



## (12)发明专利

(10)授权公告号 CN 105308188 B

(45)授权公告日 2019.03.08

(21)申请号 201480028611.5

(22)申请日 2014.05.19

(65)同一申请的已公布的文献号

申请公布号 CN 105308188 A

(43)申请公布日 2016.02.03

(30)优先权数据

2013-105481 2013.05.17 JP

(85)PCT国际申请进入国家阶段日

2015.11.17

(86)PCT国际申请的申请数据

PCT/JP2014/063214 2014.05.19

(87)PCT国际申请的公布数据

W02014/185549 JA 2014.11.20

(73)专利权人 东芝医疗系统株式会社

地址 日本栃木

(72)发明人 吉村齐湖 石原美津子 赤星英一

樱井康雄

(74)专利代理机构 中国国际贸易促进委员会专

利商标事务所 11038

代理人 郑天松

(51)Int.Cl.

C12Q 1/6897(2018.01)

C12M 1/00(2006.01)

C12N 15/09(2006.01)

(56)对比文件

CN 102286524 A, 2011.12.21,

Guillermo Barreto et al.. "Gadd45a

promotes epigenetic gene activation by repair-mediated DNA demethylation".

《Nature》.2007,第445卷

Andreas Hasse et al.. "De novo

methylation of transfected CAT gene

plasmid constructs in F9 mouse embryonal

carcinoma cells".《Biochimica et

Biophysica Acta》.1992,第1131卷(第1期),

审查员 江涵

权利要求书4页 说明书35页

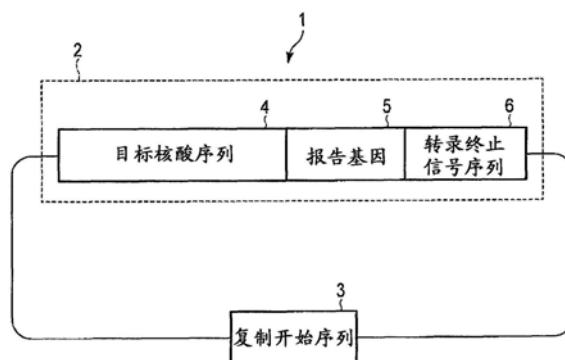
序列表10页 附图32页

### (54)发明名称

方法、自我复制载体、测定试剂盒及分析装置

### (57)摘要

本发明由实施方式提供得到细胞的表观遗传的信息的方法。此方法包括,将报告子核酸构建体导入被检细胞的核内,使报告子核酸构建体自我复制,检测在被检细胞中发生的信号的有无及/或大小,基于得到的结果,得到被检细胞的表观遗传的信息。报告子核酸构建体特征在于将被检细胞的基因组上的特定的序列中的修饰的状态抄录到被检细胞的在核内的自我复制中由官能团的取代对应的其序列上,依赖于抄录的官能团的存在状况发生可检测的信号。



1. 得到细胞的表观遗传的信息的方法, 其中

将被检细胞的基因组上的待观察的特定的序列中的修饰的状态抄录到在上述被检细胞的核内的自我复制中由官能团的取代而对应的其序列上,

使用依赖于所述抄录的官能团的存在状况发生可检测的信号自我复制载体得到细胞的表观遗传的信息, 其中, 上述自我复制载体是含下列 (a) ~ (c) 的自我复制载体:

(a) 含下列序列的第1报告子基因表达单位:

含甲基可被取代的碱基, 有依赖于上述碱基中的所述甲基的取代的程度的启动子活性, 并且对于被检细胞的基因组上的待观察的特定的序列有相同性的目标核酸序列,

在上述目标核酸序列的下游功能性地连接, 由上述目标核酸序列的活化而表达的编码报告子蛋白质的报告子基因,

在上述报告子基因的下游功能性地连接的转录终止信号序列,

(b) 在上述第1报告子基因表达单位的下游连接的含下列序列的第2报告子基因表达单位:

含甲基可被取代的碱基, 有依赖于上述碱基中的所述甲基的取代的程度的启动子活性, 并且对于被检细胞的基因组上的待观察的特定的序列有相同性的目标核酸序列,

在上述目标核酸序列的下游功能性地连接, 由上述目标核酸序列的活化而表达的编码报告子蛋白质的报告子基因,

在上述报告子基因的下游功能性地连接的转录终止信号序列, 及

(c) 与上述第1及第2报告子基因表达单位存在于相同核酸上的复制起始序列,

上述自我复制载体的上述第1报告子基因表达单位及上述第2报告子基因表达单位各自所含的上述目标核酸的甲基化状态互相不同,

所述方法包括:

(1) 向被检细胞导入上述自我复制载体,

(2) 通过在复制起始蛋白质表达的条件下培养上述被检细胞, 使上述自我复制载体复制,

(3) 检测各自来源于所述第1报告子基因表达单位及所述第2报告子基因表达单位的第1报告子蛋白质及第2报告子蛋白质及/或测定其表达量,

(4) 比较上述 (3) 中得到的上述第1报告子蛋白质和上述第2报告子蛋白质的表达量, 得到关于在所述被检细胞的基因组上的特定的序列中所含的核酸的甲基的取代的信息。

2. 权利要求1所述的方法, 其特征在于, 为了开始上述自我复制载体的自我复制, 在上述被检细胞的核内, 活化所述复制起始序列的复制起始蛋白质被表达。

3. 权利要求2所述的方法, 其特征在于, 为了开始上述自我复制载体的自我复制, 向上述被检细胞的核内导入含表达所述复制起始蛋白质的复制起始蛋白质单位的复制起始蛋白质基因表达载体。

4. 权利要求2所述的方法, 其特征在于, 上述自我复制载体还含表达所述复制起始蛋白质的复制起始蛋白质单位。

5. 权利要求2所述的方法, 其特征在于,

上述自我复制载体还含:

在上述第1报告子基因表达单位的下游功能性地连接的IRES序列、

在上述IRES序列的下游功能性地连接的编码上述复制起始蛋白质的序列，

上述第1报告子基因表达单位的上述转录终止信号序列在编码上述复制起始蛋白质的序列的下游功能性地连接，

上述复制起始序列在上述转录终止信号序列的下游功能性地连接。

6. 权利要求3所述的方法，其中上述复制起始蛋白质单位含：

编码所述复制起始蛋白质的序列，及

在其上游功能性地连接的组成型表达的启动子。

7. 权利要求1所述的方法，其特征在于，上述目标核酸序列中的官能团可被取代的所述碱基是至少1个胞嘧啶及/或鸟嘌呤。

8. 权利要求1所述的方法，其中上述官能团是甲基。

9. 权利要求1所述的方法，其中上述报告子基因选自：荧光素酶基因、 $\beta$ -半乳糖苷酶基因、一氧化氮合成酶基因、黄嘌呤氧化酶基因、蓝色荧光蛋白质基因、绿色荧光蛋白质基因、红色荧光蛋白质基因及重金属结合蛋白质基因。

10. 权利要求6所述的方法，其特征在于，编码上述复制起始蛋白质的基因和上述复制起始序列的组合选自以下的组合：

(1) 编码上述复制起始蛋白质的基因是猿猴病毒40的大T抗原基因，上述复制起始序列是来自猿猴病毒40的复制起始序列；

(2) 编码上述复制起始蛋白质的基因是爱泼斯坦-巴尔病毒的EBNA-1基因，上述复制起始序列是来自爱泼斯坦-巴尔病毒的复制起始序列；及

(3) 编码上述复制起始蛋白质的基因是小鼠多瘤病毒的大T抗原基因，上述复制起始序列是来自小鼠多瘤病毒的复制起始序列。

11. 权利要求10所述的方法，其中上述大T抗原基因是

SEQ ID NO:7、SEQ ID NO:8或SEQ ID NO:9所示的核酸、或者

由SEQ ID NO:7、SEQ ID NO:8或SEQ ID NO:9的序列中的1或数个碱基缺失、取代或附加的序列显示，且有DNA复制起始功能的核酸。

12. 权利要求1所述的方法，其中上述被检细胞是自对象采集的血液中所含的细胞。

13. 得到用于判断细胞的特性的信息的非诊断目的的方法，其基于由权利要求1所述的方法得到的被检细胞的表现遗传的信息，得到用于判断分离的上述被检细胞的特性的信息，其中上述被检细胞不是刚从生物体分离的细胞，而是培养细胞。

14. 权利要求13所述的方法，其中上述特性是药剂感受性。

15. 得到关于细胞是癌化的细胞的可能性大小的信息的非诊断目的的方法，其基于由权利要求1所述的方法得到的被检细胞的表现遗传的信息，得到关于分离的所述被检细胞是癌化的细胞的可能性大小的信息，其中上述被检细胞不是刚从生物体分离的细胞，而是培养细胞。

16. 基于由权利要求12~15之任一项所述的方法得到的结果，对于采集了所述被检细胞的对象，判断药剂感受性，或者选择在外科手术后要使用的药剂或免疫疗法剂的种类的非诊断目的的方法。

17. 得到关于细胞罹患特定的疾病的可能性大小的信息的非诊断目的的方法，其基于由权利要求1所述的方法得到的被检细胞的表现遗传的遗传信息，得到关于分离的所述被

检细胞罹患特定的疾病的可能性大小的信息,其中上述被检细胞不是刚从生物体分离的细胞,而是培养细胞。

18. 用于得到细胞的表观遗传的信息的自我复制载体,其包含:

(a) 含下列序列的第1报告子基因表达单位:

含官能团可被取代的碱基,有依赖于上述碱基中的所述甲基的取代的程度的启动子活性,并且对于被检细胞的基因组上的待观察的特定的序列有相同性的目标核酸序列,

在上述目标核酸序列的下游功能性地连接,由上述目标核酸序列的活化而表达的编码报告子蛋白质的报告子基因,

在上述报告子基因的下游功能性地连接的转录终止信号序列、

(b) 在上述第1报告子基因表达单位的下游连接的含下列序列的第2报告子基因表达单位:

含官能团可被取代的碱基,有依赖于上述碱基中的所述甲基的取代的程度的启动子活性,并且对于被检细胞的基因组上的待观察的特定的序列有相同性的目标核酸序列,

在上述目标核酸序列的下游功能性地连接,由上述目标核酸序列的活化而表达的编码报告子蛋白质的报告子基因,

在上述报告子基因的下游功能性地连接的转录终止信号序列,及

(c) 与上述第1及上述第2报告子基因表达单位存在于相同核酸上的复制起始序列,

上述第1报告子基因表达单位及上述第2报告子基因表达单位各自所含的上述目标核酸的甲基化状态互相不同。

19. 权利要求18所述的载体,其特征在于,上述第1报告子基因表达单位和上述第2报告子基因表达单位各自含有的上述报告子基因是互相不同的种类的基因。

20. 权利要求18所述的载体,其特征在于,上述报告子基因选自:荧光素酶基因、 $\beta$ -半乳糖苷酶基因、一氧化氮合成酶基因、黄嘌呤氧化酶基因、蓝色荧光蛋白质基因、绿色荧光蛋白质基因、红色荧光蛋白质基因及重金属结合蛋白质基因。

21. 测定试剂盒,其具备:

权利要求18~20之任一项所述的自我复制载体,及  
收容上述自我复制载体的容器。

22. 分析装置,其具备:

对于含被检细胞的样品导入权利要求18~20之任一项所述之上述自我复制载体的基因导入部,

通过使复制起始蛋白质表达,使来自上述基因导入部的所述样品中所含的所述自我复制载体自我复制的恒温部,

在来自上述恒温部的所述样品中,检测来源于所述报告子基因的表达而发生的信号的检测部,

基于所述检测的信号,对于所述被检细胞的表观遗传的信息进行分析的分析部。

23. 权利要求18~20之任一项所述的自我复制载体用于制造试剂盒的用途,上述试剂盒用于基于被检细胞的表观遗传的信息而判断上述被检细胞的特性。

24. 权利要求23的用途,其中上述特性是药剂感受性。

25. 权利要求18~20之任一项所述的自我复制载体用于制造试剂盒的用途,上述试剂



盒用于基于被检细胞的表观遗传的信息而确定上述被检细胞是否是癌化的细胞。

26. 权利要求18~20之任一项所述的自我复制载体用于制造试剂盒的用途, 上述试剂盒用于基于被检细胞的表观遗传的信息而确定上述被检细胞是否罹患特定的疾病。

## 方法、自我复制载体、测定试剂盒及分析装置

### 【技术领域】

[0001] 本发明的实施方式涉及得到细胞的表观遗传的信息的方法、判断细胞的特性的方法、判断药剂感受性的或选择药剂或免疫疗法剂的种类的方法、疾病的诊断方法、以及自我复制载体、测定试剂盒及分析装置。

### 【背景技术】

[0002] 在基因的表达与多种因子相关。由那样的因子的存在或不在或者存在的量的变化而对对象的健康状态被左右。另外,如果观察那样的因子的状态,则判断对象中的健康状态变得可能。

[0003] 在关于基因的表达的因子中,有伴随构成基因组的DNA的碱基序列的变化的和不伴随DNA碱基序列的变化的。在伴随DNA的碱基序列的变化的中含多型或变异。另一方面,不伴随DNA碱基序列的变化的有关于表观遗传控制的变化。关于这样的多型及变异、以及表观遗传控制的信息(即,表观遗传的信息)在医疗、农业、渔业等的各种各样的领域中有用。在表观遗传的信息中有DNA的修饰、组蛋白的化学修饰、由非翻译性RNA的控制等。

[0004] 例如,目前报告,得到对特定的碱基序列的甲基化等的修饰的有无或修饰量的变化等的表观遗传的信息对于医学有用。

[0005] 关于DNA的甲基化的信息是有用的表观遗传的信息的1例。DNA 甲基化异常在多阶段的发癌过程中自早期被确认。另外,在初期肝癌中也确认目标基因群集中的甲基化。

[0006] 基因组DNA的甲基化反应是指在细胞内由DNA甲基转移酶在基因组上的胞嘧啶附加甲基。细胞分裂时基因组复制,则复制的基因组在与复制前的基因组同位置的胞嘧啶发生甲基的附加。

[0007] 现在,在检测DNA的甲基化的手段中,例如,有硫酸氢盐处理 PCR法、测序法、利用报告子载体的方法。例如,作为利用报告子载体的方法,报告有测定由时钟基因的DNA甲基化的昼夜节律控制结构,筛选昼夜节律障碍改善剂的方法。所述筛选法是制作染色体上整合在细胞内甲基化的已知的目标核酸序列的稳定表达细胞株,利用与细胞分裂时的基因组复制同时进行的甲基化反应,筛选昼夜节律障碍改善剂的方法。

[0008] 【现有技术文献】

[0009] 【专利文献】

[0010] 专利文献1:特开2009-232761号公报

[0011] 专利文献2:特开2012-60894号公报

[0012] 【非专利文献】

[0013] 非专利文献1:Yoshiura K et al.,Proc Natl Acad Sci U S A.1995, 92,7416-7419。

[0014] 非专利文献2:Chen YL et al.,Biochem Biophys Res Commun. 2012,425(2), 290-296

[0015] 非专利文献3:Hayatsu H et al.,Biochemistry.1970,9,2858-2865。

- [0016] 非专利文献4:Gonzalgo ML et al.,Nucleic Acids Res.1997,25, 2529-2531。
- [0017] 非专利文献5:Herman JG et al.,Proc Natl Acad Sci U S A.1996, 93,9821-9826。
- [0018] 非专利文献6:Clark SJ et al.,Nucleic Acids Res.1994,22, 2990-2997。
- [0019] 非专利文献7:Frommer M et al.,Proc Natl Acad Sci U S A.1992, 89,1827-1831。
- [0020] 非专利文献8:Warnecke PM et al.,Methods.2002,27,101-107。

### 【发明内容】

#### [0021] 【发明要解决的技术课题】

[0022] 实施方式旨在提供简便地得到细胞的表观遗传的信息的方法。

#### [0023] 【解决课题的技术方案】

[0024] 由实施方式提供得到细胞的表观遗传的信息的方法。此方法包括:将报告子核酸构建体导入被检细胞的核内,使报告子核酸构建体自我复制,检测在被检细胞中发生的信号的有无及/或大小,基于得到的结果,得到被检细胞的表观遗传的信息。报告子核酸构建体的特征在于,将被检细胞的基因组上的特定的序列中的修饰的状态抄录到在被检细胞的核内的自我复制中由官能团的取代而对应的其序列上,依赖于抄录的官能团的存在状况发生可检测的信号。

### 【附图说明】

[0025] 【图1】图1是显示自我复制载体的1例的图。

[0026] 【图2】图2是显示自我复制载体的1例的图。

[0027] 【图3】图3是显示自我复制载体的1例的图。

[0028] 【图4】图4是显示自我复制载体的例的图。

[0029] 【图5】图5是显示自我复制载体的例的图。

[0030] 【图6】图6是显示自我复制载体的1例的图。

[0031] 【图7】图7是显示自我复制载体的1例的图。

[0032] 【图8】图8是显示得到细胞的表观遗传的信息的方法的1例的方案。

[0033] 【图9】图9是显示得到细胞的表观遗传的信息的方法的1例的方案。

[0034] 【图10】图10是显示得到细胞的表观遗传的信息的方法的1例的方案。

[0035] 【图11】图11是显示得到细胞的表观遗传的信息的方法的1例的方案。

[0036] 【图12】图12是显示判断被检细胞的特性的方法的1例的图。

[0037] 【图13】图13是显示用于进行得到细胞的表观遗传的信息的方法的分析装置的框图。

[0038] 【图14】图14是显示含甲基化目标序列的自我复制载体的制作工序的1例的图。

[0039] 【图15】图15是显示实验结果的坐标图。

[0040] 【图16】图16是显示含CK19启动子区域和复制起始序列的自我复制载体的图。

[0041] 【图17】图17是显示复制起始蛋白质基因表达载体的图。

[0042] 【图18】图18是显示CK19启动子区域的CCGG序列的位置的图。

- [0043] 【图19】图19是显示实验结果的图。
- [0044] 【图20】图20是显示实验结果的图。
- [0045] 【图21】图21是显示实验结果的图。
- [0046] 【图22】图22是显示实验结果的图。
- [0047] 【图23】图23是含COX2启动子区域和复制起始序列的自我复制载体的图。
- [0048] 【图24】图24是显示COX2启动子区域的CCGG序列的位置的图。
- [0049] 【图25】图25是显示实验结果的图。
- [0050] 【图26】图26是显示实验结果的图。
- [0051] 【图27】图27是显示实验结果的图。
- [0052] 【图28】图28是显示实验结果的图。
- [0053] 【图29】图29是显示实验结果的图。
- [0054] 【图30】图30是显示实验结果的图。
- [0055] 【图31】图31是显示实验结果的图。
- [0056] 【图32】图32是显示荧光素酶测定的结果的图。
- [0057] 【图33】图33是显示含甲基化目标序列的自我复制载体的制作方法的1例的图。
- [0058] 【图34】图34是显示向细胞的2种载体的共转染的模式图。
- [0059] 【图35】图35是显示自我复制载体的1例的图。
- [0060] 【图36】图36是显示实验结果的图。
- [0061] 【实施方式】

[0062] 以下,对于本发明的实施方式详细地进行说明。

[0063] 根据实施方式的得到细胞的表观遗传的信息的方法使用可自我复制的报告子核酸构建体。此报告子核酸构建体在被检细胞的核内自我复制时,将被检细胞的基因组中的特定的序列上的修饰的状态抄录到自身的序列。即,报告子核酸构建体含与要观察的特定的序列相同的序列。再者,由自我复制,基因组的特定的序列上的修饰状态被抄录(拷贝)到与报告子核酸构建体中的其相同的序列上。报告子核酸构建体依赖于抄录到所述报告子核酸构建体的官能团的存在状况,发生可检测的信号。此时、由官能团的存在状况而信号的有无或大小或量不同。由此,报告子核酸构建体报告基因组的特定的序列上的修饰状态。



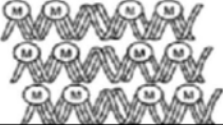
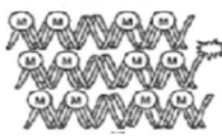

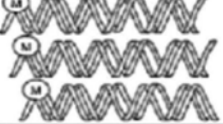




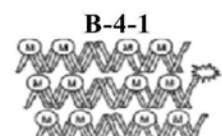

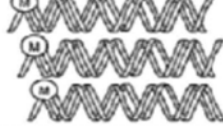
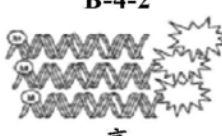
[0064] 报告子核酸构建体的报告功能通过报告子核酸构建体自我复制而初始得到。从而,报告子核酸构建体被导入置于报告子核酸构建体的自我复制开始的条件的被检细胞的核内。然后,在经过对于被导入被检细胞的核内的报告子核酸构建体自我复制必要时间后,在被检细胞内检测报告子核酸构建体发生的信号。信号的检测可对于信号的有无进行,也可对于信号的大小或量进行。由检测得到的结果是被检细胞的表观遗传的信息。其中,“细胞的表观遗传的信息”可为对于构成被检细胞的基因组上的特定的序列的碱基是否结合有官能团、是否有取代等的信息。其中“可检测的信号”可为荧光、发光及呈色等,或者也可作为蛋白质等的物质的呈示。或者其是,例如,可由本领域公知的方法检测的信号即可。另外,所述信号依赖于抄录到报告子核酸构建体的官能团的存在状况,其提示的有无或大小改变。

[0065] “报告子核酸构建体”是指用于报告要试验的细胞的表观遗传的信息的核酸构建体。以下描述对于报告子核酸构建体和基因组的特定序列中的修饰的相关的例。以下的例

是表观遗传的信息是关于基因组的特定的序列中的甲基化的信息时的1例。表1示报告子核酸构建体的报告子活性和甲基化的相关。

[0066] 【表1】

[0067]

报告子核酸构建体的当初的修饰状态	基因组上的特定序列的修饰状态	自我复制时的官能团的取代	可检测的信号的大小	报告子活性
<b>A-1</b>  含多个甲基的报告子核酸构建体	<b>A-2-1</b> 甲基化状态 高 基因组 	<b>A-3-1</b> 无 	<b>A-4-1</b>  低	<b>A-5-1</b> 低
	<b>A-2-2</b> 甲基化状态 低 基因组 	<b>A-3-2</b> 有 (脱甲基化) 	<b>A-4-2</b>  高	<b>A-5-2</b> 高
<b>B-1</b>  含少的甲基的或不 含甲基的报告子核 酸构建体	<b>B-2-1</b> 甲基化状态 高 基因组 	<b>B-3-1</b> 有 (甲基化) 	<b>B-4-1</b>  低 (缓慢地降低)	<b>B-5-1</b> 低
	<b>B-2-2</b> 甲基化状态 低 基因组 	<b>B-3-2</b> 无 	<b>B-4-2</b>  高	<b>B-5-2</b> 高

[0068] 表1所示的相关是报告子核酸构建体的信号由报告子核酸构建体的对应的序列的脱甲基化变大,由甲基化变小的1例。

[0069] 此报告子核酸构建体,例如,可在癌的检测中使用。在癌细胞的情况中,在基因组上的特定的序列中发生脱甲基化。

[0070] 这样的细胞中的表观遗传的信息在其自我复制时,被抄录到报告子核酸构建体中所含的对应的序列。例如,将报告子核酸构建体的对应的序列预先甲基化,则在自我复制时发生脱甲基化。此时、自报告子核酸构建体的可检测的信号变大。例如,以这样的大的信号的检测作为指标,检测癌变得可能。其中,“对应的序列”是指与基因组上的特定的序列、即要检测的修饰所在的序列相同的序列。

[0071] 例如,准备含多个甲基的报告子核酸构建体(A-1)。设基因组上的特定的序列含多个甲基(A-2-1)。此时,当自我复制时,所述构建体中无官能团的取代,或者少(A-3-1)。此时,检测信号小(A-4-1)。相对报告子活性低(A-5-1)。另一方面,设基因组上的特定的序列不含多个甲基(A-2-2)。此时,当自我复制时,在所述构建体中官能团被取代,引起脱甲基化(A-3-2)。此时,检测信号大(A-4-2)。相对报告子活性高(A-5-2)。

[0072] 或者,例如,准备含少的甲基的报告子核酸构建体(B-1)。设基因组上的特定的序列含多个甲基(B-2-1)。此时,当自我复制时,所述构建体中的官能团被取代,发生甲基化

(B-3-2)。此时,检测信号小(B-4-2)。而且,在此时,随时间经过而信号变小。相对报告子活性低(B-5-2)。另一方面,设基因组上的特定的序列不含多个甲基(B-2-2)。此时,当自我复制时,在所述构建体中无官能团的取代,或者少(B-3-2)。此时,检测信号大(B-4-2)。相对报告子活性高(B-5-2)。

[0073] 以上、显示报告子核酸构建体的报告子活性和甲基化的相关,但此相关仅用于例示,对其无限定。

[0074] 【1.报告子核酸构建体】

[0075] 报告子核酸构建体,例如,可为自我复制载体。接下来使用附图说明作为报告子核酸构建体的自我复制载体。再有,对于相同的构成赋予相同的符号,对于说明则省略。另外,可各实施方式的构成组合而在1个实施方式中利用,另外也可多个实施方式在1个分析中同时使用。

[0076] 【第1实施方式】

[0077] 在得到细胞的表观遗传的信息的方法中,在细胞内复制起始蛋白质表达的条件下,使用自我复制载体。自我复制载体也被称为“报告子载体”。

[0078] 使用图1,对于自我复制载体的1例进行说明。自我复制载体1含报告子基因表达单位2和复制起始序列3。复制起始序列3与报告子基因表达单位2存在于相同的核酸上。

[0079] 报告子基因表达单位2含目标核酸序列4、编码报告子蛋白质的报告子基因5和转录终止信号序列6。

[0080] 目标核酸序列4是要得到表观遗传的信息的核酸序列。目标核酸序列4与被检细胞有的特定的核酸序列,对于其序列具有相同性。再者,目标核酸序列4有被修饰碱基。然后,自我复制载体1(例如目标核酸序列4)有依赖于此被修饰碱基中的修饰的程度的启动子活性。自我复制载体1(例如目标核酸序列4)的启动子活性的活化由目标核酸序列4上的被修饰碱基的修饰控制。这样的被修饰碱基成为修饰的目标。目标核酸序列4根据希望预先选择即可。从关于目标核酸序列4得到的表观遗传的信息,得到被检细胞中的特定的核酸序列的表观遗传的信息。

[0081] 关于目标核酸序列4,“具有相同性的序列”是与基因组上的特定的序列相同的序列,或者与基因组上的特定的序列实质上相同的序列即可。基因组上的特定的序列和实质上相同的序列是指,例如,在所述特定的序列中,1个或数个碱基被取代、缺失或付加的序列。另外,“特定的序列”是要得到表观遗传的信息的序列即可。例如,“具有相同性的”、“相同的”及“相同的”是指对于基因组序列是90%以上或95%以上相同的序列即可。“基因组上的序列和相同的序列”可为与在基因组上连续的特定的序列相同的序列,也可为与在基因组上不连续的2个以上的特定的序列各自相同的序列一体化而成的序列。

[0082] 例如,目标核酸序列4可为将与在基因组上的多个位置存在的2个以上的序列各自相同的序列一体化而成的序列,也可含那样的一体化而成的序列。换言之,将在基因组上的多个位置上存在的2个以上的序列的各自作为1单位,则目标核酸序列4,对于这些多个单位的各自,可含将各自相同的序列一体化的序列,也可含那样的一体化的序列。再换言之,目标核酸序列4对应于作为在基因组上各自相同的序列的单位,也可由2个以上的单位构成。在目标核酸序列4中,例如,也可为1个单位与基因组上的特定的启动子序列的部分相同,另1个单位与含可控制所述启动子序列的活化的被修饰碱基的序列相同。

[0083] 例如,含基因组上的特定的启动子序列的部分和可控制此启动子序列的活化的被修饰碱基的序列不是在基因组上必然连续存在。即,它们之间也可含进一步的序列。这样的进一步的序列只要是对启动子序列及其活性无影响的序列,就没必要非得使目标核酸序列4含那样的进一步的序列。另外,只要有依赖于目标核酸序列4上的被修饰碱基和被修饰碱基中的修饰的程度的启动子活性,则目标核酸序列4除与被检细胞的基因组上的序列有相同性的序列之外,也可含任意的进一步的序列。任意的进一步的序列不妨碍有依赖于目标核酸序列4上的被修饰碱基和被修饰碱基中的修饰的程度的启动子活性。例如,任意的进一步的序列可在构成与被检细胞的基因组上的序列有相同性的序列的任何的碱基之间含有,也可邻接与被检细胞的基因组上的序列有相同性的序列而含有。

[0084] 在上述中,关于目标核酸序列4,记载了对于基因组上的序列“具有相同性的序列”,但关于被检细胞,也同样地理解对于目标核酸序列4“具有相同性的序列”即可。

[0085] 自我复制载体1是向被检细胞导入后复制自己的载体。目标核酸序列4是与被检细胞的基因组中所含的特定的核酸序列具有相同性的序列。由这样的构成,利用自我复制载体1的自我复制能力,将被检细胞的基因组上的相同的序列的修饰状态抄录到自我复制载体1中的目标核酸序列4变得可能。

[0086] 目标核酸序列4的例,例如,可举出SOX2基因启动子序列、GFAP 基因启动子序列、CK19基因启动子序列及COX2基因启动子序列等,对其无限定。

[0087] 目标核酸序列4的修饰,例如,是甲基化、烷基化。此时,目标核酸序列4的修饰的状态是指对于构成目标核酸序列4的碱基结合或者未结合作为官能团的甲基及/或烷基。被检细胞的“表观遗传的信息”是指可存在或不存在那样的官能团。或者其也可为何程度的量的官能团存在的信息。或者其也可为那样的修饰的状态的变化、即官能团的取代的有无或程度等的信息。另外,也可为它们之中之任何的组合。目标核酸序列4的修饰,例如,也可为对目标核酸序列4的主链中的胞嘧啶及/或鸟嘌呤的修饰。在目标核酸序列4中存在多个胞嘧啶及/或鸟嘌呤时,也可为关于它们之中之一或1个以上的修饰。例如,目标核酸序列4的修饰可为对目标核酸序列4中的胞嘧啶及/或鸟嘌呤的甲基化。

[0088] 启动子活性是指诱导与启动子序列的下游功能性地连接的基因的转录。目标核酸序列4有启动子活性时,其启动子活性可为由目标核酸序列4的序列全体来源,也可为目标核酸序列4之一部分的序列来源的。另外例如,目标核酸序列4也可含最小启动子。另外,在可显示目标核酸序列4的启动子活性的序列的全体上连续或断续地存在被修饰碱基,在可显示启动子活性的序列以外的目标核酸序列4的序列上存在被修饰碱基,也可为其组合。目标核酸序列4的启动子活化由目标核酸序列4上的被修饰碱基的修饰的状态控制。换言之,目标核酸序列4的启动子活性依赖于目标核酸序列4的序列内的碱基的修饰的程度而活化。

[0089] 目标核酸序列4的启动子活性,例如,在目标核酸序列4的被修饰碱基中的修饰的程度低时活化,或者在修饰存在时活化,或者在修饰不存在之时活化等,也可为这些之任何。对这样的启动子活性的目标核酸序列4的性质是称为“依赖于目标核酸序列4的序列内的碱基的修饰的程度而活化”或“依赖于碱基中的官能团的取代的程度而活化”的性质的例。

[0090] 有依赖于目标核酸序列4的序列内的被修饰碱基中的修饰的程度的启动子活性的序列的例,例如,是被验细胞来源的序列即可,也可为人工地合成的序列。例如,那样的序列

的例也可为神经胶质细胞纤维性酸性蛋白质 (GFAP; Glial fibrillary acidic protein) 基因的5'上游区域(-1651~+32bp) (SEQ ID NO:1) 等。GFAP基因的5'上游区域序列在目标核酸序列的甲基化不存在时活化。将其序列示于表2。或者,也可为细胞角蛋白19 (CK19) 基因的5'上游区域(-617~+61bp) (SEQ ID NO:2) 或环加氧酶2 (COX2) 基因的5'上游区域(-540~+115bp) (SEQ ID NO:3) 等。这些的序列示于表2、表3及表4。

[0091] 【表2】〈SEQ ID NO:1〉小鼠GFAP5'上游区域(-1651bp~+32bp) 序列

gtctgtaagc tgaagacctg gcagtgctga gctggtcacc cccaggacc	50
tcctttgtg cccaacgagt gactcacctt ggcatagaca taatggtcag	100
gggtgggcac gcagcctgct tcccgctgtg ctccaggcct ccttcgatgc	150
ttccgagaa gtctattgag ctgggagctt gtactgcacc cggggctgac	200
atcctggcat cctgggataa aagcagccca cggggctgcc ctgcatat	250
gcctcactgg cggcagagga caaggctcta ttcagcaagt gccctggagt	300
agacaccaga agccaagca tgggcagagg aaggcagggg ttggggggag	350
cagagctgtc tgtgtccag aagccaagg acacagatgg ctaaggcgcc	400
tgggagggac ctgagtggaa gagatagatg ggctgaagt ctaagcagc	450
aacagcctcc tccccgcat tggtaggggt ggggtttgtt tccccgacc	500
tacatatccc tcagaggcct ggtgtgtagg aatttaaagg aggtaaatct	550
cctgagagaa tgaggggtac ccaggaagac ggggtgttac agaaagactc	600
cagcatgcac agccaactca ctcaaaacta ctctgtcagg ggctgccggg	650
ggccaggctc ggggtggggg gtgggggggc aaagagaagc tggaccaggg	700
agaaatggcc cactaggctg gatatgaggc cacagagggg ctgaggaatg	750
aagcctgctg tctaccta ttaggatctg cgtgcatacc ttctgctgtg	800
cactctaaac acacagccag aggcctaagt tgaccctgga gtcacagaga	850
gggtccaac cttagccctc cactcctgaa ctccaggaat gagaagatag	900
agttggagcg attcagggga gaggactctg ttgagaatgg gggccacagg	950
aaactgtaat ataggtgat cccggaggaa gggaaatagg tctcaagtt	1000
cctagcatct cacaggcccc cagagaagga cagagtggg gtggtcctgg	1050
cttacaggct ctaagaactg gaagctgatt accccaccaa gctgtgact	1100
ctctgtctct gctctgtct ctgtgtgtc gcgctcgtc acacttatca	1150
cacaaatglt catgtgtgtg cacatagatg agttgagacc agaggtcaac	1200
ctcaggcact gttgccttg tttctgaga gacatttct ctctggacct	1250
ggaactcgcc aattagttag agccaggaag tctgtgatt ttcactgcc	1300
agcactggag ttacaagta tgcactgtca acccaggcct ttgtattca	1350
ttctgcagct agaactggg tgggtcttca tgcctgacag gcaagcaatt	1400
tatggactaa gctgttctt cggccctctc ttgaccatt taccagaaag	1450
gggttctct gatcaatggc gaagccaggc tgggttccc aagaaagcct	1500
tgactctggg tacagtgacc tcagtgggt gagaggatt ctccccttag	1550
ctgggctggg gccagcttc accccctcag gctattcagt gggggtgctt	1600
ccaggaagtc aggggcagat ttagtccaac cgttctctc ataaaggccc	1650
tgacatcca ggagccagca gaggcagggc agg	1683

