(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 111448311 B
(45)授权公告日 2024.11.19

(21)申请号 201880072772.2
(22)申请日 2018.09.07
(65)同一申请的已公布的文献号
申请公布号 CN 111448311 A
(43)申请公布日 2020.07.24
(30)优先权数据
62/556,408 2017.09.09 US
62/610,121 2017.12.22 US
62/630,808 2018.02.14 US
(85)PCT国际申请进入国家阶段日
2020.05.09
(86)PCT国际申请的申请数据
PCT/US2018/050091 2018.09.07
(87)PCT国际申请的公布数据
WO2019/051318 EN 2019.03.14
(73)专利权人 博德研究所
地址 美国马萨诸塞州
专利权人 麻省理工学院 哈佛学院院长等
(72)发明人 F·张 J·戈滕贝格 O·阿布达耶
(74)专利代理机构 北京市金杜律师事务所 11256
专利代理师 陈文平 袁元
(51)Int.Cl.
C12N 9/22(2006.01)
C12N 15/00(2006.01)
(56)对比文件
WO 2017070605 A1,2017.04.27
Alexandra East-Seletsky et al..“Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection”.《Nature》.2016,第538卷第270-273页和Methods部分.
Jonathan S. Gootenberg et al..“Nucleic acid detection with CRISPR-Cas13a/C2c2”.《Science》.2017,第1-9页.
审查员 赵二双

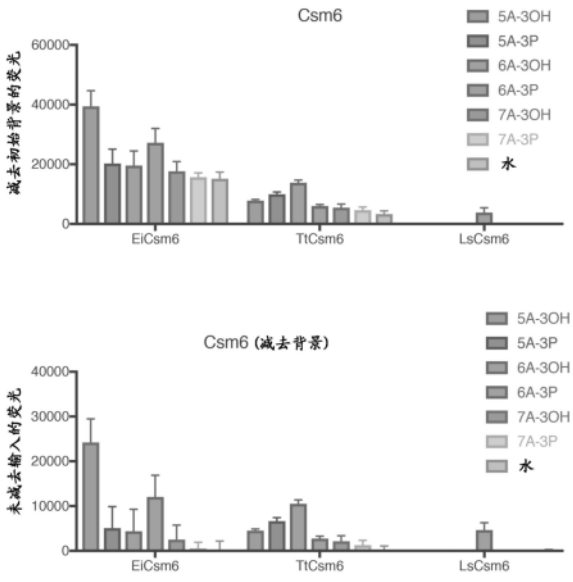
权利要求书7页 说明书92页
序列表（电子公布） 附图53页

(54)发明名称

基于多效应子CRISPR的诊断系统

(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。



CN 111448311 B

1.一种核酸检测系统,其包括:

a)检测CRISPR系统,其包括:

i.在激活后具有附带切割活性的检测VI型CRISPR效应蛋白,

ii.一种或多种设计用于杂交至相应靶核酸序列的指导RNA,和

b)在激活后具有附带切割活性的一种或多种IIIa型信号放大CRISPR Csm6效应蛋白;和

c)一种或多种基于RNA的掩蔽构建体,其包含被所述检测VI型CRISPR效应蛋白和/或所述一种或多种IIIa型信号放大CRISPR Csm6效应蛋白的附带切割活性所失活的非靶核酸序列,

其中所述检测VI型CRISPR效应蛋白和杂交至相应靶核酸序列的所述一种或多种指导RNA之间的复合物形成激活所述检测VI型CRISPR效应蛋白的附带切割活性;

其中激活的检测VI型CRISPR效应蛋白切割激活序列,其切割产物激活所述一种或多种VIIIa型信号放大CRISPR Csm6效应蛋白的附带切割活性;并且

其中所述检测VI型CRISPR效应蛋白和所述一种或多种IIIa型信号放大CRISPR Csm6效应蛋白的附带切割活性协同作用以失活所述基于RNA的掩蔽构建体以产生或掩蔽可检测的信号。

2.根据权利要求1所述的系统,其中所述Csm6效应蛋白选自EiCsm6和LsCsm6。

3.根据权利要求1所述的系统,其中所述靶核酸序列是靶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。

4.根据权利要求1所述的系统,其中所述检测VI型CRISPR效应蛋白是RNA靶向效应蛋白。

5.根据权利要求4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。

6.根据权利要求5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。

7.根据权利要求6所述的系统,其中所述RxxxxH基序包含R{N/H/K}$X_1X_2X_3$H序列。

8.根据权利要求7所述的系统,其中X_1是R、S、D、E、Q、N、G或Y,并且X_2独立地是I、S、T、V或L,并且X_3独立地是L、F、N、Y、V、I、S、D、E或A。

9.根据权利要求4所述的系统,其中所述RNA靶向效应蛋白是C2c2。

10.根据权利要求9所述的系统,其中所述C2c2效应蛋白是Leptotrichia wadeiF0279 C2c2效应蛋白。

11.根据权利要求1所述的系统,其中所述一种或多种基于RNA的掩蔽构建体包括RNA适体、RNA系连的抑制剂、干扰RNA、RNA寡核苷酸或核酶。

12.根据权利要求11所述的系统,其中所述一种或多种基于RNA的掩蔽构建体抑制可检测阳性信号的产生。

13.根据权利要求12所述的系统,其中所述一种或多种基于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。

14.根据权利要求11所述的系统,其中干扰RNA抑制由报告构建体编码的基因产物的产生,其中所述基因产物在表达时产生可检测阳性信号。

15.根据权利要求11所述的系统,其中核酶产生所述可检测阴性信号,并且其中当所述

核酶失活时产生所述可检测阳性信号。

16.根据权利要求11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。

17.根据权利要求11所述的系统,其中所述适体或RNA系连的抑制剂螯合酶,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。

18.根据权利要求11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。

19.根据权利要求18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。

20.根据权利要求19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。

21.根据权利要求11所述的系统,其中所述适体螯合一对试剂,所述试剂当从所述适体释放时组合以产生可检测信号。

22.根据权利要求11所述的系统,其中所述一种或多种基于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。

23.根据权利要求11所述的系统,其中所述一种或多种基于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。

24.根据权利要求23所述的系统,其中所述纳米颗粒是胶体金属。

25.根据权利要求24所述的系统,其中所述胶体金属是胶体金。

26.根据权利要求11所述的系统,其中所述一种或多种基于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。

27.根据权利要求11所述的系统,其中所述一种或多种基于RNA的掩蔽构建体包含与插入染料复合的RNA,其中所述插入染料在所述RNA的切割后改变吸光度。

28.根据权利要求27所述的系统,其中所述插入染料是吡罗红-Y或亚甲蓝。

29.根据权利要求22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。

30.根据权利要求11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。

31.根据权利要求1所述的系统,其中所述一种或多种基于RNA的掩蔽构建体被所述检测VI型CRISPR效应蛋白和所述一种或多种IIIa型信号放大CRISPR Csm6效应蛋白切割。

32.根据权利要求31所述的系统,其中所述一种或多种基于RNA的掩蔽构建体包括被所述检测VI型CRISPR效应蛋白切割的基于RNA的掩蔽构建体。

33.根据权利要求31所述的系统,其中所述一种或多种基于RNA的掩蔽构建体包括被所述一种或多种IIIa型信号放大CRISPR Csm6效应蛋白切割的基于RNA的掩蔽构建体。

34.根据权利要求31所述的系统,其中所述一种或多种基于RNA的掩蔽构建体包括被所

述检测VI型CRISPR效应蛋白切割的一种基于RNA的掩蔽构建体和被所述一种或多种IIIa型信号放大CRISPR Csm6效应蛋白切割的一种基于RNA的掩蔽构建体。

35.根据权利要求31所述的系统,其中被所述一种或多种IIIa型信号放大CRISPR Csm6效应蛋白切割的所述基于RNA的掩蔽构建体包含均聚物A-RNA或C-RNA。

36.根据权利要求1所述的系统,其中所述一种或多种设计用于结合相应靶核酸序列的指导RNA包含错配。

37.根据权利要求1所述的系统,其中所述一种或多种设计用于结合相应靶核酸序列的指导RNA包含合成错配。

38.根据权利要求36所述的系统,其中所述错配是在所述靶核酸序列中单核苷酸多态性或其他单核苷酸变异的上游或下游。

39.根据权利要求37所述的系统,其中所述指导RNA中的所述合成错配位于从间隔子的5'末端计数的位置3、4、5或6处。

40.根据权利要求37所述的系统,其中所述指导RNA中的所述合成错配位于从间隔子的5'末端计数的位置3处。

41.根据权利要求36所述的系统,其中所述指导RNA中的所述错配位于从间隔子的5'末端计数的位置1、2、3、4、5、6、7、8或9处。

42.根据权利要求37所述的系统,其中所述指导RNA中的所述合成错配位于从间隔子的5'末端计数的位置5处。

43.根据权利要求36所述的系统,其中所述错配是在所述指导RNA中的所述单核苷酸多态性或其他单核苷酸变异的上游或下游的1、2、3、4或5个核苷酸。

44.根据权利要求36所述的系统,其中所述错配是所述指导RNA中的所述单核苷酸多态性或其他单核苷酸变异的下游的2个核苷酸。

45.根据权利要求1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。

46.根据权利要求45所述的系统,其中所述指导RNA包含少于28个核苷酸的间隔子。

47.根据权利要求45所述的系统,其中所述指导RNA包含间隔子,所述间隔子包含在20至27个核苷酸之间并且包括20和27个核苷酸。

48.根据权利要求45所述的系统,其中所述指导RNA包含由20-25个核苷酸组成的间隔子。

49.根据权利要求45所述的系统,其中所述指导RNA包含由20-23个核苷酸组成的间隔子。

50.根据权利要求45所述的系统,其中所述指导RNA包含由20或23个核苷酸组成的间隔子。

51.根据权利要求1至50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。

52.根据权利要求1至50中任一项所述的系统,其中所述系统还包含一种或多种激活序列,并且其中每种激活序列包含聚A段和/或当被所述检测CRISPR RNA靶向效应蛋白切割时产生含有2'3'环状磷酸酯末端的六腺苷酸。

53.根据权利要求1至50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核酸序列。

54.根据权利要求53所述的系统,其中所述疾病状态是癌症。

55.根据权利要求53所述的系统,其中所述疾病状态是自身免疫疾病。

56.根据权利要求53所述的系统,其中所述疾病状态是感染。

57.根据权利要求56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。

58.根据权利要求56所述的系统,其中所述感染是病毒感染。

59.根据权利要求58所述的系统,其中所述病毒感染是由DNA病毒引起。

60.根据权利要求59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、Malocoherpesviridae、脂毛病毒科、腺病毒科、Ampullaviridae、囊泡病毒科、非洲猪瘟病毒科、杆状病毒科、Cicaudaviridae、Clavaviridae,覆盖噬菌体科、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、痘病毒科、Sphaerolipoviridae、复层病毒科、Turriviridae、Dinodnavirus、盐末端蛋白病毒属或Rhizidovirus。

61.根据权利要求59所述的系统,其中所述DNA病毒是小杆状病毒科、人类疱疹病毒、水痘带状疱疹病毒、非洲猪瘟病毒、猿猴病毒40、JC病毒、BK病毒、包括牛痘或天花。

62.根据权利要求57所述的系统,其中所述病毒是双链RNA病毒、正义RNA病毒、负义RNA病毒、逆转录病毒或其组合。

63.根据权利要求57所述的系统,其中所述病毒是冠状病毒科病毒、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。

64.根据权利要求63所述的系统,其中所述病毒是冠状病毒、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。

65.根据权利要求63所述的系统,其中所述病毒是SARS。

66.根据权利要求57所述的系统,其中所述感染是细菌感染。

67.根据权利要求66所述的系统,其中所述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、分支杆菌属种、支原体属种、诺卡氏菌属种、奈瑟氏菌属种、Pasteurelaa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。

68.根据权利要求57所述的系统,其中所述感染由真菌引起。

69.根据权利要求68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、肺孢子虫属种、葡萄穗霉属、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。

70.根据权利要求68所述的系统,其中所述真菌是荚膜组织胞浆菌、耶氏肺孢子虫、纸葡萄穗霉或其组合。

71.根据权利要求57所述的系统,其中所述感染是由原生动物引起。

72.根据权利要求71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。

73.根据权利要求57所述的系统,其中所述感染是由寄生虫引起。

74.根据权利要求73所述的系统,其中所述寄生虫是克氏锥虫、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、canthamoeba castellanii、Balamuthia madrillaris、Entamoeba histolytica、Blastocystic hominis、Babesia microti、Cryptosporidium parvum、Cyclospora cayetanensis、Plasmodium falciparum、P.vivax、P.ovale、P.malariae和Toxoplasma gondii或其组合。

75.根据权利要求73所述的系统,其中所述寄生虫是克氏锥虫、G.lamblia、G.duodenalis或其组合。

76.根据权利要求53所述的系统,其中所述靶核酸序列是药物耐受性基因。

77.根据权利要求53所述的系统,其中所述靶核酸序列是针对恶性疟原虫,间日疟原虫,卵形疟原虫,三日疟原虫,和诺氏疟原虫的治疗的药物耐受性的诊断剂。

78.根据权利要求53所述的系统,其中所述靶核酸序列是抗生素耐受性或易感性基因或转录物。

79.根据权利要求1所述的系统,其还包括核酸扩增试剂。

80.根据权利要求79所述的系统,其中所述核酸扩增试剂包括扩增靶RNA分子的试剂,所述扩增靶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)。

81.根据权利要求3所述的系统,其中所述靶核酸序列包含靶RNA分子的核糖核酸序列

或靶DNA分子的互补核糖核酸序列。

82.根据权利要求1所述的系统,其还包括富集CRISPR系统,其包含催化失活的VI型CRISPR效应蛋白。

83.根据权利要求82所述的系统,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核酸序列。

84.根据权利要求82所述的系统,其中所述催化失活的VI型CRISPR效应蛋白是催化失活的C2c2。

85.根据权利要求82至84中任一项所述的系统,其中所述富集CRISPR效应系统还包含标签,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。

86.根据权利要求85所述的系统,其中所述固体基质是流动池。

87.一种用于检测样品中的靶核酸序列的诊断装置,其包括一个或多个个体离散体积,每个个体离散体积包含限定量的权利要求1至86中任一项所述的核酸检测系统用于与预定体积的样品混合。

88.根据权利要求87所述的装置,其中所述个体离散体积是液滴。

89.根据权利要求87所述的装置,其中所述个体离散体积限定在固体基质上。

90.根据权利要求87所述的装置,其中所述个体离散体积是微孔。

91.根据权利要求87所述的装置,其中所述个体离散体积是在基质上限定的斑点。

92.根据权利要求91所述的装置,其中所述基质是柔性材料基质。

93.根据权利要求92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。

94.一种用于检测样品中的靶核酸序列的方法,其包括:

将一个或多个样品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至86中任一项所述的核酸检测系统;

在允许所述一种或多种指导RNA与一种或多种靶核酸序列结合的条件下孵育所述一个或多个样品;

通过所述一种或多种指导RNA与所述一种或多种靶核酸序列的结合激活所述检测CRISPR效应蛋白,

其中激活所述检测CRISPR效应蛋白导致所述一种或多种基于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核酸序列在所述样品中的存在,

其中所述方法不是疾病的治疗或诊断方法。

95.根据权利要求94所述的方法,其中所述靶核酸序列是靶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。

96.根据权利要求94或95所述的方法,其还包括扩增所述样品RNA。

97.根据权利要求96所述的方法,其中扩增所述样品RNA包括通过NASBA进行扩增。

98.根据权利要求96所述的方法,其中扩增所述样品RNA包括通过RPA进行扩增。

99.根据权利要求94或95所述的方法,其中所述样品是生物样品或环境样品。

100.根据权利要求99所述的方法,其中生物样品是血液、血浆、血清、尿液、粪便、粘液、淋巴液、滑液、胆汁、腹水、胸腔积液、血清肿、唾液、脑脊液、房水或玻璃体液、或从关节获得的液体、或者皮肤或粘膜表面的拭子。

101.根据权利要求99所述的方法,其中生物样品是从脓肿或任何其他的感染或炎症部位获得的液体。

102.根据权利要求100所述的方法,其中从关节获得的液体是从正常关节或受疾病影响的关节获得的液体。

103.根据权利要求102所述的方法,其中所述疾病是类风湿性关节炎、骨关节炎、痛风或化脓性关节炎。

104.根据权利要求99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。

105.根据权利要求94或95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。

106.根据权利要求94或95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。

107.根据权利要求94或95所述的方法,其还包括比较所述可检测阳性信号和标准信号。

108.一种用于检测样品中的靶核酸序列的方法,其包括:

将样品与权利要求1至86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,

其中所述方法不是疾病的治疗或诊断方法。

109.根据权利要求108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。

110.根据权利要求109所述的方法,其中所述第一和第二分子在侧流条的离散检测位点处检测。

111.根据权利要求109所述的方法,其中通过所述第一分子和所述第二分子结合识别所述第一或第二分子的抗体并且检测结合的分子来检测所述第一分子和所述第二分子。

112.根据权利要求111所述的方法,其中所述抗体是夹心抗体。

113.根据权利要求110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。

114.根据权利要求1所述的系统,其中所述检测VI型CRISPR效应蛋白是Cas13a效应蛋白、Cas13b效应蛋白或Cas13d效应蛋白。

基于多效应子CRISPR的诊断系统

[0001] 相关申请的交叉引用

[0002] 本申请要求2017年9月9日提交的美国临时申请No.62/556,408;2017年12月22日提交的美国临时申请No.62/610,121;和2018年2月14日提交的美国临时申请No.62/630,808的权益。上述申请的全部内容通过引用全部并入本文。

[0003] 关于美国联邦政府资助研究的声明

[0004] 本发明是在美国国立卫生研究院授予的拨款号MH110049的美国政府支持下完成的。美国政府拥有本发明的某些权利。

技术领域

[0005] 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

[0006] 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

[0007] 在一个方面,本发明提供了一种核酸检测系统,其包括:检测CRISPR系统,其包括:效应蛋白,一种或多种设计用于结合相应靶分子的指导RNA(guide RNA),和一种或多种信号放大CRISPR效应蛋白;和一种或多种基于RNA的掩蔽构建体(masking construct)。

[0008] 在另一个方面,本发明提供了一种多肽检测系统,其包括:检测CRISPR系统,其包括:效应蛋白,一种或多种设计用于结合触发RNA(trigger RNA)的指导RNA,和一种或多种信号放大CRISPR效应蛋白;一种或多种基于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。

[0009] 在某些实施方式中,所述一种或多种信号放大CRISPR效应蛋白包括IIIa型CRISPR蛋白。所述III型CRISPR蛋白可以是Csm6蛋白。所述Csm6蛋白可以选自EiCsm6和LsCsm6。

[0010] 在某些实施方式中,所述一种或多种信号放大CRISPR效应蛋白包括Csx28或Csx27。

[0011] 在某些实施方式中,所述一种或多种信号放大CRISPR效应蛋白包括Csm6,Csx28,

Csx27或其任何组合中的一种或多种。

[0012] 在某些实施方式中,所述系统还包括核酸扩增试剂。

[0013] 在某些实施方式中,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。

[0014] 在某些实施方式中,所述CRISPR系统效应蛋白是RNA靶向效应蛋白。所述RNA靶向效应蛋白可以包含一个或多个HEPN结构域。所述一个或多个HEPN结构域可以包含RxxxxH基序序列。所述RxxxH基序可以包含R{N/H/K]$X_1X_2X_3$H序列。X_1可以是R、S、D、E、Q、N、G或Y,并且X_2可以独立地是I、S、T、V或L,并且X_3可以独立地是L、F、N、Y、V、I、S、D、E或A。

[0015] 在某些实施方式中,所述CRISPR RNA靶向效应蛋白是C2c2。所述C2c2可以是在Cas1基因的20kb内。所述C2c2效应蛋白可以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)、Nitratifractor、支原体属(Mycoplasma)、弯曲杆菌属(Campylobacter)和毛螺菌属(Lachnospira)。所述C2c2或Cas13b效应蛋白可以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(Clostridium)]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;Lachnospiraceae bacterium NK4A144;Chloroflexus aggregans;Demequina aurantiaca;Thalassospira sp.TSL5-1;Pseudobutyrivibrio sp.OR37;Butyrivibrio sp.YAB3001;Blautia sp.Marseille-P2398;Leptotrichia sp.Marseille-P3007;Bacteroides ihuae;Porphyromonadaceae bacterium KH3CP3RA;Listeria riparia;和Insolitispirillum peregrinum。所述C2c2效应蛋白可以是L.wadei F0279或L.wadei F0279(Lw2)C2c2效应蛋白。

[0016] 在某些实施方式中,所述一种或多种基于RNA的掩蔽构建体抑制可检测阳性信号的产生。

[0017] 在某些实施方式中,所述一种或多种基于RNA的掩蔽构建体可以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。

[0018] 在某些实施方式中,所述一种或多种基于RNA的掩蔽构建体可以包含抑制由报告构建体编码的基因产物的产生的沉默RNA(silencing RNA),其中所述基因产物在表达时产

生所述可检测阳性信号。

[0019] 在某些实施方式中,所述一种或多种基于RNA的掩蔽构建体可以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。所述核酶可以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。

[0020] 在某些实施方式中,所述基于RNA的掩蔽剂是RNA适体和/或包含RNA系连的抑制剂(RNA-tethered inhibitor)。

[0021] 所述适体或RNA系连的抑制剂可以螯合酶,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。所述适体可以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。所述酶可以是凝血酶、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(calf alkaline phosphatase)。所述酶可以是凝血酶,并且所述底物可以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。

[0022] 在某些实施方式中,所述适体螯合一对试剂,所述试剂当从所述适体释放时组合以产生可检测信号。

[0023] 在某些实施方式中,所述一种或多种基于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。

[0024] 在某些实施方式中,所述一种或多种基于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。所述纳米颗粒可以是胶体金属。所述胶体金属可以是胶体金。

[0025] 在某些实施方式中,所述一种或多种基于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。

[0026] 在某些实施方式中,所述一种或多种基于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。所述插入剂可以是吡罗红-Y(pyronine-Y)或亚甲蓝(methylene blue)。

[0027] 在某些实施方式中,所述可检测配体是荧光团,并且所述掩蔽组分是猝灭分子。

[0028] 在某些实施方式中,所述一种或多种基于RNA的掩蔽构建体可以被所述CRISPR RNA靶向效应蛋白和IIIa型CRISPR蛋白切割。换句话说,所述基于RNA的掩蔽构建体可以是可以被所述CRISPR RNA靶向效应蛋白和IIIa型CRISPR蛋白切割的单一构建体,或者可以被两者切割的多于一种构建体,或者可以被任一者切割的多于一种构建体。所述一种或多种基于RNA的掩蔽构建体可以包括可以被所述CRISPR RNA靶向效应蛋白切割的基于RNA的掩蔽构建体。所述一种或多种基于RNA的掩蔽构建体可以包括可以被IIIa型CRISPR蛋白切割的基于RNA的掩蔽构建体。所述一种或多种基于RNA的掩蔽构建体可以包括可以被所述CRISPR RNA靶向效应蛋白切割的一种基于RNA的掩蔽构建体和可以被所述IIIa型CRISPR蛋白切割的一种基于RNA的掩蔽构建体。可以被所述IIIa型CRISPR蛋白切割的所述基于RNA的掩蔽构建体可以包含均聚物A(homopolymeric A)或C-RNA。

[0029] 在某些实施方式中,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。所述错配可以是所述靶分子中SNP或其他单核苷酸变异的上游或下游。

[0030] 在某些实施方式中,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。

[0031] 在某些实施方式中,所述一种或多种指导RNA被设计用于检测当被所述CRISPR RNA靶向效应蛋白切割时产生含有2'3'环状磷酸酯末端(2'3'cyclic phosphate end)的六腺苷酸(hexadenylates)的靶分子或触发RNA。所述一种或多种指导RNA可以被设计用于检测包含聚A段(poly A stretch)的靶分子或触发RNA。

[0032] 在某些实施方式中,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。所述疾病状态可以是癌症。所述疾病状态可以是自身免疫疾病。所述疾病状态可以是感染。

[0033] 在某些实施方式中,所述感染是由病毒、细菌、真菌、原生动物或寄生虫引起。

[0034] 在某些实施方式中,所述感染可以是病毒感染。所述病毒感染可以由DNA病毒引起。所述DNA病毒可以是肌病毒科(Myoviridae)、短尾病毒科(Podoviridae)、长尾病毒科(Siphoviridae)、异疱疹病毒科(Alloherpesviridae)、疱疹病毒科(Herpesviridae)(包括人类疱疹病毒和水痘带状疱疹(Varicella Zoster)病毒)、Malocoherpesviridae、脂毛病毒科(Lipothrixviridae)、小杆状病毒科(Rudiviridae)、腺病毒科(Adenoviridae)、Ampullaviridae、囊泡病毒科(Ascoviridae)、非洲猪瘟病毒科(Asfarviridae)(包括非洲猪瘟病毒)、杆状病毒科(Baculoviridae)、Cicaudaviridae、Clavaviridae,覆盖病毒科(Corticoviridae)、小纺锤形病毒科(Fuselloviridae)、球状病毒科(Globuloviridae)、Guttaviridae、Hytrosaviridae、虹彩病毒科(Iridoviridae)、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(Poxviridae)(包括牛痘和天花)、Sphaerolipoviridae、复层病毒科(Tectiviridae)、Turriviridae、Dinodnavirus、盐末端蛋白病毒属(Salterprovirus)、Rhizidovirus。所述病毒感染可以由双链RNA病毒、正义RNA病毒、负义RNA病毒、逆转录病毒或其组合引起。所述病毒感染可以由冠状病毒科病毒、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。所述病毒感染可以由冠状病毒、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。

[0035] 在某些实施方式中,所述感染是细菌感染。引起所述细菌感染的细菌可以是不动杆菌属种(Acinetobacter species)、放线杆菌属种(Actinobacillus species)、放线菌纲种(Actinomycetes species)、放线菌属种(Actinomyces species)、气球菌属种(Aerococcus species)、气单胞菌属种(Aeromonas species)、红孢子虫属种(Anaplasma species)、产碱杆菌属种(Alcaligenes species)、芽孢杆菌属种(Bacillus species)、拟杆菌属种(Bacteriodes species)、巴尔通氏体属种(Bartonella species)、双歧杆菌属种(Bifidobacterium species)、博代氏杆菌属种(Bordetella species)、包柔氏螺旋体属种(Borrelia species)、布鲁氏菌属种(Brucella species)、伯克霍尔德氏菌属种

(Burkholderia species)、弯曲杆菌属种(Campylobacter species)、嗜二氧化碳噬细胞菌属种(Capnocytophaga species)、衣原体属种(Chlamydia species)、柠檬酸杆菌属种(Citrobacter species)、柯克斯氏体属种(Coxiella species)、棒状杆菌属种(Corynbacterium species)、梭菌属种(Clostridium species)、艾肯菌属种(Eikenella species)、肠杆菌属种(Enterobacter species)、埃希氏菌属种(Escherichia species)、肠球菌属种(Enterococcus species)、埃立克体属种(Ehlichia species)、表皮癣菌属种(Epidermophyton species)、丹毒丝菌属种(Erysipelothrix species)、真细菌属种(Eubacterium species)、弗兰西斯菌属种(Francisella species)、梭杆菌属种(Fusobacterium species)、加德纳菌属种(Gardnerella species)、孪生球菌属种(Gemella species)、嗜血杆菌属种(Haemophilus species)、螺杆菌属种(Helicobacter species)、金氏菌属种(Kingella species)、克雷白氏杆菌属种(Klebsiella species)、乳酸菌属种(Lactobacillus species)、乳球菌属种(Lactococcus species)、李斯特菌属种(Listeria species)、钩端螺旋体属种(Leptospira species)、军团菌属种(Legionella species)、钩端螺旋体属种(Leptospira species)、明串珠菌属种(Leuconostoc species)、曼氏杆菌属种(Mannheimia species)、小孢子菌属种(Microsporum species)、微球菌属种(Micrococcus species)、莫拉克斯氏菌属种(Moraxella species)、摩根氏菌属种(Morganell species)、动弯杆菌属种(Mobiluncus species)、微球菌属种(Micrococcus species)、分支杆菌属种(Mycobacterium species)、支原体属种(Mycoplasm species)、诺卡氏菌属种(Nocardia species)、奈瑟氏菌属种(Neisseria species)、Pasteurelaa种、片球菌属种(Pediococcus species)、消化链球菌属种(Peptostreptococcus species)、糠疹癣菌属种(Pityrosporum species)、邻单胞菌属种(Plesiomonas species)、普氏菌属种(Prevotella species)、卟啉单胞菌属种(Porphyromonas species)、变形杆菌属种(Proteus species)、普罗维登斯菌属种(Providencia species)、假单胞菌属种(Pseudomonas species)、丙酸菌属种(Propionibacteriums species)、红球菌属种(Rhodococcus species)、立克次氏体属种(Rickettsia species)、红球菌属种(Rhodococcus species)、沙雷氏菌属种(Serratia species)、寡养单胞菌属种(Stenotrophomonas species)、沙门氏菌属种(Salmonella species)、沙雷氏菌属种(Serratia species)、志贺氏菌属种(Shigella species)、葡萄球菌属种(Staphylococcus species)、链球菌属种(Streptococcus species)、螺旋菌属种(Spirillum species)、链杆菌属种(Streptobacillus species)、密螺旋体属种(Treponema species)、Tropheryma种、毛癣菌属种(Trichophyton species)、脲原体属种(Ureaplasma species)、韦永氏球菌属种(Veillonella species)、弧菌属种(Vibrio species)、耶尔森氏鼠疫杆菌属种(Yersinia species)、黄杆菌属种(Xanthomonas species)或其组合。

[0036] 在某些实施方式中,所述感染由真菌引起。所述真菌可以是曲霉属(Aspergillus)、芽生菌属(Blastomyces)、念珠菌病(Candidiasis)、球孢子菌病(Coccidiodomycosis)、新型隐球菌(Cryptococcus neoformans)、加蒂隐球菌种(Cryptococcus gatti,sp.)、组织胞浆菌属种(Histoplasma sp.)(例如荚膜组织胞浆菌(Histoplasma capsulatum))、肺孢子虫属种(Pneumocystis sp.)(例如耶氏肺孢子虫

(Pneumocystis jirovecii))、葡萄穗霉属(Stachybotrys)(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、突脐蠕孢属(Exserohilum)、枝孢属(Cladosporium)、地丝菌属(Geotrichum)、酵母属(Saccharomyces)、汉逊酵母属种(Hansenula species)、念珠菌属种(Candida species)、克鲁维酵母属种(Kluyveromyces species)、德巴利氏酵母属种(Debaryomyces species)、毕赤酵母属种(Pichia species)、青霉菌属种(Penicillium species)、枝孢属种(Cladosporium species)、丝衣酶属种(Byssochlamys species)或其组合。

[0037] 在某些实施方式中,所述感染是由原生动物引起。所述原生动物可以是眼虫动物界(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)或其组合。

[0038] 在某些实施方式中,所述感染是由寄生虫引起。所述寄生虫可以是克氏锥虫(Trypanosoma cruzi)(查加斯病)、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、Naegleria fowleri、Giardia intestinalis(G.lamblia、G.duodenalis)、canthamoeba castellanii、Balamuthia madrillaris、Entamoeba histolytica、Blastocystic hominis、Babesia microti、Cryptosporidium parvum、Cyclospora cayetanensis、Plasmodium falciparum、P.vivax、P.ovale、P.malariae和Toxoplasma gondii或其组合。

[0039] 在某些实施方式中,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)。

[0040] 在某些实施方式中,所述系统还包括富集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。所述富集CRISPR系统可以包括催化失活的CRISPR效应蛋白。所述催化失活的CRISPR效应蛋白可以是催化失活的C2c2。所述富集CRISPR效应蛋白可以还包含标签,其中所述标签用于拉下(pull down)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。所述固体基质可以是流动池(flow cell)。

[0041] 在某些实施方式中,所述指导RNA中的所述合成错配位于间隔子(spacer)的位置3、4、5或6处,优选位置3。所述指导RNA中的所述错配可以位于间隔子的位置1、2、3、4、5、6、7、8或9处,优选位置5。所述错配可以是所述指导RNA中的所述SNP或其他单核苷酸变异的上游或下游的1、2、3、4或5个核苷酸,优选2个核苷酸,优选下游。

[0042] 在某些实施方式中,所述指导RNA可以包含相对于野生型间隔子截短的间隔子。在某些实施方式中,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。所述指导RNA可以包含由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。

[0043] 在某些实施方式中,所述掩蔽构建体包含设计用于结合G-四链体(G-quadruplex)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。

[0044] 在另一个方面,本方法提供了一种诊断装置,其包括一个或多个个体离散体积(individual descrete volume),每个个体离散体积包含本文中任何实施方式的CRISPR系统。

[0045] 在某些实施方式中,每个个体离散体积可以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。

[0046] 在某些实施方式中,每个个体离散体积还包含核酸扩增试剂。

[0047] 在某些实施方式中,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。

[0048] 在某些实施方式中,所述个体离散体积是液滴。

[0049] 在某些实施方式中,所述个体离散体积限定在固体基质上。所述个体离散体积可以是微孔。所述个体离散体积可以是在基质上限定的斑点(spot)。所述基质可以是柔性材料基质。所述柔性材料基质可以是纸基质或基于柔性聚合物的基质。

[0050] 在另一个方面,本发明提供了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(set of samples)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。

[0051] 在另一个方面,本发明提供了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、如本文所述的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。

[0052] 在某些实施方式中,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。

[0053] 在某些实施方式中,所述方法还包括扩增所述样品RNA或所述触发RNA。扩增RNA可以包括通过NASBA进行扩增。扩增RNA可以包括通过RPA进行扩增。

[0054] 在某些实施方式中,所述样品是生物样品或环境样品。所述生物样品可以是血液、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、房水或玻璃体液、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。所述环境样品可以从食品样品、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。

[0055] 在某些实施方式中,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。

[0056] 在某些实施方式中,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。

[0057] 在某些实施方式中,所述一种或多种指导RNA被设计用于结合无细胞核酸。

[0058] 在某些实施方式中,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。所述疾病状态的特征可以在于存在或不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存在。所述靶分子可以是抗生素或药物耐受性或易感性基因或转录物或多肽。

[0059] 在某些实施方式中,所述方法可以还包括比较所述可检测阳性信号和(合成的)标准信号。

[0060] 在另一方面,本发明提供了一种用于检测样品中的靶核酸的方法,其包括:将样品与如本文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。所述核酸检测系统可以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(lateral flow strip)的离散检测位点处检测。所述第一分子和所述第二分子可以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的分子来检测,优选采用夹心抗体。所述侧流条可以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。

附图说明

[0061] 图1-展示了通过聚A低聚物进行的Csm6激活。

[0062] 图2-提供了三种Csm6直系同源物(ortholog)的切割偏好。EiCsm6显示具有C或A核苷酸偏好的最强活性。LsCsm6显示具有A核苷酸偏好的活性。

[0063] 图3-提供了显示被普氏菌属种MA2016 Cas13(“Cas13b5”)和L.wadei Cas13a切割的RNA的切割末端的凝胶。在没有碱性磷酸酶的情况下,只有具有5'羟基的片段可以用染料标记并显示在凝胶上,如泳道3和7所示。

[0064] 图4-提供了显示登革热靶标的Cas13b5检测通过Csm6正反馈而显著提高的图。登革热ssRNA的Cas13b5检测被补充了LsCsm6或ttCsm6以及长度为6或7个核苷酸的聚A低聚物。这些低聚物是被Cas13b5切割并以在3'端上的2,3-环状磷酸酯结束,其能够激活用于FAM-聚-A-猝灭体报告构建体的切割的Csm6。

[0065] 图5-Csx28直系同源物的蛋白质纯化。本研究中使用的PinCsx28,PauCsx28和PguCsx28的离子交换色谱法(IEC)的色谱图。显示了相对于洗脱体积(ml)的测得的UV吸光度(mAU)。红线显示用于蛋白质洗脱的渐增NaCl浓度。显示了PinCsx28,PauCsx28和PguCsx28的浓缩Csx28直系同源物的SDS PAGE。

[0066] 图6-Csx28共孵育对Cas13切割活性的影响。(A)与PinCsx28,PauCsx28,PguCsx28

共孵育或不与Csx28共孵育的PinCas13b的切割活性。(B)与PinCsx28,PauCsx28,PguCsx28共孵育或不与Csx28共孵育的PsmCas13b的切割活性。(C)与PinCsx28,PauCsx28,PguCsx28共孵育或不与Csx28共孵育的PauCas13b的切割活性。(D)与PinCsx28,PauCsx28,PguCsx28共孵育或不与Csx28共孵育的CcaCas13b的切割活性。(E)与PinCsx28,PauCsx28,PguCsx28共孵育或不与Csx28共孵育的PguCas13b的切割活性。(F)与PinCsx28,PauCsx28,PguCsx28共孵育或不与Csx28共孵育的LwaCas13a的切割活性。

[0067] 图7-通过Csm6正反馈增强的SHERLOCK信号。(A)通过PNK标记和凝胶电泳对2,3环状磷酸酯末端的检测的示意图(B)电泳凝胶,其展示了通过ssRNA靶标2(均聚物环)的LwaCas13a或PsmCas13b切割而产生的2,3环状磷酸酯末端。Cas13酶与靶向ssRNA靶标的合适crRNA一起孵育,并将切割产物用染料IR800进行5'标记(有或没有碱性磷酸酶处理)。(C)在SHERLOCK反应中Csm6介导的正反馈的示意图。(D)通过不同长度的、结束于2,3环状磷酸酯的腺嘌呤低聚物对两种Csm6直系同源物的激活。Csm6切割是使用由A,C,G或U六均聚物组成,末端标记有荧光团和淬灭体的RNA传感器测量。(E)通过具有不同长度腺嘌呤段的腺嘌呤-尿苷活化剂的LwaCas13a切割对两种Csm6直系同源物的激活。LwaCas13a靶向合成登革热sRNA。(F)来自LwaCas13a附带切割(collateral cleavage)的消化产物的质谱分析(左)或单独的2,3环状磷酸酯活化剂(右)。主峰用质量和相应的结构进行标记。(G)Csm6直系同源物的切割基序偏好发现筛选的示意图。(H)用于EiCsm6切割活性的偏好3碱基基序的热图。热图中表示的值是在所有耗尽基序(depleted motif)中的每个3碱基的计数。如果-$\log_2$(靶标/非靶标)值大于0.5,则认为基序被耗尽。在-$\log_2$(靶标/非靶标)值中,靶标和非靶标分别表示靶标条件和非靶标条件下基序的频率。(I)用于EiCsm6切割活性的偏好序列基序的序列标识。(J)检测20nM的登革热ssRNA的渐增浓度的$(A)_6$-$(U)_5$活化剂的组合的LwaCas13a和EiCsm6信号。(K)与LwaCas13a组合的铜绿假单胞菌(P.aeruoginosa)酰基转移酶合成靶标的EiCsm6增强的LwaCas13aSHERLOCK检测。(L)铜绿假单胞菌酰基转移酶合成靶标的EiCsm6增强的LwaCas13a SHERLOCK检测的动力学。(M)与LwaCas13a组合的合成登革热RNA的EiCsm6增强的侧流检测,没有通过RPA进行的预扩增。条带强度定量在右侧显示。(N)与LwaCas13a组合的铜绿假单胞菌酰基转移酶基因的EiCsm6增强的侧流SHERLOCK。条带强度定量在右侧显示。

[0068] 图8-Csm6直系同源物的蛋白质纯化。(A)本研究中使用的EiCsm6,TtCsm6,LsCsm6和SaCsm6的尺寸排阻色谱法的色谱图。显示了相对于洗脱体积(ml)的测得的UV吸收度(mAU)。(B)在尺寸排阻色谱法之前的EiCsm6,TtCsm6和LsCsm6级分的SDS-PAGE凝胶。级分显示了在streptactin孵育后的细菌裂解物上清液(1),在使用SUMO蛋白酶的切割后的streptactin树脂(2),以及释放的、未标记的Csm6蛋白(3)。(C)在尺寸排阻色谱法后的浓缩Csm6蛋白的最终SDS-PAGE。BSA标准曲线(左)用于定量Csm6蛋白(右)。显示了BSA的五种稀释和EiCsm6、TtCsm6和LsCsm6的两种稀释。

[0069] 图9-使用具有3'2,3-环状磷酸酯末端的不同长度的活化剂,TtCsm6切割的碱基偏好。

[0070] 图10-使用具有不同长度和不同3'端基团的活化剂评估EiCsm6,TtCsm6和LsCsm6活性。(A)使用长度为5-7个核苷酸且具有3'OH或3'磷酸酯末端的腺嘌呤低聚物评估EiCsm6,TtCsm6和LsCsm6切割活性。使用荧光RNase警报传感器测量切割活性。(B)分别用长

度为5nt和6nt的3'OH腺嘌呤低聚物刺激的EiCsm6和LsCsm6的碱基偏好。用于检测RNase活性的荧光均聚物传感器长度为5nt。

[0071] 图11-Csm6切割后的各种各样基序的大小分析和表示。(A)EiCsm6样品的生物分析仪迹线显示了跟随活化剂依赖性的RNase活性的文库大小的变化。(B)箱形图显示了在5分钟和60分钟时间点,Csm6、Csm6与活化剂、Csm6与活化剂和rNTP、或背景文库的靶标与非靶标基序比率的基序分布。(C)在60分钟时间点,Csm6、Csm6与活化剂、Csm6与活化剂和rNTP、或背景文库的耗尽基序的数量。

[0072] 图12-通过随机基序文库筛选确定的Csm6条件的单碱基和两碱基偏好。(A)热图显示了在60分钟时间点,Csm6、Csm6与活化剂、和Csm6与活化剂和rNTP的单碱基偏好,如通过随机基序文库切割筛选确定的。热图中表示的值是在所有耗尽基序中每个单碱基的计数。如果-$\log_2$(靶标/非靶标)值大于0.5,则认为基序被耗尽。在-$\log_2$(靶标/非靶标)值中,靶标和非靶标分别表示靶标条件和非靶标条件下基序的频率。(B)热图显示了在60分钟时间点,Csm6、Csm6与活化剂、和Csm6与活化剂和rNTP的两碱基偏好,如通过随机基序文库切割筛选确定的。热图中表示的值是在所有耗尽基序中每个两碱基的计数。如果-$\log_2$(靶标/非靶标)值大于0.5,则认为基序被耗尽。在-$\log_2$(靶标/非靶标)值中,靶标和非靶标分别表示靶标条件和非靶标条件下基序的频率。

[0073] 图13-通过随机基序文库筛选确定的Csm6条件的三碱基偏好。(A)热图显示了在60分钟时间点,Csm6的三碱基偏好,如通过随机基序文库切割筛选确定的。热图中表示的值是在所有耗尽基序中每个三碱基的计数。如果-$\log_2$(靶标/非靶标)值大于0.5,则认为基序被耗尽。在-$\log_2$(靶标/非靶标)值中,靶标和非靶标分别表示靶标条件和非靶标条件下基序的频率。(B)热图显示了在60分钟时间点,Csm6与活化剂的三碱基偏好,如通过随机基序文库切割筛选确定的。热图中表示的值是在所有耗尽基序中每个三碱基的计数。如果-$\log_2$(靶标/非靶标)值大于0.5,则认为基序被耗尽。在-$\log_2$(靶标/非靶标)值中,靶标和非靶标分别表示靶标条件和非靶标条件下基序的频率。(C)热图显示了在60分钟时间点,Csm6与活化剂和rNTP的三碱基偏好,如通过随机基序文库切割筛选确定的。热图中表示的值是在所有耗尽基序中每个三碱基的计数。如果-$\log_2$(靶标/非靶标)值大于0.5,则认为基序被耗尽。在-$\log_2$(靶标/非靶标)值中,靶标和非靶标分别表示靶标条件和非靶标条件下基序的频率。

[0074] 图14-通过随机基序文库筛选确定的Csm6条件的单碱基和两碱基偏好。(A)热图显示了在60分钟时间点,Csm6的四碱基偏好,如通过随机基序文库切割筛选确定的。热图中表示的值是在所有耗尽基序中每个四碱基的计数。如果-$\log_2$(靶标/非靶标)值大于0.5,则认为基序被耗尽。在-$\log_2$(靶标/非靶标)值中,靶标和非靶标分别表示靶标条件和非靶标条件下基序的频率。(B)热图显示了在60分钟时间点,Csm6与活化剂的四碱基偏好,如通过随机基序文库切割筛选确定的。热图中表示的值是在所有耗尽基序中每个四碱基的计数。如果-$\log_2$(靶标/非靶标)值大于0.5,则认为基序被耗尽。在-$\log_2$(靶标/非靶标)值中,靶标和非靶标分别表示靶标条件和非靶标条件下基序的频率。(C)热图显示了在60分钟时间点,Csm6与活化剂和rNTP的四碱基偏好,如通过随机基序文库切割筛选确定的。热图中表示的值是在所有耗尽基序中每个四碱基的计数。如果-$\log_2$(靶标/非靶标)值大于0.5,则认为基序被耗尽。在-$\log_2$(靶标/非靶标)值中,靶标和非靶标分别表示靶标条件和非靶标条件下

基序的频率。

[0075] 图15-用于农业应用的Csm6增强的PsmCas13b和LwaCas13a检测。在各种各样长度的聚A活化剂下的PsmCas13b检测的EiCsm6增强。

[0076] 图16-来自LwaCas13a的切割末端的质谱分析。(A)色谱图迹线显示了LwaCas13a消化的活化剂(顶部)或2,3环状磷酸酯活化剂(底部)的洗脱曲线。分析了图5中的质谱的蓝色突出显示的峰。(B)在LwaCas13a消化的活化剂(左)或2,3环状磷酸酯活化剂(右)(SEQ ID NO:137)样品中通过质谱检测到的不同化合物的丰度表。

