCATTACTTGGATCGGATTAGCCTCACTC/iSp
- [0289] C3/iSpC3/ATTTGGTTCCATTTTACCAGCT
- [0290] 衔接子_条形码18_c-has-miR-208a-5p
/5Phos/CCTAGTTCCGCTGGGTGCACGAGTGC GTGTCAACCGTCCAGATGCTCGTG/iS
- [0291] pC3/iSpC3/GTATAACCCGGGGCCAAAAGCTC
- [0292] 衔接子_条形码19_c-has-miR-181b-5p
/5Phos/CCTAGTTCCGCTGGGGTACTGTGCGCAGTTGTCTCGGCGGAGTTGAGACTGA/iS
- [0293] pC3/iSpC3/ACCCACCGACAGCAATGAATGTT
- [0294] 衔接子_条形码20_c-has-miR-29a-3p
/5Phos/CCTAGTTCCGCTGGGGATCATGGTAGTCTTCAAGATCGAGTATGTCTGTG/iSp
- [0295] C3/iSpC3/TAACCGATTTCAAGATGGTGCTA
- [0296] 不含C3间隔子的miRNA(条形码21-22)
- [0297] 衔接子_条形码21_c-has-miR-30c-5p
/5Phos/CCTAGTTCCGCTGGGTACTGAACACAAGTTCGTCGTCGAGCAATCACAATGC
- [0298] TGAGAGTGTAGGATGTTTACA
- [0299] 衔接子_条形码22_c-has-miR-29a-3p
/5Phos/CCTAGTTCCGCTGGGGCGTGAAGAGAGTTTCATAATACGTCCAGCCGCATGTAA
- [0300] CCGATTTCAAGATGGTGCTA

[0301] 实例2

[0302] 该实例描述了通过对各个载体的唯一条形码进行排序和去重来识别各个载体。该实例也描述了通过停滞分析来识别与分析物结合的载体。

[0303] 方法

[0304] 将如实例1中所述的条形码(总浓度30nM,每个浓度相同)与连接c链以1:3的摩尔比在无核酸酶水中孵育1h。通过在4℃离心3分钟、然后在室温下孵育1小时来启动杂交。将所得混合物与10nM适体混合,并通过加入等体积的TA连接酶预混液(New England Biolabs)来执行连接,4℃离心3分钟以充分混合不同组分,同时在低温下保存连接酶。室温孵育20分钟之后,加入总体积1.4倍的Ampure XP珠(Beckmann Coutler)吸收核酸供进一步纯化。用短片段缓冲液(Oxford Nanopore Technologies)洗涤珠子,以选择性去除过量的

未杂交的连接c链和未连接的条形码。纯化之后,用无核酸酶水洗涤珠子,以洗掉含有连接的马达蛋白-条形码复合物的纯化核酸。测序实验的最终溶液由洗脱液、测序缓冲液、系链(100nM)、无核酸酶水组成,并与一定浓度的靶向分析物一起孵育至少30分钟。

[0305] 所有实验均使用纳米孔App进行分析,该App是预先存在的定制编写的MATLAB代码(由Joshua Edel在2006年至2021年期间开发)。FAST5测序文件上传到应用程序中,并在其中进行进一步处理。应用程序允许单独或批量上传和分析MinION的所有512个通道。

[0306] 使用阈值算法检测易位事件。首先,根据实验,将线性基线设置在-0.16与-2nA之间。使用1.8的步长偏移来定义事件的开始(绿线)。如果事件超过标准差50(黑线)阈值,算法就会检测到它们。通过过滤0.025s与10s之间的事件来进一步分析这些事件。该时间范围内的所有事件都被定义为事件。另外,该算法用于定位链设计内的C3或用作比对标记物的其他间隔子元件。以这种方式,间隔子的存在对于检测载体上存在或不存在分子来说不是必需的。

[0307] 使用GUPPY或/和MinKNOW软件(Oxford Nanopore Technologies)对上一步中检测到的事件进行测序和碱基识别。

[0308] 在对原始读取信号进行碱基识别之后,将测序事件与参考序列(例如,条形码序列1-10)进行比对。如果满足以下几点,则算法会将条形码分配给最高比对分数的事件:

[0309] 1) 比对分数需要高于50;

[0310] 2) 最大比对分数与第二高分数的差异需要至少为5;

[0311] 3) 最大比对分数与其他条形码的平均比对分数的差异需要至少为10;

[0312] 4) p值必须 ≤ 0.0001 。

[0313] 仅当满足所有标准时,最大比对才被分类为条形码。如果不是,则该事件被去除并且不被考虑供进一步分析。通过计算与群体的其余部分相比的最高比对分数的p值来去除误报。

[0314] 对于停滞分析,定义了C3部分以及事件的开始和结束。这对于计算所有测序读数的平均易位时间非常重要。如果事件的易位时间与平均值相比显著较长(通常使用移动标准差100个bin),则该事件被分类为停滞。此外,每次停滞需要大于20个bin才能被视为停滞。

[0315] 通过对各个载体的唯一条形码进行测序和去重来识别各个载体。

[0316] 所有序列均按上文所述进行碱基识别,并与参考序列(即条形码序列)进行比对。最大比对分数用于对条形码进行分类。如果最大比对分数和第二高分数太接近,则该事件不会被分类。此外,p值用于对事件进行分类并去除误报分类。利用该方法,识别各个条形码的准确性达到99.95%,其中使用所有记录事件的86%并对其进行分类,这意味着与本领域中的先前方法相比,浪费的数据相对较少(图4)。

[0317] 条形码序列具有大于90%的准确性。观察到仅0.0001%的假阳性率(图5(a))。该方法对错误条形码分类的偏好非常低,如图5(b)的混淆矩阵所示。改进的算法和检测技术将继续改进条形码分类。改进条形码分类的另一方法是将条形码的重复包括在载体中作为校对机制。该实例示出了可以成功区分复杂混合物中的10个条形码载体,从而实现高度多重检测。

[0318] 通过停滞分析识别与分析物结合的载体。

[0319] 在此实例中,使用停滞分析来确定靶标是否已经与条形码链结合。如果靶分析物未与条形码链结合,则当前信号不表明停滞(在停留时间、电流振幅中)。结合的分析物(此处给出了互补miRNA的实例)使载体停滞,这产生唯一的电流分布(参见图7)。停滞可能是由于以下:例如,(1)双链核苷酸结构的解链(例如,miRNA、DNA检测);(2)G-四链体或茎环适体结构的解开(例如,用于检测蛋白质、神经递质和小结合分子分析物);或(3)结合抗体-抗原的条带化。靶分子的相对浓度可以通过关联停滞载体与非停滞载体的分数来确定(参见图8)。样品中载体的绝对浓度可以通过测量各个单个分子检测事件之间的时间(事件间时间)来确定,如纳米孔文献中广泛描述。

[0320] 使用本文所述的方法确定分子浓度的另一方法是确定分子对与分子结合的载体的检测的浓度依赖性。例如,可以按照与图9所示类似的方式绘制标准曲线。

[0321] 在另一实验中,研究了对凝血酶与15-mer凝血酶结合适体之间结合的检测。与凝血酶结合后,“波形”事件显示出更长的停留时间,与400nM凝血酶结合后发现,在电流翻转方面,停滞显著增加,对应于G-四链体和适体-蛋白质相互作用的解旋。通过将凝血酶浓度从0nM增加到400nM来进一步验证凝血酶和15-mer凝血酶结合适体之间的结合的浓度依赖性,并且由于更多的波形事件与更长的停留时间,观察到停滞的增加(参见图10和图11)。

[0322] 在另一实验中,研究了使用茎环适体检测结合血清素。与血清素结合时停滞的增加归因于与血清素结合时适体从环到G-四联体的结构重组,导致环、G-四联体和适体-血清素相互作用的解旋(参见图12)。通过将血清素浓度增加到0nM与40nM之间来进一步验证血清素与适体之间结合的浓度依赖性。观察到停滞的增加表现为更多的波形事件和更长的停留时间(参见图13)。

[0323] 基于平均延迟时间与浓度也观察了血清素与适体之间的结合的浓度依赖性,如下表3所示。

[0324] 表3

浓度 (nM)	平均延迟时间 (ms)
0	95
2.5	97
5	103
10	112
20	135
40	144
60	187
80	388

[0326] 在另一实验中,研究了使用阶梯环适体检测结合乙酰胆碱。使用的条形码(下划线)-间隔子-适体(斜体)是:

[0327] TCGATACAATACA/间隔子/ATCCGTCACACCTGCTCTAGGGGATCAAAGCTATGCGACCA
TGCGAGTGGATACTGGTGTGGCTCCCGTAT

[0328] 与乙酰胆碱结合后可以看到明显的停滞事件。通过在0nM与40nM之间增加乙酰胆碱浓度来观察乙酰胆碱与适体之间的结合对延迟(停滞)百分比的浓度依赖性(图14)。

[0329] 实例3

[0330] 可以使用其他方法来识别与分析物结合的载体。

[0331] 在一些实施例中,没有必要执行特定的停滞分析。碱基识别软件程序GUPPY和MinKNOW(Oxford Nanopore Technologies)可以用于直接分析载体在孔内移动时产生的数据。

[0332] 单独地,并且如上所述,另一选择是使用酶消化来去除/消化具有未结合的分子结合区的载体的任何部分。这可以通过使用a) 核酸内切酶或b) 核酸外切酶来实现,这些酶靶向未与分析物靶标结合的载体中的分子结合区(参见图3)。因此,在消化之后,将基于条形码序列之后当前信号的存在或缺乏来区分与靶标结合的载体和未与靶标结合的载体。或者,如果分子结合区是多肽,则可以使用消化具有未结合的分子结合区的载体部分的蛋白酶,条件是,与包括与分子结合区结合的分子的载体相比,当包括未结合且酶促消化的分子的载体在孔内穿过时可以观察到信号的差异。

[0333] 实例4

[0334] 载体组装协议

[0335] 将条形码(总浓度30nM,每个浓度相同)与连接c链以1:3的摩尔比孵育。通过在4°C离心1分钟,然后在室温下孵育1h来启动杂交。将所得混合物与10nM接头混合,并通过添加等体积的TA连接酶主混合物(New England Biolabs)来执行连接。将样品在4°C离心1分钟,以充分混合不同组分,同时在低温下保存连接酶。室温孵育20分钟之后,加入总体积1.4倍的Ampure XP珠(Beckmann Coulter)吸收核酸供进一步纯化。用短片段缓冲液(Oxford Nanopore Technologies)洗涤珠子两次,以选择性去除过量的未杂交的连接c链和未连接的条形码。纯化之后,用无核酸酶水洗涤珠子,以洗掉含有连接的马达蛋白-条形码复合物的纯化核酸。测序实验的最终溶液由洗脱液、测序缓冲液、系链(100nM)、无核酸酶水组成,并与一定浓度的靶向分析物一起孵育至少30分钟。对于miRNA实验,使用的浓度为0.05nM、0.1nM、0.25nM、0.5nM、1nM、2.5nM、5nM、10nM、25nM和50nM。对于蛋白质实验,使用的浓度为10pg/mL、50pg/mL、100pg/mL、500pg/mL、1ng/mL、30ng/mL。

[0336] 实验运行协议

[0337] 所有实验均使用研究和/或客户脚本在37°C运行30分钟。

[0338] 数据分析工作流程

[0339] 所有实验均使用纳米孔App进行分析,这是一种预先存在的定制编写的MATLAB代码。将纳米孔测序设备(MinION设备;Oxford Nanopore Technologies)输出的FAST5测序文件上传到应用程序中,在其中进行进一步处理。该应用程序允许单独或批量上传和分析MinION的所有512个通道。

[0340] 事件检测

[0341] 使用阈值算法检测易位事件。首先,根据实验,将线性基线设置在-0.16nA与-2nA之间。使用1.8的步长偏移来定义事件的开始。如果事件超过标准差30的阈值,算法就会检测到它们。通过过滤大于0.1s的事件来进一步分析这些事件。该时间帧内的所有事件都被定义为事件。

[0342] 测序和碱基识别

[0343] 使用GUPPY和/或MinKNOW软件(Oxford Nanopore Technologies)对上一步中检测到的事件进行测序和碱基识别。

[0344] 比对

[0345] 在对原始读取信号进行碱基识别之后,将测序事件与参考序列(这种情况下,条形码序列1-40)进行比对。如果满足以下几点,则算法为最高比对分数的事件分配一个条形码:

[0346] 1) 序列必须以‘GGG’开头;

[0347] 2) 至少需要比对15个碱基;

[0348] 3) 前10个碱基中仅1个不匹配;

[0349] 4) 所有比对的碱基中只有1个不匹配。

[0350] 仅当满足所有标准时,最大比对才被分类为条形码。如果不是,则该事件被去除并且不被考虑供进一步分析。

[0351] 停滞

[0352] 对于停滞分析,定义了C3部分以及事件的开始和结束。这用于计算所有测序读数的平均易位时间。如果事件的易位时间与平均值相比显著较长(通常使用移动标准差75个bin),则该事件被分类为停滞。此外,每次停滞需要大于10个bin才能被视为停滞。

[0353] 结果:40个条形码的热图(图19)

[0354] 图19呈现了一个混淆矩阵,表明对错误条形码分类的偏好非常低。所有40个测试的条形码均被识别,准确性>95%。

[0355] 结果:检测多种miRNA

[0356] 本文描述的多重条形码测序方法能够检测40种不同miRNA。结果呈现在图20中。

[0357] 结果:未知miRNA浓度的定量

[0358] 本文所述的方法使用对已知miRNA浓度的多个不同样品的盲测,使得能够准确预测miRNA浓度。结果呈现在图21中。

[0359] 结果:cTnI检测

[0360] 与蛋白质心肌钙蛋白I(cTnI)的检测相关的数据如图22所示。所示肌钙蛋白适体序列如下:

[0361] AGTCTCCGCTGTCCTCCCGATGCACTTGACGTATGTCTCACTTTCTTTTCATTGACATGGGATGACGCCGTGACTG

[0362] 实例4的附录

[0363] 载体链序列-miRNA(图17)

[0364] 衔接子_条形码12_c-hsa-miR-27b-5p/5Phos/CCTAGTTCCGCTGGGATCGCTACGCCTTCGGCTCGTAATCATAGTCGAGT/iSpC3//iSpC3/GTTCACCAATCAGCTAAGCTCT

[0365] 衔接子_条形码13_c-hsa-miR-21-5p/5Phos/CCTAGTTCCGCTGGGAGCTCAGAGCAGGTC ACTCAAGATACGAGCTGCGT/iSpC3//iSpC3/TCAACATCAGTCTGATAAGCTA

[0366] 衔接子_条形码14_c-hsa-miR-221-5p/5Phos/CCTAGTTCCGCTGGGGTAAGTCTGCATCA GCGCGCGGCTGTGCGAGGATA/iSpC3//iSpC3/AAATCTACATTGTATGCCAGGT

[0367] 衔接子_条形码15_c-hsa-miR-30d-5p/5Phos/CCTAGTTCCGCTGGGCTACGACAGTACGC TAGCAAGGATAGACACTACGA/iSpC3//iSpC3/CTTCCAGTCGGGGATGTTTACA

[0368] 衔接子_条形码16_c-hsa-miR-30c-5p/5Phos/CCTAGTTCCGCTGGGTACTGAACACAAGT TCGTCGTCGAGCAATCACAAT/iSpC3//iSpC3/GCTGAGAGTGTAGGATGTTTACA

[0369] 衔接子_条形码19_c-hsa-miR-181b-5p/5Phos/CCTAGTTCGCTGGGCTAGTGCGCAGTT
GTCTCGGCGGAGTTGAGACTGA/iSpC3//iSpC3/ACCCACCGACAGCAATGAATGTT

[0370] 衔接子_条形码20_c-hsa-miR-29a-3p/5Phos/CCTAGTTCGCTGGGGATCATGGTAGTCT
TCAAGATCGAGTATGTCTGTC/iSpC3//iSpC3/TAACCGATTTTCAGATGGTGCTA

[0371] 衔接子_条形码22_c-hsa-miR-210-5p/5Phos/CCTAGTTCGCTGGGGTTCACATCAAGGT
CATACCGCGAGTTCTATTTTA/iSpC3//iSpC3/CAGTGTGCGGTGGGCAGGGGCT

[0372] 衔接子_条形码24_c-hsa-miR-126a-5p/5Phos/CCTAGTTCGCTGGGGCTTGGGGGATAG
ATGTGCCCCGCGCATCGGACCT/iSpC3//iSpC3/CGCGTACCAAAAGTAATAATG

[0373] 衔接子_条形码25_c-hsa-miR-1306-5p/5Phos/CCTAGTTCGCTGGGATGACACACGTTT
TCGATAGGGACGCCGACTTTAA/iSpC3//iSpC3/TGGACGTTTGCAGGGGAGGTGG

[0374] 衔接子_条形码26_c-hsa-miR-126-5p/5Phos/CCTAGTTCGCTGGGTGATAATAAGACCT
GACAGACAATAGGGGAGAACTC/iSpC3//iSpC3/CGCGTACCAAAAGTAATAATG

[0375] 衔接子_条形码27_c-hsa-miR-1254/5Phos/CCTAGTTCGCTGGGTAGTAATCAAGTCTG
ATCGTAATAGCTAAGTCAT/iSpC3//iSpC3/ACTGCAGGCTCCAGCTTCCAGGCT

[0376] 衔接子_条形码29_c-hsa-miR-30e-5p/5Phos/CCTAGTTCGCTGGGCCTGTTCTAATTGC
GCGGAGAGGCGAGATGTTTCT/iSpC3//iSpC3/CTTCCAGTCAAGGATGTTTACA

[0377] 衔接子_条形码30_c-hsa-miR-106a-5p/5Phos/CCTAGTTCGCTGGGTGGGATACTGACG
TGCCGGAATACCCAGACGTGCC/iSpC3//iSpC3/CTACCTGCACTGTAAGCACTTTT

[0378] 衔接子_条形码31_c-hsa-miR-199a-3p/5Phos/CCTAGTTCGCTGGGTGATCGGCACCTA
AACGCATTAGCCCTGCAATACG/iSpC3//iSpC3/TAACCAATGTGCAGACTACTGT

[0379] 衔接子_条形码32_c-hsa-miR-652-3p/5Phos/CCTAGTTCGCTGGGAGCTGCTCGGAAGC
CATAAGGTACTTTAATTTGGG/iSpC3//iSpC3/CACAACCCTAGTGGCGCCATT

[0380] 衔接子_条形码33_c-hsa-miR-26b-5p/5Phos/CCTAGTTCGCTGGG

[0381] C A C G G A T T T C T A T A T T G C T C A A C C A G G C A G C G C A A / i S p C 3 / / i S p C 3 /
ACCTATCCTGAATTACTTGAA

[0382] 衔接子_条形码34_c-hsa-miR-145-5p/5Phos/CCTAGTTCGCTGGGATTGTGCGCTTTTC
TCCATGCGTCTTAGACATCTC/iSpC3//iSpC3/AGGGATTCTGGGAAAACTGGAC

[0383] 衔接子_条形码35_c-hsa-miR-92a-3p/5Phos/CCTAGTTCGCTGGGGTGAAATCCCCGTC
TAGGTTATGGCTGGGGGATT/iSpC3//iSpC3/ACAGGCCGGGACAAGTGCAATA

[0384] 衔接子_条形码36_c-hsa-miR-146a-5p/5Phos/CCTAGTTCGCTGGGCGTAACTTATCA
CGACACAATGAACAAGCCTGCA/iSpC3//iSpC3/AACCCATGGAATTCAGTTCTCA

[0385] 衔接子_条形码37_c-hsa-miR-423-5p/5Phos/CCTAGTTCGCTGGGGCAGTGTCCGAGC
GTCCTCAATCATGAGCGATT/iSpC3//iSpC3/AAAGTCTCGCTCTCTGCCCCCTCA

[0386] 衔接子_条形码39_c-hsa-miR-27b-3p/5Phos/CCTAGTTCGCTGGGCGTCCAGACTTAAT
GTCTGCTCACTGACATCGCGA/iSpC3//iSpC3/GCAGAACTTAGCCACTGTGAA

[0387] 衔接子_条形码40_c-hsa-miR-1-3p/5Phos/CCTAGTTCGCTGGGATTAGCGGAACCAAAC
CCAGGAAGGCTTGAAGGCG/iSpC3//iSpC3/ATACATACTTCTTTACATTCCA

[0388] 衔接子_条形码42_c-hsa-miR-18a-5p/5Phos/CCTAGTTCGCTGGGAACCTTAGGGGCT
CGAATCTTTGAGACGACTAGG/iSpC3//iSpC3/CTATCTGCACTAGATGCACCTTA

- [0389] 衔接子_条形码43_c-hsa-miR-18b-5p/5Phos/CCTAGTTCCGCTGGGTAATTACTGCCCCA
CCATGACATTTTAATAGCAGT/iSpC3//iSpC3/CTAACTGCACTAGATGCACCTTA
- [0390] 衔接子_条形码45_c-hsa-miR-301a-5p/5Phos/CCTAGTTCCGCTGGGGACCTTGAGACAG
AACTTATCAATGTACAAGTAA/iSpC3//iSpC3/AGTAGTGCAATAAAGTCAGAGC
- [0391] 衔接子_条形码46_c-hsa-let7c-5p/5Phos/CCTAGTTCCGCTGGGCGAAGGATTGGCCCCC
CGATTACCACCGCCGTGAG/iSpC3//iSpC3/AACCATACAACCTACTACCTCA
- [0392] 衔接子_条形码47_c-hsa-miR-125a-5p/5Phos/CCTAGTTCCGCTGGGAATTGCCAACAGG
TCAAGCCCTGTTCTCACTGGTC/iSpC3//iSpC3/TCACAGGTAAAGGGTCTCAGGGA衔接子_条形码49_
c-hsa-miR-190a-3p/5Phos/CCTAGTTCCGCTGGGCAGTGCTTGCCCCCAGTAGAGTGTGGAAGGGCATA/
iSpC3//iSpC3/AGGAATATGTTTGATATATAG
- [0393] 衔接子_条形码50_c-hsa-miR-193b-3p/5Phos/CCTAGTTCCGCTGGGAAAGCCCACTCTC
CACACTTCAAGGTAAATGGCG/iSpC3//iSpC3/AGCGGGACTTTGAGGGCCAGTT
- [0394] 衔接子_条形码51_c-hsa-miR-193a-5p/5Phos/CCTAGTTCCGCTGGGGACTAATACAATC
GGAAGCAACTCTCACGCCGCAC/iSpC3//iSpC3/TCATCTCGCCCGCAAAGACCCA
- [0395] 衔接子_条形码52_c-hsa-miR-211-5p/5Phos/CCTAGTTCCGCTGGGCTACGATTCATGTC
TCCCCCACATATGATTGATC/iSpC3//iSpC3/AGGCGAAGGATGACAAAGGGAA
- [0396] 衔接子_条形码53_c-hsa-miR-545-5p/5Phos/CCTAGTTCCGCTGGGTCGGGCGCTTAATC
GCAATGTTTCATCCGGAACGGA/iSpC3//iSpC3/TCATCTAATAAACATTTACTGA
- [0397] 衔接子_条形码54_c-hsa-miR-550a-5p/5Phos/CCTAGTTCCGCTGGGTAAATTCACATT
ACGATAGCTGACGTCCTGGTAG/iSpC3//iSpC3/GGGCTCTTACTCCCTCAGGCACT
- [0398] 衔接子_条形码55_c-hsa-miR-638/5Phos/CCTAGTTCCGCTGGGCACCCCCTAGCCTGACC
TTAAATGATAATTCCTGT/iSpC3//iSpC3/AGGCCGCCACCCGCCGCGATCCCT
- [0399] 衔接子_条形码56_c-hsa-miR-671-5p/5Phos/CCTAGTTCCGCTGGGTTTAGCCTAACCTC
TGATGCATTGCACCGGAGTTC/iSpC3//iSpC3/CTCCAGCCCCCTCAGGGCTTCCT
- [0400] 衔接子_条形码57_c-hsa-miR-1233-5p/5Phos/CCTAGTTCCGCTGGGCCCAATATACGCT
GAACCTTCCATCCGATTTTCAG/iSpC3//iSpC3/TGCCGTGCCCTGGCCTCCCACT
- [0401] 衔接子_条形码58_c-hsa-miR-3135b/5Phos/CCTAGTTCCGCTGGGCCTTGAATTAGAAC
CGTGTGATTCTACGCCTAGG/iSpC3//iSpC3/CACCACTGCACTCGCTCCAGCC
- [0402] 衔接子_条形码59_c-hsa-miR-3908/5Phos/CCTAGTTCCGCTGGGAGGACTTCTTCGGTGT
AATCGGAGTACATCAATGT/iSpC3//iSpC3/AAACAGTCTACCTACATTGCTC
- [0403] 衔接子_条形码60_c-hsa-miR-5571-5p/5Phos/CCTAGTTCCGCTGGGAACGTACTTTGTG
GGGATAAGCTGTACAGGGCTGT/iSpC3//iSpC3/GGGAGGCTCCTTTGAGAATTG
- [0404] 载体链序列-蛋白质和神经递质(图18)
- [0405] 心肌肌钙蛋白I(cTnI) :
- [0406] 条形码65/5Phos/CCTAGTTCCGCTGGG
- [0407] ATTCAGATCTTTGACCGATCGTACTGACGTGTACG/iSpC3//iSpC3/
- [0408] AGTCTCCGCTGTCTCCCGATGCACTTGACGTATGTCTCACTTTCTTTTCATTGACATGGGATGACGCC
GTGACTG
- [0409] 心肌肌钙蛋白T(cTnT) :

- [0410] 条形码62/5Phos/CCTAGTTCGCTGGG
- [0411] CTTAAGTTCGGTCATTAACAGTGTCAATCTTGCCA/iSpC3//iSpC3/
- [0412] ATACGGGAGCCAACACCAGGACTAACATTATAAGAATTGCGAATAATCATTGGAGAGCAGGTGTGACGG
AT
- [0413] BNP:
- [0414] 条形码61/5Phos/CCTAGTTCGCTGGG
- [0415] ATGTAATTCCTGCTAGGTCCAATGTGACTGTACTA/iSpC3//iSpC3/GGCGATTCGTGATCTCTGC
TCTCGGTTTCGCGTTCGTTTCG
- [0416] 凝血酶:
- [0417] 条形码70/5Phos/CCTAGTTCGCTGGG
- [0418] TTAAGCTATTGCTAACTGTAGTCTTAGTCTAGCTA/iSpC3//iSpC3/
AGTCCGTGGTAGGGCAGGTTGGGGTGACT
- [0419] 条形码1/5Phos/CCTAGTTCGCTGGG
- [0420] GAGCTAGCATACGTGTCTAACACTGCACAGATGAT/iSpC3//iSpC3/GGTTGGTGTGGTTGG
- [0421] S-蛋白:
- [0422] 条形码7/5Phos/CCTAGTTCGCTGGG
- [0423] ACCTTTTCTAGGATGGAACATTCTA
- [0424] /iSpC3//iSpC3/
- [0425] CACGCATAACGTCTTGCGGGGCGGCGGGTTGAGAGGATGTGGGGTGGTTATGCGTG
- [0426] N-蛋白:
- [0427] 条形码6/5Phos/CCTAGTTCGCTGGG
- [0428] CTGCGACATGACTATCTAGAGTCGC/iSpC3/iSpC3/
- [0429] GCTGGATGTACCAGATTGTGCGACATCGGATTGTCTGAGTCATATGACACATCCAGC
- [0430] 血清素:
- [0431] 条形码5/5Phos/CCTAGTTCGCTGGG
- [0432] ACTAGATAAAAGGAAGGGAGCACAGTAACGTCGTT/iSpC3/iSpC3/
- [0433] CGACTGGTAGGCAGATAGGGGAAGCTGATTCGATG
- [0434] 乙酰胆碱:
- [0435] 条形码4/5Phos/CCTAGTTCGCTGGG
- [0436] TCGATACAATACA/iSpC3/iSpC3/
- [0437] ATCCGTCACACCTGCTCTAGGGGATCAAAGCTATGCGACCATGCGAGTGGATACTGGTGTGGCTCCCG
TAT
- [0438] 靶分析物
- [0439] miRNA:
- [0440] 11 hsa-miR-497-5p rCrArGrCrArGrCrArCrArCrUrGrUrGrGrUrUrUrGrU
- [0441] 12 hsa-miR-27b-5p rArGrArGrCrUrUrArGrCrUrGrArUrUrGrGrUrGrArArC
- [0442] 13 hsa-miR-21-5p rUrArGrCrUrUrArUrCrArGrArCrUrGrArUrGrUrUrGrA
- [0443] 14 hsa-miR-221-5p rArCrCrUrGrGrCrArUrArCrArArUrGrUrArGrArUrUrU
- [0444] 15 hsa-miR-30d-5p rUrGrUrArArArCrArUrCrCrCrCrGrArCrUrGrGrArArG

