(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 114127303 A
(43)申请公布日 2022.03.01

(21)申请号 201980091248.4

(22)申请日 2019.12.12

(30)优先权数据
62/778,866 2018.12.12 US
62/778,861 2018.12.12 US

(85)PCT国际申请进入国家阶段日
2021.08.04

(86)PCT国际申请的申请数据
PCT/US2019/065963 2019.12.12

(87)PCT国际申请的公布数据
WO2020/123795 EN 2020.06.18

(71)申请人 旗舰先锋创新V股份有限公司
地址 美国马萨诸塞州

(72)发明人 E.G.韦因斯坦 A.卡韦吉安
S.德拉格雷夫 N.L.尤兹维亚克
K.J.莱博 F.M.迪亚兹
D.M.纳万达 R.D.泰德斯通
J.D.皮茨

(74)专利代理机构 北京市柳沈律师事务所 11105
代理人 张文辉

(51)Int.Cl.
C12N 15/86(2006.01)

权利要求书5页 说明书318页
序列表（电子公布） 附图60页

(54)发明名称

用于递送蛋白质替代治疗方式的指环体

(57)摘要

本发明总体上涉及指环体及其组合物和用途。

CN 114127303 A

1.一种合成的指环体,所述合成的指环体包含:

(I)遗传元件,其包含:

(a)启动子元件,

(b)编码外源性效应子的核酸序列,其中所述核酸序列与所述启动子元件可操作地连接,并且其中所述外源性效应子是促红细胞生成素(EPO),和

(c)5'UTR结构域,其包含SEQ ID NO:54的核苷酸323-393的核酸序列,或与其至少90%相同的核酸序列;

(II)蛋白质外部,其包含由指环病毒ORF1核酸编码的多肽,例如ORF1分子;

其中所述遗传元件被封闭在所述蛋白质外部内;并且

其中所述合成的指环体能够将所述遗传元件递送到人细胞中。

2.如权利要求1所述的合成的指环体,其中所述ORF1分子包含SEQ ID NO:217的氨基酸序列,或与其具有至少90%同一性的氨基酸序列。

3.如前述权利要求中任一项所述的合成的指环体,其中所述ORF1分子由SEQ ID NO:54的核苷酸612-2612编码。

4.如前述权利要求中任一项所述的合成的指环体,其中所述遗传元件包含SEQ ID NO:54的核苷酸2868-2929的核酸序列,或与其具有至少85%序列同一性的核酸序列。

5.如前述权利要求中任一项所述的合成的指环体,其中所述ORF1分子包含氨基酸序列,所述氨基酸序列包含以下中的一种或多种:如表16中所列的精氨酸富集区、胶冻卷结构域、高变结构域、N22结构域和/或C末端结构域的氨基酸序列,或与其具有至少85%同一性的氨基酸序列。

6.如前述权利要求中任一项所述的合成的指环体,其中所述ORF1分子包含SEQ ID NO:58的氨基酸序列,或与其具有至少85%序列同一性的核酸序列。

7.如前述权利要求中任一项所述的合成的指环体,所述合成的指环体进一步包含包含如表16中所列的ORF2、ORF2/2、ORF2/3、ORF1/1或ORF1/2的氨基酸序列或与其具有至少85%同一性的氨基酸序列的多肽。

8.如前述权利要求中任一项所述的合成的指环体,其中所述遗传元件编码如表16中所列的ORF1、ORF2、ORF2/2、ORF2/3、ORF1/1或ORF1/2的氨基酸序列,或与其具有至少85%同一性的氨基酸序列。

9.如前述权利要求中任一项所述的合成的指环体,其中所述合成的指环体不包括包含如表16中所列的ORF2、ORF2/2、ORF2/3、TAIP、ORF1/1或ORF1/2的氨基酸序列或与其具有至少85%同一性的氨基酸序列的多肽。

10.如前述权利要求中任一项所述的合成的指环体,其中所述遗传元件不编码如表16中所列的ORF1、ORF2、ORF2/2、ORF2/3、ORF1/1或ORF1/2的氨基酸序列,或与其具有至少85%同一性的氨基酸序列。

11.如前述权利要求中任一项所述的合成的指环体,其中所述ORF1分子包含氨基酸序列$YNPX^2DXGX^2N$(SEQ ID NO:829),其中X^n各自独立地是任何n个氨基酸的连续序列。

12.如权利要求11所述的合成的指环体,其中所述ORF1分子进一步包含位于氨基酸序列$YNPX^2DXGX^2N$(SEQ ID NO:829)侧翼的第一β链和第二β链,例如,其中所述第一β链包含氨基酸序列$YNPX^2DXGX^2N$(SEQ ID NO:829)的酪氨酸(Y)残基和/或其中所述第二β链包含氨基

酸序列YNPX^{2}DXGX^{2}N(SEQ ID NO:829)的第二个天冬酰胺(N)残基(从N到C)。

13.如前述权利要求中任一项所述的合成的指环体,其中所述ORF1分子按从N末端到C末端的方向依次包含第一β链、第二β链、第一α螺旋、第三β链、第四β链、第五β链、第二α螺旋、第六β链、第七β链、第八β链和第九β链。

14.如前述权利要求中任一项所述的合成的指环体,其中所述遗传元件能够通过滚环复制在宿主细胞中扩增,例如产生至少8个拷贝。

15.如前述权利要求中任一项所述的合成的指环体,其中所述遗传元件是单链的。

16.如前述权利要求中任一项所述的合成的指环体,其中所述遗传元件是环状的。

17.如前述权利要求中任一项所述的合成的指环体,其中所述遗传元件是DNA。

18.如前述权利要求中任一项所述的合成的指环体,其中所述遗传元件是负链DNA。

19.如前述权利要求中任一项所述的合成的指环体,其中所述遗传元件以少于进入所述细胞的所述指环体的10%、8%、6%、4%、3%、2%、1%、0.5%、0.2%、0.1%的频率整合,例如,其中所述合成的指环体是非整合的。

20.如前述权利要求中任一项所述的合成的指环体,其中所述遗传元件包含表16-1中所示的共有5'UTR核酸序列的序列。

21.如前述权利要求中任一项所述的合成的指环体,其中所述遗传元件包含表16-2中所示的共有GC富集区的序列。

22.如前述权利要求中任一项所述的合成的指环体,其中所述遗传元件包含长度为至少100个核苷酸的序列,所述序列的至少70%(例如,约70%-100%、75%-95%、80%-95%、85%-95%、或85%-90%)位置由G或C组成。

23.如前述权利要求中任一项所述的合成的指环体,其中所述遗传元件包含SEQ ID NO:120的核酸序列。

24.如前述权利要求中任一项所述的合成的指环体,其中所述启动子元件对于野生型指环病毒是外源性的。

25.如前述权利要求中任一项所述的合成的指环体,其中所述启动子元件对于野生型指环病毒是内源性的。

26.如前述权利要求中任一项所述的合成的指环体,其中所述遗传元件具有约1.5-2.0、2.0-2.5、2.5-3.0、3.0-3.5、3.1-3.6、3.2-3.7、3.3-3.8、3.4-3.9、3.5-4.0、4.0-4.5或4.5-5.0kb的长度。

27.如前述权利要求中任一项所述的合成的指环体,其中所述合成的指环体能够感染人细胞,例如神经元(例如运动神经元)、肌细胞、视网膜细胞、肝细胞、血细胞、皮肤细胞、神经细胞、脂肪细胞,或内皮细胞,例如,体外。

28.如前述权利要求中任一项所述的合成的指环体,所述合成的指环体是基本上非免疫原性的,例如不诱导可检测的和/或不想要的免疫应答,例如根据实例4中所述的方法检测到的。

29.如权利要求28所述的合成的指环体,其中所述基本上非免疫原性的指环体在受试者中具有缺乏免疫应答的参考受试者中效力的至少约10%、20%、30%、40%、50%、60%、70%、80%、90%、95%或100%的效力。

30.如前述权利要求中任一项所述的合成的指环体,其中具有至少1000个所述指环体

的群体能够递送至少约100个拷贝(例如,至少1、2、3、4、5、10、20、30、40、50、100、200、300、400、500、600、700、800、900或1000个拷贝)的所述遗传元件进入一个或多个人细胞中。

31.一种合成的指环体,所述合成的指环体包含:

(I)遗传元件,其包含:

(a)启动子元件,

(b)编码外源性效应子的核酸序列,其中所述核酸序列与所述启动子元件可操作地连接,

其中所述外源性效应子包含当突变时引起人疾病的多肽,或所述多肽的功能变体,选自:

(i)治疗性酶,

(ii)结构蛋白质,或

(iii)补体蛋白质;和

(c)5'UTR结构域,其包含SEQ ID NO:54的核苷酸323-393的核酸序列,或与其至少85%相同的核酸序列;或

(II)蛋白质外部,其包含由ORF1核酸编码的多肽,例如ORF1分子;

其中所述遗传元件被封闭在所述蛋白质外部内;并且

其中所述合成的指环体能够将所述遗传元件递送到人细胞中。

32.一种药物组合物,所述药物组合物包含如前述权利要求中任一项所述的合成的指环体,以及药学上可接受的载剂或赋形剂。

33.如权利要求32所述的药物组合物,所述药物组合物包含至少10^3、10^4、105、10^6、10^7、10^8或10^9个合成的指环体。

34.如权利要求32或33所述的药物组合物,其中所述药物组合物具有预定比例的颗粒:感染性单位(例如,<300:1、<200:1、<100:1或<50:1)。

35.一种反应混合物,所述反应混合物包含:

(i)第一核酸(例如,双链或单链环状DNA),其包含如前述权利要求中任一项所述的合成的指环体的遗传元件的序列,和

(ii)第二核酸序列,其编码选自例如如表16中所列的ORF1、ORF2、ORF2/2、ORF2/3、ORF1/1或ORF1/2中的一种或多种氨基酸序列,或与其具有至少85%序列同一性的氨基酸序列。

36.如权利要求35所述的反应混合物,其中所述第一核酸和所述第二核酸在同一核酸分子中。

37.如权利要求35所述的反应混合物,其中所述第一核酸和所述第二核酸是不同的核酸分子。

38.如权利要求35所述的反应混合物,其中所述第一核酸和所述第二核酸是不同的核酸分子并且其中所述第二核酸作为双链环状DNA提供。

39.如权利要求35所述的反应混合物,其中所述第一核酸和所述第二核酸是不同的核酸分子并且其中所述第一核酸和所述第二核酸作为双链环状DNA提供。

40.如权利要求37所述的反应混合物,其中所述第二核酸序列包含在辅助细胞或辅助病毒中。

41.一种制备合成的指环体的方法,所述方法包括:

a)提供宿主细胞,其包含:

(i)第一核酸分子,其包含如前述权利要求中任一项所述的合成的指环体的遗传元件的核酸序列,和

(ii)第二核酸分子,其编码选自例如如表16中任一个中所列的ORF1、ORF2、ORF2/2、ORF2/3、ORF1/1或ORF1/2中的一种或多种氨基酸序列,或与其具有至少85%序列同一性的氨基酸序列;并且

b)在适合于制备合成的指环体的条件下孵育所述宿主细胞;

从而制备所述合成的指环体。

42.如权利要求41所述的方法,所述方法进一步包括在步骤(a)之前,将所述第一核酸分子和/或所述第二核酸分子引入所述细胞。

43.如权利要求42所述的方法,其中所述第二核酸分子在所述第一核酸分子之前、与其同时或在其之后被引入所述宿主细胞。

44.如权利要求41或42中任一项所述的方法,其中所述第二核酸分子被整合到所述宿主细胞的基因组中。

45.如权利要求41-44中任一项所述的方法,其中所述第二核酸分子是辅助(例如辅助质粒或辅助病毒的基因组)。

46.如权利要求41-44中任一项所述的方法,其中第二核酸分子编码包含氨基酸序列[W/F]X^7HX^3CX^1CX^5H(SEQ ID NO:949)的ORF2分子,其中X^n是任何n个氨基酸的连续序列。

47.一种制造合成的指环体制剂的方法,所述方法包括:

a)提供多个如权利要求1-31所述的合成的指环体、如权利要求32-34中任一项所述的药物组合物、或如权利要求35-40中任一项所述的反应混合物;

b)任选地对所述多个如权利要求1-31所述的合成的指环体评估以下一种或多种:本文所述的污染物、光密度测量值(例如OD 260)、颗粒数(例如通过HPLC)、感染性(例如颗粒:感染单位比);并且

c)例如,如果(b)的一种或多种参数满足指定的阈值,则将所述多个合成的指环体配制为例如适于施用于受试者的药物组合物。

48.一种宿主细胞,所述宿主细胞包含:

(i)第一核酸分子,其包含如前述权利要求中任一项所述的合成的指环体的遗传元件的核酸序列,和

(ii)任选地,第二核酸分子,其编码选自如表16中任一个中所列的ORF1、ORF2、ORF2/2、ORF2/3、ORF1/1或ORF1/2中的一种或多种氨基酸序列,或与其具有至少85%序列同一性的氨基酸序列。

49.一种将外源性效应子(例如,治疗性外源性效应子)递送至哺乳动物细胞的方法,所述方法包括:

(a)提供如前述权利要求中任一项所述的合成的指环体;以及

(b)使哺乳动物细胞与所述合成的指环体接触;

其中所述合成的指环体能够将所述遗传元件递送到所述哺乳动物细胞中;以及

任选地,其中在适合在宿主细胞中将所述遗传元件封闭在蛋白质外部内的条件下,通

过将所述遗传元件引入所述宿主细胞来产生所述合成的指环体；

从而将所述治疗性外源性效应子递送至所述哺乳动物细胞。

50.如权利要求1-31中任一项所述的合成的指环体或如权利要求32-34中任一项所述的药物组合物用于将所述遗传元件递送至宿主细胞的用途。

51.如权利要求1-31中任一项所述的合成的指环体或如权利要求32-34中任一项所述的药物组合物用于治疗受试者疾病或障碍的用途。

52.如权利要求1-31中任一项所述的合成的指环体或如权利要求32-34中任一项所述的药物组合物，其用于治疗受试者的疾病或障碍。

53.一种治疗受试者的疾病或障碍的方法，所述方法包括向所述受试者施用如权利要求1-31中任一项所述的合成的指环体或如权利要求32-34中任一项所述的药物组合物。

54.如权利要求1-31中任一项所述的合成的指环体或如权利要求32-34中任一项所述的药物组合物在制造用于治疗受试者的疾病或障碍的药物中的用途。

用于递送蛋白质替代治疗方式的指环体

[0001] 相关申请的交叉引用

[0002] 本申请要求2018年12月12日提交的美国临时申请号62/778,861和2018年12月12日提交的美国临时申请号62/778,866的权益。上述申请的内容通过引用整体并入本文。

[0003] 序列表

[0004] 本申请包含序列表,该序列表已以ASCII格式电子提交,并通过引用以其整体并入本文。所述ASCII副本创建于2019年12月9日,名为V2057-7003WO_SL.txt,并且大小为825,844字节。

背景技术

[0005] 持续需要开发合适的载体以向患者递送治疗性遗传物质。

发明内容

[0006] 本公开提供了一种指环体(anellosome),例如合成的指环体,其可用作递送媒介物,例如用于向真核细胞(例如人细胞或人组织)递送遗传材料、用于递送效应子(例如有效负载)用于递送治疗剂或治疗效应子(例如,蛋白质替代治疗剂)。可以递送的示例性蛋白质替代治疗剂包括例如酶或结构蛋白质。指环体可用于向患有由突变引起的疾病的受试者递送蛋白质,例如递送突变蛋白质的功能版本。

[0007] 在一些实施例中,指环体(例如颗粒,例如病毒颗粒,例如指环病毒(Anellovirus)颗粒)包含包封在蛋白质外部的遗传元件(例如,包含治疗性DNA序列的遗传元件)(例如,包含指环病毒衣壳蛋白,例如,指环病毒ORF1蛋白或由指环病毒ORF1核酸编码的多肽的蛋白质外部,例如,如本文所述),所述蛋白质外部能够将遗传元件引入细胞(例如哺乳动物细胞,例如人细胞)。在一些实施例中,指环体是包含蛋白质外部的颗粒,所述蛋白质外部包含由以下编码的多肽:指环病毒的ORF1核酸(例如,甲型细环病毒、乙型细环病毒或丙型细环病毒的ORF1核酸,例如,甲型细环病毒进化分枝1、甲型细环病毒进化分枝2、甲型细环病毒进化分枝3、甲型细环病毒进化分枝4、甲型细环病毒进化分枝5、甲型细环病毒进化分枝6或甲型细环病毒进化分枝7的ORF1,例如如本文所述)。本公开的指环体的遗传元件可以是环状和/或单链DNA分子(例如环状和单链),并且通常包括与蛋白质外部结合的由所述蛋白质外部封闭的蛋白质结合序列或附接于其上的多肽,其可以促进遗传元件被封闭在蛋白质外部内和/或遗传元件相对于其他核酸在蛋白质外部内富集。在一些情况下,遗传元件是环状或线性的。在一些情况下,遗传元件包含或编码效应子(例如,核酸效应子,例如非编码RNA,或多肽效应子,例如,蛋白质),例如,所述效应子可以在细胞中表达。在一些情况下,效应子例如对于野生型指环病毒或靶细胞是内源性效应子或外源性效应子。在一些实施例中,效应子对于野生型指环病毒或靶细胞是外源的。在一些实施例中,指环体可以通过接触细胞并将编码外效应子的遗传元件引入细胞中而将效应子递送到细胞中,从而效应子由细胞产生或表达。在某些情况下,效应子是内源性效应子(例如,对于靶细胞是内源的,但例如由指环体以增加的量提供)。在其他情况下,效应子是外源性效应子。在一些情况下,外源性效应

子可以调节细胞的功能或调节细胞中靶分子的活性或水平。例如，效应子可以降低细胞中靶蛋白的水平(例如，如实例3和4中所述)。在另一个实例中，指环体可以递送和体内表达效应子，例如外源蛋白质(例如，如实例19和28中所述)。指环体可用于例如将遗传物质递送至靶细胞、组织或受试者；将效应子递送至靶细胞、组织或受试者；或用于治疗疾病和病症，例如通过将可作为治疗剂起作用的效应子递送至所期望细胞、组织或受试者。

[0008] 本发明进一步提供合成的指环体。合成的指环病毒与野生型病毒(例如，野生型指环病毒，例如本文所述的)相比具有至少一个结构差异，例如相对于野生型病毒的缺失、插入、取代、修饰(例如，酶促修饰)。通常，合成的指环体包括封闭在蛋白质外部内的外源性遗传元件，所述蛋白质外部可用以将遗传元件，或其中编码的效应子(例如，外源性效应子或内源性效应子)(例如，多肽或核酸效应子)递送到真核细胞(例如人细胞)中。在实施例中，指环体不会引起可检测的和/或不想要的免疫或炎症应答，例如不会引起炎症的一种或多种分子标志物(例如，TNF-α、IL-6、IL-12、IFN)增加超过1％、5％、10％、15％，以及不会引起B细胞应答，例如反应性或中和抗体，例如指环体可以对靶细胞、组织或受试者基本上无免疫原性。

[0009] 在一方面，本发明的特征在于一种指环体，其包含：(a)蛋白质外部；(b)遗传元件，其包含启动子元件和编码外源性效应子的核酸序列(例如DNA序列)和蛋白质结合序列(例如外部蛋白质结合序列)；其中所述外源性效应子包含当突变时引起人疾病的多肽，或所述多肽的功能变体；其中所述遗传元件被封闭在所述蛋白质外部内；并且其中指环体被配置为将遗传元件递送到真核细胞中。任选地，遗传元件包含相对于野生型指环病毒基因组序列(例如，如本文所述)的至少一个差异(例如，突变、化学修饰或表观遗传改变)，例如插入、取代、修饰(例如酶促修饰)和/或缺失，例如结构域或其部分(例如TATA框、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号或GC富集区中的一个或多个)的缺失。

[0010] 在一方面，本发明的特征在于一种指环体，其包含：(i)遗传元件，其包含启动子元件和编码效应子(例如内源性或外源性效应子)的序列和蛋白质结合序列(例如，外部蛋白质结合序列，例如包装信号)；以及(ii)蛋白质外部；其中所述遗传元件被封闭在所述蛋白质外部(例如衣壳)内；并且其中所述指环体能够将所述遗传元件递送到真核(例如，哺乳动物，例如，人)细胞中。在一些实施例中，遗传元件是单链和/或环状DNA。可替代地或组合地，遗传元件具有以下特性中的一种、两种、三种或全部：是环状的，是单链的，它以少于进入细胞的遗传元件的约0.0001％、0.001％、0.005％、0.01％、0.05％、0.1％、0.5％、1％、1.5％、或2％频率整合进入细胞的基因组中，和/或它以小于1、2、3、4、5、6、7、8、9、10、15、20、25或30个拷贝/基因组整合进入靶细胞的基因组中。在一些实施例中，整合频率如Wang等人(2004,Gene Therapy[基因疗法]11:711-721,通过引用整体并入本文)所述确定。在一些实施例中，遗传元件被包封在蛋白质外部内。在一些实施例中，指环体能够将遗传元件递送到真核细胞中。在一些实施例中，遗传元件包含与野生型指环病毒序列(例如，野生型细环病毒(TTV)、细小环病毒(TTMV)或TTMDV序列，例如，表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中的任一个中列出的野生型指环病毒序列)具有至少75％(例如，至少75％、76％、77％、78％、79％、80％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％)序列同一性的核酸序列(例如，300-4000个核苷酸的核酸序列，例如300-3500个核苷酸、300-3000个核苷酸、300-2500个核苷酸、300-2000个核苷酸、300-1500个核苷

酸)。在一些实施例中,遗传元件包含与野生型指环病毒序列(例如,如本文所述的野生型指环病毒序列,例如表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中的任一个中列出的)具有至少75%(例如,至少75%、76%、77%、78%、79%、80%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%)序列同一性的核酸序列(例如,至少300个核苷酸、500个核苷酸、1000个核苷酸、1500个核苷酸、2000个核苷酸、2500个核苷酸、3000个核苷酸或更多的核酸序列)。在一些实施例中,核酸序列是经密码子优化的,例如用于在哺乳动物(例如人)细胞中表达。在一些实施例中,核酸序列中至少50%、60%、70%、80%、90%、95%、96%、97%、98%、99%或100%的密码子是经密码子优化,例如,用于在哺乳动物(例如,人)细胞中的表达。

[0011] 在一方面,本发明的特征在于一种感染性(对人细胞感染性)颗粒,其包含指环病毒衣壳(例如,包含指环病毒ORF,例如ORF1,多肽的衣壳),所述指环病毒衣壳包封遗传元件,所述遗传元件包含与衣壳结合的蛋白质结合序列和编码治疗性效应子的异源(对于指环病毒而言)序列。在实施例中,所述颗粒能够将遗传元件递送到哺乳动物例如人的细胞中。在一些实施例中,遗传元件与野生型指环病毒具有小于约6%(例如,小于6%、5.5%、5%、4.5%、4%、3.5%、3%、2.5%、2%、1.5%或更少)同一性。在一些实施例中,遗传元件与野生型指环病毒具有不超过1.5%、2%、2.5%、3%、3.5%、4%、4.5%、5%、5.5%或6%同一性。在一些实施例中,遗传元件与野生型指环病毒具有至少约2%到至少约5.5%(例如,2至5%、3%至5%、4%至5%)同一性。在一些实施例中,遗传元件具有大于约2000、3000、4000、4500或5000个核苷酸的非病毒序列(例如,非指环病毒基因组序列)。在一些实施例中,遗传元件具有大于约2000至5000、2500至4500、3000至4500、2500至4500、3500或4000、4500(例如,约3000至4500)个核苷酸的非病毒序列(例如,非指环病毒基因组序列)。在一些实施例中,遗传元件是单链、环状DNA。可替代地或组合地,遗传元件具有以下特性中的一种、两种或3种:是环状的,是单链的,它以少于进入细胞的遗传元件的约0.001%、0.005%、0.01%、0.05%、0.1%、0.5%、1%、1.5%、或2%的频率整合进入细胞的基因组中,它以小于1、2、3、4、5、6、7、8、9、10、15、20、25或30个拷贝/基因组整合进入靶细胞的基因组中或以少于进入细胞的遗传元件的约0.0001%、0.001%、0.005%、0.01%、0.05%、0.1%、0.5%、1%、1.5%或2%的频率整合。在一些实施例中,整合频率如Wang等人(2004,Gene Therapy[基因疗法]11:711-721,通过引用整体并入本文)所述确定。

[0012] 本文还描述了基于指环病毒的病毒载体和病毒颗粒,其可用于将药剂(例如,外源性效应子或内源性效应子,例如,治疗性效应子)递送到细胞(例如,待治疗受试者中的细胞)。在一些实施例中,指环病毒可用作有效递送媒介物,将药剂(例如本文所述的效应子)引入靶细胞,例如治疗性或预防性治疗的受试者中的靶细胞。

[0013] 在一方面,本发明的特征在于一种多肽(例如合成多肽,例如ORF1分子),其包含(例如,串联):

[0014] (i)第一区域,其包含精氨酸富集区,与本文描述的精氨酸富集区序列具有至少70%(例如,至少约70%、80%、90%、95%、96%、97%、98%、99%或100%)序列同一性的氨基酸序列,或者含有至少60%、70%或80%碱性残基(例如,精氨酸、赖氨酸或其组合)的至少约40个氨基酸的序列,

[0015] (ii)第二区域,其包含胶冻卷(jelly-roll)结构域,例如与本文所述的胶冻卷区

域序列具有至少30％(例如至少约30％、35％、40％、50％、60％、70％、80％、90％、95％、96％、97％、98％、99％或100％)序列同一性的氨基酸序列或含有至少6个β链的序列，

[0016] (iii)第三区域，其包含与本文所述N22结构域序列具有至少30％(例如至少约30％、35％、40％、50％、60％、70％、80％、90％、95％、96％、97％、98％、99％或100％)序列同一性的氨基酸序列，

[0017] (iv)第四区域，其包含与本文所述的指环病毒病毒ORF1 C末端结构域(CTD)序列具有至少70％(例如至少约70％、80％、90％、95％、96％、97％、98％、99％或100％)序列同一性的氨基酸序列，以及

[0018] (v)任选地，其中所述多肽具有与本文描述的野生型指环病毒ORF1蛋白的序列同一性小于100％、99％、98％、95％、90％、85％、80％的氨基酸序列。

[0019] 在一些实施例中，所述多肽包含与本文所述的指环病毒ORF1分子(例如，如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所列)至少约70％、80％、90％、95％、96％、97％、98％、99％或100％的序列同一性。在一些实施例中，所述多肽包含与本文所述指环病毒ORF1分子(例如，如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10)中任一个中所列)的子序列(例如，精氨酸(Arg)富集结构域、胶冻卷结构域、高可变区(HVR)、N22结构域或C末端结构域(CTD))至少约70％、80％、90％、95％、96％、97％、98％、99％或100％的序列同一性。在一个实施例中，(i)、(ii)、(iii)和(iv)区的氨基酸序列与其各自的参考具有至少90％序列同一性，并且其中所述多肽具有与本文描述的野生型指环病毒ORF1蛋白具有低于100％、99％、98％、95％、90％、85％、80％序列同一性的氨基酸序列。

[0020] 在一方面，本发明的特征在于复合物，所述复合物包含如本文所述的多肽(例如，如本文所述的指环病毒ORF1分子)和包含启动子元件和编码效应子(例如，外源性效应子或内源性效应子)的核酸序列(例如，DNA序列)以及蛋白质结合序列的遗传元件。

[0021] 本公开还提供了核酸分子(例如，包括如本文所述的遗传元件的核酸分子，或包括编码如本文所述的蛋白质外部蛋白的序列的核酸分子)。本发明的核酸分子可包括(a)本文所述的遗传元件和(b)编码本文所述的蛋白质外部蛋白的核酸序列中的一者或两者。

[0022] 在一方面，本发明的特征在于分离的核酸分子，其包含遗传元件，所述遗传元件包含与编码效应子例如有效负载的序列可操作地连接的启动子元件和外部蛋白质结合序列。在一些实施例中，外部蛋白质结合序列包括与本文公开的指环病毒的5'UTR序列至少75％(80％、85％、90％、95％、97％、100％)相同的序列。在实施例中，遗传元件是单链DNA，是环状的，以少于进入细胞的遗传元件的约0.001％、0.005％、0.01％、0.05％、0.1％、0.5％、1％、1.5％、或2％的频率整合，和/或以小于1、2、3、4、5、6、7、8、9、10、15、20、25或30个拷贝/基因组整合进入靶细胞的基因组中或以少于进入细胞的遗传元件的约0.001％、0.005％、0.01％、0.05％、0.1％、0.5％、1％、1.5％或2％的频率整合。在一些实施例中，整合频率如Wang等人(2004,Gene Therapy[基因疗法]11:711-721，通过引用整体并入本文)所述确定。在实施例中，效应子不是源自TTV，并且不是SV40-miR-S1。在实施例中，核酸分子不包含TTMV-LY2的多核苷酸序列。在实施例中，启动子元件能够指导效应子在真核(例如哺乳动物，例如人)细胞中的表达。

[0023] 在一些实施例中，核酸分子是环状的。在一些实施例中，核酸分子是线性的。在一

些实施例中，本文所述的核酸分子包含一个或多个经修饰的核苷酸(例如碱基修饰、糖修饰或骨架修饰)。

[0024] 在一些实施例中，核酸分子包含编码ORF1分子(例如，指环病毒ORF1蛋白，例如，如本文所述)的序列。在一些实施例中，核酸分子包含编码ORF2分子(例如，指环病毒ORF2蛋白，例如，如本文所述)的序列。在一些实施例中，核酸分子包含编码ORF3分子(例如，指环病毒ORF3蛋白，例如，如本文所述)的序列。在一方面，本发明的特征在于一种遗传元件，其包含以下中的一个、两个或三个：(i)启动子元件和编码效应子的序列，例如外源性或内源性效应子；(ii)与野生型指环病毒序列具有至少75％(例如，至少75％、76％、77％、78％、79％、80％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％)序列同一性的至少72个连续核苷酸(例如，至少72、73、74、75、76、77、78、79、80、90、100或150个核苷酸)；或与野生型指环病毒序列具有至少72％(例如，至少72、73、74、75％、76％、77％、78％、79％、80％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％)序列同一性的至少100(例如，至少300、500、1000、1500)个连续核苷酸；以及(iii)蛋白质结合序列，例如外部蛋白质结合序列，其中核酸构建体是单链DNA；并且其中核酸构建体是环状的，以少于进入细胞的遗传元件的约0.001％、0.005％、0.01％、0.05％、0.1％、0.5％、1％、1.5％、或2％频率整合，和/或以小于1、2、3、4、5、6、7、8、9、10、15、20、25或30个拷贝/基因组整合进入靶细胞的基因组中。在一些实施例中，编码效应子(例如，外源性或内源性效应子，例如，如本文所述)的遗传元件是经密码子优化的。在一些实施例中，遗传元件是环状的。在一些实施例中，遗传元件是线性的。在一些实施例中，遗传元件包括指环载体，例如，如本文所述。在一些实施例中，本文所述的遗传元件包含一个或多个经修饰的核苷酸(例如碱基修饰、糖修饰或骨架修饰)。在一些实施例中，遗传元件包含编码ORF1分子(例如，指环病毒ORF1蛋白，例如，如本文所述)的序列。在一些实施例中，遗传元件包含编码ORF2分子(例如，指环病毒ORF2蛋白，例如，如本文所述)的序列。在一些实施例中，遗传元件包含编码ORF3分子(例如，指环病毒ORF3蛋白，例如，如本文所述)的序列。

[0025] 在一方面，本发明的特征在于宿主细胞或辅助细胞，其包含：(a)核酸，其包含编码ORF1分子、ORF2分子或ORF3分子中的一个或多个的序列(例如，编码本文描述的指环病毒ORF1多肽的序列)，其中所述核酸是质粒，是病毒核酸，或者整合到辅助细胞染色体中；和(b)遗传元件，其中所述遗传元件包含(i)可操作地连接至编码效应子(例如，外源性效应子或内源性效应子)的核酸序列(例如，DNA序列)的启动子元件和(ii)结合(a)的多肽的蛋白质结合序列，其中任选地所述遗传元件不编码ORF1多肽(例如，ORF1蛋白质)。例如，宿主细胞或辅助细胞包括以顺式(同一核酸分子的两个部分)或以反式(不同核酸分子的每个部分)的(a)和(b)。在实施例中，(b)的遗传元件是环状单链DNA。在一些实施例中，宿主细胞是生产细胞系。在一些实施例中，宿主细胞或辅助细胞是贴壁的或悬浮的，或两者。在一些实施例中，宿主细胞或辅助细胞在微载剂中生长。在一些实施方案中，宿主细胞或辅助细胞与cGMP生产实践相容。在一些实施例中，宿主细胞或辅助细胞在适合促进细胞生长的培养基中生长。在某些实施例中，一旦宿主细胞或辅助细胞已经充分生长(例如，达到合适的细胞密度)，可以将培养基更换为适合宿主细胞或辅助细胞生产指环体的培养基。

[0026] 在一方面，本发明的特征在于一种药物组合物，其包含本文所述的指环体(例如合成的指环体)。在实施例中，药物组合物进一步包含药学上可接受的载剂或赋形剂。在实施

例中，药物组合物包含以下单位剂量，所述单位剂量包含约10^5-10^{14}基因组当量的指环体/千克目标受试者。在一些实施例中，包含制剂的药物组合物在可接受的时间和温度范围内是稳定的，和/或与所期望的施用途径和/或该施用途径需要的任何装置(例如针头或注射器)相容。在一些实施例中，药物组合物被配制用于作为单剂量或多剂量施用。在一些实施例中，药物组合物在施用地点配制，例如由医疗保健专业人员配制。在一些实施例中，药物组合物包含所期望浓度的指环体基因组或基因组等价物(例如，如由基因组数量/体积所定义)。

[0027] 在一方面，本发明的特征在于一种治疗受试者的疾病或障碍的方法，所述方法包括向所述受试者施用例如本文所述的指环体，例如合成的指环体。

[0028] 在一方面，本发明的特征在于一种将效应子或有效负载(例如，内源性或外源性效应子)递送到细胞、组织或受试者的方法，所述方法包括向所述受试者施用例如本文所述的指环体，例如合成的指环体，其中所述指环体包含编码效应子的核酸序列。在实施例中，有效负载是核酸。在实施例中，有效负载是多肽。

[0029] 在一方面，本发明的特征在于一种将指环体递送至细胞的方法，所述方法包括使指环体，例如合成的指环体(例如如本文所述)与细胞(例如，真核细胞，例如，哺乳动物细胞)接触，例如在体内或离体接触。

[0030] 在一方面，本发明的特征在于一种制备指环体，例如合成的指环体的方法。所述方法包括：

[0031] a)提供宿主细胞，其包含：

[0032] (i)第一核酸分子，其包含如本文所述的指环体，例如合成的指环体的遗传元件的核酸序列，和

[0033] (ii)所述第一核酸或第二核酸分子，所述第二核酸分子编码以下中的一种或多种：选自ORF1、ORF2、ORF2/2、ORF2/3、ORF1/1或ORF1/2的氨基酸序列(例如，如表16中任一个中所列)，或与其具有至少70％(例如，至少70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％)序列同一性的氨基酸序列；以及

[0034] b)在适合于制备指环体的条件下孵育所述宿主细胞。

[0035] 在一些实施例中，所述方法进一步包括，在步骤(a)之前，将所述第一核酸分子和/或所述第二核酸分子引入所述宿主细胞。在一些实施例中，所述第二核酸分子在所述第一核酸分子之前、与其同时或在其之后被引入所述宿主细胞。在其他实施例中，所述第二核酸分子被整合到所述宿主细胞的基因组中。在一些实施例中，所述第二核酸分子是辅助(例如辅助质粒或辅助病毒的基因组)。在另一方面，本发明的特征在于一种制备指环体组合物的方法，所述方法包括：

[0036] a)提供宿主细胞，其包含例如表达例如本文所述的指环体(例如合成的指环体)的一种或多种组分(例如，所有组分)。例如，所述宿主细胞包含(a)包含编码本文所述的指环病毒ORF1多肽的序列的核酸，其中所述核酸是质粒、病毒核酸或整合到辅助细胞染色体中；(b)遗传元件，其中所述遗传元件包含(i)可操作地连接至编码效应子(例如，外源性效应子或内源性效应子)的核酸序列(例如，DNA序列)的启动子元件和(i)结合(a)的多肽的蛋白质结合序列(例如包装序列)，其中所述宿主细胞或辅助细胞包括以顺式或以反式的(a)和(b)。在实施例中，(b)的遗传元件是环状单链DNA。在一些实施例中，宿主细胞是生产细胞

系;

[0037] b)在适合从所述宿主细胞产生指环体制剂的条件下培养所述宿主细胞,其中所述制剂的指环体包含包封所述遗传元件(例如,如本文所述)的蛋白质外部(例如,包含ORF1分子),从而制备指环体制剂;并且

[0038] 任选地,c)将指环体的制剂配制为例如适于向受试者施用的药物组合物。

[0039] 在一些实施例中,在生产时(例如,通过瞬时转染)将指环体的组分引入宿主细胞。在一些实施例中,宿主细胞稳定表达指环体的组分(例如,其中一种或多种编码指环体的组分的核酸被引入宿主细胞或其祖细胞,例如通过稳定转染)。

[0040] 在一些实施例中,所述方法进一步包括一个或多个纯化步骤(例如,通过沉降、层析和/或超滤纯化)。在一些实施例中,纯化步骤包括从制剂中去除血清、宿主细胞DNA、宿主细胞蛋白质、缺乏遗传元件的颗粒和/或酚红中的一种或多种。在一些实施例中,所得制剂或包含所述制剂的药物组合物在可接受的时间和温度范围内是稳定的,和/或与所期望的施用途径和/或该施用途径需要的任何装置(例如针头或注射器)相容。

[0041] 在一方面,本发明的特征在于一种制备指环体组合物的方法,所述方法包括:a)提供多个本文所述的指环体,或本文所述的指环体的制剂;b)将所述指环体或其制剂配制成例如适合施用于受试者的药物组合物。

[0042] 在一方面,本发明的特征在于一种制备包含指环体的宿主细胞的方法,所述宿主细胞例如第一宿主细胞或生产者细胞(例如,如图12所示),例如第一宿主细胞的群体,所述方法包括将例如本文所述的遗传元件引入宿主细胞中并在适合于生产指环体的条件下培养宿主细胞。在实施例中,所述方法进一步包括将辅助物例如辅助病毒引入宿主细胞中。在实施例中,引入包括用指环体对宿主细胞进行转染(例如化学转染)或电穿孔。

[0043] 在一方面,本发明的特征在于一种制备指环体的方法,其包括提供包含例如本文所述的指环体的宿主细胞,例如第一宿主细胞或生产者细胞(例如,如图12所示),并从宿主细胞中纯化指环体。在一些实施例中,所述方法进一步包括,在提供步骤之前,使宿主细胞与例如本文所述的指环体接触,并且在适合于生产指环体的条件下孵育宿主细胞。在实施例中,宿主细胞是在上述制备宿主细胞的方法中描述的第一宿主细胞或生产者细胞。在实施例中,从宿主细胞中纯化指环体包括裂解宿主细胞。

[0044] 在一些实施例中,所述方法进一步包括第二步骤,所述步骤使由第一宿主细胞或生产者细胞产生的指环体与第二宿主细胞,例如允许细胞(例如,如图12所示),例如第二宿主细胞群体接触。在一些实施例中,所述方法进一步包括在适合于生产指环体的条件下孵育第二宿主细胞。在一些实施例中,所述方法进一步包括从第二宿主细胞中纯化指环体,例如从而产生指环体种子群体。在实施例中,与第一宿主细胞群体相比,第二宿主细胞群体产生至少约2-100倍更多的指环体。在实施例中,从第二宿主细胞中纯化指环体包括裂解第二宿主细胞。在一些实施例中,所述方法进一步包括第二步骤,所述步骤使第二宿主细胞产生的指环体与第三宿主细胞,例如允许细胞(例如,如图12所示),例如第三宿主细胞群体接触。在一些实施例中,所述方法进一步包括在适合于生产指环体的条件下孵育第三宿主细胞。在一些实施例中,所述方法进一步包括从第三宿主细胞中纯化指环体,例如从而产生指环体储备群体。在实施例中,从第三宿主细胞中纯化指环体包括裂解第三宿主细胞。在实施例中,与第二宿主细胞群体相比,第三宿主细胞群体产生至少约2-100倍更多的指环体。

[0045] 在一些实施例中，宿主细胞在适合促进细胞生长的培养基中生长。在某些实施例中，一旦宿主细胞已经充分生长(例如，达到合适的细胞密度)，可以将培养基更换为适合宿主细胞生产指环体的培养基。在一些实施例中，由宿主细胞生产的指环体从宿主细胞分离(例如，通过裂解宿主细胞)，之后与第二宿主细胞接触。在一些实施例中，由宿主细胞生产的指环体与第二宿主细胞接触而无需介入纯化步骤。

[0046] 在一方面，本发明的特征在于一种制备药物指环体制剂的方法。所述方法包括(a)制备如本文所述的指环体制剂，(b)针对以下一种或多种药物质量控制参数评估所述制剂(例如，药物指环体制剂、指环体种子群体或指环体储备物群体)：同一性、纯度、效价、效力(例如，以每个指环体颗粒的基因组当量)和/或例如来自指环体所包含的遗传元件的核酸序列，和(c)配制用于评估的药物用途的制剂满足预定标准，例如，满足药物规范。在一些实施例中，评估同一性包括评估(例如，确认)指环体的遗传元件的序列，例如编码效应子的序列。在一些实施例中，评估纯度包括评估杂质的量，所述杂质是例如支原体、内毒素、宿主细胞核酸(例如，宿主细胞DNA和/或宿主细胞RNA)、动物来源的过程杂质(例如，血清白蛋白或胰蛋白酶)、具有复制能力的因子(RCA)，例如具有复制能力的病毒或不需要的指环体(例如，除所需的指环体例如本文所述的合成的指环体以外的指环体)、游离病毒衣壳蛋白、外源因子和聚集体。在一些实施例中，评估效价包括评估制剂中功能性相比于非功能性(例如感染性相比于非感染性)指环体的比率(例如，如通过HPLC评估的)。在一些实施例中，评估效力包括评估制剂中可检测的指环体功能水平(例如，其中编码的效应子的表达和/或功能或基因组等价物)。

[0047] 在实施例中，配制的制剂基本上不含病原体、宿主细胞污染物或杂质；具有预定水平的非感染性颗粒或预定比例的颗粒：感染单位(例如，<300:1、<200:1、<100:1或<50:1)。在一些实施例中，可以在单个批次中生产多个指环体。在实施例中，可以评估批次中生产的指环体的水平(例如，单独或一起)。

[0048] 在一方面，本发明的特征在于一种宿主细胞，其包含：

[0049] (i)第一核酸分子，其包含如本文所述的指环体的遗传元件的核酸序列，和

[0050] (ii)任选地，第二核酸分子，所述第二核酸分子编码以下中的一种或多种：选自ORF1、ORF2、ORF2/2、ORF2/3、ORF1/1或ORF1/2的氨基酸序列(如表16中任一个中所列)，或与之具有至少约70％(例如，至少约70％、80％、90％、95％、96％、97％、98％、99％或100％)序列同一性的氨基酸序列。

[0051] 在一方面，本发明的特征在于一种反应混合物，其包含本文所述的指环体和辅助病毒，其中所述辅助病毒包含多核苷酸，例如编码外部蛋白质的多核苷酸(例如，能够结合外部蛋白质结合序列和任选的脂质包膜的外部蛋白质)，编码复制蛋白(例如聚合酶)的多核苷酸或其任何组合。

[0052] 在一些实施例中，分离指环体(例如，合成的指环体)，例如从宿主细胞中分离和/或从溶液(例如，上清液)中的其他成分中分离。在一些实施例中，例如从溶液(例如上清液)中纯化指环体(例如，合成的指环体)。在一些实施例中，相对于溶液中的其他成分，富集溶液中的指环体。

[0053] 在任何前述指环体、指环载体、组合物或方法的一些实施例中，遗传元件包含指环体基因组，例如，根据实例9中所述的方法鉴定。在实施例中，指环体基因组包含TTV-tth8核

酸序列，例如表5中所示的TTV-tth8核酸序列，其具有TTV-tth8核酸序列的核苷酸3436-3707的至少10％、20％、30％、40％、50％、60％、70％、80％、90％、95％、99％或100％的缺失。在实施例中，指环体基因组包含TTMV-LY2核酸序列，例如表15中所示的TTMV-LY2核酸序列，其具有TTMV-LY2核酸序列的核苷酸574-1371、1432-2210、574-2210和/或2610-2809的至少10％、20％、30％、40％、50％、60％、70％、80％、90％、95％、99％或100％的缺失。在实施例中，指环体基因组是能够自我复制和/或自我扩增的指环体基因组。在实施例中，指环体基因组不能够自我复制和/或自我扩增。在实施例中，指环体基因组能够以反式复制和/或扩增，例如在辅助物例如辅助病毒存在下。

[0054] 任何前述指环体、指环载体、组合物或方法的附加特征包括以下列举的实施例中的一个或多个。

[0055] 本领域技术人员将认知到、或不使用过度常规实验就能够确定本案所述发明的特定实施例的许多等效实例。此类等效实施例意在由以下列举的实施例涵盖。

[0056] 列举的实施例

[0057] 1.一种指环体，所述指环体包含：

[0058] (a)蛋白质外部；

[0059] (b)遗传元件，其包含启动子元件和编码外源性效应子的核酸序列(例如DNA序列)和蛋白质结合序列(例如外部蛋白质结合序列)；

[0060] 其中所述外源性效应子包含当突变时引起人疾病的多肽，或所述多肽的功能变体；

[0061] 其中所述遗传元件被封闭在所述蛋白质外部内；并且

[0062] 其中指环体被配置为将遗传元件递送到真核细胞中；

[0063] 任选地，其中所述遗传元件包含相对于野生型指环病毒基因组序列(例如，如本文所述)的至少一个差异(例如，突变、化学修饰或表观遗传改变)，例如插入、取代、酶促修饰和/或缺失，例如结构域(例如TATA框、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号或GC富集区中的一个或多个)的缺失。

[0064] 2.一种指环体，所述指环体包含：

[0065] (a)蛋白质外部；

[0066] (b)遗传元件，其包含可操作地连接至编码效应子(例如内源性效应子或外源性效应子)的异源核酸序列(例如，DNA序列)的启动子元件；

[0067] 其中所述外源性效应子包含当突变时引起人疾病的多肽，或所述多肽的功能变体；

[0068] 其中所述遗传元件被封闭在所述蛋白质外部内；

[0069] 任选地，其中所述遗传元件包含相对于野生型指环病毒基因组序列(例如，如本文所述)的至少一个差异(例如，突变、化学修饰或表观遗传改变)，例如插入、取代、酶促修饰和/或缺失，例如结构域(例如TATA框、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号或GC富集区中的一个或多个)的缺失。

[0070] 3.一种指环体，所述指环体包含：

[0071] (a)蛋白质外部；

[0072] (b)遗传元件，其包含启动子元件和编码外源性效应子的核酸序列(例如DNA序列)

和蛋白质结合序列(例如外部蛋白质结合序列)；

[0073] 其中所述外源性效应子包含酶或结构蛋白质；

[0074] 其中所述遗传元件被封闭在所述蛋白质外部内;并且

[0075] 其中指环体被配置为将遗传元件递送到真核细胞中；

[0076] 任选地,其中所述遗传元件包含相对于野生型指环病毒基因组序列(例如,如本文所述)的至少一个差异(例如,突变、化学修饰或表观遗传改变),例如插入、取代、酶促修饰和/或缺失,例如结构域(例如TATA框、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号或GC富集区中的一个或多个)的缺失。

[0077] 4.如前述实施例中任一项所述的指环体,其中所述效应子包含作用于人细胞内源性小分子底物的酶。

[0078] 5.如前述实施例中任一项所述的指环体,其中所述效应子包含神经元或肌肉细胞内源性的多肽,或所述多肽的功能变体。

[0079] 6.如前述实施例中任一项所述的指环体,其中所述效应子包含酶、结构蛋白质或补体蛋白质。

[0080] 7.如前述实施例中任一项所述的指环体,其中所述效应子是表A的酶或其功能变体,或其中所述效应子是表B的多肽或其功能变体。

[0081] 8.如前述实施例中任一项所述的指环体,其中所述效应子包含野生型人蛋白质,例如野生型SMN、肌营养不良蛋白、UGT1A1、OCA1或C1。

[0082] 9.如前述实施例中任一项所述的指环体,其中所述人疾病是导致与健康受试者中的底物的水平相比所述底物的水平升高的疾病。

[0083] 10.如前述实施例中任一项所述的指环体,其中所述人疾病是由与健康受试者的结构蛋白质的水平相比所述结构蛋白质的水平降低引起的疾病。

[0084] 11.一种治疗受试者的疾病或障碍的方法,所述方法包括向所述受试者施用有效量的指环体组合物,其中所述指环体组合物包含多个指环体,所述多个指环体包含：

[0085] (a)蛋白质外部；

[0086] (b)遗传元件,其包含启动子元件和编码效应子(例如外源性效应子或内源性效应子)的核酸序列(例如,DNA序列),和蛋白质结合序列(例如,外部蛋白质结合序列)；

[0087] 其中所述效应子包含当突变时引起人疾病的多肽,或所述多肽的功能变体；

[0088] 其中所述遗传元件被封闭在所述蛋白质外部内;并且

[0089] 任选地,其中所述遗传元件包含相对于野生型指环病毒基因组序列(例如,如本文所述)的至少一个差异(例如,突变、化学修饰或表观遗传改变),例如插入、取代、酶促修饰和/或缺失,例如结构域(例如TATA框、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号或GC富集区中的一个或多个)的缺失。

[0090] 12.如实施例11所述的方法,其中所述疾病或障碍是遗传疾病或障碍,例如除干扰素病(interferonopathy)之外的遗传疾病。

[0091] 13.如实施例11所述的方法,其中所述疾病或障碍是神经肌肉疾病(例如,脊髓性肌萎缩症)、肌营养不良症(例如,杜氏肌营养不良症或贝克尔(Becker)肌营养不良症)、酶促不足(例如,克里格勒-纳贾尔(Crigler-Najjar)综合征)、或白化病(例如眼白化病)、或C1缺乏症。

[0092] 14.一种治疗受试者的遗传疾病或障碍的方法,所述方法包括向所述受试者施用有效量的指环体组合物,其中所述指环体组合物包含多个指环体,所述多个指环体包含:

[0093] (a)蛋白质外部;

[0094] (b)遗传元件,其包含启动子元件和编码效应子(例如外源性效应子或内源性效应子)的核酸序列(例如,DNA序列),和蛋白质结合序列(例如,外部蛋白质结合序列);

[0095] 其中所述遗传元件被封闭在所述蛋白质外部内;并且

[0096] 任选地,其中所述遗传元件包含相对于野生型指环病毒基因组序列(例如,如本文所述)的至少一个差异(例如,突变、化学修饰或表观遗传改变),例如插入、取代、酶促修饰和/或缺失,例如结构域(例如TATA框、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号或GC富集区中的一个或多个)的缺失。

[0097] 15.如实施例14所述的方法,其中所述效应子包含酶、结构蛋白质或补体蛋白质。

[0098] 16.一种向需要酶活性的受试者递送酶活性的方法,所述方法包括向所述受试者施用有效量的指环体组合物,

[0099] 其中所述指环体组合物包含多个指环体,所述多个指环体包含:

[0100] (a)蛋白质外部;

[0101] (b)遗传元件,其包含启动子元件和编码效应子(例如内源性效应子或外源性效应子)的核酸序列(例如,DNA序列),和蛋白质结合序列(例如,外部蛋白质结合序列);

[0102] 任选地,其中所述遗传元件包含相对于野生型指环病毒基因组序列(例如,如本文所述)的至少一个差异(例如,突变、化学修饰或表观遗传改变),例如插入、取代、酶促修饰和/或缺失,例如结构域(例如TATA框、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号或GC富集区中的一个或多个)的缺失;

