



(21) 申请号 201880029221.8
(22) 申请日 2018.05.02
(65) 同一申请的已公布的文献号
申请公布号 CN 110914439 A
(43) 申请公布日 2020.03.24
(30) 优先权数据
1706945.1 2017.05.02 GB
(85) PCT国际申请进入国家阶段日
2019.11.01
(86) PCT国际申请的申请数据
PCT/GB2018/051166 2018.05.02
(87) PCT国际申请的公布数据
W02018/203049 EN 2018.11.08
(73) 专利权人 英国研究与创新署
地址 英国斯温登
(72) 发明人 马尔科·特里波迪
埃内斯托·恰巴蒂
(74) 专利代理机构 北京集佳知识产权代理有限公司 11227
专利代理师 顾晋伟 尹玉峰
(51) Int.Cl.
C12N 15/86 (2006.01)
A61K 35/76 (2015.01)

(56) 对比文件
CN 1635905 A, 2005.07.06
WO 2016125364 A1, 2016.08.11
WO 2017004022 A2, 2017.01.05
余元勋 等.《中国分子神经病学》.安徽科学技术出版社, 2015, 第194-197页.
张惠展 等.《基因工程 第4版》.华东理工大学出版社, 2017, 第111-112页.
张慧展.《基因工程 第4版》.华东理工大学出版社, 2017, 第111-112页.
GRAY DANIEL C ET AL.Activation of Specific Apoptotic Caspases with an Engineered Small-Molecule-Activated Protease.《CELL》.2010, 第142卷 (第142期), 第637-646页.
D. ATASOY ET AL.A FLEX Switch Targets Channelrhodopsin-2 to Multiple Cell Types for Imaging and Long-Range Circuit Mapping.《THE JOURNAL OF NEUROSCIENCE》.2008, 第28卷 (第28期), 第7025-7030页.
FINKE S ET AL.Replication strategies of rabies virus.《VIRUS RESEARCH》.2005, 第111卷 (第111期), 第120-131页.
审查员 陈永强
权利要求书3页 说明书47页
序列表37页 附图26页

(54) 发明名称
自灭活病毒载体
(57) 摘要

本发明涉及基于来自单股负链病毒目的病毒,特别是狂犬病病毒的载体。更具体地,本发明涉及狂犬病病毒载体,其已经转染了靶细胞,可在可复制型和不可复制型形式之间转换。在其他一些应用中,本发明避免了与目前基于狂犬病病毒的载体相关的细胞毒性。

1. 狂犬病病毒载体基因组, 其包含编码复制调节蛋白的基因, 其中所述复制调节蛋白包含所述病毒基因组的复制所需的病毒蛋白部分, 所述复制调节蛋白能够采用展示降解决定子的靶向构型, 和不展示所述降解决定子的非靶向构型, 并且其中所述病毒蛋白部分为N蛋白, 并且其中所述降解决定子是PEST序列。

2. 根据权利要求1所述的载体基因组, 其中由所述载体基因组编码的复制调节蛋白是抑制性调节剂, 其包含病毒蛋白部分和调节子部分, 其中所述调节子部分包含所述降解决定子或由其组成。

3. 根据权利要求2所述的载体基因组, 其中所述调节子部分在与关联激活剂接触时可转换至非靶向构型。

4. 根据权利要求3所述的载体基因组, 其中所述激活剂从所述病毒蛋白部分切割所述调节子部分。

5. 根据权利要求4所述的载体基因组, 其中所述激活剂是蛋白酶。

6. 根据权利要求5所述的载体基因组, 其中所述调节子部分包含所述蛋白酶的切割位点, 所述切割位点位于所述病毒蛋白部分与所述降解决定子之间。

7. 根据权利要求5所述的载体基因组, 其中所述蛋白酶不作用于由所述载体基因组编码的任何其他蛋白质。

8. 根据权利要求5所述的载体基因组, 其中所述蛋白酶是病毒蛋白酶、因子Xa、肠激酶或凝血酶。

9. 根据权利要求3所述的载体基因组, 其中所述激活剂由所述载体基因组编码, 并且其中所述激活剂的表达或功能是诱导型的。

10. 根据权利要求3所述的载体基因组, 其中所述激活剂是所述调节子部分的配体。

11. 根据权利要求1所述的载体基因组, 其中由所述载体基因组编码的复制调节蛋白是可抑制的调节蛋白, 其可在与关联抑制剂接触时转换成展示降解决定子的靶向构型。

12. 根据权利要求11所述的载体基因组, 其中所述复制调节蛋白包含病毒蛋白部分和调节子部分, 并且其中所述抑制剂从所述调节蛋白切割所述调节子部分以产生或显露所述降解决定子。

13. 根据权利要求12所述的载体基因组, 其中所述抑制剂是蛋白酶。

14. 根据权利要求13所述的载体基因组, 其中所述调节子部分包含所述蛋白酶的切割位点, 所述切割位点位于所述病毒蛋白部分与所述降解决定子之间。

15. 根据权利要求13或权利要求14所述的载体基因组, 其中所述蛋白酶不作用于由所述载体基因组编码的任何其他蛋白质。

16. 根据权利要求13或14所述的载体基因组, 其中所述蛋白酶是病毒蛋白酶、因子Xa、肠激酶或凝血酶。

17. 根据权利要求11至14中任一项所述的载体基因组, 其中所述抑制剂由所述载体基因组编码, 并且其中所述抑制剂的表达或功能是诱导型的。

18. 根据权利要求1所述的载体基因组, 其中所述基因组还包含编码P蛋白、M蛋白和/或L蛋白的基因。

19. 根据权利要求1所述的载体基因组, 其中所述基因组还包含编码包膜蛋白的基因。

20. 根据权利要求19所述的载体基因组, 其中所述包膜蛋白对于单股负链病毒载体是

天然的。

21. 根据权利要求19所述的载体基因组,其中所述包膜蛋白是假型化包膜蛋白。

22. 根据权利要求1所述的载体基因组,其中所述载体基因组不编码包膜蛋白。

23. 根据权利要求1所述的载体基因组,其中所述载体基因组还包含一个或多个异源基因。

24. 根据权利要求23所述的载体基因组,其中所述异源基因编码标记蛋白、期望针对其产生免疫应答的蛋白质、重组酶、核酸酶、指导RNA (gRNA) 分子、修复模板RNA或基因表达的核酸调节剂。

25. 根据权利要求24所述的载体基因组,其中所述核酸酶是RNA指导的内切核酸酶。

26. 根据权利要求25所述的载体基因组,其还包含编码指导RNA (gRNA) 分子和/或修复模板RNA的异源基因。

27. 核糖核蛋白复合物,其包含与一种或更多种病毒蛋白缔合的根据权利要求1至26中任一项所述的载体基因组。

28. 根据权利要求27所述的核糖核蛋白复合物,其包含与N、P和L蛋白缔合的所述载体基因组。

29. 根据权利要求27所述的核糖核蛋白复合物,其是功能性病毒核衣壳,其能够在引入靶细胞的细胞质时启动转录。

30. 狂犬病病毒载体病毒粒子,其包含根据权利要求1至26中任一项所述的病毒载体基因组。

31. 根据权利要求30所述的载体病毒粒子,其包含狂犬病病毒包膜蛋白。

32. 根据权利要求30所述的载体病毒粒子,其包含假型化包膜蛋白。

33. 根据权利要求1至14和18-26中任一项所述的载体基因组、根据权利要求27至29中任一项所述的核糖核蛋白复合物或根据权利要求30至32中任一项所述的载体病毒粒子,其用于医学治疗方法。

34. 根据权利要求1至14和18-26中任一项所述的载体基因组、根据权利要求27至29中任一项所述的核糖核蛋白复合物或根据权利要求30至32中任一项所述的载体病毒粒子,其用作免疫刺激剂。

35. 正义核酸分子,其编码根据权利要求1至26中任一项所述的病毒载体基因组。

36. 包装细胞,其包含编码根据权利要求1至26中任一项所述的载体基因组的核酸构建体并且能够产生根据权利要求30至32中任一项所述的病毒粒子。

37. 将基因递送至靶细胞的体外方法,其包括使所述靶细胞与根据权利要求27至29中任一项所述的核糖核蛋白复合物或根据权利要求30至32中任一项所述的病毒粒子接触。

38. 将基因递送至靶细胞的非治疗方法,其包括使所述靶细胞与根据权利要求27至29中任一项所述的核糖核蛋白复合物或根据权利要求30至32中任一项所述的病毒粒子接触。

39. 根据权利要求37所述的方法,其中所述载体编码抑制性调节蛋白,并且所述方法包括使所述细胞与关联激活剂接触。

40. 根据权利要求39所述的方法,其中所述激活剂是蛋白质,并且所述方法包括将包含编码所述激活剂之基因的核酸引入所述靶细胞中,使得所述激活剂在所述靶细胞中表达。

41. 根据权利要求37所述的方法,其中所述载体编码可抑制的调节蛋白,并且所述方法

包括使所述细胞与关联抑制剂接触。

42. 根据权利要求41所述的方法, 其中所述抑制剂是蛋白质, 并且所述方法包括将包含编码所述抑制剂之基因的核酸引入所述靶细胞中, 使得所述抑制剂在所述靶细胞中表达。

43. 根据权利要求39至42中任一项所述的方法, 其中所述关联激活剂或所述关联抑制剂的表达和/或功能是诱导型的, 并且所述方法包括诱导所述关联激活剂或所述关联抑制剂在所述靶细胞中的表达和/或功能的步骤。

44. 根据权利要求37所述的方法, 其中:

(i) 所述载体编码抑制性调节蛋白和激活剂, 其中所述激活剂的表达或功能是诱导型的, 并且其中所述方法包括诱导所述激活剂在所述靶细胞中的表达和/或功能的步骤; 或者

(ii) 所述载体编码可抑制的调节蛋白和抑制剂, 其中所述抑制剂的表达或功能是诱导型的, 并且其中所述方法包括诱导所述抑制剂在所述靶细胞中的表达和/或功能的步骤。

45. 根据权利要求44所述的方法, 其包括使所述靶细胞与关联诱导剂接触。

46. 根据权利要求37所述的方法, 其中所述载体基因组不包含编码包膜蛋白的基因, 并且所述方法包括将包含编码包膜蛋白的基因的核酸构建体引入所述靶细胞中的步骤。

47. 根据权利要求37至42和44至46中任一项所述的方法, 其中所述靶细胞是神经细胞。

48. 试剂盒, 其包含根据权利要求1至26中任一项所述的载体基因组、根据权利要求27至29中任一项所述的核糖核蛋白复合物或根据权利要求30至32中任一项所述的病毒粒子, 以及 (a) 关联激活剂或抑制剂或者编码关联激活剂或抑制剂的核酸; 和/或 (b) 编码包膜蛋白的核酸。

49. 组合物, 其包含根据权利要求27至29中任一项所述的核糖核蛋白复合物或根据权利要求30至32中任一项所述的病毒粒子, 其任选地与赋形剂或载体混合。

50. 根据权利要求49所述的组合物, 其中所述组合物是药物组合物, 并且所述载体是可药用载体。

自灭活病毒载体

[0001] 本申请要求于2017年5月2日提交的GB 1706945.1的优先权,其内容和要素出于所有目的通过引用并入本文。

发明领域

[0002] 本发明涉及基于来自单股负链病毒目(order Mononegavirales)的病毒(例如弹状病毒(rhabdovirus))的载体,并且特别是狂犬病病毒(rabies virus)载体。更具体地,其涉及已经转染了靶细胞的载体,其可在可复制型和不可复制型形式之间转换。在其他一些应用中,本发明避免了与目前基于单股负链病毒的载体(例如狂犬病病毒载体)相关的细胞毒性。

[0003] 发明背景

[0004] 基于单股负链病毒目的病毒的载体在多种治疗、诊断和研究方面中具有相当大的潜力。尽管可对这些载体进行改造,使得其不从转染细胞中繁殖,或者能够仅以严格控制的方式繁殖,但其仍然对转染细胞显示出明显的细胞毒性,从而限制了其应用。

[0005] 例如,神经营养性病毒(neurotropic virus),特别是缺失G的狂犬病(Δ G-狂犬病)(12,13)通过沿着回路(circuit)途径在神经元之间传播,提供了获得对拓扑结构限定的(topologically defined)神经元的遗传访问的潜在工具(13-16)。

[0006] 然而,尽管基于 Δ G-狂犬病的方法在神经元回路的解剖学研究中具有变革性作用,但病毒诱导的细胞毒性阻止了在狂犬病感染之后超过5至15天的时间将其用于跟踪网络动力学或操纵体内神经网络的功能特性(17-19)。消除病毒细胞毒性的一种可能的解决方案将是在初次感染之后沉默病毒转录。对于基于DNA的病毒,例如腺相关病毒(Adeno-Associated Virus, AAV),传统上是通过以CRE重组酶依赖性方式转化病毒基因组盒来实现的;翻转基因组盒可有效开启或关闭病毒(20)。由于缺乏可靠的RNA重组酶,目前对于基于RNA的病毒(例如狂犬病病毒)尚不采用这种方法。

[0007] 因此,仍然需要解决细胞毒性问题的改良的单股负链病毒载体。

发明内容

[0008] 由于单股负链病毒目病毒中转录和复制的偶联性质(coupled nature),发明人假设病毒蛋白稳定性的条件化调节可充当被感染的细胞内病毒转录-复制周期的开关,其当满足特定条件时打开或关闭病毒。

[0009] 因此,发明人设计了这样的病毒载体,其能够转染(并因此将遗传有效载荷传递至)靶细胞,并且其可在其中可表达病毒蛋白的活性状态与其中病毒蛋白表达被抑制的失活状态之间转换。持续维持失活状态通常将导致载体从宿主细胞中被消除。

[0010] 特别地,载体编码“复制调节蛋白”,所述复制调节蛋白包含载体基因组复制所需的单股负链病毒蛋白,并且所述复制调节蛋白能够采用两种构型。

[0011] 形成复制调节蛋白的一部分的单股负链病毒蛋白在本文中被称作“病毒蛋白部分”。

[0012] 在一种被称为“靶向”构型的构型中,复制调节蛋白被靶向以在靶细胞中降解。因此,“靶向”构型是不稳定的。当复制调节蛋白维持这种构型时,来自载体基因组的表达和/或载体基因组的复制被抑制或完全抑制。

[0013] 被称为“非靶向”构型的调节蛋白的第二构型比“靶向”构型更稳定。因此,通常因为病毒蛋白部分在宿主细胞胞质溶胶中积累的水平高于靶向构型,所以其支持更高水平的来自载体的表达和/或载体基因组的复制。

[0014] 在本文中提及“稳定性”仅涉及相关蛋白在靶细胞中的半衰期及其被靶细胞的蛋白质降解途径降解的倾向。靶向构型的半衰期比非靶向构型的半衰期更短,并且因此其被认为比非靶向构型的稳定性更差。

[0015] 因此,复制调节蛋白提供了开关,所述开关使得能够严格调节载体基因组的表达和/或载体基因组的复制。

[0016] 通常,调节蛋白的“靶向”构型展示降解靶信号或“降解决定子(degron)”,即标记该蛋白在细胞内降解的部分。在本说明书中术语“降解决定子”和“降解靶信号”可互换使用。

[0017] 调节蛋白的“非靶向”构型与靶向构型的不同之处在于,其不包含或不展示导致靶向构型稳定性较低的降解决定子。因此,非靶向构型通常比靶向构型在细胞内的半衰期更长。

[0018] 下面将更详细地讨论可在本发明的上下文中使用的降解决定子。

[0019] 复制调节蛋白的默认状态可以是靶向构型。也就是说,靶向构型可由载体基因组编码,使得复制调节蛋白以靶向构型合成。通过与可被称为“激活剂”或“稳定剂”的试剂接触来刺激采用非靶向构型,因为其支持产生足够的功能性蛋白来激活载体表达和/或复制。因此,其可允许载体在细胞内持续存在或存活。

[0020] 具有靶向构型作为其默认状态的复制调节蛋白可被称为“抑制性”复制调节蛋白,因为除了在关联(cognate)激活剂存在的情况之外,来自编码这样的调节蛋白的载体的表达(或这样的载体的复制)将倾向于被抑制。因此,“抑制性”调节蛋白在构成上是不稳定的。

[0021] 当以编码具有默认靶向构型的复制调节蛋白的载体首先感染靶细胞时,蛋白质合成的初始水平可足以支持一段时期的复制并可产生感染性子代病毒粒子(取决于载体的有效载荷和通过靶细胞反式提供的任何其他蛋白质)。此后,由于复制调节蛋白的不稳定性,在不存在激活剂的情况下复制被抑制或被完全压制。

[0022] 或者,复制调节蛋白的默认状态可以是非靶向构型。也就是说,复制调节剂的非靶向构型可由载体基因组编码,使得复制调节蛋白以非靶向构型合成。通过与可被称为“抑制剂”或“去稳定剂”的试剂接触来刺激采用靶向构型,因为其存在将倾向于抑制或压制来自载体的表达和/或载体基因组的复制。

[0023] 具有非靶向构型作为其默认状态的复制调节蛋白可被称为“抑制性”复制调节蛋白,因为除了在关联抑制剂存在的情况之外,从包含这样的调节蛋白的载体的表达将倾向于继续进行。

[0024] 在许多实施方案中,复制调节蛋白包含病毒蛋白部分和调节子部分。病毒蛋白部分和调节子部分可形成调节蛋白的独特折叠的结构域,尽管这不一定总是必需的。病毒蛋白部分通常能够发挥相关病毒蛋白天然形式的所有生物学功能。

[0025] 调节子部分(当存在时)的性质和特征确定了调节蛋白是抑制性调节蛋白(inhibitory modulator protein)(即,在不存在关联激活剂的情况下抑制了载体表达或复制)还是可抑制的调节蛋白(inhibitable modulator protein)(即,在不存在关联抑制剂的情况下载体表达或复制正常进行)。

[0026] 对于抑制性调节蛋白,调节子部分包含降解决定子或由其组成,并且在调节蛋白的默认状态下,降解决定子被展示使得该蛋白被靶向以降解。因此,关联激活剂将例如通过共价或非共价修饰来发挥去除、掩蔽或以其他方式使降解子失效的作用。例如,关联激活剂可从病毒蛋白部分切割调节子部分。因此,激活剂可以是蛋白酶,其能够作用于调节蛋白以从病毒蛋白部分切割调节子部分。这种机制可被描述为复制调节蛋白(或病毒蛋白部分)的“切割诱导的稳定化”。或者,激活剂可与调节蛋白结合,例如与调节子部分结合,以掩蔽或以其他方式使降解决定子失活。因此,激活剂可以是调节蛋白的配体,例如调节蛋白的调节子部分。这种机制可被描述为复制调节蛋白(或病毒蛋白部分)的“配体诱导的稳定化”。

[0027] 对于可抑制的调节蛋白,调节子部分不包含或不展示默认状态下的降解决定子。相反,关联抑制剂将通常通过与调节子部分共价或非共价相互作用而与调节蛋白相互作用以产生或显露降解决定子。因此,关联抑制剂可从病毒蛋白部分切割调节子部分,以在病毒蛋白部分中显露或产生降解决定子。这种机制可被描述为复制调节蛋白的“切割诱导的去稳定化”(或“切割诱导的降解”)。或者,抑制剂可与复制调节蛋白结合,例如与调节子部分结合,以产生或显露降解决定子。因此,抑制剂可以是调节蛋白的配体,例如调节蛋白的调节子部分。这种机制可被描述为复制调节蛋白的“配体诱导的去稳定化”(或“配体诱导的降解”)。因此,调节子部分和抑制剂可分别代表诱导型降解决定子系统的第一和第二组成部分,其中第一和第二组成部分的缔合产生降解决定子。

[0028] 因此,本发明提供了包含编码复制调节蛋白的基因的单股负链病毒载体基因组,其中复制调节蛋白包含病毒基因组复制所需的单股负链病毒蛋白部分,复制调节蛋白能够采用展示降解决定子的靶向构型和不展示降解决定子的非靶向构型。

[0029] 载体基因组通常是负义单链RNA分子。

[0030] 如上所述,由载体基因组编码的复制调节蛋白可以是抑制性调节剂。其可包含病毒蛋白部分和调节子部分,其中调节子部分包含降解决定子或由其组成。

[0031] 复制调节蛋白可在与关联激活剂接触时转换成非靶向构型。在一些实施方案中,激活剂能够从调节蛋白切割调节子部分。在另一些实施方案中,激活剂是调节蛋白(例如调节子部分)的配体。配体可例如在空间上或通过构象变化掩蔽或以其他方式使降解决定子失活。

[0032] 或者,由载体基因组编码的复制调节蛋白可以是可抑制的调节剂,其可在与关联抑制剂接触时转换至展示降解决定子的靶向构型。其可包含病毒蛋白部分和调节子部分。

[0033] 在一些实施方案中,抑制剂能够从调节蛋白切割调节子部分以产生或显露降解决定子。在另一些实施方案中,抑制剂是调节子部分的配体,并且在与调节子部分结合时产生或显露降解决定子。

[0034] 因此,适当地,激活剂或抑制剂可通过从病毒蛋白部分切割调节子部分来发挥作用,以分别去除或显露降解决定子。在这样的实施方案中,激活剂或抑制剂通常是蛋白酶。在这样的实施方案中,调节子部分包含蛋白酶的切割位点。

[0035] 蛋白酶可与靶细胞正交(orthogonal)。也就是说,蛋白酶识别在靶细胞的蛋白质组中未发现的切割位点,即在由靶细胞编码并在其中表达的天然蛋白质中未发现该切割位点。因此,特定的蛋白酶可根据靶细胞而变化。合适的蛋白酶的实例包括病毒蛋白酶、因子Xa、肠激酶和凝血酶,所述病毒蛋白酶特别是来自不是单股负链病毒(例如烟草蚀斑病毒蛋白酶(Tobacco Etch Virus protease,TEVp)和人鼻病毒(human rhinovirus,HRV)3C蛋白酶)的蛋白酶。当靶细胞是神经细胞时,这些方法中的任何一种都可以是合适的。

[0036] 优选地,所选择的蛋白酶不作用于由载体基因组编码的任何其他蛋白质,例如,在由载体基因组编码的蛋白质中,只有复制调节剂包含被相关蛋白酶识别的切割位点。

[0037] 激活剂或抑制剂本身可由载体基因组编码。在这样的情况下,将通常可通过使包含载体基因组的细胞与适当的诱导剂接触来诱导试剂的表达或功能。例如,载体可包含编码以功能上无活性的形式表达的试剂的一种或更多种基因,其中在与诱导剂接触时诱导功能。

[0038] 因此,试剂可表达为两个或更多个分离的蛋白部分,其需要功能性诱导剂的存在以将其缔合为功能形式。每个组成的试剂部分可表达为与伴侣蛋白的融合体,其中伴侣蛋白在与功能性诱导剂接触时缔合。

[0039] 下面提供了关于诱导型的试剂的更多细节。

[0040] 任何单股负链病毒都可用作如上所述的载体的基础。具有非分段基因组的那些基因组是特别合适的。对于许多应用而言,可期望的是该病毒能够在哺乳动物细胞中感染和复制。例如,单股负链病毒可以是弹状病毒(例如狂犬病毒(例如狂犬病病毒))或水疱性病毒(例如水疱性口炎病毒(vesicular stomatitis virus,VSV))。

[0041] 复制调节剂的病毒蛋白部分可包含大多数单股负链病毒共有的任何蛋白质或由其组成。这些蛋白质包括大蛋白(large protein,L)、核蛋白(nucleoprotein,N或NP)、磷蛋白(phosphoprotein,P)、基质蛋白(matrix protein,M)或糖蛋白(glycoprotein,G;也简称为“包膜”蛋白)。该术语通常用于弹状病毒,并且在本说明书中也用于通过功能或序列同源性来指其在其他单股负链病毒中的等同物。然而,有必要鉴定提供对病毒复制的充分控制的适当的病毒蛋白部分。

[0042] 复制调节蛋白的病毒蛋白部分是从病毒基因组表达(例如从病毒基因组转录mRNA)和/或病毒基因组复制所需的蛋白。因此,病毒蛋白部分通常不是病毒包膜蛋白(例如糖蛋白或G蛋白),因为对于大多数单股负链病毒而言,其他病毒蛋白和基因组的复制在没有包膜蛋白的情况下进行。

[0043] 因此,大蛋白(L)、核蛋白(N或NP)、磷蛋白(P)或基质蛋白(M)可以是优选的,例如大蛋白(L),核蛋白(N或NP)或磷蛋白(P)。

[0044] 核蛋白(N或NP蛋白)可以是特别优选的,尤其是在弹状病毒载体(例如狂犬病病毒载体)中。实际上,本发明人已经发现,仅N蛋白对基于狂犬病病毒的载体中的病毒复制提供了足够的控制。

[0045] 无论哪种单股负链病毒蛋白代表复制调节蛋白的病毒蛋白部分,应理解的是,复制调节蛋白代表单股负链病毒蛋白在载体基因组中的唯一拷贝。换句话说,载体基因组不包含编码仅以非靶向构型存在的相同蛋白质形式的基因。

[0046] 因此,载体基因组可包含编码N蛋白、P蛋白、M蛋白和L蛋白中的一种、两种、三种或

全部四种的基因,其中这些蛋白中的一种被提供作为复制调节蛋白的病毒蛋白部分。

[0047] 载体基因组可包含编码包含N蛋白作为病毒蛋白部分的复制调节蛋白的基因,并且可任选地还包含P蛋白、M蛋白和L蛋白中的一种、两种或三种。

[0048] 载体基因组可包含编码以下的基因:

[0049] (i) 包含N蛋白作为病毒蛋白部分的复制调节蛋白;

[0050] (ii) P蛋白;

[0051] (iii) M蛋白;和

[0052] (iv) L蛋白。

[0053] 载体基因组还可包含编码包膜蛋白的基因。包膜蛋白可对于单股负链病毒载体是天然的。例如,当载体是狂犬病病毒时,其可以是狂犬病病毒G蛋白。或者,可并入不同的包膜蛋白以调节由载体产生的病毒粒子的感染性。替代包膜蛋白的并入可被称为“假型化”。在这样的情况下,包膜蛋白可包含对于载体是天然的包膜蛋白的胞内结构域和任选的跨膜结构域,以及异源胞外结构域,即胞外结构域与胞内结构域不是来自同一单股负链病毒包膜蛋白。胞外结构域可来自不同的病毒包膜蛋白(例如,来自不同单股负链病毒种、科或属的包膜蛋白,或来自不同病毒目的包膜蛋白),或者可以是能够与在靶细胞表面表达的受体结合的任何蛋白结构域,只要该包膜蛋白仍然能够介导病毒粒子感染靶细胞即可。

[0054] 例如,狂犬病病毒载体可包含至少具有来自VSV包膜蛋白、A型禽肉瘤白血病毒(avian sarcoma leukosis virus, ASLV)包膜蛋白或B型ASLV包膜蛋白的胞外结构域的包膜蛋白。其可包含来自VSV、A型ASLV或B型ASLV的完整包膜蛋白。或者,其可包含来自狂犬病病毒G蛋白的胞内结构域和任选的跨膜结构域。

[0055] 在一些实施方案中,载体基因组不编码包膜蛋白。这样的病毒粒子对于神经元细胞中所谓的“单突触示踪”特别有用。

[0056] 除了任何假型化包膜蛋白之外,载体基因组可包含一个或更多个异源基因,即编码不是对于载体基因组是天然的表达产物(通常是蛋白质或RNA分子)的基因,即不存在由与载体相同病毒类型的野生型病毒编码的表达产物。异源基因可被视为将通过载体递送至靶细胞的“有效载荷”。因此,异源基因可编码期望在靶细胞中表达的任何表达产物(无论是RNA还是蛋白质)。因此,异源基因的特性和功能将取决于载体的预期作用。在许多实施方案中,异源基因不编码单股负链病毒表达产物或病毒表达产物。

[0057] 异源基因可位于载体基因组内的任何合适的位点。在一些实施方案中,其位于编码M和L蛋白的基因(例如替换编码内源性G蛋白的基因)之间。

[0058] 异源基因可例如编码以下中的一个或更多个:

[0059] -标记蛋白,例如荧光蛋白或赋予抗生素抗性的蛋白;

[0060] -期望针对其产生免疫应答的蛋白;

[0061] -重组酶,例如CRE重组酶,例如其中靶细胞的基因组包含重组酶的识别位点;

[0062] -核酸酶,例如RNA指导的内切核酸酶,例如Cas9;

[0063] -指导RNA(gRNA)分子,例如与RNA指导的内切核酸酶组合使用;

[0064] -修复模板RNA,例如与RNA指导的内切核酸酶组合使用;

[0065] -基因表达的核酸调节剂,例如siRNA、RNAi、反义RNA或核酶。

[0066] 异源基因将处于病毒转录调节序列(例如转录起始子和终止子信号)的控制下。

[0067] 当包括两个或更多个异源基因时,其通常将各自具有其自身相关的调节序列。异源基因表达可受到以下因素的影响,例如异源基因使用的转录起始位点,以及其与位于紧邻病毒基因组上游的基因末端的间距,即在紧邻上游的基因末端与异源基因的起始之间的基因间区域(intergenic region, IGR)的长度。N蛋白在狂犬病病毒中具有最高的转录水平,因此可期望的是异源基因使用来自N蛋白的转录起始位点。转录起始位点可例如具有序列AACACCCCT(例如在毒株B19和N2C中所见)或AACACCTCT。Finke等(2000)已经证明,基因上游的IGR长度会影响其表达,并且较短的IGR序列与表达提高相关。因此,可期望的是异源基因上游的IGR序列的长度为5个核苷酸或更短,例如2个核苷酸。当IGR的长度为2个核苷酸时,其序列可以为CT。

[0068] 可期望的是包含两个或更多个异源基因,尤其是如果其具有互补功能。例如,可期望的是包含编码RNA指导的内切核酸酶的异源基因、用于该内切核酸酶的指导RNA(以指导内切核酸酶至靶细胞染色体的期望位点)、以及任选地还包含修复模板RNA以指导RNA在靶位点上要实现的特定修改。异源基因在病毒基因组中可彼此相邻。异源基因之间的IGR序列的长度可例如是5个核苷酸或更短,例如2个核苷酸。可期望的是每个异源基因使用来自N蛋白的N蛋白转录起始位点。

[0069] 内切核酸酶可以是Cas9酶,例如来自金黄色葡萄球菌(*Streptococcus aureus*) (saCas9)、酿脓链球菌(*Streptococcus pyogenes*) (SpCas9)、脑膜炎奈瑟氏球菌(*Neisseria meningitidis*) (NM Cas9)、嗜热链球菌(*Streptococcus thermophilus*) (ST Cas9)、齿垢密螺旋体(*Treponema denticola*) (TD Cas9)或其变体(例如SpCas9的D1135E、VRER、EQR或VQR变体)的Cas9酶。由于saCas9及其变体相对较小的大小,其可以是优选的。

[0070] 这里使用的术语Cas9包括Cas9内切核酸酶的功能性变体,例如Cas9切口酶和核酸酶失效的Cas9(dCas9)。

[0071] 可并入载体的异源基因的大小将取决于多种因素,例如载体所基于的特定病毒,病毒粒子的包装能力以及已经从载体中缺失的天然病毒基因的包装能力(如果有的话)。

[0072] 例如,在其中已经缺失G基因的狂犬病病毒载体中,在对包装效率没有实质性影响的情况下,可包含至少3.7kb的异源编码序列,并且甚至高至4.0kb、高至4.5kb,或高至5.0kb。参见参考文献32。

[0073] 异源基因也可与靶细胞异源,即靶细胞不天然编码或表达该基因的产物。或者,靶细胞可编码或表达基因产物但是在不同的调节控制下,使得通过引入异源基因来调节靶细胞中基因产物的表达模式。

[0074] 本发明还提供了核糖核蛋白复合物,其包含与一种或更多种病毒蛋白(例如N、P和L蛋白中的一种或更多种)缔合的载体基因组。载体基因组可与所有N、P和L蛋白缔合。核糖核蛋白复合物可以是功能性病毒核衣壳,其能够在引入靶细胞的细胞质时启动转录。其可另外包含M蛋白。

[0075] 本发明还提供了包含如本文中所述的单股负链病毒载体基因组的单股负链病毒载体病毒粒子。病毒粒子通常包含L蛋白、N蛋白、P蛋白和M蛋白,并且还包含包膜蛋白。包膜蛋白可以是天然的单股负链病毒包膜蛋白。或者,病毒粒子可以已经用如上所述的不同的包膜蛋白进行假型化以调节其感染性。

[0076] 病毒粒子通常采用核糖核蛋白复合物(包含基因组、L、N、M和P蛋白)的形式,其被

包含该包膜蛋白的膜包膜包围。

[0077] 病毒粒子能够将基因组递送至靶细胞并启动基因组的转录。

[0078] 为避免疑义,病毒粒子通常不包含复制调节蛋白;即,病毒粒子中的蛋白质不能在靶向和非靶向构型之间转换。

[0079] 本发明还提供了编码如本文中所述的病毒载体基因组的正义核酸分子。

[0080] 正义核酸可以是RNA或DNA分子。当其是DNA分子时,其可作为表达构建体的一部分提供,其包含能够指导病毒基因组分子合成的转录调节序列。表达构建体可作为附加体(例如质粒)的一部分提供或可整合到宿主细胞的染色体中。

[0081] 本发明还提供了包装细胞(packaging cell),其包含编码载体基因组的核酸构建体并能够产生如本文中所述的病毒粒子。

[0082] 因此,包装细胞表达L、N、M和P蛋白以及包膜蛋白。蛋白质能够形成还包含载体基因组的感染性病毒粒子。

[0083] 根据载体基因组本身的结构和编码内容,这些蛋白质可由载体基因组编码或由细胞内其他核酸构建体反式(in trans)提供。

[0084] 病毒粒子颗粒本身中的蛋白质通常不与降解决定子连接。通常,包装细胞将编码(并能够表达)所有病毒粒子蛋白的功能性形式,而没有降解决定子。但是,如果病毒粒子蛋白中的一种仅在如本文所述的抑制性复制调节剂的背景下表达,则包装细胞还必须包含(例如表达)关联激活剂,以使得病毒复制可以进行。

[0085] 编码要反式提供的任何蛋白质的构建体通常是DNA构建体,并且可独立地进行染色体整合或作为附加体(例如质粒)存在。

[0086] 包装细胞的性质将取决于特定的病毒载体,但是可以是哺乳动物细胞,例如灵长类动物或啮齿动物细胞。其可以是成纤维细胞或任何其他合适的细胞类型,例如HEK(人胚胎肾)293细胞或BHK(小仓鼠肾)细胞。

[0087] 狂犬病病毒载体包装细胞的构建以及病毒载体设计的其他方面在Osakada和Callaway(2013)Design and generation of rabies virus vectors,Nature Protocols 8(8):1583-1601中进行了描述。还参见Wickersham等(2010)Production of glycoprotein-deleted rabies viruses for monosynaptic tracing and high level gene expression in neurons,Nature Protocols 5(3):595-606。

[0088] 本发明还提供了将基因递送至靶细胞的方法,其包括使靶细胞与本发明的核糖核蛋白复合物或病毒粒子接触。

[0089] 当使用核糖核蛋白复合物时,可期望的是将其直接递送至细胞胞质溶胶中,例如通过显微注射,或与载体(例如聚合物或脂质)结合,以促进其转运至细胞中。

[0090] 在许多实施方案中,靶细胞不天然表达负责调节蛋白的靶向和非靶向构型之间转换的试剂。当载体编码抑制性复制蛋白时,激活剂是初始初次感染之后继续复制和繁殖载体所需的。因此,当期望进一步复制和繁殖时,将通常使载体与激活剂的递送系统或激活剂本身一起使用。

[0091] 当载体编码可抑制的调节蛋白时,可不必提供抑制剂。但是,可期望的是载体表达和复制不应该立即进行,在这样的情况下,可将载体与抑制剂的递送系统或抑制剂本身一起使用。

[0092] 因此,该方法可包括使靶细胞与关联激活剂或抑制剂适当接触的步骤。可将靶细胞与关联激活剂周期性地例如以1天至1个月的重复间隔、例如以1天至14天的间隔、例如以5至10天的间隔接触不确定的时间段。这样的定期施用对于在靶细胞内维持存活的载体而不允许足够的转录和病毒复制来损害细胞可以是有用的。

[0093] 当试剂是蛋白质时,该方法可进一步包括将包含编码关联激活剂或抑制剂的基因的核酸引入靶细胞,使得试剂在靶细胞中表达。

[0094] 可在载体之前、与载体基本同时或在载体之后将试剂引入靶细胞。当将试剂在载体之后引入时,通常将在载体的一个月内,例如在2周内、1周内、1天内、12小时内或1小时内引入。可将其周期性地,例如以1天至1个月的重复间隔、例如以1天至14天的间隔、例如以5至10天的间隔引入不确定的时间段。

[0095] 试剂的表达可以是诱导型的。例如,编码试剂的基因可包含诱导型启动子,使得试剂的转录需要将细胞与转录诱导剂接触。诱导剂可以是抗生素,例如多西环素或四环素。对这样的抗生素有响应的诱导型启动子是技术人员公知的。

[0096] 或者,所述试剂可以以功能失活形式表达,其需要使细胞与功能性诱导剂接触以恢复功能。例如,试剂可表达为两个或更多个分离的蛋白部分,其需要功能性诱导剂的存在以将其缔合为功能形式。每个组成的试剂部分可表达为与伴侣蛋白的融合体,其中伴侣蛋白在与功能性诱导剂接触时缔合。例如,TEV蛋白酶(TEVp)可表达为两个单独的亚基,每个亚基都与异源蛋白部分(FRB和FKBP)融合,其仅在雷帕霉素(所谓的“SPLIT-TEV”系统)存在下缔合(参见See Gray等(2010) Activation of Specific Apoptotic Caspases with an Engineered Small-Molecule-Activated Protease, Cell 142(4):637-646.)。也可使用其他对融合伴侣,其对缔合具有不同的要求。

[0097] 在一些实施方案中,试剂可由载体基因组本身或在靶细胞的基因组中编码。因此,载体基因组或靶细胞可包含一种或更多种编码激活剂或抑制剂的基因,其中试剂的表达或功能是诱导型的。在这样的情况下,所述方法不需要使靶细胞与激活剂或抑制剂接触,而仅需要使靶细胞与(表达或功能性)诱导剂接触。

[0098] 因此,所述方法可包括通常通过使靶细胞与诱导剂接触来诱导试剂在靶细胞中的表达和/或功能的步骤。

[0099] 可使靶细胞与诱导物周期性地,例如以1天至1个月的重复间隔、例如以1天至14天的间隔、例如以5至10天的间隔接触不确定的时间段。

[0100] 在一些实施方案中,载体基因组不包含编码包膜蛋白的基因。因此,该方法可(另外地或替代地)包括将包含编码包膜蛋白的基因的核酸构建体引入靶细胞中的步骤,使得包膜蛋白在靶细胞中表达并入质膜中。可在载体之前、与载体基本同时或在载体之后将包膜蛋白引入靶细胞。当将包膜蛋白在载体之后引入时,通常将在载体的一个月内,例如在14天内、在7天内、在1天内、在12小时内或在1小时内引入。

[0101] 包膜蛋白的表达可以是诱导型的。例如,编码包膜蛋白的基因可包含诱导型启动子,使得包膜蛋白的转录需要将细胞与转录诱导剂接触。

[0102] 可通过任何适当的方式将包含编码抑制剂或激活剂的基因和/或包膜蛋白的核酸递送至靶细胞。例如,每个核酸可以以包封形式(例如包封在脂质体中)或通过病毒载体独立地以DNA或RNA、裸核酸的形式与例如聚合物或脂质的载体(例如与载体的复合物)结合引

入。合适的病毒载体包含逆转录病毒载体(包括慢病毒载体)、腺病毒载体和腺相关病毒(AAV)载体。AAV载体可以是特别合适的。

[0103] 编码试剂和包膜蛋白的基因可分开或一起(例如作为同一核酸构建体的一部分)递送。

[0104] 发现具有抑制性调节蛋白的病毒载体可特别用作针对相同病毒的天然形式的免疫刺激剂,例如作为疫苗。将载体递送至接受者将导致病毒仅在也接受激活剂的细胞中复制。子代病毒粒子可从这样的感染的细胞中形成并释放。然而,其将不能在未接受激活剂的细胞中进一步复制。因此,将载体和激活剂共同施用至特定的选定细胞,可能会导致病毒复制和释放的控制周期短,从而刺激免疫系统并引发其识别病毒。

[0105] 可将病毒载体以其他方式进行减毒以减少对接受者的任何风险。因此,其可包含一种或更多种进一步的突变以减少或消除感染性或毒力。

[0106] 子代病毒粒子的释放可用于模拟针对天然病毒的抗体应答。

[0107] 为了产生和释放子代病毒粒子,病毒载体必须编码包膜蛋白,或者如果载体缺乏包膜蛋白,则必须将包膜蛋白反式单独提供给感染的细胞。使用缺乏包膜蛋白的载体可代表有用的安全措施。

[0108] 如果载体缺乏包膜蛋白,并且没有反式提供包膜蛋白,则在存在激活剂的情况下仍然可以有足够的蛋白表达,从而导致通过接受者细胞的MHC分子(尤其是MHC I)展示病毒抗原,并且随后刺激针对病毒的T细胞应答(尤其是CTL应答)。

[0109] 然而,在许多情况下仍然可期望使用编码包膜蛋白的载体。

[0110] 发现本发明的载体在向神经细胞的基因递送中特别有用,特别是当基于狂犬病病毒或水疱性口炎病毒时。众所周知,使用缺失G的狂犬病病毒可在中枢神经系统和外周神经系统二者中进行单突触回路示踪。然而,目前的单突触回路示踪方法使用有限,因为由于细胞内病毒蛋白的积累,标记的细胞通常仅维持存活1至2周。本载体提供了在不影响生存力的情况下遗传标记神经细胞的能力。

[0111] 本发明还提供了试剂盒,其包含本发明的载体基因组以及(a)关联激活剂或抑制剂或者编码关联激活剂或抑制剂的核酸;和/或(b)编码包膜蛋白的核酸。

[0112] 试剂盒可包含用于递送核酸的系统,例如在其他地方所述的载体。相同的系统可用于递送关联存活因子和如上所述的包膜蛋白二者。

[0113] 本发明还提供了包含本发明的载体病毒粒子的组合物,其任选地与赋形剂或载体混合。组合物可以是药物组合物,并且载体可以是可药用载体。

[0114] 单股负链病毒感染多种类型的细胞,包括植物、昆虫、鱼类和哺乳动物细胞。原则上,认为本发明可应用于任何类型的单股负链病毒,因为其符合相似的基因组组织模式并且依赖于相似的转录和复制偶联系统。弹状病毒可特别适合。

[0115] 发现本发明的载体可特别用于感染哺乳动物细胞的弹状病毒属和种,特别是狂犬病毒(lyssavirus),例如狂犬病病毒(rabies virus)(狂犬病狂犬病毒;Rabies lyssavirus,RABV)和水疱性病毒(例如水疱性口炎病毒(vesicular stomatitis virus,VSV))。狂犬病病毒对神经元具有特殊的向性,这使其成为向神经组织进行基因递送的有价值的候选基因。

[0116] 当载体基于狂犬病病毒时,可使用任何合适的毒株,包括CVS(攻击病毒标准品)及

其变体,例如CVS-N2c、PV4(巴斯德病毒(pasteur virus)),PM(皮特曼-摩尔(Pitman-Moore)),Flury-LEP(鸡胚细胞低传代株(10w egg passage)),Flury-HEP(鸡胚细胞高传代株(high egg passage)),ERA和SAD(Street-Alabama Dufferin)。在参考文献25中描述了CVS-N2c毒株,并且完整的基因组序列可在2010年12月29日在GenBank登录no.HM535790, HM535790.1版本下获得。

[0117] 因此,靶细胞类型可来自灵长类(例如旧世界猴(Old World monkey)、新世界猴(New World monkey)、猿或人)、啮齿动物(例如小鼠或大鼠)、犬类动物(canine)(例如家犬)、猫类动物(feline)(例如家猫)、马类动物(equine)(例如马)、牛类动物(bovine)(例如牛)、山羊类动物(caprine)(例如山羊)、绵羊类动物(ovine)(例如绵羊)或兔类动物(lagomorph)(例如兔)。

[0118] 靶细胞类型可以是任何期望的细胞类型。在一些实例中,其可以是神经细胞,例如神经元或胶质细胞。神经细胞可以是周围或中枢神经系统的一部分。

附图说明

[0119] 图1

[0120] 在病毒蛋白的系统蛋白酶体靶向之后筛选病毒扩增效率。(A)通过蛋白酶体靶向PEST结构域的可逆病毒蛋白去稳定化。(B)HEKGG和HEKTGG中TEV_p依赖性病毒扩增。比例尺:100 μ m。(C-H)所有重组狂犬病构建体(品红色)和对照狂犬病(蓝绿色)的扩增效率的定量(平均值 $\pm$ SD;虚线显示阈值水平)。(I-K)在HEKTGG(蓝绿色,+TEV)和HEGG(品红色,-TEV)中的扩增效率的定量。x轴,转染后天数(p.t.),y轴扩增效率。

[0121] 图2.

[0122] 体内无细胞毒性。(A)SiR生命周期。(B)SiR表达盒和实验程序。(C-E”)用SiR^{CRE}-mCherry感染并在p.i 1至3至8周成像的Rosa-LoxP-STOP-LoxP-YFP小鼠海马切片的共聚焦图像。比例尺:25 μ m。(F)将在1至2至3周和8周的YFP和mCherry阳性神经元的数量归一化至1周时间点(平均值 $\pm$ SEM)。(G)将病毒RNA(品红色)和内源YFP表达(蓝绿色)水平归一化至1周RNA水平(平均值 $\pm$ SEM)。

[0123] 图3.

[0124] SiR跨突触和逆行(retrograde)传播。(A)将AAV-TVAmCherry-GLY注射到Rosa-LoxSTOPLox-YFP小鼠的CA1中,并且在2周之后用EnVA假型SiR^{CRE}特异性靶向表达TVA的神经元。(A’) 在注射部位检测到YFP^{ON}/mCherry^{ON}起始神经元(箭头)和(A”-A”) SiR病毒的跨突触跳跃永久标记了CA3和外侧内嗅皮层(lateral enthorinal cortex,LEC)中具有YFP表达的神经元。比例尺:25 μ m。(B-B”)在p.i 3周时用AAV-TVAmCherry-GLY和SiR^{CRE}感染的CA1锥体神经元的共聚焦图像。比例尺:10 μ m。(C-C”)通过SiR^{mCherry}海马CA1注射从CA3和LEC投射到CA1的神经元的逆行示踪。(D)在上丘(Superior Colliculus,SC)的浅层中通过SiR注射感染的V1皮质神经元。比例尺:25 μ m。

[0125] 图4.

[0126] SiR感染对神经元生理没有长期影响。(A-B)在感染后(p.i.)1周和p.i 8周对CA1锥体神经元正向和负向电流阶跃的膜电位响应。(C)输入电阻(D)静息膜电位和动作电位(AP)阈值,(E)AP波幅和宽度,(F)p.i 1周(n=10,洋红色)和p.i.8周(n=8,蓝绿色)时提高

的神经元正向电流阶跃的触发频率(平均值 $\pm$ SEM)。(G)对表达ChR2的SiR^{CRE-Mcherry}感染的CA1神经元(p.i.1周)强度提高(每次扫描提高1%,直到出现峰值;0%的浅灰色变为9%的黑色)的0.2ms蓝光脉冲的膜电位响应。插入,为每次扫描递送LED功率。(H、I)对来自同一神经元记录的800ms光脉冲(2.17mW)和20Hz下的40个1.5ms长的光脉冲的膜电位响应。(J)p.i.1周(品红色)和8周(蓝绿色),在频率提高的40次光刺激之后的动作电位成功率。(K)在由DNQX(20 μ M)阻断的非ChR2表达神经元中记录到光诱发的EPSP。两种条件的平均迹线均以黑色显示。 $*P<0.05$ 。

[0127] 图5.

[0128] SiR示踪的V1神经元的方向调谐响应未改变。(A)分别在V2和V1中在Rosa-LoxP-STOP-LoxP-tdTomato小鼠中注射SiR^{CRE}和AAV-GCaMP6s的示意图。(B-B'')SiR^{CRE}注射之后,V1神经元的双光子最大投射。在表达GCaMP6s的灰色神经元中(B),在表达tdTomato的品红色神经元中(B'),在表达二者的合并神经元中(B'',合并)。比例尺:50 μ m。(C)建立了视觉刺激示意图。(D)在B小组中显示的同一视野中的有效ROI的轮廓。(E)GCaMP6s(蓝绿色)和GCaMP6s-tdTomato(品红色)神经元的代表性Ca²⁺迹线。比例尺:200s,20% dF/F₀。(F)注射SiR的4周之后活神经元的平均百分比(n=122个GCaMP6神经元(蓝绿色),n=59个GCaMP6s和GCaMP6s-tdTomato神经元(洋红色))。(G)荧光随时间的变化,反映了在每个神经元的优选方向上对漂移光栅的视觉响应。(H)V1被感染的神经元的调谐曲线的示例。比例尺:5s,10% dF/F₀。

[0129] 图6.

[0130] 在体外和体内测试 Δ G-N^{pest}狂犬病^{SPLIT-TEV-mCherry}的细胞毒性。(A)将hESCs来源的神经元用 Δ G-N^{pest}狂犬病^{SPLIT-TEV-mCherry}感染,并在16天内纵向成像。(B-B'') Δ G-N^{pest}狂犬病^{SPLIT-TEV-mCherry}和B19 Δ G-狂犬病对照(C-C'')感染的hESCs来源的神经元在感染后(p.i.)4至10天和16天成像。(D)在存在或不存在雷帕霉素的情况下,在施用CTR Δ G-狂犬病或 Δ G-N^{pest}狂犬病^{SPLIT-TEV-mCherry}后4至10和16天之后,将感染细胞的百分比归一化至第4天时间点(平均值 $\pm$ SEM)。(E)归一化至第4天时间点的 Δ G-N^{pest}狂犬病^{SPLIT-TEV-mCherry}和 Δ G-狂犬病被感染的神经元的mCherry信号强度(平均值 $\pm$ SEM,比例如D中所示)。比例尺:50 μ m。(F-G'')在p.i.1、2或3周时用 Δ G狂犬病(蓝绿色)或 Δ G-N^{pest}狂犬病^{SPLIT-TEV-mCherry}(洋红色,G-G'')体内感染的海马切片。比例尺:50 μ m(H)归一化至第7天时间点的海马中 Δ G狂犬病(黑色)或 Δ G-N^{pest}狂犬病^{SPLIT-TEV-mCherry}(灰色)在p.i.1、2或3周时的被感染的神经元的百分比(平均值 $\pm$ SEM)。

[0131] 图7.

[0132] 雷帕霉素诱导HEK细胞中的Split TEV重建和PEST结构域的切割。(A)标记病毒蛋白的药剂稳定策略。雷帕霉素诱导Split TEV蛋白的二聚化,其切割拯救病毒蛋白的降解决定子结构域。(B)HEK中的Split TEV雷帕霉素剂量反应。随着雷帕霉素浓度的提高(0至10至50nM),TEV报告物的TEV依赖性切割增加。(c)将Split TEV盒克隆到狂犬病基因组中的糖蛋白基因座中。在用表达狂犬病的Split TEV感染并用TEV报告物转染的HEK293T中观察到了明显的雷帕霉素依赖性TEV活性。

[0133] 图8.

