



(10) 申请公布号 CN 119998448 A

(43) 申请公布日 2025.05.13

(21) 申请号 202380069481.9

(22) 申请日 2023.09.29

(30) 优先权数据

2022-156368 2022.09.29 JP

(85) PCT国际申请进入国家阶段日

2025.03.27

(86) PCT国际申请的申请数据

PCT/JP2023/035799 2023.09.29

(87) PCT国际申请的公布数据

W02024/071424 JA 2024.04.04

(71) 申请人 庆应义塾

地址 日本

申请人 JSR株式会社

(72) 发明人 佐谷秀行 加藤翔一 池本笃史

(74) 专利代理机构 北京集佳知识产权代理有限公司 11227

专利代理师 金世煜 李书慧

(51) Int.Cl.

*C12N 15/09* (2006.01)

*C12N 5/10* (2006.01)

*C12N 15/113* (2006.01)

*C12N 15/12* (2006.01)

*C12N 15/55* (2006.01)

*C12N 15/63* (2006.01)

*C12N 15/86* (2006.01)

权利要求书3页 说明书34页

序列表(电子公布) 附图12页

(54) 发明名称

引起细胞内的响应的功能性分子的探索方法

(57) 摘要

用于分析功能性分子的功能的载体或载体组,包含:编码功能性分子的候补分子的表达系统的多核苷酸、编码由既定的细胞内的响应而激活的转录调控序列或翻译调控序列的多核苷酸、以及编码可操作地与该转录调控序列或翻译调控序列连接的报告基因系统的多核苷酸。一种引起细胞内的响应的功能性分子的探索方法,包括如下步骤:将该载体或载体组导入至细胞内、以及测定该细胞中的该载体或载体组所含有的报告基因系统的表达。

1. 一种载体或载体组,用于分析功能性分子的功能,  
包含:编码功能性分子的候选分子的表达系统的多核苷酸、编码由既定的细胞内的响应而激活的转录调控序列或翻译调控序列的多核苷酸、以及编码可操作地与该转录调控序列或翻译调控序列连接的报告基因系统的多核苷酸。
2. 根据权利要求1所述的载体或载体组,其中,所述功能性分子为RNA或肽。
3. 根据权利要求2所述的载体或载体组,其中,所述RNA为shRNA或者向导RNA。
4. 根据权利要求1所述的载体或载体组,其中,所述由既定的细胞内的响应而激活的转录调控序列或翻译调控序列为p53结合序列、抗氧化剂响应序列、核因子 $\kappa$ B结合序列、转录激活因子6响应序列、金属调节元件、热休克元件、缺氧反应元件、外源物质响应元件、活化T细胞核因子结合序列、TCF/LEF型DNA结合蛋白结合序列、sesRNA、以及toehold RNA。
5. 根据权利要求4所述的载体或载体组,其中,所述转录调控序列激活配置于其下游的启动子。
6. 根据权利要求5所述的载体或载体组,其中,所述报告基因系统包含所述启动子。
7. 根据权利要求1所述的载体或载体组,其中,所述报告基因系统随着所述转录调控序列或翻译调控序列的激活而表达。
8. 根据权利要求1所述的载体或载体组,其中,所述报告基因系统为发光或荧光蛋白报告基因系统、或者酶报告基因系统。
9. 根据权利要求8所述的载体或载体组,其中,所述酶报告基因系统包含具有改变靶核苷酸序列的作用的蛋白质与其目标序列的组合。
10. 根据权利要求8所述的载体或载体组,其中,所述酶报告基因系统表达DNA核酸酶、DNA重组酶、转座酶、或者整合酶。
11. 根据权利要求8所述的载体或载体组,其中,所述酶报告基因系统包含人工核酸酶与其识别序列的组合、限制酶与对应的限制酶识别序列的组合。
12. 根据权利要求8所述的载体或载体组,其中,所述酶报告基因系统包含:Cas家族与向导RNA与Cas识别序列的组合、TALEN与其识别序列的组合、ZFN与其识别序列的组合、或者碱基编辑器或先导编辑器与其识别序列的组合。
13. 根据权利要求11所述的载体或载体组,其中,所述限制酶为I—CeuI、I—SceI、PI—PspI或PI—SceI。
14. 根据权利要求8所述的载体或载体组,其中,所述酶报告基因系统包含:CRE—LoxP系统、CRE—LoxP系统中代替LoxP序列或除LoxP序列以外还具有LoxP变体的序列的系统、Vika—Vox系统、Dre—rox系统、或者Flp—FRT系统。
15. 根据权利要求1所述的载体或载体组,其中,所述载体为质粒载体、线性DNA序列、转座子载体、或病毒载体。
16. 根据权利要求15所述的载体或载体组,其中,所述病毒载体为慢病毒载体、腺相关病毒载体、杆状病毒载体、MMLV逆转录病毒载体、或MSCV逆转录病毒载体。
17. 根据权利要求1所述的载体或载体组,其中,进一步具有固有的分子条形码序列。
18. 一种引起细胞内的响应的功能性分子的探索方法,包括如下步骤:  
将权利要求1~17中任一项所述的载体或载体组导入到细胞中,以及测定在该细胞中的该载体或载体组所含有的报告基因系统的表达。

19. 根据权利要求18所述的方法,其中,所述报告基因系统的表达通过发光或荧光测定法、蛋白质定量法、定量RT-PCR、定量PCR、或测序进行定量。

20. 根据权利要求18所述的方法,其中,包括如下步骤:在细胞群中添加2种以上的载体或载体组而将各载体导入到该细胞群中的任一细胞中,

该2种以上的载体或载体组各自包含编码彼此不同的功能性分子的多核苷酸。

21. 根据权利要求18所述的方法,其中,所述细胞为采取自人或非人动物的细胞、采取自患者或实验动物的样本来源的细胞、原代培养细胞、类器官、细胞球状体、永生化培养细胞、或者实验动物个体的体内的细胞。

22. 一种参与细胞内的响应的基因的探索方法,包括如下步骤:

将多个载体或载体组导入至细胞群中,其中,该多个载体或载体组各自编码:干扰受试对象基因的RNA、由既定的细胞内的响应而激活的转录调控序列或翻译调控序列、以及可操作地与该转录调控序列或翻译调控序列连接的报告基因系统,但是,该多个载体或载体组各自所编码的该RNA相互不同;

对该细胞群中的该多个载体或载体组各自所编码的报告基因系统的表达分别进行测定;以及

基于该多个载体或载体组各自所包含的报告基因系统的表达来确定该受试对象基因是否为参与该既定的细胞内的响应的基因。

23. 根据权利要求22所述的方法,其中,基于针对所述多个载体或载体组各自所包含的报告基因系统的表达的统计值来确定所述受试对象基因是否为参与所述既定的细胞内的响应的基因。

24. 根据权利要求22所述的方法,其中,所述干扰受试对象基因的RNA为shRNA或gRNA。

25. 一种试剂盒,包含权利要求1~17中任一项所述的载体或载体组。

26. 根据权利要求25所述的试剂盒,其为疾病的诊断药或伴随诊断药。

27. 一种疾病的诊断方法,包括如下步骤:实施权利要求18~21中任一项所述的引起细胞内的响应的功能性分子的探索方法。

28. 一种伴随诊断方法,包括如下步骤:实施权利要求18~21中任一项所述的引起细胞内的响应的功能性分子的探索方法。

29. 一种用于分析对于细胞内的响应的转录调控序列或翻译调控序列的活性的方法,包括如下步骤:

制备导入了载体的细胞,其中,

该载体包含:编码转录调控序列或翻译调控序列的多核苷酸、编码可操作地与该转录调控序列或翻译调控序列连接的报告基因系统的多核苷酸、以及识别该转录调控序列或翻译调控序列的标记序列,

该报告基因系统包含具有改变靶核苷酸序列的作用的蛋白质及其目标序列,通过该报告基因系统的表达而该识别序列发生改变;

在该细胞中引发既定的细胞内的响应;

引发该响应后,对该载体中的包含该目标序列和该标记序列的区域进行分析,将该目标序列的变化进行定量;

基于该标记序列将该目标序列与该转录调控序列或翻译调控序列关联起来;

基于该目标序列的变化量,评价与该目标序列产生了关联的转录调控序列或翻译调控序列的对于该既定的细胞内的响应的活性。

30. 根据权利要求29所述的方法,其中,所述载体包含:编码转录调控序列的多核苷酸、以及识别该转录调控序列的标记序列。

31. 根据权利要求29所述的方法,其中,通过测序来分析包含所述目标序列和标记序列的区域。

32. 根据权利要求29所述的方法,其中,包括如下步骤:将多个载体导入于细胞群中,该多个载体包含编码各自不同的转录调控序列或翻译调控序列的多核苷酸。

## 引起细胞内的响应的功能性分子的探索方法

### 技术领域

[0001] 本发明涉及引起细胞内的响应的功能性分子的探索方法、该探索方法中使用的载体或载体组、以及该探索方法的各种应用。

### 背景技术

[0002] 在疾病的产生、分化等各种生理过程中阐明各基因的功能在有效的的新药靶点的探索、生理过程的分子机制的理解上非常重要。

[0003] 基因的功能分析中有在特定状态的细胞中对靶基因的mRNA量、蛋白质量进行分析、或进行蛋白质的翻译后修饰的方法。作为更直接的功能分析方法,采用了基于靶基因的人为表达抑制或表达增强的表型分析。作为一般的表型分析法,有选择性抑制或激活任意靶基因并在细胞水平、组织水平或个体水平对其效果进行分析的方法。但是,在以往的表型分析中,基本上只能在1种实验条件下调查1种基因或任意1组基因组合,因此,表现型的综合分析需要大量的精力和时间。

[0004] 近年来,为了实现综合性的表型分析,开发出了将RNAi、CRISPR/Cas9系统与大规模并行测序技术组合而成的筛选方法(非专利文献1、2)。该方法中,集合对个别基因具有抑制效果的shRNA或gRNA载体而导入细胞、组织中,由此在同一实验条件下的各个细胞中产生个别的基因抑制。然后,富集呈现表型的细胞或未呈现表型的细胞,对得到的细胞群中的shRNA或gRNA的存在量通过大规模并行测序进行定量,由此筛选参与特定表型的责任基因。或者,不进行基于表型的细胞富集而对细胞进行单细胞测序分析,同时对单细胞中的gRNA的序列和基因表达进行分析,由此来实现综合性表型分析(非专利文献3~5)。将这样的基因抑制和大规模并行测序技术组合而成的表型分析与以往的表型分析相比实现了高通量的基因的功能分析。但是,这些分析方法也由于进行细胞群的富集而需要长期培养,富集操作、单细胞测序的文库制备中产生很大工作量,因此存在进一步改良的余地。

[0005] 作为用于基因的功能分析的工具,市售有共表达编码shRNA的DNA和使将载体导入细胞可视化的报告基因的慢病毒载体(例如, System Biosciences, LLC的shRNA克隆载体等)。另外,以往已知有使用与报告基因连接的响应DNA损伤、氧化等细胞应激而激活的转录调节因子(例如p53反应元件)作为标志物的用于分析细胞的应激状态、药物响应的报告基因检测(非专利文献6)。

[0006] 非专利文献7和专利文献1中记载了一种酶活性的评价方法,将包含编码已知启动子或已知酶的基因的变体和与其连接的Canvas序列的载体导入至细胞中,利用由该变体表达的酶在细胞内促进致突变蛋白的表达来诱导该Canvas序列的突变,根据该Canvas序列的突变的多少来评价该变体所编码的酶的活性。非专利文献8中记载了如下方法:将抑制特定转录因子的表达的gRNA和与条形码序列连接的增强子导入至细胞中,基于条形码序列将gRNA与增强子活性关联起来,由此评价相对于增强子的转录因子的作用。

[0007] 现有技术文献

[0008] 专利文献

- [0009] 专利文献1:国际公开公报第2023/028476号
- [0010] 非专利文献
- [0011] 非专利文献1:Cell,2013,152(4):909-922
- [0012] 非专利文献2:Nature biotechnology,2016,34(6):634-636
- [0013] 非专利文献3:Cell,2016,167(7):1853-1866
- [0014] 非专利文献4:Cell,2016,167(7):1883-1896
- [0015] 非专利文献5:Nature methods,2017,14(3):297-301
- [0016] 非专利文献6:Mutation Research,Genetic Toxicology and Environmental Mutagenesis,2010,696(1):21-40.
- [0017] 非专利文献7:bioRxiv preprint,doi:doi.org/10.1101/2022.03.09.483646, Version 1,2022年3月9日
- [0018] 非专利文献8:bioRxiv preprint,doi:doi.org/10.1101/2020.09.30.321323, 2020

## 发明内容

- [0019] 希望提高RNA、肽等功能性分子的综合分析的效率 and 通量。
- [0020] 本发明人构建了一种载体,其包含编码要分析的RNA或肽的表达系统的多核苷酸、以及编码因可由该RNA或肽的扰动引起的细胞内的响应而激活的转录调控序列和可操作地与该转录调控序列连接的报告基因系统的多核苷酸。也可以利用将各个多核苷酸配置于不同载体的载体组。将该载体或载体组导入于细胞时,如果伴随所表达的该RNA或肽的扰动在细胞内产生该响应,则该报告基因系统会与其对应地进行表达。因此,可以基于该细胞内的该报告基因系统的表达,将该RNA或肽与该细胞内的响应产生关联,进而分析其功能。此外,发现:如果将包含各自不同的RNA或肽、或者报告基因系统的多个上述载体或载体组导入于1个细胞或细胞群,并且对来自各载体的报告基因系统的表达分别进行测定,则可以在1个细胞或细胞群中一次性分析多个RNA或肽的功能。
- [0021] 因此,本发明提供以下方案。
- [0022] (1)一种载体或载体组,用于分析功能性分子的功能,包含:编码功能性分子的候补分子的表达系统的多核苷酸、编码由既定的细胞内的响应而激活的转录调控序列或翻译调控序列的多核苷酸、以及编码可操作地与该转录调控序列或翻译调控序列连接的报告基因系统的多核苷酸。
- [0023] (2)根据(1)所述的载体或载体组,其中,上述功能性分子为RNA或肽。
- [0024] (3)根据(1)或(2)所述的载体或载体组,其中,上述RNA为shRNA或向导RNA。
- [0025] (4)根据(1)~(3)中任一项所述的载体或载体组,其中,上述由既定的细胞内的响应而激活的转录调控序列或翻译调控序列为p53结合序列、抗氧化剂响应序列、核因子 $\kappa$ B结合序列、转录激活因子6响应序列、金属调节元件(Metal Regulatory Element)、热休克元件(Heat Shock Element)、缺氧反应元件(Hypoxia Response Element)、活化T细胞核因子结合序列、TCF/LEF型DNA结合蛋白结合序列、sesRNA、或toehold RNA。
- [0026] (5)根据(1)~(4)中任一项的载体或载体组,其中,上述转录调控序列激活配置于其下游的启动子。

- [0027] (6)根据(5)所述的载体或载体组,其中,上述报告基因系统包含上述启动子。
- [0028] (7)根据(1)~(6)中任一项所述的载体或载体组,其中,上述报告基因系统随着上述转录调控序列或翻译调控序列的激活而表达。
- [0029] (8)根据(1)~(7)中任一项所述的载体或载体组,其中,上述报告基因系统为发光或荧光蛋白报告基因系统、或者酶报告基因系统。
- [0030] (9)根据(8)所述的载体或载体组,其中,上述酶报告基因系统包含具有改变靶核苷酸序列的作用的蛋白质与其目标序列的组合。
- [0031] (10)根据(8)所述的载体或载体组,其中,上述酶报告基因系统表达DNA核酸酶、DNA重组酶、转座酶、或整合酶。
- [0032] (11)根据(8)所述的载体或载体组,其中,上述酶报告基因系统包含:人工核酸酶与其识别序列的组合、限制酶与对应的限制酶识别序列的组合。
- [0033] (12)根据(8)所述的载体或载体组,其中,上述酶报告基因系统包含:Cas家族与向导RNA与Cas识别序列的组合、TALEN(转录激活因子样效应物核酸酶)与其识别序列的组合、ZFN与其识别序列的组合、或者碱基编辑器(Base Editor)或先导编辑器(Prime Editor)与其识别序列的组合。
- [0034] (13)根据(11)所述的载体或载体组,其中,上述限制酶为I—CeuI、I—SceI、PI—PspI或PI—SceI。
- [0035] (14)根据(8)所述的载体或载体组,其中,上述酶报告基因系统包含:CRE—LoxP系统、CRE—LoxP系统中代替LoxP序列或除此以外还具有LoxP变体的序列的系统、Vika—Vox系统、Dre—rox系统、或Flp—FRT系统。
- [0036] (15)根据(1)~(14)中任一项所述的载体或载体组,其中,上述载体为质粒载体、线性DNA序列、转座子载体、或病毒载体。
- [0037] (16)根据(15)所述的载体或载体组,其中,上述病毒载体为慢病毒载体、腺相关病毒载体、杆状病毒载体、MMLV逆转录病毒载体、或MSCV逆转录病毒载体。
- [0038] (17)根据(1)~(16)中任一项所述的载体或载体组,其中,进一步具有固有的分子条形码序列。
- [0039] (18)一种引起细胞内的响应的功能性分子的探索方法,包括如下步骤:
- [0040] 将(1)~(17)中任一项所述的载体或载体组导入至细胞中,以及
- [0041] 测定该细胞中的该载体或载体组所含有的报告基因系统的表达。
- [0042] (19)根据(18)所述的方法,其中,上述报告基因系统的表达通过发光或荧光测定法、蛋白质定量法、定量RT—PCR、定量PCR、或测序进行定量。
- [0043] (20)根据(18)或(19)所述的方法,其中,包括如下步骤:在细胞群中添加2种以上的载体或载体组而将各载体导入于该细胞群中的任一细胞中,该2种以上的载体或载体组各自包含编码彼此不同的功能性分子的多核苷酸。
- [0044] (21)根据(18)~(20)中任一项所述的方法,其中,上述细胞为采取自人或非人动物的细胞、采取自患者或实验动物的样本来源的细胞、原代培养细胞、类器官、细胞球状体、永生化培养细胞、或实验动物个体的体内的细胞。
- [0045] (22)一种参与细胞内的响应的基因的探索方法,包括如下步骤:
- [0046] 将多个载体或载体组导入于细胞群中,这里,该多个载体或载体组各自编码:干扰

受试对象基因的RNA、由既定的细胞内的响应而激活的转录调控序列或翻译调控序列、以及可操作地与该转录调控序列或翻译调控序列连接的报告基因系统,但是,该多个载体或载体组各自所编码的该RNA相互不同;

[0047] 对该细胞群中的该多个载体或载体组各自所编码的报告基因系统的表达分别进行测定;以及

[0048] 基于该多个载体或载体组各自所包含的报告基因系统的表达来确定该受试对象基因是否为参与该既定的细胞内的响应的基因。

[0049] (23)根据(22)所述的方法,其中,基于针对上述多个载体或载体组各自所包含的报告基因系统的表达的统计值,来确定上述受试对象基因是否为参与上述既定的细胞内的响应的基因。

[0050] (24)根据(22)或(23)所述的方法,其中,上述干扰受试对象基因的RNA为shRNA或gRNA。

[0051] (25)一种试剂盒,包含(1)~(17)中任一项所述的载体或载体组。

[0052] (26)根据(25)所述的试剂盒,其为疾病的诊断药或伴随诊断药。

[0053] (27)一种疾病的诊断方法,包括如下步骤:实施(18)~(21)中任一项所述的引起细胞内的响应的功能性分子的探索方法。

[0054] (28)一种伴随诊断方法,包括如下步骤:实施(18)~(21)中任一项所述的引起细胞内的响应的功能性分子的探索方法。

[0055] (29)(1)~(17)中任一项所述的载体或载体组在引起细胞内的响应的功能性分子的探索中的使用。

[0056] (30)(1)~(17)中任一项所述的载体或载体组在引起细胞内的响应的功能性分子的探索中的使用。

[0057] (31)(1)~(17)中任一项所述的载体或载体组在疾病诊断中的使用。

[0058] (32)(1)~(17)中任一项所述的载体或载体组在伴随诊断中的使用。

[0059] (33)(1)~(17)中任一项所述的载体或载体组在疾病的诊断药或伴随诊断药的制造中的使用。

[0060] (34)一种用于分析对于细胞内的响应的转录调控序列或翻译调控序列的活性的方法,包括如下步骤:

[0061] 制备导入了载体的细胞,其中,

[0062] 该载体包含:编码转录调控序列或翻译调控序列的多核苷酸、编码可操作地与该转录调控序列或翻译调控序列连接的报告基因系统的多核苷酸、以及识别该转录调控序列或翻译调控序列的标记序列,

[0063] 该报告基因系统包含具有改变靶核苷酸序列的作用的蛋白质及其目标序列,通过该报告基因系统的表达而该识别序列发生变化;

[0064] 在该细胞中引发既定的细胞内的响应;

[0065] 引发该响应后,对该载体中的包含该目标序列和该标记序列的区域进行分析,将该目标序列的变化进行定量;

[0066] 基于该标记序列将该目标序列与该转录调控序列或翻译调控序列关联起来,

[0067] 基于该目标序列的变化量,评价与该目标序列产生了关联的转录调控序列或翻译

调控序列的对于该既定的细胞内的响应的活性。

[0068] (35)根据(34)所述的方法,其中,上述载体包含:编码转录调控序列的多核苷酸、和识别该转录调控序列的标记序列。

[0069] (36)根据(34)或(35)所述的方法,其中,通过测序对包含上述目标序列和标记序列的区域进行分析。

[0070] (37)根据(34)~(36)中任一项所述的方法,其中,包括如下步骤:将多个载体导入于细胞群中,

[0071] 该多个载体包含编码各自不同的转录调控序列或翻译调控序列的多核苷酸。

[0072] 本发明提供对引起细胞内的响应的功能性分子的探索有用的载体或载体。该载体或载体被设计成在由编码于载体中的RNA、肽等分子在细胞内引起规定的响应时,编码于相同的载体或载体组中的报告基因系统进行表达。因此,根据使用该载体或载体的本发明的功能性分子的探索方法,可以基于该报告基因系统的表达而将该分子与该响应关联起来,对该分子的功能进行分析。即,可以利用简便的步骤来推断RNA、肽等分子的功能。另外,根据本发明,能够在1个细胞或细胞群中一次性分析多个分子的功能,分析的通量(throughput)提高。