[0103] 其中所述遗传元件被封闭在所述蛋白质外部内;并且

[0104] 从而将所述酶活性递送给所述受试者。

[0105] 17.一种恢复(至少部分恢复)受试者中的功能,例如恢复至野生型表型的方法,所述方法包括向所述受试者施用有效量的指环体组合物,

[0106] 其中所述指环体组合物包含多个指环体,所述多个指环体包含:

[0107] (a)蛋白质外部;

[0108] (b)遗传元件,其包含启动子元件和编码效应子(例如内源性效应子或外源性效应子)的核酸序列(例如,DNA序列),和蛋白质结合序列(例如,外部蛋白质结合序列);

[0109] 任选地,其中所述遗传元件包含相对于野生型指环病毒基因组序列(例如,如本文所述)的至少一个差异(例如,突变、化学修饰或表观遗传改变),例如插入、取代、酶促修饰和/或缺失,例如结构域(例如TATA框、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号或GC富集区中的一个或多个)的缺失;

[0110] 其中所述遗传元件被封闭在所述蛋白质外部内;并且

[0111] 从而恢复所述受试者中的所述功能。

[0112] 18.如实施例11-17中任一项所述的方法,其中所述效应子共价修饰与健康受试者中的底物的水平相比在所述受试者中以升高的水平存在的所述底物。

[0113] 19.如实施例11-18中任一项所述的方法,其中所述效应子是在所述受试者中突变的蛋白质的功能变体,例如,是对应于在所述受试者中突变的蛋白质的野生型蛋白质。

[0114] 20.如实施例11-19中任一项所述的方法,其中所述效应子是在所述受试者中突变的蛋白质的功能同源物(例如,直向同源物或旁系同源物)。

[0115] 21.如实施例11-20中任一项所述的方法,其中所述效应子共价修饰底物,其中所述底物也被与在所述受试者中突变的蛋白质对应的野生型蛋白质修饰。

[0116] 22.如实施例11-21中任一项所述的方法,其中所述效应子催化与所述受试者中突变的蛋白质对应的野生型蛋白质相同的化学反应。

[0117] 23.如实施例11-22中任一项所述的方法,其中所述效应子包含结构蛋白质、酶或补体蛋白质。

[0118] 24.一种向受试者递送效应子的方法,所述方法包括向所述受试者施用有效量的指环体组合物,

[0119] 其中所述指环体组合物包含多个指环体,所述多个指环体包含:

[0120] (a)蛋白质外部;

[0121] (b)遗传元件,其包含启动子元件和编码效应子(例如内源性效应子或外源性效应子)的核酸序列(例如,DNA序列),和蛋白质结合序列(例如,外部蛋白质结合序列);

[0122] 其中所述效应子包含当突变时引起人疾病的多肽,或所述多肽的功能变体;

[0123] 任选地,其中所述遗传元件包含相对于野生型指环病毒基因组序列(例如,如本文所述)的至少一个差异(例如,突变、化学修饰或表观遗传改变),例如插入、取代、酶促修饰和/或缺失,例如结构域(例如TATA框、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号或GC富集区中的一个或多个)的缺失;

[0124] 其中所述遗传元件被封闭在所述蛋白质外部内;并且

[0125] 从而将所述效应子递送给所述受试者。

[0126] 25.一种调节,例如抑制或增强受试者的生物学功能的方法,所述受试者是例如患有可通过调节所述受试者的生物学功能治疗的疾病或障碍的受试者,所述方法包括向所述受试者施用有效量的指环体组合物,

[0127] 其中所述指环体组合物包含多个指环体,所述多个指环体包含:

[0128] (a)蛋白质外部;

[0129] (b)遗传元件,其包含启动子元件和编码效应子(例如内源性效应子或外源性效应子)的核酸序列(例如,DNA序列),和蛋白质结合序列(例如,外部蛋白质结合序列);

[0130] 其中所述效应子包含当突变时引起人疾病的多肽,或所述多肽的功能变体;

[0131] 任选地,其中所述遗传元件包含相对于野生型指环病毒基因组序列(例如,如本文所述)的至少一个差异(例如,突变、化学修饰或表观遗传改变),例如插入、取代、酶促修饰和/或缺失,例如结构域(例如TATA框、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号或GC富集区中的一个或多个)的缺失;

[0132] 其中所述遗传元件被封闭在所述蛋白质外部内;并且

[0133] 从而调节,例如,抑制或增强所述受试者的生物学功能。

[0134] 26.如前述实施例中任一项所述的指环体或方法,其中所述效应子包含SMN、肌营养不良蛋白、UGT1A1、OCA1、C1或前述任一项的功能变体。

[0135] 27.如前述实施例中任一项所述的指环体或方法,其中所述效应子是分泌型多肽(例如,所述效应子包含信号序列,例如效应子内源性的信号序列,或异源信号序列),或所

述效应子是细胞内多肽。

[0136] 28.一种治疗受试者的疾病或障碍的方法,所述方法包括向所述受试者施用有效量的指环体组合物或分离的核酸分子(例如,表达载体),

[0137] 其中所述指环体组合物或分离的核酸分子包含遗传元件,所述遗传元件包含启动子元件和编码效应子(例如,外源性效应子或内源性效应子)(例如各自如本文所述)的核酸序列(例如,DNA序列),并且其中

[0138] (i)所述疾病或障碍是或包括神经肌肉疾病,例如脊髓性肌萎缩症,并且所述效应子包含SMN或其功能变体;

[0139] (ii)所述疾病或障碍是肌营养不良症(例如,杜氏肌营养不良症或贝克尔肌营养不良症),并且所述效应子包含肌营养不良蛋白或其功能变体,例如微肌营养不良蛋白;

[0140] (iii)所述疾病或障碍是酶促不足,例如克里格勒-纳贾尔综合征,并且所述效应子包含尿苷二磷酸葡糖醛酸转移酶,例如UGT1A1或其功能变体;

[0141] (iv)所述疾病或障碍是白化病,例如眼白化病,并且所述效应子包含酪氨酸酶,例如OCA1或其功能变体;

[0142] (v)所述疾病或障碍是C1补体缺乏症,并且所述效应子包含C1或其功能变体;或

[0143] (vi)所述疾病或障碍是表A的疾病或障碍,并且所述效应子包含表A的酶或其功能变体;

[0144] (vii)所述疾病或障碍是表B的疾病或障碍,并且所述效应子包含表B的多肽或其功能变体;

[0145] 从而治疗所述受试者的疾病或障碍。

[0146] 29.一种向患有疾病或障碍的受试者递送效应子的方法,所述方法包括向所述受试者施用有效量的指环体组合物或分离的核酸分子(例如,表达载体),

[0147] 其中所述指环体组合物或分离的核酸分子包含遗传元件,所述遗传元件包含启动子元件和编码效应子(例如,外源性效应子或内源性效应子)的核酸序列(例如,DNA序列);并且

[0148] 其中:

[0149] (i)所述疾病或障碍包括神经肌肉疾病,例如脊髓性肌萎缩症,并且所述效应子包含SMN或其功能变体;

[0150] (ii)所述疾病或障碍是肌营养不良症(例如,杜氏肌营养不良症或贝克尔肌营养不良症),并且所述效应子包含肌营养不良蛋白或其功能变体,例如微肌营养不良蛋白;

[0151] (iii)所述疾病或障碍是酶促不足,例如克里格勒-纳贾尔综合征,并且所述效应子包含尿苷二磷酸葡糖醛酸转移酶,例如UGT1A1或其功能变体;

[0152] (iv)所述疾病或障碍是白化病,例如眼白化病,并且所述效应子包含酪氨酸酶,例如OCA1或其功能变体;

[0153] (v)所述疾病或障碍是C1补体缺乏症,并且所述效应子包含C1或其功能变体;或

[0154] (vi)所述疾病或障碍是表A的疾病或障碍,并且所述效应子包含表A的酶或其功能变体;

[0155] (vii)所述疾病或障碍是表B的疾病或障碍,并且所述效应子包含表B的多肽或其功能变体;

[0156] 从而将所述效应子递送给所述受试者。

[0157] 30.一种制备指环体组合物的方法,所述方法包括:

[0158] a)提供宿主细胞,其包含一种或多种编码如前述实施例中任一项所述的指环体的组分的核酸分子;

[0159] b)在允许细胞产生一种或多种指环体的条件下维持(例如,培养)所述宿主细胞,从而制造指环体;以及

[0160] c)将所述指环体配制为例如适于向受试者施用的药物组合物。

[0161] 31.一种制备指环体组合物的方法,所述方法包括:

[0162] a)提供多个如前述实施例中任一项所述的指环体;

[0163] b)任选地对所述多个如前述实施例中任一项所述的指环体评估以下一种或多种:本文所述的污染物、光密度测量值(例如OD 260)、颗粒数(例如通过HPLC)、感染性(例如颗粒:感染单位比);并且

[0164] c)例如,如果(b)的一种或多种参数满足指定的阈值,则将所述多个指环体例如配制为适于施用于受试者的药物组合物。

[0165] 32.如实施例31所述的方法,其中所述指环体组合物包含至少10^5、10^6、10^7、10^8、10^9、10^{10}、10^{11}、10^{12}、10^{13}、10^{14}或10^{15}个指环体。

[0166] 33.如实施例31或32所述的方法,其中所述指环体组合物包含至少10ml、20ml、50ml、100ml、200ml、500ml、1L、2L、5L、10L、20L或50L。

[0167] 34.如前述实施例中任一项所述的指环体或方法,其中所述效应子包含在突变时引起人疾病的多肽(例如,SMN、肌营养不良蛋白、UGT1A1、OCA1或C1)或所述多肽的功能变体。

[0168] 35.如前述实施例中任一项所述的指环体或方法,其中所述效应子包含共价修饰人细胞内源性小分子底物的酶(例如尿苷二磷酸葡糖醛酸转移酶,例如UGT1A1;或酪氨酸酶,例如,OCA1)。

[0169] 36.如前述实施例中任一项的指环体或方法,其中所述效应子包含对神经元(例如,SMN)或肌肉细胞(例如,肌营养不良蛋白)而言是内源性的多肽,或所述多肽的功能变体。

[0170] 37.如前述实施例中任一项所述的指环体或方法,其中所述效应子包含补体蛋白质(例如,C1)。

[0171] 38.如前述实施例中任一项所述的指环体或方法,其中所述遗传元件包含相对于野生型指环病毒基因组序列(例如,如本文所述)的至少一个差异(例如,突变、化学修饰或表观遗传改变),例如插入、取代、酶促修饰和/或缺失,例如结构域(例如TATA框、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号或GC富集区中的一个或多个)的缺失。

[0172] 39.如前述实施例中任一项所述的指环体或方法,其中所述遗传元件包含区域,所述区域包含以下核酸序列的至少10、15、20、25、30、31、32、33、34、35或36个连续核苷酸:

[0173] (i)CGCGCTGCGCGCGCCGCCCAGTAGGGGGAGCCATGC(SEQ ID NO:160),

[0174] (ii)GCGCTX$_1$CGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:164),其中X$_1$选自T、G或A;

[0175] (iii)GCGCTTCGCGCGCCGCCCACTAGGGGGCGTTGCGCG(SEQ ID NO:165);

[0176] (iv)GCGCTGCGCGCGCCGCCCAGTAGGGGGCGCAATGCG(SEQ ID NO:166)；

[0177] (v)GCGCTGCGCGCGCGGCCCCCGGGGGAGGCATTGCCT(SEQ ID NO:167)；

[0178] (vi)GCGCTGCGCGCGCGCCGGGGGGGCGCCAGCGCCC(SEQ ID NO:168)；

[0179] (vii)GCGCTTCGCGCGCGCCGGGGGGCTCCGCCCCCCC(SEQ ID NO:169)；

[0180] (viii)GCGCTTCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:170)；

[0181] (ix)GCGCTACGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:171)；或

[0182] (x)GCGCTACGCGCGCGCCGGGGGGCTCTGCCCCCCC(SEQ ID NO:172)；

[0183] 或与其具有至少75％、76％、77％、78％、79％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。

[0184] 40.如前述实施例中任一项所述的指环体或方法，其中所述遗传元件包含序列，所述序列包含具有至少70％、71％、72％、73％、74％、75％、76％、77％、78％、79％、80％或80.6％的GC含量的至少20、25、30、31、32、33、34、35或36个连续核苷酸。

[0185] 41.如实施例40所述的指环体或方法，其中所述遗传元件包含具有至少80％的GC含量的至少20、25、30、31、32、33、34、35或36个连续核苷酸。

[0186] 42.如实施例40所述的指环体或方法，其中所述遗传元件包含具有至少70％、71％、72％、73％、74％、75％、76％、77％、78％、79％、80％或80.6％的GC含量的至少36个连续核苷酸。

[0187] 43.如实施例40所述的指环体或方法，其中所述遗传元件包含具有至少80％的GC含量的至少36个连续核苷酸。

[0188] 1000.一种多肽，例如ORF1分子，其包含以下中的一项或多项：

[0189] (a)第一区域，其包含与本文描述的精氨酸富集区序列(例如MPYYYRRRRYNYRRPRWYGRGWIRRPFRRRFRRKRRVR(SEQ ID NO:216)或MAWGWWKRRRRWWFRKRWTRGRLRRRWPRSARRRPRRRRVRRRRRW RRGRRKTRTYRRRRRFRRRGRK(SEQ ID NO:186)，或如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所列)具有至少70％(例如，至少约70％、80％、90％、95％、96％、97％、98％、99％或100％)序列同一性的氨基酸序列，或者含有至少60％、70％或80％碱性残基(例如，精氨酸、赖氨酸或其组合)的至少约40个氨基酸的序列，

[0190] (b)第二区域，其包含与本文所述的胶冻卷区域序列(例如PTYTTIPLKQWQPPYKRTCYIKGQDCLIYYSNLRLGMNSTMYEKSIVPVHWPGGGSFSVSMLTLDALYDIHKLCRNWWTSTNQDLPLVRYKGCKITFYQSTFTDYIVRIHTELPANSNKLTYPNTHPLMMMMSKYKHIIPSRQTRRKKKPYTKIFVKPPPQFENKWYFATDLYKIPLLQIHCTACNLQNPFVKPDKLSNNVTLWSLNT(SEQ ID NO:217)，或如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所列)具有至少30％(例如至少约30％、35％、40％、50％、60％、70％、80％、90％、95％、96％、97％、98％、99％或100％)序列同一性的氨基酸序列或含有至少6个(例如，至少6、7、8、9、10、11或12个)β链的序列；

[0191] (c)第三区域，其包含与本文所述N22结构域序列(例如TMALTPFNEPIFTQIQYNPDRDTGEDTQLYLLSNATGTGWDPPGIPELILEGFPLWLIYWGFADFQKNLKKVTNIDTNYMLVAKTKFTQKPGTFYLVILNDTFVEGNSPYEKQPLPEDNIKWYPQVQYQLEAQNKLLQTGPFTPNIQGQLSDNISMFYKFYFK(SEQ ID NO:219)，或如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的

任一个中所列)具有至少30％(例如至少约30％、35％、40％、50％、60％、70％、80％、90％、95％、96％、97％、98％、99％或100％)序列同一性的氨基酸序列;以及

[0192] (d)第四区域,其包含与本文所述的指环病毒病毒ORF1 C末端结构域(CTD)序列(例如WGGSPPKAINVENPAHQIQYPIPRNEHETTSLQSPGEAPESILYSFDYRHGNYTTTALSRISQDWALKDTVSKITEPDRQQLLKQALECLQISEETQEKKEKEVQQLISNLRQQQQLYRERIISLLKDQ(SEQ ID NO:220),或如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所列)具有至少30％(例如至少约30％、35％、40％、50％、60％、70％、80％、90％、95％、96％、97％、98％、99％或100％)序列同一性的氨基酸序列;

[0193] 其中所述ORF1分子包含相对于野生型ORF1蛋白(例如,如本文所述)的至少一个差异(例如,突变、化学修饰或表观遗传改变),例如插入、取代、化学或酶促修饰,和/或缺失,例如结构域(例如,精氨酸富集区、胶冻卷结构域、HVR、N22或CTD中的一个或多个,例如如本文所述)的缺失。

[0194] 1000A.如实施例1000所述的多肽,其中(a)、(b)、(c)和(d)的区域的氨基酸序列与其各自的参考具有至少90％序列同一性。

[0195] 1001.如实施例1000所述的多肽,其中所述多肽包含:

[0196] (i)所述第一区域和所述第二区域;

[0197] (ii)所述第一区域和所述第三区域;

[0198] (iii)所述第一区域和所述第四区域;

[0199] (iv)所述第二区域和所述第三区域;

[0200] (v)所述第二区域和所述第四区域;

[0201] (vi)所述第三区域和所述第四区域;

[0202] (vii)所述第一区域、所述第二区域和所述第三区域;

[0203] (viii)所述第一区域、所述第二区域和所述第四区域;

[0204] (ix)所述第一区域、所述第三区域和所述第四区域;或

[0205] (x)所述第二区域、所述第三区域和所述第四区域。

[0206] 1002.一种多肽,例如ORF1分子,其包含:

[0207] (a)第一区域,其包含与本文描述的精氨酸富集区序列(例如MPYYYRRRRYNYRRPRWYGRGWIRRPFRRRFRRKRRVR(SEQ ID NO:216)或MAWGWWKRRRRWWFRKRWTRGRLRRRWPRSARRRPRRRRVRRRRRWRRGRRKTRTYRRRRRFRRRGRK(SEQ ID NO:186),或如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所列)具有至少70％(例如,至少约70％、80％、90％、95％、96％、97％、98％、99％或100％)序列同一性的氨基酸序列,或者含有至少60％、70％或80％碱性残基(例如,精氨酸、赖氨酸或其组合)的至少约40个氨基酸的序列,

[0208] (b)第二区域,其包含与本文所述的胶冻卷区域序列(例如PTYTTIPLKQWQPPYKRTCYIKGQDCLIYYSNLRLGMNSTMYEKSIVPVHWPGGGSFSVSMLTLDALYDIHKLCRNWWTSTNQDLPLVRYKGCKITFYQSTFTDYIVRIHTELPANSNKLTYPNTHPLMMMMSKYKHIIPSRQTRRKKKPYTKIFVKPPPQFENKWYFATDLYKIPLLQIHCTACNLQNPFVKPDKLSNNVTLWSLNT(SEQ ID NO:217),或如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所列)具有至少30％(例如至少约30％、35％、40％、50％、60％、70％、80％、90％、95％、96％、97％、98％、99％或

100%)序列同一性的氨基酸序列或含有至少6个β链的序列;

[0209] (c)第三区域,其包含与本文所述N22结构域序列(例如TMALTPFNEPIFTQIQYNPDRDTGEDTQLYLLSNATGTGWDPPGIPELILEGFPLWLIYWGFADFQKNLKKVTNIDTNYMLVAKTKFTQKPGTFYLVILNDTFVEGNSPYEKQPLPEDNIKWYPQVQYQLEAQNKLLQTGPFTPNIQGQLSDNISMFYKFYFK(SEQ ID NO:219),或如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所列)具有至少30%(例如至少约30%、35%、40%、50%、60%、70%、80%、90%、95%、96%、97%、98%、99%或100%)序列同一性的氨基酸序列;以及

[0210] (d)第四区域,其包含与本文所述的指环病毒病毒ORF1 C末端结构域(CTD)序列(例如WGGSPPKAINVENPAHQIQYPIPRNEHETTSLQSPGEAPESILYSFDYRHGNYTTTALSRISQDWALKDTVSKITEPDRQQLLKQALECLQISEETQEKKEKEVQQLISNLRQQQQLYRERIISLLKDQ(SEQ ID NO:220),或如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所列)具有至少30%(例如至少约30%、35%、40%、50%、60%、70%、80%、90%、95%、96%、97%、98%、99%或100%)序列同一性的氨基酸序列;

[0211] 其中所述ORF1分子包含相对于野生型ORF1蛋白(例如,如本文所述)的至少一个差异(例如,突变、化学修饰或表观遗传改变),例如插入、取代、化学或酶促修饰,和/或缺失,例如结构域(例如,精氨酸富集区、胶冻卷结构域、HVR、N22或CTD中的一个或多个,例如如本文所述)的缺失。

[0212] 1002A.如实施例1002所述的多肽,其中(a)、(b)、(c)和(d)区域的氨基酸序列与其各自的参考具有至少90%序列同一性。

[0213] 1003.如前述实施例中任一项所述的多肽,其中:

[0214] 所述第一区域包含与表16中列出的ORF1序列的氨基酸1-38具有至少70%(例如至少约70%、80%、90%、95%、96%、97%、98%、99%或100%)序列同一性的氨基酸序列;

[0215] 所述第二区域包含与表16中列出的ORF1序列的氨基酸39-246具有至少70%(例如至少约70%、80%、90%、95%、96%、97%、98%、99%或100%)序列同一性的氨基酸序列;

[0216] 所述第二区域包含与表16中列出的ORF1序列的氨基酸375-537具有至少70%(例如至少约70%、80%、90%、95%、96%、97%、98%、99%或100%)序列同一性的氨基酸序列;和/或

[0217] 所述第四区域包含与表16中列出的ORF1序列的氨基酸538-666具有至少70%(例如至少约70%、80%、90%、95%、96%、97%、98%、99%或100%)序列同一性的氨基酸序列。

[0218] 1003A.如实施例1003所述的多肽,其中所述第一、第二、第三和第四区域的氨基酸序列与其各自的参考具有至少90%序列同一性。

[0219] 1004.如前述实施例中任一项所述的多肽,其中:

[0220] 所述第一区域包含与如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所列的精氨酸富集区序列具有至少70%(例如,至少约70%、80%、90%、95%、96%、97%、98%、99%或100%)序列同一性的氨基酸序列;

[0221] 所述第二区域包含与如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所列的胶冻卷区域序列具有至少70%(例如,至少约70%、80%、90%、95%、96%、97%、98%、99%或100%)序列同一性的氨基酸序列;

[0222] 所述第三区域包含与如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所列的N22结构域序列具有至少70％(例如，至少约70％、80％、90％、95％、96％、97％、98％、99％或100％)序列同一性的氨基酸序列；和/或

[0223] 所述第四区域包含与如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所列的CTD序列具有至少70％(例如，至少约70％、80％、90％、95％、96％、97％、98％、99％或100％)序列同一性的氨基酸序列。

[0224] 1004A.如实施例1004所述的多肽，其中所述第一、第二、第三和第四区域的氨基酸序列与其各自的参考具有至少90％序列同一性。

[0225] 1005.如前述实施例中任一项所述的多肽，其中所述多肽以N末端到C末端的顺序包含所述第一区域、所述第二区域、所述第三区域和所述第四区域。

[0226] 1006.如前述实施例中任一项所述的多肽，其中所述至少一个差异包括所述第一区域相对于野生型ORF1蛋白的精氨酸富集区的至少一个差异。

[0227] 1007.如前述实施例中任一项所述的多肽，其中所述第一区域包含来自与所述多肽或其不包括所述第一区域的部分具有最大序列同一性的野生型指环病毒以外的指环病毒的ORF1蛋白的精氨酸富集区。

[0228] 1008.如前述实施例中任一项所述的多肽，其中所述第一区域包含氨基酸序列，所述氨基酸序列与来自除野生型指环病毒(所述多肽与其具有最大序列同一性)之外的指环病毒的精氨酸富集区具有至少70％序列同一性。

[0229] 1009.如前述实施例中任一项所述的多肽，其中所述第一区域包含多肽，所述多肽与野生型指环病毒基因组(例如，如本文所述)或其与所述第一区域具有相同氨基酸长度的部分具有小于15％(例如，小于15％、14％、13％、12％、11％、10％、9％、8％、7％、6％、5％、4％、3％、2％或1％)序列同一性。

[0230] 1010.如前述实施例中任一项所述的多肽，其中所述第一区域具有DNA结合活性和/或核定位活性。

[0231] 1011.如前述实施例中任一项所述的多肽，其中所述第一区域包含DNA结合区域和/或核定位序列。

[0232] 1012.如前述实施例中任一项所述的多肽，其中所述至少一个差异包括所述第二区域中相对于野生型ORF1蛋白的胶冻卷区域的至少一个差异。

[0233] 1013.如前述实施例中任一项所述的多肽，其中所述第二区域包含来自与所述多肽或其不包括所述第二区域的部分具有最大序列同一性的野生型指环病毒以外的指环病毒的ORF1蛋白的胶冻卷区域。

[0234] 1014.如前述实施例中任一项所述的多肽，其中所述第二区域包含氨基酸序列，所述氨基酸序列与来自除野生型指环病毒(所述多肽与其具有最大序列同一性)之外的指环病毒的胶冻卷区具有至少70％序列同一性。

[0235] 1015.如前述实施例中任一项所述的多肽，其中所述第二区域包含多肽，所述多肽与野生型指环病毒基因组(例如，如本文所述)或其与所述第二区域具有相同氨基酸长度的部分具有小于15％(例如，小于15％、14％、13％、12％、11％、10％、9％、8％、7％、6％、5％、4％、3％、2％或1％)序列同一性。

[0236] 1016.如前述实施例中任一项所述的多肽，其中所述至少一个差异包括所述第三

区域中相对于野生型ORF1蛋白的N22结构域的至少一个差异。

[0237] 1017.如前述实施例中任一项所述的多肽，其中所述第三区域包含来自与所述多肽或其不包括所述第三区域的部分具有最大序列同一性的野生型指环病毒以外的指环病毒的ORF1蛋白的N22结构域。

[0238] 1018.如前述实施例中任一项所述的多肽，其中所述第三区域包含氨基酸序列，所述氨基酸序列与来自除野生型指环病毒(所述多肽与其具有最大序列同一性)之外的指环病毒的N22区具有至少70％序列同一性。

[0239] 1019.如前述实施例中任一项所述的多肽，其中所述第三区域包含多肽，所述多肽与野生型指环病毒基因组(例如，如本文所述)或其与所述第三区域具有相同氨基酸长度的部分具有小于15％(例如，小于15％、14％、13％、12％、11％、10％、9％、8％、7％、6％、5％、4％、3％、2％或1％)序列同一性。

[0240] 1020.如前述实施例中任一项所述的多肽，其中所述至少一个差异包括所述第四区域中相对于野生型ORF1蛋白的CTD结构域的至少一个差异。

[0241] 1021.如前述实施例中任一项所述的多肽，其中所述第四区域包含来自与所述多肽或其不包括所述第四区域的部分具有最大序列同一性的野生型指环病毒以外的指环病毒的ORF1蛋白的CTD结构域。

[0242] 1022.如前述实施例中任一项所述的多肽，其中所述第四区域包含氨基酸序列，所述氨基酸序列与来自除野生型指环病毒(所述多肽与其具有最大序列同一性)之外的指环病毒的CTD区具有至少70％序列同一性。

[0243] 1023.如前述实施例中任一项所述的多肽，其中所述第四区域包含多肽，所述多肽与野生型指环病毒基因组(例如，如本文所述)或其与所述第四区域具有相同氨基酸长度的部分具有小于15％(例如，小于15％、14％、13％、12％、11％、10％、9％、8％、7％、6％、5％、4％、3％、2％或1％)序列同一性。

[0244] 1024.如前述实施例中任一项所述的多肽，所述多肽进一步包含氨基酸序列，例如高变区(HVR)序列(例如，如本文所述的指环病毒ORF1分子的HVR序列)，其中所述氨基酸序列包含至少约55(例如，至少约45、50、51、52、53、54、55、56、57、58、59、60或65)个氨基酸(例如，约45-160、50-160、55-160、60-160、45-150、50-150、55-150、60-150、45-140、50-140、55-140或60-140个氨基酸)。

[0245] 1025.如实施例1024所述的多肽，其中所述HVR序列位于所述第二区域和所述第三区域之间。

[0246] 1026.如实施例1024或1025所述的多肽，其中所述HVR序列包含氨基酸序列，所述氨基酸序列与来自跟所述ORF1蛋白具有最大序列同一性的野生型指环病毒以外的指环病毒的HVR具有至少70％(例如，至少约70％、80％、90％、95％、96％、97％、98％、99％或100％)序列同一性。

[0247] 1027.如实施例1024-1026中任一项所述的多肽，其中所述HVR序列相对于所述第一区域、第二区域、第三区域和/或第四区域中的一个或多个是异源的。

[0248] 1028.如实施例1024-1027中任一项所述的多肽，其中所述至少一个差异包括所述HVR序列中相对于野生型ORF1蛋白(例如，来自野生型指环病毒基因组，例如如本文所述)的HVR序列的至少一个差异。

[0249] 1029.如实施例1024-1028中任一项所述的多肽，其中所述HVR序列包含来自与所述多肽或其不包括所述HVR类型的部分具有最大序列同一性的野生型指环病毒以外的指环病毒的ORF1蛋白的HVR。

[0250] 1030.如实施例1024-1029中任一项所述的多肽，其中所述HVR序列包含氨基酸序列，所述氨基酸序列与来自跟所述多肽具有最大序列同一性的野生型指环病毒以外的指环病毒的HVR具有至少约70％序列同一性。

[0251] 1031.如实施例1024-1030中任一项所述的多肽，其中所述HVR包含与如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所列的HVR序列具有至少70％(例如，至少约70％、80％、90％、95％、96％、97％、98％、99％或100％)序列同一性的氨基酸序列。

[0252] 1032.如实施例1024-1031中任一项所述的多肽，其中所述HVR序列包含与表16中列出的ORF1序列的氨基酸247-374的至少70％(例如至少约70％、80％、90％、95％、96％、97％、98％、99％或100％)序列同一性。

[0253] 1033.如前述实施例中任一项所述的多肽，所述多肽进一步包含异源多肽，例如以下多肽，其相对于所述第一区域、第二区域、第三区域和/或第四区域中的一个或多个是异源，和/或相对于包含所述多肽的指环体是外源的。

[0254] 1034.如实施例1033所述的多肽，其中所述多肽缺乏指环病毒HVR序列。

[0255] 1035.如实施例1033所述的多肽，其中所述异源多肽存在于所述指环体的外部。

[0256] 1036.如实施例1033所述的多肽，其中所述异源多肽存在于所述指环体的内部。

[0257] 1037.如实施例1033-1036中任一项所述的多肽，其中所述异源多肽具有对所述指环体或野生型指环病毒而言是外源的功能性。

[0258] 1038.如实施例1033-1037中任一项所述的多肽，其中所述异源多肽由约140个或更少的氨基酸(例如，100、110、120、125、130、135、136、137、138、139、140、145、150、155或160个或更少的氨基酸)。

[0259] 1039.实施例1033-1038中任一项所述的多肽，其中所述异源多肽的大小相对于例如本文所述的指环病毒的野生型HVR区为50％-150％。

[0260] 1039A.如实施例1033-1039中任一项所述的多肽，其中所述异源多肽位于所述第二区域和所述第三区域之间。

[0261] 1040.如前述实施例中任一项所述的多肽，其进一步包含在所述第一区域和所述第二区域之间的一个或多个氨基酸、在所述第二区域和所述第三区域之间的一个或多个氨基酸和/或在所述第三区域和所述第四区域之间的一个或多个氨基酸。

[0262] 1041.如前述实施例中任一项所述的多肽，其进一步包含一个或多个相对于所述第一区域位于N末端的氨基酸。

[0263] 1042.如前述实施例中任一项所述的多肽，其进一步包含一个或多个相对于所述第四区域位于C末端的氨基酸。

[0264] 1043.如前述实施例中任一项所述的多肽，其包含多个子序列，所述多个子序列具有至少四个(例如，4、5、6、7、8、9、10、15、20、25或30个)连续氨基酸，与例如如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所列的野生型指环病毒ORF1氨基酸序列的相应子序列具有100％序列同一性。

[0265] 1044.如前述实施例中任一项所述的多肽，其包含多个子序列，所述多个子序列具有至少十个(例如，10、15、20、25、30、40或50个)连续氨基酸，与例如如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所列的野生型指环病毒ORF1氨基酸序列的相应子序列具有80％序列同一性。

[0266] 1045.如前述实施例中任一项所述的多肽，其包含多个子序列，所述多个子序列具有至少二十个(例如，20、25、30、40、50、60、70、80、90或100个)连续氨基酸，与例如如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所列的野生型指环病毒ORF1氨基酸序列的相应子序列具有60％序列同一性。

[0267] 1046.如实施例1043-1045任一项所述的多肽，其中所述多个子序列位于所述第一区域、第二区域、第三区域和/或第四区域内。

[0268] 1047.如前述实施例中任一项所述的多肽，其中所述第一区域包含至少约40个氨基酸(例如，至少约50、60、70、80、90或100个氨基酸，例如，约40-100、40-90、40-80、40-70、50-100、50-70、60-100、60-90、60-80或60-70个氨基酸)。

[0269] 1048.如前述实施例中任一项所述的多肽，其中所述第一区域包含至少约70％(例如，至少约70％、75％、80％、85％、90％、95％或100％)碱性残基(例如精氨酸、赖氨酸或其组合)。

[0270] 1049.如前述实施例中任一项所述的多肽，其中所述第一区域包含至少约70％(例如，至少约70％、75％、80％、85％、90％、95％或100％)精氨酸残基。

[0271] 1050.如前述实施例中任一项所述的多肽，其中所述多肽与所述多肽的另外拷贝形成同多聚体。

[0272] 1051.如实施例1050所述的多肽，其中所述第一区域结合所述多肽的另外拷贝上的相应第一区域。

[0273] 1052.如实施例1050所述的多肽，其中所述同多聚体形成衣壳，例如包封核酸，例如遗传元件或指环病毒基因组或其部分。

[0274] 1053.如前述实施例中任一项所述的多肽，其中所述多肽是衣壳蛋白或可形成衣壳的一部分。

[0275] 1054.如前述实施例中任一项所述的多肽，其中所述多肽具有复制酶活性。

[0276] 1055.如前述实施例中任一项所述的多肽，其中所述多肽结合核酸(例如，DNA)。

[0277] 1056.一种复合物，其包含：

[0278] (a)如前述实施例中任一项所述的多肽，以及

[0279] (b)遗传元件，其包含启动子元件和编码效应子(例如外源性效应子或内源性效应子)的核酸序列(例如，DNA序列)，和蛋白质结合序列。

[0280] 1057.一种复合物，其包含：

[0281] (a)ORF1分子，和

[0282] (b)遗传元件，其包含启动子元件和编码效应子(例如外源性效应子或内源性效应子)的核酸序列(例如，DNA序列)，和蛋白质结合序列；

[0283] 其中所述ORF1分子结合(例如，非共价结合)所述遗传元件，

[0284] 其中ORF1分子、遗传元件或ORF1分子和遗传元件两者包含相对于野生型ORF1蛋白、野生型指环病毒基因组或野生型ORF1蛋白和野生型指环病毒基因组两者(例如，如本文

所述)的至少一个差异(例如,突变、化学修饰或表观遗传改变),例如插入、取代、化学或酶促修饰,和/或缺失,例如结构域(例如,例如如本文所述的精氨酸富集区、胶冻卷结构域、HVR、N22或CTD中的一个或多个)或基因组区域(例如,例如如本文所述的TATA盒、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号或GC富集区中的一个或多个)的缺失。

[0285] 1058.如实施例1056或1057所述的复合物,其中所述复合物是体外的,例如,其中所述复合物处于基本上无细胞的组合物中。

[0286] 1059.如实施例1056-1058中任一项所述的复合物,其中所述复合物在细胞中,例如宿主细胞,例如辅助细胞,例如在细胞核中。

[0287] 1060.如实施例1056-1059中任一项所述的复合物,其中所述ORF1分子是蛋白质外部的一部分。

[0288] 1061.如实施例1056-1060中任一项所述的复合物,其中所述遗传元件正在经历复制。

[0289] 1062.如实施例1056-1061中任一项所述的复合物,其中所述复合物在指环体中。

[0290] 1063.如实施例1056-1062中任一项所述的复合物,其中所述遗传元件进一步包含编码所述多肽的核酸序列。

[0291] 1064.如实施例1056-1063中任一项所述的复合物,其中所述遗传元件不包含编码所述多肽的核酸序列。

[0292] 1065.如实施例1056-1064中任一项所述的复合物,其中所述遗传元件包含GC富集区,例如如本文所述。

[0293] 1066.如实施例1065所述的复合物,其中所述GC富集区包含以下任一者的核酸序列的至少10、15、20、25、30、31、32、33、34、35或36个连续核苷酸:

[0294] (i)CGCGCTGCGCGCGCCGCCCAGTAGGGGGAGCCATGC(SEQ ID NO:160),

[0295] (ii)GCGCTX$_1$CGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:164),其中X$_1$选自T、G或A;

[0296] (iii)GCGCTTCGCGCGCCGCCCACTAGGGGGCGTTGCGCG(SEQ ID NO:165);

[0297] (iv)GCGCTGCGCGCGCCGCCCAGTAGGGGGCGCAATGCG(SEQ ID NO:166);

[0298] (v)GCGCTGCGCGCGCGGCCCCCGGGGGAGGCATTGCCT(SEQ ID NO:167);

[0299] (vi)GCGCTGCGCGCGCGCGCCGGGGGGGCGCCAGCGCCC(SEQ ID NO:168);

[0300] (vii)GCGCTTCGCGCGCGCGCCGGGGGGCTCCGCCCCCCC(SEQ ID NO:169);

[0301] (viii)GCGCTTCGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:170);

[0302] (ix)GCGCTACGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:171);或

[0303] (x)GCGCTACGCGCGCGCGCCGGGGGGCTCTGCCCCCCC(SEQ ID NO:172);

[0304] 或与其具有至少75%、76%、77%、78%、79%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。

[0305] 1067.一种指环体,所述指环体包含:

[0306] (a)蛋白质外部;

[0307] (b)如前述实施例中任一项所述的多肽或复合物;

[0308] (c)遗传元件,其包含可操作地连接至编码效应子(例如内源性效应子或外源性效应子,例如,如本文所述)的核酸序列(例如,DNA序列)的启动子元件;并且

[0309] 其中所述遗传元件被封闭在所述蛋白质外部内。

[0310] 1068.一种指环体,所述指环体包含:

[0311] (a)蛋白质外部;

[0312] (b)遗传元件,其包含:

[0313] (i)可操作地连接至编码效应子(例如内源性效应子或外源性效应子,例如,如本文所述)的核酸序列(例如,DNA序列)的启动子元件,以及

[0314] (ii)编码如前述实施例中任一项所述的多肽的核酸;并且

[0315] 其中所述遗传元件被封闭在所述蛋白质外部内。

[0316] 1069.一种指环体,所述指环体包含:

[0317] (a)蛋白质外部;

[0318] (b)ORF1分子或编码所述ORF1分子的核酸;

[0319] (c)遗传元件,其包含可操作地连接至编码效应子的异源核酸序列(例如,DNA序列)的启动子元件;并且

[0320] 其中所述遗传元件被封闭在所述蛋白质外部内。

[0321] 1070.一种指环体,所述指环体包含:

[0322] (a)蛋白质外部;

[0323] (b)ORF1分子或编码所述ORF1分子的核酸;

[0324] (c)遗传元件,其包含启动子元件,编码效应子(例如外源性效应子或内源性效应子)的核酸序列(例如,DNA序列),和含有以下核酸序列的至少10、15、20、25、30、31、32、33、34、35或36个连续核苷酸的区域:

[0325] (i)CGCGCTGCGCGCGCCGCCCAGTAGGGGGAGCCATGC(SEQ ID NO:160),

[0326] (ii)GCGCTX$_1$CGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:164),其中X$_1$选自T、G或A;

[0327] (iii)GCGCTTCGCGCGCCGCCCACTAGGGGGCGTTGCGCG(SEQ ID NO:165);

[0328] (iv)GCGCTGCGCGCGCCGCCCAGTAGGGGGCGCAATGCG(SEQ ID NO:166);

[0329] (v)GCGCTGCGCGCGCGGCCCCCGGGGGAGGCATTGCCT(SEQ ID NO:167);

[0330] (vi)GCGCTGCGCGCGCGCGCCGGGGGGGCGCCAGCGCCC(SEQ ID NO:168);

[0331] (vii)GCGCTTCGCGCGCGCCGGGGGGCTCCGCCCCCCC(SEQ ID NO:169);

[0332] (viii)GCGCTTCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:170);

[0333] (ix)GCGCTACGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:171);或

[0334] (x)GCGCTACGCGCGCGCCGGGGGGCTCTGCCCCCCC(SEQ ID NO:172);

[0335] 或与其具有至少75%、76%、77%、78%、79%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列;并且

[0336] 其中所述遗传元件包含相对于野生型指环病毒基因组序列(例如,如本文所述)的至少一个差异(例如,突变、化学修饰或表观遗传改变),例如插入、取代、酶促修饰和/或缺失,例如结构域(例如TATA框、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号或GC富集区中的一个或多个)的缺失;

[0337] 其中所述遗传元件被封闭在所述蛋白质外部内;并且

[0338] 其中指环体被配置为将遗传元件递送到真核细胞中;并且任选地,其中所述遗传

元件：

[0339] (i)不包含相对于野生型TTV-tth8基因组序列的核苷酸3436至3607的缺失，例如，如本文所述；

[0340] (ii)不包含相对于野生型TTMV-LY2基因组序列的核苷酸1432至2210的缺失，例如，如本文所述；和/或

[0341] (iii)不包含相对于野生型TTMV-LY2基因组序列的至少101个核苷酸的缺失，例如，如本文所述。

[0342] 1071.一种指环体，所述指环体包含：

[0343] (a)蛋白质外部；

[0344] (b)ORF1分子或编码所述ORF1分子的核酸；

[0345] (c)遗传元件，其包含启动子元件，编码效应子(例如外源性效应子或内源性效应子)的核酸序列(例如，DNA序列)，和含有具有至少70％(例如，至少70％、71％、72％、73％、74％、75％、76％、77％、78％、79％、80％或80.6％)的GC含量的至少20个(例如，至少20、25、30、31、32、33、34、35或36个)连续核苷酸的序列；

[0346] 其中所述遗传元件包含相对于野生型指环病毒基因组序列(例如，如本文所述)的至少一个差异(例如，突变、化学修饰或表观遗传改变)，例如插入、取代、酶促修饰和/或缺失，例如结构域(例如TATA框、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号或GC富集区中的一个或多个)的缺失；

[0347] 其中所述遗传元件被封闭在所述蛋白质外部内；并且

[0348] 其中指环体被配置为将遗传元件递送到真核细胞中；并且

[0349] 任选地，其中所述遗传元件：

[0350] (i)不包含相对于野生型TTV-tth8基因组序列的核苷酸3436至3607的缺失，例如，如本文所述；

[0351] (ii)不包含相对于野生型TTMV-LY2基因组序列的核苷酸1432至2210的缺失，例如，如本文所述；和/或

[0352] (iii)不包含相对于野生型TTMV-LY2基因组序列的至少101个核苷酸的缺失，例如，如本文所述。

[0353] 1072.一种指环体，所述指环体包含：

[0354] (a)蛋白质外部；

[0355] (b)ORF1分子或编码所述ORF1分子的核酸；

[0356] 其中：

[0357] (i)所述ORF1分子的至少30％(例如至少30％、35％、40％、45％、50％、55％、60％、65％、70％、75％、80％、90％或更多)的氨基酸是β-链的一部分；

[0358] (ii)所述ORF1分子的二级结构包含至少三(例如至少3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20)个β-链；

[0359] (iii)所述ORF1分子的二级结构包含β-链与α-螺旋的比例为至少1:1、2:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1或10:1；并且

[0360] (c)遗传元件，其包含启动子元件，编码效应子(例如外源性效应子或内源性效应子)的核酸序列(例如，DNA序列)，和蛋白质结合序列；

[0361] 其中所述遗传元件包含相对于野生型指环病毒基因组序列(例如,如本文所述)的至少一个差异(例如,突变、化学修饰或表观遗传改变),例如插入、取代、酶促修饰和/或缺失,例如结构域(例如TATA框、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号或GC富集区中的一个或多个)的缺失;

[0362] 其中所述遗传元件被封闭在所述蛋白质外部内;并且

[0363] 其中指环体被配置为将遗传元件递送到真核细胞中;并且

[0364] 任选地,其中所述遗传元件:

[0365] (i)不包含相对于野生型TTV-tth8基因组序列的核苷酸3436至3607的缺失,例如,如本文所述;

[0366] (ii)不包含相对于野生型TTMV-LY2基因组序列的核苷酸1432至2210的缺失,例如,如本文所述;和/或

[0367] (iii)不包含相对于野生型TTMV-LY2基因组序列的至少101个核苷酸的缺失,例如,如本文所述。

[0368] 1073.一种指环体,所述指环体包含:

[0369] (a)蛋白质外部;

[0370] (b)ORF1分子或编码所述ORF1分子的核酸;

[0371] (c)遗传元件,其包含启动子元件和编码效应子(例如外源性效应子或内源性效应子)的核酸序列(例如,DNA序列),和蛋白质结合序列;

[0372] 其中所述遗传元件包含相对于野生型指环病毒基因组序列(例如,如本文所述)的至少一个差异(例如,突变、化学修饰或表观遗传改变),例如插入、取代、酶促修饰和/或缺失,例如结构域(例如TATA框、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号或GC富集区中的一个或多个)的缺失;

[0373] 其中所述遗传元件被封闭在所述蛋白质外部内;并且

[0374] 其中指环体被配置为将遗传元件递送到真核细胞中;并且

[0375] 任选地,其中所述遗传元件:

[0376] (i)不包含相对于野生型TTV-tth8基因组序列的核苷酸3436至3607的缺失,例如,如本文所述;

[0377] (ii)不包含相对于野生型TTMV-LY2基因组序列的核苷酸1432至2210的缺失,例如,如本文所述;和/或

[0378] (iii)不包含相对于野生型TTMV-LY2基因组序列的至少101个核苷酸的缺失,例如,如本文所述。

[0379] 1074.一种指环体,所述指环体包含:

[0380] (a)蛋白质外部;

[0381] (b)遗传元件,其包含启动子元件,编码效应子(例如外源性效应子或内源性效应子)的核酸序列(例如,DNA序列),和含有以下核酸序列的至少10、15、20、25、30、31、32、33、34、35或36个连续核苷酸的区域:

[0382] (i)CGCGCTGCGCGCGCCGCCCAGTAGGGGGAGCCATGC(SEQ ID NO:160),

[0383] (ii)GCGCTX$_1$CGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:164),其中X$_1$选自T、G或A;

[0384] (iii)GCGCTTCGCGCGCCGCCCACTAGGGGGCGTTGCGCG(SEQ ID NO:165)；

[0385] (iv)GCGCTGCGCGCGCCGCCCAGTAGGGGGCGCAATGCG(SEQ ID NO:166)；

[0386] (v)GCGCTGCGCGCGCGGCCCCCGGGGGAGGCATTGCCT(SEQ ID NO:167)；

[0387] (vi)GCGCTGCGCGCGCGCGCCGGGGGGGCGCCAGCGCCC(SEQ ID NO:168)；

[0388] (vii)GCGCTTCGCGCGCGCCGGGGGGCTCCGCCCCCCC(SEQ ID NO:169)；

[0389] (viii)GCGCTTCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:170)；

[0390] (ix)GCGCTACGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:171)；或

[0391] (x)GCGCTACGCGCGCGCCGGGGGGCTCTGCCCCCCC(SEQ ID NO:172)；

[0392] 或与其具有至少75％、76％、77％、78％、79％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列；并且

[0393] 其中所述遗传元件包含相对于野生型指环病毒基因组序列(例如，如本文所述)的至少一个差异(例如，突变、化学修饰或表观遗传改变)，例如插入、取代、酶促修饰和/或缺失，例如结构域(例如TATA框、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号或GC富集区中的一个或多个)的缺失；

[0394] 其中所述遗传元件被封闭在所述蛋白质外部内；并且

[0395] 其中指环体被配置为将遗传元件递送到真核细胞中；并且

[0396] 任选地，其中所述遗传元件：

[0397] (i)不包含相对于野生型TTV-tth8基因组序列的核苷酸3436至3607的缺失，例如，如本文所述；

[0398] (ii)不包含相对于野生型TTMV-LY2基因组序列的核苷酸1432至2210的缺失，例如，如本文所述；和/或

[0399] (iii)不包含相对于野生型TTMV-LY2基因组序列的至少101个核苷酸的缺失，例如，如本文所述。

[0400] 1075.一种指环体，所述指环体包含：

[0401] (a)蛋白质外部；

[0402] (b)遗传元件，其包含启动子元件，编码效应子(例如外源性效应子或内源性效应子)的核酸序列(例如，DNA序列)，和含有具有至少70％、71％、72％、73％、74％、75％、76％、77％、78％、79％、80％或80.6％的GC含量的至少20、25、30、31、32、33、34、35或36个连续核苷酸的序列；以及

[0403] 其中所述遗传元件包含相对于野生型指环病毒基因组序列(例如，如本文所述)的至少一个差异(例如，突变、化学修饰或表观遗传改变)，例如插入、取代、酶促修饰和/或缺失，例如结构域(例如TATA框、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号或GC富集区中的一个或多个)的缺失；

[0404] 其中所述遗传元件被封闭在所述蛋白质外部内；并且

[0405] 其中指环体被配置为将遗传元件递送到真核细胞中；并且

[0406] 任选地，其中所述遗传元件：

[0407] (i)不包含相对于野生型TTV-tth8基因组序列的核苷酸3436至3607的缺失，例如，如本文所述；

[0408] (ii)不包含相对于野生型TTMV-LY2基因组序列的核苷酸1432至2210的缺失，例

如,如本文所述;和/或

[0409] (iii)不包含相对于野生型TTMV-LY2基因组序列的至少101个核苷酸的缺失,例如,如本文所述。

[0410] 1076.一种指环体,所述指环体包含:

[0411] (a)蛋白质外部;

[0412] (b)遗传元件,其包含启动子元件和编码效应子(例如外源性效应子或内源性效应子)的核酸序列(例如,DNA序列),

[0413] 其中所述遗传元件包含区域(例如,包装区域,例如,相对于编码所述效应子的核酸序列定位在3'处),所述区域具有:

[0414] 与以下核酸序列的至少95%(例如,至少95%、96%、97%、98%、99%或100%)序列同一性:CGCGCTGCGCGCGCCGCCCAGTAGGGGGAGCCATGC(SEQ ID NO:160);

[0415] 其中所述遗传元件被封闭在所述蛋白质外部内;并且

[0416] 其中所述指环体被配置为将遗传元件递送到真核细胞中。

[0417] 1076A.一种指环体,所述指环体包含:

[0418] (i)遗传元件,其包含启动子元件和编码治疗性外源性效应子的核酸序列,其中所述遗传元件包含与来自本文所述的指环病毒(例如,如表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中的任一个中所列)的5'UTR核苷酸序列具有至少95%序列同一性的序列;和/或

[0419] (ii)蛋白质外部,其包含与本文所述的指环病毒(例如,如表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中的任一个中所列)的ORF1基因编码的多肽具有至少95%序列同一性的多肽;

[0420] 其中所述遗传元件被封闭在所述蛋白质外部内,并且

[0421] 任选地,其中所述指环体能够将所述遗传元件递送到哺乳动物细胞中。

[0422] 1076B.一种指环体,所述指环体包含:

[0423] (I)遗传元件,其包含:(a)启动子元件,和(b)编码外源性效应子(例如,如本文所述的外源性效应子)的核酸序列,其中所述核酸序列与所述启动子元件可操作地连接;和(c)5'UTR结构域,其包含以下之一:

[0424] (c)(i)SEQ ID NO:54的核苷酸323-393的核酸序列,或与其至少85%相同的核酸序列;

[0425] (c)(ii)SEQ ID NO:113、SEQ ID NO:114、SEQ ID NO:115、SEQ ID NO:116、SEQ ID NO:117、SEQ ID NO:118、SEQ ID NO:119的核酸序列或与其至少85%相同的核酸序列;或

[0426] (c)(iii)SEQ ID NO:61的核苷酸117-187的核酸序列,或与其至少85%相同的核酸序列;

[0427] (II)包含ORF1分子的蛋白质外部;

[0428] 其中所述遗传元件被封闭在所述蛋白质外部内;并且

[0429] 其中合成的指环体能够将所述遗传元件递送到哺乳动物例如人的细胞中。

[0430] 1077.如前述实施例中任一项所述的指环体,其中所述蛋白质外部包含ORF1分子。

[0431] 1078.如前述实施例中任一项所述的指环体,其中所述蛋白质外部中至少60%(例如,至少60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98、99%或100%)的蛋

