



(12) 发明专利申请

(10) 申请公布号 CN 114174508 A

(43) 申请公布日 2022. 03. 11

(21) 申请号 202080048714.3

(22) 申请日 2020.05.01

(30) 优先权数据

62/841,770 2019.05.01 US

(85) PCT国际申请进入国家阶段日

2021.12.31

(86) PCT国际申请的申请数据

PCT/US2020/031030 2020.05.01

(87) PCT国际申请的公布数据

W02020/223634 EN 2020.11.05

(71) 申请人 猛犸生物科学公司

地址 美国加利福尼亚州

(72) 发明人 卢卡斯·本杰明·哈林顿

贾尼斯·沙·陈

伊萨克·帕特森·维特

(74) 专利代理机构 北京安信方达知识产权代理

有限公司 11262

代理人 王玮玮 郑霞

(51) Int.Cl.

C12N 15/10 (2006.01)

C12N 9/22 (2006.01)

C12N 15/90 (2006.01)

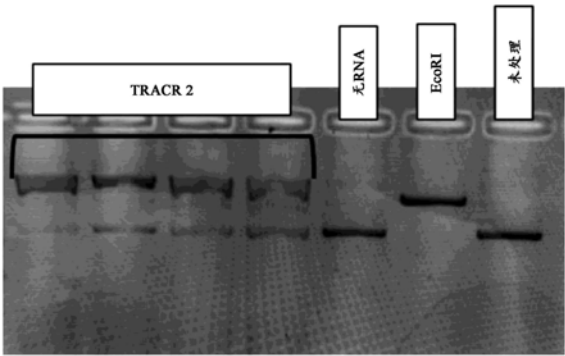
权利要求书3页 说明书86页 附图6页

(54) 发明名称

可编程核酸酶和使用方法

(57) 摘要

本文提供了可编程核酸酶和使用所述可编程核酸酶进行基因组编辑和核酸检测的方法。在一些实施方案中,该可编程核酸酶是可编程切口酶。



1. 一种在靶核酸中引入断裂的方法,所述方法包括通过使所述靶核酸与以下(a)和(b)接触以引入所述断裂:

(a) 第一指导核酸,所述第一指导核酸包含与长度不超过900个氨基酸的第一可编程切口酶结合的第一区域;

(b) 第二指导核酸,所述第二指导核酸包含与长度不超过900个氨基酸的第二可编程切口酶结合的第一区域,

其中所述第一指导核酸包含与所述靶核酸结合的第二区域,并且其中所述第二指导核酸包含与所述靶核酸结合的第二区域,并且其中所述第一指导核酸的第二区域和所述第二指导核酸的第二区域与所述靶核酸的相反链结合。

2. 根据权利要求1所述的方法,其中所述第一可编程切口酶和所述第二可编程切口酶的长度为350至900个氨基酸。

3. 根据权利要求1-2中任一项所述的方法,其中所述第一可编程切口酶和所述第二可编程切口酶的长度为480至550个氨基酸。

4. 根据权利要求1-3任一项所述的方法,其中所述第一可编程切口酶和所述第二可编程切口酶是V型CRISPR/Cas酶。

5. 根据权利要求4所述的方法,其中所述V型CRISPR/Cas酶包含三个部分RuvC结构域。

6. 根据权利要求5所述的方法,其中所述三个部分RuvC结构域为RuvC-I、RuvC-II和RuvC-III亚结构域。

7. 根据权利要求1-6中任一项所述的方法,其中所述第一可编程切口酶和所述第二可编程切口酶是Cas14蛋白。

8. 根据权利要求7所述的方法,其中所述Cas14蛋白是Cas14a蛋白、Cas14b蛋白、Cas14c蛋白、Cas14d蛋白或Cas14e蛋白。

9. 根据权利要求7-8中任一项所述的方法,其中所述Cas14蛋白是Cas14a蛋白。

10. 根据权利要求7-8中任一项所述的方法,其中所述Cas14蛋白是Cas14b蛋白。

11. 根据权利要求7-8中任一项所述的方法,其中所述Cas14蛋白是Cas14e蛋白。

12. 根据权利要求1-11中任一项所述的方法,其中所述第一可编程切口酶、所述第二可编程切口酶或两者与SEQ ID NO:1-SEQ ID NO:91或SEQ ID NO:170中的任一项具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。

13. 根据权利要求1-12中任一项所述的方法,其中所述第一可编程切口酶、所述第二可编程切口酶或两者是SEQ ID NO:1-SEQ ID NO:91或SEQ ID NO:170中的任一项。

14. 根据权利要求1-13中任一项所述的方法,其中所述第一可编程切口酶、所述第二可编程切口酶或两者与SEQ ID NO:1具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。

15. 根据权利要求1-14中任一项所述的方法,其中所述第一可编程切口酶、所述第二可编程切口酶或两者是SEQ ID NO:1。

16. 根据权利要求1-13中任一项所述的方法,其中所述第一可编程切口酶、所述第二可编程切口酶或两者与SEQ ID NO:10具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。

17. 根据权利要求1-13或16中任一项所述的方法,其中所述第一可编程切口酶、所述第

二可编程切口酶或两者是SEQ ID NO:10。

18. 根据权利要求1-13中任一项所述的方法,其中所述第一可编程切口酶、所述第二可编程切口酶或两者与SEQ ID NO:11具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。

19. 根据权利要求1-13或18中任一项所述的方法,其中所述第一可编程切口酶、所述第二可编程切口酶或两者是SEQ ID NO:11。

20. 根据权利要求1-13中任一项所述的方法,其中所述第一可编程切口酶、所述第二可编程切口酶或两者与SEQ ID NO:17具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。

21. 根据权利要求1-13或20中任一项所述的方法,其中所述第一可编程切口酶、所述第二可编程切口酶或两者是SEQ ID NO:17。

22. 根据权利要求1-13中任一项所述的方法,其中所述第一可编程切口酶、所述第二可编程切口酶或两者与SEQ ID NO:33具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。

23. 根据权利要求1-13或22中任一项所述的方法,其中所述第一可编程切口酶、所述第二可编程切口酶或两者是SEQ ID NO:33。

24. 根据权利要求1-23中任一项所述的方法,其中所述第一指导核酸是第一指导RNA。

25. 根据权利要求1-24中任一项所述的方法,其中所述第二指导核酸是第二指导RNA。

26. 根据权利要求1-25中任一项所述的方法,其中所述第一区域是重复序列并且其中所述第二区域是间隔区序列。

27. 根据权利要求1-26中任一项所述的方法,其中所述第一指导核酸和所述第二指导核酸包含crRNA和tracrRNA。

28. 根据权利要求1-26中任一项所述的方法,其中所述第一指导核酸和所述第二指导核酸包含crRNA和trancRNA。

29. 根据权利要求27-28中任一项所述的方法,其中所述crRNA包含所述重复序列和所述间隔区序列。

30. 根据权利要求26-29中任一项所述的方法,其中所述重复序列与所述tracrRNA的区段杂交。

31. 根据权利要求27-30中任一项所述的方法,其中所述tracrRNA与SEQ ID NO:98-SEQ ID NO:99、SEQ ID NO:101、SEQ ID NO:103、SEQ ID NO:105-SEQ ID NO:151中的任一项具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。

32. 根据权利要求27-31中任一项所述的方法,其中所述tracrRNA是SEQ ID NO:98-SEQ ID NO:99、SEQ ID NO:101、SEQ ID NO:103、SEQ ID NO:105-SEQ ID NO:151中的任一项。

33. 根据权利要求27-31中任一项所述的方法,其中所述tracrRNA与SEQ ID NO:99具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。

34. 根据权利要求27-31或33中任一项所述的方法,其中所述tracrRNA是SEQ ID NO:99。

35. 根据权利要求27-31中任一项所述的方法,其中所述tracrRNA与SEQ ID NO:101具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。

36. 根据权利要求27-31或35中任一项所述的方法, 其中所述tracrRNA是SEQ ID NO: 101。

37. 根据权利要求27-31中任一项所述的方法, 其中所述tracrRNA与SEQ ID NO: 103具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。

38. 根据权利要求27-31或37中任一项所述的方法, 其中所述tracrRNA是SEQ ID NO: 103。

39. 根据权利要求27-31中任一项所述的方法, 其中所述tracrRNA与SEQ ID NO: 119具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。

40. 根据权利要求27-31或39中任一项所述的方法, 其中所述tracrRNA是SEQ ID NO: 119。

41. 根据权利要求1-40中任一项所述的方法, 其中所述第一可编程切口酶和所述第二可编程切口酶表现出比双链切割活性高2倍的产生切口活性。

42. 根据权利要求1-41中任一项所述的方法, 其中所述第一可编程切口酶和所述第二可编程切口酶在两个不同位点使所述靶核酸产生切口。

43. 根据权利要求1-42中任一项所述的方法, 其中所述靶核酸包含双链DNA。

44. 根据权利要求43所述的方法, 其中所述两个不同位点在所述双链DNA的相反链上。

45. 根据权利要求1-44中任一项所述的方法, 其中所述靶核酸包含已突变的序列或与疾病相关的序列。

46. 根据权利要求45所述的方法, 其中所述疾病是癌症。

47. 根据权利要求1-46中任一项所述的方法, 其中所述方法包括向有需要的受试者施用所述第一可编程切口酶和所述第二可编程切口酶。

48. 根据权利要求45所述的方法, 其中在所述第一可编程切口酶和所述第二可编程切口酶使所述靶核酸产生切口之后移除所述已突变的序列。

49. 根据权利要求1-48中任一项所述的方法, 其中所述第一可编程切口酶和所述第二可编程切口酶相同。

50. 根据权利要求1-49中任一项所述的方法, 其中所述第一可编程切口酶和所述第二可编程切口酶不同。

可编程核酸酶和使用方法

交叉参考

[0001] 本申请要求于2019年5月1日提交的美国临时专利申请号62/841,770的优先权,其通过引用以其全文并入本文。

背景技术

[0002] 某些可编程核酸酶可用于核酸序列的基因组编辑或核酸序列的检测。需要能够在各种样品条件下工作并且可用于基因组编辑和诊断的高效、可编程切口酶。

发明内容

[0003] 在各个方面,本公开内容提供了一种在靶核酸中引入断裂的方法,所述方法包括通过使所述靶核酸与以下(a)和(b)接触来引入所述断裂:(a)第一指导核酸,所述第一指导核酸包含与长度不超过900个氨基酸的第一可编程切口酶结合的第一区域;(b)第二指导核酸,所述第二指导核酸包含与长度不超过900个氨基酸的第二可编程切口酶结合的第一区域,其中所述第一指导核酸包含与所述靶核酸结合的第二区域并且其中所述第二指导核酸包含与所述靶核酸结合的第二区域并且其中所述第一指导核酸的第二区域和所述第二指导核酸的第二区域与所述靶核酸的相反链结合。在一些方面,所述第一可编程切口酶和所述第二可编程切口酶的长度为350至900个氨基酸。在一些方面,所述第一可编程切口酶和所述第二可编程切口酶的长度为480至550个氨基酸。

[0004] 在一些方面,所述第一可编程切口酶和所述第二可编程切口酶是V型CRISPR/Cas酶。在一些方面,所述V型CRISPR/Cas酶包含三个部分RuvC结构域。在一些方面,所述三个部分RuvC结构域为RuvC-I、RuvC-II和RuvC-III亚结构域。在一些方面,其中所述第一可编程切口酶和所述第二可编程切口酶是Cas14蛋白。在一些方面,所述Cas14蛋白是Cas14a蛋白、Cas14b蛋白、Cas14c蛋白、Cas14d蛋白或Cas14e蛋白。在一些方面,所述Cas14蛋白是Cas14a蛋白。在一些方面,所述Cas14蛋白是Cas14b蛋白。在一些方面,所述Cas14蛋白是Cas14e蛋白。

[0005] 在一些方面,所述第一可编程切口酶、所述第二可编程切口酶或两者与SEQ ID NO:1-SEQ ID NO:91或SEQ ID NO:170中的任一项具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。在一些方面,所述第一可编程切口酶、所述第二可编程切口酶或两者是SEQ ID NO:1-SEQ ID NO:91或SEQ ID NO:170中的任一项。

[0006] 在一些方面,所述第一可编程切口酶、所述第二可编程切口酶或两者与SEQ ID NO:1具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。在一些方面,所述第一可编程切口酶、所述第二可编程切口酶或两者是SEQ ID NO:1。

[0007] 在一些方面,所述第一可编程切口酶、所述第二可编程切口酶或两者与SEQ ID NO:10具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。在一些方面,所述第一可编程切口酶、所述第二可编程切口酶或两者是SEQ ID NO:10。

[0008] 在一些方面,所述第一可编程切口酶、所述第二可编程切口酶或两者与SEQ ID NO:11具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。在一些方面,所述第一可编程切口酶、所述第二可编程切口酶或两者是SEQ ID NO:11。

[0009] 在一些方面,所述第一可编程切口酶、所述第二可编程切口酶或两者与SEQ ID NO:17具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。在一些方面,所述第一可编程切口酶、所述第二可编程切口酶或两者是SEQ ID NO:17。

[0010] 在一些方面,所述第一可编程切口酶、所述第二可编程切口酶或两者与SEQ ID NO:33具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。在一些方面,所述第一可编程切口酶、所述第二可编程切口酶或两者是SEQ ID NO:33。

[0011] 在一些方面,所述第一指导核酸是第一指导RNA。在一些方面,所述第二指导核酸是第二指导RNA。在一些方面,所述第一区域是重复序列并且其中所述第二区域是间隔区序列。在一些方面,所述第一指导核酸和所述第二指导核酸包含crRNA和tracrRNA。在一些方面,所述第一指导核酸和所述第二指导核酸包含crRNA和trancRNA。在一些方面,所述crRNA包含所述重复序列和所述间隔区序列。在一些方面,所述重复序列与所述tracrRNA的区段杂交。

[0012] 在一些方面,所述tracrRNA与SEQ ID NO:98-SEQ ID NO:99、SEQ ID NO:101、SEQ ID NO:103、SEQ ID NO:105-SEQ ID NO:151中的任一项具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。在一些方面,所述tracrRNA是SEQ ID NO:98-SEQ ID NO:99、SEQ ID NO:101、SEQ ID NO:103、SEQ ID NO:105-SEQ ID NO:151中的任一项。

[0013] 在一些方面,所述tracrRNA与SEQ ID NO:99具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。在一些方面,所述tracrRNA是SEQ ID NO:99。在一些方面,所述tracrRNA与SEQ ID NO:101具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。在一些方面,所述tracrRNA是SEQ ID NO:101。在一些方面,所述tracrRNA与SEQ ID NO:103具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。在一些方面,所述tracrRNA是SEQ ID NO:103。在一些方面,所述tracrRNA与SEQ ID NO:119具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。在一些方面,所述tracrRNA是SEQ ID NO:119。

[0014] 在一些方面,,所述第一可编程切口酶和所述第二可编程切口酶表现出比双链切割活性高2倍的产生切口活性。在一些方面,所述第一可编程切口酶和所述第二可编程切口酶在两个不同位点使所述靶核酸产生切口。在一些方面,所述靶核酸包含双链DNA。在一些方面,所述两个不同位点在所述双链DNA的相反链上。在一些方面,所述靶核酸包含已突变的序列或与疾病相关的序列。在一些方面,所述疾病是癌症。

[0015] 在一些方面,所述方法包括向有需要的受试者施用所述第一可编程切口酶和所述第二可编程切口酶。在一些方面,在所述第一可编程切口酶和所述第二可编程切口酶使所

述靶核酸产生切口之后移除所述已突变的序列。在一些方面,所述第一可编程切口酶和所述第二可编程切口酶相同。在一些方面,所述第一可编程切口酶和所述第二可编程切口酶不同。在各个方面,本公开内容提供了一种在靶核酸中引入链断裂的方法,所述方法包括通过使所述靶核酸与以下(a)和(b)接触来引入所述链断裂:(a)第一指导RNA,所述第一指导RNA包含与第一可编程切口酶结合的第一区域;(b)第二指导RNA,所述第二指导RNA包含与第二可编程切口酶结合的第一区域,其中所述第一指导RNA包含与所述靶核酸结合的第二区域并且其中所述第二指导RNA包含与所述靶核酸结合的第二区域并且其中所述第一指导RNA的第二区域和所述第二指导RNA的第二区域与所述靶核酸的相反链结合。

[0016] 在一些方面,所述第一可编程切口酶和所述第二可编程切口酶在两个不同的位点使所述靶核酸产生切口。在一些方面,所述靶核酸包含双链DNA。在一些方面,所述两个不同的位点在所述双链DNA的相反链上。在一些方面,所述靶核酸包含已突变的序列或与疾病相关的序列。在进一步的方面,所述疾病是癌症。

[0017] 在一些方面,所述方法包括向有需要的受试者施用所述第一可编程切口酶和所述第二可编程切口酶。在一些方面,在所述第一可编程核酸酶和所述第二可编程核酸酶使所述靶核酸产生切口之后移除所述已突变的序列。在一些方面,所述第一可编程切口酶和所述第二可编程切口酶包含Cas14蛋白。

[0018] 在各个方面,本公开内容提供了一种检测样品中靶核酸的方法,所述方法包括使所述样品与以下(a)、(b)和(c)接触:(a)可编程切口酶;(b)指导RNA,所述指导RNA包含与所述可编程切口酶结合的第一区域和与所述靶核酸结合的第二区域;(c)标记的单链DNA报告子,所述标记的单链DNA报告子不与所述指导RNA结合;切割所述标记的单链DNA报告子以释放可检测的标记;以及通过测量来自所述可检测标记的信号检测所述靶核酸。

[0019] 在一些方面,所述靶核酸是单链DNA。在一些方面,所述可编程切口酶包含Cas 14蛋白。在一些方面,所述靶核酸在样品中。在一些方面,所述样品包含磷酸盐缓冲液、Tris缓冲液或HEPES缓冲液。在进一步的方面,所述样品包含7至9的pH值。在进一步的方面,所述样品包含7.5至8的pH值。在一些方面,所述样品包含25nM至200mM的盐浓度。

[0020] 在一些方面,所述单链DNA报告子包含ssDNA-荧光猝灭DNA报告子。在进一步的方面,所述ssDNA-荧光猝灭DNA报告子是通用ssDNA-荧光猝灭DNA报告子。在一些方面,所述可编程切口酶表现出不依赖PAM的产生切口和切割。在一些方面,所述Cas14蛋白包含Cas14e蛋白。在一些方面,所述Cas14蛋白包含400至800个氨基酸残基。

[0021] 在各个方面,本公开内容提供了一种组合物,所述组合物包含可编程切口酶和指导RNA,所述指导RNA包含与所述可编程切口酶结合的第一区域和与所述靶核酸的第二区域结合。

[0022] 在一些方面,所述靶核酸包含单链DNA或双链DNA。在一些方面,所述可编程切口酶表现出不依赖PAM的产生切口和切割。在一些方面,所述可编程切口酶使所述双链DNA的单链产生切口。在一些方面,所述可编程切口酶切割单链DNA。在一些方面,所述可编程切口酶包含Cas14蛋白。在进一步的方面,所述Cas14蛋白包含Cas14e蛋白。在进一步的方面,所述Cas14蛋白包含400至800个氨基酸残基。

援引并入

[0023] 本说明书中提及的所有出版物、专利和专利申请都通过引用并入本文,其程度如

同特别地和单独地指出要通过引用并入每一个出版物、专利或专利申请。

附图说明

[0024] 本公开内容的新颖特征在所附权利要求中具体阐述。通过参考以下对其中利用到本公开内容的原理的说明性实施方案加以阐述的详细描述以及附图,将会获得对本公开内容的特征和优点的更好理解,在这些附图中:

[0025] 图1显示了示出通过可编程切口酶使dsDNA产生切口的凝胶。四种可编程切口酶(这里是四种Cas14e蛋白)与指导RNA (TRACR2) 一起独立地添加至前四个泳道,指导RNA与可编程切口酶形成复合体。当指导RNA与可编程切口酶复合并且该复合体与其靶核酸结合时,可编程切口酶的切口酶活性被激活。这在凝胶的前四个泳道中由所得的两条条带显示,其中上面的条带是有切口的靶dsDNA。第五个泳道是包含可编程切口酶但不含指导RNA的对照泳道,其中靶dsDNA保持完整。第六个泳道显示了限制性内切酶EcoRI对dsDNA的切割,产生双链断裂。第七个泳道显示了未处理的靶dsDNA(例如,无可编程切口酶、指导RNA或限制性内切酶)。

[0026] 图2显示了盐、缓冲液和温度对使用Cas14e的ssDNA DETECTR反应的影响。左上柱状图的x轴显示了各种缓冲条件和pH值,y轴显示了减去背景的荧光。荧光表明报告子的切割。荧光越强表明活性越高。中上和右上图显示了x轴所示的温度(“ON”表示添加可以与指导RNA杂交的靶ssDNA;“OFF”表示添加不与指导RNA杂交的脱靶ssDNA)相对于y轴所示的原始荧光。“OFF”荧光用于确定背景荧光。下面三个线图显示了在各种盐条件下(从左至右为25nM NaCl、100nM NaCl和200mM NaCl)随时间的荧光。荧光表明报告因子的切割。荧光越强表明活性越高。较高的线且荧光随时间增加,显示在存在靶ssDNA的情况下,Cas14e蛋白与指导RNA复合对报告子的切割。较低的线且荧光随时间保持平缓,显示在存在脱靶ssDNA的情况下,Cas14e蛋白与指导RNA复合。

[0027] 图3显示了三个图,这些图从左至右评估了均聚物荧光猝灭(FQ)报告子的切割。荧光表明报告子的切割。荧光越强表明切割活性越高。最左侧的图使用T12(12个胸腺嘧啶残基) ssDNA-FQ报告子,中间的图使用A12(12个腺嘌呤残基) ssDNA-FQ报告子,最右侧的图使用C12(12个胞嘧啶残基) ssDNA-FQ报告子。在每个图中,上线显示了在存在靶ssDNA的情况下与指导RNA复合的Cas14e蛋白,下线显示了在存在脱靶ssDNA的情况下与指导RNA复合的Cas14e蛋白。

[0028] 图4显示了使用与指导RNA偶联的Cas14e蛋白检测靶dsDNA的三个DETECTR反应的荧光随时间变化的图。荧光表明报告子的切割。荧光越强表明活性越高。最上线显示了在存在具有野生型(wt) PAM的靶dsDNA的情况下报告子的切割。最上线正下方的线显示了在存在具有突变(mut) PAM的靶dsDNA的情况下报告子的切割。最下线显示了在存在500nM脱靶ssDNA的情况下报告子的切割。结果显示Cas14e对PAM限制不敏感。

[0029] 图5显示了分别与四种不同的指导核酸复合的四种可编程切口酶的顺式切割活性测定的结果。可编程切口酶与指导核酸靶向的质粒DNA孵育60分钟。该图显示了在60分钟孵育期间产生切口(单链断裂)或双链断裂的质粒百分比,如凝胶电泳所测量的。

[0030] 图6显示了与18、16或15个单独的指导核酸复合的三种不同的可编程切口酶的顺式切割活性测定的结果。可编程切口酶与指导核酸靶向的质粒DNA孵育10分钟。该图显示了

每个可编程切口酶-指导核酸对的表现出切口(单链断裂;“切口”)或双链断裂(“切割”)的质粒百分比。

[0031] 图6A显示了使用Cas14a.3的测定结果。

[0032] 图6B显示了使用Cas14b.4的测定结果。

[0033] 图6C显示了使用Cas14b.10的测定结果。

具体实施方式

[0034] 本公开内容提供了可编程核酸酶的组合物。在一些实施方案中,可编程核酸酶是可编程DNA核酸酶。这些可编程核酸酶可以与指导RNA复合,该指导RNA可以与靶DNA结合。在某些实施方案中,当可编程核酸酶与指导RNA复合并且靶DNA与指导RNA杂交时,ssDNA(例如ssDNA报告子)的反式切割由可编程核酸酶激活。ssDNA反式切割的检测可用于确定样品中的靶DNA。在一些实施方案中,可编程核酸酶是可编程切口酶。在进一步的实施方案中,可编程核酸酶是可编程DNA切口酶。

[0035] 本文公开的可编程切口酶可表现出顺式切割活性或靶标切割活性。靶标切割活性可以指可编程切口酶对靶核酸的切割。在一些情况下,顺式切割活性导致靶核酸中的双链断裂。在某些情况下,顺式切割活性导致靶核酸的单链断裂(切口酶活性)。在一些情况下,顺式切割活性产生靶核酸中双链和单链断裂的混合物。在进一步的情况下,顺式切割双链和单链断裂形成的速率可以依赖于指导核酸的序列。在一些情况下,顺式切割双链和单链断裂形成的比例可以依赖于指导核酸的序列。

用于使靶核酸产生切口和检测靶核酸的试剂

[0036] 许多试剂与本文公开的组合物和方法一致。本文所述的试剂可用于使靶核酸产生切口和用于检测靶核酸。本文公开的试剂可以包括可编程切口酶、指导核酸、靶核酸和缓冲液。如本文所述,可以使用本文公开的可编程切口酶(例如,本文公开的Cas14a、Cas14b或Cas14e)和其他试剂修饰或检测包含DNA或RNA的靶核酸(例如,靶DNA与指导核酸杂交)。如本文所述,可以是目的核酸的扩增子,并且可以使用本文公开的可编程切口酶和其他试剂检测扩增子(例如,靶DNA与指导核酸杂交)包含DNA的靶核酸。此外,使用两种或多种可编程切口酶或者与靶向多个靶核酸的指导核酸复合的(一种)可编程切口酶和(一种)非切口酶可编程核酸酶检测多个靶核酸是可能的,其中可编程核酸酶表现出不同的对报告子的核酸序列非依赖性切割(例如,RNA报告子被第一可编程核酸酶切割,DNA报告子被第二可编程核酸酶切割)。

可编程切口酶

[0037] 在一些实施方案中,本公开内容的可编程切口酶(例如,Cas14)由于其小尺寸而特别适用于基因组编辑和在DETECTR测定中使用。这些蛋白的较小性质使它们在基因组编辑的背景下更容易被包装并以更高的效率被递送,并且在测定中更容易作为试剂掺入。在一些实施方案中,本公开内容的可编程切口酶的长度为400至800个氨基酸残基、长度为400至420个氨基酸残基、长度为420至440个氨基酸残基、长度为440至460个氨基酸残基、长度为460至480个氨基酸残基、长度为480至500个氨基酸残基、长度为500至520个氨基酸残基、长度为520至540个氨基酸残基、长度为540至560个氨基酸残基、长度为560至580个氨基酸残基、长度为580至600个氨基酸残基、长度为600至620个氨基酸残基、长度为620至640个氨基

酸残基、长度为640至660个氨基酸残基、长度为660至680个氨基酸残基、长度为680至700个氨基酸残基、长度为700至720个氨基酸残基、长度为720至740个氨基酸残基、长度为740至760个氨基酸残基、长度为760至780个氨基酸残基、长度为780至800个氨基酸残基、长度为400至500个氨基酸残基、长度为500至600个氨基酸残基、长度为600至700个氨基酸残基、长度为700至800个氨基酸残基、长度为450至550个氨基酸残基、长度为550至650个氨基酸残基、长度为650至750个氨基酸残基、或长度为750至800个氨基酸残基。在一些实施方案中,本公开内容的可编程切口酶的长度为350至900个氨基酸。在一些实施方案中,本公开内容的可编程切口酶的长度为500至550个氨基酸。在优选的实施方案中,本公开内容的可编程切口酶的长度为480至550个氨基酸残基。

[0038] 在一些实施方案中,V型CRISPR/Cas酶是可编程Cas14核酸酶。Cas14可称为CasZ。本公开内容的Cas14蛋白包括3个部分RuvC域(RuvC-I、RuvC-II和RuvC-III,本文中也称为子域),它们相对于Cas14蛋白的一级氨基酸序列是非连续的,但一旦产生蛋白质并折叠,便形成RuvC域。天然存在的Cas14蛋白用作核酸内切酶,催化靶核酸中的特定序列处的切割。可编程的Cas14核酸酶可以是Cas14a蛋白、Cas14b蛋白、Cas14c蛋白、Cas14d蛋白、Cas14e蛋白、Cas14f蛋白、Cas14g蛋白、Cas14h蛋白、Cas14j、Cas14k、Cas14l,或Cas14u蛋白。

[0039] 重要的是需注意Cas14与先前鉴定的CRISPR-Cas核酸内切酶相比短,因此使用该蛋白作为可选方案提供了编码该蛋白的核苷酸序列相对短的优势。这在例如期望编码Cas14蛋白的核酸的情况下,例如在采用病毒载体(例如,AAV载体)的情况下,用于递送至用于研究和/或临床应用的细胞例如真核细胞(例如,哺乳动物细胞、人细胞、小鼠细胞,体外、离体、体内)是有用的。此外,在它们的自然环境中,编码Cas14的DNA序列存在于也具有Cas1蛋白的基因座中。

[0040] 表1中所示的Cas14或其变体与SEQ ID NO:1-SEQ ID NO:91或SEQ ID NO:170中的任一项具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性,可以表现出产生切口活性。“同一性百分比”和“同一性%”是指在比对中两个序列(核苷酸或氨基酸)在相同位置具有相同残基的程度。例如,“氨基酸序列与SEQ ID NO:Y具有X%同一性”是指氨基酸序列与SEQ ID NO:Y的同一性%,并详细说明为氨基酸序列中X%的残基与SEQ ID NO:Y中公开的序列残基相同。通常,计算机程序用于此类计算。比较和比对序列对的示例性程序包括ALIGN(Myers和Miller,Comput Appl Biosci.1988Mar;4(1):11-7)、FASTA(Pearson和Lipman,Proc Natl Acad Sci U S A.1988Apr;85(8):2444-8;Pearson,Methods Enzymol.1990;183:63-98)和缺口BLAST(Altschul等人,Nucleic Acids Res.1997Sep 1;25(17):3389-40)、BLASTP、BLASTN或GCG(Devereux等人,Nucleic Acids Res.1984Jan 11;12(1Pt 1):387-95)。

[0041] 本公开内容的示例性可编程切口酶(例如,本文公开的Cas14a、Cas14b或Cas14e)与SEQ ID NO:1-SEQ ID NO:91或SEQ ID NO:170中的任一项具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性。与本文公开的组合物和方法一致的示例性可编程切口酶与SEQ ID NO:1具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性。与本文公开的组合物和方法一致的示例性可编程切口酶与SEQ ID NO:10具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少

99%或100%序列同一性。与本文公开的组合物和方法一致的示例性可编程切口酶与SEQ ID NO:11具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性。与本文公开的组合物和方法一致的示例性可编程切口酶与SEQ ID NO:17具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性。与本文公开的组合物和方法一致的示例性可编程切口酶与SEQ ID NO:33具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性。

表1-Cas14序列

SEQ ID NO	序列
SEQ ID NO: 1	MEVQKTVMKTLRLRPLYSQIEIEKEIEEKERRKQAGGTGELDGGFYKKLEKKHS EMFSFDRLNLLLNLQREIAKVYNHASELYIATIAQGNKSNKHYSIVYNRAYGYF YNAYIALGICSKVEANFRSNELLTQQSALPTAKSDNFPVLHKQGAEGEDGGFRIST EGSDLIFEIPIPFYEYNGENRKEPYKWVKKGGQKPVLLKLSTFRRQRNKGWAKDEG TDAEIRKVTEGKYQVSQIEINRGKKLGEHQKWFANFSIEQPIYERKPNRSIVGGDLVG IRSPLVCAINNSFSRYSVDSNDVFKFSKQVFAFRRLLSKNSLKRKGHGAHAKLEPIT EMTEKNDKFRKKIIERWAKEVTNFFVKNQVGVQIEDLSTMKDREDHFFNQYLGRF WPYYQMOTLIENKLKEYGIEVKRVQAKYTSQLCSNPNCRYWNNYFNFEYRKVNKF PKFKCEKCNLEISADYNAARNLSTPDIEKFVAKATKGINLPEK
SEQ ID NO: 2	MEEAKTVSKTLRLRPLYSAEIEKEIEEKERRKQGGKSGELDSGFYKKLEKKHT QMFGWDKLNLMLSQLQRQIARVFNQISIELYIETVIQGGKSNKHYSIVYNRAYSV FYNAYLALGITSKVEANFRSTELLMQKSSLPTAKSDNFPILLHKQKGVGEGEGGFKIS ADGNDLIFEIPIPFYEYDSANKKEPFKWKKGQKPTIKLILSTFRRQRNKGWAKDEG TDAEIRKVIEGKYQVSHIEINRGKKLGDHQQKWFVNFTIEQPIYERKLDKNIIGGIDVG KSPLVCAVNNFSARYSVDSNDVLFKFSKQAFRRRLLSKNSLKRSGHGSKNKLDPI RMTEKNDKFRKKIIERWAKEVTNFFIKNQVGTQVQIEDLSTMKDRQDNFFNQYLGRF WPYYQMQLNIENKLKEYGIETKRIKARYTSQLCSNPSCRHWNSYFSFDRHKTNNFP KFKCEKCALEISADYNAARNISTPDIEKFVAKATKGINLPDKNENVILE
SEQ ID NO: 3	MAKNITITKTLRLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKV AAYCTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEI YNQSLIELYYEIKGKGIANASSVEHYLSVCYTRAELFKNAIASGLRSKIKSNFR LKELKNMKSGLPTTKSDNFPILVKKQGGQYTGFEISNHNDSDFIIPFGRWQVKKEI DKYRPWEKFDQVQKSPKISLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQT SYIEVKRGSKIGESAWMLNLSIDVPKIDKGVDPISIGGIDVGKSPLVCAINNAFSRY SISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIER WACEIADFFIKNKVGTQVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQ YGIEIRKVAPNNTSKTCSKCGHLNNYFNFEYRKKNKFPFKCEKCNFKENADYNAA LNISNPKLKSTKEEP
SEQ ID NO: 4	MERQKVPQIRKIVRVVPLRLRPLKPYSDVIENALKKFKEKGDDTNTNDFWRAIRDRT EFFRKELNFSDEINQLERDTLFRVGLDNRVLFYSYDFDLQEKLMKDYNKIISKLFINR QSKSSFENDLTDEEVEELIEKDVTPTYGAYIGKGIKSVIKSNLGGKFIKSVKIDRETKK VTKLTAINIGLMGLPVAKSDTFPIKIKTNPDIYTFQKSTKENLQKIEDYETGIEYGDLL VQITIPWFKNENKDFSLIKTKEAIEYYKLVNGVGGKDLLNINLVLTYYHIRKKKSWQID GSSQSLVREMANGELEEKWKSFFDTFIKKYGDEGKSALVKRRVNKKSRAGKEGR ELNLDERIKRLYDSIAKSPSEINLIPENYKWKLFHSIEIPPMVNDIDSNLYGGIDFGE QNIATLCVKNIEKDDYDFLTIIYGNLKLKHAQASYARRRIMRVQDEYKARGHGKSRK TKAQEDYSERMQLRQKITERLVKQISDFFLWRNKFHMAVCSLRYEDLNTLYKGES VKAARMRQFINKQQLFNGIERKLKDYNSEIYVNSRYPHYTSRLCSKCGKLNLYDFDL KFRTKNIIIRKNPDGSEIKYMPFFICEFCGWKQAGDKNASANIADKDYQDKLNKEKE FCNIRKPKSKKEDIGEENEEERDYSRRFNRSFIYNSLKKDNKLNQEKLFDEWKNQL KRKIDGRNKFEPKEYKDRFSYLFAYYQEIKNES
SEQ ID NO: 5	MVPTELITKTLQLRVIRPLYFEEIEKELAELEKEKEFEETNSLLLESKKIDAKSLKK LKRKARSSAAVEFWKIAKEKYPDILTKPEMEFIFSEMOMMARFYNKSMTNIFIEMN NDEKVNPLSLISKASTEANQVIKCSSISSGLNRKIAGSINKTKFKQVRDGLISLPTART ETFPISFYKSTANKDEIPISKINLPSEEEADLTITLPFPFIEIKKEKKGQKAYSFYNIEKS GRSNKIDLLSTHRRQRKRGWKEEGGTSAEIRRLMEGEFDKEWEIYLGEAEKSEK AKNDLIKNMTRGKLSKDIKEQLEDIQVKYFSDNNVESWNDLSKEQKQELSKLRKKK VEELKDWKHVKEILKTRAKIGWVELKRGKRQRDRNKWVNITITRPPFINKELDDT KFGGIDLVKVPFVCAVHGSPARLIKENILQFNKMVSARNRQITKDSEQRKGRGK KNKFIKKEIFNERNELFRKKIIERWANQIVKFFEDQKCATVQIENLESFDRTSYK
SEQ	MKSDTKDKKIIHQTKTLRLRIVKPSIPMEEFTDLVRYHQMIIFPVYNNGAIDLKYL

SEQ ID NO	序列
ID NO: 6	FKAKIQKGNEARAIKYFMNKIVYAPIANTVKNSYIALGYSTKMQSSFSGKRLWDLRF GEATPPTIKADFPFYNQSGFKVSSENGEFIIGIPFGQYTKKTVSDIEKKTSAWDFK TLEDTTKKTLEILLSTKTRKMNEGWNNEGTEAEIKRVMDGTYQVTSLEILQRDDS WVFVFNIAYDSLKKQPDRLKIAGIHMGITRPLTAVIYNNKYRALSIPNTVMHLTQK QLARIKEQRTNSKYATGGHGRNAKVTGDTLSEAYRQRRKKIIEDWIASIVKFAINN EIGTIYLEDISNTNSFFAAREQKLIYLEDISNTNSFLSTYKYPISASIDTLQHKLEEKAIQ VIRKKAYYVNQICSLCGHYNKGFTYQFRRKNKFKPMKCQGCLEATSTEFNAAANV ANPDYEKLLIKHGLLQLKK
SEQ ID NO: 7	MSTITRQVRLSPTPEQSRLMAHCQQYISTVNVLVAAFDSEVLTGKVSTKDFRAALP SAVKNQALRDAQSVFKRSVELGCLPVLKKPHCQWNNQNWVEGDQLILPICKDGK TQQERFRCAVALEGKAGILRIKKKRGKWIADLTVTQEDAPESSGSAIMGVDLGIKV PAVAHIGGKGTRFFGNRSQRSRMRFFYARRKTLQKAKKLRAVRKSKGKEARWM KTINHQLSRQIVNHAHALGVGTIKIEALQGIRKGTTRKSRGAAARKNNRMTNTWSFS QLTLFITYKAQRQGITVEQVDPAYTSQDCPACRARNGAQDRTYVCSECGWRGHRD TVGAINISRRAGLSGHRRGATGA
SEQ ID NO: 8	MIAQKTIKIKLNPTKEQIIKLSIIEEYIKVSNFTAKKIAEIQESFTDSGLTQGTCECGK EKTYRKYHLLKKDNKLCITCYKRKYSQFTLQKVEFQNKTLGRNVAKLPKTYTNTN AIRFASDTFSGFDEIHKKKQNRLNSIQNRLNFWKELLYNPSNRNEIKIKVVKYAPKTD TREHPHYYSAEIKGRIKRLEKQLKKFKMPKYPEFTSETISLQRELYSWKNPDELKIS SITDKNESMNYYGKEYLKRYIDLINSQTPQILLEKENNSFYLCFPITKNIEMPKIDDTF EPVGIDWGITRNIADVVSILDSKTKKPKFVKFYSAGYILGKRKHYSRKRHFQKKRQ DKINKLGTKEDRFIDSNIHKLAFILVKEIRNHSNKPILMENITDNREEAEKSMRQNIL LHSVKSRLQNYIAYKALWNNIPTNLVKPEHTSQICNRCGHQDRENRPKGSKLFCV KCNMYMSNADFNASINIARKFYIGEYEPFYKDNEKMKSGVNSISM
SEQ ID NO: 9	LKLSEQENITTVGVKFLKLDKETSEGLNDYFDEYGKAINFAIKVIQKELAEFRFAGK VRLDENKKPLLNEGDGKKIWDFFPNEFCSCGKQVNRYVNGKSLCQECYKNKFTEYGIR KRMYSAGRKAEQDINIKNSTNKISKTHFNIAIREAFILDKSIKKQRKERFRRLREM KKKLQEFIEIRDGNKILCPKIEQQRVERYIHPSWINKEKKLEDFRGYSMSNVLGKIKIL DRNIKREEKSLKEKGQINFKARRMLDKSVKFLNDNKISFTISKNLKPEYELDLPEKE KRLNWLKEKIKIKNQKPKYAYLLRKDDNFYLQYTLETFNLKEDYSGIVGIDRGVS HIAVYTFVHNNGKNERPLFLNSSEILRLKNLQKERDRFLRRKHNNKRRKSNMRNIEK KIQLILHNSKQIVDFAKNKNAFIVFEKLEKPKKNRSKMSKKSQYKLSQFTFKKLSLSD LV DYKAKREGIKVLYISPEYTSKECSHCGEKVNTQRPFGNSSLFKCNKCGVELNAD YNASINIAKKGLNILNSTN
SEQ ID NO: 10	MEESITGVKFLRIDKETTKKLNEYFDEYGKAINFAVKIIQKELADDRFAGKAKLDQ NKNPILDENGKKIYEPDEFCSGKQVNKYVNNKPFCQECYKIRFTENGIRKRMYS KGRKAEHKINILNSTNKISKTHFNIAIREAFILDKSIKKQRKRNERLRESKKRLQQFI DMRDGKREICPTIKGQKVDRFIHPSWITDKKLEDFRGYTSLIINSKIKILDRNIKREE KSLKEKGQIIFKAKRLMLDKSIRFVGDRKVLFTISKTLKPEYELDLPSKEKRLNWLKE KIEIKNQKPKYAYLLRKNIIESEKKPNYEYYLQYTLEIKPELKDFYDGAIGIDRGINHI AVCTFISNDGKVTPPKFFSSGEILRLKNLQKERDRFLLRKHNNKRRKGNMRVIENKI NLILHRYSKQIVDMAKKNASIVFEELGRIGKSRTKMKKSQRYKLSLFIKKSDDLVD YKSRREGIRVTYVPPEYTSKECSHCGEKVNTQRPFGNSSLFKCNKCGIQLNSDYNA SINIAKKGLKIPNST
SEQ ID NO: 11	MPKQDLVTTGKFKLDVDKETRKKLDDYFDEYGKAINFAVKIIQKELADDRFAGKIA LGEDKKPLLDKDGKKIYNYPNESCSCGNQVRRYVNAKPFVDCYKLLFTENGIRKR MYSARGRKADSDINIKNSTNKISKTHFNIAIREGFILDKSLKKQRSKRIKKLLELKRK LQEFIDIRQGMVLCPIKNQVRVDKFIHPSWLKRDKKLEEFRGYSLSVVEGKIKIFNR NILREEDSLRQRGHVNFKANRIMLDKSVRFLDGGKVNFNLNKGLPKYLLDLPKKE NKLSWLNEKISLIKQKPKYAYLLRREGSFFIYTIENVPKTFSDYLGAIGIDRGISHIA VCTFVSKNGVNKAPVFFSSGEILKLKSLQKQRDLFRLRGKHNNKIRKKSNNMRNIDNKIN LILHKYSRNIVNLAKSEKAFIVFEKLEKIKKSFRKMSKSLQYKLSQFTFKKLSLSDVEY

SEQ ID NO	序列
	KAKIEGIKVDYVPPEYTSKECSHCGEKVDTPQPFNGNSSLFKCNKCRVQLNADYNASINIAKKSLNISN
SEQ ID NO: 12	MSKTTISVKLKIIDLSEKKEFLDNFYFNEYAKATTFCQLRIRRLLRNTHWLGKKEKSSKKWIFESGICDLGKGENKELVNEDRNSGEPKICKRCYNGRYGNQMIRKLFVSTKKREVENQENMDIRRVAKLNNTHYHRIPEEAFDMIKAADTAEKRRKKNVEYDKKRQMEFIE MFNDEKKRAARPKPNERETRYVHISKLESPPSKGYTLNGIKRKIDGMGKKIERAEKG LSRKKIFGYQGNRIKLDNSWVRFDLAESEITIPSLFKEMKLRTGPTNVHSGQIYFA EWFERINKQPNNYCYLIRKTSSNGKYEYYLQYTYEAEVEANKEYAGCLGVDIGCSK LAAAVYYDSKNKKAQKPIEFTNPIKKIKMRREKLKLLSRVKVRHRRRRLMQLSKT EPHIDYTCHKTARKIVEMANTAKAFISMENLETGIKQKQQAETKKQKFYRNMFLFR KLSKLEIKYKALLKGIKIVYVKPDYTSQTCSSCGADKEKTERPSQAIFRCLNPTCRYQQ RDINADFNAAVNIAKKALNNTTEVTTLL
SEQ ID NO: 13	MARAKNQPYQKLTTTTGKFKLDLSEEGKRFDEYFSEYAKAVNFCVAKVIYQLRKN LKFAGKKELAAKEWKFEISNCDFCNKQKEIYKNIANGQKVCKGCHRTNFSDNAIR KKMIPVKGRKVESKFNIHNTTKKISGTHRHWAFFEDAADIESMDKQRKEKQKRLRR EKRLKSYFFELFGDPAKRYELPKVGKQVRPYLHKIIDKDSLTKKRGYSLSYIKNKIK ISERNIERDEKSLRKASPIAFGARKIKMSKLDPKRAFDLENNVFKIPGKVIKGGYKFF GTNVANEHGGKFKYKDRISKILAGPKYFYLLRKKVAESDGNPIFEYVQWSIDTETP AITSYDNILGIDAGITNLATTVLIPKNLSAEHCSHCNNHVKPIFTKFFSGKELKAIKIK SRKQKYFLRGKHNLVKIKRIRPIEQKVDGYCHVVSQIVEMAKERNSCIALEKLEK PKSKFRQRRREKYAVSMFVFKKLATFIKYKAAREGIEIIPVEPEGTSYTCSHCKNAQ NNQRPYFKPNSKKSWSMFKCGKCGIELNSDYNAAFNIAQKALNMTSA
SEQ ID NO: 14	MDEKHFFCSYCNKELKISKNLINKISKSGSIREDEAVSKAISHNKKHSLILGIKFKLFI ENKLDKKKLNEYFDNYSKAVTFAARIFDKIRSPYKFIGLKDKNTKKWTFPAKACVF CLEEKEVAYANEKDNSKICTECYLKEFGENGIRKKIYSTRGRKVEPKYNIFNSTKELS STHYNYAIRDAFQLLDALKKQRQKKLSIFNQKLRLKEFEDIFSDPQKRIELSLKPHQ REKRYIHLKSGQESINRGYTLRFVRGKIKSLTRNIREEKSLRKKTPHFKGNRLMIF PAGIKFDFASNKVKISISKNLNPNFNFSGTNVKNHGGKSFSSRIELIKTQKPKYAYVL RKIKREYSKLRNIEIEKIRLENPNADLCDFYLQYTIETESRNNEEINGIIGIDRGITNLA CLVLLKKGDKKPSGVKFKYKGNKILGMKIAIRKHLVLLKGKRNKLRKQQRIRAIPEK INLILHQISKDIVKIAKEKNFAIALEQLEKPKKARFAQRKKEKYKLALFTFKNLSTLIE YKSKREGIPVIYVPPEKTSQMCSHCAINGDEHVDTPRPYKPKPNAQKPSYSLFKCNKC GIELNADYNAAFNIAQKGLKTLMLNHS
SEQ ID NO: 15	MLQTLVLKLDPSKEQYKMLYETMERFNEACNQIAETVFAIHSANKIEVQKTVYYPIR EKFGLSAQLTILAIRKVCEAYKRDKSIKPEFRLDGLVYDQRVLSWGLDKVSLVTL QGRQIPIKFGDYQKARMDRIRGQADLILVKGVFYLCVVVEVSEESPYDPKGVLGVD LGIKNLAVDSGGEVHSGEQTNTRERLDSLKARLQSKGTSKARHLKKLSGRMAKF SKDVNHCISKKLVAKAGTLMMSIALEDLQGIRDRTVRKAQRRLNHTWNFGLLRM FVDYKAKIAGVPLVFVDPRNTSRTCPCSGHVAKANRPTREDFRCVSCGFAGAADHI AAMNIAFRAEVSQPIVTRFFVQSQAQPSFRVG
SEQ ID NO: 16	MDEEPDSAEPNLAPISVKLKLKLDGEKLAALNDYFNEYAKAVNFCVAKVIYQLRKN LVNIRGTYLKEKKAWINQTEGCCICKIDELRCEDKNPDINGKICKKCYNGRYGNQ MIRKLFVSTNKRVPKSLDIRKVARLHNTYHRIPEEADIIKAIETAERKRRNRILFD ERRYNELKDALENEEKRVARPKPKEREVRYVPISKKDTSPSKGYTMNALVRKVSG MAKKIERAKRNLNKRKKIEYLGRRILLDKNWVRFDKSEISIPTMKEFFGEMRFEIT GPSNVMSPNGREYFTKWFDRIKAQPDNYCYLLRKESEDETDLYLQYTWRPDAHPK KDYTGCLGIDIGGSKLASAVYFDADKNRAKQPIQIFSNPIGKWKTKRQKVIKVLKSA AVRHKTKKLESRLNIEPRIDVHCHRIARKIVGMALAANAFISMENLEGGIREKQKAK ETKKQKFSRNMVFVRKLSKLEIKYKALMEGVKVVYIVPDYTSQLCSCGTNNTKRPK QAIFMCQNTTECRYFGKNINADFNAAINIAKKALNRKDIVRELS
SEQ ID	MEKNNSEQTSITGKFKLKLKDKETKEKLNNYFDEYGKAINFAVRHIIQMQLNDDRLA GKVKRDEKGPILGEDGKKILEIPNDFCSCGNQVNHVYVNGVSFCQECYKRRFSENGI