- [0445] 16 hsa-miR-30c-5p rUrGrUrArArArCrArUrCrCrUrArCrArCrUrCrUrCrArGrC
- [0446] 17 hsa-miR-133a-5p rArGrCrUrGrGrUrArArArArUrGrGrArArCrCrArArArU
- [0447] 18 hsa-miR-208a-5p rGrArGrCrUrUrUrUrGrGrCrCrCrGrGrGrUrUrArUrArC
- [0448] 19 hsa-miR-181b-5p rArArCrArUrUrCrArUrUrGrCrUrGrUrCrGrGrUrGrGrGrU
- [0449] 20 hsa-miR-29a-3p rUrArGrCrArCrCrArUrCrUrGrArArArUrCrGrGrUrUrA
- [0450] 22 hsa-miR-210-5p rArGrCrCrCrCrUrGrCrCrCrArCrCrGrCrArCrArCrUrG
- [0451] 24 hsa-miR-126a-5p rCrArUrUrArUrUrArCrUrUrUrUrGrGrUrArCrGrCrG
- [0452] 25 hsa-miR-1306-5p rCrCrArCrCrUrCrCrCrCrUrGrCrArArArCrGrUrCrCrA
- [0453] 26 hsa-miR-126-5p rCrArUrUrArUrUrArCrUrUrUrUrGrGrUrArCrGrCrG
- [0454] 27 hsa-miR-1254 rArGrCrCrUrGrGrArArGrCrUrGrGrArGrCrCrUrGrCrArGrU
- [0455] 29 hsa-miR-30e-5p rUrGrUrArArArCrArUrCrCrUrUrGrArCrUrGrGrArArG
- [0456] 30 hsa-miR-106a-5p rArArArArGrUrGrCrUrUrArCrArGrUrGrCrArGrGrUrArG
- [0457] 31 hsa-miR-199a-3p rArCrArGrUrArGrUrCrUrGrCrArCrArUrUrGrGrUrUrA
- [0458] 32 hsa-miR-652-3p rArArUrGrGrCrGrCrCrArCrUrArGrGrGrUrUrGrUrG
- [0459] 33 hsa-miR-26b-5p rUrUrCrArArGrUrArArUrUrCrArGrGrArUrArGrGrU
- [0460] 34 hsa-miR-145-5p rGrUrCrCrArGrUrUrUrUrCrCrCrArGrGrArArUrCrCrCrU
- [0461] 35 hsa-miR-92a-3p rUrArUrUrGrCrArCrUrUrGrUrCrCrCrGrGrCrCrUrGrU
- [0462] 36 hsa-miR-146a-5p rUrGrArGrArArCrUrGrArArUrUrCrCrArUrGrGrGrUrU
- [0463] 37 hsa-miR-423-5p rUrGrArGrGrGrGrCrArGrArGrArGrCrGrArGrArCrUrUrU
- [0464] 39 hsa-miR-27b-3p rUrUrCrArCrArGrUrGrGrCrUrArArGrUrUrCrUrGrC
- [0465] 40 hsa-miR-1-3p rUrGrGrArArUrGrUrArArArGrArArGrUrArUrGrUrArU
- [0466] 42 hsa-miR-18a-5p rUrArArGrGrUrGrCrArUrCrUrArGrUrGrCrArGrArUrArG
- [0467] 43 hsa-miR-18b-5p rUrArArGrGrUrGrCrArUrCrUrArGrUrGrCrArGrUrUrArG
- [0468] 45 hsa-miR-301a-5p rGrCrUrCrUrGrArCrUrUrUrArUrUrGrCrArCrUrArCrU
- [0469] 46 hsa-let7c-5p rUrGrArGrGrUrArGrUrArGrGrUrUrGrUrArUrGrGrUrU
- [0470] 47 hsa-miR-125a-5p rUrCrCrCrUrGrArGrArCrCrCrUrUrUrArArCrCrUrGrUrGrA
- [0471] 49 hsa-miR-190a-3p rCrUrArUrArUrArUrCrArArArCrArUrArUrUrCrCrU
- [0472] 50 hsa-miR-193b-3p rArArCrUrGrGrCrCrCrUrCrArArArGrUrCrCrCrGrCrU
- [0473] 51 hsa-miR-193a-5p rUrGrGrGrUrCrUrUrUrGrCrGrGrGrCrGrArGrArUrGrA
- [0474] 52 hsa-miR-211-5p rUrUrCrCrCrUrUrUrGrUrCrArUrCrCrUrUrCrGrCrCrU
- [0475] 53 hsa-miR-545-5p rUrCrArGrUrArArArUrGrUrUrUrArUrUrArGrArUrGrA
- [0476] 54 hsa-miR-550a-5p rArGrUrGrCrCrUrGrArGrGrGrArGrUrArArGrArGrCrCrC
- [0477] 55 hsa-miR-638 rArGrGrGrArUrCrGrCrGrGrGrCrGrGrGrUrGrGrCrGrGrCrCrU
- [0478] 56 hsa-miR-671-5p rArGrGrArArGrCrCrCrUrGrGrArGrGrGrCrUrGrGrArG
- [0479] 57 hsa-miR-1233-5p rArGrUrGrGrGrArGrGrCrCrArGrGrGrCrArCrGrGrCrA
- [0480] 58 hsa-miR-3135b rGrGrCrUrGrGrArGrCrGrArGrUrGrCrArGrUrGrGrUrG
- [0481] 59 hsa-miR-3908 rGrArGrCrArArUrGrUrArGrGrUrArGrArCrUrGrUrUrU
- [0482] 60 hsa-miR-5571-5p rCrArArUrUrCrUrCrArArArGrGrArGrCrCrUrCrCrC
- [0483] 蛋白质:

- [0484] 心肌肌钙蛋白I (cTnI) [Genscript]
- [0485] 心肌肌钙蛋白T (cTnT) : TNNT2人重组蛋白 | CTnT抗原 |
- [0486] ProSpec (prospecbio.com)
- [0487] BNP-32肽: [Bachem]
- [0488] 人 α -凝血酶天然蛋白, 生物素 (RP-43103) [Thermofisher]
- [0489] 参考实例1
- [0490] 不使用马达蛋白的多重测定的局限性。
- [0491] 如上所述, 载体上马达蛋白的存在允许准确识别标识符区并确定分子在分子结合区上的存在或不存在。当这些方法在不存在马达蛋白的情况下执行时, 可以获得的测量结果受到限制。
- [0492] 在一项实验中, 使用包括适体 (下划线) 和30*T螺旋链的载体通过纳米孔测量来检测凝血酶浓度的增加:
- [0493] TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAGTCCGTGGTAGGGCAGGTTGGGGTGACT (G4)。
- [0494] 通过这种方法, 可以量化凝血酶的浓度, 但无法确认误报, 也没有简单的方法来多重化检测。测量的所有均是适体和蛋白质结合适体之间信号水平的差异 (图15A)。
- [0495] 在第二实验中, 使用包括茎环适体 (下划线) 和30*T螺旋链的载体, 通过纳米孔测量来检测血清素浓度的增加:
- [0496] TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCGACTGGTAGGCAGATAGGGGAAGCTGATTCGATGCGTGGGTCG。
- [0497] 以与凝血酶实验相同的方式, 可以量化血清素的浓度, 但无法确认误报, 也没有简单的方法来多重化检测。测量的所有均是适体和神经递质结合适体之间信号水平的差异 (图15B)。
- [0498] 在第三实验中, 以多重形式检测miRNA的水平。条形码1至6 (ACGTA、GGACT、TTAAC、GCTAG、CTGAG和TAGCG) 通过针对每个条形码观察到的唯一平均电流 (分别为122.4pA、121.8pA、111.5pA、152.7pA、129.3pA、135.4pA) 来进行识别 (图14C)。然而, 这种方法的多重化能力受限, 并且受限于信号幅度波动。实际上, 多重检测中可以使用五个条形码。由于针对每个条形码观察到的幅度分布与其他条形码的特征信号重叠, 因此条形码的分配和分类不一定能够准确分类。
- [0499] 序列表
- [0500] SEQ ID NO:1-来自大肠杆菌的核酸外切酶I
- [0501] MMNDGKQQSTFLFHDYETFGTHPALDRPAQFAAIRTDFSEFNVIGEPEVFYCKPADDYLP
- [0502] QPGAVLITGITPQEARAKGENEAAFAARIHSLFTVPKTCILGYNNVRFDDEVTRNIFYRN
- [0503] FYDPYAWSWQHDNSRWDLDDVMRACYALRPEGINWPENDDGLPSFRLEHLTKANGIE
- [0504] HSNAMDAMADVYATIAMAKLVKTRQPRFLDYLFTHRNKHKLMALIDVPQMKPLVHVS
- [0505] GMFGAWRGNTSWVAPLAWHPENRNAVIMVDLAGDISPLLELSDTLRERLYTAKTDL
- [0506] GDNAAPVKLVHINKCPVLAQANTLRPEDADRLGINRQHCLDNLKILRENPPQVREKV
- [0507] AIFAEAEFPTPSDNVDAQLYNGFFSDADRAAMKIVLETEPRNLPALDITFVDKRIEKLFL
- [0508] NYRARNFPGTLDYAEQQRWLEHRRQVFTPEFLQGYADELQMLVQQYADDKEKVALLK
- [0509] ALWQYAEIIVSGSGHHHHHH

- [0510] SEQ ID NO:2-来自大肠杆菌的核酸外切酶III
[0511] MKFVSFNINGLRARPHQLEAIVEKHQPDVIGLQETKVHDDMFPLEEVAKLGYNVIFYHG
[0512] QKGHYGVALLTKETPIAVRRGFPGDDEEAQRRIIMAEIPSLLGNVTVINGYFPQGESRDH
[0513] PIKFPKAAQFYQNLQNYLETELKRDNVPLIMGDMNISPTDLDIGIGEENRKRWLRTGKC
[0514] SFLPEEREWMRLMSWGLVDTRHANPQTADRFSWFDYRSKGFDDNRGLRIDLLLAS
[0515] QPLAECCVETGIDYEIRSMEKPSDHAPVWATFRR
[0516] SEQ ID NO:3-来自嗜热链球菌的RecJ酶
[0517] MFRRKEDLDPLALLPLKGLREAAALLEEALRQGKRIRVHG DYDADGLTG TAILVRGL
[0518] AALGADVHPFIPHRLEEGYGVLMERVPEHLEASDLFLTVDCGITNHAELRELENGVE
[0519] VIVTDHHTPGKT PPPGLV VHPALTPDLKEKPTGAGVAFLLLWALHERLGLPPPLEYADL
[0520] AAVGTIADVAPLWGNRALVKEGLARIPASSWVGLRLLAEAVGYTGKAVEVAFRIAPR
[0521] INAASRLGEAEKALRLLLTDDAAEAQALVGELHRLNARRQTLEEAMLRKLLPQADPEA
[0522] KAIVLLDPEGHPGVMGIVASRILEATLRPVFLVAQGGKT VRS LAPISAVEALRSAEDLLL
[0523] RYGGHKEAAGFAMDEALFPAFKARVEAYAARFPDPVREVALLDLLPEPGLLPQVFREL
[0524] ALLEPYGEGNPEPLFL
[0525] SEQ ID NO:4-噬菌体λ核酸外切酶
[0526] MTPDIILQRTGIDVRAVEQGD DAWHKLRLGVITASEVHNVI AKPRSGKKWPD MKMSYF
[0527] HTLLAEVCTGVAPEVNAKALAWGKQYENDARTLFEFTSGVNVTESPIIYRDESMRTAC
[0528] SPDGLCSDGNGLLELKCPFTSRDFMFKRLGGFEAIKSAYMAQVQYSMWVTRKNAWYFA
[0529] NYDPRMKREGLHYV VIERDEKYMASFDEIVPEFIEKMDEALAEIGFVFGEQWR
[0530] SEQ ID NO:5-Phi29 DNA聚合酶
[0531] MKHMPRKMYS CAFETTTKVEDCRVWAYGYMNIEDHSEYKIGNSLDEFMAWVLKVQA
[0532] DLYFHNLFKDFAGFIINWLERNGFKWSADGLPNTYNTIISR MGQWY MIDICLG YKGK RK
[0533] IHTVIYDSLKKLPFPVKKI AKDFKLT VLGK DIDYHKERPVG YKITPEEYAYIKNDIQIIAE
[0534] ALLIQFKQGLDRMTAGSDSLKGFKDIITTKFKKVFPTLSLGLDKEVRYAYRGGFTWLN
[0535] DRFKEKEIGEGMVF DVNSLYPAQMYSRLLPYGEP IVFEGKYVWDEDYPLHIQHIRCEFE
[0536] LKEGYIPTIQIKRSRFYK GNEYLKSSGGEIADLWLSNVDLELMKEHYDLYNVEYISGLK
[0537] FKATTGLFKDFIDKWTYIKTTSEGAIKQLAKLMLNSLYGKFASNP DVTGKVPYLKENG
[0538] ALGFRLGEEETKDPVYTPMGVFITAWARYTTITAAQACYDRIIYCDTDSIHLTGTEIPDVI
[0539] KDIVDPKKLGYWAHESTFKRAKYLRQKTYIQDIYMKEVDGKLVEGSPDDYTDIKFSVK
[0540] CAGMTDKIKKEVTFENFKVGF SRKMKPKPVQVPGGVVLVDDTFTIKSGGSAWSHPQFE
[0541] KGGSGGGSGGSAWSHPQFEK
[0542] SEQ ID NO:6-Trwc Cba解旋酶
[0543] MLSVANVRSPSAAASYFASDNYYASADADRSGQWIGDGA KRLGLEGKVEARAFDALL
[0544] RGELPDGSSVGNPGQAHRPGTDLTFSVPKSWSLALVGKDERIIAAYREAVVEALHWA
[0545] EKNAAE TRVVEKGMVVTQATGNLAIGLFQHD TNRNQEPNLHFHAVIANVTQ GKDGK
[0546] WRTLKNDRLWQLNTTLNSIAMARFRVAVEKLG YEPGPVLKHGNFEARGISREQVMAFS
[0547] TRRKEVLEARRGPGLDAGRI AALDTRASKEGIEDRATLSKQWSEAAQSIGLDLKLPLVD
[0548] RARTKALQGMEATRIGSLVERGRAWLSRFAAHVRGDPADPLVPPSVL KQDRQTIAAA

[0549] QAVASAVRHLSQREAAFERTALYKAALDFGLPTTIADVEKRTRALVRSGDLIAGKGEHK
[0550] GWLASRDAVVTEQRILSEVAAGKGDSSPAITPQKAAASVQAAALTGQGFRLNEGQLAA
[0551] ARLILISKDRITIAVQGIAGAGKSSVLKPVAEVL RDEGHPVIGLAIQNTLVQMLERDTGIG
[0552] SQTARFLGGWNKLLDDPGNVALRAEAQASLKDHLVLDEASMVSNEDEKELVRLA
[0553] NLAGVHRLVLIGDRKQLGAVDAGKPFALLQRAGIARAEMATNL RARDPVVREAQAAA
[0554] QAGDVRKALRHLKSHTVEARGDGAQVAAETWLALDKETRARTSIYASGRAIRSAVNA
[0555] AVQQGLLASREIGPAKMKEVLDRVNTTREELRHLPAYRAGRVLEVSRKQQALGLFIG
[0556] EYRVIGQDRKGKLVEVEDKRGKRFRFDPARIRAGKGDDNLTLLEPRKLEIHEGDRI RWT
[0557] RNDHRRGLFNADQARVVEIANGKVTFTESKGDLEVELKKDDPMLKRIDLAYALNVHMA
[0558] QGLTSDRGIAVMDSRERNLSNQKTFLVTVTRLRDHLTLVVSADKLGAAVARNKGEKA
[0559] SAIEVTGSVKPTATKSGVDQPKSVEANKAEKELTRSKSKTLDFGI
[0560] SEQ ID NO:7-He1308 Mbu解旋酶
[0561] MMIRELDIPRDIIGFYEDSGIKELYPPQAEAIEMGLLEKKNLLAAIPTASGKTLLAELAMI
[0562] KAIREGGKALYIVPLRALASEKFERFKELAPFGIKVGISTGDLDSRADWLGVNDIIVATS
[0563] EKTDSLRLNGTSWMDIITTVVDEIHLSDSKNRGPTLEVTITKLMRLNPDVQVVALSA
[0564] TVGNAREMADWLGAALVLSEWRPTDLHEGVLFGDAINFPGSQKKIDRLEKDDAVNLV
[0565] LDTIKAEGQCLVFESSRRNCAGFAKTASSKVAKILDNDIMIKLAGIAEEVESTGETDTAI
[0566] VLANCIRKGVAFHHAGLNSNHRKLVENGFRQNLIKVISSTPTLAAGLNL PARRVIIRSYR
[0567] RFDSNFGMQPIPVLEYKQMAGRAGRPHLDPYGESVLLAKTYDEFAQLMENYVEADAE
[0568] DIWSKLG TENALRTHVLSTIVNGFASTRQELFDFFGATFFAYQQDKWMLEEVINDCLEF
[0569] LIDKAMVSETEDIEDASKLFLRGTRLGSLVSM LYIDPLSGSKIVDGFKDIGKSTGGNMG
[0570] SLEDDKGDDITVDTMTLLHLVCSTPDMRQLYLRNTDYTIVNEYIVAHSDEFHEIPDKLK
[0571] ETDYEFWMGEVK TAMLLEEWVTEVSAEDITRHFNVGEGDIHALADTSEWLMHAAAK
[0572] LAELLGVEYSSHAYSLEKRIRYGSGLDLMELVGIRGVGRVRARKLYNAGFVSVAKLKG
[0573] ADISVLSKLVGPKVAYNILSGIGVRVNDKHFNSAPISSNTLDTLLDKNQKTFNDFQ
[0574] SEQ ID NO:8-Dda解旋酶
[0575] MTFDDLTEGQKNAFNIVMKAIKEKKHHVTINGPAGTGKTTLT KFIIEALISTGETGIILA
[0576] APTHAAKKILSKLSGKEASTIHSILKINPVTYEENVLFEQKEVPDLAKCRVLICDEVSMY
[0577] DRKLFKILLSTIPPWCTIIGIGDNKQIRVPDPGENTAYISPFFTHKDFYQCELTEVKRSNA
[0578] PIIDVATDV RNKGWIYDKVVDGHGVRGFTGDTALRDFMVNYFSIVKSLDDL FENRVMA
[0579] FTNKSVDKLSNIIRKKIFETDKDFIVGEIIVMQEPLFKTYKIDGKPVSEIIFNNGQLVRIIE
[0580] AEYTSTFVKARGVPGEYLIRHWDLT VETYGDDEYYREKIKI ISSDEELYKFNFLGKTA
[0581] ETYKNWNKGKAPWSDFWDAKSQFSKV KALPASTFHKAQGMSVDRAFIYTPCIHYAD
[0582] VELAQQLLYVGVTGRYDV FYV

序列表

<110> 牛津纳米孔科技公开有限公司和帝国理工学院创新有限公司

<120> 方法

<130> N421601W0

<150> GB2107192.3

<151> 2021-05-19

<160> 157

<170> PatentIn版本3.5

<210> 1

<211> 485

<212> PRT

<213> 大肠杆菌

<400> 1

```

Met Met Asn Asp Gly Lys Gln Gln Ser Thr Phe Leu Phe His Asp Tyr
1           5           10           15
Glu Thr Phe Gly Thr His Pro Ala Leu Asp Arg Pro Ala Gln Phe Ala
          20           25           30
Ala Ile Arg Thr Asp Ser Glu Phe Asn Val Ile Gly Glu Pro Glu Val
          35           40           45
Phe Tyr Cys Lys Pro Ala Asp Asp Tyr Leu Pro Gln Pro Gly Ala Val
          50           55           60
Leu Ile Thr Gly Ile Thr Pro Gln Glu Ala Arg Ala Lys Gly Glu Asn
65           70           75           80
Glu Ala Ala Phe Ala Ala Arg Ile His Ser Leu Phe Thr Val Pro Lys
          85           90           95
Thr Cys Ile Leu Gly Tyr Asn Asn Val Arg Phe Asp Asp Glu Val Thr
          100          105          110
Arg Asn Ile Phe Tyr Arg Asn Phe Tyr Asp Pro Tyr Ala Trp Ser Trp
          115          120          125
Gln His Asp Asn Ser Arg Trp Asp Leu Leu Asp Val Met Arg Ala Cys
          130          135          140
Tyr Ala Leu Arg Pro Glu Gly Ile Asn Trp Pro Glu Asn Asp Asp Gly
145          150          155          160
Leu Pro Ser Phe Arg Leu Glu His Leu Thr Lys Ala Asn Gly Ile Glu
          165          170          175
His Ser Asn Ala His Asp Ala Met Ala Asp Val Tyr Ala Thr Ile Ala
          180          185          190
Met Ala Lys Leu Val Lys Thr Arg Gln Pro Arg Leu Phe Asp Tyr Leu

```

195	200	205
Phe Thr His Arg Asn Lys His Lys Leu Met Ala Leu Ile Asp Val Pro		
210	215	220
Gln Met Lys Pro Leu Val His Val Ser Gly Met Phe Gly Ala Trp Arg		
225	230	235
Gly Asn Thr Ser Trp Val Ala Pro Leu Ala Trp His Pro Glu Asn Arg		
245	250	255
Asn Ala Val Ile Met Val Asp Leu Ala Gly Asp Ile Ser Pro Leu Leu		
260	265	270
Glu Leu Asp Ser Asp Thr Leu Arg Glu Arg Leu Tyr Thr Ala Lys Thr		
275	280	285
Asp Leu Gly Asp Asn Ala Ala Val Pro Val Lys Leu Val His Ile Asn		
290	295	300
Lys Cys Pro Val Leu Ala Gln Ala Asn Thr Leu Arg Pro Glu Asp Ala		
305	310	315
Asp Arg Leu Gly Ile Asn Arg Gln His Cys Leu Asp Asn Leu Lys Ile		
325	330	335
Leu Arg Glu Asn Pro Gln Val Arg Glu Lys Val Val Ala Ile Phe Ala		
340	345	350
Glu Ala Glu Pro Phe Thr Pro Ser Asp Asn Val Asp Ala Gln Leu Tyr		
355	360	365
Asn Gly Phe Phe Ser Asp Ala Asp Arg Ala Ala Met Lys Ile Val Leu		
370	375	380
Glu Thr Glu Pro Arg Asn Leu Pro Ala Leu Asp Ile Thr Phe Val Asp		
385	390	395
Lys Arg Ile Glu Lys Leu Leu Phe Asn Tyr Arg Ala Arg Asn Phe Pro		
405	410	415
Gly Thr Leu Asp Tyr Ala Glu Gln Gln Arg Trp Leu Glu His Arg Arg		
420	425	430
Gln Val Phe Thr Pro Glu Phe Leu Gln Gly Tyr Ala Asp Glu Leu Gln		
435	440	445
Met Leu Val Gln Gln Tyr Ala Asp Asp Lys Glu Lys Val Ala Leu Leu		
450	455	460
Lys Ala Leu Trp Gln Tyr Ala Glu Glu Ile Val Ser Gly Ser Gly His		
465	470	475
His His His His His		
485		

<210> 2

<211> 268

<212> PRT

<213> 大肠杆菌

<400> 2

```

Met Lys Phe Val Ser Phe Asn Ile Asn Gly Leu Arg Ala Arg Pro His
1           5           10           15
Gln Leu Glu Ala Ile Val Glu Lys His Gln Pro Asp Val Ile Gly Leu
          20           25           30
Gln Glu Thr Lys Val His Asp Asp Met Phe Pro Leu Glu Glu Val Ala
          35           40           45
Lys Leu Gly Tyr Asn Val Phe Tyr His Gly Gln Lys Gly His Tyr Gly
          50           55           60
Val Ala Leu Leu Thr Lys Glu Thr Pro Ile Ala Val Arg Arg Gly Phe
65           70           75           80
Pro Gly Asp Asp Glu Glu Ala Gln Arg Arg Ile Ile Met Ala Glu Ile
          85           90           95
Pro Ser Leu Leu Gly Asn Val Thr Val Ile Asn Gly Tyr Phe Pro Gln
          100          105          110
Gly Glu Ser Arg Asp His Pro Ile Lys Phe Pro Ala Lys Ala Gln Phe
          115          120          125
Tyr Gln Asn Leu Gln Asn Tyr Leu Glu Thr Glu Leu Lys Arg Asp Asn
          130          135          140
Pro Val Leu Ile Met Gly Asp Met Asn Ile Ser Pro Thr Asp Leu Asp
          145          150          155          160
Ile Gly Ile Gly Glu Glu Asn Arg Lys Arg Trp Leu Arg Thr Gly Lys
          165          170          175
Cys Ser Phe Leu Pro Glu Glu Arg Glu Trp Met Asp Arg Leu Met Ser
          180          185          190
Trp Gly Leu Val Asp Thr Phe Arg His Ala Asn Pro Gln Thr Ala Asp
          195          200          205
Arg Phe Ser Trp Phe Asp Tyr Arg Ser Lys Gly Phe Asp Asp Asn Arg
          210          215          220
Gly Leu Arg Ile Asp Leu Leu Leu Ala Ser Gln Pro Leu Ala Glu Cys
          225          230          235          240
Cys Val Glu Thr Gly Ile Asp Tyr Glu Ile Arg Ser Met Glu Lys Pro
          245          250          255
Ser Asp His Ala Pro Val Trp Ala Thr Phe Arg Arg
          260          265

