



(12)发明专利申请

(10)申请公布号 CN 110506128 A

(43)申请公布日 2019.11.26

(21)申请号 201780086203.9

(72)发明人 O·阿布达耶 J·J·柯林斯

(22)申请日 2017.12.08

J·戈滕贝格 F·张 E·S·兰德

(30)优先权数据

(74)专利代理机构 北京市金杜律师事务所

62/432,553 2016.12.09 US

11256

62/456,645 2017.02.08 US

代理人 陈文平 袁元

62/471,930 2017.03.15 US

(51)Int.Cl.

62/484,869 2017.04.12 US

C12Q 1/6886(2018.01)

62/568,268 2017.10.04 US

C12Q 1/689(2018.01)

(85)PCT国际申请进入国家阶段日

C12Q 1/6895(2018.01)

2019.08.09

G01N 33/68(2006.01)

G01N 33/569(2006.01)

(86)PCT国际申请的申请数据

G01N 33/53(2006.01)

PCT/US2017/065477 2017.12.08

C12N 15/09(2006.01)

(87)PCT国际申请的公布数据

C12N 15/10(2006.01)

W02018/107129 EN 2018.06.14

C12N 15/11(2006.01)

(71)申请人 博德研究所

地址 美国马萨诸塞州

权利要求书7页 说明书116页

申请人 麻省理工学院 哈佛学院院长等

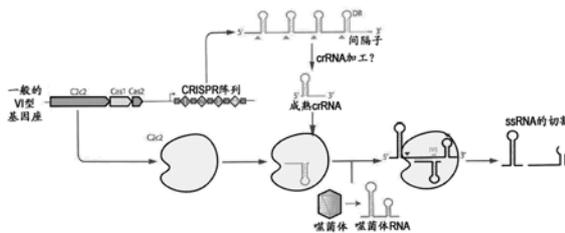
序列表(电子公布) 附图102页

(54)发明名称

基于CRISPR效应系统的诊断

(57)摘要

本文公开的实施方式利用RNA靶向效应子以提供具有阿摩尔灵敏度的稳健的基于CRISPR的诊断。本文公开的实施方式可以以相当水平的灵敏度检测DNA和RNA两者,并且可以基于单碱基对差异将靶标与非靶标区分开。此外,本文公开的实施方式可以以冷冻干燥的形式制备,以便于分配和护理点(POC)应用。这样的实施方式可用于人类健康的多种情况,包括例如病毒检测、细菌菌株分型、灵敏基因分型,和疾病相关无细胞DNA的检测。



1. 一种核酸检测系统,其包括:
检测CRISPR系统,其包括效应蛋白和一种或多种设计用于结合相应靶分子的指导RNA;
和
基于RNA的掩蔽构建体。
2. 一种多肽检测系统,其包括:
检测CRISPR系统,其包括效应蛋白和一种或多种设计用于结合触发RNA的指导RNA;
基于RNA的掩蔽构建体;和
一种或多种检测适体,所述检测适体包含掩蔽的RNA聚合酶启动子结合位点或掩蔽的引物结合位点。
3. 根据权利要求1或2所述的系统,其还包括核酸扩增试剂。
4. 根据权利要求1所述的系统,其中所述靶分子是靶DNA,并且所述系统还包括结合所述靶DNA且包含RNA聚合酶启动子的引物。
5. 根据权利要求1至4中任一项所述的系统,其中所述CRISPR系统效应蛋白是RNA靶向效应蛋白。
6. 根据权利要求5所述的系统,其中所述RNA靶向效应蛋白包含一个或多个HEPN结构域。
7. 根据权利要求6所述的系统,其中所述一个或多个HEPN结构域包含RxxxxH基序序列。
8. 根据权利要求7所述的系统,其中所述RxxxxH基序包含R {N/H/K} X₁X₂X₃H序列。
9. 根据权利要求8所述的系统,其中X₁是R、S、D、E、Q、N、G或Y,并且X₂独立地是I、S、T、V或L,并且X₃独立地是L、F、N、Y、V、I、S、D、E或A。
10. 根据权利要求1至9中任一项所述的系统,其中所述CRISPR RNA靶向效应蛋白是C2c2。
11. 根据权利要求6所述的系统,其中所述CRISPR RNA靶向效应蛋白是C2c2。
12. 根据权利要求11所述的系统,其中所述C2c2是在Cas 1基因的20kb内。
13. 根据权利要求12所述的系统,其中所述C2c2效应蛋白来自选自以下的属的生物:纤毛菌属、李斯特菌属、棒状杆菌属、萨特氏菌属、军团菌属、密螺旋体属、产线菌属、真细菌属、链球菌属、乳酸菌属、支原体属、拟杆菌属、Flaviivola、黄杆菌属、Sphaerochaeta、固氮螺菌属、葡糖醋杆菌属、奈瑟氏菌属、罗氏菌属、Parvibaculum、葡萄球菌属、Nitratifactor、支原体属、弯曲杆菌属和毛螺菌属。
14. 根据权利要求13所述的系统,其中所述C2c2或Cas13b效应蛋白来自选自以下的生物:Leptotrichia shahii;Leptotrichia wadei (Lw2);Listeria seeligeri;Lachnospiraceae bacterium MA2020;Lachnospiraceae bacterium NK4A179;[梭菌属]aminophilum DSM10710;Carnobacterium gallinarum DSM 4847;Carnobacterium gallinarum DSM 4847 (第二CRISPR基因座);Paludibacter propionicigenes WB4;Listeria weihenstephanensis FSL R9-0317;Listeriaceae bacterium FSL M6-0635;Leptotrichia wadei F0279;Rhodobacter capsulatus SB 1003;Rhodobacter capsulatus R121;Rhodobacter capsulatus DE442;Leptotrichia buccalis C-1013-b;Herbinix hemicellulosilytica;[真细菌属]rectale;Eubacteriaceae bacterium CHKCI004;Blautia sp.Marseille-P2398;Leptotrichia sp.oral taxon 879str.F0557;

Lachnospiraceae bacterium NK4A144;Chloroflexus aggregans;Demequina aurantiaca;Thalassospira sp.TSL5-1;Pseudobutyrvibrio sp.OR37;Butyrvibrio sp.YAB3001;Blautia sp.Marseille-P2398;Leptotrichia sp.Marseille-P3007;Bacteroides ihuae;Porphyromonadaceae bacterium KH3CP3RA;Listeria riparia;和 Insolitinspirillum peregrinum。

15. 根据权利要求14所述的系统,其中所述C2c2效应蛋白是L.wadeiF0279或L.wadeiF0279(Lw2)C2c2效应蛋白。

16. 根据权利要求1至15中任一项所述的系统,其中所述基于RNA的掩蔽构建体抑制可检测阳性信号的产生。

17. 根据权利要求16所述的系统,其中所述基于RNA的掩蔽构建体通过掩蔽所述可检测阳性信号或者相反产生可检测阴性信号来抑制可检测阳性信号的产生。

18. 根据权利要求16所述的系统,其中所述基于RNA的掩蔽构建体包含抑制由报告构建体编码的基因产物的产生的沉默RNA,其中所述基因产物在表达时产生所述可检测阳性信号。

19. 根据权利要求16所述的系统,其中所述基于RNA的掩蔽构建体是产生所述可检测阴性信号的核酶,并且其中当所述核酶失活时产生所述可检测阳性信号。

20. 根据权利要求19所述的系统,其中所述核酶将底物转化为第一颜色,并且其中当所述核酶失活时所述底物转化为第二颜色。

21. 根据权利要求16所述的系统,其中所述基于RNA的掩蔽剂是RNA适体和/或包含RNA系连的抑制剂。

22. 根据权利要求21所述的系统,其中所述适体或RNA系连的抑制剂整合酶,其中所述酶通过作用于底物而在从所述适体或RNA系连的抑制剂释放后产生可检测信号。

23. 根据权利要求21所述的系统,其中所述适体是抑制酶并阻止所述酶催化从底物产生可检测信号的抑制性适体,或者其中所述RNA系连的抑制剂抑制酶并阻止所述酶催化从底物产生可检测信号。

24. 根据权利要求23所述的系统,其中所述酶是凝血酶、蛋白C、中性粒细胞弹性蛋白酶、枯草杆菌蛋白酶、辣根过氧化物酶、 β -半乳糖苷酶或小牛碱性磷酸酶。

25. 根据权利要求24所述的系统,其中所述酶是凝血酶,并且所述底物是与用于凝血酶的肽底物共价连接的对硝基苯胺,或与用于凝血酶的肽底物共价连接的7-氨基-4-甲基香豆素。

26. 根据权利要求21所述的系统,其中所述适体整合一对试剂,所述试剂当从所述适体释放时组合以产生可检测信号。

27. 根据权利要求16所述的系统,其中所述基于RNA的掩蔽构建体包含可检测配体和掩蔽组分两者附着至其上的RNA寡核苷酸。

28. 根据权利要求16所述的系统,其中所述基于RNA的掩蔽构建体包含通过桥分子聚集的纳米颗粒,其中所述桥分子的至少一部分包含RNA,并且其中当所述纳米颗粒在溶液中分布时所述溶液经历颜色偏移。

29. 根据权利要求28所述的系统,其中所述纳米颗粒是胶体金属。

30. 根据权利要求29所述的系统,其中所述胶体金属是胶体金。

31. 根据权利要求16所述的系统,其中所述基于RNA的掩蔽构建体包含通过连接分子与一个或多个猝灭分子连接的量子点,其中所述连接分子的至少一部分包含RNA。

32. 根据权利要求16所述的系统,其中所述基于RNA的掩蔽构建体包含与插入剂复合的RNA,其中所述插入剂在所述RNA的切割后改变吸光度。

33. 根据权利要求32所述的系统,其中所述插入剂是吡罗红-Y或亚甲蓝。

34. 根据权利要求16所述的系统,其中所述可检测配体是荧光团,并且所述掩蔽组分是猝灭分子。

34. 根据权利要求1至34中任一项所述的系统,其中所述一种或多种设计用于结合相应靶分子的指导RNA包含(合成)错配。

35. 根据权利要求34所述的系统,其中所述错配是所述靶分子中SNP或其他单核苷酸变异的上游或下游。

36. 根据权利要求1至35中任一项所述的系统,其中所述一种或多种指导RNA被设计用于检测靶RNA或DNA中的单核苷酸多态性,或RNA转录物的剪接变体。

37. 根据权利要求1到36中任一项所述的系统,其中所述一种或多种指导RNA被设计用于结合一种或多种疾病状态诊断性靶分子。

38. 根据权利要求37所述的系统,其中所述疾病状态是癌症。

39. 根据权利要求38所述的系统,其中所述疾病状态是自身免疫疾病。

40. 根据权利要求37所述的系统,其中所述疾病状态是感染。

41. 根据权利要求40所述的系统,其中所述感染是由病毒、细菌、真菌、原生动物或寄生虫引起。

42. 根据权利要求41所述的系统,其中所述感染是病毒感染。

43. 根据权利要求42所述的系统,其中所述病毒感染是由DNA病毒引起。

44. 根据权利要求43所述的系统,其中所述DNA病毒是肌病毒科、短尾病毒科、长尾病毒科、异疱疹病毒科、疱疹病毒科(包括人类疱疹病毒和水痘带状疱疹病毒)、Malcoherpesviridae、脂毛病毒科、小杆状病毒科、腺病毒科、Ampullaviridae、囊泡病毒科、非洲猪瘟病毒科(包括非洲猪瘟病毒)、杆状病毒科、Cicaudaviridae、Clavaviridae、覆盖病毒科、小纺锤形病毒科、球状病毒科、Guttaviridae、Hytrosaviridae、虹彩病毒科、Maseilleviridae、Mimiviridae、Nudiviridae、Nimaviridae、Pandoraviridae、Papillomaviridae、Phycodnaviridae、Plasmaviridae、Polydnaviruses、Polyomaviridae(包括猿猴病毒40、JC病毒、BK病毒)、痘病毒科(包括牛痘和天花)、Sphaerolipoviridae、复层病毒科、Turriviridae、Dinodnavirus、盐末端蛋白病毒属、Rhizidovirus。

45. 根据权利要求42所述的系统,其中所述病毒感染由双链RNA病毒、正义RNA病毒、负义RNA病毒、逆转录病毒或其组合引起。

46. 根据权利要求45所述的系统,其中所述病毒感染由冠状病毒科病毒、小核糖核酸病毒科病毒、杯状病毒科病毒、黄病毒科病毒、披膜病毒科病毒、Bornaviridae、丝状病毒科、副黏液病毒科、肺病毒科、弹状病毒科、沙粒病毒科、布尼亚病毒科、正粘病毒科或Deltavirus引起。

47. 根据权利要求46所述的系统,其中所述病毒感染由冠状病毒、SARS、脊髓灰质炎病毒、鼻病毒、甲型肝炎、诺沃克病毒、黄热病病毒、西尼罗病毒、丙型肝炎病毒、登革热病毒、

寨卡病毒、风疹病毒、罗斯河病毒、辛德毕斯病毒、基孔肯亚病毒、博尔纳病病毒、埃博拉病毒、马尔堡病毒、麻疹病毒、腮腺炎病毒、尼帕病毒、亨德拉病毒、新城疫病毒、人呼吸道合胞病毒、狂犬病病毒、拉沙病毒、汉坦病毒、克里米亚-刚果出血热病毒、流感或丁型肝炎病毒引起。

48. 根据权利要求41所述的系统,其中所述感染是细菌感染。

49. 根据权利要求48所述的系统,其中引起所述细菌感染的细菌是不动杆菌属种、放线杆菌属种、放线菌纲种、放线菌属种、气球菌属种、气单胞菌属种、红孢子虫属种、产碱杆菌属种、芽孢杆菌属种、拟杆菌属种、巴尔通氏体属种、双歧杆菌属种、博代氏杆菌属种、包柔氏螺旋体属种、布鲁氏菌属种、伯克霍尔德氏菌属种、弯曲杆菌属种、嗜二氧化碳噬细胞菌属种、衣原体属种、柠檬酸杆菌属种、柯克斯氏体属种、棒状杆菌属种、梭菌属种、艾肯菌属种、肠杆菌属种、埃希氏菌属种、肠球菌属种、埃立克体属种、表皮癣菌属种、丹毒丝菌属种、真细菌属种、弗兰西斯菌属种、梭杆菌属种、加德纳菌属种、孪生球菌属种、嗜血杆菌属种、螺杆菌属种、金氏菌属种、克雷白氏杆菌属种、乳酸菌属种、乳球菌属种、李斯特菌属种、钩端螺旋体属种、军团菌属种、钩端螺旋体属种、明串珠菌属种、曼氏杆菌属种、小孢子菌属种、微球菌属种、莫拉克斯氏菌属种、摩根氏菌属种、动弯杆菌属种、微球菌属种、分支杆菌属种、支原体属种、诺卡氏菌属种、奈瑟氏菌属种、*Pasteurella*种、片球菌属种、消化链球菌属种、糠疹癣菌属种、邻单胞菌属种、普氏菌属种、卟啉单胞菌属种、变形杆菌属种、普罗维登斯菌属种、假单胞菌属种、丙酸菌属种、红球菌属种、立克次氏体属种、红球菌属种、沙雷氏菌属种、寡养单胞菌属种、沙门氏菌属种、沙雷氏菌属种、志贺氏菌属种、葡萄球菌属种、链球菌属种、螺旋菌属种、链杆菌属种、密螺旋体属种、*Tropheryma*种、毛癣菌属种、脲原体属种、韦永氏球菌属种、弧菌属种、耶尔森氏鼠疫杆菌属种、黄杆菌属种或其组合。

50. 根据权利要求41所述的系统,其中所述感染由真菌引起。

51. 根据权利要求50所述的系统,其中所述真菌是曲霉属、芽生菌属、念珠菌病、球孢子菌病、新型隐球菌、加蒂隐球菌种、组织胞浆菌属种(例如荚膜组织胞浆菌)、肺孢子虫属种(例如耶氏肺孢子虫)、葡萄穗霉属(例如纸葡萄穗霉)、毛霉菌病、孢子丝菌属、真菌性眼睛感染癣、突脐蠕孢属、枝孢属、地丝菌属、酵母属、汉逊酵母属种、念珠菌属种、克鲁维酵母属种、德巴利氏酵母属种、毕赤酵母属种、青霉菌属种、枝孢属种、丝衣酶属种或其组合。

52. 根据权利要求41所述的系统,其中所述感染是由原生动物引起。

53. 根据权利要求52所述的系统,其中所述原生动物是眼虫动物界、异叶足纲、双滴虫目、变形虫界、芽囊原虫属、顶复亚门或其组合。

54. 根据权利要求41所述的系统,其中所述感染是由寄生虫引起。

55. 根据权利要求54所述的系统,其中所述寄生虫是克氏锥虫(查加斯病)、布氏冈比亚锥虫、布氏罗德西亚锥虫、巴西利什曼原虫、*L. infantum*、*L. mexicana*、*L. major*、*L. tropica*、*L. donovani*、*Naegleria fowleri*、*Giardia intestinalis* (*G. lamblia*、*G. duodenalis*)、*canthamoeba castellanii*、*Balamuthia mandrillaris*、*Entamoeba histolytica*、*Blastocystis hominis*、*Babesia microti*、*Cryptosporidium parvum*、*Cyclospora cayetanensis*、*Plasmodium falciparum*、*P. vivax*、*P. ovale*、*P. malariae*和*Toxoplasma gondii*或其组合。

56. 根据权利要求1至55中任一项所述的系统,其中扩增靶RNA分子的试剂包括基于核

酸序列的扩增 (NASBA)、重组酶聚合酶扩增 (RPA)、环介导的等温扩增 (LAMP)、链置换扩增 (SDA)、解旋酶依赖性扩增 (HDA)、切口酶扩增反应 (NEAR)、PCR、多重置换扩增 (MDA)、滚环扩增 (RCA)、连接酶链反应 (LCR) 或分枝扩增法 (RAM)。

57. 根据权利要求1至56中任一项所述的系统, 其还包括富集CRISPR系统, 其中所述富集CRISPR系统被设计用于在用所述检测CRISPR系统检测之前结合相应的靶分子。

58. 根据权利要求57所述的系统, 其中所述富集CRISPR系统包括催化失活的CRISPR效应蛋白。

59. 根据权利要求58所述的系统, 其中催化失活的CRISPR效应蛋白是催化失活的C2c2。

60. 根据权利要求57至59中任一项所述的系统, 其中所述富集CRISPR效应蛋白还包含标签, 其中所述标签用于拉下所述富集CRISPR效应系统, 或将所述富集CRISPR系统结合到固体基质上。

61. 根据权利要求60所述的系统, 其中所述固体基质是流动池。

62. 一种诊断装置, 其包括一个或多个个体离散体积, 每个个体离散体积包含权利要求1至61中任一项所述的CRISPR系统。

63. 根据权利要求62所述的诊断装置, 其中每个个体离散体积还包含一种或多种检测适体, 所述检测适体包含掩蔽的RNA聚合酶启动子结合位点或掩蔽的引物结合位点。

64. 根据权利要求62或63所述的装置, 其中每个个体离散体积还包含核酸扩增试剂。

65. 根据权利要求62所述的装置, 其中所述靶分子是靶DNA, 并且所述个体离散体积还包含结合所述靶DNA且包含RNA聚合酶启动子的引物。

66. 根据权利要求62至65中任一项所述的装置, 其中所述个体离散体积是液滴。

67. 根据权利要求62至66中任一项所述的装置, 其中所述个体离散体积限定在固体基质上。

68. 根据权利要求67所述的装置, 其中所述个体离散体积是微孔。

69. 根据权利要求62至66中任一项所述的诊断装置, 其中所述个体离散体积是在基质上限定的斑点。

70. 根据权利要求69所述的装置, 其中所述基质是柔性材料基质。

71. 根据权利要求70所述的装置, 其中所述柔性材料基质是纸基质或基于柔性聚合物的基质。

72. 一种用于检测样品中的靶核酸的方法, 其包括:

将样品或样品组分配到一个或多个个体离散体积中, 所述个体离散体积包含权利要求中1或3至61中任一项所述的CRISPR系统;

在足以允许所述一种或多种指导RNA与一种或多种靶分子结合的条件下孵育所述样品或样品组;

通过所述一种或多种指导RNA与所述一种或多种靶分子的结合激活所述CRISPR效应蛋白, 其中激活所述CRISPR效应蛋白导致所述基于RNA的掩蔽构建体的修饰, 使得可检测阳性信号产生; 和

检测所述可检测阳性信号, 其中所述可检测阳性信号的检测指示一种或多种靶分子在所述样品中的存在。

73. 一种用于检测样品中的多肽的方法, 其包括:

将样品或样品组分配到个体离散体积的组中,所述个体离散体积包含肽检测适体、权利要求2至61中任一项所述的CRISPR系统;

在足以允许所述肽检测适体与所述一种或多种靶分子结合的条件下孵育所述样品或样品组,其中所述适体与相应靶分子的结合暴露所述RNA聚合酶结合位点或引物结合位点,导致触发RNA产生;

通过所述一种或多种指导RNA与所述触发RNA的结合激活所述RNA效应蛋白,其中激活所述RNA效应蛋白导致所述基于RNA的掩蔽构建体的修饰,使得可检测阳性信号产生;和

检测所述可检测阳性信号,其中所述可检测阳性信号的检测指示一种或多种靶分子在样品中的存在。

74. 根据权利要求72所述的方法,其中所述靶分子是靶DNA,并且所述方法还包括将所述靶DNA与包含RNA聚合酶位点的引物结合。

75. 根据权利要求72至74中任一项所述的方法,其还包括扩增所述样品RNA或所述触发RNA。

76. 根据权利要求75所述的方法,其中扩增RNA包括通过NASBA进行扩增。

77. 根据权利要求75所述的方法,其中扩增RNA包括通过RPA进行扩增。

78. 根据权利要求72至77中任一项所述的方法,其中所述样品是生物样品或环境样品。

79. 根据权利要求78所述的方法,其中生物样品是血液、血浆、血清、尿液、粪便、痰液、粘液、淋巴液、滑液、胆汁、腹水、胸腔积液、血清肿、唾液、脑脊液、水性或玻璃体液、或任何身体分泌物、渗出物、排出物(例如,从脓肿或任何其他感染或炎症部位获得的液体)、或从关节获得的液体(例如,正常关节或受疾病影响的关节,例如类风湿性关节炎、骨关节炎、痛风或化脓性关节炎)、或者皮肤或粘膜表面的拭子。

80. 根据权利要求78所述的方法,其中所述环境样品是从食品样品、纸表面、织物、金属表面、木材表面、塑料表面、土壤样品、淡水样品、废水样品、盐水样品、或其组合获得。

81. 根据权利要求72或74至80中任一项所述的方法,其中所述一种或多种指导RNA被设计用于检测靶RNA或DNA中的单核苷酸多态性,或RNA转录物的剪接变体。

82. 根据权利要求72至81中任一项所述的方法,其中所述一种或多种指导RNA被设计用于结合一种或多种疾病状态诊断性靶分子。

83. 根据权利要求81至82中任一项所述的方法,其中所述一种或多种指导RNA被设计用于结合无细胞核酸。

84. 根据权利要求82所述的方法,其中所述疾病状态是感染、器官疾病、血液疾病、免疫系统疾病、癌症、大脑和神经系统疾病、内分泌疾病、妊娠或分娩相关疾病、遗传性疾病、或环境获得性疾病。

85. 根据权利要求37所述的系统,其中所述疾病状态的特征在于存在或不在于存在抗生素或药物耐受性或易感性基因或转录物或多肽,优选在病原体或细胞中存在或不在于存在。

86. 根据权利要求37所述的系统,其中所述靶分子是抗生素或药物耐受性或易感性基因或转录物或多肽。

87. 根据权利要求35所述的系统,其中所述指导RNA中的所述合成错配位于间隔子的位置3、4、5或6处,优选位置3。

88. 根据权利要求34、35或82所述的系统,其中所述指导RNA中的所述错配位于间隔子

的位置1、2、3、4、5、6、7、8或9处,优选位置5。

89. 根据权利要求35或82所述的系统,其中所述错配是上游或下游的1、2、3、4或5个核苷酸,优选2个核苷酸,优选所述指导RNA中的所述SNP或其他单核苷酸变异的下游。

90. 根据权利要求1-56或85-89中任一项所述的系统,其中所述指导RNA包含相对于野生型间隔子截短的间隔子。

91. 根据权利要求1-56或80-85中任一项所述的系统,其中所述指导RNA包含间隔子,所述间隔子包含少于28个核苷酸,优选地在20至27个核苷酸之间并且包括20和27个核苷酸。

92. 根据权利要求1-56或80-85中任一项所述的系统,其中所述指导RNA包含由20-25个核苷酸或20-23个核苷酸组成的间隔子,例如优选地20或23个核苷酸。

93. 根据权利要求1-56或85-92中任一项所述的系统,其中所述掩蔽构建体包含设计用于结合G-四链体形成序列的RNA寡核苷酸,其中G-四链体结构是在所述掩蔽构建体的切割后由所述G-四链体形成序列形成,并且其中所述G-四链体结构产生可检测阳性信号。

94. 根据权利要求72-84中任一项所述的方法,其还包括比较所述可检测阳性信号和(合成的)标准信号。

95. 一种用于检测样品中的靶核酸的方法,其包括:

将样品与权利要求1至56中任一项所述的核酸检测系统接触;和

将所述接触的样品应用于侧流免疫色谱测定。

96. 根据权利要求94所述的方法,其中所述核酸检测系统包括包含第一和第二分子的基于RNA的掩蔽构建体,并且其中所述侧流免疫色谱测定包括检测所述第一和第二分子,优选在侧流条的离散检测位点处检测。

97. 根据权利要求95所述的方法,其中所述第一分子和所述第二分子通过结合识别所述第一或第二分子的抗体并且检测所述结合分子来检测,优选采用夹心抗体。

98. 根据权利要求94或95所述的方法,其中所述侧流条包含针对所述第一分子的上游第一抗体和针对所述第二分子的下游第二抗体,并且其中如果所述靶核酸不存在于所述样品中,则未切割的基于RNA的掩蔽构建体被所述第一抗体结合,并且其中如果所述靶核酸存在于所述样品中,则切割的基于RNA的掩蔽构建体被所述第一抗体和所述第二抗体两者结合。

基于CRISPR效应系统的诊断

[0001] 关于联邦政府资助研究的声明

[0002] 本发明是在国立卫生研究院授予的拨款号MH100706和MH110049以及由国防威胁减少机构授予的拨款号HDTRA1-14-1-0006的政府支持下完成的。政府拥有本发明的某些权利。

技术领域

[0003] 本文所公开的主题一般涉及与CRISPR效应系统的使用相关的快速诊断。

背景技术

[0004] 核酸是生物信息的通用标志。在便携式平台上以高灵敏度和单碱基特异性快速检测核酸的能力具有彻底改革许多疾病的诊断和监测的潜力,提供有价值的流行病学信息,并作为普遍的科学工具。尽管已经开发了许多用于检测核酸的方法(Du等,2017;Green等,2014;Kumar等,2014;Pardee等,2014;Pardee等,2016;Urdea等,2006),但它们不可避免地受损于在灵敏度、特异性、简单性和速度之间进行权衡。例如,qPCR方法灵敏但昂贵,并且依赖于复杂仪器,限制了对于在实验室环境中训练有素的操作员的可用性。其他方法,例如将等温核酸扩增与便携式平台相结合的新方法(Du等,2017;Pardee等,2016),在护理点(POC)环境中提供了高检测特异性,但由于灵敏度低而应用有限。随着核酸诊断变得与各种医疗保健应用越来越相关,在低成本下提供高特异性和灵敏度的检测技术在临床和基础研究环境两者中将具有很大的实用性。

发明内容

[0005] 在一个方面,本发明提供一种核酸检测系统,其包括:CRISPR系统,其包括效应蛋白和一种或多种设计用于结合相应靶分子的指导RNA;和基于RNA的掩蔽构建体;和任选地,用于扩增样品中的靶RNA分子的核酸扩增试剂。在另一个方面,实施方式提供一种多肽检测系统,其包括:CRISPR系统,其包括效应蛋白和一种或多种设计用于结合触发RNA(trigger RNA)的指导RNA,基于RNA的掩蔽构建体(masking construct);和一种或多种检测适体,所述检测适体包括掩蔽的RNA聚合酶启动子结合位点或掩蔽的引物结合位点。

[0006] 在其他实施方式中,系统还可以包括核酸扩增试剂。核酸扩增试剂可以包含含有RNA聚合酶启动子的引物。在某些实施方式中,扩增样品核酸以获得包含RNA聚合酶启动子的DNA模板,由此可通过转录产生靶RNA分子。核酸可以是DNA并通过本文所述的任何方法扩增。核酸可以是RNA并通过本文所述的逆转录方法扩增。可以在揭露(unmasking)引物结合位点时扩增适体序列,由此从扩增的DNA产物转录触发RNA。靶分子可以是靶DNA,并且系统还可以包括结合靶DNA且包含RNA聚合酶启动子的引物。

[0007] 在一个示例性实施方式中,CRISPR系统效应蛋白是RNA靶向效应蛋白。实例RNA靶向效应蛋白包括Cas13b和C2c2(现在称为Cas13a)。应理解,本文中的术语“C2c2”可与“Cas13a”互换使用。在另一个示例性实施方式中,靶向RNA的效应蛋白是C2c2。在其他实施

方式中,C2c2效应蛋白来自选自以下的属的生物:纤毛菌属(Leptotrichia)、李斯特菌属(Listeria)、棒状杆菌属(Corynebacter)、萨特氏菌属(Sutterella)、军团菌属(Legionella)、密螺旋体属(Treponema)、产线菌属(Filifactor)、真细菌属(Eubacterium)、链球菌属(Streptococcus)、乳酸菌属(Lactobacillus)、支原体属(Mycoplasma)、拟杆菌属(Bacteroides)、Flaviivola、黄杆菌属(Flavobacterium)、Sphaerochaeta、固氮螺菌属(Azospirillum)、葡糖醋杆菌属(Gluconacetobacter)、奈瑟氏菌属(Neisseria)、罗氏菌属(Roseburia)、Parvibaculum、葡萄球菌属(Staphylococcus)、Nitratifactor、支原体属(Mycoplasma)、弯曲杆菌属(Campylobacter)和毛螺菌属(Lachnospira),或者C2c2效应蛋白是选自以下的生物:Leptotrichia shahii、Leptotrichia.Wadei、Listeria seeligeri、Clostridium aminophilum、Carnobacterium gallinarum、Paludibacter propionicigenes、Listeria weihenstephanensis,或C2c2效应蛋白是L.wadei F0279或L.wadei F0279(Lw2) C2c2效应蛋白。在另一个实施方式中,设计一种或多种指导RNA以检测单核苷酸多态性、转录物的剪接变体或靶RNA或DNA中的移码突变。

[0008] 在其他实施方式中,设计一种或多种指导RNA以结合一种或多种诊断疾病状态的靶分子。在其他实施方式中,疾病状态是感染、器官疾病、血液疾病、免疫系统疾病、癌症、大脑和神经系统疾病、内分泌疾病、妊娠或分娩相关疾病、遗传性疾病、或环境获得性疾病。在其他实施方式中,疾病状态是癌症或自身免疫疾病或感染。

[0009] 在其他实施方式中,设计一种或多种指导RNA以结合包含癌症特异性体细胞突变的一种或多种靶分子。癌症特异性突变可以赋予药物抗性。可以通过用依鲁替尼(ibrutinib)、厄洛替尼(erlotinib)、伊马替尼(imatinib)、吉非替尼(gefitinib)、克唑替尼(crizotinib)、曲妥珠单抗(trastuzumab)、维罗非尼(vemurafenib)、RAF/MEK、检查点阻滞疗法、或抗雌激素疗法进行处理治疗来诱导药物抗性突变。癌症特异性突变可存在于编码选自以下的蛋白的一种或多种基因中:程序性死亡配体1(PD-L1)、雄激素受体(AR)、布鲁顿酪氨酸激酶(BTK)、表皮生长因子受体(EGFR)、BCR-Ab1、c-kit、PIK3CA、HER2、EML4-ALK、KRAS、ALK、ROS1、AKT1、BRAF、MEK1、MEK2、NRAS、RAC1和ESR1。癌症特异性突变可以是选自以下的基因中的突变:CASP8、B2M、PIK3CA、SMC1A、ARID5B、TET2、ALPK2、COL5A1、TP53、DNER、NCOR1、MORC4、CIC、IRF6、MYOCD、ANKLE1、CNKSR1、NF1、SOS1、ARID2、CUL4B、DDX3X、FUBP1、TCP11L2、HLA-A、B或C、CSNK2A1、MET、ASXL1、PD-L1、PD-L2、IDO1、IDO2、ALOX12B和ALOX15B,或拷贝数增加,排除全染色体事件,影响以下任何染色体带:6q16.1-q21、6q22.31-q24.1、6q25.1-q26、7p11.2-q11.1、8p23.1、8p11.23-p11.21(含有IDO1、IDO2)、9p24.2-p23(含有PDL1、PDL2)、10p15.3、10p15.1-p13、11p14.1、12p13.32-p13.2、17p13.1(含有ALOX12B、ALOX15B)和22q11.1-q11.21。

[0010] 在其他实施方式中,可以设计一种或多种指导RNA以结合包含杂合性丢失(LOH)标志物的一种或多种靶分子。

[0011] 在其他实施方式中,可以设计一种或多种指导RNA以结合包含单核苷酸多态性(SNP)的一种或多种靶分子。疾病可以是心脏病,并且靶分子可以是VKORC1、CYP2C9和CYP2C19。

[0012] 在其他实施方式中,疾病状态可以是妊娠或分娩相关疾病或遗传疾病。样品可以

是血液样品或粘液样品。疾病可选自13三体、16三体、18三体、Klinefelter综合症(47, XXY)、(47, XYY)和(47, XXX)、特纳综合症、唐氏综合症(21三体)、囊性纤维化、亨廷顿病、 β 地中海贫血、肌强直性营养不良、镰状细胞贫血、卟啉症、脆性X综合症、罗伯逊易位、Angelman综合症、DiGeorge综合症和Wolf-Hirschhorn综合症。

[0013] 在其他实施方式中,感染由病毒、细菌或真菌引起,或者感染是病毒感染。在特定的实施方式中,病毒感染由双链RNA病毒、正义RNA病毒、负义RNA病毒、逆转录病毒或其组合引起,或者病毒感染由冠状病毒科(Coronaviridae)病毒、小核糖核酸科(Picornaviridae)病毒、杯状病毒科(Caliciviridae)病毒、黄病毒科(Flaviviridae)病毒、披膜病毒科(Togaviridae)病毒、Bornaviridae、丝状病毒科(Filoviridae)、副黏液病毒科(Paramyxoviridae)、肺病毒科(Pneumoviridae)、弹状病毒科(Rhabdoviridae)、沙粒病毒科(Arenaviridae)、布尼亚病毒科(Bunyaviridae)、正粘病毒科(Orthomyxoviridae)或Deltavirus引起,或者病毒感染由冠状病毒(Coronavirus)、SARS、脊髓灰质炎病毒(Poliovirus)、鼻病毒(Rhinovirus)、甲型肝炎(Hepatitis A)、诺沃克病毒(Norwalk virus)、黄热病病毒(Yellow fever virus)、西尼罗病毒(West Nile virus)、丙型肝炎病毒(Hepatitis C virus)、登革热病毒(Dengue fever virus)、寨卡病毒(Zika virus)、风疹病毒(Rubella virus)、罗斯河病毒(Ross River virus)、辛德毕斯病毒(Sindbis virus)、基孔肯亚病毒(Chikungunya virus)、博尔纳病病毒(Borna disease virus)、埃博拉病毒(Ebola virus)、马尔堡病毒(Marburg virus)、麻疹病毒(Measles virus)、腮腺炎病毒(Mumps virus)、尼帕病毒(Nipah virus)、亨德拉病毒(Hendra virus)、新城疫病毒(Newcastle disease virus)、人呼吸道合胞病毒(Human respiratory syncytial virus)、狂犬病病毒(Rabies virus)、拉沙病毒(Lassa virus)、汉坦病毒(Hantavirus)、克里米亚-刚果出血热病毒(Crimean-Congo hemorrhagic fever virus)、流感(Influenza)或丁型肝炎病毒(Hepatitis D virus)引起。

[0014] 在本发明的其他实施方式中,基于RNA的掩蔽构建体抑制可检测阳性信号的产生,或者基于RNA的掩蔽构建体通过掩蔽可检测阳性信号或者相反产生可检测阴性信号来抑制可检测阳性信号的产生,或者基于RNA的掩蔽构建体包含抑制由报告构建体编码的基因产物的产生的沉默RNA,其中基因产物在表达时产生可检测阳性信号。

[0015] 在其他实施方式中,基于RNA的掩蔽构建体是产生可检测阴性信号的核酶,并且其中当核酶失活时产生可检测阳性信号,或者所述核酶将底物转化为第一颜色,并且其中当所述核酶失活时所述底物转化为第二颜色。

[0016] 在其他实施方式中,基于RNA的掩蔽剂是RNA适体,或螯合酶的适体,其中酶通过作用于底物而在从所述适体释放后产生可检测信号,或者适体螯合一对试剂,所述试剂当从所述适体释放时组合以产生可检测信号。

[0017] 在另一个实施方式中,基于RNA的掩蔽构建体包含可检测配体和掩蔽组分两者附着至其上的RNA寡核苷酸。在另一个实施方式中,可检测配体是荧光团,并且掩蔽组分是猝灭分子,或扩增靶RNA分子的试剂,例如但不限于NASBA或RPA试剂。

[0018] 在另一个方面,本发明提供一种诊断装置,其包括一个或多个个体离散体积,个体离散体积各自包含CRISPR效应蛋白、一种或多种设计用于结合相应靶分子的指导RNA、基于RNA的掩蔽构建体,并且任选地还包含核酸扩增试剂。

[0019] 在另一个方面,本发明提供了一种诊断装置,其包括一个或多个个体离散体积,个体离散体积各自包含CRISPR效应蛋白、一种或多种设计用于结合触发RNA的指导RNA、一种或多种包含掩蔽的RNA聚合酶启动子结合位点或掩蔽的引物结合位点的检测适体,并且任选地还包含核酸扩增试剂。

[0020] 在一些实施方式中,个体离散体积是液滴,或者个体离散体积限定在固体基质上,或者个体离散体积是微孔,或者个体离散体积是在基质(例如纸基质)上限定的斑点。

[0021] 在一个实施方式中,RNA靶向效应蛋白是CRISPR VI型RNA靶向效应蛋白,例如C2c2或Cas13b。在某些示例性实施方式中,C2c2效应蛋白来自选自以下的生物:纤毛菌属(*Leptotrichia*)、李斯特菌属(*Listeria*)、棒状杆菌属(*Corynebacter*)、萨特氏菌属(*Sutterella*)、军团菌属(*Legionella*)、密螺旋体属(*Treponema*)、产线菌属(*Filifactor*)、真细菌属(*Eubacterium*)、链球菌属(*Streptococcus*)、乳酸菌属(*Lactobacillus*)、支原体属(*Mycoplasma*)、拟杆菌属(*Bacteroides*)、*Flaviivola*、黄杆菌属(*Flavobacterium*)、*Sphaerochaeta*、固氮螺菌属(*Azospirillum*)、葡糖醋杆菌属(*Gluconacetobacter*)、奈瑟氏菌属(*Neisseria*)、罗氏菌属(*Roseburia*)、*Parvibaculum*、葡萄球菌属(*Staphylococcus*)、*Nitratifactor*、支原体属(*Mycoplasma*)和弯曲杆菌属(*Campylobacter*),或者C2c2效应蛋白选自:*Leptotrichia shahii*、*L.wadei*、*Listeria seeligeri*、*Lachnospiraceae bacterium*、*Clostridium aminophilum*、*Carnobacterium gallinarum*、*Paludibacter propionicigenes*、*Listeria weihenstephanensis*、*Listeriaceae bacterium*、和 *Rhodobacter capsulatus*,或者C2c2效应蛋白是*L.wadei* F0279或*L.wadei* F0279(Lw2) C2c2效应蛋白。在另一个实施方式中,一种或多种指导RNA被设计用于结合一种或多种疾病状态诊断性的靶RNA序列。

[0022] 在某些示例性实施方式中,基于RNA的掩蔽构建体抑制可检测阳性信号的产生,或者基于RNA的掩蔽构建体通过掩蔽可检测阳性信号或者相反产生可检测阴性信号来抑制可检测阳性信号的产生,或者基于RNA的掩蔽构建体包含抑制由报告构建体编码的基因产物的产生的沉默RNA,其中基因产物在表达时产生可检测阳性信号。

[0023] 在另一个示例性实施方式中,基于RNA的掩蔽构建体是产生所述可检测阴性信号的核酶,并且其中当所述核酶失活时产生所述可检测阳性信号。在一个示例性实施方式中,核酶将底物转化为第一颜色,并且其中当所述核酶失活时所述底物转化为第二颜色。在另一个示例性实施方式中,基于RNA的掩蔽剂是螯合酶的适体,其中酶通过作用于底物而在从所述适体释放后产生可检测信号,或者适体螯合一对试剂,所述试剂当从所述适体释放时组合以产生可检测信号。

[0024] 在另一个示例性实施方式中,基于RNA的掩蔽构建体包含可检测配体寡核苷酸和掩蔽组分附着至其的RNA寡核苷酸。在某些示例性实施方式中,可检测配体是荧光团,并且掩蔽组分是猝灭分子。

[0025] 在另一个方面,本发明提供一种检测样品中的靶RNA的方法,其包括:将样品或样品组分配到一个或多个个体离散体积中,所述个体离散体积包含CRISPR系统,所述CRISPR系统包含效应蛋白、一种或多种指导RNA、基于RNA的掩蔽构建体;在足以允许所述一种或多种指导RNA与一种或多种靶分子结合的条件下孵育所述样品或样品组;通过所述一种或多种指导RNA与所述一种或多种靶分子的结合激活所述CRISPR效应蛋白,其中激活所述

CRISPR效应蛋白导致所述基于RNA的掩蔽构建体的修饰,使得可检测阳性信号产生;和检测所述可检测阳性信号,其中所述可检测阳性信号的检测指示一种或多种靶分子在所述样品中的存在。

[0026] 在另一个方面,本发明提供了一种检测样品中肽的方法,其包括:将样品或样品组分配到个体离散体积的组中,所述个体离散体积包含肽检测适体、CRISPR系统,所述CRISPR系统包含效应蛋白、一种或多种指导RNA、基于RNA的掩蔽构建体,其中所述肽检测适体包含掩蔽RNA聚合酶位点并且被配置为结合一种或多种靶分子;在足以允许所述肽检测适体与所述一种或多种靶分子结合的条件下孵育所述样品或样品组,其中所述适体与相应靶分子的结合暴露所述RNA聚合酶结合位点,导致触发RNA的RNA合成;通过所述一种或多种指导RNA与所述触发RNA的结合激活所述CRISPR效应蛋白,其中激活所述CRISPR效应蛋白导致所述基于RNA的掩蔽构建体的修饰,使得可检测阳性信号产生;和检测所述可检测阳性信号,其中所述可检测阳性信号的检测指示一种或多种靶分子在样品中的存在。

[0027] 在某些示例性实施方式中,这样的方法还包括扩增样品RNA或触发RNA。在其他实施方式中,扩增RNA包括通过NASBA或RPA扩增。

[0028] 在某些示例性实施方式中,CRISPR效应蛋白是CRISPR VI型RNA靶向效应蛋白,例如C2c2或Cas13b。在其他示例性实施方式中,C2c2效应蛋白来自选自以下的生物:纤毛菌属(*Leptotrichia*)、李斯特菌属(*Listeria*)、棒状杆菌属(*Corynebacter*)、萨特氏菌属(*Sutterella*)、军团菌属(*Legionella*)、密螺旋体属(*Treponema*)、产线菌属(*Filifactor*)、真细菌属(*Eubacterium*)、链球菌属(*Streptococcus*)、乳酸菌属(*Lactobacillus*)、支原体属(*Mycoplasma*)、拟杆菌属(*Bacteroides*)、*Flaviivola*、黄杆菌属(*Flavobacterium*)、*Sphaerochaeta*、固氮螺菌属(*Azospirillum*)、葡糖醋杆菌属(*Gluconacetobacter*)、奈瑟氏菌属(*Neisseria*)、罗氏菌属(*Roseburia*)、*Parvibaculum*、葡萄球菌属(*Staphylococcus*)、*Nitratifactor*、支原体属(*Mycoplasma*)和弯曲杆菌属(*Campylobacter*),或者C2c2效应蛋白选自:*Leptotrichia shahii*、*L.wadei*、*Listeria seeligeri*、*Lachnospiraceae bacterium*、*Clostridium aminophilum*、*Carnobacterium gallinarum*、*Paludibacter propionisigenes*、*Listeria weihenstephanensis*、*Listeriaceae bacterium*、和 *Rhodobacter capsulatus*。在特定的实施方式中,C2c2效应蛋白是*L.wadei* F0279或 *L.wadei* F0279 (Lw2) C2c2效应蛋白。

[0029] 在某些示例性实施方式中,一种或多种指导RNA被设计用于结合一种或多种疾病状态诊断性靶分子。在某些其他示例性实施方式中,疾病状态是感染、器官疾病、血液疾病、免疫系统疾病、癌症、大脑和神经系统疾病、内分泌疾病、妊娠或分娩相关疾病、遗传疾病、或环境获得性疾病、癌症,或真菌感染、细菌感染、寄生虫感染或病毒感染。

[0030] 在某些示例性实施方式中,基于RNA的掩蔽构建体抑制可检测阳性信号的产生,或者所述基于RNA的掩蔽构建体通过掩蔽所述可检测阳性信号或者相反产生可检测阴性信号来抑制可检测阳性信号的产生,或者所述基于RNA的掩蔽构建体包含抑制由报告构建体编码的基因产物的产生的沉默RNA,其中所述基因产物在表达时产生所述可检测阳性信号,或所述基于RNA的掩蔽构建体是产生所述可检测阴性信号的核酶,并且其中当所述核酶失活时产生所述可检测阳性信号。在其他示例性实施方式中,核酶将底物转化为第一颜色,并且其中当所述核酶失活时所述底物转化为第二颜色,或者基于RNA的掩蔽剂是RNA适体,或螯

合酶的适体,其中酶通过作用于底物而在从所述适体释放后产生可检测信号,或者适体螯合一对试剂,所述试剂当从所述适体释放时组合以产生可检测信号。在其他实施方式中,基于RNA的掩蔽构建体包含在RNA寡核苷酸的第一末端上具有可检测配体的RNA寡核苷酸和在RNA寡核苷酸的第二末端上的掩蔽组分,或者可检测配体是荧光团,并且掩蔽组分是猝灭分子。

[0031] 在考虑以下所示示例性实施方式的详细描述之后,示例性实施方式的这些和其他方面、目的、特征和优点对于本领域普通技术人员将变得显而易见。

附图说明

[0032] 图1基于C2c2的CRISPR效应系统的实例的示意图。

[0033] 图2提供(A)来自*Leptotrichia wadei*的CRISPR/C2c2基因座的示意图。显示来自LwC2c2和LshC2c2系统的代表性crRNA结构。(SEQ. I. D. No. 142和143) (B) C2c2活性的体内细菌测定的示意图。将原型间隔子克隆到氨苄青霉素抗性质粒中的 β -内酰胺酶基因的上游,并将该构建体与靶向或非靶向间隔子一起转化到表达C2c2的大肠杆菌中。成功的转化子被计数以量化活性。(C) LwC2c2和LshC2c2体内活性的定量。(n=2个生物学重复;柱代表平均值 \pm s.e.m.) (D) LwC2c2的最终尺寸排阻凝胶过滤。(E) LwC2c2逐步纯化的考马斯蓝染色丙烯酰胺凝胶。(F) LwC2c2对不同PFS靶标的活性。将LwC2c2靶向具有位于间隔子侧翼的可变3' PFS的荧光RNA,并且在变性凝胶上显现反应产物。LwC2c2显示出对G PFS的轻微偏好。

[0034] 图3-6显示示例掩蔽构建体在不同稀释度的检测,其使用1 μ g、100ng、10ng和1ng靶标,以及4种不同量的蛋白质/crRNA(1:4、1:16、1:32、1:64),以及2个crRNA库、无crRNA条件,技术重复,在(96+48)*2=288个反应中,3小时内以5分钟间隔测量。

[0035] 图7提供根据某些示例性实施方式,使用掩蔽构建体和CRISPR效应蛋白的示例性检测方案的示意图。

[0036] 图8提供显示当使用不同的指导RNA库检测靶标时荧光随时间的变化的一组图。

[0037] 图9提供显示在不同浓度的CRISPR效应蛋白下,在不同稀释度的靶RNA中检测到的标准化荧光的图。

[0038] 图10是显示NASBA扩增反应的一般步骤的示意图。

[0039] 图11提供显示用三种不同引物组通过NASBA扩增,然后使用猝灭荧光探针进行C2c2并行检测的核酸靶标ssRNA 1的检测的图。(n=2个技术重复;柱代表平均值 \pm s.e.m.)。

[0040] 图12提供显示并行效应可用于检测慢病毒靶RNA的存在的图。

[0041] 图13提供证实并行效应和NASBA可以检测aM浓度的物质的图。

[0042] 图14提供证实并行效应和NASBA快速区分低浓度样品的图。

[0043] 图15显示在特定时间点的标准化荧光预测样品输入浓度。来自没有扩增的Cas13a检测的荧光测量与输入RNA浓度相关。(n=2个生物学重复;柱代表平均值 \pm s.e.m.)。

[0044] 图16提供RPA反应的示意图,显示了反应中的参与组分。

[0045] 图17SHERLOCK示意图;提供通过引入RPA步骤或相应引入RT-RPA步骤检测DNA或RNA靶标两者的示意图。在识别靶RNA后,并行效应导致C2c2切割切割报告子,产生荧光。单分子量的RNA或DNA可通过重组酶聚合酶扩增(RPA)扩增成DNA,并转录以产生RNA,其然后由

C2c2检测。

[0046] 图18提供通过C2c2并行检测而检测到的ssRNA靶标的示意图 (SEQ. I. D. No. 144和145)。

[0047] 图19提供了证实使用RPA的单分子DNA检测 (即在C2c2添加的15分钟内) 的一组图。

[0048] 图20提供证实将T7聚合酶混合到RPA反应中确实对DNA检测产生不利影响的一组图。

[0049] 图21提供证实将聚合酶混合到RPA反应中对DNA检测没有产生不利影响的一组图。

[0050] 图22提供证实RPA、T7转录和C2c2检测反应在同时孵育时相容并实现单分子检测的图 (n=2个技术重复;柱代表平均值±s.e.m.)。

[0051] 图23提供证实快速RPA-RNA时间孵育的功效的一组图。

[0052] 图24提供证实逐渐增加的T7聚合酶量增强对RPA-RNA的灵敏度的一组图。

[0053] 图25提供显示来自使用 $1.5\times$ 酶的一锅反应进行RPA-DNA检测测定的结果的一组图。单分子 (2aM) 检测在早至30分钟实现。

[0054] 图26提供证实RPA-DNA一锅反应证实荧光相对于输入浓度的定量降低的一组图。拟合曲线揭示了靶标输入浓度与输出荧光之间的关系。

[0055] 图27提供一组图,证实 (A) 无扩增的RNA的C2c2检测可以在低至50fM的浓度下检测ssRNA靶标。(n=2个技术重复;柱代表平均值±s.e.m.), 和 (B) RPA-C2c2反应能够进行单分子DNA检测 (n=4个技术重复;柱代表平均值±s.e.m.)。

[0056] 图28提供证实根据某些示例实施方式产生的C2c2信号可以在纸基质上检测20pM靶标的一组图。

[0057] 图29提供显示特定RNA酶抑制剂能够在纸上去除背景信号的图。

[0058] 图30是显示使用根据某些示例实施方式的系统在玻璃纤维基质上进行检测的一组图。

[0059] 图31提供一组图,提供 (A) 根据某些示例实施方式的寨卡RNA检测的示意图。用慢病毒包装被C2c2并行检测靶向的寨卡RNA或同源登革热RNA片段。48小时后收获培养基并进行热裂解、RT-RPA和C2c2检测。(B) RT-RAP-C2c2检测能够高度灵敏地检测寨卡慢病毒颗粒 (n=4个技术重复,双尾Student t-检验);****,P<0.0001;柱代表平均值±s.e.m.) (C) 根据某些示例实施方式,使用在纸上的冷冻干燥C2c2进行寨卡RNA检测的示意图。(D) 基于纸的测定能够高度灵敏地检测寨卡慢病毒颗粒 (n=4个技术重复,双尾Student t-检验);****,P<0.0001;** ,p<0.01,柱代表平均值±s.e.m.)。

[0060] 图32提供一组图,证实 (A) 从人血清中分离的寨卡RNA的C2c2检测示意图。对血清中的寨卡RNA进行逆转录、RNA的RNA酶H降解、cDNA的RPA、和C2c2检测。(B) C2c2能够高度灵敏地检测人寨卡血清样品。通过qPCR验证显示的寨卡RNA的浓度 (n=4个技术重复,双尾Student t-检验);****,p<0.0001;柱代表平均值±s.e.m.)。

[0061] 图33提供一组图,证实 (A) 冷冻干燥的C2c2能够灵敏地检测低飞摩尔范围内的ssRNA 1。C2c2能够以液体形式 (B) 或冷冻干燥 (C) 快速检测纸上的200pM ssRNA 1靶标。该反应能够灵敏地检测溶液中 (D) (n=3) 和冷冻干燥形式 (E) (n=3) 的合成寨卡RNA片段。(F) 人寨卡cDNA检测的定量曲线,显示输入浓度与检测到的荧光之间的显著相关性。(G) 在不同量的人血清存在下进行的ssRNA 1的C2c2检测 (n=2个技术重复,除非另有说明;柱代表平

均值 \pm s.e.m.)。

[0062] 图34提供(A)使用通用V3RPA引物组对,对来自细菌基因组的16S rRNA基因的C2c2检测的示意图,和(B)使用根据某些示例实施方式的测定,实现大肠杆菌或铜绿假单胞菌gDNA的灵敏且特异性的检测的能力($n=4$ 个技术重复,双尾Student t-检验;****, $p<0.0001$;柱代表平均值 \pm s.e.m.)。Ec,大肠杆菌;Kp,肺炎克雷伯氏菌;Pa,铜绿假单胞菌;Mt,结核杆菌;Sa,金黄色葡萄球菌。

[0063] 图35提供一组图,证实(A)来自四个不同的肺炎克雷伯氏菌的临床分离物的两个不同的碳青霉烯抗性基因(KPC和NDM-1)的检测,和(B)碳青霉烯抗性基因的检测(A部分)被标准化为KPC与NDM-1crRNA测定之间的信号的比率($n=2$ 个技术重复,双尾Student t-检验;****, $p<0.0001$;柱代表平均值 \pm s.e.m.)。

[0064] 图36提供了一组图,证实(A)C2c2对单错配不灵敏,但是当加载具有另外的错配的crRNA时,可以在靶标中的单核苷酸差异之间作出区分。用11个crRNA检测ssRNA靶标1-3,10个间隔子在crRNA的不同位置处含有合成错配。错配间隔子未显示靶标1的减少的切割,但显示错配靶标2和3的抑制的切割(SEQ.I.D.No.146至159)。(B)示意图显示合理设计具有合成错配的单碱基特异性间隔子的过程。合成错配置于感兴趣的SNP或碱基附近(SEQ.I.D.No.160至164)。(C)株SNP的高度特异性检测允许使用具有截短的(23个核苷酸)crRNA的C2c2检测来区分仅相差一个核苷酸的寨卡非洲RNA靶标相对于寨卡美洲RNA靶标($n=2$ 个技术重复,单尾Student t-检验;* $p<0.05$;****, $p<0.0001$;柱代表平均值 \pm s.e.m.)。

[0065] 图37提供了一组图,证实:(A)用于检测的寨卡株靶区域和crRNA序列的示意图(SEQ.I.D.No.165至170)。靶标中的SNP以红色或蓝色高亮显示,并且指导序列中的合成错配为红色。(B)株SNP的高度特异性检测允许使用SHERLOCK区分寨卡非洲RNA靶标相对于寨卡美洲RNA靶标($n=2$ 个技术重复,双尾Student t-检验;****, $p<0.0001$;柱代表平均值 \pm s.e.m.)(SEQ.I.D.No.171-176)。(C)用于检测的登革热株靶区域和crRNA序列的示意图。靶标中的SNP以红色或蓝色高亮显示,并且指导序列中的合成错配为红色。(D)株SNP的高度特异性检测允许使用SHERLOCK区分登革热株1RNA靶标相对于登革热株3RNA靶标($n=2$ 个技术重复,双尾Student t-检验;****, $p<0.0001$;柱代表平均值 \pm s.e.m.)。

[0066] 图38提供一组图,显示(A)显示用C2c2检测的人SNP的位置的圆圈图。(B)根据某些示例实施方式进行的测定可以区分人SNP。SHERLOCK可以在人类基因组中的四个不同SNP位点正确地对四个不同个体进行基因分型。每个个体的基因型和等位基因感测crRNA的身份在每个图下方注释($n=4$ 个技术重复;双尾Student t-检验;* $p<0.05$;** $p<0.01$;*** $p<0.001$;**** $p<0.0001$;柱代表平均值 \pm s.e.m.)。(C)根据某些示例实施方式检测cfDNA的方法的示意图(例如癌症突变的无细胞DNA检测)。(D)用于检测EGFR L858R和BRAF V600E的示例crRNA序列。(SEQ.I.D.No.177至182)。测定无细胞DNA中的癌症突变的两个基因组基因座的序列。显示的是其中SNP为蓝色高亮的靶标基因组序列和合成错配为红色的突变/野生型感测crRNA序列。

[0067] 图39提供一组图,证实C2c2可以检测来自EGFR L858R(C)或BRAF V600E(B)次要等位基因的模拟无细胞DNA样品中的突变次要等位基因。($n=4$ 个技术重复,双尾Student t-检验;* $p<0.05$;** $p<0.01$,****, $P<0.0001$;柱代表 \pm s.e.m.)

[0068] 图40提供一组图,证实(A)该测定可以区分rs5082处的基因型($n=4$ 个技术重

复;*, $p<0.05$;**, $p<0.01$;***, $p<0.001$;****, $p<0.0001$;柱代表平均值 \pm s.e.m.)。(B)该测定可以区分直接来自离心、变性和煮沸的唾液的gDNA在rs601338处的基因型($n=3$ 个技术重复;*, $p<0.05$;柱代表平均值 \pm s.e.m.)。

[0069] 图41提供(A)在仅与ssDNA 1相差单错配的靶标的背景中对ssDNA 1进行的示例实施方式的示意图。(B)该测定在错配背景存在下实现ssDNA 1的单核苷酸特异性检测(靶标与ssDNA仅相差单错配)。将各种浓度的靶标DNA与背景过量的具有一个错配的DNA组合,并通过该测定检测。

[0070] 图42是显示具有不同染料Cy5的掩蔽构建体也允许有效检测的图。

[0071] 图43是基于金纳米颗粒比色的测定的示意图。使用DNA接头和RNA桥的组合聚合AuNP。在添加RNA酶活性后,ssRNA桥被切割并且AuNP被释放,产生朝向红色的特征性颜色偏移。

[0072] 图44是显示检测520nm处的分散纳米颗粒的颜色偏移的能力的图。纳米颗粒是基于图43中所示的示例实施方式,并且使用不同浓度的RNA酶A的添加进行分散。

[0073] 图45是图,其显示RNA酶比色测试是定量的。

[0074] 图46是微孔板的照片,显示分散的纳米颗粒中的颜色偏移是视觉上可检测的。

[0075] 图47是图,其证实比色偏移在纸基质上是可见的。测试在37°C下在玻璃纤维934-AH上进行10分钟。

[0076] 图48是根据某些示例实施方式用于检测蛋白质或小分子的构象转换适体的示意图。连接产物(B)用作RNA靶向效应子的完整靶标,其不可检测未连接的输入产物(SEQ.I.D.No.202和424)。

[0077] 图49是凝胶图像,显示基于适体的连接可以产生RPA可检测底物。将适体与不同水平的凝血酶一起孵育,然后与探针连接。连接构建体用作3分钟RPA反应的模板。500nM凝血酶具有显著高于背景的扩增靶标的水平。

[0078] 图50显示所选C2c2直系同源物的HEPN结构域的氨基酸序列(SEQ.I.D.No.204-233)。

[0079] 图51采用RPA扩增的RNA的Cas13a检测(SHERLOCK)可以检测到浓度低至 ~ 2 aM的ssRNA靶标,比单独的Cas13a更灵敏($n=4$ 个技术重复;柱代表平均值 \pm s.e.m.)。

[0080] 图52Cas13a检测可用于感测病毒和细菌病原体。(A)从人临床样品中分离的ZIKV RNA的SHERLOCK检测的示意图。(B)SHERLOCK能够高度灵敏地检测人ZIKV阳性血清(S)或尿(U)样品。通过qPCR测定所示ZIKV RNA的近似浓度。($n=4$ 个技术重复,双尾Student t-检验;****, $p<0.0001$;柱代表平均值 \pm s.e.m.;n.d.,未检测到)。

[0081] 图53通过采用引物组2的NASBA(图11)与SHERLOCK检测ssRNA 1的比较。($n=2$ 个技术重复;柱代表平均值 \pm s.e.m.)

[0082] 图54用RPA和单反应SHERLOCK进行核酸扩增。(A)用于图1C所用稀释的ssRNA 1的数字液滴PCR定量。基于ddPCR结果用于稀释的调整浓度显示在柱状图上方。(B)用于图1D所用稀释液的ssDNA 1的数字液滴PCR定量。基于ddPCR结果用于稀释的调整浓度显示在柱状图上方。(C)RPA、T7转录和Cas13a检测反应在同时孵育时相容并实现DNA 2的单分子检测。($n=3$ 个技术重复,双尾Student t-检验;n.s.,无显著性;**, $p<0.01$;****, $p<0.0001$;柱代表平均值 \pm s.e.m.)。

[0083] 图55SHERLOCK与其他灵敏核酸检测工具的比较。(A)采用数字液滴PCR的ssDNA 1稀释系列的检测分析(n=4次技术重复,双尾Student t-检验;n.s.,无显著性;*, $p < 0.05$;**, $p < 0.01$;****, $p < 0.0001$;红线代表均值,柱代表均值 \pm s.e.m.。未显示测量拷贝/ μ L低于 10^{-1} 的样品)。(B)采用定量PCR的ssDNA 1稀释系列的检测分析(n=16个技术复制,双尾Student t-检验;n.s.,不显著;**, $P < 0.01$;****, $p < 0.0001$;红线代表平均值,柱代表平均值 \pm s.e.m.。未显示相对信号低于 10^{-10} 的样品)。(C)采用具有SYBR Green II的RPA的ssDNA 1稀释系列的检测分析(n=4个技术重复,双尾Student t-检验;*, $p < 0.05$;**, $p < 0.01$;红线代表平均值,柱代表平均值 \pm s.e.m.。未显示相对信号低于100的样品)。(D)采用SHERLOCK的ssDNA 1稀释系列的检测分析(n=4个技术重复,双尾Student t-检验;**, $p < 0.01$;****, $p < 0.0001$;红线代表平均值,柱代表平均值 \pm s.e.m.。未显示相对信号低于100的样品)。(E)四种类型的检测方法的ssDNA1稀释系列的百分比变异系数。(F)四种类型的检测方法的 $6e2$ 、 $6e1$ 、 $6e0$ 和 $6e-1$ ssDNA 1稀释的平均百分比变异系数(柱代表平均值 \pm s.e.m.)。

[0084] 图56临床细菌分离物中的碳青霉烯抗性的检测。来自肺炎克雷伯氏菌和大肠杆菌对照的五个临床分离物的两个不同的碳青霉烯抗性基因(KPC和NDM-1)的检测(n=4个技术重复,双尾Student t-检验;****, $p < 0.0001$;柱代表平均值 \pm s.e.m.;n.d.,未检测到)。

[0085] 图57LwCas13a对靶序列中的截短间隔子和单错配的灵敏度的表征。(A)在(B)-(G)中使用的截短间隔子crRNA的序列(SEQ.ID.No.425-436)。还显示ssRNA 1和2的序列,其具有以红色高亮的单个碱基对差异。显示含有合成错配的crRNA,其错配位置用红色着色。(B)在1-7位处具有合成错配的28nt间隔子crRNA对ssRNA1和2的并行切割活性(n=4个技术重复;柱代表平均值 \pm s.e.m.)。(C)(B)中测试的crRNA的特异性比率。特异性比率计算为中靶RNA(ssRNA1)并行切割与脱靶RNA(ssRNA 2)并行切割的比率(n=4个技术重复;柱代表平均值 \pm s.e.m.)。(D)在1-7位处具有合成错配的23nt间隔子crRNA对ssRNA 1和2的并行切割活性(n=4个技术重复;柱代表平均值 \pm s.e.m.)。(E)(D)中测试的crRNA的特异性比率。特异性比率计算为中靶RNA(ssRNA 1)并行切割与脱靶RNA(ssRNA 2)并行切割的比率(n=4个技术重复;柱代表平均值 \pm s.e.m.)。(F)在1-7位处具有合成错配的20nt间隔子crRNA对ssRNA 1和2的并行切割活性(n=4个技术重复;柱代表平均值 \pm s.e.m.)。(G)(F)中测试的crRNA的特异性比率。特异性比率计算为中靶RNA(ssRNA 1)并行切割与脱靶RNA(ssRNA 2)并行切割的比率(n=4个技术重复;柱代表平均值 \pm s.e.m.)。

[0086] 图58靶序列中的理想合成错配位置相对于突变的鉴定。(A)用于评估理想合成错配位置以检测ssRNA 1与ssRNA之间的突变的序列(SEQ.I.D.No.437-462)。在每个靶标上,测试在有色(红色)位置处具有合成错配的crRNA。设计每组合成错配crRNA,使得突变位置相对于间隔子的序列位置偏移。设计间隔子使得在间隔子内的3、4、5和6位处评估突变。(B)在不同位置具有合成错配的crRNA对ssRNA 1和2的并行切割活性。存在四组crRNA,突变在间隔子:靶标双链体区域内的3、4、5或6位处(n=4个技术重复;柱代表平均值 \pm s.e.m.)。(C)(B)中测试的crRNA的特异性比率。特异性比率计算为中靶RNA(ssRNA 1)并行切割与脱靶RNA(ssRNA 2)并行切割的比率(n=4个技术重复;柱代表平均值 \pm s.e.m.)。

[0087] 图59在另外的基因座处使用SHERLOCK进行基因分型,并从煮沸的唾液直接进行基因分型。SHERLOCK可以区分直接来自离心的、变性和煮沸的唾液的基因组DNA的

rs601338SNP位点处的基因型 (n=4个技术重复, 双尾Student t-检验; **, p<0.01; ****, p<0.001; 柱代表平均值±s.e.m.)。

[0088] 图60开发合成基因分型标准品以准确地对人类SNP进行基因分型。(A) 相比于PCR扩增的基因型标准品, 使用SHERLOCK对四个个体的每个在rs601338SNP位点处进行基因分型 (n=4个技术重复; 柱代表平均值±s.e.m.)。(B) 相比于PCR扩增的基因型标准品, 使用SHERLOCK对四个个体的每个在rs4363657SNP位点处进行基因分型 (n=4个技术重复; 柱代表平均值±s.e.m.)。(C) 在rs601338SNP位点处每个个体的SHERLOCK结果与合成标准品之间的计算p值的热图。显示了每个等位基因感测crRNA的热图。热图颜色图被缩放, 使得非显著性 (p>0.05) 为红色, 显著性 (p<0.05) 为蓝色 (n=4个技术重复, 单因素方差分析)。(D) 在rs4363657SNP位点处每个个体的SHERLOCK结果与合成标准品之间的计算p值的热图。显示了每个等位基因感测crRNA的热图。热图颜色图被改变大小, 使得非显著性 (p>0.05) 为红色, 显著性 (p<0.05) 为蓝色 (n=4个技术重复, 单因素方差分析)。(E) 理解SHERLOCK基因分型的p值热图结果的指南。通过选择对应于个体与等位基因合成标准品之间p值>0.05的等位基因可以容易地调用基因分型。红色区块对应于合成标准品与个体的SHERLOCK结果之间的非显著差异, 因此对应于基因型阳性结果。蓝色区块对应于合成标准品与个体的SHERLOCK结果之间的显著差异, 因此对应于基因型阴性结果。

[0089] 图61作为错配背景靶标的小分数的ssDNA 1的检测。在人类基因组DNA的背景上的ssDNA 1的稀释系列的SHERLOCK检测。注意, 在被检测的ssDNA 1靶标与背景基因组DNA之间应该没有序列相似性 (n=2个技术重复; 柱代表平均值±s.e.m.)。

[0090] 图62将来自具有寨卡病毒的患者尿液 (A) 或血清 (B) 样品在95°C (尿液) 或65°C (血清) 热灭活5分钟。根据示例性实施方式, 使用1微升灭活的尿液或血清作为2小时RPA反应的输入, 然后进行3小时C2c2/Cas13a检测反应。误差条表示基于用于检测反应的n=4个技术重复的1SD。

[0091] 图63将来自具有寨卡病毒的患者尿液样品在95°C热灭活5分钟。根据示例性实施方式, 使用1微升灭活的尿液作为30分钟RPA反应的输入, 然后进行3小时 (A) 或1小时 (B) C2c2/Cas13检测反应。误差条表示基于用于检测反应的n=4个技术重复的1SD。

[0092] 图64将来自具有寨卡病毒的患者尿液样品在95°C热灭活5分钟。根据示例性实施方式, 使用1微升灭活的尿液作为20分钟RPA反应的输入, 然后进行1小时C2c2/Cas13a检测反应。误差条表示基于用于检测反应的n=4个技术重复的1SD。

[0093] 图65将来自具有寨卡病毒的患者尿液样品在95°C热灭活5分钟。使用1微升灭活的尿液作为20分钟RPA反应的输入, 然后在存在或不存在指导RNA的情况下进行1小时C2c2/Cas13a检测反应。通过从含有指导的检测反应中减去无指导检测反应的平均荧光值而对数据进行标准化。健康人尿液用作阴性对照。误差条表示基于用于检测反应的n=4个技术重复的1SD。

[0094] 图66显示根据示例性实施方式, 用四种不同指导RNA设计检测两种疟疾特异性靶标 (SEQ. I. D. No. 463-474)。

[0095] 图67提供显示不同Cas13b直系同源物的编辑偏好的图表。密钥见表3。

[0096] 图68提供A) 使用具有不同编辑偏好的不同Cas13b直系同源物的多重测定的示意图, 和B) 证实使用Cas13b10和Cas13b5的这样的测定的可行性的数据。

[0097] 图69提供显示采用Cas13b5 (*Prevotella* sp.MA2106) 和Cas13b9 (*Prevotella intermedia*) 直系同源物的双多重化的图。效应蛋白和指导序列两者包含在同一反应中,允许在同一反应中使用不同荧光读出(聚U 530nm和聚A 485nm)进行双多重化。

[0098] 图70提供与图69相同,但在这种情况下使用Cas13a (*Leptorichia wadei* LwaCas13a) 直系同源物和Cas13b直系同源物 (*Prevotella* sp.MA2016,Cas13b5)。

[0099] 图71提供根据某些示例实施方式,用多个指导序列铺设靶序列以确定靶向的稳健性的方法 (SEQ.I.D.No.475和476)。

[0100] 图72提供杂交链式反应 (HCR) 凝胶,其显示Cas13效应蛋白可用于解锁引发剂,例如,引入如本文所述的掩蔽构建体中的引发剂,以激活杂交链式反应。

[0101] 图73提供显示在复杂裂解物中检测铜绿假单胞菌的能力的数据。

[0102] 图74提供显示根据某些示例性实施方式的某些Cas13直系同源物的离子偏好的数据。所有靶标浓度均为20nM输入,离子浓度为(1mM和10mM)。

[0103] 图75提供显示Cas13b12具有1mM硫酸锌切割偏好的数据。

[0104] 图76提供显示缓冲液优化可以增强Cas13b5对聚A报告子的信号比的数据。旧缓冲液包含40mM Tris-HCL,60mM NaCl,6mM MgCl₂,pH 7.3。新缓冲液包含20mM HEPES pH 6.8,6mM MgCl₂和60mM NaCl。

[0105] 图77提供VI-A/C型Crispr系统和VI-B1和B2型系统的示意图以及代表性Cas13b直系同源物的系统发生树。

[0106] 图78提供各种Cas13b直系同源物在不同核苷酸处和相对于LwCas13a的相对切割活性。

[0107] 图79提供显示各种实例Cas13直系同源物的相对灵敏度的图。

[0108] 图80提供显示使用示例性实施方式实现仄普托 (zepto) 摩尔 (zM) 检测水平的能力的图。

[0109] 图81提供使用具有不同编辑偏好的Cas13直系同源物和基于聚N的掩蔽构建体的多重测定的示意图。

[0110] 图82提供显示在一系列条件下的引物优化实验和假单胞菌检测的结果的数据。

[0111] 图83说明RNA指导的RNA靶向酶的Cas13b家族的生化表征和增加的灵敏度和定量SHERLOCK。A) CRISPR-Cas13基因座和crRNA结构的示意图。B) 用由A、C、G或U碱基的六聚体均聚物组成的传感器探针靶向ssRNA 1的15种Cas13b直系同源物的碱基偏好的热图。C) LwaCas13a和PsmCas13b的切割基序偏好发现筛选和优选的双碱基基序的示意图。在热图中表示的值的是所有耗尽基序中每个双碱基的计数。如果 $-1\log_2$ (靶标/无靶标) 值在LwaCas13a条件下高于1.0或在PsmCas13b条件下高于0.5,则基序被认为耗尽。在 $-1\log_2$ (靶标/无靶标) 值中,靶标和无靶标分别表示在靶标和无靶标条件中的基序的频率。D) 用U₆或A₆传感器探针靶向ssRNA 1的PsmCas13b和LwaCas13a的正交碱基偏好。E) 用LwaCas13a和PsmCas13b靶向登革热ssRNA靶标的单分子SHERLOCK检测。F) 在大反应体积中用LwaCas13a和PsmCas13b进行单分子SHERLOCK检测,用于增加的靶向ssRNA靶标1的灵敏度。G) 在各种RPA引物浓度下定量铜绿假单胞菌合成DNA。H) 铜绿假单胞菌合成DNA浓度与检测到的荧光的相关性。

[0112] 图84说明采用正交Cas13酶的样品内多重SHERLOCK。A) 使用正交Cas13酶的样品内

多重的示意图。B) 用LwaCas13a和PsmCas13b并行活性对20nM寨卡和登革热合成RNA进行样品内多重检测。C) 用LwaCas13a和PsmCas13b在逐渐降低的浓度的金黄色葡萄球菌热核酸酶和铜绿假单胞菌酰基转移酶合成靶标下进行样品内多重RPA和并行检测。D) 用LwaCas13a和CcaCas13b在rs601338处对人样品进行多重基因分型。E) 用于检测疾病等位基因、使用REPAIR进行校正和对REPAIR校正进行评估的治疗诊断时间表的示意图。F) 用LwaCas13a和PsmCas13b对来自健康和疾病模拟样品的APC等位基因进行样品内多重检测。G) 在靶向的APC突变处的REPAIR编辑效率的定量。H) 用LwaCas13a和PsmCas13b对来自REPAIR靶向和非靶向样品的APC等位基因进行样品内多重检测。

[0113] 图85提供纯化并且评估体外并行活性的15种Cas13b直系同源物的树。显示Cas13b基因(蓝色)、Csx27/Csx28基因(红色/黄色)和CRISPR阵列(灰色)。

[0114] 图86说明Cas13直系同源物的蛋白质纯化。A) 本研究中使用的Cas13b、LwCas13a和LbaCas13a的尺寸排阻色谱的色谱图。测量的UV吸光度(mAU)相对于洗脱体积(ml)显示。B) 纯化的Cas13b直系同源物的SDS-PAGE凝胶。从左到右加载14种Cas13b直系同源物。蛋白质梯子显示在左侧。C) LbaCas13a稀释(右)和BSA标准滴定(左)的最终SDS-PAGE凝胶。显示BSA的5个稀释的和LbaCas13的2个稀释。

[0115] 图87显示说明Cas13b直向同源并行切割的碱基偏好的图。A) 使用6个核苷酸长的均聚物腺嘌呤传感器靶向ssRNA 1的14种Cas13b直系同源物的切割活性。B) 使用6个核苷酸长的均聚物尿苷传感器靶向ssRNA 1的14种Cas13b直系同源物的切割活性。C) 使用6个核苷酸长的均聚物鸟嘌呤传感器靶向ssRNA 1的14种Cas13b直系同源物的切割活性。D) 使用6个核苷酸长的均聚物胞苷传感器靶向ssRNA 1的14种Cas13b直系同源物的切割活性。

[0116] 图88显示Cas13并行切割后的随机基序文库的大小分析。LwaCas13a、PsmCas13b、CcaCas13b和RNA酶A处理的文库样品的生物分析仪痕迹显示RNA酶活性后文库大小的变化。Cas13直系同源物靶向登革热ssRNA并且由于并行切割而切割随机基序文库。标志物标准品显示在第一泳道。

[0117] 图89显示RNA酶切割后各种基序的表示。A) 箱形图显示在5分钟和60分钟时间点, LwaCas13a、PsmCas13b、CcaCas13b和RNA酶A的靶标与无靶标的比率的基序分布。将RNA酶A比率与三种Cas13无靶标条件的平均值进行比较。比率也是两个切割反应重复的平均值。B) 在60分钟时间点, LwaCas13a、PsmCas13b、CcaCas13b和RNA酶A的富集基序的数量。富集基序计算为高于1 (LwaCas13a、CcaCas13b和RNA酶A) 或0.5 (PsmCas13b) 的 $-\log_2$ (靶标/无靶标) 阈值的基序。阈值为1对应于至少50%的耗尽, 而阈值为0.5对应于至少30%的耗尽。C) 由LwaCas13a、PsmCas13b和CcaCas13b的富集基序产生的序列标志。LwaCas13a和CcaCas13b显示出如所预期的强U偏好, 而PsmCas13b显示对于基序中的A碱基的独特偏好, 这与同聚物并行活性偏好一致。D) 显示LwaCas13a、PsmCas13b和CcaCas13b的正交基序偏好的热图。热图中表示的值是各个显示的基序的 $-\log_2$ (靶标/无靶标) 值。在 $-\log_2$ (靶标/无靶标) 值中, 靶标和无靶标分别表示在靶标和无靶标条件下的基序的频率。

[0118] 图90显示由随机基序文库筛选确定的RNA酶的单碱基和双碱基偏好。A) 显示在60分钟时间点由随机基序文库切割筛选所确定的LwaCas13a、PsmCas13b、CcaCas13b和RNA酶A的单碱基偏好的热图。在热图中表示的值是所有耗尽的基序中每个碱基的计数。如果 $-\log_2$ (靶标/无靶标) 值在LwaCas13a、CcaCas13b和RNA酶A条件下高于1.0或在PsmCas13b条件下

高于0.5,则基序被认为耗尽。在 $-\log_2$ (靶标/无靶标)值中,靶标和无靶标分别表示在靶标和无靶标条件下的基序的频率。B)显示由随机基序文库切割筛选确定的CcaCas13b的双碱基偏好的热图。在热图中表示的值是所有耗尽的基序中每个双碱基的计数。如果 $-\log_2$ (靶标/无靶标)值在LwaCas13a、CcaCas13b和RNA酶A条件下高于1.0或在PsmCas13b条件下高于0.5,则基序被认为耗尽。在 $-\log_2$ (靶标/无靶标)值中,靶标和无靶标分别表示在靶标和无靶标条件下的基序的频率。C)显示由随机基序文库切割筛选确定的RNA酶A的双碱基偏好的热图。在热图中表示的值是所有耗尽的基序中每个双碱基的计数。如果 $-\log_2$ (靶标/无靶标)值在LwaCas13a、CcaCas13b和RNA酶A条件下高于1.0或在PsmCas13b条件下高于0.5,则基序被认为耗尽。在 $-\log_2$ (靶标/无靶标)值中,靶标和无靶标分别表示在靶标和无靶标条件下的基序的频率。

[0119] 图91说明由随机基序文库筛选确定的RNA酶的三碱基偏好。热图显示在60分钟时间点由随机基序文库切割筛选确定的LwaCas13a、PsmCas13b、CcaCas13b和RNA酶A的三碱基偏好。在热图中表示的值是所有耗尽的基序中每个三碱基的计数。如果 $-\log_2$ (靶标/无靶标)值在LwaCas13a、CcaCas13b和RNA酶A条件下高于1.0或在PsmCas13b条件下高于0.5,则基序被认为耗尽。在 $-\log_2$ (靶标/无靶标)值中,靶标和无靶标分别表示在靶标和无靶标条件下的基序的频率。

[0120] 图92说明由随机基序文库筛选确定的RNA酶的四碱基偏好。热图显示在60分钟时间点由随机基序文库切割筛选确定的LwaCas13a、PsmCas13b、CcaCas13b和RNA酶A的四碱基偏好。在热图中表示的值是所有耗尽的基序中每个四碱基的计数。如果 $-\log_2$ (靶标/无靶标)值在LwaCas13a、CcaCas13b和RNA酶A条件下高于1.0或在PsmCas13b条件下高于0.5,则基序被认为耗尽。在 $-\log_2$ (靶标/无靶标)值中,靶标和无靶标分别表示在靶标和无靶标条件下的基序的频率。

[0121] 图93显示用聚X底物的体外切割测试Cas13直系同源物的碱基切割偏好的结果。A)与crRNA或不与crRNA一起孵育的LwaCas13a对聚U、C、G和A靶标的体外切割。B)与crRNA或不与crRNA一起孵育的CcaCas13b对聚U、C、G和A靶标的体外切割。C)与crRNA或不与crRNA一起孵育的PsmCas13b对聚U、C、G和A靶标的体外切割。

[0122] 图94显示PsmCas13b切割活性的缓冲液优化的结果。A)在靶向ssRNA 1后,测试多种缓冲液对PsmCas13b并行活性的影响。B)在不同的PsmCas13b-crRNA复合物浓度下将优化的缓冲液与原始缓冲液进行比较。

[0123] 图95说明Cas13直系同源物的并行切割的离子偏好。A)对于二价阳离子Ca、Co、Cu、Mg、Mn、Ni和Zn,采用荧光聚U传感器的PsmCas13b的切割活性。将PsmCas13b与靶向合成登革热ssRNA的crRNA一起孵育。B)对于二价阳离子Ca、Co、Cu、Mg、Mn、Ni和Zn,采用荧光聚A传感器的PsmCas13b的切割活性。将PsmCas13b与靶向合成登革热ssRNA的crRNA一起孵育。C)对于二价阳离子Ca、Co、Cu、Mg、Mn、Ni和Zn,采用荧光聚U传感器的Pin2Cas13b的切割活性。将Pin2Cas13b与靶向合成登革热ssRNA的crRNA一起孵育。D)对于二价阳离子Ca、Co、Cu、Mg、Mn、Ni和Zn,采用荧光聚A传感器的Pin2Cas13b的切割活性。将Pin2Cas13b与靶向合成登革热ssRNA的crRNA一起孵育。E)对于二价阳离子Ca、Co、Cu、Mg、Mn、Ni和Zn,采用荧光聚U传感器的CcaCas13b的切割活性。将CcaCas13b与靶向合成登革热ssRNA的crRNA一起孵育。F)对于二价阳离子Ca、Co、Cu、Mg、Mn、Ni和Zn,采用荧光聚A传感器的CcaCas13b的切割活性。将

CcaCas13b与靶向合成登革热ssRNA的crRNA一起孵育。

[0124] 图96显示具有腺嘌呤切割偏好的Cas13直系同源物的切割活性的比较。A) 与不同浓度的各自的靶向合成寨卡靶标的crRNA一起孵育的PsmCas13b和LbaCas13a的切割活性 ($n=4$ 个技术重复, 双尾Student *t*-检验; n.s., 不显著; *, $p<0.05$; **, $p<0.01$; ***, $p<0.001$; ****, $p<0.0001$; 柱代表平均值 \pm s.e.m.)。B) 与不同浓度的各自的靶向合成登革热靶标的crRNA一起孵育的PsmCas13b和LbaCas13a的切割活性 ($n=4$ 个技术重复, 双尾Student *t*-检验; n.s., 不显著; *, $p<0.05$; **, $p<0.01$; ***, $p<0.001$; ****, $p<0.0001$; 柱代表平均值 \pm s.e.m.)。

[0125] 图97说明采用SHERLOCK使用LwaCas13a和PsmCas13b对寨卡ssRNA靶标4的阿摩尔检测。A) 用LwaCas13a和聚U传感器对不同浓度的寨卡ssRNA的SHERLOCK检测。B) 用PsmCas13b和聚A传感器对不同浓度的寨卡ssRNA的SHERLOCK检测。

[0126] 图98说明采用SHERLOCK在不同浓度的CcaCas13b下对登革热ssRNA的阿摩尔检测。

[0127] 图99用铺设ssRNA 1的crRNA测试Cas13直系同源物重编程性。A) 采用铺设ssRNA 1的crRNA的LwaCas13a和CcaCas13b的切割活性。B) 采用铺设ssRNA 1的crRNA的PsmCas13b的切割活性。

[0128] 图100显示crRNA间隔子长度对Cas13直系同源切割的影响。A) 采用具有不同间隔子长度的ssRNA 1靶向crRNA的PsmCas13b的切割活性。B) 采用具有不同间隔子长度的ssRNA 1靶向crRNA的CcaCas13b的切割活性。

[0129] 图101说明用于定量SHERLOCK的引物浓度优化。A) 与不同浓度的寨卡RNA靶标和互补crRNA一起孵育的LwaCas13a在RPA引物浓度为480nM下的SHERLOCK动力学曲线。B) 与不同浓度的寨卡RNA靶标和互补crRNA一起孵育的LwaCas13a在RPA引物浓度为240nM下的SHERLOCK动力学曲线。C) 与不同浓度的寨卡RNA靶标和互补crRNA一起孵育的LwaCas13a在RPA引物浓度为120nM下的SHERLOCK动力学曲线。D) 与不同浓度的寨卡RNA靶标和互补crRNA一起孵育的LwaCas13a在RPA引物浓度为24nM下的SHERLOCK动力学曲线。E) 用四种不同的RPA引物浓度: 480nM、240nM、120nM、60nM和24nM对不同浓度的寨卡RNA的SHERLOCK检测。F) 在不同RPA引物浓度下SHERLOCK的减去背景的荧光与寨卡靶RNA浓度之间的平均 R^2 相关性。G) 在10倍稀释系列(黑点)和2倍稀释系列(红点)中对不同浓度的寨卡RNA靶标的定量SHERLOCK检测。使用120nM的RPA引物浓度。

[0130] 图102说明寨卡和登革热靶标的多重检测。A) 使用靶向寨卡ssRNA靶标的LwaCas13a和靶向登革热ssRNA靶标的PsmCas13b进行多重双色检测。两个靶标均为20nM输入。显示的所有数据代表180分钟反应时间点。B) 使用靶向寨卡ssRNA靶标的LwaCas13a和靶向登革热ssRNA靶标的PsmCas13b进行多重双色检测。两个靶标均为200pM输入。C) 用CcaCas13a和PsmCas13b并行活性对20pM寨卡和登革热合成RNA的样品内多重检测。

[0131] 图103图1说明寨卡和登革热ssRNA的样品内多重RNA检测。采用PsmCas13b和CcaCas13b对逐渐降低的浓度的寨卡和登革热合成靶标的样品内多重RPA和并行检测。

[0132] 图104说明突变和REPAIR编辑的非多重治疗诊断检测。A) 用LwaCas13a检测来自健康和疾病模拟样品的APC等位基因。B) 用LwaCas13a检测来自REPAIR靶向和非靶向样品的APC等位基因处的编辑校正。

[0133] 图105说明用金纳米颗粒聚集对RNA酶活性的比色检测。A) RNA酶活性的基于金纳

米颗粒的比色读出的示意图。在没有RNA酶活性的情况下,RNA接头聚集金纳米颗粒,导致红色损失。RNA接头的切割释放纳米颗粒并导致红色变化。B) 在不同单位的RNA酶A下RNA酶消化120分钟后比色报告子的图像。C) 在不同单位浓度的RNA酶A的消化下,AuNP比色报告子在520nm吸收处的动力学。D) 在用不同单位浓度的RNA酶A消化120分钟后,AuNP比色报告子的520nm吸收。E) 在不同单位浓度的RNA酶A的消化下,达到AuNP比色报告子的 A_{520} 最大值的一半的时间。

[0134] 图106说明来自大豆基因组DNA的CP4-EPSPS基因的定量检测。A) 在不同检测时间点,SHERLOCK减去背景的荧光与CP4-EPSPS豆百分比的平均相关 R^2 。豆百分比描述了round-up ready和野生型豆的混合物中的round-up ready豆的量。CP4-EPSPS基因仅存在于round-up ready豆中。B) 在不同豆百分比下的CP4-EPSPS抗性基因的SHLELOCK检测显示在30分钟孵育下SHERLOCK检测的定量性质。C) 在不同豆百分比下对凝集素基因的SHERLOCK检测。豆百分比描述round-up ready和野生型豆的混合物中的round-up ready豆的量。凝集素基因存在于这两种类型的豆中,因此不显示与round-up ready豆百分比的相关。

具体实施方式

[0135] 一般定义

[0136] 除非另有定义,否则本文所用的技术和科学术语具有与本公开所属领域的普通技术人员之一通常理解的相同的含义。常见术语的定义和分子生物学的技术可以在如下找到:Molecular Cloning:A Laboratory Manual,2nd edition(1989)(Sambrook,Fritsch,和Maniatis);Molecular Cloning:A Laboratory Manual,4th edition(2012)(Green和Sambrook);Current Protocols in Molecular Biology(1987)(F.M.Ausubel等eds.);the series Methods in Enzymology(Academic Press,Inc.):PCR 2:A Practical Approach(1995)(M.J.MacPherson,B.D.Hames,和G.R.Taylor eds.):Antibodies,A Laboratory Manual(1988)(Harlow和Lane,eds.):Antibodies A Laboratory Manual,2nd edition2013(E.A.Greenfield ed.);Animal Cell Culture(1987)(R.I.Freshney,ed.);Benjamin Lewin,Genes IX,published by Jones and Bartlet,2008(ISBN 0763752223);Kendrew等(eds.),The Encyclopedia of Molecular Biology,published by Blackwell Science Ltd.,1994(ISBN0632021829);Robert A.Meyers(ed.),Molecular Biology and Biotechnology:a Comprehensive Desk Reference,published by VCH Publishers, Inc.,1995(ISBN 9780471185710);Singleton等,Dictionary of Microbiology and Molecular Biology 2nd ed.,J.Wiley&Sons(New York,N.Y.1994),March,Advanced Organic Chemistry Reactions,Mechanisms and Structure 4th ed.,John Wiley&Sons(New York,N.Y.1992);和Marten H.Hofker和Jan van Deursen,Transgenic Mouse Methods and Protocols,2nd edition(2011)。

[0137] 如本文所使用的,单数形式“一个/一种(a/an)”和“所述/该(the)”包括单数和复数对象,除非上下文另有明确规定。

[0138] 术语“任选的”或“任选地”是指随后描述的事件、情况或取代可以发生或可以不发生,并且该描述包括其中事件或情况发生的例子和其中不发生的例子。

[0139] 由端点表述的数值范围包括在各自范围内包含的所有数字和分数,以及所记载的

端点。

[0140] 本文所使用的术语“约”或“大约”当指向可测量的值,例如参数、量、持续时间等时,意指包括指定值的变化或从指定值变化,例如 $\pm 10\%$ 或更少、 $\pm 5\%$ 或更少、 $\pm 1\%$ 或更少以及 $\pm 0.1\%$ 或更少的指定值的变化或从指定值变化,只要这样的变化适合于在所公开的发明中进行。应当理解,修饰语“约”或“大约”所指的本身也是特定地且优选地公开的。

[0141] 在整个说明书中提到“一个实施方式”,“实施方式”,“示例性实施方式”意味着结合该实施方式描述的特别的特征、结构或特性是包括在本发明的至少一个实施方式中。因此,在整个说明书中在各个地方出现的短语“在一个实施方式中”、“在实施方式中”或“示例性实施方式”不一定都指同一实施方式,而是可以指同一实施方式。此外,在一个或多个实施方式中,特别的特征、结构或特性可以以任何合适的方式组合,如本领域技术人员从本公开中显而易见的。此外,虽然本文描述的一些实施方式包括包括在其他实施方式中的一些但没有其他特征,不同实施方式的特征的组合是意在本发明的范围内。例如,在所附权利要求中,任何要求保护的实施方式可以以任何组合使用。

[0142] “C2c2”现在称为“Cas13a”,并且该等术语在本文中互换使用,除非另有说明。

[0143] 在本文中引用的所有的出版物、公开的专利文件和专利申请通过引用并入本文,其程度如同单独的出版物、公开的专利文件或专利申请各自被特定地和单独地指示为通过引用并入本文。

[0144] 概述

[0145] 微生物规律成簇的间隔短回文重复序列(CRISPR)和CRISPR相关(CRISPR-Cas)适应性免疫系统包括可编程的核酸内切酶,例如Cas9和Cpf1 (Shmakov等,2017;Zetsche等,2015)。虽然Cas9和Cpf1两者靶向DNA,但最近已经发现(Shmakov等,2015)和表征(Abudayyeh等,2016;Smargon等,2017)了单效应子RNA指导的RNA酶,包括C2c2,提供了用于特异性RNA感测的平台。RNA指导的RNA酶可以使用CRISPR RNA(crRNA)容易且方便地重编程以切割靶RNA。与DNA内切酶Cas9和Cpf1不同,其仅切割其DNA靶标,RNA指导的RNA酶,如C2c2,在切割其RNA靶标后保持活性,导致对邻近的非靶向RNA的“并行(collateral)”切割(Abudayyeh等,2016)。这种crRNA编程的并行RNA切割活性提供了通过触发体内程序性细胞死亡或体外非特异性RNA降解(其可以作为读出),利用RNA指导的RNA酶以检测特定RNA的存在的机会(Abudayyeh等,2016;East-Seletsky等,2016)。

[0146] 本文所公开的实施方式利用RNA靶向效应子以提供具有阿摩尔(attomolar)灵敏度的稳健的基于CRISPR的诊断。本文所公开的实施方式可以以相当的水平的灵敏度检测DNA和RNA两者,并且可以基于单碱基对差异将靶标与非靶标区分开。此外,本文所公开的实施方式可以以冷冻干燥的形式制备,以便于分配和护理点(POC)应用。这样的实施方式可用于人类健康的多种情况,包括例如病毒检测、细菌菌株分型、灵敏基因分型,和疾病相关无细胞DNA的检测。为了便于提及,本文所公开的实施方式也可称为SHERLOCK(特异性高灵敏度酶促报告子解锁,Specific High-sensitivity Enzymatic Reporter unLOCKing)。

[0147] 在一个方面,本文所公开的实施方式涉及一种核酸检测系统,其包括检测CRISPR系统、一种或多种设计用于结合相应靶分子的指导RNA、掩蔽构建体,和扩增样品中的靶核酸分子的任意的扩增试剂。在某些示例性实施方式中,系统还可以包括一种或多种检测适

体。一种或多种检测适体可以包含RNA聚合酶位点或引物结合位点。一种或多种检测适体特异性结合一种或多种靶多肽,并且被配置成使得RNA聚合酶位点或引物结合位点仅在检测适体与靶肽结合时暴露。RNA聚合酶位点的暴露有助于使用适体序列作为模板产生触发RNA寡核苷酸。因此,在这样的实施方式中,一种或多种指导RNA被配置为结合触发RNA。

[0148] 在另一个方面,本文所公开的实施方式涉及一种包括多个个体离散体积的诊断装置。个体离散体积各自包含CRISPR效应蛋白、一种或多种设计用于结合相应靶分子的指导RNA和掩蔽构建体。在某些示例性实施方式中, RNA扩增试剂可以预先加载到个体离散体积中或者在将样品添加到各个个体离散体积的同时或之后添加到个体离散体积中。装置可以是基于微流体的装置、可穿戴装置或包括在其上限定个体离散体积的柔性材料基质的装置。

[0149] 在另一个方面,本文所公开的实施方式涉及一种用于检测样品中的多肽的方法,其包括:将样品或样品组分配到个体离散体积的组中,个体离散体积各自包含CRISPR效应蛋白、一种或多种设计用于结合一种靶寡核苷酸的指导RNA,和掩蔽构建体。随后在足以允许一种或多种指导RNA与一种或多种靶分子结合的条件下维持样品组。一种或多种指导RNA与靶分子的结合进而激活CRISPR效应蛋白。一旦被激活,CRISPR效应蛋白随后使掩蔽构建体失活,例如,通过切割掩蔽构建体使得可检测阳性信号被揭露、释放或产生。个体离散体积中的阳性可检测信号的检测指示靶分子的存在。

[0150] 在另一个方面,本文所公开的实施方式涉及用于检测多肽的方法。检测多肽的方法类似于如上所述的检测靶核酸的方法。然而,还包含肽检测适体。肽检测适体如上所述起作用,并且在与靶多肽结合后促进触发寡核苷酸的产生。设计指导RNA以识别触发寡核苷酸,从而激活CRISPR效应蛋白。激活的CRISPR效应蛋白使掩蔽构建体失活导致可检测阳性信号的揭露、释放或产生。

[0151] CRISPR效应蛋白

[0152] 通常,本文和文献中使用的CRISPR-Cas或CRISPR系统,例如W02014/093622 (PCT/US2013/074667), 共同地是指涉及表达CRISPR-相关(“Cas”)基因或指导CRISPR-相关(“Cas”)基因的活性的转录物和其他元件,包括编码以下的序列:Cas基因、tracr(反式激活CRISPR)序列(例如tracrRNA或具有活性的部分tracrRNA)、tracr配对序列(包括在内源CRISPR系统的上下文中的“直接重复”和tracrRNA处理的部分直接重复)、指导序列(在内源CRISPR系统的上下文中也称为“间隔子”),或如本文所用的术语“RNA”(例如,指导Cas,例如Cas9的RNA,例如CRISPR RNA和反式激活(tracr)RNA或单指导RNA(sgRNA)(嵌合RNA))或来自CRISPR基因座的其他序列和转录物。通常,CRISPR系统的特征在于促进CRISPR复合物在靶序列(在内源CRISPR系统的上下文中也称为原型间隔子)的位点处形成的元件。当CRISPR蛋白是C2c2蛋白时,不需要tracrRNA。C2c2已经在Abudayyeh等(2016)“C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector”;*Science*;DOI:10.1126/science.aaf5573;和Shmakov等(2015)“Discovery and Functional Characterization of Diverse Class 2CRISPR-Cas Systems”,*Molecular Cell*,DOI:dx.doi.org/10.1016/j.molcel.2015.10.008中描述,其全部内容通过引用并入本文。Cas13b已经在Smargon等(2017)“Cas13b Is a Type VI-B CRISPR-Associated RNA-Guided RNases Differentially Regulated by Accessory Proteins Csx27and Csx28,”

Molecular Cell.65,1-13;dx.doi.org/10.1016/j.molcel.2016.12.023中描述,其全部内容通过引用并入本文。

[0153] 在某些实施方式中,原型间隔子邻近基序(PAM)或PAM样基序指导本文所公开的效应蛋白复合物与感兴趣的靶基因座的结合。在一些实施方式中,PAM可以是5' PAM(即位于原型间隔子的5'端的上游)。在其他实施方式中,PAM可以是3' PAM(即位于原型间隔子的5'端的下游)。术语“PAM”可与术语“PFS”或“原型间隔子侧翼位点”或“原型间隔子侧翼序列”互换使用。

[0154] 在优选的实施方式中,CRISPR效应蛋白可识别3' PAM。在某些实施方式中,CRISPR效应蛋白可识别3' PAM,其是5' H,其中H为A、C或U。在某些实施方式中,效应蛋白可为 *Leptotrichia shahii* C2c2p,更优选 *Leptotrichia shahii* DSM 19757C2c2,并且3' PAM是5' H。

[0155] 在CRISPR复合物的形成的上下文中,“靶序列”是指指导序列被设计成与其具有互补性的序列,其中靶序列与指导序列之间的杂交促进CRISPR复合物的形成。靶序列可包括RNA多核苷酸。术语“靶RNA”是指RNA多核苷酸,其是靶序列或包含靶序列。换句话说,靶RNA可以是RNA多核苷酸,或者gRNA(即指导序列)的一部分被设计为与其具有互补性并且由包含CRISPR效应蛋白和gRNA的复合物介导的效应子功能被引导至其的RNA多核苷酸的一部分。在一些实施方式中,靶序列位于细胞的细胞核或细胞质中。