SEQ ID NO	序列
NO: 17	RKRMYSAKGRKAEQDINIKNSTNKISKTHFNIAIREAFNLDKSIKKQREKRFKKLKD MKRKLQEFLEIRDGKRVICPKIEKQKVERYIHPSWINKEKKLEEFGRGYSLSIVNSKIS FDRNIQREEKSLKEKGQINFKAQRLMLDKSVKFLKDNKVSFTISKELPKTFELDLPPK EKKLNLWNEKLEIKNQKPKYAYLLRKENNIFLQYTLDSIPEIHSEYSGAVGIDRGVS HIAVYTFLDKDGKNERPFLSSSGILRLKNLQKERDKFLRKKHKNIRKKGNMRNIEQ KINLILHEYSKQIVNFAKDKNFIVFELLEPKPKSRERMSKKIQYKLSQFTFKKLSDL VDYKAKREGIKVIYVEPAYTSKDCSHCGERVNTQRPFNNGNFSLFKCNKCGIVLNSDY NASLNIARKGLNISAN
SEQ ID NO: 18	MAEEKFFFCEKCNKDIKIPKNYINKQGAEEKARAKHEHRVHALILGIKFKIYPPKKEDI SKLNDYFDEYAKAVTFTAKIVDKLKAPFLFAGKRDKDTSKKKWVFPVVKCSFCKE KTEINYRTKQGKNICNSCYLTFEGEQGLLEKIYATKGRKVSSSNLNFNSTKKLTGTH NNYVVKESLQLLDALKKQRSKRLKLSNTRRKLKQFEEMFEKEDKRFQLPLKEKQR ELRFIHSVSKDRATEFKGYTMNKIKSKIKVLRRNIEREQRSLNRKSPVFFRGTRIRLSP SVQFDDKDNKIKLTLSELKPKEYSFGSLNVANEHGRKFFAEKLLIKENSKYAYLL RRQVNNKNNKKPIYDYLYQYTFEFLPNIITNYNGILGIDRGINTLACIVLLENKKEKPSF VKFFSGKGILNLKNKRRKQLYFLKGVHNKYRKQQKIRPIEPRIDQILHDISKQIIDLAK EKRV AISLEQLEKPKPKFRQSRKAKYKLSQFNFKTLSNYIDYKAKKEGIRVIYIAPE MTSQNCSRCAMKNDLHVNTQRPYKNTSSLFKCNKCGVELNADYNAAFNIAQKGLK ILNS
SEQ ID NO: 19	MISLKLKLLPDEEQKLLDEMFWKWASICTRVGFGRADKEDLKPPKDAEGVWFSLT QLNQANTDINDLREAMKHQKHRLEYEKNRLEAQRDDTQDALKNPDRREISTKRKD LFRPKASVEKGFLKLYHQERYWVRRLEINKLIERKTKTLIKIEKGRIKFKATRITL HQGSFKIRFGDKPAFLIKALSGKNQIDAPFVVVPEQPICGSVNSKKYLDEITTNFLA YSVNAMLFGLSRSEEMLLKAKRPEKIKKKKEEKLAKKQSAFENKKELQKLLGRELT QQEEAIIETRNQFFQDFEVKITKQYSELLSKIANELKQKNDFLKVNKYPIILLRKLK KAKSKKINNLSPEWKYYLQFGVKPLLKQKSRRKSRNVLGIDRGLKHLAVTVLEP DKKTFFVWNKLYPNPITGWKWRRRKLRLSLKRLKRRISQKHETIHENQTRKKLKS QGRIDDLHNISRKIVETAKEYDAVIVVEDLQSMRQHGRSKGNRLKTLNYALSFLDY ANVMQLIKYKAGIEGIIYDVKPAGTSQNCAYCLLAQRDSHEYKRSQENSKIGVCL NPNCQNHKKQIDADLNAARVIASCYALKINDSQPFGTRKRFRKRTTN
SEQ ID NO: 20	METLSLKLKLNPSKEQLLVLDKMFWKWASICTRLGLKKAEMSDLEPPKDAEGVWF SKTQLNQANTDVNDLRKAMQHOGKRIEYELDKVENRRNEIQEMLEKPDRRDISPNR KDLFRPKAAVEKGYLKLKYHKLGYWSKELKTANKLIERKRKTLAKIDAGKMKFKP TRISLHTNSFRIKFGEEKIALSTTSKHEKIELPLITSLQRPLKTSACKSKTYLDAAIL NFLAYSTNAALFGLSRSEEMLLKAKKPEKIEKRDRKLATKRESFDKLLKTLEKLLER KLSEKEKSVFKRKQTEFFDKFCITLDETYVEALHRIAEELVSKNKYLEIKKYPVLLRK PESRLRSKLLKNLKPEDWYTYIQFGFQPLDTPKPIKTKTVLGIDRGVRHLLAVSIFD PRTKTFTFNRLYSNPVDWKWRRRKLRLSIKRLKRRLKSEKHVHLHENQFKAKLRS EGRIEDHFHNLKSKIVDLAKENNSVIVVENLGGMRQHGRGRGKWLKALNYALSHF DYAKVMQLIKYKAELAGVFVYDVAPAGTSINCAYCLLNDKASNYTRGKVIKNGK NTKIGECKTCKKEFDADLNAARVIALCYEKRLNDPQPFGRKQFKPKKP
SEQ ID NO: 21	MKALKLQIPTRKQYKILDEMFWKWASLANRVSQKGESKETLAPKKDIQKIQFNAT QLNQIEKDIKDLRGAMKEQQKQKERLLLQIQERRSTISEMLNDDNNKERDPHRPLNF RPKGWRKFHTSKHWVGELSKILQREDRVKKTIERIVAGKISFKPKRIGIWSSNYKINF FKRKISINPLNSKGFELTLMTEPTQDLIGKNGGKSVLNNKRYLDDSIKSLLMFALHSR FFGLNNTDTYLLGGKINPSLVKYYKKNQDMGEFGREIVEKFERKLKQEINEQQKKII MSQIKEQYSNRDSAFNKDYLGLINEFSEVFNRKSERAEYLLDSFEDKIKQIKQEIGE SLNISDWDFLIDEAKKAYGYEEGFTEYVYSKRYLEILNKIVKAVLITDIYFDLRKYPIL LRKPLDKIKKISNLKPDEWSYYIQFGYDSINPVQLMSTDKFLGIDRGLTHLLAYSVD KEKKEFIINQLEPNPIMGWKWKLKVKRSLQHLERRIRAQKMVKLPENQMKKKLKS IEPKIEVHYHNISRKIVNLAKDYNASIVVESLEGGGLKQHGRKKNARNRSLNYALSL FDYGKIASLIKYKADLEGVPMYEVLPAYTSQQCAKCVLEKGSFVDPEIIGYVEDIGIK

SEQ ID NO	序列
	GSLLDSLFEGTELSSIQVLKKIKNKIELSARDNHNKEINLILKYNFKGLVIVRGQDKEE IAEHPIKEINGKFAILDFVYKRGKEKVGKKGNQKVRVTGNKKVGYSKHHGQVDAD LNASRVIALCKYLDINDPILFGEQRKSFK
SEQ ID NO: 22	MVTRAIKLLDPTKNQYKLLNEMFWKWASLANRFSQKGASKETLAPKDGTOQKIQF NATQLNQIKKDVDDLRLGAMEKQKGQKERLLIQIERLLTISEILRDDSKKEKDPHRP QNFRPFGWRRFHTSAYWSSEASKLTRQVDRVRRTIERIKAGKINFKPKRIGLWSSTY KINFLKKKINISPLKSKSFELDLITEPQQKIIGKEGGKSVANSKKYLDDSIKSLIFAIS RLFGLNNKDKPLFENIITPNLVRYHKKGQEQENFKKEVIKKFENKLLKEISQKQKEII FSQIERQYENRDATFSEDYLRASEFSEIFNQRKKERAKELLNSFNEKIRQLKKEVNG NISEEDLKILEVEAEKAYNYENGFIWEYSEQFLGVLEKIARAVLISDNYFDLKKYPI LIRKPTNKSCKITNLKPEEWDYIYQFGYGLINSPMKIETKNFMGIDRGLTHLLAYSIF DRDSEKFTINQLELNPIKGWKWKLKVKRSLQHLERRMRAQKGVKLPENQMKKRL KSIEPKIESYYHNLNRKIVNLAANNASIVVESLEGGLKQHGRKKNSRHRALNYAL SLFDYGKIASLIKYSKSDLEGVPMYEVLPAYTSQQCAKCVLKKGSFVEPEIIGYIEEIGF KENLLTLLFEDTGLSSVQVLKKSCKNMTLSARDKEGKMVDLVLYNFKGLVISQEK KKEEIVEFPIKEIDGKFAVLDSAYKRGKERISKKGQKL VYTGNKKVGYSVHGQV DADLNASRVIALCKYLGINEPIVFGEQRKSFK
SEQ ID NO: 23	LDLITEPIQPHKSSSLRSKEFLEYQISDFLNFSLHSLFFGLASNEGPLVDFKIYDKIVIPK PEERFPKKESEEGKKLDSFDRVEEYYSKLEKKIERKLNTEENVIDREKTRIWGE VNKLEEIRSIIDEINEIKKQKHISEKSKLLGEKWKKVNNIQETLLSQEYVSLISNLSDEL TNKKKELLAKKYSKFDDKIKKIKEDYGLEFDENTIKKEGEKAFLNPKDFSKYQFSS YLKLIGEIARSLITYKGFLDLNKYPIHFRKPINKVKKIHNLEPDEWKYIYQFGYEQINN PKLETENILGIDRGLTHLLAYSVFEPSSKFIKLEPNPIEGWKWKLRLKLRRIQNLE RRWRAQDNVCLPENQMKKNLRSIEDKVENLYHNLSRKIVDLAKEKNACIVFEKLEG QGMKQHGRKKSDRLRGLNYKLSLFDYGKIAKLIKAEIEGPIYRIDSAYTSQNC KCVLESRRFAQPEEISCLDDFKEGDNLDRILEGTGLVEAKIYKLLKEKKEDFEIEE DIAMFDTKKVIKENKEKTVILDYVYTRRKEIIGTNHKKNIKGIKTYGNTKIGYCMK HGQVDADLNASRTIALCKNFDINNPEIWK
SEQ ID NO: 24	MSDESLVSSDKLAIKIKIVPNAEQAKMLDEMFKKWSSICNRISRGKEDIETLRPDEG KELQFNSTQLNSATMDVSDLKKAMARQGERLEAEVSKLRGRYETIDASLRDPSRRH TNPQKPSSFYPSDWDISGRLTPRFHTARHYSTELRKLKAKEDKMLKTINKIKNGKIVF KPKRITLWPSSVNMAFKGSRLLLKPFANGFEMELPIVISPQKTADGKSQKASAEYMR NALLGLAGYSINQLLFGMNRSQKMLANAKKPEKVEKFLEQMKNKDANFDKKIKAL EGKWLLDRKLKESEKSSIAVVRTKFFKSGKVELNEDYLKLLKHMANEILERDGFVN LNKYPILSRKPMKRYKQKNIDNLKPNMWKYIYQFGYEPFERKASGKPKNIMGIDRG LTHLLAVAVFSPDQKFLFNHLESNPIMHWKWKLRKIRRSIQHMERRIRAEKNKHHI EAQLKKRLGSIEEKTEQHYHIVSSKIINWAEYEAIVLESLSHMKQRGGKKSVRTRA LNYALSLFDYKVARLITYKARIRGIPVYDVLPGMTSKTCATCLLNGSQGAYVRGLE TTKAAGKATKRKNMKIGKCMVCNSSSENSMIDADLNAAARVIAICKYKNLNDPQPAG SRKVFKRF
SEQ ID NO: 25	MLALKLKIMPTEKQAEILDAMFWKWASICSRIAKMKKKVSVKENKKELSKKIPSNS DIWFSKTQLCQAEVDVGDHKKALKNFKEKRQESLLDELKYKVKAINVINDESKREI DPNNPSKFRIKDSTKKGNLNSPKFTLKKWQKILQENEKRIKKKESTIEKLKRGNIFF NPTKISLHEEYSINFGSSKLLNCFYKYNKKSGINSQLENKFNEFQNLNIICSPLQ PIRGSSKRSFEFIRNSIINFLMYSYAKLFGIPRSVKALMKSNKDENKLEEKLKKK KSSFNKTVEFEKMIKRLSDNESKILNDESKKFFEIISNNKYIPSEYKLLKDISEE IYNSNIDFKPYKYSILIRKPLSKFKSKKLYNLKPTDYKYLYQLSYEPFSKQLIATKIL GIDRGLKHLLAVSVFDPSONKFVYNKLIKNPVFKWKKRYHDLKRSIRNRERRIRALT GVHIHENQLIKLKSMMKNKINVLHNVSKNIVDLAKKYESTIVLERLENLKQHGRSK GKRYKKLNYVLSNFDYKKIESLISYKAKKEGVPVSNINPKYTSKTCACCLLEVNLQS ELKNEYNRDSKNSKIGICNIHQIDADLNAAARVIALCYSKNLNPHFK
SEQ	VINLFGYKFALYPNKTQEELLNKHLGECGWLYNKAIEQNEYKADSNIEEAQKKFE

SEQ ID NO	序列
ID NO: 26	LLPDKNSDEAKVLRGNISKDNYVYRTLVLKKKKSEINVQIRKAVVLRPAETIRNLAKV KKKGLSVGRLKFIPIREWDLVLPFKQSDQIRLEENYLILEPYGRLKFKMHRPLLGKPKT FCIKRTATDRWTISFSTEYDDSNMRKNDGGQVGDVGLKTHLRLSNENPDDEPRYPN PKIWKRYDRRLTILQRRISKS KKLGNRTRLRLRLSRLWEKIRNSRADLIQNETYEIL SENKLIAIEDLNVKGMQEKDKKGRKGRTRAQEKGLHRSISDAAFSEFRRVLEYKA KRFGESEVKPVSAIDSSKECHNCGNKKGMPLESRIYECPCGLKIDRLNSAKVILAR ATGVRPGSNARADTKISATAGASVQTEGTVSEDFRQQMETSDQKPMQGEKSKEPPM NPEHKSSGRGSKHVNIGCKNKVGLYNEDENSRSTEKQIMDENRSTTEDMVEIGALH SPVLTT
SEQ ID NO: 27	MIASIDYEAVSQALIVFEFKAKGKDSQYQAIDEAIRSYRFIRNSCLRYWMDNKKVGK YDLNKYCKVLAKQYPFANKLNSQARQSAAECWSAISRFYDNCKRKVSGKKGFPGK FKKHARSVEYKTSGWKLSNRKAITFTDKNIGIGLKLKGTGYDLHFSQLEDMKRVRL VRRADGYYVQFCISVDVKVETEPTGAIGLDVGIKYFLADSSGNTIENPQFYRKA EK KLNRRNRKSKKYIRGVKPKSKNYHKARCRYARKHLRVSRQRKEYCKRVAYCVIHI SNDVVAYEDLNVKGMVKNRHLAKSISDVAWSTFRHWLEYFAIKYGKLTIPVAPHN TSQNCSCNDKKVPKSLSTRTHICHHCYSED RDVNAAKNILKKALSTVGGTGSGLK GEIEPLLVEQSCTRKFDL
SEQ ID NO: 28	LAEENTLHLTLAMSLPLNDLPENRTRSELWRRQWLPQKKLSLLLGVNQSVRKA AA DCLRWFEFYQELLWWEPTDPDGKLLDKEGRPIKRTAGHMRVLRKLEEIAPFRGYQ LGS AVKNGLRHKVADLLLSYAKRKLD PQFTDKTSYPSIGDQFPIVWTGAFVCYEQSI TGQLYLYLPLFPRGSHQEDITNNYDPDRGPALQVFGEKEIARLSRSTSGLLPLQFDK WGEATFIRGENNPPTWKATHRRSDKKWLSEVLLREKDFQPKRVELLVRNGRIFVNV ACEIPTKPLLEVENFMGVSFGLLEHLVTVVINRDGNVHVHQRQEPARRYEKTYFARL ERLRRRGPFPSQELET FHYRQVAQIVEEALRFKSVPAVEQVGNIPKGRYNPRLNLRL SYWPFGLADLTSYKAVKEGLPKPYSVYSATAKMLCSTCGAANKEGDQPISLKGPT VYCGNCGTRHNTGFNTALNLARRAQELFVKGVVAR
SEQ ID NO: 29	MSQSLLKWHDMAGRDKDASRLQKSAVEGVLLHLTASHRVALEMLEKSVSQTVA VTMEAAQQRLVIVLEDDPTKATSRKRVISADLQFTREEFGSLPNWAQKLASTCPEIA TKYADKHINSIRIAWGVAKESTNGDAVEQKLQWQIRLLDVTMFLQQLVLQLADKA LLEQIPSSIRGGIGQEVAAQVTSHIQLLDSGTVLKAELPTISDRNSELARKQWEDAIQT VCTYALPFSRERARILDPGKYAAEDPRGDRLINIDPMWARVLKGPTVKSPLLFVSG SSIRIVKLTLPKHAAGHKHTFTATYLVLPVSREWINSPLGTVQEKVQWWKKPDVL ATQELLVGKGALKKSANTLVIPI SAGKKRFFNHILPALQRGFPLQWQRI VGRSYRRP ATHRKWF AQLTIGYTNPSSLP EMALGIHFGMKDILWWALADKQGNILKDGSIPGNSI LDFSLQEKGKIERQQKAGKNVAGKKYKGSLLNATYRVVNGVLEFSKGISA EHASQP IGLGL ETIRFVDKASGSSPVNARHSNWN YGQLSGIFANKAGPAGFSVTEITLKK AQR DLSDAEQARVLAIEATKR FASRIKRLATKRKDDTLFV
SEQ ID NO: 30	VEPVEKERFYR TYTFRLDGQPRTQNLTTQSGWGLLTKAVLDNTKHYWEIVHHARI ANQPIVFENPVIDEQGNPKLNKLGPFRFWKRPISDIVNQLRALFENQNPYQLGSSLIQ GTYWDVAENLASWYALNKEYLAGTATWGEPSFPEPHPLTEINQWMPLTFSSGKV RLLKNASGRYFIGLPILGENNPCYMRMTIEKLIPCDGKGRTSGSLILFPLVGIYAQQH RRMTDICESIRTEKGKLAWAQVSIDYVREVDKRRRMRRTRKSQGWIQGPWQEVFIL RLVLAHKAPKLYKPRCFAGISLGPKTLASCVILDQDERVVEKQQWWSGSELLSLIHQG EERLRLSREQSKPTWNAAYRKQLKSLINTQVFTIVTFLRERGA AVRLES IARVRKSTP APPVNFLLSHWAYRQITERLKD LAIRNGMPLTHSNGSYGVRFTCSQCGATNQGIKDP TKYKVDIESETFLCSICSHREIAAVNTATNLAKQLLDE
SEQ ID NO: 31	MNDTETSETLTSHRTVCAHLHVGETGSLPRLVEAALAE LITLNGRATQALLSLAKN GLVLRDKEENLIAAELTLPCKNKYADVAAKAGEPILATRINNKGKLVTKK WYGE GNSYHIVRFTPETGMFTVRVFDRYAFDEELLHLHSEVVFGSDLPKGKIAKTDSPAN FLQAVFTSFLELPFQGFDPDIVVKPAMKQAAEQLLSYVQLEAGENQQA EYPDTNERDP ELRLVEWQKSLHEL SVRTEPF EFVRARDIDY AETDRRGNR FVNITPEWTKFAESPF ARRLPLKIPPEFCILLRRKTEGHAKIPNRIYLG LQIFDGVTPDSTLGVLATAEDGKLFW

SEQ ID NO	序列
	WHDHLDEFSNLEGKPEPKLKNKPQLLMVSLEYDREQRFEESVGGDRKICLVTLKET RNFRRGWNGRILGIHFQHNPVITWALMDHDAEVLEKGFIEGNAFLGKALDKQALNE YLQKGGKWVGDRSFGNKLKGITHTLASLIVRLAREKDAWIALEEISWVQKQSADSV ANHEIVEQPHHSLTR
SEQ ID NO: 32	MNDTETSETLTSHRTVCAHLHVVGGETGSLPRLVEAALAELITLNGRATQALLSLAKN GLVLRDKEENLIAAELTLPCRKNKYADVAAKAGEPILATRINNKGKLVTKKWYGE GNSYHIVRFTPETGMFTVRVFDRYAFDEELLHLHSEVVFSGDLPGKIKAKTDSLPA FLQAVFTSFLELPGFQPDIVVKPAMKQAAEQLLSYVQLEAGENQQAEPDPTNERDP ELRLVEWQKSLHEL SVRTEPF EFVRARDIDYYAETDRRGNR FVNITPEWTKFAESPF ARRLPKIPPEFCILLRRKTEGHAKIPNRIYLGQIFDGVTPDSTLGLVATAEDGKLFW WHDHLDEFSNLEGKPEPKLKNKPQLLMVSLEYDREQRFEESVGGDRKICLVTLKET RNFRRGRHGHTRTDRLPAGNTLWRADFATSAEVAAPKWNGRILGIHFQHNPVITWA LMDHDAEVLEKGFIEGNAFLGKALDKQALNEYLQKGGKWVGDRSFGNKLKGITHT LASLIVRLAREKDAWIALEEISWVQKQSADSVANRRFSMWNYSRLATLIEWLGTDI ATRDCGTAAPLAHKVSDYLTHFTCPECGACRKAGQKKEIADTVRAGDILTCRCKGF SGPIPDNFIAEFVAKKALERMLKKKPV
SEQ ID NO: 33	MAKRNFGEKSEALYRAVRFEVRPSKEELSILLA VSEVLRMLFNSALAERQQVFTEFI ASLYAELKSASVPEEISEIRKKLREAYKEHSISLFDQINALTARRVEDEAFASVTRNW QEETLDALDGAYKSFLSLRRKGDYDAHSRSDSGFFQKIPGRSGFKIGEGRIALSCG AGRKLSFPIPDYQQGRLAETTKLKKFELYRDQPNLAKSGRFWISVVYELPKPEATTC QSEQVAFVALGASSIGVVSQRGEEVIALWRSDKHWPKEIAVEERMKRRVKGSRG WLRLNSGKRRMHMISSRQHVDEREIVDYLVRNHGSHFVVTTEL VVRSGEKLAD SSKPERGGSGLGNWAAQNTGSLSRLVRQLEEKVKEHGGSVRKHKLTLTEAPPARGA ENKLWMARKLRESFLKEV
SEQ ID NO: 34	LAKNDEKELLYQSVKFEIYPDESKIRVLTRVSNILVLVWNSALGERRARFELYIAPLY EELKKFPRKSAESNALRQKIREGYKEHIPTFFDQLKLLTPMRKEDPALLGSVPRAY QEETLNTLNGSFVSFMTLRRNNDMDAKPPKGRAEDRFHEISGRSGFKIDGSEFVLST KEQKLRFPINPYQLEKLKEAKQIKKFTLYQSRDRRFWISIAYEIELPDQRPFNPEEVIYI AFGASSIGVISPEGEKVIDFWRPDKHWKPKIKEVENRMRSCKKGSRAWKKRAAARR KMYAMTQRQQKLNHREIVASLLRLGFHFVVTETVRSKPGKLADGSNPKRGGAPQ GFNWSAQNTGSFGEFILWLKQKVKEQGGTVQTFRLVLGQSERPEKGRDNKIEMVR LLREKYLESQTIVV
SEQ ID NO: 35	MAKGKKKEGKPLYRAVRFEIFPTSDQITLFLRVSKNLQQVWNEAWQERQSCYEQFF GSIYERIGQAKKRAQEAGFSEVWENEAKKGLNKKLRQQEISMQLVSEKESLLQELSI AFQEHGVTLYDQINGLTARRIIGEFALIPRNWQEETLDSL DGSFKSFLALRKNGDPDA KPPRQRVSENSFYKIPGRSGFKVSNGQIYLSFGKIGQTLTSVIPEFQLKRLETAIKLKK FELCRDERDMAKPGRFWISVAYEIPKPEKVPVVSQITYLAIGASRLGVVSPKGEFCL NLPRSDYHWKPQINALQERLEGVVKGSRKWKRMMAACTRMFAKLGHQKQHGQ YEVVKKLLRHGVHFVVTTELKVRSPGALADASKSDRKGSPTGPNWSAQNTGNIAR LIQKLTDKASEHGGTVIKRNPPLLSLEERQLPDAQRKIFIAKKLREEFLADQK
SEQ ID NO: 36	MAKREKKDDVVLRGTKMRIYPTDRQVTLMDMWRRRCISLWNLLNLETAAYGAK NTRSKLGWRSIWARVVEENHAKALIVYQHKGCKKDGSFVLKRDGTVKHPPRERFP GDRKILLGLFDALRHTLDKGAKCKCNVNQPYALTRA WLDETGHGARTADIHAWLK DFKGECDCTAISTA AKYCPAPPTAELLTKIKRAAPADDLPVDQAILLDLFGALRGGL KQKECDHTHARTVAYFEKHELAGRAEDILAWLIAHGGTCDCKIVEEAANHCPGPRL FIWEHELAMIMARLKAEPRTIEWIGDLPSHAAQT VVKDLVKALQTMLKERAKAAAG DESARKTGFPKFKKQAYAAGSVYFPNTT MFFDVAAGRVQLPNGCGSMRCEIPRQLV AELLERNLKPGLVIGAQLGLLGRIWRQGDRWYLSQWERPQPTLLPKTGRTAGVK IAASIVFTTYDNRGQTKEYPMPPADKCLTAVHLVAGKQNSRALEAQKEKEKKLKR KERLRLGKLEKGHDPNALKPLKRPRVRRSKLFYKSAARLAACEAIERDRRDGFLHR VTNEIVHKFDVSVQKMSVAPMMRRQKQKEKQIESKKNEAKKEDNGAAKKPRNL KPVRKLLRHVAMARGRQFLEYKYNDLRGPGSVLIADRLEPEVQECSRCGTKNPQM

SEQ ID NO	序列
	KDGRRLRCIGVLPDGTDCDAVLPRNRNAARNAEKRLRKHREAHNA
SEQ ID NO: 37	MNEVLPIPAVGEDAADTIMRGSKMRIYPSVRQAATMDLWRRRCIQLWNLLLELEQA AYSGENRRTQIGWRSIWATVVEDSHAEAVRVAREGKKRKDGTRKAPSGKEIPPLD PAMLAQIQRMNGAVDVPKTGEVTPAQPRLFMWEHELQKIMARLKQAPRTHWID DLPSHAAQSVVKDLIKALQAMLRERKKRASGIGGRDTGFPKFKKNRYAAGSVYFA NTQLRFEAKRGKAGDPDAVRGEFARVKLPNGVGWMECRMPRHINAAHAYAQATL MGGRIWRQGENWYLSQWKMPKPAPLPRAGRATAIKIAAAIPITTVDNRGQTREYA MPPIDRERIAAHAAAGRAQSRALARKRRRAKKREAYAKKRHAKKLERGIAAKPPGR ARIKLSPGFYAAAALAKLEAEDANAREAWLHEITTQIVRNFVDIAVPRMEVAKLM KKPEPPEEKKEQVKAPWQGRSLKAARVMMRRTAMALIQTTLKYKAVDLRGPQ AYEEIAPLDVTAACSGCGVLKPEWKMARAKGREIMRCQEPLPGGKTCNTVLTYT RNSARVIGRELAVRLAERQKA
SEQ ID NO: 38	MTTQKTYNFCFYDQRFELSKEAGEVYSRSLEEFWKIYDETGVWLSKFDLQKHMR NKLERKLLHSDSFLGAMQQVHANLASWKQAKKVVPDACP RP KPFLOAILFKKSQI KYKNGFLRLTLGTEKEFLYLKWDINIPLPIYGSVTYSKTRGWKINLCLETEVEQKNLS ENKYLSIDLGVKRVATIFDGENTITLSGKKFMGLMHYRNKLNKGKTSRSLSHKKKGS NNYKKIQRAKRKTTDRLLNIQKEMLHKYSSFIVNYAIRNDIGNIIIGDNSSTHDSNM RGKTNQKISQNPEQKLKNYIKYKFESISGRVDIVPEPYTSRKCPHCKNIKKSSPKGRT YKCKKCGFIFDRDGVGAINIYNENVSGQIISPGRIRSLTEPIGMKFHNEIFYKSYVAA
SEQ ID NO: 39	MSVRSFQARVECDKQTMELHWRTHKVFNERLPEIHKILFKMKRGECGQNDKQKSLY KSIQSILEANAQNADYLLNSVSIKGWKPGTAKKYRNASFTWADDAKLSSQGIHV YDKKQVLGDLPGMMSQMVCQRSVEAISGHIELTKKWEKEHNEWLKEKEKWESE EHKKYLDLREKFEQFEQSIGGKITKRRGRWHLYLKWLSDNPDFAAWRGNKAVINPL SEKAQIRINKAKPNKNSVERDEFFKANPEMKALDNLHGYYERNFVRRRKTKKNPD GFDHKPTFTLPHPTIHPRWVFVNKPKNPEGYRKLILPKKAGDLGSLEMRLLTGEKN KGNYPDDWISVKFKADPRLSLIRPVKGRRVVRKGKEQGQTKETDSYEFFDKHLKK WRPAKLSGVKLIFPDKTPKAALYFTCDIPDEPLTETAKKIQLWLETGDVTKKGKKR KKKVLPHGLVSCAVDLSMRRGTTGFATLCRYENGKIHLRSRNLWVGKKEGKGCH PYRWTEGPDGLHIAHKREIRILRSKRGPVKGEESHIDLQKHIDYMGEDRFKKAAR TIVNFALNTENAASKNGFYPRADVLLLENLEGLIPDAEKERGINRALAGWNRRHLVE RVIEMAKDAGFKRRVFEIPPYGTSSQVCSKCGALGRRYSIIRENNRREIRFGYVEKLFA CPNCGYCANADHNASVNLNRRFLIEDSFKSYDWRKRLSEKKQKEEETIESKLMKD LCAMHKISRGSSISK
SEQ ID NO: 40	MHLWRTHCVFNQRLPALLKRLFAMRRGEVGGNEAQRQVYQVVAQFVLARDAKDS VDLLNAVSLRKRSANSFAKKKATISCNGQAREVTGEEVF AE AVALASKGVFAYDK DDMRAGLPDSLFQPLTRDAVACMR SHEELVATWKKEYREWRDRKSEWEAEPEHA LYLNLRPKFEEGEAARGGRFRKRAERDHAYLDWLEANPQLAAWRRKAPPAVVPID EAGKRRIARAKAWKQASVRAEEFWKRNPELHALHKIHVQYLREFVRPRTRRNKR REGFKQRPTFTMPDPVRHPRWCLFNAPQTSPQGYRLLRLPQSRRTVGSVELRLLTGP SDGAGFPDAWVNVRFKADPRLAQLRPVKVPRTVTRGKNKGAKVEADGFRYYDDQ LLIERDAQVSGVKLLFRDIRMAPFADKPIEDRLLSATPYLVFAVEIKDEARTERAKAI RFDETSELTKSGKKRKTLPAGLVSAVDLDTRGVGFLLTRAVIGVPEIQQTHHGVRLL QSRVAVGQVEARASGEAEWSPGDLAHIARHKREIRRLRQLRGKPKVGERSHVR LAHIDRMGEDRFKKAARKIVNEALRGSNPAAGDPYTRADVLLYESLETLLPDAERE RGINRALLRWNRKALIEHLKRMCDAGIRHFPVSPFGTSQVCSKCGALGRRYSLAR ENGRAVIRFGWVERLFACPNPECPGRRPDRPDRPFTCNSDHNASVNLHRVFALGDQ AVAAFRALAPRDSPARTLAVKRVEDTLRPQLMRVHKLADAGVDSPF
SEQ ID NO: 41	MATLVYRYGVRAHGSARQQDAVSDPAMLEQLRLGHELRNALVGVQHRYEDGKR AVWSGFASVAAADHRVTTGETAVAELEKQARAHSADRTAATRQGTAE SLKAAR AAVKQARADRKAAMA AVAEQAKPKIQALGDDRDAEIKDLYRRFCQDGVLLPRCG RCAGDLRSDGDCTDCGAAHEPRKLYWATYNAIREDHQTAVKLVEAKRKAGQPAR LRFRRTWDGTLTVQLQRMHGPACRCVTC AEKLTRRARKTDPQAPAVAADPAYPP

SEQ ID NO	序列
	TDPPRDPALLASGQGKWRNVLQLGTWIPPGIEWSAMSRAERRRVGRSHIGWQLGGG RQLTLPVQLHRQMPADADVAMAQLTRVRVGGRRHMSVALTAKLPDPPQVQGLPP VALHLGWRQRPDGSRLRVATWACPQLDLPPAVADVVS HGGRWGEVIMPARWLA DAEVPRLGRRDKAMEPVLEALADWLEAHEACTARMTPALVRRWRSQGRLAG LTNRWRGQPPTGSAEILTYLEAWRIQDKLLWERESHLRRRLAARRDDAWRRVASW LARHAGVLVDDADIAELRRRDDPADTDPTMPASAAQAARARAALAAPGRLRHLA TITATRDGLGVHTVASAGLTRLHRKCGHQAQPDPRYAASAVVTCPGCGNGYDQDY NAAMLMLDRQQQP
SEQ ID NO: 42	MSRVELHRAYKFRLYPTPAQVAELAEWERQLRRLYNLAHSQRLAAMQRHVRPKSP GVLKSECLSCGAVAVAEIGTDGKAKKTVKHAVGCSVLECRSCGSPDAEGRTAHT AACSFVDYRQGREMTQLLEDDQLARVVCSARQETLRDLEKAWQRWHKMPGFG KPHFKKRIDSRIYFSTPKSWAVDLGYLSFTGVASSVGRIKIRQDRVWPGDAKFSSC HVVRDVDEWYAVFPLTFTKEIEKPKGGAVGINRGAVHAIADSTGRVVDSPKFYARS LGVIHRARLLDRKVPFGRAVKPSPTKYHGLPKADIDAAAARVNASPGRLVYEARA RGSIAAAEAHLAALVLPAPRQTSQPLSEGRNRERARRFLALAHQVRRRQREWFLHN ESAHYAQSYTKIAIEDWSTKEMTSSEPRDAEEMKRVTRARNRSILDVGWYELGRQI AYKSEATGAFAKVDPLRETETHVPEAIVRERDVDVSGMLRGEAGISGTCSRCGG LLRASASGHADAECVCLHVEVGDVNAAVNVLKRAMPFGAAPPSSKEKAKVTIGIK GRKKKRAA
SEQ ID NO: 43	MSRVELHRAYKFRLYPTPVQVAELSEWERQLRRLYNLGHEQRLTLTRHLRPKSPG VLKGECLSCDSTQVQEVGADGRPKTTVRHAEQCPTLACRSCGALRDAEGRTAHTV ACAFVDYRQGREMTLLAADDQLARVVCSARQEVLRDLDAWQRWRKMPGFG KPRFKRRTDSRIYFSTPKAWKLEGGHLSFTGAATTVGAIKMRQDRNWPASVQFSS CHVVRDVDEWYAVFPLTFVAEVARPKGGAVGINRGAVHAIADSTGRVVDSPRYYA RALGVIRHARLFDKRVPSGHAVKPSPTKYRGLSAIEVDRVARATGFTPGRVVTEAL NRGGVAYAEALAAIAVLGHGPERPLTSDGRNREKARKFLALAHQVRRRQREWFL HNESAHYARTYSKIAIEDWSTKEMTASEPQGEETRRVTRSRNRSILDVGWYELGRQ LAYKTEATGAFAQVDPGLKETETNPVKAIAADARDVDVSGMLRGEAGISGTCSKCG GLLRAPASGHADAECICLNVEVGDVNAAVNVLKRAMPFGDAPPASGEKPKVSIGI KGRQKKKAA
SEQ ID NO: 44	MEAIATGMSPEERRVELGILPGSVELKRAYKFRLYPMKVQQAELSEWERQLRRLYNL AHEQRLAALLRYRDWDFQKGACPSRVAVPGVHTAACDHVDYFRQAREMTQLE VDAQLSRVICCARQEVLRDLDAWQRWRKKGGRPRFKRRTDSRIYLSSTPKHWEI AGRYLRLSGLASSVGEIRIEQDRAFPEGALLSSCSIVRDVDEWYACLPLTFTQPIERAP HRSVGLNRGVVHALADSDGRVVDSPKFFERALATVQKRSRDLARKVSGSRNAHKA RIKLAKAHQVRRRQRAAFLHQESAYYSKGFDLVALEDMSVRKMTATAGEAPEMGR GAQRDLNRGILDVGWYELARQIDYKRLAHGGELLRVDPGQTTPLACVTEEQPARGI SSACAVCGIPLARPASGNARMRCTACGSSQVGDVNAAENVLTRALSSAPSGPKSPK ASIKIKGRQKRLGTPANRAGEASGGDPPVRGPVEGGTLAYVVEPVSESQSDT
SEQ ID NO: 45	MTVRTYKYRAYPTPEQAEALTSWLRFASQLYNAALEHRKNAWGRHDAHGRGFRF WDGDAAPRKKSDPPGRWVYRGGGGAHISKNDQGKLLTEFRREHAELLPPGMPALV QHEVLARLERSMAAFFQRATKGQKAGYPRWRSEHRYDSLTFGLTSPSKERFDPETG ESLGRGKTVGAGTYHNGDLRLTGLGELRILEHRRIPMGAIPKSVIVRRSGKRWFSI AMEMPSVEPAASGRPAVGLDMGVVTWGTAFTADTSAAAALVADLRRMATDPSCD RRLEELEREAQQLSEVLAHCRARGLDPAARRCPKELTKLYRRSLHRLGELDRACA RIRRRLOAAHDIAEPVPDEAGSAVLIEGSNAGMRHARRVARTQRRVARRTRAGHAH SNRRKKAVQAYARAKERERSARGDHRHKVSRLVRQFEEISVEALDIKQLTVAPEH NPDPQPDLPADVQRRNRGELDAAWGAFFAALDYKAADAGGRVARKPAPHTTQE CARCGTLVPKISLRVHRCPCGYTAPRTVNSARNVLQRPLEEPGRAGPSGANGRG VPHAVA
SEQ ID	MNCRYRYRIYPTPGQRQSLARLFGCVRVVWNDALFLCRQSEKLPKNSLQKLCITQ AKKTEARGWLQVSAIPLQSSVADLGVAFKNFFQSRSGKRKGKKNPPRVKRRNN

SEQ ID NO	序列
NO: 46	RQGARFTRGGFKVKTskvylarigdikikwsrplpsepssvtvikdcagqyflsfvve vkpeikppknpsigidlglktfascsngekidspdy srl yrklkrcqrrlakrqrgsk rrermrvkvaklnaqirdkrkdflhklstkvvnenvialedlnvggmlknrkl raisqagwyefrslcegkaekhnrdfrvisrweptsqvcsecgyrwgkidlsvrsiv cincgvehdrddnasvniesqaglkvgvghthdskrtgsacktsngavcvepsthre yvqltlfDW
SEQ ID NO: 47	MKSRWTFRCYPTPEQEQHLARTFGCVRFVWNWALRARTDAFRAGERIGYPATDKA LTLKQQPETVWLNEVSSVCLQQALRDLQVAFSNFFDKRAAHPSFKRKEARQSANY TERGFSFDHERRILKLAKIGAIKVKWSRKAIPHPSSIRLIRTASGKYFVSLVETQPAP MPETGESVGVDGVARLATLSNGERISNPKHGAKWQRRALAFYQKRLARATKGSKR RMRIKRHVARIHEKIGNSRSDTLHLKSTDLVTRFDLICVEDLNLGMVKNHSLARSL HDASIGSAIRMIEEKAERYGKNVVKIDRWFPSSKTCSDCGHIVEQLPLNVREWTCPE CGTTHDRDANAAANILAVGQTVSAHGGTVRRSRAKASERKSQRSANRQGVNRA
SEQ ID NO: 48	KEPLNIGKTAKAVFKEIDPTSLNRAANYDASIELNCKECKFKPFKNVKRYEFNFYNN WYRCNPNSCLQSTYKAQVRKVEIGYEKLKNEILTQMYYYPWFGRLYQNFFHDERD KMTSLDEIQVIGVQNKVFFNTVEKAWREIHKRFRDNKETMETIPELKHAAGHGKR KLSNKSLLRRRFAFVQKSFKFVDSVSYRSFSNNIACVLP SRIGVDLGGVISRNPKR EYIPQEISFNAFWKQHEGLKGRNIEIQSVQYKGETVKRIEADTGEDKAWGKNRQR RFTSLILKLVPKQGGKKVWKYPEKRNENYEFPIPIEFILDSGETSIRFGGDEGEAG KQKHLVIPFNDSKATPLASQQTLLNSRFNAEVKSCIGLAIYANYFYGYARNYVISSI YHKNSKNGQAITAIYLESIAHNYVKAIERQLQNLLNLNDRDFSMESHKKELKKYFGG DLEGTGGAQKRREKEEIEKEIEQSYLPRLIRLSLTKMVTQVEM
SEQ ID NO: 49	ELIVNENKDPNIGKTAKAVFKEIDPTSLNRAANYDASIELNCKECKFKPFKNNTKRHD FSFYSNWHRCSPNSCLQSTYRAKIRKTEIGYEKLKNEILNQMYYYPWFGRLYQNFFN DQRDKMTSLDEIQVTGVQNKIFFNTVEKAWREIHKRFRDNKETMRTIPDLKNKSGH GSRKLSNKSLLRRRFAFAQKSFKLVDNSVSYRAFSNNVACVLP SKIGVDIGGIINKD LKREYIPQEITFNFWKQHDGLKGRNIEIHSVQYKGEIVKRIEADTGEDKAWGKNR QRRFTSLILKITPKQGGKKIWKPFPEKKNASDYEFPIPIEFILDNGDASIKFGGEEGEV GKQKHLVIPFNDSKATPLSSQMLLETSRFNAEVKSTIGLALYANYFVSYARNYVIK STYHKNSKKGQIVTEIYLESISQNFVRAIQRLQSLMLNLKDWGFMQTHKKELKKY FGSDLEGSKGGQKRREKEEIEKEIEASYLPRLIRLSLTKSVTKAEM
SEQ ID NO: 50	PEEKTSKLPNSINLAANYDANEKFNCKECKFHPFKNKKRYEFNFYNNLHGCKSCT KSTNNPAVKRIEIGYQKLKFEIKNQMEAYPWFGRLRINFYSDEKRKMSELNEMQVT GVKNKIFFDAIECAWREILKRFRESKETLITIPKLKNKAGHGARKHRNKKLLIRRA FMKKNFHFLDNDSISYRSFANNIACVLP SKVGVDIGGISPDVGKDIKPVDISLNLMW ASKEGIKSGRKVEIYSTQYDGNMVKKIEAETGEDKSWGKNRKRRTSLLSIPKPSK QVQEFDFKEWPRYKDIEKKVQWRGFIKIFDSNHNSIEFGTYQGGKQKVLPIPFNDS KTTPLGSKMNKLEKLRFNKSKISRLGSAIAANKFLEAARTYCVDSLYHEVSSANAIG KGKIFIEYYLEILSQNYIEAAQKQLQRFIESIEQWFVADPFQGRLLKQYFKDDLKRAKC FLCANREVQTTCYAAVKLHKSCAEKVKDKNKELAIKERNNKEDAVIKEVEASNYPR VIRLKLTKTITNKAM
SEQ ID NO: 170	SESENKIEQYYAFLYSFRDKYEKPEFKNRGDIKRLQNKWEDFLKEQNLKNDKKLS NYIFSNRNFRRSYDREENEEDGIDEKSKPKRINCFEKEKNLKDQYDKDAINASANK DGAQKWGCFECIFFPMYKIESGDPNKRRIINKTRFKLFDLYNLKGCKSCLRSTYHPY RSNVYIESNYDKLKREIGNFLQKNIFQRMRAKVSEGKYLTLNDEYRLSCVAMHF KNRWLFFDSIQKVLRETIKQRLKQMRESYDEQAKTKRSKGHGGRACYEDQVRMIRR RAYSAQAHKLLDNGYITLFDYDDKEINKVCLTAINQEGFDIGGYLNSDIDNVMPPIEI SFHLKWKYNPILNIESPFSKAKISDYLRKIREDLNLERGKEGKARSKKNVRRKVLAS KGEDGYKKIFTDFFSKWKEELEGNAMERVLSQSSGDIQWSKKKRIHYTTLVNLNLL DKKGVGNLKYIEAEKTKILSFDKNENKFWPITIQLLDGYEIGTEYDEIKQLNEKTS KQFTIYDPNTKIIKIPFTDSKAVPLGMLGINIATLTKTVKKTERDIKVSIFKGGLNSKIV SKIGKGIYAGYFPTVDKEILEEVEEDTLNDEFSSKSQRNIFLKSIIKNYDKMLKEQLFD

SEQ ID NO	序列
	FYSFLVRNDLGVRFLLTDRELQNIEDESFNLEKRFFETDRDRIARWFDNTNTDDGKEKFKKLANEIVDSYKPRILRLPVVRVIKRIQPVKQREM
SEQ ID NO: 51	KYSTRDFSELNEIQVTACKQDEFFKVIQNAWREIHKRFLNRENFIEKKIFKNKKGRGKRQESDKTIQRNRASVMKNFQLIENEKIILRAPSGHVACVFPVKVGLDIGGFKTDDLEKNIFPPRTITINVFWKNRDRQRKGRKLEVWGKARTKLIK VHKWDKLEEVKKKRLKSLEQKQEKSLDNWSEVNNSDFYKVQIDELQEKIDKSLKGRTMNKILDNKAKESKEAEGLYIEWEKDFEGEMLRRIEASTGGEEKWGKRRQRRHTSLLLDIKNNSRGSKEIINFYSYAKQGKKEKKIEFFPPLTITLDAEEESPLNIKSIPIEDKNATSKYFSPFTETRA TPLSILGDRVQKFKTKNISGAIKRNLGSSISSCKIVQNAETSAKSILSLPNVKEDNNMEIFINTMSKNYFRAMMKQMESFIFEMEPKTLIDPYKEKAIKWFEVAASSRAKRLKKLSKADIKKSELLSNTTEFEKEKQEKLEALEKEIEEFYLPRIVRLQLTKTILETPVM
SEQ ID NO: 52	KKLQLLGHKILLKEYDPNAVNAANFETSTAELCGQCKMKPFKNRRFQYTFGKNYHGCLSCIQNVYYAKKRIVQIAKEELKHQLTDSIASIPYKYSLSFNTNSIDELYILKQERAAFFSNTNSIDELYITGIENNIAFKVISAIWDEIHKRRQRYAESLTDGTGTVKANRGHGGTAYKSNTRQEKIRALQKQTLHMVTNPYISLARYKNNYIVATLPRITIGMHIGAIKDRDPQKKLSDYAINFNVFWSSDRQLIELSTVQYTGDMVRKIEAETGENNKWGENMKRTKTSLLLEILTKKTTDELTFKDWAFSTKKEIDSVTKKTYQGFPIGIFEGNESSVKFGSQNYFPLPFDKITPPTAEGFRDLWLKGSFSSQMKTSYGLAIYSNKVITNAIPAYVINKNMFYKIARAENGKQIKAKFLKKYLDIAGNNYVPFIIMQHYRVLDTFEEMPISQPKVIRLSLTKTQHIIKKDKTDSKM
SEQ ID NO: 53	NTSNLINLGKKAINISANYDANLEVGCNCKFLSSNGNFPRTQTNVKEGCHSCEKSTYEPSIYLVKIGERKAKYDVLDSLKKFTFQSLKYQSKKSMKSRNKKPKELKEFVIFANKNKAFDVIQKSYNHLILQIKKEINRMNSKKRKNHKKRRLFRDREKQLNKLRLIESSNLFLENKGNHVFYVAIHSVGRDIGVIGSYDEKLNFELETYQLYFNDDKRLLYAYKPKQNKIIEKEKLWNLKKEKEPLDLEYEKPLNKSITFSIKNDNLFKVSKDLMLRRAKFNIQKKEKLSKEERKINRDLIKIKGLVNSMSYGRFDELKKEKNIWSPHIYREVRQKEIKPCLIKNGDRIEQLKKKMERLRRFREKRQKKISKDLIFAERIAYNFHTKSIKNTSNKINIDQEA KR GKASYMRKRIGYETFKNKYCEQCLSKGNVYRNQVQKGCSCFENPFDWIKKGDENLLPKKNEDLRVKGAFRDEALEKQIVKIAFNIAKGYEDFYDNLGESTEKDLKLFKVGTINEQESLKL
SEQ ID NO: 54	TSNPIKLGGKAINISANYDSNLQIGCKNCKFLSYNGNFPRTQTNVKEGCHSCEKSTYEPVYTVRIGERRSKYDVLDSLKKFTFQSLKYQSKKMKTRSKGIRGLEEFVISANLKKAMDVIQKSYRHLILNIKNEIVRMNGKKRNKNHKKRLLFRDREKQLNKLRLIEGSSFFKPPTVKGDNSIFTCVAIHNIGRDIGIAGDYFDKLEPKIELTYQLYEYNPKKESKINKRLLYAYKPKQNKIIEKEKLWNLKKEKSPDLLEYEKPLTKSITFLVKRDGVFRISKDLMLRKAKFIIQKKEKLSKEERKINRDLIKIKSNIISLTYGRFDELKKDKTIWSPHIFRDVKQKITPCIERKGDRMDIFQQLRKKSERLRENRRKQKKISKDLIFAERIAYNFHTKSIKNTSNLINIKHEAKRGKASYMRKRIGNETFRIKYCEQCFPKNNVYKNVQKGCSCFEDPFEYIKKGNEIDLIPNKNQDLKAKGAFRDDALEKQIIKVAFNIAKGYEDFYENLKKTTTEKDIRLKFKVGTIIEEM
SEQ ID NO: 55	NNSINLSKKAINISANYDANLQVRCKNCKFLSSNGNFPRTQTNVKEGCHSCEKSTYEPVYDVKIGIEKAKYEVLDLKLKFTFQSLKYQSKKMKFRSKKIKELKEFVIFAKESKALNVINRSYKHLILNIKNDINRMNSKKRIKHNHKGRLFLDRQKQLSKLKLIEGSSFFVPAKNVGNKSVFTCVAIHSIGRDIGIAGLYDSFTKPVNEITYQIFFSGERRLLYAYKPKQLKILSIKENLWSLKNEKKPLDLYEKPLGKNLNFNVKGGDLFRVSKDLMIRNAKFNVHGRQRLSDEERLINRNFIKIKGEVVSLSYGRFEELKKDRKLWSPHIFKDVQRNKIKPCLVMQGGQRIDIFEQLKRKLELLKKIRKSQKKLSKDLIFGERIAYNFHTKSIKNTSNKINIDSDAKRGRASYMRKRIGNETFKLYCDVCFKANVYRRVQNGCSCSENPNYIKKGDKDLLPKKDEGLAIKGAFRDEKLNKQIIKVAFNIAKGYEDFYDDLKRTKEDVDLKFKIGTTVLDQKPMIEFDGIVITWL
SEQ ID	LLTTVVETNNLAKKAINVAANFDANIDRQYYRCTPNLCRFIAQSPRETKEKDAGCSSCTQSTYDPKVYVIKIGKLLAKYEILKSLKRFLFMNRYFKQKKTERAQQKQKIGTELN