```

<210> 3

<211> 425

<212> PRT

<213> 大肠杆菌

<400> 3

```

Met Phe Arg Arg Lys Glu Asp Leu Asp Pro Pro Leu Ala Leu Leu Pro
1           5           10           15
Leu Lys Gly Leu Arg Glu Ala Ala Ala Leu Leu Glu Glu Ala Leu Arg
          20           25           30
Gln Gly Lys Arg Ile Arg Val His Gly Asp Tyr Asp Ala Asp Gly Leu
          35           40           45
Thr Gly Thr Ala Ile Leu Val Arg Gly Leu Ala Ala Leu Gly Ala Asp
          50           55           60
Val His Pro Phe Ile Pro His Arg Leu Glu Glu Gly Tyr Gly Val Leu
65           70           75           80
Met Glu Arg Val Pro Glu His Leu Glu Ala Ser Asp Leu Phe Leu Thr
          85           90           95
Val Asp Cys Gly Ile Thr Asn His Ala Glu Leu Arg Glu Leu Leu Glu
          100          105          110
Asn Gly Val Glu Val Ile Val Thr Asp His His Thr Pro Gly Lys Thr
          115          120          125
Pro Pro Pro Gly Leu Val Val His Pro Ala Leu Thr Pro Asp Leu Lys
          130          135          140
Glu Lys Pro Thr Gly Ala Gly Val Ala Phe Leu Leu Leu Trp Ala Leu
145          150          155          160
His Glu Arg Leu Gly Leu Pro Pro Pro Leu Glu Tyr Ala Asp Leu Ala
          165          170          175
Ala Val Gly Thr Ile Ala Asp Val Ala Pro Leu Trp Gly Trp Asn Arg
          180          185          190
Ala Leu Val Lys Glu Gly Leu Ala Arg Ile Pro Ala Ser Ser Trp Val
          195          200          205
Gly Leu Arg Leu Leu Ala Glu Ala Val Gly Tyr Thr Gly Lys Ala Val
          210          215          220
Glu Val Ala Phe Arg Ile Ala Pro Arg Ile Asn Ala Ala Ser Arg Leu
225          230          235          240
Gly Glu Ala Glu Lys Ala Leu Arg Leu Leu Leu Thr Asp Asp Ala Ala
          245          250          255
Glu Ala Gln Ala Leu Val Gly Glu Leu His Arg Leu Asn Ala Arg Arg
          260          265          270
Gln Thr Leu Glu Glu Ala Met Leu Arg Lys Leu Leu Pro Gln Ala Asp
          275          280          285

```

Pro Glu Ala Lys Ala Ile Val Leu Leu Asp Pro Glu Gly His Pro Gly
 290 295 300
 Val Met Gly Ile Val Ala Ser Arg Ile Leu Glu Ala Thr Leu Arg Pro
 305 310 315 320
 Val Phe Leu Val Ala Gln Gly Lys Gly Thr Val Arg Ser Leu Ala Pro
 325 330 335
 Ile Ser Ala Val Glu Ala Leu Arg Ser Ala Glu Asp Leu Leu Leu Arg
 340 345 350
 Tyr Gly Gly His Lys Glu Ala Ala Gly Phe Ala Met Asp Glu Ala Leu
 355 360 365
 Phe Pro Ala Phe Lys Ala Arg Val Glu Ala Tyr Ala Ala Arg Phe Pro
 370 375 380
 Asp Pro Val Arg Glu Val Ala Leu Leu Asp Leu Leu Pro Glu Pro Gly
 385 390 395 400
 Leu Leu Pro Gln Val Phe Arg Glu Leu Ala Leu Leu Glu Pro Tyr Gly
 405 410 415
 Glu Gly Asn Pro Glu Pro Leu Phe Leu
 420 425
 <210> 4
 <211> 226
 <212> PRT
 <213> 噬菌体λ
 <400> 4
 Met Thr Pro Asp Ile Ile Leu Gln Arg Thr Gly Ile Asp Val Arg Ala
 1 5 10 15
 Val Glu Gln Gly Asp Asp Ala Trp His Lys Leu Arg Leu Gly Val Ile
 20 25 30
 Thr Ala Ser Glu Val His Asn Val Ile Ala Lys Pro Arg Ser Gly Lys
 35 40 45
 Lys Trp Pro Asp Met Lys Met Ser Tyr Phe His Thr Leu Leu Ala Glu
 50 55 60
 Val Cys Thr Gly Val Ala Pro Glu Val Asn Ala Lys Ala Leu Ala Trp
 65 70 75 80
 Gly Lys Gln Tyr Glu Asn Asp Ala Arg Thr Leu Phe Glu Phe Thr Ser
 85 90 95
 Gly Val Asn Val Thr Glu Ser Pro Ile Ile Tyr Arg Asp Glu Ser Met
 100 105 110
 Arg Thr Ala Cys Ser Pro Asp Gly Leu Cys Ser Asp Gly Asn Gly Leu
 115 120 125

Glu Leu Lys Cys Pro Phe Thr Ser Arg Asp Phe Met Lys Phe Arg Leu
 130 135 140
 Gly Gly Phe Glu Ala Ile Lys Ser Ala Tyr Met Ala Gln Val Gln Tyr
 145 150 155 160
 Ser Met Trp Val Thr Arg Lys Asn Ala Trp Tyr Phe Ala Asn Tyr Asp
 165 170 175
 Pro Arg Met Lys Arg Glu Gly Leu His Tyr Val Val Ile Glu Arg Asp
 180 185 190
 Glu Lys Tyr Met Ala Ser Phe Asp Glu Ile Val Pro Glu Phe Ile Glu
 195 200 205
 Lys Met Asp Glu Ala Leu Ala Glu Ile Gly Phe Val Phe Gly Glu Gln
 210 215 220
 Trp Arg
 225
 <210> 5
 <211> 608
 <212> PRT
 <213> 噬菌体phi-29
 <400> 5
 Met Lys His Met Pro Arg Lys Met Tyr Ser Cys Ala Phe Glu Thr Thr
 1 5 10 15
 Thr Lys Val Glu Asp Cys Arg Val Trp Ala Tyr Gly Tyr Met Asn Ile
 20 25 30
 Glu Asp His Ser Glu Tyr Lys Ile Gly Asn Ser Leu Asp Glu Phe Met
 35 40 45
 Ala Trp Val Leu Lys Val Gln Ala Asp Leu Tyr Phe His Asn Leu Lys
 50 55 60
 Phe Asp Gly Ala Phe Ile Ile Asn Trp Leu Glu Arg Asn Gly Phe Lys
 65 70 75 80
 Trp Ser Ala Asp Gly Leu Pro Asn Thr Tyr Asn Thr Ile Ile Ser Arg
 85 90 95
 Met Gly Gln Trp Tyr Met Ile Asp Ile Cys Leu Gly Tyr Lys Gly Lys
 100 105 110
 Arg Lys Ile His Thr Val Ile Tyr Asp Ser Leu Lys Lys Leu Pro Phe
 115 120 125
 Pro Val Lys Lys Ile Ala Lys Asp Phe Lys Leu Thr Val Leu Lys Gly
 130 135 140
 Asp Ile Asp Tyr His Lys Glu Arg Pro Val Gly Tyr Lys Ile Thr Pro
 145 150 155 160

Glu	Glu	Tyr	Ala	Tyr	Ile	Lys	Asn	Asp	Ile	Gln	Ile	Ile	Ala	Glu	Ala
						165				170				175	
Leu	Leu	Ile	Gln	Phe	Lys	Gln	Gly	Leu	Asp	Arg	Met	Thr	Ala	Gly	Ser
						180				185				190	
Asp	Ser	Leu	Lys	Gly	Phe	Lys	Asp	Ile	Ile	Thr	Thr	Lys	Lys	Phe	Lys
						195				200				205	
Lys	Val	Phe	Pro	Thr	Leu	Ser	Leu	Gly	Leu	Asp	Lys	Glu	Val	Arg	Tyr
						210				215				220	
Ala	Tyr	Arg	Gly	Gly	Phe	Thr	Trp	Leu	Asn	Asp	Arg	Phe	Lys	Glu	Lys
225						230				235				240	
Glu	Ile	Gly	Glu	Gly	Met	Val	Phe	Asp	Val	Asn	Ser	Leu	Tyr	Pro	Ala
						245				250				255	
Gln	Met	Tyr	Ser	Arg	Leu	Leu	Pro	Tyr	Gly	Glu	Pro	Ile	Val	Phe	Glu
						260				265				270	
Gly	Lys	Tyr	Val	Trp	Asp	Glu	Asp	Tyr	Pro	Leu	His	Ile	Gln	His	Ile
						275				280				285	
Arg	Cys	Glu	Phe	Glu	Leu	Lys	Glu	Gly	Tyr	Ile	Pro	Thr	Ile	Gln	Ile
						290				295				300	
Lys	Arg	Ser	Arg	Phe	Tyr	Lys	Gly	Asn	Glu	Tyr	Leu	Lys	Ser	Ser	Gly
305						310				315				320	
Gly	Glu	Ile	Ala	Asp	Leu	Trp	Leu	Ser	Asn	Val	Asp	Leu	Glu	Leu	Met
						325				330				335	
Lys	Glu	His	Tyr	Asp	Leu	Tyr	Asn	Val	Glu	Tyr	Ile	Ser	Gly	Leu	Lys
						340				345				350	
Phe	Lys	Ala	Thr	Thr	Gly	Leu	Phe	Lys	Asp	Phe	Ile	Asp	Lys	Trp	Thr
						355				360				365	
Tyr	Ile	Lys	Thr	Thr	Ser	Glu	Gly	Ala	Ile	Lys	Gln	Leu	Ala	Lys	Leu
						370				375				380	
Met	Leu	Asn	Ser	Leu	Tyr	Gly	Lys	Phe	Ala	Ser	Asn	Pro	Asp	Val	Thr
385						390				395				400	
Gly	Lys	Val	Pro	Tyr	Leu	Lys	Glu	Asn	Gly	Ala	Leu	Gly	Phe	Arg	Leu
						405				410				415	
Gly	Glu	Glu	Glu	Thr	Lys	Asp	Pro	Val	Tyr	Thr	Pro	Met	Gly	Val	Phe
						420				425				430	
Ile	Thr	Ala	Trp	Ala	Arg	Tyr	Thr	Thr	Ile	Thr	Ala	Ala	Gln	Ala	Cys
						435				440				445	
Tyr	Asp	Arg	Ile	Ile	Tyr	Cys	Asp	Thr	Asp	Ser	Ile	His	Leu	Thr	Gly
						450				455				460	
Thr	Glu	Ile	Pro	Asp	Val	Ile	Lys	Asp	Ile	Val	Asp	Pro	Lys	Lys	Leu

465	470	475	480
Gly Tyr Trp Ala His Glu Ser Thr Phe Lys Arg Ala Lys Tyr Leu Arg			
	485	490	495
Gln Lys Thr Tyr Ile Gln Asp Ile Tyr Met Lys Glu Val Asp Gly Lys			
	500	505	510
Leu Val Glu Gly Ser Pro Asp Asp Tyr Thr Asp Ile Lys Phe Ser Val			
	515	520	525
Lys Cys Ala Gly Met Thr Asp Lys Ile Lys Lys Glu Val Thr Phe Glu			
	530	535	540
Asn Phe Lys Val Gly Phe Ser Arg Lys Met Lys Pro Lys Pro Val Gln			
545	550	555	560
Val Pro Gly Gly Val Val Leu Val Asp Asp Thr Phe Thr Ile Lys Ser			
	565	570	575
Gly Gly Ser Ala Trp Ser His Pro Gln Phe Glu Lys Gly Gly Gly Ser			
	580	585	590
Gly Gly Gly Ser Gly Gly Ser Ala Trp Ser His Pro Gln Phe Glu Lys			
	595	600	605
<210> 6			
<211> 970			
<212> PRT			
<213> 人工序列			
<220>			
<223> Trwc Cba解旋酶			
<400> 6			
Met Leu Ser Val Ala Asn Val Arg Ser Pro Ser Ala Ala Ala Ser Tyr			
1	5	10	15
Phe Ala Ser Asp Asn Tyr Tyr Ala Ser Ala Asp Ala Asp Arg Ser Gly			
	20	25	30
Gln Trp Ile Gly Asp Gly Ala Lys Arg Leu Gly Leu Glu Gly Lys Val			
	35	40	45
Glu Ala Arg Ala Phe Asp Ala Leu Leu Arg Gly Glu Leu Pro Asp Gly			
	50	55	60
Ser Ser Val Gly Asn Pro Gly Gln Ala His Arg Pro Gly Thr Asp Leu			
65	70	75	80
Thr Phe Ser Val Pro Lys Ser Trp Ser Leu Leu Ala Leu Val Gly Lys			
	85	90	95
Asp Glu Arg Ile Ile Ala Ala Tyr Arg Glu Ala Val Val Glu Ala Leu			
	100	105	110
His Trp Ala Glu Lys Asn Ala Ala Glu Thr Arg Val Val Glu Lys Gly			

115	120	125
Met Val Val Thr Gln Ala Thr Gly Asn Leu Ala Ile Gly Leu Phe Gln		
130	135	140
His Asp Thr Asn Arg Asn Gln Glu Pro Asn Leu His Phe His Ala Val		
145	150	155
Ile Ala Asn Val Thr Gln Gly Lys Asp Gly Lys Trp Arg Thr Leu Lys		
165	170	175
Asn Asp Arg Leu Trp Gln Leu Asn Thr Thr Leu Asn Ser Ile Ala Met		
180	185	190
Ala Arg Phe Arg Val Ala Val Glu Lys Leu Gly Tyr Glu Pro Gly Pro		
195	200	205
Val Leu Lys His Gly Asn Phe Glu Ala Arg Gly Ile Ser Arg Glu Gln		
210	215	220
Val Met Ala Phe Ser Thr Arg Arg Lys Glu Val Leu Glu Ala Arg Arg		
225	230	235
Gly Pro Gly Leu Asp Ala Gly Arg Ile Ala Ala Leu Asp Thr Arg Ala		
245	250	255
Ser Lys Glu Gly Ile Glu Asp Arg Ala Thr Leu Ser Lys Gln Trp Ser		
260	265	270
Glu Ala Ala Gln Ser Ile Gly Leu Asp Leu Lys Pro Leu Val Asp Arg		
275	280	285
Ala Arg Thr Lys Ala Leu Gly Gln Gly Met Glu Ala Thr Arg Ile Gly		
290	295	300
Ser Leu Val Glu Arg Gly Arg Ala Trp Leu Ser Arg Phe Ala Ala His		
305	310	315
Val Arg Gly Asp Pro Ala Asp Pro Leu Val Pro Pro Ser Val Leu Lys		
325	330	335
Gln Asp Arg Gln Thr Ile Ala Ala Ala Gln Ala Val Ala Ser Ala Val		
340	345	350
Arg His Leu Ser Gln Arg Glu Ala Ala Phe Glu Arg Thr Ala Leu Tyr		
355	360	365
Lys Ala Ala Leu Asp Phe Gly Leu Pro Thr Thr Ile Ala Asp Val Glu		
370	375	380
Lys Arg Thr Arg Ala Leu Val Arg Ser Gly Asp Leu Ile Ala Gly Lys		
385	390	395
Gly Glu His Lys Gly Trp Leu Ala Ser Arg Asp Ala Val Val Thr Glu		
405	410	415
Gln Arg Ile Leu Ser Glu Val Ala Ala Gly Lys Gly Asp Ser Ser Pro		
420	425	430

Ala Ile Thr Pro Gln Lys Ala Ala Ala Ser Val Gln Ala Ala Ala Leu		
435	440	445
Thr Gly Gln Gly Phe Arg Leu Asn Glu Gly Gln Leu Ala Ala Ala Arg		
450	455	460
Leu Ile Leu Ile Ser Lys Asp Arg Thr Ile Ala Val Gln Gly Ile Ala		
465	470	475
Gly Ala Gly Lys Ser Ser Val Leu Lys Pro Val Ala Glu Val Leu Arg		
	485	490
Asp Glu Gly His Pro Val Ile Gly Leu Ala Ile Gln Asn Thr Leu Val		
	500	505
Gln Met Leu Glu Arg Asp Thr Gly Ile Gly Ser Gln Thr Leu Ala Arg		
	515	520
Phe Leu Gly Gly Trp Asn Lys Leu Leu Asp Asp Pro Gly Asn Val Ala		
	530	535
Leu Arg Ala Glu Ala Gln Ala Ser Leu Lys Asp His Val Leu Val Leu		
	545	550
Asp Glu Ala Ser Met Val Ser Asn Glu Asp Lys Glu Lys Leu Val Arg		
	565	570
Leu Ala Asn Leu Ala Gly Val His Arg Leu Val Leu Ile Gly Asp Arg		
	580	585
Lys Gln Leu Gly Ala Val Asp Ala Gly Lys Pro Phe Ala Leu Leu Gln		
	595	600
Arg Ala Gly Ile Ala Arg Ala Glu Met Ala Thr Asn Leu Arg Ala Arg		
	610	615
Asp Pro Val Val Arg Glu Ala Gln Ala Ala Ala Gln Ala Gly Asp Val		
	625	630
Arg Lys Ala Leu Arg His Leu Lys Ser His Thr Val Glu Ala Arg Gly		
	645	650
Asp Gly Ala Gln Val Ala Ala Glu Thr Trp Leu Ala Leu Asp Lys Glu		
	660	665
Thr Arg Ala Arg Thr Ser Ile Tyr Ala Ser Gly Arg Ala Ile Arg Ser		
	675	680
Ala Val Asn Ala Ala Val Gln Gln Gly Leu Leu Ala Ser Arg Glu Ile		
	690	695
Gly Pro Ala Lys Met Lys Leu Glu Val Leu Asp Arg Val Asn Thr Thr		
	705	710
Arg Glu Glu Leu Arg His Leu Pro Ala Tyr Arg Ala Gly Arg Val Leu		
	725	730
Glu Val Ser Arg Lys Gln Gln Ala Leu Gly Leu Phe Ile Gly Glu Tyr		

Glu Asp Ser Gly Ile Lys Glu Leu Tyr Pro Pro Gln Ala Glu Ala Ile

20	25	30
Glu Met Gly Leu Leu Glu Lys Lys Asn Leu Leu Ala Ala Ile Pro Thr		
35	40	45
Ala Ser Gly Lys Thr Leu Leu Ala Glu Leu Ala Met Ile Lys Ala Ile		
50	55	60
Arg Glu Gly Gly Lys Ala Leu Tyr Ile Val Pro Leu Arg Ala Leu Ala		
65	70	75
Ser Glu Lys Phe Glu Arg Phe Lys Glu Leu Ala Pro Phe Gly Ile Lys		
85	90	95
Val Gly Ile Ser Thr Gly Asp Leu Asp Ser Arg Ala Asp Trp Leu Gly		
100	105	110
Val Asn Asp Ile Ile Val Ala Thr Ser Glu Lys Thr Asp Ser Leu Leu		
115	120	125
Arg Asn Gly Thr Ser Trp Met Asp Glu Ile Thr Thr Val Val Val Asp		
130	135	140
Glu Ile His Leu Leu Asp Ser Lys Asn Arg Gly Pro Thr Leu Glu Val		
145	150	155
Thr Ile Thr Lys Leu Met Arg Leu Asn Pro Asp Val Gln Val Val Ala		
165	170	175
Leu Ser Ala Thr Val Gly Asn Ala Arg Glu Met Ala Asp Trp Leu Gly		
180	185	190
Ala Ala Leu Val Leu Ser Glu Trp Arg Pro Thr Asp Leu His Glu Gly		
195	200	205
Val Leu Phe Gly Asp Ala Ile Asn Phe Pro Gly Ser Gln Lys Lys Ile		
210	215	220
Asp Arg Leu Glu Lys Asp Asp Ala Val Asn Leu Val Leu Asp Thr Ile		
225	230	235
Lys Ala Glu Gly Gln Cys Leu Val Phe Glu Ser Ser Arg Arg Asn Cys		
245	250	255
Ala Gly Phe Ala Lys Thr Ala Ser Ser Lys Val Ala Lys Ile Leu Asp		
260	265	270
Asn Asp Ile Met Ile Lys Leu Ala Gly Ile Ala Glu Glu Val Glu Ser		
275	280	285
Thr Gly Glu Thr Asp Thr Ala Ile Val Leu Ala Asn Cys Ile Arg Lys		
290	295	300
Gly Val Ala Phe His His Ala Gly Leu Asn Ser Asn His Arg Lys Leu		
305	310	315
Val Glu Asn Gly Phe Arg Gln Asn Leu Ile Lys Val Ile Ser Ser Thr		
325	330	335

Pro Thr Leu Ala Ala Gly Leu Asn Leu Pro Ala Arg Arg Val Ile Ile		
340	345	350
Arg Ser Tyr Arg Arg Phe Asp Ser Asn Phe Gly Met Gln Pro Ile Pro		
355	360	365
Val Leu Glu Tyr Lys Gln Met Ala Gly Arg Ala Gly Arg Pro His Leu		
370	375	380
Asp Pro Tyr Gly Glu Ser Val Leu Leu Ala Lys Thr Tyr Asp Glu Phe		
385	390	395
Ala Gln Leu Met Glu Asn Tyr Val Glu Ala Asp Ala Glu Asp Ile Trp		
405	410	415
Ser Lys Leu Gly Thr Glu Asn Ala Leu Arg Thr His Val Leu Ser Thr		
420	425	430
Ile Val Asn Gly Phe Ala Ser Thr Arg Gln Glu Leu Phe Asp Phe Phe		
435	440	445
Gly Ala Thr Phe Phe Ala Tyr Gln Gln Asp Lys Trp Met Leu Glu Glu		
450	455	460
Val Ile Asn Asp Cys Leu Glu Phe Leu Ile Asp Lys Ala Met Val Ser		
465	470	475
Glu Thr Glu Asp Ile Glu Asp Ala Ser Lys Leu Phe Leu Arg Gly Thr		
485	490	495
Arg Leu Gly Ser Leu Val Ser Met Leu Tyr Ile Asp Pro Leu Ser Gly		
500	505	510
Ser Lys Ile Val Asp Gly Phe Lys Asp Ile Gly Lys Ser Thr Gly Gly		
515	520	525
Asn Met Gly Ser Leu Glu Asp Asp Lys Gly Asp Asp Ile Thr Val Thr		
530	535	540
Asp Met Thr Leu Leu His Leu Val Cys Ser Thr Pro Asp Met Arg Gln		
545	550	555
Leu Tyr Leu Arg Asn Thr Asp Tyr Thr Ile Val Asn Glu Tyr Ile Val		
565	570	575
Ala His Ser Asp Glu Phe His Glu Ile Pro Asp Lys Leu Lys Glu Thr		
580	585	590
Asp Tyr Glu Trp Phe Met Gly Glu Val Lys Thr Ala Met Leu Leu Glu		
595	600	605
Glu Trp Val Thr Glu Val Ser Ala Glu Asp Ile Thr Arg His Phe Asn		
610	615	620
Val Gly Glu Gly Asp Ile His Ala Leu Ala Asp Thr Ser Glu Trp Leu		
625	630	635
Met His Ala Ala Ala Lys Leu Ala Glu Leu Leu Gly Val Glu Tyr Ser		

	645		650		655
Ser His Ala Tyr Ser Leu Glu Lys Arg Ile Arg Tyr Gly Ser Gly Leu					
	660		665		670
Asp Leu Met Glu Leu Val Gly Ile Arg Gly Val Gly Arg Val Arg Ala					
	675		680		685
Arg Lys Leu Tyr Asn Ala Gly Phe Val Ser Val Ala Lys Leu Lys Gly					
	690		695		700
Ala Asp Ile Ser Val Leu Ser Lys Leu Val Gly Pro Lys Val Ala Tyr					
705		710		715	720
Asn Ile Leu Ser Gly Ile Gly Val Arg Val Asn Asp Lys His Phe Asn					
	725		730		735
Ser Ala Pro Ile Ser Ser Asn Thr Leu Asp Thr Leu Leu Asp Lys Asn					
	740		745		750
Gln Lys Thr Phe Asn Asp Phe Gln					
	755		760		
<210> 8					
<211> 439					
<212> PRT					
<213> 人工序列					
<220>					
<223> Dda解旋酶					
<400> 8					
Met Thr Phe Asp Asp Leu Thr Glu Gly Gln Lys Asn Ala Phe Asn Ile					
1	5		10		15
Val Met Lys Ala Ile Lys Glu Lys Lys His His Val Thr Ile Asn Gly					
	20		25		30
Pro Ala Gly Thr Gly Lys Thr Thr Leu Thr Lys Phe Ile Ile Glu Ala					
	35		40		45
Leu Ile Ser Thr Gly Glu Thr Gly Ile Ile Leu Ala Ala Pro Thr His					
	50		55		60
Ala Ala Lys Lys Ile Leu Ser Lys Leu Ser Gly Lys Glu Ala Ser Thr					
65		70		75	80
Ile His Ser Ile Leu Lys Ile Asn Pro Val Thr Tyr Glu Glu Asn Val					
	85		90		95
Leu Phe Glu Gln Lys Glu Val Pro Asp Leu Ala Lys Cys Arg Val Leu					
	100		105		110
Ile Cys Asp Glu Val Ser Met Tyr Asp Arg Lys Leu Phe Lys Ile Leu					
	115		120		125
Leu Ser Thr Ile Pro Pro Trp Cys Thr Ile Ile Gly Ile Gly Asp Asn					

130	135	140
Lys Gln Ile Arg Pro Val Asp Pro Gly Glu Asn Thr Ala Tyr Ile Ser		
145	150	155
Pro Phe Phe Thr His Lys Asp Phe Tyr Gln Cys Glu Leu Thr Glu Val		
165	170	175
Lys Arg Ser Asn Ala Pro Ile Ile Asp Val Ala Thr Asp Val Arg Asn		
180	185	190
Gly Lys Trp Ile Tyr Asp Lys Val Val Asp Gly His Gly Val Arg Gly		
195	200	205
Phe Thr Gly Asp Thr Ala Leu Arg Asp Phe Met Val Asn Tyr Phe Ser		
210	215	220
Ile Val Lys Ser Leu Asp Asp Leu Phe Glu Asn Arg Val Met Ala Phe		
225	230	235
Thr Asn Lys Ser Val Asp Lys Leu Asn Ser Ile Ile Arg Lys Lys Ile		
245	250	255
Phe Glu Thr Asp Lys Asp Phe Ile Val Gly Glu Ile Ile Val Met Gln		
260	265	270
Glu Pro Leu Phe Lys Thr Tyr Lys Ile Asp Gly Lys Pro Val Ser Glu		
275	280	285
Ile Ile Phe Asn Asn Gly Gln Leu Val Arg Ile Ile Glu Ala Glu Tyr		
290	295	300
Thr Ser Thr Phe Val Lys Ala Arg Gly Val Pro Gly Glu Tyr Leu Ile		
305	310	315
Arg His Trp Asp Leu Thr Val Glu Thr Tyr Gly Asp Asp Glu Tyr Tyr		
325	330	335
Arg Glu Lys Ile Lys Ile Ile Ser Ser Asp Glu Glu Leu Tyr Lys Phe		
340	345	350
Asn Leu Phe Leu Gly Lys Thr Ala Glu Thr Tyr Lys Asn Trp Asn Lys		
355	360	365
Gly Gly Lys Ala Pro Trp Ser Asp Phe Trp Asp Ala Lys Ser Gln Phe		
370	375	380
Ser Lys Val Lys Ala Leu Pro Ala Ser Thr Phe His Lys Ala Gln Gly		
385	390	395
Met Ser Val Asp Arg Ala Phe Ile Tyr Thr Pro Cys Ile His Tyr Ala		
405	410	415
Asp Val Glu Leu Ala Gln Gln Leu Leu Tyr Val Gly Val Thr Arg Gly		
420	425	430
Arg Tyr Asp Val Phe Tyr Val		
435		