[0092] [0093] 【表3】〈SEQ ID NO:2〉人CK19基因的5'上游区域(-617bp~+61bp) 序列

[0094] GCCTGGTCAACATGGTGAAACCCCTGTCTCTACTAAAAATACAAAAA TTAGCTGGGCGTGGTGGCGCG  
TGCCTGTAATCCCAGCTACTCCGG AGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGATGGAGGCTG CAGTGG  
CCCAGATCACACCACTGCACTCCAGTCTGGGCGACAG AGACTCGTCTCAAAAAACAAAAACAAATCACTGG  
GTCAGGGGT GTGTAGGAATATGACCCCAGAGGGACTGTAATTCAGTGGTGTC AAATCTGGGTGATCTTAA  
GGGTGAGGCTCAGAATGTGGCTTCC AGGGACAGGGGTGTCCAGGTATGTCCCTGACGGGGGAAAGGCCA GAACA  
GGGTCTGCAGAGAGAGAGTGGGGGAGTGCGGGTCTGAG CTTCTGCGCGACCGGGGCGGGGCACCTCTGGAGGGC  
AGGGGC CTCTGGTCTCTGGGAGGGGAGGAATTGACCAATGGGGAGAGAG CCCATATTTGCTCTCAGGAGCCTG  
CAAATTCCTCAGGGCTCAGATA TCCGCCCTGACACCATTCCTCCCTTCCCCCTCCACCGGCCGCG GGCATAA



AAGGCGCCAGGTGAGGGCCTCGCCGCTCCTCCCGCA ATCGCAGCTTCTGAGACCAGGGTTGCTCCGTCCGTGCT  
CCGCCTC GCCCTCGAG

[0095] 【表4】〈SEQ ID NO:3〉人COX2基因的5'上游区域(-540bp~ +115bp)序列

[0096] CTTAACCTTACTCGCCCCAGTCTGTCCCGACGTGACTTCCTCGACC CTCTAAAGACGTACAGACCAG  
ACACGGCGGCGGCGGCGGGAGAG GGGATTCCCTGCGCCCCCGACCTCAGGGCCGCTCAGATTCTTG GAGAGGA  
AGCCAAGTGTCTTCTGCCCTCCCCCGGTATCCCATCC AAGGCGATCAGTCCAGAACTGGCTCTCCGAAGCGCTC  
GGGCAAA GACTGCGAAGAAGAAAAGACATCTGGCGGAAACCTGTGCGCCTGG GGCGGTGGAACTCGGGAGGAG  
AGGGAGGGATCAGACAGGAGA GTGGGGACTACCCCCTCTGCTCCCAAATTGGGGCAGCTTCCTGGG TTTCCGAT  
TTTCTCATTTCCGTGGGTAAAAAACCCCTGCCCCACCG GGCTTACGCAATTTTTTTAAGGGGAGAGGAGGAAAA  
ATTTGTGG GGGGTACGAAAAGGCGGAAAGAAACAGTCATTTTCGTACATGGGC TTGGTTTTTCAGTCTTAAAA  
AGGAAGGTTCTCTCGGTTAGCGACC AATTGTCATACGACTTGCACTGAGCGTCAGGAGCACGTCCAGGAA CTCC  
TCAGCAGCGCCTCCTTCAGCTCCACAGCCAGACGCCCTCAG ACAGCAAAGCCTACCCCCGCGCTCGAG

[0097] 报告子基因5功能性地连接在目标核酸序列4的下游。即,目标核酸序列4功能性地连接在报告子基因5的上游。报告子基因5通过目标核酸序列4活化而转录。由此转录,报告子蛋白质表达。报告子蛋白质的检测可由荧光、发光及呈色等的检测进行,或者也可由其公知的蛋白质的检测手段检测。

[0098] 报告子蛋白质,例如,可为荧光素酶、β-半乳糖苷酶、一氧化氮合成酶、黄嘌呤氧化酶、蓝色荧光蛋白质、绿色荧光蛋白质、红色荧光蛋白质及重金属结合蛋白质。这时,报告子基因5可为编码它们的蛋白质的基因。报告子基因5的例是荧光素酶基因、β-半乳糖苷酶基因、一氧化氮合成酶基因、黄嘌呤氧化酶基因、蓝色荧光蛋白质基因、绿色荧光蛋白质基因、红色荧光蛋白质基因及重金属结合蛋白质基因等。

[0099] 转录终止信号序列6是终结其上游的基因的转录的序列即可,例如,可为聚(A)附加信号。

[0100] 聚(A)附加信号是发挥哺乳动物的基因的转录终止的功能即可。作为例,可举出SV40病毒之后期聚(A)附加信号(SEQ ID NO:4)或牛生长激素基因的聚(A)附加信号(SEQ ID NO:5)等。但是,在本实施方式的实施中可使用的聚(A)附加信号,对其无限定,只要不损害作为聚(A)附加信号的功能,也可使用改变基因序列的。对于这些的聚(A)附加信号的例,将其序列示于表5。

[0101] 【表5】〈SEQ ID NO:4〉SV40病毒后期聚(A)

	cagacatgat aagatacatt gatgagtgtg gacaaaccac aactagaatg	50
	cagtgaaaaa aatgctttat ttgtgaaatt tigtatgcta ttgctttatt	100
[0102]	tgtaaccatt ataagctgca ataaacaagt taacaacaac aattgcattc	150
	attttatgtt tcagggtcag ggggaggtgt gggaggttt ttaaagcaag	200
	taaaacctct acaaatgtgg ta	222

[0103] 〈SEQ ID NO:5〉牛生长激素转录终止信号序列

	gittaaacc gctgatcagc ctgactgtg ccttctagt gccagccatc	50
	tggtgtttgc cctcccccg tgccttcctt gacctggaa ggtgccactc	100
[0104]	ccactgtcct ttctaataa aatgaggaaa ttgcatcgca tigtctgagt	150
	agggtgcatt ctattctggg ggggtgggtg gggcaggaca gcaaggggga	200
	ggattgggaa gacaatagca ggcattgct	228

[0105] 关于多个序列的连接,“功能性的”是指各基因在可发挥作为目的的功能的状态下

连接的状态。例如，在编码蛋白质的序列的情况中，是指各自的碱基序列编码的氨基酸序列正确连接。换言之，是指其无氨基酸密码子框错开，在导入所述碱基序列的细胞中合成发挥作为目的的活性的功能的肽。另外在这里，“功能性地连接”的术语也可与“可操作地连接”的术语互换使用。

[0106] 转录终止信号序列6与报告子基因5的下游功能性地连接。

[0107] 自我复制载体1可在被检细胞中表达复制起始蛋白质的条件下，在被检细胞内进行自我复制。复制起始序列3是由复制起始蛋白质的结合而开始自我复制的序列。

[0108] 复制起始序列3，例如，是在哺乳类细胞中由复制起始蛋白质的结合活化而开始自我复制的序列即可，是本领域公知的复制起始序列即可。复制起始序列3，例如，也可为猿猴病毒40 (SEQ ID NO:6)、爱泼斯坦-巴尔病毒或小鼠多瘤病毒来源的复制起始序列，但对这些无限定。将复制起始序列的例的序列示于表6。

[0109] 【表6】〈SEQ ID NO:6〉猿猴病毒40复制起始序列

	<b>gctagcaata aaatatcttt atttcatta catctgtgtg ttggttttt</b>	<b>50</b>
	<b>gtgtgaatcg atagtactaa catacgctct ccatcaaaac aaaacgaaac</b>	<b>100</b>
[0110]	<b>aaaacaaact agcaaaatag gctgtcccca gtgcaagtc aggtgccaga</b>	<b>150</b>
	<b>acatttctcg ctage</b>	<b>165</b>

[0111] 自我复制载体1是可自我复制的载体即可，例如，也可为质粒载体。例如，也可利用本领域公知的质粒载体。例如，此自我复制载体的质粒载体的例公开于特开2013-42721。

[0112] 自我复制载体1的自我复制通过复制起始蛋白质结合复制起始序列3而开始。

[0113] 图2模式地示自我复制载体1的复制起始序列3结合了复制起始蛋白质7的状态。复制起始蛋白质7与复制起始序列3结合，再结合 DNA复制中涉及的多种蛋白质。由结合的多种蛋白质，在被检细胞内进行自我复制载体1的复制。

[0114] 复制起始蛋白质7“在细胞内复制起始蛋白质表达的条件”下表达。“在细胞内复制起始蛋白质表达的条件”是指使自我复制载体1的自我复制开始的复制起始蛋白质在导入有所述职体的细胞内表达的条件下即可。即，在要导入自我复制载体1的被检细胞内存在复制起始蛋白质单位即可。“复制起始蛋白质单位”是指用于表达复制起始蛋白质的单位，也可称为“复制起始蛋白质表达单位”。

[0115] 复制起始蛋白质单位是，在自我复制载体1中含可使自我复制开始的复制起始蛋白质表达的序列即可。或者复制起始蛋白质单位是，在自我复制载体1中也可含可表达使自我复制开始的复制起始蛋白质的序列。例如，复制起始蛋白质单位含控制编码复制起始蛋白质的序列的表达的启动子、与其下游功能性地连接的编码复制起始蛋白质的序列和与其下游功能性地连接的转录终止信号序列。或者例如，复制起始蛋白质单位也可由控制编码复制起始蛋白质的序列的表达的启动子、与其下游功能性地连接的编码复制起始蛋白质的序列和与其下游功能性地连接的转录终止信号序列构成。

[0116] 与复制起始序列3结合，自我复制载体1中编码使自我复制开始的复制起始蛋白质的序列的例如下。即其，例如，也可为猿猴病毒40 的大T抗原基因、爱泼斯坦-巴尔病毒的EBNA-1基因、小鼠多瘤病毒的大T抗原基因等。

[0117] 优选的大T抗原基因的例各自示于表7、表8及表9。即，其可为由SEQ ID NO:7、SEQ ID NO:8或SEQ ID NO:9各自所示的核酸。或者，可为由SEQ ID NO:7、SEQ ID NO:8或SEQ ID NO:9的序列中的1或数个碱基缺失、取代或附加的序列所示，并且有DNA复制起始功能的核

酸。其中,DNA复制起始功能是指编码可开始复制的蛋白质。

[0118] 【表7】〈SEQ ID NO:7〉猿猴病毒40的大T抗原基因(野生型)

```

atggataaag ttttaaacag agaggaatct ttgcagctaa tggaccttct      50
aggctctgaa aggagtgcct gggggaatat tcctctgatg agaaaggcat      100
alttaaaaaa atgcaaggag ttctatcctg ataaaggagg agatgaagaa      150
aaaatgaaga aaatgaatag tctgtacaag aaaatggaag atggagtaaa      200
atatgctcat caacctgact ttggaggctt ctgggatgca actgagattc      250
caacctatgg aactgatgaa tgggagcagt ggtggaatgc cttlaatgag      300
gaaaacctgt ttgtctcaga agaaatgcca tctagtgatg atgaggctac      350
tgctgactct caacattcta ctctccaaa aaagaagaga aaggtagaag      400
acccaagga ctttcttca gaattgctaa gtttttgag tcatgctgtg      450
tttagtaata gaactcttg tgcctttgt atttacacca caaaggaaaa      500
agctgcactg ctatacaga aaattatgga aaaatattct gtaaccttta      550
taagtaggca taacagttat aatcataaca tactgtttt tcttactcca      600
cacaggcata gagtgtctgc tattaataac tatgtctaaa aattgtgtac      650
cttagcttt ttaatttgta aaggggttaa taaggaatat ttgatgtata      700
gtgccttgac tagagalcca tttctgtta ttgaggaaag ttgccagggt      750
gggttaaagg agcatgattt taatccagaa gaagcagagg aaactaaaca      800
agtgtcctgg aagcttgtaa cagagtatgc aatggaaaca aaatgtgatg      850
atgtgttgtt attgcttggg atgtacttgg aatttcagta cagttttgaa      900
atgtgtttaa aatgtattaa aaaagaacag cccagccact ataagtacca      950
tgaaaagcat tatgcaaatg ctgtatatt tgctgacagc aaaaaccaa      1000
aaacatatg ccaacaggct gtgatactg tttagctaa aaagcgggtt      1050
gatagcctac aattaactag agaacaatg ttaacaaaca gattlaatga      1100
tcttttggat aggatggata taatgtttg ttctacaggc tctgtgaca      1150
tagaagaatg gatggctgga gtgtctggc tacactgtt gtgcccacaa      1200
atggattcag tgggtatga ctttttaaaa tgcatgggtt acaacattcc      1250
taaaaaaaga tactggctgt ttaaaggacc aattgatagt ggtaaaacta      1300
cattagcagc tgctttgctt gaattatgtg gggggaaagc tttaaattgt      1350
aatttgcctt tggacaggct gaactttgag ctaggagtag ctattgacca      1400
gttttagta gttttgagg atgtaaaggg cactggaggg gagtccagag      1450
atttgccttc aggtcaggga attaataacc tggacaattt aagggtattt      1500
ttgatggca gtgttaagg aaacttagaa aagaacacc taaataaaag      1550
aactcaaata ttccccctg gaatagtcac catgaatgag tacagtgtgc      1600
ctaaaacact gcaggccaga ttgtaaaac aaatagattt taggccaga      1650
gattatttaa agcatgtcct ggaacgcagt gagttttgt tagaaaagag      1700
aataattcaa agtggcattg cttgtctct tatgttaatt tggtagagac      1750
ctgtggctga gttgtctca agtatcaga gcagaattgt ggagtggaaa      1800
gagagattgg acaagaggt tagttgtca ggtatcaaaa aaatgaagtt      1850
taatgtggct atgggaattg gagttttaga ttggctaaga aacagtgatg      1900
atgatgatga agacagccag gaaaatgctg ataaaaatga agatgggtgg      1950
gagaagaaca tggaagactc agggcatgaa acaggcattg atcacagtc      2000
ccaaggctca ttacaggccc ctacgtcctc acagtctgtt catgatcata      2050
atcagccata ccacattgt agaggtttta ctgtcttaa aaaacctccc      2100
acacctcccc ctgaacctga acctgaacaa taa                        2133

```

[0120] 【表8】〈SEQ ID NO:8〉猿猴病毒40的大T抗原基因(变异型1)

	atggataaag ttttaaacag agaggaatct ttgcagctaa tggaccttct	50
	aggcttgaa aggagtgcct gggggaatat tcctctgatg agaaaggcat	100
	atttaaaaa atgcaaggag ttcatcctg ataaaggagg agatgaagaa	150
	aaaatgaaga aaatgaatac tctgtacaag aaaatggaag atggagtaaa	200
	atatgctcat caacctgact ttggaggctt ctgggatgca actgagattc	250
	caacctatgg aactgatgaa tgggagcagt ggtggaatgc cttaatgag	300
	gaaaacctgt ttgctcaga agaaatgcc aatagtgatg atgaggctac	350
	tgctgactct caacattcta ctctccaaa aaagaagaga aaggtagaag	400
	acccaagga ctttcttca gaattgctaa gtitttgag tcatgcttg	450
	tttagtaata gaactcttc ttgcttgct atttacacca caaaggaaaa	500
	agctgcactg ctatacaaga aaattatgga aaaataitct gtaacctta	550
	taagtaggca taacagtta atcataaca tactgtttt tctactcca	600
	cacaggcata gagtgtctgc tattaataac tatgctcaaa aattgtgtac	650
	cttagcttt ttaattgta aaggggttaa taaggaatat ttgatgtata	700
	gtgccttgac tagagaacca tttctgtta ttgaggaaag ttgccaggt	750
	gggttaaagg agcatgatt taatccagaa gaagcagagg aaactaaaca	800
	agtgtcctgg aagcttgtaa cagagtatgc aatggaaaca aaatgtgatg	850
	atgtgtgtt attgctggg atgtacttg aatticagta cagtttgaa	900
	atgtgttaa aatgtattaa aaaagaacag cccagccact ataagtaacca	950
	tgaaaagcat tatgcaaatg ctgctatatt tgctgacagc aaaaaccaa	1000
	aaacatag ccaacaggct gtgatactg ttttagctaa aaagcgggtt	1050
[0121]	gatagcctac aattaactag agaacaaatg ttaacaaaca gattaatga	1100
	tctttggat aggatggata taatgttgg tctacaggc tctgctgaca	1150
	tagaagaatg gatggctgga gtgcttggc tacactgtt gtgcccaca	1200
	atggagtcag tgggtatga cttttaaaa tgcattggtg acaacattcc	1250
	taaaaaaaga tactggctgt taaaggacc aattgatagt ggtaaaacta	1300
	cattagcagc tgcttgcct gaattatgtg gggggaagc ttaaatgtt	1350
	aatttgcct tggacaggct gaacttgag ctaggagtag ctattgacca	1400
	gttttagta gtittgagg atgtaaagg cactggagg gagtccagag	1450
	atttgcctc aggtcaggga attaataacc tggacaatt aagggattat	1500
	ttggatggca gtgttaagg aaacttagaa aagaaacacc taaataaaag	1550
	aactcaaata ttccccctg gaatagtcac catgaatgag tacagtgtgc	1600
	ctaaaacact gcaggccaga ttgtaaaac aaatagattt taggccaga	1650
	gattatttaa agcattgcct ggaacgcagt gatttttgt tagaaaagag	1700
	aataaitcaa agtggcattg cttgcttct tatgtlaatt tggtagagac	1750
	ctgtggctga gtttctcaa agtattcaga gcagaattgt ggagtggaaa	1800
	gagagattgg acaaagagtt tagttgtca gtgtatcaaa aaatgaagtt	1850
	taatgtggct atgggaattg gagttttaga ttggctaaga aacagtgatg	1900
	atgatgatga agacagccag gaaaatgctg ataaaaatga agatgggtggg	1950
	gagaagaaca tggagactc agggcatgaa acaggcattg attcacagtc	2000
	ccaaggcica ttacaggccc ctacgtctc acagtctgt catgatcata	2050
	atcagccata ccacattgt agaggltta cttgctttaa aaaacctccc	2100
	acacctccc ctgaacctga acctgaacaa laa	2133

[0122] 【表9】〈SEQ ID NO:9〉猿猴病毒40的大T抗原基因(变异型2)

	atggataaag ttttaaacag agaggaatct ttgcagctaa tggaccttct	50
	aggtcttgaa aggagtgcct gggggaalat tccctgatg agaaaggcat	100
	attiaaaaaa atgcaaggag ttcatccig ataaaggagg agatgaagaa	150
	aaaatgaaga aaatgaatac tctgtacaag aaaatggaag atggagtaaa	200
	atatgctcat caacctgact ttggaggctt ctgggatgca actgagattc	250
	caacctatgg aactgatgaa tgggagcagt ggtggaatgc cttaaatgag	300
	gaaaacctgt ttgtctcaga agaaatgcca tctagtgatg atgaggctac	350
	tgctgactct caacattcta ctctccaaa aaagaagaga aaggtagaag	400
	acccaagga ctttcttca gaattgctaa gtttttgag tcatgctgtg	450
	tttagtaata gaactcttgc ttgctttgct atttacacca caaaggaaaa	500
	agctgcactg ctatacaaga aaattatgga aaaatattct gtaaccttta	550
	taagtaggca taacagttat aatcataaca tactgtttt tcttactcca	600
	cacaggcata gagtgcctgc tattaataac tatgctcaaa aattgtgtac	650
	cttagcttt ttaattgta aaggggttaa taaggaatat ttgatgtata	700
	gtgccttgac tagagatcca ttctcgtta ttgaggaaag ttgccagggt	750
	gggttaaagg agcatgattt taatccagaa gaagcagagg aaactaaaca	800
	agtgctctgg aagcttgtaa cagagtatgc aatggaaaca aaatgtgatg	850
	atgtgtgtt atgcttggg atgtacttg aatttcagta cagtittgaa	900
	atgtgttaa aatgtattaa aaaagaacag ccagccact ataagtacca	950
	tgaagcat tatgcaaatg ctgtatatt tgctgacagc aaaaaccaa	1000
[0123]	aaaccatag ccaacaggct gtgatactg tttagctaa aaagcgggtt	1050
	gatagcctac aattaactag agaacaatg ttaacaaca gattaatga	1100
	tctttggat aggatggata taatgttgg ttctacaggc tctgctgaca	1150
	tagaagaatg gatggctgga gttgctggc tacactgtt gtgcccaaa	1200
	atggagtcag tgggtgatga cttttaaaa tgcattgtgt acaacattcc	1250
	taaaaaaga tactggctgt taaaggacc aattgatagt ggtaaaacta	1300
	cattagcagc tgccttgcct gaattatgtg ggggaaagc ttlaaatgtt	1350
	aatttgcctt tggacaggct gaacttgag ctaggagtag ctattgacca	1400
	gttttagta gttttgagg atgtaaagg cactggaggg gattccagag	1450
	aattgcctc aggtcaggga attaataacc tggacaattt aagggtatt	1500
	ttggatggca gtgttaagg aaacttagaa aagaacacc taaataaaag	1550
	aactcaaat tttcccttg gaatagtcac catgaatgag tacagtgtgc	1600
	ctaaaacact gcaggccaga ttgtaaaac aaatagattt taggcccaca	1650
	gattatttaa agcatgcct ggaacgcagt gagttttgt tagaaaagag	1700
	aataattcaa agtggcattg ctttgcctt tatgttaatt tggtagagac	1750
	ctgtggctga gtttgcctaa agtattcaga gcagaattgt ggagtggaaa	1800
	gagagattgg acaagagatt tagttgtca gtgtatcaa aaatgaagtt	1850
	taatgtggct atgggaattg gagttttaga ttggctaaga aacagtgatg	1900
	atgatgatga agacagccag gaaaatgctg ataaaaatga agatgggtggg	1950
	gagaagaaca tgaagactc agggcatgaa acaggcattg attcacagtc	2000
	ccaaggctca ttacaggccc ctacgtcctc acagtctgt catgatcata	2050
	atcagccata ccacattgt agaggttta ctgtcttaa aaaacctccc	2100
	acacctcccc ctgaacctga acctgaaaca taa	2133

[0124] 复制起始蛋白质单位可在自我复制载体1的核酸上存在,也可在与自我复制载体1的核酸不同的核酸上存在。在与自我复制载体1不同的核酸上存在复制起始蛋白质单位时,复制起始蛋白质单位可在被检细胞的染色体上存在,也可在被检细胞中所含的其他核酸上存在。

[0125] 在复制起始蛋白质单位在与自我复制载体1不同的核酸上时,复制起始蛋白质单位有如以下一样的构成。即,其含组成型表达的启动子、与其下游功能性地连接的编码复制起始蛋白质的序列和与其下游功能性地连接的转录终止信号序列,或者由它们的序列组成。由这样的构成,编码复制起始蛋白质的序列被转录、翻译而产生复制起始蛋白质即可。

[0126] 复制起始蛋白质单位含于自我复制载体1中时,复制起始序列3 也可含于自我复

制载体1。自我复制载体1中的报告子基因表达单位2 和复制起始蛋白质单位的表达可由1个启动子控制,其表达也可由独立的2个相同或互相不同的启动子控制。

[0127] 由互相不同的启动子控制表达时,例如,复制起始蛋白质单位在自我复制载体1的核酸上,与报告子基因表达单位2独立地存在。然后复制起始蛋白质单位具备组成型表达的启动子和与其下游功能性地连接的编码复制起始蛋白质的序列即可。另外,还可转录终止信号序列功能性地连接到编码复制起始蛋白质的序列的下游。

[0128] 在自我复制载体1中,在报告子基因表达单位2和复制起始蛋白质单位由1个启动子控制表达时,例如,可为以下一样的构成。那样的自我复制载体1的例,例如,可由与报告子基因表达单位的下游功能性地连接的IRES (internal ribosomal entry site) 序列 (SEQ ID NO: 10)、与IRES序列的下游功能性地连接的编码复制起始蛋白质的序列和与其下游功能性地连接的转录终止信号序列构成。将IRES序列的序列示于表10。

[0129] 【表10】〈SEQ ID NO:10〉脑心肌炎病毒内部核糖体进入序列

	cctctagcgg gatcaattcc gccccccccc cctaacgtta ctggccgaag	50
	ccgcttgga taaggccgggt ggcgttgt ctatatgta tttccacca	100
	tattgcgctc tttggcaat gtgagggccc ggaaacctgg cctgtcttc	150
	ttgacgagca ttctagggg tcttccct ctcgcaaag gaatgcaagg	200
	tctgtgaat gtcgtgaagg aagcagttcc tctggaagct tctgaagac	250
	aaacaacgtc tgiagcgacc cttgcaggc agcgaaccc cccacctggc	300
[0130]	gacagggtgcc tctgcggcca aaagccacgt gtataagata cacctgcaa	350
	ggcggcaciaa cccagtgcc acgtgtgag ttggatagtt gtggaagag	400
	tcaaatggct ctcctcaagc gtattcaaca aggggtgaa ggatgccag	450
	aaggtacccc attgatggg atctgatctg gggcctcggg gcacatgctt	500
	tacgtgtgt tagtcgaggt taaaaaacgt ctaggccccc cgaaccacgg	550
	ggacgtggtt ttccttgaa aaacacgata ataccatgat tgaacaaga	599

[0131] 这样的复制起始蛋白质单位含编码复制起始蛋白质的序列、与其上游功能性地连接的组成型表达的启动子、与其下游功能性地连接的编码复制起始蛋白质的序列和与其下游功能性地连接的转录终止信号序列,或者由它们的序列构成。编码复制起始蛋白质的序列被转录、翻译而产生复制起始蛋白质。即便如此,自我复制载体1在相同的载体内含复制起始序列。产生的复制起始蛋白质与载体内的复制起始序列结合。对于这样的自我复制载体,作为质粒载体的1例,在第2实施方式中进行说明。

[0132] 在复制起始蛋白质单位存在于被检细胞的基因组上时,如例如以下一样生成复制起始蛋白质。将复制起始蛋白质基因表达载体使用转染法或者病毒载体导入被检细胞,培养而使复制起始蛋白质基因表达载体整合进细胞的染色体。转染法是由生物化学方法或物理化学方法进行即可。作为生物化学方法,可举出使用阳离子脂质的脂转染法、使用磁性粒子的磁转染法或使用氯化钙的方法等。作为物理化学方法,例如,可举出电穿孔法等。不仅可单独使用上述的生物化学方法或物理化学方法,也可互相组合。例如,可将作为生物化学方法的脂转染法和磁转染法组合使用,也可将作为生物化学方法的脂转染法和作为物理化学方法的电穿孔法组合使用。

[0133] 通过复制起始蛋白质基因表达载体稳定地整合进染色体,在被检细胞内复制起始蛋白质基因被稳定地转录、翻译而生成复制起始蛋白质。稳定地表达复制起始蛋白质的细胞株,例如,有HEK293T细胞或COS细胞等。

[0134] 当复制起始蛋白质单位在与自我复制载体1的核酸不同、也与被检细胞的染色体



不同的核酸链上存在时,复制起始蛋白质单位也可含于自我复制载体1以外的第2载体中而存在。例如,那样的第2载体是在被检细胞内表达复制起始蛋白质的载体即可,也可利用本领域公知之任何的载体。这样的第2载体也称为“复制起始蛋白质基因表达载体”。

[0135] 如此,例如,复制起始蛋白质单位也可含于自我复制载体中。另外例如,被检细胞含与自我复制载体及被检细胞的基因组不同的进一步的序列,在那样的进一步的序列上也可含复制起始蛋白质单位。那样的进一步的序列可在被检细胞的染色体上含有,也可含于第2载体等的进一步的核酸链。

[0136] 编码复制起始蛋白质的基因和复制起始序列的组合是可开始复制的组合即可。那样的组合的例如下:

[0137] 编码复制起始蛋白质的基因是猿猴病毒40的大T抗原基因,上述复制起始序列是从猿猴病毒40的复制起始序列的组合;

[0138] 编码复制起始蛋白质的基因是爱泼斯坦-巴尔病毒的EBNA-1基因,复制起始序列是从爱泼斯坦-巴尔病毒的复制起始序列的组合;

[0139] 编码复制起始蛋白质的基因是小鼠多瘤病毒的大T抗原基因,复制起始序列是从小鼠多瘤病毒的复制起始序列的组合。优选如这些一样的组合,但对这些无限定。

[0140] 被检细胞,例如,可为细胞组及/或单一细胞。被检细胞可为在血液、尿、便、精液、唾液、活组织检查得到的组织、口腔内粘膜、体腔液及喀痰等中所含的细胞,或者也可培养细胞等。

[0141] 向被检细胞的自我复制载体1的导入可在体外进行,也可在体内进行。

[0142] **【第2实施方式】**

[0143] 接下来,参照图3对于质粒载体的1例进行说明。

[0144] 作为质粒载体的自我复制载体31,如图3所示,具备目标核酸序列4、与其下游功能性地连接的报告子基因5、与其下游功能性地连接的IRES32、与其下游功能性地连接的编码复制起始蛋白质的序列33、与其下游功能性地连接的转录终止信号序列34和与其下游功能性地连接的复制起始序列3。在此例中,报告子基因表达单位2的构成要素是目标核酸序列4、报告子基因5和转录终止信号序列34。另外,复制起始蛋白质单位的构成要素是目标核酸序列4、IRES32、编码复制起始蛋白质的序列33和转录终止信号序列34。

[0145] “IRES”是“内部核糖体进入序列(internal ribosomal entry site)”的略称,也被称为“内部核糖体结合序列”。IRES也可为例如脑心肌炎病毒来源的序列。由IRES32的存在,由依赖于目标核酸序列4的序列内的碱基的修饰的程度而活化的目标核酸序列4,报告子基因5的序列被转录,其后编码复制起始蛋白质的序列33被转录。编码复制起始蛋白质的序列33的序列被阅读后,由转录终止信号序列34的存在,转录停止。从编码复制起始蛋白质的序列33转录、翻译而产生的复制起始蛋白质与复制起始序列3结合,开始自我复制载体31的复制。

[0146] 在含被导入被检细胞的目标核酸序列4的自我复制载体31中,目标核酸序列4的被修饰碱基在载体导入时可无修饰,也可有修饰。

[0147] 这样的自我复制载体31中所含的目标核酸序列4的被修饰碱基,例如未被修饰时,向被检细胞导入后,如以下一样作用。首先,在被检细胞中存在的作为目的的核酸的修饰状态高时,由于目标核酸序列4的被修饰碱基被高度修饰,目标核酸序列4未活化,报告子基

因和复制起始蛋白质基因不表达。从而,自我复制载体31的自我复制停止。对此,在被检细胞中存在的作为目的的核酸的修饰状态低,或者无修饰时,由于目标核酸序列4的被修饰碱基的修饰状态维持低,目标核酸序列4被活化。在目标核酸序列4处于活化状态时,由IRES32连接的报告子基因5和编码复制起始蛋白质的序列33被转录为双顺反子的mRNA,翻译。IRES32通过整合进2个基因之间,此2个基因作为单一双顺反子性mRNA转录。核糖体不仅从双顺反子性mRNA翻译报告子基因5,还与IRES的位置结合,也翻译编码复制起始蛋白质的序列。转录终止信号序列34是编码用于终结报告子基因5和编码复制起始蛋白质的序列33的转录的信号的碱基序列。自编码复制起始蛋白质的序列33转录,由被检细胞的成分翻译的复制起始蛋白质识别复制起始序列3,与其结合。再者,被检细胞中所含的DNA复制中涉及的多种蛋白质结合。由结合的多种蛋白质,在被检细胞中进行自我复制载体31的自我复制。

【0148】 目标核酸序列4的被修饰碱基被修饰时,被导入被检细胞后,如以下一样发挥作用。首先,当被检细胞中存在的作为目的的核酸的修饰状态高时,在目标核酸序列4的被修饰碱基中,由于维持当初的修饰的程度而目标核酸序列4未活化。因此,报告子基因和复制起始蛋白质基因不表达。从而,自我复制载体31的自我复制停止。与此相比,当被检细胞中存在的作为目的的核酸的修饰状态低,或者无修饰时,由于目标核酸序列4的被修饰碱基中的修饰状态降低、或者成无修饰,从而目标核酸序列4被活化。目标核酸序列4处于活化状态时,由 IRES32连接的报告子基因5和编码复制起始蛋白质的序列33被转录为双顺反子的mRNA,翻译。IRES32通过整合进2个基因之间,此2个基因作为单一双顺反子性mRNA转录。核糖体不仅从双顺反子性 mRNA翻译报告子基因5,还与IRES的位置结合,对于编码复制起始蛋白质的序列也进行翻译。转录终止信号序列34是编码用于终结报告子基因5和编码复制起始蛋白质的序列33的转录的信号的碱基序列。自编码复制起始蛋白质的序列33转录,由被检细胞的成分翻译的复制起始蛋白质识别复制起始序列3,对于其结合。再者,被检细胞中所含的DNA复制中涉及的多种蛋白质结合。由结合的多种蛋白质,在被检细胞中进行自我复制载体31的自我复制。

【0149】 当目标核酸序列4的被修饰碱基中的修饰的程度低时,报告子基因5表达。再者,通过存在于相比其更下游的序列被阅读,自我复制载体31的自我复制开始。由自我复制载体31的自我复制发生的复制生成物,在序列上有与自我复制载体31相同的序列。但是被检细胞中存在的作为目的的核酸的修饰状态被目标核酸序列4的被修饰碱基中的修饰状态反映。结果,目标核酸序列4的被修饰碱基中的修饰的程度变高,则目标核酸序列4未活化。为此,报告子基因5不表达,再者,存在于相比其更下游的序列不被阅读。从而,自我复制载体31的自我复制停止。

### 【0150】 【第3实施方式】

【0151】 对于进一步的质粒载体的1例,参照图4(a)、(b)及(c)进行说明。

【0152】 首先,使用图4(a)对于基本的构成进行说明。作为质粒载体的自我复制载体41含报告子基因表达单位2、与其连接的复制起始蛋白质单位42和与其连接的复制起始序列3。报告子基因表达单位2含目标核酸序列4、与其下游功能性地连接的报告子基因5和与其下游功能性地连接的转录终止信号序列6。复制起始蛋白质单位42存在于与报告子基因表达单位2相同的核酸上,含组成型表达的启动子44、与其下游功能性地连接的编码复制起始蛋白质的序列45和与其下游功能性地连接的转录终止信号序列46。复制起始序列3与复制起



始蛋白质单位42的转录终止信号序列46的下游连接。

[0153] 被导入被检细胞的自我复制载体41的目标核酸序列4的被修饰碱基在所述载体41的导入时可无修饰,也可有修饰。有依赖于目标核酸序列4的被修饰碱基的修饰的程度的启动子活性的序列的例如上所述。

[0154] 这样的自我复制载体41在被导入被检细胞后,如以下一样发挥作用。由自我复制载体41的自我复制发生的复制生成物,关于序列有与自我复制载体41相同的序列。但是自我复制载体的目标核酸序列4 的被修饰碱基中的修饰状态被被检细胞中存在的作为目的的核酸的修饰状态反映。

[0155] 例如,当目标核酸序列4的被修饰碱基被修饰,并且,被检细胞中存在的作为目的的核酸的修饰状态高时,目标核酸序列4未活化。为此,报告子基因5不表达。与此相比,在当初的目标核酸序列4的被修饰碱基被修饰,被检细胞中存在的作为目的的核酸的修饰状态低,或者无修饰时,则如下。自我复制载体41的目标核酸序列4的被修饰碱基的修饰状态被被检细胞中存在的作为目的的核酸的修饰状态反映。结果,目标核酸序列4的被修饰碱基中的修饰的程度变低,目标核酸序列4被活化。由此,报告子基因5表达。

[0156] 在目标核酸序列4的被修饰碱基未被修饰时,当被检细胞中存在的作为目的的核酸的修饰状态高时,目标核酸序列4的被修饰碱基被修饰。因此,目标核酸序列4未活化。从而,报告子基因不表达。与此相比,当被检细胞中存在的作为目的的核酸的修饰状态低,或者无修饰时,由于维持当初的修饰的程度,目标核酸序列4被活化。由此,报告子基因5表达。

[0157] 编码复制起始蛋白质的序列45由组成型表达的启动子44而转录被控制。组成型表达的启动子44,由于常处于活化的状态,被重复活化。编码复制起始蛋白质的序列45的转录由转录终止信号序列46而转录停止。复制起始蛋白质,由组成型表达的启动子44的活性,无关于有依赖于目标核酸序列4的序列内的碱基的修饰的程度的启动子活性的序列的修饰的状态而重复生成。