[0134] 体内短期SiR^{mCherry}动力学。(A)SiR^{mCherry}盒设计。(B)在Rosa-LoxSTOPLox-YFP小鼠

的CA1中注射SiR mCherry-CRE。(C-E”)在p.i.3至6和9天时,用SiR^{mCherry-CRE}感染的CA1锥体神经元的共聚焦图像。比例尺:25 μ m。(F-F”)在p.i.3至6和9天时,YFP^{ON}、mCherry^{ON}和YFP^{ON}mCherry^{ON}神经元的百分比。

[0135] 图9

[0136] SiR的药理学再激活。(A)多西环素诱导的AAV的设计。rTTA反式激活剂由病毒组成型表达,在多西环素存在下,其驱动TEV蛋白酶表达。(B)在Rosa-LoxP-STOP-LoxP-YFP小鼠的海马中注射AAV-TRE-TEV,然后一周之后在相同区域中施用SiR^{CRE-mCherry}并且在另一周之后施用多西环素。(C-F)用SiR和AAV-TRE-TEV感染的海马锥体神经元在p.i.2或3周时用多西环素再激活。比例尺:50 μ m。(G)在所有YFP^{ON}被感染的神经元中对mCherry^{ON}神经元进行定量。

[0137] 图10.

[0138] Δ G-狂犬病诱导了皮质和海马的死亡。在p.i.1、2和3周时用 Δ G-狂犬病^{GFP}感染的皮质神经元(A-A”)和CA1锥体神经元(B-B”)的共聚焦图像。比例尺:50 μ m。(C)归一化至一周时间点的皮质(黑色)和海马(灰色)中 Δ G狂犬病在p.i.1、2或3周时的被感染的神经元的百分比(平均值 $\pm$ SEM)(海马,在2周时的细胞死亡率为92 $\pm$ 3%,每个时间点n=3,单因素方差分析,F=101,P=2.4 $\times 10^{-5}$;在3周时皮质85 $\pm$ 2%的细胞死亡,每个时间点n=3,单因素方差分析,F=17,P=3.2 $\times 10^{-3}$)

[0139] 图11.

[0140] 编码SiR CRE-mCherryPEST载体基因组的质粒的图;(A)T7启动子序列中显示的质粒序列以双下划线显示。开放阅读框用单下划线显示,并依次代表N-TEV-PEST、P、M、iCRE-2A-mCherryPEST和L蛋白。

[0141] 图12.

[0142] 编码SiR载体的具有多个克隆位点的代替包膜(G)蛋白基因的质粒的图。

[0143] 图13.

[0144] 质粒的完整序列示于图12中。主链质粒(即非SiR)序列以具有阴影的大写字母表示。在载体开放阅读框下划线,按顺序(从5'-3')为N-TEV-PEST蛋白、P蛋白、M蛋白、L蛋白。多个克隆位点(SbfI-NheI和PacI-AscI)以斜体大写字母显示。在N-TEV-PEST蛋白中,N序列以小写字母常规字体显示、TEV切割位点以大写斜体显示、PEST序列以大写双下划线显示、并且接头肽序列以小写字母斜体显示。

[0145] 图14.

[0146] 通过从图12和13中所示的质粒转录获得的负义载体RNA基因组序列。

[0147] 发明详述

[0148] 单股负链病毒

[0149] 单股负链病毒目包含以下科:博尔纳病毒、丝状病毒、Myomonaviridae、Nyamiviridae、副黏病毒、肺病毒、弹状病毒和Sunviridae科,以及五个未分类的属:Anphevirus、Arlivirus、Chengtivirus、Crustavirus和Wastritried。下表列出了这些科中的属和种。下表中的星号“*”表示类型种类。

[0150]

单股负链病毒目：科、属、种、和病毒			
科	属	种	病毒（缩写）
博尔纳病毒	博尔纳病毒	蛇毒类 1 型博尔纳病毒	Loveridge garter 蛇病毒 1 型 (LGSV-1)
		哺乳动物 1 型博尔纳病毒*	博尔纳病病毒 1 型 (BoDV-1)
			博尔纳病病毒 2 型 (BoDV-2)
		雀形目 1 型博尔纳病毒	金丝雀博尔纳病 毒 1 型 (CnBV-1)
			金丝雀博尔纳病 毒 2 型 (CnBV-2)
			金丝雀博尔纳病 毒 3 型 (CnBV-3)
		雀形目 2 型博尔纳病毒	雌性雀科博尔纳 病毒 1 型 (EsBV-1)
		鸚形目 1 型博尔纳病毒	鸚鵡博尔纳病毒 1 型 (PaBV-1)
			鸚鵡博尔纳病毒 2 型 (PaBV-2)
			鸚鵡博尔纳病毒 3 型 (PaBV-3)
			鸚鵡博尔纳病毒 4 型 (PaBV-4)
			鸚鵡博尔纳病毒 7 型 (PaBV-7)

[0151]

单股负链病毒目：科、属、种、和病毒			
科	属	种	病毒（缩写）
		鸚形目 2 型博尔纳病毒	鸚鵡博尔纳病毒 5 型（PaBV-5）
		水鸟 1 型博尔纳病毒	水鸟博尔纳病毒 1 型（ABBV-1）
			水鸟博尔纳病毒 2 型（ABBV-2）
丝状病毒	<i>Cuevavirus</i>	<i>Lloviu cuevavirus*</i>	Lloviu 病毒 （LLOV）
	埃博拉病毒	本迪布焦埃博拉病毒	本迪布焦病毒 （BDBV）
		雷斯顿埃博拉病毒	雷斯顿病毒 （RESTV）
		苏丹埃博拉病毒	苏丹病毒（SUDV）
		塔伊森林埃博拉病毒	塔伊森林病毒 （TAFV）
		扎伊尔埃博拉病毒*	埃博拉病毒 （EBOV）
	马尔堡病毒	马尔堡马尔堡病毒*	马尔堡病毒 （MARV）
			Ravn 病毒 （RAVV）
<i>Mymonaviridae</i>	<i>Sclerotimonavirus</i>	核盘菌 <i>sclerotimonavirus*</i>	核盘菌阴性-链状 RNA 病毒 1 型 （SsNSRV-1）
<i>Nyamiviridae</i>	尼亚病毒	<i>Midway nyavirus</i>	Midway 病毒 （MIDWV）
		<i>Nyamanini nyavirus*</i>	Nyamanini 病毒

[0152]

单股负链病毒目：科、属、种、和病毒			
科	属	种	病毒（缩写）
			（NYMV）
		内华达山脉 <i>nyavirus</i>	内华达山脉病毒 （SNVV）
	<i>Socyvirus</i>	大豆胞囊线虫 <i>socyvirus</i> *	大豆胞囊线虫病 毒 1 型（SbCNV-1）
副黏病毒	<i>Aquaparamyxovirus</i>	大西洋鲑鱼副黏病毒*	大西洋鲑鱼副黏 病毒（AsaPV）
	禽副黏病毒	禽副黏病毒 2 型	禽副黏病毒 2 型 （APMV-2）
		禽副黏病毒 3 型	禽副黏病毒 3 型 （APMV-3）
		禽副黏病毒 4 型	禽副黏病毒 4 型 （APMV-4）
		禽副黏病毒 5 型	禽副黏病毒 5 型 （APMV-5）
		禽副黏病毒 6 型	禽副黏病毒 6 型 （APMV-6）
		禽副黏病毒 7 型	禽副黏病毒 7 型 （APMV-7）
		禽副黏病毒 8 型	禽副黏病毒 8 型 （APMV-8）
		禽副黏病毒 9 型	禽副黏病毒 9 型 （APMV-9）
		禽副黏病毒 10 型	禽副黏病毒 10 型 （APMV-10）
		禽副黏病毒 11 型	禽副黏病毒 11 型 （APMV-11）
		禽副黏病毒 12 型	禽副黏病毒 12 型 （APMV-12）
		新城疫病毒*	禽副黏病毒 1 型 （APMV-1）
	<i>Ferlavirus</i>	矛头蛇副黏病毒*	矛头蛇病毒 （FDLV）
	亨尼病毒	<i>Cedar</i> 亨尼病毒	Cedar 病毒（CedV）
		<i>Ghanaian bat</i> 亨尼病毒	Kumasi 病毒（KV）
		亨德拉病毒*	亨德拉病毒（HeV）
		<i>Mojiang</i> 亨尼病毒 <i>s</i>	Mojiang 病毒

[0153]

单股负链病毒目：科、属、种、和病毒			
科	属	种	病毒（缩写）
			(MojV)
		尼帕病毒	尼帕病毒 (NiV)
	麻疹病毒	犬瘟热病毒	犬瘟热病毒 (CDV)
		鲸麻疹病毒	鲸麻疹病毒 (CeMV)
		Feline 麻疹病毒	Feline 麻疹病毒 (FeMV)
		麻疹病毒*	麻疹病毒 (MeV)
		反刍宠物瘟疫性炎症病毒	反刍宠物瘟疫性炎症病毒 (PPRV)
		海豹瘟热病毒	海豹瘟热病毒 (PDV)
		牛瘟麻疹病毒	牛瘟麻疹病毒 (RPV)
	呼吸道病毒	牛副流感病毒 3 型	牛副流感病毒 3 型 (BPIV-3)
		人副流感病毒 1 型	人副流感病毒 1 型 (HPIV-1)
		人副流感病毒 3 型	人副流感病毒 3 型
		猪副流感病毒 1 型	猪副流感病毒 1 型 (PPIV-1)
		仙台病毒*	仙台病毒 (SeV)
	腮腺炎病毒	人副流感病毒 2 型	人副流感病毒 2 型 (HPIV-2)
		人副流感病毒 4 型	人副流感病毒 4a 型 (HPIV-4a)
			人副流感病毒 4b

[0154]

单股负链病毒目：科、属、种、和病毒			
科	属	种	病毒（缩写）
			型（HPIV-4b）
		马普埃拉病毒	马普埃拉病毒 （MapV）
		腮腺炎病毒*	腮腺炎病毒（MuV）
			bat 腮腺炎病毒 （BMV）
		副流感病毒 5 型	副流感病毒 5 型 （PTV-5）
		猪麻疹病毒	La Piedad Michoacan Mexico 病毒（LPMV）
肺病毒	偏肺病毒	猿猴副流感病毒 41 型	猿猴副流感病毒 41 型（SV-41）
		禽偏肺病毒*	禽偏肺病毒 （AMPV）
	正肺病毒	人偏肺病毒	人偏肺病毒 （HMPV）
		牛呼吸道合胞病毒	牛呼吸道合胞病毒 （BRSV）
		人呼吸道合胞病毒*	人呼吸道合胞病毒 A2 型（HRSV-A2）
			人呼吸道合胞病 毒 B1 型 （HRSV-B1）
			人呼吸道合胞病毒 S2 型（HRSV-S2）
		小鼠肺病毒	小鼠肺病毒（MPV）

[0155]

单股负链病毒目：科、属、种、和病毒			
科	属	种	病毒（缩写）
弹状病毒	细胞质弹状病毒	苜蓿矮缩细胞质弹状病毒	苜蓿矮缩病毒（ADV）
		大麦黄条点花叶细胞质弹状病毒	大麦黄条点花叶病毒（BYSMV）
		硬花甘蓝坏死性黄化细胞质弹状病毒	硬花甘蓝坏死性黄化病毒（BNYV）
		羊茅叶条纹病细胞质弹状病毒	羊茅叶条纹病毒（FLSV）
		莴苣坏死黄化病细胞质弹状病毒*	莴苣坏死黄化病毒（LNYV）
		莴苣黄斑细胞质弹状病毒	莴苣黄斑病毒（LYMoV）
		北方禾谷花叶细胞质弹状病毒	北方禾谷花叶病毒（NCMV）
		苦苣菜细胞质弹状病毒	苦苣菜病毒（SonV）
		草莓皱缩细胞质弹状病毒	草莓皱缩病毒（SCV）
		美洲小麦条点花叶细胞质弹状病毒	美洲小麦条点花叶病毒（WASMV）
	双角病毒	咖啡环斑双角病毒	咖啡环斑病毒（CoRSV）
		兰花斑点双角病毒*	兰花斑点病毒（OFV）
	暂时热病毒	阿德莱德河暂时热病毒	阿德莱德河病毒（ARV）
		贝里马暂时热病毒	贝里马病毒

[0156]

单股负链病毒目：科、属、种、和病毒			
科	属	种	病毒（缩写）
			(BRMV)
		牛暂时热病毒*	牛暂时热病毒 (BEFV)
		狂犬相关暂时热病毒	狂犬相关病毒 (KOTV)
		奥博第安暂时热病毒	奥博第安病毒 (OBOV)
	狂犬病毒	阿拉万狂犬病毒	阿拉万病毒 (ARAV)
		澳洲蝙蝠狂犬病毒	澳洲蝙蝠狂犬病毒 (ABLV)
		Bokeloh 蝙蝠狂犬病毒	Bokeloh 蝙蝠狂犬病毒 (BBLV)
		杜文黑基狂犬病毒	杜文黑基病毒 (DUVV)
		欧洲蝙蝠狂犬病毒 1 型	欧洲蝙蝠狂犬病毒 1 型 (EBLV-1)
		欧洲蝙蝠狂犬病毒 2 型	欧洲蝙蝠狂犬病毒 2 型 (EBLV-2)
		<i>IkomaBBLV</i>	<i>IkomaBBLV</i> (IKOV)
		伊尔库特狂犬病毒	伊尔库特病毒 (IRKV)
		库詹德狂犬病毒	库詹德病毒 (KHUV)
		拉各斯蝙蝠狂犬病毒	拉各斯蝙蝠病毒 (LBV)

[0157]

单股负链病毒目：科、属、种、和病毒			
科	属	种	病毒（缩写）
		莫科拉狂犬病毒	莫科拉病毒 (MOKV)
		狂犬病毒*	狂犬病毒 (RABV)
		Shimoni 蝙蝠狂犬病毒	Shimoni 蝙蝠病毒 (SHIBV)
		西高加索蝙蝠狂犬病毒	西高加索蝙蝠病毒 (WCBV)
	粒外弹状病毒	牙鲆粒外弹状病毒	牙鲆弹状病毒 (HIRV)
		大麻哈鱼 1 型粒外弹状病毒*	传染性造血坏死病 毒 (IHNV)
		大麻哈鱼 2 型粒外弹状病毒	病毒出血性败血症 病毒 (VHSV)
		黑鱼粒外弹状病毒	黑鱼弹状病毒 (SHRV)
	细胞核弹状病毒	曼陀罗黄脉细胞核弹状病毒	曼陀罗黄脉病毒 (DYVV)
		茄斑驳矮化细胞核弹状病毒	茄斑驳矮化病毒 (EMDV)
		玉米细条纹细胞核弹状病毒	玉米细条纹病毒 (MSFV)
		伊朗玉米花叶细胞核弹状病毒	伊朗玉米花叶病毒 (MIMV)
		玉米花叶细胞核弹状病毒	玉米花叶病毒 (MMV)
		马铃薯黄矮细胞核弹状病毒*	马铃薯黄矮病毒 (PYDV)

[0158]

单股负链病毒目：科、属、种、和病毒			
科	属	种	病毒（缩写）
		水稻黄矮细胞核弹状病毒	水稻黄矮病毒 (RYSV)
			水稻暂黄病毒 (RTYV)
		苦苣菜黄网细胞核弹状病毒	苦苣菜黄网病毒 (SYNV)
		苦菜黄脉细胞核弹状病毒	苦菜黄脉病毒 (SYVV)
		芋头脉萎黄细胞核弹状病毒	芋头脉萎黄病毒 (TaVCV)
	<i>Perhabdovirus</i>	<i>Anguillid perhabdovirus</i>	欧洲鳗鱼病毒 X (EVEX)
		鲈鱼 <i>perhabdovirus</i> *	鲈鱼弹状病毒 (PRV)
		海鲢 <i>perhabdovirus</i>	湖鲢弹状病毒 (LTRV)
	西格玛病毒	果蝇 (<i>Drosophila affinis</i>) 西格玛病毒	果蝇西格玛病毒 (DAffSV)
		<i>Drosophila ananassae</i> 西格玛病毒	<i>Drosophila ananassae</i> 西格玛病毒 (DAnaSV)
		伊米果蝇 (<i>Drosophila immigrans</i>) 西格玛病毒	伊米果蝇西格玛病毒 (DImmSV)
		黑腹果蝇 (<i>Drosophila melanogaster</i>) 西格玛病毒*	黑腹果蝇西格玛病毒 (DMe1SV)
		暗果蝇 (<i>Drosophila obscura</i>)	暗果蝇西格玛病毒

[0159]

单股负链病毒目：科、属、种、和病毒			
科	属	种	病毒（缩写）
		西格玛病毒	(DObsSV)
		<i>Drosophila tristis</i> 西格玛病毒	<i>Drosophila tristis</i> 西格玛病毒 (DTriSV)
		厩腐蝇西格玛病毒	厩腐蝇西格玛病毒 (MStaSV)
	春病毒 (<i>Sprivivirus</i>)	鲤春病毒*	鲤春病毒血症病毒 (SVCV)
		狗鱼幼鱼春病毒	草鱼弹状病毒 (GrCRV)
			狗鱼幼鱼弹状病毒 (PFRV)
			丁鲷弹状病毒 (TenRV)
	<i>Tibrovirus</i>	卡森平原 <i>tibrovirus</i>	卡森平原病毒 (CPV)
		蒂布鲁加尔冈 <i>tibrovirus</i> *	毕文斯阿姆病毒 (BAV)
			蒂布鲁加尔冈病 毒 (TIBV)
	<i>Tupavirus</i>	<i>Durham tupavirus</i> *	Durham 病毒 (DURV)
		树鼩 <i>tupavirus</i>	树鼩病毒 (TUPV)
	巨脉病毒 (<i>Varicosavirus</i>)	莴苣巨脉伴随 巨脉病毒*	莴苣巨脉伴随病毒 (LBVaV)
	水疱性病毒	阿拉戈斯水疱性病毒	阿拉戈斯水疱性口 炎病毒 (VSAV)

[0160]

单股负链病毒目：科、属、种、和病毒			
科	属	种	病毒（缩写）
		卡拉加斯水疱性病毒	卡拉加斯病毒 (CJSV)
		金迪普拉水疱性病毒	金迪普拉病毒 (CHPV)
		科卡尔水疱性病毒	科卡尔病毒 (COCV)
		印第安纳水疱性病毒*	印第安纳水疱性口 炎病毒 (VSIV)
		伊斯法罕水疱性病毒	伊斯法罕病毒 (ISFV)
		马拉巴水疱性病毒	马拉巴病毒 (MARAV)
		新泽西水疱性病毒	新泽西水疱性口炎 病毒 (VSNJV)
		帛黎水疱性病毒	帛黎病毒 (PIRYV)
	未分类	<i>Flanders 病毒</i>	Flanders 病毒 (FLAV)
		<i>Ngaingan 病毒</i>	Ngaingan 病毒 (NGAV)
		<i>Wongabel 病毒</i>	Wongabel 病毒 (WONV)
<i>Sunviridae</i>	<i>Sunshinevirus</i>	爬行类 <i>sunshinevirus 1</i> *	Sunshine Coast 病 毒 (SunCV)
未分类	<i>Anphevirus</i>	<i>Xincheng anphevirus</i> *	Xīnchéng蚊病毒 (XcMV)
未分类	<i>Arlivirus</i>	<i>Lishi arlivirus</i> *	Líshí蛛病毒 2 型 (LsSV-2)

[0161]

单股负链病毒目：科、属、种、和病毒			
科	属	种	病毒（缩写）
未分类	<i>Chengtivirus</i>	<i>Tacheng chengtivirus</i> *	Tǎchéng 蜚病毒 6 型（TcTV-6）
未分类	<i>Crustavirus</i>	<i>Wenzhou crustavirus</i> *	Wēnzhōu 蟹病毒 1 型（WzCV-1）
未分类	<i>Wastrivirus</i>	<i>Sanxia wastrivirus</i> *	Sānxiá 水龟病毒 4 型（SxWSV-4）

[0162] 有关更多详细信息,参见“Taxonomy of the order Mononegavirales:update 2016”,Afonso等,Arch.Virol. (2016) 161:2351-2360。

[0163] 该目内的病毒(在此被称为“单股负链病毒”)是具有非感染性、线性、单链、负义RNA的基因组的包膜病毒。

[0164] 大多数的单股负链病毒具有非分段的基因组。不希望受理论束缚,认为本说明书中描述的方法和组合物更适用于非分段病毒而不是分段病毒。

[0165] 基因组具有反向互补的3' 和5' 端,并且不与任何蛋白质共价连接。

[0166] 基因组通常具有按3' UTR(非翻译区)-核心蛋白编码基因-包膜蛋白编码基因-聚合酶编码基因-5' UTR顺序的编码和非编码元件的保守布局。

[0167] 由基因组编码的核心蛋白包括:核蛋白(通常被称为“N”或“NP”)、经常被磷酸化并因此被称为磷蛋白(被称为“P”)的蛋白、和基质蛋白(“M”)。

[0168] 聚合酶是RNA依赖性的RNA聚合酶并通常是病毒编码的最大蛋白质并因此可被称为大(或“L”)蛋白质。聚合酶在该顺序内相对高度保守。P蛋白可代表L蛋白的辅助因子,即为了聚合酶活性可需要L蛋白和P蛋白一起存在。

[0169] 术语N(或NP)、P、M和L蛋白在本说明书中用于指在单股负链病毒中具有相应作用的任何蛋白,无论其常规名称是什么。

[0170] 包膜蛋白是可被糖基化的跨膜蛋白。因此,在一些单股负链病毒中,其被称为糖蛋白(或“G”蛋白)。

[0171] 病毒粒子包含螺旋核糖核蛋白核衣壳,其中基因组与N、P和L蛋白组合。该核衣壳被基质和膜包膜层包围。

[0172] 宿主细胞的感染导致病毒核衣壳释放到胞质溶胶中,在该胞质溶胶中经由依赖RNA的RNA聚合酶进行转录和复制。

[0173] 所有转录均从基因组3' 端的单个启动子发生。每个基因转录之后,聚合酶终止或继续到下游的下一个基因,产生mRNA产生梯度,最接近基因组3' 端的那些基因以最高拷贝数转录,并且朝向5' 末端的基因产生的转录本越来越少。通常,每种病毒会产生5至10个不同的mRNA。核蛋白(通常被称为“N”或“NP”)的水平决定了mRNA生成与基因组复制之间转换的时间。复制涉及通过聚合酶产生全长正义反基因组,其随后被转录成全长负义基因组拷贝以包装成病毒粒子。

[0174] 本说明书中的术语“基因”用于指编码表达产物的载体基因组中的序列,并且其通常在载体的启动子的控制下指导该表达产物表达。最终表达产物可以是RNA(在这种情况下表达仅需要从载体基因组转录)或蛋白质(在这种情况下表达需要通过病毒蛋白转录为mRNA,然后翻译为蛋白质)。

[0175] 弹状病毒,并且特别是狂犬病病毒的特征,在下面作为单股负链病毒生物学的说明性实例进行更详细描述。

[0176] 弹状病毒的基因组的长度通常为约11至15kb。其不同地感染脊椎动物(包括哺乳动物和鱼类)、昆虫和植物。

[0177] 狂犬病基因组在其末端具有短的非编码区,被称为3' 前导区(leader,le)和5' 非转录尾区(trailer,tr),其分别启动和终止基因组转录和复制。3' 和5' 端是反向互补的。该末端还包含用于转录和复制以及用于包封基因组RNA的启动子序列。

[0178] 5个结构基因是有序的(3' -5') N(核蛋白)、P(磷酸蛋白)、M(基质)、G(糖蛋白)和L(大),其具有短的非编码基因间区(IGR)。每个结构基因均包含侧接3' 转录起始信号(transcription initiation signal,TIS)(共有序列3' -U-U-G-U-R-R-n-G-A-5')和5' 转录终止多腺苷酸化(transcription termination polyadenylation,TTP)信号(共有序列3' A/U-C-U-U-U-U-U-U-G-5')的编码区。

[0179] 通过G蛋白与其在靶细胞表面上的受体结合来介导感染。然后,通过细胞的内体转运途径使病毒内化。内体中的低pH值诱导膜融合(在内体膜和病毒包膜之间)(也由G蛋白介导),将核糖核蛋白复合物释放到细胞质中。

[0180] 然后编码病毒蛋白的mRNA的转录开始,其由病毒聚合酶(包含P和L蛋白)介导。认为转录是通过终止启动机制进行的,从基因组的3' 端开始,并向5' 端进行,产生6个连续的转录本,首先是前导RNA,并且然后每个N、P、M、G和L基因依次出现。认为聚合酶在每个终止信号处与模板分离,并在下一个起始信号处重新启动。这导致mRNA产生的梯度,前导转录本的量最大,其次是N、P、M、G和L转录本。

[0181] 在感染之后,聚合酶从mRNA的产生转换到病毒基因组的复制,其通过全长的正义RNA中间体进行。正义中间体和子代负义病毒基因组二者都与蛋白质N一起包装以形成核蛋白复合物。M蛋白在决定转录和复制之间生物转换时间、以及参与将RNP核衣壳募集到宿主细胞质膜、与糖蛋白G的结合和子代病毒粒子颗粒从细胞中萌芽等方面起调节作用。

[0182] 有关综述,包括有关单个蛋白质的作用和功能的更详细信息,参见Albertini等, Rabies virus Transcription and Replication,Advances in Virus Research 79, December 2011(ISSN 0065-3527;DOI:10.1016/B978-0-12-387040-7.00001-9)。

[0183] 狂犬病N蛋白的示例性序列如下:

[0184] MDADKIVFKVNNQVVSLKPEIIVDQYEEKYPAIKDLKKPCITLGKAPDLNKAYKSVLSGM
SAAKLNPDVCSYLAAMQFFEGTCEPDWTSYGVIVARKGDKITPGSLVEIKRTDVEGNW
ALTGGMELTRDPTVPEHASLVGLLSLYRLSKISGQNTGNYKTNIADRIEQIFETAPFVK
IVEHHTLMTTHKMCANWSTIPNFRFLAGTYDMFFSRIEHLYSAIRVGTVTAYEDCSGLV
SFTGFIKQINLTAREAILYFFHKNFEEIIRMFEPGQETAVPHSYFIHFRSLGLSGKSPY
SSNAVGHVFNLIHFVGCYMGQVRSLNATVIAACAPHEMSVLGGYLGEEFFGKGTFERRFF
RDEKELQEYEAELTKTDVALADDGTVNSDDEDFSGETRSPYAVYTRIMMNGGRLKRSH
IRRYVSVSSNHQARPNSFAEFLNKTYSSDS

[0185] 狂犬病P蛋白的示例性序列如下:

[0186] MSKIFVNPSAIRAGLADLEMAEETVDLINRNIEDNQAHLQGEPIEVDNLPEDMGRHLHDD
GKSPNHGEIAKVGEKGYREDFQMDEGEDPSFLFQSYLENVGVQIVRQMRSGERFLKIWSQ
TVEEIIISYVAVNFPNPPGKSSSEDKSTQTTGRELKKETTPTPSQRESQSSKARMAAQIASG
PPALEWSATNEEDDLSEAEIAHQIAESFSKKYKFPSRSSGILLYNFEQLKMNLDIVKE
AKNVPGVTRLAHDGSKLPLRCVLGWVALANSKKFQLLVESDKLSKIMQDDLNRYSCT

[0187] 狂犬病M蛋白的示例性序列如下：

[0188] MNLLRKIVKNRRDEDTQKSSPASAPLDDDDLWLPPPEYVPLKELTGKKNMNRNFCINGRVK
VCSPNGYSFRILRHILKSFDEIYSGNHRMIGLVKVVIIGLALSGSPVEGLNWVYKLRRTF
[0189] IFQWADSRGPLEGEELEYSQEITWDDDETFVGLQIRVIAKQCHIQRVWCINMNPACQL
WSDMSLQTRSEEDKDSLLLE

[0190] 狂犬病毒L蛋白的示例性序列如下：

MLDPGEVYDDPIDPIELEAEPRGTPIVPNILRNSDYNLNSPLIEDPARLMLEWLKTGNRP
YRMTLTDNCSRSFRVLKDYFKKVDLGLSKVGGMAAQSMISLWLYGAHSESNRSRRCITDL
AHFYSSKSSPIEKLLNLTGLNRGLRIPPEGVLSCLERVDYDNAFGRYLANTYSSYLFFHVI
TLYMNALDWDEEKTILALWKDLTSVDIGKDLVKFKDQIWGLLIVTKDFVYSQSSNCLFDR
NYTLMKDLFLSRFNSLMVLLSPPEPRYSDDLISQLCQLYIAGDQVLSMCGNSGYEVIKI
LEPYVVNSLVQRAEKFRPLIHSLGDFPVFIKDKVSQLEETFGPCARRFFRALDQFDNIHD
LVFVFGCYRHWGHPYIDYRKGLSKLYDQVHLKKMIDKSYQECLASDLARRILRWGFDKYS
KWYLDLSRFLARDHPLTPYIKTQTWPPKHIVDLVGDTHWKLPITQIFEIPESMDPSEILDD
KSHSFTRTRLASWLSENRGGPVPSEKVIITALS KPPVNPREFLRSIDLGGLPDEDLIIGL
KPKERELKIEGRFFALMSWNLRLYFVITEKLLANYILPLFDALTMTDNLNKVFKKLIDRV
TGQGLLDYSRVTYAFHLDYEKWNHQRLESTEDVFSVLDQVFGKRVFSRTHFFQKAWI
YYSRSDSLIGLREDQIYCLDASNGPTCWNQDGGLEGLRQKGWSLVSLLMIDRESQIRNT
RTKILAQGDNQVLCPTYMLSPGLSQEGLLYELERISRNALSIYRAVEEGASKLGLIIKKE
ETMCSYDFLIYGKTPLFRGNILVPESKRWARVSCVSNDQIVNLANIMSTVSTNALTVAQH
SQSLIKPMRDFLLMSVQAVFHYLLFSPILKGRVYKILSAEGESFLLAMSRIIYLDPSLGG
ISGMSLGRFHIRQFSDPVSEGLSFWREIWLSSQESWIHALCQEAGNPD LGERTLESFTRL
LEDPTTLNIRGGASPTILLKDAIRKALYDEVDKVENSEFREAILLSKTHRDNFILFLISV
[0191] EPLFPRFLSELFSSSFLGIPESIIIGLIQNSRTIRRQFRKSLSKTLEESFYNSEIHGISRM
TQTPQRVGGVWPCSSERADLLREISWGRKVVGTTVPHPSEMLGLLPKSSISCTCGATGGG
NPRVSVSVLPSFDQSFFSRGPLKGYLGSSTSMSTQLFHAWEKVTNVHVVKRALSLKESIN
WFITRDSNLAQALIRNIMSLTGPDFLEAPVFKRTGSALHRFKSARYSEGGSVCNPL
LSHISVSTDTMSDLTQDGKNYDFMFQPLMLYAQTWTSELVQRDTRLRDLSTFWHLRCNRC
VRPIDDVTLETSQIFEFDPVSKRISRMVSGAVPHFQRLPDIRLRPGDFESLSGREKSHHI
GSAQGLLYSILVAIHDSGYNDGTIFPVNIYGVKVS PRDYLRGLARGVLIGSSICFLTRMTN
ININRPLELVSGVISYILLRLDNHPSLYIMLREPSLRGEI FSI PQI PAAYPTTMKEG NR
SILCYLQHVLR YEREIITASPENDWLWIFSDFRSAKMTYLSLITYQSHLLLRVERNLSK
SMRDNLRLQLSSLMRQVLGGHGEDTLESDDNIQRLLKDSLRRTRWVDQEVRAARTMTGDY
SPNKKVSRKVGCSSEWVCSAQQVAVSTSANPAPVSELDIRALSKRFQNP LISGLRVVQWAT
GAHYKLKPILDDLNVFPSLCLVVG DGSGGISRAVLNMFPDAKL VFNSLLEVNDLMASGTH
PLPPSAIMRGGNDIVSRVIDLDSIWEKPSDLRNLATWKYFQSVQKQVNMSYDLIICDAEV
TDIASINRITLLMSDFALSIDGPLYLVFKTYGTMLVNP NYKAIQHLSRAFP SVTG FITQV
TSSFSSELYLRF SKRGKFFRDAEYLTSSTLREMSLVLFNCSSPKSEMQRARSLNYQDLVR
GFPEEIIISNPYNEMIITLIDSDVESFLVHKMVDDLELQRTLSKVAIIIAIMIVFSNRVF
NVSKPLTDPSFYPPSPDKILRHFNICSTMMYLSTALGDVPSFARLHDLYNR PITYYFRK
QVIRGNVYLSWSWSNDTSVFKRVACNSSLSLSSH WIRLIYKIVKTTRLVGS IKDLSREVE
RHLHRYNRWITLEDIRSRSSLLDYSCL

[0192] 因此,当在本发明的载体中使用时,狂犬病N蛋白可具有以上所示的序列,或可与其具有至少70%同一性,例如至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、或至少99%同一性。

[0193] 狂犬病P蛋白可具有以上所示的序列,或可与其具有至少70%同一性,例如至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、或至少99%同一性。

[0194] 狂犬病M蛋白可具有以上所示的序列,或可与其具有至少70%同一性,例如至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、或至少99%同一性。

[0195] 狂犬病L蛋白可具有以上所示的序列,或可与其具有至少70%同一性,例如至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、或至少99%同一性。

[0196] 狂犬病G蛋白(毒株B19G)的示例性序列如下:

MVPQALLFVPLLVFPLCFGKFPIYTIIPDKLGPWSPIDIHHLSCPNNLVVEDEGCTNLSGFSY
MELKVGYILAIKVNGFTCTGVVTEAETYTNFVGIVTTTTFKRKHFRPTPDACRAAYNWKMGAD
PRYEESLHNPPDYRWLRTVKTTKESLVIISPSVADLDPYDRSLHSRVFPSPGKCSGVAVSST
YCSTNHDTIWMPEPNRLGMSCDIFTNSRGRASKGSETCGFVDERGLYKSLKGACKLKLGG
[0197] VLGLRLMDGTWVSMQTSNETKWCPDKLVNLHDFRSDEIEHLVVEELVRKREECLDALESIM
TTKSVSFRRLSHLRKLVPFGFKAYTIFNKTLMEADAHYKSVRTWNEILPSKGCLRVGGGRCHP
HVNGVFFNGIILGPDGNVLIPEMQSSLLQQHMELLESSVIPLVHPLADPSTVFKDGEAEDF
VEVHLDPDVHNQVSGVDLGLPNWGKYVLLSAGALTALMLIIIFLMTCCRRVNRSEPTQHNLRG
TGREVSVTPQSGKIISWESHKSGGETRL

[0198] 其中胞外结构域以常规字体显示、对跨膜结构域下划线、并且胞内结构域以斜体显示。

[0199] 作为替代的所谓的“优化的”G蛋白具有B19毒株胞内结构域但胞外结构域来自不同毒株,其具有以下序列:

MVPQALLFVPLLVFPLCFGKFPIYTIIPDKLGPWSPIDIHHLSCPNNLVVEDEGCTNLSGFSY
MELKVGYISAIKMNGFTCTGVVTEAETYTNFVGIVTTTTFKRKHFRPTPDACRAAYNWKMGAD
PRYEESLHNPPDYHWRVKTTKESLVIISPSVADLDPYDRSLHSPVFPNGNCSGVAVSST
YCSTNHDTIWMPEPNRLGMSCDIFTNSRGRASKGSETCGFVDERGLYKSLKGACKLKLGG
[0200] VLGLRLMDGTWVAMQTSNETKWCPGQLVNLHDFRSDEIEHLVVEELVKKREECLDALESIM
TTKSVSFRRLSHLRKLVPFGFKAYTIFNKTLMEADAHYKSVRTWNEIIPSKGCLRVGGGRCHP
HVNGVFFNGIILGPDGNVLIPEMQSSLLQQHMELLVSSVIPLMHPLADPSTVFKNGDEAEDF
VEVHLDPDVHERISGVDLGLPNWGKYVLLSAGALTALMLIIIFLMTCWRRVNRSEPTQHNLRG
TGREVSVTPQSGKIISWESHKSGGETRL

[0201] 其中胞外结构域以常规字体显示、对跨膜结构域下划线、并且胞内结构域以斜体显示。

[0202] 狂犬病G蛋白可具有以上显示的序列中的任意一个,或可与其具有至少70%同一性,例如至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、或至少99%同一性。

[0203] 与基于狂犬病病毒的载体一起使用的假型化包膜蛋白可具有来自以上显示的G蛋

白序列中的任意一个的胞内结构域和任选的跨膜结构域,或者可与其具有至少70%同一性,例如至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、或至少99%同一性。胞外结构域可与以上显示的胞外结构域序列中的一个具有至少70%同一性,例如至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、或至少99%同一性,或者可根据载体将靶向的细胞进行选择。

[0204] 由载体编码的每个基因(或开放阅读框)通常与适当的调节信号结合以确保表达(即转录和翻译)。这适用于载体携带的任何异源基因,以及编码复制调节蛋白和其他病毒蛋白的基因。因此,每个基因可包含转录起始信号和转录终止信号。转录起始信号通常是长度为约8个碱基的富含C的序列,例如ACATCCCT和ACACCCCT。终止信号通常是聚(A)片段,例如AAAAAAA。每个基因还可包含Kozak序列以确保适当的转录起始。还可存在另外的调节序列,包括非单股负链病毒元件,例如内部核糖体进入位点(internal ribosomal entry site, IRES)以确保期望的基因表达水平,特别是对于任何异源基因。

[0205] 相对于参考序列的百分比(%)氨基酸序列同一性被定义为在比对序列并引入缺口(如果必要的话)以实现最大百分比序列同一性之后并且不考虑作为序列同一性的一部分的任何保守替换的情况下,候选序列中与参考序列中的氨基酸残基相同的氨基酸残基的百分比。同一性%值可通过WU-BLAST-2确定(Altschul等,Methods in Enzymology,266:460-480(1996))。WU-BLAST-2使用数个检索参数,其中大多数参数设置为默认值。可调的参数设置为以下值:重叠跨度=1,重叠分数=0.125,字阈值(T)=11。如下确定氨基酸序列同一性%值:通过用如由WU-BLAST-2确定的匹配相同残基的数目除以参考序列残基的总数(忽略通过WU-BLAST-2引入到参考序列中以使比对得分最大化的空位),再乘以100。

[0206] 单突触回路示踪

[0207] 缺乏功能性糖蛋白G的狂犬病病毒载体用于映射神经元之间的突触连接是公知的。参见,例如,US 8,334,095 B2。

[0208] 狂犬病病毒仅在逆行方向上在突触连接的神经元之间选择性传播(除了一些感觉神经元之外)。缺乏G糖蛋白的狂犬病病毒不能从感染细胞的表面萌发。因此,基于狂犬病病毒(其中已缺失G基因或将其灭活)的载体(Δ G-狂犬病或RVdG)只能在反式(例如通过辅助病毒(例如腺相关病毒(adeno-associated viral, AAV)载体)提供编码G糖蛋白(或其他假型化包膜蛋白)的基因)的情况下向邻近细胞传播。如果将 Δ G-狂犬病病毒载体和辅助病毒小心地靶向同一细胞,则狂犬病病毒能够向突触邻近的细胞传播,但不能进一步传播(因为邻近细胞不表达进一步传播所期望的包膜蛋白)。因此,这使得突触连接能够被可视化和映射,例如通过 Δ G-狂犬病病毒载体携带的荧光标记蛋白。

[0209] 基本标记方法的许多变化是已知的。在哺乳动物用途的一些实例中,用来自A型或B型禽肉瘤白血病毒(ASLV)的包膜蛋白(EnvA、EnvB),其分别对受体TVA和TVB具有特异性,来对 Δ G-狂犬病进行假型化。由于这些蛋白质不是在哺乳动物中天然表达的,因此相关假型化载体的转导可限于特定工程化以表达TVA或TVB的细胞。

[0210] 用于突触映射的一些方法使用已经被工程化以表达病毒重组酶的动物模型,所述病毒重组酶可靶向狂犬病载体或辅助病毒基因组中的识别位点。通过将重组酶基因置于接受者动物的诱导型启动子的转录控制下,可达到进一步的控制水平。或者,接受者动物可已经被工程化为具有这样的重组酶的基因组识别位点,以便在用编码重组酶蛋白的载体进行

转导之后发生细胞基因组重组事件。

[0211] 通常,使用重组酶和识别位点的Cre-Lox配对,其中Cre重组酶作用于Lox识别位点。

[0212] 然而,G糖蛋白的缺失或失活不干扰接受者细胞内病毒转录、蛋白质合成和核心复制的正常过程。病毒RNA和蛋白质在细胞内的积累具有这样的作用,即标记的细胞通常仅维持存活1至2周,这是由于病毒蛋白质、针对感染细胞的免疫应答或其他机制直接损害了细胞活力。

[0213] 可转换载体

[0214] 本发明的载体利用单股负链病毒转录和复制的偶联性质为感染细胞内的病毒转录-复制循环提供开关,根据特定激活剂或抑制剂的存在(和活性)来打开或关闭病毒。

[0215] 如上所述,载体编码复制调节蛋白,所述复制调节蛋白可在展示靶向该蛋白以降解(例如通过蛋白酶体)的降解决定子的构型和不展示降解决定子并因此更稳定的构型之间转换。复制调节蛋白的这些构型被称为“靶向的”和“非靶向的”以反映降解决定子的存在或不存在。

[0216] 因此,虽然复制调节蛋白主要以靶向构型存在,但是病毒不能以功能性形式积累大量的相关蛋白。虽然在初次感染之后可能会立刻出现短期的病毒转录和复制,但此后病毒转录和复制将停止。如果这种构型维持足够长的时间段,病毒最终将从细胞清除,因为其他病毒蛋白和RNA基因组本身将会被细胞的正常机制分解。

[0217] 病毒将以靶向构型在感染的细胞中持续存在的时间长度将根据所讨论的特定细胞和病毒而变化。但是,一般来说,这种基于狂犬病病毒的载体很可能在约两周到一个月的时期内从神经细胞清除。

[0218] 一般来说,构型之间的转换由调节子部分与激活剂或抑制剂(视情况而定)之间的相互作用产生。相互作用可涉及调节子部分的共价或非共价修饰。

[0219] 共价修饰可通过酶作用来实现,即激活剂或抑制剂是酶,并且调节子部分是该酶的底物。共价修饰的一个实例是蛋白水解切割。

[0220] 非共价修饰可通过使激活剂或抑制剂与调节子部分结合来实现,即激活剂或抑制剂可以是调节子部分的配体。

[0221] 因此,激活剂可通过“切割诱导的稳定化”或“配体诱导的稳定化”使复制调节蛋白稳定。相反,抑制剂可通过“切割诱导的去稳定化”或“配体诱导的去稳定化”来使复制调节蛋白去稳定。(这些也可被称为“切割诱导的降解”或“配体诱导的降解”。)其他机制也是可以的。

[0222] 明显的是,取决于复制调节蛋白以及激活剂或抑制剂的性质,靶向和非靶向构型之间的转换可以是可逆的或不可逆的。例如,如果激活剂或抑制剂是调节蛋白的配体,则两种构型之间的转换可以是可逆的。如果构型之间的转换是通过共价修饰(例如蛋白水解)介导的,则转换可以是不可逆的。

[0223] 在本发明的上下文中,降解决定子可以是赋予蛋白质特别短的半衰期的任何特征,例如通过标记其以降解。已知许多不同类型的降解决定子。一些通过标记蛋白质进行泛素化而发挥作用(泛素依赖性降解决定子),而另一些则不依赖泛素。

[0224] 降解决定子可以是PEST序列。PEST序列是长度通常为至少12个氨基酸的亲水性并

且富含脯氨酸、谷氨酸、丝氨酸和苏氨酸的肽序列基序。Li等(Generation of Destabilized Green Fluorescent Protein as a Transcription Reporter, J.Biol.Chem.1998;273(52):34970-5)描述了小鼠鸟氨酸脱羧酶的第423至450位残基的PEST序列,其序列为HGFPPEVEEQDDGTLPMSCAQESGMDRH及其变体,所述变体也具有去稳定活性。在下面的实施例中使用该蛋白质第422至461位残基的区域(“mODC(422至461)”),并且其序列为SHGFPPEVEEQDDGTLPMSCAQESGMDRHPAACASARINV。

[0225] NPDC-1(神经增殖和分化控制蛋白-1)的C端区域也具有PEST基序,其序列为KELD TASSDEENEDGDFTVYECPLAPTGE MEVR。

[0226] 参见(NPDC-1, a Novel Regulator of Neuronal Proliferation, Is Degraded by the Ubiquitin/Proteasome System through a PEST Degradation Motif, Spencer等, J.Biol.Chem.2004;279,37069-37078)。

[0227] 因此,复制调节蛋白可包含PEST序列。包含PEST序列的复制调节蛋白通常将是抑制性调节蛋白,因为其将倾向于被降解直到去除PEST序列。

[0228] PEST序列将通常位于病毒蛋白部分的C端,例如在调节蛋白的C端。

[0229] 因此,复制调节蛋白通常将包含病毒蛋白部分和位于病毒蛋白部分C端的调节子部分,其中调节子部分包含蛋白酶切割位点和PEST序列。

[0230] 以下氨基酸序列是包含狂犬病病毒N蛋白作为病毒蛋白部分,以及包含TEV蛋白酶切割位点和PEST序列(“N-TEV-PEST”)的调节子部分的复制调节蛋白的实例:

[0231] MDADKIVFKVNNQVVS LKPEI IVDQY EYKYP AIKDLKKPCITLGKAPDLNKAYKSVLSGM
SAAKLNPDVCSYLAAMQFFEGT CPEDWTSYGIV IARKGDKITPGSLVEIKRTDVEGNW
ALTGGMELTRDPTVPEHASLVGLLLSLYRLSKISGQNTGNYKTNIADRIEQIFETAPFVK
IVEHHTLMTTHKMCANWSTIPNFRFLAGTYDMFFSRIEHLYSAIRVGT VVTAYEDCSGLV
SFTGFIKQINLTAREAILYFFHKNFEEEIIRMFEPGQETAVPHSYFIHFRSLGLSGKSPY
SSNAVGHVFNLIHFVGCYMGQVRSLNATVIAACAPHEMSVLGGYLGEFFGKGTFERRFF
RDEKELQEYEAELTKTDVALADDGTVNSDDEDFSGETRSP EAVYTRIMMNGGRLKRSH
IRRYVSVSSNHQARPNSFAEFLNKTYSSDSGSGENLYFQSGSGSHGFPPEVEEQDDGTL P
MSCAQESGMDRHPAACASARINV

[0232] N蛋白病毒蛋白部分以常规字体显示、调节子部分以斜体显示、TEV序列用单下划线、PEST序列用双下划线。

[0233] 降解决定子可以是N端氨基酸或短N端序列基序。公认的是,在蛋白质N端具有除Met、Gly或Val以外的残基的蛋白质倾向于比具有Met、Gly或Val的蛋白质更不稳定。这被称为N端规则。参见,例如,Varshavsky, A. (2011) “The N-end rule pathway and regulation by proteolysis”; Protein Science 20:1298-1345。

[0234] 多种机制都涉及通过蛋白质N端确定其稳定性,其包括所谓的Arg/N末端规则途径(其涉及蛋白质底物的N端精氨酰化)和Ac/N末端规则途径(其涉及具有Met、Ala、Val、Ser、Thr或Cys的某些蛋白质的共翻译N端乙酰化。在某些情况下,蛋白质的原始N端将被内源性细胞酶(例如Met-氨基肽酶)加工。在这样的情况下,加工之后存在的N端残基可代表降解决定子。

[0235] 作为准则,哺乳动物细胞中的稳定性顺序(最不稳定/最短半衰期至最稳定/最长半衰期)大致为Gln(Q)、Arg(R)、Glu(E)、Phe(F)、Asp(D)、Cys(C)、Lys(K)、Asn(N)、Ser(S)、Tyr(Y)、Trp(W)、His(H)、Ala(A)、Leu(L)、Thr(T)、Ile(I)、Pro(P)、Gly(G)、Met(M)、Val(V)。

[0236] 可针对任何给定的蛋白质测试不同的替代物,以鉴定在切割之前和之后在N端展示的残基的合适组合,以实现所需的蛋白质稳定性差异。

[0237] 因此,调节蛋白可包含N端调节子部分和在调节子部分下游的病毒蛋白部分,其中调节子部分在其N端包含第一残基以及关联蛋白酶的切割位点,并且其中在通过关联蛋白酶切割之后,病毒蛋白部分在其N端具有第二残基,该残基具有与第一残基不同的半衰期。

[0238] 因此,抑制性调节蛋白可包含第一N端残基,并且可被切割(例如通过蛋白酶)以暴露第二N端残基,该第二N端残基赋予比第一N端残基更强的稳定性。例如,第一N端残基可以是Arg或Lys(或序列Arg-Lys),并且第二N端残基可以是Val。

[0239] 相反,可抑制的调节蛋白可与第一N端残基一起表达,并且可被切割(例如通过蛋白酶)以暴露第二N端残基,该第二N端残基赋予比第一N端残基更低的稳定性。例如,第一N端残基可以是Met或Val,或序列Met-Val,并且第二N端残基可以是Arg或Lys(或序列Arg-Lys)。

[0240] 在两种情况下,调节子部分将位于分子的N端,并且可在紧邻所选第二N端残基的上游包含关联蛋白酶的切割位点。

[0241] 通常,调节蛋白将包含一个或更多个表面暴露的Lys残基,以充当泛素化的位点。

[0242] 上述的降解决定子技术提供了切割诱导的稳定化和切割诱导的去稳定化的实例。

[0243] 其他条件的降解决定子技术在Kanemaki等,Eur.J.Physiol.(2013) 465:419-425和其中引用的参考文献中进行了描述,并且以下提供了其中一些技术的更详细信息。

[0244] 配体诱导的去稳定化的一个实例是植物生长素诱导型(auxin-inducible)的降解决定子。这利用植物生长素来诱导靶向降解。植物生长素(例如吲哚-3-乙酸(IAA)或1-萘乙酸(NAA))与植物蛋白TIR1(被称为SCF的泛素连接酶组分)的结合使TIR1与AUX/IAA家族的蛋白质相互作用,导致AUX/IAA蛋白泛素化并且随后其被蛋白酶体降解。SCF泛素连接酶的其他组分在大多数真核生物中均保存良好。因此,AUX/IAA蛋白例如IAA17可在表达TIR1的细胞(例如已经被工程化以表达TIR1的真核非植物细胞)中用作调节子部分。向细胞添加植物生长素将导致包含AUX/IAA部分的任何蛋白质降解。因此,可抑制的调节子部分可包含AUX/IAA蛋白。

[0245] DD-FKBP和LID-FKBP降解决定子系统基于FK506结合蛋白12(FKBP12),并分别提供了配体诱导的稳定化和配体诱导的去稳定化的实例。

[0246] FKBP12的去稳定化结构域(“DD-FKBP”)被连续降解,但是通过被称为“Shield-1”的配体的存在而稳定,所述配体是雷帕霉素的细胞可渗透的类似物。(参见Banaszynski等,Cell.2006September 8;126(5):995-1004.)因此DD-FKBP可用作抑制性调节子部分,其可通过引入Shield-1来稳定。DD-FKBP部分可位于调节蛋白中病毒蛋白部分的N端或C端。

[0247] LID-FKBP系统是配体诱导的去稳定化(或降解)系统,其中具有序列TRGVVEVAEGVLLRRRGN的合成的19个氨基酸的降解决定子融合到FKBP12蛋白(或被称为FKBP12^{F36V}的其Phe36Val变体)的C端。该肽与FKBP12的配体结合袋结合,从而将其隔离。在相同的结合位点添加另一种配体,例如Shield-1,会诱导降解决定子肽的释放和暴露,其从而标记整个蛋白质以降解。因此,调节子部分可包含LID-FKBP部分。包含LID-FKBP部分的调节子部分将通常位于复制调节蛋白中病毒蛋白部分的C端,并且理想地位于调节子蛋白的C端。

[0248] 疏水性配体与调节子部分的结合也可用于标记蛋白质以降解。例如,所谓的“HaloTag”系统使用修饰的卤代烷脱氢酶(HaloTag)和与修饰的卤代烷脱氢酶的活性位点共价结合的小分子疏水性配体(例如HyT13)。似乎配体与蛋白质的结合标记蛋白质以通过蛋白酶体降解。因此,可抑制的调节子部分可包含HaloTag,其可位于病毒蛋白部分的N端或C端。