[0077] 图17-报告体和活化剂优化对LwaCas13a活性的Csm6增强的影响。(A)用于Cas13a活性的Csm6增强的不同活化剂设计的示意图(SEQ ID NO:137、138和139)。(B)针对不同的活化剂设计,LwaCas13a检测的EiCsm6增强的性能。

[0078] 图18-报告体和活化剂浓度对LwaCas13a活性的Csm6增强的影响。(A)在不同比率的聚A和聚U报告体下,LwaCas13a检测的EiCsm6增强。(B)在各种各样浓度的$(A)_6$-$(U)_5$活化剂下,LwaCas13a检测的EiCsm6增强。

[0079] 图19-用于农业应用的Csm6增强的PsmCas13b和LwaCas13a检测。(A)采用LwaCas13a,在Roundup-Ready或WT大豆粗提取物中CP4-EPSPS除草剂抗性基因或凝集素的Csm6增强的检测。(B)采用LwaCas13a,在Roundup-Ready或WT大豆粗提取物中CP4-EPSPS除草剂抗性基因或凝集素的Csm6增强的检测的动力学。

[0080] 图20-(A)-(D)Csm6增强了来自U切割性CcaCas13b酶的信号。

[0081] 图21-体外转录组分对Csm6活性的影响。(A)在IVT组分以及有和没有2,3环状磷酸酯活化剂的存在下的EiCsm6活性。组分包括3mM的另外的$MgCl_2$、1mM的rNTP混合物,30U的T7聚合酶。(B)在各种浓度的核糖核苷酸存在下,具有$(A)_6$-$(U)_5$活化剂和聚A报告体的EiCsm6和LwaCas13a活性。(C)在各种各样浓度的核糖核苷酸存在下,采用$(A)_6$-$(U)_5$活化剂和聚A/RNaseAlert报告体组合,组合的EiCsm6和LwaCas13a活性。(D)在各种各样浓度的核糖核苷酸存在下,采用$(A)_6$-$(U)_5$活化剂和聚A/5x RNaseAlert报告体组合,组合的EiCsm6和LwaCas13a活性。(E)在各种各样浓度的核糖核苷酸存在下,采用5x$(A)_6$-$(U)_5$活化剂和聚A/RNaseAlert报告体组合,组合的EiCsm6和LwaCas13a活性。(F)在各种各样浓度的核糖核苷酸存在下,采用环状磷酸酯活化剂和聚A/RNaseAlert报告体组合,组合的EiCsm6和LwaCas13a活性。

[0082] 图22-各种各样的rNTP对Csm6活性的抑制。在存在1mM各种各样的核糖核苷酸的情况下,或在不存在各自为1mM的rNTP混合物的情况下,与2,3环状磷酸酯在一起的EiCsm6的RNase活性。

[0083] 图23-除草剂抗性基因的三步SHERLOCK检测。在存在1mM各种各样核糖核苷酸的情况下,或在不存在各自为1mM的rNTP混合物的情况下,与2,3环状磷酸酯在一起的EiCsm6的RNase活性。

[0084] 图24-具有不同报告体组合的侧流Csm6增强SHERLOCK。(A)使用各种报告体设计的Csm6增强SHERLOCK的侧流检测。sA:短聚A传感器;lA:长聚A传感器;sC:短聚C传感器;lC:长聚C传感器;sA/C:短聚A/C传感器;lA/C:长聚A/C传感器;M:混合RNase警报样传感器。(B)条带强度从(A)中的检测的定量。

[0085] 图25-使用SHERLOCK和Csm6的单分子定量和增强的信号。(A)用于铜绿假单胞菌合

成DNA的定量的DNA反应方案的示意图。(B)铜绿假单胞菌合成DNA在各种各样RPA引物浓度下的定量。值表示平均值±S.E.M.。(C)铜绿假单胞菌合成DNA浓度与检测到的荧光的相关性。值表示平均值+/-S.E.M.。(D)采用正交报告体的LwaCas13a和Csm6切割活性的独立读出(readout)的示意图。(E)通过具有不同长度腺嘌呤段的腺嘌呤-尿苷332活化剂的LwaCas13a切割对EiCsm6的激活。LwaCas13a靶向合成DENV ssRNA。值表示平均值+/-S.E.M.。(F)渐增浓度的检测20nM DENV ssRNA的$(A)_6$-$(U)_5$活化剂的组合的LwaCas13a和EiCsm6信号。值表示平均值+/-S.E.M.。(G)铜绿假单胞菌酰基转移酶合成靶标的EiCsm6增强的LwaCas13a SHERLOCK检测的动力学。

[0086] 图26-使SHERLOCK适应侧流检测。(A)使用SHERLOCK的侧流检测的示意图。(B)使用侧流SHERLOCK和1小时的LwaCas13a反应,合成ZIKV ssRNA的检测。(C)来自(B)中的检测的条带强度的定量。(D)来自患者液体活检样品的治疗相关EGFR突变的侧流检测的示意图。(E)在具有L858R或WT癌症突变的患者来源无细胞DNA样品中的EGFR L858R突变的检测。值表示平均值+/-S.E.M.。(F)在具有L858R或WT等位基因的患者来源无细胞DNA样品中的EGFR L858R突变的侧流检测。(G)来自(E)中的检测的条带强度的定量。(H)在具有外显子19缺失或WT等位基因的患者来源无细胞DNA样品中的EGFR外显子19缺失突变的检测。值表示平均值+/-S.E.M.。(I)在具有外显子19缺失或WT等位基因的患者来源无细胞DNA样品中的EGFR外显子19缺失突变的侧流检测。(J)来自(H)中的检测的条带强度的定量。(K)DENV ssRNA的EiCsm6增强的LwaCas13a检测的侧流读出的示意图。(L)与LwaCas13a组合的合成DENV RNA的EiCsm6增强的侧流检测,没有通过RPA进行的预扩增。条带强度定量在右侧显示。

[0087] 图27-Cas13a/b酶的二核苷酸偏好。(A)靶向ssRNA 1的10种Cas13a/b直系同源物的二核苷酸碱基偏好的热图,除非另有说明,报告体由A,C,G或U RNA碱基的双核苷酸组成,(*)表示具有高背景活性的未减去背景的(non-background subtracted)直系同源物。(B)对指示的靶标测试的高背景活性直系同源物LbuCas13a和PinCas13b的二核苷酸碱基偏好的热图。(C)在有和没有20nM DENV ssRNA靶标的情况下,用不同间隔子长度测试的LbuCas13a对AU二核苷酸基序的切割活性。(D)在有和没有20nM DENV ssRNA靶标的情况下,用不同间隔子长度测试的LbuCas13a对UG二核苷酸基序的切割活性。

[0088] 图28-对LwaCas13a和PsmCas13b产生的切割末端的分析。(A)通过PNK标记和凝胶电泳对2,3环状磷酸酯末端的检测的示意图。(B)电泳凝胶,其展示了通过ssRNA靶标2或3(均聚物环)的LwaCas13a或PsmCas13b切割产生的2,3环状磷酸酯末端。将Cas13酶与靶向所述ssRNA靶标的合适crRNA一起孵育,并将切割产物用染料IR800进行5'标记,有或没有碱性磷酸酶处理。

[0089] 图29-Csm6直系同源物的蛋白质纯化。(A)本研究中使用的EiCsm6,TtCsm6,LsCsm6和SaCsm6的尺寸排阻色谱法的色谱图。显示了相对于洗脱体积(ml)的测得的UV吸光度(mAU)。(B)在尺寸排阻色谱法之前,EiCsm6,TtCsm6和LsCsm6级分的SDS-PAGE凝胶。级分显示了在streptactin孵育后的细菌裂解物上清液(1),在使用SUMO蛋白酶的切割后的streptactin树脂(2),以及释放的、未标记的Csm6蛋白(3)。(C)在尺寸排阻色谱法后,浓缩Csm6蛋白的最终SDS-PAGE。BSA标准曲线(左)用于定量Csm6蛋白(右)。显示了BSA的五种稀释和EiCsm6、TtCsm6和LsCsm6的两种稀释。

[0090] 图30-Csm6直系同源物的碱基偏好和激活。(A)在SHERLOCK反应中Csm6介导的正反

馈的示意图。(B)通过不同长度的、结束于2',3'-环状磷酸酯的腺嘌呤低聚物对EiCsm6的激活。使用由A,C,G或U均聚物组成,末端标记有荧光团和淬灭体的RNA报告体测量Csm6切割。(C)通过结束于2',3'-环状磷酸酯的腺嘌呤低聚物激活的LsCsm6切割的碱基偏好。(D)通过结束于2',3'-环状磷酸酯的腺嘌呤低聚物激活的TtCsm6切割的碱基偏好。

[0091] 图31-在Csm6切割后的各种各样基序的大小分析和表示。(A)Csm6直系同源物的切割基序偏好发现筛选的示意图。(B)EiCsm6样品的生物分析仪迹线显示了跟随活化剂依赖性的RNase活性的文库大小的变化。(C)箱形图显示了在5分钟和60分钟时间点,Csm6、Csm6与活化剂、Csm6与活化剂和rNTP、或背景文库的靶标与非靶标基序比率的基序分布。(D)在60分钟时间点,Csm6、Csm6与活化剂、Csm6与活化剂和rNTP、或背景文库的耗尽基序的数量。

[0092] 图32-通过随机基序文库筛选确定的Csm6条件的单碱基和两碱基偏好。(A)用于EiCsm6切割活性的偏好序列基序的序列标识。(B)热图显示了在60分钟时间点,Csm6、Csm6与活化剂、和Csm6与活化剂和rNTP的单碱基偏好,如通过随机基序文库切割筛选确定的。热图中表示的值是在所有耗尽基序中每个单碱基的计数。如果-log2(靶标/非靶标)值大于0.5,则认为基序被耗尽。在-log2(靶标/非靶标)值中,靶标和非靶标分别表示靶标条件和非靶标条件下基序的频率。(C)热图显示在60分钟时间点,Csm6、Csm6与活化剂、和Csm6与活化剂和rNTP的两碱基偏好,如通过随机基序文库切割筛选确定的。热图中表示的值是在所有耗尽基序中每个两碱基的计数。如果-log2(靶标/非靶标)值大于0.5,则认为基序被耗尽。在-log2(靶标/非靶标)值中,靶标和非靶标分别表示靶标条件和非靶标条件下基序的频率。(D)用于EiCsm6切割活性的偏好3-碱基基序的热图。热图中表示的值是在所有耗尽基序中每个3碱基的计数。如果-log2(靶标/非靶标)值大于0.5,则认为基序被耗尽。在-log2(靶标/非靶标)值中,靶标和非靶标分别表示靶标条件和非靶标条件下基序的频率。(E)EiCsm6对源自随机基序文库筛选的排在前面的报告序列的切割活性。(F)通过具有不同长度腺嘌呤段的腺嘌呤-尿苷活化剂的LwaCas13a切割对LsCsm6的激活。LwaCas13a靶向合成DENV ssRNA。

[0093] 图33-来自LwaCas13a的切割末端的质谱分析。(A)LwaCas13a消化和验证切割产物的质谱分析的示意图。(B)来自LwaCas13a附带切割的消化产物(左)或单独的2,3环状磷酸酯活化剂(右)的质谱分析。主峰用质量和相应的结构进行标记。(C)色谱图迹线显示了LwaCas13a消化的活化剂(顶部)或2,3环状磷酸酯活化剂(底部)的洗脱曲线。(D)在LwaCas13a消化的活化剂(左)或2,3环状磷酸酯活化剂(右)样品中通过质谱检测到的不同化合物的丰度表。

[0094] 图34-报告体和活化剂优化对LwaCas13a活性的Csm6增强的影响。(A)用于Cas13a活性的Csm6增强的不同活化剂设计的示意图(B)针对不同的活化剂设计,LwaCas13a检测的EiCsm6增强的性能。

[0095] 图35-报告体和活化剂浓度对LwaCas13a活性的Csm6增强的影响。(A)在聚A和聚U报告体的各种各样比率下,LwaCas13a检测的EiCsm6增强。(B)在各种各样浓度的$(A)_6$-$(U)_5$活化剂下,LwaCas13a检测的EiCsm6增强。

[0096] 图36-体外转录组分对Csm6活性的影响。(A)在IVT组分以及有和没有2,3环状磷酸酯活化剂的存在下的EiCsm6活性。组分包括3mM的另外的$MgCl_2$、1mM的rNTP混合物,30U的T7聚合酶。(B)在各种各样浓度的核糖核苷酸存在下,采用$(A)_6$-$(U)_5$活化剂和聚A报告体,

EiCsm6和LwaCas13a活性。(C)在各种各样浓度的核糖核苷酸存在下,采用$(A)_6$-$(U)_5$活化剂和聚A/RNaseAlert报告体组合,组合的EiCsm6和LwaCas13a活性。(D)在各种各样浓度的核糖核苷酸存在下,采用$(A)_6$-$(U)_5$活化剂和聚A/5x RNaseAlert报告体组合,组合的EiCsm6和LwaCas13a活性。(E)在各种各样浓度的核糖核苷酸存在下,采用5x$(A)_6$-$(U)_5$活化剂和聚A/RNaseAlert报告体组合,组合的EiCsm6和LwaCas13a活性。(F)在各种各样浓度的核糖核苷酸存在下,采用环状磷酸酯活化剂和聚A/RNaseAlert报告体组合,组合的EiCsm6和LwaCas13a活性。

[0097] 图37-具有不同报告体组合的侧流Csm6增强SHERLOCK。(A)使用各种各样报告体设计的Csm6增强SHERLOCK的侧流检测。sA:短聚A传感器;lA:长聚A传感器;sC:短聚C传感器;lC:长聚C传感器;sA/C:短聚A/C传感器;lA/C:长聚A/C传感器;M:混合RNase警报样传感器。(B)来自(A)中的检测的条带强度的定量。(C)具有分开的RPA和IVT步骤的酰基转移酶ssDNA的EiCsm6增强的LwaCas13a SHERLOCK检测的侧流读出的示意图。(D)与LwaCas13a组合的铜绿假单胞菌酰基转移酶基因的EiCsm6增强侧流SHERLOCK。条带强度定量在右侧显示。

[0098] 图38-本研究中使用的Csm6活化剂的表格。显示的是SEQ ID NO:140(用于U切割器5A的Csm6聚A聚U探针),SEQ ID NO:141(用于U切割器6A的Csm6聚A聚U探针),SEQ ID NO:142(用于U切割器7A的Csm6聚A聚U探针),SEQ ID NO:143(5'聚U/聚A 6A探针2,3环状磷酸酯),和SEQ ID NO:144(5'聚A/聚U/聚A 6A探针2,3环状磷酸酯)。该表中的其余序列短于10个核苷酸,未分配序列标识符。

[0099] 图39-在本研究中纯化的Cas13和Csm6蛋白质的表格。

[0100] 图40-在本研究中使用的crRNA的表格。显示的是SEQ ID NO:145-519,其中SEQ ID NO:145-147分别表示用于ssRNA/ssDNA 1crRNA2的完整crRNA序列,间隔子,和直接重复序列。其余的序列标识符遵循相同的模式。

[0101] 图41-在本研究中使用的RNA和DNA靶标的表格。显示的是SEQ ID NO:520-533。

[0102] 图42-在本研究中使用的RPA引物的表格。显示的是SEQ ID NO:534-563,其中SEQ ID NO:534-536分别表示用于DENV ssRNA的正向引物序列,正向引物序列(具有T7 RNAP启动子)和反向引物序列。其余的序列标识符遵循相同的模式。

[0103] 图43-在本研究中使用的切割报告体的表格。显示的是SEQ ID NO:564(具有FAM/生物素的侧流报告体),SEQ ID NO:565(用于碱基偏好筛选的RNA基序文库),SEQ ID NO:566(金纳米颗粒接头),SEQ ID NO:567(磁性珠缀合物低聚物),SEQ ID NO:568(用于侧流的长聚A),SEQ ID NO:569(用于侧流的长聚C),和SEQ ID NO:570(用于侧流的长聚A/C)。该表格中列出的其余序列短于10个核苷酸,未分配序列标识符。

具体实施方式

[0104] 一般定义

[0105] 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:ALaboratory 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(ISBN 0632021829);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)。

[0106] 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。

[0107] 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。

[0108] 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。

[0109] 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+/-10%或更少、+/-5%或更少、+/-1%或更少以及+/-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”所指的值本身也是特定地且优选地公开的。

[0110] 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。

[0111] 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。

[0112] 概述

[0113] 所公开的实施方式提供了核酸检测组合物、系统和方法。在某些实施方式中,核酸检测组合物包含至少两种CRISPR蛋白。在某些示例性实施方式中,所述组合物和系统包含一种或多种检测蛋白和一种或多种信号放大蛋白。检测蛋白是通过与相应指导序列结合至

靶序列而被激活,所述激活最终导致可检测信号的产生。信号放大蛋白通过检测蛋白的活性而被激活,并进一步放大可检测信号。在某些示例性实施方式中,核酸检测组合物包含至少一种。

[0114] 在一个方面,本文公开的实施方式涉及核酸检测组合物。组合物包含至少一种CRISPR检测效应蛋白和至少一种CRISPR信号放大效应蛋白。在一个示例性实施方式中,CRISPR检测效应蛋白是VI型CRISPR效应蛋白。在一个示例性实施方式中,VI型CRISPR效应蛋白是Cas13a效应蛋白。在另一个示例性实施方式中,VI型CRISPR效应蛋白是Cas13b效应蛋白。在某些示例性实施方式中,CRISPR信号放大蛋白是III型CRISPR蛋白。在某些示例性实施方式中,III型CRISPR蛋白是Csm6蛋白。在某些示例性实施方式中,组合物还包含用于检测CRISPR效应子蛋白的一种或多种指导序列,其中所述指导序列设计用于结合一种或多种靶序列。在某些示例性实施方式中,组合物可以进一步包含激活信号放大CRISPR蛋白的激活序列。激活序列不同于指导序列和靶序列。在某些示例性实施方式中,激活序列是聚A核苷酸序列。聚A核苷酸序列可以具有可变的长度,这取决于所使用的CRISPR信号放大效应蛋白的类型。在某些示例性实施方式中,聚A激活序列的长度可以是5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24或25个核苷酸。在某些示例性实施方式中,组合物包含指导序列和激活序列两者。在某些示例性实施方式中,组合物可进一步包含如下文更详细描述的报告构建体。

[0115] 在某些其他方面,本文公开的实施方式涉及包含至少一种检测CRISPR效应蛋白和至少一种信号放大CRISPR检测蛋白,一种或多种指导序列,一种或多种指导激活序列,和一种或多种报告构建体的系统。上述元件各自在下文进一步详细描述。

[0116] 在某些其他方面,本文所述的组合物和系统可以被并入装置。合适的装置平台在下文进一步详细描述。

[0117] 在某些其他方面,本文公开的实施方式涉及使用本文公开的组合物、系统和装置检测靶核酸序列,和/或在某些示例性实施方式中检测蛋白质的方法。

[0118] 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)。

[0119] CRISPR检测效应蛋白

[0120] 通常,本文和文献中使用的CRISPR-Cas或CRISPR系统,例如WO2014/093622(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 asingle-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-BCRISPR-Associated RNA-Guided RNases Differentially Regulated by Accessory Proteins Csx27 and 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。

[0121] 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。

[0122] 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。

[0123] 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。

[0124] 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WO2014/093622(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.orjp/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。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。

[0125] 在一个示例性实施方式中,检测效应蛋白包含含有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。

[0126] 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(SEQ ID NO:571)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(SEQ ID NO:572)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(SEQ ID NO:573)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。

[0127] 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。

[0128] 在特定的实施方式中,CRISPR检测效应蛋白是VI型RNA靶向Cas酶。在某些示例性实施方式中,VI型RNA靶向Cas酶是Cas 13a,本文中也称为C2c2。在其他示例性实施方式中,VI型RNA靶向Cas酶是Cas 13b。在特定的实施方式中,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 NK4A179 C2c2、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 NK4A179 C2c2、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)。

[0129] 在某些其他示例性实施方式中,CRISPR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)。C2c2 HEPN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C2c2 CRISPR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-C2c2 enable guide-RNA processing and RNA detection”Nature doi:10/1038/nature19802和Abudayyeh等“C2c2 is a single-component programmable RNA-guided RNA targeting CRISPR effector”bioRxiv doi:10.1101/054742。

[0130] 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。

[0131] 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)、Nitratifractor、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。

[0132] 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tracr序列。tracr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tracr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tracr配对序列但与直接重复序列或tracr配对序列具有超过50%同一性的任何序列作为潜在的tracr序列。取潜在的tracr序列并分析与其相关的转录终止子序列。

[0133] 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)、Nitratifractor、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。

[0134] 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。

[0135] 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。

[0136] 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。

[0137] 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Cas酶的催化结构域的实例可以包括但不限于HEPN结构域。

[0138] 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。

[0139] 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)、Nitratifractor、支原体属(Mycoplasma)、弯曲杆菌属(Campylobacter)。

[0140] 在某些实施方式中,效应蛋白可以是李斯特菌属C2c2p,优选Listeria seeligeria C2c2p,更优选Listeria seeligeria serovar 1/2b str.SLCC3954 C2c2p,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。

[0141] 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。

[0142] 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。

[0143] 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;Pseudobutyrivibrio sp.OR37;Butyrivibrio sp.YAB3001;Blautia sp.Marseille-P2398;Leptotrichia sp.Marseille-P3007;Bacteroides ihuae;Porphyromonadaceae bacterium KH3CP3RA;Listeria riparia;和Insolitispirillum peregrinum。

[0144] 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。

[0145] 在某些示例性实施方式中,C2c2效应蛋白选自下表1。

[0146] 表1

[0147]

C2c2 直系同源物	编码	多字母
Leptotrichia shahii	C2-2	Lsh
L. wadei F0279 (Lw2)	C2-3	Lw2
Listeria seeligeri	C2-4	Lse
Lachnospiraceae bacterium MA2020	C2-5	LbM

[0148]

Lachnospiraceae bacterium NK4A179	C2-6	LbNK179
Clostridium aminophilum DSM 10710	C2-7	Ca
Carnobacterium gallinarum DSM 4847	C2-8	Cg
Carnobacterium gallinarum DSM 4847	C2-9	Cg2
Paludibacter propionicigenes WB4	C2-10	Pp
Listeria weihenstephanensis FSL R9-0317	C2-11	Lwei
Listeriaceae bacterium FSL M6-0635	C2-12	LbFSL
Leptotrichia wadei F0279	C2-13	Lw
Rhodobacter capsulatus SB 1003	C2-14	Rc
Rhodobacter capsulatus R121	C2-15	Rc
Rhodobacter capsulatus DE442	C2-16	Rc
Leptotrichia buccalis C-1013-b	C2-17	LbuC2c2
Herbinix hemicellulosilytics	C2-18	HheC2c2
Eubacterium rectale	C2-19	EreC2c2
Eubacteriaceae bacterium CHKC1004	C2-20	EbaC2c2
Blautia sp. Marseille-P2398	C2-21	BsmC2c2
Leptotrichia sp. oral taxon 879 str. F0557	C2-22	LspC2c2
Lachnospiraceae bacterium NK4a144		
Chloroflexus aggregans		
Demequina aurantiaca		
Thalassospira sp. TSL5-1		
Pseudobutyrivibrio sp. OR37		
Butyrivibrio sp. YAB3001		
Blautia sp. Marseille-P2398		
Leptotrichia sp. Marseille-P300		
Bacteroides ihuae		
Porphyromonadaceae bacterium KH3CP3RA		
Listeria riparia		
Insolitispirillum peregrinum		

[0149] 表2

[0150] 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。

[0151]

C2c2-2		*L. shahii* (Lsh) (SEQ ID No: 1)
c2c2-3		*L. wadei* (Lw2) (SEQ ID No: 2)
c2c2-4		*Listeria seeligeri* (SEQ ID No: 3)
c2c2-5	1	*Lachnospiraceae* bacterium MA2020 (SEQ ID No: 4)
c2c2-6	2	*Lachnospiraceae* bacterium NK4A179 (SEQ ID No: 5)
c2c2-7	3	*Clostridium aminophilum* DSM 10710 (SEQ ID No: 6)
c2c2-8	5	*Carnobacterium gallinarum* DSM 4847 (SEQ ID No: 7)
c2c2-9	6	*Carnobacterium gallinarum* DSM 4847 (SEQ ID No: 8)
c2c2-10	7	*Paludibacter propionicigenes* WB4 (SEQ ID No: 9)
c2c2-11	9	*Listeria weihenstephanensis* FSL R9-0317 (SEQ ID No: 10)
c2c2-12	10	*Listeriaceae* bacterium FSL M6-0635 = *Listeria newyorkensis* FSL M6-0635 (SEQ ID No: 11)
c2c2-13	12	*Leptotrichia wadei* F0279 (SEQ ID No: 12)
c2c2-14	15	*Rhodobacter capsulatus* SB 1003 (SEQ ID No: 13)
c2c2-15	16	*Rhodobacter capsulatus* R121 (SEQ ID No: 14)
c2c2-16	17	*Rhodobacter capsulatus* DE442 (SEQ ID No: 15)
LbuC2c2		Leptorichia buccalis C-1013-b (SEQ ID NO: 16)
HheC2c2		Herbinix hemicellulosilytica (SEQ ID NO:17)
EreC2c2		Eubacterium rectale (SEQ ID NO: 18)
EbaC2C2		Eubacteriaceae bacterium CHKCI004 (SEQ ID NO: 19)
C2c2 NK4A144		Lachnospiraceae bacterium NK4A144 (SEQ ID No: 20)
C2c2 Chloro_agg		RNA 结合蛋白 S1 Chloroflexus aggregans (SEQ ID No: 21)
C2c2 Dem_Aur		Demequina aurantiaca (SEQ ID No: 22)
C2c2 Thal_Sp_TSL5		Thalassospira sp. TSL5-1 (SEQ. I.D. No 23)
C2c2 Pseudo_sp		Pseudobutyrivibrio sp. OR37 (SEQ ID No: 24)

[0152]

C2c2_Buty_sp		Butyrivibrio sp. YAB3001 (SEQ ID No: 25)
C2c2_Blautia_sp		Blautia sp. Marseille-P2398(SEQ ID No: 26)
C2c2_Lepto_sp_Marseille		Leptotrichia sp. Marseille-P3007 (SEQ. ID No. 27)
C2c2_Bacteroides_ihuae		Bacteroides ihuae (SEQ. I.D. No 28)
C2c2_Porph_bacterium		Porphyromonadaceae bacterium KH3CP3RA(SEQ ID No: 29)
C2c2_Listeria_riparia		Listeria riparia (SEQ ID No: 30)
C2c2_insolitis_peregrinum		Insolitispirillum peregrinum (SEQ ID No: 31)

[0153] 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 NK4A179 C2c2、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。

[0154] 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。

[0155] 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:SEQ ID NO:32.

[0156] 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;1954;I958;R961;Y965;E970;R971;D972;R1046(HEPN),H1051(HEPN),Y1075;D1076;K1078;K1080;I1083;I1090。

[0157] 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-BCRISPR-associated RNA-Guided RNase differentially regulated by accessory proteins Csx27 and Csx28”Molecular Cell,65,1-13(2017);dx.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,检测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。

[0158] 表3

[0159]

Bergeyella zoohelcum	1 (SEQ ID NO: 33)
Prevotella intermedia	2 (SEQ ID NO: 34)

[0160]

Prevotella buccae	3 (SEQ ID NO: 35)
Porphyromonas gingivalis	4 (SEQ ID NO: 36)
Bacteroides pyogenes	5 (SEQ ID NO: 37)
Alistipes sp. ZOR0009	6 (SEQ ID NO: 38)
Prevotella sp. MA2016	7a (SEQ ID NO: 39)
Prevotella sp. MA2016	7b (SEQ ID NO: 40)
Riemerella anatipestifer	8 (SEQ ID NO: 41)
Prevotella aurantiaca	9 (SEQ ID NO: 42)
Prevotella saccharolytica	10 (SEQ ID NO: 43)
HMPREF9712_03108 [Myroides odoratimimus CCUG 10230]	11 (SEQ ID NO: 44)
Prevotella intermedia	12 (SEQ ID NO: 45)
Capnocytophaga canimorsus	13 (SEQ ID NO: 46)
Porphyromonas gulae	14 (SEQ ID NO: 47)
Prevotella sp. P5-125	15 (SEQ ID NO: 48)
Flavobacterium branchiophilum	16 (SEQ ID NO: 49)
Myroides odoratimimus	17 (SEQ ID NO: 50)
Flavobacterium columnare	18 (SEQ ID NO: 51)
Porphyromonas gingivalis	19 (SEQ ID NO: 52)
Porphyromonas sp. COT-052 OH4946	20 (SEQ ID NO: 53)
Prevotella intermedia	21 (SEQ ID NO: 54)
PIN17_0200 [Prevotella intermedia 17]	AFJ07523 (SEQ ID NO: 55)
Prevotella intermedia	BAU18623 (SEQ ID NO: 56)
HMPREF6485_0083 [Prevotella buccae ATCC 33574]	EFU31981 (SEQ ID NO: 57)
HMPREF9144_1146 [Prevotella pallens ATCC 700821]	EGQ18444 (SEQ ID NO: 58)
HMPREF9714_02132 [Myroides odoratimimus CCUG 12901]	EHO08761 (SEQ ID NO: 59)
HMPREF9711_00870 [Myroides odoratimimus CCUG 3837]	EKB06014 (SEQ ID NO: 60)
HMPREF9699_02005 [Bergeyella zoohelcum ATCC 43767]	EKB54193 (SEQ ID NO: 61)
HMPREF9151_01387 [Prevotella saccharolytica F0055]	EKY00089 (SEQ ID NO: 62)
A343_1752 [Porphyromonas gingivalis JCVI SC001]	EOA10535 (SEQ ID NO: 63)
HMPREF1981_03090 [Bacteroides pyogenes F0041]	ERI81700 (SEQ ID NO: 64)
HMPREF1553_02065 [Porphyromonas gingivalis F0568]	ERJ65637 (SEQ ID NO: 65)
HMPREF1988_01768 [Porphyromonas gingivalis F0185]	ERJ81987 (SEQ ID NO: 66)
HMPREF1990_01800 [Porphyromonas gingivalis W4087]	ERJ87335 (SEQ ID NO: 67)
M573_117042 [Prevotella intermedia ZT]	KJJ86756 (SEQ ID NO: 68)
A2033_10205 [Bacteroidetes bacterium GWA2_31_9]	OFX18020.1 (SEQ ID NO: 69)
SAMN05421542_0666 [Chryseobacterium jejuense]	SDI27289.1 (SEQ ID NO: 70)
SAMN05444360_11366 [Chryseobacterium carnipullorum]	SHM52812.1 (SEQ ID NO: 71)

[0161]

SAMN05421786_1011119 [Chryseobacterium ureilyticum]	SIS70481.1 (SEQ ID NO: 72)
Prevotella buccae	WP_004343581 (SEQ ID NO: 73)
Porphyromonas gingivalis	WP_005873511 (SEQ ID NO: 74)
Porphyromonas gingivalis	WP_005874195 (SEQ ID NO: 75)
Prevotella pallens	WP_006044833 (SEQ ID NO: 76)
Myroides odoratimimus	WP_006261414 (SEQ ID NO: 77)
Myroides odoratimimus	WP_006265509 (SEQ ID NO: 78)
Prevotella sp. MSX73	WP_007412163 (SEQ ID NO: 79)
Porphyromonas gingivalis	WP_012458414 (SEQ ID NO: 80)
Paludibacter propionicigenes	WP_013446107 (SEQ ID NO: 81)
Porphyromonas gingivalis	WP_013816155 (SEQ ID NO: 82)
Flavobacterium columnare	WP_014165541 (SEQ ID NO: 83)
Psychroflexus torquis	WP_015024765 (SEQ ID NO: 84)
Riemerella anatipestifer	WP_015345620 (SEQ ID NO: 85)
Prevotella pleuritidis	WP_021584635 (SEQ ID NO: 86)
Porphyromonas gingivalis	WP_021663197 (SEQ ID NO: 87)
Porphyromonas gingivalis	WP_021665475 (SEQ ID NO: 88)
Porphyromonas gingivalis	WP_021677657 (SEQ ID NO: 89)
Porphyromonas gingivalis	WP_021680012 (SEQ ID NO: 90)
Porphyromonas gingivalis	WP_023846767 (SEQ ID NO: 91)
Prevotella falsenii	WP_036884929 (SEQ ID NO: 92)
Prevotella pleuritidis	WP_036931485 (SEQ ID NO: 93)
[Porphyromonas gingivalis	WP_039417390 (SEQ ID NO: 94)
Porphyromonas gulae	WP_039418912 (SEQ ID NO: 95)
Porphyromonas gulae	WP_039419792 (SEQ ID NO: 96)

[0162]

Porphyromonas gulae	WP_039426176 (SEQ ID NO: 97)
Porphyromonas gulae	WP_039431778 (SEQ ID NO: 98)
Porphyromonas gulae	WP_039437199 (SEQ ID NO: 99)
Porphyromonas gulae	WP_039442171 (SEQ ID NO: 100)
Porphyromonas gulae	WP_039445055 (SEQ ID NO: 101)
Capnocytophaga cynodegmi	WP_041989581 (SEQ ID NO: 102)
Prevotella sp. P5-119	WP_042518169 (SEQ ID NO: 103)
Prevotella sp. P4-76	WP_044072147 (SEQ ID NO: 104)
Prevotella sp. P5-60	WP_044074780 (SEQ ID NO: 105)
Phaeodactylibacter xiamenensis	WP_044218239 (SEQ ID NO: 106)
Flavobacterium sp. 316	WP_045968377 (SEQ ID NO: 107)
Porphyromonas gulae	WP_046201018 (SEQ ID NO: 108)
WP_047431796	Chryseobacterium sp. YR477 (SEQ ID NO: 109)
Riemerella anatipestifer	WP_049354263 (SEQ ID NO: 110)
Porphyromonas gingivalis	WP_052912312 (SEQ ID NO: 111)
Porphyromonas gingivalis	WP_058019250 (SEQ ID NO: 112)
Flavobacterium columnare	WP_060381855 (SEQ ID NO: 113)
Porphyromonas gingivalis	WP_061156470 (SEQ ID NO: 114)
Porphyromonas gingivalis	WP_061156637 (SEQ ID NO: 115)
Riemerella anatipestifer	WP_061710138 (SEQ ID NO: 116)
Flavobacterium columnare	WP_063744070 (SEQ ID NO: 117)
Riemerella anatipestifer	WP_064970887 (SEQ ID NO: 118)
Sinomicrobium oceani	WP_072319476.1 (SEQ ID NO: 119)
Reichenbachiella agariperforans	WP_073124441.1 (SEQ ID NO: 120)

[0163] 指导序列

[0164] 如本文所用,术语“crRNA”或“指导RNA”或“单指导RNA”,“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的任何多核苷酸序列的多核苷酸,并指导包含

gRNA和CRISPR效应蛋白的nRNA靶向复合物与靶核酸序列的序列特异性结合。通常,gRNA可以是(i)能够与CRISPR效应蛋白形成复合物,和(ii)包含与靶多核苷酸序列具有足够互补性以与靶序列杂交并且指导CRISPR复合物与靶序列的序列特异性结合的序列的任何多核苷酸序列。如本文所用,术语“能够与CRISPR效应蛋白形成复合物”是指gRNA具有允许CRISPR效应蛋白与gRNA特异性结合使得形成能够以序列特异性方式结合靶RNA且可以对所述靶RNA发挥功能的复合物的结构。gRNA的结构组分可包括直接重复序列和指导序列(或间隔子)。与靶RNA的序列特异性结合是由与靶RNA互补的gRNA的一部分(“指导序列”)介导。在本发明的实施方式中,术语指导RNA,即能够将Cas引导至靶基因座的RNA,如在先前引用的文献如WO2014/093622(PCT/US2013/074667)中使用。如本文所用,术语“其中指导序列能够杂交”是指gRNA的子部分(subsection)与靶序列具有足够互补性以与之杂交并介导CRISPR复合物与靶RNA的结合。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复合物与靶序列的序列特异性结合的任何多核苷酸序列。在一些实施方式中,当使用合适的比对算法最佳地比对时,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,SanDiego,CA),SOAP(可从soap.genomics.org.cn获得)和Maq(可从maq.sourceforge.net获得)。

[0165] 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。

[0166] 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。

[0167] 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CRISPR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序列的宿主细胞,例如通过用编码CRISPR序列的组分的载体来转染,然后评估在靶序列内的优先切割,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RNA的切割。其他测定是可能的,并且由本领域技术人员想到。

[0168] 指导序列,并因此核酸靶向指导RNA可以被选择为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。靶序列可以包含RNA多核苷酸。术语"靶RNA"是指为或包含靶序列的RNA多核苷酸。换句话说,靶RNA可以是gRNA的一部分(即指导序列)被设计成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功能指向其的RNA多核苷酸或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。

[0169] 在一些实施方式中,选择核酸靶向指导RNA以降低RNA靶向指导RNA内的二级结构程度。在一些实施方式中,核酸靶向指导RNA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)。

[0170] 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')。

[0171] 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。

[0172] 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,优选至少18nt,如至少19、20、21、22或更多nt。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。

[0173] 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的间隔子长度为至少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。

[0174] 在经典CRISPR-Cas系统中,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。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,在实施例中,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。

[0175] 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。

[0176] 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,并且可检测信号将被产生。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,并且没有可检测信号将被产生。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。

[0177] 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。

[0178] 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。

[0179] 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)。

[0180] 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)。

[0181] 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。

[0182] 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。

[0183] 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。

[0184] 信号放大CRISPR效应蛋白

[0185] 在某些示例性实施方式中，信号放大CRISPR效应蛋白是III-A型CRISPR-Cas系统效应蛋白。在某些示例性实施方式中，III-A型CRISPR-Cas效应蛋白是Csm6。Csm6与多蛋白Csm效应子复合体一起发挥功能，但不是该复合体的一部分(参见，例如，US20170198286A1；

WO2016035044A1;M.Kazlauskiene等,Science10.1126/science.aao0100(2017);和Niewoehner等,2017,2017年6月23日首次在线公布的bioRxiv预印本;doi:dx.doi.org/10.1101/153262)。

[0186] 在表皮葡萄球菌(Staphylococcus epidermidis)中,Csm复合物(SeCsm)包含Cas10,Csm2,Csm3,Csm4,和Csm5蛋白质。使用嗜热链球菌(Streptococcus thermophilus)(St)的Csm复合物,证明了III-A型CRISPR-Cas系统在体外和在细胞中均具有RNA切割活性(参见,例如,US20170198286A1)。

[0187] III-A型CRISPR-Cas系统包括嗜热链球菌(GenBank KM222358),DGCC7710(GenBank AWVZ01000003),LMD-9(GenBank NC008532),表皮葡萄球菌RP62a(GenBank NC002976),意大利肠球菌DSM15952(GenBank AEPV01000074),乳酸乳球菌DGCC7167(GenBank JX524189)和硫磺矿硫化叶菌(Sulfolobus solfataricus)P2(GenBank AE006641)。DGCC8004的III-A型系统含有在CRISPR2阵列侧翼的10个cas基因,包括cast,cas2,cas6,cas10,csm2,csm3,csm4,csm5,csm6和csm6'基因。DGCC8004CRISPR2基因座与DGCC7710(GenBank AWVZ00000000,(Horvath和Barrangou,2010))和LMD-9(GenBank NC_008532,(Makarova等,2006))的基因座分享相似的基因排列。主要区别在于DGCC8004中的另外的csm6'基因。DGCC8004中的Csm6'蛋白包含386个氨基酸,与428个氨基酸的Csm6蛋白具有34%的氨基酸同一性,启示可能的古老基因复制事件,之后是序列趋异。相比之下,DGCC7710在csm6前面仅含有短117-nt ORF。DGCC8004中与CRISPR2相关的Cas/Csm蛋白与DGCC7710和LMD-9中的相应蛋白同源(氨基酸同一性超过90%,除了对于Csm2蛋白,其分享约70%的同一性)。其他实验表征的III-A型系统包括表皮葡萄球菌RP62a(GenBank NC002976,(Marraffini和Sontheimer,2008)),意大利肠球菌DSM15952(GenBank AEPV01000074,(Millen等,2012))和乳酸乳球菌DGCC7167(GenBank JX524189,(Millen等,2012))与DGCC8004分享cas10-csm2-csm3-csm4-csm5-csm6基因簇的保守排列,而cas6和cast/cas2基因的位置在某些菌株中有所不同。硫磺矿硫化叶菌P2(GenBank AE006641)中的III-A型CRISPR-Cas基因座具有不同的基因组织(gene organization),显示与DGCC8004中的Cas/Csm直系同源物的低蛋白质序列相似性。值得注意的是,Csm3蛋白在不同菌株中在Cas/Csm蛋白中最为保守,并且5个拷贝的Csm3旁系同源物(paralogues)存在于硫磺矿硫化叶菌中。表皮葡萄球菌、意大利肠球菌和乳酸乳球菌中的重复序列具有相同长度(36nt),然而核苷酸保守性限于重复序列的回文部分和3'-端。重复序列的8-nt 3'-端序列(其可以有助于crRNA 5'-把手)显示了嗜热链球菌,表皮葡萄球菌,意大利肠球菌和乳酸乳球菌之间的ACGRRAAC共有性,但与硫磺矿硫化叶菌的不同(AUUGAAG(Rouillon等,2013))。

[0188] Csm6已经显示是ssRNA特异性内切核糖核酸酶,并且确定了该活性的结构基础(Niewoehner和Jinek,2016,Structural basis for the endoribonuclease activity of the type III-A CRISPR-associated protein Csm6.RNA 22:318-329)。

[0189] 在一些实施方式中,本发明的核酸靶向系统的一个或多个元件衍生自包含内源性CRISPR RNA靶向系统的特定生物。在某些实施方式中,CRISPR RNA靶向系统包含Csm6蛋白,Csm6直系同源物或Csm6样蛋白。如本文所用,对Csm6的讨论也涉及Csm6蛋白,Csm6直系同源物或Csm6样蛋白。Csm6直系同源物可见于本文所述和本领域已知的生物(参见,例如,WO2016035044A1和Niewoehner和Jinek,2016)。示例性Csm6直系同源物包括但不限于嗜热

栖热菌(T.thermophilus)(TtCsm6,GI:55978335),表皮葡萄球菌(SeCsm6,GI:488416649),变形链球菌(SmCsm6,GI:24379650),嗜热链球菌(StCsm6,GI:585230687)和激烈火球菌(P.furiosus)Csx1(PfCsx1,GI:33359545)。在某些实施方式中,可用于本发明的Csm6蛋白包含至少一个N末端CARF(CRISPR相关的罗斯曼(Rossman)折叠)结构域和至少一个C末端HEPN结构域(高级真核生物和原核生物核苷酸结合结构域)。在某些实施方式中,Csm6蛋白形成二聚体。在某些实施方式中,HEPN结构域的二聚化导致核糖核酸酶活性位点的形成。在某些实施方式中,CARF结构域的二聚体界面包括电正性口袋。不受理论的束缚,该口袋可以起到用于核糖核酸酶活性的变构控制的配体结合位点的功能。

[0190] 在某些示例性实施方式中,本文所述的基于CRISPR的检测系统包含Csm6蛋白,其包含至少一个HEPN结构域,包括但不限于本文所述的HEPN结构域,本领域已知的HEPN结构域(Niewoehner和Jinek,2016),以及通过与共有序列基序相比较而被认为是HEPN结构域的结构域。本文提供了多种这样的结构域。在一个非限制性实例中,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,Csm6蛋白包含单个HEPN结构域。在某些其他示例性实施方式中,Csm6蛋白包含两个HEPN结构域。

[0191] 在一个示例性实施方式中,Csm6蛋白包含一个或多个包含RxxxxH基序序列的HEPN结构域。RxxxxH基序序列可以但不限于来自本文描述的HEPN结构域或本领域已知的HEPN结构域。RxxxxH基序序列还包括通过结合两个或更多个HEPN结构域的部分而产生的基序序列。如所指出的,共有序列可以源自本文公开的直系同源物的序列。在某些实施方式中,HEPN结构域包含保守的R-X4-6-H基序(Anantharaman等,Biol Direct..,2013年6月15日;8:15;和Kim等,Proteins,2013年2月;81(2):261-70)。

[0192] 在本发明的一个实施方式中,HEPN结构域包含至少一个RxxxxH基序,其包含R{N/H/K}X1X2X3H的序列(SEQ ID NO:571)。在本发明的一个实施方式中,HEPN结构域包含RxxxxH基序,其包含R{N/H}X1X2X3H的序列(SEQ ID NO:572)。在本发明的一个实施方式中,HEPN结构域包含R{N/K}X1X2X3H的序列(SEQ ID NO:573)。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。

[0193] 已经描述了CARF结构域和CARF结构域的共有序列(参见,例如,Makarova等,Front Genet.,2014;5:102)。在某些实施方式中,Csm6包含至少一个CARF结构域,其包含核心结构域,该核心结构域包含六链Rossmann样折叠,且核心链5和链6形成β-发夹。序列保守性的主要区域与核心结构域的链1和链4相关联。在某些实施方式中,链1的末端的特征在于极性残基,通常具有醇性侧链(S/T)。在某些实施方式中,在链4的紧靠的下游是高度保守的碱性残基(K/R),其优选与[DN]X[ST]XXX[RK](SEQ ID NO:574)标志相关联。在某些实施方式中,将Csm6截短以去除N-端CARF结构域(例如,TtCsm6的氨基酸1-190或直系同源Csm6蛋白中的等价残基)。