白质包含ORF1分子。

[0432] 1079.如前述实施例中任一项所述的指环体,其中所述蛋白质外部中不超过1%(例如,不超过1%、2%、3%、4%、5%、10%、15%、20%、25%、30%、35%或40%)的蛋白质包含ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3、ORF2t/3和/或ORF3分子。

[0433] 1080.如前述实施例中任一项所述的指环体,其中所述ORF1分子包含与表A1-A12、B1-B5、C1-C5、1-18、20-37、或D1-D10中的任一个中所列的或由其中的序列编码的ORF1蛋白具有至少70%(例如,至少70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)同一性的氨基酸序列。

[0434] 1081.如前述实施例中任一项所述的指环体,其中所述ORF1分子包含前述实施例中任一项所述的多肽。

[0435] 1082.如前述实施例中任一项所述的指环体,其中所述遗传元件进一步包含编码所述ORF1分子的核酸序列。

[0436] 1083.如前述实施例中任一项所述的指环体,其中所述遗传元件不包含编码所述ORF1分子的核酸序列。

[0437] 1084.如前述实施例中任一项所述的指环体,其中所述遗传元件包含具有至少80%的GC含量的至少20、25、30、31、32、33、34、35或36个连续核苷酸。

[0438] 1085.如前述实施例中任一项所述的指环体,其中所述遗传元件包含具有至少70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%或80.6%的GC含量的至少36个连续核苷酸。

[0439] 1086.如前述实施例中任一项所述的指环体,其中所述遗传元件包含具有至少80%的GC含量的至少36个连续核苷酸。

[0440] 1087.一种分离的核酸组合物(例如,包含一种、两种或更多种核酸分子),其包含编码如前述实施例中任一项所述的多肽的核酸;

[0441] 任选地,其中所述分离的核酸组合物进一步包含相对于野生型指环病毒基因组序列(例如,如本文所述)的至少一个差异(例如,突变、化学修饰或表观遗传改变),例如插入、取代、酶促修饰和/或缺失,例如结构域(例如TATA框、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号或GC富集区中的一个或多个)的缺失;并且

[0442] 任选地,其中所述核酸分子不包含:

[0443] (i)相对于野生型TTV-tth8基因组序列的核苷酸3436至3607的缺失,例如,如本文所述;

[0444] (ii)相对于野生型TTMV-LY2基因组序列的核苷酸1432至2210的缺失,例如,如本文所述;和/或

[0445] (iii)相对于野生型TTMV-LY2基因组序列的至少101个核苷酸的缺失,例如,如本文所述。

[0446] 1088.一种分离的核酸组合物(例如,包含一种、两种或更多种核酸分子),其中所述分离的核酸组合物包含编码ORF1分子的遗传元件;

[0447] 其中:

[0448] (i)所述ORF1分子的至少30%(例如至少30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、90%或更多)的氨基酸是β-折叠的一部分;

[0449] (ii)所述ORF1分子的二级结构包含至少三(例如至少3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20)个β-折叠；

[0450] (iii)所述ORF1分子的二级结构包含β-折叠与α-螺旋的比例为至少1:1、2:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1或10:1；并且

[0451] 其中所述遗传元件包含启动子元件，编码效应子(例如外源性效应子或内源性效应子)的核酸序列，和蛋白质结合序列；

[0452] 其中所述遗传元件包含相对于野生型指环病毒基因组序列(例如，如本文所述)的至少一个差异(例如，突变、化学修饰或表观遗传改变)，例如插入、取代、酶促修饰和/或缺失，例如结构域(例如TATA框、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号或GC富集区中的一个或多个)的缺失；并且

[0453] 任选地，其中所述核酸分子不包含：

[0454] (i)相对于野生型TTV-tth8基因组序列的核苷酸3436至3607的缺失，例如，如本文所述；

[0455] (ii)相对于野生型TTMV-LY2基因组序列的核苷酸1432至2210的缺失，例如，如本文所述；和/或

[0456] (iii)相对于野生型TTMV-LY2基因组序列的至少101个核苷酸的缺失，例如，如本文所述。

[0457] 1089.一种分离的核酸组合物(例如，包含一种、两种或更多种核酸分子)，其包含：

[0458] (a)编码ORF1分子的遗传元件；

[0459] (b)以下核酸序列的至少10、15、20、25、30、31、32、33、34、35或36个连续核苷酸：

[0460] (i)CGCGCTGCGCGCGCCGCCCAGTAGGGGGAGCCATGC(SEQ ID NO:160)，

[0461] (ii)GCGCTX$_1$CGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:164)，其中X$_1$选自T、G或A；

[0462] (iii)GCGCTTCGCGCGCCGCCCACTAGGGGGCGTTGCGCG(SEQ ID NO:165)；

[0463] (iv)GCGCTGCGCGCGCCGCCCAGTAGGGGGCGCAATGCG(SEQ ID NO:166)；

[0464] (v)GCGCTGCGCGCGCGGCCCCCGGGGGAGGCATTGCCT(SEQ ID NO:167)；

[0465] (vi)GCGCTGCGCGCGCGCGCCGGGGGGGCGCCAGCGCCC(SEQ ID NO:168)；

[0466] (vii)GCGCTTCGCGCGCGCCGGGGGGCTCCGCCCCCCC(SEQ ID NO:169)；

[0467] (viii)GCGCTTCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:170)；

[0468] (ix)GCGCTACGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:171)；或

[0469] (x)GCGCTACGCGCGCGCCGGGGGGCTCTGCCCCCCC(SEQ ID NO:172)；

[0470] 或与其具有至少75％、76％、77％、78％、79％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列；并且

[0471] (c)相对于野生型指环病毒基因组序列(例如，如本文所述)的至少一个差异(例如，突变、化学修饰或表观遗传改变)，例如插入、取代、酶促修饰和/或缺失，例如结构域(例如TATA框、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号或GC富集区中的一个或多个)的缺失；

[0472] 任选地，其中所述核酸分子不包含：

[0473] (i)相对于野生型TTV-tth8基因组序列的核苷酸3436至3607的缺失，例如，如本文

所述；

[0474] (ii)相对于野生型TTMV-LY2基因组序列的核苷酸1432至2210的缺失，例如，如本文所述；和/或

[0475] (iii)相对于野生型TTMV-LY2基因组序列的至少101个核苷酸的缺失，例如，如本文所述。

[0476] 1090.一种分离的核酸组合物(例如，包含一种、两种或更多种核酸分子)，其中所述分离的核酸组合物包含：

[0477] (a)编码ORF1分子的遗传元件；

[0478] (b)具有至少70％、71％、72％、73％、74％、75％、76％、77％、78％、79％、80％或80.6％的GC含量的至少20、25、30、31、32、33、34、35或36个连续核苷酸；并且

[0479] 其中所述分离的核酸组合物包含相对于野生型指环病毒基因组序列(例如，如本文所述)的至少一个差异(例如，突变、化学修饰或表观遗传改变)，例如插入、取代、酶促修饰和/或缺失，例如结构域(例如TATA框、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号或GC富集区中的一个或多个)的缺失；并且

[0480] 任选地，其中所述核酸分子不包含：

[0481] (i)相对于野生型TTV-tth8基因组序列的核苷酸3436至3607的缺失，例如，如本文所述；

[0482] (ii)相对于野生型TTMV-LY2基因组序列的核苷酸1432至2210的缺失，例如，如本文所述；和/或

[0483] (iii)相对于野生型TTMV-LY2基因组序列的至少101个核苷酸的缺失，例如，如本文所述。

[0484] 1090A.一种分离的核酸组合物(例如，包含一种、两种或更多种核酸分子)，其中所述分离的核酸组合物包含遗传元件，所述遗传元件包含来自本文所述的指环病毒(例如，如表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中的任一个中所列)的5'UTR核苷酸序列。

[0485] 1091.如实施例1089-1090中任一项所述的分离的核酸组合物，其中(a)和(b)是相同核酸的部分。

[0486] 1092.如实施例1089-1091中任一项所述的分离的核酸组合物，其中(a)和(b)是不同核酸的部分。

[0487] 1093.如前述实施例中任一项所述的分离的核酸组合物，其中所述遗传元件进一步包含以下中的一种或多种：TATA盒、起始元件、帽位点、转录起始位点、5'UTR保守结构域、ORF1编码序列、ORF1/1编码序列、ORF1/2编码序列、ORF2编码序列、ORF2/2编码序列、ORF2/3编码序列、ORF2/3t编码序列、三开放阅读框区域、聚(A)信号和/或来自本文所述的指环病毒的GC富集区(例如，如表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中的任一个中所列)，或与其具有至少70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的序列。

[0488] 1094.如前述实施例中任一项所述的分离的核酸组合物，其中所述遗传元件进一步包含指环病毒基因组序列(例如，如本文所述，例如，如表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中的任一个中所列)，或与其具有至少70％、75％、80％、85％、90％、

95％、96％、97％、98％、99％或100％序列同一性的序列。

[0489] 1095.如实施例1094所述的分离的核酸组合物，所述分离的核酸组合物进一步包含指环病毒基因组序列的至少一个另外拷贝或与其具有至少70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的序列(例如，总共1、2、3、4、5或6个拷贝)。

[0490] 1096.如前述实施例中任一项所述的分离的核酸组合物，所述分离的核酸组合物进一步包含所述遗传元件的至少一个另外拷贝(例如，总共1、2、3、4、5或6个拷贝)。

[0491] 1097.一种分离的核酸组合物(例如，包含一种、两种或更多种核酸分子)，其包含以下核酸序列的至少10、15、20、25、30、31、32、33、34、35或36个连续核苷酸：

[0492] (i)CGCGCTGCGCGCGCCGCCCAGTAGGGGGAGCCATGC(SEQ ID NO:160)，

[0493] (ii)GCGCTX$_1$CGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:164)，其中X$_1$选自T、G或A；

[0494] (iii)GCGCTTCGCGCGCCGCCCACTAGGGGGCGTTGCGCG(SEQ ID NO:165)；

[0495] (iv)GCGCTGCGCGCGCCGCCCAGTAGGGGGCGCAATGCG(SEQ ID NO:166)；

[0496] (v)GCGCTGCGCGCGCGGCCCCCGGGGGAGGCATTGCCT(SEQ ID NO:167)；

[0497] (vi)GCGCTGCGCGCGCGCGCCGGGGGGGCGCCAGCGCCC(SEQ ID NO:168)；

[0498] (vii)GCGCTTCGCGCGCGCGCCGGGGGGCTCCGCCCCCCC(SEQ ID NO:169)；

[0499] (viii)GCGCTTCGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:170)；

[0500] (ix)GCGCTACGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:171)；或

[0501] (x)GCGCTACGCGCGCGCGCCGGGGGGCTCTGCCCCCCC(SEQ ID NO:172)；

[0502] 或与其具有至少75％、76％、77％、78％、79％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列；并且

[0503] 相对于野生型指环病毒基因组序列(例如，如本文所述)的至少一个差异(例如，突变、化学修饰或表观遗传改变)，例如插入、取代、酶促修饰和/或缺失，例如结构域(例如TATA框、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号或GC富集区中的一个或多个)的缺失；

[0504] 任选地，其中所述核酸分子不包含：

[0505] (i)相对于野生型TTV-tth8基因组序列的核苷酸3436至3607的缺失，例如，如本文所述；

[0506] (ii)相对于野生型TTMV-LY2基因组序列的核苷酸1432至2210的缺失，例如，如本文所述；和/或

[0507] (iii)相对于野生型TTMV-LY2基因组序列的至少101个核苷酸的缺失，例如，如本文所述。

[0508] 1098.一种分离的核酸组合物(例如，包含一种、两种或更多种核酸分子)，其中所述分离的核酸组合物包含具有至少70％、71％、72％、73％、74％、75％、76％、77％、78％、79％、80％或80.6％的GC含量的至少20、25、30、31、32、33、34、35或36个连续核苷酸；并且

[0509] 其中所述分离的核酸组合物包含相对于野生型指环病毒基因组序列(例如，如本文所述)的至少一个差异(例如，突变、化学修饰或表观遗传改变)，例如插入、取代、酶促修饰和/或缺失，例如结构域(例如TATA框、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号或GC富集区中的一个或多个)的缺失；并且

[0510] 任选地，其中所述核酸分子不包含：

[0511] (i)相对于野生型TTV-tth8基因组序列的核苷酸3436至3607的缺失，例如，如本文所述；

[0512] (ii)相对于野生型TTMV-LY2基因组序列的核苷酸1432至2210的缺失，例如，如本文所述；和/或

[0513] (iii)相对于野生型TTMV-LY2基因组序列的至少101个核苷酸的缺失，例如，如本文所述。

[0514] 1099.如前述实施例中任一项所述的分离的核酸组合物，其中所述ORF1分子包含前述实施例中任一项所述的多肽。

[0515] 1100.如前述实施例中任一项所述的分离的核酸组合物，其包含具有至少80％的GC含量的至少20、25、30、31、32、33、34、35或36个连续核苷酸。

[0516] 1101.如前述实施例中任一项所述的分离的核酸组合物，其包含具有至少70％、71％、72％、73％、74％、75％、76％、77％、78％、79％、80％或80.6％的GC含量的至少36个连续核苷酸。

[0517] 1102.如前述实施例中任一项所述的分离的核酸组合物，其包含具有至少80％的GC含量的至少36个连续核苷酸。

[0518] 1103.如前述实施例中任一项所述的分离的核酸组合物，其进一步包含以下中的一种或多种：启动子元件、编码效应子(例如外源性效应子或内源性效应子)的核酸序列和/或蛋白质结合序列(例如，外部蛋白质结合序列)。

[0519] 1104.如前述实施例中任一项所述的分离的核酸组合物，其包含野生型指环病毒基因组序列的至少约100、150、200、250、300、350、400、450或500个连续核苷酸或与其具有至少75％、76％、77％、78％、79％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。

[0520] 1105.一种分离的核酸分子(例如，表达载体)，其包含与以下核酸序列具有至少95％(例如，至少95％、96％、97％、98％、99％或100％)序列同一性的核酸序列：

[0521] (i)CGCGCTGCGCGCGCCGCCCAGTAGGGGGAGCCATGC(SEQ ID NO:160)，

[0522] (ii)GCGCTX$_1$CGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:164)，其中X$_1$选自T、G或A；

[0523] (iii)GCGCTTCGCGCGCCGCCCACTAGGGGGCGTTGCGCG(SEQ ID NO:165)；

[0524] (iv)GCGCTGCGCGCGCCGCCCAGTAGGGGGCGCAATGCG(SEQ ID NO:166)；

[0525] (v)GCGCTGCGCGCGCGGCCCCCGGGGGAGGCATTGCCT(SEQ ID NO:167)；

[0526] (vi)GCGCTGCGCGCGCGCGCCGGGGGGGCGCCAGCGCCC(SEQ ID NO:168)；

[0527] (vii)GCGCTTCGCGCGCGCGCCGGGGGGCTCCGCCCCCCC(SEQ ID NO:169)；

[0528] (viii)GCGCTTCGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:170)；

[0529] (ix)GCGCTACGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:171)；或

[0530] (x)GCGCTACGCGCGCGCGCCGGGGGGCTCTGCCCCCCC(SEQ ID NO:172)。

[0531] 1106.如前述实施例中任一项所述的分离的核酸组合物，其中所述分离的核酸分子是环状的。

[0532] 1107.一种分离的细胞，例如宿主细胞，所述分离的细胞包含：

[0533] (a)编码如前述实施例中任一项所述的多肽的核酸,其中所述核酸是质粒、病毒核酸或整合到细胞染色体中,和

[0534] (b)遗传元件,其中所述遗传元件包含启动子元件和编码效应子(例如外源性效应子或内源性效应子)的核酸序列(例如,DNA序列),以及蛋白质结合序列,其中任选地所述遗传元件不编码ORF1多肽(例如,ORF1蛋白)。

[0535] 1108.一种分离的细胞,例如宿主细胞,所述分离的细胞包含:

[0536] (a)编码ORF1分子的核酸,其中所述核酸是质粒、病毒核酸或整合到细胞染色体中,和

[0537] (b)遗传元件,其中遗传元件包含启动子元件和编码效应子(例如外源性效应子或内源性效应子)的核酸序列(例如,DNA序列),和蛋白质结合序列。

[0538] 1109.一种分离的细胞,例如宿主细胞,所述分离的细胞包含:

[0539] (a)编码ORF1分子的核酸(例如,其中所述核酸是质粒、病毒核酸或整合到细胞染色体中),和

[0540] (b)不编码ORF1分子的遗传元件,其中所述遗传元件包含启动子元件和编码效应子(例如外源性效应子或内源性效应子)的核酸序列(例如,DNA序列),和蛋白质结合序列。

[0541] 1109A.一种分离的细胞,例如宿主细胞,所述分离的细胞包含:

[0542] (i)一种核酸分子(例如,第一核酸分子),其包含如本文所述的指环体的遗传元件(例如,不编码ORF1分子的遗传元件)的核酸序列,和

[0543] (ii)任选地,核酸分子,例如第二核酸分子,所述第二核酸分子编码以下中的一种或多种:选自ORF1、ORF2、ORF2/2、ORF2/3、ORF1/1或ORF1/2的氨基酸序列(例如,如表16中任一个中所列),或与其具有至少70%(例如,至少70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)序列同一性的氨基酸序列。

[0544] 1110.如前述实施例中任一项所述的分离的细胞,其中不编码ORF1分子的遗传元件编码ORF1分子的片段,例如不形成衣壳的片段,例如少于1000、900、800、700、600、500、400、300、200、100、50、20或10个核苷酸的片段。

[0545] 1111.一种分离的细胞,例如宿主细胞,所述分离的细胞包含编码ORF1分子的核酸(例如,其中所述核酸是质粒,是病毒核酸,或者整合到细胞染色体中),其中所述分离的细胞不包含ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3、ORF2t/3和/或ORF3分子中的一种或多种。

[0546] 1112.一种分离的细胞,例如宿主细胞,所述分离的细胞包含如前述实施例中任一项所述的核酸组合物。

[0547] 1113.一种编码ORF1分子的辅助核酸(例如,质粒或病毒核酸),其中所述分离的细胞不包含ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3、ORF2t/3和/或ORF3分子中的一种或多种。

[0548] 1114.一种组合物,所述组合物包含:

[0549] (a)本文所述的分离的细胞,和

[0550] (b)本文所述的指环体。

[0551] 1115.一种组合物,所述组合物包含:

[0552] (a)细胞,其编码ORF1分子的核酸(例如,其中所述核酸是质粒、病毒核酸或整合到细胞染色体中),和

[0553] (b)不编码ORF1分子的遗传元件(例如,在细胞内或在细胞外,例如细胞培养基

中),其中所述遗传元件包含启动子元件和编码效应子(例如外源性效应子或内源性效应子)的核酸序列(例如,DNA序列),和蛋白质结合序列。

[0554] 1116.一种药物组合物,其包含如前述实施例中任一项所述的多肽、复合物、指环体或分离的核酸以及药学上可接受的载剂和/或赋形剂。

[0555] 1117.一种制造ORF1分子的方法,所述方法包括:

[0556] (a)提供包含编码如前述实施例中任一项所述的多肽的核酸的宿主细胞(例如,本文所述的宿主细胞),并且

[0557] (b)将所述宿主细胞维持在允许细胞生产所述多肽的条件下;

[0558] 从而制造所述ORF1分子。

[0559] 1118.一种制造ORF1分子的方法,所述方法包括:

[0560] (a)提供包含如前述实施例中任一项所述的核酸组合物的宿主细胞(例如,本文所述的宿主细胞),并且

[0561] (b)将所述宿主细胞维持在允许细胞生产所述多肽的条件下;

[0562] 从而制造所述ORF1分子。

[0563] 1119.如实施例1117或1118所述的方法,其中所述宿主细胞是辅助细胞。

[0564] 1120.如实施例1119所述的方法,其中所述辅助细胞包含一种或多种另外的编码一种或多种另外的ORF(例如,ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3、ORF2t/3和/或ORF3)的野生型指环病毒,例如,如本文所述。

[0565] 1121.如实施例1117-1120中任一项所述的方法,其中所述核酸整合到所述宿主细胞的基因组中。

[0566] 1122.如实施例1117-1121中任一项所述的方法,其中所述宿主细胞产生至少约10、20、30、40、50、60、70、80、90、100、200、300、400、500、1000、10,000、50,000、100,000、500,000或1,000,000个拷贝(例如至少约60个拷贝)的所述多肽/宿主细胞。

[0567] 1123.如实施例1117-1122中任一项所述的方法,其中所述宿主细胞产生至少约1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、1000、10,000或100,000个拷贝(例如至少约60个拷贝)的所述多肽/由所述宿主细胞产生的指环体。

[0568] 1124.如实施例1117-1123中任一项所述的方法,其中所述方法包括提供多个宿主细胞,并将所述宿主细胞维持在允许产生至少1000个拷贝的所述多肽/细胞的条件下。

[0569] 1125.如实施例1124所述的方法,其中所述多个宿主细胞产生至少约1×10^{5}、$1x10^{6}$、$1x10^{7}$、$1x10^{8}$、$9x10^{8}$、$1x10^{9}$、$1x10^{10}$、$1x10^{11}$或$1x10^{12}$个拷贝的所述多肽。

[0570] 1126.一种制备指环体组合物的方法,所述方法包括:

[0571] (a)提供辅助细胞,例如本文所述的辅助细胞;

[0572] (b)在允许细胞生产指环体的条件下将遗传元件引入所述辅助细胞,并且

[0573] (c)配制所述指环体,例如,作为适合施用给受试者的药物组合物,从而制备所述指环体组合物。

[0574] 1127.一种制备指环体组合物的方法,所述方法包括:

[0575] (a)提供宿主细胞;

[0576] (b)将辅助核酸引入所述宿主细胞;

[0577] (c)在允许细胞生产指环体的条件下将遗传元件引入所述宿主细胞(例如,在(b)

之前、之后或与其同时)；并且

[0578] (d)配制所述指环体，例如，作为适合施用给受试者的药物组合物；

[0579] 从而制备所述指环体组合物。

[0580] 1128.一种制备指环体组合物的方法，所述方法包括：

[0581] (a)提供辅助细胞，其包含编码ORF1分子的核酸(例如，其中所述核酸是质粒、病毒核酸或整合到辅助细胞染色体中)；

[0582] (b)在允许细胞生产指环体的条件下将遗传元件引入所述辅助细胞，其中所述遗传元件不编码ORF1分子，其中所述遗传元件包含启动子元件和编码效应子(例如外源性效应子或内源性效应子)的核酸序列(例如，DNA序列)，和蛋白质结合序列；并且

[0583] (c)配制所述指环体，例如，作为适合施用给受试者的药物组合物；

[0584] 从而制备所述指环体组合物。

[0585] 1129.一种制备指环体组合物的方法，所述方法包括：

[0586] (a)提供宿主细胞；

[0587] (b)将编码ORF1分子的辅助核酸(例如，其中所述核酸是质粒或病毒核酸)引入所述宿主细胞中；并且

[0588] (c)在允许细胞生产指环体的条件下将遗传元件引入所述宿主细胞(例如，在(b)之前、之后或与其同时)，其中所述遗传元件包不编码ORF1分子，其中所述遗传元件包含启动子元件和编码效应子(例如外源性效应子或内源性效应子)的核酸序列(例如，DNA序列)，和蛋白质结合序列，

[0589] 从而制备所述指环体。

[0590] 1130.如前述实施例中任一项所述的方法，所述方法进一步包括将所述指环体与所述辅助细胞或宿主细胞分离。

[0591] 1131.如前述实施例中任一项所述的方法，其中提供辅助细胞包括将辅助核酸引入所述宿主细胞，例如，其中所述辅助核酸编码ORF1分子(例如，其中所述核酸是质粒或病毒核酸)。

[0592] 1132.如前述实施例中任一项所述的方法，其中所述辅助细胞包含所述ORF1分子。

[0593] 1133.如前述实施例中任一项所述的方法，其中所述核酸包含以下中的一种或多种：TATA盒、起始元件、帽位点、转录起始位点、5'UTR保守结构域、ORF1编码序列、ORF1/1编码序列、ORF1/2编码序列、ORF2编码序列、ORF2/2编码序列、ORF2/3编码序列、ORF2/3t编码序列、三开放阅读框区域、聚(A)信号和/或来自本文所述的指环病毒的GC富集区(例如，如表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中的任一个中所列)，或与其具有至少70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的序列。

[0594] 1134.如前述实施例中任一项所述的方法，其中所述核酸包含指环病毒基因组序列(例如，如本文所述，例如，如表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中的任一个中所列)，或与其具有至少70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的序列。

[0595] 1135.如前述实施例中任一项所述的方法，其中所述核酸包含指环病毒基因组序列的至少一个另外拷贝或与其具有至少70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的序列(例如，总共1、2、3、4、5或6个拷贝)。

[0596] 1136.如前述实施例中任一项所述的方法,其中所述宿主细胞或辅助细胞包含所述核酸的至少一个另外拷贝(例如,总共1、2、3、4、5或6个拷贝)。

[0597] 1137.如前述实施例中任一项所述的方法,其中所述核酸是环状的。

[0598] 1137A.一种制备指环体,例如合成指环体的方法,所述方法包括:

[0599] a)提供宿主细胞,其包含:

[0600] (i)核酸分子,例如第一核酸分子,其包含如本文所述的指环体,例如合成的指环体的遗传元件的核酸序列,和

[0601] (ii)核酸分子,例如第二核酸分子,所述第二核酸分子编码以下中的一种或多种:选自ORF1、ORF2、ORF2/2、ORF2/3、ORF1/1或ORF1/2的氨基酸序列(例如,如表16中任一个中所列),或与其具有至少70%(例如,至少70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)序列同一性的氨基酸序列;并且

[0602] b)在适合于制备指环体的条件下培养所述宿主细胞。

[0603] 1137B.如实施例1137A所述的方法,所述方法进一步包括,在步骤(a)之前,将所述第一核酸分子和/或所述第二核酸分子引入所述宿主细胞。

[0604] 1137C.如实施例1137A或1137B所述的方法,其中所述第二核酸分子在所述第一核酸分子之前、与其同时或在其之后被引入所述宿主细胞。

[0605] 1137D.如实施例1137C所述的方法,其中所述第二核酸分子被整合到所述宿主细胞的基因组中。

[0606] 1137E.如实施例1137C所述的方法,其中所述第二核酸分子是辅助(例如辅助质粒或辅助病毒的基因组)。

[0607] 1137F.如实施例1137A-1137E中任一项所述的方法,其中所述第一核酸包含以下中的一种或多种:TATA盒、起始元件、帽位点、转录起始位点、5'UTR保守结构域和/或来自本文所述的指环病毒的GC富集区(例如,如表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中的任一个中所列),或与其具有至少70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的序列。

[0608] 1138.一种向受试者递送效应子的方法,所述方法包括向所述受试者施用包含以下的指环体:

[0609] (a)包含ORF1分子的蛋白质外部;

[0610] (b)遗传元件,其包含启动子元件和编码所述效应子(例如外源性效应子或内源性效应子)的核酸序列(例如,DNA序列),和含有以下核酸序列的至少10、15、20、25、30、31、32、33、34、35或36个连续核苷酸的区域:

[0611] (i)CGCGCTGCGCGCGCCGCCCAGTAGGGGGAGCCATGC(SEQ ID NO:160),

[0612] (ii)GCGCTX$_1$CGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:164),其中X$_1$选自T、G或A;

[0613] (iii)GCGCTTCGCGCGCCGCCCACTAGGGGGCGTTGCGCG(SEQ ID NO:165);

[0614] (iv)GCGCTGCGCGCGCCGCCCAGTAGGGGGCGCAATGCG(SEQ ID NO:166);

[0615] (v)GCGCTGCGCGCGCGGCCCCCGGGGGAGGCATTGCCT(SEQ ID NO:167);

[0616] (vi)GCGCTGCGCGCGCGCCGGGGGGGCGCCAGCGCCC(SEQ ID NO:168);

[0617] (vii)GCGCTTCGCGCGCGCCGGGGGGCTCCGCCCCCCC(SEQ ID NO:169);

[0618] (viii)GCGCTTCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:170)；

[0619] (ix)GCGCTACGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:171)；或

[0620] (x)GCGCTACGCGCGCGCCGGGGGGCTCTGCCCCCCC(SEQ ID NO:172)；

[0621] 或与其具有至少75％、76％、77％、78％、79％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列；并且

[0622] 其中所述遗传元件被封闭在所述蛋白质外部内；并且

[0623] 任选地，其中所述遗传元件：

[0624] (i)不包含相对于野生型TTV-tth8基因组序列的核苷酸3436至3607的缺失，例如，如本文所述；

[0625] (ii)不包含相对于野生型TTMV-LY2基因组序列的核苷酸1432至2210的缺失，例如，如本文所述；和/或

[0626] (iii)不包含相对于野生型TTMV-LY2基因组序列的至少101个核苷酸的缺失，例如，如本文所述，

[0627] 从而将所述效应子递送给受试者。

[0628] 1139.一种向受试者递送效应子的方法，所述方法包括向所述受试者施用包含以下的指环体：

[0629] (a)包含ORF1分子的蛋白质外部；

[0630] (b)遗传元件，其包含启动子元件，编码所述效应子(例如外源性效应子或内源性效应子)的核酸序列(例如，DNA序列)，和含有具有至少70％、71％、72％、73％、74％、75％、76％、77％、78％、79％、80％或80.6％的GC含量的序列的至少20、25、30、31、32、33、34、35或36个连续核苷酸；

[0631] 其中所述遗传元件被封闭在所述蛋白质外部内；并且

[0632] 任选地，其中所述遗传元件：

[0633] (i)不包含相对于野生型TTV-tth8基因组序列的核苷酸3436至3607的缺失，例如，如本文所述；

[0634] (ii)不包含相对于野生型TTMV-LY2基因组序列的核苷酸1432至2210的缺失，例如，如本文所述；和/或

[0635] (iii)不包含相对于野生型TTMV-LY2基因组序列的至少101个核苷酸的缺失，例如，如本文所述，

[0636] 从而将所述效应子递送给受试者。

[0637] 1140.一种向受试者递送效应子的方法，所述方法包括向所述受试者施用包含以下的指环体：

[0638] (a)包含ORF1分子的蛋白质外部；

[0639] (b)遗传元件，其包含启动子元件和编码所述效应子(例如外源性效应子或内源性效应子)的核酸序列(例如，DNA序列)，和蛋白质结合序列；

[0640] 其中所述遗传元件被封闭在所述蛋白质外部内；并且

[0641] 任选地，其中所述遗传元件：

[0642] (i)不包含相对于野生型TTV-tth8基因组序列的核苷酸3436至3607的缺失，例如，如本文所述；

[0643] (ii)不包含相对于野生型TTMV-LY2基因组序列的核苷酸1432至2210的缺失,例如,如本文所述;和/或

[0644] (iii)不包含相对于野生型TTMV-LY2基因组序列的至少101个核苷酸的缺失,例如,如本文所述,

[0645] 从而将所述效应子递送给受试者。

[0646] 1141.一种将效应子递送至靶细胞的方法,所述非包括使所述靶细胞与指环体接触,所述指环体包含:

[0647] (a)包含ORF1分子的蛋白质外部;

[0648] (b)遗传元件,其包含启动子元件和编码所述效应子(例如外源性效应子或内源性效应子)的核酸序列(例如,DNA序列),和含有以下核酸序列的至少10、15、20、25、30、31、32、33、34、35或36个连续核苷酸的区域:

[0649] (i)CGCGCTGCGCGCGCCGCCCAGTAGGGGGAGCCATGC(SEQ ID NO:160),

[0650] (ii)GCGCTX$_1$CGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:164),其中X$_1$选自T、G或A;

[0651] (iii)GCGCTTCGCGCGCCGCCCACTAGGGGGCGTTGCGCG(SEQ ID NO:165);

[0652] (iv)GCGCTGCGCGCGCCGCCCAGTAGGGGGCGCAATGCG(SEQ ID NO:166);

[0653] (v)GCGCTGCGCGCGCGGCCCCCGGGGGAGGCATTGCCT(SEQ ID NO:167);

[0654] (vi)GCGCTGCGCGCGCGCGCCGGGGGGGCGCCAGCGCCC(SEQ ID NO:168);

[0655] (vii)GCGCTTCGCGCGCGCGCCGGGGGGCTCCGCCCCCCC(SEQ ID NO:169);

[0656] (viii)GCGCTTCGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:170);

[0657] (ix)GCGCTACGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:171);或

[0658] (x)GCGCTACGCGCGCGCGCCGGGGGGCTCTGCCCCCCC(SEQ ID NO:172);

[0659] 或与其具有至少75%、76%、77%、78%、79%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列;并且

[0660] 其中所述遗传元件被封闭在所述蛋白质外部内;并且

[0661] 任选地,其中所述遗传元件:

[0662] (i)不包含相对于野生型TTV-tth8基因组序列的核苷酸3436至3607的缺失,例如,如本文所述;

[0663] (ii)不包含相对于野生型TTMV-LY2基因组序列的核苷酸1432至2210的缺失,例如,如本文所述;和/或

[0664] (iii)不包含相对于野生型TTMV-LY2基因组序列的至少101个核苷酸的缺失,例如,如本文所述,

[0665] 从而将所述效应子递送至所述靶细胞。

[0666] 1142.一种将效应子递送至靶细胞的方法,所述非包括使所述靶细胞与指环体接触,所述指环体包含:

[0667] (a)包含ORF1分子的蛋白质外部;

[0668] (b)遗传元件,其包含启动子元件,编码所述效应子(例如外源性效应子或内源性效应子)的核酸序列(例如,DNA序列),和含有具有至少70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%或80.6%的GC含量的序列的至少20、25、30、31、32、33、34、35或

36个连续核苷酸;

[0669] 其中所述遗传元件被封闭在所述蛋白质外部内;并且

[0670] 任选地,其中所述遗传元件:

[0671] (i)不包含相对于野生型TTV-tth8基因组序列的核苷酸3436至3607的缺失,例如,如本文所述;

[0672] (ii)不包含相对于野生型TTMV-LY2基因组序列的核苷酸1432至2210的缺失,例如,如本文所述;和/或

[0673] (iii)不包含相对于野生型TTMV-LY2基因组序列的至少101个核苷酸的缺失,例如,如本文所述,

[0674] 从而将所述效应子递送至所述靶细胞。

[0675] 1143.一种将效应子递送至靶细胞的方法,所述非包括使所述靶细胞与指环体接触,所述指环体包含:

[0676] (a)包含ORF1分子的蛋白质外部;

[0677] (b)遗传元件,其包含启动子元件和编码所述效应子(例如外源性效应子或内源性效应子)的核酸序列(例如,DNA序列),和蛋白质结合序列;

[0678] 其中所述遗传元件被封闭在所述蛋白质外部内;并且

[0679] 任选地,其中所述遗传元件:

[0680] (i)不包含相对于野生型TTV-tth8基因组序列的核苷酸3436至3607的缺失,例如,如本文所述;

[0681] (ii)不包含相对于野生型TTMV-LY2基因组序列的核苷酸1432至2210的缺失,例如,如本文所述;和/或

[0682] (iii)不包含相对于野生型TTMV-LY2基因组序列的至少101个核苷酸的缺失,例如,如本文所述,

[0683] 从而将所述效应子递送至所述靶细胞。

[0684] 1143A.一种将效应子递送至靶细胞的方法,所述非包括使所述靶细胞与指环体接触,所述指环体包含:

[0685] (i)遗传元件,其包含启动子元件和编码治疗性外源性效应子的核酸序列,其中所述遗传元件包含与来自本文所述的指环病毒(例如,如表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中的任一个中所列)的5'UTR核苷酸序列具有至少95%序列同一性的序列;和/或

[0686] (ii)蛋白质外部,其包含与本文所述的指环病毒(例如,如表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中的任一个中所列)的ORF1基因编码的多肽具有至少95%序列同一性的多肽;

[0687] 其中所述遗传元件被封闭在所述蛋白质外部内;并且

[0688] 任选地,其中所述遗传元件:

[0689] (i)不包含相对于野生型TTV-tth8基因组序列的核苷酸3436至3607的缺失,例如,如本文所述;

[0690] (ii)不包含相对于野生型TTMV-LY2基因组序列的核苷酸1432至2210的缺失,例如,如本文所述;和/或

[0691] (iii)不包含相对于野生型TTMV-LY2基因组序列的至少101个核苷酸的缺失，例如，如本文所述，

[0692] 从而将所述效应子递送至所述靶细胞。

[0693] 1144.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件不编码NCBI登录号A7XCE8.1的氨基酸序列。

[0694] 1145.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述ORF1分子包含与表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所列的ORF1序列具有至少70％(例如至少70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％)序列同一性的氨基酸序列。

[0695] 1146.(i)如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述ORF1分子的至少30％(例如至少30％、35％、40％、45％、50％、55％、60％、65％、70％、75％、80％、90％或更多)的氨基酸是β-折叠的一部分。

[0696] 1147.(ii)如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述ORF1分子的二级结构包含至少三(例如至少3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20)个β-折叠。

[0697] 1148.(iii)如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述ORF1分子的二级结构包含β-折叠与α-螺旋的比例为至少1:1、2:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1或10:1。

[0698] 1149.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述ORF1分子包含精氨酸富集区(例如，与表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所列的精氨酸富集区序列具有至少70％、80％、85％、90％、95％、96％、97％、98％、99％或100％同一性)。

[0699] 1150.如实施例1149所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述精氨酸富集区包含至少15、20、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、45或50个连续核苷酸，所述连续核苷酸包含至少40％(例如至少40％、41％、42％、43％、44％、45％、46％、47％、48％、49％、50％、55％、60％、65％、66％、67％、68％、69％、70％、75％、80％、85％、90％或95％)精氨酸残基。

[0700] 1151.如实施例1149或1150所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述精氨酸富集区位于所述ORF1分子的N末端或C末端。

[0701] 1152.如1149-1151中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述精氨酸富集区与氨基酸序列TVVRRRGRSPRRRTPSPRRRRSQSPRRRRSQSRESQC(SEQ ID NO:808)、RRRYARPYRRRHIRRYRRRRRHFRRRR(SEQ ID NO:809)、MPYYYRRRRYNYRRPRWYGRGWIRRPFRRRFRRKRRVR(SEQ ID NO:216)或MAWGWWKRRRRWWFRKRWTRGRLRRRWPRSARRRPRRRRVRRRRRW RRGRRKTRTYRRRRRFRRRGRK(SEQ ID NO:186)具有至少70％(例如，至少70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％)序列同一性。

[0702] 1153.如1149-1152中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述精氨酸富集区与表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所列的精氨酸富集区序列具有至少70％、80％、85％、90％、95％、96％、97％、98％、99％或100％同一性。

[0703] 1154.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述ORF1分子包含胶冻卷结构域，例如与本文所述ORF1分子的胶冻卷结构域的氨基酸序列具有至少30％(例如至少约30％、35％、40％、50％、60％、70％、80％、90％、95％、96％、97％、98％、99％或100％)序列同一性，例如，具有氨基酸序列PTYTTIPLKQWQPPYKRTCYIKGQDCLIYYSNLRLGMNSTMYEKSIVPVHWPGGGSFSVSMLTLDALYDIHKLCRNWWTSTNQDLPLVRYKGCKITFYQSTFTDYIVRIHTELPANSNKLTYPNTHPLMMMMSKYKHIIPSRQTRRKKKPYTKIFVKPPPQFENKWYFATDLYKIPLLQIHCTACNLQNPFVKPDKLSNNVTLWSLNT(SEQ ID NO:217)的胶冻卷结构域，或表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所列的胶冻卷结构域序列。

[0704] 1155.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述ORF1分子包含N22结构域，例如与本文所述ORF1分子的N22结构域的氨基酸序列具有30％(例如至少约30％、35％、40％、50％、60％、70％、80％、90％、95％、96％、97％、98％、99％或100％)序列同一性，例如，具有氨基酸序列TMALTPFNEPIFTQIQYNPDRDTGEDTQLYLLSNATGTGWDPPGIPELILEGFPLWLIYWGFADFQKNLKKVTNIDTNYMLVAKTKFTQKPGTFYLVILNDTFVEGNSPYEKQPLPEDNIKWYPQVQYQLEAQNKLLQTGPFTPNIQGQLSDNISMFYKFYFK(SEQ ID NO:219)的N22结构域，或表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所列的N22结构域序列。

[0705] 1156.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述ORF1分子定位于所述细胞的核。

[0706] 1157.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件或分离的核酸分子相对于例如如本文所述的野生型指环病毒基因组序列的约500、1000、1100、1200、1210或1219个连续核苷酸包含不超过50％、60％、70％、80％、90％、95％、96％、97％、98％、99％或100％序列同一性。

[0707] 1158.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件或分离的核酸分子相对于例如如本文所述的野生型指甲型细环病毒(甲型细环病毒)(例如甲型细环病毒的进化分枝1、2或3)基因组序列的约500、1000、1500、2000、2100、2200、2300、2400、2500、2600、2700、2800、2900、3000、3100、3200、3300、3400、3450、3460、3470、3480、3490、3500、3510、3520、3530、3540、3550、3560、3570或3580个连续核苷酸包含不超过50％、60％、70％、80％、90％、95％、96％、97％、98％、99％或100％序列同一性。

[0708] 1159.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件或分离的核酸分子相对于例如如本文所述的野生型乙型细环病毒(Betatorquevirus)基因组序列的约500、1000、1100、1200、1210或1219个连续核苷酸包含不超过50％、60％、70％、80％、90％、95％、96％、97％、98％、99％或100％序列同一性。

[0709] 1160.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件或分离的核酸分子相对于例如如本文所述的野生型丙型细环病毒(Gammatorquevirus)基因组序列的约500、1000、1500、2000、2100、2200、2300、2400、2500、2600、2700、2800、2900、3000、3100、3120、3130、3140、3141或3142个连续核苷酸包含

不超过50%、60%、70%、80%、90%、95%、96%、97%、98%、99%或100%序列同一性。

[0710] 1161.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件或分离的核酸分子相对于例如如本文所述的野生型甲型细环病毒(例如甲型细环病毒的进化分枝1、2或3)基因组序列的至少约500，1000，1500，2000，2100，2200，2300，2400，2500，2600，2700，2800，2900，3000，3100，3200，3300，3400，3450，3460，3470，3480，3490，3500，3510，3520、3530、3540、3550、3560、3570或3580个连续核苷酸(例如，约500-3580、1000-3580、1500-3580、2000-3580或3000-3580个连续核苷酸)包含至少50%、60%、70%、80%、90%、95%、96%、97%、98%、99%或100%序列同一性。

[0711] 1162.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件或分离的核酸分子相对于例如如本文所述的野生型乙型细环病毒基因组序列的至少约500、1000、1100、1200、1210或1219个连续核苷酸(例如，约500-1000、500-1100、500-1200、500-1219、1000-1100、1000-1200或1000-1219个连续核苷酸)包含至少50%、60%、70%、80%、90%、95%、96%、97%、98%、99%或100%序列同一性。

[0712] 1163.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件或分离的核酸分子相对于例如如本文所述的野生型丙型细环病毒基因组序列的至少约500、1000、1500、2000、2100、2200、2300、2400、2500、2600、2700、2800、2900、3000、3100、3120、3130、3140、3141或3142个连续核苷酸(例如，约500-3142、1000-3142、1500-3142、2000-3142或2500-3142个连续核苷酸)包含至少50%、60%、70%、80%、90%、95%、96%、97%、98%、99%或100%序列同一性。

[0713] 1164.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件或分离的核酸分子相对于例如如本文所述的野生型TTMV-LY2基因组序列的约500、1000、1100、1200、1210或1219个连续核苷酸包含不超过50%、60%、70%、80%、90%、95%、96%、97%、98%或99%序列同一性。

[0714] 1165.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件或分离的核酸分子相对于例如如本文所述的野生型TTV-tth8基因组序列的约500、1000、1500、2000、2100、2200、2300、2400、2500、2600、2700、2800、2900、3000、3100、3200、3300、3400、3500、3550、3560、3570、3580或3581个连续核苷酸包含不超过50%、60%、70%、80%、90%、95%、96%、97%、98%或99%序列同一性。

[0715] 1166.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件或分离的核酸分子相对于例如如本文所述的野生型指环病毒基因组序列包含至少1578、1579、1580、1590、1600、1650、1700、1750或2000个核苷酸的缺失。

[0716] 1167.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件或分离的核酸分子相对于例如如本文所述的野生型指环病毒基因组序列包含1至99、1至90、1至80、1至70、1至60、1至50、10至99、10至90、10至80、10至70、10至60、10至50、20至99、20至90、20至80、20至70、20至60、20至50、30至99、30至90、30至80、30至70、30至60、30至50、40至99、40至90、40至80、40至70、40至60、或40至50个核苷酸的缺失。

[0717] 1168.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组

合物或方法，其中所述遗传元件或分离的核酸分子相对于例如如本文所述的野生型指环病毒基因组序列不具有100个核苷酸缺失、172个核苷酸缺失或1577个核苷酸。

[0718] 1169.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件或分离的核酸分子包含相对于例如如本文所述的野生型指环病毒基因组序列的三个或更多个缺失。

[0719] 1170.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件或分离的核酸分子包含与以下核酸序列具有至少75％(例如，至少75％、76％、77％、78％、79％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％)序列同一性的区域：

[0720] (i)CGCGCTGCGCGCGCCGCCCAGTAGGGGGAGCCATGC(SEQ ID NO:160)，

[0721] (ii)GCGCTX$_1$CGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:164)，其中X$_1$选自T、G或A；

[0722] (iii)GCGCTTCGCGCGCCGCCCACTAGGGGGCGTTGCGCG(SEQ ID NO:165)；

[0723] (iv)GCGCTGCGCGCGCCGCCCAGTAGGGGGCGCAATGCG(SEQ ID NO:166)；

[0724] (v)GCGCTGCGCGCGCGGCCCCCGGGGGAGGCATTGCCT(SEQ ID NO:167)；

[0725] (vi)GCGCTGCGCGCGCGCGCCGGGGGGGCGCCAGCGCCC(SEQ ID NO:168)；

[0726] (vii)GCGCTTCGCGCGCGCGCCGGGGGGCTCCGCCCCCCC(SEQ ID NO:169)；

[0727] (viii)GCGCTTCGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:170)；

[0728] (ix)GCGCTACGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:171)；或

[0729] (x)GCGCTACGCGCGCGCGCCGGGGGGCTCTGCCCCCCC(SEQ ID NO:172)。

[0730] 1171.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件或分离的核酸分子包含与以下核酸序列具有至少95％(例如，至少95％、96％、97％、98％、99％或100％)序列同一性的区域：

[0731] (i)CGCGCTGCGCGCGCCGCCCAGTAGGGGGAGCCATGC(SEQ ID NO:160)，

[0732] (ii)GCGCTX$_1$CGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:164)，其中X$_1$选自T、G或A；

[0733] (iii)GCGCTTCGCGCGCCGCCCACTAGGGGGCGTTGCGCG(SEQ ID NO:165)；

[0734] (iv)GCGCTGCGCGCGCCGCCCAGTAGGGGGCGCAATGCG(SEQ ID NO:166)；

[0735] (v)GCGCTGCGCGCGCGGCCCCCGGGGGAGGCATTGCCT(SEQ ID NO:167)；

[0736] (vi)GCGCTGCGCGCGCGCGCCGGGGGGGCGCCAGCGCCC(SEQ ID NO:168)；

[0737] (vii)GCGCTTCGCGCGCGCGCCGGGGGGCTCCGCCCCCCC(SEQ ID NO:169)；

[0738] (viii)GCGCTTCGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:170)；

[0739] (ix)GCGCTACGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:171)；或

[0740] (x)GCGCTACGCGCGCGCGCCGGGGGGCTCTGCCCCCCC(SEQ ID NO:172)。

[0741] 1172.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件或分离的核酸分子包含与核酸序列CCGCCATCTTAAGTAGTTGAGGCGGACGGTGGCGTGAGTTCAAAGGTC ACCATCAGCCACACCTACTCAAAATGGTGG(SEQ ID NO:161)具有至少75％(例如，至少75％、76％、77％、78％、79％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％)序列同一性的区域。

[0742] 1173.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件或分离的核酸分子包含与核酸序列CTTAAGTAGTTGAGGCGGACGGTGGCGTGAGTTCAAAGGTCACCATCAGCCACACCTACTCAAAATGGTGGACAATTTCTTCCGGGTCAAAGGTTACAGCCGCCATGTTAAAACACGTGACGTATGACGTCACGGCCGCCATTTTGTGACACAAGATGGCCGACTTCCTTCC(SEQ ID NO:162)具有至少75％(例如，至少75％、76％、77％、78％、79％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％)序列同一性的区域。

[0743] 1174.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件或分离的核酸分子包含具有至少80％的GC含量的至少20、25、30、31、32、33、34、35或36个连续核苷酸。

[0744] 1175.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件或分离的核酸分子包含具有至少70％、71％、72％、73％、74％、75％、76％、77％、78％、79％、80％或80.6％的GC含量的至少36个连续核苷酸。

[0745] 1176.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件或分离的核酸分子包含具有至少80％的GC含量的至少36个连续核苷酸。

[0746] 1177.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其进一步包含编码例如如本文所述的指环病毒，例如野生型指环病毒的ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3、ORF2t/3和/或ORF3的核酸序列。

[0747] 1178.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述启动子元件、编码所述效应子的核酸序列或蛋白质结合序列对应地与表A1-A12、B1-B5、C1-C5或1-18中的任一个中的指环病毒的启动子元件、编码效应子的核酸序列或蛋白质结合序列具有至少75％(例如，至少75％、76％、77％、78％、79％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％)序列同一性。

[0748] 1179.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件或分离的核酸分子包含相对于编码所述效应子的核酸序列位于3’的包装区域。

[0749] 1180.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件或分离的核酸分子包含相对于编码所述效应子的核酸序列位于5’的包装区域。

[0750] 1181.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件或分离的核酸分子包含编码指环病毒蛋白的核酸序列，所述指环病毒蛋白与本文所述的指环病毒的ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3、ORF2t/3和/或ORF3的氨基酸序列具有至少75％(例如，至少75％、76％、77％、78％、79％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％)序列同一性。

[0751] 1182.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件或分离的核酸分子包含单链DNA。

[0752] 1183.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件或分离的核酸分子是环状的和/或少于进入细胞的遗传元

件的约0.001％、0.005％、0.01％、0.05％、0.1％、0.5％、1％、1.5％、或2％的频率整合到真核细胞的基因组中。

[0753] 1184.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件或分离的核酸与野生型指环病毒序列(例如，野生型细环病毒(TTV)、细环微小病毒(TTMV)或TTMDV序列，例如野生型指环病毒序列，例如表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中的任一个中所列)，或其由约50、60、70、80、90、100、150、200、250、300、350、400、450、500、600、700、800、900、1000、1100、1200、1300、1400、1500、1600、1700、1800、1900、2000、2100、2200、2300、2400、2500、2600、2700、2800、2900或3000个连续核苷酸组成的部分具有至少75％(例如，至少75％、76％、77％、78％、79％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％)序列同一性。

[0754] 1185.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述蛋白质结合序列表20中所示的共有5'UTR序列具有至少75％(例如，至少75％、76％、77％、78％、79％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％)序列同一性。

[0755] 1186.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述蛋白质结合序列与表21中所示的共有富GC序列具有至少75％(例如，至少75％、76％、77％、78％、79％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％)序列同一性。

[0756] 1187.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述蛋白质结合序列与表38中所示的5'UTR序列和表39中所示的富GC序列具有至少75％(例如，至少75％、76％、77％、78％、79％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％)序列同一性。

[0757] 1188.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件或分离的核酸分子包含与表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中的任一个中的核酸序列的指环病毒5'UTR保守结构域具有至少85％序列同一性的序列。

[0758] 1189.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件或分离的核酸分子包含与表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中的任一个中的核酸序列的指环病毒富GC区域具有至少85％序列同一性的序列。