#### 附图说明

[0073] 图1A是质粒1和2的结构。

[0074] 图1B是质粒3的结构。

[0075] 图1C是质粒1~3中的报告基因系统的行为。

[0076] 图1D是质粒4的结构。

[0077] 图1E是线性DNA1的结构。

[0078] 图1F是质粒5的结构。

[0079] 图2是导入了质粒1的细胞中的基于阿霉素处理的荧光表达。纵轴:红色荧光阳性细胞/绿色荧光阳性细胞,横轴:阿霉素浓度和阿霉素处理后的时间。各点表示由各培养孔得到的红色荧光阳性细胞数除以绿色荧光阳性细胞数所得的值。条形表示平均值。

[0080] 图3示出阿霉素处理后的质粒1中的mKate2序列的倒位率(RQ)。

[0081] 图4示出阿霉素处理后的导入了质粒1和质粒2的细胞中的红色荧光阳性细胞/绿色荧光阳性细胞比。DMSO:未添加阿霉素,Doxo:添加阿霉素.shCon:导入shScramble(质粒1)的细胞,shRAD51:导入shRAD51(质粒2)的细胞。各点表示由各培养孔得到的红色荧光阳性细胞数除以绿色荧光阳性细胞数所得的值。条形表示平均值。

[0082] 图5示出TBHQ处理后的质粒3中的mKate2序列的倒位率。Flip:倒位序列比率,Original:非倒位序列比率。横轴为TBHQ浓度。

[0083] 图6示出CHIR99021处理后6小时和24小时时的细胞内Wnt通路的激活。上:AXIN2表达,下:SP5表达。表达量为相对于CHIR996021未处理(0 $\mu$ M)的相对值。横轴:CHIR99021浓度。

[0084] 图7示出CHIR99021处理后的载体中的标志物序列倒位量。Linear:线性DNA1,Plasmid:质粒4。横轴为CHIR99021浓度。表达量为相对于CHIR99021未处理(0 $\mu$ M)的相对值。

[0085] 图8示出导入病毒载体的细胞中的shRNA所诱导的mKate2序列的倒位量。Scramble:导入病毒载体1A的细胞,shRAD51:导入病毒载体1B的细胞。上:MOI 0.4的在细胞

中的表达量,下:MOI 2的在细胞中的表达量。

[0086] 图9是实施例7中倒位DNA序列的定量所使用的扩增子的示意图。shCon:包含shScramble的扩增子,shRad\_\_2和shRad\_\_4:包含shRAD51\_\_2的扩增子和包含shRAD51\_\_4的扩增子。

[0087] 图10示出导入了质粒集合的细胞中的各质粒中的mKate2序列的倒位率(倒位/非倒位)。左:未添加阿霉素,右:添加1 $\mu$ M阿霉素。横轴的编号为孔编号。shControl:包含shScramble的质粒,shRad51\_\_2和shRad51\_\_4:包含shRAD51\_\_2的质粒和包含shRAD51\_\_4的质粒。

[0088] 图11A示出导入了shRNA集合化载体线性DNA2的细胞来源的DNA中的各载体上的标志物序列的倒位几率。scramble:导入了shScramble的细胞,Targeting:导入了基因靶向shRNA的细胞。横轴是表达的shRNA所靶向的基因和以各基因为靶点的shRNA的ID号。

[0089] 图11B是针对每种shRNA的靶基因,将导入了线性DNA2的细胞来源的DNA中的各载体上的标志物序列的倒位几率平均化的结果。横轴是shRNA所靶向的基因。图表中的图形为在与各shRNA相同的载体上的标志物序列的倒位几率。

[0090] 图12是导入了shRNA集合化载体线性DNA3的细胞来源的DNA中的各载体上的标志物序列的倒位和缺失几率。横轴为表达的shRNA所靶向的基因和以各基因为靶点的shRNA的ID号。

[0091] 图13是线性DNA4的结构。

[0092] 图14A是线性DNA4A的制备顺序。

[0093] 图14B是线性DNA4B的制备顺序。

[0094] 图15是质粒6的结构。

[0095] 图16是导入了shRNA集合化载体线性DNA4的细胞来源的DNA中的各载体上的标志物序列的倒位几率。横轴是表达的shRNA所靶向的基因和以各基因为靶点的shRNA的ID号。左:无UMI的过滤,右:有UMI的过滤。

[0096] 图17是导入了minP变体序列的细胞中的mKate2序列的倒位几率。Original:原始minP,Variant:minP变体。

## 具体实施方式

[0097] 本说明书中引用的所有专利文献、非专利文献和其它发行物整体作为参考援引于本说明书中。

[0098] 本说明书中例示的各成分只要没有特别说明,可以分别单独使用1种,或者可以组合2种以上使用。

[0099] 本说明书中,“A~B”等表示数值范围的记载与“A以上B以下”含义相同,该数值范围包含A和B。

[0100] 本说明书中例示的蛋白质在与突变前具有同等功能时包含其氨基酸序列或编码该氨基酸序列的碱基序列进行了置换、重复、缺损、插入或移码的突变蛋白质。

[0101] 本说明书中,编码RNA、肽、报告基因等的序列与包含启动子等转录调控序列和翻译调控序列的调控序列的“可操作地连接”是指上述RNA、肽、报告基因等以可在该调控序列的控制下表达的方式进行功能性连接。表达的序列与调控序列的“可操作地连接”的步骤是

本领域技术人员所公知的。例如，“转录调控序列与报告基因可操作地连接”是指报告基因在转录调控序列的控制下被转录，另外，“翻译调控序列与报告基因可操作地连接”是指报告基因在翻译调控序列的控制下被翻译。因此，本说明书中的上述编码序列与上述调控序列的“可操作地连接”不仅是指该编码序列被存在于同一核苷酸链中的调控序列所控制表达的情况，也包含该编码序列和控制其表达的调控序列存在于不同的核苷酸链中的情况。

[0102] 本说明书中，基因或其表达产物的“表达”包含基因的转录和转录物的翻译。例如，“表达”包含从基因转录为mRNA、从mRNA翻译为蛋白质、以及蛋白质本身和编码该蛋白质的mRNA的表达。

[0103] 本说明书中，涉及核苷酸序列的“上游”和“下游”分别是指该核苷酸序列的DNA正义链的5'侧和3'侧的区域。

[0104] 本说明书中，细胞内的“扰动”是指通过与靶标（例如基因、其mRNA）互补的RNA的表达而诱导的该靶标的敲低、敲除或表达激活，或者肽等功能性分子的细胞内的表达。作为诱导扰动的RNA，可举出shRNA、gRNA等。

[0105] 本说明书中，细胞内的“响应”是指由扰动引起的细胞内事件，例如，可举出由上述RNA、肽的表达而引起的细胞内的DNA损伤、氧化等应激、既定的基因的表达量或活性的变化或者由其引起的细胞内环境的变化等。

[0106] 本说明书中的“功能性分子”是指在细胞内诱导扰动的分子。该扰动引起细胞内的响应时，该功能性分子引起细胞内的响应。

[0107] 本说明书中的“多核苷酸”只要能够保有所编码的RNA或肽的信息，就可以为单链DNA、双链DNA、单链RNA、双链RNA、或其它的人工核酸。其中，从能够稳定保存信息的观点出发，优选双链DNA。

[0108] （引起细胞内的响应的功能性分子的探索方法）

[0109] 本发明提供引起细胞内的响应的功能性分子的探索方法。作为本方法所要探索的目标即引起细胞内的响应的功能性分子（以下，也称为“目标分子”）的例子，可举出低分子量发夹RNA（shRNA）、向导RNA（guide RNA, gRNA）、增强子RNA、小激活RNA（small activating RNA）、长链非编码RNA（long noncoding RNA）等RNA和肽。

[0110] 本发明的方法中所探索的shRNA是具有茎环结构的RNA，并且是该茎环结构由40~100个核苷酸构成、其茎部分由19~45个核苷酸构成、其环部分由4~25个核苷酸构成的RNA。优选为茎环结构由45~75个核苷酸构成、其茎部分由19~30个核苷酸构成、其环部分由6~15个核苷酸构成的RNA。该shRNA的该茎部分的序列与细胞内的核苷酸序列（mRNA、基因组DNA、基因组RNA等）具有互补性，可以结合于该核苷酸序列。

[0111] 本发明的方法中所探索的向导RNA（gRNA）是指与目标序列互补的具有10~250个碱基的序列的RNA。

[0112] 本发明的方法中所探索的增强子RNA的长度为15~2000个核苷酸。

[0113] 本发明的方法中所探索的小激活RNA是指具有与靶基因的从转录起始位点的2000bp上游到转录起始位点的2000bp下游之间互补的14~30核苷酸的序列的RNA。

[0114] 本发明的方法中所探索的长链非编码RNA是指200个核苷酸以上的不具有编码区域的RNA序列。

[0115] 作为本发明的方法中所探索的肽，其分子尺寸没有特别限定，可举出蛋白质、多肽

或寡肽。本发明的方法中所探索的目标肽例如为可以在细胞内作为信号因子、酶、辅酶等发挥作用并引起细胞内扰动的肽。

[0116] 本发明的上述功能性分子的探索方法基本上包括对作为上述目标分子的候补的RNA、肽等分子进行筛选。从它们中确定作为目标的引起特定的细胞内的响应的功能性分子。该方法的特征在于使用由本发明提供的载体或载体组,该载体或载体组包含:编码作为上述目标分子的候补的(即与细胞内响应的相关性未知)功能性分子(以下,也称为“候补分子”)的表达系统的多核苷酸、编码可由该目标分子引起的(即想要调查与该候补分子的相关性的)、由既定的细胞内的响应而激活的调控序列(转录调控序列或翻译调控序列)的多核苷酸、以及编码可操作地与该调控序列连接的报告基因系统的多核苷酸。

[0117] 本发明中,编码上述候补分子的表达系统的多核苷酸、编码上述由既定的细胞内的响应而激活的调控序列的多核苷酸、以及编码可操作地与编码调控序列连接的报告基因系统的多核苷酸可以配置于同一载体上,或者也可以配置于不同载体上。各个载体上的编码该候补分子的表达系统的多核苷酸、编码由该既定的细胞内的响应而激活的调控序列的多核苷酸、或者编码该报告基因系统的多核苷酸的配置只要该候补分子表达,由该细胞内的响应而激活该调控序列,并且由该调控序列的激活而促进该报告基因系统的表达,就没有特别限定。

[0118] 一个实施方式中,提供一种载体,其配置有编码上述候补分子的表达系统的多核苷酸、以及从上游依次排列的编码上述调控序列的多核苷酸和编码上述报告基因系统的多核苷酸。另一实施方式中,提供一种载体组,其包含配置有编码上述候补分子的表达系统的多核苷酸的载体以及从上游起依次配置有编码上述调控序列的多核苷酸和编码上述报告基因系统的多核苷酸的载体。另一实施方式中,提供一种载体组,其包含配置有编码上述候补分子的表达系统的多核苷酸和编码上述报告基因系统的一部分(例如,后述的酶识别序列、LoxP或其变体和标志物序列、或者由转座酶或整合酶转入的序列)的多核苷酸的载体,以及从上游起依次配置有编码上述调控序列的多核苷酸和编码上述报告基因系统的剩余部分的多核苷酸的载体。

[0119] 一个实施方式中,作为编码上述候补分子的表达系统的多核苷酸,可举出编码RNA表达系统的多核苷酸。例如,该表达系统包含:编码包含作为上述候补分子的RNA序列的双链RNA(例如shRNA)的序列,和与其可操作地连接的RNAPo1 III启动子(例如U6启动子、H1启动子等)或RNAPo1 II启动子(CMV启动子等)等转录调控序列。上述候补分子为gRNA的情况下,也可以共表达用于使该gRNA发挥功能的蛋白质(例如,Cas9蛋白质、Cas9切口酶、dCas9/VP64、dCas9/VPR、dCas9/KRAB、dCas9/KRAB/MeCP2、碱基编辑器等)。另一实施方式中,编码上述候补分子的表达系统的多核苷酸编码肽的表达系统,所述肽的表达系统包含编码作为上述候补分子的肽的序列和与其可操作地连接的启动子等转录调控序列。编码目标RNA或目标肽的核苷酸序列、以及包含这些表达系统的载体可以通过市售的服务而合成和购入。

[0120] 作为上述由既定的细胞内的响应而激活的调控序列,可举出转录调控序列和翻译调控序列。作为该转录调控序列,可举出通过想要调查是否由上述候补分子引起的细胞内的响应、例如DNA损伤响应、氧化应激反应或蛋白质表达、例如Wnt通路等信号通路激活等而直接或间接激活的转录调控序列。作为这样的转录调控序列的更具体的例子,可举出p53结合序列(p53RE)、抗氧化剂响应序列(ARE)、核因子 $\kappa$ B(NF- $\kappa$ B)结合序列、转录激活因子6响

应序列(ATF6 ERSE)、金属调节元件(MRE)、热休克元件(HSE)、缺氧反应元件(HRE)、外源物质响应元件(XRE, Xenobiotic Responsive Element)、活化T细胞核因子(NFAT)结合序列、和TCF/LEF型DNA结合蛋白结合序列。这些转录调控序列激活配置于其下游的启动子,该启动子促进后述的报告基因系统的表达。该启动子只要能够受到该转录调控序列的控制而激活,则其种类就没有限定。该启动子可以包含于编码该转录调控序列的多核苷酸中,或者也可以包含于后述的报告基因系统中。

[0121] 上述由既定的细胞内的响应而激活的翻译调控序列例如是随着想要调查是否由上述候补分子引起的细胞内的响应而激活、并促进下游的编码RNA的翻译的序列。作为该翻译调控序列的更具体的例子,可举出sesRNA。sesRNA是通过任意的mRNA分子和细胞内的ADAR蛋白质的活性来使下游的编码RNA表达的模块,是mRNA表达的传感器(参照Nature, 2022, 610:713—721, Nature Biotechnology, 2023, 41:698—707, Nature Biotechnology, 2023, 41:482—487, WO/2023/077135, WO/2023/077138, US20230123513)。sesRNA包含与特定的目标mRNA互补的RNA序列。如果细胞内的ADAR蛋白质作用于sesRNA与该mRNA分子的双链,则下游所连接的编码RNA可以翻译。如果在sesRNA编码序列的下游配置编码报告基因系统的序列,则该报告基因系统的表达反映该目标mRNA的表达。

[0122] 作为上述翻译调控序列的其它具体例,可举出toehold RNA。toehold RNA包含与目标mRNA互补的序列、核糖体结合序列,并且与下游的编码RNA连接。toehold RNA通常形成发夹环,因此编码RNA不进行翻译,但如果与目标mRNA结合,则发夹消失,进行翻译(参照Toehold switch, Nature Biotechnology, 2022, 40:53—545)。如果在toehold RNA编码序列的下游配置编码报告基因系统的序列,则该报告基因系统的表达反映该目标mRNA的表达。

[0123] 本发明提供的载体所编码的上述报告基因系统可操作地与上述调控序列连接,通过该调控序列的激活而表达。因此,该报告基因系统的表达表明通过上述候补分子而产生了既定的细胞内的响应。通过测定该报告基因系统的表达,可以确定该候补分子是否是引起该既定的细胞内的响应的物质(即,是否为探索的目标分子)。

[0124] 作为上述报告基因系统,可举出发光或荧光蛋白报告基因系统、酶报告基因系统等。该发光或荧光蛋白报告基因系统表达发光或荧光蛋白。表达的发光或荧光蛋白作为标志物被检测。该酶报告基因系统表达生成可检测的标志物分子的酶。通过表达的该酶的作用而生成的分子作为标志物被检测。作为该酶报告基因系统所表达的酶的例子,可举出DNA核酸酶、DNA重组酶、水解酶(例如 $\beta$ -葡萄糖醛酸酶)、转座酶、整合酶等。其中,优选具有改变靶核苷酸序列的作用的蛋白质(DNA核酸酶、DNA重组酶、转座酶、整合酶等)。因此,作为该酶报告基因系统的一个例子,可举出具有改变该靶核苷酸序列的作用的蛋白质与该目标序列的组合。

[0125] 作为上述DNA核酸酶的例子,可举出Cas9核酸酶等Cas家族、转录激活因子样效应核酸酶(TALEN)、锌指核酸酶(Zinc Finger Nuclease, ZFN)等人工核酸酶、以及碱基编辑器、先导编辑器等;限制酶等。作为该限制酶,优选在切割位点形成粘性末端的限制酶,例如,可举出I—CeuI、I—SceI、PI—PspI和PI—SceI。

[0126] 作为使用了人工核酸酶的酶报告基因系统的例子,可举出人工核酸酶与其识别序列的组合,例如Cas9核酸酶等Cas家族与向导RNA(gRNA或sgRNA)与Cas识别序列的组合、改

造上述的人工核酸酶的核酸识别蛋白而碱基编辑器或先导编辑器与其识别序列的组合、TALEN与其识别序列的组合、或者ZFN与其识别序列的组合。优选该TALEN或ZFN的识别序列由30~50个核苷酸构成。作为使用限制酶的酶报告基因系统的例,可举出限制酶与对应的限制酶识别序列的组合。这些系统中,可以将由人工核酸酶或限制酶切割的DNA序列作为标志物进行检测。上述人工核酸酶的识别序列或限制酶识别序列可以配置在与含有编码该人工核酸酶或限制酶的多核苷酸的载体相同的载体上,或者也可以配置在不同载体上。例如,该识别序列可以配置在含有编码上述候补分子的表达系统的多核苷酸的载体上。作为检测被切割的DNA序列的方法,可举出大规模并行测序等下一代测序、PCR等。

[0127] 作为使用了上述DNA重组酶的酶报告基因系统的例子,可举出CRE—LoxP系统。该系统由CRE重组酶、2个LoxP序列和配置于它们之间的标志物序列构成。该LoxP序列和标志物序列可以配置在与含有编码该CRE重组酶的多核苷酸的载体相同的载体上,或者也可以配置在不同载体上。例如,该LoxP序列和标志物序列可以配置在含有编码上述候补分子的表达系统的多核苷酸的载体上。这2个LoxP序列可以在相同方向串联配置,或者也可以在相反方向串联配置。LoxP序列以相同方向配置时,其间的标志物序列在CRE的作用下被切出,另一方面,LoxP序列以相反方向配置时,其间的标志物序列在CRE的作用下发生倒位。因此,只要将切出的序列本身作为标志物序列进行检测,或者将通过切除序列而在载体上产生的序列作为标志物序列进行检测,或者对倒位的标志物序列进行检测即可。作为检测标志物序列的方法,可举出大规模并行测序等下一代测序、PCR等。

[0128] 或者,在配置成相反方向的2个LoxP序列之间与转录方向相反的方向上配置标志物序列,并将启动子连接在载体上的该LoxP序列的上游。在CRE不发挥作用的状态下,不表达该标志物,但在CRE的作用下标志物序列在LoxP序列间倒位,即在可转录的方向重新配置,则该标志物会在该启动子的作用下进行表达。只要检测所表达的标志物即可。该情况下,编码发光或荧光蛋白的序列由于容易检测而优选为该标志物序列。作为理想的启动子与标志物序列的组合的例子,可举出EF1启动子与mKate2编码序列的组合。

[0129] 作为具有与上述CRE—LoxP系统同样作用的报告基因系统的例子,可举出CRE—LoxP系统中代替LoxP序列或除此以外还具有LoxP变体的序列的系统、Vika—Vox系统、Dre—Rox系统、和Flp—FRT系统。作为LoxP变体,可举出Lox71、Lox66、Lox5171、Lox2722、和Loxm2。本说明书中,也将该CRE—LoxP系统和具有与其同样作用的报告基因系统统称为“CRE—LoxP样系统”。

[0130] 作为使用了上述转座酶的酶报告基因系统的例子,可举出piggyBac转座子系统。该系统包含piggyBac转座酶与2个反向重复序列(末端反向重复序列,Inverted Terminal Repeat Sequence; ITR)的组合。这2个ITR以彼此相反的方向串联配置,并且在其间可配置任意标志物序列。该ITR和标志物序列可以配置在与含有编码该转座酶的多核苷酸的载体相同的载体上,或者也可以配置在不同载体上。例如,该ITR和标志物序列可以配置在含有编码上述候补分子的表达系统的多核苷酸的载体上。只要检测在该转座酶的作用下的这2个ITR间的标志物序列的易位(该标志物序列从该ITR间消失)即可。或者,只要在载体中设置TTAA序列并检测易位到该TTAA序列的位置的标志物序列即可。又或者,只要检测易位到基因组上的TTAA序列的标志物序列即可。一个实施方式中,配置编码上述候补分子的表达系统的多核苷酸作为该标志物序列,并检测该多核苷酸从ITR间消失、或该多核苷酸易位到

该TTAA序列的位置即可。作为检测标志物序列的方法,可举出大规模并行测序等下一代测序、PCR等。

[0131] 作为使用了上述整合酶的酶报告基因系统的例子,可举出phiC31/attP系统。该系统包含:包含phiC31整合酶的载体、以及包含attB位点的供体载体。phiC31整合酶对宿主基因组的attP位点和供体载体的attB位点进行重组,将供体载体整合到基因组中。以供体载体的序列为标志物,对易位到基因组中的标志物序列进行检测即可。作为检测标志物序列的方法,可举出大规模并行测序等下一代测序、PCR等。

[0132] 本发明中使用的载体或载体组可以进一步具有固有的分子条形码(唯一分子标识符,Unique Molecular Identifier;UMI)序列。例如,各个载体分子无论所编码的候补分子、调控序列或报告基因系统如何,均通过固有的UMI来识别。通过进行考虑到UMI的分析,可以消除在分析过程中产生的序列的扩增(细胞分裂、文库创建时的PCR等)的影响,减少偏差。例如,可以仅分析被赋予了唯一的UMI的报告基因系统的表达量、或者基于UMI的检测数量来校正与该UMI相关的报告基因的表达量。

[0133] 本发明提供的载体的种类只要可以在导入该载体的宿主细胞内表达所携带的上述候补分子、报告基因系统即可,没有特别限定。作为该载体的例子,可举出质粒载体、线性DNA序列、转座子载体、病毒载体等。该病毒载体可以为DNA载体,也可以为RNA载体,作为例子,可举出慢病毒载体、腺相关病毒载体、杆状病毒载体、MMLV逆转录病毒载体、MSCV(鼠干细胞病毒,Murine stem cell virus)逆转录病毒载体等。该载体可以存在于所导入的细胞内的细胞核外,也可以整合到基因组中而存在。可以按照常规方法将上述多核苷酸整合到载体中。