[0249] 从以上讨论将清楚的是,复制调节蛋白的靶向和非靶向构型之间的转换(反之亦然)可通过从病毒蛋白部分切割调节子部分来实现。这通常通过蛋白酶来完成。在这样的情况下,调节子部分通常包含降解决定子和蛋白酶的切割位点,其中蛋白酶切割位点位于病毒蛋白部分和降解决定子之间。复制调节蛋白可包含位于病毒蛋白部分和调节子部分之间的接头肽。另外地或替代地,调节子部分可在蛋白酶切割位点和降解决定子之间包含接头肽。

[0250] 除非蛋白质的组分中的一个具有任何特定的功能性要求,否则复制调节剂的部分可以处于任何适当的取向。例如,病毒蛋白可位于调节子部分的N端,或者调节子部分可位于病毒蛋白的N端,并且在它们之间具有接头(如果存在的话)。然而,某些降解决定子序列确实需要位于蛋白质的N或C端。

[0251] 肽接头长度通常为3至30个氨基酸,小的亲水性氨基酸残基(例如甘氨酸和丝氨酸)的比例高,以在不损害分子的水溶解度的情况下提供所需的柔性。其还可包含蛋白酶作用的切割位点。除了切割位点以外的残基(以及蛋白酶识别所需的任何其他序列)可包含至少50%的甘氨酸和丝氨酸残基、至少60%的甘氨酸和丝氨酸残基、至少70%的甘氨酸和丝氨酸残基、至少80%的甘氨酸和丝氨酸残基、或至少90%的甘氨酸和丝氨酸残基。

[0252] 蛋白酶

[0253] 蛋白酶可与靶细胞正交,这就是说蛋白酶识别在由靶细胞编码并在靶细胞中表达的天然蛋白质中(即在靶细胞的蛋白质组中)未发现的切割位点。

[0254] 因此,具体的蛋白酶可根据载体将要递送至的预期靶细胞而变化。靶细胞将通常是神经元,在这种情况下,蛋白酶不应作用于该神经元内表达的任何天然细胞蛋白。

[0255] 明显的是,所选蛋白酶不应作用于载体编码的其他蛋白质。

[0256] 合适的蛋白酶的实例包括:

[0257] -病毒蛋白酶,例如烟草蚀斑病毒蛋白酶(TEVp)和人鼻病毒(HRV)

[0258] 3C蛋白酶;

[0259] -因子Xa;

[0260] -肠激酶;

[0261] -颗粒酶B;

[0262] -凝血酶。

[0263] 这些蛋白酶的共有切割位点如下,其中“\”表示切割的肽键的位置:

[0264]	蛋白酶	关联切割位点
	TEVp	ENLYFQ\G ENLYFQ\S
	HRV 3C 蛋白酶	LEVLFQ\GP
	因子 Xa	IEGR\
	肠激酶	DDDDK\
	凝血酶	LVPR\GS

[0265] 药物组合物 and 治疗方法

[0266] 本文中描述的药物可配制成药物组合物。除上述物质之一之外,这些组合物还可包含可药用赋形剂、载体、缓冲剂、稳定剂或本领域技术人员公知的其他材料。这样的材料应该是无毒的并且不应干扰活性成分的效力。载体或其他材料的精确性质可取决于施用途径,其可以是任何合适的途径,并且可以是经口或肠胃外的。由于用经口递送肽试剂所遇到的困难,肠胃外施用可被证明是最合适的。合适的肠胃外途径包括但不限于静脉内、肌内、腹膜内、经皮肤、皮下、透皮和其他黏膜途径,例如经鼻、含服、经直肠和经阴道途径。合适的组合物和施用方法的一些实例在Esseku和Adeyeye (2011) 以及Van den Mooter G. (2006) 中提供。

[0267] 用于经口施用的药物组合物可以是片剂、胶囊剂、散剂或液体形式。片剂可包含固体载体,例如明胶或辅料。液体药物组合物通常包含液体载体,例如水、石油、动物或植物油、矿物油或者合成的油。可包含生理盐水溶液、右旋糖或其他糖溶液,或者二醇类(例如乙二醇、丙二醇或聚乙二醇)。

[0268] 用于静脉内、经皮肤或皮下注射,或在病痛部位注射,活性成分将是肠胃外可接受的水溶液的形式,其是无热原的并且具有合适的pH、等张性和稳定性。本领域相关技术人员能够很好地使用例如等张载剂(例如氯化钠注射液、林格注射液、乳酸林格注射液)制备合适的溶液。可根据需要包含防腐剂、稳定剂、缓冲剂、抗氧化剂和/或其他添加剂。

[0269] 无论待给予个体的活性剂的性质如何(例如病毒粒子、封装的核酸分子或根据本发明的其他药学上可用的试剂),优选以“预防有效量”或“治疗有效量”施用(视情况而定,尽管预防可被视为治疗),这足以显示对个体的益处。实际施用量以及施用的速率和时间过程将取决于接受治疗的情况的性质和严重程度。治疗处方(例如剂量决定等)在普通医师和其他医生的职责范围内,并且通常考虑待治疗的疾病、个体患者的情况、递送部位、施用方法以及医师了解的其他因素。以上提及的技术和方案的一些实例可见于Remington's Pharmaceutical Sciences,第20版,2000,公开,Lippincott,Williams&Wilkins。

实施例

[0270] 我们将靶向蛋白酶体的结构域与狂犬病病毒的每种蛋白质融合(单独或组合),以将其靶向蛋白酶体(图1A-B,表1中的构建体的完整列表)。为了对病毒蛋白稳定性实施水平条件控制,将烟草蚀斑病毒切割位点(TEV)插入病毒蛋白和蛋白酶体靶向结构域之间。烟草

蚀斑病毒蛋白酶 (TEVp) 选择性切割TEV接头,将病毒蛋白从靶向蛋白酶体的结构域中分离出来,使其免于降解(图1A)。由TEVp和TEVs构成的二元系统可通过药理调节以调节病毒产生和体内病毒蛋白降解的程度和时间窗口。即,该病毒仅当存在TEVp时才能够转录和复制,从而产生了系统,其中除非提供TEVp,否则病毒的转录和复制组成性关闭。

[0271] 我们筛选了所有生成的病毒构建体对病毒产生和TEV依赖性的适应性(图1)。筛选的结果指出,N蛋白是唯一的病毒蛋白,其条件去稳定化足以以TEVp依赖性方式可逆地抑制病毒转录复制周期(图1F、K)。所有其他病毒蛋白的去稳定化,无论是单独还是组合使用,都会使不能够体外扩增或在存在或不存在TEVp时都扩增的病毒产生(图1)。经过对病毒盒设计的进一步改进之后(补充文本,图6、7),我们能够基于具有期望的TEV依赖的ON-OFF动力学的N蛋白去稳定化狂犬病盒,生产出一种自灭活的狂犬病(SiR)。

[0272] 然后,我们生成了CRE重组酶和mCherry^{PEST}的SiR编码(SiR^{CRE-mCherry},图2B)。我们通过在Rosa-LoxP-STOP-LoxP-YFP小鼠的CA1锥体层中注射SiR^{CRE-mCherry}在体内测试了SiR转录-复制动力学和细胞毒性(图2B)。由SiR驱动的CRE重组酶的瞬时表达应确保Rosa基因座和YFP表达的永久重组,即使在病毒完全转录关闭之后也是如此(图2A)。确实,虽然感染后(p.i.) 3天仅能检测到病毒编码的mCherry(图8C-C’),但在感染后6天时,病毒编码的CRE重组酶诱导了条件鼠报道基因盒的重组和在所有被感染的神经元中YFP的表达(图8D-D’)。去稳定化的mCherry^{PEST}标记具有高时间分辨率(半衰期约2小时)的活性病毒的存在(21)。

[0273] 然后,我们评估了SiR^{CRE-mCherry}感染的CA1锥体神经元的存活(通过在8周时期内跟踪YFP^{ON}神经元),并监测了被感染的神经元中mCherry的关闭,以此作为病毒转录复制周期的代表。感染之后一周,SiR^{CRE-mCherry}在一部分被感染的神经元中开始关闭($29 \pm 3\%$ mCherry^{OFF} YFP^{ON},图2C-C’’,F)。感染之后三周,转换已完成,几乎所有YFP^{ON}神经元均没有显示病毒编码的mCherry的表达($98 \pm 2\%$ YFP^{ON} mCherry^{OFF},图2E-E’’,F,N=3)。更重要的是,在8周的观察期内,我们没有检测到SiR感染之后明显的神经元丢失(在8周时期内,YFP^{ON}神经元没有明显减少,单因素方差分析, $F=0.19$, $P=0.90$, $N=3$ /时间点,图2E-E’’,F)。相反,在 ΔG -狂犬病感染之后,海马中大部分被感染的神经元在初次感染之后2周内死亡(海马,在2周时 $92 \pm 3\%$ 细胞死亡, $n=3$ /时间点,单因素方差分析, $F=101$, $P=2.4 \times 10^{-5}$,图92 $\pm 3\%$)。为了在没有mCherry^{PEST}蛋白半衰期的警告的情况下获得更好的病毒转录复制周期的时间分辨率,我们在感染之后1、2、3和8周通过实时PCR对受感染动物的脑提取物测量了病毒RNA滴度。与mCherry表达结果一致,感染之后两周病毒滴度下降至接近背景水平(图2G)。总的来说,这些结果表明,SiR^{CRE-mCherry}转录复制动力学提供了足够的时间,以在病毒消失(图2G)之前产生早期CRE重组事件(图8D-D’’),其确保了在不影响神经元存活的情况下对被感染的神经元的永久遗传访问(图2F)。

[0274] 体外实验表明,通过条件蛋白酶体降解来调节病毒稳定性足以调节病毒转录-复制周期。为了评估是否可在体内实现对病毒蛋白降解的条件控制,我们设计了AAV病毒以在多西环素诱导型启动子下表达TEVp(AAV^{TRE::TEVp},图7A)(22)。这提供了体内TEVp表达的药理学控制手段,并且因此也提供了病毒转录复制周期的药理学控制手段。

[0275] 为了解决在感染之后是否可以以及直到感染之后的哪个阶段可通过多西环素施用来再激活SiR,我们在海马C1神经元中注射了AAV^{TRE::rEVp},随后1周后进行了SiR^{CRE-mCherry}感染。在一周之后或2周之后的两个时间点通过强饲法施用多西环素(100mg/Kg) 2天(图9B)。

与RNA分析一致,当病毒在SiR感染之后1周具有转录活性时,施用多西环素可使YFP^{ON} mCherry^{ON}神经元的百分比提高一倍(mCherry^{ON} YFP^{ON}-多西环素 $22 \pm 3\%$,+多西环素 38 ± 2 , $P=0.02$,图9G),而在感染之后2周通过施用多西环素没有检测到效果(mCherry^{ON} YFP^{ON}-多西环素 $1 \pm 1\%$,+多西环素 $1 \pm 1\%$,图9G)。

[0276] 考虑到SiR^{CRE-mCherry}的生命周期较短,其是否保留了在病毒关闭之前从初次感染的神经元跨突触传播的能力。为了对此进行测试,我们首先在Rosa-LoxP-STOP-LoxP-YFP小鼠的CA1锥体层中注射了表达TVA和B19-G的AAV,然后感染了EnvA假型SiR病毒。正如从CA1和CA3之间的已知解剖学连接所期望的那样,我们在CA3的锥体层中鉴定了由SiR标记的神经元(图3A-A”),这表明特定的跨突触传播。突触前神经元也被鉴定在内嗅皮层(图3A”)。更重要的是,表达TVA和G的被SiR感染的起始细胞在整个感染期间维持存活(图3B-B”)。此外,当用最近开发的优化狂犬病糖蛋白(oG) (17) 进行假型化时,SiR还可用作高效的逆行(非跨突触)示踪剂(图3C,D)。

[0277] 为了证实SiR对神经元生理学没有任何长期影响,我们在Rosa-LoxP-STOP-LoxP-ChR2YFP中的CA1的锥体层中注射了SiR^{CRE-mCherry}。然后,我们比较了感染之后1周和2个月被感染的神经元的电生理特性。在急性海马脑片中以全细胞膜片钳模式记录了CA1中的YFP^{ON}神经元。记录的所有锥体CA1神经元均显示规则的峰形分布(图4A,B),输入电阻无明显差异(在p.i.一周时为 $226 \pm 13 \text{M}\Omega$, $n=10$,而在两个月之后为 $251 \pm 18 \text{M}\Omega$, $n=8$;图4C),静息膜电位($-57.0 \pm 1.9 \text{mV}$ 与 $-56.7 \pm 1.9 \text{mV}$;图4D),动作电位波幅($96.4 \pm 3.6 \text{mV}$ 与 $94.2 \pm 2.9 \text{mV}$)和动作电位半宽($2.4 \pm 0.1 \text{ms}$ 与 $2.4 \pm 0.1 \text{ms}$;图4E)。峰形阈值中只发现很小的差异($-46.0 \pm 1.0 \text{mV}$ 与 $-41.3 \pm 2.0 \text{mV}$, $P=0.0394$;双尾双样本Student t检验,图4D);但是,这种差异不会影响神经元的瞬时放电频率(图4F)。ChR2在SiR^{CRE-mCherry}转导的神经元中成功表达(图4G-I)。在感染之后的一周和两个月以上时,神经元都可以以不同的频率被激活,具有相似的可靠性(图4J)。一个关键的发现是,感染SiR^{CRE-mCherry}的神经元的光激活在感染之后两个月时在其突触后伴侣中引起DNQX敏感的兴奋性突触后电位(EPSP)(图4K),表明功能性连接持续存在,并且对突触功能无不利影响。

[0278] 缺乏细胞毒性和未改变的电生理响应支持SiR用于长期回路操作。SiR感染的神经元之间存在功能连通性,并且对突触功能无不利影响,这也表明在SiR感染之后网络功能很可能得以保留。为了直接测试与网络依赖的计算在SiR映射回路中是否确实不受影响,我们用SiR示踪了投射到V2的V1神经元并表征了其方向调谐偏好,作为网络依赖的计算的典型示例(18)。我们首先通过在Rosa-LoxP-STOP-LoxP-tdTomato小鼠中注射oG假型SiR^{CRE}来靶向V2投射神经元。同时,我们在同侧V1中注射AAV::GCaMP6s。SiR^{CRE}的逆行传播诱导永久标记投射到V2的V1神经元的Rosa基因座的重组($V1^{>V2}$)(图5B-B”)。然后,我们在双光子显微镜下,在p.i.4周时在体内监测了被SiR感染的 $V1^{>V2}$ 神经元的Ca²⁺动力学,同时将麻醉的动物在视野中暴露于不同方向的移动光栅中(图5C)(19)。被感染的V1神经元在特定的光栅方向上显示出荧光的显著提高,导致显示出方向或方向选择性的调谐曲线(图5G,H)。值得注意的是,记录的Ca²⁺响应以及活性神经元的百分比在SiR示踪的神经元(GCaMP6s^{ON}-tdTomato^{OFF})和邻近的非SiR V1神经元(GCaMP6s^{ON}-tdTomato^{OFF})之间相似(图5E,F)。这些数据表明,SiR示踪的网络保留了不变的计算特性,并且SiR可与GCaMP6s结合使用来监测Ca²⁺的动态,而对光学研究的时间窗口没有上限。

[0279] 单突触限制的狂犬病病毒的开发在神经回路的研究中具有变革性作用。然而,直到现在,伴随狂犬病病毒感染的细胞毒性大致上已经严重地限制了其在神经回路解剖映射中的应用。诱导的细胞毒性与病毒的转录活性有关,后者操纵了细胞转录机制以维持病毒复制(24)。因此,任何可复制型的狂犬病病毒最终都会危害细胞生理。为了克服这一限制并获得对拓扑结构限定的网络元件的终生遗传和功能性访问,我们开发了自灭活狂犬病病毒,该病毒在人类胚胎干细胞(human Embryonic Stem cell, hESC)衍生的神经元(补充结果,图6)中和小鼠体内二者中以TEVp依赖性方式在初次感染之后转录关闭。与病毒的完全转录沉默相一致,感染和维持突触功能和回路完整性几个月之后,我们观察到被感染的神经元的电生理特征没有变化,这是生理和行为研究的关键。更重要的是,我们还表明,病毒感染之后的几个月内,被感染的神经元的更高水平的回路依赖性计算(例如其方向调整为运动刺激)在体内仍不受影响。

[0280] 反过来,这也表明SiR可用于使用钙成像监测体内的网络活动。对于典型的B19 Δ G-狂犬病,神经元光学成像的最佳时间窗口通常为狂犬病感染之后5至7天(17)。最近引入的 Δ G-狂犬病病毒株变体CVS-N2c ^{Δ G}的使用可将有用的时间窗口进一步推至感染之后长至17天进行成像(25)。使用SiR,对网络元件进行光学研究的时间窗口没有上限。这些属性使SiR成为长期监测神经网络活动以及神经回路的功能和遗传操作的最有价值的解决方案(26-29)。另外, SiR独特的瞬时复制特性,提供了在生理或病理结构可塑性之后,例如在学习时、神经退行性疾病期间或创伤性脑损伤后进行回路重塑的可能性,并且其可能为在这种情况下在网络水平进行功能性干预铺平了道路。

[0281] 总体而言,自灭活 Δ G狂犬病的发展首次提供了在不会对神经元生理和回路功能产生不利影响的情况下对拓扑结构限定的网络元件的永久遗传访问。

[0282] 补充案文

[0283] 第一代 Δ G-N^{PEST}狂犬病。体内和体外细胞毒性测试

[0284] 为了获得病毒蛋白质稳定性的条件化调节,在每种病毒蛋白质的C端添加SPLIT-TEV盒(30)(AG-VN^{PEST}狂犬病^{SPLIT-TEV-mCherry})。另外,标签(myc、FLAG或V5)与每种病毒蛋白质的N端融合,以监测蛋白质表达水平。SPLIT-TEV二聚体蛋白酶仅在雷帕霉素的存在下才有活性,并且可潜在地提供用于生产期间和体内病毒蛋白质稳定性的外源调节的工具。我们首先测试了由质粒表达的SPLIT-TEV切割HEK细胞中TEV活性报告物的能力(图6B),然后我们测试了病毒表达的盒切割TEV活性报告物的能力(图6C)。

[0285] 为了探测蛋白质去稳定化对神经元存活的影响,我们用 Δ G-N^{PEST}狂犬病^{SPLIT-TEV-mCherry}感染人胚胎干细胞(hESC)来源的神经元。我们对感染的神经元的存活进行了纵向研究,并将存活率与对照 Δ G狂犬病进行了比较。感染神经元并在4-10-16天纵向成像以评估细胞死亡(图7A, B-C'')。使用表达GFP的慢病毒来使感染率归一化,以解释由于长时间操作和重复的过夜成像环节导致的细胞死亡。在感染后16天,仅 $26 \pm 4\%$ 的对照 Δ G狂犬病感染的神经元仍然可检测(每种条件N=3, n=781;图7C-C'', D)。相比之下, Δ G-N^{PEST}狂犬病^{SPLIT-TEV-mCherry}病毒在16天后没有明显的细胞损失,为 $94 \pm 6\%$ (每种条件N=3, n=917,图4B-B'', C),以及与 Δ G狂犬病对照相比细胞存活显著提高($P = 3.2 \times 10^{-5}$;配对双尾Student t检验)。

[0286] 为了理解 Δ G-N^{PEST}狂犬病^{SPLIT-TEV-mCherry}降低的细胞毒性是否与病毒转录的降低相

关,我们监测了用具有相同表达盒的对照 ΔG 狂犬病和 ΔG -N^{PEST} 狂犬病^{SPLIT-TEV-mCherry} 病毒感染的神经元中表达的报告物的强度(图7E)。随着时间的流逝, ΔG -N^{PEST} 狂犬病^{SPLIT-TEV-mCherry} 感染的细胞的平均mCherry信号明显低于对照(10天时的mCherry强度, ΔG 狂犬病为 $138 \pm 3\%$, ΔG -N^{PEST} 狂犬病^{SPLIT-TEV-mCherry} 为 $87 \pm 7\%$, $P=0.01$;每种条件 $N=3$, $n=584$;配对双尾 Student t 检验)。

[0287] 我们然后在体内测试了 ΔG -N^{PEST} 狂犬病^{SPLIT-TEV-mCherry} 病毒的性能。

[0288] 我们将 ΔG -N^{PEST} 狂犬病^{SPLIT-TEV-mCherry} 病毒中的mCherry基因替换为CRE重组酶(ΔG -N^{PEST} 狂犬病^{SPLIT-TEV-CRE})。这确保了在病毒完全转录关闭之后可以永久标记感染的神经元,允许区分病毒沉默和细胞死亡。我们在海马的CA1中在Rosa-LoxP-STOP-LoxP-tdtomato报告物小鼠谱系的CA1中注射 ΔG -N^{PEST} 狂犬病^{SPLIT-TEV-CRE}。与用对照 ΔG 狂犬病感染后观察到的相比,我们观察到在 ΔG -N^{PEST} 狂犬病^{SPLIT-TEV-CRE} 感染之后神经元存活率显著提高(在2周时 ΔG -N^{PEST} 狂犬病^{SPLIT-TEV-CRE} 的为 $25 \pm 2\%$,在2周时 ΔG 狂犬病的为 $8 \pm 3\%$, $P=7 \times 10^{-3}$,图7H),然而,在 ΔG -N^{PEST} 狂犬病^{SPLIT-TEV-CRE} 感染之后,仍然存在显著的神经元丢失(在3周时为 $76 \pm 3\%$,第3周, $P=9 \times 10^{-4}$)。

[0289] ΔG -N^{PEST} 狂犬病^{SPLIT-TEV-CRE} 的残余细胞毒性可与SPLIT-TEV盒的组成型低基础二聚化和活性有关,并且可起源于转录活性病毒颗粒。与此假设相符,我们观察到在存在或不存在雷帕霉素的情况下,对神经元存活和mCherry表达水平无显著影响(接种后10天的mCherry表达, ΔG -N^{PEST} 狂犬病^{SPLIT-TEV-CRE} 为 $87 \pm 7\%$, ΔG -N^{PEST} 狂犬病^{SPLIT-TEV+CRE} 为 $85 \pm 6\%$, $P=0.21$;每种条件 $N=3$, $n=793$;配对双尾 Student t 检验;图7E),并且没有对细胞存活的显著影响与hESC来源神经元中的雷帕霉素施用相关(在第16天; ΔG -狂犬病^{+RAP} 为 $26 \pm 4\%$, ΔG -狂犬病^{-RAP} 为 $30 \pm 7\%$, $P=0.69$; ΔG -N^{PEST} 狂犬病^{SPLIT-TEV+CRE} 为 $88\% \pm 11\%$, ΔG -N^{PEST} 狂犬病^{SPLIT-TEV-CRE} 为 $94\% \pm 6\%$, $p=0.79$;每种条件 $N=3$, $n=833$;配对双尾 Student t 检验,图7D)。此外,我们观察到在不存在雷帕霉素的情况下HEK细胞中TEV活性的组成型低水平(图6B,第2行),表明不依赖雷帕霉素的SPLIT-TEV二聚化的基础水平。总的来说,这些结果表明,与 ΔG -狂犬病相比, ΔG -N^{PEST} 狂犬病^{SPLIT-TEV} 在hESC来源的神经元中和体内具有降低的细胞毒性。然而,其在感染之后不能完全关闭,导致显著延长但依然存在的神经元细胞毒性和神经元丧失。由于这些原因,我们通过移除泄漏的SPLIT-TEV并用mCherry-CRE盒替代产生了第二代 ΔG -N^{PEST} 狂犬病,得到了具有期望的开关和TEV依赖动力学的自灭活狂犬病病毒(Self inactivating Rabies virus,SIR)(正文)。

[0290] 材料与方法

[0291] 表1.

[0292] 构建体的完整列表。

[0293]

狂犬病	N-ter 标签	C-ter PEST
pSAD-F3-mCherry	-	-
pSAD-F3-N _{ter} TAGs-mCherry	N ^{Myc} -pFLAG-M ^{V5}	-
pSAD-F3-N _{ter} TAGs-N ^{PEST} -mCherry	N ^{Myc} -pFLAG-M ^{V5}	N ^{PEST}
pSAD-F3-N _{ter} TAGs-M ^{PEST} -mCherry	N ^{Myc} -pFLAG-M ^{V5}	M ^{PEST}
pSAD-F3-N _{ter} TAGs-P ^{PEST} -mCherry	N ^{Myc} -pFLAG-M ^{V5}	P ^{PEST}
pSAD-F3-N _{ter} TAGs-L ^{PEST} -mCherry	N ^{Myc} -pFLAG-M ^{V5}	L ^{PEST}
pSAD-F3-N ^{PEST} -mCherry	-	N ^{PEST}
pSAD-F3-N _{ter} TAGs-(P+L) ^{PEST} -mCherry	N ^{Myc} -pFLAG-M ^{V5}	P ^{PEST} L ^{PEST}
pSAD-F3-(P+L+N) ^{PEST} -mCherry	-	P ^{PEST} L ^{PEST} N ^{PEST}
pSAD-F3-N ^{PEST} -iCRE-2A-mCherryPEST	-	N ^{PEST}
pSAD-F3-N ^{PEST} -cTEV-FKBP-2A-FRB-NTEV-iCRE	-	N ^{PEST}
慢病毒	基因 A	基因 B
pLenti-H2B-GFP-2A-GlySAD	H2B-GFP	GlySAD
pLenti-puro-2A-TEV	嘌呤霉素	TEV
pLenti-GFP	GFP	-
pLenti-H2B-GFP-2A-oG	H2B-GFP	oG
AAVs	基因 A	基因 B
AAV-CMV-TVAmCherry-2A-Gly	TVAmCherry	Gly
AAV-TRE _{tight} -TEV-CMV-rTTA	TEV (多西环素依赖性)	rTTA
AAV-CAG-GCaMP6s	GCaMP6s	-

[0294] 动物品系

[0295] 使用了C57BL/6野生型(WT)小鼠和以下转基因品系:Rosa-LoxP-STOP-LoxP-tdtomato(Jackson:Gt(ROSA)26Soffml4(CAG tdTomato,Rosa-LoxP-STOP-LoxP-YFP(Jackson:Gt(ROSA)26Sor<tml(EYFP)Cos>)。所有程序均根据1986年英国动物法(科学程序)(UK Animals(Scientific procedures)Act 1986)和欧洲共同体动物保护指令(European Community Council Directive on Animal Care)进行。动物饲养在12小时光照/黑暗周期下,随意取用食物和水。

[0296] ΔG狂犬病和慢病毒质粒的设计和产生

[0297] 使用pSAD-ΔG-F3质粒(21)作为起始材料通过吉布森克隆(Gibson cloning)产生表1中列出的所有减毒狂犬病质粒。简单地说,将狂犬病基因组在从待标记的蛋白质开始的两个片段中PCR扩增。然后将这些片段与标签和/或通过寡核苷酸退火获得的PEST结构域混合,并使用吉布森主混合物(Gibson master mix)组装。

[0298] 用于产生包装细胞的慢病毒载体来源于第三代慢病毒转移载体(来自Michael Hastings“361polilinker”的惠赠,最初是pCCL-SIN-18PPT.Pgk.EGFP-WPRE)。所有慢病毒载体均通过吉布斯组装产生,通过XbaI和KpnI消化打开骨架并且对CMV启动子和不同插入片段进行PCR扩增。

[0299] 细胞系

[0300] HEK293T细胞和BHK购自ATCC。表达狂犬病糖蛋白的HEK293T包装细胞(HEK-GG)通过Lenti-H2B-GFP-2A-GlySAD(表1)的慢病毒感染产生,并且在3代之后进行表达GFP的细胞的

荧光激活细胞分选 (FACS)。表达狂犬病糖蛋白和TEV蛋白酶的HEK293T包装细胞 (HEK-TGG) 通过利用Lenti-puro-2A-TEV的慢病毒感染从HEK-GG产生,并且在3代之后利用添加到培养基中的1 μ g/mL嘌呤霉素进行选择1周。

[0301] 具有优化的G的假型狂犬病病毒的BHK包装细胞 (BHK-TGoG) 利用与HEK-TGG相同的程序产生,首先用pLenti-_{H2B}-GFP-2A-oG感染,随后用pLenti-puro-2A-TEV感染。具有EnVA受体的假型狂犬病病毒的BHK包装细胞 (BHK-T-EnVA) 通过用Lenti-puro-2A-TEV感染BHK-EnVA并且用嘌呤霉素进行选择产生。

[0302] 病毒筛选

[0303] 为了筛选减毒的 Δ G狂犬病病毒,用狂犬病基因组载体pcDNA-T7、pcDNA-SADB19N、pcDNA-SADB19P、pcDNA-SADB19L和pcDNA-SADB19G (21) 共转染HEK-GG或HEK-TGG细胞,并且在补充有10%FBS (Gibco) 和100u/ml青霉素-链霉素的DMEM中保持在37 $^{\circ}$ C、5%CO₂的潮湿气氛下。转染次日以及随后每3天,将细胞用PBS洗涤,用0.05%胰蛋白酶处理,并且以1:3的比例重新铺在新培养皿中。分割后,将细胞保持在37 $^{\circ}$ C、5%CO₂下1天,然后在35 $^{\circ}$ C、3%CO₂下2天。每三天,固定细胞,并通过细胞的mCherry表达的FACS分选来评估病毒的传播。

[0304] 病毒产生

[0305] 为了回收高滴度的 Δ G-狂犬病,将分别用于对照或减毒狂犬病的HEK-GG或HEK-TGG在10cm培养皿中在70-80%汇合度时用如病毒筛选部分中所述获得1ml上清液进行感染。在感染次日分割细胞,并且在37 $^{\circ}$ C、5%CO₂下保持1或2天,每天检查病毒传播。当70-80%的细胞表达病毒标志物时,将培养基替换为2%FBS DMEM,并在35 $^{\circ}$ C、3%CO₂下保持2天。然后,如先前所述收集病毒上清液,通过以2500rpm离心10分钟除去细胞碎片,然后用0.45 μ m过滤器过滤,然后通过用蔗糖垫超速离心浓缩病毒 (22)。

[0306] 通过在10cm培养皿中用1mL病毒上清感染BHK-T-oG细胞来产生用oG假型化的狂犬病病毒。在感染次日分割细胞,并且在37 $^{\circ}$ C、5%CO₂下保持1或2天,每天检查病毒传播。当70-80%的细胞表达病毒标志物时,将培养基替换为2%FBS DMEM,并在35 $^{\circ}$ C、3%CO₂下保持2天。然后,如先前所述收集上清液并进行处理 (20)。

[0307] 使用BVA-T-EnVA细胞代替BHK-EnVA细胞,如先前所述 (22) 产生用EnVA假型化的狂犬病病毒。

[0308] 体外细胞毒性分析

[0309] 人胚胎干细胞 (hESC) 来源的神经元由Rick Livesey博士友情提供。将细胞铺在24孔玻璃底板中,并且用减毒或对照 Δ G-狂犬病病毒上清液以可比的MOI过夜感染,获得约5%的感染细胞。感染后每4天将细胞在Hibernate®-A培养基 (Invitrogen) 中在37 $^{\circ}$ C加热的Leica SP8共聚焦显微镜中过夜成像,每孔成像5个随机视野。计算细胞存活率,将每种条件相对于在相同条件下成像和处理的对照慢病毒-GFP感染的hESC来源神经元的死亡率进行归一化。

[0310] 病毒注射

[0311] 所有使用活体动物的程序均由内政部和LMB生物安全委员会 (Home Office and the LMB Biosafety committee) 批准。对于所有实验,使用6-12周龄的小鼠。小鼠用异氟烷麻醉,异氟烷在开始诱导时以2L/min的氧气中的3%的流量递送,然后保持在2L/min的氧气中的1-2%。将麻醉的动物放入立体定位仪 (David Kopf Instruments) 中,并皮下 (s.c.) 施

用力莫敌(RimadV1) (2mg/kg体重)作为抗炎药。钻出小孔(直径为500 μ m),并使用汉密尔顿神经注射器(Hamilton neurosyringe)注射病毒。在撤回之前,将注射器留在脑中5分钟。以下列滴度注射病毒:狂犬病病毒为 3×10^8 个感染单位/ml,AAV为 2×10^{12} 个基因组拷贝/ml,慢病毒为 3×10^8 感染单位/ml。在以下脑区注射至多400nl体积的病毒:CA1(从前卤AP:2.3mm,ML:1.65mm且DV:1.45mm),V1(从前卤AP:2.3mm,ML:1.65mm且DV:1.45mm),V2(从前卤AP:3.6mm,ML:1.2mm且DV:0.6mm)。

[0312] 体内细胞毒性分析

[0313] 为了测试体内病毒的细胞毒性,将400nl相同滴度(3×10^8 个感染单位/ml)的减毒和对照 Δ G-狂犬病注射到2个半球对侧海马的CA1中。在接种后1-2-3或8周,在低温恒温器中将脑切片(35 μ m)。使用安装有10x(0.45NA)空气物镜的机器人辅助尼康(Nikon)HCA显微镜在对整个海马取样(每4个切片一次)的情况下对感染的神经元成像,并使用尼康HCA分析软件对海马荧光进行计数。计算减毒和对照 Δ G-狂犬病的细胞存活,并且将每个时间点相对于使用相同注射方案的对照慢病毒-GFP感染的海马的死亡率进行归一化。

[0314] 药物诱导的体内SiR病毒的再活化

[0315] 向Rosa-LoxP-STOP-LoxP-YFP动物在海马的CH1中注射在多西环素诱导型启动子的控制下组成型表达rTTA和TEV蛋白酶的AAV(表1)。在感染后1周,向相同区域再次注射SiR^{mCherry-CRE}并且在SiR注射后1或2周施用多西环素(Santa Cruz Biotechnology,100mg/Kg)。在药物施用后1周,收集脑并在低温恒温器中切片(35 μ m)。使用机器人辅助尼康HCA显微镜在对整个海马取样(每4个切片采集一次)的情况下对感染的神经元成像并计数。

[0316] 体内SiR基因组拷贝的分析

[0317] 为了评估感染动物中随时间推移的SiR病毒的基因组拷贝,向Rosa-LoxP-STOP-LoxP-YFP动物的海马的CA1区注射SiR^{mCherry-CRE}。在1、2、3或8周后,处死小鼠然后立即手动切下注射的海马。使用Tissuelyser II(Qiagen)将海马均质化,并按照生产说明书用RNeasy试剂盒(Qiagen)进行相应处理,以提取总RNA。使用superscript IV试剂盒(Invitrogen)对每个海马的500ng RNA进行逆转录(retrotranscribed),并通过定量PCR(Rotorgene sybr-green)分析GADPH、YFP和mCherry的表达。将Livak方法用于定量。将YFP和mCherry的表达相对于GADPH管家基因的表达进行归一化($DCT = CT_{\text{基因}} - CT_{\text{GADPH}}$),并且随着时间的变化作为1周时间点的倍数变化(2^{-DDCT}) ($DDCT = DCT_{\text{时间点}} - DCT_{1\text{周}}$)。

[0318] 电生理学

[0319] 对于电生理记录,将SiR^{mCherry-CRE}双侧注射到1月龄的Rosa-LoxP-STOP-LoxP-ChR2YFP小鼠的CA1中。在感染后1周或者2至3个月进行记录。

[0320] 使用振动切片机(7000smz-2,Campden Instruments LTD,Loughborough,UK)在用碳合气(carbogen gas)(95%O₂,5%CO₂)充氧并且具有以下组成(单位mM)的冰冷的基于蔗糖的切割溶液中制备冠状海马切片(350 μ m):KC1 3,NaH₂PO₄ 1.25,MgSO₄ 2,MgCl₂ 1,CaCl₂ 1,NaHCO₃ 26.4,葡萄糖10,蔗糖206,抗坏血酸0.40,犬尿酸(kynurenic acid)1。将切片在具有渗量调节至280-300mOsm/L的充氧的人工脑脊液(aCSF;单位mM:NaCl 126,KCl 3,NaH₂PO₄ 1.25,MgSO₄ 2,CaCl₂ 2,NaHCO₃ 26.4,葡萄糖10)的浸没式容纳室中在37℃下孵育30分钟,然后在室温下在同一容纳室中储存至少1小时。然后将切片分别转移到记录室中,并且在室温下以约2mL/min的流量用充氧的aCSF浇灌。

[0321] 使用从硼硅酸盐玻璃毛细管(1.5mm OD×0.86mm ID)拉出的6-9MΩ移液管,从CA1神经元获得全细胞电流钳记录。用人工细胞内溶液填充移液管,所述人工细胞内溶液含有(单位mM):葡萄糖酸K 150,HEPES 10,NaCl 4,ATP-MR 4,GTP-Na 0.3和EGTA 0.2;调节至pH 7.2和渗量270-290mOsm/L。使用Axon Multiclamp 700B放大器(Molecular Devices,Union City,CA,USA)记录数据,并且在运行pClamp的PC上使用数字化仪(Axon Digidata 1550A, Molecular Devices,Union City,CA,USA)对信号以2kHz进行低通滤波,并且以5kHz采集。通过450X水浸物镜(0.8NA),使用450-490nm LED光(pE-300coolLED system,Scientifica Ltd,Uckfield,UK)引起SiR病毒感染的神经元的光诱发反应。

[0322] 药理学

[0323] 将AMPA受体拮抗剂DNQX (20μM;Sigma-Aldrich,Dorset,UK)用在电生理学记录的子集中,以探测SiR病毒感染的神经元与邻近神经元之间的突触连通性。

[0324] 体内2光子成像

[0325] 注射的Rosa-LoxP-STOP-LoxP-tdTomato小鼠(见病毒注射部分)用2%异氟烷麻醉。测试动物捏缩和眼睑反射以测定麻醉深度。皮下注射力莫敌(2mg/kg体重)作为抗炎药。在实验期间,向两只眼睛覆盖眼药膏以防止角膜干燥。将动物头部固定,并且将金属头部杆固定在头骨上。在VI皮质上钻出直径为4mm的穿颅术。除去颅骨后,用皮质缓冲液保持皮质表面湿润,该缓冲液包含:125mM NaCl,5mM KCl,10mM葡萄糖,10mM HEPES,2mM MgSO₄和2mM CaCl₂,调节至pH 7.3。然后将皮质用定制的塞子盖玻片(23)覆盖,并用Super Glue和牙科胶黏剂密封。将小鼠用2%的异氟烷麻醉,并固定在双光子激光扫描显微镜(Multiphoton Imaging System,Scientifica Ltd.,Uckfield,United Kingdom)下,该显微镜配备有调谐在920nm处的Ti:蓝宝石锁模激光器(Mai Tai-Series,Spectra Physics)。通过水浸式透镜(Nikon,16X,0.8NA)在变焦2或4的256×256像素的分辨率下进行成像,分别得到了390×390μm和195×195μm的视场。以3.5Hz采集数据。物镜用黑色织物锥形物遮盖,该锥形物配有固定在顶板上的塑料O形圈(24)。使用Python中的定制GUI(基于PsychoPy工具箱)控制视觉刺激,并使用距小鼠左眼15cm放置的LED屏幕进行。移动方波光栅在12方向以30度步进呈现,并使用光电二极管检测每个刺激的开始和结束时间。每个光栅方向以随机顺序出现5次,并与空白条件交替。光栅的空间频率为0.04圈/度(cpd),并且时间频率为1Hz。成像和视觉刺激是使用Arduino微控制器板一起触发。成像过程持续长至2小时,并且样品的功率控制在30-40mW的范围内。数据分析在ImageJ和Matlab中进行,并且仅限于细胞体。使用Suite2p进行感兴趣区域(ROI)的检测。荧光的相对变化计算为 $dF/F_0 = (F(t) - F_c) / (F_0)$ 。通过获得整个刺激期的每个方向的平均响应产生方向调整曲线。响应幅度表示为刺激期间荧光与刺激前基线相比的相对变化(dF/F)。所有数据均以平均值±SE表示。

[0326] ***

[0327] 虽然本发明已结合上述示例性实施方案进行描述,但是当给出本公开内容时,许多等同的修改和变化对于本领域技术人员将是明显的。因此,所阐述的本发明的一些示例性实施方案被认为是举例说明性的而非限制性的。在不脱离本发明的精神和范围的情况下,可对所描述的实施方案进行多种改变。本文中引用的所有文献均明确地通过引用并入。

[0328] 本申请中所有参考文献的教导,包括专利申请和授权专利,均通过引用完全并入本文。本申请要求其优先权的任何专利申请均以本文所述的出版物和参考文献的方式通过

引用整体并入本文。

[0329] 为了避免疑义,发明人意图使本文中的术语“包含”及其任何形式在每种情况下分别任选地可被术语“由...组成”及其任何形式替代。所有数值中的术语“约”(或“大约”)允许5%的变化,即,约1.25%的值意指为1.11至1.31%。

[0330] 应理解,本文描述的特定实施方案是作为举例说明示出,而不作为本发明的限制。本发明的主要特征可用在多个实施方案中,而不脱离本发明的范围。本领域的技术人员将意识到或者仅仅利用常规的研究能够断定,与本文所描述的特定操作等同的许多实施方案。说明书中提及的所有出版物和专利申请表示本发明所属领域的技术人员的技术水平。所有出版物和专利申请通过引用并入本文,其程度如同每个单独的出版物或专利申请具体地和单独地指出通过引用并入。

[0331] 当在权利要求书和/或说明书中与术语“包括/包含”结合使用时,使用没有数量词修饰的名词可意指“一个/种”,但是其还符合“一个/种或更多个/种”、“至少一个/种”和“一个/种或多于一个/种”的含义。除非明确地指出仅指替代方案或者替代方案相互排斥,否则权利要求书中术语“或/或者”的使用用于意指“和/或”,但是本公开内容支持指仅替代方案和“和/或”的定义。贯穿本申请中,术语“约”用于表示值包括用于装置、用来确定所述值的方法的固有误差变异,或者所述研究对象之间存在的变异。

[0332] 如在本说明书和权利要求书中使用的,词语“包含”(以及任何形式的包含)、“具有”(以及任何形式的具有)、“包括”(以及任何形式的包括)或“含有”(以及任何形式的含有)是包含性的或开放式的,不排除另外的未记载的要素或方法步骤。

[0333] 如本文中使用的,术语“或其组合”是指在该术语之前列出的项目的所有排列和组合。例如,“A、B、C或其组合”旨在包括以下中的至少一种:A、B、C、AB、AC、BC或ABC,并且如果在特定情况下顺序是重要因素的话,也包括BA、CA、CB、CBA、BCA、ACB、BAC或CAB。继续这一示例,明确包括包含一个或更多个项目或术语的重复的组合,例如BB、AAA、MB、BBC、AAABCCCC、CBBAAB、CABABB等。本领域技术人员将理解,通常对于任意组合中的项目或术语的数量没有限制,除非从上下文来看明显。

[0334] 根据本公开内容,不需要过度实验就可以制备和实施本文公开和要求保护的所有组合物和/或方法。尽管按照一些优选实施方案描述了本发明的组合物和方法,然而对于本领域技术人员明显的是,可以对本文所述组合物和/或方法以及方法的步骤或步骤的顺序进行改变,而不脱离本发明的理念、精神和范围。对于本领域技术人员显而易见的所有此类相似的替代和改变被视为在由所附权利要求书所限定的本发明的精神、范围和概念内。

[0335] 参考文献

[0336] 1.A.R.McIntosh,Towards a network theory of cognition.Neural networks: the official journal of the International Neural Network Society 13,861(Oct-Nov,2000) .

[0337] 2.G.Buzsaki,W.Freeman,Editorial overview:brain rhythms and dynamic coordination.Curr Opin Neurobiol 31,v(Apr,2015) .

[0338] 3.V.Mnih et al.,Human-level control through deep reinforcement learning.Nature 518,529(Feb 26,2015) .

[0339] 4.D.Silver et al.,Mastering the game of Go with deep neural networks

and tree search. *Nature* 529,484(Jan 28,2016) .

[0340] 5.M.Rubinov,E.Bullmore,Fledgling pathoconnectomics of psychiatric disorders. *Trends in cognitive sciences* 17,641(Dec,2013) .

[0341] 6.F.Roselli,P.Caroni,A circuit mechanism for neurodegeneration. *Cell* 151,250(Oct 12,2012) .

[0342] 7.E.M.Callaway,Transneuronal circuit tracing with neurotropic viruses. *Current Opinion in Neurobiology* 18,617(Dec 01,2008) .

[0343] 8.M.P.van den Heuvel,E.T.Bullmore,o.Sporns,Comparative Connectomics. *Trends in cognitive sciences* 20,345(May,2016) .

[0344] 9.A.Cardona et al.,An integrated micro-and macroarchitectural analysis of the Drosophila brain by computer-assisted serial section electron microscopy. *PLoS Biol* 8, (2010) .

[0345] 10.J.Livet et al.,Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature* 450,56(Nov 01,2007) .

[0346] 11.P.Osten,T.W.Margrie,Mapping brain circuitry with a light microscope. *Nat Methods* 10,515(Jun,2013) .

[0347] 12.T.Mebatsion,M.Konig,K.K.Conzelmann,Budding of rabies virus particles in the absence of the spike glycoprotein. *Cell* 84,941(Apr 22,1996) .

[0348] 13.I.R.Wickersham et al.,Monosynaptic restriction of transsynaptic tracing from single,genetically targeted neurons. *Neuron* 53,639(Apr 01,2007) .

[0349] 14.M.Tripodi,A.E.Stepien,S.Arber,Motor antagonism exposed by spatial segregation and timing of neurogenesis. *Nature*,1(Oct 19,2011) .

[0350] 15.A.E.Stepien,M.Tripodi,S.Arber,Monosynaptic rabies virus reveals premotor network organization and synaptic specificity of cholinergic partition cells. *Neuron* 68,456(Nov 04,2010) .

[0351] 16.K.Miyamichi et al.,Cortical representations of olfactory input by trans-synaptic tracing. *Nature* 472,191(Dec 22,2010) .

[0352] 17.A.Wertz et al.,PRESYNAPTIC NETWORKS.single-cell-initiated monosynaptic tracing reveals layer-specific cortical network modules. *Science* 349,70(Jul 3,2015) .

[0353] 18.P.Namburi et al.,A circuit mechanism for differentiating positive and negative associations. *Nature* 520,675(Apr 30,2015) .

[0354] 19.J.Tian et al.,Distributed and Mixed Information in Monosynaptic Inputs to Dopamine Neurons. *Neuron* 91,1374(Sep 21,2016) .

[0355] 20.D.Atasoy,Y.Aponte,H.H.Su,S.M.Sternson,A FLEX switch targets Channelrhodopsin-2 to multiple cell types for imaging and long-range circuit mapping. *Journal of Neuroscience* 28,7025(Jul 09,2008) .

[0356] 21.X.Li et al.,Generation of Destabilized Green Fluorescent Protein as a Transcription Reporter*.273,34970(1998) .

[0357] 22.P.E.Shockett,D.G.Schatz,Diverse strategies for tetracycline-regulated inducible gene expression.Proceedings of the National Academy of Sciences of the United States of America 93,5173(1996) .

[0358] 23.A.Benucci,D.L.Ringach,M.Carandini,Coding of stimulus sequences by population responses in visual cortex.Nat Neurosci 12,1317(Oct,2009) .

[0359] 24.A.V.Komarova et al.,Rabies virus matrix protein interplay with eIF3,new insights into rabies virus pathogenesis.Nucleic Acids Res 35,1522 (2007) .

[0360] 25.T.R.Reardon et al.,Rabies Virus CVS-N2c Strain Enhances Retrograde Synaptic Transfer and Neuronal Viability.Neuron, (Jan 19,2016) .

[0361] 26.A.M.Packer,B.Roska,M.Hausser,Targeting neurons and photons for optogenetics.Nat Neurosci 16,805(Jul,2013) .

[0362] 27.A.M.Packer,L.E.Russell,H.W.Dalgleish,M.Hausser,Simultaneous all-optical manipulation and recording of neural circuit activity with cellular resolution in vivo.Nat Methods 12,140(Feb,2015) .

[0363] 28.L.Luo,E.M.Callaway,K.Svoboda,Genetic dissection of neural circuits.Neuron 57,634(Apr 13,2008) .

[0364] 29.V.Emiliani,A.E.Cohen,K.Deisseroth,M.Hausser,All-Optical Interrogation of Neural Circuits.J Neurosci 35,13917(Oct 14,2015) .

[0365] 31.S.FinKe,J.H.Cox,K-K Conzelmann,Differential Transcription Attenuation of Rabies Virus Genes by Intergenic Regions:Generation of Recombinant Viruses Overexpressing the Polymerase Gene.J.Virol.74(16) ,7261 (2000)

[0366] 32.Osakada and Callaway(2013)Design and generation of rabies virus vectors,Nature Protocols 8(8) :1583-1601.

