



(12) 发明专利申请

(10) 申请公布号 CN 112368396 A

(43) 申请公布日 2021.02.12

(21) 申请号 201980043102.2

0·阿布达耶

(22) 申请日 2019.06.26

(74) 专利代理机构 北京市金杜律师事务所

11256

(30) 优先权数据

代理人 陈文平 袁元

62/690,160 2018.06.26 US

62/767,077 2018.11.14 US

(51) Int.Cl.

(85) PCT国际申请进入国家阶段日

C12Q 1/6844 (2018.01)

2020.12.25

(86) PCT国际申请的申请数据

PCT/US2019/039195 2019.06.26

(87) PCT国际申请的公布数据

W02020/006049 EN 2020.01.02

(71) 申请人 博德研究所

地址 美国马萨诸塞州

申请人 哈佛学院院长等 麻省理工学院

(72) 发明人 F·张 M·凯尔纳 J·戈滕贝格

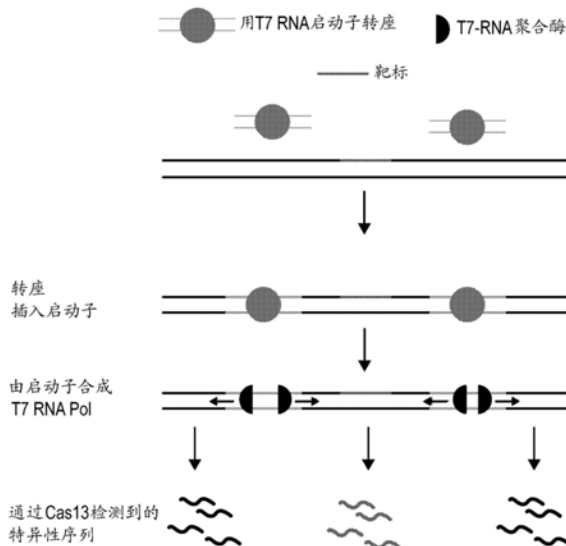
权利要求书3页 说明书26页 附图2页

(54) 发明名称

基于CRISPR/CAS和转座酶的扩增组合物、系统和方法

(57) 摘要

本文提供了用于检测靶核酸序列的方法和系统。所述方法包括使包含所述靶核酸序列的寡核苷酸与转座子复合物接触；使用转座酶将一个或多个T7 RNA启动子插入所述寡核苷酸中；以及(c) 扩增所述靶核酸序列。所述转座子复合物可包含转座酶和包含一个或多个T7 RNA启动子的转座子序列。可通过借助由所述插入的一个或多个T7 RNA启动子转录而生成包含所述靶核酸序列的RNA寡核苷酸来扩增所述靶核酸序列。可使用基于CRISPR Cas13的检测系统来检测所述扩增的靶核酸。



1. 一种检测靶核酸序列的方法,所述方法包括:
 - (a) 使包含所述靶核酸序列的寡核苷酸与转座子复合物接触,其中所述转座子复合物包含转座酶和包含一个或多个T7 RNA聚合酶启动子的转座子序列;
 - (b) 使用所述转座酶将所述一个或多个T7 RNA聚合酶启动子插入所述寡核苷酸中;
 - (c) 通过借助由所述插入的一个或多个T7 RNA聚合酶启动子转录而生成包含所述靶核酸序列的RNA寡核苷酸来扩增所述靶核酸序列。
2. 如权利要求1所述的方法,所述方法还包括检测所述扩增的靶核酸。
3. 如权利要求2所述的方法,其中使用基于CRISPR Cas13的检测系统来检测所述扩增的靶核酸。
4. 如权利要求1至3中任一项所述的方法,其中所述转座酶是Tn5转座酶、Mu转座酶或Tn7转座酶。
5. 如权利要求4所述的方法,其中所述转座酶是Tn5转座酶。
6. 如权利要求5所述的方法,其中所述转座子序列包含两个19个碱基对的嵌合末端(ME) Tn5转座酶识别序列。
7. 如权利要求1至3所述的方法,其中所述转座酶是工程化的转座酶。
8. 如权利要求7所述的方法,其中所述工程化的转座酶经优化以在30°C与45°C之间发挥作用。
9. 如权利要求8所述的方法,其中所述工程化的转座酶经优化以在35°C与40°C之间发挥作用。
10. 如权利要求7至9中任一项所述的方法,其中与野生型转座酶相比,所述工程化的转座酶经优化以便以更快的速率从所述寡核苷酸释放。
11. 如权利要求1至10中任一项所述的方法,其中将所述一个或多个T7 RNA聚合酶启动子插入所述寡核苷酸中是随机的。
12. 如权利要求1至10中任一项所述的方法,其中所述寡核苷酸是双链的。
13. 如权利要求12所述的方法,其中将所述一个或多个T7 RNA聚合酶启动子插入所述双链寡核苷酸中发生在所述双链寡核苷酸的GC富集区中。
14. 如权利要求3所述的方法,其中所述基于CRISPR Cas13的检测系统包含Cas13酶,所述Cas13酶来自选自以下组成的组的属的生物体:纤毛菌属、李斯特菌属、棒状杆菌属、萨特氏菌属、军团菌属、密螺旋体属、产线菌属、真杆菌属、链球菌属、乳杆菌属、支原体属、拟杆菌属、弗维菌属、黄杆菌属、单丝壳属、固氮螺菌属、葡糖醋杆菌属、奈瑟氏菌属、罗氏菌属、细小棒菌属、葡萄球菌属、硝酸盐裂解菌属、支原体属、弯曲杆菌属和毛螺菌属。
15. 如权利要求14所述的方法,其中所述Cas13酶来自选自以下组成的组的生物体:沙氏纤毛菌;韦德纤毛菌(Lw2);斯氏李斯特菌;毛螺菌科细菌MA2020;毛螺菌科细菌NK4A179;嗜氨[梭菌]DSM10710;鸡肉杆菌DSM 4847;鸡肉杆菌DSM 4847(第二CRISPR基因座);产丙酸沼杆菌WB4;韦氏李斯特菌FSL R9-0317;李斯特菌科细菌FSL M6-0635;韦德纤毛菌F0279;荚膜红细菌SB 1003;荚膜红细菌R121;荚膜红细菌DE442;口腔纤毛菌C-1013-b;解半纤维素赫氏菌;直肠[真杆菌];真杆菌科细菌CHKI004;布劳特氏菌属种马赛-P2398;纤毛菌属种口腔分类群879菌株F0557;毛螺菌科细菌NK4A144;聚集绿屈挠菌;桔红色去甲基醌菌;海旋菌属种TSL5-1;假丁酸弧菌属种OR37;丁酸弧菌属种YAB3001;布劳特氏

菌属种马赛-P2398;纤毛菌属种马赛-P3007;爱华拟杆菌;紫单孢菌科细菌KH3CP3RA;崖李斯特菌;和陌生非适应螺菌。

16.如权利要求1或2所述的方法,其中使用基于CRISPR Cas12的检测系统来检测所述扩增的靶核酸。

17.如权利要求16所述的方法,其中所述基于Cas12的检测系统包含Cas12a或Cas12b酶。

18.如权利要求16或17所述的方法,其中所述基于Cas12的检测系统包含Cas12酶,所述Cas12酶来自选自由以下组成的组的属的生物体:土拉弗朗西斯菌、易北普雷沃氏菌、毛螺菌科细菌、解蛋白丁酸弧菌、异域菌门菌、俭菌超门细菌、史密斯氏菌属种、氨基酸球菌属种、毛螺菌科细菌、候选白蚁甲烷支原体、挑剔真杆菌、牛眼莫拉氏菌、稻田氏钩端螺旋体、狗口腔卟啉单胞菌、解糖脲普雷沃氏菌和猕猴卟啉单胞菌、溶糊精琥珀酸弧菌、解糖脲普雷沃氏菌、嗜鳃黄杆菌、孔兹氏创伤球菌、真细菌属种、微基因组菌(罗兹曼菌)、黄杆菌属种、短普雷沃氏菌、山羊莫拉氏菌、口腔拟杆菌、犬嘴卟啉单胞菌、琼氏互养菌、布氏普雷沃氏菌、厌氧弧菌属种、溶纤维丁酸弧菌、候选甲烷嗜甲基菌、丁酸弧菌属种、口腔无芽孢厌氧菌属种、瘤胃假丁酸弧菌和产丁酸菌。

19.如权利要求1至18中任一项所述的方法,其中所述方法不需要加热所述寡核苷酸。

20.如权利要求19所述的方法,其中在恒温下进行所述靶核酸序列的扩增。

21.如前述权利要求中任一项所述的方法,其中所述靶核酸序列的长度为约20-30个、约30-40个、约40-50个或约50-100个核苷酸。

22.如前述权利要求中任一项所述的方法,其中所述靶核酸序列的长度为约100-200个、约100-500个或约100-1000个核苷酸。

23.如前述权利要求中任一项所述的方法,其中所述靶核酸序列的长度为约1000-2000个、约2000-3000个、约3000-4000个或约4000-5000个核苷酸。

24.如前述权利要求中任一项所述的方法,所述方法还包括通过选自由以下组成的组的方法来检测所述扩增的核酸:凝胶电泳、嵌入染料检测、PCR、实时PCR、荧光、荧光共振能量转移(FRET)、质谱、实时RPA、实时LAMP、实时NEAR、实时HDA、实时TMA、实时NASBA和CRISPR-SHERLOCK。

25.如前述权利要求中任一项所述的方法,其中以飞摩尔级灵敏度检测所述靶核酸。

26.如前述权利要求中任一项所述的方法,其中以渺摩尔级灵敏度检测所述靶核酸。

27.如前述权利要求中任一项所述的方法,其中包含所述靶核酸的所述寡核苷酸选自由以下组成的组:基因组DNA、线粒体DNA、病毒DNA、质粒DNA和合成的双链DNA。

28.如前述权利要求中任一项所述的方法,其中包含所述靶核酸的所述寡核苷酸是RNA寡核苷酸,并且其中所述方法包括在使所述寡核苷酸与所述转座子复合物接触之前的逆转录步骤。

29.如前述权利要求中任一项所述的方法,其中所述靶核酸序列来自样品,并且其中所述样品是生物样品或环境样品。

30.如权利要求29所述的方法,其中所述生物样品是血液、血浆、血清、尿液、粪便、痰液、粘液、淋巴液、滑液、胆汁、腹水、胸腔积液、血清肿、唾液、脑脊髓液、水状液或玻璃体液样品,或任何身体分泌物、渗出物、渗出液,或获自关节的流体,或皮肤或粘膜表面的拭子。

31. 如权利要求30所述的方法,其中所述样品是获自人类患者的血液、血浆或血清。
32. 如权利要求29所述的方法,其中所述样品是植物样品。
33. 如前述权利要求中任一项所述的方法,其中所述样品是粗样品。
34. 如权利要求1-32中任一项所述的方法,其中所述样品是纯化的样品。
35. 一种用于检测样品中的靶核酸序列的系统,所述系统包括:
 - a. 转座子复合物,其中所述转座子复合物包含转座酶和包含一个或多个T7 RNA聚合酶启动子的转座子序列,其中所述转座酶将所述一个或多个T7 RNA聚合酶启动子插入包含所述靶核酸序列的寡核苷酸中;
 - b. T7 RNA聚合酶,其中所述T7 RNA聚合酶由所述插入的一个或多个T7启动子转录所述靶核酸;和
 - c. 基于CRISPR的检测系统,所述基于CRISPR的检测系统用于检测所述转录的靶核酸。
36. 如权利要求35所述的系统,其中所述基于CRISPR的检测系统是基于CRISPR-Cas13的检测系统。
37. 如权利要求35所述的系统,其中所述转座酶是Tn5转座酶。
38. 如权利要求35所述的系统,其中所述转座酶是Tn5转座酶的变体。
39. 如权利要求35至38中任一项所述的系统,其中所述转座子序列包含两个19个碱基对的ME Tn5转座酶识别序列。
40. 如权利要求35所述的系统,其中所述基于CRISPR Cas13的检测系统是基于Cas13的CRISPR-SHERLOCK系统。
41. 如权利要求35所述的系统,其中所述基于CRISPR的检测系统是基于CRISPR Cas12的检测系统。
42. 如权利要求35至41中任一项所述的系统,所述系统还包括核酸扩增试剂。
43. 如权利要求35至42中任一项所述的系统,所述系统还包括逆转录试剂。

基于CRISPR/CAS和转座酶的扩增组合物、系统和方法

[0001] 相关申请的交叉引用

[0002] 本申请要求2018年6月26日提交的美国临时申请号62/690,160和2018年11月14日提交的美国临时申请号62/767,077的权益。以上确认的申请的全部内容特此以引用方式完全并入本文。

[0003] 关于联邦资助研究的声明

[0004] 本发明是根据由美国国立卫生研究院 (National Institutes of Health) 授予的授权号MH100706、MH110049和HL141201在政府支持下完成的。政府享有本发明的某些权利。

技术领域

[0005] 本文所公开的主题总体上涉及与CRISPR效应系统的使用有关的核酸扩增方法、系统和快速诊断。

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等人,2017;Green等人,2014;Kumar等人,2014;Pardee等人,2014;Pardee等人,2016;Urdea等人,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。举例来说,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。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。

[0007] 许多核酸扩增方法可用于各种检测平台。其中已经开发出等温核酸扩增方法,无需剧烈温度循环和复杂仪器即可进行扩增。这些方法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) 或切口酶扩增反应 (NEAR)。然而,这些等温扩增方法可能仍需要初始变性步骤和多组引物。此外,将等温核酸扩增与便携式平台相结合的新方法 (Du等人,2017;Pardee等人,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些局限。

发明内容

[0008] 在一方面,本发明提供了一种检测靶核酸序列的方法。所述方法包括:(a) 使包含所述靶核酸序列的寡核苷酸与转座子复合物接触;(b) 使用转座酶将一个或多个RNA聚合酶启动子插入所述寡核苷酸中;以及(c) 扩增所述靶核酸序列。所述转座子复合物可以包含转座酶和包含一个或多个RNA聚合酶启动子的转座子序列。可以通过借助由所述插入的一个或多个RNA聚合酶启动子转录而生成包含所述靶核酸序列的RNA寡核苷酸来扩增所述靶核酸序列。