[0134] 本发明的功能性分子的探索方法中,将上述载体或载体组导入到细胞中,接着,对该细胞中的该载体或载体组所包含的上述报告基因系统的表达进行测定。

[0135] 导入上述载体的细胞(宿主细胞)可以为哺乳动物细胞,也可以为非哺乳动物细胞。优选该细胞为人、小鼠、大鼠等哺乳动物的细胞。另外优选该细胞为微生物细胞。优选该细胞为培养细胞。作为该细胞的例子,可举出采取自人或非人动物(例如小鼠、大鼠等)的细胞、采取自患者或实验动物的样本来源的细胞、原代培养细胞、类器官、细胞球状体、永生化培养细胞、样本来源的组织切片中的细胞、大肠杆菌等微生物细胞等。也可以使用实验动物个体的体内的细胞。

[0136] 载体向宿主细胞中的导入可以按照例如磷酸钙法、电穿孔法、脂质体法、基因枪法、PEG法等常规方法来实施。载体向细胞内导入用的试剂和仪器有市售。

[0137] 对上述报告基因系统的表达进行测定的方法可以根据该报告基因系统的种类而适当地选择。优选为可以对由该报告基因系统产生的标志物进行定量的方法。一个例子中,在该标志物为蛋白质的情况下,可举出公知的蛋白质定量法、例如显色法、吸光度测定、色谱法、免疫分析法等、以及对编码蛋白质的mRNA进行定量的方法、例如定量RT-PCR、测序等。另一例子中,该标志物为发光或荧光蛋白的情况下,作为其测定方法,可举出发光或荧光成像、发光或荧光光度测定等发光或荧光测定法。

[0138] 另一例子中,从上述报告基因系统中产生的标志物为由人工核酸酶或限制酶切割的DNA序列、由人工核酸酶或限制酶的切割所诱导的DNA序列的突变(切割时在切割位点产生的碱基的置换、缺失)、或利用碱基编辑器、先导编辑器等使识别序列突变的蛋白质而突

变的DNA序列的情况下,作为对上述报告基因系统的表达进行测定的方法,可举出定量PCR、大规模并行测序等下一代测序等可以检测该被切割的序列、或该突变的DNA序列的存在量的方法。例如,在上述载体上配置有人工核酸酶的识别序列或限制酶识别序列的情况下,通过利用PCR对包含切割位点的载体序列的一部分进行扩增和定量,或者以包含切割位点的载体序列的一部分为参照序列对序列读段进行定量,可以对报告基因系统的表达量进行定量。

[0139] 另一例子中,上述报告基因系统为上述的CRE—LoxP系统的情况下,作为上述测定方法,可举出对从以相同方向配置的LoxP序列之间切割出的标志物序列进行定量的方法、对该序列切割后残留的载体序列进行定量的方法、以及对以相反方向配置的LoxP序列之间的倒位的标志物序列进行定量的方法。该序列的定量可以使用定量PCR、大规模并行测序等下一代测序等。或者,如果如上所述将编码蛋白质的标志物序列以相反方向连接在以相反方向配置的LoxP序列之间,则可以利用上述的蛋白质定量法、发光或荧光测定法对由因CRE而倒位的标志物序列表达的标志物蛋白质进行定量。为其它CRE—LoxP样系统的情况也同样。

[0140] 另一例子中,上述报告基因系统为使用了上述的转座酶或整合酶的系统的情况下,作为上述测定方法,可举出对易位的标志物序列或缺失标志物序列的序列进行定量的方法。该序列的定量可以使用定量PCR、大规模并行测序等下一代测序等。

[0141] 上述报告基因系统的表达表明通过上述候补分子而产生了使上述载体或载体组中包含的上述调控序列激活的既定的细胞内的响应。通过对该报告基因系统的表达进行测定,可以确定该候补分子是否是引起该既定的细胞内的响应的物质(即,是否为探索的目标分子)。

[0142] 一个实施方式中,本发明的方法包括如下步骤:将至少1种载体或载体组导入到1个细胞中,并对来自该载体或载体组的报告基因系统的表达进行测定。

[0143] 本说明书中的“将至少1种载体导入到1个细胞中”是指最低向1个细胞中导入至少1种共同的载体,包含向单一细胞中导入1种或2种以上的载体的情况、向多个细胞中导入至少1种或2种以上的共同的载体的情况、以及向多个细胞的各个导入不同种类的载体的情况。任一情况下向每1个细胞中导入的载体的总数可以为1个,也可以为2个以上。该导入的载体各自包含:编码上述候补分子的表达系统的多核苷酸、编码上述调控序列的多核苷酸、以及编码上述报告基因系统的多核苷酸。

[0144] 另外,本说明书中的“向1个细胞中导入至少1种载体组”是指最低向1个细胞中导入至少1种共同的载体组,包含向单一细胞中导入1种或2种以上的载体组的情况、向多个细胞中导入至少1种或2种以上共同的载体组的情况、以及向多个细胞的各自导入不同种类的载体组的情况。任一情况下向每1个细胞中导入的载体组的总数可以为1个,也可以为2个以上。该导入的载体组包含:编码上述候补分子的表达系统的多核苷酸、编码上述调控序列的多核苷酸、以及编码上述报告基因系统的多核苷酸,这些多核苷酸包含于构成该载体组的2个以上的载体中的任一者中。

[0145] 优选在本发明的方法中对由2种以上的载体或2种以上的载体组分别包含的候补分子所引起的细胞内的响应进行分析。例如,在本发明的方法中,在1个细胞群中添加2种以上的载体或2种以上的载体组而将各载体或载体组导入至该细胞群中的任一细胞后,对来

自各载体或载体组的报告基因系统的表达进行测定。

[0146] 以下,对使用2种以上的载体(各自包含编码上述候补分子的表达系统的多核苷酸、编码上述调控序列的多核苷酸、以及编码上述报告基因系统的多核苷酸)时的本发明的方法的构成和步骤进行说明。使用2种以上的载体组(各个组包含编码上述候补分子的表达系统的多核苷酸、编码上述调控序列的多核苷酸、以及编码上述报告基因系统的多核苷酸)来代替该2种以上的载体时也可以以同样的构成和步骤来实施本发明的方法,这对本领域技术人员而言是清楚的。

[0147] 一个实施方式中,上述2种以上的载体各自包含编码分别由不同的细胞内的响应而激活的彼此不同的调控序列的多核苷酸、以及编码可操作地与该调控序列连接的彼此不同的报告基因系统的多核苷酸,另一方面,各载体中编码的候补分子可以相同,也可以彼此不同,优选在各载体中彼此不同。优选该报告基因系统为发光或荧光蛋白报告基因系统。

[0148] 另一实施方式中,上述报告基因系统为人工核酸酶与其识别序列的组合、或者限制酶和对应的限制酶识别序列的组合。

[0149] 一个例子中,上述2种以上的载体包含:编码彼此相同或不同的候补分子的多核苷酸、编码彼此由不同的细胞内的响应而激活的彼此不同的调控序列的多核苷酸、以及编码可操作地与该调控序列连接的彼此不同的报告基因系统的多核苷酸。另一个例子中,上述2种以上的载体编码共同的候补分子,另一方面,编码彼此不同的调控序列(彼此由不同的细胞内的响应而激活)和报告基因系统的序列。另一个例子中,上述2种以上的载体分别编码不同的候补分子,另一方面,编码共同的调控序列和报告基因系统的序列。

[0150] 本实施方式中,随着细胞内的响应而生成由酶切割的载体。接着,对由于各候补分子的作用而被切割的载体片段进行定量。利用PCR进行定量时,使用分别结合于载体的推测切割位点(酶识别序列)的上游和下游的2种反向引物。正向引物可以使用与编码各候补分子的序列结合的引物。或者,正向引物也可以是结合于与编码各个候补分子的序列对应的条形码序列(例如,与shRNA编码序列邻接的对于每个shRNA各不相同的序列)的引物。由结合于该推测切割位点上游的反向引物而产生的扩增产物的量除以由结合于下游的反向引物而产生的扩增产物的量所得的值与没有发生切割的对照模版中的值相比为高值时,认为是由于候补分子的存在而使载体中产生了切割。检测到载体的切割则表明各候补分子引起了细胞内的响应。

[0151] 另一实施方式中,上述载体的报告基因系统是碱基编辑器或先导编辑器与其识别序列的组合。

[0152] 一个例子中,上述2种以上的载体包含:编码彼此相同或不同的候补分子的多核苷酸、编码彼此由不同的细胞内的响应而激活的彼此不同的调控序列的多核苷酸、以及编码可操作地与该调控序列连接的彼此不同的报告基因系统的多核苷酸。另一个例子中,上述2种以上的载体编码共同的候补分子,另一方面,编码彼此不同的调控序列(分别由不同的细胞内的响应而激活)和报告基因系统的序列。另一个例子中,上述2种以上的载体分别编码不同的候补分子,另一方面,编码共同的调控序列和报告基因系统的序列。

[0153] 本实施方式中,随着细胞内的响应而向载体上的序列中导入突变。接着,对由于各候补分子的作用而发生突变的载体片段进行定量。利用PCR进行定量时,与上述同样地利用PCR分别对包含各候补分子序列或与其有关联的标记序列和突变的序列的载体片段进行扩

增和定量。在PCR中,可以使用结合于与上述同样的候补分子序列或与该候补分子序列有关联的标记序列上的正向引物、以及结合于该突变的序列的反向引物。扩增产物的检测表明各候补分子引起了细胞内的响应。以测序进行定量时,作为参照序列,使用包含各候补分子序列或者上述标记序列和突变的序列的载体片段的序列。该突变的序列例如可以为上述具有改变靶核苷酸序列的作用的蛋白质所靶向的序列和其周边序列。将细胞内的各种载体集中进行测序,对与各参照序列对应的读段进行检测或定量。存在与参照序列对应的读段时,例如,在匹配的读段中,上述所靶向的序列或其周边序列区域出现序列变化的情况下,表明各候补分子引起了细胞内的响应。

[0154] 另一实施方式中,上述报告基因系统包含CRE—LoxP样系统。

[0155] 一个例子中,上述2种以上的载体包含:编码彼此相同或不同的候补分子的多核苷酸、编码彼此由不同细胞内的响应而激活的彼此不同的调控序列的多核苷酸、以及编码可操作地与该调控序列连接的彼此不同的报告基因系统的多核苷酸。另一个例子中,上述2种以上的载体编码共同的候补分子,另一方面,编码彼此不同的调控序列(彼此由不同的细胞内的响应而激活)和报告基因系统的序列。该CRE—LoxP样系统选自CRE—LoxP(和/或LoxP变体)系统、Vika—Vox系统、Dre—Rox系统和Flp—FRT系统,可使用对于每个载体为不同种类的CRE—LoxP样系统。

[0156] 另一个例子中,上述2种以上的载体编码彼此不同的候补分子,另一方面,编码共同的调控序列和报告基因系统(CRE—LoxP样系统)的序列。

[0157] 该报告基因系统的2个LoxP样序列(包含标志物序列的2个LoxP或其变体的序列、2个Vox序列、2个Rox序列、或2个FRT序列)可以配置在编码候补分子的载体上,此时不同种类的LoxP样部分可以配置在相同的载体上。例如,这2个LoxP样序列相互反向配置,其间配置有标志物序列。标志物序列的种类针对各载体可以相同,也可以不同。

[0158] 本实施方式中,随着细胞内的响应而LoxP间的标志物序列发生倒位。接着,对包含编码各候补分子的序列和倒位的标志物序列的载体片段进行定量。利用PCR进行定量时,与上述同样将包含各候补分子序列或与其有关联的标记序列和该倒位的标志物序列的载体片段利用PCR分别进行扩增和定量。PCR中,可以使用与上述同样的结合于各候补分子序列或与其有关联的标记序列上的正向引物、以及结合于该倒位的标志物序列的反向引物。扩增产物的存在表明各候补分子引起了细胞内的响应。另一个例子中,以各候补分子序列或与其有关联的标记序列和进行退火的正向引物和以增加包含标志物序列的序列的方式设计的反向引物使用序列,对由细胞得到的载体序列进行扩增。扩增产物中的标志物序列的倒位量可以通过进行定量PCR来定量。在测序中进行定量时,作为参照序列,使用包含编码候补分子的序列和标志物序列的载体片段的序列。此时,准备涵盖了各候补分子序列或上述标记序列与倒位标志物序列的组合的一系列的参照序列。对细胞内的各种载体集中进行测序,对与各参照序列对应的读段进行检测或定量。与各参照序列对应的读段的存在、例如标志物序列自原始序列发生了变化的与各参照序列对应的读段的存在表明各候补分子引起了细胞内的响应。

[0159] 另一实施方式中,上述报告基因系统包含piggyBac转座子系统或phiC31/attP系统系统。

[0160] 一个例子中,上述2种以上的载体包含:编码彼此相同或不同的候补分子的多核苷

酸、编码彼此由不同细胞内的响应而激活的彼此不同的调控序列的多核苷酸、以及编码可操作地与该调控序列连接的彼此不同的报告基因系统的多核苷酸。另一个例子中,上述2种以上的载体编码共同的候补分子,另一方面,编码彼此不同的调控序列(彼此由不同的细胞内的响应而激活)和报告基因系统的序列。因该报告基因系统而被易位的标志物序列的种类在各载体中不同。

[0161] 另一个例子中,上述2种以上的载体编码彼此不同的候补分子,另一方面,编码共同的调控序列和报告基因系统的序列。

[0162] 本实施方式中,随着细胞内的响应而使既定的标志物序列易位。接着,对编码各候补分子的序列和易位的标志物序列或该标志物序列易位后残留的载体序列进行定量。优选预先构建序列以使得通过标志物序列的易位而使包含候补分子或与其有关联的标记序列的载体的序列发生变化,接下来,对编码候补分子的序列和该序列中的因标志物序列的易位而产生的序列变化的有无进行定量。

[0163] 定量在例如利用PCR进行定量的情况下,对各候补分子序列和标志物序列进行扩增。扩增产物的存在表明各候补分子引起了细胞内的响应。在测序中进行定量时,作为参照序列,将包含各候补分子序列、标志物序列的序列作为参照序列,对与各参照序列对应的读段进行检测或定量。与各参照序列对应的读段的存在、例如标志物序列自原始序列发生了变化的与各参照序列对应的读段的存在表明各候补分子引起了细胞内的响应。

[0164] 上述的一系列实施方式中,各载体编码共同的候补分子的情况下,可以基于表达的报告基因系统的种类来推测该候补分子的功能。即,可以推测与表达的报告基因系统相关的细胞内的响应(使可操作地连接于该报告基因系统的调控序列激活的细胞内的响应)是由该候补分子引起的。另一方面,在各载体编码不同候补分子、另外编码共同的调控序列的情况下,可以筛选引起特定的细胞内的响应的功能性分子。

[0165] 载体分子包含固有的分子条形码(UMI)的情况下,可以个别地识别各载体分子。UMI的序列可以通过测序进行调查。通过将UMI与候补分子、报告基因相关联,能够减少伴随着分析过程中产生的序列扩增的报告基因的表达量的偏差。

[0166] 根据以上的本发明的方法,可以根据报告基因的表达而容易地检测由候补分子引起的细胞内的响应,另外,可以容易地将候补分子与由它引起的细胞内的响应关联起来。因此,根据本发明的方法,RNA、肽等功能性分子的功能分析的效率提高。此外,在本发明的方法中,可以在1个细胞群中导入2种以上的载体或载体组,并对各个载体或载体组的行为进行分析。该情况下,可以以1个细胞群(例如1个孔中)一次性地调查2个以上的候补分子,或者一次性地从不同方面对1个候补分子的功能进行调查,因此能够提高功能性分子的综合分析的通量。

[0167] 本发明的引起细胞内的响应的功能性分子的探索方法可应用于新药、诊断等。例如,将本方法应用于新药筛选的情况下,将本发明的载体或载体组所编码的候补分子作为药剂的候补分子,并且以编码通过由目标药剂效果引起的或者带来该目标药剂效果的细胞内的响应而激活的调控序列和可操作地连接于该调控序列的报告基因系统的方式构建载体或载体组。另外,例如,将本方法应用于药剂的靶基因的探索时,将本发明的载体或载体组所编码的候补分子作为特异性抑制新药靶候补基因的功能性分子,并且将该载体或载体组所编码的由细胞内的响应而激活的调控序列以及可操作地连接于该调控序列的报告基因系统

作为检测与药效相关的信号通路的系统。作为1个例子,在探索癌症的治疗靶点的情况下,将抑制各种基因的多种的shRNA作为候补分子,作为报告基因系统使用与认为与癌症的药效相关的在癌细胞的应激响应时激活的调控序列(p53结合序列(p53RE)、抗氧化剂响应序列(ARE)等)结合的报告基因。

[0168] (诊断方法、试剂盒)

[0169] 1个实施方式中,本发明提供疾病的诊断方法,该方法包括如下步骤:对患者或者采取自患者的样本来源的细胞给予本发明的载体或载体组,以及测定来自该给予的载体或载体组的上述报告基因系统的表达。优选该载体或载体组所编码的候补分子为特异性抑制诊断对象的致病基因(例如突变基因、遗传多态等)的shRNA。该载体或载体组所编码的由细胞内的响应而激活的调控序列使用响应因致病基因的抑制而产生的变化的序列。作为具体的调控序列的例子,可举出p53结合序列(p53RE)、抗氧化剂响应序列(ARE)、以及核因子 $\kappa$ B(NF- $\kappa$ B)结合序列、转录激活因子6响应序列(ATF6 ERSE)、金属调节元件(MRE)、热休克元件(HSE)、缺氧反应元件(HRE)、外源物质响应元件(XRE)、活化T细胞核因子(NFAT)结合序列、TCF/LEF型DNA结合蛋白结合序列、sesRNA、或toehold RNA等。来自该载体或载体组的报告基因系统的表达表明该患者的对象致病基因的激活或者异常表达。

[0170] 一个实施方式中,本发明提供一种伴随诊断方法,该方法包括如下步骤:对患者或采取自患者的样本来源的细胞给予本发明的载体或载体组,以及测定来自该给予的载体或载体组的上述报告基因系统的表达。一个实施方式中,该载体或载体组所编码的候补分子是与要调查在该患者中的效果的药物的靶基因互补性结合而抑制其表达的shRNA。另一实施方式中,该载体或载体组所编码的候补分子是与该靶基因的突变体(遗传多态等)互补性结合而抑制其表达的shRNA。另一实施方式中,该载体或载体组所编码的候补分子是跟与要调查在该患者中的效果的药物的代谢相关的基因互补性结合而抑制其表达的shRNA。该载体或载体组所编码的由细胞内的响应而激活的调控序列选择响应如下生物信号的序列,所述生物信号的激活可以期待反映治疗效果。作为具体的调控序列的例子,可举出p53结合序列(p53RE)、抗氧化剂响应序列(ARE)、以及核因子 $\kappa$ B(NF- $\kappa$ B)结合序列、转录激活因子6响应序列(ATF6 ERSE)、金属调节元件(MRE)、热休克元件(HSE)、缺氧反应元件(HRE)、外源物质响应元件(XRE)、活化T细胞核因子(NFAT)结合序列、TCF/LEF型DNA结合蛋白结合序列、sesRNA、和toehold RNA等。实际的诊断中,例如用于癌症治疗的伴随诊断中,为了弄清楚对作为目标的癌细胞带来抑制效果的治疗目标,考虑使用p53响应序列、ATF6 ERSE等响应细胞内应激的序列。来自该载体或载体组的报告基因系统的表达是预测在该患者的组织中功能性分子的抑制或激活作为靶点的基因时的治疗效果的指标。例如,在用于上述癌症治疗的伴随诊断的例子中,可以在该患者组织中确定用于对癌细胞给予应激的靶基因。该测定结果可以作为对癌症患者开具靶基因的分子靶向药时的层别化标志。

[0171] 本发明另外提供包含上述本发明的载体或载体组的试剂盒,即,包含编码候补分子的表达系统的多核苷酸、编码由既定的细胞内的响应而激活的调控序列的多核苷酸、以及编码可操作地与该调控序列连接的报告基因系统的多核苷酸的载体或载体组。该试剂盒只要包含本发明的载体即可,可以根据需要进一步包含用于检测来自该载体或载体组的报告基因系统的表达的试剂、用于将该载体或载体组导入至细胞内的试剂、用于将该载体或载体组向患者给予的药学上可接受的担载体等。

[0172] 一个实施方式中,上述试剂盒为疾病的诊断药。一个实施方式中,该诊断药中使用的载体或载体组所编码的候补分子是与诊断对象的致病基因(例如突变基因、遗传多态等)互补性结合而抑制其表达的shRNA或siRNA。一个实施方式中,作为该载体或载体组所编码的调控序列,可举出p53结合序列(p53RE)、抗氧化剂响应序列(ARE)、以及核因子 $\kappa$ B(NF-kB)结合序列、转录激活因子6响应序列(ATF6 ERSE)、金属调节元件(MRE)、热休克元件(HSE)、缺氧反应元件(HRE)、外源物质响应元件(XRE)、活化T细胞核因子(NFAT)结合序列、TCF/LEF型DNA结合蛋白结合序列、sesRNA、和toehold RNA等。

[0173] 一个实施方式中,上述试剂盒为伴随诊断药。一个实施方式中,该伴随诊断药中使用的载体或载体组所编码的候补分子是与要调查患者体内的效果的药物的靶基因或该靶基因的突变体(遗传多态等)互补性结合而抑制其表达的shRNA或siRNA。另一实施方式中,该伴随诊断药中使用的载体或载体组所编码的候补分子是跟与要调查患者体内的效果的药物的代谢相关的基因互补性结合而抑制其表达的shRNA或siRNA。一个实施方式中,作为该载体或载体组所编码的调控序列,可举出p53结合序列(p53RE)、抗氧化剂响应序列(ARE)、以及核因子 $\kappa$ B(NF-kB)结合序列、转录激活因子6响应序列(ATF6 ERSE)、金属调节元件(MRE)、热休克元件(HSE)、缺氧反应元件(HRE)、外源物质响应元件(XRE)、活化T细胞核因子(NFAT)结合序列、TCF/LEF型DNA结合蛋白结合序列、sesRNA、和toehold RNA等。