[0367] 本发明的方面

[0368] 以下编号的段落(段落(paras.))包含本文中公开的发明技术特征的广泛组合的陈述:

[0369] 1.单股负链病毒载体基因组,其包含编码复制调节蛋白的基因,其中所述复制调节蛋白包含所述病毒基因组复制所需的单股负链病毒蛋白部分,所述复制调节蛋白能够采用展示降解决定子的靶向构型,和不展示所述降解决定子的非靶向构型。

[0370] 2.根据第1段所述的载体基因组,其中由所述载体基因组编码的所述复制调节蛋白是抑制性调节剂,其包含病毒蛋白部分和调节子部分,其中所述调节子部分包含所述降解决定子或由其组成。

[0371] 3.根据第2段所述的载体基因组,其中所述调节子部分可在与关联激活剂接触时转换至非靶向构型。

[0372] 4.根据第3段所述的载体基因组,其中所述激活剂从所述病毒蛋白部分切割所述调节子部分。

[0373] 5.根据第4段所述的载体基因组,其中所述激活剂是蛋白酶。

[0374] 6. 根据第5段所述的载体基因组,其中所述调节子部分包含所述蛋白酶的切割位点,所述切割位点位于所述病毒蛋白部分和所述降解决定子之间。

[0375] 7. 根据第5段或第6段所述的载体基因组,其中所述蛋白酶不作用于由所述载体基因组编码的任何其他蛋白质。

[0376] 8. 根据第5至7段中任一项所述的载体基因组,其中所述蛋白酶是病毒蛋白酶、因子Xa、肠激酶或凝血酶。

[0377] 9. 根据第2至8段中任一项所述的载体基因组,其中所述降解决定子是PEST序列。

[0378] 10. 根据第2至9段中任一项所述的载体基因组,其中所述调节蛋白包含第一N端残基,并且可被所述关联蛋白酶切割以暴露第二N端残基,所述第二N端残基赋予比所述第一N端残基更强的稳定性。

[0379] 11. 根据第3至10段中任一项所述的载体基因组,其中所述激活剂由所述载体基因组编码,并且其中所述激活的表达或功能是诱导型的。

[0380] 12. 根据第3段所述的载体基因组,其中所述激活剂是所述调节子部分的配体。

[0381] 13. 根据第12段所述的载体基因组,其中所述降解决定子是DD-FKBP序列,并且所述激活剂是其配体。

[0382] 14. 根据第1段所述的载体基因组,其中由所述载体基因组编码的复制调节蛋白是可抑制的调节蛋白,其可在与关联抑制剂接触时展示降解决定子的靶向构型。

[0383] 15. 根据第14段所述的载体基因组,其中所述复制调节蛋白包含病毒蛋白部分和调节子部分,并且其中所述抑制剂从所述调节蛋白切割所述调节子部分以产生或显露所述降解决定子。

[0384] 16. 根据第15段所述的载体基因组,其中所述激活剂是蛋白酶。

[0385] 17. 根据第14段所述的载体基因组,其中所述调节子部分包含所述蛋白酶的切割位点,所述切割位点位于所述病毒蛋白部分和所述降解决定子之间。

[0386] 18. 根据第14段或第15段所述的载体基因组,其中所述蛋白酶不作用于由所述载体基因组编码的任何其他蛋白质。

[0387] 19. 根据第16至18段中任一项所述的载体基因组,其中所述蛋白酶是病毒蛋白酶、因子Xa、肠激酶或凝血酶。

[0388] 20. 根据第16至19段中任一项所述的载体基因组,其中所述调节蛋白包含第一N端残基,并且可被所述蛋白酶切割以暴露第二N端残基,所述第二N端残基赋予比所述第一N端残基更低的稳定性。

[0389] 21. 根据第14段所述的载体基因组,其中所述复制调节蛋白包含病毒蛋白部分和调节子部分,其中所述抑制剂是所述调节子部分的配体,并且其中所述调节子部分和所述抑制剂是诱导型降解决定子系统的组成部分。

[0390] 22. 根据第21段所述的载体基因组,其中所述调节子部分包含HaloTag序列,并且所述抑制剂是其配体。

[0391] 23. 根据第21段所述的载体基因组,其中所述调节子部分包含LID-FKBP序列,并且所述抑制剂是其配体。

[0392] 24. 根据第21段所述的载体基因组,其中所述调节子部分包含植物生长素诱导型的降解决定子序列,并且所述抑制剂是其配体。

- [0393] 25. 根据第14至24段中任一项所述的载体基因组,其中所述抑制剂由所述载体基因组编码,并且其中所述抑制剂的表达或功能是诱导型。
- [0394] 26. 根据前述段落中任一项所述的载体基因组,其是弹状病毒载体基因组。
- [0395] 27. 根据第26段所述的载体基因组,其中所述载体基因组是狂犬病毒载体基因组或水疱性病毒载体基因组。
- [0396] 28. 根据第27段所述的载体基因组,其中所述狂犬病毒载体基因组是狂犬病病毒载体基因组,或者其中所述水疱性病毒载体基因组是水疱性口炎病毒载体基因组。
- [0397] 29. 根据前述段落中任一项所述的载体基因组,其中所述复制调节剂的所述病毒蛋白部分包含单股负链病毒N(核蛋白)蛋白或由其组成。
- [0398] 30. 根据第29段所述的载体基因组,其中所述基因组还包含编码P蛋白、M蛋白和/或L蛋白的基因。
- [0399] 31. 根据前述段落中任一项所述的载体基因组,其中所述载体基因组还包含编码包膜蛋白的基因。
- [0400] 32. 根据第31段所述的载体基因组,其中所述包膜蛋白对于单股负链病毒载体是天然的。
- [0401] 33. 根据第31段所述的载体基因组,其中所述包膜蛋白是假型化包膜蛋白。
- [0402] 34. 根据第1至30段中任一项所述的载体基因组,其中所述载体基因组不编码包膜蛋白。
- [0403] 35. 根据前述段落中任一项所述的载体基因组,其中所述载体基因组还包含一个或多个异源基因。
- [0404] 36. 根据第35段所述的载体基因组,其中所述异源基因编码标记蛋白、期望针对其产生免疫应答的蛋白质、重组酶、核酸酶、指导RNA(gRNA)分子、修复模板RNA或基因表达的核酸调节剂。
- [0405] 37. 根据第36段所述的载体基因组,其中所述核酸酶是RNA指导的内切核酸酶。
- [0406] 38. 根据第37段所述的载体基因组,其包含编码RNA指导的内切核酸酶加指导RNA(gRNA)分子和/或修复模板RNA的异源基因。
- [0407] 39. 核糖核蛋白复合物,其包含与一种或更多种病毒蛋白缔合的根据第1至38段中任一项所述的载体基因组。
- [0408] 40. 根据第39段所述的核糖核蛋白复合物,其包含与N、P和L蛋白缔合的所述载体基因组。
- [0409] 41. 根据第39段或第40段所述的核糖核蛋白复合物,其是功能性病毒核衣壳,其能够在引入靶细胞的细胞质时启动转录。
- [0410] 42. 单股负链病毒载体病毒粒子,其包含根据第1至38段中任一项所述的单股负链病毒载体基因组。
- [0411] 43. 根据第42段所述的载体病毒粒子,其包含天然单股负链病毒包膜蛋白。
- [0412] 44. 根据第42段所述的载体病毒粒子,其包含假型化包膜蛋白。
- [0413] 45. 根据第1至38段中任一项所述的载体基因组、根据第39至41段中任一项所述的核糖核蛋白复合物或根据第42至44段中任一项所述的载体病毒粒子,其用于医学治疗方法。

- [0414] 46. 根据第1至38段中任一项所述的载体基因组、根据第39至41段中任一项所述的核糖核蛋白复合物或根据第42至44段中任一项所述的载体病毒粒子,其用作免疫刺激剂。
- [0415] 47. 正义核酸分子,其编码根据第1至38段中任一项所述的病毒载体基因组。
- [0416] 48. 包装细胞,其包含编码根据第1至38段中任一项所述的载体基因组的核酸构建体并且能够产生根据第42至44段中任一项所述的病毒粒子。
- [0417] 49. 将基因递送至靶细胞的方法,其包括使所述靶细胞与根据第39至41段中任一项所述的核糖核蛋白复合物或根据第42至44段中任一项所述的病毒粒子接触。
- [0418] 50. 根据第49段所述的方法,其中所述载体编码抑制性调节蛋白,并且所述方法包括使所述细胞与关联激活剂接触。
- [0419] 51. 根据第50段所述的方法,其中所述激活剂是蛋白质,并且所述方法包括将包含编码所述激活剂的基因的核酸引入所述靶细胞中,使得所述激活剂在所述靶细胞中表达。
- [0420] 52. 根据第49段所述的方法,其中所述载体编码可抑制的调节蛋白,并且所述方法包括使所述细胞与关联抑制剂接触。
- [0421] 53. 根据第52段所述的方法,其中所述抑制剂是蛋白质,并且所述方法包括将包含编码所述抑制剂的基因的核酸引入所述靶细胞中,使得所述抑制剂在所述靶细胞中表达。
- [0422] 54. 根据第50至53段中任一项所述的方法,其中所述试剂的表达和/或功能是诱导型的,并且所述方法包括诱导所述试剂在所述靶细胞中的表达和/或功能的步骤。
- [0423] 55. 根据第49段所述的方法,其中:
- [0424] (i) 所述载体编码抑制性调节蛋白和激活剂,其中所述激活剂的表达或功能是诱导型的,并且其中所述方法包括诱导所述激活剂在所述靶细胞中的表达和/或功能的步骤;或者
- [0425] (ii) 所述载体编码可抑制的调节蛋白和抑制剂,其中所述抑制剂的表达或功能是诱导型的,并且其中所述方法包括诱导所述抑制剂在所述靶细胞中的表达和/或功能的步骤。
- [0426] 56. 根据第55段所述的方法,其包括使所述靶细胞与关联诱导剂接触。
- [0427] 57. 根据第49至56段中任一项所述的方法,其中所述载体基因组不包含编码包膜蛋白的基因,并且所述方法包括将包含编码包膜蛋白的基因的核酸构建体引入所述靶细胞中的步骤。
- [0428] 58. 根据第49至57段中任一项所述的方法,其中所述靶细胞是神经细胞。
- [0429] 59. 试剂盒,其包含根据第1至38段中任一项所述的载体基因组、根据第39至41段中任一项所述的核糖核蛋白复合物或根据第42至44段中任一项所述的病毒粒子,以及 (a) 关联激活剂或抑制剂或者编码关联激活剂或抑制剂的核酸;和/或 (b) 编码包膜蛋白的核酸。
- [0430] 60. 组合物,其包含根据第39至41段中任一项所述的核糖核蛋白复合物或根据第42至44段中任一项所述的病毒粒子,其任选地与赋形剂或载体混合。
- [0431] 61. 根据第60段所述的组合物,其中所述组合物是药物组合物,并且所述载体是可药用载体。

[0001] 序列表

[0002] <110> UNITED KINGDOM RESEARCH AND INNOVATION

[0003] <120> 自灭活病毒载体

[0004] <130> GRF/CP7361819

[0005] <140> PCT/GB2018/051166

[0006] <141> 2018-05-02

[0007] <150> GB1706945.1

[0008] <151> 2017-05-02

[0009] <160> 21

[0010] <170> PatentIn version 3.5

[0011] $\langle 210 \rangle$ 1

[0012] $\langle 211 \rangle$ 10

[0013] $\langle 212 \rangle$ RNA

[0014] <213> 人工序列

[0015] $\langle 220 \rangle$

[0016] <223> 转录起始信号共有序列

[0017] <220>

[0018] $\langle 221 \rangle$ misc feature

[0019] $\langle 222 \rangle$ (10) .. (10)

[0020] $\langle 223 \rangle$ N = A或U

[0021] <400> 1

[0022] guuuuuuuuucn 10

[0023] $\langle 210 \rangle$ 2

[0024] $\langle 211 \rangle$ 450

[0025] <212> PRT

[0026] <213> 人工序列

[0027] $\langle 220 \rangle$

[0028] <223> 狂犬病N蛋白

[0029] <400> 2

[0030] Met Asp Ala Asp Lys Ile Val Phe Lys Val Asn Asn Gln Val Val Ser

[0031] 1 5 10 15

[0032] Leu Lys Pro Glu Ile Ile Val Asp Gln Tyr Glu Tyr Lys Tyr Pro Ala

[0033] 20 25 30

[0034] Ile Lys Asp Leu Lys Lys Pro Cys Ile Thr Leu Gly Lys Ala Pro Asp

[0035] 35 40 45

[0036] Leu Asn Lys Ala Tyr Lys Ser Val Leu Ser Gly Met Ser Ala Ala Lys

[0037] 50 55 60

[0038] Leu Asn Pro Asp Asp Val Cys Ser Tyr Leu Ala Ala Ala Met Gln Phe

[0039] 65 70 75 80

[0040] Phe Glu Gly Thr Cys Pro Glu Asp Trp Thr Ser Tyr Gly Ile Val Ile

[0041] 85 90 95

[0042]	Ala Arg Lys Gly Asp Lys Ile Thr Pro Gly Ser Leu Val Glu Ile Lys
[0043]	100 105 110
[0044]	Arg Thr Asp Val Glu Gly Asn Trp Ala Leu Thr Gly Gly Met Glu Leu
[0045]	115 120 125
[0046]	Thr Arg Asp Pro Thr Val Pro Glu His Ala Ser Leu Val Gly Leu Leu
[0047]	130 135 140
[0048]	Leu Ser Leu Tyr Arg Leu Ser Lys Ile Ser Gly Gln Asn Thr Gly Asn
[0049]	145 150 155 160
[0050]	Tyr Lys Thr Asn Ile Ala Asp Arg Ile Glu Gln Ile Phe Glu Thr Ala
[0051]	165 170 175
[0052]	Pro Phe Val Lys Ile Val Glu His His Thr Leu Met Thr Thr His Lys
[0053]	180 185 190
[0054]	Met Cys Ala Asn Trp Ser Thr Ile Pro Asn Phe Arg Phe Leu Ala Gly
[0055]	195 200 205
[0056]	Thr Tyr Asp Met Phe Phe Ser Arg Ile Glu His Leu Tyr Ser Ala Ile
[0057]	210 215 220
[0058]	Arg Val Gly Thr Val Val Thr Ala Tyr Glu Asp Cys Ser Gly Leu Val
[0059]	225 230 235 240
[0060]	Ser Phe Thr Gly Phe Ile Lys Gln Ile Asn Leu Thr Ala Arg Glu Ala
[0061]	245 250 255
[0062]	Ile Leu Tyr Phe Phe His Lys Asn Phe Glu Glu Glu Ile Arg Arg Met
[0063]	260 265 270
[0064]	Phe Glu Pro Gly Gln Glu Thr Ala Val Pro His Ser Tyr Phe Ile His
[0065]	275 280 285
[0066]	Phe Arg Ser Leu Gly Leu Ser Gly Lys Ser Pro Tyr Ser Ser Asn Ala
[0067]	290 295 300
[0068]	Val Gly His Val Phe Asn Leu Ile His Phe Val Gly Cys Tyr Met Gly
[0069]	305 310 315 320
[0070]	Gln Val Arg Ser Leu Asn Ala Thr Val Ile Ala Ala Cys Ala Pro His
[0071]	325 330 335
[0072]	Glu Met Ser Val Leu Gly Gly Tyr Leu Gly Glu Glu Phe Phe Gly Lys
[0073]	340 345 350
[0074]	Gly Thr Phe Glu Arg Arg Phe Phe Arg Asp Glu Lys Glu Leu Gln Glu
[0075]	355 360 365
[0076]	Tyr Glu Ala Ala Glu Leu Thr Lys Thr Asp Val Ala Leu Ala Asp Asp
[0077]	370 375 380
[0078]	Gly Thr Val Asn Ser Asp Asp Glu Asp Tyr Phe Ser Gly Glu Thr Arg
[0079]	385 390 395 400
[0080]	Ser Pro Glu Ala Val Tyr Thr Arg Ile Met Met Asn Gly Gly Arg Leu
[0081]	405 410 415
[0082]	Lys Arg Ser His Ile Arg Arg Tyr Val Ser Val Ser Ser Asn His Gln
[0083]	420 425 430

[0084]	Ala Arg Pro Asn Ser Phe Ala Glu Phe Leu Asn Lys Thr Tyr Ser Ser
[0085]	435 440 445
[0086]	Asp Ser
[0087]	450
[0088]	<210> 3
[0089]	<211> 297
[0090]	<212> PRT
[0091]	<213> 人工序列
[0092]	<220>
[0093]	<223> 狂犬病P蛋白
[0094]	<400> 3
[0095]	Met Ser Lys Ile Phe Val Asn Pro Ser Ala Ile Arg Ala Gly Leu Ala
[0096]	1 5 10 15
[0097]	Asp Leu Glu Met Ala Glu Glu Thr Val Asp Leu Ile Asn Arg Asn Ile
[0098]	20 25 30
[0099]	Glu Asp Asn Gln Ala His Leu Gln Gly Glu Pro Ile Glu Val Asp Asn
[0100]	35 40 45
[0101]	Leu Pro Glu Asp Met Gly Arg Leu His Leu Asp Asp Gly Lys Ser Pro
[0102]	50 55 60
[0103]	Asn His Gly Glu Ile Ala Lys Val Gly Glu Gly Lys Tyr Arg Glu Asp
[0104]	65 70 75 80
[0105]	Phe Gln Met Asp Glu Gly Glu Asp Pro Ser Phe Leu Phe Gln Ser Tyr
[0106]	85 90 95
[0107]	Leu Glu Asn Val Gly Val Gln Ile Val Arg Gln Met Arg Ser Gly Glu
[0108]	100 105 110
[0109]	Arg Phe Leu Lys Ile Trp Ser Gln Thr Val Glu Glu Ile Ile Ser Tyr
[0110]	115 120 125
[0111]	Val Ala Val Asn Phe Pro Asn Pro Pro Gly Lys Ser Ser Glu Asp Lys
[0112]	130 135 140
[0113]	Ser Thr Gln Thr Thr Gly Arg Glu Leu Lys Lys Glu Thr Thr Pro Thr
[0114]	145 150 155 160
[0115]	Pro Ser Gln Arg Glu Ser Gln Ser Ser Lys Ala Arg Met Ala Ala Gln
[0116]	165 170 175
[0117]	Ile Ala Ser Gly Pro Pro Ala Leu Glu Trp Ser Ala Thr Asn Glu Glu
[0118]	180 185 190
[0119]	Asp Asp Leu Ser Val Glu Ala Glu Ile Ala His Gln Ile Ala Glu Ser
[0120]	195 200 205
[0121]	Phe Ser Lys Lys Tyr Lys Phe Pro Ser Arg Ser Ser Gly Ile Leu Leu
[0122]	210 215 220
[0123]	Tyr Asn Phe Glu Gln Leu Lys Met Asn Leu Asp Asp Ile Val Lys Glu
[0124]	225 230 235 240
[0125]	Ala Lys Asn Val Pro Gly Val Thr Arg Leu Ala His Asp Gly Ser Lys

[0126]		245		250		255
[0127]	Leu Pro Leu Arg Cys Val Leu Gly Trp Val Ala Leu Ala Asn Ser Lys					
[0128]		260		265		270
[0129]	Lys Phe Gln Leu Leu Val Glu Ser Asp Lys Leu Ser Lys Ile Met Gln					
[0130]		275		280		285
[0131]	Asp Asp Leu Asn Arg Tyr Thr Ser Cys					
[0132]		290		295		
[0133]	<210> 4					
[0134]	<211> 202					
[0135]	<212> PRT					
[0136]	<213> 人工序列					
[0137]	<220>					
[0138]	<223> 狂犬病M蛋白					
[0139]	<400> 4					
[0140]	Met Asn Leu Leu Arg Lys Ile Val Lys Asn Arg Arg Asp Glu Asp Thr					
[0141]	1	5		10		15
[0142]	Gln Lys Ser Ser Pro Ala Ser Ala Pro Leu Asp Asp Asp Asp Leu Trp					
[0143]		20		25		30
[0144]	Leu Pro Pro Pro Glu Tyr Val Pro Leu Lys Glu Leu Thr Gly Lys Lys					
[0145]		35		40		45
[0146]	Asn Met Arg Asn Phe Cys Ile Asn Gly Arg Val Lys Val Cys Ser Pro					
[0147]		50		55		60
[0148]	Asn Gly Tyr Ser Phe Arg Ile Leu Arg His Ile Leu Lys Ser Phe Asp					
[0149]	65	70		75		80
[0150]	Glu Ile Tyr Ser Gly Asn His Arg Met Ile Gly Leu Val Lys Val Val					
[0151]		85		90		95
[0152]	Ile Gly Leu Ala Leu Ser Gly Ser Pro Val Pro Glu Gly Leu Asn Trp					
[0153]		100		105		110
[0154]	Val Tyr Lys Leu Arg Arg Thr Phe Ile Phe Gln Trp Ala Asp Ser Arg					
[0155]		115		120		125
[0156]	Gly Pro Leu Glu Gly Glu Glu Leu Glu Tyr Ser Gln Glu Ile Thr Trp					
[0157]		130		135		140
[0158]	Asp Asp Asp Thr Glu Phe Val Gly Leu Gln Ile Arg Val Ile Ala Lys					
[0159]	145	150		155		160
[0160]	Gln Cys His Ile Gln Gly Arg Val Trp Cys Ile Asn Met Asn Pro Arg					
[0161]		165		170		175
[0162]	Ala Cys Gln Leu Trp Ser Asp Met Ser Leu Gln Thr Gln Arg Ser Glu					
[0163]		180		185		190
[0164]	Glu Asp Lys Asp Ser Ser Leu Leu Leu Glu					
[0165]		195		200		
[0166]	<210> 5					
[0167]	<211> 2127					

[0168]	<212>	PRT
[0169]	<213>	人工序列
[0170]	<220>	
[0171]	<223>	狂犬病L蛋白
[0172]	<400>	5
[0173]	Met Leu Asp Pro Gly Glu Val Tyr Asp Asp Pro Ile Asp Pro Ile Glu	
[0174]	1 5 10 15	
[0175]	Leu Glu Ala Glu Pro Arg Gly Thr Pro Ile Val Pro Asn Ile Leu Arg	
[0176]	20 25 30	
[0177]	Asn Ser Asp Tyr Asn Leu Asn Ser Pro Leu Ile Glu Asp Pro Ala Arg	
[0178]	35 40 45	
[0179]	Leu Met Leu Glu Trp Leu Lys Thr Gly Asn Arg Pro Tyr Arg Met Thr	
[0180]	50 55 60	
[0181]	Leu Thr Asp Asn Cys Ser Arg Ser Phe Arg Val Leu Lys Asp Tyr Phe	
[0182]	65 70 75 80	
[0183]	Lys Lys Val Asp Leu Gly Ser Leu Lys Val Gly Gly Met Ala Ala Gln	
[0184]	85 90 95	
[0185]	Ser Met Ile Ser Leu Trp Leu Tyr Gly Ala His Ser Glu Ser Asn Arg	
[0186]	100 105 110	
[0187]	Ser Arg Arg Cys Ile Thr Asp Leu Ala His Phe Tyr Ser Lys Ser Ser	
[0188]	115 120 125	
[0189]	Pro Ile Glu Lys Leu Leu Asn Leu Thr Leu Gly Asn Arg Gly Leu Arg	
[0190]	130 135 140	
[0191]	Ile Pro Pro Glu Gly Val Leu Ser Cys Leu Glu Arg Val Asp Tyr Asp	
[0192]	145 150 155 160	
[0193]	Asn Ala Phe Gly Arg Tyr Leu Ala Asn Thr Tyr Ser Ser Tyr Leu Phe	
[0194]	165 170 175	
[0195]	Phe His Val Ile Thr Leu Tyr Met Asn Ala Leu Asp Trp Asp Glu Glu	
[0196]	180 185 190	
[0197]	Lys Thr Ile Leu Ala Leu Trp Lys Asp Leu Thr Ser Val Asp Ile Gly	
[0198]	195 200 205	
[0199]	Lys Asp Leu Val Lys Phe Lys Asp Gln Ile Trp Gly Leu Leu Ile Val	
[0200]	210 215 220	
[0201]	Thr Lys Asp Phe Val Tyr Ser Gln Ser Ser Asn Cys Leu Phe Asp Arg	
[0202]	225 230 235 240	
[0203]	Asn Tyr Thr Leu Met Leu Lys Asp Leu Phe Leu Ser Arg Phe Asn Ser	
[0204]	245 250 255	
[0205]	Leu Met Val Leu Leu Ser Pro Pro Glu Pro Arg Tyr Ser Asp Asp Leu	
[0206]	260 265 270	
[0207]	Ile Ser Gln Leu Cys Gln Leu Tyr Ile Ala Gly Asp Gln Val Leu Ser	
[0208]	275 280 285	
[0209]	Met Cys Gly Asn Ser Gly Tyr Glu Val Ile Lys Ile Leu Glu Pro Tyr	

[0210]	290	295	300
[0211]	Val Val Asn Ser Leu Val Gln Arg Ala Glu Lys Phe Arg Pro Leu Ile		
[0212]	305	310	315
[0213]	His Ser Leu Gly Asp Phe Pro Val Phe Ile Lys Asp Lys Val Ser Gln		
[0214]		325	330
[0215]	Leu Glu Glu Thr Phe Gly Pro Cys Ala Arg Arg Phe Phe Arg Ala Leu		335
[0216]		340	345
[0217]	Asp Gln Phe Asp Asn Ile His Asp Leu Val Phe Val Phe Gly Cys Tyr		350
[0218]		355	360
[0219]	Arg His Trp Gly His Pro Tyr Ile Asp Tyr Arg Lys Gly Leu Ser Lys		365
[0220]	370	375	380
[0221]	Leu Tyr Asp Gln Val His Leu Lys Lys Met Ile Asp Lys Ser Tyr Gln		
[0222]	385	390	395
[0223]	Glu Cys Leu Ala Ser Asp Leu Ala Arg Arg Ile Leu Arg Trp Gly Phe		400
[0224]		405	410
[0225]	Asp Lys Tyr Ser Lys Trp Tyr Leu Asp Ser Arg Phe Leu Ala Arg Asp		415
[0226]		420	425
[0227]	His Pro Leu Thr Pro Tyr Ile Lys Thr Gln Thr Trp Pro Pro Lys His		430
[0228]		435	440
[0229]	Ile Val Asp Leu Val Gly Asp Thr Trp His Lys Leu Pro Ile Thr Gln		445
[0230]	450	455	460
[0231]	Ile Phe Glu Ile Pro Glu Ser Met Asp Pro Ser Glu Ile Leu Asp Asp		
[0232]	465	470	475
[0233]	Lys Ser His Ser Phe Thr Arg Thr Arg Leu Ala Ser Trp Leu Ser Glu		480
[0234]		485	490
[0235]	Asn Arg Gly Gly Pro Val Pro Ser Glu Lys Val Ile Ile Thr Ala Leu		495
[0236]		500	505
[0237]	Ser Lys Pro Pro Val Asn Pro Arg Glu Phe Leu Arg Ser Ile Asp Leu		510
[0238]		515	520
[0239]	Gly Gly Leu Pro Asp Glu Asp Leu Ile Ile Gly Leu Lys Pro Lys Glu		525
[0240]	530	535	540
[0241]	Arg Glu Leu Lys Ile Glu Gly Arg Phe Phe Ala Leu Met Ser Trp Asn		
[0242]	545	550	555
[0243]	Leu Arg Leu Tyr Phe Val Ile Thr Glu Lys Leu Leu Ala Asn Tyr Ile		560
[0244]		565	570
[0245]	Leu Pro Leu Phe Asp Ala Leu Thr Met Thr Asp Asn Leu Asn Lys Val		575
[0246]		580	585
[0247]	Phe Lys Lys Leu Ile Asp Arg Val Thr Gly Gln Gly Leu Leu Asp Tyr		590
[0248]		595	600
[0249]	Ser Arg Val Thr Tyr Ala Phe His Leu Asp Tyr Glu Lys Trp Asn Asn		605
[0250]	610	615	620
[0251]	His Gln Arg Leu Glu Ser Thr Glu Asp Val Phe Ser Val Leu Asp Gln		

[0252]	625	630	635	640
[0253]	Val Phe Gly Leu Lys Arg Val Phe Ser Arg Thr His Glu Phe Phe Gln			
[0254]		645	650	655
[0255]	Lys Ala Trp Ile Tyr Tyr Ser Asp Arg Ser Asp Leu Ile Gly Leu Arg			
[0256]		660	665	670
[0257]	Glu Asp Gln Ile Tyr Cys Leu Asp Ala Ser Asn Gly Pro Thr Cys Trp			
[0258]		675	680	685
[0259]	Asn Gly Gln Asp Gly Gly Leu Glu Gly Leu Arg Gln Lys Gly Trp Ser			
[0260]		690	695	700
[0261]	Leu Val Ser Leu Leu Met Ile Asp Arg Glu Ser Gln Ile Arg Asn Thr			
[0262]	705	710	715	720
[0263]	Arg Thr Lys Ile Leu Ala Gln Gly Asp Asn Gln Val Leu Cys Pro Thr			
[0264]		725	730	735
[0265]	Tyr Met Leu Ser Pro Gly Leu Ser Gln Glu Gly Leu Leu Tyr Glu Leu			
[0266]		740	745	750
[0267]	Glu Arg Ile Ser Arg Asn Ala Leu Ser Ile Tyr Arg Ala Val Glu Glu			
[0268]		755	760	765
[0269]	Gly Ala Ser Lys Leu Gly Leu Ile Ile Lys Lys Glu Glu Thr Met Cys			
[0270]		770	775	780
[0271]	Ser Tyr Asp Phe Leu Ile Tyr Gly Lys Thr Pro Leu Phe Arg Gly Asn			
[0272]	785	790	795	800
[0273]	Ile Leu Val Pro Glu Ser Lys Arg Trp Ala Arg Val Ser Cys Val Ser			
[0274]		805	810	815
[0275]	Asn Asp Gln Ile Val Asn Leu Ala Asn Ile Met Ser Thr Val Ser Thr			
[0276]		820	825	830
[0277]	Asn Ala Leu Thr Val Ala Gln His Ser Gln Ser Leu Ile Lys Pro Met			
[0278]		835	840	845
[0279]	Arg Asp Phe Leu Leu Met Ser Val Gln Ala Val Phe His Tyr Leu Leu			
[0280]		850	855	860
[0281]	Phe Ser Pro Ile Leu Lys Gly Arg Val Tyr Lys Ile Leu Ser Ala Glu			
[0282]	865	870	875	880
[0283]	Gly Glu Ser Phe Leu Leu Ala Met Ser Arg Ile Ile Tyr Leu Asp Pro			
[0284]		885	890	895
[0285]	Ser Leu Gly Gly Ile Ser Gly Met Ser Leu Gly Arg Phe His Ile Arg			
[0286]		900	905	910
[0287]	Gln Phe Ser Asp Pro Val Ser Glu Gly Leu Ser Phe Trp Arg Glu Ile			
[0288]		915	920	925
[0289]	Trp Leu Ser Ser Gln Glu Ser Trp Ile His Ala Leu Cys Gln Glu Ala			
[0290]		930	935	940
[0291]	Gly Asn Pro Asp Leu Gly Glu Arg Thr Leu Glu Ser Phe Thr Arg Leu			
[0292]	945	950	955	960
[0293]	Leu Glu Asp Pro Thr Thr Leu Asn Ile Arg Gly Gly Ala Ser Pro Thr			

[0294]	965							970					975				
[0295]	Ile	Leu	Leu	Lys	Asp	Ala	Ile	Arg	Lys	Ala	Leu	Tyr	Asp	Glu	Val	Asp	
[0296]	980							985					990				
[0297]	Lys	Val	Glu	Asn	Ser	Glu	Phe	Arg	Glu	Ala	Ile	Leu	Leu	Ser	Lys	Thr	
[0298]	995							1000					1005				
[0299]	His	Arg	Asp	Asn	Phe	Ile	Leu	Phe	Leu	Ile	Ser	Val	Glu	Pro	Leu		
[0300]	1010							1015					1020				
[0301]	Phe	Pro	Arg	Phe	Leu	Ser	Glu	Leu	Phe	Ser	Ser	Ser	Phe	Leu	Gly		
[0302]	1025							1030					1035				
[0303]	Ile	Pro	Glu	Ser	Ile	Ile	Gly	Leu	Ile	Gln	Asn	Ser	Arg	Thr	Ile		
[0304]	1040							1045					1050				
[0305]	Arg	Arg	Gln	Phe	Arg	Lys	Ser	Leu	Ser	Lys	Thr	Leu	Glu	Glu	Ser		
[0306]	1055							1060					1065				
[0307]	Phe	Tyr	Asn	Ser	Glu	Ile	His	Gly	Ile	Ser	Arg	Met	Thr	Gln	Thr		
[0308]	1070							1075					1080				
[0309]	Pro	Gln	Arg	Val	Gly	Gly	Val	Trp	Pro	Cys	Ser	Ser	Glu	Arg	Ala		
[0310]	1085							1090					1095				
[0311]	Asp	Leu	Leu	Arg	Glu	Ile	Ser	Trp	Gly	Arg	Lys	Val	Val	Gly	Thr		
[0312]	1100							1105					1110				
[0313]	Thr	Val	Pro	His	Pro	Ser	Glu	Met	Leu	Gly	Leu	Leu	Pro	Lys	Ser		
[0314]	1115							1120					1125				
[0315]	Ser	Ile	Ser	Cys	Thr	Cys	Gly	Ala	Thr	Gly	Gly	Gly	Asn	Pro	Arg		
[0316]	1130							1135					1140				
[0317]	Val	Ser	Val	Ser	Val	Leu	Pro	Ser	Phe	Asp	Gln	Ser	Phe	Phe	Ser		
[0318]	1145							1150					1155				
[0319]	Arg	Gly	Pro	Leu	Lys	Gly	Tyr	Leu	Gly	Ser	Ser	Thr	Ser	Met	Ser		
[0320]	1160							1165					1170				
[0321]	Thr	Gln	Leu	Phe	His	Ala	Trp	Glu	Lys	Val	Thr	Asn	Val	His	Val		
[0322]	1175							1180					1185				
[0323]	Val	Lys	Arg	Ala	Leu	Ser	Leu	Lys	Glu	Ser	Ile	Asn	Trp	Phe	Ile		
[0324]	1190							1195					1200				
[0325]	Thr	Arg	Asp	Ser	Asn	Leu	Ala	Gln	Ala	Leu	Ile	Arg	Asn	Ile	Met		
[0326]	1205							1210					1215				
[0327]	Ser	Leu	Thr	Gly	Pro	Asp	Phe	Pro	Leu	Glu	Glu	Ala	Pro	Val	Phe		
[0328]	1220							1225					1230				
[0329]	Lys	Arg	Thr	Gly	Ser	Ala	Leu	His	Arg	Phe	Lys	Ser	Ala	Arg	Tyr		
[0330]	1235							1240					1245				
[0331]	Ser	Glu	Gly	Gly	Tyr	Ser	Ser	Val	Cys	Pro	Asn	Leu	Leu	Ser	His		
[0332]	1250							1255					1260				
[0333]	Ile	Ser	Val	Ser	Thr	Asp	Thr	Met	Ser	Asp	Leu	Thr	Gln	Asp	Gly		
[0334]	1265							1270					1275				
[0335]	Lys	Asn	Tyr	Asp	Phe	Met	Phe	Gln	Pro	Leu	Met	Leu	Tyr	Ala	Gln		

[0336]	1280	1285	1290
[0337]	Thr Trp Thr Ser Glu Leu Val	Gln Arg Asp Thr Arg	Leu Arg Asp
[0338]	1295	1300	1305
[0339]	Ser Thr Phe His Trp His Leu	Arg Cys Asn Arg Cys	Val Arg Pro
[0340]	1310	1315	1320
[0341]	Ile Asp Asp Val Thr Leu Glu	Thr Ser Gln Ile Phe	Glu Phe Pro
[0342]	1325	1330	1335
[0343]	Asp Val Ser Lys Arg Ile Ser	Arg Met Val Ser Gly	Ala Val Pro
[0344]	1340	1345	1350
[0345]	His Phe Gln Arg Leu Pro Asp	Ile Arg Leu Arg Pro	Gly Asp Phe
[0346]	1355	1360	1365
[0347]	Glu Ser Leu Ser Gly Arg Glu	Lys Ser His His Ile	Gly Ser Ala
[0348]	1370	1375	1380
[0349]	Gln Gly Leu Leu Tyr Ser Ile	Leu Val Ala Ile His	Asp Ser Gly
[0350]	1385	1390	1395
[0351]	Tyr Asn Asp Gly Thr Ile Phe	Pro Val Asn Ile Tyr	Gly Lys Val
[0352]	1400	1405	1410
[0353]	Ser Pro Arg Asp Tyr Leu Arg	Gly Leu Ala Arg Gly	Val Leu Ile
[0354]	1415	1420	1425
[0355]	Gly Ser Ser Ile Cys Phe Leu	Thr Arg Met Thr Asn	Ile Asn Ile
[0356]	1430	1435	1440
[0357]	Asn Arg Pro Leu Glu Leu Val	Ser Gly Val Ile Ser	Tyr Ile Leu
[0358]	1445	1450	1455
[0359]	Leu Arg Leu Asp Asn His Pro	Ser Leu Tyr Ile Met	Leu Arg Glu
[0360]	1460	1465	1470
[0361]	Pro Ser Leu Arg Gly Glu Ile	Phe Ser Ile Pro Gln	Lys Ile Pro
[0362]	1475	1480	1485
[0363]	Ala Ala Tyr Pro Thr Thr Met	Lys Glu Gly Asn Arg	Ser Ile Leu
[0364]	1490	1495	1500
[0365]	Cys Tyr Leu Gln His Val Leu	Arg Tyr Glu Arg Glu	Ile Ile Thr
[0366]	1505	1510	1515
[0367]	Ala Ser Pro Glu Asn Asp Trp	Leu Trp Ile Phe Ser	Asp Phe Arg
[0368]	1520	1525	1530
[0369]	Ser Ala Lys Met Thr Tyr Leu	Ser Leu Ile Thr Tyr	Gln Ser His
[0370]	1535	1540	1545
[0371]	Leu Leu Leu Gln Arg Val Glu	Arg Asn Leu Ser Lys	Ser Met Arg
[0372]	1550	1555	1560
[0373]	Asp Asn Leu Arg Gln Leu Ser	Ser Leu Met Arg Gln	Val Leu Gly
[0374]	1565	1570	1575
[0375]	Gly His Gly Glu Asp Thr Leu	Glu Ser Asp Asp Asn	Ile Gln Arg
[0376]	1580	1585	1590
[0377]	Leu Leu Lys Asp Ser Leu Arg	Arg Thr Arg Trp Val	Asp Gln Glu

[0378]	1595	1600	1605
[0379]	Val Arg His Ala Ala Arg Thr Met Thr Gly Asp Tyr Ser Pro Asn		
[0380]	1610	1615	1620
[0381]	Lys Lys Val Ser Arg Lys Val Gly Cys Ser Glu Trp Val Cys Ser		
[0382]	1625	1630	1635
[0383]	Ala Gln Gln Val Ala Val Ser Thr Ser Ala Asn Pro Ala Pro Val		
[0384]	1640	1645	1650
[0385]	Ser Glu Leu Asp Ile Arg Ala Leu Ser Lys Arg Phe Gln Asn Pro		
[0386]	1655	1660	1665
[0387]	Leu Ile Ser Gly Leu Arg Val Val Gln Trp Ala Thr Gly Ala His		
[0388]	1670	1675	1680
[0389]	Tyr Lys Leu Lys Pro Ile Leu Asp Asp Leu Asn Val Phe Pro Ser		
[0390]	1685	1690	1695
[0391]	Leu Cys Leu Val Val Gly Asp Gly Ser Gly Gly Ile Ser Arg Ala		
[0392]	1700	1705	1710
[0393]	Val Leu Asn Met Phe Pro Asp Ala Lys Leu Val Phe Asn Ser Leu		
[0394]	1715	1720	1725
[0395]	Leu Glu Val Asn Asp Leu Met Ala Ser Gly Thr His Pro Leu Pro		
[0396]	1730	1735	1740
[0397]	Pro Ser Ala Ile Met Arg Gly Gly Asn Asp Ile Val Ser Arg Val		
[0398]	1745	1750	1755
[0399]	Ile Asp Leu Asp Ser Ile Trp Glu Lys Pro Ser Asp Leu Arg Asn		
[0400]	1760	1765	1770
[0401]	Leu Ala Thr Trp Lys Tyr Phe Gln Ser Val Gln Lys Gln Val Asn		
[0402]	1775	1780	1785
[0403]	Met Ser Tyr Asp Leu Ile Ile Cys Asp Ala Glu Val Thr Asp Ile		
[0404]	1790	1795	1800
[0405]	Ala Ser Ile Asn Arg Ile Thr Leu Leu Met Ser Asp Phe Ala Leu		
[0406]	1805	1810	1815
[0407]	Ser Ile Asp Gly Pro Leu Tyr Leu Val Phe Lys Thr Tyr Gly Thr		
[0408]	1820	1825	1830
[0409]	Met Leu Val Asn Pro Asn Tyr Lys Ala Ile Gln His Leu Ser Arg		
[0410]	1835	1840	1845
[0411]	Ala Phe Pro Ser Val Thr Gly Phe Ile Thr Gln Val Thr Ser Ser		
[0412]	1850	1855	1860
[0413]	Phe Ser Ser Glu Leu Tyr Leu Arg Phe Ser Lys Arg Gly Lys Phe		
[0414]	1865	1870	1875
[0415]	Phe Arg Asp Ala Glu Tyr Leu Thr Ser Ser Thr Leu Arg Glu Met		
[0416]	1880	1885	1890
[0417]	Ser Leu Val Leu Phe Asn Cys Ser Ser Pro Lys Ser Glu Met Gln		
[0418]	1895	1900	1905
[0419]	Arg Ala Arg Ser Leu Asn Tyr Gln Asp Leu Val Arg Gly Phe Pro		

[0420]	1910	1915	1920
[0421]	Glu Glu Ile Ile Ser Asn Pro Tyr Asn Glu Met Ile Ile Thr Leu		
[0422]	1925	1930	1935
[0423]	Ile Asp Ser Asp Val Glu Ser Phe Leu Val His Lys Met Val Asp		
[0424]	1940	1945	1950
[0425]	Asp Leu Glu Leu Gln Arg Gly Thr Leu Ser Lys Val Ala Ile Ile		
[0426]	1955	1960	1965
[0427]	Ile Ala Ile Met Ile Val Phe Ser Asn Arg Val Phe Asn Val Ser		
[0428]	1970	1975	1980
[0429]	Lys Pro Leu Thr Asp Pro Ser Phe Tyr Pro Pro Ser Asp Pro Lys		
[0430]	1985	1990	1995
[0431]	Ile Leu Arg His Phe Asn Ile Cys Cys Ser Thr Met Met Tyr Leu		
[0432]	2000	2005	2010
[0433]	Ser Thr Ala Leu Gly Asp Val Pro Ser Phe Ala Arg Leu His Asp		
[0434]	2015	2020	2025
[0435]	Leu Tyr Asn Arg Pro Ile Thr Tyr Tyr Phe Arg Lys Gln Val Ile		
[0436]	2030	2035	2040
[0437]	Arg Gly Asn Val Tyr Leu Ser Trp Ser Trp Ser Asn Asp Thr Ser		
[0438]	2045	2050	2055
[0439]	Val Phe Lys Arg Val Ala Cys Asn Ser Ser Leu Ser Leu Ser Ser		
[0440]	2060	2065	2070
[0441]	His Trp Ile Arg Leu Ile Tyr Lys Ile Val Lys Thr Thr Arg Leu		
[0442]	2075	2080	2085
[0443]	Val Gly Ser Ile Lys Asp Leu Ser Arg Glu Val Glu Arg His Leu		
[0444]	2090	2095	2100
[0445]	His Arg Tyr Asn Arg Trp Ile Thr Leu Glu Asp Ile Arg Ser Arg		
[0446]	2105	2110	2115
[0447]	Ser Ser Leu Leu Asp Tyr Ser Cys Leu		
[0448]	2120	2125	
[0449]	<210> 6		
[0450]	<211> 524		
[0451]	<212> PRT		
[0452]	<213> 人工序列		
[0453]	<220>		
[0454]	<223> 狂犬病G蛋白		
[0455]	<400> 6		
[0456]	Met Val Pro Gln Ala Leu Leu Phe Val Pro Leu Leu Val Phe Pro Leu		
[0457]	1 5 10 15		
[0458]	Cys Phe Gly Lys Phe Pro Ile Tyr Thr Ile Pro Asp Lys Leu Gly Pro		
[0459]	20 25 30		
[0460]	Trp Ser Pro Ile Asp Ile His His Leu Ser Cys Pro Asn Asn Leu Val		
[0461]	35 40 45		

[0462]	Val	Glu	Asp	Glu	Gly	Cys	Thr	Asn	Leu	Ser	Gly	Phe	Ser	Tyr	Met	Glu
[0463]	50						55				60					
[0464]	Leu	Lys	Val	Gly	Tyr	Ile	Leu	Ala	Ile	Lys	Val	Asn	Gly	Phe	Thr	Cys
[0465]	65					70					75					80
[0466]	Thr	Gly	Val	Val	Thr	Glu	Ala	Glu	Thr	Tyr	Thr	Asn	Phe	Val	Gly	Tyr
[0467]					85					90					95	
[0468]	Val	Thr	Thr	Thr	Phe	Lys	Arg	Lys	His	Phe	Arg	Pro	Thr	Pro	Asp	Ala
[0469]					100					105					110	
[0470]	Cys	Arg	Ala	Ala	Tyr	Asn	Trp	Lys	Met	Ala	Gly	Asp	Pro	Arg	Tyr	Glu
[0471]					115					120					125	
[0472]	Glu	Ser	Leu	His	Asn	Pro	Tyr	Pro	Asp	Tyr	Arg	Trp	Leu	Arg	Thr	Val
[0473]					130					135					140	
[0474]	Lys	Thr	Thr	Lys	Glu	Ser	Leu	Val	Ile	Ile	Ser	Pro	Ser	Val	Ala	Asp
[0475]	145					150					155					160
[0476]	Leu	Asp	Pro	Tyr	Asp	Arg	Ser	Leu	His	Ser	Arg	Val	Phe	Pro	Ser	Gly
[0477]					165					170						175
[0478]	Lys	Cys	Ser	Gly	Val	Ala	Val	Ser	Ser	Thr	Tyr	Cys	Ser	Thr	Asn	His
[0479]					180					185						190
[0480]	Asp	Tyr	Thr	Ile	Trp	Met	Pro	Glu	Asn	Pro	Arg	Leu	Gly	Met	Ser	Cys
[0481]					195					200						205
[0482]	Asp	Ile	Phe	Thr	Asn	Ser	Arg	Gly	Lys	Arg	Ala	Ser	Lys	Gly	Ser	Glu
[0483]					210					215						220
[0484]	Thr	Cys	Gly	Phe	Val	Asp	Glu	Arg	Gly	Leu	Tyr	Lys	Ser	Leu	Lys	Gly
[0485]	225					230					235					240
[0486]	Ala	Cys	Lys	Leu	Lys	Leu	Cys	Gly	Val	Leu	Gly	Leu	Arg	Leu	Met	Asp
[0487]					245					250						255
[0488]	Gly	Thr	Trp	Val	Ser	Met	Gln	Thr	Ser	Asn	Glu	Thr	Lys	Trp	Cys	Pro
[0489]					260					265						270
[0490]	Pro	Asp	Lys	Leu	Val	Asn	Leu	His	Asp	Phe	Arg	Ser	Asp	Glu	Ile	Glu
[0491]					275					280						285
[0492]	His	Leu	Val	Val	Glu	Glu	Leu	Val	Arg	Lys	Arg	Glu	Glu	Cys	Leu	Asp
[0493]					290					295						300
[0494]	Ala	Leu	Glu	Ser	Ile	Met	Thr	Thr	Lys	Ser	Val	Ser	Phe	Arg	Arg	Leu
[0495]	305					310					315					320
[0496]	Ser	His	Leu	Arg	Lys	Leu	Val	Pro	Gly	Phe	Gly	Lys	Ala	Tyr	Thr	Ile
[0497]					325					330						335
[0498]	Phe	Asn	Lys	Thr	Leu	Met	Glu	Ala	Asp	Ala	His	Tyr	Lys	Ser	Val	Arg
[0499]					340					345						350
[0500]	Thr	Trp	Asn	Glu	Ile	Leu	Pro	Ser	Lys	Gly	Cys	Leu	Arg	Val	Gly	Gly
[0501]					355					360						365
[0502]	Arg	Cys	His	Pro	His	Val	Asn	Gly	Val	Phe	Phe	Asn	Gly	Ile	Ile	Leu
[0503]					370					375						380

[0504]	Gly	Pro	Asp	Gly	Asn	Val	Leu	Ile	Pro	Glu	Met	Gln	Ser	Ser	Leu	Leu
[0505]	385						390				395					400
[0506]	Gln	Gln	His	Met	Glu	Leu	Leu	Glu	Ser	Ser	Val	Ile	Pro	Leu	Val	His
[0507]						405					410					415
[0508]	Pro	Leu	Ala	Asp	Pro	Ser	Thr	Val	Phe	Lys	Asp	Gly	Asp	Glu	Ala	Glu
[0509]					420					425					430	
[0510]	Asp	Phe	Val	Glu	Val	His	Leu	Pro	Asp	Val	His	Asn	Gln	Val	Ser	Gly
[0511]			435						440						445	
[0512]	Val	Asp	Leu	Gly	Leu	Pro	Asn	Trp	Gly	Lys	Tyr	Val	Leu	Leu	Ser	Ala
[0513]		450						455						460		
[0514]	Gly	Ala	Leu	Thr	Ala	Leu	Met	Leu	Ile	Ile	Phe	Leu	Met	Thr	Cys	Cys
[0515]	465						470					475				480
[0516]	Arg	Arg	Val	Asn	Arg	Ser	Glu	Pro	Thr	Gln	His	Asn	Leu	Arg	Gly	Thr
[0517]						485						490				495
[0518]	Gly	Arg	Glu	Val	Ser	Val	Thr	Pro	Gln	Ser	Gly	Lys	Ile	Ile	Ser	Ser
[0519]					500					505					510	
[0520]	Trp	Glu	Ser	His	Lys	Ser	Gly	Gly	Glu	Thr	Arg	Leu				
[0521]				515						520						
[0522]	<210>	7														
[0523]	<211>	524														
[0524]	<212>	PRT														
[0525]	<213>	人工序列														
[0526]	<220>															
[0527]	<223>	优化的G蛋白														
[0528]	<400>	7														
[0529]	Met	Val	Pro	Gln	Ala	Leu	Leu	Phe	Val	Pro	Leu	Leu	Val	Phe	Pro	Leu
[0530]	1				5					10					15	
[0531]	Cys	Phe	Gly	Lys	Phe	Pro	Ile	Tyr	Thr	Ile	Pro	Asp	Lys	Leu	Gly	Pro
[0532]					20					25					30	
[0533]	Trp	Ser	Pro	Ile	Asp	Ile	His	His	Leu	Ser	Cys	Pro	Asn	Asn	Leu	Val
[0534]					35					40					45	
[0535]	Val	Glu	Asp	Glu	Gly	Cys	Thr	Asn	Leu	Ser	Gly	Phe	Ser	Tyr	Met	Glu
[0536]		50							55					60		
[0537]	Leu	Lys	Val	Gly	Tyr	Ile	Ser	Ala	Ile	Lys	Met	Asn	Gly	Phe	Thr	Cys
[0538]	65						70					75				80
[0539]	Thr	Gly	Val	Val	Thr	Glu	Ala	Glu	Thr	Tyr	Thr	Asn	Phe	Val	Gly	Tyr
[0540]						85						90				95
[0541]	Val	Thr	Thr	Thr	Phe	Lys	Arg	Lys	His	Phe	Arg	Pro	Thr	Pro	Asp	Ala
[0542]						100						105			110	
[0543]	Cys	Arg	Ala	Ala	Tyr	Asn	Trp	Lys	Met	Ala	Gly	Asp	Pro	Arg	Tyr	Glu
[0544]					115							120			125	
[0545]	Glu	Ser	Leu	His	Asn	Pro	Tyr	Pro	Asp	Tyr	His	Trp	Leu	Arg	Thr	Val

[0546]	130	135	140
[0547]	Lys Thr Thr Lys Glu Ser Leu Val Ile Ile Ser Pro Ser Val Ala Asp		
[0548]	145	150	155 160
[0549]	Leu Asp Pro Tyr Asp Arg Ser Leu His Ser Pro Val Phe Pro Gly Gly		
[0550]		165	170 175
[0551]	Asn Cys Ser Gly Val Ala Val Ser Ser Thr Tyr Cys Ser Thr Asn His		
[0552]		180	185 190
[0553]	Asp Tyr Thr Ile Trp Met Pro Glu Asn Pro Arg Leu Gly Met Ser Cys		
[0554]		195	200 205
[0555]	Asp Ile Phe Thr Asn Ser Arg Gly Lys Arg Ala Ser Lys Gly Ser Glu		
[0556]	210	215	220
[0557]	Thr Cys Gly Phe Val Asp Glu Arg Gly Leu Tyr Lys Ser Leu Lys Gly		
[0558]	225	230	235 240
[0559]	Ala Cys Lys Leu Lys Leu Cys Gly Val Leu Gly Leu Arg Leu Met Asp		
[0560]		245	250 255
[0561]	Gly Thr Trp Val Ala Met Gln Thr Ser Asn Glu Thr Lys Trp Cys Pro		
[0562]		260	265 270
[0563]	Pro Gly Gln Leu Val Asn Leu His Asp Phe Arg Ser Asp Glu Ile Glu		
[0564]		275	280 285
[0565]	His Leu Val Val Glu Glu Leu Val Lys Lys Arg Glu Glu Cys Leu Asp		
[0566]		290	295 300
[0567]	Ala Leu Glu Ser Ile Met Thr Thr Lys Ser Val Ser Phe Arg Arg Leu		
[0568]	305	310	315 320
[0569]	Ser His Leu Arg Lys Leu Val Pro Gly Phe Gly Lys Ala Tyr Thr Ile		
[0570]		325	330 335
[0571]	Phe Asn Lys Thr Leu Met Glu Ala Asp Ala His Tyr Lys Ser Val Arg		
[0572]		340	345 350
[0573]	Thr Trp Asn Glu Ile Ile Pro Ser Lys Gly Cys Leu Arg Val Gly Gly		
[0574]		355	360 365
[0575]	Arg Cys His Pro His Val Asn Gly Val Phe Phe Asn Gly Ile Ile Leu		
[0576]		370	375 380
[0577]	Gly Pro Asp Gly Asn Val Leu Ile Pro Glu Met Gln Ser Ser Leu Leu		
[0578]	385	390	395 400
[0579]	Gln Gln His Met Glu Leu Leu Val Ser Ser Val Ile Pro Leu Met His		
[0580]		405	410 415
[0581]	Pro Leu Ala Asp Pro Ser Thr Val Phe Lys Asn Gly Asp Glu Ala Glu		
[0582]		420	425 430
[0583]	Asp Phe Val Glu Val His Leu Pro Asp Val His Glu Arg Ile Ser Gly		
[0584]		435	440 445
[0585]	Val Asp Leu Gly Leu Pro Asn Trp Gly Lys Tyr Val Leu Leu Ser Ala		
[0586]		450	455 460
[0587]	Gly Ala Leu Thr Ala Leu Met Leu Ile Ile Phe Leu Met Thr Cys Trp		