[0156] 编码CRISPR效应蛋白的核酸分子,特别是C2c2,有利地是密码子优化的CRISPR效应蛋白。密码子优化序列的例子在这种情况下是被优化用于在真核生物中表达的序列,例如人类(即优化用于在人类中表达),或用于本文讨论的另外的真核生物、动物或哺乳动物;参见例如W02014/093622(PCT/US2013/074667)中的SaCas9人类密码子优化的序列。虽然这是优选的,但应当理解,其他实例是可能的,并且对于除人类以外的宿主物种的密码子优化,或对于特定器官的密码子优化是已知的。在一些实施方式中,编码CRISPR效应蛋白的酶编码序列是用于在特定细胞(例如真核细胞)中表达而优化的密码子。真核细胞可以是衍生自特定生物的那些真核细胞,例如植物或哺乳动物,包括但不限于人类,或非人类真核生物,或本文所讨论的动物或哺乳动物,例如小鼠、大鼠、兔、狗、牲畜或非人类哺乳动物或灵长类动物。在一些实施方式中,用于修饰人的种系遗传特性的方法和/或用于修饰动物的遗传特性的方法,其可能导致它们遭受痛苦而对人或动物没有任何实质性医学益处,以及由这样的方法产生的动物,可以被排除。通常,密码子优化是指通过将天然序列的至少一个密码子(例如,约或多于约1、2、3、4、5、10、15、20、25、50或更多个密码子)替换为在宿主细胞的基因中更频繁或最频繁使用的密码子同时保持天然氨基酸序列来修饰核酸序列以增强在感兴趣的宿主细胞中的表达的过程。各种物种对特定氨基酸的某些密码子表现出特定的偏好。密码子偏好(生物体之间密码子使用的差异)通常与信使RNA(mRNA)的翻译效率相关,而信使RNA(mRNA)的翻译效率又被认为取决于被翻译的密码子的特性和特定的转移RNA(tRNA)分子的可用性。所选择的tRNA在细胞中的优势通常是肽合成中最频繁使用的密码子的反映。因此,可以基于密码子优化来定制基因用于在给定生物中的最佳基因表达。密码子使用表很容易获得,例如,在可在kazusa.or.jp/codon/获得的“密码子使用数据库”,这些表可以通过多种方式进行调整。见Nakamura,Y等“Codon usage tabulated from the international DNA sequence databases:status for the year 2000”Nucl.Acids

Res. 28:292 (2000)。用于密码子优化特定序列以在特定宿主细胞中表达的计算机算法也是可获得的,例如Gene Forge (Aptagen; Jacobus, PA) 也是可获得的。在一些实施方式中,在编码Cas的序列中的一个或多个密码子(例如,1、2、3、4、5、10、15、20、25、50或更多个或所有密码子)对应于对特定氨基酸最频繁使用的密码子。

[0157] 在某些实施方式中,本文所述的方法可以包括提供Cas转基因细胞,特别是C2c2转基因细胞,其中提供或引入在细胞中与包含感兴趣的一种或多种基因的启动子的调控元件可操作地连接的编码一种或多种指导RNA的一种或多种核酸。如本文所使用,术语“Cas转基因细胞”是指细胞,例如真核细胞,其中Cas基因已经基因组整合。根据本发明,细胞的性质、类型或来源不是特别限制的。Cas转基因在细胞中引入的方式也可以变化,并且可以是本领域已知的任何方法。在某些实施方式中,通过在分离的细胞中引入Cas转基因获得Cas转基因细胞。在某些其他实施方式中,通过从Cas转基因生物中分离细胞获得Cas转基因细胞。举例而非限制,本文提及的Cas转基因细胞可以衍生自Cas转基因真核生物,例如Cas敲入真核生物。参考W02014/093622 (PCT/US13/74667),其通过引用并入本文。转让给Sangamo BioSciences, Inc. 的美国专利公开号20120017290和20110265198涉及靶向Rosa基因座的方法可以被修改以利用本发明的CRISPR Cas系统。转让给Cellestis的美国专利公开号20130236946涉及靶向Rosa基因座的方法也可以被修改以利用本发明的CRISPR Cas系统。通过进一步的实例,参考Platt等 (Cell; 159 (2) :440-455 (2014)), 描述了Cas9敲入小鼠,其通过引用并入本文。该Cas转基因还可以包含LOX-Stop-PolyA-LOX (LSL) 盒从而使Cas表达可被Cre重组酶诱导。或者,可以通过在分离的细胞中引入Cas转基因来获得Cas转基因细胞。转基因的递送系统在本领域中是公知的。举例来说,Cas转基因可以通过载体(例如,AAV、腺病毒、慢病毒)和/或颗粒和/或纳米颗粒递送在例如真核细胞中递送,如本文其他地方所述。

[0158] 本领域技术人员将理解,除了具有整合的Cas基因或由当与能够指导Cas至靶基因座的RNA复合时Cas的序列特异性作用引起的突变之外,本文提及的细胞(例如Cas转基因细胞)可以包含进一步的基因组改变。

[0159] 在某些方面,本发明涉及载体,例如用于向细胞中递送或引入Cas和/或能够指导Cas至靶基因座的RNA(即指导RNA),但也用于繁殖这些组分(例如在原核细胞中)。本文使用的“载体”是允许或促进实体从一个环境转移到另一个环境的工具。它是复制子,例如质粒、噬菌体或粘粒,其中可以插入另外的DNA片段以引起插入的片段的复制。通常,当与适当的控制元件相关联时,载体能够复制。通常,术语“载体”是指能够转运已经连接至其的另外的核酸的核酸分子。载体包括但不限于单链、双链或部分双链的核酸分子;包含一个或多个游离末端、没有游离末端(例如环状)的核酸分子;包含DNA、RNA或两者的核酸分子;和本领域已知的其他多核苷酸种类。一种类型的载体是“质粒”,其是指例如通过标准分子克隆技术另外的DNA区段可以插入其中的环状双链DNA环。另一种类型的载体是病毒载体,其中病毒衍生的DNA或RNA序列存在于载体中以包装成病毒(例如逆转录病毒、复制缺陷逆转录病毒、腺病毒、复制缺陷型腺病毒和腺相关病毒(AAV))。病毒载体还包括由病毒携带的多核苷酸用于转染到宿主细胞中。某些载体能够在引入它们的宿主细胞中自主复制(例如,具有细菌复制起点的细菌载体和附加型哺乳动物载体)。其他载体(例如,非附加型哺乳动物载体)在引入宿主细胞后整合到宿主细胞的基因组中,从而与宿主基因组一起复制。此外,某些载体

能够指导它们可操作地连接的基因的表达。这样的载体在本文中称为“表达载体”。在重组DNA技术中有用的常见表达载体通常是质粒的形式。

[0160] 重组表达载体可以包含以适于在宿主细胞中表达核酸的形式本发明的核酸,这意味着重组表达载体包含一种或多种可以基于用于表达的宿主细胞而选择的与待表达的核酸序列可操作地连接的细胞调控元件。在重组表达载体内,“可操作地连接”意指感兴趣的核苷酸序列以允许表达核苷酸序列的方式与调控元件连接(例如在体外转录/翻译系统中,或当载体导入宿主细胞时在宿主细胞中)。关于重组和克隆方法,提及2004年9月2日公布为US2004-0171156A1的美国专利申请10/815,730,其全部内容通过引用并入本文。因此,本文所公开的实施方式还可以包括包含CRISPR效应系统的转基因细胞。在某些示例性实施方式中,转基因细胞可以作为个体离散体积起作用。换言之,可以将包含掩蔽构建体的样品递送至细胞,例如在合适的递送囊泡中,并且如果靶标存在于递送囊泡中,则激活CRISPR效应子并产生可检测信号。

[0161] 载体可包含调控元件,例如启动子。载体可以包含Cas编码序列,和/或单个,但可能也可以包含至少3或8或16或32或48或50个指导RNA(例如,sgRNA)编码序列,例如1-2、1-3、1-4-5、3-6、3-7、3-8、3-9、3-10、3-8、3-16、3-30、3-32、3-48、3-50个RNA(例如,sgRNA)。在单个载体中,可以对每个RNA(例如,sgRNA)有启动子,有利地,当有多达约16个RNA时;并且,当单个载体提供超过16个RNA时,一个或多个启动子可以驱动一个以上RNA的表达,例如,当存在32个RNA时,每个启动子可以驱动两个RNA的表达,并且当存在48个RNA时,每个启动子可以驱动三个RNA的表达。通过简单的算术和完善的克隆方案以及本公开内容的教导,本领域技术人员可以就对于合适的示例性载体(例如AAV)和合适的启动子(例如U6启动子)的RNA容易地实施本发明。例如,AAV的包装限制为~4.7kb。单个U6-gRNA(加上用于克隆的限制性位点)的长度为361bp。因此,技术人员可以容易地在单个载体中装入约12-16个,例如13个U6-gRNA盒。这可以通过任何合适的方式组装,例如用于TALE组装的金门策略(genome-engineering.org/taleffectors/)。本领域技术人员还可以使用串联指导策略将U6-gRNA的数量增加约1.5倍,例如,从12-16例如13增加至约18-24例如约19个U6-gRNA。因此,本领域技术人员可以容易地在单个载体(例如AAV载体)中达到约18-24个,例如约19个启动子-RNA,例如U6-gRNA。用于增加载体中启动子和RNA数量的另一种方法是使用单个启动子(例如,U6)来表达由可切割序列分开的RNA阵列。并且,用于增加载体中启动子-RNA数量的更进一步的方法是表达由编码序列或基因的内含子中的可切割序列分开的启动子-RNA阵列;在这种情况下,使用可以具有增加的表达并且能够以组织特异性方式转录长RNA的聚合酶II启动子是有利的。(参见,例如,nar.oxfordjournals.org/content/34/7/e53.short和nature.com/mt/journal/v16/n9/abs/mt2008144a.html)。在有利的实施方式中,AAV可以包装靶向多达约50个基因的U6串联gRNA。因此,根据本领域的知识和本公开的教导,技术人员可以容易地制造和使用表达在一个或多个启动子的控制下或可操作地或功能性地连接的RNA或指导(特别是关于本文所讨论的RNA和指导的数量)的载体,例如单个载体,而无需任何过度实验。

[0162] 指导RNA编码序列和/或Cas编码序列可以功能性地或可操作地连接至调控元件,并且因此调控元件驱动表达。启动子可以是组成型启动子和/或条件型启动子和/或诱导型启动子和/或组织特异性启动子。启动子可选自RNA聚合酶、pol I、pol II、pol III、T7、U6、

H1、逆转录病毒Rous肉瘤病毒 (RSV) LTR启动子、巨细胞病毒 (CMV) 启动子、SV40启动子、二氢叶酸还原酶启动子、 β -肌动蛋白启动子、磷酸甘油激酶 (PGK) 启动子和EF1 α 启动子。有利的启动子是启动子U6。

[0163] 在一些实施方式中,核酸靶向系统的一个或多个元件衍生自包含内源CRISPR RNA靶向系统的特定生物。在某些示例性实施方式中,效应蛋白CRISPR RNA靶向系统包含至少一个HEPN结构域,包含但不限于本文所述的HEPN结构域、本领域已知的HEPN结构域,以及通过与共有序列基序比较而被认为是HEPN结构域的结构域。本文提供了多个这样的结构域。在一个非限制性实施方式中,共有序列可以衍生自本文提供的C2c2或Cas13b直系同源物的序列。在某些示例性实施方式中,效应蛋白包含单个HEPN结构域。在某些其他示例性实施方式中,效应蛋白包含两个HEPN结构域。

[0164] 在一个示例性实施方式中,效应蛋白包含含有RxxxxH基序序列的一个或多个HEPN结构域。RxxxxH基序序列可以是,但不限于,来自本文所述的HEPN结构域或本领域已知的HEPN结构域。RxxxxH基序序列还包含通过组合两个或更多个HEPN结构域的部分而产生的基序序列。如上所述,共有序列可以从以下中公开的直系同源物的序列中得到:题为“Novel CRISPR Enzymes and Systems”的美国临时专利申请62/432,240,于2017年3月15日提交的题为“Novel Type VI CRISPR Orthologs and Systems”的美国临时专利申请62/471,710,和标记为代理人案卷号47627-05-2133并与2017年4月12日提交的题为“Novel Type VI CRISPR Orthologs and Systems”的美国临时专利申请。

[0165] 在本发明的实施方式中,结构域HEPN包含至少一个包含R{N/H/K}X1X2X3H的序列的RxxxxH基序。在本发明的实施方式中,HEPN结构域包含含有R{N/H}X1X2X3H的序列的RxxxxH基序。在本发明的实施方式中,HEPN结构域包含R{N/K}X1X2X3H的序列。在某些实施方式中,X1是R、S、D、E、Q、N、G、Y或H。在某些实施方式中,X2是I、S、T、V或L。在某些实施方式中,X3是L、F、N、Y、V、I、S、D、E或A。

[0166] 根据本发明使用的其他效应子可以通过它们与cas1基因的接近来鉴定,例如,但不限于,在距离cas1基因的起点20kb和距离cas1基因的末端20kb的区域内。在某些实施方式中,效应蛋白包含至少一个HEPN结构域和至少500个氨基酸,并且其中C2c2效应蛋白天然存在于Cas基因或CRISPR阵列上游或下游20kb内的原核基因组中。Cas蛋白的非限制性实例包括Cas1、Cas1B、Cas2、Cas3、Cas4、Cas5、Cas6、Cas7、Cas8、Cas9(也称为Csn1和Csx12)、Cas10、Csy1、Csy2、Csy3、Cse1、Cse2、Csc1、Csc2、Csa5、Csn2、Csm2、Csm3、Csm4、Csm5、Csm6、Cmr1、Cmr3、Cmr4、Cmr5、Cmr6、Csb1、Csb2、Csb3、Csx17、Csx14、Csx10、Csx16、CsaX、Csx3、Csx1、Csx15、Csf1、Csf2、Csf3、Csf4,其同源物或其修饰版本。在某些示例性实施方式中,C2c2效应蛋白天然存在于原核生物基因组中Cas1基因上游或下游20kb内。术语“直系同源物”(在本文中也称为“直系同源物”)和“同源物”(在本文中也称为“同源物”)是本领域公知的。通过进一步的指导,本文使用的蛋白的“同源物”是相同物种的蛋白,其执行与其是同源物的蛋白相同或相似的功能。同源蛋白可以但不需要在结构上相关,或仅在结构上部分相关。本文使用的蛋白的“直系同源物”是不同物种的蛋白质,其执行与其是直系同源物的蛋白相同或相似的功能。直系同源蛋白可以但不需要在结构上相关,或仅在结构上部分相关。

[0167] 在特定的实施方式中,VI型RNA靶向Cas酶是C2c2。在其他示例性实施方式中,VI型RNA靶向Cas酶是Cas13b。在特定的实施方式中,本文提及的VI型蛋白例如C2c2的同源物或

直系同源物与VI型蛋白例如C2c2具有至少30%、或至少40%、或至少50%、或至少60%、或至少70%、或至少80%、更优选至少85%、甚至更优选至少90%，例如至少95%的序列同源性或同一性(例如，基于Leptotrichia shahii C2c2、Lachnospiraceae bacterium MA2020 C2c2、Lachnospiraceae bacterium NK4A179C2c2、Clostridium aminophilum(DSM 10710) C2c2、Carnobacterium gallinarum(DSM 4847) C2c2、Paludibacter propionicigenes (WB4) C2c2、Listeria weihenstephanensis (FSL R9-0317) C2c2、李斯特细菌(FSL M6-0635) C2c2、Listeria newyorkensis (FSL M6-0635) C2c2、Leptotrichia wadei (F0279) C2c2、Rhodobacter capsulatus (SB 1003) C2c2、Rhodobacter capsulatus (R121) C2c2、Rhodobacter capsulatus (DE442) C2c2、Leptotrichia wadei (Lw2) C2c2或Listeria seeligeri C2c2任一者的野生型序列)。在另外的实施方式中，本文提及的VI型蛋白例如C2c2的同源物或直系同源物与野生型C2c2具有至少30%、或至少40%、或至少50%、或至少60%、或至少70%、或至少80%、更优选至少85%、甚至更优选至少90%，例如至少95%的序列同一性(例如，基于Leptotrichia shahii C2c2、Lachnospiraceae bacterium MA2020 C2c2、Lachnospiraceae bacterium NK4A179C2c2、Clostridium aminophilum(DSM 10710) C2c2、Carnobacterium gallinarum(DSM 4847) C2c2、Paludibacter propionicigenes (WB4) C2c2、Listeria weihenstephanensis (FSL R9-0317) C2c2、李斯特细菌(FSL M6-0635) C2c2、Listeria newyorkensis (FSL M6-0635) C2c2、Leptotrichia wadei (F0279) C2c2、Rhodobacter capsulatus (SB 1003) C2c2、Rhodobacter capsulatus (R121) C2c2、Rhodobacter capsulatus (DE442) C2c2、Leptotrichia wadei (Lw2) C2c2或Listeria seeligeri C2c2任一者的野生型序列)。

[0168] 在某些其他示例性实施方式中，CRISPR系统效应蛋白是C2c2核酸酶。C2c2的活性可以取决于两个HEPN结构域的存在。已经显示这些是RNA酶结构域，即切割RNA的核酸酶(特别是核酸内切酶)。C2c2HEPN也可以靶向DNA，或潜在的DNA和/或RNA。基于C2c2的HEPN结构域至少能够结合并以其野生型形式切割RNA，优选C2c2效应蛋白具有RNA酶功能。关于C2c2CRISPR系统，参考2016年6月17日提交的美国临时申请62/351,662和2016年8月17日提交的美国临时申请62/376,377。还参考2016年6月17日提交的美国临时申请62/351,803。还参考载有Broad Institute No.10035.PA4和代理人案卷号47627.03.2133的于2016年12月8日提交的题为“Novel Crispr Enzymes and Systems”的美国临时申请。进一步参考East-Seletsky等“Two distinct RNase activities of CRISPR-C2c2enable guide-RNA processing and RNA detection”Nature doi:10/1038/nature19802和Abudayyeh等“C2c2is a single-component programmable RNA-guided RNA targeting CRISPR effector”bioRxiv doi:10.1101/054742。

[0169] 在CRISPR系统中的RNA酶功能是已知的，例如mRNA靶向已被报道用于某些III型CRISPR-Cas系统(Hale等,2014,Genes Dev,vol.28,2432-2443;Hale等,2009,Cell,vol.139,945-956;Peng等,2015,Nucleic acids research,vol.43,406-417)并提供了显著的优点。在表皮葡萄球菌III-A型系统中，遍及靶标的转录导致由Cas10-Csm核糖核蛋白效应蛋白复合物内的独立活性位点介导的靶DNA及其转录物的切割(参见Samai等,2015,Cell,vol.151,1164-1174)。因此提供了通过本效应蛋白靶向RNA的CRISPR-Cas系统、组合物或方法。

[0170] 在一个实施方式中,Cas蛋白可以包括但不限于以下的属的生物的C2c2直系同源物:纤毛菌属(*Leptotrichia*)、李斯特菌属(*Listeria*)、棒状杆菌属(*Corynebacter*)、萨特氏菌属(*Sutterella*)、军团菌属(*Legionella*)、密螺旋体属(*Treponema*)、产线菌属(*Filifactor*)、真细菌属(*Eubacterium*)、链球菌属(*Streptococcus*)、乳酸菌属(*Lactobacillus*)、支原体属(*Mycoplasma*)、拟杆菌属(*Bacteroides*)、*Flaviivola*、黄杆菌属(*Flavobacterium*)、*Sphaerochaeta*、固氮螺菌属(*Azospirillum*)、葡糖醋杆菌属(*Gluconacetobacter*)、奈瑟氏菌属(*Neisseria*)、罗氏菌属(*Roseburia*)、*Parvibaculum*、葡萄球菌属(*Staphylococcus*)、*Nitratifactor*、支原体属(*Mycoplasma*)和弯曲杆菌属(*Campylobacter*)。这样的属的生物的种可以如本文中另外讨论的。

[0171] 鉴定CRISPR-Cas系统酶的直系同源物的一些方法可以包括鉴定在感兴趣的基因组中的序列。序列的鉴定可以涉及以下步骤:在数据库中搜索直接重复序列或配对序列以鉴定包含CRISPR酶的CRISPR区域。在有义和反义两者的方向上搜索CRISPR酶侧翼的CRISPR区域中的同源序列。寻找转录终止子和二级结构。鉴定不是直接重复序列或配对序列但与直接重复序列或配对序列具有超过50%同一性的任何序列作为潜在的序列。取潜在的序列并分析与其相关的转录终止子序列。

[0172] 应当理解的是,本文所述的任何功能性可以被工程化到来自其他直系同源物的CRISPR酶中,包括包含来自多个直系同源物的片段的嵌合酶。这样的直系同源物的实例在本文其他地方描述。因此,嵌合酶可以包含生物的CRISPR酶直系同源物的片段,所述生物包括但不限于纤毛菌属(*Leptotrichia*)、李斯特菌属(*Listeria*)、棒状杆菌属(*Corynebacter*)、萨特氏菌属(*Sutterella*)、军团菌属(*Legionella*)、密螺旋体属(*Treponema*)、产线菌属(*Filifactor*)、真细菌属(*Eubacterium*)、链球菌属(*Streptococcus*)、乳酸菌属(*Lactobacillus*)、支原体属(*Mycoplasma*)、拟杆菌属(*Bacteroides*)、*Flaviivola*、黄杆菌属(*Flavobacterium*)、*Sphaerochaeta*、固氮螺菌属(*Azospirillum*)、葡糖醋杆菌属(*Gluconacetobacter*)、奈瑟氏菌属(*Neisseria*)、罗氏菌属(*Roseburia*)、*Parvibaculum*、葡萄球菌属(*Staphylococcus*)、*Nitratifactor*、支原体属(*Mycoplasma*)和弯曲杆菌属(*Campylobacter*)。嵌合酶可以包含第一片段和第二片段,并且片段可以是本文提及的属或本文提及的种的生物的CRISPR酶直系同源物;有利地,片段来自不同种的CRISPR酶直系同源物。

[0173] 在实施方式中,本文所称的C2c2蛋白也涵盖C2c2的功能变体或其同源物或直系同源物。如本文所用的蛋白的“功能变体”是指至少部分保留该蛋白的活性的这样的蛋白的变体。功能变体可以包括突变体(其可以是插入、缺失或替换突变体),包括多晶型物等。功能变体中还包括这样的蛋白与另一种通常不相关的核酸、蛋白质、多肽或肽的融合产物。功能变体可以是天然存在的或可以是人造的。有利的实施方式可以涉及工程化或非天然存在的VI型RNA靶向效应蛋白。

[0174] 在一个实施方式中,编码C2c2的一种或多种核酸分子或其直系同源物或同源物可以被密码子优化用于在真核细胞中表达。真核生物可如本文所述。一种或多种核酸分子可以是工程化的或非天然存在的。

[0175] 在一个实施方式中,C2c2或其直系同源物或同源物可以包含一个或多个突变(并且因此编码其的核酸分子可以具有一个或多个突变。突变可以是人工引入的突变并且可以

包括但不限于催化结构域中的一个或多个突变。参考Cas9酶的催化结构域的实例可以包括但不限于RuvC I、RuvC II、RuvC III和HNH结构域。

[0176] 在一个实施方式中,C2c2或其直系同源物或同源物可包含一个或多个突变。突变可以是人工引入的突变,并且可以包括但不限于催化结构域中的一个或多个突变。参考Cas9酶的催化结构域的实例可以包括但不限于HEPN结构域。

[0177] 在一个实施方式中,C2c2或其直系同源物或同源物可以用作与功能结构域融合或可操作地连接的通用核酸结合蛋白。示例性功能结构域可以包括但不限于翻译起始子、翻译激活子、翻译抑制子、核酸酶,特别是核糖核酸酶,剪接体、珠子、光诱导/可控结构域或化学诱导/可控结构域。

[0178] 在某些示例性实施方式中,C2c2效应蛋白可以来自选自以下的生物:纤毛菌属(*Leptotrichia*)、李斯特菌属(*Listeria*)、棒状杆菌属(*Corynebacter*)、萨特氏菌属(*Sutterella*)、军团菌属(*Legionella*)、密螺旋体属(*Treponema*)、产线菌属(*Filifactor*)、真细菌属(*Eubacterium*)、链球菌属(*Streptococcus*)、乳酸菌属(*Lactobacillus*)、支原体属(*Mycoplasma*)、拟杆菌属(*Bacteroides*)、*Flaviivola*、黄杆菌属(*Flavobacterium*)、*Sphaerochaeta*、固氮螺菌属(*Azospirillum*)、葡糖醋杆菌属(*Gluconacetobacter*)、奈瑟氏菌属(*Neisseria*)、罗氏菌属(*Roseburia*)、*Parvibaculum*、葡萄球菌属(*Staphylococcus*)、*Nitratifactor*、支原体属(*Mycoplasma*)、弯曲杆菌属(*Campylobacter*)。

[0179] 在某些实施方式中,效应蛋白可以是李斯特菌属C2c2p,优选*Listeria seeligeri* C2c2p,更优选*Listeria seeligeri* serovar 1/2b str.SLCC3954C2c2p,并且crRNA序列长度可以是44至47个核苷酸,具有5' 29-nt直接重复(DR)和15-nt至18-nt间隔子。

[0180] 在某些实施方式中,效应蛋白可以是纤毛菌属C2c2p,优选*Leptotrichia shahii* C2c2p,更优选*Leptotrichia shahii* DSM 19757C2c2p,并且crRNA序列长度可为42至58个核苷酸,具有至少24nt的5'直接重复,例如5' 24-28-nt直接重复(DR),和至少14nt的间隔子,例如14-nt至28-nt间隔子,或至少18nt的间隔子,例如19、20、21、22或更多nt,例如18-28、29-28、20-28、21-28或22-28nt。

[0181] 在某些示例性实施方式中,效应蛋白可以是纤毛菌属、*Leptotrichia wadei* F0279、或李斯特菌属,优选地*Listeria newyorkensis* FSL M6-0635。

[0182] 在某些示例性实施方式中,本发明的C2c2效应蛋白包括但不限于以下21种直系同源物种(包括多个CRISPR基因座:*Leptotrichia shahii*; *Leptotrichia wadei* (Lw2); *Listeria seeligeri*; *Lachnospiraceae* bacterium MA2020; *Lachnospiraceae* bacterium NK4A179; [梭菌属] *aminophilum* DSM 10710; *Carnobacterium gallinarum* DSM 4847; *Carnobacterium gallinarum* DSM 4847 (第二CRISPR基因座); *Paludibacter propionicigenes* WB4; *Listeria weihenstephanensis* FSL R9-0317; *Listeriaceae* bacterium FSL M6-0635; *Leptotrichia wadei* F0279; *Rhodobacter capsulatus* SB 1003; *Rhodobacter capsulatus* R121; *Rhodobacter capsulatus* DE442; *Leptotrichia buccalis* C-1013-b; *Herbinix hemicellulosilytica*; [真细菌属] *rectale*; *Eubacteriaceae* bacterium CHKCI004; *Blautia* sp.Marseille-P2398; *Leptotrichia* sp.oral taxon 879str.F0557。另外十二(12)个非限制性实例是:*Lachnospiraceae*

bacterium NK4A144;Chloroflexus aggregans;Demequina aurantiaca;Thalassospira sp.TSL5-1;Pseudobutyrvibrio sp.OR37;Butyrvibrio sp.YAB3001;Blautia sp.Marseille-P2398;Leptotrichia sp.Marseille-P3007;Bacteroides ihuae;Porphyromonadaceae bacterium KH3CP3RA;Listeria riparia;和Insolitispirillum peregrinum。

[0183] 在某些实施方式中,根据本发明的C2c2蛋白是或衍生自下表所述的直系同源物之一,或者是如下表所述的两种或更多种直系同源物的嵌合蛋白,或者是下表所述的直系同源物之一的突变体或变体(或嵌合突变体或变体),包括死亡C2c2、分裂C2c2、去稳定化C2c2等,如本文其他地方所定义的,有或没有与异源/功能结构域融合。

[0184] 在某些示例性实施方式中,C2c2效应蛋白选自下表1。

[0185] 表1

[0186]

C2c2 直系同源物	编码	多字母
<i>Leptotrichia shahii</i>	C2-2	Lsh
<i>L. wadei</i> F0279 (Lw2)	C2-3	Lw2
<i>Listeria seeligeri</i>	C2-4	Lse
<i>Lachnospiraceae</i> bacterium MA2020	C2-5	LbM
<i>Lachnospiraceae</i> bacterium NK4A179	C2-6	LbNK179
<i>Clostridium aminophilum</i> DSM 10710	C2-7	Ca
<i>Carnobacterium gallinarum</i> DSM 4847	C2-8	Cg
<i>Carnobacterium gallinarum</i> DSM 4847	C2-9	Cg2
<i>Paludibacter propionicigenes</i> WB4	C2-10	Pp
<i>Listeria weihenstephanensis</i> FSL R9-0317	C2-11	Lwei
<i>Listeriaceae</i> bacterium FSL M6-0635	C2-12	LbFSL
<i>Leptotrichia wadei</i> F0279	C2-13	Lw
<i>Rhodobacter capsulatus</i> SB 1003	C2-14	Rc
<i>Rhodobacter capsulatus</i> R121	C2-15	Rc
<i>Rhodobacter capsulatus</i> DE442	C2-16	Rc
<i>Leptotrichia buccalis</i> C-1013-b	C2-17	LbuC2c2
<i>Herbinix hemicellulosilytics</i>	C2-18	HheC2c2
<i>Eubacterium rectale</i>	C2-19	EreC2c2
<i>Eubacteriaceae</i> bacterium CHKC1004	C2-20	EbaC2c2
<i>Blautia</i> sp. Marseille-P2398	C2-21	BsmC2c2

[0187]

<i>Leptotrichia</i> sp. oral taxon 879 str. F0557	C2-22	LspC2c2
<i>Lachnospiraceae</i> bacterium NK4a144		
<i>Chloroflexus aggregans</i>		
<i>Demequina aurantiaca</i>		
<i>Thalassospira</i> sp. TSL5-1		
<i>Pseudobutyrvibrio</i> sp. OR37		
<i>Butyrvibrio</i> sp. YAB3001		
<i>Blautia</i> sp. Marseille-P2398		
<i>Leptotrichia</i> sp. Marseille-P300		
<i>Bacteroides ihuae</i>		
<i>Porphyromonadaceae</i> bacterium KH3CP3RA		
<i>Listeria riparia</i>		
<i>Insolitispirillum peregrinum</i>		

[0188] 上述物种的野生型蛋白质序列列于下表2中。在某些实施方式中,提供了编码C2c2蛋白的核酸序列。

[0189] 表2

[0190]

C2c2-2		<i>L. shahii</i> (Lsh) (SEQ. I.D. No. 1)
C2c2-2		<i>L. shahii</i> (Lsh) (SEQ. I.D. No. 477) WP_018451595.1
c2c2-3		<i>L. wadei</i> (Lw2) (SEQ. I.D. No. 2)
c2c2-4		<i>Listeria seeligeri</i> (SEQ. I.D. No. 3)
c2c2-5	1	<i>Lachnospiraceae</i> bacterium MA2020 (SEQ. I.D. No. 4)
c2c2-6	2	<i>Lachnospiraceae</i> bacterium NK4A179 (SEQ. I.D. No. 5)
c2c2-7	3	<i>Clostridium aminophilum</i> DSM 10710 (SEQ. I.D. No. 6)
c2c2-8	5	<i>Carnobacterium gallinarum</i> DSM 4847 (SEQ. I.D. No. 7)
c2c2-9	6	<i>Carnobacterium gallinarum</i> DSM 4847 (SEQ. I.D. No. 8)
c2c2-10	7	<i>Paludibacter propionicigenes</i> WB4 (SEQ. I.D. No. 9)
c2c2-11	9	<i>Listeria weihenstephanensis</i> FSL R9-0317 (SEQ. I.D. No. 10)
c2c2-12	10	<i>Listeriaceae</i> bacterium FSL M6-0635 = <i>Listeria newyorkensis</i> FSL M6-0635

[0191]

		(SEQ. I.D. No. 11)
c2c2-13	12	<i>Leptotrichia wadei</i> F0279 (SEQ. I.D. No. 12)
c2c2-14	15	<i>Rhodobacter capsulatus</i> SB 1003 (SEQ. I.D. No. 13)
c2c2-15	16	<i>Rhodobacter capsulatus</i> R121 (SEQ. I.D. No. 14)
c2c2-16	17	<i>Rhodobacter capsulatus</i> DE442 (SEQ. I.D. No. 15)
LbuC2c2 (C2-17)		<i>Leptotrichia buccalis</i> C-1013-b (SEQ ID NO: 309)
HheC2c2 (C2-18)		<i>Herbinix hemicellulosilytica</i> (SEQ ID NO:310)
EreC2c2 (C2-19)		<i>Eubacterium rectale</i> (SEQ ID NO: 311)
EbaC2C2 (C2-20)		Eubacteriaceae bacterium CHKCI004 (SEQ ID NO: 312)
C2c2 (C2-21)		<i>Blautia</i> sp. Marseille-P2398 (SEQ. I.D. No 319)
C2c2 (C2-22)		<i>Leptotrichia</i> sp. Oral taxon 879 str. F0557 (SEQ. I.D. No. 579)
C2c2 NK4A144 (C2-23)		Lachnospiraceae bacterium NK4A144 (SEQ. I.D. No. 313)
C2c2 Chloro_agg (C2-24)		RNA 结合蛋白 S1 <i>Chloroflexus aggregans</i> (SEQ. I.D. No. 314)
C2c2 Dem_Aur (C2-25)		<i>Demequina aurantiaca</i> (SEQ. I.D. No. 315)
C2c2 Thal_Sp_TSL5 (C2-26)		<i>Thalassospira</i> sp. TSL5-1 (SEQ. I.D. No 316)
C2c2 Pseudo_sp (C2-27)		<i>Pseudobutyrvibrio</i> sp. OR37 (SEQ. I.D. No. 317)
C2c2_Buty_sp (C2-28)		<i>Butyrvibrio</i> sp. YAB3001 (SEQ. I.D. No. 318)
C2c2_Blautia_sp (C2-29)		<i>Blautia</i> sp. Marseille-P2398(SEQ. I.D. No. 478)
C2c2_Lepto_sp_Marseille (C2-30)		<i>Leptotrichia</i> sp. Marseille-P3007 (SEQ. ID No. 320)

[0192]

C2c2_Bacteroides_ihuae (C2-31)	Bacteroides ihuae (SEQ. I.D. No 321)
C2c2_Porph_bacterium (C2-32)	Porphyromonadaceae bacterium KH3CP3RA(SEQ. I.D. No. 322)
C2c2_Listeria_riparia (C2-33)	Listeria riparia (SEQ. I.D. No. 323)
C2c2_insolitiss_peregrinum (C2-34)	Insolitispirillum peregrinum (SEQ. I.D. No. 324)

[0193] 在本发明的一个实施方式中,提供效应蛋白,其包含与Leptotrichia shahii C2c2、Lachnospiraceae bacterium MA2020 C2c2、Lachnospiraceae bacterium NK4A179C2c2、Clostridium aminophilum (DSM 10710) C2c2、Carnobacterium gallinarum (DSM 4847) C2c2、Paludibacter propionicigenes (WB4) C2c2、Listeria weihenstephanensis (FSL R9-0317) C2c2、Listeriaceae bacterium (FSL M6-0635) C2c2、Listeria newyorkensis (FSL M6-0635) C2c2、Leptotrichia wadei (F0279) C2c2、Rhodobacter capsulatus (SB 1003) C2c2、Rhodobacter capsulatus (R121) C2c2、Rhodobacter capsulatus (DE442) C2c2、Leptotrichia wadei (Lw2) C2c2或Listeria seeligeri C2c2任一者的野生型序列具有至少80%的序列同源性的氨基酸序列。

[0194] 在本发明的实施方式中,效应蛋白包含与VI型效应蛋白共有序列包括但不限于本文所述的共有序列具有至少80%序列同源性的氨基酸序列。

[0195] 根据本发明,可以从多个C2c2直系同源物中产生共有序列,其可帮助定位保守氨基酸残基、和基序,包括但不限于在介导C2C2功能的C2C2直系同源物中的催化残基和HEPN基序。使用基因比对从上文提到的33种直系同源物产生的一种这样的共有序列是:

[0196] MKISKVXXXVKKXXXGKLLKXVNERNRXAKRLSNXLBKYIXXIDKIXKKEXXXKFXAXEEITLKLNQ
 XXXBXLXKAXXDLRKDNYSXJKKILHNEDINXEEEXELLINDXLEKLXKIESXKYSYQKXXXNYXMSVQEHSKKSI
 XRIXESAKRNKEALDKFLKEYAXLDPRMEXLAKLRKLELYFYFKNDXIXXEEEXNVXXHKXLKENHPDFVEXXXN
 KENAELNXYAIEXXXJLKYYPKXAKNSNDKIFEKQELKXWIHQJENAVERILLXXGKVXYKLQXGYLAELWKI
 RINEIFIKYIXVKGAVAXFALRNXXKBENDILGGKIXKKNLNGITSFYXKIKAEELQREXAVEVAFANXLYAXD
 LXXIRXSILQFFGGASNWDXFLFFHFATSXISDKKWAELIXXKJGLVIREKLYSNNVAMFYSKDDLEKLLNXLX
 XFXLRASQVPSFKKVYVRXBFPQNLKKFNDEKDDEAYSAXYLLKEIYYNXFLPYFSANNXFFFVKNLVLKANK
 DKFXXAFXDIREMNXGSPIEYLXXTQXNXNEGRKKEEKEXDFIKFLLQIFXKGFDDYLKNNXXFILKFIPEPTX
 IEIXXELQAWYIVGKFLNARKXNLLGXFXYLKLDDIELRALRNENIKYQSSNXEKEVLEXCLELIGLLSLDLND
 YFBDEXDFAXYJGKXLDPEKXMKDLAELXPYDQNDGENPIVNRNIXLAKKYGTNLLEKJXDKVSEKEIKEYYEL
 KKEIEEYXXXGEELHEEWXQXKNRVEXRDILEYXEELXGQIINYNXLXNKVLLYFQLGLHYLLLDILGRLVGYTGI
 WERDAXLYQIAAMYXNGLPEYIXXKKNKYKDGQIVGXKINXFKXDKKXLYNAGLELFENXNEHKNIXIRNYIAHF
 NYLSKAESSLLXYSENLRXLFSYDRKLKNAVXKSLINILLRHGMVLKFKFGTDKKSXVIRSXXKIXHLKSIACKLY
 YPEVXVSKEYCKLVKXLLKYK (SEQ ID NO:325)

[0197] 在另一个非限制性实例中,协助共有序列的生成和保守残基的鉴定的序列比对工具是MUSCLE比对工具(www.ebi.ac.uk/Tools/msa/muscle/)。例如,使用MUSCLE,可以在 *Leptotrichia wadei* C2c2中鉴定C2c2直系同源物中保守的下列氨基酸位置:K2;K5;V6;E301;L331;I335;N341;G351;K352;E375;L392;L396;D403;F446;I466;I470;R474 (HEPN);H475;H479 (HEPN),E508;P556;L561;I595;Y596;F600;Y669;I673;F681;L685;Y761;L676;L779;Y782;L836;D847;Y863;L869;I872;K879;I933;I954;I958;R961;Y965;E970;R971;D972;R1046 (HEPN),H1051 (HEPN),Y1075;D1076;K1078;K1080;I1083;I1090。

[0198] 图50显示表示高度保守的残基的HEPN结构域的示例性序列比对。

[0199] 在某些示例性实施方式中,RNA靶向效应蛋白是VI型B效应蛋白,例如Cas13b和29组或30组的蛋白。在某些示例性实施方式中,RNA靶向效应蛋白包含一个或多个HEPN结构域。在某些示例性实施方式中,RNA靶向效应蛋白包含C端HEPN结构域、N端HEPN结构域或两者。关于可以在本发明的上下文中使用的示例性VI-B效应蛋白,参考2016年10月21日提交的题为“Novel CRISPR Enzymes and Systems”的美国申请号15/331,792,2016年10月16日提交的题为“Novel CRISPR Enzymes and Systems”的国际专利申请号PCT/US2016/058302,和Smargon等“Cas13b is a Type VI-B CRISPR-associated RNA-Guided RNase differentially regulated by accessory proteins Csx27and Csx28”*Molecular Cell*, 65,1-13 (2017);[dx.doi.org/10.1016/j.molcel.2016.12.023](https://doi.org/10.1016/j.molcel.2016.12.023),和美国临时申请号待指定的于2017年3月15日提交的题为“Novel Cas13b Orthologues CRISPR Enzymes and System”的申请。在特定的实施方式中,Cas13b酶是来自 *Bergeyella zoohelcum*。在某些其他示例性实施方式中,效应蛋白是或包含与表3中列出的任何序列具有至少80%序列同源性的氨基酸序列。

[0200] 表3

[0201]

B-01	<i>Bergeyella zoohelcum</i>
B-02	<i>Prevotella intermedia</i>
B-03	<i>Prevotella buccae</i>
B-04	<i>Alistipes</i> sp.ZOR0009
B-05	<i>Prevotella</i> sp.MA2016
B-06	<i>Riemerella anatipestifer</i>
B-07	<i>Prevotella aurantiaca</i>
B-08	<i>Prevotella saccharolytica</i>
B-09	<i>Prevotella intermedia</i>
B-10	<i>Capnocytophaga canimorsus</i>
B-11	<i>Porphyromonas gulae</i>
B-12	<i>Prevotella</i> sp.P5-125
B-13	<i>Flavobacterium branchiophilum</i>
B-14	<i>Porphyromonas gingivalis</i>
B-15	<i>Prevotella intermedia</i>

[0202] 在某些示例性实施方式中,Cas13b直系同源物的野生型序列见于下表4a或4b中。

[0203] 表4a

[0204]

Bergeyella zoohelcum (SEQ. I.D. No. 326)	1
Prevotella intermedia (SEQ. I.D. No. 327)	2
Prevotella buccae (SEQ. I.D. No. 328)	3
Porphyromonas gingivalis (SEQ. I.D. No. 329)	4
Bacteroides pyogenes (SEQ. I.D. No. 330)	5
Alistipes sp. ZOR0009 (SEQ. I.D. No. 331)	6
Prevotella sp. MA2016 (SEQ. I.D. No. 332)	7a
Prevotella sp. MA2016 (SEQ. I.D. No. 333)	7b
Riemerella anatipestifer (SEQ. I.D. No. 334)	8
Prevotella aurantiaca (SEQ. I.D. No. 335)	9
Prevotella saccharolytica (SEQ. I.D. No. 336)	10
HMPREF9712_03108 [Myroides odoratimimus CCUG 10230] (SEQ. I.D. No. 337)	11
Prevotella intermedia (SEQ. I.D. No. 338)	12
Capnocytophaga canimorsus (SEQ. I.D. No. 339)	13
Porphyromonas gulae (SEQ. I.D. No. 340)	14
Prevotella sp. P5-125 (SEQ. I.D. No. 341)	15
Flavobacterium branchiophilum (SEQ. I.D. No. 342)	16
Myroides odoratimimus (SEQ. I.D. No. 343)	17
Flavobacterium columnare (SEQ. I.D. No. 344)	18
Porphyromonas gingivalis (SEQ. I.D. No. 345)	19
Porphyromonas sp. COT-052 OH4946 (SEQ. I.D. No. 346)	20
Prevotella intermedia (SEQ. I.D. No. 347)	21
PIN17_0200 [Prevotella intermedia 17] (SEQ. I.D. No. 348)	AFJ07523
Prevotella intermedia (SEQ. I.D. No. 349)	BAU18623
HMPREF6485_0083 [Prevotella buccae ATCC 33574] (SEQ. I.D. No. 350)	EFU31981
HMPREF9144_1146 [Prevotella pallens ATCC 700821] (SEQ. I.D. No. 351)	EGQ18444
HMPREF9714_02132 [Myroides odoratimimus CCUG 12901] (SEQ. I.D. No. 352)	EHO08761
HMPREF9711_00870 [Myroides odoratimimus CCUG 3837] (SEQ. I.D. No. 353)	EKB06014
HMPREF9699_02005 [Bergeyella zoohelcum ATCC 43767] (SEQ. I.D. No. 354)	EKB54193
HMPREF9151_01387 [Prevotella saccharolytica F0055] (SEQ. I.D. No. 355)	EKY00089
A343_1752 [Porphyromonas gingivalis JCVI SC001] (SEQ. I.D. No. 356)	EOA10535
HMPREF1981_03090 [Bacteroides pyogenes F0041] (SEQ. I.D. No. 357)	ERI81700
HMPREF1553_02065 [Porphyromonas gingivalis F0568] (SEQ. I.D. No. 358)	ERJ65637
HMPREF1988_01768 [Porphyromonas gingivalis F0185] (SEQ. I.D. No. 359)	ERJ81987
HMPREF1990_01800 [Porphyromonas gingivalis W4087] (SEQ. I.D. No. 360)	ERJ87335
M573_117042 [Prevotella intermedia ZT] (SEQ. I.D. No. 361)	KJJ86756
A2033_10205 [Bacteroidetes bacterium GWA2_31_9] (SEQ. I.D. No. 362)	OFX18020.1
SAMN05421542_0666 [Chryseobacterium jejuense] (SEQ. I.D. No. 363)	SDI27289.1
SAMN05444360_11366 [Chryseobacterium carnipullorum] (SEQ. I.D. No. 364)	SHM52812.1
SAMN05421786_1011119 [Chryseobacterium ureilyticum] (SEQ. I.D. No. 365)	SIS70481.1
Prevotella buccae (SEQ. I.D. No. 366)	WP_004343581
Porphyromonas gingivalis (SEQ. I.D. No. 367)	WP_005873511
Porphyromonas gingivalis (SEQ. I.D. No. 368)	WP_005874195
Prevotella pallens (SEQ. I.D. No. 369)	WP_006044833
Myroides odoratimimus (SEQ. I.D. No. 370)	WP_006261414
Myroides odoratimimus (SEQ. I.D. No. 371)	WP_006265509
Prevotella sp. MSX73 (SEQ. I.D. No. 372)	WP_007412163
Porphyromonas gingivalis (SEQ. I.D. No. 373)	WP_012458414
Paludibacter propionicigenes (SEQ. I.D. No. 374)	WP_013446107
Porphyromonas gingivalis (SEQ. I.D. No. 375)	WP_013816155
Flavobacterium columnare (SEQ. I.D. No. 376)	WP_014165541
Psychroflexus torquis (SEQ. I.D. No. 377)	WP_015024765
Riemerella anatipestifer (SEQ. I.D. No. 378)	WP_015345620
Prevotella pleuritidis (SEQ. I.D. No. 379)	WP_021584635

[0205]

Porphyromonas gingivalis (SEQ. I.D. No. 380)	WP_021663197
Porphyromonas gingivalis (SEQ. I.D. No. 381)	WP_021665475
Porphyromonas gingivalis (SEQ. I.D. No. 382)	WP_021677657
Porphyromonas gingivalis (SEQ. I.D. No. 383)	WP_021680012
Porphyromonas gingivalis (SEQ. I.D. No. 384)	WP_023846767
Prevotella falsenii (SEQ. I.D. No. 385)	WP_036884929
Prevotella pleuritidis (SEQ. I.D. No. 386)	WP_036931485
[Porphyromonas gingivalis (SEQ. I.D. No. 387)	WP_039417390
Porphyromonas gulae (SEQ. I.D. No. 388)	WP_039418912
Porphyromonas gulae (SEQ. I.D. No. 389)	WP_039419792
Porphyromonas gulae (SEQ. I.D. No. 390)	WP_039426176
Porphyromonas gulae (SEQ. I.D. No. 391)	WP_039431778
Porphyromonas gulae (SEQ. I.D. No. 392)	WP_039437199
Porphyromonas gulae (SEQ. I.D. No. 393)	WP_039442171
Porphyromonas gulae (SEQ. I.D. No. 394)	WP_039445055
Capnocytophaga cynodegmi (SEQ. I.D. No. 395)	WP_041989581
Prevotella sp. P5-119 (SEQ. I.D. No. 396)	WP_042518169
Prevotella sp. P4-76 (SEQ. I.D. No. 397)	WP_044072147
Prevotella sp. P5-60 (SEQ. I.D. No. 398)	WP_044074780
Phaeodactylibacter xiamenensis (SEQ. I.D. No. 399)	WP_044218239
Flavobacterium sp. 316 (SEQ. I.D. No. 400)	WP_045968377
Porphyromonas gulae (SEQ. I.D. No. 401)	WP_046201018
WP_047431796 (SEQ. I.D. No. 402)	Chryseobacterium sp. YR477
Riemerella anatipestifer (SEQ. I.D. No. 403)	WP_049354263
Porphyromonas gingivalis (SEQ. I.D. No. 404)	WP_052912312
Porphyromonas gingivalis (SEQ. I.D. No. 405)	WP_058019250
Flavobacterium columnare (SEQ. I.D. No. 406)	WP_060381855
Porphyromonas gingivalis (SEQ. I.D. No. 407)	WP_061156470
Porphyromonas gingivalis (SEQ. I.D. No. 408)	WP_061156637
Riemerella anatipestifer (SEQ. I.D. No. 409)	WP_061710138
Flavobacterium columnare (SEQ. I.D. No. 410)	WP_063744070
Riemerella anatipestifer (SEQ. I.D. No. 411)	WP_064970887
Sinomicrobium oceani (SEQ. I.D. No. 412)	WP_072319476.1
Reichenbachiella agariperforans (SEQ. I.D. No. 413)	WP_073124441.1

[0206] 表4b

[0207]

名称或登录号
WP_015345620 (SEQ. I.D. No. 479)
WP_049354263 (SEQ. I.D. No. 480)
WP_061710138 (SEQ. I.D. No. 481)
6 (SEQ. I.D. No. 482) Alistipes sp. ZOR0009
SIS70481.1
15 Prevotella sp. (SEQ. I.D. No. 484)
WP_042518169 (SEQ. I.D. No. 485)
WP_044072147 (SEQ. I.D. No. 486)
WP_044074780 (SEQ. I.D. No. 487)

[0208]

8 (modified) (SEQ. I.D. No. 488)
WP_064970887 (SEQ. I.D. No. 489)
5 (SEQ. I.D. No. 490)
ERI81700 (SEQ. I.D. No. 491)
WP_036931485 (SEQ. I.D. No. 492)
19 (SEQ. I.D. No. 493)
WP_012458414 (SEQ. I.D. No. 494)
WP_013816155 (SEQ. I.D. No. 495)
WP_039417390 (SEQ. I.D. No. 496)
WP_039419792 (SEQ. I.D. No. 497)
WP_039426176 (SEQ. I.D. No. 498)
WP_039437199 (SEQ. I.D. No. 499)
WP_061156470 (SEQ. I.D. No. 500)
12 (SEQ. I.D. No. 501)
9 (SEQ. I.D. No. 502)
EGQ18444 (SEQ. I.D. No. 503)
KJJ86756 (SEQ. I.D. No. 504)
WP_006044833 (SEQ. I.D. No. 505)
2 (SEQ. I.D. No. 506)
3 (SEQ. I.D. No. 507)
EFU31981 (SEQ. I.D. No. 508)
WP_004343581 (SEQ. I.D. No. 509)
WP_007412163 (SEQ. I.D. No. 510)
WP_044218239 (SEQ. I.D. No. 511)
21 (SEQ. I.D. No. 512)
BAU18623 (SEQ. I.D. No. 513)
WP_036884929 (SEQ. I.D. No. 514)
WP_073124441.1 (SEQ. I.D. No. 515)
AFJ07523 (SEQ. I.D. No. 516)
4 (SEQ. I.D. No. 517)
ERJ65637 (SEQ. I.D. No. 518)
ERJ81987 (SEQ. I.D. No. 519)
ERJ87335 (SEQ. I.D. No. 520)
WP_005873511 (SEQ. I.D. No. 521)
WP_021663197 (SEQ. I.D. No. 522)
WP_021665475 (SEQ. I.D. No. 523)
WP_021677657 (SEQ. I.D. No. 524)
WP_021680012 (SEQ. I.D. No. 525)
WP_023846767 (SEQ. I.D. No. 526)
WP_039445055 (SEQ. I.D. No. 527)
WP_061156637 (SEQ. I.D. No. 528)
WP_021584635 (SEQ. I.D. No. 529)
WP_015024765 (SEQ. I.D. No. 530)
WP_047431796 (SEQ. I.D. No. 531)
WP_072319476.1 (SEQ. I.D. No. 532)
16 (SEQ. I.D. No. 533)
EKY00089 (SEQ. I.D. No. 534)
10 (SEQ. I.D. No. 535)
WP_013446107 (SEQ. I.D. No. 536)

[0209]

WP_045968377 (SEQ. I.D. No. 537)
SHM52812.1 (SEQ. I.D. No. 538)
EHO08761 (SEQ. I.D. No. 539)
EKB06014 (SEQ. I.D. No. 540)
WP_006261414 (SEQ. I.D. No. 541)
WP_006265509 (SEQ. I.D. No. 542)
11 (SEQ. I.D. No. 543)
17 (SEQ. I.D. No. 544)
OFX18020.1 (SEQ. I.D. No. 545)
SDI27289.1 (SEQ. I.D. No. 546)
WP_039442171 (SEQ. I.D. No. 547)
14 (SEQ. I.D. No. 548)
20 (SEQ. I.D. No. 549)
EOA10535 (SEQ. I.D. No. 550)
WP_005874195 (SEQ. I.D. No. 551)
WP_039418912 (SEQ. I.D. No. 552)
WP_039431778 (SEQ. I.D. No. 553)
WP_046201018 (SEQ. I.D. No. 554)
WP_052912312 (SEQ. I.D. No. 555)
WP_058019250 (SEQ. I.D. No. 556)
WP_014165541 (SEQ. I.D. No. 557)
13 (SEQ. I.D. No. 558)
WP_060381855 (SEQ. I.D. No. 559)
WP_063744070 (SEQ. I.D. No. 560)
18 (SEQ. I.D. No. 561)
WP_041989581 (SEQ. I.D. No. 562)
1 (SEQ. I.D. No. 563)
EKB54193 (SEQ. I.D. No. 564)
7 (modified) (SEQ. I.D. No. 565)
7 (modified) - residues only (SEQ. I.D. No. 566)

[0210] 在某些示例性实施方式中，RNA靶向效应蛋白是Cas13c效应蛋白，如2017年6月26日提交的美国临时专利申请号62/525,165和2017年8月16日提交的PCT申请号US2017/047193中公开的。在下表5中提供了Cas13c的示例性野生型直向同源序列。

[0211] 表5

[0212]

名称
EHO19081 (SEQ. I.D. No. 567)
WP_094899336 (SEQ. I.D. No. 568)
WP_040490876 (SEQ. I.D. No. 569)
WP_047396607 (SEQ. I.D. No. 570)
WP_035935671 (SEQ. I.D. No. 571)
WP_035906563 (SEQ. I.D. No. 572)
WP_042678931 (SEQ. I.D. No. 573)
WP_062627846 (SEQ. I.D. No. 574)

[0213]

WP 005959231 (SEQ. I.D. No. 575)
WP 027128616 (SEQ. I.D. N. 576)
WP 062624740 (SEQ. I.D. No. 577)
WP 096402050 (SEQ. I.D. No. 578)

[0214] 在某些示例性实施方式中,Cas13蛋白可选自以下任何一种。

[0215] 表6

[0216]

ID	物种	Seq. ID. No:
Cas13a1	Leptotrichia shahii	580
Cas13a2	Leptotrichia wadei (Lw2)	581
Cas13a3	Listeria seeligeri	582
Cas13a4	Lachnospiraceae bacterium MA2020	583
Cas13a5	Lachnospiraceae bacterium NK4A179	584
Cas13a6	[Clostridium] aminophilum DSM 10710	585
Cas13a7	Carnobacterium gallinarum DSM 4847	586
Cas13a8	Carnobacterium gallinarum DSM 4847	587
Cas13a9	Paludibacter propionicigenes WB4	588
Cas13a10	Listeria weihenstephanensis FSL R9-0317	589
Cas13a11	Listeriaceae bacterium FSL M6-0635	590
Cas13a12	Leptotrichia wadei F0279	591
Cas13a13	Rhodobacter capsulatus SB 1003	592
Cas13a14	Rhodobacter capsulatus R121	593
Cas13a15	Rhodobacter capsulatus DE442	594
Cas13a16	Leptotrichia buccalis C-1013-b	595
Cas13a17	Herbinix hemicellulosilytica	596

[0217]

Cas13a18	[Eubacterium] rectale	597
Cas13a19	Eubacteriaceae bacterium CHKCI004	598
Cas13a20	Blautia sp. Marseille-P2398	599
Cas13a21	Leptotrichia sp. oral taxon 879 str. F0557	600
Cas13b1	Bergeyella zoohelcum	601
Cas13b2	Prevotella intermedia	602
Cas13b3	Prevotella buccae	603
Cas13b4	Alistipes sp. ZOR0009	604
Cas13b5	Prevotella sp. MA2016	605
Cas13b6	Riemerella anatipestifer	606
Cas13b7	Prevotella aurantiaca	607
Cas13b8	Prevotella saccharolytica	608
Cas13b9	Prevotella intermedia	609
Cas13b10	Capnocytophaga canimorsus	610
Cas13b11	Porphyromonas gulae	611
Cas13b12	Prevotella sp. P5-125	612
Cas13b13	Flavobacterium branchiophilum	613
Cas13b14	Porphyromonas gingivalis	614
Cas13b15	Prevotella intermedia	615
Cas13c1	Fusobacterium necrophorum subsp. funduliforme ATCC 51357 contig00003	616
Cas13c2	Fusobacterium necrophorum DJ-2 contig0065, whole genome shotgun sequence	617
Cas13c3	Fusobacterium necrophorum BFTR-1 contig0068	618
Cas13c4	Fusobacterium necrophorum subsp. funduliforme 1_1_36S cont1.14	619
Cas13c5	Fusobacterium perfoetens ATCC 29250 T364DRAFT_scaffold00009.9_C	620
Cas13c6	Fusobacterium ulcerans ATCC 49185 cont2.38	621
Cas13c7	Anaerosalibacter sp. ND1 genome assembly Anaerosalibacter massiliensis ND1	622

[0218] 指导序列

[0219] 如本文所用,术语“指导序列”、“crRNA”、“指导RNA”或“单指导RNA”或“gRNA”是指包含与靶核酸序列具有足够互补性以与靶核酸序列杂交和引导包含指导序列和CRISPR效应蛋白的RNA靶向复合物结合到靶核酸序列的任何多核苷酸序列的多核苷酸。在一些示例性实施方式中,当使用合适的比对算法最佳地对准时,互补程度是约或大于约50%、60%、75%、80%、85%、90%、95%、97.5%、99%、或者更多。可以使用用于比对序列的任何合适的算法来确定最佳比对,其非限制性实例包括Smith-Waterman算法、Needleman-Wunsch算法、基于Burrows-Wheeler变换的算法(例如,Burrows Wheeler Aligner)、ClustalW、

Clustal X、BLAT、Novoalign (Novocraft Technologies; 可从www.novocraft.com获得)、ELAND (Illumina, San Diego, CA), SOAP (可从soap.genomics.org.cn获得) 和Maq (可从maq.sourceforge.net获得)。指导序列(在核酸靶向指导RNA内) 引导核酸靶向复合物与靶核酸序列的序列特异性结合的能力可以通过任何合适的测定来评估。例如, 足以形成核酸靶向复合物的核酸靶向CRISPR系统的组分, 包括待测试的指导序列, 可以提供给具有相应靶核酸序列的宿主细胞, 例如通过用编码核酸靶向复合物的组分的载体转染, 然后评估在靶核酸序列内的优先靶向(例如, 切割), 例如通过本文所述的Surveyor测定法。类似地, 可以通过提供靶核酸序列、核酸靶向复合物的组分, 包括待测试的指导序列和不同于测试指导序列的对照指导序列, 并且比较测试与对照指导序列反应之间在靶序列处的结合或切割速率而在测试试管中评估靶多核酸序列的切割。其他测定是可能的, 并且由本领域技术人员想到。可以选择指导序列, 并因此选择核酸靶向指导以靶向任何靶核酸序列。靶序列可以是DNA。靶序列可以是任何RNA序列。在一些实施方式中, 靶序列可以是RNA分子内的序列, 其选自信使RNA (mRNA)、前mRNA、核糖体RNA (rRNA)、转移RNA (tRNA)、微RNA (miRNA)、小干扰RNA (siRNA)、小核RNA (snRNA)、小核仁RNA (snoRNA)、双链RNA (dsRNA)、非编码RNA (ncRNA)、长非编码RNA (lncRNA) 和小细胞质RNA (scRNA)。在一些优选的实施方式中, 靶序列可以是RNA分子内的序列, 其选自mRNA、前mRNA和rRNA。在一些优选的实施方式中, 靶序列可以是选自ncRNA和lncRNA的RNA分子内的序列。在一些更优选的实施方式中, 靶序列可以是mRNA分子或前mRNA分子内的序列。

[0220] 在一些实施方式中, 选择核酸靶向指导以降低核酸靶向指导内的二级结构程度。在一些实施方式中, 核酸靶向指导的约或小于约75%、50%、40%、30%、25%、20%、15%、10%、5%、1%或更少的核苷酸当在最佳折叠时参与自我互补的碱基配对。最佳折叠可以通过任何合适的多核苷酸折叠算法确定。一些程序基于计算最小吉布斯自由能。如Zuker和Stiegler (Nucleic Acids Res. 9 (1981), 133-148) 所述, 一种这样的算法的实例是mFold。另一示例性折叠算法是使用质心结构预测算法在维也纳大学理论化学研究所开发的在线网络服务器RNAfold (参见例如A. R. Gruber等, 2008, Cell 106 (1) :23-24; 和PA Carr和GM Church, 2009, Nature Biotechnology 27 (12) :1151-62)。

[0221] 在某些实施方式中, 指导RNA或crRNA可以包含直接重复 (DR) 序列和指导序列或间隔序列, 基本上由以上组成或由以上组成。在某些实施方式中, 指导RNA或crRNA可以包含与指导序列或间隔子序列融合或连接的直接重复序列, 基本上由以上组成或由以上组成。在某些实施方式中, 直接重复序列可位于指导序列或间隔序列的上游(即5')。在其他实施方式中, 直接重复序列可以位于指导序列或间隔子序列的下游(即3')。

[0222] 在某些实施方式中, crRNA包含茎环, 优选单茎环。在某些实施方式中, 直接重复序列形成茎环, 优选单茎环。

[0223] 在某些实施方式中, 指导RNA的间隔子长度为15至35nt。在某些实施方式中, 指导RNA的间隔子长度为至少15个核苷酸。在某些实施方式中, 间隔子长度为15至17nt, 例如15、16或17nt, 17至20nt, 例如17、18、19或20nt, 20至24nt, 例如, 20、21、22、23或24nt, 23至25nt, 例如23、24或25nt, 24至27nt, 例如24、25、26或27nt, 27至30nt, 例如, 27、28、29或30nt, 30至35nt, 例如30、31、32、33、34或35nt, 或35nt或更长。

[0224] 通常, CRISPR-Cas、CRISPR-Cas9或CRISPR系统可以是如在上述文件例如WO 2014/

093622 (PCT/US2013/074667) 中所使用的,并共同地是指涉及表达CRISPR相关(“Cas”)基因或引导CRISPR相关(“Cas”)基因的活性的转录物和其他元件,包括编码以下的序列:Cas基因,特别是在CRISPR-Cas9的情况下的Cas9基因、tracr(反式激活CRISPR)序列(例如tracrRNA或具有活性的部分tracrRNA)、tracr配对序列(在内源CRISPR系统的上下文中包括“直接重复”和tracrRNA处理的部分直接重复)、指导序列(在内源CRISPR系统的上下文中也称为“间隔子”)、或如本文所用的术语“RNA”(例如,指导Cas9的RNA,例如CRISPR RNA和反式激活(tracr)RNA或单指导RNA(sgRNA)(嵌合RNA))或来自CRISPR基因座的其他序列和转录物。通常,CRISPR系统的特征在于促进CRISPR复合物在靶序列(在内源CRISPR系统的上下文中也称为原型间隔子)位点处形成的元件。在CRISPR复合物的形成的上下文中,“靶序列”是指指导序列被设计与其具有互补性的序列,其中靶序列与指导序列之间的杂交促进CRISPR复合物的形成。指导序列的部分,通过该部分与靶序列的互补性对于切割活性重要,在本文中称为种子序列。靶序列可包含任何多核苷酸,例如DNA或RNA多核苷酸。在一些实施方式中,靶序列位于细胞的细胞核或细胞质中,并且可包含在或来自细胞内存在的线粒体、细胞器、囊泡、脂质体或颗粒中的核酸。在一些实施方式中,特别是对于非核用途,NLS不是优选的。在一些实施方式中,CRISPR系统包含一个或多个核输出信号(NES)。在一些实施方式中,CRISPR系统包含一个或多个NLS和一个或多个NES。在一些实施方式中,直接重复可以通过搜索满足以下任何或所有标准的重复基序经由计算机模拟来鉴定:1.在II型CRISPR基因座侧翼的基因组序列的2Kb窗口中发现;2.跨度20至50bp;和3.间隔20至50bp。在一些实施方式中,可以使用这些标准中的2个,例如1和2、2和3、或1和3。在一些实施方式中,可以使用所有3个标准。

[0225] 在本发明的实施方式中,术语指导序列和指导RNA,即能够将Cas指导至靶标的RNA基因组基因座的RNA,如前述引用文献例如W02014/093622 (PCT/US2013/074667) 中可互换使用。通常,指导序列是与靶核酸序列具有足够互补性以与靶核酸序列杂交和引导CRISPR复合物与靶序列序列特异性结合的任何多核苷酸序列。在一些示例性实施方式中,当使用合适的比对算法最佳地对准时,指导序列与其相应靶序列之间的互补程度是约或大于约50%、60%、75%、80%、85%、90%、95%、97.5%、99%、或者更多。可以使用用于比对序列的任何合适的算法来确定最佳比对,其非限制性实例包括Smith-Waterman算法、Needleman-Wunsch算法、基于Burrows-Wheeler变换的算法(例如,Burrows Wheeler Aligner)、ClustalW、Clustal X、BLAT、Novoalign (Novocraft Technologies;可从www.novocraft.com获得)、ELAND (Illumina, San Diego, CA), SOAP (可从soap.genomics.org.cn获得)和Maq (可从maq.sourceforge.net获得)。在一些实施方式中,指导序列长度为约或大于约5、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、35、40、45、50、75或更多个核苷酸。在一些实施方式中,指导序列的长度小于约75、50、45、40、35、30、25、20、15、12或更少的核苷酸。优选地,指导序列长10-30个核苷酸。指导序列引导CRISPR复合物与靶序列序列特异性结合的能力可以通过任何合适的测定来评估。例如,足以形成CRISPR复合物的组分,包括待测试的指导序列,可以提供给具有相应靶核酸序列的宿主细胞,例如通过用编码CRISPR序列的组分的载体转染,然后评估在靶核酸序列内的优先切割,例如通过本文所述的Surveyor测定法。类似地,可以通过提供靶核酸序列、CRISPR复合物的组分,包括待测试的指导序列和不同于测试指导序列的对照指导序

列,并且比较测试与对照指导序列反应之间在靶序列处的结合或切割速率而在测试试管中评估靶多核酸序列的切割。其他测定是可能的,并且由本领域技术人员想到。

[0226] 在CRISPR-Cas系统的一些实施方式中,指导序列与其相应靶序列之间的互补程度可以是约或大于约50%、60%、75%、80%、85%、90%、95%、97.5%、99%或100%;指导或RNA或sgRNA的长度可以是约或大于约5、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、35、40、45、50、75或更多个核苷酸;或者指导或RNA或sgRNA的长度可小于约75、50、45、40、35、30、25、20、15、12或更少的核苷酸;并且有利地, tracr RNA的长度为30或50个核苷酸。然而,本发明的一个方面是减少脱靶相互作用,例如,减少指导与具有低互补性的靶序列的相互作用。的确,在实施例,显示本发明涉及突变,其导致CRISPR-Cas系统能够区分具有大于80%至约95%互补性,例如83%-84%或88-89%或94-95%互补性的靶序列和脱靶序列(例如,区分具有18个核苷酸的靶标与具有1、2或3个错配的18个核苷酸的脱靶标)。因此,在本发明的上下文中,指导序列与其相应靶序列之间的互补程度大于94.5%或95%或95.5%或96%或96.5%或97%或97.5%或98%或98.5%或99%或99.5%或99.9%,或100%。脱靶是在序列与指导之间小于100%或99.9%或99.5%或99%或99%或98.5%或98%或97.5%或97%或96.5%或96%或95.5%或95%或94.5%或94%或93%或92%或91%或90%或89%或88%或87%或86%或85%或84%或83%或82%或81%或80%的互补性,有利的是,脱靶是在序列与指导之间100%或99.9%或99.5%或99%或99%或98.5%或98%或97.5%或97%或96.5%或96%或95.5%或95%或94.5%的互补性。

[0227] 指导修饰

[0228] 在某些实施方式中,本发明的指导包含非天然存在的核酸和/或非天然存在的核苷酸和/或核苷酸类似物,和/或化学修饰。非天然存在的核酸可以包括例如天然和非天然存在的核苷酸的混合物。非天然存在的核苷酸和/或核苷酸类似物可以在核糖、磷酸和/或碱基部分修饰。在本发明的实施方式中,指导核酸包含核糖核苷酸和非核糖核苷酸。在一个这样的实施方式中,指导包含一种或多种核糖核苷酸和一种或多种脱氧核糖核苷酸。在本发明的实施方式中,指导包含一种或多种非天然存在的核苷酸或核苷酸类似物,例如具有硫代磷酸酯键、硼烷磷酸键的核苷酸、包含核糖环2'和4'碳指导之间的亚甲基桥的锁核酸(LNA)核苷酸、或桥接核酸(BNA)。修饰的核苷酸的其他实例包括2'-O-甲基类似物、2'-脱氧类似物、2-硫尿苷类似物、N6-甲基腺苷类似物或2'-氟代类似物。经修饰的碱基的其它实例包括但不限于,2-氨基嘌呤、5-溴-尿苷、假尿苷(Ψ)、N¹-甲基假尿苷($me^1\Psi$)、5-甲氧基尿苷(5moU)、肌苷、7-甲基鸟苷。指导RNA化学修饰的实例包括但不限于在一个或多个末端核苷酸处引入2'-O-甲基(M)、2'-O-甲基-3'-硫代磷酸酯(MS)、硫代磷酸酯(PS)、S-受约束的乙基(cEt)或2'-O-甲基-3'-硫代PACE(MSP)。这样的化学修饰指导与未修饰的指导相比可以包含增加的稳定性和增加的活性,尽管中靶相对于脱靶的特异性不是可预测的(参见, Hendel, 2015, Nat Biotechnol. 33 (9): 985-9, doi:10.1038/nbt.3290, 于2015年6月29日在线公布; Ragdarm等, 0215, PNAS, E7110-E7111; Allerson等, J. Med. Chem. 2005, 48: 901-904; Bramsen等, Front. Genet., 2012, 3: 154; Deng等, PNAS, 2015, 112: 11870-11875; Sharma等, MedChemComm., 2014, 5: 1454-1471; Hendel等, Nat. Biotechnol. (2015) 33 (9): 985-989; Li等, Nature Biomedical Engineering, 2017, 1, 0066 DOI: 10.1038/s41551-017-0066)。在一

些实施方式中,指导RNA的5'和/或3'端被多种功能部分修饰,包括荧光染料、聚乙二醇、胆固醇、蛋白质或检测标签。(参见Kelly等,2016,2016,J.Biotech.233:74-83)。在某些实施方式中,指导包含在与靶DNA结合的区域中的核糖核苷酸和与Cas9、Cpf1或C2c1结合的区域中的一个或多个脱氧核糖核苷酸和/或核苷酸类似物。在本发明的一个实施方式中,将脱氧核糖核苷酸和/或核苷酸类似物引入工程化指导结构中,例如但不限于5'和/或3'端、茎-环区和种子区。在某些实施方式中,修饰不在茎环区的5'柄中。指导的茎-环区域的5'柄中的化学修饰可以消除其功能(参见Li等,Nature Biomedical Engineering,2017,1:0066)。在某些实施方式中,指导的至少1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、35、40、45、50或75个核苷酸是化学修饰的。在一些实施方式中,在指导的3'或5'端的3-5个核苷酸被化学修饰。在一些实施方式中,仅在种子区域中引入微小修饰,例如2'-F修饰。在一些实施方式中,在指导的3'端引入2'-F修饰。在某些实施方式中,在指导的5'和/或3'端有三至五个核苷酸用2'-O-甲基(M)、2'-O-甲基-3'-硫代磷酸酯(MS)、S-受约束乙基(cEt)或2'-O-甲基-3'-硫代PACE(MSP)化学修饰。这样的修饰可以增强基因组编辑效率(参见Hendel等,Nat.Biotechnol.(2015)33(9):985-989)。在某些实施方式中,指导的所有磷酸二酯键都被硫代磷酸酯(PS)取代,用于增强基因破坏的水平。在某些实施方式中,在指导的5'和/或3'端超过5个核苷酸用2'-O-Me、2'-F或S-受约束的乙基(cEt)化学修饰。这样化学修饰的指导可以介导增强的基因破坏水平(参见Ragdarm等,0215,PNAS,E7110-E7111)。在本发明的一个实施方式中,修饰指导以在其3'和/或5'端包含化学部分。这样的部分包括但不限于胺、叠氮化物、炔烃、硫代、二苯并环辛炔(DBCO)或罗丹明。在某些实施方式中,化学部分通过接头(例如烷基链)与指导物缀合。在某些实施方式中,修饰的指导的化学部分可以用于将指导附着到另外的分子,例如DNA、RNA、蛋白质或纳米颗粒。这样化学修饰的指导可用于鉴定或富集由CRISPR系统一般编辑的细胞(参见Lee等,eLife,2017,6:e25312,DOI:10.7554)。

[0229] 在某些实施方式中,本文提供的CRISPR系统可利用包含指导序列的crRNA或类似多核苷酸,其中多核苷酸是RNA、DNA或RNA和DNA的混合物,和/或其中多核苷酸包含一种或多种核苷酸类似物。序列可包含任何结构,包括但不限于天然crRNA的结构,例如凸起、发夹或茎环结构。在某些实施方式中,包含指导序列的多核苷酸与可以是RNA或DNA序列的第二多核苷酸序列形成双链体。

[0230] 在某些实施方式中,使用化学修饰的指导RNA。指导RNA化学修饰的实例包括但不限于在一个或多个末端核苷酸处引入2'-O-甲基(M)、2'-O-甲基3'硫代磷酸酯(MS)或2'-O-甲基3'硫代PACE(MSP)。这样化学修饰的指导RNA与未修饰的指导RNA相比可以包含增加的稳定性和增加的活性,尽管中靶相对于脱靶的特异性不是可预测的。(参见,Hendel,2015,Nat Biotechnol.33(9):985-9,doi:10.1038/nbt.3290,2015年6月29日在线出版)。化学修饰的指导RNA还包括但不限于具有硫代磷酸酯键和包含核糖环的2'和4'碳之间的亚甲基桥的锁核酸(LNA)核苷酸的RNA。

[0231] 在一些实施方式中,指导序列的长度为约或大于约5、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、35、40、45、50、75或更多个核苷酸。在一些实施方式中,指导序列的长度小于约75、50、45、40、35、30、25、20、15、12或更少的核苷酸。优选地,指导序列长10至30个核苷酸。指导序列引导CRISPR复合物与靶核酸序列序列特异性结

合的能力可以通过任何合适的测定来评估。例如,足以形成CRIPSR复合物的CRISPR系统的组分,包括待测试的指导序列,可以提供给具有相应靶核酸序列的宿主细胞,例如通过用编码CRISPR序列的组分的载体转染,然后评估在靶核酸序列内的优先切割,例如通过Surveyor测定法。类似地,可以通过提供靶序列、CRIPSR复合物的组分,包括待测试的指导序列和不同于测试指导序列的对照指导序列,并且比较测试与对照指导序列反应之间靶序列处的结合或切割速率而在测试试管中评估靶RNA的切割。其他测定是可能的,并且由本领域技术人员想到。

[0232] 在一些实施方式中,对指导的修饰是化学修饰、插入、缺失或分裂。在一些实施方式中,化学修饰包括但不限于引入2'-O-甲基(M)类似物、2'-脱氧类似物、2-硫尿苷类似物、N6-甲基腺苷类似物、2'-氟代类似物、2-氨基嘌呤、5-溴-尿苷、假尿苷(Ψ)、N¹-甲基假尿苷($\text{me}^1\Psi$)、5-甲氧基尿苷(5moU)、肌苷、7-甲基鸟苷、2'-O-甲基-3'-硫代磷酸酯(MS)、S-受约束的乙基(cEt)、硫代磷酸酯(PS)或2'-O-甲基-3'-硫代PACE(MSP)。在一些实施方式中,指导包含一个或多个硫代磷酸酯修饰。在某些实施方式中,指导的至少1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20或25个核苷酸是化学修饰的。在某些实施方式中,种子区域中的一个或多个核苷酸是化学修饰的。在某些实施方式中,3'-末端的一个或多个核苷酸是化学修饰的。在某些实施方式中,5'-柄中的核苷酸均未经化学修饰。在一些实施方式中,种子区域中的化学修饰是微小修饰,例如引入2'-氟代类似物。在一个特定的实施方式中,种子区域的一个核苷酸被2'-氟类似物取代。在一些实施方式中,3'端中的5或10个核苷酸是化学修饰的。在Cpf1CrRNA的3'端的这样的化学修饰提高了基因切割效率(参见Li等,Nature Biomedical Engineering,2017,1:0066)。在一个特定的实施方式中,3'-末端的5个核苷酸被2'-氟代类似物取代。在一个特定的实施方式中,3'-末端的10个核苷酸被2'-氟代类似物取代。在一个特定的实施方式中,3'-末端的5个核苷酸被2'-O-甲基(M)类似物取代。

[0233] 在一些实施方式中,修饰指导的5'柄的环。在一些实施方式中,将指导的5'-柄的环修饰为具有缺失、插入、分裂或化学修饰。在某些实施方式中,环包含3、4或5个核苷酸。在某些实施方式中,环包含UCUU、UUUU、UAUU或UGUU的序列。

[0234] 可以选择指导序列,并因此选择核酸靶向指导RNA以靶向任何靶核酸序列。在CRISPR复合物的形成的上下文中,“靶序列”是指指导序列被设计成与其具有互补性的序列,其中靶序列与指导序列之间的杂交促进CRISPR复合物的形成。靶序列可包含RNA多核苷酸。术语“靶RNA”是指RNA多核苷酸,其是靶序列或包含靶序列。换句话说,靶RNA可以是RNA多核苷酸,或者gRNA的一部分(即指导序列)被设计为与其具有互补性并且由包含CRISPR效应蛋白和gRNA的复合物介导的效应子功能被引导至其的RNA多核苷酸的一部分。在一些实施方式中,靶序列位于细胞的细胞核或细胞质中。靶序列可以是DNA。靶序列可以是任何RNA序列。在一些实施方式中,靶序列可以是RNA分子内的序列,其选自信使RNA(mRNA)、前mRNA、核糖体RNA(rRNA)、转移RNA(tRNA)、微RNA(miRNA)、小干扰RNA(siRNA)、小核RNA(snRNA)、小核仁RNA(snoRNA)、双链RNA(dsRNA)、非编码RNA(ncRNA)、长非编码RNA(lncRNA)和小细胞质RNA(scRNA)。在一些优选的实施方式中,靶序列可以是RNA分子内的序列,其选自mRNA、前mRNA和rRNA。在一些优选的实施方式中,靶序列可以是选自ncRNA和lncRNA的RNA分子内的序列。在一些更优选的实施方式中,靶序列可以是mRNA分子或前mRNA分子内的序列。

[0235] 在某些实施方式中,指导RNA的间隔子长度小于28个核苷酸。在某些实施方式中,指导RNA的间隔长度为至少18个核苷酸且小于28个核苷酸。在某些实施方式中,指导RNA的间隔子长度为19至28个核苷酸。在某些实施方式中,指导RNA的间隔子长度为19至25个核苷酸。在某些实施方式中,指导RNA的间隔子长度为20个核苷酸。在某些实施方式中,指导RNA的间隔子长度为23个核苷酸。在某些实施方式中,指导RNA的间隔子长度为25个核苷酸。

[0236] 在某些实施方式中,切割效率的调节可以通过在间隔子序列和靶序列之间,包括沿着间隔子/靶标的错配的位置,引入错配,例如1个或多个错配,例如1个或2个错配。例如双错配越居中(即不是3'或5'),切割效率越受影响。因此,通过选择沿着间隔子的错配位置,可以调节切割效率。举例来说,如果期望小于100%的靶标切割(例如在细胞群中),在间隔子和靶序列之间的1个或多个,例如优选2个错配,可以在间隔子序列中引入。错配位置沿着间隔子越居中,切割百分比越低。

[0237] 在某些示例性实施方式中,切割效率可被利用来设计可以区分单个核苷酸变化的两种或更多种靶标的单指导,例如单核苷酸多态性(SNP)、变异、或(点)突变。CRISPR效应子可以具有对SNP(或其他单核苷酸变异)的降低的灵敏度,并继续以一定水平的效率切割SNP靶标。因此,对于两种靶标,或靶标组,可以设计具有与靶标之一即中靶SNP互补的核苷酸序列的指导RNA。指导RNA被进一步设计为具有合成错配。如本文所使用的,“合成错配”是指在天然存在的SNP的上游或下游引入的非天然存在的错配,例如上游或下游至多5个核苷酸,例如上游或下游4、3、2或1个核苷酸,优选上游或下游至多3个核苷酸,更优选上游或下游至多2个核苷酸,最优选上游或下游1个核苷酸(即邻近SNP)。当CRISPR效应子与中靶SNP结合时,只有单个错配将与合成错配形成,并且CRISPR效应子将继续被激活,并且可检测信号产生。当指导RNA与脱靶SNP杂交时,将形成两个错配,来自SNP的错配和合成错配,并且可检测信号不产生。因此,本文所公开的系统可以设计用于区分群体内的SNP。例如,本文所公开的系统可用于区分相差单个SNP的致病菌株或检测某些疾病特异性SNP,例如但不限于疾病相关SNP,例如但不限于癌症相关SNP。

[0238] 在某些实施方式中,指导RNA被设计为使得SNP位于间隔子序列的(从5'端开始)位置1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29或30。在某些实施方式中,指导RNA被设计为使得SNP位于间隔子序列(从5'端开始)的位置1、2、3、4、5、6、7、8或9。在某些实施方式中,指导RNA被设计为使得SNP位于间隔子序列(从5'端开始)的位置2、3、4、5、6或7。在某些实施方式中,指导RNA被设计为使得SNP位于间隔子序列(从5'端开始)的位置3、4、5或6。在某些实施方式中,指导RNA被设计为使得SNP位于间隔子序列(从5'端开始)的位置3。

[0239] 在某些实施方式中,指导RNA被设计为使得错配(例如合成错配,即除SNP之外的另外的突变)位于间隔子序列(从5'端开始)的位置1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29或30。在某些实施方式中,指导RNA被设计为使得错配位于间隔子序列(从5'端开始)的位置1、2、3、4、5、6、7、8或9。在某些实施方式中,指导RNA被设计为使得错配位于间隔子序列(从5'端开始)的位置4、5、6或7。在某些实施方式中,指导RNA被设计为使得错配位于间隔子序列(从5'端开始)的位置5。

[0240] 在某些实施方式中,指导RNA被设计为使得错配位于SNP的上游2个核苷酸(即一个中间核苷酸)。

[0241] 在某些实施方式中,指导RNA被设计为使得错配位于SNP的下游2个核苷酸(即一个中间核苷酸)。

[0242] 在某些实施方式中,指导RNA被设计为使得错配位于间隔子序列(从5'端开始)的位置5并且SNP位于间隔子序列(从5'端开始)的位置3。

[0243] 本文所述的实施方式包括在本文所讨论的真核细胞(体外,即在分离的真核细胞中)中诱导一个或多个核苷酸修饰,包括向细胞递送如本文所讨论的载体。突变可以包括通过指导RNA在细胞的各个靶序列处引入、缺失或置换一个或多个核苷酸。突变可以包括通过指导RNA在所述细胞的各个靶序列处引入、缺失或置换1-75个核苷酸。突变可包括通过指导RNA在所述细胞的各个靶序列处引入、缺失或置换1、5、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、35、40、45、50或75个核苷酸。突变可以包括通过指导RNA在所述细胞的各个靶序列处引入、缺失或置换5、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、35、40、45、50或75个核苷酸。突变包括通过指导RNA在所述细胞的各个靶序列处引入、缺失或置换10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、35、40、45、50或75个核苷酸。突变可以包括通过指导RNA在所述细胞的各个靶序列处引入、缺失或置换20、21、22、23、24、25、26、27、28、29、30、35、40、45、50或75个核苷酸。突变可以包括通过指导RNA在所述细胞的各个靶序列处引入、缺失或置换40、45、50、75、100、200、300、400或500个核苷酸。

[0244] 典型地,在内源CRISPR系统的上下文中,CRISPR复合物的形成(包含杂交至靶序列并与一种或多种Cas蛋白复合的指导序列)导致在靶序列内或附近(例如距离靶序列1、2、3、4、5、6、7、8、9、10、20、50或更多个碱基对内)的切割,但可以取决于例如二级结构,特别是在RNA靶标的情况下。

[0245] 基于RNA的掩蔽构建体

[0246] 如本文所用,“掩蔽构建体”是指可通过本文所述的激活的CRISPR系统效应蛋白切割或以其它方式失活的分子。术语“掩蔽构建体”替代地也可称为“检测构建体”。在某些示例性实施方式中,掩蔽构建体是基于RNA的掩蔽构建体。基于RNA的掩蔽构建体包含可被CRISPR效应蛋白切割的RNA元件。RNA元件的切割释放试剂或产生构象变化,其允许可检测信号产生。下面描述了证实RNA元件如何用于阻止或掩盖可检测信号的产生的示例性构建体,并且本发明的实施方式包括其变体。在切割之前,或当掩蔽构建体处于“活性”状态时,掩蔽构建体阻断阳性可检测信号的产生或检测。应当理解,在某些示例性实施方式中,可以在活性RNA掩蔽构建体的存在下产生很小背景信号。阳性可检测信号是可以使用本领域已知的光学、荧光、化学发光、电化学或其他检测方法检测的任何信号。术语“阳性可检测信号”用于区分在掩蔽构建体存在下可检测的其他可检测信号。例如,在某些实施方式中,当存在掩蔽剂时检测到第一信号(即阴性可检测信号),其然后在检测到靶分子和掩蔽剂被激活的CRISPR效应蛋白切割或失活时转化为第二信号(例如阳性可检测信号)。

[0247] 在某些示例性实施方式中,掩蔽构建体可以抑制基因产物的产生。基因产物可以由添加到样品中的报告构建体编码。掩蔽构建体可以是参与RNA干扰通路的干扰RNA,例如短发夹RNA(shRNA)或小干扰RNA(siRNA)。掩蔽构建体还可以包含microRNA(miRNA)。当存在时,掩蔽构建体抑制基因产物的表达。基因产物可以是在不存在掩蔽构建体时可被标记的探针、适体或抗体检测的荧光蛋白或其他RNA转录物或蛋白。在激活效应蛋白后,掩蔽构建

体被切割或以其他方式沉默而允许作为阳性可检测信号的基因产物的表达和检测。

[0248] 在某些示例性实施方式中,掩蔽构建体可以整合产生可检测阳性信号所需要的一种或多种试剂,使得从掩蔽构建体释放一种或多种试剂导致可检测阳性信号的产生。一种或多种试剂可以组合以产生比色信号、化学发光信号、荧光信号或任何其他可检测信号,并且可以包括已知适合于这样的目的的任何试剂。在某些示例性实施方式中,一种或多种试剂被结合一种或多种试剂的RNA适体整合。当检测到靶分子时效应蛋白被激活并且RNA适体被降解时,释放一种或多种试剂。

[0249] 在某些示例性实施方式中,掩蔽构建体可以以个体离散体积(下文进一步定义)固定在固体基质上并且整合单个试剂。例如,试剂可以是包含染料的珠子。当被固定的试剂整合时,个体珠子过于扩散而不能产生可检测信号,但是从掩蔽构建体释放后能够产生可检测信号,例如在溶液浓度中通过聚集或简单增加。在某些示例性实施方式中,固定的掩蔽剂是基于RNA的适体,其可以在检测到靶分子时被激活的效应蛋白切割。

[0250] 在某些其他示例性实施方式中,掩蔽构建体与溶液中的固定的试剂结合,从而阻断试剂与溶液中游离的单独标记结合配偶体结合的能力。因此,在对样品施加洗涤步骤时,可以在不存在靶分子的情况下将标记的结合配偶体从样品中洗掉。然而,如果效应蛋白被激活,则掩蔽构建体被切割至足以干扰掩蔽构建体结合试剂的能力的程度,从而允许标记的结合配偶体与固定的试剂结合。因此,标记的结合配偶体在洗涤步骤后保留,表明样品中存在靶分子。在某些方面,结合固定的试剂的掩蔽构建体是RNA适体。固定的试剂可以是蛋白质,并且标记的结合配偶体可以是标记的抗体。或者,固定的试剂可以是链霉亲和素,并且标记的结合配偶体可以是标记的生物素。在上述实施方式中使用的结合配偶体上的标记可以是本领域已知的任何可检测标记。另外,可以根据本文所述的总体设计使用其他已知的结合配偶体。

[0251] 在某些示例性实施方式中,掩蔽构建体可包含核酶。核酶是具有催化特性的RNA分子。天然核酶和工程化核酶两者包含可以被本文所公开的效应蛋白靶向的RNA或由以上组成。可以选择或工程化核酶以催化产生阴性可检测信号或防止产生阳性对照信号的反应。在通过活化的效应蛋白失活核酶后,产生阴性对照信号或防止产生阳性可检测信号的反应被移除,从而允许阳性可检测信号产生。在一个示例性实施方式中,核酶可以催化比色反应,使溶液显示为第一颜色。当核酶失活时,溶液变成第二颜色,第二颜色是可检测阳性信号。Zhao等“Signal amplification of glucosamine-6-phosphate based on ribozyme glmS,”*Biosens Bioelectron.* 2014;16:337-42描述了核酶如何用于催化比色反应的实例,并提供如何修饰这样的系统以在本文所公开的实施方式的上文中工作的实例。或者,核酶当存在时可以产生例如RNA转录物的切割产物。因此,阳性可检测信号的检测可以包括检测仅在不存在核酶的情况下产生的未切割的RNA转录物。

[0252] 在某些示例性实施方式中,一种或多种试剂是能够促进产生可检测信号,例如比色、化学发光或荧光信号的蛋白,例如酶,其被抑制或整合使得通过一种或多种RNA适体与蛋白的结合,蛋白不能产生可检测信号。在本文所公开的效应蛋白激活时,RNA适体被切割或降解到一个程度即它们不再抑制蛋白产生可检测信号的能力。在某些示例性实施方式中,适体是凝血酶抑制剂适体。在某些示例性实施方式中,凝血酶抑制剂适体具有GGGAACAAAGCUGAAGUACUUACCC (SEQ ID NO:414)的序列。当该适体被切割时,凝血酶将变得

具有活性并且将切割肽比色或荧光底物。在某些示例性实施方式中,比色底物是与用于凝血酶的肽底物共价连接的对硝基苯胺(pNA)。在被凝血酶切割后,pNA被释放并变成黄色并且容易被眼睛看到。在某些示例性实施方式中,荧光底物是可以使用荧光检测器检测的7-氨基-4-甲基香豆素蓝色荧光团。抑制性适体也可用于辣根过氧化物酶(HRP)、 β -半乳糖苷酶或小牛碱性磷酸酶(CAP),并且在上述一般原理内。

[0253] 在某些实施方式中,RNA酶活性用比色法通过酶抑制适体的切割检测。将RNA酶活性转化为比色信号的一种潜在模式是将RNA适体的切割与能够产生比色输出的酶的再活化偶联。在没有RNA切割的情况下,完整的适体将与酶靶标结合并抑制其活性。该读出系统的优点在于酶提供了另外的扩增步骤:一旦通过并行活性(例如Cas13a并行活性)从适体释放,比色酶将继续产生比色产物,导致信号的扩增。

[0254] 在某些实施方式中,使用用比色读出抑制酶的现有适体。存在多种具有比色读出的适体/酶对,例如凝血酶、蛋白C、中性粒细胞弹性蛋白酶和枯草杆菌蛋白酶。这些蛋白酶具有基于pNA的比色底物,并且可商购获得。在某些实施方式中,使用靶向共同比色酶的新型适体。常见且稳健的酶,例如 β -半乳糖苷酶、辣根过氧化物酶或小牛肠碱性磷酸酶,可以通过由选择策略(例如SELEX)设计的工程化适体靶向。这样的策略允许快速选择具有纳摩尔结合效率的适体,并且可以用于开发用于比色读出的另外的酶/适体对。

[0255] 在某些实施方式中,RNA酶活性用比色法通过RNA系连的抑制剂的切割检测。许多常见的比色酶具有竞争性的可逆抑制剂:例如, β -半乳糖苷酶可被半乳糖抑制。许多这些抑制剂很弱,但是它们的效果可以通过局部浓度的增加来增加。通过将局部浓度的抑制剂与RNA酶活性联系起来,比色酶和抑制剂对可以被工程化为RNA酶传感器。基于小分子抑制剂的比色RNA酶传感器涉及三个组分:比色酶、抑制剂,和与抑制剂和酶共价连接以将抑制剂系连在酶上的桥接RNA。在未切割的构型中,酶被小分子的增加的局部浓度所抑制;当RNA被切割时(例如通过Cas13a并行切割),抑制剂将释放并且比色酶将被激活。

[0256] 在某些实施方式中,通过比色法由G-四链体的形成和/或活化检测到RNA酶活性。DNA中的G四链体可与血红素(铁(III)-原卟啉IX)复合以形成具有过氧化物酶活性的DNA酶。当提供过氧化物酶底物(例如ABTS:(2,2'-氮杂双[3-乙基苯并噻唑啉-6-磺酸]-二铵盐))时,在过氧化氢存在下G-四链体-血红素复合物引起底物的氧化。然后在溶液中形成绿色。G四链体形成DNA序列的实例是:GGGTAGGGCGGGTTGGGA (SEQ. I. D. No. 415)。通过将RNA序列与该DNA适体杂交,G四链体结构的形成将受到限制。在RNA酶并行激活(例如C2c2-复合物并行激活)后,RNA短链将被切割,从而允许G四链体形成并且与血红素结合。该策略特别有吸引力,因为颜色形成是酶促的,这意味着除了RNA酶活化之外还存在另外外的扩大。

[0257] 在某些示例性实施方式中,掩蔽构建体可以以个体离散体积(下文进一步定义)固定在固体基质上并且螯合单个试剂。例如,试剂可以是包含染料的珠子。当被固定的试剂螯合时,个体珠子过于扩散而不能产生可检测信号,但是从掩蔽构建体释放后能够产生可检测信号,例如在溶液浓度中通过聚集或简单增加。在某些示例性实施方式中,固定的掩蔽剂是基于RNA的适体,其可以在检测到靶分子时被激活的效应蛋白切割。

[0258] 在一个示例性实施方式中,掩蔽构建体包含检测剂,其根据检测剂是否聚集或分散在溶液中而改变颜色。例如,某些纳米颗粒,例如胶体金,当它们从聚集体移动到分散的颗粒时,经历可见的紫色到红色的色移。因此,在某些示例性实施方式中,这样的检测剂可

以通过一种或多种桥分子聚集。参见例如图43。桥分子的至少一部分包含RNA。在激活本文所公开的效应蛋白后，桥分子的RNA部分被切割，允许检测剂分散并导致相应的颜色变化。见例如图46。在某些示例性实施方式中，桥分子是RNA分子。在某些示例性实施方式中，检测剂是胶体金属。胶体金属材料可以包括水不溶性金属颗粒或分散在液体、水溶胶或金属溶胶中的金属化合物。胶体金属可选自周期表中IA、IB、IIB和IIIB族的金属，以及过渡金属，尤其是VIII族的那些。优选的金属包括金、银、铝、钪、钛、钒、铬、锰、钴、铜、镓、锗、铟、铊、铅、铋、钨、铼、铂和钯。金属优选以离子形式提供，衍生自适当的金属化合物，例如 Al^{3+} 、 Ru^{3+} 、 Zn^{2+} 、 Fe^{3+} 、 Ni^{2+} 和 Ca^{2+} 离子。

[0259] 当RNA桥被活化的CRISPR效应子切割时，是观察到前述的色移。在某些示例性实施方式中，颗粒是胶体金属。在某些其他示例性实施方式中，胶体金属是胶体金。在某些示例性实施方式中，胶体纳米粒子为15nm金纳米粒子(AuNP)。由于独特的胶体金纳米颗粒的表面特性，当在溶液中完全分散时在520nm处观察到最大吸光度，并且肉眼看起来呈红色。在AuNP聚集时，它们表现出最大吸光度的红移并且颜色看起来更暗，最终作为深紫色聚集体从溶液中沉淀出来。在某些示例性实施方式中，修饰纳米颗粒以包括从纳米颗粒表面延伸的DNA接头。个体颗粒通过在RNA的每个末端与DNA接头的至少一部分杂交的单链RNA(ssRNA)桥连接在一起。因此，纳米颗粒将形成连接的颗粒和聚集体的网状物，呈现为暗沉淀物。在激活本文所公开的CRISPR效应子后，ssRNA桥将被切割，从连接网格释放AuNPS并产生可见的红色。下面列出了实例DNA接头和RNA桥接序列。DNA接头末端的硫醇接头可用于与AuNPS的表面缀合。可以使用其他形式的缀合。在某些示例性实施方式中，可以产生两个AuNP群，每个DNA接头一个。这将有助于促进具有正确方向的ssRNA桥的正确结合。在某些示例性实施方式中，第一DNA接头通过3'端缀合，而第二DNA接头通过5'端缀合。

[0260] C2c2 比色 DNA1	TTATAACTATTCCTAAAAA/3ThioMC3-D/ (SEQ. I.D. No. 183)
C2c2 比色 DNA2	/5ThioMC6-D/AAAAAAAAAACTCCCCTAATAAC AAT (SEQ. I.D. No. 184)
C2c2 比色桥	GGGUAGGAUAGUUAUAAUUUCCCUUCCCA UUGUUAUUAGGGAG (SEQ. I.D. No. 185)

[0261] 在某些其他示例性实施方式中，掩蔽构建体可以包含可检测标记与其连接的RNA寡核苷酸和该可检测标记的掩蔽剂。这样的可检测标记/掩蔽剂对的实例是荧光团和荧光团的猝灭剂。由于荧光团与另一荧光团或非荧光分子之间形成非荧光复合物，可发生荧光团的猝灭。这种机制称为基态复合物形成、静态淬灭或接触淬灭。因此，可以设计RNA寡核苷酸，使得荧光团和猝灭剂足够接近以发生接触猝灭。荧光团及其关联猝灭剂是本领域已知的，并且可由本领域普通技术人员为此目的选择。特定的荧光团/猝灭剂在本发明的上下文中并不重要，只有荧光团/猝灭剂对的选择才能确保荧光团的遮蔽。在激活本文所公开的效应蛋白后，RNA寡核苷酸被切割，从而切断维持接触猝灭效应所需的荧光团和猝灭剂之间的接近度。因此，荧光团的检测可用于确定样品中靶分子的存在。

[0262] 在某些其它示例性实施方式中,掩蔽构建体可以包含一种或多种金属纳米颗粒例如金颗粒附着至其的一种或多种RNA寡核苷酸。在一些实施方式中,掩蔽构建体包含由形成闭环的多种RNA寡核苷酸交联的多种金属纳米颗粒。在一个实施方式中,掩蔽构建体包含由形成闭环的三种RNA寡核苷酸交联的三种金纳米颗粒。在一些实施方式中,CRISPR效应蛋白切割RNA寡核苷酸导致由金属纳米颗粒产生的可检测信号。

[0263] 在某些其他示例性实施方式中,掩蔽构建体可以包含一种或多种量子点附着至其的一种或多种RNA寡核苷酸。在一些实施方式中,CRISPR效应蛋白对RNA寡核苷酸的切割导致由量子点产生的可检测信号。

[0264] 在一个示例性实施方式中,掩蔽构建体可包含量子点。量子点可具有附着于表面的多个接头分子。接头分子的至少一部分包含RNA。接头分子在一端附着至量子点,并沿着接头的长度或末端附着至一种或多种猝灭剂,使得猝灭剂保持足够接近以发生量子点的猝灭。接头可以是分支的。如上所述,量子点/猝灭剂对并不重要,只有选择量子点/猝灭剂对才能确保荧光团的掩蔽。量子点及其关联猝灭剂是本领域已知的,并且可以由本领域普通技术人员在本文所公开的效应蛋白激活后为此目的选择,接头分子的RNA部分被切割,从而消除了维持猝灭效应所需的量子点与一种或多种猝灭剂之间的接近度。在某些示例性实施方式中,量子点是链霉亲和素缀合的。RNA是通过生物素接头附着并用序列/5Bio_{sg}/UCUCGUACGUUC/3IAbrQSp/(SEQ ID NO.416)或/5Bio_{sg}/UCUCGUACGUUCUCUCGUACGUUC/3IAbrQSp/(SEQ ID NO.417)募集猝灭分子,其中/5Bio_{sg}/是生物素标签并且/3IAbrQSp/是爱荷华黑猝灭剂。在切割后,通过本文所公开的激活的效应子,量子点将可见地发荧光。

[0265] 以类似的方式,荧光能量转移(FRET)可用于产生可检测阳性信号。FRET是非辐射过程,通过该过程,来自能量激发的荧光团(即“供体荧光团”)的光子将另一分子(即“受体”)中的电子的能态提升到激发单重态的更高振动水平。供体荧光团返回基态而不发出该荧光团的荧光特征。受体可以是另一荧光团或非荧光分子。如果受体是荧光团,则转移的能量作为荧光团的荧光特征发射。如果受体是非荧光分子,则吸收的能量作为热量而损失。因此,在实施方式的上下文中如本文所公开的,荧光团/猝灭剂对被附着到寡核苷酸分子的供体荧光团/受体对替换。当完整时,如通过受体发射的荧光或热检测到的,掩蔽构建体产生第一信号(阴性可检测信号)。在激活本文所公开的效应蛋白后,RNA寡核苷酸被切割并且FRET被破坏,从而现在检测到供体荧光团的荧光(阳性可检测信号)。