<210> 9
<211> 35
<212> DNA
<213> 人工序列
<220>
<223> 条形码1
<400> 9
tgctactctc ctcataagca gtccggtgta tcgat 35
<210> 10
<211> 35
<212> DNA
<213> 人工序列
<220>
<223> 条形码2
<400> 10
atcgctacgc cttcggctcg taatcatagt cgagt 35
<210> 11
<211> 35
<212> DNA
<213> 人工序列
<220>
<223> 条形码3
<400> 11
agctcagagc aggtcactca agatacgagc tgcgt 35
<210> 12
<211> 35
<212> DNA
<213> 人工序列
<220>
<223> 条形码4
<400> 12
gtaagtctgc atcagcgcgcg ggctgtgcga ggata 35
<210> 13
<211> 35
<212> DNA
<213> 人工序列
<220>
<223> 条形码5
<400> 13

ctacgacagt acgctagcaa ggatagacac tacga 35

<210> 14

<211> 35

<212> DNA

<213> 人工序列

<220>

<223> 条形码6

<400> 14

tactgaacac aagttcgtcg tcgagcaatc acaat 35

<210> 15

<211> 35

<212> DNA

<213> 人工序列

<220>

<223> 条形码7

<400> 15

agtctaccat tacttggatc ggattagcct cactc 35

<210> 16

<211> 35

<212> DNA

<213> 人工序列

<220>

<223> 条形码8

<400> 16

tgcacgagtg cgtgtcaacc gtccagatgc tcgtg 35

<210> 17

<211> 35

<212> DNA

<213> 人工序列

<220>

<223> 条形码9

<400> 17

ctagtgcgca gttgtctcgg cggagttgag actga 35

<210> 18

<211> 35

<212> DNA

<213> 人工序列

<220>

<223> 条形码10

<400> 18
gatcatggta gtcttcaaga tcgagtatgt ctgtc 35
<210> 19
<211> 21
<212> RNA
<213> 人工序列
<220>
<223> miRNA 1
<400> 19
cagcagcaca cugugguuug u 21
<210> 20
<211> 22
<212> RNA
<213> 人工序列
<220>
<223> miRNA 2
<400> 20
agagcuuagc uguuugguga ac 22
<210> 21
<211> 21
<212> RNA
<213> 人工序列
<220>
<223> miRNA 3
<400> 21
uagcuuauca gacugauguu g 21
<210> 22
<211> 22
<212> RNA
<213> 人工序列
<220>
<223> miRNA 4
<400> 22
accuggcaua caauguagau uu 22
<210> 23
<211> 22
<212> RNA
<213> 人工序列
<220>

<223> miRNA 5
<400> 23
uguaaacauc cccgacugga ag 22
<210> 24
<211> 23
<212> RNA
<213> 人工序列
<220>
<223> miRNA 6
<400> 24
uguaaacauc cuacacucuc agc 23
<210> 25
<211> 22
<212> RNA
<213> 人工序列
<220>
<223> miRNA 7
<400> 25
agcugguaaa auggaaccaa au 22
<210> 26
<211> 22
<212> RNA
<213> 人工序列
<220>
<223> miRNA 8
<400> 26
gagcuuuugg cccggguuau ac 22
<210> 27
<211> 23
<212> RNA
<213> 人工序列
<220>
<223> miRNA 9
<400> 27
aacaaucauu gcugucggug ggu 23
<210> 28
<211> 22
<212> RNA
<213> 人工序列

<220>
<223> miRNA 10
<400> 28
uagcaccauc ugaaaucggu ua 22
<210> 29
<211> 67
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 29
cctagttccg ctggggagct agcatacgtg tctaactg cacagatgat nnggttggtg 60
tggttgg 67
<210> 30
<211> 67
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 30
cctagttccg ctgggatata ggcataaagt aaaatcgta ccaactcaatc nnggttggtg 60
tggttgg 67
<210> 31
<211> 67
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)

<223> 2 x iSpC3
<400> 31
cctagttccg ctgggctcgg cgttgtgtgt caaatggcgt agatctggat nnggttggtg 60
tggttg 67
<210> 32
<211> 66
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (36) .. (38)
<223> 2 x iSp18
<400> 32
cctagttccg ctgggcacag ccccatgtaa cccatnnagt ccgtggtagg gcaggttggg 60
gtgact 66
<210> 33
<211> 87
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 33
cctagttccg ctgggactag ataaaaggaa gggagcacag taacgtcggt nncgactggt 60
aggcagatag gggaagctga ttcgatg 87
<210> 34
<211> 96
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)

<223> 2 x iSpC3
<400> 34
cctagttccg ctgggtctcc cgtatccgtg gctaaacgcc ttcaatctta nnggatattg 60
cgcgattccg gtcggcagct taggaagtgc ggtgtc 96
<210> 35
<211> 97
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (41) .. (42)
<223> 2 x iSpC3
<400> 35
cctagttccg ctgggacctt ttctaggatg gaacattcta nnacgcataa cgtcttgcgg 60
ggcggcgggt tgagaggatg tcgggtggtt atgcgtg 97
<210> 36
<211> 74
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (52) .. (53)
<223> 2 x iSpC3
<400> 36
cctagttccg ctgggtgcta ctcttctca taagcagtc ggtgtatcga tnnacaaacc 60
acagtgtgct gctg 74
<210> 37
<211> 74
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)

<223> 2 x iSpC3
<400> 37
cctagttccg ctgggatcgc tacgccttcg gctcgtaatc atagtcgagt nngttcacca 60
atcagctaag ctct 74
<210> 38
<211> 74
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 38
cctagttccg ctgggagctc agagcaggtc actcaagata cgagctgcgt nntcaacatc 60
agtctgataa gcta 74
<210> 39
<211> 74
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 39
cctagttccg ctggggtaag tctgcatcag cgcgcggctg tgcgaggata nnaaatctac 60
attgtatgcc aggt 74
<210> 40
<211> 74
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)

<223> 2 x iSpC3
<400> 40
cctagttccg ctgggctacg acagtacgct agcaaggata gacactacga ncttccagt 60
cggggatggtt taca 74
<210> 41
<211> 75
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 41
cctagttccg ctgggtactg aacacaagtt cgtcgtcgag caatcacaat nngctgagag 60
tgtaggatgt ttaca 75
<210> 42
<211> 74
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 42
cctagttccg ctgggagtct accattactt ggatcggatt agcctcactc nnatttggtt 60
ccattttacc agct 74
<210> 43
<211> 74
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)

<223> 2 x iSpC3
<400> 43
cctagttccg ctgggtgcac gagtgcgtgt caaccgtcca gatgctcgtg nngtataacc 60
cgggccaaaaa gctc 74
<210> 44
<211> 75
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 44
cctagttccg ctgggctagt gcgcagttgt ctcggcggag ttgagactga nnaccaccg 60
acagcaatga atgtt 75
<210> 45
<211> 74
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 45
cctagttccg ctggggatca tggtagtctt caagatcgag tatgtctgtc nntaaccgat 60
ttcagatggt gcta 74
<210> 46
<211> 73
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 46
cctagttccg ctgggtactg aacacaagtt cgtcgtcgag caatcacaat gctgagagtg 60
taggatgttt aca 73

<210> 47
<211> 72
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 47
cctagttccg ctgggcgtga agagagtttc ataatacgtc cagccgcatg taaccgattt 60
cagatgggtgc ta 72
<210> 48
<211> 86
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (14) .. (14)
<223> 间隔子
<400> 48
tcgatacaat acanattccgt cacacctgct ctaggggatc aaagctatgc gaccatgcga 60
gtggatactg gtgttggctc ccgtat 86
<210> 49
<211> 76
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 49
agtctccgct gtcctcccga tgcacttgac gtatgtetca ctttcttttc attgacatgg 60
gatgacgccg tgactg 76
<210> 50
<211> 74
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature

<222> (51) .. (52)
<223> 2 x iSpC3
<400> 50
cctagttccg ctgggatcgc tacgccttcg gctcgtaatc atagtcgagt nngttcacca 60
atcagctaag ctct 74
<210> 51
<211> 74
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 51
cctagttccg ctgggagctc agagcaggtc actcaagata cgagctgcgt nntcaacatc 60
agtctgataa gcta 74
<210> 52
<211> 74
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 52
cctagttccg ctggggtaag tctgcatcag cgcgcggctg tgcgaggata nnaaatctac 60
attgtatgcc aggt 74
<210> 53
<211> 74
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature

<222> (51) .. (52)
<223> 2 x iSpC3
<400> 53
cctagttccg ctgggctacg acagtacgct agcaaggata gacactacga ncttccagt 60
cggggatggt taca 74
<210> 54
<211> 75
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 54
cctagttccg ctgggtactg aacacaagtt cgtcgtcgag caatcacaat nngctgagag 60
tgtaggatgt ttaca 75
<210> 55
<211> 75
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 55
cctagttccg ctgggctagt gcgcagttgt ctcggcggag ttgagactga nnaccaccg 60
acagcaatga atggt 75
<210> 56
<211> 74
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature

<222> (51) .. (52)
<223> 2 x iSpC3
<400> 56
cctagttccg ctggggatca tggtagtctt caagatcgag tatgtctgtc nntaaccgat 60
ttcagatggg gcta 74
<210> 57
<211> 74
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 57
cctagttccg ctgggggttca catcaaggtc ataccgcgag ttctatttta nncagtgtgc 60
ggtgggcagg ggct 74
<210> 58
<211> 73
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 58
cctagttccg ctggggcttg ggggatagat gtgccccgcg catcggacct nncgcgtacc 60
aaaagtaata atg 73
<210> 59
<211> 74
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature

<222> (51) .. (52)
<223> 2 x iSpC3
<400> 59
cctagttccg ctgggatgac acacgttttc gatagggacg ccgactttaa nntggacgtt 60
tgcaggggag gtgg 74
<210> 60
<211> 73
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 60
cctagttccg ctgggtgata ataagacctg acagacaata gggagaactc nncgcgtacc 60
aaaagtaata atg 73
<210> 61
<211> 76
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 61
cctagttccg ctgggttagt aatcaagtct gatcgtaata gctaagtcatt nactgcagg 60
ctccagcttc caggct 76
<210> 62
<211> 74
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature

<222> (51) .. (52)
<223> 2 x iSp3C
<400> 62
cctagttccg ctgggcctgt tctaattgcg cggagaggcg agatgtttct ncttccagt 60
caaggatggt taca 74
<210> 63
<211> 75
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 63
cctagttccg ctgggtggga tactgacgtg ccggaatacc cagacgtgcc nctacctgc 60
actgtaagca ctttt 75
<210> 64
<211> 74
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 64
cctagttccg ctgggtgate ggcacctaaa cgcattagcc ctgcaatacg ntaaccaat 60
gtgcagacta ctgt 74
<210> 65
<211> 73
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature

<222> (51) .. (52)
<223> 2 x iSpC3
<400> 65
cctagttccg ctgggagctg ctcggaagcc ataaggtact ttaatttggg nncacaaccc 60
tagtggcgcc att 73
<210> 66
<211> 73
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 66
cctagttccg ctgggcacgg atttctatat tgctcaacca ggcagcgcaa nnacctatcc 60
tgaattactt gaa 73
<210> 67
<211> 75
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 67
cctagttccg ctgggattgt gcgcttttct ccatgcgtct tagacatctc nnagggttc 60
ctgggaaaac tggac 75
<210> 68
<211> 74
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature

<222> (51) .. (52)
<223> 2 x iSpC3
<400> 68
cctagttccg ctggggtgaa atccccgtct aggttatggc tggggggatt nnacaggccg 60
ggacaagtgc aata 74
<210> 69
<211> 74
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 69
cctagttccg ctgggcgtaa acttatcacg acacaatgaa caagcctgca nnaacccatg 60
gaattcagtt ctca 74
<210> 70
<211> 75
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 70
cctagttccg ctgggggcag tgtccgagcg tcttcaatca tgagcgattc nnaaagtctc 60
gctctctgcc cctca 75
<210> 71
<211> 73
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature

<222> (51) .. (52)
<223> 2 x iSpC3
<400> 71
cctagttccg ctgggctcc agacttaatg tctgctcact gacatcgca nngcagaact 60
tagccactgt gaa 73
<210> 72
<211> 74
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 72
cctagttccg ctgggattag cggaaccaa cccaggaagg cttgaaggcg nnatacatac 60
ttctttacat tcca 74
<210> 73
<211> 75
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 73
cctagttccg ctgggaacct taggggctc gaatctttga gacgactagg nctatctgc 60
actagatgca cctta 75
<210> 74
<211> 75
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature

<222> (51) .. (52)
<223> 2 x iSpC3
<400> 74
cctagttccg ctgggtaatt actgccccac catgacattt taatagcagt nnctaactgc 60
actagatgca cctta 75
<210> 75
<211> 74
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 75
cctagttccg ctggggacct tgagacagaa cttatcaatg tacaactgaa nnagtagtgc 60
aataaagtca gagc 74
<210> 76
<211> 74
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 76
cctagttccg ctgggcgaag gattggcccc ccgattacca ccgccgtgag nnaaccatac 60
aacctactac ctca 74
<210> 77
<211> 76
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature

<222> (51) .. (52)

<223> 2 x iSpC3

<400> 77

cctagttccg ctgggaattg ccaacaggtc aagccctgtt ctcaactggc nntcacaggt 60
taaagggtct caggga 76

<210> 78

<211> 73

<212> DNA

<213> 人工序列

<220>

<223> 合成构建体

<220>

<221> misc_feature

<222> (51) .. (52)

<223> 2 x iSpC3

<400> 78

cctagttccg ctgggcagtg cttgccccca gtagagtgtg gaaggcata nnaggaatat 60
gtttgatata tag 73

<210> 79

<211> 74

<212> DNA

<213> 人工序列

<220>

<223> 合成构建体

<220>

<221> misc_feature

<222> (51) .. (52)

<223> 2 x iSpC3

<400> 79

cctagttccg ctgggaaagc ccactctcca cacttcaagg ttaaattggcg nnagcgggac 60
tttgagggcc agtt 74

<210> 80

<211> 74

<212> DNA

<213> 人工序列

<220>

<223> 合成构建体

<220>

<221> misc_feature

<222> (51) .. (52)
<223> 2 x iSpC3
<400> 80
cctagttccg ctggggacta atacaatcgg aagcaactct cacgccgcac nntcatctcg 60
cccgcaaaga ccca 74
<210> 81
<211> 74
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 81
cctagttccg ctgggctacg attcatgtct cccccacat atgattgac nnaggcgaag 60
gatgacaaag ggaa 74
<210> 82
<211> 74
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 82
cctagttccg ctgggtcggg cgcttaatcg caatgttcat ccggaacgga nntcatctaa 60
taaacattta ctga 74
<210> 83
<211> 75
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature

<222> (51) .. (52)
<223> 2 x iSpC3
<400> 83
cctagttccg ctgggtaaatt tccacattac gatagctgac gtcctggtag nngggctctt 60
actccctcag gcact 75
<210> 84
<211> 77
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 84
cctagttccg ctgggcaccc cctagcctga ccttaaatga taattcctgt nnaggccgcc 60
acccgcccgc gatccct 77
<210> 85
<211> 75
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 85
cctagttccg ctgggttttag cctaacctct gatgcattgc accggagttc nntccagcc 60
cctccagggc ttcct 75
<210> 86
<211> 74
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature

<222> (51) .. (52)
<223> 2 x iSpC3
<400> 86
cctagttccg ctgggcccac tatacgtga acctccatc cgattttcag nntgccgtgc 60
cctggcctcc cact 74
<210> 87
<211> 74
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 87
cctagttccg ctgggccttg gaattagaac cgtgtgattc tacgcctagg nncaccactg 60
cactcgtcc agcc 74
<210> 88
<211> 74
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 88
cctagttccg ctgggaggac ttcttcggtg taatcggagt acatcaatgt nnaaacagtc 60
tacctacatt gctc 74
<210> 89
<211> 73
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature

<222> (51) .. (52)
<223> 2 x iSpC3
<400> 89
cctagttccg ctgggaacgt actttgtggg gataagctgt acagggctgt nngggaggct 60
cctttgagaa ttg 73
<210> 90
<211> 128
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 90
cctagttccg ctgggattca gatctttgac cgatcgtact gacgtgtacg nnagtctccg 60
ctgtcctccc gatgcacttg acgtatgtct cactttcttt tcattgacat gggatgacgc 120
cgtgactg 128
<210> 91
<211> 123
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 91
cctagttccg ctgggcttaa gttcgggtcat taacagtgtc aatcttgcca nnatacggga 60
gccaacacca ggactaacat tataagaatt gcgaataatc attggagagc aggtgtgacg 120
gat 123
<210> 92
<211> 92
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体

<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 92
cctagttccg ctgggatgta attcctgcta ggtccaatgt gactgtacta nnggcgattc 60
gtgatctctg ctctcggttt cgcgttcggt cg 92
<210> 93
<211> 81
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 93
cctagttccg ctggggttaag ctattgctaa ctgtagtcct agtctagcta nnagtccgtg 60
gtagggcagg ttgggggtgac t 81
<210> 94
<211> 67
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 94
cctagttccg ctggggagct agcatacgtg tctaactg cacagatgat nnggttggtg 60
tggttg 67
<210> 95
<211> 98
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体

<220>
<221> misc_feature
<222> (41) .. (42)
<223> 2 x iSpC3
<400> 95
cctagttccg ctgggacctt ttctaggatg gaacattcta nncacgcata acgtcttgcg 60
gggcggcggg ttgagaggat gtggggtggt tatgcgtg 98
<210> 96
<211> 100
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (41) .. (42)
<223> 2 x iSpC3
<400> 96
cctagttccg ctgggctgcg acatgactat ctagatcgc nngctggatg tcaccagatt 60
gtcggacatc ggattgtctg agtcatatga cacatccagc 100
<210> 97
<211> 87
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (51) .. (52)
<223> 2 x iSpC3
<400> 97
cctagttccg ctgggactag ataaaaggaa gggagcacag taacgtcggt nncgactggt 60
aggcagatag gggaagctga ttcgatg 87
<210> 98
<211> 102
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体

<220>
<221> misc_feature
<222> (29) .. (30)
<223> 2 x iSpC3
<400> 98
cctagttccg ctgggtcgat acaatacann atccgtcaca cctgctctag gggatcaaag 60
ctatgcgacc atgcgagtgg atactggtgt tggctcccgt at 102
<210> 99
<211> 21
<212> RNA
<213> 人工序列
<220>
<223> miRNA 11
<400> 99
cagcagcaca cugugguuug u 21
<210> 100
<211> 22
<212> RNA
<213> 人工序列
<220>
<223> miRNA 12
<400> 100
agagcuuagc ugauugguga ac 22
<210> 101
<211> 22
<212> RNA
<213> 人工序列
<220>
<223> miRNA 13
<400> 101
uagcuuauca gacugauguu ga 22
<210> 102
<211> 22
<212> RNA
<213> 人工序列
<220>
<223> miRNA 14
<400> 102
accuggcaua caauguagau uu 22

<210> 103
<211> 22
<212> RNA
<213> 人工序列
<220>
<223> miRNA 15
<400> 103
uguaaacauc cccgacugga ag 22
<210> 104
<211> 23
<212> RNA
<213> 人工序列
<220>
<223> miRNA 16
<400> 104
uguaaacauc cuacacucuc agc 23
<210> 105
<211> 22
<212> RNA
<213> 人工序列
<220>
<223> miRNA 17
<400> 105
agcugguaaa auggaaccaa au 22
<210> 106
<211> 22
<212> RNA
<213> 人工序列
<220>
<223> miRNA 18
<400> 106
gagcuuuugg cccggguuau ac 22
<210> 107
<211> 23
<212> RNA
<213> 人工序列
<220>
<223> miRNA 19
<400> 107

aacauucauu gcugucggug ggu 23

<210> 108

<211> 22

<212> RNA

<213> 人工序列

<220>

<223> miRNA 20

<400> 108

uagcaccauc ugaaaucggu ua 22

<210> 109

<211> 22

<212> RNA

<213> 人工序列

<220>

<223> miRNA 22

<400> 109

agccccugcc caccgcacac ug 22

<210> 110

<211> 21

<212> RNA

<213> 人工序列

<220>

<223> miRNA 24

<400> 110

cauuauuacu uuugguacgc g 21

<210> 111

<211> 22

<212> RNA

<213> 人工序列

<220>

<223> miRNA 25

<400> 111

ccaccucccc ugcaaacguc ca 22

<210> 112

<211> 21

<212> RNA

<213> 人工序列

<220>

<223> miRNA 26

<400> 112
cauuauuacu uuugguacgc g 21
<210> 113
<211> 24
<212> RNA
<213> 人工序列
<220>
<223> miRNA 27
<400> 113
agccuggaag cuggagccug cagu 24
<210> 114
<211> 22
<212> RNA
<213> 人工序列
<220>
<223> miRNA 29
<400> 114
uguaaacauc cuugacugga ag 22
<210> 115
<211> 23
<212> RNA
<213> 人工序列
<220>
<223> miRNA 30
<400> 115
aaaagugcuu acagugcagg uag 23
<210> 116
<211> 22
<212> RNA
<213> 人工序列
<220>
<223> miRNA 31
<400> 116
acaguagucu gcacauuggu ua 22
<210> 117
<211> 21
<212> RNA
<213> 人工序列
<220>

<223> miRNA 32
<400> 117
aauggcgcca cuagguugu g 21
<210> 118
<211> 21
<212> RNA
<213> 人工序列
<220>
<223> miRNA 33
<400> 118
uucaaguaau ucaggauagg u 21
<210> 119
<211> 23
<212> RNA
<213> 人工序列
<220>
<223> miRNA 34
<400> 119
guccaguuuu cccaggaauc ccu 23
<210> 120
<211> 22
<212> RNA
<213> 人工序列
<220>
<223> miRNA 35
<400> 120
uauugcacuu gucccggccu gu 22
<210> 121
<211> 22
<212> RNA
<213> 人工序列
<220>
<223> miRNA 36
<400> 121
ugagaacuga auuccauggg uu 22
<210> 122
<211> 23
<212> RNA
<213> 人工序列

<220>
<223> miRNA 37
<400> 122
ugaggggcag agagcgagac uuu 23
<210> 123
<211> 21
<212> RNA
<213> 人工序列
<220>
<223> miRNA 39
<400> 123
uucacagugg cuaaguucug c 21
<210> 124
<211> 22
<212> RNA
<213> 人工序列
<220>
<223> miRNA 40
<400> 124
uggaauguaa agaaguaugu au 22
<210> 125
<211> 23
<212> RNA
<213> 人工序列
<220>
<223> miRNA 42
<400> 125
uaaggugcau cuagugcaga uag 23
<210> 126
<211> 23
<212> RNA
<213> 人工序列
<220>
<223> miRNA 43
<400> 126
uaaggugcau cuagugcagu uag 23
<210> 127
<211> 22
<212> RNA

<213> 人工序列
<220>
<223> miRNA 45
<400> 127
gcucugacuu uauugcacua cu 22
<210> 128
<211> 22
<212> RNA
<213> 人工序列
<220>
<223> miRNA 46
<400> 128
ugagguagua gguuguaugg uu 22
<210> 129
<211> 24
<212> RNA
<213> 人工序列
<220>
<223> miRNA 47
<400> 129
ucccugagac ccuuuaaccu guga 24
<210> 130
<211> 21
<212> RNA
<213> 人工序列
<220>
<223> miRNA 49
<400> 130
cuauauauca aacauauucc u 21
<210> 131
<211> 22
<212> RNA
<213> 人工序列
<220>
<223> miRNA 50
<400> 131
aacuggcccu caaagucccg cu 22
<210> 132
<211> 22

<212> RNA
<213> 人工序列
<220>
<223> miRNA 51
<400> 132
ugggucuuug cgggcgagau ga 22
<210> 133
<211> 22
<212> RNA
<213> 人工序列
<220>
<223> miRNA 52
<400> 133
uuccuuugu cauccuucgc cu 22
<210> 134
<211> 22
<212> RNA
<213> 人工序列
<220>
<223> miRNA 53
<400> 134
ucaguaaaug uuuaauagau ga 22
<210> 135
<211> 23
<212> RNA
<213> 人工序列
<220>
<223> miRNA 54
<400> 135
agugccugag ggaguaagag ccc 23
<210> 136
<211> 25
<212> RNA
<213> 人工序列
<220>
<223> miRNA 55
<400> 136
agggaucgcg ggcggguggc ggccu 25
<210> 137

<211> 23
<212> RNA
<213> 人工序列
<220>
<223> miRNA 56
<400> 137
aggaagcccu ggaggggcug gag 23
<210> 138
<211> 22
<212> RNA
<213> 人工序列
<220>
<223> miRNA 57
<400> 138
agugggaggc cagggcacgg ca 22
<210> 139
<211> 22
<212> RNA
<213> 人工序列
<220>
<223> miRNA 58
<400> 139
ggcuggagcg agugcagugg ug 22
<210> 140
<211> 22
<212> RNA
<213> 人工序列
<220>
<223> miRNA 59
<400> 140
gagcaaugua gguagacugu uu 22
<210> 141
<211> 21
<212> RNA
<213> 人工序列
<220>
<223> miRNA 60
<400> 141
caauucuaa aggagccucc c 21

<210> 142
<211> 59
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 142
tttttttttt tttttttttt tttttttttt agtccgtggt agggcagggtt ggggtgact 59
<210> 143
<211> 74
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 143
tttttttttt tttttttttt tttttttttt cgactggttag gcagataggg gaagctgatt 60
cgatgcgtgg gtcg 74
<210> 144
<211> 39
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 144
gggtgctact cttcctcata agcagtccgg tgtatcgat 39
<210> 145
<211> 38
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 145
gggatcgcta cgccttcggc tcgtaatcat agtcgagt 38
<210> 146
<211> 38
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体

<400> 146
gggagctcag agcaggtcac tcaagatacg agctgcgt 38
<210> 147
<211> 38
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 147
ggggtaagtc tgcatacagc cgcggtgtg cgaggata 38
<210> 148
<211> 38
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 148
gggctacgac agtacgctag caaggataga cactacga 38
<210> 149
<211> 38
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 149
gggtactgaa cacaagttcg tcgtcgagca atcacaat 38
<210> 150
<211> 38
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 150
gggagtctac cattacttgg atcggattag cctcactc 38
<210> 151
<211> 38
<212> DNA
<213> 人工序列
<220>

<223> 合成构建体
<400> 151
gggtgcacga gtgcgtgtca accgtccaga tgctcgtg 38
<210> 152
<211> 38
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 152
gggctagtgc gcagttgtct cggcggagtt gagactga 38
<210> 153
<211> 38
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 153
ggggatcatg gtagtcttca agatcgagta tgtctgtc 38
<210> 154
<211> 81
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 154
atattgtact tccgtttcga atttacgtat tgctcctagt tccgctggta ctgaacacaa 60
gttcgtcgtg caatacagtg t 81
<210> 155
<211> 72
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (36) .. (37)
<223> 2 x iSpC3
<400> 155

actagataaa aggaagggag cacagtaacg tcgttnncga ctggtaggca gataggggaa 60
gctgattcga tg 72
<210> 156
<211> 87
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> misc_feature
<222> (14) .. (15)
<223> 2 x iSpC3
<400> 156
tcgatacaat acannatccg tcacacctgc tctaggggat caaagctatg cgaccatgcg 60
agtggataact ggtgttggct cccgtat 87
<210> 157
<211> 16
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 157
cccagcggaa ctagga 16