[0158] 组成型表达的启动子44的例可为巨细胞病毒启动子、劳斯肉瘤病毒启动子、猿猴病毒初期启动子及猿猴病毒后期启动子等。

[0159] 图4(a)所示的自我复制载体41是,复制起始蛋白质单位42连接到报告子基因表达单位2的下游。然后,复制起始序列3连接到其下游。但是,自我复制载体41的构成不限于图4所示的实施方式。例如,也可为复制起始序列3连接到报告子基因表达单位2的下游,其下游再连接复制起始蛋白质单位42。或者,也可为报告子基因表达单位2连接到复制起始蛋白质单位42的下游,其下游再连接复制起始序列3。或者,也可为复制起始序列3连接到复制起始蛋白质单位42的下游,其下游再连接报告子基因表达单位2。

[0160] 另外,报告子基因表达单位2及复制起始蛋白质单位42的各自的方向也不限于图4(a)所示的实施方式。在图4(a)所示的自我复制载体41中,由箭头示序列被阅读的方向(图4(b))。此时,报告子基因表达单位2自左侧向右侧方向被阅读。同样地,复制起始蛋白质表达单位42也自左侧向右侧被阅读。再者,例如,报告子基因表达单位2和复制起始蛋白质单位42设计成互相相反的方向而构成的1 例示于图4(c)。在图4(c)的例的情况中,报告子基因表达单位2 自左侧向右侧被阅读。与此相比,复制起始蛋白质表达单位42自右侧向左侧被阅读。但是,报告子基因表达单位2及复制起始蛋白质单位 42的各自的方向不限于这些。复制起始蛋白质单位42,例如,也可配置在报告子基因表达单位2的上游或下游。然后,复制起

始蛋白质单位42也可连接为,在其任何的情况下,复制蛋白质基因以与报告子基因转录的方向逆向地转录。或者,复制起始蛋白质单位42也可连接为,在其任何的情况下,复制蛋白质基因以与报告子基因转录的方向相同的方向转录。

**[0161] 【第4实施方式】**

[0162] 对于进一步的质粒载体的例,使用图5进行说明。图5(a)显示导入被检细胞50的自我复制载体51。自我复制载体51具备报告子基因表达单位2和复制起始序列3。报告子基因表达单位2具备目标核酸序列4、与其下游功能性地连接的报告子基因5和与其下游功能性地连接的转录终止信号序列6。在被检细胞50中,复制起始蛋白质单位52与自我复制载体51独立地存在。复制起始蛋白质单位52具备组成型表达的启动子53、与其下游功能性地连接的编码复制起始蛋白质的序列54和与其下游功能性地连接的转录终止信号序列55。

[0163] 图5(b)显示复制起始蛋白质单位52在进一步的载体上存在的例。这时,自我复制载体51和含复制起始蛋白质单位52的载体56,作为各自第1载体51和第2载体56,一同被导入到1个被检细胞50。这些的导入方法可为如上述的生物化学方法及物理化学方法。被检细胞中被导入第2载体56,则编码复制起始蛋白质的序列54表达,转录、翻译而生成复制起始蛋白质。复制起始蛋白质识别复制起始序列3,对于其结合。再者,被检细胞中所含的DNA复制中涉及的多种蛋白质结合。由结合的多种蛋白质,在被检细胞内进行自我复制载体51的复制。

[0164] 其中在第2载体56也可被称为复制起始蛋白质基因表达载体。此载体含用于使复制起始蛋白质表达的复制起始蛋白质单位即可。另外,复制起始蛋白质基因表达载体只要在要导入自我复制载体的细胞中,可复制起始蛋白质表达载体即可。那样的载体,例如,可为质粒载体或病毒载体等的本领域公知的载体,或者也可为其自身有自我复制的功能的载体。

[0165] 自我复制载体和复制起始蛋白质基因表达载体的向被检细胞的导入也可为同时。或者,先于自我复制载体的导入,也可复制起始蛋白质基因表达载体被导入被检细胞。或者先于复制起始蛋白质基因表达载体的导入,也可自我复制载体被导入到被检细胞。

[0166] 如此通过联用自我复制载体和复制起始蛋白质基因表达载体,例如,即使在启动子活性低时或被检细胞增殖性低的细胞时,也可进行报告子蛋白质的有效的表达或自我复制载体的有效的自我复制。

**[0167] 【第5实施方式】**

[0168] 在上述的第1实施方式~第4实施方式中,在1个自我复制载体中含报告子基因表达单位。但是,1个自我复制载体中所含的报告子基因表达单位的数不限于1个。换言之,自我复制载体也可含至少1个报告子基因表达单位和复制起始序列。其中,作为进一步的实施方式,作为含多个报告子基因表达单位的1个例显示含2个报告子基因表达单位的例。

[0169] 使用图6,对于自我复制载体的进一步的1例进行说明。自我复制载体61含作为第1报告子基因表达单位的报告子基因表达单位2、作为第2报告子基因表达单位的报告子基因表达单位62和复制起始序列3。复制起始序列3存在于与报告子基因表达单位2及62相同的核酸上。

[0170] 报告子基因表达单位62含目标核酸序列64、编码报告子蛋白质的报告子基因65和转录终止信号序列66。

[0171] 目标核酸序列4和64可为相同的核酸序列,也可有互相不同的序列。报告子基因表达单位中所含的报告子基因5和65可为相同的报告子基因,也可有互相不同的报告子基因。

[0172] 转录终止信号序列6和66可为相同的序列,也可有互相不同的序列。

[0173] 图6所示的自我复制载体61是报告子基因表达单位62连接到报告子基因表达单位2的下游。然后,其下游连接复制起始序列3。但是,自我复制载体61的构成不限于图6所示的实施方式。例如,也可有复制起始序列3连接到报告子基因表达单位2的下游,其下游再连接报告子基因表达单位62。或者,也可有报告子基因表达单位2连接到报告子基因表达单位62的下游,其下游再连接复制起始序列3。或者,也可有复制起始序列3连接到报告子基因表达单位62的下游,其下游再连接报告子基因表达单位2。另外,报告子基因表达单位和复制起始序列的方向也不限于图6所示的实施方式。例如,也可连接为报告子基因表达单位62的报告子基因以报告子基因表达单位2和报告子基因表达单位62互相相反的方向,即与在报告子基因表达单位2中报告子基因转录的方向逆向转录。或者,也可连接为报告子基因表达单位62的报告子基因以在报告子基因表达单位2中与报告子基因转录的方向相同的方向转录。再者,复制起始序列3转录的方向可为与报告子基因表达单位2及/或报告子基因表达单位62相同的方向,也可有逆向的方法。

[0174] 图6中的报告子基因表达单位2、报告子基因表达单位62、及同时使用的复制起始蛋白质单位52也可全部由1个启动子控制表达。或者,它们也可由独立的3个启动子控制表达。或者,也可这3个构成之中2个由1个启动子控制,其余之一由独立的进一步的启动子控制。

[0175] 优选的1例如下。设计自我复制载体61为目标核酸序列4和64含相同的序列。此时,使目标核酸序列4和64中的官能团的状态互相不同。例如,在报告子基因表达单位2中,配置核酸的官能团未被取代的目标核酸序列4,在报告子基因表达单位62中,配置核酸的官能团被取代的目标核酸序列64。再者,此时,设计为报告子基因表达单位2和62各自有各自编码互相不同的报告子蛋白质的报告子基因。

[0176] 在这样的1例中,报告子基因表达单位2和62的差异是由目标核酸序列4中所含的核酸的官能团的取代的有无或程度、报告子基因的种类。即,在自我复制载体61中,目标核酸序列4中所含的核酸是其官能团未被取代,但目标核酸序列64中所含的核酸的官能团预先被取代。使用这样的自我复制载体,可自报告子基因表达单位2和62得到各自不同的2种检测信号。由这样的构成可识别假阳性及假阴性的信号。由此,可以更高精度进行测定。

[0177] 对报告子基因表达单位62中所含的各要素的详细是只要转用上述的对报告子基因表达单位2中所含的各要素的说明即可。

[0178] 另外,这样的第2报告子基因表达单位也可在第1实施方式~第4实施方式中含有。

[0179] **【第6实施方式】**

[0180] 对于进一步的质粒载体的例,使用图7进行说明。图7显示一同使用第4实施方式中记载的复制起始蛋白质基因表达载体和第5实施方式的例。这些的详细如在第4实施方式及第5实施方式中说明。

[0181] **【2.得到被检细胞的表现遗传的遗传信息的方法】**

[0182] 作为进一步的实施方式,提供得到细胞的表现遗传的遗传信息的方法。对于得到

被检细胞的表现遗传的遗传信息的方法,使用图8进行说明。

[0183] 图8的方法使用可自我复制的报告子核酸构建体。如上所述,此报告子核酸构建体,在被检细胞的核内的自我复制中,将其细胞的基因组上的特定的序列中的修饰的状态由官能团的取代抄录到其自身的序列。然后,依赖于抄录到报告子核酸构建体的官能团的存在状况发生可检测的信号。所述方法是,首先,将这样的报告子核酸构建体导入到被检细胞(S81)。接下来使报告子核酸构建体自我复制(S82)。检测在被检细胞中发生的信号的有无及/或大小(S83)。基于得到的结果,得到被检细胞的表现遗传的信息(S84)。由此简便地得到细胞的表现遗传的信息。

[0184] 对于进一步的实施方式,使用图9进行说明。图9显示作为报告子核酸构建体使用自我复制载体的例。首先,向被检细胞内、例如,核内导入自我复制载体(S91)。接下来,通过在复制起始蛋白质表达的条件下培养被检细胞而使自我复制载体自我复制(S92)。在经过指定的时间后,检测自报告子基因表达的报告子蛋白质(S93)。此检测可为报告子蛋白质的有无及/或表达量。另外,“指定的时间”,例如,可为对于被检细胞分裂必要的时间。自我复制载体的自我复制在被检细胞分裂时发生。从而,在指定的时间被检细胞至少分裂1次即可。从此检测结果得到被检细胞的表现遗传的信息(S94)。此方法也可通过另外还预先与对应的表核对,判断被检细胞是否处于何状态。那样的表,例如,可为细胞的状态或病态和对应报告子蛋白质的有无及/或表达量的表。或者,也可从得到的检测结果判断被检细胞的表现遗传的信息具体是何表现遗传的信息。这样的方法也可在体外或体内之任何中进行。

[0185] 对于进一步的例,使用图10进行说明。在此方法中,首先,将复制起始蛋白质基因表达载体导入被检细胞的核内(S101)。接下来,向此被检细胞的核内导入自我复制载体(S102)。温育此被检细胞。由此使自我复制载体自我复制(S103)。在经过指定的时间后,检测自报告子基因表达的报告子蛋白质(S104)。从此检测结果得到被检细胞的表现遗传的信息(S105)。

[0186] 对于进一步的例,使用图11进行说明。在此方法中,在该细胞的核内导入自我复制载体的对象中,通过维持在使复制起始蛋白质表达的条件下,使导入的上述自我复制载体自我复制(S111)。接下来,对于对象,检测自自我复制载体发生的报告子蛋白质及/或测定其表达量(S112)。基于结果,得到对在所述对象中所含的细胞中的基因组上的特定的序列中的修饰的信息(S113)。

[0187] 报告子蛋白质,例如,可为荧光素酶、 $\beta$ -半乳糖苷酶、一氧化氮合成酶、黄嘌呤氧化酶、蓝色荧光蛋白质、绿色荧光蛋白质、红色荧光蛋白质及重金属结合蛋白质。根据选择的报告子蛋白质的种类,可选择适宜的检测及/或测定方法。

[0188] 接下来,对于报告子蛋白质的检测详细地进行说明。作为如上所述报告子基因,可选择发光蛋白质基因或荧光蛋白质基因、显色蛋白质基因、活性氧生成酶基因、重金属结合蛋白质基因等。

[0189] 发光蛋白质基因是编码催化发光反应的酶蛋白质的基因。作为发光蛋白质基因,例如,可举出荧光素酶基因。作为荧光素酶基因的翻译产物的荧光素酶使用作为底物之一的荧光素进行发光反应。

[0190] 荧光蛋白质基因是编码荧光蛋白质的基因。作为荧光蛋白质基因,例如,可举出蓝色荧光蛋白质基因、绿色荧光蛋白质基因、红色荧光蛋白质基因。

[0191] 显色蛋白质基因是编码催化显色反应的酶蛋白质的基因。作为显色蛋白质基因，例如，可举出 $\beta$ -半乳糖苷酶基因。作为 $\beta$ -半乳糖苷酶基因的翻译产物的 $\beta$ -半乳糖苷酶，使用5-溴-4-氯-3-吲哚基- $\beta$ -D-吡喃半乳糖(X-Gal)或o-硝基苯基- $\beta$ -D-吡喃半乳糖(ONPG)等的底物进行显色反应。

[0192] 发光蛋白质基因，荧光蛋白质基因或显色蛋白质基因的翻译产物可由光学检测装置检测。具体而言，在报告子基因是荧光素酶基因的情况下，从被检细胞提取荧光素酶蛋白质而加底物之后，使用光度计测定溶液的发光强度即可。在报告子基因是荧光蛋白质基因的情况下，例如，向被检细胞照射特定波长的光，也可在被检细胞内由荧光测定装置测定自荧光蛋白质发生的荧光的强度。或者，在报告子基因是荧光蛋白质基因的情况下，向含自被检细胞提取的荧光蛋白质的提取液照射特定波长的光，也可由荧光测定装置测定提取液中的自荧光蛋白质发生的荧光的强度。在报告子基因是 $\beta$ -半乳糖苷酶基因的情况下，提取 $\beta$ -半乳糖苷酶蛋白质而加底物之后，使用吸光度测定装置测定溶液的吸光度即可。

[0193] 活性氧生成酶基因是编码生成活性氧的酶的基因。作为活性氧生成酶基因，例如，可举出一氧化氮合成酶基因或黄嘌呤氧化酶基因等。作为一氧化氮合成酶基因的翻译产物的一氧化氮合成酶和作为黄嘌呤氧化酶基因的翻译产物的黄嘌呤氧化酶生成活性氧。

[0194] 活性氧可用电子自旋共振装置(ESR装置)等检测。用由ESR装置的方法可直接测定活性氧的发生量，也可利用特异性的自旋捕获剂测定。利用自旋捕获剂时，首先，将活性氧生成酶用自旋捕获剂捕获后，测定自捕获的活性氧生成酶发生的活性氧的发生量。

[0195] 重金属结合蛋白质基因是编码与重金属特异性地结合的蛋白质的基因。与重金属结合的蛋白质的生成量由磁共振成像装置、核医学诊断装置或X射线计算机断层摄影装置进行即可。重金属结合蛋白质的发生量，例如，如以下一样进行。首先，将可测定的重金属预先加到被检细胞的培养液。接下来，根据需要清洗被检细胞后，从被检细胞提取重金属结合蛋白质。接下来，将提取的重金属结合蛋白质用对应于向培养液加的重金属的图像诊断装置摄像，测定重金属结合蛋白质的量。再有，可使重金属结合蛋白质在被检细胞之内部表达，也可使在被检细胞之外表面表达。

[0196] 重金属结合蛋白质基因的例是编码金属化合物结合肽的碱基序列即可。例如，金属化合物结合蛋白质是对于特定的金属化合物特异性地结合的肽、寡肽、多肽及/或蛋白质即可。

[0197] 例如，编码结合金属化合物的肽的序列可利用已知与希望得到的金属化合物结合的抗体基因或单链抗体基因的碱基序列，基于那样的碱基序列设计即可。对于那样的碱基序列，例如，也可通过在维持与金属化合物的结合的范围内的一或几个碱基的取代、缺失及附加等的修饰及/或改变、根据适用的对象的修饰及/或改变而设计。

[0198] 编码单链抗体肽的基因可自结合金属化合物的抗体的氨基酸序列设计。

[0199] 例如，在MRI摄像中，为了造影效果高，优选使用钆化合物。钆化合物，例如，是钆、钆离子、钆络合物、它们的盐及它们的衍生物、含这些之任何的衍生物、或者由钆化合物的类似物等组成的金属化合物即可。

[0200] 另外，也可提示到细胞外金属化合物结合能的报告子基因。那样的报告子基因可为编码向细胞外提示的金属化合物结合肽的碱基序列。那样的碱基序列，例如，构成为在细胞中被转录，翻译，产生金属化合物结合肽，产生的金属化合物结合肽向细胞膜移动，向

细胞外提示金属化合物结合能力即可。这样的报告子基因的例,例如,公开于特开2012-200245。如在该文献中公开,那样的报告子基因含编码金属化合物结合肽的碱基序列、编码将上述金属化合物结合肽输送到细胞膜的信号肽的碱基序列和编码将由上述信号肽输送到上述细胞膜的上述金属化合物结合肽固定化到上述细胞膜的锚肽的碱基序列即可。

**[0201] 【3.判断被检细胞的特性的方法】**

[0202] 也可利用得到被检细胞的表现遗传的遗传信息的方法判断被检细胞的特性。判断被检细胞的特性的方法包括基于被检细胞的表现遗传的遗传信息,确定被检细胞的特性。

[0203] 被检细胞的特性,例如,可为药剂感受性、药剂适应性等。

[0204] 另外,确定被检细胞的特性也可确定被检细胞是否处于特定的状况。确定被检细胞是否处于特定的状况也可在体外或体内进行。

[0205] 例如,判断细胞的特性的方法也可为,用于对于所述被检细胞来源的对象,判断药剂感受性的方法、或者用于对于所述被检细胞来源的对象,选择在外科手术后要使用的药剂或免疫疗法剂的种类的方法、或者用于判断患者的预后良好与否的方法。另外例如,判断细胞的特性的方法也可为得到用于进行这些的判断的信息的方法。

[0206] 另外,其中“特定的状况”可为特定的健康状态、例如罹患特定的疾病的状态等。特定的疾病,例如,可为癌、精神疾病、生活习惯病、神经疾病、自身免疫疾病、及循环器疾病等。也可为用于在再生医疗中确认自干细胞分化诱导的器官的甲基化状态是正常与否的方法。

[0207] 对于使用第2实施方式判断基于表观基因组的遗传信息的细胞的特性的方法的1例,使用图12进行说明。首先,准备具备希望得到的目标核酸序列4的自我复制载体31(a)。接下来,将自我复制载体31导入有核71的活的被检细胞50(b)。其后,在被检细胞50中自我复制载体31根据条件维持可自我复制的状态(b)。例如,在复制起始蛋白质存在的状态下被检细胞可分裂增殖的培养条件下,经任意期间培养被检细胞。

[0208] 在核71的自我复制载体31中,被检细胞50的核71中的作为目的的核酸序列的修饰状态被目标核酸序列4的被修饰碱基的修饰的状态反映。结果,例如,当作为目的的核酸序列的甲基化的程度高时,自我复制载体31的自我复制的程度变低(c)。另一方面,例如,当作为目的的核酸序列的甲基化的程度低时,自我复制载体31的自我复制的程度变高(d)。

[0209] 接下来,测定报告子蛋白质的量(e)。结果,当报告子蛋白质量高时,被检细胞50由于目标核酸序列4的相同的序列中的被修饰碱基未被甲基化,或者无修饰,判断为是甲基化强度低的细胞(f)。或者,结果,当报告子蛋白质量低时,被检细胞50判断为是目标核酸序列4的相同的序列中的被修饰碱基被甲基化的细胞(f)。

[0210] 实施方式的自我复制载体,在自我复制之间,根据被检细胞中存在的作为目的的核酸序列的修饰状态,其修饰状态被目标核酸序列4的被修饰碱基的修饰状态反映。目标核酸序列4的修饰状态的判断可通过检测或定量报告子蛋白质来进行。

[0211] 作为判断基于表观基因组的遗传信息的细胞的特性的方法的进一步的1例,说明使用图6的自我复制载体61的方法。首先,准备具备希望得到的目标核酸序列4及64的自我复制载体61。接下来,将自我复制载体61导入被检细胞50。其后,在被检细胞50中自我复制载体61被维持在根据条件可自我复制的状态。例如,在复制起始蛋白质存在的状态下被检细胞可分裂增殖的培养条件下,经任意期间培养被检细胞。

[0212] 当自我复制载体61在被检细胞内复制时,被检细胞50的基因组上的核酸序列的取代状态被目标核酸序列4及64的被修饰碱基的修饰的状态反映。结果,例如,当作为目的的核酸序列的甲基化的程度高时,从含未甲基化的目标核酸序列4的报告子基因表达单位2的报告子蛋白质(报告子基因5)的表达量缓慢地降低。对此,从含甲基化的目标核酸序列64的报告子基因表达单位62的报告子蛋白质(报告子基因65)的表达量不变化。

[0213] 接下来,测定报告子蛋白质的量,取从报告子基因5及65的报告子蛋白质质量的比。当此比小时,被检细胞50被判断为,与目标核酸序列4或64相同的序列中的被修饰碱基被甲基化,或者被修饰。因此,再者,判断为是甲基化强度高的细胞。

[0214] 实施方式的自我复制载体61,在自我复制之间,根据被检细胞中存在的作为目的的核酸序列的修饰状态,其修饰状态被目标核酸序列4及64的被修饰碱基的修饰状态反映。目标核酸序列的修饰状态的判断可通过检测或定量从报告子基因5的报告子蛋白质和从报告子基因65的报告子蛋白质后,取两者的比来实现。

[0215] 另外,也可提供通过还利用如上述一样得到的细胞的表现遗传的遗传信息,判断细胞所来源的对象罹患特定的疾病的可能性高,或者低的方法。另外,也可基于如上述一样得到的细胞的表现遗传的遗传信息,收集用于进行最终的诊断的信息。

[0216] **【4. 诊断方法】**

[0217] 也可基于如上述一样得到的细胞的表现遗传的遗传信息,诊断对象中的特定的疾病。一般而言,那样的诊断会最终包括由医师或兽医师等的专家的判断。

[0218] 其中“对象”主要可为含人的哺乳类、家畜、宠物及产业用动物等的个体,也可为自对象得到的器官、器官、组织、细胞组及/或单一细胞。

[0219] 诊断方法是对于作为检测对象的个体、器官、器官、组织、细胞组及/或单一细胞等导入自我复制载体。接下来检测或定量报告子蛋白质即可。

[0220] 在诊断方法中,在含发光蛋白质基因作为报告子基因的自我复制载体的情况中,例如,也可进行如以下一样判断的方法。对于作为检测对象的个体、器官、器官、组织、细胞组或单一细胞等,由生物化学方法或物理化学方法导入自我复制载体。在与所述导入同时或所述导入之前后,依赖于所述被修饰碱基的修饰的状态检测作为荧光素酶基因的翻译产物的荧光素酶。从检测对象提取荧光素酶蛋白质而加作为底物之一的荧光素之后,使用光度计测定溶液的发光强度即可。即,由此检测方法,可将荧光素酶的翻译量作为指标特异性地检测处于特定的状况的细胞。这样的诊断方法是侵袭性低的方法。

[0221] 其中,荧光素酶的检测可在单一时间点、多个时间点或继续进行,或者也可经时进行。另外,诊断方法不限于利用荧光素酶的检测,也可使用发生其他可检测的信号的报告子核酸构建体进行。

[0222] 根据进一步的实施方式,接下来记使用自个体采集的器官、器官、组织、细胞组或单一细胞判断检测对象罹患特定的疾病的与否的方法。

[0223] 这样的方法,例如,也可包括:

[0224] (1) 在表达复制起始蛋白质的条件下,对于含细胞的检测对象,提取含发生可检测的信号基因作为报告子基因的自我复制载体,

[0225] (2) 在作为检测对象的细胞或检测对象中所含的细胞可分裂增殖的培养条件下,经任意期间培养检测对象,

- [0226] (3) 检测自上述检测对象发生的所述信号，
- [0227] (4) 基于上述检测的信号，判断检测对象罹患特定的疾病与否。
- [0228] 另外，这样的方法，例如，也可包括：
- [0229] (1) 在表达复制起始蛋白质的条件下，对于含细胞的检测对象，提取含荧光素酶基因作为报告子基因的自我复制载体，
- [0230] (2) 在作为检测对象的细胞或检测对象中所含的细胞可分裂增殖的培养条件下，经任意期间培养检测对象，
- [0231] (3) 从上述检测对象提取荧光素酶蛋白质，
- [0232] (4) 使与作为上述荧光素酶蛋白质的底物的荧光素接触，用光度计检测发光，
- [0233] (5) 基于关于检测的发光的信息，判断检测对象罹患特定的疾病与否。
- [0234] 在上述 (3) 的荧光素酶蛋白质的提取前，也可进行细胞的清洗等。由这样的实施方式，根据细胞的状况检测翻译的荧光素酶蛋白质变得可能。
- [0235] 关于基于荧光素酶的检测的对象的判断可基于荧光素酶的检测的有无、荧光素酶的检测值相比预先设定的阈值更大或更小，检测值的增减等的信息进行。基于那样的信息，例如，判断检测对象罹患以特定的条件作为指标的疾病与否，疾病的重症度是多少等即可。
- [0236] 在诊断方法中，在含向细胞外提示金属化合物结合能的报告子基因作为报告子基因的自我复制载体的情况中，例如，可进行如以下一样判断的方法。对于作为检测对象的个体、器官、器官、组织、细胞组及/或单一细胞等由生物化学方法或物理化学方法导入自我复制载体。在与所述导入同时及/或所述导入之前后，依赖于所述目标核酸序列的修饰的状态，使和与在细胞外提示的报告子蛋白质形成结合对的金属化合物接触。与金属化合物的接触前后及/或接触同时，检测与报告子蛋白质特异性地结合的金属化合物。即，由此检测方法可以金属化合物的存在作为指标，特异性地检测处于特定的状况的细胞。
- [0237] 金属化合物的检测可在单一时间点、多个时间点及/或继续进行，及/或也可经时进行。
- [0238] 另外，金属化合物的检测方法可根据金属化合物中所含的金属原子的种类，使用本领域公知之任何方法进行。例如，可为应用利用金属原子的化学特性及/或物理的本领域公知的用于化学、物理学、物理化学及/或生物化学地检测金属原子的任何方法。
- [0239] 再者，例如，也可从可用例如MRI、PET、SPECT、CT等的图像诊断装置测定用于与报告子蛋白质特异性地使结合对形成的金属化合物的种类选择。由此，可与上述的图像诊断装置组合使用所述自我复制载体。此时，更高灵敏度并且更特异性地检测细胞变得可能。
- [0240] 根据进一步的实施方式，判断检测对象罹患特定的疾病与否的方法包括：
- [0241] (1) 对于作为检测对象的细胞或含细胞的检测对象，提取含向细胞外提示金属化合物结合能的报告子基因作为报告子基因的自我复制载体，
- [0242] (2) 向上述细胞外提示金属化合物结合肽，
- [0243] (3) 对于上述金属化合物结合肽，使可与上述金属化合物结合肽形成结合对的金属接触，
- [0244] (4) 检测与上述金属化合物结合肽结合的金属，
- [0245] (5) 基于所述检测的结果，判断检测对象罹患特定的疾病与否。
- [0246] 在与上述 (3) 的金属的接触后及/或与接触同时，也可进行用于除去未结合的剩余



的金属的操作、例如,清洗、濯洗,稀释及/或环流除去等。

[0247] 另外,由这样的实施方式,根据细胞的状况,由向细胞外提示的金属化合物结合能而在细胞表面结合希望得到的金属变得可能。利用向细胞外提示的金属化合物结合能,通过对于在其中结合及/或蓄积金属化合物,可检测处于特定的条件的细胞。

[0248] 基于金属的检测结果,判断检测对象罹患特定的疾病与否的方法,例如如下。即,其基于金属的检测的有无、金属的检测值相比预先设定的阈值更大或更小,检测值的增减等的信息进行即可。基于它们的信息,再者,判断检测对象罹患以特定的条件作为指标的疾病与否,疾病的重症度是多少等即可。

#### [0249] 【5.测定试剂盒】

[0250] 作为进一步的实施方式,提供测定试剂盒。测定试剂盒含至少1种报告子核酸构建体、例如,根据实施方式的至少1种自我复制载体。测定试剂盒可含1种自我复制载体,也可组合2种以上的自我复制载体含有。另外测定试剂盒也可再含负载自我复制载体的本领域公知的载体及/或收容自身载体复制载体的容器、复制起始蛋白质基因表达载体。再者,测定试剂盒也可含对于进行得到细胞的表观遗传的信息的方法或细胞特性的判断方法必要的任何试剂。

#### [0251] 【6.组合物】

[0252] 另外,报告子核酸构建体也可作为组合物提供。那样的组合物,例如,含根据实施方式的至少1种自我复制载体。组合物的1例可含1种自我复制载体,也可组合2种以上的自我复制载体含有。另外组合物,例如,也可含负载自我复制载体的本领域公知的载体、复制起始蛋白质基因表达载体、赋形剂、稳定剂及/或其他有希望得到的效果的成分等。所述组合物也可含于容器而提供。

#### [0253] 【7.装置】

[0254] 作为进一步的实施方式,提供用于进行如上述一样的方法、例如得到表观遗传的信息的方法、以及使用其的细胞特性的判断方法等的分析装置。

[0255] 对于分析装置的1例,使用图13进行说明。分析装置含基因导入部、与其邻接的恒温部、与其邻接的检测部和对于这些电连接的分析部。在基因导入部中,进行向被检细胞的核内的自我复制载体的导入。在恒温部中进行在被检细胞的核内的自我复制载体的自我复制。在检测部中进行在细胞内发生的报告子蛋白质的检测。在检测部得到的信息被送到分析部。分析部具备处理器、显示部及输入部。在分析部中,由处理器对在检测部得到的信息进行数据处理,根据需要表示在显示部。即,在检测部得到的结果被送到分析部的处理器,根据预先确定的顺序进行分析、计算及加工,表示于显示部。

[0256] 基因导入部、恒温部及检测部之内部是其内部由连络窗相通至可进行连续处理。在被检细胞从1个部过渡到下一部后,连络窗由遮蔽板闭合。

[0257] 由分析装置进行得到表观遗传的信息的方法的顺序如下。试验者将放入被检细胞和自我复制载体的容器放入基因导入部。其后自输入部输入用于开始分析的指示。自输入部的指示被送到处理器,由处理器的指示,在基因导入部中,进行用于可使自我复制载体被导入被检细胞的处理。在经过指定的时间后,结束导入处理,将收容到容器的被检细胞送到恒温部。在恒温部中,经指定的时间,在指定的条件下维持被检细胞。由此,进行自我复制载体的自我复制。经过指定的时间后,收容到容器的被检细胞从恒温部被送到检测部。在检测

部中,进行在被检细胞内表达的报告子蛋白质的检测。检测的信息被送到分析部,由处理器进行数据处理。由此得到的信息示于显示部。向基因导入部、恒温部及检测部的容器的移动由带式输送机、可动式臂及/或可动式托盘等的移动手段进行。这些的全部的移动全部在控制器的控制之下,根据预先设定的程序的指示进行。基因导入部具备对于由选择的方法的基因导入必要的构成。恒温部具备对于满足对于自我复制载体的自我复制必要的条件必要的构成。检测部具备对于根据检测的信号,检测报告子蛋白质必要的构成。

[0258] 这样的分析装置对于简便地得到细胞的表现遗传的信息有用。

### 【实施例】

[0259] 【实施例1:质粒型自我复制载体的制作】

[0260] 如图14显示概要,制作含甲基化的目标核酸序列的质粒型自我复制载体。作为目标核酸序列,选择GFAP基因启动子序列。通过进行模板核酸的扩增和由限制性内切酶的切断,作为目标核酸序列制备 GFAP基因启动子序列。得到的GFAP基因启动子序列是含被修饰碱基,并且有启动子活性的核酸序列。进行由限制性内切酶的切断,接下来,将此目标核酸序列由甲基化酶SssI甲基化。通过进行将得到的甲基化目标核酸序列由连接酶的连接,整合进pM53-R550K载体。pM53-R550K载体是从上游至下游以此顺序具备报告子基因、IRES、编码复制起始蛋白质的序列、转录终止信号序列和复制起始序列的载体。将得到的载体作为含甲基化的目标核酸序列的质粒型自我复制载体。接下来记各工序的详细。

[0261] 【(1)作为目标核酸序列的GFAP基因启动子序列的制作】

[0262] 将小鼠基因组DNA作为模板使用进行PCR。在小鼠基因组DNA 中扩增的序列是表2所示的SEQ ID NO:1。

[0263] 另外,PCR中使用的引物的碱基序列是如以下一样的SEQ ID NO: 11和12;

[0264] 正向引物:

[0265] 5'-CGACGCGTGTCTGTAAGCTGAAGACCTGGC-3' (SEQ ID NO:11)

[0266] 反向引物:

[0267] 5'-AAAAGTACTCCTGCCCTGCCTCTGCTGGCTC-3' (SEQ ID NO:12)。

[0268] 使用这些正向引物和反向引物,以小鼠基因组DNA作为模板进行PCR,得到SEQ ID NO:1所示的GFAP基因启动子序列。得到的GFAP基因启动子序列是,目标核酸序列含被修饰碱基,有依赖于其修饰的程度的启动子活性。此GFAP基因启动子序列是小鼠的GFAP 基因的5'上游区域(-1651bp~+32bp)。接下来,进行GFAP基因启动子序列的向PGV-B2 (TOYO B-Net)载体的克隆。

[0269] 将PGV-B2 (TOYO B-Net)载体用限制性内切酶Sma I切断。使其脱磷酸化。对于此载体,将磷酸化的SEQ ID NO:1所示的多核苷酸用T4连接酶连接。由此,制作PGV-B2-GFAPp载体。

[0270] 将得到的PGV-B2-GFAPp载体,在接下来继续的工序中作为用于得到GFAP基因启动子序列的材料使用的同时,作为 PGV-B2-GFAPp载体,为了比较也在后述的检测试验中使用。另外,通过由后述的工序将PGV-B2-GFAPp载体中所含的GFAP基因启动子序列甲基化,也制作甲基化PGV-B2-GFAPp载体。这也为比较而在后述的检测试验中使用。甲基化或非甲基化的PGV-B2-GFAPp载体均是不自我复制的载体。

[0271] 为了得到GFAP基因启动子序列,将PGV-B2-GFAPp载体用限制性内切酶Mlu I和限制性内切酶Xho I切断。由此,制作附加了限制性内切酶Mlu I和限制性内切酶Xho I识别序列的GFAP基因启动子序列。将此序列作为目标核酸序列而在继续的工序中使用。

[0272] **【(2) 被修饰碱基的甲基化】**

[0273] 将GFAP基因启动子序列甲基化,将其中所含的被修饰碱基甲基化。

[0274] 对于上述(1)中制作的GFAP基因启动子序列核酸的溶液40μL,添加16μL的10×NE Buffer2、8μL的甲基化酶SssI (New • England • Biolabs • 日本株式会社)、8μL的S-腺苷甲硫氨酸 (SAM) (New • England • Biolabs • 日本株式会社)、及88μL的ddH<sub>2</sub>O。由此制作反应液。使其于37℃进行6小时酶反应。甲基化反应后,使用热块于65℃、温育20分钟。由此使Sss I失活。再者,将其使用 QIA快速PCR纯化试剂盒(株式会社Qiagen)纯化。由此,得到纯化的GFAP基因启动子序列。得到的序列中的被甲基化碱基被甲基化。

[0275] 另一方面,对于上述反应液,除了未添加Sss I以外,由与上述同样的方法,将被甲基化碱基未被甲基化的纯化的GFAP基因启动子序列制作为非甲基化的对照序列。