SEQ ID NO	序列
NO: 56	EMSIFAKATNAMEVIKRA TKHCTYDIIPETKSLQMLKRRRHRVKVRSLLKILKERRM KIKKIPNTFIEIPKQAKKNKSDYYVAAALKSCGIDVGLCGAYEKNAEVEAEYTYQLY YEEKGNSSTKRILYCYNPNQKNIREFWFAFYIQGSKSHVNTPGTIRLKMFKLSPITIE SEALDFRVWNSDLKIRNGQYGFIIKRS LGKEAREIKKGMGDIKRKIGNLTYGKSPSE LKSIVHYRTERENPKKPRAARKKEDNFMEIFEMQRKKDYEVNKKRRKEATDAAKI MDFAEPIRHYHTNNLKAVRRIDMNEQVERKKT SVFLKRIMQNGYRGNYCRKCIKA PEGSNRDENVLEKNEGCLDCIGSEFIWKKSSKEKKGLWHTNRLRRIRLQCFTTAKA YENFYNDLFEKKESSLDIIKLKVSITTKSM
SEQ ID NO: 57	ASTMNLAKQAINFAANYDSNLEIGCKGCKFMSTWSKSNPKFYPRQNNQANKCHS CTYSTGEPEVPIIEIGERA AKYKIFTALKKFVFM SVAYKERRRRQRFKSKPKELKELA ICSNREKAMEVIQKSVVHCYGDVKQEIPRIRIKVLKNHKGRLFYKQKRSKIKIAKLE KGSFFKTFIPKVHNNGCHSCHEASLNKPILVTTALNTIGADIGLINDYSTIAPTETDIS WQVYYEFIPNGDSEAVKKRLLYFYKPKGALIKSIRDKYFKKKGHENAVNTGFFKYQG KIVKGPIKFVNNELDFARKPDLKSMKIKRAGFAIPSAKRLSKEDREINRESIKIKNKIY SLSYGRKKTLSDKDIIKHL YRPVRQKGVPKPLEYRKAPDGFLEFFYSLKRKERRLRKQ KEKRQKDMSEIIDA ADEF AWHRHTGSIKKT TNHINFKSEVKRGKVPIMKKRIANDSF NTRHCGKCVKQGNAINKYIEKQKNCFCDCNSIEFKWEKA ALEKKGAFKLNKRLQYI VKACFNVAKAYESFYEDFRKGEEESLDLKFKIGTTTTLKQYPQNKARAM
SEQ ID NO: 58	HSHNMLTKLGKQAINFAANYDANLEIGCKNCKFLSYSPKQANPKKYPRQTDVHED GNIACHSCMQSTKEPPVYIVPIGERKSKYEILTSLNKFTFLALKYKEKKRQAFRAKKP KELQELAI AFNKEKA I KVIDKSIQHLILNIKPEIARIQRQKRLKNRKGLLYLHKRYAI KMGLIKNGKYFKVGSPKKGKLLVLCALNTIGRDIGIIGNIEENNRSETEITYQLYF DCLDANPNELRIKEIEYNRLKSYERKIKRLVYAYKPKQTKILEIRSKFFSKGHENKVN TGSFNFNENPLNKSISIKVKN SADF KIGAPFIMLRNGKFHIPTKKRLSKEEREINRTLSK IKGRVFR LTYGRNISEQGS KSLHIYRKERQHPKLSLEIRKQPD SFIDEFEKLR LKQNFIS KLKKQRQKKLADLLQFADRIAYNYHTSSLEKTSNFINYKPEVKRGRTSYIKKRIGNE GFEKLYCETCIKSNDKENAYAVEKEELCFVCKAKPFTWKKTNKDKLGIFKYPRIKD FIRAAFTVAKSYNDFYENLKKKDLKNEIFLKFKIGLILSHEKKNHISIAKSAEDERIS GKSIKNILNKSILEKNCYSCFFHKEDM
SEQ ID NO: 59	SLERVIDKRNLAKKAINIAANFDANINKGFYRCETNQCMFIAQKPRKTNNTGCSSCL QSTYDPVIYVVKVGEMLA KYEILKSLKRFVFMNRSFKQKKTEKAKQKERIGGELNE MSIFANAALAMGVIKRAIRHCHVDIRPEINRLSELKKT KHRVA AKSLVKIVKQRKTK WKGIPNSFIQIPQKARNKDADFYVASALKSGGIDIGLCGT YDKKPHADPRWTYQLYF DTEDESEKRLLYCYNDPQAKIRDFWKT FYERGNPSMVNSPGTIEFRMEGFFEKMTPI SIESKDFDFRVWNKD L LIRRG LYEIKKRKNLNRKAREIKKAMG SVKRVLANMTY GK SPTDKKSIPVYRVEREKPKPRAVRKEENELADKLENYRREDFLIRNRRKREATEIA KIIDA AEPIRHYHTNHLRAVKRIDLSKP VARKNTSVFLKRIMQNGYRGNYCKKCIK GNIDPNKDECRLEDIKKICCEGTQNIWAKKEKLYTGRINVLNKRKIQMKLECFNVA KAYENFYDNLAALKEGDLKVLKLV SIPALNPEASDPEEDM
SEQ ID NO: 60	NASINLGKRAINLSANYDSNLVIGCKNCKFLSFNGNFPRQTNVREGCHSCDKSTYAP EVYIVKIGERKAKYDVLDSLKKFTFQSLKYQIKKSMRERSKKPKELLE FVIFANKDK AFNVIQKSYEHLILNIKQEINRMNGKKRIKNHKKRLFKDREKQLNKLRLIGSSSLFFP RENKGDKDLFTYVAIHSVGRDIGVAGSYESHIEPISDLTYQLFINNEKRLLYAYKPKQ NKIIELKENLWNLKKEKKPLDLEFTKPLEKSITFSVKNDKLFK VSKDLMLRQAKFNI QGKEKLSKEERQINRDFSKIKSNVISLSYGRFEELKKEKNIWSPHIYREV KQKEIKPCI VRKGDRIELFEQLKRKMDK LKKFRKERQKKISKDLNFAERIAYNFHTKSIKNTSNKI NIDQEAKRGKASYMRKRIGNESFRKKYCEQCFSVGNVYHNVQNGCSCFDNPIELIK KGDEGLIPKGKEDRKYKGALRDDNLQMQUIRVAFNIAKGYEDFYNNLKEKTEKDLK LKFKIGTTISTQESNNKEM
SEQ ID NO:	SNLIKLGKQAINFAANYDANLEVGCKNCKFLSSTNKYPRQTNVHLDNKMACRSCN QSTMEPAIYIVRIGEKKAKYDIYNSLT KFNQSLKYKAKRSQRFPKPKQPKELQELSIA VRKEKALDIIQKSIDHLIQDIRPEIPRIKQOKRYKNHVGKLFY LQKRRKNKLNLIKGK

SEQ ID NO	序列
61	SFFKVFSPKEKKNELLVICALTNIGRDI GLIGNYNTIINPLFEV TYQLYDYIPKKNNK NVQRRLLYAYKSKNEKILKLEAFFKRGHENAVNLGSFSYEKPLEKSLTLKIKNDKD DFQVSPSLRIRTGRFFVPSKRNLSRQEREINRRLVKIKSKIKNMTYKGFETARDKQSV HIFRLERQKEKLPLQFRKDEKEFMEEFQKLKRRTNSLKKLRKSQKKLADLLQLSEK VVYNNHTGTLKKT SNFLNFSSSVKRGKTAYIKELLGQEGFETLYCSNCINKGQKTRY NIETKEKCFSCKDVPFVWKKKSTDKDRKGAF LPAKLKDV IKAFTTVAKAYEDFYD NLKSIDKKPYIKFKIGLILAHVRHEHKARAKEEAGQKNIYNKPIKIDKNCKE CFFFK EEAM
SEQ ID NO: 62	NTTRKKFRKRTGFPQSDNIKLAYCSAIVRAANLDADIQKKHNQCNPNLCVGIKSNEQ SRKYEHS DRQALLCYACNQSTGAPKVDYIQIGEIGAKYKILQMVNAYDFLSLAYNL TKLRNGKSRGHQRMSQLDEVVIVADYEKATEVIKRSINHLLDDIRGQLSKLKKRTQ NEHITEHKQSKIRRLKRLSRLLKRRRWK WGTIPNPYLKNWVFTKKDPELVTVALL HKLGRDIGLVNRSKRRSKQKLLPKVGFQLYYKWESP SLNNIKSKAKKLPKRLLIPY KNVKLFDNKQKLENAIKSLLESYQKTIKVEFDQFFQNRTEEIIAEEQQTLE RGLLKQL EKKKNEFASQKKALKEEKKKIKEPRKAKLLMEESRSLGFLMANVSYALFNTTIEDL YKKS NVVSGCIPQEPV VVPADIQNGSLAKILFAPKDGFRIFSGQHLTIRTA KFKIR GKEIKILT KTKREILKNIEKLRRVWYREQHYK LKLFGKEVSAKPRFLDKR KTSIERRD PNKLADQTDDRQAE LRNKEYELRHKQHKMAERLDNIDTNAQNLQTL SFWVGEAD KPKLDEK DARGFGVRTCISAWK WFMEDLLKKQEEDPLLKLKLSIM
SEQ ID NO: 63	PKPKFKQKRTGFPQPDNLRKEYCLAIVRAANLDADFEKKCTKCEGIKTNKKGNIVK GRTYNSADKDNLLCYACNISTGAPAVDYVFGALEAKYKILQMVKAYDFHSLAYN LAKLWKGRGRGHQRMGGLNEVVIVSNNEKALDVIEKSLNHFHDEIRGELSRLKAKF QNEHLHVH KESKLRRKLRLKISRLKRRRWKWDVIPNSYL RNF TTKTRPDFISVALL HRVGRDIGLVTKTKIPKPTDLLPQFGFIYYTWDEPKLNLKKSRLRSEPKRLLVPY KKIELYKNKSVLEEAI RHLAEVYTEDLTICFKDFFETQKRKFVSKEKESL KRELLKEL TKLKKDFSERKTALKRDRKEIKEPKKAKLLMEESRSLGFLAANTS YALFNLIAADLY TKSKKACSTKLPRQLSTILPEIKEHKSTTSLAIKPEEGFKIRFSNTHLSIRTPKFKMKG ADIKALTKRKREILKNATKLEKSWYGLKHYK LKLYGKEVA AKPRFLDKRNPSIDRR DPKELMEQIENRRNEVKDLEYEIRKGQH QMAKRLDNVD TNAQNLQTKSFWVGEAD KPPELDSMEAKKLGLRTCISAWK WFMKDLVLLQEKSPNLKLKLSLTEM
SEQ ID NO: 64	KFSKRQEGFLIPDNIDLYKCLAIVRSANLDADVQGHKSCYGVKKNGTYRVKQNGKK GVKEKGRKYVFDLIAFKGNIEKIPHEAIEEKDQGRVIVLGKFNYKLILNIEKNHNDRA SLEIKNKIKLVQISSLETGEFLSDLGKIGIDEVYGIIEPDVFS GKELVCKACQSTY APLVEYMPVGELDAKYKILSAIKGYDFLSLAYNLSRNRANKKRGHQKLGGGELSEV VISANYDKALNVIKRSINH YHVEIKPEISK LKKKMQNEPLKVMKQARIRRELHQLSR KVKRLKWKWGMIPNPELQNIIFEKKEKDFVSYALLHTLGRDIGL FKDT SMLQVPNIS DYGFQIYYSWEDPKLNSIKKIKDLPKRLLIPYKRLDFYIDTILVAKVIKNLIELYRKS Y VYETFGEEYGYAKKAEDILFDWDSINLSEGIEQKIQKIKDEFSDLLYEARESKRQNFV ESFENILGLYDKNFASDRNSYQEKIQSMIKKQQENIEQKLKREFKEVIERGFEGMDQ NKKYYKVLSPNIKGGLLYTD TNLGGFRSHLAFMLLSKISDDL YRKNNLVSKGGNK GILDQTPETMLTLEFGKSNLPNISIKRKF FNIKYNSSWIGIRKPKFSIKGAVIREITKKV RDEQRLIKSLEGVWHKSTHFKRWGKPRFNLPRHPDREKNND DNLMESITSRREQIQL LLREKQKQQEKMAGRLDKIDKEIQNLQTANFQIKQIDKKPALTEKSE GKQSVRNALS AWKWFMEDLIKYQKRTPIQLKLAKM
SEQ ID NO: 65	KFSKRQEGFVIPENIGLYKCLAIVRSANLDADVQGHVSCYGVKKNGTYVLKQNGKK SIREKGRKYASDLVAFKGDIEKIPFEVIEEKKKEQSIVLGKFNYKLVLDMKGEKDR ASLTMKNKSKKL VQVSSLGTDEFLLTLLNEKFGIEEYGIIEPEVFS GKKL VCKACQQ STYAPLVEYMPVGELDSKYKILSAIKGYDFLSLAYNLARHRSNKKRGHQKLGGGEL SEVVISANNAKALNVIKRSLNHY YSEIKPEISKLRKKMQNEPLKVGKQARMRRELH QLSRKVKRLKWKWGKIPNLELQNTFKESDRDFISYALLHTLGRDIGM FNKTEIKMP SNILGYGFQIYYDWE EPKLNTIKKSKNTPKRILIPYKKLDFYND SILVARAIKELVGLF QESYEWEIFGNEYNYAKEAEVELIKLDEESINGNVEKKLQRIKENFSNLLEKAREKK

SEQ ID NO	序列
	RQNFIESFESIARLYDESFTADRNEYQREIQSFHIEKQKQKSIEKKLKNEFKKIVEKKFNE QEQQGKKHYRVLNPTINEFLPKDKNNLGFLRSKIAFILLSKISDDLYKKSSNAVSKGGE KGIKQQPETILDLEFSKSKLPSINIKKKLFNIKYTSSWLGRKPKFNIKGAKIREITRRV RDVQRTLKSAESSWYASTHFRRWGFPRFNQPRHPDKEKKSDDRLIESITLLREQIQL LREKQKGQKEMAGRLDDVDKKIQNLQTANFQIKQTGDKPALTEKSAGKQSFERNAL SAWKWFMENLLKYQNKTPDLKLKIARTVM
SEQ ID NO: 66	KWIEPNNIDFNKCLAITRSANLDADVQGHKMCYGIKTNGTYKAIGKINKKHNTGIIIE KRRTYVYDLIVTKEKNEKIVKKTDMAIDEEIEFDEKKEKLLKKYIKAEVLGTGELIR KDLNDGEKFDDLCSEEPQAFRRSELVCKACNQSTYASDIRYIPIGEIEAKYKILKAIK GYDFLSLKYNLGRLRDSKKRGHQKMGQGELKEFVICANKEKALDVIKRSNLNHYLN EVKDEISRLNKKMQNEPLKVNDQARWRRELNQISRRLKRLKWKWGEIPNPELKNLI FKSSRPEFVSIALIHTLGRDIGLINETELKPNNIQEYGFQIYYKWEDPELNHIKKVKNI PKRFIIPYKNLDFGKYTILSRAIEGILKLYSSSFQYKSFKDPNLFKEGEGKKITNEDFE LGYDEKIKKIKDDFKSYKKALLEKKKNTLEDLSNLSVYEQSLLTEQINNKKWKE GLLKSKESEIHKQKKIENIEDIISRIEELKNVEGWIRTKERDIVNKEETNLKREIKKELKD SYEEVRKDFSDLKKGESEKKPFREEPKPIVKDYIKFDVLPGENSALGFFLSHLSFN LFDIQYELFEKSRLSSSKHPQIPETILD
SEQ ID NO: 67	FRKFVKRSGAPQPDNLNKKYKCIIVRAANLDADIMSNESNCVMCKGIKMNRKTA KGAAKTTELGRVYAGQSGNLLCTACTKSTMGPLVDYVPIGRIRAKYTILRAVKEYD FLSLAYNLARTRVSKKGGRRQKMHSLSSELVIAAEYEIAWNIKSSVIHYHQUETKEEISG LRKKLQAEHIIHKNKEARIRREMHQISRRIKRLKWKWHMIPNSELHNFLFKQQDPSFV AVALLHTLGRDIGMINKPKGSAKREFIPEYGFQIYYKWMNPKLNDINKQKYRKMPK RSLIPYKNLNVFGDRELIENAMHKLKLYDENLEVKGSKFFKTRVVAISSKESEKLK RDLLWKGELAKIKKDFNADKNKMQELFKEVKEPKKANALMKQSRNMGFLLQNISY GALGLLANRMYEASAKQSKGDATKQPSIVIPLEMEFGNAFPKLLLRSGKFAMNVSS PWLTIKPKFVIKGNKIKNITKLMKDEKAKLRLETSYHRATHFRPTLRGSIDWDSP YFSSPKQPNTHRRSPDRLSADITEYRGRLKSVEAELREGQRAMAKKLDSDVDMTASN LQTSNFQLEKGEDPRLTEIDEKGRSIRNCISSWKKFMEDLMKAQEANPVIKIKIALKD ESSVLSSEDSM
SEQ ID NO: 68	KFHPENLNKSYCLAIVRAANLDADIQGHINCIGIKSNKSDRNYENKLESQNVELLC KACTKSTYKPNINSVPVGEKKAKYSILSEIKKYDFNSLVYNLKKYRKGKSRGHQKL NELRELVTSEYKKALDVINKSVNHYLVNIKKNMSKLLKILQNEHIIHVGTLARIRRE RNRISRKL DHYRKKWKFPVNKILKNYVFNQSPDFVSVALHLKLRDIGLITKTAIL QKSFPEYSLQLYYKYDTPKLNLYKKSKFKSLPKRILISYKYPKFDINSNYIEESIDKLL KLYEESPIYKNNSKIIEFFKKSSEDNLKSENDSLKRGIMKEFEKVTKNFSSKKKKLKEE LKLKNEDKNSKMLAKVSRPIGFLKAYLSYMLFNIISNRIFEFSRKSSGRIPQLPSCIINL GNQFENFKNELQDSNIGSKKNYKYFCNLLKSSGFNISYEEHLSIKTPNFFINGRKL KEITSEKKKIRKENEQLIKQWKLTFFKPSNLNGKKTSDKIRFKSPNPNPIERKSEDNI VENIAKVYKLEDLLSEQRKEFNKLAKKHGDGVDVEAQCLQTKSFWDNSNPIKKSLE KKNEKVSVKKKMKAIRSCISAWKWFMA DLIEAQKETPMIKLKLALM
SEQ ID NO: 69	TTLVPSHLAGIEVMDETTSRNEDMIQKETSRSNEDENYLGVKNCGINVHKSGRGSS KHEPNMPPEKSGEGQMPKQDSTEMQQRFDSESVTGETQVSAGATASIKTDARANS GP RVGTARALIVKASNLDRDIKLGSCKPCEYIRSELP MGKKNGCNHCEKSSDIASVPKVE SGFRKAKYELVRRFESFAADSISRHLGKEQARTRGKRGKKDKKEQMKGKVNLEIAI LKNESLIEYTENQILDARSNRIKEWLRSLRLRLRTRNKGLKKSIRRQLITLRRDYR KWIKPNPYRPDEDPNENSLRLHTKLGV DIGVQGGDNKRMNSDDYETSF SITWRDTA TRKICFTKPKGLLPRHMKFKLRGYPELILYNEELRIQDSQKFPLVDWERIPIFKLRGVS LGKKKV KALNRITEAPRLVVAKRIQVNIESKKKKVLTRYVYNDKSINGRLVKAEDS NKDPLLEFKKQAEIINSDAKYYENQEIAKNYLWGCEGLHKNLLEEQTKNPYLAFKY GFLNIV
SEQ ID	LDFKRTCSQELVLLPEIEGLKLSGTQGVTS LAKKLINKAANVDRDESYGCHHC IHTR TSLSKPVKKDCNSCNQSTNHPAVPITLKG YKIAFYELWHRFTSWAVDSISKALHRNK

SEQ ID NO	序列
NO: 70	VMGKVNLD EYAVVDNSHIVCYAVRKCYEKRQRSVRLHKRAYRCRAKHYNKSQPK VGRIYKKSKRRNARNLKEAKRYFQNEITNGSSDALFYKIGVDLGIAGTPETEVK VDVSICFQVYYGDARRVLRVRKMDLQSFHLDYTGKLLKLGIGNKDTFTIAKRNES LKWGSTKYEVSR AHKKFKPFGKKGSVKRKCNDYFRSIAWSCEAASQRAQSNLKN AFPYQKALVKCYKNLDYKGVKKNDMWYRLCSNRIFRYSRIAEDIAQYQSDKGKAK FEFVILAQSVAEYDISAIM
SEQ ID NO: 71	VFLTDDKRKTALRKIRSAFRKTAEIALVRAQEADSLDRQAKKLTIVTSFGAPGAKN AFIGSLQGYNWN SHRANVPSSGSAKDVFRI TELGLGIPQSAHEASIGKSFELVGNVVR YTANLLSKGYKKGAVNKGAKQQREIKGKEQLSFDLISNGPISGDKLINGQKDALAW WLIDKMGFHI GLAMEPLSSPNTYGITLQAFWKRHTAPRRYSRGVIRQWQLPFGRL APLIHNFRKKGASIPVLTNASKKLAGKGVLLEQTALVDPKKWWQVKEQVTGPLS NIWERSVPLVLYTATFTHKHGA AHKRPLTLK VIRISSGVSFLLPLSKVTPGKLVRAW MPDINILRDGRPDEAA YKGPDLIRARERSFPLAYTCVTQIADEWQKRALESNRDSITP LEAKLVTGSDLLQIHSTVQQAVEQIGGRISPIQELLAKDALQLVLQQLFMTVDLL RIQWQLKQEVADGNTSEKAVGWAIRISNIHKDAYKTAIEPCTSALKQAWNPLSGFEE RTFQLDASIVRKRSTAKTPDDELVIVLRQQA AEMTVAVTQSVSKELMELAVRHSAT LHLLVGEVASKQLSRSADKDRGAMDHWKLLSQSM
SEQ ID NO: 72	EDLLQKALNTATNVAAIERHSCISCLFTESEIDVKYKTPDKIGQNTAGCQSCTFRVGY SGNSHTLPMGNRIALDKLRETIQRYAWHSLLFNVPPAPTSKRVR AISELRVAAGRER LFTVITFVQTNILSKLQKRYAANWTPKSQERLSRLREEGQHILSLLES GS WQKEVV REDQDLIVCSALTKPGLSIGAFCRPKYLKPAKHALVRLIFVEQWPGQIWGQSKRTR RMRRRKDVERVYDISVQAWALKGKETRISECIDTMRRHQAYIGVLPFLILSGSTVR GKGDCPILKEITRMRYCPNNEGLIPLGIFYRGSANKLLRVVKGSSFTLPMWQNIETLP HPEPFSPEGWTATGALYEKNLAYWSALNEAVDWYTGQILSSGLQYPNQNEFLARLQ NVIDSIPRKWFRPQGLKNLKPNGQEDIVPNEFVIPQNAIRAHHVIEWYHKTNDLVAK TLLGWGSQTTLNQTRPQGD LRFTYTRYF FREKEVPEV
SEQ ID NO: 73	VPKKKLMRELAKKAVFEAIFNDPIPGSFSGCKRCTLIDGARVTD AIEKKQGAKRCAGC EPCTFHTLYDSVKHALPAATGCDRTAIDTGLWEILTALRSYNWMSFRRNAVSDASQ KQVWSIEELAIWADKERALRVILSALHTHTIGKLKNGFSRDGVWKG GKL YENLAQK DLAKGLFANGEIFGKELVEADHDMLAWTIVPNHQFHIGLIRGNWKPA AVEASTAFD ARWLTNGAPLRDTRTHGHRGRFRNRTEKLTVLCIKRDGGVSEEFQERDYELSVML LQPKNKLKPEPKGELNSFEDLHDHWWFLKGDEATALVGLTSDPTVGD F IQLGLYIR NPIKAHGETKRRLLICFEPPIKLPLRRAFPSEAFKTWEPTINVFRNGRRDTEAYYDIDR ARVFEPETRVSL EHL SKQWEVLRLEPDRENTDPYEAQQNEGAELQVYSLQEEAAQ KMAPKVVIDPFGQFPLELFSTFVAQLFNAPLSDTKAKIGKPLDSGFV VESHLHLEED FAYRDFVRVTFMGTEPTFRVIHYSNGEGYWKKT V LKGKNNIRTAL IPEGAKAAVDA YKNKRCPLTLEAAILNEEKDRRLVLGNKALSLLAQTARGNL TILEALAAEVLRLPSG TEGVVHLHACVTRHSTLTESTETDNM
SEQ ID NO: 74	VEKLFSERLKRAMWLKNEAGRPPAETLTLKHKRVS GGHEK VKEELQRVLRSLSGT NQAAWNLGLSGGREPKSSDALKGEKSRVVLET VVFHSGHNRVLYDVIEREDQVHQ RSSIMHMRRKGSNLLRLWGRSGKVRRKMREEVAEIKPVWHKDSRWLAIVEEGRQS VVGISSAGLAVFAVQESQCTTAEPKPLEYVVS I WFRGSKALNPQDRYLEFKKLKTE ALRGQQYDPIPFSLKRGAGCSLAIRGEGIKFGSRGPIKQFFGSDRSRPSHADYDGKRR LSLFSKYAGDLADLTEEQWNRTVSAFAEDEVR RATLANIQDFLSISHEKYAERLKKR IESIEEPVSASKLEAYLSAIFETVQQREALASNFLMRLVESVALLISLEEKSPRVEFRV ARYLAESKEGFNRKAM
SEQ ID NO: 75	VVITQSELYKERLLRVMEIKNDRGRKEPRESQGLVLRFTQVTGGQEKVKQKLWLIFE GFSGTNQASWNFGQPAGGRKPNSGDALKGPKSRVTYETVVFHFGRLRLLSAVIERHN LKQQRQTMAYMKRRAAARKKWARSGKKCSMRNEVEKIKPKWHKDPRWFDIVK EGEPSIVGISSAGFAIYIVEEPNFRQDPLEIEY AISIWFRDRSQYLTFKKIQKAEKLL ELQYNPIPFRLKQEKTSLVFESGDIKFGSRGSIEHFRDEARGKPPKADMDNNRRLTMF SVFSGNLTNLTEEQYARPVSGLLAPDEKRMPTLLKKLQDFFTPIHEKYGERIKQRLA

SEQ ID NO	序列
	NSEASKRPFKKLEEYLPAYLEFRARREGLASNWVVLINSVRTLVRIKSEDOPYIEFKVSQYLLEKEDNKAL
SEQ ID NO: 76	KQDALFEERLKKAFIKRQADPLQREELSLPPNRKIVTGGHESAKDTLKQILRAINGTNQASWNPPTSGKRDSSADALAGPKSRVKLETVVHVGHRLLKKVVEYQGHQKQQHGLKAFMRTCAAMRKKWKRSGKVVGELREQLANIQPKWHYDSRPLNLCFEGKPSVVGLRSAGIALYTIQKSVVPVKEPKPIEYAVSIWFRGPKAMDREDRCLFEKKLKIA TELRKLQFEPVSTLTQGIKGFSLYIQGNSVKFGSRGPIKYFSNESVRQRPKADPDGNKRLALFSKFSGDLSDLTEEQWNRPIAFEGHRRATLGNIQDYLTVGHEQFAISLEQLLSEKESVLQMSIEQQRLKKNLGKKAENEWVESFGAEQARKKAQGIREYISGFFQEYCSQREQWAENWVQQLNKSURLFTIQDSTPFIEFRVARYLPKGEKKKGKAM
SEQ ID NO: 77	ANHAERHKRLRKEANRAANRNRPLVADCDTGDPLVGICRLLRGDKMQPNKTGCRSCEQVEPELRDAILVSGPGRLDNYKYELFQRGRAMAVHRLLRVPKLNRPKKAAGNDEKKAENKKSEIQKEKQKQRRMMPAVSMKQVSVADFKHVIENTVRHLFGDRDRDRIAECALRAASKYFLKSRRVRPRKLPKLANPDHGKELKGLRLREKRAKLKKEKEKQAEALARSNQKGAVLHVATLKKDAPPMPEYKTQGRNDYTTFVISAIAIKVGATRGTKPLLTPQPREWQCSLYWRDQQRWIRGGLLGLQAGIVLGPKNRELLEAVLQRPICRMSCGCGNPLQVRGAADVDFMTTNPFFYVSGAAYAQQKFKPFGTKRASEDGAAGAKAREKLMTQLAKVLDKVVTAHSPLDGIWETRPEAKLRAMIMALEHEWIFLRPGPCHNA AEEVIKCDCTGGHAILWALIDEARGALEHKEFYAVTRAHHDCEKQKLGGRLAGFLDLLIAQDVPLDDAPAARKIKTLLEATPPAPCYKAATSIATCDCEGKFDKLWAIDATRAGHGTEDLWARTLAYPQNVNCKCKAGKDLTHRLADFLGLLIKRDGPFPRERPPHKVTGDRKLVFSGDKCKKGHQYVILAKAHNEEVVRAWISRWGLKSRTNKAGYAATELNLLNLWLSICRRRWMDMLTVQRDTPYIRMTGRLVVDKKEKAM
SEQ ID NO: 78	AKQREALRVALERGIVRASNRITYTLVTNCTKGGPLPEQCRMIEGKARAMKWEPKLVGCGSCAAATVDLPAIEEYAQPGRLDVAKYKLTTQILAMATRRMMVRAAKLSRRKGQWPAKVQEEKEEPPPEPKMLKAVEMRPVAIVDFNRVIQTTIEHLWAERANADEAELKALKAAAAFYGPSLKIRARGPPKAAIGRELKKAHRKKAYAERKKARRKRAELARSQARGAAAHAAIRERDIPPMAYERTQGRNDVTTPIAAAIAAATRGARPLPAPKPMKWQCSLYWNEGQRWIRGGMLTAQAYAHAANIHRPMRCMVGWGNPLKVRAFERVADPDGAKGRKAERLQTNAFYVSGAAYRNKKFKPFGTDRGGIGSARKKRERLMAQLAKILDKVVSQAHSPLDDIWHTRPAQKLRAKQLEHEWMFLRPQAPTVEG TKPDVDVAGNMQRQIKALMAPDLPIEKGSPAKRFTGDKRKKGERAVRVAEAHSD EVVTAWISRWGIQTRRNEGSYAAQELELLLNWLQICRRRWLDMTAAQRVSPYIRMKSGRMITDAADEGVAPIPLVENM
SEQ ID NO: 79	KSISGRSIKHMACKDMLKSEITEIEEKQKKESLRKWDYYSKFSDEILFRNLNVSANHDANACYGCNPAFLKEVYGFRIERRNERIISYRRGLAGCKSCVQSTGYPIEFVRRKFGADKAMEIVREVLHRRNWGALARNIGREKEADPILGELNELLLVDARPYFGNKS AANETNLAFNVITRAAKKFRDEGMYDIHKQLDIHSEEGKVPKGRKSRLIRIERKHKA IHGLDPGETWRYPHCGKGEKYGVWLNRSRLIHIKNEYRCLTAFGTTGRRMSLDVACSVLGHPLVKKRKKKGKKTVDGTELWQIKKATETLPEDPIDCTFYLYAAKPTKDPFILKVGSLKAPRWKKLHKDFFEYSDTEKTQGQEKGRVVRGKVPRILSLRPDAKFKVSIWDDPYNGKNKEGTLLRMELSGLDGAKKPLILKRYGEPNTKPKNFVFWRPHITPHPLTFTPKHDFGDPNKKTKRRRVFNREYYGHLNDLAKMEPNKFFEDREVSNKKNPKAKNIRIQAKESLPNIVAKNGRWAAFDPNDSLWKLYLHWRGRRKTIKGGISQEFQEFKERLDLYKKHEDESEWKEKEKLWENHEKEWKKTLEIHGSIAEVSQRCVMQSMGMP LDGLVQKKDYVHIGQSSLKAADDWTFSANRYKKATGPKWGKISVSNLLYDANQANAELISQSISKYLSKQKDNQGCEGRMKFLIKIIEPLRENFVKHTRWLHEMTQKDC EVRAQFSRVSM
SEQ ID NO: 80	FPSDVGADALKHVRMLQPRLTDEVKVALTRAPSDRPALARFAAVAQDGLAFVRHLNVSANHDSNCTFPRDPRDPRRGPCPNPAFLREVWGFRIARGNERALSYYRGLAGCKSCVQSTGFPSVPFHRIGADDCMRKLHEILKARNWRLLARNIGREREADPLLTE LSEYLLVDARTYPDGAAPNSGRLAENVIKRAAKKFRDEGMRDIHAQLRVHSREGKV

SEQ ID NO	序列
	PKGRLQRLRRIERKHRAIHALDPGPSWEAEGSARAQVQGVAVYRSQLLRVGHHTQQ IEPVGIVARTLFGVGRDLDVAVSVLGAPLTKRKKGSKTLESTEDFRIAKARETRA DKIEVAFVLYPTASLLRDEIPKDAFPAMRIDRFLKVGSVQADREILLQDDYYRFGD AEVKAGKNKGRTVTRPVKVPRLQALRPDAKFRVNVWADPFGAGDSPGTLLRLEVS GVTRRSQPLRLLRYGQPSTQPANFLCWRPHRVPDPMFTTPRQKFGERRKNRRTRRP RVFERLYQVHIKHLAHLERNKWFEEARVSAQKWAKARAIIRKGAEDIPVVPAPAK RRWAALQPNALWDLAHDREARKRFRGGRAAEGEEFKPRLNLYLAHEPEAEWES KRDRWERYEKKWTAVLEEHSRMCAVADRTLPOFLSDPLGARMDDKDYAFVGKSA LAVAEAFVEEGTVERAQGNCSITAKKKFASNASKRLSVANLLDVSDKADRALVFQ AVRQYVQRQAENGVEGRMAFLRKLAPLRQNFVCHTRWLHM
SEQ ID NO: 81	AARKKKRGKIGITVKAKEKSPPAAGPFMARKLVNVAANVDGVEVHLCVECEADAH GSASARLLGGCRSCTGSIGAEGRLMGSDVDVRERVIAEPVHTETERLGPDVKAFAEAG TAESKYAIQRGLEYWGVDLISNRNARTVRKMEEADRPESSTMEKTSWDEIAIKTYSQ AYHASENHLFWERQRRVRQHALALFRRARERNRGESPLQSTQRPAPLVLAALHAEA AAISGRARAQVYVLRGPSANVRAAAADIDAKPLGHYKTPSPKVARGFPVKRDLLRAR HRIVGLSRAYFKPSDVVRGTSDAIAHVAGRNIGVAGGKPKIEKTFTLFPVAYWEDV DRVVHCSSFADGPWVRDQRIKIRGVSSAVGTFSLYGLDVAWSKPTSFYIRCSDIRK KFHPKGFPGMKHWRQWAKELDRLEQRASCVVRLQDDEELLQTMERGQRYDDV FSCAATHATRGEADPSGGCSRCELVSCGVAHKVTKKAKGDTGIEAVAVAGCSLCES KLVGSPKPRVHRQMAALRQSHALNYLRRLQREWEALEAVQAPTPLYRKFYARHLE VRSM
SEQ ID NO: 82	AAKKKKQRGKIGISVKPKEGSAPPADGPFMARKLVNVAANVDGVEVNLCECEAD AHGSAPARLLGGCKSCTGSIGAEGRLMGSDVDVRADAIAPVNTETETKLGPDVQAF EAGTAETKYALQRGLEYWGVDLISNRNRSRTVRRTEEGQPESATMEKTSWDEIAIKS YTRAYHASENHLFWERQRRVRQHALALFKRAKERNRGDSTLPREPGHGLVAIAAL ACEAYAVGGRNLAETVVRGPTFGTARAVRDVEIASLGRYKTPSPKVAHGSPVKRDF LRARHRIVGLARAYYRPSDVVRGTSDAIAHVAGRNIGVAGGKPRAVEAVFTLFPVA YWEDVDRVVHCSSFQVSAPWNRDQRMKIAGVTTAAGTFSLHGGLKWKAPTSFYI RCSLTRRKFRPKGFPGMKRWRQWAKDLDRLEQRASCVVRLQDDAALLETMER GQRYDDVFACAVTHATRGEADRLAGCSRCLTPCQEAHRVTTKPRGDAGVEQVQT SDCSLCEGKLVGSPKPRHLRTLTLRQEHGLNYLRRLQREWESLEAVQVPTPLYRKF YARHLEVRSM
SEQ ID NO: 83	TDSQSESVPEVVYALTGGEVPGRVPPDGGSAGARNAPTGLRKQRGKIKISAKPSKP GSPASSLARTLVNEAANVDGVQSSGCATCRMRRANGSAPRALPIGCVACASSIGRAP QEETVCALPTTQGPVRLLEGGHALRKYDIQRALEYWGVDLIGNRLDRQAGRGME PAEGATATMKRVSMDELAVLDFGKSYASEQHLFAARQRRVRQHAKALKIRAKHA NRSGSVKRALDRSRKQVTALAREFFKPSDVVRGSDALAHVVGRNLGVSRHPAREI PQTFTLPLCAYWEDVDRVISCSLLAGEPFARDQEIRIEGVSSALGSLRLYRGAIEWH KPTSLYIRCSLTRRKFRPRGGLKKRWRQWAKDLDRLEQRACCIVRSLQADVLLQ TMERAQRFYDVHDCATHVGPVAVRCSPCAGKQFDWDRYRLLAALRQEHALNYL RRLQREWESLEAQVVKMPYLRKFYARKLEVSGPLIGLEVRREPSMGTIAIEM
SEQ ID NO: 84	AGTAGRRHGSGLGARRSINIAGVTDRHGRWGCESCVYTRDQAGNRARCAPCDQSTY APDVQEVITIGQRQAKYTIFLTLSQFSWTNMTMRNNKRAAAGRSKRTTGKRIGQLAEIK ITGVGLAHAHNVQRSLQHNITKMWRNAEKGSKRVARLKKAKQLTKRRAYFRMRM SRQSRGNGFFRTGKGGIHAVAPVKIGLDVGMASGSSEPADEQVTLDIAIWKGRKK KIRLIGAKGELAVAACRFREQQTKGDKCIPLILQDGEVRWNQNNWQCHPKKLVLPLC GLEVSRKFVSQADRLAQNKVASPLAARFDKTSVKGTLVESDFAAVLVNVTISIYQQC HAMLRSQEPTPSLRVQRTITSM
SEQ ID NO: 85	GVRFSPAQSQVFFRTVIPQSVEARFAINMAAIIHDAAGAFGCSVCRFEDRTPRNAKAV HGCSPCTRSTNRPDVFVLPVGAIAKAKYDVFMRLLGFNWTNLNRRQAKRVTVRDRIG QLDELAISMLTGKAKAVLKKSICHNVDSFKAMRGSLLKLRKASKTGKSQLRAK LSDLRERTNTTQEGSHVEGSDVALNKIGLDVGLVGKPDYPSEESVEVVVCLYFVG

SEQ ID NO	序列
	KVLILDAQGRIRDMRAKQYDGFKIPIIQRGQLTVLSVKDLGKWSLVRQDYVLAGDL RFEPKISKDRKYAECVKRIALITLQASLGFKERIPYYVTKQVEIKNASHIAFVTEAIQN CAENFREMTEYLMKYQEKSPDLKVLTLQML
SEQ ID NO: 86	RAVVGVFLEQARRALNLTNFGTNHRTGCNGCYVTPGKLSIPQDGEKNAAGCTS CLMKATASYVSYPKPLGEKVAKYSTLDALKGFPWYSLRLNLRPNYRGKPINGVQEV APVSKFRLAEEVIQAVQRYHFTELEQSFPGRRRLRELAFYTKEYRRAPEQRQHV NGDRNIVVTVLHELGFVGMFNEVELLPKTPIECAVNVFIRGNRVLLEVRKPQFDK ERLLVESLWKKDSRRHTAKWTPPNNEGRIFTAEGWKDFQLPLLLGSTSRSLRAIEKE GFVQLAPGRDPDYNNTIDEQHSRPFLLPLYLYLQGTISQYCVFAGTWVIPFQDGISP YSTKDTFQPDLLKRAYSLLLDVVKHRLGNKVASGLQYGRFPAIEELKRLVRMHGAT RKIPRGEKDLLKKGDPDTPPEWWLLEQYPEFWRLCDAAAKRVSQNVGLLLSLKKQP LWQRRWLESRTNEPLDNLPLSMALTLHLTNEEAL
SEQ ID NO: 87	AAVYSKFYIENHFKMGIPETLSRIRGPSIIQGFSVNENYINIAGVGDRDFIFGCKKCKY TRGKPSKKINKCHPCKRSTYPEVIDVRGSISEFKYKIYNKLKQEPNQSICKQNTKGR MNPSDHTSSNDGIIINGIDNRIAYNVIFSSYKHLMEKQINLLRDTTKRKARQIKKYNN SGKKKHSLRSQTKGNLKNRYHMLGMFKKGSLLITNEGDFITAVRKVGLDISLYKNE SLNKQEVETELCLNIKWGRTKSYTVSGYIPLPINIDWKL YLFKETGLTLRLFGNKYK IQSKKFLIAQLFKPKRPPCADPVVKAQKWSALNAHVQQMAGLFSDSLKRELKN RMHKQLDFKSLWVGTEYIKWFEELSRSYVEGAESLEFFRQDYFCFNYTKQTTM
SEQ ID NO: 88	PQQQRDLMLMAANYDQDYGNCGCPCTVVASAAYRPDPQAQHGCKRHLRTLGA VTHVGLGDRATITLHRLRGAALAAARAAAQAASAPMTPDTPDAPDDRRRLEAID ADDVVLGAHRLWSAVRRWADRRRAALRRRLHSEREWLLKDQIRWAELYTLIE ASGTPPQGRWRNTLGALRGQSRWRRLAPTMRTATCAETHAELWDALAEVPEMA KDRRGLLRPPVEADALWRAPMIVEGWRGGHVVVDAVAPPLDLPQPCAWTAVRLS GDPRQRWGLHLAVPPLGQVQPPDLKATLAVSMRHRGGVRVRTLQAMAVDADAP MQRHLQVPLTLQRGGGLQWGIHSRGVRRREARSMASWEGPPIWTGLQLVNRWKG QGSALLAPDRPPDTPPYAPDAVAPAQPDTKRARRTLKEACTVCRCAPGHMRQLQ VTLTGDTWRRLRAPQGAQRKAELVKVATQHDERIANYTAWYLRPEHAAGC DTCGDGSRDLGACRGCRPLLVDGQCFRRYLDKIEADRDGGLAQKPKAQEAVAAM AAKRDARAQKVAARAALSEATGQRTAATRDASHEARAQKELEAVATEGTTVRH DAAAVSAFGSWVARKGDEYRHQVGVLANRLEHGLRLQELMAPDSVVADQQRASG HARVGYRYVLTAM
SEQ ID NO: 89	AVAHPVGRGNAGSPGARGPEELPRQLVNRSNVTRPATYGCAPCRHVRLSIPKPV TGCRACEQTHPAPKRAVRGGADAAKYDLAFAFFAGWAADLEGRNRRRQVHAPLD PQDPNHEPAVTLQKIDLAEVSIEEFQVRLARSVKHRHDGRASREREKARAYAQVA KKRRNSHAHGARTTRAVRRQTRAVRRRAHRMGANSGEILVASGAEDPVPEAIDHAA QLRRRIRACARDLEGLRHLSRRYLKTELEKPCRRRPAPDLGRARCHALVESLQAAERE LEELRRCDSPDTAMRRDLAVLAAAATDADFATGWTVVGMDLGVAPRGSAAPES PMEMASVFWRKGSRRVIVSKPIAGMPIRRHELIRLEGLGLRLDGNHYTGAGVTKG RGLSEGTEPDFREKSPSTLGLFTLSDYRHESRWRPYGAKQGKTARQFFAAMSRELRA LVEHQVLAPMGPPLEAHERRFETLLKGQDNKSIHAGGGGRYVWRGPPDSKKRPA ADGDWFRFGRGHADHRGWANKRHELAANYLQSAFRLWSTLAEAEQPTPYARYKY TRVTM
SEQ ID NO: 90	WDFLTLQVYERHTSPEVCVAGNSTKCSGTRKSDHTHGVGVKLGAQEINVSANDD RDHEVGCNICVISRVSLDIKWRYGCESCQSTPEWRSIVRFDRNHKEAKGECLSRF EYWGAQSIARSLKRNLKMGVNLDELAIVQNENVVKTSKHLFDKRKDRIQANLK AVKVRMRERRKSGRQRKALRRQCRKLKRYLRSYDPSDIKEGNSCSAFTKLGLDIGIS PNKPPKIEPKVEVVFSLFYQGACDKIVTVSSPESPLPRSWKIKIDGIRALYVKSTKV GGRTFRAGQRNNRRKVRPPNVKKGKRKGSRSQFFNKFAVGLDAVSQQLPIASVQGL WGRAETKKAQTICLKQLESNKPLKESQRCLFLADNWVVRVCGFLRALSQRQGP IRYRYRCNM
SEQ	ARNVGQRNASRQSKRESAKARSRRVTGGHASVTQGVALINAAANADRDHTTGCEP

SEQ ID NO	序列
ID NO: 91	CTWERNLPLQEVIHGCDSCTKSSPFWRDIKVVNKGYREAKEEIMRIASGISADHLS RALSHNKVMGRLNLDEVCLDFRTVLDTSKHLTDSRSNGIKEHIRAVHRKIRMRRK SGKTARALRKQYFALRRQWKAGHKPNSIREGNSLTALRAVGFDVGVSEGTEPMPAP QTEVVLSVFYKGSATRILRISSPHPIAKRSWKVKIAGIKALKLIRREHDFSFGRETYNA SQRAEKRFSPHAARKDFFNSFAVQLDRLAQQLCVSSVENLWVTEPQQKLLTLAKD TAPYGIREGARFADTRARLAWNWWFRVCGFTRALHQEQEPTPYCRFTWRSKM

[0042] 在一些实施方案中,可编程核酸酶可以是可编程切口酶(例如,Cas14a、Cas14b或Cas14e)。本公开内容提供了能够在双链DNA(dsDNA)的单链中引入断裂(“切口”)的可编程切口酶的组合物。在一些实施方案中,可编程切口酶是可编程DNA切口酶。所述可编程切口酶包括与靶向dsDNA中特定目的区域的指导核酸偶联的切口酶。在一些实施方案中,两种可编程切口酶被组合并一起递送以产生两个链断裂。例如,第一可编程切口酶可以靶向并切口dsDNA的第一区域,并且第二可编程切口酶可以靶向并切口相反链上相同dsDNA的第二区域。当组合并一起递送以在dsDNA的相反链上产生切口时,可以在dsDNA中产生两条链断裂。链断裂可以通过非同源末端连接(NHEJ)或同源定向修复(HDR)进行修复和重新连接。因此,本文公开的两种可编程切口酶可以组合以选择性编辑核酸序列。这可用于任何用于治疗疾病或病症的治疗应用或农业应用的基因组编辑。