[0009] 在一些实施方案中,所述方法还可以包括检测所述扩增的靶核酸。可以使用基于

CRISPR Cas13的检测系统来检测所述扩增的靶核酸。在替代实施方案中,可以使用基于CRISPR Cas12的检测系统来检测所述靶核酸。

[0010] 在一些实施方案中,所述转座酶可以是Tn5转座酶、Mu转座酶、Tn7转座酶或工程化的转座酶。Tn5转座酶可以包含两个19个碱基对的嵌合末端(ME) Tn5转座酶识别序列。所述工程化的转座酶可经优化以在30°C与45°C之间发挥作用。所述工程化的转座酶可经优化以在35°C与40°C之间发挥作用。与野生型转座酶相比,所述工程化的转座酶可经优化以便以更快的速率从寡核苷酸释放。

[0011] 将所述一个或多个T7 RNA启动子插入所述寡核苷酸中可以是随机的,或者可以发生在GC富集区中。

[0012] 所述基于CRISPR Cas13的检测系统可以包含Cas13酶,所述Cas13酶来自选自由以下组成的组的属的生物体: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*)、真杆菌属(*Eubacterium*)、链球菌属(*Streptococcus*)、乳杆菌属(*Lactobacillus*)、支原体属(*Mycoplasma*)、拟杆菌属(*Bacteroides*)、弗维菌属(*Flaviivola*)、黄杆菌属(*Flavobacterium*)、单丝壳属(*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*)、细小棒菌属(*Parvibaculum*)、葡萄球菌属(*Staphylococcus*)、硝酸盐裂解菌属(*Nitratifactor*)、支原体属(*Mycoplasma*)、弯曲杆菌属(*Campylobacter*)和毛螺菌属(*Lachnospira*)。

[0013] 所述基于CRISPR Cas13的检测系统可以包含Cas13酶,所述Cas13酶来自选自由以下组成的组:沙氏纤毛菌(*Leptotrichia shahii*);韦德纤毛菌(*Leptotrichia wadei*) (Lw2);斯氏李斯特菌(*Listeria seeligeri*);毛螺菌科细菌(*Lachnospiraceae bacterium*) MA2020;毛螺菌科细菌NK4A179;嗜氨[梭菌] (*[Clostridium] aminophilum*) DSM 10710;鸡肉杆菌(*Carnobacterium gallinarum*) DSM 4847;鸡肉杆菌DSM 4847(第二CRISPR基因座);产丙酸沼杆菌(*Paludibacter propionigenes*) WB4;韦氏李斯特菌(*Listeria weihenstephanensis*) FSL R9-0317;李斯特菌科细菌(*Listeriaceae bacterium*) FSL M6-0635;韦德纤毛菌F0279;荚膜红细菌(*Rhodobacter capsulatus*) SB 1003;荚膜红细菌R121;荚膜红细菌DE442;口腔纤毛菌(*Leptotrichia buccalis*) C-1013-b;解半纤维素赫氏菌(*Herbinix hemicellulosilytica*);直肠[真杆菌] (*[Eubacterium] rectale*);真杆菌科细菌(*Eubacteriaceae bacterium*) CHKCI004;布劳特氏菌属种(*Blautia* sp.) 马赛-P2398;纤毛菌属种口腔分类群(*Leptotrichia* sp.oral taxon) 879菌株F0557;毛螺菌科细菌NK4A144;聚集绿屈挠菌(*Chloroflexus aggregans*);桔红色去甲基醜菌(*Demequina aurantiaca*);海旋菌属种(*Thalassospira* sp.) TSL5-1;假丁酸弧菌属种(*Pseudobutyrvibrio* sp.) OR37;丁酸弧菌属种(*Butyrvibrio* sp.) YAB3001;布劳特氏菌属种马赛-P2398;纤毛菌属种马赛-P3007;爱华拟杆菌(*Bacteroides ihuae*);紫单孢菌科细菌(*Porphyromonadaceae bacterium*) KH3CP3RA;崖李斯特菌(*Listeria riparia*);和陌生非适应螺菌(*Insolitispirillum peregrinum*)。

[0014] 所述基于CRISPR Cas12的检测系统可以包含Cas12a或Cas12b酶。所述基于Cas12的检测系统可以包含Cas12酶,所述Cas12酶来自选自由以下组成的组的属的生物体:土拉

弗朗西斯菌 (*Francisella tularensis*)、易北普雷沃氏菌 (*Prevotella albensis*)、毛螺菌科细菌、解蛋白丁酸弧菌 (*Butyrivibrio proteoclasticus*)、异域菌门菌 (*Peregrinibacteria bacterium*)、俭菌超门细菌 (*Parcubacteria bacterium*)、史密斯氏菌属种 (*Smithella sp.*)、氨基酸球菌属种 (*Acidaminococcus sp.*)、毛螺菌科细菌、候选白蚁甲烷支原体 (*Candidatus Methanoplasma termitum*)、挑剔真杆菌 (*Eubacterium eligens*)、牛眼莫拉氏菌 (*Moraxella bovoculi*)、稻田氏钩端螺旋体 (*Leptospira inadai*)、狗口腔卟啉单胞菌 (*Porphyromonas crevioricanis*)、解糖脲普雷沃氏菌 (*Prevotella disiens*) 和猕猴卟啉单胞菌 (*Porphyromonas macacae*)、溶糊精琥珀酸弧菌 (*Succinivibrio dextrinosolvens*)、解糖脲普雷沃氏菌、嗜鳃黄杆菌 (*Flavobacterium branchiophilum*)、孔兹氏创伤球菌 (*Helcococcus kunzii*)、真细菌属种 (*Eubacterium sp.*)、微基因组菌 (罗兹曼菌) (*Microgenomates (Roizmanbacteria) bacterium*)、黄杆菌属种、短普雷沃氏菌 (*Prevotella brevis*)、山羊莫拉氏菌 (*Moraxella caprae*)、口腔拟杆菌 (*Bacteroidetes oral*)、犬嘴卟啉单胞菌 (*Porphyromonas cansulci*)、琼氏互养菌 (*Synergistes jonesii*)、布氏普雷沃氏菌 (*Prevotella bryantii*)、厌氧弧菌属种 (*Anaerovibrio sp.*)、溶纤维丁酸弧菌 (*Butyrivibrio fibrisolvens*)、候选甲烷嗜甲基菌 (*Candidatus Methanomethylophilus*)、丁酸弧菌属种 (*Butyrivibrio sp.*)、口腔无芽孢厌氧菌属种 (*Oribacterium sp.*)、瘤胃假丁酸弧菌 (*Pseudobutyrvibrio ruminis*) 和产丁酸菌 (*Proteocatella sphenisci*)。

[0015] 在某些实施方案中,所述方法可能不需要加热所述寡核苷酸。

[0016] 在某些实施方案中,可以在恒温下进行所述靶核酸序列的扩增。

[0017] 在一些实施方案中,所述靶核酸序列的长度可以为约20-30个、约30-40个、约40-50个或约50-100个核苷酸。所述靶核酸序列的长度可以为约100-200个、约100-500个或约100-1000个核苷酸。所述靶核酸序列的长度可以为约1000-2000个、约2000-3000个、约3000-4000个或约4000-5000个核苷酸。

[0018] 在一些实施方案中,所述方法还可以包括通过选自以下组成的组的方法来检测所述扩增的核酸:凝胶电泳、嵌入染料检测、PCR、实时PCR、荧光、荧光共振能量转移 (FRET)、质谱、实时RPA、实时LAMP、实时NEAR、实时HDA、实时TMA、实时NASBA和基于CRISPR/Cas的检测方法。

[0019] 所述方法可以允许以飞摩尔级灵敏度检测所述靶核酸。所述方法可以允许以渺摩尔级灵敏度检测所述靶核酸。

[0020] 所述靶核酸可以是基因组DNA、线粒体DNA、病毒DNA、质粒DNA或合成的双链DNA。

[0021] 包含所述靶核酸的所述寡核苷酸可以是RNA寡核苷酸,并且所述方法可以包括在使所述寡核苷酸与所述转座子复合物接触之前的逆转录步骤。

[0022] 在一些实施方案中,所述样品可以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、脑脊髓液、水状液或玻璃体液,或任何身体分泌物、渗出物、渗出液样品,或获自关节的流体,或皮肤或粘膜表面的拭子。在特定实施方案中,所述样品是获自人类患者的血液、血浆或血清。

[0023] 在一些实施方案中,所述样品可以是植物样品。在一些实施方案中,所述样品可以

是粗样品。在一些实施方案中,所述样品可以是纯化的样品。

[0024] 在另一方面,本发明提供了一种用于检测样品中的靶核酸序列的系统。所述系统可以包括转座子复合物、T7 RNA聚合酶和用于检测转录的靶核酸的基于CRISPR Cas13的检测系统。在替代实施方案中,所述检测系统可以是基于CRISPR Cas12的检测系统。

[0025] 在一些实施方案中,所述转座子复合物可以包含转座酶和包含一个或多个T7 RNA启动子的转座子序列。所述转座酶可以将一个或多个T7RNA启动子插入包含所述靶核酸序列的双链多核苷酸中。

[0026] 在一些实施方案中,所述T7聚合酶可以由插入的一个或多个T7启动子转录所述靶核酸。

[0027] 在特定实施方案中,所述转座酶可以是Tn5转座酶或其变体。在一些实施方案中,所述转座子序列可以包含两个19个碱基对的ME Tn5转座酶识别序列。

[0028] 在特定实施方案中,所述基于CRISPR Cas13的检测系统可以是基于Cas13的检测系统。

[0029] 在一些实施方案中,所述用于检测靶核酸序列的系统还可以包括核酸扩增试剂。在一些实施方案中,所述系统还可以包括用于进行逆转录的试剂。

[0030] 结合以下对所说明的示例性实施方案的详细说明,示例性实施方案的这些和其他方面、目的、特征和优点对于本领域普通技术人员而言将变得显而易见。

附图说明

[0031] 通过参考阐述了其中可能利用了本发明原理的说明性实施方案的以下详细说明及其附图,将获得对本发明的特征和优点的理解:

[0032] 图1是示出示例性CRISPR/Cas-转座酶扩增工作流程的示意图。

[0033] 图2是显示各种浓度和温度下的扩增靶序列检测的图。

[0034] 本文中的附图仅用于说明目的,而不一定按比例绘制。

具体实施方式

[0035] 一般定义

[0036] 除非另有规定,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通常所理解的相同的含义。在分子生物学中常用的术语和技术的定义可以在以下文献中找到:Molecular Cloning:A Laboratory Manual,第2版(1989) (Sambrook、Fritsch和Maniatis);Molecular Cloning:A Laboratory Manual,第4版(2012) (Green和Sambrook);Current Protocols in Molecular Biology(1987) (F.M.Ausubel等人编辑);the series Methods in Enzymology(Academic Press,Inc.):PCR2:A Practical Approach(1995) (M.J.MacPherson、B.D.Hames和G.R.Taylor编辑):Antibodies,A Laboratory Manual(1988) (Harlow和Lane编辑):Antibodies A Laboratory Manual,第2版2013 (E.A.Greenfield编辑);Animal Cell Culture(1987) (R.I.Freshney编辑);Benjamin Lewin,Genes IX,Jones and Bartlet出版,2008 (ISBN 0763752223);Kendrew等人(编辑),The Encyclopedia of Molecular Biology,Blackwell Science Ltd.出版,1994 (ISBN 0632021829);Robert A.Meyers(编辑),Molecular Biology and Biotechnology:a

Comprehensive Desk Reference, VCH Publishers, Inc. 出版, 1995 (ISBN 9780471185710); Singleton 等人, Dictionary of Microbiology and Molecular Biology 第2版, J. Wiley & Sons (New York, N.Y. 1994), 三月, Advanced Organic Chemistry Reactions, Mechanisms and Structure 第4版, John Wiley & Sons (New York, N.Y. 1992); 和 Marten H. Hofker 和 Jan van Deursen, Transgenic Mouse Methods and Protocols, 第2版 (2011)。

[0037] 除非上下文另有明确指示, 否则如本文所用的单数形式“一个”、“一种”和“所述”包括单数和复数指示物。

[0038] 术语“任选的”或“任选地”意指后续描述的事件、情形或替代物可能发生或可能不发生, 并且该描述包括事件或情形发生的情况和不发生的情况。

[0039] 通过端点表述的数值范围包括在对应范围内的所有数值和分数, 以及所表述的端点。

[0040] 如本文所用的术语“约”或“近似”当涉及诸如参数、量、时距等可测量的值时, 有意涵盖指定值的变化和从指定值的变化, 诸如指定值和从指定值 $\pm 10\%$ 或更小、 $\pm 5\%$ 或更小、 $\pm 1\%$ 或更小和 $\pm 0.1\%$ 或更小的变化, 只要此类变化适于在所公开的发明中执行即可。应当理解, 修饰语“约”或“近似”所涉及的值本身也是具体地且优选地公开的。

[0041] 在本说明书通篇提及“一些实施方案”、“实施方案”、“示范性实施方案”意指结合实施方案描述的特定特征、结构或特性包括于本发明的至少一个实施方案中。因此, 在本说明书通篇各处出现短语“在一些实施方案中”、“在实施方案中”或“示范性实施方案”不一定全部指代同一实施方案, 但也有可能如此。此外, 如本领域技术人员从本公开将明显了解的, 在一个或多个实施方案中, 特定特征、结构或特性可以按任何适合的方式组合。此外, 虽然本文所述的一些实施方案包括其他实施方案中所包括的一些特征而非其他特征, 但不同实施方案的特征的组合有意处于本发明的范围内。举例来说, 在所附权利要求中, 任何所要求保护的实施方案可以按任何组合使用。

[0042] 如本文所用, “生物样品”可以含有全细胞和/或活细胞和/或细胞碎片。生物样品可以含有(或来源于)“体液”。本发明涵盖以下实施方案, 其中体液选自羊水、房水、玻璃体液、胆汁、血清、母乳、脑脊液、耳屎(耳垢)、乳糜、食糜、内淋巴液、外淋巴液、渗出液、粪便、女性射液、胃酸、胃液、淋巴液、粘液(包括鼻引流和痰)、心包液、腹膜液、胸膜液、脓液、炎性分泌物、唾液、皮脂(皮油)、精液、涎液、滑液、汗液、眼泪、尿液、阴道分泌物、呕吐物以及这些物质中一者或多者的混合物。生物样品包括细胞培养物、体液、来自体液的细胞培养物。体液可以例如通过穿刺或其他收集或采样程序从哺乳生物体获得。