[0276] 另外,对于上述(1)中得到的PGV-B2-GFAPp载体中所含的 GFAP基因启动子序列的甲基化也与上述同样地进行。由此得到甲基化PGV-B2-GFAPp载体。

[0277] **【(3) 质粒型自我复制载体pM53-R550K和作为纯化的目标核酸序列的GFAP基因启动子序列的功能性的连接】**

[0278] 将自我复制型载体pM53-R550K载体用限制性内切酶Mlu I和限制性内切酶Xho I切断。对此,将上述(2)中得到的纯化的GFAP 基因启动子序列由T4连接酶连接。由此,得到pM53-R550K-GFAPp 载体。此pM53-R550K-GFAPp载体含目标核酸序列。在其中,目标核酸序列被甲基化。

[0279] 由同样的方法,对于上述(2)中得到的非甲基化的对照序列也整合进所述载体。由此作为对照载体得到非甲基化的 pM53-R550K-GFAPp载体。

[0280] **【实施例2:目标核酸序列的甲基化及脱甲基化的检测】**

[0281] 进行用于依赖于pM53-R550K-GFAPp载体中所含的目标核酸序列的碱基的修饰的状态,确认启动子活化的程度变化的试验。

[0282] **【(1) 试验中使用的载体】**

[0283] 作为实施例,使用由上述的方法得到的甲基化的质粒型自我复制载体pM53-R550K-GFAPp载体和非甲基化的质粒型自我复制载体 pM53-R550K-GFAPp载体。这些是可自我复制的质粒载体。

[0284] 作为比较例,使用由上述的方法得到的甲基化的PGV-B2-GFAPp 载体和非甲基化的PGV-B2-GFAPp载体。这些是无法自我复制的质粒载体。

[0285] 作为内部标准用的载体,使用表达β-半乳糖苷酶的载体 pcDNA4/V5-His/lacZ载体(Life Technologies公司)。

[0286] **【(2) 载体的向神经胶质瘤C6细胞的导入】**

[0287] **【(a) 向细胞的载体的导入】**

[0288] 将pcDNA4/V5-His/lacZ载体、甲基化或非甲基化的 PGV-B2-GFAPp载体及甲基化或非甲基化的pM53-R550K-GFAPp载体各自导入神经胶质瘤C6细胞。此导入用脂转染法进行。在这些载体的各导入中,作为导入试剂使用脂质体2000(Life Technologies公司)。脂

转染的操作根据由试剂的製造方的手册。其概略如下。在微管中混合悬浮于25 $\mu$ L的Opti-MEM的0.45 $\mu$ L的阳离子脂质(脂质体 2000)和含任何载体的25 $\mu$ L的Opti-MEM。由此形成脂质体/载体的复合物。其中,25 $\mu$ L的Opti-MEM中所含的各载体的量如下。对于 pM53-R550K-GFAPp或PGV-B2-GFAPp,在25 $\mu$ L的Opti-MEM中含0.066 $\mu$ g的DNA。对于pcDNA4/V5-His/lacZ,25 $\mu$ L的Opti-MEM 中含0.033 $\mu$ g的DNA。

[0289] 【(b) 接种】

[0290] 对于各载体形成脂质体/载体复合物后,将各自的复合物对于培养板(96孔板)每1孔各添加50 $\mu$ L。向其中还每1孔各接种100 $\mu$ L悬浮于RPMI1640培养基(含有10%非动物化FCS的)的神经胶质瘤C6 细胞溶液( $5 \times 10^5$  cell/mL)。

[0291] 【(c) 培养】

[0292] 其后,将板稳定搅拌1分钟,将反应液和细胞溶液混合。其后,在37 $^{\circ}$ C、5%CO<sub>2</sub>气氛下培养24小时、在37 $^{\circ}$ C、5%CO<sub>2</sub>气氛下进行附着培养。

[0293] 【(2) 由5-氮杂-2-脱氧胞苷的被修饰碱基的脱甲基化】

[0294] 向神经胶质瘤C6细胞各自导入各载体后,对于各自的载体中所含的目标核酸序列进行脱甲基化反应。脱甲基化反应使用作为甲基转移酶的抑制剂的5-氮杂-2-脱氧胞苷(5-氮杂-dC)(船越株式会社)进行。

[0295] 如以下一样制备5-氮杂-dC试剂。向PBS中添加5-氮杂-dC至成 500 $\mu$ M的浓度而制作5-氮杂-dC稀释液。将此5-氮杂-dC稀释液添加到新鲜的RPMI1640培养基而制作含5 $\mu$ M的5-氮杂-dC的溶液。将得到的溶液作为5-氮杂-dC试剂使用。

[0296] 如下进行脱甲基化。首先,(c)的培养后,去除各孔的培养基。其后,添加各200 $\mu$ L的5-氮杂-dC试剂。作为未进行脱甲基化反应的对照,代替5-氮杂-dC试剂而添加200 $\mu$ L添加PBS的RPMI1640培养基。使其再于37 $^{\circ}$ C、5%CO<sub>2</sub>气氛下进行附着培养48小时。

[0297] 【(3) 报告子基因测定(荧光素酶测定)】

[0298] 从5-氮杂-dC的添加起48小时后,从培养板去除培养基。将各细胞用PBS清洗2次后,从它们的细胞各自提取荧光素酶(报告子蛋白质)。具体而言,如下进行提取。添加作为提取试剂的细胞溶解液(PicaGene Cell lysis buffer LC $\beta$ ,TOYO B-Net公司),于室温经15分钟温育细胞和细胞溶解液的悬浮液。其后,将此悬浮液以15,000rpm用离心机离心分离5分钟。由此,从悬浮液去除细胞残渣。由此得到上清。向此上清加荧光素酶底物溶液(PicaGene LT2.0,TOYO B-Net 公司)。接下来,添加作为检测试剂的荧光素酶基质溶液(荧光素溶液)。用光度计(Mithras LB940,Berthold公司)测定发生的发光强度。

[0299] 【(4) 结果】

[0300] 结果示于图15(a)及(b)。这些坐标图表示从载体导入细胞的报告子蛋白质信号强度的由脱甲基化反应的变化。

[0301] 图16(a)对于导入PGV-B2-GFAPp载体的细胞显示测定的数据。在图16(a)之左端,将自导入含非甲基化目标核酸序列的 PGV-B2-GFAPp载体,未受脱甲基化处理的细胞得到的信号强度记载为100%。在中央显示自导入含甲基化目标核酸序列的 PGV-B2-GFAPp载体,未受脱甲基化处理的细胞得到的相对信号强度。此相对信号强度是3.4%。在右端显示自导入含甲基化的被修饰碱基的PGV-B2-GFAPp载体,受脱甲基化处理的细胞得到的相对信号强度。此相对信号强度是2.4%。

[0302] 图15 (b) 对于导入pM53-R550K-GFAPp载体的细胞显示测定的数据。在图15 (b) 之左端将自导入含非甲基化的被修饰碱基的 pM53-R550K-GFAPp载体, 未受脱甲基化处理的细胞得到的信号强度记载为100%。在中央显示自导入含甲基化的被修饰碱基的 pM53-R550K-GFAPp载体, 未受脱甲基化处理的细胞得到的相对信号强度。此相对信号强度是34.1%。在右端显示自导入含甲基化的被修饰碱基的pM53-R550K-GFAPp载体, 受脱甲基化处理的细胞得到的相对信号强度。此相对信号强度是74.6%。

[0303] 如此, 用作为非复制型的质粒载体的PGV-B2-GFAPp载体, 对于目标核酸序列无法检测在细胞内发生的脱甲基化反应。另一方面, 通过使用作为自我复制型的质粒载体的pM53-R550K-GFAPp载体, 可对于目标核酸序列检测在细胞内发生的脱甲基化反应。再有, 作为目标核酸序列的GFAP基因启动子序列和作为被检细胞的神经胶质瘤 C6细胞中的对应的序列的相同性是90.26%。

[0304] 从这些的结果得知, 通过使用作为实施方式的1例的作为质粒型的自我复制载体的pM53-R550K-GFAPp载体, 使以下成为可能。即得知, 由其使用, 使以报告子蛋白质的信号强度作为指标检测被修饰碱基中的脱甲基化成为可能。即, 由这些结果示, 依赖于 pM53-R550K-GFAPp载体中所含的被修饰碱基的修饰的状态, 启动子活化的程度发生差异。此活化程度的差异可以报告子蛋白质的信号强度作为指标检测。如此, 由实施方式, 利用有与细胞所具有的特定的序列相同的目标核酸序列的自我复制载体, 得到对所述特定的序列的表观遗传的信息是可能的。另外, 含与被检细胞的特定的序列相同的序列, 并且作为质粒型的自我复制载体的pM53-R550K-GFAPp载体在被检细胞中自我复制时发生以下情况。即, 被检细胞的特定的序列内的被修饰碱基中的修饰的状态被pM53-R550K-GFAPp载体上的目标核酸序列中的被修饰碱基中的修饰的状态反映。

[0305] 从这些的结果得知, 通过使用自我复制载体, 以报告子蛋白质的信号强度作为指标, 得到被检细胞的特定的序列中的表观遗传的信息是可能的。

[0306] 以往, 为了由外来核酸得到基因组上的表观遗传的信息, 外来核酸与基因组的复制同时复制, 必受表观遗传的修饰反应。因此, 利用未整合进基因组的报告子载体判断细胞的特性是困难的。但是, 由实施方式, 可不向基因组上整合外来核酸而得到作为目的的核酸的修饰的状态等的表观遗传的信息。

[0307] **【实施例3: 基因组DNA和报告子载体中的CK19基因启动子区域(目标核酸序列) 的甲基化率的比较】**

[0308] 构建作为实施方式之一例的图5所示的报告子载体(自我复制载体51) 和复制起始蛋白质基因表达载体(有复制起始蛋白质单位52 的载体)。向报告子载体作为目标核酸序列整合CK19基因的启动子区域(-617~+61bp、SEQ ID NO:2), 比较导入所述载体的被检细胞的基因组DNA和报告子载体的目标核酸序列的甲基化率。

[0309] **【(1) 报告子载体的构建】**

[0310] 将作为目标核酸序列的CK19基因的启动子区域(-617~+61、SEQ ID NO:2) 以人的基因组DNA作为模板用PCR进行扩增。PCR 使用PrimeSTAR GLX DNA聚合酶(TaKaRa BIO) 进行。接下来示 PCR中使用的引物的碱基序列。

[0311] 正向引物, 5'-GCCTGGTCAACATGGTGAAAC-3' (SEQ ID NO: 13)

[0312] 反向引物, 5'-TGGGCTAGCCCAAGAAGCTGGTTCTGAGAGG-3' (SEQ ID NO:14)。

[0313] 将用PCR得到的目标核酸序列并入有猿猴病毒40的复制起始序列的报告子载体的荧光素酶基因的上游,制作图16所示的报告子载体 p19sLo。

[0314] 【(2) 复制起始蛋白质基因表达载体的构建】

[0315] 复制起始蛋白质基因作为猿猴病毒40的大T抗原基因 (SV40LT)。将同基因用使用 PrimeSTAR GLX DNA聚合酶 (TaKaRa BIO) 的PCR扩增后,整合进基因表达载体的巨细胞病毒初期启动子的下游。由此制作图17所示的表达复制起始蛋白质 (SV40LT) 基因的用载体 pCMV-LT。接下来显示PCR中使用的引物的碱基序列。

[0316] 正向引物, 5' -GGGGTACCAGATGGATAAAGTTTTAAACAGAGAGGAA-3' (SEQ ID NO:15)

[0317] 反向引物, 5' -GGGAATTCTTATGTTTCAGGTTTCAGGTTTCAGGGGGAG-3' (SEQ ID NO:16)。

[0318] 【(3) 转染】

[0319] 向100μL的Opti-MEM (Life Technologies) 单独加0.5μg的报告子载体,或者一同加0.4μg的报告子载体和0.1μg的复制起始蛋白质表达载体pCMV-LT。其后,加0.5μL的增强子试剂 (Plus reagent, Life Technologies)。使其于室温温育5分钟。向此溶液加1.25μL的 Lipofectamine LTX而使良好悬浮。其后,于室温温育25分钟。将溶液移到24孔板,加Huh-7细胞的悬浮液200μL而稳定混合。其后,在37℃、5%CO<sub>2</sub>气氛的温育器内培养细胞。

[0320] 【(4) 来自细胞的基因组DNA及报告子载体的调节】

[0321] 从转染起96小时后,从24孔板去除培养液,用磷酸缓冲生理盐水 (PBS) 清洗细胞。其后,加200μL的胰蛋白酶-EDTA于室温静置 5分钟。加300μL的PBS,将细胞回收到1.5ml管。其后,由离心使细胞沉淀。使细胞悬浮于200μL的PBS,使用DNeasy血和组织试剂盒 (Qiagen),得到含基因组DNA和报告子载体的溶液。

[0322] 【(5) 基因组DNA的甲基化率的检测】

[0323] 如下研究基因组DNA的CK19基因的启动子区域 (-617~+61bp、SEQ ID NO:2) 中所含的CCGG序列甲基化率 (第2胞嘧啶的5-甲基化的有无)。即,对于其,用由甲基化感受性限制性内切酶HpaII 和甲基化非感受性限制性内切酶MspI的甲基化检测法进行研究。HpaII和MspI均是识别CCGG序列的限制性内切酶。HpaII无法切断含甲基化胞嘧啶的CCGG序列 (C<sup>m</sup>CGG), MspI无关于甲基化胞嘧啶而切断CCGG序列。将上述 (4) 中调节的含基因组DNA和报告子载体的溶液用限制性内切酶HpaII或者MspI切断。其后,用QIA 快速PCR纯化试剂盒 (Qiagen) 纯化。将其作为限制性内切酶处理 DNA。以此DNA和限制性内切酶未处理DNA作为模板,使用下述引物进行PCR。由此,扩增含基因组DNA的COX2基因启动子区域的 CCGG序列的区域 (图18)。

[0324] 正向引物1 (G1), 5' -TCAGAGGGGACAAAAGGGGAGTTGG-3' (SEQ ID NO:17)

[0325] 反向引物1 (C1), 5' -AGAGGCCCTGCCCTCCAGAGGT-3' (SEQ ID NO:18)

[0326] 正向引物2 (C2), 5' -GCAAATTCCTCAGGGCTCAGATA-3' (SEQ ID NO:19)

[0327] 反向引物2 (G2), 5' -CCAGGCCTCCGAAGGACGACGTGGC-3' (SEQ ID NO:20)

[0328] 将PCR反应液用琼脂糖凝胶 (含溴化乙锭) 电泳。其后,由UV 照射可视化凝胶中的PCR扩增产物,拍摄照片。然后,将得到的照片中的各自的条带强度用图像解析软件Image J数值化。再者,由得到的数值和以下的式算出对A及B的全DNA的CCGG序列甲基化的比例。

[0329] 【数1】

[0330]

$$\text{甲基化率} = \frac{\text{HpaII 处理的 PCR (信号强度)} - \text{MspI 处理的 PCR (信号强度)}}{\text{限制性内切酶未处理 DNA 的 PCR (信号强度)} - \text{MspI 处理的 PCR (信号强度)}}$$

[0331] 这些结果示于图19a及图19(b)。其中,由上述摄影得到的照片均为背景黑,条带白。将此照片直接由图像解析软件数值化条带强度。但是,将其图像直接示于附图,则难见条带。从而,为了使条带变得更易见到,在图19(a)中,将由上述摄影得到的照片提取到计算机,使用图像加工软件,不损害条带强度及与其他数据的相对性地显示反转白黑色的像。以下,同样地摄影的照片均同样地反转白黑色。

[0332] 在图19a中,泳道N是限制性内切酶未处理DNA的PCR扩增产物的条带,表示供于PCR的全DNA量。泳道H是HpaII处理DNA的PCR扩增产物的条带。这在供于PCR的DNA之中,表示CCGG的第2胞嘧啶被甲基化的DNA量。泳道M是MspI处理DNA的PCR扩增产物,表示检测背景。从图19a的结果得知,在CCGG序列A及B的泳道H检测到PCR扩增产物的条带,其条带强度是B比A更高。

[0333] 将关于由上述的计算得到的CCGG序列甲基化的比例的结果示于图19b。从图19b的结果知,A和B的CCGG序列甲基化的比例各自是,A是2%,B是87%。

[0334] 其中,在上述式中,从对于PCR扩增产物的条带的信号强度算出甲基化率,但对于由其他扩增方法得到的扩增产物也可同样地自信号强度算出甲基化率。另外,甲基化率的算出方法不限于仅由上述式得到。

[0335] 【(6) 非复制报告子载体的甲基化率的检测】

[0336] 非复制报告子载体的甲基化率用上述(3)所示的方法进行。与报告子载体pC2sLo一同,或者p19sLo单独转染人肝癌细胞株Huh-7。96小时后,使用提取的DNA溶液检测甲基化率。在报告子载体的复制中,有复制起始序列和复制起始蛋白质在相同细胞内共存的必要。从而,将p19sLo单独导入细胞时,p19sLo在细胞内不复制。用与上述(5)同样的方法,研究报告子载体p19sLo的CK19基因的启动子区域(-617~+61bp、SEQ ID NO:2)中所含的CCGG序列甲基化率(第2胞嘧啶的5-甲基化的有无)。将上述(4)中调制的DNA用限制性内切酶HpaII、或者MspI切断。其后,用QIA快速PCR纯化试剂盒(Qiagen)纯化。将其作为限制性内切酶处理DNA。以这些DNA和限制性内切酶未处理的DNA作为模板,进行使用下述引物的PCR。由此扩增含p19sLo的CK19基因启动子区域的CCGG序列的区域(图18)。

[0337] 正向引物1(P1),5'-TAGGCTGTCCCCAGTGCAAGT-3'(SEQ ID NO:21)

[0338] 反向引物1(C1),5'-AGAGGCCCTGCCCTCCAGAGGT-3'(SEQ ID NO:18)

[0339] 正向引物2(C2),5'-GCAAATTCCTCAGGGCTCAGATA-3'(SEQ ID NO:19)

[0340] 反向引物2(P2),5'-AATGCCAAGCTTACTTAGATCG-3'(SEQ ID NO:22)。

[0341] 将PCR反应液的琼脂糖凝胶(含溴化乙锭)电泳的结果示于图20a。从图20a的结果知,在A和B的泳道H检测不到PCR扩增产物。从而,CCGG序列的甲基化的比例是,A及B均是0%(图20b)。另外,在非复制报告子载体中不发生A和B的甲基化。

[0342] 【(7) 复制报告子载体的甲基化率的检测】

[0343] 将报告子载体p19sLo和pCMV-LT共导入Huh-7,96小时后,使用提取的DNA检测复制报告子载体的甲基化率。将P19sLo和pCMV-LT共导入细胞,则由于复制起始序列和复制起始蛋白质在细胞内共存,p19sLo在细胞内复制。以将DNA用HpaII、或者MspI处理的限制性

内切酶处理DNA和限制性内切酶未处理DNA作为模板,用与上述(6)相同的引物进行PCR。由此扩增含p19sLo的CK19 基因启动子区域的CCGG序列的区域。将PCR反应液的琼脂糖凝胶(含溴化乙锭)电泳的结果示于图21a。结果,在A和B的泳道H检测到PCR扩增产物。CCGG序列的甲基化的比例是,A是0%,B是16%(图21b)。再者,组合DpnI法(在细胞内复制的载体的检测法)和由HpaII和MspI的甲基化检测法,研究仅在细胞内复制的报告子载体的CK19基因启动子区域CCGG序列的甲基化率(图22a)。结果,在A和B的泳道H中检测到PCR扩增产物。各CCGG序列的甲基化的比例是,A是0%、B是24%(图22b)。报告子载体的CK19 基因启动子区域以与基因组DNA的CK19基因启动子区域(目标核酸序列)大致相同的比率被甲基化。

[0344] **【实施例4:基因组DNA和报告子载体中的COX2基因启动子区域(目标核酸序列)的甲基化率的比较】**

[0345] 作为实施方式之一例,构建图5所示的报告子载体(自我复制载体51)和复制起始蛋白质基因表达载体(有复制起始蛋白质单位52 的载体)。向报告子载体作为目标核酸序列并入COX2基因的启动子区域(-540~+115bp、SEQ ID NO:3)。比较导入此载体的被检细胞的基因组DNA和报告子载体的目标核酸序列的甲基化率。

[0346] **【(1)报告子载体的构建】**

[0347] 将作为目标核酸序列的COX2基因的启动子区域(-540~+115bp、SEQ ID NO:3),以人的基因组DNA作为模板用PCR扩增。PCR 使用PrimeSTAR GLX DNA聚合酶(TaKaRa BIO)进行。接下来示 PCR中使用的引物的碱基序列。

[0348] 正向引物,5'-CTTAACCTTACTCGCCCCAGTCT-3'(SEQ ID NO:23)

[0349] 反向引物,5'-AGGCTCGAGCGCGGGGTAGGCTTTGCTGTCTGAG-3'(SEQ ID NO:24)。

[0350] 将用PCR得到的目标核酸序列整合到有猿猴病毒40的复制起始序列的报告子载体的荧光素酶基因的上游。由此制作图23所示的报告子载体pC2sLo。

[0351] **【(2)复制起始蛋白质基因表达载体】**

[0352] 使用与上述的例3的(2)相同的复制起始蛋白质(SV40LT)基因表达用载体pCMV-LT。

[0353] **【(3)报告子载体的甲基化】**

[0354] 调节由报告子载体和反应缓冲液组成的反应溶液(50mM NaCl、10mM Tris-HCl, 10mM MgCl<sub>2</sub>,1mM DTT,160μM S-腺苷基甲硫氨酸)。向此加DNA甲基化酶:SssI CpG甲基转移酶(New England Biolabs)于37℃反应一晚,将报告子载体的CG序列的胞嘧啶甲基化。

[0355] **【(4)转染】**

[0356] 用与实施例3的(3)同样的方法进行转染。

[0357] **【(5)来自细胞的基因组DNA及报告子载体的调节】**

[0358] 来自细胞的基因组DNA及报告子载体的调节用与上述实施例3 的(4)同样的方法进行。

[0359] **【(6)基因组DNA的甲基化率的检测】**

[0360] 如以下一样研究基因组DNA的COX2基因启动子(-540~ +115bp、SEQ ID NO:3)中所含的CCGG序列甲基化率(第2胞嘧啶的5-甲基化的有无)。即,对于其,用由甲基化感受性限制性内切酶HpaII和甲基化非感受性限制性内切酶MspI的甲基化检测法进行研究。将(5)中调制的含基因组DNA和报告子载体的溶液用限制性内切酶HpaII、或者MspI切断后,用QIA



快速PCR纯化试剂盒 (Qiagen) 纯化, 将其作为限制性内切酶处理DNA。以此DNA和限制性内切酶未处理DNA作为模板使用, 再进行使用下述引物的PCR。由此扩增含基因组DNA的COX2基因启动子区域的CCGG序列的区域(图24)。

[0361] 正向引物1 (C1), 5' -GGAAGCCAAGTGTCTTCTGC-3' (SEQ ID NO:25)

[0362] 反向引物1 (C2), 5' -GGGCAGGGTTTTTACCCAC-3' (SEQ ID NO:26)

[0363] 正向引物2 (C3), 5' -AGCTTCCTGGGTTTCCGATTT-3' (SEQ ID NO:27)

[0364] 反向引物2' (G2), 5' -GCCAGGTACTCACCTGTATGGCTG-3' (SEQ ID NO:28)。

[0365] 将PCR反应液用琼脂糖凝胶(含溴化乙锭)电泳。其后, 由UV 照射可视化凝胶中的PCR扩增产物而拍摄照片。将结果示于图25a。从图25a的结果知, 在CCGG序列A和B的泳道H中检测到PCR 扩增产物的条带。其条带强度是A比B更高。用与实施例3的(5) 同样的方法算出的A和B的CCGG序列甲基化的比例各自是29%和 7% (图25b)。

[0366] 【(7) 非复制非甲基化报告子载体的甲基化率的检测】

[0367] 非复制报告子载体的甲基化率是, 用上述(3)所示的方法, 与报告子载体pC2sLo一同, 或者单独转染人肝癌细胞株Huh-7。其96小时后, 使用提取的DNA溶液检测。用与上述(5)同样的方法, 研究报告子载体pC2sLo的COX2基因的启动子区域(-540~+115bp、SEQ ID NO:3)中所含的CCGG序列甲基化率。将在上述(5)调节的DNA 用限制性内切酶HpaII、或者MspI切断。其后, 用QIA快速PCR纯化试剂盒 (Qiagen) 纯化。将其作为限制性内切酶处理DNA。以这些的DNA和限制性内切酶未处理的DNA作为模板, 进行使用下述引物的PCR。扩增含pC2sLo的COX2基因启动子区域的CCGG序列的区域(图24)。

[0368] 正向引物1 (C1), 5' -GGAAGCCAAGTGTCTTCTGC-3' (SEQ ID NO:25)

[0369] 反向引物1 (C2), 5' -GGGCAGGGTTTTTACCCAC-3' (SEQ ID NO:26)

[0370] 正向引物2 (C3), 5' -AGCTTCCTGGGTTTCCGATTT-3' (SEQ ID NO:27)

[0371] 反向引物2' (P2), 5' -AATGCCAAGCTTACTTAGATCG-3' (SEQ ID NO:29)。

[0372] 将PCR反应液的琼脂糖凝胶(含溴化乙锭)电泳的结果示于图 26a。从图26a的结果知, 在A及B的泳道H中检测不到PCR扩增产物。CCGG序列的甲基化的比例是, A及B均是0%, 在非复制报告子载体中不发生A及B的甲基化(图26b)。

[0373] 【(8) 复制非甲基化报告子载体的甲基化率的检测】

[0374] 复制报告子载体的甲基化率是, 将报告子载体pC2sLo和复制起始蛋白质基因表达载体pCMV-LT共导入Huh-7。其96小时后, 使用提取的DNA溶液进行检测。以将DNA用HpaII或者MspI处理的限制性内切酶处理DNA和限制性内切酶未处理DNA作为模板, 用与上述(7)相同的引物进行PCR。由此扩增含pC2sLo的COX2基因启动子区域的CCGG序列的区域。将PCR反应液的琼脂糖凝胶(含溴化乙锭)电泳的结果示于图27a。结果, 在A及B的泳道H中检测到PCR扩增产物。CCGG序列甲基化的比例是, A是14%, B是24% (图27b)。再者, 通过组合DpnI法(在细胞内复制的载体的检测法) 和由HpaII和MspI的甲基化检测法, 研究仅在细胞内复制的报告子载体的COX2基因启动子区域CCGG序列的甲基化率(图28a)。结果, 在A及B的泳道H中检测到PCR扩增产物。各CCGG序列的甲基化的比例是, A是57%, B是23% (图28b)。复制的报告子载体的COX2基因启动子区域以与基因组DNA的COX2基因启动子区域(目标核酸序列)大致相同的比率被甲基化。

[0375] 【(9) 非复制甲基化报告子载体的甲基化率的检测】

[0376] 在非复制甲基化报告子载体的甲基化率的检测,使甲基化报告子载体pC2sLo单独转染人肝癌细胞株Huh-7。其96小时后,使用由上述(5)的方法提取的DNA溶液。将DNA用限制性内切酶HpaII、或者MspI切断。其后,将其用QIA快速PCR纯化试剂盒(Qiagen)纯化,作为限制性内切酶处理DNA。以这些DNA和限制性内切酶未处理的DNA作为模板,进行使用下述引物的PCR。由此扩增含pC2sLo的COX2基因启动子区域的CCGG序列的区域。将PCR反应液的琼脂糖凝胶(含溴化乙锭)电泳的结果示于图29a。从图29a的结果知,在A及B的泳道H中检测到PCR扩增产物。CCGG序列的甲基化的比例是,A是100%,B是86%(图29b)。非复制甲基化报告子载体的甲基化率与宿主细胞的基因组DNA的甲基化率不一致。

[0377] 【(10) 复制甲基化报告子载体的甲基化率的检测】

[0378] 在复制甲基化报告子载体的甲基化率的检测中,将报告子载体 pC2sLo和复制起始蛋白质基因表达载体pCMV-LT共导入Huh-7, 96小时后,使用提取的DNA溶液。将用HpaII、或者MspI处理的限制性内切酶处理DNA和限制性内切酶未处理DNA作为模板使用。用与上述(6)相同的引物进行PCR。由此扩增含pC2sLo的COX2基因启动子区域的CCGG序列的区域。将PCR反应液的琼脂糖凝胶(含溴化乙锭)电泳的结果示于图30a。结果,在A及B的泳道H中检测到PCR扩增产物。CCGG序列甲基化的比例是,A是45%,B是79%(图31b)。再者,组合DpnI法(在细胞内复制的载体的检测法)和由HpaII和MspI的甲基化检测法,研究仅在细胞内复制的报告子载体的COX2基因启动子区域CCGG序列的甲基化率(图31a)。结果,在A及B的泳道H中检测到PCR扩增产物。各CCGG序列的甲基化的比例是,A是95%,B是35%(图31b)。比较在细胞内复制的报告子载体和基因组DNA的COX2基因启动子区域内CCGG序列的甲基化率,则B被强脱甲基化,A的甲基化的比例更高。

[0379] 【实施例5:由荧光素酶测定的目标核酸序列(COX2基因启动子区域)的脱甲基化的检测】

[0380] 用与上述实施例4同样的方法,将甲基化/非甲基化报告子载体 pC2sLo单独、或者与复制起始蛋白质表达载体pCMV-LT一同转染到人肝癌细胞株Huh-7。其72小时培养后,测定pC2sLo来源的报告子活性。从24孔板去除培养液,用PBS清洗。其后,加150μL的细胞溶解液(Promega)于-80℃静置30分钟以上而冷冻。将细胞溶解液于室温解冻后,将溶解液回收离心管。由此离心使细胞残渣沉淀。将上清10μL分注到96孔板(Nunc)。向其加50μL的荧光素酶底物溶液(One-Glo Luciferase Assay System,Promega)而测定细胞溶解液的发光强度。将结果示于图32。图32的坐标图的纵轴显示发光强度的恢复率。此发光强度的恢复率是脱甲基化的指标。发光强度的恢复率是以自导入非甲基化报告子载体的细胞得到的发光强度作为100%时的自导入甲基化报告子DNA的细胞得到的发光强度的比例。在坐标图之左端显示非复制报告子载体的发光强度的恢复率,在右端显示复制报告子载体的恢复率。从此坐标图知,发光强度的恢复率是复制报告子载体比非复制报告子载体更高。由此,在复制报告子载体中,以高灵敏度检测在目标核酸序列上发生的脱甲基化变得可能。从结果知,通过使用在细胞内复制的报告子载体,可以报告子活性作为指标,得到被检细胞的特定的序列中的表观遗传的信息。

[0381] 【实施例6:有2种报告子基因表达单位(含官能团被取代的目标核酸序列的单位和含官能团未被取代的目标核酸序列的单位)及复制起始序列的自我复制载体】

[0382] 【(1) 载体的制作】

[0383] 构建作为实施方式之一例的图6所示的报告子载体(自我复制载体61)。将整合了虾来源的荧光素酶的报告子载体pNL1.1(Promega 公司)用限制性内切酶KpnI和BamHI切断。将其DNA末端用T4DNA 聚合酶平滑化。其后,纯化含虾荧光素酶和SV40转录终止序列的DNA 片段。将此片段用限制性内切酶SspI切断的p19sLo由T4DNA连接酶连接而构建自我复制载体p19sLo-NL。在此p19sLo-NL上,将CpG 的胞嘧啶掺入甲基化的COX2基因启动子(SEQ ID NO:3)。以人的基因组作为模板,由PCR扩增COX2基因启动子。接下来示PCR中使用的引物序列。

[0384] 正向引物, 5'-GGGGCTAGCCTTAACCTTACTCGCCCCAGTCT-3' (SEQ ID NO:30)

[0385] 反向引物, 5'-AGGCTCGAGCGCGGGGGTAGGCTTTGCTGTCTGAG-3' (SEQ ID NO:24)。

[0386] 纯化PCR扩增产物(COX2基因启动子)。其后,调节含扩增产物和反应缓冲液的反应溶液(50mM NaCl,10mM Tris-HCl,10mM MgCl<sub>2</sub>,1mM DTT,160μM S-腺苷基甲硫氨酸)。向其中加DNA甲基化酶:SssI CpG甲基转移酶(New England Biolabs)于37℃反应一晚。由此将COX2基因启动子的CG序列的胞嘧啶甲基化。将此甲基化的DNA片段(甲基化COX2基因启动子)由T4DNA连接酶连接到用限制性内切酶KpnI和XhoI处理的自我复制报告子载体p19sLo-NL。由此,构建含2种报告子基因表达单位(包含含甲基化 COX2基因启动子的报告子基因表达单位和含非甲基化CK19基因启动子的报告子基因表达单位)和复制起始序列的自我复制报告子载体 p19sLo-mC2sNL(图33)。

[0387] 【(2) 转染】

[0388] 向100μL的Opti-MEM(Life Technologies)单独加、或者与0.4μg 的p19sLo-mC2sNL和0.1μg的复制起始蛋白质表达载体pCMV-LT 一同加0.5μg的报告子载体。其后,加0.5μL的增强子试剂(Plus reagent,Life Technologies)而于室温温育5分钟。向此溶液加1.25μL 的Lipofectamine LTX而良好悬浮。其后,于室温温育25分钟。将溶液移到24孔板。向其中加Huh-7细胞的悬浮液200μL而稳定混合。由此将载体导入细胞(图34)。

[0389] 【实施例7:有报告子基因表达单位、及表达复制起始基因的单位、及复制起始序列的自我复制载体】

[0390] 【(1) 载体的制作】

[0391] 构建作为实施方式的1例的图4所示的报告子载体(自我复制载体41)。从pCMV-LT(实施例3(2))将复制起始蛋白质基因表达单位用限制性内切酶BglI和BamHI切断。将其DNA末端用T4DNA 聚合酶平滑化。其后,纯化复制起始蛋白质基因表达单位的DNA片段。将此DNA片段与用限制性内切酶SspI切断的pC2sLo(实施例4(1))连接而构建自我复制载体pC2sLo-CMVL(图35)。

[0392] 【(2) 转染】

[0393] 向100μL的Opti-MEM(Life Technologies)加0.5μg的 pC2sLo-CMVL。其后,加0.5μL的增强子试剂(Plus reagent,Life Technologies),于室温温育5分钟。向此溶液加1.25μL的 Lipofectamine LTX而使良好悬浮。其后,于室温温育25分钟。将溶液移到24孔板。向其中加Huh-7细胞的悬浮液200μL而稳定混合。其后,在37℃、5%CO<sub>2</sub>气氛的温育器内培养细胞。

[0394] 【(3) 来自细胞的基因组DNA及报告子载体的调节】

[0395] 基因组DNA及报告子载体的调节用与上述实施例3的(4)同样的方法进行。

## [0396] 【(4) 报告子载体的甲基化率的检测】

[0397] 报告子载体的甲基化率是,将自我复制报告子载体 pC2sLo-CMVLT导入Huh-7,其72小时后,使用提取的DNA溶液用与上述实施例4(7)同样的方法进行研究。将在上述(3)调制的溶液用限制性内切酶HpaII、或者MspI切断。其后,将其用QIA快速PCR纯化试剂盒(Qiagen)纯化。以这些的DNA和限制性内切酶未处理DNA作为模板,用与上述实施例4(7)相同的引物进行PCR。扩增含pC2sL-CMVLT的COX2基因启动子区域(-540~+115bp、SEQ ID NO:3)的CCGG序列的区域(图24)。将PCR反应液的琼脂糖凝胶(含溴化乙锭)电泳的结果示于图36a。从图36a的结果知,在A及B的泳道H中检测到PCR扩增产物。CCGG序列的甲基化的比例是,A是9%,B是0.5%。在自我复制报告子载体pC2sLo-CMVLT的COX2基因启动子区域中,与基因组DNA的COX2基因启动子区域(例4(6)、图25)同样地,A的甲基化率高。