[0043] 在一些实施方案中,可编程切口酶可以是能够在dsDNA的单链上产生切口的Cas蛋白。与本文公开的可编程切口酶一致的Cas蛋白包括Cas14,在本文中也称为CasZ。在特定实施方案中,与本文公开的可编程切口酶一致的Cas蛋白包括Cas14e,在本文中也称为CasZe。本文公开的Cas14e可编程切口酶可用于基因组编辑目的以产生链断裂以切除DNA区域或随后引入DNA区域。

[0044] 使靶核酸产生切口的方法可以包括使靶核酸与第一指导核酸接触(例如,包含与长度不超过900个氨基酸的第一可编程切口酶结合的第一区域的指导核酸)和第二指导核酸(例如,包含与长度不超过900个氨基酸的第二可编程切口酶结合的第一区域的指导核酸)。第一指导核酸可以包含与靶核酸结合的第二区域,并且第二指导核酸可以包含与靶核酸结合的第二区域。第一指导核酸的第二区域和第二指导核酸的第二区域可以与靶核酸的相反链结合。

[0045] 在一些实施方案中,本文公开的可编程切口酶(例如,Cas14e)可用于DNA核酸内切酶靶向CRISPR TransReporter(DETECTR)测定。DETECTR分析利用一些可编程核酸酶和可编程切口酶(例如,CRISPR-Cas效应蛋白)的反式切割能力来实现对样品中靶DNA的快速、高保真检测。从生物样品中提取DNA后,与靶DNA序列互补的crRNA可以与靶DNA结合,通过可编程核酸酶或可编程切口酶(例如,可编程切口酶,例如Cas14e)启动无差别的ssDNase活性。在一些实施方案中,在使提取的DNA接触与指导RNA复合的可编程切口酶之前,通过PCR或等温扩增反应扩增该提取的DNA。与靶DNA杂交后,可编程切口酶的反式切割活性被激活,然后可以切割ssDNA荧光猝灭(FQ)报告分子。报告分子的切割可以提供荧光读数,表明样品中存在靶DNA。在一些实施方案中,在DETECTR测定中本文公开的可编程切口酶(例如,Cas14e)可以与其他可编程核酸酶或其他可编程切口酶组合或复用。

[0046] 本公开内容的可编程切口酶可以在存在靶DNA的某些条件下显示出活性增强,如通过ssDNA-FQ报告子的切割增强所测量。例如,本公开内容的可编程切口酶可以具有基于

缓冲剂制剂、pH水平、温度或盐的可变水平的活性。与本公开内容一致的缓冲液包括磷酸盐缓冲液、Tris缓冲液和HEPES缓冲液。本公开内容的可编程切口酶(例如,Cas14e)可以在磷酸盐缓冲液、Tris缓冲液和HEPES缓冲液中显示最佳活性。可编程切口酶还可以在不同pH水平下表现出不同水平的切割。例如,在pH 7至pH 9可以观察到增强的切割。在一些实施方案中,本公开内容的可编程切口酶在约pH 7、约pH 7.1、约pH 7.2、约pH 7.3、约pH 7.4、约pH 7.5、约pH 7.6、约pH 7.7、约pH 7.8、约pH 7.9、约pH 8、约pH 8.1、约pH 8.2、约pH 8.3、约pH 8.4、约pH 8.5、约pH 8.6、约pH 8.7、约pH 8.8、约pH 8.9、约pH 9、pH 7至7.5、pH 7.5至8、pH 8至8.5、pH 8.5至9、或pH 7至8.5表现处增强的切割。

[0047] 在一些实施方案中,本公开内容的可编程切口酶(例如,Cas14e)在存在靶DNA的情况下在25℃至50℃的温度下表现出ssDNA-FQ报告子的增强的切割。例如,本公开内容的可编程切口酶(例如,Cas14e)可以在约25℃、约26℃、约27℃、约28℃、约29℃、约30℃、约31℃、约32℃、约33℃、约34℃、约35℃、约36℃、约37℃、约38℃、约39℃、约40℃、约41℃、约42℃、约43℃、约44℃、约45℃、约46℃、约47℃、约48℃、约49℃、约50℃、30℃至40℃、35℃至45℃或35℃至40℃表现出ssDNA-FQ报告子的增强的切割。

[0048] 本公开内容的可编程切口酶(例如,Cas14e)可以对存在靶DNA的样品中的盐浓度不敏感。有利地,所述可编程切口酶在25nM盐至200mM盐的不同盐浓度下可以具有活性并且能够切割ssDNA-FQ-报告子序列。各种盐与本文公开的可编程切口酶的这种特性一致,包括NaCl或KCl。本公开内容的可编程切口酶可以在25nM至500nM盐、500nM至1000nM盐、1000nM至2000nM盐、2000nM至3000nM盐、3000nM至4000nM盐、4000nM至5000nM盐、5000nM至6000nM盐、6000nM至7000nM盐、7000nM至8000nM盐、8000nM至9000nM盐、9000nM至0.01mM盐、0.01mM至0.05mM盐、0.05mM至0.1mM盐、0.1mM至10mM盐、10mM至100mM盐、或100mM至500mM盐的盐浓度下具有活性。因此,本公开内容的可编程切口酶(例如,Cas14e)可以表现出样品中盐浓度非依赖性切割活性。

[0049] 本公开内容的可编程切口酶(例如,Cas14e)能够切割任何ssDNA-FQ报告子,无论其序列如何。因此,本文提供的可编程切口酶可以能够切割通用ssDNA FQ报告子。在一些实施方案中,本文提供的可编程切口酶切割包含5至20个腺嘌呤、5至20个胸腺嘧啶、5至20个胞嘧啶、或5至20个鸟嘌呤的均聚物ssDNA-FQ报告子。因此,本公开内容的可编程切口酶能够切割也被可编程核酸酶切割的ssDNA-FQ报告子,如本文别处所公开的,实现在具有单一ssDNA-FQ报告子的单一测定中容易地复用多种可编程切口酶和可编程核酸酶。

[0050] 本公开内容的可编程切口酶(例如,Cas14e)可以与靶DNA中的野生型前间区序列邻近基序(PAM)或突变PAM结合。在一些实施方案中,本公开内容的可编程切口酶对PAM不敏感并且无论靶DNA中的PAM序列如何都可以与靶DNA结合。在一些实施方案中,本公开内容的可编程切口酶不依赖于PAM并且无论靶DNA中是否存在PAM序列都可以与靶DNA结合。

[0051] 在一些实施方案中,本公开内容的可编程切口酶(例如,Cas14)由于其小尺寸而特别适用于基因组编辑和在DETECTR测定中使用。这些蛋白的较小性质使它们在基因组编辑的背景下更容易被包装并以更高的效率被递送,并且在测定中更容易作为试剂掺入。在一些实施方案中,本公开内容的可编程切口酶的长度为400至800个氨基酸、长度为400至420个氨基酸、长度为420至440个氨基酸、长度为440至460个氨基酸、长度为460至480个氨基酸、长度为480至500个氨基酸、长度为500至520个氨基酸、长度为520至540个氨基酸、长度

为540至560个氨基酸、长度为560至580个氨基酸、长度为580至600个氨基酸、长度为600至620个氨基酸、长度为620至640个氨基酸、长度为640至660个氨基酸、长度为660至680个氨基酸、长度为680至700个氨基酸、长度为700至720个氨基酸、长度为720至740个氨基酸、长度为740至760个氨基酸、长度为760至780个氨基酸、长度为780至800个氨基酸、长度为400至500个氨基酸、长度为500至600个氨基酸、长度为600至700个氨基酸、长度为700至800个氨基酸、长度为450至550个氨基酸、长度为550至650个氨基酸、长度为650至750个氨基酸、或长度为750至800个氨基酸。在一些实施方案中,本公开内容的可编程切口酶的长度为350至900个氨基酸残基。在一些实施方案中,本公开内容的可编程切口酶的长度为500至550个氨基酸残基。在优选的实施方案中,本公开内容的可编程Cas14切口酶的长度为480至550个氨基酸。

[0052] 可编程切口酶(例如,Cas14a、Cas14b或Cas14e可编程切口酶)和其他试剂(例如,指导核酸)可以在本文公开的缓冲液中配制。多种缓冲溶液与本文公开的方法、组合物、试剂、酶和试剂盒兼容。缓冲液与本文所述的不同可编程切口酶兼容。本文公开的任何方法、组合物、试剂、酶或试剂盒可以包含缓冲液。这些缓冲液可以与本文所述的其他试剂、样品和支持介质兼容,用于检测病痛,例如疾病、癌症或遗传病症或遗传信息,例如用于表型分析、基因分型或确定血统。如本文所述,缓冲液可以提高本文所述的任何可编程切口酶的顺式或反式切割速率。缓冲液可以增加可编程切口酶对靶核酸的辨别力。如本文所述的方法可以在缓冲液中进行。

[0053] 在一些实施方案中,缓冲液可以包含缓冲剂、盐、拥挤剂或去污剂中的一种或多种或其任何组合。缓冲液可以包含还原剂。缓冲液可以包括竞争剂。示例性缓冲剂包括HEPES、TRIS、MES、ADA、PIPES、ACES、MOPSO、BIS-TRIS丙烷、BES、MOPS、TES、DISO、Trizma、TRICINE、GLY-GLY、HEPPS、BICINE、TAPS、A MPD、A MPDSO、CHES、CAPSO、AMP、CAPS、磷酸盐、柠檬酸盐、乙酸盐、咪唑或其任何组合。缓冲剂可以与可编程切口酶兼容。与可编程切口酶兼容的缓冲液可以包含浓度为1mM至200mM的缓冲剂。与可编程切口酶兼容的缓冲液可以包含浓度为10mM至30mM的缓冲剂。与可编程切口酶兼容的缓冲液可以包含浓度为约20mM的缓冲剂。组合物(例如,包含可编程切口酶的组合物)可以具有2.5至3.5的pH值。组合物(例如,包含可编程切口酶的组合物)可以具有3至4的pH值。组合物(例如,包含可编程切口酶的组合物)可以具有3.5至4.5的pH值。组合物(例如,包含可编程切口酶的组合物)可以具有4至5的pH值。组合物(例如,包含可编程切口酶的组合物)可以具有4.5至5.5的pH值。组合物(例如,包含可编程切口酶的组合物)可以具有5至6的pH值。组合物(例如,包含可编程切口酶的组合物)可以具有5.5至6.5的pH值。组合物(例如,包含可编程切口酶的组合物)可以具有6至7的pH值。组合物(例如,包含可编程切口酶的组合物)可以具有6.5至7.5的pH值。组合物(例如,包含可编程切口酶的组合物)可以具有7至8的pH值。组合物(例如,包含可编程切口酶的组合物)可以具有7.5至8.5的pH值。组合物(例如,包含可编程切口酶的组合物)可以具有8至9的pH值。组合物(例如,包含可编程切口酶的组合物)可以具有8.5至9.5的pH值。组合物(例如,包含可编程切口酶的组合物)可以具有9至10的pH值。组合物(例如,包含可编程切口酶的组合物)可以具有9.5至10.5的pH值。

[0054] 缓冲液可以包含盐。示例性盐包括NaCl、KCl、乙酸镁、乙酸钾、CaCl₂和MgCl₂。缓冲液可以包含乙酸钾、乙酸镁、氯化钠、氯化镁或其任何组合。与可编程切口酶兼容的缓冲液

可以包含浓度为5mM至100mM的盐。与可编程切口酶兼容的缓冲液可以包含浓度为5mM至10mM的盐。在一些实施方案中,与可编程切口酶兼容的缓冲液包含1mM至60mM的盐。在一些实施方案中,与可编程切口酶兼容的缓冲液包含1mM至10mM的盐。在一些实施方案中,与可编程切口酶兼容的缓冲液包含约105mM的盐。在一些实施方案中,与可编程切口酶兼容的缓冲液包含约55mM的盐。在一些实施方案中,与可编程切口酶兼容的缓冲液包含约7mM的盐。在一些实施方案中,与可编程切口酶兼容的缓冲液包含盐,其中该盐包含乙酸钾和乙酸镁。在一些实施方案中,与可编程切口酶兼容的缓冲液包含盐,其中该盐包含氯化钠和氯化镁。在一些实施方案中,与可编程切口酶兼容的缓冲液包含盐,其中该盐包含氯化钾和氯化镁。

[0055] 缓冲液可以包含拥挤剂。示例性拥挤剂包括甘油和牛血清白蛋白。缓冲液可以包含甘油。拥挤剂可以减少溶液中其他分子可用的溶剂体积,从而增加所述分子的有效浓度。与可编程切口酶兼容的缓冲液可以包含浓度为0.01% (v/v) 至10% (v/v) 的拥挤剂。与可编程切口酶兼容的缓冲液可以包含浓度为0.5% (v/v) 至10% (v/v) 的拥挤剂。

[0056] 缓冲液可以包含去污剂。示例性去污剂包括吐温、Triton-X和IGEPAL。缓冲液可以包括吐温、Triton-X或其任何组合。与可编程切口酶兼容的缓冲液可以包含Triton-X。与可编程切口酶兼容的缓冲液可以包括IGEPAL CA-630。在一些实施方案中,与可编程切口酶兼容的缓冲液包含浓度为2% (v/v) 或更低去污剂。与可编程切口酶兼容的缓冲液可以包含浓度为2% (v/v) 或更低去污剂。与可编程切口酶兼容的缓冲液可以包含浓度为0.00001% (v/v) 至0.01% (v/v) 的去污剂。与可编程切口酶兼容的缓冲液可以包含浓度为约0.01% (v/v) 的去污剂。

[0057] 缓冲液可以包含还原剂。示例性还原剂包含二硫苏糖醇 (DTT)、 β -巯基乙醇 (BME) 或三(2-羧乙基) 膦 (TCEP)。与可编程切口酶兼容的缓冲液可以包含DTT。与可编程切口酶兼容的缓冲液可以包含浓度为0.01mM至100mM的还原剂。与可编程切口酶兼容的缓冲液可以包含浓度为0.1mM至10mM的还原剂。与可编程切口酶兼容的缓冲液可以包含浓度为0.5mM至2mM的还原剂。与可编程切口酶兼容的缓冲液可以包含浓度为0.01mM至100mM的还原剂。与可编程切口酶兼容的缓冲液可以包含浓度为0.1mM至10mM的还原剂。与可编程切口酶兼容的缓冲液可以包含浓度为约1mM的还原剂。

[0058] 与可编程切口酶兼容的缓冲液可以包含竞争剂。示例性竞争剂与靶核酸或报告核酸竞争可编程切口酶的切割。示例性竞争剂包括肝素、咪唑和鲑鱼精子DNA。与可编程切口酶兼容的缓冲液可以包含浓度为1 μ g/mL至100 μ g/mL的竞争剂。与可编程切口酶兼容的缓冲液可以包含浓度为40 μ g/mL至60 μ g/mL的竞争剂。

指导核酸

[0059] 本公开内容的试剂可以包含指导核酸。如本文所述,指导核酸可以与单链靶核酸或其部分结合。例如,如本文所述,指导核酸可以与靶核酸,例如来自病毒或细菌或者引起疾病的其他药剂的核酸或其扩增子结合。如本文所述,指导核酸可以与靶核酸,例如来自细菌、病毒、寄生虫、原生动物、真菌或引起疾病的其他药剂的核酸或其扩增子结合,并且进一步包含可以赋予对治疗(例如抗生素治疗)的抗性的突变,例如单核苷酸多态性(SNP)。如本文所述,指导核酸可以与靶核酸,例如来自癌症基因或与遗传疾病相关的基因的核酸,优选为DNA或其扩增子结合。指导核酸包含与靶核酸反向互补的核酸片段。通常,指导核酸与靶核酸特异性结合。靶核酸可以是逆转录的RNA、DNA、DNA扩增子或合成核酸。靶核酸可以是目

的核酸的单链DNA或DNA扩增子。指导核酸可以是非天然存在的指导核酸。非天然存在的指导核酸可以包含具有与目的靶核酸序列杂交的重复序列和间隔区的工程化序列。非天然存在的指导核酸可以被重组表达或化学合成。

[0060] 指导核酸(gRNA)序列可以与靶核酸的靶序列杂交。在一些实施方案中,gRNA是gRNA系统(例如,包含crRNA和tracrRNA或crRNA和trancRNA)。crRNA可以包含与tracrRNA的区域杂交的重复区。tracrRNA可以与可编程核酸酶(例如,本公开内容的可编程切口酶)结合。在一些实施方案中,重复区可以包含相对于pre-crRNA中的重复序列的突变或截短。crRNA的重复序列可以与tracrRNA相互作用,tracrRNA可以与可编程核酸酶(例如,可编程切口酶)相互作用,从而实现crRNA、tracrRNA和可编程核酸酶形成复合体。该复合体可以称为核蛋白。crRNA可以包含间隔区序列。间隔区序列可以与靶核酸的靶序列杂交,其中靶序列是靶核酸的区段。间隔区序列可以与靶序列反向互补。在一些情况下,间隔区序列可以与靶序列充分反向互补,然而,不一定是100%反向互补,以实现杂交。在一些实施方案中,可编程核酸酶(例如,可编程切口酶)可以切割前体RNA(“pre-crRNA”)以产生gRNA,也称为“成熟指导RNA”。切割pre-crRNA以产生成熟指导RNA的可编程核酸酶(例如,可编程切口酶)被称为具有pre-crRNA加工活性。

[0061] 指导核酸(例如,gRNA系统的crRNA)可以包含至少部分地与靶核酸的序列反向互补的序列。指导核酸可以是非天然存在的指导核酸。非天然存在的指导核酸可以包含具有与目的靶核酸序列杂交的重复序列和间隔区的工程化序列。非天然存在的指导核酸可以重组表达或化学合成。指导核酸可以包含crRNA和tracrRNA或crRNA和trancRNA。有时,指导核酸包含crRNA和tracrRNA。指导核酸可以与靶核酸特异性结合。具体地,指导核酸的crRNA可以包含重复区和间隔区。重复区与tracrRNA的序列杂交,并且间隔区与靶核酸中的靶序列杂交。

[0062] 在一些实施方案中,tracrRNA序列与SEQ ID NO:98-SEQ ID NO:99、SEQ ID NO:101、SEQ ID NO:103、SEQ ID NO:105-SEQ ID NO:151中的任一项包含至少70%、至少80%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。在一些实施方案中,tracrRNA序列与SEQ ID NO:99包含至少70%序列同一性。在一些实施方案中,tracrRNA序列与SEQ ID NO:99包含至少75%序列同一性。在一些实施方案中,tracrRNA序列与SEQ ID NO:99包含至少80%序列同一性。在一些实施方案中,tracrRNA序列与SEQ ID NO:99包含至少85%序列同一性。在一些实施方案中,tracrRNA序列与SEQ ID NO:99包含至少90%序列同一性。在一些实施方案中,tracrRNA序列与SEQ ID NO:99包含至少95%序列同一性。在一些实施方案中,与本公开内容的可编程切口酶复合使用的tracr序列包含SEQ ID NO:99。在一些实施方案中,tracrRNA序列与SEQ ID NO:101包含至少70%序列同一性。在一些实施方案中,tracrRNA序列与SEQ ID NO:101包含至少75%序列同一性。在一些实施方案中,tracrRNA序列与SEQ ID NO:101包含至少80%序列同一性。在一些实施方案中,tracrRNA序列与SEQ ID NO:101包含至少85%序列同一性。在一些实施方案中,tracrRNA序列与SEQ ID NO:101包含至少90%序列同一性。在一些实施方案中,tracr序列与SEQ ID NO:101包含至少95%序列同一性。在一些实施方案中,与本公开内容的可编程切口酶复合使用的tracr序列包含SEQ ID NO:101。在一些实施方案中,tracrRNA序列与SEQ ID NO:103包含至少70%序列同一性。在一些实施方案中,tracrRNA序列与SEQ ID NO:103包含至少75%序列同一

性。在一些实施方案中, tracrRNA序列与SEQ ID NO:103包含至少80%序列同一性。在一些实施方案中, tracrRNA序列与SEQ ID NO:103包含至少85%序列同一性。在一些实施方案中, tracrRNA序列与SEQ ID NO:103包含至少90%序列同一性。在一些实施方案中, tracrRNA序列与SEQ ID NO:103包含至少95%序列同一性。在一些实施方案中, 与本公开内容的可编程切口酶复合使用的tracrRNA序列包含SEQ ID NO:103。在一些实施方案中, tracrRNA序列与SEQ ID NO:119包含至少70%序列同一性。在一些实施方案中, tracrRNA序列与SEQ ID NO:119包含至少75%序列同一性。在一些实施方案中, tracrRNA序列与SEQ ID NO:119包含至少80%序列同一性。在一些实施方案中, tracrRNA序列与SEQ ID NO:119包含至少85%序列同一性。在一些实施方案中, tracrRNA序列与SEQ ID NO:119包含至少90%序列同一性。在一些实施方案中, tracrRNA序列与SEQ ID NO:119包含至少95%序列同一性。在一些实施方案中, 与本公开内容的可编程切口酶复合使用的tracrRNA序列包含SEQ ID NO:119。

[0063] 在一些实施方案中, 可编程切口酶包含与SEQ ID NO:10具有70%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:10具有75%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:10具有80%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:10具有85%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:10具有90%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:10具有95%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:11具有70%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:11具有75%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:11具有80%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:11具有85%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:11具有90%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:11具有95%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:33具有70%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:33具有75%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:33具有80%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:33具有85%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:33具有90%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:33具有95%同一性的序列。

[0064] 在一些情况下, 指导核酸不是天然存在的, 并且是通过原本分开的序列区段的人工组合而制得。通常, 人工组合是通过化学合成、基因工程化技术或分离的核酸区段的人工操作来进行的。在一些情况下, 包含与靶核酸反向互补的序列的指导核酸的区段的长度为20个核苷酸。指导核酸可以具有至少10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29或30个与靶核酸反向互补的核苷酸。在一些情况下, 指导核酸的长度可以是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。例如, 指导核酸可以是至少10个碱基。在一些实施方案中, 指导核酸可以是10至50个碱基。在一些实施方案中, 指导核酸可以是至少25个碱基。在一些情况下, 指导核酸具有与靶核酸反向互补的正好或约12个核苷酸(nt)至约80nt、约12nt至约50nt、12nt至约45nt、12nt至约40nt、12nt至约35nt、12nt至约30nt、12nt至约25nt、12nt至约20nt、12nt至约19nt、19nt至

约20nt、19nt至约25nt、19nt至约30nt、19nt至约35nt、19nt至约40nt、19nt至约45nt、19nt至约50nt、19nt至约60nt、20nt至约25nt、20nt至约30nt、20nt至约35nt、20nt至约40nt、20nt至约45nt、20nt至约50nt、或约20nt至约60nt。在一些情况下,指导核酸具有与靶核酸反向互补的约10nt至约60nt、约20nt至约50nt、或约30nt至约40nt。应当理解,指导核酸的序列不需要与其靶核酸的序列100%反向互补以特异性地可杂交、可杂交或特异性结合。指导核酸可以具有在核酸残基5至20的区域中包含至少一个尿嘧啶的序列,该序列与靶核酸中的修饰可变区域反向互补。在一些情况下,指导核酸具有在核酸残基5至9、10至14或15至20的区域中包含至少一个尿嘧啶的序列,该序列与靶核酸中的修饰可变区域反向互补。指导核酸可以具有在核酸残基5至20的区域中包含至少一个尿嘧啶的序列,该序列与靶核酸中的甲基化可变区域反向互补。在一些情况下,指导核酸具有在核酸残基5至9、10至14或15至20的区域中包含至少一个尿嘧啶的序列,该序列与靶核酸中的甲基化可变区域反向互补。指导核酸可以与靶核酸杂交。

[0065] 指导核酸(例如,非天然存在的指导核酸)可以选自指导核酸组,该组指导核酸已经与感染株或目的基因组基因座的核酸序列平铺。指导核酸可以选自指导核酸组,该组指导核酸已经相对于HPV 16或HPV18毒株的核酸平铺。通常,已经相对于感染株或感兴趣基因组基因座的核酸平铺的指导核酸可以合并(pooling)以用于本文所述的方法。通常,这些指导核酸被合并用于在单个测定中检测靶核酸。相对于单个靶核酸平铺的指导核酸的合并可以增强使用本文所述的方法对靶核酸的检测。相对于单个靶核酸平铺的指导核酸的合并可以确保使用本文所述的方法在单个反应内对靶核酸的广泛覆盖。例如,平铺是沿着靶核酸顺序进行的。平铺有时是沿着靶核酸重叠的。在一些情况下,平铺包括沿着靶核酸平铺的指导核酸之间的间隙。在一些实例中,指导核酸的平铺是非连续的。通常,用于检测靶核酸的方法包括使靶核酸与指导核酸池和可编程切口酶(例如,本文公开的Cas14a、Cas14b或Cas14e)接触,其中指导核酸池的指导核酸序列具有选自与靶核酸的核酸序列相对应的平铺指导核酸组的序列;以及测定通过切割报告子的核酸群中的至少一些报告子所产生的信号。指导核酸的合并可以确保在单个反应内对靶物种的广谱鉴定或广泛覆盖。这可以在可能由多种生物体引起的疾病或适应症如脓毒症中特别有用。

[0066] 本文公开的组合物可以包含gRNA系统。如本文所述,gRNA系统可以包含crRNA和tracrRNA或trancRNA。在gRNA系统中,crRNA和tracrRNA或trancRNA可以是不同的多核糖核苷酸。

[0067] 在一些实施方案中,gRNA系统中的crRNA包含重复区和间隔区。重复区可以与tracrRNA或trancRNA的区域杂交。间隔区可以与靶核酸的区域杂交。

[0068] gRNA系统中的tracrRNA或trancRNA可以包含与crRNA杂交的区域和与可编程核酸酶(例如,可编程切口酶)相互作用的区域。

[0069] 本公开内容的可编程切口酶(例如,Cas14蛋白)可以在核糖核蛋白(RNP)与靶核酸结合(可编程切口酶和gRNA的复合体)时被激活以表现出切割活性(例如,靶核酸的顺式切割或侧支核酸的反式切割),其中gRNA的crRNA的间隔区与靶核酸杂交。

[0070] 在一些实施方案中,可以使用trancRNA代替tracrRNA。本公开内容的组合物和方法可以包括CasZ反式激活非编码RNA(“trancRNA”;在本文中也称为“CasZ trancRNA”)。在一些情况下,trancRNA与本公开内容的CasZ多肽和CasZ指导RNA形成复合体。trancRNA可以

被鉴定为由CasZ基因座中存在的核苷酸序列编码的高转录RNA。编码trancRNA的序列可以位于cas基因和CasZ基因座阵列(重复区)之间(例如,可以位于与重复序列相邻的位置)。以下实例展示了CasZ trancRNA的检测。在一些情况下,CasZ trancRNA与CasZ多肽免疫共沉淀(形成复合体)。在一些情况下,系统功能需要存在CasZ trancRNA。trancRNA相关数据(例如,它们的表达及其在天然存在的阵列上的位置)在以下实施例章节中提供。

[0071] 在一些情况下,CasZ trancRNA的长度为60个核苷酸(nt)至270nt(例如,60-260、70-270、70-260或75-255nt)。在一些情况下,CasZ trancRNA(例如,CasZa trancRNA)的长度为60-150nt(例如,60-140、60-130、65-150、65-140、65-130、70-150、70-140或70-130nt)。在一些情况下,CasZ trancRNA(例如,CasZa trancRNA)的长度为70-130nt。在一些情况下,CasZ trancRNA(例如,CasZa trancRNA)的长度为约80nt。在一些情况下,CasZ trancRNA(例如,CasZa trancRNA)的长度为约90nt。在一些情况下,CasZ trancRNA(例如,CasZa trancRNA)的长度为约120nt。

[0072] 在一些情况下,CasZ trancRNA(例如,CasZb trancRNA)的长度为85-240nt(例如,85-230、85-220、85-150、85-130、95-240、95-230、95-220、95-150或95-130nt)。在一些情况下,CasZ trancRNA(例如,CasZb trancRNA)的长度为95-120nt。在一些情况下,CasZ trancRNA(例如,CasZb trancRNA)的长度为约105nt。在一些情况下,CasZ trancRNA(例如,CasZb trancRNA)的长度为约115nt。在一些情况下,CasZ trancRNA(例如,CasZb trancRNA)的长度为约215nt。

[0073] 在一些情况下,CasZ trancRNA(例如,CasZc trancRNA)的长度为80-275nt(例如,85-260nt)。在一些情况下,CasZ trancRNA(例如,CasZc trancRNA)的长度为80-110nt(例如,85-105nt)。在一些情况下,CasZ trancRNA(例如,CasZc trancRNA)的长度为235-270nt(例如,240-260nt)。在一些情况下,CasZ trancRNA(例如,CasZc trancRNA)的长度为约95nt。在一些情况下,CasZ trancRNA(例如,CasZc trancRNA)的长度为约250nt。

[0074] 本公开内容的组合物和方法包括Cas14反式激活非编码RNA(“trancRNA”;在本文中也称为“Cas14 trancRNA”)。在一些情况下,trancRNA与本公开内容的Cas14多肽和Cas14指导RNA形成复合体。trancRNA可以被鉴定为由Cas14基因座中存在的核苷酸序列编码的高转录RNA。编码trancRNA的序列通常位于cas基因和Cas14基因座阵列(重复区)之间(例如,可以位于与重复序列相邻的位置)。以下实例展示了Cas14 trancRNA的检测。在某些情况下,Cas14 trancRNA与CasZ多肽免疫共沉淀(形成复合体)。在一些情况下,系统功能需要存在CasZ trancRNA。

[0075] 在一些情况下,Cas14 trancRNA的长度为60个核苷酸(nt)至270nt(例如,60-260、70-270、70-260或75-255nt)。在一些情况下,Cas14 trancRNA(例如,Cas14a trancRNA)的长度为60-150nt(例如,60-140、60-130、65-150、65-140、65-130、70-150、70-140或70-130nt)。在一些情况下,Cas14 trancRNA(例如,Cas14a trancRNA)的长度为70-130nt。在一些情况下,Cas14 trancRNA(例如,Cas14a trancRNA)的长度为约80nt。在一些情况下,Cas14 trancRNA(例如,Cas14a trancRNA)的长度为约90nt。在一些情况下,Cas14 trancRNA(例如,Cas14a trancRNA)的长度为约120nt。

[0076] 在一些情况下,Cas14 trancRNA(例如,Cas14b trancRNA)的长度为85-240nt(例如,85-230、85-220、85-150、85-130、95-240、95-230、95-220、95-150或95-130nt)。在一些

情况下,Cas14 trancRNA(例如,Cas14b trancRNA)的长度为95-120nt。在一些情况下,Cas14trancRNA(例如,Cas14b trancRNA)的长度为约105nt。在一些情况下,Cas14 trancRNA(例如,Cas14b trancRNA)的长度为约115nt。在一些情况下,Cas14 trancRNA(例如,Cas14b trancRNA)的长度为约215nt。

[0077] 在一些情况下,Cas14 trancRNA(例如,Cas14c trancRNA)的长度为80-275nt(例如,85-260nt)。在一些情况下,Cas14 trancRNA(例如,Cas14c trancRNA)的长度为80-110nt(例如,85-105nt)。

在一些情况下,Cas14 trancRNA(例如,Cas14c trancRNA)的长度为235-270nt(例如,240-260nt)。在一些情况下,Cas14 trancRNA(例如,Cas14c trancRNA)的长度为约95nt。在一些情况下,Cas14trancRNA(例如,Cas14c trancRNA)的长度为约250nt。

样品

[0078] 大量样品与本文公开的组合物和方法兼容。如本文所述,样品可用于在本文公开的使靶核酸产生切口的方法中。如本文所述,样品可用于本文公开的DETECTR测定方法中。如本文所述,样品与本文公开的任何可编程切口酶以及所述可编程切口酶在检测靶核酸的方法中的用途兼容。如本文所述,样品与包含可编程切口酶和缓冲液的任何组合物兼容。本文描述的是含有脱氧核糖核酸(DNA)、核糖核酸(RNA)或两者的样品,可以使用本公开内容的可编程切口酶对其进行修饰或检测。如本文所述,可编程切口酶在指导核酸与靶核酸杂交后在与样品中目的靶核酸结合后被激活。随后,激活的可编程切口酶表现出报告子中的核酸的序列非依赖性切割。报告子另外包括可检测部分,其在报告子中的核酸的序列非依赖性切割时被释放。可检测部分发出可检测信号,该信号可以通过各种方法(例如,分光光度法、荧光测量、电化学测量)进行测量。

[0079] 包含目的靶核酸的各种样品类型与本公开内容一致。这些样品可以包含用于检测的靶核酸序列。在一些实施方案中,靶核酸的检测表明病痛,例如疾病、癌症或遗传病症或遗传信息,例如用于表型分析、基因分型或确定血统并且与本文所述的试剂和支持介质兼容。通常,可以获取来自个体或动物的样品或环境样品以测试是否存在疾病、癌症、遗传病或任何感兴趣的突变。来自个体的生物样品可以是血液、血清、血浆、唾液、尿液、粘膜样品、腹膜样品、脑脊髓液、胃分泌物、鼻分泌物、痰、咽溢泌物、尿道或阴道分泌物、溢泌物、积液或组织。组织样品可以在应用于本公开内容的检测系统之前被解离或液化。来自环境的样品可以是土壤、空气或水。在一些情况下,环境样品作为来自感兴趣表面的拭子获取或直接从感兴趣的表面获取。在一些情况下,将原始样品施加至检测系统。在一些情况下,在施加至检测系统之前,将样品用缓冲液或流体稀释或浓缩或者纯净地施加至检测系统。有时,样品容纳在不超过20 μ l中。在一些情况下,样品容纳在不超过1、5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、90、100、200、300、400、500 μ l或1 μ l至500 μ l、10 μ l至200 μ l、50 μ l至200 μ l的任意值中,有时,样品容纳在超过500 μ l中。

[0080] 在一些实施方案中,靶核酸是单链DNA。本文公开的方法、试剂、酶和试剂盒可以使能够直接检测编码目的序列的DNA,特别是编码目的序列的单链DNA,而无需例如,通过使用RNA聚合酶将DNA转录成RNA。本文公开的组合物和方法可以使能够检测靶核酸,该靶核酸是目的核酸的扩增核酸。在一些实施方案中,靶核酸是cDNA、基因组DNA、基因组DNA的扩增子或RNA的DNA扩增子。核酸可以编码来自基因组基因座的序列。在一些情况下,与指导核

酸结合的靶核酸的长度为5至100、5至90、5至80、5至70、5至60、5至50、5至40、5至30、5至25、5至20、5至15、或5至10个核苷酸。核酸的长度可以为10至90、20至80、30至70、或40至60个核苷酸。核酸的长度可以为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、30、31、32、33、34、35、36、37、38、39、40、45、50、60、70、80、90或100个核苷酸。靶核酸可以编码与指导核酸序列反向互补的序列。

[0081] 在一些情况下,样品取自单细胞真核生物体;植物或植物细胞;藻细胞;真菌细胞;动物细胞、组织或器官;无脊椎动物的细胞、组织或器官;脊椎动物如鱼、两栖动物、爬行动物、鸟和哺乳动物的细胞、组织、流体或器官;哺乳动物如人,非人类灵长动物、有蹄类动物、猫、牛、绵羊和山羊的细胞、组织、流体或器官。在一些情况下,样品取自线虫、原生动物、蠕虫或疟疾寄生虫。在一些情况下,样品包含来自真核细胞、哺乳动物细胞、人类细胞、原核细胞或植物细胞的细胞裂解物的核酸。在一些情况下,样品包含从细胞表达的核酸。

[0082] 本文所述的样品可以包含至少一种靶核酸。靶核酸包含与指导核酸的区段反向互补的区段。通常,样品包含靶核酸的区段和与靶核酸的区段包含至少50%序列同一性的至少一种核酸。有时,所述至少一种核酸包含与靶核酸的区段包含至少60%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的区段。通常,样品包含靶核酸的区段和至少一种核酸区段,所述至少一种核酸区段包含与靶核酸小于100%序列同一性但与靶核酸的区段不少于50%序列同一性。有时,样品包含靶核酸的区段和至少一种核酸区段,所述至少一种核酸区段包含与靶核酸小于100%序列同一性但与靶核酸的区段不少于60%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性。例如,与包含与靶核酸的区段包含小于100%序列同一性但与靶核酸的区段包含不小于50%序列同一性的区段的至少一种核酸相比,靶核酸的区段包含突变。有时,与包含与靶核酸的区段包含小于100%序列同一性但与靶核酸的区段包含不少于60%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的区段的至少一种核酸相比,靶核酸的区段包含突变。通常,与包含与靶核酸的区段包含小于100%序列同一性但与靶核酸的区段包含不小于50%序列同一性的区段的至少一种核酸相比,靶核酸的区段包含突变。突变可以是1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20或更多个核苷酸的突变。通常,突变是单核苷酸突变。单核苷酸突变可以是单核苷酸多态性(SNP),其是少于1%的群体中存在的DNA序列中的单碱基对变异。有时,靶核酸包含单核苷酸突变,其中单核苷酸突变包含SNP的野生型变体。单核苷酸突变或SNP可以与样品的表型或取样的生物体的表型相关。在一些情况下,SNP与相对于野生型表型的表型改变有关。通常,与包含与靶核酸的区段包含小于100%序列同一性但与靶核酸的区段包含不小于50%序列同一性的区段的至少一种核酸相比,靶核酸序列的区段包含缺失。突变可以是1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20个或更多个核苷酸缺失。突变可以是约5、约10、约15、约20、约25、约30、约35、约40、约45、约50、约55、约60、约65、约70、约75、约80、约85、约90、约95、约100、约200、约300、约400、约500、约600、约700、约800、约900或约1000个核苷酸缺失。突变可以是1至5、5至10、10至15、15至20、20至25、25至30、30至35、35至40、40至45、45至50、50至55、55至60、60至65、65至70、70至75、75至80、80至85、85至90、90至95、95至100、100至200、200至300、300至400、400至500、500至600、600至700、700至800、800至900、900至1000、1至50、1至100、25至50、25至100、50至100、100

至500、100至1000、或500至1000个核苷酸缺失。与本文所述方法的指导核酸结合的靶核酸区段包含突变,例如SNP或缺失。突变可以是单核苷酸突变或SNP。SNP可以是同义置换或非同义置换。非同义置换可以是错义置换或无义点突变。同义置换可以是沉默置换。突变可以是一个或多个核苷酸缺失。通常,单核苷酸突变、SNP或缺失与疾病(例如癌症或遗传病症)有关。突变,例如单核苷酸突变、SNP或缺失,可以在来自生物体种系的靶核酸的序列中被编码或者可以在来自患病细胞的靶核酸中(例如癌细胞)被编码。

[0083] 用于疾病检测的样品可以包含能够与本文所述试剂的指导核酸结合的至少一种靶核酸。用于疾病检测的样品可以至少包含目的核酸,其被扩增以产生可以与本文所述试剂的指导核酸结合的靶核酸。目的核酸可以包含DNA、RNA或其组合。

[0084] 靶核酸(例如:靶DNA)可以是导致样品中的疾病的病毒或细菌或其他物剂的核酸的一部分。靶核酸可以是来自样品的疾病中表达的基因的核酸的一部分。在一些情况下,该序列是靶核酸序列的区段。靶核酸序列的区段可以来自基因组基因座、转录mRNA或逆转录cDNA。靶核酸序列的区段的长度可以是5至100、5至90、5至80、5至70、5至60、5至50、5至40、5至30、5至25、5至20、5至15、或5至10个核苷酸。靶核酸序列的区段的长度可以是5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、45、50、60、70、80、90或100个核苷酸。靶核酸区段的序列可以与指导核酸序列的区段反向互补。相对于标准样品,靶核酸可以包含与疾病表型或疾病倾向相关的遗传变异(例如,单核苷酸多态性)。靶核酸可以是RNA的部分的扩增子,可以是DNA或者可以是来自样品中任何生物体的DNA扩增子。

[0085] 在一些实施方案中,靶核酸序列包含病毒或细菌或者样品中导致疾病的其他因子的核酸序列。在一些实施方案中,靶核酸包含在使用本文公开的组合物、系统和方法通过可编程切口酶(例如,本文公开的Cas14a、Cas14b或Cas14e)检测之前使用逆转录酶从RNA逆转录的DNA。在一些情况下,靶核酸是样品中来自性传播感染或传染病的核酸的一部分。在一些情况下,靶核酸是以下至少一项中的来自基因组基因座的核酸的一部分或任何DNA扩增子,例如来自基因基因座的逆转录mRNA或cDNA、转录的mRNA或来自基因基因座的逆转录cDNA:人类免疫缺陷病毒(HIV)、人类乳头瘤病毒(HPV)、衣原体、淋病、梅毒、滴虫、性传播感染、疟疾、登革热、埃博拉、基孔肯雅病和利什曼病。病原体包括病毒、真菌、蠕虫、原生动物、疟疾寄生虫、疟原虫属寄生虫、弓形虫属寄生虫和血吸虫属寄生虫。蠕虫包括蛔虫、心丝虫和吞噬线虫、吸虫、棘头虫和绦虫。原生动物感染包括来自贾第虫属种、滴虫属种、非洲锥虫病、阿米巴痢疾、巴贝虫病、小袋虫痢疾、南美锥虫病、球虫病、疟疾和弓形虫病的感染。诸如寄生虫/原生动物病原体之类的病原体的示例包括但不限于:恶性疟原虫(*Plasmodium falciparum*)、间日疟原虫(*P. vivax*)、克氏锥虫(*Trypanosoma cruzi*)和鼠弓形虫(*Toxoplasma gondii*)。真菌病原体包括但不限于新型隐球菌(*Cryptococcus neoformans*)、荚膜组织胞浆菌(*Histoplasma capsulatum*)、粗球孢子菌(*Coccidioides immitis*)、皮炎芽生菌(*Blastomyces dermatitidis*)、砂眼衣原体(*Chlamydia trachomatis*)和白色念珠菌(*Candida albicans*)。病原性病毒包括但不限于免疫缺陷病毒(例如,HIV);流感病毒;登革热;西尼罗病毒;疱疹病毒;黄热病病毒;丙型肝炎病毒;甲型肝炎病毒;乙型肝炎病毒;乳头瘤病毒;等等。病原体包括例如HIV病毒、结核分枝杆菌(*Mycobacterium tuberculosis*)、无乳链球菌(*Streptococcus agalactiae*)、耐甲氧西林

金黄色葡萄球菌 (*methicillin-resistant Staphylococcus aureus*)、嗜肺军团菌 (*Legionella pneumophila*)、酿脓链球菌 (*Streptococcus pyogenes*)、大肠杆菌 (*Escherichia coli*)、淋病奈瑟氏球菌 (*Neisseria gonorrhoeae*)、脑膜炎奈瑟氏球菌 (*Neisseria meningitidis*)、肺炎球菌 (*Pneumococcus*)、新型隐球菌 (*Cryptococcus neoformans*)、荚膜组织胞浆菌 (*Histoplasma capsulatum*)、乙型流感嗜血杆菌 (*Hemophilus influenzae B*)、苍白密螺旋体 (*Treponema pallidum*)、莱姆病螺旋体、铜绿假单胞菌 (*Pseudomonas aeruginosa*)、麻风分枝杆菌 (*Mycobacterium leprae*)、流产布鲁氏菌 (*Brucella abortus*)、狂犬病病毒、流感病毒、巨细胞病毒、单纯疱疹病毒I、单纯疱疹病毒II、人血清细小病毒样病毒、呼吸道合胞病毒 (RSV)、生殖器支原体 (*M.genitalium*)、阴道鞭毛虫 (*T.vaginalis*)、水痘带状疱疹病毒、乙型肝炎病毒、丙型肝炎病毒、麻疹病毒、腺病毒、人类T细胞白血病病毒、爱泼斯坦-巴尔病毒、鼠白血病病毒、腮腺炎病毒、水疱性口炎病毒、辛德比斯病毒、淋巴细胞性脉络膜脑膜炎病毒、疣病毒、蓝舌病毒、仙台病毒、猫白血病病毒、呼肠孤病毒、脊髓灰质炎病毒、猿猴病毒40、小鼠乳腺肿瘤病毒、登革热病毒、风疹病毒、西尼罗河病毒、恶性疟原虫 (*Plasmodium falciparum*)、间日疟原虫 (*Plasmodium vivax*)、鼠弓形虫 (*Toxoplasma gondii*)、让氏锥虫 (*Trypanosoma rangeli*)、克鲁氏锥虫 (*Trypanosoma cruzi*)、罗得西亚锥虫 (*Trypanosoma rhodesiense*)、布氏锥虫 (*Trypanosoma brucei*)、曼氏裂体吸虫 (*Schistosoma mansoni*)、日本血吸虫 (*Schistosoma japonicum*)、牛巴贝虫 (*Babesia bovis*)、盲肠型球虫 (*Eimeria tenella*)、旋盘尾丝虫 (*Onchocerca volvulus*)、热带利什曼原虫 (*Leishmania tropica*)、结核分枝杆菌 (*Mycobacterium tuberculosis*)、旋毛虫 (*Trichinella spiralis*)、小泰勒虫 (*Theileria parva*)、胞状绦虫 (*Taenia hydatigena*)、羊绦虫 (*Taenia ovis*)、牛肉绦虫 (*Taenia saginata*)、细粒棘球绦虫 (*Echinococcus granulosus*)、科氏中殖孔绦虫 (*Mesocestoides corti*)、关节炎支原体 (*Mycoplasma arthritidis*)、猪鼻炎支原体 (*M.hyorhinis*)、口腔支原体 (*M. orale*)、精氨酸支原体 (*M.arginini*)、莱氏无胆甾原体 (*Acholeplasma laidlawii*)、唾液支原体 (*M.salivarium*) 和肺炎支原体 (*M.pneumoniae*)。在一些情况下,靶核酸是来自样品中导致疾病的细菌或其他物剂的来自基因组基因座的核酸、从基因组转录的mRNA或逆转录的cDNA的一部分。

[0086] 用于癌症测试的样品可以包含可以与本文所述试剂的指导核酸结合的至少一个靶核酸区段。在一些情况下,靶核酸区段是来自以下基因的核酸的一部分:与癌症相关的突变的基因、其过表达与癌症有关的基因、抑癌基因、癌基因、检查点抑制剂基因、与细胞生长相关的基因、与细胞代谢相关的基因或与细胞周期相关的基因。有时,靶核酸编码癌症生物标志物,例如前列腺癌生物标志物或非小细胞肺癌。在一些情况下,该测定可用于检测靶核酸中可预测肺癌的“热点”。

在一些情况下,靶核酸是与血热相关的核酸的一部分。在一些情况下,靶核酸区段是来自基因组基因座、从基因座转录的mRNA或逆转录的cDNA的核酸的一部分,该基因座是以下至少一个:ALK、APC、ATM、AXIN2、BAP1、BARD1、BLM、BMPR1A、BRCA1、BRCA2、BRIP1、CASR、CDC73、CDH1、CDK4、CDKN1B、CDKN1C、CDKN2A、CEBPA、CHEK2、CTNNA1、DICER1、DIS3L2、EGFR、EPCAM、FH、FLCN、GATA2、GPC3、GREM1、HOXB13、HRAS、KIT、MAX、MEN1、MET、MITF、MLH1、MSH2、MSH3、MSH6、MUTYH、NBN、NF1、NF2、NTHL1、PALB2、PDGFRA、PHOX2B、PMS2、POLD1、POLE、POT1、

PRKAR1A、PTCH1、PTEN、RAD50、RAD51C、RAD51D、RB1、RECQL4、RET、RUNX1、SDHA、SDHAF2、SDHB、SDHC、SDHD、SMAD4、SMARCA4、SMARCB1、SMARCE1、STK11、SUFU、TERC、TERT、TMEM127、TP53、TSC1、TSC2、VHL、WRN和WT1。可以使用本文公开的组合物和方法探测前述基因座的任何区域的突变或缺失。例如,在EGFR基因座中,本文公开的组合物和检测方法可用于检测单核苷酸多态性或缺失。SNP或缺失可以发生在非编码区或编码区。SNP或缺失可以发生在外显子例如外显子19中。