[0043] 术语“受试者”、“个体”和“患者”在本文中可互换使用, 是指脊椎动物, 优选地是哺乳动物, 更优选地是人类。哺乳动物包括但不限于鼠、猴、人类、农场动物、运动型动物和宠物。还包括在体内获得或在体外培养的生物实体的组织、细胞和他们的子代。

[0044] 术语“熔融”、“解链”或“变性”是指分离核酸双链体的两条互补链的全部或部分。

[0045] “核酸酶”和“内切核酸酶”在本文中可互换使用, 意指对DNA切割具有内切核酸催化活性的酶。

[0046] 术语“直系同源物(orthologue)”(在本文中也称为“直系同源物(ortholog)”)和“同系物(homologue)”(本文中也称为“同系物(homolog)”)在本领域中是众所周知的。作为

进一步指导,如本文所用的蛋白质的“同系物”是与作为其同系物的蛋白质发挥相同或类似功能的相同种类的蛋白质。同源蛋白质可以是但不需要是结构上相关的,或仅是部分结构上相关的。如本文所用的蛋白质的“直系同源物”是与作为其直系同源物的蛋白质发挥相同或类似功能的不同种类的蛋白质。直系同源蛋白质可以是但不需要是结构上相关的,或仅是部分结构上相关的。同系物和直系同源物可以通过同源建模(参见例如Greer, Science第228卷(1985)1055和Blundell等人Eur J Biochem vol 172(1988),513)或“结构BLAST”(Dey F,Cliff Zhang Q,Petrey D,Honig B.Toward a“structural BLAST”:using structural relationships to infer function.Protein Sci.2013年4月;22(4):359-66.doi:10.1002/pro.2225.)。另参见Shmakov等人(2015)了解在CRISPR-Cas基因座领域中的申请。同源蛋白质可以是但不需要是结构上相关的,或仅是部分结构上相关的。

[0047] 现在将“C2c2”称为“Cas13a”,除非另有说明,否则这些术语在本文中可互换使用。术语“第29组”、“第30组”和Cas13b在本文中可互换使用。术语“Cpf1”和“Cas12a”在本文中可互换使用。术语“C2c1”和“Cas12b”在本文中可互换使用。

[0048] 在下文中描述各种实施方案。应当指出的是,具体实施方案不旨在作为详尽的描述或作为对本文所论述的更广泛方面的限制。结合特定实施方案描述的一个方面不必限于该实施方案,而是可以与任何其他一个或多个实施方案一起实践。在本说明书通篇提及“一个实施方案”、“实施方案”、“示例性实施方案”意指结合实施方案描述的特定特征、结构或特性包括于本发明的至少一个实施方案中。因此,在本说明书通篇各处出现短语“在一个实施方案中”、“在实施方案中”或“示例性实施方案”不一定全部指代同一实施方案,但也有可能如此。此外,如本领域技术人员从本公开将明显了解的,在一个或多个实施方案中,特定特征、结构或特性可以按任何适合的方式组合。此外,虽然本文所述的一些实施方案包括其他实施方案中所包括的一些特征而非其他特征,但不同实施方案的特征的组合有意处于本发明的范围内。举例来说,在所附权利要求中,任何所要求保护的实施方案可以按任何组合使用。

[0049] 本文所引用的所有公布、公布的专利文献和专利申请特此以引用方式并入,如同每个单独公布、公布的专利文献或专利申请被确切地且单独地指明为以引用方式整体并入。

[0050] 综述

[0051] 本文所公开的实施方案提供了通过使含有靶核酸序列的寡核苷酸与转座子复合物接触来等温扩增所述靶核酸序列的系统和方法。寡核苷酸可以是单链或双链RNA、DNA或RNA/DNA杂合寡核苷酸。转座子复合物包含转座酶和包含一个或多个RNA聚合酶启动子的转座子序列。转座酶促进所述一个或多个RNA聚合酶启动子插入寡核苷酸中。然后,RNA聚合酶启动子可以由插入的一个或多个RNA聚合酶启动子转录靶核酸序列。此系统的一个优点是无需加热或熔融双链DNA模板,因为RNA聚合酶聚合酶需要双链模板。这种等温扩增快速且简单,不需要复杂而昂贵的仪器进行变性和冷却。在某些示例性实施方案中,RNA聚合酶启动子是修饰的T7 RNA启动子的天然产物。

[0052] 检测靶核酸序列的方法

[0053] 在某些示例性实施方案中,本发明提供了一种检测靶核酸序列的方法,所述方法包括:(a)使包含所述靶核酸序列的寡核苷酸与转座子复合物接触,其中所述转座子复合物

包含转座酶和包含一个或多个RNA聚合酶启动子的转座子序列；(b) 使用所述转座酶将一个或多个T7 RNA启动子插入所述寡核苷酸中；以及(c) 通过借助由所述插入的一个或多个T7RNA启动子转录而生成包含所述靶核酸序列的RNA寡核苷酸来扩增所述靶核酸序列。

[0054] 本文所提供的方法可以包括使包含所述靶核酸序列的寡核苷酸与转座子复合物接触。如本文所用，术语“寡核苷酸”可以指但不必限于被设计成与DNA或RNA序列特异性杂交的核酸聚合物。如本文所述，术语“接触”是指以直接物理缔合的形式，包括固体形式或液体形式两者放置。如本文所用，术语“转座子”是指由转座酶或整合酶识别并且是能够进行转座的功能性核酸-蛋白质复合物(即，转座体)的基本组分的核酸区段。如本文所用，术语“转座酶”是指一种酶，它是能够进行转座的功能性核酸-蛋白质复合物的组分并介导转座。术语“转座酶”还指来自逆转录转座子或起源于逆转录病毒的整合酶。转座子复合物在转座酶和双链DNA片段之间形成，所述片段含有对该酶的特异性结合序列，称为“转座子末端”。转座子结合位点的序列可以在某些位置用其他碱基修饰，而不影响转座子复合物形成可以有效地转座成靶DNA的稳定结构的能力。

[0055] 术语“靶核酸序列”或“靶DNA或RNA”或“靶核酸”是指是或包含靶序列的DNA或RNA多核苷酸。换句话说，靶DNA或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所针对的DNA或RNA多核苷酸或DNA或RNA多核苷酸的一部分。在一些实施方案中，靶序列位于细胞的细胞核或细胞质中。在一些实施方案中，寡核苷酸可以是双链的。在一些实施方案中，将T7 RNA聚合酶启动子插入双链寡核苷酸中发生在双链寡核苷酸的GC富集区中。

[0056] 在一些实施方案中，靶核酸序列可以来自样品。在某些实施方案中，样品可以包括但不必限于生物样品或环境样品。生物样品可以包括但不限于血液、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、脑脊髓液、水状液或玻璃体液，或任何身体分泌物、渗出物、渗出液，或获自关节的流体，或皮肤或粘膜表面的拭子。在特定实施方案中，样品可以是获自人类患者的血液、血浆或血清。在一些实施方案中，样品可以是植物样品。在一些实施方案中，样品可以是粗样品。在一些实施方案中，样品可以是纯化的样品。

[0057] 在一些实施方案中，包含靶核酸的寡核苷酸可以是基因组DNA、线粒体DNA、病毒DNA、质粒DNA或合成的双链DNA。在一些实施方案中，包含靶核酸的寡核苷酸可以是RNA寡核苷酸，并且所述方法可以包括在使寡核苷酸与转座子复合物接触之前的逆转录步骤。

[0058] 在本文所提供的实施方案中，转座子复合物包含转座酶和包含一个或多个RNA聚合酶启动子的转座子序列。术语“启动子”是指参与结合RNA聚合酶以起始转录的DNA区域。在特定实施方案中，RNA聚合酶启动子可以是T7 RNA聚合酶启动子。可以使用转座酶将T7 RNA启动子插入双链多核苷酸中。在一些实施方案中，将T7 RNA聚合酶启动子插入寡核苷酸中可以是随机的。

[0059] 大多数转座子的转座频率非常低，他们使用复杂的机制而限制了活性。举例来说，Tn5转座酶利用次优的DNA结合序列，并且该转座酶的C端干扰DNA结合。Reznikoff和同事们对Tn5转座所涉及的机制进行了仔细的表征。Tn5通过剪切和粘贴机制进行转座。转座子具有两对由转座酶利用的19bp元件：外部元件(OE)和内部元件(IE)。一个转座酶单体与所利用的两个元件中的每一个结合。单体与转座子的每个末端结合后，两个单体二聚化，形成突

触。具有至少200bp但小于1000bp的供体主链的载体在细菌中转座的功能最强。转座子切割是通过反式催化进行的,并且仅当与每个DNA末端结合的单体处于突触复合物中时才发生。Tn5转座时具有宽松的靶位点选择,因此可以几乎无靶序列特异性地插入靶DNA中。

[0060] Tn5转座的自然下调可以通过选择超活性转座酶并优化转座酶结合元件来克服[Yorket等人1998]。通过修饰野生型OE的三个碱基制成的嵌合元件(ME)使得细菌以及无细胞系统中的转座事件增加到50倍。估计优化的ME和超活性突变转座酶的联合作用使得转座活性提高到100倍。Goryshin等人表明,可以通过电穿孔将预先形成的Tn5转座复合物功能性地引入细菌或酵母菌中[Goryshin等人2000]。将DNA线性化,以使反向重复序列精确定位在转座子的两端,使得Goryshin和同事们绕开了转座的切割步骤,从而提高了转座效率。

[0061] 在一些实施方案中,转座酶可用于标记包含靶序列的寡核苷酸序列。术语“标记”是指如所述的运用测序的转座酶可接近染色质测定(ATAC-seq)中的步骤。(参见,Buenrostro,J.D.、Giresi,P.G.、Zaba,L.C.、Chang,H.Y.、Greenleaf,W.J.,Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin,DNA-binding proteins and nucleosome position.Nature methods 2013;10(12):1213-1218)。具体而言,在体外加载高通量DNA测序衔接子的超活性Tn5转座酶可以同时通过测序衔接子对基因组进行片段化和标记。在一个实施方案中,衔接子与本文所述的方法相容。

[0062] 在一些实施方案中,转座酶可以是Tn5转座酶。在一些实施方案中,转座酶可以是Tn5转座酶的变体或工程化的转座酶。可以使用本领域中已知的任何方法来工程化转座酶。工程化的转座酶可经优化以在范围为30°C至45°C、35°C至40°C的温度或介于两者之间的任何温度下发挥作用。与野生型转座酶相比,工程化的转座酶可经优化以便以更快的速率从寡核苷酸释放。

[0063] 在一些实施方案中,转座酶可以是Tn5转座酶、Mu转座酶或Tn7转座酶。体外转座效率可能会因所使用的转座子系统而异。一般来讲,Tn5和Mu转座酶可实现更高的转座效率水平。在一些实施方案中,插入可以是随机的。在一些实施方案中,插入可发生在靶序列的GC富集区中。

[0064] 在一些实施方案中,转座子序列可以包含两个19个碱基对的嵌合末端(ME) Tn5转座酶识别序列。Tn5转座酶通常将转座在此类短的19个碱基对的ME Tn5转座酶识别序列之间所含的任何DNA序列。

[0065] 在一些实施方案中,转座酶的使用允许在不存在加热或熔融的情况下分离双链多核苷酸。

[0066] 扩增

[0067] 在某些示例性实施方案中,本发明提供了用于扩增靶核酸序列的方法。如本文别处所述,术语“靶核酸序列”或“靶DNA或RNA”或“靶核酸”是指是或包含靶序列的DNA或RNA多核苷酸。换句话说,靶DNA或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所针对的DNA或RNA多核苷酸或DNA或RNA多核苷酸的一部分。在一些实施方案中,靶序列位于细胞的细胞核或细胞质中。

[0068] 在本文所提供的实施方案中,检测靶核酸序列的方法可还包括使用任何插入的T7

RNA启动子转录靶核酸序列。可通过生成包含靶核酸序列的RNA寡核苷酸来扩增靶核酸序列。

[0069] 本发明的一个实施方案可以包括扩增靶核酸序列。扩增可以是等温的,并针对温度进行选择。在一个实施方案中,扩增以37度快速进行。在其他实施方案中,可以通过选择可在不同温度下操作的聚合酶来选择等温扩增的温度。聚合酶可以选自由以下组成的组: Bst 2.0DNA聚合酶、Bst 2.0WarmStart DNA聚合酶、Bst 3.0DNA聚合酶、全长Bst DNA聚合酶、大片段Bst DNA聚合酶、大片段Bsu DNA聚合酶、phi29 DNA聚合酶、T7 DNA聚合酶和测序酶DNA聚合酶。在特定实施方案中,在恒温下进行靶核酸序列的扩增。

[0070] 该扩增可以用于扩增具有不同长度的靶核酸序列。举例来说,靶核酸序列的长度可以是约10-20个、约20-30个、约30-40个、约40-50个、约50-100个、约100-200个、约100-500个、约100-1000个、约1000-2000个、约2000-3000个、约3000-4000个或约4000-5000个核苷酸。靶核酸可以是DNA,例如基因组DNA、线粒体DNA、病毒DNA、质粒DNA或合成的双链DNA。靶核酸可以是单链核酸,例如RNA分子。可以在碱基扩增之前将单链核酸转化为双链核酸。举例来说,可以在扩增之前通过逆转录将RNA分子转化为双链DNA。单链核酸可以选自由以下组成的组:单链病毒DNA、病毒RNA、信使RNA、核糖体RNA、转移RNA、微小RNA、短干扰RNA、小核RNA、合成RNA和合成单链DNA。

[0071] 可以将等温扩增方法与多种检测方法相结合来检测扩增的核酸产物。举例来说,检测方法可以包括凝胶电泳、嵌入染料检测、PCR、实时PCR、荧光、荧光共振能量转移(FRET)、质谱、实时RPA、实时LAMP、实时NEAR、实时HDA、实时转录介导扩增(TMA)、实时NASBA和CRISPR-SHERLOCK。扩增和检测相结合可以实现渺摩尔级灵敏度或飞摩尔级灵敏度。