[0174] (对调控序列的活性进行分析的方法)

[0175] 本发明另外提供一种对调控序列(转录调控序列或翻译调控序列)的活性进行分析的方法。本分析方法使用包含编码作为分析对象的调控序列的多核苷酸、编码可操作地与该调控序列连接的报告基因系统的多核苷酸、以及识别该调控序列的标记序列的载体。本分析方法中,在导入了该载体的细胞中,诱发既定的细胞内的响应,接着,对该载体所编码的报告基因系统的表达进行定量。基于该报告基因系统的表达和该标记序列评价针对该既定的细胞内的响应的该调控序列的活性。

[0176] 上述载体所编码的调控序列只要是想要分析针对既定的细胞内的响应的活性的转录调控序列或翻译调控序列即可,没有特别限定。该既定的细胞内的响应的种类没有特别限定,例如,可举出DNA损伤响应、氧化应激反应、特定的蛋白质或mRNA表达。优选该调控序列为转录调控序列。优选该转录调控序列为启动子或启动子变体。

[0177] 作为上述载体所包含的识别上述调控序列的标记序列,例如,可举出按照每一种来识别调控序列的条形码序列以及固有识别调控序列的分子条形码序列(Unique Molecular Identifier; UMI)。例如,对每个调控序列的种类、结构或序列赋予不同条形码序列。或者,也可以在调控序列的区域内存在该标记序列。例如,序列具有突变的调控序列变体自身可以成为标记序列。

[0178] 作为上述载体所编码的报告基因系统的例子,可举出发光或荧光蛋白报告基因系统、酶报告基因系统等。作为该酶报告基因系统所表达的酶的例子,可举出DNA核酸酶、DNA重组酶、水解酶(例如 $\beta$ -葡萄糖醛酸酶)、转座酶、整合酶等。该酶报告基因系统的具体例如上所述。

[0179] 一个优选的实施方式中,该报告基因系统包含具有改变靶核苷酸序列的作用的蛋白质及其目标序列,通过该报告基因系统的表达而使该目标序列发生变化。通过测量目标序列的变化量,可以对该报告基因系统的表达进行定量。作为这样的报告基因系统的优选

例,可举出上述的CRE—LoxP样系统、piggyBac转座子系统、phiC31/attP系统、碱基编辑器或先导编辑器和其识别序列等。CRE—LoxP样系统由CRE重组酶和2个LoxP样序列以及配置于它们之间的标志物序列构成。这2个LoxP样序列可以在相同方向串联配置,或者也可以在相反方向串联配置。CRE表达时,LoxP样序列配置于相同方向的情况下,其间的标志物序列被从载体上切出,另一方面,LoxP序列配置于相反方向的情况下,载体上的标志物序列倒位。piggyBac转座子系统由piggyBac转座酶、2个ITR和其间的标志物序列构成。该转座酶表达时,ITR间的标志物序列进行易位(从该ITR间消失)。因此,通过调查上述载体中的该目标序列(该标志物序列、酶识别序列)的变化、缺失,能够对该报告基因系统的表达进行定量。

[0180] 具体而言,本分析方法包括如下步骤:制备导入了上述载体的细胞;在该细胞中引发既定的细胞内的响应;引发该响应后,对该载体中的包含该目标序列和该标记序列的区域进行分析,将该目标序列的变化进行定量;

[0181] 基于该标记序列将该目标序列与该转录调控序列或翻译调控序列关联起来,

[0182] 基于该目标序列的变化量,对该目标序列产生了关联的转录调控序列或翻译调控序列的针对该既定的细胞内的响应的活性进行评价。

[0183] 作为本分析方法中可以使用的载体的种类、导入载体的细胞(宿主细胞)的种类、向细胞中导入载体的方法的例子,可举出与上述的引起细胞内的响应的功能性分子的探索方法中可以使用的例子相同的例子。

[0184] 作为本分析方法中的引发上述既定的细胞内的响应的方法,没有特别限定,例如,可举出应用诱导DNA损伤、氧化应激的药剂或刺激,应用Wnt通路等信号通路激活剂,抑制参与DNA修复、抗氧化的基因、蛋白质,激活特定的信号通路所涉及的基因、蛋白质,激活或抑制目标蛋白质、mRNA表达所涉及的基因等。这样的药剂、刺激、或者激活或抑制靶基因、蛋白质的方法为该领域所公知。

[0185] 作为对上述载体中的包含目标序列和标记序列的区域进行分析的方法,只要是可以对核苷酸序列的变化进行检测、定量的方法即可,例如,可举出定量RT—PCR、定量PCR、测序等。

[0186] 目标序列的变化量反映报告基因系统的表达量、乃至调控序列的活性。另一方面,与该目标序列存在于相同区域的标记序列由于确定带来该目标序列的变化的调控序列,因此可以基于该标记序列而将目标序列与调控序列产生关联。因此,可以基于该目标序列的变化量而对与该目标序列产生关联的调控序列的针对上述既定的细胞内的响应的活性进行评价。例如,既定的细胞内的响应后产生的目标序列的变化的量越大,可以评价为与该目标序列产生关联的该调控序列针对该既定的细胞内的响应所产生的活性更强。

[0187] 本分析方法包括如下步骤:向1个细胞中导入至少1种载体,并对来自该载体的标记或报告基因系统的表达进行测定。该“在1个细胞中导入至少1种载体”如上述所定义。优选在本分析方法中使用包含编码彼此不同的调控序列的多核苷酸的多核(2种以上)的载体的集合池,对各个调控序列的活性进行综合性分析。例如,在本发明的方法中,在1个细胞群中添加该载体的集合池而将各载体导入到该细胞群中的任一细胞中后,对各个载体中的目标序列的变化进行综合定量。例如,通过对来自不同载体的包含目标序列和标记序列的区域进行集中测序,可以对各个目标序列的变化进行定量,并且基于标记序列来确定带来各个目标序列的变化的调控序列。

[0188] (参与细胞内的响应的基因的探索方法)

[0189] 本发明另外提供一种参与细胞内的响应的基因的探索方法。本方法包括以下步骤:

[0190] 将多个载体或载体组导入于细胞群中,这里,该多个载体或载体组各自编码:干扰受试对象基因的RNA、由既定的细胞内的响应而激活的调控序列、以及可操作地与该调控序列连接的报告基因系统,其中,该多个载体或载体组各自所编码的该RNA相互不同;

[0191] 对该细胞群中的该多个载体或载体组各自所编码的报告基因系统的表达分别进行测定;以及

[0192] 基于该多个载体或载体组各自所包含的报告基因系统的表达来确定该受试对象基因是否为参与该既定的细胞内的响应的基因。

[0193] 因此,本方法中使用的载体或载体组编码:干扰上述受试对象基因的RNA、由要调查与上述受试对象基因的关系的既定的细胞内的响应而激活的调控序列、以及可操作地与该调控序列连接的报告基因系统。换言之,该载体或载体组包含:编码该RNA的表达系统的多核苷酸、编码该调控序列的多核苷酸、以及编码该报告基因系统的多核苷酸。这些多核苷酸可以配置于同一载体上,或者也可以配置于不同载体上。各个载体上的编码该RNA的表达系统的多核苷酸、编码该调控序列的多核苷酸、或者编码该报告基因系统的多核苷酸的配置只要该RNA表达,并且通过该调控序列的激活而促进该报告基因系统的表达,就没有特别限定。

[0194] 作为上述载体或载体组所编码的干扰上述受试对象基因的RNA,可举出shRNA、gRNA、miRNA、saRNA、lncRNA、eRNA等。其中,优选shRNA和gRNA。该RNA的表达系统可以包含编码该RNA的序列和可操作地与该序列连接的启动子等转录调控序列。该受试对象基因只要是要调查与上述既定的细胞内的响应的关系的基因即可,没有特别限定。该既定的细胞内的响应的种类没有限定,例如可举出DNA损伤响应、氧化应激反应、特定的蛋白质或mRNA表达。

[0195] 上述载体或载体组所编码的上述调控序列只要是由要调查与上述受试对象基因的关系的既定的细胞内的响应而激活的转录调控序列或翻译调控序列即可,没有特别限定。另外,该载体或载体组所编码的上述报告基因系统也只要可以在该调控序列的控制下表达即可,没有特别限定。作为该调控序列和报告基因系统的例子,可举出上述作为引起细胞内的响应的功能性分子的探索方法中使用的载体所可编码的调控序列和报告基因系统所例示的例子。该调控序列的活性和其所控制的报告基因系统的表达的水平反映由上述RNA的导入而引起的该既定的细胞内的响应的大小。

[0196] 本方法中使用的上述多个载体或载体组优选3种以上、更优选4种以上的载体或载体组。该多个载体或载体组各自编码的上述RNA对共同的受试对象基因进行干扰,但彼此不同。例如,上述RNA是以共同的受试对象基因为靶、但由彼此不同的序列构成的shRNA或gRNA。

[0197] 上述多个载体或载体组所编码的上述报告基因系统构建成可以识别是否是由于该多个载体或载体组所编码的上述RNA的任一者而表达的序列。优选该多个载体或载体组所编码的报告基因系统将其表达标记为与相同的载体或载体组所编码的RNA产生关联。

[0198] 本方法中,对上述多个载体或载体组各自编码的报告基因系统的表达分别进行测

定。接着,基于测定的各个报告基因系统的表达来确定上述受试对象基因是否为参与该既定的细胞内的响应的基因。

[0199] 一个实施方式中,本方法对由上述多个载体或载体组各自编码的上述RNA的导入所引起的既定的细胞内的响应的大小进行综合性分析。例如,本发明的方法中,在1个细胞群中添加该多个载体或载体组的集合池将各载体或载体组导入到该细胞群中的任一细胞中后,对来自各载体或载体组的报告基因的表达进行测定。基于上述报告基因系统与上述RNA的关联,可以对由上述RNA各自引起的既定的细胞内的响应的变化进行综合分析。

[0200] 无论被导入的RNA的种类如何都产生了上述既定的细胞内的响应的变化(例如响应的促进或抑制)时,可以确定该RNA所靶向的受试对象基因参与了该既定的细胞内的响应(促进或抑制)。另一方面,仅在特定种类的RNA中产生了该既定的细胞内的响应的变化时,RNA所靶向的受试对象基因没有参与该既定的细胞内的响应的可能性较高。

[0201] 一个实施方式中,基于上述多个载体或载体组各自所包含的报告基因系统的表达(换言之,由于与该报告基因系统产生了关联的RNA的导入而引起的上述既定的细胞内的响应的变化)的统计值,来确定上述受试对象基因是否为参与上述既定的细胞内的响应的基因。例如,可以算出所有报告基因系统的表达的平均值、中位数,并基于该值而对该受试对象基因进行排名。或者,由于过半数的报告基因系统的表达以相同趋势(例如促进或抑制)发生了变化,因此可以确定该受试对象基因参与了该既定的细胞内的响应。

[0202] 作为本方法中可以使用的载体的种类、导入载体的细胞(宿主细胞)的种类、将载体向细胞中导入的方法、测定报告基因系统的表达的方法的例子,可举出与上述的引起细胞内的响应的功能性分子的探索方法中可以使用的例子相同的例子。

[0203] 实施例

[0204] 以下,根据实施例对本发明进行更详细的说明。但是,本发明不受这些实施例等的任何限定。

[0205] 将以下的实施例中使用的PCR用的引物和模版示于表1~7和表A1~J。

[0206] [表1]

[0207] 表1质粒1制作用引物

序列号	引物	序列 (5'-3')	模版
1	正向	GTTAAGCCAGCCCGACACAATAAAATATCTTTAITTTCATTACATCTGTGT GITGG	p53RE-CRE
2	反向	ctttgacaagaagctgggtCTAATCGCCATCTTCCAGC	
3	正向	GCTGGAAGATGGCGATTAGaccagctttctgtacaag	SV40 late pA
4	反向	CACTGACGGGCACCGGAGCCgcgtaattaaTACCACATTTG	
[0208]	5 正向	GTAttaattaacgcGGCTCCGGTGCCCGTCAGTG	EF1A 启动子-Lox2272-LoxP-mKate2 (反向互
	6 反向	gaatcatgggaatagccctcTACCACATTTGTAGAGGTTTTAC	补)-Lox2272-LoxP-SV40 late pA
	7 正向	GTAAAACCTCTACAAATGTGGTAagaggcctatttccatgattc	U6-shScramble-终止子-PGK 启动子-EGFP-rGB
	8 反向	CTAIGCGGCATCAGAGCAGgattcctacataagagaagagg	pA
	9 正向	cctctcttattatggagatcCTGCTCTGATGCCGCATAG	OriC
	10 反向	GAAAATAAAGATAITTTATTGTGTCGGGGCTGGCTTAAC	

[0209] [表2]

[0210] 表2质粒2制作用引物

序列号	引物	序列 (5'-3')	模版
1	正向	GTTAAGCCAGCCCCGACACAATAAAATATCTTTATTTTCATTACATCTGTGT GTTGG	p53RE-CRE
2	反向	ctttgtacaagaagctgggtCTAATCGCCATCTCCAGC	
3	正向	GCTGGAAGATGGCGATTAGaccagctttctgtacaag	SV40 late pA
4	反向	CACTGACGGGCACCGGAGCCcggttaattaaTACCACATTG	
5	正向	GTAttaattaacgCGCTCCGGTGCCCCGTCAGTG	EF1A 启动子-Lox2272-LoxP-mKate2 (反向互补)
6	反向	gaatcatgggaaatagccctcTACCACATTGTAGAGGTTTAC	-Lox2272-LoxP-SV40 late pA
7	正向	GTA AACCTCTACAAATGTGGTAgaggcctatttcccatgattc	
11	反向	aagttgatgatcgaaattcaaaaaGAAGCTATGTTCCGCAATTACTCGAGTAATGGCGAAC ATAGCTTCcgggtgttctcttcc	U6-shRAD51_1-终止子
7	正向	GTA AACCTCTACAAATGTGGTAgaggcctatttcccatgattc	
12	反向	aagttgatgatcgaaattcaaaaaGCCAACGATGTGAAGAAATTTCTCGAGAATTTCTTCA CATCGTTGGCccgggtgttctcttcc	U6-shRAD51_2-终止子
7	正向	GTA AACCTCTACAAATGTGGTAgaggcctatttcccatgattc	
13	反向	aagttgatgatcgaaattcaaaaaAAGCTATGTTCCGCAATTACTCGAGTTAATGGCGAAC CATAGCTTCcgggtgttctcttcc	U6-shRAD51_3-终止子
7	正向	GTA AACCTCTACAAATGTGGTAgaggcctatttcccatgattc	
14	反向	aagttgatgatcgaaattcaaaaaGCAGTGATGTTCTGGATAACTCGAGTTATCCAGGAC ATCACTGCcgggtgttctcttcc	U6-shRAD51_4-终止子
15	正向	ttttgaattcagatcatcaactttagatag	PGK 启动子-EGFP-rGB pA
8	反向	CTATGCGGCATCAGAGCAGgatctccataagagaaggg	
9	正向	ecttctcttatggagatcCTGCTCTGATGCCGCATAG	OriC
10	反向	GAAAATAAAGATATTTTATTGTGTCGGGGCTGGCTTAAC	

[0211] [表3]

[0213] 表3质粒3制作用引物

序列号	引物	序列 (5'-3')	模版
1	正向	GTTAAGCCAGCCCCGACACAATAAAATATCTTTATTTTCATTACATCTGTGT GTTGG	载体 1 的一部分
16	反向	CAGTTAGGCCAGAGAAATGTTCTGGC	
17	正向	GGCTGTCCCAGTGCAAGTG	ARE
18	反向	GGTGGCTTACCAACAGTACCG	
19	正向	GCTCGACTTCCAGCTTGGC	载体 1 的一部分
4	反向	CACTGACGGGCACCGGAGCCcggttaattaaTACCACATTG	
5	正向	GTAttaattaacgeGGCTCCGGTGCCCCGTCAGTG	EF1A 启动子-Lox2272-LoxP-mKate2 (反向互补)
6	反向	gaatcatgggaaatagccctcTACCACATTGTAGAGGTTTAC	-Lox2272-LoxP-SV40 late pA
7	正向	GTA AACCTCTACAAATGTGGTAgaggcctatttcccatgattc	U6-shScramble-终止子-PGK 启动子-EGFP-rGB pA
8	反向	CTATGCGGCATCAGAGCAGgatctccataagagaaggg	
9	正向	ecttctcttatggagatcCTGCTCTGATGCCGCATAG	OriC
10	反向	GAAAATAAAGATATTTTATTGTGTCGGGGCTGGCTTAAC	

[0215] [表4]

[0216] 表4mKate2序列的倒位率的计算用引物

目标	引物	序列 (5'-3')	序列号
原始序列	正向	ATTTTCAGGTGTCGTGACAAGTTTGTAC	20
	反向	TCCCTAGCAAACCTGGGGCAC	21
倒位序列	正向	ATTTTCAGGTGTCGTGACAAGTTTGTAC	20
	反向	GAAGTGGTGGTTGTTACCGG	22
主干序列 (OriC 序列特异性引物)	正向	TCAGGAAGCTTGGATCAACCGG	23
	反向	TGATCGGTGATCCTGGACCG	24

[0218] [表5]

[0219] 表5扩增子制作用引物(质粒1,2共同)

引物	序列 (5'-3')	序列号
正向	GGCCAGCTTGGCACTTGATG	25
反向	GGACGTGAAGAATGTGCGAG	26

[0221] [表6]

[0222] 表6每种shRNA的扩增子制作用引物

扩增子	引物	序列 (5'-3')	序列号
包含 hScramble 的片段	正向	GGCCAGCTTGGCACTTGATG	25
	反向	GAGGGCGACTTAAACCTTAGG	27
包含 shRAD51_2 的片段	正向	GGCCAGCTTGGCACTTGATG	25
	反向	AATTTCTTCACATCGTTGGC	28
包含 shRAD51_4 的片段	正向	GGCCAGCTTGGCACTTGATG	25
	反向	TTATCCAGGACATCACTGC	29

[0224] [表7]

[0225] 表7mKate2序列的倒位率的计算用引物

目标	引物	序列 (5'-3')	序列号
原始序列	正向	ATTCAGGTGTCGTGACAAGTTTGTAC	20
	反向	TCCCTAGCAAACCTGGGGCAC	21
倒位序列	正向	ATTCAGGTGTCGTGACAAGTTTGTAC	20
	反向	GAAGTGGTGGTTGTTCACGG	22

[0227] [表A1]

[0228] 表A1线性DNA1制作用引物

序列号	引物	序列 (5'-3')
30	正向	GTTAAGCCAGCCCCGACACTGTTACGTAGAGGAGGTTATCC
31	反向	CTATGCGGCATCAGAGCAGGTACCGTTATATGCCGGCGG

[0230] [表A2]

[0231] 表A2线性DNA2A制作用引物

序列号	引物	序列 (5'-3')
30	正向	GTTAAGCCAGCCCCGACACTGTTACGTAGAGGAGGTTATCC
32	反向	CGGTGTTTCGTCCTTCCACAAGATATATAAAGCC
33	正向	tagtacagtagtaatagtatatgtgatggaTC
31	反向	CTATGCGGCATCAGAGCAGGTACCGTTATATGCCGGCGG

[0233] [表B]

[0234] 表B线性DNA2A/线性DNA3B用shRNA (下划线表示目标序列的相同区域)

shRNA ID	靶基因	序列 (5'-3')	序列号
USP7_1	USP7	TCTTGGCTTTATATACTTTGTGGAAAGGACGAAACACCGGCCAGCT AAGTATCAAAGGAAACTCGAGTTTCCTTTGATACTTAGCTGGTTTT TTAGTACAGTAGTAATAGTATAATGTGATGGATCT	34
USP7_2	USP7	TCTTGGCTTTATATACTTTGTGGAAAGGACGAAACACCGGCCCTGGA TTTGTGGTTACGTTACTCGAGTAACGTAACCACAAATCCAGGTTTT TTAGTACAGTAGTAATAGTATAATGTGATGGATCT	35
USP7_3	USP7	TCTTGGCTTTATATACTTTGTGGAAAGGACGAAACACCGGGTGTCC TATATCCAGTGTAAACTCGAGTTTACACTGGATATAGGACACTTTTT TAGTACAGTAGTAATAGTATAATGTGATGGATCT	36
USP7_4	USP7	TCTTGGCTTTATATACTTTGTGGAAAGGACGAAACACCGGCGTGGT GTCAAGGTGTAATACTCGAGTTAGTACACCTTGACACCACGTTT TTAGTACAGTAGTAATAGTATAATGTGATGGATCT	37
DHX29_1	DHX29	TCTTGGCTTTATATACTTTGTGGAAAGGACGAAACACCGGCGTGTGA CCTTTGGAGGAATTACTCGAGTAATTCCTCCAAAGGTACACGTTTT TTAGTACAGTAGTAATAGTATAATGTGATGGATCT	38
DHX29_2	DHX29	TCTTGGCTTTATATACTTTGTGGAAAGGACGAAACACCGGCCTAAG TATCAGAAACTTCTACTCGAGTAGAAGTTTCTGATACTTAGGTTTT TTAGTACAGTAGTAATAGTATAATGTGATGGATCT	39
DHX29_3	DHX29	TCTTGGCTTTATATACTTTGTGGAAAGGACGAAACACCGGCGTGGT AGTATGCCCTACAAICTCGAGATTGTAGGGCATACTACCAGTTTT TTAGTACAGTAGTAATAGTATAATGTGATGGATCT	40
DHX29_4	DHX29	TCTTGGCTTTATATACTTTGTGGAAAGGACGAAACACCGGCCTCAA ATACAAGCCACTATCTCGAGAATAGTGGCTTGTATTTGAGGTTTT TTAGTACAGTAGTAATAGTATAATGTGATGGATCT	41
SETDB1_1	SETDB1	TCTTGGCTTTATATACTTTGTGGAAAGGACGAAACACCGGCGTGTGAC TTCATAGAGGAGTATCTCGAGATACTCCTCTATGAAGTCACGTTTT TTAGTACAGTAGTAATAGTATAATGTGATGGATCT	42
SETDB1_2	SETDB1	TCTTGGCTTTATATACTTTGTGGAAAGGACGAAACACCGGGCTCAG ATGATAACTTCTGTACTCGAGTACAGAAGTTATCATCTGAGCTTTT TTAGTACAGTAGTAATAGTATAATGTGATGGATCT	43
SETDB1_3	SETDB1	TCTTGGCTTTATATACTTTGTGGAAAGGACGAAACACCGGAGTTAG AGACATGGGTAATACCCTCGAGGTATTACCCATGTCTCTAACTTTTT TAGTACAGTAGTAATAGTATAATGTGATGGATCT	44