[0087] 用于遗传病测试的样品可以包含可以与本文所述试剂的指导核酸结合的至少一个靶核酸区段。在一些实施方案中,遗传病是血友病、镰状细胞性贫血、 β -地中海贫血,杜氏肌营养不良、严重联合免疫缺陷或囊性纤维化。在一些情况下,靶核酸区段是来自以下基因的核酸的一部分:具有与遗传病相关的突变的基因、其过表达与遗传病相关的基因、与导致遗传病的异常细胞生长相关的基因或与导致遗传病的异常细胞代谢相关的基因。在一些情况下,靶核酸区段是来自基因组基因座、从基因座转录的mRNA或逆转录的cDNA的核酸的一部分,该基因座是以下至少一个:CFTR、FMR1、SMN1、ABCB11、ABCC8、ABCD1、ACAD9、ACADM、ACADVL、ACAT1、ACOX1、ACSF3、ADA、ADAMTS2、ADGRG1、AGA、AGL、AGPS、AGXT、AIRE、ALDH3A2、ALDOB、ALG6、ALMS1、ALPL、AMT、AQP2、ARG1、ARSA、ARSB、ASL、ASNS、ASPA、ASS1、ATM、ATP6V1B1、ATP7A、ATP7B、ATRX、BBS1、BBS10、BBS12、BBS2、BCKDHA、BCKDHB、BCS1L、BLM、BSND、CAPN3、CBS、CDH23、CEP290、CERKL、CHM、CHRNE、CIITA、CLN3、CLN5、CLN6、CLN8、CLRN1、CNGB3、COL27A1、COL4A3、COL4A4、COL4A5、COL7A1、CPS1、CPT1A、CPT2、CRB1、CTNS、CTSK、CYBA、CYBB、CYP11B1、CYP11B2、CYP17A1、CYP19A1、CYP27A1、DBT、DCLRE1C、DHCR7、DHDDS、DLD、DMD、DNAH5、DNAI1、DNAI2、DYSF、EDA、EIF2B5、EMD、ERCC6、ERCC8、ESCO2、ETFA、ETFDH、ETHE1、EVC、EVC2、EYS、F9、FAH、FAM161A、FANCA、FANCC、FANCG、FH、FKRP、FKTN、G6PC、GAA、GALC、GALK1、GALT、GAMT、GBA、GBE1、GCDH、GFM1、GJB1、GJB2、GLA、GLB1、GLDC、GLE1、GNE、GNPTAB、GNPTG、GNS、GRHPR、HADHA、HAX1、HBA1、HBA2、HBB、HEXA、HEXB、HGSNAT、HLCS、HMGCL、HOGA1、HPS1、HPS3、HSD17B4、HSD3B2、HYAL1、HYLS1、IDS、IDUA、IKBKAP、IL2RG、IVD、KCNJ11、LAMA2、LAMA3、LAMB3、LAMC2、LCA5、LDLR、LDLRAP1、LHX3、LIFR、LIPA、LOXHD1、LPL、LRPPRC、MAN2B1、MCOLN1、MED17、MESP2、MFSD8、MKS1、MLC1、MMAA、MMAB、MMACHC、MMADHC、MPI、MPL、MPV17、MTHFR、MTM1、MTRR、MTTP、MUT、MYO7A、NAGLU、NAGS、NBN、NDRG1、NDUFAF5、NDUFS6、NEB、NPC1、NPC2、NPHS1、NPHS2、NR2E3、NTRK1、OAT、OPA3、OTC、PAH、PC、PCCA、PCCB、PCDH15、PDHA1、PDHB、PEX1、PEX10、PEX12、PEX2、PEX6、PEX7、PFKM、PHGDH、PKHD1、PMM2、POMGNT1、PPT1、PROP1、PRPS1、PSAP、PTS、PUS1、PYGM、RAB23、RAG2、RAPSN、RARS2、RDH12、RMRP、RPE65、RPGRIP1L、RS1、RTEL1、SACS、SAMHD1、SEPSecs、SGCA、SGCB、SGCG、SGSH、SLC12A3、SLC12A6、SLC17A5、SLC22A5、SLC25A13、SLC25A15、SLC26A2、SLC26A4、SLC35A3、SLC37A4、SLC39A4、SLC4A11、SLC6A8、SLC7A7、SMARCA1、SMPD1、STAR、SUMF1、TAT、TCIRG1、TECPR2、TFR2、TGM1、TH、TMEM216、TPP1、TRMU、TSFM、TPPA、TYMP、USH1C、USH2A、VPS13A、VPS13B、VPS45、VRK1、VSX2、WNT10A、XPA、XPC和ZFYVE26。

[0088] 用于表型分析测试的样品可以包含可以与本文所述试剂的指导核酸结合的至少一个靶核酸区段。在一些情况下,靶核酸区段是来自与表型性状有关的基因的核酸的一部分。

[0089] 用于基因分型测试的样品可以包含可以与本文所述试剂的指导核酸结合的至少

一个靶核酸区段。在一些情况下,靶核酸区段是来自与基因型有关的基因的核酸的一部分。

[0090] 用于祖先测试的样品可以包含可以与本文所述试剂的指导核酸结合的至少一个靶核酸区段。在一些情况下,靶核酸区段是来自与起源或种族的地理区域有关的基因的核酸的一部分。

[0091] 样品可用于鉴定疾病状态。例如,样品是本文所述的任何样品,并且是从受试者获得的,用于鉴定受试者的疾病状态。该疾病可以是癌症或遗传病。有时,方法包括从受试者获得血清样品;并鉴定受试者的疾病状态。通常,疾病状态是前列腺疾病状态。

[0092] 在一些情况下,靶核酸是单链核酸。替代地或组合地,靶核酸是双链核酸,并且在与试剂接触之前或之后被制备成单链核酸。靶核酸可以是RNA、DNA、合成核酸或在生物或环境样品中发现的核酸。靶核酸包括但不限于mRNA、rRNA、tRNA、非编码RNA、长非编码RNA和微小RNA(miRNA)。在一些情况下,靶核酸是mRNA。在一些情况下,靶核酸来自本文所述的病毒、寄生虫或细菌。在一些情况下,靶核酸由本文所述的基因转录。

[0093] 许多靶核酸与本文公开的方法和组合物一致。本文描述的一些方法可以检测作为靶核酸群体,以各种浓度或量存在于样品中的靶核酸。在一些情况下,样品具有至少2个靶核酸。在一些情况下,样品具有至少3、5、10、20、30、40、50、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000或10000个靶核酸。在一些情况下,样品具有1至10,000、100至8000、400至6000、500至5000、1000至4000或2000至3000个靶核酸。在一些情况下,该方法检测以每 10^1 个非靶核酸、 10^2 个非靶核酸、 10^3 个非靶核酸、 10^4 个非靶核酸、 10^5 个非靶核酸、 10^6 个非靶核酸、 10^7 个非靶核酸、 10^8 个非靶核酸、 10^9 个非靶核酸或 10^{10} 个非靶核酸至少一个拷贝存在的靶核酸。通常,靶核酸可以占样品中总核酸的0.05%至20%。有时,靶核酸占样品中总核酸的0.1%至10%。在一些情况下,靶核酸占样品中总核酸的0.1%至5%。靶核酸还可以占样品中总核酸的0.1%至1%。靶核酸可以是DNA或RNA。靶核酸可以是小于样品中总核酸的100%的任何量。靶核酸可以占样品中总核酸的100%。

[0094] 在一些实施方案中,样品包含浓度小于1nM、小于2nM、小于3nM、小于4nM、小于5nM、小于6nM、小于7nM、小于8nM、小于9nM、小于10nM、小于20nM、小于30nM、小于40nM、小于50nM、小于60nM、小于70nM、小于80nM、小于90nM、小于100nM、小于200nM、小于300nM、小于400nM、小于500nM、小于600nM、小于700nM、小于800nM、小于900nM、小于1 μ M、小于2 μ M、小于3 μ M、小于4 μ M、小于5 μ M、小于6 μ M、小于7 μ M、小于8 μ M、小于9 μ M、小于10 μ M、小于100 μ M、或小于1mM的靶核酸。在一些实施方案中,样品包含浓度为1nM至2nM、2nM至3nM、3nM至4nM、4nM至5nM、5nM至6nM、6nM至7nM、7nM至8nM、8nM至9nM、9nM至10nM、10nM至20nM、20nM至30nM、30nM至40nM、40nM至50nM、50nM至60nM、60nM至70nM、70nM至80nM、80nM至90nM、90nM至100nM、100nM至200nM、200nM至300nM、300nM至400nM、400nM至500nM、500nM至600nM、600nM至700nM、700nM至800nM、800nM至900nM、900nM至1 μ M、1 μ M至2 μ M、2 μ M至3 μ M、3 μ M至4 μ M、4 μ M至5 μ M、5 μ M至6 μ M、6 μ M至7 μ M、7 μ M至8 μ M、8 μ M至9 μ M、9 μ M至10 μ M、10 μ M至100 μ M、100 μ M至1mM、1nM至10nM、1nM至100nM、1nM至1 μ M、1nM至10 μ M、1nM至100 μ M、1nM至1mM、10nM至100nM、10nM至1 μ M、10nM至10 μ M、10nM至100 μ M、10nM至1mM、100nM至1 μ M、100nM至10 μ M、100nM至100 μ M、100nM至1mM、1 μ M至10 μ M、1 μ M至100 μ M、1 μ M至1mM、10 μ M至100 μ M、10 μ M至1mM、或100 μ M至1mM的靶核酸序列。在一些实施方案中,样品包含浓度为20nM至200 μ M、50nM至100 μ M、200nM至50 μ M、500nM至20 μ M、或

2 μ M至10 μ M的靶核酸。在一些实施方案中,样品中不存在靶核酸。

[0095] 在一些实施方案中,样品包含少于10个拷贝、少于100个拷贝、少于1000个拷贝、少于10,000个拷贝、少于100,000个拷贝、或少于1,000,000个拷贝的靶核酸序列。在一些实施方案中,样品包含10个拷贝至100个拷贝、100个拷贝至1000个拷贝、1000个拷贝至10,000个拷贝、10,000个拷贝至100,000个拷贝、100,000个拷贝至1,000,000个拷贝、10个拷贝至1000个拷贝、10个拷贝至10,000个拷贝、10个拷贝至100,000个拷贝、10个拷贝至1,000,000个拷贝、100个拷贝至10,000个拷贝、100个拷贝至100,000个拷贝、100个拷贝至1,000,000个拷贝、1,000个拷贝至100,000个拷贝、或1,000个拷贝至1,000,000个拷贝的靶核酸序列。在一些实施方案中,样品包含10个拷贝至500,000个拷贝、200个拷贝至200,000个拷贝、500个拷贝至100,000个拷贝、1000个拷贝至50,000个拷贝、2000个拷贝至20,000个拷贝、3000个拷贝至10,000个拷贝、或4000个拷贝至8000个拷贝。在一些实施方案中,样品中不存在靶核酸。

[0096] 许多靶核酸群体与本文所述的方法和组合物一致。本文所述的一些方法可以检测以不同浓度或量存在于样品中的两个或更多个靶核酸群体。在一些情况下,样品具有至少2个靶核酸群体。在一些情况下,样品具有至少3、4、5、6、7、8、9、10、20、30、40或50个靶核酸群体。在一些情况下,样品具有3至50、5至40或10至25个靶核酸群体。在一些情况下,该方法检测以每10¹个非靶核酸、10²个非靶核酸、10³个非靶核酸、10⁴个非靶核酸、10⁵个非靶核酸、10⁶个非靶核酸、10⁷个非靶核酸、10⁸个非靶核酸、10⁹个非靶核酸或10¹⁰个非靶核酸至少一个拷贝存在的靶核酸群体。靶核酸群体可以以不同的浓度或量存在于样品中。

[0097] 在一些实施方案中,本文公开的靶核酸可以激活可编程切口酶以启动基于核酸的报告子(例如,包含DNA序列的报告子、包含RNA序列的报告子、或包含DNA和RNA的报告子)的序列非依赖性切割。例如,本公开内容的可编程切口酶被靶DNA激活以切割具有RNA的报告子(本文也称为“RNA报告子”)。可选地,本公开内容的可编程切口酶被靶RNA激活以切割具有RNA的报告子。或者,本公开内容的可编程切口酶被靶DNA激活以切割具有DNA的报告子(本文也称为“DNA报告子”)。RNA报告子可以包含用检测部分标记的单链RNA或者可以是如本文公开的任何RNA报告子。DNA报告子可以包含用检测部分标记的单链DNA或者可以是本文公开的任何DNA报告子。

[0098] 在一些实施方案中,本文方法中描述的靶核酸最初不包含PAM序列。然而,可以使用本文所述的方法产生任何目的靶核酸以包含PAM序列,并且因此为PAM靶核酸。如本文所用,PAM靶核酸是指已经扩增以插入被CRISPR/Cas系统识别的PAM序列的靶核酸。

[0099] 任何上述公开的样品与本文公开的方法、组合物、试剂、酶和试剂盒一致,可用作本文公开的任何疾病的伴随诊断或可用于试剂盒、即时诊断或非处方诊断。

使靶核酸产生切口的方法

[0100] 本文公开了将断裂引入靶核酸的方法。在一些实施方案中,断裂可以是单链断裂(例如,切口)。本文公开的可编程切口酶(例如,本文公开的Cas14a、Cas14b或Cas14e)和gRNA系统(例如,包含crRNA和tracrRNA的gRNA或包含crRNA和tracrRNA的gRNA)可用于将断裂引入靶核酸。将断裂引入靶核酸的方法可以包括使靶核酸与第一指导核酸(例如,包含与长度不超过900个氨基酸的第一可编程切口酶结合的第一区域的指导核酸)和第二指导核酸(例如,包含与长度不超过900个氨基酸的第二可编程切口酶结合的第一区域的指导核

酸)接触。第一指导核酸可以包含与靶核酸结合的第二区域,并且第二指导核酸可以包含与靶核酸结合的第二区域。第一指导核酸的第二区域和第二指导核酸的第二区域可以结合靶核酸的相反链。

[0101] 本文所述的方法(例如,将切口引入靶核酸的方法)可用于修饰靶核酸。修饰靶核酸的方法可以使用本文所述的包含可编程切口酶(例如,Cas14蛋白)和gRNA系统(例如,包含crRNA和tracrRNA的gRNA系统或包含crRNA和trancRNA的gRNA系统)的组合物。修饰靶核酸可以包括切割靶核酸、使靶核酸的一个或多个核苷酸缺失、将一个或多个核苷酸插入靶核酸、使靶核酸的一个或多个核苷酸突变或修饰(例如,甲基化、去甲基化、脱氨基或氧化)靶核酸的一个或多个核苷酸。靶核酸可以包含基因组、染色体、质粒、基因、启动子、非翻译区、开放阅读框、内含子、外显子或操纵子中的一种或多种。靶核酸可以包含基因组、染色体、质粒、基因、启动子、非翻译区、开放阅读框、内含子、外显子或操纵子中的一种或多种的区段。在一些实施方案中,靶核酸可以是细胞或生物体的部分。在一些实施方案中,靶核酸可以是无细胞遗传组分。在一些实施方案中,修饰靶核酸包括基因组编辑。基因组编辑可以包括修饰细胞或生物体的基因组、染色体、质粒或其他遗传物质。在一些实施方案中,细胞或生物体的基因组、染色体、质粒或其他遗传物质在体内经修饰。在一些实施方案中,细胞或生物体的基因组、染色体、质粒或其他遗传物质在细胞中经修饰。在一些实施方案中,在体外修饰细胞或生物体的基因组、染色体、质粒或其他遗传物质。例如,可以使用本文所述的组合物在体外修饰质粒并将其引入细胞或生物体。在一些实施方案中,修饰靶核酸可以包括从靶核酸中删除序列。例如,可以从靶核酸中移除已突变的序列或与疾病相关的序列。在一些实施方案中,修饰靶核酸可以包括用第二序列置换靶核酸中的序列。例如,已突变的序列或与疾病相关的序列可以被不存在突变或与疾病不相关的第二序列置换。在一些实施方案中,修饰靶核酸可以包括将序列引入靶核酸中。例如,可以将有益序列或者可减轻或消除疾病的序列插入靶核酸。

[0102] 修饰靶核酸可以包括在靶核酸中引入断裂(例如,单链断裂)。在一些实施方案中,可以通过使靶核酸与可编程切口酶(例如,Cas14可编程切口酶)指导核酸接触来引入断裂。指导核酸可以与可编程切口酶结合并与靶核酸的区域杂交,从而将可编程切口酶募集至靶核酸区域。可编程切口酶与指导核酸和靶核酸区域的结合可以激活可编程切口酶,并且可编程切口酶可以在靶核酸的区域中引入断裂(例如,单链断裂)。在一些实施方案中,修饰靶核酸可以包括在靶核酸的第一区域引入第一断裂和在靶核酸的第二区域引入第二断裂。例如,修饰靶核酸可以包括使靶核酸与结合第一可编程切口酶并与靶核酸的第一区域杂交的第一指导核酸和结合第二可编程切口酶并与靶核酸的第二区域杂交的第二指导核酸接触。第一可编程切口酶可以在靶核酸的第一区域的第一链中引入第一断裂,并且第二可编程切口酶可以在靶核酸的第二区域的第二链中引入第二断裂。在一些实施方案中,可以移除第一断裂和第二断裂之间的靶核酸区段,从而修饰靶核酸。在一些实施方案中,可以置换(例如,用插入序列)第一断裂和第二断裂之间的靶核酸区段,从而修饰靶核酸。

[0103] 与双链切割活性相比,用于修饰靶核酸的可编程切口酶可以具有更高的产生切口活性。在一些实施方案中,,可编程切口酶可以表现出比双链切割活性高至少约1.1倍、至少约1.2倍、至少约1.3倍、至少约1.4倍、至少约1.5倍、至少约1.6倍、至少约1.7倍、至少约1.8倍、至少约1.9倍、至少约2倍、至少约2.1倍、至少约2.2倍、至少约2.3倍、至少约2.4倍、至少

约2.5倍、至少约2.6倍、至少约2.7倍、至少约2.8倍、至少约2.9倍、至少约3倍、至少约3.5倍、至少约4倍、至少约4.5倍、至少约5倍、至少约5.5倍、至少约6倍、至少约7倍、至少约8倍、至少约9倍、至少约10倍、至少约15倍、至少约20倍、至少约30倍、至少约40倍、或至少约50倍的产生切口活性。

[0104] 在其他情况下,与产生切口活性相比,用于修饰靶核酸的可编程切口酶可以具有更高双链切割活性。在一些实施方案中,可编程切口酶可以表现出比产生切口活性高至少约1.1倍、至少约1.2倍、至少约1.3倍、至少约1.4倍、至少约1.5倍、至少约1.6倍、至少约1.7倍、至少约1.8倍、至少约1.9倍、至少约2倍、至少约2.1倍、至少约2.2倍、至少约2.3倍、至少约2.4倍、至少约2.5倍、至少约2.6倍、至少约2.7倍、至少约2.8倍、至少约2.9倍、至少约3倍、至少约3.5倍、至少约4倍、至少约4.5倍、至少约5倍、至少约5.5倍、至少约6倍、至少约7倍、至少约8倍、至少约9倍、至少约10倍、至少约15倍、至少约20倍、至少约30倍、至少约40倍、或至少约50倍的双链切割活性。

[0105] 在一些实施方案中,可编程切口酶的产生切口活性和双链切割活性依赖于含有可编程切口酶的样品中存在的条件和种类。在一些情况下,可编程切口酶的产生切口活性和双链切割活性对存在的tracrRNA的序列有反应。在一些情况下,可以通过改变存在的tracrRNA的序列来调节产生切口活性和双链切割活性的比率。在一些情况下,可编程切口酶的产生切口活性和双链切割活性对温度、pH、渗透压、缓冲液、靶核酸浓度、离子强度和抑制剂浓度产生不同反应。在一些实施方案中,可以通过调节样品条件和tracrRNA序列来主动控制可编程切口酶的产生切口活性与切割活性的比率。

[0106] 本文所述的组合物和方法可用于治疗、预防或抑制受试者的病痛。例如,包括通过使靶核酸与包含可编程切口酶的组合物接触将切口引入靶核酸的方法可用于治疗、预防或抑制受试者的病痛。该病痛可以包括疾病、癌症、遗传病症、瘤形成和感染。在一些情况下,病痛与一种或多种基因序列有关,该病痛包括但不限于11-羟化酶缺乏症;17,20-碳链酶缺乏症;17-羟化酶缺乏症;3-羟基异丁酸尿症;3-羟基类固醇脱氢酶缺乏症;46,XY性腺发育不全;AAA综合征;ABCA3缺乏症;ABCC8相关的胰岛功能亢进;血浆铜蓝蛋白缺乏症;2型软骨成长不全;肢端脱皮综合征;肠病性肢端皮炎;肾上腺皮质微结节增生;肾上腺脑白质营养不良;肾上腺脊髓神经病;Aicardi-Goutieres综合征;Alagille病;Alpers综合征; α -甘露糖苷贮积症;Alstrom综合征;阿尔茨海默病;牙釉质发育不全;阿米什人小头畸形;肌萎缩性脊髓侧索硬化症;厌食性发育不良;雄激素不敏感综合征;Antley-Bixler综合征;APECED、Apert综合征、泪腺和唾液腺发育不全、精氨酸血症、致心律失常性右心室发育不良、Arts综合征、ARVD2、芳基硫酸酯酶缺乏型异染性脑白质营养不良、共济失调毛细血管扩张症、自身免疫性淋巴增生综合征;1型自身免疫性多腺体综合征;常染色体显性遗传无汗性外胚层发育不良;常染色体显性多囊肾病;常染色体隐性小耳畸形;常染色体隐性肾糖尿;常染色体内脏异位性;Bardet-Biedl综合征;Bartter综合征;基底细胞癌综合征;Batten病;良性复发性肝内胆汁淤积症; β -甘露糖苷贮积症;Bethlem病;Blackfan-Diamond贫血;睑裂狭小;Byler病;C综合征;CADASIL;氨基甲酰磷酸合成酶缺乏症;心面皮肤综合征;胃肠道间质瘤;肉碱棕榈酰转移酶缺乏症;软骨毛发发育不全;cb1C型联合甲基丙二酸尿症;CD18缺乏症;CD3Z相关的原发性T细胞免疫缺陷;CD40L缺乏症;CDAGS综合征;CDG1A;CDG1B;CDG1M;CDG2C;CEDNIK综合征;中央轴空病;中央核肌病;脑毛细血管畸形;4型脑面部

骨骼综合征;脑脊髓神经骨骼综合征;脑腱黄瘤病;CHARGE联合畸形;巨颌症;CHILD综合征;慢性肉芽肿病;慢性复发性多灶性骨髓炎;希特林蛋白缺乏症;经典血色沉着病;CNPPB综合征;钴胺素C病;Cockayne综合征;辅酶Q10缺乏症;Coffin-Lowry综合征;Cohen综合征;凝血因子V联合缺乏症;常见变异免疫缺陷;完全雄激素不敏感;锥杆营养不良;构象疾病;1型先天性胆汁酸合成缺陷;2型先天性胆汁酸合成缺陷;先天性胆汁酸合成缺陷型;先天性红细胞生成性卟啉症;先天性全身性骨硬化;Cornelia de Lange综合征;Cousin综合征;Cowden病;COX缺乏症;Crigler-Najjar病;1型Crigler-Najjar综合征;Crisponi综合征;Currarino综合征;Curth-Macklin型鱼鳞病;皮肤松弛症;胱氨酸病;d-2-羟基戊二酸尿症;DDP综合征;Dejerine-Sottas病;Denys-Drash综合征;结蛋白心肌病;结蛋白肌病;DGUOK相关的线粒体DNA耗竭;谷氨酸代谢紊乱;5型远端脊髓性肌萎缩症;DNA修复疾病;显性视神经萎缩;Doyme蜂窝状视网膜营养不良;杜氏肌营养不良症;先天性角化不良;4型Ehlers-Danlos综合征;Ehlers-Danlos综合征;Elejalde病;Ellis-van Creveld病;Emery-Dreifuss肌营养不良症;脑肌病mtDNA耗竭综合征;酶性疾病;EPCAM相关的先天性簇状肝病;大疱性表皮松解症伴幽门闭锁;运动引起的低血糖症;面肩肱型肌营养不良症;Faisalabad组织细胞增生症;家族性非典型分枝杆菌病;家族性毛细血管畸形-动静脉;家族性食管失弛缓症;家族性球静脉畸形;家族性噬血细胞性淋巴组织细胞增生症;家族性地中海热;家族性巨藻;家族性神经鞘瘤病;家族性脊柱裂;家族性脾脏无脾/发育不全;家族性血栓性血小板减少性紫癜;Fanconi病;Feingold综合征;FENIB;进展性骨化纤维发育不良;FKTN;Francois-Neetens斑点角膜营养不良;Frasier综合征;Friedreich共济失调;FTDP-17;岩藻糖苷贮积症;G6PD缺乏症;半乳糖唾液酸贮积症;Galloway综合征;Gardner综合征;Gaucher病;Gitelman综合征;GLUT1缺乏症;1b型糖原贮积病;2型糖原贮积病;3型糖原贮积病;4型糖原贮积病;9a型糖原贮积病;糖原贮积症;GM1-神经节苷脂沉积症;格林伯格综合征;Greig端部多发性并指综合征;头发遗传病;HANAC综合征;丑角型先天性鱼鳞病;HDR综合征;3型血色素沉着症;4型血色素沉着症;血友病A;3型遗传性血管性水肿;遗传性血管性水肿;遗传性出血性毛细血管扩张症;遗传性低纤维蛋白原血症;遗传性骨内血管畸形;遗传性平滑肌瘤病和肾细胞癌;遗传性神经痛性肌萎缩症;遗传性感觉和自主神经病变类型;Hermansky-Pudlak病;HHH综合征;HHT2;1型汗性外胚层发育不良;汗性外胚层发育不良;HNF4A相关的高胰岛素血症;HNPPC;小头畸形的人类免疫缺陷;亨廷顿病;高IgD综合征;高胰岛素血症-高氨血症综合征;视网膜色素上皮肥大;软骨症;少汗性外胚层发育不良;ICF综合征;特发性先天性假性肠梗阻;1型超IgM免疫缺陷;3型IgM超免疫缺陷;4型超IgM免疫缺陷;5型超IgM的免疫缺陷;甲状腺代谢异常;婴儿内脏肌病;婴儿X连锁脊髓性肌萎缩症;妊娠期肝内胆汁淤积症;IPEX综合征;IRAK4缺乏症;孤立先天性无脾;Jeune综合征影像;Johanson-Blizzard综合征;Joubert综合征;JP-HHT综合征;血色病;幼年透明质纤维瘤病;少年型肾单位肾病;Kabuki面具综合征;Kallmann综合征;Kartagener综合征;KCNJ11相关的胰岛素过多;Kearns-Sayre综合征;Kostmann病;Kozlowski型脊柱干骺端发育不良;Krabbe病;LADD综合征;晚婴儿型神经元蜡样脂褐质沉积症;LCK缺乏症;LDHCP综合征;Legius综合征;Leigh综合征;致死性先天性挛缩综合征2;致死性先天性挛缩综合征;3型致死性挛缩综合征;2型致死性新生儿CPT缺乏症;致死性骨硬化性骨发育不良;LIG4综合征;无脑回畸形1型影像;无脑回畸形3型;Loeys-Dietz综合征;低磷脂相关性胆石症;赖氨酸尿

蛋白不耐受;Maffucci综合征;Majeed综合征;甘露糖结合蛋白缺乏症;Marfan病;Marshall综合征;MASA综合征;MCAD缺乏症;McCune-Albright综合征;MCKD2;Meckel综合征;Meesmann角膜营养不良;巨膀胱-小结肠-肠蠕动减退;1型巨幼红细胞性贫血;MEHMO;MELAS;Melnick-Needles综合征;MEN2;Menkes病;异染性脑白质营养不良;甲基丙二酸尿症;甲基戊酸尿症;小球胆-先天性肾病综合征;微绒毛萎缩;线粒体神经胃肠脑肌病;念珠形发;X单体性;镶嵌型9三体综合征;Mowat-Wilson综合征;2型粘脂贮积症;Ma型粘脂贮积症;IV型粘脂贮积症;粘多糖贮积症;3A型粘多糖贮积症;3C型粘多糖贮积症;4B型粘多糖贮积症;多微小轴空病;多酰基辅酶A脱氢缺乏症;多发性皮肤和粘膜静脉曲张;1型多发性内分泌肿瘤;多种硫酸酯酶缺乏症;NAIC;指甲-髌骨综合征;线状体肌病;新生儿糖尿病;新生儿表面活性剂缺乏;肾炎;Netherton病;神经纤维瘤病;1型神经纤维瘤病;A型Niemann-Pick病;B型Niemann-Pick病;C型Niemann-Pick病;NKX2E;Noonan综合征;北美印第安儿童肝硬化;NR0B1重复相关DSD;眼部遗传病;眼耳综合征;OLEDAID;肾单位稀少巨大症;肾单位稀少巨大症肾发育不全;Ollier病;Opitz-Kaveggia综合征;1型口面部指趾综合征;2型口面部指趾综合征;骨性Paget病;2型耳腭指综合征;OXPHOS疾病;掌跖角化过度;肾脏细菌瘤;Parkes-Weber综合征;Parkinson综合征;21q22.2-q22.3部分缺失;Pearson综合征;Pelizaeus-Merzbacher病;Pendred综合征;Cantrell五联症;过氧化物酶体酰基辅酶A-氧化酶缺乏症;Peutz-Jeghers综合征;Pfeiffer综合征;Pierson综合征;色素性结节性肾上腺皮质疾病;哌啶酸血症;Pitt-Hopkins综合征;缩醛磷脂缺乏症;胸膜肺母细胞瘤和囊性肾瘤;多囊性脂膜性骨发育不良;卟啉症;卵巢早衰;原发性红斑肢痛症;原发性血色素沉着症;原发性高草酸尿症;进行性家族性肝内胆汁淤积症;丙酸血症;丙酮酸脱羧酶缺乏症;RAPADILINO综合征;肾胱氨酸病;横纹肌瘤易感综合征;Rieger综合征;环状染色体4;Roberts综合征;Robinow-Sorauf综合征;Rothmund-Thomson综合征;SCID;Saethre-Chotzen综合征;Sandhoff病;SC phocomelia综合征;SCAS;Schinzel phocomelia综合征;1型短肋多指综合征;4型短肋多指综合征;2型短肋多指综合征;3型短肋多指综合征;Shwachman病;Shwachman-Diamond病;镰状细胞性贫血;Silver-Russell综合征;Simpson-Golabi-Behmel综合征;Smith-Lemli-Opitz综合征;SPG7相关的遗传性痉挛性截瘫;球形红细胞增多症;手足畸形伴长骨缺陷;脊柱肋骨发育不良;散发性内脏肌病伴包涵体;贮积症;STRA6相关综合征;Tay-Sachs病;死因性发育不良;甲状腺代谢疾病;Tourette综合征;转甲状腺素蛋白相关淀粉样变性;13三体性;22三体性;三体性2p综合征;结节性硬化症;簇状肠病;尿素循环疾病;Van Den Ende-Gupta综合征;Van der Woude综合征;杂色镶嵌非整倍体综合征;VLCAD缺乏症;von Hippel-Lindau病;Waardenburg综合征;WAGR综合征;Walker-Warburg综合征;Werner综合征;Wilson病;Wolcott-Rallison综合征;Wolfram综合征;X连锁无丙种球蛋白血症;X连锁慢性特发性假性肠梗阻;X连锁腭裂伴舌炎;X连锁显性点状软骨发育不良;X连锁外胚层发育不良;X连锁Emery-Dreifuss肌营养不良症;X连锁无脑回畸形;X连锁淋巴组织增生性疾病;X连锁内脏异位性;1型黄嘌呤尿症;2型黄嘌呤尿症;着色性干皮病;XPV;以及Zellweger病。

[0107] 在一些实施方案中,治疗、预防或抑制受试者的病痛可以包括使与特定疾病相关的靶核酸与可编程切口酶(例如,Cas14a、Cas14b或Cas14e可编程切口酶)接触。在一些方面,治疗、预防或抑制疾病的方法可以包括移除、修饰、置换、转座或影响对其有需要的患者

的基因组序列的调节。在一些实施方案中,治疗、预防或抑制疾病的方法可以涉及调节基因表达。在一些实施方案中,治疗、预防或抑制疾病的方法可以包括将与病原体(例如病毒或细菌)相关的核酸序列靶向至本公开内容的可编程切口酶。

[0108] 本文所述的组合物和方法可用于治疗、预防、诊断或鉴别受试者的癌症。例如,包括通过使靶核酸与包含可编程切口酶的组合物接触将切口引入靶核酸的方法可用于治疗、预防、诊断或鉴别受试者的癌症。在一些方面,该方法可以靶向细胞或组织。在一些实施方案中,该方法可以应用于受试者,例如人。如本文所用,术语“癌症”是指特征在于异常或不受调节的细胞生长或活性的生理状况。在一些情况下,癌症可能涉及在受试者的各种组织之间表现出异常或不受调节的生长或活性的细胞的扩散。在一些方面,癌症可以是遗传病状。癌症的实例包括,但不限于急性淋巴细胞白血病、急性髓性白血病、肾上腺皮质癌、肛门癌、星形细胞瘤、胆管癌、膀胱癌、骨癌、脑癌、乳腺癌、支气管癌、Burkitt淋巴瘤、癌(Carcinoma)、心脏肿瘤、宫颈癌、脊索瘤、慢性淋巴细胞白血病、慢性髓细胞白血病、慢性骨髓增生性肿瘤、结肠癌、结直肠癌、颅咽管瘤、皮肤T细胞淋巴瘤、导管癌、胚胎肿瘤、子宫内膜癌、室管膜瘤、食道癌、上皮神经母细胞瘤、尤文肉瘤、颅外生殖细胞肿瘤、性腺外生殖细胞肿瘤、输卵管癌、纤维组织细胞瘤、胆囊癌、胃癌、胃肠道癌、胃肠道类癌、胃肠道间质瘤、妊娠期毛细胞瘤、颈细胞病癌症、心脏肿瘤、肝细胞癌、组织细胞增生症、霍奇金淋巴瘤、下咽癌、眼内黑色素瘤、胰岛细胞肿瘤、卡波西肉瘤、肾癌、朗格汉斯细胞组织细胞增生症、喉癌、白血病、唇和口腔癌、肝癌、肺癌、淋巴瘤、恶性纤维组织细胞瘤、黑色素瘤、Merkel细胞癌、转移性鳞状颈癌、中线癌、口腔癌、多发性内分泌肿瘤综合征、多发性骨髓瘤、蕈样真菌病、骨髓增生异常综合征、骨髓性白血病、髓样白血病、骨髓增生性肿瘤、鼻腔和鼻窦癌、鼻咽癌、神经母细胞瘤、非霍奇金淋巴瘤、非小细胞肺癌、口腔癌、骨肉瘤、卵巢癌、胰腺癌、胰腺神经内分泌肿瘤、乳头瘤病、副神经节瘤、鼻窦和鼻腔癌、甲状旁腺癌、阴茎癌、咽癌、嗜铬细胞瘤、垂体瘤、浆细胞肿瘤、胸膜肺母细胞瘤、原发性中枢神经系统(CNS)淋巴瘤、原发性腹膜癌、前列腺癌、直肠癌、复发性癌症、肾细胞癌、视网膜母细胞瘤、横纹肌肉瘤、唾液腺癌、Sézary综合征、皮肤癌、小细胞肺癌、小肠癌、软组织肉瘤、鳞状细胞癌、原发性隐匿性鳞状颈癌、胃癌、T细胞淋巴瘤、睾丸癌、喉癌、胸腺瘤和胸腺癌、甲状腺癌、气管支气管癌、肾盂和输尿管移行细胞癌、输尿管癌、肾盂癌、尿道癌、子宫癌、子宫肉瘤、阴道癌、血管瘤、外阴癌和Wilms肿瘤。

[0109] 在一些情况下,癌症与特定的生物标志物有关。生物标志物是指示可用作细胞或有机体状态(例如,疾病存在或不存在)的标志的化学种类或特征。生物标志物的非限制性实例包括生物分子、核酸序列、蛋白质、代谢物、核酸、蛋白质修饰。生物标志物可以指一个种类或多个种类,例如细胞表面特征。

[0110] 本公开内容的方法(例如,修饰靶核酸的方法)可以包括用可编程切口酶(例如,Cas14a、Cas14b或Cas14e可编程切口酶)靶向生物标志物或与生物标志物相关的核酸。在一些情况下,生物标志物是与癌症相关的基因。与癌症相关的基因的非限制性实例包括ABL、AF4/HRX、AKT-2、ALK、ALK/NPM、AML1、AML1/MTG8、APC、ATM、AXIN2、AXL、BAP1、BARD1、BCL-2、BCL-3、BCL-6、BCR/ABL、BLM、BMP1A、BRCA1、BRCA2、BRIP1、c-MYC、CASR、CDC73、CDH1、CDK4、CDKN1B、CDKN1C、CDKN2A、CEBPA、CHEK2、CTNNA1、DBL、DEK/CAN、DICER1、DIS3L2、E2A/PBX1、EGFR、ENL/HRX、EPCAM、ERG/TLS、ERBB、ERBB-2、ETS-1、EWS/FLI-1、FH、FLCN、FMS、FOS、FPS、

GATA2、GLI、GPGSP、GREM1、HER2/neu、HOX11、HOXB13、HST、IL-3、INT-2、JUN、KIT、KS3、K-SAM、LBC、LCK、LM01、LM02、L-MYC、LYL-1、LYT-10、LYT-10/C α 1、MAS、MAX、MDM-2、MEN1、MET、MITF、MLH1、MLL、MOS、MSH1、MSH2、MSH3、MSH6、MTG8/AML1、MUTYH、MYB、MYH11/CBFB、NBN、NEU、NF1、NF2、N-MYC、NTHL1、OST、PALB2、PAX-5、PBX1/E2A、PDGFRA、PHOX2B、PIM-1、PMS2、POLD1、POLE、POT1、PRAD-1、PRKAR1A、PTCH1、PTEN、RAD50、RAD51C、RAD51D、RAF、RAR/PML、RAS-H、RAS-K、RAS-N、RB1、RECQL4、REL/NRG、RET、RHOA、RHOB、ROS、RUNX1、SDHA、SDHAF、SDHB、SDHC、SDHD、SET/CAN、SIS、SKI、SMAD4、SMARCA4、SMARCB1、SMARCE1、SRC、STK11、SUFU、TAL1、TAL2、TAN-1、TIAM1、TERC、TERT、TMEM127、TP53、TSC1、TSC2、TRK、VHL、WRN和WT1。在一些情况下，癌症的基因生物标志物将携带一个或多个突变。在一些情况下，癌症的基因生物标志物相对于未患癌症的患者或样本将被上调或下调。

靶核酸的检测

[0111] 本公开内容提供了能够通过可编程切口酶平台(例如DNA核酸内切酶靶向CRISPR TransReporter (DETECTR) 平台)进行靶DNA检测的方法和组合物。在一些实施方案中,在分析报告子的切割之前扩增来自样品的靶DNA。靶DNA可以通过PCR或等温扩增技术进行扩增。与DETECTR技术兼容的DNA扩增方法可用于可编程核酸酶,例如可编程切口酶。例如,可以扩增ssDNA。ssDNA而非dsDNA的扩增使得通过具有dsDNA激活的反式切割所要求PAM的蛋白的不依赖PAM的核酸检测成为可能,与某些Cas14蛋白的情况一样。

[0112] 某些可编程核酸酶(例如,Cas14可编程切口酶)表现出无差别ssDNA反式切割,使其可用于检测样品中的DNA。在一些实施方案中,从多种核酸模板(RNA、ss/dsDNA)产生靶ssDNA以在DETECTR平台中实现FQ报告子的切割。某些可编程切口酶(例如Cas14a1)被ssDNA激活,在此基础上它们可以表现出ssDNA反式切割,并且由此可以用于切割DETECTR系统中的ssDNA FQ报告分子。这些可编程切口酶靶向样品中存在的ssDNA或从任意数量的核酸模板(RNA、ssDNA或dsDNA)生成和/或扩增。

[0113] 在一些实施方案中,本文公开的可编程切口酶与tracrRNA结合使用。在一些实施方案中,tracrRNA序列与SEQ ID NO:98-SEQ ID NO:99、SEQ ID NO:101、SEQ ID NO:103、SEQ ID NO:105-SEQ ID NO:151中的任一项包含至少70%、至少80%、至少90%、至少92%、至少95%、至少97%、或至少99%序列同一性。在一些实施方案中,tracrRNA序列与SEQ ID NO:99包含至少70%序列同一性。在一些实施方案中,tracrRNA序列与SEQ ID NO:99包含至少75%序列同一性。在一些实施方案中,tracrRNA序列与SEQ ID NO:99包含至少80%序列同一性。在一些实施方案中,tracrRNA序列与SEQ ID NO:99包含至少85%序列同一性。在一些实施方案中,tracrRNA序列与SEQ ID NO:99包含至少90%序列同一性。在一些实施方案中,tracrRNA序列与SEQ ID NO:99包含至少95%序列同一性。在一些实施方案中,与本公开内容的可编程切口酶复合使用的tracrRNA序列包含SEQ ID NO:99。在一些实施方案中,tracrRNA序列与SEQ ID NO:101包含至少70%序列同一性。在一些实施方案中,tracrRNA序列与SEQ ID NO:101包含至少75%序列同一性。在一些实施方案中,tracrRNA序列与SEQ ID NO:101包含至少80%序列同一性。在一些实施方案中,tracrRNA序列与SEQ ID NO:101包含至少85%序列同一性。在一些实施方案中,tracrRNA序列与SEQ ID NO:101包含至少90%序列同一性。在一些实施方案中,tracrRNA序列与SEQ ID NO:101包含至少95%序列同一性。在一些实施方案中,与本公开内容的可编程切口酶复合使用的tracrRNA序列包

含SEQ ID NO:101。在一些实施方案中, tracrRNA序列与SEQ ID NO:103包含至少70%序列同一性。在一些实施方案中, tracrRNA序列与SEQ ID NO:103包含至少75%序列同一性。在一些实施方案中, tracrRNA序列与SEQ ID NO:103包含至少80%序列同一性。在一些实施方案中, tracrRNA序列与SEQ ID NO:103包含至少85%序列同一性。在一些实施方案中, tracrRNA序列与SEQ ID NO:103包含至少90%序列同一性。在一些实施方案中, tracrRNA序列与SEQ ID NO:103至少95%序列同一性。在一些实施方案中, 与本公开内容的可编程切口酶复合使用的tracrRNA序列包含SEQ ID NO:103。在一些实施方案中, tracrRNA序列与SEQ ID NO:119包含至少70%序列同一性。在一些实施方案中, tracrRNA序列与SEQ ID NO:119包含至少75%序列同一性。在一些实施方案中, tracrRNA序列与SEQ ID NO:119包含至少80%序列同一性。在一些实施方案中, tracrRNA序列与SEQ ID NO:119包含至少85%序列同一性。在一些实施方案中, tracrRNA序列与SEQ ID NO:119包含至少90%序列同一性。在一些实施方案中, tracrRNA序列与SEQ ID NO:119包含至少95%序列同一性。在一些实施方案中, 与本公开内容的可编程切口酶复合使用的tracrRNA序列包含SEQ ID NO:119。

[0114] 在一些实施方案中, 可编程切口酶包含与SEQ ID NO:10具有70%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:10具有75%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:10具有80%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:10具有85%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:10具有90%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:10具有95%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:11具有70%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:11具有75%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:11具有80%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:11具有85%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:11具有90%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:11具有95%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:33具有70%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:33具有75%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:33具有80%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:33具有85%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:33具有90%同一性的序列。在一些实施方案中, 可编程切口酶包含与SEQ ID NO:33具有95%同一性的序列。

[0115] 本文公开的组合物、试剂盒和方法可以在测定靶核酸的方法中实施。在一些实施方案中, 测定样品中靶核酸的方法包括: 使样品与包含指导核酸的复合体接触, 该指导核酸包含与靶核酸的区段和可编程切口酶(例如, 本文公开的Cas14a、Cas14b或Cas14e)反向互补的区段, 该可编程切口酶在形成包含与靶核酸区段结合的指导核酸区段的复合体时表现出序列非依赖性切割, 其中样品包含至少一种核酸, 该至少一种核酸与靶核酸的区段包含至少50%序列同一性; 以及分析报告核酸群体中至少一种报告核酸的切割, 其中该切割表示样品中存在靶核酸, 其中不存在切割表示样品中不存在靶核酸。靶核酸可以占样品中总核酸的0.05%至20%。有时, 靶核酸占样品中总核酸的0.1%至10%。在一些情况下, 靶核酸占样品中总核酸的0.1%至5%。通常, 样品包含靶核酸的区段和至少一种核酸, 该至少一种

核酸与靶核酸的区段包含小于100%序列同一性但与靶核酸的区段包含不小于50%序列同一性。例如,与至少一种核酸相比,靶核酸的区段包含突变,该至少一种核酸与靶核酸的区段包含小于100%序列同一性但与靶核酸的区段包含不少于50%序列同一性。通常,与至少一种核酸相比,靶核酸的区段包含单核苷酸突变,该至少一种核酸与靶核酸的区段包含小于100%序列同一性但与靶核酸的区段包含不小于50%序列同一性。

[0116] 可编程切口酶DETECTR反应混合物中各种试剂的浓度可以根据反应的特定规模而变化。例如,可编程切口酶的终浓度可以在以下范围变化:1pM至1nM、1pM至10pM、10pM至100pM、100pM至1nM、1nM至10nM、10nM至20nM、20nM至30nM、30nM至40nM、40nM至50nM、50nM至60nM、60nM至70nM、70nM至80nM、80nM至90nM、90nM至100nM、100nM至200nM、200nM至300nM、300nM至400nM、400nM至500nM、500nM至600nM、600nM至700nM、700nM至800nM、800nM至900nM、900nM至1000nM。与靶核酸互补的sgRNA的终浓度可以为1pM至1nM、1pM至10pM、10pM至100pM、100pM至1nM、1nM至10nM、10nM至20nM、20nM至30nM、30nM至40nM、40nM至50nM、50nM至60nM、60nM至70nM、70nM至80nM、80nM至90nM、90nM至100nM、100nM至200nM、200nM至300nM、300nM至400nM、400nM至500nM、500nM至600nM、600nM至700nM、700nM至800nM、800nM至900nM、900nM至1000nM。ssDNA-FQ报告子的浓度可以为1pM至1nM、1pM至10pM、10pM至100pM、100pM至1nM、1nM至10nM、10nM至20nM、20nM至30nM、30nM至40nM、40nM至50nM、50nM至60nM、60nM至70nM、70nM至80nM、80nM至90nM、90nM至100nM、100nM至200nM、200nM至300nM、300nM至400nM、400nM至500nM、500nM至600nM、600nM至700nM、700nM至800nM、800nM至900nM、900nM至1000nM。

[0117] Cas14 DETECTR反应的实例由20 μ L总反应体积中的终浓度为100nM Cas14、125nM sgRNA和50nM ssDNA-FQ报告子组成。反应在荧光酶标仪(Tecan Infinite Pro 200M Plex)中在37℃下孵育2小时,每30秒进行一次荧光测量(例如, λ_{ex} :485nm; λ_{em} :535nm)。检测到的荧光波长可因报告分子而异。

[0118] 本文描述的是包含单链报告核酸的试剂,该单链报告核酸包含检测部分,其中在使用本文公开的方法从核酸模板产生和扩增ssDNA后报告核酸(例如,上述ssDNA-FQ报告子)能够被可编程切口酶切割,从而产生第一可检测信号。

[0119] 因此,本文公开的方法包括从样品中的目的靶核酸模板(例如,cDNA、ssDNA或dsDNA)产生和扩增ssDNA,使ssDNA与ssDNA激活的可编程切口酶孵育,导致报告核酸(也称为ssDNA-FQ报告子)的无差别不依赖PAM的切割,以产生可检测的信号,以及定量可检测信号以检测目的靶核酸序列。

报告子

[0120] 本文描述了包含报告子的试剂。报告子可以包含单链核酸和检测部分(例如,标记的单链DNA报告子),其中核酸能够被激活的可编程切口酶切割,释放检测部分,并产生可检测信号。如本文所用,“报告子(reporter)”与“报告核酸”或“报告分子(reporter molecule)”可互换使用。在指导RNA与靶核酸杂交时被激活的本文公开的可编程切口酶可以切割报告子。切割“报告子”在本文中可称为切割“报告核酸”、“报告分子”或“报告子的核酸”。

[0121] 本文公开的装置和方法的主要优势是针对未扩增或扩增的样品中的总核酸(不包括报告子的核酸)设计过量的报告子。总核酸可以包括靶核酸和非靶核酸,不报告报告子的

核酸。非靶核酸可以来自经裂解或未经裂解的原始样品。非靶核酸还可以是扩增的副产物。因此,非靶核酸可包括来自经裂解或未经裂解的原始样品以及来自扩增的样品的非靶核酸。由于存在大量非靶核酸,激活的可编程核酸酶的结合和切割报告子序列的能力可能受抑制。这是因为激活的可编程核酸酶并行地切割任何核酸。如果存在大量总核酸,它们对可编程核酸酶的竞争力可能大于报告子。本文公开的装置和方法被设计成相对于总核酸具有过量的报告子,使得来自切割反应(例如,DETECTR反应)的可检测信号是特别突出的。在一些实施方案中,相对于总核酸,报告子以至少1.5倍、至少2倍、至少3倍、至少4倍、至少5倍、至少6倍、至少7倍、至少8倍、至少9倍、至少10倍、至少11倍、至少12倍、至少13倍、至少14倍、至少15倍、至少16倍、至少17倍、至少18倍、至少19倍、至少20倍、至少30倍、至少40倍、至少50倍、至少60倍、至少70倍、至少80倍、至少90倍、至少100倍、1.5倍至100倍、2倍至10倍、10倍至20倍、20倍至30倍、30倍至40倍、40倍至50倍、50倍至60倍、60倍至70倍、70倍至80倍、80倍至90倍、90倍至100倍、1.5倍至10倍、1.5倍至20倍、10倍至40倍、20倍至60倍或10倍至80倍过量存在。