## [0398] 【符号的说明】

[0399]

1、31、41、51、61	自我复制载体
2、62	报告子基因表达单位
3	复制起始序列
4、64	目标核酸序列
5、65	报告子基因
6、34、46、55、66	转录终止信号序列
7	复制起始蛋白质
32	IRES
33、45、54	编码复制起始蛋白质的序列
42	复制起始蛋白质表达单位
44、53	组成型表达的启动子
45、50	被检细胞
51	第1载体
56	第2载体

[0001]

## 序列表

&lt;110&gt; Kabushiki Kaisha Toshiba

&lt;120&gt; 得到细胞的表现遗传的信息的方法、判断细胞的特性的方法、判断药剂感受性的或选择药剂或免疫疗法剂的种类的方法、疾病的诊断方法、以及自我复制载体、测定试剂盒及分析装置

&lt;130&gt; 14S0091PCT

&lt;150&gt; JP 2013-105481

&lt;151&gt; 2013/05/17

&lt;160&gt; 31

&lt;170&gt; PatentIn version 3.4

&lt;210&gt; 1

&lt;211&gt; 1683

&lt;212&gt; DNA

&lt;213&gt; 鼠

&lt;400&gt; 1

```

gtctgtaagc tgaagacctg gcagtgtga gctggtcacc cccaggacc tcctttgtg      60
cccaacgagt gactcacctt ggcatagaca taatggtcag gggtagggcac gcagcctgct      120
tcccgtgtg ctccaggcct ccttcgatgc ttccgagaa gtctattgag ctgggagctt      180
gtactgcacc cggggctgac atectggcat cctgggataa aagcagccca cggggctgcc      240
cttgccatat gcctcactgg cggcagagga caaggctcta ttcagcaagt gccctggagt      300
agacaccaga agcccaagca tgggcagagg aaggcagggg ttggggggag cagagctgtc      360
tgtgttccag aagcccaagg acacagatgg ctaaggcgcc tgggagggac ctgagtggaa      420
gagatagatg ggctgaagt ctcaagcagc aacagcctcc tccccccat tggtaggggt      480
ggggttttgt ttcccgacc tacatatccc tcagaggcct ggtgtgtagg aatttaaagg      540
aggtaaactc cctgagagaa tgaggggtac ccaggaagac ggggtgttac agaaagactc      600
cagcatgcac agccaactca ctcaaaacta ctctgtcagg ggctgccggg ggccaggctc      660
ggggtggggg gtgggggggg aaagagaagc tggaccaggg agaaatggcc cactagctg      720
gatatgaggc cacagagggg ctcaggaatg aagcctgctg tcttacccta ttaggatctg      780
cgtgcatacc ttctgtgtg cactctaaac acacagccag aggtcaagt tgacctgga      840
gtcacagaga gggctccaac cttagccctc cactcctgaa ctccaggaat gagaagatag      900
agttggagcg attcagggga gaggactctg ttgagaatgg gggccacagg aaactgtaat      960
ataggttgat cccggaggaa gggaataggt tcttcaagtt ctagcatct cacaggcccc      1020
cagagaagga cagagttggg gtgttcctgg cttacaggct ctaagaactg gaagctgatt      1080
acccaccacaa gctgtgcact ctctgtctct gtctctgtct ctgtgtgtgc gcgctcgtgc      1140
acacttatca cacaatgtt catgtgtgtg cacatagatg agttgagacc agaggtcaac      1200
ctcaggcact gttgccttgg tttctgaga gagcatctct ctctggacct ggaactcgcc      1260
aattagttag agccaggaag tctgctgatt ttactgccc agcactggag ttacaagta      1320
tgcactgtca acccaggcct ttgtattca ttctgcagct agaacttggg tgggtcttca      1380
tgcttgacag gcaagcaatt tatggactaa gctgtttcct cggccctctc ttgaccatt      1440
taccagaaag ggggttcctt gatcaatggc gaagccaggc tgggtttccc aagaagcct      1500

```

	tgactctggg tacagtgacc tcagtggggt gagaggagtt ctccccttag ctgggctggg	1560
	gccagcttc accccctcag gctattcagt gggggtgctt ccaggaagtc aggggcagat	1620
	ttagtccaac ccgttctcc ataaaggccc tgacatccca ggagccagca gaggcagggc	1680
	agg	1683
	<210> 2	
	<211> 678	
	<212> DNA	
	<213> 智人 (Homo Sapiens)	
	<400> 2	
	gcctgtcaa catggtgaaa cctgtctct actaaaaata caaaaattag ctgggcgtgg	60
	tggcgcgtgc ctgtaatccc agctactccg gaggtgagg caggagaatc gcttgaaccc	120
	gggagatgga ggctgcagtg ggccgagatc acaccactgc actccagtct gggcgacaga	180
	gactcgtctc caaaaaaaca aaaacaaaat cactgggtca ggggtgtgta ggaatatgac	240
	cccagaggga ctgtaattcc cagtgggtgc aaactctggg tgatcttaaa gggtagaggct	300
	cagaatgtgg ctccaggga cagggtgtgc caggatgtgc cctgacgggg gaaaggccag	360
	aacagggtct gcagagagag agtgggggag tgcgggtcgg agcttctcgc cggaccgggg	420
	cggggcacct ctggagggca ggggcctctg gtctctggga ggggagggaa ttgaccaatg	480
	gggagagagc ccatatttgc tctcaggagc ctgcaaattc ctcagggtc agatatccgc	540
	ccctgacacc attcctccct tccccctcc accggccgcg ggcataaaag gcgccaggtg	600
	agggcctcgc cgctcctccc gcgaatcgca gcttctgaga ccagggttgc tccgtccgtg	660
[0002]	ctccgcctcg ccctcgag	678
	<210> 3	
	<211> 655	
	<212> DNA	
	<213> 智人 (Homo Sapiens)	
	<400> 3	
	cttaacctta ctgccccag tctgtccga cgtgacttcc tgcacctct aaagacgtac	60
	agaccagaca cggcggcggc ggccggagag gggattccct gcgccccgg acctcagggc	120
	cgctcagatt cctggagagg aagccaagtg tccttctgcc ctccccgggt atcccatcca	180
	aggcgatcag tccagaactg gctctcgga gcgctcgggc aaagactgcg aagaagaaaa	240
	gacatctggc ggaacctgt gcgcctgggg cggtggaact cggggaggag agggagggat	300
	cagacaggag agtggggact acccctctg ctcccaaatt ggggcagctt cctgggtttc	360
	cgattttctc atttccgtgg gtaaaaaacc ctgccccac cgggcttac caatttttt	420
	aaggggagag gagggaaaaa ttgtggggg gtacgaaaag gcggaaagaa acagtcattt	480
	cgtcacatgg gcttggtttt cagtcttata aaaaggaagg ttctctcgg tagcgaccaa	540
	ttgtcatacg acttgacgtg agcgtcagga gcacgtccag gaactctca gcagcgctc	600
	cttcagctcc acagccagac gccctcagac agcaaagcct acccccgcg tcgag	655
	<210> 4	
	<211> 222	
	<212> DNA	
	<213> 猿猴病毒 40	
	<400> 4	
	cagacatgat aagatacatt gatgagtttg gacaaaccac aactagaatg cagtgaaaa	60

	aatgctttat ttgtgaaatt tgtgatgcta ttgctttatt tgtaaccatt ataagctgca	120
	ataaacaagt taacaacaac aattgcattc attttatgtt tcagggttcag ggggagggtgt	180
	gggagggtttt ttaaagcaag taaaacctct acaaatgtgg ta	222
	<210> 5	
	<211> 228	
	<212> DNA	
	<213> 牛	
	<400> 5	
	gtttaaaccc gctgatcagc ctgcactgtg ccttctagtt gccagccatc tgttgtttgc	60
	ccctcccccg tgccttctt gaccttgaa ggtgcactc ccactgtcct ttcctaataa	120
	aatgaggaata ttgcatcgca ttgtctgagt aggtgtcatt ctattctggg gggtaggggtg	180
	gggcaggaca gcaaggggga ggattgggaa gacaatagca ggcatgct	228
	<210> 6	
	<211> 165	
	<212> DNA	
	<213> 猿猴病毒 40	
	<400> 6	
	gctagcaata aaatatcttt attttcatta catctgtgtg ttggtttttt gtgtgaatcg	60
	atagtactaa catagctct ccatcaaac aaaacgaaac aaaacaaact agcaaaatag	120
	gctgtcccca gtgcaagtgc aggtgccaga acatttctcg ctage	165
	<210> 7	
[0003]	<211> 2133	
	<212> DNA	
	<213> 猿猴病毒 40	
	<400> 7	
	atggataaag ttttaaacag agaggaatct ttgcagctaa tggaccttct aggtcttgaa	60
	aggagtgcct gggggaatat tcctctgatg agaaaggcat atttaaaaa atgcaaggag	120
	tttcatcctg ataaaggagg agatgaagaa aaaatgaaga aaatgaatac tctgtacaag	180
	aaaaatggaag atggagtaaa atatgctcat caacctgact ttggaggctt ctgggatgca	240
	actgagattc caacctatgg aactgatgaa tgggagcagt ggtggaatgc ctttaatgag	300
	gaaaacctgt tttgctcaga agaaatgcca tctagtgatg atgaggctac tgctgactct	360
	caacattcta ctctccaaa aaagaagaga aaggtagaag accccaagga ctttccttca	420
	gaattgctaa gttttttgag tcatgctgtg tttagtaata gaactcttgc ttgctttgct	480
	atttacacca caaaggaaaa agctgcaatg ctatacaaga aaattatgga aaaatattct	540
	gtaaccttta taagtaggca taacagtat aatcataaca tactgttttt tcttactcca	600
	cacaggcata gagtgtctgc tattaataac tatgctcaaa aattgtgtac ctttagcttt	660
	ttaatttgta aaggggttaa taaggaatat ttgatgtata gtgccttgac tagagatcca	720
	ttttctgtta ttgaggaaag ttgtccaggt ggggttaaagg agcatgattt taatccagaa	780
	gaagcagagg aaactaaaca agtgccttgg aagcttgtaa cagagtatgc aatggaaaca	840
	aaatgtgatg atgtgttgtt attgcttggg atgtacttgg aatttcagta cagttttgaa	900
	atgtgtttaa aatgtattaa aaaagaacag cccagccact ataagtacca tgaagagcat	960
	tatgcaaatg ctgctatatt tgctgacagc aaaaacaaa aaacatatg ccaacaggct	1020

```

gttgatactg ttttagctaa aaagcgggtt gatagcctac aattaactag agaacaatg 1080
ttaacaaaca gatttaatga tcttttgat aggatggata taatgtttgg ttctacaggc 1140
tcctgctgaca tagaagaatg gatggctgga gttgcttggc tacactgttt gttgcccaaa 1200
atggattcag tgggtatga ctttttaaaa tgcattgtgt acaacattcc taaaaaaga 1260
tactggctgt ttaaaggacc aattgatagt ggtaaaacta cattagcagc tgctttgctt 1320
gaattatgtg ggggaaagc tttaaatgtt aatttgcctt tggacaggct gaactttgag 1380
ctaggagtag ctattgacca gtttttagta gtttttgagg atgtaaaggg cactggaggg 1440
gagtcagag atttgccttc aggtcaggga attaataacc tggacaattt aagggtattt 1500
ttggatggca gtgttaaggt aaacttagaa aagaaacacc taaataaaag aactcaaata 1560
tttccccctg gaatagtcac catgaatgag tacagtgtgc ctaaaacact gcaggccaga 1620
tttgtaaaac aaatagattt taggccaga gattatttaa agcattgcct ggaacgcagt 1680
gagtttttgt tagaaaagag aataattcaa agtggcattg ctttgccttct tatgttaatt 1740
tggtacagac ctgtggctga gtttgcctaa agtattcaga gcagaattgt ggagtggaaa 1800
gagagattgg acaagaggtt tagtttgcga gtgtatcaaa aaatgaagtt taatgtggct 1860
atgggaattg gagttttaga ttggctaaga aacagtgatg atgatgatga agacagccag 1920
gaaaatgctg ataaaaatga agatggtggg gagaagaaca tggaagactc agggcatgaa 1980
acaggcattg attcacagtc ccaaggctca tttcagggcc ctcagtcctc acagtctgtt 2040
catgatcata atcagccata ccacatttgt agaggtttta cttgctttaa aaaacctccc 2100
acacctcccc ctgaacctga acctgaacaa taa 2133

```

[0004]

```

<210> 8
<211> 2133
<212> DNA
<213> 猿猴病毒 40

```

```

<400> 8
atggataaag ttttaaacag agaggaatct ttgcagctaa tggaccttct aggtcttgaa 60
aggatgcctt gggggaatat tcctctgatg agaaaggcat atttaaaaaa atgcaaggag 120
tttcattctg ataaaggagg agatgaagaa aaaatgaaga aaatgaatac tctgtacaag 180
aaaaatggaag atggagtaaa atatgctcat caacctgact ttggaggett ctgggatgca 240
actgagattc caacctatgg aactgatgaa tgggagcagt ggtggaatgc ctttaatgag 300
gaaaacctgt ttgtctcaga agaaatgcca tctagtgatg atgaggctac tgctgactct 360
caacattcta ctctccaaa aaagaagaga aaggtagaag accccaagga ctttcttca 420
gaattgctaa gttttttgag tcatgctgtg tttagtaata gaactcttgc ttgctttgct 480
atttacacca caaagaaaa agctgcactg ctatacaaga aaattatgga aaaatattct 540
gtaaccttta taagtaggca taacagtat aatcataaca tactgttttt tcttactcca 600
cacaggcata gattgtctgc tattaataac tatgtcctaa aattgtgtac ctttagcttt 660
ttaatttgta aaggggttaa taaggaatat ttgatgtata gtgccttgac tagagatcca 720
tttctgttta ttgaggaaaag ttgcccagg ggtttaaagg agcatgattt taatccagaa 780
gaagcagagg aaactaaaca agtgtcctgg aagcttgtaa cagagtatgc aatggaacaa 840
aaatgtgatg atgtgttgtt attgcttggg atgtacttgg aatttcagta cagttttgaa 900
atgtgtttaa aatgtattaa aaaagaacag ccagccactc ataagtacca tgaaaagcat 960

```



	tatgcaaatg ctgctatatt tgctgacagc aaaaaccaa aaacatgatg ccaacaggct	1020
	gttgatactg ttttagctaa aaagcgggtt gatagcctac aattaactag agaacaaatg	1080
	ttaacaaaca gatttaatga tcttttgat aggatggata taatgtttgg ttctacaggc	1140
	tctgctgaca tagaagaatg gatggctgga gttgcttggc tacactgttt gttgccccaa	1200
	atggagtcag tgggtatga ctttttaaaa tgcattggtgt acaacattcc taaaaaaga	1260
	tactggctgt ttaaaggacc aattgatagt ggtaaaacta cattagcagc tgctttgctt	1320
	gaattatgtg ggggaaagc tttaaatgtt aatttgcctt tggacaggct gaactttgag	1380
	ctaggagtag ctattgacca gtttttagta gtttttgagg atgtaaaggg cactggaggg	1440
	gagtcagag atttgccttc aggtcaggga attaataacc tggacaattt aagggattat	1500
	ttggatggca gtgttaaggt aaacttagaa aagaaacacc taaataaaag aactcaaata	1560
	ttccccctg gaatagtcac catgaatgag tacagtgtgc ctaaaacact gcaggccaga	1620
	tttgtaaac aaatagattt taggccaga gattatttaa agcattgcct ggaacgcagt	1680
	gagttttgt tagaaaagag aataattcaa agtggcattg ctttgcttct tatgttaatt	1740
	tggtagacac ctgtggctga gtttgctcaa agtattcaga gcagaattgt ggagtggaaa	1800
	gagagattgg acaagagtt tagtttgtca gtgtatcaaa aaatgaagtt taatgtggct	1860
	atgggaattg gagttttaga ttggctaaga aacagtgatg atgatgatga agacagccag	1920
	gaaaatgctg ataaaaatga agatggtggg gagaagaaca tggaagactc agggcatgaa	1980
	acaggcattg attcacagtc ccaaggctca ttccaggccc ctgagtcctc acagtctgtt	2040
[0005]	catgatcata atcagccata ccacatttgt agaggtttta cttgctttaa aaaacctccc	2100
	acacctcccc ctgaacctga acctgaaaca taa	2133
	<210> 9	
	<211> 2133	
	<212> DNA	
	<213> 猿猴病毒 40	
	<400> 9	
	atggataaag ttttaaacag agaggaatct ttgcagctaa tggaccttct aggtcttgaa	60
	aggagtgcct gggggaatat tcctctgatg agaaaggcat atttaaaaa atgcaaggag	120
	tttcatcctg ataaaggagg agatgaagaa aaaatgaaga aaatgaatac tctgtacaag	180
	aaaatggaag atggagtaaa atatgctcat caacctgact ttggaggctt ctgggatgca	240
	actgagattc caacctatgg aactgatgaa tgggagcagt ggtggaatgc ctttaatgag	300
	gaaaacctgt tttgctcaga agaaatgcca tctagtgatg atgaggctac tgctgactct	360
	caacattcta ctctccaaa aaagaagaga aaggtagaag accccaagga ctttccttca	420
	gaattgctaa gttttttgag tcatgctgtg tttagtaata gaactcttgc ttgctttgct	480
	atttacacca caaaggaaaa agctgcactg ctatacaaga aaattatgga aaaatattct	540
	gtaaccttta taagtaggca taacagttat aatcataaca tactgttttt tcttactcca	600
	cacaggcata gagtgtctgc tattaataac tatgctcaaa aattgtgtac ctttagcttt	660
	ttaatttgta aaggggttaa taaggaatat ttgatgtata gtcccttgac tagagatcca	720
	ttttctgta ttgaggaaag ttigccaggt gggttaaagg agcatgattt taatccagaa	780
	gaagcagagg aaactaaaca agtgtcctgg aagcttgtaa cagagtatgc aatggaaaca	840

	aaatgtgatg atgtgttgtt attgcttggg atgtacttgg aatttcagta cagttttgaa	900
	atgtgttttaa aatgtattaa aaaagaacag cccagccact ataagtacca tgaaaagcat	960
	tatgcaaatg ctgctatatt tgctgacagc aaaaaccaa aaaccatatt ccaacaggct	1020
	gttgatactg ttttagctaa aaagcgggtt gatagcctac aattaactag agaacaatg	1080
	ttacaaaca gatttaata tcttttggat aggatggata taatgtttgg ttctacaggc	1140
	tctgctgaca tagaagaatg gatggctgga gttgcttggc tacactgttt gttgcccata	1200
	atggagtcag tgggtgtatga ctttttaaaa tgcattgtgt acaacattcc taaaaaaga	1260
	tactggctgt ttaaaggacc aattgatagt ggtaaaacta cattagcagc tgctttgctt	1320
	gaattatgtg gggggaaagc tttaaatgtt aatttgcctt tggacaggct gaactttgag	1380
	ctaggagtag ctattgacca gtttttagta gtttttgagg atgtaaaggg cactggaggg	1440
	gagtcacagag atttgccttc aggtcaggga attaataacc tggacaattt aagggtattat	1500
	ttggatggca gtgttaaggt aaacttagaa aagaacacc taaataaaag aactcaaata	1560
	tttccccctg gaatagtcac catgaatgag tacagtgtgc ctaaaacact gcaggccaga	1620
	tttgtaaaac aaatagattt taggcccata gattatttaa agcattgcct ggaacgcagt	1680
	gagtttttgt tagaaaagag aataattcaa agtggcattg ctttgccttct tatgttaatt	1740
	tgttacagac ctgtggctga gtttgcitaa agtattcaga gcagaattgt ggagtggaaa	1800
	gagagattgg acaaagagtt tagtttgcitaa gtgtatcaaa aaatgaagtt taatgtggct	1860
	atgggaattg gagttttaga ttggctaaga aacagtgatg atgatgatga agacagccag	1920
	gaaaatgctg ataaaaatga agatgggtggg gagaagaaca tggaagactc agggcatgaa	1980
[0006]	acaggcattg attcacagtc ccaaggctca tttcaggccc ctacgtcctc acagtctgtt	2040
	catgatcata atcagccata ccacatttgt agagggttta cttgctttaa aaaacctccc	2100
	acacctcccc ctgaacctga acctgaaaca taa	2133
	<210> 10	
	<211> 599	
	<212> DNA	
	<213> 脑心肌炎病毒	
	<400> 10	
	cctctagcgg gatcaattcc gcccccccc cctaacttta ctggccgaag ccgcttggaa	60
	taaggccggt gtgcgtttgt ctatatgtta ttttccacca tattgccgtc ttttggcaat	120
	gtgagggccg ggaacactgg cctgtcttc ttgacgagca ttcttagggg tctttccctt	180
	ctcgccaaag gaatgcaagg tctgttgaat gtcgtgaagg aagcagttcc tctggaagct	240
	tcttgaagac aaacaacgtc tgtagcgacc ctttgcaggc agcggaaccc cccacctggc	300
	gacaggtgcc tctgcggcca aaagccactg gtataagata cacctgcaaa ggcggcacaa	360
	ccccagtgcc acgttgtgag ttggatagtt gtggaaagag tcaaatggct ctctcaagc	420
	gtattcaaca aggggctgaa ggatgccagc aaggtacccc attgtatggg atctgatctg	480
	gggcctcggt gcacatgctt tacgtgtgtt tagtcgaggt taaaaaacgt ctaggccccc	540
	cgaaccacgg ggacgtggtt ttcctttgaa aaacacgata ataccatgat tgaacaaga	599
	<210> 11	
	<211> 30	
	<212> DNA	

[0007]	<213> 人工的		
	<220>		
	<223> 引物		
	<400> 11		
	cgacgcgtgt ctgtaagctg aagacctggc	30	
	<210> 12		
	<211> 31		
	<212> DNA		
	<213> 人工的		
	<220>		
	<223> 引物		
	<400> 12		
	aaaagtactc ctgccctgcc tctgctggct c	31	
	<210> 13		
	<211> 21		
	<212> DNA		
	<213> 人工的		
	<220>		
	<223> 引物		
	<400> 13		
	gcctggtcaa catggtgaaa c	21	
	<210> 14		
	<211> 31		
	<212> DNA		
	<213> 人工的		
	<220>		
	<223> 引物		
	<400> 14		
	tgggctagcc caagaagctg gttctgagag g	31	
	<210> 15		
	<211> 37		
	<212> DNA		
	<213> 人工的		
	<220>		
	<223> 引物		
	<400> 15		
	ggggtaccag atggataaag ttttaaacag agaggaa	37	
	<210> 16		
	<211> 37		
	<212> DNA		
	<213> 人工的		
	<220>		
	<223> 引物		
	<400> 16		
	gggaattctt atgtttcagg ttcaggttca gggggag	37	
	<210> 17		
	<211> 25		
	<212> DNA		
	<213> 人工的		
	<220>		

[0008]	<223> 引物	
	<400> 17	
	tcagagggga caaaaggga gttgg	25
	<210> 18	
	<211> 23	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 引物	
	<400> 18	
	agaggcccct gccctccaga ggt	23
	<210> 19	
	<211> 23	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 引物	
	<400> 19	
	gcaaattcct cagggtcag ata	23
	<210> 20	
	<211> 25	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 引物	
	<400> 20	
	ccaggcctcc gaaggacgac gtggc	25
	<210> 21	
	<211> 21	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 引物	
	<400> 21	
	taggtgtcc ccagtgaag t	21
	<210> 22	
	<211> 22	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 引物	
	<400> 22	
	aatgccaagc ttacttagat cg	22
	<210> 23	
	<211> 23	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 引物	
	<400> 23	
	cttaacctta ctgccccag tct	23

	<210> 24	
	<211> 35	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 引物	
	<400> 24	
	aggctcgagc gcggggtag gctttgctgt ctgag	35
	<210> 25	
	<211> 21	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 引物	
	<400> 25	
	ggaagccaag tgtccttctg c	21
	<210> 26	
	<211> 20	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 引物	
	<400> 26	
	gggcagggtt ttttaccac	20
[0009]	<210> 27	
	<211> 21	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 引物	
	<400> 27	
	agcttcctgg gtttccgatt t	21
	<210> 28	
	<211> 24	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 引物	
	<400> 28	
	gccaggtact cacctgtatg gctg	24
	<210> 29	
	<211> 22	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 引物	
	<400> 29	
	aatgccaagc ttacttagat cg	22
	<210> 30	
	<211> 32	

	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 引物	
	<400> 30	
[0010]	ggggctagcc ttaaccttac tcgccccagt ct	32
	<210> 31	
	<211> 4	
	<212> DNA	
	<213> 智人 (Homo Sapiens)	
	<400> 31	
	ccgg	4