[0235]

[0236] [表C1]

[0237] 表C1线性DNA3A制作用引物

[0238]

序列号	引物	序列 (5'-3')
45	正向	TGTTACGTAGAGGAGTTATCC
46	反向	GTACCGTTATATGCCGGCGG

[0239] [表C2]

[0240] 表C2线性DNA3B制作用引物

[0241]

序列号	引物	序列 (5'-3')
45	正向	TGTTACGTAGAGGAGTTATCC
47	反向	CCGGTGTTCGTCCTTTCCACAAG
48	正向	TAGTACAGTAGTAATAGTATAATGTGATGGATC
46	反向	GTACCGTTATATGCCGGCGG

[0242] [表D]

[0243] 表D Wnt通路相关基因分析用引物

序列号	引物	序列 (5'-3')	靶基因	
[0244]	49	正向	GTCTCCTCTGACTTCAACAGCG	GAPDH
	50	反向	ACCACCCTGTTGCTGTAGCCAA	
	51	正向	CTGGCTTTGGTGAAGTGTG	AXIN2
	52	反向	AGTTGCTCACAGCCAAGACA	
	53	正向	GAAACAGTGCTCGGGTTTTC	SP5
	54	反向	TAGCTCTGCATGGAGCTGAA	

[0245] [表E]

[0246] 表E Lox间序列分析用引物

序列号	引物	序列 (5'-3')	靶基因	
[0247]	55	正向	TGATCAACCGTGGTTACAAGCAGG	Lox
	56	反向	CTGGACTGTACCTCGTTGTACGG	
	57	正向	GGCCTATTCCCATGATTCCTTC	主干序列
	58	反向	CCAAGAAATTACTTTCTACGTCAC G	

[0248] [表F]

[0249] 表F Lox间序列扩增用引物

序列号	引物	序列 (5'-3')	
[0250]	25	正向	GGCCAGCTTGGCACTTGATG
	59	反向	AAAGCCATACGGGAAGCAATAG

[0251] [表G]

[0252] 表G Lox间序列-shRNA编码序列扩增用引物

序列号	引物	序列 (5'-3')	
[0253]	60	正向	AATAGGGCCGGTGGAAAGGACGAAACACCG
	61	反向	TGACTCGTCGCCAGTTCCTACATTCGCTCTC
	62	正向	CAACGGCTACGTGGAAAGGACGAAACACCG
	63	反向	CTCCATGTTGCCAGTTCCTACATTCGCTCTC
	64	正向	CAGCGAATGCGTGGAAAGGACGAAACACCG
	65	反向	CTACAAGGATCCAGTTCCTACATTCGCTCTC

[0254] [表H1]

[0255] 表H1线性DNA4制作用引物

序列号	引物	序列 (5'-3')	
[0256]	45	正向	TGTTACGTTAGAGGAGGTTATCC
	66	反向	AGATCCATCACATATACTATTACTACTGTAC
	45	正向	TGTTACGTTAGAGGAGGTTATCC
	47	反向	CCGGTGTTCGTCCTTTCCACAAG

67	正向	TGCGACTGACGGTTGCAACG
46	反向	GTACCGTTATATGCCGGCGG

[0257] [表H2]

[0258] 表H2线性DNA4扩增用引物

序列号	引物	序列 (5'-3')
45	正向	TGTTACGTTAGAGGAGTTATCC
46	反向	GTACCGTTATATGCCGGCGG

[0260] [表I]

[0261] 表I质粒7制作用引物

序列号	引物	序列 (5'-3')	模版
1	正向	GTAAAGCCAGCCCCGACACAATAAAATATCTTTAATTTCAITACATCTGTGT GTTGG	p53RE-CRE
68	反向	AGTGICTAAGCTTGGCCGC	
69	正向	GGATAACAAGATCTGGCCTCG	minP 变体
70	反向	GGGCCATGGTGGCTTAC	
71	正向	CTTGCCAATCCGGTACTGTTG	CRE
2	反向	ctttgtacaagaagctgggtCTAATCGCCAICTTCCAGC	
3	正向	GCTGGAAGATGGCGATTAGaccagctttctgtacaag	SV40 late pA
4	反向	CACTGACGGGACCCGGAGCCgcgtaattaaTACCACAITTG	
5	正向	GTAAaattaacgcGGCTCCGGTGCCGTCAGTG	EF1A 启动子-Lox2272-LoxP-mKate2 (反向互 补)-Lox2272-LoxP-SV40 late pA
6	反向	gaatcatgggaataggccctTACCACAITTGTAGAGGTTTTAC	
7	正向	GTAAAACCTCTACAAATGTGGTAagggcctatttccatgattc	U6-shScramble-终止子-PGK 启动子-EGFP-rGB pA
8	反向	CTATGCGGCATCAGAGCAGgatccataagaagagg	
9	正向	ccttctcttatggagatcTGCTCTGATGCCGALAG	
10	反向	GAAAATAAAGATAITTTAATTGTGTCGGGGCTGGCTTAAAC	OriC

[0262]

[0263] [表J]

[0264] 表J minP变体序列-mKate2序列扩增用引物

序列号	引物	序列 (5'-3')
69	正向	GGATATCAAGATCTGGCCTCG
26	反向	GGACGTGAAGAATGTGCGAG

[0266] 实施例1载体的制作

[0267] 1) 质粒1的制作

[0268] 以表1所示的模版、引物的组合来实施PCR。PCR使用Platinum™ SuperFi II PCR Master Mix (ThermoFisher)。对PCR产物进行琼脂糖凝胶电泳。电泳中检测到单一条带时，向Wizard™ SV Gel and PCR Clean-Up System (Promega) 的柱子中直接添加PCR反应溶液，按照制造商推荐的规程对PCR片段进行纯化。电泳中检测到非特异性序列时，切出含有目标片段的凝胶而加入到Wizard™ SV Gel and PCR Clean-Up System (Promega) 中，对PCR片段进行纯化。纯化的片段使用无细胞克隆系统 (Cell-Free Cloning System, OriCiro) 按照制造商记载的规程进行组装，在体外进行扩增。将得到的反应溶液直接添加到Wizard™ SV Gel and PCR Clean-Up System (Promega) 的柱子中，按照制造商推荐的规程对质粒载体 (质粒1) 进行纯化。将得到的质粒利用XbaI (New England Biolabs) 进行切割，通过琼脂糖凝胶电泳来确认分子量。本质粒在功能性分子的区域具有编码不具备靶点的Scramble shRNA的序列 (sh Scramble)，具有p53RE作为转录调控序列，并且在该转录调控序列的下游具有作为报告基因系统的CRE-loxP样序列。该CRE-loxP样序列具有CRE基因序列和相互反向配置的2个Lox2722—LoxP序列以及配置于其间的红色荧光蛋白mKate2编码序列 (mKate2序列) (图1A)。该mKate2序列在原始状态下配置成与转录方向相反的方向，不表达。转录调控序列激

活时,可以通过CRE的作用使mKate2序列倒位并在上游的EF1启动子的作用下进行转录(图1C)。mKate2表达了的细胞用红色荧光标记,这表明由功能性分子产生了使p53RE激活的细胞内的响应。该质粒具有与恒定表达的启动子连接的GFP,具有该质粒的转化体由绿色荧光标记。

#### [0269] 2) 质粒2的制作

[0270] 以表2所示的模版、引物的组合来实施PCR,除此以外,按照与质粒1同样的步骤来制作质粒载体(质粒2),并确认分子量。本质粒在功能性分子的区域具有编码以RAD51为靶的shRNA的序列(shRAD51),并且如上所述,具有包含p53RE和mKate2序列作为转录调控序列和报告基因系统的CRE—loxP样序列(图1A)。该质粒中的shRAD51为表2所示的4种(shRAD51\_\_1~shRAD51\_\_4)中的任一种。将具有各个shRAD51序列的4种质粒进行集合用作质粒2。质粒2中的mKate2序列与质粒1同样通过转录调控序列的激活而进行倒位并表达(图1C),由此细胞用红色荧光标记。该质粒具有与恒定表达的启动子连接的GFP,具有该质粒的转化体用绿色荧光标记。

#### [0271] 3) 质粒3的制作

[0272] 以表3所示的模版、引物的组合来实施PCR,除此以外,按照与质粒1同样的步骤来制作质粒载体(质粒3),并确认分子量。本质粒在功能性分子的区域具有shScramble,具有抗氧化剂响应序列(ARE)作为转录调控序列,并且在该转录调控序列的下游具有作为报告基因系统的CRE—loxP样序列。该CRE—loxP样序列具有CRE基因序列和相互反向配置的2个Lox2722—LoxP序列以及配置于其间的mKate2编码序列(图1B)。质粒3中的mKate2序列与质粒1同样地通过转录调控序列的激活而倒位并表达(图1C),由此细胞由红色荧光标记。该质粒具有与恒定表达的启动子连接的GFP,具有该质粒的转化体由绿色荧光标记。

#### [0273] 4) 质粒4的制作

[0274] 由GenScript Japan株式会社制作具有图1D的结构 of 的质粒载体。本质粒在功能性分子的区域具有shScramble,具有TCF/LEF DNA结合蛋白结合序列作为转录调控序列,并且在该转录调控序列的下游具有作为报告基因系统的CRE—loxP样序列。该CRE—loxP样序列具有CRE基因序列以及相互反向配置的Lox71序列和Lox66序列(Lee等, Gene, 1998, 216(1): 55—65)。该Lox序列间的间隔序列使用m2变体(Parrish等, BioMed Research International, 2011, doi.org/10.1155/2011/924068)。另外,本质粒具有在相同方向配置的Rox序列作为CRE—loxP样序列的一部分。Rox序列一般已知为Dre重组酶的目标序列,但报道了也与Cre重组酶具有交叉性(Nat. Methods, 11(2014), pp. 763—772)。质粒4中的由Lox序列夹住的标志物序列在由转录调控序列的激活而表达诱导的CRE的活性下发生倒位。转录调控序列为TCF/LEF DNA结合蛋白结合序列,通过与TCF/LEF结合而被激活,TCF/LEF的活性通过Wnt/ $\beta$ -Catenin信号通路的激活而增加。该标志物序列的倒位用来反映导入了该质粒的细胞中的Wnt/ $\beta$ -Catenin信号通路的激活。

#### [0275] 5) 病毒载体的制作

[0276] 制作慢病毒载体1A、1B、1C。慢病毒载体1A具有与质粒1相同的shRNA(shScramble)和作为报告基因系统的loxP样序列。慢病毒载体1B具有与质粒2相同的shRNA(shRAD51)和作为报告基因系统的loxP样序列。这些载体的loxP样序列中包含相互反向配置的2个Lox2722—LoxP序列和在其间倒位配置红色荧光蛋白mKate2编码序列。慢病毒载体1C具

有p53RE转录调控序列和CRE基因序列。构建将慢病毒载体1A或1B与慢病毒载体1C组合而成的载体组。该载体组以如下方式发挥作用：利用载体1A或1B的shRNA而诱导细胞内的DNA损伤响应时，感应到该DNA损伤响应而激活载体1C的p53RE来表达CRE，由此产生载体1A或1B的loxP样序列的变化（mKate2序列的倒位）。病毒载体1A、1B和1C由Vector Builder Japan株式会社进行质粒合成和慢病毒包装来制作。

[0277] 6) 线性DNA1的制作

[0278] 以上述质粒4为模版，利用表A1所示的引物来实施PCR，将得到的片段使用Wizard™SV Gel and PCR Clean—Up System (Promega) 进行纯化。将得到的DNA片段作为线性DNA1。线性DNA1具有包含与质粒4同样的shScramble的功能性分子的区域、转录调控序列、报告基因系统和用于使这些序列在细胞内发挥功能的序列（图1E）。

[0279] 7) 线性DNA2的制作

[0280] 以上述线性DNA1为基础，制作包含编码各种shRNA的序列的线性DNA，将它们集合作为线性DNA2。作为shRNA，选择以涉及Wnt通路的基因为靶的shRNA。具体的序列制作按照以下顺序进行。以上述质粒4为模版并利用表A2所示的引物来实施PCR，将得到的PCR产物利用Wizard™SV Gel and PCR Clean—Up System (Promega) 进行纯化，制作从线性DNA1中删除了包含shRNA的区域的DNA片段。将包含编码以涉及Wnt通路的基因（USP7、DHX29或SETDB1）为靶的shRNA的序列（表B）的单链寡核苷酸DNA片段的集合池委托Integrated DNATechnologies株式会社来制作。上述涉及Wnt通路的基因是通过敲除而激活Wnt通路的基因（Evron等，Oncogenesis, 2021, 10(9):63）。这些DNA片段被设计为相互重叠的。将这些DNA片段混合作为模版，通过PCR使片段结合。将得到的PCR产物在1%琼脂糖凝胶中进行电泳后，切出该片段的条带利用Wizard™SV Gel and PCR Clean—Up System (Promega) 进行纯化，得到线性DNA2A。将纯化的线性DNA2A的溶液和上述线性DNA1的溶液混和，制成线性DNA2。线性DNA2具有包含编码以Wnt通路为靶的shRNA的序列或shScramble的功能性分子区域、与质粒4同样的转录调控序列和报告基因系统、以及用于使这些序列在细胞内发挥功能的序列，功能性分子区域为异源DNA片段集合。

[0281] 8) 线性DNA3的制作

[0282] 将编码sesRNA的DNA整合到线性DNA2的转录控制区域中，制作线性DNA3。sesRNA是通过任意的mRNA分子和细胞内的ADAR蛋白质的活性而使下游的报告基因表达的模块（Nature, 2022, 610:713—721, Nature Biotechnology, 2023, 41:698—707, Nature Biotechnology, 2023, 41:482—487, WO/2023/077135, WO/2023/077138, US20230123513）。设计了一种编码感应AXIN2基因的mRNA的sesRNA的DNA。AXIN2由于通过Wnt通路激活而表达增加，因此该sesRNA因细胞内的Wnt通路的激活而增加下游基因的表达。

[0283] 将质粒4的转录控制区域用编码该sesRNA的DNA置换来制作质粒5（图1F）。质粒5由GenScript Japan株式会社制作。将质粒5通过具有表C1的引物的PCR而线性化，得到线性DNA3A。接着，以质粒5为模版，利用表C2所示的引物来实施PCR，将得到的PCR产物利用Wizard™SV Gel and PCR Clean—Up System (Promega) 进行纯化，制作从线性DNA3A中删除包含shRNA的区域的DNA片段。将包含编码以涉及Wnt通路的基因为靶的shRNA（表B）的序列的单链DNA片段的集合池委托Integrated DNA Technologies株式会社来制作。这些DNA片段被设计为相互重叠的。将这些DNA片段由NEBuilder (New England Biolabs Japan株式

会社)来结合。以结合的片段为模版来进行PCR,将PCR产物利用Wizard™SV Gel and PCR Clean—Up System (Promega)进行纯化,得到线性DNA3B。将线性DNA3A的溶液与线性DNA3B的溶液混合得到线性DNA3。线性DNA3具有包含编码以Wnt通路为靶的shRNA的序列或shScramble的功能性分子区域、编码与质粒5同样的sesRNA翻译调控序列和报告基因系统的序列、以及用于使这些序列在细胞内发挥功能的序列,功能性分子区域为异源DNA片段集合。

[0284] 实施例2质粒针对DNA损伤响应的响应

[0285] 为了验证基于实施例1所制作的质粒所带来的DNA损伤响应的检测能力,调查质粒1对人为诱导的DNA损伤响应的反应。将HEK293细胞以15000cells/孔的密度接种于96孔板,用EMEM(添加NEAA)培养基培养1天。使用Lipofectamine 3000kit(ThermoFisher)将100ng/孔的质粒1导入至细胞中。导入所使用的P3000 reagent和Lipofectamine 3000reagent的量均为0.2 $\mu$ L/孔。转染起40.5小时后,将具有DNA损伤效果的阿霉素(Sigma—Aldrich)以终浓度10 $\mu$ M、5 $\mu$ M、1 $\mu$ M或0 $\mu$ M(DMSO)进行添加。在阿霉素添加前以及添加后8小时、24小时和33小时,使用荧光显微镜BZ—X710(KEYENCE)以4倍物镜(BZ—PF04P,KEYENCE)对细胞进行观察。对基于EGFP的绿色荧光以激发波长470 $\pm$ 20nm、荧光波长525 $\pm$ 25nm的滤波器组进行检测,将基于mKate2的红色荧光用激发波长545 $\pm$ 12.5nm、荧光波长605 $\pm$ 35nm的滤波器组进行检测。荧光的阳性或阴性的判断使用安装于KEYENCE制的BZ—x Analyzer软件的混合细胞计数功能来实施。分析的结果,红色荧光阳性细胞数/绿色荧光阳性细胞数的值依赖于阿霉素的处理时间和浓度地增加(图2),暗示了通过由阿霉素所致的DNA损伤响应的诱导而激活质粒中的p53RE和CRE而使mKate2序列进行了倒位。根据该结果,表明了实施例1的质粒可以检测细胞的DNA损伤响应。

[0286] 为了确认mKate2序列的倒位,在阿霉素处理33小时后从培养物中分离细胞,由该细胞利用苯酚·氯仿提取、乙醇沉淀而得到DNA溶液。使用表4所示的引物和TB Green™Premix Ex Taq™II(Takara),利用SYBR Green法来实施qPCR分析,对倒位的序列进行定量。其结果,通过阿霉素处理而使mKate2序列的倒位率(RQ)增大(图3),是与上述的荧光分析对应的结果。

[0287] 实施例3由质粒所编码的shRNA所致的DNA损伤响应的增强

[0288] 调查由实施例1中制作的质粒所编码的shRNA所致的DNA损伤响应的增强效果和该质粒的该效果的检测能力。作为质粒,使用具有Scramble shRNA的质粒1和具有shRAD51的质粒2。质粒1、2均搭载了p53RE,由于DNA损伤响应而使CRE重组酶表达。质粒2具有的shRNA所抑制的RAD51是与DNA修复有关的蛋白质,通过其表达抑制来增强DNA损伤响应(Journal of medicinal chemistry,2012,55(7):3011—3020;Oncology Reports,2019,42(6):2426—2434)。

[0289] 将HEK293细胞以25000cells/孔的密度接种于96孔板,用EMEM(添加NEAA)培养基培养1天。使用Lipofectamine 3000kit(ThermoFisher)将100ng/孔的质粒1和质粒2分别导入不同孔的细胞中。导入所使用的P3000 reagent和Lipofectamine 3000reagent的量均为0.2 $\mu$ L/孔。转染起36小时后,以终浓度5 $\mu$ M或0 $\mu$ M(DMSO)添加具有DNA损伤效果的阿霉素(Sigma—Aldrich)。在阿霉素的添加前以及添加后26小时和74小时,按照与实施例2相同的顺序对细胞进行观察,对荧光阳性或阴性的细胞进行计数。分析的结果,添加阿霉素74小时

后,含有导入了shRAD51(质粒2)的细胞的孔中的红色荧光阳性细胞的比例与含有导入了shScramble(质粒1)的细胞的孔相比增加了(图4)。根据该结果,表明了通过质粒所编码的shRAD51而增强了由阿霉素所致的DNA损伤响应,通过质粒而检测到该DNA损伤响应的增强。

[0290] 实施例4抗氧化应激响应的检测

[0291] 使用实施例1中制作的具有响应氧化应激的ARE作为由细胞内的响应而激活的转录调控序列的质粒3,调查质粒对人为诱导的细胞内氧化应激的响应。将HEK293细胞以25000cells/孔的密度接种于96孔板,在EMEM(添加NEAA)培养基中培养1天。使用Lipofectamine 3000kit(ThermoFisher),将100ng/孔的质粒3导入于细胞中。导入中使用的P3000 reagent和Lipofectamine 3000reagent的量均为0.2 $\mu$ L/孔。转染起66小时后,以终浓度100 $\mu$ M、50 $\mu$ M和0 $\mu$ M(DMSO)添加叔丁基对苯二酚(TBHQ,东京化成),诱导氧化应激。在TBHQ添加后26小时、75小时和99小时以与实施例2相同的顺序对细胞进行观察。荧光阳性或阴性细胞的判断使用CellProfiler(Broad Institute[URL:cellprofiler.org/interfaces])来实施。分析的结果,红色荧光阳性细胞的比率依赖于TBHQ的处理时间和浓度而增加(图5),暗示了通过基于TBHQ的氧化应激而质粒中的ARE激活而使mKate2序列进行了倒位。表明质粒3可以检测细胞的氧化应激。

[0292] 实施例5Wnt信号通路激活的检测

[0293] 使用实施例1中制作的具有响应Wnt通路的激活的TCF/LEF DNA结合蛋白结合序列作为转录调控序列的质粒4和线性DNA1,调查载体对于细胞内的Wnt通路的激活的响应。将HEK293细胞以15000cells/孔的密度接种于96孔板,用EMEM(添加NEAA)培养基培养1天。使用Lipofectamine 3000kit(ThermoFisher)将25ng/孔的质粒4和线性DNA1分别导入细胞中。导入中使用的P3000 reagent和Lipofectamine 3000reagent的量均为0.2 $\mu$ L/孔。转染起3天后,以终浓度16 $\mu$ M、8 $\mu$ M、4 $\mu$ M、2 $\mu$ M和0 $\mu$ M(DMSO)添加作为GSK3 $\beta$ 抑制剂的CHIR99021(Cayman Chemical),激活Wnt通路。

[0294] 确认药剂处理后的细胞内的Wnt通路的激活。药剂处理6小时后和24小时后,使用SuperPrep<sup>TM</sup>Cell Lysis&RT Kit for qPCR(东洋纺株式会社)从细胞中提取RNA,合成cDNA。利用RT-qPCR对已知作为Wnt通路的下游基因的AXIN2和SP5的mRNA表达量进行定量。本分析使用TB Green<sup>TM</sup>Premix Ex Taq<sup>TM</sup>II(Takara),利用SYBR Green法来实施。使用GAPDH作为内标基因。分析中使用的引物在表D中示出。分析的结果,在药剂处理后6小时和24小时这两个时间点以终浓度16 $\mu$ M、8 $\mu$ M添加CHIR99021的条件下,看到已知作为Wnt通路的下游基因的AXIN2和SP5的mRNA表达量增加,确认通过药剂添加而在细胞内产生了Wnt通路的激活(图6)。