[0194] 在某些实施方式中,Csm6包含至少一个6H结构域(Niewoehner和Jinek,2016)。TtCsm6多肽链的6H结构域(残基191-292)由五个α螺旋组成,形成右手螺线管结构域。不受理论的束缚,因为某些直系同源物可以不具有6H结构域,所以该结构域对于本发明的Csm6蛋白的活性不是需要的。

[0195] Csm6已经显示通过充当降解入侵者RNA转录物的独立核糖核酸酶而对干扰作出贡

献。Csm6蛋白被III型干扰复合物产生的第二信使激活。在靶RNA被III型干扰复合物结合后,Cas10亚基将ATP转化为环状寡聚腺苷酸产物,该产物通过与其CARF结构域结合而变构地激活Csm6。消除变构激活的CARF结构域突变体内抑制Csm6活性,以及Csm6的Cas10 Palm结构域拟表型丧失中的突变(M.Kazlauskiene等,2017;和Niewoehner等,2017)。

[0196] 在某些示例性实施方式中,当激活的CRISPR检测蛋白切割激活序列时,信号放大CRISPR效应蛋白被激活。激活序列在下文进一步详细描述。激活序列的切割产物激活信号放大CRISPR效应蛋白的单独活性,如RNA核酸酶活性。例如,Csm6一旦被激活,则类似于Cas13酶的附带作用不加选择地切割RNA。因此,除了报告构建体的检测效应子修饰之外,被激活的信号放大CRISPR效应蛋白也修饰报告构建体以进一步增强信号产生。在某些实施方式中,当与另外的CRISPR酶(例如,Cas13)结合提供时,Csm6被激活。在某些实施方式中,Csm6当与Cas13结合使用时可以产生协同作用,使得Cas13附带活性被大大提高。不受理论约束,当Csm6也包括在测定中(例如,护理点测定)时,Cas13的浓度可以在测定中大大减少。因此,Csm6添加到Cas13诊断测定法可以用于增加测定法的灵敏度并降低成本。

[0197] 在某些示例性实施方式中,一种或多种信号放大效应蛋白选自表4。

[0198] 表4

[0199]

EiCsm6（SEQ ID NO: 121）	WP_007208953.1	意大利肠球菌
TtCsm6（SEQ ID NO: 122）	WP_011229148.1	嗜热栖热菌
StCsm6（SEQ ID NO: 123）	WP_000865879.1	葡萄球菌
ShCsm6（SEQ ID NO: 124）	WP_085050120.1	溶血性葡萄球菌

[0200]

PtCsm6（SEQ ID NO: 125）	WP_078807318.1	*Pilibacter termitis*
SaCsm6（SEQ ID NO: 126）	EHO90787.1	金黄色葡萄球菌亚种 金黄色葡萄球菌 21252
ThCsm6（SEQ ID NO: 127）	WP_094243908.1	嗜盐四联球菌（*Tetragenococcus halophilus*）
FsCsm6（SEQ ID NO: 128）	WP_069876671.1	梭菌属种（Fusibacter sp.） 3D3
LaCsm6（SEQ ID NO: 129）	WP_056988115.1	*Lactobacillus acidipiscis*
LsCsm6（SEQ ID NO: 130）	WP_081509150.1	唾液乳酸菌（*Lactobacillus salvarius*）

[0201] 激活序列

[0202] 激活序列与CRISPR信号放大效应蛋白结合使用。在某些示例性实施方式中,一旦CRISPR检测效应蛋白被激活,则激活序列被CRISPR检测效应蛋白切割。激活序列的切割产生激活CRISPR信号放大效应蛋白的切割片段。如上所述,在某些示例性实施方式中,CRISPR信号放大效应蛋白的激活导致可检测信号的协同产生。在某些示例性实施方式中,激活序列是均聚物寡核苷酸。在某些示例性实施方式中,均聚物寡核苷酸的切割产生至少一种具有3'2,3环状磷酸酯的切割产物,其刺激CRISPR信号放大效应蛋白的激活。在某些示例性实施方式中,均聚物是聚A寡核苷酸。在某些示例性实施方式中,激活序列的长度为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个核苷酸。在某些示例性实施方式中,激活序列的长度为5至10个核苷酸,长度为5至15个核苷酸,长度为5至20个核苷酸,长度为5至25个核苷酸,长度为5至30个核苷酸,长度为10至15个核苷酸,长度为10至20个核苷酸,长度为10至25个核苷酸,长度为10至30个核苷酸,长度为15至20个核苷酸,长度为15至25个核苷酸,长度为15至30个核苷酸,长度为20至25个核苷酸,长度为20至30个核苷酸,或长度为25至30个核苷酸。

[0203] 示例性激活序列在表5中显示。

[0204] 表5

[0205]

rArArArArA	5A-3OH
rArArArArA/3Phos/	5A-3P
rArArArArArA	6A-3OH

rArArArArArA/3Phos/	6A-3P
rArArArArArArA	7A-3OH
rArArArArArArA/3Phos/	7A-3P

[0206] 报告构建体

[0207] 如本文所用,“报告构建体(reporter construct)”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“报告构建体”替代地也可称为“检测构建体”或“掩蔽构建体”。在某些示例性实施方式中,报告构建体是基于RNA的报告构建体。报告构建体配置成使得阳性可检测信号的产生或检测直到CRISPR效应系统被激活才实现。阳性可检测信号可以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。报告构建体可以阻止可检测阳性信号的产生或掩蔽可检测阳性信号的存在直到报告构建体被CRISPR效应蛋白活性修饰。术语“阳性可检测信号”用于与在报告构建体存在下可以是可检测的其他可检测信号相区别。例如,在某些实施方式中,当存在未修饰的报告构建体时可以检测到第一信号(即阴性可检测信号),其然后在激活的CRISPR效应蛋白修饰报告构建体后转化为第二信号(例如阳性可检测信号)。

[0208] 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shRHN或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或抗体检测到的荧光蛋白或其他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。

[0209] 在某些示例性实施方式中,掩蔽构建体可以螯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所述一种或多种试剂的RNA适体螯合。当检测到靶分子后效应蛋白被激活时,释放所述一种或多种试剂。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,其被抑制或螯合使得蛋白不可以通过一种或多种RNA适体与所述蛋白的结合,产生可检测信号。在激活本文公开的效应蛋白后,RNA适体被切割或降解至它们不再抑制所述蛋白产生可检测信号的能力的程度。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的序列(SEQ ID NO:131)。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7-氨基-4-甲基香豆素,其是可以使用荧光检测器检测的蓝色荧光团。抑制性适体也可以在上述一般原理内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)。

[0210] 在某些实施方式中,RNA酶活性用比色法通过酶抑制性适体的切割来检测。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相结合。在不存在RNA切割的情况下,完整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:一旦通过附带活性(例如Cas13a附带活性)从适体释放,比色酶将继续产生比色产物,导致信号的增加。

[0211] 在某些实施方式中,使用抑制具有比色读出的酶的现有适体。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,并且是可商购的。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,可以被通过选择策略(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。

[0212] 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来检测。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,和与抑制剂和酶两者共价连接、将抑制剂系连在酶上的桥接RNA。在未切割的构型中,酶被增加的局部浓度的小分子抑制;当RNA被切割时(例如通过Cas13a附带切割),抑制剂将被释放并且比色酶将被激活。

[0213] 在某些实施方式中,RNA酶活性通过比色法由G-四链体的形成和/或激活来检测。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(DNAzyme)。当被供应了过氧化物酶底物(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,其然后在溶液中形成绿色。G四链体形成性DNA序列的实例是:GGGTAGGGCGGGTTGGGA(SEQ ID NO:132)。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酶附带激活(例如C2c2-复合物附带激活)后,RNA短链将被切割,从而允许G四链体形成并且血红素结合。该策略特别有吸引力,因为颜色形成是酶促的,这意味着在RNA酶激活之外还存在另外的扩增。

[0214] 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。

[0215] 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。

[0216] 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分子失活核酶后,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被检测到。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下文中工作的实例。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。

[0217] 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。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。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、钌、锌、铁、镍和钙。其他合适的金属还包括以下它们的各种氧化态:锂、钠、镁、钾、钪、钛、钒、铬、锰、钴、铜、镓、锶、铌、钼、钯、铟、锡、钨、铼、铂和钆。金属优选以离子形式提供,衍生自适当的金属化合物,例如A13+、Ru3+、Zn2+、Fe3+、Ni2+和Ca2+离子。

[0218] 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、静态淬火或接触淬火。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。

[0219] 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。

[0220] 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。

[0221] 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/5Biosg/UCUCGUACGUUC/3IAbRQSp/(SEQ ID NO.133)或/5Biosg/UCUCGUACGUUCUCUCGUACGUUC/3IAbRQSp/(SEQ ID NO.134)募集猝灭分子,其中/5Biosg/是生物素标签并且/3lAbRQSp/是爱荷华黑淬灭剂。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。

[0222] 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射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)。

[0223] 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。

[0224] 扩增

[0225] 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)。

[0226] 在某些示例性实施方式中,基于RNA或DNA扩增核酸序列的扩增是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℃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。

[0227] 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℃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。

[0228] 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。

[0229] 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。

[0230] 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。

[0231] 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。

[0232] 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,使得可以同时进行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。

[0233] 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。

[0234] 基于指导释放和靶标生成的信号放大

[0235] 在某些示例性实施方式中,本文公开的组合物、系统和方法可以进一步包括添加第二靶标。第二靶标不同于主要靶标,并且可以在测定中是通用的。第二靶标可以高浓度添加到每个测定中。对应于第二靶标的指导序列被包含在每个测定体积中。次级靶标指导序列会被保护,使得其不能结合次级靶标或检测CRISPR效应蛋白,例如Cas13蛋白。保护性基团或结构配置成使得其可以在检测CRISPR效应蛋白的激活后被切割。一旦保护性基团或结构被去除,释放的指导序列能够与溶液中的检测CRISPR效应蛋白形成复合物,并触发检测CRISPR效应蛋白的进一步激活,由此通过报告构建体的修饰而导致可检测信号的进一步产生。

[0236] 或者,可以将第二指导序列与受保护的靶标一起添加到每个测定体积中。通过激活的检测CRISPR效应蛋白将保护者基团或结构从靶标上切割下来,会允许另外的CRISPR效应蛋白/指导序列/第二靶序列复合物形成,从而进一步增加附带作用。

[0237] 在某些示例性实施方式中,保护性基团或结构可以是被附带活性切割掉的阻断性第二结构环。

[0238] 在另一方面,被激活的附带作用可以切割受保护的或环化的引物,其会被释放以在模板上针对指导序列,靶序列或两者进行扩增反应。在某些示例性实施方式中,扩增反应是等温扩增反应,例如重组酶聚合酶扩增或滚环扩增。扩增的模板的后续转录会产生更多的指导序列和/或靶标,从而允许另外的检测CRISPR效应蛋白激活。

[0239] 靶RNA/DNA富集

[0240] 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。

[0241] 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。在某些示例性实施方式中,富集可以在20-37℃之间发生。在某些示例性实施方式中,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。

[0242] 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。

[0243] 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),WO 04/0918497,U.S.7,057,026;WO 91/06678;WO 07/123744;US 7,329,492;US 7,211,414;US 7,315,019;U.S.7,405,281,和US 2008/0108082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,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)。

[0244] 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。

[0245] 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。

[0246] 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。

[0247] 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,一个或多个修饰的或突变的氨基酸残基是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。

[0248] 在特定的实施方式中,一个或多个修饰的或突变的氨基酸残基是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,一个或多个修饰的或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,一个或多个修饰的或突变的氨基酸残基是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。

[0249] 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。

[0250] 蛋白质的检测

[0251] 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。在结合靶标之前,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结合靶标后,适体的结构经历构象变化,使得RNA聚合酶启动子被暴露。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。

[0252] 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,和在足以允许肽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使RNA聚合酶启动子结合位点暴露,导致通过RNA聚合酶与RNA聚合酶启动子结合位点的结合,合成触发RNA。

[0253] 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。

[0254] 装置

[0255] 本文所述的系统可以在诊断装置上具体化。可以使用能够在装置内限定多个个体离散体积的装置的许多基质和配置。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任意限定的体积或空间,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,或者由可以含有靶分子和可索引核酸标识符(例如核酸条形码)的其他手段例如化学、扩散速率限制、电磁或光照射或其任何组合限定的。“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种类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)、具有由以有序阵列或随机图案的沉积试剂定义的区域的显微镜载玻片、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。

[0256] 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,本文的系统和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):950-54Pardee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WO/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”Cell134(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。

[0257] 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微生物有关的信息,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。

[0258] 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。在一个示例性实施方式中,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。

[0259] 在其他示例性实施方式中,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。

[0260] 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.2013Jun;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。

[0261] 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SHERLOCK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。

[0262] 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。

[0263] 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出和/或允许在单个实验中对数百万个细胞的灵敏和定量的测量,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。在凝胶单体聚合时以在液滴内形成珠子。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。可以通过偶联杂交链反应(HCR引发剂)扩增的检测构建体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。

[0264] 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)。

[0265] 本文描述的系统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。

[0266] 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。

[0267] 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。

[0268] 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。

[0269] 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标签的融合蛋白。亲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)。

[0270] 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)。

[0271] 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。

[0272] 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。

[0273] 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。

[0274] 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。

[0275] 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。

[0276] 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。

[0277] 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,Diagnostics 2014,4(3),104-128;mReader from Mobile Assay;and Holomic Rapid Diagnostic Test Reader)。

[0278] 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时,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。

[0279] 实例方法和应用

[0280] 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。

[0281] 在某些示例性实施方式中,将特异于单个靶标的单个指导RNA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RNA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,而其他优先切割C、U或T。因此可以产生都是单一核苷酸或包含大部分的单一核苷酸的指导RNA,各自具有不同荧光团。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靶标。

[0282] 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。参见例如下文描述的实施例。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。

[0283] 微生物应用

[0284] 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作指导治疗方案,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。

[0285] 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。

[0286] 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。

[0287] 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。

[0288] 微生物检测

[0289] 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RNA,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WO2015/085194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。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。

[0290] 检测单核苷酸变体

[0291] 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。

[0292] 基于rRNA序列的检测

[0293] 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,扩增引物组可以设计用于侧接核糖体RNA序列的恒定区,并且指导RNA可以设计用于通过可变内部区域区分各物种。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。

[0294] 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。

[0295] 筛选药物抗性

[0296] 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。

[0297] 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.standford.edu。本文所公开的实施方式可用于检测这样的变体等。

[0298] 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),以及埃博拉的感染性

增加(Diehl等Cell.2016,167(4):1088-1098。

[0299] 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)。

[0300] 组覆盖方法

[0301] 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。这样的方法描述于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WO2017/040316中公开的方法进行比较,所述方法通过引用并入本文。如WO2017040316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):e3 doi10.1093/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。

[0302] 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。

[0303] 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。

[0304] 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)。

[0305] 监测微生物爆发

[0306] 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。

[0307] 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。

[0308] 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)。

[0309] 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)。

[0310] 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)。

[0311] 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)。

[0312] 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)。

[0313] 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)。

[0314] 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的机制,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。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。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 Leone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)。

[0315] 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)。

[0316] 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。

[0317] 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。

[0318] 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。

[0319] 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。

[0320] 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。

[0321] 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,可以同时在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。

[0322] 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。

[0323] 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。

[0324] 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。

[0325] 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。

[0326] 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。

[0327] 筛选微生物遗传扰动

[0328] 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。

[0329] 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。

[0330] 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。

[0331] 筛选环境样品

[0332] 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的,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。

[0333] 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。

[0334] 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。

[0335] 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。

[0336] 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。

[0337] 样品类型

[0338] 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,生物样品可以是从例如血液、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。

[0339] 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)。

[0340] 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。

[0341] 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 Hou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WO2011109762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。

[0342] 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 Hou等,Pathogen Isolation from Whole Blood Using Spiral Microchannel,CaseNo.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。

[0343] 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。

[0344] 示例性微生物

[0345] 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。

[0346] 细菌

[0347] 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 baumanii、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 baumanii、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(Corynbacterium)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 chafeensia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(Morganell)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、Mycobacteriumintracellulare、Mycobacterium avium、Mycobacterium bovis、和Mycobacterium marinum)、支原体属(Mycoplasm)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 cholerasuis和Salmonella typhimurium)、沙雷氏菌属(Serratia)种(例如Serratia marcesans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 moniliformi、密螺旋体属(Treponema)种(例如Treponema carateum、Treponema petenue、Treponema pallidum和Treponema endemicum、Trichophyton rubrum、T.mentagrophytes、Tropheryma whippelii、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。

[0348] 真菌

[0349] 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(Coccidiodomycosis)、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。

[0350] 原生动物

[0351] 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 madrillaris、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。

[0352] 寄生虫

[0353] 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)。

[0354] 病毒

[0355] 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

mmarenavirus、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、Kyasanur 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 leukoenchalitis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、Newcastledisease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 Zviral 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(Newcastledisease 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)

[0356] 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(SPSVV)、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。

[0357] 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)。

[0358] 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 Zozter virus)、Malocoherpesviridae、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、屎肠球菌、奇异变形杆菌、无乳链球菌、或嗜麦芽糖葡萄球菌或其组合。

[0359] 疟疾检测和监测

[0360] 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。

[0361] 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。

[0362] 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。

[0363] 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。

[0364] 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))。

[0365] 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。

[0366] 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:

[0367] • 到2030年将疟疾病例发病率降低至少90%。

[0368] • 到2030年将疟疾死亡率降低至少90%。

[0369] • 到2030年消除至少35个国家的疟疾。

[0370] • 防止所有无疟疾国家的疟疾再次出现。

[0371] 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:

[0372] • 确保普遍获得疟疾预防、诊断和治疗;

[0373] • 加快努力消除和实现无疟疾地位;和

[0374] • 将疟疾监测转变为核心干预措施。

[0375] 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、本芴醇、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(atemisins)药

物;及其组合。

[0376] 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。

[0377] 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。

[0378] 另外的靶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。

[0379] 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(pfcrt)、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(pfnhe)、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(pfdhps)、二氢叶酸还原酶活性(pfdhpr)和二氢叶酸还原酶胸苷酸合酶(pfdhfr)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异源基因。

[0380] 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。

[0381] 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。

[0382] 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)。

[0383] 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-氨苯砜联合疗法抗性相关。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。

[0384] 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。

[0385] 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。

[0386] 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芴醇、青蒿素、奎宁、氟喹、卤代蒽醌和氯喹的灵敏度。此外,PfMDR1的扩增与对苯芴醇、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。

[0387] 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。

[0388] 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。

[0389] 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、M133I和G280D可能与阿托伐醌抗性相关。

[0390] 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。

[0391] 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。

[0392] 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 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中描述的;其内容通过引用并入本文。

[0393] 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。

[0394] 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。

[0395] 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。

[0396] 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。

[0397] 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。

[0398] 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。

[0399] 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(pfcrt)、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(pfnhe)、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异源基因。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。

[0400] 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。

[0401] 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。

[0402] 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)。

[0403] 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。

[0404] 生物标志物的检测

[0405] 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)。

[0406] 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:

[0407] 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;

[0408] 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;

[0409] 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

[0410] 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。

[0411] 生物标志物样品类型

[0412] 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。

[0413] 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。

[0414] 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。

[0415] 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。

[0416] 循环肿瘤细胞

[0417] 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 US 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)。

[0418] 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。

[0419] 无细胞染色质

[0420] 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 jOncology 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在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。

[0421] 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。

[0422] 无细胞DNA(cfDNA)

[0423] 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。

[0424] 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。

[0425] 外来体

[0426] 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WO2016172598A1)。

[0427] SNP检测和基因分型

[0428] 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。

[0429] 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)。

[0430] 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:

[0431] 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;

[0432] 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;

[0433] 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

[0434] 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。

[0435] 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。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。

[0436] 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.2013Mar;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。

[0437] 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。然后可以将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。

[0438] 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。

[0439] 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。

[0440] LOH检测

[0441] 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。

[0442] “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。

[0443] 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 ID No:135)是二核苷酸微卫星,GTCGTCGTCGTCGTC(SEQ ID No:136)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。

[0444] 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.Science264,436-440(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,Science271,350-353(1996);和Miozzo等,CancerRes.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)。

[0445] 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,本发明可用于以与任何在先的方法相比具有改善的灵敏度检测LOH标志物,从而提供突变事件的早期检测。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。

[0446] 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。

[0447] 检测表观遗传修饰

[0448] 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。

[0449] 产前筛查

[0450] 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。

[0451] 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将流产,尽管医生和患者也发现早期诊断高风险妊娠是有用的,使得分娩可以安排在宝宝可以接受适当护理的三级护理医院。

[0452] 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涉及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。

[0453] 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。

[0454] 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℃。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。

[0455] 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。

[0456] 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)。

[0457] 癌症和癌症药物抗性检测

[0458] 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.2016May 20;7:11589;Landau DA,等,Mutations driving CLL and their evolution in progression and relapse.Nature.2015Oct22;526(7574):

525-30;Landau DA,等,Clonal evolution in hematological malignancies and therapeutic implications.Leukemia.2014Jan;28(1):34-43;和Landau DA,等,Evolution and impact of subclonal mutations in chronic lymphocytic leukemia.Cell.2013Feb14;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。

[0459] 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.2015January 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WO2016100975A1)。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、IDO1、IDO2、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。

[0460] 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。

[0461] 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。

[0462] 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。

[0463] 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。

[0464] 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。

[0465] 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。

[0466] 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 oligodendroglioma.Nature.2016Nov 10;539(7628):309-313.doi:10.1038/nature20123.Epub 2016Nov 2;和国际专利公开序列号WO 2017004153 A1)。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。

[0467] 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。

[0468] 免疫疗法应用

[0469] 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。

[0470] 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.2016Sep8;166(6):1500-1511.e9.doi:10.1016/j.cell.2016.08.052)。

[0471] 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。

[0472] 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。

[0473] 在一些实施方式中,本文所公开的系统和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。

[0474] 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。

[0475] 检测核酸标记的物品

[0476] 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WO/2014/047561中。

[0477] 酶

[0478] 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。

[0479] 实施例

[0480] 实施例1:用另外的CRISPR相关蛋白提高CAS13活性

[0481] CRISPR效应子经常与另外的组分相互作用以调节活性,申请人寻求利用这些相互作用以提高SHERLOCK的灵敏度和速度。VI-B型CRISPR系统通常具有干扰调节性蛋白Csx27和Csx28,且Csx28共表达已被证明体内提高Cas13b蛋白的干扰活性,意味着它们可以能够体外增加Cas13b的核酸内切酶活性。由于Csx28在我们手中不稳定,因此申请人从三种VI-B型系统纯化了Csx28-Sumo融合蛋白(图5),并测试了Csx28补充是否增加了Cas13a和Cas13b蛋白的活性。申请人发现,Csx28蛋白或者降低了Cas13的活性,或者增加了靶标非依赖性(target-independent)切割(图6A-F)。

[0482] 实施例2:Csm6切割活性的表征

[0483] 被证明CRISPR III型效应核酸酶Csm6的基于核酸的变构激活的近期研究激起了兴趣,申请人想知道Cas13的内切核酸活性是否能够产生有效的Csm6活化剂,其包含具有2,3环状磷酸酯末端的线性RNA腺苷酸(图7C,7D)。申请人因此寻求通过对含有A或U均聚物环的合成ssRNA 2进行Cas13体外切割测定,表征Cas13衍生的切割产物的RNA末端化学性质(图7A)。体外切割反应的切割后荧光标记证实了LwCas13a和PsmCas13b产生了具有5'羟基化和2'3'环状磷酸酯RNA末端的切割产物(图7B)。该结果引导我们表达和纯化已知的和新颖的Csm6直系同源物两者,用于探索Csm6与Cas13一起用于正反馈信号放大的用途(图7C和8)。通过测试具有不同长度和3'末端修饰的RNA腺苷酸,申请人发现EiCsm6和LsCsm6被含有2'3'环状磷酸酯末端的六腺苷酸有效地激活(图7D和10)。而且申请人发现,Csm6直系同源物具有对A和C均聚物RNA传感器的强切割偏好(图7d)。为了获得对于Csm6的切割特异性的更全面的洞见,申请人对用于Cas13的相同6聚简并RNA报告体文库进行了体外切割测定和

RNA测序。出人意料地,耗尽序列基序分析显示了对鸟苷的强偏好(图7F-7G和12-14),这是申请人由于均聚物G-RNA传感器的低效率切割而没有预料到的结果。然而,鉴于对均聚物A-RNA传感器的强切割活性,申请人将这种设计用于未来的Csm6实验,这也由于它们不同的切割偏好而允许我们独立地测量LwCas13a和EiCsm6。

[0484] 实施例3:采用CRISPR-Csm6的正反馈信号放大

[0485] 为了使Cas13的活性与Csm6激活相结合,申请人设计了会在Cas13切割后产生最佳Csm6刺激的RNA活化剂。申请人将PsmCas13b与较长的聚A活化剂一起孵育,所述聚A活化剂的长度会因切割而缩短,以生成短的、结束于环状磷酸酯的活化剂。活化剂的PsmCas13b消化导致LsCsm6活性的适度增加(图14A),可能是由于从切割产生的一系列子活化剂(sub-activator)。

[0486] 申请人通过设计含有两种碱基特性(具有最佳活化剂长度的聚A段,然后是可以被U靶向性Cas13酶切割的聚U段)的RNA活化剂对这种方法进行了改进。申请人发现,在添加靶标后,LwaCas13a能够消化这些活化剂并产生针对EiCsm6和LsCsm6的最佳活化剂,并且该活化需要正确长度的聚A(图7H)。使用质谱法,申请人验证了LwaCas13a产生了用于Csm6激活的、预期的、结束于环状磷酸酯的产物(图7I和16A,B)。激活对于具有3'保护与聚U的设计是最有效的,而包括5'保护与聚U和内部聚U段的其他激活设计,仅在靶RNA的存在下在激活Csm6时较不有效(图17A,B)。通过在同一反应中组合用于Csm6和Cas13两者的报告体,申请人能够测定这两种酶的组合切割活性。申请人发现,增加活化剂浓度增加了被Cas13激活的Csm6的协同益处(图7J),并且增加Csm6特异性聚A报告体也增加了Csm6信号,导致在活化剂添加后更大的增加(图18A,B)。Csm6增强的LwaCas13a增加了合成乙酰转移酶基因以及来自大豆基因组DNA的除草剂抗性的整体信号和动力学(图7K,L和19A,B)。Csm6也可以增强来自U切割性CcaCas13b酶的信号(图20A-D)。

[0487] 因为将Csm6增强与RPA预扩增结合会增加信号,所以申请人在与RPA结合所必需的体外转录组分的存在下测试了Csm6的活性。申请人发现,镁和游离rNTP两者降低了Csm6在环状磷酸酯活化剂的存在下的核酸酶活性(图21A)。降低溶液中rNTP的量降低了转录的RNA的量,因此对来自Cas13a的Csm6活性具有负面影响(图21B-E),甚至是在增加的报告体或活化剂的存在下。与这两个竞争因素相对应,环状磷酸酯活化剂条件,减少rNTP恢复了Csm6特异性信号(图21F)。测试的所有rNTP种类导致了Csm6抑制,表明Csm6和rNTP解决方案会需要在不同反应中进行(图22)。申请人因此将转录步骤分离到单独反应中,从而允许足够的扩增,同时在Csm6反应中稀释所得rNTP,这个被分离的反应能够稳健地检测除草剂抗性基因(图23)。

[0488] 申请人假设,将Csm6与Cas13检测在侧流上结合,可通过使用报告体的组合来减少视觉读出(visual readout)的背景,从而防止单一报告序列的切割和假阳性读出。申请人测试了在Csm6和活化剂的存在下,具有各种各样的序列和长度的侧流报告体,发现长A/C报告体可以具有信号,但维持来自LwaCas13的正交性(orthogonality)(图24A,B)。申请人将这种报告体与RNaseAlert报告体结合使用,以在不存在RPA扩增的情况下,和仅在Csm6的存在下,检测登革热样本(图7M)。申请人随后将RPA、Csm6和侧流结合以检测酰基转移酶靶标,发现当使用混合探针时Csm6对于检测是必要的,并且背景被降低(图7N)。总体而言,这种技术组合导致了快速且稳健的视觉检测。

[0489] 实施例4:增加合成酰基转移酶基因检测信号的Csm6-增强的LwaCas13A

[0490] 为了放大检测信号,申请人利用了CRISPR III型效应核酸酶Csm6,其被终止于2',3'-环状磷酸酯的环状腺苷酸分子或线性腺嘌呤均聚物激活。LwaCas13a和PsmCas13b附带活性(collateral activity)产生了具有羟基化5'端和2',3'-环状磷酸酯末端的切割产物(图28),表明Cas13附带活性可以产生Csm6激活性种类物质,这会允许在SHERLOCK测定中的放大的信号的检测。通过测试具有不同长度和3'末端修饰的RNA腺苷酸分子(图29和图30A;图38),申请人发现,来自意大利肠球菌(Enterococcus italicus)的Csm6(EiCsm6)和来自唾液乳杆菌(Lactobacillus salivarius)的Csm6(LsCsm6)被含有2',3'-环状磷酸酯末端的六腺苷酸有效地激活(图30B-C)。此外,EiCsm6,LsCsm6和来自嗜热栖热菌(Thermus thermophilus)的Csm6(TtCsm6)基于传感器筛选展示了对富含A和C的传感器的强切割偏好,从而使得能够在单独的通道中独立地测量LwaCas13a和Csm6切割活性(图25D和图30B-D,图3,图32A-E)。

[0491] 为了使Cas13的活性与Csm6激活相结合,申请人设计了受保护的RNA活化剂,其含有聚A段,然后是保护性聚U段,其可以被尿嘧啶优先性Cas13酶切割,原理是LwaCas13a可以降解所有尿苷直至均聚物A段,因为它对UU和AU两碱基基序具有稳健的活性(图27)。申请人发现,在添加靶标和LwaCas13a-crRNA复合物后,EiCsm6和LsCsm6被$(A)_6$-$(U)_5$活化剂激活,这与A_6活化剂对于Csm6激活为最优的发现一致且通过质谱法确认(图25E和图32F,图33和图34)。申请人在同一荧光通道内的同一反应中组合了用于Csm6和Cas13两者的报告体,发现对于DENV ssRNA检测,增加活化剂浓度提高了Cas13对Csm6的协同激活(图25F),并且增加Csm6特异性聚A报告体也增加了Csm6信号,导致在活化剂添加后信号的更大的增加(图35A-B)。在优化后(图36)申请人发现,Csm6增强的LwaCas13a增加了通过SHERLOCK进行的合成酰基转移酶基因检测的总体信号和动力学(图25G)。

[0492] 为了改善检测的稳健性并减少假阳性读出的可能性,申请人将Csm6与Cas13检测在侧流上结合(图26K)。申请人在存在Csm6和活化剂的情况下测试了具有各种各样的序列和长度的侧流报告体,发现长A-C报告体显示了强切割信号(图37A-B)。申请人将该报告体与Cas13侧流报告体结合用于快速检测DENV ssRNA,对于扩增仅依赖于Csm6(即,在不存在RPA的情况下)(图26L)。申请人随后将RPA,Cas13/Csm6和侧流读出结合以检测酰基转移酶靶标,发现由Csm6赋予的信号的增加允许通过侧流进行的更快速的检测(图37C-D),具有降低的背景。

[0493] 材料和方法

[0494] Cas13和Csm6直系同源物的蛋白表达和纯化

[0495] LwaCas13a表达和纯化如前所述进行,稍作,并在下文详细描述。表达LbuCas13a,LbaCas13a,Cas13b和Csm6直系同源物并用改良方案纯化。简而言之,将细菌表达载体转化到RosettaTM 2(DE3)pLysS Singles Competent Cells(Millipore)中。将12.5mL启动培养物(starter culture)在Terrific Broth 4生长培养基(Sigma)(TB)中过夜生长,其用于接种4L的TB以在37℃和300RPM下生长直到OD600为0.5。此时,通过补充IPTG(Sigma)到终浓度为500μM,诱导蛋白质表达,并将细胞冷却至18℃达16h以进行蛋白质表达。然后将细胞在4℃下以5000g离心15分钟。收获细胞沉淀物并储存在-80℃下以用于随后的纯化。

[0496] 蛋白质纯化的所有后续步骤均在4℃下进行。压碎细胞沉淀并重悬于补充有蛋白

酶抑制剂(Complete Ultra EDTA-free tablets),溶菌酶(500μg/1ml)和benzonase的裂解缓冲液(20mM Tris-HCl,500mM NaCl,1mM DTT,pH 8.0)中,然后使用LM20 Microfluidizer系统以27,000PSI进行高压细胞破坏。通过在4℃下以10,000g离心1小时来使裂解物澄清。将上清液施加到5mL的StrepTactin Sepharose(GE)上,旋转孵育1小时,然后在裂解缓冲液中洗涤结合了蛋白的StrepTactin树脂3次。将该树脂与250单位的SUMO蛋白酶(250mg/ml)一起重悬于SUMO消化缓冲液(30mM Tris-HCl,500mM NaCl,1mM DTT,0.15%Igepal(NP-40),pH8.0)中,并于4℃下旋转孵育过夜。将悬浮液施加到柱,以通过重力流从树脂洗脱和分离。用1柱体积的裂解缓冲液洗涤树脂两次以最大化蛋白质洗脱。将洗脱液在阳离子交换缓冲液(20mM HEPES,1mM DTT,5%甘油,pH 7.0;对于LbuCas13a,LbaCas13a,EiCsm6,LsCsm6,TtCsm6为pH 7.5)中稀释以降低盐浓度至250mM以为阳离子交换色谱法作准备。

[0497] 对于阳离子交换和凝胶过滤纯化,将蛋白质通过FPLC(AKTA PURE,GE Healthcare Life Sciences)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 HEPES,1mM DTT,5%甘油,pH 7.0;对于LbuCas13a,LbaCas13a为pH 7.5)中在250mM至2M NaCl的盐梯度上洗脱。通过SDS-PAGE测试所得级分中重组蛋白的存在,合并含有该蛋白的级分,并通过Centrifugal Filter Unit(Millipore 50MWCO)在S200缓冲液(10mM HEPES,1M NaCl,5mM $MgCl_2$,2mM DTT,pH 7.0)中浓缩至1mL。通过FPLC将浓缩蛋白质上样到凝胶过滤柱(Superdex®200Increase10/300GL,GE Healthcare Life Sciences)上。通过SDS-PAGE分析凝胶过滤所得的馏分,合并含蛋白质的馏分,将缓冲液交换至存储缓冲液(600mM NaCl,50mM Tris-HCl,pH 7.5,5%甘油,2mM DTT)中,并于-80℃冷冻用于存储。

[0498] 在本研究中纯化的所有蛋白质的登录号和质粒图谱可在图39中获得。

[0499] 核酸靶标和crRNA制备

[0500] 用NEBNext PCR主混合物(master mix)PCR扩增用于Cas12a和基因组DNA检测的核酸靶标,凝胶提取,并使用MinElute凝胶提取试剂盒(Qiagen)纯化。对于基于RNA的检测,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30℃孵育过夜,并使用MEGAclear Transcription Clean-up试剂盒(Thermo Fisher)纯化RNA。

[0501] crRNA制备如前所述进行,稍作修改,并在下文详细描述。对于crRNA的制备,构建体订购为具有附加T7启动子序列的超聚DNA(Integrated DNA Technologies)。将crRNA DNA退火至短T7引物(终浓度为10uM),并使用HiScribe T7 Quick High Yield RNA Synthesis试剂盒(New England Biolabs)在37℃下与T7聚合酶一起孵育过夜。使用RNAXP清洁珠(Beckman Coulter)以2倍的珠子与反应体积的比率纯化crRNA,并另外添加1.8倍的异丙醇(Sigma)。

[0502] 在该研究中使用的所有crRNA序列可在图40中获得。该研究中使用的所有DNA和RNA靶序列可在图41中获得。

[0503] 使用NCBIPrimer-BLAST使用默认参数设计RPA引物,除了扩增子大小(在100和140nt之间),引物解链温度(在54℃和67℃之间)和引物大小(在30和35nt之间)。引物然后订购为DNA(Integrated DNA Technologies)。

[0504] RPA和RT-RPA反应分别按照TwistAmp®Basic或TwistAmp®Basic RT(TwistDx)指示的进行,除了在输入模板之前添加了280mM MgAc。除非另外说明,否则反应在37℃下以1μL输入进行1小时。

[0505] 对于核酸的SHERLOCK定量,在标准浓度(480nM最终)和更低(240nM,120nM,60nM,24nM)下测试RPA引物浓度以找到最佳浓度。RPA反应进一步进行20分钟。

[0506] 当用RPA扩增多个靶标时,将引物浓度调节至480nM的最终浓度。即,对于两个引物对添加120nM的每个引物用于双链体检测。

[0507] 在该研究中使用的所有RPA引物可在图42中获得。

[0508] 荧光切割测定

[0509] 检测测定如前所述进行,稍作修改,该程序在下面详述。使用45nM纯化Cas13,22.5nM crRNA,淬灭的荧光RNA报告体(125nM RNAse Alert v2,Thermo Scientific,均聚物和二核苷酸报告体(IDT);对于聚A Trilink报告体为250nM),0.5μL鼠RNase抑制剂(New England Biolabs),25ng的背景总人类RNA(从HEK293FT培养物纯化),和不同量的输入核酸靶标进行检测测定,除非另有说明,否则是在核酸酶测定缓冲液(20mM HEPES,60mM NaCl,6mM $MgCl_2$,pH 6.8)中。对于Csm6荧光切割反应,将蛋白质以10nM的终浓度与500nM的2',3'环状磷酸酯低聚腺苷酸,250nM的荧光报告体和0.5μL鼠RNase抑制剂一起在核酸酶测定缓冲液(20mM HEPES,60mM NaCl,6mM $MgCl_2$,pH 6.8)中使用。使反应在荧光板读数器(BioTek)上在37℃(除非另有说明)进行1-3小时,每5分钟测量荧光动力学。在涉及AsCas12a的反应中,使用来自IDT的重组蛋白,45nM AsCas12a被包括。在多重反应的情况下,反应中使用45nM的每种蛋白质和22.5nM的每种crRNA。

[0510] 在本研究中使用的所有切割报告体可在图43中获得。

[0511] SHERLOCK核酸检测

[0512] 用45nM纯化Cas13,22.5nM crRNA,淬灭的荧光RNA报告体(125nM RNAse Alert v2,Thermo Scientific,均聚物和二核苷酸报告体(IDT),对于聚A Trilink报告体为250nM),0.5μL鼠RNase抑制剂(New England Biolabs),25ng背景总人类RNA(从HEK293FT培养物纯化)和1uL RPA反应在核酸酶测定缓冲液(20mM HEPES,60mM NaCl,6mM $MgCl_2$,pH 6.8),rNTP混合物(1mM最终,NEB),0.6μL T7聚合酶(Lucigen)和3mM $MgCl_2$中进行检测测定。使反应在荧光板读数器(BioTek)上在37℃(除非另有说明)进行1-3小时,每5分钟测量荧光动力学。

[0513] 对于一锅式核酸检测,检测测定如前所述进行,稍作修改。单个100μL组合反应测定包括0.48μM正向引物,0.48μM反向引物,1×RPA再水合缓冲液,不同量的DNA输入,45nM LwCas13a重组蛋白,22.5nM crRNA,125ng背景人类总RNA,125nM底物报告体(RNase Alert v2),2.5μL鼠RNase抑制剂(New England Biolabs),2mM ATP,2mM GTP,2mM UTP,2mM CTP,1μL T7聚合酶混合物(Lucigen),5mM $MgCl_2$和14mM MgAc。使反应在荧光板读数器(BioTek)上在37℃(除非另有说明)进行1-3小时,每5分钟测量荧光动力学。对于侧流读出,将20uL的组合反应添加到100uL的HybriDetect1测定缓冲液(Milenia)中,并在HybriDetect 1侧流条(Milenia)上运行。

[0514] 用于切割片段分析的核酸标记

[0515] 如上所述,从dsDNA模板体外转录靶RNA并纯化。体外切割反应如前文荧光切割反

应所述进行,具有以下修改。1μg RNA靶标代替荧光报告体,并且没有使用背景RNA。切割反应在37℃下进行5分钟(LwaCas13a)或1小时(PsmCas13b)。使用RNA clean&concentrator-5试剂盒(Zymo Research)纯化切割反应,并在10uL UltraPure水(Gibco)中洗脱。按照5' EndTag labeling Reaction(Vector Laboratories)试剂盒方案,进一步用10μg的马来酰亚胺IRDye 800CW(Licor)标记切割反应。为了确定由Cas13切割产生的5'末端,对方案进行了修改以执行碱性磷酸酶(AP)处理或用UltraPure水替代以仅标记含有5'-OH的RNA种类,而未消化的三磷酸(PPP)RNA种类为仅在执行AP处理时标记的。

[0516] 用于高分辨率切割片段分析的质谱

[0517] 为了通过质谱确定由Cas13附带RNase活性产生的切割末端,体外切割反应如上所述进行,并具有以下修改。使用1nM最终浓度的Cas13 RNA靶标,3μM最终浓度的Csm6活化剂,不使用背景RNA。对于对照反应,用UltraPure水代替Cas13靶标,或在不存在Cas13靶标,Cas13蛋白和Cas13 crRNA的情况下,将标准体外切割反应与含有2',3'环状磷酸酯活化剂的六腺苷酸一起孵育。切割反应在37℃下进行1小时,并使用New England Biolabs siRNA纯化方案纯化。简而言之,加入十分之一体积的3M NaOAc,2μL不含RNase的Glycoblue(Thermofisher)和三体积的95%冷乙醇,在-20℃放置2小时,并在14,000g离心15分钟。除去上清液,加入两体积的80%EtOH,并在室温下孵育10分钟。倒出上清液,并将样品以14,000g离心5分钟。将沉淀风干后,加入50μL UltraGrade水并送到干冰上用于质谱分析。

[0518] 对于质谱分析,将样品用UltraGrade水按1:1稀释,并在与Agilent1290HPLC偶联的Bruker Impact II q-TOF质谱仪上以负离子模式进行分析。将10μL注射到使用在水中的0.1%氢氧化铵v/v作为流动相A和乙腈作为流动相相B的PLRP-S柱(50mm,5um粒径,1000埃孔径PLRP-S柱,2.1mm ID)上。始终保持流速恒定为0.3ml/分钟。流动相组成始于0%B并保持前2分钟。此后,组成在接下来的8分钟改变为100%B并保持一分钟。组成然后在0.1分钟回到0%B,然后保持接下来的4.9分钟以使柱重新平衡至起始条件。质谱仪针对大MW离子进行了调谐,并在m/z 400-5000之间采集了数据。使用甲酸钠的注射,通过m/z校准质谱仪的整个数据集。使用具有MaxEnt解卷积算法许可的Bruker Compass Data Analysis 4.3分析数据,以从带负电的离子数据生成计算出的中性质谱。

[0519] 从人唾液中提取基因组DNA

[0520] 唾液DNA提取如前所述进行,稍作修改,并在下文详述。从在收集前30分钟限制消耗食物或饮料的志愿者收集了2mL唾液。然后按照试剂盒方案建议的,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)添加到100μL志愿者唾液中,并以1800g离心5分钟。倒出上清液,将沉淀重悬在含0.2%Triton X-100(Sigma)的磷酸盐缓冲盐水中,然后在95℃孵育5分钟。1μL样品用作RPA反应中的直接输入。

[0521] 数字液滴PCR定量

[0522] ddPCR定量如前所述进行,稍作修改,并在下文详述。为了确认靶标稀释液的浓度,我们进行了数字液滴PCR(ddPCR)。对于DNA定量,使用具有针对靶标序列设计的PrimeTime qPCR探针/引物测定(IDT)的ddPCR Supermix for Probes(无dUTP)(BioRad)制备液滴。对于RNA定量,使用用于具有针对靶标序列设计的PrimeTime qPCR探针/引物测定的探针的一步法RT-ddPCR试剂盒制备液滴。在任一情况下使用QX200液滴发生器(BioRad)产生液滴,并

将其转移到PCR板中。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液滴读取器上进行测量来确定核酸浓度。

[0523] Cas13-Csm6荧光切割测定

[0524] Cas13-Csm6组合荧光切割测定如标准Cas13荧光切割反应所述进行,具有以下修改。除非另有说明,否则Csm6蛋白添加至10nM最终浓度,400nM Csm6荧光报告体和500nM Csm6活化剂。为了区分Cas13与Csm6附带RNase活性,使用了两种不同的荧光团进行荧光检测(FAM和HEX)。由于rNTP对Csm6活性的干扰,IVT在RPA预扩增步骤中进行,然后将1μL的该反应作为输入添加到Cas13-Csm6切割测定。

[0525] 在我们测试了三步Cas13-Csm6切割测定的情况下,RPA通常按上文讨论的进行不同的时间,然后用作不同时间的正常IVT反应的输入。然后将1μL的IVT用作上段所述的Cas13-Csm6反应的输入。本研究中使用的所有Csm6活化剂可在图38中获得。

[0526] 采用文库的基序发现筛选

[0527] 为了筛选Cas13切割偏好,体外RNA切割反应如上所述建立,具有以下修改。Cas13靶标在20nM使用,1μM的DNA-RNA寡核苷酸(IDT)替代荧光报告体,所述DNA-RNA寡核苷酸含有两侧是DNA把手的6聚段的随机化核糖核苷酸,用于NGS文库制备。反应在37℃下进行60分钟(除非另有说明)。使用Zymo oligo-clean and concentrator-5试剂盒(Zymo research)纯化反应,并使用15μL UltraPure水洗脱。使用与DNA把手结合的基因特异性引物,将10μL的纯化反应用于逆转录。