[0266] 在某些示例性实施方式中,掩蔽构建体包括使用插入染料,其响应于长RNA切割为短核苷酸而改变它们的吸光度。存在多种这样的染料。例如,吡喃-Y将与RNA复合并形成在572nm处具有吸光度的复合物。RNA的切割导致吸光度损失和颜色变化。亚甲蓝可以以类似的方式使用,RNA切割后在688nm处吸光度变化。因此,在某些示例性实施方式中,掩蔽构建体包含RNA和插入染料复合物,其在通过本文所公开的效应蛋白切割RNA时改变吸光度。

[0267] 在某些示例性实施方式中,掩蔽构建体可包含用于HCR反应的引发剂。参见例如Dirks和Pierce.PNAS 101,15275-15728(2004)。HCR反应利用两种发夹物质中的势能。当具有与发夹之一上的相应区域互补的部分的单链引发剂被释放到先前稳定的混合物中,它打开一个物质的发夹。该过程进而暴露出打开其他物质的发夹的单链区域。该过程进而暴露出与原始引发剂相同的单链区域。所产生的链反应可导致形成带切口的双螺旋,其生长直至发夹供应耗尽。所得产物的检测可在凝胶上或比色法进行。示例性比色检测方法包括例

如在Lu等“Ultra-sensitive colorimetric assay system based on the hybridization chain reaction-triggered enzyme cascade amplification ACS Appl Mater Interfaces, 2017, 9 (1) :167-175, Wang等“An enzyme-free colorimetric assay using hybridization chain reaction amplification and split aptamers”Analyst 2015, 150, 7657-7662, 和Song等“Non covalent fluorescent labeling of hairpin DNA probe coupled with hybridization chain reaction for sensitive DNA detection.” Applied Spectroscopy, 70 (4) :686-694 (2016) 中公开的那些。

[0268] 在某些示例性实施方式中,掩蔽构建体可包含HCR引发剂序列和可切割结构元件,例如环或发夹,其防止从引发剂引发HCR反应。在通过激活的CRISPR效应蛋白切割结构元件后,然后释放引发剂以触发HCR反应,其检测表明样品中存在一种或多种靶标。在某些示例性实施方式中,掩蔽构建体包含具有RNA环的发夹。当活化的CRISPR效应蛋白切割RNA环时,可以释放引发剂以引发HCR反应。

[0269] 靶标扩增

[0270] 在某些示例性实施方式中,可以在激活CRISPR效应蛋白之前扩增靶RNA和/或DNA。可以使用任何合适的RNA或DNA扩增技术。在某些示例性实施方式中, RNA或DNA扩增是等温扩增。在某些示例性实施方式中,等温扩增可以是基于核酸测序的扩增(NASBA)、重组酶聚合酶扩增(RPA)、环介导的等温扩增(LAMP)、链置换扩增(SDA)、解旋酶依赖性扩增(HDA)、或切口酶扩增反应(NEAR)。在某些示例性实施方式中,可以使用非等温扩增方法,其包括但不限于PCR、多重置换扩增(MDA)、滚环扩增(RCA)、连接酶链反应(LCR)、或衍生物扩增方法(RAM)。

[0271] 在某些示例性实施方式中, RNA或DNA的扩增是NASBA,其开始于通过序列特异性反向引物逆转录靶RNA而产生RNA/DNA双链体。然后使用RNA酶H降解RNA模板,允许含有启动子(例如T7启动子)的正向引物结合并引发互补链的延伸,产生双链DNA产物。然后RNA聚合酶启动子介导的DNA模板的转录产生靶RNA序列的拷贝。重要的是,新的靶RNA各自都可以被指导RNA检测,因此进一步增强了测定的灵敏度。然后,指导RNA对靶RNA的结合导致CRISPR效应蛋白的激活,并且方法如上所述进行。NASBA反应具有能够在适度的等温条件下,例如,在约41°C进行的额外优点,使得它适用于在现场和远离临床实验室进行早期和直接检测的系统和装置。

[0272] 在某些其他示例性实施方式中,重组酶聚合酶扩增(RPA)反应可用于扩增靶核酸。RPA反应使用能够将序列特异性引物与双链体DNA中的同源序列配对的重组酶。如果存在靶DNA,则启动DNA扩增,不需要其他样品操纵,例如热循环或化学熔融。整个RPA扩增系统作为干燥制剂是稳定的,并且可以安全运输而无需冷藏。RPA反应也可以在等温温度下在37-42°C的最佳反应温度下进行。序列特异性引物被设计成扩增包含待检测的靶核酸序列的序列。在某些示例性实施方式中,将RNA聚合酶启动子(例如T7启动子)添加至引物之一。这导致包含靶序列和RNA聚合酶启动子的扩增的双链DNA产物。在RPA反应之后或期间,添加将从双链DNA模板产生RNA的RNA聚合酶。然后可以通过CRISPR效应系统检测扩增的靶RNA。以这种方式,使用本文所公开的实施方式可以检测靶DNA。RPA反应也可用于扩增靶RNA。首先使用逆转录酶将靶RNA转化为cDNA,然后进行第二链DNA合成,此时如上所述进行RPA反应。

[0273] 因此,在某些示例性实施方式中,本文所公开的系统可以包括扩增试剂。本文描述

了可用于扩增核酸的不同组分或试剂。例如，如本文所述的扩增试剂可包括缓冲液，例如 Tris 缓冲液。Tris 缓冲液可以以适合于期望的应用或使用的任何浓度使用，例如包括但不限于 1mM、2mM、3mM、4mM、5mM、6mM、7mM、8mM、9mM、10mM、11mM、12mM、13mM、14mM、15mM、25mM、50mM、75mM、1M 等的浓度。本领域技术人员将能够确定用于本发明的缓冲液例如 Tris 的适当浓度。

[0274] 盐，例如氯化镁 ($MgCl_2$)、氯化钾 (KCl) 或氯化钠 (NaCl)，可以被包含在扩增反应例如 PCR 中以改善核酸片段的扩增。尽管盐浓度将取决于特定的反应和应用，但在一些实施方式中，特定大小的核酸片段可在特定的盐浓度下产生最佳结果。较大的产物可能需要改变盐浓度，通常是较低的盐，以产生期望的结果，而较小的产物的扩增可以在较高的盐浓度下产生更好的结果。本领域技术人员将理解，盐的存在和/或浓度以及盐浓度的改变可以改变生物或化学反应的严格性，因此可以使用为本发明和如本文所述的反应提供适当的条件的任何盐。

[0275] 生物或化学反应的其他组分可以包括破开或裂解细胞以用于其中材料的分析的细胞裂解组分。细胞裂解组分可包括但不限于洗涤剂，如上所述的盐，例如 NaCl、KCl、硫酸铵 $[(NH_4)_2SO_4]$ 等。可适用于本发明的洗涤剂可包括 TritonX-100、十二烷基硫酸钠 (SDS)、CHAPS (3-[(3-胆烷基丙基) 二甲基铵]-1-丙磺酸盐)、乙基三甲基溴化铵、壬基苯氧基聚乙氧基乙醇 (NP-40)。洗涤剂的浓度可取决于具体应用，并且在某些情况下可以特异于反应。扩增反应可以包括以适合于本发明的任何浓度使用的 dNTP 和核酸引物，例如包括但不限于 100nM、150nM、200nM、250nM、300nM、350nM、400nM、450nM、500nM、550nM、600nM、650nM、700nM、750nM、800nM、850nM、900nM、950nM、1mM、2mM、3mM、4mM、5mM、6mM、7mM、8mM、9mM、10mM、20mM、30mM、40mM、50mM、60mM、70mM、80mM、90mM、100mM、150mM、200mM、250mM、300mM、350mM、400mM、450mM、500mM 等的浓度。同样地，根据本发明有用的聚合酶可以是本领域已知的和对本发明有用的任何特异性或通用聚合酶，包括 Taq 聚合酶、Q5 聚合酶等。

[0276] 在一些实施方式中，如本文所述的扩增试剂可适用于热启动扩增。在一些实施方式中，热启动扩增可能是有益的，以减少或消除衔接子分子或寡聚物的二聚化，或以其他方式防止不需要的扩增产物或假象并获得期望的产物的最佳扩增。本文描述的用于扩增的许多组分也可用于热启动扩增。在一些实施方式中，可以适当地使用适合与热启动扩增一起使用的试剂或组分代替一种或多种组合物组分。例如，可以使用在特定温度或其他反应条件下表现出期望的活性的聚合酶或其他试剂。在一些实施方式中，可以使用设计或优化用于热启动扩增的试剂，例如，聚合酶可以在转座后或在达到特定温度后被激活。这样的聚合酶可以是基于抗体的或基于适体的。如本文所述的聚合酶是本领域已知的。这样的试剂的实例可包括但不限于热启动聚合酶、热启动 dNTP 和光笼形 (photo-caged) dNTP。这样的试剂是本领域已知的并且是可获得的。本领域技术人员将能够确定适合于个体试剂的最佳温度。

[0277] 可以使用特定的热循环机械或装置来执行核酸的扩增，并且可以在单个的反应中或大量进行，使得可以同时任何期望数量的反应。在一些实施方式中，可以使用微流体或机器人装置进行扩增，或者可以使用温度的手动改变来进行扩增以实现期望的扩增。在一些实施方式中，可以进行优化以获得特定应用或材料的最佳反应条件。本领域技术人员将理解并能够优化反应条件以获得足够的扩增。

[0278] 在某些实施方式中,用本发明的方法或系统检测DNA需要在检测之前将(扩增的)DNA转录成RNA。

[0279] 靶RNA/DNA富集

[0280] 在某些示例性实施方式中,可在检测或扩增靶RNA或DNA之前首先富集靶RNA或DNA。在某些示例性实施方式中,这种富集可以通过CRISPR效应系统结合靶核酸来实现。

[0281] 目前的靶特异性富集方案在用探针杂交之前需要单链核酸。在各种优点中,本发明的实施方式可以跳过该步骤并且能够直接靶向双链DNA(部分或完全双链)。此外,本文所公开的实施方式是酶驱动的靶向方法,其提供允许等温富集的更快的动力学和更容易的工作流程。在某些示例实施方式中,富集可以发生在低至20-37°C的温度。在某些示例实施方式中,在单个测定中使用针对不同靶核酸的指导RNA组,允许检测多个靶标和/或单个靶标的多个变体。

[0282] 在某些示例性实施方式中,死亡CRISPR效应蛋白可以结合在溶液中的靶核酸并且随后从所述溶液中分离。例如,可以使用特异性结合死亡CRISPR效应蛋白的抗体或其他分子(例如适体)从溶液中分离与靶核酸结合的死亡CRISPR效应蛋白。

[0283] 在其他示例性实施方式中,死亡CRISPR效应蛋白可以与固体基质结合。固定的基质可以指适合于或可以被修饰以适合多肽或多核苷酸的附着的任何材料。可能的基质包括但不限于玻璃和改性的功能化玻璃、塑料(包括丙烯酸树脂、聚苯乙烯和苯乙烯与其他材料的共聚物、聚丙烯、聚乙烯、聚丁烯、聚氨酯、Teflon™等)、多糖、尼龙或硝化纤维素、陶瓷、树脂、二氧化硅或二氧化硅基材料(包括硅和改性硅)、碳、金属、无机玻璃、塑料、光纤纤维束、以及各种其他聚合物。在一些实施方式中,固体支持物包含适合于以有序图案固定分子的图案化表面。在某些实施方式中,图案化表面是指固体载体的暴露层中或上的不同区域的排列。在一些实施方式中,固体支持物包含表面中的孔阵列或凹陷。固体支持物的组成和几何形状可随其用途而变化。在一些实施方式中,固体支持物是平面结构,例如载玻片、芯片、微芯片和/或阵列。这样,基质的表面可以是平面层的形式。在一些实施方式中,固体支持物包含流通池的一个或多个表面。本文使用的术语“流通池”指包含一种或多种流体试剂可以流动跨越的固体表面的腔室。例如,在Bentley等Nature 456:53-59(2008),W004/0918497,U.S.7,057,026;W0 91/06678;W0 07/123744;US7,329,492;US 7,211,414;US 7,315,019;U.S.7,405,281,和US2008/0108082中描述了可以容易地在本公开的方法中使用的示例性流通池和相关的流体系统和检测平台。在一些实施方式中,固体支持物或其表面是非平面的,例如管或容器的内表面或外表面。在一些实施方式中,固体支持物包括微球或珠子。“微球”、“珠子”、“颗粒”旨在表示在固体基质的上下文中,意指由各种材料制成的小的离散颗粒,包括但不限于塑料、陶瓷、玻璃和聚苯乙烯。在某些实施方式中,微球是磁性的微球或珠子。或者或另外地,珠子可以是多孔的。珠子的尺寸范围从纳米(例如100nm)到毫米(例如1mm)。

[0284] 含有或怀疑含有靶核酸的样品然后可以暴露于基质以允许靶核酸与结合的死亡CRISPR效应蛋白的结合。然后可以洗去非靶分子。在某些示例性实施方式中,然后可以使用本文所公开的方法从CRISPR效应蛋白/指导RNA复合物释放靶核酸用于进一步检测。在某些示例性实施方式中,可以首先如本文所述扩增靶核酸。

[0285] 在某些示例性实施方式中,CRISPR效应子可以用结合标签标记。在某些示例性实

施方式中,可以对CRISPR效应子进行化学标签化。例如,CRISPR效应子可以是化学生物素化的。在另一个示例性实施方式中,可以通过向CRISPR效应子添加编码融合体的另外的序列来产生融合体。这样的融合的一个实例是AviTag™,它采用在独特的15个氨基酸肽标签上的单个生物素的高度靶向酶促缀合。在某些实施方式中,CRISPR效应子可以用捕获标签标记,例如但不限于GST、Myc、血凝素(HA)、绿色荧光蛋白(GFP)、flag、His标签、TAP标签和Fc标签。无论是融合体、化学标签还是捕获标签,结合标签一旦已经结合靶核酸或将CRISPR效应系统固定在固体基质上,就可以用于拉下CRISPR效应系统。

[0286] 在某些示例性实施方式中,指导RNA可以用结合标签标记。在某些示例性实施方式中,可以使用引入一种或多种生物素化核苷酸(例如生物素化的尿嘧啶)的体外转录(IVT)标记整个指导RNA。在一些实施方式中,可以将生物素化学或酶促地添加至指导RNA,例如将一个或多个生物素基团添加至指导RNA的3'端。结合标签可用于在结合发生后拉下指导RNA/靶核酸复合物,例如,通过将指导RNA/靶核酸暴露于链霉亲和素包被的固体基质。

[0287] 因此,在某些示例性实施方式中,工程化或非天然存在的CRISPR效应子可用于富集目的。在一个实施方式中,修饰可以包括效应蛋白的一个或多个氨基酸残基的突变。一个或多个突变可以在效应蛋白的一个或多个催化活性结构域中。效应蛋白与缺乏所述一个或多个突变的效应蛋白相比可具有降低或消除的核酸酶活性。效应蛋白可能不指导靶基因座处RNA链的切割。在优选的实施方式中,一个或多个突变可包括两个突变。在优选的实施方式中,一个或多个氨基酸残基在C2c2效应蛋白(例如工程化或非天然存在的效应蛋白或C2c2)中被修饰。在特定的实施方式中,一个或多个修饰的突变氨基酸残基是C2c2中对应于R597、H602、R1278和H1283例如突变R597A、H602A、R1278A和H1283A(参考LshC2c2氨基酸)的那些中的一个或多个,或LshC2c2直系同源物中相应的氨基酸残基。

[0288] 在特定的实施方式中,一个或多个修饰的突变氨基酸残基是C2c2中对应于K2、K39、V40、E479、L514、V518、N524、G534、K535、E580、L597、V602、D630、F676、L709、I713、R717(HEPN)、N718、H722(HEPN)、E773、P823、V828、I879、Y880、F884、Y997、L1001、F1009、L1013、Y1093、L1099、L1111、Y1114、L1203、D1222、Y1244、L1250、L1253、K1261、I1334、L1355、L1359、R1362、Y1366、E1371、R1372、D1373、R1509(HEPN)、H1514(HEPN)、Y1543、D1544、K1546、K1548、V1551、I1558(根据到C2c2共同编号)的那些中的一个或多个。在某些实施方式中,一个或多个修饰的突变的氨基酸残基是C2c2中对应于R717和R1509的那些中的一个或多个。在某些实施方式中,一个或多个修饰的突变氨基酸残基是C2c2中对应于K2、K39、K535、K1261、R1362、R1372、K1546和K1548的那些中的一个或多个。在某些实施方式中,所述突变导致蛋白质具有改变或修饰的活性。在某些实施方式中,所述突变导致蛋白质具有降低的活性,例如降低的特异性。在某些实施方式中,所述突变导致蛋白质不具有催化活性(即“死亡的”C2c2)。在实施方式中,所述氨基酸残基对应于LshC2c2氨基酸残基,或来自不同物种的C2c2蛋白的相应氨基酸残基。可以促进这些步骤的装置。在一些实施方式中,为了减小Cas13b效应子和一个或多个功能结构域的融合蛋白的大小,可以截短Cas13b效应子的C末端,同时仍保持其RNA结合功能。例如,至少20个氨基酸、至少50个氨基酸、至少80个氨基酸、或至少100个氨基酸、或至少150个氨基酸、或至少200个氨基酸、或至少250个氨基酸、或至少300个氨基酸、或至少350个氨基酸、或至多120个氨基酸、或至多140个氨基酸、或至多160个氨基酸、或至多180个氨基酸、或至多200个氨基酸、或至多250个氨基酸、或至多300个

氨基酸、或至多350个氨基酸、或最多400个氨基酸可以在Cas13b效应子的C末端截短。Cas13b截短的具体实例包括C-末端 Δ 984-1090、C-末端 Δ 1026-1090和C-末端 Δ 1053-1090、C-末端 Δ 934-1090、C-末端 Δ 884-1090、C-末端 Δ 834-1090、C-末端 Δ 784-1090、和C-末端 Δ 734-1090,其中氨基酸位置对应于普氏菌属种P5-125Cas13b蛋白的氨基酸位置。

[0289] 上述富集系统也可用于耗尽某些核酸的样品。例如,可以设计指导RNA以结合非靶RNA以从样品中除去非靶RNA。在一个示例性实施方式中,可以设计指导RNA以结合确实携带特定核酸变异的核酸。例如,在给定样品中可以预期更高拷贝数的非变体核酸。因此,本文所公开的实施方式可用于从样品中去除非变体核酸,以提高检测CRISPR效应系统可检测给定样品中的靶变体序列的效率。

[0290] 放大和/或增强可检测阳性信号

[0291] 在某些示例性实施方式中,可以引入进一步放大可检测阳性信号的进一步的修饰。例如,激活的CRISPR效应蛋白并行激活可用于产生次要靶标或另外的指导序列,或两者。在一个示例性实施方式中,反应溶液将含有以高浓度掺入的次要靶标。次要靶标可以不同于主要靶标(即测定被设计用于检测的靶标),并且在某些情况下可以在所有反应体积中是共同的。可以保护次要靶标的第二指导序列,例如通过二级结构特征,例如具有RNA环并且不能结合第二靶标或CRISPR效应蛋白质的发夹。通过激活的CRISPR效应蛋白切割保护基团(即在通过与溶液中的主要靶标形成复合物激活后)和与溶液中的游离CRISPR效应蛋白的复合物的形成和从掺入的次要靶标中的激活。在某些其他示例性实施方式中,在针对次要靶标序列的第二指导序列的情况下,使用类似的概念。次要靶标序列可以受到次要靶标上的结构特征或保护基团的保护。然后将保护基团切割脱离次要靶标允许形成另外的CRISPR效应蛋白/第二指导序列/次要靶标复杂物。在又一个示例性实施方式中,可以使用主要靶标激活的CRISPR效应蛋白来切割受保护或环化的引物,然后将其释放以在编码第二指导序列、次要靶标序列或两者的模板上进行等温扩增反应,例如本文所公开的那些。随后转录该扩增的模板将产生更多的第二指导序列和/或次要靶标序列,然后是另外的CRISPR效应蛋白并行激活。

[0292] 蛋白质的检测

[0293] 除了检测核酸之外,本文所公开的系统、装置和方法通过掺入特定配置的多肽检测适体还可以适于检测多肽(或其它分子)。多肽检测适体不同于上面讨论的掩蔽构建体适体。首先,适体被设计为特异性结合一种或多种靶分子。在一个示例性实施方式中,靶分子是靶多肽。在另一个示例性实施方式中,靶分子是靶化学化合物,例如靶向治疗分子。用于设计和选择对给定靶标具有特异性的适体的方法,例如SELEX,是本领域已知的。除了对给定靶标的特异性外,适体还被进一步设计为引入RNA聚合酶启动子结合位点。在某些示例性实施方式中,RNA聚合酶启动子是T7启动子。在将适体结合至靶标之前,RNA聚合酶位点不能被RNA聚合酶接近或以其他方式识别。然而,适体被配置成使得在靶标结合后,适体的结构经历构象变化,使得RNA聚合酶启动子随后暴露。RNA聚合酶启动子下游的适体序列充当用于通过RNA聚合酶产生触发RNA寡核苷酸的模板。因此,适体的模板部分可以进一步整合识别给定适体及其靶标的条形码或其他识别序列。然后可以设计如上所述的指导RNA以识别这些特异性触发寡核苷酸序列。指导RNA与触发寡核苷酸的结合激活CRISPR效应蛋白,其继续使掩蔽构建体失活并产生如前所述的阳性可检测信号。

[0294] 因此,在某些示例性实施方式中,本文所公开的方法包括另外的步骤:将样品或样品组分配到个体离散体积的组中,个体离散体积各自包含肽检测适体、CRISPR效应蛋白、一种或多种指导RNA、掩蔽构建体,和在足以允许检测适体与一种或多种靶分子结合的条件下孵育样品或样品组,其中适体与相应靶标的结合导致RNA聚合酶启动子结合位点的暴露,使得通过RNA聚合酶与RNA聚合酶启动子结合位点的结合启动触发RNA的合成。

[0295] 在另一个示例性实施方式中,适体的结合可以在适体与靶多肽结合后暴露引物结合位点。例如,适体可以暴露RPA引物结合位点。因此,引物的添加或包含将进入扩增反应,例如上文概述的RPA反应。

[0296] 在某些示例性实施方式中,适体可以是构象转换适体,其在结合感兴趣的靶标后可以改变二级结构并暴露单链DNA的新区域。在某些示例性实施方式中,这些单链DNA的新区域可以用作连接的底物,延伸适体并产生可以使用本文所公开的实施方式特异性地检测的更长的ssDNA分子。适体设计可以进一步与三元复合物结合用于检测低表位靶标,例如葡萄糖(Yang等,2015年:<http://pubs.acs.org/doi/abs/10.1021/acs.analchem.5b01634>)。实例构象转换适体和相应的指导RNA(crRNA)如下所示。

[0297]

凝血酶适体	(SEQ. I.D. No. 186)
凝血酶连接探针	(SEQ. I.D. No. 187)
凝血酶 RPA 正向 1 引物	(SEQ. I.D. No. 188)
凝血酶 RPA 正向 2 引物	(SEQ. I.D. No. 189)
凝血酶 RPA 反向 1 引物	(SEQ. I.D. No. 190)
凝血酶 crRNA1	(SEQ. I.D. No. 191)
凝血酶 crRNA2	(SEQ. I.D. No. 192)
凝血酶 crRNA3	(SEQ. I.D. No. 193)
PTK7 全长扩增子控制	(SEQ. I.D. No. 194)
PTK7 适体	(SEQ. I.D. No. 195)
PTK7 连接探针	(SEQ. I.D. No. 196)
PTK7RPA 正向 1 引物	(SEQ. I.D. No. 197)
PTK7RPA 反向 1 引物	(SEQ. I.D. No. 198)
PTK7 crRNA1	(SEQ. I.D. No. 1991)

[0298]

PTK7 crRNA2	(SEQ. I.D. No. 200)
PTK7 crRNA3	(SEQ. I.D. No. 201)

[0299] 装置

[0300] 本文所述的系统可以在诊断装置上具体化。可以使用许多基质和配置。装置可能能够在装置内定义多个个体离散体积。如本文所用“个体离散体积”是指离散空间,例如容器(container)、容器(receptacle)或可以由防止和/或抑制靶分子迁移的性质限定的其他限定体积或空间,例如由物理性质例如壁限定的体积或空间,例如孔、管的壁或液滴的表面,其可以是不可渗透的或半渗透的,或者由例如化学、扩散速率限制、电磁或光照射或其任何组合的其他手段限定的可以在限定空间内包含样品。个体离散体积可以通过分子标签鉴定,例如核酸条形码。“扩散速率受限”(例如扩散限定的体积)是指仅某些分子或反应可以接近,因为扩散约束有效地限定了空间或体积,如同其中扩散将限制靶分子从一个流到另一个流的迁移的两个平行层流的情况。“化学”定义的体积或空间是指由于其化学或分子性质(例如大小)例如通过珠子的表面电荷、基质尺寸或可以允许选择可以进入珠子内部的其他物理性质而仅某些靶分子可以存在的空间,其中例如凝胶珠可以排除某些种类而不是其他种类进入珠子。“电-磁”定义的体积或空间是指其中靶分子或其支持物的电磁特性(例如电荷或磁性)可用于限定空间中的某些区域(例如在磁场内或直接在磁铁上捕获磁性粒子)的空间。“光学上”限定的体积是指可以通过用可见光、紫外线、红外线或其他波长的光照射它来限定的任何空间区域,使得仅可以标记限定的空间或体积内的靶分子。使用非壁或半透性离散体积的一个优点是一些试剂,例如缓冲液、化学活化剂或其他试剂可以通过离散体积,而其他材料例如靶分子可以保持在离散体积或空间。通常,离散体积将包括在允许标记的条件适合于用可转位核酸标识符标记靶分子的流体介质(例如,水溶液、油、缓冲液和/或能够支持细胞生长的培养基)。在所公开的方法中有用的示例性离散体积或空间包括液滴(例如微流体液滴和/或乳液液滴)、水凝胶珠或其他聚合物结构(例如聚乙二醇二丙烯酸酯珠或琼脂糖珠)、组织载玻片(例如具有由化学、光学或物理手段定义的特定区域、体积或空间的固定福尔马林石蜡包埋的组织载玻片)、具有由有序阵列或随机图案的沉积试剂定义的区域、显微镜载玻片、管(例如离心管、微量离心管、试管、比色杯、锥形管等)、瓶子(例如玻璃瓶、塑料瓶、陶瓷瓶、锥形瓶、闪烁瓶等)、孔(例如板中的孔)、板、移液管或移液管尖端等。在某些实施方式中,隔室是油包水乳液中的水性液滴。在特定的实施方式中,本文所述的需要精确或均匀体积的任何应用、方法或系统可以使用声学液体分配器。

[0301] 在某些示例性实施方式中,装置包括在其上可以限定多个斑点的柔性材料基质。适用于诊断和生物传感的柔性基质材料在本领域中是已知的。柔性基质材料可以由植物衍生的纤维制成,例如纤维素纤维,或者可以由柔性聚合物制成,例如柔性聚酯薄膜和其他聚合物类型。在各个限定的斑点内,将本文所述的系统的试剂应用于个体斑点。各个斑点可以包含相同的试剂,除了不同的指导RNA或指导RNA组,或者在适用的情况下,不同的检测适体,以一次筛选多种靶标。因此,本文的系统 and 装置可以能够筛选来自多个来源的样品(例如来自不同个体的多个临床样品)中相同靶标或有限数量的靶标的存在,或者筛选单个样品的等分试样(或来自相同来源的多个样品)中样品中多种不同靶标的存在。在某些示例性实施方式中,将本文所述的系统的元件冷冻干燥到纸或布基质上。Pardee等Cell. 2016, 165 (5):1255-66和Pardee等Cell. 2014, 159 (4):950-54 Pardee公开了可用于某些示例性装置的示例性基于柔性材料的基质。在Shevkoplyas等的题为“Paper based diagnostic test”

的国际专利申请公开号W0/2013/071301, Siegel等的题为“Paper-based microfluidic systems”的美国专利申请公开号2011/0111517, 和Shafiee等“Paper and Flexible Substrates as Materials for Biosensing Platforms to Detect Multiple Biotargets”Scientific Reports 5:8719 (2015) 中公开了与生物流体(包括血液)一起使用的合适的基于柔性材料的基质。Wang等“Flexible Substrate-Based Devices for Point-of-Care Diagnostics”Cell 134 (11) :909-21 (2016) 公开了另外的基于柔性的材料, 包括适用于可穿戴诊断装置的那些。其他柔性基材料可包括硝化纤维素、聚碳酸酯、甲基乙基纤维素、聚偏二氟乙烯(PVDF)、聚苯乙烯或玻璃(参见例如US20120238008)。在某些实施方式中, 离散体积被疏水表面分开, 例如但不限于蜡、光致抗蚀剂或固体墨水。

[0302] 在一些实施方式中, 可以提供用作传感器或指示器的放射量测定器或徽章, 其使得将对某些微生物或其他试剂的暴露通知给穿戴者。例如, 本文所述的系统可用于检测特定病原体。同样地, 上文公开的基于适体的实施方式可用于检测多肽以及特定适体可结合的其他试剂(例如化学试剂)两者。这样的装置可用于监视士兵或其他军事人员, 以及临床医生、研究人员、医院工作人员等, 以便尽快提供与暴露于潜在危险因素有关的信息, 例如用于生物或化学战剂检测。在其他实施方式中, 这种监视徽章可用于防止免疫受损患者、烧伤患者、接受化疗的患者、儿童或老年人暴露于危险微生物或病原体。

[0303] 可使用本文所述的系统和装置进行分析的样品来源包括受试者或环境样品的生物样品。环境样品可包括表面或流体。生物样品可包括但不限于唾液、血液、血浆、血清、粪便、尿液、痰液、粘液、淋巴液、滑液、脊髓液、脑脊髓液、皮肤或粘膜的拭子, 或其组合。在一个示范性实施方式中, 环境样品取自固体表面, 例如用于制备食物或其他敏感组合物和材料的表面。

[0304] 在其他示范性实施方式中, 本文所述的系统的元件可以放置在一次性使用的基质上, 例如用于擦拭表面或样品流体的拭子或布。例如, 该系统可用于通过擦拭食品(例如水果或蔬菜)的表面来测试食物上病原体的存在。类似地, 一次性使用的基质可用于擦拭其他表面以检测某些微生物或试剂, 例如用于安全性筛选。一次性使用的基质也可以在取证中具有应用, 其中CRISPR系统被设计用于检测, 例如鉴定可用于鉴定嫌疑人的DNA SNP或某些组织或细胞标志物以确定样品中存在的生物物质的类型。同样地, 一次性使用的基质可用于从患者收集样品-例如来自口腔的唾液样品-或皮肤的拭子。在其他实施方式中, 可以从肉制品中取出样品或拭子, 以检测在肉制品上或肉制品内是否存在污染物。

[0305] 食品、临床、工业和其他环境设置需要近实时微生物诊断(参见例如LuTK, Bowers J和Koeris MS., Trends Biotechnol. 2013 Jun; 31 (6) :325-7)。在某些实施方式中, 本发明用于使用对病原体(例如, 空肠弯曲杆菌、产气荚膜梭菌、沙门氏菌属、大肠杆菌、蜡样芽孢杆菌、单核细胞增生性李斯特氏菌、志贺氏菌属、金黄色葡萄球菌、金黄色葡萄球菌肠炎、链球菌、霍乱弧菌、副溶血性弧菌、创伤弧菌、小肠结肠炎耶尔森菌和假结核耶尔森氏菌、布鲁氏菌属、溃疡性棒状杆菌、伯氏疟原虫或Plesiomonas shigelloides)特异的指导RNA快速检测食源性病原体。

[0306] 在某些实施方式中, 装置是流动条或包括流动条。例如, 侧向流动条允许通过颜色检测RNA酶(例如C2c2)。修饰RNA报告子以具有连接至5'端的第一分子(例如FITC)和连接至3'端的第二分子(例如生物素)(或反之亦然)。侧向流动条被设计为具有两条捕获线, 在第

一线处杂交的抗第一分子抗体(例如抗FITC)和在第二下游线处杂交的抗第二分子(例如抗生物素)抗体。当反应沿着条带向下流动时,未切割的报告子将在第一个捕获线处结合抗第一分子抗体,而切割的报告子将释放第二分子并允许在第二捕获线处结合第二分子。第二分子夹心抗体,例如与纳米颗粒缀合,例如金纳米颗粒,将在第一或第二线处结合任何第二分子并导致强读出/信号(例如颜色)。随着更多的报告子被切割,更多的信号将在第二捕获线处累积,并且在第一线处将出现更少的信号。在某些方面,本发明涉及如本文所述的流动条带用于检测核酸或多肽的用途。在某些方面,本发明涉及用本文定义的流动条带检测核酸或多肽的方法,例如(侧)流测试或(侧)流免疫色谱测定。

[0307] 在某些示例性实施方式中,装置是产生和/或合并不同液滴的微流体装置(即个体离散体积)。例如,可以形成包含待筛选的样品的第一组液滴,并且形成包含本文所述的系统的元件的第二组液滴。然后合并第一组和第二组液滴,然后在合并的液滴组上执行如本文所述的诊断方法。本文所公开的微流体装置可以是基于硅酮的芯片并且可以使用多种技术制造,包括但不限于热压花、弹性体模制、注射成型、LIGA、软光刻、硅制造和相关的薄膜加工技术。用于制造微流体装置的合适材料包括但不限于环烯烃共聚物(COC)、聚碳酸酯、聚(二甲基硅氧烷)(PDMS)和聚(甲基丙烯酸酯)(PMMA)。在一个实施方式中,PDMS中的软光刻可用于制备微流体装置。例如,可以使用限定基板内流动通道、阀门和过滤器的位置的光刻制造模具。将基质材料倒入模具中并使其凝固以形成印模。然后将印模密封到固体支撑物上,例如但不限于玻璃。由于一些聚合物例如PDMS吸收一些蛋白质并且可以抑制某些生物过程的疏水性质,钝化剂可以是必要的(Schoffner等*Nucleic Acids Research*,1996,24:375-379)。合适的钝化剂是本领域已知的,包括但不限于硅烷、聚对二甲苯、n-十二烷基-b-D-甘露糖苷(DDM)、普朗尼克、Tween-20、其他类似表面活性剂、聚乙二醇(PEG)、白蛋白、胶原蛋白和其他类似的蛋白质和多肽。

[0308] 在某些示例性实施方式中,系统和/或装置可以适于转换为流式细胞术读数或在单个实验中允许对数百万个细胞进行所有灵敏和定量测量,并改进现有的基于流动的方法,例如PrimeFlow测定。在某些示例性实施方式中,可以将细胞形成含有未聚合的凝胶单体的液滴,然后将其形成适于通过流式细胞术分析的单细胞液滴。可以将包含荧光可检测标记的检测构建体投到包含未聚合的凝胶单体的液滴中。在凝胶单体聚合时在液滴内形成珠子。因为凝胶聚合是通过自由基形成的,荧光报告子变得与凝胶共价结合。可以进一步修饰检测构建体以包含接头,例如胺。可在凝胶形成后添加猝灭剂,并通过接头与报告构建体结合。因此,当报告子被CRISPR效应蛋白切割时,猝灭剂不与凝胶结合并且自由扩散。可以通过偶联检测构建体至杂交链反应(HCR引发剂)来实现液滴中信号的放大。DNA/RNA杂合发夹可以引入可以包含具有RNA酶敏感结构域的发夹环凝胶中。通过保护具有RNA酶敏感结构域的发夹环内的链置换位点,可以在CRISPR效应蛋白切割发夹环后选择性去保护HCR引发剂。在通过位点介导的链置换对HCR引发剂进行去保护后,可以将荧光HCR单体洗涤到凝胶中以使得信号放大,其中引发剂被去保护。

[0309] 可以在本发明的上下文中使用的微流体装置的实例描述于Hour等“Direct Detection and drug-resistance profiling of bacteremias using inertial microfluidics”*Lap Chip*.15(10):2297-2307(2016)。

[0310] 在本文描述的系统中,可以进一步结合到评估临床环境外的受试者的生物样品

(例如生物流体)的可穿戴医疗装置中,并将该测定的结果远程报告给医疗护理专业人员可访问的中央服务器。该装置可以具有自动采样血液的能力,例如在Peeters等题为“Needle-free Blood Draw”的美国专利申请公开号2015/0342509和Andrew Conrad.的题为“Nanoparticle Phoresis”的美国专利申请公开号2015/0065821中描述的装置。

[0311] 在某些示例性实施方式中,装置可包括个体的孔,例如微孔板孔。微孔板孔的尺寸可以是标准6、24、96、384、1536、33456或9600尺寸孔的尺寸。在某些示例性实施方式中,可以将本文所述的系统的元件冷冻干燥并在分配和使用之前施加到孔的表面。

[0312] 本文所公开的装置还可包括入口和出口端口,或开口,其又可连接到阀、管、通道、腔室和注射器和/或泵用于将流体引入装置和从装置中抽出流体。这些装置可以连接到允许流体在微流体装置内定向移动的流体流动致动器。示例性致动器包括但不限于注射泵、机械致动的再循环泵、电渗泵、灯泡、波纹管、隔膜或旨在迫使流体运动的气泡。在某些示例性实施方式中,装置连接到具有一起工作以使流体移动通过装置的可编程阀的控制器。在某些示例性实施方式中,装置连接到下文进一步详细讨论的控制器。这些装置可通过终止于金属销中以插入装置上的入口端口的管道连接到流动致动器、控制器和样品加载装置。

[0313] 如本文所示的系统的元件当冷冻干燥时是稳定的,因此也考虑不需要支持装置的实施方式,即系统可被应用到支持本文所公开的反应并且允许用于检测来自表面或溶液的阳性可检测信号的任何表面或流体。除冷冻干燥外,系统还可以稳定储存并以粒状形式使用。用于形成合适的粒化形式的聚合物是本领域已知的。

[0314] 在某些实施方式中,CRISPR效应蛋白与装置中的个体离散体积结合。个体离散体积可包含对不同靶分子特异的不同指导RNA。在某些实施方式中,将样品暴露于包含多于一个离散体积的固体基质,离散体积各自包含对靶分子特异的指导RNA。不受理论束缚,每种指导RNA将从样品中捕获其靶分子,并且样品不需要分成单独的测定。因此,可以保留有价值的样品。效应蛋白可以是包含亲和和标签的融合蛋白。亲和和标签是本领域公知的(例如,HA标签、Myc标签、Flag标签、His标签、生物素)。效应蛋白可以与生物素分子连接,离散体积可以包含链霉亲和素。在其他实施方式中,CRISPR效应蛋白由对效应蛋白特异的抗体结合。先前已经描述了结合CRISPR酶的方法(参见,例如US20140356867A1)。

[0315] 本文所公开的装置还可以包括本领域中已知的用于通过其它方法分析样品的护理点(POC)的元件的装置。参见,例如St John和Price,“Existing and Emerging Technologies for Point-of-Care Testing”(Clin Biochem Rev.2014Aug;35(3):155-167)。

[0316] 本发明可与无线实验室芯片(LOC)诊断传感器系统一起使用(参见例如,美国专利号9470699“Diagnostic radio frequency identification sensors and applications thereof”)。在某些实施方式中,本发明在由无线装置(例如手机、个人数字助理(PDA)、平板电脑)控制的LOC中执行,并且将结果报告给所述装置。

[0317] 射频识别(RFID)标签系统包括RFID标签,其发送数据以供RFID读取器接收(也称为询问机)。在典型的RFID系统中,单个物体(例如,商店商品)配备有包含应答器的相对小的标签。应答器具有被赋予独特的电子产品代码的存储器芯片。RFID读取器通过使用通信协议发出激活标签内的应答器的信号。因此,RFID读取器能够读取和写入标签的数据。另外,RFID标签读取器根据RFID标签系统应用处理数据。目前,存在无源和有源型RFID标签。

无源型RFID标签不包含内部电源,而是由从RFID读取器接收的射频信号供电。或者,有源型RFID标签包含内部电源,这使有源型RFID标签具有更大的传输范围和存储容量。无源标签与有源标签的使用取决于特定应用。

[0318] 芯片实验室技术在科学文献中有很充分的描述,由多个微流体通道、输入或化学孔组成。孔中的反应可以使用射频识别(RFID)标签技术来测量,因为来自RFID电子芯片的导电引线可以直接连接到每个测试孔。天线可以印刷或安装在电子芯片的另一层中或直接安装在装置的背面。此外,引线、天线和电子芯片可以嵌入LOC芯片中,从而防止电极或电子器件的短路。由于LOC允许复杂的样品分离和分析,因此该技术允许LOC测试独立于复杂或昂贵的读取器完成。而是可以使用例如蜂窝电话或PDA的简单无线装置。在一个实施方式中,无线装置还控制微流体通道的分离和控制,以进行更复杂的LOC分析。在一个实施方式中,LED和其他电子测量或感测装置包括在LOC-RFID芯片中。不受理论束缚,这种技术是一次性的,允许在实验室外进行需要分离和混合的复杂的测试。

[0319] 在优选的实施方式中,LOC可以是微流体装置。LOC可以是无源芯片,其中芯片通过无线装置供电和控制。在某些实施方式中,LOC包括用于保持试剂的微流体通道和用于引入样品的通道。在某些实施方式中,来自无线装置的信号将功率传递给LOC并激活样品和测定试剂的混合。具体地,在本发明的情况下,系统可包括掩蔽剂、CRISPR效应蛋白和对靶分子特异的指导RNA。在激活LOC后,微流体装置可以混合样品和测定试剂。在混合时,传感器检测信号并将结果发送到无线装置。在某些实施方式中,去掩蔽剂是导电RNA分子。导电RNA分子可以附着到导电材料上。导电分子可以是导电纳米颗粒、导电蛋白、附着在蛋白质或乳胶或其他导电珠上的金属颗粒。在某些实施方式中,如果使用DNA或RNA,则导电分子可以直接连接到匹配的DNA或RNA链。可以在传感器上检测导电分子的释放。该测定可以是一步过程。

[0320] 由于可以精确地测量表面区域的电导率,因此可以在一次性无线RFID电测定中获得定量结果。此外,测试区域可以非常小,允许在给定的区域中进行更多测试,因此节省了成本。在某些实施方式中,使用各自与不同的CRISPR效应蛋白结合的单独的传感器和固定在传感器上的指导RNA来检测多个靶分子。不受理论束缚,可以通过无线装置区分不同传感器的激活。

[0321] 除了本文所述的导电方法之外,可以使用依赖于RFID或蓝牙作为用于一次性RFID测定的基本低成本通信和电力平台的其他方法。例如光学装置可以被用于评估给定的靶分子的存在和水平。在某些实施方式中,光学传感器检测荧光掩蔽剂的去掩蔽。

[0322] 在某些实施方式中,本发明的装置可包括用于诊断性读取测定的手持便携式装置(参见例如Vashist等,Commercial Smartphone-Based Devices and Smart Applications for Personalized Healthcare Monitoring and Management,Diagnosics 2014,4(3),104-128;mReader from Mobile Assay;and Holomic Rapid Diagnostic Test Reader)。

[0323] 如本文所指出的,某些实施方式允许通过比色变化的检测,当实施方式在POC情况下或在其中获得更复杂的检测装置以读出信号可能是有限的使用的资源贫乏的环境中使用,其具有某些附带的益处。然而,本文所公开的便携式实施方式还可以与能够检测可见光范围之外的信号的手持式分光光度计结合。Das等“Ultra-portable,wireless smartphone spectrophotometer for rapid,non-destructive testing of fruit ripeness.”Nature Scientific Reports.2016,6:32504,DOI:10.1038/srep32504描述了

可以与本发明结合使用的手持式分光光度计装置的实例。最后,在利用基于量子点的掩蔽构建体的某些实施方式中,由于量子点提供的接近完全的量子产率,可以成功地使用手持UV光或其他合适的装置来检测信号。

[0324] 实例方法和测定

[0325] 该测定平台的低成本和适应性将自身提供给许多应用,包括(i)一般RNA/DNA/蛋白质定量,(ii)快速、多重RNA/DNA和蛋白质表达检测,以及(iii)临床和环境样品的靶核酸、肽和蛋白质的灵敏检测。另外,本文所公开的系统可以适用于检测生物环境,例如细胞中的转录物。鉴于本文所述的CRISPR效应子的高度特异性,可以追踪转录物的等位基因特异性表达或活细胞中的疾病相关突变。

[0326] 在某些示例性实施方式中,将特异于单个靶标的单个指导序列放置在单独的体积中。然后各个体积可以接收不同样品或相同样品的等分试样。在某些示例性实施方式中,可以将各自针对单独的靶标的多种指导序列置于单个孔中,使得可以在不同的孔中筛选多种靶标。为了在单个体积中检测多种指导RNA,在某些示例性实施方式中,可以使用具有不同特异性的多种效应蛋白。例如,可以使用具有不同序列特异性的不同直系同源物。例如,一种直系同源物可以优先切割A,而其他直系同源物优先切割C、G、U/T。因此,可以产生完全包含单个核苷酸或包含单个核苷酸的实质部分的掩蔽构建体,其各自具有可以在不同波长下检测的不同的荧光团。以这种方式,可以在单个个体离散体积中筛选多达四种不同的靶标。在某些示例性实施方式中,可以使用来自相同类CRISPR效应蛋白的不同直系同源物,例如两种Cas13a直系同源物,两种Cas13b直系同源物,或两种Cas13c直系同源物。各种Cas13蛋白的核苷酸偏好如图67所示。在某些其他示例性实施方式中,可以使用具有不同核苷酸编辑偏好的不同直系同源物,例如Cas13a和Cas13b直系同源物,或Cas13a和Cas13c直系同源物,或Cas13b直系同源物和Cas13c直系同源物等。在某些示例性实施方式中,使用具有聚U偏好的Cas13蛋白和具有聚A偏好的Cas13蛋白。在某些示例性实施方式中,具有聚U偏好的Cas13蛋白是*Prevotella intermedia* Cas13b。具有聚A偏好的Cas13蛋白是*Prevotella* sp.MA2106Cas13b蛋白(PsmCas13b)。在某些示例性实施方式中,具有聚U偏好的Cas13蛋白是*Leptotrichia wadei* Cas13a(LwaCas13a)蛋白,并且具有聚A偏好的Cas13蛋白是*Prevotella* sp.MA2106Cas13b蛋白。在某些示例性实施方式中,具有聚U偏好的Cas13蛋白是*Capnocytophaga canimorsus* Cas13b蛋白(CcaCas13b)。

[0327] 除了单碱基编辑偏好,可以基于Cas13直系同源物的其他基序切割偏好来设计另外的检测构建体。例如,Cas13直系同源物可以优先切割二核苷酸序列、三核苷酸序列或包含4、5、6、7、8、9或10个核苷酸基序的更复杂基序。因此,使用本文所公开的实施方式的多重测定的上限主要受到可区分可检测标记的数量的限制。用于鉴定这样的基序的示例性方法在以下工作实施例中进一步公开。

[0328] 如本文所证实的,CRISPR效应系统能够检测低至阿摩尔浓度的靶分子。参见例如图13、14、19、22和下文描述的工作实施例。由于所述系统的灵敏度,需要快速和灵敏检测的许多应用可受益于本文所公开的实施方式,并且考虑在本发明的范围内。实例测定和应用在下文进一步详细描述。

[0329] 微生物应用

[0330] 在某些示例性实施方式中,本文所公开的系统、装置和方法涉及检测样品(例如从

受试者获得的生物样品)中一种或多种微生物剂的存在。在某些示例性实施方式中,微生物可以是细菌、真菌、酵母、原生动物、寄生虫或病毒。因此,本文所公开的方法可以适用于其他方法(或组合),所述其他方法需要快速鉴定微生物种类,监测微生物蛋白质(抗原)、抗体,抗体基因的存在,检测某些表型(例如细菌药物抗性),监测疾病进展和/或疾病暴发,以及抗生素筛查。由于本文所公开的实施方式的快速且灵敏的诊断能力,检测低至单核苷酸差异的微生物种类型,以及作为POC装置部署的能力,本文所公开的实施方式可用于指导治疗方案,例如作为适当的抗生素或抗病毒的选择。本文所公开的实施方式还可用于筛选环境样品(空气、水、表面、食物等)中微生物污染的存在。

[0331] 公开了鉴定微生物种的方法,例如细菌、病毒、真菌、酵母或寄生物种等。本文所公开的具体实施方式描述了识别和区分单个样品内或多个样品中的微生物种的方法和系统,允许识别许多不同的微生物。本方法允许通过检测样品中靶核酸序列的存在在生物或环境样品中检测病原体并且区分一种或多种生物(例如,细菌、病毒、酵母、原生动物和真菌或其组合)的两种或多种物种。从样品获得的阳性信号表明存在微生物。通过使用多于一种效应蛋白,使用本发明的方法和系统可以同时鉴定多种微生物,其中效应蛋白各自靶向特定的微生物靶序列。以这种方式,可以对特定受试者进行多水平分析,其中可以一次检测任何数量的微生物。在一些实施方式中,可以使用可以鉴定一种或多种微生物种的探针组进行多种微生物的同时检测。

[0332] 样品的多重分析可以大规模检测样品,减少分析的时间和成本。然而,多重分析通常受到生物样品可用性的限制。然而,根据本发明,可以进行多重分析的替代方案,使得可以将多种效应蛋白添加到单个样品中,并且掩蔽构建体可以各自与单独的猝灭染料组合。在这种情况下,可以分别从每种猝灭染料获得阳性信号以在单个样品中进行多次检测。

[0333] 本文公开了用于区分样品中一种或多种生物的两种或更多种物种的方法。该方法也适用于检测样品中一种或多种生物的一种或多种物种。

[0334] 微生物检测

[0335] 在一些实施方式中,提供了用于检测样品中微生物的方法,包括将样品或样品组分配到一个或多个个体离散体积中,所述个体离散体积包含如本文所述的CRISPR系统;在足以允许一种或多种指导RNA与一种或多种微生物特异性靶标结合的条件下孵育样品或样品组;通过一种或多种指导RNA与一种或多种靶分子结合来激活CRISPR效应蛋白,其中激活CRISPR效应蛋白导致基于RNA的掩蔽构建体的修饰,从而产生可检测阳性信号;检测可检测阳性信号,其中可检测阳性信号的检测表明样品中存在一种或多种靶分子。一种或多种靶分子可以是mRNA、gDNA(编码或非编码)、trRNA或rRNA,其包含可用于将两种或更多种微生物物种/菌株彼此区分的靶核苷酸潮汐序列。可以设计指导RNA以检测靶序列。本文所公开的实施方式还可以利用某些步骤来改善指导RNA与靶RNA序列之间的杂交。用于增强核糖核酸杂交的方法公开在题为“Enhanced Methods of Ribonucleic Acid Hybridization”的W02015/085194中,其通过引用并入本文。微生物特异性靶标可以是RNA或DNA或蛋白质。如果DNA方法可以进一步包括使用引入如本文所述的RNA聚合酶启动子的DNA引物。如果靶标是蛋白质,则该方法将利用适体和本文所述的蛋白质检测特异性步骤。

[0336] 检测单核苷酸变体

[0337] 在一些实施方式中,可以使用特异于并结合本文所述的靶序列的指导RNA来检测

一种或多种鉴定的靶序列。本发明的系统和方法甚至可以区分不同微生物物种之间存在的单核苷酸多态性,因此,根据本发明使用多种指导RNA可以进一步扩展或改善可用于在物种之间区分的靶序列的数量。例如,在一些实施方式中,一种或多种指导RNA可以区分种、属、科、目、纲、门、界或表型或其组合的微生物。

[0338] 基于rRNA序列的检测

[0339] 在某些示例性实施方式中,本文所公开的装置、系统和方法可用于区分样品中的多种微生物物种。在某些示例性实施方式中,鉴定可以基于核糖体RNA序列,包括16S、23S和5S亚基。用于鉴定相关rRNA序列的方法公开于美国专利申请公开号2017/0029872中。在某些示例性实施方式中,指导RNA组可以设计为通过各物种或菌株独特的可变区来区分各物种。还可以设计指导RNA以靶向区分属、科、目、纲、门、界水平的微生物或其组合的RNA基因。在使用扩增的某些示例性实施方式中,可以设计扩增引物组以侧接核糖体RNA序列的恒定区和设计用于通过可变内部区域区分各物种的指导RNA。在某些示例性实施方式中,引物和指导RNA可以被设计为分别在16S亚基中保守和可变区。也可以使用在物种或物种子集中唯一可变的其他基因或基因组区域,例如RecA基因家族、RNA聚合酶 β 亚基。其他合适的系统发育标志物及其鉴定方法例如在Wu等arXiv:1307.8690[q-bio.GN]中讨论。

[0340] 在某些示例性实施方式中,设计方法或诊断以同时筛选跨多种系统发育和/或表型水平的微生物。例如,该方法或诊断可包括使用具有不同指导RNA的多种CRISPR系统。第一组指导RNA可以区分例如分枝杆菌、革兰氏阳性菌和革兰氏阴性菌。这些一般种类可以进一步细分。例如,可以设计指导RNA并将其用于区分革兰氏阴性细菌中肠道和非肠道的方法或诊断中。可以设计第二组指导RNA以区分属或种水平的微生物。因此,可以产生识别所有分枝杆菌、革兰氏阳性、革兰氏阴性(进一步分为肠道和非肠道)的基质,其中在给定的样品中鉴定的每种细菌的种的属落入这些种类之一。前述仅用于示例目的。还考虑了用于分类其他微生物类型的其他手段,并且其将遵循上述总体结构。

[0341] 筛选药物抗性

[0342] 在某些示例性实施方式中,本文所公开的装置、系统和方法可用于筛选感兴趣的微生物基因,例如抗生素和/或抗病毒抗性基因。可以设计指导RNA以区分已知的感兴趣基因。然后可以使用本文所公开的用于检测这些基因的实施方式筛选样品,包括临床样品。在POC处筛查药物抗性的能力将在选择适当的治疗方案方面具有巨大的益处。在某些示例性实施方式中,抗生素抗性基因是碳青霉烯酶,包括KPC、NDM1、CTX-M15、OXA-48。其他抗生素抗性基因是已知的,并且可以在例如Comprehensive Antibiotic Resistance Database (Jia等“CARD 2017:expansion and model-centric curation of the Comprehensive Antibiotic Resistance Database.”Nucleic Acids Research,45,D566-573)中找到。

[0343] 利巴韦林是一种有效的抗病毒药物,可以抵抗许多RNA病毒。一些临床上重要的病毒已经进化对利巴韦林具有抗性,包括口蹄疫病毒DOI:10.1128/JVI.03594-13;脊髓灰质炎病毒(Pfeifer and Kirkegaard.PNAS,100(12):7289-7294,2003);和丙型肝炎病毒((Pfeiffer and Kirkegaard,J.Virol.79(4):2346-2355,2005)。许多其它持续RNA病毒,例如肝炎和HIV,已经进化对现有的抗病毒药物具有抗性:乙型肝炎病毒(拉米夫定、替诺福韦、恩替卡韦)doi:10/1002/hep22900;丙肝病毒(替拉瑞韦,BILN2061,ITMN-191,Sch6,波西普韦,AG-021541,ACH-806)doi:10.1002/hep.22549;和HIV(许多药物抗性突变)

hivb.stanford.edu。本文所公开的实施方式可用于检测这样的变体等。

[0344] 除了药物抗性之外,还有许多临床相关的突变可用本文所公开的实施方式检测,例如LCMV中的持续性与急性感染([doi:10.1073/pnas.1019304108](https://doi.org/10.1073/pnas.1019304108)),以及埃博拉的感染性增加(Diehl等Cell.2016,167(4):1088-1098)。

[0345] 如本文其他地方所述,可以通过在gRNA中引入合成错配来区分密切相关的微生物物种(例如,在给定的靶序列中仅具有单个核苷酸差异)。

[0346] 组覆盖方法

[0347] 在特定的实施方式中,设计了可以鉴定例如限定组的微生物内的所有微生物物种的指导RNA组。在某些示例性实施方式中,可以将如本文所述的用于产生指导RNA的方法与W02017/040316中公开的方法进行比较,所述方法通过引用并入本文。如W02017040316中所述,组覆盖溶液可以鉴定覆盖整个靶序列或靶序列组(例如基因组序列组)所需的最小数量的靶序列探针或指导RNA。先前已使用组覆盖方法来鉴定引物和/或微阵列探针,通常在20至50碱基对范围内。参见,例如,Pearson等cs.virginia.edu/~robins/papers/primers_dam11_final.pdf., Jabado等Nucleic Acids Res.2006 34(22):6605-11, Jabado等Nucleic Acids Res.2008,36(1):e3doi10.1093/nar/gkm1106, Duitama等Nucleic Acids Res.2009,37(8):2483-2492, Phillippy等BMC Bioinformatics.2009,10:293doi:10.1186/1471-2105-10-293。然而,这样的方法通常涉及将每个引物/探针作为k聚体处理并且使用后缀阵列搜索精确匹配或允许不精确匹配。另外,该方法通常采用二元方法通过选择引物或探针来检测杂交,使得每种输入序列仅需要被一种引物或探针结合,并且该结合沿着序列的位置是无关紧要的。替代方法可以将靶基因组划分为预定义的窗口,并且在二元方法下有效地将每个窗口视为单独的输入序列——即它们确定给定的探针或指导RNA是否在各窗口内结合,并且要求所有窗口都被某些探针或指导RNA的状态所束缚。实际上,这些方法将组覆盖问题中“通用”的各元件视为整个输入序列或输入序列的预定义窗口,并且如果探针或指导RNA的起始在元件内结合,则各元件被认为是“覆盖的”。这些方法限制了允许不同探针或指导RNA设计覆盖给定靶序列的流动性。

[0348] 相比之下,本文所公开的实施方式涉及检测更长的探针或指导RNA长度,例如,在适合于杂交选择测序的70bp至200bp的范围中。此外,本文W02017/040316所公开的方法可以应用于采用能够定义探针或指导RNA组的泛靶序列方法,其可以在大的和/或可变的靶序列组中鉴定和促进所有物种和/或菌株序列的检测测序。例如,本文所公开的方法可用于在单个测定中鉴定给定病毒或多种不同病毒的所有变体。此外,本文所公开的方法将组覆盖问题中“通用”的各元件视为靶序列的核苷酸,并且只要探针或指导RNA结合至包含该元件的靶基因组的一些区段,则各元件被认为是“覆盖的”。可以使用这些类型的组覆盖方法代替先前方法的二元方法,本文所公开的方法更好地模拟探针或指导RNA如何与靶序列杂交。不仅仅询问给定的指导RNA序列是否与给定的窗口结合,这些方法可以用于检测杂交模式——即其中给定的探针或指导RNA与一种或多种靶序列结合——然后从这些杂交模式中确定覆盖靶序列集至足以使样品富集和测序任何和所有靶序列两者的程度所需的最小数量的探针或指导RNA。可以通过定义使损失函数最小化的某些参数来确定这些杂交模式,从而使得能够以允许参数针对各物种而变化的方式(例如反映各物种的多样性)以及使用组覆盖解决方案的直接应用无法实现的计算有效方式(例如先前在探针或指导RNA设计背景

中应用的那些) 鉴定最小探针或指导RNA组。

[0349] 检测多种转录物丰度的能力可以允许产生指示特定表型的独特微生物特征。可以使用各种机器学习技术来导出基因标签。因此,CRISPR系统的指导RNA可用于鉴定和/或定量由基因标签定义的生物标志物的相对水平,以检测某些表型。在某些示例性实施方式中,基因标签表示对抗生素的灵敏度、对抗生素的抗性或其组合。

[0350] 在本发明的一个方面,方法包括检测一种或多种病原体。以这种方式,可以获得个体微生物对受试者的感染之间的区分。在一些实施方式中,这样的区分可以使临床医生能够检测或诊断特定疾病,例如疾病的不同变体。优选地,病原体序列是病原体的基因组或其片段。该方法还可以包括确定病原体的进化。确定病原体的进化可包括鉴定病原体突变,例如核苷酸缺失、核苷酸插入、核苷酸取代。在后者中,存在非同义、同义和非编码替换。在爆发期间,突变更常见于非同义。该方法可以进一步包括确定如上所述分析的两种病原体序列之间的置换率。无论突变是有害的还是适应性的,都需要进行功能分析,然而,非同义突变的发生率表明,这种流行病的持续发展可以为病原体的适应提供机会,强调了快速遏制的必要性。因此,该方法可以进一步包括评估病毒适应的风险,其中确定非同义突变的数量。(Gire,等,Science 345,1369,2014)。

[0351] 监测微生物爆发

[0352] 在一些实施方式中,如本文所述的CRISPR系统或其使用方法可用于确定病原体爆发的演变。该方法可以包括从来自一个或多个受试者的多个样品中检测一种或多种靶序列,其中靶序列是来自导致爆发的微生物的序列。这样的方法可以进一步包括确定病原体传播的模式,或由病原体引起的疾病暴发中涉及的机制。

[0353] 病原体传播的模式可以包括来自病原体的天然宿主的持续的新传播或者在从天然库的单次传播之后的受试者-受试者传播(例如,人-人传播)或两者的混合。在一个实施方式中,病原体传播可以是细菌或病毒传播,在这样的情况下,靶序列优选是微生物基因组或其片段。在一个实施方式中,病原体传播的模式是病原体传播的早期模式,即在病原体爆发之初。确定爆发开始时病原体传播的模式增加了尽早停止疫情的可能性,从而降低了本国和国际传播的可能性。

[0354] 确定病原体传播的模式可以包括根据本文描述的方法检测病原体序列。确定病原体传播的模式可以进一步包括检测受试者之间的病原体序列的共享的宿主内变异,并确定共享的宿主内变异是否显示时间模式。观察到的宿主内和宿主间变异的模式提供了关于传播和流行病学的重要见解(Gire,等,2014)。

[0355] 在显示时间模式的受试者之间检测共享的宿主内变异是受试者之间(特别是人之间)的传递链接的指示,因为它可以通过来自多个来源的受试者感染(重复感染)、样品污染反复突变(有或没有平衡选择以增强突变)、或者通过传播链中较早发生的突变而产生的稍微不同的病毒的共同传播(Park等,Cell 161(7):1516-1526,2015)来解释。检测受试者之间共享的宿主内变异可包括检测位于共同单核苷酸多态性(SNP)位置的宿主内变体。位于共同(SNP)位置的宿主内变体的阳性检测指示重复感染和污染作为宿主内变体的主要解释。可以基于作为宿主间变体出现的SNP频率来分离重复感染和污染(Park等,2015)。否则可以排除重复感染和污染。在后一种情况下,检测受试者之间共享的宿主内变异可以进一步包括评估同义和非同义变体的频率并将同义和非同义变体的频率相互比较。非同义突变

是改变蛋白质的氨基酸、可能导致经受自然选择的微生物的生物学变化的突变。同义置换不改变氨基酸序列。同义和非同义变体的等频率指示宿主内变体中性进化。如果同义和非同义变体的频率是不同的,则可能通过平衡选择来维持宿主内变体。如果同义和非同义变体的频率较低,则指示复发突变。如果同义和非同义变体的频率很高,则指示共传播(Park等,2015)。

[0356] 与埃博拉病毒一样,拉沙病毒(LASV)可引起出血热,病死率高。Andersen等从临床和啮齿动物储库样品中产生了近200个LASV序列的基因组目录(Andersen等,Cell Volume162,Issue 4,p738-750,2015年8月13日)。Andersen等显示,2013-2015年EVD流行病是由人对人的传播推动,而LASV感染主要是由于库对人的感染。Andersen等阐明了LASV在西非的传播,并表明这种迁移伴随着LASV基因组丰度、死亡率、密码子适应和转化效率的变化。该方法可以进一步包括在系统发育上比较第一病原体序列与第二病原体序列,以及确定在第一和第二病原体序列之间是否存在系统发育关系。第二病原体序列可以是较早的参考序列。如果存在系统发育关联,则该方法可以进一步包括将第一病原体序列的系统发育生根于第二病原体序列。因此,可以构建第一病原体序列的谱系。(Park等,2015)。

[0357] 该方法可以进一步包括确定突变是有害的还是适应性的。有害突变指示传播受损的病毒和死端感染,因此通常仅存在于个体受试者中。对一个个体受试者独特的突变是在系统发育树的外部分支上发生的突变,而内部分支突变是存在于多个样品中的突变(即在多个受试者中)。更高的非同义置换率是系统发育树的外部分支的特征(Park等,2015)。

[0358] 在系统发育树的内部分支中,选择有更多机会过滤掉有害的突变体。根据定义,内部分支产生了多个后代谱系,因此不太可能包含具有适应性成本的突变。因此,较低的非同义置换率表明内部分支(Park等,2015)。

[0359] 可能对健康影响较小的同义突变以具可比性的频率发生在内部和外部分支上(Park等,2015)。

[0360] 通过分析测序的靶序列,例如病毒基因组,可以发现导致流行病发作严重程度机制,例如2014年埃博拉病毒爆发期间。例如,Gire等对2014年爆发的基因组与早期爆发的所有20个基因组进行了系统发育比较,表明2014年西非病毒可能在过去十年中从中非传播开来。使用与其他埃博拉病毒基因组的差异来植根系统发育是有问题的(6,13)。然而,在最古老的爆发上植根树,显示样品日期与根尖距离之间存在很强的相关性,每年每个位点的置换率为 8×10^{-4} (13)。这表明,最近三次爆发的谱系在2004年左右大致同时与共同祖先不同,这支持了这样的假设:每次爆发都代表了来自天然库中同一遗传多样性病毒群体的独立人畜共患事件。他们还发现,2014年EBOV爆发可能是由于天然库的单次传播造成的,随后是爆发期间人对人的传播。他们的研究结果还表明,Sierra Leon的流行病可能源于同时从几内亚引入两种遗传上不同的病毒(Gire等,2014)。

[0361] 确定拉沙病毒是如何从其起源点扩散出来的也是可能的,特别是由于人对人的传播,甚至回溯了400年前这种传播的历史(Andersen等,Cell 162(4):738-50,2015)。

[0362] 关于2013-2015EBOV爆发期间所需的工作以及医疗人员在爆发现场遇到的困难,并且更一般地,本发明的方法使得可以使用较少的选择的探针进行测序,使得可以加速测序,从而缩短从采样到结果采购所需的时间。此外,可以将试剂盒和系统设计成可在现场使用,从而可以容易地进行患者的诊断,而无需将样品发送或运送到国家或世界的另一部分。

[0363] 在上述任何方法中,对靶序列或其片段进行测序可以使用上述任何测序过程。此外,对靶序列或其片段进行测序可以是近实时测序。可以根据先前描述的方法(实验程序:Matranga等,2014;和Gire等,2014)进行靶序列或其片段的测序。对靶序列或其片段进行测序可以包括多个靶序列的平行测序。对靶序列或其片段进行测序可包括Illumina测序。

[0364] 分析与一种或多种所选探针杂交的靶序列或其片段可以是鉴定分析,其中所选探针与靶序列或其片段的杂交表明样品内存在靶序列。

[0365] 目前,初级诊断基于患者的症状。然而,各种疾病可能具有相同的症状,因此诊断依赖于统计学。例如,疟疾引发流感样症状:头痛、发烧、发抖、关节疼痛、呕吐、溶血性贫血、黄疸、尿液中的血红蛋白、视网膜损伤和抽搐。这些症状也常见于败血症、肠胃炎和病毒性疾病。在后者中,埃博拉出血热具有以下症状:发热、喉咙痛、肌肉疼痛、头痛、呕吐、腹泻、皮疹、肝脏和肾脏功能下降、内部和外部出血。

[0366] 当患者被提供到医疗单位,例如在热带非洲,基本诊断将结束于疟疾,因为统计上,疟疾是非洲该地区内最可能的疾病。因此,患者接受疟疾治疗,尽管患者实际上可能没有感染疾病并且患者最终没有得到正确治疗。缺乏正确的治疗可能会危及生命,特别是当患者感染的疾病呈现快速进化时。在医务人员意识到给予患者的治疗无效并且进行正确的诊断并对患者进行适当的治疗之前可能为时已晚。

[0367] 本发明的方法对这种情况提供了解决方案。实际上,因为指导RNA的数量可以显著减少,这使得可以在单个芯片上提供分成组的所选择的探针(每组对一种疾病是特异性的),使得多种疾病例如病毒感染可以是同时被诊断出来。由于本发明,可以同时单个芯片上诊断超过3种疾病,优选地超过4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20种疾病,优选地是在给定地理区域的人口中最常发生的疾病。由于每个组中选择的探针是特定于诊断的疾病之一,可以进行更准确的诊断从而减少了对患者进行错误治疗的风险。

[0368] 在其它情况下,疾病例如病毒感染可能没有发生任何症状,或已造成的症状,但在患者被呈现给医务人员之前它们逐渐消失。在这样的情况下,患者不会寻求任何医疗帮助,或者由于在呈现当天没有症状而诊断很复杂。

[0369] 本发明也可以与其他方法一起用于诊断疾病,鉴定病原体和优化基于核酸检测的治疗,例如在粗的、未纯化的样品的mRNA。

[0370] 本发明的方法还提供了解决这种情况的强大工具。实际上,由于多组各自特异于在给定区域的群体内发生的最常见疾病之一的所选择的指导RNA被包含在单个诊断中,医务人员仅需要将取自患者的生物样品与芯片接触。阅读芯片可以发现患者感染的疾病。

[0371] 在某些情况下,患者呈现给医务人员为特定症状的诊断。本发明的方法不仅可以识别哪种疾病引起这些症状,而且可以确定患者是否患有他不知道的另一种疾病。

[0372] 当搜索爆发的机制时,该信息可能是最重要的。实际上,具有相同病毒的患者组也显示出时间模式,表明受试者对受试者的传播联系。

[0373] 筛选微生物遗传扰动

[0374] 在某些示例性实施方式中,本文所公开的CRISPR系统可用于筛选微生物遗传扰动。这样的方法可以是有用的,例如用于绘制微生物通路和功能网络。可以对微生物细胞进行遗传修饰,然后在不同的实验条件下进行筛选。如上所述,本文所公开的实施方式可以以多重方式筛选单个样品中的多种靶分子,或在单个个体离散体积中的单靶标。可以修饰遗

传修饰的微生物以包括鉴定由特定微生物细胞或微生物细胞群携带的特定遗传修饰核酸条形码序列。条形码是用作标识符的短核苷酸序列(例如,DNA、RNA或其组合)。核酸条形码可以具有4-100个核苷酸的长度,并且可以是单链或双链的。用条形码鉴定细胞的方法是本领域已知的。因此,本文所述的CRISPR效应系统的指导RNA可用于检测条形码。阳性可检测信号的检测指示样品中存在特定的遗传修饰。本文所公开的方法可以与用于检测互补基因型或表型读数(指示在测试的实验条件下遗传修饰的影响)的其他方法组合。待筛选的遗传修饰可包括但不限于基因敲入、基因敲除、倒位、易位、转座、或一个或多个核苷酸插入、缺失、置换、突变或添加编码具有功能结果的表位的核酸,例如改变蛋白质稳定性或检测。以类似的方式,本文所述的方法可以用于合成生物学应用,以筛选基因调控元件和基因表达模块的特定排列的功能。

[0375] 在某些示例性实施方式中,该方法可用于筛选亚等位基因。产生的亚等位基因及其在鉴定关键细菌功能基因和鉴定新抗生素治疗剂中的用途,如2016年11月4日提交的题为“Multiplex High-Resolution Detection of Micro-organism Strains,Related Kits,Diagnostic Methods and Screening Assays”的PCT/US2016/060730中所公开,其通过引用并入本文。

[0376] 不同的实验条件可以包括将微生物细胞暴露于不同的化学试剂、化学试剂的组合、不同浓度的化学试剂或化学试剂的组合、暴露于化学试剂或化学试剂的组合的不同的持续时间、不同的物理参数或两者。在某些示例性实施方式中,化学试剂是抗生素或抗病毒剂。待筛选的不同物理参数可包括不同温度、大气压、不同大气和非大气气体浓度、不同pH水平、不同培养基组合物或其组合。

[0377] 筛选环境样品

[0378] 本文所公开的方法还可以用于通过检测靶核酸或多肽的存在来筛选环境样品的污染物。例如,在一些实施方式中,本发明提供了检测微生物的方法,包括:将如本文所述的CRISPR系统暴露于样品;通过将一种或多种指导RNA与一种或多种微生物特异性靶RNA或一种或多种触发RNA结合来激活RNA效应蛋白,从而产生可检测阳性信号。可以检测阳性信号并指示样品中存在一种或多种微生物。在一些实施方式中,CRISPR系统可以在如本文所述的基质上,并且基质可以暴露于样品。在其他实施方式中,可以将相同的CRISPR系统和/或不同的CRISPR系统应用于基质上的多个离散位置。在其他实施方式中,不同CRISPR系统可以在每个位置检测不同的微生物。如上面进一步详细描述,基质可以是柔性材料基质,例如,包括但不限于纸基质,织物基质或柔性聚合物基质。

[0379] 根据本发明,可通过在待样品化的流体中暂时浸渍基质、通过将待测试的流体应用于基质、或通过将待测试的表面接触基质将基质被动地暴露于样品。可以适当地使用将样品引入基质的任何方法。

[0380] 如本文所述,用于本发明中使用的样品可以是生物或环境样品,例如食物样品(新鲜水果或蔬菜、肉类)、饮料样品、纸表面、织物表面、金属表面、木材表面、塑料表面、土壤样品、淡水样品、废水样品、盐水样品、暴露于大气或其他气体样品、或其组合。例如,可以擦拭由任何材料制成的家用/商用/工业表面,包括但不限于金属、木材、塑料、橡胶等,并测试污染物。可以测试土壤样品中是否存在致病细菌或寄生虫或其他微生物,用于环境目的和/或用于人类、动物或植物疾病测试。可以评估水样品,例如淡水样品、废水样品或盐水样品的

清洁度和安全性、和/或可饮用性,以检测例如隐孢子虫 (*Cryptosporidium parvum*)、兰伯氏贾第虫 (*Giardia lamblia*) 或其他微生物污染物的存在。在其他实施方式中,生物样品可以从以下来源获得,包括但不限于组织样品、唾液、血液、血浆、血清、粪便、尿液、痰液、粘液、淋巴液、滑液、脑脊髓液、腹水、胸腔积液、血清肿,脓或皮肤或粘膜表面拭子。在一些特定的实施方式中,环境样品或生物样品可以是粗制样品和/或在应用该方法之前可以不从样品中纯化或扩增一种或多种靶分子。微生物的鉴定对于任何数量的应用可能是有用的和/或需要的,因此可以根据本发明使用来自本领域技术人员认为合适的任何来源的任何类型的样品。

[0381] 在一些实施方式中,检查餐馆或其他食品提供者中的细菌(例如大肠杆菌)对食物的污染;食物表面;检测沙门氏菌、弯曲杆菌或大肠杆菌等病原体的水分;还要检查制造商和监管机构的食品质量以确定肉类来源的纯度;识别军团菌等病原体的空气污染;检查啤酒是否受到小球菌和乳酸杆菌等病原体的污染或破坏;在制造过程中被细菌或真菌污染巴氏杀菌或未巴氏杀菌的干酪。

[0382] 根据本发明的微生物可以是病原微生物或导致食物或消费产品变质的微生物。病原微生物可能是致病的或对人类、动物或植物不利。对于人类或动物的目的,微生物可能导致疾病或导致生病。本发明的动物或兽医应用可以鉴定被微生物感染的动物。例如,本发明的方法和系统可鉴定具有病原体的伴侣动物,包括但不限于犬舍咳、狂犬病病毒和心丝虫。在其他实施方式中,本发明的方法和系统可用于育种目的的亲子鉴定。植物微生物可能导致植物受到伤害或疾病、产量降低或改变性状,例如果味、味道、稠度、气味、食物或消费品污染目的,微生物可能对食品或消费品的味道、气味、颜色、一致性或其他商业性质产生不利影响。在某些示例性实施方式中,微生物是细菌物种。细菌可以是嗜冷菌 (*psychotroph*)、大肠杆菌、乳酸菌、或孢子形成细菌。在某些示例性实施方式中,细菌可以是引起疾病或生病,或以其他方式导致不期望的产品或性状的任何细菌物种。根据本发明的细菌可以对人、动物或植物具有致病性。

[0383] 样品类型

[0384] 用于本文所公开的方法的合适样品包括从生物体或其部分获得的任何常规生物样品,例如植物、动物、细菌等。在特定的实施方式中,生物样品获自动物受试者,例如人受试者。生物样品是从任何生物体获得、排泄或分泌的任何固体或流体样品,包括但不限于单细胞生物,例如细菌、酵母、原生动物和变形虫等,多细胞生物(例如植物或动物,包括来自健康或表面健康的人类受试者或受待诊断或调查的病症或疾病影响的人类患者的样品,例如病原微生物例如病原细菌或病毒的感染)。例如,生物样品可以是例如血液、血浆、血清、尿液、粪便、痰液、粘液、淋巴液、滑液、胆汁、腹水、胸腔积液、血清肿、唾液、脑脊液、水性或玻璃体液、或任何身体分泌物、渗出液、渗出液(例如,从脓肿或任何其他感染或炎症部位获得的液体)中获得的生物液体或从关节(例如,正常关节或受疾病影响的关节,例如类风湿性关节炎、骨关节炎、痛风或脓毒性关节炎)获得的液体,或皮肤或粘膜表面的拭子。

[0385] 样品也可以是从任何器官或组织获得的样品(包括活组织检查或尸体解剖标本,例如肿瘤活检)或者可以包含细胞(原代细胞或培养的细胞)或由任何细胞、组织或器官调理的培养基。示例性的样品包括但不限于,细胞、细胞裂解物、血涂片、细胞离心制剂、细胞学涂片、体液(例如血液、血浆、血清、唾液、痰、尿、支气管肺泡灌洗、精液等)、组织活检(例

如肿瘤活组织检查)、细针抽吸物和/或组织切片(例如低温恒温器组织切片和/或石蜡包埋的组织切片)。在其他实例中,样品包括循环肿瘤细胞(其可通过细胞表面标志物鉴定)。在特定的实例中,样品直接使用(例如新鲜或冷冻),或可在使用前操作,例如通过固定(例如使用福尔马林)和/或嵌入蜡中(例如福尔马林固定的石蜡包埋(FFPE)组织样品)。应当理解,可以使用从受试者获得组织的任何方法,并且所使用的方法的选择将取决于各种因素,例如组织的类型、受试者的年龄或从业者可用的程序。获得这样的样品的标准技术可在本领域中获得。参见,例如Schluger等,J.Exp.Med.176:1327-33(1992);Bigby等,Am.Rev.Respir.Dis.133:515-18(1986);Kovacs等,NEJM 318:589-93(1988);和Ognibene等,Am.Rev.Respir.Dis.129:929-32(1984)。

[0386] 在其它实施方式中,样品可以是环境样品,例如水、土壤、或作为工业或医学表面这样的表面。在一些实施方式中,例如美国专利公开号2013/0190196中公开的方法可以用于高度灵敏度和特异性地直接从粗细胞样品中检测核酸标签,特别是RNA水平。通过BLAST软件比较来自感兴趣病原体的编码序列与其他生物体中的所有编码序列可以鉴定或选择对各感兴趣病原体特异的序列。

[0387] 本公开内容的多个实施方式涉及使用本领域中公知的程序和方法以成功分级临床血样。参见,例如在Han Wei Hour等,Microfluidic Devices for Blood Fractionation,Micromachines 2011,2,319-343;Ali Asgar S.Bhagat等Dean Flow Fractionation(DFF) Isolation of Circulating Tumor Cells(CTCs) from Blood,15th International Conference on Miniaturized Systems for Chemistry and Life Sciences,October 2-6,2011,Seattle,WA;和国际专利公开号W02011109762中公开的程序,其全部公开内容通过引用并入本文。血液样品通常在培养中扩增以增加样品大小以用于测试目的。在本发明的一些实施方式中,血液或其他生物样品可以用于本文所述的方法中,而不需要在培养中扩增。

[0388] 此外,本公开内容的多个实施方式涉及使用本领域已知的程序和方法使用螺旋微通道成功地从全血分离病原体,如在Han Wei Hour等,Pathogen Isolation from Whole Blood Using Spiral Microchannel,Case No.15995JR,Massachusetts Institute of Technology中所述。文稿在准备中,其全部公开内容通过引用并入本文。

[0389] 由于本文公开的实施方式的增加的灵敏度,在某些示例性实施方式中,测定和方法可以在粗制样品或其中待检测的靶分子不从所述样品进一步分级或纯化的样品中进行。

[0390] 示例性微生物

[0391] 本文所公开的实施方式可用于检测许多不同的微生物。本文所用的术语微生物包括细菌、真菌、原生动物、寄生虫和病毒。

[0392] 细菌

[0393] 以下提供可能使用本文所公开的实施方式进行检测的类型的微生物的实例列表。在某些示例性实施方式中,微生物是细菌。可根据所公开的方法检测的细菌的实例包括但不限于以下的一种或多种(或其组合):*Acinetobacter baumannii*、放线杆菌属(*Actinobacillus*)种、放线菌纲(*Actinomycetes*)种、放线菌属(*Actinomyces*)种(例如*Actinomyces israelii*和*Actinomyces naeslundii*)、气单胞菌属(*Aeromonas*)种(例如*Aeromonas hydrophila*、*Aeromonas veronii biovar sobria*(*Aeromonas sobria*)和

Aeromonas caviae)、*Anaplasma phagocytophilum*、*Anaplasma marginale* *Alcaligenes xylosoxidans*、*Acinetobacter baumannii*、*Actinobacillus actinomycetemcomitans*、芽孢杆菌属 (*Bacillus*) 种 (例如 *Bacillus anthracis*、*Bacillus cereus*、*Bacillus subtilis*、*Bacillus thuringiensis* 和 *Bacillus stearothermophilus*)、拟杆菌属 (*Bacteriodes*) 种 (例如 *Bacteroides fragilis*)、巴尔通氏体属 (*Bartonella*) 种 (例如 *Bartonella bacilliformis* 和 *Bartonella henselae*)、双歧杆菌属 (*Bifidobacterium*) 种、博代氏杆菌属 (*Bordetella*) 种 (例如 *Bordetella pertussis*、*Bordetella parapertussis* 和 *Bordetella bronchiseptica*)、包柔氏螺旋体属 (*Borrelia*) 种 (例如 *Borrelia recurrentis* 和 *Borrelia burgdorferi*)、布鲁氏菌属 (*Brucella*) 种 (例如 *Brucella abortus*、*Brucella canis*、*Brucella melintensis* 和 *Brucella suis*)、伯克霍尔德氏菌属 (*Burkholderia*) 种 (例如 *Burkholderia pseudomallei* 和 *Burkholderia cepacia*)、弯曲杆菌属 (*Campylobacter*) (例如 *Campylobacter jejuni*、*Campylobacter coli*、*Campylobacter lari* 和 *Campylobacter fetus*)、嗜二氧化碳噬细胞菌属 (*Capnocytophaga*) 种、*Cardiobacterium hominis*、*Chlamydia trachomatis*、*Chlamydophila pneumoniae*、*Chlamydophila psittaci*、*Citrobacter* sp. *Coxiella burnetii*、棒状杆菌属 (*Corynebacterium*) 种 (例如 *Corynebacterium diphtheriae*、*Corynebacterium jeikeum* 和 *Corynebacterium*)、梭菌属 (*Clostridium*) 种 (例如 *Clostridium perfringens*、*Clostridium difficile*、*Clostridium botulinum* 和 *Clostridium tetani*)、*Eikenella corrodens*、肠杆菌属 (*Enterobacter*) 种 (例如 *Enterobacter aerogenes*、*Enterobacter agglomerans*、*Enterobacter cloacae* 和大肠杆菌, 包括机会性大肠杆菌, 例如 enterotoxigenic *E. coli*、enteroinvasive *E. coli*、enteropathogenic *E. coli*、enterohemorrhagic *E. coli*、enteroaggregative *E. coli* 和 uropathogenic *E. coli*)、肠球菌属 (*Enterococcus*) 种 (例如 *Enterococcus faecalis* 和 *Enterococcus faecium*) 埃立克体属 (*Ehrlichia*) 种 (例如 *Ehrlichia chafeensis* 和 *Ehrlichia canis*)、*Epidermophyton floccosum*、*Erysipelothrix rhusiopathiae*、真细菌属 (*Eubacterium*) 种、*Francisella tularensis*、*Fusobacterium nucleatum*、*Gardnerella vaginalis*、*Gemella morbillorum*、嗜血杆菌属 (*Haemophilus*) 种 (例如 *Haemophilus influenzae*、*Haemophilus ducreyi*、*Haemophilus aegyptius*、*Haemophilus parainfluenzae*、*Haemophilus haemolyticus* 和 *Haemophilus parahaemolyticus*)、螺杆菌属 (*Helicobacter*) 种 (例如 *Helicobacter pylori*、*Helicobacter cinaedi* 和 *Helicobacter fennelliae*)、*Kingella kingii*、克雷白氏杆菌属 (*Klebsiella*) 种 (例如 *Klebsiella pneumoniae*、*Klebsiella granulomatis* 和 *Klebsiella oxytoca*)、乳酸菌属 (*Lactobacillus*) 种、*Listeria monocytogenes*、*Leptospira interrogans*、*Legionella pneumophila*、*Leptospira interrogans*、消化链球菌属 (*Peptostreptococcus*) 种、*Mannheimia hemolytica*、*Microsporum canis*、*Moraxella catarrhalis*、摩根氏菌属 (*Morganella*) 种、动弯杆菌属 (*Mobiluncus*) 种、微球菌属 (*Micrococcus*) 种、分支杆菌属 (*Mycobacterium*) 种 (例如 *Mycobacterium leprae*、*Mycobacterium tuberculosis*、*Mycobacterium paratuberculosis*、*Mycobacterium intracellulare*、*Mycobacterium avium*、*Mycobacterium bovis*、和 *Mycobacterium marinum*)、支原体属 (*Mycoplasma*) 种 (例如 *Mycoplasma pneumoniae*、*Mycoplasma hominis*、和 *Mycoplasma genitalium*)、诺卡氏菌属

(*Nocardia*) 种(例如 *Nocardia asteroides*、*Nocardia cyriacigeorgica* 和 *Nocardia brasiliensis*)、奈瑟氏菌属 (*Neisseria*) 种(例如 *Neisseria gonorrhoeae* 和 *Neisseria meningitidis*)、*Pasteurella multocida*、*Pityrosporum orbiculare* (*Malassezia furfur*)、*Plesiomonas shigelloides*、*Prevotella* sp.、*Porphyromonas* sp.、*Prevotella melaninogenica*、变形杆菌属 (*Proteus*) 种(例如 *Proteus vulgaris* 和 *Proteus mirabilis*)、普罗维登斯菌属 (*Providencia*) 种(例如 *Providencia alcalifaciens*、*Providencia rettgeri* 和 *Providencia stuartii*)、铜绿假单胞菌、*Propionibacterium acnes*、*Rhodococcus equi*、*Rickettsia* sp. (例如 *Rickettsia rickettsii*、*Rickettsia akari* 和 *Rickettsia prowazekii*、*Orientia tsutsugamushi* (之前: *Rickettsia tsutsugamushi*) 和 *Rickettsia typhi*)、红球菌属 (*Rhodococcus*) 种、*Serratia marcescens*、*Stenotrophomonas maltophilia*、沙门氏菌属 (*Salmonella*) 种(例如 *Salmonella enterica*、*Salmonella typhi*、*Salmonella paratyphi*、*Salmonella enteritidis*、*Salmonella choleraesuis* 和 *Salmonella typhimurium*)、沙雷氏菌属 (*Serratia*) 种(例如 *Serratia marcescens* 和 *Serratia liquifaciens*)、志贺氏菌属 (*Shigella*) 种(例如 *Shigella dysenteriae*、*Shigella flexneri*、*Shigella boydii* 和 *Shigella sonnei*)、葡萄球菌属 (*Staphylococcus*) 种(例如 *Staphylococcus aureus*、*Staphylococcus epidermidis*、*Staphylococcus hemolyticus*、*Staphylococcus saprophyticus*)、链球菌属 (*Streptococcus*) 种(例如肺炎链球菌 (*Streptococcus pneumoniae*) (例如氯霉素抗性血清型4肺炎链球菌、奇放线菌素抗性血清型6B肺炎链球菌、链霉素抗性血清型9V肺炎链球菌、红霉素抗性血清型14肺炎链球菌、奥普托欣抗性血清型14肺炎链球菌、利福平抗性血清型18C肺炎链球菌、四环素抗性血清型19F肺炎链球菌、青霉素抗性血清型19F肺炎链球菌、和甲氧苄氨嘧啶抗性血清型23F肺炎链球菌、氯霉素抗性血清型4肺炎链球菌、奇放线菌素抗性血清型6B肺炎链球菌、链霉素抗性血清型9V肺炎链球菌、奥普托欣抗性血清型14肺炎链球菌、利福平抗性血清型18C肺炎链球菌、青霉素抗性血清型19F肺炎链球菌、或甲氧苄氨嘧啶抗性血清型23F肺炎链球菌)、*Streptococcus agalactiae*、*Streptococcus mutans*、*Streptococcus pyogenes*、A组链球菌、化脓性链球菌、组B链球菌、*Streptococcus agalactiae*、组C链球菌、*Streptococcus anginosus*、*Streptococcus equismilis*、组D链球菌、*Streptococcus bovis*、组F链球菌、和 *Streptococcus anginosus* 组G链球菌)、*Spirillum minus*、*Streptobacillus moniliformis*、密螺旋体属 (*Treponema*) 种(例如 *Treponema carateum*、*Treponema petenue*、*Treponema pallidum* 和 *Treponema endemicum*、*Trichophyton rubrum*、*T. mentagrophytes*、*Tropheryma whippellii*、*Ureaplasma urealyticum*、韦永氏球菌属 (*Veillonella*) 种、弧菌属 (*Vibrio*) 种(例如 *Vibrio cholerae*、*Vibrio parahemolyticus*、*Vibrio vulnificus*、*Vibrio parahaemolyticus*、*Vibrio vulnificus*、*Vibrio alginolyticus*、*Vibrio mimicus*、*Vibrio hollisae*、*Vibrio fluvialis*、*Vibrio metchnikovii*、*Vibrio damsela* 和 *Vibrio furnisii*)、耶尔森氏鼠疫杆菌属 (*Yersinia*) 种(例如 *Yersinia enterocolitica*、*Yersinia pestis*、和 *Yersinia pseudotuberculosis*) 和 *Xanthomonas maltophilia* 等。

[0394] 真菌

[0395] 在某些示例性实施方式中,微生物是真菌或真菌物种。可以根据所公开的方法检

测真菌的实例包括但不限于以下任何一种或多种(或任意组合):曲霉属(*Aspergillus*)、芽生菌属(*Blastomyces*)、念珠菌病(*Candidiasis*)、球孢子菌病(*Coccidioidomycosis*)、新型隐球菌(*Cryptococcus neoformans*)、加蒂隐球菌(*Cryptococcus gatti*)种、组织胞浆菌属(*Histoplasma*)种(例如荚膜组织胞浆菌(*Histoplasma capsulatum*))、肺孢子虫属(*Pneumocystis*)种(例如耶氏肺孢子菌(*Pneumocystis jirovecii*))、葡萄穗霉属(例如纸葡萄穗霉(*Stachybotrys chartarum*))、毛霉菌病(*Mucroymcosis*)、孢子丝菌属(*Sporothrix*)、真菌性眼睛感染癣(*fungal eye infections ringworm*)、突脐蠕孢属(*Exserohilum*)、枝孢属(*Cladosporium*)、在某些示例性实施方式中,真菌是酵母。可以根据所公开的方法检测的酵母的实例包括但不限于以下中的一种或多种(或任何组合):曲霉属种(例如*Aspergillus fumigatus*,*Aspergillus flavus*和*Aspergillus clavatus*)、隐球菌属种(例如新型隐球菌、加蒂隐球菌、*Cryptococcus laurentii*和*Cryptococcus albidus*)、地丝菌属种、酵母属种、汉逊酵母属种、念珠菌属种(例如*Candida albicans*)、克鲁维酵母属种、德巴利氏酵母属种、毕赤酵母属种。在某些示例性实施方式中,真菌是霉菌。示例性霉菌包括但不限于青霉属种、枝孢属种、丝衣酶属种或其组合。

[0396] 原生动物

[0397] 在某些示例性实施方式中,微生物是原生动物。可根据所公开的方法和装置检测的原生动物的实例包括但不限于以下中的任何一种或多种(或任何组合):眼虫动物界(*Euglenozoa*)、异叶足纲(*Heterolobosea*)、双滴虫目(*Diplomonadida*)、变形虫界(*Amoebozoa*)、芽囊原虫属(*Blastocystic*)、顶复亚门(*Apicomplexa*)。示例性眼虫动物界包括但不限于克氏锥虫(*Trypanosoma cruzi*) (查加斯病(*Chagas disease*))、布氏冈比亚锥虫(*T.brucei gambiense*)、布氏罗德西亚锥虫(*T.brucei rhodesiense*)、巴西利什曼原虫(*Leishmania braziliensis*)、*L.infantum*、*L.mexicana*、*L.major*、*L.tropica*、*L.donovani*。示例性异叶足纲包括但不限于*Naegleria fowleri*。示例性双滴虫目包括但不限于*Giardia intestinalis* (*G.lamblia*、*G.duodenalis*)。示例性变形虫界包括但不限于*canthamoeba castellanii*、*Balamuthia madriilaris*、*Entamoeba histolytica*。示例性芽囊原虫属包括但不限于*Blastocystic hominis*。示例性顶复亚门包括但不限于*Babesia microti*、*Cryptosporidium parvum*、*Cyclospora cayetanensis*、*Plasmodium falciparum*、*P.vivax*、*P.ovale*、*P.malariae*和*Toxoplasma gondii*。

[0398] 寄生虫

[0399] 在某些示例性实施方式中,微生物是寄生虫。可以根据公开的方法检测的寄生虫的实例包括但不限于,盘尾属(*Onchocerca*)种和疟原虫(*Plasmodium*)种中的一种或多种(或任何组合)。

[0400] 病毒

[0401] 在某些示例性实施方式中,本文所公开的系统、装置和方法涉及检测样品中的病毒。本文所公开的实施方式可用于检测病毒感染(例如受试者或植物)、或测定病毒株,包括因单核苷酸多态性而不同的病毒株。病毒可以是DNA病毒、RNA病毒或逆转录病毒。可用于本发明的病毒的非限制性实例包括但不限于埃博拉、麻疹、SARS、基孔肯雅热、肝炎、马尔堡、黄热病、MERS、登革热、拉沙、流感、弹状病毒或HIV。肝炎病毒可包括甲型肝炎、乙型肝炎或丙型肝炎。流感病毒可包括例如甲型流感或乙型流感。HIV可能包括HIV 1或HIV 2。在某些

示例性实施方式中,病毒序列可以是人呼吸道合胞病毒、苏丹埃博拉病毒、Bundibugyo病毒、Tai Forest埃博拉病毒、Reston埃博拉病毒、Achimota、Aedes flavivirus、Aguacate病毒、Akabane病毒、Alethinophid reptarenavirus、Allpahuayo mammarenavirus、Amapari mammarenavirus、Andes病毒、Apoi病毒、Aravan病毒、Aroa病毒、Arumwot病毒、大西洋鲑副粘病毒、澳大利亚蝙蝠狂犬病病毒、Avian bornavirus、Avian metapneumovirus、Avian paramyxoviruses、企鹅或福克兰群岛病毒、BK多瘤病毒、Bagaza病毒、Banna病毒、蝙蝠疱疹病毒、Bat sapovirus、Bear Canon mammarenavirus、Beilong病毒、Betacoronavirus、Betapapillomavirus 1-6、Bhanja病毒、Bokeloh bat lyssavirus、波尔钠病病毒、Bourbon病毒、Bovine hepacivirus、Bovine parainfluenza病毒3、牛呼吸道合胞体病毒、Brazoran病毒、Bunyamwera病毒、Caliciviridae病毒、California encephalitis病毒、Candiru病毒、Canine distemper病毒、Canine pneumovirus、Cedar病毒、细胞融合剂病毒、Cetacean morbillivirus、Chandipura病毒、Chaoyang病毒、Chapare mammarenavirus、Chikungunya病毒、Colobus monkey papillomavirus、Colorado tick fever病毒、Cowpox病毒、Crimean-Congo hemorrhagic fever病毒、Culex flavivirus、Cupixi mammarenavirus、登革热病毒、Dobrava-Belgrade病毒、Donggang病毒、Dugbe病毒、Duvenhage病毒、Eastern equine encephalitis病毒、Entebbe bat病毒、Enterovirus A-D、European bat lyssavirus 1-2、Eyach病毒、Feline morbillivirus、Fer-de-Lance paramyxovirus、Fitzroy River病毒、Flaviviridae病毒、Flexal mammarenavirus、GB病毒C、Gairo病毒、Gemycircularvirus、Goose paramyxovirus SF02、Great Island病毒、Guanarito mammarenavirus、Hantaan病毒、Hantavirus Z10、Heartland病毒、Hendra病毒、肝炎A/B/C/E、Hepatitis delta病毒、人类博卡病毒、人类冠状病毒、人内源性逆转录病毒K、人类肠道冠状病毒、人类生殖器相关环状DNA病毒-1、人类疱疹1-8、人类免疫缺陷病毒1/2、Human mastadenovirus A-G、人乳头瘤病毒、人类副流感病毒1-4、Human paraechovirus、Human picornavirus、Human smacovirus、Ikoma lyssavirus、Ilheus病毒、流感A-C、Ippy mammarenavirus、Irkut病毒、J-病毒、JC polyomavirus、Japanese encephalitis病毒、Junin mammarenavirus、KI polyomavirus、Kadipiro病毒、Kamiti River病毒、Kedougou病毒、Khujand病毒、Kokobera病毒、Kysanur forest disease病毒、Lagos bat病毒、Langat病毒、Lassa mammarenavirus、Latino mammarenavirus、Leopards Hill病毒、Liao ning病毒、Ljungan病毒、Lloviu病毒、Louping ill病毒、Lujo mammarenavirus、Luna mammarenavirus、Lunk病毒、Lymphocytic choriomeningitis mammarenavirus、Lyssavirus Ozernoe、MSSI2\、.225病毒、Machupo mammarenavirus、Mamastrovirus 1、Manzanilla病毒、Mapuera病毒、Marburg病毒、Mayaro病毒、Measles病毒、Menangle病毒、Mercadeo病毒、Merkel cell polyomavirus、中东呼吸综合征冠状病毒、Mobala mammarenavirus、Modoc病毒、Moijang病毒、Mokolo病毒、Monkeypox病毒、Montana myotis leukoenthalitis病毒、Mopeia lassa virus reassortant 29、Mopeia mammarenavirus、Morogoro病毒、Mossman病毒、Mumps病毒、Murine pneumonia病毒、Murray Valley encephalitis病毒、Nariva病毒、Newcastle disease病毒、Nipah病毒、Norwalk病毒、Norway rat hepacivirus、Ntaya病毒、O'nyong-nyong病毒、Oliveros mammarenavirus、Omsk hemorrhagic fever病毒、Oropouche病毒、Parainfluenza病毒5、Parana

mammarenavirus、Parramatta River病毒、Peste-des-petits-ruminants病毒、Pichande mammarenavirus、Picornaviridae病毒、Pirital mammarenavirus、Piscihepevirus A、猪副流感病毒1、porcine rubulavirus、Powassan病毒、灵长类动物的T细胞白血病毒1-2、Primate erythroparvovirus 1、Punta Toro病毒、Puumala病毒、Quang Binh病毒、Rabies病毒、Razdan病毒、Reptile bornavirus 1、Rhinovirus A-B、Rift Valley fever病毒、Rinderpest病毒、Rio Bravo病毒、Rodent Torque Teno病毒、Rodent hepacivirus、Ross River病毒、Rotavirus A-I、Royal Farm病毒、Rubella病毒、Sabia mammarenavirus、Salem病毒、Sandfly fever Naples病毒、Sandfly fever Sicilian病毒、Sapporo病毒、Sathuperi病毒、Seal anellovirus、Semliki Forest病毒、Sendai病毒、Seoul病毒、Sepik病毒、严重急性呼吸综合症相关冠状病毒、严重发热伴血小板减少综合症病毒、Shamonda病毒、Shimoni bat病毒、Shuni病毒、Simbu病毒、Simian torque teno病毒、Simian病毒40-41、Sin Nombre病毒、Sindbis病毒、Small anellovirus、Sosuga病毒、Spanish goat encephalitis病毒、Spondweni病毒、St.Louis encephalitis病毒、Sunshine病毒、TTV-like mini病毒、Tacaribe mammarenavirus、Taila病毒、Tamana bat病毒、Tamiami mammarenavirus、Tembusu病毒、Thogoto病毒、Thottapalayam病毒、Tick-borne encephalitis病毒、Tioman病毒、Togaviridae病毒、Torque teno canis病毒、Torque teno douroucouli病毒、Torque teno felis病毒、Torque teno midi病毒、Torque teno sus病毒、Torque teno tamarin病毒、Torque teno病毒、Torque teno zalophus病毒、Tuhoko病毒、Tula病毒、Tupaia paramyxovirus、Usutu病毒、Uukuniemi病毒、Vaccinia病毒、Variola病毒、Venezuelan equine encephalitis病毒、Vesicular stomatitis Indiana病毒、WU Polyomavirus、Wesselsbron病毒、West Caucasian bat病毒、West Nile病毒、Western equine encephalitis病毒、Whitewater Arroyo mammarenavirus、黄热病病毒、Yokose病毒、Yug Bogdanovac病毒、Zaire ebolavirus、寨卡病毒、或Zygosaccharomyces bailii virus Z viral sequence。可以检测的RNA病毒的实例包括以下中的一种或多种(或任何组合):冠状病毒科(Coronaviridae)病毒、小核糖核酸科(Picornaviridae)病毒、杯状病毒科(Caliciviridae)病毒、黄病毒科(Flaviviridae)病毒、披膜病毒科(Togaviridae)病毒、Bornaviridae、丝状病毒科(Filoviridae)、副黏液病毒科(Paramyxoviridae)、肺病毒科(Pneumoviridae)、弹状病毒科(Rhabdoviridae)、沙粒病毒科(Arenaviridae)、布尼亚病毒科(Bunyaviridae)、正粘病毒科(Orthomyxoviridae)或Deltavirus。在某些实施方式中,病毒是冠状病毒(Coronavirus)、SARS、脊髓灰质炎病毒(Poliovirus)、鼻病毒(Rhinovirus)、甲型肝炎(Hepatitis A)、诺沃克病毒(Norwalk virus)、黄热病病毒(Yellow fever virus)、西尼罗病毒(West Nile virus)、丙型肝炎病毒(Hepatitis C virus)、登革热病毒(Dengue fever virus)、寨卡病毒(Zika virus)、风疹病毒(Rubella virus)、罗斯河病毒(Ross River virus)、辛德毕斯病毒(Sindbis virus)、基孔肯亚病毒(Chikungunya virus)、博尔纳病病毒(Borna disease virus)、埃博拉病毒(Ebola virus)、马尔堡病毒(Marburg virus)、麻疹病毒(Measles virus)、腮腺炎病毒(Mumps virus)、尼帕病毒(Nipah virus)、亨德拉病毒(Hendra virus)、新城疫病毒(Newcastle disease virus)、人呼吸道合胞病毒(Human respiratory syncytial virus)、狂犬病病毒(Rabies virus)、拉沙病毒(Lassa virus)、汉坦病毒(Hantavirus)、克里米亚-刚果出血热病毒(Crimean-