[0759] 1190.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述启动子元件包括RNA聚合酶II依赖性启动子、RNA聚合酶III依赖性启动子、PGK启动子、CMV启动子、EF-1α启动子、SV40启动子、CAGG启动子或UBC启动子、TTV病毒启动子、组织特异性U6(pollIII)、具有激活蛋白(TetR-VP16、Gal4-VP16、dCas9-VP16等)的上游DNA结合位点的最小CMV启动子。

[0760] 1191.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述效应子编码治疗剂，例如治疗性肽或多肽或治疗性核酸。

[0761] 1192.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组

合物或方法，其中所述效应子包含调节性核酸，例如，miRNA、siRNA、mRNA、lncRNA、RNA、DNA、反义RNA、gRNA；荧光标签或标记、抗原、肽、天然生物活性肽的合成或类似物肽、激动剂或拮抗剂肽、抗微生物肽、成孔肽、双环肽、靶向或细胞毒性肽、降解或自毁肽、小分子、免疫效应子(例如，影响对免疫应答/信号的敏感性)、死亡蛋白(例如，凋亡或坏死的诱导物)、肿瘤的非裂解性抑制剂(例如，癌蛋白抑制剂)、表观遗传修饰剂、表观遗传酶、转录因子、DNA或蛋白质修饰酶、DNA嵌入剂、外排泵抑制剂、核受体激活剂或抑制剂、蛋白酶体抑制剂、酶的竞争性抑制剂、蛋白质合成效应子或抑制剂、核酸酶、蛋白质片段或结构域、配体、抗体、受体或CRISPR系统或组分。

[0762] 1193.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述指环体能够自主复制。

[0763] 1194.如前述实施例中任一项的分离的核酸分子，其中所述表达载体选自由以下组成的组：质粒、粘粒、人工染色体、噬菌体和病毒。

[0764] 1195.一种分离的细胞，所述分离的细胞包含如前述实施例中任一项所述的分离的核酸或指环体。

[0765] 1196.如实施例195所述的分离的细胞，所述分离的细胞进一步包含例如如本文所述的例如野生型指环病毒的ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3、ORF2t/3和/或ORF3。

[0766] 1197.一种向受试者递送效应子的方法，所述方法包括向所述受试者施用如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、分离的细胞或组合物；其中所述遗传元件或分离的核酸分子编码效应子，并且其中所述效应子在所述受试者中表达。

[0767] 1198.一种治疗有需要的受试者的疾病或障碍的方法，所述方法包括向所述受试者施用如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、分离的细胞或组合物；其中所述遗传元件或分离的核酸分子编码治疗剂，并且其中所述治疗剂在所述受试者中表达。

[0768] 1199.一种将效应子递送至离体细胞或细胞群(例如，从受试者获得的细胞或细胞群)的方法，所述方法包括向所述细胞或细胞群引入如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、分离的细胞或组合物；其中所述遗传元件或分离的核酸分子编码效应子，并且其中所述效应子在所述细胞或细胞群中表达。

[0769] 1200.如前述实施例中任一项所述的指环体，其中所述遗传元件是单链DNA，并且具有以下特性中的一种或两种：呈环状和/或以少于进入细胞的遗传元件的约0.001％、0.005％、0.01％、0.05％、0.1％、0.5％、1％、1.5％或2％的频率整合到真核细胞的基因组中。

[0770] 1201.如前述实施例中任一项所述的指环体，其中所述遗传元件与野生型指环病毒序列(例如，野生型细环病毒(TTV)、细小环病毒(TTMV)或TTMDV序列，例如野生型指环病毒序列，例如表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中任一个所列)具有至少75％(例如，至少75％、76％、77％、78％、79％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％)序列同一性。

[0771] 1202.如前述实施例中任一项所述的指环体，其中所述蛋白质结合序列与表38中所示的共有5'UTR序列，或与表39中所示的共有富GC序列，或与表38中所示的共有5'UTR序列和表39中所示的共有富GC序列具有至少75％(例如至少75％、76％、77％、78％、79％、

80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％)序列同一性。

[0772] 1203.如前述实施例中任一项所述的指环体,其中所述启动子元件包括RNA聚合酶II依赖性启动子、RNA聚合酶III依赖性启动子、PGK启动子、CMV启动子、EF-1α启动子、SV40启动子、CAGG启动子或UBC启动子、TTV病毒启动子、组织特异性U6(pollIII)、具有激活蛋白(TetR-VP16、Gal4-VP16、dCas9-VP16等)的上游DNA结合位点的最小CMV启动子。

[0773] 1204.如前述实施例中任一项所述的指环体,其中所述启动子元件包含TATA盒。

[0774] 1205.如前述实施例中任一项所述的指环体,其中所述启动子元件对于野生型指环病毒,例如表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、6、9、11、13、15或17中的任一个中所列的野生型指环病毒序列是内源性的。

[0775] 1206.如前述实施例中任一项所述的指环体,其中所述启动子元件对于野生型指环病毒,例如表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、6、9、11、13、15或17中的任一个中所列的野生型指环病毒序列是外源性的。

[0776] 1207.如前述实施例中任一项所述的指环体,其中所述效应子编码治疗剂,例如治疗性肽或多肽或治疗性核酸。

[0777] 1208.如前述实施例中任一项所述的指环体,其中所述效应子包含调节性核酸,例如,miRNA、siRNA、mRNA、lncRNA、RNA、DNA、反义RNA、gRNA;荧光标签或标记、抗原、肽、天然生物活性肽的合成或类似物肽、激动剂或拮抗剂肽、抗微生物肽、成孔肽、双环肽、靶向或细胞毒性肽、降解或自毁肽、小分子、免疫效应子(例如,影响对免疫应答/信号的敏感性)、死亡蛋白(例如,凋亡或坏死的诱导物)、肿瘤的非裂解性抑制剂(例如,癌蛋白抑制剂)、表观遗传修饰剂、表观遗传酶、转录因子、DNA或蛋白质修饰酶、DNA嵌入剂、外排泵抑制剂、核受体激活剂或抑制剂、蛋白酶体抑制剂、酶的竞争性抑制剂、蛋白质合成效应子或抑制剂、核酸酶、蛋白质片段或结构域、配体、抗体、受体或CRISPR系统或组分。

[0778] 1209.如前述实施例中任一项所述的指环体,其中所述效应子包含miRNA。

[0779] 1210.如前述实施例中任一项所述的指环体,其中所述效应子例如miRNA靶向宿主基因,例如调节基因的表达,例如增加或减少基因的表达。

[0780] 1211.如前述实施例中任一项所述的指环体,其中所述效应子包含miRNA,并降低宿主基因的表达。

[0781] 1212.如前述实施例中任一项所述的指环体,其中所述效应子包含长度为约20-200、30-180、40-160、50-140或60-120个核苷酸的核酸序列。

[0782] 1213.如前述实施例中任一项所述的指环体,其中编码所述效应子的核酸序列的长度为约20-200、30-180、40-160、50-140或60-120个核苷酸。

[0783] 1214.如前述实施例中任一项所述的指环体,其中编码所述效应子的序列的大小为至少约100个核苷酸。

[0784] 1215.如前述实施例中任一项所述的指环体,其中编码所述效应子的序列的大小为约100至约5000个核苷酸。

[0785] 1216.如前述实施例中任一项所述的指环体,其中编码所述效应子的序列的大小为约100-200、200-300、300-400、400-500、500-600、600-700、700-800、800-900、900-1000、1000-1500或1500-2000个核苷酸。

[0786] 1217.如前述实施例中任一项所述的指环体，其中编码所述效应子的序列位于ORF1基因座(例如，ORF1基因座的C末端)、miRNA基因座、TATA盒上游的5'非编码区、5'UTR、聚A区下游的3'非编码区、或所述遗传元件的GC富集区上游的非编码区中一个或多个处、之内或附近(例如，5'或3')。

[0787] 1218.如实施例1217所述的指环体，其中编码所述效应子的序列位于所述遗传元件的聚A区和GC富集区之间。

[0788] 1219.如前述实施例中任一项所述的指环体，其中所述蛋白质结合序列包含与野生型指环病毒(例如表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、6、9、11、13、15或17中任一个所列的野生型指环病毒序列)的5'UTR保守结构域或GC富集结构域具有至少75％(例如，至少75％、76％、77％、78％、79％、80％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％)序列同一性的核酸序列。

[0789] 1220.如前述实施例中任一项所述的指环体，其中所述遗传元件(例如所述遗传元件的蛋白质结合序列)与以下具有至少约75％(例如，至少75％、80％、85％、90％、95％、96％、97％、98％、99％或100％)同一性：

[0790] (i)表38中所示的共有5'UTR核酸序列；

[0791] (ii)表38中所示的示例性TTV 5'UTR核酸序列；

[0792] (iii)表38所示的TTV-CT30F 5'UTR核酸序列；

[0793] (iv)表38所示的TTV-HD23a 5'UTR核酸序列；

[0794] (v)表38中所示的TTV-JA20 5'UTR核酸序列；

[0795] (vi)表38所示的TTV-TJN02 5'UTR核酸序列；

[0796] (vii)表38所示的TTV-tth8 5'UTR核酸序列；

[0797] (viii)表39所示的共有GC富集区；

[0798] (ix)表39中所示的示例性TTV GC富集区；

[0799] (x)表39中所示的TTV-CT30F GC富集区；

[0800] (xi)表39中所示的TTV-JA20 GC富集区；

[0801] (xii)表39中所示的TTV-TJN02 GC富集区；

[0802] (xiii)表39中所示的TTV-HD23a GC富集区；或

[0803] (xiv)表39中所示的TTV-tth8 GC富集区。

[0804] 1221.如前述实施例中任一项所述的指环体，其中所述蛋白质外部包含能够与蛋白质结合序列特异性结合的外部蛋白质。

[0805] 1222.如前述实施例中任一项所述的指环体，其中所述蛋白质外部包含以下一种或多种：一个或多个糖基化蛋白、亲水性DNA结合区、精苏氨酸富集区、谷氨酰胺富集区、N末端聚精氨酸序列、可变区、C末端聚谷氨酰胺/谷氨酸序列、以及一个或多个二硫键。

[0806] 1223.如前述实施例中任一项所述的指环体，其中所述蛋白质外部包含以下一种或多种特征：二十面体对称、识别和/或结合与一个或多个宿主细胞分子相互作用以介导进入宿主细胞的分子、缺乏脂质分子、缺乏碳水化合物、pH和温度稳定、耐洗涤剂、以及在宿主中基本上是非免疫原性的或基本上是非致病性的。

[0807] 1224.如前述实施例中任一项所述的指环体，其中所述蛋白质外部包含至少一个提供一个或多个功能的功能结构域，所述功能是例如，物种和/或组织和/或细胞选择性、遗

传元件结合和/或包装、免疫逃逸(基本上非免疫原性的和/或耐受性)、药代动力学、内吞作用和/或细胞附着、核进入、细胞内调节和定位、胞吐作用调节、繁殖和核酸保护。

[0808] 1225.如前述实施例中任一项所述的指环体,其中除所述效应子之外的遗传元件部分的组合大小为约2.5kb-5kb(例如,约2.8kb-4kb、约2.8kb-3.2kb、约3.6kb-3.9kb或约2.8kb-2.9kb)、小于约5kb(例如,小于约2.9kb、3.2kb、3.6kb、3.9kb或4kb)或至少100个核苷酸(例如,至少1kb)。

[0809] 1226.如前述实施例中任一项所述的指环体,其中所述遗传元件是单链的。

[0810] 1227.如前述实施例中任一项所述的指环体,其中所述遗传元件是环状的。

[0811] 1228.如前述实施例中任一项所述的指环体,其中所述遗传元件是DNA。

[0812] 1229.如前述实施例中任一项所述的指环体,其中所述遗传元件是负链DNA。

[0813] 1230.如前述实施例中任一项所述的指环体,其中所述遗传元件包括附加体。

[0814] 1231.如前述实施例中任一项所述的指环体,其中所述指环体的脂质含量按重量计小于10%、5%、2%或1%,例如不包含脂质双层。

[0815] 1232.如前述实施例中任一项所述的指环体,其中相对于包含外部脂质双层的病毒颗粒例如逆转录病毒,所述指环体抵抗被洗涤剂降解(例如,温和洗涤剂,例如,胆盐,例如脱氧胆酸钠)。

[0816] 1233.如实施例1232所述的指环体,其中至少约50%(例如,至少约50%、60%、70%、80%、85%、90%、95%、96%、97%、98%、99%、99.5%或99.9%)的指环体在37℃下与洗涤剂(例如,按重量计0.5%的洗涤剂)孵育30分钟后不会被降解。

[0817] 1234.如前述实施例中任一项所述的指环体,其中所述遗传元件包括相对于野生型指环病毒序列,例如野生型TTV序列或野生型TTMV序列的至少一个元件的缺失,例如表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中的任一个中所列的元件。

[0818] 1235.如实施例1234所述的指环体,其中所述遗传元件包含缺失,所述缺失包含对应于以下的核酸序列:

[0819] (i)TTV-tth8序列的核苷酸3436-3607,例如表5中所示的核酸序列;

[0820] (ii)TTMV-LY2序列的核苷酸574-1371和/或核苷酸1432-2210,例如表15中所示的核酸序列;

[0821] (iii)TTMV-LY2序列的核苷酸1372-1431,例如表15中所示的核酸序列;或

[0822] (iv)TTMV-LY2序列的核苷酸2610-2809,例如表15中所示的核酸序列。

[0823] 1236.如前述实施例中任一项所述的指环体,其中所述遗传元件包括野生型指环病毒序列(例如,野生型细环病毒(TTV)、细小环病毒(TTMV)或TTMDV序列,例如表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中任一个所列的序列)的至少72个核苷酸(例如至少73、74、75等个核苷酸,任选地小于基因组的全长)。

[0824] 1237.如前述实施例中任一项所述的指环体,其中所述遗传元件进一步包括以下序列中的一个或多个:编码一个或多个miRNA的序列、编码一个或多个复制蛋白的序列、编码外源基因的序列、编码治疗剂序列的序列、调节性序列(例如启动子、增强子)、编码一个或多个靶向内源基因的调节性序列(siRNA、lncRNA、shRNA)的序列、编码治疗性mRNA或蛋白质的序列以及编码细胞溶解/细胞毒性RNA或蛋白质的序列。

[0825] 1238.如前述实施例中任一项所述的指环体,其中所述指环体进一步包括第二遗

传元件，例如封闭在所述蛋白质外部内的第二遗传元件。

[0826] 1239.如实施例1238所述的指环体，其中所述第二遗传元件包含蛋白质结合序列，例如外部蛋白质结合序列，例如包装信号，例如5'UTR保守结构域或GC富集区，例如本文所述。

[0827] 1240.如前述实施例中任一项所述的指环体，其中所述指环体不能可检测地感染细菌细胞，例如感染少于1％、0.5％、0.1％或0.01％的细菌细胞。

[0828] 1241.如前述实施例中任一项所述的指环体，其中所述指环体能够例如在体外感染哺乳动物细胞，例如人细胞，例如免疫细胞、肝细胞、上皮细胞。

[0829] 1242.如前述实施例中任一项所述的指环体，其中所述遗传元件以少于进入细胞的指环体的10％、8％、6％、4％、3％、2％、1％、0.5％、0.2％、0.1％的频率整合，例如，其中所述指环体是非整合的。

[0830] 1243.如前述实施例中任一项所述的指环体，其中所述遗传元件能够复制(例如通过滚环复制)，例如能够每个细胞产生至少1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、10^2、$2x10^2$、$5x10^2$、10^3、$2x10^3$、$5x10^3$或10^4个基因组当量的遗传元件，例如通过定量PCR测定法测定。

[0831] 1244.如前述实施例中任一项所述的指环体，其中所述遗传元件能够复制(例如通过滚环复制)，例如与将所述遗传元件递送到细胞中之前的指环体中所存在量相比，能够在所述细胞中多产生至少1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、10^2、$2x10^2$、$5x10^2$、10^3、$2x10^3$、$5x10^3$或10^4个遗传元件的基因组当量，例如通过定量PCR测定法测定。

[0832] 1244A.如实施例1243或1244所述的指环体，其中相对于所述遗传元件以顺式和/或以反式提供所述蛋白质外部。

[0833] 1244B.如实施例1243-1244A中任一项所述的指环体，其中所述细胞中的辅助核酸(例如辅助病毒)编码所述蛋白质外部或其一部分(例如ORF1分子)。

[0834] 1244C.如实施例1243-1244B中任一项所述的指环体，其中一种或多种复制因子(例如，复制酶)相对于所述遗传元件以顺式和/或反式提供。

[0835] 1244D.如实施例1244C所述的指环体，其中所述细胞中的辅助核酸(例如辅助病毒)编码所述一种或多种复制因子。

[0836] 1245.如前述实施例中任一项所述的指环体，其中所述遗传元件不能够复制，例如其中所述遗传元件在复制起点发生改变或缺乏复制起点。

[0837] 1246.如前述实施例中任一项所述的指环体，其中所述遗传元件不能够自我复制，例如在不整合到宿主细胞基因组中的情况下不能够复制。

[0838] 1247.如前述实施例中任一项所述的指环体，其中所述指环体基本上是非致病性的，例如，在受试者中不诱导可检测到的有害症状(例如，相对于未暴露于指环体的受试者，增加的细胞死亡或毒性)。

[0839] 1248.如前述实施例中任一项所述的指环体，其中所述指环体基本上是非免疫原性的，例如不诱导可检测的和/或不想要的免疫应答，例如根据实例4中所述的方法检测到的。

[0840] 1249.如实施例1248所述的指环体，其中所述基本上非免疫原性的指环体在受试者中具有缺乏免疫应答的参考受试者中效力的至少约10％、20％、30％、40％、50％、60％、

70％、80％、90％、95％或100％的效力。

[0841] 1250.如实施例1248或1249所述的指环体，其中所述免疫应答包含一种或多种对所述指环体或其部分，或由其核酸编码的产物具有特异性的抗体；针对所述指环体或包含所述指环体的细胞的细胞应答(例如免疫效应细胞(例如，T细胞或NK细胞)应答)；或巨噬细胞对所述指环体或包含所述指环体的细胞的吞噬。

[0842] 1251.如前述实施例中任一项所述的指环体，其中所述指环体的免疫原性低于AAV，例如如通过本文所述的测定法测定的引发的免疫应答低于同等量的AAV所检测到的免疫应答，如通过本文所述的测定法测定的诱导的抗体流行率小于70％(例如小于约60％、50％、40％、30％、20％或10％抗体流行率)，或者基本上是非免疫原性的。

[0843] 1252.如前述实施例中任一项所述的指环体，其中具有至少1000个所述指环体的群体能够递送至少约100个拷贝(例如，至少1、2、3、4、5、10、20、30、40、50、100、200、300、400、500、600、700、800、900或1000个拷贝)的所述遗传元件进入所述真核细胞中的一个或多个。

[0844] 1253.如前述实施例中任一项所述的指环体，其中所述指环体的群体(例如，每个细胞至少1、2、3、4、5、10、20、30、40、50、100、200、300、400、500、600、700、800、900或1000基因组当量的所述遗传元件)能够将所述遗传元件递送到所述真核细胞的群体的至少10％、20％、30％、40％、50％、60％、70％、80％、90％、95％、99％或更多中，例如，其中所述真核细胞是HEK293T细胞，如实例22所述。

[0845] 1254.如前述实施例中任一项所述的指环体，其中所述指环体的群体(例如，每个细胞至少1、2、3、4、5、10、20、30、40、50、100、200、300、400、500、600、700、800、900或1000基因组当量的所述遗传元件)能够将每个细胞至少1、2、3、4、5、6、7、8、9、10、20、50、100、200、500、1000、2000、5000、8,000、$1x10^4$、$1x10^5$、$1x10^6$、$1x10^7$或更多个拷贝的所述遗传元件递送到所述真核细胞的群体中，例如，其中所述真核细胞是HEK293T细胞，如实例22所述。

[0846] 1255.如前述实施例中任一项所述的指环体，其中所述指环体的群体(例如，每个细胞至少1、2、3、4、5、10、20、30、40、50、100、200、300、400、500、600、700、800、900或1000基因组当量的所述遗传元件)能够将每个细胞1-3、1-4、1-5、1-6、1-7、1-8、1-9、1-10、5-10、10-20、20-50、50-100、100-1000、1000-10^4、$1x10^4$-$1x10^5$、$1x10^4$-$1x10^6$、$1x10^4$-$1x10^7$、$1x10^5$-$1x10^6$、$1x10^5$-$1x10^7$或$1x10^6$-$1x10^7$个拷贝的所述遗传元件递送到所述真核细胞的群体中，例如，其中所述真核细胞是HEK293T细胞，如实例22所述。

[0847] 1256.如前述实施例中任一项所述的指环体，其中所述指环体在至少两次传代后而存在。

[0848] 1257.如前述实施例中任一项所述的指环体，其中所述指环体是通过包括至少两次传代的方法而产生的。

[0849] 1258.如前述实施例中任一项所述的指环体，其中所述指环体将所述效应子选择性地递送至期望的细胞类型、组织或器官(例如，骨髓、血液、心脏、GI、皮肤、视网膜中的光感受器、上皮衬里或胰腺)或在其中以更高的水平存在(例如，优选地在其中累积)。

[0850] 1259.如前述实施例中任一项所述的指环体，其中所述真核细胞是哺乳动物细胞，例如人细胞。

[0851] 1260.如前述实施例中任一项所述的指环体，其中所述指环体或其拷贝在递送到

细胞中后24小时(例如1天、2天、3天、4天、5天、6天、1周、2周、3周、4周、30天或1个月)可在细胞中检测到。

[0852] 1261.如前述实施例中任一项所述的指环体,其中例如相对于感染细胞所用的指环体的量,在感染后3-4天后,例如,使用感染性测定法,例如根据实例7的测定法,所述指环体在细胞沉淀和上清液中以至少约10^8倍(例如,约10^5倍、10^6倍、10^7倍、10^8倍、10^9倍或10^{10}倍)基因组当量/mL产生。

[0853] 1262.一种组合物,其包括如前述实施例中任一项所述的指环体。

[0854] 1263.一种药物组合物,其包括如前述实施例中任一项所述的指环体,以及药学上可接受的载剂或赋形剂。

[0855] 1264.如实施例1262或1263所述的组合物或药物组合物,其包括至少50%、60%、70%、80%、90%、95%、96%、97%、98%、99%或更多指环体,例如合成的指环体。

[0856] 1265.如实施例1262-1264中任一项所述的组合物或药物组合物,其包括至少10^3、10^4、10^5、10^6、10^7、10^8或10^9个合成的指环体。

[0857] 1266.如实施例1262-1265中任一项所述的组合物或药物组合物,其具有以下一种或多种特征:

[0858] a)所述药物组合物满足药物或良好生产规范(GMP)标准;

[0859] b)所述药物组合物是根据良好生产规范(GMP)制成的;

[0860] c)所述药物组合物具有低于预定参考值的病原体水平,例如基本上不含病原体;

[0861] d)所述药物组合物具有低于预定参考值的污染物水平,例如,基本上不含污染物;

[0862] e)所述药物组合物具有预定水平的非感染性颗粒或预定比例的颗粒:感染性单位(例如,<300:1、<200:1、<100:1或<50:1),或

[0863] f)所述药物组合物具有低免疫原性或基本上是非免疫原性的,例如,如本文所述。

[0864] 1267.如实施例1262-1266中任一项所述的组合物或药物组合物,其中所述药物组合物具有低于预定参考值的污染物水平,例如,基本上不含污染物。

[0865] 1268.如实施例1267所述的组合物或药物组合物,其中所述污染物选自由以下组成的组:支原体、内毒素、宿主细胞核酸(例如,宿主细胞DNA和/或宿主细胞RNA)、动物来源的过程杂质(例如,血清白蛋白或胰蛋白酶)、具有复制能力的因子(RCA),例如具有复制能力的病毒或不需要的指环体(例如,除所需的指环体例如本文所述的合成的指环体以外的指环体)、游离病毒衣壳蛋白、外源因子和聚集体。

[0866] 1269.如实施例1268所述的组合物或药物组合物,其中所述污染物是宿主细胞DNA,并且阈值量是约10ng宿主细胞DNA/所述剂量药物组合物。

[0867] 1270.如实施例1262-1269中任一项所述的组合物或药物组合物,其中所述药物组合物包含按重量计小于10%(例如小于约10%、5%、4%、3%、2%、1%、0.5%或0.1%)的污染物。

[0868] 1271.如前述实施例中任一项所述的指环体、组合物或药物组合物用于治疗受试者的疾病或障碍(例如,如本文所述)的用途。

[0869] 1272.如前述实施例中任一项所述的指环体、组合物或药物组合物,其用于治疗受试者的疾病或障碍(例如,如本文所述)。

[0870] 1273.一种治疗受试者的疾病或障碍(例如,如本文所述)的方法,所述方法包括向

所述受试者施用人前述实施例中任一项所述的指环体(例如,合成的指环体)或药物组合物。

[0871] 1274.一种调节,例如增强或抑制受试者的生物学功能(例如,如本文所述)的方法,所述方法包括向所述受试者施用如前述实施例中任一项所述的指环体(例如,合成的指环体)或药物组合物。

[0872] 1275.如实施例1273-1274中任一项所述的方法,其中所述指环体不包括外源性效应子。

[0873] 1276.如实施例1273-1275中任一项所述的方法,其中所述指环体包含例如如本文所述的野生型野生型指环病毒。

[0874] 1277.如实施例1273-1276中任一项所述的方法,其中施用所述指环体例如合成的指环体导致将所述遗传元件递送到所述受试者中至少10%、20%、30%、40%、50%、60%、70%、80%、90%、95%、99%或更多的靶细胞群体中。

[0875] 1278.如实施例1273-1277中任一项所述的方法,其中施用所述指环体例如合成的指环体导致将所述效应子递送到所述受试者中至少10%、20%、30%、40%、50%、60%、70%、80%、90%、95%、99%或更多的靶细胞群体中。

[0876] 1279.如实施例1277或1278所述的方法,其中所述靶细胞包括哺乳动物细胞,例如人细胞,例如神经元(例如运动神经元)、肌细胞、视网膜细胞、肝细胞、血细胞、皮肤细胞、神经细胞、脂肪细胞,或内皮细胞,例如,体外。

[0877] 1280.如实施例1277-1279中任一项所述的方法,其中所述靶细胞存在于脊髓、肌肉、视网膜或肝脏中。

[0878] 1281.如实施例1277-1280中任一项所述的方法,其中所述遗传元件被递送到其中的靶细胞各自接受所述遗传元件的至少10、50、100、500、1000、10,000、50,000、100,000或更多个拷贝。

[0879] 1282.如实施例1273-1281中任一项所述的方法,其中所述效应子包含miRNA,并且其中所述miRNA降低(例如所述指环体被递送到其中的)细胞或细胞群体中靶蛋白或RNA的水平例如至少10%、20%、30%、40%或50%。

[0880] 1283.一种将指环体(例如,合成的指环体)递送至细胞的方法,所述方法包括将如前述实施例中任一项所述的指环体与细胞(例如,真核细胞,例如,哺乳动物细胞)接触。

[0881] 1284.如实施例1283所述的方法,其进一步包括使辅助病毒与所述细胞接触,其中所述辅助病毒包含多核苷酸,例如编码外部蛋白质例如能够结合所述外部蛋白质结合序列的外部蛋白质和任选的脂质包膜的多核苷酸。

[0882] 1285.如实施例1284所述的方法,其中使所述辅助病毒在所述指环体与所述细胞接触之前、同时或之后与所述细胞接触。

[0883] 1286.如实施例1283所述的方法,其进一步包括使辅助多核苷酸与所述细胞接触。

[0884] 1287.如实施例1286所述的方法,其中所述辅助多核苷酸包含编码外部蛋白质例如能够结合所述外部蛋白质结合序列的外部蛋白质和脂质包膜的序列多核苷酸。

[0885] 1288.如实施例1286所述的方法,其中所述辅助多核苷酸是RNA(例如,mRNA)、DNA、质粒、病毒多核苷酸或其任何组合。

[0886] 1289.如实施例1286-1288中任一项所述的方法,其中使所述辅助多核苷酸在所述

指环体与所述细胞接触之前、同时或之后与所述细胞接触。

[0887] 1290.如实施例1283-1289中任一项所述的方法，其进一步包括使辅助蛋白(例如，生长因子)与所述细胞接触。

[0888] 1291.如实施例1290所述的方法，其中所述辅助蛋白包括病毒复制蛋白或衣壳蛋白。

[0889] 1292.一种宿主细胞，其包括如前述实施例中任一项所述的指环体。

[0890] 1293.一种核酸分子，其包括启动子元件、编码效应子(例如有效负载)的序列和外部蛋白质结合序列，

[0891] 其中所述核酸分子是单链DNA，并且其中所述核酸分子呈环状和/或以少于进入细胞的核酸分子的约0.001％、0.005％、0.01％、0.05％、0.1％、0.5％、1％、1.5％或2％的频率整合；

[0892] 其中所述效应子不是源自TTV，并且不是SV40-miR-S1；

[0893] 其中所述核酸分子不包含TTMV-LY的多核苷酸序列；

[0894] 其中所述启动子元件能够指导效应子在真核细胞中的表达。

[0895] 1294.一种遗传元件，其包含：

[0896] (i)启动子元件和编码效应子例如有效负载的序列，任选地，其中所述效应子相对于野生型指环病毒序列是外源的；

[0897] (ii)与野生型指环病毒序列具有至少75％序列同一性的至少72个连续核苷酸(例如，至少72个、73个、74个、75个、76个、77个、78个、79个、80个、90个、100个或150个核苷酸)；或与野生型指环病毒序列具有至少72％(例如，至少72％、73％、74％、75％、76％、77％、78％、79％、80％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％)序列同一性的至少100个连续核苷酸；以及

[0898] (iii)蛋白质结合序列，例如外部蛋白质结合序列，以及

[0899] 其中所述核酸构建体是单链DNA；以及

[0900] 其中所述核酸构建体呈环状和/或以少于进入细胞的遗传元件的约0.001％、0.005％、0.01％、0.05％、0.1％、0.5％、1％、1.5％或2％的频率整合。

[0901] 1295.一种制造指环体组合物的方法，所述方法包括：

[0902] a)提供宿主细胞，其包含一种或多种编码指环体的组分的核酸分子，例如本文所述的合成的指环体，例如，其中所述指环体包含蛋白质外部和遗传元件，例如包含启动子元件、编码效应子的序列(例如内源性或外源性效应子)和蛋白质结合序列(例如外部蛋白质结合序列，例如包装信号)的遗传元件；

[0903] b)从所述宿主细胞产生指环体，从而制造指环体；并且

[0904] c)将所述指环体配制为例如适于向受试者施用的药物组合物。

[0905] 1296.一种制造合成的指环体组合物的方法，所述方法包括：

[0906] a)提供多个如前述实施例中任一项所述的指环体、组合物或药物组合物；

[0907] b)任选地对所述多个如前述实施例中任一项所述的指环体、组合物或药物组合物评估以下一种或多种：本文所述的污染物、光密度测量值(例如OD 260)、颗粒数(例如通过HPLC)、感染性(例如颗粒:感染单位比，例如，通过荧光和/或ELISA确定)；并且

[0908] c)例如，如果(b)的一种或多种参数满足指定的阈值，则将所述多个指环体例如配

制为适于施用于受试者的药物组合物。

[0909] 1297.如实施例1296所述的方法，其中所述指环体组合物包含至少10^5、10^6、10^7、10^8、10^9、10^{10}、10^{11}、10^{12}、10^{13}、10^{14}或10^{15}个指环体，或其中所述指环体组合物包含至少10^5、10^6、10^7、10^8、10^9、10^{10}、10^{11}、10^{12}、10^{13}、10^{14}或10^{15}个指环体基因组/mL。

[0910] 1298.如实施例1296或1297所述的方法，其中所述指环体组合物包含至少10ml、20ml、50ml、100ml、200ml、500ml、1L、2L、5L、10L、20L或50L。

[0911] 1299.一种反应混合物，所述反应混合物包含如前述实施例中任一项所述的指环体和辅助病毒，其中所述辅助病毒包含多核苷酸，例如编码外部蛋白质例如能够结合所述外部蛋白质结合序列的外部蛋白质和任选的脂质包膜的多核苷酸。

[0912] 1300.一种反应混合物，所述反应混合物包含如前述实施例中任一项所述的指环体和编码选自表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16或18、20-37或D1-D10中任一个的ORF2、ORF2/2、ORF2/3、ORF2t/3、ORF1、ORF1/1或ORF1/2的氨基酸序列或与其具有至少75％(例如75％、80％、85％、90％、95％、96％、97％、98％、99％或100％)序列同一性的氨基酸序列中一个或多个的第二核酸序列。

[0913] 1301.如实施例1300所述的反应混合物，其中所述第二核酸序列是所述遗传元件的一部分。

[0914] 1302.如实施例1301所述的反应混合物，其中所述第二核酸序列不是所述遗传元件的一部分，例如，所述第二核酸序列包含在辅助细胞或辅助病毒中。

[0915] 1303.一种合成的指环体，所述合成的指环体包含：

[0916] 遗传元件，其包含(i)编码非致病性外部蛋白质的序列，(ii)使所述遗传元件与所述非致病性外部蛋白质结合的外部蛋白质结合序列，和(iii)编码效应子例如调节性核酸的序列；以及

[0917] 与所述遗传元件相关联(例如包封或封闭所述遗传元件)的蛋白质外部。

[0918] 1304.一种药物组合物，其包含：

[0919] a)指环体，其包含：

[0920] 遗传元件，其包含(i)编码非致病性外部蛋白质的序列，(ii)使所述遗传元件与所述非致病性外部蛋白质结合的外部蛋白质结合序列，和(iii)编码效应子例如调节性核酸的序列；以及

[0921] 与所述遗传元件相关联(例如包封或封闭所述遗传元件)的蛋白质外部；以及

[0922] b)药物赋形剂。

[0923] 1305.一种药物组合物，其包含：

[0924] a)至少10^3、10^4、10^5、10^6、10^7、10^8或10^9个指环体(例如，本文所述的合成的指环体)，其包含：

[0925] 遗传元件，其包含(i)编码非致病性外部蛋白质的序列，(ii)使所述遗传元件与所述非致病性外部蛋白质结合的外部蛋白质结合序列，和(iii)编码效应子例如调节性核酸的序列；以及

[0926] 与所述遗传元件相关联(例如包封或封闭所述遗传元件)的蛋白质外部；

[0927] b)药物赋形剂，以及任选地，

[0928] c)少于预定量的：支原体、内毒素、宿主细胞核酸(例如，宿主细胞DNA和/或宿主细

胞RNA)、动物来源的过程杂质(例如,血清白蛋白或胰蛋白酶)、具有复制能力的因子(RCA),例如具有复制能力的病毒或不需要的指环体、游离病毒衣壳蛋白、外源因子、内源因子和/或聚集体。

[0929] 1306.如前述实施例中任一项所述的指环体或组合物,所述指环体或组合物进一步包括以下特征中的至少一个:所述遗传元件是单链DNA;所述遗传元件呈环状;所述指环体是非整合的;所述指环体具有基于指环病毒或其他非致病性病毒的序列、结构和/或功能,并且所述指环体是非致病性的。

[0930] 1307.如前述实施例中任一项所述的指环体或组合物,其中所述蛋白质外部包含非致病性外部蛋白质。

[0931] 1308.如前述实施例中任一项所述的指环体或组合物,其中所述蛋白质外部包含以下一种或多种:一个或多个糖基化蛋白、亲水性DNA结合区、精氨酸富集区、苏氨酸富集区、谷氨酰胺富集区、N末端聚精氨酸序列、可变区、C末端聚谷氨酰胺/谷氨酸序列、以及一个或多个二硫键。

[0932] 1309.如前述实施例中任一项所述的指环体或组合物,其中所述蛋白质外部包含以下一种或多种特征:二十面体对称、识别和/或结合与一个或多个宿主细胞分子相互作用以介导进入宿主细胞的分子、缺乏脂质分子、缺乏碳水化合物、包含一种或多种所期望的碳水化合物(例如,糖基化)、pH和温度稳定、耐洗涤剂、以及在宿主中是非免疫原性的或是非致病性的。

[0933] 1310.如前述实施例中任一项所述的指环体或组合物,其中编码所述非致病性外部蛋白质的序列包含与表19中所列的一个或多个序列或其片段具有至少70%、80%、85%、90%、95%、96%、97%、98%、99%同一性的序列。

[0934] 1311.如前述实施例中任一项所述的指环体或组合物,其中所述非致病性外部蛋白质包含至少一个提供一个或多个功能的功能结构域,所述功能是例如物种和/或组织和/或细胞嗜性、病毒基因组结合和/或包装、免疫逃逸(非免疫原性和/或耐受性)、药代动力学、内吞作用和/或细胞附着、核进入、细胞内调节和定位、胞吐作用调节、繁殖和核酸保护。

[0935] 1312.如前述实施例中任一项所述的指环体或组合物,其中所述效应子包含调节性核酸,例如,miRNA、siRNA、mRNA、lncRNA、RNA、DNA、反义RNA、gRNA;治疗剂,例如荧光标签或标记、抗原、肽治疗剂、天然生物活性肽的合成或类似物肽、激动剂或拮抗剂肽、抗微生物肽、成孔肽、双环肽、靶向或细胞毒性肽、降解或自毁肽、以及多种降解或自毁肽、小分子、免疫效应子(例如,影响对免疫应答/信号的敏感性)、死亡蛋白(例如,凋亡或坏死的诱导物)、肿瘤的非裂解性抑制剂(例如,癌蛋白抑制剂)、表观遗传修饰剂、表观遗传酶、转录因子、DNA或蛋白质修饰酶、DNA嵌入剂、外排泵抑制剂、核受体激活剂或抑制剂、蛋白酶体抑制剂、酶的竞争性抑制剂、蛋白质合成效应子或抑制剂、核酸酶、蛋白质片段或结构域、配体或受体、以及CRISPR系统或组分。

[0936] 1312b.如前述实施例中任一项所述的指环体或组合物,其中所述效应子包含因子IX或其功能变体。

[0937] 1313.如前述实施例中任一项所述的指环体或组合物,其中所述效应子包含与本文所述的一种或多种miRNA序列具有至少70%、80%、85%、90%、95%、96%、97%、98%、99%或100%同一性的序列。

[0938] 1314.如前述实施例所述的指环体或组合物,其中所述效应子例如miRNA靶向宿主基因,例如调节所述基因的表达。

[0939] 1315.如前述实施例的指环体或组合物,其中所述miRNA包含与本文所述的一种或多种miRNA序列具有至少70%、80%、85%、90%、95%、96%、97%、98%、99%或100%同一性的序列。

[0940] 1316.如前述实施例中任一项所述的指环体或组合物,其中所述遗传元件进一步包含以下一个或多个序列:编码一个或多个miRNA的序列、编码一个或多个复制蛋白的序列、编码外源基因的序列、编码治疗剂序列的序列、调节性序列(例如启动子、增强子)、编码一个或多个靶向内源基因的调节性序列(siRNA、lncRNA、shRNA)的序列、编码治疗性mRNA或蛋白质的序列以及编码细胞溶解/细胞毒性RNA或蛋白质的序列。

[0941] 1317.如前述实施例中任一项所述的指环体或组合物,其中所述遗传元件具有以下一个或多个特征:与宿主细胞的基因组不整合、是附加体核酸、是单链DNA、是约1kb至10kb、存在于细胞核内、能够与内源蛋白结合,以及产生靶向宿主基因的微小RNA。

[0942] 1318.如前述实施例中任一项所述的指环体或组合物,其中所述遗传元件包含至少一个病毒序列或与表23中列出的一个或多个序列或其片段(例如,编码ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3、ORF2t/3和/或ORF3分子的片段,和/或包含TATA盒、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号或GC富集区中的一个或多个的片段)具有至少70%、80%、85%、90%、95%、96%、97%、98%、99%同一性。

[0943] 1319.前述实施例的指环体或组合物,其中所述病毒序列来自以下至少一种:单链DNA病毒(例如指环病毒、比那病毒、圆环病毒、双生病毒、基诺病毒、丝状病毒、微小病毒、纳米病毒、细小病毒和斯派拉病毒)、双链DNA病毒(例如腺病毒、瓶状病毒、囊泡病毒、非洲猪瘟病毒、杆状病毒、福斯罗病毒、球状病毒、滴状病毒、腺炎病毒、疱疹病毒、虹彩病毒、脂毛病毒、线性病毒和痘病毒)、RNA病毒(例如甲病毒、真菌传杆状病毒、肝炎病毒、大麦病毒、烟草花叶病毒、烟草脆裂病毒、三角病毒、风疹病毒、双RNA病毒、囊状病毒、分体病毒和呼肠病毒)。

[0944] 1320.如前述实施例所述的指环体或组合物,其中所述病毒序列来自一种或多种非指环病毒,例如腺病毒、疱疹病毒、pox病毒、痘苗病毒、SV40、乳头瘤病毒、RNA病毒(例如逆转录病毒、例如慢病毒)、单链RNA病毒(例如肝炎病毒)或双链RNA病毒(例如轮状病毒)。

[0945] 1321.如前述实施例中任一项所述的指环体或组合物,其中所述蛋白质结合序列与所述蛋白质外部的精氨酸富集区相互作用。

[0946] 1322.如前述实施例中任一项所述的指环体或组合物,其中所述指环体能够在哺乳动物细胞例如人细胞中复制。

[0947] 1323.如前述实施例所述的指环体或组合物,其中所述指环体在宿主细胞中是非致病性的和/或非整合的。

[0948] 1324.如前述实施例中任一项所述的指环体或组合物,其中所述指环体在宿主中是非免疫原性的。

[0949] 1325.如前述实施例中任一项所述的指环体或组合物,其中所述指环体在宿主或宿主细胞中抑制/增强一种或多种病毒特性,例如选择性,例如感染性,例如免疫抑制/激活。

[0950] 1326.如前述实施例所述的指环体或组合物,其中所述指环体的量足以调节(例如表型、病毒水平、基因表达、与其他病毒竞争、疾病状态等至少约5%、10%、15%、20%、25%、30%、35%、40%、45%、50%或更高)。

[0951] 1327.如前述实施例中任一项所述的组合物,所述组合物进一步包含至少一种病毒或包含所述病毒的基因组的载体,例如所述指环体的变体,例如普通/天然病毒。

[0952] 1328.如前述实施例中任一项所述的组合物,所述组合物进一步包含异源部分、至少一个小分子、抗体、多肽、核酸、靶向剂、显像剂、纳米颗粒及其组合。

[0953] 1329.一种载体,所述载体包含遗传元件,所述遗传元件包含(i)编码非致病性外部蛋白质的序列,(ii)使所述遗传元件与所述非致病性外部蛋白质结合的外部蛋白质结合序列,和(iii)编码效应子例如调节性核酸的序列。

[0954] 1330.如前述实施例所述的载体,其中所述遗传元件无法与宿主细胞的基因组整合。

[0955] 1331.如前述实施例中任一项所述的载体,其中所述遗传元件能够在哺乳动物细胞例如人细胞中复制。

[0956] 1332.如前述实施例中任一项所述的载体,所述载体进一步包含外源核酸序列,例如,所述外源核酸序列被选择来调节基因例如人基因的表达。

[0957] 1333.一种药物组合物,其包含如前述实施例中任一项所述的载体和药物赋形剂。

[0958] 1334.如前述实施例所述的组合物,其中所述载体在宿主细胞中是非致病性的和/或非整合的。

[0959] 1335.如前述实施例中任一项所述的组合物,其中所述载体在宿主中是非免疫原性的。

[0960] 1336.如前述实施例所述的组合物,其中所述载体的量足以调节(例如表型、病毒水平、基因表达、与其他病毒竞争、疾病状态等至少约5%、10%、15%、20%、25%、30%、35%、40%、45%、50%或更高)。

[0961] 1337.如前述实施例中任一项所述的组合物,所述组合物进一步包含至少一种病毒或包含所述病毒的基因组的载体,例如所述指环体的变体,普通/天然病毒、辅助病毒、非指环病毒。

[0962] 1338.如前述实施例中任一项所述的组合物,所述组合物进一步包含异源部分、至少一个小分子、抗体、多肽、核酸、靶向剂、显像剂、纳米颗粒及其组合。

[0963] 1339.一种产生、繁殖和收获如前述实施例中任一项所述的指环体的方法。

[0964] 1340.一种设计和制造如前述实施例中任一项所述的载体的方法。

[0965] 1341.一种向受试者施用有效量的如前述实施例中任一项所述的组合物的方法。

[0966] 1342.一种将核酸或蛋白质有效负载递送至靶细胞、组织或受试者的方法,所述方法包括使所述靶细胞、组织或受试者与核酸组合物接触,所述核酸组合物包含(a)衍生自病毒的第一DNA序列,其中所述第一DNA序列足以产生能够感染所述靶细胞、组织或受试者的颗粒,以及(a)编码所述核酸或蛋白质有效负载的第二DNA序列,改进包括:

[0967] 所述第一DNA序列包含与表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中任一个中所列的相应序列具有至少80%(至少85%、90%、95%、97%、99%、100%)序列同一性的至少500个(至少600、700、800、900、1000、1200、1400、1500、1600、1800、2000)个

核苷酸,或

[0968] 所述第一DNA序列编码与表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中所列的ORF具有至少80%(至少85%、90%、95%、97%、99%、100%)序列同一性的序列,或

[0969] 所述第一DNA序列包含与表19中所列的共有序列具有至少90%(至少95%、97%、99%、100%)序列同一性的序列。

[0970] 1343.一种将核酸或蛋白质效应子递送至靶细胞、组织或受试者的方法,所述方法包括将所述靶细胞、组织或受试者与如前述实施例中任一项所述的指环体或核酸组合物接触,所述核酸组合物包含(a)衍生自病毒的第一DNA序列,其中所述第一DNA序列足以产生能够感染所述靶细胞、组织或受试者的如前述实施例中任一项所述的指环体,以及(a)编码所述核酸或蛋白质效应子的第二DNA序列。

[0971] 1344.一种经密码子优化的核酸分子,所述经密码子优化的核酸分子编码与野生型指环病毒ORF1、ORF2或ORF3氨基酸序列具有至少70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%同一性的氨基酸序列。

[0972] 1345.如实施例1344所述的经密码子优化的核酸分子,所述经密码子优化的核酸分子编码与例如如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所列的野生型指环病毒ORF1氨基酸序列具有至少70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%同一性的氨基酸序列。

[0973] 1346.一种药物组合物,所述药物组合物包含:

[0974] (a)指环体,例如如前述实施例中任一项所述的指环体,和

[0975] (b)载剂,其选自囊泡、脂质纳米颗粒(LNP)、红细胞、外泌体(例如哺乳动物或植物外泌体)或融合体。

[0976] 2001.一种指环体,所述指环体包含:

[0977] (a)蛋白质外部;

[0978] (b)遗传元件,其包含启动子元件和编码效应子(例如内源性效应子或外源性效应子)的核酸序列(例如DNA序列)和蛋白质结合序列(例如外部蛋白质结合序列),

[0979] 其中所述遗传元件:

[0980] (i)与如表A1所列的指环病毒序列具有至少72.2%(例如,至少72.2%、72.3%、72.4%、72.5%、73%、74%、75%、76%、77%、78%、79%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%)序列同一性;

[0981] (ii)与如表A3所列的指环病毒序列具有至少68.4%(例如,至少68.4%、68.5%、68.6%、68.7%、68.8%、68.9%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%)序列同一性;

[0982] (iii)与如表A5所列的指环病毒序列具有至少81.7%(例如,至少81.7%、81.8%、81.9%、82%、83%、84%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%)序列同一性;

[0983] (iv)与如表A7所列的指环病毒序列具有至少92.6%(例如,至少92.6%、92.7%、92.8%、92.9%、93%、94%、95%、96%、97%、98%、99%或100%)序列同一性;

[0984] (v)与如表A9所列的指环病毒序列具有至少65％(例如,至少65％、66％、67％、68％、69％、70％、75％、76％、77％、78％、79％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％)序列同一性;或

[0985] (vi)与如表A11所列的指环病毒序列具有至少65％(例如,至少65％、66％、67％、68％、69％、70％、75％、76％、77％、78％、79％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％)序列同一性;

[0986] 任选地,其中所述遗传元件包含相对于野生型指环病毒基因组序列(例如,如本文所述)的至少一个差异(例如,突变、化学修饰或表观遗传改变),例如插入、取代、酶促修饰和/或缺失,例如结构域(例如TATA框、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号或GC富集区中的一个或多个)的缺失;

[0987] 其中所述遗传元件被封闭在所述蛋白质外部内;并且

[0988] 其中所述指环体被配置为将遗传元件递送到真核细胞中。

[0989] 2002.一种指环体,所述指环体包含:

[0990] (a)蛋白质外部;

[0991] (b)遗传元件,其包含启动子元件和编码效应子(例如内源性效应子或外源性效应子)的核酸序列(例如DNA序列)和蛋白质结合序列(例如外部蛋白质结合序列),

[0992] 其中所述遗传元件包含不超过约:

[0993] (i)1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、25、30、35、40、45、50、60、70、80、90、100、125、150、175、200、250、300、350、400、450、500、600、700、800、900、1000、1010、1011、1012、1013、1014、1015、1016或1017个核苷酸差异,例如取代、插入或缺失,这是相对于如表A1所列的指环病毒序列;

[0994] (ii)1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、25、30、35、40、45、50、60、70、80、90、100、125、150、175、200、250、300、350、400、450、500、600、700、800、900、1000、1100、1110、1120、1130、1140、1150、11160、1170、1171、1172、1173或1174个核苷酸差异,例如取代、插入或缺失,这是相对于如表A3所列的指环病毒序列;

[0995] (iii)1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、25、30、35、40、45、50、60、70、80、90、100、125、150、175、200、250、300、350、400、450、500、600、610、620、630、640、650、660、670、671或672个核苷酸差异,例如取代、插入或缺失,这是相对于如表A5所列的指环病毒序列;

[0996] (iv)1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、25、30、35、40、45、50、60、70、80、90、100、125、150、175、200、250、260、270、271、272、273、274、275、276、277、278、279或280个核苷酸差异,例如取代、插入或缺失,这是相对于如表A7所列的指环病毒序列;

[0997] (v)1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、25、30、35、40、45、50、60、70、80、90、100、125、150、175、200、250、300、350、400、450、500、600、700、800、900或1000个核苷酸差异,例如取代、插入或缺失,这是相对于如表A9所列的指环病毒序列;或

[0998] (vi)1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、25、30、35、40、45、50、60、70、80、90、100、125、150、175、200、250、300、350、400、450、500、600、700、800、900或1000个核苷酸差异,例如取代、插入或缺失,这是相对于如表A11所列的指环病毒序列;

[0999] 任选地,其中所述遗传元件包含相对于野生型指环病毒基因组序列(例如,如本文

所述)的至少一个差异(例如,突变、化学修饰或表观遗传改变),例如插入、取代、酶促修饰和/或缺失,例如结构域(例如TATA框、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号或GC富集区中的一个或多个)的缺失;

[1000] 其中所述遗传元件被封闭在所述蛋白质外部内;并且

[1001] 其中所述指环体被配置为将遗传元件递送到真核细胞中。

[1002] 2002.一种指环体,所述指环体包含:

[1003] (a)蛋白质外部;

[1004] (b)遗传元件,其包含启动子元件和编码效应子(例如内源性效应子或外源性效应子)的核酸序列(例如DNA序列)和蛋白质结合序列(例如外部蛋白质结合序列),

[1005] 其中所述遗传元件包含不超过约:

[1006] (i)1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、25、30、35、40、45、50、60、70、80、90、100、125、150、175、200、250、300、350、400、450、500、600、700、800、900、1000、1010、1011、1012、1013、1014、1015、1016、或1017个核苷酸差异,例如取代、插入或缺失,这是相对于如表B1所列的指环病毒序列;

[1007] (ii)1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、25、30、35、40、45、50、60、70、80、90、100、125、150、175、200、250、300、350、400、450、500、600、700、800、900、1000、1100、1110、1120、1130、1140、1150、11160、1170、1171、1172、1173、或1174个核苷酸差异,例如取代、插入或缺失,这是相对于如表B2所列的指环病毒序列;

[1008] (iii)1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、25、30、35、40、45、50、60、70、80、90、100、125、150、175、200、250、300、350、400、450、500、600、610、620、630、640、650、660、670、671、或672个核苷酸差异,例如取代、插入或缺失,这是相对于如表B3所列的指环病毒序列;