[0528] 根据qScript Flex cDNA-试剂盒(quantabio)方案,于42℃进行45分钟的逆转录(RT)。为了评估切割效率和产物纯度,将RT-反应物在水中以1:10稀释,并加载到Small RNA试剂盒中,并在Bioanalyzer 2100(Agilent)上运行。第一轮的NGS文库制备使用四微升的RT反应。NEBNext(NEB)用于扩增第一链cDNA,使用625nM的正向引物和625nM的反向引物的混合物,15个循环,在98℃初始变性3分钟,在98℃循环变性10s,在63℃退火10s,在72℃延伸20s,和在72℃最终延伸2分钟。

[0529] 将两微升的第一轮PCR反应用于第二轮PCR扩增,以连接用于NGS测序的Illumina相容性索引(NEB)。相同的NEBNext PCR方案用于扩增。通过琼脂糖凝胶电泳(2%Sybr Gold E-Gel Invitrogen系统)分析PCR产物,合并5μL的每个反应。凝胶提取合并的样品,用Qubit DNA 2.0DNA高灵敏度试剂盒定量并针对4nM终浓度进行标准化。将最终文库稀释至2pM,并使用75循环试剂盒在NextSeq500Illumina系统上测序。

[0530] 基序筛选分析

[0531] 为了分析来自随机基序文库筛选的偏好基序的消耗,从序列数据中提取了6聚区域,并针对每个样品的总读取计数进行标准化。然后将标准化的读取计数用于在实验条件和匹配的对照之间生成对数比率,具有伪计数(pseudocount)调整。对于Cas13实验,匹配的对照不具有添加的靶RNA;对于Csm6和RNase A实验,匹配的对照不具有酶。使用对数比率分布形状来确定富集的基序的临界值。然后使用富集的基序确定1、2或3核苷酸组合的出现。基序标识是使用Weblogo3生成。

[0532] Cas13蛋白和crRNA直接重复序列的系统发生分析

[0533] 为了研究直系同源物聚类,在具有MUSCLE的Geneious中用Cas13a和Cas13b蛋白序列产生多序列比对,然后使用具有heatmap.2功能的R中的欧几里德距离来聚类。为了研究

直接重复序列聚类,使用Geneious算法在Geneious中用Cas13a和Cas13b直接重复序列生成多序列比对,然后使用具有heatmap.2功能的R中的欧几里德距离来聚类。为了研究基于二核苷酸基序偏好的直系同源物的聚类,使用heatmap.2功能,使用R中的欧几里德距离对切割活性矩阵进行聚类。

[0534] 金纳米粒子比色法

[0535] 从IDT合成RNA低聚物,在5'和3'端具有巯基(序列见图43)。为了使巯基脱保护,将终浓度为20mM的低聚物在含有100mM DTT的150mM磷酸钠缓冲液中室温还原2小时。然后使用sephadex NAP-5柱(GE Healthcare)将低聚物纯化至最终体积为700μL的水中。如前所述,将10μM的还原低聚物以280μL的体积添加到600μL的2.32nM 15nm金纳米颗粒(TedPella)中,其是低聚物与纳米颗粒的2000:1比率。随后,将10μL的pH 8.3的1M Tris-HCl和90μL的1M NaCl添加到低聚物-纳米颗粒混合物中,并在室温下旋转孵育18小时。18小时后,加入另外的1M Tris-HCl(5μL、pH 8.3)和5M NaCl(50μL),这在室温下再旋转孵育15小时。孵育后,将最终溶液以22,000g离心25分钟。弃去上清液,将缀合的纳米颗粒重悬于50μL的200mM NaCl中。

[0536] 使用RNase A测定法测试纳米颗粒的RNase敏感性。将不同量的RNase A(Thermo Fischer)添加到20μL总反应体积中的1×RNase A缓冲液和6μL的缀合的纳米颗粒。使用平板分光光度计每5分钟监测520nm处的吸光度,持续3小时。

[0537] 使用FAM-生物素报告体,Cas13活性的侧流读出

[0538] 对于基于FAM-RNA-生物素报告体的切割的侧流,将非RPA LwaCas13a反应或SHERLOCK-LwaCas13a反应运行1小时,除非另有说明,FAM-RNA-生物素报告体的最终浓度为1uM。孵育后,将20uL LwaCas13a反应上清液添加到100uL HybriDetect 1测定缓冲液(Milenia)中,并在HybriDetect 1侧流条(Milenia)上运行。

[0539] REPAIR构建体的克隆,哺乳动物细胞转染,RNA分离和用于REPAIR的NGS文库制备

[0540] 如前所述,克隆了用于模拟APC突变的回复的构建体和用于REPAIR的指导构建体。简而言之,设计了以APC:c.1262G>A突变为中心的96nt序列,在表达载体下金门(golden gate)克隆,并将相应的指导序列金门克隆到用于PspCas13b指导体的U6表达载体中。为了模拟患者样品,将300ng突变体或野生型APC表达载体用Lipofectamine 2000(Invitrogen)转染到HEK293FT细胞中,并按照生产商的说明用Qiamp DNA Blood Midi Kit(Qiagen)收获转染后两天的DNA。将20ng DNA用作SHERLOCK-LwaCas13a反应中的输入。

[0541] 如前所述,使用REPAIR系统进行RNA校正:将150ng的dPspCas13b-ADAR(DD)E488Q,200ng的指导载体和30ng的APC表达载体共转染,按照制造商的说明使用RNeasy Plus Mini Kit(Qiagen)收获转染后两天的RNA。将30ng RNA用作SHERLOCK-LwaCas13a反应中的输入。

[0542] 如前所述,通过NGS独立地确定RNA编辑级分。使用带有序列特异性引物的qScript Flex试剂盒(Quanta Biosciences)对RNA进行反转录。用NEBNext High Fidelity 2X PCR Mastermix(New England Biosciences)扩增第一链cDNA,使用625nM最终的正向引物和625nM的反向引物的混合物,15个循环,在98℃下初始变性3分钟,在98℃下循环变性10秒,在65℃下退火30秒,在72℃下延伸30秒,和在72℃下最终延伸2分钟。使用NEBNext,使用具有18个循环的相同方案,两微升的第一轮PCR反应用于第二轮PCR扩增,以连接用于NGS测序的Illumina相容性索引。通过琼脂糖凝胶电泳(2%Sybr Gold E-Gel Invitrogen)分析PCR

产物,合并5μL的每个反应。凝胶提取合并的样品,用Qubit DNA 2.0DNA高灵敏度试剂盒定量并针对4nM最终浓度进行标准化,并用300cycle v2 MiSeq试剂盒(Illumina)读出。

[0543] SHERLOCK荧光数据的分析

[0544] SHERLOCK荧光分析如前所述进行,略有修改,并在下文详细描述。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。

[0545] 如下计算用于SNP区分的crRNA比率,以调整样品之间(sample-to-sample)的总体变化:

[0546] $$crRNA\ A_i\ 比率 = \frac{(m+n)A_i}{\sum_{i=1}^{m} A_i + \sum_{i=1}^{n} B_i}$$

[0547] 其中,对于给定个体,Ai和Bi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。由于我们通常具有每个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。因为在SHERLOCKv2中,我们通过测量不同颜色通道中的A_i和B_i来完成基因分型,所以我们将530颜色通道缩放6倍以匹配480颜色通道中的强度值。

[0548] 在不存在靶标的情况下Cas13直系同源物的混杂切割

[0549] Cas13家族的一些成员,例如PinCas13b和LbuCas13a,在存在或不存在靶标的情况下显示混杂切割,并且该背景活性是二核苷酸报告体依赖性的。该背景活性对于LbuCas13a也是间隔子依赖性的。在一些报告体中,U和A碱基偏好在蛋白质或DR相似性内聚类。有趣的是,此处鉴定的二核苷酸偏好与从直接重复序列相似性或蛋白质相似性聚类的Cas13家族不对应。

[0550] 用于PsmCas13b和CcaCas13b的crRNA设计的表征

[0551] 为了确定对于PsmCas13b和CcaCas13b检测最佳的crRNA,我们测试了34-12nt的crRNA间隔子长度,发现PsmCas13b在30的间隔子长度具有峰值灵敏度,而CcaCas13b在高于28nt的间隔子长度具有相等的灵敏度,证明使用30nt间隔子评估Cas13活性是有道理的。为了进一步探索与LwaCas13a相比CcaCas13b和PsmCas13b的靶向的稳健性,我们设计了在ssRNA 1上平均间隔开的11个不同crRNA,发现LwaCas13a附带活性对crRNA设计是稳健的,而CcaCas13b和PsmCas13b均显示了活性在不同crRNA中的更多的变化性。

[0552] 针对另外的正交基序的随机文库基序筛选

[0553] 为了进一步探索Cas13a和Cas13b直系同源物的切割偏好的多样性,我们开发了基于文库的方法以表征用于附带核酸内切酶活性的偏好基序。我们使用了两侧为恒定DNA把手的简并6聚RNA报告体,其允许未切割的序列的扩增和读出。将该文库与Cas13酶一起孵育,产生了可检测的切割模式,该切割模式取决于靶RNA的添加,并且对来自这些反应的耗尽基序的测序揭示了该文库的偏态在消化时间期间的增加,其指示一群用于切割的偏好基序。来自高度耗尽的基序的序列标识和成对碱基偏好重现了对LwaCas13a和CcaCas13b观察到的U偏好以及PsmCas13b的A偏好。我们从如从该筛选确定的排在前面的基序合成了报告

体以验证这些发现,发现LwaCas13a,CcaCas13a和PsmCas13b均切割了它们最高度偏好的基序。我们还发现了仅对一个直系同源物而不对其他直系同源物显示切割的多个序列,这可以允许来自二核苷酸基序的可供选择的正交读出。与不同靶标一起孵育的LwaCas13a产生相似的切割基序偏好,表明无论靶标序列如何,附带活性的碱基偏好都是恒定的。

[0554] 在LwaCas13a切割后,活化剂产物的验证

[0555] 使用质谱法,我们验证了LwaCas13a消化产生了用于Csm6激活的、预期的、终止于环状磷酸酯的产物。激活对具有3'保护与聚U的设计是最有效的,而包括5'保护与聚U和内部聚U段的其他激活设计,排他地在存在靶RNA的情况下在激活Csm6时较不有效(可能是因为LwaCas13a对UA基序几乎没有活性,而5'保护对阻止Csm6的激活无效)。

[0556] 对组合RPA和Csm6反应的优化

[0557] 因为将Csm6增强与RPA预扩增结合会增加信号和灵敏度,所以我们在存在与RPA结合所必需的体外转录组分的情况下测试了Csm6的活性。我们发现镁和游离rNTP都在存在环状磷酸酯活化剂的情况下降低了Csm6的核酸酶活性。减少溶液中rNTP的量减少了转录的RNA的量,因此对Cas13a对Csm6的激活具有负面影响,即使在增加的报告体或活化剂浓度的存在下。

[0558] 通过以下编号的段落进一步描述本发明:

[0559] 1.一种组合物,其包含一种或多种检测CRISPR效应蛋白和一种或多种信号放大CRISPR效应蛋白。

[0560] 2.段落1所述的组合物,其中所述一种或多种检测CRISPR效应蛋白是VI型CRISPR效应蛋白。

[0561] 3.段落2所述的组合物,其中所述VI型CRISPR效应蛋白是Cas13a,Cas13b或两者。

[0562] 4.前述段落中任一项所述的组合物,其中所述一种或多种信号放大CRISPR效应蛋白是IIIa型CRISPR蛋白。

[0563] 5.段落4所述的组合物,其中所述III型CRISPR蛋白是Csm6蛋白。

[0564] 6.前述段落中任一项所述的组合物,其还包含设计用于结合相应靶分子的一种或多种指导序列。

[0565] 7.段落7所述的组合物,其还包含一种或多种激活序列,所述激活序列设计用于在所述激活序列的切割后激活所述一种或多种信号放大CRISPR效应蛋白。

[0566] 8.前述段落中任一项所述的组合物,其还包含报告构建体。

[0567] 9.前述段落中任一项所述的组合物,其还包含受保护的第二指导序列和第二靶序列,其中所述第二指导序列设计用于在所述受保护的第二指导序列上的保护性元件的去除后结合所述第二靶序列。

[0568] 10.前述段落中任一项所述的组合物,其还包含受保护的第二靶序列和第二指导序列,其中所述第二指导序列设计用于在所述受保护的第二靶序列上的保护性元件的去除后结合所述第二靶序列。

[0569] 11.前述段落中任一项所述的组合物,其还包含受保护的或环化的引物和模板编码指导序列和/或靶序列,其中所述受保护的或环化的引物设计用于在保护性元件的去除或所述环化的引物的切割后引发所述模板的扩增反应。

[0570] 12.前述段落中任一项所述的组合物,其还包含扩增试剂。

[0571] 13.一种诊断装置,其包括段落1至12中任一项所述的组合物。

[0572] 14.一种用于检测样品中的靶分子的方法,其包括:

[0573] 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段落8至12中任一项所述的组合物;

[0574] 在足以允许所述一种或多种指导序列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;

[0575] 通过所述一种或多种指导序列与所述一种或多种靶分子的结合来激活所述检测CRISPR效应蛋白,其中激活所述检测CRISPR效应蛋白导致所述激活序列的切割,使得所述信号放大CRISPR效应蛋白被激活,并且其中激活的检测CRISPR效应蛋白和激活的信号放大CRISPR效应蛋白两者修饰报告构建体,使得可检测阳性信号产生;和

[0576] 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所述样品中一种或多种靶分子的存在。

[0577] ***

[0578] 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。

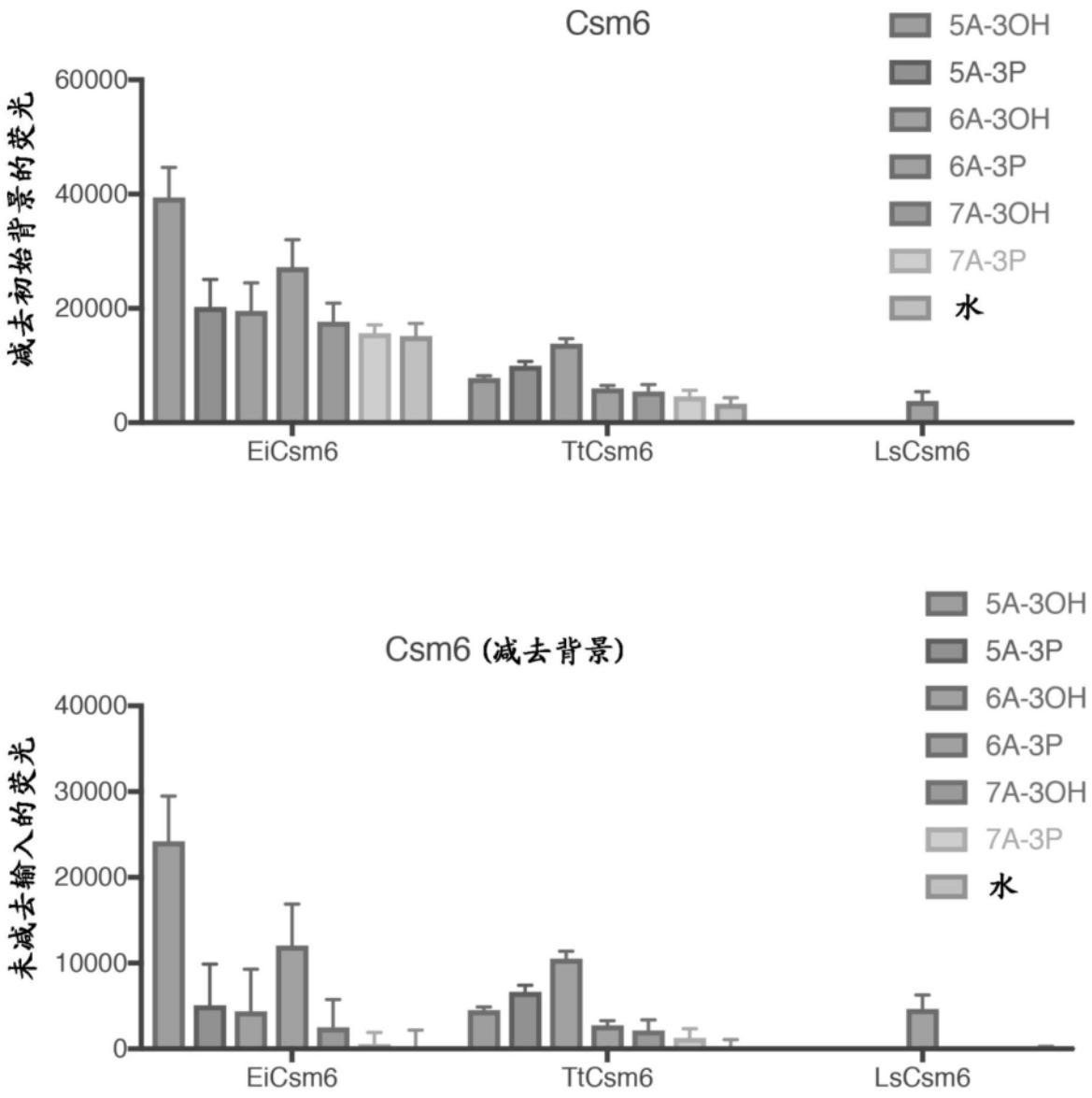

Csm6
减去初始背景的荧光
60000
40000
20000
0
EiCsm6
TtCsm6
LsCsm6
5A-3OH
5A-3P
6A-3OH
6A-3P
7A-3OH
7A-3P
水
Csm6 (减去背景)
未减去输入的荧光
40000
30000
20000
10000
0
EiCsm6
TtCsm6
LsCsm6
5A-3OH
5A-3P
6A-3OH
6A-3P
7A-3OH
7A-3P
水

图1

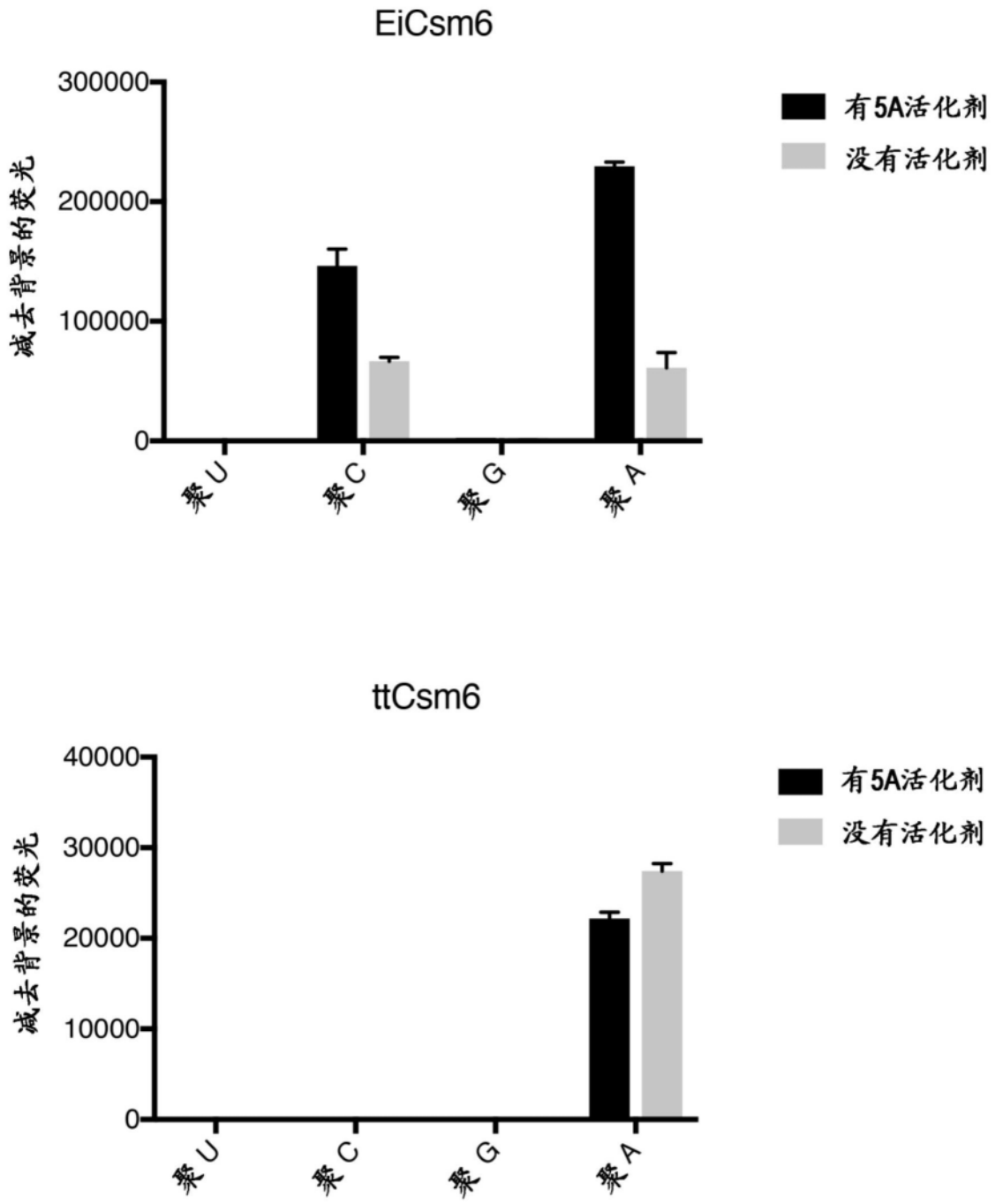

EiCsm6
减去背景的荧光
300000
200000
100000
0
有5A活化剂
没有活化剂
靶U
靶C
靶G
靶A
ttCsm6
减去背景的荧光
40000
30000
20000
10000
0
有5A活化剂
没有活化剂
靶U
靶C
靶G
靶A

图2

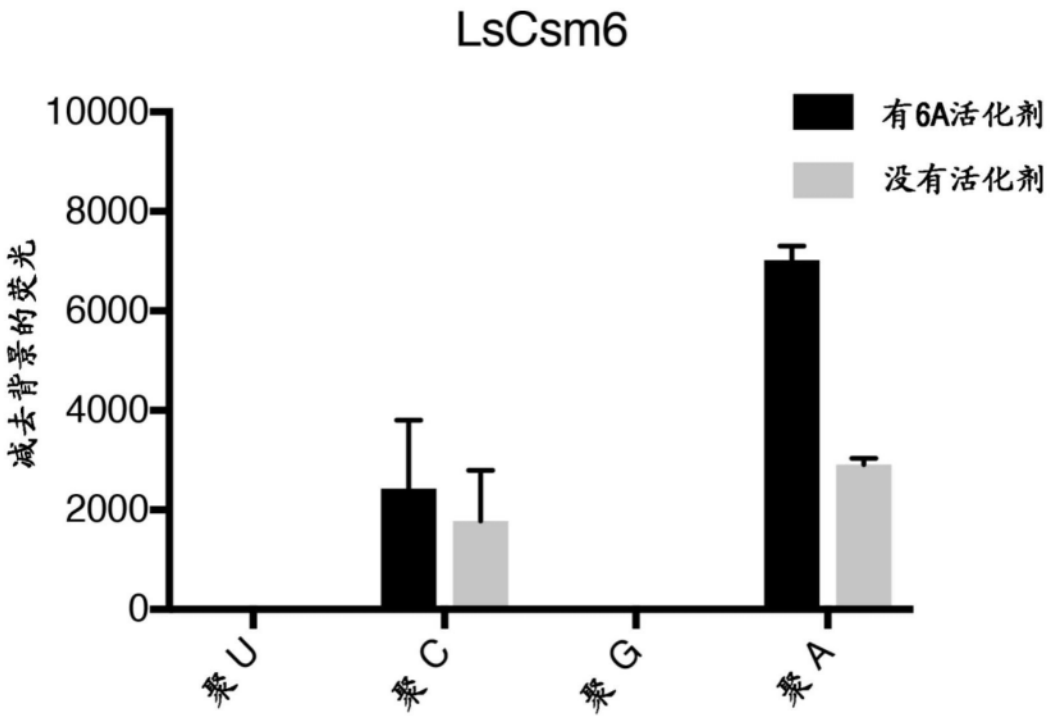

LsCsm6
有6A活化剂
没有活化剂
减去背景的荧光
10000
8000
6000
4000
2000
0
靶U
靶C
靶G
靶A

图2续

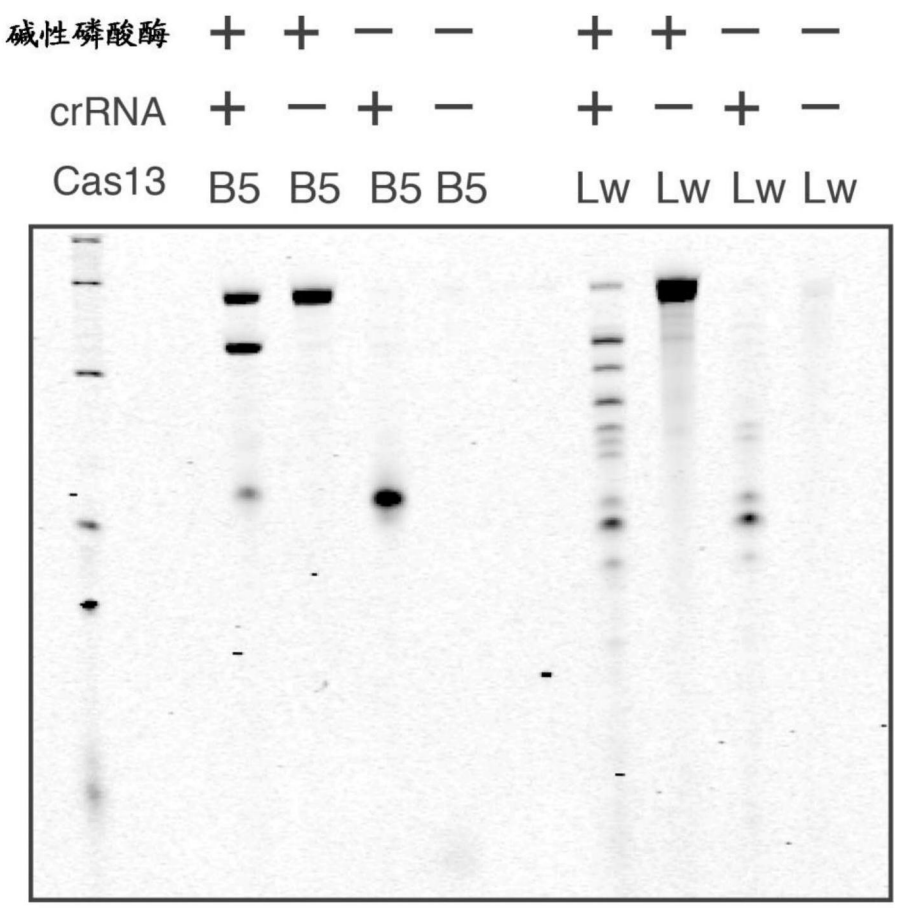


图3

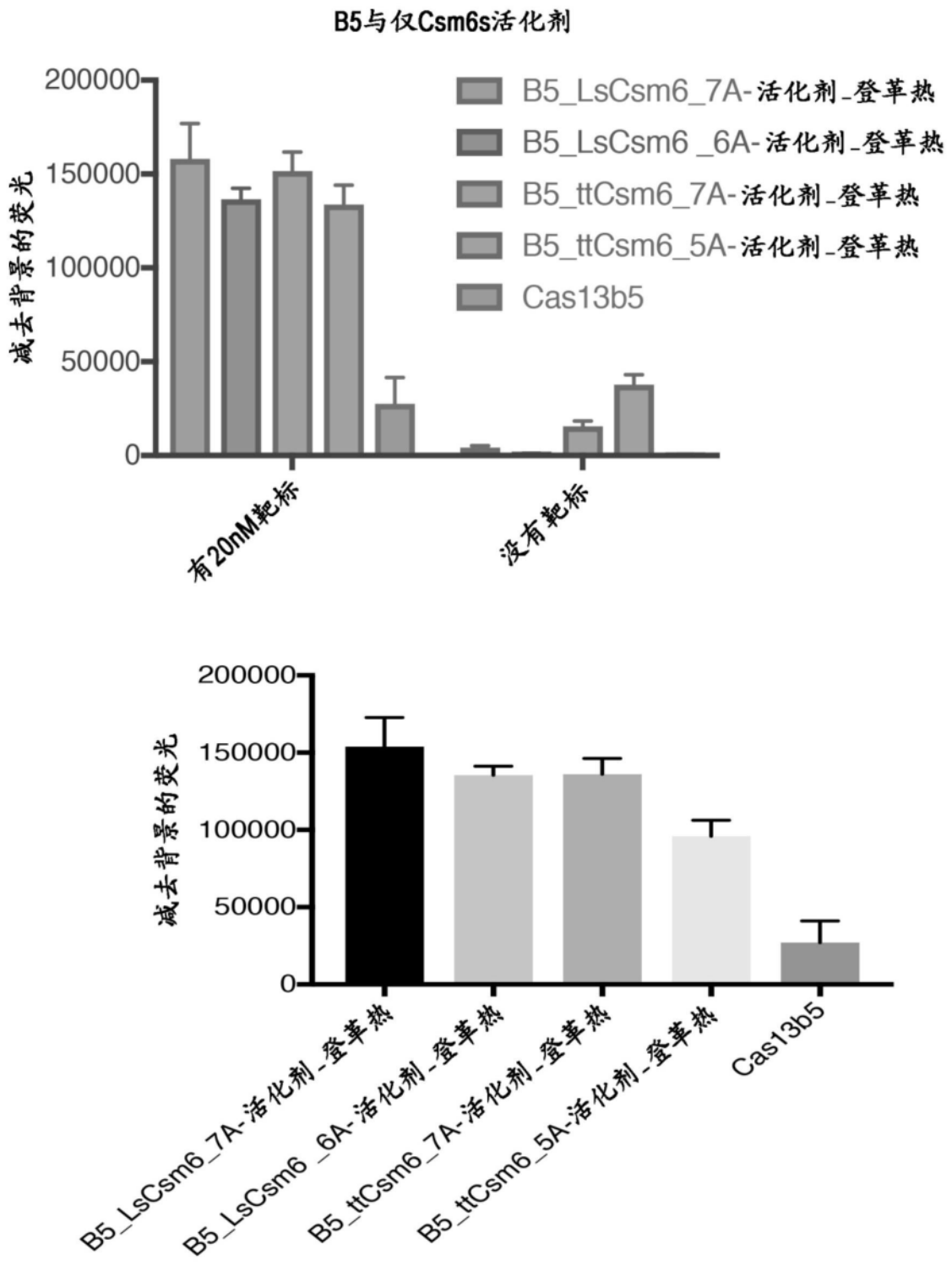

B5与仅Csm6s活化剂
200000
150000
100000
50000
0
减去背景的荧光
B5_LsCsm6_7A-活化剂_登革热
B5_LsCsm6 _6A-活化剂_登革热
B5_ttCsm6_7A-活化剂_登革热
B5_ttCsm6_5A-活化剂_登革热
Cas13b5
有20nM靶标
没有靶标
200000
150000
100000
50000
0
减去背景的荧光
B5_LsCsm6_7A-活化剂_登革热
B5_LsCsm6 _6A-活化剂_登革热
B5_ttCsm6_7A-活化剂_登革热
B5_ttCsm6_5A-活化剂_登革热
Cas13b5

图4

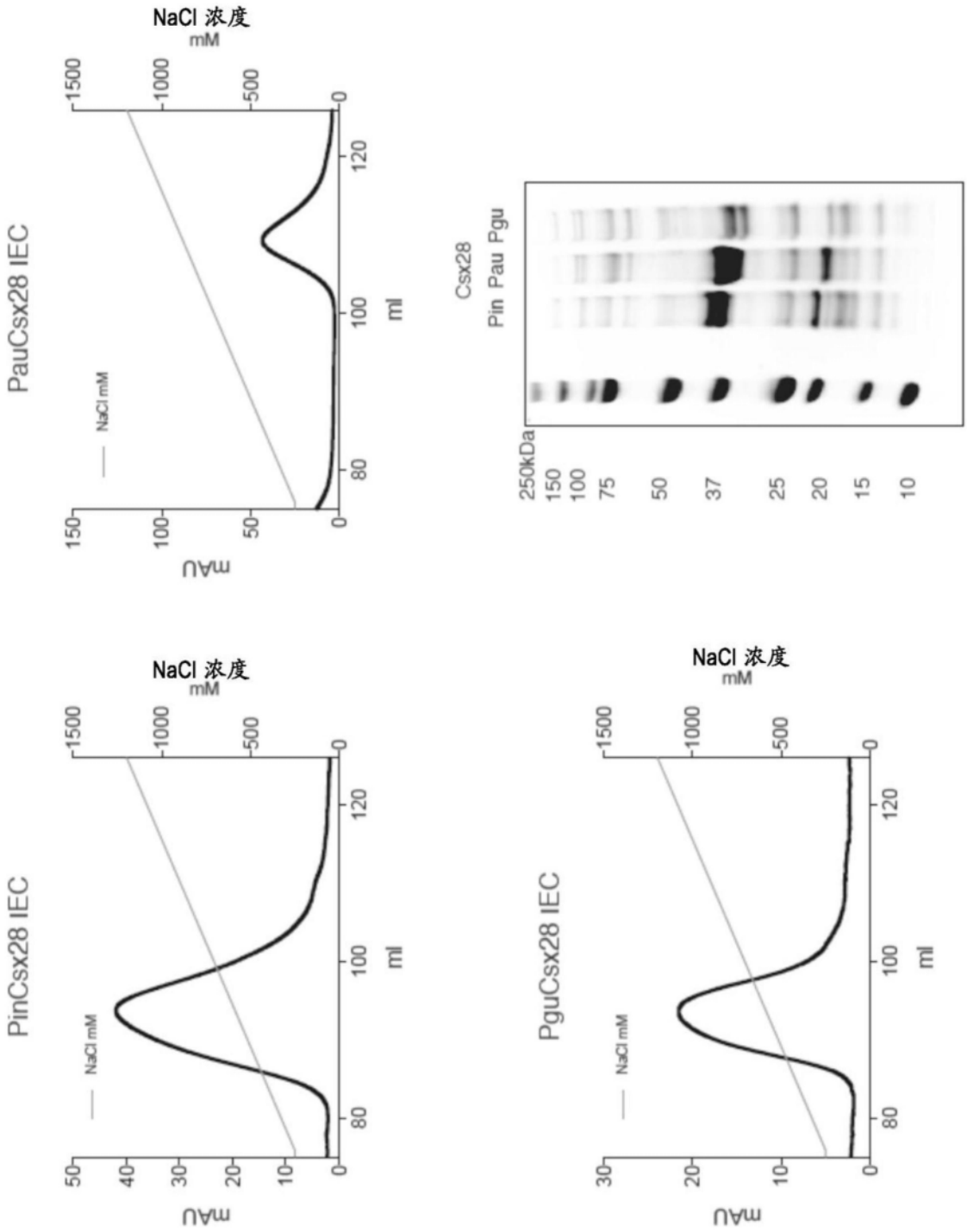

NaCl 浓度
mM
PauCsx28 IEC
NaCl mM
mAU
ml
Csx28
Pin Pau Pgu
250kDa
150
100
75
50
37
25
20
15
10
PinCsx28 IEC
PguCsx28 IEC

图5

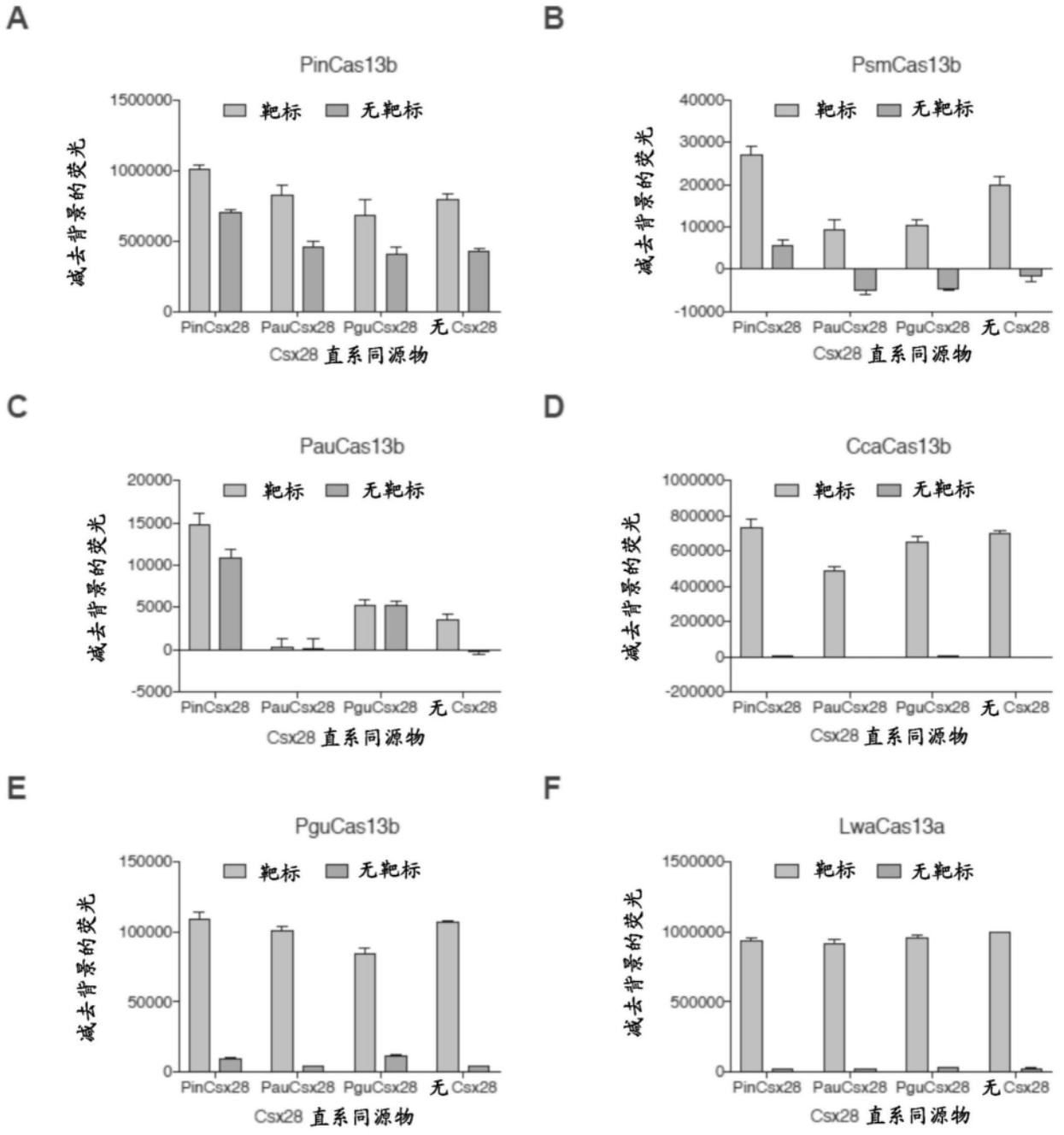

A
PinCas13b
减去背景的荧光
1500000
1000000
500000
0
靶标
无靶标
PinCsx28
PauCsx28
PguCsx28
无Csx28
Csx28直系同源物
B
PsmCas13b
减去背景的荧光
40000
30000
20000
10000
0
-10000
靶标
无靶标
PinCsx28
PauCsx28
PguCsx28
无Csx28
Csx28直系同源物
C
PauCas13b
减去背景的荧光
20000
15000
10000
5000
0
-5000
靶标
无靶标
PinCsx28
PauCsx28
PguCsx28
无Csx28
Csx28直系同源物
D
CcaCas13b
减去背景的荧光
1000000
800000
600000
400000
200000
0
-200000
靶标
无靶标
PinCsx28
PauCsx28
PguCsx28
无Csx28
Csx28直系同源物
E
PguCas13b
减去背景的荧光
150000
100000
50000
0
靶标
无靶标
PinCsx28
PauCsx28
PguCsx28
无Csx28
Csx28直系同源物
F
LwaCas13a
减去背景的荧光
1500000
1000000
500000
0
靶标
无靶标
PinCsx28
PauCsx28
PguCsx28
无Csx28
Csx28直系同源物

图6

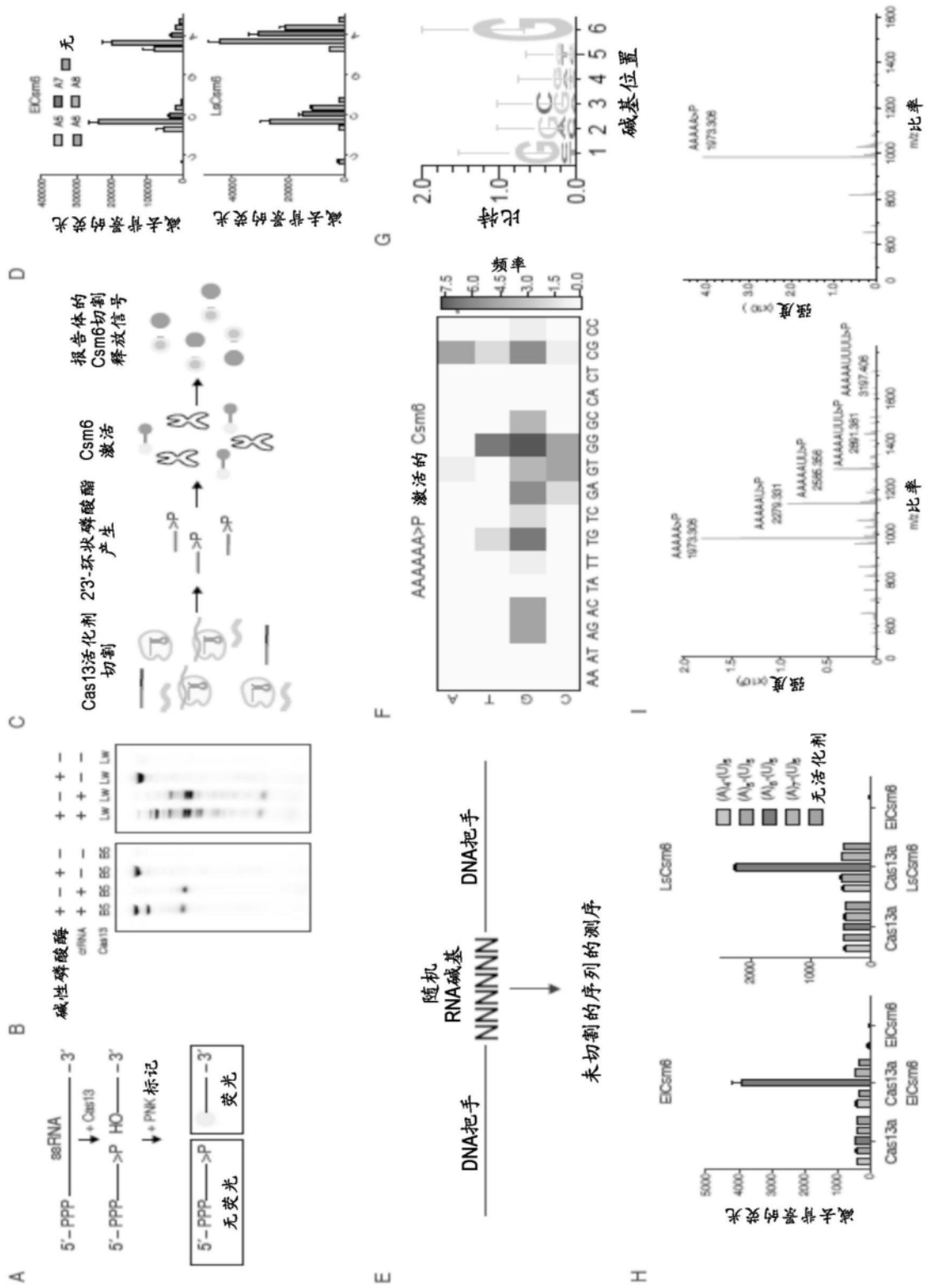

A
B
C
D
E
F
G
H
I
碱性磷酸酶
ssRNA
crRNA
Cas13
+ Cas13
+ PNK 标记
无荧光
荧光
Cas13活化剂
切割
2'3'-环状磷酸酯
产生
Csm6
激活
报告体的
Csm6切割
释放信号
随机
RNA碱基
NNNNNN
DNA把手
未切割的序列的测序
AAAAAA>P 激活的 Csm6
碱基位置
频率
无活化剂
m/z比率
LsCsm6
EiCsm6
Cas13a