[0072] 这些方法可以但不必限于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)。

[0073] 因此,在某些示例性实施方案中,本文所公开的系统可以包括扩增试剂。本文描述了可用于核酸扩增的不同组分或试剂。举例来说,如本文所述的扩增试剂可以包括缓冲液,诸如Tris缓冲液。Tris缓冲液可以在适于所需应用或用途的任何浓度下使用,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等浓度。本领域技术人员将能够确定用于本发明的缓冲液(诸如Tris)的适当浓度。

[0074] 为了改善核酸片段的扩增,可以在扩增反应(诸如PCR)中包括盐,诸如氯化镁($MgCl_2$)、氯化钾(KCl)或氯化钠(NaCl)。尽管盐浓度将取决于特定反应和应用,但在一些实施方案中,特定大小的核酸片段在特定盐浓度下可能产生最佳结果。较大产物可能需要改变的盐浓度,通常是较低的盐,以产生所需结果,而较小产物的扩增在较高盐浓度下可能产生较佳结果。本领域技术人员将了解,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度,并且因此可以使用为本发明和如本文所述的反应提供适当条件的任何盐。

[0075] 生物或化学反应的其他组分可以包括细胞裂解组分以使细胞破开或溶解以供分析其中的物质。细胞裂解组分可以包括但不限于洗涤剂;如上所述的盐,诸如NaCl、KCl、硫酸铵 $[(NH_4)_2SO_4]$;或其他。可以适于本发明的洗涤剂可以包括Triton X-100、十二烷基硫酸

钠(SDS)、CHAPS(3-[3-胆酰胺基丙基)二甲铵基]-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)。洗涤剂的浓度可以取决于特定应用,并且在一些情况下可以特定针对反应。扩增反应可以包括在适于本发明的任何浓度下使用的dNTP和核酸引物,诸如包括但不限于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等浓度。同样地,根据本发明可用的聚合酶可以是本领域中已知并且可用于本发明的任何特定或一般聚合酶,包括Taq聚合酶、Q5聚合酶等。

[0076] 在一些实施方案中,如本文所述的扩增试剂可以适用于热启动扩增中。热启动扩增在一些实施方案中可以有益于减少或消除衔接分子或寡核苷酸的二聚,或以其他方式防止不合需要的扩增产物或人造物并且获得所需产物的最佳扩增。本文所述的用于扩增中的许多组分也可以用于热启动扩增中。在一些实施方案中,视情况而定,适用于热启动扩增的试剂或组分可以替代组成组分中的一种或多种而使用。举例来说,可以使用在特定温度或其他反应条件下表现出所需活性的聚合酶或其他试剂。在一些实施方案中,可以使用经过设计或优化以用于热启动扩增中的试剂,例如,在转座之后或在达到特定温度之后可以使聚合酶激活。此类聚合酶可以是基于抗体或基于适体的。如本文所述的聚合酶在本领域中是已知的。此类试剂的实例可以包括但不限于热启动聚合酶、热启动dNTP和光笼化dNTP。此类试剂在本领域中是已知且可得的。本领域技术人员将能够确定适于个别试剂的最佳温度。

[0077] 核酸扩增可以使用特定热循环机器或设备进行,并且可以在单个反应中或成批进行,以便可以同时任何所需数目的反应。在一些实施方案中,扩增可以使用微流体或机器人装置进行,或可以使用温度的手动改变来进行以达成所需扩增。在一些实施方案中,可以进行优化以获得用于特定应用或材料的最佳反应条件。本领域技术人员将了解并且能够优化反应条件以获得足够的扩增。

[0078] 在某些实施方案中,使用本发明的方法或系统的DNA检测需要在检测之前将(扩增的)DNA转录成RNA。

[0079] 在某些实施方案中,可以通过借助由插入的一个或多个T7 RNA启动子转录而生成包含所述靶核酸序列的RNA寡核苷酸来扩增靶核酸序列。

[0080] 在某些实施方案中,靶核酸序列可以包含RNA。在此类情况下,所述方法可以包括在使靶寡核苷酸与转座子复合物接触之前逆转录靶序列。

[0081] 显然,本发明的检测方法可以涉及核酸扩增和检测程序的各种组合。待检测的核酸可以是任何天然存在的或合成的核酸,包括但不限于DNA和RNA,核酸可以通过任何合适的方法来扩增以提供可以检测的中间产物。对中间产物的检测可以通过任何合适的方法来进行,所述方法包括但不限于结合并激活CRISPR蛋白,所述CRISPR蛋白通过直接或附带活性产生可检测信号部分。

[0082] 检测

[0083] 在某些示例性实施方案中,可以使用检测方法进一步检测扩增的靶核酸。检测扩增的核酸可以包括使用嵌入染料、PCT、实时PCR、荧光探针、FRET或质谱进行检测。在某些示例性实施方案中,可以通过基于CRISPR-Cas的检测系统检测扩增的靶核酸序列。在某些示

例性实施方案中,基于CRISPR-Cas的检测系统可以包括基于CRISPR Cas13的检测系统。示例性的基于CRISPR-Cas13的检测系统描述于Gootenberg等人Science 360,439-444(2018)和国际专利申请号WO/2018/107129,所述文献以引用方式并入本文。示例性的基于CRISPR-Cas13的检测系统还描述于国际专利申请号WO 2018/170340中,所述申请以引用方式整体并入本文,特别着重于第[0157]至[0261]段。

[0084] CRISPR系统

[0085] 一般来讲,如本文中诸如WO 2014/093622(PCT/US2013/074667)的文献中所用的微生物成簇规律间隔短回文重复序列(CRISPR)或CRISPR相关(CRISPR-Cas)或系统共同地涉及CRISPR相关的(“Cas”)基因的表达中所涉及或引导所述基因的活性的转录物和其他元件,包括编码Cas基因的序列、tracr(反式激活CRISPR)序列(例如tracrRNA或活性部分tracrRNA)、tracr配对序列(在内源性CRISPR系统的情形中涵盖“正向重复序列”和tracrRNA加工的部分正向重复序列)、指导序列(在内源性CRISPR系统的情形中也称为“间隔区”),或如本文所用的那个术语“一种或多种RNA”(例如用以导向Cas诸如Cas9的一种或多种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。一般来讲,CRISPR系统由促进在靶序列的位点处CRISPR复合物形成的元件表征(在内源性CRISPR系统的情形中也称为原间隔区)。当CRISPR蛋白质是Cpf1蛋白质时,不需要tracrRNA。

[0086] 本文所提供的实施方案利用靶向RNA的效应子以提供具有渺摩尔级灵敏度的稳健的基于CRISPR的诊断。本文所公开的实施方案可以相当的灵敏度水平检测RNA,并且可基于单碱基对差异将靶标与非靶标区分开。为了易于参考,本文所公开的实施方案也可以称为SHERLOCK(特异性高灵敏度酶促报告子解锁,Specific High-sensitivity Enzymatic Reporter unLOCKing)。

[0087] 本文所述的实施方案提供了使用CRISPR-SHERLOCK方法检测转录的靶核酸。

[0088] 尽管Cas9与Cpf1均靶向DNA,但近来已经发现(Shmakov等人,2015)并表征(Abudayyeh等人,2016;Smargon等人,2017)了单效应RNA指导的RNA酶,包括C2c2,这提供了用于特异性RNA传感的平台。可以使用CRISPR RNA(crRNA)容易且便利地对RNA指导的RNA酶进行重新编程以切割靶RNA。不同于仅切割其DNA靶标的DNA内切核酸酶Cas9和Cpf1,RNA指导的RNA酶(如C2c2)在切割其RNA靶标后保持活性,引起附近的非靶向RNA的“附带”切割(Abudayyeh等人,2016)。这种crRNA编程的附带RNA切割活性使得有机会使用RNA指导的RNA酶通过触发可以充当读出的体内程序性细胞死亡或体外非特异性RNA降解来检测特异性RNA的存在(Abudayyeh等人,2016;East-Seletsky等人,2016)。

[0089] 在某些实施方案中,原间隔区相邻基序(PAM)或PAM样基序引导如本文所公开的效应蛋白复合物结合至目标靶基因座。在一些实施方案中,PAM可以是5' PAM(即,位于原间隔区的5'末端上游)。在其他实施方案中,PAM可以是3' PAM(即,位于原间隔区的5'末端下游)。术语“PAM”可以与术语“PFS”或“原间隔区侧接位点”或“原间隔区侧接序列”互换使用。

[0090] 在优选实施方案中,CRISPR效应蛋白可以识别3' PAM。在某些实施方案中,CRISPR效应蛋白可以识别作为5' H的3' PAM,其中H是A、C或U。在某些实施方案中,效应蛋白可以是沙氏纤毛菌C2c2p,更优选地是沙氏纤毛菌DSM 19757C2c2,并且3' PAM为5' H。

[0091] 在形成CRISPR复合物的情形中,“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”是指是或包含靶序列的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所针对的RNA多核苷酸或RNA多核苷酸的一部分。在一些实施方案中,靶序列位于细胞的细胞核或细胞质中。

[0092] 编码CRISPR效应蛋白,特别是C2c2的核酸分子有利地是密码子优化的CRISPR效应蛋白。在这种情况下,密码子优化的序列的实例是对于在真核生物,例如人类中表达而优化(即,对于在人类中表达而优化),或对于如本文所论述的另一种真核生物、动物或哺乳动物中表达而优化的序列;参见例如WO 2014/093622 (PCT/US2013/074667) 中的SaCas9人类密码子优化的序列。虽然这是优选的,但将了解,其他实例可能存在,并且对于除人类以外的宿主种类的密码子优化或对于特定器官的密码子优化是已知的。在一些实施方案中,编码CRISPR效应蛋白的酶编码序列对于在特定细胞,诸如真核细胞中表达进行密码子优化。真核细胞可以是特定生物体的那些或来源于特定生物体的那些,所述生物体诸如植物或哺乳动物,包括但不限于人类,或如本文所论述的非人类真核生物或动物或哺乳动物,例如小鼠、大鼠、兔、犬、家畜或非人类哺乳动物或灵长类动物。在一些实施方案中,可以排除有可能不会对人类或动物带来任何实质性医学益处的修改人类的种系遗传身份的过程和/或修改动物的遗传身份的过程,以及由此类过程产生的动物。一般来讲,密码子优化是指修饰核酸序列用于增强在目标宿主细胞中的表达的过程,这个过程是通过用在宿主细胞的基因中较频繁或最频繁使用的密码子替代原生序列的至少一个密码子(例如约或大于约1个、2个、3个、4个、5个、10个、15个、20个、25个、50个或更多个密码子),同时维持原生氨基酸序列。不同种类对特定氨基酸的某些密码子展现特定偏性。密码子偏性(生物体之间密码子使用的差异)常常与信使RNA (mRNA) 的翻译效率相关,据信所述效率继而尤其取决于所翻译的密码子的特性和特定转移RNA (tRNA) 分子的可用性。细胞中所选tRNA的主导性一般反映了肽合成中最频繁使用的密码子。因此,可以基于密码子优化来调整基因用于给定的生物体中最佳基因表达。密码子使用表可容易得到,例如,在kazusa.or.jp/codon/上可得的“密码子使用数据库(Codon Usage Database)”中,并且这些表可以按许多方式进行改编。参见Nakamura,Y.,等人“Codon usage tabulated from the international DNA sequence databases:status for the year 2000”Nucl.Acids Res.28:292(2000)。对于在特定宿主细胞中表达对特定序列进行密码子优化的计算机算法也可得到,诸如Gene Forge (Aptagen;Jacobus,PA) 也可得到。在一些实施方案中,编码Cas的序列中的一个或多个密码子(例如1个、2个、3个、4个、5个、10个、15个、20个、25个、50个或更多个或所有密码子)对应于对于特定氨基酸最频繁使用的密码子。

[0093] 在一个示例性实施方案中,效应蛋白包含一个或多个包含RxxxxH基序序列的HEPN结构域。RxxxxH基序序列可以是但不限于来自本文所述的HEPN结构域或本领域中已知的HEPN结构域。RxxxxH基序序列还包括通过组合两个或更多个HEPN结构域的部分而建立的基序序列。如所指出的,共有序列可以来源于题为“新颖的CRISPR酶和系统(Novel CRISPR Enzymes and Systems)”的美国临时专利申请62/432,240、2017年3月15日提交的题为“新颖的VI型CRISPR直系同源物和系统(Novel Type VI CRISPR Orthologs and Systems)”的美国临时专利申请62/471,710和标记为代理人案号47627-05-2133且2017年4月12日提交

的题为“新颖的VI型CRISPR直系同源物和系统 (Novel Type VI CRISPR Orthologs and Systems)”的美国临时专利申请中公开的直系同源物的序列。

[0094] 在特定实施方案中,靶向RNA的VI型Cas酶是C2c2或Cas13a。在其他示例性实施方案中,靶向RNA的VI型Cas酶是Cas13b。在特定实施方案中,如本文所提及的VI型蛋白质,诸如C2c2的同系物或直系同源物与VI型蛋白质,诸如C2c2(例如,基于以下任一种的野生型序列:沙氏纤毛菌C2c2、毛螺菌科细菌MA2020 C2c2、毛螺菌科细菌NK4A179 C2c2、嗜氨梭菌(DSM 10710) C2c2、鸡肉杆菌(DSM 4847) C2c2、产丙酸沼杆菌(WB4) C2c2、韦氏李斯特菌(FSL R9-0317) C2c2、李斯特菌科细菌(FSL M6-0635) C2c2、纽约李斯特菌(*Listeria newyorkensis*) (FSL M6-0635) C2c2、韦德纤毛菌(F0279) C2c2、荚膜红细菌(SB 1003) C2c2、荚膜红细菌(R121) C2c2、荚膜红细菌(DE442) C2c2、韦德纤毛菌(Lw2) C2c2或斯氏李斯特菌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%,更优选地至少85%,甚至更优选地至少90%,例如至少95%的序列同源性或同一性。在其他实施方案中,如本文所提及的VI型蛋白质,诸如C2c2的同系物或直系同源物与野生型C2c2(例如,基于以下任一种的野生型序列:沙氏纤毛菌C2c2、毛螺菌科细菌MA2020 C2c2、毛螺菌科细菌NK4A179C2c2、嗜氨梭菌(DSM 10710) C2c2、鸡肉杆菌(DSM 4847) C2c2、产丙酸沼杆菌(WB4) C2c2、韦氏李斯特菌(FSL R9-0317) C2c2、李斯特菌科细菌(FSL M6-0635) C2c2、纽约李斯特菌(FSL M6-0635) C2c2、韦德纤毛菌(F0279) C2c2、荚膜红细菌(SB 1003) C2c2、荚膜红细菌(R121) C2c2、荚膜红细菌(DE442) C2c2、韦德纤毛菌(Lw2) C2c2或斯氏李斯特菌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%,更优选地至少85%,甚至更优选地至少90%,例如至少95%的序列同一性。