[0588]	465	470	475	480
[0589]	Arg Arg Val Asn Arg Ser Glu Pro Thr Gln His Asn Leu Arg Gly Thr			
[0590]		485	490	495
[0591]	Gly Arg Glu Val Ser Val Thr Pro Gln Ser Gly Lys Ile Ile Ser Ser			
[0592]		500	505	510
[0593]	Trp Glu Ser His Lys Ser Gly Gly Glu Thr Arg Leu			
[0594]		515	520	
[0595]	<210>	8		
[0596]	<211>	28		
[0597]	<212>	PRT		
[0598]	<213>	人工序列		
[0599]	<220>			
[0600]	<223>	PEST序列		
[0601]	<400>	8		
[0602]	His Gly Phe Pro Pro Glu Val Glu Glu Gln Asp Asp Gly Thr Leu Pro			
[0603]	1	5	10	15
[0604]	Met Ser Cys Ala Gln Glu Ser Gly Met Asp Arg His			
[0605]		20	25	
[0606]	<210>	9		
[0607]	<211>	40		
[0608]	<212>	PRT		
[0609]	<213>	人工序列		
[0610]	<220>			
[0611]	<223>	mODC(422-461)		
[0612]	<400>	9		
[0613]	Ser His Gly Phe Pro Pro Glu Val Glu Glu Gln Asp Asp Gly Thr Leu			
[0614]	1	5	10	15
[0615]	Pro Met Ser Cys Ala Gln Glu Ser Gly Met Asp Arg His Pro Ala Ala			
[0616]		20	25	30
[0617]	Cys Ala Ser Ala Arg Ile Asn Val			
[0618]		35	40	
[0619]	<210>	10		
[0620]	<211>	34		
[0621]	<212>	PRT		
[0622]	<213>	人工序列		
[0623]	<220>			
[0624]	<223>	NPDC-1的C端区域		
[0625]	<400>	10		
[0626]	Lys Glu Leu Asp Thr Ala Ser Ser Asp Glu Glu Asn Glu Asp Gly Asp			
[0627]	1	5	10	15
[0628]	Phe Thr Val Tyr Glu Cys Pro Gly Leu Ala Pro Thr Gly Glu Met Glu			
[0629]		20	25	30

[0630]	Val Arg		
[0631]	<210>	11	
[0632]	<211>	503	
[0633]	<212>	PRT	
[0634]	<213>	人工序列	
[0635]	<220>		
[0636]	<223>	N-TEV-PEST	
[0637]	<400>	11	
[0638]	Met Asp Ala Asp Lys Ile Val Phe Lys Val Asn Asn Gln Val Val Ser		
[0639]	1	5	10 15
[0640]	Leu Lys Pro Glu Ile Ile Val Asp Gln Tyr Glu Tyr Lys Tyr Pro Ala		
[0641]		20	25 30
[0642]	Ile Lys Asp Leu Lys Lys Pro Cys Ile Thr Leu Gly Lys Ala Pro Asp		
[0643]		35	40 45
[0644]	Leu Asn Lys Ala Tyr Lys Ser Val Leu Ser Gly Met Ser Ala Ala Lys		
[0645]		50	55 60
[0646]	Leu Asn Pro Asp Asp Val Cys Ser Tyr Leu Ala Ala Ala Met Gln Phe		
[0647]	65	70	75 80
[0648]	Phe Glu Gly Thr Cys Pro Glu Asp Trp Thr Ser Tyr Gly Ile Val Ile		
[0649]		85	90 95
[0650]	Ala Arg Lys Gly Asp Lys Ile Thr Pro Gly Ser Leu Val Glu Ile Lys		
[0651]		100	105 110
[0652]	Arg Thr Asp Val Glu Gly Asn Trp Ala Leu Thr Gly Gly Met Glu Leu		
[0653]		115	120 125
[0654]	Thr Arg Asp Pro Thr Val Pro Glu His Ala Ser Leu Val Gly Leu Leu		
[0655]		130	135 140
[0656]	Leu Ser Leu Tyr Arg Leu Ser Lys Ile Ser Gly Gln Asn Thr Gly Asn		
[0657]	145	150	155 160
[0658]	Tyr Lys Thr Asn Ile Ala Asp Arg Ile Glu Gln Ile Phe Glu Thr Ala		
[0659]		165	170 175
[0660]	Pro Phe Val Lys Ile Val Glu His His Thr Leu Met Thr Thr His Lys		
[0661]		180	185 190
[0662]	Met Cys Ala Asn Trp Ser Thr Ile Pro Asn Phe Arg Phe Leu Ala Gly		
[0663]		195	200 205
[0664]	Thr Tyr Asp Met Phe Phe Ser Arg Ile Glu His Leu Tyr Ser Ala Ile		
[0665]		210	215 220
[0666]	Arg Val Gly Thr Val Val Thr Ala Tyr Glu Asp Cys Ser Gly Leu Val		
[0667]	225	230	235 240
[0668]	Ser Phe Thr Gly Phe Ile Lys Gln Ile Asn Leu Thr Ala Arg Glu Ala		
[0669]		245	250 255
[0670]	Ile Leu Tyr Phe Phe His Lys Asn Phe Glu Glu Glu Ile Arg Arg Met		
[0671]		260	265 270

[0672]	Phe Glu Pro Gly Gln Glu Thr Ala Val Pro His Ser Tyr Phe Ile His		
[0673]	275	280	285
[0674]	Phe Arg Ser Leu Gly Leu Ser Gly Lys Ser Pro Tyr Ser Ser Asn Ala		
[0675]	290	295	300
[0676]	Val Gly His Val Phe Asn Leu Ile His Phe Val Gly Cys Tyr Met Gly		
[0677]	305	310	315 320
[0678]	Gln Val Arg Ser Leu Asn Ala Thr Val Ile Ala Ala Cys Ala Pro His		
[0679]	325	330	335
[0680]	Glu Met Ser Val Leu Gly Gly Tyr Leu Gly Glu Glu Phe Phe Gly Lys		
[0681]	340	345	350
[0682]	Gly Thr Phe Glu Arg Arg Phe Phe Arg Asp Glu Lys Glu Leu Gln Glu		
[0683]	355	360	365
[0684]	Tyr Glu Ala Ala Glu Leu Thr Lys Thr Asp Val Ala Leu Ala Asp Asp		
[0685]	370	375	380
[0686]	Gly Thr Val Asn Ser Asp Asp Glu Asp Tyr Phe Ser Gly Glu Thr Arg		
[0687]	385	390	395 400
[0688]	Ser Pro Glu Ala Val Tyr Thr Arg Ile Met Met Asn Gly Gly Arg Leu		
[0689]	405	410	415
[0690]	Lys Arg Ser His Ile Arg Arg Tyr Val Ser Val Ser Ser Asn His Gln		
[0691]	420	425	430
[0692]	Ala Arg Pro Asn Ser Phe Ala Glu Phe Leu Asn Lys Thr Tyr Ser Ser		
[0693]	435	440	445
[0694]	Asp Ser Gly Ser Gly Glu Asn Leu Tyr Phe Gln Ser Gly Ser Gly Ser		
[0695]	450	455	460
[0696]	His Gly Phe Pro Pro Glu Val Glu Glu Gln Asp Asp Gly Thr Leu Pro		
[0697]	465	470	475 480
[0698]	Met Ser Cys Ala Gln Glu Ser Gly Met Asp Arg His Pro Ala Ala Cys		
[0699]	485	490	495
[0700]	Ala Ser Ala Arg Ile Asn Val		
[0701]	500		
[0702]	<210> 12		
[0703]	<211> 19		
[0704]	<212> PRT		
[0705]	<213> 人工序列		
[0706]	<220>		
[0707]	<223> 合成降解决定子		
[0708]	<400> 12		
[0709]	Thr Arg Gly Val Glu Glu Val Ala Glu Gly Val Val Leu Leu Arg Arg		
[0710]	1	5	10 15
[0711]	Arg Gly Asn		
[0712]	<210> 13		
[0713]	<211> 7		

[0714]	<212>	PRT
[0715]	<213>	人工序列
[0716]	<220>	
[0717]	<223>	TEVp的共有切割位点
[0718]	<400>	13
[0719]	Glu Asn Leu Tyr Phe Gln Gly	
[0720]	1	5
[0721]	<210>	14
[0722]	<211>	7
[0723]	<212>	PRT
[0724]	<213>	人工序列
[0725]	<220>	
[0726]	<223>	TEVp的共有切割位点
[0727]	<400>	14
[0728]	Glu Asn Leu Tyr Phe Gln Ser	
[0729]	1	5
[0730]	<210>	15
[0731]	<211>	8
[0732]	<212>	PRT
[0733]	<213>	人工序列
[0734]	<220>	
[0735]	<223>	HRV 3C蛋白酶的共有切割位点
[0736]	<400>	15
[0737]	Leu Glu Val Leu Phe Gln Gly Pro	
[0738]	1	5
[0739]	<210>	16
[0740]	<211>	4
[0741]	<212>	PRT
[0742]	<213>	人工序列
[0743]	<220>	
[0744]	<223>	因子Xa的共有切割位点
[0745]	<400>	16
[0746]	Ile Glu Gly Arg	
[0747]	1	
[0748]	<210>	17
[0749]	<211>	5
[0750]	<212>	PRT
[0751]	<213>	人工序列
[0752]	<220>	
[0753]	<223>	肠激酶的共有切割位点
[0754]	<400>	17
[0755]	Asp Asp Asp Asp Lys	

[0756]	1	5
[0757]	<210>	18
[0758]	<211>	6
[0759]	<212>	PRT
[0760]	<213>	人工序列
[0761]	<220>	
[0762]	<223>	凝血酶的共有切割位点
[0763]	<400>	18
[0764]	Leu Val Pro Arg Gly Ser	
[0765]	1	5
[0766]	<210>	19
[0767]	<211>	17680
[0768]	<212>	DNA
[0769]	<213>	人工序列
[0770]	<220>	
[0771]	<223>	SiR CRE-mCherryPEST
[0772]	<400>	19
[0773]	gacattgatt attgactagt tattaatagt aatcaattac ggggtcatta gttcatagcc 60	
[0774]	catatatgga gttccgcgtt acataactta cggtaaattgg cccgcctggc tgaccgcccc 120	
[0775]	acgacccccg cccattgacg tcaataatga cgtatgttcc catagtaacg ccaataggga 180	
[0776]	ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc 240	
[0777]	aagtgtatca tatgccaagt acgcccccta ttgacgtcaa tgacggtaaa tggcccgccct 300	
[0778]	ggcattatgc ccagtacatg accttatggg actttcctac ttggcagtac atctacgtat 360	
[0779]	tagtcatgcg tattaccatg gtgatgcggt tttggcagta catcaatggg cgtggatagc 420	
[0780]	ggtttgactc acggggattt ccaagtctcc accccattga cgtcaatggg agtttgtttt 480	
[0781]	ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa ctccgcccc ttagcgcaaa 540	
[0782]	tgggcggtag gcgtgtacgg tgggaggtct atataagcag agctctctgg ctaactagag 600	
[0783]	aaccactgac ttactggctt atcgaaatta atacgactca ctataggag acccaagctg 660	
[0784]	gctagattaa gcgtctgatg agtccgtgag gacgaaaccc ggagtcccg gtcacgctta 720	
[0785]	acaaccagat caaagaaaaa acagacattg tcaattgcaa agcaaaaatg taacaccctt 780	
[0786]	acaatggatg ccgacaagat tgtattcaaa gtcaataatc aggtggtctc tttgaagcct 840	
[0787]	gagattatcg tggatcaata tgagtacaag taccctgcca tcaaagattt gaaaaagccc 900	
[0788]	tgtataaccc taggaaaggc tcccgaattt aataaagcat acaagtcagt tttgtcaggc 960	
[0789]	atgagcgccg ccaaacttaa tcctgacgat gtatgttcct atttggcagc ggcaatgcag 1020	
[0790]	ttttttgagg ggacatgtcc ggaagactgg accagctatg gaattgtgat tgcacgaaaa 1080	
[0791]	ggagataaga tcaccccagg ttctctggtg gagataaaac gtactgatgt agaagggaat 1140	
[0792]	tgggctctga caggaggcat ggaactgaca agagaccccc ctgtccctga gcatgcgtcc 1200	
[0793]	ttagtcggtc ttctcttgag tctgtatagg ttgagcaaaa tatccgggca aaacactggt 1260	
[0794]	aactataaga caaacattgc agacaggata gagcagattt ttgagacagc cccttttgtt 1320	
[0795]	aaaatcgtgg aacaccatac tctaatagaca actcacaaaa tgtgtgctaa ttggagtact 1380	
[0796]	ataccaaact tcagattttt ggccggaacc tatgacatgt ttttctccc gattgagcat 1440	
[0797]	ctatatcag caatcagagt gggcacagtt gtcactgctt atgaagactg ttcaggactg 1500	

[0798]	gtatcattta ctgggttcat aaaacaaatc aatctcaccg ctagagaggc aatactatat	1560
[0799]	ttcttccaca agaactttga ggaagagata agaagaatgt ttgagccagg gcaggagaca	1620
[0800]	gctgttcttc actcttattt catccacttc cgttcactag gcttgagtgg gaaatctcct	1680
[0801]	tattcatcaa atgctgttgg tcacgtgttc aatctcattc actttgtagg atgctatatg	1740
[0802]	ggccaagtca gatccctaaa tgcaacggtt attgctgcat gtgctcctca tgaaatgtct	1800
[0803]	gttctagggg gctatctggg agaggaattc ttcgggaaag ggacatttga aagaagattc	1860
[0804]	ttcagagatg agaaagaact tcaagaatac gaggcggctg aactgacaaa gactgacgta	1920
[0805]	gcactggcag atgatggaac tgtcaactct gacgacgagg actacttttc aggtgaaacc	1980
[0806]	agaagtccgg aggctgttta tactcgaatc atgatgaatg gaggtcgact aaagagatct	2040
[0807]	cacatacggg gatatgtctc agtcagttcc aatcatcaag cccgtccaaa ctcatctgcc	2100
[0808]	gagtttctaa acaagacata ttgagtgac tcaggttccg gagagaacct ctacttccaa	2160
[0809]	tcgggatccg gtagccatgg cttcccgcg gaggtggagg agcaggatga tggcacgctg	2220
[0810]	cccatgtctt gtgcccagga gacgggatg gaccgtcacc ctgcagcctg tgcttctgct	2280
[0811]	aggatcaatg tgtaagaagt tgaataacaa aatgccggaa atctacggat tgtgtatatc	2340
[0812]	catcatgaaa aaaactaaca cccctccttt cgaaccatcc caaacatgag caagatcttt	2400
[0813]	gtcaatccta gtgctattag agccggtctg gccgatcttg agatggctga agaaactgtt	2460
[0814]	gatctgatca atagaaatat cgaagacaat caggctcacc tccaaggga acccatagag	2520
[0815]	gtggacaatc tccctgagga tatggggcga cttcacctgg atgatggaaa atcgcccaac	2580
[0816]	catggtgaga tagccaaggt gggagaaggc aagtatcgag aggactttca gatggatgaa	2640
[0817]	ggagaggatc ctagcttcct gtccagtc aacctggaaa atgttggagt ccaaatagtc	2700
[0818]	agacaaatga ggtcaggaga gagatttctc aagatatggt cacagaccgt agaagagatt	2760
[0819]	atatcctatg tcgcgtcaa ctttcccaac cctccaggaa agtcttcaga ggataaatca	2820
[0820]	accagacta ctggccgaga gctcaagaag gagacaacac ccactcctc tcagagagaa	2880
[0821]	agccaatcat cgaaagccag gatggcggct caaattgctt ctggccctcc agcccttgaa	2940
[0822]	tggtcggeta ccaatgaaga ggatgatcta tcagtggagg ctgagatcgc tcaccagatt	3000
[0823]	gcagaaagtt tctccaaaaa atataagttt ccctctcgat cctcaggat actcttgtat	3060
[0824]	aattttgagc aattgaaaat gaaccttgat gatatagtta aagaggcaaa aaatgtacca	3120
[0825]	ggtgtgaccc gtttagccca tgacgggtcc aaactcccc taagatgtgt actgggatgg	3180
[0826]	gtcgttttgg ccaactctaa gaaattccag ttgttagtcg aatccgacaa gctgagtaaa	3240
[0827]	atcatgcaag atgacttgaa tcgctataca tcttgctaac cgaacctctc ccctcagtc	3300
[0828]	ctctagacaa taaaatccga gatgtcccaa agtcaacatg aaaaaaacag gcaacaccac	3360
[0829]	tgataaatcg atgaacctcc tacgtaagat agtgaaaaac cgcagggacg aggacactca	3420
[0830]	aaaatcctct cccgcgtcag cccctctgga tgacgatgac ttgttgcttc cacccttga	3480
[0831]	atacgtcccg ctgaaagaac ttacaggcaa gaagaacatg aggaactttt gtatcaacgg	3540
[0832]	aagggttaaa gtgtgtagcc cgaatggtta ctcgttcagg atcctgcggc acattctgaa	3600
[0833]	atcatctgac gagatatatt ctgggaatca taggatgac gggttagtca aagtggttat	3660
[0834]	tggactggct ttgtcaggat ctccagtcct tgagggcctg aactgggtat acaaatgag	3720
[0835]	gagaaccttt atcttccagt gggctgattc caggggccct cttgaagggg aggagtggga	3780
[0836]	atactctcag gagatcactt gggatgatga tactgagttc gtcggattgc aaataagagt	3840
[0837]	gattgcaaaa cagtgtcata tccagggcag agtctggtgt atcaacatga acccgagagc	3900
[0838]	atgtcaacta tggctctgaca tgtctcttca gacacaaagg tccgaagagg acaagattc	3960
[0839]	ctctctgctt ctagaataac gtacgtcaga ttatatcccc caaatttatc acttgtttac	4020

[0840]	ctctggagga gagaacatat gggctcaact ccaacccttg ggagcaatat aacaaaaaac	4080
[0841]	atgttatggt gccattaaac cgctgcattt catcaaagtc aagttgatta cttttacatt	4140
[0842]	ttgatcctct tggatgtgaa aaaaactatt aacatccctc aaaggacctg caggtacgcg	4200
[0843]	gccgcggtac cgccaccatg gtgccaaga agaagaggaa agtctccaac ctgctgactg	4260
[0844]	tgcacaaaaa cctgcctgcc ctccctgtgg atgccacctc tgatgaagtc aggaagaacc	4320
[0845]	tgatggacat gttcaggac aggcaggcct tctctgaaca cacctggaag atgctcctgt	4380
[0846]	ctgtgtgcag atcctgggct gcctggtgca agctgaacaa caggaaatgg ttccctgctg	4440
[0847]	aacctgagga tgtgaggac tacctcctgt acctgcaagc cagaggcctg gctgtgaaga	4500
[0848]	ccatccaaca gcacctgggc cagctcaaca tgctgcacag gagatctggc ctgcctcgcc	4560
[0849]	cttctgactc caatgctgtg tccctggtga tgaggagaat cagaaaggag aatgtggatg	4620
[0850]	ctggggagag agccaagcag gccctggcct ttgaacgcac tgactttgac caagtcagat	4680
[0851]	ccctgatgga gaactctgac agatgccagg acatcaggaa cctggccttc ctgggcattg	4740
[0852]	cctacaacac cctgctgcgc attgccgaaa ttgccagaat cagagtgaag gacatctccc	4800
[0853]	gcaccgatgg tgggagaatg ctgatccaca ttggcaggac caagaccctg gtgtccacag	4860
[0854]	ctggtgtgga gaaggccctg tccctggggg ttaccaagct ggtggagaga tggatctctg	4920
[0855]	tgtctggtgt ggctgatgac cccaacaact acctgttctg ccgggtcaga aagaatggtg	4980
[0856]	tggctgcccc ttctgccacc tcccaactgt ccacccgggc cctggaaggg atctttgagg	5040
[0857]	ccaccaccg cctgatctat ggtgccagg atgactctgg gcagagatac ctggcctggt	5100
[0858]	ctggccactc tgccagagtg ggtgctgcca gggacatggc cagggtggt gtgtccatcc	5160
[0859]	ctgaaatcat gcaggtggt ggctggacca atgtgaacat tgtgatgaac tacatcagaa	5220
[0860]	acctggactc tgagactggg gccatggtga ggctgctcga ggatggggac ggcagtggag	5280
[0861]	gatccggagc cacgaacttc tctctgtaa agcaagcagg agacgtggaa gaaaaccccg	5340
[0862]	gtcctaccgg tgtgagcaag ggcgaggagg ataacatggc catcatcaag gagttcatgc	5400
[0863]	gcttcaaggt gcacatggag ggctccgtga acggccacga gttcgagatc gagggcgagg	5460
[0864]	gcgagggcg cccctacgag ggcaccaga ccgccaaact gaaggtgacc aagggtggcc	5520
[0865]	ccctgccctt cgctgggac atcctgtccc ctacgttcat gtacggctcc aaggcctacg	5580
[0866]	tgaagcacc cgccgacatc cccgactact tgaagctgtc cttccccgag ggcttcaagt	5640
[0867]	gggagcgct gatgaacttc gaggacggcg gcgtggtgac cgtgaccag gactcctccc	5700
[0868]	tgcaggacgg cgagttcatc tacaaggtga agctgcgcgg caccaacttc cctccgacg	5760
[0869]	gccccgtaat gcagaagaag accatgggct gggaggcctc ctccgagcgg atgtaccccg	5820
[0870]	aggacggcgc cctgaagggc gagatcaagc agaggctgaa gctgaaggac ggcggccact	5880
[0871]	acgacgctga ggtcaagacc acctacaagg ccaagaagcc cgtgcagctg cccggcgct	5940
[0872]	acaacgtcaa catcaagttg gacatcacct ccacaaacga ggactacacc atcgtggaac	6000
[0873]	agtacgaac cgccgagggc cgccactcca ccggcggcat ggacgagctg tacaagggat	6060
[0874]	atctcagcca tggcttccc cgagggtgg aggagcagga tgatggcacg ctgcccattg	6120
[0875]	cttgtccca ggagagcggg atggaccgtc accctgcagc ctgtgcttct gctaggatca	6180
[0876]	atgtgtgact cgaggcgcg cctaccgcg gtagcttttc agtcgagaaa aaaacattag	6240
[0877]	atcagaagaa caactggcaa cactttctca acctgagact tacttcaaga tgctcgatcc	6300
[0878]	tggagaggtc tatgatgacc ctattgacct aatcgagtta gaggctgaac ccagaggaac	6360
[0879]	ccccattgtc cccaacatct tgaggaaactc tgactacaat ctcaactctc ctttgataga	6420
[0880]	agatcctgct agactaatgt tagaatggtt aaaaacaggg aatagacctt atcggatgac	6480
[0881]	tctaacagac aattgtctca ggtctttcag agttttgaaa gattatttca agaaggtaga	6540

[0882]	tttgggttct ctcaaggtgg gcggaatggc tgcacagtca atgattttctc tctggttata	6600
[0883]	tggtgcccac tctgaatcca acaggagccg gagatgtata acagacttgg cccattttcta	6660
[0884]	ttccaagtcg tccccatag agaagctgtt gaatctcacg ctaggaaata gagggctgag	6720
[0885]	aatcccccca gagggagtgt taagttgcct tgagagggtt gattatgata atgcatttgg	6780
[0886]	aaggtatctt gccaacacgt attcctctta cttgttcttc catgtaatca cttatacat	6840
[0887]	gaacgcccta gactgggatg aagaaaagac catcctagca ttatggaaag atttaacctc	6900
[0888]	agtggacatc gggaaggact tggtaaagtt caaagaccaa atatggggac tgctgatcgt	6960
[0889]	gacaaaggac tttgtttact cccaaagttc caattgtctt ttgacagaa actacacact	7020
[0890]	tatgctaaaa gatcttttct tgtctcgctt caactcctta atggctcttg tctctcccc	7080
[0891]	agagccccga tactcagatg acttgatata tcaactatgc cagctgtaca ttgctgggga	7140
[0892]	tcaagtcttg tctatgtgtg gaaactccg ctagaagtc atcaaaatat tggagccata	7200
[0893]	tgctgtgaat agtttagtcc agagagcaga aaagtttagg cctctcattc attccttggg	7260
[0894]	agactttcct gtatttataa aagacaaggt aagtcaactt gaagagacgt tcggctccctg	7320
[0895]	tgcaagaagg ttctttaggg ctctggatca attcgacaac atacatgact tggtttttgt	7380
[0896]	gtttggctgt tacaggcatt gggggcaccc atatatagat tatcgaaagg gtctgtcaaa	7440
[0897]	actatatgat caggttcacc ttaaaaaaat gatagataag tcctaccagg agtgcttagc	7500
[0898]	aagcgacctc gccaggagga tccttagatg gggttttgat aagtactcca agtggtatct	7560
[0899]	ggattcaaga ttctagccc gagaccaccc cttgactcct tatatcaaaa cccaaacatg	7620
[0900]	gccacccaaa catattgtag acttggtggg ggatacatgg cacaagctcc cgatcacgca	7680
[0901]	gatctttgag attcctgaat caatggatcc gtcagaaaata ttggatgaca aatcacattc	7740
[0902]	tttcaccaga acgagactag cttcttggct gtcagaaaac cgaggggggc ctgttcttag	7800
[0903]	cgaaaaagtt attatcacgg ccctgtctaa gccgcctgtc aatccccgag agtttctgag	7860
[0904]	gtctatagac ctcgaggatg tgccagatga agacttgata attggcctca agccaaagga	7920
[0905]	acgggaattg aagattgaag gtcgattctt tgctctaatt tcatggaatc taagattgta	7980
[0906]	ttttgtcatc actgaaaaac tcttggccaa ctacatcttg ccactttttg acgcgctgac	8040
[0907]	tatgacagac aacctgaaca aggtgtttta aaagctgacg gacagggtca ccgggcaagg	8100
[0908]	gcttttggac tattcaaggg tcacatatgc atttcacctg gactatgaaa agtggacaaa	8160
[0909]	ccatcaaaga ttagagtcaa cagaggatgt attttctgtc ctagatcaag tgtttggatt	8220
[0910]	gaagagagtg ttttctagaa cacacgagtt ttttcaaaag gcctggatct attattcaga	8280
[0911]	cagatcagac ctcatcgggt tacgggagga tcaaatatac tgcttagatg cgtccaacgg	8340
[0912]	cccaacctgt tggaatggcc aggatggcgg gctagaaggc ttacggcaga agggctggag	8400
[0913]	tctagtcagc ttattgatga tagatagaga atctcaaac aggaacacaa gaacaaaaat	8460
[0914]	actagctcaa ggagacaacc aggttttatg tccgacatac atgttgtcgc cagggtctatc	8520
[0915]	tcaagagggg ctctctatg aattggagag aatatcaagg aatgcacttt cgatatacag	8580
[0916]	agccgtcgag gaaggggcat ctaagctagg gctgatcatc aagaaagaag agaccatgtg	8640
[0917]	tagttatgac ttctcatct atggaaaaac ccctttgttt agaggtaaca tatttggtgc	8700
[0918]	tgagtccaaa agatgggcca gagtctcttg cgtctctaag gaccaaatag tcaacctcgc	8760
[0919]	caatataatg tcgacagtgt ccaccaatgc gctaacagtg gcacaacact ctcaatcttt	8820
[0920]	gatcaaaccg atgagggatt ttctgtcat gtcagtacag gcagtcttct actacctgct	8880
[0921]	atttagccca atcttaaagg gaagagtta caagattctg agcgctgaag gggagagctt	8940
[0922]	tctcctagcc atgtcaagga taatctatct agatccttct ttgggaggga tatctggaat	9000
[0923]	gtccctcgga agattccata tacgacagtt ctgaccct gtctctgaag ggttatcctt	9060

[0924]	ctggagagag atctgggttaa gctcccaaga gtcctggatt cacgcgttgt gtcaagaggc	9120
[0925]	tggaaaccca gatcttgag agagaacact cgagagcttc actgccttc tagaagatcc	9180
[0926]	gaccacctta aatatcagag gaggggccag tectaccatt ctactcaagg atgcaatcag	9240
[0927]	aaaggcttta tatgacgagg tggacaaggt ggaaaattca gagtttcgag aggcaatcct	9300
[0928]	gttgccaag acccatagag ataattttat actcttctta atatctgttg agcctctgtt	9360
[0929]	tcctcgattt ctcagtgagc tattcagttc gtcttttttg ggaatccccg agtcaatcat	9420
[0930]	tggattgata caaaactccc gaacgataag aaggcagttt agaaagagtc tctcaaaaac	9480
[0931]	tttagaagaa tccttctaca actcagagat ccacgggatt agtcggatga cccagacacc	9540
[0932]	tcagagggtt gggggggtgt ggccttgctc ttcagagagg gcagatctac ttagggagat	9600
[0933]	ctcttgggga agaaaagtgg taggcacgac agttcctcac ctttctgaga tgttgggatt	9660
[0934]	acttcccaag tcctctatct cttgcacttg tggagcaaca ggaggaggca atcctagagt	9720
[0935]	ttctgtatca gtactcccgt cttttgatca gtcatttttt tcacgaggcc ccctaaagg	9780
[0936]	atacttgggc tcgtccacct ctatgtcgac ccagctattc catgcatggg aaaaagtcac	9840
[0937]	taatgttcat gtggtgaaga gagctctatc gttaaaagaa tctataaact ggttcattac	9900
[0938]	tagagattcc aacttggtc aagctctaatt taggaacatt atgtctctga caggccctga	9960
[0939]	tttccctcta gaggaggccc ctgtcttcaa aaggacgggg tcagccttgc ataggttcaa	10020
[0940]	gtctgccaga tacagcgaag gaggtattc ttctgtctgc ccgaacctcc tctctcatat	10080
[0941]	ttctgttagt acagacacca tgtctgattt gacccaagac gggaagaact acgatttcat	10140
[0942]	gttccagcca ttgatgcttt atgcacagac atggacatca gagctggtac agagagacac	10200
[0943]	aaggetaaga gactctacgt ttcatggca cctccgatgc aacagggtgtg tgagacccat	10260
[0944]	tgacgacgtg accctggaga cctctcagat cttcgagttt ccggatgtgt cgaaaagaat	10320
[0945]	atccagaatg gtttctgggg ctgtgcctca cttccagagg cttcccgata tccgtctgag	10380
[0946]	accaggagat tttgaatctc taagcggtag agaaaagtct caccatatcg gatcagctca	10440
[0947]	ggggtcttta tactcaatct tagtggcaat tcacgactca ggatacaatg atggaacct	10500
[0948]	cttccctgtc aacatatacg gcaaggtttc ccttagagac tatttgagag ggctcgcaag	10560
[0949]	gggagtattg ataggatcct cgatttgctt cttgacaaga atgacaaata tcaatattaa	10620
[0950]	tagacctctt gaattggtct caggggtaat ctcatatatt ctctgaggc tagataacca	10680
[0951]	tcctccttg tacataatgc tcagagaacc gtctcttaga ggagagatat tttctatccc	10740
[0952]	tcagaaaatc cccgccgtt atccaaccac tatgaaagaa ggcaacagat caatcttgtg	10800
[0953]	ttatctccaa catgtgtac gctatgagcg agagataatc acggcgtctc cagagaatga	10860
[0954]	ctggctatgg atcttttcag actttagaag tgccaaaatg acgtacctat ccctcattac	10920
[0955]	ttaccagtct catcttctac tccagagggt tgagagaaac ctatctaaga gtatgagaga	10980
[0956]	taacctgcga caattgagtt ctttgatgag gcagggtgtg ggcgggcacg gagaagatac	11040
[0957]	cttagagtca gacgacaaca ttcaacgact gctaaaagac tctttacgaa ggacaagatg	11100
[0958]	ggtggatcaa gaggtgcgcc atgcagctag aacctgact ggagattaca gcccacaa	11160
[0959]	gaagggtgcc cgtaaggtag gatgttcaga atgggtctgc tctgctcaac aggttgacgt	11220
[0960]	ctctacctca gcaaaccgg ccctgtctc ggagcttgac ataaggggcc tctctaagag	11280
[0961]	gttccagaac cttttgatct cgggcttgag agtggttcag tgggcaaccg gtgctcatta	11340
[0962]	taagcttaag cctattctag atgatctcaa tgttttccca tctctctgcc ttgtagttag	11400
[0963]	ggacgggtca ggggggatat caagggcagt cctcaacatg tttccagatg ccaagcttgt	11460
[0964]	gttcaacagt cttttagagg tgaatgacct gatggcttcc ggaacacatc cactgcctcc	11520
[0965]	ttcagcaatc atgaggggag gaaatgatat cgtctccaga gtgatagatc ttgactcaat	11580

[0966]	ctgggaaaaa ccgtccgact tgagaaactt ggcaacctgg aaatacttcc agtcagtcca	11640
[0967]	aaagcaggtc aacatgtcct atgacctcat tatttgcgat gcagaagtta ctgacattgc	11700
[0968]	atctatcaac cggatcaccc tgttaatgtc cgatttttga ttgtctatag atggaccact	11760
[0969]	ctatttggtc ttcaaaactt atgggactat gctagtaaat ccaaactaca aggctattca	11820
[0970]	acacctgtca agagcgttcc cctcggtcac aggggtttatc acccaagtaa cttcgtcttt	11880
[0971]	ttcatctgag ctctacctcc gattctccaa acgagggaag tttttcagag atgctgagta	11940
[0972]	cttgacctct tccacccttc gagaaatgag ctttgtgtta ttcaattgta gcagccccaa	12000
[0973]	gagtgaatg cagagagctc gttccttgaa ctatcaggat cttgtgagag gatttctga	12060
[0974]	agaaatcata tcaaatcctt acaatgagat gatcataact ctgattgaca gtgatgtaga	12120
[0975]	atcttttcta gtccacaaga tggttgatga tcttgagtta cagaggggaa ctctgtctaa	12180
[0976]	agtggctatc attatagcca tcatgatagt tttctccaac agagtcttca acgtttccaa	12240
[0977]	accctaact gaccctcgt tctatccacc gtctgatccc aaaatcctga ggcacttcaa	12300
[0978]	catatgttgc agtactatga tgtatctatc tactgttcta ggtgacgtcc ctagcttcgc	12360
[0979]	aagacttcac gacctgtata acagacctat aacttattac ttcagaaagc aagtcattcg	12420
[0980]	agggaaactt tatctatctt ggagttggtc caacgacacc tcagtgttca aaagggtagc	12480
[0981]	ctgtaattct agcctgagtc tgtcatctca ctggatcagg ttgatttaca agatagttaa	12540
[0982]	gactaccaga ctctgttgca gcatcaagga tctatccaga gaagtggaaa gacaccttca	12600
[0983]	taggtacaac aggtggatca ccctagagga tatcagatct agatcatccc tactagacta	12660
[0984]	cagttgcctg tgaaccggat actcctggaa gcctgcccac gctaagactc ttgtgtgatg	12720
[0985]	tatcttgaag aaaacaagat caccggatac tcttgggaag ctgcccacgc taagactctt	12780
[0986]	gtgtgatgta tcttgaaaaa aacaagatcc taaatctgaa cttttggttg ttgtattgtt	12840
[0987]	tttctcattt ttgttgttta ttgtttaagc gtgggtcggc atggcatctc cacctcctcg	12900
[0988]	cgggtccgacc tgggcacccg aaggaggacg cacgtccact cggatggcta agggagaagg	12960
[0989]	gcgaattcca gcacactggc ggccgttact agtggatccg agctcggtag caagcttaag	13020
[0990]	tttaaacgcg tgatcagcct cgactgtgcc ttctagtgtc cagccatctg ttgtttgccc	13080
[0991]	ctccccgtg ctttcttga ccctggaagg tgccactccc actgtccttt cctaataaaa	13140
[0992]	tgaggaaatt gcatcgcat gtctgagtag gtgtcattct attctggggg gtgggggtggg	13200
[0993]	gcaggacagc aagggggagg attgggaaga caatagcagg catgctgggg atgcgggtggg	13260
[0994]	ctctatggct tctgaggcgg aaagaaccag ctggggctct aggggggtatc cccacgcgc	13320
[0995]	ctgtagcggc gcattaagcg cggcgggtgt ggtggttacg cgcagcgtga ccgctacact	13380
[0996]	tgccagcgcc ctagcgccc ctcctttcgc tttcttcctc tcctttctcg ccacgttcgc	13440
[0997]	cggctttccc cgtcaagctc taaatcgggg gctcccttta gggttccgat ttagtgcttt	13500
[0998]	acggcacctc gaccccaaaa aacttgatta gggatgatgt tcacgtagt ggccatcgcc	13560
[0999]	ctgatagacg gtttttcgcc ctttgacgtt ggagtccacg ttctttaata gtggactctt	13620
[1000]	gttccaaact ggaacaacac tcaaccctat ctcggtctat tcttttgatt tataagggat	13680
[1001]	tttgccgatt tcggcctatt ggttaaaaaa tgagctgatt taacaaaaat ttaacgcga	13740
[1002]	ttaattctgt ggaatgtgtg tcagttaggg tgtggaaagt cccaggctc cccagcaggc	13800
[1003]	agaagtatgc aaagcatgca tctcaattag tcagcaacca ggtgtggaaa gtccccaggc	13860
[1004]	tccccagcag gcagaagtat gcaagcatg catctcaatt agtcagcaac catagtcccc	13920
[1005]	cccctaactc cgcccatccc gccctaact ccgcccagtt ccgcccattc tccgcccatt	13980
[1006]	ggctgactaa ttttttttat ttatgcagag gccgaggccg cctctgcctc tgagctattc	14040
[1007]	cagaagtagt gaggaggctt ttttgaggc ctaggctttt gcaaaaagct cccgggagct	14100

[1008]	tgtatatcca	ttttcggatc	tgatcaagag	acaggatgag	gatacgtttcg	catgattgaa	14160
[1009]	caagatggat	tgcacgcagg	ttctccggcc	gcttgggtgg	agaggctatt	cggctatgac	14220
[1010]	tgggcacaac	agacaatcgg	ctgctctgat	gccgccgtgt	tccggctgtc	agcgcagggg	14280
[1011]	cgcgcggttc	tttttgtcaa	gaccgacctg	tccggtgccc	tgaatgaact	gcaggacgag	14340
[1012]	gcagcgcggc	tatcgtggct	ggccacgacg	ggcgttcctt	gcgcagctgt	gctcgacgtt	14400
[1013]	gtcactgaag	cgggaaggga	ctggctgcta	ttgggcgaag	tgccggggca	ggatctcctg	14460
[1014]	tcactcacc	ttgctcctgc	cgagaaagta	tccatcatgg	ctgatgcaat	gcggcggtg	14520
[1015]	catacgttg	atccggtac	ctgcccattc	gaccaccaag	cgaacatcg	catcgagcga	14580
[1016]	gcacgtactc	ggatggaagc	cggctctgtc	gatcaggatg	atctggacga	agagcatcag	14640
[1017]	gggctcgcgc	cagccgaact	gttcgccagg	ctcaaggcgc	gcatgcccga	cggcgaggat	14700
[1018]	ctcgtcgtga	cccatggcga	tgctgtcttg	ccgaatatca	tggtggaaaa	tgccgcgttt	14760
[1019]	tctggattca	tcgactgtgg	cgggtgggt	gtggcggacc	gctatcagga	catagcgttg	14820
[1020]	gctaccctg	atattgctga	agagcttggc	ggcgaatggg	ctgaccgctt	cctcgtgctt	14880
[1021]	tacggtatcg	ccgctcccga	ttcgcagcgc	atcgcttctt	atcgcttctt	tgacgagttc	14940
[1022]	ttctgagcgg	gactctgggg	ttcgaaatga	ccgaccaagc	gacgcccac	ctgccatcac	15000
[1023]	gagatttcca	ttccaccgcc	gccttctatg	aaaggttggg	cttcggaatc	gttttccggg	15060
[1024]	acgccggctg	gatgatcctc	cagcgcgggg	atctcatgct	ggagtctctc	gcccacccca	15120
[1025]	actgttttat	tgcagcttat	aatggttaca	aataaagcaa	tagcatcaca	aatttcacaa	15180
[1026]	ataaagcatt	tttttactg	cattctagtt	gtggtttgtc	caaactcatc	aatgtatctt	15240
[1027]	atcatgtctg	tataccgtcg	acctctagct	agagcttggc	gtaatcatgg	tcatagctgt	15300
[1028]	ttctgtgtg	aaattgttat	ccgtcacaca	ttccacacaa	catacgagcc	ggaagcataa	15360
[1029]	agtgtaaagc	ctgggtgcc	taatgagtga	gctaactcac	attaattgcg	ttgcgctcac	15420
[1030]	tgcccgcctt	ccagtcggga	aacctgtcgt	gccagctgca	ttaatgaatc	ggccaacgcg	15480
[1031]	cggggagagg	cggtttgcgt	attgggcgct	cttccgcttc	ctcgtcact	gactcgtgc	15540
[1032]	gctcggtcgt	tcggctcgg	cgagcggtat	cagctcactc	aaaggcggt	atacggttat	15600
[1033]	ccacagaatc	aggggataac	gcaggaaaga	acatgtgagc	aaaaggccag	caaaaggcca	15660
[1034]	ggaaccgtaa	aaaggccgcg	ttgctggcgt	ttttccatag	gctccgccc	cctgacgagc	15720
[1035]	atcacaaaaa	tcgacgtca	agtcagaggt	ggcgaaaccc	gacaggacta	taaagatacc	15780
[1036]	aggcgtttcc	ccctggaagc	tcctcgtgc	gctctcctgt	tccgacctg	ccgcttaccg	15840
[1037]	gatacctgtc	cgcctttctc	ccttcgggaa	gcgtggcgt	ttctcatagc	tcacgtgtga	15900
[1038]	ggtatctcag	ttcgggtgag	gtcgttcgct	ccaagctggg	ctgtgtgcac	gaaccccccg	15960
[1039]	ttcagcccga	ccgtgcgcc	ttatccggt	actatcgtct	tgagtccaac	ccggtgaagc	16020
[1040]	acgacttata	gccactggca	gcagccactg	gtaacaggat	tagcagagcg	aggtatgtag	16080
[1041]	gcggtgctac	agagttcttg	aagtgggtgc	ctaactacgg	ctacactaga	agaacagtat	16140
[1042]	ttggtatctg	cgctctgctg	aagccagtta	ccttcggaaa	aagagttggt	agctcttgat	16200
[1043]	ccggcaaaaa	aaccaccgct	ggtagcgtg	gtttttttgt	ttgcaagcag	cagattacgc	16260
[1044]	gcagaaaaaa	aggatctcaa	gaagatcctt	tgatcttttc	tacgggtctt	gacgtcagt	16320
[1045]	ggaacgaaaa	ctcacgttaa	gggattttgg	tcatgagatt	atcaaaaagg	atcttcacct	16380
[1046]	agatcctttt	aaattaaaaa	tgaagtttta	aatcaatcta	aagtatatat	gagtaaactt	16440
[1047]	ggtctgacag	ttaccaatgc	ttaatcagt	aggcacctat	ctcagcgatc	tgtctatttc	16500
[1048]	gttcatccat	agttgcctga	ctccccgtcg	tgtagataac	tacgatacgg	gagggccttac	16560
[1049]	catctggccc	cagtgtgca	atgataccgc	gagaccacg	ctcaccggt	ccagatttat	16620

[1050]	cagcaataaa ccagccagcc ggaagggccg agcgcagaag tggtcctgca actttatccg	16680
[1051]	cctccatcca gtctattaat tgttgccggg aagctagagt aagtagttcg ccagttaata	16740
[1052]	gtttgcgcaa cgttgttgcc attgctacag gcatcgtggt gtcacgctcg tcgttttgta	16800
[1053]	tggcttcatt cagctccggt tcccaacgat caaggcgagt tacatgatcc cccatgtttgt	16860
[1054]	gcaaaaaagc ggtagctcc ttcggctctc cgatcgttgt cagaagtaag ttggccgcag	16920
[1055]	tgttatcact catggttatg gcagcactgc ataattctct tactgtcatg ccatccgtaa	16980
[1056]	gatgcttttc tgtgactggt gagtactcaa ccaagtcatc ctgagaatag tgtatgcggc	17040
[1057]	gaccgagttg ctcttgcccg gcgtcaatac gggataatac cgcgccacat agcagaactt	17100
[1058]	taaaagtgct catcattgga aaacgttctt cggggcgaaa actctcaagg atcttaccgc	17160
[1059]	tgttgagatc cagttcgatg taaccactc gtgcacccaa ctgatcttca gcattcttta	17220
[1060]	ctttcaccag cgtttctggg tgagcaaaaa caggaaggca aaatgccgca aaaaaggga	17280
[1061]	taaggcgac acggaaatgt tgaatactca tactcttcct ttttcaatat tattgaagca	17340
[1062]	tttatcaggg ttattgtctc atgagcggat acatatttga atgtatttag aaaaataaac	17400
[1063]	aaataggggt tccgcgcaca tttcccgaa aagtgccacc tgacgtcgac ggatcgggag	17460
[1064]	atctcccgat cccctatggt gcactctcag tacaatctgc tctgatgccg catagttaag	17520
[1065]	ccagtatctg ctccctgctt gtgtgttgga ggtcgtgag tagtgcgca gcaaaattta	17580
[1066]	agctacaaca aggcaaggct tgaccgacaa ttgcatgaag aatctgctta gggtaggcg	17640
[1067]	ttttgcgctg cttcgcgatg tacgggccag atatacgct	17680
[1068]	<210>	20
[1069]	<211>	15678
[1070]	<212>	DNA
[1071]	<213>	人工序列
[1072]	<220>	
[1073]	<223>	编码SiR载体的质粒
[1074]	<400>	20
[1075]	gacggatcgg gagatctccc gatcccctat ggtgcactct cagtacaatc tgctctgatg	60
[1076]	ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg	120
[1077]	cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc	180
[1078]	ttagggtag gcgttttgcg ctgcttcgag atgtacgggc cagatatacg cgttgacatt	240
[1079]	gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata	300
[1080]	tggagtccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc	360
[1081]	cccgccatt gagtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc	420
[1082]	attgacgtca atgggtggag tattttacggt aaactgcca cttggcagta catcaagtgt	480
[1083]	atcatatgcc aagtacgcc cctattgacg tcaatgacgg taaatggccc gcctggcatt	540
[1084]	atgccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca	600
[1085]	tcgctattac catggatgat cggttttggc agtacatcaa tgggcgtgga tagcggtttg	660
[1086]	actcacggg atttcaaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc	720
[1087]	aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg	780
[1088]	gtaggcgtgt acgggtggag gtctatataa gcagagctct ctggctaact agagaacca	840
[1089]	ctgcttactg gcttatcgaa attaatacga ctactatag ggagacccaa gctggctaga	900
[1090]	ttaagcgtct gatgagtcg tgaggacgaa acccgagtc ccgggtcacg cttacaacc	960
[1091]	agatcaaaga aaaaacagac attgtcaatt gcaaagcaaa aatgtaacac ccctacaatg	1020

[1092]	gatgccgaca agattgtatt caaagtcaat aatcaggtgg tctctttgaa gcctgagatt	1080
[1093]	atcgtggatc aatatgagta caagtaccct gccatcaaag atttgaaaaa gccctgtata	1140
[1094]	accctaggaa aggtctccga tttaaataaa gcatacaagt cagttttgtc aggcatgagc	1200
[1095]	gccgccaaac ttaatcctga cgatgtatgt tcctatttgg cagcggcaat gcagtttttt	1260
[1096]	gaggggacat gtccggaaga ctggaccagc tatggaattg tgattgcacg aaaaggagat	1320
[1097]	aagatcaccc caggttctct ggtggagata aaacgtactg atgtagaagg gaattgggct	1380
[1098]	ctgacaggag gcatggaact gacaagagac cccactgtcc ctgagcatgc gtccttagtc	1440
[1099]	ggtcttctct tgagtctgta taggttgagc aaaatatccg ggcaaaacac tggtaactat	1500
[1100]	aagacaaaca ttgcagacag gatagagcag atttttgaga cagccccctt tgttaaaatc	1560
[1101]	gtggacacc atactcta at gacaactcac aaaatgtgtg ctaattggag tactatacca	1620
[1102]	aacttcagat ttttggccg aacctatgac atgtttttct cccgattga gcatctatat	1680
[1103]	tcagcaatca gagtgggac agttgtcact gcttatgaag actgttcagg actggtatca	1740
[1104]	tttactgggt tcataaaaca aatcaatctc accgctagag aggcaatact atatttcttc	1800
[1105]	cacaagaact ttgaggaaga gataagaaga atgtttgagc cagggcagga gacagctgtt	1860
[1106]	cctcactctt atttcatcca ctccgttca ctaggcttga gtgggaaatc tccttattca	1920
[1107]	tcaaatgtg ttggtcacgt gttcaatctc attcactttg taggatgcta tatgggtcaa	1980
[1108]	gtcagatccc taaatgcaac ggttattgct gcatgtgctc ctcatgaaat gtctgttcta	2040
[1109]	gggggctatc tgggagagga attcttcggg aaagggacat ttgaaagaag attcttcaga	2100
[1110]	gatgagaaag aacttcaaga atacgaggcg gctgaactga caaagactga cgtagcactg	2160
[1111]	gcagatgatg gaactgtcaa ctctgacgac gaggactact tttcaggtga aaccagaagt	2220
[1112]	ccggaggtg tttatactcg aatcatgatg aatggaggtc gactaaagag atctcacata	2280
[1113]	cggagatatg tctcagtcag ttccaatcat caagcccgct caaactcatt cgccgagttt	2340
[1114]	ctaaacaaga catattcgag tgactcaggt tccggagaga acctctactt ccaatcgga	2400
[1115]	tccggtagcc atggcttccc gccggagggt gaggagcagg atgatggcac gctgcccag	2460
[1116]	tcttgtgccc aggagagcgg gatggaccgt caccctgcag cctgtgcttc tgctaggatc	2520
[1117]	aatgtgtaag aagtgaata acaaaatgcc ggaaatctac ggattgtgta tatccatcat	2580
[1118]	gaaaaaact aacacccctc ctttcgaacc atcccaaaca tgagcaagat ctttgtcaat	2640
[1119]	cctagtgtga ttagagccgg tctggccgat cttgagatgg ctgaagaaac tgttgatctg	2700
[1120]	atcaatagaa atatgaaga caatcaggt catctccaag gggaacccat agagggtggac	2760
[1121]	aatctccctg aggatatggg gcgacttcac ctggatgatg gaaaatcgcc caaccatggt	2820
[1122]	gagatagcca aggtgggaga aggcaagtat cgagaggact ttcagatgga tgaaggagag	2880
[1123]	gatcctagct tcctgttcca gtcatactg gaaaatgttg gagtccaaat agtcagacaa	2940
[1124]	atgaggtcag gagagagatt tctcaagata tggtcacaga ccgtagaaga gattatatcc	3000
[1125]	tatgtcggg tcaactttcc caaccctcca ggaaagtctt cagaggataa atcaaccag	3060
[1126]	actactggcc gagagctcaa gaaggagaca acaccactc cttctcagag agaaagccaa	3120
[1127]	tcacgaaag ccaggatggc ggctcaaatt gcttctggcc ctccagccct tgaatggtcg	3180
[1128]	gctaccaatg aagaggatga tctatcagtg gaggtgaga tcgctcacca gattgcagaa	3240
[1129]	agtttctcca aaaaatataa gtttcctct cgatcctcag ggatactctt gtataatttt	3300
[1130]	gagcaattga aatgaacct tgatgatata gttaaagagg caaaaaatgt accaggtgtg	3360
[1131]	accggttag cccatgacgg gtccaaactc cccctaagat gtgtactggg atgggtcgct	3420
[1132]	ttggccaact ctaagaaatt ccagttgtta gtcgaatccg acaagctgag taaaatcatg	3480
[1133]	caagatgact tgaatcgcta tacatcttgc taaccgaacc tctccctca gtcctctag	3540