Congo hemorrhagic fever virus)、流感(Influenza)或丁型肝炎病毒(Hepatitis D virus)

[0402] 在某些示例性实施方式中,病毒可以是选自以下的植物病毒:烟草花叶病毒(TMV)、番茄斑萎病毒(TSWV)、黄瓜花叶病毒(CMV)、马铃薯Y病毒(PVY)、RT病毒花椰菜花叶病毒(CaMV)、梅花痘病毒(PPV)、雀麦花叶病毒(BMV)、马铃薯病毒X(PVX)、柑橘衰退病毒(CTV)、大麦黄矮病毒(BYDV)、马铃薯卷叶病毒(PLRV)、番茄丛生特技病毒(TBSV)、水稻球茎病毒(RTSV)、水稻黄色斑驳病毒(RYMV)、水稻灰白色病毒(RHBV)、玉米雷亚朵菲纳病毒(MRFV)、玉米矮花叶病毒(MDMV)、甘蔗花叶病毒(SCMV)、甘薯羽毛斑驳病毒(SPFMV)、甘薯沉降脉状线虫病毒(PSV)、葡萄扇叶病毒(GFLV)、葡萄病毒A(GVA)、葡萄病毒B(GVB)、葡萄斑点病毒(GFkV)、葡萄卷叶病毒相关病毒-1、-2和-3、(GLRaV-1、-2和-3)、南芥菜花叶病毒(ArMV)、或落叶松茎麻点相关病毒(RSPaV)。在优选的实施方式中,靶RNA分子是所述病原体的一部分或从所述病原体的DNA分子转录。例如,靶序列可以包含在RNA病毒的基因组中。如果所述病原体感染或已经感染所述植物,则进一步优选CRISPR效应蛋白在所述植物中水解所述病原体的所述靶RNA分子。因此,当CRISPR系统(或其完成所需的部分)在治疗上应用时,优选CRISPR系统能够从植物病原体切割靶RNA分子,即在感染发生或预防后,即在感染发生之前。

[0403] 在某些示例性实施方式中,病毒可以是逆转录病毒。可使用本文所公开的实施方式检测的逆转录病毒的实例包括以下属的病毒中的一种或多种或任何组合:反转录病毒属(Alpharetrovirus)、乙型反录病毒属(Betaretrovirus)、丙型反录病毒属(Gammaretrovirus)、逆转录病毒属(Deltaretrovirus)、Epsilonretrovirus、慢病毒(Lentivirus)、泡沫病毒属(Spumavirus)或Metaviridae、假病毒科(Pseudoviridae)和逆转录病毒科(Retroviridae)(包括HIV)、肝脱氧核糖核酸病毒科(Hepadnaviridae)(包括乙型肝炎病毒)和花椰菜花叶病毒科(Caulimoviridae)(包括花椰菜花叶病毒)。

[0404] 在某些示例性实施方式中,病毒是DNA病毒。可使用本文所公开的实施方式检测的实例DNA病毒包括以下中的一种或多种病毒(或任何组合):Family Myoviridae、Podoviridae、Siphoviridae、Alloherpesviridae、疱疹病毒科(包括人疱疹病毒和Varicella Zoster virus)、Malcoherpesviridae、脂毛噬菌体科、Rudiviridae、腺病毒、Ampullaviridae、囊泡病毒科、非洲猪瘟病毒科(包括非洲猪瘟病毒)、杆状病毒、Cicaudaviridae、Clavaviridae、Corticoviridae、微小纺锤形噬菌体科、Globuloviridae、滴状病毒科、Hytrosaviridae、虹彩病毒科、Maseilleviridae、Mimiviridae、Nudiviridae、Nimaviridae、Pandoraviridae、乳头瘤病毒、藻类去氧核糖核酸病毒科、Plasmaviridae、Polydnaviruses、Polyomaviridae(包括猿猴病毒40、JC病毒、BK病毒)、痘病毒科(包括牛痘和天花)、Sphaerolipoviridae、Tectiviridae、Turriviridae、Dinodnavirus、Salterprovirus、Rhizidovirus等。在一些实施方式中,诊断怀疑患有细菌感染的受试者中的物种特异性细菌感染的方法被描述为获得包含来自受试者的细菌核糖体核糖核酸的样品;使样品与一种或多种所述探针接触,并检测样品中存在的细菌核糖体核糖核酸序列与探针之间的杂交,其中杂交检测表明受试者感染大肠杆菌、肺炎克雷伯氏菌、铜绿假单胞菌、金黄色葡萄球菌、鲍氏不动杆菌、白色念珠菌、阴沟肠杆菌、粪肠球菌、尿肠球菌、奇异变形杆菌、无乳链球菌、或嗜麦芽糖葡萄球菌或其组合。

[0405] 疟疾检测和监测

[0406] 疟疾是由疟原虫寄生虫引起的蚊子传播病理学。寄生虫通过被感染的雌性按蚊蚊子的叮咬传播给人们。五个疟原虫物种引起人类疟疾：恶性疟原虫、间日疟原虫、卵形疟原虫、三日疟原虫和诺氏疟原虫。其中，根据世界卫生组织 (WHO)，恶性疟原虫和间日疟原虫是造成最大威胁的原因。恶性疟原虫是非洲大陆上最普遍的疟疾寄生虫，是全球大多数疟疾相关死亡的原因。间日疟原虫是撒哈拉以南非洲以外大多数国家的主要疟疾寄生虫。

[0407] 2015年，91个国家和地区持续传播疟疾。根据世卫组织最新估计，2015年有2.12亿疟疾病例，429000例死亡。在疟疾传播率高的地区，5岁以下儿童特别容易受到感染、疾病和死亡；超过三分之二 (70%) 的疟疾死亡发生在这个年龄组。2010年至2015年，全球5岁以下疟疾死亡率下降了29%。然而，疟疾仍然是五岁以下儿童的主要杀手，每两分钟就会夺走一个孩子的生命。

[0408] 如WHO所述，疟疾是一种急性发热性疾病。在非免疫个体中，在感染性蚊虫叮咬后7天或更长时间出现症状。最初的症状——发烧、头痛、发冷和呕吐——可能是轻微的，难以识别为疟疾，但如果不在24小时内治疗，恶性疟原虫疟疾可能会发展成严重疾病，往往会导致死亡。

[0409] 重症疟疾的儿童经常发展以下一种或多种症状：严重贫血、代谢性酸中毒引起的呼吸窘迫、或脑型疟。在成人中，多器官受累也很常见。在疟疾流行地区，人们可能会产生部分免疫力，从而导致无症状感染。

[0410] 快速和有效的诊断测试的发展与公共健康高度相关。事实上，疟疾的早期诊断和治疗不仅可以减少疾病并预防死亡，还有助于减少疟疾的传播。根据世界卫生组织的建议，所有疑似疟疾病例都应在接受治疗前使用基于寄生虫的诊断检测 (特别是使用快速诊断检测) 进行确认 (参见“世界卫生组织疟疾治疗指南”，第三版，2015年4月出版)。

[0411] 抗疟疗法的抗性成为关键的健康问题，其极大地减少治疗策略。实际上，正如世界卫生组织网站上所报道的那样，恶性疟原虫对前几代药物例如氯喹和磺胺多辛/乙胺嘧啶 (SP) 的抗性在20世纪50年代和60年代变得普遍，破坏了疟疾控制工作并扭转了儿童生存的增长。因此，世界卫生组织建议对抗疟疾药物的抗性进行常规监测。实际上，准确的诊断可以避免不适当的治疗并限制对抗疟疾药物的抗性的延长。

[0412] 在此背景下，世界卫生大会于2015年5月通过的“2016-2030年世界卫生组织疟疾技术战略”为所有疟疾流行国家提供了技术框架。它旨在指导和支持区域和国家计划，因为它们致力于控制和消除疟疾。该战略制定了雄心勃勃但可实现的全球目标，包括：

[0413] • 到2030年将疟疾病例发病率降低至少90%。

[0414] • 到2030年将疟疾死亡率降低至少90%。

[0415] • 到2030年消除至少35个国家的疟疾。

[0416] • 防止所有无疟疾国家的疟疾再次出现。

[0417] 该战略是经过2年的广泛协商进程的结果，涉及来自70个成员国的400多名技术专家的参与。它基于3个关键轴：

[0418] • 确保普遍获得疟疾预防、诊断和治疗；

[0419] • 加快努力消除和实现无疟疾地位；和

[0420] • 将疟疾监测转变为核心干预措施。

[0421] 对疟原虫的治疗包括芳基氨基醇例如奎宁或奎宁衍生物例如氯喹、阿莫地喹、甲氟喹、哌喹、本苄醇、伯氨喹；亲脂性羟基萘醌类似物，例如阿托伐醌；抗叶酸药物，例如磺胺类药物磺胺多辛、氨苯砜和乙胺嘧啶；氯胍；阿托伐醌/氯胍的组合；青蒿素 (artemisinins) 药物；及其组合。

[0422] 用于诊断由蚊子传播病原体的存在的靶序列包括诊断疟原虫，尤其是影响人类的疟原虫物种例如恶性疟原虫、间日疟原虫、卵形疟原虫、三日疟原虫和诺氏疟原虫存在的序列，包括来自其基因组的序列。

[0423] 靶序列用于诊断监测对治疗疟原虫，尤其是影响人类的疟原虫物种例如恶性疟原虫，间日疟原虫，卵形疟原虫，三日疟原虫，和诺氏疟原虫的药物抗性。

[0424] 另外的靶序列包括序列，所述序列包含编码参与疟原虫寄生虫的基本生物学过程的蛋白质的靶分子/核酸分子，特别是转运蛋白，例如来自药物/代谢物转运蛋白家族的蛋白质，参与底物易位的ATP结合盒 (ABC) 蛋白，例如ABC转运蛋白C亚家族或Na⁺/H⁺交换蛋白、膜谷胱甘肽S-转移酶；参与叶酸途径的蛋白，例如二氢蝶酸合成酶、二氢叶酸还原酶活性或二氢叶酸还原酶-胸苷酸合成酶；和参与质子跨越线粒体内膜的蛋白，特别是细胞色素b复合物的转运。另外的靶标还可以包括编码血红素聚合酶的基因。

[0425] 另外的靶序列包括编码参与基本生物学过程的蛋白的靶分子/核酸分子，其可选自恶性疟原虫氯喹抗性转运蛋白基因 (pf crt)、恶性疟原虫多药物抗性转运蛋白1 (pfmdr1)、恶性疟原虫多药物抗性相关蛋白基因 (Pfmrp)、恶性疟原虫的Na⁺/H⁺交换蛋白基因 (pf nhe)、编码恶性疟原虫输出蛋白1的基因、恶性疟原虫的Ca²⁺转运ATP酶6 (pfatp6)；恶性疟原虫二氢蝶合酶 (pf dhps)、二氢叶酸还原酶活性 (pf dhpr) 和二氢叶酸还原酶胸苷酸合酶 (pf dhfr) 基因、细胞色素b基因、GTP环化水解酶和Kelch13 (K13) 基因，以及它们在其他疟原虫物种中的功能性同源基因。

[0426] 许多突变，特别是单点突变，已经在它们的当前治疗靶标和特定抗性表型相关的蛋白中鉴定。因此，本发明允许检测蚊子传播的寄生虫例如疟原虫的各种抗性表型。

[0427] 本发明允许检测靶核酸/分子的一个或多个突变，并且特别是一个或多个单核苷酸多态性。因此，下面的任何一个突变或它们的组合可以用作药物抗性标志物，并且可以根据本发明进行检测。

[0428] 恶性疟原虫K13中的单点突变包括在以下位置的单点突变：位置252、441、446、449、458、493、539、543、553、561、568、574、578、580、675、476、469、481、522、537、538、579、584和719，特别是突变E252Q、P441L、F446I、G449A、N458Y、Y493H、R539T、I543T、P553L、R561H、V568G、P574L、A578S、C580Y、A675V、M476I；C469Y；A481V；S522C；N537I；N537D；G538V；M579I；D584V；和H719N。这些突变通常与青蒿素药物抗性表型相关 (Artemisinin and artemisinin-based combination therapy resistance, April 2016WHO/HTM/GMP/2016.5)。

[0429] 在恶性疟原虫二氢叶酸还原酶 (DHFR) (PFDHFR-TS, PFD0830w) 中，重要的多态性包括位置108、51、59和164的突变，特别是调节对乙胺嘧啶的抗性的108D、164L、51I和59R。其他多态性还包括与对磺胺多辛的抗性相关的437G、581G、540E、436A和613S。另外观察到的突变包括Ser108Asn、Asn51Ile、Cys59Arg、Ile164Leu、Cys50Arg、Ile164Leu、Asn188Lys、Ser189Arg和Val213Ala、Ser108Thr和Ala16Val。突变Ser108Asn、Asn51Ile、Cys59Arg、

Ile164Leu、Cys50Arg、Ile164Leu特别与基于乙胺嘧啶的疗法和/或氯鸟嘌呤-氨苯砒联合疗法抗性相关。环胍抗性似乎与双突变Ser108Thr和Ala16Val相关。dhfr的扩增也可能与治疗抗性特别是乙胺嘧啶抗性高度相关。

[0430] 在恶性疟原虫二氢蝶合酶 (DHPS) (PfdHPS, PF08_0095) 中, 重要的多态性包括在位置436、437、581和613的Ser436Ala/Phe、Ala437Gly、Lys540Glu、Ala581Gly和Ala613Thr/Ser的突变。位置581和/或613的多态性也与对磺胺多辛-乙胺嘧啶碱疗法的抗性相关。

[0431] 在恶性疟原虫氯喹抗性转运蛋白 (PfCRT), 位置76的多态性, 特别是突变Lys76Thr, 与对氯喹抗性相关。进一步的多态性包括可能与氯喹抗性相关的Cys72Ser、Met74Ile、Asn75Glu、Ala220Ser、Gln271Glu、Asn326Ser、Ile356Thr和Arg371Ile。PfCRT也在残基S33、S411和T416处被磷酸化, 其可以调节蛋白质的转运活性或特异性。

[0432] 在恶性疟原虫多药物抗性转运蛋白1 (PfMDR1) (PFE1150w) 中, 在位置86、184、1034、1042的多态性, 尤其是Asn86Tyr、Tyr184-PHE、Ser1034Cys、Asn1042Asp和Asp1246Tyr已经鉴定并报道影响对苯苄醇、青蒿素、奎宁、氟喹、卤代蒽醌和氯喹的灵敏度有影响。此外, PfMDR1的扩增与对苯苄醇、青蒿素、奎宁、甲氟喹和卤代蒽醌的灵敏度降低有关, 并且PfMDR1的失调导致氯喹抗性增加。也可以检测pfmdr1的扩增。PfMDR1的磷酸化状态也具有高度相关性。

[0433] 在恶性疟原虫多药物抗性相关蛋白 (PfMRP) (基因参考PFA0590w) 中, 位置191和/或437的多态性例如Y191H和A437S已被鉴定, 并氯喹抗性表型相关。

[0434] 在恶性疟原虫NA⁺/H⁺交换蛋白 (PfNHE) (参考PF13_0019) 中, 在微卫星ms4670的DNNND的增加的重复可以是用于奎宁抗性的标志物。

[0435] 改变由细胞色素b基因 (Cytb, mal_mito_3) 编码的细胞色素b蛋白的泛醇结合位点的突变是与阿托伐醌抗性相关。位置26、268、276、133和280的突变特别是Tyr26Asn、Tyr268Ser、M1331和G280D可能与阿托伐醌抗性相关。

[0436] 例如在间日疟原虫中, PvMDR1 (PfMDR1的同系物) 的突变已与氯喹抗性有关, 特别是位置976的多态性, 例如突变Y976F。

[0437] 上面的突变以蛋白质的序列来定义。然而, 技术人员能够确定以核酸靶序列鉴定的相应的突变, 包括SNP。

[0438] 其他鉴定的药物抗性标志物是本领域已知的, 例如, 如在“Susceptibility of Plasmodium falciparum to antimalarial drugs (1996-2004)”; WHO; Artemisinin and artemisinin-based combination therapy resistance (April 2016 WHO/HTM/GMP/2016.5); “Drug-resistant malaria: molecular mechanisms and implications for public health” FEBS Lett. 2011 Jun 6; 585 (11): 1551-62. doi: 10.1016/j.febslet.2011.04.042. Epub 2011 Apr 23. Review. PubMed PMID: 21530510中描述的; 其内容通过引用并入本文。

[0439] 至于可根据本发明检测的多肽, 本文中提及的所有基因的基因产物可被用作靶标。相应地, 预期这样的多肽可用于物种鉴定, 分型和/或药物抗性检测。

[0440] 在某些示例性实施方式中, 本文公开的系统、装置和方法涉及检测样品 (例如从受试者获得的生物样品) 中一种或多种蚊子传播的寄生虫的存在。在某些示例性实施方式中, 寄生虫可以从选自物种恶性疟原虫、间日疟原虫、卵形疟原虫、三日疟原虫或诺氏疟原虫。

因此,本文所公开的方法可以适用于其他方法(或组合),其他方法需要快速鉴定寄生虫物种,监测寄生虫和寄生虫形式(例如对应于感染的各个阶段和寄生虫生命-周期,例如外红细胞周期、红细胞周期、孢子周期;寄生虫形式包括裂殖子、子孢子、裂殖体、配子体)的存在;检测某些表型(例如病原体药物抗性)、监测疾病进展和/或爆发,以及治疗(药物)筛查。此外,在疟疾的情况下,在感染性咬伤之后可能经过很长时间,即在此期间患者不表现出症状的长的潜伏期。同样,预防性治疗可以延缓症状的出现,并且在复发之前也可以观察到长的无症状期。这种延迟很容易导致误诊或延误诊断,从而损害治疗的有效性。

[0441] 由于本文所公开的实施方式的快速且灵敏的诊断能力,低至单核苷酸差异的寄生虫类型的检测,以及作为POC装置部署的能力,本文所公开的实施方式可用于指导治疗方案,例如选择适当的疗程。本文所公开的实施方式还可用于筛选环境样品(蚊子群体等)以确定寄生虫的存在和分型。还可以修改实施方式以同时检测蚊子传播的寄生虫和其他蚊子传播的病原体。在某些情况下,疟疾和其他由蚊子传播的病原体最初可能出现类似的症状。因此,快速区分感染类型的能力可以指导重要的治疗决策。可与疟疾一起检测的其他蚊子传播的病原体包括登革热、西尼罗河病毒、基孔肯雅热、黄热病,丝虫病、日本脑炎、圣路易斯脑炎、西部马脑炎、东部马脑炎、委内瑞拉马脑炎、拉克罗斯脑炎和寨卡。

[0442] 在某些示例性实施方式中,本文所公开的装置、系统和方法可用于区分样品中蚊子传播寄生虫物种。在某些示例性实施方式中,鉴定可以基于核糖体RNA序列,包括18S、16S、23S和5S亚基。在某些示例性实施方式中,鉴定可以基于在基因组中以多个拷贝存在的基因序列,例如线粒体基因例如CYTB。在某些示例性实施方式中,鉴定可以基于高度表达和/或高度保守的基因序列,例如GAPDH、组蛋白H2B、烯醇酶或LDH。用于鉴定相关rRNA序列的方法公开于美国专利申请公开号2017/0029872中。在某些示例性实施方式中,指导RNA组可以被设计为通过各物种或菌株独特的可变区来区分各物种。指导RNA还可以被设计为靶向区分在属、科、目、纲、门、界水平的微生物的RNA基因或其组合。在使用扩增的某些示例性实施方式中,可以设计扩增引物组以侧接核糖体RNA序列的恒定区和设计用于通过可变内部区域区分各物种的指导RNA。在某些示例性实施方式中,引物和指导RNA可以被设计为分别在16S亚基中保守和可变区。也可以使用在物种或物种子集中唯一可变的其他基因或基因组区域,例如RecA基因家族、RNA聚合酶 β 亚基。其他合适的系统发育标志物及其鉴定方法例如在Wu等arXiv:1307.8690[q-bio.GN]中讨论。

[0443] 在某些示例性实施方式中,可以基于以多拷贝存在于基因组中的基因例如线粒体基因如CYTB基因来执行物种鉴定。在某些示例性实施方式中,可以基于高度表达和/或高度保守的基因(例如GAPDH、组蛋白H2B、烯醇酶或LDH)进行物种鉴定。

[0444] 在某些示例性实施方式中,方法或诊断被设计为同时跨多个系统发育和/或表型的水平筛选蚊子传播的寄生虫。例如,方法或诊断可包括使用具有不同指导RNA的多种CRISPR系统。第一组指导RNA可以区分,例如,恶性疟原虫或间日疟原虫。这些一般纲可以进一步细分。例如,指导RNA可以被设计并用于区分药物抗性菌株(通常或针对特定药物或药物的组合)的方法或诊断中。可以设计第二组指导RNA以区分物种水平的微生物。因此,可以产生识别所有蚊子传播的寄生虫物种或亚种的基质,进一步根据药物抗性进行划分。前述仅用于示例目的。还考虑了用于分类其他类型的蚊子传播的寄生虫的其他手段,并且其将遵循上述一般结构。

[0445] 在某些示例性实施方式中,本文所公开的装置、系统和方法可用于筛选感兴趣的蚊子传播的寄生虫基因,例如药物抗性基因。指导RNA可以被设计以区分已知的感兴趣的基因。然后可以使用本文所公开的实施方式筛选样品,包括临床样品,用于检测一种或多种这样的基因。在POC筛选药物抗性的能力将在选择适当的治疗方案方面具有巨大的益处。在某些示例性实施方式中,药物抗性基因是编码蛋白质的基因,例如转运蛋白,例如来自药物/代谢物转运蛋白家族的蛋白质,参与底物易位的ATP结合盒(ABC)蛋白质,例如ABC转运蛋白C亚家族或Na⁺/H⁺交换蛋白;参与叶酸途径的蛋白,例如二氢蝶酸合成酶、二氢叶酸还原酶活性或二氢叶酸还原酶-胸苷酸合成酶;和参与跨越线粒体内膜的质子的转运的蛋白,特别是细胞色素b复合物。另外的靶标还可以包括编码血红素聚合酶的基因。在某些示例性实施方式中,药物抗性基因选自恶性疟原虫氯喹抗性转运蛋白基因(pfcr1)、恶性疟原虫多药物抗性转运蛋白1(pfmdr1)、恶性疟原虫多药物抗性相关蛋白基因(Pfmrp)、恶性疟原虫Na⁺/H⁺交换基因(pfche)、恶性疟原虫Ca²⁺转运ATP酶6(pfatp6)、恶性疟原虫二氢蝶呤合成酶(pfdhps)、二氢叶酸还原酶活性(pfdhpr)和二氢叶酸还原酶-胸苷酸合成酶(pfdhfr)的基因、细胞色素b基因、GTP环化水解酶和Kelch13(K13)基因,以及它们在其他疟原虫物种中的功能性同源基因。其他鉴定的药物抗性标志物是本领域已知的,例如在“Susceptibility of Plasmodium falciparum to antimalarial drugs (1996-2004)”;WHO;Artemisinin and artemisinin-based combination therapy resistance (April 2016WHO/HTM/GMP/2016.5);“Drug-resistant malaria:molecular mechanisms and implications for public health”FEBS Lett.2011Jun 6;585(11):1551-62.doi:10.1016/j.febslet.2011.04.042.Epub 2011Apr 23.Review.PubMed PMID:21530510中所描述的;其内容通过引用并入本文。

[0446] 在一些实施方式中,本文所述的CRISPR系统、检测系统或其使用方法可用于确定蚊子传播的寄生虫爆发的进化。该方法可以包括从来自一个或多个受试者的多个样品中检测一种或多种靶序列,其中靶序列是来自传播或引起爆发的蚊子传播的寄生虫的序列。这样的方法可以进一步包括确定蚊子传播的寄生虫传播的模式,或由蚊子传播的寄生虫引起的疾病暴发中涉及的机制。样品可以衍生自一个或多个人,和/或衍生自一只或多只蚊子。

[0447] 病原体传播的模式可以包括从蚊子传播的寄生虫的天然库的持续新传播或在从天然库的单个传播后的其他传播或两者的混合。在一个实施方式中,靶序列优选是在蚊子传播的寄生虫基因组中的序列或其片段。在一个实施方式中,蚊子传播的寄生虫传播的模式是蚊子传播的寄生虫传播的早期模式,即在蚊子传播的寄生虫爆发之初。在爆发开始时确定蚊子传播寄生虫传播的模式增加了尽早停止疫情的可能性,从而减少了本地和国际传播的可能性。

[0448] 确定蚊子传播的寄生虫传播的模式可包括根据本文所述的方法检测蚊子传播的寄生虫的序列。确定病原体传播的模式可以进一步包括检测受试者之间的蚊子传播的寄生虫的序列的共享的宿主内变异,并确定共享的宿主内变异是否显示时间模式。观察到的宿主内和宿主间变异的模式提供了关于传播和流行病学的重要见解(Gire,等,2014)。

[0449] 除了本文所公开的其它样品类型,样品可以从一只或多只蚊子衍生,例如,样品可以包含蚊子的唾液。

[0450] 生物标志物的检测

[0451] 在某些示例性实施方式中,本文所公开的系统、装置和方法可用于生物标志物的检测。例如,本文所公开的系统、装置和方法可用于SNP检测和/或基因分型。本文所公开的系统、装置和方法还可用于检测以异常基因表达为特征的任何疾病状态或病症。异常基因表达包括表达基因的畸变、表达的位置和表达水平。可以检测与心血管、免疫紊乱和癌症以及其他疾病相关的多种转录物或蛋白质标志物。在某些示例性实施方式中,本文所公开的实施方式可用于涉及裂解的疾病的无细胞DNA检测,例如肝纤维化和限制性/阻塞性肺病。在某些示例性实施方式中,实施方式可用于更快和更便携的无细胞DNA的产前检测。本文所公开的实施方式可以用于筛选与心血管健康、脂质/代谢特征、种族识别、亲子鉴定、人类ID(例如,将嫌疑人与SNP标签的犯罪数据库匹配)相关联的不同SNP的小组。本文所公开的实施方式还可以用于与癌症肿瘤相关和从癌症肿瘤释放的突变的无细胞DNA的检测。本文所公开的实施方式还可用于检测肉质,例如,通过提供给定肉制品中不同动物来源的快速检测。本文所公开的实施方式还可用于检测GMO或与DNA相关的基因编辑。如本文其他地方所述,可以通过在gRNA中引入合成错配来区分密切相关的基因型/等位基因或生物标志物(例如,在给定的靶序列中仅具有单个核苷酸差异)。

[0452] 在一个方面,本发明涉及用于检测样品中的靶核酸的方法,包括:

[0453] a. 将样品或样品组分配到一个或多个个体离散体积中,所述个体离散体积包含根据本文所述的本发明的CRISPR系统;

[0454] b. 在足以允许所述一种或多种指导RNA与一种或多种靶分子结合的条件下孵育所述样品或样品组;

[0455] c. 通过所述一种或多种指导RNA与所述一种或多种靶分子的结合激活所述CRISPR效应蛋白,其中激活所述CRISPR效应蛋白导致所述基于RNA的掩蔽构建体的修饰,使得可检测阳性信号产生;和

[0456] d. 检测所述可检测阳性信号,其中所述可检测阳性信号的检测指示一种或多种靶分子在所述样品中的存在。

[0457] 生物标志物样品类型

[0458] 本文所述的测定的灵敏度是非常适合于在多种生物样品的类型中检测靶核酸,包括其中靶核酸是稀的或样品材料是有限的的样品类型。生物标志物筛选可以在许多样品类型上进行,包括但不限于唾液、尿液、血液、粪便、痰液和脑脊髓液。本文所公开的实施方式还可用于检测基因的上调和/或下调。例如,因为样品可以连续稀释,使得仅过表达的基因保持高于测定的检测极限阈值。

[0459] 在某些实施方式中,本发明提供了获得生物流体(例如尿、血浆或血清、痰、脑脊髓液)的样品并提取DNA的步骤。待检测的突变核苷酸序列可以是较大分子的一部分或最初可以作为离散分子存在。

[0460] 在某些实施方式中,从癌症患者的血浆/血清中分离DNA。为了比较,可以从同一患者的非肿瘤组织(对照)中分离从肿瘤组织分离的DNA样品和第二样品,例如淋巴细胞。非肿瘤组织可以与肿瘤组织是相同类型或来自不同的器官来源。在某些实施方式中,收集血液样品并通过离心立即将血浆与血细胞分离。可以过滤血清并冷冻保存直至DNA提取。

[0461] 在某些示例性实施方式中,靶核酸是从粗的或未经处理的样品直接检测,例如血液、血清、唾液、脑脊液、痰、或尿。在某些示例性实施方式中,靶核酸是无细胞DNA。

[0462] 循环肿瘤细胞

[0463] 在一个实施方式中,可以用本发明测定循环细胞(例如,循环肿瘤细胞(CTC))。可以进行用于本文所述任何方法的循环肿瘤细胞(CTC)的分离。可用于本发明达到特异性和灵敏检测和捕获循环细胞的示例性技术已经描述(Mostert B,等,Circulating tumor cells (CTCs):detection methods and their clinical relevance in breast cancer.Cancer Treat Rev.2009;35:463-474;和Talasaz AH,等,Isolating highly enriched populations of circulating epithelial cells and other rare cells from blood using a magnetic sweeper device.Proc Natl Acad Sci U S A.2009;106:3970-3975)。在CTC105-106外周血单核细胞的背景中可以找到少至一个CTC(Ross A A,等, Detection and viability of tumor cells in peripheral blood stem cell collections from breast cancer patients using immunocytochemical and clonogenic assay techniques.Blood.1993,82:2605-2610)。**CellSearch®**平台使用涂有抗上皮细胞粘附分子(EpCAM)抗体的免疫磁珠来富集表达EPCAM的上皮细胞,然后进行免疫染色以确认细胞角蛋白染色的存在和白细胞标志物CD45的缺失,以确认捕获的细胞是上皮肿瘤细胞(Momburg F,等,Immunohistochemical study of the expression of a Mr 34,000human epithelium-specific surface glycoprotein in normal and malignant tissues.Cancer Res.1987;47:2883-2891;和Allard WJ,等,Tumor cells circulate in the peripheral blood of all major carcinomas but not in healthy subjects or patients with nonmalignant diseases.Clin Cancer Res.2004;10:6897-6904)。已经前瞻性地证明捕获的细胞数对具有晚期疾病的乳腺癌、结肠直肠癌和前列腺癌患者具有预后意义(Cohen SJ,等,J Clin Oncol.2008;26:3213-3221;Cristofanilli M,等N Engl J Med.2004;351:781-791;Cristofanilli M,et al.,J Clin Oncol.2005;23:1420-1430;和 de Bono JS,等Clin Cancer Res.2008;14:6302-6309)。

[0464] 本发明还提供了用CTC-Chip技术分离CTC。CTC-Chip是一种基于微流体的CTC捕获装置,其中血液流过包含数千个包被CTC结合的抗EpCAM抗体的微球的腔室(Nagrath S,等 Isolation of rare circulating tumor cells in cancer patients by microchip technology.Nature.2007;450:1235-1239)。与**CellSearch®**系统相比,CTC-Chip显著提高了CTC计数和纯度(Maheswaran S,等Detection of mutations in EGFR in circulating lung-cancer cells,N Engl J Med.2008;359:366-377),两个平台均可用于下游分子分析。

[0465] 无细胞染色质

[0466] 在某些实施方式中,根据本发明分离和分析无细胞染色质片段。可以在健康个体(Stroun等Annals of the New York Academy of Sciences 906:161-168(2000))以及患有疾病状态的个体的血清中检测核小体。此外,患有良性和恶性疾病(例如癌症和自身免疫疾病)的患者的核小体的血清浓度相当高(Holdenrieder等(2001) Int J Cancer95,1 14-120,Trejo-Becerril et al(2003) Int J Cancer 104,663-668;Kuroi等1999Breast Cancer 6,361-364;Kuroi et al(2001) Int j Oncology 19,143-148;Amoura et al (1997) Arth Rheum 40,2217-2225;Williams et al(2001) J Rheumatol 28,81-94)。不受理论束缚,携带肿瘤的患者中高浓度的核小体源自增殖的肿瘤中自发发生的细胞凋亡。

在血液中循环的核小体含有独特的修饰组蛋白。例如,美国专利公开号2005/0069931(2005年3月31日)涉及使用针对特定组蛋白N-末端修饰的抗体作为疾病的诊断指标,使用这样的组蛋白特异性抗体从患者的血液或血清样品中分离核小体以便于纯化和分析伴随的DNA用于诊断/筛选目的。因此,本发明可以使用染色质结合的DNA来检测和监测例如肿瘤突变。与修饰组蛋白相关的DNA的鉴定可以作为疾病和先天性缺陷的诊断标志物。

[0467] 因此,在另一个实施方式中,分离的染色质片段衍生自循环染色质,优选循环单和寡核小体。分离的染色质片段可以衍生自生物样品。生物样品可以来自受试者或有需要的患者。生物样品可以是血清、血浆、淋巴液、血液、血液级分、尿液、滑液、脊髓液、唾液、循环肿瘤细胞或粘液。

[0468] 无细胞DNA(cfDNA)

[0469] 在某些实施方式中,本发明可用于检测无细胞DNA(cfDNA)。血浆或血清中的无细胞DNA可用作非侵入性诊断工具。例如,已经研究并优化了无细胞胎儿DNA,用于测试相容的RhD因子、X连锁遗传病症的性别测定、单基因病症的测试、先兆子痫的鉴定。例如,对母体血浆中cfDNA的胎儿细胞级分进行测序是检测与胎儿染色体非整倍性相关的拷贝数变化的可靠方法。另一个实例,从癌症患者中分离的cfDNA已被用于检测与治疗决策相关的关键基因的突变。

[0470] 在某些示例性实施方式中,本公开提供了直接从患者样品中检测cfDNA。在某些其他示例性实施方式中,本公开使用上文公开的富集实施方式并且在检测靶cfDNA之前提供富集的cfDNA。

[0471] 外来体

[0472] 在一个实施方式中,可以用本发明测定外来体。外来体是小的细胞外囊泡,已被证明含有RNA。通过超速离心、过滤、化学沉淀、尺寸排阻色谱和微流体分离外来体是本领域已知的。在一个实施方式中,使用外来体生物标志物纯化外来体。可以通过任何已知方法进行来自生物样品的外来体的分离和纯化(参见例如W02016172598A1)。SNP检测和基因分型

[0473] 在某些实施方式中,本发明可用于检测生物样品中单核苷酸多态性(SNP)的存在。SNP可以与产妇测试(例如性别决定、胎儿缺陷)有关。它们可以与刑事调查有关。在一个实施方式中,可以通过本发明识别刑事调查中的嫌疑人。不受基于理论核酸的法医证据的约束,可以需要可获得的最灵敏的测定来检测嫌疑人或受害者的遗传物质,因为测试的样品可能是有限的。

[0474] 在其它实施方式中,与疾病相关的SNP为本发明所涵盖。与疾病相关的SNP是本领域公知的,并且本领域技术人员可以应用本发明的方法来设计合适的指导RNA(参见例如www.ncbi.nlm.nih.gov/clinvar?term=human%5Borgn%5D)。

[0475] 在一个方面,本发明涉及用于基因分型的方法,例如SNP基因分型,其包括:

[0476] a) 将样品或样品组分配到一个或多个个体离散体积中,所述个体离散体积包含如本文所述的根据本发明的CRISPR系统;

[0477] b) 在足以允许一种或多种指导RNA与一种或多种靶分子结合的条件下孵育样品或样品组;

[0478] c) 通过所述一种或多种指导RNA与所述一种或多种靶分子结合激活CRISPR效应蛋白,其中激活所述CRISPR效应蛋白导致所述基于RNA的掩蔽构建体的修饰,使得可检测阳性

信号产生;和

[0479] d) 检测所述可检测阳性信号,其中所述可检测阳性信号的检测指示为特定基因型特征的一种或多种靶分子在样品中的存在。

[0480] 在某些实施方式中,将可检测信号与(例如通过比较信号强度)一种或多种标准品信号进行比较,优选合成标准信号,例如图60中的示例性实施方式中所示。在某些实施方式中,标准品是或对应于特定基因型。在某些实施方式中,标准品包含特定的SNP或其他(单个)核苷酸变异。在某些实施方式中,标准品是(PCR扩增的)基因型标准品。在某些实施方式中,标准品是DNA或包含DNA。在某些实施方式中,标准品是或包含RNA。在某些实施方式中,标准品是或包含从DNA转录的RNA。在某些实施方式中,标准品是或包含从RNA逆转录的DNA。在某些实施方式中,将可检测信号与一种或多种标准品进行比较,每种标准品对应于已知的基因型,例如SNP或其他(单个)核苷酸变异。在某些实施方式中,将可检测信号与一个或多个标准品信号进行比较,并且比较包括统计分析,例如通过参数或非参数统计分析,例如通过单向或双向ANOVA等。在某些实施方式中,可检测信号是与一种或多种标准品信号相比较,并且当可检测信号不(统计学上)显著偏离标准品时,基因型被确定为对应于所述标准品的基因型。

[0481] 在其它实施方式中,本发明允许用于紧急药物基因组学的快速基因分型。在一个实施方式中,可以使用单个护理点对进入急诊室的患者进行基因分型。可能怀疑患者患有血凝块,并且急诊医师需要确定血液稀释剂剂量管理。在示例性实施方式中,本发明可以基于例如VKORC1、CYP2C9和CYP2C19的标志物的基因分型提供在心肌梗塞或中风治疗期间施用血液稀释剂的指导。在一个实施方式中,血液稀释剂是抗凝血剂华法林(Holford, NH (1986年12月) "Clinical Pharmacokinetics and Pharmacodynamics of Warfarin Understanding the Dose-Effect Relationship". Clinical Pharmacokinetics. Springer International Publishing. 11 (6) :483-504)。与血液凝固相关的基因是本领域已知的(参见例如US20060166239A1; Litin SC, Gastineau DA (1995) "Current concepts in anticoagulant therapy". Mayo Clin. Proc. 70 (3) :266-72; 和 Rusdiana等, Responsiveness to low-dose warfarin associated with genetic variants of VKORC1, CYP2C9, CYP2C19, and CYP4F2 in an Indonesian population. Eur J Clin Pharmacol. 2013 Mar; 69 (3) :395-405)。具体地,在VKORC1 1639(或3673)单核苷酸多态性中,共有("野生型")G等位基因被A等位基因取代。具有A等位基因(或"A单倍型")的人比具有G等位基因(或"非A单倍型")的人产生更少的VKORC1。这些变异体的流行程度也因种族而异,37%的白种人和14%的非洲人携带A等位基因。最终结果是凝血因子的数量减少,因此凝血能力降低。

[0482] 在某些示例性实施方式中,用于检测患者中的SNP的遗传物质的可用性允许在不扩增DNA或RNA样品的情况下检测SNP。在基因分型的情况下,容易获得测试的生物样品。在某些示例性实施方式中,可以缩短本发明的孵育时间。可以在酶促反应发生所需的一段时间内进行测定。本领域技术人员可以在5分钟内(例如5分钟连接)进行生化反应。本发明可以使用自动DNA提取装置从血液中获得DNA。然后将DNA添加到产生效应蛋白的靶分子的反应中。在产生靶分子后,可以立即切割掩蔽剂并检测信号。在示例性实施方式中本发明允许POC快速诊断用于在施用药物(例如血液稀释剂)之前确定基因型。在使用扩增步骤的

情况下,所有反应在一步法中在相同的反应中发生。在优选的实施方式中,POC测定可在少于1小时、优选10分钟、20分钟、30分钟、40分钟或50分钟内进行。

[0483] 在某些实施方式中,本文所公开的系统、装置和方法可用于检测长非编码RNA (lncRNA) 的存在或表达水平。某些lncRNA的表达与疾病状态和/或药物抗性相关。特别地,某些lncRNA (例如,TCONS_00011252,NR_034078,TCONS_00010506,TCONS_00026344,TCONS_00015940,TCONS_00028298,TCONS_00026380,TCONS_0009861,TCONS_00026521,TCONS_00016127,NR_125939,NR_033834,TCONS_00021026,TCONS_00006579,NR_109890和NR_026873) 与癌症治疗的抗性相关,例如对用于治疗黑素瘤 (例如结节性黑素瘤、恶性黑斑病、恶性黑色素瘤、肢端黑色素瘤、浅表扩散黑素瘤、粘膜黑色素瘤、息肉样黑素瘤、促纤维增生性黑素瘤、无黑色素瘤黑色素瘤和软组织黑素瘤) 的一种或多种BRAF抑制剂 (例如Vemurafenib、Dabrafenib、索拉非尼、GDC-0879、PLX-4720和LGX818) 的抗性。使用本文描述的各种实施方式检测lncRNA可以促进疾病诊断和/或治疗选择的选择。

[0484] 在一个实施方式中,本发明可以指导DNA或RNA靶向疗法 (例如CRISPR、TALE、锌指蛋白、RNAi),特别是在其中设置其中疗法的快速施用对治疗结果很重要的环境下。

[0485] LOH检测

[0486] 癌细胞与正常细胞相比经历遗传物质 (DNA) 的丧失。这种几乎所有 (如果不是全部) 癌症经历遗传物质的丧失被称为“杂合性丧失” (LOH)。杂合性丧失 (LOH) 是一种严重的染色体事件,导致整个基因和周围染色体区域的丧失。杂合性丧失在癌症中是常见事件,其中它可以指示不存在在丧失区域中的功能性肿瘤抑制基因。然而,丧失可以是沉默的,因为在染色体对的另一条染色体上仍然存在一个功能基因。肿瘤抑制基因的剩余拷贝可以通过点突变失活,导致肿瘤抑制基因的丧失。遗传物质从癌细胞丧失可导致染色体上的特定基因座处对细胞活力或细胞生长至关重要的基因的两个或更多个等位基因之一的选择性丧失。

[0487] “LOH标志物”是当与正常细胞相比时与癌症或其它疾病相关的来自微卫星基因座、缺失、改变、或扩增的DNA。LOH标志物通常与肿瘤抑制基因或另一种 (通常是肿瘤相关的) 基因的丧失有关。

[0488] 术语“微卫星”是指广泛分布于人类基因组的DNA的短重复序列。微卫星是串联重复 (即相邻) DNA基序的区域,其长度范围为2至5个核苷酸,并且通常是重复5-50次。例如,序列TATATATATA (SEQ. I. D. No. 418) 是二核苷酸微卫星,GTCGTCGTCGTCGTC (SEQ. I. D. No. 419) 是三核苷酸微卫星 (A是腺嘌呤,G鸟嘌呤,C胞嘧啶和T胸腺嘧啶)。已经显示这样的微卫星的重复长度的体细胞改变代表肿瘤的特征性特征。指导RNA可以被设计为检测这样的微卫星。此外,本发明可用于检测重复长度的改变,以及基于可检测信号的定量的扩增和缺失。某些微卫星位于基因的调控侧翼或内含子区域,或直接位于基因的密码子中。在这样的情况下微卫星突变可导致表型变化和疾病,尤其是三联体扩张性疾病,例如脆性X综合症和亨廷顿舞蹈病。

[0489] 在许多种恶性肿瘤中已经报道了在特定染色体区域频繁的杂合性丧失 (LOH)。特定染色体区域的等位基因丧失是在各种恶性肿瘤中观察到的最常见的遗传改变,因此微卫星分析已用于检测来自体液的样品中的癌细胞的DNA,例如肺癌的痰和膀胱癌的尿液。(Rouleau, et al. Nature 363, 515-521 (1993); and Latif, et al. Science 260, 1317-1320

(1993))。此外,已经确定,在患有癌症和一些其他疾病的个体的血浆中存在显著增加的可溶性DNA浓度,表明无细胞血清或血浆可用于检测具有微卫星异常的癌症DNA。(Kamp, et al. Science 264, 436-440 (1994); and Steck, et al. Nat Genet. 15 (4), 356-362 (1997))。两组已经报告了有限数量的小细胞肺癌或头颈癌患者的血浆或血清中的微卫星改变。(Hahn 等, Science 271, 350-353 (1996); 和 Miozzo 等, Cancer Res. 56, 2285-2288 (1996))。先前已经显示检测肿瘤和黑素瘤患者血清中的杂合性丧失(参见,例如,美国专利号US6465177B1)。

[0490] 因此,有利的是从或处于癌症风险中的受试者检测LOH标志物。本发明可用于检测肿瘤细胞中的LOH。在一个实施方式中,循环肿瘤细胞可用作生物样品。在优选的实施方式中,从血清或血浆中获得的无细胞DNA用于非侵入性地检测和/或监测LOH。在其他实施方式中,生物样品可以是本文所述的任何样品(例如用于膀胱癌的尿液样品)。不受理论束缚,本发明可用于与任何在先的方法相比具有改善的灵敏度检测LOH标志物,从而提供突发事件的早期检测。在一个实施方式中,在生物流体中检测LOH,其中LOH的存在与癌症的发生相关。本文描述的方法和系统通过提供用于检测与癌症相关的特定等位基因的LOH的非侵入性、快速且准确的方法,代表了优于现有技术(例如PCR或组织活检)的显著进步。因此,本发明提供了一种方法和系统,其可用于筛选高风险群体并监测经历化学预防、化学疗法、免疫疗法或其他治疗的高风险患者。

[0491] 由于本发明的方法要求只从体液例如血液提取DNA,它可以在任何时间执行,并且重复地针对单个患者。在手术前或手术后;治疗前、治疗期间和治疗后,例如化疗、放射治疗、基因治疗或免疫治疗;或在治疗后对疾病进展,稳定或复发的随访期间可采集血液并监测LOH。不受理论束缚,本发明的方法还可用于检测亚临床疾病的存在或复发,其中LOH标志物对该患者具有特异性,因为LOH标志物对个体患者的肿瘤是特异性的。所述方法还可以使用肿瘤特异性LOH标志物检测是否存在多个转移。

[0492] 检测表观遗传修饰

[0493] 指示癌症或癌症进展的组蛋白变体、DNA修饰和组蛋白修饰可用于本发明。例如,美国专利公开20140206014描述了癌症样品与健康受试者相比具有升高的核小体H2AZ、macroH2A1.1, 5-甲基胞嘧啶、P-H2AX (Ser139) 水平。由于癌细胞的凋亡增加,个体中癌细胞的存在可以在血液中产生更高水平的无细胞核小体。在一个实施方式中,针对与凋亡相关的标志物的抗体,例如H2BSer14 (P),可用于鉴定已从凋亡的肿瘤细胞释放的单个核小体。因此,根据本发明可以有利地以高灵敏度和准确度分析由肿瘤细胞产生的DNA。

[0494] 产前筛查

[0495] 在某些实施方式中,本发明的方法和系统可用于产前筛查。在某些实施方式中,无细胞DNA用于产前筛选的方法中。在某些实施方式中,可以用本发明检测与单核小体或寡核小体相关的DNA。在优选的实施方式中,检测与单核小体或寡核小体相关的DNA用于产前筛选。在某些实施方式中,无细胞染色质片段用于产前筛选的方法中。

[0496] 产前诊断或产前筛查是指在胎儿或胚胎出生前检测疾病或病症。目的是检测出生缺陷,例如神经管缺陷、唐氏综合症、染色体异常、遗传性疾病和其他病症,例如脊柱裂、腭裂、泰萨克斯病、镰状细胞性贫血、地中海贫血、囊性纤维化、肌营养不良和脆性X综合症。筛查也可用于产前性别辨别。常见的测试程序包括羊膜穿刺术、超声检查,包括颈部半透明超声检查、血清标志物检测或遗传筛查。在某些情况下,进行测试以确定胎儿是否会中止,尽

管医生和患者也发现早期诊断高风险妊娠是有用的,这样就可以安排宝宝可以得到适当的护理的分娩护理医院。

[0497] 已经认识到母体血液中存在胎儿细胞,并且这些细胞为产前基于DNA的诊断提供了胎儿染色体的潜在来源。另外,胎儿DNA的范围是母体血液中总DNA的约2-10%。目前可用的产前基因检测通常涉及侵入性程序。例如,在妊娠10-12周左右对孕妇进行绒毛膜绒毛取样(CVS),并且在14-16周左右进行羊膜穿刺术均包含侵入性程序以获得用于测试胎儿染色体异常的样品。通过这些采样程序获得的胎儿细胞通常使用细胞遗传学或荧光原位杂交(FISH)分析来测试染色体异常。早在妊娠第6周,无细胞胎儿DNA就已存在于孕妇血浆和血清中,妊娠期间浓度升高,分娩前达到峰值。因为这些细胞在怀孕的早期出现,它们可以形成准确、无创、妊娠早期测试的基础。不受理论束缚,本发明在检测少量胎儿DNA方面提供了前所未有的灵敏度。不受理论束缚,大量的母体DNA通常与感兴趣的胎儿DNA一起同时恢复,因此降低了胎儿DNA定量和突变检测的灵敏度。本发明通过该测定的出乎意料的高灵敏度克服了这些问题。

[0498] H3类组蛋白由四种不同的蛋白质类型组成:主要类型,H3.1和H3.2;替换类型,H3.3;和睾丸特异性变体,H3t。尽管H3.1和H3.2密切相关,仅在Ser96上不同,H3.1在至少5个氨基酸位置上与H3.3不同。此外,H3.1相比于其在包括肝、肾和心脏的成体组织中的存在存在在胎肝中高度富集。在成人组织中,H3.3变体比H3.1变体更丰富,而对于胎肝则相反。本发明可以使用这些差异来检测包含胎儿和母体细胞两者和/或胎儿核酸的母体生物样品中的胎儿核小体和胎儿核酸。

[0499] 在一个实施方式中,胎儿核小体可以从血液中获得。在其他实施方式中,胎儿核小体获自宫颈粘液样品。在某些实施方式中,通过在妊娠中期的早期或妊娠早期的晚期从孕妇擦拭或灌洗获得宫颈粘液样品。可以将样品置于培养箱中以释放被困在粘液中的DNA。培养箱可以设定在37°C。可以将样品摇动约15至30分钟。为了释放DNA,可以用粘蛋白酶进一步溶解粘液。如本领域所公知的,样品也可以经受条件,例如化学处理等,以诱导细胞凋亡以释放胎儿核小体。因此,可以用诱导细胞凋亡的试剂处理宫颈粘液样品,由此释放胎儿核小体。关于循环胎儿DNA的富集,参考美国专利公开号20070243549和20100240054。当将所述方法和系统应用于产前筛选时,本发明是特别有利的,其中只有一小部分核小体或DNA可以来源于胎儿。

[0500] 根据本发明的产前筛查可以用于疾病,包括但不限于13三体、16三体、18三体、Klinefelter综合症(47,XXY)、(47,XYY)和(47,XXX)、特纳综合症、唐氏综合症(21三体)、囊性纤维化、亨廷顿病、 β 地中海贫血、肌强直性营养不良、镰状细胞贫血、卟啉症、脆性X综合症、罗伯逊易位、Angelman综合症、DiGeorge综合症和Wolf-Hirschhorn综合症。

[0501] 本发明的另外多个方面涉及诊断、预后和/或治疗与多种遗传疾病相关的缺陷,这些缺陷在国立卫生研究院的网站上以遗传性疾病为主题进一步描述(health.nih.gov/topic/Genetic Disorders的网站)。

[0502] 癌症和癌症药物抗性检测

[0503] 在某些实施方式中,本发明可用于检测与癌症相关的基因和突变。在某些实施方式中,检测与抗性相关的突变。抗性肿瘤细胞的扩增或抗性突变在肿瘤细胞的克隆群体中出现可以在治疗期间出现(参见,例如Burger JA,等,Clonal evolution in patients

with chronic lymphocytic leukaemia developing resistance to BTK inhibition. *Nat Commun.* 2016 May 20; 7:11589; Landau DA, 等, Mutations driving CLL and their evolution in progression and relapse. *Nature.* 2015 Oct 22; 526 (7574): 525-30; Landau DA, 等, Clonal evolution in hematological malignancies and therapeutic implications. *Leukemia.* 2014 Jan; 28 (1): 34-43; 和 Landau DA, 等, Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell.* 2013 Feb 14; 152 (4): 714-26)。因此, 检测这样的突变需要高度灵敏的测定, 并且监测需要重复的活组织检查。重复的活组织检查是不方便的、侵入性的和昂贵的。使用本领域已知的现有方法, 难以在血液样品或其他非侵入性收集的生物样品 (例如血液、唾液、尿液) 中检测抗性突变。抗性突变可以是指与化疗、靶向治疗或免疫疗法的抗性相关的突变。

[0504] 在某些实施方式中, 突变发生在可用于检测癌症进展的个体癌症中。在一个实施方式中, 已经表征了与针对肿瘤的T细胞溶细胞活性相关的突变, 并且可以通过本发明检测 (参见例如, Rooney等, *Molecular and genetic properties of tumors associated with local immune cytolytic activity, Cell.* 2015 January 15; 160 (1-2): 48-61)。可以基于这些突变的检测为患者开发个性化疗法 (参见例如 W02016100975A1)。在某些实施方式中, 与细胞溶解活性相关的癌症特异性突变可以是选自 CASP8、B2M、PIK3CA、SMC1A、ARID5B、TET2、ALPK2、COL5A1、TP53、DNER、NCOR1、MORC4、CIC、IRF6、MYOCD、ANKLE1、CNKSR1、NF1、SOS1、ARID2、CUL4B、DDX3X、FUBP1、TCP11L2、HLA-A、B或C、CSNK2A1、MET、ASXL1、PD-L1、PD-L2、ID01、ID02、ALOX12B和ALOX15B的基因中的突变、或拷贝数增加, 排除全染色体事件, 影响以下任何染色体条带: 6q16.1-q21、6q22.31-q24.1、6q25.1-q26、7p11.2-q11.1、8p23.1、8p11.23-p11.21 (含有ID01、ID02)、9p24.2-p23 (含有PDL1、PDL2)、10p15.3、10p15.1-p13、11p14.1、12p13.32-p13.2、17p13.1 (含有ALOX12B、ALOX15B) 和22q11.1-q11.21。

[0505] 在某些实施方式中, 本发明用于在治疗过程中和治疗完成后检测癌症突变 (例如抗性突变)。本发明的灵敏度可以允许非侵入性检测治疗期间出现的克隆突变, 并且可以用于检测疾病中的复发。

[0506] 在某些示例性实施方式中, 检测差异表达的microRNA (miRNA) 的miRNA和/或miRNA标签可用于检测或监测癌症的进展和/或检测对癌症疗法的药物抗性。例如, Nadal等 (*Nature Scientific Reports*, (2015) doi:10.1038/srep12464) 描述可用于检测非小细胞肺癌 (NSCLC) 的mRNA特征。

[0507] 在某些示例性实施方式中, 细胞克隆亚群中抗性突变的存在可用于确定治疗方案。在其他实施方式中, 可以基于常见的肿瘤突变施用用于治疗患者的个体化疗法。在某些实施方式中, 常见突变响应于治疗而出现并导致药物抗性。在某些实施方式中, 本发明可用于监测患者中获得具有这样的药物抗性突变的细胞的突变或扩增的细胞。

[0508] 用各种化学治疗剂治疗, 特别是用靶向治疗例如酪氨酸激酶抑制剂, 经常导致抵抗治疗剂的活性的靶分子中的新突变。正在评估克服这种抗性的多种策略, 包括开发不受这些突变影响的第二代疗法和用多种药剂治疗, 包括那些在抗性突变下游起作用的药物。在示例性实施方式中, 依鲁替尼 (靶向布鲁顿酪氨酸激酶 (BTK) 并用于CLL和某些淋巴瘤的分子) 的常见突变是481位的丝氨酸至丝氨酸变化 (BTK/C481S)。厄洛替尼 (靶向表皮生长因

子受体 (EGFR) 的酪氨酸激酶结构域) 通常用于治疗肺癌, 并且在治疗后总是发展出抗性肿瘤。在抗性克隆中发现的常见突变是790位的苏氨酸至甲硫氨酸突变。

[0509] 癌症患者的群体之间共享的非沉默突变和可用本发明检测的普遍抗性突变是本领域已知的 (参见例如, WO/2016/187508)。在某些实施方式中, 可以通过用依鲁替尼、厄洛替尼、伊马替尼、吉非替尼、克唑替尼、曲妥珠单抗、维罗非尼、RAF/MEK、检查点阻断疗法或抗雌激素疗法治疗诱导药物抗性突变。在某些实施方式中, 癌症特异性突变存在于编码选自程序性死亡-配体1 (PD-L1)、雄激素受体 (AR)、布鲁顿酪氨酸激酶 (BTK)、表皮生长因子受体 (EGFR)、BCR-Abl、c-kit、PIK3CA、HER2、EML4-ALK、KRAS、ALK、ROS1、AKT1、BRAF、MEK1、MEK2、NRAS、RAC1和ESR1的蛋白质的一种或多种基因中。

[0510] 免疫检查点是减慢或停止免疫反应, 并阻止免疫细胞的不受控制的活动导致过度组织损伤的抑制途径。在某些实施方式中, 靶向的免疫检查点是程序性死亡-1 (PD-1或CD279) 基因 (PDCD1)。在其他实施方式中, 靶向的免疫检查点是细胞毒性T淋巴细胞相关抗原 (CTLA-4)。在另外的实施方式中, 靶向的免疫检查点是CD28和CTLA4Ig超家族的另一成员, 例如BTLA、LAG3、ICOS、PDL1或KIR。在其他另外的实施方式中, 靶向的免疫检查点是TNFR超家族的成员, 例如CD40、OX40、CD137、GITR、CD27或TIM-3。

[0511] 最近, 在肿瘤及其微环境的基因表达已经在单细胞水平被表征 (参见例如, Tirosh, 等, Dissecting the multicellular ecosystem of metastatic melanoma by single cell RNA-seq. *Science* 352, 189-196, doi:10.1126/science.aad0501 (2016)); Tirosh等, Single-cell RNA-seq supports a developmental hierarchy in human oligodendrogloma. *Nature*. 2016 Nov 10; 539 (7628): 309-313. doi:10.1038/nature20123. Epub 2016 Nov 2; 和国际专利公开序列号WO 2017004153A1)。在某些实施方式中, 可以使用本发明检测基因标签。在一个实施方式中, 在肿瘤微环境中监测或检测补体基因。在一个实施方式中, 监视或检测MITF和AXL程序。在一个实施方式中, 检测肿瘤特异性干细胞或祖细胞标签。这样的标签表明免疫应答的状态和肿瘤的状态。在某些实施方式中, 可以检测肿瘤在增殖、抗性治疗和免疫细胞丰度方面的状态。

[0512] 因此, 在某些实施方式中, 本发明提供低成本的、快速、多重癌症检测组用于循环DNA, 例如肿瘤DNA, 特别是用于监测疾病复发或共同抗性突变的发展。

[0513] 免疫疗法应用

[0514] 本文所公开的实施方式还可用于其他免疫疗法环境中。例如, 在一些实施方式中, 在受试者中诊断、预后和/或分期免疫应答的方法包括检测一种或多种生物标志物的第一表达水平、活性和/或功能, 并将检测到的水平与对照水平进行比较, 其中检测水平和对照水平的差异表明受试者中存在免疫应答。

[0515] 在某些实施方式中, 本发明可用于确定肿瘤浸润淋巴细胞 (TIL) 的功能障碍或活化。可以使用已知方法从肿瘤中分离TIL。可以分析TIL以确定它们是否应该用于过继细胞转移疗法。另外, 可以在将它们施用于受试者之前分析嵌合抗原受体T细胞 (CART细胞) 的功能障碍或活化的标签。已经描述了功能障碍和激活T细胞的示例性标签 (参见, 例如Singer M等, A Distinct Gene Module for Dysfunction Uncoupled from Activation in Tumor-Infiltrating T Cells. *Cell*. 2016 Sep 8; 166 (6): 1500-1511. e9. doi:10.1016/j.cell.2016.08.052)。

[0516] 在一些实施方式中,C2c2用于评估免疫细胞的状态,例如T细胞(例如,CD8+和/或CD4+T细胞)。特别地,可以例如基于与一种或多种T细胞状态相关的基因或基因标签来确定T细胞活化和/或功能障碍。以这种方式,c2c2可用于确定一个或多个T细胞亚群的存在。

[0517] 在一些实施方式中,C2c2可以在诊断测定法中使用,或者可以被用作确定患者是否是适用于施用免疫疗法或另一种类型的治疗的方法。例如,可以通过c2c2进行基因或生物标志物标签的检测,以确定患者是否对给定的治疗有反应,或者如果患者没有反应,则可能是由于T细胞功能障碍。这样的检测提供了关于患者最适合接受的治疗类型的信息。例如,患者是否应该接受免疫疗法。

[0518] 在一些实施方式中,本文所公开的系统 and 测定可以允许临床医生识别患者对疗法的反应(例如,过继细胞转移(ACT)疗法)是否是由于细胞功能障碍,并且如果是,则生物标志物标签的上调和下调的水平将使问题得以解决。例如,如果接受ACT的患者是无反应性的,则可以通过本文所公开的测定法测定作为ACT的一部分施用的细胞,以确定已知与细胞活化和/或功能障碍状态相关的生物标记物标签的相对表达水平。如果特定的抑制性受体或分子在ACT细胞中上调,则可以用该受体或分子的抑制剂治疗患者。如果特定刺激受体或分子在ACT细胞中下调,则可以用该受体或分子的激动剂治疗患者。

[0519] 在某些示例性实施方式中,本文描述的系统、方法和装置可用于筛选识别特定细胞类型、细胞表型或细胞状态的基因标签。同样地,通过使用例如压缩感测(compressed sensing)的方法,本文所公开的实施方式可用于检测转录组。基因表达数据是高度结构化的,使得一些基因的表达水平预测其他基因的表达水平。基因表达数据高度结构化的知识允许假设系统中的自由度数是小的,这允许假设相对基因丰度的计算基础是稀疏的。有可能做出一些生物学上允许申请人在采样不足的情况下恢复非线性相互作用项而不具备哪些基因可能相互作用的具体知识的动机假设。特别是,如果申请人假设遗传相互作用是低等级,稀疏或这些的组合,那么真实的自由度数相对于完整的组合扩展很小,这使得申请人能够相对地推断出具有相对小数量的扰动的完整的非线性景观。围绕这些假设,矩阵完成和压缩感知的分析理论可用于设计采样不足的组合扰动实验。此外,核心学习框架可用于通过建立组合扰动的预测函数而不直接学习任何个体相互作用共系数来使用样品不足。压缩感知提供鉴定待检测的最小数量的靶转录物的方法以便获得全面的基因表达图谱。用于压缩感测的方法在2016年10月27日提交的PCT/US2016/059230“Systems and Methods for Determining Relative Abundances of Biomolecules”中公开,其通过引用并入本文。已经使用例如压缩感测的方法来鉴定最小转录物靶标组,然后可以设计相应的指导RNA组以检测所述转录物。因此,在某些示例性实施方式中,用于获得细胞的基因表达谱的方法包括检测、使用本文所公开的实施方式、提供细胞或细胞群的基因表达图谱的最小转录组。

[0520] 检测基因编辑和/或脱靶效应

[0521] 本文所公开的实施方式可以与其他基因编辑工具组合使用,以确认一个或多个期望的基因编辑是成功的和/或检测任何脱靶效应的存在。可以使用针对一个或多个靶基因座的一种或多种指导筛选已编辑的细胞。由于本文所公开的实施方式使用CRISPR系统,因此也设想了治疗诊断应用。例如,本文所公开的基因分型实施方式可用于选择合适的靶基因座或鉴定需要靶编辑的细胞或细胞群。然后可以使用相同或单独的系统来确定编辑效率。如下文的工作实施方式中所述,本文所公开的实施方式可用于在短至一周内设计流线

型治疗诊断途径。

[0522] 检测核酸标记的物品

[0523] 或者,本文所述的实施方式可用于检测核酸标识符。核酸标识符是可用于鉴定特定物品的非编码核酸。示例性核酸标识符,例如DNA水印,在Heider和Barnekow“DNA watermarks:A proof of concept”BMC Molecular Biology 9:40(2008)中描述。核酸标识符也可以是核酸条形码。基于核酸的条形码是用作相关分子的标识符(例如靶分子和/或靶核酸)的核苷酸的短序列(例如DNA、RNA或其组合)。核酸条形码的长度可以为至少,例如,4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、35、40、45、50、60、70、80、90或100个核苷酸,并且可以是单链或双链形式。可以将一个或多个核酸条形码连接或“标记”至靶分子和/或靶核酸。这种连接可以是直接的(例如条形码与靶分子的共价或非共价结合)或间接的(例如通过另外的分子,例如特异性结合剂,例如抗体(或其他蛋白质)或条形码接收衔接子(或其他核酸分子)。靶分子和/或靶核酸可以用组合方式的多个核酸条形码标记,例如核酸条形码多联体。通常,核酸条形码用于将靶分子和/或靶核酸鉴定为来自特定区室(例如离散体积)、具有特定物理性质(例如亲和力、长度、序列等),或已经受某些处理条件限制。靶分子和/或靶核酸可以与多个核酸条形码相关联以提供关于所有这些特征(和更多)的信息。产生核酸-条形码的方法公开于例如国际专利申请公开号W0/2014/047561中。

[0524] 酶

[0525] 申请还提供了C2c2的直系同源物,其表现出强大的活性,使得它们特别适合于RNA切割和检测的不同应用。这些应用包括但不限于本文所述的那些。更具体地,证实具有比其他测试的更强的活性的直系同源物是从生物体Leptotrichia wadei(LwC2c2)鉴定的C2c2直系同源物。因此,本申请提供了用于修饰感兴趣的基因座的方法,包括向所述基因座递送包含C2c2效应蛋白的非天然存在的或工程化的组合物,更具体地,如本文所述的具有增加的活性的C2c2效应蛋白和一种或多种核酸组分,其中至少一种或多种核酸组分被工程化,一种或多种核酸组分将复合物引导至感兴趣的靶标,并且效应蛋白与一种或多种核酸组分形成复合物,并且复合物与感兴趣的靶基因座结合。在特定的实施方式中,感兴趣的靶基因座包含RNA。申请进一步提供具有增加的活性的C2c2效应蛋白在RNA序列特异性干扰、RNA序列特异性基因调控、RNA或RNA产物或lincRNA或非编码RNA的筛选、或核RNA或mRNA、诱变、荧光原位杂交或育种中的用途。

[0526] 本发明将在下文的实施例中进一步描述,其不限制在权利要求中描述的本发明的范围。

[0527] 工作实施例

[0528] 实施例1总体方案

[0529] 提供两种对DNA和RNA进行C2c2诊断测试的方法。该方案也可以与在递送检测适体后蛋白质检测变体一起使用。第一种是两步反应(a two step reaction),其中扩增和C2c2检测分开进行。第二种是在一个反应中将所有组合在一起,这称为两步反应(a two-step reaction)。重要的是要记住,对于更高浓度的样品,扩增可能不是必要的,因此最好有单独的没有内置扩增的C2c2方案。

[0530] 仅CRISPR效应子-无扩增:

	组分	体积(μL)
[0531]	蛋白质 (最终 44 nM)	2
	crRNA (最终 12 nM)	1
	背景靶标 (总共 100 ng)	1
[0532]	靶 RNA (可变)	1
	RNA 传感器探针 (125 nM)	4
	MgCl ₂ (最终 6 mM)	2
	反应缓冲液 10×	2
	RNA 酶抑制剂 (来自 NEB 的鼠的)	2
	H ₂ O	5
	总共	20

[0533] 反应缓冲液是:40mM Tris-HCl,60mM NaCl,pH 7.3

[0534] 在37℃下进行该反应20分钟至3小时。用激发:485nm/20nm,发射:528nm/20nm读出。可以在20分钟时开始检测单分子灵敏度的信号,但是当然灵敏度对于更长的反应时间更高。

[0535] 两步反应:

[0536] RPA扩增混合物

[0537]

组分	体积(μL)
引物A (100μM)	0.48
引物B (100μM)	0.48
RPA缓冲液	59
MgAc	5
靶标(可变浓度)	5
ATP (100μM, 来自NEB试剂盒)	2
GTP (100μM, 来自NEB试剂盒)	2
UTP (100μM, 来自NEB试剂盒)	2
CTP (100μM, 来自NEB试剂盒)	2
T7聚合酶(来自NEB试剂盒)	2
H ₂ O	25
总共	104.96

[0538] 一起混合该反应,然后再悬浮冻干酶混合物的两到三管。向混合物中加入5μL的

280mM MgAc以开始反应。进行反应10-20分钟。每个反应都是20 μ L,因此这足以进行多达五次反应。

[0539] C2c2检测混合物

组分	体积 (μ L)
蛋白质 (最终 44 nM)	2
crRNA (最终 12 nM)	1
背景靶标 (总共 100 ng)	1
RPA 反应	1
[0540] RNA 传感器探针 (125 nM)	4
MgCl ₂ (最终 6 mM)	2
反应缓冲液 10 \times	2
RNA 酶抑制剂 (来自 NEB 的鼠的)	2
H ₂ O	5
总共	20

[0541] 反应缓冲液是:40mM Tris-HCL,60mM NaCl,pH 7.3

[0542] 进行该反应20分钟至3小时。最短检测时间为约20分钟以观察单分子灵敏度。进行反应更长时间只提高灵敏度。

[0543] 一锅反应:

组分	体积 (μ L)
[0544] 引物 A (100 μ M)	0.48
引物 B (100 μ M)	0.48
RPA 缓冲液	59

	MgAc	5
	Lw2C2c2 (最终 44 nM)	2
	crRNA (最终 12 nM)	2
	背景 RNA (从 250 ng/ μ L)	2
	RNA 酶警报底物 (在 20 μ L 中重悬后)	5
	鼠 RNA 酶抑制剂, 来自 NEB	10
[0545]	靶标 (可变浓度)	5
	ATP (100 μ M 来自 NEB 试剂盒)	2
	GTP (100 μ M 来自 NEB 试剂盒)	2
	UTP (100 μ M 来自 NEB 试剂盒)	2
	CTP (100 μ M 来自 NEB 试剂盒)	2
	T7 聚合酶 (来自 NEB 试剂盒)	2
	H ₂ O	4
	总共	104.96

[0546] 提到的NEB试剂盒是HighScribe T7高产量试剂盒。为了重悬缓冲液,使用1.5x浓度:将三管冻干的底物重悬于59 μ L缓冲液中,并在上述混合物中使用。每个反应为20 μ L,因此这足以进行5次反应。早至30-40分钟已经观察到该反应的单分子灵敏度。

[0547] 实施例2-来自LEPTOTRICHIA WADEI的C2C2介导DNA和RNA的高灵敏性和特异性检测

[0548] 快速、便宜和灵敏的核酸检测可有助于护理点的病原体检测,基因分型和疾病监测。RNA指导的、RNA靶向的CRISPR效应子Cas13a(以前称为C2c2)在靶标识别时表现出混杂的(promiscuous)RNA酶活性的“并行(collateral)效应”。申请人将Cas13a的并行效应与等温扩增组合以建立基于CRISPR的诊断(CRISPR-Dx),提供具有阿摩尔(attomolar)灵敏度和单碱基错配特异性的快速DNA或RNA检测。申请人使用这种基于Cas13a的分子检测平台(称为SHERLOCK(特异性高灵敏度酶报告子解锁))以检测寨卡病毒和登革热病毒的特定株、区分病原菌、基因分型人类DNA,和鉴定无细胞肿瘤DNA突变。此外,SHERLOCK反应试剂可以冻干用于冷链独立和长期储存,并且易于在纸上重构用于现场应用。

[0549] 在便携式平台上以高灵敏度和单碱基特异性快速检测核酸的能力可有助于疾病诊断和监测、流行病学和一般实验室任务。尽管存在用于检测核酸的方法(1-6),但它们在灵敏度、特异性、简单性、成本和速度之间权衡。微生物规律成簇的间隔短回文重复序列(CRISPR)和CRISPR相关(CRISPR-Cas)适应性免疫系统含有可用于基于CRISPR的诊断(CRISPR-Dx)的可编程核酸内切酶。虽然一些Cas酶靶向DNA(7,8),单效应子RNA指导的RNA

酶例如Cas13a (以前称为C2c2) (8) 可以用CRISPR RNA (crRNA) 重编程(9-11) 以提供用于特定RNA感测的平台。在识别其RNA靶标后, 激活的Cas13a参与附近的非靶向RNA的“并行”切割(10)。这种crRNA编程的并行切割活性允许Cas13a在体内通过触发程序性细胞死亡(10) 或在体外通过经标记RNA (10、12) 的非特异性降解来检测特定RNA的存在。这里申请人描述SHERLOCK (特异性高灵敏度酶报告子解锁), 其是基于商业化报告子RNA的核酸扩增和3Cas13a介导的并行切割(12) 的具有阿摩尔灵敏度的体外核酸检测平台, 允许靶标的实时检测(图17)。

[0550] 方法

[0551] C2c2基因座的克隆和用于表达的蛋白质

[0552] 对于细菌体内效率测定, 订购来自Leptotrichia wadei F0279和Leptotrichia shahii的C2c2蛋白作为用于哺乳动物表达的密码子优化的基因(金斯瑞, 中国江苏) 并与在 β -内酰胺酶靶向或非靶向间隔子的侧翼的相应的直接重复序列一起克隆到pACYC184主链中。间隔子表达由J23119启动子驱动。

[0553] 对于蛋白质纯化, 哺乳动物密码子优化的C2c2蛋白被克隆到用于蛋白质纯化的细菌表达载体(6 \times His/Twin Strep SUMO, 作为来自Ilya Finkelstein的礼物接受的基于pET的表达载体)中。

[0554] 细菌体内C2c2效率测定

[0555] 将LwC2c2和LshC2c2体内效率质粒和先前描述的 β -内酰胺酶质粒(Abudayyeh 2016) 以分别为90ng和25ng共转化到NovaBlue Singles感受态细胞(Millipore)中。转化后, 将细胞稀释液涂在氨苄青霉素和胆固醇LB-琼脂平板上, 并在37 $^{\circ}$ C孵育过夜。第二天对菌落进行计数。

[0556] 核酸靶标和crRNA制备

[0557] 使用KAPA Hifi Hot Start (Kapa Biosystems) PCR扩增核酸靶标, 并使用MinElute凝胶提取试剂盒(Qiagen) 进行凝胶提取和纯化。使用HiScribe T7Quick High Yield RNA Synthesis试剂盒(New England Biolabs) 将纯化的dsDNA与T7聚合酶一起在30 $^{\circ}$ C孵育过夜, 并用MEGAclean Transcription Clean-up试剂盒(Thermo Fisher) 纯化RNA。

[0558] 对于crRNA制备, 订购作为具有附加T7启动子序列的DNA的构建体(Integrated DNA Technologies)。将crRNA DNA退火至短T7引物(终浓度10 μ M), 并使用HiScribe T7Quick High Yield RNA Synthesis Kit(New England Biolabs) 在37 $^{\circ}$ C下将其与T7聚合酶一起孵育过夜。使用RNAXP清洁珠(Beckman Coulter) 以2 \times 比率的珠子对反应体积, 另外补充1.8 \times 的异丙醇(Sigma), 纯化crRNA。

[0559] NASBA等温扩增

[0560] [Pardee 2016]中描述了NASBA反应的细节。对于20 μ L总反应体积, 6.7 μ L反应缓冲液(Life Sciences, NECB-24)、3.3 μ L核苷酸混合物(Life Sciences, NECN-24)、0.5 μ L无核酸酶水、0.4 μ L 12.5 μ M NASBA引物、0.1 μ L RNA酶抑制剂(Roche, 03335402001) 和4 μ L RNA扩增子(或用于阴性对照的水) 在4 $^{\circ}$ C组合并在65 $^{\circ}$ C孵育2分钟, 然后在41 $^{\circ}$ C孵育10分钟。向每个反应中加入5 μ L酶混合物(LifeSciences, NEC-1-24), 并将反应混合物在41 $^{\circ}$ C下孵育2小时。使用的NASBA引物为5' -AATTCTAATACGACTCACTATAGGGGGATCCTCTAGAAATAT GGATT-3' (SEQ ID NO.16) 和5' -CTCGTATGTTGTGTGGAATTGT-3' (SEQ ID NO.17), 下划线部分指示T7启

动子序列。

[0561] 重组酶聚合酶扩增

[0562] RPA引物使用NCBI Primer blast (Ye等, BMC Bioinformatics 13, 134 (2012) 使用默认参数设计, 除了扩增子大小 (100和140nt之间)、引物的熔融温度 (54°C和67°C之间) 和引物大小 (30和35nt之间)。订购引物作为DNA (Integrated DNA Technologies)。

[0563] RPA和RT-RPA反应的运行是分别根据 **TwistAmp® Basic** 或 **TwistAmp® Basic RT (TwistDx)** 指示, 除了280mM MgAc是在输入模板之前加入。除非另有说明, 否则反应在1μL输入下在37°C下进行2小时。

[0564] LwC2c2蛋白纯化

[0565] 将C2c2细菌表达载体转化到Rosetta™2 (DE3) pLysS Singles Competent Cells (Millipore) 中。在Terrific Broth 4生长培养基 (12g/L胰蛋白胨、24g/L酵母提取物、9.4g/L K₂HPO₄、2.2g/L KH₂PO₄, Sigma) (TB) 中培养16mL起始培养物用于接种4L TB, 其在37°C, 300RPM孵育直至OD₆₀₀为0.6。此时, 通过补充IPTG (Sigma) 至终浓度为500μM而诱导蛋白质表达, 并将细胞冷却至18°C 16小时以进行蛋白质表达。然后将细胞在4°C下5200g离心15分钟。收获细胞沉淀并储存在-80°C下用于后续纯化。

[0566] 蛋白质纯化的所有后续步骤都在4°C进行。将细胞沉淀压碎并重悬于补充有蛋白酶抑制剂 (Complete Ultra无EDTA片剂)、溶菌酶和核酸酶 (benzonase) 的裂解缓冲液 (20mM Tris-HCl、500mM NaCl、1mM DTT、pH 8.0) 中, 然后在以下条件进行超声 (Sonifier 450, Branson, Danbury, CT): 100的振幅开启1秒并关闭2秒, 总超声处理时间为10分钟。通过在4°C于10,000g离心1小时来澄清裂解物, 并通过Stericup 0.22微米过滤器 (EMD Millipore) 过滤上清液。将过滤的上清液应用于Strep Tactin Sepharose (GE) 并旋转孵育1小时, 然后在裂解缓冲液中洗涤蛋白质结合的Strep Tactin树脂3次。将树脂与250单位SUMO蛋白酶 (Thermo Fisher) 一起重悬于SUMO消化缓冲液 (30mM Tris-HCl、500mM NaCl 1mM DTT、0.15% Igepal (NP-40), pH 8.0) 中, 并在4°C下旋转孵育过夜。通过SDS-PAGE和Commassie Blue染色确认消化, 并通过旋下树脂来分离蛋白质洗脱液。通过FPLC (AKTA PURE, GE Healthcare Life Sciences) 将蛋白质加载到5mL HiTrap SP HP阳离子交换柱 (GE Healthcare Life Sciences) 上, 并在洗脱缓冲液 (20mM Tris-HCl, 1mM DTT, 5%甘油, pH 8.0) 中通过130mM至2M NaCl的盐梯度洗脱。通过SDS-PAGE测试所得级分中LwC2c2的存在, 合并含有蛋白质的级分, 并通过Centrifugal Filter Unit在S200缓冲液 (10mM HEPES、1M NaCl、5mM MgCl₂、2mM DTT、1mM DTT, pH 7.0) 中浓缩至1mL。通过FPLC将浓缩的蛋白质加载到凝胶过滤柱 (**Superdex® 200 Increase 10/300GL**, GE Healthcare Life Sciences) 上。通过SDS-PAGE分析来自凝胶过滤的所得级分, 合并含有LwC2c2的级分并将缓冲液交换到储存缓冲液 (600mM NaCl、50mM Tris-HCl、pH 7.5、5%甘油、2mM DTT) 中并在-80°C冷冻储存。

[0567] LwC2c2并行检测

[0568] 检测测定法采用在核酸酶测定缓冲液 (40mM Tris-HCl、60mM NaCl、6mM MgCl₂, pH 7.3) (除非另有说明) 中的45nM纯化的LwC2c2、22.5nM crRNA、125nM底物报告子 (Thermo Scientific RNA酶警报v2)、2μL鼠RNA酶抑制剂、100ng的背景总RNA, 和不同量的输入核酸

靶标进行。如果输入是包含来自RPA反应的T7启动子的扩增的DNA,则上述C2c2反应被改变为包含1mM ATP、1mM GTP、1mM UTP、1mM CTP和0.6 μ L T7聚合酶混合物(NEB)。使反应在荧光板读数器(BioTek)上在37 $^{\circ}$ C下进行1-3小时(除非另有说明),每5分钟测量荧光动力学。

[0569] 组合RPA-DNA扩增、DNA至RNA的T7聚合酶转化和C2c2检测的一锅反应通过集成上述具有RPA扩增混合物的反应条件进行。简而言之,在由0.48 μ M正向引物、0.48 μ M反向引物、1 \times RPA再水合缓冲液、不同量的DNA输入、45nM LwC2c2重组蛋白、22.5nM crRNA、250ng背景总RNA、200nM底物报告子(RNA酶警报v2)、4 μ L RNA酶抑制剂、2mM ATP、2mM GTP、2mM UTP、2mM CTP、1 μ L T7聚合酶混合物、5mM MgCl₂和14mM MgAc组成的50 μ L一锅测定中。

[0570] 使用TaqMan探针进行定量PCR(qPCR)分析

[0571] 为了比较SHERLOCK定量与其他已建立的方法,对ssDNA 1的稀释系列进行qPCR。针对ssDNA 1设计TaqMan探针和引物组(下面的序列)并用IDT合成。使用TaqMan Fast Advanced Master Mix(Thermo Fisher)进行测定,并在Roche LightCycler 480上测量。

[0572] qPCR引物/探针序列。

[0573]

名称	序列
正向引物	GTG GAA TTG TGA GCG GAT AAA C (SEQ ID NO: 420)
反向引物	AAC AGC AAT CTA CTC GAC CTG (SEQ ID NO: 421)
TaqMan 探针	/56-FAM/AGGAAACAG/ZEN/CTATGACCATGATTACGCC/3IABkFQ/ (SEQ ID NO: 422 和 423)

[0574] 使用SYBR Green II的实时RPA

[0575] 为了比较SHERLOCK量化与其他已建立的方法,申请人对ssDNA 1的稀释系列进行RPA。为了实时定量DNA的累积,申请人将1 \times SYBR Green II(Thermo Fisher)加入如上描述的典型RPA反应混合物中,其提供与核酸的量相关的荧光信号。使反应在37 $^{\circ}$ C下在荧光板读数器(BioTek)上进行1小时,每5分钟测量荧光动力学。

[0576] 慢病毒制备和加工

[0577] 慢病毒制备和加工是基于先前已知的方法。简言之,使用HeBS-CaCl₂方法将包含寨卡或登革热RNA片段的10 μ g pSB700衍生物、7.5 μ g psPAX2和2.5 μ g pMD2.G转染至HEK293FT细胞(Life Technologies,R7007)。在更换培养基(补充有10%FBS、1%青霉素-链霉素和4mM GlutaMAX(ThermoFisher Scientific)的DMEM)后28小时,使用0.45 μ m注射器过滤器过滤上清液。使用ViralBind Lentivirus Purification Kit(Cell Biolabs,VPK-104)和Lenti-X Concentrator(Clontech,631231)从上清液中纯化和制备慢病毒。使用QuickTiter Lentivirus Kit(CellBiolabs,VPK-112)定量病毒浓度。将病毒样品掺入7%人血清(Sigma,H4522),将其加热至95 $^{\circ}$ C并保持2分钟,并用作RPA的输入。

[0578] 寨卡人血清样品的分离和cDNA纯化

[0579] 用AVL缓冲液(Qiagen)失活疑似寨卡阳性人血清或尿液样品,并且通过QIAamp Viral RNA minikit(Qiagen)实现RNA的分离。通过混合随机引物、dNTP和样品RNA,然后在70 $^{\circ}$ C下热变性7分钟,将分离的RNA转化为cDNA。然后用Superscript III(Invitrogen)逆转录变性的RNA,在22-25 $^{\circ}$ C孵育10分钟,50 $^{\circ}$ C孵育45分钟,55 $^{\circ}$ C孵育15分钟,80 $^{\circ}$ C孵育10分钟。然后将cDNA与RNA酶H(New England Biolabs)在37 $^{\circ}$ C下一起孵育20分钟以破坏RNA:cDNA杂

交体中的RNA。

[0580] 从人类唾液中提取基因组DNA

[0581] 2mL唾液是从在收集前30分钟限制消耗食物或饮料的志愿者收集。然后如试剂盒方案推荐的,使用QIAamp® DNA Blood Mini Kit (Qiagen) 处理样品。对于煮沸的唾液样品,将400µL磷酸盐缓冲盐水 (Sigma) 加入到100µL志愿者唾液中并以1800g离心5分钟。轻轻倒出上清液,并将沉淀重悬于含有0.2% Triton X-100 (Sigma) 的磷酸盐缓冲盐水中,然后在95°C孵育5分钟。将1µL样品用作RPA反应的直接输入。

[0582] 冷冻干燥和纸沉积

[0583] 将玻璃纤维滤纸 (Whatman, 1827-021) 高压灭菌90分钟 (Consolidated Stills and Sterilizers, MKII), 并在5%无核酸酶BSA (EMD Millipore, 126609-10GM) 中封闭过夜。用无核酸酶水 (Life technologies, AM9932) 将纸张润洗一次后,将它们与4% RNasecure™ (Life technologies, AM7006) 在60°C下一起孵育20分钟,并用无核酸酶水再润洗三次。在使用之前,将处理过的纸在80°C下在热板 (Cole-Parmer, IKAC-MagHS7) 上干燥20分钟。将如前所述的1.8µL C2c2反应混合物放在置于黑色透明底384孔板 (Corning, 3544) 中的盘 (2mm) 上。对于冷冻干燥试验,将含有反应混合物盘的板在液氮中快速冷冻,并如Pardee等 (2) 所述冷冻干燥过夜。将RPA样品在无核酸酶水中1:10稀释,并将1.8µL混合物上样到纸盘上,并使用读板器 (BioTek Neo) 在37°C下孵育。

[0584] 细菌基因组DNA提取

[0585] 对于涉及CRE检测的实验,细菌培养物在溶原性肉汤 (LB) 中生长至对数中期,然后沉淀并使用Qiagen DNeasy Blood and Tissue Kit进行gDNA提取和纯化,视情况而定使用制造商用于革兰氏阴性或革兰氏阳性细菌的方案。在Qubit荧光计上通过Quant-It dsDNA测定法定量gDNA,并通过Nanodrop分光光度计上的200-300nm吸光度光谱评估其量。

[0586] 对于在大肠杆菌和铜绿假单胞菌之间作区分的实验,将细菌培养物在Luria-Bertani (LB) 肉汤中生长至早期稳定期。使用便携式PureLyse bacteria gDNA Extraction Kit (Claremont BioSolutions) 处理1.0mL的大肠杆菌和铜绿假单胞菌两者。将1X结合缓冲液加入细菌培养物中,然后通过电池供电的裂解盒3分钟。将在水中的0.5×结合缓冲液用作洗涤溶液,然后用150µL水洗脱。

[0587] 数字液滴PCR定量

[0588] 为了确认在图1C中使用的ssDNA 1和ssRNA 1标准稀释的浓度,申请人进行数字液滴PCR (ddPCR)。对于DNA定量,使用ddPCR Supermix for Probes (无dUTP) 采用设计用于靶向ssDNA 1序列的PrimeTime qPCR探针/引物测定来制备液滴。对于RNA定量,使用用于探针的一步RT-ddPCR试剂盒采用设计用于靶向ssRNA 1序列的PrimeTime qPCR探针/引物测定来制备液滴。在任一情况下使用QX200液滴发生器 (BioRad) 产生液滴并转移至PCR板。如试剂盒方案中所述,在热循环仪上进行基于液滴的扩增,随后通过在QX200液滴读数器上的测量来测定核酸浓度。

[0589] 用于人类基因分型的合成标准品

[0590] 为了创造用于人类样品基因型的准确调用的标准品,申请人围绕SNP靶标设计引物以扩增代表两个纯合子基因型中的每一个的来自人类基因组DNA的~200bp区域。然后通过以1:1的比率混合纯合标准品来制备杂合标准品。然后将这些标准品稀释至相等的基因

组浓度(~0.56fg/μL),并与真实人类样品一起用作SHERLOCK的输入。

[0591] 肿瘤突变无细胞DNA(cfDNA)的检测

[0592] 模拟实际患者cfDNA样品的模拟cfDNA标准品购自商业供应商(Horizon Discovery Group)。这些标准品作为BRAF V600E和EGFR L858R突变体两者的四种等位基因分数(100%WT和0.1%、1%和5%突变体)提供。提供3μL的这些标准品作为SHERLOCK的输入。

[0593] 荧光数据的分析

[0594] 为了计算减去背景的荧光数据,减去样品的初始荧光以允许在不同条件之间比较。从样品减去背景条件(无输入或无crRNA条件)的荧光以产生减去背景的荧光。

[0595] 用于SNP或菌株区分的指导比率是通过将每个指导除以指导值的总和来计算,以调整样品对样品的整体变化。计算SNP或菌株区分的crRNA比率以调整样品对样品的总体变化,如下:

$$[0596] \quad crRNA A_i \text{ 比率} = \frac{(m+n)A_i}{\sum_{i=1}^m A_i + \sum_{i=1}^n B_i}$$

[0597] 其中,对于给定个体,A_i和B_i分别是指crRNA感测等位基因A或等位基因B的技术复制i的SHERLOCK强度值。由于测定通常具有每个crRNA四个技术重复,因此m和n等于4,并且分母等于给定SNP基因座和个体的所有八个crRNA SHERLOCK强度值的总和。因为有两个crRNA,所以个体的各crRNA的crRNA比率平均将一直总和为2。因此,在纯合性的理想情况下,阳性等位基因crRNA的平均crRNA比率将为2,并且阴性等位基因crRNA的平均crRNA比率将为零。在杂合性的理想情况下,两种crRNA中的每一种的平均crRNA比率将是1。

[0598] LwCas13a切割要求的表征。

[0599] 原型间隔子侧翼位点(PFS)是Cas13a的稳健核糖核酸酶活性所需的靶位点附近存在的特定基序。PFS位于靶位点的3'端,之前被我们的小组对LshCas13a表征为H(非G)(1)。尽管该基序类似于原型间隔子邻近基序(PAM)(DNA靶向2类系统的序列限制),但它在功能上不同,因为它不参与阻止内源系统中CRISPR基因座的自靶向。Cas13a的未来结构研究可能阐明PFS对于Cas13a:crRNA靶复合物形成和切割活性的重要性。

[0600] 申请人纯化来自大肠杆菌的重组LwCas13a蛋白(图2D-E)并测定其以每个可能的间隔子侧翼位点(PFS)核苷酸(A,U,C或G)切割173-nt ssRNA的能力(图2F)。与LshCas13a类似,LwCas13a可以稳健地切割具有A、U或C PFS的靶标,对具有G PFS的ssRNA具有较少的活性。虽然观察到针对具有G PFS的ssRNA1的较弱活性,但申请人仍然看到对具有G PFS基序的两个靶位点的稳健检测(表3;rs601338crRNA和寨卡靶向crRNA 2)。很可能H PFS不是在所有情况下都需要的,并且在许多情况下,强效切割或并行活性可以使用G PFS实现。

[0601] 重组酶聚合酶扩增(RPA)和其他等温扩增策略的讨论。

[0602] 重组酶聚合酶扩增(RPA)是由三种必需酶组成的等温扩增技术:重组酶、单链DNA结合蛋白(SSB)和链置换聚合酶。RPA通过不需要温度调节而克服了其他扩增策略,特别是聚合酶链反应(PCR)中存在的许多技术困难,因为酶都在约37°C的恒定温度下起作用。RPA用由体内DNA复制和修复启发的酶促方法代替了用于双链模板的全局熔融和引物退火的温度循环。重组酶-引物复合物扫描双链DNA并促进互补位点处的链交换。链交换通过SSB稳定,允许引物保持结合。重组酶的自发解体以其ADP结合状态发生,允许链置换聚合酶侵入

并延伸引物,允许扩增而无需在护理点和野外设置中不可得的复杂仪器。在37-42°C的温度范围内循环重复该过程导致指数DNA扩增。公布的原始制剂使用枯草芽孢杆菌Pol I (Bsu)作为链置换聚合酶,T4uvsX作为重组酶,和T4gp32作为单链DNA结合蛋白(2),尽管不清楚本研究中使用的由TwistDx销售的当前制剂中是哪些成分。

[0603] 另外,RPA具有许多限制:

[0604] 1) 尽管Cas13a检测是定量的(图15),但实时RPA定量可以困难,因为当重组酶使用所有可用的ATP时其的快速饱和。虽然实时PCR是定量的,但由于循环扩增的能力,RPA没有机制来严格控制扩增速率。可以进行某些调整以降低扩增速率,例如降低可用的镁或引物浓度,降低反应温度,或设计低效的引物。尽管观察到定量SHERLOCK的一些实例,例如图31、32和52,但情况并非总是如此,并且可以取决于模板。

[0605] 2) RPA效率可以对引物设计敏感。制造商通常建议设计更长的引物以确保平均GC含量(40-60%)下的有效重组酶结合,并筛选多达100个引物对以找到高度敏感的引物对。申请人已发现使用SHERLOCK,只需要设计两个引物对以实现具有单分子灵敏度的阿摩尔测试。这种稳健性可能是由于抵消了扩增子扩增中的任何低效率的组成型活性Cas13a并行活性对信号的额外扩增。这种特性对于我们在图34中的细菌病原体鉴定特别重要。在扩增高度结构化区域(例如细菌基因组中的16S rRNA基因位点)时遇到了问题,因为RPA中没有涉及熔融步骤。因此,引物中的二级结构变成问题,限制了扩增效率并因此限制了灵敏度。尽管存在这些RPA特定的问题,但由于来自Cas13a的额外信号扩增,本文所公开的实施方式被认为是成功的。

[0606] 3) 对于有效的RPA,扩增序列长度必须短(100-200bp)。对于大多数应用,这不是一个重要问题,可能甚至是有利的(例如,cfDNA检测,其中平均片段大小为160bp)。有时大的扩增子长度是重要的,例如当期望通用引物用于细菌检测并且用于区分的SNP分布在大区域时。

[0607] SHERLOCK的模块化允许任何扩增技术,甚至非等温方法,在T7转录和Cas13a检测之前使用。这种模块性是通过在单个反应中T7和Cas13a步骤的相容性使成为可能,允许对具有T7启动子的任何扩增的DNA输入进行检测。在使用RPA之前,尝试将基于核酸序列的扩增(NASBA)(3、4)用于我们的检测测定(图10)。然而,NASBA没有大幅改善Cas13a的灵敏度(图11和53)。检测前可采用的其他扩增技术包括PCR、环介导等温扩增(LAMP)(5)、链置换扩增(SDA)(6)、解旋酶依赖性扩增(HDA)(7)和切口酶扩增反应(NEAR)(8)。与任何等温技术交换的能力允许SHERLOCK克服任何一种扩增技术的特定限制。

[0608] 工程化错配的设计。

[0609] 申请人证实,LshCas13a靶标切割在靶标:crRNA双链体中存在两个或更多个错配时减少,但相对不受单个错配影响,这是申请人对LwCas13a并行切割确认的观察(图36A)。申请人假设,通过在crRNA间隔子序列中引入另外的突变,申请人会使针对具有另外的错配(总共两个错配)的靶标的并行切割不稳定,同时保留中靶并行切割,因为会仅存在单个错配。为了测试工程化增加的特异性的可能性,申请人设计了靶向ssRNA 1的多个crRNA并且包含在crRNA的长度中的错配(图36A)以优化中靶并行切割并最小化相差单个错配的靶标的并行切割。申请人观察到这些错配没有减少ssRNA 1的并行切割,但显著降低了包含另外的错配的靶标(ssRNA 2)的信号。在ssRNA 1和2之间最好地区分的设计的crRNA包含接近

ssRNA 2错配的合成错配,实际上在杂交RNA中产生“泡”或变形。由靶标中存在的合成错配和另外的错配的协调(即,双错配)引起的灵敏度损失与LshCas13a和LwCas13a对连续或附近的双错配的灵敏度一致,并且为合理设计能够实现单核苷酸区分的crRNA提供了基础(图36B)。

[0610] 对于ZIKV和DENV株的错配检测,我们的全长crRNA含有两个错配(图37A、B)。由于株之间的高序列差异,申请人不能够找到在两个基因组之间仅具有单个核苷酸差异的28nt连续段。然而,申请人预测较短crRNA仍然具有功能性,并设计针对两个ZIKV株中的靶标的、包含间隔子序列中的合成错配和靶序列中的仅一个错配的较短23nt crRNA。这些crRNA仍然可以区分ZIKV的非洲和美洲株(图36C)。对23nt和20nt crRNA进行的后续测试表明,间隔子长度的减少降低了活性,但维持或增强了区分单个错配的能力(图57A-G)。为了更好地理解如何可以引入合成错配以促进单核苷酸突变区分,申请人在三种不同间隔子长度:28、23和20nt的前七个位置中铺设(tile)合成错配(图57A)。在在第三位置处具有突变的靶标上,当合成错配位于间隔子的第5位时,LwCas13a显示最大特异性,在较短间隔子长度具有改善的特异性,尽管中靶活性水平较低(图57B-G)。申请人还将靶标突变移位到第3-6位,并且在突变周围的间隔子中铺设合成错配(图58)。

[0611] 使用合成标准品用SHERLOCK进行基因分型。

[0612] 从SNP基因座的PCR扩增产生的合成标准品的评估允许基因型的准确识别(图60A、B)。通过计算个体的样品的SHERLOCK结果与合成标准品之间的所有比较(ANOVA),可以通过找到具有最相似的SHERLOCK检测强度的合成标准品来识别每个个体的基因型(图60C、D)。这种SHERLOCK基因分型方法可通用于任何SNP基因座(图60E)。

[0613] SHERLOCK是可承受的、适应性的CRISPR-Dx平台。

[0614] 对于SHERLOCK的成本分析,省略了确定为具有可忽略不计的成本的试剂,包括用于合成crRNA的DNA模板、RPA中使用的引物、常见缓冲液(MgCl₂、Tris HCl、甘油、NaCl、DTT)、玻璃微纤维滤纸,和RNasecure试剂。对于DNA模板,来自IDT的超聚合成(ultramersynthesis)提供40份体外转录反应的材料(每个足以用于~10,000个反应),为约70美元,这为crRNA合成增加了可忽略不计的成本。对于RPA引物,可以以约10美元购买30nt DNA引物的25nmole IDT合成,提供足以用于5000次SHERLOCK反应的材料。玻璃微纤维纸可以0.50美元/张获得,其足以用于数百次SHERLOCK反应。4%RNasecure试剂的成本为7.20美元/mL,其足以用于500次测试。

[0615] 此外,对于所有实验,除基于纸的测定外,使用384孔板(Corning3544),成本为0.036美元/反应。由于成本可以忽略不计,这不包括在整体成本分析中。此外,SHERLOCK-POC不需要使用塑料容器,因为它可以容易地在纸上进行。本文使用的SHERLOCK的读出方法是配备有过滤器组或单色器的读板器。作为资本投资,读板器的成本不包括在计算中,因为随着越来越多的反应在仪器上进行,成本急剧下降,并且可以忽略不计。对于POC应用,可以使用更便宜和便携的替代品,例如手持式分光光度计(9)或便携式电子读数器(4),其将仪器的成本降低至<200美元。虽然与体积较大的仪器相比,这些更便携的解决方案将降低读出的速度和易用性,但它们允许更广泛的使用。