[0095] 在某些其他示例性实施方案中,CRISPR系统效应蛋白是C2c2核酸酶。C2c2的活性可以取决于两个HEPN结构域的存在。这些已经显示为RNA酶结构域,即,切割RNA的核酸酶(特别是内切核酸酶)。C2c2 HEPN还可以靶向DNA,或潜在地DNA和/或RNA。在C2c2的HEPN结构域至少能够结合至RNA并且以其野生型形式切割RNA的基础上,则优选的是C2c2效应蛋白具有RNA酶功能。关于C2c2 CRISPR系统,参考2016年6月17日提交的美国临时案62/351,662和2016年8月17日提交的美国临时案62/376,377。还参考2016年6月17日提交的美国临时案62/351,803。还参考2016年12月8日提交的题为“新颖的Crispr酶和系统 (Novel Crispr Enzymes and Systems)”的美国临时案,带有博德研究所(Broad Institute)编号10035.PA4和代理人案号47627.03.2133。进一步参考East-Seletsky等人“Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection” Nature doi:10.1038/nature19802和Abudayyeh等人“C2c2 is a single-component programmable RNA-guided RNA targeting CRISPR effector” bioRxiv doi:10.1101/054742。

[0096] 在一个实施方案中,Cas蛋白质可以是以下属的生物体的C2c2直系同源物:包括但不限于纤毛菌属、李斯特菌属、棒状杆菌属、萨特氏菌属、军团菌属、密螺旋体属、产线菌属、真杆菌属、链球菌属、乳杆菌属、支原体属、拟杆菌属、弗维菌属、黄杆菌属、单丝壳属、固氮螺菌属、葡糖醋杆菌属、奈瑟氏菌属、罗氏菌属、细小棒菌属、葡萄球菌属、硝酸盐裂解菌属、支原体属、弯曲杆菌属和毛螺菌属。这种属的生物体的种类可以如本文其他方面所论述。

[0097] 在某些示例性实施方案中,本发明的C2c2效应蛋白包括但不限于以下21种直系同

源物种类(包括多个CRISPR基因座):沙氏纤毛菌;韦德纤毛菌(Lw2);斯氏李斯特菌;毛螺菌科细菌MA2020;毛螺菌科细菌NK4A179;嗜氨[梭菌]DSM 10710;鸡肉杆菌DSM 4847;鸡肉杆菌DSM4847(第二CRISPR基因座);产丙酸沼杆菌WB4;韦氏李斯特菌FSL R9-0317;李斯特菌科细菌FSL M6-0635;韦德纤毛菌F0279;荚膜红细菌SB 1003;荚膜红细菌R121;荚膜红细菌DE442;口腔纤毛菌C-1013-b;解半纤维素赫氏菌;直肠[真杆菌];真杆菌科细菌CHKCI004;布劳特氏菌属种马赛-P2398;和纤毛菌属种口腔分类群879菌株F0557。另外十二(12)种非限制性实例是:毛螺菌科细菌NK4A144;聚集绿屈挠菌;桔红色去甲基醌菌;海旋菌属种TSL5-1;假丁酸弧菌属种OR37;丁酸弧菌属种YAB3001;布劳特氏菌属种马赛-P2398;纤毛菌属种马赛-P3007;爱华拟杆菌;紫单孢菌科细菌KH3CP3RA;崖李斯特菌;和陌生非适应螺菌。

[0098] 在实施方案中,如本文所提及的C2c2蛋白质还涵盖C2c2或其同系物或直系同源物的功能变体。如本文所用的蛋白质的“功能变体”是指这种蛋白质的变体,其至少部分保留所述蛋白质的活性。功能变体可以包括突变体(其可以是插入、缺失或置换突变体),包括多形体等。功能变体还包括这种蛋白质与另一种通常无关的核酸、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。有利的实施方案可以涉及工程化的或非天然存在的靶向RNA的VI型效应蛋白。

[0099] 在一个实施方案中,编码C2c2或其直系同源物或同系物的一个或多个核酸分子可以对于在真核细胞中表达进行密码子优化。真核生物可以如本文所论述。一个或多个核酸分子可以是工程化的或非天然存在的。

[0100] 在一个实施方案中,C2c2或其直系同源物或同系物可以包含一个或多个突变,并且因此编码其的一个或多个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。关于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。

[0101] CRISPR-SHERLOCK方法可以是基于Cas13的CRISPR-SHERLOCK方法(参见Gootenberg等人Science 356(6336):438-442(2017);Gootenberg等人Science 360:439-444(2018);Myhrvold等人Science 360:444-448(2018);PCT/US17/65477)。

[0102] 在一些实施方案中,还可以使用基于CRISPR Cas12的检测系统来检测扩增的靶核酸。基于Cas12的检测系统可以包含Cas12a、Cas12b或Cas12c酶。

[0103] 基于Cas12的检测系统可以包含Cas12酶,所述Cas12酶来自选自以下组成的组的属的生物体:土拉弗朗西斯菌、易北普雷沃氏菌、毛螺菌科细菌、解蛋白丁酸弧菌、异域菌门菌、俭菌超门细菌、史密斯氏菌属种、氨基酸球菌属种、毛螺菌科细菌、候选白蚁甲烷支原体、挑剔真杆菌、牛眼莫拉氏菌、稻田氏钩端螺旋体、狗口腔卟啉单胞菌、解糖脲普雷沃氏菌和猕猴卟啉单胞菌、溶糊精琥珀酸弧菌、解糖脲普雷沃氏菌、嗜鳃黄杆菌、孔兹氏创伤球菌、真细菌属种、微基因组菌(罗兹曼菌)、黄杆菌属种、短普雷沃氏菌、山羊莫拉氏菌、口腔拟杆菌、犬嘴卟啉单胞菌、琼氏互养菌、布氏普雷沃氏菌、厌氧弧菌属种、溶纤维丁酸弧菌、候选甲烷嗜甲基菌、丁酸弧菌属种、口腔无芽孢厌氧菌属种、瘤胃假丁酸弧菌和产丁酸菌。

[0104] 指导序列

[0105] 如本文所用的术语“指导序列”、“crRNA”、“指导RNA”或“单指导RNA”或“gRNA”是指包含与靶核酸序列具有足够互补性的任何多核苷酸序列的多核苷酸以与靶核酸序列杂交并且引导包含指导序列和CRISPR效应蛋白的靶向RNA的复合物序列特异性地结合至靶核酸

序列。在一些示例性实施方案中,当使用合适的比对算法最佳比对时,互补程度为约或大于约50%、60%、75%、80%、85%、90%、95%、97.5%、99%或更大。最佳比对可以借助于用于比对序列的任何合适算法来确定,其非限制性实例包括史密斯-沃特曼算法(Smith-Waterman algorithm)、尼德曼-翁施算法(Needleman-Wunsch algorithm)、基于巴罗斯-维勒变换(Burrows-Wheeler Transform)的算法(例如巴罗斯-维勒比对仪(Burrows Wheeler Aligner))、ClustalW、Clustal X、BLAT、Novoalign(Novocraft Technologies;在www.novocraft.com上可得)、ELAND(Illumina, San Diego, CA)、SOAP(在soap.genomics.org.cn上可得)和Maq(在maq.sourceforge.net上可得)。可以通过任何合适的测定来评定指导序列(在核酸靶向指导RNA内)引导核酸靶向复合物序列特异性地结合至靶核酸序列的能力。举例来说,足以形成核酸靶向复合物的核酸靶向CRISPR系统的组分,包括有待测试的指导序列,可以提供给具有相应靶核酸序列的宿主细胞,诸如通过用编码核酸靶向复合物的组分的载体转染,继而诸如通过如本文所述的Surveyor测定评定靶核酸序列内的优先靶向(例如切割)。类似地,可以在试管中通过提供靶核酸序列、核酸靶向复合物的组分(包括有待测试的指导序列)和不同于测试指导序列的对照指导序列,以及在测试指导序列与对照指导序列反应之间比较靶序列处的结合或切割速率来评估靶核酸序列的切割。其他测定可能存在,并且将为本领域技术人员所想到。可以选择指导序列并且因此选择核酸靶向指导以靶向任何靶核酸序列。靶序列可以是DNA。靶序列可以是任何RNA序列。在一些实施方案中,靶序列可以是选自由以下组成的组的RNA分子内的序列:信使RNA(mRNA)、前体mRNA、核糖体RNA(rRNA)、转移RNA(tRNA)、微小RNA(miRNA)、小干扰RNA(siRNA)、小细胞核RNA(snRNA)、小细胞核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案中,靶序列可以是选自由mRNA、前体mRNA和rRNA组成的组的RNA分子内的序列。在一些优选的实施方案中,靶序列可以是选自由ncRNA和lncRNA组成的组的RNA分子内的序列。在一些更优选的实施方案中,靶序列可以是mRNA分子或前体mRNA分子内的序列。

[0106] 在一些实施方案中,选择核酸靶向指导物以降低核酸靶向指导物内的二级结构程度。在一些实施方案中,当最佳折叠时,核酸靶向指导物的约或小于约75%、50%、40%、30%、25%、20%、15%、10%、5%、1%或更少的核苷酸参与自身互补碱基配对。最佳折叠可以通过任何合适的多核苷酸折叠算法来确定。一些程序是基于计算最小吉布斯自由能(Gibbs free energy)。一种此类算法的实例是如Zuker和Stiegler(Nucleic Acids Res. 9 (1981), 133-148)所描述的mFold。另一个示例性折叠算法是维也纳大学(University of Vienna)的理论化学研究所(Institute for Theoretical Chemistry)开发的使用质心结构预测算法的在线网络服务器RNAfold(参见例如A.R.Gruber等人,2008,Cell 106(1):23-24;以及PA Carr和GM Church,2009,Nature Biotechnology 27(12):1151-62)。

[0107] 在某些实施方案中,指导RNA或crRNA可以包含正向重复(DR)序列和指导序列或间隔区序列,基本上由其组成,或由其组成。在某些实施方案中,指导RNA或crRNA可以包含融合或连接至指导序列或间隔区序列的正向重复序列,基本上由其组成,或由其组成。在某些实施方案中,正向重复序列可以位于指导序列或间隔区序列上游(即,5')。在其他实施方案中,正向重复序列可以位于指导序列或间隔区序列下游(即3')。

[0108] 在某些实施方案中,crRNA包含茎环,优选单个茎环。在某些实施方案中,正向重复

序列形成茎环,优选单个茎环。

[0109] 在某些实施方案中,指导RNA的间隔区长度为15至35nt。在某些实施方案中,指导RNA的间隔区长度为至少15个核苷酸。在某些实施方案中,间隔区长度为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;27-30nt,例如27、28、29或30nt;30-35nt,例如30、31、32、33、34或35nt;或35nt或更长。

[0110] 一般来讲,CRISPR-Cas、CRISPR-Cas9或CRISPR系统可以如在诸如WO 2014/093622 (PCT/US2013/074667) 的前述文献中那样使用并且共同地涉及CRISPR相关的(“Cas”)基因的表达中所涉及或引导所述基因的活性的转录物和其他元件,包括编码Cas基因(特别地,在CRISPR-Cas9的情况下为Cas9基因)的序列、tracr(反式激活CRISPR)序列(例如tracrRNA或活性部分tracrRNA)、tracr配对序列(在内源性CRISPR系统的情形中涵盖“正向重复序列”和tracrRNA加工的部分正向重复序列)、指导序列(在内源性CRISPR系统的情形中也称为“间隔区”),或如本文所用的那个术语“一种或多种RNA”(例如用以导向Cas9的一种或多种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。一般来讲,CRISPR系统由促进在靶序列的位点处CRISPR复合物形成的元件表征(在内源性CRISPR系统的情形中也称为原间隔区)。在形成CRISPR复合物的情形中,“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。与靶序列的互补对于切割活性很重要的指导序列的部分在本文中称为种子序列。靶序列可以包含任何多核苷酸,诸如DNA或RNA多核苷酸。在一些实施方案中,靶序列位于细胞的细胞核或细胞质中,并且可以包括处于存在于细胞内的线粒体、细胞器、囊泡、脂质体或粒子中或来自其的核酸。在一些实施方案中,特别是对于非核用途,NLS不是优选的。在一些实施方案中,CRISPR系统包含一个或多个核输出信号(NES)。在一些实施方案中,CRISPR系统包含一个或多个NLS和一个或多个NES。在一些实施方案中,可以通过搜索满足以下任何或所有条件的重复基序,在计算机上鉴定正向重复序列:1. 在II型CRISPR基因座侧翼的2Kb基因组序列窗口中;2. 跨度为20至50bp;和3. 间隔20至50bp。在一些实施方案中,可以使用这些标准中的2个,例如1和2、2和3或1和3。在一些实施方案中,可以使用所有3个标准。

[0111] 在本发明的实施方案中,术语指导序列和指导RNA,即,能够将Cas导向靶基因组基因座的RNA,如前面引用的文献诸如WO2014/093622 (PCT/US2013/074667) 中所述互换使用。一般来讲,指导序列是与靶多核苷酸序列具有足够互补性以与靶序列杂交并且引导CRISPR复合物序列特异性地结合至靶序列的任何多核苷酸序列。在一些实施方案中,当使用合适的比对算法最佳比对时,指导序列与其相应靶序列之间的互补程度为约或大于约50%、60%、75%、80%、85%、90%、95%、97.5%、99%或更大。最佳比对可以借助于用于比对序列的任何合适算法来确定,其非限制性实例包括史密斯-沃特曼算法(Smith-Waterman algorithm)、尼德曼-翁施算法(Needleman-Wunsch algorithm)、基于巴罗斯-维勒变换(Burrows-Wheeler Transform)的算法(例如巴罗斯-维勒比对仪(Burrows Wheeler Aligner))、ClustalW、Clustal X、BLAT、Novoalign(Novocraft Technologies;在www.novocraft.com上可得)、ELAND(Illumina, San Diego, CA)、SOAP(在soap.genomics.org.cn上可得)和Maq(在maq.sourceforge.net上可得)。在一些实施方案