[1134] acaataaaat ccgagatgtc ccaaagtcaa catgaaaaaa acaggcaaca ccaactgataa 3600
[1135] atcgatgaac ctctacgta agatagtga aaaccgcagg gacgaggaca ctcaaaaatc 3660
[1136] ctctcccgcg tcagcccctc tggatgacga tgacttgtgg cttccacccc ctgaatacgt 3720
[1137] cccgctgaaa gaacttacag gcaagaagaa catgaggaac ttttgtatca acggaagggt 3780
[1138] taaagtgtgt agcccgaatg gttactcgtt caggatcctg cggcacattc tgaaatcatt 3840
[1139] cgacgagata tattctggga atcataggat gatcgggtta gtcaaagtgg ttattggact 3900
[1140] ggctttgtca ggatctccag tccctgaggg cctgaactgg gtatacaaat tgaggagaac 3960
[1141] ctttatcttc cagtgggctg attccagggg ccctcttgaa ggggaggagt tggaatactc 4020
[1142] tcaggagatc acttgggatg atgatactga gttcgtcgga ttgcaaataa gagtgattgc 4080
[1143] aaaacagtgt catatccagg gcagagtctg gtgtatcaac atgaacccga gagcatgtca 4140
[1144] actatggctc gacatgtctc ttcagacaca aagggtccga gaggacaaag attcctctct 4200
[1145] gcttctagaa taacgtacgt cagattatat cccgcaaatt tatcacttgt ttacctctgg 4260
[1146] aggagagaac atatgggctc aactccaacc cttgggagca atataacaaa aaacatgtta 4320
[1147] tggtgccatt aaaccgctgc atttcatcaa agtcaagttg attaccttta cattttgatc 4380
[1148] ctcttggatg tgaaaaaac tattaacatc cctcaaagga cctgcaggta cgcggccgct 4440
[1149] acgcccgggc tacgctagca tgaaaaaac taacaccctt ccttaattaa tacggcgcgc 4500
[1150] ctaccgcggg tagcttttca gtcgagaaaa aaacattaga tcagaagaac aactggcaac 4560
[1151] actttctcaa cctgagactt acttcaagat gctcgatcct ggagaggtct atgatgacct 4620
[1152] tattgacca atcgagttag aggctgaacc cagaggaacc ccattgttc ccaacatctt 4680
[1153] gaggaactct gactacaatc tcaactctcc tttgatagaa gatcctgcta gactaatgtt 4740
[1154] agaatggta aaaacagga atagacctta tcggatgact ctaacagaca attgctccag 4800
[1155] gtctttcaga gttttgaaag attatttcaa gaaggtagat ttgggttctc tcaaggtggg 4860
[1156] cggaatggct gcacagtcaa tgatttctct ctggttatat ggtgcccact ctgaatccaa 4920
[1157] caggagccgg agatgtataa cagacttggc ccatttctat tccaagtcgt ccccataga 4980
[1158] gaagctgttg aatctcacgc taggaaatag agggctgaga atccccccag agggagtgtt 5040
[1159] aagttgcctt gagagggttg attatgataa tgcatttga aggtatcttg ccaacacgta 5100
[1160] ttctcttac ttgttcttc atgtaatcac cttatacatg aacgccctag actgggatga 5160
[1161] agaaaagacc atcctagcat tatggaaaga tttaacctca gtggacatcg ggaaggactt 5220
[1162] ggtaaagtgc aaagaccaa tatggggact gctgatcgtg acaaaggact ttgtttactc 5280
[1163] ccaaagttcc aattgtcttt ttgacagaaa ctacacactt atgctaaaag atcttttctt 5340
[1164] gtctcgcttc aactccttaa tggcttctct ctctcccca gagccccgat actcagatga 5400
[1165] cttgatattc caactatgcc agctgtacat tgctggggat caagtcttgt ctatgtgtgg 5460
[1166] aaactccggc tatgaagtca tcaaaatatt ggagccatat gtcgtgaata gtttagtcca 5520
[1167] gagagcagaa aagtttaggc ctctcattca ttccttggga gactttcctg tatttataaa 5580
[1168] agacaaggta agtcaacttg aagagacgtt cggctccctg gcaagaaggt tctttagggc 5640
[1169] tctggatcaa ttcgacaaca tacatgactt ggttttgtg tttggctgtt acaggcattg 5700
[1170] ggggcacca tatatagatt atcgaaaggg tctgtcaaaa ctatatgatc aggttcacct 5760
[1171] taaaaaatg atagataagt cctaccagga gtgcttagca agcgacctag ccaggaggat 5820
[1172] ccttagatgg ggttttgata agtactccaa gtggtatctg gattcaagat tcctagcccc 5880
[1173] agaccacccc ttgactcctt atatcaaac ccaacatgg ccacccaaac atattgtaga 5940
[1174] cttgggtggg gatacatggc acaagctccc gatcacgcag atctttgaga ttcctgaatc 6000
[1175] aatggatccg tcagaaatat tggatgacaa atcacattct ttcaccagaa cgagactagc 6060

[1176] ttcttggctg tcagaaaacc gaggggggcc tgttcctagc gaaaaagtta ttatcacggc 6120
[1177] cctgtctaag ccgcctgtca atccccgaga gtttctgagg tctatagacc tcggaggatt 6180
[1178] gccagatgaa gacttgataa ttggcctcaa gccaaaggaa cgggaattga agattgaagg 6240
[1179] tcgattcttt gctctaattgt catggaatct aagattgtat tttgtcatca ctgaaaaact 6300
[1180] cttggccaac tacatcttgc cactttttga cgcgtgact atgacagaca acctgaacaa 6360
[1181] ggtgtttaaa aagctgatcg acagggtcac cgggcaaggc cttttggact attcaagggt 6420
[1182] cacatatgca tttcacctgg actatgaaaa gtggaacaac catcaaagat tagagtcaac 6480
[1183] agaggatgta ttttctgtcc tagatcaagt gtttgattg aagagagtgt tttctagaac 6540
[1184] acacgagttt tttcaaaagg cctggatcta ttattcagac agatcagacc tcatcgggtt 6600
[1185] acgggaggat caaatatact gcttagatgc gtccaacggc ccaacctgtt ggaatggcca 6660
[1186] ggatggcggg ctagaaggct tacggcagaa gggctggagt ctagtcagct tattgatgat 6720
[1187] agatagagaa tctcaaatca ggaacacaag aaccaaaata ctagctcaag gagacaacca 6780
[1188] ggttttatgt ccgacataca tgttgcgcc agggctatct caagaggggc tcctctatga 6840
[1189] attggagaga atatcaagga atgcactttc gatatacaga gccgtcgagg aaggggcatc 6900
[1190] taagctaggg ctgatcatca agaaagaaga gaccatgtgt agttatgact tcctcatcta 6960
[1191] tggaaaaacc cttttgttta gaggtaacat attggtgcct gagtccaaaa gatgggccag 7020
[1192] agtctcttgc gtctctaatt accaaatagt caacctgcc aatataatgt cgacagtgtc 7080
[1193] caccaatgcg ctaacagtgg cacaacactc tcaatctttg atcaaaccga tgagggattt 7140
[1194] tctgctcatg tcagtacagg cagtctttca ctacctgcta tttagcccaa tcttaaaggg 7200
[1195] aagagtttac aagattctga gcgctgaagg ggagagcttt ctctagcca tgtcaaggat 7260
[1196] aatctatcta gatccttctt tgggagggat atctggaatg tccctcggaa gattccatat 7320
[1197] acgacagttc tcagaccctg tctctgaagg gttatcctc tggagagaga tctggttaag 7380
[1198] ctccaagag tcctggattc acgcgttgtg tcaagaggct ggaaaccag atcttggaga 7440
[1199] gagaacactc gagagcttca ctgccttct agaagatccg accaccttaa atatcagagg 7500
[1200] aggggccagt cctaccattc tactcaagga tgcaatcaga aaggctttat atgacgaggt 7560
[1201] ggacaagggt gaaaattcag agtttcgaga ggcaatcctg ttgtccaaga cccatagaga 7620
[1202] taattttata ctcttcttaa tatctgttga gcctctgtt cctcgattc tcagttagct 7680
[1203] attcagttcg tcttttttgg gaatccccga gtcaatcatt ggattgatac aaaactccc 7740
[1204] aacgataaga aggcagttta gaaagagtct ctcaaaaact ttagaagaat cttctacaa 7800
[1205] ctgagagatc cacgggatta gtcggatgac ccagacacct cagagggttg ggggggtgtg 7860
[1206] gccttgctct tcagagaggg cagatctact tagggagatc tcttggggaa gaaaagtgt 7920
[1207] aggcacgaca gttcctcacc cttctgagat gttgggatta cttccaagt cctctatttc 7980
[1208] ttgcaattgt ggagcaacag gaggaggcaa tcctagagtt tctgtatcag tactcccgtc 8040
[1209] ctttgatcag tcattttttt cagaggccc cctaaaggga tacttgggct cgtccacctc 8100
[1210] tatgtcgacc cagctattcc atgcatggga aaaagtcact aatgttcatg tggatgaag 8160
[1211] agctctatcg ttaaaagaat ctataaactg gttcattact agagattcca acttggctca 8220
[1212] agctctaatt aggaacatta tgtctctgac aggcctgat ttccctctag aggaggcccc 8280
[1213] tgtcttcaaa aggacgggt cagccttgca taggttcaag tctgccagat acagcgaagg 8340
[1214] aggttattct tctgtctgcc cgaacctcct ctctcatatt tctgttagta cagacacat 8400
[1215] gtctgatttg acccaagacg ggaagaacta cgatttcatg ttccagccat tgatgcttta 8460
[1216] tgcacagaca tggacatcag agctgttaca gagagacaca aggctaagag actctacgtt 8520
[1217] tcattggcac ctccgatgca acaggtgtgt gagaccatt gacgacgtga ccctggagac 8580

[1218]	ctctcagatc	ttcgagtttc	cggatgtgtc	gaaaagaata	tccagaatgg	tttctggggc	8640
[1219]	tgtgcctcac	ttccagaggc	ttcccgatat	ccgtctgaga	ccaggagatt	ttgaatctct	8700
[1220]	aagcggtaga	gaaaagtctc	accatatcgg	atcagctcag	gggctcttat	actcaatctt	8760
[1221]	agtggcaatt	cacgactcag	gatacaatga	tggaaccatc	ttccctgtca	acatatacgg	8820
[1222]	caaggtttcc	cctagagact	atttgagagg	gctcgcaagg	ggagtattga	taggatcctc	8880
[1223]	gatttgcttc	ttgacaagaa	tgacaaatat	caatattaat	agacctcttg	aattgggtctc	8940
[1224]	aggggtaatc	tcatatattc	tcttgaggct	agataacccat	ccctccttgt	acataatgct	9000
[1225]	cagagaaccg	tctcttagag	gagagatatt	ttctatccct	cagaaaatcc	ccgccgctta	9060
[1226]	tccaaccact	atgaaagaag	gcaacagatc	aatcttgtgt	tatctccaac	atgtgctacg	9120
[1227]	ctatgagcga	gagataatca	cggcgtctcc	agagaatgac	tggctatgga	tcttttcaga	9180
[1228]	ctttagaagt	gccaaaatga	cgtacctatc	cctcattact	taccagtctc	atcttctact	9240
[1229]	ccagagggtt	gagagaaacc	tatctaagag	tatgagagat	aacctgcgac	aattgagttc	9300
[1230]	tttgatgagg	caggtgctgg	gcgggcacgg	agaagatacc	ttagagtcag	acgacaacat	9360
[1231]	tcaacgactg	ctaaaagact	ctttacgaag	gacaagatgg	gtggatcaag	aggtgcgcca	9420
[1232]	tgcagctaga	accatgactg	gagattacag	ccccaacaag	aaggtgtccc	gtaaggtagg	9480
[1233]	atgttcagaa	tgggtctgct	ctgctcaaca	ggttgcaatc	tctacctcag	caaaccggcg	9540
[1234]	ccctgtctcg	gagcttgaca	taagggccct	ctctaagagg	ttccagaacc	ctttgatctc	9600
[1235]	gggcttgaga	gtggttcagt	gggcaaccgg	tgctcattat	aagcttaagc	ctattctaga	9660
[1236]	tgatctcaat	gttttcccat	ctctctgcct	tgtagttggg	gacgggtcag	gggggatatc	9720
[1237]	aagggcagtc	ctcaacatgt	ttccagatgc	caagcttgtg	ttcaacagtc	tttttagaggt	9780
[1238]	gaatgacctg	atggcttccg	gaacacatcc	actgcctcct	tcagcaatca	tgaggggagg	9840
[1239]	aaatgatatc	gtctccagag	tgatagatct	tgactcaatc	tgggaaaaac	cgtccgactt	9900
[1240]	gagaaacttg	gcaacctgga	aatacttcca	gtcagtccaa	aagcagggtc	acatgtccta	9960
[1241]	tgacctcatt	atttgcgatg	cagaagttac	tgacattgca	tctatcaacc	ggatcacccct	10020
[1242]	gttaatgtcc	gattttgcat	tgtctataga	tggaccactc	tatttgggtct	tcaaaactta	10080
[1243]	tgggactatg	ctagtaaata	caaaactaca	ggctattcaa	cacctgtcaa	gagcggtccc	10140
[1244]	ctcggtcaca	gggtttatca	cccaagtaac	ttcgtctttt	tcatctgagc	tctacctccg	10200
[1245]	atttccaaa	cgagggaagt	ttttcagaga	tgctgagtac	ttgacctctt	ccacccttcg	10260
[1246]	agaaatgagc	cttgtgttat	tcaattgtag	cagccccaag	agtgagatgc	agagagctcg	10320
[1247]	ttccttgaa	tatcaggatc	ttgtgagagg	atttcctgaa	gaaatcatat	caaatectta	10380
[1248]	caatgagatg	atcataactc	tgattgacag	tgatgtagaa	tcttttctag	tccacaagat	10440
[1249]	ggttgatgat	cttgagttac	agagggaac	tctgtctaaa	gtggctatca	ttatagccat	10500
[1250]	catgatagtt	ttctccaaca	gagtcttcaa	cgtttccaaa	cccctaactg	acccctcggt	10560
[1251]	ctatccaccg	tctgatccca	aaatcctgag	gcacttcaac	atatgttgca	gtactatgat	10620
[1252]	gtatctatct	actgctttag	gtgacgtccc	tagcttcgca	agacttcacg	acctgtataa	10680
[1253]	cagacctata	acttattact	tcagaaagca	agtcattcga	gggaacgttt	atctatcttg	10740
[1254]	gagttgggtc	aacgacacct	cagtgttcaa	aagggtagcc	tgtaattcta	gcctgagtct	10800
[1255]	gtcatctcac	tggatcagg	tgattttaca	gatagtgaag	actaccagac	tcgttggcag	10860
[1256]	catcaaggat	ctatccagag	aagtggaaag	acaccttcat	aggtacaaca	ggtggatcac	10920
[1257]	cctagaggat	atcagatcta	gatcatccct	actagactac	agttgcctgt	gaaccggata	10980
[1258]	ctcctggaag	cctgcccatt	ctaagactct	tgtgtgatgt	atcttgaaaa	aaacaagatc	11040
[1259]	ctaaatctga	acctttggtt	gtttgattgt	ttttctcatt	tttgttggtt	atttggttaag	11100

[1260]	cgtgggtcgg catggcatct ccacctctc gcggtccgac ctgggcatcc gaaggaggac	11160
[1261]	gcacgtccac tcggatggct aagggagaag gggaattcc agcacactgg cggccgttac	11220
[1262]	tagtggatcc gagctcggt ccaagctta gtttaaaccg ctgatcagcc tcgactgtgc	11280
[1263]	cttctagttag ccagccatct gttgtttgcc cctccccgt gccttccttg accctggaag	11340
[1264]	gtgccactcc cactgtcctt tcctaataaa atgaggaaat tgcacgcat tgtctgagta	11400
[1265]	ggtgtcattc tattctgggg ggtgggggtg ggcaggacag caagggggag gattgggaag	11460
[1266]	acaatagcag gcatgctggg gatgcggtg gctctatggc ttctgaggcg gaaagaacca	11520
[1267]	gctggggctc taggggtat cccacgcgc cctgtagcgg cgcattaagc gcggcgggtg	11580
[1268]	tggtggttac gcgcagcgtg accgtacac ttgccagcgc cctagcggc gctcctttcg	11640
[1269]	ctttcttccc ttctttctc gccacgttcg ccggctttcc ccgtcaagct ctaaactcggg	11700
[1270]	ggctcccttt agggttccga tttagtgtt tacggcacct cgaccccaaa aaacttgatt	11760
[1271]	agggtgatgg ttcacgtagt gggccatcgc cctgatagac ggtttttcgc cctttgacgt	11820
[1272]	tggagtccac gttctttaat agtggactct tgttccaaac tggaacaaca ctcaacccta	11880
[1273]	tctcggctta ttctttgat ttataaggga ttttgccgat ttcgccctat tggttaaaaa	11940
[1274]	atgagctgat ttaacaaaaa tttaacgcga attaatctg tggaatgtgt gtcagttagg	12000
[1275]	gtgtggaaag tccccaggt cccagcagg cagaagtatg caaagcatgc atctcaatta	12060
[1276]	gtcagcaacc aggtgtggaa agtccccagg ctccccagca ggcagaagta tgcaaagcat	12120
[1277]	gcatctcaat tagtcagcaa ccatagtcgc gccctaact ccgcccattc cgcccctaac	12180
[1278]	tccgcccagt tccgcccatt ctccgcccc tggctgacta atttttttta tttatgcaga	12240
[1279]	ggccgaggcc gcctctgcct ctgagctatt ccagaagtag tgaggaggt tttttggagg	12300
[1280]	cctaggtctt tgcaaaaagc tcccgggagc ttgtatatcc attttcggat ctgatcaaga	12360
[1281]	gacaggatga ggatcgtttc gcatgattga acaagatgga ttgcacgcag gttctccggc	12420
[1282]	cgcttgggtg gagaggctat tcggctatga ctgggcacaa cagacaatcg gctgctctga	12480
[1283]	tgccccgtg ttccggctgt cagcgcagg ggcgccggtt ctttttgtca agaccgacct	12540
[1284]	gtccggtgcc ctgaatgaac tgcaggacga ggcagcggc ctatcgtggc tggccacgac	12600
[1285]	ggcgcttct tgccgagctg tgctcgactg tgtcactgaa gcgggaagg actggctgct	12660
[1286]	attgggcgaa gtgccggggc aggatctcct gtcactcac cttgctcctg ccgagaaagt	12720
[1287]	atccatcatg gctgatgcaa tgcggcggt gcatagctt gatccggcta cctgcccatt	12780
[1288]	cgaccacca gcgaacatc gcatcgagc agcacgtact cggtatggaag ccggtcttgt	12840
[1289]	cgatcaggat gatctggacg aagagcatca ggggctcgc ccagccgaac tgttcgccag	12900
[1290]	gctcaaggcg cgcatgccc acggcgagga tctcgtcgtg acccatggcg atgcctgctt	12960
[1291]	gccgaatata atggtgaaa atggccgctt ttctggattc atcgactgtg gccggctggg	13020
[1292]	tgtggcggac cgctatcagg acatagcgtt ggctaccgt gatattgtg aagagcttgg	13080
[1293]	cggcgaatgg gctgaccgt tctcgtgct ttacggatc gccgtcccc attcgacgcg	13140
[1294]	catgccttc tatgccttc ttgacgagtt cttctgagcg ggactctggg gttcgaaatg	13200
[1295]	accgaccaag cgacgcccc cctgccatca cgagatttcg attccaccgc cgccttctat	13260
[1296]	gaaaggttgg gcttcggaat cgttttccgg gacgccggt ggatgatcct ccagcgcggg	13320
[1297]	gatctcatgc tggagtctt cgcaccccc aacttgttta ttgcagctta taatggttac	13380
[1298]	aaataaagca atagcatcac aaatttcaca aataaagcat ttttttact gcattctagt	13440
[1299]	tgtggtttgt ccaactcat caatgtatct tatcatgtct gtataccgtc gacctctagc	13500
[1300]	tagagcttgg cgtaatcatg gtcatagtct tttcctgtgt gaaattgtta tccgctcaca	13560
[1301]	attccacaca acatacgagc cggaagcata aagtgtaaag cctgggggtgc ctaatgagt	13620

[1302]	agctaactca cattaattgc gttgcgctca ctgcccgtt tccagtcggg aaacctgtcg	13680
[1303]	tgccagctgc attaatgaat cggccaacgc gcggggagag gcggtttgcg tattgggcgc	13740
[1304]	tcttccgctt cctcgtcac tgactcgtg cgctcggtcg ttcggctgcg gcgagcggtg	13800
[1305]	tcagctcact caaaggcgg aatacggta tccacagaat caggggataa cgcaggaaag	13860
[1306]	aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg	13920
[1307]	tttttccata ggctccgccc ccctgacgag catcacaaaa atcgacgctc aagtcagagg	13980
[1308]	tggcgaaacc cgacaggact ataaagatac caggcgtttc cccctggaag ctccctcgtg	14040
[1309]	cgctctcctg ttccgaccct gccgcttacc ggatacctgt ccgcctttct cccctcgga	14100
[1310]	agcgtggcgc tttctcatag ctacgctgt aggtatctca gttcggtgta ggtcgttcgc	14160
[1311]	tccaagctgg gctgtgtgca cgaaccccc gttcagccc accgctgcgc cttatccggt	14220
[1312]	aactatcgtc ttgagtccaa cccggttaaga cagacttat cgccactggc agcagccact	14280
[1313]	ggtaacagga ttagcagagc gaggtatgta ggcggtgcta cagagttctt gaagtgttg	14340
[1314]	cctaactacg gctacactag aagaacagta tttggtatct gcgctctgct gaagccagtt	14400
[1315]	accttcggaa aaagagttgg tagctcttga tccggcaaac aaaccaccgc tggtagcgg	14460
[1316]	ggtttttttg tttgcaagca gcagattacg cgcagaaaaa aaggatctca agaagatcct	14520
[1317]	ttgatctttt ctacggggtc tgacgctcag tggaacgaaa actcacgtta agggattttg	14580
[1318]	gtcatgagat tatcaaaaag gatcttcacc tagatccttt taaattaaaa atgaagtttt	14640
[1319]	aatcaatct aaagtatata tgagtaaact tggcttgaca gttaccaatg cttaatcagt	14700
[1320]	gaggcaccta tctcagcgat ctgtctatit cgttcatcca tagttgcctg actccccgc	14760
[1321]	gtgtagataa ctacgatacg ggagggtta ccatctggcc ccagtgcgc aatgataccg	14820
[1322]	cgagaccac gtcaccggc tccagattta tcagcaataa accagccagc cggaaggcc	14880
[1323]	gagcgcagaa gtggtcctgc aactttatcc gcctccatcc agtctattaa ttgttgccg	14940
[1324]	gaagctagag taagtagttc gccagttaat agtttgca acgttggtgc cattgctaca	15000
[1325]	ggcatcgtgg tgacacgctc gtcgtttggt atggcttcat tcagctccg ttcccaacga	15060
[1326]	tcaaggcgag ttacatgac ccccatgtt tgcaaaaaag cggttagtc ctteggctcct	15120
[1327]	ccgatcgttg tcagaagtaa gttggccgca gtgttatcac tcatggttat ggcagcactg	15180
[1328]	cataattctc ttactgtcat gccatccgta agatgcttt ctgtgactgg tgagtactca	15240
[1329]	accaagtcac tctgagaata gtgtatgcg cgaccgagtt gctcttgccc ggcgtcaata	15300
[1330]	cgggataata ccgcgccaca tagcagaact ttaaaagtgc tcatcattgg aaaacgttct	15360
[1331]	tcggggcgaa aactctcaag gatcttaccg ctgttgagat ccagttcgat gtaaccact	15420
[1332]	cgtgcacca actgatctc agcatcttt actttcacca gcgtttctgg gtgagcaaaa	15480
[1333]	acaggaaggc aaaatgccgc aaaaaaggga ataaggcgca cacggaatg ttgaatactc	15540
[1334]	atactcttcc tttttcaata ttattgaagc atttatcagg gttattgtct catgagcgga	15600
[1335]	tacatatttg aatgtattta gaaaaataaa caaatagggg ttccgcgcac atttccccga	15660
[1336]	aaagtgccac ctgacgctc	15678
[1337]	<210>	21
[1338]	<211>	10156
[1339]	<212>	RNA
[1340]	<213>	人工序列
[1341]	<220>	
[1342]	<223>	负义载体RNA基因组
[1343]	<400>	21

[1344]	acgcuuaaca aaauaacaac aaaaauagaga aaaacaauca aacaaccaa gguucagauu	60
[1345]	uaggauucuug uuuuuuucac gauacauacac acaagagucu uagcaugggc aggcuccag	120
[1346]	gaguauccgg uucacaggca acuguagucu aguagggau gcuagauucu gauauccucu	180
[1347]	aggguagucc accuguugua ccuaugaagg ugucuuucca cuucucugga uagaucuuug	240
[1348]	augcugccaa cgagucuggu agucuuacac aucuuguaaa ucaaccugau ccagugagau	300
[1349]	gacagacuca ggcuagaauu acaggcuacc cuuuugaaca cugagguguc guuggacca	360
[1350]	cuccaagaua gauaaacguu cccucgaaug acuugcuuuc ugaaguaaau aguuauaggu	420
[1351]	cuguuauaca ggucgugaag ucuugcgaag cuaggacgu caccuaaagc aguagauaga	480
[1352]	uacaucauag uacugcaaca uauuguagag ugccucagga uuuugggaur agacggugga	540
[1353]	uagaacgagg ggucaguug ggguuuggaa acguugaaga cucuguugga gaaaacuauc	600
[1354]	augauggcuu uauagauagc cacuuuagac agaguucucc ucuguaacuc aagaucua	660
[1355]	accuauugu ggacuagaaa agauucuaa ucacugucua ucagaguauu gaucaucua	720
[1356]	uuguaaggau uugauaugau uucuuacagga aaucucucua caagaucug auaguuaag	780
[1357]	gaacgagcuc ucugcaucuc acucuuuggg cugcuacaau ugaauaacac aaggcuauu	840
[1358]	ucucgaaggg uggaagaggu caaguacua gcaucucua aaaacuucc ucguuuggag	900
[1359]	aaucggaggu agagcucaga ugaaaaagac gaaguuaauu gggugauaaa ccugugacc	960
[1360]	gagggaacg cucuugacag guguuagaa gccuuguagu uuggauuuac uagcauaguc	1020
[1361]	ccauaaguuu ugaagacca auagaguggu ccaucuaug acaaugcaa aucggacauu	1080
[1362]	aacaggguga uccgguuugau agaugcaaug ucaguaacuu cugcaucgca aauaagagg	1140
[1363]	ucauaggaca uguugaccug cuuuuggacu gacuggaagu auuuccaggu ugccaaguuu	1200
[1364]	cucaagucgg acgguuuuuc ccagauugag ucaagaucua ucacucugga gacgauaua	1260
[1365]	uuuccucucc ucaugauugc ugaaggaggc aguggaugug uuccggaagc caucaggua	1320
[1366]	uucaccucua aaagacuguu gaacacaagc uggcaucug gaaacauguu gaggacugcc	1380
[1367]	cuugauaucc cccugaccc gucccaacu acaaggcaga gagaugggaa acauugaga	1440
[1368]	ucaucuaaga uaggcuuag cuuauaaua gaccggug cccacugaac cacucuaag	1500
[1369]	cccagauca aaggguucug gaaccucua gagagggccc uuaugucaag cuccgagaca	1560
[1370]	ggggccgggu uugcugaggu agagacugca accuguugag cagagcagac ccauucugaa	1620
[1371]	cauccuaccu uacgggacac cuucuuugug gggcuguaau cuccagucua gguucuauc	1680
[1372]	gcauggcga ccucuuagc caccuauuu guccuucgua aagagucuuu uagcagucgu	1740
[1373]	ugaauuguu cgucugacuc uaagguauu ucuccugucc cgcccagcac cugccuacuc	1800
[1374]	aaagaucua auugucgag guuauucuc auacucuuag auagguuuu cucaaccuc	1860
[1375]	uggaguaga gaugagacug guaaguauu agggauaggu acguauuuu ggcacuucua	1920
[1376]	aagucugaaa agaucauag ccagucuuu ucuggagacg ccgugauuu cuucgcua	1980
[1377]	uagcguagca cauguuggag auaacacaag auugaucug ugccuucuuu cauaguggu	2040
[1378]	ggauaagcgg cggggauuu cugagggaau gaaaauaucu cuccucuaag agacgguuu	2100
[1379]	cugagcauuu uguacaagga ggaugguuu ucuagccua ggagaauua ugagauuacc	2160
[1380]	ccugagacca auucaagagg ucuauuaau uugauuuu ucauucuuu caagaagca	2220
[1381]	aucgaggau cuaucaauu ucccuugcg agccucua aauagucuc aggggaaacc	2280
[1382]	uugccguaua uguugacagg gaagauguu ccauauuu auccugaguc gugaauugcc	2340
[1383]	acuaagauu aguauaagag cccugagcu gaucggaau ggugagacu uucucuaacc	2400
[1384]	cuuagagauu caaaucucc uggucucaga cggauaucgg gaagccucug gaagugaggc	2460
[1385]	acagcccag aaaccauuu ggauauucuu uucgacacau ccggaacuc gaagaucua	2520

[1386]	gaggucucca	gggucacguc	gucaaugggu	cucacacacc	uguugcaucg	gaggugccaa	2580
[1387]	ugaaacguag	agucucuuag	ccuugugucu	cucuguacca	gcucugaugu	ccaugucugu	2640
[1388]	gcuaaaagca	ucaauggcug	gaacaugaaa	ucguaguucu	ucccgucuuu	ggucaaaauca	2700
[1389]	gacauggugu	cuguacuaac	agaaauauga	gagaggaggu	ucgggcagac	agaagaauac	2760
[1390]	ccuccuucgc	uguaucuggc	agacuugaac	cuaugcaagg	cugaccccg	ccuuuugaag	2820
[1391]	acaggggccu	ccucuagagg	gaaaucaggg	ccugucagag	acauaauguu	ccuaauuaga	2880
[1392]	gcuugagcca	aguuggaauc	ucuaguaaug	aaccaguuuu	uagauucuuu	uaacgauaga	2940
[1393]	gcucucuuca	ccacaugaac	auuagugacu	uuuucccaug	cauggaauag	cugggucgac	3000
[1394]	auagaggugg	acgagcccaa	guauccuuu	agggggccuc	gugaaaaaaa	ugacugauca	3060
[1395]	aaggacggga	guacugauac	agaaacucua	ggauugccuc	cuccuguugc	uccacaagug	3120
[1396]	caagaaauag	aggacuuggg	aaguuaucuu	aacaucucag	aaggguagg	aacugucgug	3180
[1397]	ccuaccacuu	uuuuuuuu	agagaucucc	cuaaguagau	cugcccucuc	ugaagagcaa	3240
[1398]	ggccacaccc	ccccaacccu	cugagguguc	ugggucaucc	gacuaauccc	guggaucucu	3300
[1399]	gaguuguaga	aggauucuu	uaaaguuuu	gagagacucu	uucuaaacug	ccuucuuau	3360
[1400]	guucgggagu	uuuguaucaa	uccaauugu	gacucgggga	uucccaaaaa	agacgaacug	3420
[1401]	aauagcucac	ugagaaucg	aggaaacaga	ggcucaacag	auauuaagaa	gaguauaaaa	3480
[1402]	uuauucua	gggucuuuga	caacaggauu	gccucucgaa	acucugaauu	uuccaccuug	3540
[1403]	uccaccucgu	cauauaaagc	cuuucugauu	gcauccuuga	guagaauugu	aggacuggcc	3600
[1404]	ccuccucuga	uauuuuaggu	ggucgggacu	ucuagaaggc	gagugaagcu	cucgaguguu	3660
[1405]	cucucuccaa	gaucuggguu	uccagccucu	ugacacaacg	cgugaaucca	ggacucuuug	3720
[1406]	gagcuuaacc	agaucucucu	ccagaaggau	aaccuuucag	agacagggu	ugagaacugu	3780
[1407]	cguauaugga	aucuuccgag	ggacauucca	gauauccuc	ccaaagaagg	aucuagauag	3840
[1408]	auuauccuug	acauggcuag	gagaaagcuc	ucccuucag	cgcucagaau	cuuguaaacu	3900
[1409]	cuuccuuua	agauugggcu	aaauagcagg	uagugaaaga	cugccuguac	ugacaugagc	3960
[1410]	agaaaaucc	ucaucgguuu	gaucuaagau	ugagaguguu	gugccacugu	uagcgcauug	4020
[1411]	guggacacug	ucgacauuau	auuggcgagg	uugacuauuu	ggucuuuaga	gacgcaagag	4080
[1412]	acucuggccc	aucuuuugga	cucaggcacc	aaauuguuac	cucuaaaca	agggguuuu	4140
[1413]	ccauagauga	ggaagucua	acuacacaug	gucucuucuu	ucuugaugau	cagcccuagc	4200
[1414]	uuagaugccc	cuuccucgac	ggcucugua	aucgaaagug	cauuccuuga	uauucucucc	4260
[1415]	aaucacua	ggagccccuc	uugagauagc	ccuggcgaca	acauguaugu	cggacauaaa	4320
[1416]	accugguugu	cuccuugagc	uaguauuuu	guucuuugu	uccugauuu	agauucucua	4380
[1417]	ucuaucua	auaagcugac	uagacuccag	ccuucugcc	guaagccuuc	uagcccgcca	4440
[1418]	uccuggccau	uccaacaggu	ugggccguug	gacgcaucua	agcaguauau	uugaucucc	4500
[1419]	cguaacccga	ugaggucuga	ucugucugaa	uaauagauc	aggccuuuu	aaaaaacucg	4560
[1420]	uguguucuag	aaaacacucu	cuucaaucca	aacacuugau	cuaggacaga	aaauacaucc	4620
[1421]	ucuguugacu	cuaauucuu	augguuguuc	cacuuuuc	aguccaggug	aaugcauau	4680
[1422]	gugacccuug	aaauagucca	aagcccuugc	ccggugaccc	ugucgaucag	cuuuuuuac	4740
[1423]	accuuguuca	gguugucugu	cauagucagc	gcgucaaaaa	guggcaagau	guaguuggcc	4800
[1424]	aagaguuuu	cagugaugac	aaaauacaau	cuuagauucc	augacauuag	agcaaagaau	4860
[1425]	cgaccuuc	ucuucaauuc	ccguuccuuu	ggcuugaggc	cauuuauca	gucuucacu	4920
[1426]	ggcaauccuc	cgaggucua	agaccucaga	aacucucggg	gauugacagg	cggcuuagac	4980
[1427]	agggccguga	uaauaacuuu	uucgcuagga	acaggccccc	cucgguuuuc	ugacagccaa	5040

[1428]	gaagcuaguc ucguucuggu gaaagaau guuuuguc au ccaauuuuc ugacggaucc	5100
[1429]	auugauucag gaauucuaa gaucugcgug aucgggagcu ugugccaugu auccccacc	5160
[1430]	aagucuacaa uauuuuggg uggccaugu uggguuuuga uauaaggagu caaggggugg	5220
[1431]	ucucgggcua ggaaucuuga auccagauac cacuuggagu acuuaucaaa accccaucua	5280
[1432]	aggauccucc uggcuagguc guugcuaag cacuccuggu aggacuauac uaucauuuu	5340
[1433]	uuuaggugaa ccugaucaua uaguuuugac agaccuuuc gaaauucua auuagguguc	5400
[1434]	ccccaugcc uguaacagcc aaacacaaaa accaaguc au guauguuguc gaauugaucc	5460
[1435]	agagcccuua agaaccuuu ugcacaggga ccgaacgucu cuucaaguug acuuaccuug	5520
[1436]	ucuuuuauaa auacaggaaa gucuccaag gaauaaua gaggccuaa cuuuucugcu	5580
[1437]	cucuggacua aacuaucac gacauauggc uccaauuuu ugaugacuuc auagccggag	5640
[1438]	uuuccacaca uagacaagac uugauccca gcaauguaca gcuggcauag uugagaua	5700
[1439]	aagucaucug aguaucggg cucuggggga gagagcaaga ccauuaggau guugaagcga	5760
[1440]	gacaagaaaa gaucuuuuag cauaagugug uaguuuucgu caaaaagaca auuggaacu	5820
[1441]	ugggaguaaa caaaguccu ugucacgauc agcagucucc auuuuuguc uuugaacuu	5880
[1442]	accaaguccu ucccgauguc cacugagggu aaauuuucc auuauugcu gauggucuu	5940
[1443]	ucuuauccc agucuaaggc guucaugau aaggugaua cauggaagaa caaguaagag	6000
[1444]	gaauacgugu uggcaagau ccuuccaaa gcauuauca auucaaccu cucaaggca	6060
[1445]	cuuaacacuc ccucuggggg gauucucag ccucuaauu cuagcgugag auucaacagc	6120
[1446]	uucucuaugg gggacgacu ggaauagaaa ugggccaagu cuguuauaca ucuccggc	6180
[1447]	cuguuggauu cagagugggc accauuaac cagagagaaa uauuugacug ugcagccau	6240
[1448]	ccgcccaccu ugagagaacc caaauuacc uucuuagaa auuauucaa aacucugaa	6300
[1449]	gaccuggagc auuugucug uagagucac cgauaagguc uauucccug uuuuacac	6360
[1450]	ucuaacaua gucuagcagg aucuucua auaggagagu ugagauugua gucagagu	6420
[1451]	cucaagauu uggggacaa gggguuccu cuggguucag ccucuaacuc gauugggu	6480
[1452]	auagguc auuagaccuc uccaggau agcaucua aguaaguc agguugagaa	6540
[1453]	aguuugcca guuuuuu ugaucuaa uuuuuuuc gacugaaa cuaccgagg	6600
[1454]	uaggcgcc guuuuuu aggggggug uuuuuuu uauugcuag guagccggg	6660
[1455]	cguagcgcc gcguaccugc agguccuu agggauuu auuuuuu ucacauca	6720
[1456]	gaggaucaaa auguaaaggu auucaacu acuuuga auuagcagg uuuuauagg	6780
[1457]	ccauaacaug uuuuuuu uauugcucc aaggguuga guugagcca uauuucuc	6840
[1458]	ccuccaggg uaaacaag auuuuuu gggauuuu cuagcuagc uuauucua	6900
[1459]	agcagagagg auuuuuguc cucuucggc cuuuguguc gaagagacu gucagacc	6960
[1460]	aguugac auuucgggu cauuugua caccagac ugcuccggu augacacug	7020
[1461]	uuugcauca cuuuuuu cauuccagc aacucagau caucaucca agugaucuc	7080
[1462]	ugagaguuu ccaacuccc ccuucaga gggcccuag auuagccca cuugaagau	7140
[1463]	aagguucucc ucauuuuga uaccaguc agggccuag ggacuggaga uccugacaa	7200
[1464]	gccagucca uaaccacuu gacuaaccg auuauccu auuuccaga auuauucg	7260
[1465]	ucgaauuu ucagaauug ccgcaggau cugaacgag auuauucg gcuaacacu	7320
[1466]	uuuaccuuc cuuugaua aaaguucuc auguucuu ugcuguaag uuuuucagc	7380
[1467]	gggacguuu cagggggg aagccaca ucaucguc ccagagggg ugacgaggga	7440
[1468]	gaggauuuu gaguuccuc guccucggg uuuuucac ucuuacgu gagguuac	7500
[1469]	gauuuaucag ugguguugc uuuuuuuu auguugacu ugggacacu cgauuuu	7560

[1470]	ugucuagagg	gacugagggg	agagguucgg	uuagcaagau	guauagcgau	ucaagucauc	7620
[1471]	uugcaugauu	uuacucagcu	ugucggauuc	gacuaacaac	uggaauuucu	uagaguuggc	7680
[1472]	caaagcgacc	caucccagua	cacauuuuag	ggggaguuug	gacccgucan	gggcuaaacg	7740
[1473]	ggucacaccu	gguacauuuu	uugccucuuu	aacuaauuca	ucaagguuca	uuuucanug	7800
[1474]	cuaaaaaua	uacaagagua	ucccugagga	ucgagaggga	aacuuauuuu	uuuuggagaa	7860
[1475]	acuuucugca	aucuggugag	cgaucucagc	cuccacugau	agaucanccu	cuucanuggu	7920
[1476]	agccgaccan	ucaagggcug	gagggccaga	agcaauuuga	gccgccancc	uggcuuucga	7980
[1477]	ugauuggcuu	ucucucugag	aaggaguggg	uguugucucc	uuuugagcu	cucggccagu	8040
[1478]	agucuggguu	gauuuauccu	cugaagacuu	uccuggaggg	uugggaaagu	ugaccgcgac	8100
[1479]	auaggauaua	aucucuucua	cggucuguga	ccauaucuug	agaaanucuu	cuccugaccu	8160
[1480]	cauuugucug	acuauuugga	cuccaacaau	uuccaggua	gacuggaaca	ggaagcuagg	8220
[1481]	auccucuccu	ucauccanccu	gaaaguccuc	ucgauacuug	ccuucuccca	ccuuggcuau	8280
[1482]	cucaccaugg	uugggcgauu	uuccancauc	caggugaagu	cgcgccanau	ccucagggag	8340
[1483]	auuguccacc	ucuanuggguu	ccccuuggag	augagccuga	uugucuucga	uauuucanuu	8400
[1484]	gaucagauca	acaguuucuu	cagccanuc	aagaucggcc	agaccggcuc	uaauagcacu	8460
[1485]	aggauugaca	aagaucuuuc	ucauuuugg	gaugguucga	aaggaggggu	guuaguuuuu	8520
[1486]	uucaugaugg	auauacacaa	uccguagauu	uccggcauuu	uguuauucaa	cuucuuacac	8580
[1487]	auugauccua	gcagaagcac	aggcugcagg	gugacggucc	aucccgucuu	ccugggcaca	8640
[1488]	agacaugggc	agcgugccan	cauccugcuc	cuccaccucc	ggcgggagc	cauggcuacc	8700
[1489]	ggaucccga	uggaaguaga	gguucucucc	ggaaccugag	ucacucgaau	augucuuguu	8760
[1490]	uagaaacucg	gcgaauagau	uuggacgggc	uugaugauug	gaacugacug	agacauaucu	8820
[1491]	ccguauugua	gaucucuuaa	gucgaccucc	auucaucaug	auucgaguau	aaacagccuc	8880
[1492]	cggacuucug	guuucaccug	aaaaguaguc	cucgucguca	gaguugacag	uuccancauc	8940
[1493]	ugccagugcu	acgucagucu	uugucaguuc	agccgccucg	uauucuuuga	guucuuucuc	9000
[1494]	aucucugaag	aaucuuucuu	caaauugucc	uuucccgag	aaauccucuc	ccagauagcc	9060
[1495]	cccuagaaca	gacauuucan	gaggagcaca	ugcagcaaua	accguugcan	uuagggauc	9120
[1496]	gacuugaccc	auauagcauc	cuacaaagug	aaugagauug	aacacgugac	caacagcanu	9180
[1497]	ugaugaauaa	ggagauuucc	cacucaagcc	uagugaacgg	aaguggauga	aaauagagug	9240
[1498]	aggaacagcu	gucuccugcc	cuggcucaaa	cauucuuuc	aucucuuccu	caaaguucuu	9300
[1499]	guggaagaaa	uauaguauug	ccucucuagc	ggugagauug	auuuguuuua	ugaaccagcu	9360
[1500]	aaaugauacc	aguccugaac	agucuucan	agcagugaca	acugugccca	cucugauugc	9420
[1501]	ugaauauaga	ugcucaaucc	gggagaaaaa	caugucanag	guuccggcca	aaaauucgaa	9480
[1502]	guuugguaua	guacuccaan	uagcacacau	uuugugaguu	gucauuagag	uauggguuuc	9540
[1503]	cagauuuua	acaaaagggg	cugucucaaa	aaucugcucu	auccugucug	caauuuugu	9600
[1504]	cuuauaguua	ccaguguuuu	gcccggauau	uuugcucaac	cuauacagac	ucaagagaag	9660
[1505]	accgacuaag	gacgcaugcu	cagggacagu	ggggucucuu	gucaguucca	ugccuccugu	9720
[1506]	cagagcccaa	uucccuucua	caucaguacg	uuuuauucc	accagagaa	cuggggugau	9780
[1507]	cuuauccucc	uuucgugcaa	ucacauuucc	auagcugguc	cagucuuccg	gacauugccc	9840
[1508]	cuaaaaaaac	ugcauugccg	cugccaaaau	ggaacauaca	ucgucaggau	uaaguugggc	9900
[1509]	ggcgcucaug	ccugacaaaa	cugacuugua	ugcuuuauuu	aaaucgggag	ccuuuccuag	9960
[1510]	gguuauacag	ggcuuuuua	aaucuuugau	ggcagggua	uuguacucan	auugauccac	10020
[1511]	gauaaucuca	ggcuucaag	agaccaccug	auuauugacu	uugaauacaa	ucuugucggc	10080

[1512]	auccauugua gggguguuac auuuuugcuu ugcaauugac aaugucuguu uuuuuuuga	10140
[1513]	ucugguuguu aagcgu	10156