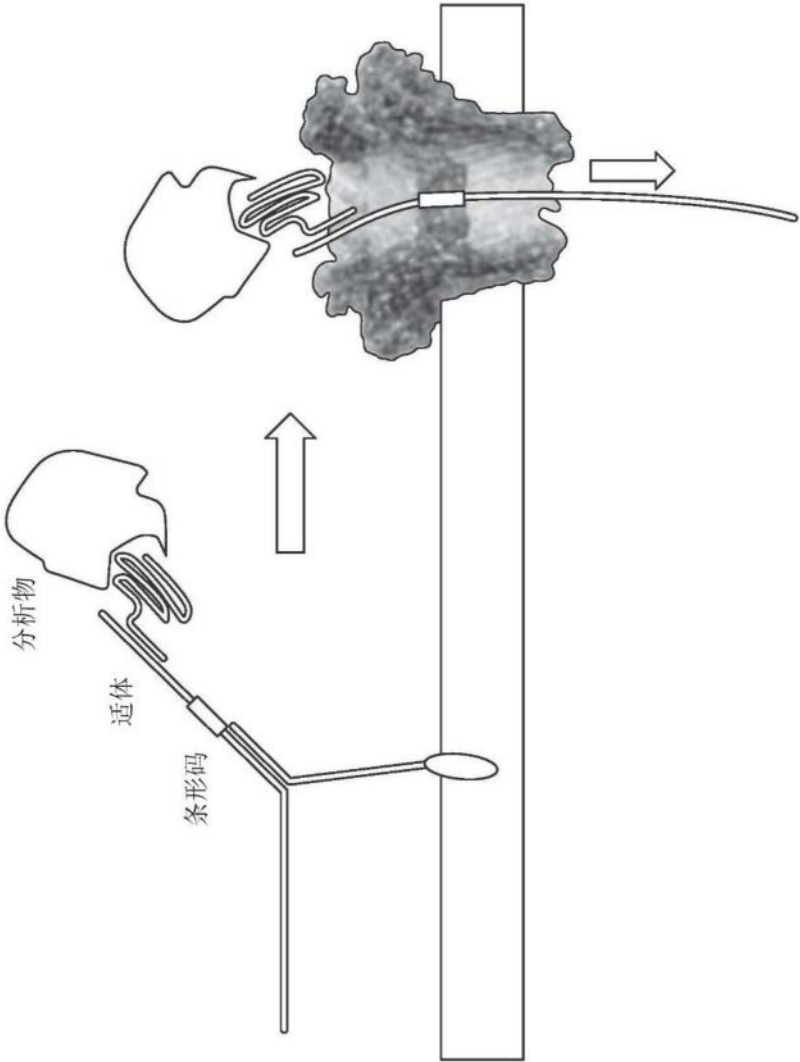


图1

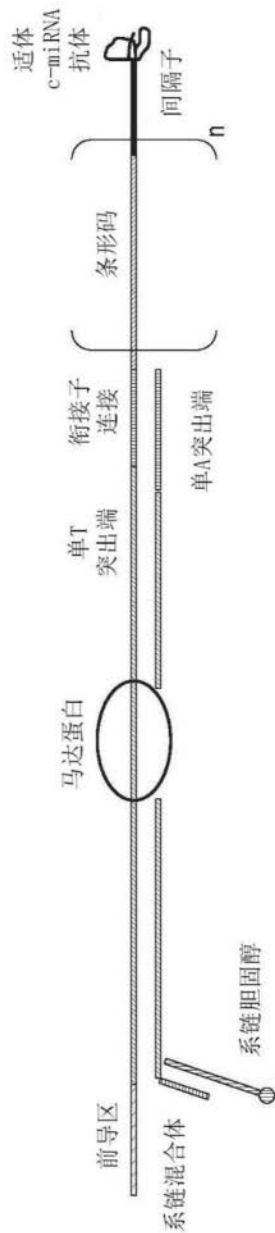


图2A

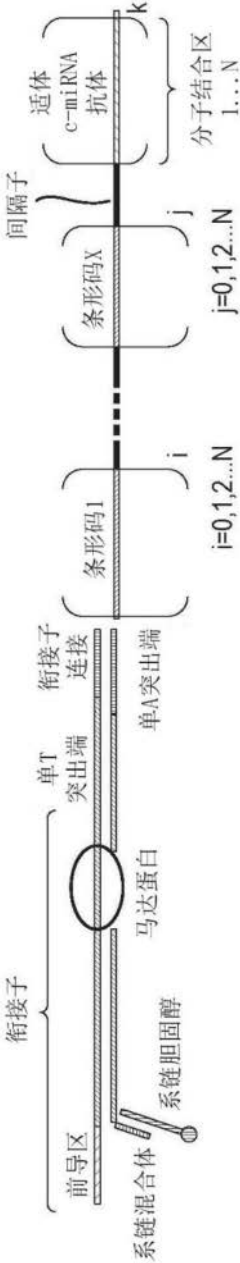


图2B

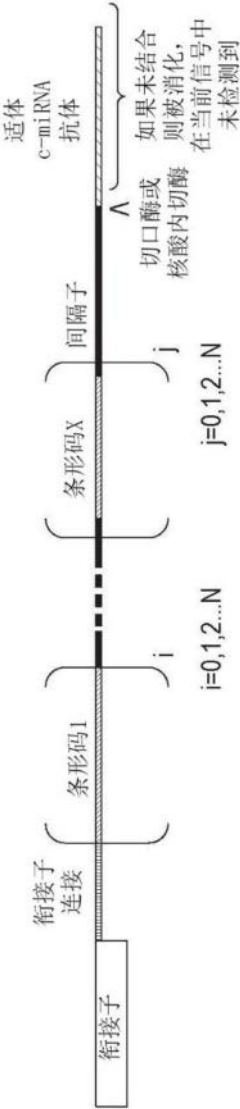


图3A

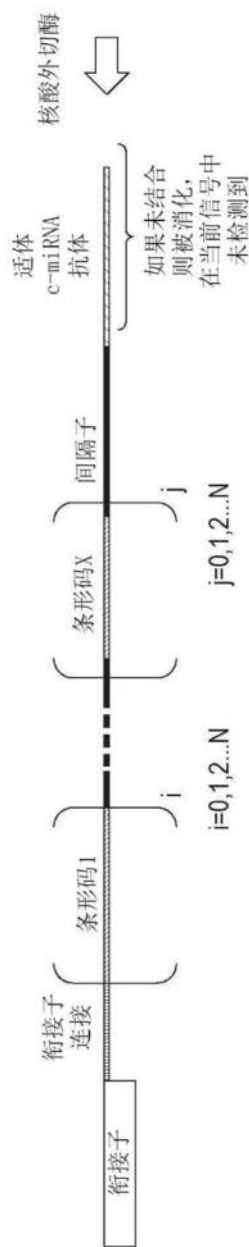


图3B

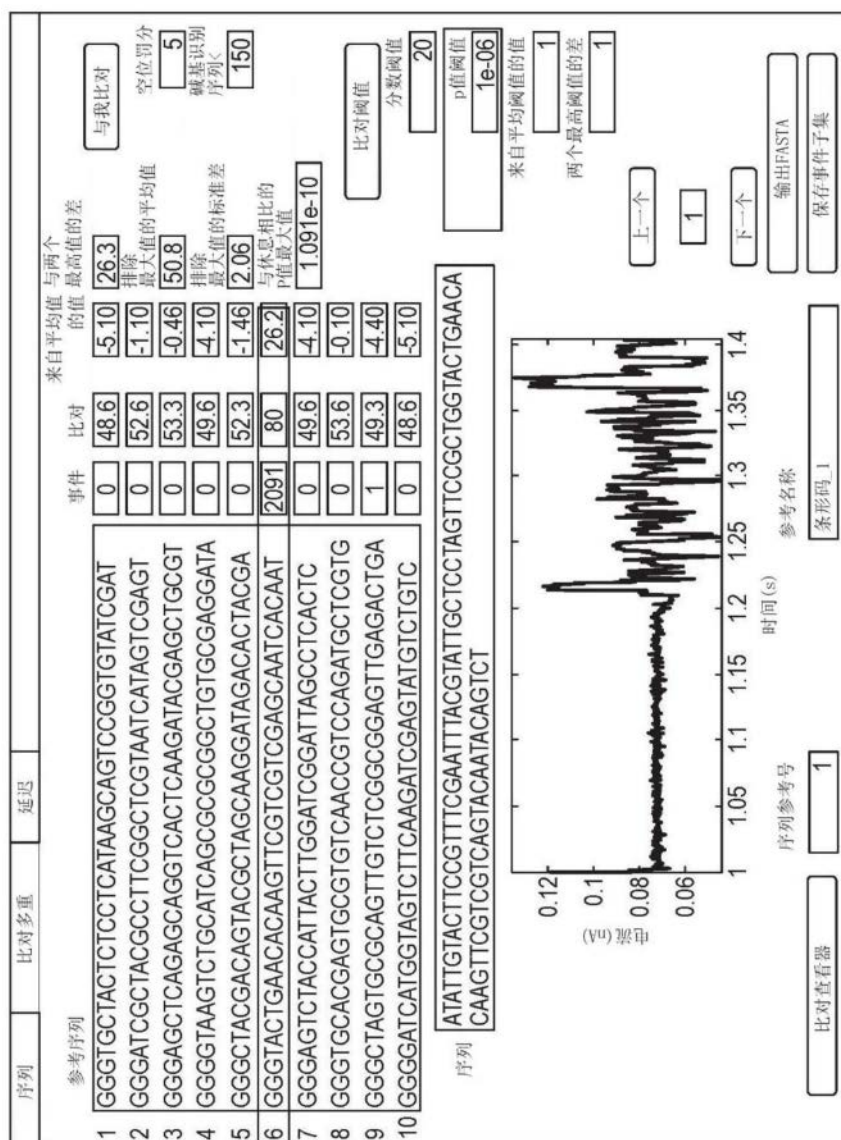


图4A

条形码1	1	1	3	3	0	0
条形码2	3	3	2	2	0	0
条形码3	5	5	3	0	0	0
条形码4	2	2	1	0	0	0
条形码5	3	3	2	0	0	0
条形码6	2407	2406	2404	2381	2319	2100
条形码7	0	0	0	0	0	0
条形码8	12	12	10	7	1	0
条形码9	4	4	3	2	1	1
条形码10	0	0	0	0	0	0
p值	0.1	0.01	0.001	0.0001	0.00001	0.000001

图4B

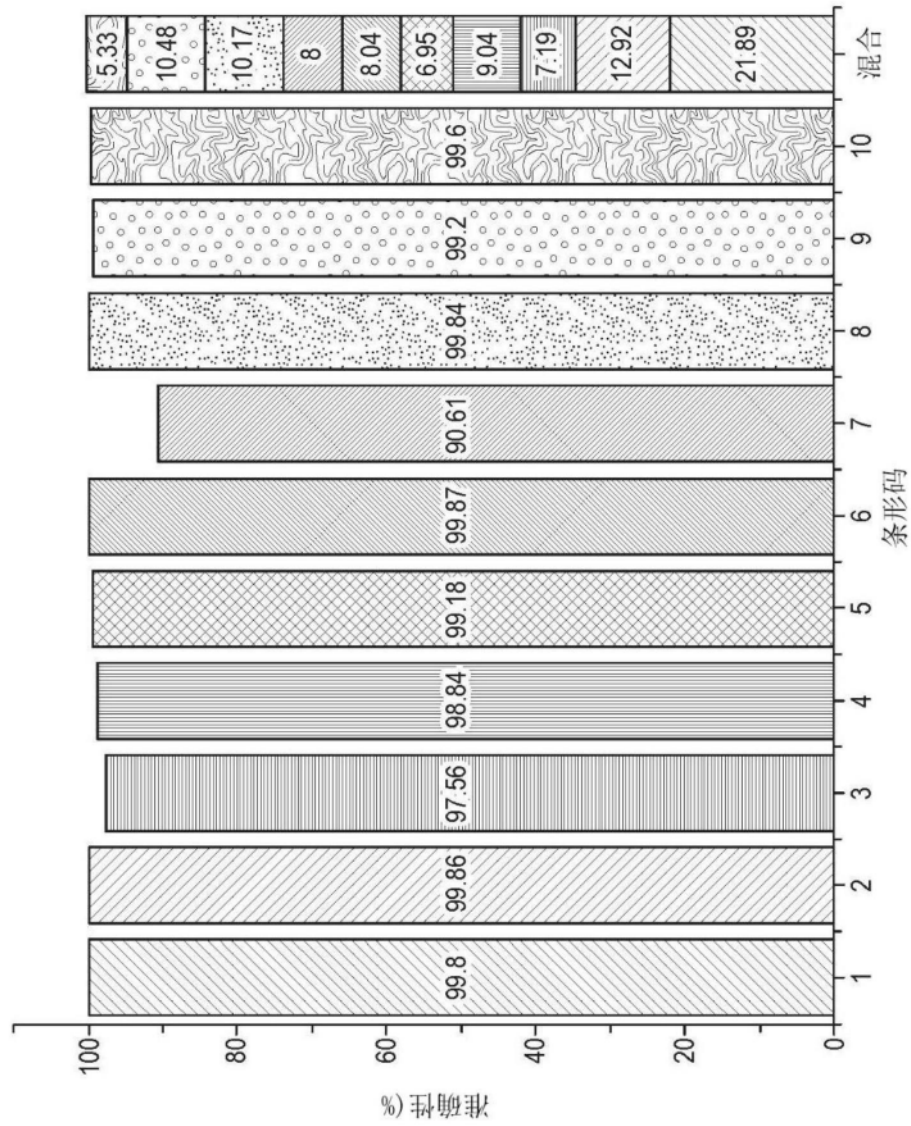


图5A

测试条形码

	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10
B1	99.8	0.08	0	0	0.04	0	0	0	0.23	0
B2	0	99.86	2.03	0	0.04	0	0	0	0	0
B3	0	0	97.56	0	0.04	0	0	0	0.1	0.04
B4	0	0	0	98.84	0.04	0	0	0.05	0.03	0
B5	0	0	0	0	99.1	0	0	0.05	0.07	0
B6	0.17	0.05	0.41	0	0.25	99.87	0.84	0	0.27	0
B7	0.03	0	0	0.58	0.14	0	90.61	0.03	0	0
B8	0	0	0	0.58	0.14	0.09	8.56	99.84	0.1	0.04
B9	0	0	0	0	0.11	0.04	0	0.03	99.2	0.07
B10	0	0	0	0	0.04	0	0	0	0	99.6

分类条形码

图5B

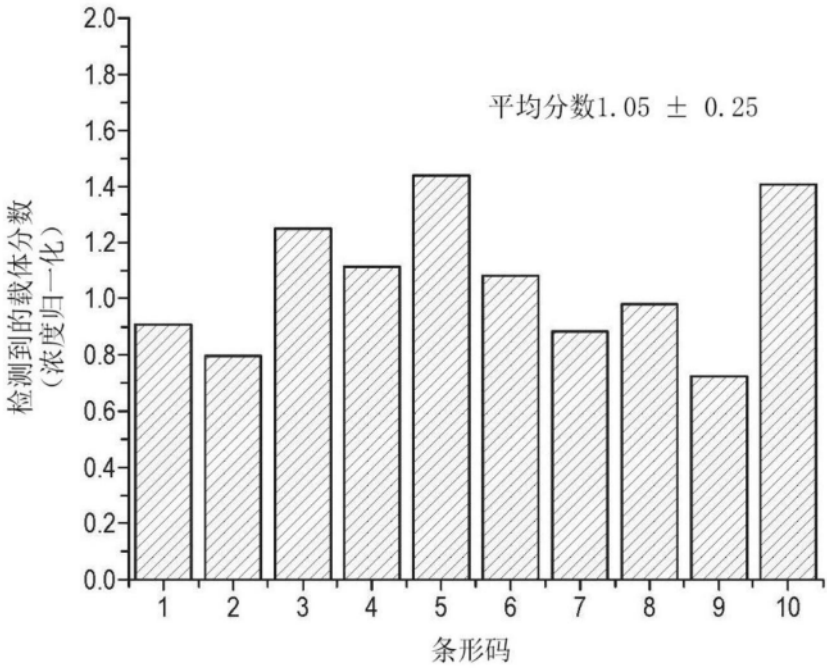


图6

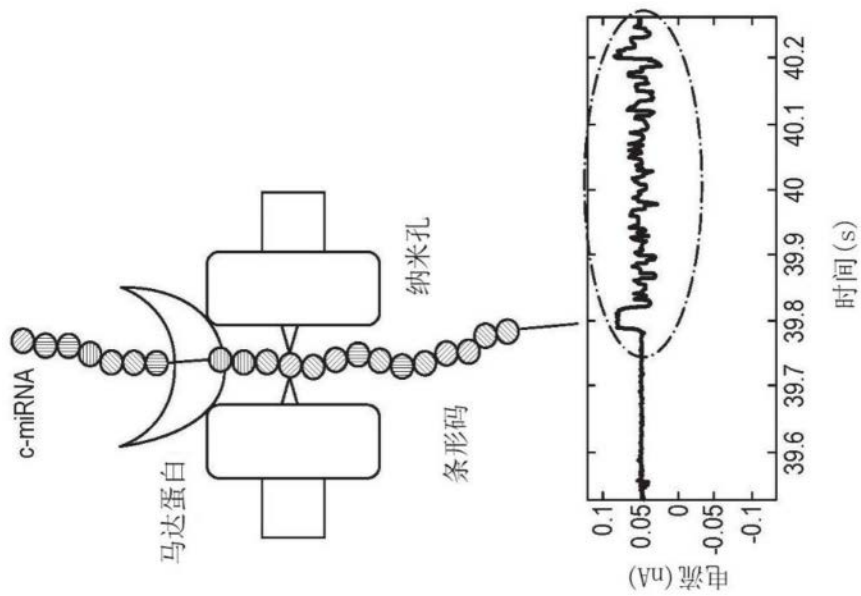


图7A

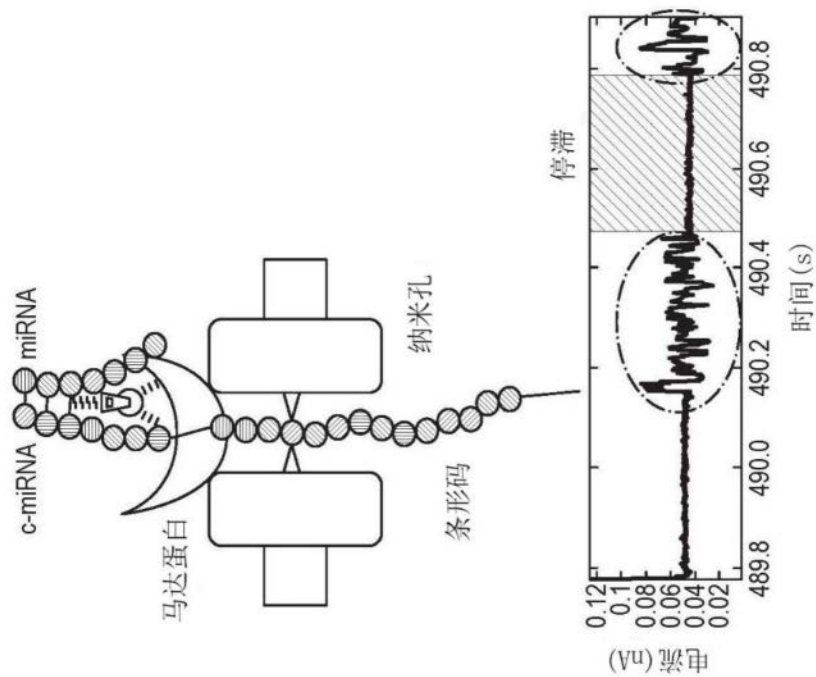


图7B

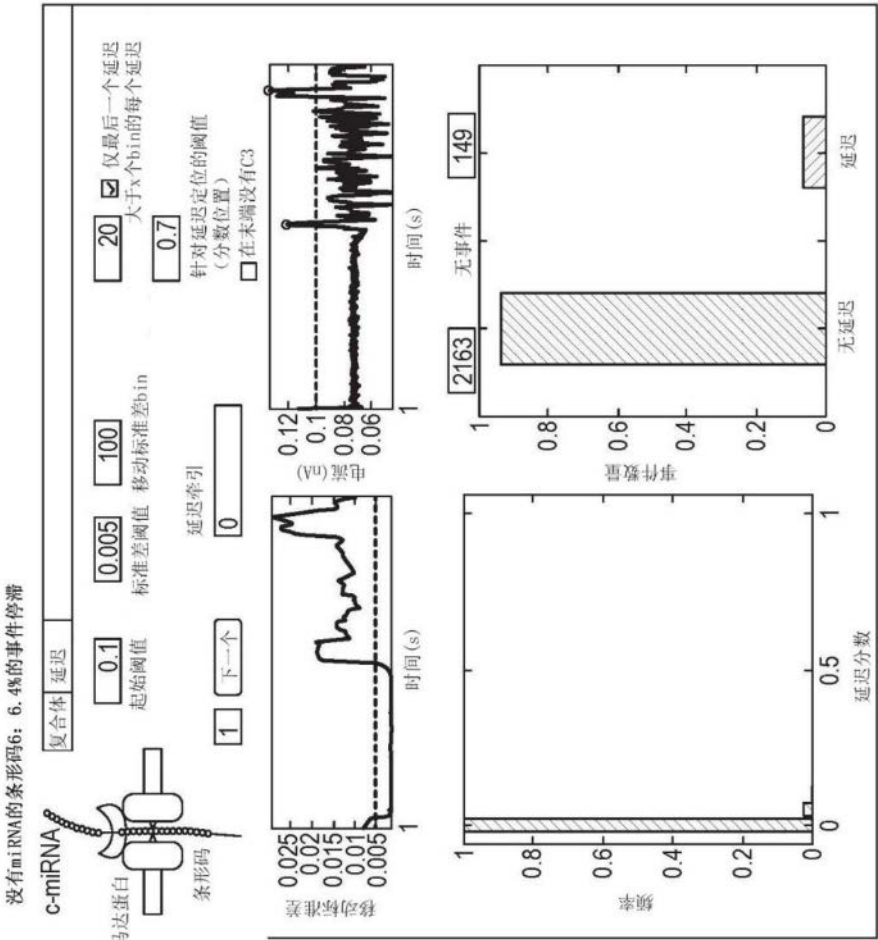


图8

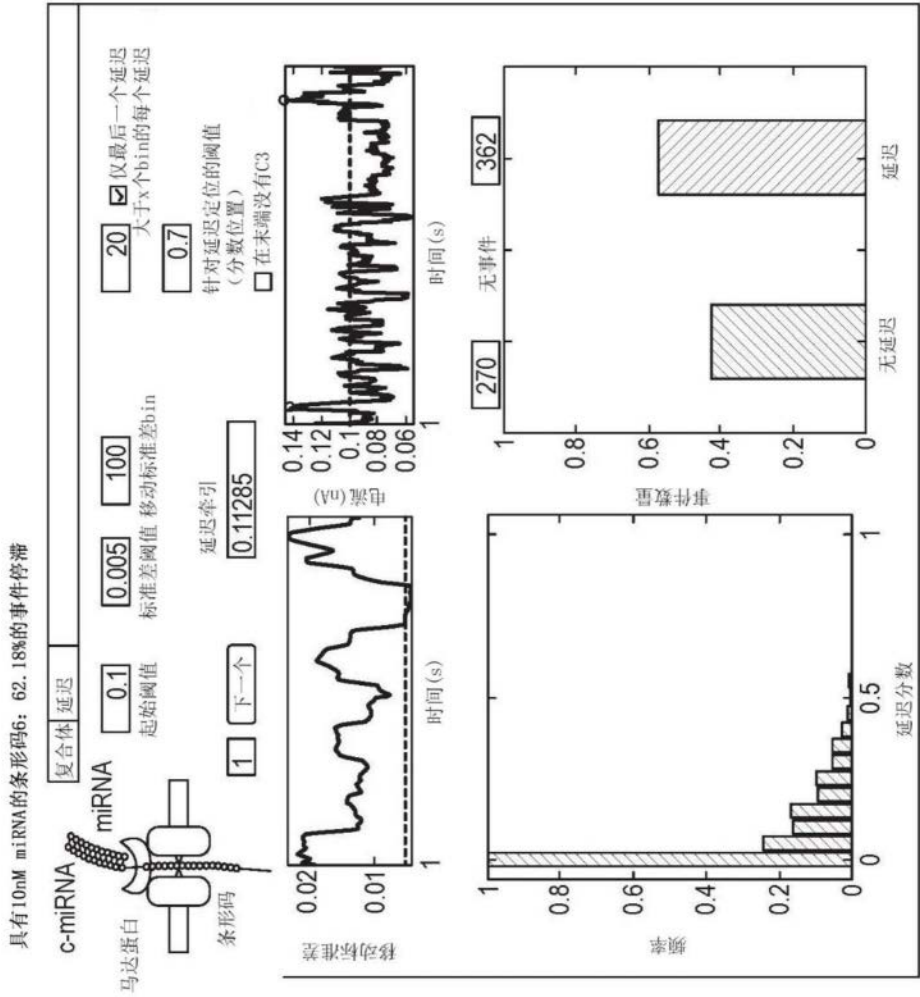


图8(续)