中,指导序列的长度为约或大于约5个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个、20个、21个、22个、23个、24个、25个、26个、27个、28个、29个、30个、35个、40个、45个、50个、75个或更多个核苷酸。在一些实施方案中,指导序列的长度为小于约75个、50个、45个、40个、35个、30个、25个、20个、15个、12个或更少个核苷酸。优选地,指导序列长度为10-30个核苷酸。可以通过任何适合的测定来评定指导序列引导CRISPR复合物序列特异性地结合至靶序列的能力。举例来说,足以形成CRISPR复合物的CRISPR系统的组分,包括有待测试的指导序列,可以提供给具有相应靶序列的宿主细胞,诸如通过用编码CRISPR序列的组分的载体转染,继而诸如通过如本文所述的Surveyor测定评定靶序列内的优先切割。类似地,可以在试管中通过提供靶序列、CRISPR复合物的组分(包括有待测试的指导序列)和不同于测试指导序列的对照指导序列,以及在测试指导序列与对照指导序列反应之间比较靶序列处的结合或切割速率来评估靶多核苷酸序列的切割。其他测定可能存在,并且将为本领域技术人员所想到。

[0112] 在CRISPR-Cas系统的一些实施方案中,指导序列与其相应靶序列之间的互补程度可以为约或大于约50%、60%、75%、80%、85%、90%、95%、97.5%、99%或100%;指导物或RNA或sgRNA的长度可以为约或大于约5个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个、20个、21个、22个、23个、24个、25个、26个、27个、28个、29个、30个、35个、40个、45个、50个、75个或更多个核苷酸;或者指导物或RNA或sgRNA的长度可以小于约75个、50个、45个、40个、35个、30个、25个、20个、15个、12个或更少个核苷酸;并且有利地tracr RNA的长度为30或50个核苷酸。然而,本发明的一个方面是减少脱靶相互作用,例如,减少指导物与具有低互补性的靶序列相互作用。实际上,在实施例显示,本发明涉及使得CRISPR-Cas系统能够区分靶序列与具有大于80%至约95%互补性,例如83%-84%或88-89%或94-95%互补性的脱靶序列(例如,区分具有18个核苷酸的靶标与具有1个、2个或3个错配的18个核苷酸的脱靶)的突变。因此,在本发明的情形中,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%的序列与指导物之间的互补性,有利的是,脱靶为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%的序列与指导物之间的互补性。

[0113] 指导物修饰

[0114] 在某些实施方案中,本发明的指导物包含非天然存在的核酸和/或非天然存在的核苷酸和/或核苷酸类似物和/或化学修饰。非天然存在的核酸可以包括例如天然和非天然存在的核苷酸的混合物。非天然存在的核苷酸和/或核苷酸类似物可在核糖、磷酸和/或碱基部分被修饰。在本发明的实施方案中,指导核酸包含核糖核苷酸和非核糖核苷酸。在一个这样的实施方案中,指导物包含一种或多种核糖核苷酸和一种或多种脱氧核糖核苷酸。在本发明的实施方案中,指导物包含一种或多种非天然存在的核苷酸或核苷酸类似物,诸如具有硫代磷酸酯键联、硼酸磷酸酯键联的核苷酸、包含在核糖环的2'和4'碳原子之间的亚甲基桥的锁定核酸(LNA)或桥接核酸(BNA)。修饰的核苷酸的其他实例包括2'-O-甲基类似

物、2'-脱氧类似物、2-硫代尿苷类似物、N6-甲基腺苷类似物或2'-氟类似物。修饰的碱基的其他实例包括但不限于2-氨基嘌呤、5-溴-尿苷、假尿苷(Ψ)、N¹-甲基假尿苷($me^1\Psi$)、5-甲氧基尿苷(5moU)、肌苷、7-甲基鸟苷。指导RNA化学修饰的实例包括但不限于在一个或多个末端核苷酸处并入2'-O-甲基(M)、2'-O-甲基-3'-硫代磷酸酯(MS)、硫代磷酸酯(PS)、S-约束乙基(cEt),或2'-O-甲基-3'-硫代PACE(MSP)。此类化学修饰的指导物与未修饰的指导物相比可以包含增加的稳定性和增加的活性,不过中靶对脱靶特异性不可预测。(参见Hendel,2015,Nat Biotechnol.33(9):985-9,doi:10.1038/nbt.3290,2015年6月29日在线发布;Ragdarm等人,0215,PNAS,E7110-E7111;Allerson等人,J.Med.Chem.2005,48:901-904;Bramsen等人,Front.Genet.,2012,3:154;Deng等人,PNAS,2015,112:11870-11875;Sharma等人,MedChemComm.,2014,5:1454-1471;Hendel等人,Nat.Biotechnol.(2015)33(9):985-989;Li等人,Nature Biomedical Engineering,2017,1,0066DOI:10.1038/s41551-017-0066)。在一些实施方案中,指导RNA的5'和/或3'端被包括荧光染料、聚乙二醇、胆固醇、蛋白质或检测标签在内的多种功能性部分修饰。(参见Kelly等人,2016,J.Biotech.233:74-83)。在某些实施方案中,指导物在结合至靶DNA的区域中包含核糖核苷酸,并在结合至Cas9、Cpf1或C2c1的区域中包含一个或多个脱氧核糖核苷酸和/或核苷酸类似物。在本发明的实施方案中,将脱氧核糖核苷酸和/或核苷酸类似物并入工程化的指导结构(诸如但不限于5'端和/或3'端、茎环区和种子区)中。在某些实施方案中,修饰不在茎环区的5'柄(5'-handle)中。指导物的茎环区的5'柄中的化学修饰可能会废除其功能(参见Li等人,Nature Biomedical Engineering,2017,1:0066)。在某些实施方案中,指导物的至少1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个、20个、21个、22个、23个、24个、25个、26个、27个、28个、29个、30个、35个、40个、45个、50个或75个核苷酸经化学修饰。在一些实施方案中,指导物的3'或5'端的3-5个核苷酸经化学修饰。在一些实施方案中,在种子区中仅引入较小的修饰,诸如2'-F修饰。在一些实施方案中,在指导物的3'端引入2'-F修饰。在某些实施方案中,指导物的5'和/或3'端的3至5个核苷酸用2'-O-甲基(M)、2'-O-甲基-3'-硫代磷酸酯(MS)、S-约束乙基(cEt)或2'-O-甲基-3'-硫代PACE(MSP)化学修饰。这样的修饰可以提高基因组编辑效率(参见Hendel等人,Nat.Biotechnol.(2015)33(9):985-989)。在某些实施方案中,指导物的所有磷酸二酯键被硫代磷酸酯(PS)取代以增强基因破坏的水平。在某些实施方案中,指导物的5'和/或3'端的多于5个核苷酸用2'-O-Me、2'-F或S-约束乙基(cEt)化学修饰。这种化学修饰的指导物可以介导增强的基因破坏水平(参见Ragdarm等人,0215,PNAS,E7110-E7111)。在本发明的一个实施方案中,指导物被修饰成在其3'和/或5'端包含化学部分。这样的部分包括但不限于胺、叠氮化物、炔、硫代基、二苯并环辛炔(DBCO)或若丹明。在某些实施方案中,化学部分通过接头诸如烷基链缀合至指导物。在某些实施方案中,修饰的指导物的化学部分可用于将指导物衔接至另一分子,诸如DNA、RNA、蛋白质或纳米粒子。这种化学修饰的指导物可用于识别或富集一般由CRISPR系统编辑的细胞(参见Lee等人,eLife,2017,6:e25312,DOI:10.7554)。

[0115] 在某些实施方案中,如本文所提供的CRISPR系统可以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与DNA的混合物,并且/或者其中多核苷酸包含一种或多种核苷酸类似物。序列可以包含任何结构,包括但不限于原生crRNA的结构,

诸如凸环、发夹或茎环结构。在某些实施方案中,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。

[0116] 在某些实施方案中,利用化学修饰的指导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并入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)。此类化学修饰的指导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,不过中靶对脱靶特异性不可预测。(参见Hendel, 2015, *Nat Biotechnol.* 33(9):985-9, doi:10.1038/nbt.3290, 2015年6月29日在线发布)。化学修饰的指导RNA还包括但不限于具有硫代磷酸酯键的RNA和在核糖环的2'与4'碳之间包含亚甲基桥的锁核酸(LNA)核苷酸。

[0117] 在一些实施方案中,指导序列的长度为约或大于约5个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个、20个、21个、22个、23个、24个、25个、26个、27个、28个、29个、30个、35个、40个、45个、50个、75个或更多个核苷酸。在一些实施方案中,指导序列的长度为小于约75个、50个、45个、40个、35个、30个、25个、20个、15个、12个或更少个核苷酸。优选地,指导序列长度为10至30个核苷酸。可以通过任何适合的测定来评定指导序列引导CRISPR复合物序列特异性地结合至靶序列的能力。举例来说,足以形成CRISPR复合物的CRISPR系统的组分,包括有待测试的指导序列,可以提供给具有相应靶序列的宿主细胞,诸如通过用编码CRISPR序列的组分的载体转染,继而诸如通过Surveyor测定评定靶序列内的优先切割。类似地,可以在试管中通过提供靶序列、CRISPR复合物的组分(包括有待测试的指导序列)和不同于测试指导序列的对照指导序列,以及在测试指导序列与对照指导序列反应之间比较靶序列处的结合或切割速率来评估靶RNA的切割。其他测定可能存在,并且将为本领域技术人员所想到。

[0118] 在一些实施方案中,对指导物的修饰是化学修饰、插入、缺失或拆分。在一些实施方案中,化学修饰包括但不限于并入2'-O-甲基(M)类似物、2'-脱氧类似物、2-硫代尿苷类似物、N6-甲基腺苷类似物、2'-氟类似物、2-氨基嘌呤、5-溴-尿苷、假尿苷(Ψ)、N¹-甲基假尿苷($me^1\Psi$)、5-甲氧基尿苷(5moU)、肌苷、7-甲基鸟苷、2'-O-甲基-3'-硫代磷酸酯(MS)、S-约束乙基(cEt)、硫代磷酸酯(PS)或2'-O-甲基-3'-硫代PACE(MSP)。在一些实施方案中,指导物包含一种或多种硫代磷酸酯修饰。在某些实施方案中,指导物的至少1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个、20个或25个核苷酸经化学修饰。在某些实施方案中,种子区中的一个或多个核苷酸经化学修饰。在某些实施方案中,在3'端的一个或多个核苷酸经化学修饰。在某些实施方案中,5'柄中的核苷酸均未经化学修饰。在一些实施方案中,种子区中的化学修饰是次要修饰,诸如并入2'-氟类似物。在具体实施方案中,种子区的一个核苷酸被2'-氟类似物替代。在一些实施方案中,3'端中的5或10个核苷酸经化学修饰。在Cpf1 CrRNA的3'端处的此类化学修饰提高基因切割效率(参见Li等人, *Nature Biomedical Engineering*, 2017, 1:0066)。在具体实施方案中,3'端中的5个核苷酸被2'-氟类似物替代。在具体实施方案中,3'端中的10个核苷酸被2'-氟类似物替代。在具体实施方案中,3'端中的5个核苷酸被2'-O-甲基(M)类似物替代。

[0119] 在一些实施方案中,指导物的5'柄的环经修饰。在一些实施方案中,指导物的5'柄的环被修饰为具有缺失、插入、拆分或化学修饰。在某些实施方案中,环包含3、4或5个核苷酸。在某些实施方案中,环包含序列UCUU、UUUU、UAUU或UGUU。

[0120] 可以选择指导序列并且因此选择核酸靶向指导RNA以靶向任何靶核酸序列。在形成CRISPR复合物的情形中，“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”是指是或包含靶序列的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所针对的RNA多核苷酸或RNA多核苷酸的一部分。在一些实施方案中，靶序列位于细胞的细胞核或细胞质中。靶序列可以是DNA。靶序列可以是任何RNA序列。在一些实施方案中，靶序列可以是选自由以下组成的组的RNA分子内的序列：信使RNA (mRNA)、前体mRNA、核糖体RNA (rRNA)、转移RNA (tRNA)、微小RNA (miRNA)、小干扰RNA (siRNA)、小细胞核RNA (snRNA)、小细胞核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案中，靶序列可以是选自由mRNA、前体mRNA和rRNA组成的组的RNA分子内的序列。在一些优选的实施方案中，靶序列可以是选自由ncRNA和lncRNA组成的组的RNA分子内的序列。在一些更优选的实施方案中，靶序列可以是mRNA分子或前体mRNA分子内的序列。

[0121] 在某些实施方案中，指导RNA的间隔区长度小于28个核苷酸。在某些实施方案中，指导RNA的间隔区长度为至少18个核苷酸并且小于28个核苷酸。在某些实施方案中，指导RNA的间隔区长度在19个与28个核苷酸之间。在某些实施方案中，指导RNA的间隔区长度在19个与25个核苷酸之间。在某些实施方案中，指导RNA的间隔区长度为20个核苷酸。在某些实施方案中，指导RNA的间隔区长度为23个核苷酸。在某些实施方案中，指导RNA的间隔区长度为25个核苷酸。

[0122] 在某些实施方案中，可以通过在间隔区序列与靶序列之间，包括沿着间隔区/靶标的错配位置引入错配，例如1个或多个错配，诸如1个或2个错配来探索切割效率的调节。举例来说，双重错配越处于中心（即，不是3'或5'），切割效率越受影响。因此，通过选择沿着间隔区的错配位置，可以调节切割效率。作为实例，如果需要小于100%的靶标切割（例如在细胞群体中），那么可以在间隔区序列中引入1个或多个，诸如优选2个在间隔区与靶序列之间的错配。错配位置沿着间隔区越处于中心，切割百分比越低。