图7

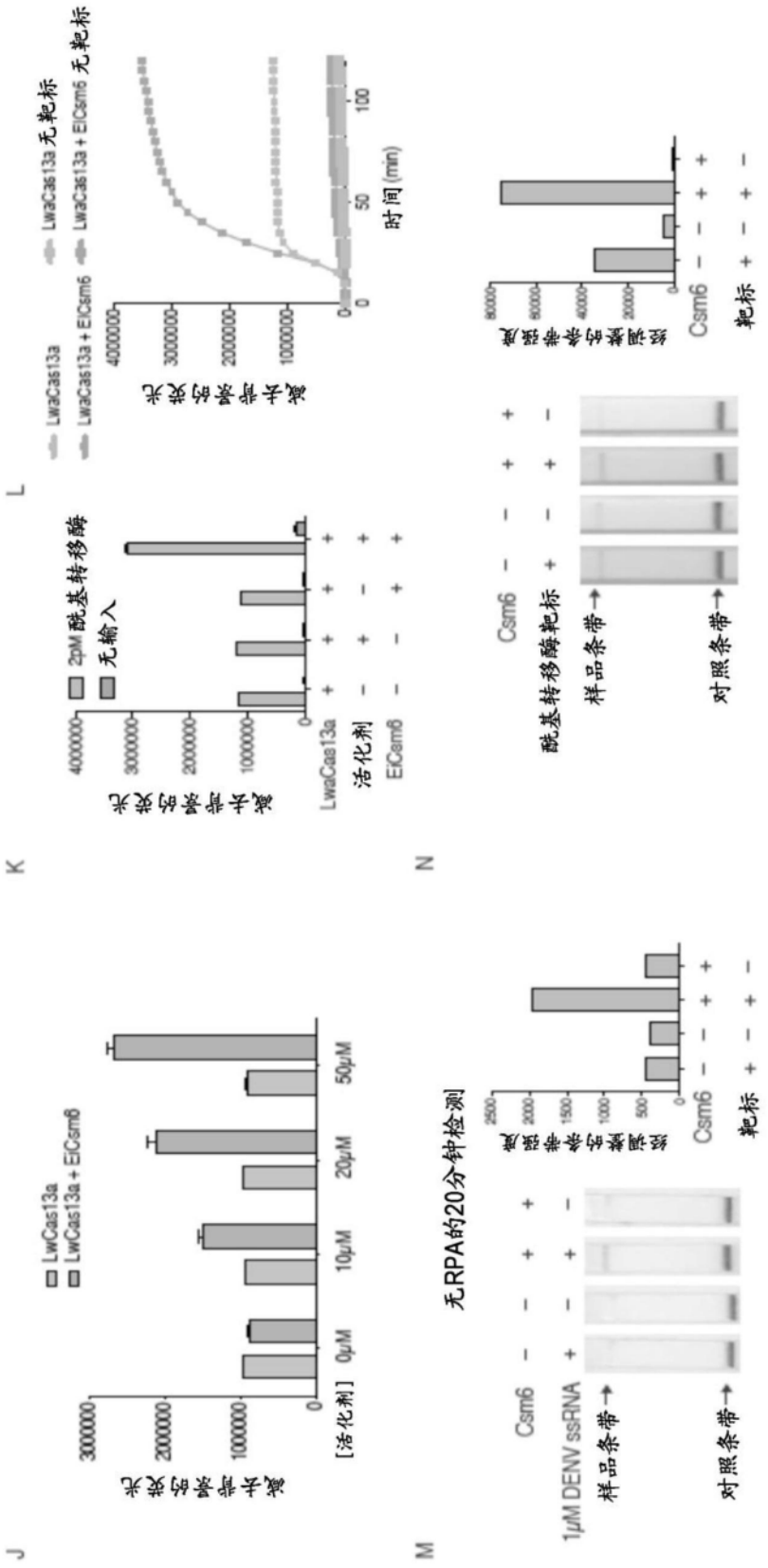

J
LwCas13a
LwCas13a + EiCsm6
荧光背景减除
[活化剂]
0μM
10μM
20μM
50μM
K
2pM 酰基转移酶
无输入
LwaCas13a
EiCsm6
活化剂
L
LwaCas13a
LwaCas13a + EiCsm6
LwaCas13a 无靶标
LwaCas13a + EiCsm6 无靶标
时间 (min)
M
无RPA的20分钟检测
1μM DENV ssRNA
Csm6
样品条带→
对照条带→
靶标
N
酰基转移酶靶标
样品条带→
对照条带→
Csm6
靶标

图7 (续)

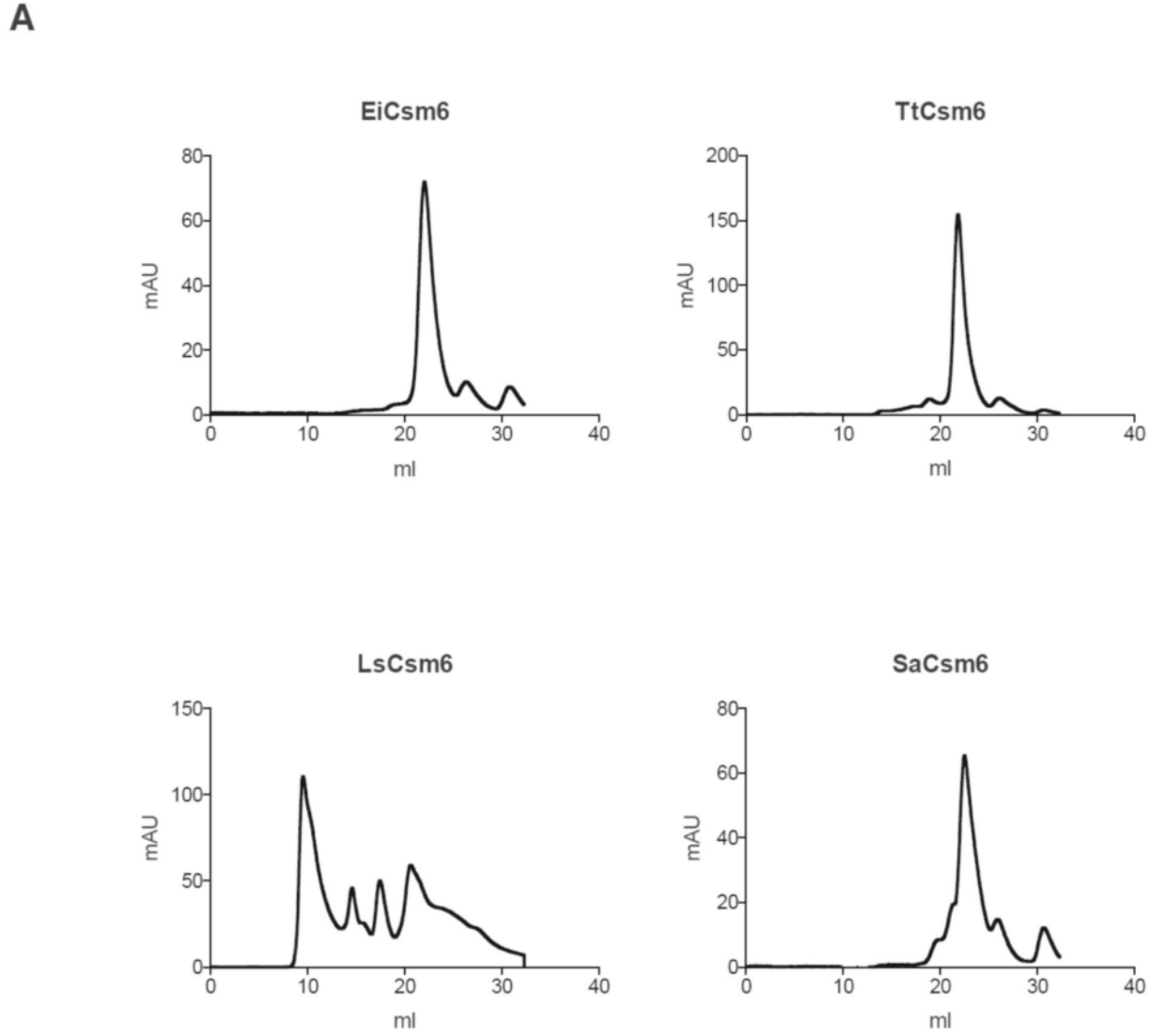
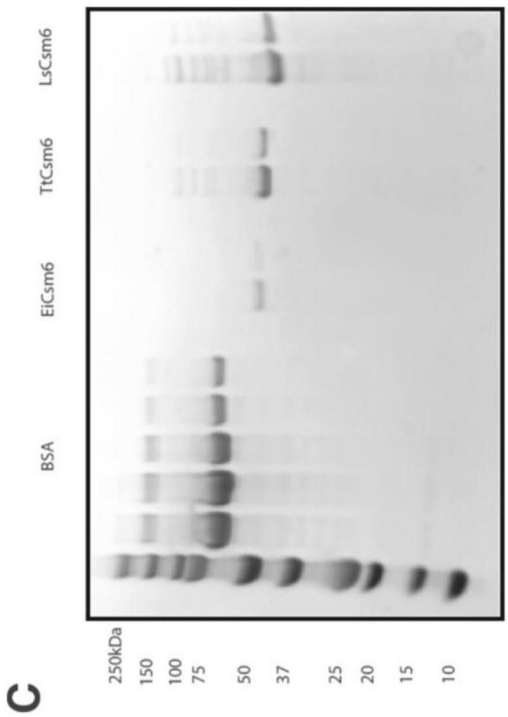

A
EiCsm6
TtCsm6
LsCsm6
SaCsm6
mAU
ml
0
10
20
30
40
0
20
40
60
80
0
50
100
150
200

图8


C
BSA
EiCsm6
TtCsm6
LsCsm6
250kDa
150
100
75
50
37
25
20
15
10

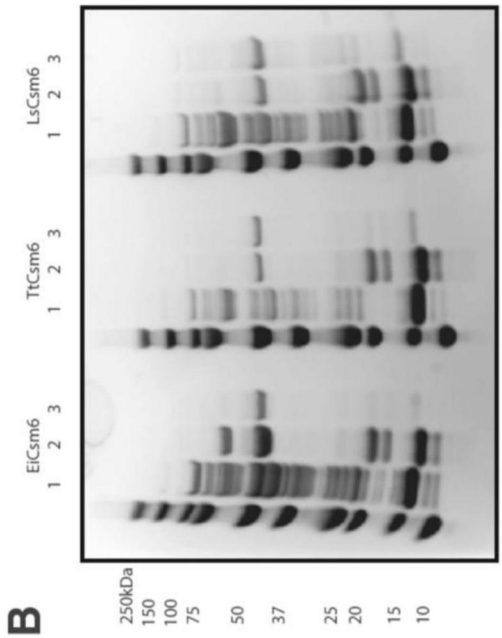

B
EiCsm6
1 2 3
TtCsm6
1 2 3
LsCsm6
1 2 3
250kDa
150
100
75
50
37
25
20
15
10

图8(续)

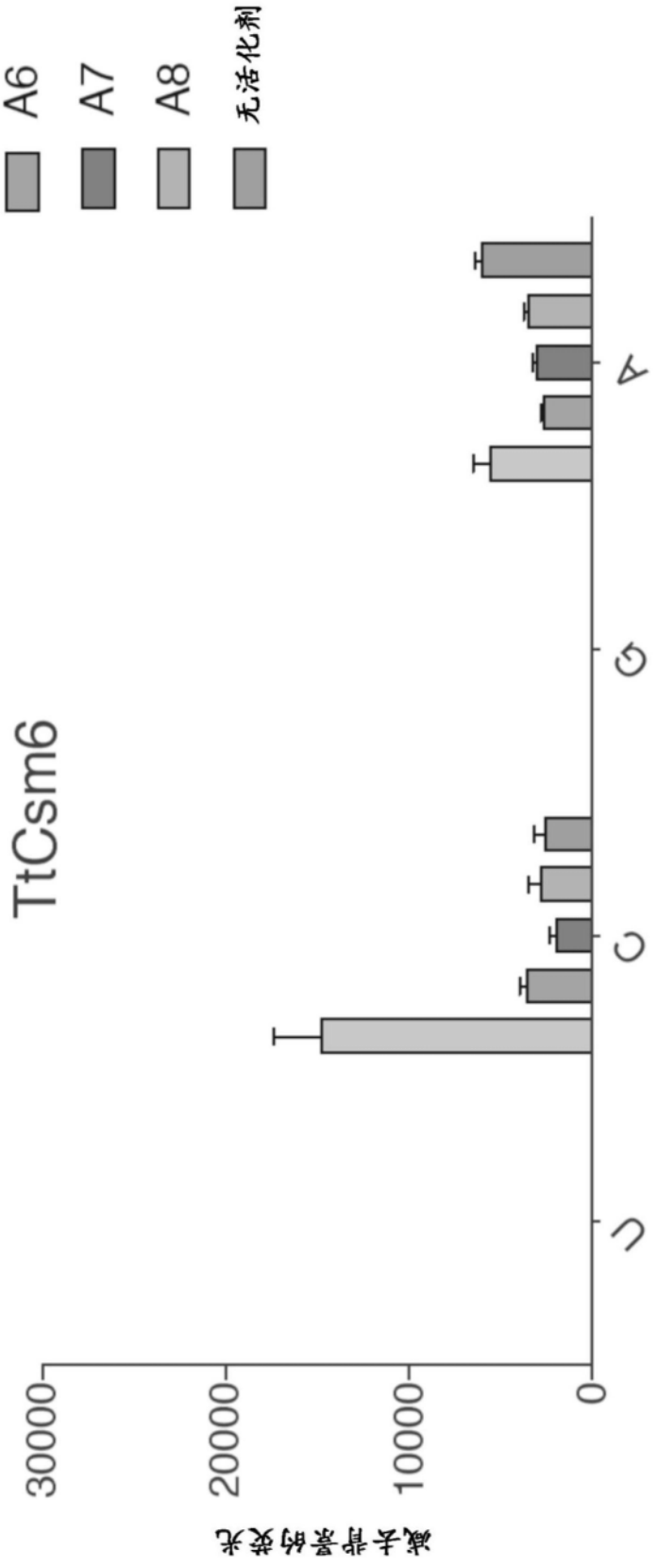


图9

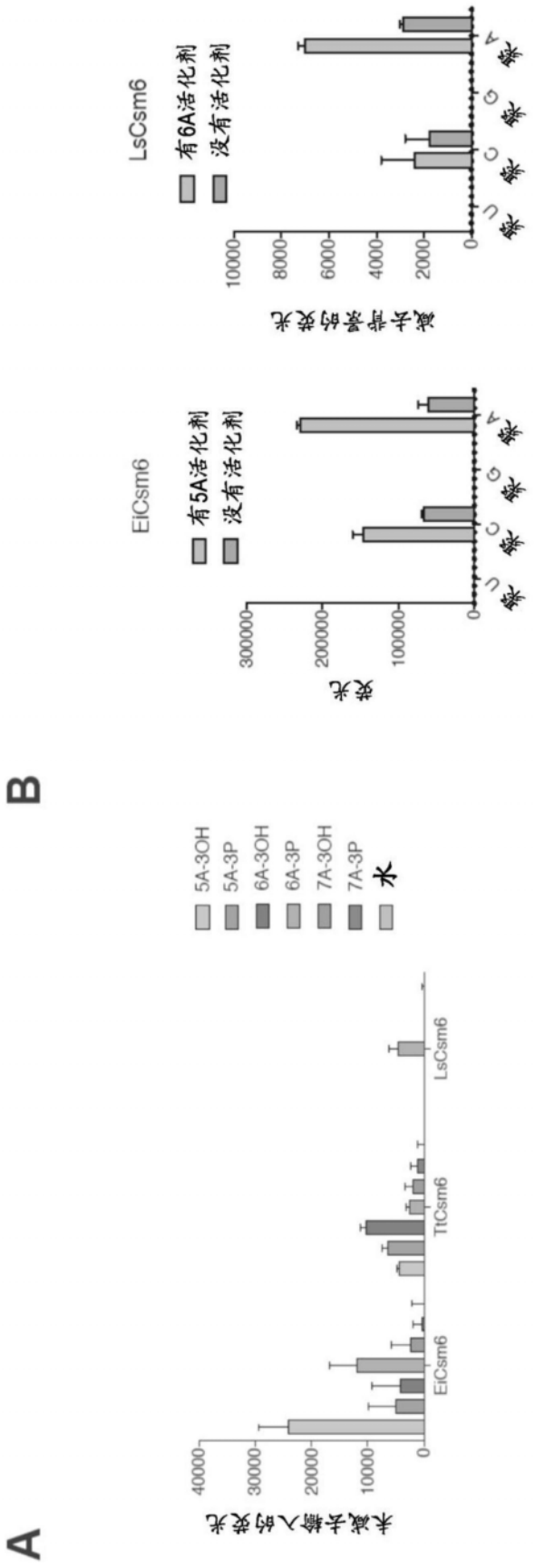

A
未裂去撒入的荧光
40000
30000
20000
10000
0
EiCsm6
TtCsm6
LsCsm6
5A-3OH
5A-3P
6A-3OH
6A-3P
7A-3OH
7A-3P
水
B
EiCsm6
有5A活化剂
没有活化剂
荧光
300000
200000
100000
0
LsCsm6
有6A活化剂
没有活化剂
荧光
10000
8000
6000
4000
2000
0

图10

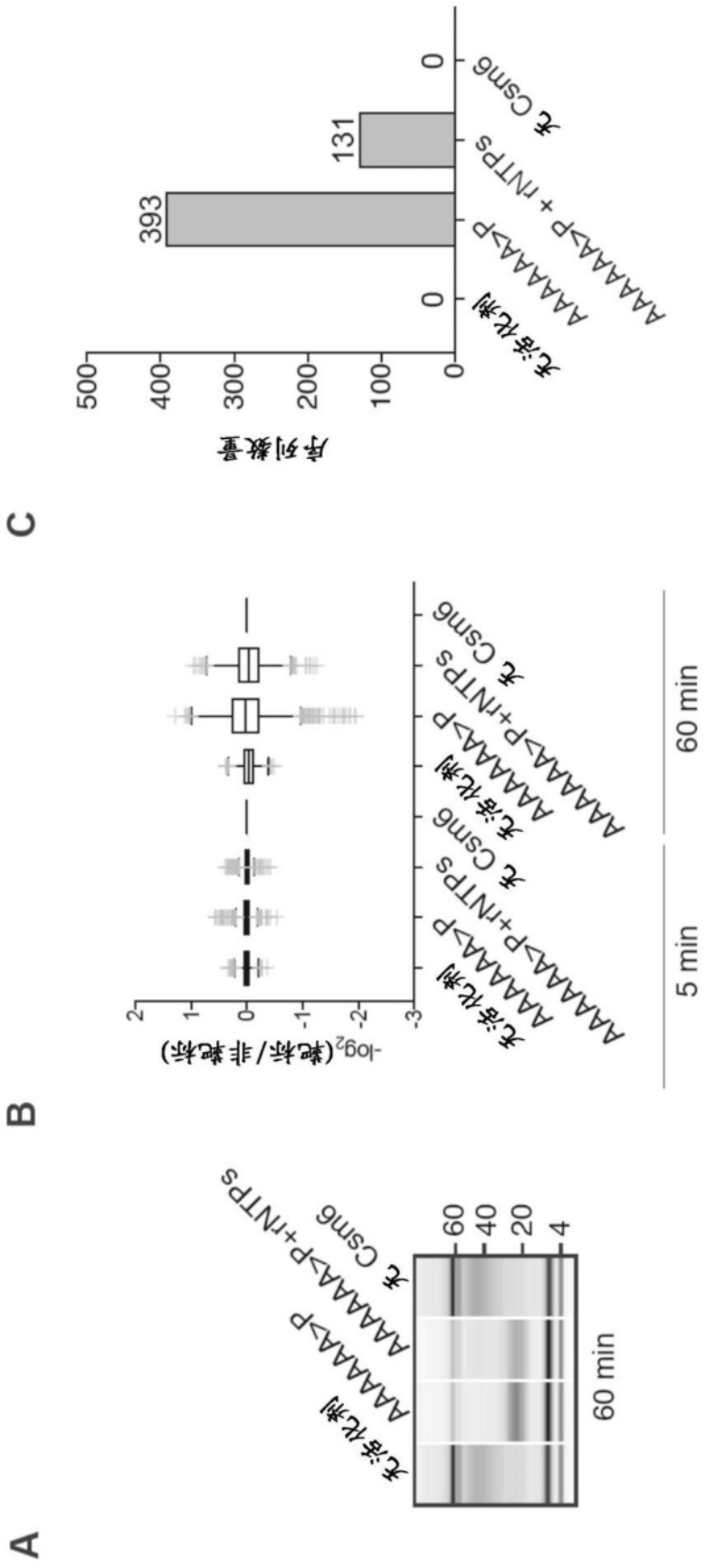

A
未活化剂
AAAAAA>P
AAAAAA>P+rNTPs
未 Csm6
60
40
20
4
60 min
B
未活化剂
AAAAAA>P
AAAAAA>P+rNTPs
未 Csm6
未活化剂
AAAAAA>P
AAAAAA>P+rNTPs
未 Csm6
5 min
60 min
-log2(靶标/非靶标)
2
1
0
-1
-2
-3
C
未活化剂
AAAAAA>P
AAAAAA>P + rNTPs
未 Csm6
0
393
0
131
0
序列数量
500
400
300
200
100
0

图11

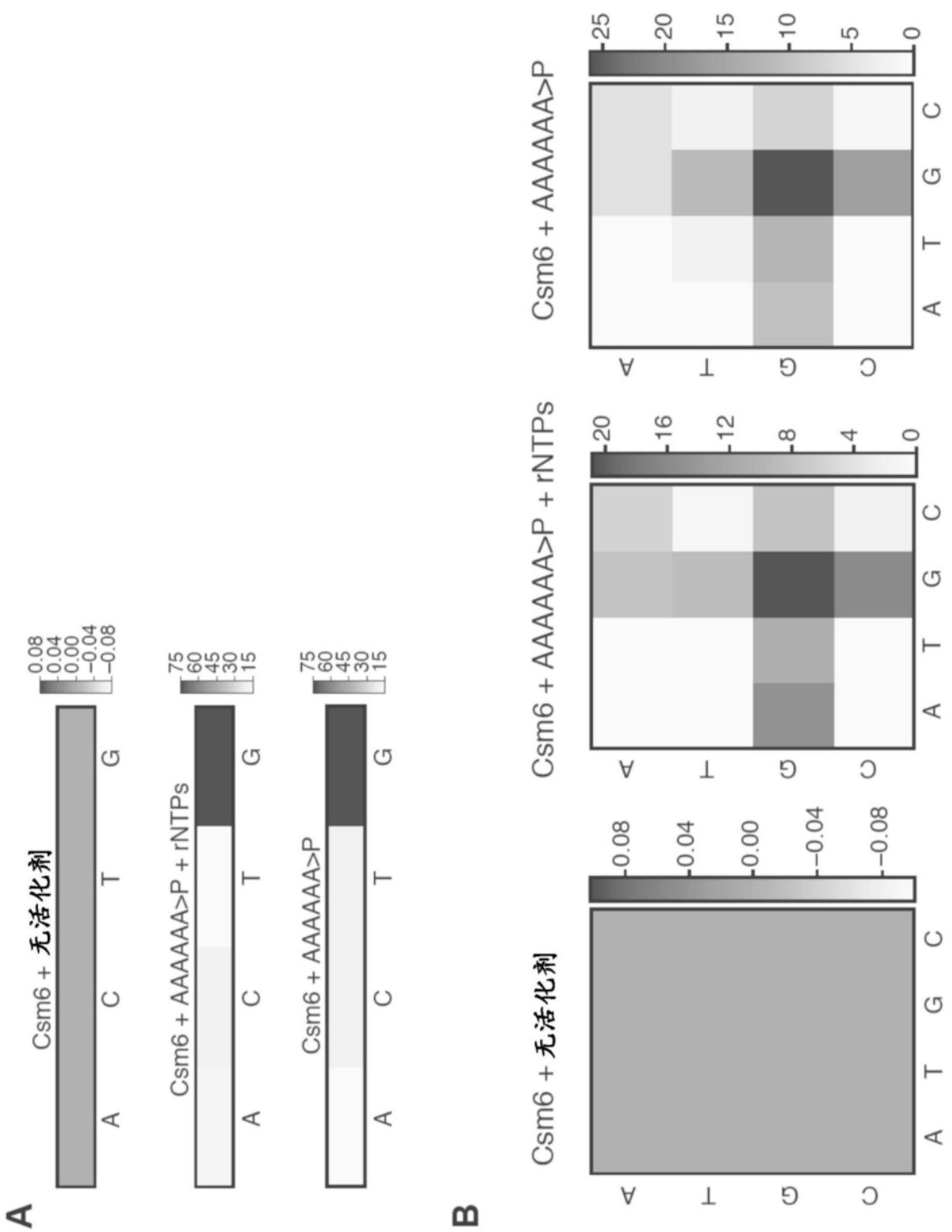

A
Csm6 + 无活化剂
0.08
0.04
0.00
−0.04
−0.08
A C T G
Csm6 + AAAAAA>P + rNTPs
75
60
45
30
15
A C T G
Csm6 + AAAAAA>P
75
60
45
30
15
A C T G
B
Csm6 + 无活化剂
0.08
0.04
0.00
−0.04
−0.08
A T G C
Csm6 + AAAAAA>P + rNTPs
20
16
12
8
4
0
A T G C
Csm6 + AAAAAA>P
25
20
15
10
5
0
A T G C

图12

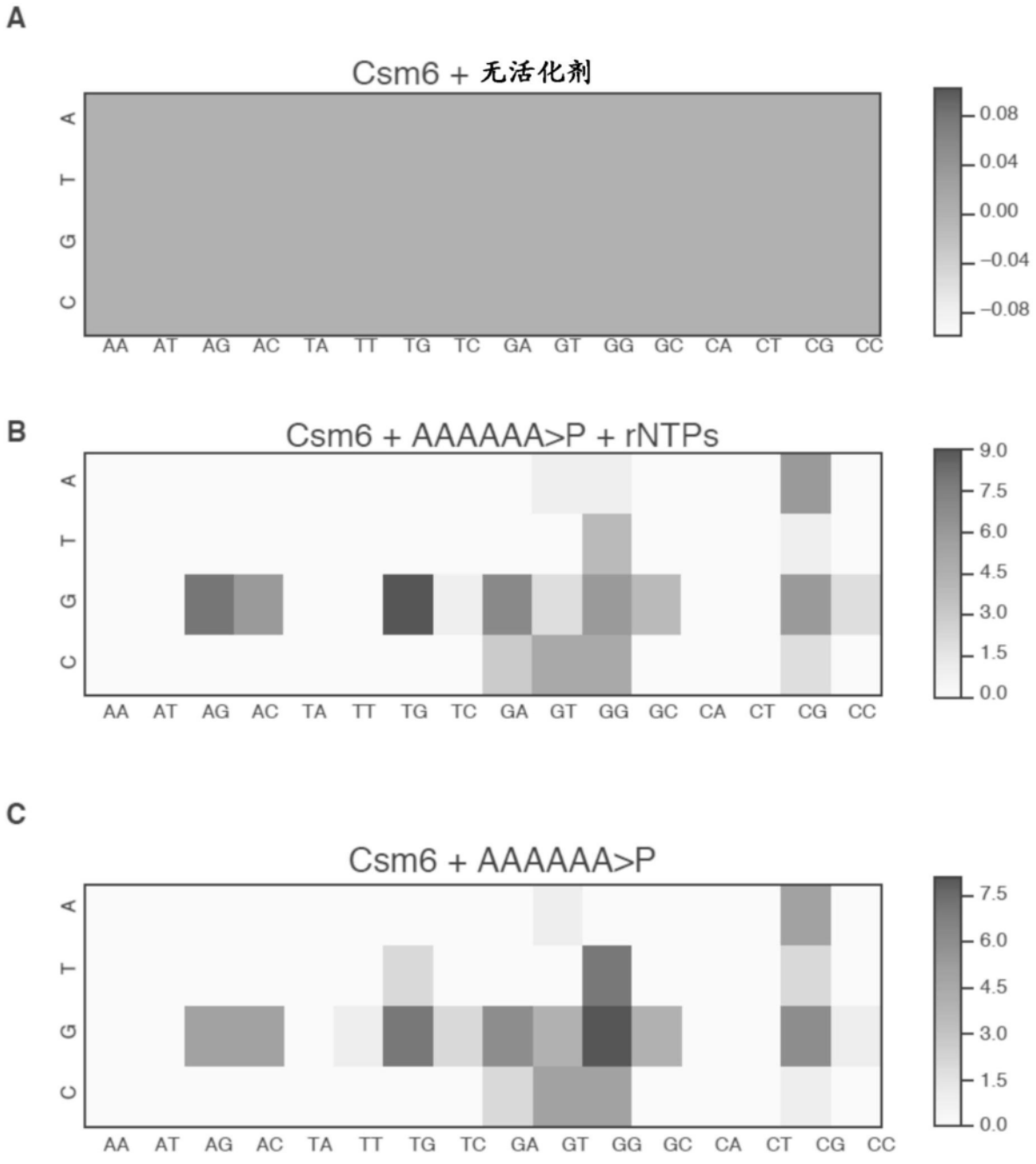

A
Csm6 + 无活化剂
A
T
G
C
AA AT AG AC TA TT TG TC GA GT GG GC CA CT CG CC
0.08
0.04
0.00
−0.04
−0.08
B
Csm6 + AAAAAA>P + rNTPs
9.0
7.5
6.0
4.5
3.0
1.5
0.0
C
Csm6 + AAAAAA>P
7.5
6.0
4.5
3.0
1.5
0.0

图13

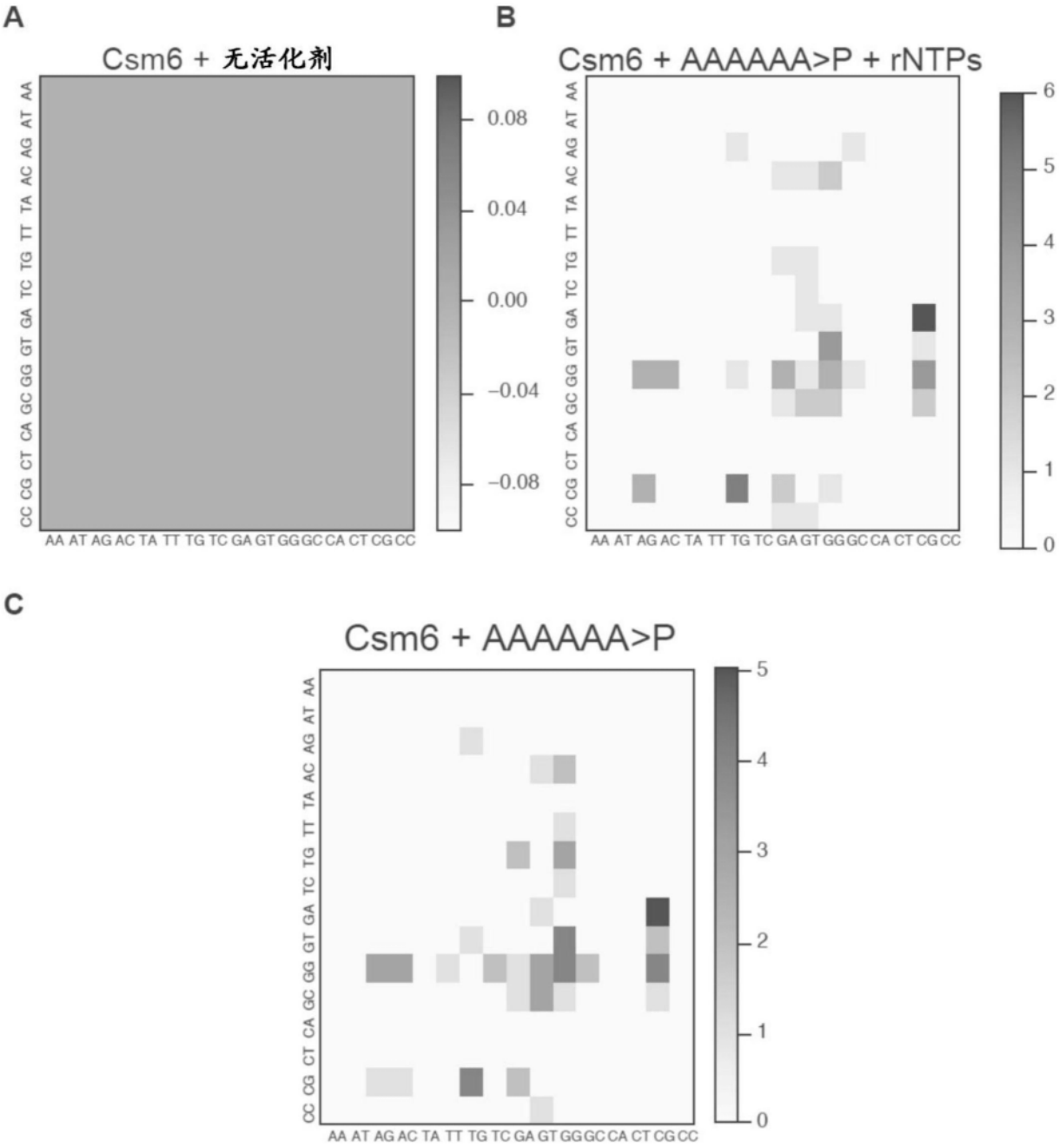

A
Csm6 + 无活化剂
B
Csm6 + AAAAAA>P + rNTPs
C
Csm6 + AAAAAA>P
AA AT AG AC TA TT TG TC GA GT GG GC CA CT CG CC
0.08
0.04
0.00
−0.04
−0.08
6
5
4
3
2
1
0

图14

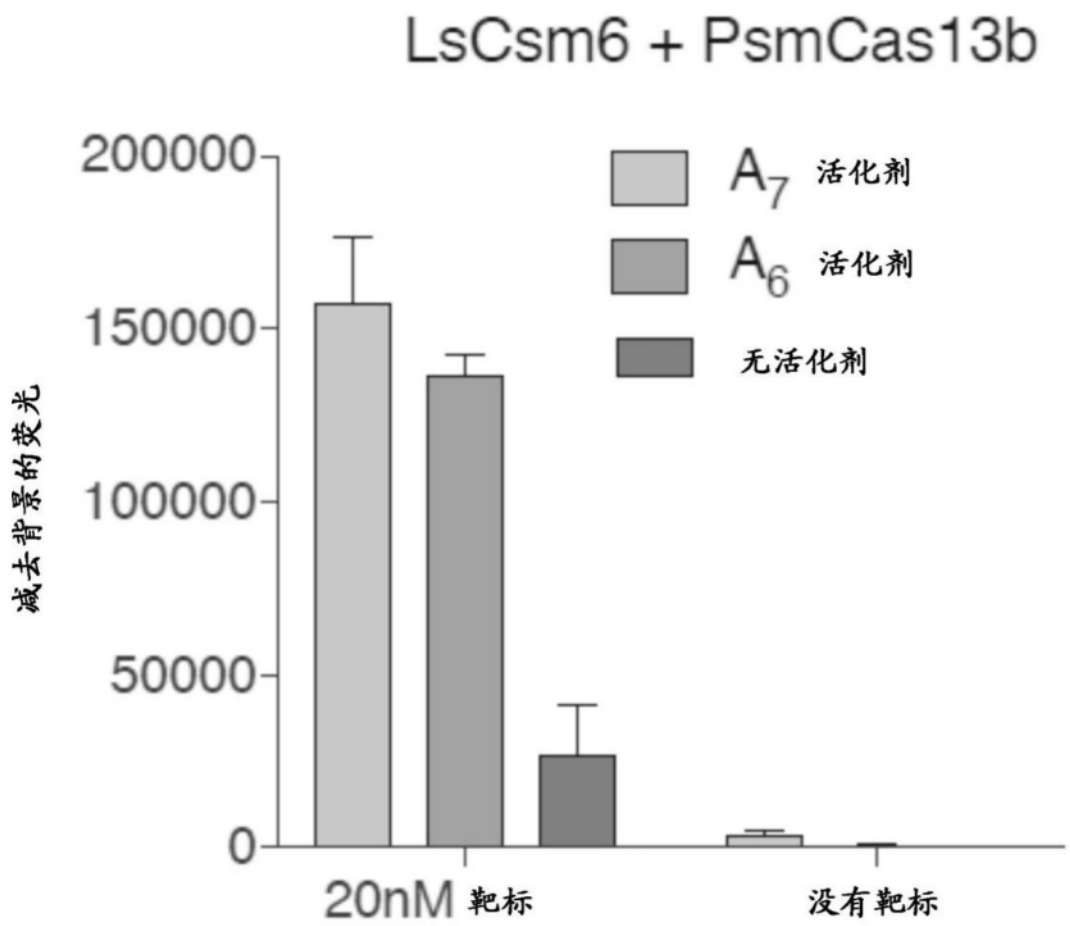
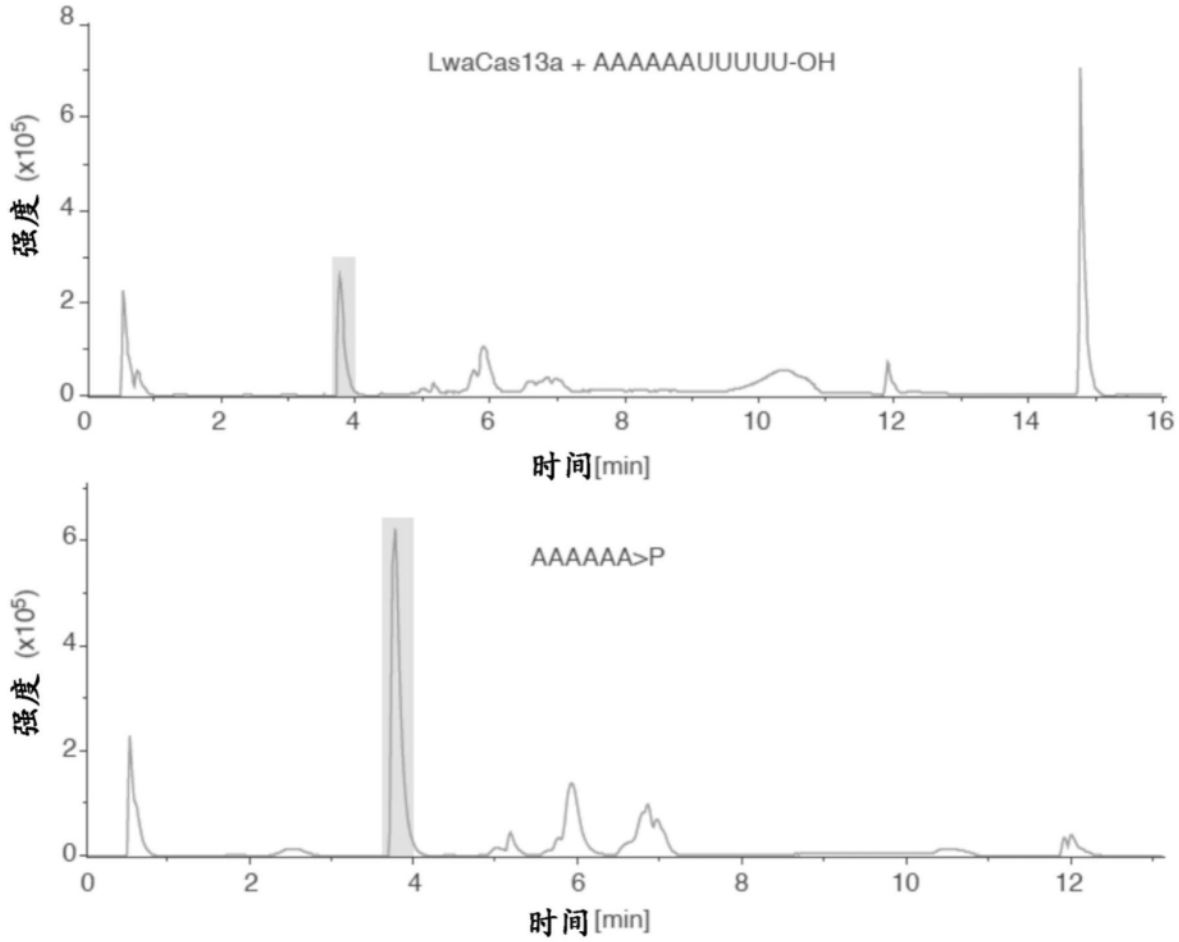

LsCsm6 + PsmCas13b
200000
150000
100000
50000
0
减去背景的荧光
A7 活化剂
A6 活化剂
无活化剂
20nM 靶标
没有靶标

图15

A



B

LwaCas13a + AAAAAAUUUUU-OH

Cmpt.	质量	分子	丰度	Abd.[%]	Std.Dev.
A	1973.306	[M -H]-	196313	100	0.014056
B	2279.331	[M -H]-	107038	54.52	0.003627
C	2585.356	[M -H]-	62759	31.97	0.005015
D	2891.381	[M -H]-	32484	16.55	0.006336
E	3197.406	[M -H]-	19313	9.84	0.007767
F	3441.476	[M -H]-	14841	7.56	0.00678
G	1995.287	[M -H]-	13024	6.63	0.004689
H	1644.254	[M -H]-	11876	6.05	0.001244
I	2071.281	[M -H]-	11697	5.96	0.005988
J	2377.307	[M -H]-	10195	5.19	0.002957
K	2301.313	[M -H]-	7095	3.61	0.003939
L	2683.332	[M -H]-	7360	3.75	0.008262
M	2607.338	[M -H]-	5076	2.59	0.00323

AAAAAA>P

Cmpt.	质量	分子	丰度	Abd.[%]	Std.Dev.
A	1973.306	[M -H]-	439056	100	0.003272
B	1644.254	[M -H]-	65030	14.81	0.001781
C	2063.374	[M -H]-	49525	11.28	0.003049
D	1995.287	[M -H]-	29564	6.73	0.003413
E	2071.282	[M -H]-	27554	6.28	0.003287

图16

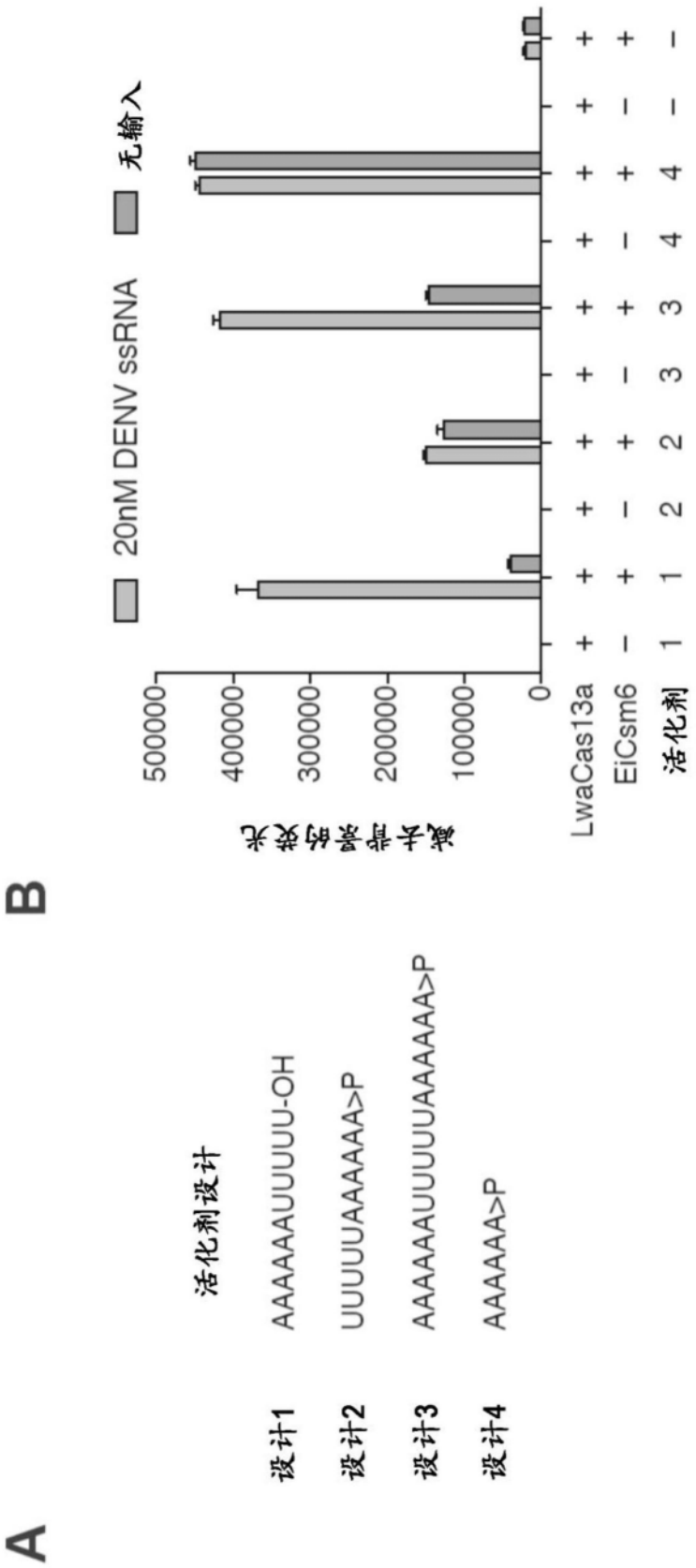

A
活化剂设计
设计1 AAAAAAUUUUU-OH
设计2 UUUUUAAAAAA>P
设计3 AAAAAAUUUUUAAAAAA>P
设计4 AAAAAA>P
B
20nM DENV ssRNA
无输入
荧光背景校正读数
500000
400000
300000
200000
100000
0
LwaCas13a + + + + + + + + + +
EiCsm6 − + − + − + − + − +
活化剂 1 1 2 2 3 3 4 4 − −

图17


A
20μM聚A/2μM 聚U传感器
2pM 酰基转移酶
无输入
减去背景的荧光
4000000
3000000
2000000
1000000
0
LwaCas13a
LwaCas13a + 冻化剂
EiCsm6 + LwaCas13a
EiCsm6 + LwaCas13a + 冻化剂
8μM聚A/2μM 聚U传感器
2pM 酰基转移酶
无输入
减去背景的荧光
LwaCas13a
LwaCas13a + 冻化剂
EiCsm6 + LwaCas13a
EiCsm6 + LwaCas13a + 冻化剂
4μM聚A/2μM 聚U传感器
2pM 酰基转移酶
无输入
减去背景的荧光
LwaCas13a
LwaCas13a + 冻化剂
EiCsm6 + LwaCas13a
EiCsm6 + LwaCas13a + 冻化剂
B
2/4μM
RnaseAlert/聚A
LwaCas13
LwaCas13 + EiCsm6
减去背景的荧光
3000000
2000000
1000000
0
0μM
10μM
20μM
50μM
活化剂浓度
2/8μM
RnaseAlert/聚A
LwaCas13
LwaCas13 + EiCsm6
减去背景的荧光
活化剂浓度
2/20μM
RnaseAlert/聚A
LwaCas13
LwaCas13 + EiCsm6
减去背景的荧光
活化剂浓度