[0295] 接下来,调查载体针对Wnt通路的激活的响应。药剂处理24小时后,从培养物中分离细胞,根据手册使用Template Prepper for DNA(日本基因)试剂盒由该细胞得到DNA溶液。将得到的DNA溶液作为模版,使用TB Green<sup>TM</sup>Premix Ex Taq<sup>TM</sup>II(Takara)利用SYBR Green法通过qPCR分析对标志物序列(m2变体)的倒位进行定量。分析中使用的引物示于表E中。标准化使用主干(Backbone)序列的定量值。分析的结果,在质粒4导入组和线性DNA1导入组这两者中检测到了依赖于CHIR99021浓度的倒位序列的增加(图7)。表明本发明的载体可以检测细胞内的Wnt通路激活。另外,根据实施例2~5的结果,表明通过变更质粒中的转录调控序列的种类而能够检测各种各样的细胞内响应。

[0296] 实施例6基于病毒载体的细胞内响应的检测和shRNA功能分析

[0297] 针对实施例1中制作的病毒载体组,调查shRNA所致的细胞内响应诱导能力和该细胞内响应的检测能力。将HEK293细胞以15000cells/孔的密度接种于96孔板,用EMEM(添加NEAA)培养基培养1天。对各培养孔以感染复数(MOI)达到0.4或2的方式导入病毒载体1A或1B,并且以达到MOI10的方式导入病毒载体1C。转导72小时后和144小时后,从培养物中分离细胞,根据手册使用Template Prepper for DNA(日本基因)试剂盒由该细胞得到DNA溶液。通过以得到的DNA溶液为模版、使用表F的引物的PCR对载体内的Lox序列间的包含mKate2的区域进行扩增。使用TB Green™Premix Ex Taq™II(Takara)利用SYBR Green法通过qPCR分析对PCR产物内的mKate2区域的倒位进行定量。定量使用表4中记载的引物,算出倒位的mKate2区域与未倒位的mKate2区域的量比。作为结果,在转导后72小时到144小时之间,在含有shRAD51的载体中mKate2序列的倒位增加了(图8)。根据其结果,表明搭载于病毒载体的shRAD51抑制RAD51的表达而诱导细胞内的DNA损伤响应,以及由载体组的报告基因系统检测到了该细胞内响应。另外,根据实施例2~6的结果,表明可以用各种各样的载体实施由shRNA产生的细胞内响应的检测。

[0298] 实施例7由质粒池中含有的各种shRNA所致的DNA损伤响应的增强效果的综合性检测

[0299] 验证了即便集合不同的质粒而导入细胞,也可以分别检测依赖于各质粒所编码的shRNA的报告基因信号。作为质粒,使用具有Scramble shRNA的质粒1和包含具有不同shRAD51序列的4种质粒的集合池的质粒2。

[0300] 制备以等摩尔比包含质粒1和质粒2的质粒混合液。将HEK293细胞以25000cells/孔的密度接种于96孔板,用EMEM(添加NEAA)培养基培养1天。使用Lipofectamine 3000kit(ThermoFisher),将合计含有25ng/孔的质粒1和2的上述质粒混合液导入细胞中。导入中使用的P3000reagent和Lipofectamine 3000reagent的量均为0.2μL/孔。转染起72小时后以终浓度1μM和0μM(DMSO)添加具有DNA损伤效果的阿霉素(Sigma—Aldrich)。

[0301] 阿霉素处理96.5小时后从培养物中分离细胞,根据手册使用Template Prepper for DNA(日本基因)试剂盒由该细胞得到DNA溶液。接着,由该DNA溶液,通过使用表5的引物的PCR对包含shRNA区域和mKate2编码序列的片段进行扩增。接下来,通过使用表6的引物的PCR对每个shRNA区域制作包含mKate2序列的扩增子(图9)。使用表7的引物和TB Green™Premix Ex Taq™II(Takara)利用SYBR Green法通过qPCR分析对含有shScramble、shRAD51\_\_2和shRAD51\_\_4的扩增子中包含的mKate2序列的倒位量进行定量。分析的结果,在添加了阿霉素的细胞中,含有shRAD51\_\_2和shRAD51\_\_4的扩增子中,与含有shScramble的扩增子相比mKate2序列的倒位率增加了(图10)。

[0302] 为了通过与qPCR不同的分析方法来确认结果,使用大规模并行测序对上述的包含shRNA区域和mKate2序列的片段进行分析。使用连接测序试剂盒(Ligation Sequencing Kit,Nanopore,SQK—LSK109)对上述扩增子附加接头,制作Nanopore分析用文库后,利用Flongle Flow Cell(Nanopore)来实施测序。接下来,制作涵盖原理上可产生的mKate2序列与shRNA序列的组的参照序列数据,通过BWA—MEM(arXiv:1303.3997[q—bio.GN],2013)来进行读段的映射。根据得到的结果,对每一种shRNA计数倒位和未倒位的mKate2序列,算出各shRNA的mKate2序列的倒位的几率。接下来,算出将shScramble的几率设为1时的各

shRNA的相对几率。分析的结果,在与shRAD51相同的读段上存在的mKate2序列的倒位率高(表8)。根据这些结果,表明即便集合包含各自不同的shRNA的多个质粒而导入细胞中,也可以一次性检测由各个shRNA引起的细胞内的响应的效果。

[0303] [表8]

[0304] 表8在与各shRNA相同的读段中存在的mKate2序列的分析结果

shRNA	原始储存 序列 (计数)	倒位储存 序列 (计数)	几率	几率比
shScramble	8389	37030	4.414	1.000
shRAD51_1	11370	51238	4.506	1.021
shRAD51_2	11835	54862	4.636	1.050
shRAD51_3	14252	66833	4.689	1.062
shRAD51_4	10122	47320	4.675	1.059

[0305]

[0306] 实施例8使用载体池的引起细胞内响应的shRNA的探索

[0307] 验证本发明的载体可以探索、确定激活Wnt通路的shRNA。作为载体,使用实施例1中制作的线性DNA2。线性DNA2是包含分别具有抑制与Wnt通路相关的基因(SETDB1、DHX29或USP7,通过它们的抑制而激活Wnt通路)的shRNA的线性DNA和具有shScramble的线性DNA的集合化的DNA片段。将HEK293细胞以15000cells/孔的密度接种于96孔板,用EMEM(添加NEAA)培养基培养1天。使用Lipofectamine 3000kit(ThermoFisher),将合计25ng/孔的线性DNA2导入细胞。导入中使用的P3000 reagent和Lipofectamine 3000reagent的量均为0.2 $\mu$ L/孔。

[0308] 脂质体转染72小时后,从培养物中分离细胞,根据手册使用Template Prepper for DNA(日本基因)试剂盒由该细胞得到DNA溶液。以该DNA溶液为模版,进行使用表G的引物的PCR对包含标志物序列(Lox间序列)和shRNA编码序列的区域进行扩增。由于使用的引物的末端附加了索引序列,因此得到的扩增子被赋予了与各培养孔对应的索引。使用连接测序试剂盒(Ligation Sequencing Kit,Nanopore,SQK-LSK110)对得到的扩增子附加接头而制作Nanopore分析文库,通过Flongle Flow Cell(Nanopore)来实施测序。对得到的读段进行基于Minibar([github.com/calacademy-research/minibar](https://github.com/calacademy-research/minibar))的索引的确定和删除。接下来,制作涵盖原理上可产生的标志物序列和shRNA序列的组合的一系列参照序列数据,利用BWA-MEM针对该参照序列来进行读段的映射。根据得到的结果,对每一种shRNA计数标志物序列的倒位的有无,算出各shRNA的倒位的几率。分析的结果,在与以涉及Wnt通路的基因为靶的shRNA相同的读段上存在的标志物序列跟在与Scramble shRNA相同的读段上存在的标志物序列相比倒位的几率高(图11)。根据这些结果,表明可以通过本发明的载体来确定激活Wnt通路的shRNA,即便集合包含各自不同的shRNA的多个载体而导入细胞中,也可以一次性检测由各个shRNA引起的细胞内的响应的效果。

[0309] 另外,通过对每种靶基因汇总在与shRNA相同的分子内存在的标志物序列的得分

来调查基因扰动对细胞内的响应的效果。将图11A的数据针对每种靶基因进行汇总,并算出平均值(图11B)。其结果,抑制各靶基因(特别是USP7和SETDB1)的载体的倒位几率的平均值与对照相比增加了。暗示这些靶基因有助于抑制细胞内的Wnt/ $\beta$ -Catenin信号通路的激活。综上,通过对每种作为功能性分子的靶的基因汇总分析结果,可以对与细胞内的响应相关的基因进行评价和排序。

[0310] 实施例9使用具有不同检测系统的载体池的引起细胞内响应的shRNA的探索

[0311] 验证实施例1中制作的线性DNA3可以探索、确定激活Wnt通路的shRNA。线性DNA3与线性DNA2同样地是包含具有抑制与Wnt通路相关的基因的shRNA的线性DNA和具有shScramble的线性DNA的集合化的DNA片段。线性DNA3具有编码sesRNA的序列作为翻译调控区域,通过由线性DNA3表达的sesRNA随着位于Wnt通路的下游的AXIN2的mRNA表达而增加下游基因(CRE)的表达从而检测出Wnt通路的激活。将HEK293细胞以15000cells/孔的密度接种于96孔板,用EMEM(添加NEAA)培养基培养1天。使用Lipofectamine 3000kit(ThermoFisher),将合计5ng/孔的线性DNA3导入细胞中。导入中使用的P3000 reagent和Lipofectamine 3000reagent的量均为0.2 $\mu$ L/孔。

[0312] 脂质体转染72小时后,从培养物中分离细胞,根据手册使用Template Prepper for DNA(日本基因)试剂盒由该细胞得到DNA溶液。以该DNA溶液为模版,进行使用表G的引物的PCR对包含标志物序列(Lox间序列)和shRNA序列的区域进行扩增。由于使用的引物的末端附加了索引序列,因此得到的扩增子附加了与实验条件对应的索引。使用连接测序试剂盒(Ligation Sequencing Kit,Nanopore,SQK-LSK110)对得到的扩增子附加接头,制作Nanopore分析用文库,通过Flongle Flow Cell(Nanopore)来实施测序。对得到的读段进行基于Minibar的索引的确定和删除。接下来,制作涵盖原理上可产生的标志物序列和shRNA序列的组合的一系列参照序列数据,利用BWA-MEM针对该参照序列来进行读段的映射。根据得到的结果,对每种shRNA计数标志物序列的倒位的有无,算出各shRNA的由Lox序列夹着的区域的倒位几率和由Rox序列夹着的序列的缺失几率。分析的结果,在与以涉及Wnt通路的基因为靶的shRNA(DHX29\_\_1~DHX29\_\_4)相同的读段上存在的标志物序列跟在与Scramble shRNA相同的读段上存在的标志物序列相比倒位和缺失的几率高(图12)。根据这些结果,表明可以利用本实施例中使用的载体池来探索可激活细胞内Wnt通路的shRNA序列或者与Wnt通路相关的基因。另外,根据实施例7~8,表明了本发明的载体池可以使用各种各样的转录调控序列和报告基因系统用以检测由shRNA引起的细胞内的响应的效果。

[0313] 实施例10通过导入UMI序列来提高检测能力

[0314] 1) 线性DNA4的制作

[0315] 将载体中包含的报告基因系统由条形码(Unique Molecular Identifier;UMI)序列来识别。制作在线性DNA1或2的shRNA编码序列与报告基因系统序列之间导入了50bp的UMI序列的线性DNA4(图13)。作为UMI序列,委托株式会社FASMAC制作随机的单链寡DNA的集合池。

[0316] 委托Integrated DNATechnologies株式会社制作包含编码以涉及Wnt通路的基因为靶的shRNA(表B)的序列的单链DNA片段的集合池。以上述质粒4为模版,利用表H1中示出的引物进行PCR,将得到的PCR产物利用Wizard<sup>TM</sup>SV Gel and PCR Clean-Up System(Promega)进行纯化,制作DNA片段。这些DNA片段被设计为相互重叠的。将这些DNA片段和包

含UMI的DNA片段以图14A和B所示的组合进行混和,利用NEBuilder(New England Biolabs Japan株式会社)进行结合而得到包含shScramble的线性DNA4A和包含以Wnt相关基因为靶的shRNA的线性DNA4B。使用表H2的引物进行以线性DNA4A和4B分别为模版的PCR,将PCR产物利用Wizard™SV Gel and PCR Clean-Up System(Promega)进行纯化。混合线性DNA4A的溶液和线性DNA4B的溶液而得到线性DNA4。线性DNA4具有:包含编码以Wnt通路为靶的shRNA的序列或shScramble的功能性分子区域、与质粒4同样的转录调控序列和报告基因系统、以及用于使这些序列在细胞内发挥功能的序列,功能性分子区域为异源DNA片段集合。线性DNA4中的各DNA片段包含了识别各个片段的UMI序列。

[0317] 2) 质粒6的制作

[0318] 如图15所示地设计hyPBase(高活性piggyBac转移酶,hyperactive piggyback transposase)表达质粒,由Vector Builder Japan株式会社制作。将得到的质粒作为质粒6。

[0319] 3) UMI序列的效果的验证

[0320] 将HEK293细胞以15000cells/孔的密度接种于96孔板,用EMEM(添加NEAA)培养基培养1天。使用Lipofectamine 3000kit(ThermoFisher),将合计5ng/孔的线性DNA4和95ng/孔的质粒6同时导入细胞中。线性DNA直链4具有ITR序列(图13),该序列通过由质粒6表达的hyPBase的活性而整合到基因组中。线性DNA4是将分别具有抑制与Wnt相关的基因的shRNA和Scramble shRNA的线性DNA集合化的DNA片段。导入中使用的P3000 reagent和Lipofectamine 3000reagent的量均为0.2 $\mu$ L/孔。

[0321] 脂质体转染后72小时后从培养物中分离细胞,根据手册使用Template Prepper for DNA(日本基因)试剂盒由该细胞得到DNA溶液。以该DNA溶液为模版作为,使用表G的引物,对包含标志物序列(Lox间序列)、UMI和shRNA序列的区域进行扩增。由于使用的引物的末端附加了索引序列,因此得到的扩增子附加了与实验条件对应的索引。使用连接测序试剂盒(Ligation Sequencing Kit,Nanopore,SQK-LSK110)对得到的扩增子附加接头,制作Nanopore分析文库,通过Flongle Flow Cell(Nanopore)来实施测序。对得到的读段进行基于Minibar的索引的确定和删除。接下来,制作涵盖原理上可产生的标志物序列和shRNA序列的组合的一系列参照序列数据,利用BWA-MEM针对该参照序列来进行读段的映射。另外,根据由映射得到的SAM文件和读段的信息来确定UMI区域,并通过CD-HIT来对UMI序列进行聚类,由此判断得到的UMI是否为唯一的。分成不利用UMI序列进行过滤的情况和仅用具有唯一的UMI序列的读段进行分析的情况,对各读段计数标志物序列的倒位的有无,算出各shRNA(shScramble、DHX29\_\_1~DHX29\_\_4)的倒位几率。分析的结果,在与以Wnt相关基因为靶的shRNA(DHX29\_\_1~DHX29\_\_4)相同的读段上存在的标志物序列跟在与Scramble shRNA相同的读段上存在的标志物序列相比倒位几率高,该趋势在仅分析具有唯一的UMI的读段的情况下更明显(图16)。根据这些结果,认为通过在报告基因系统中导入UMI,能够避免伴随分析时的序列扩增的偏差,提高基于本发明载体的目标shRNA的检测能力。

[0322] 实施例11转录调控序列的活性评价

[0323] 1) 质粒7的制作

[0324] 以表I所示的模版、引物的组合来实施PCR,除此以外,按照与质粒1同样的步骤来制作质粒载体(质粒7),确认分子量。作为模版而准备的minP变体序列是包含原始序列的94

种minP变体序列的集合池。本质粒在功能性分子的区域具有shScramble,具有p53RE作为转录调控序列,具有CRE-loxP样序列作为报告基因系统。该CRE-loxP样序列具有CRE基因序列和相互反向配置的2个Lox2722-LoxP序列以及配置于其间的mKate2序列。转录调控序列激活时,该mKate2序列在CRE的作用下倒位而进行转录、表达(参照图1B)。质粒7具有位于p53RE与CRE之间的具有各种多样性的最小启动子(minP变体)序列,是分别具有共计94种minP变体序列的载体的集合池。

[0325] 2) 转录调控序列的活性评价

[0326] 使用质粒7来评价转录调控序列针对细胞内响应的活性,对这样的评价系统进行研究。将HEK293细胞以25000cells/孔的密度接种于96孔板,用EMEM(添加NEAA)培养基培养1天。使用Lipofectamine 3000kit(ThermoFisher),将合计25ng/孔的质粒7导入细胞。导入中使用的P3000reagent和Lipofectamine 3000reagent的量均为0.2 $\mu$ L/孔。转染起4天后以终浓度5 $\mu$ M和0 $\mu$ M(DMSO)添加具有DNA损伤效果的阿霉素(Sigma-Aldrich)。

[0327] 阿霉素处理24小时后从培养物中分离细胞,根据手册使用Template Prepper for DNA(日本基因)试剂盒由该细胞得到DNA溶液。以该DNA溶液为模版,使用表J的引物通过PCR对包含minP变体序列和mKate2序列的片段进行扩增。使用连接测序试剂盒(Ligation Sequencing Kit,Nanopore,SQK-LSK110)将接头附加于得到的扩增子,制作Nanopore分析文库后,通过Flongle Flow Cell(Nanopore)来实施测序。制作涵盖原理上可产生的mKate2序列和minP变体序列的组合的一系列参照序列数据,利用BWA-MEM针对该参照序列来进行读段的映射。根据得到的结果,对各minP变体序列计数倒位和未倒位的mKate2序列,算出各minP变体序列中的mKate2序列的倒位的几率。分析的结果,检测到与原始minP序列相比对阿霉素的刺激反映更敏锐的minP变体序列(图17)。根据这些结果,表明可以构建用于评价针对细胞内响应的转录调控序列等载体成分的活性的载体系统。