[0123] 在某些示例性实施方案中，可以探索切割效率以设计单指导物，所述单指导物可以区分两种或更多种因单核苷酸，诸如单核苷酸多态性 (SNP)、变异或（点）突变而变化的靶标。CRISPR效应子可能对SNP（或其他单核苷酸变异）的灵敏度降低并且继续以某一效率水平切割SNP靶标。因此，对于两种靶标或一组靶标，指导RNA可以被设计成具有与靶标中的一种，即，中靶SNP互补的核苷酸序列。指导RNA被进一步设计成具有合成错配。如本文所用，“合成错配”是指在天然存在的SNP上游或下游，诸如上游或下游至多5个核苷酸，例如上游或下游4个、3个、2个或1个核苷酸，优选地上游或下游至多3个核苷酸，更优选地上游或下游至多2个核苷酸，最优选地上游或下游1个核苷酸（即，邻近SNP）处引入的非天然存在的错配。当CRISPR效应子结合至中靶SNP时，仅单错配将与合成错配一起形成，并且将继续激活CRISPR效应子并产生可检测信号。当指导RNA与脱靶SNP杂交时，将形成两个错配，即来自SNP的错配和合成错配，并且不产生可检测信号。因此，本文所公开的系统可以被设计成区分群体内的SNP。举例来说，所述系统可以用于区分因单个SNP而不同的病原性株系或检测某些疾病特异性SNP，诸如但不限于疾病相关SNP，诸如但不限于癌症相关SNP。

[0124] 在某些实施方案中,指导RNA被设计成使得SNP位于间隔区序列的位置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上(以5'端为起点)。在某些实施方案中,指导RNA被设计成使得SNP位于间隔区序列的位置1、2、3、4、5、6、7、8或9上(以5'端为起点)。在某些实施方案中,指导RNA被设计成使得SNP位于间隔区序列的位置2、3、4、5、6或7上(以5'端为起点)。在某些实施方案中,指导RNA被设计成使得SNP位于间隔区序列的位置3、4、5或6上(以5'端为起点)。在某些实施方案中,指导RNA被设计成使得SNP位于间隔区序列的位置3上(以5'端为起点)。

[0125] 在某些实施方案中,指导RNA被设计成使得错配(例如合成错配,即,除SNP以外的其他突变)位于间隔区序列的位置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上(以5'端为起点)。在某些实施方案中,指导RNA被设计成使得错配位于间隔区序列的位置1、2、3、4、5、6、7、8或9上(以5'端为起点)。在某些实施方案中,指导RNA被设计成使得错配位于间隔区序列的位置4、5、6或7上(以5'端为起点)。在某些实施方案中,指导RNA被设计成使得错配位于间隔区序列的位置5上(以5'端为起点)。

[0126] 在某些实施方案中,指导RNA被设计成使得错配位于SNP上游2个核苷酸(即,一个间插核苷酸)。

[0127] 在某些实施方案中,指导RNA被设计成使得错配位于SNP下游2个核苷酸(即,一个间插核苷酸)。

[0128] 在某些实施方案中,指导RNA被设计成使得错配位于间隔区序列的位置5上(以5'端为起点)并且SNP位于间隔区序列的位置3上(以5'端为起点)。

[0129] 本文所述的实施方案涵盖如本文所论述的在真核细胞中(在体外,即,在分离的真核细胞中)诱导一个或多个核苷酸修饰,包括向细胞递送如本文所论述的载体。一个或多个突变可以包括经由一个或多个指导RNA在细胞的每个靶序列处一个或多个核苷酸的引入、缺失或取代。突变可以包括经由一个或多个指导RNA在所述一个或多个细胞的每个靶序列处1-75个核苷酸的引入、缺失或取代。突变可以包括经由一个或多个指导RNA在所述一个或多个细胞的每个靶序列处1个、5个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个、20个、21个、22个、23个、24个、25个、26个、27个、28个、29个、30个、35个、40个、45个、50个或75个核苷酸的引入、缺失或取代。突变可以包括经由一个或多个指导RNA在所述一个或多个细胞的每个靶序列处5个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个、20个、21个、22个、23个、24个、25个、26个、27个、28个、29个、30个、35个、40个、45个、50个或75个核苷酸的引入、缺失或取代。突变包括经由一个或多个指导RNA在所述一个或多个细胞的每个靶序列处10个、11个、12个、13个、14个、15个、16个、17个、18个、19个、20个、21个、22个、23个、24个、25个、26个、27个、28个、29个、30个、35个、40个、45个、50个或75个核苷酸的引入、缺失或取代。突变可以包括经由一个或多个指导RNA在所述一个或多个细胞的每个靶序列处20个、21个、22个、23个、24个、25个、26个、27个、28个、29个、30个、35个、40个、45个、50个或75个核苷酸的引入、缺失或取代。突变可以包括经由一个或多个指导RNA在所述一个或多个细胞的每个靶序列处40个、45个、50个、75个、100个、200个、300个、400个或500个核苷酸的引入、缺失或取代。

[0130] 通常,在内源CRISPR系统的情形中,CRISPR复合物(包含杂交至靶序列并且与一种

或多种Cas蛋白质复合的指导序列)的形成使得靶序列中或附近(例如距其1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、20个、50个或更多个碱基对以内)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。

[0131] 关于CRISPR/Cas系统、其组分、及此类组分的递送(包括方法、材料、递送媒介物、载体、粒子、以及其制造和使用(包括关于量和配制品)),连同表达CRISPR-Cas的真核细胞、表达CRISPR-Cas的真核生物(诸如小鼠)的一般信息,参考以下:美国专利号8,999,641、8,993,233、8,697,359、8,771,945、8,795,965、8,865,406、8,871,445、8,889,356、8,889,418、8,895,308、8,906,616、8,932,814和8,945,839;美国专利公布US 2014-0310830(美国申请序列号14/105,031)、US 2014-0287938 A1(美国申请序列号14/213,991)、US 2014-0273234 A1(美国申请序列号14/293,674)、US2014-0273232 A1(美国申请序列号14/290,575)、US 2014-0273231(美国申请序列号14/259,420)、US 2014-0256046 A1(美国申请序列号14/226,274)、US 2014-0248702 A1(美国申请序列号14/258,458)、US 2014-0242700A1(美国申请序列号14/222,930)、US 2014-0242699 A1(美国申请序列号14/183,512)、US 2014-0242664 A1(美国申请序列号14/104,990)、US 2014-0234972 A1(美国申请序列号14/183,471)、US 2014-0227787 A1(美国申请序列号14/256,912)、US 2014-0189896 A1(美国申请序列号14/105,035)、US 2014-0186958(美国申请序列号14/105,017)、US 2014-0186919 A1(美国申请序列号14/104,977)、US 2014-0186843 A1(美国申请序列号14/104,900)、US 2014-0179770 A1(美国申请序列号14/104,837)和US 2014-0179006 A1(美国申请序列号14/183,486)、US 2014-0170753(美国申请序列号14/183,429);US 2015-0184139(美国申请序列号14/324,960);14/054,414欧洲专利申请EP 2 771 468(EP13818570.7)、EP 2 764 103(EP13824232.6)和EP 2 784 162(EP14170383.5);以及PCT专利公布W02014/093661(PCT/US2013/074743)、W02014/093694(PCT/US2013/074790)、W02014/093595(PCT/US2013/074611)、W02014/093718(PCT/US2013/074825)、W02014/093709(PCT/US2013/074812)、W02014/093622(PCT/US2013/074667)、W02014/093635(PCT/US2013/074691)、W02014/093655(PCT/US2013/074736)、W02014/093712(PCT/US2013/074819)、W02014/093701(PCT/US2013/074800)、W02014/018423(PCT/US2013/051418)、W02014/204723(PCT/US2014/041790)、W02014/204724(PCT/US2014/041800)、W02014/204725(PCT/US2014/041803)、W02014/204726(PCT/US2014/041804)、W02014/204727(PCT/US2014/041806)、W02014/204728(PCT/US2014/041808)、W02014/204729(PCT/US2014/041809)、W02015/089351(PCT/US2014/069897)、W02015/089354(PCT/US2014/069902)、W02015/089364(PCT/US2014/069925)、W02015/089427(PCT/US2014/070068)、W02015/089462(PCT/US2014/070127)、W02015/089419(PCT/US2014/070057)、W02015/089465(PCT/US2014/070135)、W02015/089486(PCT/US2014/070175)、W02015/058052(PCT/US2014/061077)、W02015/070083(PCT/US2014/064663)、W02015/089354(PCT/US2014/069902)、W02015/089351(PCT/US2014/069897)、W02015/089364(PCT/US2014/069925)、W02015/089427(PCT/US2014/070068)、W02015/089473(PCT/US2014/070152)、W02015/089486(PCT/US2014/070175)、W02016/049258(PCT/US2015/051830)、W02016/094867(PCT/US2015/065385)、W02016/094872(PCT/US2015/065393)、W02016/094874(PCT/US2015/065396)、W02016/106244(PCT/US2015/067177)。

[0132] 还提到了2015年6月17日提交的美国申请62/180,709,受保护的指导RNA (PROTECTED GUIDE RNAs (PGRNAS));2014年12月12日提交的美国申请62/091,455,受保护的指导RNA (PROTECTED GUIDE RNAs (PGRNAS));2014年12月24日提交的美国申请62/096,708,受保护的指导RNA (PROTECTED GUIDE RNAs (PGRNAS));2014年12月12日提交的美国申请62/091,462、2014年12月23日提交的美国申请62/096,324、2015年6月17日提交的美国申请62/180,681和2015年10月5日提交的美国申请62/237,496,CRISPR转录因子的死亡指导物 (DEAD GUIDES FOR CRISPR TRANSCRIPTION FACTORS);2014年12月12日提交的美国申请62/091,456和2015年6月17日提交的美国申请62/180,692,用于CRISPR-CAS系统的受护航的和官能化的指导物 (ESCORTED AND FUNCTIONALIZED GUIDES FOR CRISPR-CAS SYSTEMS);2014年12月12日提交的美国申请62/091,461,用于关于造血干细胞 (HSC) 的基因组编辑的CRISPR-CAS系统和组合物的递送、用途以及治疗应用 (DELIVERY, USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS FOR GENOME EDITING AS TO HEMATOPOETIC STEM CELLS (HSCs));2014年12月19日提交的美国申请62/094,903,通过基因组范围的插入捕获测序对双链断裂和基因组重排的无偏鉴定 (UNBIASED IDENTIFICATION OF DOUBLE-STRAND BREAKS AND GENOMIC REARRANGEMENT BY GENOME-WISE INSERT CAPTURE SEQUENCING);2014年12月24日提交的美国申请62/096,761,用于序列操纵的系统、方法及优化的酶以及指导物支架的工程化 (ENGINEERING OF SYSTEMS, METHODS AND OPTIMIZED ENZYME AND GUIDE SCAFFOLDS FOR SEQUENCE MANIPULATION);2014年12月30日提交的美国申请62/098,059、2015年6月18日提交的美国申请62/181,641和2015年6月18日提交的美国申请62/181,667,RNA靶向系统 (RNA-TARGETING SYSTEM);2014年12月24日提交的美国申请62/096,656和2015年6月17日提交的62/181,151,具有稳定化结构域或与稳定化结构域相关联的CRISPR (CRISPR HAVING OR ASSOCIATED WITH DESTABILIZATION DOMAINS);2014年12月24日提交的美国申请62/096,697,具有AAV或与AAV相关联的CRISPR (CRISPR HAVING OR ASSOCIATED WITH AAV);2014年12月30日提交的美国申请62/098,158,工程化CRISPR复合物插入靶向系统 (ENGINEERED CRISPR COMPLEX INSERTIONAL TARGETING SYSTEMS);2015年4月22日提交的美国申请62/151,052,用于胞外核外报告的细胞靶向 (CELLULAR TARGETING FOR EXTRACELLULAR EXOSOMAL REPORTING);2014年9月24日提交的美国申请62/054,490,使用粒子递送组分靶向病症和疾病的CRISPR-CAS系统和组合物的递送、用途以及治疗应用 (DELIVERY, USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS FOR TARGETING DISORDERS AND DISEASES USING PARTICLE DELIVERY COMPONENTS);2014年2月12日提交的美国申请61/939,154,用于用优化的功能性CRISPR-CAS系统进行序列操纵的系统、方法和组合物 (SYSTEMS, METHODS AND COMPOSITIONS FOR SEQUENCE MANIPULATION WITH OPTIMIZED FUNCTIONAL CRISPR-CAS SYSTEMS);2014年9月25日提交的美国申请62/055,484,用于用优化的功能性CRISPR-CAS系统进行序列操纵的系统、方法和组合物 (SYSTEMS, METHODS AND COMPOSITIONS FOR SEQUENCE MANIPULATION WITH OPTIMIZED FUNCTIONAL CRISPR-CAS SYSTEMS);2014年12月4日提交的美国申请62/087,537,用于用优化的功能性CRISPR-CAS系统进行序列操纵的系统、方法和组合物 (SYSTEMS, METHODS AND COMPOSITIONS FOR SEQUENCE MANIPULATION WITH OPTIMIZED FUNCTIONAL CRISPR-CAS