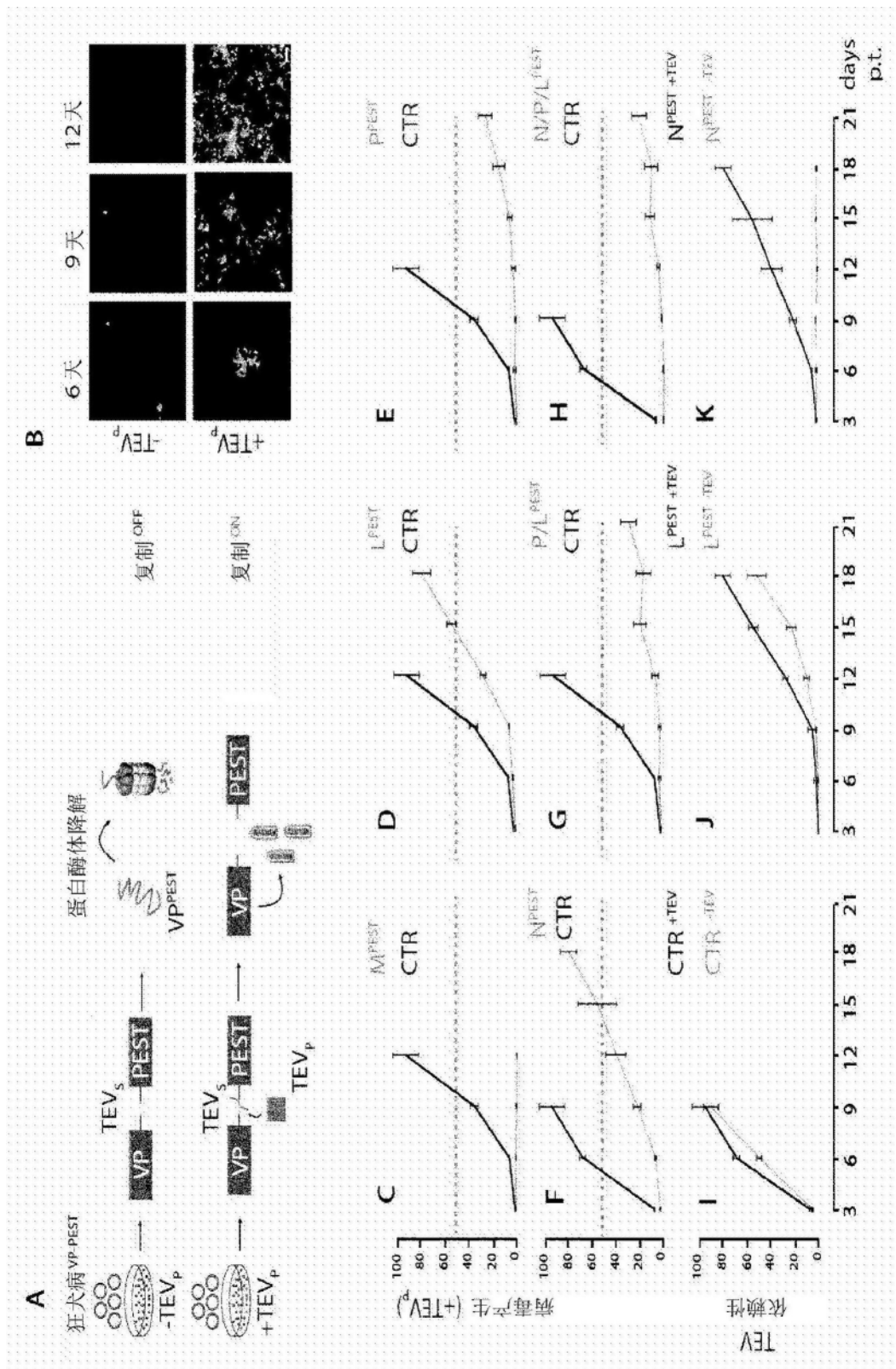


图1

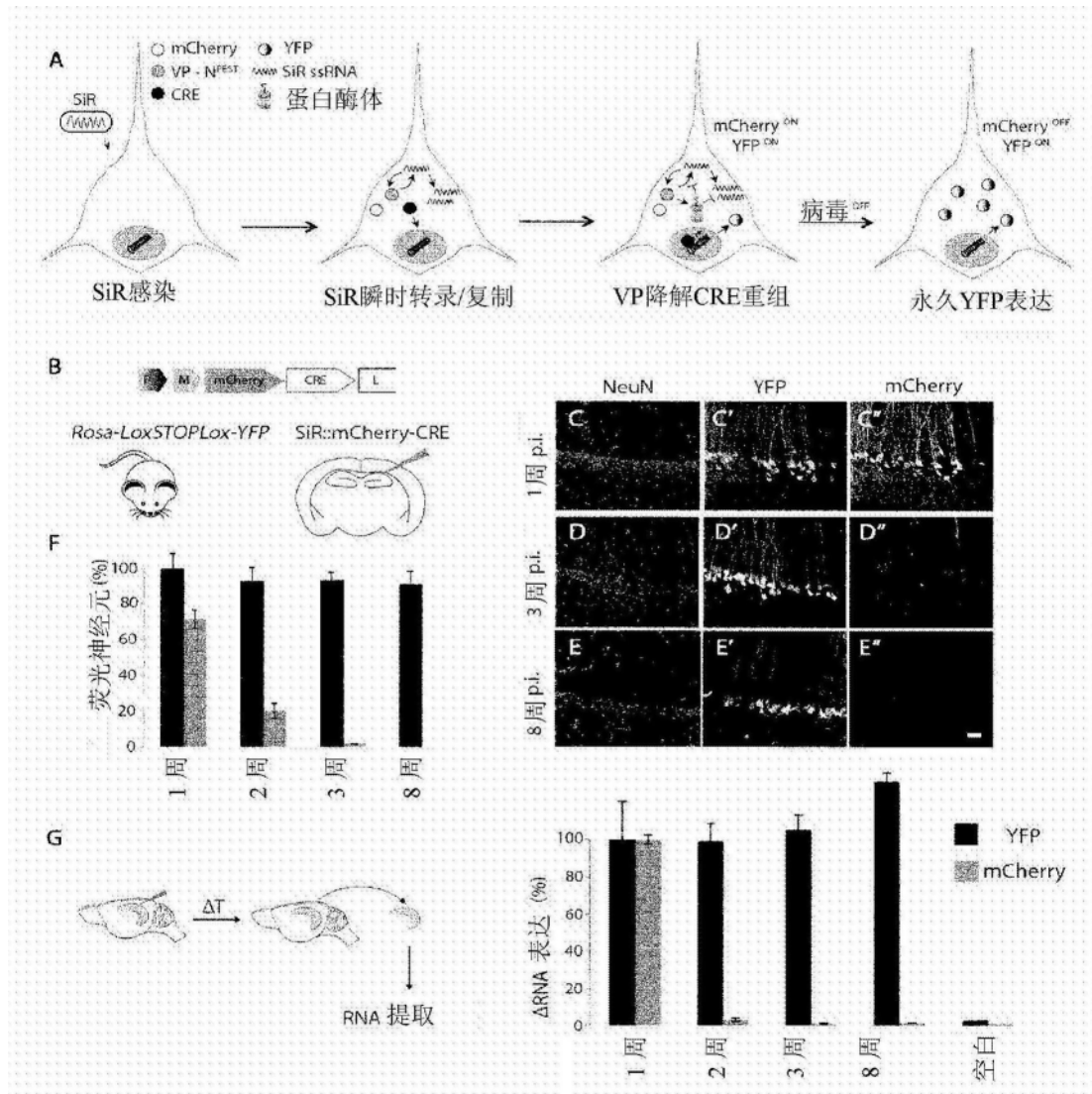


图2

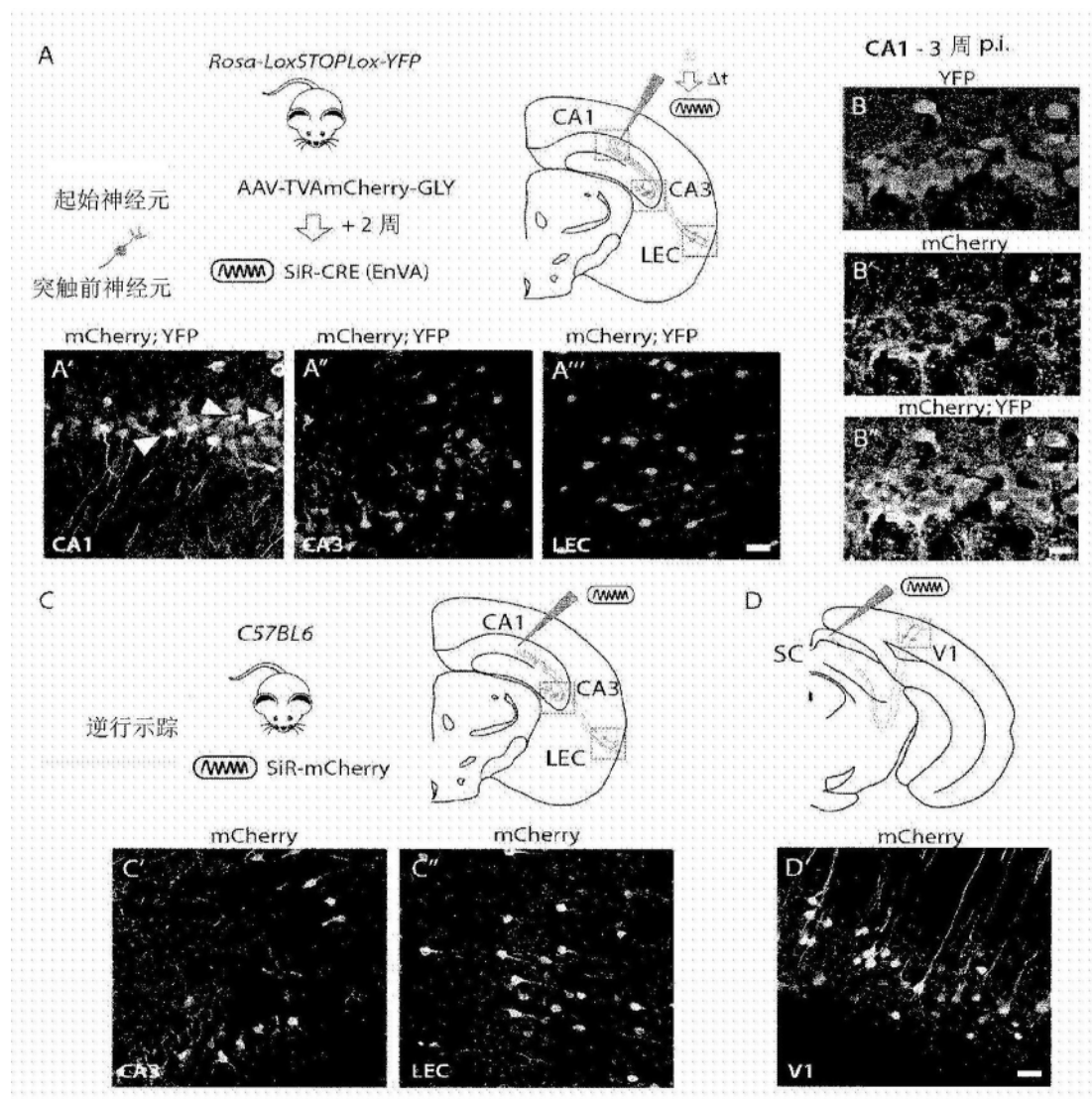


图3

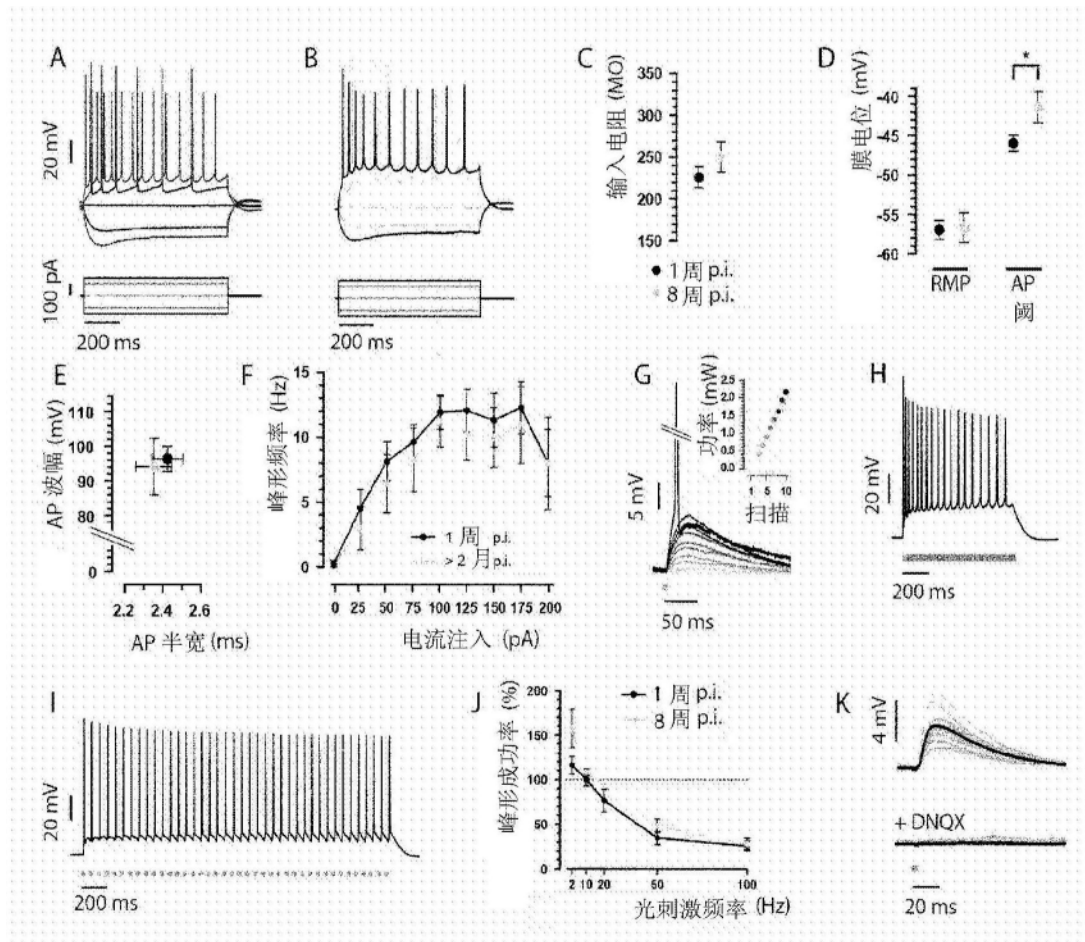


图4

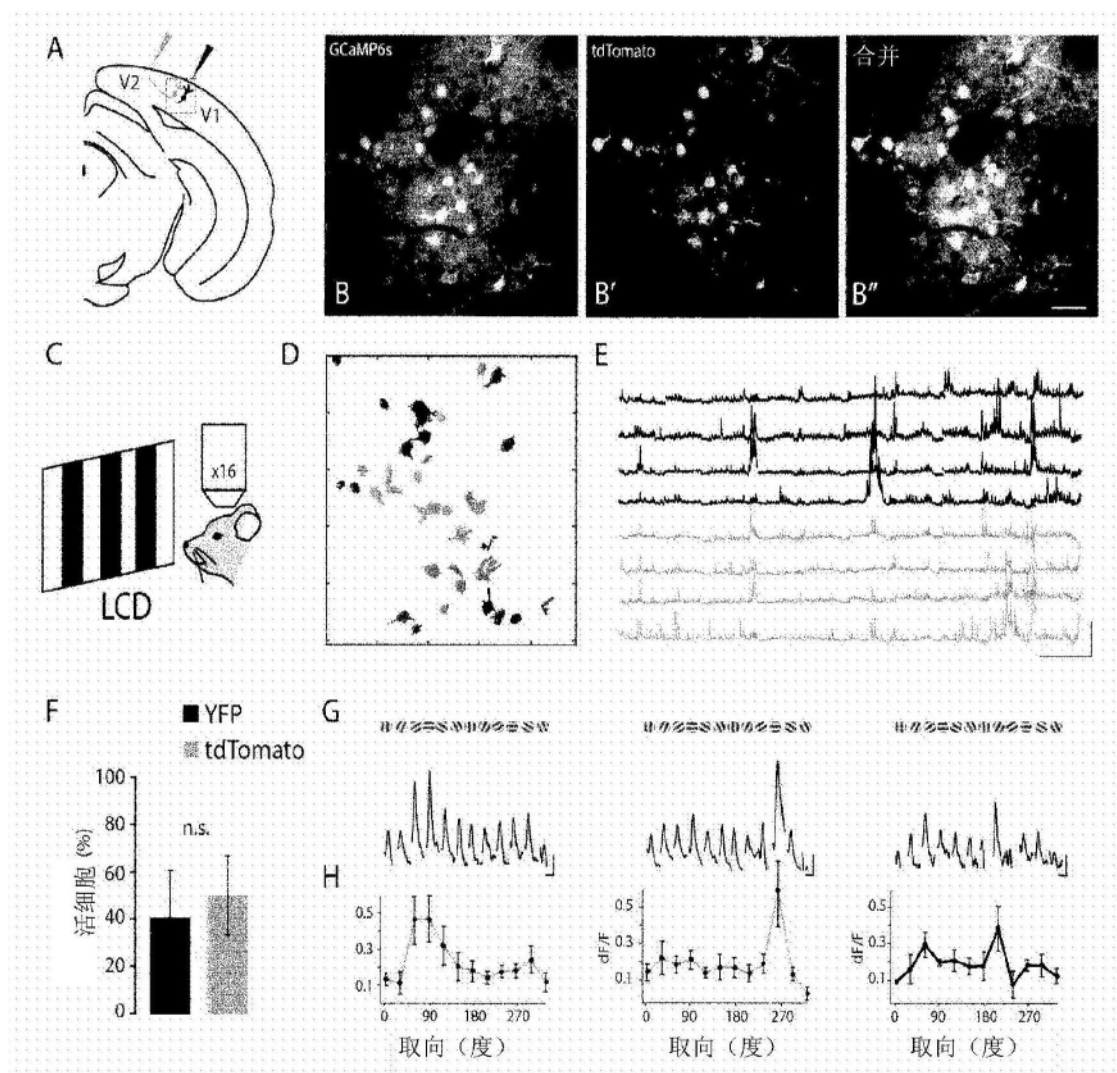


图5

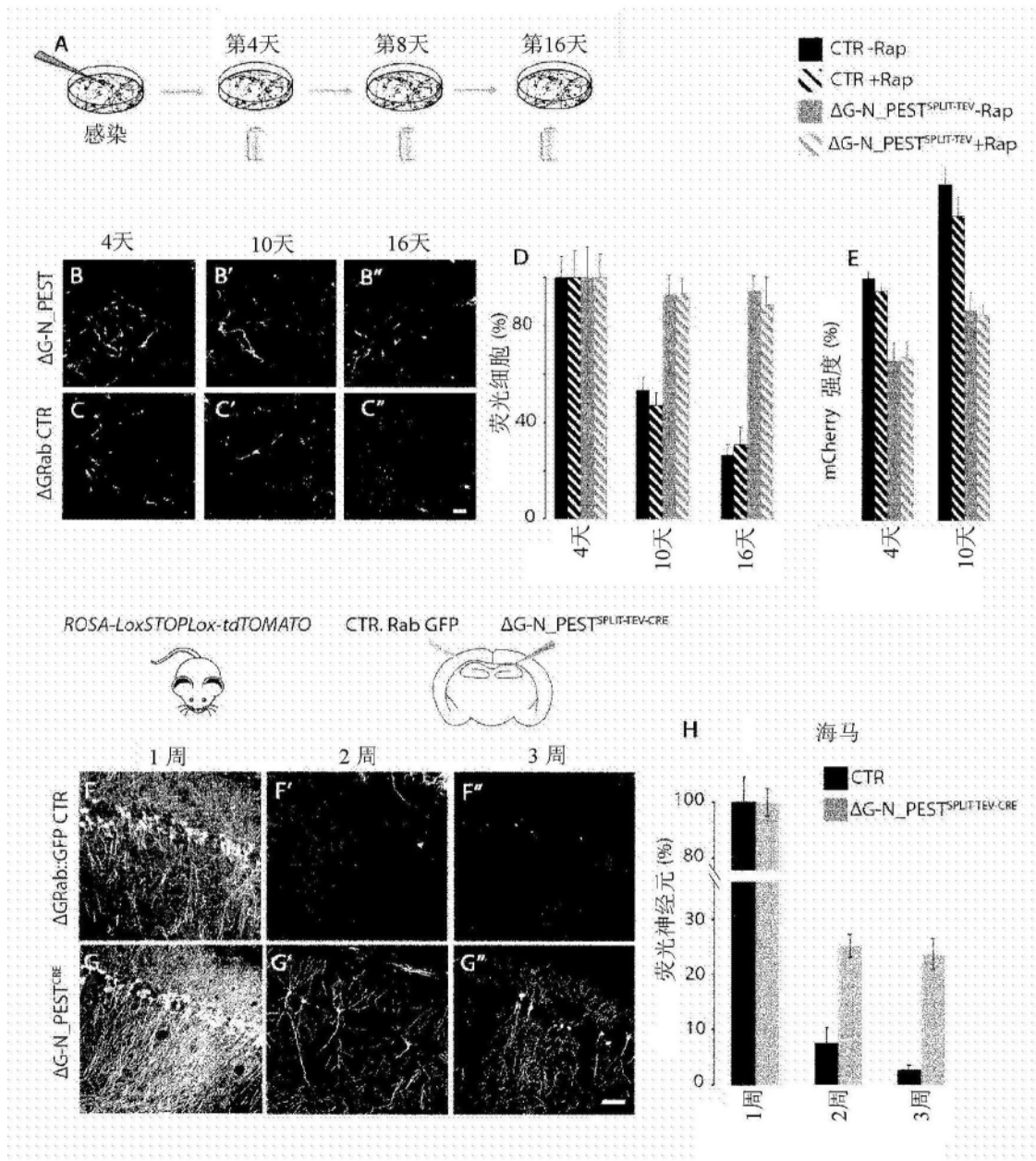


图6

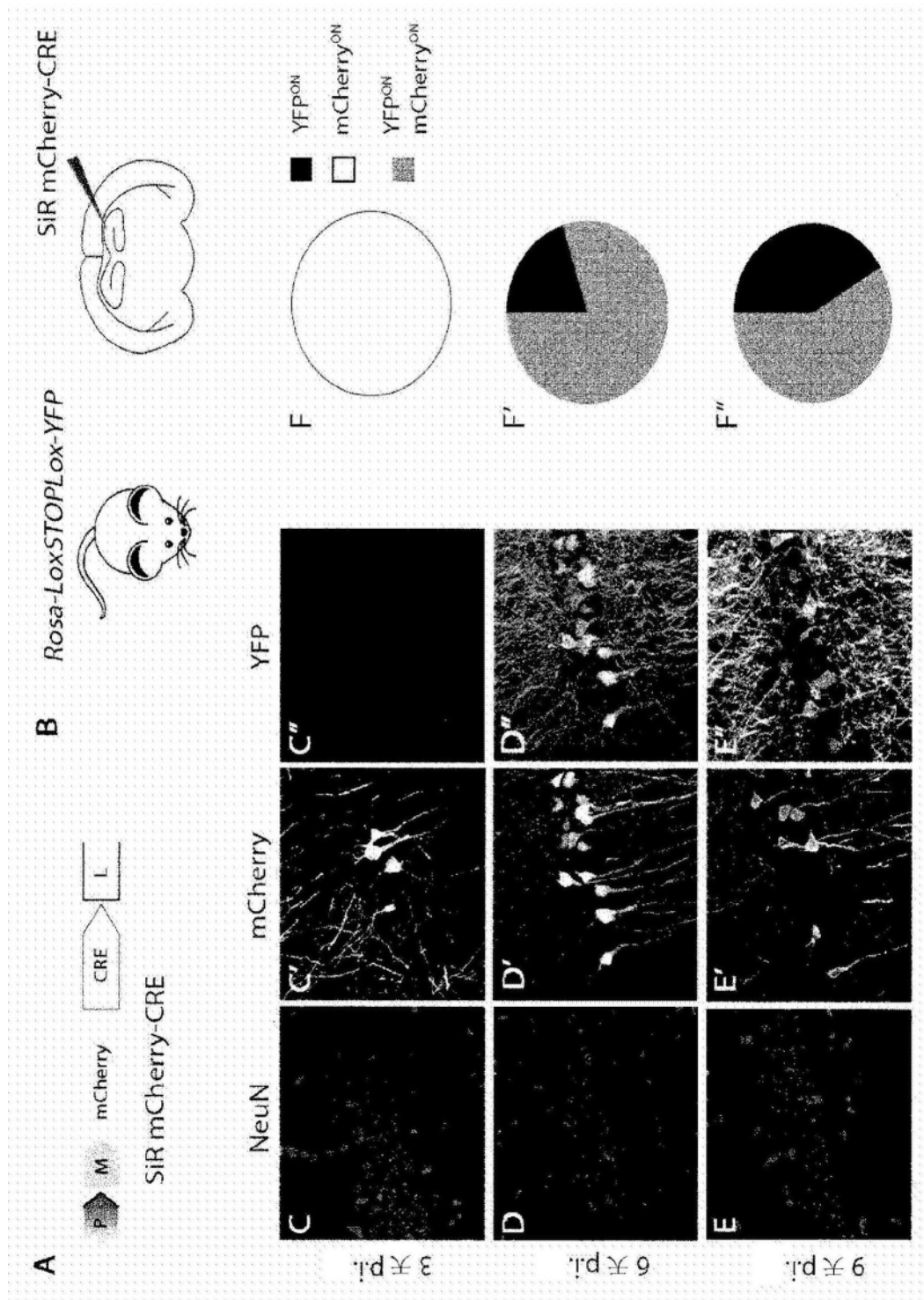


图8

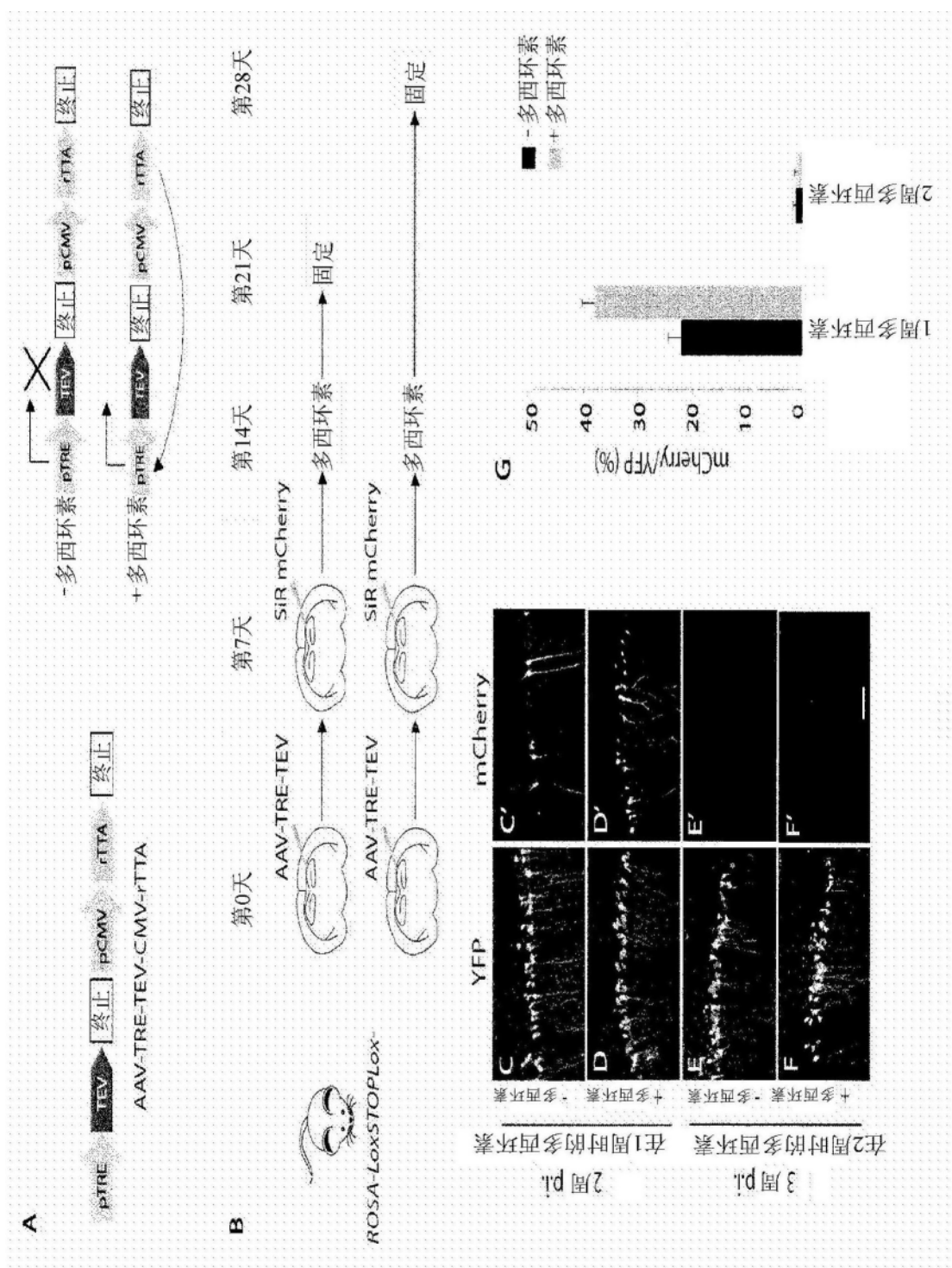


图9

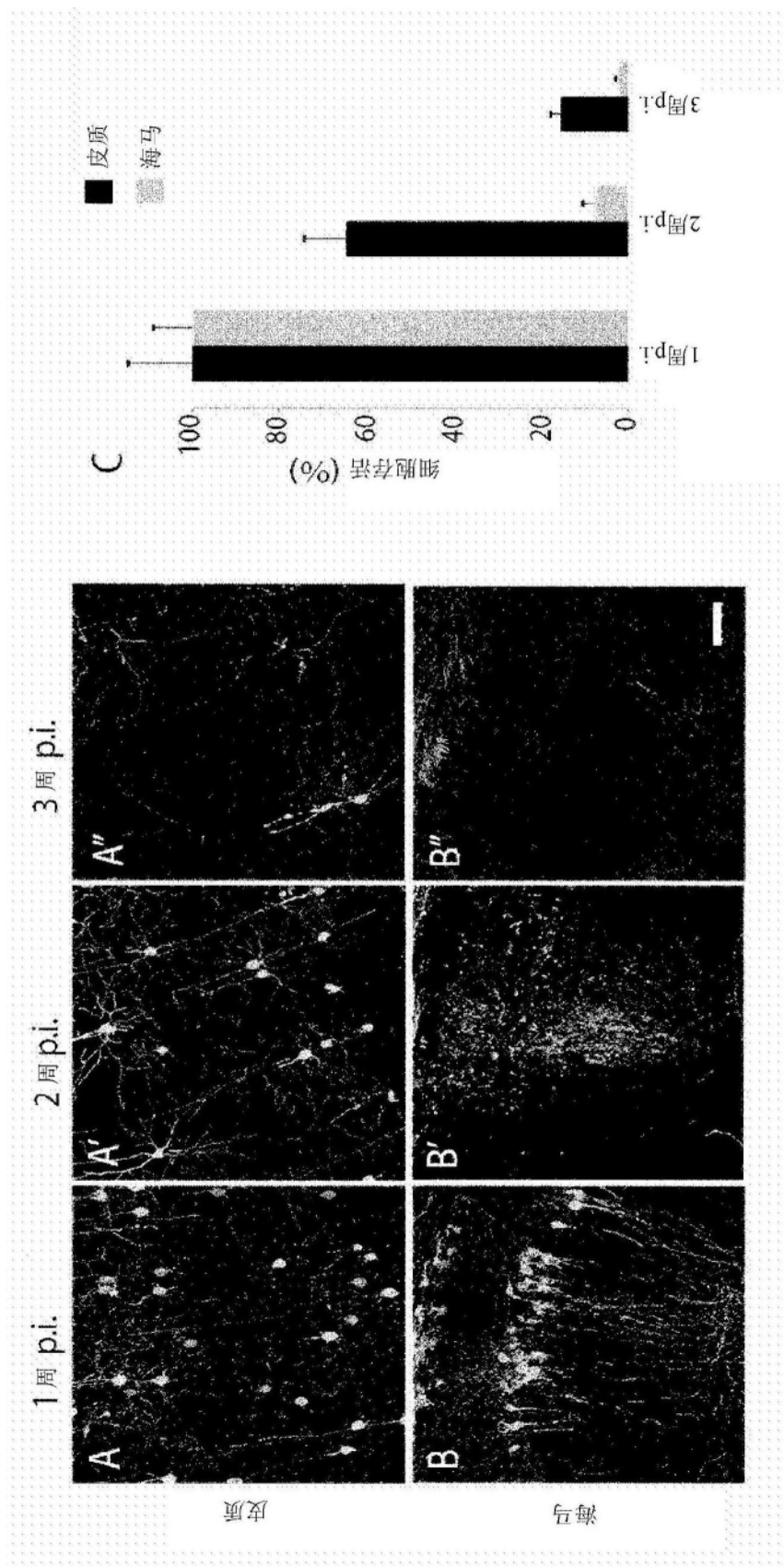


图10

(B)

GACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGA
GTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCGCCATTGACG
TCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCATTGACGTCAATGGGTGGAGTATT
TACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAA
TGACGGTAAATGGCCCGCCTGGCATTATGCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTAC
ATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGC
GGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTGTTTGGCACCAAAA
TCAACGGGACTTTCCAAAATGTCGTAACAACCTCCGCCCCATTGACGCCAATGGGCGGTAGGCGTGTACGG
TGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAAGTAGAGAACCCTGCTTACTGGCTTATCGAAATTA
ATACGACTCACTATAGGGAGACCCAAAGCTGGCTAGATTAGCGTCTGATGAGTCCGTGAGGACGAAACCC
GGAGTCCCGGGTCACGCTTAACAACCCAGATCAAAGAAAAACAGACATTGTCAATTGCAAAGCAAAAAATG
TAACACCCCTACAATGGATGCCGACAAGATTGTATTCAAAGTCAATAATCAGGTGGTCTCTTTGAAGCCT
GAGATTATCGTGGATCAATATGAGTACAAGTACCCTGCCATCAAAGATTTGAAAAAGCCCTGTATAACCC
TAGGAAAGGCTCCCGATTTAAATAAAGCATACAAGTCAGTTTTTGTGAGGCATGAGCGCCGCCAAACTTAA
TCCTGACGATGTATGTTCTTATTTGGCAGCGGCAATGCAGTTTTTTGAGGGACATGTCCGGAAGACTGG
ACGAGCTATGGAATTGTGATTGCACGAAAAGGAGATAAGATCACCCACAGTTCTCTGGTGGAGATAAAAC
GTACTGATGTAGAAGGGAATTGGGCTCTGACAGGAGGCATGGAAGTACAAAGAGACCCCACTGTCCCTGA
GCATGCGTCTCTAGTCGGTCTTCTCTGAGTCTGTATAGGTTGAGCAAAATATCCGGGCAAAACACTGGT
AACTATAAGACAAACATTGCAGACAGGATAGAGCAGATTTTGGAGACAGCCCTTTTGTAAATCGTGG
AACACCATACTCTAATGACAACTCACAAAATGTGTGCTAATTGGAGTACTATACCAAACCTCAGATTTTT
GGCCGGAACCTATGACATGTTTTCTCCCGGATTGAGCATCTATATTCAGCAATCAGAGTGGGCACAGTT
GTCAGTCTTATGAAGACTGTTGAGGACTGGTATCATTTACTGGGTTCAAAAACAAATCAATCTCACCG
CTAGAGAGGCAATACTATATTTCTCCACAAGAAGTTTGGAGGAGAGATAAGAAGAATGTTTGGCCAGG
GCAGGAGACAGCTGTTCTCTACTCTTATTTTCATCCACTTCCGTTCACTAGGCTTGAGTGGGAAATCTCCT
TATTCATCAAATGCTGTTGGTCACGTGTTCAATCTCATTCACTTTGTAGGATGCTATATGGGTCAAGTCA
GATCCCTAAATGCAACGGTTATTGCTGCATGTGCTCCTCATGAAATGCTGTCTAGGGGGCTATCTGGG
AGAGGAATTCTTCGGGAAAGGGACATTTGAAAGAAGATTCTTCAGAGATGAGAAAGAACTTCAAGAATAC
GAGGCGCTGAACCTGACAAAGACTGACGTAGCACTGGCAGATGATGGAAGTGTCAACTCTGACGACGAGG
ACTACTTTTCAGGTGAAACCAGAAGTCCGGAGGCTGTTTATCTCGAATCATGATGAATGGAGTCTGACT
AAAGAGATCTCACATACGGAGATATGTCTCAGTCAGTTCCAATCATCAAGCCCGTCCAAACTCATTCGCC
GAGTTTCTAAACAAGACATATTCGAGTGACTCAGGTTCCGGAGAGAACCTCTACTTCCAATCGGGATCCG
GTAGCCATGGCTTCCCGCCGGAGGTGGAGGAGCAGGATGATGGCACGCTGCCCATGTCTTGTGCCCAGGA
GAGCGGGATGGACCGTCAACCTGCAGCCTGTGCTTCTGCTAGGATCAATGTGTAAGAAGTTGAATAACAA
AATGCCGGAATCTACGGATTGTGTATATCCATCATGAAAAAACTAACACCCCTCCTTTTGAACCATCC
CAACATGAGCAAGATCTTTGTCAATCCTAGTGCTATTAGAGCCGGTCTGGCCGATCTTGAGATGGCTGA
AGAACTGTTGATCTGATCAATAGAAATATCGAAGACAATCAGGCTCATCTCCAAGGGGAACCCATAGAG
GTGGACAATCTCCCTGAGGATATGGGGCGACTTCACCTGGATGATGGAAAATCGCCCAACCATGGTGAGA
TAGCCAAGGTGGGAGAAGGCAAGTATCGAGAGGACTTTTCAGATGGATGAAGGAGAGGATCCTAGCTTCCT
GTTCCAGTCATACCTGGAAAATGTTGGAGTCCAAATAGTCAGACAAATGAGGTGAGGAGAGAGATTTCTC
AAGATATGGTCACAGACCGTAGAAGAGATTATATCCTATGTCGCGGTCAACTTTCCCAACCTCCAGGAA
AGTCTTCAGAGGATAAATCAACCCAGACTACTGGCCGAGAGCTCAAGAAGGAGACAACACCCACTCCTTC
TCAGAGAGAAAGCCAATCATCGAAAGCCAGGATGGCGGCTCAAATGCTTCTGGCCCTCCAGCCCTTGAA
TGGTCGGCTACCAATGAAGAGGATGATCTATCAGTGGAGGCTGAGATCGCTCACCAGATTGCAGAAAGTT
TCTCCAAAAAATATAAGTTTCCCTCTCGATCCTCAGGGATACTCTTGATAAATTTGAGCAATTGAAAAT
GAACCTTGATGATATAGTTAAAGAGGCAAAAAATGTACCAGGTGTGACCCGTTTAGCCCATGACGGGTCC
AACTCCCCCTAAGATGTGTACTGGGATGGGTGCTTTGGCCAACTTAAGAAATTCAGTTGTTAGTCG
AATCCGACAAGCTGAGTAAATCATGCAAGATGACTTGAATCGTATACATCTTGCTAACCGAACCTCTC
CCCTCAGTCCCTCTAGACAATAAAATCCGAGATGTCCAAAG

图11(续)

TCAACATGAAAAAACAGGCAACACCACTGATAAATCGATGAACCTCCTACGTAAGATAGTGAAAAACCG
CAGGGACGAGGACACTCAAAAATCCTCTCCCGCGTCAGCCCCTCTGGATGACGATGACTTGTGGCTTCCA
CCCCCTGAATACGTCCCGCTGAAAGAACTTACAGGCAAGAAGAACATGAGGAACCTTTGTATCAACGGAA
GGGTTAAAGTGTGTAGCCCGAATGGTTACTCGTTTCAGGATCCTGCGGCACATTCTGAAATCATTCGACGA
GATATATTCTGGGAATCATAGGATGATCGGGTTAGTCAAAGTGGTTATTGGACTGGCTTTGTTCAGGATCT
CCAGTCCCTGAGGGCCTGAACTGGGTATACAAATTGAGGAGAACCTTTATCTTCCAGTGGGCTGATTCCA
GGGGCCCTCTTGAAGGGGAGGAGTTGGAATACTCTCAGGAGATCACTTGGGATGATGATACTGAGTTCGT
CGGATTGCAAATAAGAGTGATTGCAAAACAGTGTCATATCCAGGGCAGAGTCTGGTGTATCAACATGAAC
CCGAGAGCATGTCAACTATGGTCTGACATGTCTCTTCAGACACAAAGGTCCGAAGAGGACAAAGATTCCCT
CTCTGCTTCTAGAATAACGTACGTACGATTATATCCCGCAAATTTATCACTTGTTTACCTCTGGAGGAGA
GAACATATGGGCTCAACTCCAACCCCTGGGAGCAATATAACAAAAACATGTTATGGTGCCATTAAACCG
CTGCATTTTCATCAAAGTCAAGTTGATTACCTTTACATTTTGATCCTCTTGGATGTGAAAAAACTATTAA
CATCCCTCAAAGGACCTGCAGGTACGCGGCCCGGTCACGCCACCATGGTGCCCAAGAAGAAGAGGAAAG
TCTCCAACCTGCTGACTGTGCACCAAAACCTGCCTGCCCTCCCTGTGGATGCCACCTCTGATGAAGTCAG
GAAGAACCTGATGGACATGTTTCAGGGACAGGCAGGCCTTCTCTGAACACACCTGGAAGATGCTCCTGTCT
GTGTGCAGATCCTGGGCTGCCCTGGTGCAAGCTGAACAACAGGAAATGGTTCCCTGCTGAACCTGAGGATG
TGAGGGACTACCTCCTGTACCTGCAAGCCAGAGGCCTGGCTGTGAAGACCATCCAACAGCACCTGGGCCA
GCTCAACATGCTGCACAGGAGATCTGGCCTGCCTCTGACTCCAATGCTGTGTCCCTGGTGTG
AGGAGAATCAGAAAGGAGAATCTGGATGCTGGGGAGAGAGCCAAGCAGGCCCTGGCCTTTGAACGCCTG
ACTTTGACCAAGTCAGATCCCTGATGGAGAACTCTGACAGATGCCAGGACATCAGGAACCTGGCCTTCCT
GGGCATTGCCTACAACACCCCTGCTGCGCATTGCCGAAATTGCCAGAATCAGAGTGAAGGACATCTCCCGC
ACCGATGGTGGGAGAATGCTGATCCACATTGGCAGGACCAAGACCCTGGTGTCCACAGCTGGTGTGGAGA
AGGCCCTGTCCCTGGGGGTACCAAGCTGGTGGAGAGATGGATCTCTGTGTCTGGTGTGGCTGATGACCC
CAACAACCTACCTGTTCTGCGGGTCAGAAAGAATGGTGTGGCTGCCCTTCTGCCACCTCCCAACTGTCC
ACCCGGGCCCTGGAAGGGATCTTTGAGGCCACCCACCGCCTGATCTATGGTGCCAAGGATGACTCTGGGC
AGAGATACCTGGCCTGGTCTGGCCACTCTGCCAGAGTGGGTGCTGCCAGGGACATGGCCAGGGCTGGTGT
GTCCATCCCTGAAATCATGCAGGCTGGTGGCTGGACCAATGTGAACATTGTGATGAACTACATCAGAAAC
CTGGACTCTGAGACTGGGGCCATGGTGAGGCTGCTCGAGGATGGGGACGGCAGTGGAGGATCCGGAGCCA
CGAACTTCTCTCTGTTAAAGCAAGCAGGAGACGTGGAAGAAAACCCCGGTCCTACCGGTGTGAGCAAGGG
CGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGAGGGCTCCGTGAAC
GGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCTACGAGGGCACCCAGACCGCCAAGCTGA
AGGTGACCAAGGGTGGCCCCCTGCCCTTCGCCTGGGACATCCTGTCCCCTCAGTTCATGTACGGCTCCAA
GGCCTACGTGAAGCACCCCGCCGACATCCCCGACTACTTGAAGCTGTCCTTCCCCGAGGGCTTCAAGTGG
GAGCGCGTGATGAACTTCGAGGACGGCGGCGTGGTGACCGTGACCCAGGACTCCTCCCTGCAGGACGGCG
AGTTCATCTACAAGGTGAAGCTGCGCGGCACCAACTTCCCTCCGACGGCCCCGTAATGCAGAAGAAGAC
CATGGGCTGGGAGGCCTCCTCCGAGCGGATGTACCCCGAGGACGGCGCCCTGAAGGGCGAGATCAAGCAG
AGGCTGAAGCTGAAGACGGCGGCCACTACGACGCTGAGGTCAAGACCACCTACAAGGCCAAGAAGCCCG
TGCAGCTGCCCCGGCGCTACAACGTCAACATCAAGTTGGACATCACCTCCACAACGAGGACTACACCAT
CGTGAACAGTACGAACGCGCCGAGGGCCGCCACTCCACCGCGGCATGGACGAGCTGTACAAGGGATAT
CTCAGCCATGGCTTCCCGCCGAGGTGGAGGAGCAGGATGATGGCACGCTGCCCATGTCTTGTGCCCAGG
AGAGCGGGATGGACCGTCACCTGCAGCCTGTGCTTCTGCTAGGATCAATGTGTGA

图11 (续)

CTCGAGGGCGCGCTACCCGCGGTAGCTTTTCAGTCGAGAAAAAACATTAGATCAGAAGAACAACCTGGC
AACACTTTCTCAACCTGAGACTTACTTCAAGATGCTCGATCCTGGAGAGGTCTATGATGACCCTATTGAC
CCAATCGAGTTAGAGGCTGAACCCAGAGGAACCCCATTTGTCCCAACATCTTGAGGAACCTGACTACA
ATCTCAACTCTCCTTTGATAGAAGATCCTGCTAGACTAATGTTAGAATGGTTAAAAACAGGGAATAGACC
TTATCGGATGACTCTAACAGACAATTGCTCCAGGTCTTTTCAGAGTTTTGAAAGATTATTTCAAGAAGGTA
GATTTGGGTTCTCTCAAGGTGGGCGGAATGGCTGCACAGTCAATGATTTCTCTCTGGTTATATGGTGCCC
ACTCTGAATCCAACAGGAGCCGGAGATGTATAACAGACTTGGCCCATTCTATTCCAAGTCGTCCCCCAT
AGAGAAGCTGTTGAATCTCACGCTAGGAAATAGAGGGCTGAGAATCCCCCAGAGGGAGTGTTAAGTTGC
CTTGAGAGGGTTGATTATGATAATGCATTGGAAGGTATCTTGCCAACACGTATTCCTCTTACTTGTCTCT
TCCATGTAATCACCTTATACATGAACGCCCTAGACTGGGATGAAGAAAAGACCATCCTAGCATTATGGAA
AGATTTAACCTCAGTGGACATCGGGAAGGACTTGGTAAAGTTCAAAGACCAAATATGGGGACTGCTGATC
GTGACAAAGGACTTTGTTTACTCCCAAAGTCCAATTGTCTTTTTGACAGAACTACACACTTATGCTAA
AAGATCTTTCTTGCTCGCTTCAACTCCTTAATGGTCTTGCTCTCTCCCCCAGAGCCCCGATACTCAGA
TGACTTGATATCTCAACTATGCCAGCTGTACATTGCTGGGGATCAAGTCTTGCTATGTGTGGAACTCC
GGCTATGAAGTCATCAAAATATTGGAGCCATATGTCGTGAATAGTTTAGTCCAGAGAGCAGAAAAGTTTA
GGCCTCTCATTCACTTCTTGGGAGACTTTCTGTATTTATAAAAGACAAGGTAAGTCAACTTGAAGAGAC
GTTTCGGTCCCTGTGCAAGAAGTTCTTTAGGGCTCTGGATCAATTCGACAACATACATGACTTGGTTTTT
GTGTTTTGGCTGTTACAGGCATTGGGGGCACCCATATAGATTATCGAAAGGGTCTGTCAAAACTATATG
ATCAGGTTACCTTAAAAAATGATAGATAAGTCTTACCAGGAGTGCTTAGCAAGCGACCTAGCCAGGAG
GATCCTTAGATGGGGTTTTGATAAGTACTCCAAGTGGTATCTGGATTCAAGATTCTAGCCCCGAGACCAC
CCCTTGACTCCTTATATCAAAACCCAAACATGGCCACCCAAACATATTGTAGACTTGGTGGGGGATACAT
GGCACAAGCTCCCGATCACGCAGATCTTTGAGATTCTGAATCAATGGATCCGTGAGAAATATTGGATGA
CAAATCACATTCTTTCACCAGAACGAGACTAGCTTCTTGGCTGTGAGAAAACCGAGGGGGGCTGTTCCT
AGCGAAAAAGTTATTATCACGGCCCTGTCTAAGCCGCCTGTCAATCCCCGAGAGTTTCTGAGGTCTATAG
ACCTCGGAGGATTGCCAGATGAAGACTTGATAATTGGCCTCAAGCCAAAGGAACGGGAATTGAAGATTGA
AGGTCGATCTTTGCTCTAATGTCTATGGAATCTAAGATTGTATTTGTCTACTGAAAAACTCTTGGCC
AACTACATCTTGCCACTTTTTGACGCGCTGACTATGACAGACAACCTGAACAAGGTGTTTAAAAAGCTGA
TCGACAGGGTCACCGGGCAAGGGCTTTTGGACTATTCAAGGGTACATATGCAATTCACCTGGACTATGA
AAAGTGGAACAACCATCAAAGATTAGAGTCAACAGAGGATGTATTTTCTGTCTAGATCAAGTGTGTTGGA
TTGAAGAGAGTGTTTTCTAGAACACACGAGTTTTTTCAAAGGCCCTGGATCTATTATTCAGACAGATCAG
ACCTCATCGGGTTACGGGAGGATCAATATACTGCTTAGATGCGTCCAACGGCCCAACCTGTTGGAATGG
CCAGGATGGCGGGCTAGAAGGCTTACGGCAGAAGGGCTGGAGTCTAGTCAGCTTATTGATGATAGATAGA
GAATCTCAAAATCAGGAACACAAGAACCAAAATACTAGCTCAAGGAGACAACAGGTTTTATGTCCGACAT
ACATGTTGTCGCCAGGGCTATCTCAAGAGGGGCTCCTCTATGAATTGGAGAGAATATCAAGGAATGCACT
TTCGATATACAGAGCCGTCGAGGAAGGGGCATCTAAGCTAGGGCTGATCATCAAGAAAGAAGAGACCATG
TGAGTTATGACTTCTCTATCTATGGAAAAACCCCTTTGTTTAGAGGTAACATATTGGTGCCTGAGTCCA
AAAGATGGGCCAGAGTCTCTTGCGTCTCTAATGACCAATAGTCAACCTCGCCAATATAATGTCGACAGT
GTCCACCAATGCGCTAACAGTGGCACAACTCTCAATCTTTGATCAAAACCGATGAGGGATTTTCTGTCT
ATGTCAGTACAGGCAGTCTTTCACTACCTGCTATTTAGCCCAATCTTAAAGGGAAGAGTTTACAAGATTC
TGAGCGCTGAAGGGGAGAGCTTTCTCCTAGCCATGTCAAGGATAATCTATCTAGATCCTTCTTTGGGAGG
GATATCTGGAATGTCCCTCGGAAGATTCCATATACGACAGTTCTCAGACCCTGTCTCTGAAGGGTTATCC
TTCTGGAGAGAGATCTGGTTAAGCTCCCAAGAGTCTTGATTACGCGTTGTGTCAAGAGGCTGGAAACC
CAGATCTTGAGAGAGAACTCGAGAGCTTCACTCGCCTTCTAGAAGATCCGACCACCTTAAATATCAG
AGGAGGGGCCAGTCTTACCATTCTACTCAAGGATGCAATCAGAAAGGCTTTATATGACGAGGTGGACAAG
GTGGAATAATCAGAGTTTCGAGAGGCAATCCTGTTGTCCAAGACCCATAGAGATAATTTTATACTCTTCT
TAATATCTGTTGAGCCTCTGTTTCCTCGATTTCTCAGTGAGCTATTCAAGTTCGTCTTTTTTGGGAATCCC
CGAGTCAATCATTGGATTGATACAAAACCTCCCGAACGATAAGAAGGCAGTTTAGAAAGAGTCTCTCAAAA
ACTTTAGAAGAATCCTTCTACAACCTCAGAGATCCACGGGATTAGTCGGATGACCCAGACACCTCAGAGGG
TTGGGGGGGTGTGGCCTTGCTCTTCAGAGAGGGCAGATCTACTTAGGGAGATCTCTTGGGGAAGAAAAGT
GGTAGGCACGACAGTTCCTCACCTTCTGAGATGTTGGGATTACTTCCAAGTCCTCTATTTCTTGCACT
TGTGGAGCAACAGGAGGAGGCAATCCTAGAGTTTCT

图11 (续)

GTATCAGTACTCCCGTCTTTGATCAGTCATTTTTTTTACGAGGCCCCCTAAAGGGATACTTGGGCTCGT
CCACCTCTATGTCGACCCAGCTATTCATGCATGGGAAAAAGTCACTAATGTTTCATGTGGTGAAGAGAGC
TCTATCGTTAAAAGAATCTATAAACTGGTTCATTACTAGAGATTCCAACCTTGGCTCAAGCTCTAATTAGG
AACATTATGTCTCTGACAGGCCCTGATTTCCCTCTAGAGGAGGCCCTGTCTTCAAAGGACGGGGTCAG
CCTTGCATAGGTTCAAGTCTGCCAGATACAGCGAAGGAGGGTATTCTTCTGTCTGCCCCGAACCTCCTCTC
TCATATTTCTGTTAGTACAGACACCATGTCTGATTTGACCAAGACGGGAAGAACTACGATTTTCATGTTT
CAGCCATTGATGCTTTATGCACAGACATGGACATCAGAGCTGGTACAGAGAGACACAAGGCTAAGAGACT
CTACGTTTCATTGGCACCTCCGATGCAACAGGTGTGTGAGACCCATTGACGACGTGACCCCTGGAGACCTC
TCAGATCTTCGAGTTTCCGGATGTGTGCAAAAAGAATATCCAGAATGGTTTCTGGGGCTGTGCCTCACTTC
CAGAGGCTTCCCGATATCCGTCTGAGACCAGGAGATTTTGAATCTCTAAGCGGTAGAGAAAAGTCTCACC
ATATCGGATCAGCTCAGGGGCTCTTATACTCAATCTTAGTGGCAATTCACGACTCAGGATACAATGATGG
AACCATCTTCCCTGTCAACATATACGGCAAGGTTTCCCCTAGAGACTATTTGAGAGGGCTCGCAAGGGGA
GTATTGATAGGATCCTCGATTTGCTTCTTGACAAGAATGACAAATATCAATATTAATAGACCTCTTGAAT
TGGTCTCAGGGGTAATCTCATATATTCTCCTGAGGCTAGATAACCATCCCTCCTTGTACATAATGCTCAG
AGAACCGTCTCTTAGAGGAGAGATATTTCTATCCCTCAGAAAATCCCCGCCGCTTATCCAACCACTATG
AAGAAGGCAACAGATCAATCTTGTGTATCTCCAACATGTGCTACGCTATGAGCGAGAGATTAATCAGCG
CGTCTCCAGAGAATGACTGGCTATGGATCTTTTCAGACTTTTGAAGTGCCAAAATGACGTACCTATGCCCT
CATTACTTACCAGTCTCATCTTCTACTCCAGAGGGTTGAGAGAAACCTATCTAAGAGTATGAGAGATAAC
CTGCGACAATTGAGTTCTTTGATGAGGCAGGTGCTGGGCGGGCACGGAGAAGATACCTTAGAGTCAGACG
ACAACATTCAACGACTGCTAAAAGACTCTTTACGAAGGACAAGATGGGTGGATCAAGAGGTGCGCCATGC
AGCTAGAACCATGACTGGAGATTACAGCCCCAACAGAAGGTGTCCGTAAGGTAGGATGTTCAGAATGG
GTCTGCTCTGCTCAACAGGTTGCAGTCTCTACCTCAGCAAACCCGCCCTGTCTCGGAGCTTGACATAA
GGGCCCTCTCTAAGAGGTTCCAGAACCCCTTTGATCTCGGGCTTGAGAGTGGTTTCAGTGGGCAACCGGTGC
TCATTATAAGCTTAAGCCTATTCTAGATGATCTCAATGTTTTCCCATCTCTCTGCCTTGTAGTTGGGGAC
GGGTGAGGGGGGATATCAAGGGCAGTCTCAACATGTTTTCAGATGCCAAGCTTGTGTTCAACAGTCTTT
TAGAGGTGAACTGATGGCTTCCGGAACACATCCACTGCCTCCTCAGCAATCATGAGGGGAGGAAA
TGATATCGTCTCCAGAGTGATAGATCTTGACTCAATCTGGGAAAAACCGTCCGACTTGAGAAACTTGGCA
ACCTGGAAATACTTCCAGTCAGTCCAAAAGCAGGTCAACATGTCTATGACCTCATTATTTGCGATGCAG
AAGTTACTGACATTGCATCTATCAACCGGATCACCCTGTTAATGTCCGATTTTGCAATTGTCTATAGATGG
ACCACTCTATTTGGTCTTCAAACCTTATGGGACTATGCTAGTAAATCCAACTACAAGGCTATTCAACAC
CTGTCAAGAGCGTTCCCTCGGTACAGGGTTTATCACCCAAGTAACTTCGTCTTTTTCATCTGAGCTCT
ACCTCCGATTCTCCAAACGAGGGAAGTTTTTCAGAGATGCTGAGTACTTGACCTCTTCCACCCTTCGAGA
AATGAGCCTTGTGTTATTCAATTGTAGCAGCCCCAAGAGTGAGATGCAGAGAGCTCGTTTCCTTGAACAT
CAGGATCTTGTGAGAGGATTTCCCTGAAGAAATCATATCAAATCCTTACAATGAGATGATCATAACTCTGA
TTGACAGTGATGTAGAATCTTTCTAGTCCACAAGATGGTTGATGATCTTGAGTTACAGAGGGGAACTCT
GTCTAAAGTGGCTATCATTATAGCCATCATGATAGTTTTCTCCAACAGAGTCTTCAACGTTTTCCAAACCC
CTAACTGACCCCTCGTTCTATCCACCGTCTGATCCAAAATCCTGAGGCACTTCAACATATGTTGCAGTA
CTATGATGTATCTATCTACTGCTTTAGGTGACGTCCCTAGCTTCGCAAGACTTCACGACCTGTATAACAG
ACCTATAACTTATTACTTCAGAAAGCAAGTCATTGAGGGGAACGTTTATCTATCTTGGAGTTGGTCCAAC
GACACCTCAGTGTTCAAAGGGTAGCCTGTAATCTAGCCTGAGTCTGTCTCATCTCACTGGATCAGGTTGA
TTTACAAGATAGTGAAGACTACCAGACTCGTTGGCAGCATCAAGGATCTATCCAGAGAAGTGGAAGACA
CCTTCATAGGTACAACAGGTGGATCACCTAGAGGATATCAGATCTAGATCATCCCTACTAGACTACAGT
TGCTGTGAACCGGATACTCCTGGAAGCCTGCCCATGCTAAGACTCTTGTGT