图18

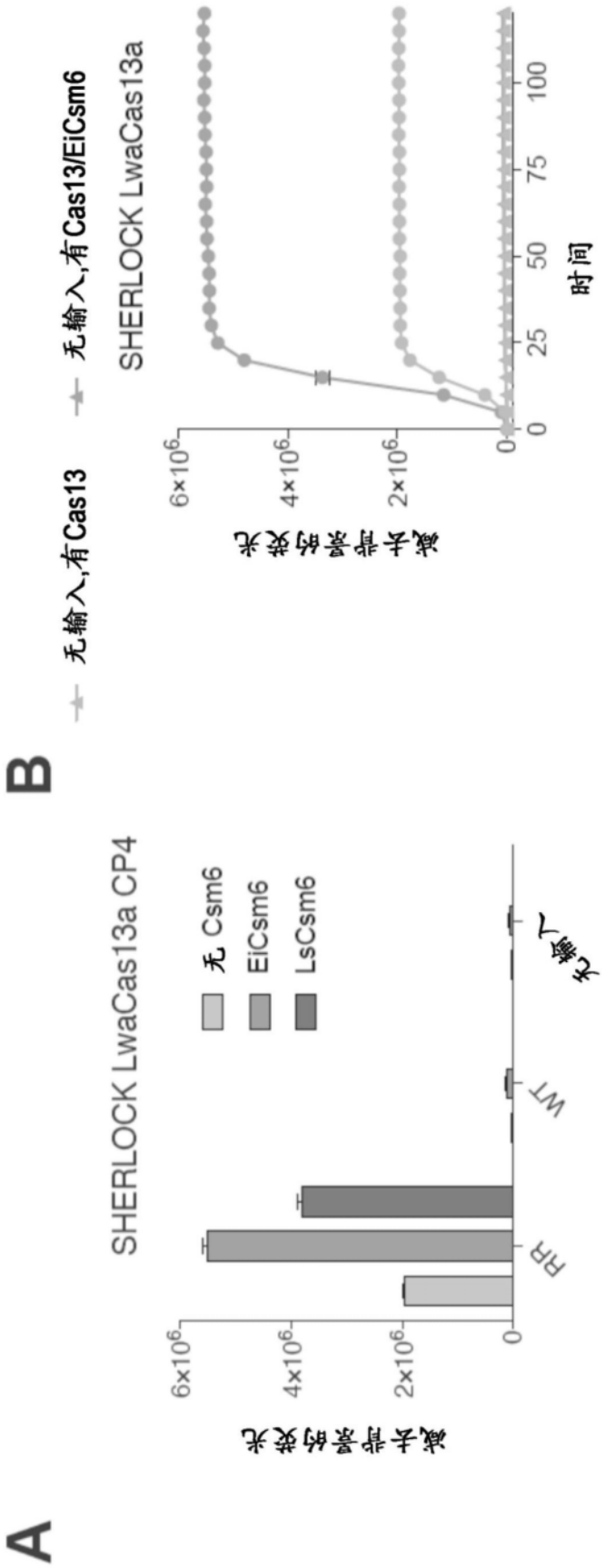

A
SHERLOCK LwaCas13a CP4
无 Csm6
EiCsm6
LsCsm6
减去背景的荧光
6×10^6
4×10^6
2×10^6
0
RR
WT
无输入
B
SHERLOCK LwaCas13a
无输入,有Cas13
无输入,有Cas13/EiCsm6
减去背景的荧光
6×10^6
4×10^6
2×10^6
0
0
25
50
75
100
时间

图19

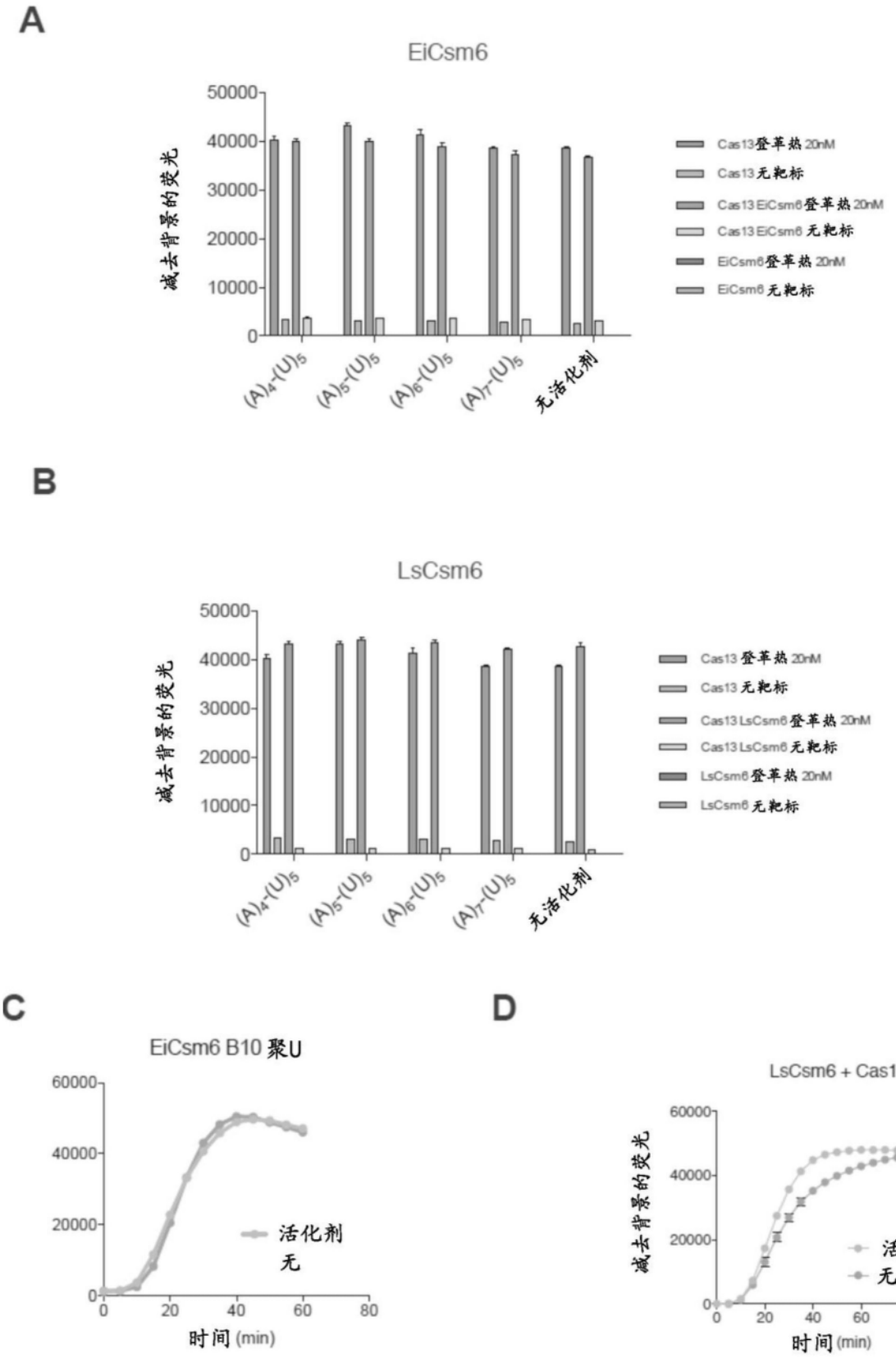

A
EiCsm6
减去背景的荧光
50000
40000
30000
20000
10000
0
(A)4-(U)5
(A)5-(U)5
(A)6-(U)5
(A)7-(U)5
无活化剂
Cas13登革热 20nM
Cas13 无靶标
Cas13 EiCsm6登革热 20nM
Cas13 EiCsm6 无靶标
EiCsm6登革热 20nM
EiCsm6 无靶标
B
LsCsm6
减去背景的荧光
50000
40000
30000
20000
10000
0
(A)4-(U)5
(A)5-(U)5
(A)6-(U)5
(A)7-(U)5
无活化剂
Cas13 登革热 20nM
Cas13 无靶标
Cas13 LsCsm6登革热 20nM
Cas13 LsCsm6 无靶标
LsCsm6登革热 20nM
LsCsm6 无靶标
C
EiCsm6 B10 聚U
60000
40000
20000
0
0
20
40
60
80
时间 (min)
活化剂
无
D
LsCsm6 + Cas13b10
减去背景的荧光
60000
40000
20000
0
0
20
40
60
80
100
时间 (min)
活化剂
无活化剂

图20

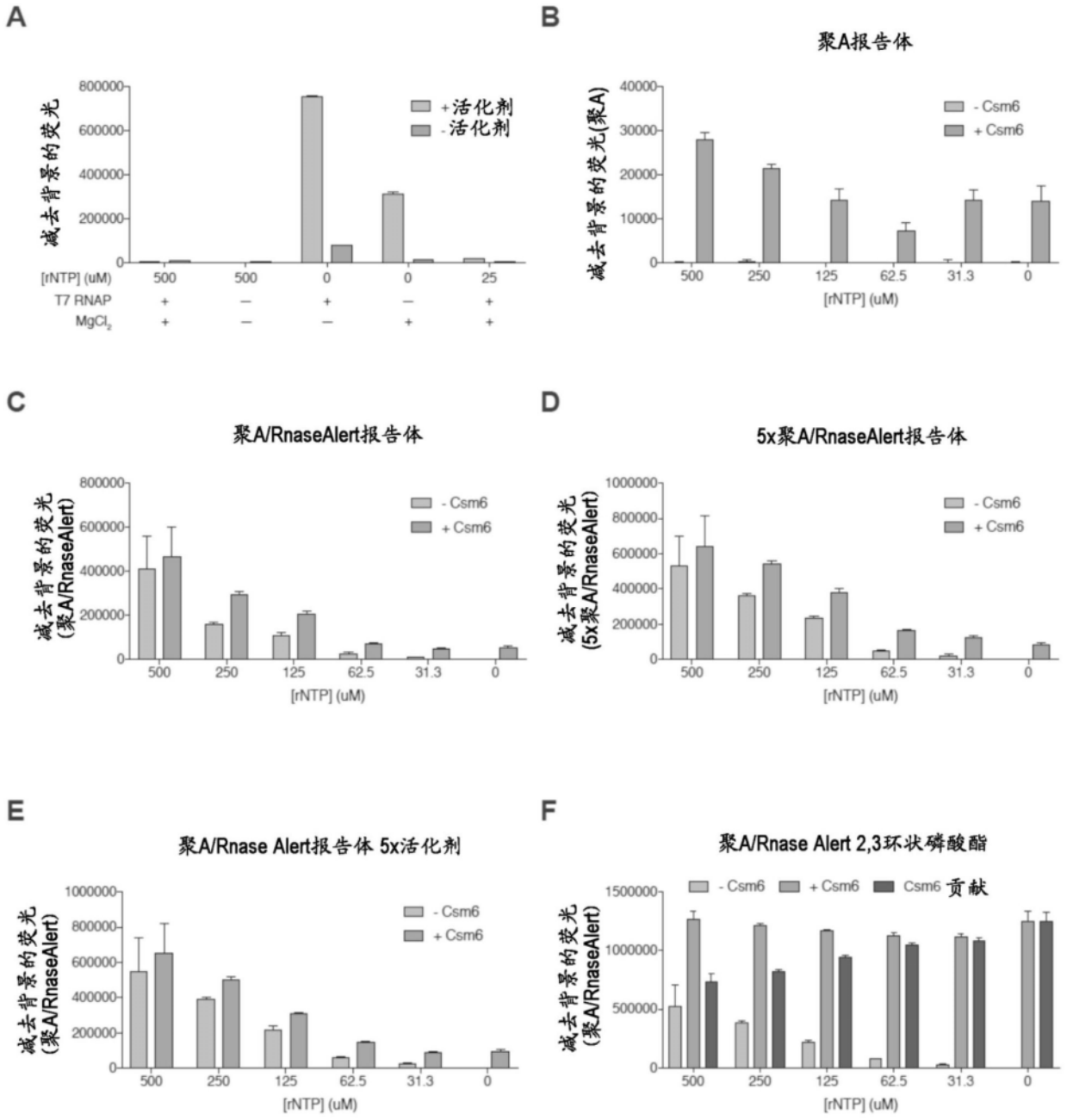

A
减去背景的荧光
800000
600000
400000
200000
0
+活化剂
-活化剂
[rNTP] (uM) 500 500 0 0 25
T7 RNAP + — + — +
MgCl2 + — — + +
B
聚A报告体
减去背景的荧光(聚A)
40000
30000
20000
10000
0
- Csm6
+ Csm6
500 250 125 62.5 31.3 0
[rNTP] (uM)
C
聚A/RnaseAlert报告体
减去背景的荧光(聚A/RnaseAlert)
800000
600000
400000
200000
0
- Csm6
+ Csm6
500 250 125 62.5 31.3 0
[rNTP] (uM)
D
5x聚A/RnaseAlert报告体
减去背景的荧光(5x聚A/RnaseAlert)
1000000
800000
600000
400000
200000
0
- Csm6
+ Csm6
500 250 125 62.5 31.3 0
[rNTP] (uM)
E
聚A/Rnase Alert报告体 5x活化剂
减去背景的荧光(聚A/RnaseAlert)
1000000
800000
600000
400000
200000
0
- Csm6
+ Csm6
500 250 125 62.5 31.3 0
[rNTP] (uM)
F
聚A/Rnase Alert 2,3环状磷酸酯
减去背景的荧光(聚A/RnaseAlert)
1500000
1000000
500000
0
- Csm6
+ Csm6
Csm6 贡献
500 250 125 62.5 31.3 0
[rNTP] (uM)

图21

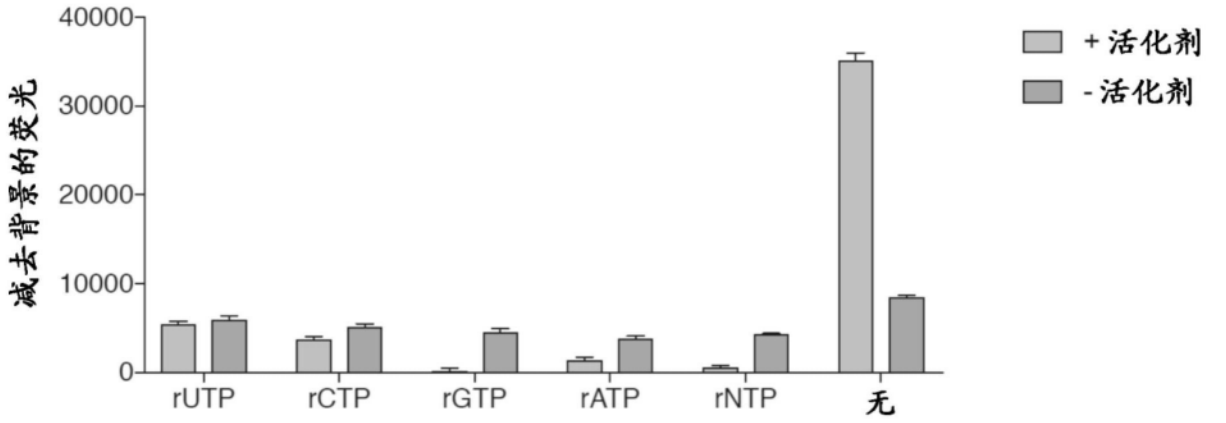

40000
30000
20000
10000
0
减去背景的荧光
rUTP
rCTP
rGTP
rATP
rNTP
无
+活化剂
-活化剂

图22

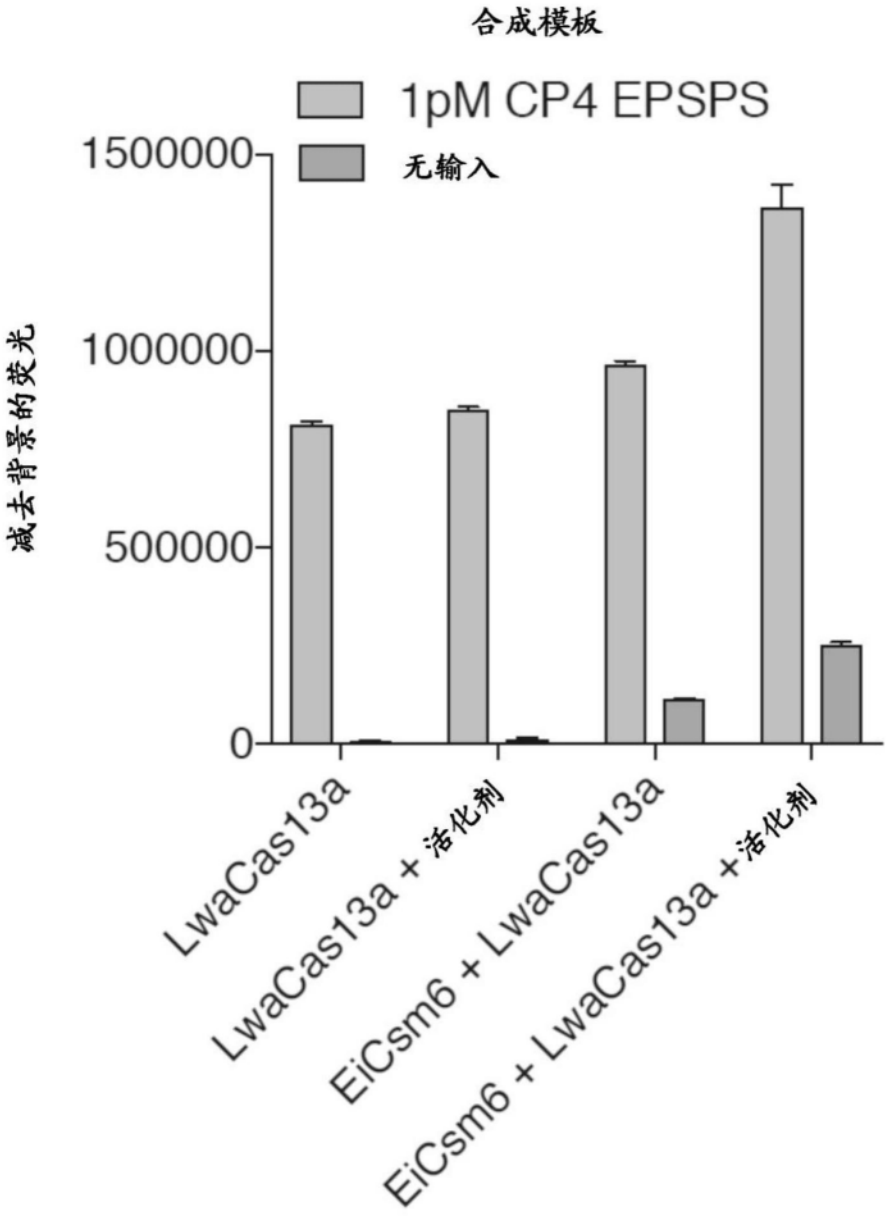

合成模板
1pM CP4 EPSPS
无输入
减去背景的荧光
1500000
1000000
500000
0
LwaCas13a
LwaCas13a + 活化剂
EiCsm6 + LwaCas13a
EiCsm6 + LwaCas13a +活化剂

图23

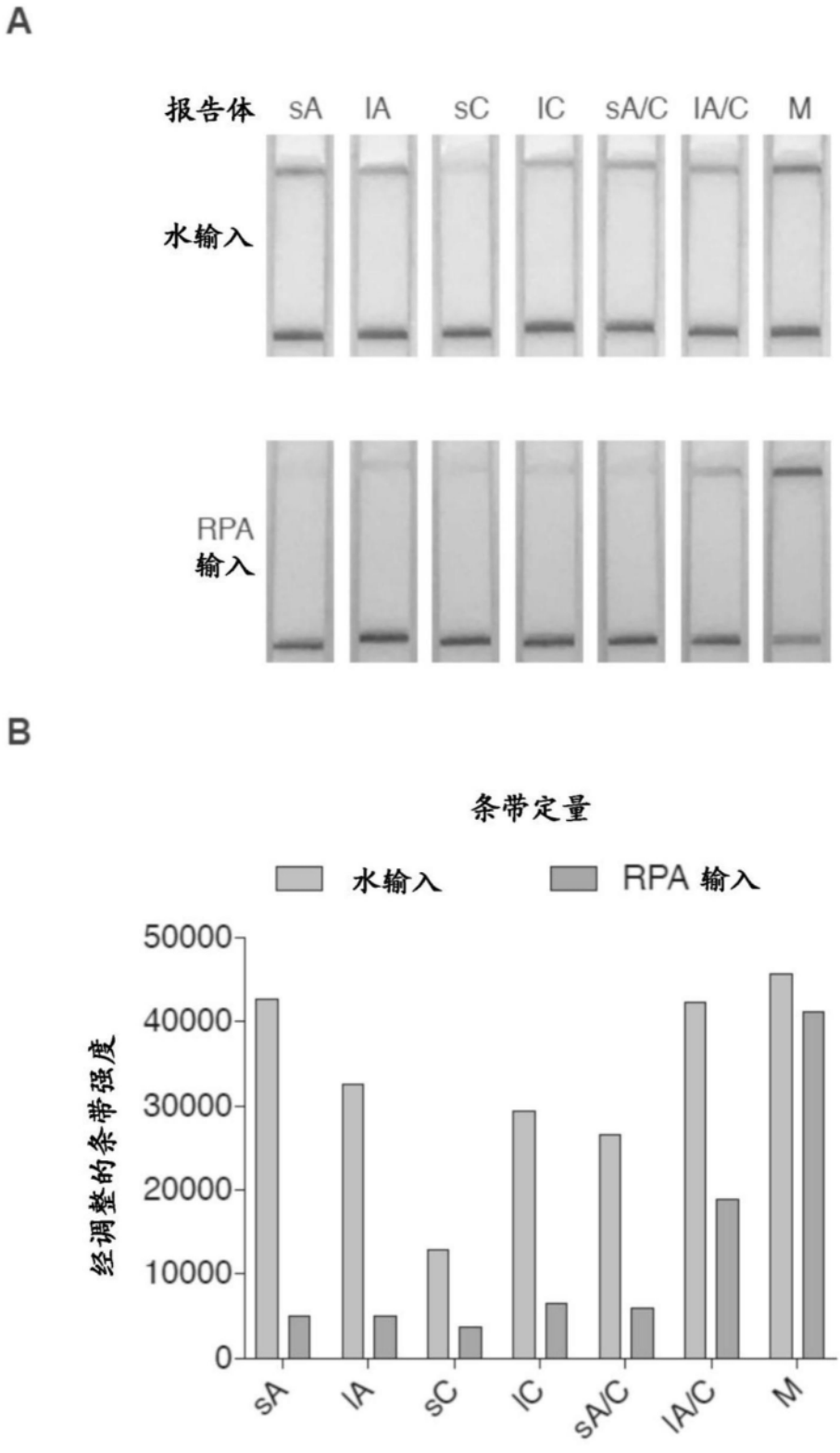

A
报告体
sA
IA
sC
IC
sA/C
IA/C
M
水输入
RPA
输入
B
条带定量
水输入
RPA 输入
经调整的条带强度
50000
40000
30000
20000
10000
0
sA
IA
sC
IC
sA/C
IA/C
M

图24

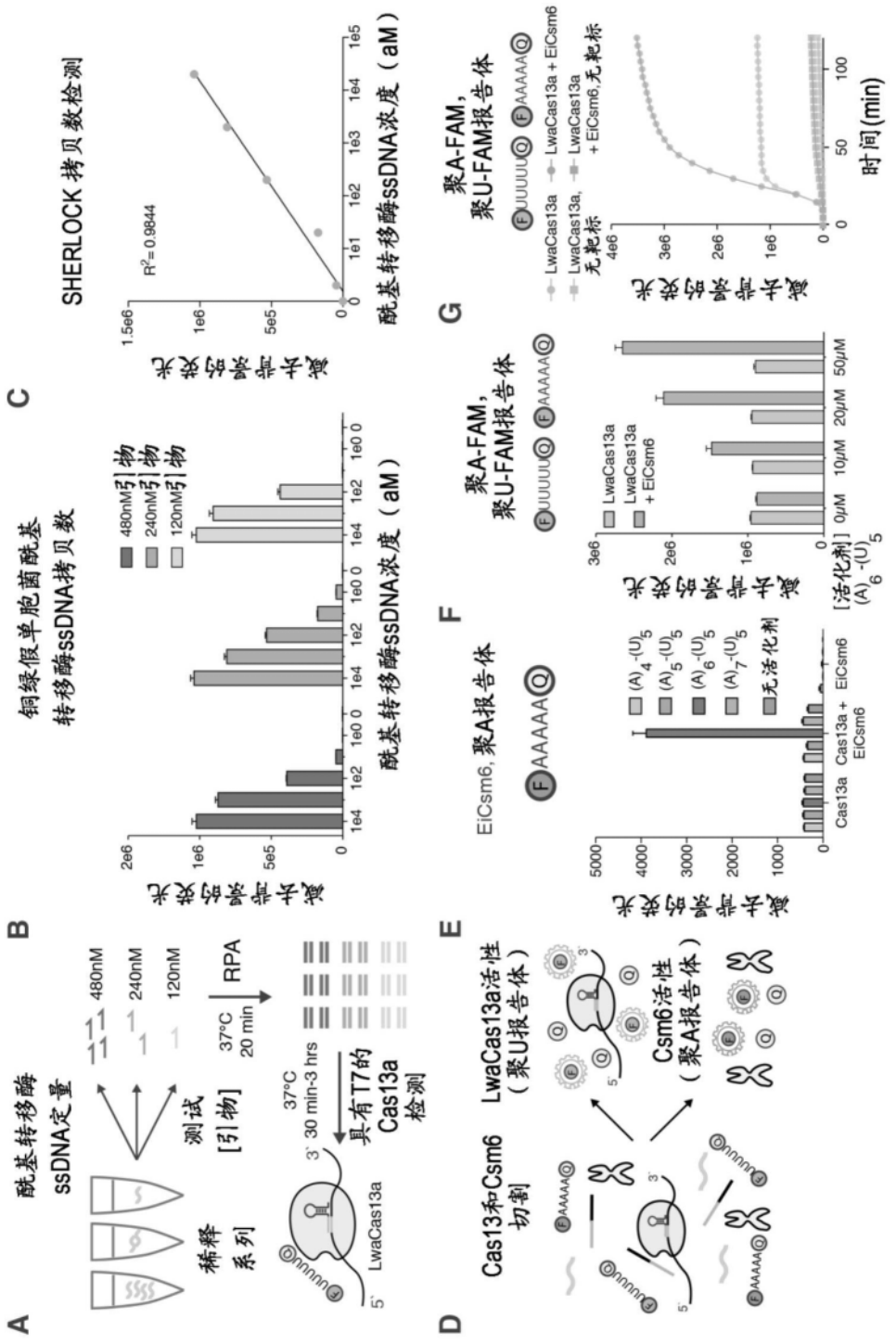

A
酰基转移酶
ssDNA定量
稀释
系列
480nM
240nM
120nM
测试
[引物]
37°C
20 min
RPA
37°C
30 min-3 hrs
具有T7的
Cas13a
检测
LwaCas13a
B
铜绿假单胞菌酰基
转移酶ssDNA拷贝数
480nM引物
240nM引物
120nM引物
酰基转移酶ssDNA浓度（aM）
C
SHERLOCK 拷贝数检测
R²= 0.9844
酰基转移酶ssDNA浓度（aM）
D
Cas13和Csm6
切割
LwaCas13a活性
（聚U报告体）
Csm6活性
（聚A报告体）
E
EiCsm6,
聚A报告体
Cas13a
Cas13a + EiCsm6
EiCsm6
无活化剂
F
聚A-FAM,
聚U-FAM报告体
LwaCas13a
LwaCas13a + EiCsm6
[活化剂]
0μM
10μM
20μM
50μM
G
聚A-FAM,
聚U-FAM报告体
LwaCas13a
LwaCas13a + EiCsm6
LwaCas13a,
无靶标
LwaCas13a
+ EiCsm6,无靶标
时间(min)

图25

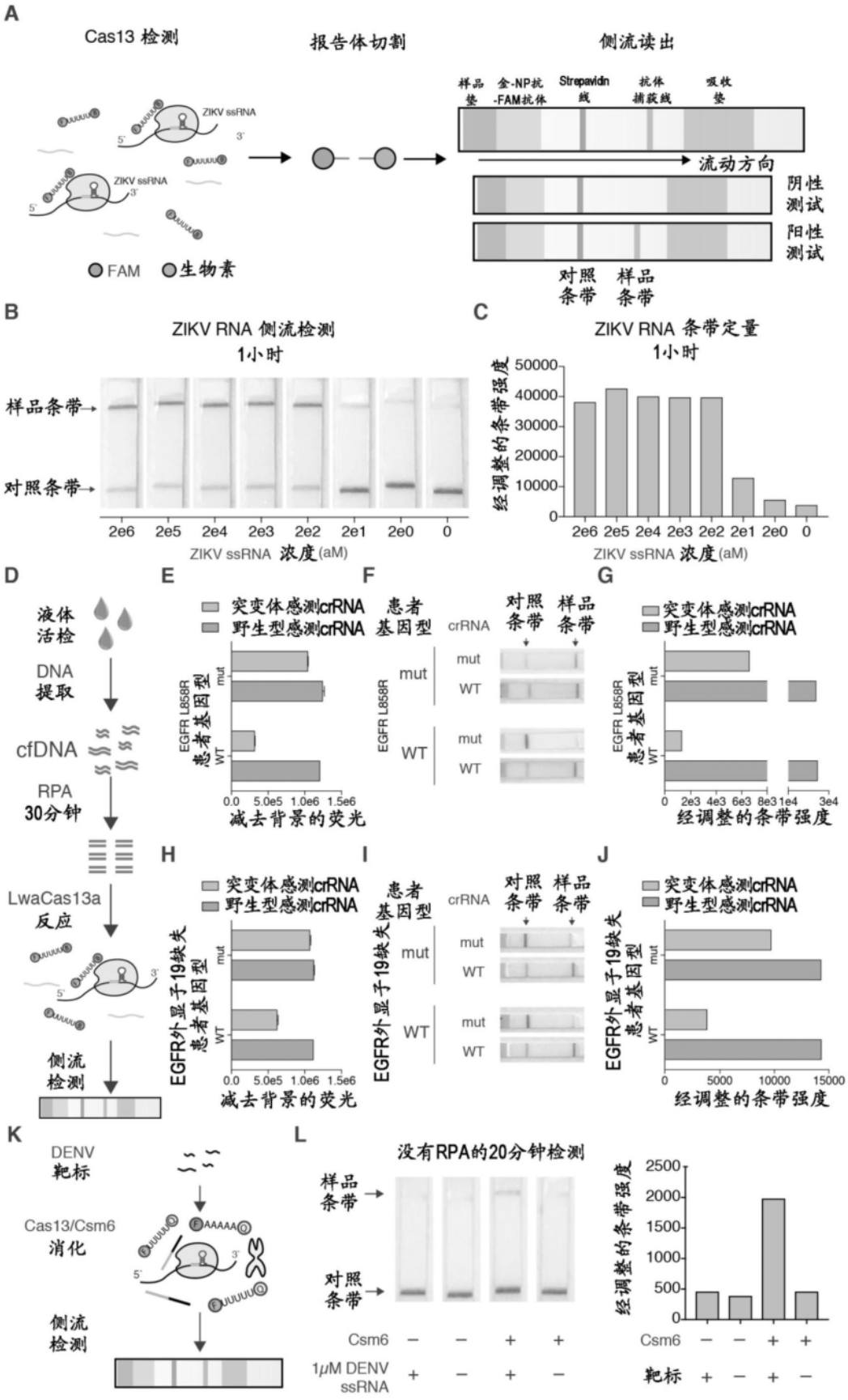

A
Cas13 检测
报告体切割
侧流读出
ZIKV ssRNA
FAM
生物素
样品垫
金-NP抗-FAM抗体
Streptavidin线
抗体捕获线
吸收垫
流动方向
阴性测试
阳性测试
对照条带
样品条带
B
ZIKV RNA 侧流检测
1小时
样品条带
对照条带
2e6 2e5 2e4 2e3 2e2 2e1 2e0 0
ZIKV ssRNA 浓度(aM)
C
ZIKV RNA 条带定量
1小时
经调整的条带强度
50000
40000
30000
20000
10000
0
2e6 2e5 2e4 2e3 2e2 2e1 2e0 0
ZIKV ssRNA 浓度(aM)
D
液体活检
DNA提取
cfDNA
RPA
30分钟
LwaCas13a反应
侧流检测
E
突变体感测crRNA
野生型感测crRNA
EGFR L858R患者基因型
mut
WT
0.0 5.0e5 1.0e6 1.5e6
减去背景的荧光
F
患者基因型
crRNA
对照条带
样品条带
mut
WT
G
突变体感测crRNA
野生型感测crRNA
EGFR L858R患者基因型
0 2e3 4e3 6e3 8e3 1e4 3e4
经调整的条带强度
H
突变体感测crRNA
野生型感测crRNA
EGFR外显子19缺失患者基因型
0.0 5.0e5 1.0e6 1.5e6
减去背景的荧光
I
患者基因型
crRNA
对照条带
样品条带
EGFR外显子19缺失
J
突变体感测crRNA
野生型感测crRNA
EGFR外显子19缺失患者基因型
0 5000 10000 15000
经调整的条带强度
K
DENV靶标
Cas13/Csm6消化
侧流检测
L
没有RPA的20分钟检测
样品条带
对照条带
Csm6 − − + +
1μM DENV ssRNA + − + −
经调整的条带强度
2500
2000
1500
1000
500
0
Csm6 − − + +
靶标 + − + −

图26

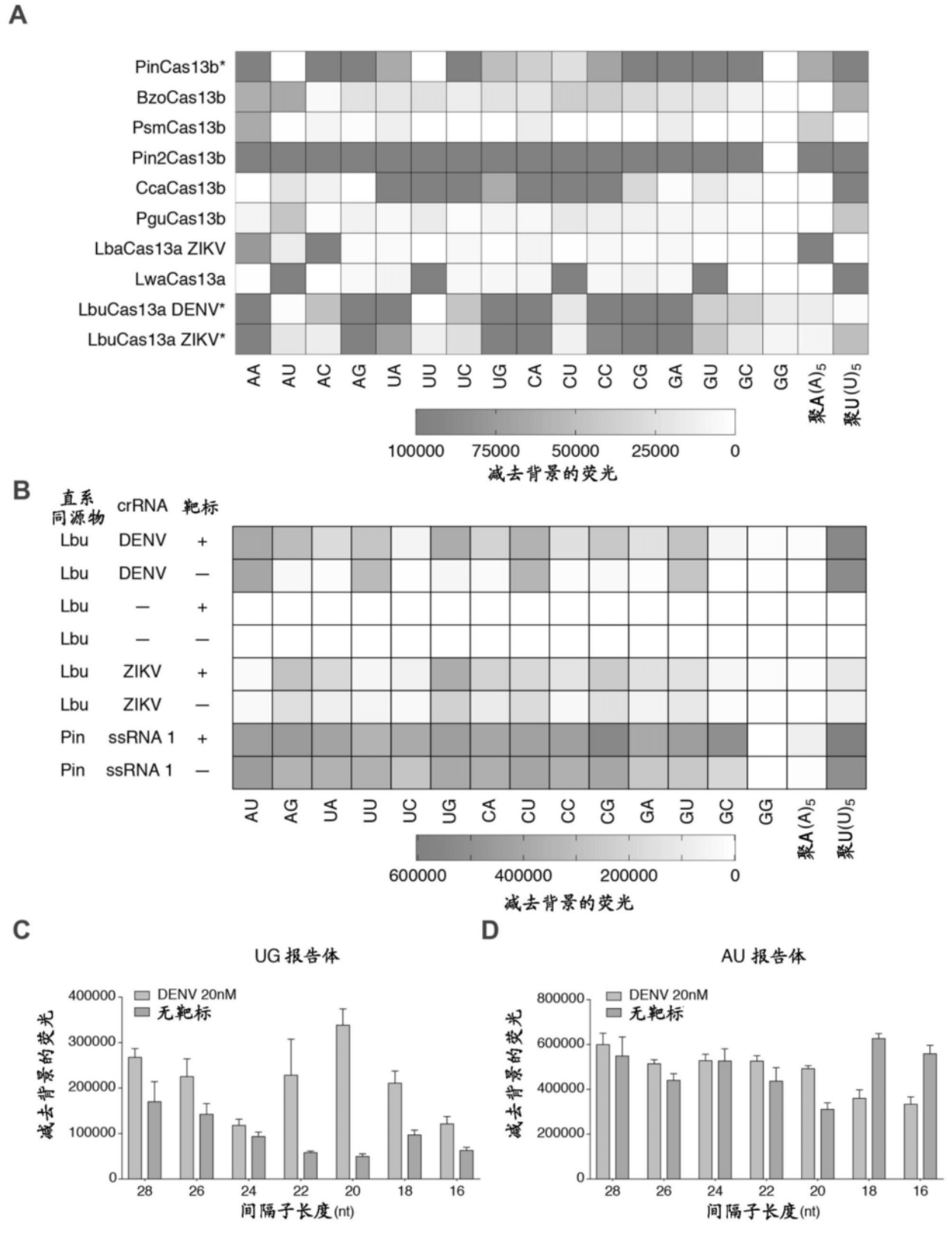

A
PinCas13b*
BzoCas13b
PsmCas13b
Pin2Cas13b
CcaCas13b
PguCas13b
LbaCas13a ZIKV
LwaCas13a
LbuCas13a DENV*
LbuCas13a ZIKV*
AA AU AC AG UA UU UC UG CA CU CC CG GA GU GC GG 聚A(A)5 聚U(U)5
100000 75000 50000 25000 0
减去背景的荧光
B
直系同源物 crRNA 靶标
Lbu DENV +
Lbu DENV —
Lbu — +
Lbu — —
Lbu ZIKV +
Lbu ZIKV —
Pin ssRNA 1 +
Pin ssRNA 1 —
AU AG UA UU UC UG CA CU CC CG GA GU GC GG 聚A(A)5 聚U(U)5
600000 400000 200000 0
减去背景的荧光
C
UG 报告体
DENV 20nM
无靶标
减去背景的荧光
400000 300000 200000 100000 0
28 26 24 22 20 18 16
间隔子长度(nt)
D
AU 报告体
DENV 20nM
无靶标
减去背景的荧光
800000 600000 400000 200000 0
28 26 24 22 20 18 16
间隔子长度(nt)

图27

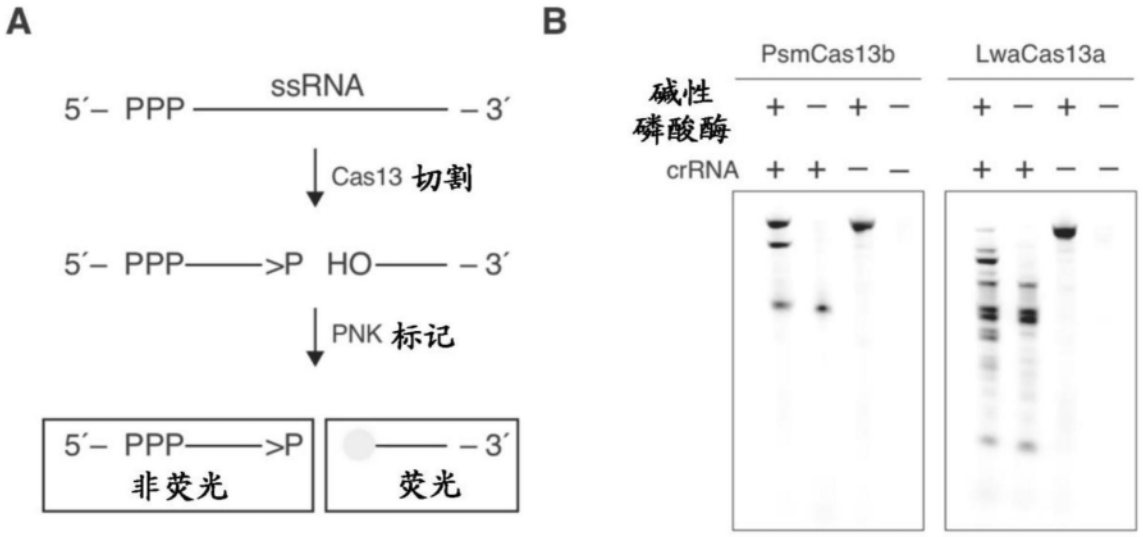

A
5′– PPP
ssRNA
– 3′
Cas13 切割
5′– PPP——>P HO—— – 3′
PNK 标记
5′– PPP——>P
非荧光
——– 3′
荧光
B
PsmCas13b
LwaCas13a
碱性磷酸酶
crRNA
+ – + – + – + –
+ + – – + + – –

图28

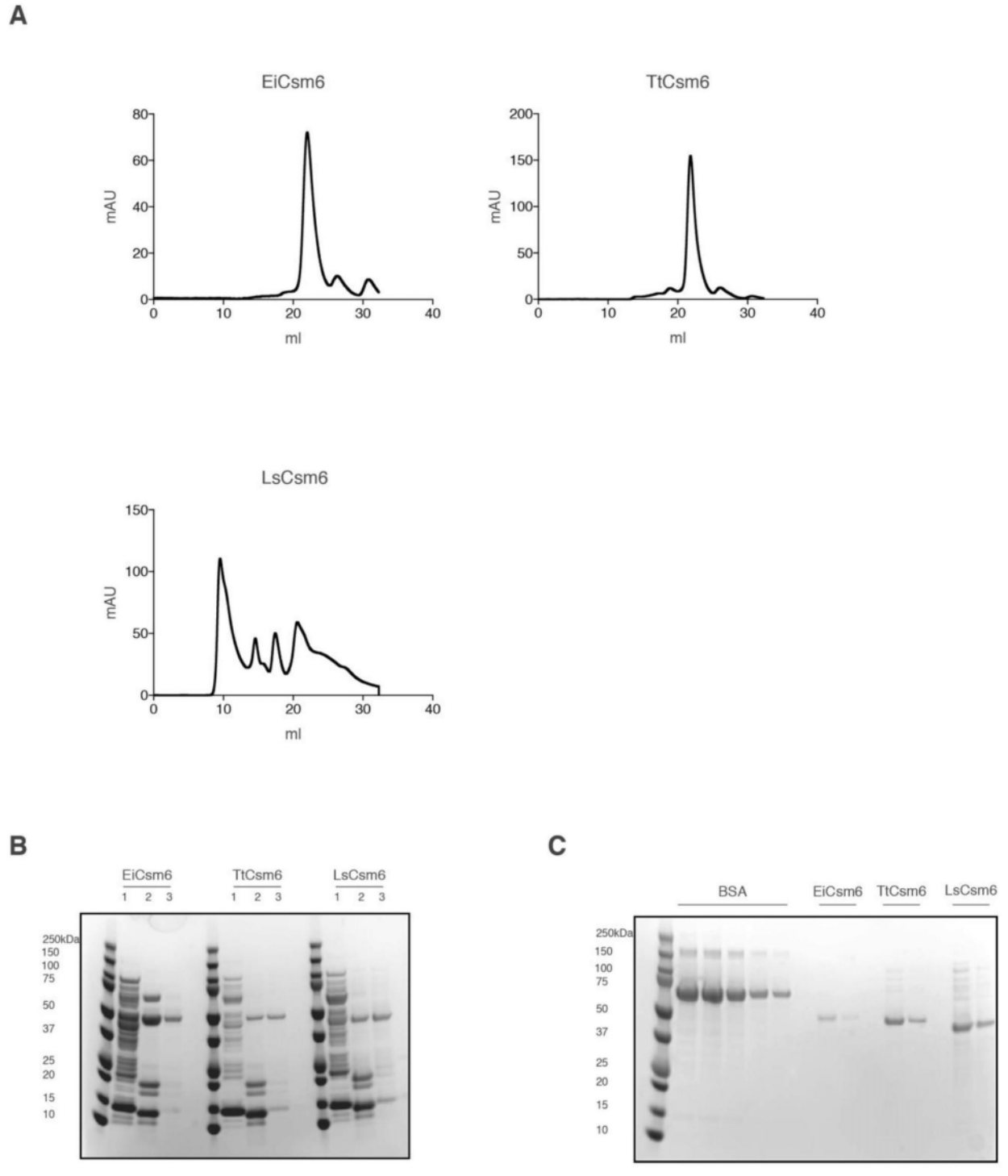

A
EiCsm6
mAU
ml
TtCsm6
mAU
ml
LsCsm6
mAU
ml
B
EiCsm6
TtCsm6
LsCsm6
1 2 3
250kDa
150
100
75
50
37
25
20
15
10
C
BSA
EiCsm6
TtCsm6
LsCsm6
250kDa
150
100
75
50
37
25
20
15
10