SYSTEMS);2014年9月24日提交的美国申请62/054,651,用于对体内多种癌症突变的竞争建模的CRISPR-CAS系统和组合物的递送、用途以及治疗应用(DELIVERY,USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS FOR MODELING COMPETITION OF MULTIPLE CANCER MUTATIONS IN VIVO);2014年10月23日提交的美国申请62/067,886,用于对体内多种癌症突变的竞争建模的CRISPR-CAS系统和组合物的递送、用途以及治疗应用(DELIVERY,USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS FOR MODELING COMPETITION OF MULTIPLE CANCER MUTATIONS IN VIVO);2014年9月24日提交的美国申请62/054,675和2015年6月17日提交的美国申请62/181,002,在神经元细胞/组织中CRISPR-CAS系统和组合物的递送、用途以及治疗应用(DELIVERY,USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS IN NEURONAL CELLS/TISSUES);2014年9月24日提交的美国申请62/054,528,在免疫疾病或病症中CRISPR-CAS系统和组合物的递送、用途以及治疗应用(DELIVERY,USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS IN IMMUNE DISEASES OR DISORDERS);2014年9月25日提交的美国申请62/055,454,用于使用细胞穿透肽(CPP)靶向病症和疾病的CRISPR-CAS系统和组合物的递送、用途以及治疗应用(DELIVERY,USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS FOR TARGETING DISORDERS AND DISEASES USING CELL PENETRATION PEPTIDES(CPP));2014年9月25日提交的美国申请62/055,460,多功能性CRISPR复合物和/或优化的酶连接的功能性CRISPR复合物(MULTIFUNCTIONAL-CRISPR COMPLEXES AND/OR OPTIMIZED ENZYME LINKED FUNCTIONAL-CRISPR COMPLEXES);2014年12月4日提交的美国申请62/087,475和2015年6月18日提交的62/181,690,用优化的功能性CRISPR-CAS系统进行功能性筛选(FUNCTIONAL SCREENING WITH OPTIMIZED FUNCTIONAL CRISPR-CAS SYSTEMS);2014年9月25日提交的美国申请62/055,487,用优化的功能性CRISPR-CAS系统进行功能性筛选(FUNCTIONAL SCREENING WITH OPTIMIZED FUNCTIONAL CRISPR-CAS SYSTEMS);2014年12月4日提交的美国申请62/087,546和2015年6月18日提交的62/181,687,多功能性CRISPR复合物和/或优化的酶连接的功能性CRISPR复合物(MULTIFUNCTIONAL CRISPR COMPLEXES AND/OR OPTIMIZED ENZYME LINKED FUNCTIONAL-CRISPR COMPLEXES);2014年12月30日提交的美国申请62/098,285,对肿瘤生长和转移的CRISPR介导的体内建模以及遗传筛选(CRISPR MEDIATED IN VIVO MODELING AND GENETIC SCREENING OF TUMOR GROWTH AND METASTASIS)。

[0133] 提到了2015年6月18日提交的美国申请62/181,659和2015年8月19日提交的美国申请62/207,318,用于序列操纵的CAS9直系同源物和变体的系统、方法、酶和指导物支架的工程化和优化(ENGINEERING AND OPTIMIZATION OF SYSTEMS,METHODS,ENZYME AND GUIDE SCAFFOLDS OF CAS9 ORTHOLOGS AND VARIANTS FOR SEQUENCE MANIPULATION)。提到了2015年6月18日提交的美国申请62/181,663和2015年10月22日提交的美国申请62/245,264,新颖的CRISPR酶和系统(NOVEL CRISPR ENZYMES AND SYSTEMS);2015年6月18日提交的美国申请62/181,675、2015年10月22日提交的美国申请62/285,349、2016年2月17日提交的62/296,522和2016年4月8日提交的美国申请62/320,231,新颖的CRISPR酶和系统(NOVEL CRISPR ENZYMES AND SYSTEMS);2015年9月24日提交的美国申请62/232,067、2015年12月

18日提交的美国申请14/975,085、欧洲申请号16150428.7、2015年8月16日提交的美国申请62/205,733、2015年8月5日提交的美国申请62/201,542、2015年7月16日提交的美国申请62/193,507和2015年6月18日提交的美国申请62/181,739,其各自的标题为新颖的CRISPR酶和系统(NOVEL CRISPR ENZYMES AND SYSTEMS);以及2015年10月22日提交的美国申请62/245,270,新颖的CRISPR酶和系统(NOVEL CRISPR ENZYMES AND SYSTEMS)。还提到了2014年12月12日提交的美国申请61/939,256和2014年12月12日提交的WO 2015/089473(PCT/US2014/070152),其各自的标题为用于序列操纵的具有新架构的系统、方法和优化的指导物组合物的工程化(ENGINEERING OF SYSTEMS, METHODS AND OPTIMIZED GUIDE COMPOSITIONS WITH NEW ARCHITECTURES FOR SEQUENCE MANIPULATION)。还提到了2015年8月15日提交的PCT/US2015/045504、2015年6月17日提交的美国申请62/180,699和2014年8月17日提交的美国申请62/038,358,其各自的标题为使用CAS9切口酶进行基因组编辑(GENOME EDITING USING CAS9 NICKASES)。

[0134] 这些专利、专利公布和申请的每一者,以及在其中或在它们的审查程序期间引用的所有文献(“申请引用文献”)以及在这些申请引用文献中引用或参考的所有文献,连同其中提到的或在其中任何文献中提到并以引用方式并入本文的针对任何产品的任何说明书、说明、产品规格和产品表,特此以引用方式并入本文,并且可以在本发明的实践中采用。所有文献(例如这些专利、专利公布和申请以及申请引用文献)在如同每个单独文献被确切地且单独地指明为以引用方式并入的相同程度上以引用方式并入本文。

[0135] 用于检测靶核酸序列的系统

[0136] 在某些示例性实施方案中,本发明提供了一种用于检测样品中的靶核酸序列的系统。所述系统可以包括如本文所述的转座子复合物。如本文所述,转座子复合物可以包含转座酶和包含一个或多个T7 RNA聚合酶启动子的转座子序列。转座酶可以将一个或多个T7 RNA聚合酶启动子插入包含靶核酸序列的寡核苷酸中。

[0137] 在一些实施方案中,所述系统还可以包括T7 RNA聚合酶。如本文所述,T7 RNA聚合酶可以由插入的T7启动子之一转录靶核酸。

[0138] 所述系统还可以包括如本文所述的用于检测转录的靶核酸的基于CRISPR的检测系统。在一些实施方案中,基于CRISPR的检测系统可以是基于CRISPR-Cas13的检测系统。基于CRISPR Cas13的检测系统可以是CRISPR-SHERLOCK检测系统。在一些实施方案中,基于CRISPR的检测系统可以是如本文所述的基于CRISPR-Cas12的检测系统。

[0139] 在某些示例性实施方案中,转座酶可以是Tn5转座酶或其变体。转座酶可以是如本文所述的工程化的转座酶。在某些示例性实施方案中,转座子序列可以包含如本文所述的两个19个碱基对的ME Tn5转座酶识别序列。

[0140] 在某些示例性实施方案中,CRISPR-SHERLOCK系统可以是如本文所述的基于Cas13的CRISPR-SHERLOCK系统。

[0141] 在某些示例性实施方案中,用于检测靶核酸序列的系统还可以包含如本文所述的用于核酸扩增的试剂。

[0142] 在某些示例性实施方案中,用于检测靶核酸序列的系统还可以包含用于逆转录包含RNA的靶序列的试剂。

[0143] ***

[0144] 在不脱离本发明的范围和精神的情况下,本发明所述的方法、药物组合物和药盒的各种修改和变型对于本领域技术人员而言是显而易见的。尽管已结合具体实施方案描述了本发明,但是应当理解,本发明能够进行进一步修改,并且所要求保护的本发明不应该不当地受到这些具体实施方案的限制。实际上,对于本领域技术人员而言显而易见的用于实施本发明的所述模式的各种变型旨在落入本发明的范围内。本申请旨在覆盖本发明的任何变型、用途或改型,这些变型、用途或改型总体遵循本发明的原理且包括在本发明所属领域内已知习惯性实践内的与本公开内容有所偏离的内容,并且可以在陈述之前应用于本文的基本特征。

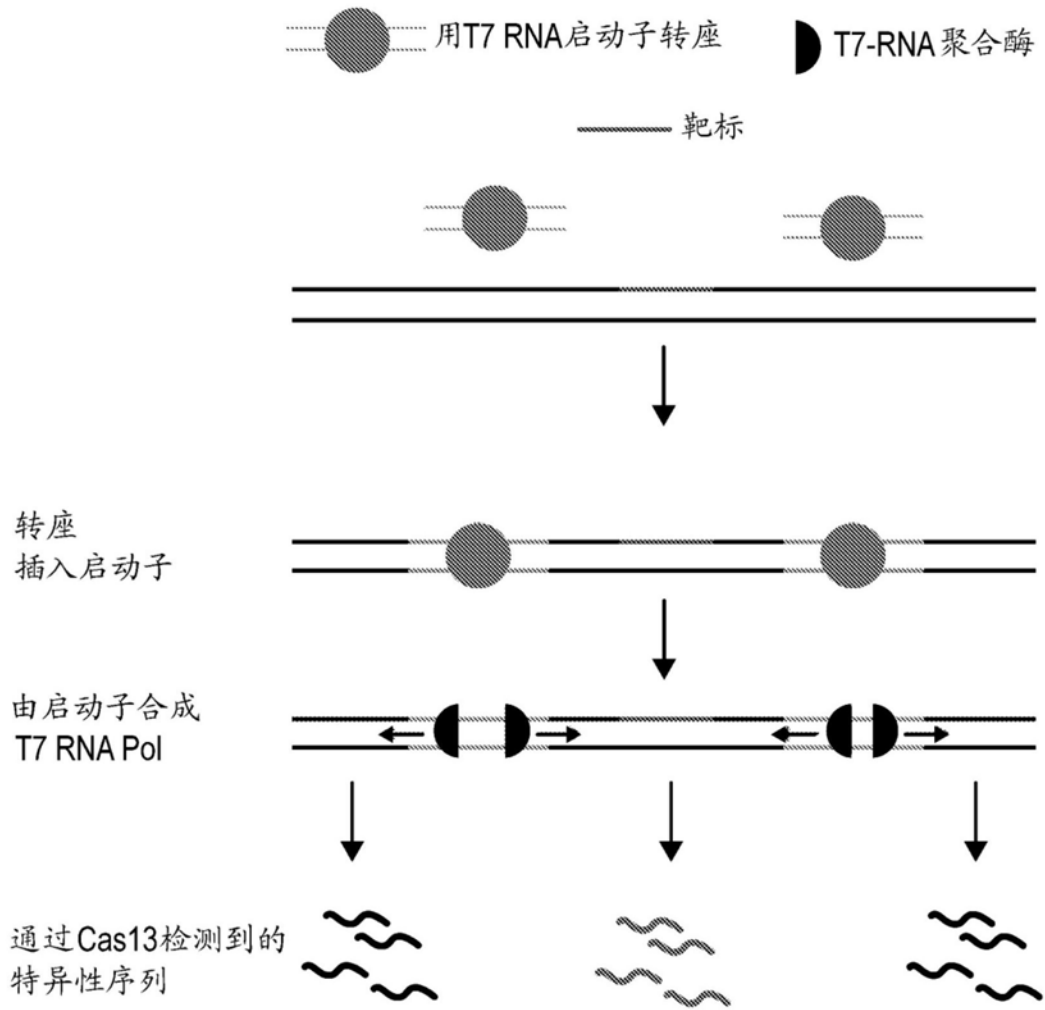


图1

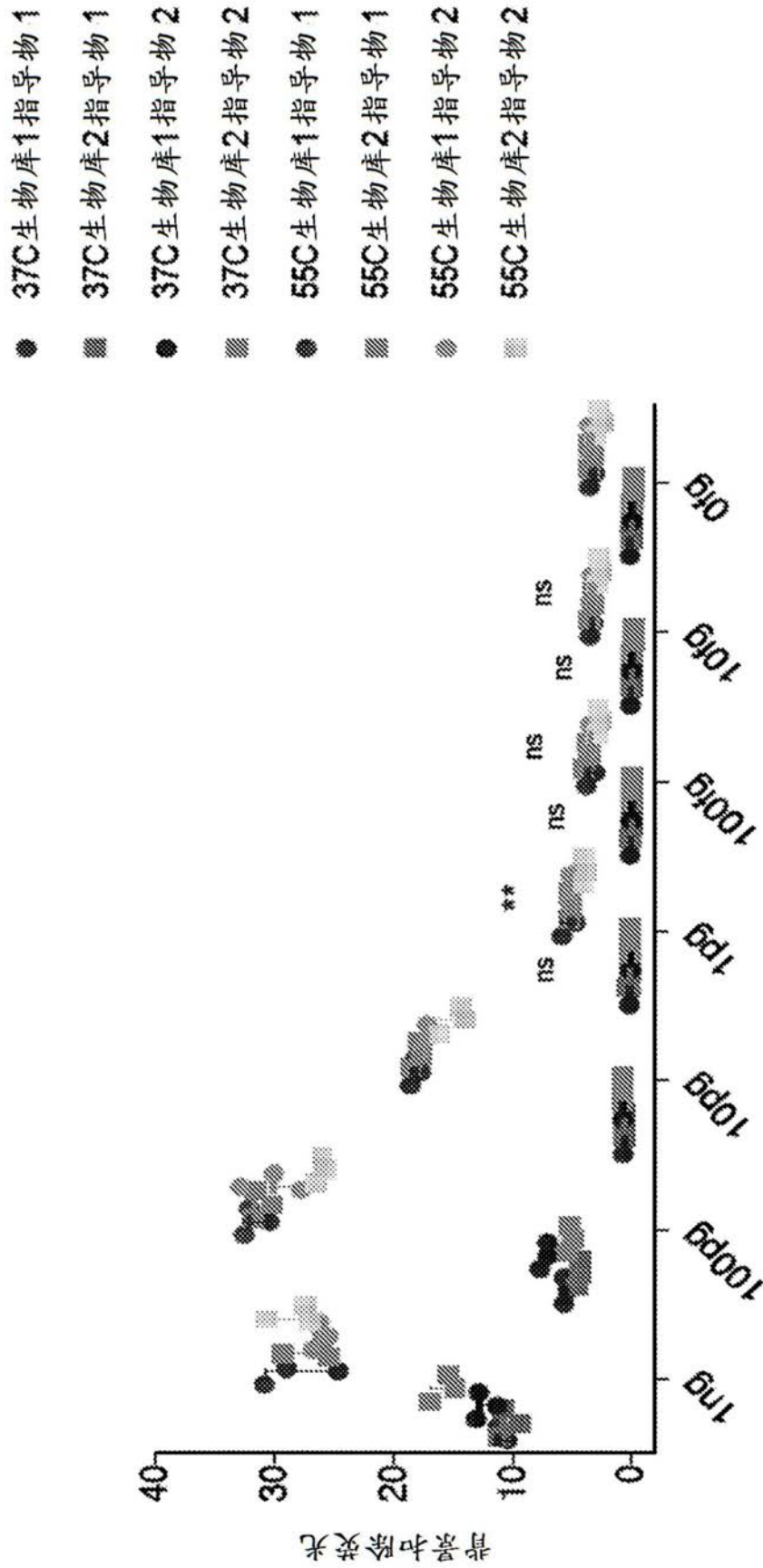


图2