[1009] (iv)1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、25、30、35、40、45、50、60、70、80、90、100、125、150、175、200、250、260、270、271、272、273、274、275、276、277、278、279、或280个核苷酸差异,例如取代、插入或缺失,这是相对于如表B4所列的指环病毒序列;或

[1010] (v)1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、25、30、35、40、45、50、60、70、80、90、100、125、150、175、200、250、300、350、400、450、500、600、700、800、900、或1000个核苷酸差异,例如取代、插入或缺失,这是相对于如表B5所列的指环病毒序列;

[1011] 任选地,其中所述遗传元件包含相对于野生型指环病毒基因组序列(例如,如本文所述)的至少一个差异(例如,突变、化学修饰或表观遗传改变),例如插入、取代、酶促修饰和/或缺失,例如结构域(例如TATA框、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号或GC富集区中的一个或多个)的缺失;

[1012] 其中所述遗传元件被封闭在所述蛋白质外部内;并且

[1013] 其中所述指环体被配置为将遗传元件递送到真核细胞中。

[1014] 2003.如前述实施例中任一项所述的指环体,其中所述遗传元件不是天然存在的序列(例如,相对于野生型指环病毒序列(例如,野生型细环病毒(TTV)、细环微小病毒(TTMV)或TTMDV序列,例如野生型指环病毒序列,例如,如表B1-B5、A1、A3、A5、A7、A9、A11、1、3、5、7、9、11或13中的任一个中所列)包含至少一个差异(例如,突变、化学修饰或表观遗传改变),例如插入、取代、酶促修饰和/或缺失,例如结构域(例如,TATA盒、帽位点、转录起始

位点、5'UTR、开放阅读框(ORF)、聚(A)信号中或GC富集区中的一个或多个)的缺失)。

[1015] 2004.如前述实施例中任一项所述的指环体,指环体包含多肽,所述多肽包含与指环病毒ORF1分子的氨基酸序列(例如,表C1-C5、A2、A4、A6、A8、A10或A12中任一个中所列的指环病毒ORF1序列)具有至少70%(例如,至少约70、80、90、95%、96%、97%、98%、99%或100%)序列同一性的氨基酸序列。

[1016] 2005.如实施例2004所述的指环体,其中所述蛋白质外部包含所述多肽。

[1017] 2006.如实施例2005所述的指环体,其中所述蛋白质外部中至少60%(例如至少60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)的蛋白质包含所述多肽。

[1018] 2007.如前述实施例中任一项所述的指环体,其中所述蛋白质外部中至少60%(例如,至少60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98、99%或100%)的蛋白质包含ORF1分子。

[1019] 2008.如前述实施例中任一项所述的指环体,所述指环体包含编码氨基酸序列的核酸分子(例如,在遗传元件中),所述氨基酸序列与指环病毒ORF1分子的氨基酸序列(例如表C1-C5、A2、A4、A6、A8、A10或A12中任一个中所列的指环病毒ORF1序列)具有至少70%(例如,至少约70%、80%、90%、95%、96%、97%、98%、99%或100%)序列同一性)。

[1020] 2009.如前述实施例中任一项所述的指环体,其中所述遗传元件包含区域,所述区域包含以下核酸序列的至少10、15、20、25、30、31、32、33、34、35或36个连续核苷酸:

[1021] (i)CGCGCTGCGCGCGCCGCCCAGTAGGGGGAGCCATGC(SEQ ID NO:160),

[1022] (ii)GCGCTX$_1$CGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:164),其中X$_1$选自T、G或A;

[1023] (iii)GCGCTTCGCGCGCCGCCCACTAGGGGGCGTTGCGCG(SEQ ID NO:165);

[1024] (iv)GCGCTGCGCGCGCCGCCCAGTAGGGGGCGCAATGCG(SEQ ID NO:166);

[1025] (v)GCGCTGCGCGCGCGGCCCCCGGGGGAGGCATTGCCT(SEQ ID NO:167);

[1026] (vi)GCGCTGCGCGCGCGCCGGGGGGGCGCCAGCGCCC(SEQ ID NO:168);

[1027] (vii)GCGCTTCGCGCGCGCCGGGGGGCTCCGCCCCCCC(SEQ ID NO:169);

[1028] (viii)GCGCTTCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:170);

[1029] (ix)GCGCTACGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:171);或

[1030] (x)GCGCTACGCGCGCGCCGGGGGGCTCTGCCCCCCC(SEQ ID NO:172);

[1031] 或与其具有至少75%、76%、77%、78%、79%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。

[1032] 2010.如前述实施例中任一项所述的指环体,其中所述遗传元件包含如本文所述的5'UTR区和/或GC富集区(例如,分别列于表38或39中)。

[1033] 2011.一种分离的核酸分子(例如,表达载体),所述分离的核酸分子包含遗传元件,所述遗传元件:

[1034] (i)与如表A1所列的指环病毒序列具有至少72.2%(例如,至少72.2%、72.3%、72.4%、72.5%、73%、74%、75%、76%、77%、78%、79%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%)序列同一性;

[1035] (ii)与如表A3所列的指环病毒序列具有至少68.4%(例如,至少68.4%、68.5%、

68.6%、68.7%、68.8%、68.9%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%)序列同一性;

[1036] (iii)与如表A5所列的指环病毒序列具有至少81.7%(例如,至少81.7%、81.8%、81.9%、82%、83%、84%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%)序列同一性;

[1037] (iv)与如表A7所列的指环病毒序列具有至少92.6%(例如,至少92.6%、92.7%、92.8%、92.9%、93%、94%、95%、96%、97%、98%、99%或100%)序列同一性;

[1038] (v)与如表A9所列的指环病毒序列具有至少65%(例如,至少65%、66%、67%、68%、69%、70%、75%、76%、77%、78%、79%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%)序列同一性;或

[1039] (vi)与如表A11所列的指环病毒序列具有至少65%(例如,至少65%、66%、67%、68%、69%、70%、75%、76%、77%、78%、79%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%)序列同一性;

[1040] 任选地,其中所述遗传元件包含相对于野生型指环病毒基因组序列(例如,如本文所述)的至少一个差异(例如,突变、化学修饰或表观遗传改变),例如插入、取代、酶促修饰和/或缺失,例如结构域(例如TATA框、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号或GC富集区中的一个或多个)的缺失。

[1041] 2012.一种分离的核酸分子(例如,表达载体),所述分离的核酸分子包含遗传元件,所述遗传元件包含不超过约:

[1042] (i)1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、25、30、35、40、45、50、60、70、80、90、100、125、150、175、200、250、300、350、400、450、500、600、700、800、900、1000、1010、1011、1012、1013、1014、1015、1016或1017个核苷酸差异,例如取代、插入或缺失,这是相对于如表A1所列的指环病毒序列;

[1043] (ii)1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、25、30、35、40、45、50、60、70、80、90、100、125、150、175、200、250、300、350、400、450、500、600、700、800、900、1000、1100、1110、1120、1130、1140、1150、11160、1170、1171、1172、1173或1174个核苷酸差异,例如取代、插入或缺失,这是相对于如表A3所列的指环病毒序列;

[1044] (iii)1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、25、30、35、40、45、50、60、70、80、90、100、125、150、175、200、250、300、350、400、450、500、600、610、620、630、640、650、660、670、671或672个核苷酸差异,例如取代、插入或缺失,这是相对于如表A5所列的指环病毒序列;

[1045] (iv)1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、25、30、35、40、45、50、60、70、80、90、100、125、150、175、200、250、260、270、271、272、273、274、275、276、277、278、279或280个核苷酸差异,例如取代、插入或缺失,这是相对于如表A7所列的指环病毒序列;

[1046] (v)1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、25、30、35、40、45、50、60、70、80、90、100、125、150、175、200、250、300、350、400、450、500、600、700、800、900或1000个核苷酸差异,例如取代、插入或缺失,这是相对于如表A9所列的指环病毒序列;或

[1047] (vi)1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、25、30、35、40、45、50、60、70、

80、90、100、125、150、175、200、250、300、350、400、450、500、600、700、800、900或1000个核苷酸差异，例如取代、插入或缺失，这是相对于如表A11所列的指环病毒序列；

[1048] 任选地，其中所述遗传元件包含相对于野生型指环病毒基因组序列(例如，如本文所述)的至少一个差异(例如，突变、化学修饰或表观遗传改变)，例如插入、取代、酶促修饰和/或缺失，例如结构域(例如TATA框、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号或GC富集区中的一个或多个)的缺失。

[1049] 2012A.一种分离的核酸分子(例如，表达载体)，所述分离的核酸分子包含遗传元件，所述遗传元件包含不超过约：

[1050] (i)1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、25、30、35、40、45、50、60、70、80、90、100、125、150、175、200、250、300、350、400、450、500、600、700、800、900、1000、1010、1011、1012、1013、1014、1015、1016、或1017个核苷酸差异，例如取代、插入或缺失，这是相对于如表B1所列的指环病毒序列；

[1051] (ii)1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、25、30、35、40、45、50、60、70、80、90、100、125、150、175、200、250、300、350、400、450、500、600、700、800、900、1000、1100、1110、1120、1130、1140、1150、11160、1170、1171、1172、1173、或1174个核苷酸差异，例如取代、插入或缺失，这是相对于如表B2所列的指环病毒序列；

[1052] (iii)1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、25、30、35、40、45、50、60、70、80、90、100、125、150、175、200、250、300、350、400、450、500、600、610、620、630、640、650、660、670、671、或672个核苷酸差异，例如取代、插入或缺失，这是相对于如表B3所列的指环病毒序列；

[1053] (iv)1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、25、30、35、40、45、50、60、70、80、90、100、125、150、175、200、250、260、270、271、272、273、274、275、276、277、278、279、或280个核苷酸差异，例如取代、插入或缺失，这是相对于如表B4所列的指环病毒序列；或

[1054] (v)1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、25、30、35、40、45、50、60、70、80、90、100、125、150、175、200、250、300、350、400、450、500、600、700、800、900、或1000个核苷酸差异，例如取代、插入或缺失，这是相对于如表B5所列的指环病毒序列；

[1055] 任选地，其中所述遗传元件包含相对于野生型指环病毒基因组序列(例如，如本文所述)的至少一个差异(例如，突变、化学修饰或表观遗传改变)，例如插入、取代、酶促修饰和/或缺失，例如结构域(例如TATA框、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号或GC富集区中的一个或多个)的缺失。

[1056] 2013.如前述实施例中任一项所述的分离的核酸分子，其中所述遗传元件不是天然存在的序列(例如，相对于野生型指环病毒序列(例如，野生型细环病毒(TTV)、细环微小病毒(TTMV)或TTMDV序列，例如野生型指环病毒序列，例如，如表B1-B5、A1、A3、A5、A7、A9、A11、1、3、5、7、9、11或13中的任一个中所列)包含至少一个差异(例如，突变、化学修饰或表观遗传改变)，例如插入、取代、酶促修饰和/或缺失，例如结构域(例如，TATA盒、帽位点、转录起始位点、5'UTR、开放阅读框(ORF)、聚(A)信号中或GC富集区中的一个或多个)的缺失)。

[1057] 2014.如前述实施例中任一项所述的分离的核酸分子，其中所述分离的核酸分子包含编码ORF1分子(例如，表C1-C5、A2、A4、A6、A8、A10或A12中的任一个中所列的ORF1分子，或与其具有至少70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同

一性的多肽)；

[1058] 其中：

[1059] (i)所述ORF1分子的至少30％(例如至少30％、35％、40％、45％、50％、55％、60％、65％、70％、75％、80％、90％或更多)的氨基酸是β-折叠的一部分；

[1060] (ii)所述ORF1分子的二级结构包含至少三(例如至少3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20)个β-折叠；

[1061] (iii)所述ORF1分子的二级结构包含β-折叠与α-螺旋的比例为至少1:1、2:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1或10:1；并且

[1062] 2015.如前述实施例中任一项所述的分离的核酸分子，所述分离的核酸分子包含以下核酸序列的至少10、15、20、25、30、31、32、33、34、35或36个连续核苷酸：

[1063] (i)CGCGCTGCGCGCGCCGCCCAGTAGGGGGAGCCATGC(SEQ ID NO:160)，

[1064] (ii)GCGCTX$_1$CGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:164)，其中X$_1$选自T、G或A；

[1065] (iii)GCGCTTCGCGCGCCGCCCACTAGGGGGCGTTGCGCG(SEQ ID NO:165)；

[1066] (iv)GCGCTGCGCGCGCCGCCCAGTAGGGGGCGCAATGCG(SEQ ID NO:166)；

[1067] (v)GCGCTGCGCGCGCGGCCCCCGGGGGAGGCATTGCCT(SEQ ID NO:167)；

[1068] (vi)GCGCTGCGCGCGCGCGCCGGGGGGGCGCCAGCGCCC(SEQ ID NO:168)；

[1069] (vii)GCGCTTCGCGCGCGCCGGGGGGCTCCGCCCCCCC(SEQ ID NO:169)；

[1070] (viii)GCGCTTCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:170)；

[1071] (ix)GCGCTACGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:171)；或

[1072] (x)GCGCTACGCGCGCGCCGGGGGGCTCTGCCCCCCC(SEQ ID NO:172)；

[1073] 或与其具有至少75％、76％、77％、78％、79％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。

[1074] 2016.如前述实施例中任一项所述的分离的核酸分子，所述分离的核酸分子包含具有至少70％、71％、72％、73％、74％、75％、76％、77％、78％、79％、80％或80.6％的GC含量的至少20、25、30、31、32、33、34、35或36个连续核苷酸。

[1075] 2017.如前述实施例中任一项所述的分离的核酸分子，其中所述遗传元件进一步包含以下中的一种或多种：TATA盒、起始元件、帽位点、转录起始位点、5'UTR保守结构域、ORF1编码序列、ORF1/1编码序列、ORF1/2编码序列、ORF2编码序列、ORF2/2编码序列、ORF2/3编码序列、ORF2/3t编码序列、三开放阅读框区域、聚(A)信号和/或来自本文所述的指环病毒的GC富集区(例如，如表B1-B5、A1、A3、A5、A7、A9或A11中的任一个中所列)，或与其具有至少70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的序列。

[1076] 2018.如前述实施例中任一项所述的分离的核酸分子，其中所述遗传元件进一步包含指环病毒基因组序列(例如，如本文所述，例如，如表B1-B5、A1、A3、A5、A7、A9、A11、1、3、5、7、9、11、13、15或17中的任一个中所列)的至少一个或两个拷贝(例如，1、2、3、4、5或6个拷贝)，或与其具有至少70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的序列。

[1077] 2019.如前述实施例中任一项所述的分离的核酸分子，所述分离的核酸分子进一步包含所述遗传元件的至少一个另外拷贝(例如，总共1、2、3、4、5或6个拷贝)。

[1078] 2020.如前述实施例中任一项所述的分离的核酸分子,其中所述分离的核酸分子是环状的。

[1079] 2021.一种分离的核酸组合物(例如,包含一种、两种或更多种核酸分子),所述分离的核酸组合物包含编码如前述实施例中任一项所述的分离的核酸。

[1080] 2022.如前述实施例中任一项所述的分离的核酸,其中所述遗传元件进一步包含启动子元件、编码效应子(例如内源性效应子或外源性效应子)的核酸序列(例如,DNA序列)、和/或蛋白质结合序列(例如,外部蛋白质结合序列)。

[1081] 2022A.如前述实施例中任一项所述的分离的核酸分子,其中所述遗传元件包含所述ORF1的高变结构域(HVD)中的插入或取代。

[1082] 2023.如前述实施例中任一项所述的指环体或分离的核酸分子,其中所述遗传元件包含以下中的一种或多种:TATA盒、起始位点、5'UTR保守结构域、ORF1、ORF2、ORF2下游序列、ORF2、ORF3和/或GC富集区,或与其具有至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%同一性的序列,例如,如表B1-B5、A1、A3、A5、A7、A9或A11中的任一个中所示。

[1083] 2024.如前述实施例中任一项所述的指环体或分离的核酸,所述指环体或分离的核酸包含(例如,在蛋白质外部中)或编码一种或多种包含以下的多肽:选自表C1-C5、A2、A4、A6、A8、A10或A12中任一个中的ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3、ORF2t/3和/或ORF3的氨基酸序列,或与其具有至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。

[1084] 2025.如前述实施例中任一项所述的指环体或分离的核酸,其中所述遗传元件包含序列,所述序列包含具有至少70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%或80.6%的GC含量的至少20、25、30、31、32、33、34、35或36个连续核苷酸的。

[1085] 2026.如实施例2025所述的指环体或分离的核酸,其中所述遗传元件包含具有至少80%的GC含量的至少20、25、30、31、32、33、34、35或36个连续核苷酸。

[1086] 2027.如实施例2025所述的指环体或分离的核酸,其中所述遗传元件包含具有至少70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%或80.6%的GC含量的至少36个连续核苷酸。

[1087] 2028.如实施例2025所述的指环体或分离的核酸,其中所述遗传元件包含具有至少80%的GC含量的至少36个连续核苷酸。

[1088] 2029.如前述实施例中任一项所述的指环体或分离的核酸,其中所述遗传元件包含区域(例如包装区域),所述区域包含以下核酸序列的至少10、15、20、25、30、31、32、33、34、35或36个连续核苷酸:

[1089] (i)CGCGCTGCGCGCGCCGCCCAGTAGGGGGAGCCATGC(SEQ ID NO:160),

[1090] (ii)GCGCTX$_1$CGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:164),其中X$_1$选自T、G或A;

[1091] (iii)GCGCTTCGCGCGCCGCCCACTAGGGGGCGTTGCGCG(SEQ ID NO:165);

[1092] (iv)GCGCTGCGCGCGCCGCCCAGTAGGGGGCGCAATGCG(SEQ ID NO:166);

[1093] (v)GCGCTGCGCGCGCGGCCCCCGGGGGAGGCATTGCCT(SEQ ID NO:167);

[1094] (vi)GCGCTGCGCGCGCGCGCCGGGGGGGCGCCAGCGCCC(SEQ ID NO:168);

[1095] (vii)GCGCTTCGCGCGCGCCGGGGGGCTCCGCCCCCCC(SEQ ID NO:169)；

[1096] (viii)GCGCTTCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:170)；

[1097] (ix)GCGCTACGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:171)；或

[1098] (x)GCGCTACGCGCGCGCCGGGGGGCTCTGCCCCCCC(SEQ ID NO:172)；

[1099] 或与其具有至少75％、76％、77％、78％、79％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。

[1100] 2030.如实施例2029所述的指环体或分离的核酸，其中所述包装区域相对于编码所述效应子的核酸序列位于3’。

[1101] 2031.一种多肽，所述多肽包含以下中的一个或多个：

[1102] (a)第一区域，其包含与本文所述的指环病毒ORF1分子的精氨酸富集区序列(例如，表C1-C5、A2、A4、A6、A8、A10或A12中任一个中所列的指环病毒ORF1序列)具有至少70％(例如，至少约70、80、90、95％、96％、97％、98％、99％或100％)序列同一性的氨基酸序列；

[1103] (b)第二区域，其包含与本文所述的指环病毒ORF1分子的胶冻卷区序列(例如，表C1-C5、A2、A4、A6、A8、A10或A12中任一个中所列的指环病毒ORF1序列)具有至少30％(例如，至少约30％、35％、40％、50％、60％、70％、80％、90％、95％、96％、97％、98％、99％或100％)序列同一性的氨基酸序列；

[1104] (c)第三区域，其包含与本文所述的指环病毒ORF1分子的N22结构域序列(例如，表C1-C5、A2、A4、A6、A8、A10或A12中任一个中所列的指环病毒ORF1序列)具有至少30％(例如，至少约30％、35％、40％、50％、60％、70％、80％、90％、95％、96％、97％、98％、99％或100％)序列同一性的氨基酸序列；和/或

[1105] (d)第四区域，其包含与本文所述的指环病毒ORF1分子的指环病毒ORF1 C末端结构域(CTD)序列(例如，表C1-C5、A2、A4、A6、A8、A10或A12中任一个中所列的指环病毒ORF1序列)具有至少30％(例如，至少约30％、35％、40％、50％、60％、70％、80％、90％、95％、96％、97％、98％、99％或100％)序列同一性的氨基酸序列；

[1106] 其中所述ORF1分子包含相对于野生型ORF1蛋白(例如，如本文所述)的至少一个差异(例如，突变、化学修饰或表观遗传改变)，例如插入、取代、化学或酶促修饰，和/或缺失，例如结构域(例如，精氨酸富集区、胶冻卷结构域、HVR、N22或CTD中的一个或多个，例如如本文所述)的缺失。

[1107] 2031A.如实施例2031所述的多肽，所述多肽包含以下中的一个或多个：

[1108] (a)第一区域，其包含与本文所述的指环病毒ORF1分子的精氨酸富集区序列(例如，表C1-C5、A2、A4、A6、A8、A10或A12中任一个中所列的指环病毒ORF1序列)具有至少90％序列同一性的氨基酸序列；

[1109] (b)第二区域，其包含与本文所述的指环病毒ORF1分子的胶冻卷区序列(例如，表C1-C5、A2、A4、A6、A8、A10或A12中任一个中所列的指环病毒ORF1序列)具有至少90％序列同一性的氨基酸序列；

[1110] (c)第三区域，其包含与本文所述的指环病毒ORF1分子的N22结构域序列(例如，表C1-C5、A2、A4、A6、A8、A10或A12中任一个中所列的指环病毒ORF1序列)具有至少90％序列同一性的氨基酸序列；和/或

[1111] (d)第四区域，其包含与本文所述的指环病毒ORF1分子的指环病毒ORF1 C末端结

构域(CTD)序列(例如,表C1-C5、A2、A4、A6、A8、A10或A12中任一个中所列的指环病毒ORF1序列)具有至少90％序列同一性的氨基酸序列;

[1112] 其中所述ORF1分子包含相对于野生型ORF1蛋白(例如,如本文所述)的至少一个差异(例如,突变、化学修饰或表观遗传改变),例如插入、取代、化学或酶促修饰,和/或缺失,例如结构域(例如,精氨酸富集区、胶冻卷结构域、HVR、N22或CTD中的一个或多个,例如如本文所述)的缺失。

[1113] 2032.如实施例2031所述的多肽,其中所述多肽包含:

[1114] (i)所述第一区域和所述第二区域;

[1115] (ii)所述第一区域和所述第三区域;

[1116] (iii)所述第一区域和所述第四区域;

[1117] (iv)所述第二区域和所述第三区域;

[1118] (v)所述第二区域和所述第四区域;

[1119] (vi)所述第三区域和所述第四区域;

[1120] (vii)所述第一区域、所述第二区域和所述第三区域;

[1121] (viii)所述第一区域、所述第二区域和所述第四区域;

[1122] (ix)所述第一区域、所述第三区域和所述第四区域;或

[1123] (x)所述第二区域、所述第三区域和所述第四区域。

[1124] 2033.如实施例2031或2032所述的多肽,其中所述多肽以N末端到C末端的顺序包含所述第一区域、所述第二区域、所述第三区域和所述第四区域。

[1125] 2034.如前述实施例中任一项所述的多肽,所述多肽进一步包含氨基酸序列,例如高变区(HVR)序列(例如,如本文所述的指环病毒ORF1分子的HVR序列),其中所述氨基酸序列包含至少约55(例如,至少约45、50、51、52、53、54、55、56、57、58、59、60或65)个氨基酸(例如,约45-160、50-160、55-160、60-160、45-150、50-150、55-150、60-150、45-140、50-140、55-140或60-140个氨基酸)。

[1126] 2035.如实施例2034所述的多肽,其中所述HVR包含与本文所述的指环病毒ORF1分子的指环病毒ORF1 HVR序列(例如,表C1-C5、A2、A4、A6、A8、A10或A12中任一个中所列的指环病毒ORF1序列)具有至少30％(例如,至少约30％、35％、40％、50％、60％、70％、80％、90％、95％、96％、97％、98％、99％或100％)序列同一性的氨基酸序列。

[1127] 2036.如实施例2034或2035所述的多肽,其中所述HVR序列位于所述第二区域和所述第三区域之间。

[1128] 2037.如实施例2034-2036中任一项所述的多肽,其中所述HVR包含如本文所述的HVR的一个或多个特征。

[1129] 2038.一种多肽,所述多肽包含表C1-C5、A2、A4、A6、A8、A10或A12中任一个中的ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3、ORF2t/3和/或ORF3的氨基酸序列,或与其具有至少75％、80％、85％、90％、95％、96％、97％、98％、99％或100％的序列同一性,并且其中所述多肽进一步包含相对于野生型指环病毒ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3、ORF2t/3和/或ORF3序列(例如,如本文所述,例如,如表C1-C5、A2、A4、A6、A8、A10或A12中的任一个中所列)的至少一个差异(例如,突变或化学修饰),例如缀合、添加、插入、取代和/或缺失,例如结构域的缺失。

[1130] 2039.一种多肽，所述多肽包含表C1-C5、A2、A4、A6、A8、A10或A12中任一个中的ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3、ORF2t/3和/或ORF3的氨基酸序列，或与其具有至少75％、80％、85％、90％、95％、96％、97％、98％、99％或100％的序列同一性。

[1131] 2040.一种多肽，所述多肽与选自表C1-C5、A2、A4、A6、A8、A10或A12中任一个中的ORF1、ORF2、ORF2或ORF3的氨基酸序列具有至少75％、80％、85％、90％、95％、96％、97％或98％但不超过99％的序列同一性。

[1132] 2041.一种多肽，所述多肽相对于选自表C1-C5、A2、A4、A6、A8、A10或A12中任一个中的ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3、ORF2t/3和/或ORF3的氨基酸序列具有至少1个但不超过2、5、10、20、50或100个氨基酸差异，例如取代、插入或缺失。

[1133] 2042.如前述实施例中任一项所述的多肽，其中所述多肽是分离的多肽。

[1134] 2043.一种复合物，其包含：

[1135] (a)如前述实施例中任一项所述的多肽，以及

[1136] (b)遗传元件，其包含启动子元件和编码效应子(例如外源性效应子或内源性效应子)的核酸序列(例如，DNA序列)，和蛋白质结合序列。

[1137] 2044.如实施例2043所述的复合物，其中所述复合物包含如本文所述的复合物的一个或多个特征。

[1138] 2045.一种融合蛋白，所述融合蛋白包含第一氨基酸序列以及异源性部分，所述第一氨基酸序列选自表C1-C5、A2、A4、A6、A8、A10或A12中任一个中的ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3、ORF2t/3和/或ORF3分子，或与其具有至少75％、80％、85％、90％、95％、96％、97％、98％、99％或100％的序列同一性。

[1139] 2046.一种融合蛋白，所述融合蛋白包含第一氨基酸序列以及异源性部分，所述第一氨基酸序列选自表C1-C5、A2、A4、A6、A8、A10或A12中任一个中的ORF1分子，或与其具有至少75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性。

[1140] 2047.如前述实施例中任一项所述的融合蛋白，其中所述异源性部分包含靶向部分。

[1141] 2048.如前述实施例中任一项所述的融合蛋白，其中所述第一氨基酸序列包含相对于野生型指环病毒ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3、ORF2t/3和/或ORF3序列(例如，如本文所述，例如，如表C1-C5、A2、A4、A6、A8、A10或A12中的任一个中所列)的至少一个差异(例如，突变或化学修饰)，例如缀合、添加、插入、取代和/或缺失，例如结构域的缺失。

[1142] 2049.一种宿主细胞，所述宿主细胞包含如前述实施例中任一项所述的指环体、分离的核酸、融合蛋白或多肽。

[1143] 2050.一种反应混合物，所述反应混合物包含如前述实施例中任一项所述的指环体和辅助病毒，其中所述辅助病毒包含多核苷酸，例如编码外部蛋白质例如结合所述外部蛋白质结合序列的外部蛋白质和任选的脂质包膜的多核苷酸。

[1144] 2051.一种治疗受试者的疾病或障碍的方法，所述方法包括向所述受试者施用如前述实施例中任一项所述的指环体、分离的核酸分子、融合蛋白或多肽或如前述实施例中任一项所述的药物组合物。

[1145] 2052.如实施例2051所述的方法，其中所述疾病或障碍选自免疫障碍、感染性疾

病、炎性障碍、自身免疫病、癌症(例如实体瘤)和胃肠道障碍。

[1146] 2053.如前述实施例中任一项所述的指环体、分离的核酸、融合蛋白或多肽用于治疗受试者的疾病或障碍的用途。

[1147] 2054.如实施例2053所述的用途,其中所述疾病或障碍选自免疫障碍、感染性疾病、炎性障碍、自身免疫病、癌症(例如实体瘤,例如肺癌)和胃肠道障碍。

[1148] 2055.如前述实施例中任一项所述的指环体、分离的核酸、组合物或药物组合物,其用于治疗受试者的疾病或障碍。

[1149] 2055A.如前述实施例中任一项所述的指环体、分离的核酸、组合物或药物组合物,其用于用作药物。

[1150] 2056.一种调节,例如抑制或增强受试者的生物学功能的方法,所述方法包括向所述受试者施用如前述实施例中任一项所述的指环体、分离的核酸、融合蛋白或多肽或如前述实施例中任一项所述的药物组合物。

[1151] 2057.一种将指环体递送至细胞的方法,所述方法包括使如前述实施例中任一项所述的指环体、分离的核酸、融合蛋白或多肽与细胞例如真核细胞例如哺乳动物细胞接触。

[1152] 2058.如实施例2057所述的方法,其进一步包括使辅助病毒与所述细胞接触,其中所述辅助病毒包含多核苷酸,例如编码外部蛋白质例如结合所述外部蛋白质结合序列的外部蛋白质和任选的脂质包膜的多核苷酸。

[1153] 2059.如实施例2058所述的方法,其中使所述辅助病毒在所述指环体与所述细胞接触之前、同时或之后与所述细胞接触。

[1154] 2060.如实施例2057所述的方法,其进一步包括使辅助多核苷酸与所述细胞接触。

[1155] 2061.如实施例2060所述的方法,其中所述辅助多核苷酸包含编码外部蛋白质例如结合所述外部蛋白质结合序列的外部蛋白质和脂质包膜的序列多核苷酸。

[1156] 2062.如实施例2060所述的方法,其中所述辅助多核苷酸是RNA(例如,mRNA)、DNA、质粒、病毒多核苷酸或其任何组合。

[1157] 2063.如实施例2060-2062中任一项所述的方法,其中使所述辅助多核苷酸在所述指环体与所述细胞接触之前、同时或之后与所述细胞接触。

[1158] 2064.如实施例2057-2063中任一项所述的方法,其进一步包括使辅助蛋白与所述细胞接触。

[1159] 2065.如实施例2064所述的方法,其中所述辅助蛋白包括病毒复制蛋白或衣壳蛋白。

[1160] 2066.一种将核酸或蛋白质效应子递送至靶细胞、组织或受试者的方法,所述方法包括使所述靶细胞、组织或受试者与核酸组合物接触,所述核酸组合物包含(a)衍生自病毒的第一DNA序列,其中所述第一DNA序列足以产生感染所述靶细胞、组织或受试者的颗粒,以及(a)编码所述核酸或蛋白质效应子的第二DNA序列,改进包括:

[1161] 所述第一DNA序列包含与表B1-B5、A1、A3、A5、A7、A9或A11中任一个中所列的相应序列具有至少80%(至少85%、90%、95%、97%、99%、100%)序列同一性的至少500个(至少600、700、800、900、1000、1200、1400、1500、1600、1800、2000)个核苷酸,或

[1162] 所述第一DNA序列编码与指环病毒ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3、ORF2t/3和/或ORF3分子(例如,表C1-C5、A2、A4、A6、A8、A10或A12中的任一个中所列)具有至

少80％(至少85％、90％、95％、97％、99％、100％)序列同一性的序列。

[1163] 2067.一种制造指环体组合物的方法,所述方法包括:

[1164] a)提供宿主细胞,其包含一种或多种编码如前述实施例中任一项所述的指环体的组分的核酸分子,例如本文所述的合成的指环体,例如,其中所述指环体包含蛋白质外部和遗传元件,例如包含启动子元件、编码效应子的序列(例如内源性效应子或外源性效应子)和蛋白质结合序列(例如外部蛋白质结合序列,例如包装信号)的遗传元件;

[1165] b)从所述宿主细胞产生指环体,从而制造指环体;并且

[1166] c)将所述指环体配制为例如适于向受试者施用的药物组合物;

[1167] 任选地,其中所述一种或多种核酸分子编码辅助蛋白。

[1168] 2068.一种制造指环体组合物的方法,所述方法包括:

[1169] a)提供多个如前述实施例中任一项所述的指环体;

[1170] b)任选地对所述多个如前述实施例中任一项所述的指环体评估以下一种或多种:本文所述的污染物、光密度测量值(例如OD 260)、颗粒数(例如通过HPLC)、感染性(例如颗粒:感染单位比);并且

[1171] c)例如,如果(b)的一种或多种参数满足指定的阈值,则将所述多个指环体例如配制为适于施用于受试者的药物组合物。

[1172] 2069.如实施例2068所述的方法,其中所述指环体组合物包含至少10^5、10^6、10^7、10^8、10^9、10^{10}、10^{11}、10^{12}、10^{13}、10^{14}或10^{15}个指环体。

[1173] 2070.如实施例2068或2069所述的方法,其中所述指环体组合物包含至少10ml、20ml、50ml、100ml、200ml、500ml、1L、2L、5L、10L、20L或50L。

[1174] 2071.如前述实施例中任一项所述的指环体或分离的核酸,其中所述遗传元件被配置为在哺乳动物细胞例如人细胞中复制。

[1175] 2072.如前述实施例中任一项所述的指环体或分离的核酸,其中遗传元件进一步包含外源核酸序列,例如,选择用于调节基因,例如人基因的表达。

[1176] 2073.如前述实施例中任一项所述的指环体或分离的核酸,其中所述蛋白质结合序列的至少60％(例如至少60％、65％、70％、75％、80％、85％、90％、95％或100％)由G或C组成。

[1177] 2074.如前述实施例中任一项所述的指环体或分离的核酸,其中所述遗传元件包含长度为至少80、90、100、110、120、130或140个核苷酸的序列,所述序列的至少70％(例如,至少70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％)或约70％-100％、75％-95％、80％-95％、85％-95％、或85％-90％位置由G或C组成。

[1178] 2075.如前述实施例中任一项所述的指环体或分离的核酸,其中所述蛋白质结合序列结合所述蛋白质外部的精氨酸富集区。

[1179] 2076.如前述实施例中任一项所述的指环体或分离的核酸,其中所述蛋白质外部包含与所述蛋白质结合序列特异性结合的外部蛋白质。

[1180] 2077.如前述实施例中任一项所述的指环体或分离的核酸,其中除所述效应子之外的遗传元件部分的组合大小为约2.5kb-5kb(例如,约2.8kb-4kb、约2.8kb-3.2kb、约3.6kb-3.9kb或约2.8kb-2.9kb)、小于约5kb(例如,小于约2.9kb、3.2kb、3.6kb、3.9kb或4kb)或至少100个核苷酸(例如,至少1kb)。

[1181] 2078.如前述实施例中任一项所述的指环体或分离的核酸,其中所述遗传元件是单链的。

[1182] 2079.如前述实施例中任一项所述的指环体或分离的核酸,其中所述遗传元件是环状的。

[1183] 2080.如前述实施例中任一项所述的指环体或分离的核酸,其中所述遗传元件是DNA。

[1184] 2081.如前述实施例中任一项所述的指环体或分离的核酸,其中遗传元件是负链DNA。

[1185] 2082.如前述实施例中任一项所述的指环体或分离的核酸,其中所述遗传元件包含附加体。

[1186] 2083.如前述实施例中任一项所述的指环体或分离的核酸,其中所述指环体在所需器官或组织中以相对于其他器官或组织更高的水平存在(例如,优先积累在其中)。

[1187] 2084.如前述实施例中任一项所述的指环体或分离的核酸,其中所述真核细胞是哺乳动物细胞,例如人细胞。

[1188] 2085.一种组合物,所述组合物包含如前述实施例中任一项所述的指环体或分离的核酸。

[1189] 2086.一种药物组合物,所述药物组合物包含指环体或分离的核酸以及药学上可接受的载剂或赋形剂。

[1190] 2087.一种药物组合物,其包含:

[1191] a)至少10^3、10^4、10^5、10^6、10^7、10^8或10^9个如前述实施例中任一项所述的指环体;

[1192] b)药物赋形剂,以及任选地,

[1193] c)少于预定量的:支原体、内毒素、宿主细胞核酸(例如,宿主细胞DNA和/或宿主细胞RNA)、动物来源的过程杂质(例如,血清白蛋白或胰蛋白酶)、具有复制能力的因子(RCA),例如具有复制能力的病毒或不需要的指环体、游离病毒衣壳蛋白、外源因子和/或聚集体。

[1194] 2088.如实施例2085或2086所述的组合物或药物组合物,其包含至少50%、60%、70%、80%、90%、95%、96%、97%、98%、99%或更多指环体,例如合成的指环体。

[1195] 2089.如实施例2085-2088中任一项所述的组合物或药物组合物,其包含至少10^3、10^4、10^5、10^6、10^7、10^8或10^9个指环体。

[1196] 2090.一种药物组合物,其包含:

[1197] a)至少10^3、10^4、10^5、10^6、10^7、10^8或10^9个如前述实施例中任一项所述的指环体;

[1198] b)药物赋形剂,以及任选地,

[1199] c)少于预定量的:支原体、内毒素、宿主细胞核酸(例如,宿主细胞DNA和/或宿主细胞RNA)、动物来源的过程杂质(例如,血清白蛋白或胰蛋白酶)、具有复制能力的因子(RCA),例如具有复制能力的病毒或不需要的指环体、游离病毒衣壳蛋白、外源因子和/或聚集体。

[1200] 2091.如实施例2085-2090中任一项所述的组合物或药物组合物,其具有以下一种或多种特征:

[1201] a)所述药物组合物满足药物或良好生产规范(GMP)标准;

[1202] b)所述药物组合物是根据良好生产规范(GMP)制成的;

[1203] c)所述药物组合物具有低于预定参考值的病原体水平,例如基本上不含病原体;

[1204] d)所述药物组合物具有低于预定参考值的污染物水平,例如,基本上不含污染物;

[1205] e)所述药物组合物具有预定水平的非感染性颗粒或预定比例的颗粒:感染性单位(例如,<300:1、≤200:1、≤100:1或<50:1),或

[1206] f)所述药物组合物具有低免疫原性或基本上是非免疫原性的,例如,如本文所述。

[1207] 2092.如实施例2085-2091中任一项所述的组合物或药物组合物,其中所述药物组合物具有低于预定参考值的污染物水平,例如,基本上不含污染物。

[1208] 2093.如实施例92所述的组合物或药物组合物,其中所述污染物选自由以下组成的组:支原体、内毒素、宿主细胞核酸(例如,宿主细胞DNA和/或宿主细胞RNA)、动物来源的过程杂质(例如,血清白蛋白或胰蛋白酶)、具有复制能力的因子(RCA),例如具有复制能力的病毒或不需要的指环体(例如,除所需的指环体例如本文所述的合成的指环体以外的指环体)、游离病毒衣壳蛋白、外源因子和聚集体。

[1209] 2094.如实施例2093所述的组合物或药物组合物,其中所述污染物是宿主细胞DNA,并且阈值量是约500ng宿主细胞DNA/所述剂量药物组合物。

[1210] 2095.如实施例2085-2094中任一项所述的组合物或药物组合物,其中所述药物组合物包含按重量计小于10%(例如小于约10%、5%、4%、3%、2%、1%、0.5%或0.1%)的污染物。

[1211] 2096.如前述实施例中任一项所述的方法,其中所述指环体不包含外源性效应子。

[1212] 2097.如前述实施例中任一项所述的方法,其中施用所述指环体例如合成的指环体导致将所述遗传元件递送到所述受试者中至少10%、20%、30%、40%、50%、60%、70%、80%、90%、95%、99%或更多的靶细胞群体中。

[1213] 2098.如前述实施例中任一项所述的方法,其中施用所述指环体例如合成的指环体导致将所述外源性效应子递送到所述受试者中至少10%、20%、30%、40%、50%、60%、70%、80%、90%、95%、99%或更多的靶细胞群体中。

[1214] 2099.如实施例2097或2098所述的方法,其中所述靶细胞包括例如体外的哺乳动物细胞,例如人细胞,例如免疫细胞、肝细胞、肺上皮细胞。

[1215] 2100.如实施例2097-2099中任一项所述的方法,其中所述靶细胞存在于肝或肺中。

[1216] 2101.如实施例2097-2100中任一项所述的方法,其中所述遗传元件被递送到其中的靶细胞各自接受所述遗传元件的至少10、50、100、500、1000、10,000、50,000、100,000或更多个拷贝。

[1217] 2102.如前述实施例中任一项所述的方法,其中所述效应子包含miRNA,并且任选地其中所述miRNA降低(例如所述指环体被递送到其中的)细胞或细胞群体中靶蛋白或RNA的水平例如至少10%、20%、30%、40%或50%。

[1218] 2103.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法,其中所述遗传元件(例如,遗传元件的5'UTR)物理地关联(例如,结合)至所述蛋白质外部(例如,蛋白质外部的ORF1分子)。

[1219] 2104.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法,其中封闭在所述蛋白质外部的遗传元件对核酸内切酶消化具有抗性,例如根据Martin等人(2013,Hum.Gene Ther.Methods[人类基因疗法与方法]24(4):253-269,通过

引用以其整体并入本文)中描述的方法确定;任选地,其中使用的DNA酶的量为约60U/ml或约300U。

[1220] 2105.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法,其中所述遗传元件包含长度至少为100个核苷酸的序列,所述序列在至少80%的位置上由G或C组成。

[1221] 2106.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法,其中所述遗传元件是环状单链DNA。

[1222] 2107.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法,其中所述遗传元件不包含一个或多个细菌质粒元件(例如,细菌复制起点或可选择标记,例如,细菌抗性基因)。

[1223] 2108.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法,其中所述遗传元件以少于进入哺乳细胞的指环体的1%的频率整合。

[1224] 2109.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法,其中所述启动子元件对于野生型指环体病毒是外源的或内源的。

[1225] 2110.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法,其中所述外源性效应子是治疗性外源性效应子,例如治疗性肽、治疗性多肽或治疗性核酸(例如,miRNA)。

[1226] 2111.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法,其中具有至少1000(例如至少1000、1500、2000、3000、4000、5000、6000、7000、8000、9000、10,000、20,000、50,000、75,000、100,000、200,000、500,000、1,000,000或更多)个所述指环体的群体将至少100(例如至少100、150、200、250、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、10,000、50,000、100,000或更多)个拷贝的所述遗传元件递送到一个或更多个哺乳动物细胞中。

[1227] 2112.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法,其中所述指环体包含一种或多种多肽,所述多肽包含以下中的一种或多种:选自指环病毒ORF2、ORF2/2、ORF2/3、ORF1、ORF1/1或ORF1/2(例如,如本文所述)的氨基酸序列或与其具有至少95%序列同一性的氨基酸序列。

[1228] 2113.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法,其中所述遗传元件包含编码以下的核酸序列:选自指环病毒ORF1、ORF2、ORF2/2、ORF2/3、ORF1/1或ORF1/2(例如,如本文所述)的氨基酸序列,或与其具有至少95%序列同一性的氨基酸序列。

[1229] 2114.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法,其中所述指环体不包含编码复制因子和衣壳蛋白之一或两者的多核苷酸,或其中所述指环体是复制缺陷的。

[1230] 2115.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法,其中使所述指环体与细胞在体外或体内接触。

[1231] 2116.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法,其中所述指环体不包含与指环病毒ORF2、ORF2/2、ORF2/3、ORF1/1或ORF1/2(例如,如本文所述)具有至少95%序列同一性的多肽。

[1232] 2117.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件能够通过滚环复制被扩增(例如，在细胞中，例如宿主细胞，例如哺乳动物细胞，例如人细胞，例如HEK293T或A549细胞)，例如产生至少2、4、8、16、32、64、128、256、518或1024个拷贝。

[1233] 2118.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件由双链环状DNA分子产生。

[1234] 2119.如实施例2118所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述双链环状DNA分子通过体外环化产生。

[1235] 2118.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件由包含所述遗传元件的核酸序列的两个拷贝的DNA分子产生。

[1236] 2119.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件的核酸序列的两个拷贝在所述DNA分子中串联排列。

[1237] 2120.一种核酸分子，所述核酸分子包含核酸序列的两个拷贝，所述核酸序列包含指环体遗传元件(例如，前述实施例中任一项所述的遗传元件)的5'UTR。

[1238] 2121.一种核酸分子，所述核酸分子包含启动子元件；编码外源性效应子的核酸序列；如表B1-B5中任一个中所列的5'UTR序列，或与其具有至少85％(例如，至少85％、90％、95％、96％、97％、98％、99％、或100％)同一性的核酸序列；和表B1-B5中任一个中所列的GC富集区，或与其具有至少85％(例如，至少85％、90％、95％、96％、97％、98％、99％，或100％)同一性的核酸序列。

[1239] 2122.如实施例2121所述的核酸分子，其中所述核酸分子是单链或双链的。

[1240] 2123.如实施例2121所述的核酸分子，其中所述核酸分子是环状的。

[1241] 2124.如前述实施例中任一项所述的多肽、复合物、指环体、分离的核酸、细胞、组合物或方法，其中所述遗传元件包含含有以下核酸序列的5'UTR：

[1242] $CGGGAGCCX_1CGAGGTGAGTGAAACCACCGAGGTCTAGGGGCAATTCGGGCTAGGGCAGTCTAGCGGAACGGG$，其中$X_1$是C或不存在，

[1243] 或与其至少95％相同的核酸序列。

[1244] 3001.一种合成的指环体，所述合成的指环体包含：

[1245] (I)遗传元件，其包含：

[1246] (a)启动子元件，

[1247] (b)编码外源性效应子的核酸序列，其中所述核酸序列与所述启动子元件可操作地连接，

[1248] 其中所述外源性效应子包含当突变时引起人疾病的多肽，或所述多肽的功能变体，选自：

[1249] (i)治疗性酶(例如，因子IX、UGT1A1或OCA1)，

[1250] (ii)结构蛋白质(例如SMN、肌营养不良蛋白)，或

[1251] (iii)补体蛋白质(例如C1)；并且

[1252] (c)5'UTR结构域，其包含：

[1253] SEQ ID NO:54的核苷酸323-393的核酸序列，或与其至少85％相同的核酸序列；

[1254] SEQ ID NO:113、SEQ ID NO:114、SEQ ID NO:115、SEQ ID NO:116、SEQ ID NO:117、SEQ ID NO:118、SEQ ID NO:119的核酸序列或与其至少85％相同的核酸序列;或

[1255] SEQ ID NO:61的核苷酸117-187的核酸序列,或与其至少85％相同的核酸序列;

[1256] (II)包含ORF1分子的蛋白质外部;

[1257] 其中所述遗传元件被封闭在所述蛋白质外部内;并且

[1258] 其中所述合成的指环体能够将所述遗传元件递送到人细胞中。

[1259] 3002.如实施例3001所述的合成的指环体,其中所述外源性效应子包含因子IX或其功能变体。

[1260] 3003.如实施例3001所述的合成的指环体,其中外源性效应子是表A的酶或其功能变体。

[1261] 3004.如实施例3001所述的合成的指环体,其中所述效应子是表B的多肽或其功能变体。

[1262] 3005.如实施例3001所述的合成的指环体,其中所述ORF1分子包含SEQ ID NO:217的氨基酸序列,或与其具有至少90％同一性的氨基酸序列。

[1263] 3006.如前述实施例中任一项所述的合成的指环体,其中所述ORF1分子由SEQ ID NO:54的核苷酸612-2612编码。

[1264] 3007.如前述实施例中任一项所述的合成的指环体,其中所述遗传元件包含SEQ ID NO:54的核苷酸2868-2929的核酸序列,或与其具有至少85％序列同一性的核酸序列。

[1265] 3008.如前述实施例中任一项所述的合成是指环体,其中所述ORF1分子包含氨基酸序列,所述氨基酸序列包含以下中的一种或多种:如表16中所列的精氨酸富集区、胶冻卷结构域、高变结构域、N22结构域和/或C末端结构域的氨基酸序列,或与其具有至少85％同一性的氨基酸序列。

[1266] 3009.如前述实施例中任一项所述的合成的指环体,其中所述ORF1分子包含SEQ ID NO:58的氨基酸序列,或与其具有至少85％序列同一性的核酸序列。

[1267] 3010.如前述实施例中任一项所述的合成的指环体,其进一步包含含有如表16中所列的ORF2、ORF2/2、ORF2/3、ORF1/1或ORF1/2的氨基酸序列或与其具有至少85％同一性的氨基酸序列的多肽。

[1268] 3011.如前述实施例中任一项所述的合成的指环体,其中所述遗传元件编码如表16中所列的ORF1、ORF2、ORF2/2、ORF2/3、ORF1/1或ORF1/2的氨基酸序列,或与其具有至少85％同一性的氨基酸序列。

[1269] 3012.如前述实施例中任一项所述的合成的指环体,其中所述合成的指环体不包含含有如表16中所列的ORF2、ORF2/2、ORF2/3、TAIP、ORF1/1或ORF1/2的氨基酸序列或与其具有至少85％同一性的氨基酸序列的多肽。

[1270] 3013.如前述实施例中任一项所述的合成的指环体,其中所述遗传元件不编码如表16中所列的ORF1、ORF2、ORF2/2、ORF2/3、ORF1/1或ORF1/2的氨基酸序列,或与其具有至少85％同一性的氨基酸序列。

[1271] 3014.如前述实施例中任一项所述的合成的指环体,其中所述ORF1分子包含氨基酸序列$YNPX^2DXGX^2N$(SEQ ID NO:829),其中X^n各自独立地是任何n个氨基酸的连续序列。

[1272] 3015.如实施例3014所述的合成的指环体,其中所述ORF1分子进一步包含位于氨

基酸序列YNPX^{2}DXGX^{2}N(SEQ ID NO:829)侧翼的第一β链和第二β链,例如,其中所述第一β链包含氨基酸序列YNPX^{2}DXGX^{2}N(SEQ ID NO:829)的酪氨酸(Y)残基和/或其中所述第二β链包含氨基酸序列YNPX^{2}DXGX^{2}N(SEQ ID NO:829)的第二个天冬酰胺(N)残基(从N到C)。

[1273] 3016.如前述实施例中任一项所述的合成的指环体,其中所述ORF1分子按从N末端到C末端方向的顺序包含第一β链、第二β链、第一α螺旋、第三β链、第四β链、第五β链、第二α螺旋、第六β链、第七β链、第八β链和第九β链。

[1274] 3017.如前述实施例中任一项所述的合成的指环体,其中所述遗传元件能够通过滚环复制在宿主细胞中扩增,例如产生至少8个拷贝。

[1275] 3018.如前述实施例中任一项所述的合成的指环体,其中所述遗传元件是单链的。

[1276] 3019.如前述实施例中任一项所述的合成的指环体,其中所述遗传元件是环状的。

[1277] 3020.如前述实施例中任一项所述的合成的指环体,其中所述遗传元件是DNA。

[1278] 3021.如前述实施例中任一项所述的合成的指环体,其中所述遗传元件是负链DNA。

[1279] 3022.如前述实施例中任一项所述的合成的指环体,其中所述遗传元件以少于进入细胞的指环体的10%、8%、6%、4%、3%、2%、1%、0.5%、0.2%、0.1%的频率整合,例如,其中所述合成的指环体是非整合的。

[1280] 3023.如前述实施例中任一项所述的合成的指环体,其中所述遗传元件包含表16-1中所示的共有5'UTR核酸序列的序列。

[1281] 3024.如前述实施例中任一项所述的合成的指环体,其中所述遗传元件包含表16-2中所示的共有GC富集区的序列。

[1282] 3025.如前述实施例中任一项所述的合成的指环体,其中所述遗传元件包含长度为至少100个核苷酸的序列,所述序列的至少70%(例如,约70%-100%、75%-95%、80%-95%、85%-95%、或85%-90%)位置由G或C组成。