图29

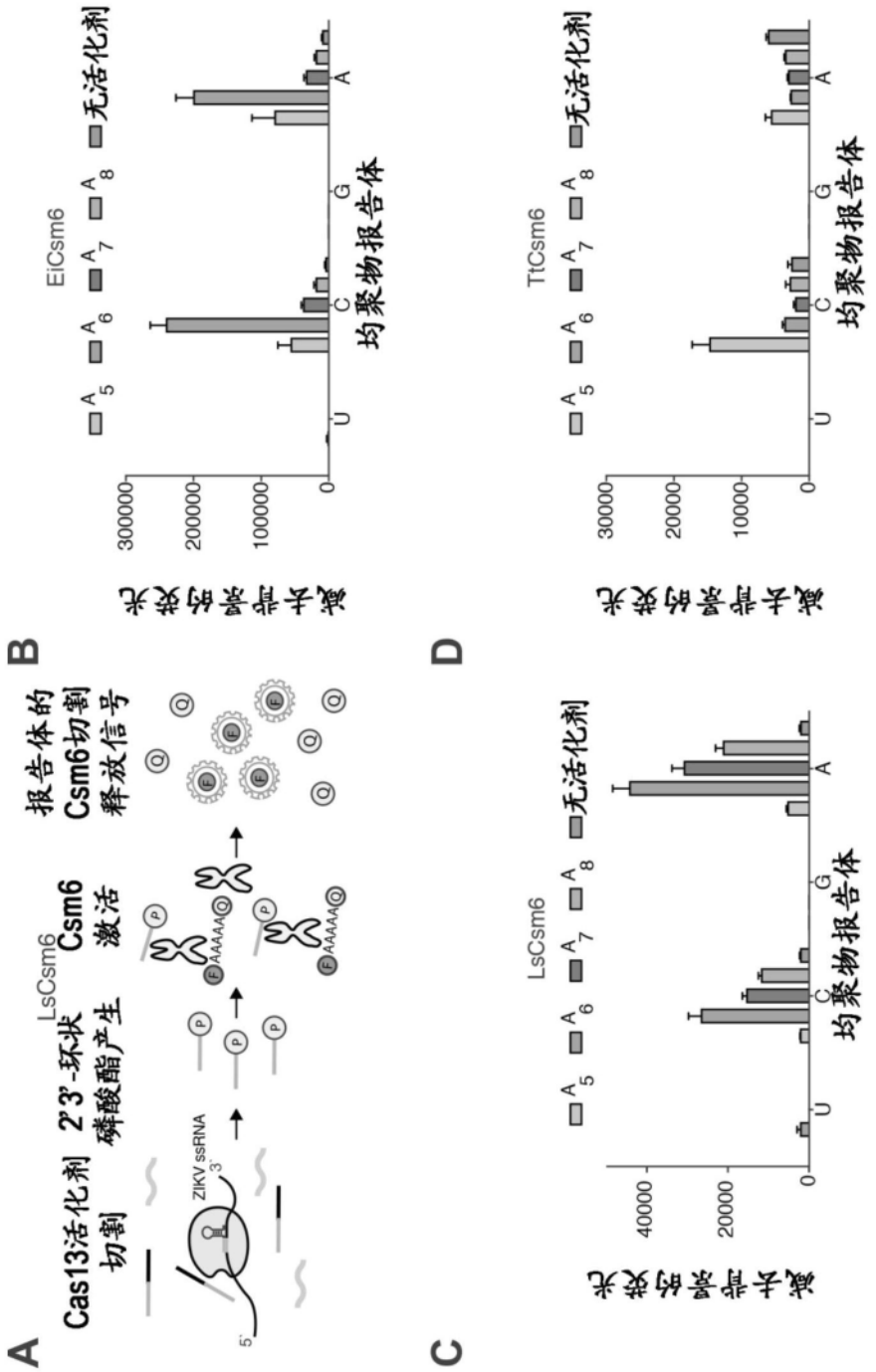

A
Cas13活化剂
切割
2'3'-环状
磷酸酯产生
LsCsm6
Csm6
激活
ZIKV ssRNA
报告体的
Csm6切割
释放信号
B
EiCsm6
无活化剂
均聚物报告体
C
LsCsm6
无活化剂
均聚物报告体
D
TtCsm6
无活化剂
均聚物报告体

图30

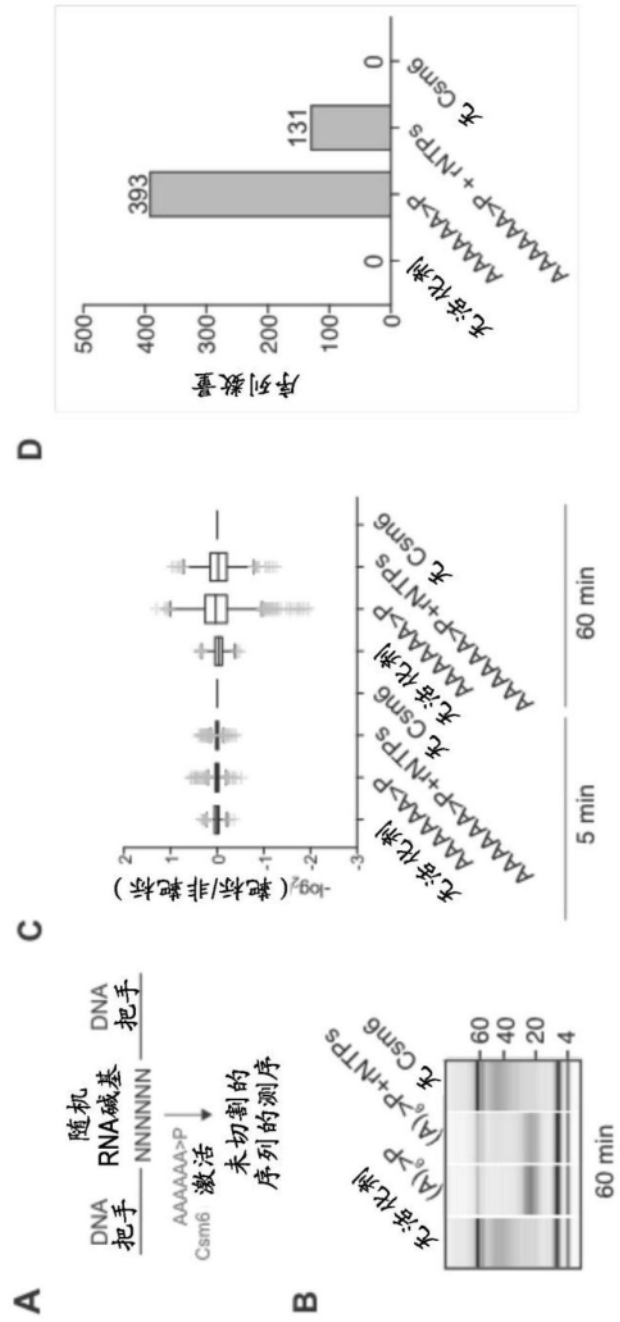

A
DNA 把手
随机 RNA碱基 NNNNNN
DNA 把手
AAAAAA>P
Csm6 激活
未切割的序列的测序
B
靶标
非靶标
Csm6
AAAAAA>P
AAAAAA>P+rNTPs
5 min
60 min
-log2(靶标/非靶标)
C
D
事件数量
393
131
0
500
400
300
200
100
0

图31

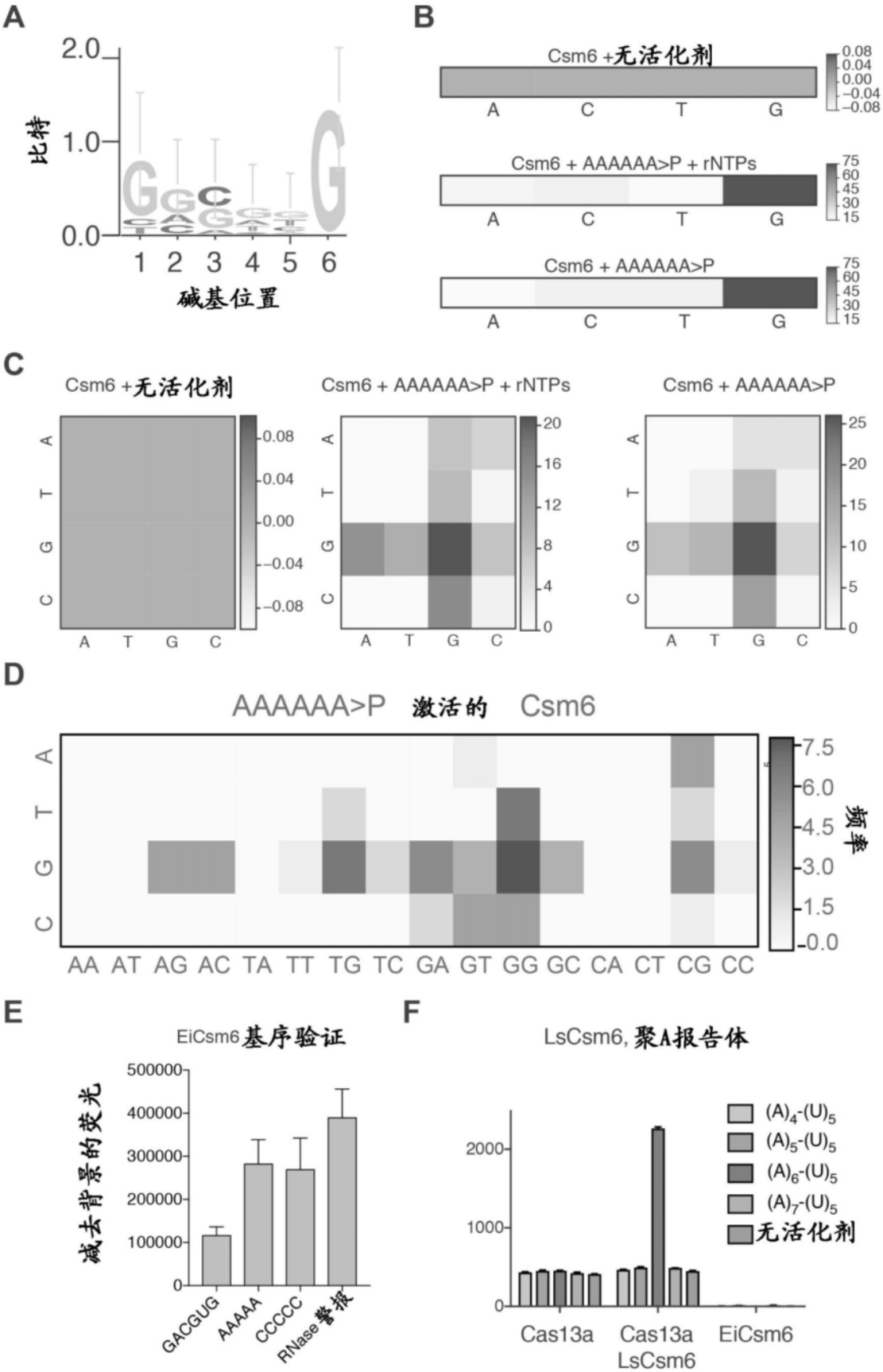
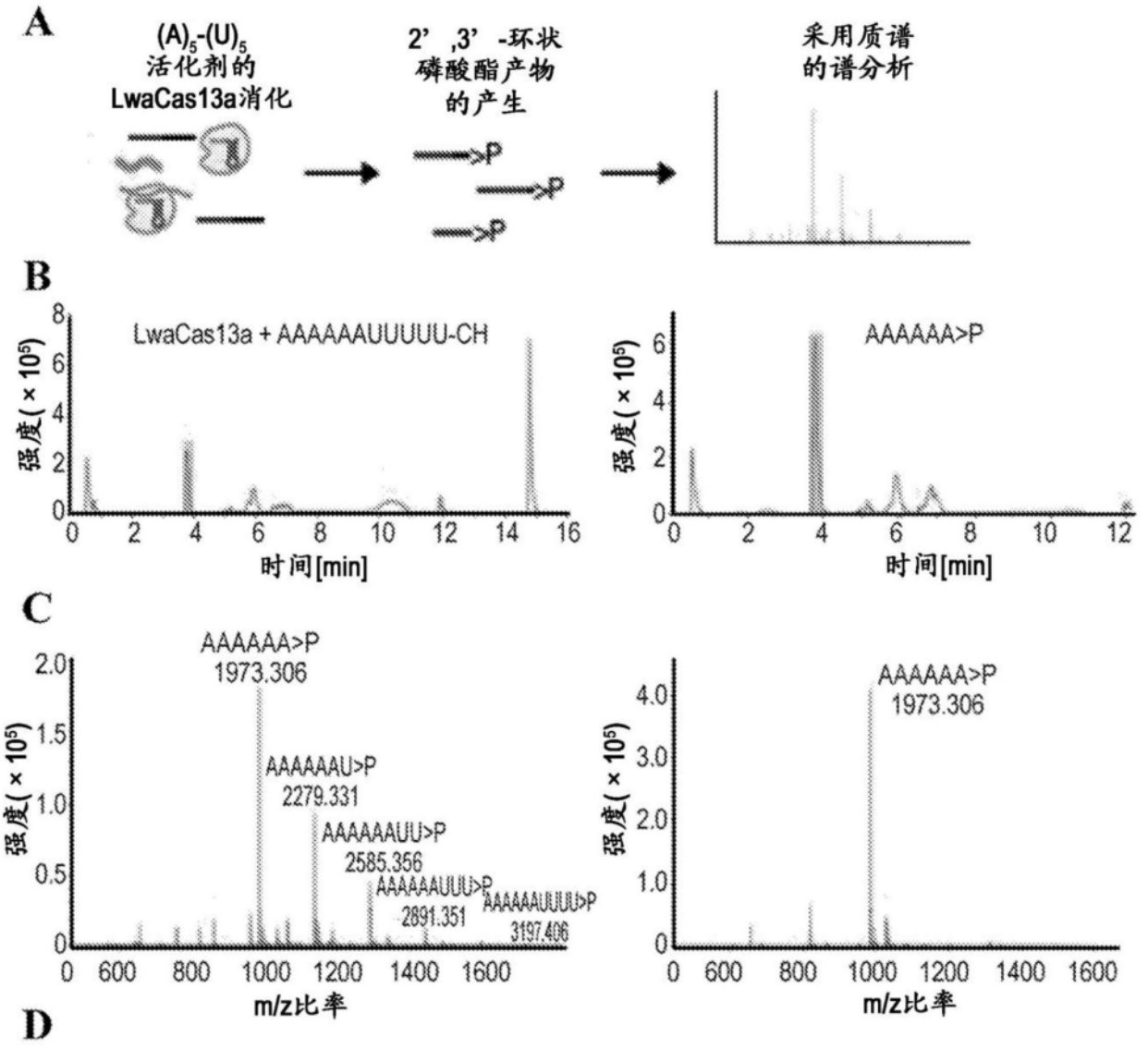

A
2.0
1.0
0.0
比特
1 2 3 4 5 6
碱基位置
B
Csm6 +无活化剂
0.08
0.04
0.00
−0.04
−0.08
A C T G
Csm6 + AAAAAA>P + rNTPs
75
60
45
30
15
A C T G
Csm6 + AAAAAA>P
75
60
45
30
15
A C T G
C
Csm6 +无活化剂
Csm6 + AAAAAA>P + rNTPs
Csm6 + AAAAAA>P
A T G C
0.08
0.04
0.00
−0.04
−0.08
20
16
12
8
4
0
25
20
15
10
5
0
D
AAAAAA>P 激活的 Csm6
A T G C
AA AT AG AC TA TT TG TC GA GT GG GC CA CT CG CC
7.5
6.0
4.5
3.0
1.5
0.0
频率
E
EiCsm6基序验证
减去背景的荧光
500000
400000
300000
200000
100000
0
GACGUG
AAAAA
CCCCC
RNase警报
F
LsCsm6, 聚A报告体
2000
1000
0
(A)4-(U)5
(A)5-(U)5
(A)6-(U)5
(A)7-(U)5
无活化剂
Cas13a
Cas13a LsCsm6
EiCsm6

图32



LwaCas13a + AAAAAAUUUUU-OH

Cmpt.	质量	分子	丰度	Abd.[%]	Std.Dev.
A	1973.305	[M-H]-	196313	100	0.014056
B	2279.331	[M-H]-	107038	54.52	0.003627
C	2585.356	[M-H]-	62759	31.97	0.005015
D	2591.381	[M-H]-	32484	18.55	0.005335
E	3197.405	[M-H]-	19313	9.84	0.007767
F	3441.475	[M-H]-	14841	7.55	0.00678
G	1995.287	[M-H]-	13024	6.53	0.004689
H	1644.254	[M-H]-	11876	6.05	0.001244
I	2071.281	[M-H]-	11897	5.98	0.005988
J	2377.307	[M-H]-	10185	5.19	0.002857
K	2301.313	[M-H]-	7095	3.61	0.003839
L	2583.332	[M-H]-	7380	3.75	0.008262
M	2507.338	[M-H]-	5076	2.59	0.00323

AAAAAA>P

Cmpt.	质量	分子	丰度	Abd.[%]	Std.Dev.
A	1973.305	[M-H]-	439356	100	0.003272
B	1644.254	[M-H]-	65030	14.81	0.001781
C	2063.374	[M-H]-	49525	11.28	0.003049
D	1995.287	[M-H]-	29584	6.73	0.003413
E	2071.282	[M-H]-	27554	6.28	0.003257

图33

A

活化剂设计	
设计 1	AAAAAAUUUUU-OH
设计 2	UUUUUAAAAAA>P
设计 3	AAAAAAUUUUUAAAAAA>P
设计 4	AAAAAA>P

B

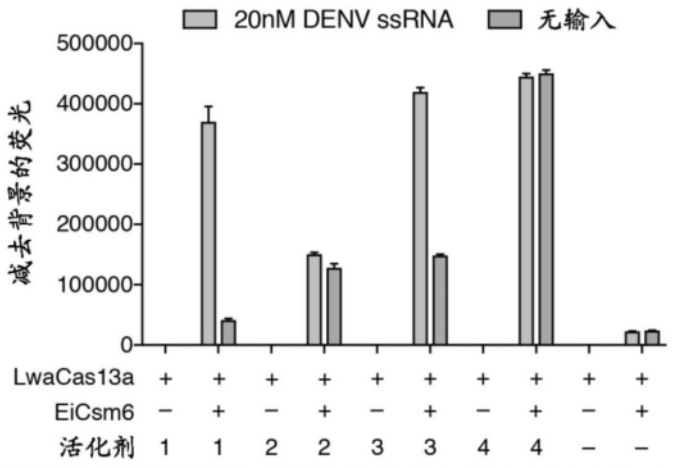


图34

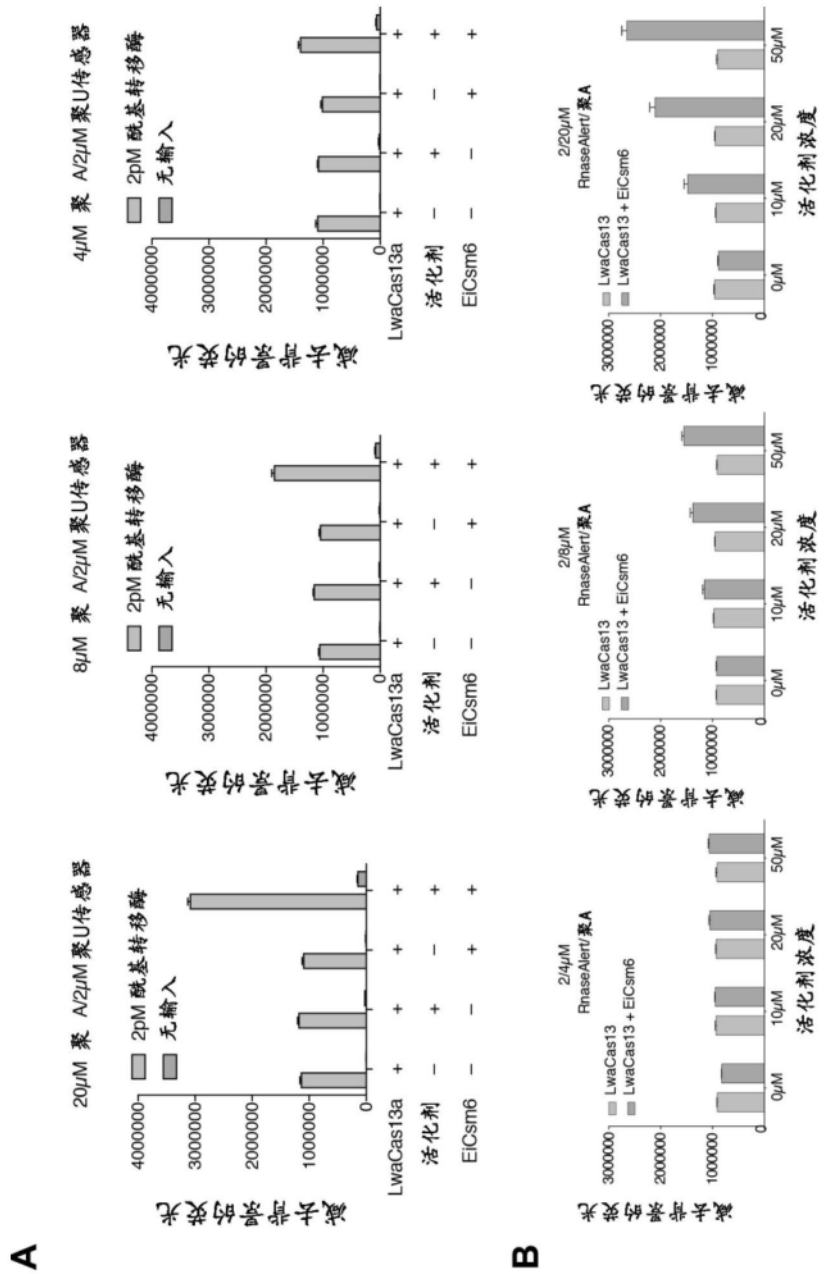

A
20μM 聚 A/2μM 聚U传感器
4μM 聚 A/2μM 聚U传感器
8μM 聚 A/2μM 聚U传感器
2pM 酰基转移酶
无输入
减去背景的荧光
4000000
3000000
2000000
1000000
0
LwaCas13a
EiCsm6
活化剂
B
2/4μM
2/8μM
2/20μM
RnaseAlert/聚A
LwaCas13
LwaCas13 + EiCsm6
0μM
10μM
20μM
50μM
活化剂浓度

图35

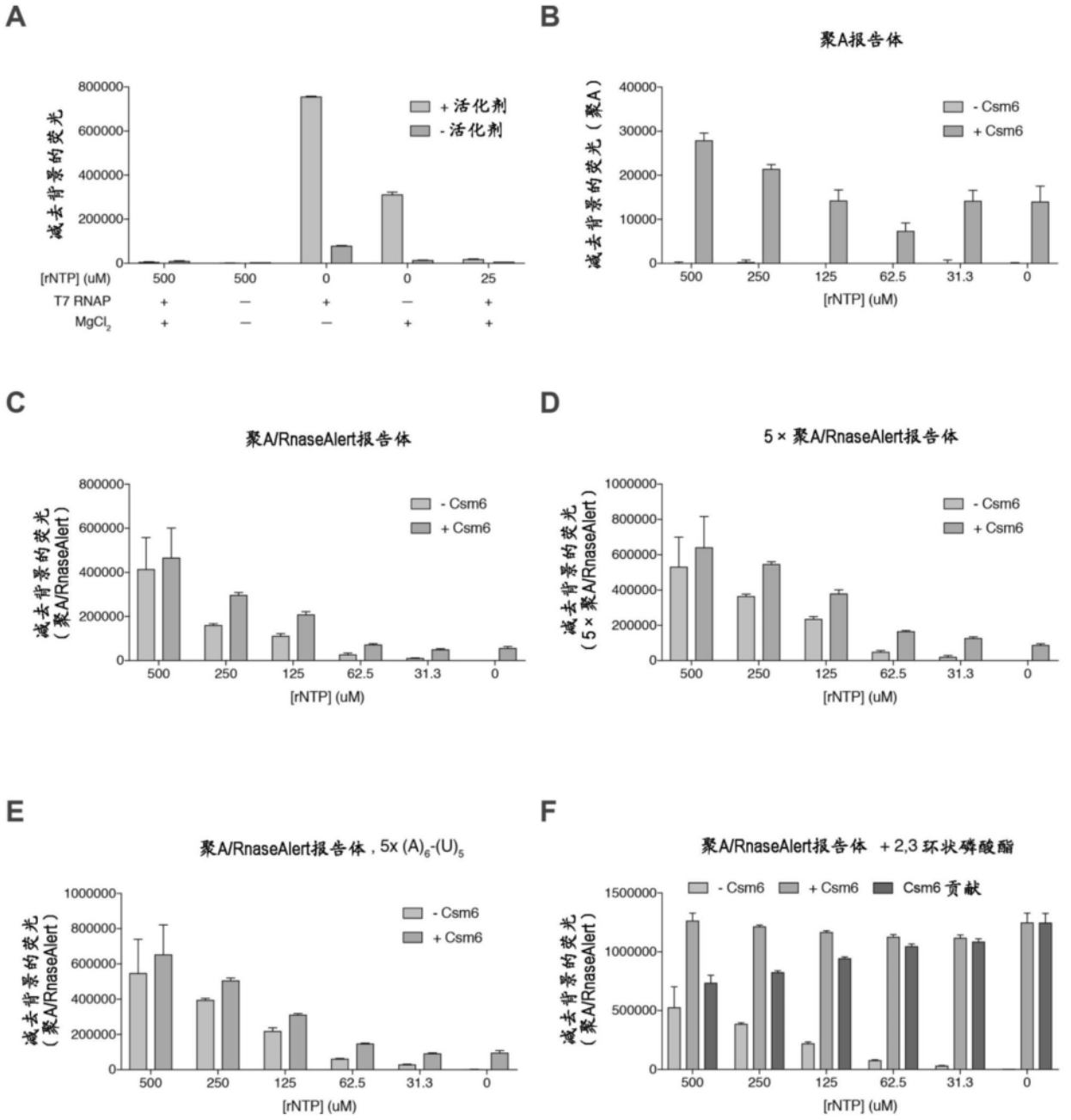

A
减去背景的荧光
800000
600000
400000
200000
0
+ 活化剂
- 活化剂
[rNTP] (uM) 500 500 0 0 25
T7 RNAP + — + — +
MgCl2 + — — + +
B
聚A报告体
减去背景的荧光（聚A）
40000
30000
20000
10000
0
- Csm6
+ Csm6
500 250 125 62.5 31.3 0
[rNTP] (uM)
C
聚A/RnaseAlert报告体
减去背景的荧光（聚A/RnaseAlert）
800000
600000
400000
200000
0
- Csm6
+ Csm6
500 250 125 62.5 31.3 0
[rNTP] (uM)
D
5 × 聚A/RnaseAlert报告体
减去背景的荧光（5×聚A/RnaseAlert）
1000000
800000
600000
400000
200000
0
- Csm6
+ Csm6
500 250 125 62.5 31.3 0
[rNTP] (uM)
E
聚A/RnaseAlert报告体，5x (A)6-(U)5
减去背景的荧光（聚A/RnaseAlert）
1000000
800000
600000
400000
200000
0
- Csm6
+ Csm6
500 250 125 62.5 31.3 0
[rNTP] (uM)
F
聚A/RnaseAlert报告体 +2,3 环状磷酸酯
减去背景的荧光（聚A/RnaseAlert）
1500000
1000000
500000
0
- Csm6
+ Csm6
Csm6 贡献
500 250 125 62.5 31.3 0
[rNTP] (uM)

图36

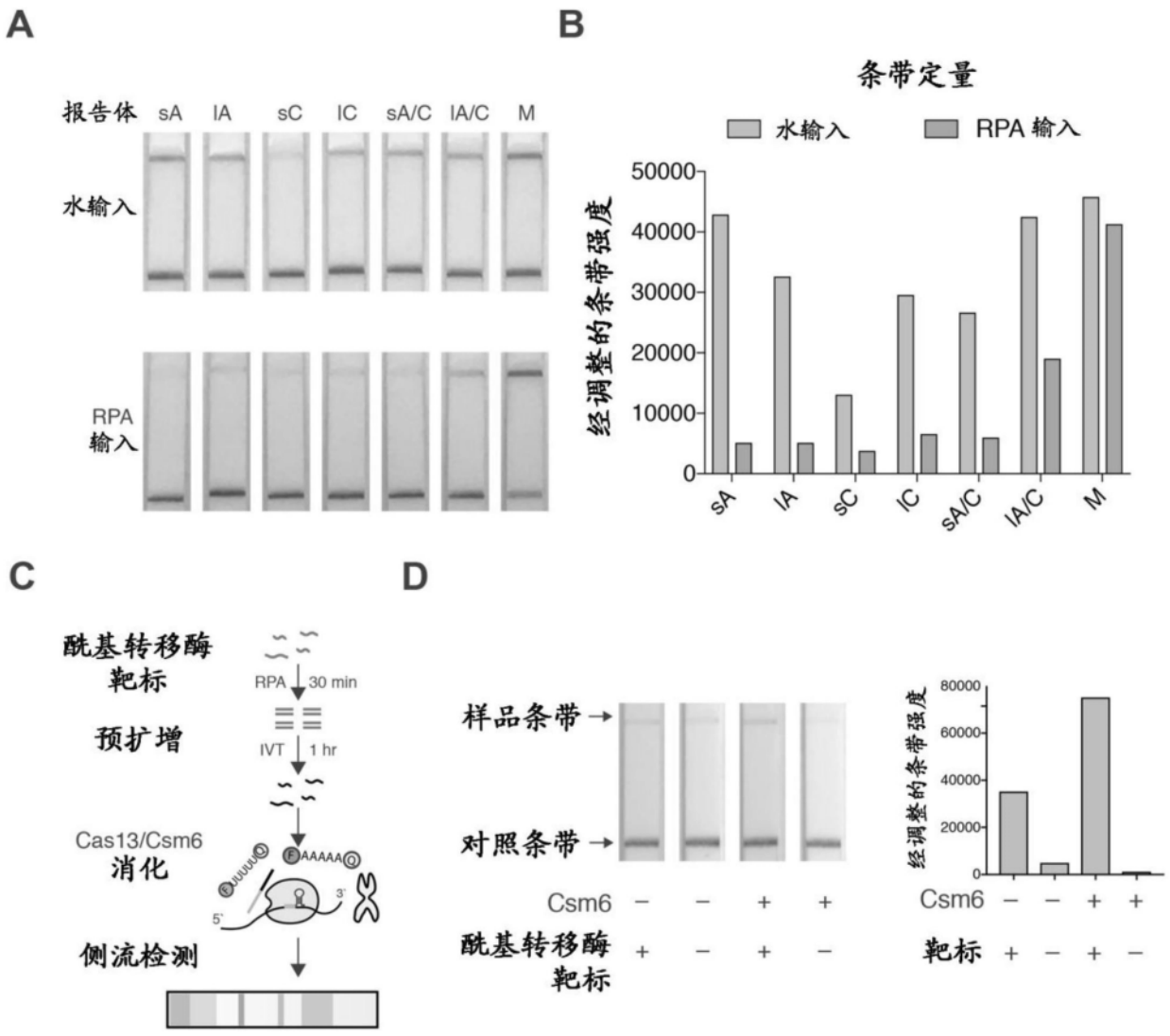

A
报告体
sA
IA
sC
IC
sA/C
IA/C
M
水输入
RPA
输入
B
条带定量
水输入
RPA 输入
经调整的条带强度
50000
40000
30000
20000
10000
0
sA
IA
sC
IC
sA/C
IA/C
M
C
酰基转移酶
靶标
RPA
30 min
预扩增
IVT
1 hr
Cas13/Csm6
消化
AAAAA
5'
3'
侧流检测
D
样品条带
对照条带
Csm6
酰基转移酶
靶标
经调整的条带强度
80000
60000
40000
20000
0
Csm6
靶标

图37

名称	序列	1st Fig.
聚A (n=5) 2',3' 环状磷酸酯末端	rArArArArA-(2,3-cyclic phosphate)	Fig. 4
聚A (n=6) 2',3' 环状磷酸酯末端	rArArArArArA-(2,3-cyclic phosphate)	Fig. 4
聚A (n=7) 2',3' 环状磷酸酯末端	rArArArArArArA-(2,3-cyclic phosphate)	Fig. 4
聚A (n=8) 2',3' 环状磷酸酯末端	rArArArArArArArA-(2,3-cyclic phosphate)	Fig. 4
用于U切割器4A的Csm6聚A聚U探针	rArArArArUrUrUrUrU	Fig. 4
用于U切割器5A的Csm6聚A聚U探针	rArArArArArUrUrUrUrU	Fig. 4
用于U切割器6A的Csm6聚A聚U探针	rArArArArArArUrUrUrUrU	Fig. 4
用于U切割器7A的Csm6聚A聚U探针	rArArArArArArArUrUrUrUrU	Fig. 4
5’ 聚U/聚A 6A探针2,3环状磷酸酯	rUrUrUrUrUrArArArArArArA-(2,3-cyclic phosphate)	fig. S31
5’ 聚A/聚U/聚A 6A探针2,3环状磷酸酯	rArArArArArArUrUrUrUrUrArArArArArArA-(2,3-cyclic phosphate)	fig. S31

图38

缩写	蛋白质名称	菌株名称	Benchling链接	登录号
Lwa	LwaCas13a	Leptotrichia wadei	https://benchling.com/s/seq-66CfLwu7sLMQMbcXe7Ih	WP_021746774.1
Lba	LbaCas13a	Lachnospiraceae bacterium NK4A179	https://benchling.com/s/seq-xdOysFgbmqAsTRoTiERc	WP_022785443.1
Lbu	LbuCas13a	Leptotrichia buccalis C-1013-b	https://benchling.com/s/seq-e0aUn6uEVvWXntoggf60	WP_015770004.1
Bzo	BzoCas13b	Bergeyella zoohelcum	https://benchling.com/s/seq-mA3sJ4GIi4x0JB5q7KHK	WP_002664492
Pin	PinCas13b	Prevotella intermedia	https://benchling.com/s/seq-iA58bdz9mHOZmbFLj92f	WP_036860899
Pbu	PbuCas13b	Prevotella buccae	https://benchling.com/s/seq-nNv4KSgZDFtdPX88zSS2	WP_004343973
Asp	AspCas13b	Alistipes sp. ZOR0009	https://benchling.com/s/seq-lHs6D7J5Z2NkCbbkggek	WP_047447901
Psm	PsmCas13b	Prevotella sp. MA2016	https://benchling.com/s/seq-v7Q1TzaZzAyNZIGKNnH3	WP_036929175
Ran	RanCas13b	Riemerella anatipestifer	https://benchling.com/s/seq-HIhclUZszBOQAdW5rlmW	WP_004919755
Pau	PauCas13b	Prevotella aurantiaca	https://benchling.com/s/seq-Se9MuspJQek3x4vvR1BF	WP_025000926
Psa	PsaCas13b	Prevotella saccharolytica	https://benchling.com/s/seq-NXtrOPLbhpyc9nZk1seg	WP_051522484
Pin2	Pin2Cas13b	Prevotella intermedia	https://benchling.com/s/seq-mSXhS57arjPDuvnQjZOn	WP_061868553
Cca	CcaCas13b	Capnocytophaga canimorsus	https://benchling.com/s/seq-BNVzFUQjqSnkYLARxLwE	WP_013997271
Pgu	PguCas13b	Porphyromonas gulae	https://benchling.com/s/seq-GVOv8zBVIta2utHyuTSR	WP_039434803
Psp	PspCas13b	Prevotella sp. P5-125	https://benchling.com/s/seq-XmnWQgXrpvVAwXoNtJGw	WP_044065294
Pig	PigCas13b	Porphyromonas gingivalis	https://benchling.com/s/seq-hxdDNJtJmA5axRvcxm0p	WP_053444417
Pin3	Pin3Cas13b	Prevotella intermedia	https://benchling.com/s/seq-GlaCfl5cDw4sKXz6LM11	WP_050955369
Ei	EiCsm6	Enterococcus italicus	https://benchling.com/s/seq-YrP8xiVG3rBwxYMgCUH0	WP_007208953.1
Ls	LsCsm6	Lactobacillus salivarius	https://benchling.com/s/seq-duuAaForfhsBc53zLY5z	WP_081509150.1
Tt	TtCsm6	Thermus thermophilus	https://benchling.com/s/seq-esibVH1rmHPjHYXxKWja	WP_011229148.1

图39

名称	直系同源物	完整crRNA序列	间隔子	直接重复序列	靶标	1st Fig.
ssRNA/ssDNA 1 crRNA 2	LwaCas13a	GATTTAGACTACCCCAAAA ACGAAGGGGACTAAAACCT ACCAAGTAATCCATATTTC TAGAGGATC	CTACCAAG TAATCCAT ATTTCTAG AGGATC	GATTTAGACTAC CCCAAAAACGAA GGGGACTAAAAC	ssRNA 1	Fig. 1B/fig. S3
BzoCas13b ssRNA/ssDNA crRNA 2	BzoCas13b	CTACCAAGTAATCCATATT TCTAGAGGATCGTTGGAAC TGCTCTCATTTTGGAGGGT AATCACAAC	CTACCAAG TAATCCAT ATTTCTAG AGGATC	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	Fig. 1B/fig. S3
PinCas13b ssRNA/ssDNA crRNA 2	PinCas13b	CTACCAAGTAATCCATATT TCTAGAGGATCGTTGCATC TGCCTGCTGTTTGCAAGGT AAAAACAAC	CTACCAAG TAATCCAT ATTTCTAG AGGATC	GTTGCATCTGCC TGCTGTTTGCAA GGTAAAAACAAC	ssRNA 1	Fig. 1B/fig. S3
PbuCas13b ssRNA/ssDNA crRNA 2	PbuCas13b	CTACCAAGTAATCCATATT TCTAGAGGATCGTTGCATC TGCCTTCTTTTTGAAAGGT AAAAACAAC	CTACCAAG TAATCCAT ATTTCTAG AGGATC	GTTGCATCTGCC TTCTTTTTGAAA GGTAAAAACAAC	ssRNA 1	Fig. 1B/fig. S3
AspCas13b ssRNA/ssDNA crRNA 2	AspCas13b	CTACCAAGTAATCCATATT TCTAGAGGATCGCTGTTAT ATCCTTACCTTTGTAAGGG AAGTACAGC	CTACCAAG TAATCCAT ATTTCTAG AGGATC	GCTGTTATATCC TTACCTTTGTAA GGGAAGTACAGC	ssRNA 1	Fig. 1B/fig. S3
PsmCas13b ssRNA/ssDNA crRNA 2	PsmCas13b	CTACCAAGTAATCCATATT TCTAGAGGATCGTTGTAGA AGCTTATCGTTTGGATAGG TATGACAAC	CTACCAAG TAATCCAT ATTTCTAG AGGATC	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	Fig. 1B/fig. S3
RanCas13b ssRNA/ssDNA crRNA 2	RanCas13b	CTACCAAGTAATCCATATT TCTAGAGGATCGTTGGGAC TGCTCTCACTTTGAAGGGT ATTCACAAC	CTACCAAG TAATCCAT ATTTCTAG AGGATC	GTTGGGACTGCT CTCACTTTGAAG GGTATTCACAAC	ssRNA 1	Fig. 1B/fig. S3
PauCas13b ssRNA/ssDNA crRNA 2	PauCas13b	CTACCAAGTAATCCATATT TCTAGAGGATCGTTGTATC TGCCTTCTGTTTGAAAGGT AAAAACAAC	CTACCAAG TAATCCAT ATTTCTAG AGGATC	GTTGTATCTGCC TTCTGTTTGAAA GGTAAAAACAAC	ssRNA 1	Fig. 1B/fig. S3
PsaCas13b ssRNA/ssDNA crRNA 2	PsaCas13b	CTACCAAGTAATCCATATT TCTAGAGGATCGTTGTGTC TACCTCCTTTTTGAGAGGT AAAAACAGC	CTACCAAG TAATCCAT ATTTCTAG AGGATC	GTTGTGTCTACC TCCTTTTTGAGA GGTAAAAACAGC	ssRNA 1	Fig. 1B/fig. S3
Pin2Cas13b ssRNA/ssDNA crRNA 2	Pin2Cas13b	CTACCAAGTAATCCATATT TCTAGAGGATCGTTGCATC TGCCTGCTGTTTGCAAGGT AAAAACAAC	CTACCAAG TAATCCAT ATTTCTAG AGGATC	GTTGCATCTGCC TGCTGTTTGCAA GGTAAAAACAAC	ssRNA 1	Fig. 1B/fig. S3
CcaCas13b ssRNA/ssDNA crRNA 2	CcaCas13b	CTACCAAGTAATCCATATT TCTAGAGGATCGTTGGAAC TGCTCTCATTTTGGAGGGT AATCACAAC	CTACCAAG TAATCCAT ATTTCTAG AGGATC	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	Fig. 1B/fig. S3
PguCas13b ssRNA/ssDNA crRNA 2	PguCas13b	CTACCAAGTAATCCATATT TCTAGAGGATCGTTGGATC TACCCTCTATTTGAAGGGT ACACACAAC	CTACCAAG TAATCCAT ATTTCTAG AGGATC	GTTGGATCTACC CTCTATTTGAAG GGTACACACAAC	ssRNA 1	Fig. 1B/fig. S3
PspCas13b ssRNA/ssDNA crRNA 2	PspCas13b	CTACCAAGTAATCCATATT TCTAGAGGATCGTTGTGGA AGGTCCAGTTTTGAGGGGC TATTACAAC	CTACCAAG TAATCCAT ATTTCTAG AGGATC	GTTGTGGAAGGT CCAGTTTTGAGG GGCTATTACAAC	ssRNA 1	Fig. 1B/fig. S3
PigCas13b ssRNA/ssDNA crRNA 2	PigCas13b	CTACCAAGTAATCCATATT TCTAGAGGATCGTTGGATC TACCCTCTATTCGAAGGGT ACACACAAC	CTACCAAG TAATCCAT ATTTCTAG AGGATC	GTTGGATCTACC CTCTATTCGAAG GGTACACACAAC	ssRNA 1	Fig. 1B/fig. S3
Pin3Cas13b ssRNA/ssDNA crRNA 2	Pin3Cas13b	CTACCAAGTAATCCATATT TCTAGAGGATCGTTGCATC TGCCTGCTGTTTGCAAGGT AAAAACAAC	CTACCAAG TAATCCAT ATTTCTAG AGGATC	GTTGCATCTGCC TGCTGTTTGCAA GGTAAAAACAAC	ssRNA 1	Fig. 1B/fig. S3
DENV crRNA LwaCas13a	LwaCas13a	GATTTAGACTACCCCAAAA ACGAAGGGGACTAAAACTG CTTCTGTCCAGTGAGCATG GTCTTCG	TGCTTCTG TCCAGTGA GCATGGTC TTCG	GATTTAGACTAC CCCAAAAACGAA GGGGACTAAAAC	DENV ssRNA	Fig. 1D
DENV crRNA PsmCas13b	PsmCas13b	TTTGCTTCTGTCCAGTGAG CATGGTCTTCGGTTGTAGA AGCTTATCGTTTGGATAGG TATGACAAC	TTTGCTTC TGTCCAGT GAGCATGG TCTTCG	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	DENV ssRNA	Fig. 1D

图40

ssDNA 1 crRNA Cas12a	AsCas12a	TAATTTCTACTCTTGTAGA TCTGTGTTTATCCGCTCAC AA	CTGTGTTT ATCCGCTC ACAA	TAATTTCTACTC TTGTAGAT	ssDNA 1	Fig. 1D
热核酸酶 crRNA PsmCas13b	PsmCas13b	ATGCTTTGTTTCAGGTGTA TCAACCAATAAGTTGTAGA AGCTTATCGTTTGGATAGG TATGACAAC	ATGCTTTG TTTCAGGT GTATCAAC CAATAA	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	Thermon uclease ssDNA	Fig. 1H
酰基转移酶 LwaCas13a crRNA	LwaCas13a	GATTTAGACTACCCCAAAA ACGAAGGGGACTAAAACAG CACGCTCACCCGCGGGTTG CCTTCGG	AGCACGCT CACCCGCG GGTTGCCT TCGG	GATTTAGACTAC CCCAAAAACGAA GGGGACTAAAAC	Acyltrans ferase ssDNA	Fig. 1H
ZIKV LwaCas13a crRNA	LwaCas13a	GATTTAGACTACCCCAAAA ACGAAGGGGACTAAAACAC TCCCTAGAACCACGACAGT TTGCCTT	ACTCCCTA GAACCACG ACAGTTTG CCTT	GATTTAGACTAC CCCAAAAACGAA GGGGACTAAAAC	ZIKV ssRNA	Fig. 3B
EGFR L858R 野生型感测 crRNA	LwaCas13a	GATTTAGACTACCCCAAAA ACGAAGGGGACTAAAACCC AGGCCAAAATCTGTGATCT TGACATG	CCAGGCCA AAATCTGT GATCTTGA CATG	GATTTAGACTAC CCCAAAAACGAA GGGGACTAAAAC	EGFR L858L WT	Fig. 3E
EGFR L858R 突变体感测 crRNA	LwaCas13a	GATTTAGACTACCCCAAAA ACGAAGGGGACTAAAACCC CGGCCAAAATCTGTGATCT TGACATG	CCCGGCCA AAATCTGT GATCTTGA CATG	GATTTAGACTAC CCCAAAAACGAA GGGGACTAAAAC	EGFR L858R mutation	Fig. 3E
外显子19 缺失突变体 感测crRNA	LwaCas13a	GATTTAGACTACCCCAAAA ACGAAGGGGACTAAAACGT TGGCTTTCGGAGATGTCTT GATAGCG	GTTGGCTT TCGGAGAT GTCTTGAT AGCG	GATTTAGACTAC CCCAAAAACGAA GGGGACTAAAAC	EGFR Exon 19 deletion	Fig. 3H
外显子19 缺失野生型 感测crRNA	LwaCas13a	GATTTAGACTACCCCAAAA ACGAAGGGGACTAAAACGA TGTTGCTTCTCTTAATTCC TTGATAG	GATGTTGC TTCTCTTA ATTCCTTG ATAG	GATTTAGACTAC CCCAAAAACGAA GGGGACTAAAAC	EGFR Exon 19 WT	Fig. 3H
A-等位基因（疾病）感测crRNA APC基因 (NM_000038.5) crRNA	LwaCas13a	GATTTAGACTACCCCAAAA ACGAAGGGGACTAAAACCC TATCAGGTTTCACAGTAAG CGCGTAT	CCTATCAG GTTTCACA GTAAGCGC GTAT	GATTTAGACTAC CCCAAAAACGAA GGGGACTAAAAC	APC synthetic mutation	Fig. 5D
G-等位基因（健康）感测crRNA APC基因 (NM_000038.5) crRNA	PsmCas13b	CCTGGTTCATGAGCTTCCT GCCACTGCCAAGTTGTAGA AGCTTATCGTTTGGATAGG TATGACAAC	CCTGGTTC ATGAGCTT CCTGCCAC TGCCAA	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	APC synthetic WT	Fig. 5D
DENV crRNA LbaCas13a	LbaCas13a	GTTGATGAGAAGAGCCCAA GATAGAGGGCAATAACTGC TTCTGTCCAGTGAGCATGG TCTTCG	TGCTTCTG TCCAGTGA GCATGGTC TTCG	GTTGATGAGAAG AGCCCAAGATAG AGGGCAATAAC	DENV ssRNA	fig. S6A
ZIKV crRNA PsmCas13b	PsmCas13b	TGACTCCCTAGAACCACGA CAGTTTGCCTTGTTGTAGA AGCTTATCGTTTGGATAGG TATGACAAC	TGACTCCC TAGAACCA CGACAGTT TGCCTT	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ZIKV ssRNA	fig. S6B
ZIKV crRNA LbaCas13a	LbaCas13a	GTTGATGAGAAGAGCCCAA GATAGAGGGCAATAACACT CCCTAGAACCACGACAGTT TGCCTT	ACTCCCTA GAACCACG ACAGTTTG CCTT	GTTGATGAGAAG AGCCCAAGATAG AGGGCAATAAC	ZIKV ssRNA	fig. S6B
DENV LbuCas13a 28nt 间隔子	LbuCas13a	GACCACCCCAAAAATGAAG GGGACTAAAACATGCTTCT GTCCAGTGAGCATGGTCTT CG	TGCTTCTG TCCAGTGA GCATGGTC TTCG	GACCACCCCAAA AATGAAGGGGAC TAAAACA	DENV ssRNA	fig. S7A
ZIKV LbuCas13a	LbuCas13a	GACCACCCCAAAAATGAAG GGGACTAAAACAACTCCCT AGAACCACGACAGTTTGCC TT	ACTCCCTA GAACCACG ACAGTTTG CCTT	GACCACCCCAAA AATGAAGGGGAC TAAAACA	ZIKV ssRNA	fig. S7A
DENV LbuCas13a 26nt 间隔子	LbuCas13a	GACCACCCCAAAAATGAAG GGGACTAAAACATGCTTCT GTCCAGTGAGCATGGTCTT	TGCTTCTG TCCAGTGA GCATGGTC TT	GACCACCCCAAA AATGAAGGGGAC TAAAACA	DENV ssRNA	fig. S7C
DENV LbuCas13a 24nt 间隔子	LbuCas13a	GACCACCCCAAAAATGAAG GGGACTAAAACATGCTTCT GTCCAGTGAGCATGGTC	TGCTTCTG TCCAGTGA GCATGGTC	GACCACCCCAAA AATGAAGGGGAC TAAAACA	DENV ssRNA	fig. S7C