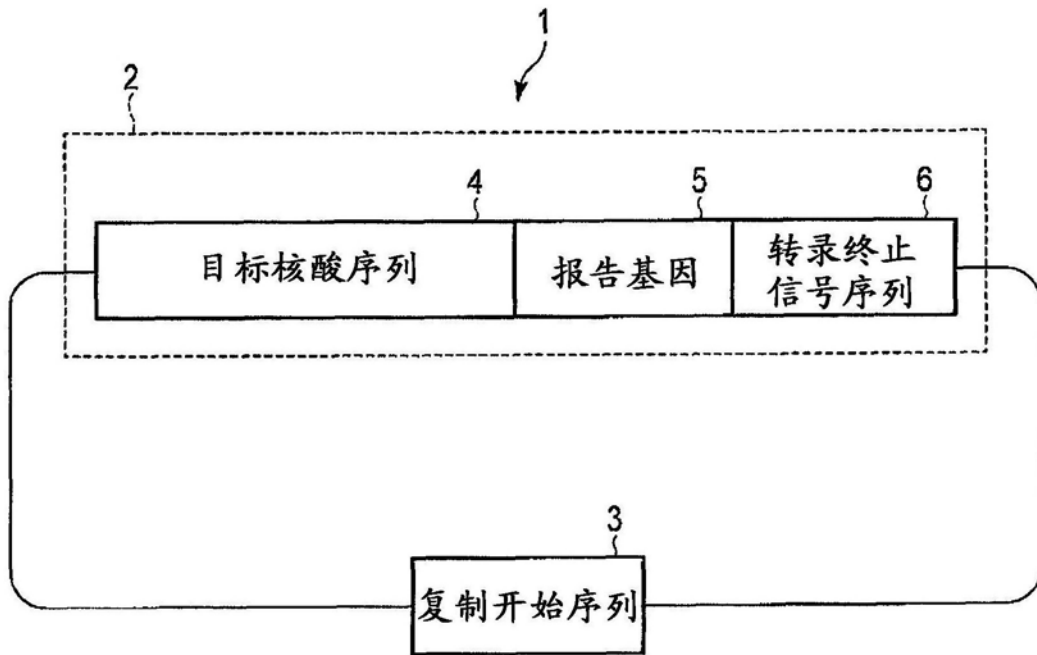


图1

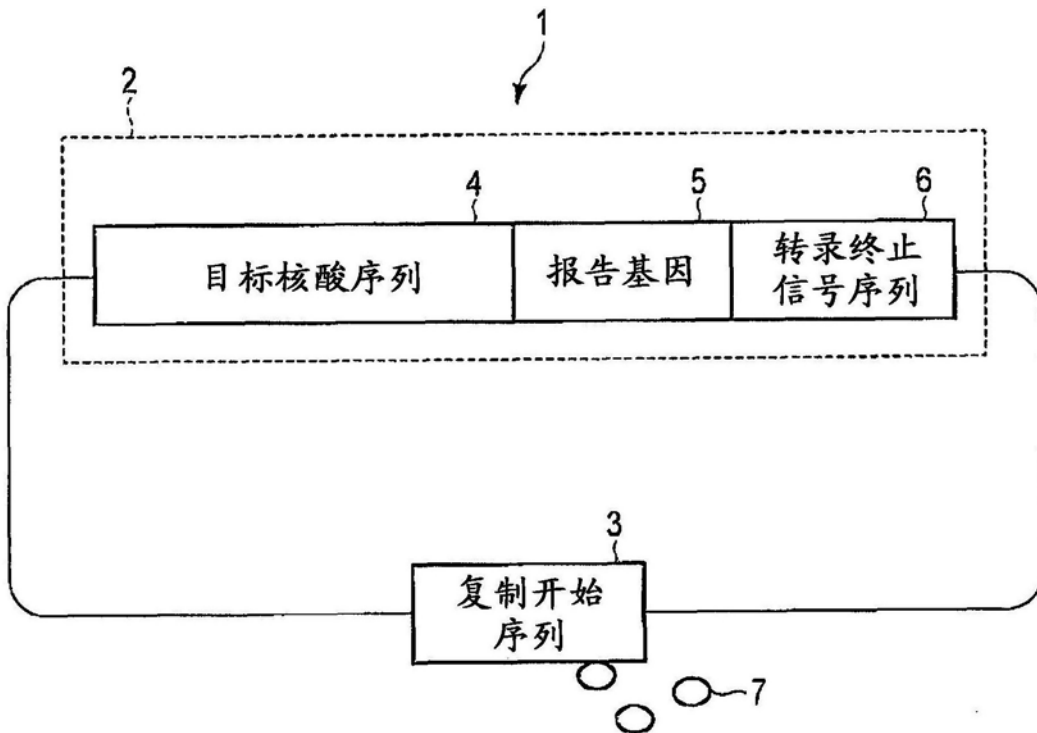


图2

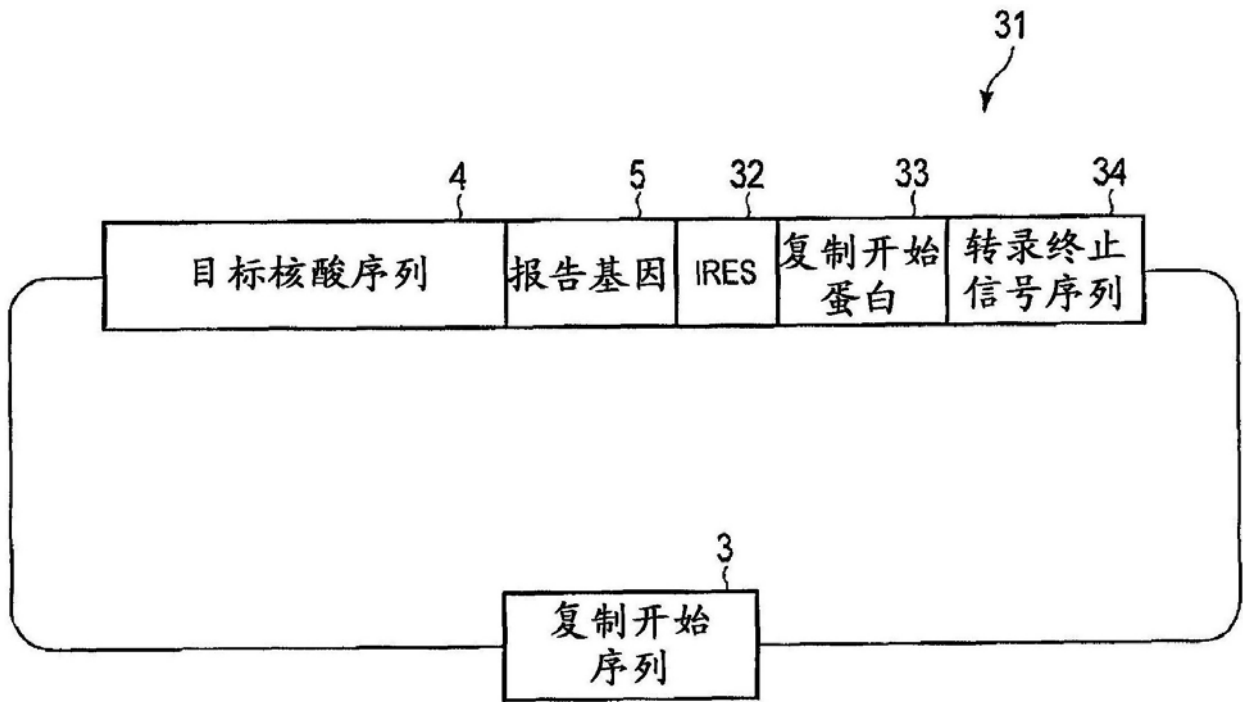


图3



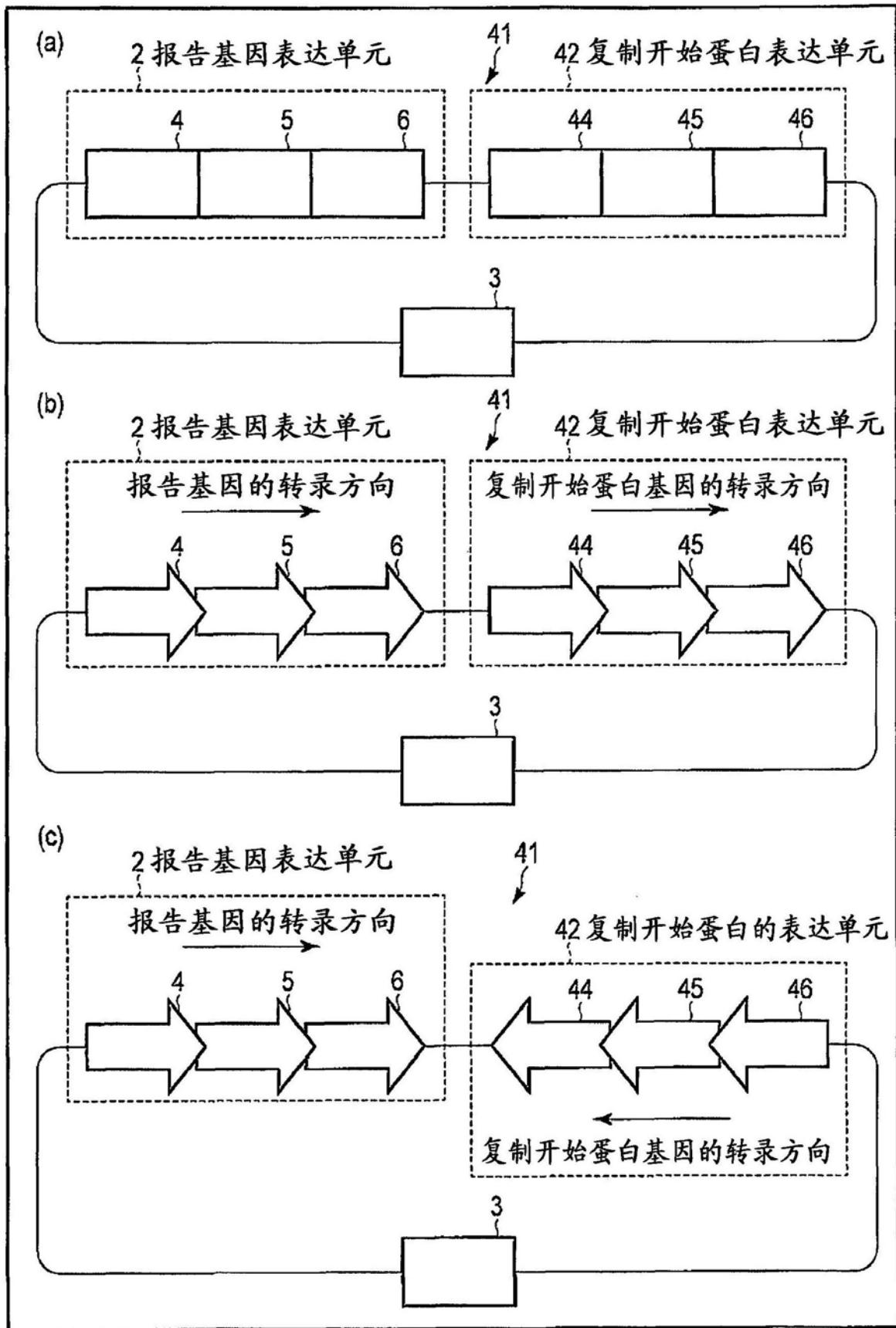


图4

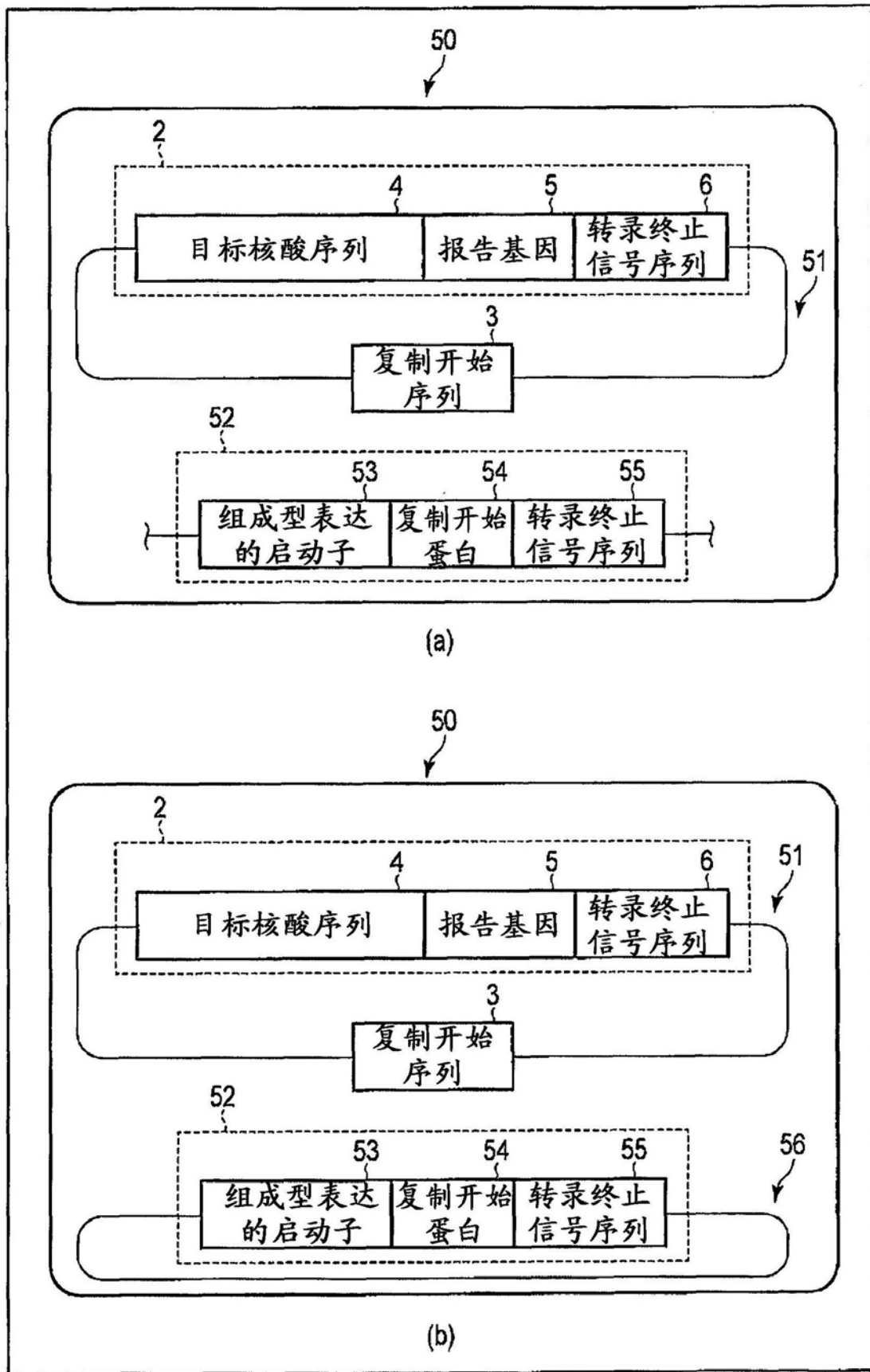


图5

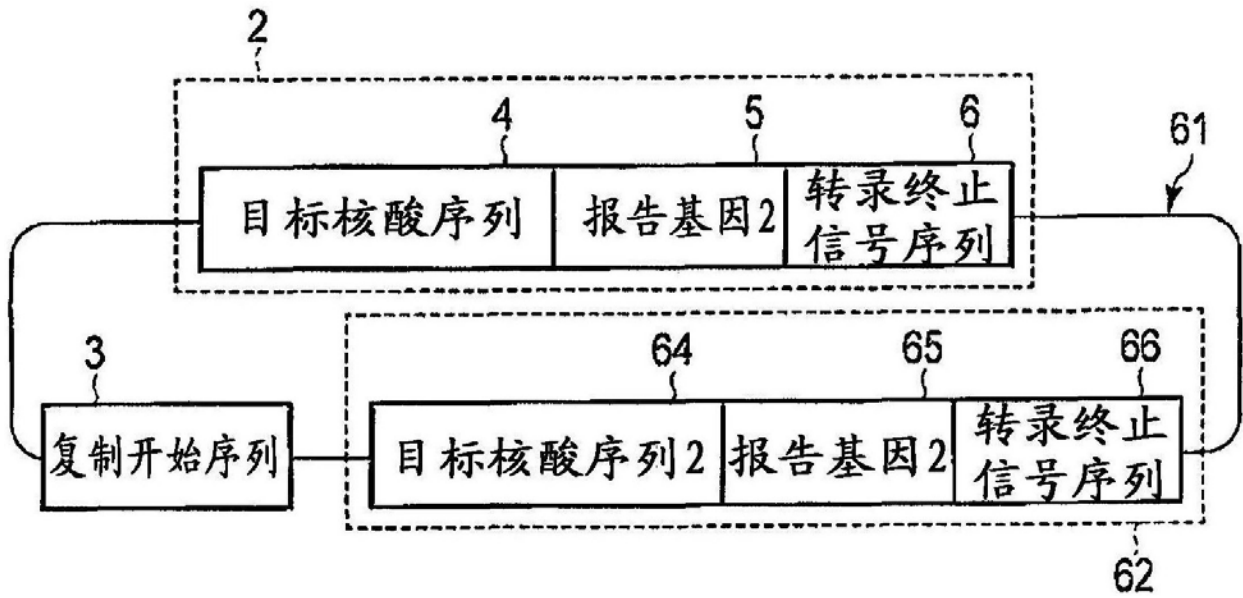


图6

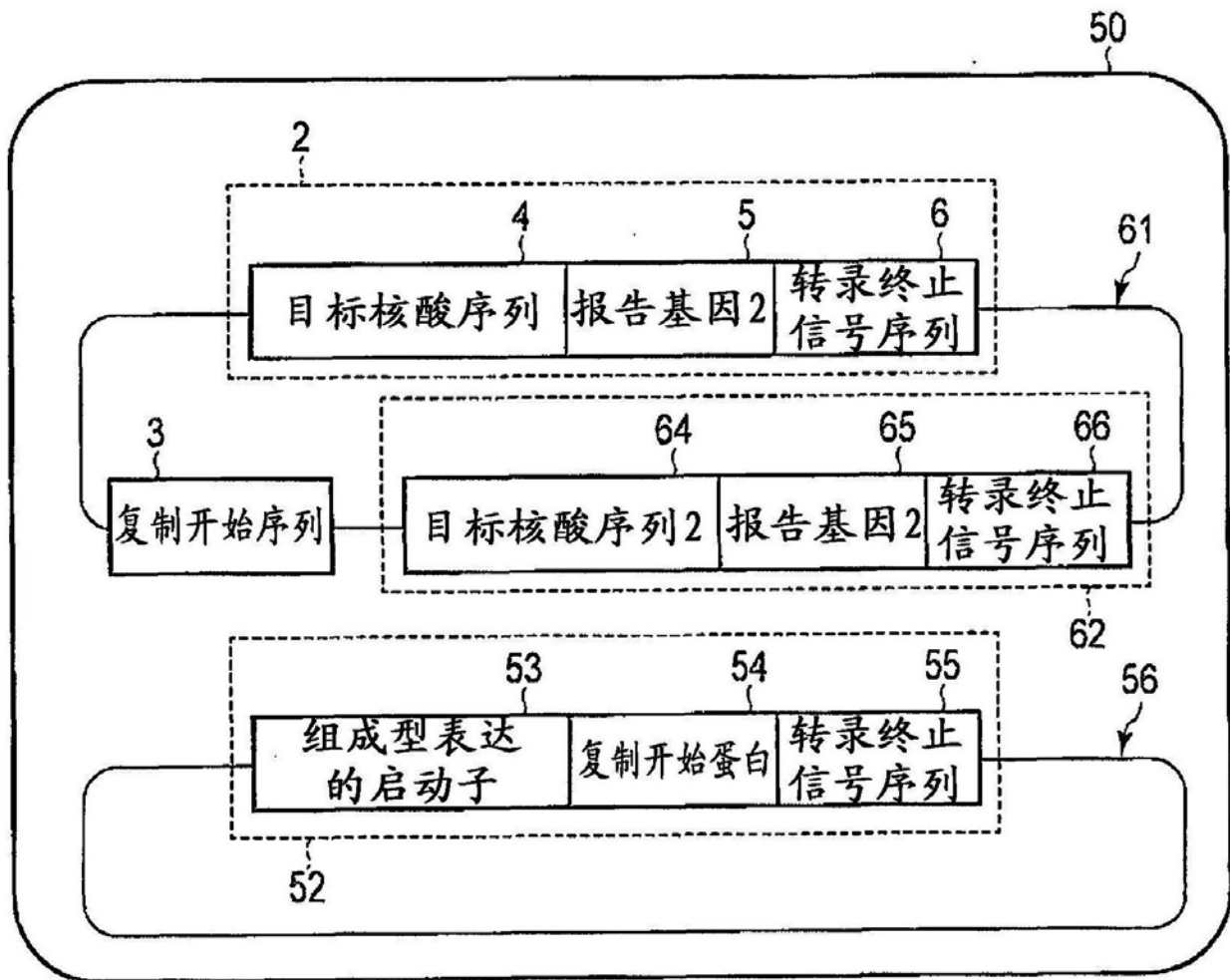


图7

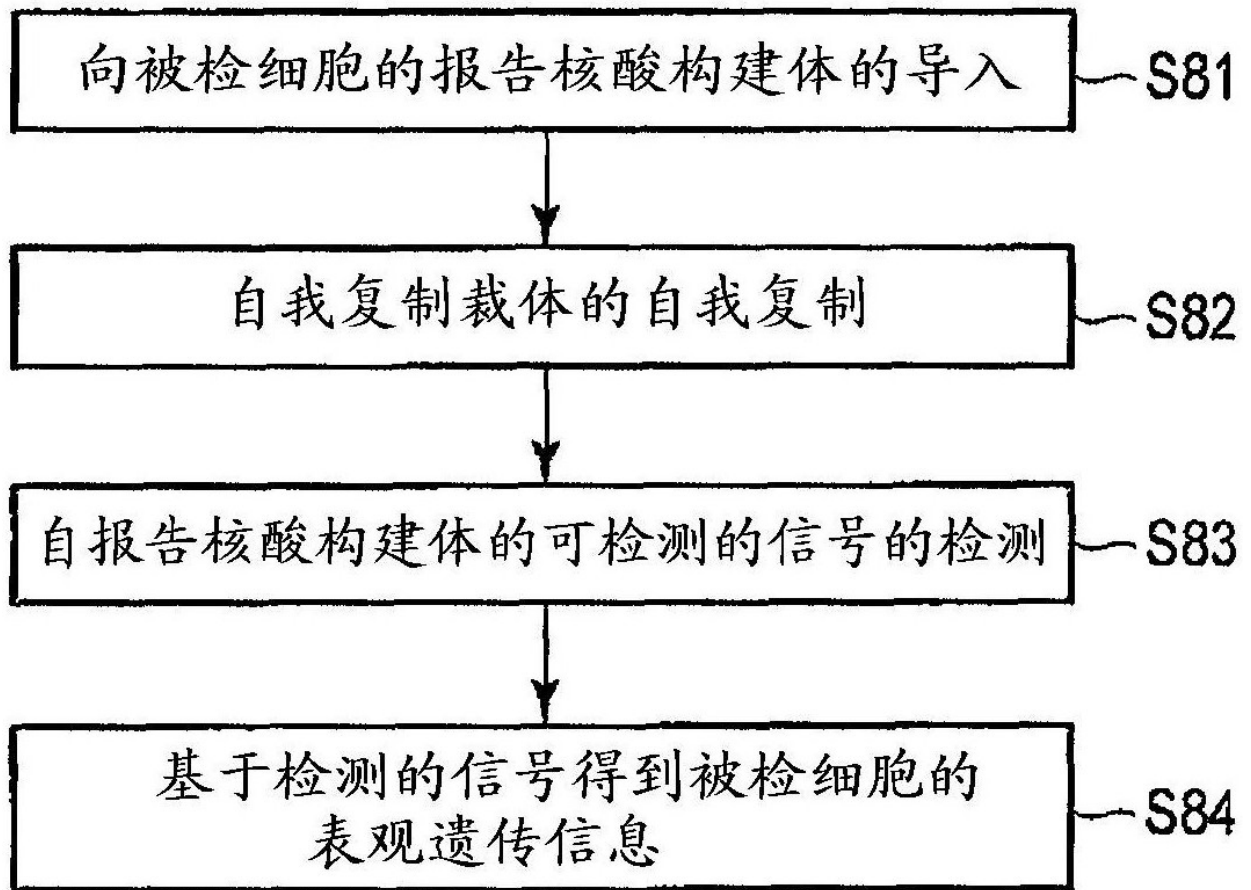


图8

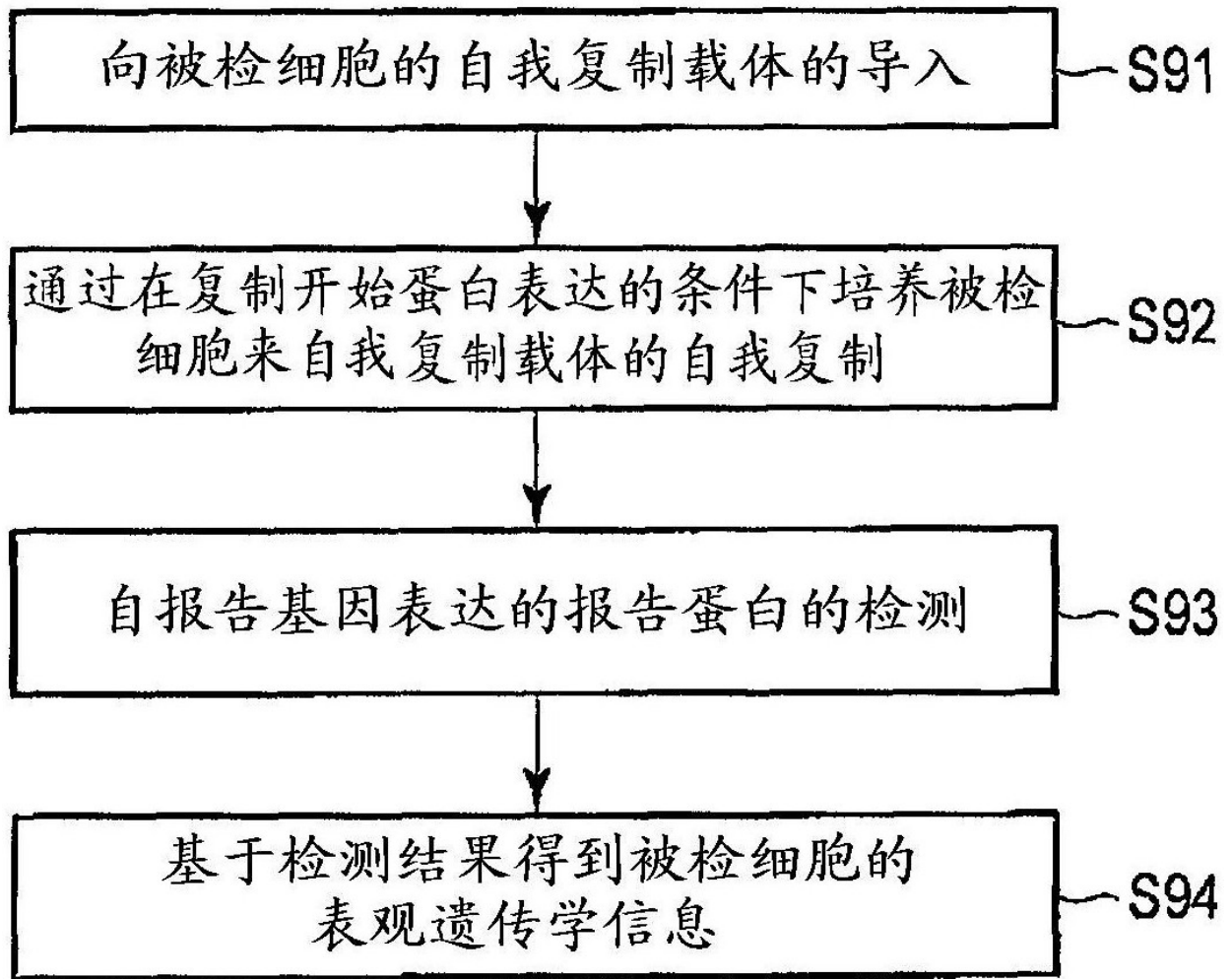


图9

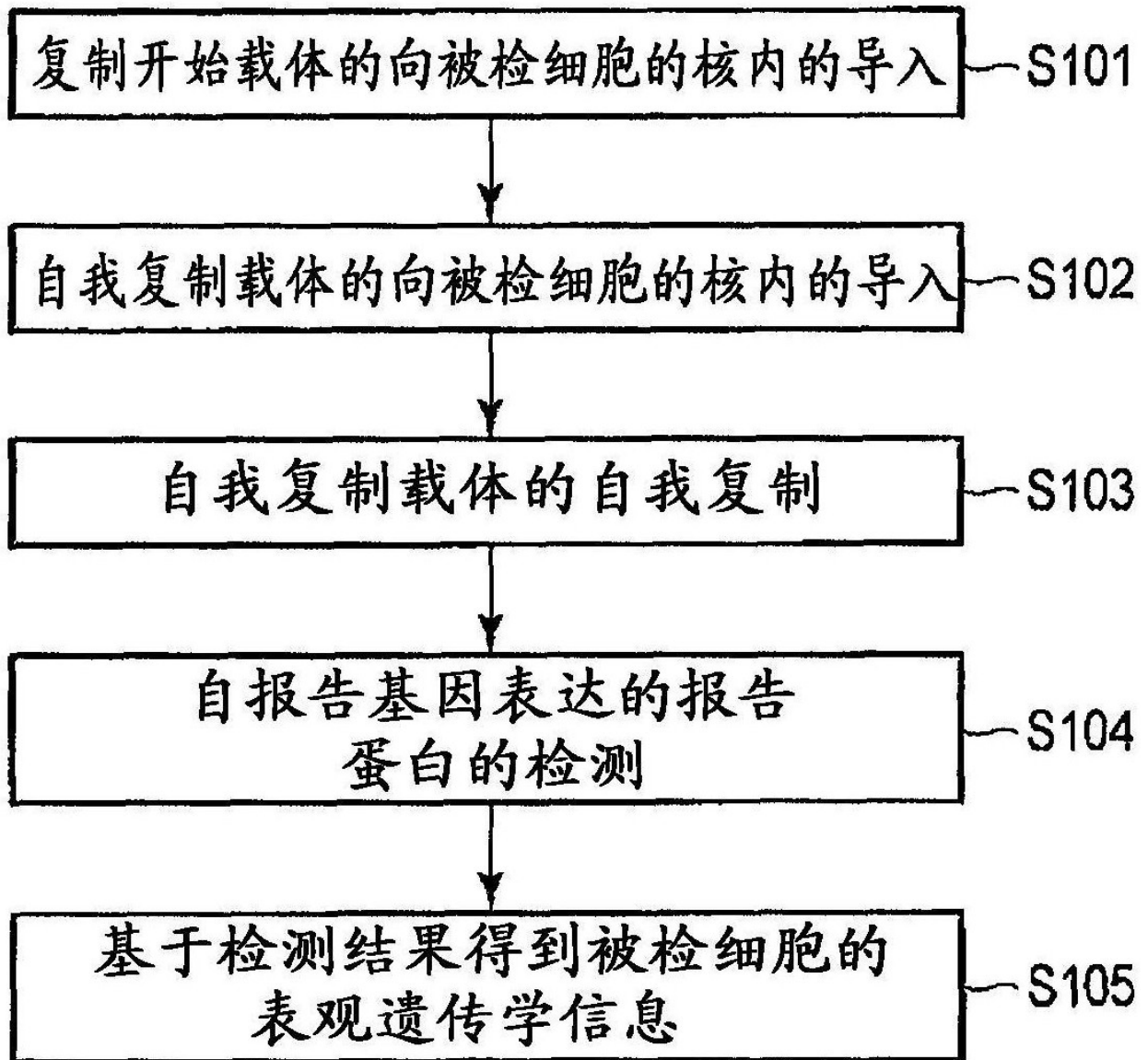


图10

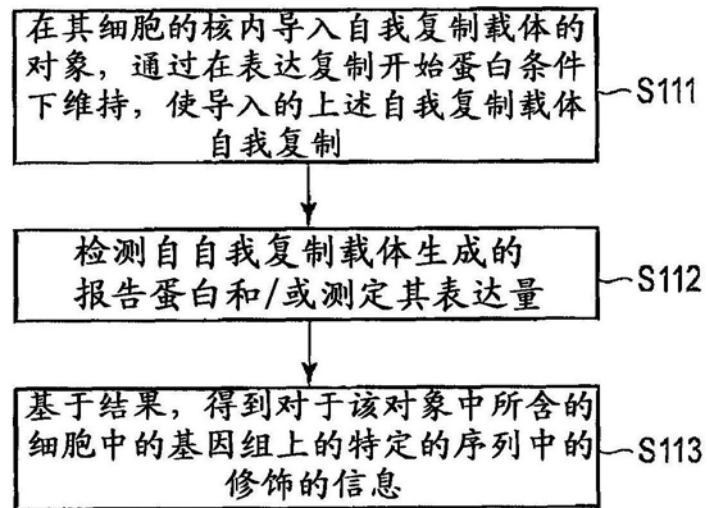


图11

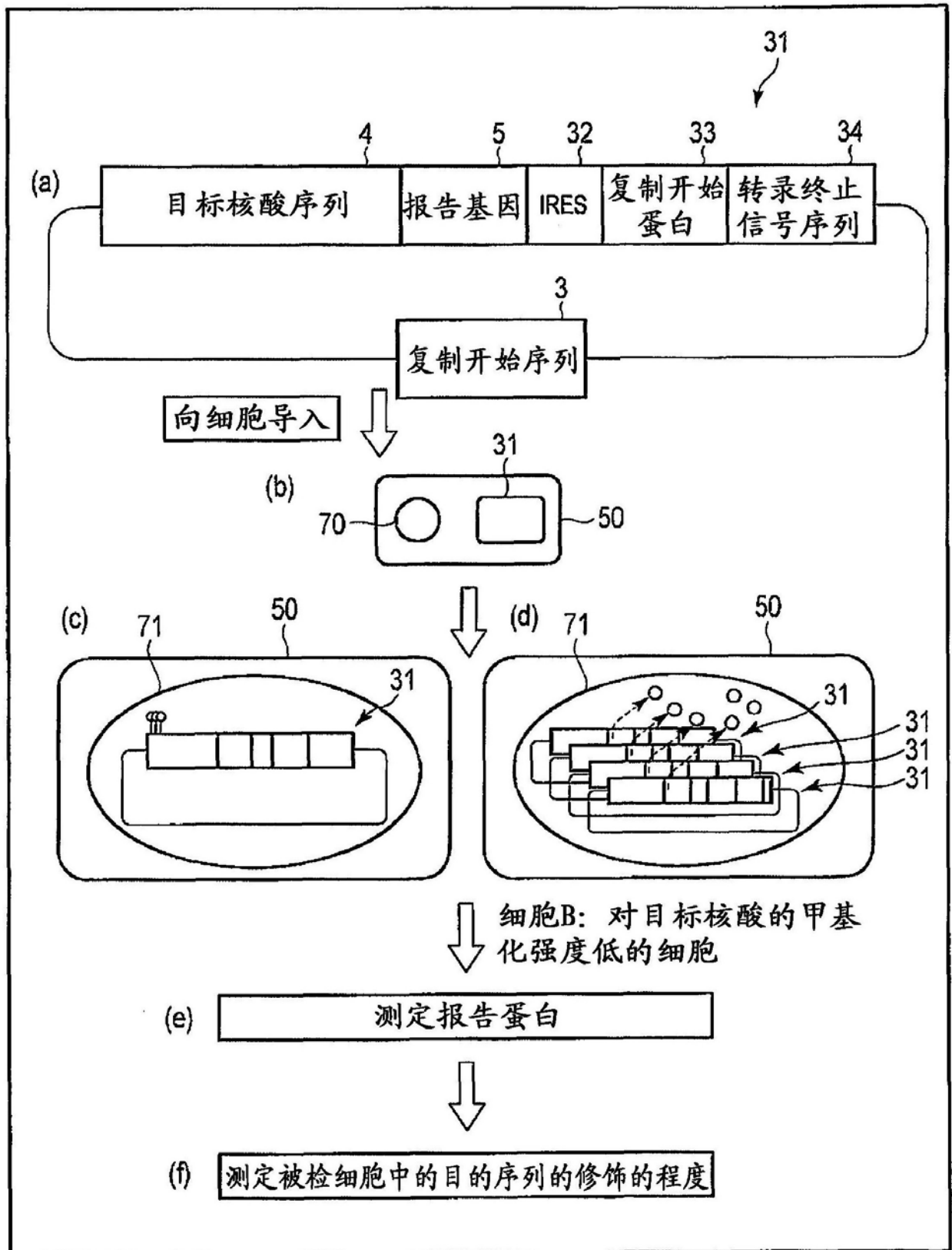


图12