[1283] 3026.如前述实施例中任一项所述的合成的指环体,其中所述遗传元件包含SEQ ID NO:120的核酸序列。

[1284] 3027.如前述实施例中任一项所述的合成的指环体,其中所述启动子元件对于野生型指环病毒是外源的。

[1285] 3028.如前述实施例中任一项所述的合成的指环体,其中所述启动子元件对于野生型指环病毒是内源的。

[1286] 3029.如前述实施例中任一项所述的合成的指环体,其中所述外源性效应子包含肽、天然生物活性肽的合成或类似肽、激动剂肽或拮抗剂肽、酶的竞争性抑制剂、配体、抗体或受体。

[1287] 3030.如前述实施例中任一项所述的合成的指环体,其中编码所述外源性效应子的核酸序列是约20-200、30-180、40-160、50-140、60-120、200-2000、200-500、500-1000、1000-1500或1500-2000个核苷酸长度。

[1288] 3031.如前述实施例中任一项所述的合成的指环体,其中所述遗传元件的长度为约1.5-2.0、2.0-2.5、2.5-3.0、3.0-3.5、3.1-3.6、3.2-3.7、3.3-3.8、3.4-3.9、3.5-4.0、4.0-4.5或4.5-5.0kb。

[1289] 3032.如前述实施例中任一项所述的合成的指环体,其中所述合成的指环体能够

感染人细胞，例如神经元(例如运动神经元)、肌细胞、视网膜细胞、肝细胞、血细胞、皮肤细胞、神经细胞、脂肪细胞，或内皮细胞，例如，体外。

[1290] 3033.如前述实施例中任一项所述的合成的指环体，所述合成的指环体是基本上非免疫原性的，例如不诱导可检测的和/或不想要的免疫应答，例如根据实例4中所述的方法检测到的。

[1291] 3034.如实施例3033所述的合成的指环体，其中所述基本上非免疫原性的指环体在受试者中具有缺乏免疫应答的参考受试者中效力的至少约10％、20％、30％、40％、50％、60％、70％、80％、90％、95％或100％的效力。

[1292] 3035.如前述实施例中任一项所述的合成的指环体，其中具有至少1000个所述指环体的群体能够递送至少约100个拷贝(例如，至少1、2、3、4、5、10、20、30、40、50、100、200、300、400、500、600、700、800、900或1000个拷贝)的所述遗传元件进入一个或多个人细胞中。

[1293] 3036.一种药物组合物，其包含如前述实施例中任一项所述的合成的指环体，以及药学上可接受的载剂或赋形剂。

[1294] 3037.如实施例3036所述的药物组合物，其包含至少10^3、10^4、105、10^6、10^7、10^8或10^9个合成的指环体。

[1295] 3038.如实施例3036或3037所述的药物组合物，其中所述药物组合物具有预定比例的颗粒:感染性单位(例如，<300:1、<200:1、<100:1或<50:1)。

[1296] 3039.一种反应混合物，所述反应混合物包含：

[1297] (i)第一核酸(例如，双链或单链环状DNA)，其包含前述实施例中任一项所述的合成的指环体的遗传元件的序列，和

[1298] (ii)第二核酸序列，其编码选自例如如表16中所列的ORF1、ORF2、ORF2/2、ORF2/3、ORF1/1或ORF1/2中的一种或多种氨基酸序列，或与其具有至少85％序列同一性的氨基酸序列。

[1299] 3040.如实施例3039所述的反应混合物，其中所述第一核酸和所述第二核酸在同一核酸分子中。

[1300] 3041.如实施例3039所述的反应混合物，其中所述第一核酸和所述第二核酸是不同的核酸分子。

[1301] 3042.如实施例3039所述的反应混合物，其中所述第一核酸和所述第二核酸是不同的核酸分子并且其中所述第二核酸作为双链环状DNA提供。

[1302] 3043.如实施例3039所述的反应混合物，其中所述第一核酸和所述第二核酸是不同的核酸分子并且其中所述第一核酸和所述第二核酸作为双链环状DNA提供。

[1303] 3044.如实施例3041所述的反应混合物，其中所述第二核酸序列包含在辅助细胞或辅助病毒中。

[1304] 3045.一种制备合成的指环体的方法，所述方法包括：

[1305] a)提供宿主细胞，其包含：

[1306] (i)第一核酸分子，其包含如前述实施例中任一项所述的合成的指环体的遗传元件的核酸序列，和

[1307] (ii)第二核酸分子，其编码选自例如如表16中任一个中所列的ORF1、ORF2、ORF2/2、ORF2/3、ORF1/1或ORF1/2中的一种或多种氨基酸序列，或与其具有至少85％序列同一性

的氨基酸序列;并且

[1308] b)在适合于制备合成的指环体的条件下孵育所述宿主细胞;

[1309] 从而制备所述合成的指环体。

[1310] 3046.如实施例3045所述的方法,所述方法进一步包括,在步骤(a)之前,将所述第一核酸分子和/或所述第二核酸分子引入所述细胞。

[1311] 3047.如实施例3046所述的方法,其中所述第二核酸分子在所述第一核酸分子之前、与其同时或在其之后被引入所述宿主细胞。

[1312] 3048.如实施例3045或3046中任一项所述的方法,其中所述第二核酸分子被整合到所述宿主细胞的基因组中。

[1313] 3049.如实施例3045-3048中任一项所述的方法,其中所述第二核酸分子是辅助物(例如辅助质粒,或辅助病毒的基因组)。

[1314] 3050.如实施例3045-3048中任一项所述的方法,其中第二核酸分子编码包含氨基酸序列[W/F]X^7HX^3CX^1CX^5H(SEQ ID NO:949)的ORF2分子,其中X^n是任何n个氨基酸的连续序列。

[1315] 3051.一种制造合成的指环体制剂的方法,所述方法包括:

[1316] a)提供多个如实施例3001-3035所述的合成的指环体、如实施例3036-3038任一项所述的药物组合物、或如实施例3039-3044任一项所述的反应混合物;

[1317] b)任选地对所述多个如实施例3001-3035所述的合成的指环体评估以下一种或多种:本文所述的污染物、光密度测量值(例如OD 260)、颗粒数(例如通过HPLC)、感染性(例如颗粒:感染单位比);并且

[1318] c)例如,如果(b)的一种或多种参数满足指定的阈值,则将所述多个合成的指环体配制为例如适于施用于受试者的药物组合物。

[1319] 3052.一种宿主细胞,所述宿主细胞包含:

[1320] (i)第一核酸分子,其包含如前述实施例中任一项所述的合成的指环体的遗传元件的核酸序列,和

[1321] (ii)任选地,第二核酸分子,其编码选自如表16中任一个中所列的ORF1、ORF2、ORF2/2、ORF2/3、ORF1/1或ORF1/2中的一种或多种氨基酸序列,或与其具有至少85%序列同一性的氨基酸序列。

[1322] 3053.一种将外源性效应子(例如,治疗性外源性效应子)递送至哺乳动物细胞的方法,所述方法包括:

[1323] (a)提供如前述实施例中任一项所述的合成的指环体;并且

[1324] (b)使哺乳动物细胞与所述合成的指环体接触;

[1325] 其中所述合成的指环体能够将所述遗传元件递送到所述哺乳动物细胞中;以及

[1326] 任选地,其中在适合在宿主细胞中将所述遗传元件封闭在蛋白质外部内的条件下,通过将所述遗传元件引入所述宿主细胞来产生所述合成的指环体;

[1327] 从而将所述治疗性外源性效应子递送至所述哺乳动物细胞。

[1328] 3054.如实施例3001-3035中任一项所述的合成的指环体或如实施例3036-3038中任一项所述的药物组合物用于将所述遗传元件递送至宿主细胞的用途。

[1329] 3055.如实施例3001-3035中任一项所述的合成的指环体或如实施例3036-3038中

任一项所述的药物组合物用于治疗受试者的疾病或障碍的用途。

[1330] 3056.如实施例3055所述的用途，其中所述疾病或障碍是神经肌肉疾病(例如，脊髓性肌萎缩症)、肌营养不良症(例如，杜氏肌营养不良症或贝克尔(Becker)肌营养不良症)、酶促不足(例如，克里格勒-纳贾尔(Crigler-Najjar)综合征)、或白化病(例如眼白化病)、或C1缺乏症。

[1331] 3057.如实施例3001-3035任一项所述的合成的指环体或如实施例3036-3038任一项所述的药物组合物，用于治疗受试者的疾病或障碍。

[1332] 3058.一种治疗受试者的疾病或障碍的方法，所述方法包括向所述受试者施用如实施例3001-3035中任一项所述的合成的指环体或如实施例3036-3038中任一项所述的药物组合物，其中所述疾病或障碍是神经肌肉疾病(例如，脊髓性肌萎缩症)、肌营养不良症(例如，杜氏肌营养不良症或贝克尔(Becker)肌营养不良症)、酶促不足(例如，克里格勒-纳贾尔(Crigler-Najjar)综合征)、或白化病(例如眼白化病)、或C1缺乏症。

[1333] 3059.如实施例3001-3035中任一项所述的合成的指环体或如实施例3036-3038中任一项所述的药物组合物在制造用于治疗受试者的疾病或障碍的药物中的用途，任选地其中所述疾病或障碍是神经肌肉疾病(例如，脊髓性肌萎缩症)、肌营养不良症(例如，杜氏肌营养不良症或贝克尔(Becker)肌营养不良症)、酶促不足(例如，克里格勒-纳贾尔(Crigler-Najjar)综合征)、或白化病(例如眼白化病)、或C1缺乏症。

[1334] 根据说明书和附图并且根据权利要求书，本发明的其他特征、目标和优点将是显而易见的。

[1335] 除非另外定义，否则本文所用的所有技术和科学术语均具有与本发明所属领域的普通技术人员通常所理解的相同的含义。所有的公开物、专利申请、专利、以及在此提及的其他参考文献以其全文通过引用结合。另外，材料、方法和实施例仅是说明性的而不旨在限制。

附图说明

[1336] 当结合附图阅读时，将更好地理解以下详细描述的本发明的实施例。出于说明本发明的目的，在附图中示出了本发明示例的实施例。然而，应当理解，本发明不限于附图中所示实施例的精确安排和手段。

[1337] 本专利或申请文件包含至少一个彩色附图。应请求并且支付必要的费用后，具有一个或多个彩色附图的本专利或专利申请公开的副本将由专利局提供。

[1338] 图1A是示出衣壳蛋白序列的氨基酸区的百分比序列相似性的图示。

[1339] 图1B是示出衣壳蛋白序列的百分比序列相似性的图示。

[1340] 图2是示出指环体的一个实施例的图示。

[1341] 图3描绘了编码TTMiniV的LY1菌株(“指环体1”)的卡那霉素载体的示意图。

[1342] 图4描绘了编码TTMiniV的LY2菌株(“指环体2”)的卡那霉素载体的示意图。

[1343] 图5描述了293T和A549细胞中合成的指环体的转染效率。

[1344] 图6A和6B描述了定量PCR结果，所述结果说明了合成的指环体对293T细胞的成功感染。

[1345] 图7A和7B描述了定量PCR结果，所述结果说明了合成的指环体对A549细胞的成功

感染。

[1346] 图8A和8B描述了定量PCR结果,所述结果说明了合成的指环体对Raji细胞的成功感染。

[1347] 图9A和9B描述了定量PCR结果,所述结果说明了合成的指环体对Jurkat细胞的成功感染。

[1348] 图10A和10B描述了定量PCR结果,所述结果说明了合成的指环体对Chang细胞的成功感染。

[1349] 图11A-11B是一系列图表,其显示了用TTMV-LY2Δ574-1371、Δ1432-2210、2610::nLuc转染或感染的细胞中萤光素酶的表达。在受感染的细胞中观察到发光,这表明成功复制和包装。

[1350] 图11C是描绘了甲型细环病毒(细环病毒;TTV)的系统进化树的图,其中进化分枝突出显示。代表了至少100个指环病毒株。来自几个进化分枝的示例性序列在本文中提供,例如在表A1-A12、B1-B5、C1-C5和1-18中。

[1351] 图12是示出了用于产生指环体(例如,本文所述的具有复制能力的指环体或复制缺陷型指环体)的示例性工作流程的示意图。

[1352] 图13是示出了针对设计用于定量TTV和TTMV基因组当量的引物组的引物特异性的图。基于SYBR绿色化学的定量PCR显示在编码各个基因组的质粒上,使用TTMV或TTV特异性引物组(如所示)对每个扩增产物都有一个不同的峰。

[1353] 图14是一系列图,其显示了通过qPCR定量TTV基因组当量的PCR效率。浓度增加的引物和固定浓度的水解探针(250nM)与两种不同的商业qPCR预混合物一起使用。90%-110%的效率产生定量过程中最低的误差传播。

[1354] 图15是图,其示出了利用7个基因组当量浓度log10,线性扩增TTMV(靶标1)或TTV(靶标2)的示例性扩增图。用7个10倍稀释以高PCR效率和线性对基因组当量进行定量(R^2 TTMV:0.996;R^2 TTV:0.997)。

[1355] 图16A-16B是一系列图,其显示了对指环体储备物中TTMV基因组当量的定量。(A)两种储备物的扩增图,每种按1:10稀释并一式两份运行。(B)与图A相同的两个样品,此处在线性范围的情况下显示。显示的是两个代表性样品的上限和下限。PCR效率:99.58%,R^2:0988。

[1356] 图17是显示用指定质粒转染的HEK293T细胞中miR-625表达倍数变化的图。

[1357] 图18是示出了来自每个甲型细环病毒进化分枝的代表性序列的比对的成对同一性的图。将TTV-CT30F、TTV-P13-1、TTV-tth8、TTV-HD20a、TTV-16、TTV-TJN02和TTV-HD16d的DNA序列进行比对。沿着比对长度显示了50-bp滑动窗口中的成对百分比同一性。上方的括号指示了成对同一性的非编码和编码区,其中指示了成对同一性。下面的括号表示高或低序列保守的区域。

[1358] 图19是示出了七个甲型细环病毒进化分枝中推定蛋白质的氨基酸比对的成对同一性的图。将来自TTV-CT30F、TTV-P13-1、TTV-tth8、TTV-HD20a、TTV-16、TTV-TJN02和TTV-HD16d的推定蛋白质的氨基酸序列进行比对。沿着每个比对的长度显示了15-aa滑动窗口中的成对百分比同一性。指示了开放阅读框DNA序列和蛋白质氨基酸序列的成对同一性。(*)针对TTV-CT30F、TTV-tth8、TTV-16和TTV-TJN02,比对推定的ORF2t/3氨基酸序列。

[1359] 图20是示出了七个甲型细环病毒进化分枝中5'UTR内结构域高度保守的图(分别为SEQ ID NO 810-817,按出现顺序)。将每个代表性甲型细环病毒的71-bp 5'UTR保守结构域序列进行比对。所述序列在七个进化分枝之间具有95.2%的成对同一性。

[1360] 图21是示出了来自七个甲型细环病毒进化分枝的GC富集结构域的比对的图。每种指环病毒在ORF的下游都有GC含量大于70%的区。显示的是来自TTV-CT30F、TTV-P13-1、TTV-tth8、TTV-HD20a、TTV-16、TTV-TJN02和TTV-HD16d的GC富集区的比对。这些区的长度不同,但是在比对时,它们显示出75.4%的成对同一性。

[1361] 图22是显示用编码靶向n-myc相互作用蛋白(NMI)的miRNA的指环体感染Raji B细胞的图。显示的是在用NMI miRNA编码的指环体感染Raji B细胞(箭头)或对照细胞后检测到的指环体基因组当量的量化。

[1362] 图23是显示用编码靶向n-myc相互作用蛋白(NMI)的miRNA的指环体感染Raji B细胞的图。蛋白质印迹显示编码针对NMI的miRNA的指环体降低了Raji B细胞中NMI蛋白的表达,而用缺乏所述miRNA的指环体感染的Raji B细胞显示出与对照相当的NMI蛋白表达。

[1363] 图24是一系列图,显示了在用包含内源miRNA编码序列的指环体和相应的其中内源miRNA编码序列被缺失的指环体感染后宿主细胞中产生的指环体颗粒的量化。

[1364] 图25A-25C是一系列图,显示了来自与纳米萤光素酶融合的TTMV-LY2的ORF的细胞内定位。(A)在Vero细胞中,ORF2(顶行)似乎定位于细胞质,而ORF1/1(底行)似乎定位于细胞核。(B)在HEK293细胞中,ORF2(顶行)似乎定位于细胞质,而ORF1/1(底行)似乎定位于细胞核。(C)细胞中ORF1/2和ORF2/2的定位模式。

[1365] 图26是一系列图,显示了TTV-tth8的3'非编码区(NCR)中的顺序缺失对照。顶行显示野生型TTV-tth8指环病毒的结构。第二行显示TTV-tth8,其在3'NCR的GC富集区中缺失36个核苷酸(Δ36nt(GC))。第三行显示缺失36个核苷酸和另外缺失miRNA序列的TTV-tth8,导致总共缺失78个核苷酸(Δ36nt(GC)ΔmiR)。第四行显示从3'NCR缺失171个核苷酸的TTV-tth8,3'NCR包括36个核苷酸缺失区和miRNA序列(Δ3'NCR)。

[1366] 图27A-27D是一系列图,显示TTV-tth8的3'NCR中的顺序缺失对指环病毒ORF转录物水平具有显著影响。显示的是第2天的ORF1和ORF2(A)、第2天的ORF1/1和ORF2/2(B)、第2天的ORF1/2和ORF2/3(C)和第2天的ORF2t3(D)的表达。

[1367] 图28A-28B是一系列图,显示用于产生表达纳米萤光素酶的指环体的构建体(A)和用于转染细胞的一系列指环体/质粒组合(B)

[1368] 图29A-29C是一系列图,显示注射了指环体的小鼠中纳米萤光素酶的表达。(A)注射后第0-9天小鼠中的纳米萤光素酶表达。(B)注射了各种指环体/质粒构建体组合的小鼠中的纳米萤光素酶表达,如所示。(C)注射后在小鼠体内检测到的纳米萤光素酶发光的定量。A组接受TTMV-LY2载体±纳米萤光素酶。B组接受纳米萤光素酶蛋白和TTMV-LY2 ORF。

[1369] 图29D是来自七个不同甲型细环病毒进化分枝的代表性指环体的基因组组织的示意图。将TTV-CT30F、TTV-P13-1、TTV-tth8、TTV-HD20a、TTV-16、TTV-TJN02和TTV-HD16d的序列进行比对,其中标注了关键区域。推定的开放阅读框(ORF)以浅灰色表示,TATA盒以深灰色表示,关键的推定调控区以中灰色表示,包括起始元件、5'UTR保守结构域和GC富集区(例如,如所示)。

[1370] 图30是显示用于确定指环病毒前miRNA的内源性靶的示例性工作流程的示意图。

[1371] 图31A-31B是一系列图，显示串联的指环病毒质粒可以增加指环病毒或指环病毒的产生。(A)示例性串联指环病毒质粒的质粒图谱。(B)用串联指环病毒质粒转染HEK293T细胞，产生的病毒基因组数目是携带单拷贝的质粒的四倍。

[1372] 图31C是显示TTMV-LY2质粒pVL46-063和pVL46-240的环化的凝胶电泳图像。

[1373] 图31D是显示线性和环状TTMV-LY2构建体的拷贝数的色谱图，如通过尺寸排阻色谱法(SEC)确定的。

[1374] 图32是显示来自九个指环病毒基因组序列的36个核苷酸的GC富集区和基于其的共有序列的比对的图(分别为SEQ ID NO 818-827，按出现顺序)。

[1375] 图33是一系列图，显示了来自指环病毒株LY2和CBD203的ORF1结构。推定结构域被标记：精氨酸富集区(arg富集)，核心区域，包括胶冻卷结构域、高变区(HVR)、N22区和C末端结构域(CTD)，如所示。

[1376] 图34是显示来自乙型细环病毒株CBS203的ORF1结构的图。指出了在一组110种乙型细环病毒之间显示出高度相似性的残基。指示的是所有评估的株中60％-79.9％相似性的残基、80％-99.9％相似性的残基和100％相似性的残基。

[1377] 图35是显示以下的图：来自甲型细环病毒的258个序列的比对的共有序列(SEQ ID NO:828)，具有高相似性得分的残基突出显示深灰色(100％)、中灰色(80％-99.9％)、浅灰色(60％-80％)。推定结构域在框中指示。共有序列下方的框图也指示了同一性百分比，其中中灰色框表示100％同一性，浅灰色框表示30％-99％同一性，深灰色框表示低于30％同一性。

[1378] 图36是示意图，显示了指环病毒ORF1分子的结构域和将被来自不同指环病毒的高变结构域替换的高变区。

[1379] 图37是示意图，显示了ORF1的结构域和将被来自非指环病毒来源的目的蛋白质或肽(POI)替换的高变区。

[1380] 图38是一系列图，显示了基于指环病毒基因组的示例性指环体遗传元件的设计。从指环病毒基因组(左)中缺失了蛋白质编码区，留下了指环病毒非编码区(NCR)，包括病毒启动子、5'UTR保守结构域(5CD)和GC富集区。有效负载DNA被插入到蛋白质编码基因座的非编码区(右)。由此产生的指环体含有有效负载DNA(包括开放阅读框、基因、非编码RNA等)和必需的指环病毒顺式复制和包装元件，但缺乏用于复制和包装的必需蛋白质元件。

[1381] 图39是柱状图，显示了包含编码外源人免疫粘附素的遗传元件的指环体成功转导人肺衍生细胞系EKVX。

[1382] 图40是显示了以下的图：基于tth8或LY2被设计为包含编码人促红细胞生成素(hEpo)的序列的指环体可以将功能性转基因递送至哺乳动物细胞。

[1383] 图41A和41B是一系列图，显示了在静脉注射后七天可检测到施用给小鼠的工程化指环体。

[1384] 图42是显示了以下的图：在静脉内施用编码hGH的工程化指环体后七天在全血的细胞级分中检测到hGH mRNA。

[1385] 图43A-43D是一系列图，说明了指环病毒ORF2中高度保守的基序。图43公开了SEQ ID NO:949。

[1386] 图44A和44B是一系列图，显示了人组织中全长ORF1 mRNA表达的证据。

[1387] 图45是显示了以下的图:体外环化(IVC)的TTV-tth8基因组(IVC TTV-tth8)与质粒中的TTV-tth8基因组相比在HEK293T细胞中以预期密度产生TTV-tth8基因组拷贝的能力。

[1388] 图46是一系列图,显示了体外环化(IVC)的LY2基因组(WT LY2 IVC)和质粒(WT LY2质粒)中的野生型LY2基因组在Jurkat细胞中以预期密度产生LY2基因组拷贝的能力。

[1389] 图47是显示了以下的图:来自甲型细环病毒、乙型细环病毒和丙型细环病毒(SEQ ID NO:950-975)的指环病毒ORF1蛋白的胶冻卷结构域的二级结构的比对。这些二级结构元件是高度保守的。

[1390] 图48是显示了以下的图:位于N22结构域中的ORF1基序的保守序列和二级结构(按出现顺序分别为SEQ ID NO 976-1000和851)。人TTV ORF1的保守YNPXXDXGXXN(SEQ ID NO:829)基序具有保守的二级结构。特别是,基序中的酪氨酸破坏了β链,第二β链从基序的末端天冬酰胺开始。

具体实施例

[1391] 定义

[1392] 本发明将针对特定实施例并参考某些附图进行描述,但是本发明不限于此,而是仅由权利要求书来限定。除非另有说明,否则下文陈述的术语通常应以其共识来理解。

[1393] 在本说明书和权利要求中使用术语“包含”时,不排除其他要素。出于本发明的目的,术语“由……组成”被认为是术语“包含”的优选实施例。如果在下文中将组定义为包含至少一定数量的实施例,则这将被理解为优选地还公开了仅由这些实施例组成的组。

[1394] 当指代单数名词时使用不定冠词或定冠词,例如“一个(a)”、“一种(an)”或“所述(the)”,除非另有说明,否则这包括该名词的复数形式。

[1395] 词语“用于治疗、调节等的化合物、组合物、产物等”应理解为是指本身适合于所指示的治疗、调节等目的的化合物、组合物、产物等。词语“用于治疗、调节等的化合物、组合物、产物等”另外作为一个实施例披露了这种化合物、组合物、产物等用于治疗、调节等。

[1396] 词语“用于…的化合物、组合物、产物等”、“化合物、组合物、产物等用于制造用以…的药物、药物组合物、兽用组合物、诊断组合物等的用途”或“用于用作药物……的化合物、组合物、产品等”表示这些化合物、组合物、产物等将用于可在人体或动物身上实施的治疗方法中。它们被认为是涉及治疗方法等的实施例和权利要求的等同披露。如果实施例或权利要求因此是指“用于治疗疑似患有疾病的人或动物的化合物”,则这也被认为是披露“化合物在制造用于治疗疑似患有疾病的人或动物的药物中的用途”或“通过向疑似患有疾病的人或动物施用化合物的治疗方法”。词语“用于治疗、调节等的化合物、组合物、产物等”应理解为是指本身适合于所指示的治疗、调节等目的的化合物、组合物、产物等。

[1397] 如果在下文中在括号中提供术语、值、数量等的实例,则这将被理解为表示在括号中提及的实例可以构成实施例。例如,如果指出“在实施例中,核酸分子包含与表1的指环病毒ORF1编码核苷酸序列(例如,表1的核酸序列的核苷酸571-2613)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列”,则一些实施例涉及包含与表1的核酸序列的核苷酸571-2613具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列的核酸分子。

[1398] 如本文所用，术语“指环体”是指包含封闭在蛋白质外部中的遗传元件，例如附加体，例如环状DNA的媒介物。如本文所用，“合成的指环体”通常是指非天然存在的指环体，例如具有相对于野生型病毒(例如，如本文所述的野生型指环病毒)不同的序列。在一些实施例中，合成的指环体被工程化或重组，例如，包含遗传元件，所述遗传元件包含相对于野生型病毒基因组(例如，如本文所述的野生型指环病毒基因组)的差异或修饰。在一些实施例中，封闭在蛋白质外部中包含被蛋白质外部100％覆盖，以及小于100％覆盖，例如95％、90％、85％、80％、70％、60％、50％或更少。例如，在蛋白质外部可以存在间隙或不连续(例如，使得所述蛋白质外部可透过水、离子、肽或小分子)，只要遗传元件例如进入宿主细胞之前保留在蛋白质外部中即可，。在一些实施例中，将指环体纯化，例如，将其从其原始来源中分离和/或基本上不含(>50％、>60％、>70％、>80％、>90％)其他组分。

[1399] 如本文所用，术语“指环载体(anellovector)”是指包含足够的核酸序列(所述核酸序列来源于或高度类似于(例如，至少85％、90％、95％、96％、97％、98％、99％或100％相同)指环病毒基因组序列或其连续部分以允许包装到蛋白质外部(例如，衣壳)中)并进一步包含异源序列的载体。在一些实施例中，指环载体是病毒载体或裸核酸。在一些实施例中，指环载体包含天然指环病毒序列的至少约50、60、70、71、72、73、74、75、80、90、100、150、200、300、400、500、600、700、800、900、1000、1100、1200、1300、1400、1500、1600、1700、1800、1900、2000、2500、3000或3500个连续核苷酸与其或高度相似(至少85％、90％、95％、96％、97％、98％、99％或100％相同)的序列。在一些实施例中，指环载体进一步包含指环病毒ORF1、ORF2或ORF3中的一种或多种。在一些实施例中，异源序列包含多克隆位点、包含异源启动子、包含治疗性蛋白质的编码区或编码治疗性核酸。在一些实施例中，衣壳是野生型指环病毒衣壳。在实施例中，指环载体包含本文所述的遗传元件，例如，包含遗传元件，所述遗传元件包含启动子、编码治疗效应子的序列和衣壳结合序列。

[1400] 如本文所用，术语“抗体分子”是指包含至少一个免疫球蛋白可变结构域序列的蛋白质，例如免疫球蛋白链或其片段。术语“抗体分子”包括全长抗体和抗体片段(例如，scFv)。在一些实施例中，抗体分子是多特异性抗体分子，例如，所述抗体分子包含多个免疫球蛋白可变结构域序列，其中该多个中的第一免疫球蛋白可变结构域序列对第一表位具有结合特异性并且该多个中的第二免疫球蛋白可变结构域序列对第二表位具有结合特异性。在实施例中，多特异性抗体分子是双特异性抗体分子。双特异性抗体分子的特征通常在于对第一表位具有结合特异性的第一免疫球蛋白可变结构域序列和对第二表位具有结合特异性的第二免疫球蛋白可变结构域序列。

[1401] 如本文所用，“编码”的核酸是指编码氨基酸序列或功能性多核苷酸(例如非编码RNA，例如siRNA或miRNA)的核酸序列。

[1402] 如本文所用，“外源性”因子(例如，效应子、核酸(例如，RNA)、基因、有效负载、蛋白质)是指不被相应的野生型病毒(例如，本文所述的指环病毒)包括或编码的因子。在一些实施例中，外源性因子不是天然存在的，例如具有相对于天然存在的蛋白质或核酸(例如，通过插入、缺失或置换)发生改变的序列的蛋白质或核酸。在一些实施例中，外源性因子不天然存在于宿主细胞中。在一些实施例中，外源性因子天然存在于宿主细胞中，但是对于病毒是外源的。在一些实施例中，外源性因子天然存在于宿主细胞中，但是不以期望的水平或在期望的时间存在。

[1403] “异源”药剂或元件(例如,效应子、核酸序列、氨基酸序列),如本文关于另一种药剂或元件(例如,效应子、核酸序列、氨基酸序列)所使用,指不是天然地一起发现的药剂或元件,例如,在野生型病毒,例如指环病毒中。在一些实施例中,异源核酸序列可以与天然存在的核酸序列(例如,天然存在于指环病毒中的序列)存在于同一核酸中。在一些实施例中,异源药剂或元件相对于指环病毒是外源性的,指环体的其他(例如,其余的)元件基于该指环病毒。

[1404] 如本文所用,术语“遗传元件”是指通常在指环体中的核酸序列。应当理解,遗传元件可以作为裸DNA产生,并且任选地进一步装配到蛋白质外部中。还应理解,指环体可将其遗传元件插入细胞中,从而导致遗传元件存在于细胞中,而蛋白质外部不一定进入细胞。

[1405] 如本文所用,术语“ORF1分子”是指具有指环病毒ORF1蛋白(例如,如本文所述的指环病毒ORF1蛋白,例如,如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所列)的活性和/或结构特征的多肽,或其功能片段。在一些情况下,ORF1分子可以包含以下中的一种或多种(例如,1、2、3或4种):包含至少60%碱性残基(例如,至少60%精氨酸残基)的第一区域,包含至少约六个β链(例如,至少4、5、6、7、8、9、10、11或12个β链)的第二区域,包含指环病毒N22结构域(例如,如本文所述,例如来自如本文所述的指环病毒ORF1蛋白的N22结构域)的结构或活性的第三区域,和/或包含指环病毒C末端结构域(CTD)(例如,如本文所述,例如来自如本文所述的指环病毒ORF1蛋白的CTD)的结构或活性的第四区域。在一些情况下,ORF1分子以N末端到C末端的顺序包含所述第一、第二、第三和第四区域。在一些情况下,指环体包含ORF1分子,所述分子以N末端到C末端的顺序包含所述第一、第二、第三和第四区域。在一些情况下,ORF1分子可以包含由指环病毒ORF1核酸(例如,如表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中的任一个中所列)编码的多肽。在一些情况下,ORF1分子可进一步包含异源序列,例如高变区(HVR),例如来自例如本文所述的指环病毒ORF1蛋白的HVR。如本文所用,“指环病毒ORF1蛋白”是指由指环病毒基因组(例如,野生型指环病毒基因组,例如,如本文所述)编码的ORF1蛋白,例如具有表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所列的氨基酸序列的ORF1蛋白,或由表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中的任一个中所列的ORF1基因编码的ORF1蛋白。

[1406] 如本文所用,术语“ORF2分子”是指具有指环病毒ORF2蛋白(例如,如本文所述的指环病毒ORF2蛋白,例如,如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所列)的活性和/或结构特征的多肽,或其功能片段。如本文所用,“指环病毒ORF2蛋白”是指由指环病毒基因组(例如,野生型指环病毒基因组,例如,如本文所述)编码的ORF2蛋白,例如具有表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所列的氨基酸序列的ORF2蛋白,或由表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中的任一个中所列的ORF2基因编码的ORF2蛋白。

[1407] 如本文所用,术语“蛋白质外部”是指主要(例如>50%、>60%、>70%、>80%、>90%)是蛋白质的外部组分。

[1408] 如本文所用,术语“调节性核酸”是指修饰编码表达产物的DNA序列的表达,例如转录和/或翻译的核酸序列。在实施例中,表达产物包含RNA或蛋白质。

[1409] 如本文所用,术语“调节性序列”是指修饰靶基因产物的转录的核酸序列。在一些

实施例中，调节性序列是启动子或增强子。

[1410] 如本文所用，术语“复制蛋白”是指在感染、病毒基因组复制/表达、病毒蛋白合成和/或病毒组分装配过程中使用的蛋白，例如病毒蛋白。

[1411] 如本文所用，“基本上非致病性的”生物体、颗粒或组分是指例如在宿主生物例如哺乳动物例如人中不会引起或诱导可检测的疾病或致病性状况的生物、颗粒(例如，病毒或指环体，例如本文所述)或其组分。在一些实施例中，向受试者施用指环体可导致轻微的反应或副作用，其作为护理标准的一部分是可接受的。

[1412] 如本文所用，术语“非致病性的”是指例如在宿主生物例如哺乳动物例如人中不会引起或诱导可检测的疾病或致病性状况的生物体或其组分。

[1413] 如本文所用，“基本上非整合的”遗传元件是指一种遗传元件，例如病毒或指环体中的遗传元件，例如如本文所述，其中进入宿主细胞(例如，真核细胞)或生物体(例如，哺乳动物，例如人)的遗传元件中的少于约0.01％、0.05％、0.1％、0.5％或1％整合到基因组中。在一些实施例中，遗传元件没有可检测地整合到例如宿主细胞的基因组中。在一些实施例中，可以使用本文所述的技术，例如核酸测序、PCR检测和/或核酸杂交，检测遗传元件整合到基因组中。

[1414] 如本文所用，“基本上非免疫原性的”生物体、颗粒或组分是指例如在宿主组织或生物体(例如哺乳动物，例如人)中不会引起或诱导不期望或未靶向的免疫应答的生物体、颗粒(例如，病毒或指环体，例如本文所述)或其组分。在实施例中，基本上非免疫原性的生物体、颗粒或组分不产生可检测的免疫应答。在实施例中，基本上非免疫原性的指环体对包含表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中任一个所示的氨基酸序列或由表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中任一个所示的核酸序列编码的蛋白质不产生可检测的免疫应答。在实施例中，通过测定受试者中抗体的存在或水平(例如，抗指环体抗体的存在或水平，例如针对本文所述的指环体的抗体的存在或水平)来检测免疫应答(例如，不期望的或未靶向的免疫应答)，例如，根据Tsuda等人(1999；J.Virol.Methods[病毒学方法杂志]77:199-206；通过引用并入本文)中所述的抗TTV抗体检测方法和/或Kakkola等人(2008；Virology[病毒学]382:182-189；通过引用并入本文)中所述的测定抗TTV IgG水平的方法。针对指环病毒或基于其的指环体的抗体还可以通过本领域中检测抗病毒抗体的方法来检测，例如检测抗AAV抗体的方法，例如Calcedo等人(2013；Front.Immunol.[免疫学前沿]4(341):1-7；通过引用并入本文)。

[1415] 如本文所用的“子序列”是指分别包含在较大核酸序列或氨基酸序列中的核酸序列或氨基酸序列。在一些情况下，子序列可以包含较大序列的结构域或功能片段。在一些情况下，子序列可以包含较大序列的片段，所述片段当从所述较大序列分离时能够形成与当与所述较大序列的其余部分一起存在时形成的二级和/或三级结构相似的二级和/或三级结构。在一些情况下，子序列可以被另一序列替换(例如，包含外源序列的子序列或与较大序列的其余部分异源的序列，例如来自不同指环病毒的相应子序列)。

[1416] 如本文所用，“治疗(treatment)”，“治疗(treating)”及其同源词是指旨在改善、改进、稳定、预防或治愈疾病、病理状况或障碍的受试者的医学管理。该术语包括积极治疗(旨在改善疾病、病理状况或障碍的治疗)、因果治疗(针对相关疾病、病理状况或障碍的原因的治疗)、姑息治疗(旨在缓解症状而设计的治疗)、预防性治疗(旨在预防、最小化或部分

或完全抑制相关疾病、病理状况或障碍进展的治疗);以及支持治疗(用于补充另一治疗的治疗)。

[1417] 如本文所用,术语"病毒组"是指在特定环境中,例如身体的一部分,例如在生物体中,例如在细胞中,例如在组织中的病毒。

[1418] 本发明一般涉及指环体,例如合成的指环体,及其用途。本发明提供指环体,包含指环体的组合物,以及制备或使用指环体的方法。指环体通常用作递送媒介物,例如用于将治疗剂递送至真核细胞。通常,指环体将包括遗传元件,所述遗传元件包含封闭在蛋白质外部内的核酸序列(例如,编码效应子,例如外源性效应子或内源性效应子)。相对于指环病毒序列(例如,如本文所述),指环体可包括一个或多个序列(例如,如本文所述的区域或结构域)的缺失。指环体可以用作基本上非免疫原性的媒介物,用于将遗传元件或其中编码的效应子(例如,多肽或核酸效应子,例如如本文所述)递送到真核细胞中,例如用以治疗包含所述细胞的受试者的疾病或障碍。

[1419] 内容列表

[1420] I.指环体

[1421] A.指环病毒

[1422] B.ORF1分子

[1423] C.ORF2分子

[1424] D.遗传元件

[1425] E.蛋白质结合序列

[1426] F.5'UTR区域

[1427] G.GC富集区

[1428] H.效应子

[1429] I.蛋白质外部

[1430] II.载体

[1431] III.组合物

[1432] IV.宿主细胞

[1433] V.使用方法

[1434] VI.生产方法

[1435] VII.施用/递送

[1436] I.指环体

[1437] 在一些方面,本文所述的发明包括使用和制备指环体、指环体制剂和治疗性组合物的组合物和方法。在一些实施例中,指环体具有基于指环病毒(例如,本文所述的指环病毒,例如,包含含有表A1-A12、B1-B5、C1-C5、1-18、20-37或D1-D10中任一个中所示的序列的核酸或多肽的指环病毒)或其片段、或其他基本上非致病性的病毒(例如共生病毒(symbiotic virus)、普通病毒(commensal virus)、天然病毒)的序列、结构和/或功能。在一些实施例中,基于指环病毒的指环体包含至少一个对于指环病毒是外源的元件,例如,外源性效应子或编码安排在指环体的遗传元件内的外源性效应子的核酸序列。在一些实施例中,基于指环病毒的指环体包含至少一个与来自该指环病毒的另一元件异源的元件,例如,与另一连接的核酸序列异源的效应子编码核酸序列,例如启动子元件。在一些实施例中,指

环体包含遗传元件(例如,环状DNA,例如单链DNA),其包含至少一个相对于遗传元件的其余部分和/或蛋白质外部是异源的元件(例如,编码效应子的外源元件,例如,如本文所述)。指环体可以是用于将有效负载进入宿主例如人的递送媒介物(例如,基本上非致病性的递送媒介物)。在一些实施例中,指环体能够在真核细胞例如哺乳动物细胞例如人细胞中复制。在一些实施例中,指环体在哺乳动物(例如人)细胞中是基本上非致病性的和/或基本上非整合的。在一些实施例中,指环体在哺乳动物例如人中是基本上非免疫原性的。在一些实施例中,指环体是复制缺陷型。在一些实施例中,指环体具有复制能力。

[1438] 在一些实施例中,指环体包含指环体或其组分(例如,遗传元件,例如,包含编码效应子的序列和/或蛋白质外部),例如,如PCT申请号PCT/US2018/037379所描述,其通过引用整体并入本文。

[1439] 在一个方面,本发明包括指环体,所述指环体包含(i)遗传元件,其包含启动子元件,编码效应子的序列(例如内源性效应子或外源性效应子,例如有效负载)和蛋白质结合序列(例如外部蛋白质结合序列,例如包装信号),其中遗传元件是单链DNA,并具有以下一种或两种性质:呈环状和/或以少于进入细胞的遗传元件的约0.001%、0.005%、0.01%、0.05%、0.1%、0.5%、1%、1.5%或2%的频率整合到真核细胞的基因组中;以及(ii)蛋白质外部;其中所述遗传元件被封闭在所述蛋白质外部内;并且其中指环体能够将遗传元件递送到真核细胞中。

[1440] 在本文描述的指环体的一些实施例中,遗传元件以少于进入细胞的遗传元件的约0.001%、0.005%、0.01%、0.05%、0.1%、0.5%、1%、1.5%或2%的频率整合。在一些实施例中,施用给受试者的多个指环体的少于约0.01%、0.05%、0.1%、0.5%、1%、2%、3%、4%或5%的遗传元件将整合到受试者的一个或多个宿主细胞的基因组中。在一些实施例中,例如本文所述的指环体的群体的遗传元件以低于相当的AAV病毒群体的频率的频率整合到宿主细胞的基因组中,例如以比相当的AAV病毒群体的频率低约50%、60%、70%、75%、80%、85%、90%、95%、100%或更低的频率。

[1441] 在一个方面,本发明包括指环体,其包含:(i)遗传元件,其包含启动子元件和编码效应子(例如内源性效应子或外源性效应子,例如有效负载)的序列,以及蛋白质结合序列(例如外部蛋白质结合序列),其中遗传元件与野生型指环病毒序列(例如,野生型细环病毒(TTV)、细小环病毒(TTMV)或TTMDV序列,例如,表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中的任一个中所列的野生型指环病毒序列)具有至少75%(例如,至少75%、76%、77%、78%、79%、80%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%)序列同一性;以及(ii)蛋白质外部;其中所述遗传元件被封闭在所述蛋白质外部内;并且其中指环体能够将遗传元件递送到真核细胞中。

[1442] 在一个方面,本发明包括指环体,其包含:

[1443] a)遗传元件,其包含(i)编码外部蛋白质(例如,非致病性外部蛋白质)的序列,(ii)使所述遗传元件与所述非致病性外部蛋白质结合的外部蛋白质结合序列,和(iii)编码效应子(例如内源性效应子或外源性效应子)的序列;以及

[1444] b)与所述遗传元件相关联(例如包封或封闭所述遗传元件)的蛋白质外部。

[1445] 在一些实施例中,所述指环体包括来自非包膜、环状、单链DNA病毒的序列或表达产物(或与其具有>70%、75%、80%、85%、90%、95%、97%、98%、99%、100%同源性)。动

物环状单链DNA病毒通常是指感染真核非植物宿主并具有环状基因组的单链DNA(ssDNA)病毒亚组。因此,动物环状ssDNA病毒可区别于感染原核生物的ssDNA病毒(即微病毒科和丝状噬菌体科)和感染植物的ssDNA病毒(即双生病毒科和矮化病毒科)。它们也可区别于感染非植物真核生物的线性ssDNA病毒(即细小病毒科)。

[1446] 在一些实施例中,所述指环体例如瞬时或长期地调节宿主细胞功能。在某些实施例中,细胞功能发生稳定地改变,例如调节持续存在至少约1小时至约30天,或至少约2小时、6小时、12小时、18小时、24小时、2天、3天、4天、5天、6天、7天、8天、9天、10天、11天、12天、13天、14天、15天、16天、17天、18天、19天、20天、21天、22天、23天、24天、25天、26天、27天、28天、29天、30天、60天或更长时间或其间的任何时间。在某些实施例中,细胞功能发生瞬时地改变,例如调节持续存在不超过约30分钟至约7天、或不超过约1小时、2小时,3小时、4小时、5小时、6小时、7小时、8小时、9小时、10小时、11小时、12小时、13小时、14小时、15小时、16小时、17小时、18小时、19小时、20小时、21小时、22小时、24小时、36小时、48小时、60小时、72小时、4天、5天、6天、7天或其间的任何时间。

[1447] 在一些实施例中,遗传元件包括启动子元件。在实施例中,启动子元件选自RNA聚合酶II依赖性启动子、RNA聚合酶III依赖性启动子、PGK启动子、CMV启动子、EF-1α启动子、SV40启动子、CAGG启动子或UBC启动子、TTV病毒启动子、组织特异性U6(pollIII)、具有激活蛋白(TetR-VP16、Gal4-VP16、dCas9-VP16等)的上游DNA结合位点的最小CMV启动子。在实施例中,启动子元件包含TATA盒。在实施例中,启动子元件对于野生型指环病毒(例如本文所述)是内源性的。

[1448] 在一些实施例中,遗传元件包括以下一种或多种特征:单链、环状、负链和/或DNA。在实施例中,遗传元件包括附加体。在一些实施例中,除效应子之外的遗传元件部分的组合大小为约2.5kb-5kb(例如,约2.8kb-4kb、约2.8kb-3.2kb、约3.6kb-3.9kb或约2.8kb-2.9kb)、小于约5kb(例如,小于约2.9kb、3.2kb、3.6kb、3.9kb或4kb)或至少100个核苷酸(例如,至少1kb)。

[1449] 在一些情况下,本文所述的指环体、包含指环体的组合物、使用所述指环体的方法等部分基于实例,这些实例说明了不同的效应子(例如miRNA(例如针对IFN或miR-625)、shRNA等)和蛋白质结合序列(例如与衣壳蛋白例如Q99153结合的DNA序列)如何与蛋白质外部(例如Arch Virol[病毒学文献](2007)152:1961-1975中披露的衣壳)组合,产生可随后用于将效应子递送至细胞(例如动物细胞,例如人细胞或非人动物细胞,例如猪或小鼠细胞)的指环体。在实施例中,效应子可以沉默诸如干扰素的因子的表达。实例进一步描述了如何通过将效应子插入例如衍生自指环病毒的序列中来制备指环体。基于这些实例,下文中的描述考虑了实例中所考虑的具体发现和组合的各种变型。例如,技术人员将从实例中理解,特定miRNA仅用作效应子的实例,而其他效应子可以是例如其他调节性核酸或治疗性肽。类似地,实例中使用的特定衣壳可以替换为下文所述的基本上非致病性的蛋白质。在实例中描述的特定指环病毒序列也可以替换为下文描述的指环病毒序列。这些考虑类似地适用于蛋白质结合序列、调节性序列例如启动子等。独立于此,本领域技术人员将特别考虑与实例紧密相关的这些实施例。

[1450] 在一些实施例中,将指环体或指环体中包含的遗传元件引入细胞(例如人细胞)中。在一些实施例中,例如由指环体的遗传元件编码的效应子(例如,RNA,例如miRNA)在细

胞(例如,人细胞)中表达,例如,一旦将指环体或遗传元件引入细胞中。在实施例中,将指环体或其中包含的遗传元件引入细胞中例如通过改变细胞中靶分子的表达水平,来调节(例如,增加或减少)靶分子(例如,靶核酸,例如RNA,或靶多肽)在细胞中的水平。在实施例中,引入指环体或其中包含的遗传元件降低了细胞产生的干扰素的水平。在实施例中,将指环体或其中包含的遗传元件引入细胞中调节(例如,增加或减少)细胞的功能。在实施例中,将指环体或其中包含的遗传元件引入细胞中调节(例如,增加或降低)细胞的存活力。在实施例中,将指环体或其中包含的遗传元件引入细胞中降低了细胞(例如癌细胞)的存活力。

[1451] 在一些实施例中,本文描述的指环体(例如,合成的指环体)诱导小于70%的抗体流行率(例如小于约60%、50%、40%、30%、20%或10%抗体流行率)。在实施例中,根据本领域已知的方法测定抗体流行率。在实施例中,通过检测生物样品中的针对指环病毒(例如本文所述)或基于其的指环体的抗体来测定抗体流行率,例如,根据Tsuda等人(1999;J.Virol.Methods[病毒学方法杂志]77:199-206;通过引用并入本文)中所述的抗TTV抗体检测方法和/或Kakkola等人(2008;Virology[病毒学]382:182-189;通过引用并入本文)中所述的测定抗TTV IgG血清阳性率的方法。针对指环病毒或基于其的指环体的抗体还可以通过本领域中检测抗病毒抗体的方法来检测,例如检测抗AAV抗体的方法,例如Calcedo等人(2013;Front.Immunol.[免疫学前沿]4(341):1-7;通过引用并入本文)。

[1452] 在一些实施例中,复制不足、复制缺陷或复制无能力遗传元件不编码遗传元件复制所需的所有必要机器或组分。在一些实施例中,复制缺陷遗传元件不编码复制因子。在一些实施例中,复制缺陷遗传元件不编码一个或多个ORF(例如,ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3和/或ORF2t/3,例如,如本文所述)。在一些实施例中,可以以反式提供不由遗传元件编码的机器或组分(例如,使用辅助病毒,例如辅助病毒或辅助质粒,或在由宿主细胞包含的核酸中编码,例如整合到宿主细胞的基因组中),例如,使得遗传元件可以在以反式提供的机器或组分的存在下进行复制。

[1453] 在一些实施例中,包装不足、包装缺陷或包装无能的遗传元件不能被包装到蛋白质外部(例如,其中蛋白质外部包含衣壳或其部分,例如包含由ORF1核酸编码的多肽,例如,如本文所述)中。在一些实施例中,与野生型指环病毒(例如,如本文所述)相比,包装缺陷遗传元件以小于10%(例如,小于10%、9%、8%、7%、6%、5%、4%、3%、2%、1%、0.5%、0.1%、0.01%或0.001%)的效率包装进入蛋白质外部中。在一些实施例中,即使在允许包装野生型指环病毒的遗传元件(例如,如本文所述)的因子(例如,ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3、或ORF2t/3)存在下,包装缺陷遗传元件也不能被包装到蛋白质外部中。在一些实施例中,即使在允许包装野生型指环病毒的遗传元件(例如,如本文所述)的因子(例如,ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3或ORF2t/3)存在下,与野生型指环病毒(例如,如本文所述)相比,包装缺陷遗传元件以小于10%(例如,小于10%、9%、8%、7%、6%、5%、4%、3%、2%、1%、0.5%、0.1%、0.01%或0.001%)的效率包装进入蛋白质外部中。

[1454] 在一些实施例中,包装感受态遗传元件可被包装到蛋白质外部(例如,其中蛋白质外部包含衣壳或其部分,例如包含由ORF1核酸编码的多肽,例如,如本文所述)中。在一些实施例中,与野生型指环病毒(例如,如本文所述)相比,包装感受态遗传元件以至少20%(例如,至少20%、30%、40%、50%、60%、70%、80%、85%、90%、95%、96%、97%、98%、99%、

100％或更高)的效率包装进入蛋白质外部中。在一些实施例中，在允许包装野生型指环病毒的遗传元件(例如，如本文所述)的因子(例如，ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3、或ORF2t/3)存在下，包装感受态遗传元件能被包装到蛋白质外部中。在一些实施例中，在允许包装野生型指环病毒的遗传元件(例如，如本文所述)的因子(例如，ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3或ORF2t/3)存在下，与野生型指环病毒(例如，如本文所述)相比，包装感受态遗传元件以至少20％(例如，至少20％、30％、40％、50％、60％、70％、80％、85％、90％、95％、96％、97％、98％、99％、100％或更高)的效率包装进入蛋白质外部中。

[1455] 指环病毒

[1456] 在一些实施例中，例如本文所述的指环体包括衍生自指环病毒的序列或表达产物。在一些实施例中，指环体包括相对于指环病毒是外源的一种或多种序列或表达产物。在一些实施例中，指环体包括相对于指环病毒是内源的一种或多种序列或表达产物。在一些实施例中，指环体包括相对于指环体中的一种或多种其他序列或表达产物是异源的一种或多种序列或表达产物。指环病毒通常具有负极性的单链环状DNA基因组。指环病毒通常与任何人疾病无关联。但是，在一个或多个对照队列群体中无症状的指环病毒病毒血症发生率高，指环病毒病毒科内显著的基因组多样性，病史上无法在体外传播所述因子以及缺乏指环病毒疾病的一种或多种动物模型妨碍了将指环病毒感染与人疾病相关联的尝试(Yzebe等人，Panminerva Med.(2002)44:167-177;Biagini,P.,Vet.Microbiol.[兽医微生物学](2004)98:95-101)。