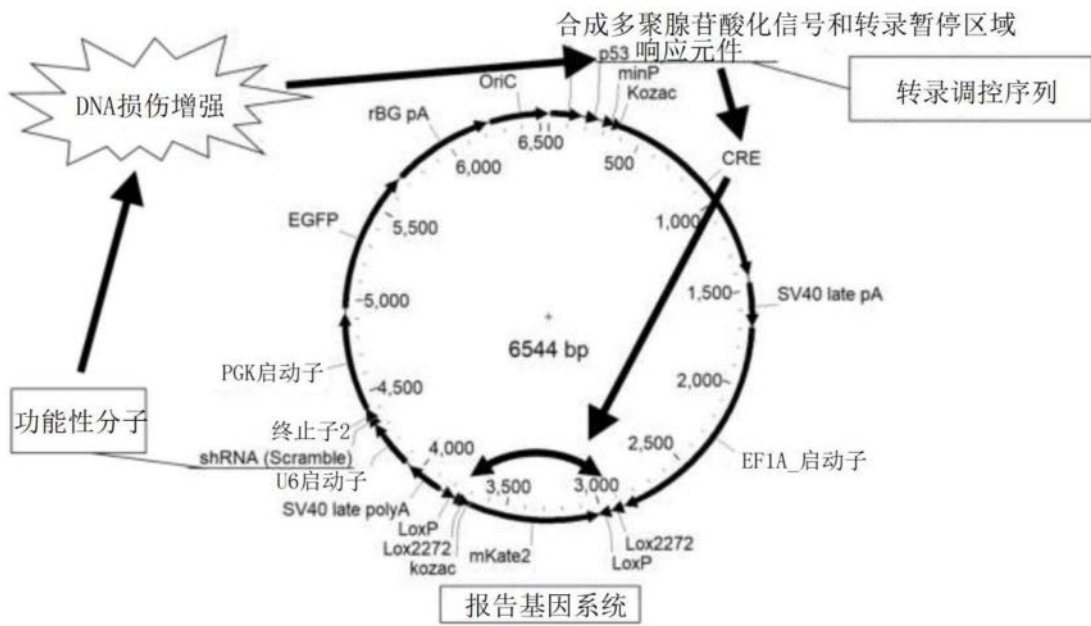


图1A

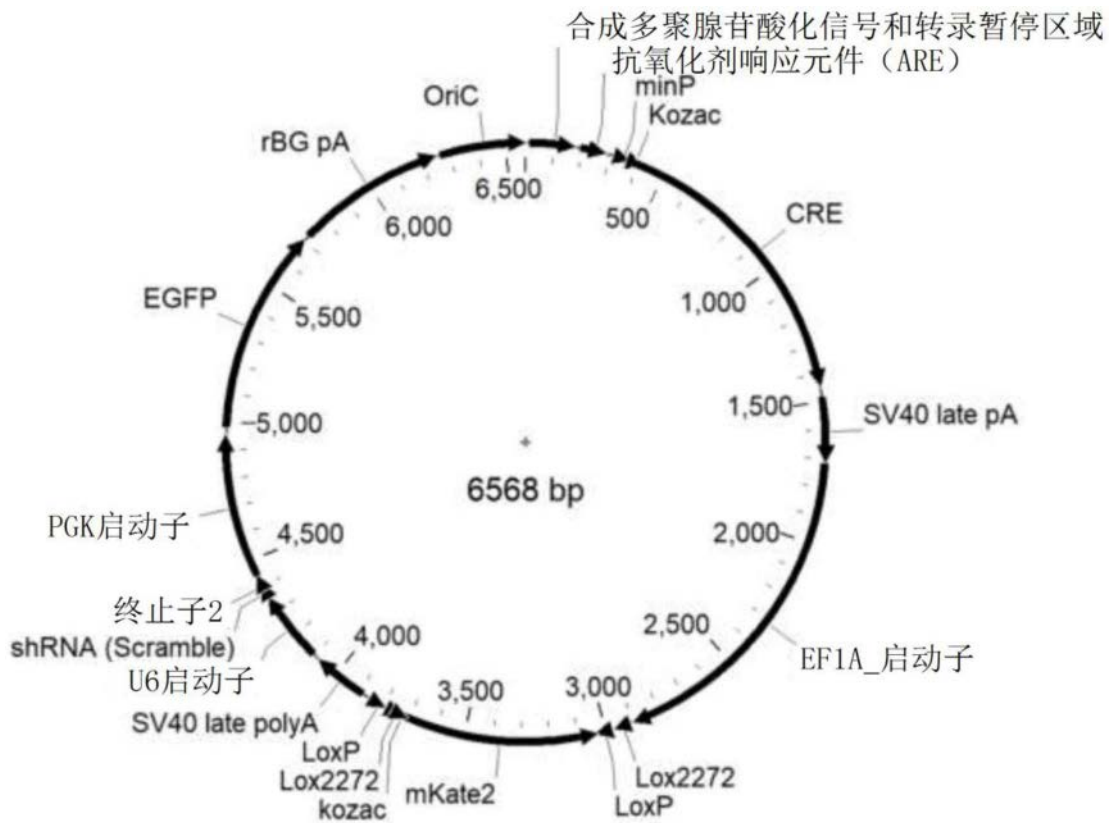


图1B

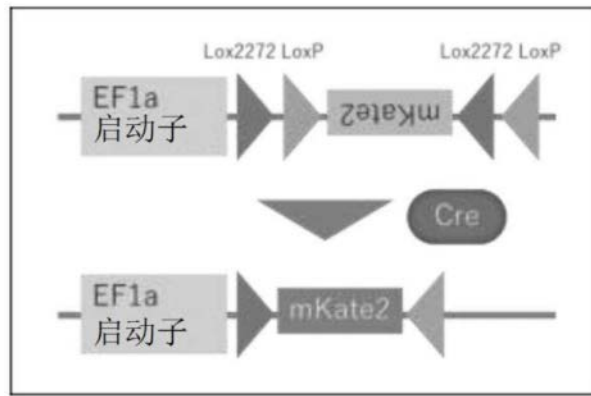


图1C

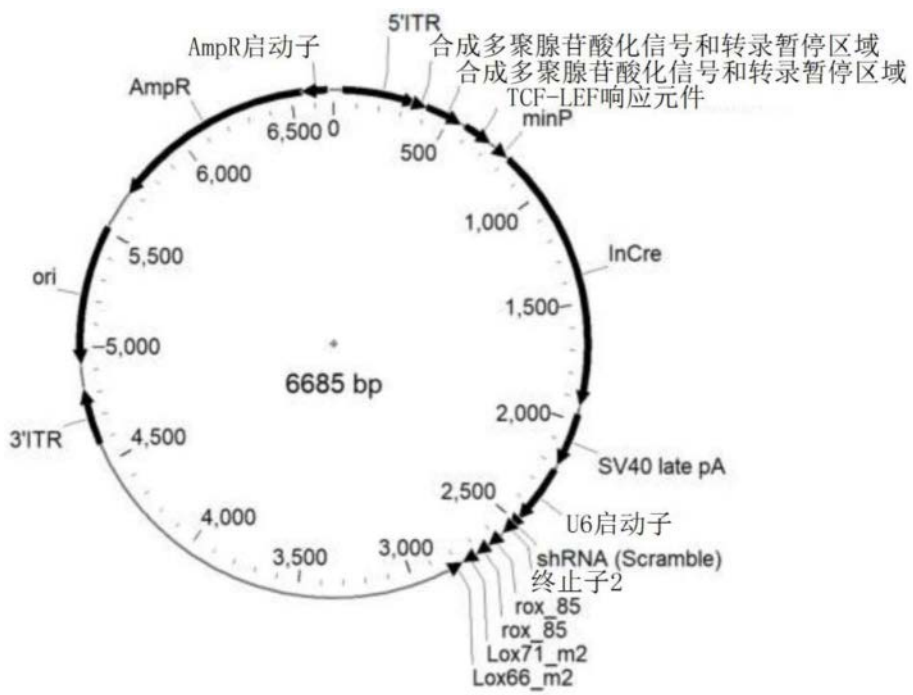


图1D



图1E

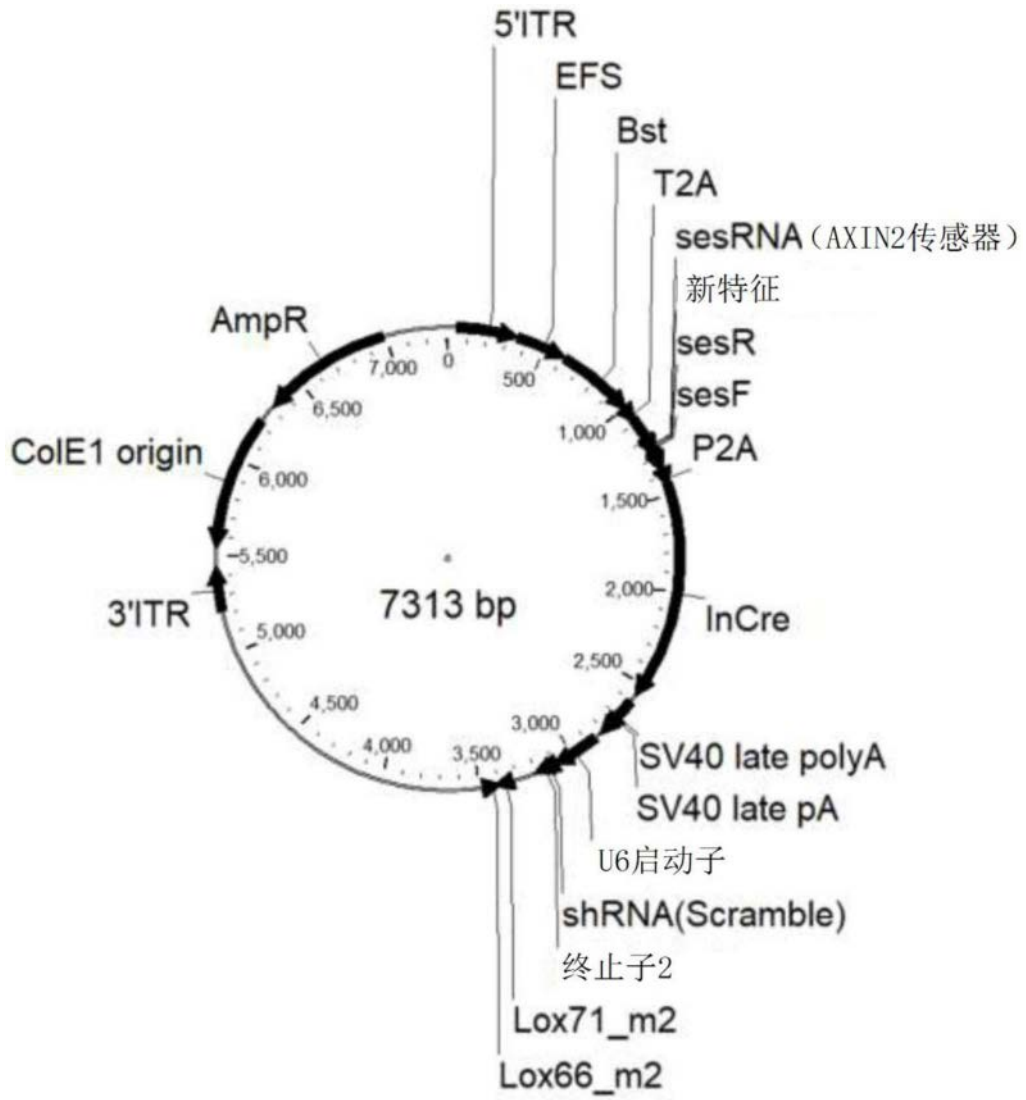


图1F

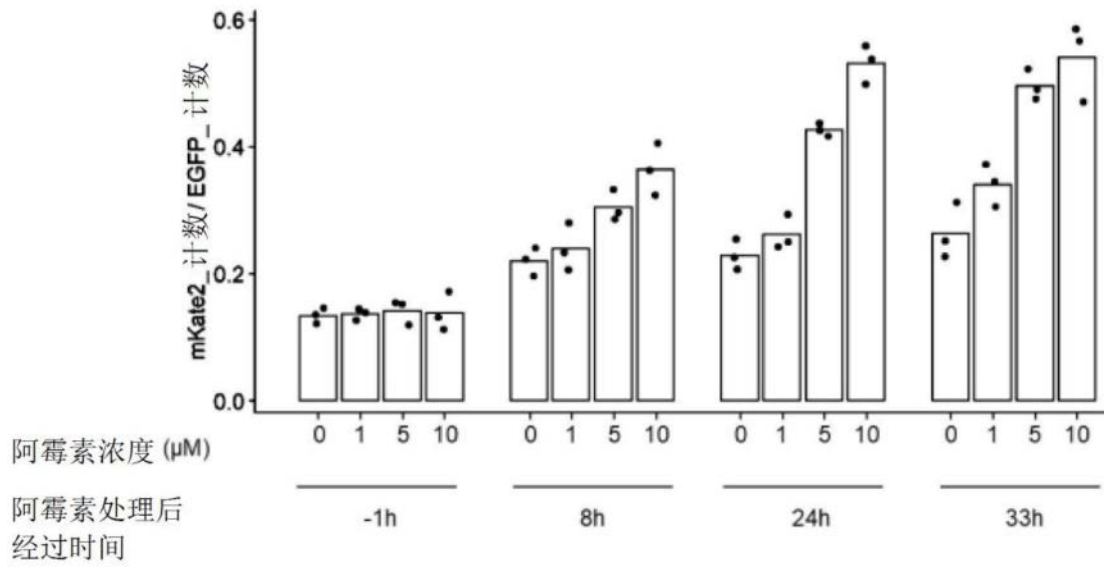


图2

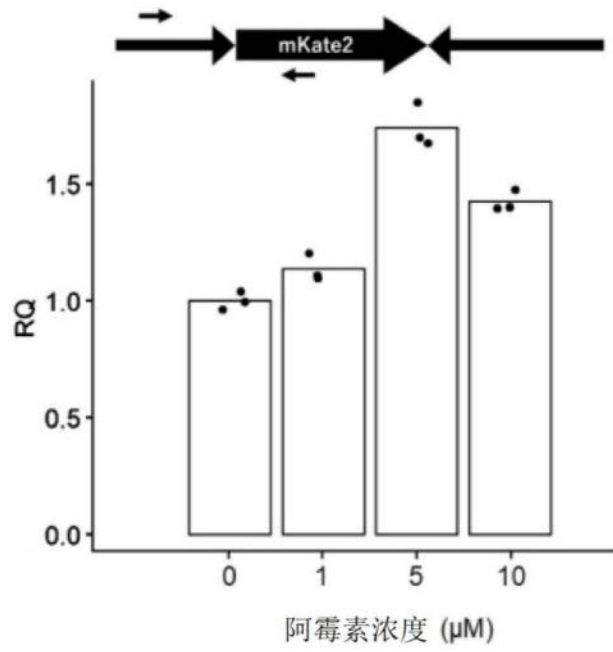


图3

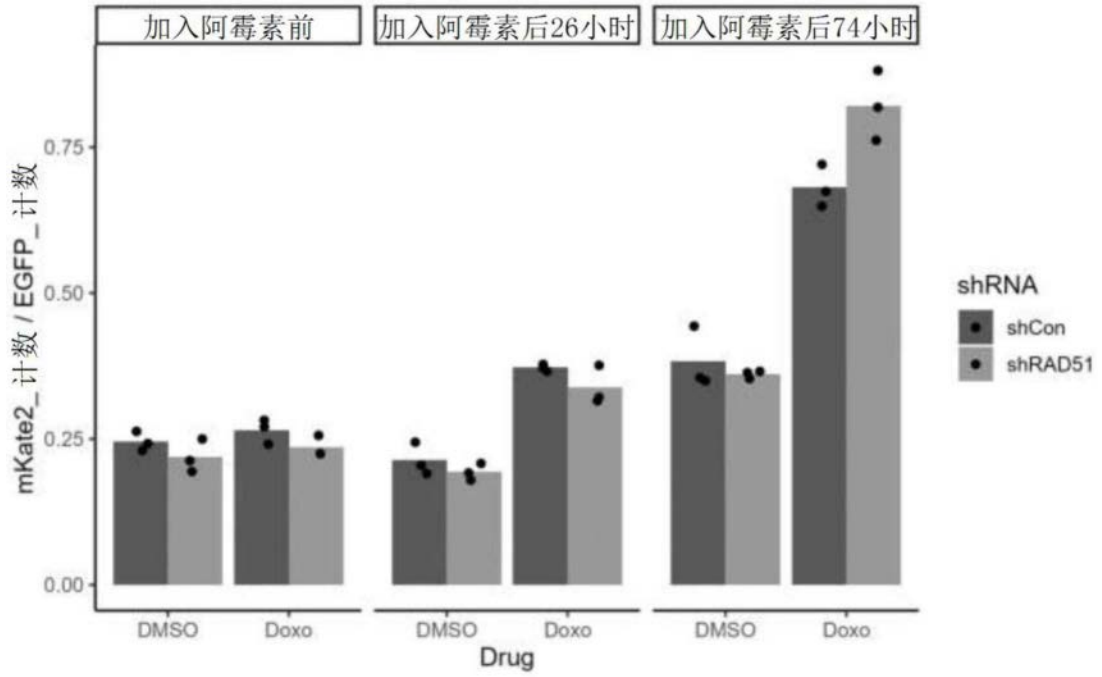


图4

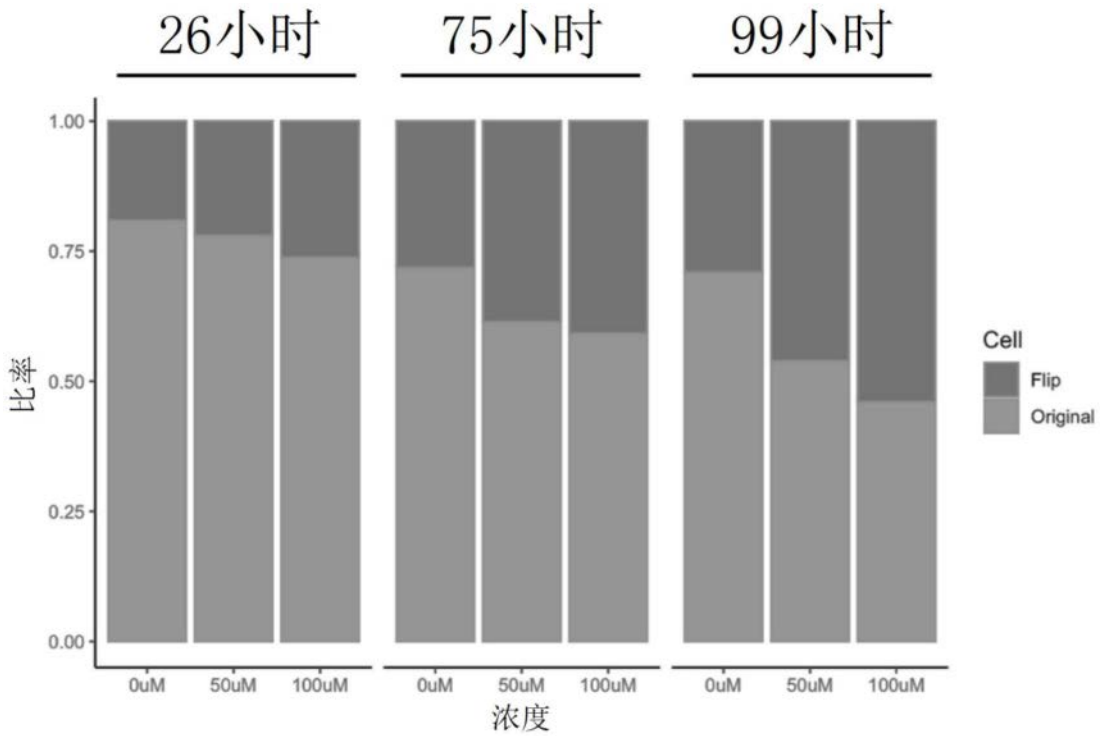


图5

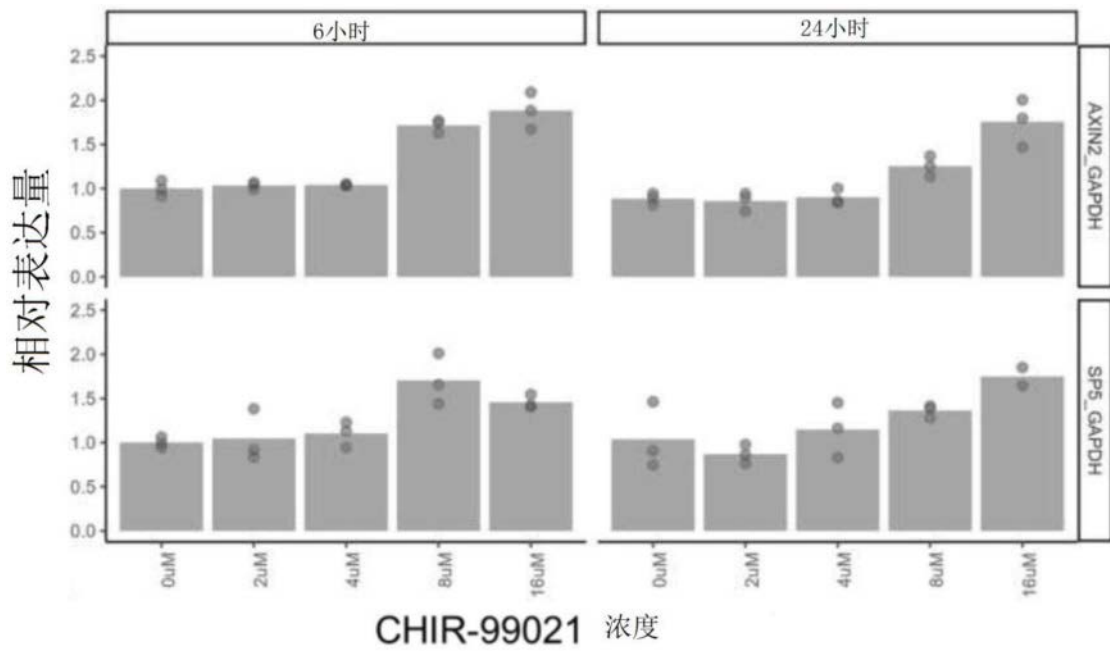


图6

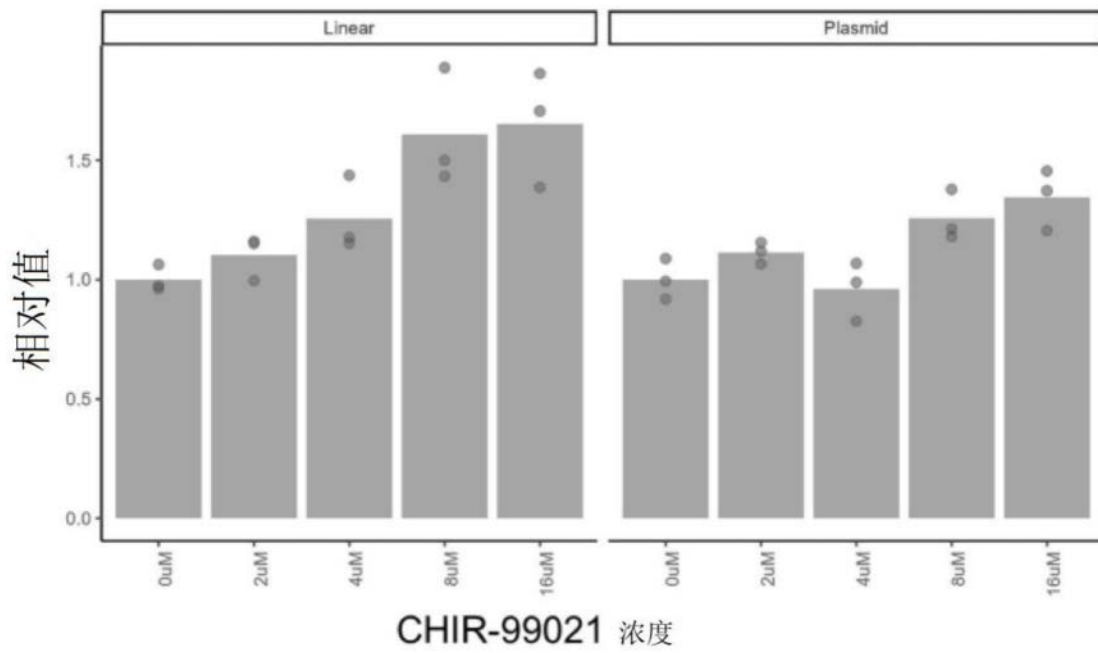


图7