[0616] 结果

[0617] 本文描述的测定和系统通常可以包括扩增和检测的两步法。在第一步期间,扩增

核酸样品 (RNA或DNA), 例如通过等温扩增。在第二步期间, 将扩增的DNA转录成RNA, 随后与CRISPR效应子 (例如C2c2) 和编程以检测靶核酸序列的存在的crRNA一起孵育。为了能够进行检测, 将已经用猝灭荧光团标记的报告子RNA添加到反应中。报告子RNA的并行切割导致荧光团的去猝灭, 并允许实时检测核酸靶标 (图17A)。

[0618] 为了实现稳健的信号检测, 从生物 *Leptotrichia wadei* (LwC2c2) 鉴定C2c2的直系同源物并评估。LwC2c2蛋白的活性通过在大肠杆菌中与合成CRISPR阵列一起表达并编程其以切割导致细菌在氨苄青霉素选择下死亡的 β -内酰胺酶mRNA内的靶位点来评估 (图2B)。对LwC2c2基因座观察到比LshC2c2基因座更少的存活的大肠杆菌菌落, 证实LwC2c2直系同源物的更高切割活性 (图2C)。然后从大肠杆菌中纯化人类密码子优化的LwC2c2蛋白 (图2D-E), 并且用不同原型间隔子侧翼位点 (PFS) 核苷酸测定其切割173-nt ssRNA的能力 (图2F)。LwC2c2能够切割可能的四个PFS靶标中的每一个, 对具有G PFS的ssRNA的活性略微更低。

[0619] 使用可商购RNA荧光读板器测量LwC2c2RNA酶并行活性的实时测量 (图17A)。为了确定LwC2c2活性的基线灵敏度, 将LwC2c2与ssRNA靶标1 (ssRNA 1) 和与ssRNA靶标内的位点互补的crRNA以及RNA传感器探针一起孵育 (图18)。这产生了 ~ 50 fM的灵敏度 (图27A), 其尽管比其他近期核酸检测技术 (Pardee等, 2014) 更灵敏, 但对于需要亚飞摩尔检测性能的许多诊断应用 (Barletta等, 2004; Emmadi等, 2011; Rissin等, 2010; Song等, 2013) 不够灵敏。

[0620] 为了增加灵敏度, 等温扩增步骤在LwC2c2孵育之前增加。将LwC2c2介导的检测与先前使用的等温扩增方法如基于核酸序列的扩增 (NASBA) (Compton, 1991; Pardee等, 2016) 结合, 在一定程度上改善了灵敏度 (图11)。测试了可用于在两小时内指数扩增DNA的替代等温扩增方法, 即重组酶聚合酶扩增 (RPA) (Piepenburg等, 2006)。通过在RPA引物上添加T7RNA聚合酶启动子, 可以将扩增的DNA转化为RNA用于随后通过LwC2c2进行检测 (图17)。因此, 在某些示例性实施方式中, 测定包括通过RPA进行的扩增、DNA至RNA的T7RNA聚合酶转变, 以及通过C2c2从猝灭报告子解锁荧光的后续RNA检测的组合。

[0621] 使用ssRNA 1的合成DNA版本的实例方法, 可以在每个反应1-10个分子的范围内实现阿摩尔灵敏度 (图27B, 左)。为了验证检测的准确性, 用数字液滴PCR定量输入DNA的浓度, 并确认最低可检测靶标浓度 ($2\mu\text{M}$) 是每微升单个分子的浓度。通过增加逆转录步骤, RPA还可以将RNA扩增为dsDNA形式, 允许我们对ssRNA 1实现阿摩尔灵敏度 (27B, 右)。类似地, 通过数字液滴PCR确认RNA靶标的浓度。为了评估实例方法作为POC诊断测试起作用的可行性, 对所有组分在单一反应中起作用的能力 (RPA、T7聚合酶扩增和LwC2c2检测) 进行测试, 并发现采用该测定的一锅法版本的阿摩尔灵敏度 (图22)。

[0622] 该测定能够在液体中或纸上进行灵敏的病毒检测

[0623] 接下来确定该测定在需要高灵敏度并且可以获益于便携式诊断的传染病应用中是否是有效的。为了测试在模型系统中的检测, 产生了包含寨卡病毒基因组和相关黄病毒登革热的RNA片段的慢病毒 (Dejnirattisai等, 2016), 并且定量病毒颗粒的数量 (图31A)。检测到模拟病毒的水平低至 2aM 。同时, 也可以在含有寨卡和登革热RNA片段的这些代理病毒 (proxy virus) 之间显示出明显的区别 (图31B)。为了确定该测定是否与冷冻干燥相容以去除对用于分配的冷链的依赖性, 将反应组分冷冻干燥。在使用样品以使冻干组分再水化后, 检测到 20fM 的ssRNA 1 (图33A)。由于资源贫乏和POC环境会获益于纸质测试以便于使用, 还评估了C2c2检测在玻璃纤维纸上的活性, 并发现纸点C2c2反应能够进行靶标检测 (图

33B)。组合在一起,冷冻干燥和纸点C2c2检测反应导致ssRNA 1的灵敏检测(图33C)。相似水平的灵敏度也在溶液中的LwC2c2和冷冻干燥的LwC2c2之间对合成寨卡病毒RNA片段的检测观察到,证实了冷冻干燥SHERLOCK的稳健性和快速POC寨卡病毒诊断的潜力(图33D-E)。为此,测试了该测定的POC变体的能力以确定区分寨卡RNA与登革热RNA的能力(图31C)。虽然纸点和冻干略微降低了读出的绝对信号,但与具有登革热对照序列的模拟病毒的检测相比,该测定在浓度低至20aM时仍然显著地检测到模拟寨卡病毒(图31D)。

[0624] 在人类中的寨卡病毒RNA水平已报道为低至在患者唾液中为 3×10^6 拷贝/mL (4.9fM)和在患者血清中为 7.2×10^5 拷贝/mL (1.2fM) (Barzon等,2016;Gourinat等,2015;Lanciotti等,2008)。从获得的患者样品中,观察到浓度低至 1.25×10^3 拷贝/mL (2.1aM)。为了评估该测定是否能够进行低滴度临床分离物的寨卡病毒检测,从患者中提取病毒RNA并进行逆转录,并将得到的cDNA用作测定的输入(图32A)。在低至1.25拷贝/ μ L (2.1aM)的浓度下观察到对寨卡人血清样品的显著检测(图32B)。此外,来自患者样品的信号预测寨卡病毒RNA拷贝数,并可用于预测病毒负载(图31F)。为了测试对不能进行核酸纯化的疾病情况的广泛适用性,测试掺入人血清中的ssRNA 1的检测,并确定该测定在血清水平低于2%时有活性(图33G)。

[0625] 细菌病原体区别和基因区别

[0626] 为了确定该测定是否可用于区分细菌病原体,选择16S V3区域作为初始靶标,因为保守侧翼区域允许通用RPA引物用于细菌物种之间,并且可变内部区域允许物种的区别。设计了一组5种可能的靶向crRNA用于致病菌株和分离的大肠杆菌和铜绿假单胞菌gDNA(图34A)。该测定能够区分大肠杆菌或铜绿假单胞菌gDNA,并显示其他物种的crRNA的低背景信号(图34A、B)。

[0627] 该测定还可以适于快速检测和区分感兴趣的细菌基因,例如抗生素抗性基因。耐碳青霉烯肠杆菌(CRE)是重大的新兴公共卫生挑战(Gupta等,2011)。评估了该测定检测碳青霉烯抗性基因的能力,以及该测试是否可以区分不同的碳青霉烯抗性基因。从含有肺炎克雷伯氏菌(*Klebsiella pneumoniae*)碳青霉烯酶(KPC)或新德里金属 β -内酰胺酶1(NDM-1)抗性基因的临床分离物获得肺炎克雷伯氏菌(*Klebsiella pneumoniae*)并且设计crRNA来区分基因。所有CRE对缺乏这些抗性基因的细菌具有显著信号(图35A),并且我们可以显著区分KPC和NDM-1抗性株(图35B)。

[0628] CRISPR RNA指导的RNA酶的单碱基错配特异性

[0629] 已经显示某些CRISPR RNA指导的RNA酶直系同源物,例如LshC2c2,不容易区分单碱基错配(Abudayyeh等,2016)。如本文所证实的,LwC2c2也共享该特征(图37A)。为了增加LwC2c2切割的特异性,开发了用于在crRNA:靶标双链体中引入合成错配的系统,其增加了对错配的总灵敏度并且使得单碱基错配灵敏度成为可能。设计用于靶标1的多个crRNA,并且其包含在crRNA的长度中的错配(图37A)以优化中靶切割并最小化相差单个错配的靶标的切割。这些错配不降低ssRNA靶标1的切割效率,但显著降低包含另外的错配的靶标(ssRNA靶标2)的信号。在靶标1和2之间最好地区分的设计的crRNA包含接近靶标2错配的合成错配,实际上产生“泡”。由在靶标中存在的合成错配和另外的错配(即,双错配)的协调导致的灵敏度丧失与LshC2c2对连续或附近双错配的灵敏度(Abudayyeh等,2016)一致,并提出了用于合理设计使得能够实现单核苷酸区别的crRNA的安排(图37B)。

[0630] 已经证实C2c2可以被工程化以识别单碱基错配,确定该工程化特异性是否可用于区分密切相关的病毒病原体。设计多个crRNA以检测寨卡病毒的非洲或美洲株(图37A)和登革热病毒的株1或3(图37C)。这些crRNA包含在间隔子序列中的合成错配,由于合成错配而导致在与中靶株形成双链时形成单个泡(bubble)。然而,当合成错配间隔子与脱靶株形成双链时,由于合成错配和SNP错配而形成两个泡。合成错配crRNA以显著高于脱靶株的信号检测到它们相应的株,允许稳健的株区分(图37B、37D)。由于株之间显著的序列相似性,不可能找到两个基因组之间仅具有单个核苷酸差异的28nt连续段以证明真正的单核苷酸株区别。然而,预测较短crRNA仍可起作用,因为它们伴随LshC2c2(they are with LshC2c2)(Abudayyeh等,2016),并且因此针对两个寨卡株中的靶标设计了包含在间隔子序列中的合成错配和在靶序列中的仅一个错配的较短23-nt crRNA。这些crRNA仍然能够以高灵敏度区分寨卡的非洲和美洲株(图36C)。

[0631] 使用从唾液纯化的DNA进行快速基因分型

[0632] 从人唾液快速基因分型可用于紧急药物基因组学情况或用于家庭诊断。为了证实本文所公开的实施方式用于基因分型的潜力,使用23andMe基因分型数据作为金标准(Eriksson等,2010),选择五个基因座以设置C2c2检测基准(图38A)。这五个基因座跨越了广泛的功能关联,包括对药物例如他汀类或对乙酰氨基酚的敏感性,诺如病毒易感性,和心脏病风险(表7)。

[0633] 表7:测试的SNP变体

ID	基因	类别
rs5082	APOA2	饱和脂肪消耗和体重增加
rs1467558	CD44	对乙酰氨基酚代谢
[0634] rs2952768	靠近 CREB1	吗啡依赖
rs4363657	SLCO1B1	他汀类药物使用者的肌病风险增加 4.5 倍
rs601338	FUT2	对诺如病毒的抗性

[0635] 收集来自四个人受试者的唾液,并在不到一小时内使用简单的商业试剂盒纯化基因组DNA。这四个受试者在五个基因座中具有多样化基因型组,为对该测定用于基因分型设置基准提供足够宽的样品空间。对于五个SNP基因座中的每一个,使用RPA与适当的引物扩增受试者的基因组DNA,然后用LwC2c2和设计用于特异性检测两个可能等位基因之一的crRNA对进行检测(图38B)。该测定的特异性足以以高显著性区分等位基因并推断纯合子和杂合子基因型两者。因为在检测之前对唾液进行了DNA提取方案,测试该测定以确定是否可以通过使用加热至95°C 5分钟的唾液而不进一步提取以使其甚至更适用于POC基因分型。该测定能够正确地对两名患者进行基因分型,所述患者的唾液仅经受加热5分钟,然后进行后续扩增和C2c2检测(图40B)。

[0636] 检测在低等位基因分数下的cfDNA中的癌突变

[0637] 由于该测定对在靶标中的单核苷酸差异具有高度特异性,设计测试以确定该测定

是否足够灵敏以检测无细胞DNA (cfDNA) 中的癌症突变。cfDNA片段是野生型cfDNA片段的小百分比(0.1%至5%) (Bettegowda等,2014;Newman等,2014;Olmedillas Lopez等,2016;Qin等,2016)。cfDNA领域中的重大挑战是检测这些突变,因为鉴于血液中的背景中发现的高水平的非突变DNA,它们通常难以发现(Bettegowda等,2014;Newman等,2014;Qin等,2016)。POC cfDNA癌症测试也可用于定期筛查癌症存在,特别是对于处于缓解风险的患者。

[0638] 该测定检测野生型背景中的突变DNA的能力通过在具有在crRNA靶向位点中的单个突变的ssDNA1的背景中稀释dsDNA靶标1来确定(图41A-B)。LwC2c2能够感测低至背景dsDNA的0.1%和在dsDNA 1的阿摩尔浓度内的水平的dsDNA 1。该结果显示,背景突变体dsDNA 1的LwC2c2切割足够低以允许对中靶dsDNA在0.1%等位基因分数下的稳健检测。在低于0.1%的水平下,背景活性可能是问题,阻止对正确靶标的任何进一步显著检测。

[0639] 由于该测定法可以感测具有在临床相关范围中的等位基因分数的合成靶标,评估该测定是否能够在cfDNA中检测癌症突变。设计了针对两种不同癌症突变EGFR L858R和BRAF V600E的RPA引物,并且使用类似于实际人cfDNA样品的具有5%、1%和0.1%的等位基因分数的商业cfDNA标准品进行测试。使用可以区分突变等位基因与野生型等位基因的crRNA对(图38C),实现了对0.1%等位基因分数的这两个突变基因座的检测(图39A-B)。

[0640] 讨论

[0641] 通过将C2c2的自然特性与等温扩增和淬灭荧光探针结合,本文所公开的测定和系统已被证实为检测RNA和DNA的通用的、稳健的方法,并且适用于各种快速诊断,包括传染性应用和快速基因分型。本文所公开的测定和系统的主要优点是可以在几天内重新设计和合成新的POC测试,低至0.6美元/测试。

[0642] 由于许多人类疾病应用需要检测单错配的能力,开发了合理方法以通过在crRNA的间隔子序列中引入合成错配来工程化crRNA以对靶序列中的单错配具有高度特异性。用于使用CRISPR效应子实现特异性的其他方法依赖于在数十种指导设计上的基于筛选的方法(Chavez等,2016)。使用设计的错配crRNA,证实了在相差单错配的位点对寨卡和登革热病毒株的区分,来自人唾液gDNA的SNP的快速基因分型,以及cfDNA样品中癌症突变的检测。

[0643] 测定平台的低成本和适应性将自身提供给进一步的应用,包括(i) 替代特定qPCR测定,例如Taqman的一般RNA/DNA定量经验,(ii) 类似于微阵列的快速、多重RNA表达检测,和(iii) 其他灵敏度检测应用,例如检测食品中来自其他来源的核酸污染。此外,C2c2可以潜在地用于检测生物环境内的转录物,例如细胞中,并且鉴于C2c2检测的高度特异性,追踪转录物的等位基因特异性表达或活细胞中的疾病相关突变可以是可能的。随着适体的广泛可用性,通过将用适体检测蛋白质与揭示RPA的隐蔽扩增位点、然后进行C2c2检测相结合以感测蛋白质也可能是可能的。

[0644] 使用CRISPR-Cas13a/C2c2的核酸检测:阿摩尔灵敏度和单核苷酸特异性

[0645] 为了实现稳健的信号检测,申请人从Leptotrichia wadei鉴定了Cas13a的直系同源物(LwCas13a),其表现出相比于Leptotrichia shahii Cas13a(LshCas13a)(10)更高的RNA指导的RNA酶活性(图2,也参见上文“LwCas13a切割要求的表征”)。LwCas13a与ssRNA靶标1(ssRNA 1)、crRNA和报告子(淬灭荧光RNA)一起孵育(图18)(13),产生了~50fM的检测灵敏度(图51、15),这对于许多诊断应用不够灵敏(12、14-16)。申请人因此探索了将基于

Cas13a的检测与不同的等温扩增步骤相结合(图10、11、53、16)(17、18)。在所探索的方法中,重组酶聚合酶扩增(RPA)(18)提供了最大的灵敏度,并且可以与T7转录相结合以将扩增的DNA转化为RNA用于通过LwCas13a进行后续检测(也参见上文“重组酶聚合酶扩增(RPA)和其他等温扩增策略的讨论”)。申请人将这种通过RPA进行扩增、扩增的DNA到RNA的T7RNA聚合酶转录,以及通过Cas13a并行RNA切割介导的报告子信号释放对靶RNA的检测的组合称为SHERLOCK。

[0646] 申请人首先确定了SHERLOCK用于检测RNA(当与逆转录结合时)或DNA靶标的灵敏度。申请人实现了RNA和DNA两者的单分子灵敏度,如通过数字液滴PCR(ddPCR)验证的(图27,51,54A、B)。当所有SHERLOCK组分在单一反应中组合时,保持了阿摩尔灵敏度,证实该平台作为护理点(POC)诊断的可行性(图54C)。SHERLOCK具有与两种已建立的灵敏核酸检测方法ddPCR和定量PCR(qPCR)相似水平的灵敏度,而单独的RPA对于检测低水平的靶标不够灵敏(图55A-D)。而且,SHERLOCK显示出比ddPCR、qPCR和RPA更小的变异,如通过重复之间的变异系数所测量的(图55E-F)。

[0647] 申请人接下来检查了SHERLOCK是否在需要高灵敏度的传染病应用中有效。申请人产生了携带寨卡病毒(ZIKV)或相关黄病毒登革热(DENV)的基因组片段的慢病毒(19)(图31A)。SHERLOCK检测到低至2aM的病毒颗粒,并且可以区分ZIKV和DENV(图31B)。为了探索SHERLOCK在该领域的潜在用途,申请人首次证实,冻干并随后再水合的Cas13acrRNA复合物(20)可以检测20fM的非扩增ssRNA 1(图33A),并且靶标检测在玻璃纤维纸上也是可能的(图33B)。SHERLOCK的其他组分也适合冷冻干燥:RPA作为冻干试剂在环境温度下提供,并且申请人先前证实T7聚合酶耐受冷冻干燥(2)。组合在一起,冷冻干燥和纸点Cas13a检测反应产生了与水溶液反应水平相当的ssRNA 1的灵敏度检测(图33C-E)。尽管纸点和冻干略微降低了读出的绝对信号,但SHERLOCK(图31C)可以在低至20aM的浓度下容易地检测到模拟ZIKV病毒(图31D)。SHERLOCK还能够检测临床分离物(血清,尿液或唾液)中的ZIKV,其中滴度可低至 2×10^3 拷贝/mL(3.2aM)(21)。从患者血清或尿液样品中提取并反转录成cDNA(图32E和52A)的ZIKV RNA可以在低至 1.25×10^3 拷贝/mL(2.1aM)的浓度下检测到,如通过qPCR验证的(图32F和52B)。此外,来自患者样品的信号预测ZIKV RNA拷贝数,并可用于预测病毒负载(图33F)。为了模拟没有核酸纯化的样品检测,申请人测量了掺入人血清中的ssRNA 1的检测,并且发现Cas13a可以检测含有多达2%的血清的反应中的RNA(图33G)。本文所公开的实施方式的另一个重要的流行病学应用是细菌病原体的鉴定和特定细菌基因的检测。申请人靶向16S rRNA基因V3区域,其中保守侧翼区域允许通用RPA引物在细菌物种之间使用,并且可变内部区域允许物种的区别。在针对不同致病菌株和从大肠杆菌和铜绿假单胞菌分离的gDNA的一组五种可能的靶向crRNA中(图34A),SHERLOCK对菌株进行正确的基因分型,并显示低交叉反应性(图34B)。此外,申请人能够使用SHERLOCK区分具有两种不同的抗性基因:肺炎克雷伯氏菌碳青霉烯酶(KPC)和新德里金属- β -内酰胺酶1(NDM-1)(22)的肺炎克雷伯氏菌的临床分离株(图56)。

[0648] 为了增加SHERLOCK的特异性,申请人在crRNA:靶双链体中引入合成错配,其使得LwCas13a能够区分相差单碱基错配的靶标(FIG36A、B;也参见上文“工程化错配的设计”)。申请人设计了在间隔子序列中具有合成错配的多个crRNA以检测ZIKV的非洲或美洲株(图37A)和DENV的株1或3(图37C)。合成错配crRNA以显著高于(双尾Student t-检验; $p < 0.01$)

脱靶株的信号检测到它们相应的株,允许基于单错配的稳健的株区分(图37B、D,36C)。进一步的表征揭示Cas13a检测当突变位于间隔子的位置3并且合成错配位于位置5时,实现最大特异性,同时保持中靶灵敏度(图57和58)。检测单碱基差异的能力开启了使用SHERLOCK进行快速人类基因分型的机会。申请人选择了跨越一系列健康相关单核苷酸多态性(SNP)的五个基因座(表1),并使用23andMe基因分型数据作为在这些SNP处的金标准(23)对SHERLOCK检测设置基准(图38A)。申请人从感兴趣的基因座中具有多样化基因型的四个人受试者收集唾液,并通过商业柱纯化或直接加热五分钟提取基因组DNA(20)。SHERLOCK以高显著性和足够特异性区分等位基因以推断纯合和杂合基因型两者(图38B、40、59、60;还参见上文“使用合成标准品用SHERLOCK进行基因分型”)。最后,申请人试图确定SHERLOCK是否可以检测无细胞(cf)DNA片段中的低频癌症突变,这因为患者血液中的高水平野生型DNA而具有挑战性(24-26)。申请人首先发现SHERLOCK可以检测到稀释到基因组DNA背景中的阿摩尔浓度的ssDNA 1(图61)。接下来,申请人发现SHERLOCK还能够检测到水平低至背景DNA的0.1%(其在临床相关范围内)的含有单核苷酸多态性(SNP)的等位基因(图41A、B)。申请人随后证实了SHERLOCK可以在具有低至0.1%的等位基因分数的模拟cfDNA样品中检测到两种不同的癌症突变,即EGFR L858R和BRAF V600E(图38、39)(20)。

[0649] SHERLOCK平台将自身提供给进一步的应用,包括:(i)代替特异性qPCR测定,例如TaqMan的一般RNA/DNA定量,(ii)快速、多重RNA表达检测,和(iii)其他灵敏检测应用,例如检测核酸污染。此外,Cas13a可以潜在地检测生物环境内的转录物并跟踪转录物的等位基因特异性表达或活细胞中的疾病相关突变。SHERLOCK是检测RNA和DNA的通用的、稳健的方法,适用于快速诊断,包括传染病应用和灵敏基因分型。SHERLOCK纸测试可以有信心在几天内重新设计和合成,低至0.61美元/测试(参见上文“SHERLOCK是可承受的、适应性的CRISPR-Dx平台”),因为测试的几乎所有crRNA都产生高灵敏度和特异性。这些特性突出了CRISPR-Dx的强大功能,为快速、稳健和灵敏的生物分子检测开辟了新的道路。

[0650] 表8:使用的RPA引物

[0651]

名称	序列	第 1 图
RP0683 - RPA ssDNA/ssRNA 1 F	(SEQ. I.D. No. 18)	图 27B
RP0684 - RPA ssDNA/ssRNA 1 R	(SEQ. I.D. No. 19)	图 27B
AMPL-25 Zika 8B long-rpa3-f	(SEQ. I.D. No. 20)	图 31B
AMPL-26 Zika 8B long-rpa3-r	(SEQ. I.D. No. 21)	图 31B
RP819 - zika 区域 8 F	(SEQ. I.D. No. 22)	图 31C
RP821 - zika 区域 8 R	(SEQ. I.D. No. 23)	图 31C
517 细菌 V3 F	(SEQ. I.D. No. 24)	图 34B
RP758 细菌 V3 R	(SEQ. I.D. No. 25)	图 34B
wR0074 A2 rs5082 F	(SEQ. I.D. No. 26)	图 38B

[0652]

wR0074 E2 rs5082 R	AA (SEQ. I.D. No. 27)	图 38B
wR0074 A4 rs1467558 F	(SEQ. I.D. No. 28)	图 38B
wR0074 E4 rs1467558 R	(SEQ. I.D. No. 29)	图 38B
wR0074 A5 rs2952768 F	(SEQ. I.D. No. 30)	图 38B
wR0074 E5 rs2952768 R	(SEQ. I.D. No. 31)	图 38B
wR0074 A9 rs4363657 F	(SEQ. I.D. No. 32)	图 38B
wR0074 E9 rs4363657 R	(SEQ. I.D. No. 33)	图 38B
wR0074 A11 rs601338 F	(SEQ. I.D. No. 34)	图 38B
wR0074 E11 rs601338 R	(SEQ. I.D. No. 35)	图 38B
RP824 BRAFV600E cfDNA F	(SEQ. I.D. No. 36)	图 39A
RP769 BRAFV600E cfDNA R	(SEQ. I.D. No. 37)	图 39A
RP826 EGFR858R cfDNA F	(SEQ. I.D. No. 38)	图 39B
RP804 EGFR858R cfDNA R	(SEQ. I.D. No. 39)	图 39B
AMPL-31 T1-nasba1-f	(SEQ. I.D. No. 40)	图 11
AMPL-32 T1-nasba1-r	(SEQ. I.D. No. 41)	图 11
AMPL-33 T1-nasba2-f	(SEQ. I.D. No. 42)	图 11
AMPL-34 T1-nasba2-r	(SEQ. I.D. No. 43)	图 11
AMPL-35 T1-nasba3-f	(SEQ. I.D. No. 44)	图 11
AMPL-36 T1-nasba3-r	(SEQ. I.D. No. 45)	图 11
wR0075 A1 KPC F	(SEQ. I.D. No. 46)	图 35A
wR0075 B1 KPC R	(SEQ. I.D. No. 47)	图 35A
wR0075 A3 NDM F	(SEQ. I.D. No. 48)	图 35A
wR0075 B3 NDM R	(SEQ. I.D. No. 49)	图 35A

[0653] 表9:使用的crRNA序列

[0654]

名称	完整 crRNA 序列	间隔子序列	第 1 图	PFS
靶标 1 crRNA	(SEQ. I.D. No. 50)	(SEQ. I.D. No. 51)	图 2F	C
Zika 靶向 crRNA 1	(SEQ. I.D. No. 52)	(SEQ. I.D. No. 53)	图 31A	U
Zika 靶向 crRNA 2	(SEQ. I.D. No. 54)	(SEQ. I.D. No. 55)	图 33D	G
大肠杆菌检测 crRNA	(SEQ. I.D. No. 56)	(SEQ. I.D. No. 57)	图 22B	U
肺炎克雷伯氏菌 检测 crRNA	(SEQ. I.D. No. 58)	(SEQ. I.D. No. 59)	图 34B	U
铜绿假单胞菌检 测 crRNA	(SEQ. I.D. No. 60)	(SEQ. I.D. No. 61)	图 34B	U

[0655]

结核杆菌检测 crRNA	(SEQ. I.D. No. 62)	(SEQ. I.D. No. 63)	图 34B	U
金黄色葡萄球菌 检测 crRNA	(SEQ. I.D. No. 64)	(SEQ. I.D. No. 65)	图 34B	G
KPC crRNA	(SEQ. I.D. No. 66)	(SEQ. I.D. No. 67)	图 35A	U
NDM crRNA	(SEQ. I.D. No. 68)	(SEQ. I.D. No. 69)	图 35A	C
错配 crRNA 1	(SEQ. I.D. No. 70)	(SEQ. I.D. No. 71)	图 36A	C
错配 crRNA 2	(SEQ. I.D. No. 72)	(SEQ. I.D. No. 73)	图 36A	C
错配 crRNA 3	(SEQ. I.D. No. 74)	(SEQ. I.D. No. 75)	图 36A	C
错配 crRNA 4	(SEQ. I.D. No. 76)	(SEQ. I.D. No. 77)	图 36A	C
错配 crRNA 5	(SEQ. I.D. No. 78)	(SEQ. I.D. No. 79)	图 36A	C
错配 crRNA 6	(SEQ. I.D. No. 80)	(SEQ. I.D. No. 81)	图 36A	C
错配 crRNA 7	(SEQ. I.D. No. 82)	(SEQ. I.D. No. 83)	图 36A	C
错配 crRNA 8	(SEQ. I.D. No. 84)	(SEQ. I.D. No. 85)	图 36A	C
错配 crRNA 9	(SEQ. I.D. No. 86)	(SEQ. I.D. No. 87)	图 36A	C
错配 crRNA 10	(SEQ. I.D. No. 88)	(SEQ. I.D. No. 89)	图 36A	C
非洲 crRNA 1	(SEQ. I.D. No. 90)	(SEQ. I.D. No. 91)	图 38A	C
非洲 crRNA 2	(SEQ. I.D. No. 92)	(SEQ. I.D. No. 93)	图 38A	C
美洲 crRNA 1	(SEQ. I.D. No. 94)	(SEQ. I.D. No. 95)	图 38A	U
美洲 crRNA 2	(SEQ. I.D. No. 96)	(SEQ. I.D. No. 97)	图 38A	U
登革热株 3 crRNA 1	(SEQ. I.D. No. 98)	(SEQ. I.D. No. 99)	图 38C	A
登革热株 3 crRNA 2	(SEQ. I.D. No. 100)	(SEQ. I.D. No. 101)	图 38C	A
登革热株 1 crRNA 1	(SEQ. I.D. No. 102)	(SEQ. I.D. No. 103)	图 38C	A
登革热株 1 crRNA 2	(SEQ. I.D. No. 104)	(SEQ. I.D. No. 105)	图 38C	A
较短非洲 crRNA 1	(SEQ. I.D. No. 106)	(SEQ. I.D. No. 107)	图 36C	C
较短非洲 crRNA 2	(SEQ. I.D. No. 108)	(SEQ. I.D. No. 109)	图 36C	C
较短美洲 crRNA 1	(SEQ. I.D. No. 110)	(SEQ. I.D. No. 111)	图 36C	U
较短美洲 crRNA 2	(SEQ. I.D. No. 112)	(SEQ. I.D. No. 113)	图 36C	U
rs1467558 crRNA C	(SEQ. I.D. No. 114)	(SEQ. I.D. No. 115)	图 38B	C
rs1467558 crRNA T	(SEQ. I.D. No. 116)	(SEQ. I.D. No. 117)	图 38B	C

[0656]

rs2952768 crRNA C	(SEQ. I.D. No. 118)	(SEQ. I.D. No. 119)	图 38B	A
rs2952768 crRNA T	(SEQ. I.D. No. 120)	(SEQ. I.D. No. 121)	图 38B	A
rs4363657 crRNA C	(SEQ. I.D. No. 122)	(SEQ. I.D. No. 123)	图 38B	A
rs4363657 crRNA T	(SEQ. I.D. No. 124)	(SEQ. I.D. No. 125)	图 38B	A
rs601338 crRNA A	(SEQ. I.D. No. 126)	(SEQ. I.D. No. 127)	图 38B	G
rs601338 crRNA G	(SEQ. I.D. No. 128)	(SEQ. I.D. No. 129)	图 38B	G
rs5082 crRNA G	(SEQ. I.D. No. 130)	(SEQ. I.D. No. 131)	图 40A	A
rs5082 crRNA A	(SEQ. I.D. No. 132)			A
EGFR L858R 野 生型 crRNA	(SEQ. I.D. No. 134)	(SEQ. I.D. No. 135)	图 38C	C
EGFR L858R 突 变体 crRNA	(SEQ. I.D. No. 136)	(SEQ. I.D. No. 137)	图 38C	C
BRAF V600E 野 生型 crRNA	(SEQ. I.D. No. 138)	(SEQ. I.D. No. 139)	图 38C	A
BRAF V600E 突 变体 crRNA	(SEQ. I.D. No. 140)	(SEQ. I.D. No. 141)	图 38C	A
23 nt 错配 crRNA 1	(SEQ. I.D. No. 303)	(SEQ. I.D. No. 304)	图 57D	C
23 nt 错配 crRNA 2	(SEQ. I.D. No. 305)	(SEQ. I.D. No. 306)	图 57D	C
23 nt 错配 crRNA 4	(SEQ. I.D. No.307)	(SEQ. I.D. No. 308)	图 57D	C
23 nt 错配 crRNA 5	(SEQ. I.D. No. 234)	(SEQ. I.D. No. 235)	图 57D	C
23 nt 错配 crRNA 6	(SEQ. I.D. No. 236)	(SEQ. I.D. No. 237)	图 57D	C
23 nt 错配 crRNA 7	(SEQ. I.D. No. 238)	(SEQ. I.D. No. 239)	图 57D	C
20 nt 错配 crRNA 1	(SEQ. I.D. No. 240)	(SEQ. I.D. No. 241)	图 57F	C
20 nt 错配 crRNA 2	(SEQ. I.D. No. 242)	(SEQ. I.D. No. 243)	图 57F	C
20 nt 错配 crRNA 4	(SEQ. I.D. No. 244)	(SEQ. I.D. No. 245)	图 57F	C
20 nt 错配 crRNA 5	(SEQ. I.D. No. 246)	(SEQ. I.D. No. 247)	图 57F	C

[0657]

20 nt 错配 crRNA 6	(SEQ. I.D. No. 248)	(SEQ. I.D. No. 249)	图 57F	C
20 nt 错配 crRNA 7	(SEQ. I.D. No. 250)	(SEQ. I.D. No. 251)	图 57F	C
靶标错配 4 错配 crRNA 1	(SEQ. I.D. No. 252)	(SEQ. I.D. No. 253)	图 58B	C
靶标错配 4 错配 crRNA 2	(SEQ. I.D. No. 254)	(SEQ. I.D. No. 255)	图 58B	C
靶标错配 4 错配 crRNA 3	(SEQ. I.D. No. 256)	(SEQ. I.D. No. 257)	图 58B	C
靶标错配 4 错配 crRNA 5	(SEQ. I.D. No. 258)	(SEQ. I.D. No. 259)	图 58B	C
靶标错配 4 错配 crRNA 6	SEQ. I.D. No. 260)	(SEQ. I.D. No. 261)	图 58B	C
靶标错配 4 错配 crRNA 7	(SEQ. I.D. No. 262)	(SEQ. I.D. No. 263)	图 58B	C
靶标错配 5 错配 crRNA 2	(SEQ. I.D. No. 264)	(SEQ. I.D. No. 265)	图 58B	C
靶标错配 5 错配 crRNA 3	(SEQ. I.D. No. 266)	(SEQ. I.D. No. 267)	图 58B	C
靶标错配 5 错配 crRNA 4	(SEQ. I.D. No. 268)	(SEQ. I.D. No. 269)	图 58B	C
靶标错配 5 错配 crRNA 6	(SEQ. I.D. No. 270)	(SEQ. I.D. No. 271)	图 58B	C
靶标错配 5 错配 crRNA 7	(SEQ. I.D. No. 272)	(SEQ. I.D. No. 273)	图 58B	C
靶标错配 5 错配 crRNA 8	(SEQ. I.D. No. 274)	(SEQ. I.D. No. 275)	图 58B	C
靶标错配 6 错配 crRNA 3	(SEQ. I.D. No. 276)	(SEQ. I.D. No. 277)	图 58B	C
靶标错配 6 错配 crRNA 4	(SEQ. I.D. No. 278)	(SEQ. I.D. No. 279)	图 58B	C
靶标错配 6 错配 crRNA 5	(SEQ. I.D. No. 280)	(SEQ. I.D. No. 281)	图 58B	C
靶标错配 6 错配 crRNA 7	(SEQ. I.D. No. 282)	(SEQ. I.D. No. 283)	图 58B	C
靶标错配 6 错配 crRNA 8	(SEQ. I.D. No. 284)	(SEQ. I.D. No. 285)	图 58B	C
靶标错配 6 错配 crRNA 9	(SEQ. I.D. No. 286)	(SEQ. I.D. No. 287)	图 58B	C

[0658] 表10:本实施例中使用的RNA和DNA靶标

[0659]

名称	序列	第 1 图
ssRNA 1 (C PFS)	(SEQ. I.D. No. 288)	图 2F
ssRNA 1 (G PFS)	(SEQ. I.D. No. 289)	图 2F
ssRNA 1 (A PFS)	(SEQ. I.D. No. 290)	图 2F
ssRNA 1 (U PFS)	(SEQ. I.D. No. 291)	图 2F
ssDNA 1	(SEQ. I.D. No. 292)	图 27
DNA 2	(SEQ. I.D. No. 293)	图 54B
在慢病毒中的 ZIKV	(SEQ. I.D. No. 294)	图 31B
在慢病毒中的 DENV	(SEQ. I.D. No. 295)	图 31B
合成 ZIKV 靶标	(SEQ. I.D. No. 296)	图 33D
合成非洲 ZIKV 靶标	(SEQ. I.D. No. 297)	图 37A
合成美洲 ZIKV 靶标	(SEQ. I.D. No. 298)	图 37A
合成登革热株 1 靶标	(SEQ. I.D. No. 299)	图 37C
合成登革热株 3 靶标	(SEQ. I.D. No. 300)	图 37C
ssRNA 2	(SEQ. I.D. No. 301)	图 36A
ssRNA 3	(SEQ. I.D. No. 302)	图 36A

表 11: 本实施例中使用的质粒

质粒名称	描述	质粒图的链接
pC004	β -内酰胺酶筛选靶标	https://benchling.com/s/IPJ1cCwR
pC009	LshCas13a基因座到具有靶向间隔子的pACYC184中	https://benchling.com/s/seqylkMuglYmiG4A3VhShZg
pC010	LshCas13a基因座到具有非靶向间隔子的pACYC184中	https://benchling.com/s/seq-2WApFr3zni1GOACyQY8a
pC011	LwCas13a基因座到具有靶向间隔子的pACYC184中	https://benchling.com/s/seq-Vyk8qK2fyhzezfNgLJHM
pC012	LwCas13a基因座到具有非靶向间隔子的pACYC184中	https://benchling.com/s/seq-RxZAgPBzBUGQThkxR2Kx
pC013	Twinstrep-SUMO-huLwCas13a用于细菌表达	https://benchling.com/s/seq-66CfLwu7sLMQMbcXe7Ih

[0660] 实施例3具有正交碱基偏好的Cas13b直系同源物的表征

[0661] 申请人生物化学地表征了最近定义的RNA指导的RNA靶向酶的VI型CRISPR-Cas13b家族的14个直系同源物来寻找用于改善SHERLOCK检测技术的新候选者(图83A和85)。申请人能够在大肠杆菌中异源表达14个Cas13b直系同源物并且纯化蛋白质用于体外RNA酶活性测定(图86)。因为不同的Cas13直系同源物可能具有不同的碱基偏好以获得最佳的切割活性,申请人生成由5个A、G、C或U组成的荧光RNA酶均聚物传感器以评估正交切割偏好。申请人将每个直系同源物与其靶向合成的173nt ssRNA 1的关联crRNA一起孵育,并使用均聚物荧光传感器测量并行切割活性(图83B和87)。

[0662] 实施例4用文库进行基序发现筛选

[0663] 为了进一步探索各种Cas13a和Cas13b直系同源物的切割偏好的多样性,申请人开发了基于文库的方法以表征响应于并行活性的内切核酸酶活性偏好的基序。申请人使用了侧翼为恒定DNA柄(handle)的简并6聚体RNA报告子,这允许扩增和读出未切割的序列(图83C)。孵育该文库与并行活化的Cas13酶导致可检测的切割并依赖于靶RNA的添加(图88)。对耗尽的基序的测序揭示了在消化时间内的文库的偏态的增加(图89A),指示碱基偏好,并且选择高于阈值比率的序列产生了与酶的切割相对应的富集序列(图89B)。来自富集基序的序列标志再现了对LwaCas13a和CcaCas13b观察到的U偏好和PsmCas13b的A偏好(图89C)。申请人还确定了显示仅一个直系同源物的切割但不显示其他直系同源物的切割的多个序列以允许独立的读出(图89D)。

[0664] 为了理解酶偏好的特定子基序,申请人分析了耗尽的基序的单碱基偏好(图90A),其与测试的均聚物基序一致,以及对于两碱基基序(图83C和90B)。这些两碱基基序揭示了更复杂的偏好,特别是对于LwaCas13a和PsmCas13b,其偏好TA、GA和AT双碱基序列的。更高阶的基序还揭示另外的偏好(图91和92)。

[0665] 申请人利用靶标的体外切割确认了LwaCas13a、PsmCas13b和CcaCas13b的并行偏好(图93)。为了改善PsmCas13b的弱消化,申请人优化了缓冲液组成和酶浓度(图94A、B)。在PsmCas13b和Cas13b直系同源物上测试的其他双阳离子不具有大的效应(图95A-F)。申请人还就两个RNA靶标比较了PsmCas13b与先前表征的A偏好Cas13家族成员,并且发现相当的或改善的灵敏度(图96A、B)。根据这些结果,申请人在具有独立报告子的单独反应中比较了LwaCas13a和PsmCas13b的动力学,并发现两个通道之间的低水平串扰(图83D)。

[0666] 实施例5用LwaCas13a、PsmCas13b和CcaCas13b进行单分子检测

[0667] SHERLOCK技术的关键特征是它使得能够通过LwaCas13a并行RNA酶活性进行单分子检测(2aM或1分子/ μ L)。为了表征Cas13b酶的灵敏度,申请人用PsmCas13b和CcaCas13b进行SHERLOCK,CcaCas13b是另一种具有尿苷偏好的高活性Cas13b酶(图83E)。申请人发现LwaCas13a、PsmCas13b和CcaCas13b能够实现两种不同的RNA靶标ssRNA 1和合成寨卡sRNA的2aM检测(图83E;97、和98)。为了调查用这三种酶靶向的稳健性,申请人设计了在ssRNA 1中均匀间隔的11个不同的crRNA,并且发现LwaCas13a最一致地实现信号检测,而CcaCas13b和Psmcas13b两者均显示了在crRNA之间的检测中多得多的变异(图99)。为了鉴定用于检测的最佳crRNA,申请人从34nt到12nt变化PsmCas13b和CcaCas13b的间隔子长度,并且发现PsmCas13b在间隔子长度为30时具有峰值灵敏度,而CcaCas13b在高于28nt间隔子长度时具有等同的灵敏度(图100)。申请人还测试了是否可以推动检测限超过2aM,允许更大的样品体积输入到SHERLOCK中。通过放大预扩增RPA步骤,申请人发现LwaCas13a和PsmCas13b两者都可以为200、20和2zM的输入样品给出显著的检测信号,并允许250 μ L和540 μ L的体积输入。

[0668] 实施例6采用RPA的定量SHERLOCK

[0669] 由于SHERLOCK依赖于指数扩增,因此准确定量核酸可以是困难的。申请人假设降低RPA步骤的效率可以改善在输入量与SHERLOCK反应的信号之间的相关性。申请人观察到,SHERLOCK检测的动力学在一系列样品浓度中对引物浓度非常敏感(图101A-D)。申请人稀释了引物浓度,这增加了信号和定量准确度两者(图83G和101E)。该观察结果可以是由于引物二聚体形成的减少,允许更有效的扩增,同时防止饱和。120nM的引物浓度表现出在信号与

输入之间的最大相关性(图101F)。该准确度在低至阿摩尔范围的大范围浓度中是可持续的(图83H和101G)。

[0670] 实施例7用正交Cas13直系同源物进行双颜色多重

[0671] 核酸诊断的有利特征是同时检测多个样品输入的能力,允许多重检测组或样品内对照。Cas13酶的正交碱基偏好提供具有多重SHERLOCK的机会。申请人可以通过具有不同碱基特性和荧光团颜色的荧光均聚物传感器测定同一反应中的不同Cas13酶的并行活性,使得能够同时测量多个靶标(图84A)。为了证实该概念,申请人设计了针对寨卡病毒ssRNA的LwaCas13a crRNA和针对登革热病毒ssRNA的PsmCas13b crRNA。申请人发现在同一反应中具有这两组Cas13-crRNA复合物的该测定能够鉴定反应中是否存在寨卡或登革热RNA或两者(图84B)。申请人还发现,由于CcaCas13b与PsmCas13b之间的正交偏好,这两种酶也可以用于寨卡和登革热靶标的多重检测(图102)。申请人成功地使能够将这个概念向含有多重RPA引物和Cas13-crRNA复合物两者的整个SHERLOCK反应拓展。申请人设计了针对铜绿假单胞菌的LwaCas13a crRNA和针对金黄色葡萄球菌的PsmCas13b crRNA并且能够检测低至阿摩尔范围的这两个DNA靶标(图84C)。类似地,使用PsmCas13b和LwaCas13a两者,申请人能够使用SHERLOCK实现寨卡和登革热RNA的阿摩尔多重检测(图103)。

[0672] 申请人已经显示LwaCas13a使得能够进行单核苷酸变体检测,并且这可以应用于从人类唾液进行快速基因分型,但检测需要两个单独的反应:每个等位基因感测crRNA一个。为了使得能够进行单反应SHERLOCK基因分型,申请人设计了针对rs601338SNP(与诺如病毒抗性相关的 $\alpha(1,2)$ -岩藻糖基转移酶FUT2基因中的变体)的G等位基因的LwaCas13a crRNA和针对rs601338SNP的A等位基因的PsmCas13b crRNA。使用这种单样品多重方法,申请人能够成功地使用4个不同人类受试者的唾液对他们进行基因分型,以及准确鉴定他们是纯合的还是杂合的。

[0673] 为了进一步展示Cas13酶家族的多功能性,申请人模拟涉及Cas13作为伴随诊断和疗法自身两者的治疗方法。申请人最近开发了PspCas13b用于转录物的可编程RNA编辑,其可用于使用称为用于可编程A至I替换的RNA编辑(REPAIR)的系统,校正遗传疾病中的突变。由于诊断可以在与疗法配对以指导治疗决策或监测治疗结果时非常有用,因此申请人认为SHERLOCK可用于基因分型以指导REPAIR治疗,也可作为对已编辑RNA的读出来追踪疗法的编辑效率(图84E)。申请人选择证实这个治疗诊断概念以校正在家族性腺瘤性息肉病1(涉及大肠和直肠中的癌症的遗传性疾病)中的APC突变(APC:c.1262G>A)。申请人设计了APC基因的健康和突变cDNA并将这些转染到HEK293FT细胞中。申请人能够从这些细胞中收获DNA并使用具有LwaCas13a和PsmCas13b的单样品多重SHERLOCK对正确样品成功地进行基因分型(图84F)。同时,申请人为REPAIR系统设计并克隆了指导RNA,并用指导RNA和dPspCas13b-ADAR2_{dd}(E488Q) REPAIR系统转染具有疾病基因型的细胞。48小时后,申请人收获了RNA,申请人将其分开以输入到SHERLOCK中以检测编辑结果和进行下一代测序(NGS)分析以确认编辑率。测序揭示申请人用REPAIR系统实现了43%编辑(图84G)并且能够用SHERLOCK检测到这个,因为健康感测crRNA显示比非靶向指导对照状态更高的信号,而疾病感测crRNA显示信号降低(图84H和104)。总体而言,用于该测定的试剂的设计和合成花费了3天,基因分型花费了1天,并且使用REPAIR进行校正和感测编辑速率花费了3天,产生了持续仅7天的总体治疗诊断途径。

[0674] 申请人已经证实,使用来自*Leptotrichia wadei*的VI型RNA指导的RNA靶向CRISPR-Cas13a直系同源物对核酸的高度灵敏和特异的检测。申请人已经进一步显示,Cas13b酶家族在生物化学上是具有活性的并且具有使得它们适合于通过SHERLOCK对核酸进行多重检测的独特性质。通过表征Cas13b酶的正交碱基偏好,申请人发现PsmCas13b识别而LwaCas13a不识别的荧光RNA传感器的特定序列。申请人能够利用这些碱基偏好以使得对两个不同靶标进行样品内多重检测成为可能,并显示该特征用于区分病毒株和基因分型个体的实用性。另外,通过预扩增步骤的工程化,可以使SHERLOCK定量,允许对输入核酸浓度的近似或者定量。申请人已经另外显示正交PsmCas13b能够进行单分子检测,并且通过放大体积,申请人可以进行上至~0.5毫升和下至2zM浓度的样品的检测。

[0675] 采用SHERLOCK进行多重检测通过在空间上进行多个反应而是可能的,但凭借正交碱基偏好进行样品内多重允许许多靶标大规模检测且成本更便宜。虽然申请人已经在本文中显示两输入多重,但切割基序筛选使得能够设计另外的正交切割传感器(图90)。LwaCas13a和CcaCas13b都切割相同的尿苷均聚物并因此不是如通过均聚物传感器测量的正交的(图83B),通过基序筛选显示非常独特的切割偏好(图90)。通过筛选另外的Cas13a、Cas13b和Cas13c直系同源物,很可能许多直系同源物将揭示独特的6聚体基序偏好,这可以理论上允许仅受到光谱独特荧光传感器的数量限制的高度多重化SHERLOCK。高度多重化SHERLOCK使得能够进行许多技术应用,特别是涉及复杂输入感测和逻辑运算的那些。

[0676] 基于Cas13的检测对视觉的、更灵敏的和多重的读出的这些另外的改进使得能够进行用于核酸检测的增加的应用,尤其是在便携和无仪器分析有必要的环境中。快速多重基因分型可以为药物基因组学决策提供信息,测试田间多种作物性状,或评估共存病原体的存在。快速、等温读出增加了这种用于电源或便携式读数器不可得,甚至是对于稀有物质如循环DNA的环境的检测的可及性。改进的基于CRISPR的核酸测试使得更容易了解核酸在农业、病原体检测和慢性疾病中的存在。

[0677]

[0678] 在不脱离本发明的范围和精神的情况下,本发明所描述的方法、药物组合物和试剂盒的各种修改和变化对于本领域技术人员而言是显而易见的。尽管已经结合具体实施方式描述了本发明,但是应当理解,它能够作进一步修改,并且要求保护的本发明不应当不适当地受限于这样的特定实施方式。实际上,对于本领域技术人员显而易见的用于实施本发明的所描述的模式的各种修改是旨在落入本发明的范围内。本申请旨在涵盖总体遵循本发明的原理且包括从本公开到本发明所属领域内的已知惯用实践内并且可以应用于本文前文说明的必要特征的这样的偏离的本发明的任何变化、用途或改编。

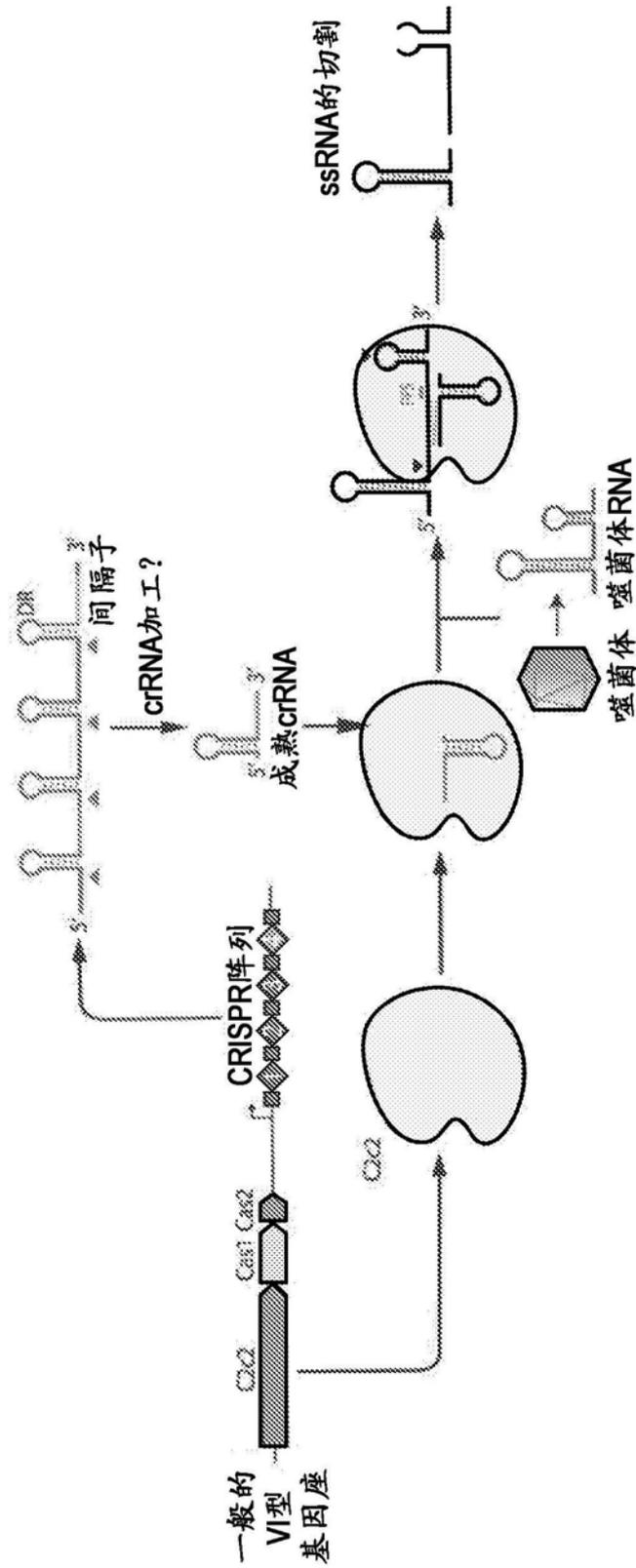


图1

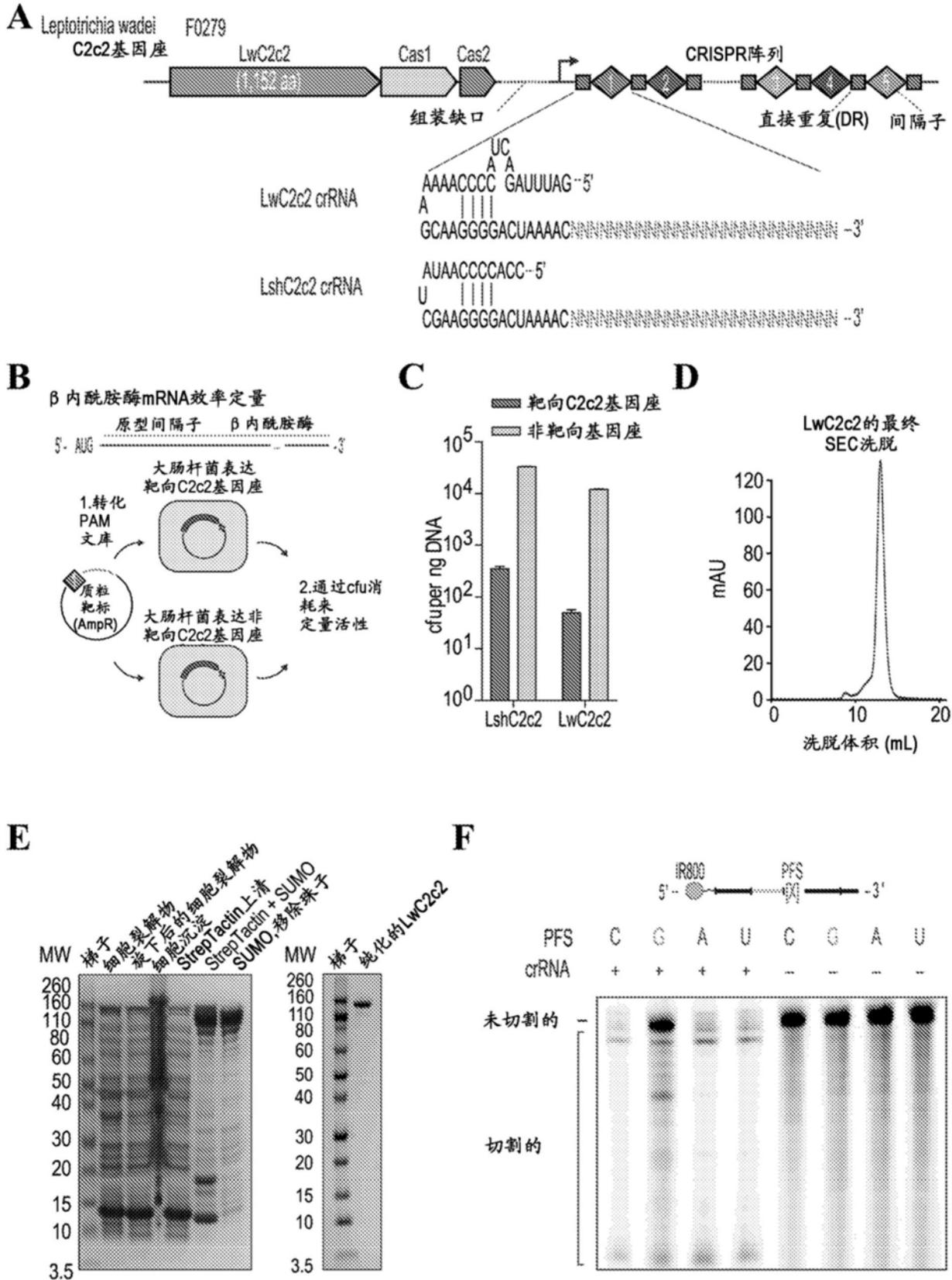


图2

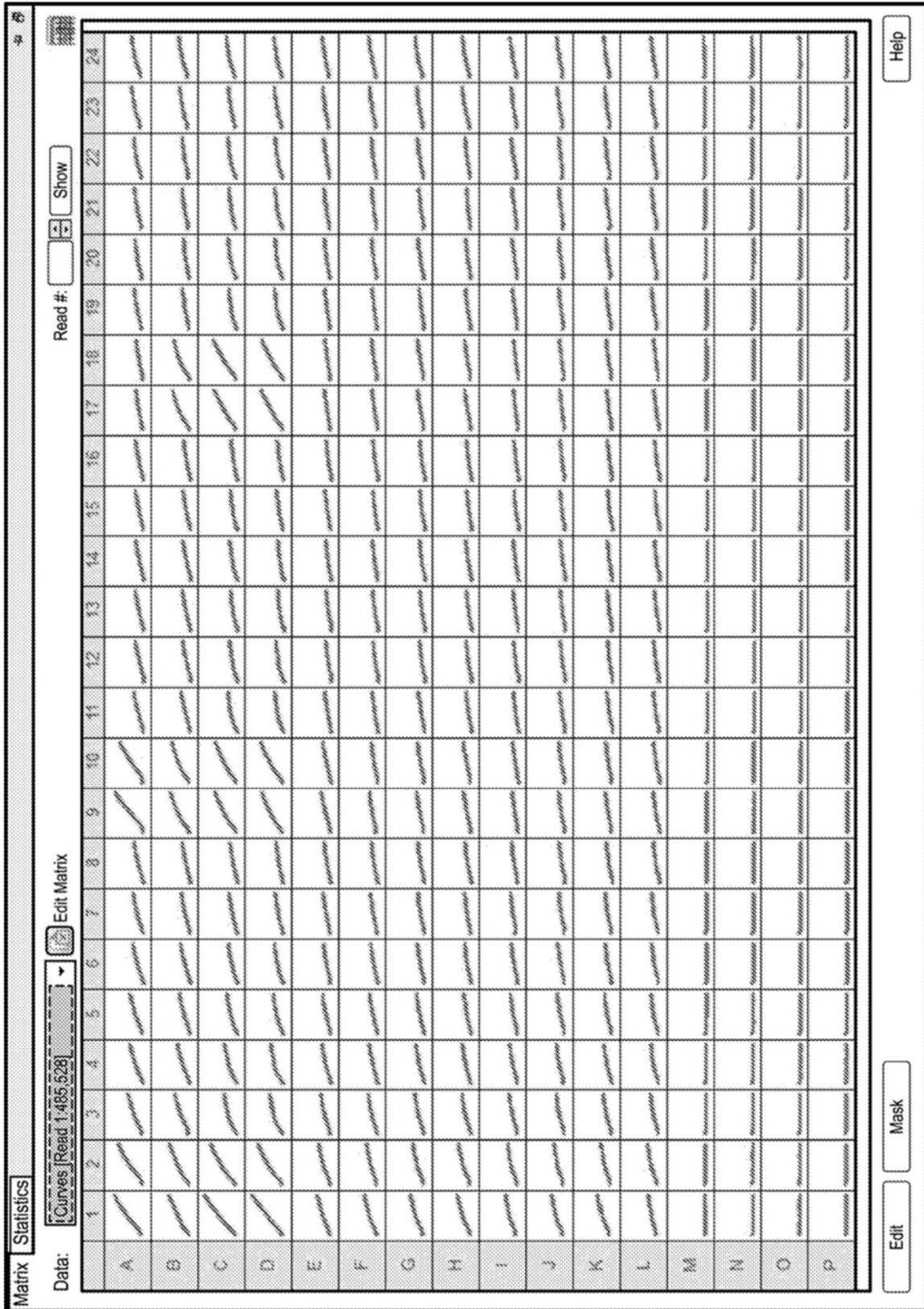


图3

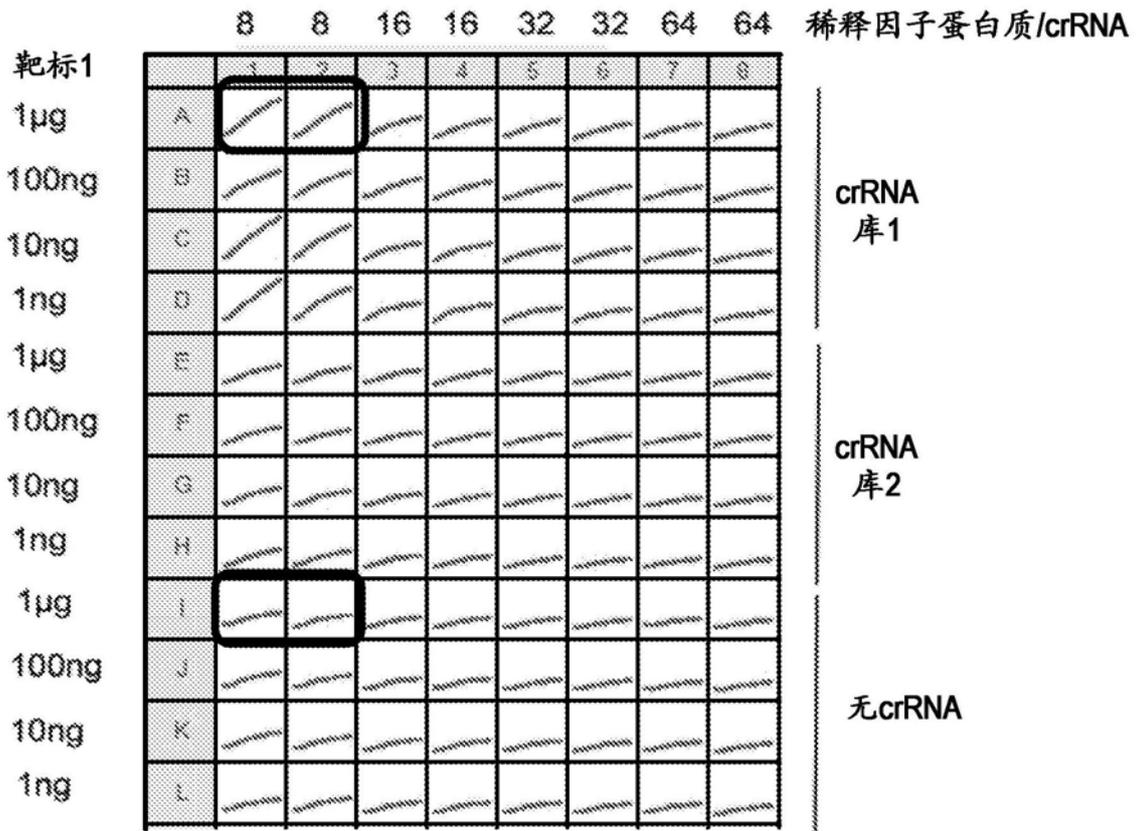


图4

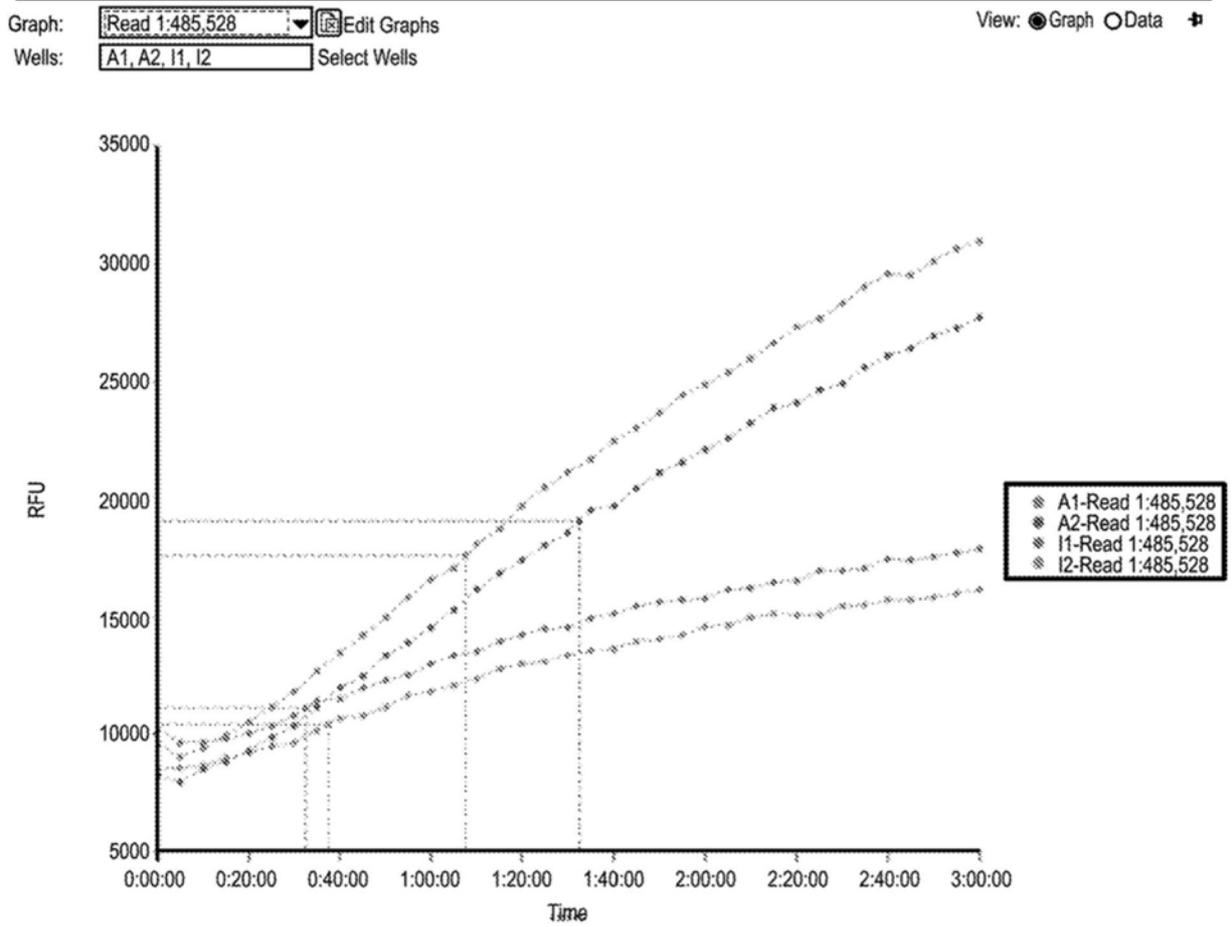


图5

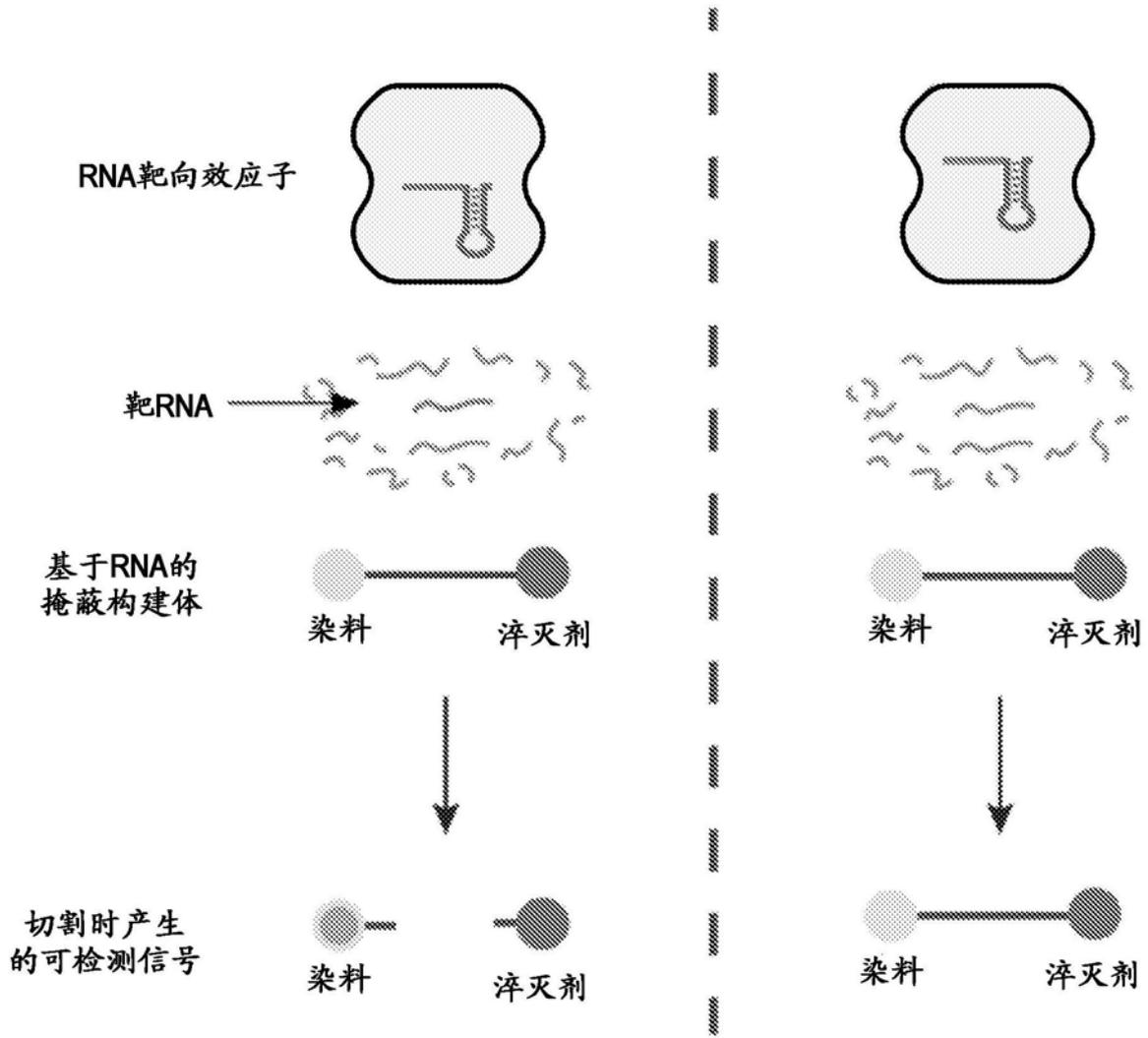


图7

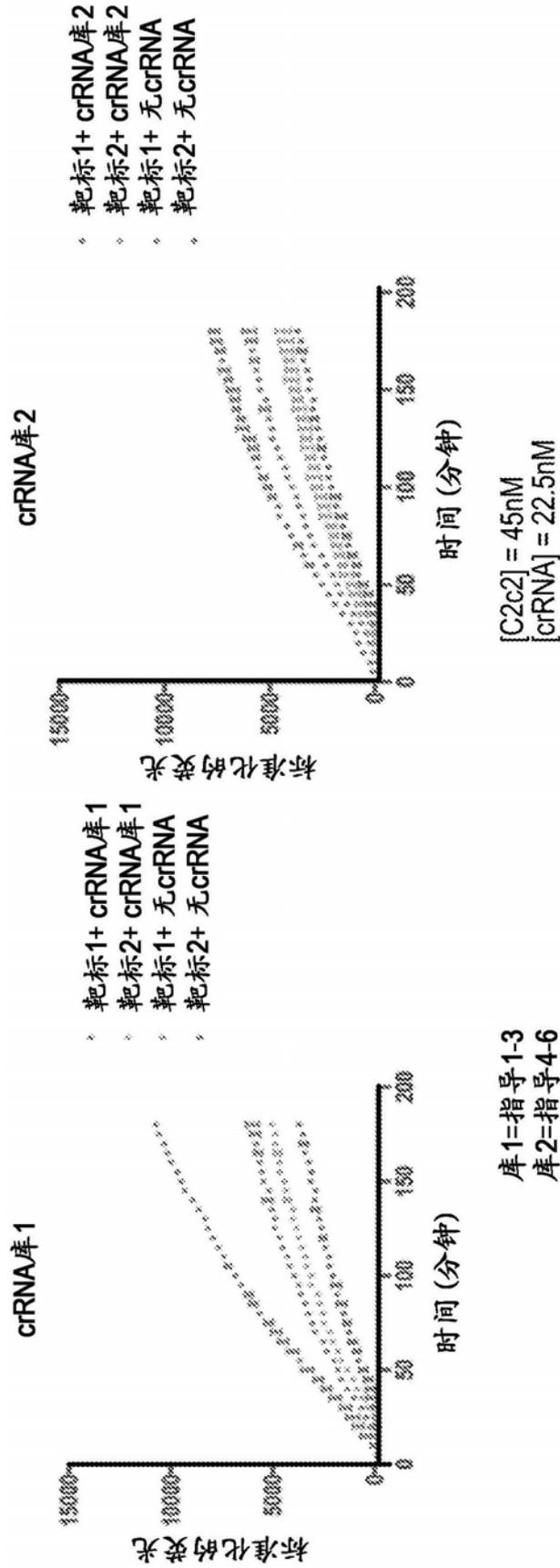


图8

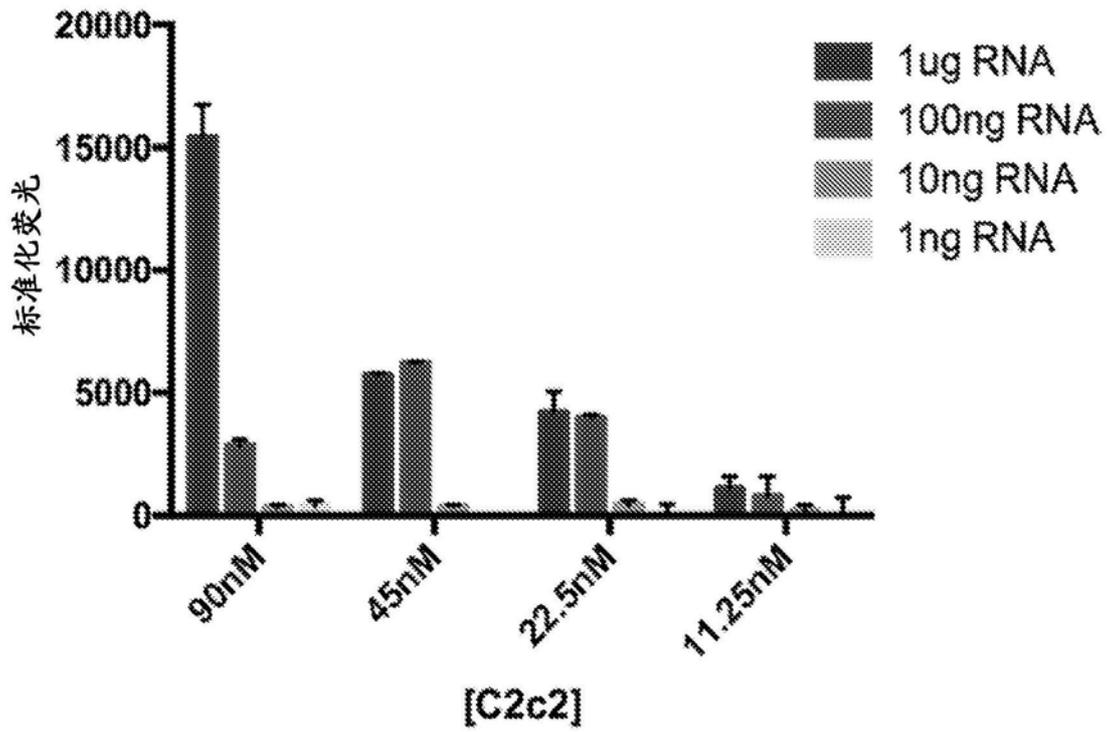


图9

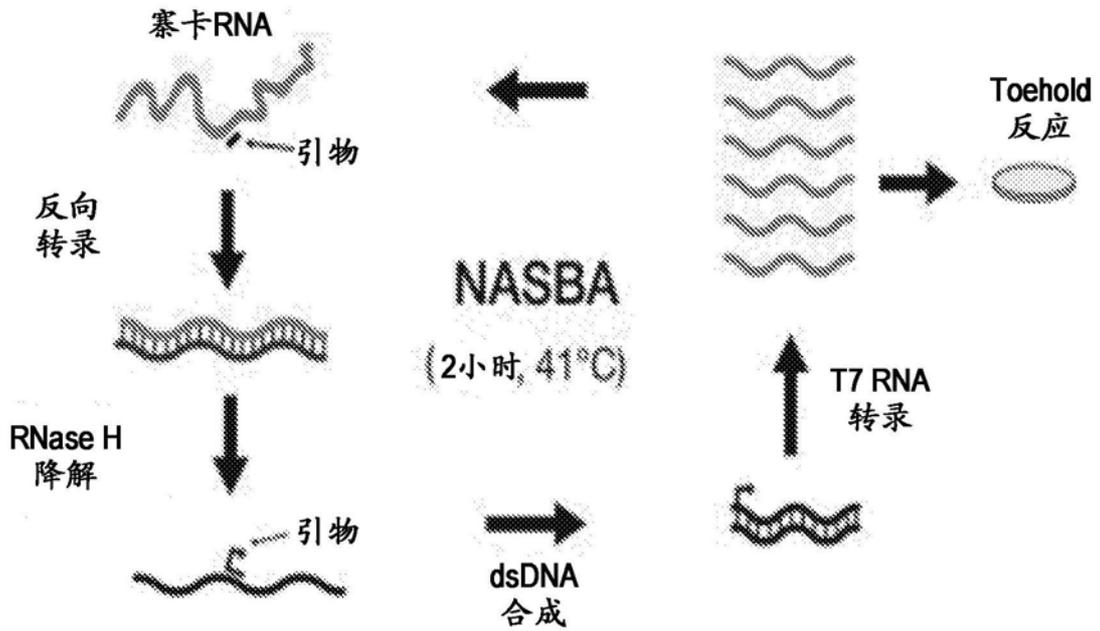


图10

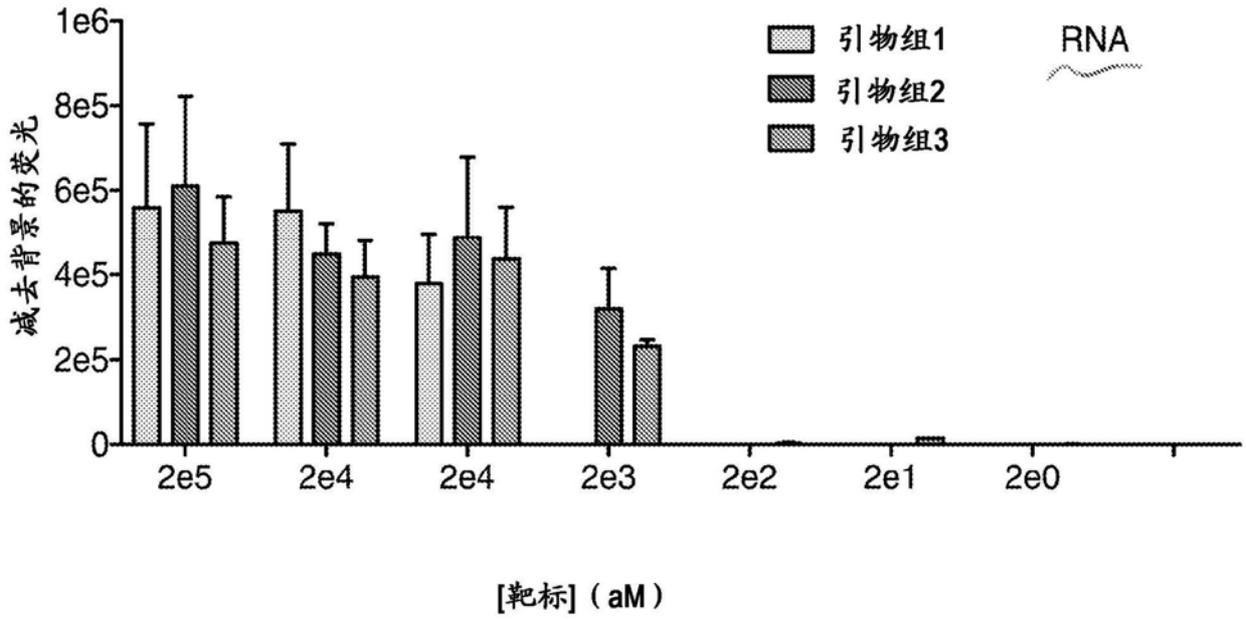


图11

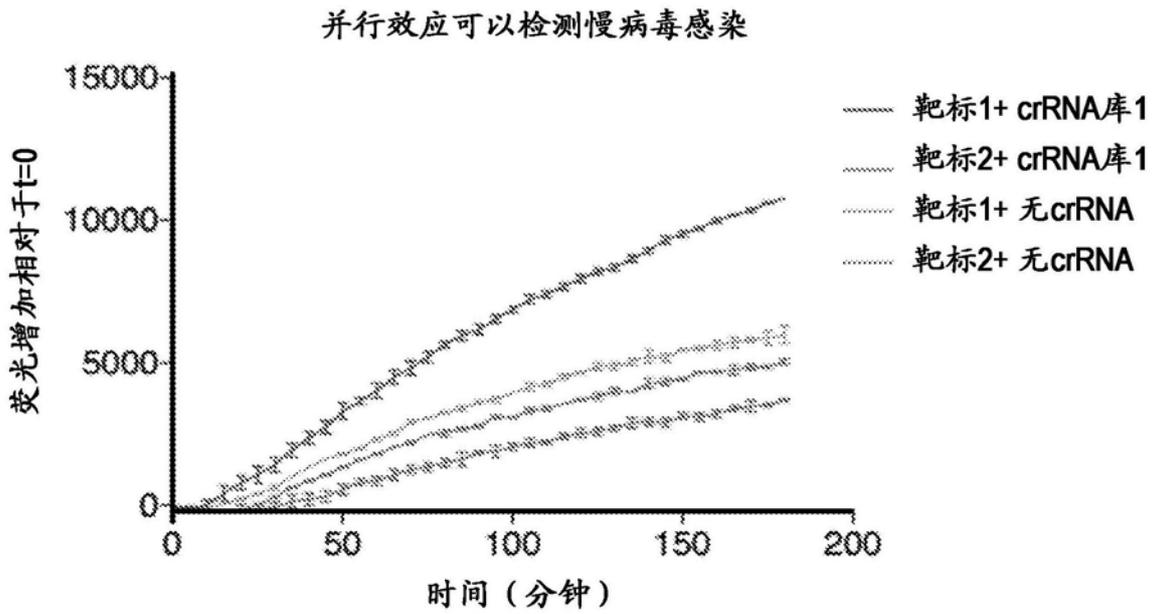


图12

并行和NASBA可以检测
aM浓度的物质

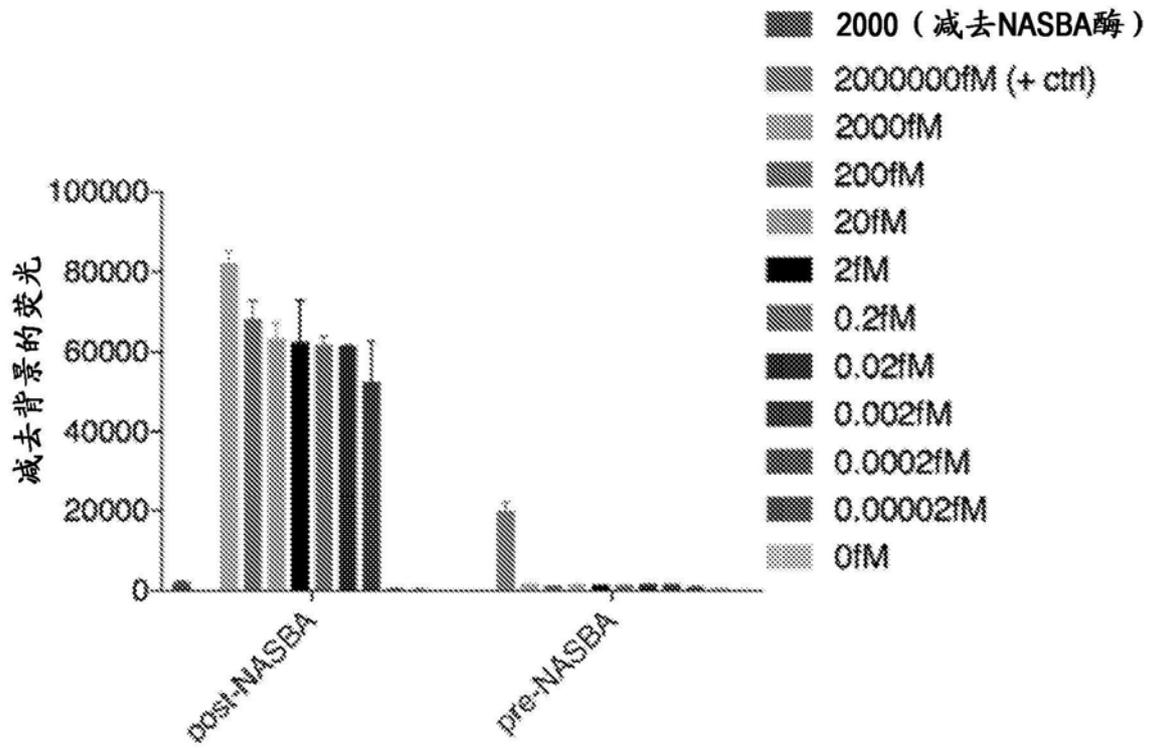


图13

并行和NASBA可以快速区分
低浓度样品

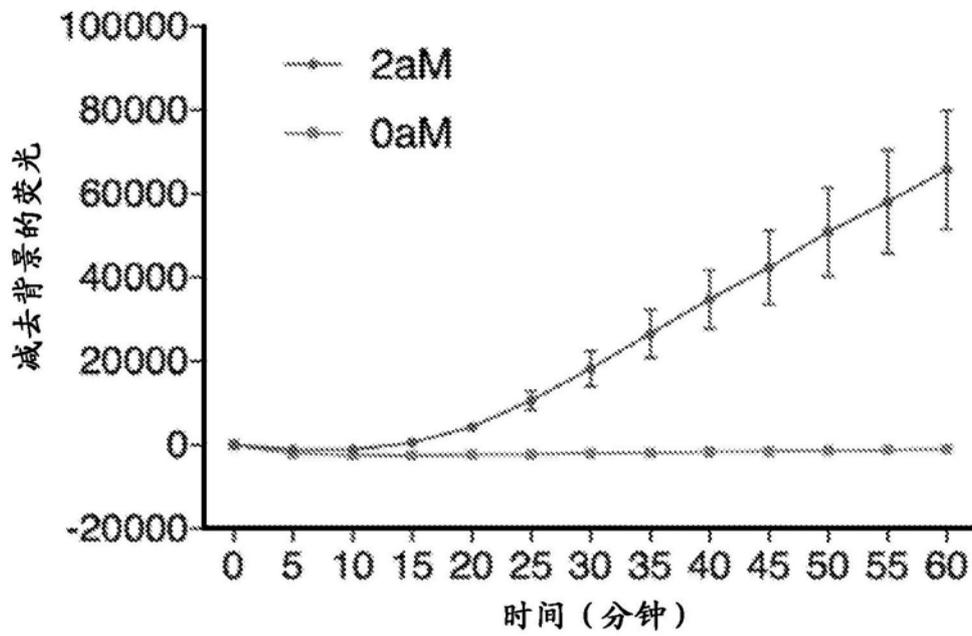


图14

在时间点处的标准化荧光
预测样品浓度

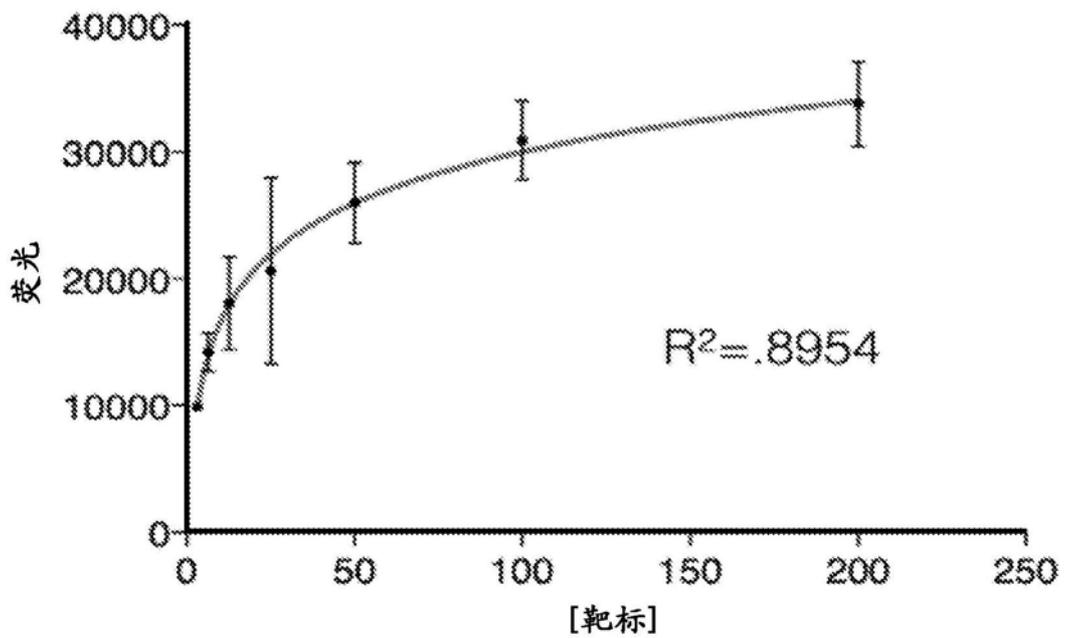


图15

A

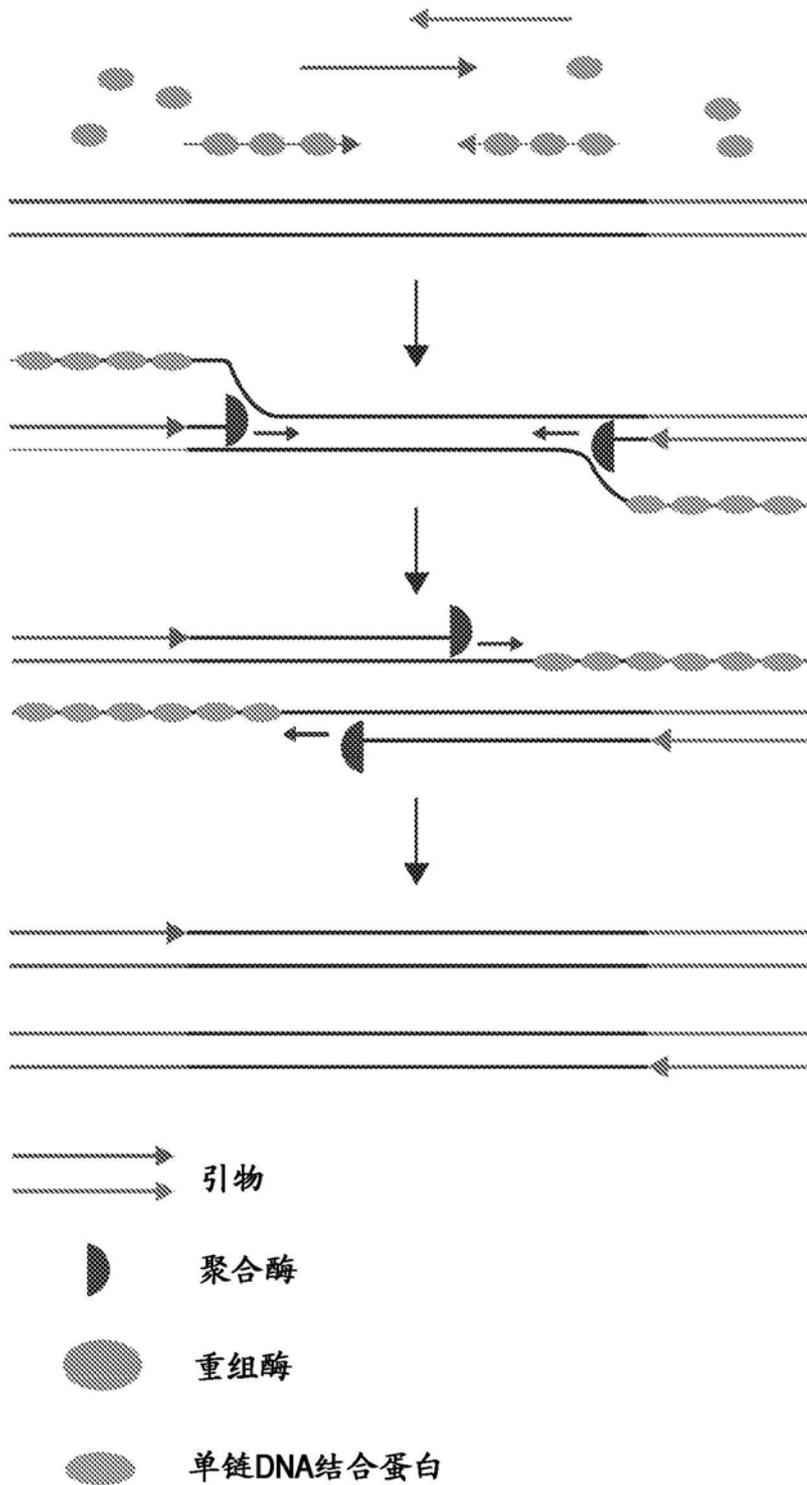


图16

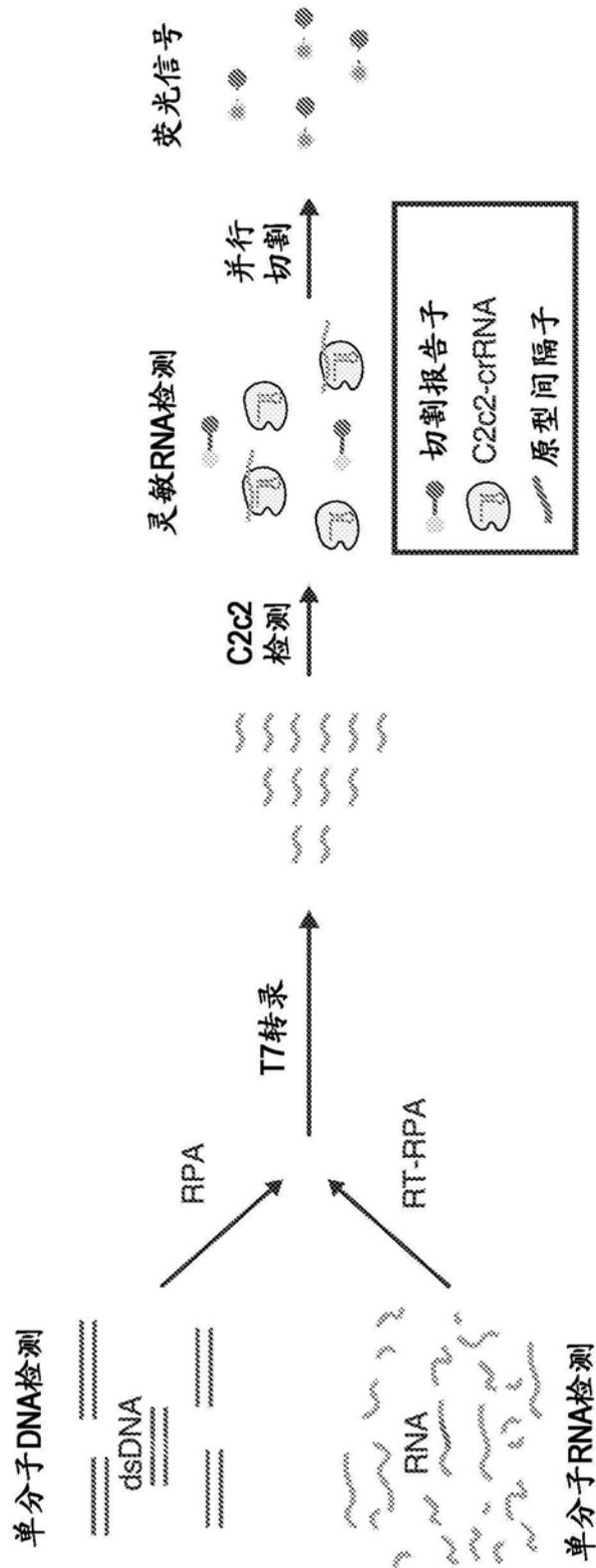


图17

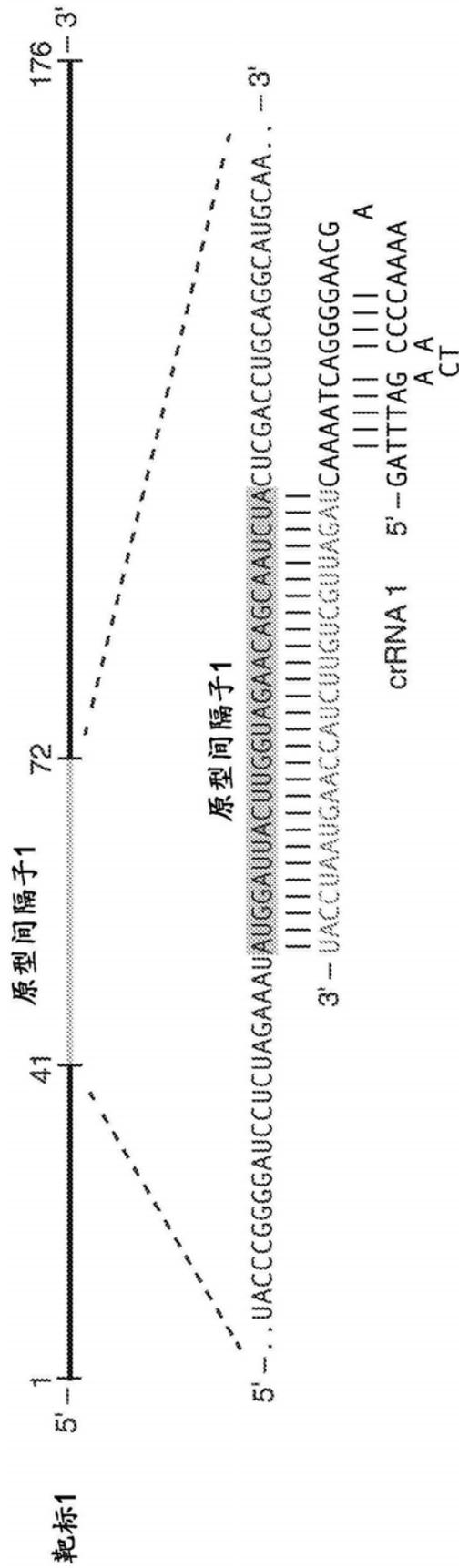


图18

使用RPA的单分子DNA检测

注：这是20分钟的RPA和15分钟的C2c2
(非常快!)