[1457] 指环病毒通常是通过口鼻或粪口感染、母婴传播和/或子宫内传播而传播的(Gerner等人，Ped.Infect.Dis.J.[小儿传染病杂志](2000)19:1074-1077)。在一些情况下，感染者的特征可以是延长(数月至数年)的指环病毒属病毒血症。人可以共同感染一种以上的基因群或菌株(Saback，等人，Scad.J.Infect.Dis.[斯堪的纳维亚感染病杂志](2001)33:121-125)。有建议认为这些基因群可以在感染者内重组(Rey等人，Infect.[感染](2003)31:226-233)。在诸如肝脏、外周血单核细胞和骨髓的几种组织中发现了双链同工型(复制性)中间体(Kikuchi等人，J.Med.Virol.[医学病毒学杂志](2000)61:165-170;Okamoto等人，Biochem.Biophys.Res.Commun.[生物化学和生物物理研究通讯](2002)270:657-662;Rodriguez-lnigo等人，Am.J.Pathol.[美国病理学杂志](2000)156:1227-1234)。

[1458] 在一些实施例中，遗传元件包含编码以下的核苷酸序列：氨基酸序列或其功能片段或与本文所述的氨基酸序列中任一个(例如，指环病毒氨基酸序列)具有至少约60％、70％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的序列。

[1459] 在一些实施例中，本文所述的指环体包含一个或多个核酸分子(例如本文所述的遗传元件)，其包含与例如本文所述的指环病毒序列或其片段具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的序列。在实施例中，指环体包含选自以下的核酸序列：表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中的任一个中所示的序列，或与其具有至少70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的序列。在实施例中，指环体包含多肽，所述多肽包含如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中任一个中所示的序列，或与其具有至少70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的序列。

[1460] 在一些实施例中，本文所述的指环体包含一个或多个核酸分子(例如，如本文所述的遗传元件)，其包含与本文所述的任一指环病毒(例如，表A1-A12、B1-B5、C1-C5或1-18中任一个中注释或所列序列编码的指环病毒序列)的TATA盒、帽位点、起始元件、转录起始位点、5'UTR保守结构域、ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3、ORF2t/3、三开放阅读框区、聚(A)信号、GC富集区、或其任何组合中的一个或多个具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的序列。在一些实施例中，核酸分子包含编码衣壳蛋白的序列，例如本文所述的任一指环病毒(例如，表A1-A12或1-18中任一个中注释或所列序列编码的指环病毒序列)的ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3、ORF2t/3序列。在实施例中，核酸分子包含编码衣壳蛋白的序列，所述衣壳蛋白包含与指环病毒ORF1或ORF2蛋白(例如，表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所示的ORF1或ORF2氨基酸序列，或表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中的任一个中所示的核酸序列编码的ORF1或ORF2氨基酸序列)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，核酸分子包含编码衣壳蛋白的序列，所述衣壳蛋白包含与指环病毒ORF1蛋白(例如，表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所示的ORF1氨基酸序列，或表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中的任一个中所示的核酸序列编码的ORF1氨基酸序列)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。

[1461] 在实施例中，核酸分子包含与表A1的指环病毒ORF1核苷酸序列(例如，表A1的核酸序列的核苷酸574-2775)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A1的指环病毒ORF1/1核苷酸序列(例如表A1的核酸序列的核苷酸574-699和/或2326-2775)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A1的指环病毒ORF1/2核苷酸序列(例如表A1的核酸序列的核苷酸574-699和/或2552-2759)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A1的指环病毒ORF2核苷酸序列(例如，表A1的核酸序列的核苷酸335-703)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A1的指环病毒ORF2/2核苷酸序列(例如表A1的核酸序列的核苷酸335-699和/或2326-2759)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A1的指环病毒ORF2/3核苷酸序列(例如表A1的核酸序列的核苷酸335-699和/或2552-2957)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A1的指环病毒ORF2t/3核苷酸序列(例如表A1的核酸序列的核苷酸335-465和/或2552-2957)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A1的指环病毒TATA盒核苷酸序列(例如，表A1的核酸序列的核苷酸77-81)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A1的指环病毒起始元件核苷酸序列(例如，表A1的核酸序列的核苷酸95-110)具有

至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A1的指环病毒转录起始位点核苷酸序列(例如，表A1的核酸序列的核苷酸105)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A1的指环病毒5'UTR保守结构域核苷酸序列(例如，表A1的核酸序列的核苷酸165-235)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A1的指环病毒三开放阅读框区核苷酸序列(例如，表A1的核酸序列的核苷酸2535-2746)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A1的指环病毒聚(A)信号核苷酸序列(例如，表A1的核酸序列的核苷酸2953-2958)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A1的指环病毒GC富集核苷酸序列(例如，表A1的核酸序列的核苷酸3620-3648)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。

[1462] 在实施例中，核酸分子包含与表A3的指环病毒ORF1核苷酸序列(例如，表A3的核酸序列的核苷酸599-2887)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A3的指环病毒ORF1/1核苷酸序列(例如，表A3的核酸序列的核苷酸599-724和/或2414-2887)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A3的指环病毒ORF1/2核苷酸序列(例如，表A3的核酸序列的核苷酸599-724和/或2643-2849)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A3的指环病毒ORF2核苷酸序列(例如，表A3的核酸序列的核苷酸342-728)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A3的指环病毒ORF2/2核苷酸序列(例如，表A3的核酸序列的核苷酸342-724和/或2414-2849)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A3的指环病毒ORF2/3核苷酸序列(例如，表A3的核酸序列的核苷酸342-724和/或2643-3057)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A3的指环病毒TATA盒核苷酸序列(例如，表A3的核酸序列的核苷酸87-91)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A3的指环病毒起始元件核苷酸序列(例如，表A3的核酸序列的核苷酸105-120)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A3的指环病毒转录起始位点核苷酸序列(例如，表A3的核酸序列的核苷酸115)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A3的指环病毒5'UTR保守结构域核苷酸序列(例如，表A3的核酸序列的核苷酸175-245)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A3的指环病毒

三开放阅读框区核苷酸序列(例如,表A3的核酸序列的核苷酸2626-2846)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表A3的指环病毒聚(A)信号核苷酸序列(例如,表A3的核酸序列的核苷酸3052-3058)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。

[1463] 在实施例中,核酸分子包含与表A5的指环病毒ORF1核苷酸序列(例如,表A5的核酸序列的核苷酸556-2904)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表A5的指环病毒ORF1/1核苷酸序列(例如,表A5的核酸序列的核苷酸556-687和/或2422-2904)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表A5的指环病毒ORF1/2核苷酸序列(例如,表A5的核酸序列的核苷酸556-687和/或2564-2878)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表A5的指环病毒ORF2核苷酸序列(例如,表A5的核酸序列的核苷酸305-691)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表A5的指环病毒ORF2/2核苷酸序列(例如,表A5的核酸序列的核苷酸305-687和/或2422-2878)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表A5的指环病毒ORF2/3核苷酸序列(例如,表A5的核酸序列的核苷酸305-687和/或2564-3317)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表A5的指环病毒ORF2t/3核苷酸序列(例如,表A5的核酸序列的核苷酸305-360和/或2564-3317)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表A5的指环病毒TATA盒核苷酸序列(例如,表A5的核酸序列的核苷酸50-55)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表A5的指环病毒起始元件核苷酸序列(例如,表A5的核酸序列的核苷酸68-83)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表A5的指环病毒转录起始位点核苷酸序列(例如,表A5的核酸序列的核苷酸78)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表A5的指环病毒5'UTR保守结构域核苷酸序列(例如,表A5的核酸序列的核苷酸138-208)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表A5的指环病毒三开放阅读框区核苷酸序列(例如,表A5的核酸序列的核苷酸2626-2846)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表A5的指环病毒聚(A)信号核苷酸序列(例如,表A5的核酸序列的核苷酸3316-3319)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。

[1464] 在实施例中,核酸分子包含与表A7的指环病毒ORF1核苷酸序列(例如,表A7的核酸序列的核苷酸589-2889)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、

99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A7的指环病毒ORF1/1核苷酸序列(例如，表A7的核酸序列的核苷酸589-711和/或2362-2889)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A7的指环病毒ORF1/2核苷酸序列(例如，表A7的核酸序列的核苷酸589-711和/或2555-2863)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A7的指环病毒ORF2核苷酸序列(例如，表A7的核酸序列的核苷酸353-715)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A7的指环病毒ORF2/2核苷酸序列(例如，表A7的核酸序列的核苷酸353-711和/或2362-2863)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A7的指环病毒ORF2/3核苷酸序列(例如，表A7的核酸序列的核苷酸353-711和/或2555-3065)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A7的指环病毒ORF2t/3核苷酸序列(例如，表A7的核酸序列的核苷酸353-432和/或2555-3065)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A7的指环病毒TATA盒核苷酸序列(例如，表A7的核酸序列的核苷酸86-90)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A7的指环病毒起始元件核苷酸序列(例如，表A7的核酸序列的核苷酸104-119)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A7的指环病毒转录起始位点核苷酸序列(例如，表A7的核酸序列的核苷酸114)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A7的指环病毒5'UTR保守结构域核苷酸序列(例如，表A7的核酸序列的核苷酸174-244)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A7的指环病毒三开放阅读框区核苷酸序列(例如，表A7的核酸序列的核苷酸2555-2863)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A7的指环病毒聚(A)信号核苷酸序列(例如，表A7的核酸序列的核苷酸3062-3066)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A7的指环病毒GC富集核苷酸序列(例如，表A7的核酸序列的核苷酸3720-3742)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。

[1465] 在实施例中，核酸分子包含与表A9的指环病毒ORF1核苷酸序列(例如，表A9的核酸序列的核苷酸511-2793)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A9的指环病毒ORF1/1核苷酸序列(例如，表A9的核酸序列的核苷酸511-711和/或2326-2793)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A9的指环病毒ORF1/2核苷酸序列(例如，表A9的核酸序列的核苷

酸511-711和/或2525-2767)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A9的指环病毒ORF2核苷酸序列(例如，表A9的核酸序列的核苷酸272-637)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A9的指环病毒ORF2/2核苷酸序列(例如，表A9的核酸序列的核苷酸272-633和/或2326-2767)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A9的指环病毒ORF2/3核苷酸序列(例如，表A9的核酸序列的核苷酸272-633和/或2525-2984)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A9的指环病毒ORF2t/3核苷酸序列(例如，表A9的核酸序列的核苷酸272-633和/或2525-2984)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A9的指环病毒TATA盒核苷酸序列(例如，表A9的核酸序列的核苷酸12-17)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A9的指环病毒起始元件核苷酸序列(例如，表A9的核酸序列的核苷酸30-45)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A9的指环病毒转录起始位点核苷酸序列(例如，表A9的核酸序列的核苷酸40)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A9的指环病毒5'UTR保守结构域核苷酸序列(例如表A9的核酸序列的核苷酸100-171)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A9的指环病毒三开放阅读框区核苷酸序列(例如，表A9的核酸序列的核苷酸2525-2767)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A9的指环病毒聚(A)信号核苷酸序列(例如，表A9的核酸序列的核苷酸2981-2985)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。

[1466] 在实施例中，核酸分子包含与表A11的指环病毒ORF1核苷酸序列(例如，表A11的核酸序列的核苷酸704-3001)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A11的指环病毒ORF1/1核苷酸序列(例如，表A11的核酸序列的核苷酸704-826和/或2534-3001)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A11的指环病毒ORF1/2核苷酸序列(例如，表A11的核酸序列的核苷酸704-826和/或2721-2975)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A11的指环病毒ORF2核苷酸序列(例如，表A11的核酸序列的核苷酸465-830)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A11的指环病毒ORF2/2核苷酸序列(例如，表A11的核酸序列的核苷酸465-826和/或2534-2975)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表A11的指环病毒ORF2/

3核苷酸序列(例如,表A11的核酸序列的核苷酸465-826和/或2721-3192)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表A11的指环病毒ORF2t/3核苷酸序列(例如,表A11的核酸序列的核苷酸465-595和/或2721-3192)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表A11的指环病毒TATA盒核苷酸序列(例如,表A11的核酸序列的核苷酸206-210)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表A11的指环病毒起始元件核苷酸序列(例如,表A11的核酸序列的核苷酸224-239)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表A11的指环病毒转录起始位点核苷酸序列(例如,表A11的核酸序列的核苷酸234)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表A11的指环病毒5'UTR保守结构域核苷酸序列(例如,表A11的核酸序列的核苷酸294-364)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表A11的指环病毒三开放阅读框区核苷酸序列(例如,表A11的核酸序列的核苷酸2721-2975)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表A11的指环病毒聚(A)信号核苷酸序列(例如,表A11的核酸序列的核苷酸3189-3193)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表A11的指环病毒GC富集核苷酸序列(例如,表A11的核酸序列的核苷酸3844-3895)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。

[1467] 在实施例中,核酸分子包含与表B1的指环病毒ORF1核苷酸序列(例如,表B1的核酸序列的核苷酸574-2775)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表B1的指环病毒ORF1/1核苷酸序列(例如表B1的核酸序列的核苷酸574-699和/或2326-2775)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表B1的指环病毒ORF1/2核苷酸序列(例如表B1的核酸序列的核苷酸574-699和/或2552-2759)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表B1的指环病毒ORF2核苷酸序列(例如,表B1的核酸序列的核苷酸335-703)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表B1的指环病毒ORF2/2核苷酸序列(例如表B1的核酸序列的核苷酸335-699和/或2326-2759)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表B1的指环病毒ORF2/3核苷酸序列(例如表B1的核酸序列的核苷酸335-699和/或2552-2957)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表B1的指环病毒ORF2t/3核苷酸序列(例如表B1的核酸序列的核苷酸335-465和/或2552-2957)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、

99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B1的指环病毒TATA盒核苷酸序列(例如，表B1的核酸序列的核苷酸77-81)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B1的指环病毒起始元件核苷酸序列(例如，表B1的核酸序列的核苷酸95-110)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B1的指环病毒转录起始位点核苷酸序列(例如，表B1的核酸序列的核苷酸105)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B1的指环病毒5'UTR保守结构域核苷酸序列(例如，表B1的核酸序列的核苷酸165-235)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B1的指环病毒三开放阅读框区核苷酸序列(例如，表B1的核酸序列的核苷酸2535-2746)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B1的指环病毒聚(A)信号核苷酸序列(例如，表B1的核酸序列的核苷酸2953-2958)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B1的指环病毒GC富集核苷酸序列(例如，表B1的核酸序列的核苷酸3620-3648)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。

[1468] 在实施例中，核酸分子包含与表B2的指环病毒ORF1核苷酸序列(例如，表B2的核酸序列的核苷酸574-2775)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B2的指环病毒ORF1/1核苷酸序列(例如表B2的核酸序列的核苷酸574-699和/或2326-2775)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B2的指环病毒ORF1/2核苷酸序列(例如表B2的核酸序列的核苷酸574-699和/或2552-2759)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B2的指环病毒ORF2核苷酸序列(例如，表B2的核酸序列的核苷酸335-703)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B2的指环病毒ORF2/2核苷酸序列(例如表B2的核酸序列的核苷酸335-699和/或2326-2759)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B2的指环病毒ORF2/3核苷酸序列(例如表B2的核酸序列的核苷酸335-699和/或2552-2957)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B2的指环病毒ORF2t/3核苷酸序列(例如表B2的核酸序列的核苷酸335-465和/或2552-2957)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B2的指环病毒TATA盒核苷酸序列(例如，表B2的核酸序列的核苷酸77-81)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B2的指环病毒起始元件核苷酸序列(例如，表B2的核酸序列的核苷酸95-110)具有

至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B2的指环病毒转录起始位点核苷酸序列(例如，表B2的核酸序列的核苷酸105)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B2的指环病毒5'UTR保守结构域核苷酸序列(例如，表B2的核酸序列的核苷酸165-235)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B2的指环病毒三开放阅读框区核苷酸序列(例如，表B2的核酸序列的核苷酸2535-2746)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B2的指环病毒聚(A)信号核苷酸序列(例如，表B2的核酸序列的核苷酸2953-2958)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B2的指环病毒GC富集核苷酸序列(例如，表B2的核酸序列的核苷酸3620-3648)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。

[1469] 在实施例中，核酸分子包含与表B3的指环病毒ORF1核苷酸序列(例如，表B3的核酸序列的核苷酸574-2775)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B3的指环病毒ORF1/1核苷酸序列(例如表B3的核酸序列的核苷酸574-699和/或2326-2775)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B3的指环病毒ORF1/2核苷酸序列(例如表B3的核酸序列的核苷酸574-699和/或2552-2759)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B3的指环病毒ORF2核苷酸序列(例如，表B3的核酸序列的核苷酸335-703)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B3的指环病毒ORF2/2核苷酸序列(例如表B3的核酸序列的核苷酸335-699和/或2326-2759)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B3的指环病毒ORF2/3核苷酸序列(例如表B3的核酸序列的核苷酸335-699和/或2552-2957)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B3的指环病毒ORF2t/3核苷酸序列(例如表B3的核酸序列的核苷酸335-465和/或2552-2957)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B3的指环病毒TATA盒核苷酸序列(例如，表B3的核酸序列的核苷酸77-81)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B3的指环病毒起始元件核苷酸序列(例如，表B3的核酸序列的核苷酸95-110)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B3的指环病毒转录起始位点核苷酸序列(例如，表B3的核酸序列的核苷酸105)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B3的指环病毒

5'UTR保守结构域核苷酸序列(例如,表B3的核酸序列的核苷酸165-235)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表B3的指环病毒三开放阅读框区核苷酸序列(例如,表B3的核酸序列的核苷酸2535-2746)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表B3的指环病毒聚(A)信号核苷酸序列(例如,表B3的核酸序列的核苷酸2953-2958)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表B3的指环病毒GC富集核苷酸序列(例如,表B3的核酸序列的核苷酸3620-3648)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。

[1470] 在实施例中,核酸分子包含与表B4的指环病毒ORF1核苷酸序列(例如,表B4的核酸序列的核苷酸574-2775)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表B4的指环病毒ORF1/1核苷酸序列(例如表B4的核酸序列的核苷酸574-699和/或2326-2775)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表B4的指环病毒ORF1/2核苷酸序列(例如表B4的核酸序列的核苷酸574-699和/或2552-2759)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表B4的指环病毒ORF2核苷酸序列(例如,表B4的核酸序列的核苷酸335-703)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表B4的指环病毒ORF2/2核苷酸序列(例如表B4的核酸序列的核苷酸335-699和/或2326-2759)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表B4的指环病毒ORF2/3核苷酸序列(例如表B4的核酸序列的核苷酸335-699和/或2552-2957)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表B4的指环病毒ORF2t/3核苷酸序列(例如表B4的核酸序列的核苷酸335-465和/或2552-2957)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表B4的指环病毒TATA盒核苷酸序列(例如,表B4的核酸序列的核苷酸77-81)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表B4的指环病毒起始元件核苷酸序列(例如,表B4的核酸序列的核苷酸95-110)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表B4的指环病毒转录起始位点核苷酸序列(例如,表B4的核酸序列的核苷酸105)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表B4的指环病毒5'UTR保守结构域核苷酸序列(例如,表B4的核酸序列的核苷酸165-235)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表B4的指环病毒三开放阅读框区核苷酸序列(例如,表B4的核酸序列的核苷酸2535-2746)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、

99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B4的指环病毒聚(A)信号核苷酸序列(例如，表B4的核酸序列的核苷酸2953-2958)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B4的指环病毒GC富集核苷酸序列(例如，表B4的核酸序列的核苷酸3620-3648)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。

[1471] 在实施例中，核酸分子包含与表B5的指环病毒ORF1核苷酸序列(例如，表B5的核酸序列的核苷酸574-2775)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B5的指环病毒ORF1/1核苷酸序列(例如表B5的核酸序列的核苷酸574-699和/或2326-2775)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B5的指环病毒ORF1/2核苷酸序列(例如表B5的核酸序列的核苷酸574-699和/或2552-2759)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B5的指环病毒ORF2核苷酸序列(例如，表B5的核酸序列的核苷酸335-703)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B5的指环病毒ORF2/2核苷酸序列(例如表B5的核酸序列的核苷酸335-699和/或2326-2759)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B5的指环病毒ORF2/3核苷酸序列(例如表B5的核酸序列的核苷酸335-699和/或2552-2957)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B5的指环病毒ORF2t/3核苷酸序列(例如表B5的核酸序列的核苷酸335-465和/或2552-2957)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B5的指环病毒TATA盒核苷酸序列(例如，表B5的核酸序列的核苷酸77-81)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B5的指环病毒起始元件核苷酸序列(例如，表B5的核酸序列的核苷酸95-110)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B5的指环病毒转录起始位点核苷酸序列(例如，表B5的核酸序列的核苷酸105)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B5的指环病毒5'UTR保守结构域核苷酸序列(例如，表B5的核酸序列的核苷酸165-235)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B5的指环病毒三开放阅读框区核苷酸序列(例如，表B5的核酸序列的核苷酸2535-2746)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B5的指环病毒聚(A)信号核苷酸序列(例如，表B5的核酸序列的核苷酸2953-2958)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表B5的指环病毒GC富集核苷酸序列(例如，表B5的核酸序列的核苷酸3620-

3648)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。

[1472] 在实施例中，核酸分子包含与表1的指环病毒ORF1核苷酸序列(例如，表1的核酸序列的核苷酸571-2613)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中，核酸分子包含与表1的指环病毒ORF1/1核苷酸序列(例如，表1的核酸序列的核苷酸571-587和/或2137-2613)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中，核酸分子包含与表1的指环病毒ORF1/2核苷酸序列(例如，表1的核酸序列的核苷酸571-687和/或2339-2659)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中，核酸分子包含与表1的指环病毒ORF2核苷酸序列(例如，表1的核酸序列的核苷酸299-691)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中，核酸分子包含与表1的指环病毒ORF2/2核苷酸序列(例如，表1的核酸序列的核苷酸299-687和/或2137-2659)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中，核酸分子包含与表1的指环病毒ORF2/3核苷酸序列(例如，表1的核酸序列的核苷酸299-687和/或2339-2831)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中，核酸分子包含与表1的指环病毒ORF2t/3核苷酸序列(例如，表1的核酸序列的核苷酸299-348和/或2339-2831)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中，核酸分子包含与表1的指环病毒TATA盒核苷酸序列(例如，表1的核酸序列的核苷酸84-90)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中，核酸分子包含与表1的指环病毒帽位点核苷酸序列(例如，表1的核酸序列的核苷酸107-114)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中，核酸分子包含与表1的指环病毒转录起始位点核苷酸序列(例如，表1的核酸序列的核苷酸114)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中，核酸分子包含与表1的指环病毒5'UTR保守结构域核苷酸序列(例如，表1的核酸序列的核苷酸177-247)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中，核酸分子包含与表1的指环病毒三开放阅读框区核苷酸序列(例如，表1的核酸序列的核苷酸2325-2610)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中，核酸分子包含与表1的指环病毒聚(A)信号核苷酸序列(例如，表1的核酸序列的核苷酸2813-2818)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中，核酸分子包含与表1的指环病毒GC富集核苷酸序列(例如，表1的核酸序列的核苷酸3415-3570)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。

[1473] 在实施例中，核酸分子包含与表3的指环病毒ORF1核苷酸序列(例如，表3的核酸序列的核苷酸729-2972)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中，核酸分子包含与表3的指环病毒ORF1/1

核苷酸序列(例如,表3的核酸序列的核苷酸729-908和/或2490-2972)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表3的指环病毒ORF1/2核苷酸序列(例如,表3的核酸序列的核苷酸729-908和/或2725-3039)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表3的指环病毒ORF2核苷酸序列(例如,表3的核酸序列的核苷酸412-912)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表3的指环病毒ORF2/2核苷酸序列(例如,表3的核酸序列的核苷酸412-908和/或2490-3039)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表3的指环病毒ORF2/3核苷酸序列(例如,表3的核酸序列的核苷酸412-908和/或2725-3208)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表3的指环病毒TATA盒核苷酸序列(例如,表3的核酸序列的核苷酸112-119)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表3的指环病毒起始元件核苷酸序列(例如,表3的核酸序列的核苷酸128-148)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表3的指环病毒转录起始位点核苷酸序列(例如,表3的核酸序列的核苷酸148)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表3的指环病毒5'UTR保守结构域核苷酸序列(例如,表3的核酸序列的核苷酸204-273)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表3的指环病毒三开放阅读框区核苷酸序列(例如,表3的核酸序列的核苷酸2699-2969)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表3的指环病毒聚(A)信号核苷酸序列(例如,表3的核酸序列的核苷酸3220-3225)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表3的指环病毒GC富集核苷酸序列(例如,表3的核酸序列的核苷酸3302-3541)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。

[1474] 在实施例中,核酸分子包含与表5的指环病毒ORF1核苷酸序列(例如,表5的核酸序列的核苷酸599-2830)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表5的指环病毒ORF1/1核苷酸序列(例如,表5的核酸序列的核苷酸599-715和/或2363-2830)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表5的指环病毒ORF1/2核苷酸序列(例如,表5的核酸序列的核苷酸599-715和/或2565-2789)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表5的指环病毒ORF2核苷酸序列(例如,表5的核酸序列的核苷酸336-719)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子

包含与表5的指环病毒ORF2/2核苷酸序列(例如,表5的核酸序列的核苷酸336-715和/或2363-2789)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表5的指环病毒ORF2/3核苷酸序列(例如,表5的核酸序列的核苷酸336-715和/或2565-3015)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表5的指环病毒ORF2t/3核苷酸序列(例如,表5的核酸序列的核苷酸336-388和/或2565-3015)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表5的指环病毒TATA盒核苷酸序列(例如,表5的核酸序列的核苷酸83-88)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表5的指环病毒帽位点核苷酸序列(例如,表5的核酸序列的核苷酸104-111)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表5的指环病毒转录起始位点核苷酸序列(例如,表5的核酸序列的核苷酸111)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表5的指环病毒5'UTR保守结构域核苷酸序列(例如,表5的核酸序列的核苷酸170-240)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表5的指环病毒三开放阅读框区域核苷酸序列(例如,表5的核酸序列的核苷酸2551-2786)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表5的指环病毒聚(A)信号核苷酸序列(例如,表5的核酸序列的核苷酸3011-3016)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表5的指环病毒GC富集核苷酸序列(例如,表5的核酸序列的核苷酸3632-3753)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。

[1475] 在实施例中,核酸分子包含与表7的指环病毒ORF1核苷酸序列(例如,表7的核酸序列的核苷酸586-2928)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表7的指环病毒ORF1/1核苷酸序列(例如,表7的核酸序列的核苷酸586-717和/或2446-2928)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表7的指环病毒ORF1/2核苷酸序列(例如,表7的核酸序列的核苷酸586-717和/或2675-2902)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表7的指环病毒ORF2核苷酸序列(例如,表7的核酸序列的核苷酸335-721)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表7的指环病毒ORF2/2核苷酸序列(例如,表7的核酸序列的核苷酸335-717和/或2446-2902)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表7的指环病毒ORF2/3核苷酸序列(例如,表7的核酸序列的核苷酸335-717和/或2675-3109)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸

分子包含与表7的指环病毒TATA盒核苷酸序列(例如,表7的核酸序列的核苷酸82-87)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表7的指环病毒起始元件核苷酸序列(例如,表7的核酸序列的核苷酸95-115)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表7的指环病毒转录起始位点核苷酸序列(例如,表7的核酸序列的核苷酸115)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表7的指环病毒5'UTR保守结构域核苷酸序列(例如,表7的核酸序列的核苷酸170-238)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表7的指环病毒三开放阅读框区域核苷酸序列(例如,表7的核酸序列的核苷酸2640-2899)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表7的指环病毒聚(A)信号核苷酸序列(例如,表7的核酸序列的核苷酸3106-3114)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表7的指环病毒GC富集核苷酸序列(例如,表7的核酸序列的核苷酸3768-3878)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。

[1476] 在实施例中,核酸分子包含与表9的指环病毒ORF1核苷酸序列(例如,表9的核酸序列的核苷酸588-2873)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表9的指环病毒ORF1/1核苷酸序列(例如,表9的核酸序列的核苷酸588-722和/或2412-2873)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表9的指环病毒ORF1/2核苷酸序列(例如,表9的核酸序列的核苷酸588-722和/或2638-2847)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表9的指环病毒ORF2核苷酸序列(例如,表9的核酸序列的核苷酸331-726)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表9的指环病毒ORF2/2核苷酸序列(例如,表9的核酸序列的核苷酸331-722和/或2412-2847)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表9的指环病毒ORF2/3核苷酸序列(例如,表9的核酸序列的核苷酸331-722和/或2638-3058)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表9的指环病毒ORF2t/3核苷酸序列(例如,表9的核酸序列的核苷酸331-380和/或2638-3058)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表9的指环病毒TATA盒核苷酸序列(例如,表9的核酸序列的核苷酸82-86)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,核酸分子包含与表9的指环病毒起始元件核苷酸序列(例如,表9的核酸序列的核苷酸100-115)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实

施例中，核酸分子包含与表9的指环病毒转录起始位点核苷酸序列(例如，表9的核酸序列的核苷酸115)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表9的指环病毒5'UTR保守结构域核苷酸序列(例如，表9的核酸序列的核苷酸170-240)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表9的指环病毒三开放阅读框区核苷酸序列(例如，表9的核酸序列的核苷酸2699-2969)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表9的指环病毒聚(A)信号核苷酸序列(例如，表9的核酸序列的核苷酸3220-3225)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表9的指环病毒GC富集核苷酸序列(例如，表9的核酸序列的核苷酸3302-3541)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。

[1477] 在实施例中，核酸分子包含与表11的指环病毒ORF1核苷酸序列(例如，表11的核酸序列的核苷酸599-2839)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表11的指环病毒ORF1/1核苷酸序列(例如，表11的核酸序列的核苷酸599-727和/或2381-2839)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表11的指环病毒ORF1/2核苷酸序列(例如，表11的核酸序列的核苷酸599-727和/或2619-2813)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表11的指环病毒ORF2核苷酸序列(例如，表11的核酸序列的核苷酸357-731)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表11的指环病毒ORF2/2核苷酸序列(例如，表11的核酸序列的核苷酸357-727和/或2381-2813)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表11的指环病毒ORF2/3核苷酸序列(例如，表11的核酸序列的核苷酸357-727和/或2619-3021)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表11的指环病毒ORF2t/3核苷酸序列(例如，表11的核酸序列的核苷酸357-406和/或2619-3021)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表11的指环病毒TATA盒核苷酸序列(例如，表11的核酸序列的核苷酸89-90)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表11的指环病毒帽位点核苷酸序列(例如，表11的核酸序列的核苷酸107-114)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表11的指环病毒转录起始位点核苷酸序列(例如，表11的核酸序列的核苷酸114)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表11的指环病毒5'UTR保守结构域核苷酸序列(例如，表11的核酸序列的核苷酸174-244)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实

施例中，核酸分子包含与表11的指环病毒三开放阅读框区核苷酸序列(例如，表11的核酸序列的核苷酸2596-2810)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表11的指环病毒聚(A)信号核苷酸序列(例如，表11的核酸序列的核苷酸3017-3022)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表11的指环病毒GC富集核苷酸序列(例如，表11的核酸序列的核苷酸3691-3794)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。

[1478] 在实施例中，核酸分子包含与表13的指环病毒ORF1核苷酸序列(例如，表13的核酸序列的核苷酸599-2896)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表13的指环病毒ORF1/1核苷酸序列(例如，表13的核酸序列的核苷酸599-724和/或2411-2896)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表13的指环病毒ORF1/2核苷酸序列(例如，表13的核酸序列的核苷酸599-724和/或2646-2870)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表13的指环病毒ORF2核苷酸序列(例如，表13的核酸序列的核苷酸357-728)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表13的指环病毒ORF2/2核苷酸序列(例如，表13的核酸序列的核苷酸357-724和/或2411-2870)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表13的指环病毒ORF2/3核苷酸序列(例如，表13的核酸序列的核苷酸357-724和/或2646-3081)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表13的指环病毒TATA盒核苷酸序列(例如，表13的核酸序列的核苷酸82-86)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表13的指环病毒起始元件核苷酸序列(例如，表13的核酸序列的核苷酸94-115)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表13的指环病毒转录起始位点核苷酸序列(例如，表13的核酸序列的核苷酸115)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表13的指环病毒5'UTR保守结构域核苷酸序列(例如，表13的核酸序列的核苷酸170-240)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表13的指环病毒三开放阅读框区核苷酸序列(例如，表13的核酸序列的核苷酸2629-2867)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表13的指环病毒聚(A)信号核苷酸序列(例如，表13的核酸序列的核苷酸3076-3086)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表13的指环病毒GC富集核苷酸序列(例如，表13的核酸序列的核苷酸3759-3866)具有至少约70％、75％、80％、85％、90％、

95％、96％、97％、98％、99％或100％序列同一性的核酸序列。

[1479] 在实施例中，核酸分子包含与表15的指环病毒ORF1核苷酸序列(例如，表15的核酸序列的核苷酸612-2612)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表15的指环病毒ORF1/1核苷酸序列(例如，表15的核酸序列的核苷酸612-719和/或2274-2612)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表15的指环病毒ORF1/2核苷酸序列(例如，表15的核酸序列的核苷酸612-719和/或2449-2589)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表15的指环病毒ORF2核苷酸序列(例如，表15的核酸序列的核苷酸424-723)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表15的指环病毒ORF2/2核苷酸序列(例如，表15的核酸序列的核苷酸424-719和/或2274-2589)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表15的指环病毒ORF2/3核苷酸序列(例如，表15的核酸序列的核苷酸424-719和/或2449-2812)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表15的指环病毒TATA盒核苷酸序列(例如，表15的核酸序列的核苷酸237-243)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表15的指环病毒帽位点核苷酸序列(例如，表15的核酸序列的核苷酸260-267)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表15的指环病毒转录起始位点核苷酸序列(例如，表15的核酸序列的核苷酸267)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表15的指环病毒5'UTR保守结构域核苷酸序列(例如，表15的核酸序列的核苷酸323-393)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表15的指环病毒三开放阅读框区核苷酸序列(例如，表15的核酸序列的核苷酸2441-2586)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表15的指环病毒聚(A)信号核苷酸序列(例如，表15的核酸序列的核苷酸2808-2813)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表15的指环病毒GC富集核苷酸序列(例如，表15的核酸序列的核苷酸2868-2929)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。

[1480] 在实施例中，核酸分子包含与表17的指环病毒ORF1核苷酸序列(例如，表17的核酸序列的核苷酸432-2453)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表17的指环病毒ORF1/1核苷酸序列(例如，表17的核酸序列的核苷酸432-584和/或1977-2453)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表17的指环病毒ORF1/2核苷酸序列(例如，表17的核酸序列的核苷

酸432-584和/或2197-2388)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表17的指环病毒ORF2核苷酸序列(例如，表17的核酸序列的核苷酸283-588)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表17的指环病毒ORF2/2核苷酸序列(例如，表17的核酸序列的核苷酸283-584和/或1977-2388)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表17的指环病毒ORF2/3核苷酸序列(例如，表17的核酸序列的核苷酸283-584和/或2197-2614)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表17的指环病毒TATA盒核苷酸序列(例如，表17的核酸序列的核苷酸21-25)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表17的指环病毒帽位点核苷酸序列(例如，表17的核酸序列的核苷酸42-49)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表17的指环病毒转录起始位点核苷酸序列(例如，表17的核酸序列的核苷酸49)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表17的指环病毒5'UTR保守结构域核苷酸序列(例如，表17的核酸序列的核苷酸117-187)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表17的指环病毒三开放阅读框区核苷酸序列(例如，表17的核酸序列的核苷酸2186-2385)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表17的指环病毒聚(A)信号核苷酸序列(例如，表17的核酸序列的核苷酸2676-2681)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，核酸分子包含与表17的指环病毒GC富集核苷酸序列(例如，表17的核酸序列的核苷酸3054-3172)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。

[1481] 在实施例中，核酸分子包含编码与表A2的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A2的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A2的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A2的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A2的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A2的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A2的指环

病毒ORF2t/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。

[1482] 在实施例中，核酸分子包含编码与表A4的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A4的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A4的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A4的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A4的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A4的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。

[1483] 在实施例中，核酸分子包含编码与表A6的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A6的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A6的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A6的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A6的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A6的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A6的指环病毒ORF2t/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。

[1484] 在实施例中，核酸分子包含编码与表A8的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A8的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A8的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A8的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A8的指环病毒ORF2/

2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A8的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A8的指环病毒ORF2t/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。

[1485] 在实施例中，核酸分子包含编码与表A10的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A10的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A10的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A10的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A10的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A10的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A10的指环病毒ORF2t/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。

[1486] 在实施例中，核酸分子包含编码与表A12的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A12的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A12的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A12的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A12的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A12的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表A12的指环病毒ORF2t/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。

[1487] 在实施例中，核酸分子包含编码与表C1的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表C1的指环病毒ORF1/1氨基酸序列具有

至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表C1的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表C1的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表C1的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表C1的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表C1的指环病毒TAIP氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。

[1488] 在实施例中，核酸分子包含编码与表C2的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表C2的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表C2的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表C2的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表C2的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表C2的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表C2的指环病毒TAIP氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。

[1489] 在实施例中，核酸分子包含编码与表C3的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表C3的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表C3的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表C3的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表C3的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表C3的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或

100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表C3的指环病毒TAIP氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。

[1490] 在实施例中，核酸分子包含编码与表C4的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表C4的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表C4的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表C4的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表C4的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表C4的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表C4的指环病毒TAIP氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。

[1491] 在实施例中，核酸分子包含编码与表C5的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表C5的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表C5的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表C5的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表C5的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表C5的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表C5的指环病毒TAIP氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。

[1492] 在实施例中，核酸分子包含编码与表2的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表2的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表2的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的

氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表2的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表2的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表2的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表2的指环病毒ORF2t/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。

[1493] 在实施例中，核酸分子包含编码与表4的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表4的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表4的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表4的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表4的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表4的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。

[1494] 在实施例中，核酸分子包含编码与表6的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表6的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表6的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表6的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表6的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表6的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表6的指环病毒ORF2t/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。

[1495] 在实施例中，核酸分子包含编码与表8的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序

列的核酸序列。在实施例中，核酸分子包含编码与表8的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表8的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表8的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表8的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表8的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。

[1496] 在实施例中，核酸分子包含编码与表10的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表10的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表10的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表10的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表10的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表10的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表10的指环病毒ORF2t/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。

[1497] 在实施例中，核酸分子包含编码与表12的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表12的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表12的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表12的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表12的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表12的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表12的指环

病毒ORF2t/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。

[1498] 在实施例中，核酸分子包含编码与表14的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表14的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表14的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表14的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表14的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表14的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。

[1499] 在实施例中，核酸分子包含编码与表16的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表16的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表16的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表16的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表16的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表16的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。

[1500] 在实施例中，核酸分子包含编码与表18的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表18的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表18的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表18的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表18的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。在实施例中，核酸分子包含编码与表18的指环病毒

ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的核酸序列。

[1501] 在实施例中，本文所述的指环体包含具有与表A2的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A2的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A2的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A2的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A2的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A2的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A2的指环病毒ORF2t/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在一些实施例中，ORF1分子(例如，包含在指环体中)包含由表A1的核酸序列的核苷酸574-2775的指环病毒ORF1核酸序列编码的多肽。在一些实施例中，ORF1分子(例如，包含在指环体中)包含表A2的指环病毒ORF1蛋白或其剪接变体或翻译后加工(例如，蛋白水解加工)变体。

[1502] 在实施例中，本文所述的指环体包含具有与表A4的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A4的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A4的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A4的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A4的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A4的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在一些实施例中，ORF1分子(例如，包含在指环体中)包含由表A3的核酸序列的核苷酸599-2887的指环病毒ORF1核酸序列编码的多肽。在一些实施例中，ORF1分子(例如，包含在指环体中)包含表A4的指环病毒ORF1蛋白或其剪接变体或翻译后加工(例如，蛋白水解加工)变体。

[1503] 在实施例中，本文所述的指环体包含具有与表A6的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨

基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A6的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A6的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A6的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A6的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A6的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A6的指环病毒ORF2t/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在一些实施例中，ORF1分子(例如，包含在指环体中)包含由表A5的核酸序列的核苷酸556-2904的指环病毒ORF1核酸序列编码的多肽。在一些实施例中，ORF1分子(例如，包含在指环体中)包含表A6的指环病毒ORF1蛋白或其剪接变体或翻译后加工(例如，蛋白水解加工)变体。

[1504] 在实施例中，本文所述的指环体包含具有与表A8的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A8的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A8的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A8的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A8的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A8的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A8的指环病毒ORF2t/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在一些实施例中，ORF1分子(例如，包含在指环体中)包含由表A7的核酸序列的核苷酸589-2889的指环病毒ORF1核酸序列编码的多肽。在一些实施例中，ORF1分子(例如，包含在指环体中)包含表A8的指环病毒ORF1蛋白或其剪接变体或翻译后加工(例如，蛋白水解加工)变体。

[1505] 在实施例中，本文所述的指环体包含具有与表A10的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A10的指环病毒ORF1/1

氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A10的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A10的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A10的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A10的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A10的指环病毒ORF2t/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在一些实施例中，ORF1分子(例如，包含在指环体中)包含由表A9的核酸序列的核苷酸511-2793的指环病毒ORF1核酸序列编码的多肽。在一些实施例中，ORF1分子(例如，包含在指环体中)包含表A10的指环病毒ORF1蛋白或其剪接变体或翻译后加工(例如，蛋白水解加工)变体。

[1506] 在实施例中，本文所述的指环体包含具有与表A12的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A12的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A12的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A12的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A12的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A12的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表A12的指环病毒ORF2t/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在一些实施例中，ORF1分子(例如，包含在指环体中)包含由表A11的核酸序列的核苷酸704-3001的指环病毒ORF1核酸序列编码的多肽。在一些实施例中，ORF1分子(例如，包含在指环体中)包含表A12的指环病毒ORF1蛋白或其剪接变体或翻译后加工(例如，蛋白水解加工)变体。

[1507] 在实施例中，本文所述的指环体包含具有与表C1的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表C1的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序

列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表C1的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表C1的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表C1的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表C1的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表C1的指环病毒ORF2t/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表C1的指环病毒TAIP氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在一些实施例中，ORF1分子(例如，包含在指环体中)包含由表B1的核酸序列的核苷酸512-2545的指环病毒ORF1核酸序列编码的多肽。在一些实施例中，ORF1分子(例如，包含在指环体中)包含表C1的指环病毒ORF1蛋白或其剪接变体或翻译后加工(例如，蛋白水解加工)变体。

[1508] 在实施例中，本文所述的指环体包含具有与表C2的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表C2的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表C2的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表C2的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表C2的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表C2的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表C2的指环病毒ORF2t/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表C2的指环病毒TAIP氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在一些实施例中，ORF1分子(例如，包含在指环体中)包含由表B2的核酸序列的核苷酸501-2489的指环病毒ORF1核酸序列编码的多肽。在一些实施例中，ORF1分子(例如，包含在指环体中)包含表C2的指环病毒ORF1蛋白或其剪接变体或翻译后加工(例如，蛋白水解加工)变体。

[1509] 在实施例中，本文所述的指环体包含具有与表C3的指环病毒ORF1氨基酸序列具有

至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表C3的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表C3的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表C3的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表C3的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表C3的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表C3的指环病毒ORF2t/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表C3的指环病毒TAIP氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在一些实施例中，ORF1分子(例如，包含在指环体中)包含由表B3的核酸序列的核苷酸572-2758的指环病毒ORF1核酸序列编码的多肽。在一些实施例中，ORF1分子(例如，包含在指环体中)包含表C3的指环病毒ORF1蛋白或其剪接变体或翻译后加工(例如，蛋白水解加工)变体。

[1510] 在实施例中，本文所述的指环体包含具有与表C4的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表C4的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表C4的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表C4的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表C4的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表C4的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表C4的指环病毒ORF2t/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表C4的指环病毒TAIP氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在一些实施例中，ORF1分子(例如，包含在指环体中)包含由表B4的核酸序列的核苷酸581-2884的指环病毒ORF1核酸序列编码的多肽。在一些实施例中，ORF1

分子(例如,包含在指环体中)包含表C4的指环病毒ORF1蛋白或其剪接变体或翻译后加工(例如,蛋白水解加工)变体。

[1511] 在实施例中,本文所述的指环体包含具有与表C5的指环病毒ORF1氨基酸序列具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列的蛋白质。在实施例中,本文所述的指环体包含具有与表C5的指环病毒ORF1/1氨基酸序列具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列的蛋白质。在实施例中,本文所述的指环体包含具有与表C5的指环病毒ORF1/2氨基酸序列具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列的蛋白质。在实施例中,本文所述的指环体包含具有与表C5的指环病毒ORF2氨基酸序列具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列的蛋白质。在实施例中,本文所述的指环体包含具有与表C5的指环病毒ORF2/2氨基酸序列具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列的蛋白质。在实施例中,本文所述的指环体包含具有与表C5的指环病毒ORF2/3氨基酸序列具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列的蛋白质。在实施例中,本文所述的指环体包含具有与表C5的指环病毒ORF2t/3氨基酸序列具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列的蛋白质。在实施例中,本文所述的指环体包含具有与表C5的指环病毒TAIP氨基酸序列具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列的蛋白质。在一些实施例中,ORF1分子(例如,包含在指环体中)包含由表B5的核酸序列的核苷酸614-2911的指环病毒ORF1核酸序列编码的多肽。在一些实施例中,ORF1分子(例如,包含在指环体中)包含表C5的指环病毒ORF1蛋白或其剪接变体或翻译后加工(例如,蛋白水解加工)变体。

[1512] 在实施例中,本文所述的指环体包含具有与表2的指环病毒ORF1氨基酸序列具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列的蛋白质。在实施例中,本文所述的指环体包含具有与表2的指环病毒ORF1/1氨基酸序列具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列的蛋白质。在实施例中,本文所述的指环体包含具有与表2的指环病毒ORF1/2氨基酸序列具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列的蛋白质。在实施例中,本文所述的指环体包含具有与表2的指环病毒ORF2氨基酸序列具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列的蛋白质。在实施例中,本文所述的指环体包含具有与表2的指环病毒ORF2/2氨基酸序列具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列的蛋白质。在实施例中,本文所述的指环体包含具有与表2的指环病毒ORF2/3氨基酸序列具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列的蛋白质。在实施例中,本文所述的指环体包含具有与表2的指环病毒ORF2t/3氨基酸序列具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列的蛋白质。在一些实施例中,ORF1分子(例如,包含在指环体中)包含由表1的核酸序列的核苷酸

571-2613的指环病毒ORF1核酸序列编码的多肽。在一些实施例中，ORF1分子(例如，包含在指环体中)包含表2的指环病毒ORF1蛋白或其剪接变体或翻译后加工(例如，蛋白水解加工)变体。

[1513] 在实施例中，本文所述的指环体包含具有与表4的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表4的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表4的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表4的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表4的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表4的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在一些实施例中，ORF1分子(例如，包含在指环体中)包含由表3的核酸序列的核苷酸729-2972的指环病毒ORF1核酸序列编码的多肽。在一些实施例中，ORF1分子(例如，包含在指环体中)包含表4的指环病毒ORF1蛋白或其剪接变体或翻译后加工(例如，蛋白水解加工)变体。

[1514] 在实施例中，本文所述的指环体包含具有与表6的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表6的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表6的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表6的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表6的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表6的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表6的指环病毒ORF2t/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在一些实施例中，ORF1分子(例如，包含在指环体中)包含由表5的核酸序列的核苷酸599-2830的指环病毒ORF1核酸序列编码的多肽。在一些实施例中，ORF1分子(例如，包含在指环体中)包含表6的指环病毒ORF1蛋白或其剪接变体或翻译后加工(例如，蛋白水解加工)变体。

[1515] 在实施例中，本文所述的指环体包含具有与表8的指环病毒ORF1氨基酸序列具有

至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表8的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表8的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表8的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表8的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表8的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在一些实施例中，ORF1分子(例如，包含在指环体中)包含由表7的核酸序列的核苷酸586-2928的指环病毒ORF1核酸序列编码的多肽。在一些实施例中，ORF1分子(例如，包含在指环体中)包含表8的指环病毒ORF1蛋白或其剪接变体或翻译后加工(例如，蛋白水解加工)变体。

[1516] 在实施例中，本文所述的指环体包含具有与表10的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表10的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表10的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表10的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表10的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表10的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表10的指环病毒ORF2t/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在一些实施例中，ORF1分子(例如，包含在指环体中)包含由表9的核酸序列的核苷酸588-2873的指环病毒ORF1核酸序列编码的多肽。在一些实施例中，ORF1分子(例如，包含在指环体中)包含表10的指环病毒ORF1蛋白或其剪接变体或翻译后加工(例如，蛋白水解加工)变体。

[1517] 在实施例中，本文所述的指环体包含具有与表12的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表12的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表12的指环

病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表12的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表12的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表12的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表12的指环病毒ORF2t/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在一些实施例中，ORF1分子(例如，包含在指环体中)包含由表11的核酸序列的核苷酸599-2839的指环病毒ORF1核酸序列编码的多肽。在一些实施例中，ORF1分子(例如，包含在指环体中)包含表12的指环病毒ORF1蛋白或其剪接变体或翻译后加工(例如，蛋白水解加工)变体。

[1518] 在实施例中，本文所述的指环体包含具有与表14的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表14的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表14的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表14的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表14的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表14的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在一些实施例中，ORF1分子(例如，包含在指环体中)包含由表13的核酸序列的核苷酸599-2896的指环病毒ORF1核酸序列编码的多肽。在一些实施例中，ORF1分子(例如，包含在指环体中)包含表14的指环病毒ORF1蛋白或其剪接变体或翻译后加工(例如，蛋白水解加工)变体。

[1519] 在实施例中，本文所述的指环体包含具有与表16的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表16的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表16的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表16的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环

体包含具有与表16的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表16的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在一些实施例中，ORF1分子(例如，包含在指环体中)包含由表15的核酸序列的核苷酸612-2612的指环病毒ORF1核酸序列编码的多肽。在一些实施例中，ORF1分子(例如，包含在指环体中)包含表16的指环病毒ORF1蛋白或其剪接变体或翻译后加工(例如，蛋白水解加工)变体。

[1520] 在实施例中，本文所述的指环体包含具有与表18的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表18的指环病毒ORF1/1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表18的指环病毒ORF1/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表18的指环病毒ORF2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表18的指环病毒ORF2/2氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在实施例中，本文所述的指环体包含具有与表18的指环病毒ORF2/3氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列的蛋白质。在一些实施例中，ORF1分子(例如，包含在指环体中)包含由表17的核酸序列的核苷酸432-2453的指环病毒ORF1核酸序列编码的多肽。在一些实施例中，ORF1分子(例如，包含在指环体中)包含表18的指环病毒ORF1蛋白或其剪接变体或翻译后加工(例如，蛋白水解加工)变体。

[1521] 在一些实施例中，本文所述的多肽包含与本文所述的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与表2的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与表4的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与表6的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与表8的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与表10的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与表12的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与表14的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多

肽包含与表16的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与表18的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与表A2的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与表A4的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与表A6的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与表A8的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与表A10的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与表A12的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与表C1的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与表C2的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与表C3的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与表C4的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与表C5的指环病毒ORF1氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。

[1522] 在一些实施例中，本文所述的多肽包含与由本文所述的指环病毒ORF1核酸编码的ORF1分子具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与由表1中所列的指环病毒ORF1核酸编码的ORF1分子具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与由表3中所列的指环病毒ORF1核酸编码的ORF1分子具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与由表5中所列的指环病毒ORF1核酸编码的ORF1分子具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与由表7中所列的指环病毒ORF1核酸编码的ORF1分子具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与由表9中所列的指环病毒ORF1核酸编码的ORF1分子具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与由表11中所列的指环病毒ORF1核酸编码的ORF1分子具有