[0122] 本文公开的装置和方法的第二重要优势是设计了包含指导核酸、可编程核酸酶和报告子的过量体积,其接触包含具有感兴趣靶核酸的样品的较小体积。包含样品的较小体积可以是未裂解的样品、裂解的样品或已经历逆转录、扩增和体外转录的任意组合的裂解样品。粗样品、未裂解样品、裂解样品或裂解并扩增的样品中存在的各种试剂,如缓冲剂、硫酸镁、盐、pH、还原剂、引物、dNTP、NTP、细胞裂解物、非靶核酸、引物或其他组合,可以抑制可编程核酸酶发现并切割报告子的核酸的能力。这可能是由于并非报告子的核酸对可编程核酸酶的竞争力大于报告子的核酸。替代地,样品中的各种试剂可能直接抑制可编程核酸酶的活性。因此,本文提供的使包含指导核酸、可编程核酸酶和报告子的过量体积与包含具有感兴趣靶核酸的样品的较小体积接触的装置和方法通过确保可编程核酸酶能够找到并切割报告子的核酸来更好地检测靶核酸。在一些实施方案中,包含指导核酸、可编程核酸酶和报告子的体积(可以成为“第二体积”)比包含样品的体积(可以成为“第一体积”)大4倍。在一些实施方案中,包含指导核酸、可编程核酸酶和报告子的体积(可以成为“第二体积”)比包含样品的体积(可以成为“第一体积”)大至少1.5倍、至少2倍、至少3倍、至少4倍、至少5倍、至少6倍、至少7倍、至少8倍、至少9倍、至少10倍、至少11倍、至少12倍、至少13倍、至少14倍、至少15倍、至少16倍、至少17倍、至少18倍、至少19倍、至少20倍、至少30倍、至少40倍、至少50倍、至少60倍、至少70倍、至少80倍、至少90倍、至少100倍、1.5倍至100倍、2倍至10倍、10倍至20倍、20倍至30倍、30倍至40倍、40倍至50倍、50倍至60倍、60倍至70倍、70倍至80倍、80倍至90倍、90倍至100倍、1.5倍至10倍、1.5倍至20倍、10倍至40倍、20倍至60倍或10倍至80倍。在一些实施方案中,包含样品的体积是至少0.5 μ L、至少1 μ L、至少1 μ L、至少2 μ L、至少3 μ L、至少4 μ L、至少5 μ L、至少6 μ L、至少7 μ L、至少8 μ L、至少9 μ L、至少10 μ L、至少11 μ L、至少12 μ L、至少13 μ L、至少14 μ L、至少15 μ L、至少16 μ L、至少17 μ L、至少18 μ L、至少19 μ L、至少20 μ L、至少25 μ L、至少30 μ L、至少35 μ L、至少40 μ L、至少45 μ L、至少50 μ L、至少55 μ L、至少60 μ L、至少65 μ L、至少70 μ L、至少75 μ L、至少80 μ L、至少85 μ L、至少90 μ L、至少95 μ L、至少100 μ L、0.5 μ L至5 μ L、5 μ L至10 μ L、10 μ L至15 μ L、15 μ L至20 μ L、20 μ L至25 μ L、25 μ L至30 μ L、30 μ L至35 μ L、35 μ L至40 μ L、40 μ L至45 μ L、45 μ L至50 μ L、10 μ L至20 μ L、5 μ L至20 μ L、1 μ L至40 μ L、2 μ L至10 μ L或1 μ L至10 μ L。在一些实施方案中,包含可编程核酸酶、指导核酸和报告子的体积是至少10 μ L、至少11 μ L、

至少12μL、至少13μL、至少14μL、至少15μL、至少16μL、至少17μL、至少18μL、至少19μL、至少20μL、至少21μL、至少22μL、至少23μL、至少24μL、至少25μL、至少26μL、至少27μL、至少28μL、至少29μL、至少30μL、至少40μL、至少50μL、至少60μL、至少70μL、至少80μL、至少90μL、至少100μL、至少150μL、至少200μL、至少250μL、至少300μL、至少350μL、至少400μL、至少450μL、至少500μL、10μL至15μL、15μL至20μL、20μL至25μL、25μL至30μL、30μL至35μL、35μL至40μL、40μL至45μL、45μL至50μL、50μL至55μL、55μL至60μL、60μL至65μL、65μL至70μL、70μL至75μL、75μL至80μL、80μL至85μL、85μL至90μL、90μL至95μL、95μL至100μL、100μL至150μL、150μL至200μL、200μL至250μL、250μL至300μL、300μL至350μL、350μL至400μL、400μL至450μL、450μL至500μL、10μL至20μL、10μL至30μL、25μL至35μL、10μL至40μL、20μL至50μL、18μL至28μL或17μL至22μL。

[0123] 在一些情况下,报告核酸是包含脱氧核糖核苷酸的单链核酸序列。在其他情况下,报告核酸是包含核糖核苷酸的单链核酸序列。报告子的核酸可以是包含至少一种脱氧核糖核苷酸和至少一种核糖核苷酸的单链核酸序列。在一些情况下,报告子的核酸是单链核酸,其在作为切割位点的内部位置包含至少一个核糖核苷酸残基。在一些情况下,报告子的核酸在内部位置包含至少2、3、4、5、6、7、8、9或10个核糖核苷酸残基。在一些情况下,报告子的核酸在内部位置包含2至10、3至9、4至8、或5至7个核糖核苷酸残基。有时核糖核苷酸残基是连续的。可选地,核糖核苷酸残基散布在非核糖核苷酸残基之间。在一些情况下,报告子的核酸仅具有核糖核苷酸残基。在一些情况下,报告子的核酸仅具有脱氧核糖核苷酸残基。在一些情况下,核酸包含对本文所述的可编程切口酶的切割具有抗性的核苷酸。在一些情况下,报告子的核酸包含合成核苷酸。在一些情况下,报告子的核酸包含至少一个核糖核苷酸残基和至少一个非核糖核苷酸残基。在一些情况下,报告子的核酸长度为5-20、5-15、5-10、7-20、7-15或7-10个核苷酸。在一些情况下,报告子的核酸长度为3至20、4至10、5至10、或5至8个核苷酸。在一些情况下,报告子的核酸包含至少一个尿嘧啶核糖核苷酸。在一些情况下,报告子的核酸包含至少两个尿嘧啶核糖核苷酸。有时报告子的核酸仅具有尿嘧啶核糖核苷酸。在一些情况下,报告子的核酸包含至少一种腺嘌呤核糖核苷酸。在一些情况下,报告子的核酸包含至少两个腺嘌呤核糖核苷酸。在一些情况下,报告子的核酸仅具有腺嘌呤核糖核苷酸。在一些情况下,报告子的核酸包含至少一个胞嘧啶核糖核苷酸。在一些情况下,报告子的核酸包含至少两个胞嘧啶核糖核苷酸。在一些情况下,报告子的核酸包含至少一个鸟嘌呤核糖核苷酸。在一些情况下,报告子的核酸包含至少两个鸟嘌呤核糖核苷酸。报告子的核酸可以仅包含未修饰的核糖核苷酸、仅包含未修饰的脱氧核糖核苷酸或其组合。在一些情况下,报告子的核酸长度为5至12个核苷酸。在一些情况下,报告核酸的长度为至少2、至少3、至少4、至少5、至少6、至少7、至少8、至少9、至少10、至少11、至少12、至少13、至少14、至少15、至少16、至少17、至少18、至少19、至少20、至少21、至少22、至少23、至少24、至少25、至少26、至少27、至少28、至少29或至少30个核苷酸。在一些情况下,报告核酸的长度为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个核苷酸。

[0124] 报告子的单链核酸包含能够产生第一可检测信号的检测部分。有时,报告核酸包含能够产生信号的蛋白。信号可以是量热、电位、电流、光学(例如荧光、比色等)或压电信号。在一些情况下,检测部分在切割位点的一侧。任选地,猝灭部分在切割位点的另一侧。猝

灭部分有时是荧光猝灭部分。在一些情况下,猝灭部分在切割位点的5'处,而检测部分在切割位点的3'处。在一些情况下,检测部分在切割位点的5'处,而猝灭部分在切割位点的3'处。猝灭部分有时在检测核酸的5'末端。检测部分有时在检测核酸的3'末端。在一些情况下,检测部分在检测核酸的5'末端。在一些情况下,猝灭部分位于报告子核酸的3'末端。在一些情况下,报告子的单链核酸是能够产生第一可检测信号的至少一个单链核酸群体。在一些情况下,报告子的单链核酸是能够产生第一可检测信号的单链核酸群体。任选地,存在多于一个报告子的单链核酸群体。在一些情况下,存在2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、30、40、50、或大于50或任何数量的涵盖能够产生可检测信号的报告子的不同单链核酸群体的列表的范围。在一些情况下,能够产生可检测信号的报告子有2至50、3至40、4至30、5至20、或6至10个不同的单链核酸群体。

表2-报告子中单链核酸的实例

5'检测部分*	序列 (SEQ ID NO)	3'猝灭剂*
/56-FAM/	rUrUrUrUrU (SEQ ID NO: 153)	/3IABkFQ/
/5IRD700/	rUrUrUrUrU (SEQ ID NO: 153)	/3IRQC1N/
/5TYE665/	rUrUrUrUrU (SEQ ID NO: 153)	/3IAbRQSp/

5'检测部分*	序列 (SEQ ID NO)	3'猝灭剂*
/5Alex594N/	rUrUrUrUrU (SEQ ID NO: 153)	/3IAbRQSp/
/5ATTO633N/	rUrUrUrUrU (SEQ ID NO: 153)	/3IAbRQSp/
/56-FAM/	rUrUrUrUrUrUrUrU (SEQ ID NO: 154)	/3IABkFQ/
/5IRD700/	rUrUrUrUrUrUrUrU (SEQ ID NO: 154)	/3IRQC1N/
/5TYE665/	rUrUrUrUrUrUrUrU (SEQ ID NO: 154)	/3IAbRQSp/
/5Alex594N/	rUrUrUrUrUrUrUrU (SEQ ID NO: 154)	/3IAbRQSp/
/5ATTO633N/	rUrUrUrUrUrUrUrU (SEQ ID NO: 154)	/3IAbRQSp/
/56-FAM/	rUrUrUrUrUrUrUrUrU (SEQ ID NO: 155)	/3IABkFQ/
/5IRD700/	rUrUrUrUrUrUrUrUrU (SEQ ID NO: 155)	/3IRQC1N/
/5TYE665/	rUrUrUrUrUrUrUrUrU (SEQ ID NO: 155)	/3IAbRQSp/
/5Alex594N/	rUrUrUrUrUrUrUrUrU (SEQ ID NO: 155)	/3IAbRQSp/
/5ATTO633N/	rUrUrUrUrUrUrUrUrU (SEQ ID NO: 155)	/3IAbRQSp/
/56-FAM/	TTTTTrUrUTTTT (SEQ ID NO: 156)	/3IABkFQ/
/5IRD700/	TTTTTrUrUTTTT (SEQ ID NO: 156)	/3IRQC1N/
/5TYE665/	TTTTTrUrUTTTT (SEQ ID NO: 156)	/3IAbRQSp/
/5Alex594N/	TTTTTrUrUTTTT (SEQ ID NO: 156)	/3IAbRQSp/
/5ATTO633N/	TTTTTrUrUTTTT (SEQ ID NO: 156)	/3IAbRQSp/
/56-FAM/	TTrUrUTT (SEQ ID NO: 157)	/3IABkFQ/
/5IRD700/	TTrUrUTT (SEQ ID NO: 157)	/3IRQC1N/
/5TYE665/	TTrUrUTT (SEQ ID NO: 157)	/3IAbRQSp/
/5Alex594N/	TTrUrUTT (SEQ ID NO: 157)	/3IAbRQSp/
/5ATTO633N/	TTrUrUTT (SEQ ID NO: 157)	/3IAbRQSp/
/56-FAM/	TArArUGC (SEQ ID NO: 158)	/3IABkFQ/
/5IRD700/	TArArUGC (SEQ ID NO: 158)	/3IRQC1N/
/5TYE665/	TArArUGC (SEQ ID NO: 158)	/3IAbRQSp/
/5Alex594N/	TArArUGC (SEQ ID NO: 158)	/3IAbRQSp/
/5ATTO633N/	TArArUGC (SEQ ID NO: 158)	/3IAbRQSp/
/56-FAM/	TArUrGGC (SEQ ID NO: 159)	/3IABkFQ/
/5IRD700/	TArUrGGC (SEQ ID NO: 159)	/3IRQC1N/
/5TYE665/	TArUrGGC (SEQ ID NO: 159)	/3IAbRQSp/
/5Alex594N/	TArUrGGC (SEQ ID NO: 159)	/3IAbRQSp/
/5ATTO633N/	TArUrGGC (SEQ ID NO: 159)	/3IAbRQSp/
/56-FAM/	rUrUrUrUrU (SEQ ID NO: 160)	/3IABkFQ/
/5IRD700/	rUrUrUrUrU (SEQ ID NO: 160)	/3IRQC1N/
/5TYE665/	rUrUrUrUrU (SEQ ID NO: 160)	/3IAbRQSp/
/5Alex594N/	rUrUrUrUrU (SEQ ID NO: 160)	/3IAbRQSp/
/5ATTO633N/	rUrUrUrUrU (SEQ ID NO: 160)	/3IAbRQSp/
/56-FAM/	TTATTATT (SEQ ID NO: 161)	/3IABkFQ/
/56-FAM/	TTATTATT (SEQ ID NO: 161)	/3IABkFQ/
/5IRD700/	TTATTATT (SEQ ID NO: 161)	/3IRQC1N/
/5TYE665/	TTATTATT (SEQ ID NO: 161)	/3IAbRQSp/
/5Alex594N/	TTATTATT (SEQ ID NO: 161)	/3IAbRQSp/

5'检测部分*	序列 (SEQ ID NO)	3'猝灭剂*
/5ATTO633N/	TTATTATT (SEQ ID NO: 161)	/3IAbRQSp/
/56-FAM/	TTTTTT (SEQ ID NO: 162)	/3IABkFQ/
/56-FAM/	TTTTTTTTT (SEQ ID NO: 163)	/3IABkFQ/
/56-FAM/	TTTTTTTTTTT (SEQ ID NO: 164)	/3IABkFQ/
/56-FAM/	TTTTTTTTTTTTT (SEQ ID NO: 165)	/3IABkFQ/
/56-FAM/	TTTTTTTTTTTTTTT (SEQ ID NO: 166)	/3IABkFQ/
/56-FAM/	AAAAAA (SEQ ID NO: 167)	/3IABkFQ/
/56-FAM/	CCCCC (SEQ ID NO: 168)	/3IABkFQ/
/56-FAM/	GGGGG (SEQ ID NO: 169)	/3IABkFQ/
/56-FAM/	TTATTATT (SEQ ID NO: 161)	/3IABkFQ/

/56-FAM/:5'-荧光素(集成DNA技术公司)

/3IABkFQ/:3'Iowa Black FQ(集成DNA技术公司)

/5IRD700/:5'IRDye 700(集成DNA技术公司)

/5TYE665/:5'TYE 665(集成DNA技术公司)

/5Alex594N/:5'Alexa Fluor 594(NHS酯)(集成DNA技术公司)

/5ATTO633N/:5'ATTO TM 633(NHS酯)(集成DNA技术公司)

/3IRQC1N/:3'IRDye QC-1Quencher(Li-Cor)

/3IAbRQSp/:3'Iowa Black RQ(集成DNA技术公司)

rU:尿嘧啶核糖核苷酸

rG:鸟嘌呤核糖核苷酸

*该表格将检测部分和猝灭部分称为其商品名,并标识了其来源。但是,也可以使用其他来源中具有相似功能的替代物、仿制品或非商标部分。

[0125] 检测部分可以是近红外荧光团。检测部分可以是发射500nm至720nm范围内的荧光的荧光团。检测部分可以是发射500nm至720nm范围内的荧光的荧光团。在一些情况下,检测部分发射波长为700nm或更高的荧光。在其他情况下,检测部分发射约660nm或约670nm的荧光。在一些情况下,检测部分发射500至520、500至540、500至590、590至600、600至610、610至620、620至630、630至640、640至650、650至660、660至670、670至680、680至690、690至700、700至710、710至720或720至730nm范围内的荧光。在一些情况下,检测部分发射450nm至750nm、500nm至650nm或550nm至650nm范围内的荧光。检测部分可以是发射与6-荧光素、IRDye 700、TYE 665、Alexa Fluor或ATTO TM 633(NHS酯)在相同范围内的荧光的荧光团。检测部分可以是荧光素亚酰胺、6-荧光素、IRDye 700、TYE 665、Alexa Fluor 594或ATTO TM 633(NHS酯)。检测部分可以是发射与6-荧光素(集成DNA技术公司)、IRDye 700(集成DNA技术公司)、TYE665(集成DNA技术公司)、Alexa Fluor 594(集成DNA技术公司)或ATTO TM 633(NHS酯)(集成DNA技术公司)在相同范围内的荧光的荧光团。检测部分可以是荧光素亚酰胺、6-荧光素(集成DNA技术公司)、洋地黄毒苷、IRDye 700(集成DNA技术公司)、TYE 665(集成DNA技术公司)、Alexa Fluor 594(集成DNA技术公司)或ATTO TM 633(NHS酯)(集成DNA技术公司)。本文所述的任何检测部分均可以来自任何市售来源,可以是具有所列出的检测部分的相似功能、通用名或非商品名的替代物。

[0126] 检测部分可以基于待测试的样品类型而选用。例如,为近红外荧光团的检测部分与尿液样品一起使用。作为另一示例,具有发射520nm左右的荧光的荧光团的SEQ ID NO:

153被用于测试非尿液样品,而具有具有发射700nm左右的荧光的荧光团的SEQ ID NO:160被用于测试尿液样品。

[0127] 猝灭部分可以基于其猝灭检测部分的能力来选择。猝灭部分可以是非荧光的荧光猝灭剂。猝灭部分可以猝灭发射500nm至720nm范围内的荧光的检测部分。猝灭部分可以猝灭发射500nm至720nm范围内的荧光的检测部分。在一些情况下,猝灭部分猝灭发射波长为700nm或更高的荧光的检测部分。在其他情况下,猝灭部分猝灭发射约660nm或约670nm的荧光的检测部分。在一些情况下,猝灭部分猝灭发射500至520、500至540、500至590、590至600、600至610、610至620、620至630、630至640、640至650、650至660、660至670、670至680、680至690、690至700、700至710、710至720或720至730nm范围内的荧光的检测部分。在一些情况下,猝灭部分猝灭发射450nm至750nm、500nm至650nm或550nm至650nm范围内的荧光的检测部分。猝灭部分可以猝灭荧光素亚酰胺、6-荧光素、IRDye 700、TYE 665、Alexa Fluor 594或ATTO TM 633(NHS酯)。猝灭部分可以是Iowa Black RQ、Iowa Black FQ或IRDye QC-1Quencher。猝灭部分可以猝灭荧光素亚酰胺、6-荧光素(集成DNA技术公司)、IRDye 700(集成DNA技术公司)、TYE 665(集成DNA技术公司)、Alexa Fluor 594(集成DNA技术公司)或ATTO TM 633(NHS酯)(集成DNA技术公司)。猝灭部分可以是Iowa Black RQ(集成DNA技术公司)、Iowa Black FQ(集成DNA技术公司)或IRDye QC-1猝灭剂(LiCor)。本文所述的任何猝灭部分均可以来自任何市售来源,可以是具有所列出的猝灭部分的相似功能、通用或非商品名的替代物。

[0128] 由检测部分的释放而产生的可检测信号表明,由可编程核酸酶的切割已经发生并且样品含有靶核酸。在一些情况下,检测部分包含荧光染料。检测部分有时包含荧光共振能量转移(FRET)对。在一些情况下,检测部分包含近红外(IR)染料。在一些情况下,检测部分包含紫外(UV)染料。替代地或组合地,检测部分包含多肽。检测部分有时包含生物素。检测部分有时包含抗生物素蛋白或链霉亲和素的至少一种。在一些情况下,检测部分包含多糖、聚合物或纳米颗粒。在一些情况下,检测部分包含金纳米颗粒或乳胶纳米颗粒。

[0129] 检测部分可以是能够产生量热、电位、电流、光学(例如,荧光、比色等)或压电信号的任何部分。有时,报告子的核酸是蛋白质-核酸,其能够在核酸裂解时产生量热、电位、电流、光学(例如,荧光、比色等)或压电信号。量热信号通常是在报告子的核酸切割后产生的热量。有时,热量信号是在报告子的核酸切割后吸收的热。例如,电位信号是在报告子的核酸切割后产生的电势。安培计信号可以是报告子的核酸切割后产生的电子移动。通常,信号是光信号,例如比色信号或荧光信号。光信号是例如在报告子的核酸裂解后产生的光输出。有时,光信号是报告子的核酸切割前后光吸收的变化。通常,压电信号是报告子的核酸切割前后之间的质量变化。

[0130] 可检测信号可以是比色信号或肉眼可见的信号。在一些情况下,可检测信号可以是荧光的、电的、化学的、电化学的或磁性的。在一些情况下,第一检测信号可以通过检测部分与检测区域中的捕获分子的结合产生,其中第一检测信号表明样品含有靶核酸。有时该系统能够检测多于一种类型的靶核酸,其中该系统包含多于一种类型的指导核酸和多于一种类型的报告核酸。在一些情况下,可检测信号可以由切割事件直接产生。可选地或组合地,可检测信号可以由信号事件间接产生。有时可检测信号不是荧光信号。在一些情况下,可检测信号可以是比色或基于颜色的信号。在一些情况下,可基于其在支持介质的检测区

域上的空间位置来识别检测到的靶核酸。在一些情况下,第二可检测信号可以在与第一产生的信号不同的空间位置产生。

[0131] 蛋白质-核酸通常是酶-核酸。酶在存在于酶-核酸中时可能在空间上受阻,而在从核酸切割后具有功能性。酶与底物产生反应。酶可以是转化酶。通常,转化酶的底物是蔗糖,并且DNS试剂可用于监测转化酶活性。在一些情况下,优选使用异双功能接头经由磺基SMCC化学将核酸(例如,DNA)与转化酶缀合。蛋白质-核酸有时是底物-核酸。底物通常是与酶产生反应的底物。

[0132] 蛋白质-核酸可以连接至固体支持物上。固体支持物可以是表面。表面可以是电极。固体支持物有时是珠粒。珠粒通常是磁珠。切割时,蛋白质从固体释放并且与其他混合物相互作用。例如,蛋白质是酶,并且在酶-核酸的核酸被切割后,酶流过腔室进入包含底物的混合物中。当酶与酶底物相遇时,发生反应,例如比色反应,然后检测该反应。作为另一示例,蛋白质是酶底物,并且在酶底物-核酸的核酸被切割后,酶流过腔室进入包含酶的混合物中。当酶底物与酶相遇时,发生反应,例如量热反应,然后检测该反应。

[0133] 通常,信号是比色信号或肉眼可见的信号。在一些情况下,信号是荧光、电、化学、电化学或磁信号。信号可以是量热、电位、安培、光学(例如荧光,比色等)或压电信号。在一些情况下,可检测信号是比色信号或肉眼可见的信号。在一些情况下,可检测信号是荧光、电、化学、电化学或磁信号。在一些情况下,第一检测信号由检测部分与检测区域中的捕获分子的结合产生,其中该第一检测信号表明样品包含靶核酸。有时该系统能够检测超过一种类型的靶核酸,并且该系统包含超过一种类型的指导核酸和超过一种类型的检测核酸。在一些情况下,可检测信号由切割事件直接产生。替代地或组合地,可检测信号由信号事件间接产生。可检测信号有时不是荧光信号。在一些情况下,可检测信号是比色信号或基于颜色的信号。在一些情况下,检测的靶核酸是基于其在支持介质的检测区域上的空间位置而鉴定的。在一些情况下,第二可检测信号与产生的第一信号在空间上不同的位置产生。

[0134] 在一些情况下,对于检测样品中单链靶核酸的主题方法,检测阈值小于或等于10nM。本文中术语“检测阈值”用于描述样品中为了进行检测而必须存在的最小量的靶核酸。例如,当检测阈值为10nM时,当靶核酸以10nM或更高的浓度存在于样品中时,可以检测到信号。在一些情况下,检测阈值小于或等于5nM、1nM、0.5nM、0.1nM、0.05nM、0.01nM、0.005nM、0.001nM、0.0005nM、0.0001nM、0.00005nM、0.00001nM、10pM、1pM、500fM、250fM、100fM、50fM、10fM、5fM、1fM、500阿摩尔(attomolar,aM)、100aM、50aM、10aM或1aM。在一些情况下,检测阈值范围为1aM至1nM、1aM至500pM、1aM至200pM、1aM至100pM、1aM至10pM、1aM至1pM、1aM至500fM、1aM至100fM、1aM至1fM、1aM至500aM、1aM至100aM、1aM至50aM、1aM至10aM、10aM至1nM、10aM至500pM、10aM至200pM、10aM至100pM、10aM至10pM、10aM至1pM、10aM至500fM、10aM至100fM、10aM至1fM、10aM至500aM、10aM至100aM、10aM至50aM、100aM至1nM、100aM至500pM、100aM至200pM、100aM至100pM、100aM至10pM、100aM至1pM、100aM至500fM、100aM至100fM、100aM至1fM、100aM至500aM、500aM至1nM、500aM至500pM、500aM至200pM、500aM至100pM、500aM至10pM、500aM至1pM、500aM至500fM、500aM至100fM、500aM至1fM、1fM至1nM、1fM至500pM、1fM至200pM、1fM至100pM、1fM至10pM、1fM至1pM、10fM至1nM、10fM至500pM、10fM至200pM、10fM至100pM、10fM至10pM、10fM至1pM、500fM至1nM、500fM至500pM、500fM至200pM、500fM至100pM、500fM至10pM、500fM至1pM、800fM至1nM、800fM至500pM、

800fM至200pM、800fM至100pM、800fM至10pM、800fM至1pM、1pM至1nM、1pM至500pM、1pM至200pM、1pM至100pM或1pM至10pM。在一些情况下,检测阈值范围为800fM至100pM、1pM至10pM、10fM至500fM、10fM至50fM、50fM至100fM、100fM至250fM或250fM至500fM。在一些情况下,检测阈值范围为2aM至100pM、20aM至50pM、50aM至20pM、200aM至5pM或500aM至2pM。在一些情况下,样本中检测到单链靶核酸的最小浓度的范围为1aM至1nM、10aM至1nM、100aM至1nM、500aM至1nM、1fM至1nM、1fM至500pM、1fM至200pM、1fM至100pM、1fM至10pM、1fM至1pM、10fM至1nM、10fM至500pM、10fM至200pM、10fM至100pM、10fM至10pM、10fM至1pM、500fM至1nM、500fM至500pM、500fM至200pM、500fM至100pM、500fM至10pM、500fM至1pM、800fM至1nM、800fM至500pM、800fM至200pM、800fM至100pM、800fM至10pM、800fM至1pM、1pM至1nM、1pM至500pM、1pM至200pM、1pM至100pM或1pM至10pM。在一些情况下,样本中检测到单链靶核酸的最小浓度的范围为2aM至100pM、20aM至50pM、50aM至20pM、200aM至5pM或500aM至2pM。在一些情况下,样品中可检测到单链靶核酸的最小浓度的范围为1aM至100pM。在一些情况下,样品中可检测到单链靶核酸的最小浓度的范围为1fM至100pM。在一些情况下,样品中可检测到单链靶核酸的最小浓度的范围为10fM至100pM。在一些情况下,样品中可检测到单链靶核酸的最小浓度的范围为800fM至100pM。在一些情况下,样品中可检测到单链靶核酸的最小浓度的范围为1pM至10pM。在一些情况下,本文所述的装置、系统、流体装置、试剂盒和方法检测包括多种核酸(例如多种非靶核酸)的样品中的靶单链核酸,其中靶单链核酸的浓度低至1aM、10aM、100aM、500aM、1fM、10fM、500fM、800fM、1pM、10pM、100pM或1pM。

[0135] 在一些实施方案中,在切割反应中靶核酸以约10nM、约20nM、约30nM、约40nM、约50nM、约60nM、约70nM、约80nM、约90nM、约100nM、约200nM、约300nM、约400nM、约500nM、约600nM、约700nM、约800nM、约900nM、约1 μ M、约10 μ M、或约100 μ M的浓度存在。在一些实施方案中,在切割反应中靶核酸以10nM至20nM、20nM至30nM、30nM至40nM、40nM至50nM、50nM至60nM、60nM至70nM、70nM至80nM、80nM至90nM、90nM至100nM、100nM至200nM、200nM至300nM、300nM至400nM、400nM至500nM、500nM至600nM、600nM至700nM、700nM至800nM、800nM至900nM、900nM至1 μ M、1 μ M至10 μ M、10 μ M至100 μ M、10nM至100nM、10nM至1 μ M、10nM至10 μ M、10nM至100 μ M、100nM至1 μ M、100nM至10 μ M、100nM至100 μ M、或1 μ M至100 μ M的浓度存在。在一些实施方案中,在切割反应中靶核酸以20nM至50 μ M、50nM至20 μ M、或200nM至5 μ M的浓度存在。

[0136] 在一些情况下,本文所述的装置、系统、流体装置、试剂盒和方法检测样品中的靶单链核酸,其中样品与试剂接触预定时间长度,该预定时间长度足以发生反式切割或切割反应完成。在一些情况下,本文所述的装置、系统、流体装置、试剂盒和方法检测样品中的靶单链核酸,其中样品与试剂接触不超过60分钟。有时,样品与试剂接触不超过120分钟、110分钟、100分钟、90分钟、80分钟、70分钟、60分钟、55分钟、50分钟、45分钟、40分钟、35分钟、30分钟、25分钟、20分钟、15分钟、10分钟、5分钟、4分钟、3分钟、2分钟或1分钟。有时,样品与试剂接触至少120分钟、110分钟、100分钟、90分钟、80分钟、70分钟、60分钟、55分钟、50分钟、45分钟、40分钟、35分钟、30分钟、25分钟、20分钟、15分钟、10分钟、5分钟。在一些情况下,样品与试剂接触5分钟至120分钟、5分钟至100分钟、10分钟至90分钟、15分钟至45分钟或20分钟至35分钟。在一些情况下,本文所述的装置、系统、流体装置、试剂盒和方法可以在小于10小时、小于9小时、小于8小时、小于7小时、小于6小时、小于5小时、小于4小时、小于3小时、小于2小时、小于1小时、小于50分钟、小于45分钟、小于40分钟、小于35分钟、小于30分

钟、小于25分钟、小于20分钟、小于15分钟、小于10分钟、小于9分钟、小于8分钟、小于7分钟、小于6分钟、或小于5分钟内检测到样品中的靶核酸。在一些情况下,本文所述的装置、系统、流体装置、试剂盒和方法可以在5分钟至10小时、10分钟至8小时、15分钟至6小时、20分钟至5小时、30分钟至2小时、或45分钟至1小时内检测到样品中的靶核酸。

[0137] 当指导核酸与靶核酸结合时,可以启动可编程切口酶的反式切割活性,并且可以切割报告子的核酸,从而导致检测到荧光。指导核酸可以是非天然存在的指导核酸。非天然存在的指导核酸可以包含具有与目的靶核酸序列杂交的重复区和间隔区的工程化序列。可以重组表达或化学合成非天然存在的指导核酸。核酸报告子可以包含检测部分,其中核酸报告子可以被激活的可编程切口酶切割,从而产生信号。如本文所述的一些方法可以是测定样品中靶核酸的方法,该方法包括使样品与包含指导核酸和可编程切口酶的复合体接触,该指导核酸包含与靶核酸的区段反向互补的区段,该可编程切口酶在形成包含与靶核酸片段结合的指导核酸片段的复合体后表现出序列非依赖性切割;以及测定指示蛋白质-核酸群体的至少一些蛋白质-核酸的切割的信号,其中该信号指示样品中存在靶核酸并且其中不存在该信号指示样品中不存在靶核酸。使用可编程切口酶切割报告子的核酸可以以50%的效率切割,如通过信号变化所测量的,信号是量热、电位、电流、光学(例如荧光、比色等)或压电,作为非限制性实例。如本文所述的一些方法可以是检测样品中靶核酸的方法,包括使包含靶核酸的样品与以下接触:靶向靶核酸区段的指导核酸、当与指导核酸和靶核酸区段复合时能够被激活的可编程切口酶、包含检测部分的报告子的单链核酸,其中报告子的核酸能够被激活的可编程切口酶切割,从而产生第一可检测信号;使用通过颜色变化测量切割的可编程切口酶切割报告子的单链核酸,并且测量支持介质上的第一可检测信号。使用可编程切口酶切割报告子的单链核酸可以以50%的效率切割,如通过颜色变化所测量的。在一些情况下,通过颜色变化测量的切割效率为至少40%、50%、60%、70%、80%、90%或95%。颜色变化可以是可检测的比色信号或肉眼可见的信号。颜色的变化可以作为第一可检测信号进行测量。第一可检测信号可以在使包含靶核酸的样品与靶向靶核酸区段的指导核酸、当与指导核酸和靶核酸复合时能够被激活的可编程切口酶部分、以及包含检测部分的报告子的单链核酸接触后5分钟内可检测到,其中报告子的核酸能够被激活的核酸酶切割。第一可检测信号可以在与样品接触的1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、25、30、35、40、45、50、55、60、70、80、90、100、110或120分钟内可检测到。在一些实施方案中,第一可检测信号可在与样品接触的1至120、5至100、10至90、15至80、20至60、或30至45分钟内可检测到。

[0138] 在一些情况下,本文描述的方法、试剂、酶和试剂盒在样品中用可编程切口酶和报告子的单链核酸检测靶单链核酸,其中样品与试剂接触足以反式切割报告子的单链核酸的预定时长。

[0139] 如本文所述的一些方法可以是检测样品中靶核酸的方法,包括使包含靶核酸的样品与靶向靶序列的指导核酸、当与指导核酸和靶序列复合时能够被激活的可编程切口酶、包含检测部分的单链报告核酸接触,其中报告核酸能够被激活的核酸酶切割,从而产生第一可检测信号;使用通过颜色变化测量的可编程切口酶切割单链报告核酸,并且测量支持介质上的第一可检测信号。如通过颜色变化所测量的,使用可编程切口酶切割单链报告核酸可以以50%的效率切割。在一些情况下,通过颜色变化测量的切割效率为至少40%、至少

50%、至少60%、至少70%、至少80%、至少90%、或至少95%。颜色变化可以是可检测的比色信号或肉眼可见的信号。颜色的变化可以作为第一可检测信号进行测量。第一可检测信号可以在使包含靶核酸的样品与靶向靶序列的指导核酸、当与指导核酸和靶序列复合时能够被激活的可编程切口酶和包含检测部分的单链报告核酸接触后5分钟内可检测到,其中报告核酸能够被激活的核酸酶切割。第一可检测信号可以在与样品接触的1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、25、30、35、40、45、50、55、60、70、80、90、100、110或120分钟内可检测到。

复用可编程核酸酶和可编程切口酶

[0140] 本文描述了包含与指导核酸和靶核酸分子复合时能够被激活的可编程核酸酶的试剂。此外,这些试剂可以与不同类型的可编程核酸酶一起使用,例如,用于复用可编程核酸酶。在一些实施方案中,可编程切口酶(例如,Cas14可编程切口酶)可以与另外的可编程核酸酶复用。例如,Cas14可编程切口酶可以与另外的可编程核酸酶复用以修饰或检测靶核酸。在一些实施方案中,可编程切口酶可以是Cas14a可编程切口酶、Cas14b可编程切口酶、Cas14c可编程切口酶、Cas14d可编程切口酶或Cas14e可编程切口酶。在一些实施方案中,第一可编程切口酶(例如,Cas14可编程切口酶)可以与第二可编程切口酶复用。

[0141] 在一些实施方案中,用于复用的另外的可编程核酸酶是任何可编程核酸酶。有时,可编程核酸酶是任何Cas蛋白(在本文中也称为Cas核酸酶)。在一些情况下,可编程核酸酶是Cas13。在一些实施方案中,Cas13是Cas13a、Cas13b、Cas13c、Cas13d或Cas13e。在一些情况下,可编程核酸酶可以是Mad7或Mad2。在一些情况下,可编程核酸酶是Cas12。有时Cas12是Cas12a、Cas12b、Cas12c、Cas12d或Cas12e。在一些情况下,可编程核酸酶是Csm1、Cas9、C2c4、C2c8、C2c5、C2c10、C2c9或CasZ。有时,Csm1也可称为smCms1、miCms1、obCms1或suCms1。有时CasZ也可以称为smCms1、miCms1、obCms1或suCms1。Cas13a有时也可以称为C2c2。CasZ有时也可以称为Cas14a、Cas14b、Cas14c、Cas14d、Cas14e、Cas14f、Cas14g或Cas14h。可编程核酸酶有时可以是V型CRISPR-Cas系统。在一些情况下,可编程核酸酶有时可以是VI型CRISPR-Cas系统。可编程核酸酶有时可以是III型CRISPR-Cas系统。在一些情况下,可编程核酸酶可以来自沙氏纤毛菌(*Leptotrichia shahii*, Lsh)、斯氏李斯特杆菌(*Listeria seeligeri*, Lse)、口腔纤毛菌(*Leptotrichia buccalis*, Lbu)、维氏纤毛菌(*Leptotrichia wadei*, Lwa)、荚膜红细菌(*Rhodobacter capsulatus*, Rca)、Herbinix hemicellulosilytica (Hhe)、浅红慢生盐场菌(*Paludibacter propionigenes*, Ppr)、毛螺科菌(*Lachnospiraceae bacterium*, Lba)、直肠真杆菌([*Eubacterium*] rectale, Ere)、*Listeria newyorkensis* (Lny)、嗜胺梭菌(*Clostridium aminophilum*, Cam)、普氏菌属(*Prevotella* sp., Psm)、犬咬二氧化碳嗜纤维菌(*Capnocytophaga canimorsus*, Cca)、毛螺科菌(Lba)、动物溃疡伯格菌(*Bergeyella zoohelcum*, Bzo)、中间普雷沃菌(*Prevotella intermedia*, Pin)、颊普雷沃菌(*Prevotella buccae*, Pbu)、另枝菌属(*Alistipes* sp., Asp)、鸭疫里默氏杆菌(*Riemerella anatipestifer*, Ran)、橙色普雷沃菌(*Prevotella aurantiaca*, Pau)、解糖普雷沃菌(*Prevotella saccharolytica*, Psa)、中间普雷沃菌(Pin2)、犬咬二氧化碳嗜纤维菌(Cca)、古卟啉单胞菌(*Porphyromonas gulae*, Pgu)、普雷沃菌属(*Prevotella* sp., Psp)、牙龈卟啉单胞菌(*Porphyromonas gingivalis*, Pig)、中间普雷沃菌(Pin3)、意大利肠球菌(*Enterococcus italicus*, Ei)、唾液乳杆菌(*Lactobacillus*

salivarius, Ls) 或嗜热栖热菌 (*Thermus thermophilus*, Tt) 中的至少一种。可编程核酸酶的任何组合均可用于复用。在一些实施方案中, 可编程核酸酶的复用在单个反应体积中发生。在其他实施方案中, 可编程核酸酶的复用在单个装置的不同反应体积中发生。

靶核酸扩增

[0142] 本文公开了使用本文所述的任何方法、试剂、试剂盒或装置扩增用于检测的靶核酸的方法。如本文所述, 用于扩增靶核酸的组合物及其使用方法与本文公开的DETECTR测定方法兼容。如本文所述, 用于扩增靶核酸的组合物及其使用方法与本文公开的任何可编程切口酶和所述可编程切口酶在检测靶核酸的方法中的用途兼容。靶核酸可以是扩增的目的核酸。目的核酸可以是本文公开的任何核酸或来自本文公开的任何样品。这种扩增可以是热扩增(例如, 使用PCR)或等温扩增。样品的这种核酸扩增可以提高靶核酸检测的灵敏度、特异性或准确度中的至少一种。用于核酸扩增的试剂可以包括重组酶、寡核苷酸引物、单链DNA结合(SSB)蛋白和聚合酶。核酸扩增可以是转录介导的扩增(TMA)。核酸扩增可以是解旋酶依赖性扩增(HDA)或环状解旋酶依赖性扩增(cHDA)。在另外的情况下, 核酸扩增是链置换扩增(SDA)。核酸扩增可以是重组酶聚合酶扩增(RPA)。核酸扩增可以是环介导的扩增(LAMP)或指数扩增反应(EXPAR)中的至少一种。在某些情况下, 核酸扩增是通过滚环扩增(RCA)、连接酶链反应(LCR)、扩增RNA靶标的简单方法(SMART)、单引物等温扩增(SPIA)、多置换扩增(MDA)、基于核酸序列的扩增(NASBA)、铰链引发的引物依赖性核酸扩增(hinge-initiated primer-dependent amplification of nucleic acids, HIP)、切口酶扩增反应(NEAR)或改进的多置换扩增(IMDA)。核酸扩增可以进行不超过1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、40、50或60分钟。核酸扩增反应有时在20-45℃左右的温度下进行。核酸扩增反应可以在不超过20℃、25℃、30℃、35℃、37℃、40℃、45℃的温度下进行。核酸扩增反应可以在至少20℃、25℃、30℃、35℃、37℃、40℃或45℃的温度下进行。

[0143] 如本文所述, 用于扩增靶核酸的组合物及其使用方法, 与包含可编程切口酶和缓冲液的任何组合物, 其已被开发以改进可编程切口酶的功能, 以及所述组合物在检测靶核酸的方法中的用途兼容。如本文所述, 用于扩增靶核酸的组合物及其使用方法, 与本文公开的任何方法兼容, 包括测定靶核酸序列中至少一个碱基差异的方法(例如, 测定SNP或碱基突变)、通过扩增靶核酸序列以引入PAM测定缺乏PAM的靶核酸的方法以及用于通过扩增将PAM引入靶核酸序列的组合物。在一些情况下, 靶核酸的扩增可以提高检测反应的灵敏度。在一些情况下, 靶核酸的扩增可以增加检测反应的特异性。相对于不对应于靶核酸的核酸浓度, 靶核酸的扩增可以增加样品中靶核酸的浓度。在一些实施方案中, 靶核酸的扩增可用于修饰靶核酸的序列。例如, 扩增可用于将PAM序列插入缺乏PAM序列的靶核酸中。在一些情况下, 扩增可用于增加靶核酸序列的同质性。例如, 扩增可用于移除靶核酸序列中的非目的核酸变异。

[0144] 扩增的靶核酸可以以相对于可编程切口酶的量存在于DETECTR反应中。在一些实施方案中, 相对于可编程切口酶的量, 扩增的靶核酸以至少1倍、2倍、3倍、4倍、5倍、10倍、25倍、50倍、100倍、500倍、1000倍、10,000倍、或100,000倍摩尔过量存在。在一些实施方案中, 相对于可编程切口酶的量, 扩增的靶核酸以不超过1倍、2倍、3倍、4倍、5倍、10倍、25倍、50倍、100倍、500倍、1000倍、10,000倍、或100,000倍摩尔过量存在。在一些实施方案中, 相对于可编程切口酶的量, 扩增的靶核酸以1倍至2倍、1倍至3倍、1倍至4倍、1倍至5倍、1倍

至10倍、1倍至25倍、1倍至50倍、1倍至100倍、1倍至500倍、1倍至1000倍、1倍至10,000倍、1倍至100,000倍、5倍至10倍、5倍至25倍、5倍至50倍、5倍至100倍、5倍至500倍、5倍至1000倍、5倍至10,000倍、5倍至100,000倍、10倍至25倍、10倍至50倍、10倍至100倍、10倍至500倍、10倍至1000倍、10倍至10,000倍、10倍至100,000倍、100倍至500倍、100倍至1000倍、100倍至10,000倍、100倍至100,000倍、1000倍至10,000倍、1000倍至100,000倍、或10,000倍至100,000倍摩尔过量存在。在一些实施方案中,相对于靶核酸的量,可编程切口酶以至少1倍、2倍、3倍、4倍、5倍、10倍、25倍、50倍、100倍、500倍、1000倍、10,000倍、或100,000倍摩尔过量存在。在一些实施方案中,相对于靶核酸的量,可编程切口酶以不超过1倍、2倍、3倍、4倍、5倍、10倍、25倍、50倍、100倍、500倍、1000倍、10,000倍、或100,000倍摩尔过量存在。在一些实施方案中,相对于靶核酸的量,可编程切口酶以1倍至2倍、1倍至3倍、1倍至4倍、1倍至5倍、1倍至10倍、1倍至25倍、1倍至50倍、1倍至100倍、1倍至500倍、1倍至1000倍、1倍至10,000倍、1倍至100,000倍、5倍至10倍、5倍至25倍、5倍至50倍、5倍至100倍、5倍至500倍、5倍至1000倍、5倍至10,000倍、5倍至100,000倍、10倍至25倍、10倍至50倍、10倍至100倍、10倍至500倍、10倍至1000倍、10倍至10,000倍、10倍至100,000倍、100倍至500倍、100倍至1000倍、100倍至10,000倍、100倍至100,000倍、1000倍至10,000倍、1000倍至100,000倍、或10,000倍至100,000倍摩尔过量存在。在一些实施方案中,样品中不存在靶核酸。

[0145] 扩增的靶核酸可以以相对于指导核酸的量的量存在于DETECTR反应中。在一些实施方案中,相对于指导核酸的量,扩增的靶核酸以至少1倍、2倍、3倍、4倍、5倍、10倍、25倍、50倍、100倍、500倍、1000倍、10,000倍、或100,000倍摩尔过量存在。在一些实施方案中,相对于指导核酸的量,扩增的靶核酸以不超过1倍、2倍、3倍、4倍、5倍、10倍、25倍、50倍、100倍、500倍、1000倍、10,000倍、或100,000倍摩尔过量存在。在一些实施方案中,相对于靶核酸的量,扩增的靶核酸以1倍至2倍、1倍至3倍、1倍至4倍、1倍至5倍、1倍至10倍、1倍至25倍、1倍至50倍、1倍至100倍、1倍至500倍、1倍至1000倍、1倍至10,000倍、1倍至100,000倍、5倍至10倍、5倍至25倍、5倍至50倍、5倍至100倍、5倍至500倍、5倍至1000倍、5倍至10,000倍、5倍至100,000倍、10倍至25倍、10倍至50倍、10倍至100倍、10倍至500倍、10倍至1000倍、10倍至10,000倍、10倍至100,000倍、100倍至500倍、100倍至1000倍、100倍至10,000倍、100倍至100,000倍、1000倍至10,000倍、1000倍至100,000倍、或10,000倍至100,000倍摩尔过量存在。在一些实施方案中,相对于靶核酸的量,指导核酸以至少1倍、2倍、3倍、4倍、5倍、10倍、25倍、50倍、100倍、500倍、1000倍、10,000倍、或100,000倍摩尔过量存在。在一些实施方案中,相对于靶核酸的量,指导核酸以不超过1倍、2倍、3倍、4倍、5倍、10倍、25倍、50倍、100倍、500倍、1000倍、10,000倍、或100,000倍摩尔过量存在。在一些实施方案中,相对于靶核酸的量,指导核酸以1倍至2倍、1倍至3倍、1倍至4倍、1倍至5倍、1倍至10倍、1倍至25倍、1倍至50倍、1倍至100倍、1倍至500倍、1倍至1000倍、1倍至10,000倍、1倍至100,000倍、5倍至10倍、5倍至25倍、5倍至50倍、5倍至100倍、5倍至500倍、5倍至1000倍、5倍至10,000倍、5倍至100,000倍、10倍至25倍、10倍至50倍、10倍至100倍、10倍至500倍、10倍至1000倍、10倍至10,000倍、10倍至100,000倍、100倍至500倍、100倍至1000倍、100倍至10,000倍、100倍至100,000倍、1000倍至10,000倍、1000倍至100,000倍、或10,000倍至100,000倍摩尔过量存在。在一些实施方案中,样品中不存在靶核酸。

试剂盒

[0146] 本文公开了用于使用上述方法检测或修饰本文公开的靶核酸的试剂盒。在一些实施方案中,试剂盒包含可编程切口酶系统、试剂和支持介质。试剂和可编程切口酶系统可以在试剂室中或支持介质上提供。可选地,试剂和可编程切口酶系统可以被使用试剂盒个体放入试剂室或支持介质中。任选地,试剂盒进一步包括缓冲液和滴管。试剂室可以是检测孔或容器。试剂室的开口可以足够大以容纳支持介质。缓冲液可以装在滴管瓶中,以便于分配。滴管可以是一次性的并且可以转移固定体积。滴管可用于将样品放入试剂室或支持介质上。