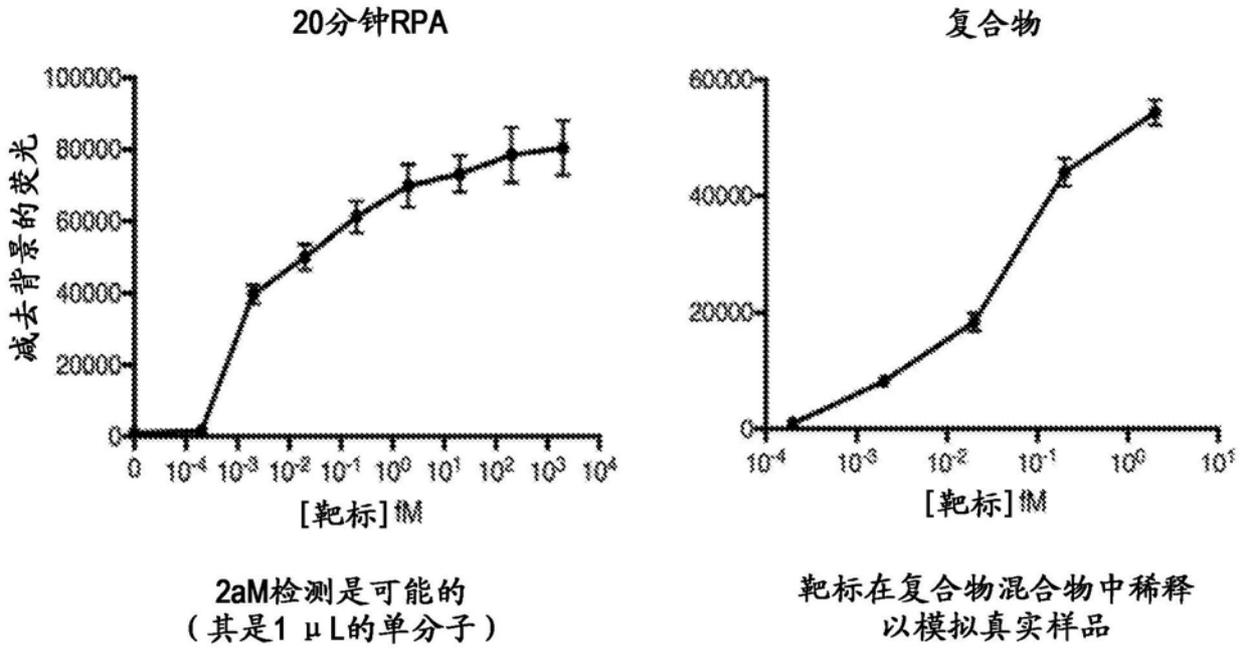
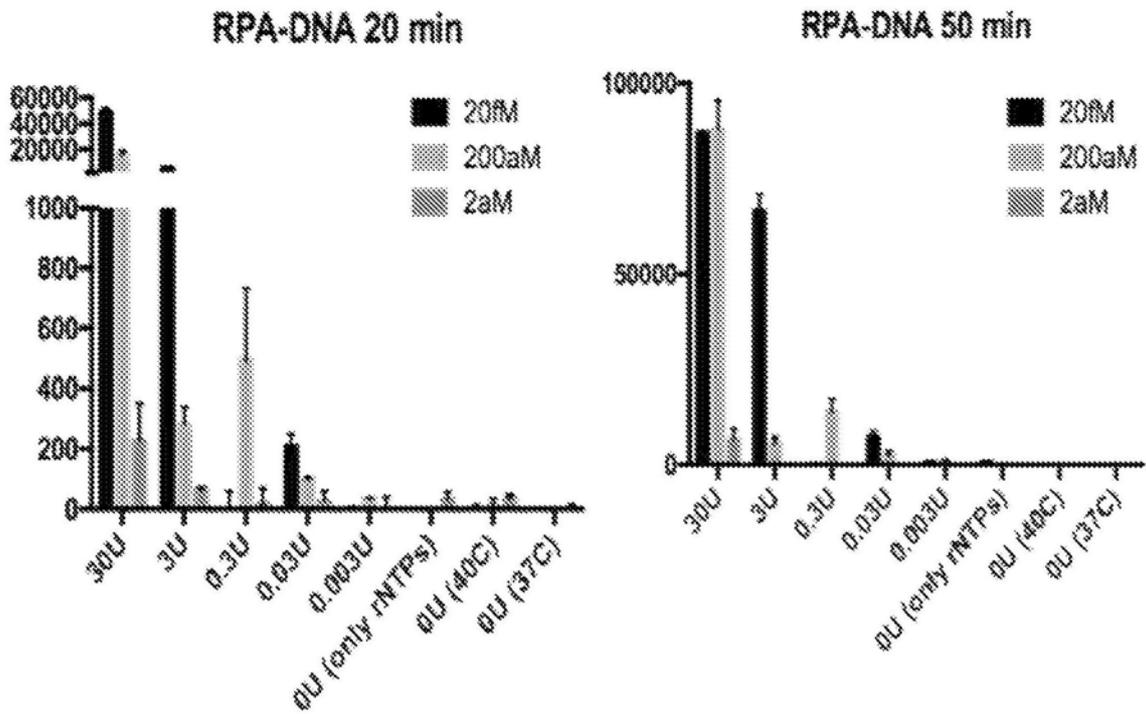


图19

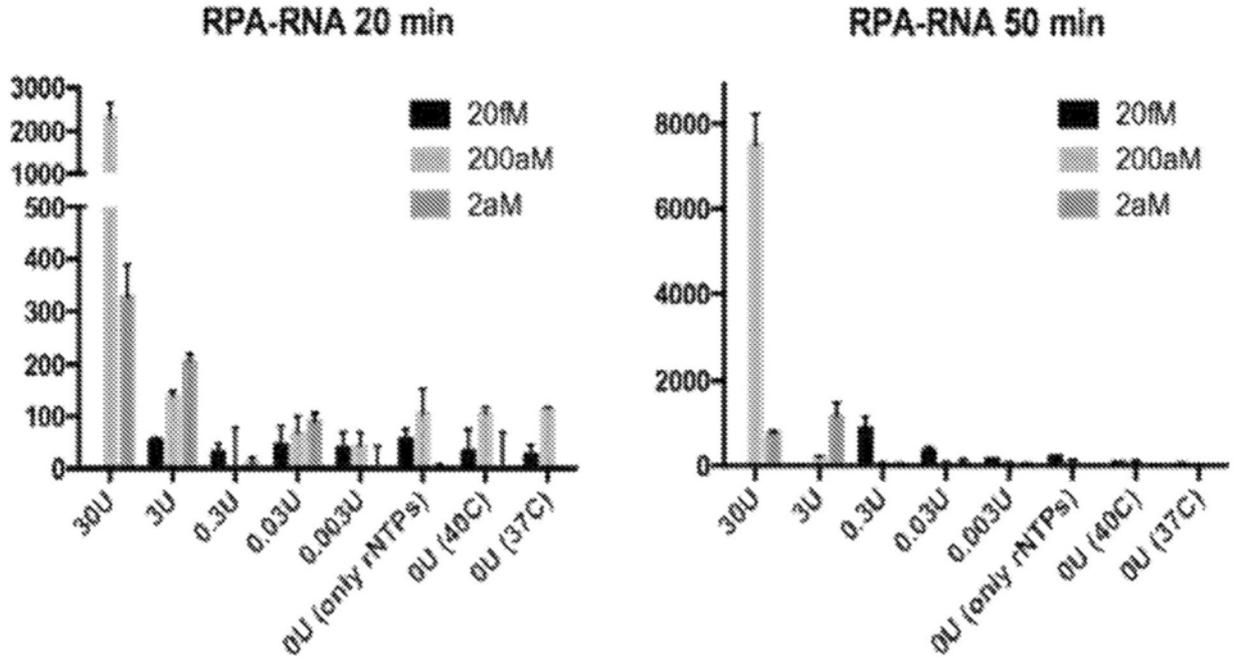
将T7聚合酶混合到反应中仍然对于DNA检测生效



早至C2c2孵育20分钟时DNA的单分子检测。更多的T7聚合酶是更好的 (X轴是T7的单位)

图20

将T7聚合酶混合到反应中仍然对于RNA检测生效



早至C2c2解育20分钟时RNA的单分子检测。

图21

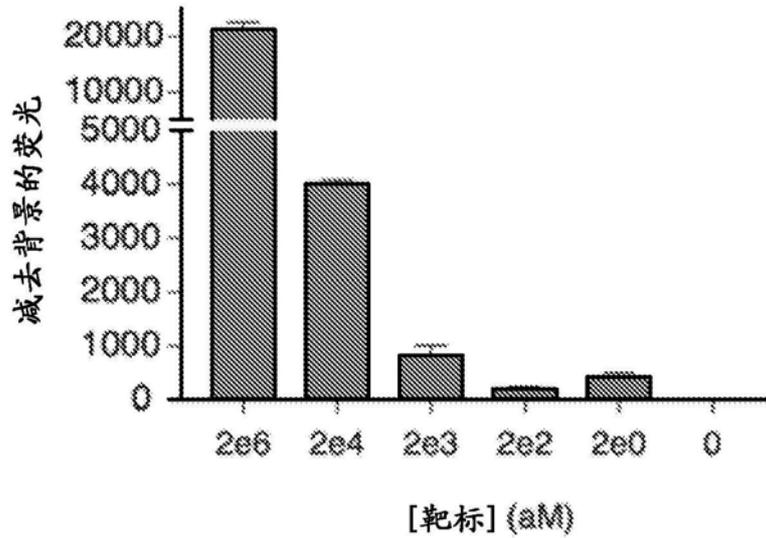


图22

快速RPA-RNA时间解育仍然生效

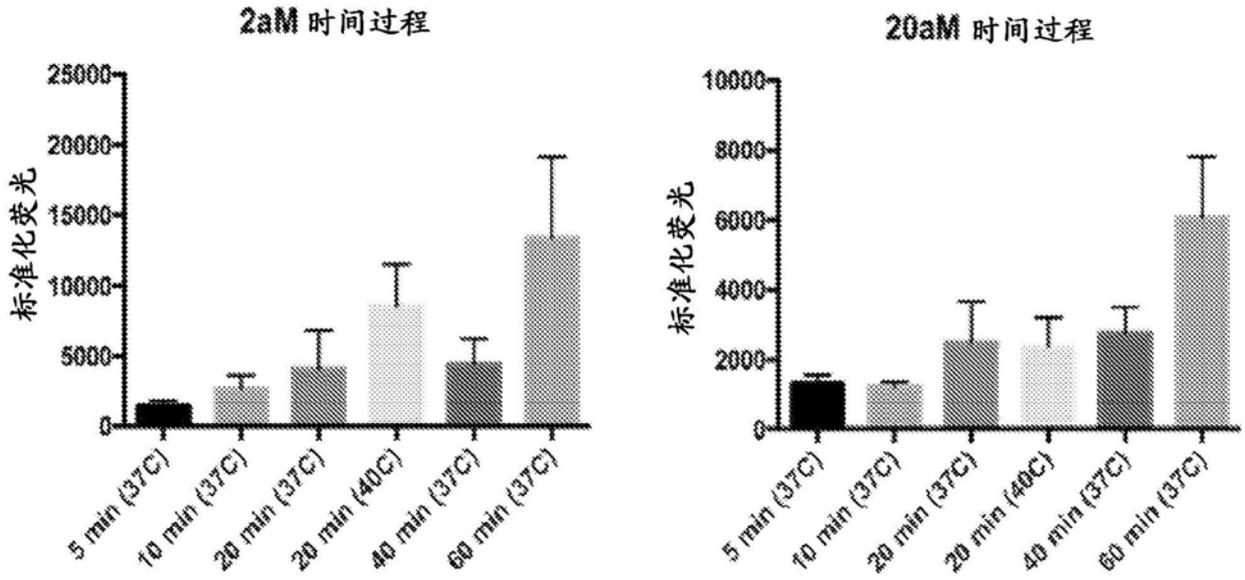


图23

增加T7聚合酶的量也增强RPA-RNA的灵敏度

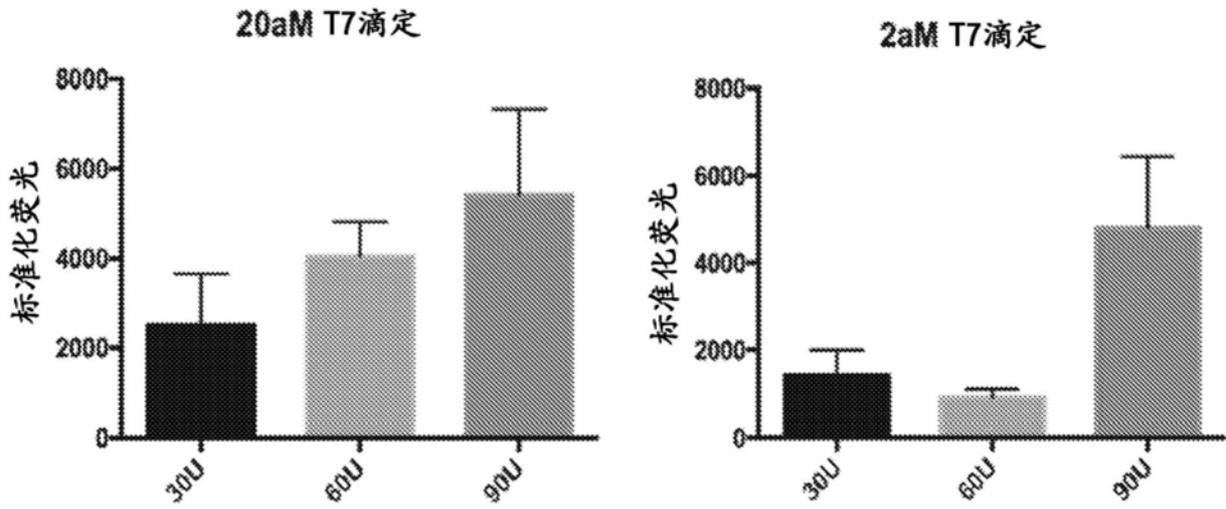


图24

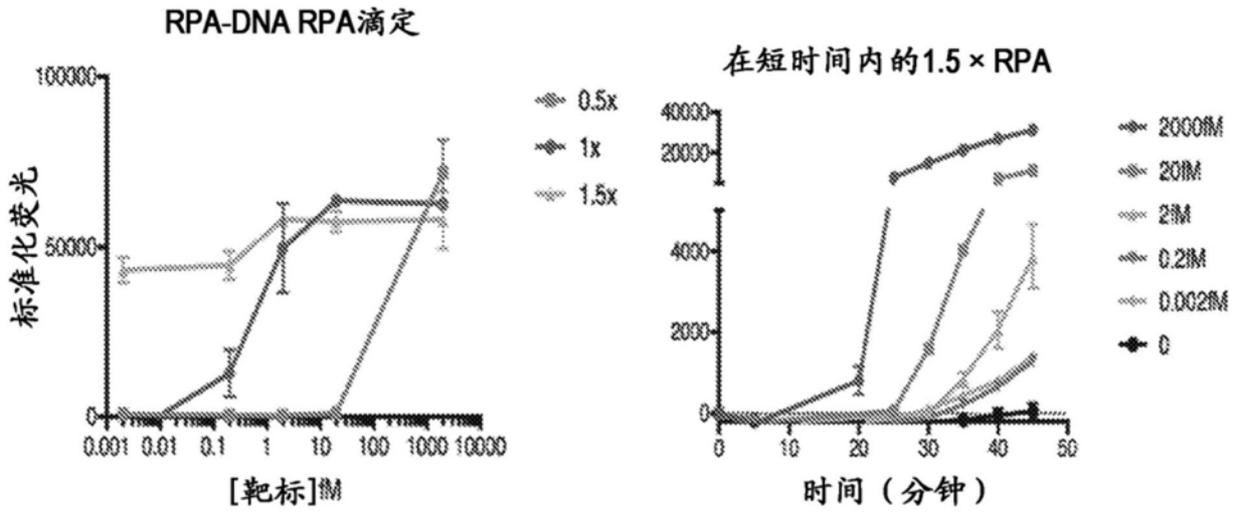


图25

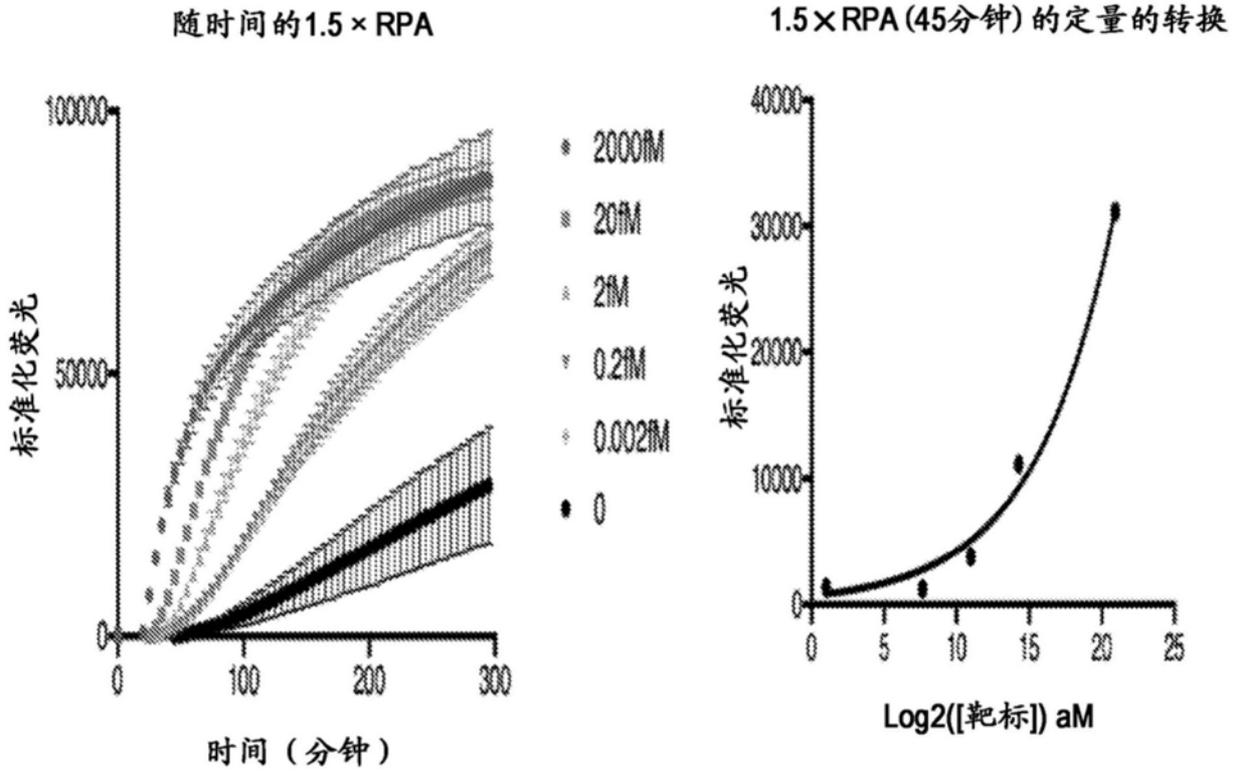


图26

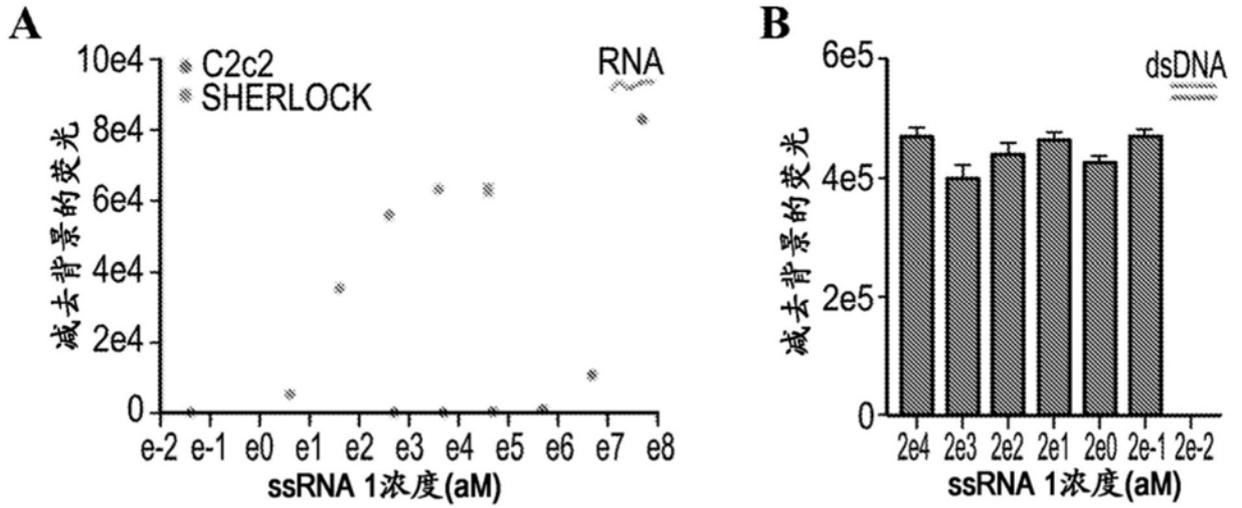
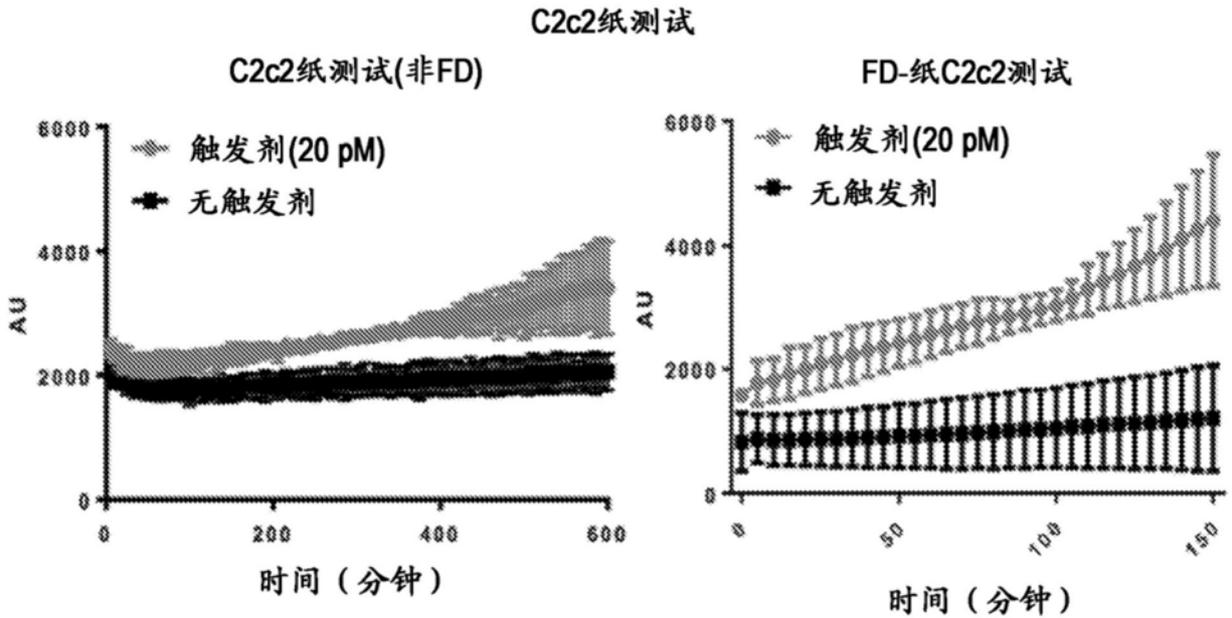


图27



这是20pM靶标在纸上的C2c2信号。注意在实际增强信号前的冻干(右图)。

图28

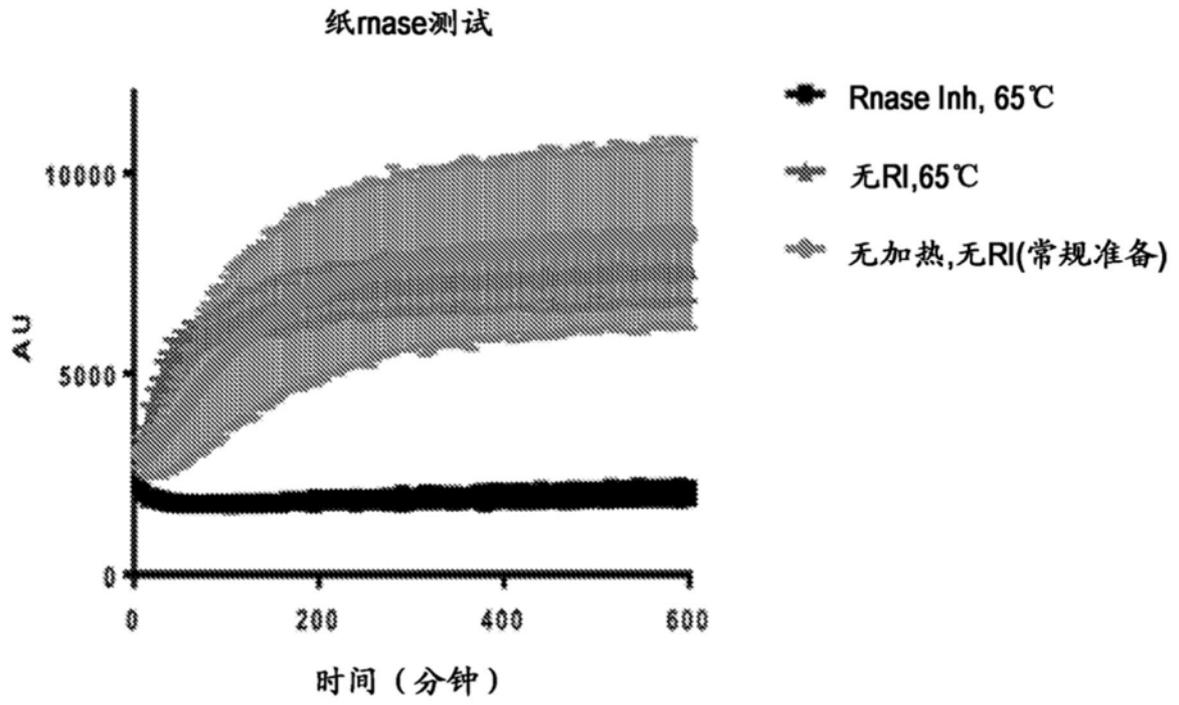


图29

C2c2玻璃纤维测试

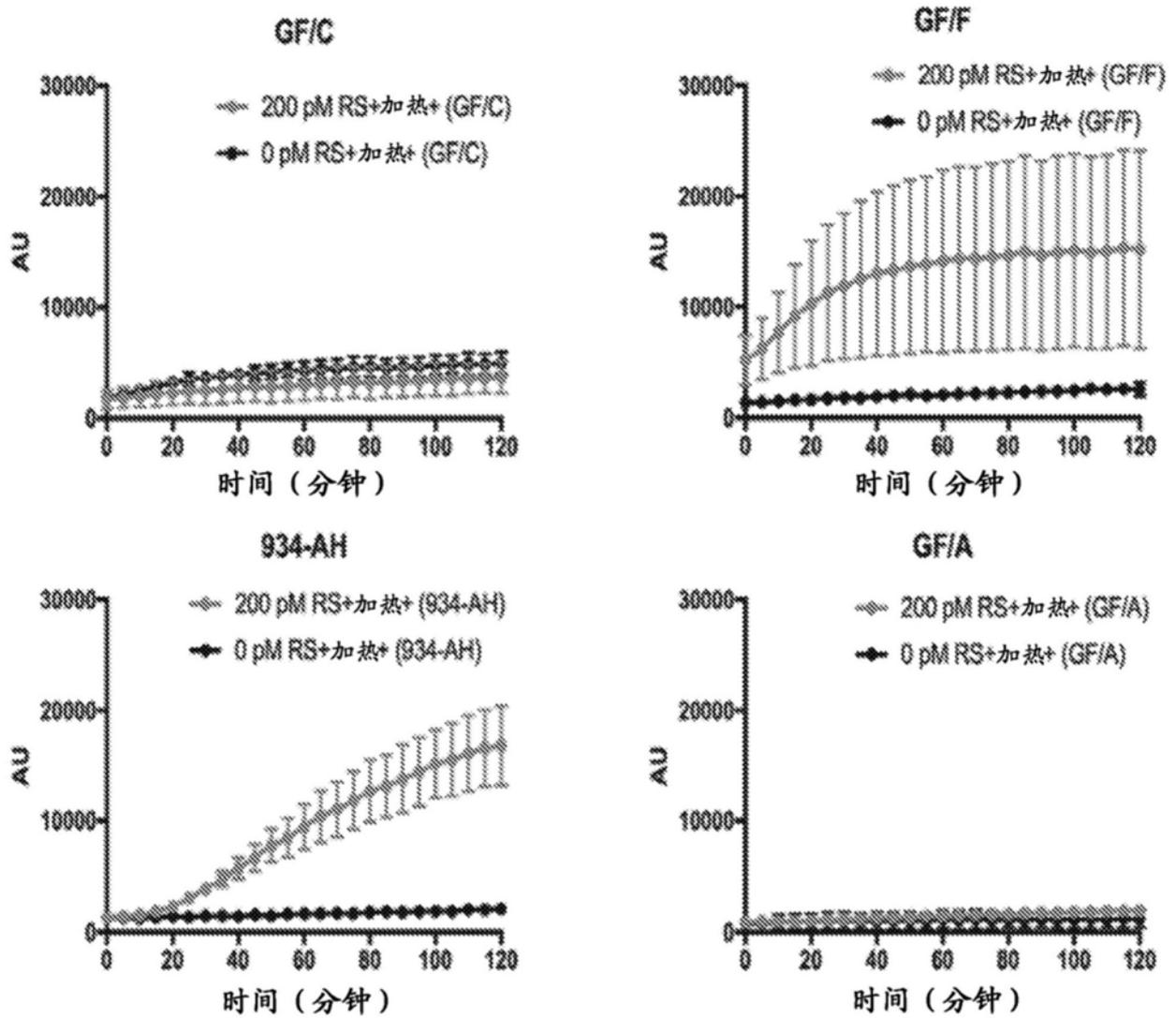
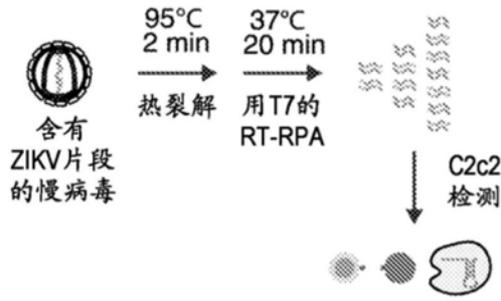
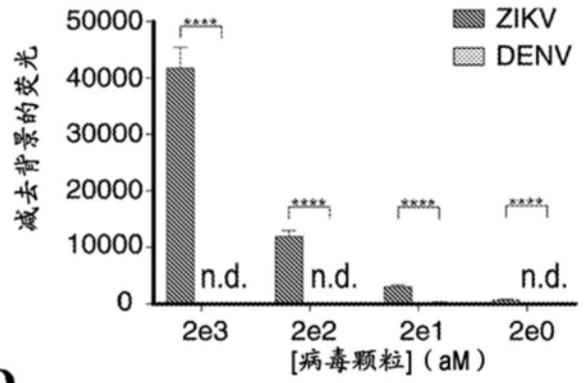


图30

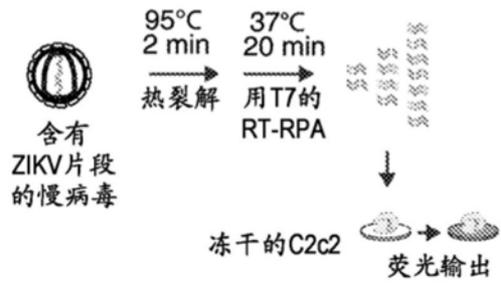
A



B



C



D

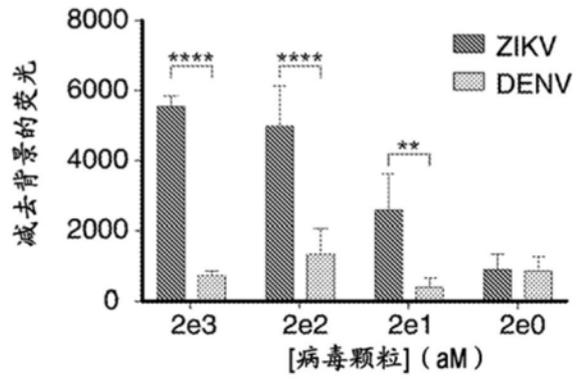
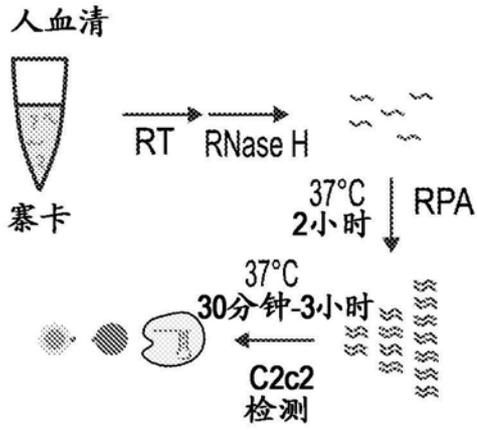


图31

A



B

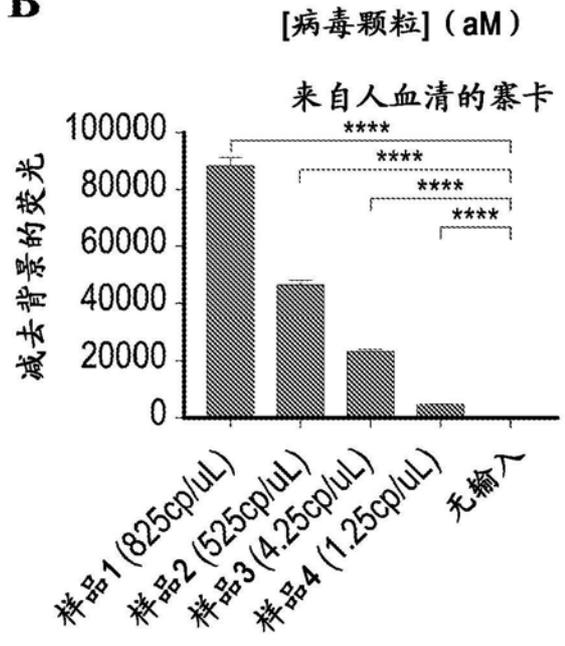


图32

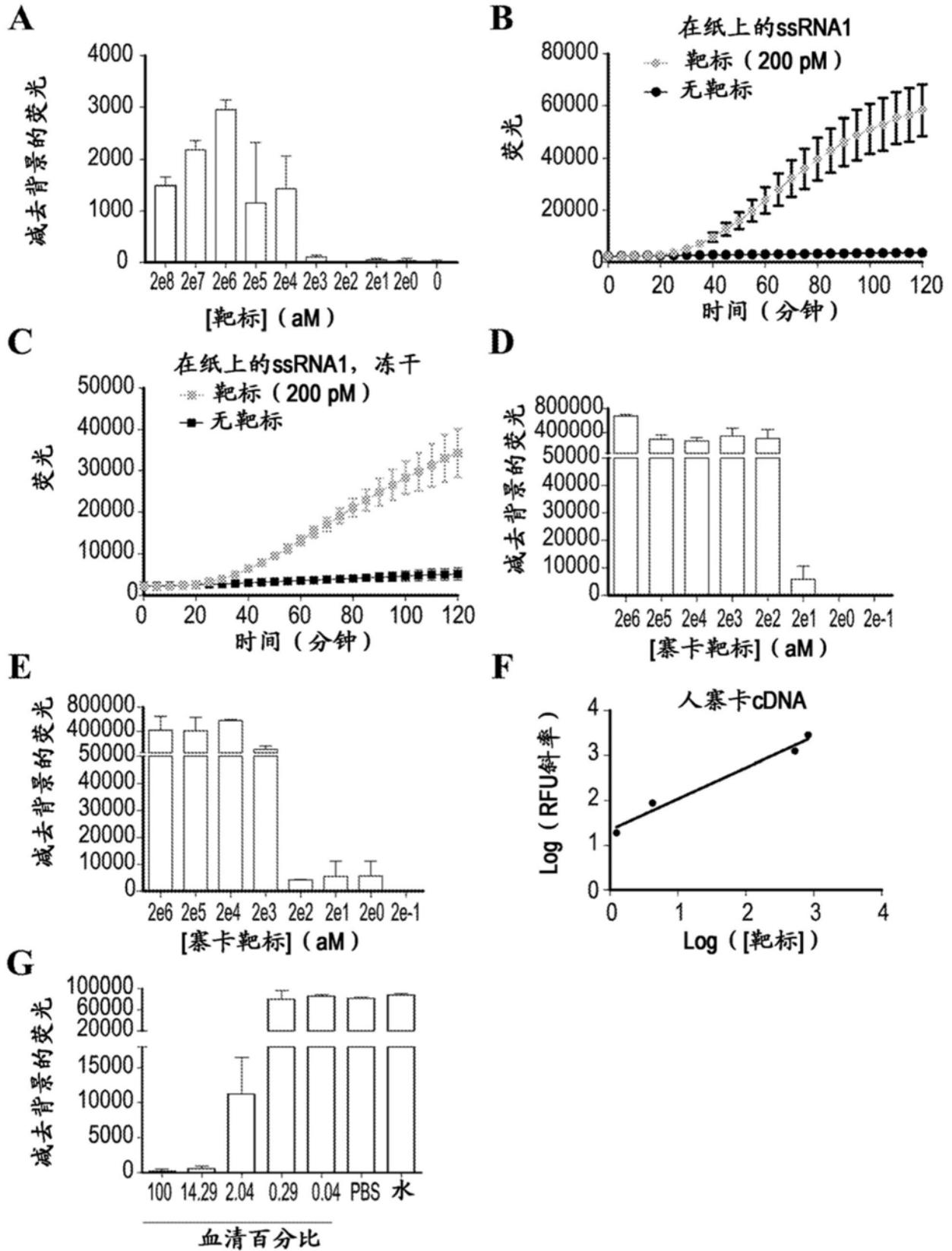


图33

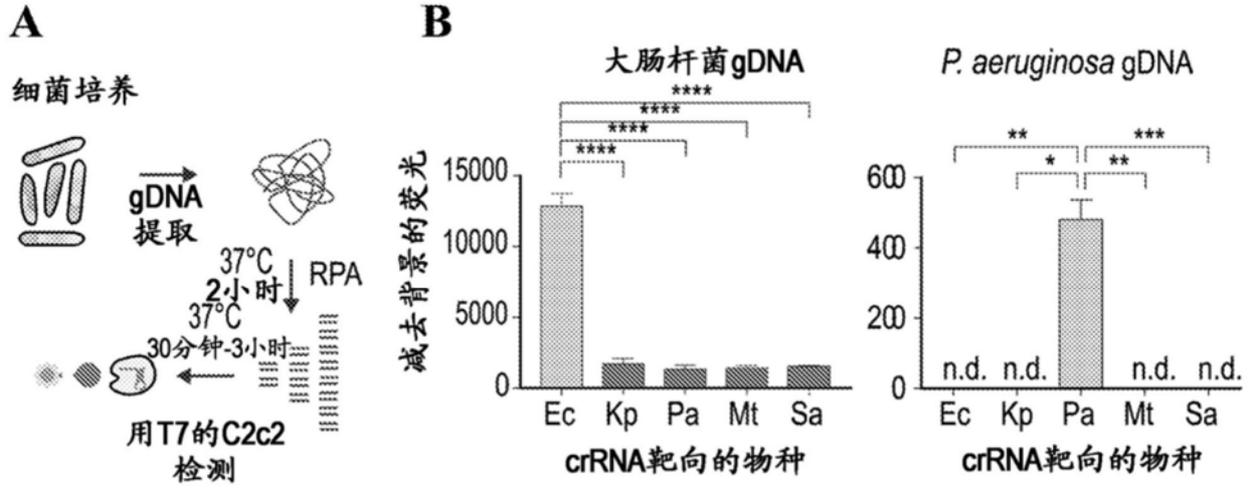


图34

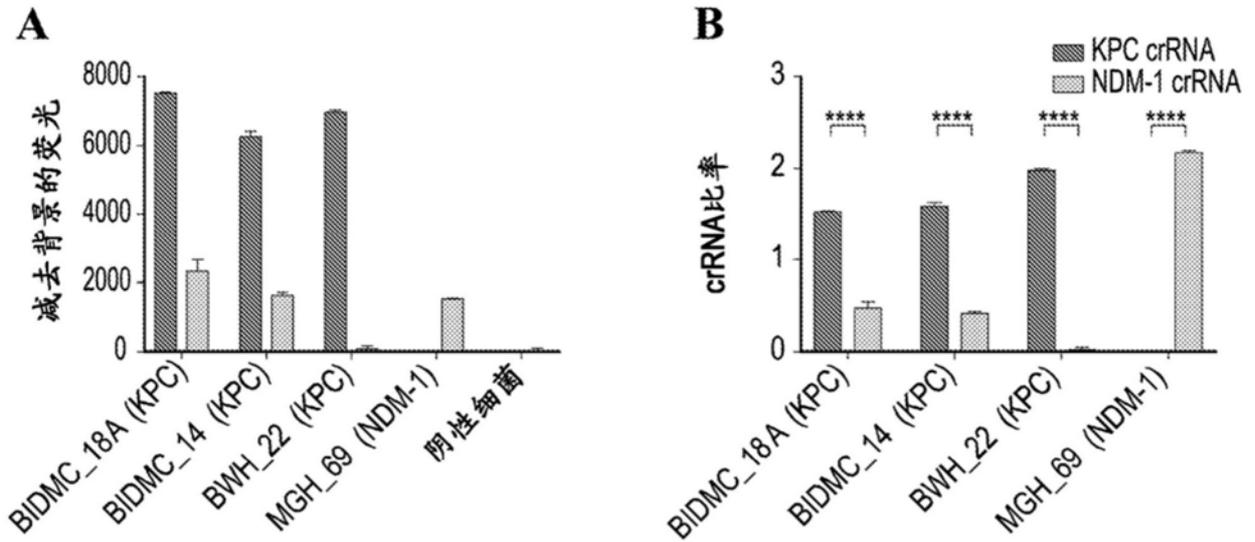
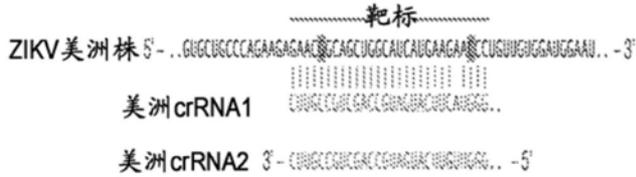
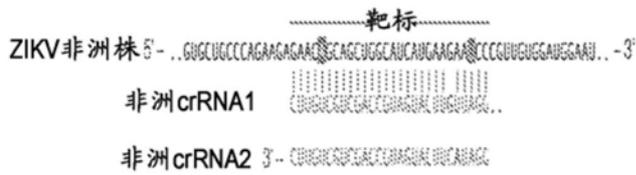
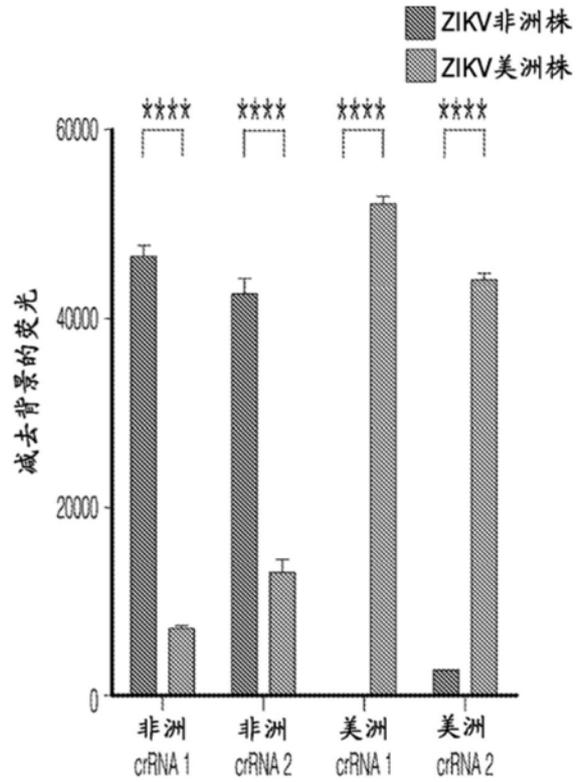


图35

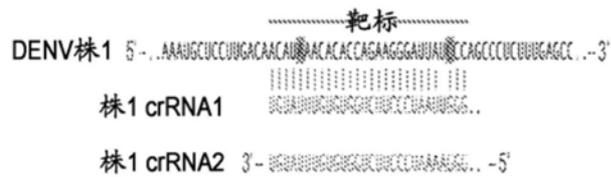
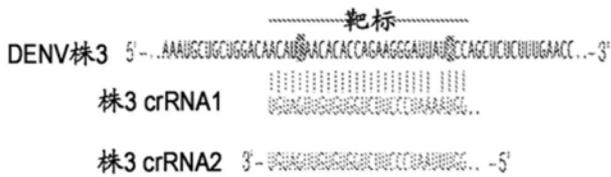
A



B



C



D

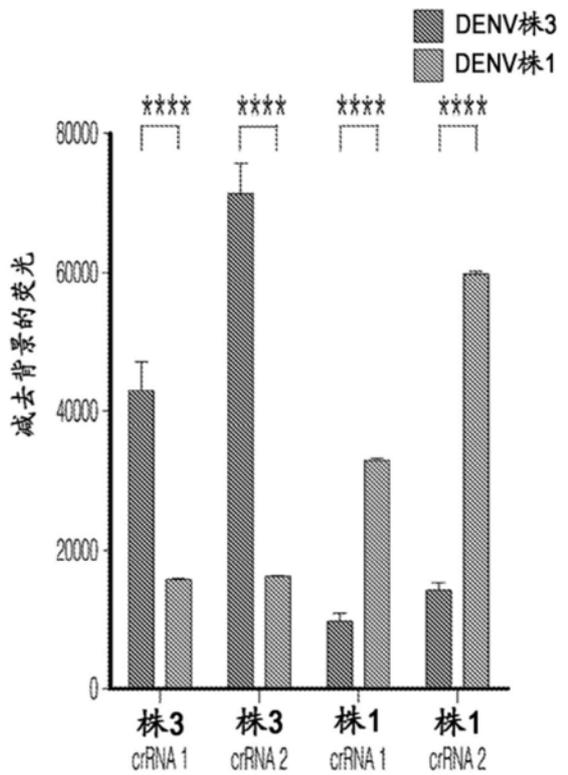


图37

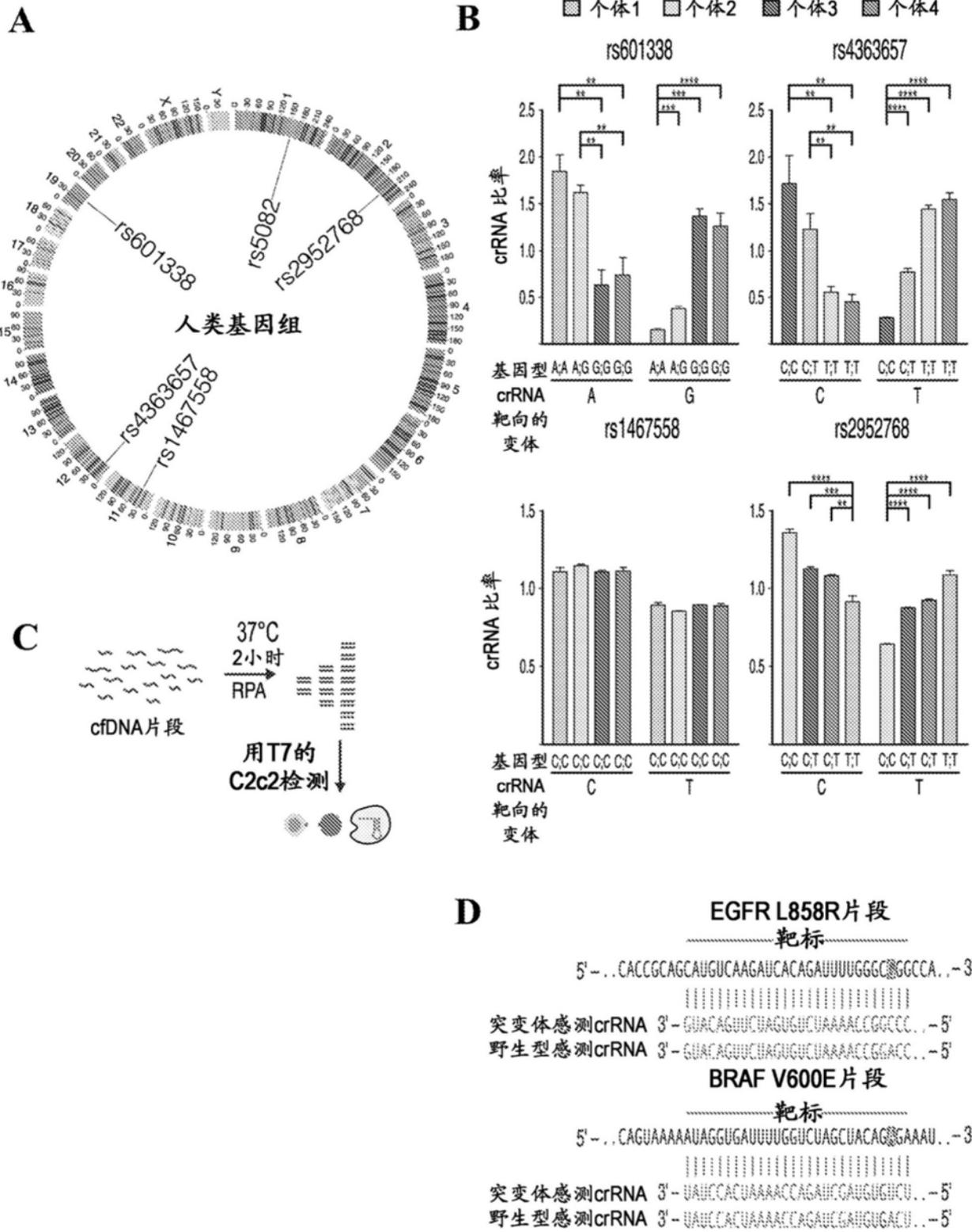


图38

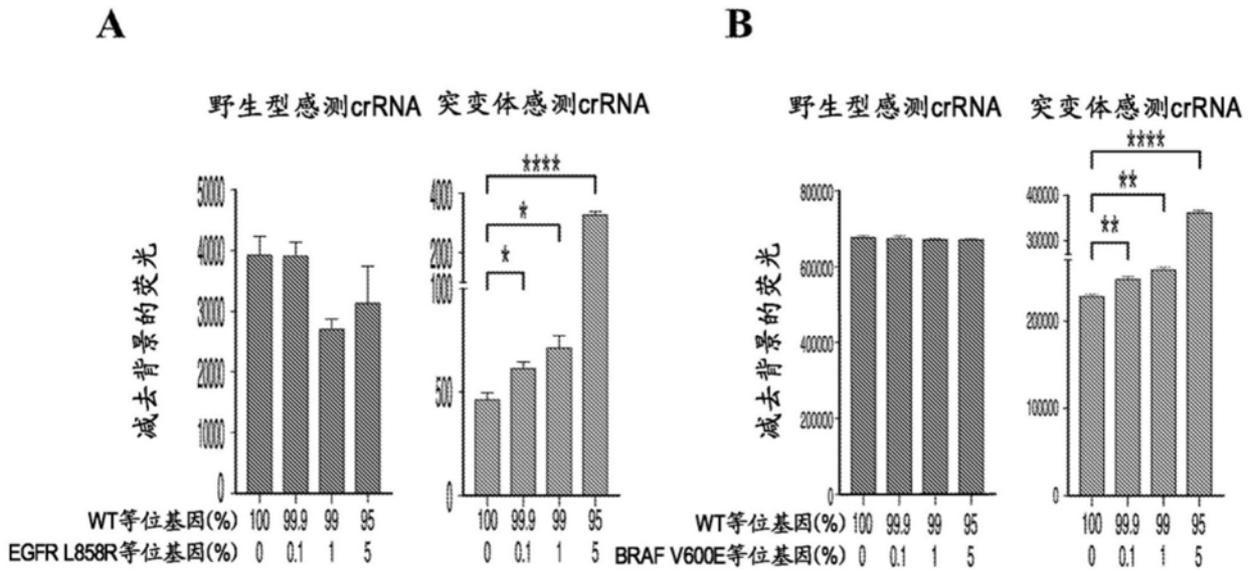


图39

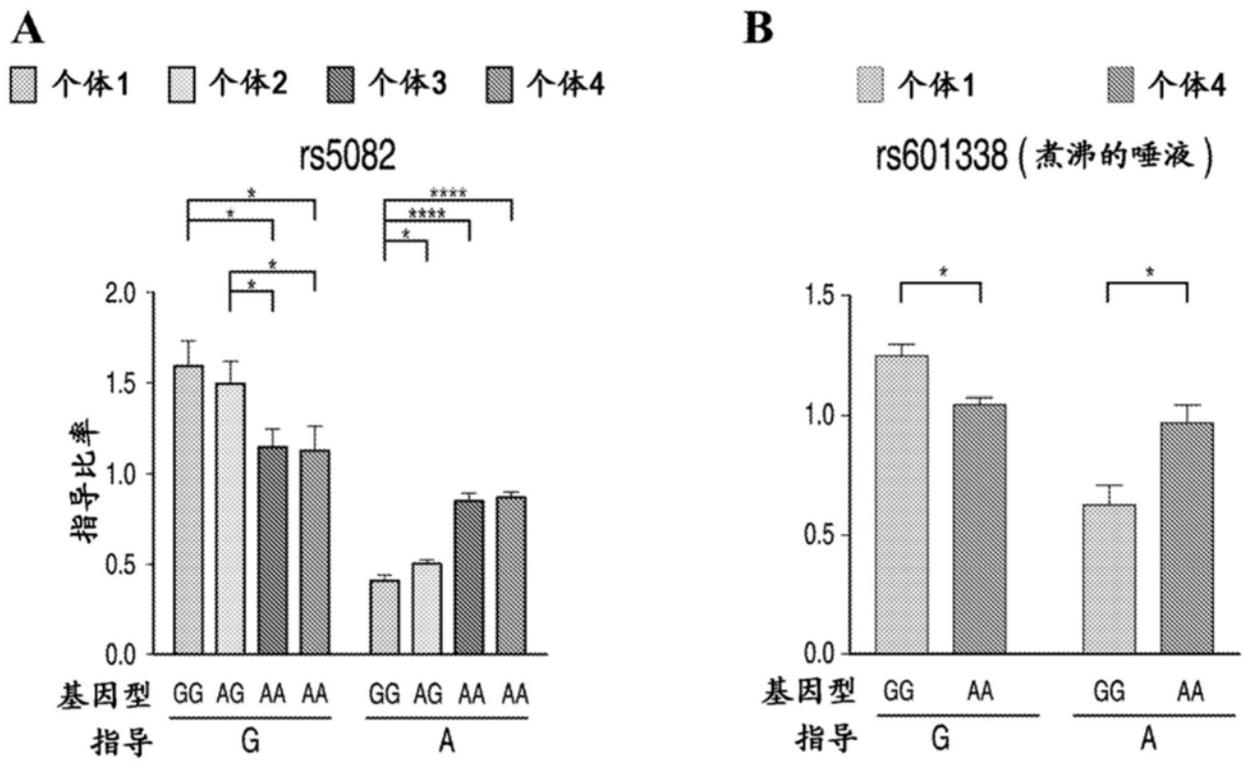
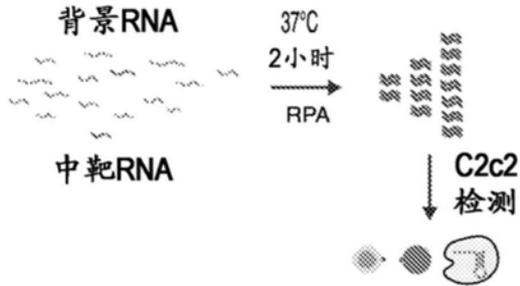


图40

A



B

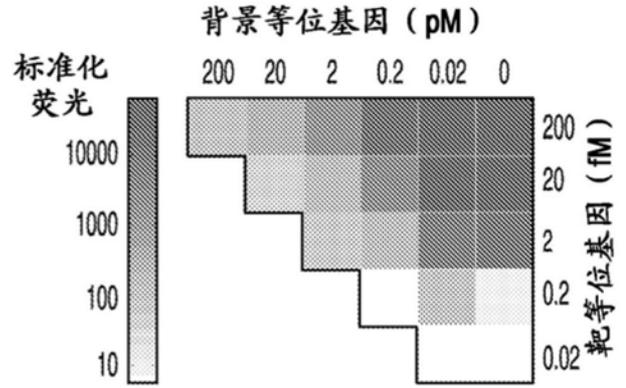


图41

用不同的染料Cy5合成的
传感器RNA探针也工作

Cy5 Ex/Em 638/668nm

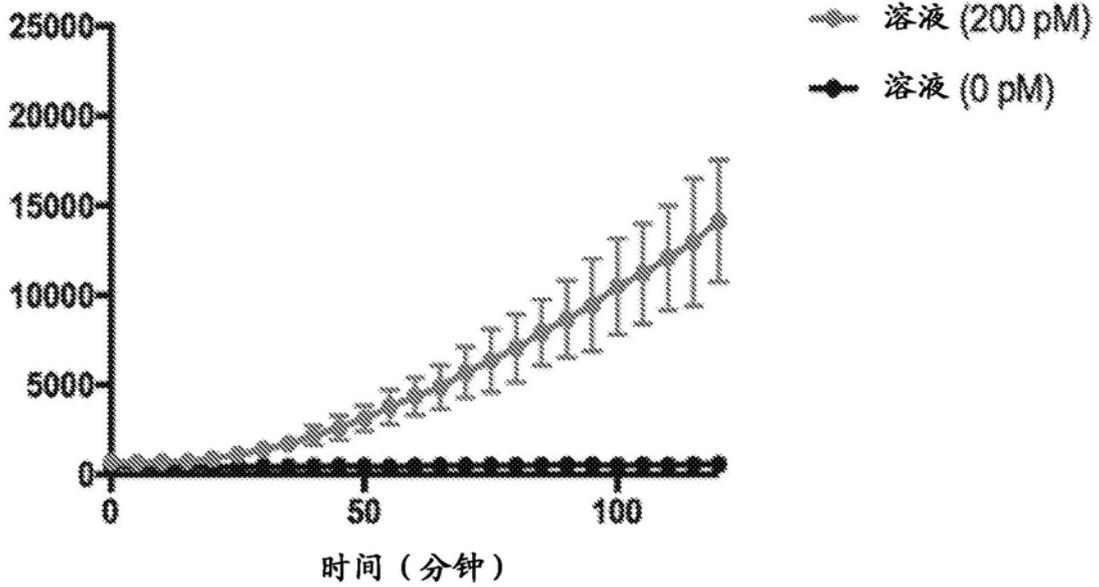


图42

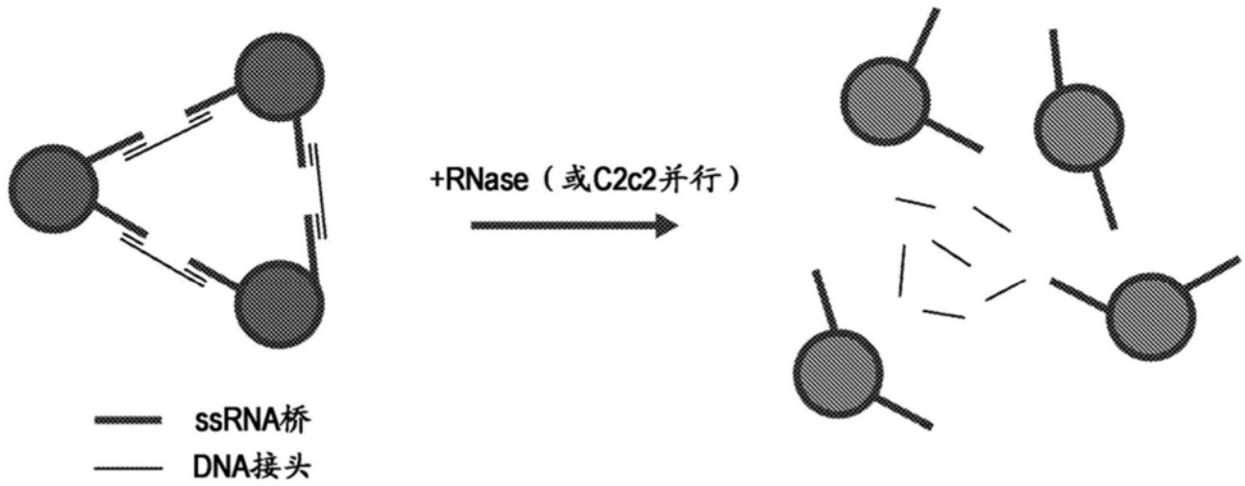


图43

比色测定

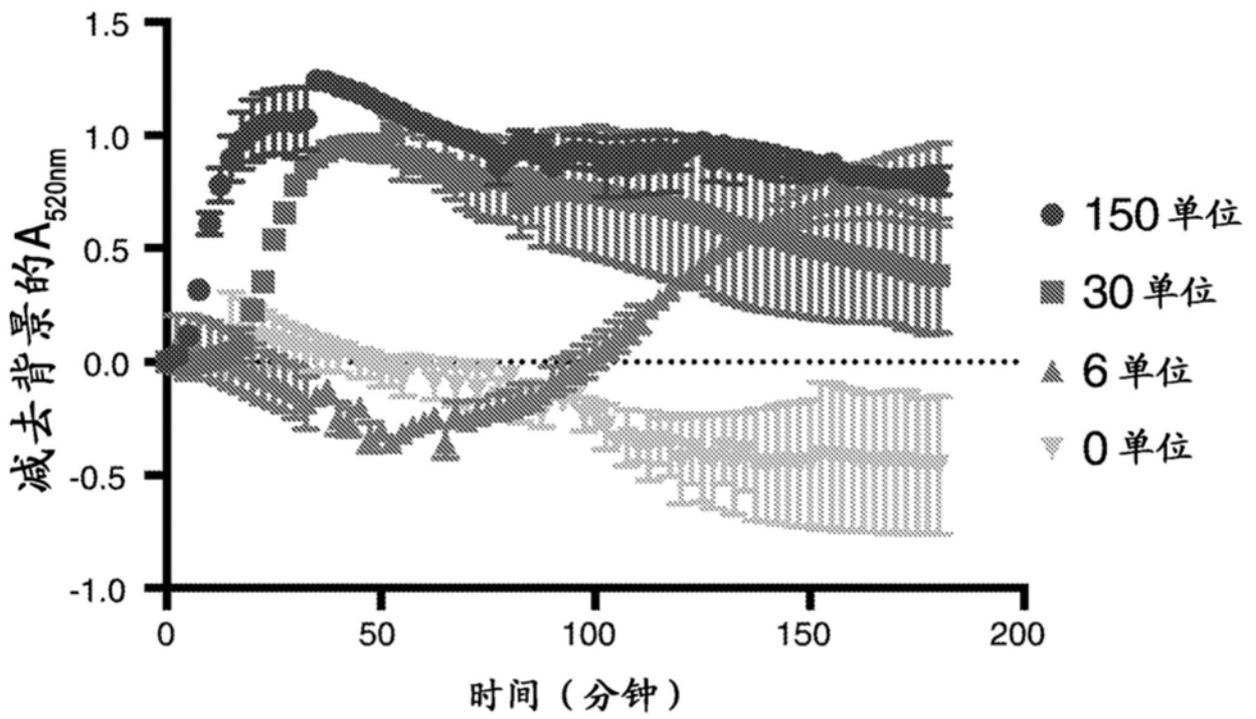


图44

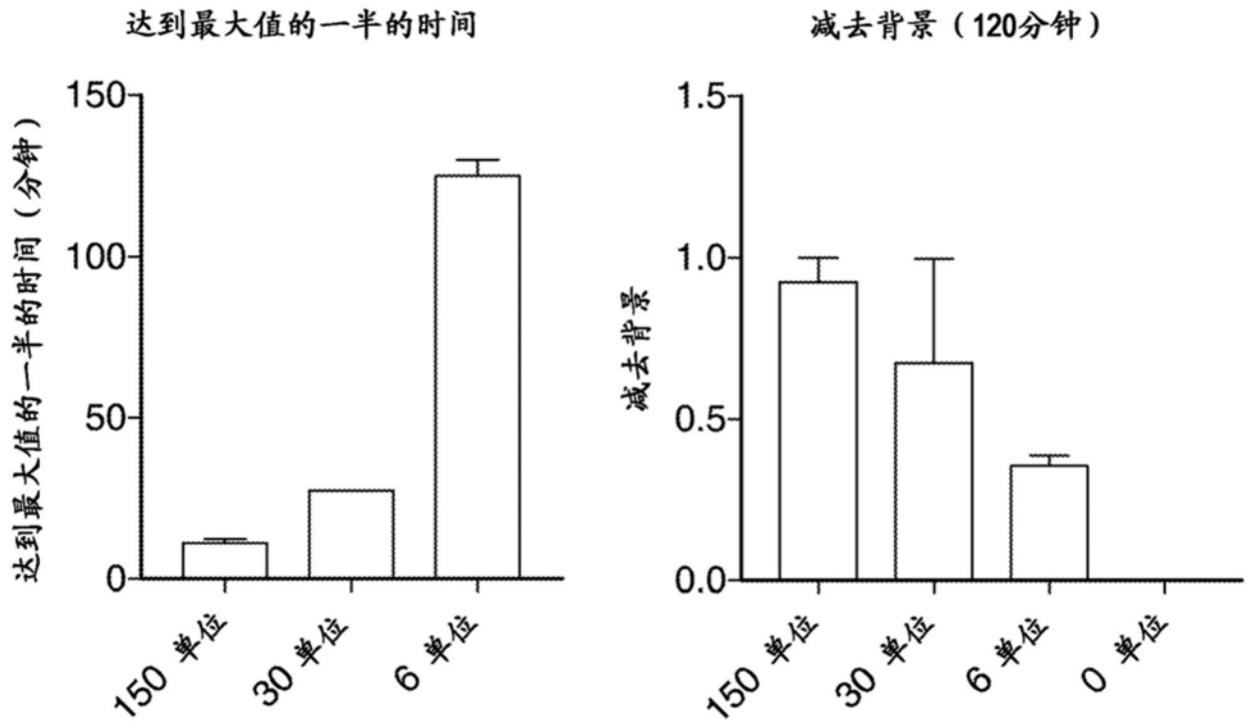


图45

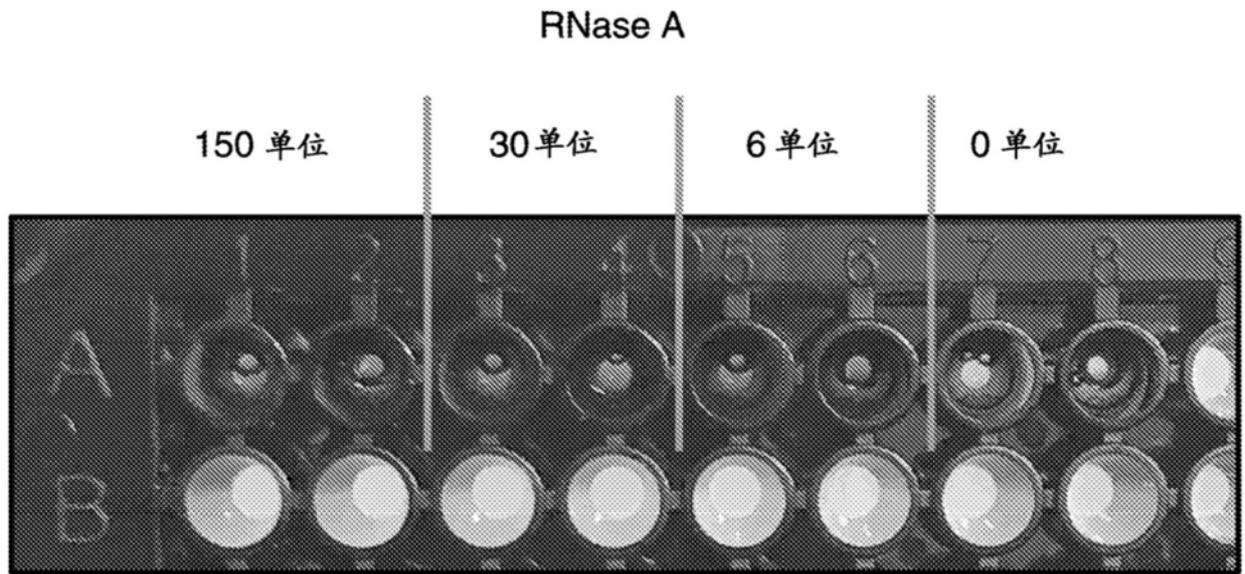
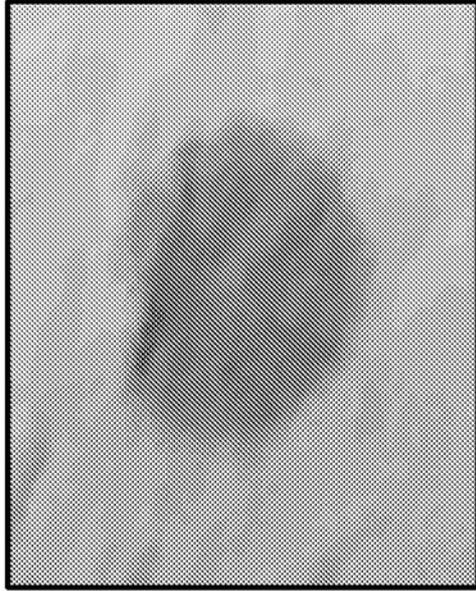


图46



30单位的RNase A



无RNase A

图47

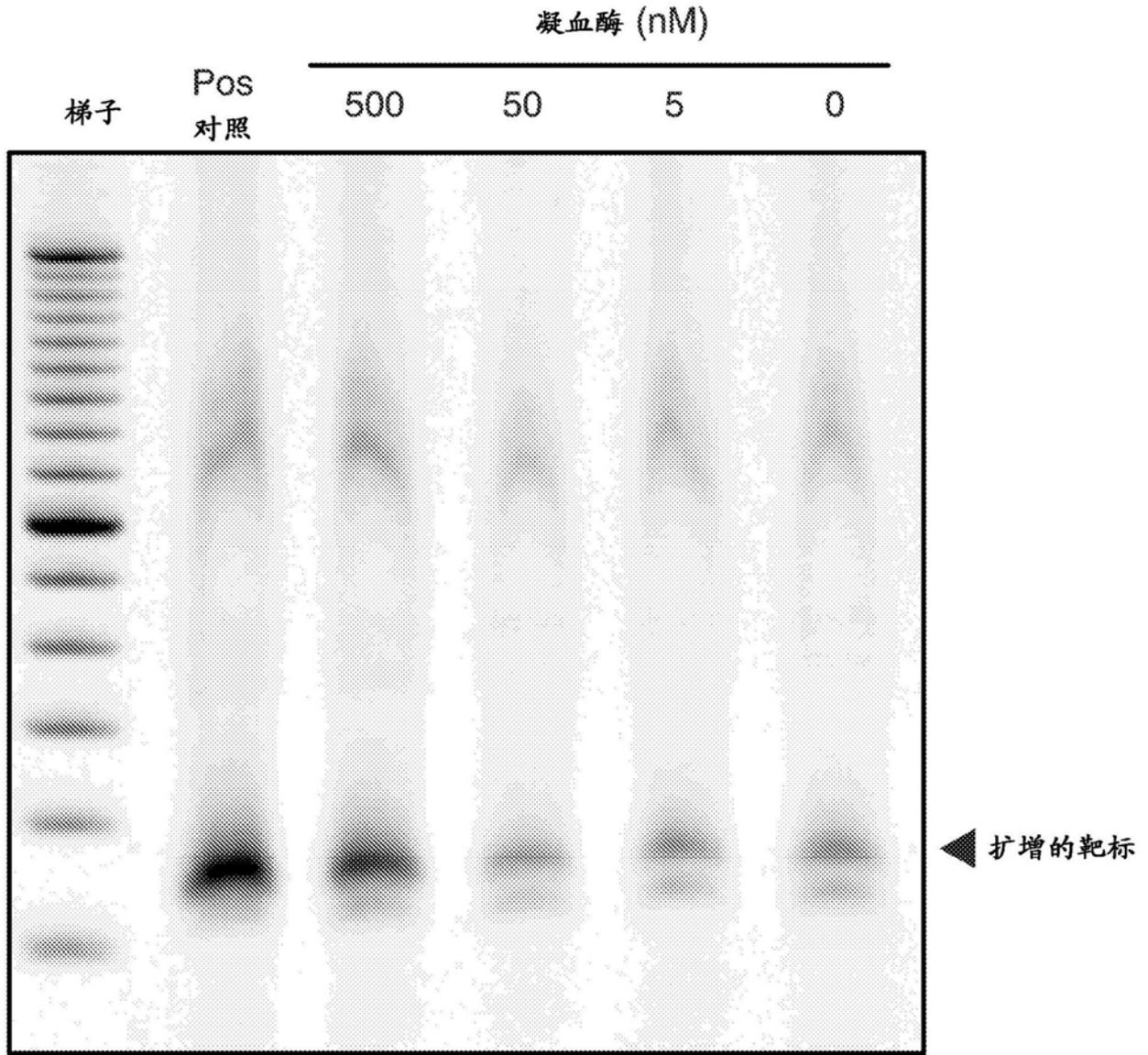


图49

	586	693	1276	1298
c2c2_Leptotnchia shahii	IRKFTKIGTN	ERNRILHA	SIRNYISHFYIVRNP	
c2c2-5 Lachnospiraceae bacterium MA2020	LYSLKSMLYS	MRNSSFHF	IFRNEIDHFHYFYDR	
c2c2-6 Lachnospiraceae	LTDLKDVIYS	MRNDSFHY	ELRNYIEHFYRYSSF	
c2c2-7 [Clostridium] ammophilum DSM 10710	ADDLRKAIYS	LRNETFHF	DVRKYVDHFYKYATS	
c2c2-8 Camobacterium gallinarum DSM 4847	IWALRGSVQQ	IRNEIFHS	KIRNQTAHLSVLQLE	
c2c2-9 Camobacterium gallinarum DSM 4847	LWAIIRGAVQR	VRNQIFHQ	EIRNNIAHLHVLNRND	
c2c2-10 Paludibacter propionicigenes WB4	LWGIRGAVQQ	IRNNVNHY	DIRNHIAHFNYLTKD	
c2c2-11 Listena weihenstephanensis FSL R9-0317	IWAIRGSIQQ	IRNEVYHC	NARNHIAHLNLYSLK	
c2c2-12 Listenaceae bacterium FSL M6-0635	IWAIRGSIQQ	IRNEVYHC	NARNHIAHLNLYSLK	
c2c2-13 Leptotrichia wadei F0279	FANIDEAISS	IRHGIVHF	YIRNYIAHFNYIPHA	
c2c2-14 Rhodobacter capsulatus SB 1003	VFALLRYLRG	CRNQTFHL	QTRKDLAHFNVLDR	
c2c2-15 Rhodobacter capsulatus R121	VFALLRYLRG	CRNQTFHL	QTRKDLAHFNVLDR	
c2c2-16 Rhodobacter capsulatus DE442	VFALLRYLRG	CRNQTFHL	QTRKDLAHFNVLDR	
c2-3 L wadei (Lw2)	FANIDEAISS	IRHGIVHF	YIRNYIAHFNYIPHA	
c2-4 Listeria seeligeri	SWGLRGAIAP	IRNEIHL	EKRNNISHFNLYNGQ	
		↑↑ ↑	↑↑ ↑	

图50

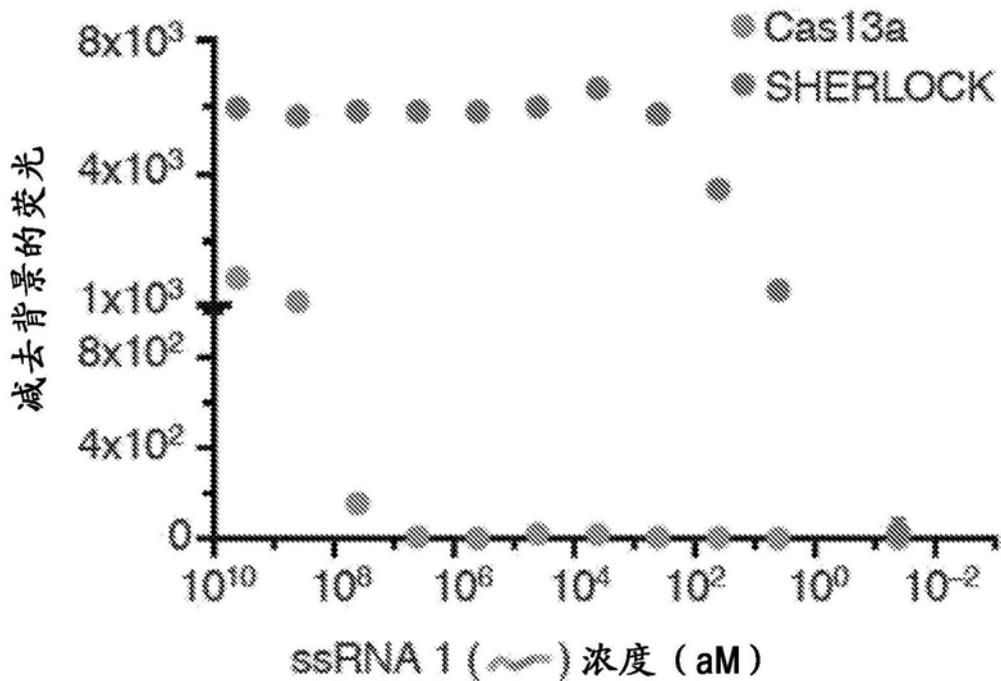


图51

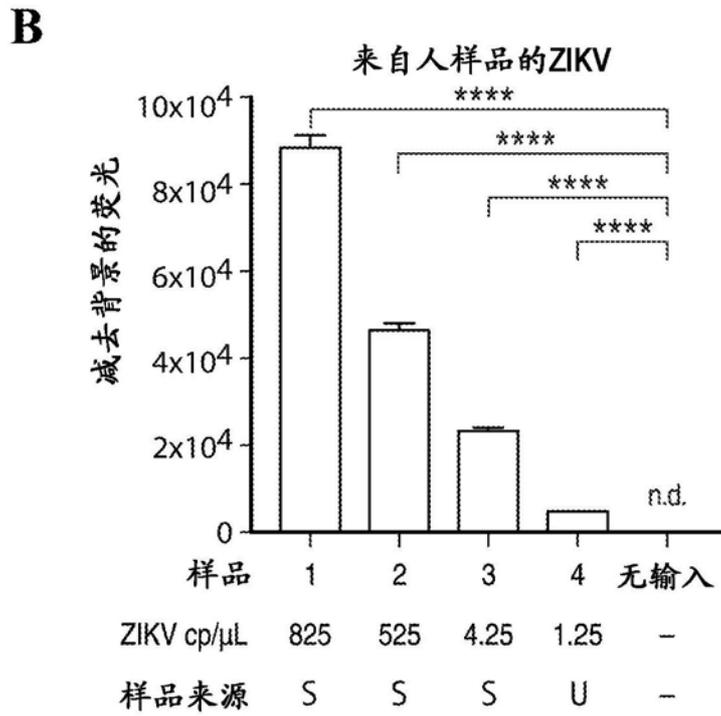
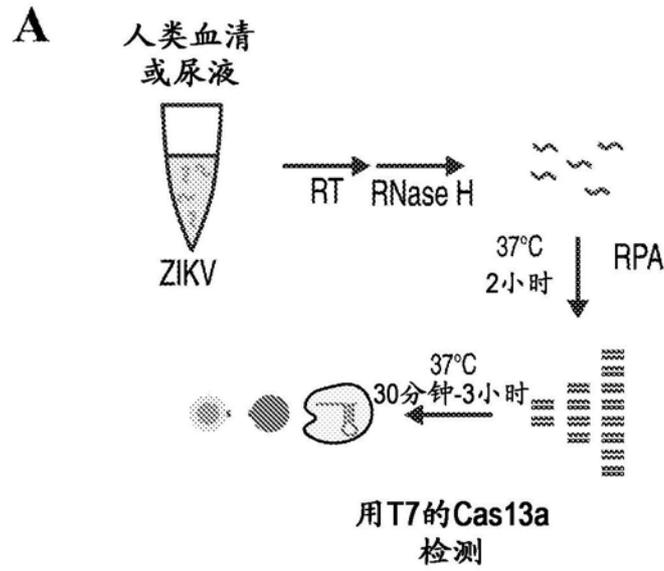


图52

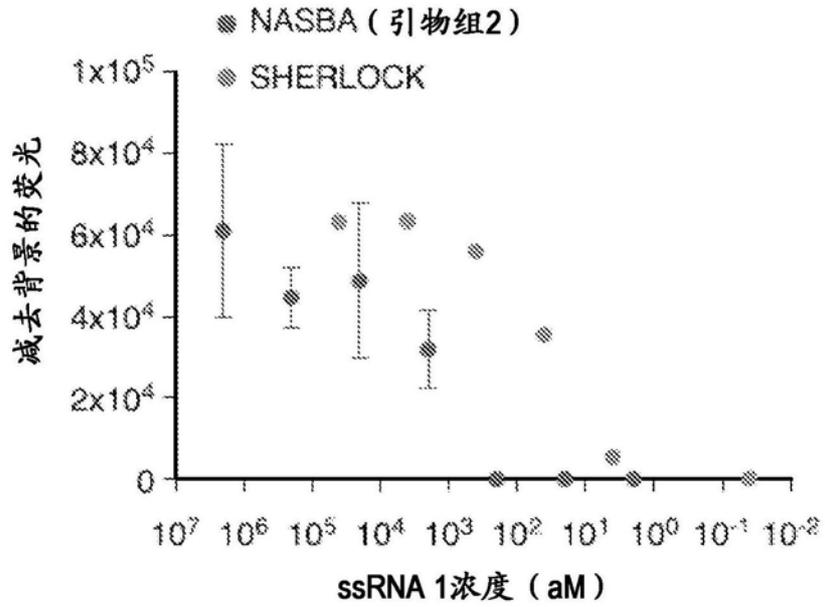


图53

A

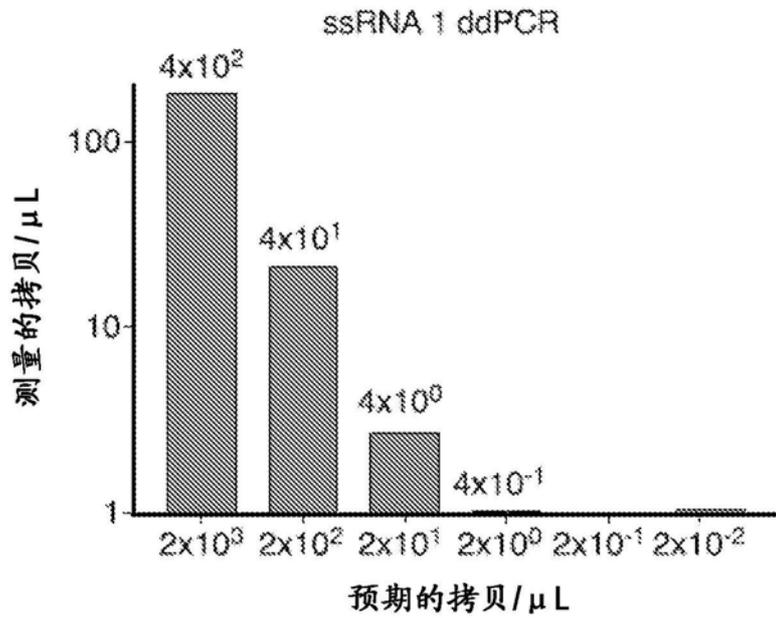
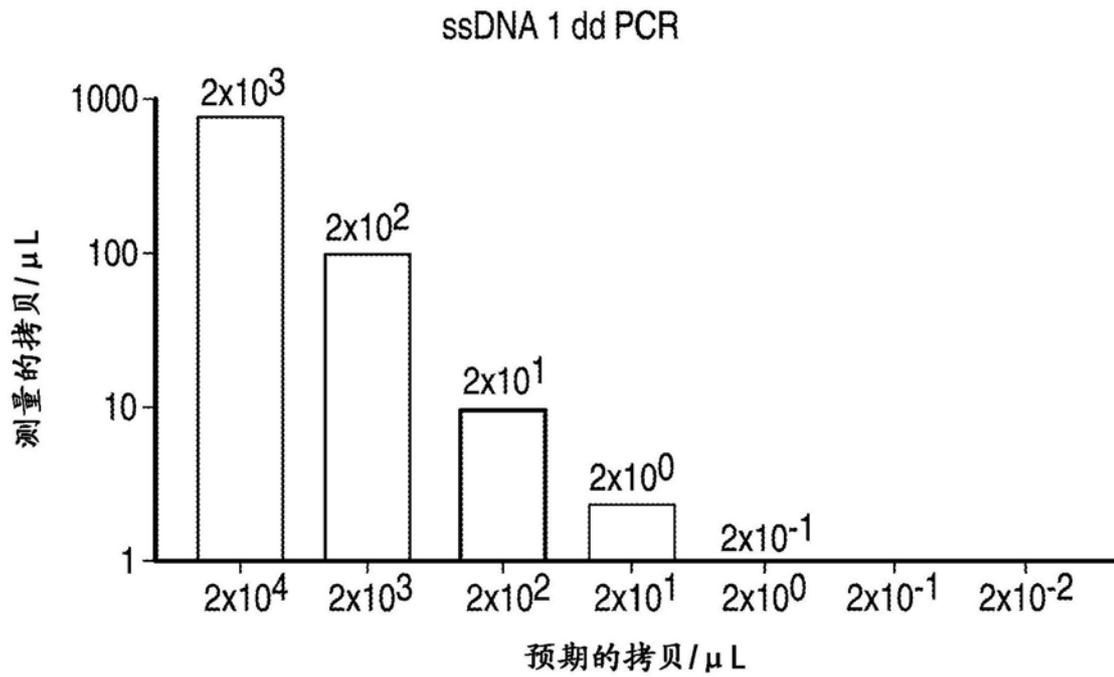


图54

B



C

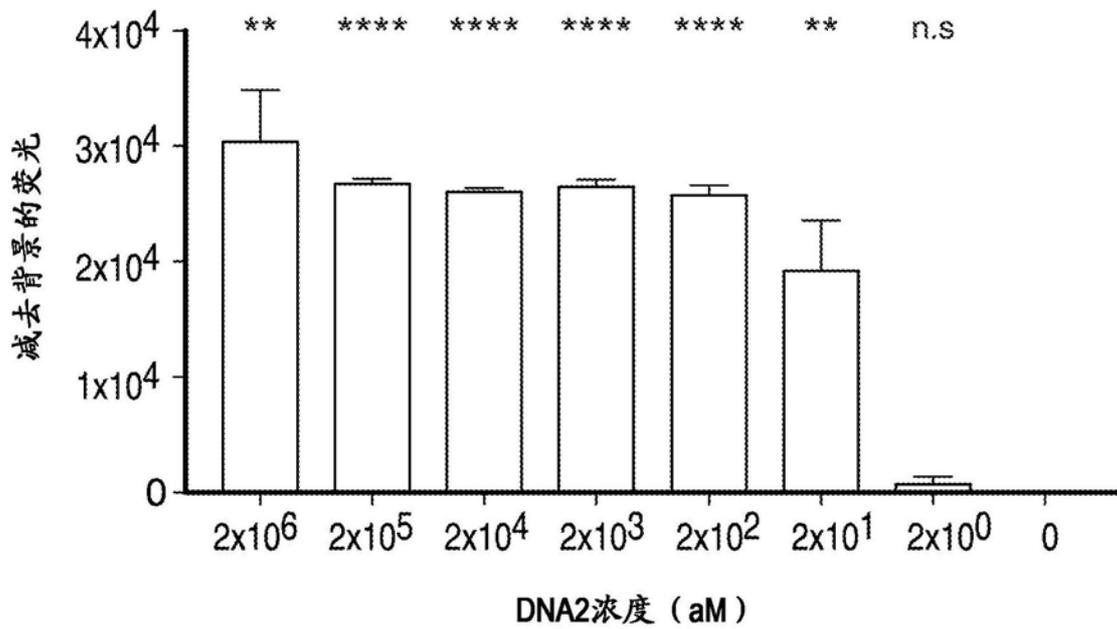


图54(续)

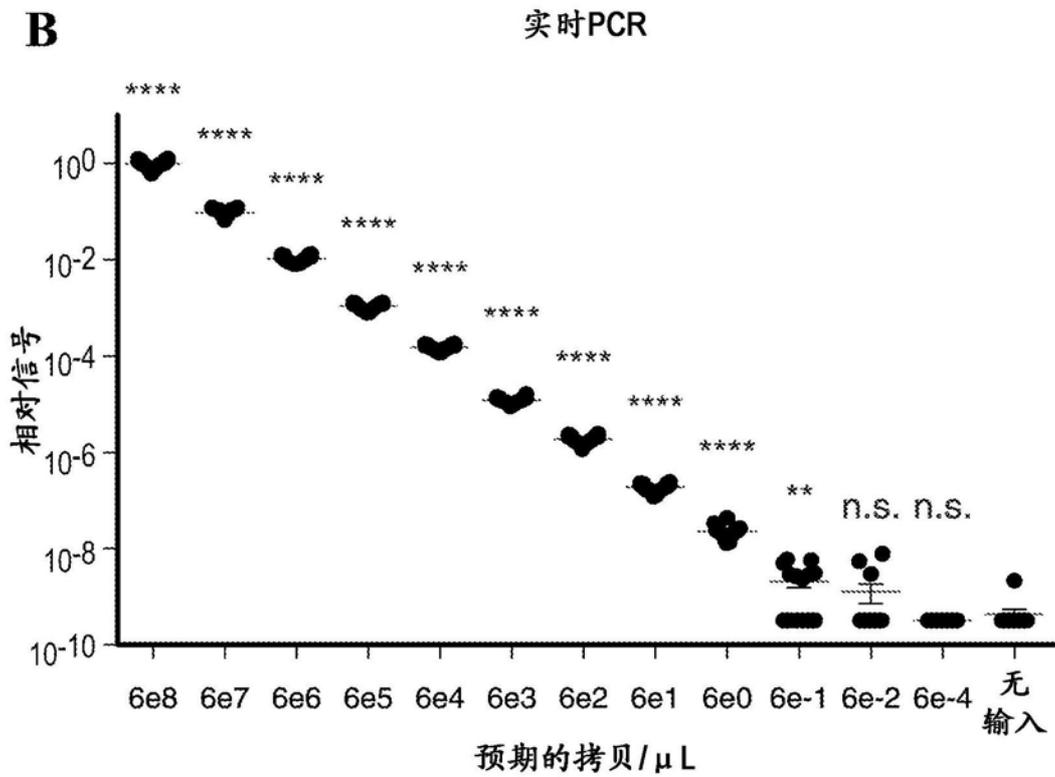
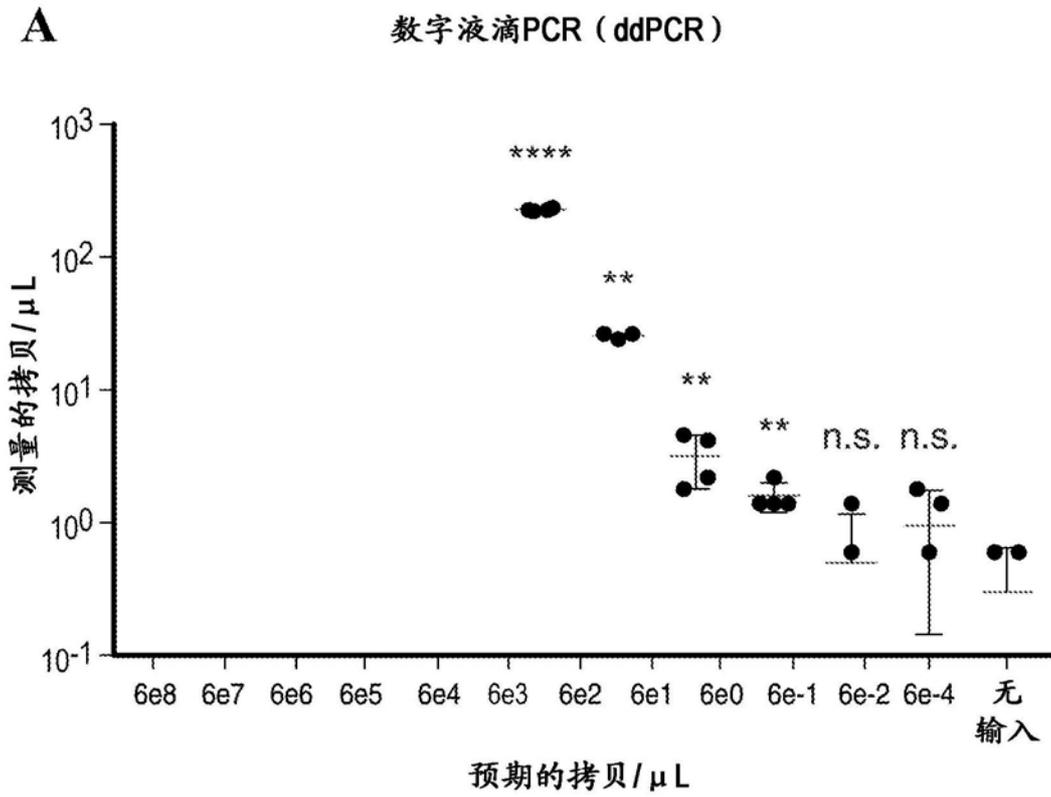


图55

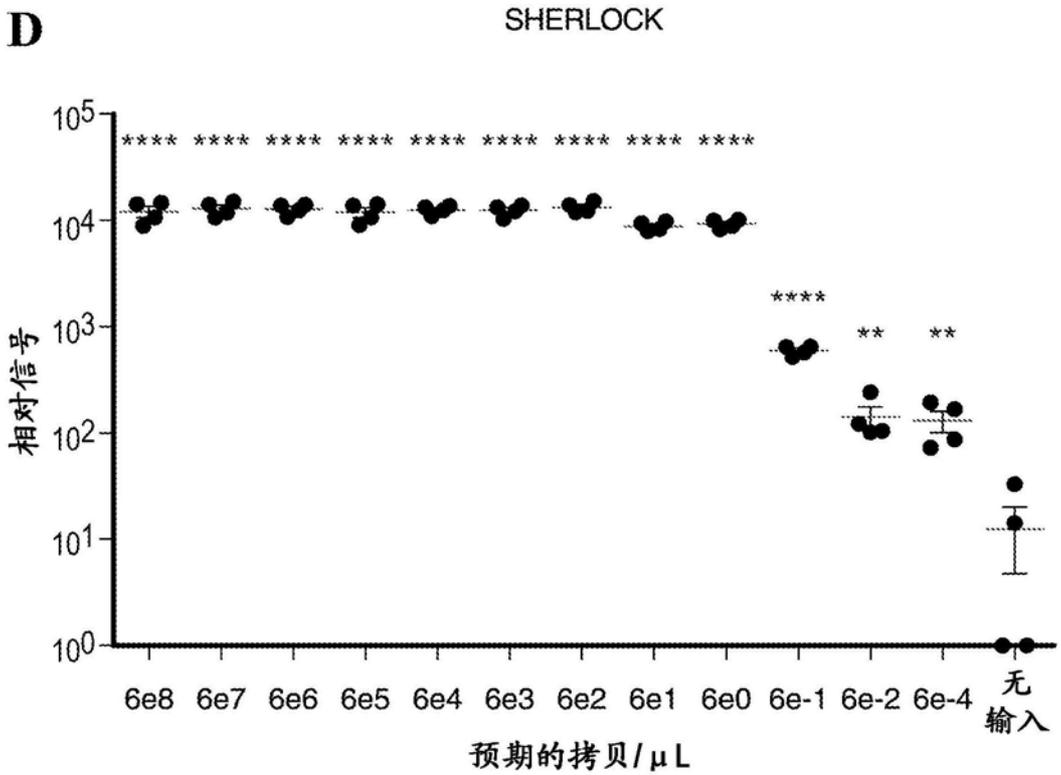
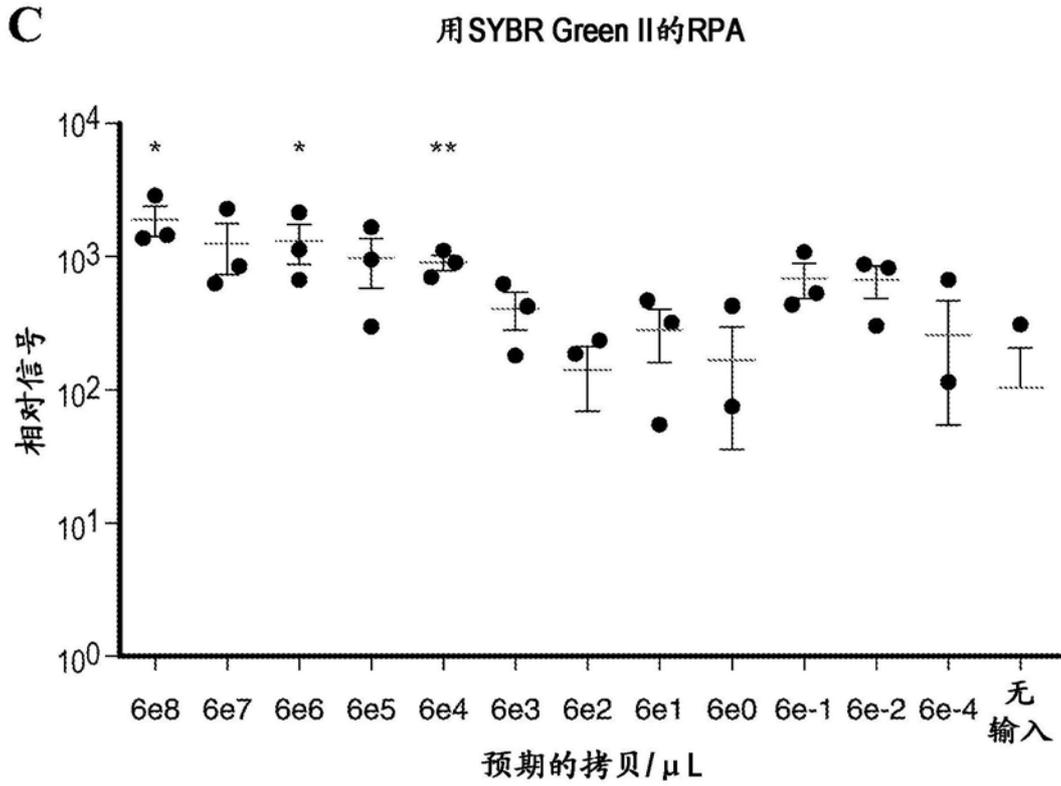
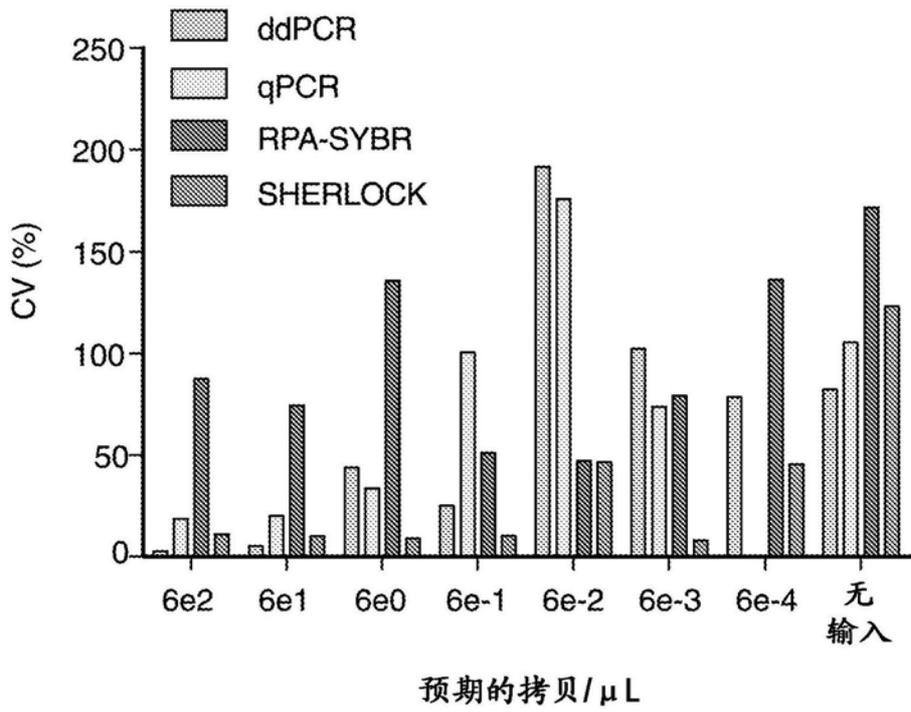


图55 (续)

E

变异系数



F

平均CV分析

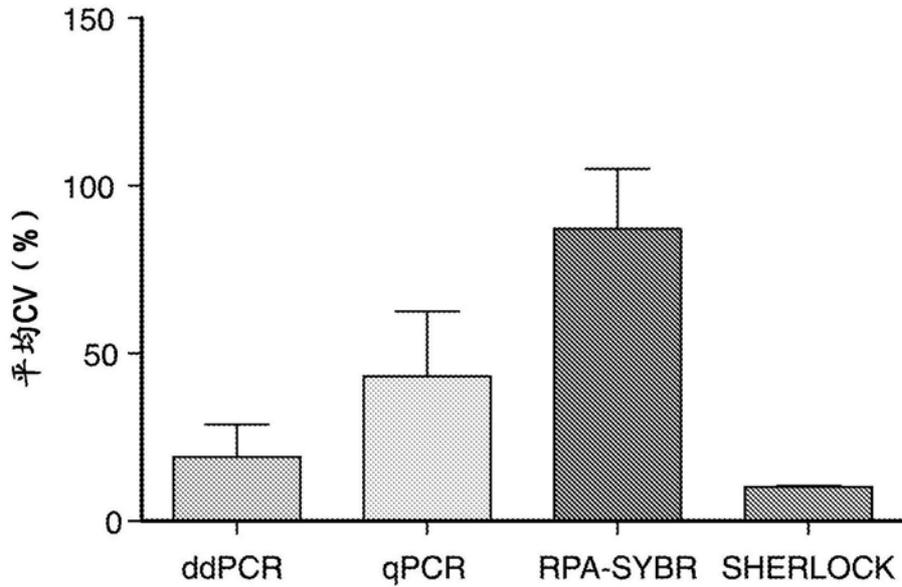


图55 (续)