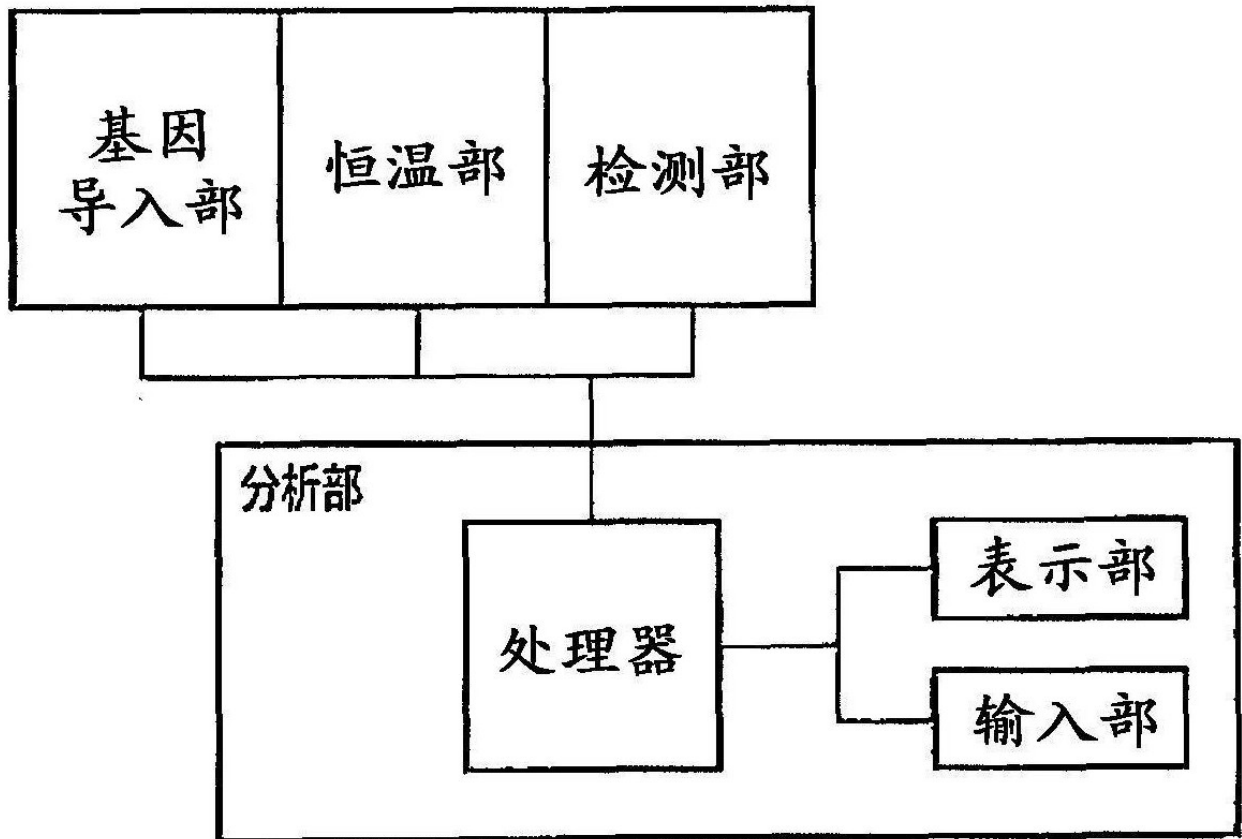


图13

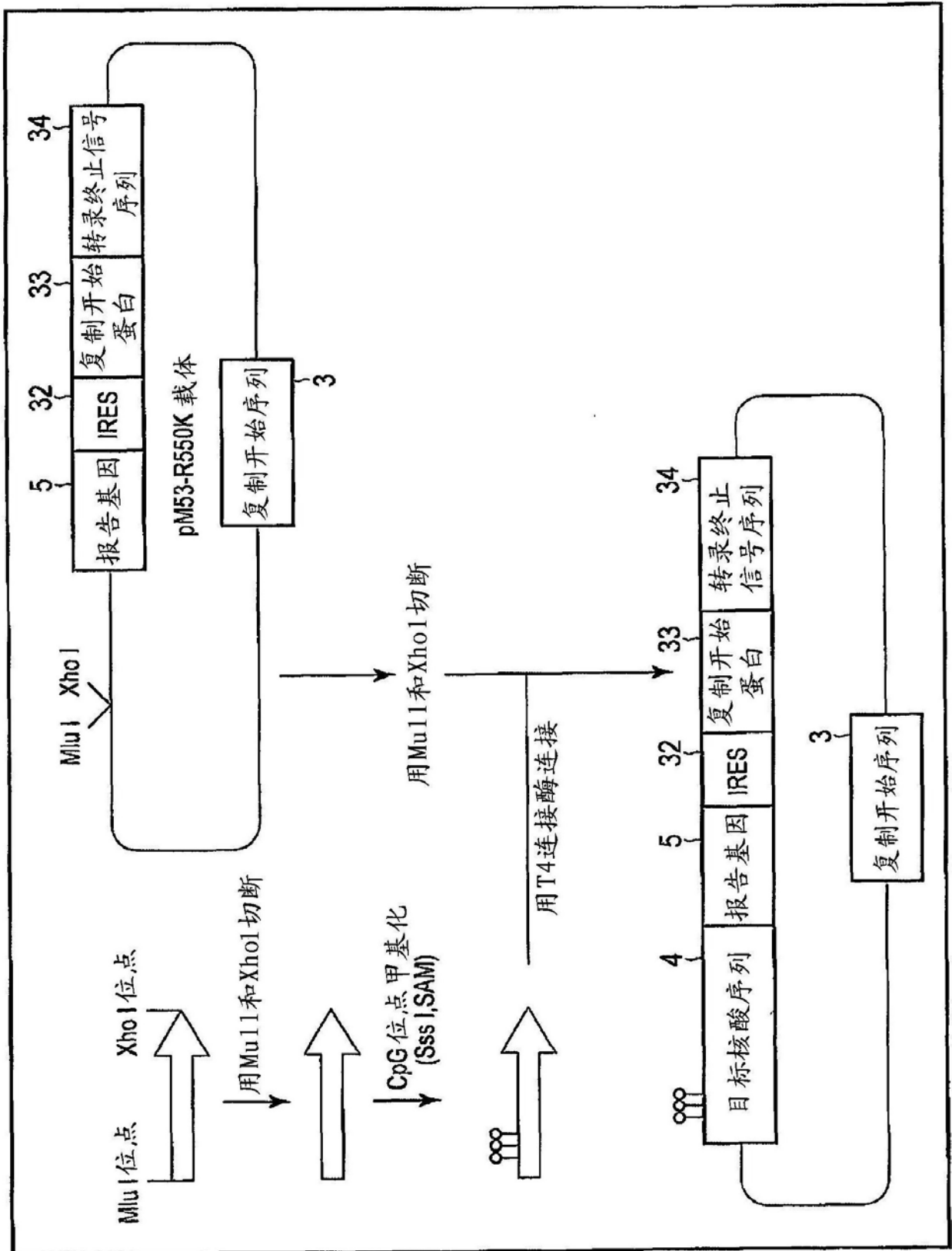


图14

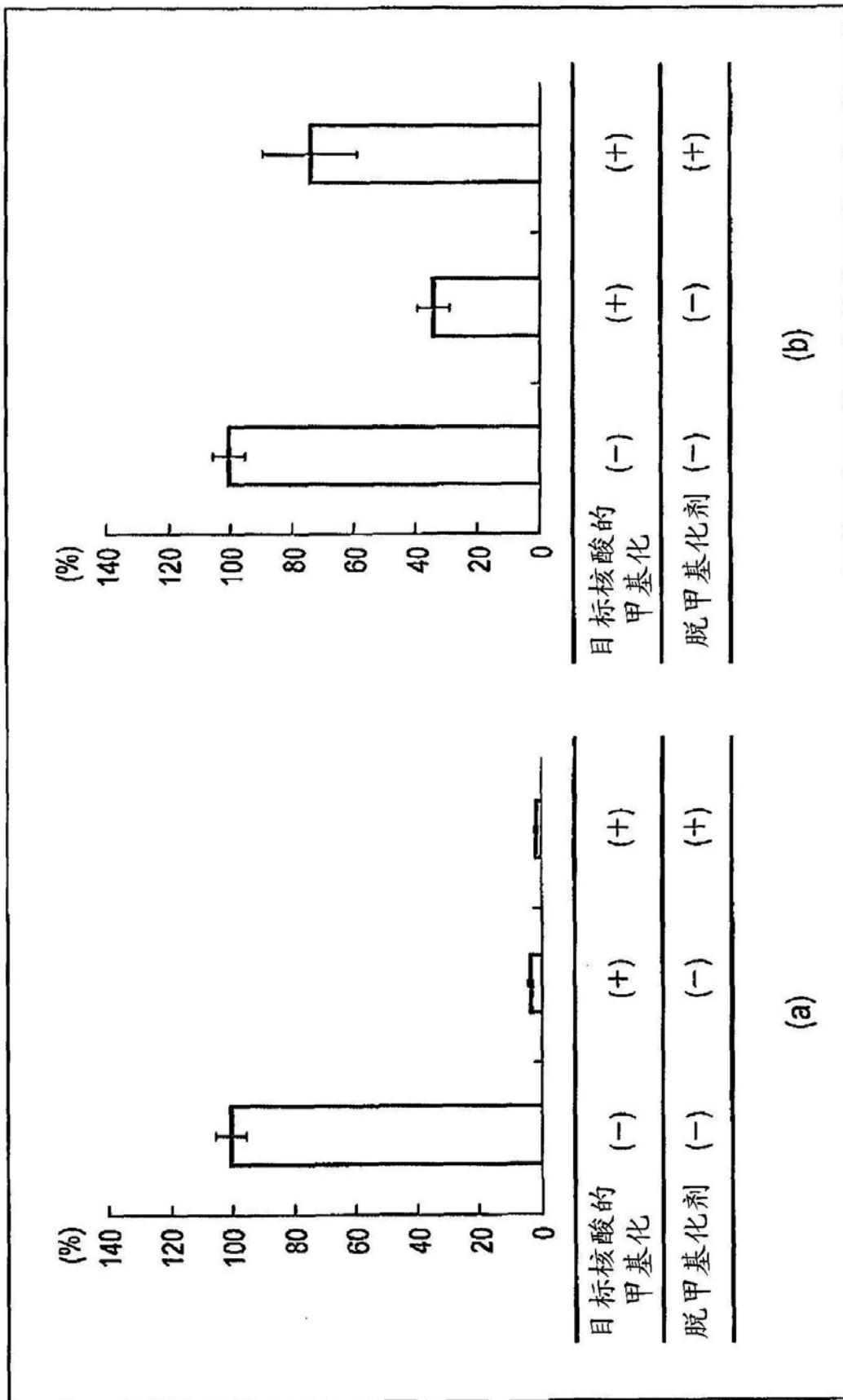


图15

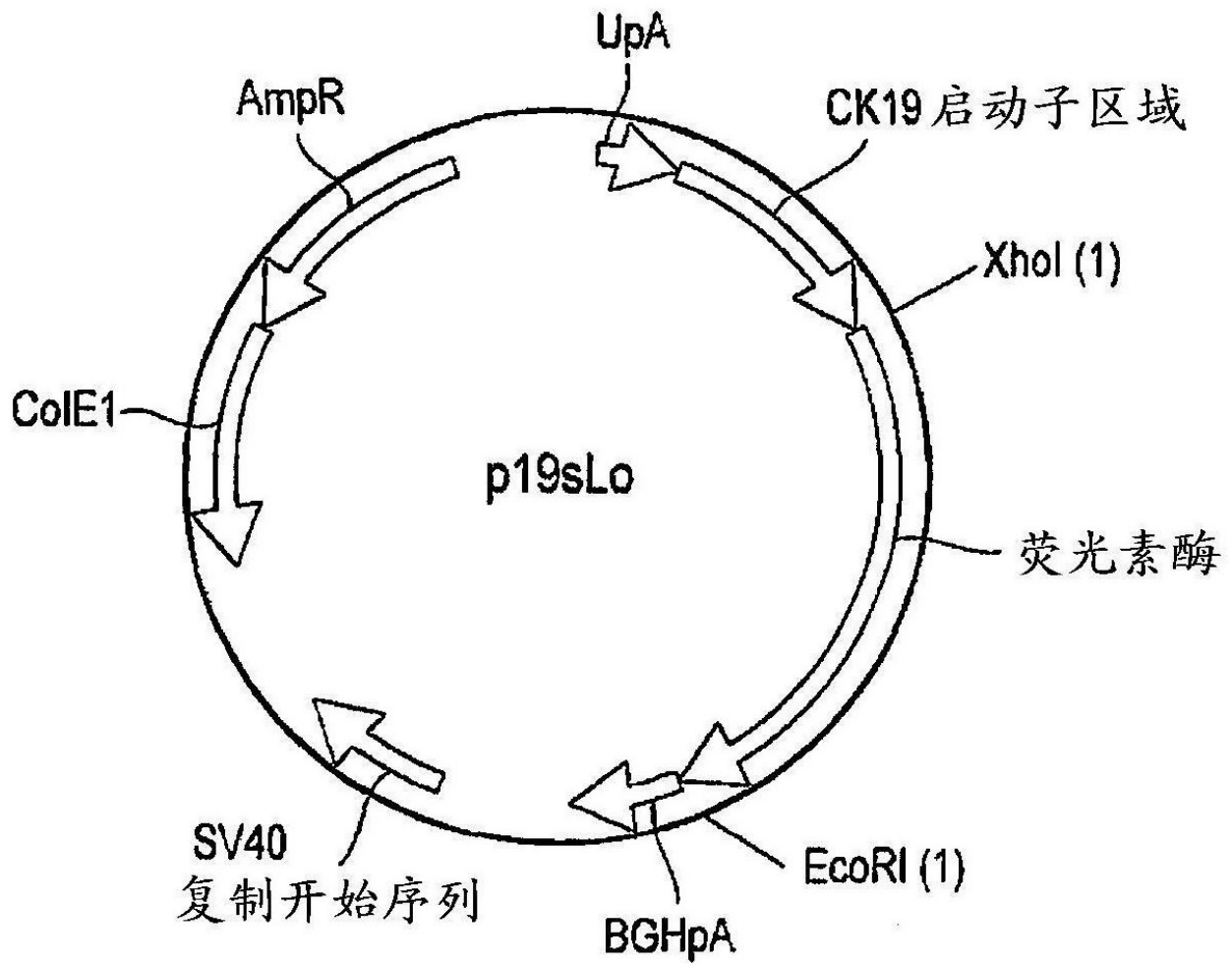


图16

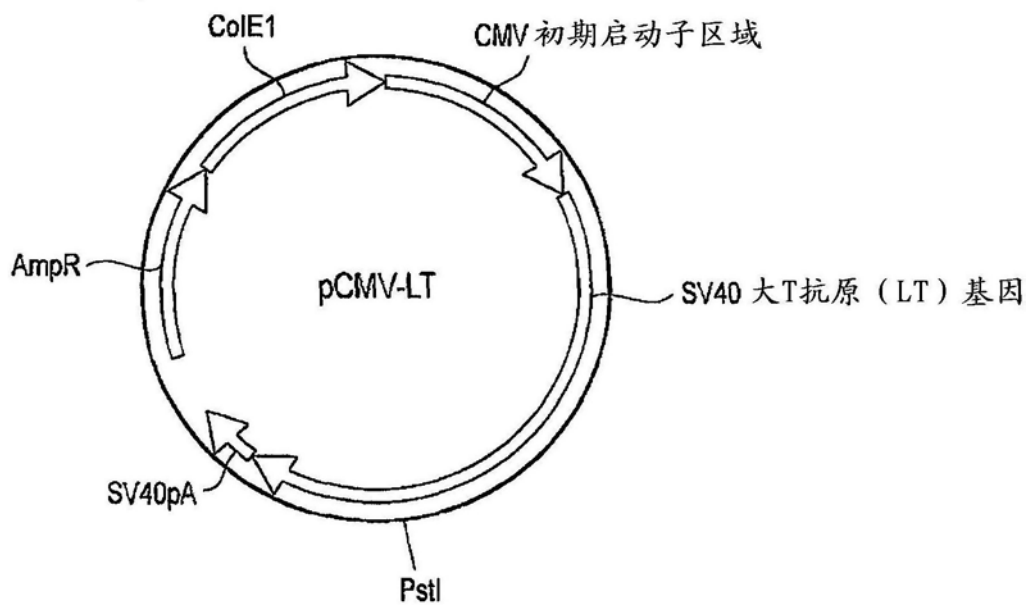


图17

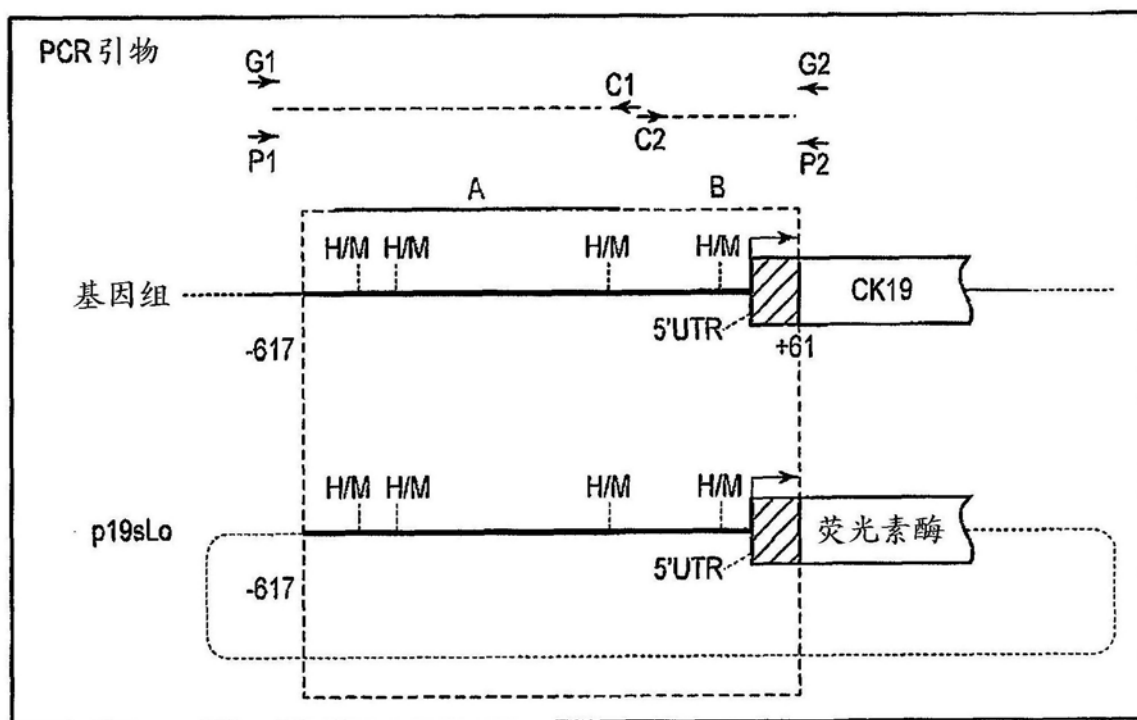


图18

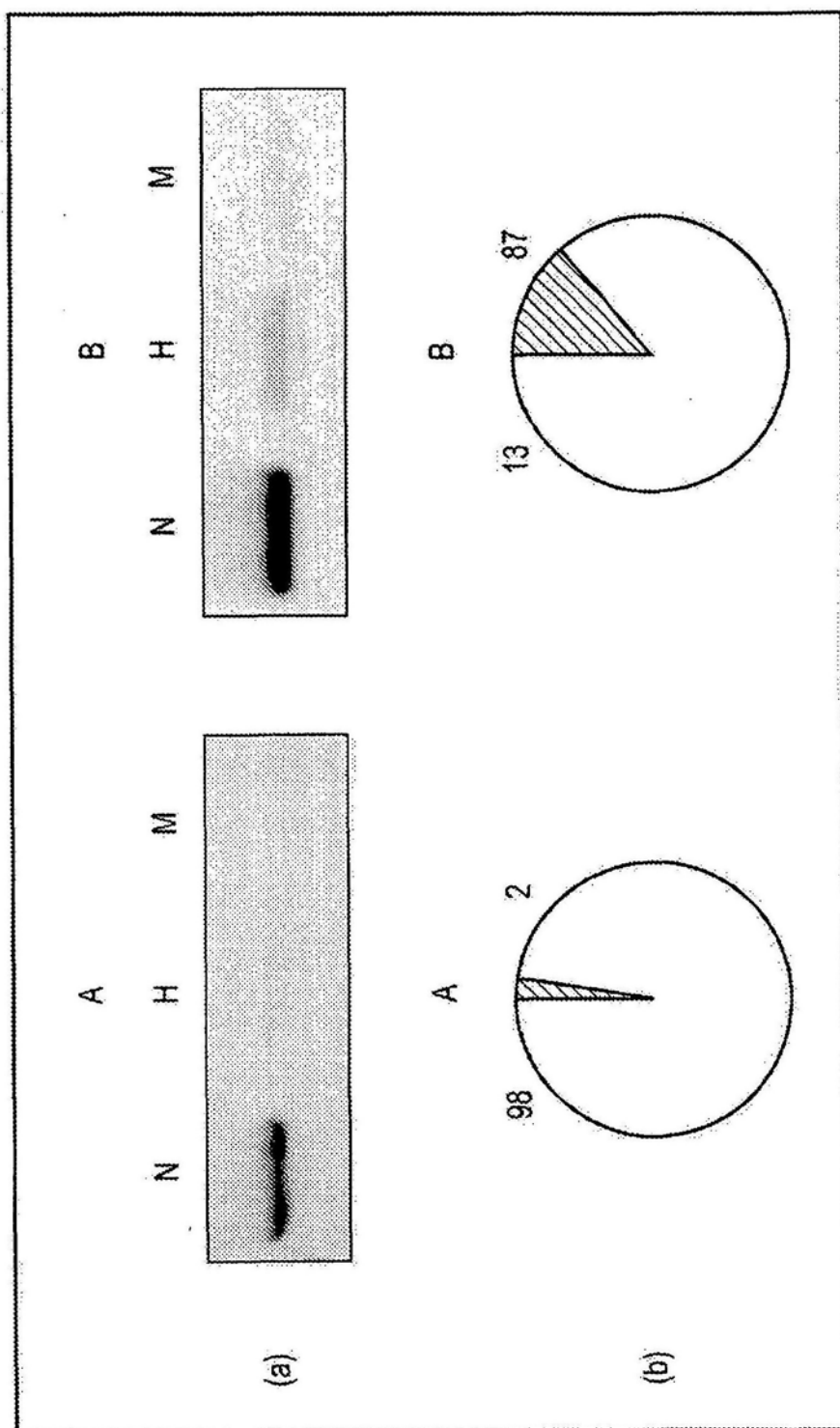


图19

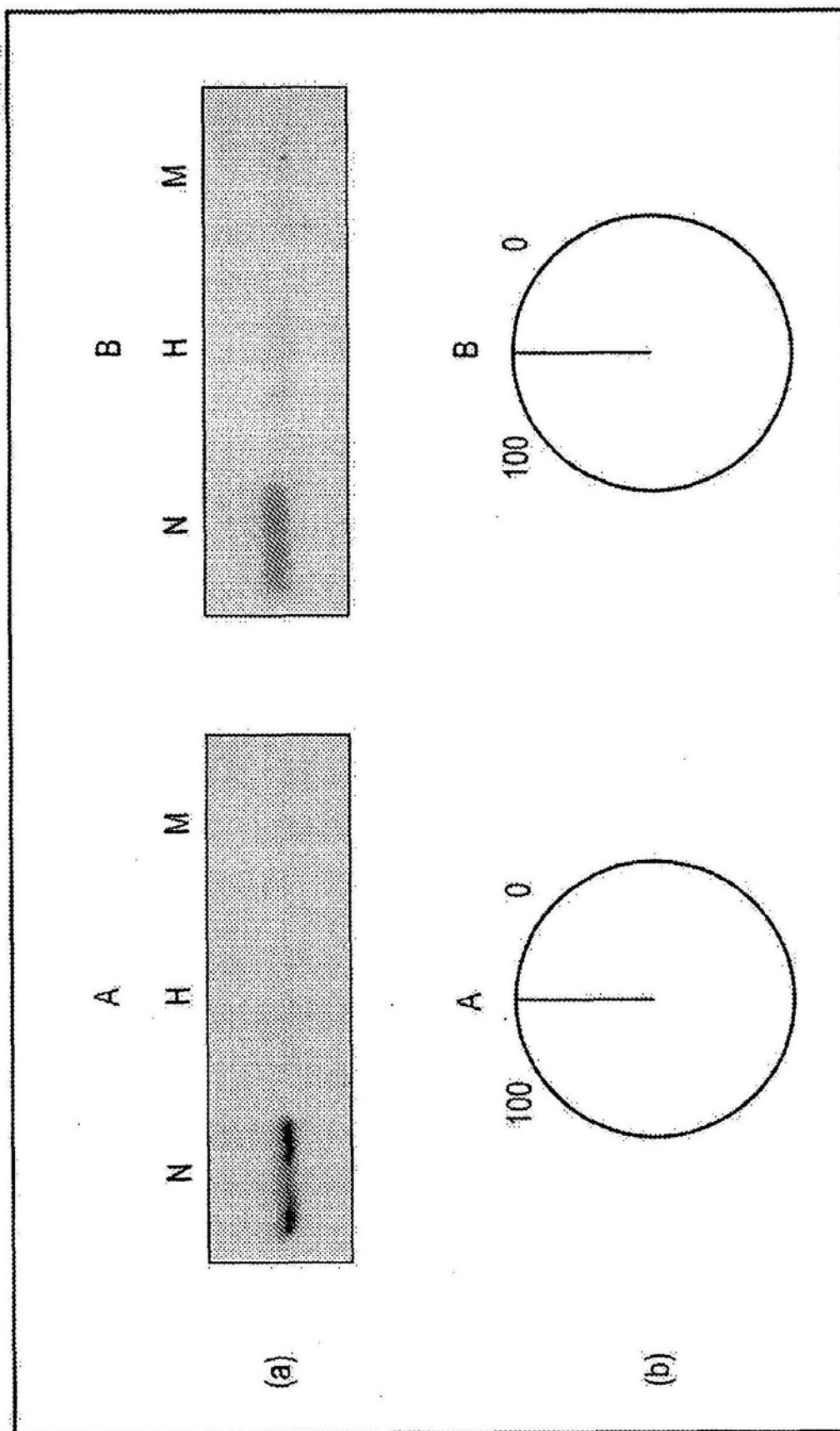


图20

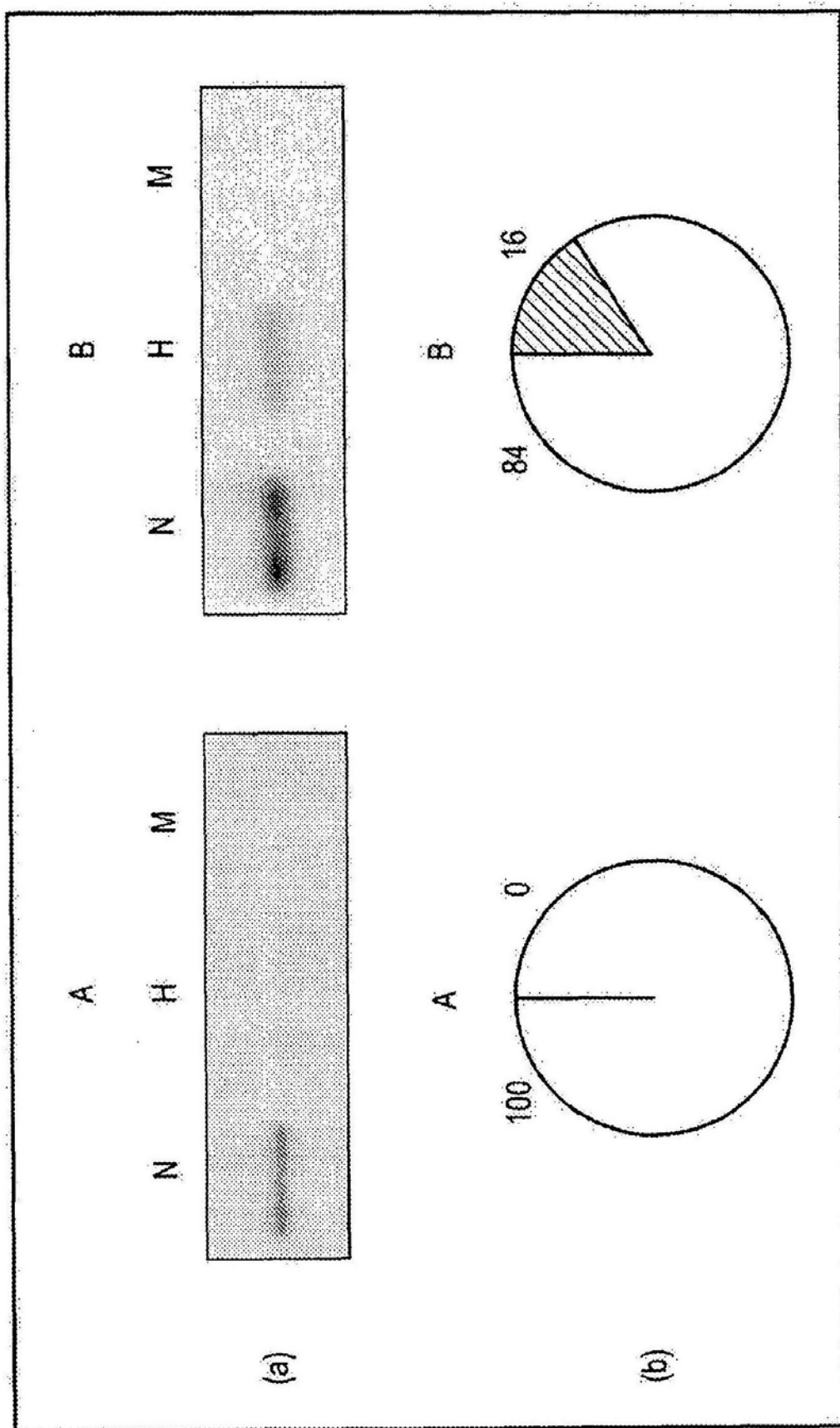


图21



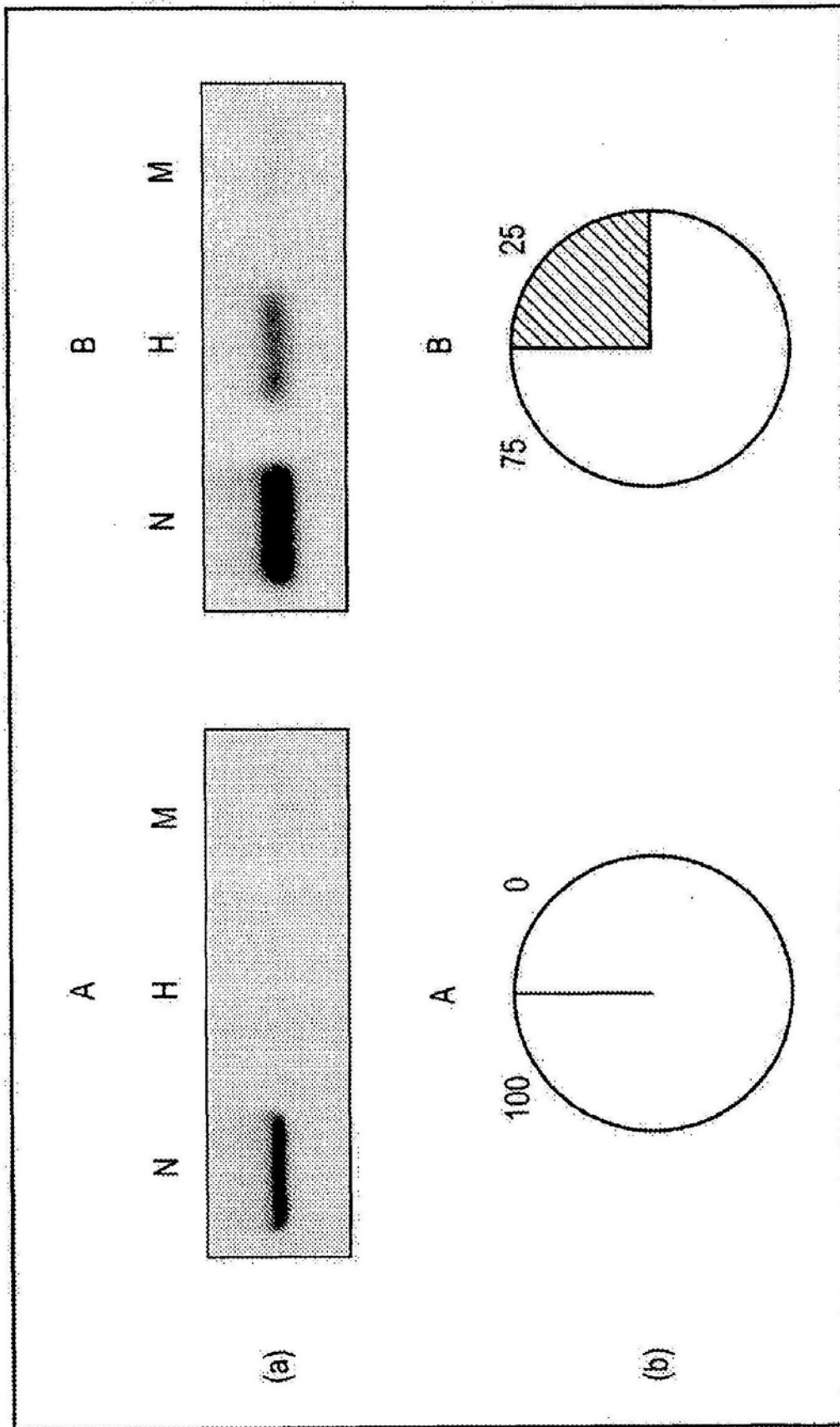


图22

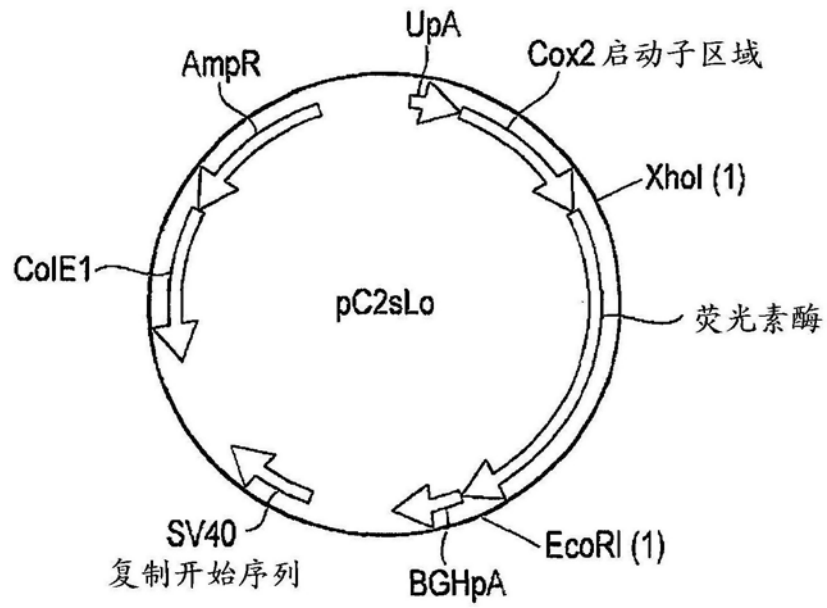


图23

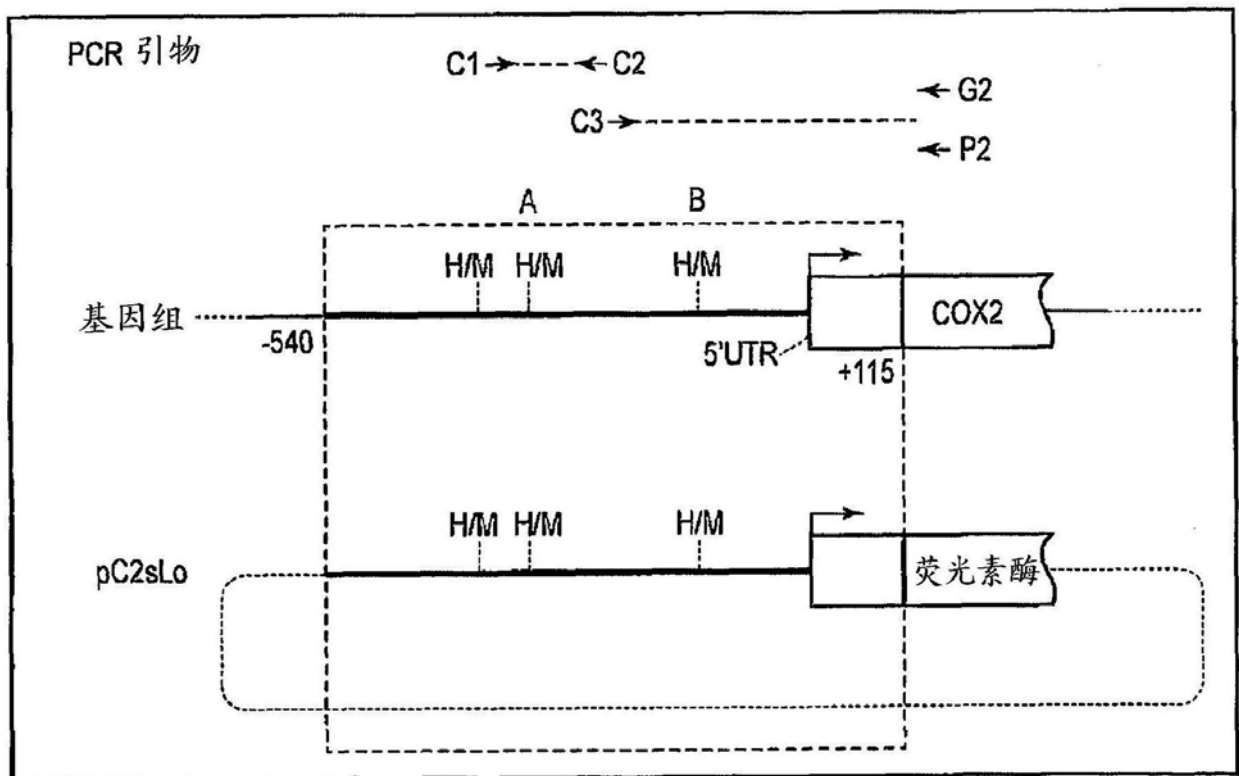


图24