[0147] 用于检测本文所述的靶核酸的试剂盒或系统还包括用于样品中的靶核酸的核酸扩增的试剂。等温核酸扩增允许在偏远地区或资源匮乏的环境中使用试剂盒或系统,而无需专门的用于扩增的装置。通常,用于核酸扩增的试剂包括重组酶、寡核苷酸引物、单链DNA结合(SSB)蛋白和聚合酶。有时,样品的核酸扩增提高了检测靶核酸的测定中的灵敏度、特异性或准确性的至少一种。在一些情况下,核酸扩增在支持介质上的核酸扩增区域中进行。替代地或组合地,在试剂腔室中执行核酸扩增,并且将所得到的样品施加在支持介质上。有时,核酸扩增是等温核酸扩增。在一些情况下,核酸扩增是转录介导扩增(TMA)。核酸扩增是解旋酶依赖性扩增(HDA)或在其他情况下是环状解旋酶依赖性扩增(cHDA)。在其他情况下,核酸扩增是链置换扩增(SDA)。在一些情况下,核酸扩增通过重组酶聚合酶扩增(RPA)进行。在一些情况下,核酸扩增通过环介导扩增(LAMP)或指数式扩增反应(EXPAR)中的至少一个进行。在一些情况下,核酸扩增通过滚环扩增(RCA)、连接酶链反应(LCR)、扩增RNA靶标的简单方法(SMART)、单引物等温扩增(SPIA)、多置换扩增(MDA)、基于核酸序列的扩增(NASBA)、铰链引发的引物依赖性核酸扩增(HIP)、切口酶扩增反应(NEAR)或改进的多置换扩增(IMDA)进行。在一些情况下,LAMP扩增可允许同时对Cas12和Cas13进行单个扩增步骤。在一些实施方案中,LAMP可以允许同时扩增针对至多三个或三个以上可编程核酸酶的靶核酸。在一些实施方案中,使用RPA,需要的引物更少,并且多路复用可以增加至三个或六个可编程核酸酶。通常,核酸扩增进行不超过1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、40、50或60分钟或在1至60分钟之间的任意值。有时,核酸扩增进行1至60分钟、5至55分钟、10至50分钟、15至45分钟、20至40分钟或25至35分钟。有时,核酸扩增反应在20-45℃左右进行。在一些情况下,核酸扩增反应在不大于20℃、25℃、30℃、35℃、37℃、40℃、45℃或20℃至45℃的任意值的温度下进行。在一些情况下,核酸扩增反应在至少20℃、25℃、30℃、35℃、37℃、40℃或45℃或20℃至45℃的任意值的温度下进行。在一些情况下,核酸扩增反应在20℃至45℃、25℃至40℃、30℃至40℃或35℃至40℃的温度下进行。

[0148] 在一些实施方案中,用于检测靶核酸的试剂盒包含支持介质;靶向靶序列的指导核酸;当与指导核酸和靶序列复合时能够被激活的可编程切口酶;以及包含检测部分的报告核酸,其中报告核酸能够被激活的核酸酶切割,从而产生第一可检测信号。通常,试剂盒进一步包含用于扩增目的靶核酸以产生PAM靶核酸的引物。

[0149] 在一些实施方案中,用于检测靶核酸的试剂盒包含PCR板;靶向靶序列的指导核酸;当与指导核酸和靶序列复合时能够被激活的可编程切口酶;以及包含检测部分的单链报告核酸,其中报告核酸能够被激活的核酸酶切割,从而产生第一可检测信号。PCR板的孔可由靶向靶核酸区段的指导核酸、当与指导核酸和靶序列复合时能够被激活的可编程核酸酶以及至少一个包含检测部分的单链检测核酸群体预先等分。因此,用户可以将感兴趣的

生物样品添加到预等分的PCR板的孔中,并使用荧光读取器或可见光读取器测量可检测信号。

[0150] 在一些实施方案中,用于修饰靶核酸的试剂盒包含支持介质;靶向靶序列的指导核酸;以及当与指导核酸和靶序列复合时能够被激活的可编程切口酶。

[0151] 在一些实施方案中,用于修饰靶核酸的试剂盒包含PCR板;靶向靶序列的指导核酸;以及当与指导核酸和靶序列复合时能够被激活的可编程切口酶。PCR板的孔可以预先等分有靶向靶序列的指导核酸和当与指导核酸和靶序列复合时能够被激活的可编程切口酶。因此,使用者可以将目的生物样品添加至预等分的PCR板的孔中。

[0152] 在一些情况下,此类试剂盒可包括包装、载体或容器,其被分隔以容纳一个或多个容器,例如小瓶、试管等,每个容器包括将在本文所述方法中使用的一个单独元件。合适的容器包括,例如,测试孔、瓶、小瓶和试管。在一个实施方案中,容器由例如玻璃、塑料或聚合物的多种材料形成。

[0153] 本文描述的试剂盒或系统包含包装材料。包装材料的实例包括但不限于适合预期使用模式的小袋、泡罩包装、瓶、管、袋、容器、瓶和任何包装材料。

[0154] 试剂盒通常包括列出内容物和/或使用说明的标签,以及带有使用说明的包装说明书。通常还包括一组说明书。在一个实施方案中,标签在容器上或与容器相关联。在一些情况下,当构成标签的字母、数字或其他字符被附着、模制或蚀刻到容器本身时,标签就在容器上;当标签出现在同样容纳容器的大容器或载体内时,标签就与容器相关联,例如作为包装说明书。在一个实施方案中,标签用于指示所述内容物将用于特定治疗应用。标签还指示内容物的使用指导,例如在本文所描述的方法中。

[0155] 在包装成型产品并打包或装箱以保持无菌屏障后,可通过热灭菌、气体灭菌、 γ 射线辐照或电子束灭菌对产品进行最终灭菌。或者,可以通过无菌处理制备和包装产品。

[0156] 除非另有定义,否则本文所使用的所有技术术语具有本发明所属领域中的普通技术人员通常理解的相同含义。除非上下文明确规定,如本说明书和所附权利要求书中所使用的单数形式“一个(a)”、“一(an)”和“所述(the)”包括复数指代。除非另有说明,否则本文中任何提及的“或”旨在包括“和/或”。

[0157] 如本文所用,术语“包括”及其语法等同项规定了所述特征、整数、步骤、操作、元件和/或组分的存在,但不排除存在或添加一个或多个其他特征、整数、步骤、操作、元件、组分和/或其集合。如本文所使用的,术语“和/或”包括一个或多个相关联的所列项目的任何和所有组合。

[0158] 除非上下文有明确说明或明显说明,如本文所用,提及数字或数字范围的术语“约”应理解为表示所述数字或其数字的 $\pm 10\%$ 或对于范围所列值而言低于所列下限的 10% 且高于所列上限的 10% 。

[0159] 如本文所用,术语“个体”、“受试者”和“患者”可互换使用,包括动物界的任何成员,包括人类。

[0160] 如本文所用,术语“抗体”是指但不限于单克隆抗体、合成抗体、多克隆抗体、多特异性抗体(包括双特异性抗体)、人类抗体、人源化抗体、嵌合抗体、单链(scFv)(包括双特异性scFv)、单链抗体、Fab片段、F(ab')₂片段、F(ab')₂片段、二硫键连接Fvs(sdFv)或其表位结合片段。在一些情况下,抗体是免疫球蛋白分子或免疫球蛋白分子的免疫活性部分。在一些

情况下,抗体来源于包括鸟类和哺乳动物的动物。或者,抗体是人或人源化单克隆抗体。

编号实施方案

[0161] 以下实施方案叙述了本文公开的特征组合的非限制性排列。还预期了特征组合的其他排列。特别地,这些编号的实施方案中的每个方案均被预期为依赖或涉及每个先前或随后编号的实施方案,而不依赖它们所列出的顺序。1.一种在靶核酸中引入断裂的方法,所述方法包括通过使所述靶核酸与以下(a)和(b)接触来引入所述断裂:(a)第一指导核酸,所述第一指导核酸包含与长度不超过900个氨基酸的第一可编程切口酶结合的第一区域;(b)第二指导核酸,所述第二指导核酸包含与长度不超过900个氨基酸的第二可编程切口酶结合的第一区域,其中所述第一指导核酸包含与所述靶核酸结合的第二区域并且其中所述第二指导核酸包含与所述靶核酸结合的第二区域并且其中所述第一指导核酸的第二区域和所述第二指导核酸的第二区域与所述靶核酸的相反链结合。2.实施方案1的方法,其中所述第一可编程切口酶和所述第二可编程切口酶的长度为350至900个氨基酸。3.实施方案1-2中任一项的方法,其中所述第一可编程切口酶和所述第二可编程切口酶的长度为480至550个氨基酸。4.实施方案1-3中任一项的方法,其中所述第一可编程切口酶和所述第二可编程切口酶是V型CRISPR/Cas酶。5.实施方案4的方法,其中所述V型CRISPR/Cas酶包含三个部分RuvC结构域。6.实施方案5的方法,其中所述三个部分RuvC结构域为RuvC-I、RuvC-II和RuvC-III亚结构域。7.实施方案1-6中任一项的方法,其中所述第一可编程切口酶和所述第二可编程切口酶是Cas14蛋白。8.实施方案7的方法,其中所述Cas14蛋白是Cas14a蛋白、Cas14b蛋白、Cas14c蛋白、Cas14d蛋白或Cas14e蛋白。9.实施方案7-8中任一项的方法,其中所述Cas14蛋白是Cas14a蛋白。10.实施方案7-8中任一项的方法,其中所述Cas14蛋白是Cas14b蛋白。11.实施方案7-8中任一项的方法,其中所述Cas14蛋白是Cas14e蛋白。12.实施方案1-11中任一项的方法,其中所述第一可编程切口酶、所述第二可编程切口酶或两者与SEQ ID NO:1-SEQ ID NO:91或SEQ ID NO:170中的任一项具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。13.实施方案1-12中任一项的方法,其中所述第一可编程切口酶、所述第二可编程切口酶或两者是SEQ ID NO:1-SEQ ID NO:91或SEQ ID NO:170中的任一项。14.实施方案1-13中任一项的方法,其中所述第一可编程切口酶、所述第二可编程切口酶或两者与SEQ ID NO:1具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。15.实施方案1-14中任一项的方法,其中所述第一可编程切口酶、所述第二可编程切口酶或两者是SEQ ID NO:1。16.实施方案1-13中任一项的方法,其中所述第一可编程切口酶、所述第二可编程切口酶或两者与SEQ ID NO:10具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。17.实施方案1-13或16中任一项的方法,其中所述第一可编程切口酶、所述第二可编程切口酶或两者是SEQ ID NO:10。18.实施方案1-13中任一项的方法,其中所述第一可编程切口酶、所述第二可编程切口酶或两者与SEQ ID NO:11具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。19.实施方案1-13或18中任一项的方法,其中所述第一可编程切口酶、所述第二可编程切口酶或两者是SEQ ID NO:11。20.实施方案1-13中任一项的方法,其中所述第一可编程切口酶、所述第二可编程切口酶或两者与SEQ ID NO:17具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。21.实施方案1-13或20中任一项的方法,其中所述第一可

编程切口酶、所述第二可编程切口酶或两者是SEQ ID NO:17。22. 实施方案1-13中任一项的方法,其中所述第一可编程切口酶、所述第二可编程切口酶或两者与SEQ ID NO:33具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。23. 实施方案1-13或22中任一项的方法,其中所述第一可编程切口酶、所述第二可编程切口酶或两者是SEQ ID NO:33。24. 实施方案1-23中任一项的方法,其中所述第一指导核酸是第一指导RNA。25. 实施方案1-24中任一项的方法,其中所述第二指导核酸是第二指导RNA。26. 实施方案1-25中任一项的方法,其中所述第一区域是重复序列并且其中所述第二区域是间隔区序列。27. 实施方案1-26中任一项的方法,其中所述第一指导核酸和所述第二指导核酸包含crRNA和tracrRNA。28. 实施方案1-26中任一项的方法,其中所述第一指导核酸和所述第二指导核酸包含crRNA和trancRNA。29. 实施方案27-28中任一项的方法,其中所述crRNA包含所述重复序列和所述间隔区序列。30. 实施方案26-29中任一项的方法,其中所述重复序列与所述tracrRNA的区段杂交。31. 实施方案27-30中任一项的方法,其中所述tracrRNA与SEQ ID NO:98-SEQ ID NO:99、SEQ ID NO:101、SEQ ID NO:103、SEQ ID NO:105-SEQ ID NO:151中的任一项具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。32. 实施方案27-31中任一项的方法,其中所述tracrRNA是SEQ ID NO:98-SEQ ID NO:99、SEQ ID NO:101、SEQ ID NO:103、SEQ ID NO:105-SEQ ID NO:151中的任一项。33. 实施方案27-31中任一项的方法,其中所述tracrRNA与SEQ ID NO:99具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。34. 实施方案27-31或33中任一项的方法,其中所述tracrRNA是SEQ ID NO:99。35. 实施方案27-31中任一项的方法,其中所述tracrRNA与SEQ ID NO:101具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。36. 实施方案27-31或35中任一项的方法,其中所述tracrRNA是SEQ ID NO:101。37. 实施方案27-31中任一项的方法,其中所述tracrRNA与SEQ ID NO:103具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。38. 实施方案27-31或37中任一项的方法,其中所述tracrRNA是SEQ ID NO:103。39. 实施方案27-31中任一项的方法,其中所述tracrRNA与SEQ ID NO:119具有至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%序列同一性。40. 实施方案27-31或39中任一项的方法,其中所述tracrRNA是SEQ ID NO:119。41. 实施方案1-40中任一项的方法,其中所述第一可编程切口酶和所述第二可编程切口酶表现出比双链切割活性高2倍的产生切口活性。42. 实施方案1-41中任一项的方法,其中所述第一可编程切口酶和所述第二可编程切口酶在两个不同位点使所述靶核酸产生切口。43. 实施方案1-42中任一项的方法,其中所述靶核酸包含双链DNA。44. 实施方案43的方法,其中所述两个不同位点在所述双链DNA的相反链上。45. 实施方案1-44中任一项的方法,其中所述靶核酸包含已突变的序列或与疾病相关的序列。46. 实施方案45的方法,其中所述疾病是癌症。47. 实施方案1-46中任一项的方法,其中所述方法包括向有需要的受试者施用所述第一可编程切口酶和所述第二可编程切口酶。48. 实施方案45的方法,其中在所述第一可编程切口酶和所述第二可编程切口酶使所述靶核酸产生切口之后移除所述已突变的序列。49. 实施方案1-48中任一项的方法,其中所述第一可编程切口酶和所述第二可编程切口酶相同。50. 实施方案1-49中任一项的方法,其中所述第一可编程切口酶和所述第二可编程切口酶不同。1. 一种在靶核酸中引入断裂的方法,所述方法包括通过使所述靶核酸与以下(a)和

(b) 接触来引入所述断裂：(a) 第一指导RNA，所述第一指导RNA包含与第一可编程切口酶结合的第一区域；(b) 第二指导RNA，所述第二指导RNA包含与第二可编程切口酶结合的第一区域，其中所述第一指导RNA包含与所述靶核酸结合的第二区域并且其中所述第二指导RNA包含与所述靶核酸结合的第二区域并且其中所述第一指导RNA的第二区域和所述第二指导RNA的第二区域与所述靶核酸的相反链结合。2. 实施方案1的方法，其中所述第一可编程切口酶和所述第二可编程切口酶在两个不同的位点使所述靶核酸产生切口。3. 实施方案1的方法，其中所述靶核酸包含双链DNA。4. 实施方案3的方法，其中所述两个不同的位点在所述双链DNA的相反链上。5. 实施方案1的方法，其中所述靶核酸包含已突变的序列或与疾病相关的序列。6. 实施方案5的方法，其中所述疾病是癌症。7. 实施方案1的方法，其中所述方法包括向有需要的受试者施用所述第一可编程切口酶和所述第二可编程切口酶。8. 实施方案5的方法，其中在所述第一可编程核酸酶和所述第二可编程核酸酶使所述靶核酸产生切口之后移除所述已突变的序列。9. 实施方案1的方法，其中所述第一可编程切口酶和所述第二可编程切口酶包含Cas14蛋白。10. 一种检测样品中靶核酸的方法，所述方法包括使所述样品与以下(a)、(b)和(c)接触：(a) 可编程切口酶；(b) 指导RNA，所述指导RNA包含与所述可编程切口酶结合的第一区域和与所述靶核酸结合的第二区域；和(c) 标记的单链DNA报告子，所述标记的单链DNA报告子不与所述指导RNA结合；切割所述标记的单链DNA报告子以释放可检测的标记；以及通过测量来自所述可检测标记的信号来检测所述靶核酸。11. 实施方案10的方法，其中所述靶核酸是单链DNA。12. 实施方案10的方法，其中所述可编程切口酶包含Cas 14蛋白。13. 实施方案1或10的方法，其中所述靶核酸在样品中。14. 实施方案13的方法，其中所述样品包含磷酸盐缓冲液、Tris缓冲液或HEPES缓冲液。15. 实施方案13的方法，其中所述样品包含7至9的pH值。16. 实施方案13的方法，其中所述样品包含7.5至8的pH值。17. 实施方案13的方法，其中所述样品包含25nM至200mM的盐浓度。18. 实施方案10的方法，其中所述单链DNA报告子包含ssDNA-荧光猝灭DNA报告子。19. 实施方案18的方法，其中所述ssDNA-荧光猝灭DNA报告子是通用ssDNA-荧光猝灭DNA报告子。20. 实施方案1或10的方法，其中所述可编程切口酶表现出不依赖PAM的切口和切割。21. 实施方案9或12的方法，其中所述Cas14蛋白包含Cas14e蛋白。22. 实施方案9或12的方法，其中所述Cas14蛋白包含400至800个氨基酸残基。23. 一种组合物，所述组合物包含可编程切口酶和指导RNA，所述指导RNA包含与所述可编程切口酶结合的第一区域和与所述靶核酸结合的第二区域。24. 实施方案23的组合物，其中所述靶核酸包含单链DNA或双链DNA。25. 实施方案23的组合物，其中所述可编程切口酶表现出不依赖PAM的产生切口和切割。26. 实施方案23的组合物，其中所述可编程切口酶使所述双链DNA的单链产生切口。27. 实施方案23的组合物，其中所述可编程切口酶切割单链DNA。28. 实施方案23的组合物，其中所述可编程切口酶包含Cas14蛋白。29. 实施方案28的组合物，其中所述Cas14蛋白包含Cas14e蛋白。30. 实施方案29的组合物，其中所述Cas14蛋白包含400至800个氨基酸残基。

实施例

[0162] 包括以下实施例以进一步描述本公开内容的一些方面，并且不应用于限制本发明的范围。

实施例1

Cas14e是可编程切口酶

[0163] 本实施例显示了Cas14e是可编程切口酶。图1显示了说明由四种不同Cas14e蛋白对dsDNA产生切口的凝胶。在凝胶中,四种不同Cas14e蛋白与指导RNA (TRACR2) 一起独立添加至前四个泳道,指导RNA与每种Cas14e蛋白形成复合体。当指导RNA与Cas14e蛋白复合并且该复合体与其靶核酸结合时,Cas14e蛋白的切口酶活性被激活。这在凝胶的前四个泳道中由所得的两条条带显示,其中上面的条带是带有切口的靶dsDNA。第五个泳道是包含Cas14e蛋白但不含指导RNA的对照泳道,其中靶dsDNA保持完整。第六个泳道显示了限制酶EcoRI对dsDNA的切割,该酶在靶dsDNA中产生双链断裂。第七个泳道显示未处理的靶dsDNA (例如,无可编程切口酶、指导RNA或限制性内切酶)。因此,Cas14e是可编程切口酶。

实施例2

使用可编程切口酶在dsDNA中引入链断裂

[0164] 本实施例描述了使用可编程切口酶 (例如,Cas14,例如Cas14e) 在dsDNA中引入断裂。两种可编程切口酶,例如靶向dsDNA的两条不同链的两种Cas14e,被共同递送。第一Cas14e蛋白与靶向dsDNA的第一区域的第一指导RNA结合,并且第二Cas14e蛋白与靶向位于dsDNA的相反链上的第二区域的第二指导RNA结合。Cas14e蛋白使相反的靶DNA链产生切口,并在dsDNA中产生两个断裂。这两条链断裂通过非同源末端连接 (NHEJ) 或同源定向修复 (HDR) 修复和重新连接。因此,两种可编程切口酶被组合以选择性编辑靶核酸分子的序列。

实施例3

使用缓冲液、pH值和温度调整Cas14e可编程切口酶活性

[0165] 本实施例描述了在DETECTR测定中使用缓冲液、pH值和温度调整Cas14e可编程切口酶活性。Cas14e可编程切口酶与含有靶ssDNA的样品孵育,这可激活Cas14e可编程切口酶无差别切割ssDNA-FQ报告子。切割的ssDNA-FW报告子释放可检测的标记,该标记通过荧光读数进行测量。在各种缓冲液、pH值和温度条件下运行该DETECTR测定,使用中靶指导RNA和脱靶指导RNA对照。

[0166] 图2显示了盐、缓冲液和温度对使用Cas14e的ssDNA DETECTR反应的影响。左上柱状图的x轴显示了各种缓冲条件和pH值,在y轴显示了减去背景的荧光。荧光表明报告子的切割。荧光越强表明活性越高。中上和右上图显示了x轴所示的温度 (“ON”表示添加可以与指导RNA杂交的靶ssDNA; “OFF”表示添加不与指导RNA杂交的脱靶ssDNA) 相对于y轴所示的原始荧光。荧光表明报告子的切割。

实施例4

Cas14e反式切割活性不依赖盐浓度

[0167] 本实施例显示了在DETECTR测定中,Cas14e反式切割活性不依赖盐浓度。将Cas14e蛋白与含有靶ssDNA的样品孵育,这可激活Cas14e反式切割活性无差别切割ssDNA-FQ报告子。切割的ssDNA-FQ报告子释放可检测的标记,该标记通过荧光读数进行测量。在存在靶ssDNA和存在脱靶ssDNA的情况下,在各种盐浓度 (包括25nM NaCl、100nM NaCl和200mM NaCl) 下运行该DETECTR测定。

[0168] 图2的下面三个线图显示了在各种盐条件 (从左至右为25nM NaCl、100nM NaCl和200mM NaCl) 下随时间的荧光。荧光表明报告子的切割。较高的线且荧光随时间增加,显示在存在靶ssDNA的情况下,Cas14e蛋白与指导RNA复合对报告子的切割。较低的线显示了在存在脱靶ssDNA的情况下,与指导RNA复合的Cas14e蛋白对报告子的背景切割。令人惊讶的

是,发现Cas14e可编程切口酶即使在非常高的盐浓度下也能发挥作用,这表明这些切口酶的功能不依赖盐浓度。

实施例5

ssDNA-FQ报告子序列非依赖性Cas14e

[0169] 本实施例显示了在DETECTR测定中Cas14e蛋白的ssDNA-FQ报告子序列的非依赖性活性。Cas14e蛋白与含有靶ssDNA的样品孵育,这可激活Cas14e蛋白以无差别切割ssDNA-FQ报告子的三种不同均聚物(T12、A12或C12)中的一种。切割的ssDNA-FQ报告子释放可检测的标记,该标记通过荧光读数进行测量。

[0170] 图3显示了三个图,这些图从左至右评估了均聚物荧光猝灭(FQ)报告子的切割。荧光表明报告子的切割。荧光越强表明活性越高。最左侧的图使用T12(12个胸腺嘧啶残基)ssDNA-FQ报告子,中间的图使用A12(12个腺嘌呤残基)ssDNA-FQ报告子,最右侧的图使用C12(12个胞嘧啶残基)ssDNA-FQ报告子。在每个图中,上线显示了在存在靶ssDNA的情况下与指导RNA复合的Cas14e蛋白对报告子的切割,下线显示了在存在脱靶ssDNA的情况下与指导RNA复合的Cas14e蛋白对报告子的切割。

实施例6

Cas14e PAM-不敏感性

[0171] 本实施例描述了Cas14e PAM-不敏感性。图4显示了使用与指导RNA偶联的Cas14e蛋白检测靶dsDNA的三个DETECTR反应的荧光随时间变化的图。荧光表明报告子的切割。荧光越强表明活性越高。最上线显示了在存在具有野生型(wt)PAM的靶dsDNA的情况下的报告子的切割。最上线正下方的线显示了在存在具有突变(mut)PAM的靶dsDNA的情况下报告子的切割。在一些实施方案中,突变PAM与天然PAM的区别在于单个核苷酸。在一些实施方案中,突变PAM与天然PAM的区别在于多个核苷酸。在一些实施方案中,突变PAM比天然PAM短。在一些实施方案中,突变PAM比天然PAM长。最下线显示了在存在500nM脱靶dsDNA的情况下报告子的切割。结果显示Cas14e可以在没有PAM限制的情况下检测靶dsDNA。

实施例7

Cas14a和Cas14b产生切口和切割活性

[0172] 本实施例描述了测量与多种指导核酸复合的多种可编程切口酶的产生切口和切割活性的测定。通过单独测量与不同指导核酸复合的SEQ ID NO:1、SEQ ID NO:10、SEQ ID NO:11和SEQ ID NO:17的四种不同Cas14直系同源物的活性检测不同tracr序列和Cas14序列对靶核酸的产生切口和切割的影响。Cas14可编程切口酶与靶向SEQ ID NO:96的crRNA序列和表3(下表)中所示的多种tracr序列复合。每个测定中使用的Cas14、gRNA对的序列如表3所示。

表3-指导核酸、靶核酸和可编程切口酶序列

物种	SEQ ID NO	序列
Cas14a.3 可编程切口酶	SEQ ID NO: 1	MEVQKTVMKTLRLRPLYSQEIEKEIKEEK ERRKQAGGTGELDGGFYKKLEKXHSEMFSF DRLNLLLNQLQREIAKVYNHAISELYIATIAQ GNKSNKHYYISSIVYNRAYGYFYNAIYALGICS KVEANFRSNELLTQQSALPTAKSDNFPIVLHK QKGAEGEDGGFRISTEGSDLIFEIPIPFYEYNG ENRKEPYKWVKKGGQKPVLLKLILSTFRRQRN KGWAKDEGTD AEIRKVTEGKYQVSQIEINRG KKLGEHQKWFANFSIEQPIYERKPNRSIVGGL DVGIRSPVCAINNSFSRYSVDSNDVFKFSKQ VFAFRRLLSKNSLKRKGHGAHKLPEITEM TEKNDKFRKKIIRWAKEVTNFFVKNQVGVQ QIEDLSTMKDREDHFFNQYLRGFWPYQMQ TLIENKLKEYGIEVKRVQAKYTSQLCSNPNC RYWNNYFNFEYRKVNKFPKFKCEKCNLEISA DYNAARNLSTPDIEKFVAKATKGINLPEK

物种	SEQ ID NO	序列
编码 Cas14a.3 可编程切口酶的 DNA	SEQ ID NO: 92	ATGGAAGTGCAGAAAACCGTGATGAAAAC CCTGAGCCTGCGGATCCTGCGGCCTCTGT CTCTCAAGAGATCGAGAAAGAGATCAAAG AGGAAAAAGAGCGCCGGAAGCAGGCTGGC GGAACAGGCGAACTTGATGGCGGCTTCTAC AAGAAGCTGGAAGAAACACAGCGAGAT GTTCAGCTTCGACCGGCTGAACCTGCTGCT GAATCAGCTGCAGCGGGAATCGCCAAGG TGTACAACCACGCCATCAGCGAGCTGTATA TCGCCACAATCGCCAGGGCAACAAGAGC AACAAGCACTACATCAGCAGCATCGTGTAC AACAGAGCCTACGGCTACTTCTACAACGCC TATATCGCCCTGGGCATCTGCAGCAAGGTG GAAGCCAACTTCAGAAGCAACGAGCTGCT GACCCAGCAGAGCGCACTGCCTACAGCCA AGAGCGACAACCTCCCCATCGTGTGTCACA AGCAGAAAGGCGCCGAAGGCGAGGATGGC GGATTTCAGAATCAGCACCGAGGGCAGCGA CCTGATCTTCGAGATCCCCATTCCATTCTAC GAGTACAACGGCGAGAACCGGAAAGAACC CTACAAATGGGTCAAGAAAGGCGGGCAGA AACAGTGCTGAAGCTGATCCTGAGCACCT TCCGGCGGCAGAGAAACAAAGGCTGGGCC AAAGATGAGGGCACCGACGCCGAGATCAG AAAAGTGACAGAGGGCAAGTACCAGGTGT CCCAGATCGAGATCAACCGGGGCAAGAAA CTGGGCGAGCACCAGAAGTGGTTTCGCCAAT TTCAGCATCGAGCAGCCCATCTACGAGCGG AAGCCCAACAGATCTATCGTCGGCGGCCTG GACGTGGGCATTAGATCTCCACTCGTGTGC GCCATCAACAACAGCTTCTCCCGGTACAGC GTGGACAGCAACGACGTGTTCAAGTTCAGC AAACAGGTGTTTCGCCTTCAGACGGCGGCTG CTGAGCAAGAACAGCCTGAAGAGAAAAGG CCACGGCGCTGCCACAAGCTGGAACCTAT CACCGAGATGACCGAGAGAAGACGACAAGT TCCGGAAGAAAATCATCGAGCGCTGGGCT AAAGAAGTGACCAACTTCTTCGTGAAGAAT CAAGTGGGCATTGTGCAGATCGAGGACCTG TCCACCATGAAGGACAGAGAGGACCACTT CTTCAACCAGTACCTGAGAGGCTTCTGGCC CTACTACCAGATGCAGACCTGATCGAGAA CAAGCTGAAAGAATACGGCATCGAAGTGA AGCGCGTGCAGGCCAAGTACACCAGCCAG CTGTGCAGCAACCCCAACTGCCGGTACTGG AACAACTACTTCAACTTCGAGTATCGGAAA GTGAACAAGTTCCCCAAGTTCAAGTGCAG AAGTGCAACCTGGAATCAGCGCCGACTA CAATGCCGCCAGAAACCTGAGCACACCCG ACATCGAGAAGTTCGTGGCCAAGGCCACC AAGGCATCAACCTGCCTGAGAAGTGA

物种	SEQ ID NO	序列
Cas14b.3 可编程切口酶	SEQ ID NO: 10	MEESIITGVKFKLRIDKETTKKLNEYFDEYGK AINFAVKIIQKELADDRFAGKAKLDQNKNPIL DENGKKIYEFDEFCSGKQV NKYVNNK PFC QECYKIRFTENGIRKRMYS AKGRKA EHKINIL NSTNKISKTHFN YAI REAFILDKSIKKQRKKR NERLRESKKRLQQFIDMRDGKREICPTIKGQK VDRFIHPSWITKDKKLEDFRGYTLSIINSKIKIL DRNIKREEKSLKEKGQIIFKAKRLMLDKSIRF VGDRKVLFTISK TLPKEYELD LPSKEKRLNW LKEKIEIKNQKPKYAYLLRKNIESEKKPNYE YYLQYTLEIKPELKDFYDGAIGIDRGINHIAV CTFISNDGKVTPPKFFSSGEILRLKNLQKERD RFLLRKHNNRKKGNMRVIENKINLILHRY S KQIVDMAKKLNASIVFEELGRIGKSRTKMKK SQRYKLSLFIFKKLSDLVDYKSRREGIRVTYV PPEYTSKECSHCGEKVNTQRPFGNYS LFKC NKCGIQLNSDYNASINI AKKGLKIPNST
编码 Cas14b.3 可编程切口酶的 DNA	SEQ ID NO: 93	ATGGAAGAGTCCATCATCACCGGCGTGAA GTTCAAGCTGCGGATCGACAAAGAAACCA CCAAGAAGCTGAACGAGTACTTCGACGAG TACGGCAAGGCCATCAACTTCGCCGTGAAG ATCATCCAGAAAGAGCTGGCCGACGACAG ATTTCGCCGAAAGGCCAAGCTGGACCAGA ACAAGAACCCCATCCTGGACGAGAATGGC AAGAAGATCTACGAGTTCCCCGACGAGTTC TGCAGCTGCGGCAAGCAAGTGAACAAATA CGTGAACAACAAGCCCTTCTGCCAAGAGTG CTACAAGATCCGGTTCACCGAGAACGGCAT CCGGAAGAGAATGTACAGCGCCAAGGGCA GAAAGGCCGAGCACAAAGATCAACATCCTG AACAGCACCAACAAGATCAGCAAGACCCA CTTCAACTACGCCATCAGAGAGGCCTTCAT CCTGGATAAGAGCATCAAGAAGCAGCGGA AGAAGCGCAACGAGCGGCTGAGAGAGAGC AAGAAGAGACTGCAGCAGTTCATCGACAT GCGCGACGGCAAGAGAGAGATCTGCCCTA CCATCAAGGGCCAGAAGGTGGACCGGTTC ATTACCCCAAGCTGGATCACCAAGGACAAG AAGCTCGAGGACTTCCGGGGCTACACCCTG AGCATCATCAACAGCAAGATCAAGATTCTG GACCGGAACATCAAGCGGGAAGAGAAGTC CCTGAAAGAGAAGGGGCAGATCATCTTCA AGGCCAAGAGACTGATGCTGGACAAGTCC ATCAGATTTCGTGGGCGACAGAAAGGTGCT GTTTACCATCTCCAAGACACTGCCCAAAGA GTACGAGCTGGACCTGCCTAGCAAAGAGA AGCGGCTGAACTGGCTGAAAGAAAAGATC GAGATCATCAAGAACCAGAAGCCGAAGTA CGCCTACCTGCTGCGGAAGAACATCGAGA GCGAGAAGAAGCCCAATTACGAGTACTAC CTGCAGTACACCCTGGAAATCAAGCCCGAG CTGAAGGACTTCTACGACGGCGCCATCGGC

物种	SEQ ID NO	序列
		ATCGACAGAGGCATCAATCACATTGCCGTG TGCACCTTCATCAGCAACGACGGCAAAGTG ACCCTCCTAAGTTCTTCAGCAGCGGCGAG ATCCTGAGACTGAAGAACCTGCAGAAAGA ACGCGACCGGTTCTTGCTGAGAAAGCACA ACAAGAATCGGAAGAAAGGCAACATGCGC GTGATCGAGAACAAGATTAACCTGATCCTG CACCGGTACAGCAAGCAGATCGTGGACAT GGCCAAAAAGCTGAATGCCAGCATCGTGT CGAGGAAGTGGGCAGAATCGGCAAGAGCC GGACCAAGATGAAGAAGTCCCAGCGGTAC AAGCTGAGCCTGTTTCATCTTTAAGAACTG AGCGACCTGGTGGACTACAAGTCTCGGAG AGAAGGCATCAGAGTGACCTACGTGCCAC CAGAGTACACCAGCAAAGAGTGCTCTCACT GCGGAGAGAAAGTGAACACCCAGCGGCCT TTCAACGGCAACTACTCCCTGTTCAAGTGC AACAAGTGCGGCATCCAGCTGAACAGCGA CTACAACGCCAGCATCAATATCGCCAAGAA GGGCCTGAAGATCCCCAACTCCACCTGA
Cas14b.4 可编程切口酶	SEQ ID NO: 11	MPKQDLVTTGIKFKLDVDKETRKKLDDYFD EYGKAINFAVKIIQKNLKEDRFAGKIALGEDK KPLLDKDGKKIYNYPNESCSCGNQVRRYVN AKPFCVDCYKLKFTENGIRKRMYSARGRKA DSDINIKNSTNKISKTHFNIAIREGFILDKSLK KQRSKRIKKLLELKRKLQEFIDIRQGQMVLC KIKNQRVDKFIHPSWLKRDKKLEEFRGYSLS VVEGKIKIFNRNLRREDSLRQRGHVNFKANR IMLDKSVRFLDGGKVNFNLNKGLPKEYLLDL PKKENKLSWLNEKISLIKLPKYAYLLRRE GSFFIQYTIENVPKTFSDYLGAIGIDRGISHIAV CTFVSKNGVNKAPVFFSSGEILKLKSLQKQR DLFLRGKHNKIRKKSNNRDNINLILHKYS RNIVNLAKSEKAFIVFEKLEKIKKSFRKMSKS LQYKLSQFTFKLSDLVEYKAKIEGIKVDYV PPEYTSKECSHCGEKVDTRPFNGNSSLFKC NKCRVQLNADYNASINIAKKSLNISN
编码 Cas14b.4 可编程切口酶的 DNA	SEQ ID NO: 94	ATGCCCAAGCAGGATCTGGTCACCACCGGC ATCAAGTTCAAGCTGGACGTGGACAAAGA GACACGGAAGAACTGGACGACTACTTCG ACGAGTACGGCAAGGCCATCAACTTCGCCG TGAAGATCATCCAGAAGAACCTGAAAGAG GACCGCTTCGCCGGCAAGATTGCCCTGGGC GAAGATAAGAAGCCCCTGCTGGACAAGGA CGGCAAGAAGATCTACAACCTACCCCAACG AGAGCTGCTCCTGCGGCAATCAAGTGCGGA GATACGTGAACGCCAAGCCTTTCTGCGTGG ACTGCTACAAGCTGAAGTTCACCGAGAACG GCATCCGGAAGCGGATGTACTCTGCCAGAG GAAGAAAGGCCGACAGCGACATCAACATC AAGAACAGCACCAACAAGATCAGCAAGAC CCACTTCAACTACGCCATCAGAGAGGGCTT

物种	SEQ ID NO	序列
		CATCCTGGACAAGAGCCTGAAGAAGCAGC GGAGCAAGCGGATCAAGAAGCTGCTGGAA CTGAAGCGGAAGCTGCAAGAGTTCATCGA CATCCGGCAGGGCCAGATGGTGCTGTGCC CAAGATCAAGAACCAGAGAGTGGACAAGT TCATTACCCCCAGCTGGCTGAAGAGAGACA AAAAGCTGGAAGAATTCCGGGGCTACAGC CTGAGCGTGGTGGGAAGGCAAGATTAAGAT CTTCAACCGGAACATCCTGCGCGAAGAGG ACTCCCTGAGACAGAGGGGCCACGTGAAC TTTAAGGCCAACCGGATCATGCTGGATAAG AGCGTGCGGTTCTTGACGGCGGCAAGT GAATTTCAACCTGAACAAGGGCCTGCCGAA AGAGTACCTGCTGGATCTGCCCAAGAAAG AGAACAAGCTGTCTGGCTGAACGAAAAG ATCAGCCTGATCAAGCTGCAGAAGCCTAAG TACGCCTACCTGCTGAGAAGAGAGGGCAG CTTTTTCATCCAGTACACCATCGAGAACGT GCCCAAGACCTTCAGCGATTACCTGGGCGC CATCGGCATCGACAGAGGCATCTCTCACAT TGCCGTGTGCACCTTCGTGTCCAAGAACGG CGTGAACAAGGCCCTGTGTTCTTCAGCTC TGCGGAGATCCTGAAGCTGAAAAGCCTGC AGAAACAGAGGGACCTGTTCTGCGGGGC AAGCACAACAAAATCCGGAAGAAAAGCAA CATGCGGAACATCGACAACAAGATTAACCT GATCCTGCACAAGTACAGCCGCAACATCGT GAACCTGGCCAAGAGCGAGAAGGCCTTTA TCGTGTTGAGAAGCTCGAGAAGATCAAA AAGTCCCGGTTCAAGATGAGCAAGTCCCTG CAGTATAAGCTGAGCCAGTTCACCTTCAAG AAAGTGAAGCGACCTGGTCGAGTACAAGGC CAAGATCGAGGGCATCAAGGTGGACTACG TGCCACCTGAGTACACCAGCAAAGAGTGCT CTCACTGCGGCGAGAAGGTGGACACCCAG AGGCCTTTCAACGGCAACAGCAGCCTGTTC AAGTGTAACAAGTGCCGGGTGCAGCTGAA CGCCGACTACAACGCCAGCATCAATATCGC CAAGAAGTCCCTGAACATCAGCAACAAC TGA
Cas14.b10 可编程切口酶	SEQ ID NO: 17	MEKNNSEQTSITTGIKFKLKDKEKELNN YFDEYGKAINFAVRHMQMLNDDRLAGKYKR DEKGKPILGEDGKKILEIPNDFCSCGNQVNH VNGVSFCQECYKKRFSNGIRKRMYSAGR KAEQDINIKNSTNKISKTHFNIAIREAFNL SIKKQREKRFFKKLKDMMKRKLQEFLEIRD GKR VICPKIEKQKVERYIHPSWINKEKKLE EFRGY SLSIVNSKISFDRNIQREEKSLKEK GQINFKA QRLMLDKSVKFLKDNKVSFTIS KELPKTFELD LPKKEKKNLWLNKLEIKN QPKYAYLLRK ENNIFLQYTLDSEIHS EYSGAVGIDRGVSHI AVYTFLDKDGK NERPFLSSSGILRLKNLQKE RDKFLR KKHNRKIRKKGNMNRNIEQKINLILHE

物种	SEQ ID NO	序列
		YSKQIVNFAKDKNAFIVFELLEKPKKSRRMS KKIQYKLSQFTFKLSDLVDYKAKREGIKVIY VEPAYTSKDCSHCGERVNTQRPFGNFSLFK CNKCGIVLNSDYNASLNIARKGLNISAN
编码 Cas14b.10 可编程切口酶 的 DNA	SEQ ID NO: 95	ATGGAAAAGAACAACAGCGAGCAGACCAG CATCACCACCGGCATCAAGTTCAAGCTGAA GCTGGACAAAAGAGACAAAAGAGAAGCTGA ACAACACTTTCGACGAGTACGGCAAGGCC ATCAACTTCGCCGTGCGGATCATCCAGATG CAGCTGAACGACGATAGACTGGCCGGCAA GTACAAGCGGGACGAGAAGGGAAAGCCTA TCCTGGGCGAAGATGGCAAGAAGATCCTG GAAATCCCCAACGACTTCTGCAGCTGCGGC AATCAAGTGAACCACTACGTGAACGGCGT GTCCTTCTGCCAAGAGTGCTACAAGAAGCG GTTTCAGCGAGAACGGCATCCGGAAGAGAA TGTCAGCGCCAAGGGCAGAAAGGCCGAG CAGGACATCAACATCAAGAACAGCACCAA CAAGATCAGCAAGACCCACTTCAACTACGC CATCAGAGAGGCCTTCAACCTGGACAAGA GCATCAAGAAGCAGAGGGAAAAGCGCTTC AAGAACTGAAGGACATGAAGCGGAAGCT GCAAGAGTTCTCGAGATCCGCGACGGCA AGAGAGTGATCTGCCCCAAGATCGAGAAG CAGAAGGTGGAACGGTACATTACCCCCAG CTGGATCAACAAAGAGAAGAAGCTGGAAG AATTCGGGGCTACAGCCTGAGCATCGTGA ACAGCAAGATCAAGAGCTTCGACCGGAAC ATCCAGCGCGAGGAAAAGAGCCTGAAAGA GAAGGGCCAGATCAACTTCAAGGCCACG GGCTGATGCTGGATAAGAGCGTGAAGTTCC TCAAGGACAACAAGGTGTCCTTACCATCA GCAAAGAGCTGCCCAAGACCTTCGAGCTG GACCTGCCTAAGAAAGAGAAAAAACTGAA CTGGCTGAACGAGAAGCTCGAGATCATTA GAACCAGAAGCCGAAGTACGCCTACCTGCT GCGCAAAGAGAAACAACATCTTCTGCAGTA CACCTGGACAGCATCCCCGAGATCCACAG CGAATATTCTGGCGCCGTGGGCATCGATAG AGGCGTGTCACATATCGCCGTGTACACCTT CCTGGATAAGGACGGAAAGAACGAGCGGC CATTCTTCTGAGCAGCAGCGGCATCCTGC GGCTGAAGAACCTGCAGAAAGAGCGGGAC AAGTTCCTGCGGAAGAAGCACAAACAAAT CCGAAAAAAGGGCAACATGCGGAACATCG AGCAGAAGATCAACCTGATCCTGCACGAGT ACTCCAAGCAGATCGTGAACCTTGCCAAGG ACAAGAACGCCTTCATCGTGTTCGAGCTGC TGAAAAAGCCCAAGAAAAGCCGCGAGCGG ATGAGCAAGAAAATCCAGTACAAGCTGTC CCAGTTCACCTTCAAAAAGCTGAGCGACCT GGTGGACTACAAGGCAAGCGCGAGGGCA TCAAAGTGATCTACGTGGAACCCGCCTACA

物种	SEQ ID NO	序列
		CCAGCAAGGACTGTTCTCACTGTGGCGAGA GAGTGAACACCCAGCGGCCTTTCAACGGCA ACTTCAGCCTGTTCAAGTGCAACAAGTGCG GCATCGTGCTGAACAGCGACTACAACGCCA GCCTGAATATCGCCCGGAAGGGCCTGAAC ATCAGCGCCAATTGA
超螺旋质粒 DNA 靶序列	SEQ ID NO: 96	TATTAATACTCGTATTGCTGTTTCGATT AT
Cas14a.3 crRNA	SEQ ID NO: 97	GUUGCAGAACCCGAAUAGACGAAUGA AGGAAUGCAACUAUUAUUACUCGUA UUGCUGU
Cas14a.3 tracr3	SEQ ID NO: 98	CGAUUCCUCCCUACAGUAGUUAGGUA UAGCCGAAAGGUAGAGACUAAAUCUG UAGUUGGAGUGGGCCGCUUGCAUCGG CCUAAAGUUGAGAAGUGUCAGACUCU GAUAACCCUCAACGACGAUAUUCUUU AUUUC
Cas14a.3 tracr4	SEQ ID NO: 99	CGAUUCCUCCCUACAGUAGUUAGGUA UAGCCGAAAGGUAGAGACUAAAUCUG UAGUUGGAGUGGGCCGCUUGCAUCGG CCUAAAGUUGAGAAGUGUCAGACUCU GAUAACCCUCAACGACGAUAUUCUUU AUUUCGGUUCAAAGUUCUGCACAAAA CAGGUGAGUCCUUAUAAACCGGUG
Cas14b.3 crRNA	SEQ ID NO: 100	CUUUCAUACUCAGAACAAAGGGAUUA AGGAAUGCAACUAUUAUUACUCGUA UUGCUGU
Cas14b.3 tracr 2	SEQ ID NO: 101	GAACAGACCAAUCUUUAAUUCGCUUC UGAUUUAAAAAAUCAGAAUCUCUUUA UAAAUAGUAUUACAAAAAGUGUACAU UCCAAAAUCCGAAAGCAGAAUUGACC UUUUUAAG
Cas14b.3 tracr 3	SEQ ID NO: 151	AUGCGGAAGAUUUGGCGUUGUUGUAA CGCAAUAAGGGGUAACCCUGAAAAGG UUUGAAAUCAUAUAAACCUAGUUUUA UUUGAGUUUAGGCUCAGAUAAAUGA ACAGACCAAUCUUUAAUUCGCUUCUG AUUUAAAAAAUCAGAAUCUCUUUUAU AAUAGUAUUACAAAAAGUGUACAUUC CAAAAUCCGAAAGCAGAAUUGACC
Cas14b.4 crRNA	SEQ ID NO: 102	AUUUCAUACUCAGAACAAAGGGAUUA AGGAAUGCAACUAUUAUUACUCGUA UUGCUGU
Cas14b.4 tracr1	SEQ ID NO: 103	UUGGUUAAGCCAAGAU AUGGAAUGCC AUUGUAAUAUUAUGGUGUUGACUAG

物种	SEQ ID NO	序列
		UUUAGAUUUAAAACAAUCUUCGAUGGC UAUAUGCGGAAGGUUUGGCGUCGUUG UAACGCAAUAGGGGGCGACCCCGAAA AGGUUUGAUUCAUAUCAAACUUAGU UUUGUUUAAGUUAAGGCUUAUUCAAA AUGAACAGACCAAUUCUAAUACCUU UAUCUGACU
Cas14b.10 crRNA	SEQ ID NO: 104	GUUGC GCGAAUAGAAUAAAGGAAUUA AGGAAUGCAACUAUUAAAUACUCGUA UUGCUGU
Cas14b.10 tracr1	SEQ ID NO: 105	AGUGUAAGUUGAAGUGUGAGCUUAUG GAUUAUUUUUACAAAUAUUACUGA CUUACUAAGAUUCUUGAGGGUAUAC CCAAAAGAUUGGCGUUGUUGCAACG CAUAAGAUGUAAAUCUGAAAAGGUU UGAAAUCAUAUAAGUAAUUUUUUAUUUG AGUUUCGGCUUGAGUAAAUGAACAG ACCAAUUUUUAAUUUCGUUCAUAUCA UCGCAACUA

[0173] 在存在含有紧邻TTTA PAM的3'位点的SEQ ID NO:96靶序列的超螺旋质粒DNA的情况下,Cas14-指导RNA核糖核蛋白复合体在Tris,pH 7.9缓冲液(50mM乙酸钾、20mM tris-乙酸盐、10mM乙酸镁、100μg/ml BSA)中于37℃下孵育60分钟。超螺旋质粒DNA的序列是

GTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCA
GTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGAT
TTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAA
GTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTT
GCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGC GC
AACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTC
GTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGC
GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCC
TTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTT
ATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCA
TGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACC

AAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTG
CCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACT
TTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAAC
CTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCA
CTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCG
TTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAA
GGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTC
CTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATG
AGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAG
GGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAA
GAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTAT
CACGAGGCCCTTTCGTCTCGCGCGTTTCGGTGATGACGGTGAAA
ACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTG
CCATGGACATGTTTATTATTAAATACTCGTATTGCTGTTGATT
ATGACCGAATTCCCTGTCGTGCCAGCTGCATTAATGAATCGGCC
AACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCT
TCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTTCGGCTGCGGCG
AGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACA
GAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGC
CAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGT
TTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGA
CGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGAT
ACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTT
CCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCG
GGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAG
TTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAAC
CCCCGTTTCAGCCCGACCGCTGCGCCTTATCCGGTAACCTATCGT
CTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGC
AGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGT

GCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAG
AAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCT
TCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCAC
CGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGC
GCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACG
GGGTCTGACGCTCAGTGGAACGAAAACCTCACGTTAAGGGATTTT
GGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAA
ATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAA
ACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTAT
CTCAGCGATCTGTCTATTTCGTTTCATCCATAGTTGCCTGACTCCC
CGTC (SEQ ID NO:152;靶序列用粗体和下划线表示)。在60分钟的孵育过程中,