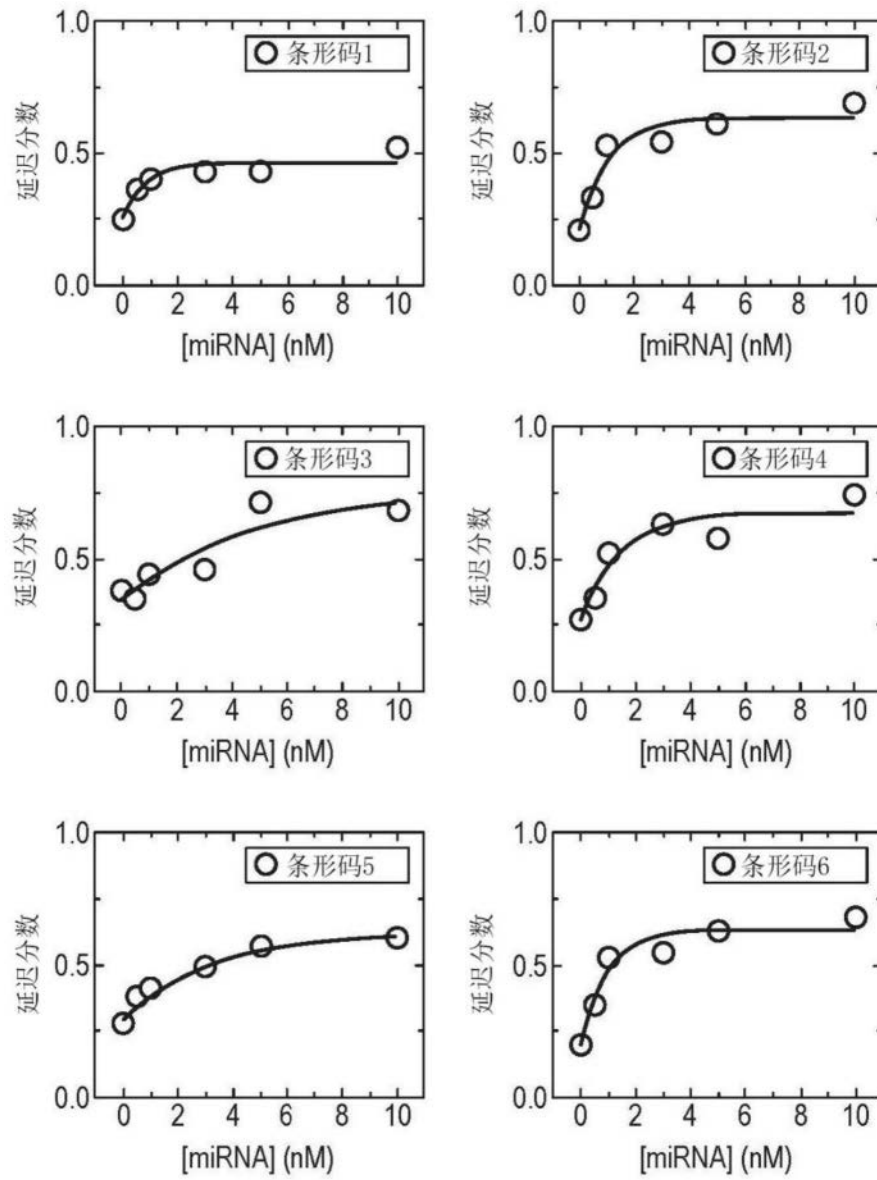


图9

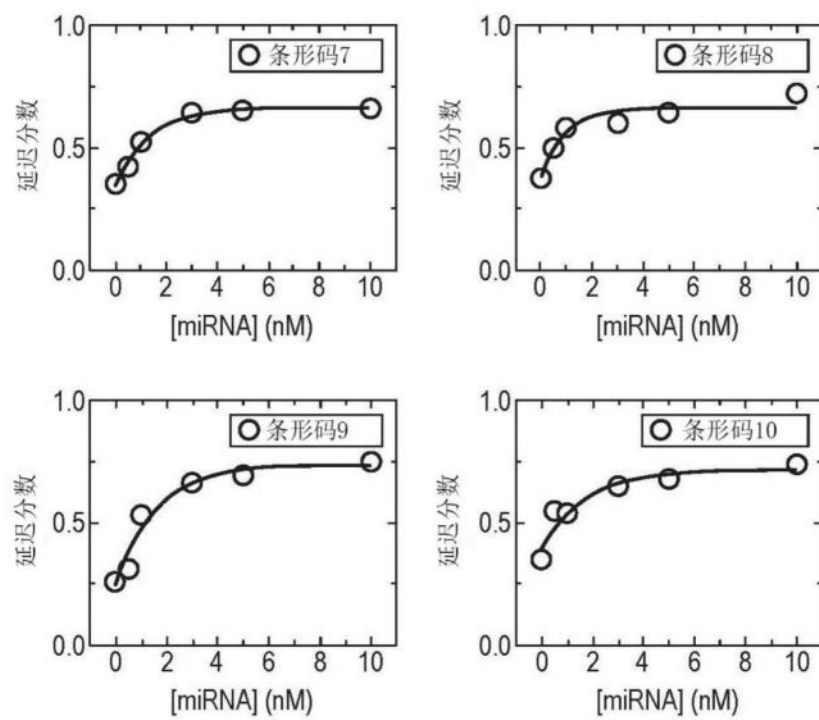


图9 (续)

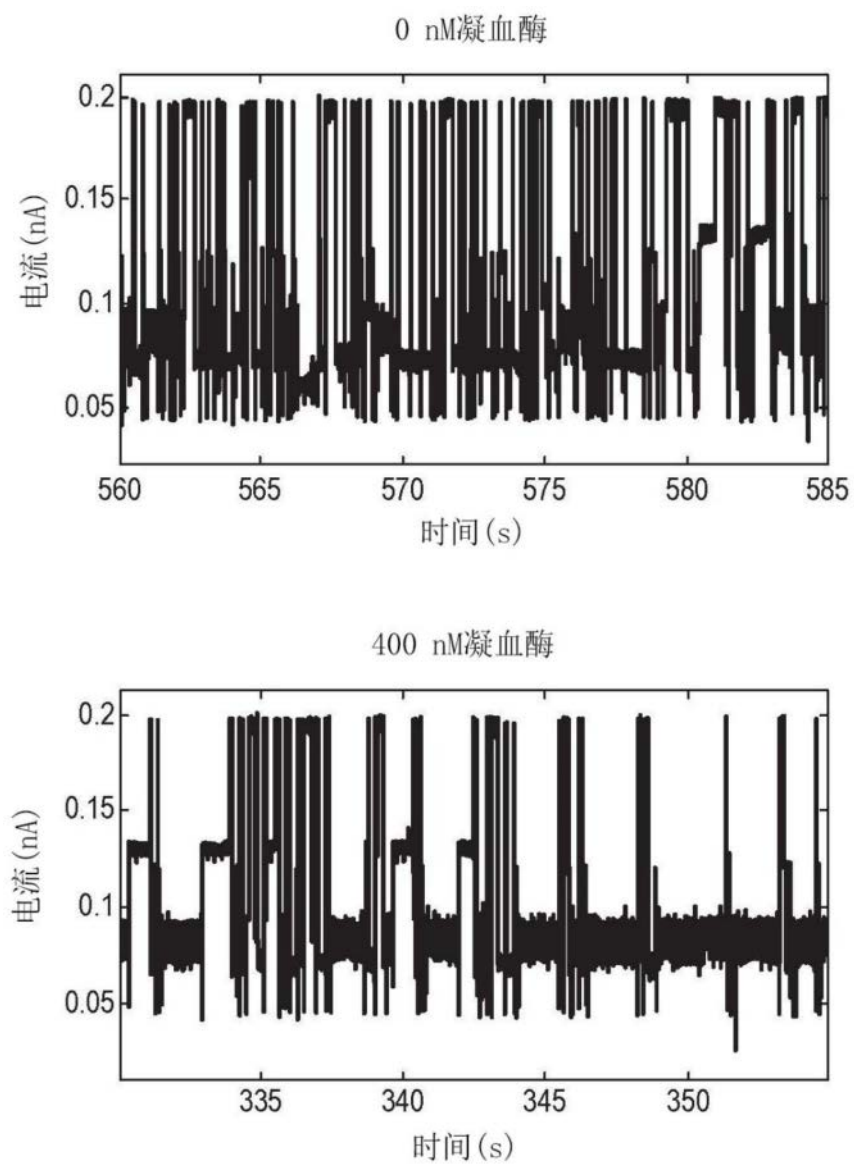


图10

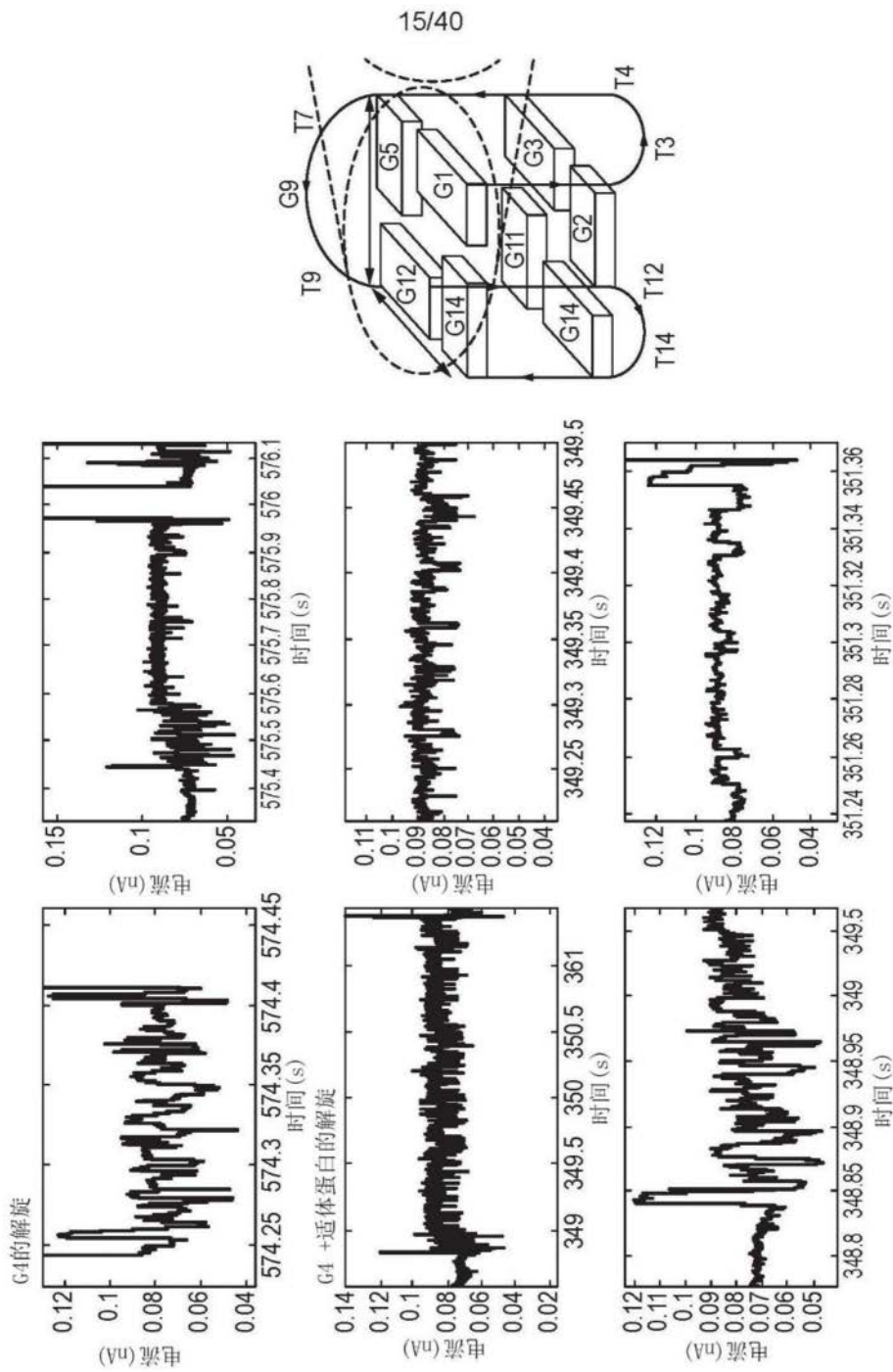


图10(续)

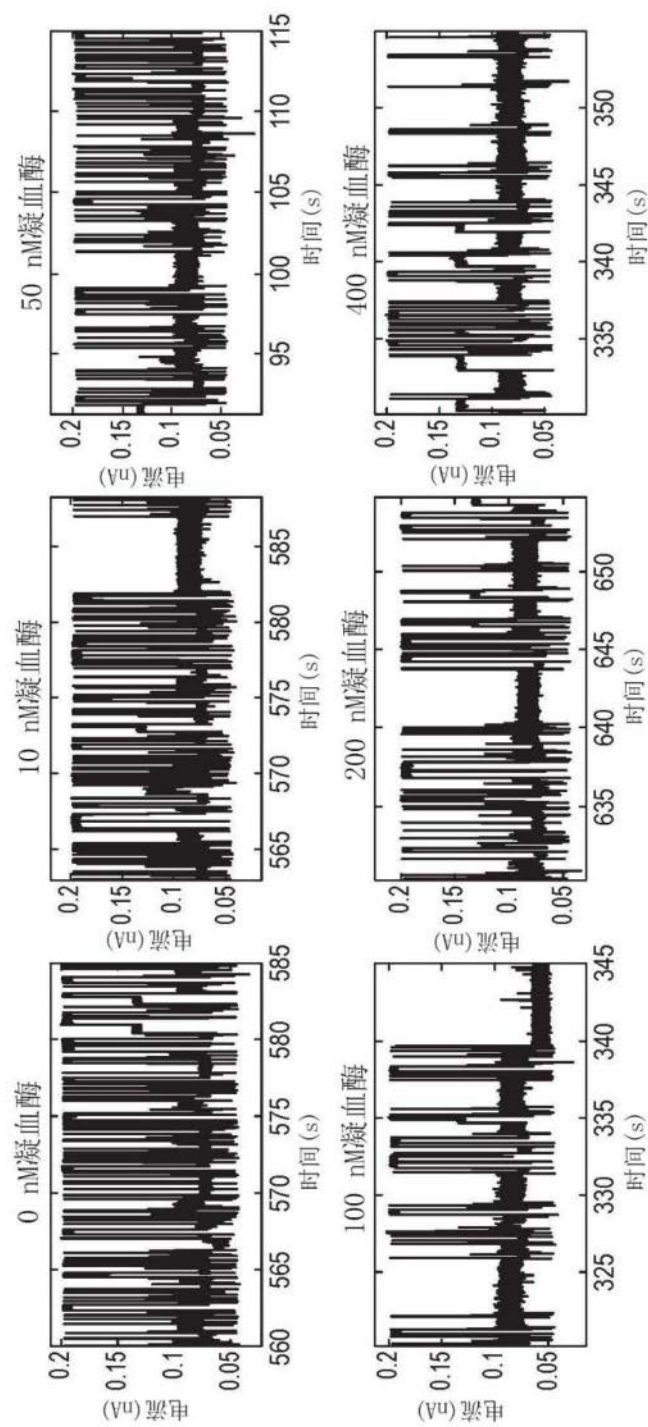


图11

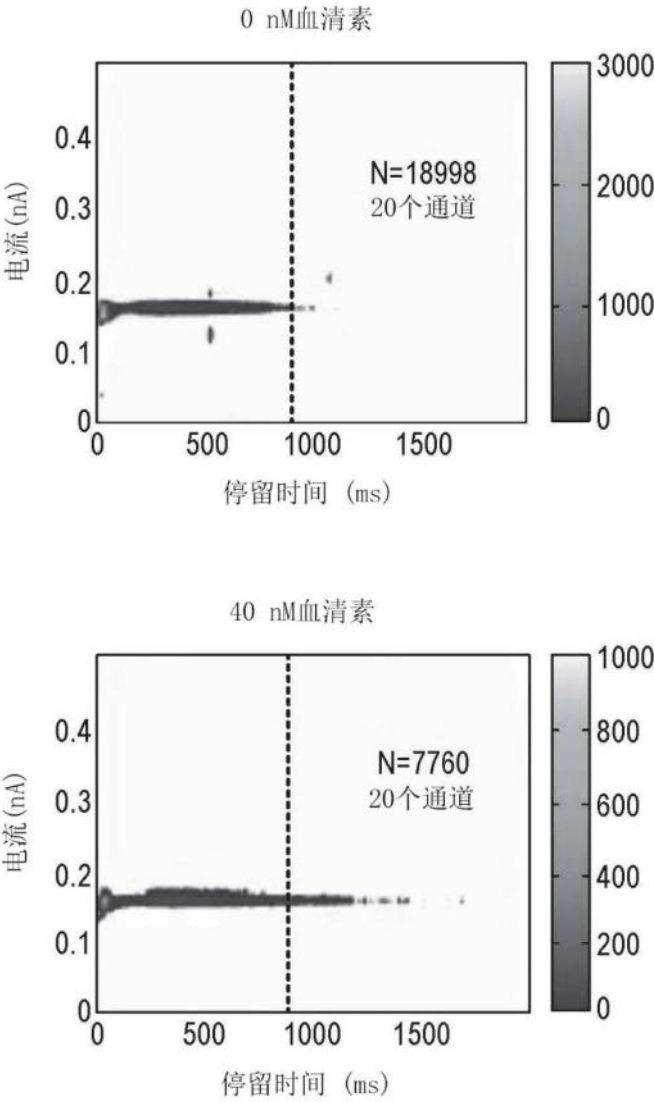


图13A

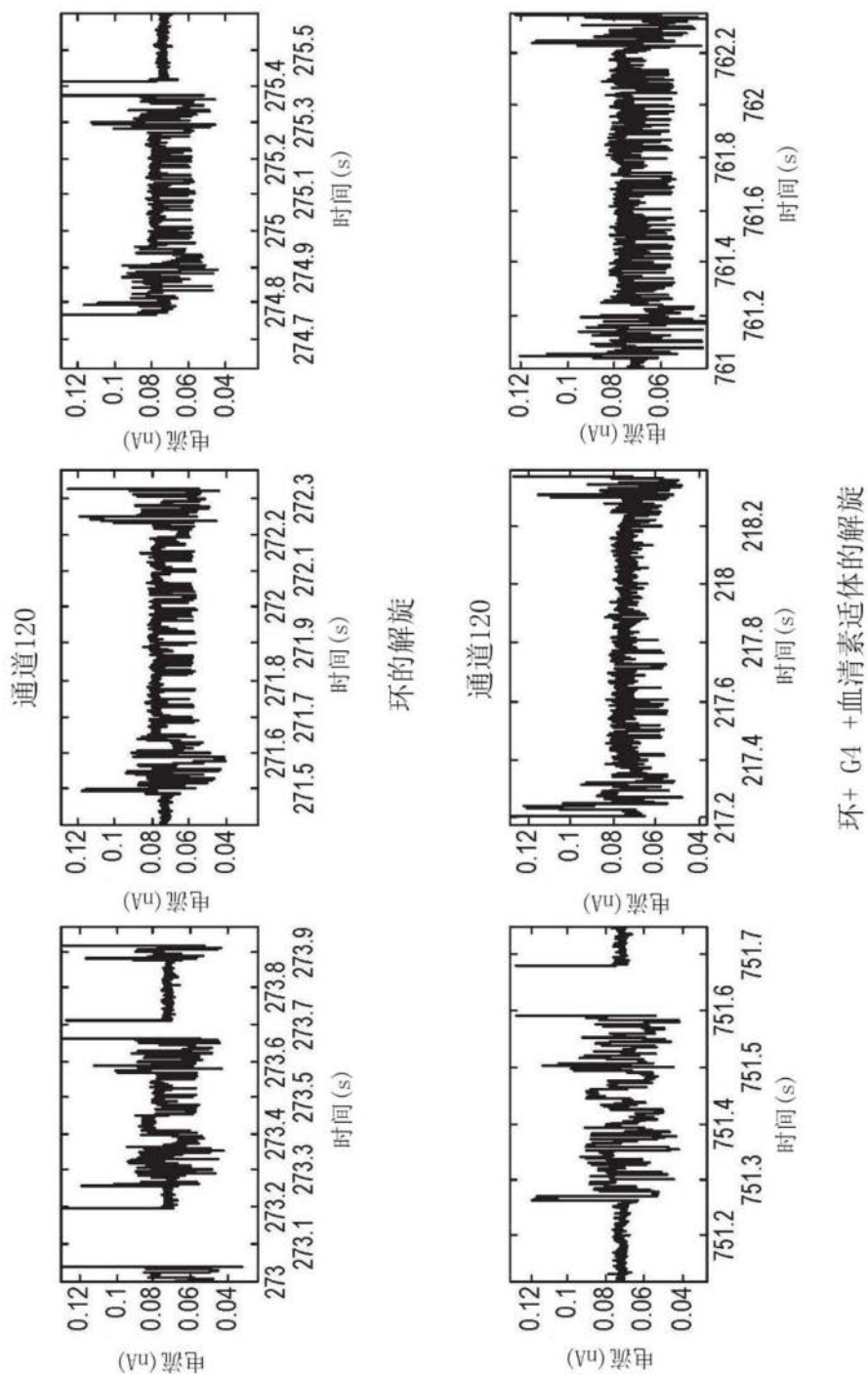


图13A(续)

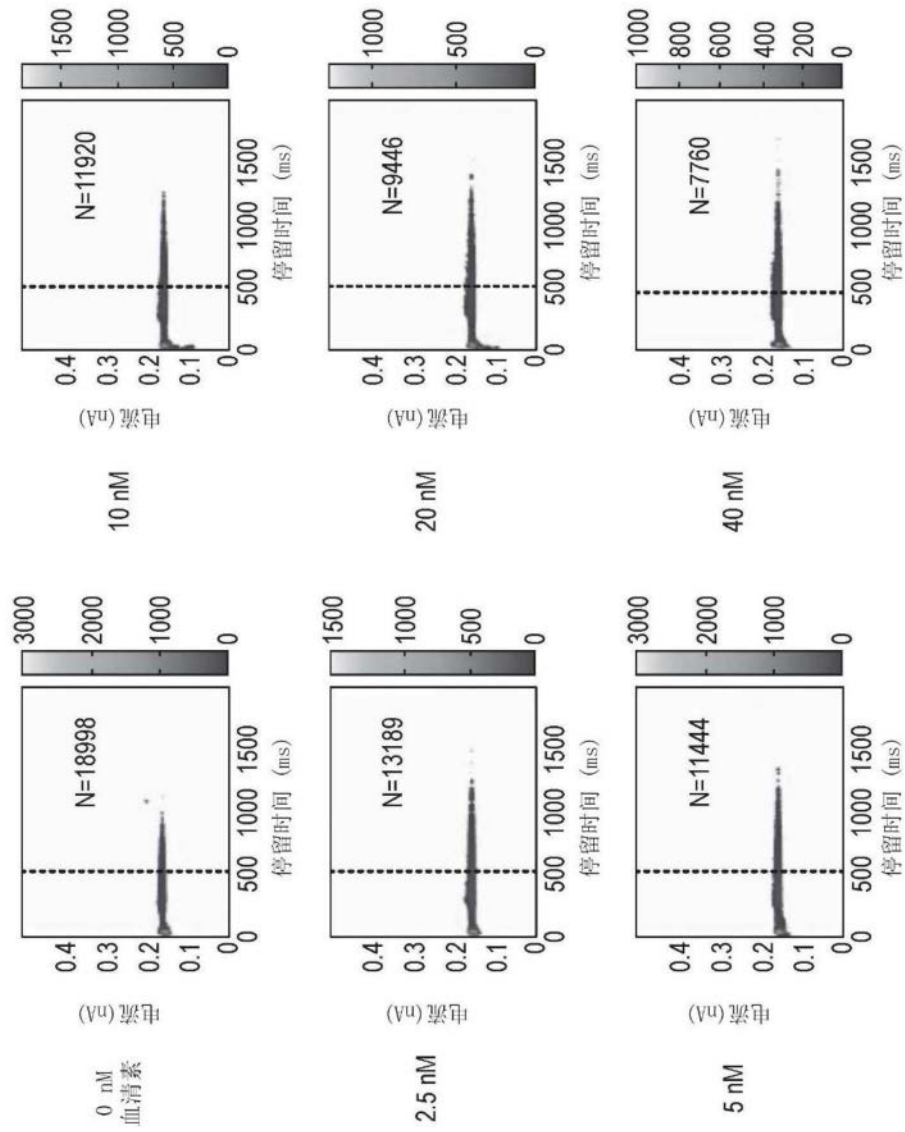


图13B

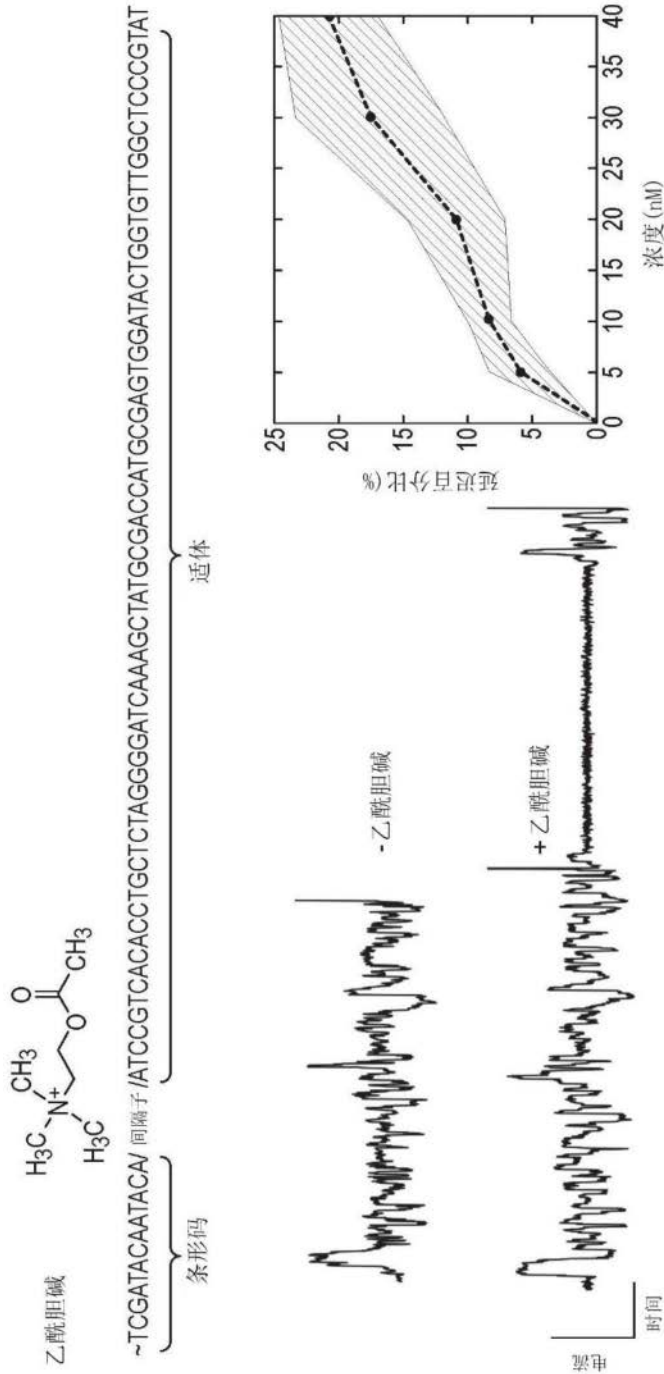


图14

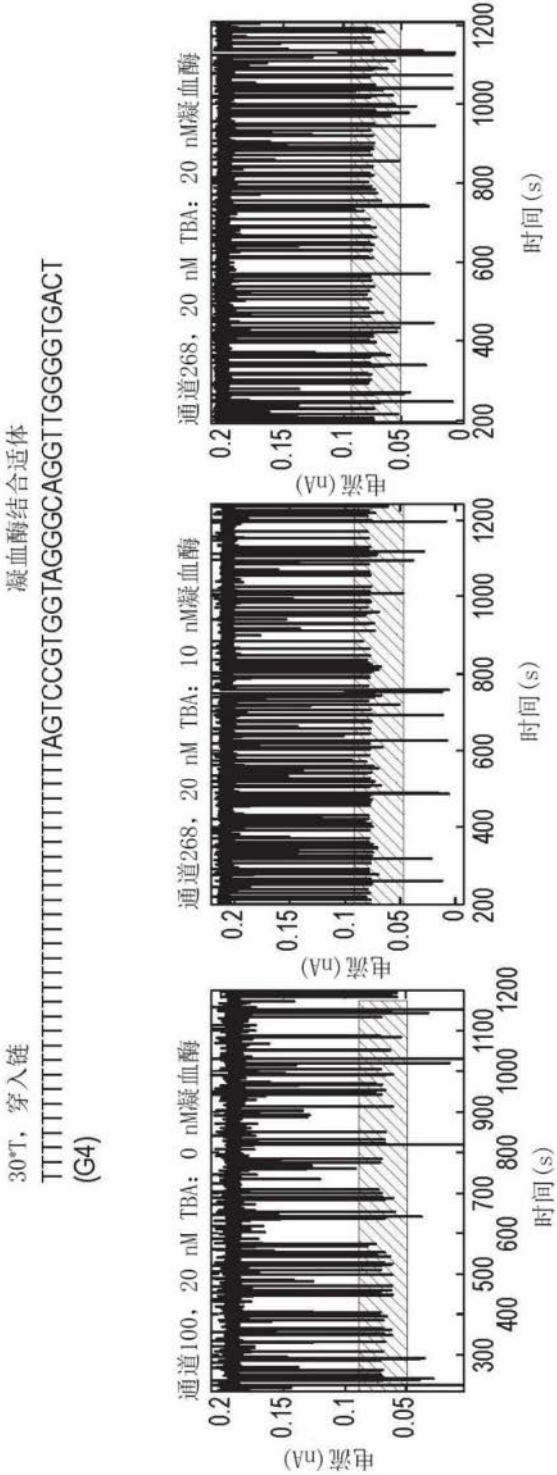


图15A

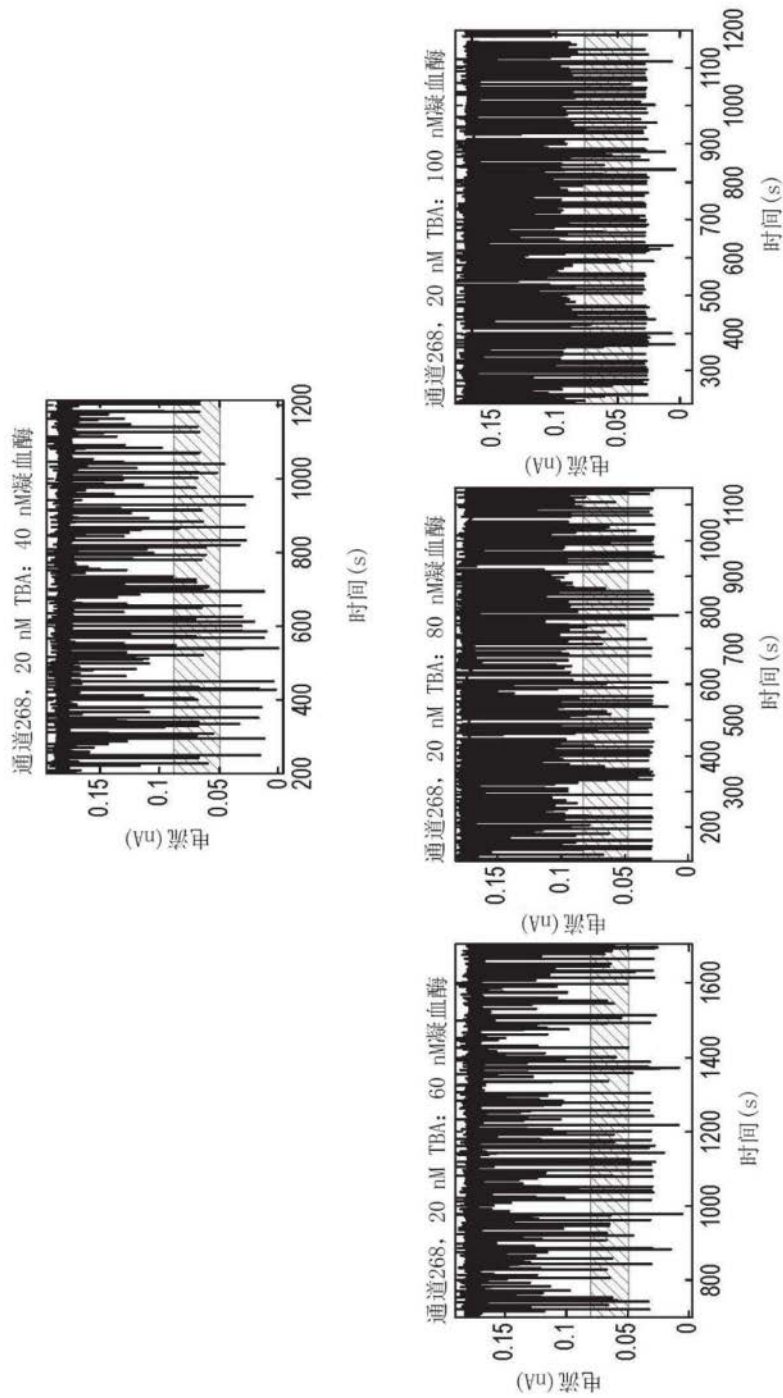


图15A(续)