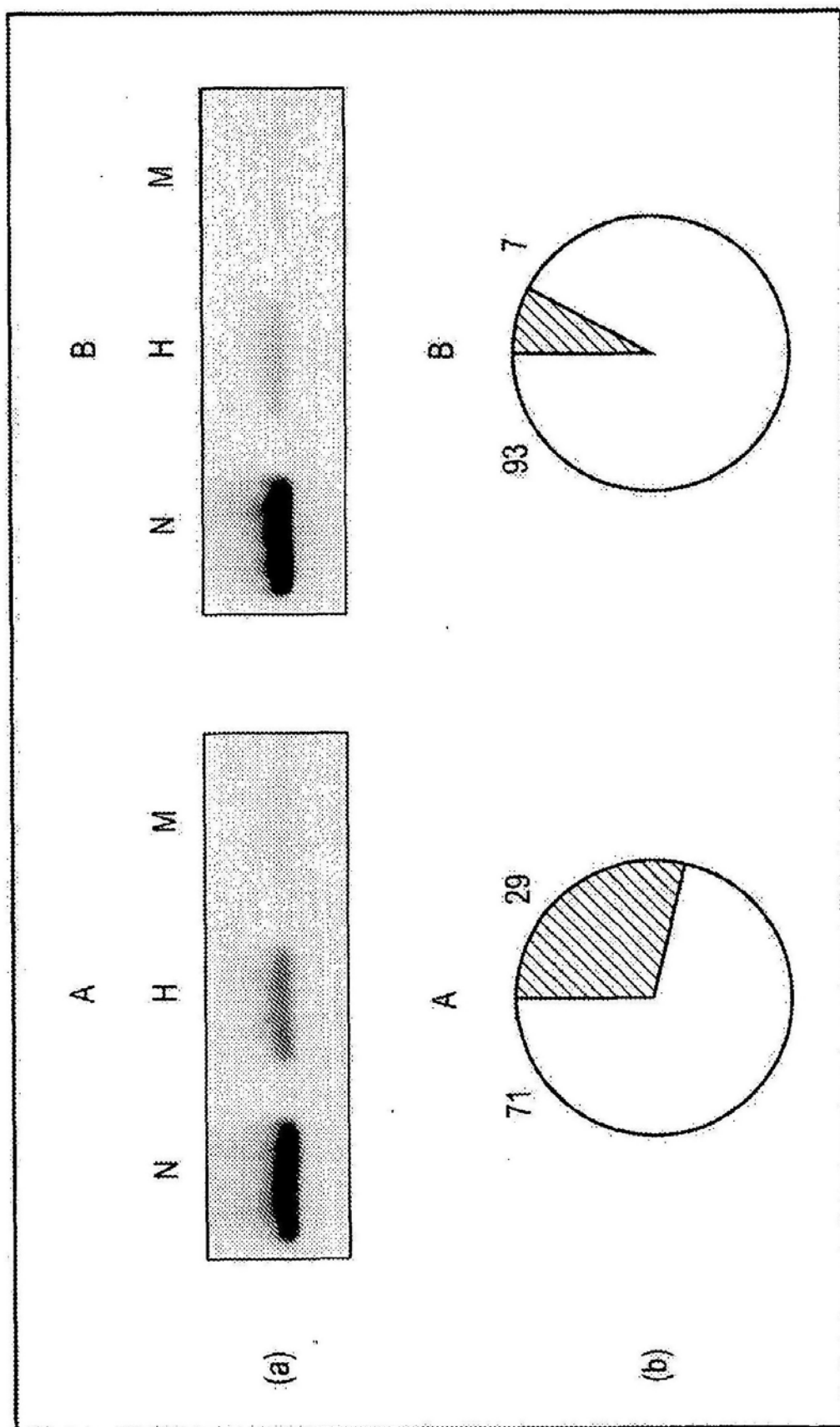


图25

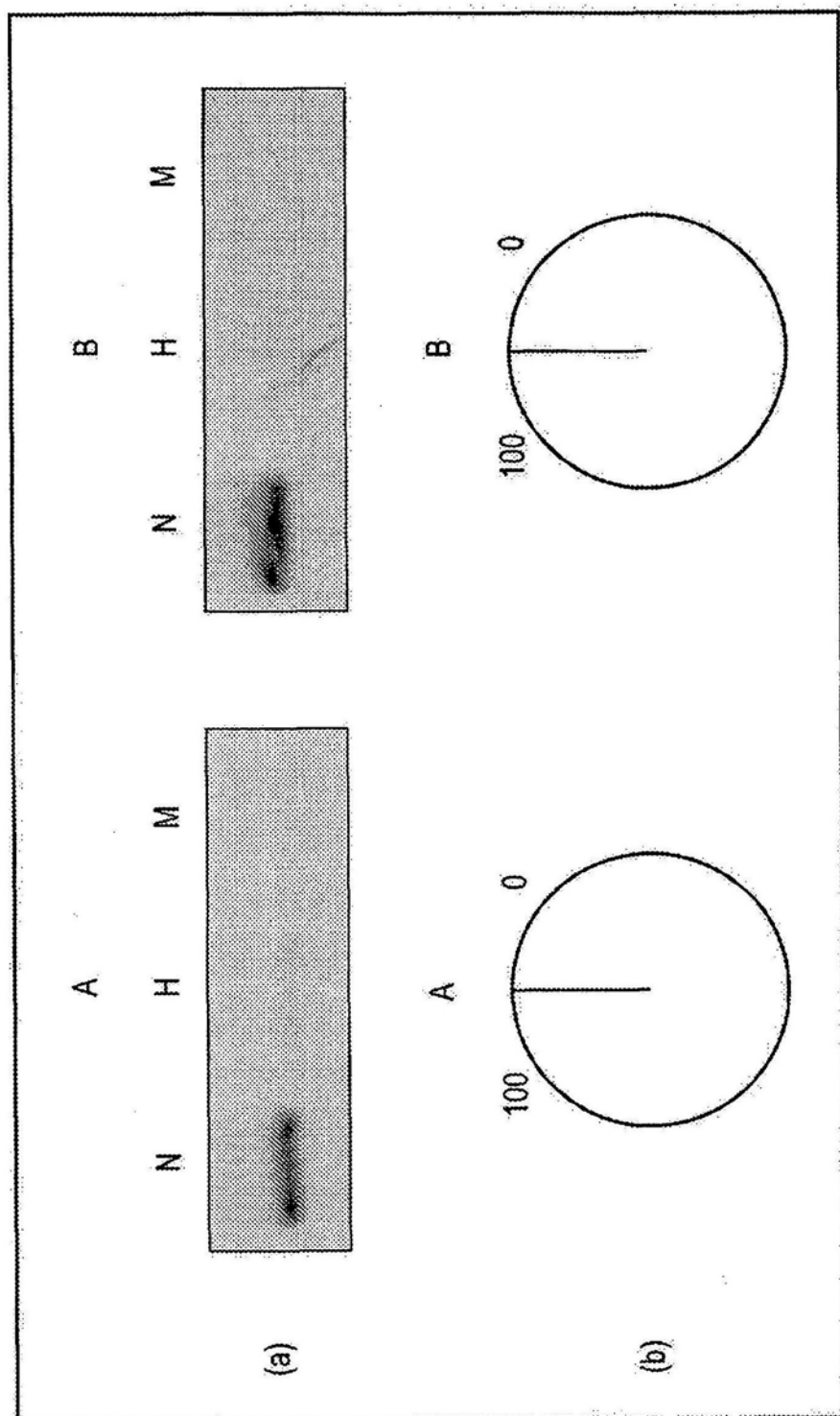


图26

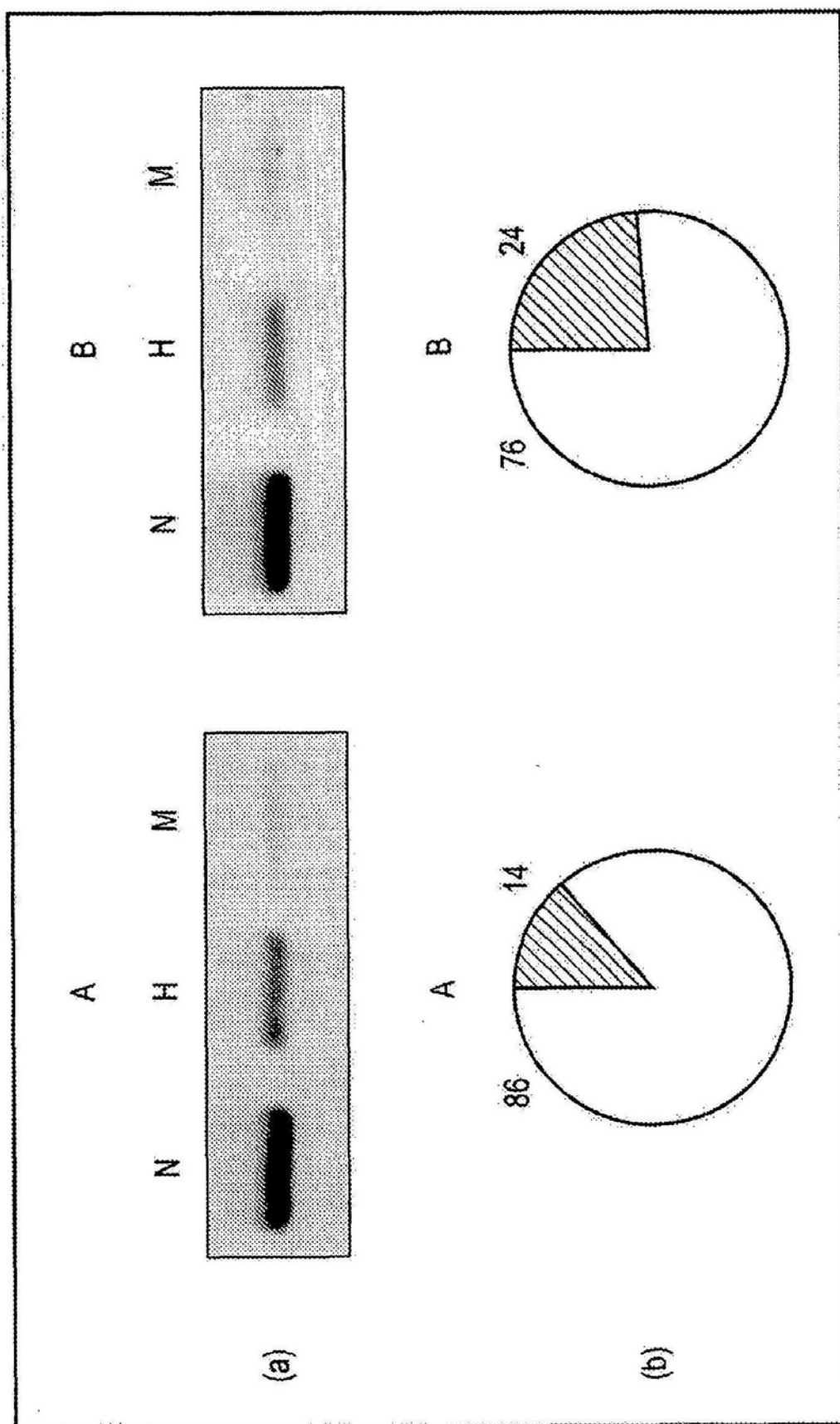


图27

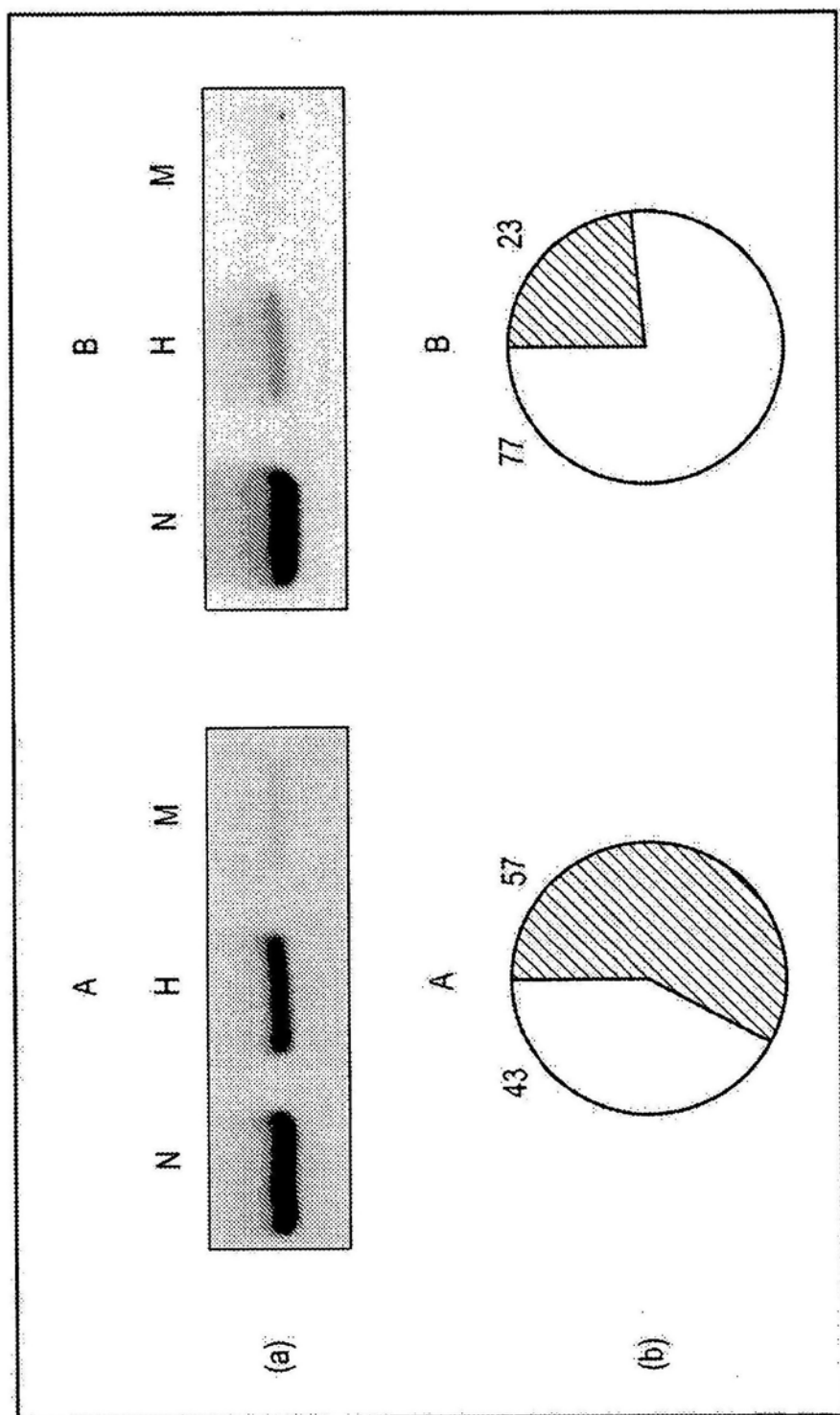


图28

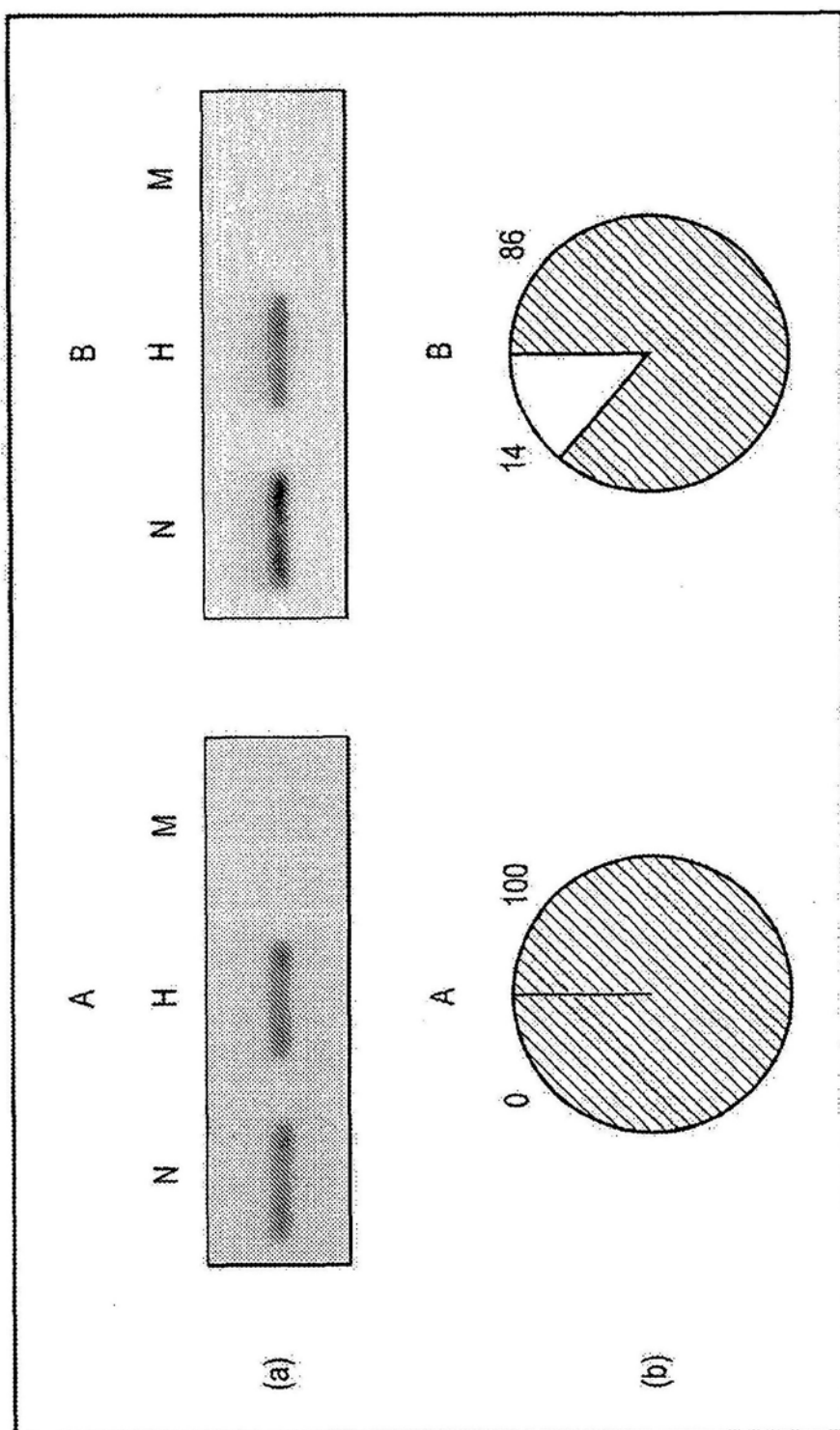


图29

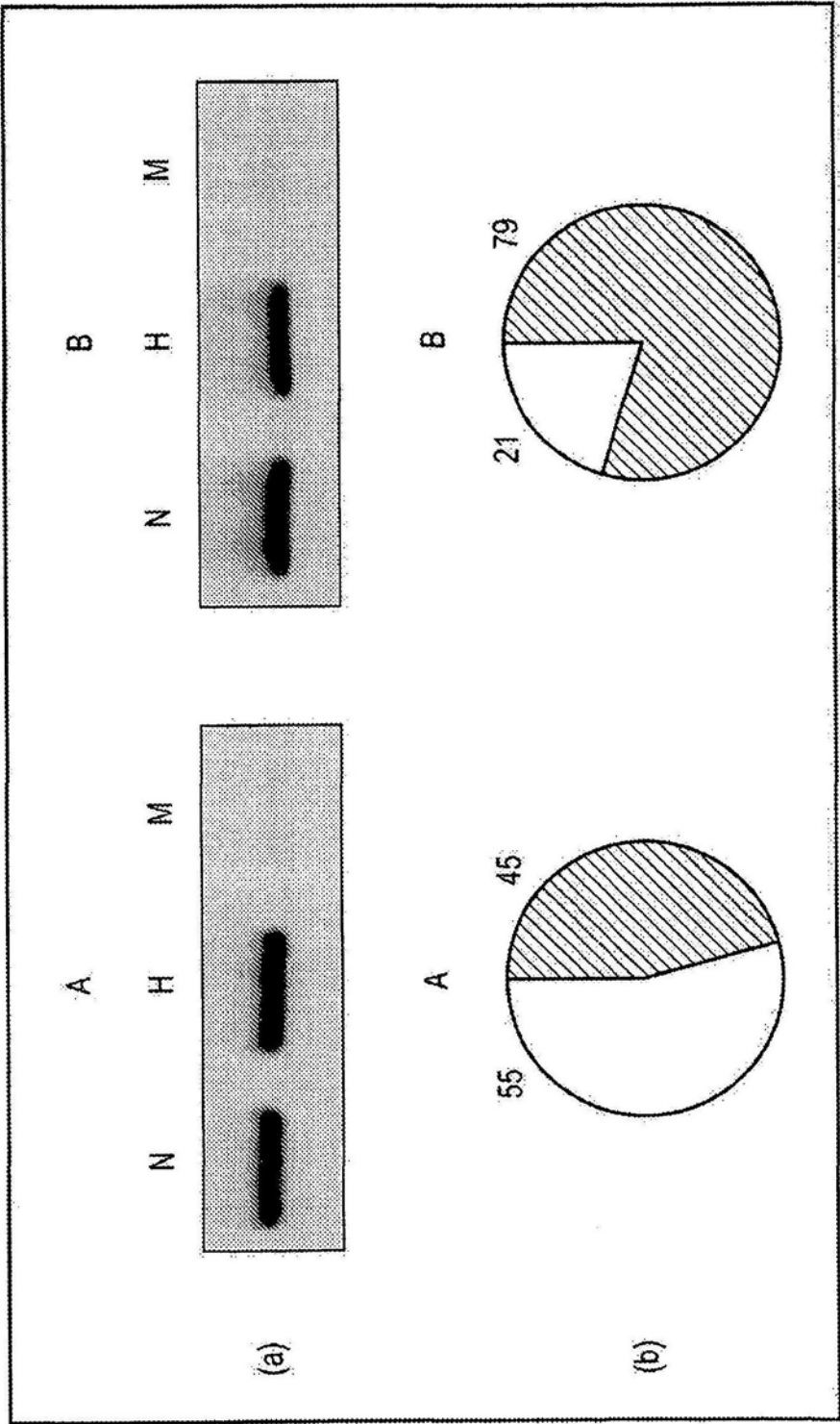


图30



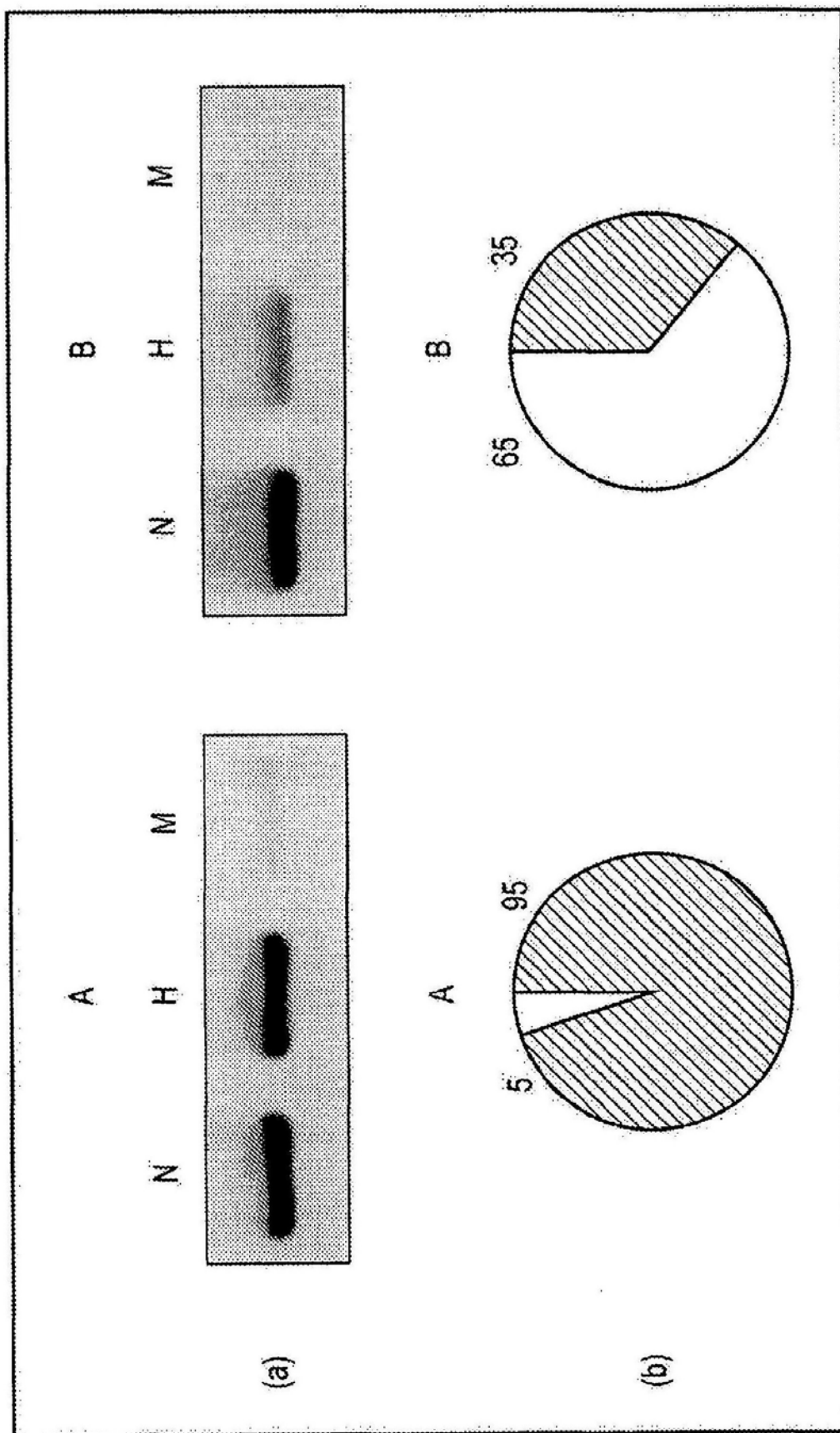


图31

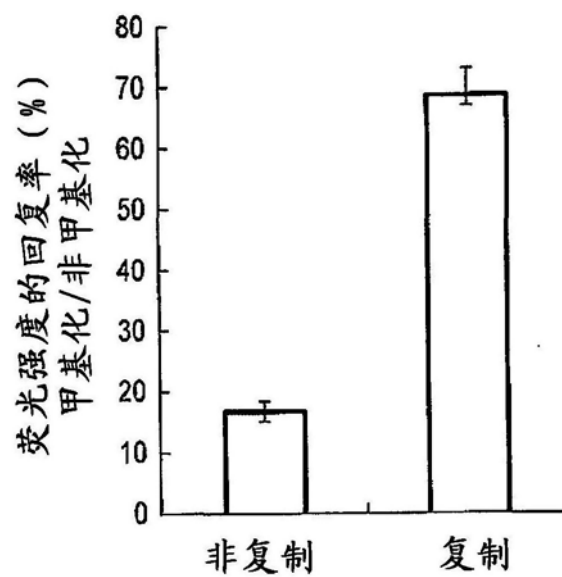


图32

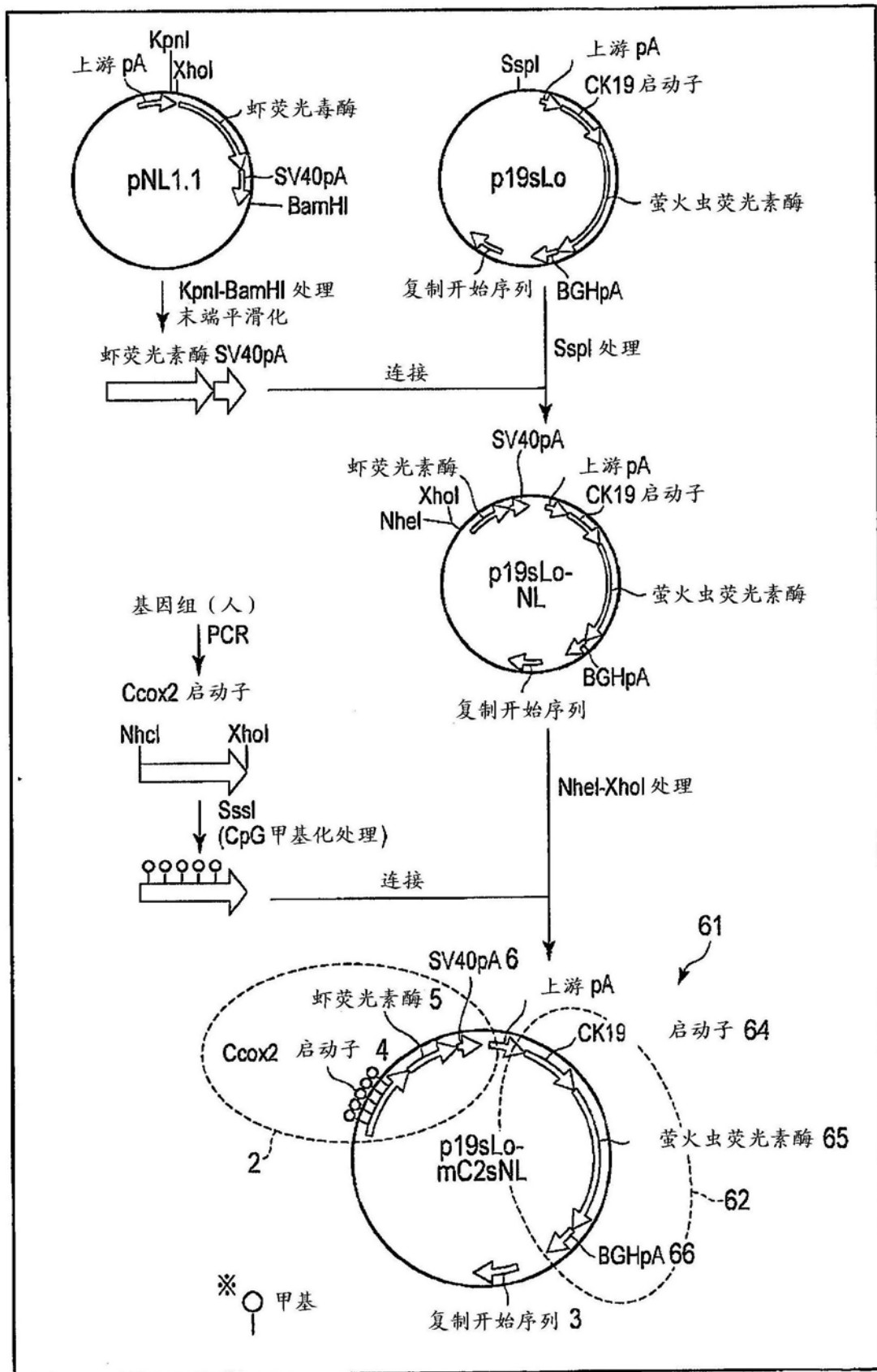


图33

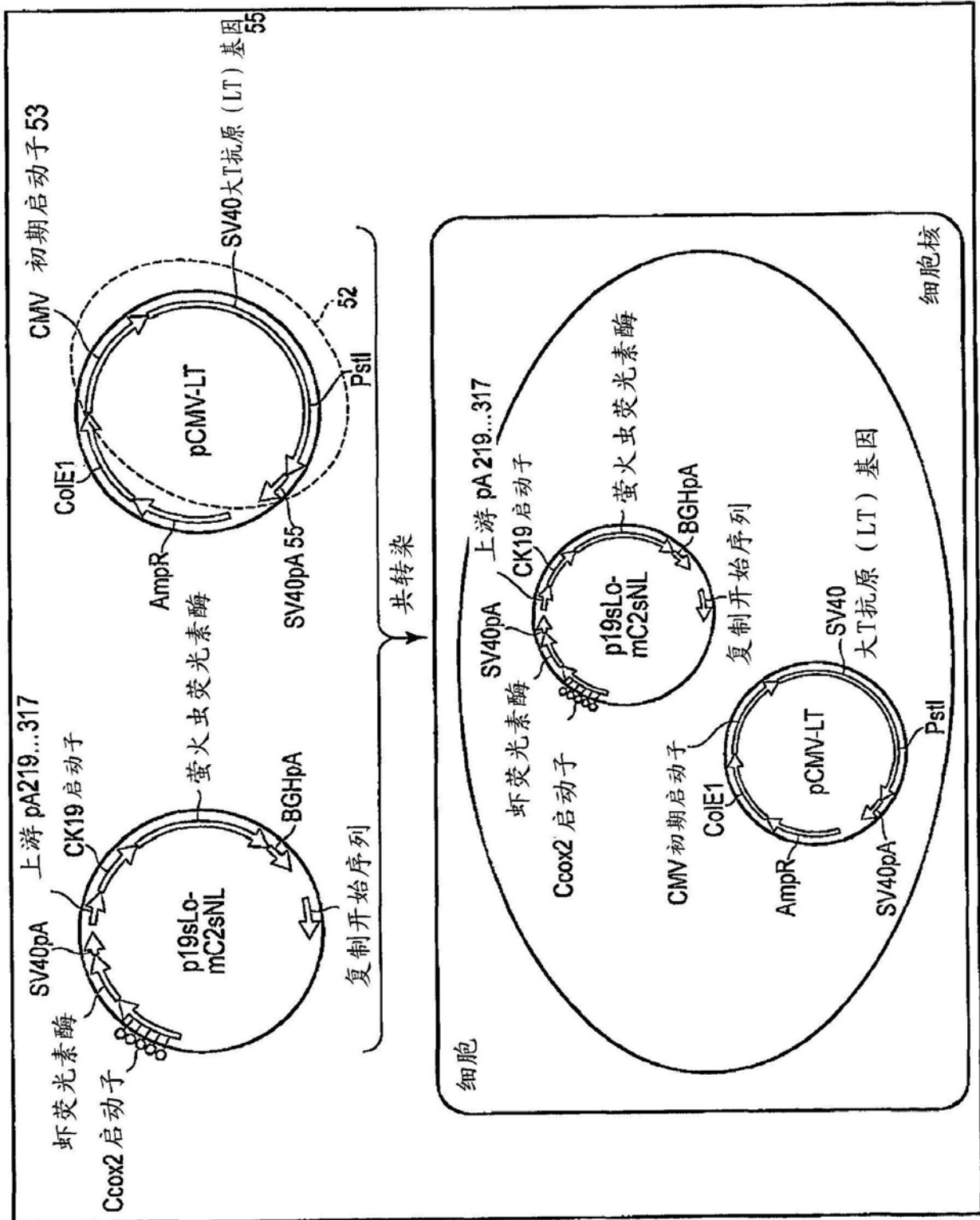


图34

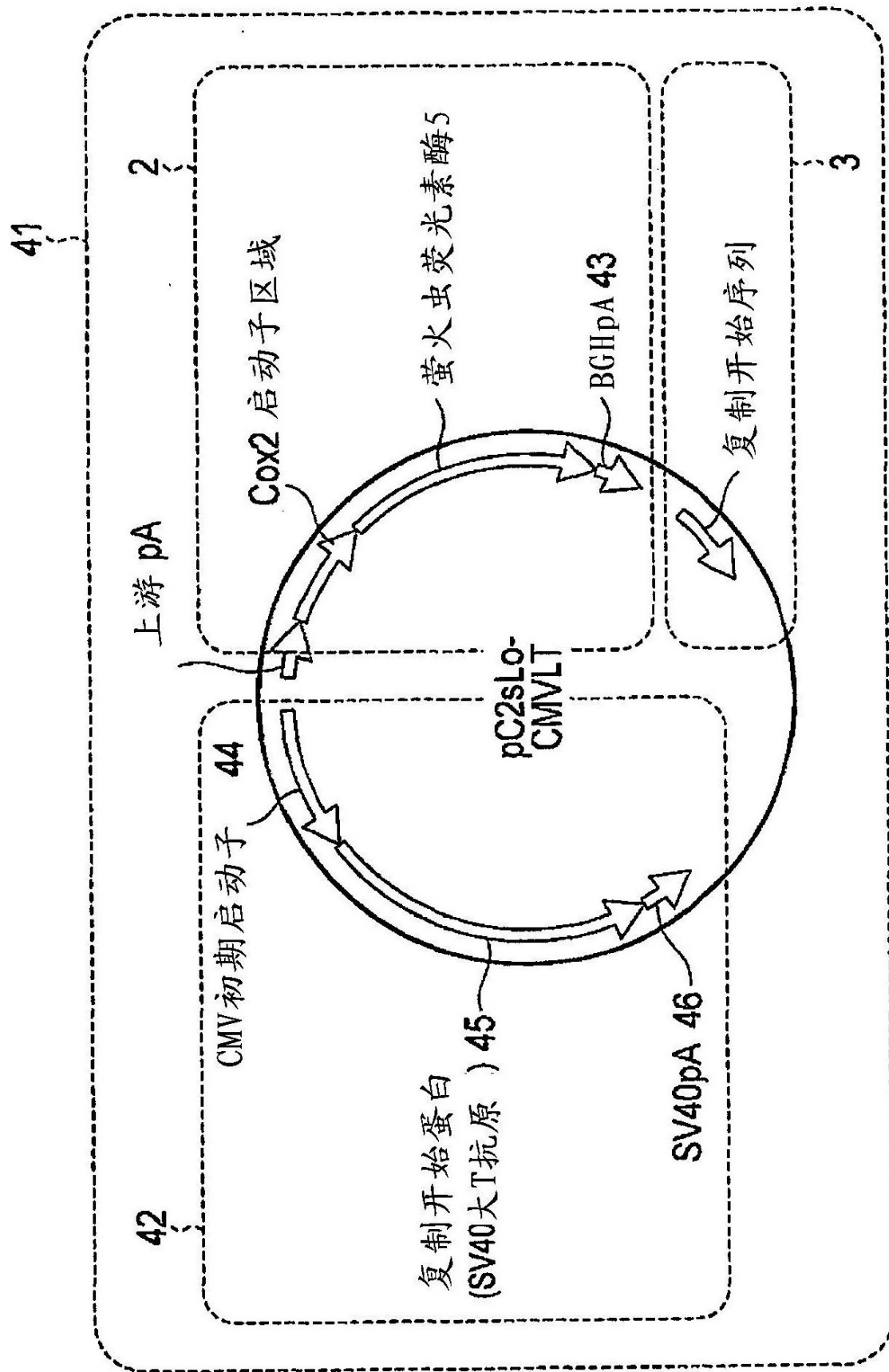


图35

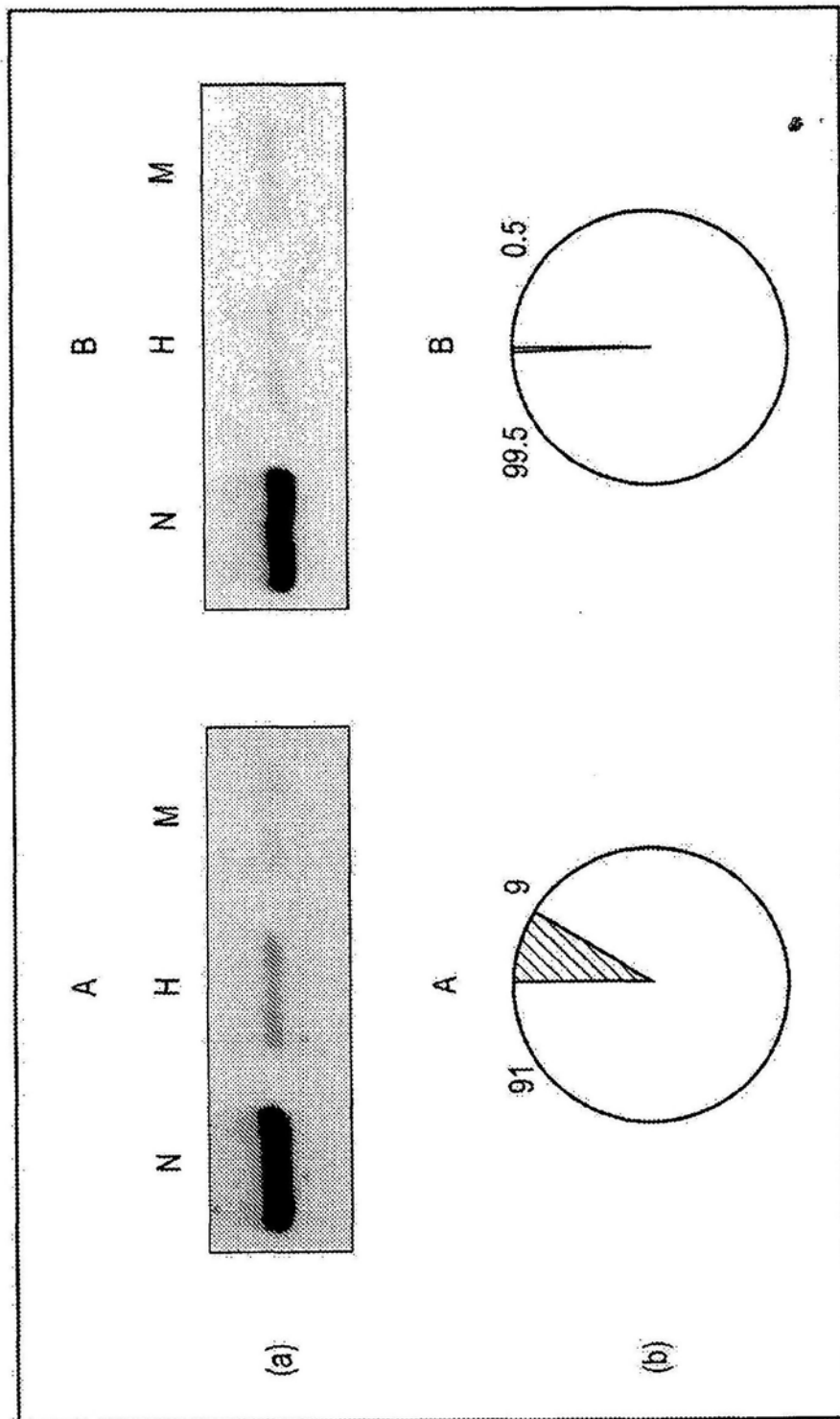


图36