图11 (续)

GATGTATCTTGAAAAAACAAGATCACCGGATACTCCTGGAAGCCTGCCATGCTAAGACTCTTGTGTGA
TGTATCTTGAAAAAACAAGATCCTAAATCTGAACCTTTGGTTGTTTGATTGTTTTCTCATTTTTGTTG
TTTATTTGTTAAGCGTGGGTGCGCATGGCATCTCCACCTCCTCGCGGTCCGACCTGGGCATCCGAAGGAG
GACGCACGTCCACTCGGATGGCTAAGGGAGAAGGGCGAATTCCAGCACACTGGCGGCCGTTACTAGTGA
TCCGAGCTCGGTACCAAGCTTAAGTTTAAACCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCA
TCTGTTGTTTGCCCCCTCCCCCGTGCCCTTCCTTGACCCTGGAAGGTGCCACTCCCACCTGTCCTTTCTAAT
AAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCTATTCTATTTCTGGGGGGTGGGGTGGGCAGGA
CAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAG
GCGGAAAGAACAGCTGGGGCTCTAGGGGTATCCCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGG
GTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTTCGCTTCTCT
CCCTTCCTTTCTCGCCACGTTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTC
CGATTTAGTGCTTTACGGCACCTCGACCCCAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCAT
CGCCCTGATAGACGGTTTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTC
AACTGGAACAACACTCAACCTATCTCGGTCTATTCTTTTGATTATAAGGGATTTTGGCGATTTCCGGC
TATTGGTTAAAAAATGAGCTGATTTAACAATAAATTAACGCGAATTAATTCTGTGGAATGTGTGTGAGTT
AGGGTGTGGAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCA
ACCAGGTGTGGAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAG
CAACCATAGTCCCGCCCTAACTCCGCCATCCCGCCCTAACTCCGCCAGTTCGCCCATTTCTCCGCC
CCATGGCTGACTAATTTTTTTTTATTTATGAGAGGCCGAGGCCGCTCTGCCTCTGAGCTATTCCAGAAG
TAGTGAGGAGCTTTTTTGGAGGCTTAGGCTTTTGCAAAAGCTCCCGGGAGCTTGTATATCCATTTTCG
GATCTGATCAAGAGACAGGATGAGGATCGTTTCGCATGATTGAACAAGATGGATTGCACGCAGGTTCTCC
GGCCGCTTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCC
GTGTTCCGGCTGTACGCGCAGGGGCGCCGGTTCTTTTTGTCAGACCGACCTGTCCGGTGCCCTGAATG
AACTGCAGGACGAGGCAGCGCGCTATCGTGCTGGCTGGCCACGACGGGCTTCCTTGCGCAGCTGTCTCGA
CGTTGTCACTGAAGCGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCTCT
CACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGG
CTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTAATCGGATGGAAGCCGGTCT
TGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAAGTGTTCGCCAGGCTCAAG
GCGCGCATGCCCGACGCGGAGGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGG
AAAATGGCCGCTTTTCTGGATTATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGC
GTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCTGCTGCTTTACGGT
ATCGCCGCTCCCGATTGCGAGCGCATCGCTTCTATCGCTTCTTTGACGAGTTCTTCTGAGCGGGACTCT
GGGGTTTCGAAATGACCGACCAAGCGACGCCCAACCTGCCATCACGAGATTTGATTCCACCGCCGCTTTC
TATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCA
TGCTGGAGTTCTTCGCCACCCCAACTTGTTTATTGACGCTTATAATGGTTACAAATAAAGCAATAGCAT
CACAAATTTACAAATAAAGCATTTTTTCACTGCATTCTAGTTGTGGTTGTCCAAACTCATCAATGTA
TCTTATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTG
TGTGAAATTGTTATCCGCTCACAAATCCACACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGG
TGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCAGTCGGGAAACCTG
TCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGCGGTTTGCGTATTGGGCGCTCTTCG
CTTCTCGCTCACTGACTCGCTGCGCTCGGTGCTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGC
GGTAATACGGTTATCCACAGAATCAGGGGATAACGAGGAAGAACATGTGAGCAAAAGGCCAGCAAAAG
GCCAGGAACCGTAAAAAGGCCGCTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACA
AAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTTCCCTGG
AAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCG
GGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGGTGTAGGTCGTTTCGCTCCAAGC
TGGGCTGTGTGCACGAACCCCCGTTACGCCGACCGCTGCGCCTTATCCGGTAACATATCGTCTTGAGTC
CAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTAT
GTAGGCGGTGTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTGGTA
TCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGTGGTAGCTCTTGATCCGGCAAACAAACCAC
CGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCA

图11 (续)

GATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGG
AACGAAAACACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAA
ATTAAAAATGAAGTTTAAATCAATCTAAAGTATATATGAGTAACTTGGTCTGACAGTTACCAATGCTT
AATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGACTCCCCGTCGTG
TAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCT
CACCGGCTCCAGATTTATCAGCAATAAACAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCTTGCAAC
TTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGT
TTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTCGTTTGGTATGGCTTCATTCA
GCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTT
CGGTCCCTCCGATCGTTGTGAGAAGTAAGTTGGCCGCGAGTGTTATCACTCATGGTTATGGCAGCACTGCAT
AATTCTCTTACTGTTCATGCCATCCGTAAGATGCTTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCT
GAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAG
CAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACCTCAAGGATCTTACCGCTG
TTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCG
TTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAGGGAATAAGGGCGACACGGAATGTTG
AATACTCATACTCTTCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATAC
ATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTG
ACGTCGACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACCTCTCAGTACAATCTGCTCTGATGCCGCA
TAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAG
CTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCT
TCGCGATGTACGGGCCAGATATACGCGT

图11(续)

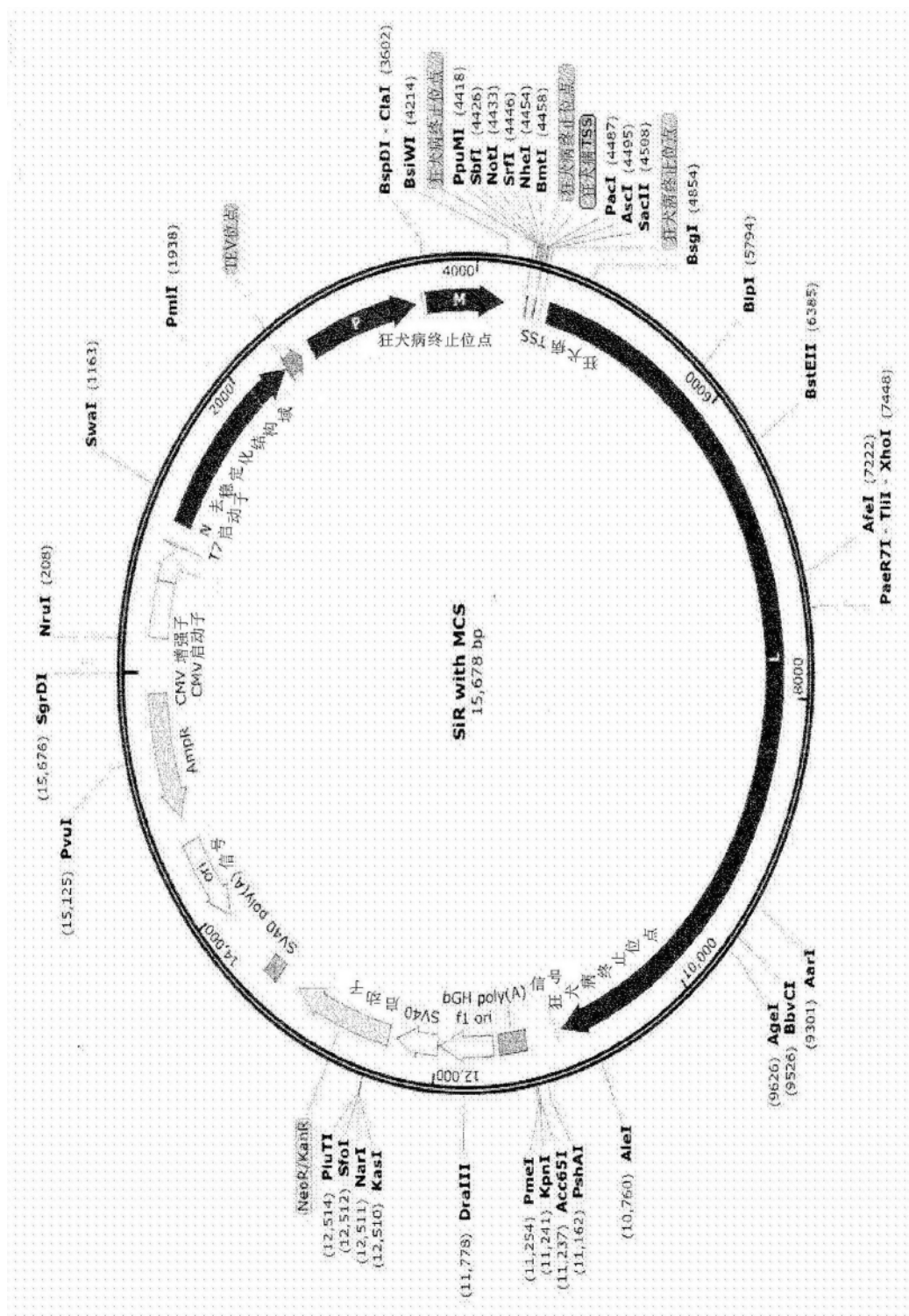


图12

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACCTCTCAGTACAATCTGCTCTGATGCCGCATAGTT
AAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACA
ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTCGCTGCTTCGCG
ATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTC
ATTAGTTCATAGCCCATATATGGAGTTCGCGCTTACATAACTTACGCTAAATGGCCCGCTGGCTGACCG
CCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCC
ATTGACGTCAATCGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCC
AAGTACGCCCCCTATTGACGTCAATCACGGTAAATGGCCCGCTGGCATTATGCCAGTACATGACCTTA
TGGGACTTTTCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGC
AGTACATCAATGGGCGTGGATAGCGGTTTACTCACGGGGATTTCGAAGTCTCCACCCCATGACGTCAA
TGGGAGTTTGTGTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCATGACG
CAAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCA
CTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCCAAGCTGGCTAGATTAAGCGTCT
GATGAGTCCGTGAGGACGAAACCCGGAGTCCCGGGTCacgcttaacaaccagatcaaaagaaaaaacagac
attgtcaattgcaaagcaaaaatgtaacacccctacaatggatgccgacaagattgtattcaaaagtcaat
aatcagggtggtctctttgaagcctgagattatcgtggatcaatatgagtacaagtaacctgccatcaaaag
atttgaaaaagccctgtataaccctaggaaaggctcccgatttaataaaagcatacaagtcagttttgtc
aggcatgagcgccgcaaaacttaactcctgacgatgtatgttctctatttggcagcggaatgcagttttt
gaggggacatgtccggaagactggaccagctatggaattgtgattgcacgaaaaggagataagatcacc
cagggttctctggtggagataaaaactgactgatgtagaagggaattgggctctgacaggaggcatggaact
gacaagagacccactgtccctgagcatgcgtccttagtcggtcttctctttagttctgtataggttgagc
aaaatatccgggcaaaacactggttaactataagacaaacattgcagacaggatagagcagatttttgaga
cagccccctttgttaaaatcgtggaacaccatactctaatagacaactcacaatatgtgtgctaattggag
tactataccaaacttcagatttttggccggaacctatgacatgtttttctcccgattgagcatctatat
tcagcaatcagagtgggcacagttgtcactgcttatgaagactgttcaggactggtatcatttactgggt
tcataaaacaaatcaatctcaccgctagagaggcaatactatatattcttccacaagaactttgaggaaga
gataagaagaatgtttgagccagggcaggagacagctgttctcactcttatttcatccacttccgttca
ctaggcttgagtgggaaaatctccttattcatcaaatgctgttggtcacgtgttcaatctcattcatttg
taggatgctatatgggtcaagtcagatccctaaatgcaacgggtattgtctgcatgtctcctcatgaaat
gtctgttctagggggcttatctgggagaggaattcttcgggaaagggaatttgaaagaagattcttcaga
gatgagaaagaacttcaagaatacagggcggtgaaactgacaaagactgacgtagcactggcagatgatg
gaactgtcaactctgacgacgaggactacttttcagggtgaaaccagaagtcggagggtgtttatactcg
aatcatgatgaatggaggtcgactaaagagatctcacatacggagatatgtctcagtcagttccaatcat
caagcccggtccaaactcattcggcgagtttctaaacaagacatattcgagtgactcaGGTTCCGGAGaga
acctctacttccaatcgGGATCCGGTAGCCATGGCTTCCCGCCGGAGCTGGAGGAGCAGGATGATGGCAC
GCTGCCCATGTCTTGTGCCAGGAGAGCGGGATGGACCGTCACCCTGCAGCCTGTGCTTCTGCTAGGATC
AATGTGtaagaagttgaataacaaaatgccggaatctacggattgtgtatatccatcatgaaaaaact
aacacccctcctttcgaaacctcccaaacatgagcaagatccttgtcaatcctagtgtctattagagccgg
tctggccgatcttgagatggctgaagaaaactgttgatctgatcaatagaaatatcgaagacaatcaggct
catctccaaggggaacctatagaggtggacaatctcctgaggataggggcgacttcacctggatgatg
gaaaatcgcccaacctggtgagatagccaaggtgggagaaggcaagtatogagaggactttcagatgga
tgaaggagaggatcctagcttctgttccagtcatacctggaaaatgttgaggtccaaatagtcagacaa
atgaggtcaggagagagatttctcaagatatggtcacagaccgtagaagagattatatcctatgtcggg
tcaactttcccaacctccaggaaaagtcttcagaggataaatcaaccagactactggccgagagctcaa
gaaggagacaacacccactccttctcagagagaaagccaatcatcgaaagccaggatggcggctcaaatt
gcttctggccctccagccctgaatggtcggctaccaatgaagaggatgatctatcagtgagggtgaga
tcgctcaccagattgcagaaagtcttccaaaaaatataagtttccctctcgatcctcagggaactctt
gtataattttgagcaattgaaaatgaaccttgatagatagtttaagaggcaaaaaatgtaccagggtg
accggttagcccatgacggtccaaactccccataagatgtgtactgggatgggtcgctttggccaact
ctaagaaattccagttgttagtcgaatccgacaag

图13

ctgagtaaaatcatgcaagatgacttgaatcgctatacatcttgcataacggaacctctcccctcagtcoc
tctagacaataaaaatccgagatgtcccaaagtcaacatgaaaaaacaggcaacaccactgataaaTCGa
tgaacctcctacgttaagatagtgaaaaaccgcagggacgaggacactcaaaaatcctctcccggtcagc
ccctctggatgacgatgacttgtggttccacccccctgaatacgtcccgctgaaagaacttacaggcaag
aagaacatgaggaacttttgtatcaacggaagggttaaagtgtgtagcccggaatgggttaactcgttcagga
tctcgggcacattctgaaatcattcgacgagatataattctgggaatcataggatgatcgggttagtcaa
agtgggttattggactggctttgtcaggatctccagtcctgaggcctgaactgggtatacaaaattgagg
agaacctttatcttccagtggtgattccaggggccctcttgaagggaggagttggaatactctcagg
agatcacttgggatgatgatactgagttcgtcggattgcaaataagagtgattgcaaacagtgatcatat
ccagggcagagtgctggtgtatcaacatgaacccgagagcatgtcaactatggtctgacatgtctcttcag
acacaaaggtccgaaggagacaaagattcctctctgcttctagaataaogtacgtcagatttatatccgc
aaatttatcacttgtttacctctggaggagagaacatattgggtcaactccaaccttgggagcaatata
acaaaaaacatggttattggtgccattaaaccgctgcatttcatcaaaagtcaggttgattacctttacattt
tgatcctcttggatgtgaaaaaaactattaacatccctcaagggaCTGCAGGTACGCGGCCCTACGCC
CGGGCTACGCTAGCatgaaaaaaactaacacccctctTTAATTAATACGGCGCGCCTaccgcggttagct
tttcagtcgagaaaaaacattagatcagaagaacaactggcaacactttctcaacctgagaacttaactc
aagatgctcgatcctggagaggtctatgatgacctattgacccaatcgaggttagaggctgaacccagag
gaacccccattgtcccaacatcttgaggaaactctgactacaatctcaactctccttggatagaagatcc
tgctagactaatgttagaatgggttaaaaacagggaatagaccttatcggtgactctaacagacaattgc
tccaggtctttcagagttttgaaagattatttcaagaaggtagatttgggttctctcaagggtggcgga
tggtctgcacagtcaatgatttctctggttatattggtgccactctgaatccaacaggagccggagatg
tataacagacttggccatttctattccaagtctgcccccatagagaagctgttgaaatctacgctagga
aatagagggtgagaaatccccccagaggaggtgttaagttgccttgagagggttgattatgataatgcat
ttggaaggtatcttgccaacacgtattcctcttacttgttcttccatgtaataccttatcatatgaacgc
cctagactgggatgaagaaaagaccatcctagcattatggaagatttaacctcagtgacatcggaag
gaacttggttaaagttcaagaccaaataatggggactgctgatcgtgacaaaggacttgtttactccaaa
gttccaattgtcttttgacagaaactacacacttatgctaaaagatcttttcttgtctcgcttcaactc
cttaattggtcttctctctcccccagagccccgatactcagatgacttgatatctcaactatgccagctg
tacattgctggggatcaagtcttgtctatgtgtggaactccggctatgaagtcatacaaatattggagc
catatgtcgtgaatagtttagtccagagagcagaaaaagtttaggcctctcattcattccttgggagactt
tctgtattttataaaagacaaggttaagtcaacttgaagagacgttcgggtccctgtgcaagaaggttcttt
agggctctggatcaattcgacaacatacatgacttgggtttttgtgtttggctgttacaggcattgggggc
acccatataatagattatcgaaagggtctgtcaaaactatatgatcagggttcaccttaaaaaaatgataga
taagtcctaccaggagtgcttagcaagcgacctagccaggaggatccttagatggggttttgataagtac
tccaagtggtatctggattcaagattcctagcccgagaccaccttgactccttatatcaaaacccaaa
catggccacccaaaacatattgttagacttgggtgggggatacatggcacaagctcccgatcacgcagatctt
tgagattcctgaatcaatggatccgtcagaaaatattggatgacaaatcacattctttcaccagaacgaga
ctagcttcttggctgtcagaaaaaccgagggggcctgttccctagcgaaaaagttattatcacggccctgt
ctaagccgcctgtcaatccccgagagtttctgaggtctatagacctcgaggattgccagatgaagactt
gataattggcctcaagccaaaggaacgggaattgaagattgaaggtcgattctttgtctaatgtcatgg
aatctaagattgtattttgtcatcactgaaaaactcttggccaactacatcttggcaactttttgacgcgc
tgactatgacagacaacctgaacaaggtgtttaaaaagctgatcgacagggtcacccgggcaagggtttt
ggactattcaagggtcacatatgcatttcacctggactatgaaaagtggaaacaacctcaagatttagag
tcaacagaggatgtattttctgtctagatcaagtggttggattgaagagagtggtttctagaacacacg
agttttttcaaaaaggcctggatctattattcagacagatcagacctcatcggttacgggaggtcaaat
atactgcttagatgcgtccaacggcccaacctgttgggaatggccaggatggcggttagaaggcttacgg
cagaagggtggagctagtcagcttattgatgatagatagagaatctcaaatcaggaaacacaagaacca
aaatactagctcaaggagacaaccaggttttatgtccgacatacatgttgtcgccagggttatctcaaga
ggggctcctctatgaattggagagaatatcaaggaatgcactttcgatatacagagccgtcgagggaaggg
gcattcaagctagggctgatcatcaagaaagaagagaccatgtgtagttatgacttctcatctatggaa
aaacccctttgttttagaggtaacatattgggtgcctgagtcacaaaagatgggacagagtccttgcgtctc

图13(续)

图13 (续)

ttgcagtactatgatgtatctatctactgctttaggtgacgtccctagcttcgcaagacttcacgacctg
tataacagacctataaacttattacttcagaaagcaagtcattcgagggaacggtttatctatcttggagtt
ggtccaacgacacctcagtggttcaaaagggtagcctgtaattctagcctgagtcctcatctcactggat
caggttgatttacaagatagtgaagactaccagactcgttggcagcatcaaggatctatccagagaagtg
gaaagacaccttcataaggtacaacaggtggatcaccttagaggatatcagatctagatcatccctactag
actacagttgcctgtgaacccggatactcctggaagcctgccatgctaagactcttgtgtgatgtatctt
gaaaaaaacaagatcctaatactgaacctttggttggttgattgtttttctcatttttggttgtttatttg
ttaagcgtGGGTGGCATGGCATCTCCACCTCCTCGCGGTCCGACCTGGGCATCCGAAGGAGGACGCACG
TCCACTCGGATGGCTAAGGGAGAAGGGCGAATTCCAGCACACTGGCGGCCGTTACTAGTGGATCCGAGCT
CGGTACCAAGCTTAAGTTTAAACCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTGT
TTGCCCCCTCCCCGTGCGCTTCTTGACCTGGAAGGTGCCACTCCCACTGTCTTTCTTAATAAAATGAG
GAAATTGCAATCGCATTTGCTGAGTAGGTGTCTATTCTTGGGGGTGGGGTGGGGCAGGACAGCAAGG
GGGAGGATTGGGAAGACAATAGCAGGCATGTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAG
AACCAGCTGGGGCTCTAGGGGGTATCCCCACGGCCCTGTAGCGGCGCATTAAGCGCGGGTGTGGTG
GTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCCTTCCCTTCT
TTCTCGCCACGTTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTAG
TGCTTTACGGCACCTCGACCCCAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGA
TAGACGGTTTTTTCGCCCTTTGACGTTGGAGTCCAGGTTETTTAATAGTGGACTCTTGTTCAAACGGAA
CAACACTCAACCTATCTCGGTCTATTCTTTGATTATATAAGGGATTTTGCCGATTTCCGGCTATGGTT
AAAAATGAGCTGATTAAACAAAAATTAACCGCAATTAATTCTGTGGATGTGTGTGAGTTAGGGTGTG
GAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAGCATGCATCTCAATTAGTCAGCAACCAGGTG
TGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAGCATGCATCTCAATTAGTCAGCAACCATATA
GTCCCGCCCCCTAATCCGCCCCATCCCGCCCCCTAACTCCGCCCCAGTTCGCCCCATTCTCCGCCCCATGGCT
GACTAATTTTTTTTATTTATGCAGAGGCGAGGCGCGCTCTGCCTCTGAGCTATTCCAGAAGTAGTGAGG
AGGCTTTTTTTGGAGGCTTAGGCTTTTGCAAAAAGCTCCCGGGAGCTTGATATCCATTTTCGGATCTGAT
CAAGAGACAGGATGAGGATCGTTTCGCATGATTAACAAGATGGATTGCAGCAGGTTCTCCGGCCGCTT
GGGTGGAGAGGCTATTCCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGCTGTTCCG
GCTGTGACGCGCAGGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTGCCTGAATGAATGCGAG
GACGAGGCGAGCGCGCTATCGTGGCTGGCCACGACGGCGGTTTCTTGGCAGCTGTGCTCGACGTTGTCA
CTCAAGCGGGGAAGGACTGGCTGCTATTTGGGCGAAGTGGCGGGGCGAGGATCTCTGCTCATCTACCTGCA
TCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGGTACCTGC
CCATTCCACCACCAAGCGAAACATCGCATCGAGCGAGCAGTACTCGCATGGAAGCGGGTCTTGTGATC
AGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACGTGTTGCCAGGCTCAAGGCGCGCAT
GCCCGACGGCGAGGATCTGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGCC
CGCTTTTCTGCATTTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGCTA
CCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCGTGCTTTACGGTATCGCCGC
TCCCGATTCCGACGGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCCG
AAATGACCGACCAAGCGACGCCCCAACCTGCCATCAGGAGATTTCGATTCCACCGCCGCTTCTATGAAAG
GTTGGGCTTCGGAATCGTTTTCCGGGACGCGCGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAG
TTCTTCGCCCAGCCCAACTTGTTTATGTCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATT
TCACAAATAAAGCATTTTTTCACTGCAATCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCA
TGCTGTATACCGTCCACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCTGTGTGAAAT
TGTTATCCGCTCACAAATCCACACAACATACGAGCCGGAAGCATAAACTGTAAAGCTGGGGTGCCTAAT
GAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCTGCCA
GCTGCATTAATGAATCGGCCAACCGCGGGGAGAGGCGGTTTGGCTATTGGGCGCTCTTCCGCTTCTCTCG
CTCAGTCACTCGCTCGCTCGGTCGTTTCGGCTGCGCGAGCGGTATCAGCTCACTCAAAGCGGTATATAC
GGTTATCCACAGAATCAGGGGATAACGACGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAA
CCGTAAAAAGCGCGCTTGTGCGGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGA
CGCTCAAGTCAGAGGTGGCGAABCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAGCTCCC
TCGTGCGCTCTCTGTTCGACCTGCCGCTTACCGGATACCTGTCCGCTTCTCTCCTTCGGGAAGCGT
GGCGCTTCTCATAGCTCAGCTGTAGGTATCTCA

图13(续)

```
GTTCCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGGTTTCAGCCCGACCGCTGCCG  
CTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACT  
GGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCCTTGAAGTGGTGGCCTAACTAAG  
GCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGG  
TAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTgggtTTTTTTGTTTGCAAGCAGCAGATTACG  
CGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTGTGACGCTCAGTGGAAACGAAA  
ACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAA  
ATGAAGTTTTAAATCAATCTAAAGTATATATGACTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGT  
GAGGCACCTATCTCAGCGATCTGTCTATTTCTGTTTCATCCATAGTTGCCTGAATCCTCCCGTCTGTAGATAA  
CTACGATACGGGAGGGCTTACCATCTGGCCCGAGTGTGCAATGATACCGCGAGACCCACGCTCACCGGC  
TCCAGATTTATCAGCAATAAACCCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTIATCC  
GCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCCGCCAGTTAATAGTTTGGCGA  
ACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTCTGTTTGGTATGGCTTCATTACAGCTCCGG  
TTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCT  
CCGATCGTTGTGAGAAGTAAGTTGGCCGAGTGTATCACTCATGGTTATGCCAGCACTGCATAATTCTC  
TTACTGTCTATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAAGTACTCAACCAAGTCATTCTGAGAATA  
GTGTATGCGGGGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGGGCCACATAGCAGAACT  
TTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGAT  
CCAGTTCGATGTAACCCACTCGTGCACCCAACTCATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGG  
GTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTC  
ATACTCTTCCTTTTCAATATTATTGAAGCATTATCAGGGTTATTGTCTCATGAGCGGATACATATTTG  
AATGATTTTAGAAAAATAAACAAATAGGGGTTCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC
```

图13(续)

5' -
ACGCUUAACAAUAAACAACAAAAUGAGAAAAACAAUCAAACAACCAAAGGUUCAGAUUUAGGAUCUUG
UUUUUUUCAAGAUACAUCACACAAGAGUCUUAGCAUUGGGCAGGCUUCCAGGAGUAUCCGGUUCACAGGCA
ACUGUAGUCUAGUAGGGGAUGAUCUAGAUUCCUCUAGGGUGAUCCACCUGUUGUACCUAUGAAGG
UGUCUUUCCACUUCUCUGGAUAGAUCCUUGAUGCUGCCAACGAGUCUGGUAGUCUUCACUAUCUUGUAAA
UCAACCUGAUCCAGUGAGAUGACAGACUCAGGCUAGAAUUAACAGGCUACCCUUUUGAACACUGAGGUGUC
GUUGGACCAACUCCAAGAUAGAUAAACGUUCCUCGAAUGACUUGCUCUUGAAGUAAUAGUUAUAGGU
CUGUUUAACAGGUCGUGAAGUCUUGCGAAGCUAGGGACGUCACCUAAGCAGUAGAUAGAUACAUCAUAG
UACUGCAACAUAGUUGAAGUGCCUCAGGAUUUUGGGAUCAGACGGUGGAUAGAACGAGGGGUCAGUUG
GGGUUUGGAAACGUUGAAGACUCUGUUGGAGAAAACUAUCAUGAUGGCUAUAUAGAUAGCCACUUAGAC
AGAGUUCUUUCUGUAACUCAAGAUCAUACAACCAUCUUGUGGACUAGAAAAGAUUCUACAUCACUGUCA
UCAGAGUUAUGAUCAUCUCUAGUUAAGGAUUUGAUUGAUUUUCUUCAGGAAUCCUCUCACAAGAUCCUG
AUAGUUAAGGAACGAGCUCUCUGCAUCUCACUCUUGGGGUGCUACAAGUAAUAAACAAGGCUCAUU
UCUCGAAGGGUGGAAGAGGUCAGUACUCAGCAUCUCUGAAAAACUUCUCCUCGUUUUGGAGAAUCGGAGGU
AGAGCUCAGAUAGAAAAGACGAAGUUAUCUUGGGUGAUAAACCCUGUGACCGAGGGGAACGCUCUUGACAG
GUGUUGAAUAGCCUUGUAGUUUGGAUUUACUAGCAUAGUCCCAUAGUUUUGAAGACCAAUAGAGUGGU
CCAUCUAUAGACAAUGCAAAUCCGACAUUAAACAGGUGAUCGGUUGAUAGUAGCAUAGUACUUAACUU
CUGCAUCGCAAAUAAUGAGGUCUAGGACAUUGUAGCCUGCUUUUGGACUGACUGGAAGUAAUUCAGGU
UGCCAAGUUUCUAGUCCGACGGUUUUUCCAGAUUGAGUCAAGAUCUACUCUGGAGACCAUAUCA
UUUCCUCCCUCAUGAUUGCUGAAGGAGGAGUGGAUGUGUUCGGAAGCCAUAGGUCAUUCACCUCA
AAAGACUGUUGAACACAAGCUUGGCAUCUGGAAACAUUGUAGGACUGCCUUGAUUACCCUUGACCC
GUCCCCAACUACAAGGCAGAGAGUUGGAAAACAUUGAGAUCAUCUAGAAUAGGCUUAAAGCUUAAUAGA
GCACCGGUUGCCACUGAACCACUCUACAAGCCCGAGAUCAAAGGGUUCUGGAACCUUAGAGAGGGCCC
UUAUGUCAAGCUCGAGACAGGGGCCGGGUUUGCUGAGGUAGAGACUGCAACCUGUUGAGCAGAGCAGAC
CCAUCUGAACAUCCUACCUUACGGGACACCUUCUUGUUGGGGUGUAUUCUCCAGUCAUGGUUCUAGCU
GCAUGGCGCACCUCUUGAUCCACCAUCUUGUCCUUGUAAAGAGUCUUUAGCAGUCGUUGAAUGUUGU
CGUCUGACUCUAAGGUUUCUUCUCCGUGCCCGCCAGCACCUGCCUCAUCAAAGACUCAAUUGUCGAG
GUUAUCUCUAUACUCUUAAGAUAGGUUUCUCUCAACCCUCUGGAGUAGAAGAUAGAGACUGGUAAGUAAUG
AGGGAUAGGUACGUCAUUUUGGCACUUCUAAAGUCUGAAAAGAUCCAAGCCAGUCUUAUCUGGAGACG
CCGUGAUUAUCUCUCGCUCAUAGCGUAGCACAUGUUGGAGAUAAACAAGAUUGAUUGUUGCCUUCUUU
CAUAGUGGUUGGAUAAGCGGCGGGGAUUUUCUGAGGGAUAGAAAUAUCUCUCCUCUAAAGAGACGGUUCU
CUGAGCAUUAUGUACAAGGAGGGAUGGUUAUCUAGCCUCAGGAGAAUAUAGAGAUUACCCUGAGACCA
AUUCAAGAGGUCUUAUAAUUAUGAUUUUGUCAUUCUUGUCAAGAAGCAAUUCGAGGAUCCUAUCAAUAC
UCCCCUUGCGAGCCUCUCAAAUAGUCUCUAGGGGAAACCUUGCCGUUAUUGUUGACAGGGAAGAUUGGU
CCAUCAUUGUAUCCUGAGUCGUGAAUUGCCACUAAGAUUGAGUAUAAAGAGCCCUAGAGCUGAUCCGAUUA
GGUGAGACUUUUCUUAACCGCUUAGAGAUUCAAUUCUCCUGGUCUCAGACGGAUUACGGGAAGCCUCUG
GAAGUGAGGCACAGCCCCAGAAACAUUCUGGAUUAUUCUUUUGACACAUCCGGAACUCGAAGAUUGA
GAGGUCUCCAGGGUCACGUCGUCAAUGGGUCUCACACACCUGUUGCAUCGGAGGUGCCAAUGAAACGUAG
AGUCUCUAGCCUUGUGUCUCUCUGUACCAGCUCUGAUGUCCAUGUCUGUGCAUAAAGCAUCAAUGGCUG
GAACAUGAAAUCGUAGUUCUUCUCCGUCUUGGGUCUCAAUACAGACAUGGUGUCUGUACUAAACAGAAAUAUGA
GAGAGGAGGUUCGGGCAGACAGAAGAAUACCCUCCUUCGUGUAUCUGGCAGACUUGAACCUAUGCAAGG
CUGACCCCGUCCUUUUGAAGACAGGGGCCUCCUCUAGAGGGAAUACAGGGCCUGUCAGAGACAUAUUGU
CCUAAUAGAGCUUGAGCCAAGUUGGAAUCUCUAGUAAUGAACCAGUUUAUAGAUUCUUUUAACGAUAGA
GCUCUCUUCACCACAUGAACAUUAGUGACUUUUUCCCAUGCAUGGAAUAGCUGGGUCGACAUAAGGGUGG
ACGAGCCCAAGUAUCCCUUJAGGGGGCCUCUGUAAAAAAUAGACUGAUCAAAGGACGGGAGUACUGAUAC
AGAAACUCUAGGAUUGCCUCCUCCUGUUGCUCACAAAGUGCAAGAAUAGAGGACUUGGGAAGUAAUCCC
AACAUUCUAGAAGGGUGAGGAACUGUCGUGCCUACCACUUUUCUCCCAAGAGAUUCCCUAAGUAGAU
CUGCCUCUCUGAAGAGCAAGGCCACCCCCCAACCCUCUGAGGUGUCUGGGUCAUCCGACUAAUCCC
GUGGAUCUCUGAGUUGAAGGAUUCUUAAGUUUUUGAGAGACUCUUUUAACUGCCUUCUUAUC
GUUCGGGAGUUUUGUAUCAAUCCAUGAUUGACUGGGGAUUCUCCAAAAAGACGAACUGAAUAGCUCAC
UGAGAAAUCGAGGAAACAGAGGCUCAACAGAUUU

图14

AAGAAGAGUAUAAAAUUAUCUCUAUGGGUCUUGGACAACAGGAUUGCCUCUCGAAACUCUGAAUUUCCA
CCUUGUCCACCUCGUCUAUAAAGCCUUUCUGAUUGCAUCCUUGAGUAGAAUGGUAGGACUGGCCCCUCC
UCUGAUUUUAAAGGUGGUCGGAUCUUCUAGAAGGCGAGUGAAGCUCUCGAGUGUUCUCUCUCCAAGAUUCU
GGGUUUCAGCCUCUUGACACAACGCGUGAAUCCAGGACUCUUGGGAGCUUAACCAGAUUCUCUCUCCAGA
AGGAUAACCCUUCAGAGACAGGGUCUGAGAACUGUCGUAUUAUGGAUUCUCCGAGGGACAUUCCAGAUAU
CCUCCCAAAGAAGGAUCUAGAUAGAUUAUCCUUGACAUGGCUAGGAGAAAGCUCUCCCCUUCAGCGCUC
AGAAUCUUGUAAACUCUCCUUAAGAUUGGGCUAAAUAGCAGGUAGUGAAAGACUGCCUGUACUGACA
UGAGCAGAAAAUCCUCAUCGCUUGAUCAAAGAUUGAGAGUGUUGGCCACUGUAGCGCAUUGGUGGA
CACUGUCGACAUAUAUUGGCGAGGUUGACUAUUUGGUCAUUAAGAGACGCAAGAGACUCUGGCCCAUCUU
UUGGACUCAGGCACCAUAUUGUACCUCUAAACAAAGGGUUUUUCCAUAUGAGGGAAGUCAUAACUAC
ACAUGGUCUCUUCUUUCUUGAUGAUCAGCCCUAGCUUAGAUGCCCCUCCUCGACGGCUCUGUAUAUCGA
AAGUGCAUUCUUGAUUAUUCUCCAAUUCUAAGAGGAGCCCCUCUUGAGAUAGCCCUGGCGACAACAUG
UAUGUCGGACAUAUAAACCUGGUUGUCUCCUUGAGCUAGUAUUUUGGUUCUUGUGUCCUGAUUUUGAGAUU
CUCUAUCUAUCAUAAUAGCUGACUAGACUCCAGCCCUUCUGCCGUAAGCCUUCUAGCCCGCAUCCUG
GCCAUUCCAACAGGUUGGGCCGUUGGACGCAUCUAAGCAGUAUAUUGAUCCUCCCGUAACCCGAUGAGG
UCUGAUCUGUCUGAAUAAUAGAUCCAGGCCUUUUGAAAAAACUCUGUGUUCUAGAAAAACUCUCUUA
AUCCAAACACUUGAUCUAGGACAGAAAAUACAUCUUCUGUUGACUCUAACUUGAUGGUUGUCCACUU
UUCAUAGUCCAGGUGAAAUGCAUUGUGACCCUUGAAUAGUCCAAAAGCCUUGCCCGGUGACCCUGUCG
AUCAGCUUUUUAACACCCUUGUUCAGGUUGUCUGUCAUAGUCAGCGCGUCAAAGUGGCAAGAUUGAGU
UGGCCAAGAGUUUUUCAGUGAUGACAAAAUACAUCUUGAUUCCAUGACAUAAGAGCAAAGAAUCCGACC
UUCAAUCUUCAAUUCUCCGUUCUUGGCUUGAGGCCAAUUAUCAAGUCUUCUUGGCAUCCUCCGAGG
UCUAUAGACCUCAGAAACUCUCGGGAUUGACAGGCGGCUUAGACAGGGCCGUGAUAAUAAUUCUUUCGC
UAGGAACAGGCCCCCUCGGUUUCUGACAGCCAAGAAGCUAGUCUCGUUCUGGUGAAAGAAUGUGAUUU
GUCAUCCAUAUUUCUGACGGAUCCAUGAUUCAGGAUUCUCAAAGAUUCGCGUAUCCGGGAGCUUGUGC
CAUGUAUCCCCACCAAGUCUACAUAUUGUUUGGUGGCCAUGUUUGGUUUUGAUUAAGGAGUCAAGG
GGUGGUCUCGGGCUAGGAUUCUUGAAUCCAGAUACCACUUGGAGUACUUAUCAAACCCCAUCUAAGGAU
CCUCCUGSCUAGGUCGCUUGCUAAGCACUCCUGGUAGGACUUAUCUAUCAUUUUUUAAGGUGAACCGUA
UCAUAUAGUUUGACAGACCCUUCGAUAAUUAUAUUGGGUGCCCCCAAUGCCUGUAACAGCCAAACA
CAAAAACCAAGUCAUGUAUGUUGUCGAAUUGAUCCAGAGCCCUAAAGAACCUCUUGCACAGGGACCGAA
CGUCUCUUCAGUUGACUUAACUUGUCUUUUUAUAAUACAGGAAGUCUCCAAGGAUUGAAUGAGAGGC
CUAAACUUUCUGCUCUCUGGACUAAACUAUUCACGACAUUAUGGCUCCAAUUAUUUGAUGACUUCAUAGC
CGGAGUUUCCACACAUAGACAAGACUUGAUCCCCAGCAAUGUACAGCUGGCAUAGUUGAGAUUAUAGUC
AUCUGAGUAUCGGGGCUCUGGGGGAGAGCAAGACCAUUAAGGAGUUGAAGCGAGACAAGAAAAGAUUCU
UUUAGCAUAAGUGUGAUGUUUCUGUCAAAGAGCAAUUGGAACUUUGGAGUAAACAAAGUCCUUUGUCA
CGAUCAGCAGUCCCCAUUAUUGGUCUUUGAACUUUACCAAGUCCUCCCGAUGUCCACUGAGGUUAAUUC
UUUCCAUAUAGCUAGGAUGGUCUUUUCUUAUCCAGUCUAGGGCGUUAUGUAUAAGGUGAUUACAUGG
AAGAACAAGUAAGAGGAUACGUGUUGGCAAGAUACCUUCCAAUUGCAUUAUCAUAUACAACCCUCUCAA
GGCAACUUAACACUCCUUCUGGGGGGAUUCUAGCCCUUAUUUCCUAGCGUGAGAUUCAACAGCUUCUC
UAUGGGGACGACUUGGAAUAGAAAUGGGCCAAGUCUGUUAUACAUCUCCGGCUCUUGUUGGAUUCAGAG
UGGGCACCAUAUAACCAGAGAGAAUUAUUGACUGUGCAGCCAUUCCGCCACCUUGAGAGAACCCAAU
CUACCUUCUUGAAAUAUUCUUUCAAACUCUGAAAGACCUGGAGCAAUUGUCUGUUGAGAGUCAUCCGAUA
AGGUCUAUUCUUGUUUUUAACCAUUCUAACAUUAGUCUAGCAGGAUCUUCUAUCAAAGGAGAGUUGAGA
UUGUAGUCAGAGUUCUUAAGAUGUUGGGGACAAUGGGGGUUCUUCUGGGUUCAGCCUUAACUCGAUUG
GGUCAUAGGGUACAUAAGACCUCUCCAGGAUCGAGCAUCUUGAAGUAAGUCUAGGUUGAGAAAGUGU
UGCCAGUUGUUCUUGAUCUAUUGUUUUUUCUGCAGUAAAAGCUACCGCGGGUAGGCGCGCGUAUU
AAUUAAGGAGGGGUGUAGUUUUUUAUGCUAGCGUAGCCCGGGCGUAGCGGCCGUAACUUGCAGGUC
CUUUGAGGGAUGUUAUAGUUUUUUAACAUCUCAAAGAGGAUCAAAUUGUAAGGUAAUACAUCUAGCUUU
GAUGAAAUGCAGCGGUUUAUUGGCACCAUAACAUGUUUUUGUUAUAUUGCUCCCAAGGGUUGGAGUUGA
GCCCAUAGUUCUCUCCUCCAGAGGUAAACAAGUGAUAUUAUUGCGGGAUUAUUCUGACGUACGUUAUU
CUAGAAGCAGAGAGGAUUCUUGUCCUCUUCGGACCUUGUGUCUGAAGAGACAUGUCAGACCAUAGUUG
ACAUGCUCUCGGGUUCAUGUUGAUACACCAGACUCUGCCUGGAUAUGACACUGUUUUGCAAUCACUCUU
AUUUGCAAUCCGACGAACUCAGUAUCAUCCCAAGUGAUCUCCUGAGAGUAUCCAACUCCUCCCUU

图14(续)

CAAGAGGGCCCCUGGAAUCAGCCACUGGAAGUAAAAGGUUCUCCUCAAUUUGUAUACCCAGUUCAGGCC
CUCAGGGACUGGAGAUCCUGACAAAGCCAGUCCAAUAACCACUUGACUAACCCGAUCAUCCUAUGAUUC
CCAGAAUAUAUCUCGUGCGAAUGAUUUCAGAAUGUGCCGCAGGAUCCUGAACGAGUAACCAUUCGGGCUAC
ACACUUUAACCCUCCGUUGAUACAAAAGUCCUCAUGUUCUUCUUGCCUGUAAGUUCUUUCAGCGGGAC
GUAUUCAGGGGGUGGAAGCCACAAGUCAUCGUCAUCCAGAGGGGUGACGCGGGAGAGGAUUUUUGAGUG
UCCUCGUCUCCUGCGGUUUUACUAUCUACGUAGGAGGUUCAUCGAUUUAUCAGUGGUGUUGCCUGUUU
UUUUCUAUGUUGACUUUGGGACAUCUCGGAUUUUAUUGUCUAGAGGGACUGAGGGGAGAGGUUCGGUUAGC
AAGAUGUAUAGCGAUUCAAGUCAUCUUGCAUGAUUUUACUCAGCUUGUCGGAUUCGACUAACAACUGGAA
UUUCUUAAGAUUGGCCAAAGCGACCCAUCCAGUACACAUCUUAAGGGGGAGUUUGGACCCGUAUGGGCU
AAACGGGUCACACCUGGUACAUUUUUUGCCUUAUUAACUAUAUCAUCAAGGUUCAUUUUCAAUUGCUCAA
AAUUAUACAAGAGUAUCCUGAGGAUCGAGAGGCAACUUAUAUUUUUGGAGAAACUUUCGCAAUCUG
GUGAGCGAUCUCAGCCUCCACUGAUAGAUAUCCUUCUUAUUGGUAGCCGACCAUUCAGGGCUGGAGGG
CCAGAAGCAAUUUGAGCCGCCAUCCUGGCUUUCGAUGAUUGGCUUUCUCUCUGAGAAGGAGUGGGUGUUG
UCUCCUUCUUGAGCUCUCGGCCAGUAGUCUGGGUUGAUUUUAUCCUCUGAAGACUUUCCUGGAGGGUUGG
AAAGUUGACCGGCAUAGGAUAUAUCUCUUCUACGGUCUGUGACCAUAUCUUGAGAAUCUCUCUCCU
GACCUCAUUUGUCUGACUAUUUGGACUCCACAUAUUUCCAGGUUAGACUGGAACAGGAAGCUAGGAUCCU
CUCCUUAUCCAUCUGAAAGUCCUCUGAUACUUGCCUUCUCCACCUCUUGGCUAUCUCACCAUGGUUGGG
CGAUUUUCCAUCAUCCAGGUGAAGUCGCCCCAUUCCUCAGGGAGAUUGUCCACCUCUAUGGGUUCUCCU
UGGAGAUGAGCCUGAUUGUCUUCGAUAUUUCUAUUGAUCAGAUCAACAGUUUCUUCAGCCAUCUCAAGAU
CGGCCAGACCGGCUCAAUAGCACUAGGAUUGACAAAGAUCUUGCUCAUGUUUGGGAUGGUUCCGAAAGGA
GGGGUGUUAAGUUUUUUAUCAUGGAUAUACACAUAUCCGUAUAUUUCCGGCAUUUUGUUAUUAACUUCU
UACACAUUGAUCCUAGCAGAAGCACAGGCUGCAGGGUGACGGUCCAUCCCGCUCUCCUGGGCACAAGACA
UGGGCAGCGUGCCAUAUCCUGCUCCUCCACCUCGCGGGAAGCCAUGGCUACCGGAUCCCGAUUGGAA
GUAGAGGUUCUCUCCGGAACCUAGUCACUCGAAUAUGUCUUGUUUAGAAACUCGGCGAAUGAGUUUGGA
CGGGCUUGAUGAUUGGAACUGACUGAGACAUAUCUCCGUAUGUGAGAUUCUUUAGUCGACCUCCAUUA
UCAUGAUUCGAGUAUAAACAGCCUCCGGACUUCUGGUUUCACCUGAAAAGUAGUCCUCGUCGUCAGAGUU
GACAGUUCCAUCAUCUGCCAGUGCUACGUCAGUCUUGUCAGUUCAGCCGCCUCGUAUUCUUGAAGUUCU
UUCUCAUCUCUGAAGAAUCUUCUUCAAAUGUCCUUCUCCGAAAGAAUCCUCUCCAGAUAGCCCCUA
GAACAGACAUUUCAUGAGGAGCACAUUGCAGCAAUAACCGUUGCAUUUAGGGAUUCGACUUGACCCAUUA
GCAUCCUACAAGUGAAUGAGAUUGAACAGUGACCAACAGCAUUGAUGAAUAAGGAGAUUUCCACUC
AAGCCUAGUGAACGGAAGUGGAUGAAUAAGAGUGAGGAACAGCUGUCUCCUGCCUGGCUCAAACAUUC
UUCUUUAUCUCUCCUCAAAGUUCUUGUGGAAGAAUAUAGUAUUGCCUCUCUAGCGGUGAGAUUGAUUG
UUUUAUGAACCCAGUAAAUGAUACCAGUCCUGAACAGUCUUAUAAGCAGUGACAACUGUGCCCAUCUG
AUUGCUGAAUAUAGAUUCUAAUCCGGGAGAAAAACAUGUCAUAGGUUCCGGCCAAAAUCUGAAGUUUG
GUUAUGUACUCCAAUAGCACACAUAUUUGAGUUGUCAUUAAGAUUGGUGUCCACGAUUUUAACAAA
AGGGGCGUGUCUCAAUAUCGUCUUAUCCUGUCUGCAAUGUUGUCUUAUAGUUACCAGUGUUUUGCCCG
GAUAUUUUGCUCAACCUUAUACAGACUCAAGAGAAGACCGACUAAGGACGCAUGCUCAGGGACAGUGGGGU
CUCUUGUCAGUUCUCCUGUCAGAGCCCAAUCCCUUCUACAUCAGUACGUUUUAUCUCCACCAG
AGAACCUGGGGUGAUUCUUAUCUCCUUUCGUGCAAUACAAUCCAUAGCUGGUCAGUCUUCGGACAU
GUCCCCUAAAAACUGCAUUGCCGUGCCAAUAGGAACAUAUACUAGGUAUAGUUUGGCGGCGC
UCAUGCCUGACAAAACUGACUUGUAUGCUUUAUUUAAAUCGGGAGCCUUCUUAAGGGUUAACAGGGCU
UUUCAAAUCUUUGAUGGCAGGGUACUUGUACUCAUAUUGAUCCACGAUAAUCUCAGGCUUCAAGAGACC
ACCUGAUUAUGACUUUGAAUACAUCUUGUCGGCAUCCAUGUAGGGGUGUUAUUAUUGCUUUGCAA
UUGACAAUGUCUGUUUUUUCUUGAUUGGUUUGUAAGCGU-3'

图14(续)