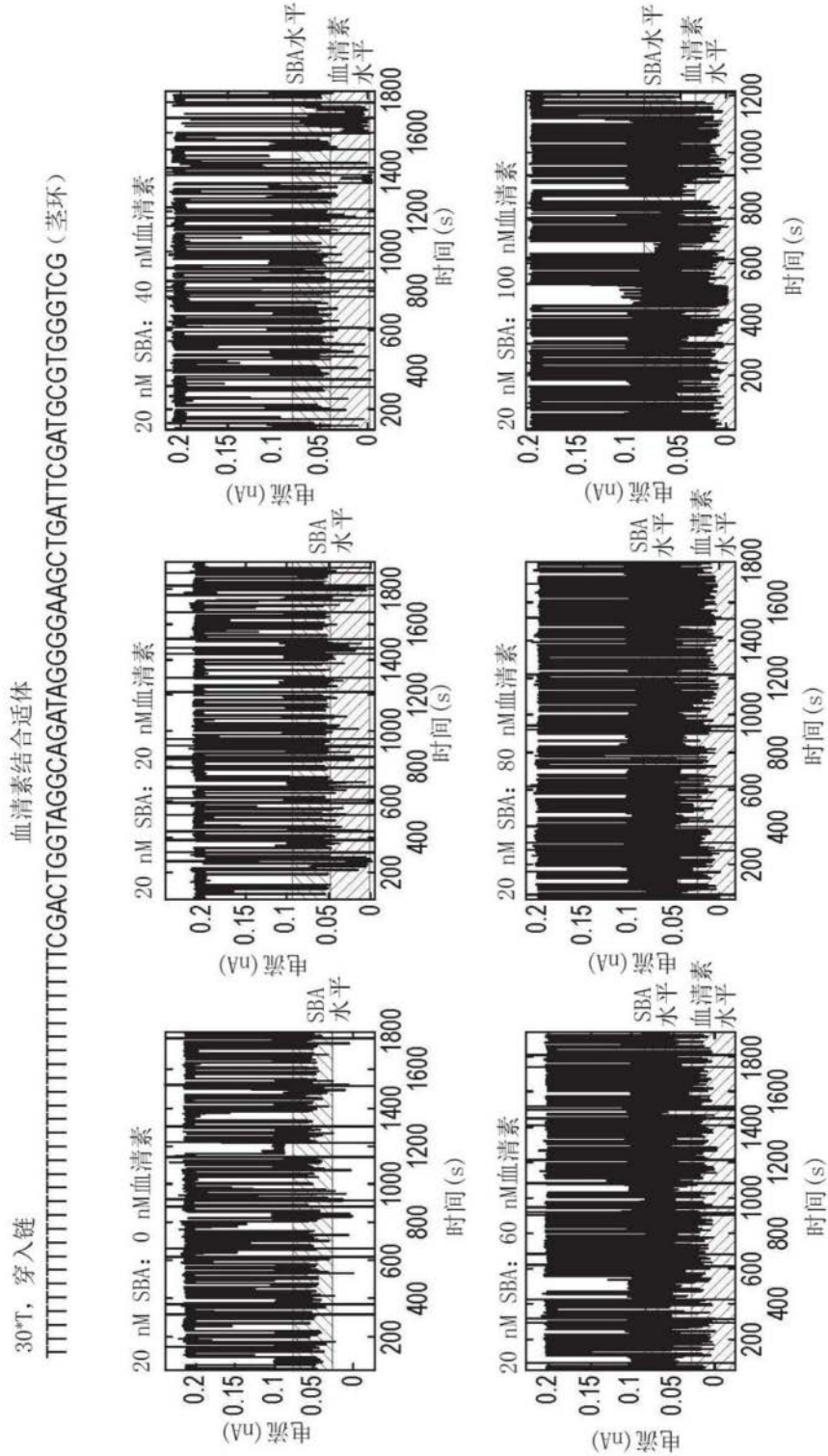


图15B

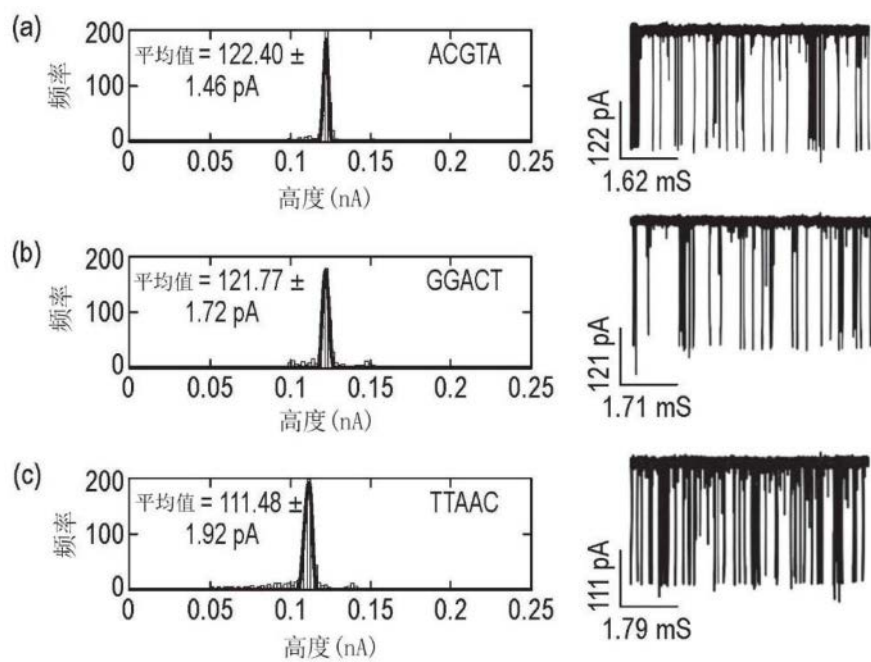


图15C

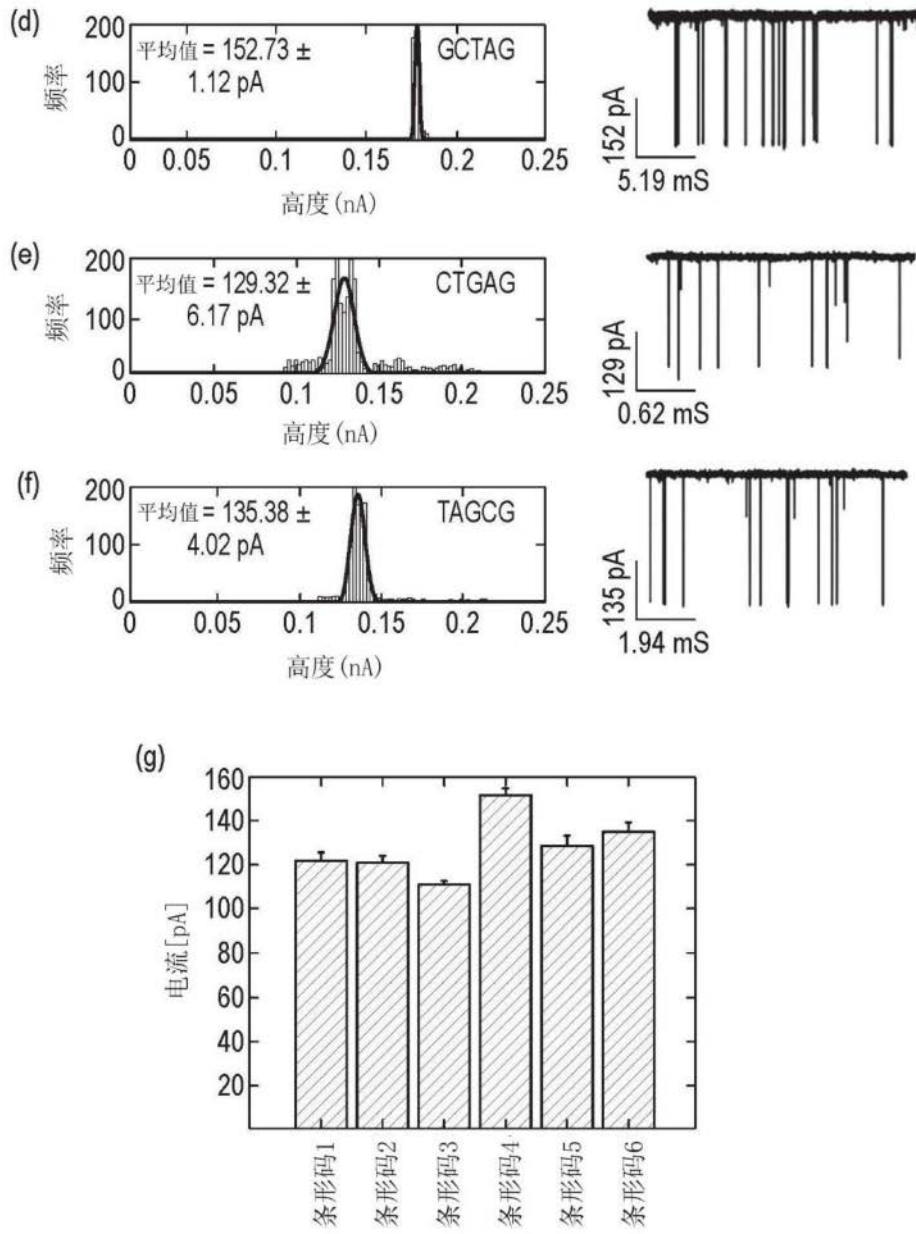


图15C (续)

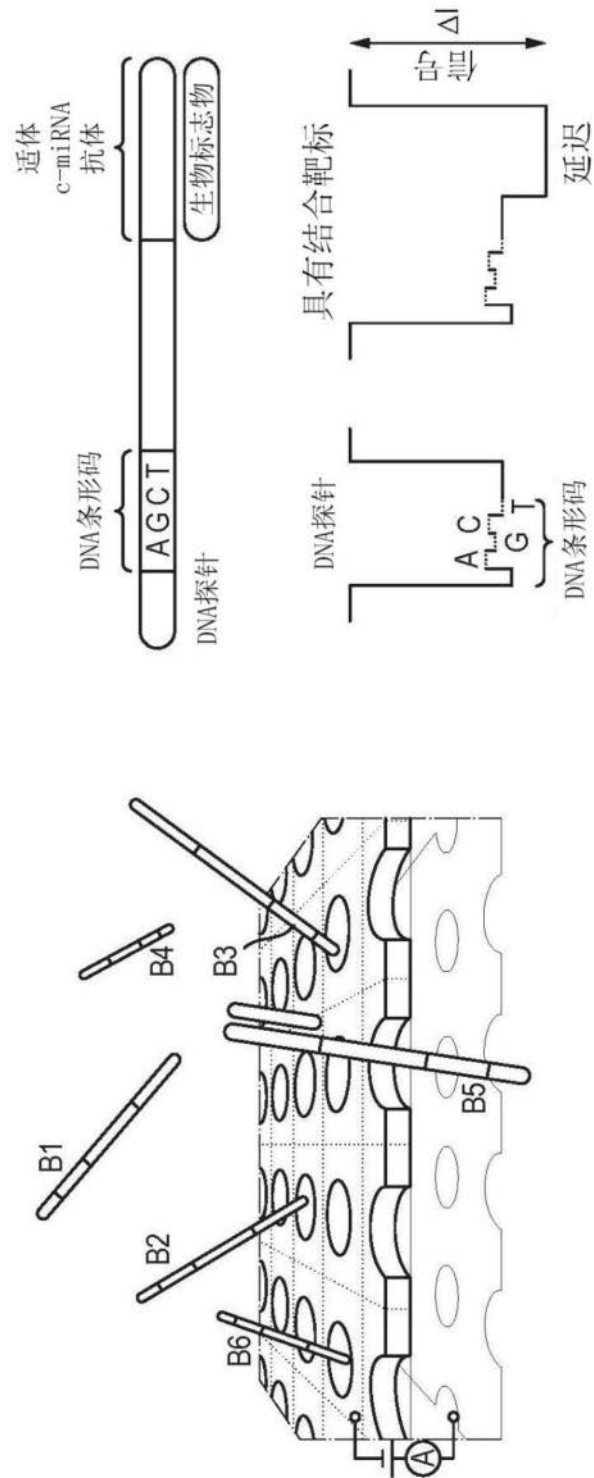


图16

载体链序列-MIRNA

适体_条形码12	_c-hsa-miR-27b-5p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码13	_c-hsa-miR-21-5p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码14	_c-hsa-miR-221-5p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码15	_c-hsa-miR-30d-5p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码16	_c-hsa-miR-30c-5p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码19	_c-hsa-miR-181b-5p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码20	_c-hsa-miR-29a-3p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码22	_c-hsa-miR-210-5p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码24	_c-hsa-miR-126a-5p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码25	_c-hsa-miR-1306-5p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码26	_c-hsa-miR-126-5p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码27	_c-hsa-miR-1254	/5Phos/	CCTAGTCCGCTGGG
适体_条形码29	_c-hsa-miR-30e-5p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码30	_c-hsa-miR-106a-5p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码31	_c-hsa-miR-199a-3p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码32	_c-hsa-miR-652-3p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码33	_c-hsa-miR-26b-5p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码34	_c-hsa-miR-145-5p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码35	_c-hsa-miR-92a-3p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码36	_c-hsa-miR-146a-5p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码37	_c-hsa-miR-423-5p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码39	_c-hsa-miR-27b-3p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码40	_c-hsa-miR-1-3p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码42	_c-hsa-miR-18a-5p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码43	_c-hsa-miR-18b-5p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码45	_c-hsa-miR-301a-5p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码46	_c-hsa-let7c-5p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码47	_c-hsa-miR-125a-5p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码49	_c-hsa-miR-190a-3p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码50	_c-hsa-miR-193b-3p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码51	_c-hsa-miR-193a-5p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码52	_c-hsa-miR-211-5p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码53	_c-hsa-miR-545-5p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码54	_c-hsa-miR-550a-5p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码55	_c-hsa-miR-638	/5Phos/	CCTAGTCCGCTGGG
适体_条形码56	_c-hsa-miR-671-5p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码57	_c-hsa-miR-1233-5p	/5Phos/	CCTAGTCCGCTGGG
适体_条形码58	_c-hsa-miR-3135b	/5Phos/	CCTAGTCCGCTGGG
适体_条形码59	_c-hsa-miR-3908	/5Phos/	CCTAGTCCGCTGGG
适体_条形码60	_c-hsa-miR-5571-5p	/5Phos/	CCTAGTCCGCTGGG

图17

ATCGCTACGCCTTCGGCTCGTAATCATAGTCGAGT	/iSpC3//iSpC3/	GTTACCAATCAGCTAAGCTCT
AGCTCAGAGCAGGTCACTCAAGATACGAGCTGCGT	/iSpC3//iSpC3/	TCAACATCAGTCTGATAAGCTA
GTAAGTCTGCATCAGCGCGCGGCTGTGCGAGGATA	/iSpC3//iSpC3/	AAATCTACATTGTATGCCAGGT
CTACGACAGTACGCTAGCAAGGATAGACTACGA	/iSpC3//iSpC3/	CTTCCAGTCGGGGATGTTTACA
TACTGAACACAAGTTCGTCGTCGAGCAATCACAAT	/iSpC3//iSpC3/	GCTGAGAGTGTAGGATGTTTACA
CTAGTGCGCAGTTGTCTCGGCGGAGTTGAGACTGA	/iSpC3//iSpC3/	ACCCACCGACAGCAATGAATGTT
GATCATGGTAGTCTTCAAGATCGAGTATGTCTGTC	/iSpC3//iSpC3/	TAACCGATTTTCAGATGGTGCTA
GTTACATCAAGGTCATACCGCGAGTTCTATTTTA	/iSpC3//iSpC3/	CAGTGTGCGGTGGGCGAGGGGCT
GCTTGGGGGATAGATGTGCCCGCGCATCGGACCT	/iSpC3//iSpC3/	CGCGTACCAAAAAGTAATAATG
ATGACACACGTTTTCGATAGGGGACGCCGACTTTAA	/iSpC3//iSpC3/	TGGACGTTTGCAGGGGAGGTGG
TGATAATAAGACCTGACAGACAATAGGGAGAATC	/iSpC3//iSpC3/	CGCGTACCAAAAAGTAATAATG
TTAGTAATCAAGTCTGATCGTAATAGCTAAGTCAT	/iSpC3//iSpC3/	ACTGCAGGCTCCAGCTTCCAGGCT
CCTGTTCTAATTGCGCGGAGAGGCGAGATGTTTCT	/iSpC3//iSpC3/	CTTCCAGTCAAGGATGTTTACA
TGGGATACTGACGTGCCGGAATACCCAGACGTGCC	/iSpC3//iSpC3/	CTACCTGCACTGTAAGCACTTTT
TGATCGGCACCTAAACGCATTAGCCCTGCAATACG	/iSpC3//iSpC3/	TAACCAATGTGCAGACTACTGT
AGCTGCTCGGAAGCCATAAGGTACTTTAATTTGGG	/iSpC3//iSpC3/	CACAACCTAGTGGCGCCATT
CACGGATTTCTATATTGCTCAACCAGGCAGCGCAA	/iSpC3//iSpC3/	ACCTATCCTGAATTACTTGAA
ATTGTGCGCTTTTCTCCATGCGTCTTAGACATCTC	/iSpC3//iSpC3/	AGGGATTCTGGGAAAAGTGGAA
GTGAAATCCCCGTCTAGGTTATGGCTGGGGGGATT	/iSpC3//iSpC3/	ACAGGCCGGGACAAGTGCAATA
CGTAAACTTATCACGACACAATGAACAAGCCTGCA	/iSpC3//iSpC3/	AACCATGGAATTCAAGTTCTCA
GGCAGTGTCCGAGCGTCTCAATCATGAGCGATTCT	/iSpC3//iSpC3/	AAAGTCTCGTCTCTGCCCTCA
CGTCCAGACTTAATGTCTGCTCACTGACATCGCGA	/iSpC3//iSpC3/	GCAGAACTTAGCCACTGTGAA
ATTAGCGGAACCAACCCAGGAAGGCTTGAAGGCG	/iSpC3//iSpC3/	ATACATACTTCTTTACATTCCA
AACCTTAGGGGCTCGAATCTTTGAGACGACTAGG	/iSpC3//iSpC3/	CTATCTGCACTAGATGCACCTTA
TAATTACTGCCCCACCATGACATTTTAATAGCAGT	/iSpC3//iSpC3/	CTAACTGCACTAGATGCACCTTA
GACCTTGAGACAGAACTTATCAATGTACAAGTGA	/iSpC3//iSpC3/	AGTAGTGAATAAAGTCAGAGC
CGAAGGATTGGCCCCCGATTACCACCGCCGTGAG	/iSpC3//iSpC3/	AACCATACAACCTACTACCTCA
AATTGCCAACAGGTCAAGCCCTGTTCTCACTGGTC	/iSpC3//iSpC3/	TCACAGGTTAAAGGGTCTCAGGGA
CAGTGCTTGCCCCAGTAGAGTGTGGAAGGGCATA	/iSpC3//iSpC3/	AGGAATATGTTTGATATATAG
AAAGCCCACTCTCCACACTTCAAGGTTAAATGGCG	/iSpC3//iSpC3/	AGCGGGACTTTGAGGGCCAGTT
GACTAATACAATCGGAAGCAACTCTACGCCGCAC	/iSpC3//iSpC3/	TCATCTCGCCGCAAAGACCCA
CTACGATTCATGTCTCCCCCACATATGATTGATC	/iSpC3//iSpC3/	AGGCGAAGGATGACAAAGGGAA
TCGGGCGCTTAATCGCAATGTTTATCCGGAACGGA	/iSpC3//iSpC3/	TCATCTAATAAACATTTACTGA
TAAATTCACATTACGATAGCTGACGTCCTGGTAG	/iSpC3//iSpC3/	GGGCTTTACTCCCTCAGGCACT
CACCCCTAGCCTGACCTTAAATGATAATTCCTGT	/iSpC3//iSpC3/	AGGCCGCCACCCGCCGCGATCCCT
TTTAGCCTAACCTCTGATGCATTGCACCGGAGTTC	/iSpC3//iSpC3/	CTCCAGCCCTCAGGGCTTCTCT
CCCAATATACGCTGAACCTTCCATCCGATTTTCAG	/iSpC3//iSpC3/	TGCCGTGCCCTGGCCTCCCACT
CCTTGGAATTAGAACCGTGTGATTCTACGCCTAGG	/iSpC3//iSpC3/	CACCACTGCACTCGTCCAGCC
AGGACTTCTTCGGTGTAAATCGGAGTACATCAATGT	/iSpC3//iSpC3/	AAACAGTCTACCTACATTGCTC
AACGTACTTTGTGGGGATAAGCTGTACAGGGCTGT	/iSpC3//iSpC3/	GGGAGGCTCCTTGAGAATTG

图17(续)

载体链序列-蛋白质和神经递质

心肌钙蛋白I (cTnI):			
条形码65	/SPHos/	CCTAGTTCGGCTGGG	ATTCAGATCTTTGACCGATCGTACTGACGTGACG /SpC3/ /SpC3/ AGTCTCCGCTGTCCTCCCGATGCAC TTGACGTAGTCTCAGTCTTTTTCATTGACATGGGATGAGCCGCTGACTG
心肌钙蛋白T (cTnT):			
条形码62	/SPHos/	CCTAGTTCGGCTGGG	CTTAAGTTCGGGTATTAACAGTGTCAATCTTGCCA /SpC3/ /SpC3/ ATACGGGAGCCAAACACAGGACTAAGATTATTAAGATTGCCAATATCATTGGAGAGCAGGTGTGTGACGGAT
BNP:			
条形码61	/SPHos/	CCTAGTTCGGCTGGG	ATGTAATTCCTGCTAGGTCCAAATGTGACTGTACTA /SpC3/ /SpC3/ GGGGATTCGTGATCTCTGCTCTCGGTTTCGGGTTGCTTCG
凝血酶:			
条形码70	/SPHos/	CCTAGTTCGGCTGGG	TTAAGCTATTGCTAACTGTAGTCCCTAGTCTAGCTA /SpC3/ /SpC3/ AGTCCGTGTTAGGGCAGGTTGGGGTGACT
条形码1	/SPHos/	CCTAGTTCGGCTGGG	GAGCTAGCATACGTGTCTAAACACTGCACAGATGAT /SpC3/ /SpC3/ GGTGGTGTGGTTGG
S-蛋白:			
条形码7	/SPHos/	CCTAGTTCGGCTGGG	ACCTTTTCTAGGATGGAAACATCTA /SpC3/ /SpC3/ CACGGATACGCTCTTCGGGGGGGGGTTGAGAGGATGTGGGGTGTATGCCGTG
N-蛋白:			
条形码6	/SPHos/	CCTAGTTCGGCTGGG	CTGGGACATGACTATCTAGAGTCGC /SpC3/ /SpC3/ GCTGGGATGCACCAATTTGCGGACATCGGATTGTCTGAGTCAATGACACATCCAGC
血清素:			
条形码5	/SPHos/	CCTAGTTCGGCTGGG	ACTAGATAAAGGAGGGGAGCAGACAGTAACTCGTT /SpC3/ /SpC3/ CGACTGGTAGGCAAGTAGGGGAAGCTGATTCGATG
乙酰胆碱:			
条形码4	/SPHos/	CCTAGTTCGGCTGGG	TGATACATACA /SpC3/ /SpC3/ ATCCGTACACCTGCTCTAGGGGATCAAGCTATGCCACCATGCGAGTGGATCTGTTGGCTCCCGTAT