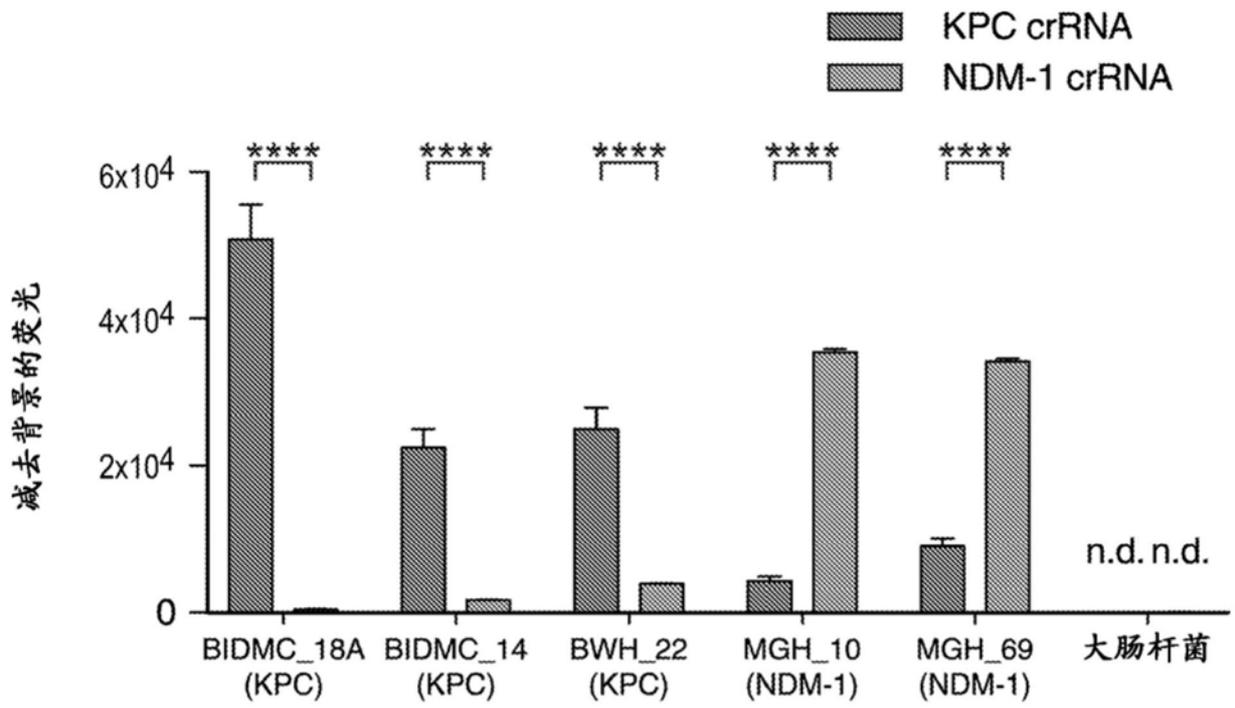


图56

A

```

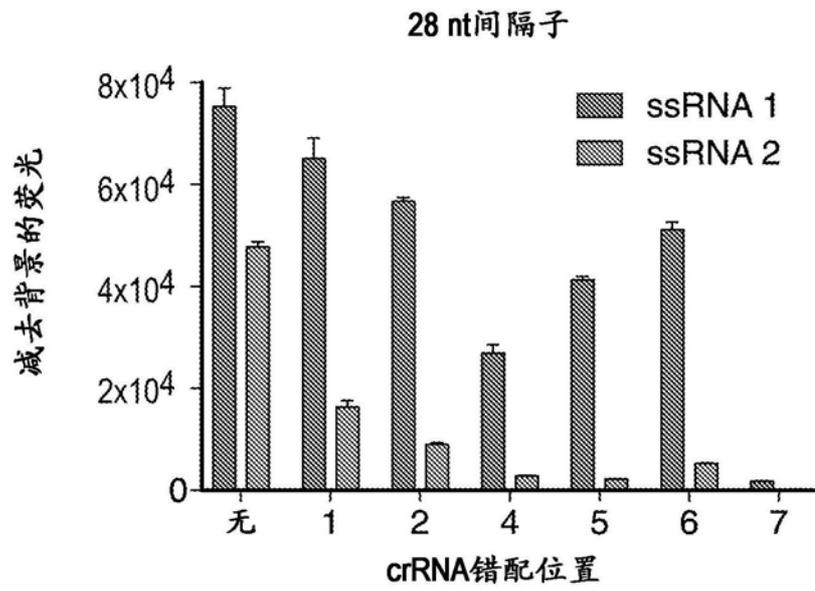
      UC
      A A
AAAACCCC GAUUUAG -5'
A   |||| ||||
GCAAGGGGACUAAAACUAGAUUGCUGUUCUACCAAGUAAUCCAU -3' 28nt
                                UAGAUUGCUGUUCUACCAAGUAA -3' 23nt
                                UAGAUUGCUGUUCUACCAAG -3' 20nt
    
```

3'-. .AUCUAACGACAAGAUGGUUCAUUAGGUA. . - 5' ssRNA 1
 3'-. .AUUAACGACAAGAUGGUUCAUUAGGUA. . - 5' ssRNA 2

UAGAUUGCUGUUCUACC. .	无	
AAGAUUGCUGUUCUACC. .	1	
UUGAUUGCUGUUCUACC. .	2	
UAGUUUGCUGUUCUACC. .	4	crRNA
UAGAAUGCUGUUCUACC. .	5	错配位置
UAGAUAGCUGUUCUACC. .	6	
UAGAUUCCUGUUCUACC. .	7	

图57

B



C

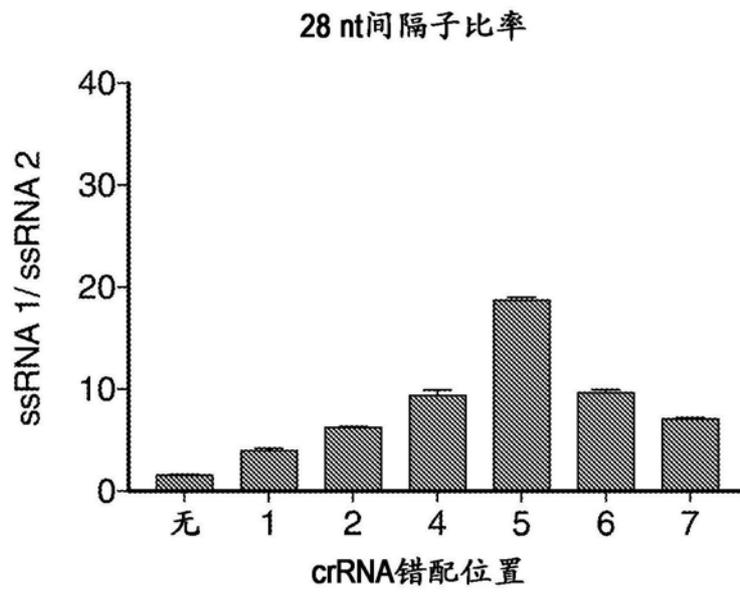
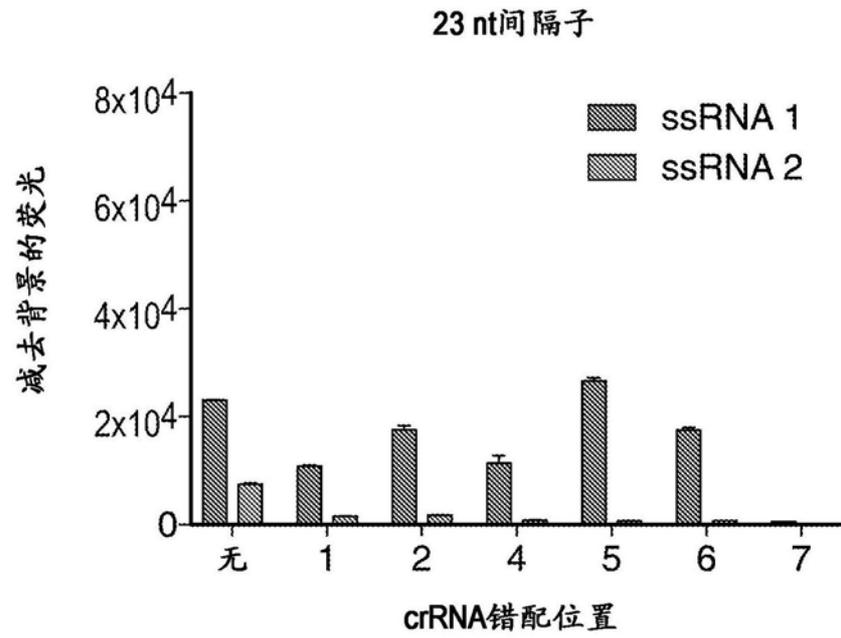


图57 (续)

D



E

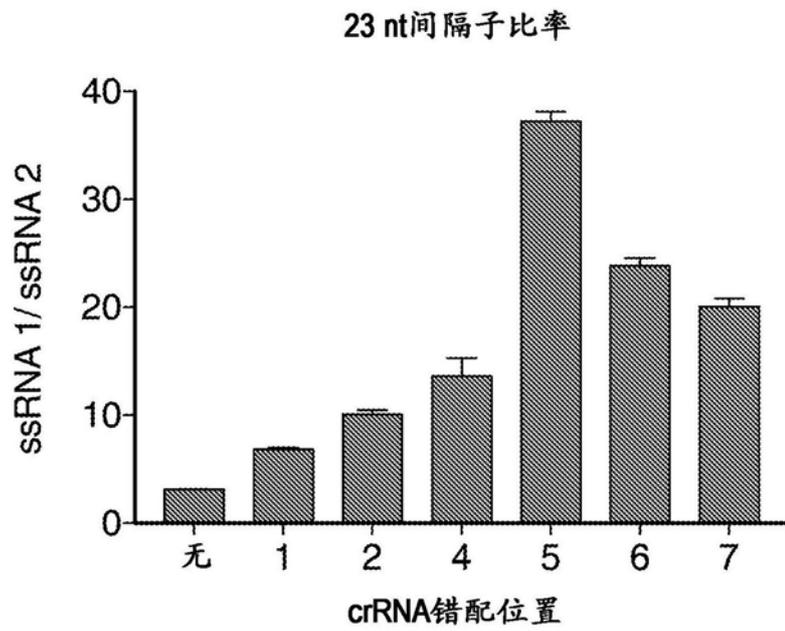
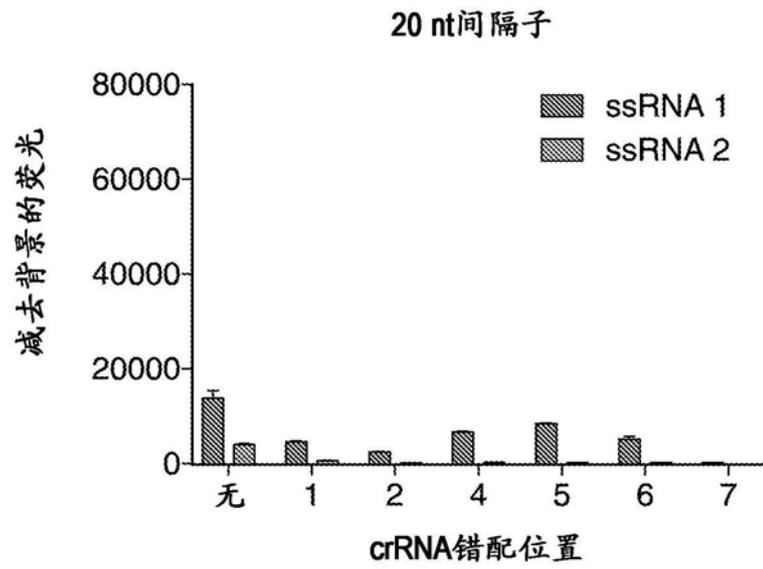


图57 (续)

F



G

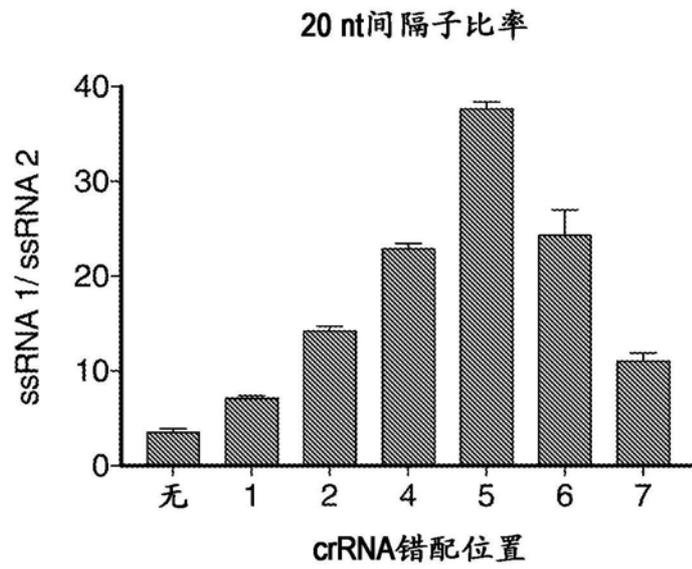


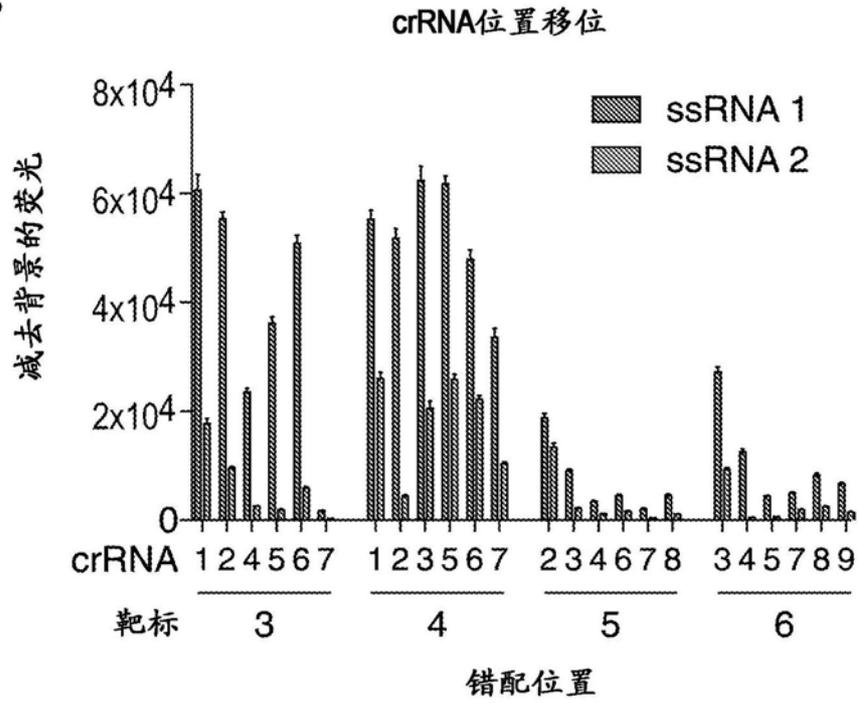
图57(续)

A



图58

B



C

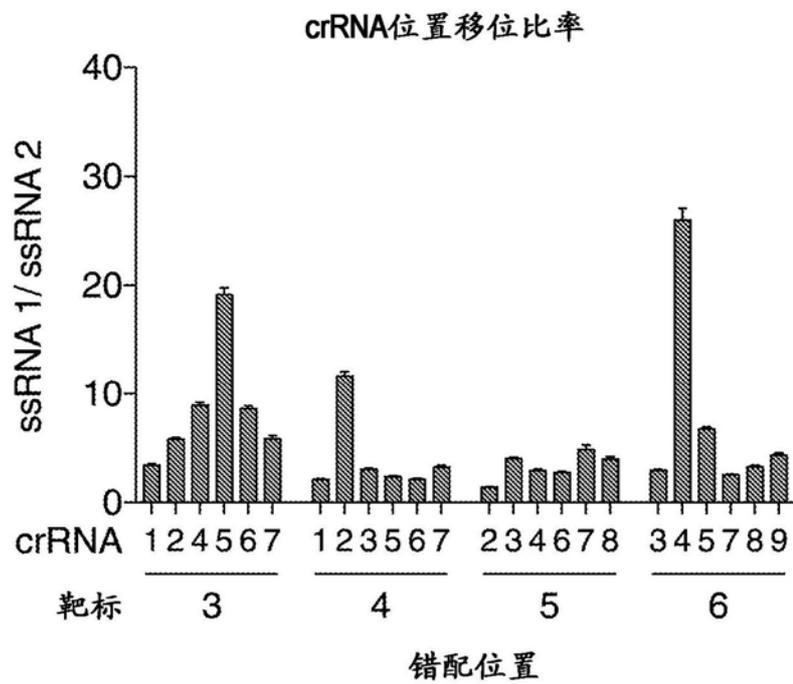


图58 (续)

B

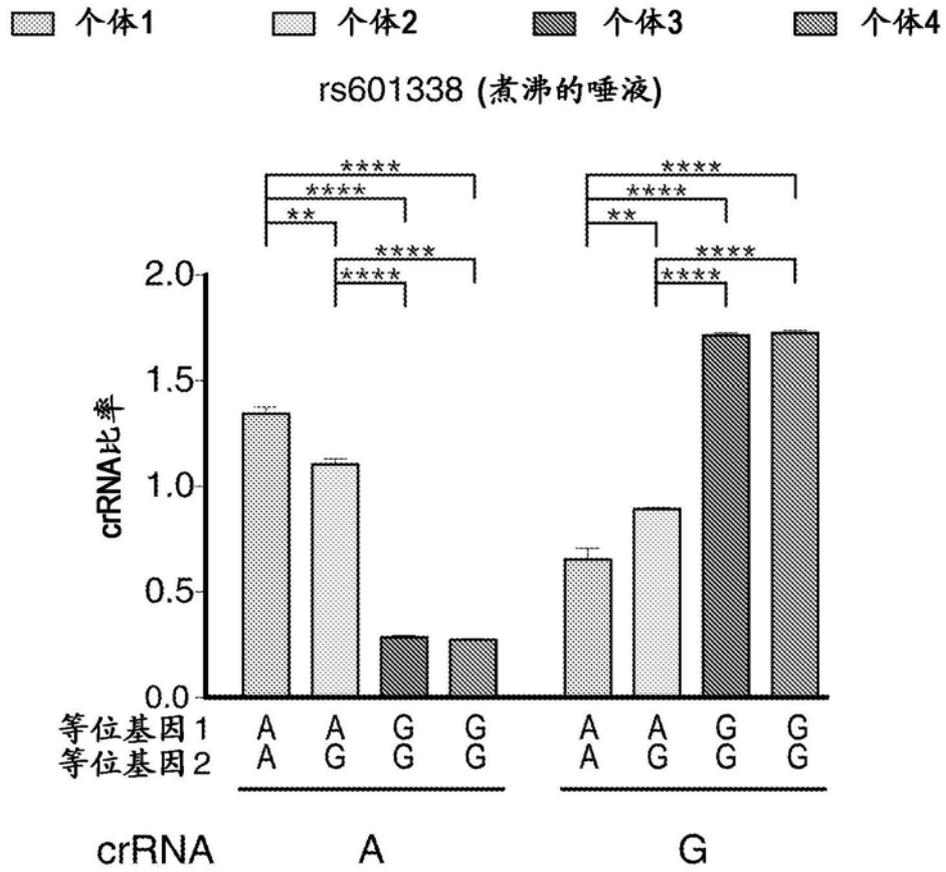
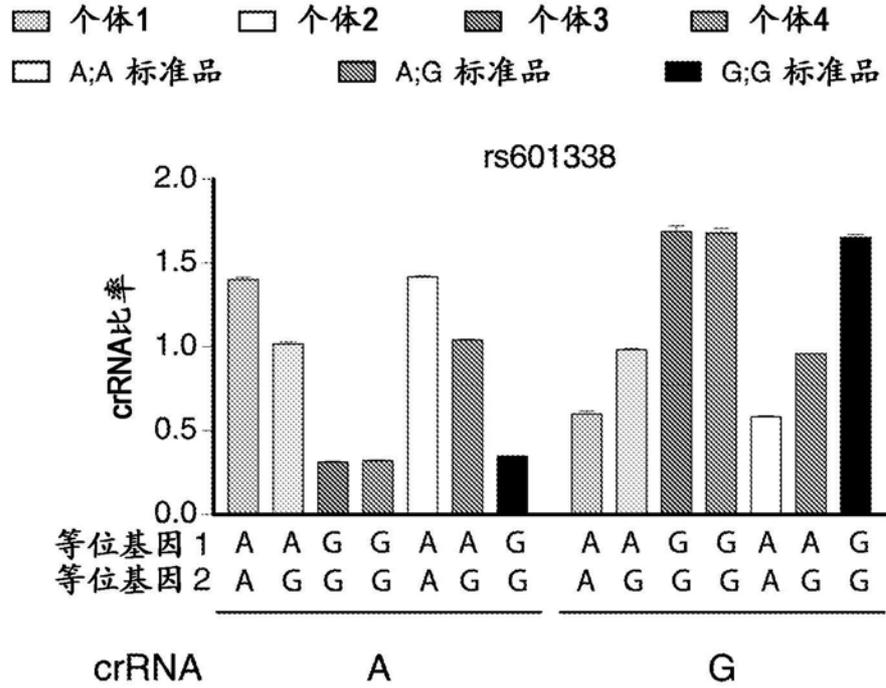


图59

A



B

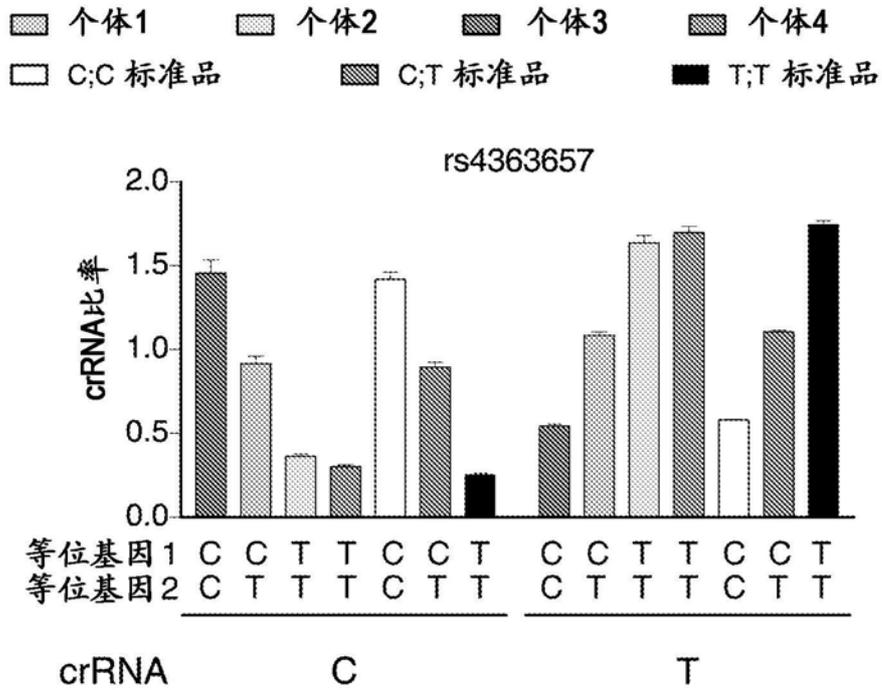


图60

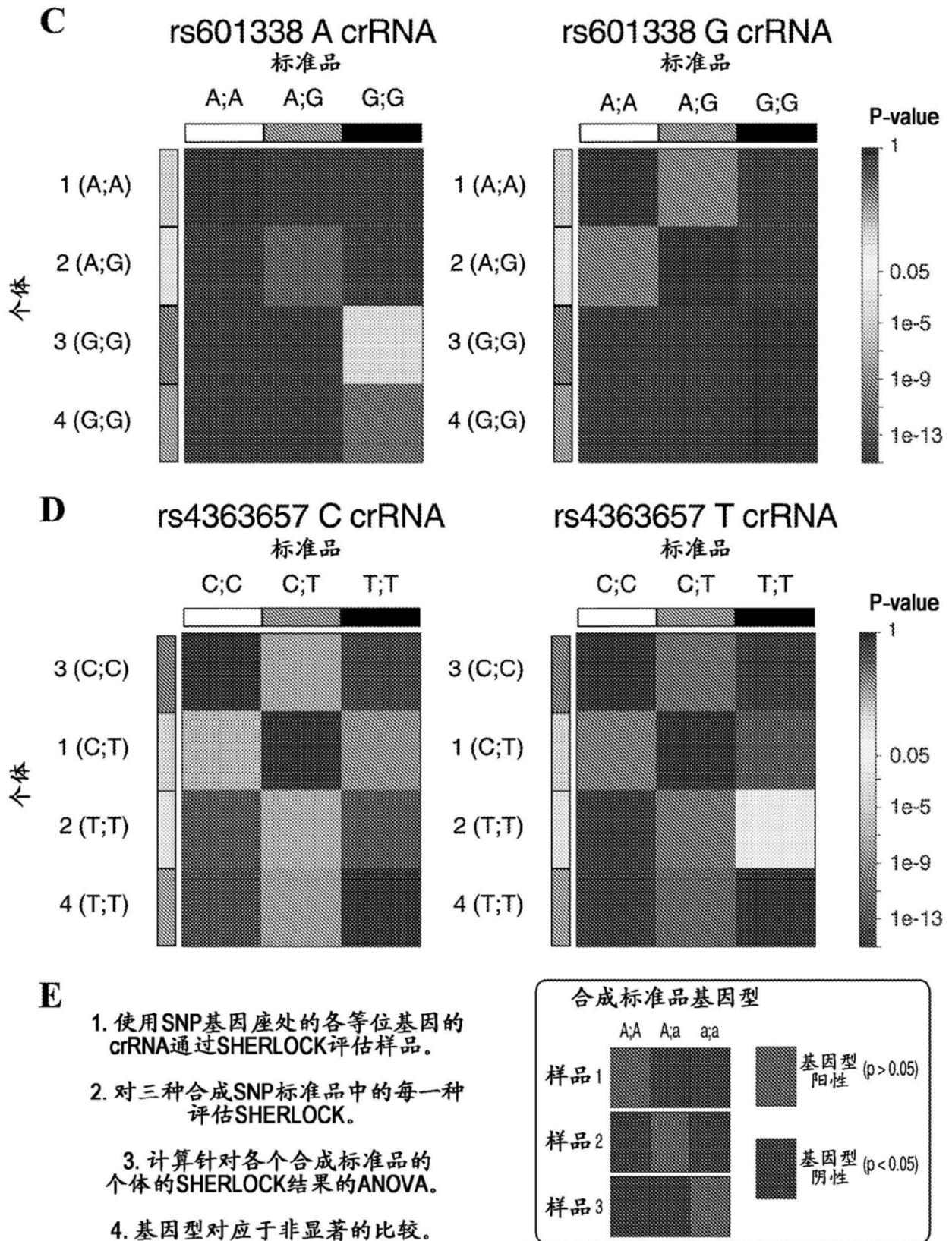


图60 (续)

具有基因组DNA背景的SHERLOCK

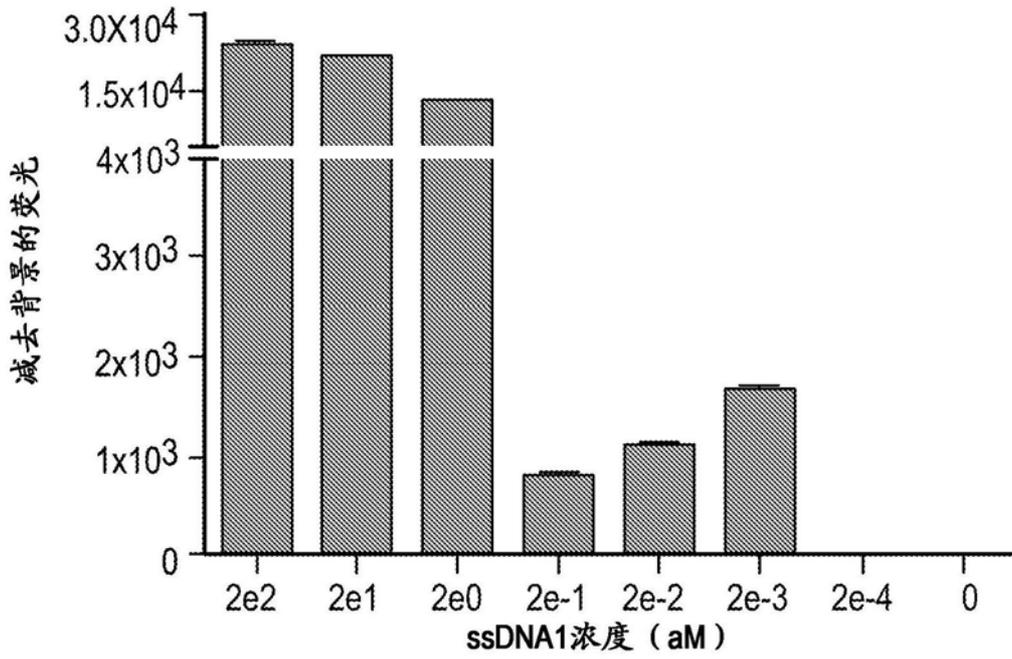


图61

直接尿液
2小时RPA, 3小时检测

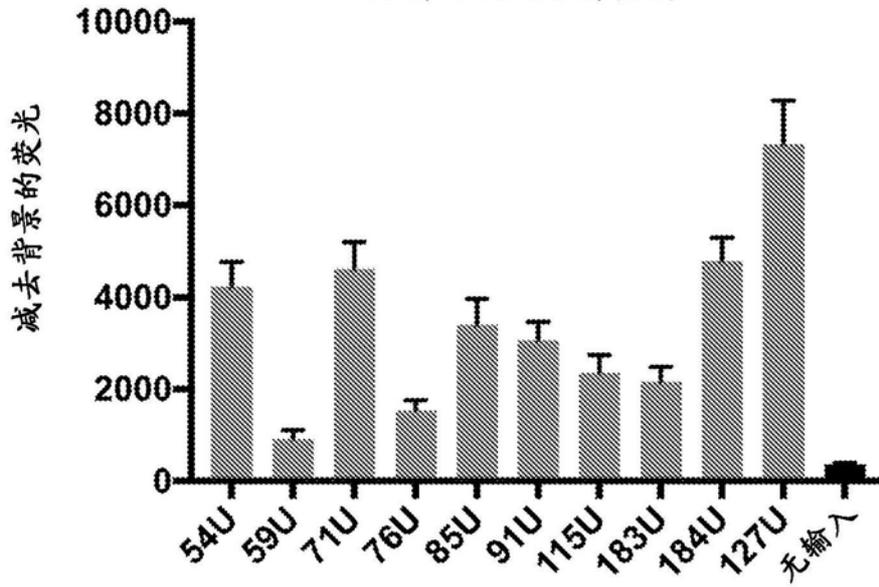


图62A

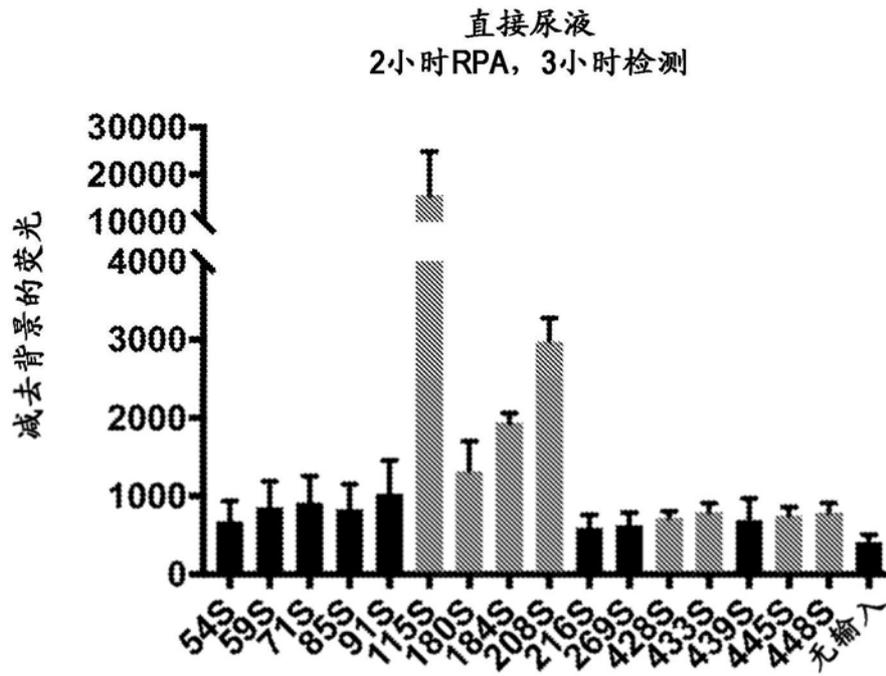


图62B

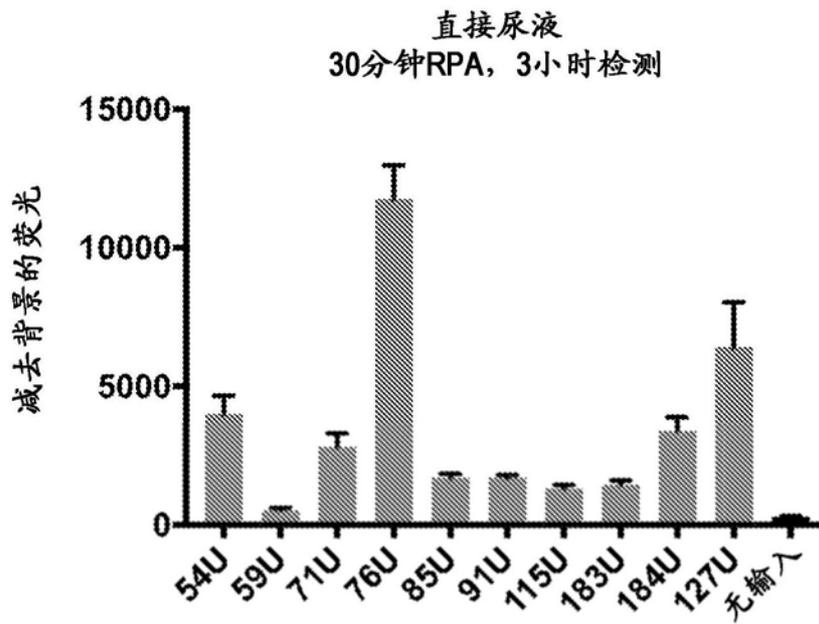


图63A

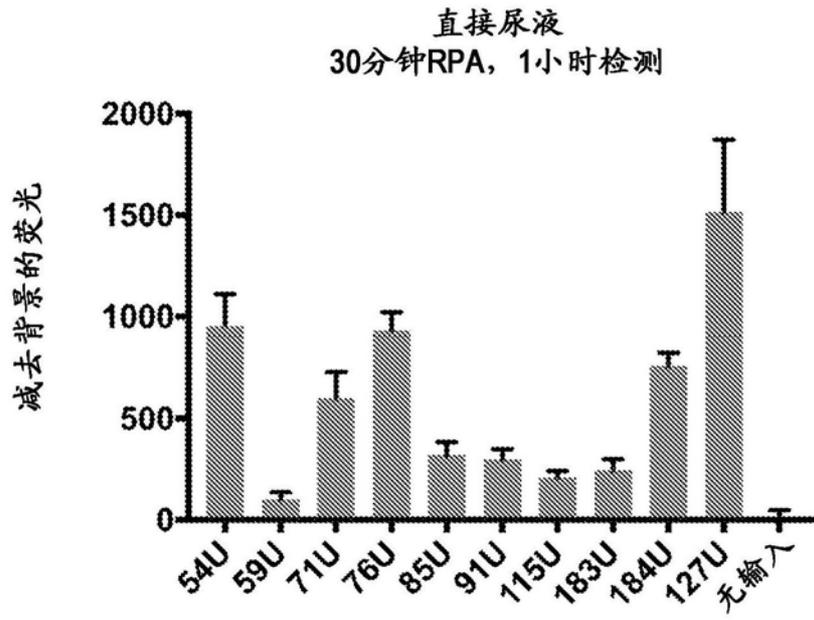


图63B

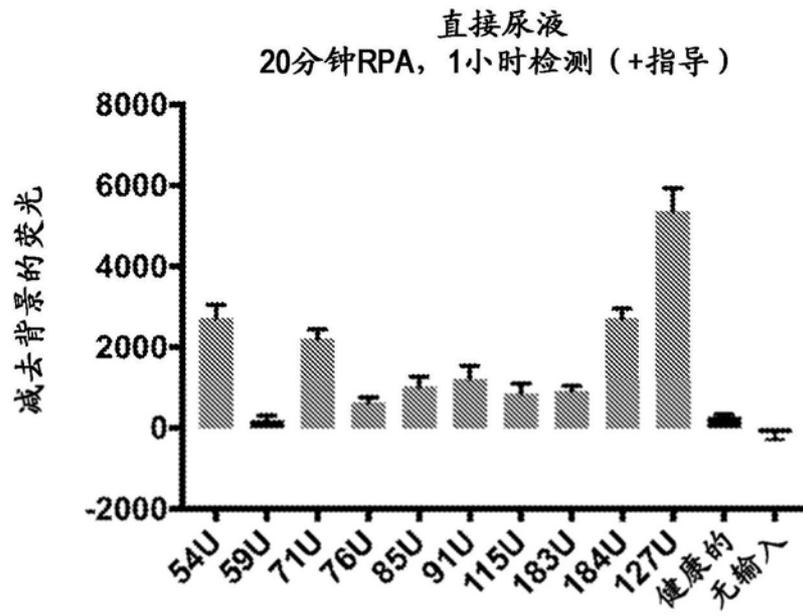


图64

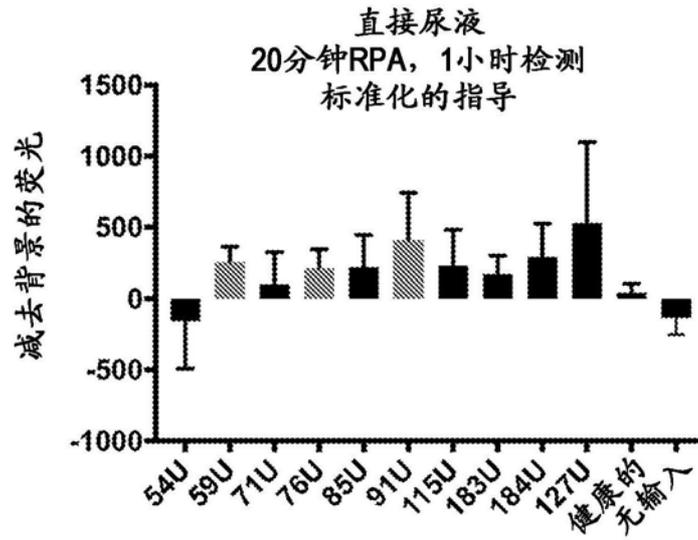


图65A

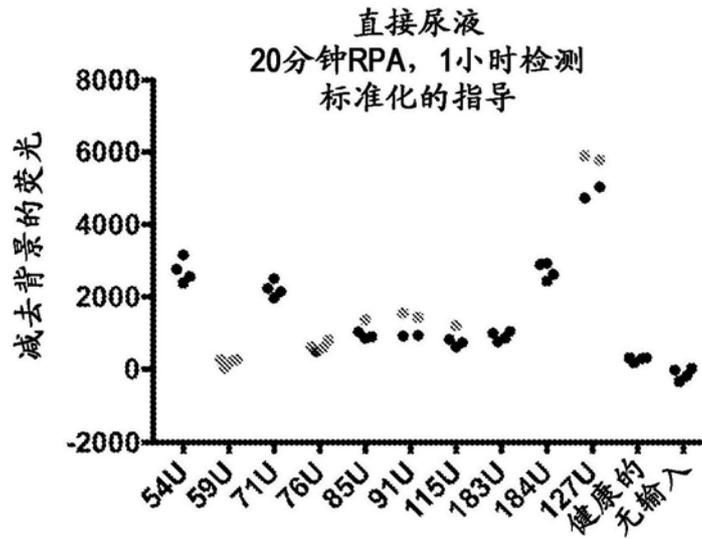
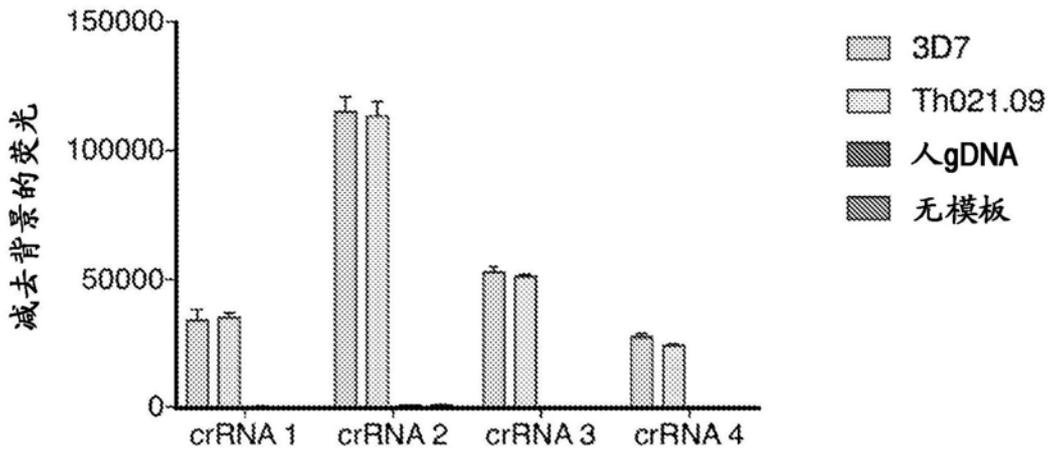


图65B



gaaattaatcagactcactatagggTTATTGCAATTATTAATCTTG AACGAGGAATG	<i>P. falciparum</i> 靶标1引物组1F
gaaattaatcagactcactatagggGATTGACAGATTAATAGCTCT TTCTTGATTTC	<i>P. falciparum</i> 靶标1引物组2F
ATTTTTCTTGCCAAACAATTCATCATATCTT	<i>P. falciparum</i> 靶标1引物组1R
TTCAATTCAAATAAGAATATAGTACTCGC	<i>P. falciparum</i> 靶标1引物组2R
gaaattaatcagactcactatagggTTCTTATTAGCAGAACAAAGA AGTTTAAACAAC	<i>P. falciparum</i> 靶标2引物组1F
gaaattaatcagactcactatagggATTTTATGCAATGTTAAAAAC TGTTCCAAGTA	<i>P. falciparum</i> 靶标2引物组2F
TAATTGACATCCAATCCATAATAAAGCATAGA	<i>P. falciparum</i> 靶标2引物组1R
GAATTATAGTTGTAAACTCTTTGTTCTGCT	<i>P. falciparum</i> 靶标2引物组2R
cctagtaagcatgattcatcagatttggttttagtccccttcggt tttgggtagtctaaatccccctatagtgagtcgtattaatttc	<i>P. falciparum</i> 靶标1引物组1crRNA
ggatggtgatgcatgcccgttttagtggtttagtccccttcggt tttgggtagtctaaatccccctatagtgagtcgtattaatttc	<i>P. falciparum</i> 靶标1引物组2crRNA
caattaaaaatgatitgggtgctagaggtttagtccccttcggt tttgggtagtctaaatccccctatagtgagtcgtattaatttc	<i>P. falciparum</i> 靶标2引物组1crRNA
gctggttagtaattgtattattatcatgttttagtccccttcggt tttgggtagtctaaatccccctatagtgagtcgtattaatttc	<i>P. falciparum</i> 靶标2引物组2crRNA

图66

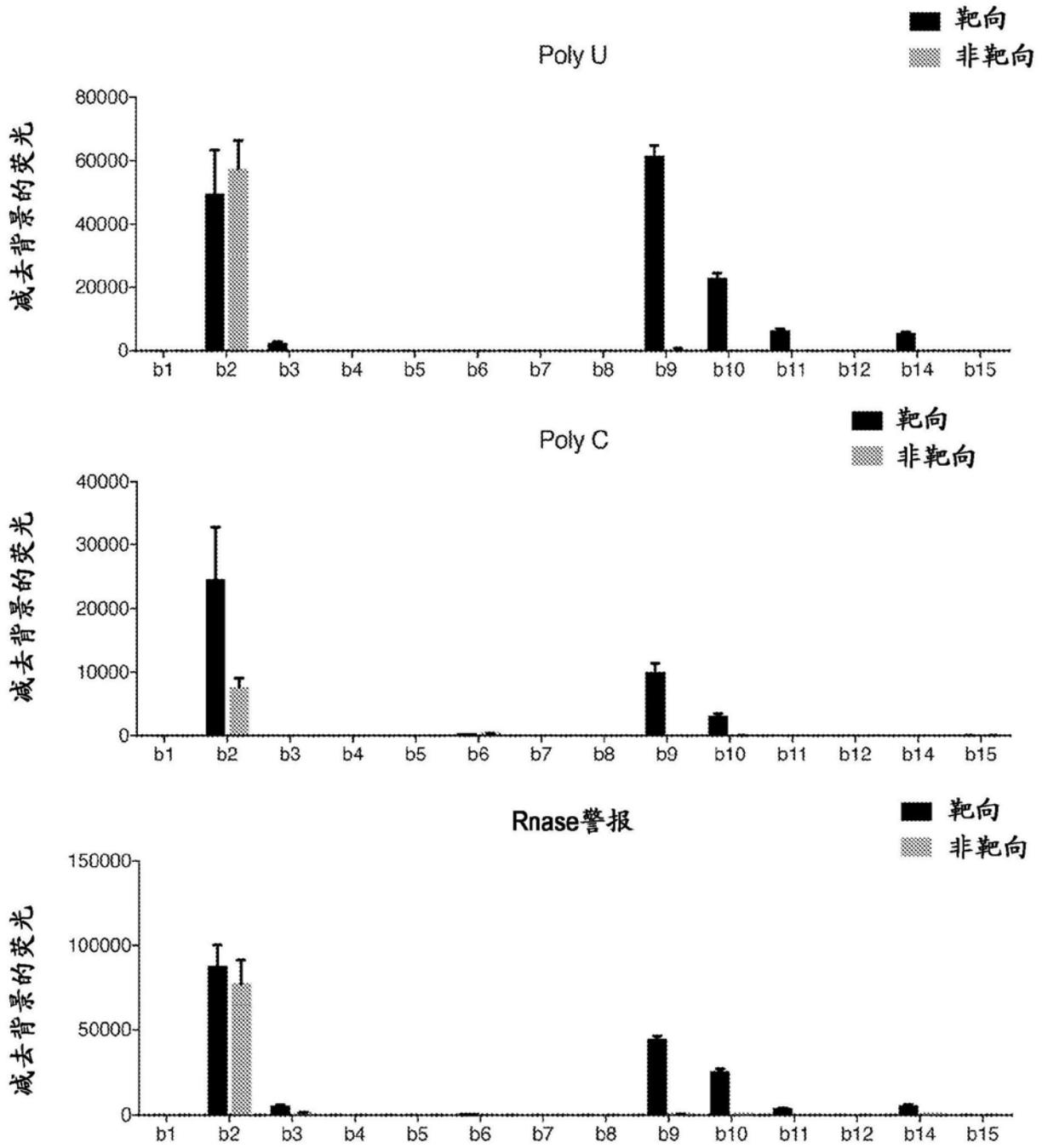


图67

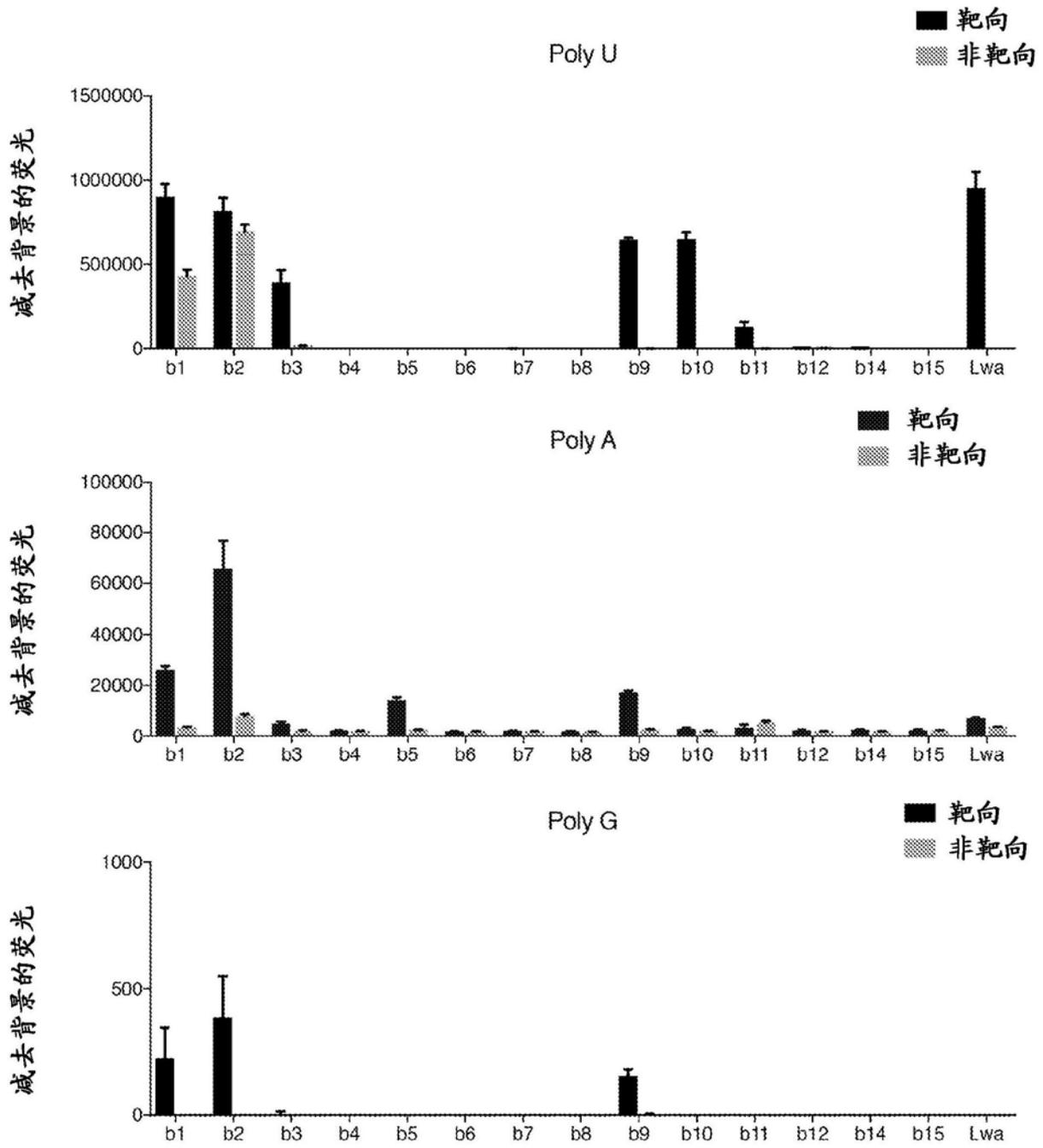


图67(续)

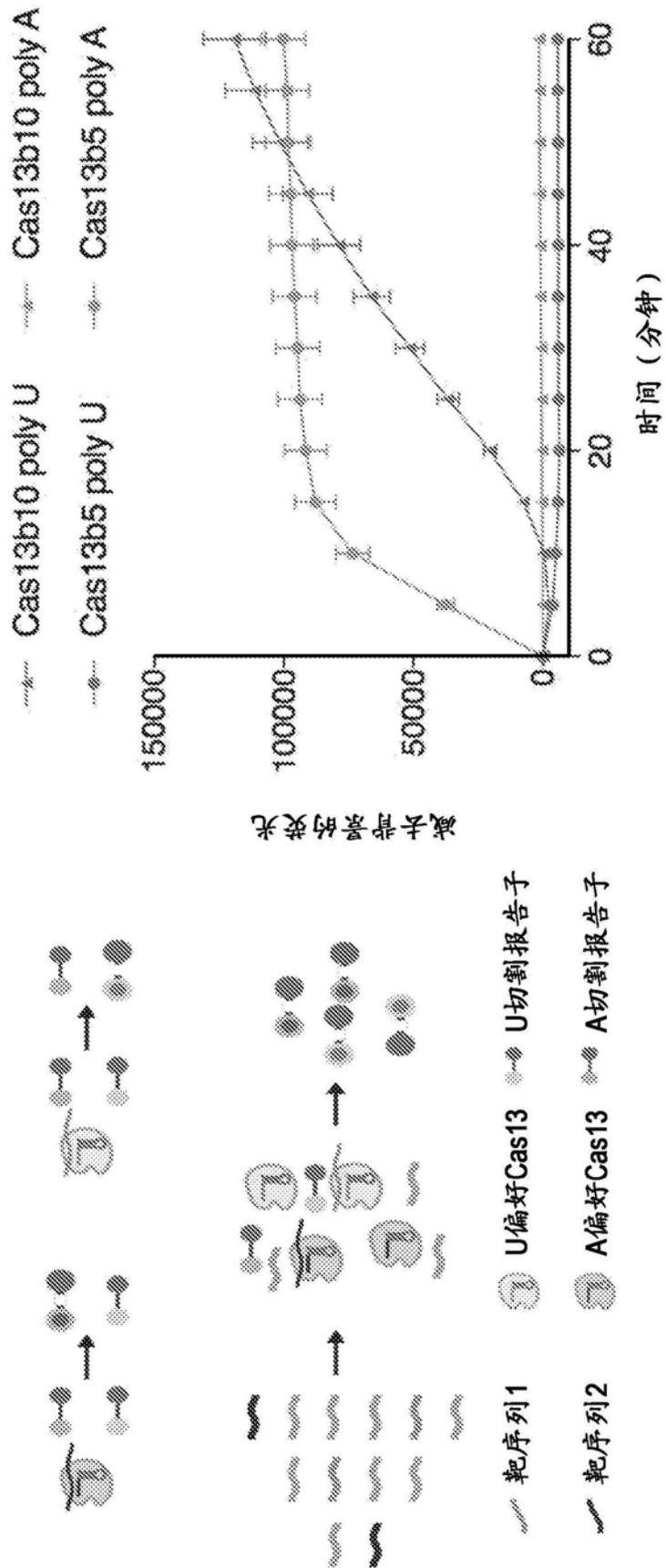


图68

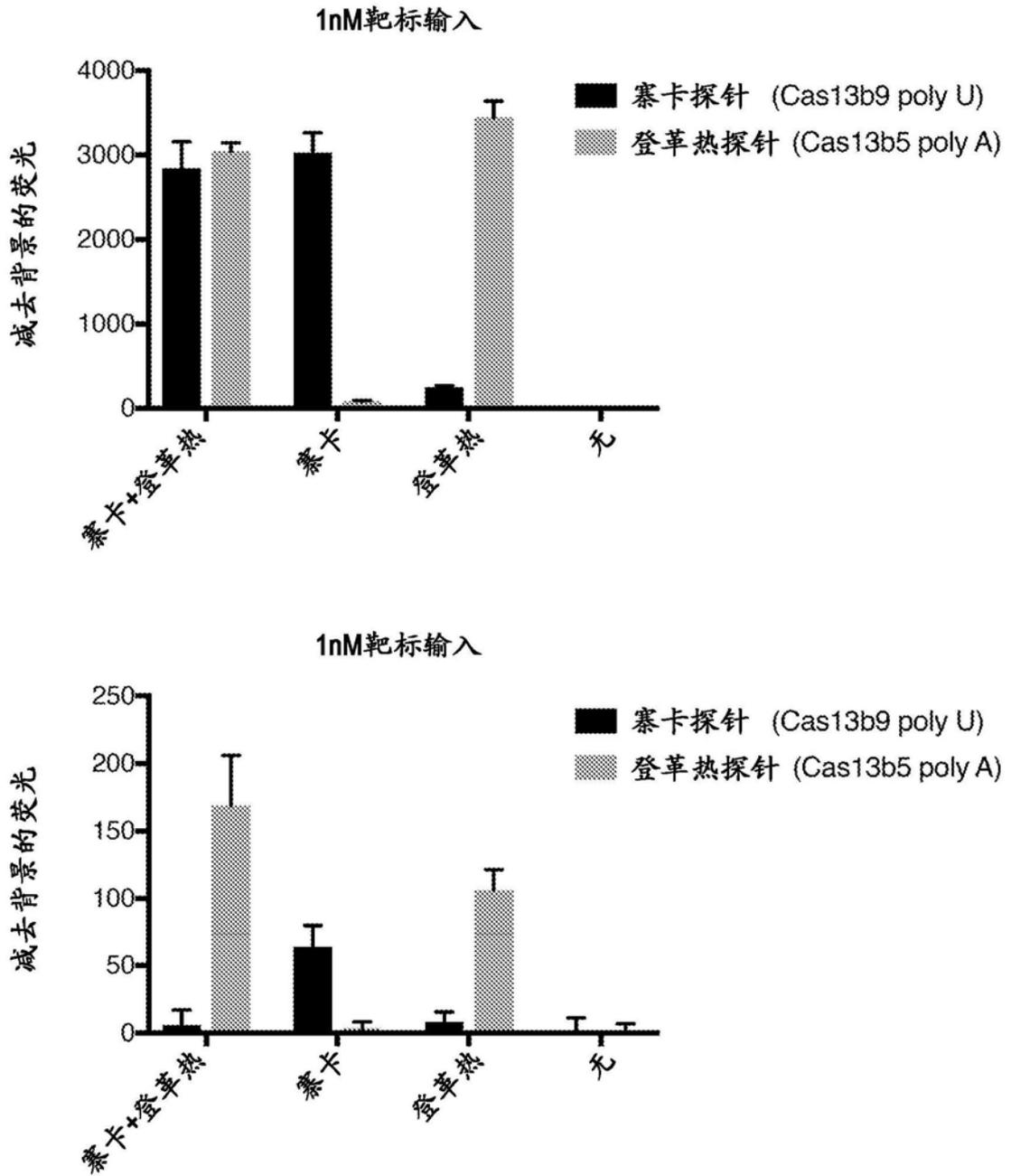


图69

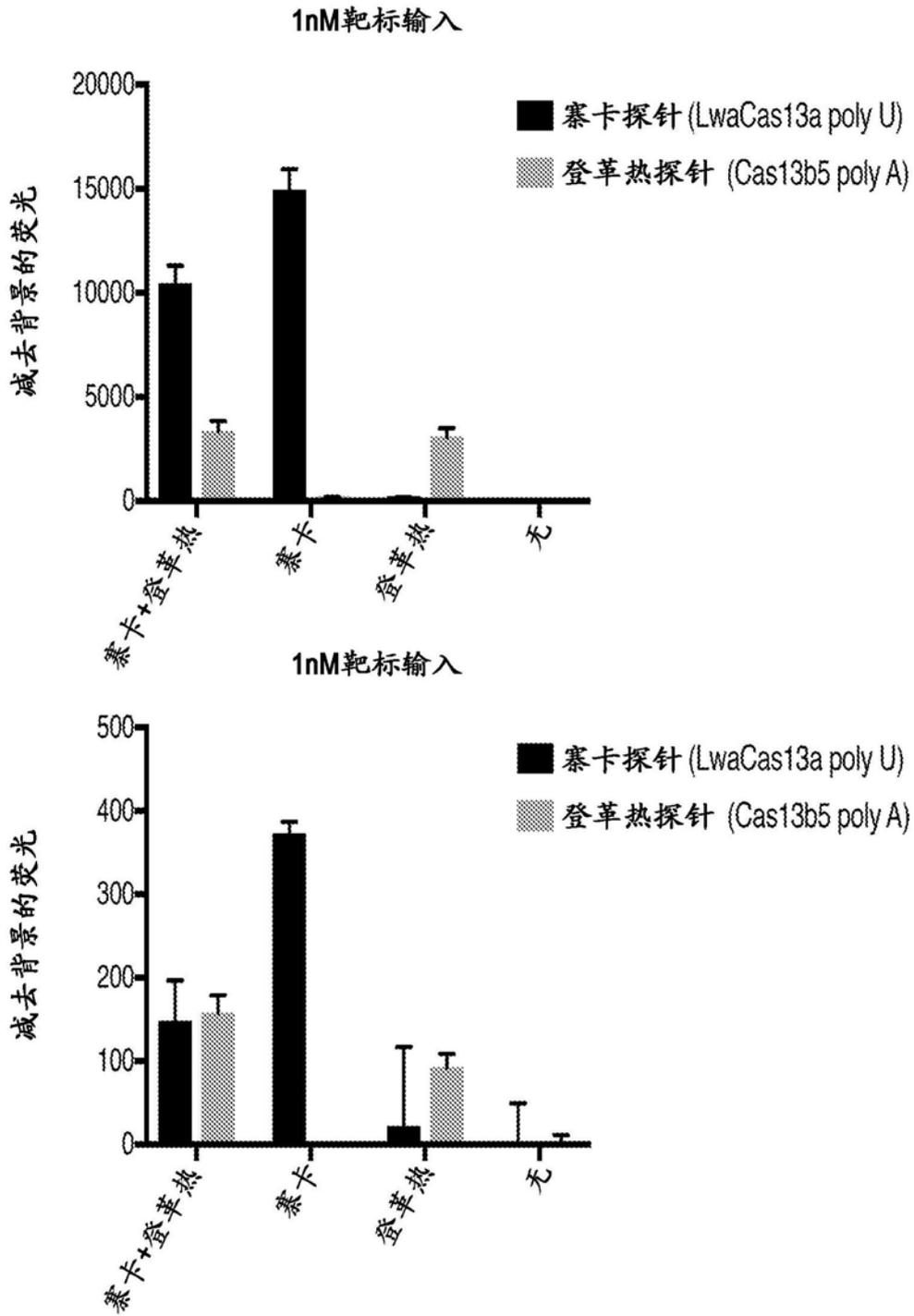


图70

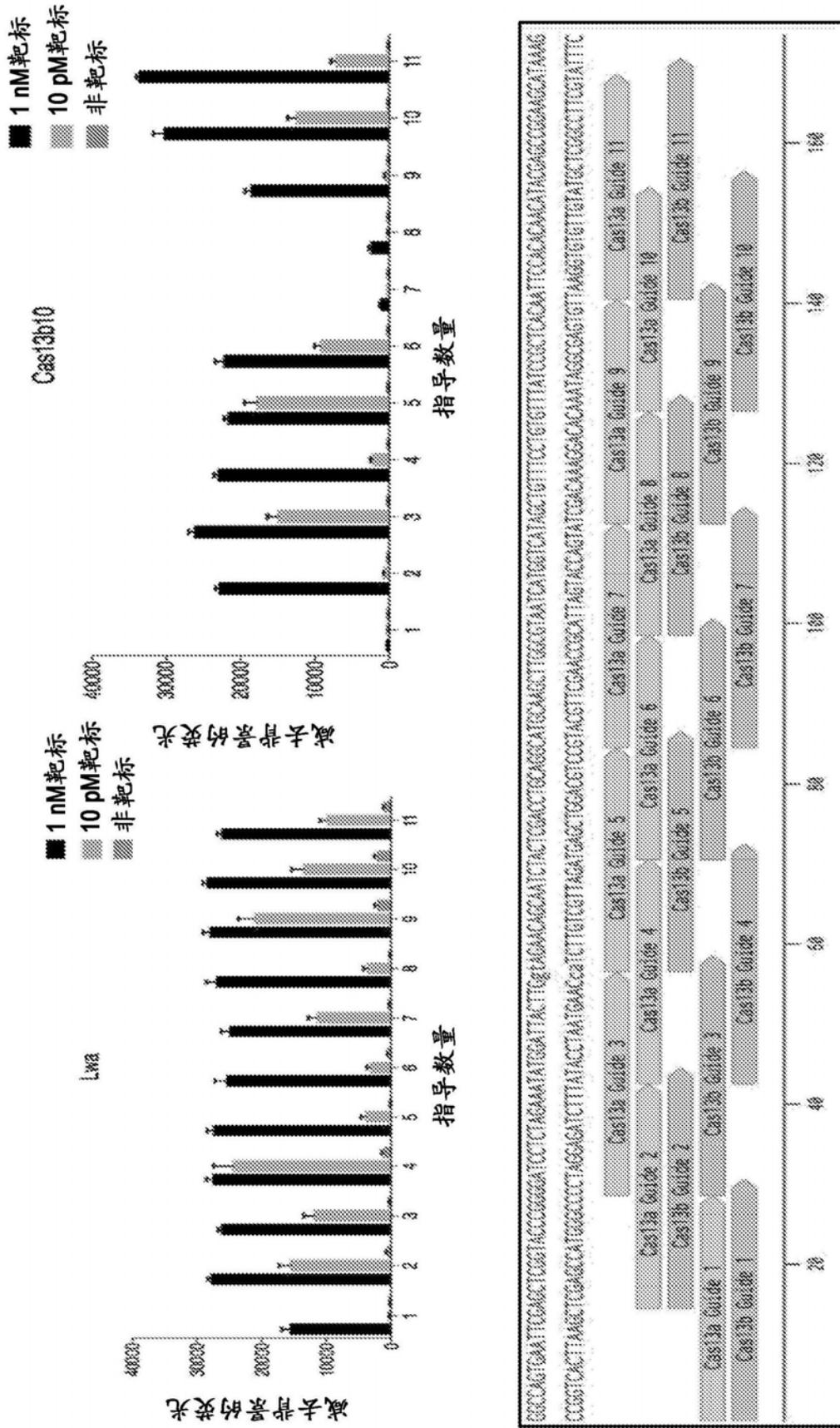


图71

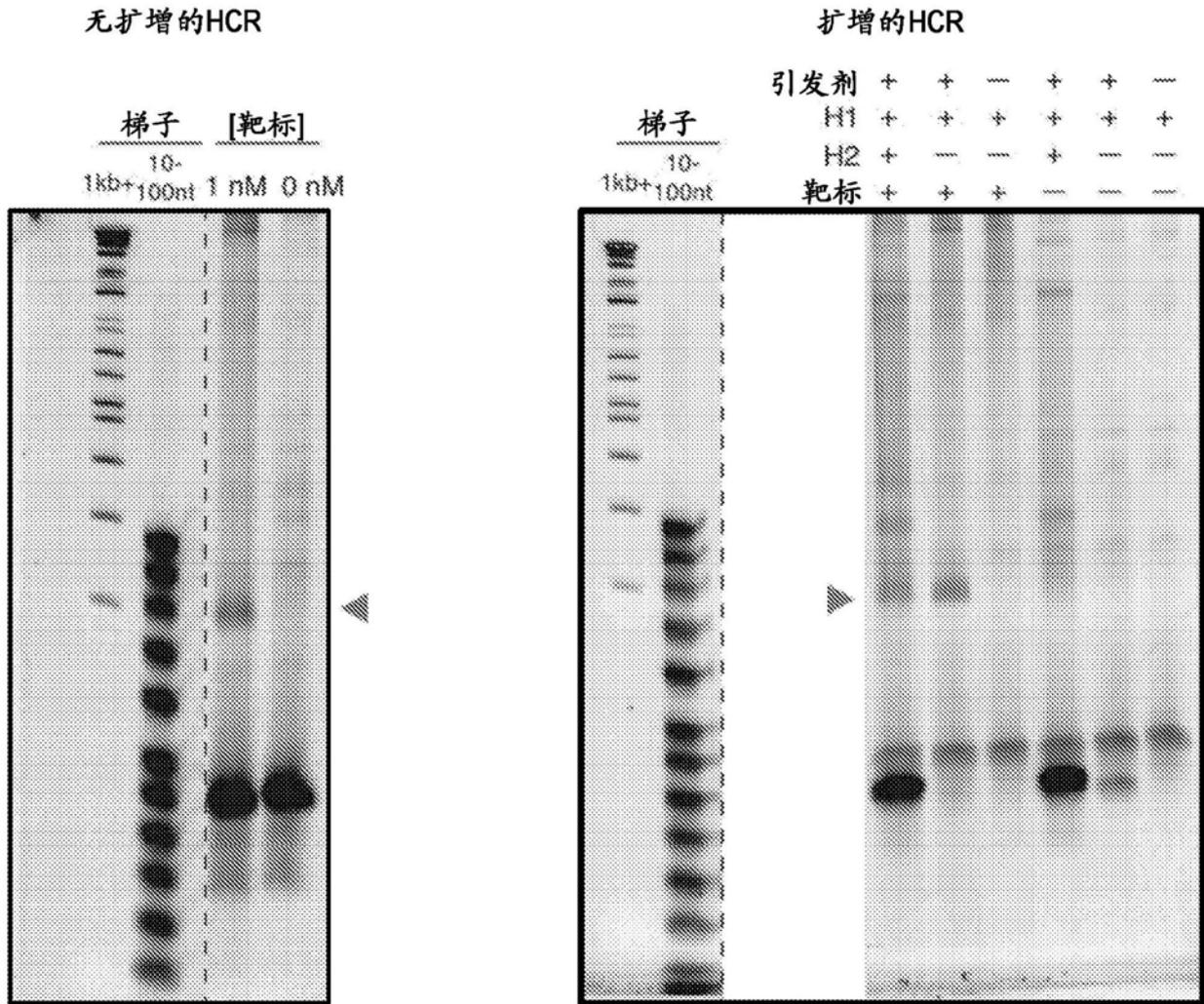


图72

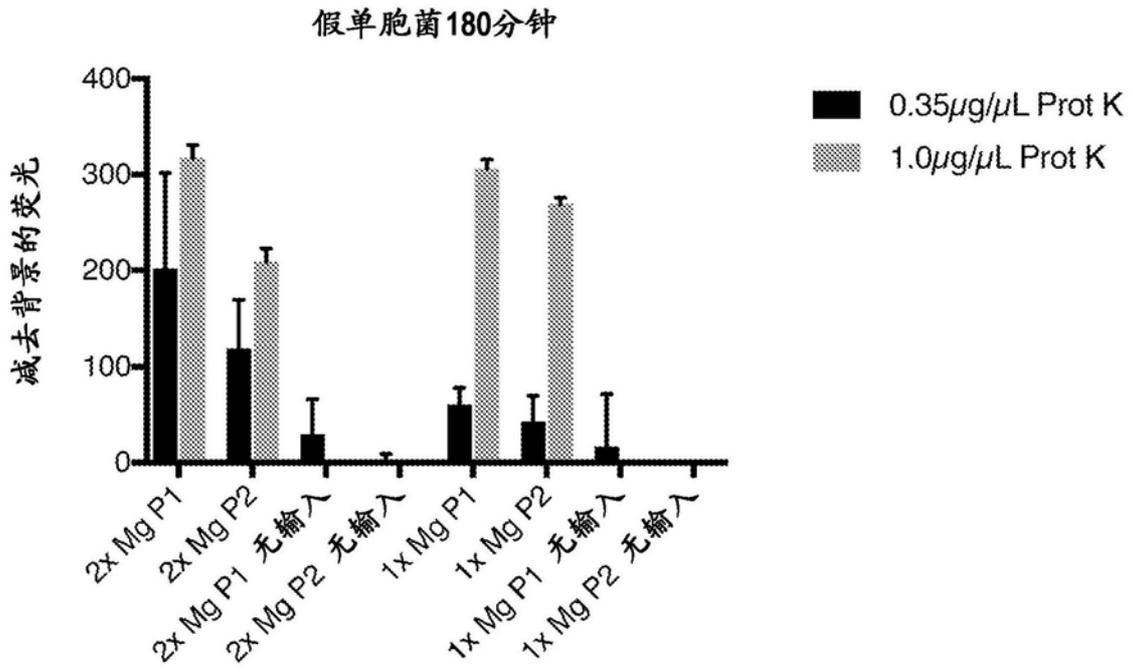


图73

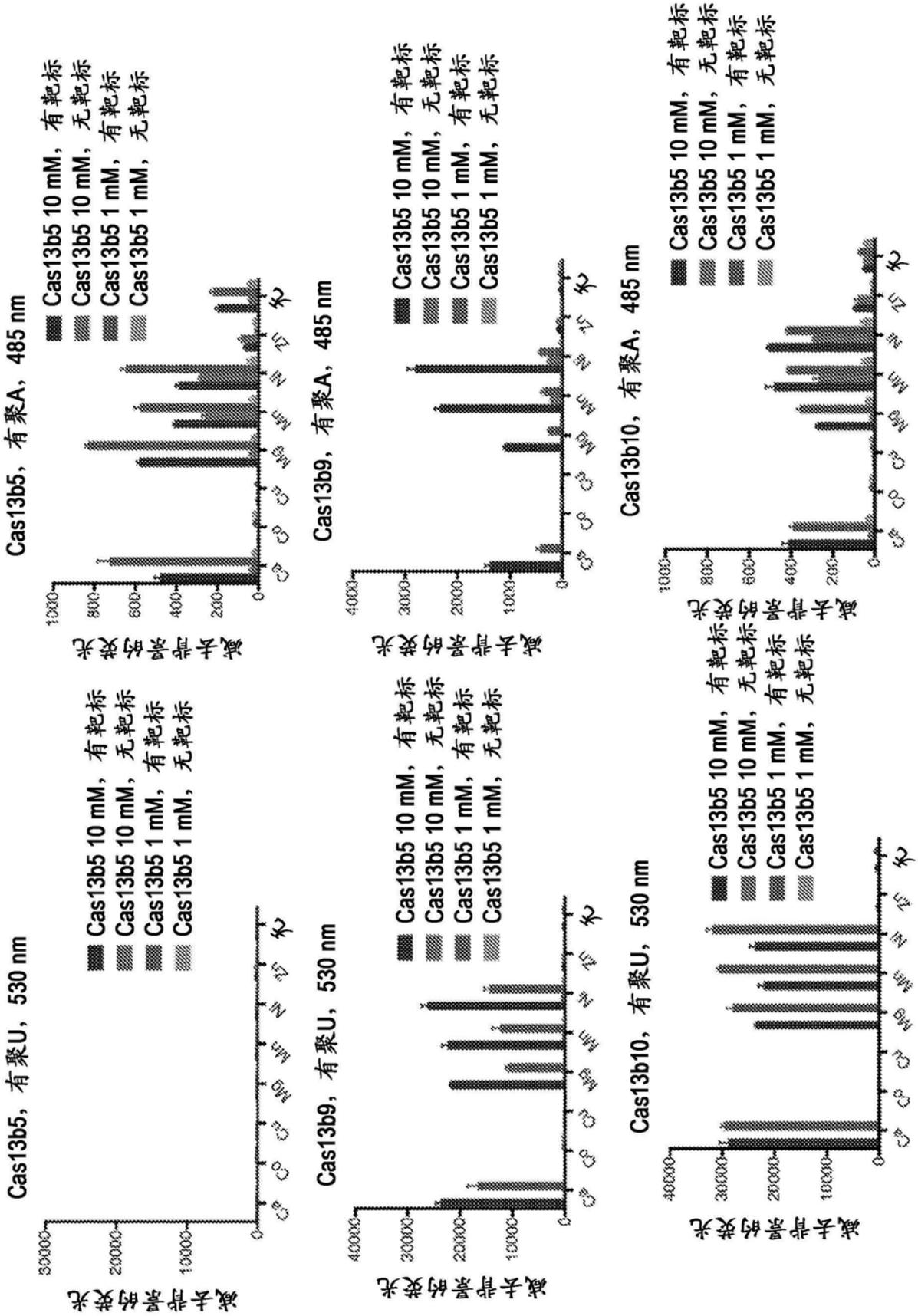


图74

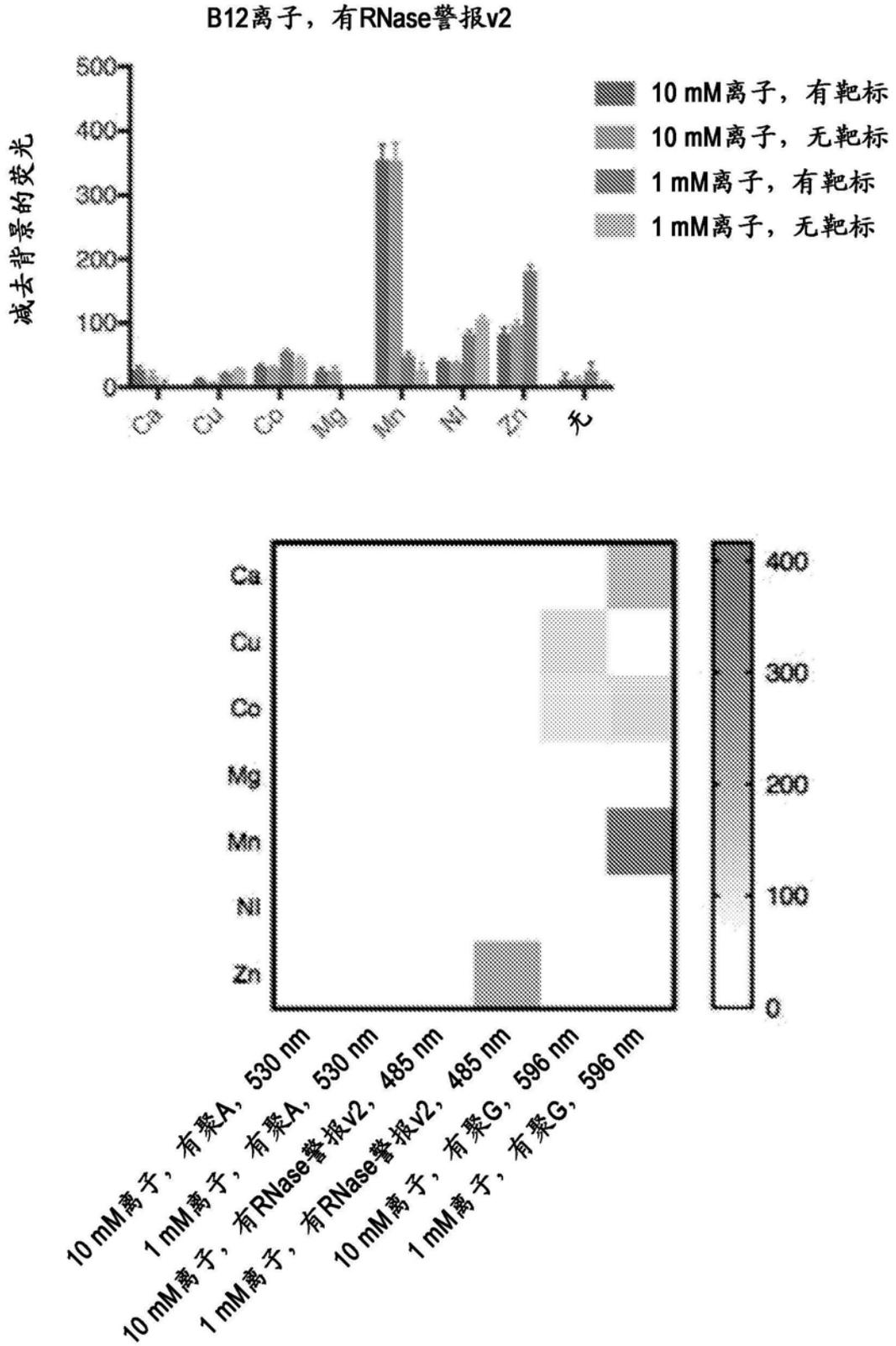


图75

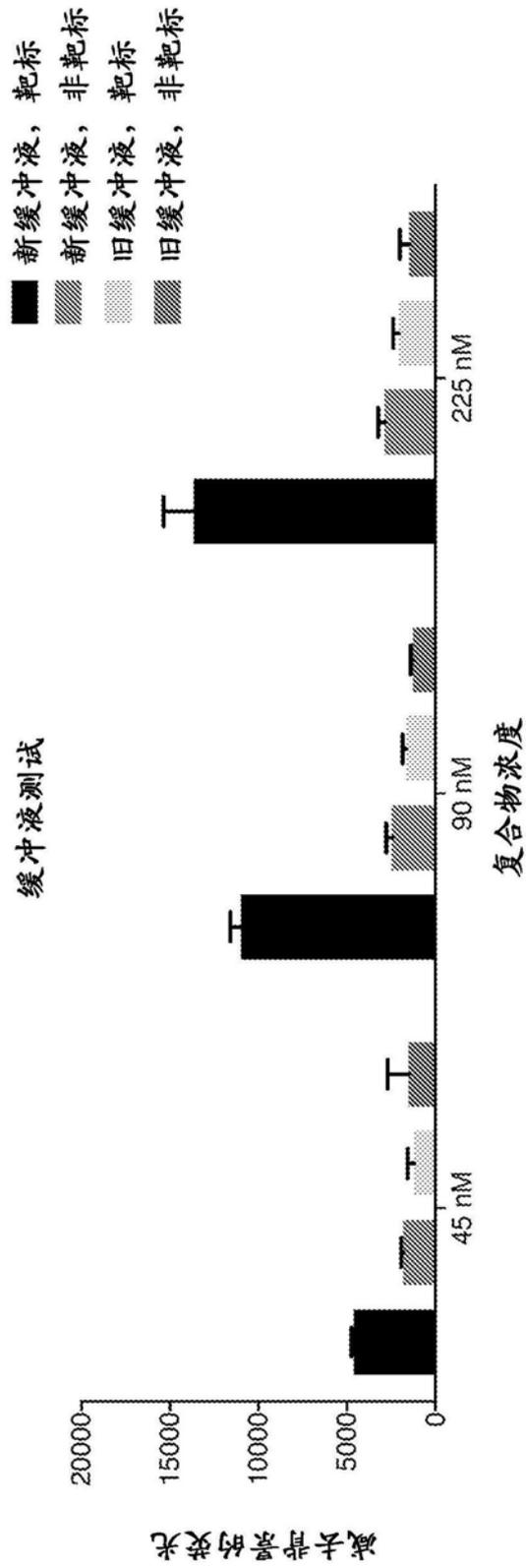


图76

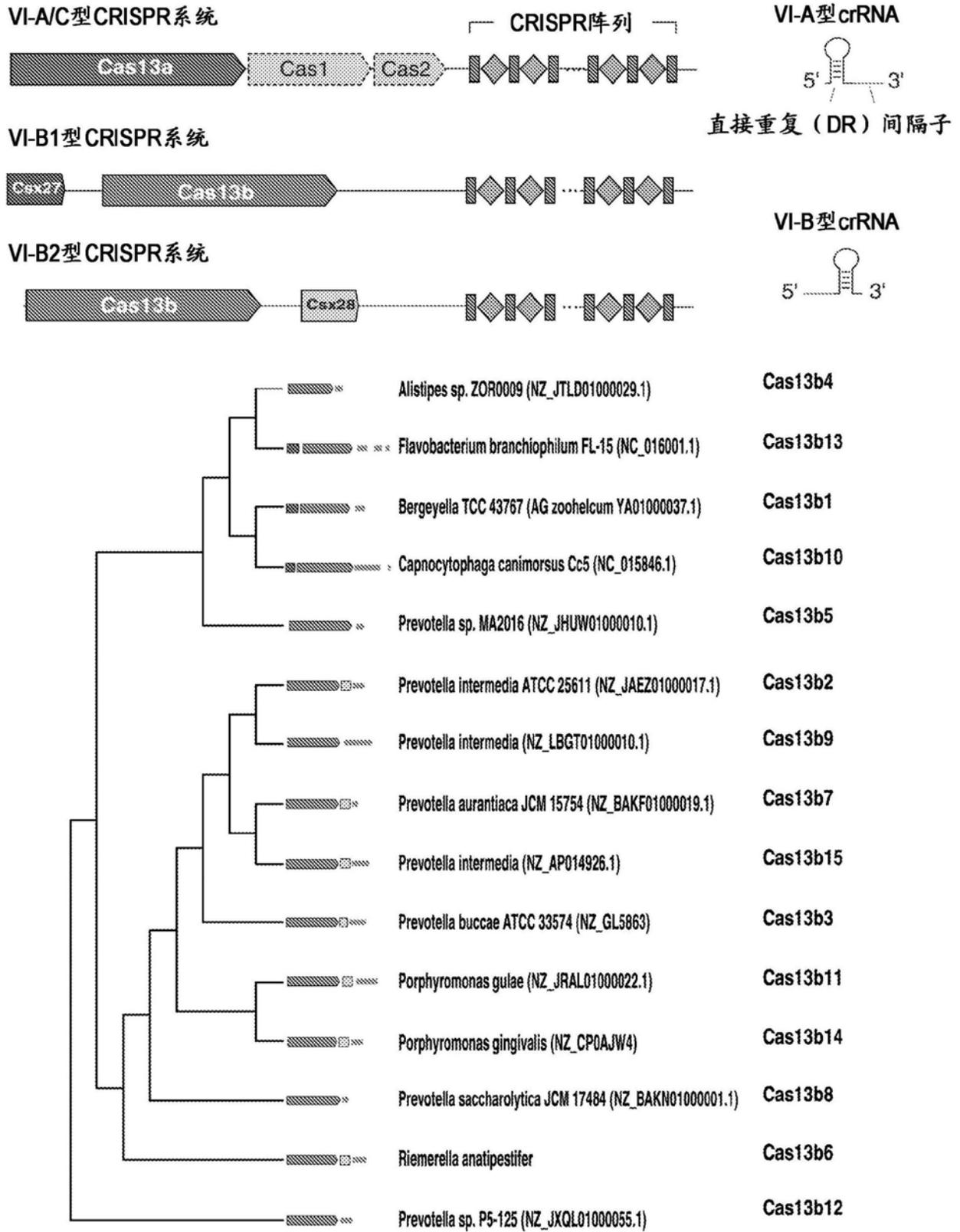


图77

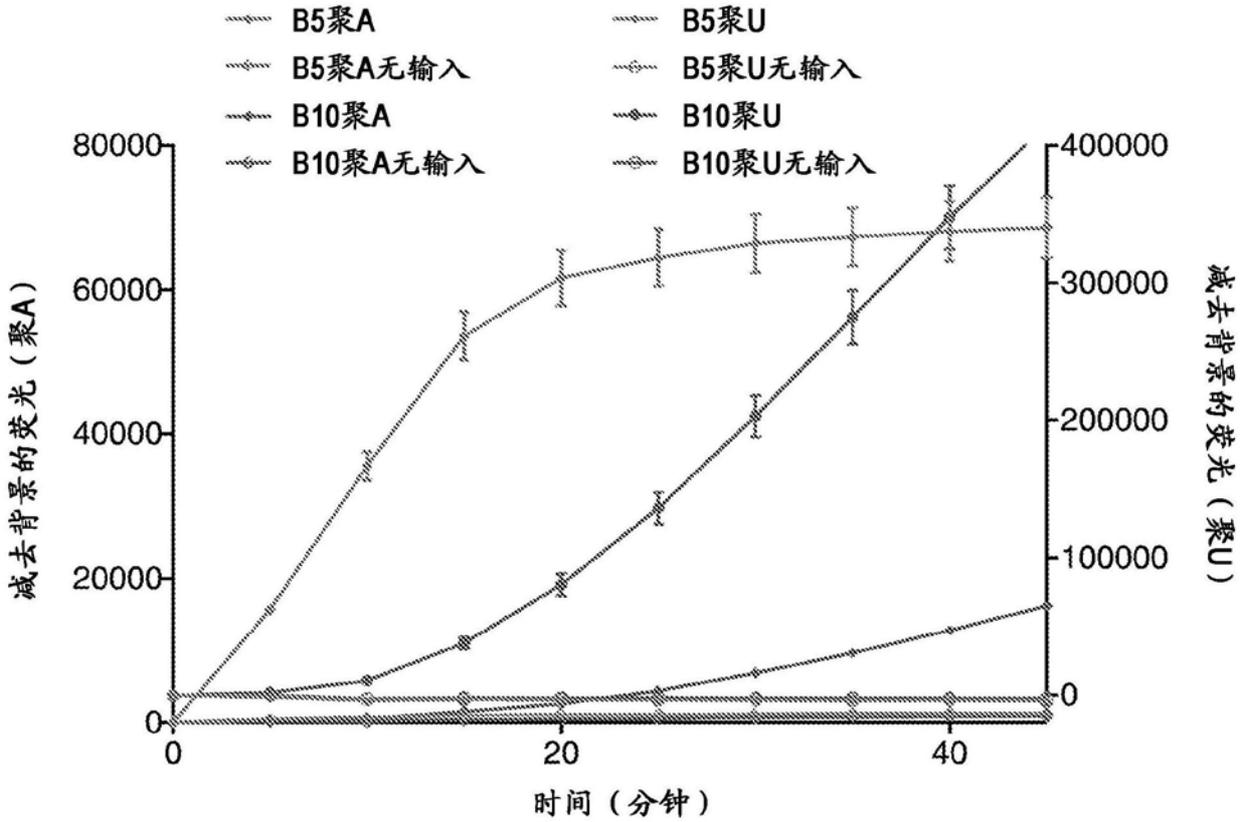
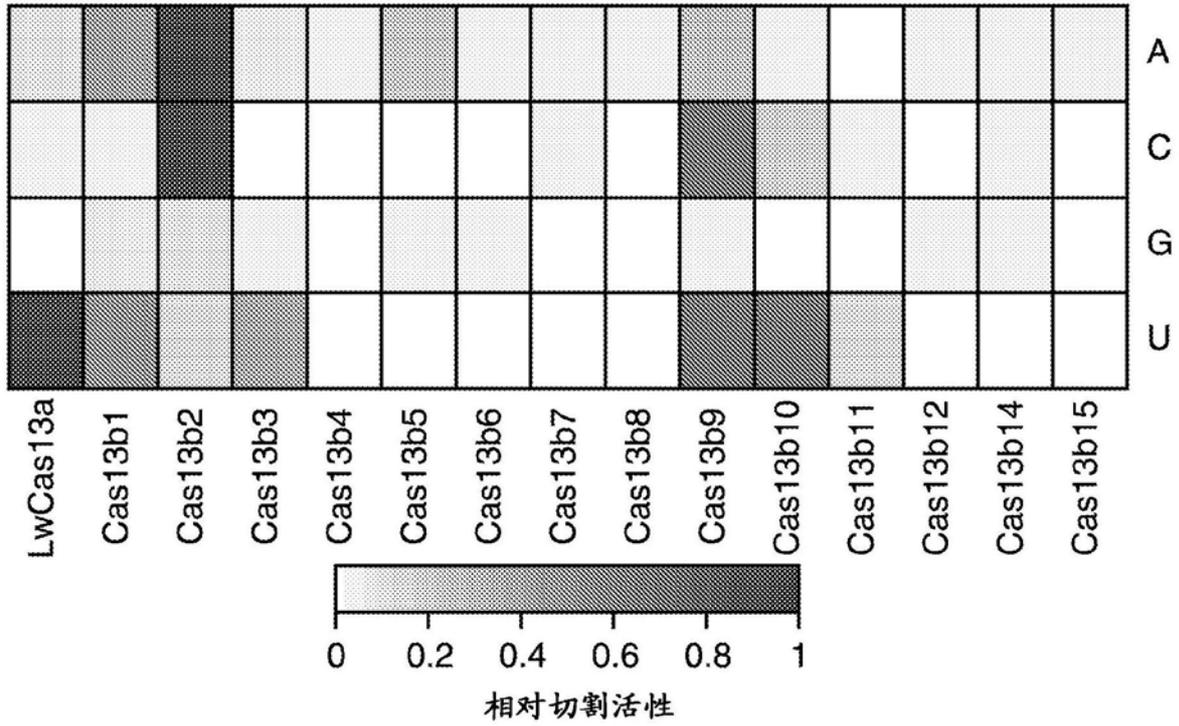


图78

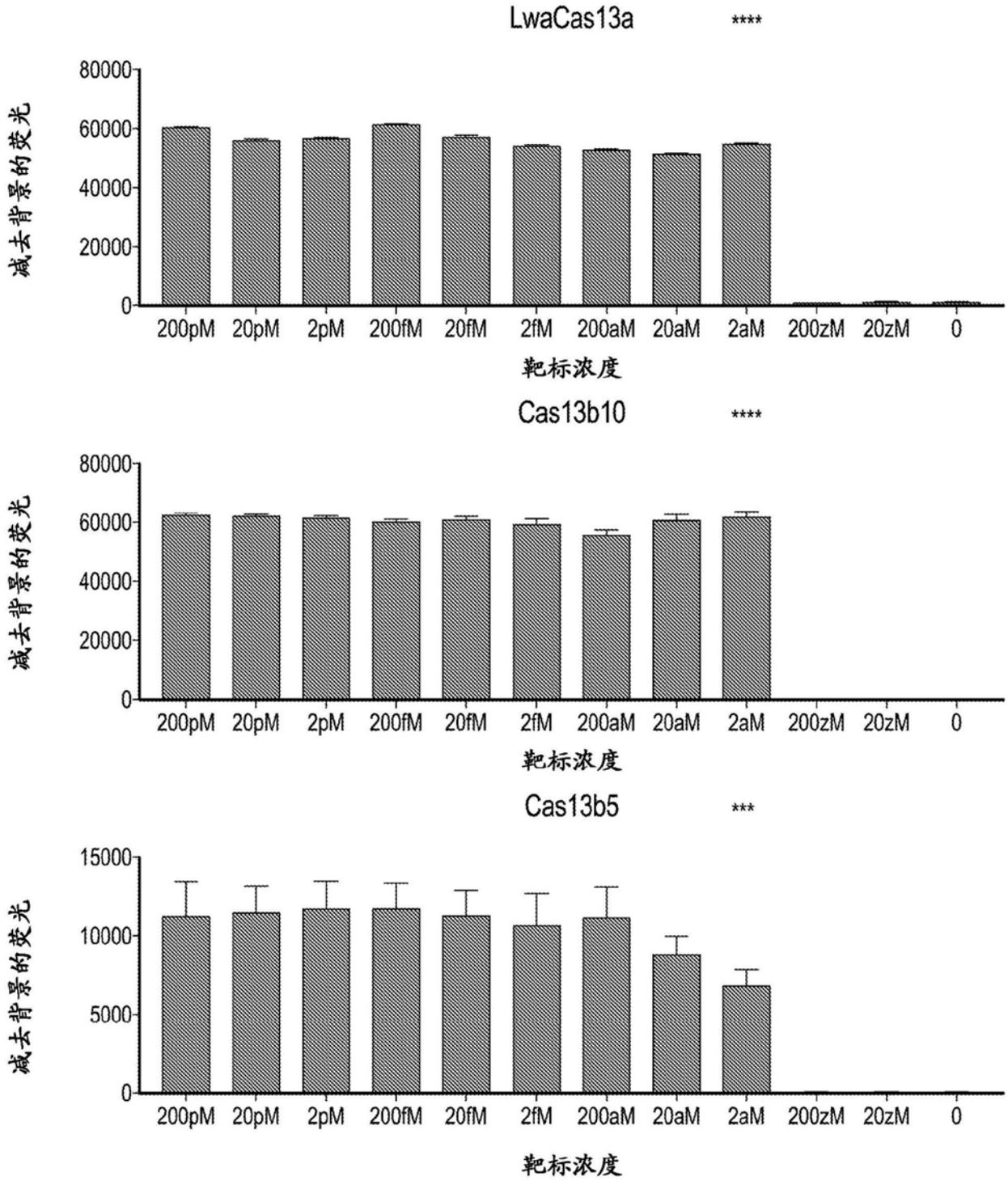


图79

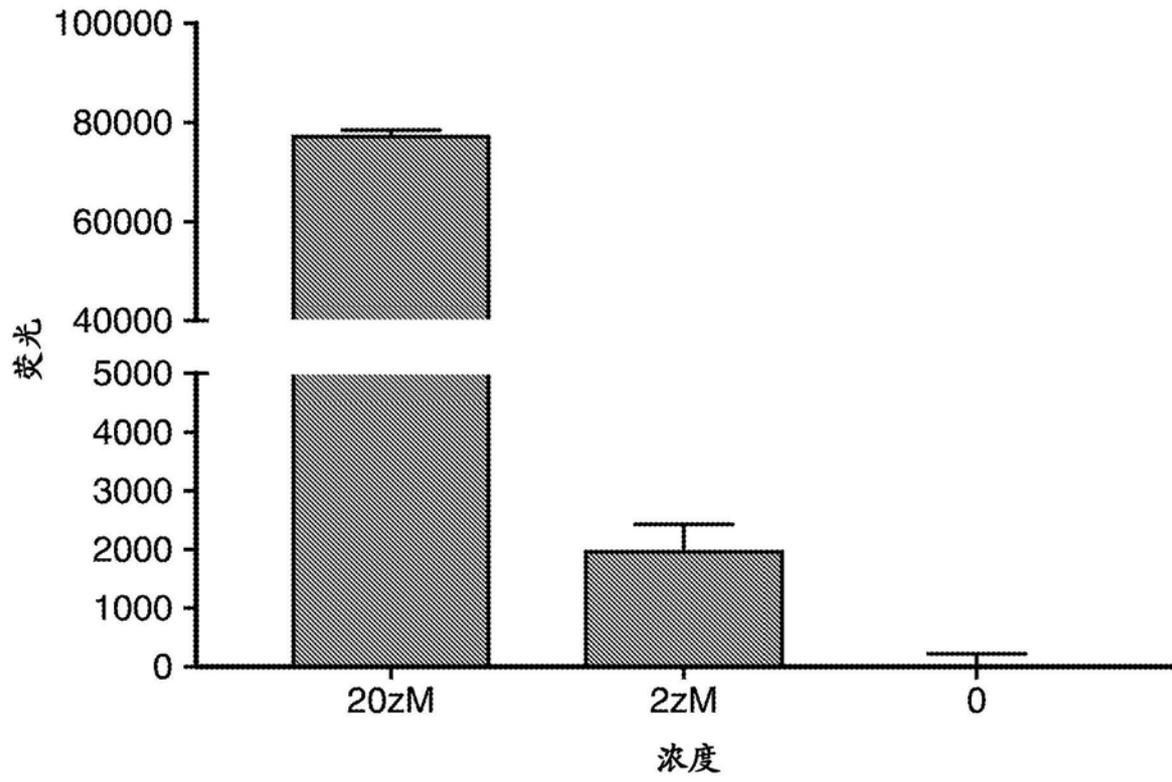


图80

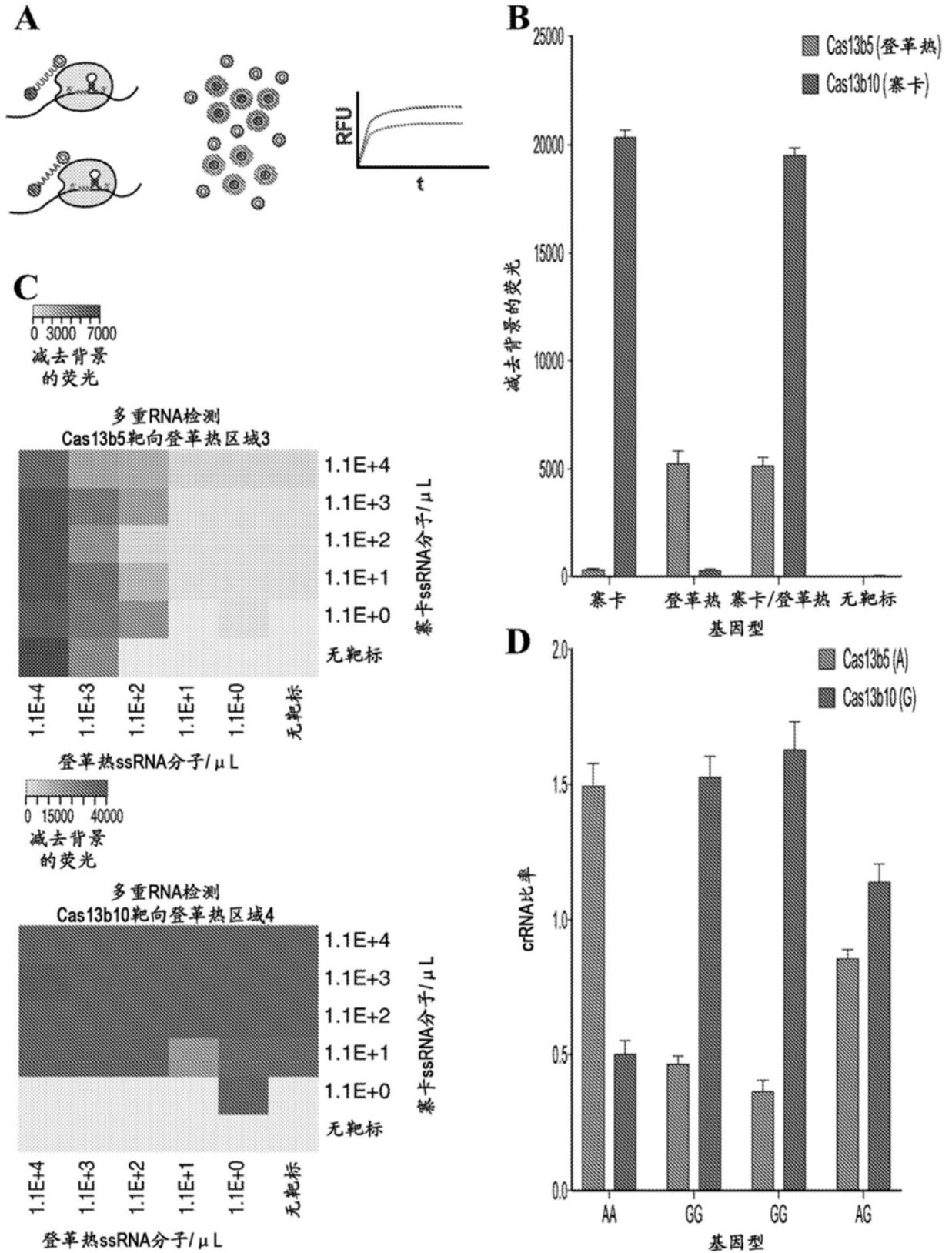


图81

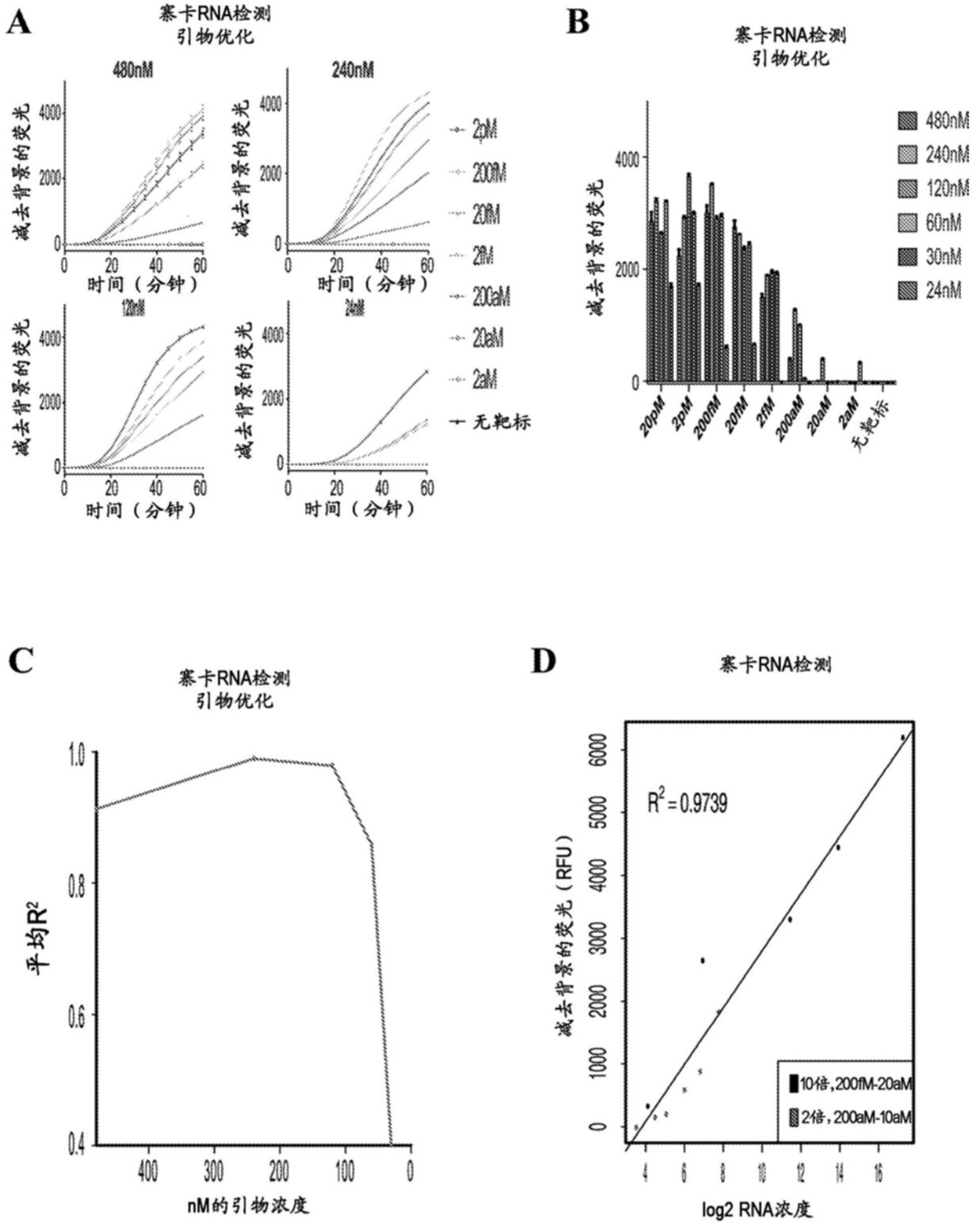
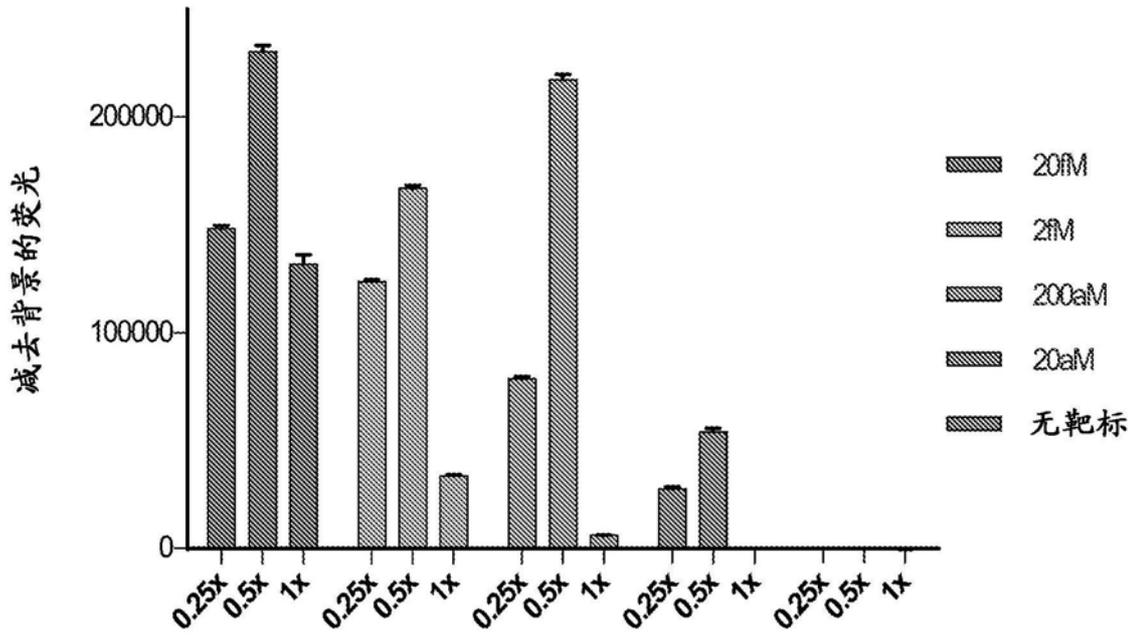