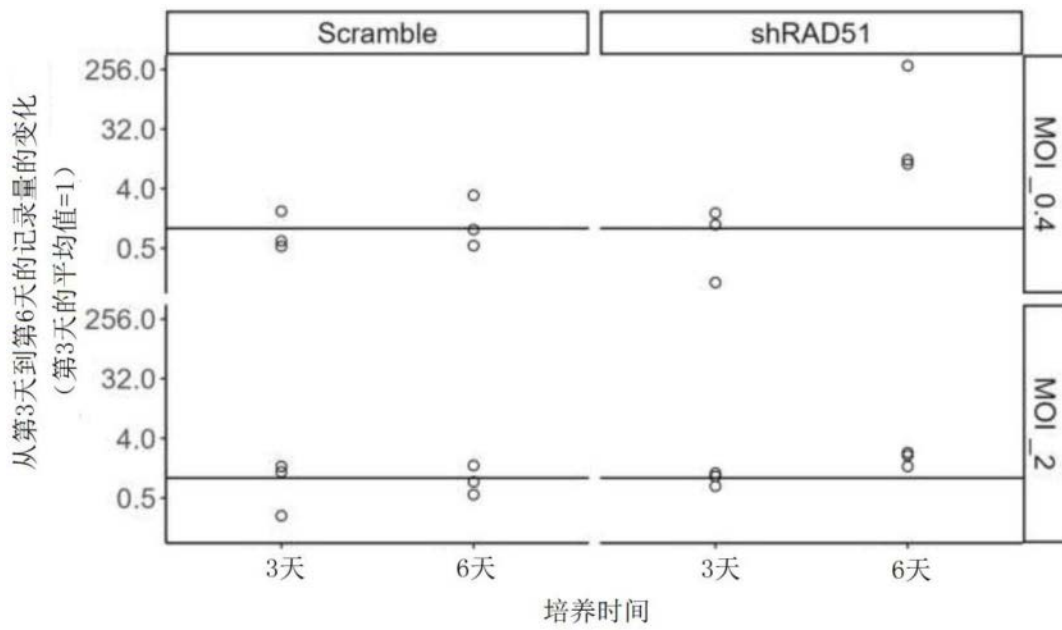


图8

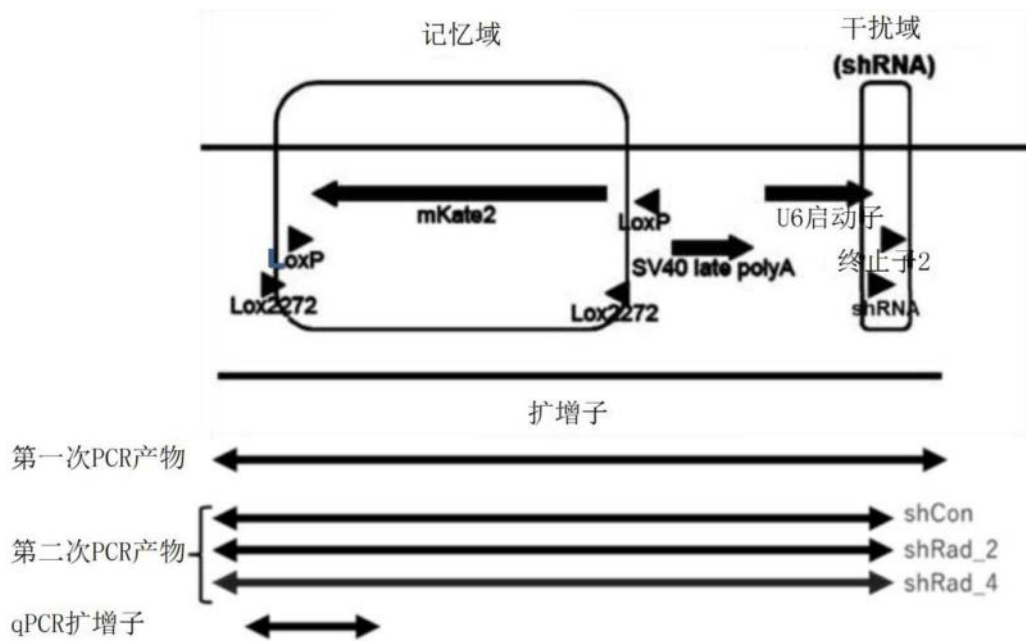


图9

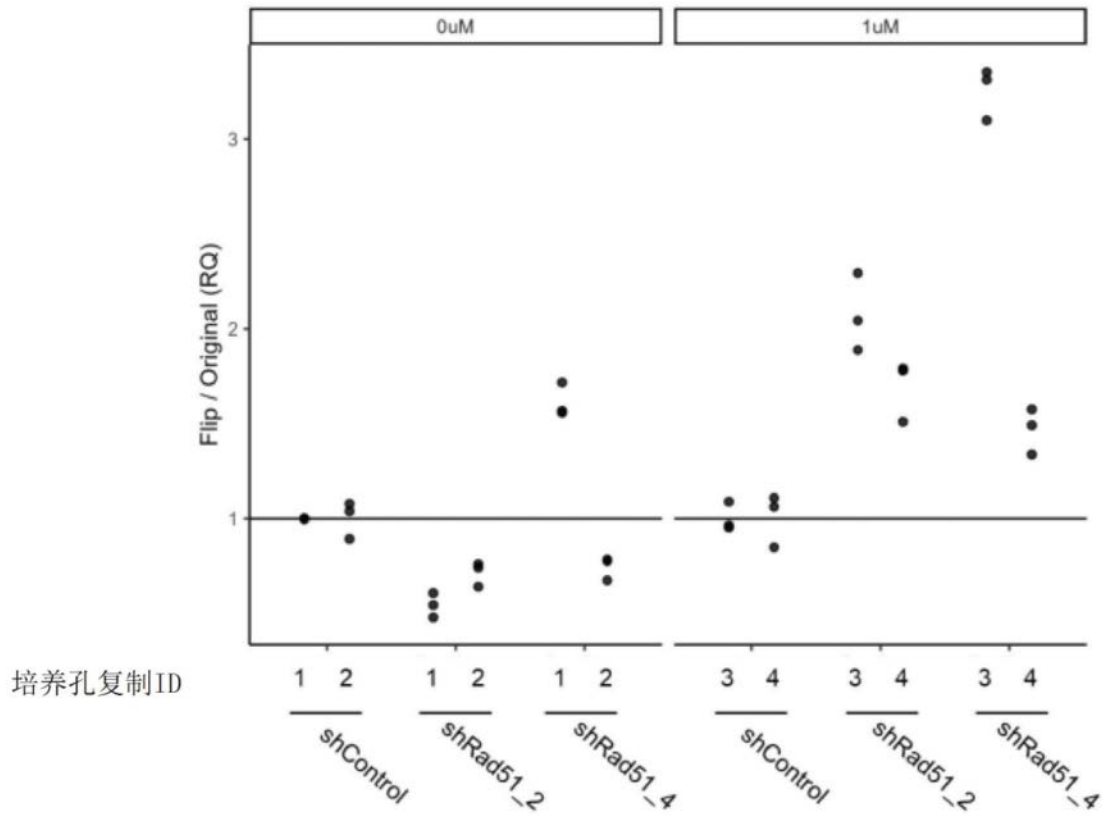


图10

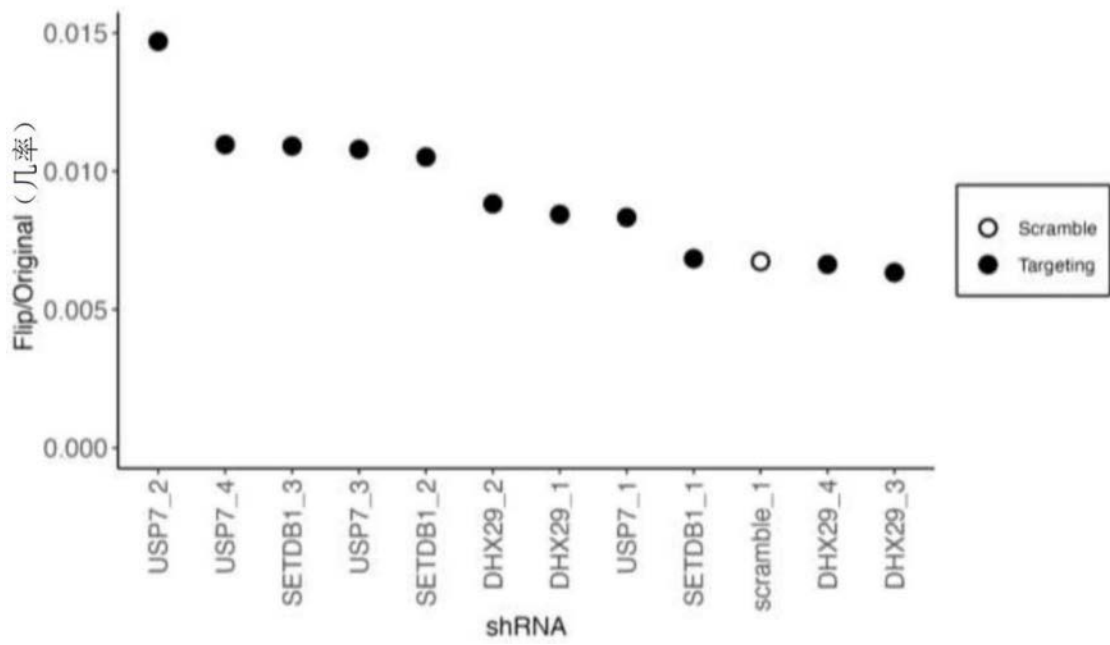


图11A

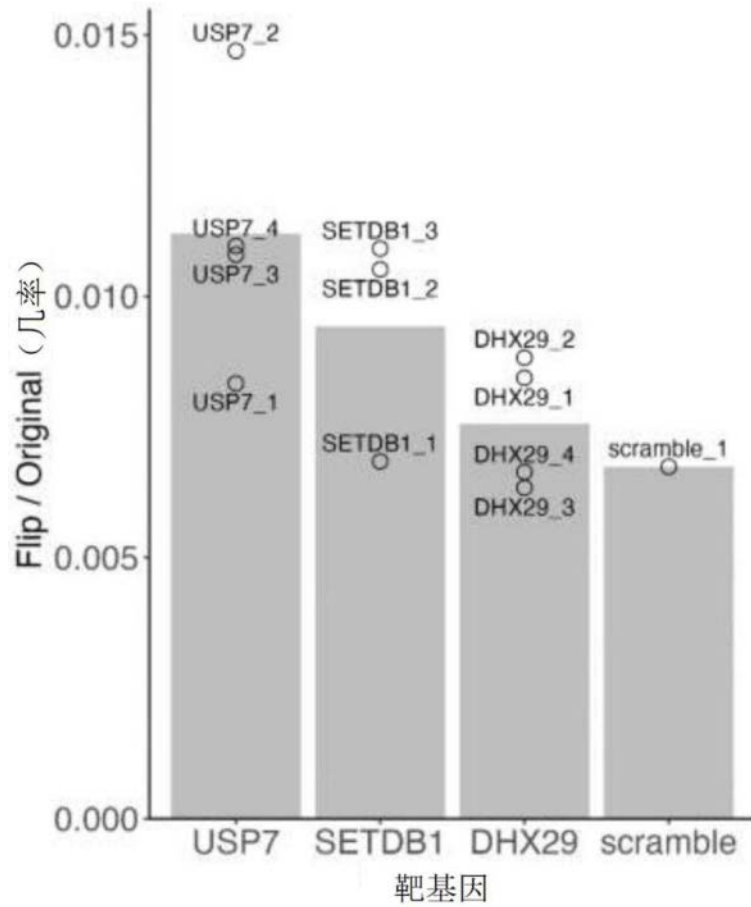


图11B

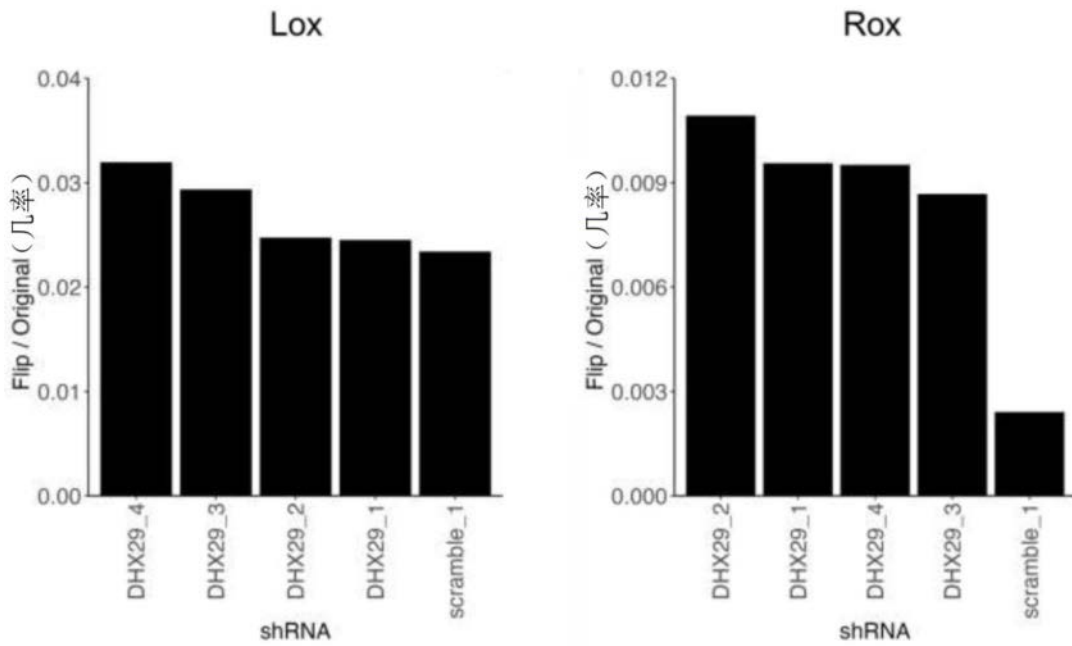


图12

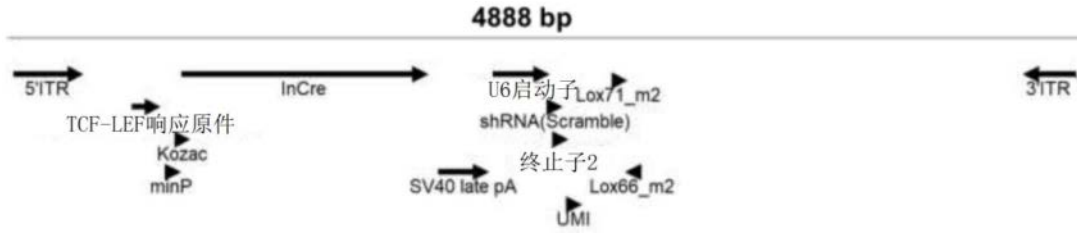


图13

线性DNA4A (混合物1)

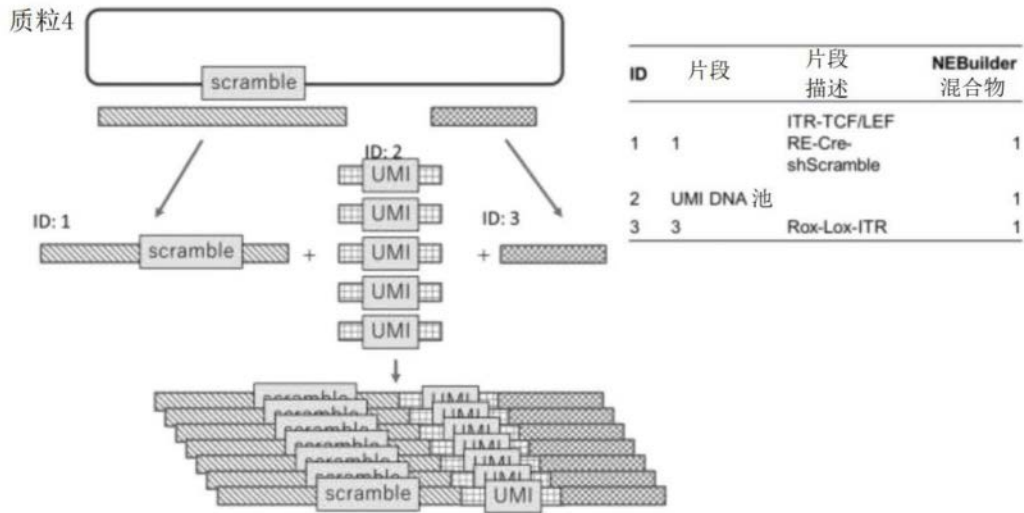


图14A

线性DNA4B (混合物2)

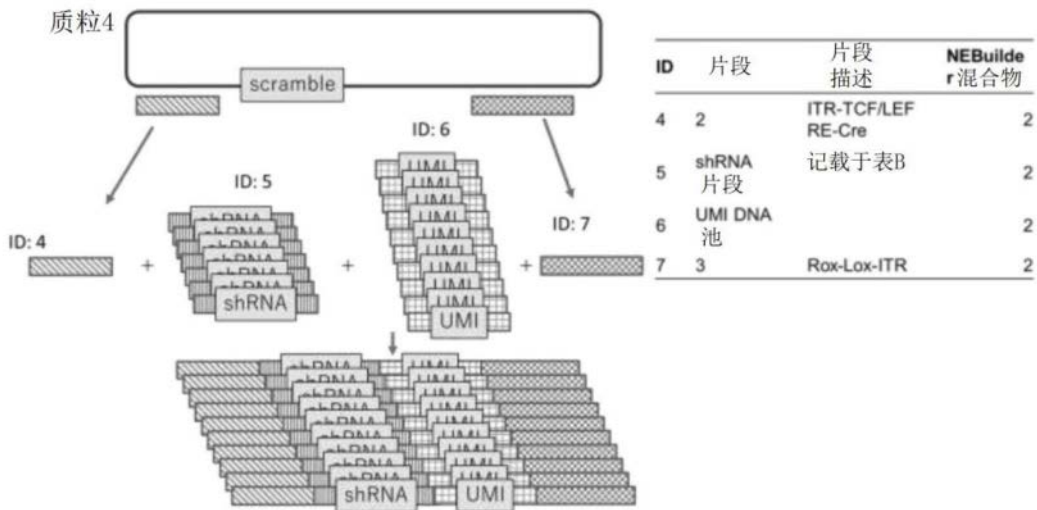


图14B

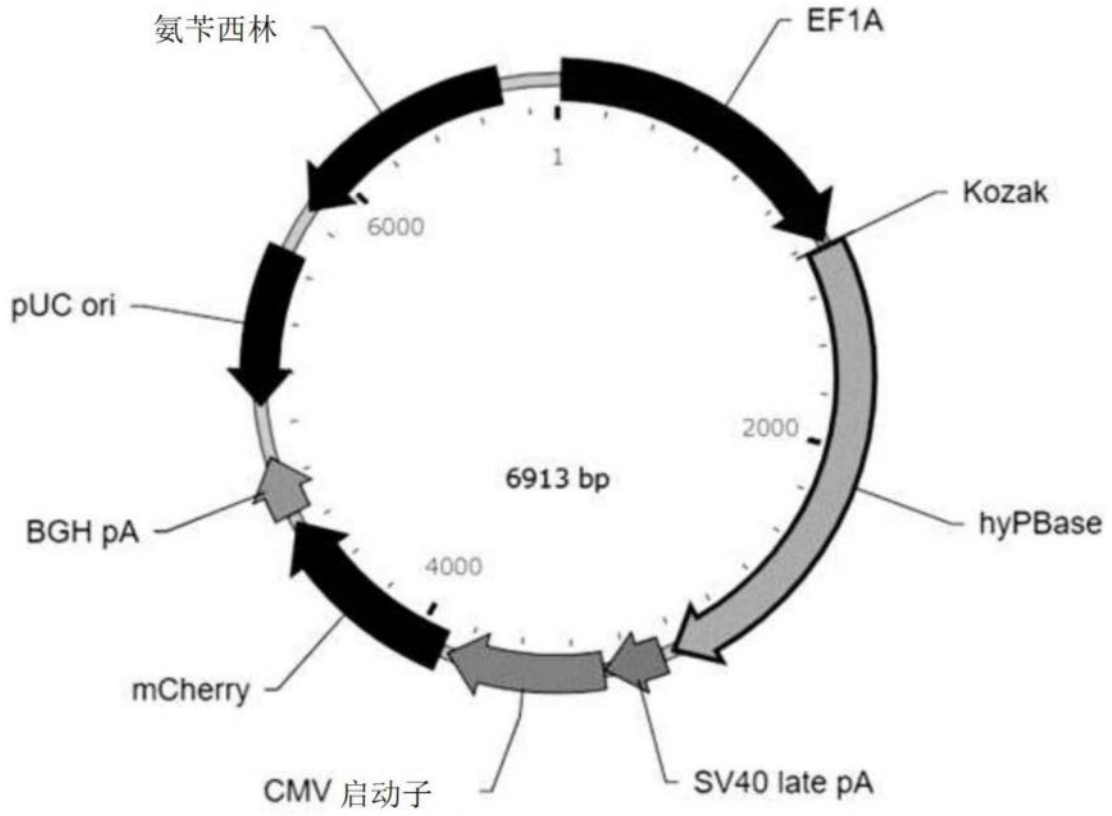


图15

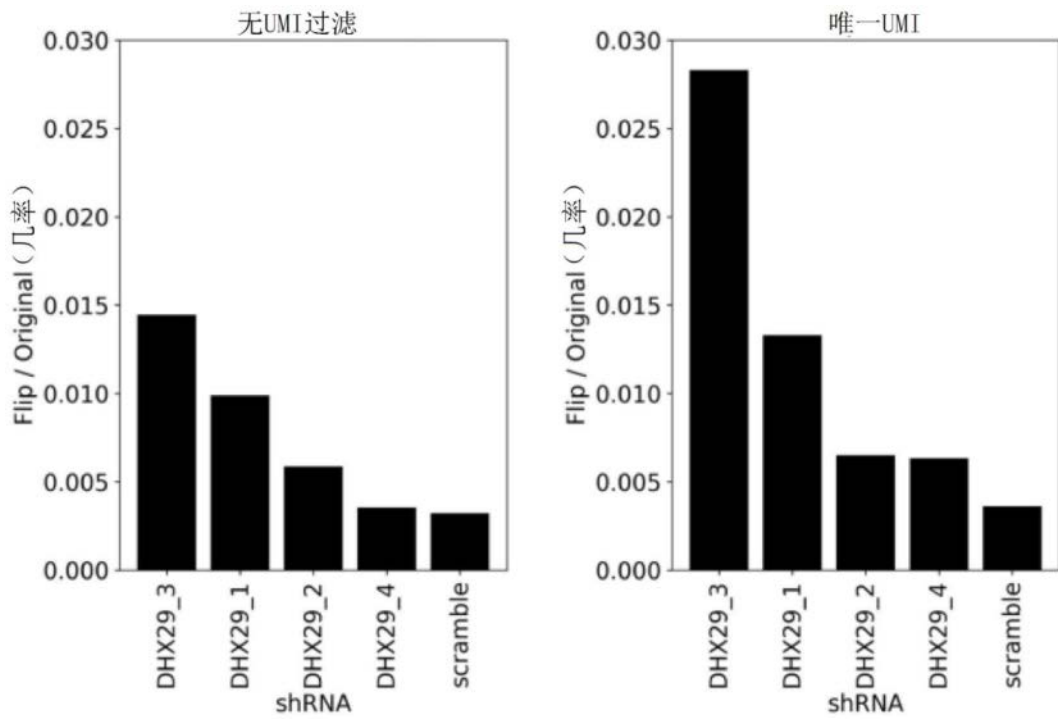


图16