图18

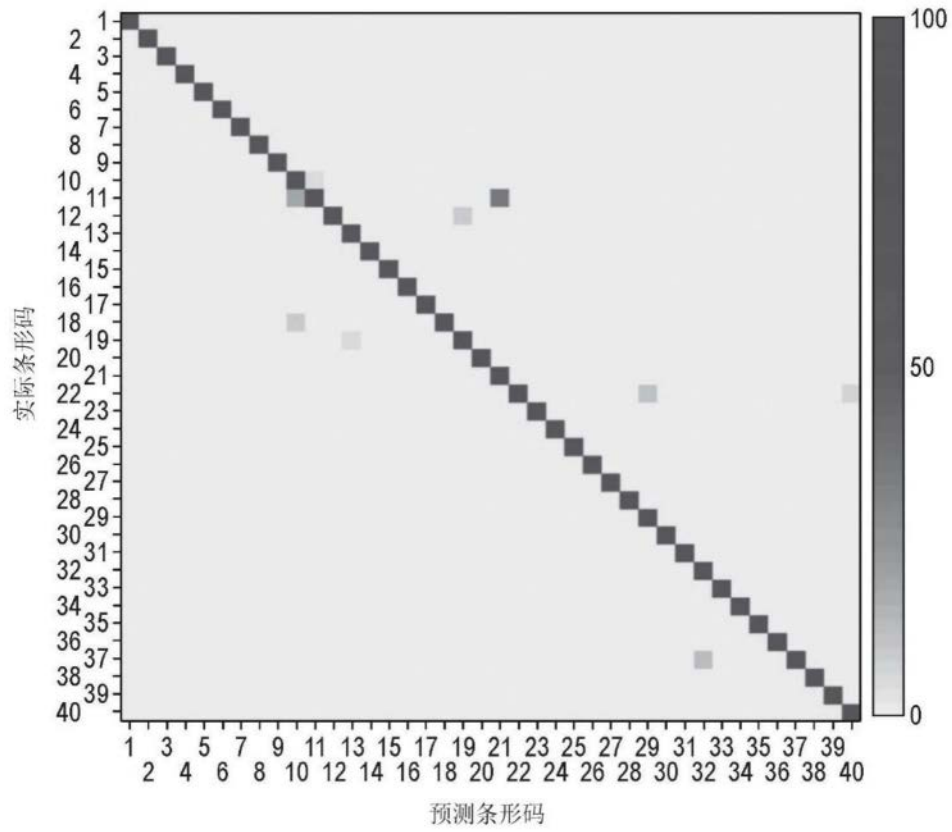


图19

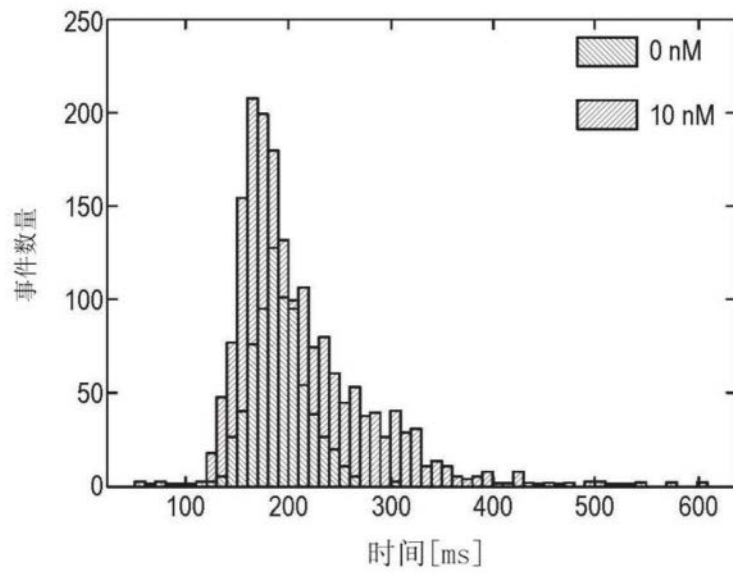


图20A

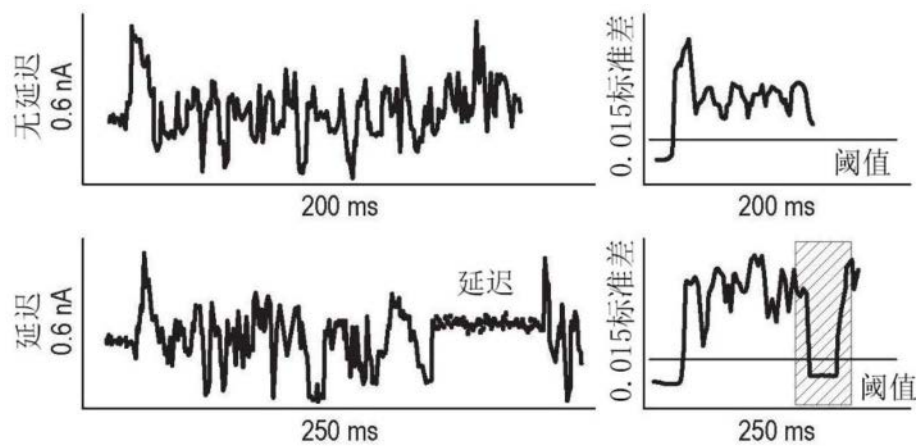


图20B

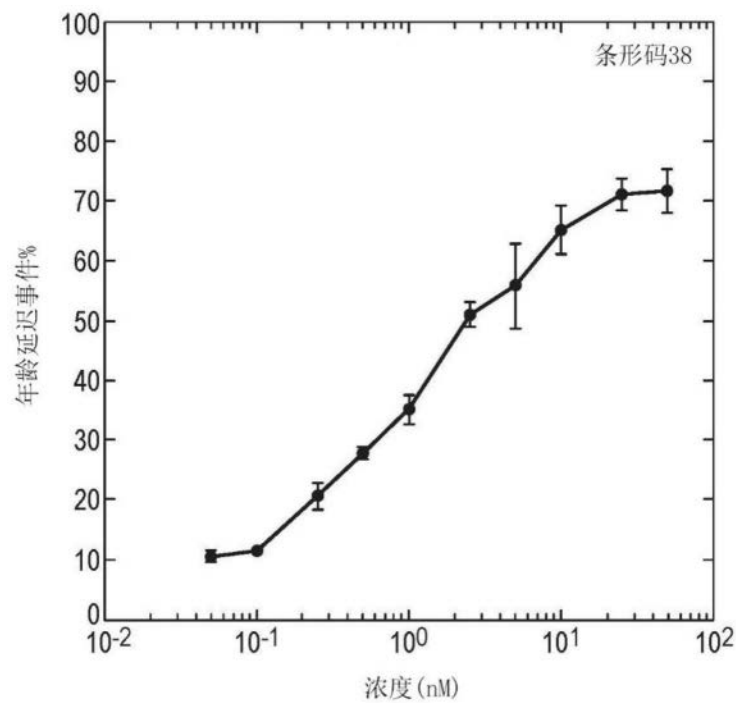


图20C

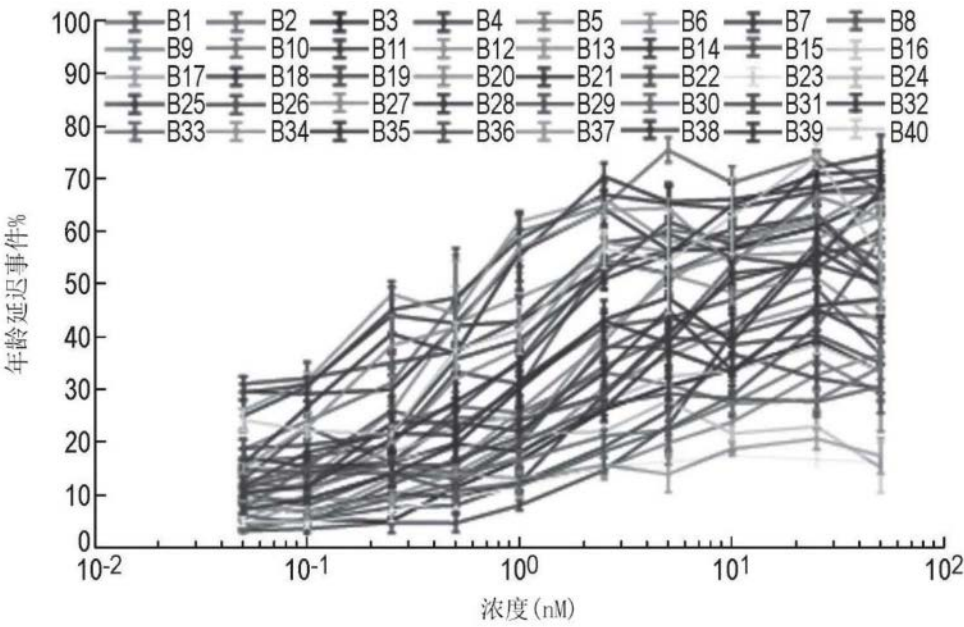


图20D

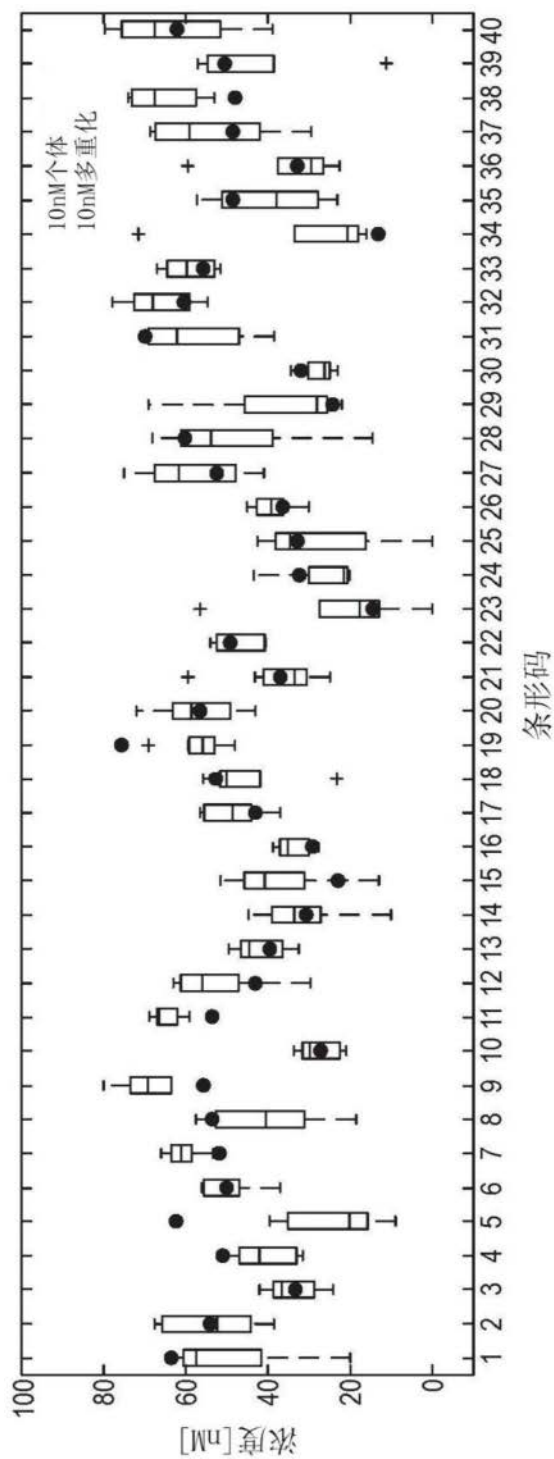


图20E

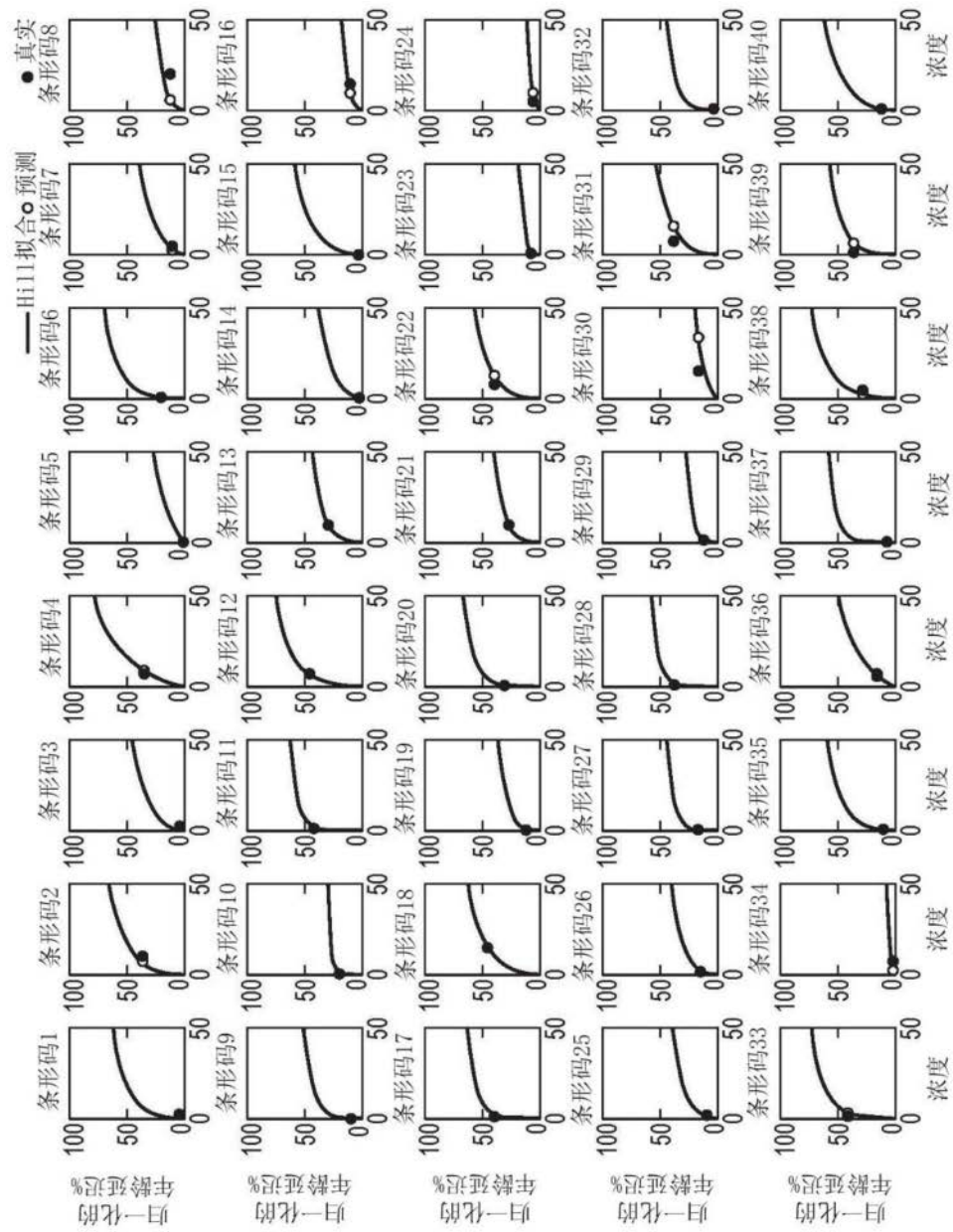


图21A

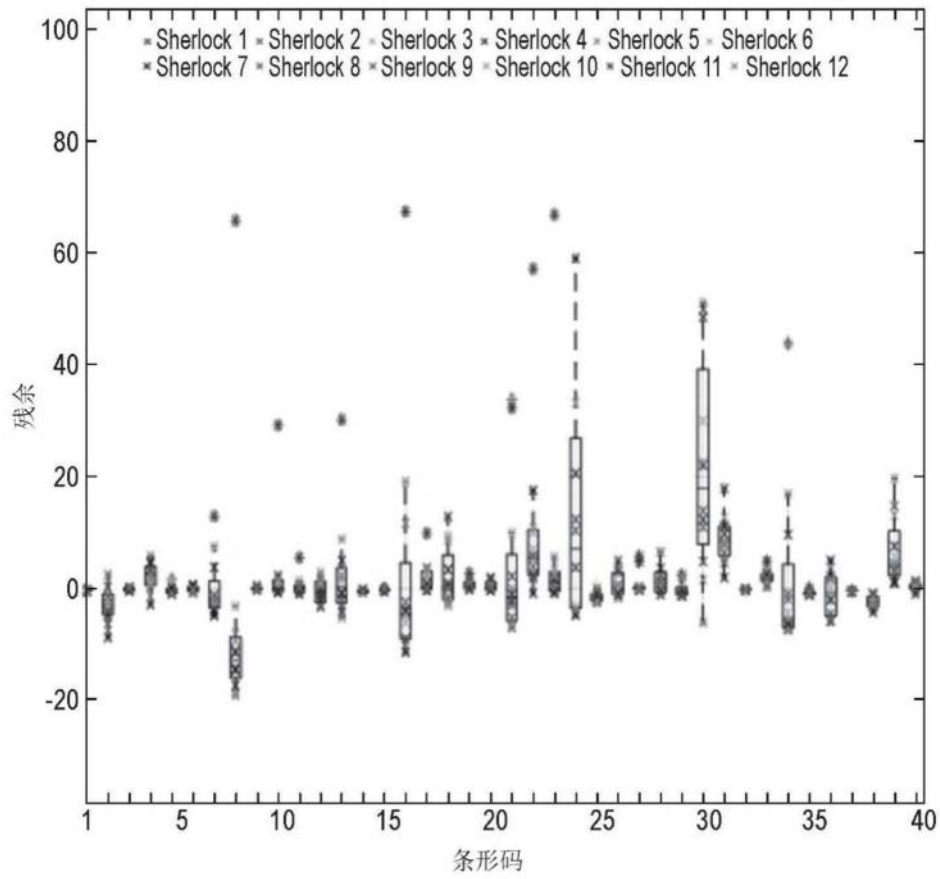


图21B

AGTCTCCGCTGTCTCCCGATGCACTTGACGTATGTCTCACTTTCTTTTCATTGACATGGGATGACGCCGTGACTG

图22A

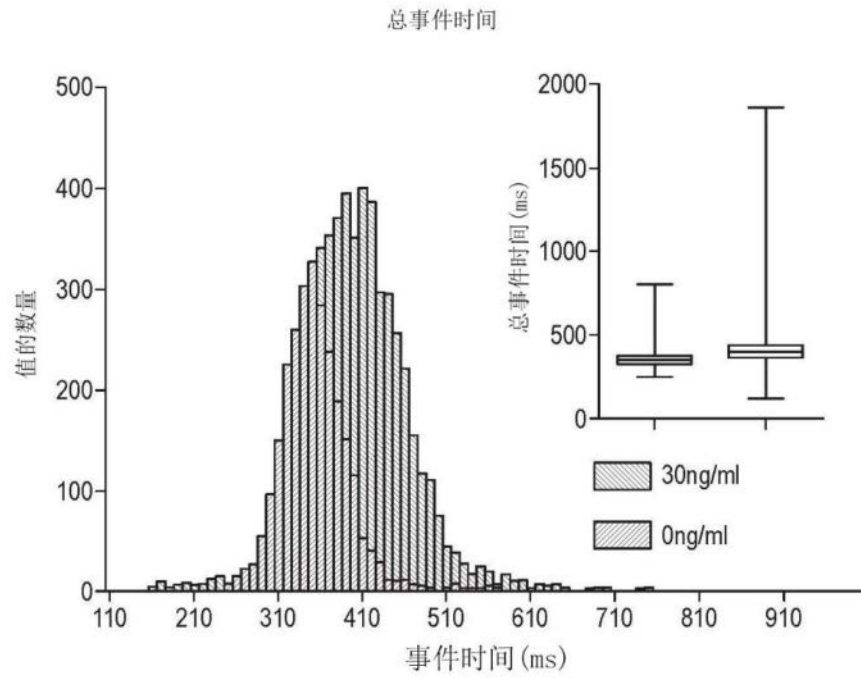


图22B

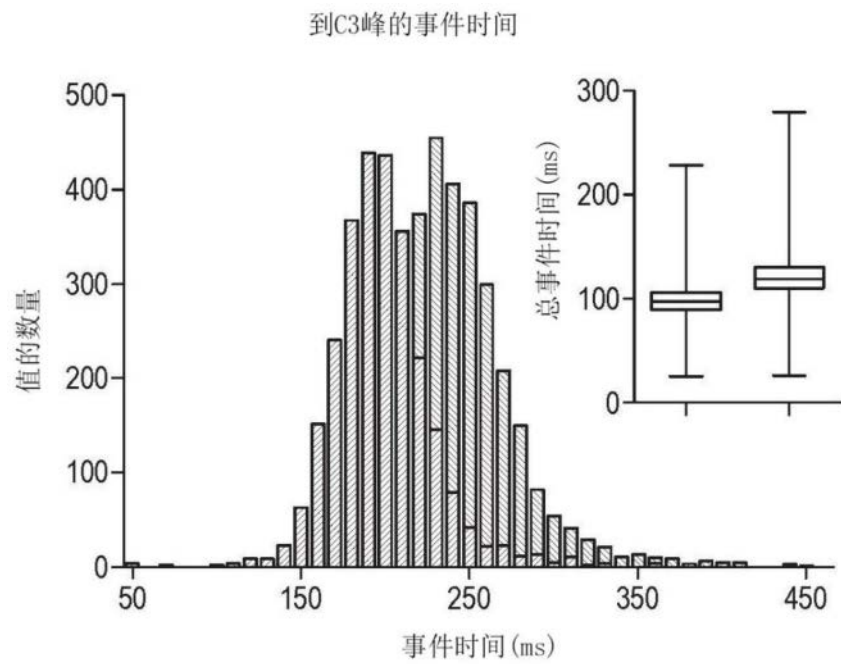


图22C

到C3峰直至事件结束的事件时间

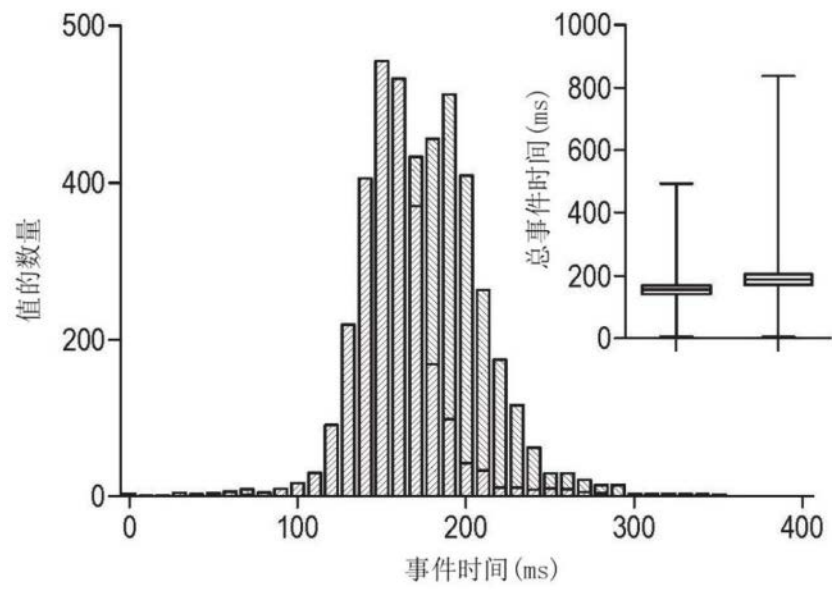


图22D

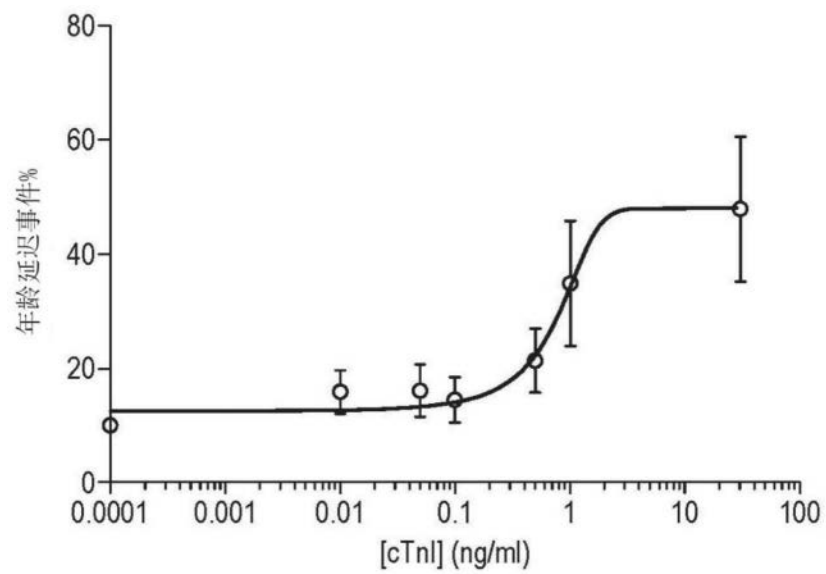


图22E

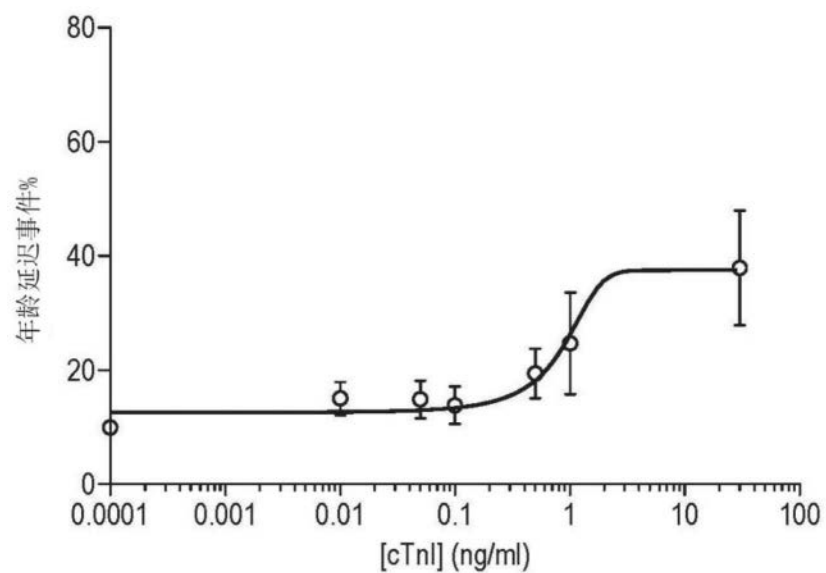


图22F

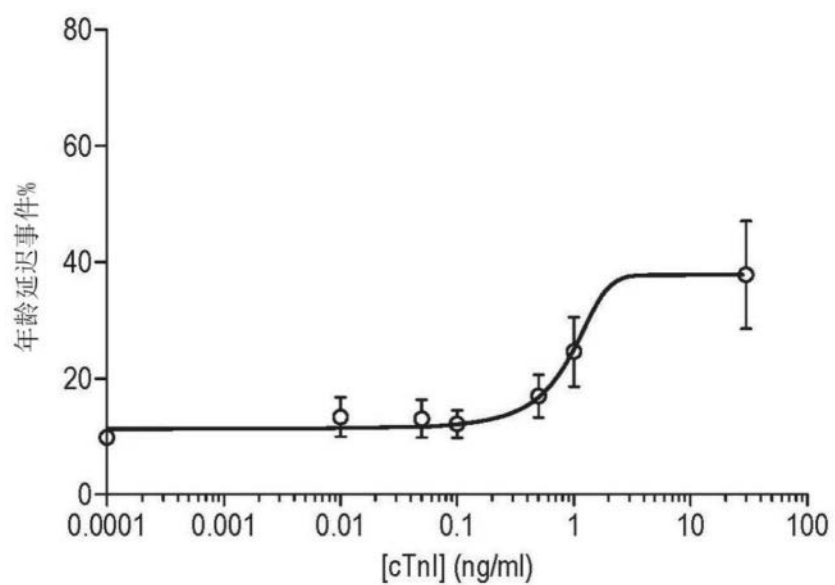


图22G