核糖核蛋白使质粒产生切口并进行切割。60分钟后,用1mg/ml蛋白酶K、0.08%SDS和15mM EDTA猝灭反应。

[0174] 然后通过琼脂糖凝胶上的凝胶电泳分析质粒的产生切口和切割。在每个测定中经过顺式和反式切割的质粒百分比如图5所示。测定显示,Cas14a和Cas14b直系同源物之间的产生切口和切割活性不同,并且可编程切口酶产生切口和切割活性依赖于tracrRNA序列。

实施例8

Cas14a和Cas14b产生切口和切割活性对tracrRNA序列的依赖性

[0175] 本实施例显示了不同可编程切口酶的产生切口切割活性对tracrRNA序列的依赖性。SEQ ID NO:1、SEQ ID NO:11和SEQ ID NO:17分别与十八种不同的指导RNA复合。指导RNA含有靶向SEQ ID NO:96的相同间隔区序列和不同的tracr序列。表4总结了分析中使用的可编程切口酶和指导RNA的序列。

表4-指导核酸、靶核酸和可编程切口酶序列

物种	SEQ ID NO	序列
Cas14a.3 可编程切口酶	SEQ ID NO: 1	MEVQKTVMKTLSLRILRPLYSQEIEKEIKE EKERRKQAGGTGELDGGFYKKLEKKHSE MFSFDRLNLLNQLQREIAKVYNHAISEL YIATIAQGNKSNKHYYISSIVYNRAYGYFYN AYIALGICSKVEANFRSNELLTQQSALPTA

物种	SEQ ID NO	序列
		KSDNFPIVLHKQKGAEGEDGGFRISTEGSD LIFEIPIPFYEYNGENRKEPYKWVKKGGQK PVLKLILSTFRRQRNKGWAKDEGTDAEIR KVTEGKYQVSQIEINRGKKLGEHQKWFA NFSIEQPIYERKPNRSIVGGLDVGIRSPVC AINNSFSRYSVDSNDVFKFSKQVFAFRRL LSKNSLKRKGHGAHKLPIITEMTEKNDK FRKKIIRWAKEVTNFFVKNQVGIVQIEDL STMKDREDHFFNQYLRFWPPYYQMQLI ENKLKEYGIEVKRVQAKYTSQLCSNPNCR YWNMYFNFEYRKVNKFPKFKCEKCNLEIS ADYNAARNLSTPDIEKFVAKATKGINLPE K
Cas14.b4 可编程 切口酶	SEQ ID NO: 11	MPKQDLVTTGIKFKLDVDKETRKKLDDY FDEYGKAINFAVKIIQKNLKEDRFAGKIAL GEDKKPLLDKDGKKIYNYPNESCSCGNQV RRYVNAKPFVDCYKLKFTENGIRKRM SARGRKADSDINIKNSTNKISKTHFNIAIR EGFILDKSLKKQRSKRIKKLLELKRKLQEF IDIRQGMVLCPIKNQRVDKFIHPSWLK RDKKLEEFRGYSLSVVEGKIKIFNRNIRE EDSLRQRGHVNFKANRIMLDKSVRFLDG GKVNFNLNKGLPKYLLDLPKKENKLSW LNEKISLIKLPKYAYLLRREGSFFIYTI ENVPKTFSDYLGAIGIDRGISHIAVCTFVSK NGVNKAPVVFSSGEILKLKSLQKQRDFLR GKHNKIRKKSNNMRNIDNKINLILHKYSRNI VNLAKESEAFIVFEKLEKIKKSRFKMSKSL QYKLSQFTFKKLSDLVEYKAKIEGIKVDY VPPEYTSKECSHCGEKVDTPRPFNGNSSLF KCNKCRVQLNADYNASINIAKSLNISN
Cas14.b10 可编程 切口酶	SEQ ID NO: 17	MEKNNSEQTSITTGIKFKLKDKEKELN NYFDEYGKAINFAVRIIQQMLNDDRLAGK YKRDEKGKPILGEDGKKILEIPNDFCSCGN QVNHYVNGVSFCQECYKKRFSENGIRKR MYSAGRKAEQDINIKNSTNKISKTHFNIAIR AFNLDKSIKKQREKRFFKLLKDMKRK LQEFLEIRDGKRVICPKIEKQKVERYIHPS WINKEKKLEEFRGYSLIVNSKIKSFDRNI QREEKSLKEKGQINFKAQRLMLDKSVKFL KDNKVSFTISKELPKTFELDLPKKEKKNL WLNEKLEIKNQPKYAYLLRKENNIFLQ YTLDSIPEIHSEYSGAVGIDRGVSHIAVYTF LDKDGKNERPFFLSSSGILRLKNLQKERDK FLRKKHKNIRKKGNNMRNIEQKINLILHEYS KQIVNFAKDKNFIVFELLEKPKKSRERM SKKIYKLSQFTFKKLSDLVDYKAKREGI KVIYVEPAYTSKDCSHCGERVNTQRPFG NFSLFKCNKCGIVLNSDYNASLNIARKGL NISAN

物种	SEQ ID NO	序列
超螺旋质粒 DNA 靶序列	SEQ ID NO: 96	TATTAAATACTCGTATTGCTGTTCTGA TTAT
Cas14a.3 crRNA	SEQ ID NO: 97	GUUGCAGAACCCGAAUAGACGAAU GAAGGAAUGCAACUAUUAUUACU CGUAUUGCUGU
Cas14a.3 tracr1	SEQ ID NO: 106	CGAUUCCUCCCUACAGUAGUUAGGU AUAGCCGAAAGGUAGAGACUAAAU CUGUAGUUGGAGUGGGCCGCUUGC AUCGGCCUAAAGU
Cas14a.3 tracr2	SEQ ID NO: 107	CGAUUCCUCCCUACAGUAGUUAGGU AUAGCCGAAAGGUAGAGACUAAAU C
Cas14a.3 tracr3	SEQ ID NO: 98	CGAUUCCUCCCUACAGUAGUUAGGU AUAGCCGAAAGGUAGAGACUAAAU CUGUAGUUGGAGUGGGCCGCUUGC AUCGGCCUAAAGUUGAGAAGUGUC AGACUCUGAUAAACCCUCAACGACGA UAUUCUUUAUUUC
Cas14a.3 tracr4	SEQ ID NO: 99	CGAUUCCUCCCUACAGUAGUUAGGU AUAGCCGAAAGGUAGAGACUAAAU CUGUAGUUGGAGUGGGCCGCUUGC AUCGGCCUAAAGUUGAGAAGUGUC AGACUCUGAUAAACCCUCAACGACGA UAUUCUUUAUUUCGGUUCAAAGUU CUGCACAAAACAGGUGAGUCCUUAU AAACCGGUG
Cas14a.3 tracr5	SEQ ID NO: 108	GUUGGAGUGGGCCGCUUGCAUCGGC CUAAAGUUGAGAAGUGUCAGACUC UGAUAAACCCUCAACGACGAUAUUCU UUAUUUCGGUUC
Cas14a.3 tracr6	SEQ ID NO: 109	CGAUUCCUCCCUACAGUAGUUAGGU AUAGCCGAAAGGUAGAGACUAAAU CUGUAGUUGGAGUGGGCCGCUUGC AUCGGCCUAAAGUUGAGAAGUGUC
Cas14a.3 tracr7	SEQ ID NO: 110	CAUCGGCCUAAAGUUGAGAAGUGU CAGACUCUGAUAAACCCUCAACGACG AUUUCUUUAUUUCGGUUCAAAGU UCUGCACAAAACAGGUGAGUCCUUA UAAACCGGUGUGCAGAACGCCGGCU CACCUUUUCCUUCUUC
Cas14a.3 tracr8	SEQ ID NO: 111	UAGGUUAUAGCCGAAAGGUAGAGAC UAAUUCUGUAGUUGGAGUGGGCCG CUUGCAUCGGCCUAAAGUUGAGAA

物种	SEQ ID NO	序列
		GUGUCAGACUCUGAUAAACCCUCAAC GACGAUAUUCUUUAUUUC
Cas14a.3 tracr9	SEQ ID NO: 112	CUACAGUAGUUAGGUUAUAGCCGAA AGGUAGAGACUAAAUCUGUAGUUG GAGUGGGCCGCUUGCAUCGGCCUAA AGUUGAGAAGUGUCAGACUCUGAU AACCCUCAACGACGAUAUUCUUUAU UUC
Cas14a.3 tracr10	SEQ ID NO: 113	CUCCCUACAGUAGUUAGGUUAUAGCC GAAAGGUAGAGACUAAAUCUGUAG UUGGAGUGGGCCGCUUGCAUCGGCC UAAAGUUGAGAAGUGUCAGACUCU GAUAACCCUCAACGACGAUAUUCUU UAUUUC
Cas14a.3 tracr11	SEQ ID NO: 114	GCACACGAUUCCUCCCUACAGUAGU UAGGUUAUAGCCGAAAGGUAGAGAC UAAAUCUGUAGUUGGAGUGGGCCG CUUGCAUCGGCCUAAAGUUGAGAA GUGUCAGACUCUGAUAAACCCUCAAC GACGAUAUUCUUUAUUUC
Cas14a.3 tracr12	SEQ ID NO: 115	UUGAUGCACACGAUUCCUCCCUACA GUAGUUAGGUUAUAGCCGAAAGGUA GAGACUAAAUCUGUAGUUGGAGUG GGCCGCUUGCAUCGGCCUAAAGUUG AGAAGUGUCAGACUCUGAUAAACCCU CAACGACGAUAUUCUUUAUUUC
Cas14a.3 tracr13	SEQ ID NO: 116	CGAUUCCUCCCUACAGUAGUUAGGU AUAGCCGAAAGGUAGAGACUAAAU CUGUAGUUGGAGUGGGCCGCUUGC AUCGGCCUAAAGUUGAGAAGUGUC AGACUCUGAUAAACCCUCAACGACGA UAUUCUUU
Cas14a.3 tracr14	SEQ ID NO: 117	CGAUUCCUCCCUACAGUAGUUAGGU AUAGCCGAAAGGUAGAGACUAAAU CUGUAGUUGGAGUGGGCCGCUUGC AUCGGCCUAAAGUUGAGAAGUGUC AGACUCUGAUAAACCCUCAACGACGA UAUUCUUUAUUUCGGUUC
Cas14a.3 tracr15	SEQ ID NO: 118	CGAUUCCUCCCUACAGUAGUUAGGU AUAGCCGAAAGGUAGAGACUAAAU CUGUAGUUGGAGUGGGCCGCUUGC AUCGGCCUAAAGUUGAGAAGUGUC AGACUCUGAUAAACCCUCAACGACGA UAUUCUUUAUUUCGGUUCAAAGUU CUGCACAAA

物种	SEQ ID NO	序列
Cas14a.3 tracr16	SEQ ID NO: 119	GUUGGAGUGGGCCGCUUGCAUCGGC CUAAAGUUGAGAAGUGUCAGACUC UGAUAACCCUCAACGACGAUAUUCU UUAUUUC
Cas14a.3 tracr17	SEQ ID NO: 120	CGAUUCCUCCCUACAGUAGUUAGGU AUAGCCGAAAGGUAGAGACUAAAU CUGUAGUUGGAGUGGGCCGCUUGC AUCGGCCUAAAGUUGAGAAGUGUC AGACUCUGAUAACCCUCAACGACGA UAUUCUUUAUUUCGGUUCAAAGUU CUGCACAAAACAGGUGAGUCCUUA
Cas14a.3 tracr18	SEQ ID NO: 121	CGAUUCCUCCCUACAGUAGUUAGGU AUAGCCGAAAGGUAGAGACUAAAU CUGUAGUUGGAGUGGGCCGCUUGC AUCGGCCUAAAGUUGAGAAGUGUC AGACUCUGAUAACCCUCAACGACGA UAUUCUUUAUUUCGGUUCAAAGUU CUGCACAAAACAGGUGAGUCCUUAU AAACCGGUGUGCAGAACG
Cas14b.4 crRNA	SEQ ID NO: 102	AUUUCAUACUCAGAACAAAGGGAU UAAGGAAUGCAACUAUUAAAUACU CGUAUUGCUGU
Cas14b.4 tracr 1	SEQ ID NO: 103	UUGGUUAAGCCAAGAU AUGGAAUG CCAUUGUAAUAUUAUGGUGUUGAC UUAGUUUAGAUUUAAACAAUCUUC GAUGGCUAU AUGCGGAAGGUUUGG CGUCGUUGUAACGCAAUAGGGGGC GACCCCGAAAAGGUUUGAUUAUCAU AUCAAACUUAGUUUUGUUUAAGUU AAGGCUUAUUCAAAAUGAACAGAC CAAUUCUUAUACCUUUAUCUGACU
Cas14b.4 tracr 2	SEQ ID NO: 122	UUGGUUAAGCCAAGAU AUGGAAUG CCAUUGUAAUAUUAUGGUGUUGAC UUAGUUUAGAUUUAAACAAUCUUC GAUGGCUAU AUGCGGAAGGUUUGG CGUCGUUGUAACGCAAUAGGGGGC GACCCCGAAAAGGUUUGAUUAUCAU AUCAAACUUAGUUUUGUUUAAGUU AA
Cas14b.4 tracr 3	SEQ ID NO: 123	UUGGUUAAGCCAAGAU AUGGAAUG CCAUUGUAAUAUUAUGGUGUUGAC UUAGUUUAGAUUUAAACAAUCUUC GAUGGCUAU AUGCGGAAGGUUUGG CGUCGUUGUAACGCAAUAGGGGGC

物种	SEQ ID NO	序列
		GACCCCGAAAAGGUUUGAUUAUCAU AUCAAAC
Cas14b.4 tracr 4	SEQ ID NO: 124	AUUGUAAUAUUAUGGUGUUGACUU AGUUUAGAUUUAAAACAAUCUUCGA UGGCUAU AUGCGGAAGGUUUGGCG UCGUUGUAACGCAAUAGGGGGCGA CCCCGAAAAGGUUUGAUUAUCAUAUC AAAC
Cas14b.4 tracr 5	SEQ ID NO: 125	AGCUCUUUUGUUGGUUAAGCCAAG AUAUGGAAUGCCAUGUAAUAUUA UGGUGUUGACUUAGUUUAGAUUUUA AACAAUCUUCGAUGGCUAU AUGCG GAAGGUUUGGCGUCGUUGUAACGC AAUAGGGGGCGACCCCGAAAAGGU UUGAUUAUCAUAUCAAACUUAGUUU UGUUUAAGUUAAGGCUUAUUCAAA AUGAACAGACCAAUUCUUAUACCU UUAUCUGACU
Cas14b.4 tracr 6	SEQ ID NO: 126	UCCAAUAACUAGCUCUUUUGUUGG UUAAGCCAAGAU AUGGAAUGCCA UGUAAUAUUAUGGUGUUGACUUAG UUUAGAUUUAAAACAAUCUUCGAUG GCUAU AUGCGGAAGGUUUGGCGUC GUUGUAACGCAAUAGGGGGCGACCC CGAAAAGGUUUGAUUAUCAUAUCAA ACUUAGUUUUGUUUAAGUUAAGGC UUAUUCAAA AUGAACAGACCAAU CUAAUACCUUUAUCUGACU
Cas14b.4 tracr 7	SEQ ID NO: 127	CAAGAU AUGGAAUGCCAUGUAAU AUUAUGGUGUUGACUUAGUUUAGA UUUAAAACAAUCUUCGAUGGCUAU UGCGGAAGGUUUGGCGUCGUUGUA ACGCAAUAGGGGGCGACCCCGAAAA GGUUUGAUUAUCAUAUCAAACUUAG UUUUGUUUAAGUUAAGGCUUAUUC AAA AUGAACAGACCAAUUCUUAU ACCUUUAUCUGACU
Cas14b.4 tracr 8	SEQ ID NO: 128	AAUGCCAUGUAAUAUUAUGGUGU UGACUUAGUUUAGAUUUAAAACAAU CUUCGAUGGCUAU AUGCGGAAGGU UUGGCGUCGUUGUAACGCAAUAGG GGCGACCCCGAAAAGGUUUGAUUA UCAUAUCAAACUUAGUUUUGUUUA AGUUAAGGCUUAUUCAAA AUGAAC AGACCAAUUCUUAUACCUUUAUCU GACU

物种	SEQ ID NO	序列
Cas14b.4 tracr 9	SEQ ID NO: 129	UAAUAUUAUGGUGUUGACUAGUU UAGAUUUAAACAAUCUUCGAUGGC UAUAUGCGGAAGGUUUGGCGUCGU UGUAAACGCAAUAGGGGGCGACCCCG AAAAGGUUUGAUUCAUAUCAAAAC UUAGUUUUGUUUAAGUUAAGGCUU AUUCAAAAUGAACAGACCAAUUCU UAAUACCUUUAUCUGACU
Cas14b.4 tracr 10	SEQ ID NO: 130	UUGGUUAAGCCAAGAU AUGGAAUG CCAUUGUAAUAUUAUGGUGUUGAC UUAGUUUAGAUUUAAACAAUCUUC GAUGGCUAU AUGCGGAAGGUUUGG CGUCGUUGUAAACGCAAUAGGGGGC GACCCCGAAAAGGUUUGAUUCAU AUCAAACUUAGUUUUGUUUAAGUU AAGGCUUAUUCAAAUGAACAGAC CAAUUCUUAUACCUUUAUCUGACU AAACGUCAGA
Cas14b.4 tracr 11	SEQ ID NO: 131	UUGGUUAAGCCAAGAU AUGGAAUG CCAUUGUAAUAUUAUGGUGUUGAC UUAGUUUAGAUUUAAACAAUCUUC GAUGGCUAU AUGCGGAAGGUUUGG CGUCGUUGUAAACGCAAUAGGGGGC GACCCCGAAAAGGUUUGAUUCAU AUCAAACUUAGUUUUGUUUAAGUU AAGGCUUAUUCAAAUGAACAGAC CAAUUCUUAUACCUUUAUCUGACU AAACGUCAGAACAUCUUUAU
Cas14b.4 tracr 12	SEQ ID NO: 132	UUGGUUAAGCCAAGAU AUGGAAUG CCAUUGUAAUAUUAUGGUGUUGAC UUAGUUUAGAUUUAAACAAUCUUC GAUGGCUAU AUGCGGAAGGUUUGG CGUCGUUGUAAACGCAAUAGGGGGC GACCCCGAAAAGGUUUGAUUCAU AUCAAACUUAGUUUUGUUUAAGUU AAGGCUUAUUCAAAUGAACAGAC CAAUUCUUAUACCUUUAUCUGACU AAACGUCAGAACAUCUUUAUAAAU CAAUUU
Cas14b.4 tracr 13	SEQ ID NO: 133	UUGGUUAAGCCAAGAU AUGGAAUG CCAUUGUAAUAUUAUGGUGUUGAC UUAGUUUAGAUUUAAACAAUCUUC GAUGGCUAU AUGCGGAAGGUUUGG CGUCGUUGUAAACGCAAUAGGGGGC GACCCCGAAAAGGUUUGAUUCAU AUCAAACUUAGUUUUGUUUAAGUU

物种	SEQ ID NO	序列
		AAGGCUUAUUCAAAAUGAACAGAC CAAUUCUUAUACCUUUAUCUGACU AAACGUCAGAACAUCUUUAUAAU CAUUUACAAAAUGA
Cas14b.4 tracr 14	SEQ ID NO: 134	UUGGUUAAGCCAAGAU AUGGAAUG CCAUUGUAAUUAU AUGGUGUUGAC UUAGUUUAGAUUUAACAAUCUUC GAUGGCUAU AUGCGGAAGGUUUGG CGUCGUUGUAACGCAAUAGGGGGC GACCCCGAAAAGGUUUGAUUAUCAU AUCAAACUUAGUUUUGUUUAAGUU AAGGCUUAUUCAAAAUGAACAGAC CAAUUCUUAUACCU
Cas14b.4 tracr 15	SEQ ID NO: 135	UUGGUUAAGCCAAGAU AUGGAAUG CCAUUGUAAUUAU AUGGUGUUGAC UUAGUUUAGAUUUAACAAUCUUC GAUGGCUAU AUGCGGAAGGUUUGG CGUCGUUGUAACGCAAUAGGGGGC GACCCCGAAAAGGUUUGAUUAUCAU AUCAAACUUAGUUUUGUUUAAGUU AAGGCUUAUUCAAAAUGAACAGAC CAAUU
Cas14b.4 tracr 16	SEQ ID NO: 136	UUGGUUAAGCCAAGAU AUGGAAUG CCAUUGUAAUUAU AUGGUGUUGAC UUAGUUUAGAUUUAACAAUCUUC GAUGGCUAU AUGCGGAAGGUUUGG CGUCGUUGUAACGCAAUAGGGGGC GACCCCGAAAAGGUUUGAUUAUCAU AUCAAACUUAGUUUUGUUUAAGUU AAGGCUUAUUCAAA
Cas14b.10 crRNA	SEQ ID NO: 104	GUUGCGCGAAUAGAAUAAAGGAAU UAAGGAAUGCAACUAUUAAAUACU CGUAUUGCUGU
Cas14b.10 tracr 1	SEQ ID NO: 105	AGUGUAAGUUGAAGUGUGAGCUUA UGGAUUUAUUUUACAAAAUAAUA CUGACUUACUAAGAUUUCUUGAGG GUUAACCCAAAAAGAUUGGCGUUG UUGCAACGCAAUAAGAUGUAAAUC UGAAAAGGUUUGAAUCAUAUAAG UAAUUUUUUUUGAGUUUCGGCUUG AGUAAAUGAACAGACCAAUUUUU AAUUUCGUUCAUAUCAUCGCAACUA
Cas14b.10 tracr 2	SEQ ID NO: 137	AUUUUUAUUUGAGUUUCGGCUUGAG UAAAUGAACAGACCAAUUUUUAA UUUCGUUCAUAUCAUCGCAACUAAG AAUCCUUUAAAAAGGAAAUGCAG

物种	SEQ ID NO	序列
		AAAAUGGACAUUCCAAAAUUCAAA ACAAAAUUCGGUUUUUAAGACCA GAAAG
Cas14b.10 tracr 3	SEQ ID NO: 138	ACUGACUUACUAAGAUUUCUUGAG GGUAUACCCAAAAAGAUUGGCGUU GUUGCAACGCAAUAAGAUGUAAAU CUGAAAAGGUUUGAAAUCAUAUAA GUAAUUUUUAAUUUGAGUUUCGGCUU GAGUAAAAUGAACAGACCAAUUUU UAAUUUCGUUCAUAUCAUCGCAACU A
Cas14b.10 tracr 4	SEQ ID NO: 139	GAAGUGUGAGCUUAUGGAUUAUUA UUUACAAAAUAAUACUGACUUACU AAGAUUUCUUGAGGGUAUACCCAA AAAGAUUGGCGUUGUUGCAACGCA AUAAGAUGUAAAUCUGAAAAGGUU UGAAAUCAUAUAAGUAAUUUUAAU UGAGUUUCGGCUUGAGUAAAAUGA ACAGACCAAUUUUUAAUUUCGUUC AUAUCAUCGCAACUA
Cas14b.10 tracr 5	SEQ ID NO: 140	CUUAUGGAUUAUUAUUUACAAAAU AAUACUGACUUACUAAGAUUUCUU GAGGGUAUACCCAAAAAGAUUGGC GUUGUUGCAACGCAAUAAGAUGUA AAUCUGAAAAGGUUUGAAAUCAUA UAAGUAAUUUUAAUUUGAGUUUCGG CUUGAGUAAAAUGAACAGACCAAU UUUUAAUUUCGUUCAUAUCAUCGC AACUA
Cas14b.10 tracr 6	SEQ ID NO: 141	AUUAUUUACAAAAUAAUACUGACU UACUAAGAUUUCUUGAGGGUAUAC CCAAAAAGAUUGGCGUUGUUGCAA CGCAAUAAGAUGUAAAUCUGAAAA GGUUUGAAAUCAUAUAAGUAAUUU UAUUUGAGUUUCGGCUUGAGUAAA AUGAACAGACCAAUUUUUAAUUUC GUUCAUAUCAUCGCAACUA
Cas14b.10 tracr 7	SEQ ID NO: 142	AAAUAAUACUGACUUACUAAGAU UCUUGAGGGUAUACCCAAAAAGAU UGGCGUUGUUGCAACGCAAUAAGA UGUAAAUCUGAAAAGGUUUGAAAU CAUAUAAGUAAUUUUAAUUUGAGUU UCGGCUUGAGUAAAAUGAACAGAC CAAUUUUUAAUUUCGUUCAUAUCA UCGCAACUA

物种	SEQ ID NO	序列
Cas14b.10 tracr 8	SEQ ID NO: 143	UGAACAUUCAGCGCCAAUUGAAGUG UAAGUUGAAGUGUGAGCUUAUGGA UUUUUUUUACAAAUAUACUGA CUUACUAAGAUUCUUGAGGGUUAU ACCCAAAAGAUUGGCGUUGUUGC AACGCAAUAAGAUGUAAAUCUGAA AAGGUUUGAAAUCAUAUAAGUAAU UUUUUUUGAGUUUCGGCUUGAGUA AAUGAACAGACCAAUUUUUAAUU UCGUUCAUAUCAUCGCAACUA
Cas14b.10 tracr 9	SEQ ID NO: 144	AGUGUAAGUUGAAGUGUGAGCUUA UGGAUUUUUUUUACAAAUAUUA CUGACUUACUAAGAUUUCUUGAGG GUUAACCCAAAAGAUUGGCGUUG UUGCAACGCAAUAAGAUGUAAAUC UGAAAAGGUUUGAAAUCAUAUAAG UAAUUUUUUUUGAGUUUCGGCUUG AGUAAAUGAACAGACCAAUUUUU AAUUUCGUUCAUAUCAUCGCAACUA AGAAUUCUUU
Cas14b.10 tracr 10	SEQ ID NO: 145	AGUGUAAGUUGAAGUGUGAGCUUA UGGAUUUUUUUUACAAAUAUUA CUGACUUACUAAGAUUUCUUGAGG GUUAACCCAAAAGAUUGGCGUUG UUGCAACGCAAUAAGAUGUAAAUC UGAAAAGGUUUGAAAUCAUAUAAG UAAUUUUUUUUGAGUUUCGGCUUG AGUAAAUGAACAGACCAAUUUUU AAUUUCGUUCAUAUCAUCGCAACUA AGAAUUCUUUAAAAGGAA
Cas14b.10 tracr 11	SEQ ID NO: 146	AGUGUAAGUUGAAGUGUGAGCUUA UGGAUUUUUUUUACAAAUAUUA CUGACUUACUAAGAUUUCUUGAGG GUUAACCCAAAAGAUUGGCGUUG UUGCAACGCAAUAAGAUGUAAAUC UGAAAAGGUUUGAAAUCAUAUAAG UAAUUUUUUUUGAGUUUCGGCUUG AGUAAAUGAACAGACCAAUUUUU AAUUUCGUUCAUAUCAUCGCAACUA AGAAUUCUUUAAAAGGAAUUC AGAAA
Cas14b.10 tracr 12	SEQ ID NO: 147	AGUGUAAGUUGAAGUGUGAGCUUA UGGAUUUUUUUUACAAAUAUUA CUGACUUACUAAGAUUUCUUGAGG GUUAACCCAAAAGAUUGGCGUUG UUGCAACGCAAUAAGAUGUAAAUC

物种	SEQ ID NO	序列
		UGAAAAGGUUGAAAUCAUAUAAG UAAUUUUAUUUGAGUUUCGGCUUG AGUAAAUGAACAGACCAAUUUUU AAUUUCGUUCAUAUCAUCGCAACUA AGAAUUCUUUAAAAAGGAAAUGC AGAAAUGGACAUUCCAAAAUUCA AAACAAAUAUCGGUUUUUUAAGAC CAGAAAG
Cas14b.10 tracr 13	SEQ ID NO: 148	AGUGUAAGUUGAAGUGUGAGCUUA UGGAUUAUUUUUACAAAAUAAUA CUGACUUACUAAGAUUUCUUGAGG GUAUACCCAAAAAGAUUGGCGUUG UUGCAACGCAUAAGAUGUAAAUC UGAAAAGGUUGAAAUCAUAUAAG UAAUUUUAUUUGAGUUUCGGCUUG AGUAAAUGAACAGACCAAUUUUU AAUUUCGUUCAUAUC
Cas14b.10 tracr 14	SEQ ID NO: 149	AGUGUAAGUUGAAGUGUGAGCUUA UGGAUUAUUUUUACAAAAUAAUA CUGACUUACUAAGAUUUCUUGAGG GUAUACCCAAAAAGAUUGGCGUUG UUGCAACGCAUAAGAUGUAAAUC UGAAAAGGUUGAAAUCAUAUAAG UAAUUUUAUUUGAGUUUCGGCUUG AGUAAAUGAACAGACCAAUUUUU AAUUU
Cas14b.10 tracr 15	SEQ ID NO: 150	AGUGUAAGUUGAAGUGUGAGCUUA UGGAUUAUUUUUACAAAAUAAUA CUGACUUACUAAGAUUUCUUGAGG GUAUACCCAAAAAGAUUGGCGUUG UUGCAACGCAUAAGAUGUAAAUC UGAAAAGGUUGAAAUCAUAUAAG UAAUUUUAUUUGAGUUUCGGCUUG AGUAAAUGAACAGACCAA

[0176] 然后在存在含有紧邻TTTA PAM的3'位点的SEQ ID NO:96靶序列的超螺旋质粒DNA的情况下,Cas14-指导RNA复合体在Tris,pH 7.9缓冲液(50mM乙酸钾、20mM tris-乙酸盐、10mM乙酸镁、100μg/ml BSA)中于37℃下孵育。超螺旋质粒DNA的序列是(SEQ ID NO:152;靶序列以粗体和下划线显示)。在此期间,可编程切口酶使质粒产生切口并切割。10分钟后,用1mg/ml蛋白酶K、0.08% SDS和15mM EDTA猝灭反应。

[0177] 然后通过琼脂糖凝胶上的凝胶电泳分析质粒的产生切口和切割。图6显示了在每个测定中产生切口的质粒百分比和被切割的质粒百分比。图6A显示了Cas14a.3的结果。图6B显示了Cas14b.4的结果。图6C显示了Cas14b.10的结果。结果表明,产生切口和切割活性因可编程切口酶类型而异,并且可编程切口酶进行的产生切口和切割速率可以通过优化tracr序列进行控制。

[0178] 虽然本文已经示出和描述了本发明的优选实施方案,但是对于本领域技术人员来说显而易见的是,这些实施方案仅作为实例提供。在不背离本发明的情况下,本领域技术人

员现将想到许多变化、改变和置换。应当理解,在实践本发明时可以采用对本文所述的本发明的实施方案的各种替代方案。以下权利要求旨在限定本发明的范围以及由此涵盖在这些权利要求及其同等内容范围内的方法和结构。

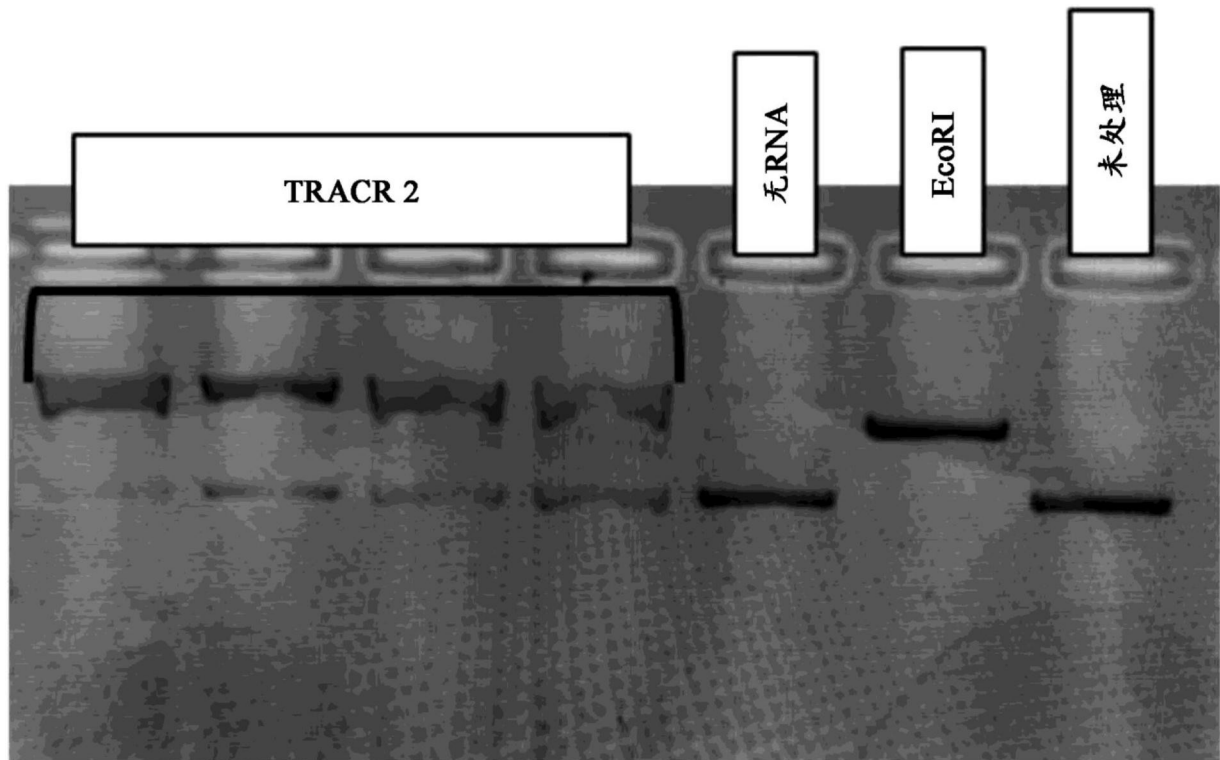


图1

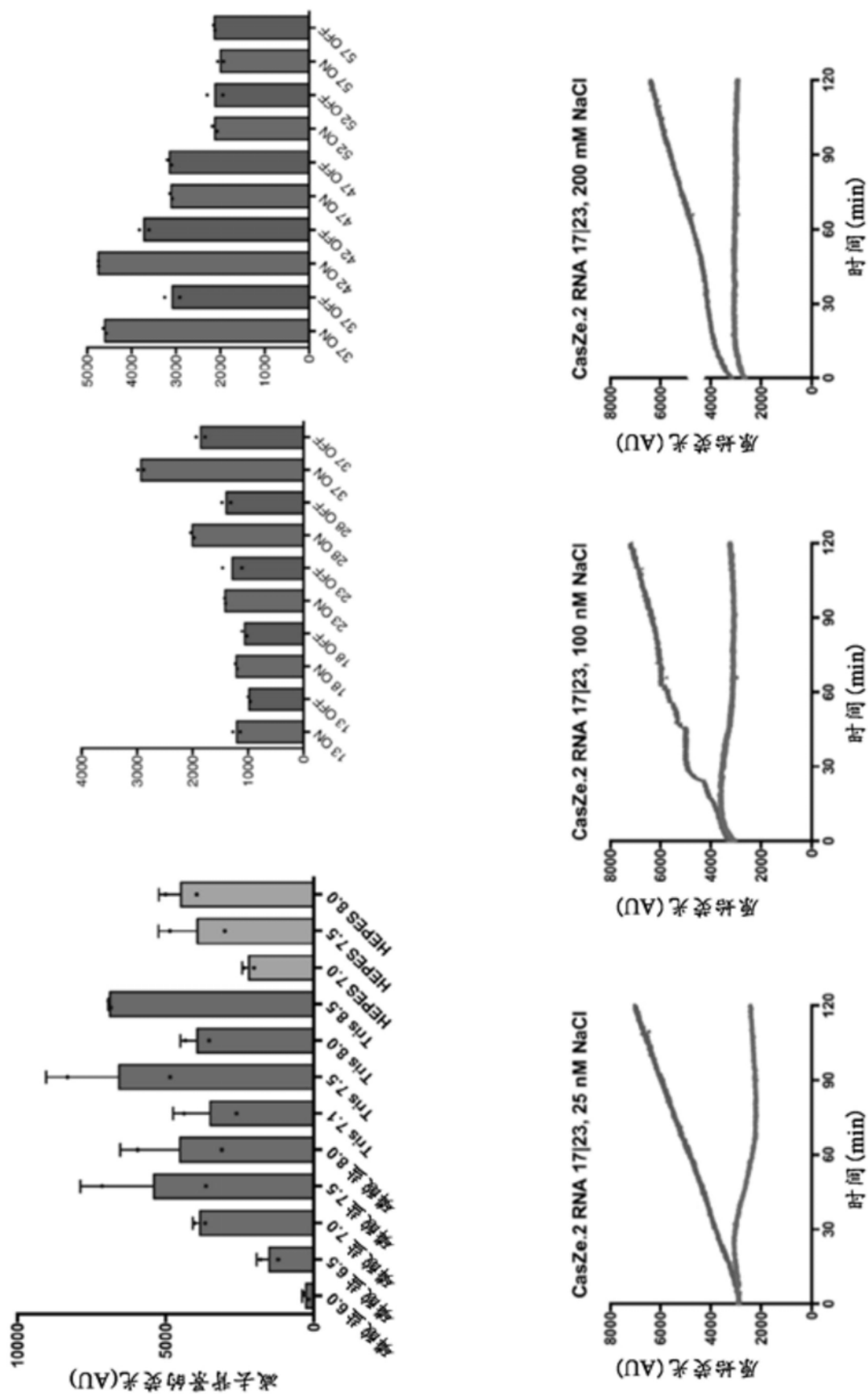


图2

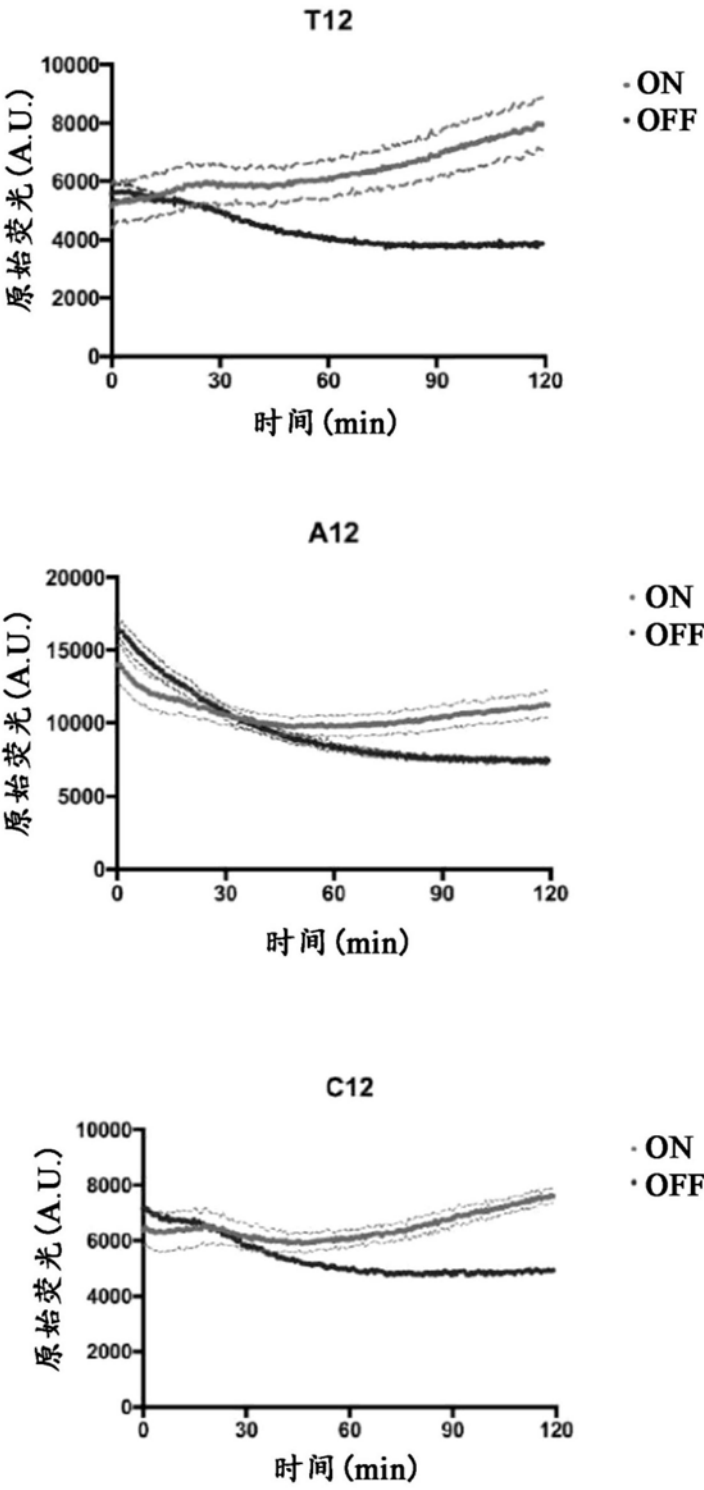


图3

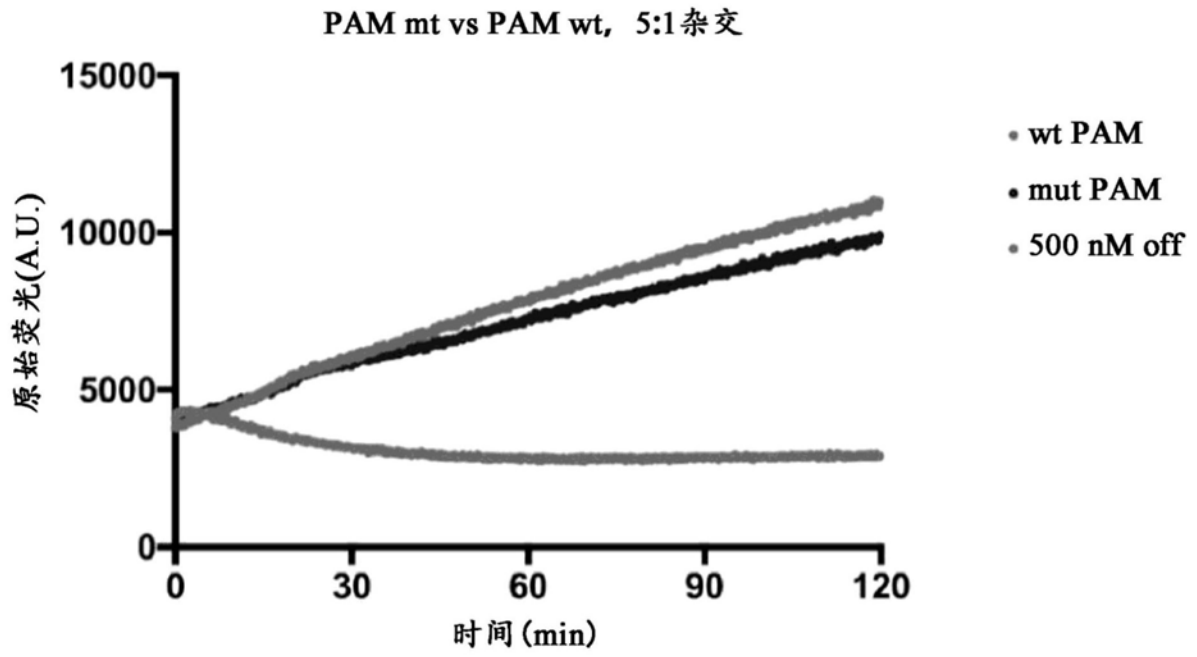


图4

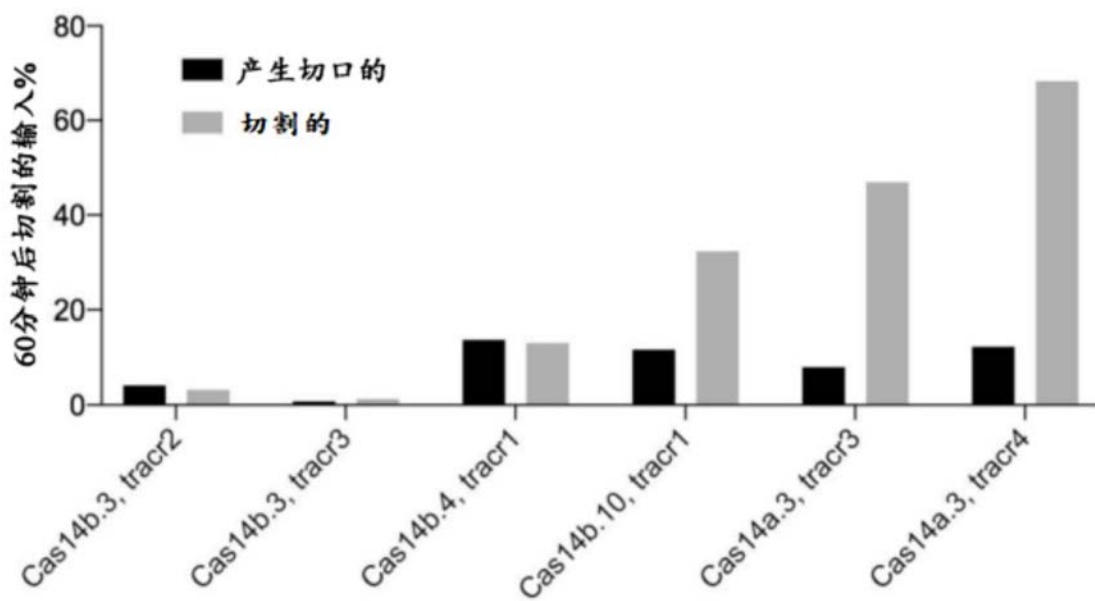


图5

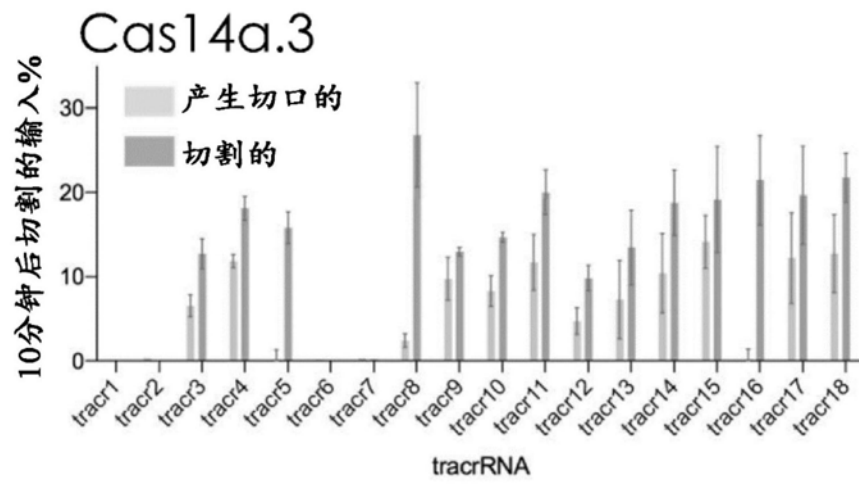


图6A

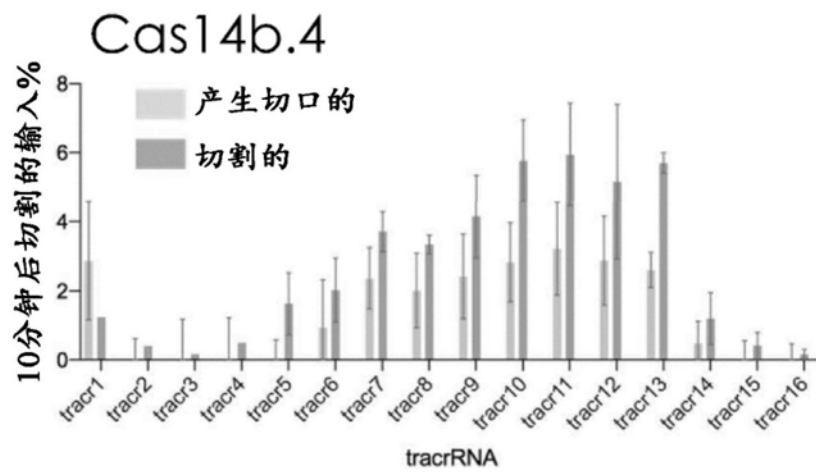


图6B

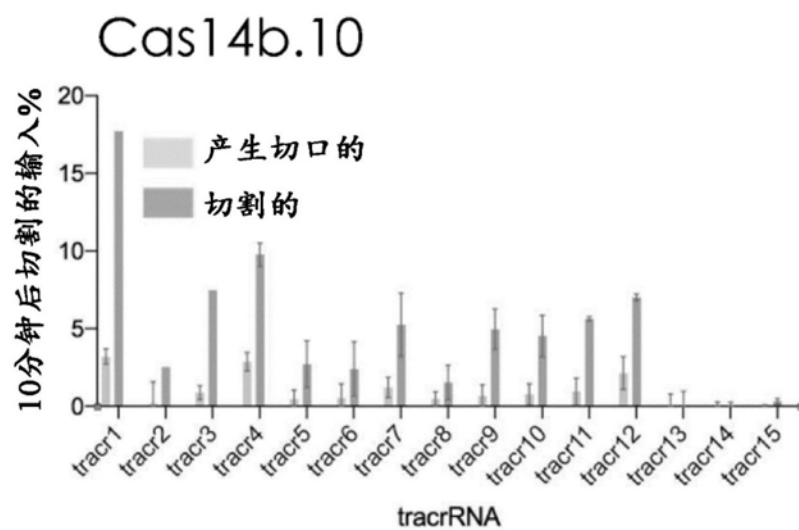


图6C