图82

E

假单胞菌



F

假单胞菌

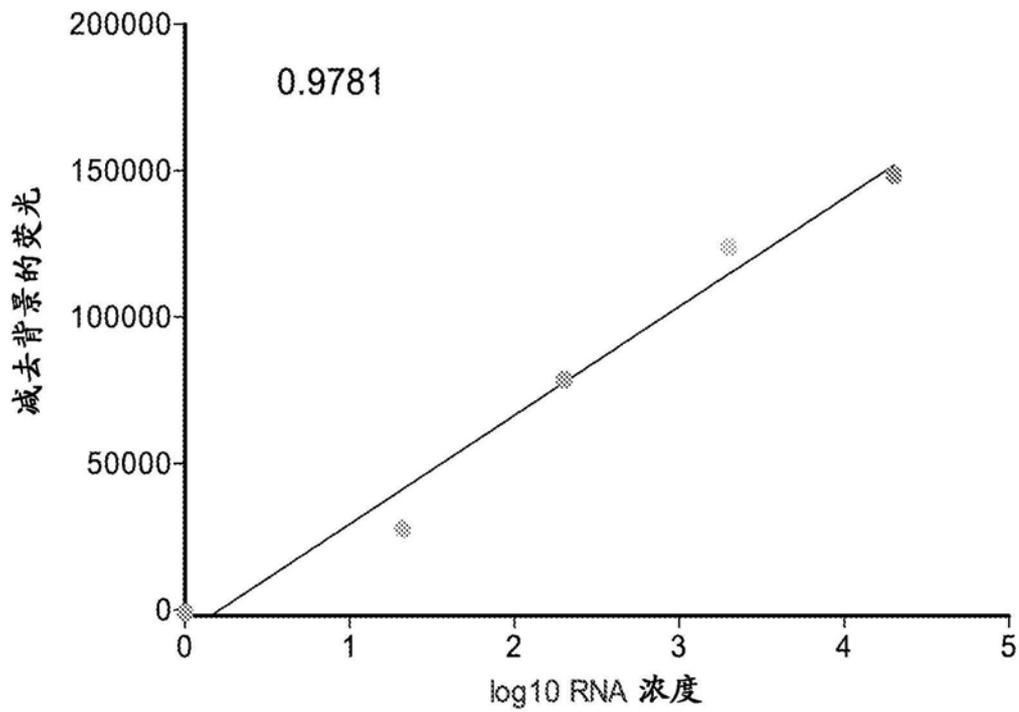


图82(续)

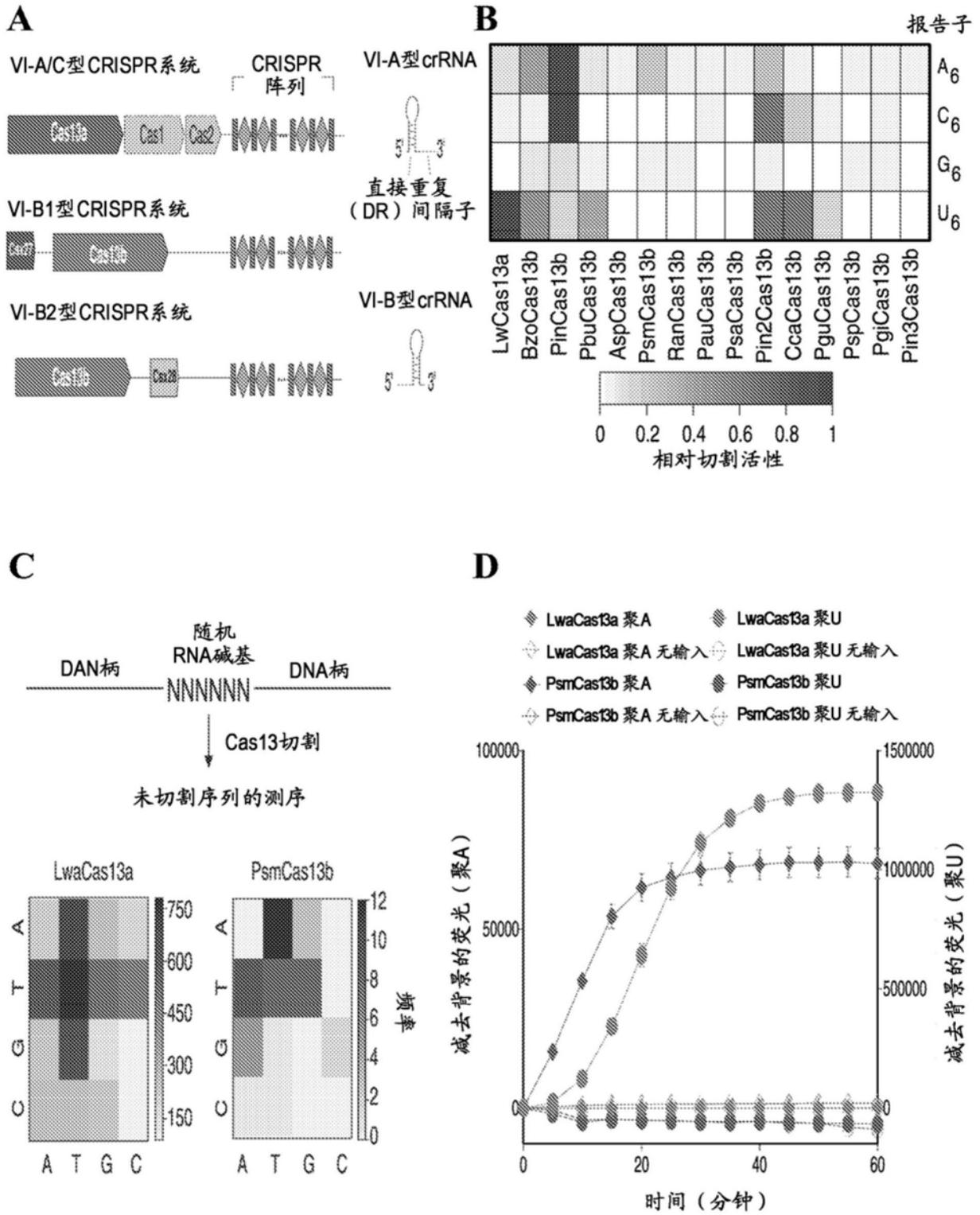


图83

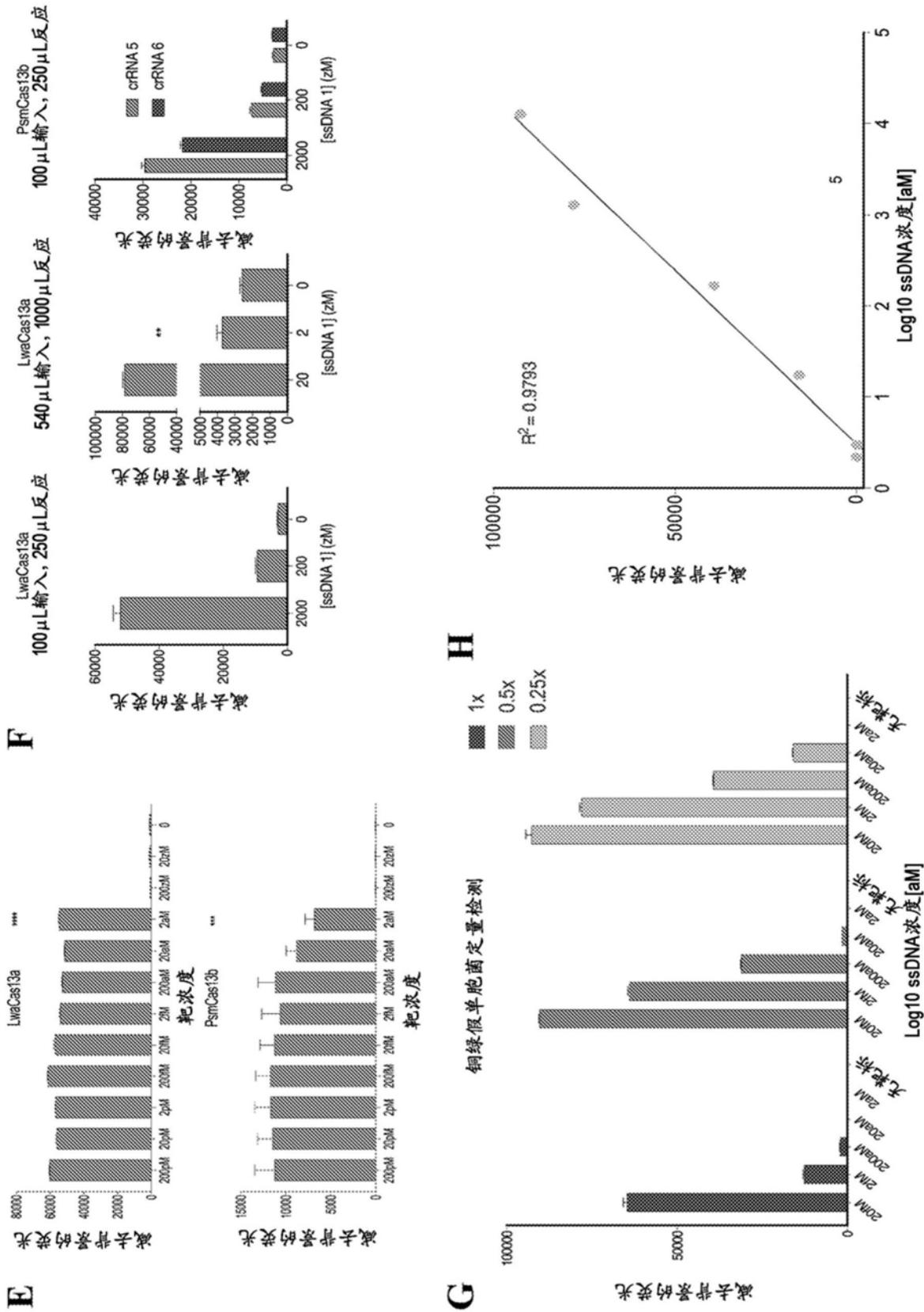


图83 (续)

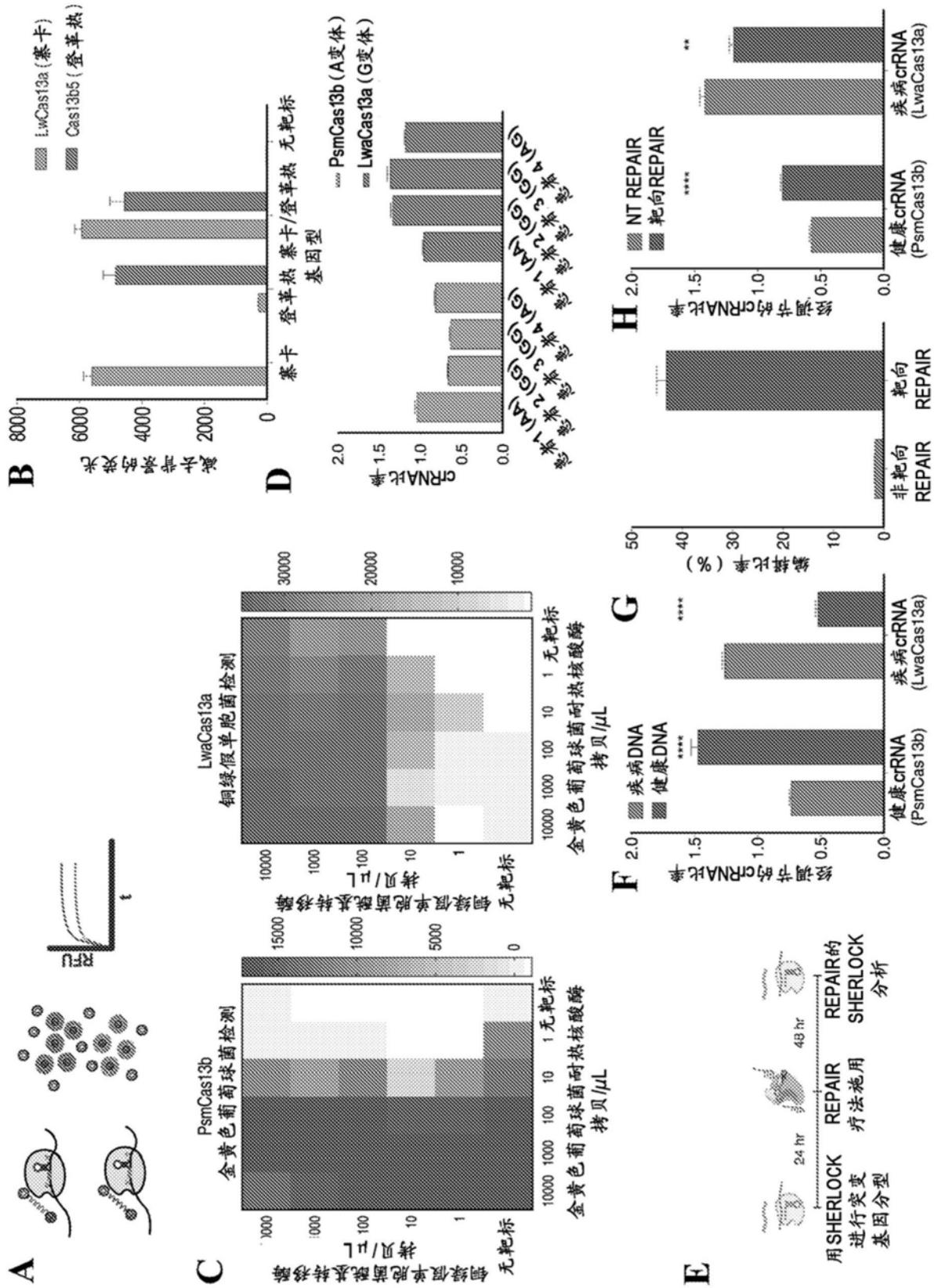


图84

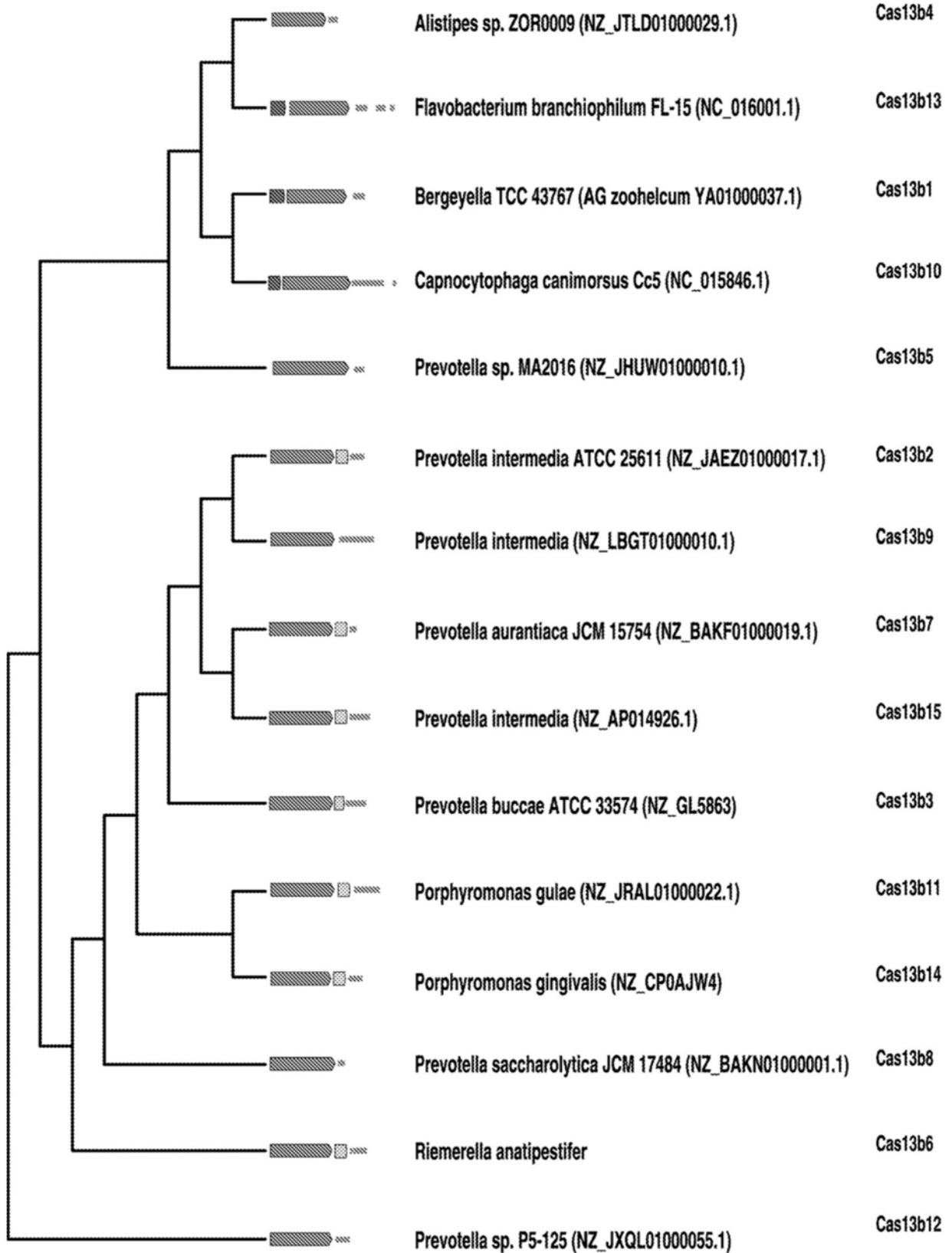


图85

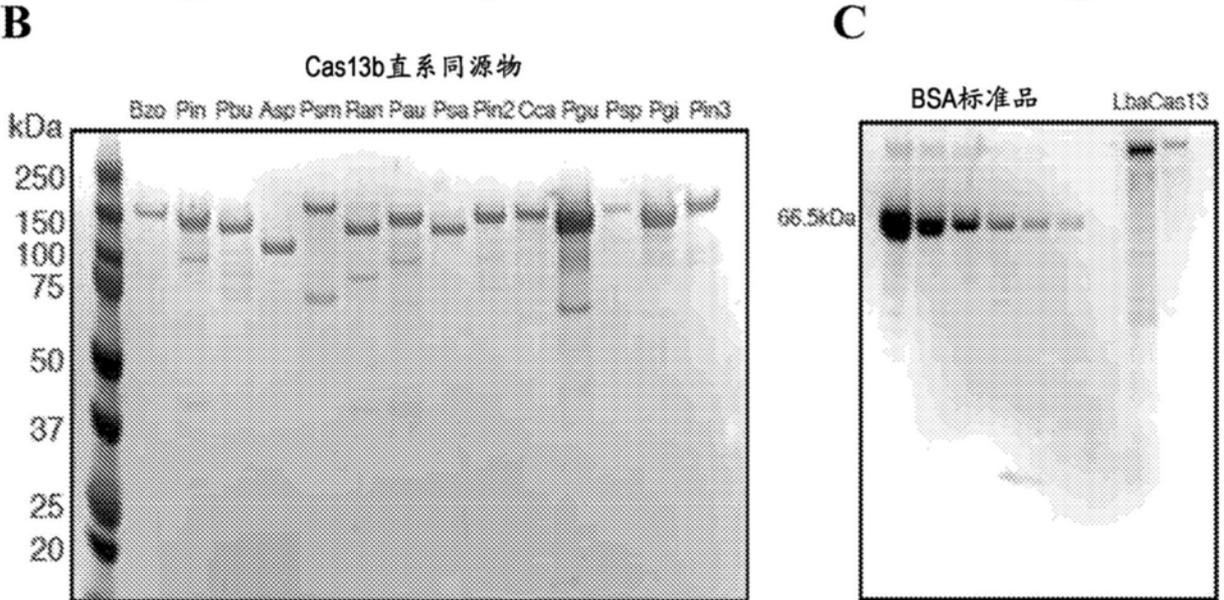
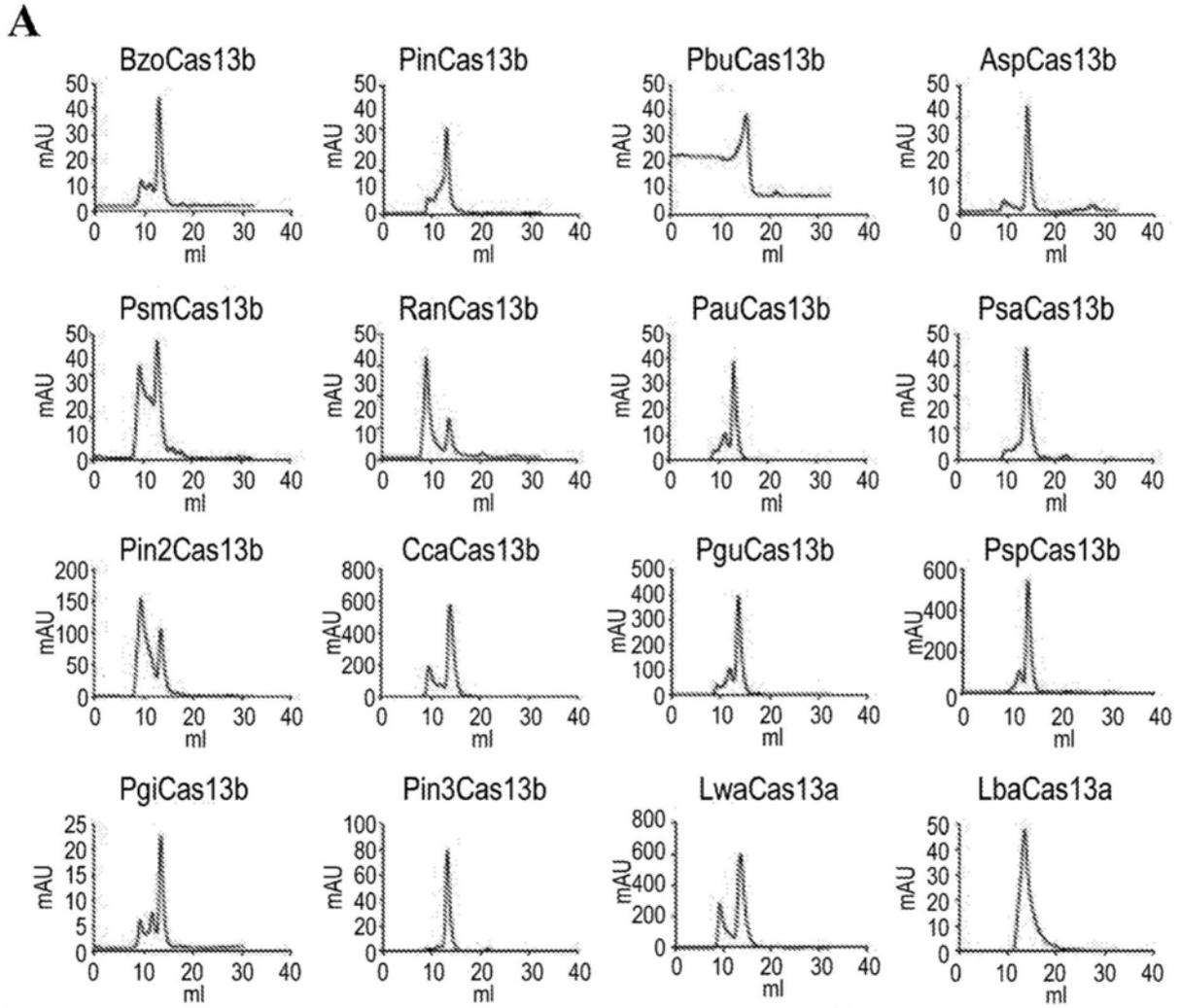


图86

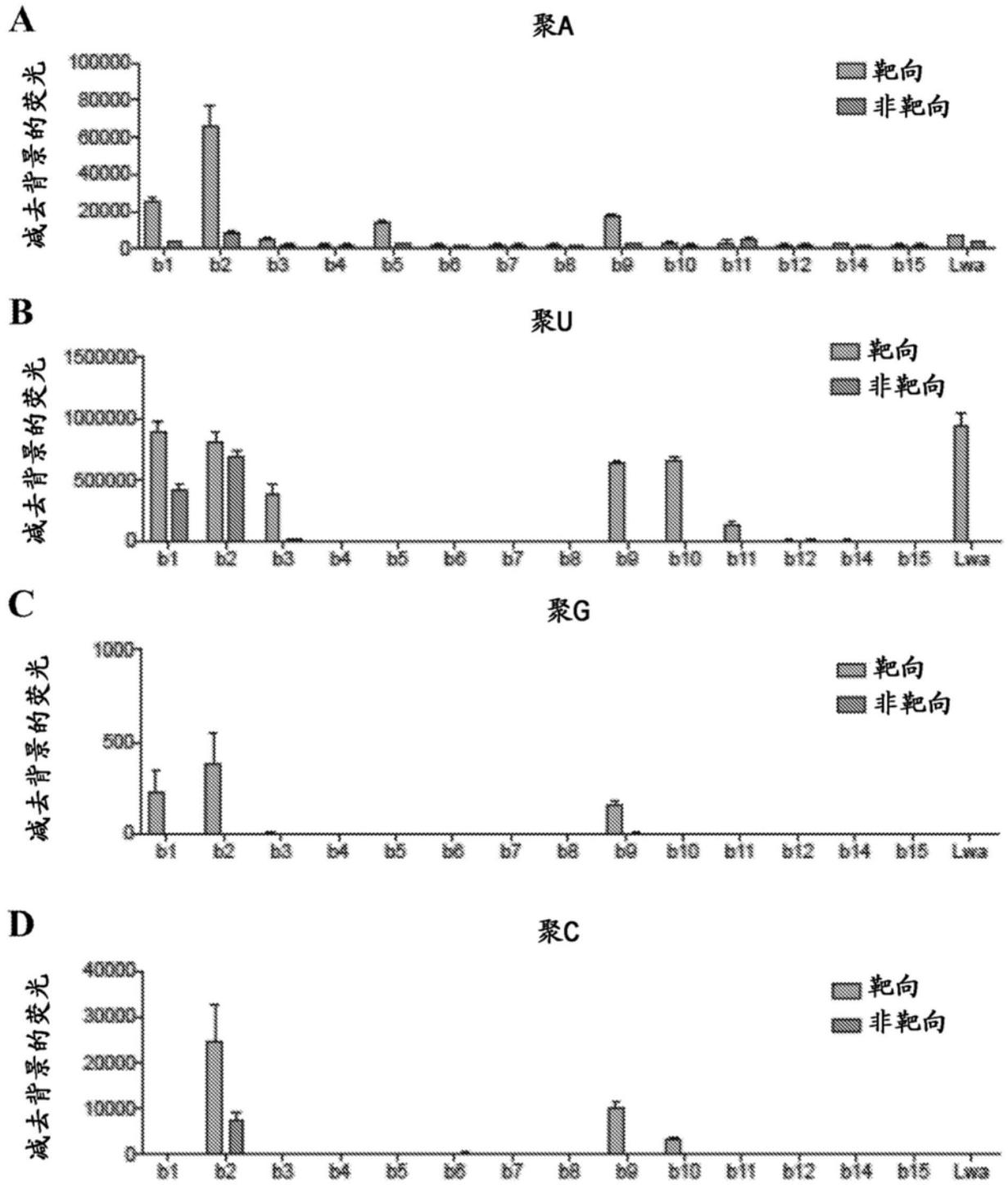


图87

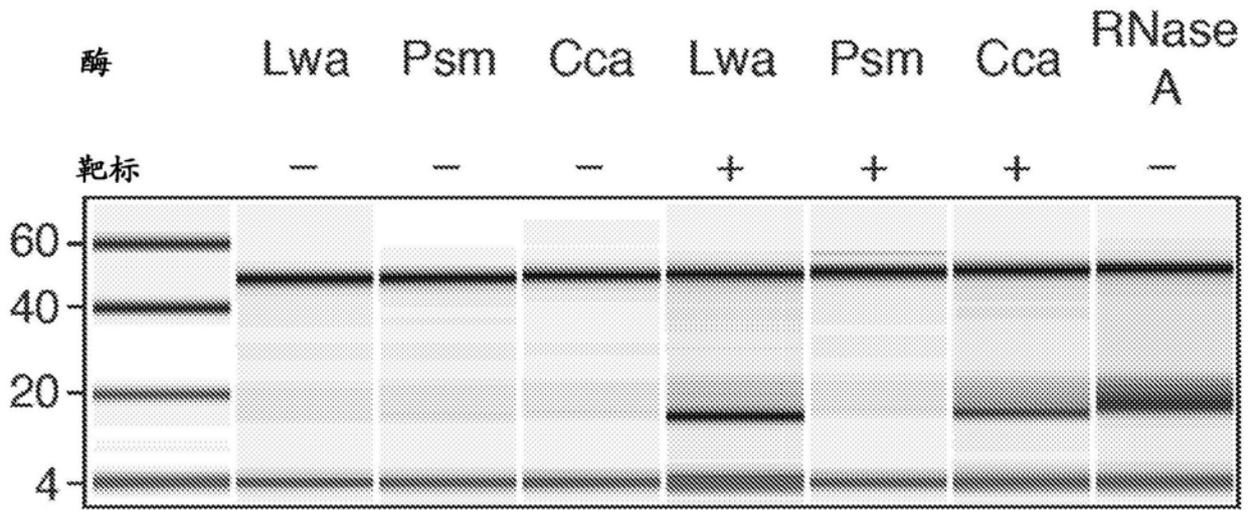


图88

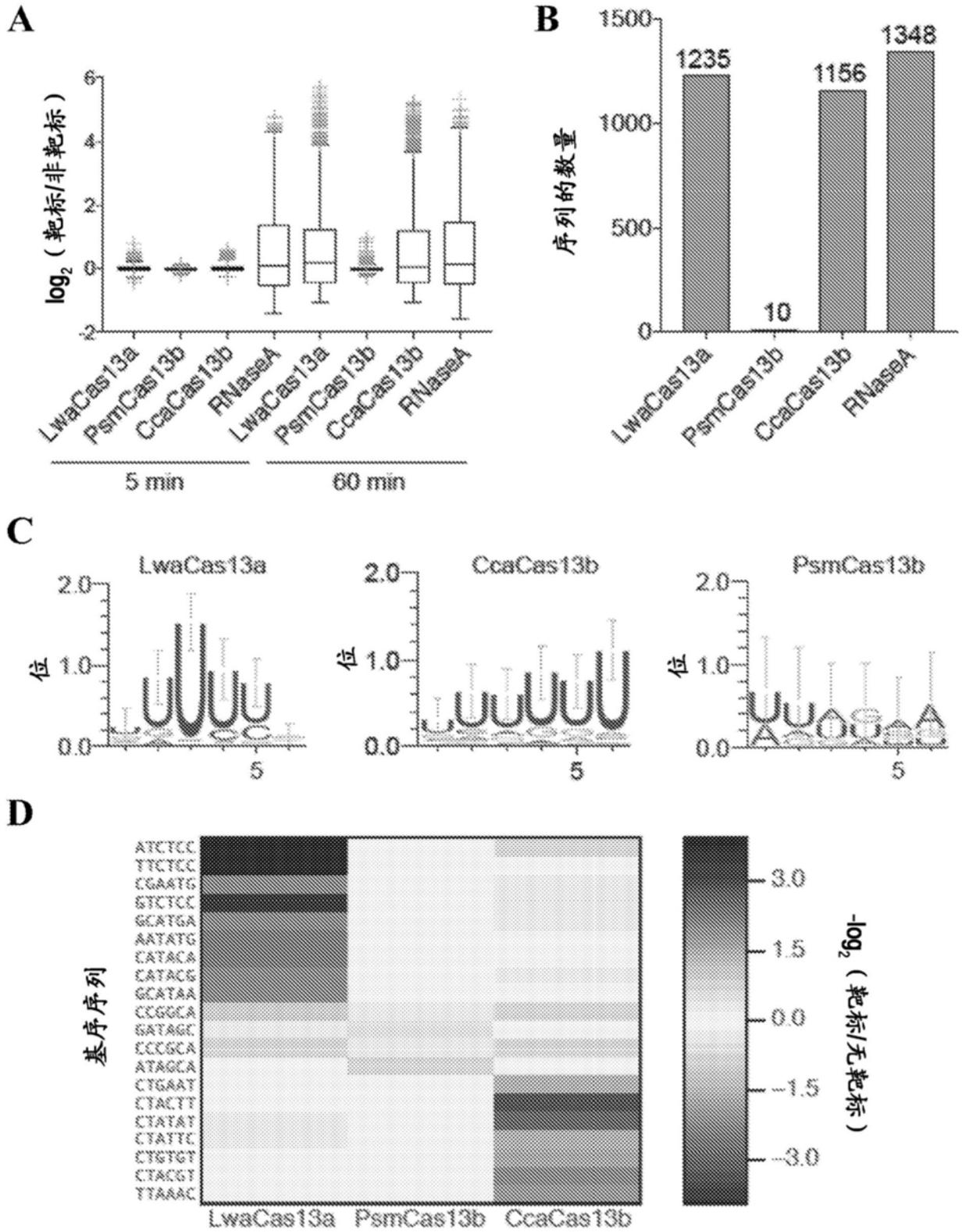


图89

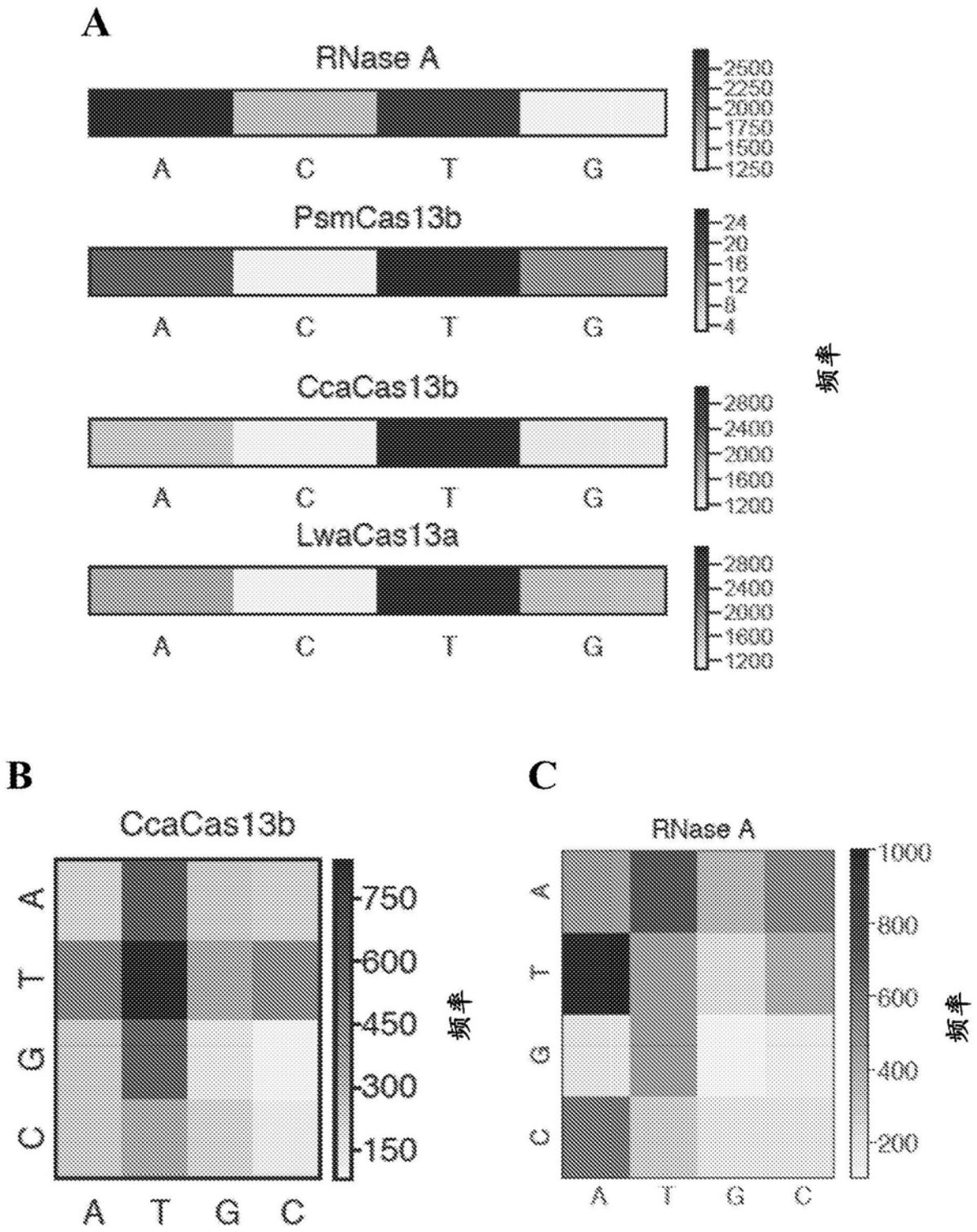


图90

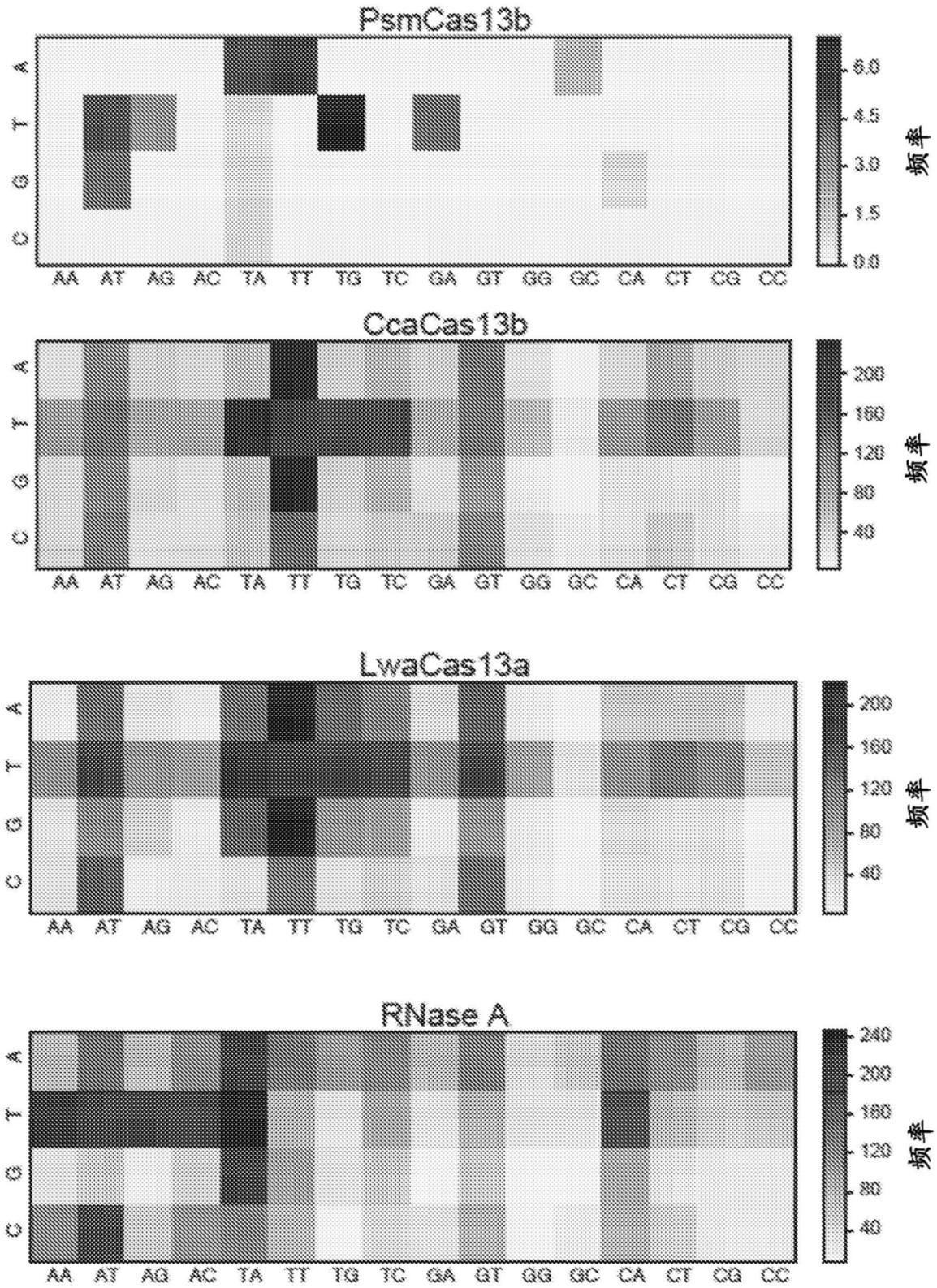


图91

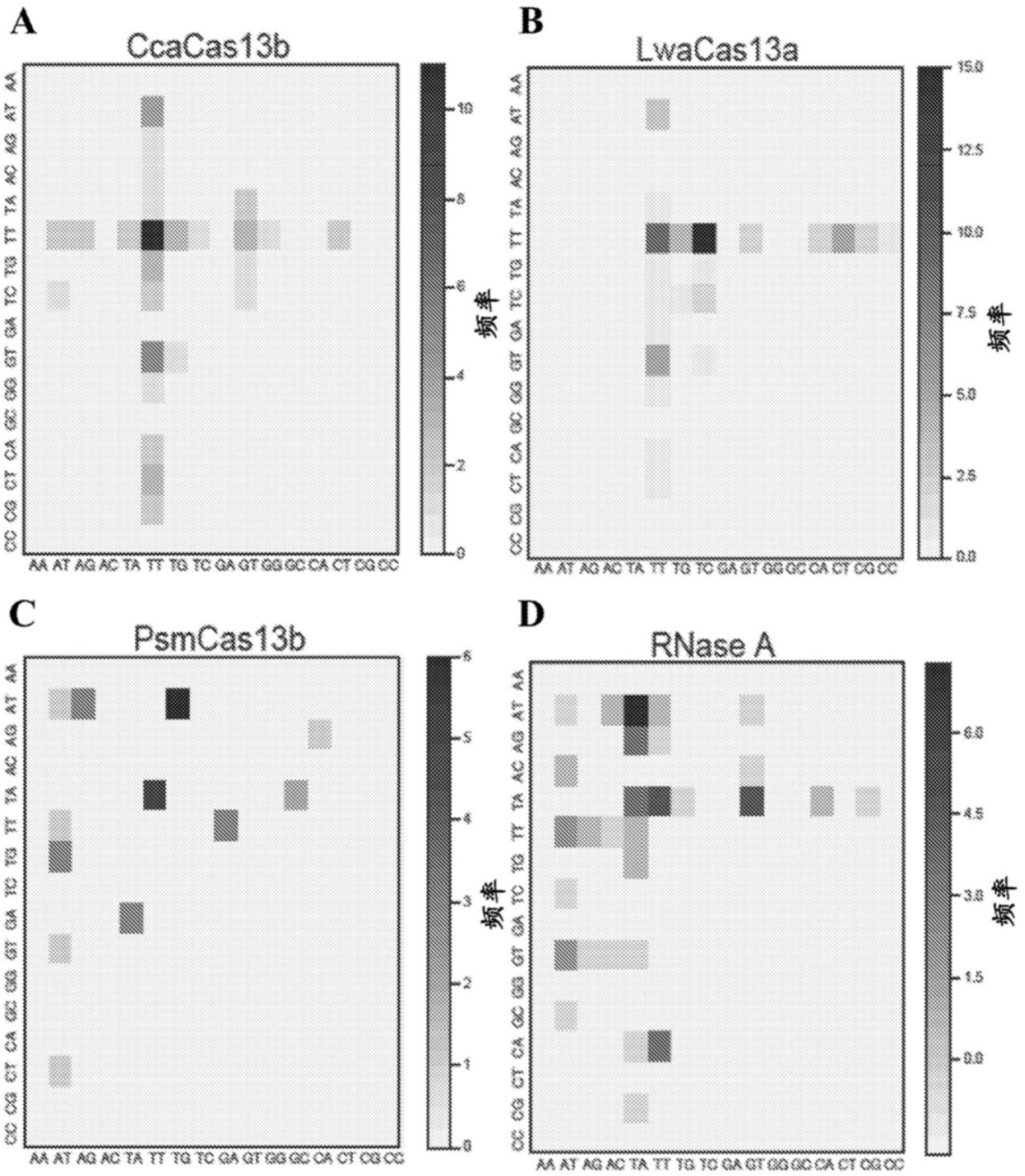


图92

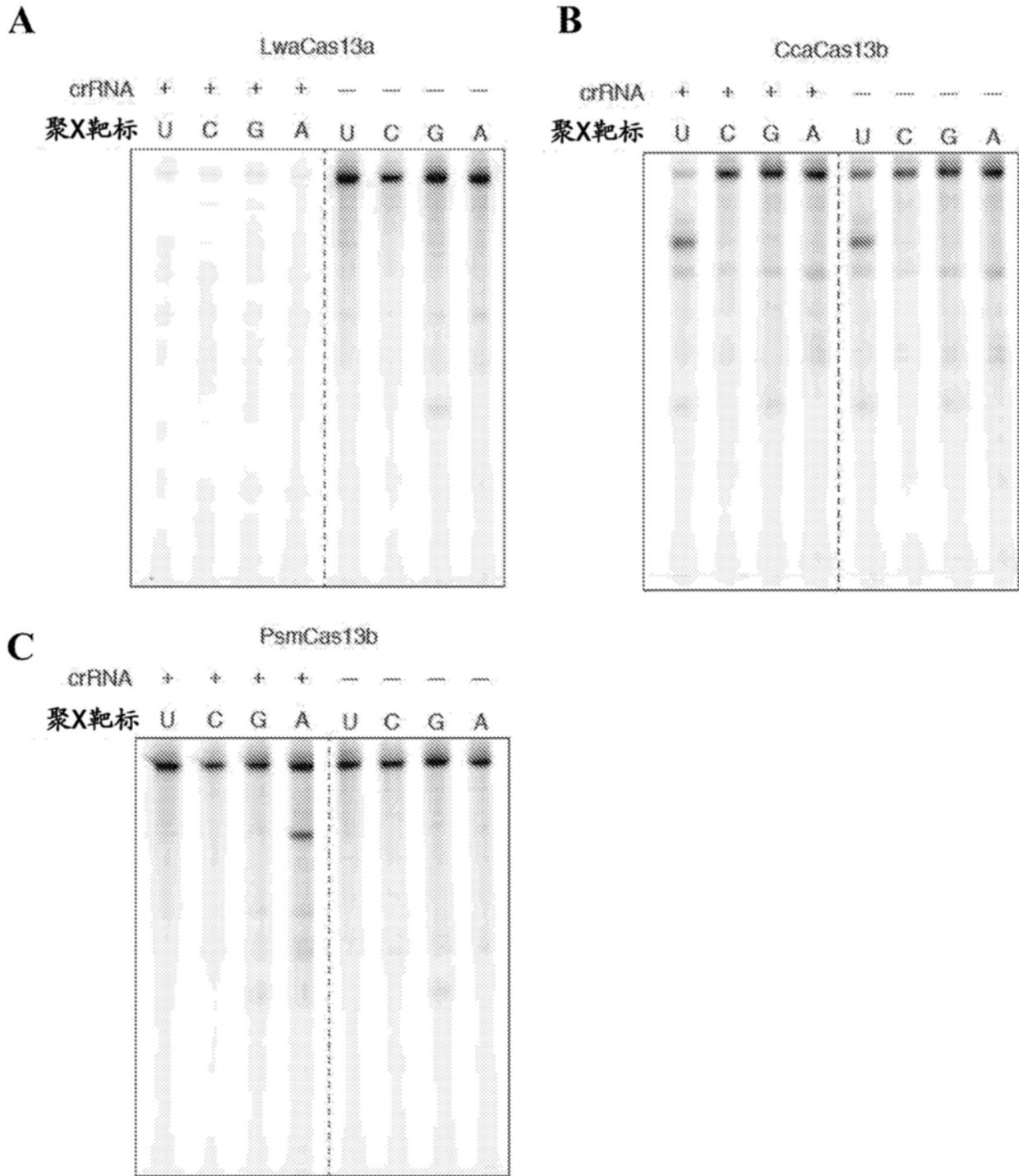


图93

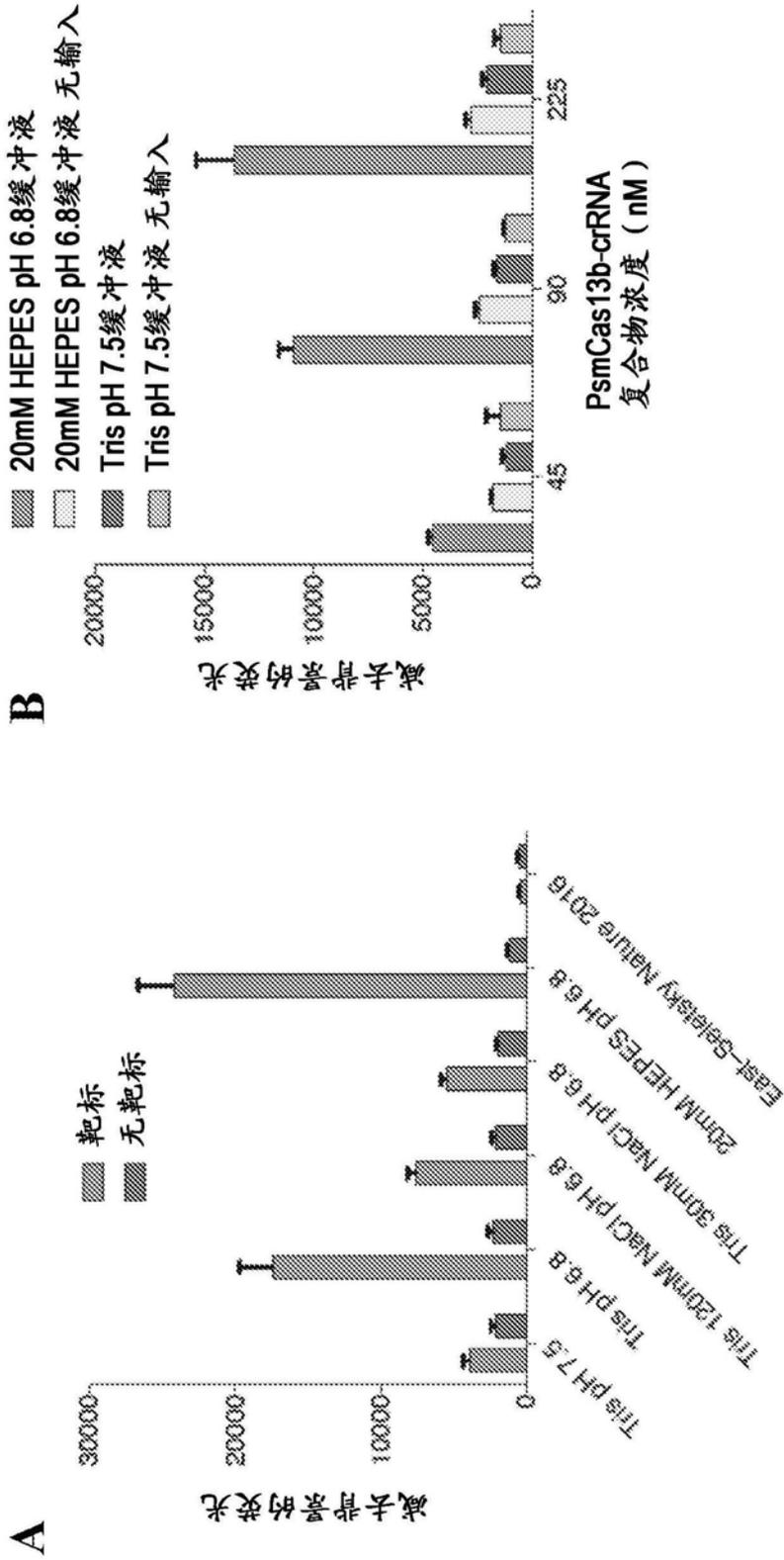


图94

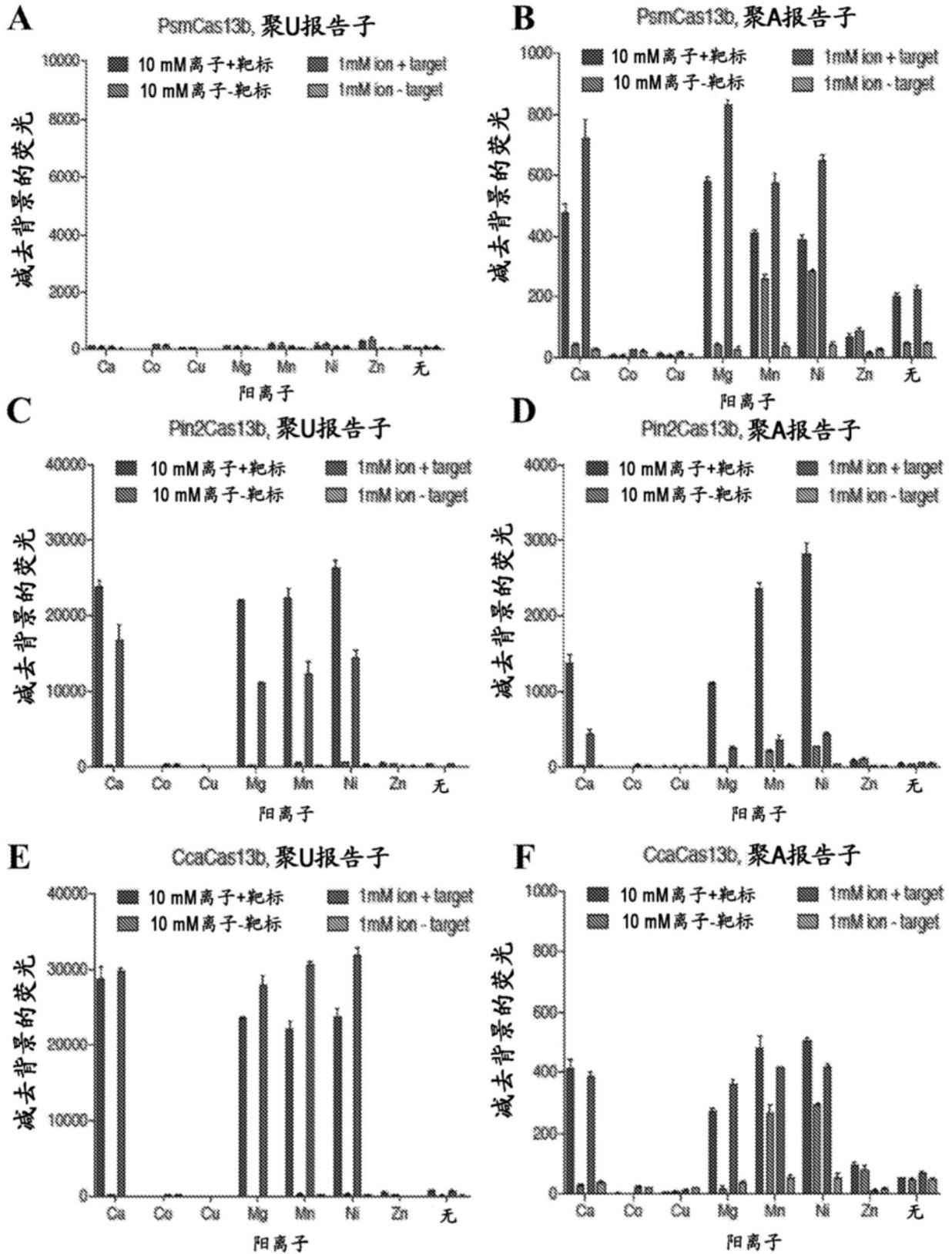


图95

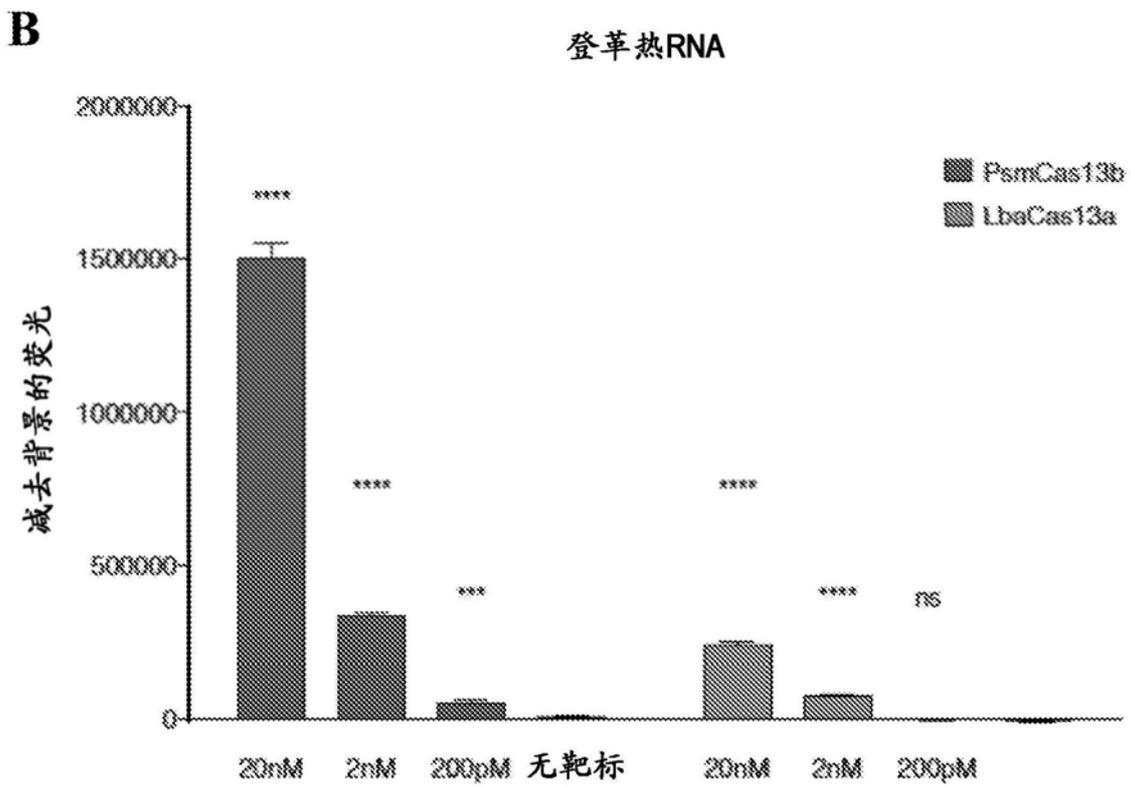
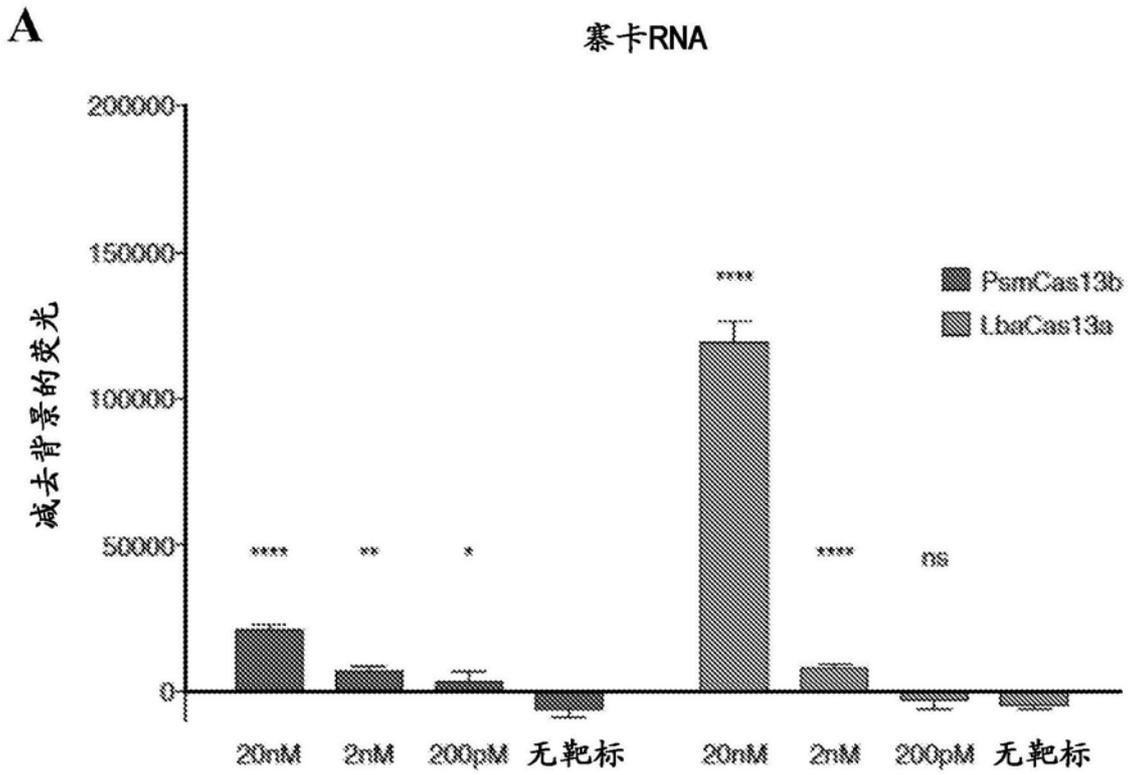
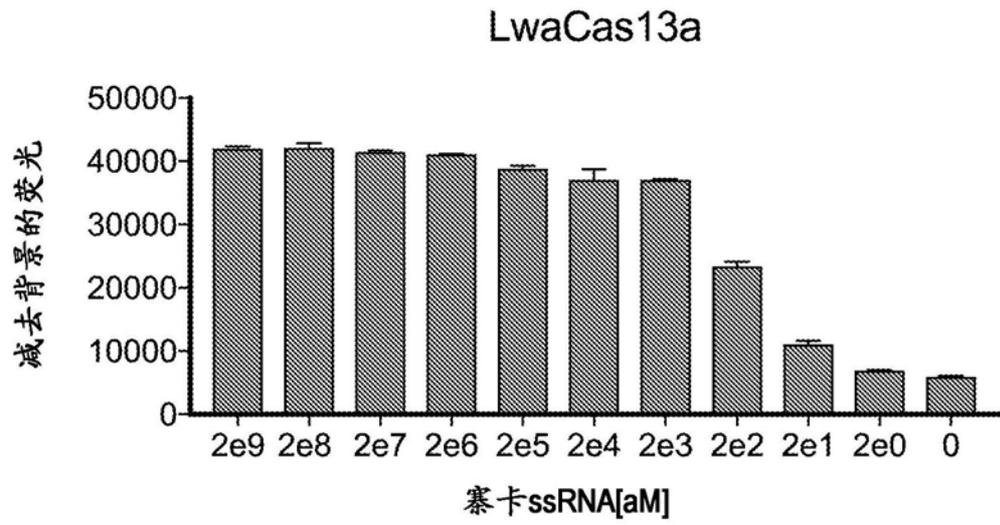


图96

A



B

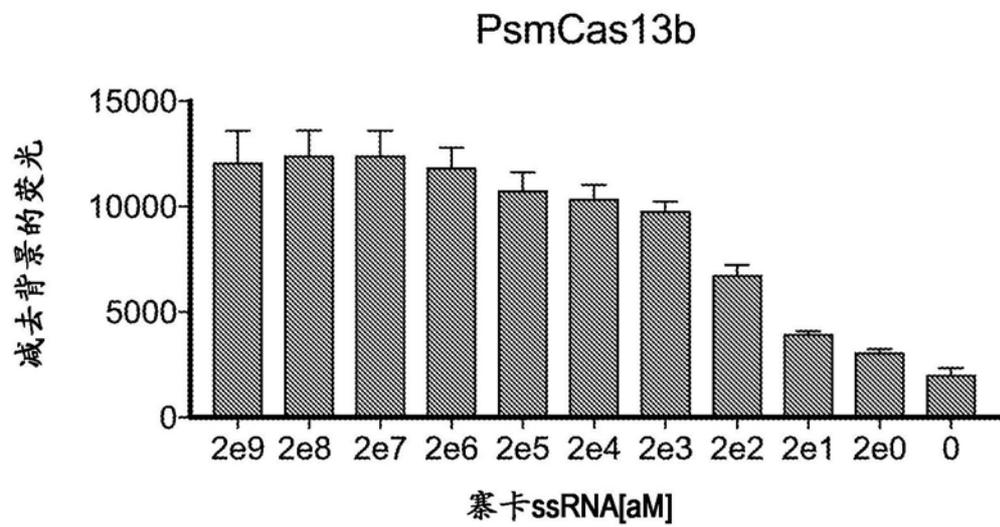


图97

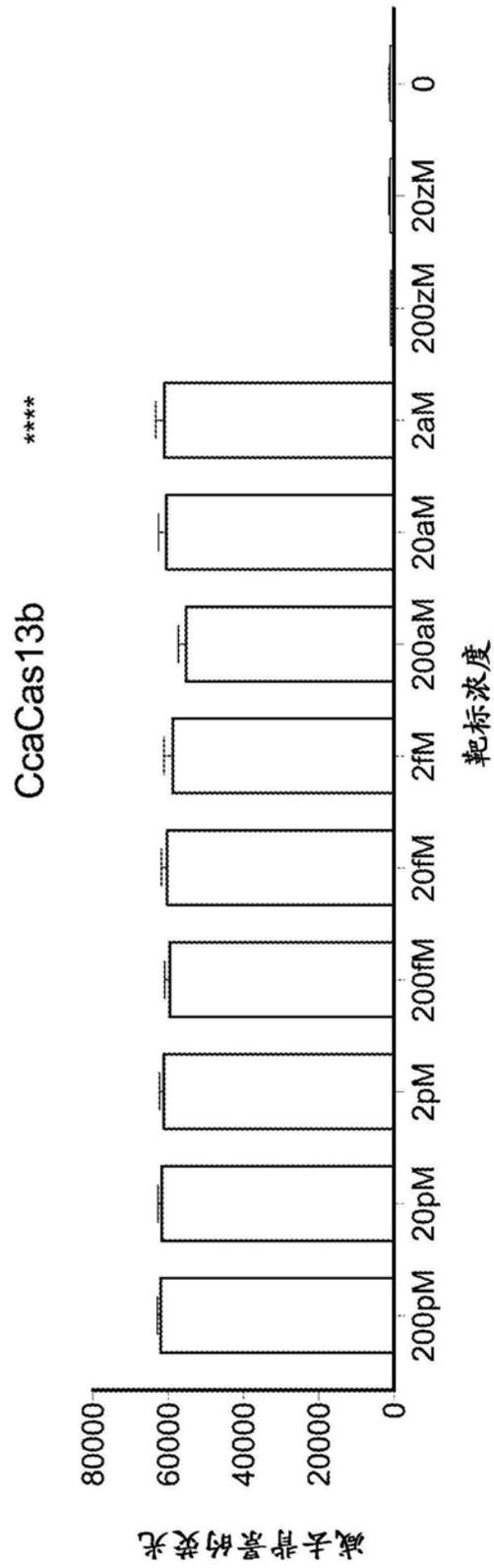


图98

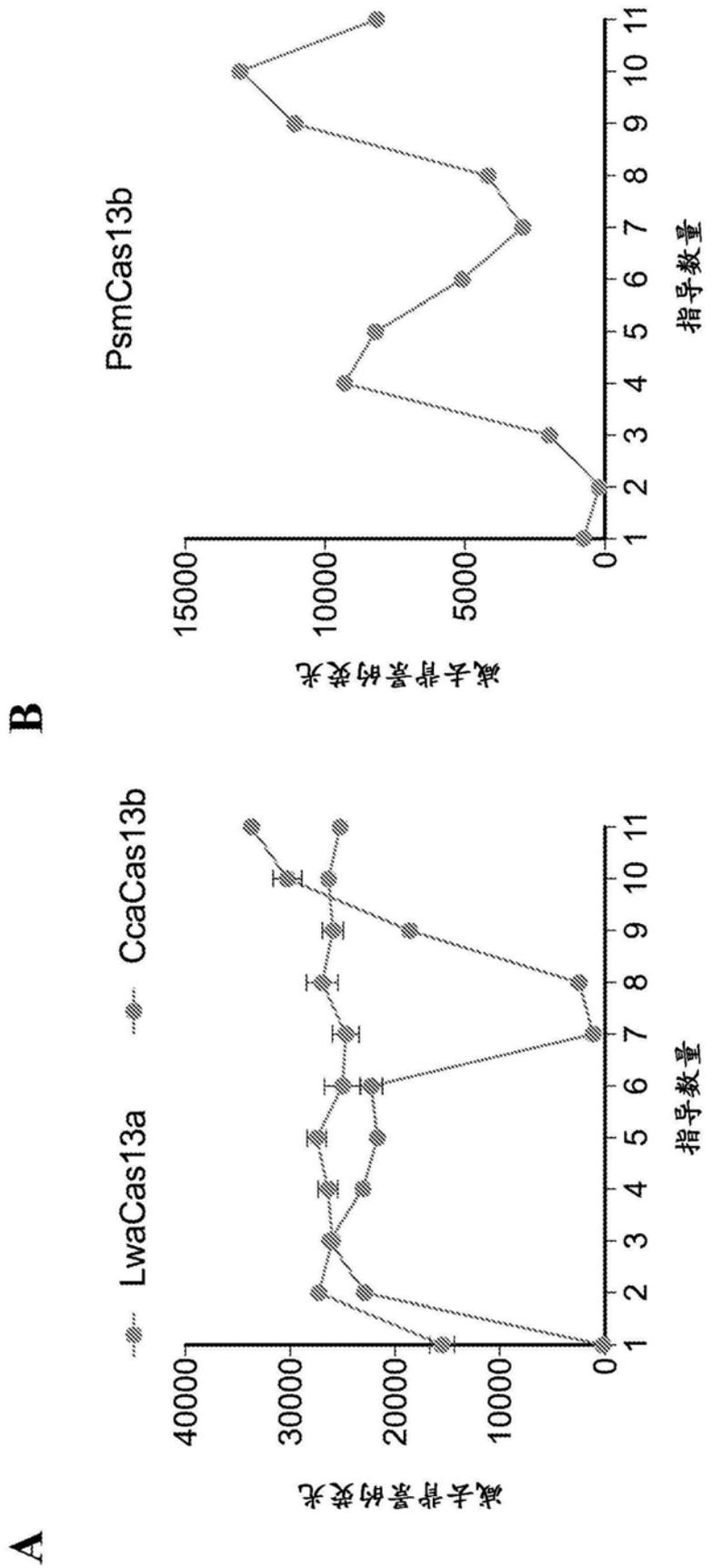


图99

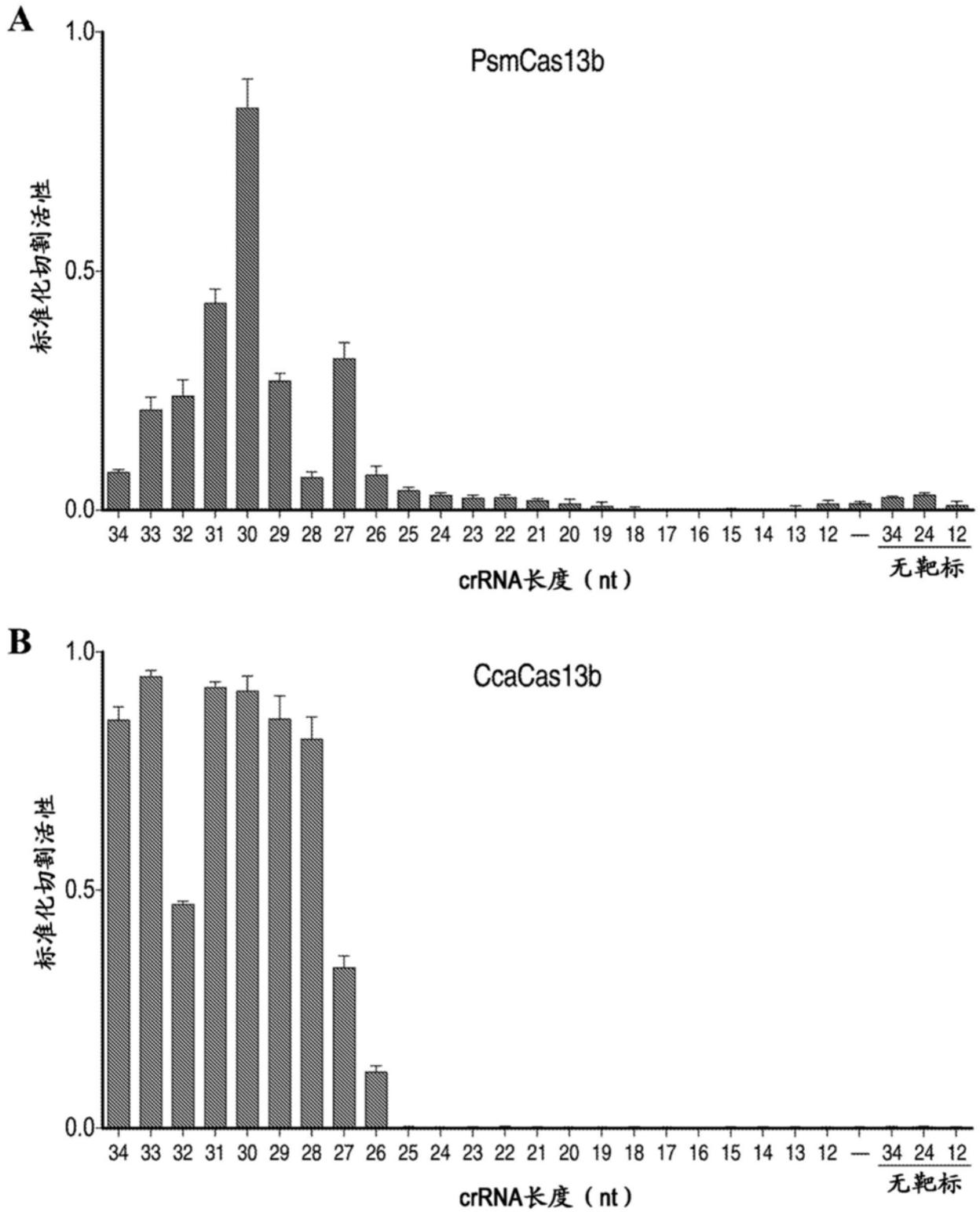


图100

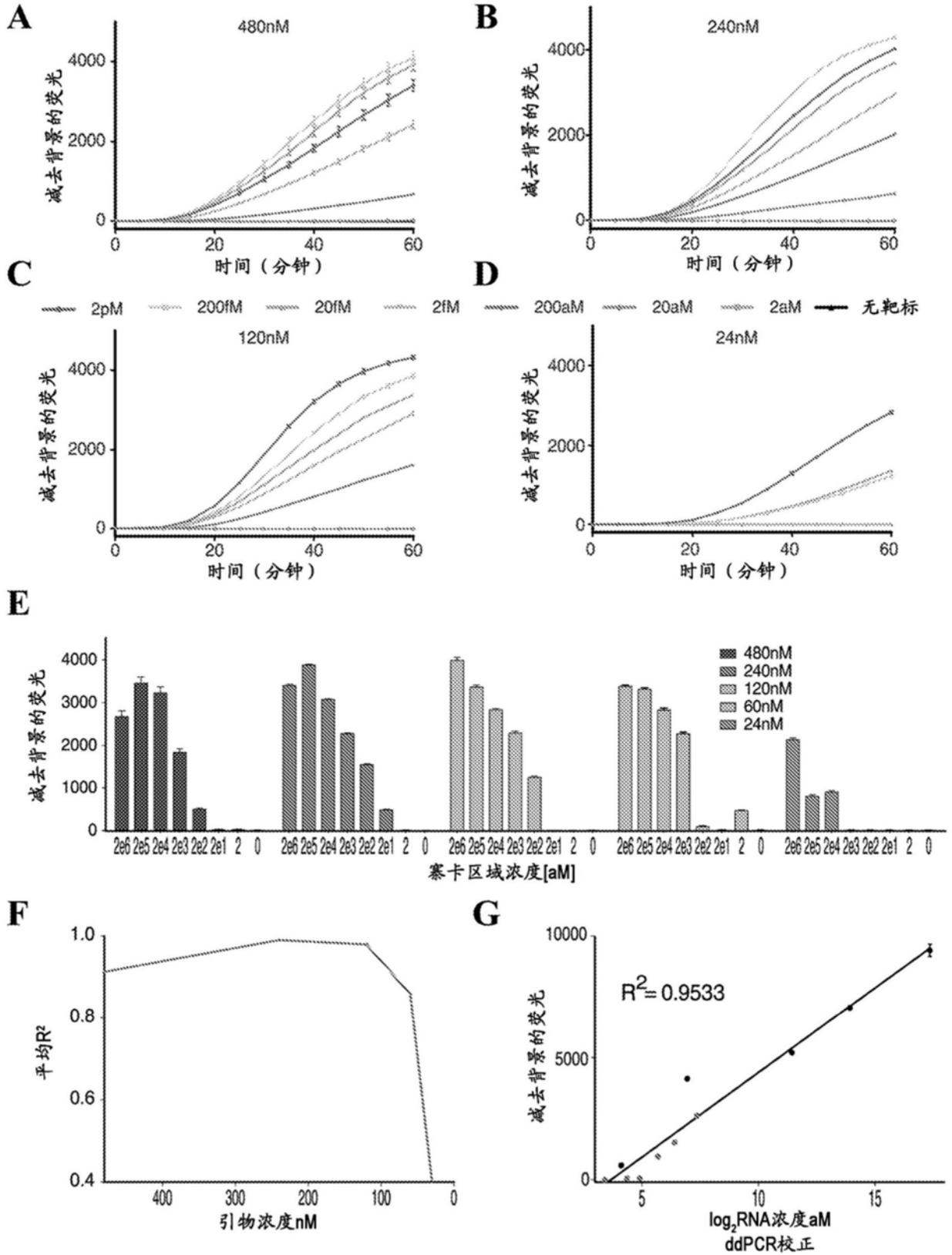
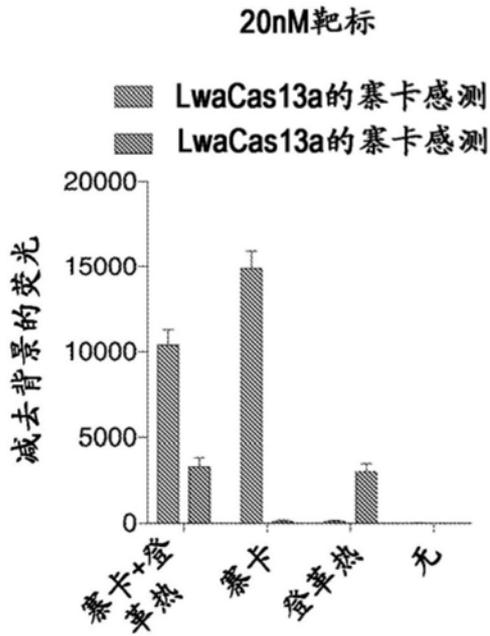
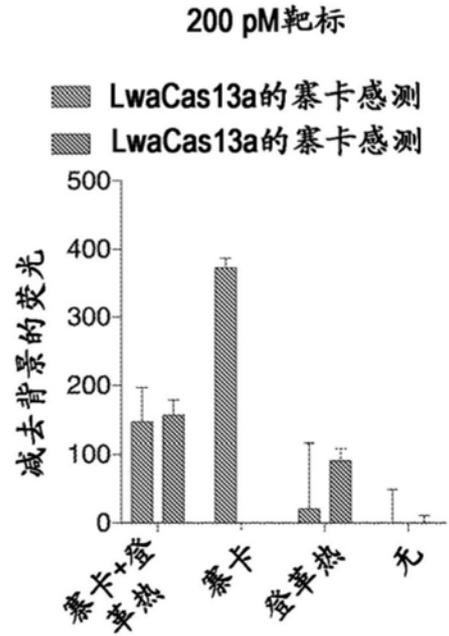


图101

A



B



C

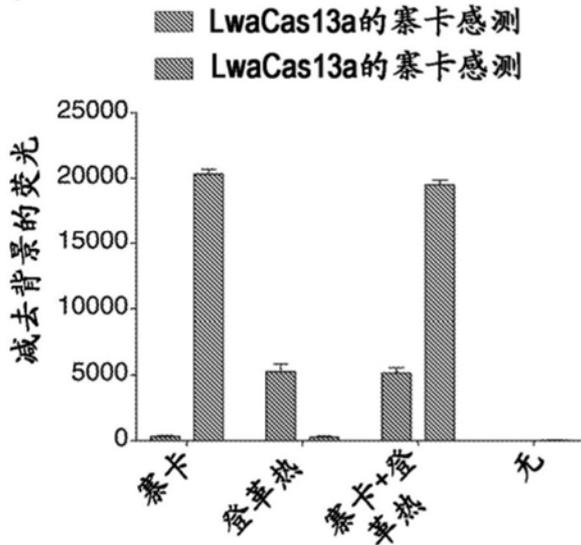
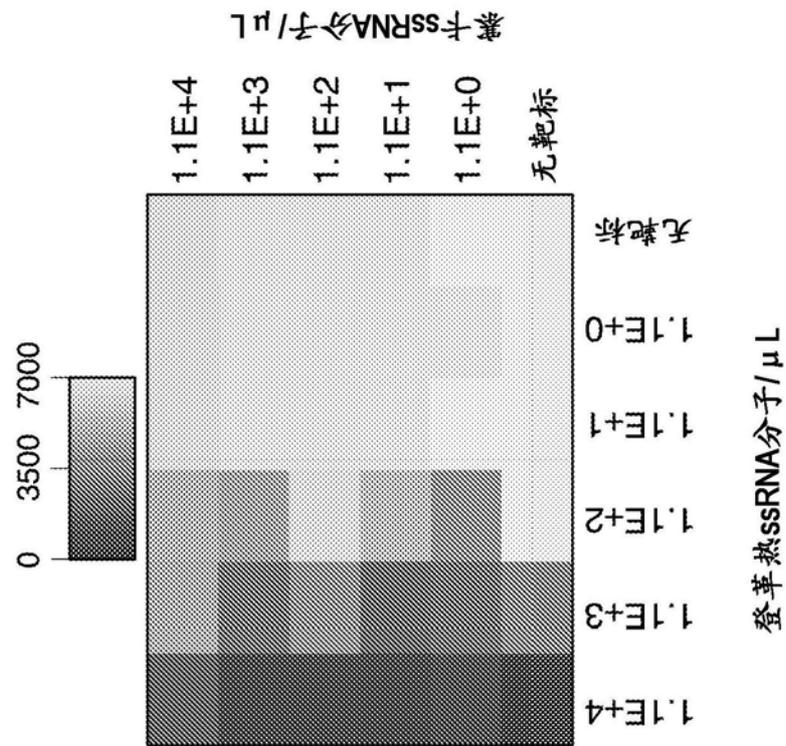


图102

A

PsmCas13通道-登革热



B

LwaCas13a通道-寨卡区域4

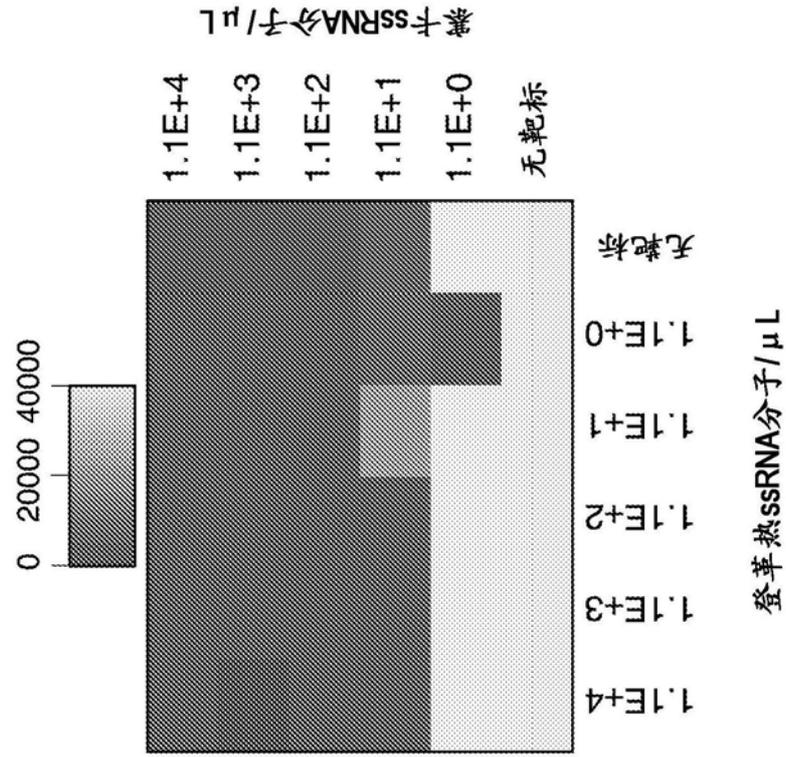


图103

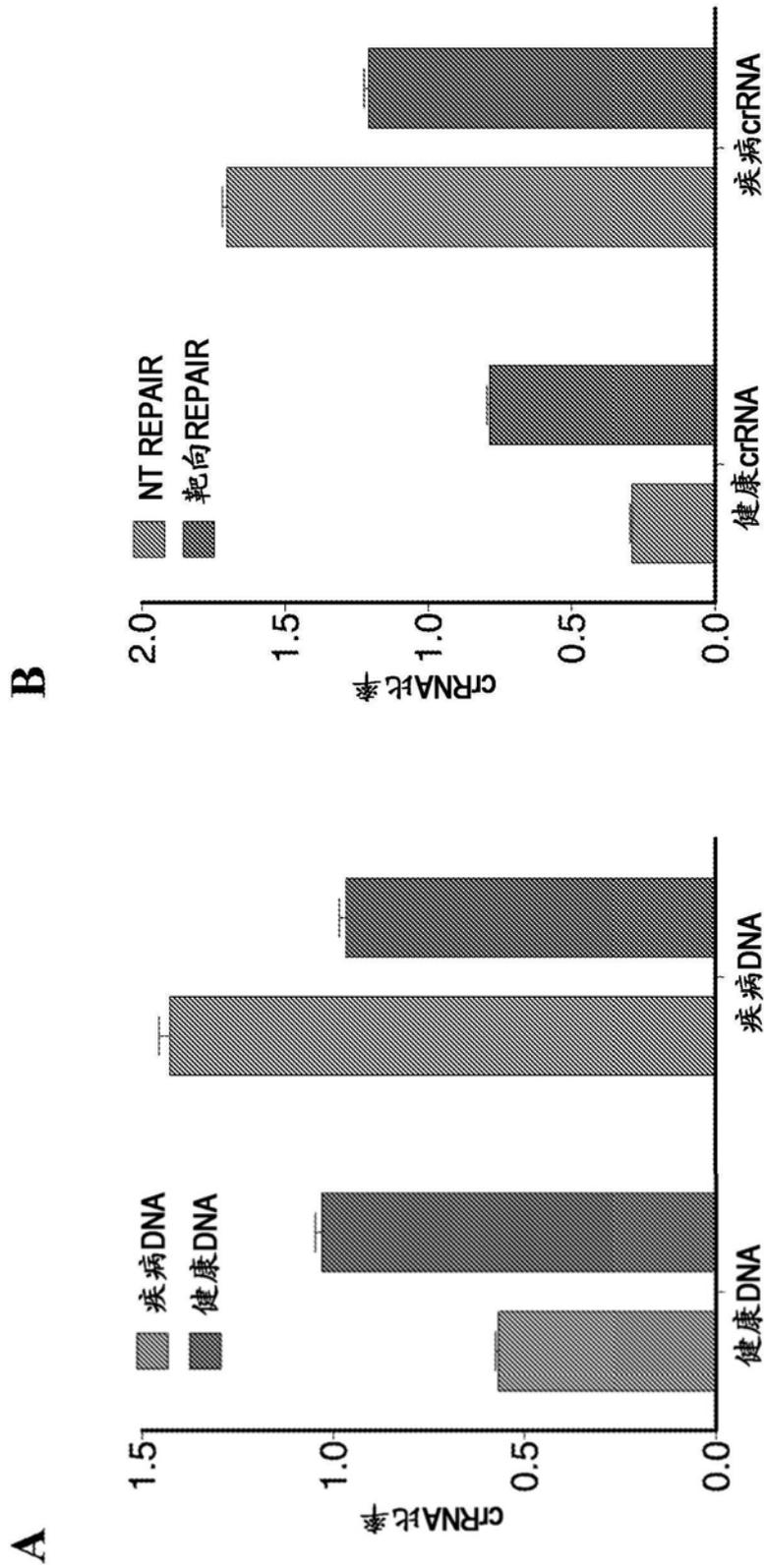


图104

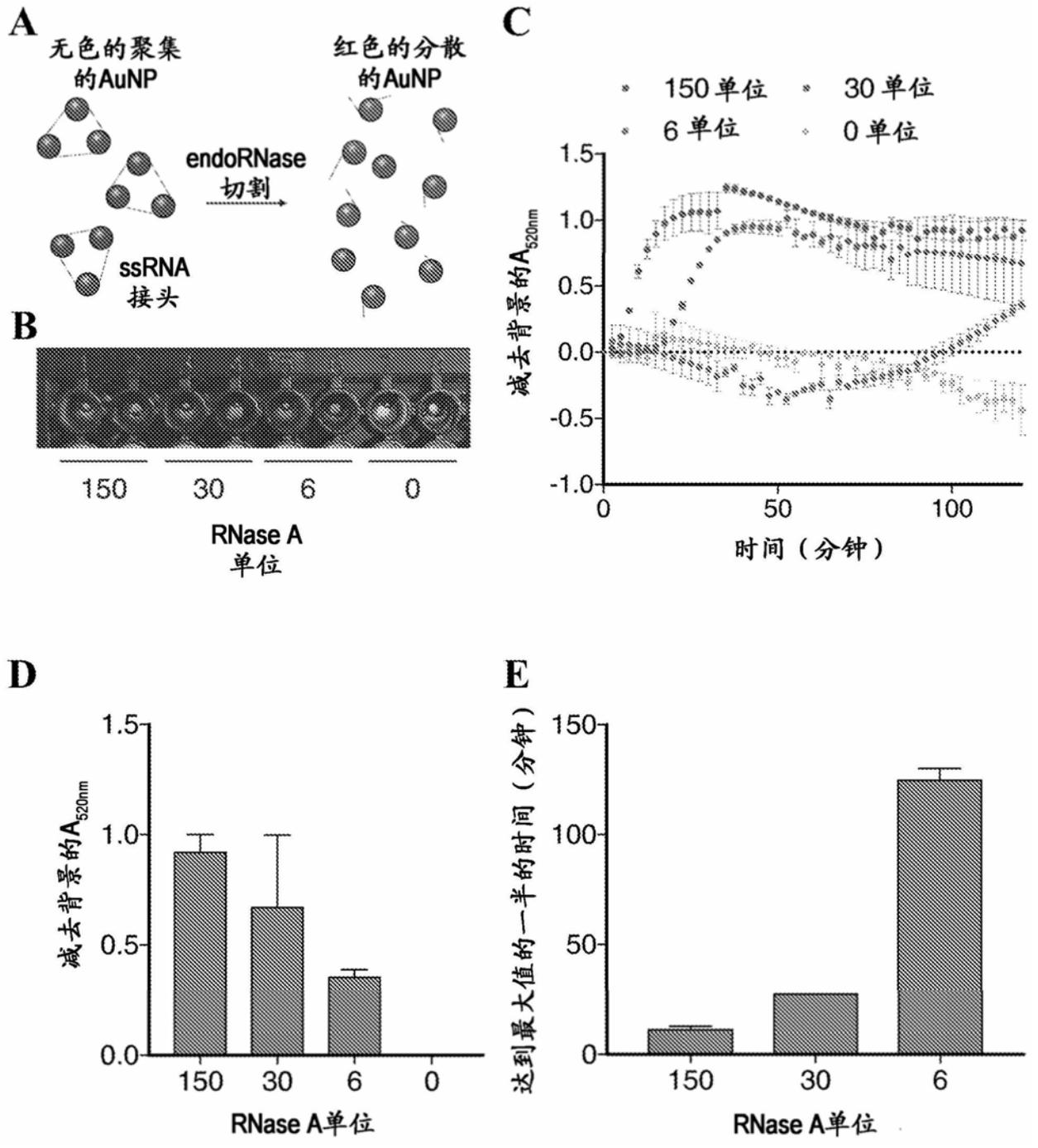


图105

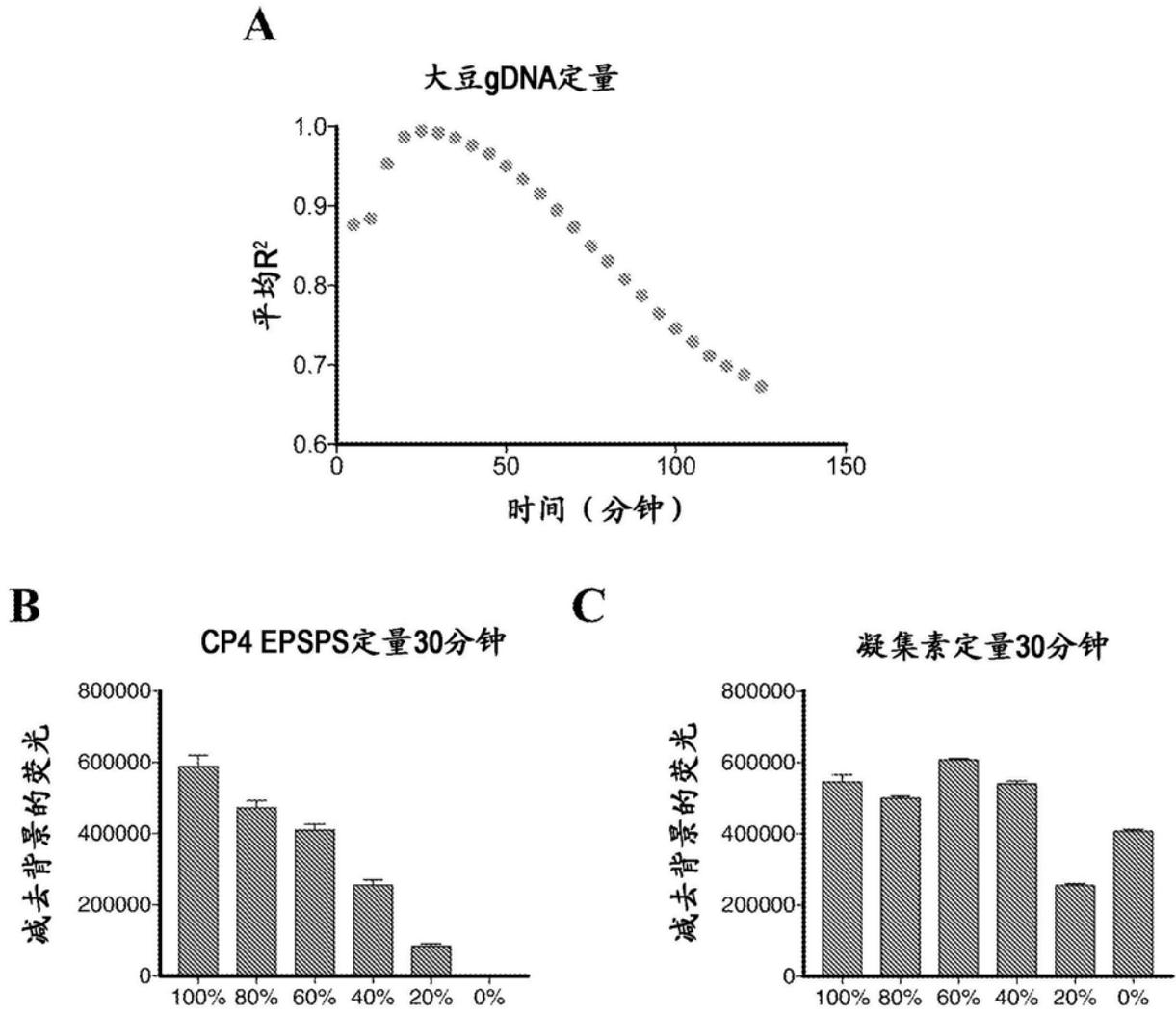


图106