图40续

DENV LbuCas13a 22nt 间隔子	LbuCas13a	GACCACCCCAAAAATGAAG GGGACTAAAACATGCTTCT GTCCAGTGAGCATGG	TGCTTCTG TCCAGTGA GCATGG	GACCACCCCAAA AATGAAGGGGAC TAAAACA	DENV ssRNA	fig. S7C
DENV LbuCas13a 20nt 间隔子	LbuCas13a	GACCACCCCAAAAATGAAG GGGACTAAAACATGCTTCT GTCCAGTGAGCAT	TGCTTCTG TCCAGTGA GCAT	GACCACCCCAAA AATGAAGGGGAC TAAAACA	DENV ssRNA	fig. S7C
DENV LbuCas13a 18nt 间隔子	LbuCas13a	GACCACCCCAAAAATGAAG GGGACTAAAACATGCTTCT GTCCAGTGAGC	TGCTTCTG TCCAGTGA GC	GACCACCCCAAA AATGAAGGGGAC TAAAACA	DENV ssRNA	fig. S7C
CcaCas13b 间隔子测试34 nt	CcaCas13b	TGTTCTACCAAGTAATCCA TATTTCTAGAGGATCGTTG GAACTGCTCTCATTTTGGA GGGTAATCACAAC	TGTTCTAC CAAGTAAT CCATATTT CTAGAGGA TC	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S10A
CcaCas13b 间隔子测试33 nt	CcaCas13b	GTTCTACCAAGTAATCCAT ATTTCTAGAGGATCGTTGG AACTGCTCTCATTTTGGAG GGTAATCACAAC	GTTCTACC AAGTAATC CATATTTC TAGAGGAT C	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S10A
CcaCas13b 间隔子测试32 nt	CcaCas13b	TTCTACCAAGTAATCCATA TTTCTAGAGGATCGTTGGA ACTGCTCTCATTTTGGAGG GTAATCACAAC	TTCTACCA AGTAATCC ATATTTCT AGAGGATC	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S10A
CcaCas13b 间隔子测试31 nt	CcaCas13b	TCTACCAAGTAATCCATAT TTCTAGAGGATCGTTGGAA CTGCTCTCATTTTGGAGGG TAATCACAAC	TCTACCAA GTAATCCA TATTTCTA GAGGATC	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S10A
CcaCas13b 间隔子测试30 nt	CcaCas13b	CTACCAAGTAATCCATATT TCTAGAGGATCGTTGGAAC TGCTCTCATTTTGGAGGGT AATCACAAC	CTACCAAG TAATCCAT ATTTCTAG AGGATC	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S10A
CcaCas13b 间隔子测试29 nt	CcaCas13b	TACCAAGTAATCCATATTT CTAGAGGATCGTTGGAACT GCTCTCATTTTGGAGGGTA ATCACAAC	TACCAAGT AATCCATA TTTCTAGA GGATC	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S10A
CcaCas13b 间隔子测试28 nt	CcaCas13b	ACCAAGTAATCCATATTTC TAGAGGATCGTTGGAACTG CTCTCATTTTGGAGGGTAA TCACAAC	ACCAAGTA ATCCATAT TTCTAGAG GATC	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S10A
CcaCas13b 间隔子测试27 nt	CcaCas13b	CCAAGTAATCCATATTTCT AGAGGATCGTTGGAACTGC TCTCATTTTGGAGGGTAAT CACAAC	CCAAGTAA TCCATATT TCTAGAGG ATC	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S10A
CcaCas13b 间隔子测试26 nt	CcaCas13b	CAAGTAATCCATATTTCTA GAGGATCGTTGGAACTGCT CTCATTTTGGAGGGTAATC ACAAC	CAAGTAAT CCATATTT CTAGAGGA TC	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S10A
CcaCas13b 间隔子测试25 nt	CcaCas13b	AAGTAATCCATATTTCTAG AGGATCGTTGGAACTGCTC TCATTTTGGAGGGTAATCA CAAC	AAGTAATC CATATTTC TAGAGGAT C	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S10A
CcaCas13b 间隔子测试24 nt	CcaCas13b	AGTAATCCATATTTCTAGA GGATCGTTGGAACTGCTCT CATTTTGGAGGGTAATCAC AAC	AGTAATCC ATATTTCT AGAGGATC	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S10A
CcaCas13b 间隔子测试23 nt	CcaCas13b	GTAATCCATATTTCTAGAG GATCGTTGGAACTGCTCTC ATTTTGGAGGGTAATCACA AC	GTAATCCA TATTTCTA GAGGATC	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S10A
CcaCas13b 间隔子测试22 nt	CcaCas13b	TAATCCATATTTCTAGAGG ATCGTTGGAACTGCTCTCA TTTTGGAGGGTAATCACAA C	TAATCCAT ATTTCTAG AGGATC	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S10A
CcaCas13b 间隔子测试21 nt	CcaCas13b	AATCCATATTTCTAGAGGA TCGTTGGAACTGCTCTCAT TTTGGAGGGTAATCACAAC	AATCCATA TTTCTAGA GGATC	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S10A
CcaCas13b 间隔子测试20 nt	CcaCas13b	ATCCATATTTCTAGAGGAT CGTTGGAACTGCTCTCATT TTGGAGGGTAATCACAAC	ATCCATAT TTCTAGAG GATC	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S10A
CcaCas13b 间隔子测试19 nt	CcaCas13b	TCCATATTTCTAGAGGATC GTTGGAACTGCTCTCATTT TGGAGGGTAATCACAAC	TCCATATT TCTAGAGG ATC	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S10A
CcaCas13b 间隔子测试18 nt	CcaCas13b	CCATATTTCTAGAGGATCG TTGGAACTGCTCTCATTTT GGAGGGTAATCACAAC	CCATATTT CTAGAGGA TC	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S10A
CcaCas13b 间隔子测试17 nt	CcaCas13b	CATATTTCTAGAGGATCGT TGGAACTGCTCTCATTTTG GAGGGTAATCACAAC	CATATTTC TAGAGGAT C	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S10A

图40续

CcaCas13b 间隔子测试 16 nt	CcaCas13b	ATATTTCTAGAGGATCGTT GGAACTGCTCTCATTTTGG AGGGTAATCACAAC	ATATTTCT AGAGGATC	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S10A
CcaCas13b 间隔子测试 15 nt	CcaCas13b	TATTTCTAGAGGATCGTTG GAACTGCTCTCATTTTGGA GGGTAATCACAAC	TATTTCTA GAGGATC	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S10A
CcaCas13b 间隔子测试 14 nt	CcaCas13b	ATTTCTAGAGGATCGTTGG AACTGCTCTCATTTTGGAG GGTAATCACAAC	ATTTCTAG AGGATC	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S10A
CcaCas13b 间隔子测试 13 nt	CcaCas13b	TTTCTAGAGGATCGTTGGA ACTGCTCTCATTTTGGAGG GTAATCACAAC	TTTCTAGA GGATC	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S10A
CcaCas13b 间隔子测试 12 nt	CcaCas13b	TTCTAGAGGATCGTTGGAA CTGCTCTCATTTTGGAGGG TAATCACAAC	TTCTAGAG GATC	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S10A
PsmCas13b 间隔子测试 34 nt	PsmCas13b	TGTTCTACCAAGTAATCCA TATTTCTAGAGGATCGTTG TAGAAGCTTATCGTTTGGA TAGGTATGACAAC	TGTTCTAC CAAGTAAT CCATATTT CTAGAGGA TC	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S10B
PsmCas13b 间隔子测试 33 nt	PsmCas13b	GTTCTACCAAGTAATCCAT ATTTCTAGAGGATCGTTGT AGAAGCTTATCGTTTGGAT AGGTATGACAAC	GTTCTACC AAGTAATC CATATTTC TAGAGGAT C	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S10B
PsmCas13b 间隔子测试 32 nt	PsmCas13b	TTCTACCAAGTAATCCATA TTTCTAGAGGATCGTTGTA GAAGCTTATCGTTTGGATA GGTATGACAAC	TTCTACCA AGTAATCC ATATTTCT AGAGGATC	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S10B
PsmCas13b 间隔子测试 31 nt	PsmCas13b	TCTACCAAGTAATCCATAT TTCTAGAGGATCGTTGTAG AAGCTTATCGTTTGGATAG GTATGACAAC	TCTACCAA GTAATCCA TATTTCTA GAGGATC	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S10B
PsmCas13b 间隔子测试 30 nt	PsmCas13b	CTACCAAGTAATCCATATT TCTAGAGGATCGTTGTAGA AGCTTATCGTTTGGATAGG TATGACAAC	CTACCAAG TAATCCAT ATTTCTAG AGGATC	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S10B
PsmCas13b 间隔子测试 29 nt	PsmCas13b	TACCAAGTAATCCATATTT CTAGAGGATCGTTGTAGAA GCTTATCGTTTGGATAGGT ATGACAAC	TACCAAGT AATCCATA TTTCTAGA GGATC	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S10B
PsmCas13b 间隔子测试 28 nt	PsmCas13b	ACCAAGTAATCCATATTTC TAGAGGATCGTTGTAGAAG CTTATCGTTTGGATAGGTA TGACAAC	ACCAAGTA ATCCATAT TTCTAGAG GATC	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S10B
PsmCas13b 间隔子测试 27 nt	PsmCas13b	CCAAGTAATCCATATTTCT AGAGGATCGTTGTAGAAGC TTATCGTTTGGATAGGTAT GACAAC	CCAAGTAA TCCATATT TCTAGAGG ATC	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S10B
PsmCas13b 间隔子测试 26 nt	PsmCas13b	CAAGTAATCCATATTTCTA GAGGATCGTTGTAGAAGCT TATCGTTTGGATAGGTATG ACAAC	CAAGTAAT CCATATTT CTAGAGGA TC	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S10B
PsmCas13b 间隔子测试 25 nt	PsmCas13b	AAGTAATCCATATTTCTAG AGGATCGTTGTAGAAGCTT ATCGTTTGGATAGGTATGA CAAC	AAGTAATC CATATTTC TAGAGGAT C	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S10B
PsmCas13b 间隔子测试 24 nt	PsmCas13b	AGTAATCCATATTTCTAGA GGATCGTTGTAGAAGCTTA TCGTTTGGATAGGTATGAC AAC	AGTAATCC ATATTTCT AGAGGATC	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S10B
PsmCas13b 间隔子测试 23 nt	PsmCas13b	GTAATCCATATTTCTAGAG GATCGTTGTAGAAGCTTAT CGTTTGGATAGGTATGACA AC	GTAATCCA TATTTCTA GAGGATC	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S10B
PsmCas13b 间隔子测试 22 nt	PsmCas13b	TAATCCATATTTCTAGAGG ATCGTTGTAGAAGCTTATC GTTTGGATAGGTATGACAA C	TAATCCAT ATTTCTAG AGGATC	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S10B
PsmCas13b 间隔子测试 21 nt	PsmCas13b	AATCCATATTTCTAGAGGA TCGTTGTAGAAGCTTATCG TTTGGATAGGTATGACAAC	AATCCATA TTTCTAGA GGATC	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S10B
PsmCas13b 间隔子测试 20 nt	PsmCas13b	ATCCATATTTCTAGAGGAT CGTTGTAGAAGCTTATCGT TTGGATAGGTATGACAAC	ATCCATAT TTCTAGAG GATC	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S10B
PsmCas13b 间隔子测试 19 nt	PsmCas13b	TCCATATTTCTAGAGGATC GTTGTAGAAGCTTATCGTT TGGATAGGTATGACAAC	TCCATATT TCTAGAGG ATC	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S10B
PsmCas13b 间隔子测试 18 nt	PsmCas13b	CCATATTTCTAGAGGATCG TTGTAGAAGCTTATCGTTT GGATAGGTATGACAAC	CCATATTT CTAGAGGA TC	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S10B
PsmCas13b 间隔子测试 17 nt	PsmCas13b	CATATTTCTAGAGGATCGT TGTAGAAGCTTATCGTTTG GATAGGTATGACAAC	CATATTTC TAGAGGAT C	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S10B

图40续

PsmCas13b 间隔子测试16 nt	PsmCas13b	ATATTTCTAGAGGATCGTT GTAGAAGCTTATCGTTTGG ATAGGTATGACAAC	ATATTTCT AGAGGATC	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S10B
PsmCas13b 间隔子测试15 nt	PsmCas13b	TATTTCTAGAGGATCGTTG TAGAAGCTTATCGTTTGGA TAGGTATGACAAC	TATTTCTA GAGGATC	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S10B
PsmCas13b 间隔子测试14 nt	PsmCas13b	ATTTCTAGAGGATCGTTGT AGAAGCTTATCGTTTGGAT AGGTATGACAAC	ATTTCTAG AGGATC	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S10B
PsmCas13b 间隔子测试13 nt	PsmCas13b	TTTCTAGAGGATCGTTGTA GAAGCTTATCGTTTGGATA GGTATGACAAC	TTTCTAGA GGATC	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S10B
PsmCas13b 间隔子测试12 nt	PsmCas13b	TTCTAGAGGATCGTTGTAG AAGCTTATCGTTTGGATAG GTATGACAAC	TTCTAGAG GATC	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S10B
LwaCas13a 铺瓦式crRNA1	LwaCas13a	GATTTAGACTACCCCAAAA ACGAAGGGGACTAAAACCC GGGTACCGAGCTCGAATTC ACTGGCC	CCGGGTAC CGAGCTCG AATTCACT GGCC	GATTTAGACTAC CCCAAAAACGAA GGGGACTAAAAC	ssRNA 1	fig. S11
LwaCas13a 铺瓦式crRNA2	LwaCas13a	GATTTAGACTACCCCAAAA ACGAAGGGGACTAAAACTT CTCTAGAGGATCCCCGGGTA CCGAGCT	TTTCTAGA GGATCCCC GGGTACCG AGCT	GATTTAGACTAC CCCAAAAACGAA GGGGACTAAAAC	ssRNA 1	fig. S11
LwaCas13a 铺瓦式crRNA3	LwaCas13a	GATTTAGACTACCCCAAAA ACGAAGGGGACTAAAACCC AAGTAATCCATATTTCTAG AGGATCC	CCAAGTAA TCCATATT CTCTAGAGG ATCC	GATTTAGACTAC CCCAAAAACGAA GGGGACTAAAAC	ssRNA 1	fig. S11
LwaCas13a 铺瓦式crRNA4	LwaCas13a	GATTTAGACTACCCCAAAA ACGAAGGGGACTAAAACAG ATTGCTGTTCTACCAAGTA ATCCATA	AGATTGCT GTTCTACC AAGTAATC CATA	GATTTAGACTAC CCCAAAAACGAA GGGGACTAAAAC	ssRNA 1	fig. S11
LwaCas13a 铺瓦式crRNA5	LwaCas13a	GATTTAGACTACCCCAAAA ACGAAGGGGACTAAAACCC TGCAGGTCGAGTAGATTGC TGTTCTA	CCTGCAGG TCGAGTAG ATTGCTGT TCTA	GATTTAGACTAC CCCAAAAACGAA GGGGACTAAAAC	ssRNA 1	fig. S11
LwaCas13a 铺瓦式crRNA6	LwaCas13a	GATTTAGACTACCCCAAAA ACGAAGGGGACTAAAACGC CAAGCTTGCATGCCTGCAG GTCGAGT	GCCAAGCT TGCATGCC TGCAGGTC GAGT	GATTTAGACTAC CCCAAAAACGAA GGGGACTAAAAC	ssRNA 1	fig. S11
LwaCas13a 铺瓦式crRNA7	LwaCas13a	GATTTAGACTACCCCAAAA ACGAAGGGGACTAAAACAT GACCATGATTACGCCAAGC TTGCATG	ATGACCAT GATTACGC CAAGCTTG CATG	GATTTAGACTAC CCCAAAAACGAA GGGGACTAAAAC	ssRNA 1	fig. S11
LwaCas13a 铺瓦式crRNA8	LwaCas13a	GATTTAGACTACCCCAAAA ACGAAGGGGACTAAAACCA CAGGAAACAGCTATGACCA TGATTAC	CACAGGAA ACAGCTAT GACCATGA TTAC	GATTTAGACTAC CCCAAAAACGAA GGGGACTAAAAC	ssRNA 1	fig. S11
LwaCas13a 铺瓦式crRNA9	LwaCas13a	GATTTAGACTACCCCAAAA ACGAAGGGGACTAAAACTG TGAGCGGATAAACACAGGA AACAGCT	TGTGAGCG GATAAACA CAGGAAAC AGCT	GATTTAGACTAC CCCAAAAACGAA GGGGACTAAAAC	ssRNA 1	fig. S11
LwaCas13a 铺瓦式crRNA10	LwaCas13a	GATTTAGACTACCCCAAAA ACGAAGGGGACTAAAACAT GTTGTGTGGAATTGTGAGC GGATAAA	ATGTTGTG TGGAATTG TGAGCGGA TAAA	GATTTAGACTAC CCCAAAAACGAA GGGGACTAAAAC	ssRNA 1	fig. S11
LwaCas13a 铺瓦式crRNA11	LwaCas13a	GATTTAGACTACCCCAAAA ACGAAGGGGACTAAAACTG CTTCCGGCTCGTATGTTGT GTGGAAT	TGCTTCCG GCTCGTAT GTTGTGTG GAAT	GATTTAGACTAC CCCAAAAACGAA GGGGACTAAAAC	ssRNA 1	fig. S11
CcaCas13b 铺瓦式crRNA1	CcaCas13b	CCCCGGGTACCGAGCTCGA ATTCACTGGCCGTTGGAAC TGCTCTCATTTTGGAGGGT AATCACAAC	CCCCGGGT ACCGAGCT CGAATTCA CTGGCC	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S11
CcaCas13b 铺瓦式crRNA2	CcaCas13b	TATTTCTAGAGGATCCCCG GGTACCGAGCTGTTGGAAC TGCTCTCATTTTGGAGGGT AATCACAAC	TATTTCTA GAGGATCC CCGGGTAC CGAGCT	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S11
CcaCas13b 铺瓦式crRNA3	CcaCas13b	TACCAAGTAATCCATATTT CTAGAGGATCCGTTGGAAC TGCTCTCATTTTGGAGGGT AATCACAAC	TACCAAGT AATCCATA TTTCTAGA GGATCC	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S11
CcaCas13b 铺瓦式crRNA4	CcaCas13b	GTAGATTGCTGTTCTACCA AGTAATCCATAGTTGGAAC TGCTCTCATTTTGGAGGGT AATCACAAC	GTAGATTG CTGTTCTA CCAAGTAA TCCATA	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S11
CcaCas13b 铺瓦式crRNA5	CcaCas13b	TGCCTGCAGGTCGAGTAGA TTGCTGTTCTAGTTGGAAC TGCTCTCATTTTGGAGGGT AATCACAAC	TGCCTGCA GGTCGAGT AGATTGCT GTTCTA	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S11
CcaCas13b 铺瓦式crRNA6	CcaCas13b	ACGCCAAGCTTGCATGCCT GCAGGTCGAGTGTTGGAAC TGCTCTCATTTTGGAGGGT AATCACAAC	ACGCCAAG CTTGCATG CCTGCAGG TCGAGT	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S11
CcaCas13b 铺瓦式crRNA7	CcaCas13b	CTATGACCATGATTACGCC AAGCTTGCATGGTTGGAAC	CTATGACC ATGATTAC	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S11

图40续

		TGCTCTCATTTTGGAGGGT AATCACAAC	GCCAAGCT TGCATG			
CcaCas13b 铺瓦式crRNA8	CcaCas13b	AACACAGGAAACAGCTATG ACCATGATTACGTTGGAAC TGCTCTCATTTTGGAGGGT AATCACAAC	AACACAGG AAACAGCT ATGACCAT GATTAC	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S11
CcaCas13b 铺瓦式crRNA9	CcaCas13b	ATTGTGAGCGGATAAACAC AGGAAACAGCTGTTGGAAC TGCTCTCATTTTGGAGGGT AATCACAAC	ATTGTGAG CGGATAAA CACAGGAA ACAGCT	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S11
CcaCas13b 铺瓦式crRNA10	CcaCas13b	GTATGTTGTGTGGAATTGT GAGCGGATAAAGTTGGAAC TGCTCTCATTTTGGAGGGT AATCACAAC	GTATGTTG TGTGGAAT TGTGAGCG GATAAA	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S11
CcaCas13b 铺瓦式crRNA11	CcaCas13b	TATGCTTCCGGCTCGTATG TTGTGTGGAATGTTGGAAC TGCTCTCATTTTGGAGGGT AATCACAAC	TATGCTTC CGGCTCGT ATGTTGTG TGGAAT	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ssRNA 1	fig. S11
PsmCas13b 铺瓦式crRNA1	PsmCas13b	CCCCGGGTACCGAGCTCGA ATTCACTGGCCGTTGTAGA AGCTTATCGTTTGGATAGG TATGACAAC	CCCCGGGT ACCGAGCT CGAATTCA CTGGCC	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S11
PsmCas13b 铺瓦式crRNA2	PsmCas13b	TATTTCTAGAGGATCCCCG GGTACCGAGCTGTTGTAGA AGCTTATCGTTTGGATAGG TATGACAAC	TATTTCTA GAGGATCC CCGGGTAC CGAGCT	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S11
PsmCas13b 铺瓦式crRNA3	PsmCas13b	TACCAAGTAATCCATATTT CTAGAGGATCCGTTGTAGA AGCTTATCGTTTGGATAGG TATGACAAC	TACCAAGT AATCCATA TTTCTAGA GGATCC	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S11
PsmCas13b 铺瓦式crRNA4	PsmCas13b	GTAGATTGCTGTTCTACCA AGTAATCCATAGTTGTAGA AGCTTATCGTTTGGATAGG TATGACAAC	GTAGATTG CTGTTCTA CCAAGTAA TCCATA	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S11
PsmCas13b 铺瓦式crRNA5	PsmCas13b	TGCCTGCAGGTCGAGTAGA TTGCTGTTCTAGTTGTAGA AGCTTATCGTTTGGATAGG TATGACAAC	TGCCTGCA GGTCGAGT AGATTGCT GTTCTA	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S11
PsmCas13b 铺瓦式crRNA6	PsmCas13b	ACGCCAAGCTTGCATGCCT GCAGGTCGAGTGTTGTAGA AGCTTATCGTTTGGATAGG TATGACAAC	ACGCCAAG CTTGCATG CCTGCAGG TCGAGT	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S11
PsmCas13b 铺瓦式crRNA7	PsmCas13b	CTATGACCATGATTACGCC AAGCTTGCATGGTTGTAGA AGCTTATCGTTTGGATAGG TATGACAAC	CTATGACC ATGATTAC GCCAAGCT TGCATG	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S11
PsmCas13b 铺瓦式crRNA8	PsmCas13b	AACACAGGAAACAGCTATG ACCATGATTACGTTGTAGA AGCTTATCGTTTGGATAGG TATGACAAC	AACACAGG AAACAGCT ATGACCAT GATTAC	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S11
PsmCas13b 铺瓦式crRNA9	PsmCas13b	ATTGTGAGCGGATAAACAC AGGAAACAGCTGTTGTAGA AGCTTATCGTTTGGATAGG TATGACAAC	ATTGTGAG CGGATAAA CACAGGAA ACAGCT	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S11
PsmCas13b 铺瓦式crRNA10	PsmCas13b	GTATGTTGTGTGGAATTGT GAGCGGATAAAGTTGTAGA AGCTTATCGTTTGGATAGG TATGACAAC	GTATGTTG TGTGGAAT TGTGAGCG GATAAA	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S11
PsmCas13b 铺瓦式crRNA11	PsmCas13b	TATGCTTCCGGCTCGTATG TTGTGTGGAATGTTGTAGA AGCTTATCGTTTGGATAGG TATGACAAC	TATGCTTC CGGCTCGT ATGTTGTG TGGAAT	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 1	fig. S11
ZIKV CcaCas13b	CcaCas13b	CTTGAACTCTACCAGTGCT TCTTTGTTGTTGTTGGAAC TGCTCTCATTTTGGAGGGT AATCACAAC	CTTGAACT CTACCAGT GCTTCTTT GTTGTT	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	ZIKV ssRNA	fig. S16B
DENV crRNA CcaCas13b	CcaCas13b	TTTGCTTCTGTCCAGTGAG CATGGTCTTCGGTTGGAAC TGCTCTCATTTTGGAGGGT AATCACAAC	TTTGCTTC TGTCCAGT GAGCATGG TCTTCG	GTTGGAACTGCT CTCATTTTGGAG GGTAATCACAAC	DENV ssRNA	fig. S17A
人类ID rs601338 A-等位基因感测 PsmCas13b	PsmCas13b	CCGCTTCACCGGCTACCCC TGCTCCAAGAGTTGTAGAA GCTTATCGTTTGGATAGGT ATGACAAC	CCGCTTCA CCGGCTAC CCCTGCTC CAAGA	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	Human locus rs601338	fig. S18C
人类ID rs601338 G-等位基因感测 LwaCas13a	LwaCas13a	GATTTAGACTACCCCAAAA ACGAAGGGGACTAAAACCT GCACCTTCTACCACCACCT CCGCCAG	CTGCACCT TCTACCAC CACCTCCG CCAG	GATTTAGACTAC CCCAAAAACGAA GGGGACTAAAAC	Human locus rs601338	fig. S18C
ssRNA/ssDNA 1 crRNA 1	LwaCas13a	GATTTAGACTACCCCAAAA ACGAAGGGGACTAAAACTA	TAGATTGC TGTTCTAC	GATTTAGACTAC CCCAAAAACGAA GGGGACTAAAAC	ssRNA 1	fig. S20

图40续

		GATTGCTGTTCTACCAAGT AATCCAT	CAAGTAAT CCAT			
T790M突变体感测等位基因crRNA	LwaCas13a	GATTTAGACTACCCCAAAA ACGAAGGGGACTAAAACGC AAGATGAGCTGCACGGTGG AGGTGAG	GCAAGATG AGCTGCAC GGTGGAGG TGAG	GATTTAGACTAC CCCAAAAACGAA GGGGACTAAAAC	EGFR T790M mutant synthetic ssDNA	fig. S24
T790M野生型感测等位基因crRNA	LwaCas13a	GATTTAGACTACCCCAAAA ACGAAGGGGACTAAAACGC GTCATGAGCTGCACGGTGG AGGTGAG	GCGTCATG AGCTGCAC GGTGGAGG TGAG	GATTTAGACTAC CCCAAAAACGAA GGGGACTAAAAC	EGFR T790M WT synthetic ssDNA	fig. S24
ssRNA 3 (PsmCas13b靶标）crRNA	PsmCas13b	TAGATTGCTGTTCTACCAA GTAATCCATATGTTGTAGA AGCTTATCGTTTGGATAGG TATGACAAC	TAGATTGC TGTTCTAC CAAGTAAT CCATAT	GTTGTAGAAGCT TATCGTTTGGAT AGGTATGACAAC	ssRNA 3	fig. S25
ssRNA 2 (LwaCas13a靶标）crRNA	LwaCas13a	GATTTAGACTACCCCAAAA ACGAAGGGGACTAAAACGA TTGCTGTTCTACCAAGTAA TCCATAT	GATTGCTG TTCTACCA AGTAATCC ATAT	GATTTAGACTAC CCCAAAAACGAA GGGGACTAAAAC	ssRNA 2	fig. S25

图40续

名称	序列	核酸	1st Fig.
DENV ssRNA	AGUACAUAUUCAGGGGCCAACCUCUCAACAAUGACGAAGACCAUGCUC ACUGGACAGAAGCAAAAAUGCUGCUGGACAACAUCAACACACCAGAAG GGAUUAUACCAGCUCUCUUUGAACCAGAAAGGGAGAAGUCAGCCGCCA UAGACGGUGAAUACCGCCUGAAGGGU	RNA	Fig. 1B
ssDNA 1	GGCCAGTGAATTCGAGCTCGGTACCCGGGGATCCTCTAGAAATATGGA TTACTTGgtAGAACAGCAATCTACTCGACCTGCAGGCATGCAAGCTTG GCGTAATCATGGTCATAGCTGTTTCCTGTGTTTATCCGCTCACAATTC CACACAACATACGAGCCGGAAGCATAAAG	DNA	Fig. 1F
ZIKV ssRNA	GACACCGGAACUCCACACUGGAACAACAAAGAAGCACUGGUAGAGUUC AAGGACGCACAUGCCAAAAGGCAAACUGUCGUGGUUCUAGGGAGUCAA GAAGGAGCAGUUCACACGGCCCUUGCUGGAGCUCUGGAGGCUGAGAUG GAUGGUGCAAAGGGAAGGCUGUCCUCUGGC	RNA	Fig. 1F
热核酸酶	TTAATTAAAGCGATTGATGGTGATACTGTTAAATTAATGTACAAAGGT CAACCAATGACATTCAGACTATTATTGGTTGATACACCTGAAACAAAG CATCCTAAAAAAGGTGTAGAGAAATATGGTCCTGAAGCAAGTGCATTT ACGAAAAAGATGGTAGAAAATGCAAAGAAAATTGAAGTCGAGTTTG	DNA	Fig. 1H
酰基转移酶	GGGGAGGATGTCGGGCGCGCACGTTTTCCCTTCGCTGAGCACGCTGCG CGCGTCGCCTACGTGAATGCGCTGTTCGATGCGTTGGCCGAAGGCAAC CCGCGGGTGAGCGTGCTCGACCCCTCCAGCGTGCTCTGCGATGGCCTG GATTGTTTCGCCGAACGTGATGGCTGGTCGCTGTACATGGATAACA	DNA	Fig. 1H
ssRNA 1	GGCCAGUGAAUUCGAGCUCGGUACCCGGGGAUCCUCUAGAAAUAUGGA UUACUUGgUAGAACAGCAAUCUACUCGACCUGCAGGCAUGCAAGCUUG GCGUAAUCAUGGUCAUAGCUGUUUCCUGUGUUUAUCCGCUCACAAUUC CACACAACAUACGAGCCGGAAGCAUAAAG	RNA	fig. S3
随机基序文库	TTCCTGTGAAGCTAAAGAAGGAGAATGrNrNrNrNrNrNTATTGATAG CAGCTGTGGCACCTGCAC	Mixed DNA/RNA	fig. S12
EGFR外显子19缺失突变体合成ssDNA	TGCCAGTTAACGTCTTCCTTCTCTCTCTGTCATAGGGACTCTGGATCC CAGAAGGTGAGAAAGTTAAAATTCCCGTCGCTATCAAGACATCTCCGA AAGCCAACAAGGAAATCCTCGATGTGAGTTTCTGCTTTGCTGTGTGGG GGTCCATGGCTCTGAACCTCAGGCCCACCTTTTCTCAT	DNA	fig. S24A
EGFR外显子19缺失WT合成ssDNA	TGCCAGTTAACGTCTTCCTTCTCTCTCTGTCATAGGGACTCTGGATCC CAGAAGGTGAGAAAGTTAAAATTCCCGTCGCTATCAAGGAATTAAGAG AAGCAACATCTCCGAAAGCCAACAAGGAAATCCTCGATGTGAGTTTCT GCTTTGCTGTGTGGGGGTCCATGGCTCTGAACCTCAGGCCCACCTTTT CTCAT	DNA	fig. S24A
EGFR T790M突变体合成ssDNA	CCTCCCTCCAGGAAGCCTACGTGATGGCCAGCGTGGACAACCCCCACG TGTGCCGCCTGCTGGGCATCTGCCTCACCTCCACCGTGCAGCTCATCA TGCAGCTCATGCCCTTCGGCTGCCTCCTGGACTATGTCCGGGAACACA AAGACAATATTGGCTCCCAGTACCTGCTCAACTGGTGTGTGCAGATCG CA	DNA	fig. S24E
EGFR T790M WT合成ssDNA	CCTCCCTCCAGGAAGCCTACGTGATGGCCAGCGTGGACAACCCCCACG TGTGCCGCCTGCTGGGCATCTGCCTCACCTCCACCGTGCAGCTCATCA CGCAGCTCATGCCCTTCGGCTGCCTCCTGGACTATGTCCGGGAACACA AAGACAATATTGGCTCCCAGTACCTGCTCAACTGGTGTGTGCAGATCG CA	DNA	fig. S24E
ssRNA 2 (LwaCas13a 靶标	UAGGUGUUCCACAGGGUAGCCAGCAGCAUCCUGCGAUGCAAAUAUGGA UUACUUGGUAGAACAGCAAUCUAAUCCGGAACAUAAUGGUGCAGGGCG CUGACUUCCGCGUUUGUUUUAAAUCAAACACGGAAACCGAAGACCAUU CAUGUUGUUGCUGCCGGAAGCAUAAAG	RNA	fig. S25B
ssRNA 3 (PsmCas13b 靶标	UAGGUGUUCCACAGGGUAGCCAGCAGCAUCCUGCGAUGCAAAUAUGGA UUACUUGGUAGAACAGCAAUCUAAUCCGGAACAUAAUGGUGCAGGGCG	RNA	fig. S25B
	CUGACUUCCGCGUUUGAAAAAAAACAAACACGGAAACCGAAGACCAUU CAUGUUGUUGCUGCCGGAAGCAUAAAG		

图41

靶标	正向引物序列	正向引物序列（具有T7 RNAP启动子）	反向引物序列	1st Fig.
DENV ssRNA	GTACATATTCAGGGGCCAACCTCTC	gaaattaatacgactcactatagggGTACATATTCAGGGGCCAACCTCTC	TTTCTGGTTCAAAGAGAGCTGGTAT	Fig. 1D
热核酸酶ssDNA	TGTACAAAGGTCAACCAATGACATTCAG	gaaatTAATACGACTCACTATAGGGTGTACAAAGGTCAACCAATGACATTCAG	TGCACTTGCTTCAGGACCATATTTC	Fig. 1H
酰基转移酶	CTACGTGAATGCGCTGTTCGATG	gaaatTAATACGACTCACTATAGGGCTACGTGAATGCGCTGTTCGATG	GAAACAATCCAGGCCATCGCAGAG	Fig. 1H
EGFR L858R	TCTGGATCCCAGAAGGTGAGAAAGTTAAAA	gaaatTAATACGACTCACTATAGGGTCTGGATCCCAGAAGGTGAGAAAGTTAAAA	CCACACAGCAAAGCAGAAACTCACATCGAG	Fig. 3E
EGFR 外显子19缺失	TCTGGATCCCAGAAGGTGAGAAAGTTAAAA	gaaatTAATACGACTCACTATAGGGTCTGGATCCCAGAAGGTGAGAAAGTTAAAA	CCACACAGCAAAGCAGAAACTCACATCGAG	Fig. 3H
治疗诊断学APC靶标(NM_000038.5)	AGGGCCGCCACTCCACCGGCGGCATGGATGAG	gaaatTAATACGACTCACTATAGGGAGGGCCGCCACTCCACCGGCGGCATGGATGAG	GAAGAGTTCTTCACCTTTACTCACggaTCCtcc	Fig. 5B
ZIKV ssRNA	CCACACTGGAACAACAAAGAAGCAC	gaaatTAATACGACTCACTATAGGGCCACACTGGAACAACAAAGAAGCAC	ACAGCCTTCCCTTTGCACCATCCATCTCAG	fig. S6
基因座rs601338	ATAGTCCCCTCGGCGAACATGGACCCCTACAA	gaaattaatacgactcactatagggATAGTCCCCTCGGCGAACATGGACCCCTACAA	GAGTACGTCCGCTTCACCGGCTACCCCTGCTC	fig. S18C
ssDNA/ssRNA 1	ATCCTCTAGAAATATGGATTACTTGGTAGAACAG	AATTCTAATACGACTCACTATAGGGATCCTCTAGAAATATGGATTACTTGGTAGAACAG	GATAAACACAGGAAACAGCTATGACCATGATTACG	fig. S20
EGFR T790M	CCCCACGTGTGCCGCCTGCTGGGCATCTGC	gaaatTAATACGACTCACTATAGGGCCCCACGTGTGCCGCCTGCTGGGCATCTGC	ATATTGTCTTTGTGTTCCCGGACATAGTCC	fig. S24E

图42

名称	序列	荧光团	1st Fig.
聚U报告体	/56-FAM/rUrUrUrUrU/3IABkFQ/	FAM	Fig. 1
聚A报告体	/56-FAM/rArArArArA/3IABkFQ/	FAM	Fig. 1
聚U报告体，用于多重化	/5HEX/rUrUrUrUrU/3IABkFQ/	HEX	Fig. 1
rArA报告体，用于测试二碱基偏好	/56-FAM/TArArAGC/3IABkFQ/	FAM	Fig. 1 and fig. S7
rArU报告体，用于测试二碱基偏好	/56-FAM/TArArUGC/3IABkFQ/	FAM	Fig. 1 and fig. S7
rArC报告体，用于测试二碱基偏好	/56-FAM/TArArCGC/3IABkFQ/	FAM	Fig. 1 and fig. S7
rArG报告体，用于测试二碱基偏好	/56-FAM/TArArGGC/3IABkFQ/	FAM	Fig. 1 and fig. S7
rUrA报告体，用于测试二碱基偏好	/56-FAM/TArUrAGC/3IABkFQ/	FAM	Fig. 1 and fig. S7
rUrU报告体，用于测试二碱基偏好	/56-FAM/TArUrUGC/3IABkFQ/	FAM	Fig. 1 and fig. S7
rUrC报告体，用于测试二碱基偏好	/56-FAM/TArUrCGC/3IABkFQ/	FAM	Fig. 1 and fig. S7
rUrG报告体，用于测试二碱基偏好	/56-FAM/TArUrGGC/3IABkFQ/	FAM	Fig. 1 and fig. S7
rCrA报告体，用于测试二碱基偏好	/56-FAM/TArCrAGC/3IABkFQ/	FAM	Fig. 1 and fig. S7
rCrU报告体，用于测试二碱基偏好	/56-FAM/TArCrUGC/3IABkFQ/	FAM	Fig. 1 and fig. S7
rCrC报告体，用于测试二碱基偏好	/56-FAM/TArCrCGC/3IABkFQ/	FAM	Fig. 1 and fig. S7
rCrG报告体，用于测试二碱基偏好	/56-FAM/TArCrGGC/3IABkFQ/	FAM	Fig. 1 and fig. S7
rGrA报告体，用于测试二碱基偏好	/56-FAM/TArGrAGC/3IABkFQ/	FAM	Fig. 1 and fig. S7
rGrU报告体，用于测试二碱基偏好	/56-FAM/TArGrUGC/3IABkFQ/	FAM	Fig. 1 and fig. S7
rGrC报告体，用于测试二碱基偏好	/56-FAM/TArGrCGC/3IABkFQ/	FAM	Fig. 1 and fig. S7
rGrG报告体，用于测试二碱基偏好	/56-FAM/TArGrGGC/3IABkFQ/	FAM	Fig. 1 and fig. S7
聚U Cy5，用于多重化	/5Cy5/rUrUrUrUrU/3IAbRQSp/	FAM	Fig. 1
具有FAM/生物素的侧流报告体	/56-FAM/mArArUrGrGrCmAmArArUrGrGrCmA/3Bio/	N/A	Fig. 3
聚C报告体	/56-FAM/rCrCrCrCrC/3IABkFQ/	FAM	fig. S3
聚G报告体	/56-FAM/rGrGrGrGrG/3IABkFQ/	FAM	fig. S3
RNA基序文库，用于碱基偏好筛选	TTCCTGTGAAGCTAAAGAAGGAGAATGrNrNrNrNrNrNTATTGATAGCAGCTGTGGCACCTGCAC	N/A	fig. S12

图43

LwaCas13a 验证基序 1	/56-FAM/TrGrUrUrUrUrC/3IABkFQ/	FAM	fig. S13
LwaCas13a 验证基序 2	/56-FAM/TrUrUrUrUrUrC/3IABkFQ/	FAM	fig. S13
LwaCas13a 验证基序 3	/56-FAM/TrCrArUrUrUrG/3IABkFQ/	FAM	fig. S13
PsmCas13b 验证基序 1	/56-FAM/TrUrArUrUrGrA/3IABkFQ/	FAM	fig. S13
PsmCas13b 验证基序 2	/56-FAM/TrArUrUrGrArU/3IABkFQ/	FAM	fig. S13
PsmCas13b 验证基序 3	/56-FAM/TrUrUrGrArUrA/3IABkFQ/	FAM	fig. S13
CcaCas13b 验证基序 1	/56-FAM/TrUrUrUrGrUrU/3IABkFQ/	FAM	fig. S13
CcaCas13b 验证基序 2	/56-FAM/TrUrGrUrUrUrU/3IABkFQ/	FAM	fig. S13
CcaCas13b 验证基序 3	/56-FAM/TrArUrUrUrUrU/3IABkFQ/	FAM	fig. S13
Lwa 正交基序1	/56-FAM/TrCrGrArArUrG/3IABkFQ/	FAM	fig. S14
Lwa 正交基序2	/56-FAM/TrGrUrCrUrCrC/3IABkFQ/	FAM	fig. S14
Lwa 正交基序3	/56-FAM/TrGrCrArUrGrA/3IABkFQ/	FAM	fig. S14
Lwa 正交基序4	/56-FAM/TrCrArUrArCrA/3IABkFQ/	FAM	fig. S14
Lwa 正交基序5	/56-FAM/TrCrArUrArCrG/3IABkFQ/	FAM	fig. S14
Lwa 正交基序6	/56-FAM/TrGrCrArUrArA/3IABkFQ/	FAM	fig. S14
CcaCas13b 正交基序 1	/56-FAM/TrCrUrArCrUrU/3IABkFQ/	FAM	fig. S14
CcaCas13b 正交基序 2	/56-FAM/TrCrUrArCrGrU/3IABkFQ/	FAM	fig. S14
CcaCas13b 正交基序 3	/56-FAM/TrUrUrArArArC/3IABkFQ/	FAM	fig. S14
金纳米颗粒接头	/5ThioMC6-D/rCrUrCrCrCrUrArArUrArArCrArArUrUrUrArUrArArCrUrArUrUrUrCrCrUrArCrCrCrUrUrUrCrCrCrArArArArArArA/3ThioMC3-D/	N/A	fig. S21
磁珠缀合物低聚物	/5AmMC12/AGAGCATCACCATGATCCTGrUrUrUrUrUrUrUTG/iBiodT/CTCGGATATCTCGACTA/36-FAM/	N/A	fig. S22
EiCsm6 验证基序1	/56-FAM/TrGrArCrGrUrG/3IABkFQ/	N/A	fig. S29
用于侧流的短聚A	/FamCE/rArArArArArA/BioBB/	N/A	fig. S34A
用于侧流的长聚A	/FamCE/rArArArArArArArArArArA/BioBB/	N/A	fig. S34A
用于侧流的短聚C	/56-FAM/rCrCrCrCrCrC/3Bio/	N/A	fig. S34A
用于侧流的长聚C	/56-FAM/rCrCrCrCrCrCrCrCrCrCrC/3Bio/	N/A	fig. S34A
用于侧流的短聚A/C	/56-FAM/rArCrArCrArC/3Bio/	N/A	fig. S34A
用于侧流的长聚A/C	/56-FAM/rArCrArCrArCrArCrArCrArC/3Bio/	N/A	fig. S34A

图43续