至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与由表13中所列的指环病毒ORF1核酸编码的ORF1分子具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与由表15中所列的指环病毒ORF1核酸编码的ORF1分子具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与由表17中所列的指环病毒ORF1核酸编码的ORF1分子具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与由表A1中所列的指环病毒ORF1核酸编码的ORF1分子具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与由表A3中所列的指环病毒ORF1核酸编码的ORF1分子具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与由表A5中所列的指环病毒ORF1核酸编码的ORF1分子具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与由表A7中所列的指环病毒ORF1核酸编码的ORF1分子具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与由表A9中所列的指环病毒ORF1核酸编码的ORF1分子具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与由表A11中所列的指环病毒ORF1核酸编码的ORF1分子具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与由表B1中所列的指环病毒ORF1核酸编码的ORF1分子具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与由表B2中所列的指环病毒ORF1核酸编码的ORF1分子具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与由表B3中所列的指环病毒ORF1核酸编码的ORF1分子具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与由表B4中所列的指环病毒ORF1核酸编码的ORF1分子具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，本文所述的多肽包含与由表B5中所列的指环病毒ORF1核酸编码的ORF1分子具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。

[1523] 在一些实施例中，多肽包含如表2、4、6、8、10、12、14、16或18中任一个中所示的氨基酸序列(例如，ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3、ORF2t/3序列)，或与其具有至少70％、75％、80％、85％、90％、95％、96％、97％、98％、99的序列％或100％序列同一性的序列。

[1524] 表A1.新颖指环病毒核酸序列(甲型细环病毒)

[1525] 名称 TTV-RTx1

[1526] 属/进化分枝 甲型细环病毒，进化分枝6

[1527] 登录号 SRR2167793

[1528] 全序列:3648bp

1 10 20 30 40 50

CGTCACTAACCACGTGACTCCCACAGGCCAACCACAGTGTACGTGATTCA
CTTCCTGGGAGTGGTTTACATTATAATATAAGCAACTGCACTTCCGAATG
GCTGAGTTTTCCACGCCCGTCCGCAGCGAGAACACCACGGAGGGGAGTCC
GCGCGTCCCGTGGGCGGGTGCCGAAGGTGAGTTTACACACCGCAGTCAAG
GGGCAATTCGGGCACGGGACTGGCCGGGCTATGGGCAAGGCTCTTAAAAA
GCTATGTTTCTTGGTAGGCCGTACCGAAAGAAAAGGAAACTGCTACTGCT
ACCACTGCATTCTACACCGAAAACTAGCCGGGTTATGAGCTGGTCTAGGC
CTGTACATAATGCCACAGGCATTGAAAGAAACTGGTGGGAGTCCTGTCTT
AGATCCCACGCAAGTTCTTGTGGCTGCGGTAATTTTGTTAATCATATTAA
TGTACTGGCTAATCGGTATGGCTTTGCTGGTTCCACGGAGACGCCGGGTA
ATCCTCGGCCGAGGCCCCCGGTACTGAGCTCCACCACCAGCACTCCTACC
GATCAATCCAGACCAGCTCTACCATGGCATGGGGATACTGGTGGAGAAGG
CGCTTCTGGAGACCCCGCAGGAGATGGAGAACGTGGCGCCGCAGAAGGAG
ACTACGGCCCAGAAGATCTAGACGCACTTTTCGACGCACTCGACGAAGAG
TAAGGAGGCGACGGTGGGGGAGGCGTGCACGCAGGCGGGGATGGCGACGC
AGGACTTATATTAGAGCCAGGCGACGCAGGAGACGAAAAAGACTTGTACT
GACTCAGTGGCATCCCGCAGTTAGAAGAAAATGTAAAATTACAGGCTACA
TGCCTATAGTATACTGTGGACATGGCAGAGCTAGTTTTAACTATGCCTGG
CACTCTGATGACTGTATAAAACAACCACTACCCTTTGGAGGCTCACTATC
TACAGTGTCCTTCAACCTAAAAGTACTATTTGACGAAAACCAAAGAGGAC
TAAACAAATGGAGCTACCCAAATGACCAACTAGACCTCGCCAGATACAAA
GGCTGTAGACTAACATTTTACAGAAAAAAAAACACAGACTACATAGCTCA
ATATGACATATCAGAACCTTATCAACTAGACAAATATAGCTGTGCAAACT
ATCACCCCTCAAAAATGATGTTTGCAAAAAACAAAATTTTAATTCCTAGC

[1529] TATGATACAAAACCTAGAGGCAGACAAAGAGTTAGAGTTAGAATAGGGCC
CCCTAAACTATTTACAGACAAGTGGTACAGTCAATCAGACTTATGCAAGG
TAAACCTTGTGTCACTTGCGGTTTCTGCGGCTTCCTTTCTCCACCCATTC
GGCTCACCACAAACTGCCAACTTTTGTGCAACCTTCCAGGTGCTGCAACC
GTTCTACTACCAGGCTATAGGCATTAGTTCTACAAAACACTCAGAAGTTA
TAGACATTTTATATAAGAAAAATACATACTGGCAAAGCAACATTACCTCT
TGGTTTTTAACTAATGTTAAAAACCCAAAAAATATGTCCACAAAAATGTT
TGAGGACATTAATGTTAAATCAAACAAAGACAGTAATTATGACTGGTTTC
CATTTACCCCATACACTACAGAAAACTATTCAAAAATTCAAAATGCAGCT
CAAGAATACTGGAAATATTTAACTAGTGACCACCCACAAGCTACTAATAG
CAATGAAGGCCTAGTACAACCATGGACTAATGCCACTATAAAACAATATG
AATACCACCTCGGTATGTTTAGTCCTATATTTATAGGACCTACCAGAGCT
AAAACTAAATTTAAAACAGCATACTTTGACTGCACTTATAACCCACTACT
AGACAAAGGAATGGGAAACAGAATATGGTATCAATACGCAACCAAAGCTG
ACACACAAATATCAAAAACAGGGTGCTACTGCATGTTAGAAGACATTCCA
ATATATGCAGCATTTTATGGATACGTAGACTTTATAGAAATGGAAATAGG
TAAAGGACAAGACATTAAAGAGAACGGACTTATTTGCTGCATATGTAGAT
ACACAGACCCCCCAATGTACAATGAACAACATCCAGACATGGGATTTGTA
TTTTATAACACTAACTTTGGAAATGGAAAATGGATAGATGGACGGGGCGA
CATACCTACTTACTGGATGCAAAGATGGAGACCTGTTGTATTATTTCAAA
CTGATGTTATTAGAGACTTAGTAGAAACTGGACCTTTTAGTTACAAAGAT
GACCTAGCAAATACCTCACTGACTATGAAATATGAATTCTATTTTACCTG
GGGCGGAAACCAGGCGTACCACCAGACAATCAAAAACCCTTGTAAAGACG
AAGGTACCGGACCCCATAGACAGCCTAGAGACGTACAAGTTACGGACCCG
ACAACCGTGGGACCTGAATATGTGTTCCACGCGTGGGACTGGAGACGGGG
CTTCCTTAGCGAGCGAGCTCTCAGACGCATGTTCGAAAAACCTCTCAACT
ATGATGAGTATTCTAAAAAACCAAAAAGACCTAGAATATTTCCTCCAACA

GAAACAGAGTCCCGAAACCAAGAGCTCGAAGAAAGCTCGCTTTCAGAGGA
AGAAAAGTCGCTACTCTCCACAGAAGAGATCCAGAAAGAGGAGATACAGC
GACAGTTCAAGCGACAGCTCAAGCGACAGCTGCGCCTCGGGCAGCAGCTC
AAACTCCTCCAACAACAACTCCTCAAGACGCAAGCGGGCCTGCACCTAAA
CCCCCTTTCATATTTCCCGCAATAAATAAAGTGTACCTGTTCCCAGACAG
AGCTCCAAAACCTAAACCCACCTCTGGAGACTGGGAAACAGAGTATGCAG
CTTGCAGTGCCTTTGACAGACCCGCTAGAACCAACCTTAGCTCACCCCCT
TACTACCCAGGAGTACCTACTCCCTGGCAAGTAAAATTCAGCCTTAAATT
TCAATAAAGTGCATTTTTACTACAGCTGGGCCGTGGGAGTTTCACTTGTC
GGTGTCTACCTCTTAAGGTCACTAAGCACTCCGAGCGCAGCGAGGAGTGC
GACCCTTAACCCTGGGTCAACGCCTTCGGAGCCGCGCGCTACGCCTTCGG
[1530] CTGCGCGCGGCACCTCAGACCCCCGCTCGTGCTGACGCGCTTGCGCGCGT
CAGACCACTTCGGGCTCGCGGGGGTCGGGAACTTTGCTAACAGACTCCGA
GGTGCCATTGGACACAGAGTGGGCGTTCAGCAACGAAAGTGAGTGGGGCC
AGACTTCGCCATAAGGCCTTTATCTTCTTGCCATTTGTCAGTATAAGGGG
TTGCCATAGGCTTCGGCCTCAATTTTAGGCCTTCCGGACTACCAAAATGG
CCGATTTAGTGACGTCACGGCGGCCATTTTAAGTAAGGCGGAAGTAACTC
CACTATTTACAAAATGGCGGCGGAGCACTTCCGGCTTGCCCAAAATGGCG
GCAAAAAACATCCGGGTCAAAGGTCGTTACCACGTCACAAGTCACGTGGG
AGGGTGGTGCTGTAAACCCGGAAGCAATCCTCTCACGTGGCTAGTCACGT
GACTAACACGTCACACCCGCCATTTTGTTTTACAAAATGGCCGACTTCCT
TCCGCTTTTTTAAAAATAACGGCTCAGCGGCGGCGCGCGCTACGCG
(SEQ ID NO: 830)

[1531] 注释：

推定结构域	*碱基范围*
TATA 盒	77-81
起始元件	95-110
转录起始位点	105
5’UTR 保守结构域	165 - 235
ORF2	335 - 703
ORF2/2	335 - 699; 2326 - 2759
[1532] ORF2/3	335 - 699; 2552 - 2957
ORF2t/3	335-465; 2552 - 2957
ORF1	574 - 2775
ORF1/1	574 - 699; 2326 - 2775
ORF1/2	574 - 699; 2552 - 2759
三开放阅读框区	2535 - 2746
聚(A)信号	2953 - 2958
GC 富集区**	3620 - 3648

[1533] 表A2.新颖指环病毒氨基酸序列(甲型细环病毒,进化分枝6)

[1534]

TTV-RTx1 (甲型细环病毒进化分枝6)	
ORF2	MSWSRPVHNATGIERNWWESCLRSHASSCGCGNFVNHINVLANRYGFAGS TETPGNPRPRPPVLSSTTSTPTDQSRPALPWHGDTGGEGASGDPAGDGERGA AEGDYGPEDLDALFDALDEE (SEQ ID NO: 831)
ORF2/2	MSWSRPVHNATGIERNWWESCLRSHASSCGCGNFVNHINVLANRYGFAGS TETPGNPRPRPPVLSSTTSTPTDQSRPALPWHGDTGGEGASGDPAGDGERGA AEGDYGPEDLDALFDALDEEQSKTLVKTKVPDPIDSLETYKLRTRQPWDLN MCSTRGTGDGASLASELSDACSKNLSTMMSILKNQKDLEYFLQQKQSPETK SSKKARFQRKKSRYSPQKRSRKRRYSDSSSDSSSDSCASGSSSNSSNNNSSRR KRACT (SEQ ID NO: 832)
ORF2/3	MSWSRPVHNATGIERNWWESCLRSHASSCGCGNFVNHINVLANRYGFAGS TETPGNPRPRPPVLSSTTSTPTDQSRPALPWHGDTGGEGASGDPAGDGERGA AEGDYGPEDLDALFDALDEENRVPKPRARRKLAFRGRKVATLHRRDPERG DTATVQATAQATAAPRAAAQTPPTTTPQDASGPAPKPPFIFPAINKVYLFPD RAPKPKPTSGDWETEYAACSAFDRPARTNLSSPPYYPGVPTPWQVKFSLKF Q (SEQ ID NO: 833)
ORF2t/3	MSWSRPVHNATGIERNWWESCLRSHASSCGCGNFVNHINVLANRNRVPKP RARRKLAFRGRKVATLHRRDPERGDTATVQATAQATAAPRAAAQTPPTTTP QDASGPAPKPPFIFPAINKVYLFPDRAPKPKPTSGDWETEYAACSAFDRPART NLSSPPYYPGVPTPWQVKFSLKFQ (SEQ ID NO: 834)
ORF1	MAWGYWWRRRFWRPRRRWRTWRRRRRLRPRRSRRTFRRTRRRVRRRRW GRRARRRGWRRRTYIRARRRRRRKRLVLTQWHPAVRRKCKITGYMPIVYC GHGRASFNYAWHSDDCIKQPLPFGGSLSTVSFNLKVLFDENQRGLNKWSYP NDQLDLARYKGCRLTFYRKKNTDYIAQYDISEPYQLDKYSCANYHPSKMM FAKNKILIPSYDTKPRGRQRVRVRIGPPKLFTDKWYSQSDLCKVNLVSLAVS AASFLHPFGSPQTANFCATFQVLQPFYYQAIGISSTKHSEVIDILYKKNTYWQ SNITSWFLTNVKNPKNMSTKMFEDINVKSNKDSNYDWFPFTPYTTENYSKIQ NAAQEYWKYLTSDHPQATNSNEGLVQPWTNATIKQYEYHLGMFSPIFIGPT RAKTKFKTAYFDCTYNPLLDKGMGNRIWYQYATKADTQISKTGCYCMLED IPIYAAFYGYVDFIEMEIGKGQDIKENGLICCICRYTDPPMYNEQHPDMGFVF YNTNFGNGKWIDGRGDIPTYWMQRWRPVVLFQTDVIRDLVETGPFSYKDD

[1535]

	LANTSLTMKYEFYFTWGGNQAYHQTIKNPCKDEGTGPHRQPRDVQVTDPT TVGPEYVFHAWDWRRGFLSERALRRMFEKPLNYDEYSKKPKRPRIFPPTET ESRNQELEESSLSEEEKSLLSTEEIQKEEIQRQFKRQLKRQLRLGQQLKLLQQ QLLKTQAGLHLNPLSYFPQ (SEQ ID NO:835)
ORF1/1	MAWGYWWRRRFWRPRRRWRTWRRRRRLRPRRSRRTFRRTRRRTIKNPCK DEGTGPHRQPRDVQVTDPTTVGPEYVFHAWDWRRGFLSERALRRMFEKPL NYDEYSKKPKRPRIFPPTETESRNQELEESSLSEEEKSLLSTEEIQKEEIQRQFK RQLKRQLRLGQQLKLLQQQLLKTQAGLHLNPLSYFPQ (SEQ ID NO: 836)
ORF1/2	MAWGYWWRRRFWRPRRRWRTWRRRRRLRPRRSRRTFRRTRRRKQSPETK SSKKARFQRKKSRYSPQKRSRKRRYSDSSSDSSSDSCASGSSSNSSNNNSSRR KRACT (SEQ ID NO: 837)

[1536] 表A3.新颖指环病毒核酸序列(甲型细环病毒)

[1537] 名称 TTV-RTx2

[1538] 属/进化分枝 甲型细环病毒,进化分枝6

[1539] 登录号 SRR3479021

[1540] 全序列:3704bp

[1541]

```
1         10         20         30         40         50
|         |          |          |          |          |
CCCCGAAGTCCGTCACTAACCACGTGACTCCCACAGGCCAATCAGATGCT
ATGTCGTGCACTTCCTGGGCTGTGTCTACGTCCTCATATAAGTAACTGCA
CTTCCGAATGGCTGAGTTTTCCACGCCCGTCCGCAGCGGCAGCACCACGG
AGGGTGATCCCCGCGTCCCGTGGGCGGGTGCCGGAGGTGAGTTTACACAC
CGCAGTCAAGGGGCAATTCGGGCACGGGACTGGCCGGGCTATGGGCAAGG
CTCTTAAAAAGCTATGTTCTTCGGTAGGTGCTGGAGAAAGAAAAGGAAAG
TGCTTCTGCAAGATCTGTCAACTCCACCGAAAAAACCTGCTATGAGTGTG
TGGCTTCCTCCCATAGACAATGTTACCGAGCGTGAGAGGAGCTGGCTCTC
TAGCATTCTTCAGTCTCACAGAGCTTTTTGTGGGTGCCATGATGCTATCT
ATCATCTTAGCAGTCTGGCTGCTCGCTTTAATATGCAACCAGGGCCGTCG
CCGGGTGGTGATTCTAGGCCGCCGCGACCGCCACTAAGACGCCTGCCCGC
GCTCCCGGGTCCCAGAGACCCCCCTAGCGACACCAACAACCGCAGGTCAT
GGCCTACTGGGGATGGTGGAGACGGAGGCGCTGGCCAAGGCGCAGGTGGA
GGCGCTACCGCTACCGAAGAAGACTACCGCGCCGAAGACCTAGACGAGCT
GTACGCCGCCCTCGAAGGAGACGAGTAAGGAGGCGCCGCGGTAGGGGGTG
GTACAGAGGGCGACGCTACTCCCGCAGACGGTACAGACGTAGATATGTGA
GGCGAAAGAGAAAGACTCTAGTTTGGAGACAGTGGCAGCCTCAAAATATC
AGAAAATGCAGGATCAGGGGCATAATTCCCATCCTGATATGCGGACACGG
GAGGGGGGCCAGAAACTATGCGCTCCACAGCGACGACATAACCCCCCAGA
```

[1542]

ACACCCCCTTCGGGGGAGGACTGAGCACCACCTCCTGGAGCCTAAAAGTG
CTATATGACCAGCACACCAGGGGACTCAACAGGTGGTCTGCCAGTAACGA
GAGCCTAGACCTTGCCAGATACAATGGCTGTAGTTTCACTTTCTACAGAG
ACAAAAAGACTGACTTTATAGTGACCTATGACACCTCTGCTCCCTACAAA
CTAGACAAATACAGCTCCCCCAGCTACCACCCAGGGTCCATGATGCTCAT
GACAAAACACAAAATCCTGATCCCCAGTTTTGACACAAAACCCAAAGGTC
CTGCCAAAATTAGAGTCAGAATCAAGCCCCCCAAAATGTTCTTAGATAAA
TGGTACACTCAAGACGACCTCTGTTCCGTTAATCTTGTGTCACTTGCGGT
TAGCGCAGCTTCCTTTACACATCCGTTCTGCCCACCACTAACTGACACTC
CTTGTGTAACGCTGCAGGTGTTGAAAGACTTCTACTACACAACCATAGGC
TACTCCTCTAATGCAGACAAAGTAGAGTCTGTATTCACTAACACTCTCTA
CAAACACTGCTGCTACTATCAGTCCTTTCTCACCACTCAATTTATAGCCA
AAATCACTCGCACACCAGATGGACAACCAGTAGCCACATTCTCTCCTCCT
ACCTCTTTCCCTGGCACAACTGTAACAAAAAGTTCCATAGAATCATTTAA
CCAATGGGTAACTTCCACAGGTACAAGTGGCTGGCTAACAAATGCAAACC
AACACTTTCATTTCTGTAACTATAAACCAGATGCCACAAAGCTAAAATGG
CTCAGACAGTACTACTTTGACTGGGAAACATACAAATTAGCAGATGTAAA
GCCAGACGGCCTTACACCCTCAGTAAACTGGTATGAGTACAGAATAGGCC
TCTTTAGTCCTATTTTCCTGAGCCCCTTCAGATCTAGCAGTCTAGACTTT
CCCAGAGCCTACCAGGATGTGAACTACAACCCCCTGGTAGACAAAGGAGT
GGGCAACATCATATGGTTCCAATACAACACAAAACCAGACACACAGCTGT
CAGTACCCAGCTGCAAGTGTGTCATAGAAGACAAACCCCTATGGGCAGCC
TTCTATGGCTACAGTGACTTTGTACAACAAGAGATAGGAGACTACACAGA
CGCAGAGGCCGTGGGCTTCGTCTGTGTCATCTGTCCATACACCAAACCCC
CTCTAAAAAACCCAGACAACCCCATGCAAGGGTTCATATTCTATGACAGC
CTTTTTGGCAATGGCAAGTGGATAGATGGCACGGGGCACGTCCCCCTTTA
CTGGCAGAGCAGGTGGAGGCCAGAGATGCTCTTCCAAGAAAACACCATGA
GAGACATCACACTATCTGGGCCCTTCAGCTACAAGGACGACTATAAGAAC
TGTGTACTGACTTGCAAATACAAATTTAACTTTCGATTCGGGGGCAATCT
TCTCCACGAACAGACGATCAGAAACCCATGCCCCACGGACGGACATCCCA
GTACCGGTAGACAGCCTAGAGACGTACAAGTGGTTGACCCGATCAAAGTG
GGCCCCCGGTTCGTGTTCCACTCCTGGGACTGGCGCAGAGGCTACCTTAG
CCCAGCAGCTCTCAAAAGAATTGGAGAGCAACCGCTCGATTATGAAGCTT
ATTCGTACCGCCCAAAGAGACCTAGAATCTTTCCTCCCACAGAAGGAGAC
CAGCTCGCCCGAAGTCGAGAAGAAGACTCATTTTCAGAGGAAGAAAGTCC
CCATATCTCGTTCGAAGAGGGGCAGGAACCGAAAGCCCAGGCGGTACAGC
AGCACCTCCTCCGACACCTCAGAAAGCAGCGAGAACTCCGAAAGCGACTC
CGAGCCCTGTTCCAAAGCCTCCAAAAGACGCAGGCGGGTCTCCACGTAAA
TCCATTATTATTCAACCAGCCTGCAATCAGGTTCTGATGTTCCCAGAGAT
GGGGCCTAAGCCAGCTCCCACTGCCCAAGACTGGCAGTGCGAATACGAGA
CATGTAAGCACTGGGATAGACCCCCCAGAAAGTTTCTCACAGACCCCCCT
TTCTATCCCTGGGCCCCTACTACTTACAATGTATCTTTCAAGCTAAACTT
CAAATAAACTAGGCCGTGGGAGTCTCACTTGTCGGTGTCTACCTCTTAAG
GTCACTAAGCACTCCGAGCGTCAGCGAGGAGTGCGACCCTTCCCCCTGGT
GCAACGCCCTCGGCGGCCGCGCGCTACGCCTTCGGCTGCGCGCGGCACCT
CGGACCCCCGCTCGTGCTGACGCGCTCGCGCGCGTCAGACCACTTCGGGC
TCGCGGGGGTCGGGAAATTTGCTAAACAGACTCCGAGTTGCCATTGGACA
CAGGAGCTGTGAATCAGTAACGAAAGTGAGTGGGGCCAGACTTCGCCATA
AGGCCTTTATCTTCTTGCCATTTGTCCGTGAGGAGGGGTCGCCAAGACGC
GGACCCCGTTTTCGGACCTTCCGAACTACCAAAATGGCCGATTCAGTGAC
GTCACGGCAGCCATTTTGTGTAAGCACCGCCCAGGACAGACGTCACAGTT
CAAAGGTCATCCTCGAGCGGAACTTACAGAAAATGGCGGTCAATTGCTTC
CGGGTCAAAGGTCACGCCTACGTCATAAGTCACGTGGTGGAGGCTACTGC
GCATACACGGAAGTAGGCCCCGCCACGTGACCGACCACGTGGGTGCTGCG
TCACGGCCGCCATTTTGTATCACAAAATGGCCGACTTCCTTCCTCTTTTT
CAAA (SEQ ID NO: 838)

[1543] 注释：

推定结构域	*碱基范围*
TATA 盒	87 - 91
起始元件	105-120
转录起始位点	115
5' UTR 保守结构域	175 - 245
ORF2	342 - 728
ORF2/2	342 - 724; 2414 - 2849
ORF2/3	342 - 724; 2643 - 3057
ORF1	599 - 2887
ORF1/1	599 - 724; 2414 - 2887
ORF1/2	599 - 724; 2643 - 2849
三开放阅读框区	2626 - 2846
聚(A)信号	3052 - 3058

[1544]

[1545] 表A4.新颖指环病毒氨基酸序列(甲型细环病毒,进化分枝6)

TTV-RTx2 (甲型细环病毒进化分枝 6)	
ORF2	MSVWLPPIDNVTERERSWLSSILQSHRAFCGCHDAIYHLSSLAARFNMQPGP SPGGDSRPPRPPLRRLPALPGPRDPPSDTNNRRSWPTGDGGDGGAGQGAGG GATATEEDYRAEDLDELYAALEGDE (SEQ ID NO: 839)
ORF2/2	MSVWLPPIDNVTERERSWLSSILQSHRAFCGCHDAIYHLSSLAARFNMQPGP SPGGDSRPPRPPLRRLPALPGPRDPPSDTNNRRSWPTGDGGDGGAGQGAGG GATATEEDYRAEDLDELYAALEGDERSETHAPRTDIPVPVDSLETYKWLTR SKWAPGSCSTPGTGAEATLAQQLSKELESNRSIMKLIRTAQRDLESFLPQKE TSSPEVEKKTHFQRKKVPISRSKRGRNRKPRRYSSTSSDTSESSENSESDSEPC SKASKRRRRVST (SEQ ID NO: 840)
ORF2/3	MSVWLPPIDNVTERERSWLSSILQSHRAFCGCHDAIYHLSSLAARFNMQPGP SPGGDSRPPRPPLRRLPALPGPRDPPSDTNNRRSWPTGDGGDGGAGQGAGG GATATEEDYRAEDLDELYAALEGDERRPARPKSRRRLIFRGRKSPYLVRRG AGTESPGGTAAPPPTPQKAARTPKATPSPVPKPPKDAGGSPRKSIIIQPACNQ

[1546]

[1547]

	VLMFPEMGPKPAPTAQDWQCEYETCKHWDRPPRKFLTDPPFYPWAPTTYN VSFKLNFK (SEQ ID NO: 841)
ORF1	MAYWGWWRRRRWPRRRWRRYRYRRRLPRRRPRRAVRRPRRRRVRRRRG RGWYRGRRYSRRRYRRRYVRRKRKTLVWRQWQPQNIRKCRIRGIIPILICGH GRGARNYALHSDDITPQNTPFGGGLSTTSWSLKVLYDQHTRGLNRWSASNE SLDLARYNGCSFTFYRDKKTDFIVTYDTSAPYKLDKYSSPSYHPGSMMLMT KHKILIPSFDTKPKGPAKIRVRIKPPKMFLDKWYTQDDLCSVNLVSLAVSAA SFTHPFCPPLTDTPCVTLQVLKDFYYTTIGYSSNADKVESVFTNTLYKHCCY YQSFLTTQFIAKITRTPDGQPVATFSPPTSFPGTTVTKSSIESFNQWVTSTGTS GWLTNANQHFHFCNYKPDATKLKWLRQYYFDWETYKLADVKPDGLTPSV NWYEYRIGLFSPIFLSPFRSSSLDFPRAYQDVNYNPLVDKGVGNIIWFQYNTK PDTQLSVPSCKCVIEDKPLWAAFYGYSDFVQQEIGDYTDAEAVGFVCVICPY TKPPLKNPDNPMQGFIFYDSLFGNGKWIDGTGHVPLYWQSRWRPEMLFQE NTMRDITLSGPFSYKDDYKNCVLTCKYKFNFRFGGNLLHEQTIRNPCPTDGH PSTGRQPRDVQVVDPIKVGPRFVFHSWDWRRGYLSPAALKRIGEQPLDYEA YSYRPKRPRIFPPTEGDQLARSREEDSFSEEESPHISFEEGQEPKAQAVQQHLL RHLRKQRELRKRLRALFQSLQKTQAGLHVNPLLFNQPAIRF (SEQ ID NO: 842)
ORF1/1	MAYWGWWRRRRWPRRRWRRYRYRRRLPRRRPRRAVRRPRRRRTIRNPCP TDGHPSTGRQPRDVQVVDPIKVGPRFVFHSWDWRRGYLSPAALKRIGEQPL DYEAYSYRPKRPRIFPPTEGDQLARSREEDSFSEEESPHISFEEGQEPKAQAV QQHLLRHLRKQRELRKRLRALFQSLQKTQAGLHVNPLLFNQPAIRF (SEQ ID NO: 843)
ORF1/2	MAYWGWWRRRRWPRRRWRRYRYRRRLPRRRPRRAVRRPRRRRKETSSPE VEKKTHFQRKKVPISRSKRGRNRKPRRYSSTSSDTSESSENSESDSEPCSKAS KRRRRVST (SEQ ID NO: 844)

[1548] 表A5.新颖指环病毒核酸序列(甲型细环病毒)

[1549] 名称 TTV-RTx3

[1550] 属/进化分枝 甲型细环病毒,进化分枝4

[1551] 登录号 SRR3479781

[1552] 全序列:3653bp

```
1         10         20         30         40         50
|         |          |          |          |          |
CCAACCAGAGTCTATGTCGTGCACTTCCTGGGCATGGTCTACGTAATAAT
ATAAAGCGGTGCACTTCCGAATGGCTGAGTTTTCCACGCCCGTCCGCAGC
GAGATCGCGACGGAGGAGCGATCGAGCGTCCCGAGGGCGGGTGCCGGAGG
TGAGTTTACACACCGCAGTCAAGGGGCAATTCGGGCTCGGGACTGGCCGG
GCTATGGGCAAGGCTCTTAAAAAGCCATGTTTCTCGGTAAACTTTACAGG
CAGAAAAGGAAACTGCTACTGCAGCCTGTGCGTGCTCCACAGACGCCATC
TTCCATGAGCTCTACCTGGCGAGTGCCCCGCGGCGATGTCTCCGCCCGCG
AGCTATGTTGGTACCGCTCAGTTCGAGAGAGCCACGATGCTTTTTGTGGC
TGTCGTGATCCTGTTTTTCATCTTTCTCGTCTGGCTGCACGTTCTAACCA
TCAGGGACCTCCGACGCCCCCCACGGACGAGCGCCCGTCGGCGTCTACCC
CAGTGAGGCGCCTGCTGCCGCTGCCCTCCTACCCCGGCGAGGGTCCCCAG
GCTAGATGGCCTGGTGGGGATGGAGAAGGCGCTGGTGGCGCCCGCGGAGG
CGCTGGAGATGGCGGCGCCCGCGCAGGCGAAGAAGAGTACCGGCCCGAAG
ACCTCGACGAGCTGTTCGACGCTATCGAACAAGAACAGTAAGGAGACGGA
GGCGAGGGTGGCGGAGGGGCTACAGGCGCCGTTACAGACTGAGACGCTAC
CGTAGAAGGGGCAGGCGACGCAAAAAAATAGTACTGACTCAGTGGAACCC
CCAGACTGTCAGAAAGTGCTTTATCAGAGGACTGATGCCAGTACTATGGG
CGGGCATGGGCACGGGGGGCCACAACTACGCCGTCCGCTCAGATGACTTT
GTGGTAGACAGAGGCTTCGGGGGCTCCTTCGCCACAGAAACTTTCTCCCT
GAGGGTCCTCTTTGACCAGTACCAGAGAGGATTTAATAGGTGGTCTCACA
CCAACGAAGACCTAGACCTGGCCCGCTACACGGGCTGCAAATGGACATTT
TACAGACACCAAGACACAGACTTTATAGTGTACTTTACAAACAATCCCCC
[1553] CATGAAAACCAACCAGCACACAGCCCCTCTCACAACTCCAGGCATGCTCA
TGAGGAGCAAGTATAAAATACTAGTGCCCAGTTTTAAAACAAGACCAAAG
GGCAGAAAAACAGTGTCAGTGAGAGTTAGACCCCCCAAACTGTTTCAGGA
CAAATGGTATACTCAACAGGACCTCTGTCCAGTACCCCTCGTCCAACTGA
ACGTGACCGCAGCGGATTTCACACATCCGTTCGGCTCACCACTAACTGAC
ACGCCTTGCATAAGATTCCAAGTTTTAGGGAACTTATACAACAAGTGCCT
AAATATAGATCTTCCGCAATTTGATGAGGACGGTGAGATACTCACTTCAA
CACCTTATAACAGAGAAAACAAAGAAGATCTTAAAAAGCTTTATAAAACT
CTATTTGTAGATGAACACGCAGGCAATTATTGGCAGACATTCTTAACCAA
CACAATGGTAAAGTCACACATAGATGCAAACCAAGCAAAGACATACGATC
AAGAAAAAACTGCTGCAGAACAAGGTAAAGACCCCTTCCCAACAAACCCA
CCAAAAGACCAATTCACTACCTGGAACAAGAAACTAGTAGACCCTAGAGA
CAGCAACTTTCTCTTTGCCACATATCACCCAAAAAACATTAAAAAAGCTA
TAAAAACCATGAGAGACAACAACTTTGCTCTCACCACAGGCAAAAATGAC
ATATATGGAGACTACACCGCGGCCTACACCAGAAACACCCACATGCTAGA
CTACTACCTAGGCTTTTATAGCCCCATATTTCTTTCCAGCGGTAGGTCCA
ACACAGAGTTCTGGACCGCCTACAGAGACATAGTATATAATCCCCTCTTA
GACAAAGGCACAGGCAACATGATCTGGTTCCAATATCACACAAAAACAGA
CAATATATACAAAAAACCAGAGTGCCACTGGGAGATACTAGACATGCCCC
TGTGGGCCCTCTGCAACGGGTATGTAGAGTACCTAGAGAGCCAAATAAAG
TACGGGGACATCCTAGTAGAGGGCAAAGTCCTCATCAGATGCCCCTACAC
CAAACCCGCACTGGTAGACCCCAATAACAGCCTAGCTGGTTACGTGGTAT
TCAACACCACCTTCGGCCAGGGAAAATGGATAGATGGCAAAGGCTACATC
CCCCTACACGAGAGGAGCAAGTGGTACGTCATGCTCAGATACCAGACCGA
CGTACTCCATGACATAGTGACTTGTGGACCCTGGCAGTACAGAGACGATA
```

ACAAAAACTCTCAGCTAATAGCCAAGTACAGATTCAAGTTCTACTGGGGA
GGTAACATGGTACATTCTCAGGTCATCAGAAACCCGTGCAAAGACACCCA
AGTATCCGGACCCCGTCGACAGCCTCGCGAAGTACAAGTCGTTGACCCGC
AACTCATTACGCCGCCGTGGGTCCTCCACTCGTTCGACCAGAGACGAGGA
ATGTTTACTGCAGGAGCTATCAAACGTCTGCTCAAGCAACCAATACCTGG
CGAGTATGCTCCTACACCACTCAGGGTCCCGCTCCTCTTTCCCTCCTCAG
AGTTCCAGCGAGAGGGAGAAGATGCAGAAAGCGGCTCAGGTTCACCACCC
AAGAGACCGCGACTCTGGCAGGAAGAGGCCAACCAGACGCAAACGGAGTC
CTCGGAGGGGCCGGCGGAGACGACGAGGGAGCTCCTCGAGCGAAAGCTCA
GAGAGCAGCGAGTCCTCAACCTCCAACTCCAGCATGTCGCAGTACAACTC
GCCAAAACCCAAGCGAACCTCCACATAAACCCCCTATTATACTCCCAGCC
TTAAACAAAGTGTATCTATTCCCCCCTGACAAGCCCACTCCCATACAGNN
NNNNNNNNNNNNNNNNAACACAGAGTTCGAAGCCTGCCAGGCCTTCGACA
[1554] GACCACCTAGAAAATACCTCTCAGACACACCTACCTACCCTTGGCTCCCC
GTCCCCAATCCTGAAATAAAGGTCAGCTTTAAGCTCGGTTTCAAATCTTA
CAAGGCCGTGGGAGTTTCACTGGTCGGTGTCTACCTCTTAAGGTCACTAA
GCACTCCGAGCGTCAGCGAGGAGTGCGACCCTTCCCCCTGGTGCAACGCC
CTCGGCGGCCGCGCGCTACGCCTTCGGCTGCGCGCGGCACCTCGGACCCC
CGCTCGTGCTGACGCGCTCGCGCGCGTCAGACCACTTCGGGCTCGCGGGG
GTCGGGAATTTTGCTAAACAGACTCCGAGTTGCCATTGGACACTGTAGCT
GTGAATCAGTAACGAAAGTGAGTGGGGCCAGACTTCGCCATAAGGCCTTT
ATCTTCTTGCCATTGGTCCGTGTAGGGGGTCGCCATAGGCTTCGGGTTCG
GTTTTAGGCCTTCCGGACTACAAAAATGGCGGATTTAGTGACGTCACGGC
CGCCATTTTAAGTAGGTGCCGTCCAGGACTGCTGTTCCGGGTCACAGGGC
ATCCTCGGCGGAACTTACACAAAATGGCGGTCAAAAACATCCGGGTCAAA
GGTCGCAGCTACGTCATAAGTCACGTGCAGGGGTCCTGCTGCGTCATATG
CGG (SEQ ID NO: 845)

[1555] 注释：

推定结构域	*碱基范围*
TATA 盒	50 - 55
起始元件	68-83
转录起始位点	78
5' UTR 保守结构域	138 - 208
ORF2	305 - 691
[1556] ORF2/2	305 - 687; 2422 - 2878
ORF2/3	305 - 687; 2564 - 3317
ORF2t/3	305 - 360; 2564 - 3317
ORF1	556 - 2904
ORF1/1	556 - 687; 2422 - 2904
ORF1/2	556 - 687; 2564 - 2878
三开放阅读框区	2626 - 2846

[1557] 聚(A)信号 3316 - 3319

[1558] 表A6.新颖指环病毒氨基酸序列(甲型细环病毒,进化分枝4)

[1559]

TTV-RTx3 (甲型细环病毒进化分枝 4)	
ORF2	MSSTWRVPRGDVSARELCWYRSVRESHDAFCGCRDPVFHLSRLAARSNHQ GPPTPPTDERPSASTPVRRLLPLPSYPGEGPQARWPGGDGEGAGGARGGAG DGGARAGEEEYRPEDLDELFDAIEQEQ (SEQ ID NO: 846)
ORF2/2	MSSTWRVPRGDVSARELCWYRSVRESHDAFCGCRDPVFHLSRLAARSNHQ GPPTPPTDERPSASTPVRRLLPLPSYPGEGPQARWPGGDGEGAGGARGGAG DGGARAGEEEYRPEDLDELFDAIEQEQSSETRAKTPKYPDPVDSLAKYKSLT RNSLRRRGSSTRSTRDEECLLQELSNVCSSNQYLASMLLHHSGSRSSFPPQSS SEREKMQKAAQVHHPRDRDSGRKRPTRRKRSPRRGRRRRRGSSSSESSESSE SSTSNSSMSQYNSPKPKRTST (SEQ ID NO: 847)
ORF2/3	MSSTWRVPRGDVSARELCWYRSVRESHDAFCGCRDPVFHLSRLAARSNHQ GPPTPPTDERPSASTPVRRLLPLPSYPGEGPQARWPGGDGEGAGGARGGAG DGGARAGEEEYRPEDLDELFDAIEQEQSYQTSAQATNTWRVCSYTTQGPAP LSLLRVPARGRRCRKRLRFTTQETATLAGRGQPDANGVLGGAGGDDEGAP RAKAQRAASPQPPTPACRSTTRQNPSEPPHKPPIILPALNKVYLFPPDKPTPIQ XXXXXXNTEFEACQAFDRPPRKYLSDTPTYPWLPVPNPEIKVSFKLGFKSYK AVGVSLVGVYLLRSLSTPSVSEECDPSPWCNALGGRALRLRLRAAPRTPAR ADALARVRPLRARGGREFC (SEQ ID NO: 848)
ORF2t/3	MSSTWRVPRGDVSARELCWSYQTSAQATNTWRVCSYTTQGPAPLSLLRVP ARGRRCRKRLRFTTQETATLAGRGQPDANGVLGGAGGDDEGAPRAKAQR AASPQPPTPACRSTTRQNPSEPPHKPPIILPALNKVYLFPPDKPTPIQXXXXXX NTEFEACQAFDRPPRKYLSDTPTYPWLPVPNPEIKVSFKLGFKSYKAVGVSL VGVYLLRSLSTPSVSEECDPSPWCNALGGRALRLRLRAAPRTPARADALAR VRPLRARGGREFC (SEQ ID NO: 849)
ORF1	MAWWGWRRRWWRPRRRWRWRRPRRRRRVPARRPRRAVRRYRTRTVRR RRRGWRRGYRRRYRLRRYRRRGRRRKKIVLTQWNPQTVRKCFIRGLMPVL WAGMGTGGHNYAVRSDDFVVDRGFGGSFATETFSLRVLFDQYQRGFNRW SHTNEDLDLARYTGCKWTFYRHQDTDFIVYFTNNPPMKTNQHTAPLTTPG MLMRSKYKILVPSFKTRPKGRKTVSVRVRPPKLFQDKWYTQQDLCPVPLVQ LNVTAADFTHPFGSPLTDTPCIRFQVLGNLYNKCLNIDLPQFDEDGEILTSTP

[1560]

	YNRENKEDLKKLYKTLFVDEHAGNYWQTFLTNTMVKSHIDANQAKTYDQ EKTAAEQGKDPFPTNPPKDQFTTWNKKLVDPRDSNFLFATYHPKNIKKAIKT MRDNNFALTTGKNDIYGDYTAAYTRNTHMLDYYLGFYSPIFLSSGRSNTEF WTAYRDIVYNPLLDKGTGNMIWFQYHTKTDNIYKKPECHWEILDMPLWAL CNGYVEYLESQIKYGDILVEGKVLIRCPYTKPALVDPNNSLAGYVVFNTTFG QGKWIDGKGYIPLHERSKWYYVMLRYQTDVLHDIVTCGPWQYRDDNKNSQ LIAKYRFKFYWGGNMVHSQVIRNPCKDTQVSGPRRQPREVQVVDPQLITPP WVLHSFDQRRGMFTAGAIKRLLKQPIPGEYAPTPLRVPLLFPSSEFQREGED AESGSGSPPKRPRLWQEEANQTQTESSEGPAETTRELLERKLREQRVLNLQL QHVAVQLAKTQANLHINPLLYSQP (SEQ ID NO: 850)
ORF1/1	MAWWGWRRRWWRPRRRWRWRRPRRRRRVPARRPRRAVRRYRTRTVIRN PCKDTQVSGPRRQPREVQVVDPQLITPPWVLHSFDQRRGMFTAGAIKRLLK QPIPGEYAPTPLRVPLLFPSSEFQREGEDAESGSGSPPKRPRLWQEEANQTQT ESSEGPAETTRELLERKLREQRVLNLQLQHVAVQLAKTQANLHINPLLYSQP (SEQ ID NO: 852)
ORF1/2	MAWWGWRRRWWRPRRRWRWRRPRRRRRVPARRPRRAVRRYRTRTELSN VCSSNQYLASMLLHHSGSRSSFPPQSSSEREKMQKAAQVHHPRDRDSGRKR PTRRKRSPRRGRRRRRGSSSSESSESSESSTSNSSMSQYNSPKPKRTST (SEQ ID NO: 853)

[1561] 表A7.新颖指环病毒核酸序列(甲型细环病毒)

[1562] 名称 TTV-RTx4

[1563] 属/进化分枝 甲型细环病毒,进化分枝4

[1564] 登录号 SRR3481579

[1565] 全序列:3742bp

[1566]
```
1         10         20         30         40         50
|         |          |          |          |          |
AAAGTGCTACGTCACTAACCACGTGACACCCACAGGCCAACCGAATGCTA
TGTCGTGCACTTCCTGGGCCGGGTCTACGTCCTCATATAACTACCTGCAC
TTCCGAATGGCTGAGTTTTCCACGCCCGTCCGCAGCGGTGAAGCCACGGA
GGGAGATCAGCGCGTCCCGAGGGCGGGTGCCGAAGGTGAGTTTACACACC
GAAGTCAAGGGGCAATTCGGGCTCGGGACTGGCCGGGCTATGGGCAAGGC
```

TCTGAAAAAAGCATGTTTATTGGCAGGCATTACAGAAAGAAAAGGGCGCT
GCCACTGTGTGCTGTGCGATCAACAAAGAAGGCTTGCAAACTACTAATAG
TAATGTGGACCCCACCTCGCAATGACCAACAGTACCTTAACTGGCAATGG
TACTCAAGTATACTTAGCTCCCACGCTGCTATGTGCGGGTGTCCCGACGT
TGTTGCTCATTTTAATCATCTTGCTTCTGTGCTTCGCGCCCCGCAAAATC
CACCCCCACCCGGTCCCCAGCGAAACCTGCCCCTCCGACGGCTGCCGGCT
CTCCCGGCTGCGCCAGAGGCGCCCGGAGATAGAGCACCATGGCCTATGGC
TGGTGGCGCCGGAGGAGAAGACGGTGGCGCAGGTGGAGACGCAGACCATG
GAGGCGCCGCTGGAGGACCAGAAGACGCAGACCTGTTAGACGCCGTGGCC
GCCGCAGAAACGTAAGGAGACGCCGCAGAGGAGGGAGGTGGAGGAGGAGG
TACAGGAGATGGAAAAGAAAGGGCAGACGCAGAAAAAAAGCTAAAATAAT
AATAAGACAATGGCAACCTAACTACAGAAGGAGATGTAACATAGTAGGCT
ATATTCCTGTACTGATATGTGGCGAAAATACTGTCAGCAGAAACTATGCC
ACACACTCAGACGATACTAACTACCCAGGACCCTTTGGGGGGGGTATGAC
TACAGACAAATTTACCTTAAGAATTCTGTATGACGAGTACAAAAGGTTTA
TGAACTATTGGACAGCATCTAATGAAGACCTAGACCTCTGTAGATATCTA
GGAGTAAACCTGTACTTTTTTAGACACCCAGAAGTAGACTTTATTATAAA
AATAAATACCATGCCCCCTTTTCTAGACACAGAACTAACAGCTCCTAGCA
TACACCCAGGAATGCTAGCCTTAGACAAAAGAGCAAGATGGATACCTAGC
TTAAAATCTAGACCAGGAAAAAAACACTATATTAAAATAAGAGTAGGGGC
GCCTAAAATGTTCACAGATAAATGGTACCCCCAAACAGATCTTTGTGACA
TGGTGCTGCTAACTGTCTATGCAACCGCAGCGGATATGCAATATCCGTTC
GGCTCACCACTAACTGACTCTGTGGTTGTGAACTTCCAGGTTCTGCAATC
CATGTATGATGAAACCATTAGCATATTACCAGATCAAAAGGAGAAAAGAA
TAACGCTGCTCACTAGTATAGCCTTTTATAACACCACACAAACTATAGCC
CAATTAAAGCCATTTATAGATGCAGGCAATATGACTTCAACTACAACAGC
AACAACATGGGGATCATACATAAACACAACCAAATTTAATACAGCAGCCA
CTACAACATACACATACCCAGGCAGTACTACAACTACAGTAACTATGTTA
[1567] ACTTGTAATGACTCCTGGTACAGAGGAACAGTATATAACGACCAAATTAA
AAATTTACCAAAGGAAGCAGCTCAATTATACTTAAAAGCAACAAAAACCT
TACTAGGAAACACCTTCACAAATGACGACCACACACTAGAATACCATGGA
GGACTGTACAGCTCAATTTGGCTGTCCCCCGGCAGATCTTACTTTGAAAC
ACCAGGAGCATACACAGACATAAAATACAACCCATTTACAGACAGAGGAG
AAGGAAACATGCTATGGATAGACTGGCTAAGCAAAAAAAATATGAACTAT
GACAAACTACAAAGTAAATGTTTAATATCAGACCTACCTTTATGGGCAGC
AGCATATGGATATTTAGAATTTTGTGCAAAAAGTACAGGAGACCAAAATA
TACACATGAATGCCAGACTACTAATAAGAAGTCCCTTTACAGACCCCCAA
CTACTAGTACACACAAACCCCACAAAAGGCTTTGTTCCCTACTCTTTAAA
CTTTGGAAATGGTAAAATGCCAGGAGGTAGTAGTAATGTTCCTATTAGAA
TGAGAGCTAAATGGTATCCAACATTGTTTCACCAGCAAGAAGTACTAGAG
GCCTTAGCACAGTCAGGCCCCTTTGCATACCACTCAGACATTAAAAAAGT
ATCTCTGGGTATGAAATACCGTTTTAAGTGGATCTGGGGTGGAAACCCCG
TTCGCCAACAGGTTGTTAGAAATCCCTGCAAAGACTCCCACTCCTCGGTC
AATAGAGTCCCTAGAAGCTTACAAATCGTTGACCCGAAATACAACTCACC
GGAACTCACATTCCATACGTGGGACTTCAGACGTGGCCTCTTTGGCCAGA
AAGCTATTGAGAGAATGCAACAACAACCAACAACTACTGACATTTTTTCA
GCAGGCCGCAAGAGACCCAGGAGGGACACCGAGGTGTACCACTCCAGCCA
AGAAGGGGAGCAAAAAGAAAGCTTACTTTTCCCCCCAGTCAAGCTCCTCA
GACGAGTCCCCCCGTGGGAAGACTCGCAGCAGGAGGAAAGCGGGTCGCAA
AGCTCAGAGGAAGAGACGCAGACCGTCTCCCAGCAGCTCAAGCAGCAGCT
GCAGCAACAGCGAATCCTGGGAGTCAAACTCATACTCCTGTTCAACCAAG
TCCAAAAAATCCAACAAAATCAAGATATCAACCCTACCTTGTTACCAAGG
GGGGGGGATCTAGCATCCTTATTTCAAATAGCACCATAAACATGTTTGGA
GACCCCAAACCTTACAACCCTTCCAGTAATGACTGGAAAGAGGAGTATGA
GGCCTGTAGAATATGGGACAGACCCCCAAGAGGCAATCTAAGAGACACCC
CCTTTTACCCCTGGGCCCCCAAAGAAAACCAGTACCGTGTAAACTTTAAA
CTTGGATTTCAATAAAGCTAGGCCGTGGGACTTTCACTTGTCGGTGTCTG

CTTATAAAAGTAACCAAGCACTCCGAGCGAAGCGAGGAGTGCGACCCTTG
GGGGCTCAACGACTTCGGAGCCGCGCGTTAAGCCTTCGGCTGCGCGCGGC
ACCTCAGACCCCCGCTCGTGCTGACACGCTTGCGCGTGTCAGACCACTTC
GGGCTCGCGGGGGTCGGGAAATTTATTAAACAGACTCCGAGTTGCCATTG
GACACAGTAGTCTATGAACAGCAACGAAAGTGAGTGGGGCCAGACTTCGC
CATAAGGCCTTTATCTTCTTGCCATTTGTCAGTATAGAGGGTCGCCATAG

[1568] GCTTCGGTCTCCATTTTAACCTGTAAAAACTACCAAAATGGCCGTTCCAG
TGACGTGACAGCCGCCATTTTAAGTAGCTGACGTCAAGGATTGACGTAAA
GGTTAAAGGTCATCCTCGGCGGAAGCTACACAAAATGGTGGACAACATCT
TCCGGGTCAAAGGTCGTGCACACGTCAAAAGTCACGTGGTGGGGACCCGC
TGTAACCCGGAAGTAGGCCCCGTCACGTGATTTGTCACGTGTGTACACGT
CACAGCCGCCATTTTGTTTTACAAAATGGCTGACTTCCTTCCTCTTTTTT
GAAAAAAGGCGCCAAAAAAGGCTCCGCCCCCCGGCCCCCCCC (SEQ ID NO: 854)

[1569] 注释：

推定结构域	*碱基范围*
TATA 盒	86 - 90
起始元件	104 - 119
转录起始位点	114
5' UTR 保守结构域	174 - 244
ORF2	353 - 715
ORF2/2	353 - 711; 2362 - 2863
ORF2/3	353 - 711; 2555 - 3065
ORF2t/3	353 - 432; 2555 - 3065
ORF1	589 - 2889
ORF1/1	589 - 711; 2362 - 2889
ORF1/2	589 - 711; 2555 - 2863
三开放阅读框区	2555 - 2863
聚(A)信号	3062 - 3066
GC 富集区或其一部分**	3720 - 3742

[1570]

[1571] 表A8.新颖指环病毒氨基酸序列(甲型细环病毒,进化分枝4)

TTV-RTx4 (甲型细环病毒进化分枝 4)	
ORF2	MWTPPRNDQQYLNWQWYSSILSSHAAMCGCPDVVAHFNHLASVLRAPQN PPPPGPQRNLPLRRLPALPAAPEAPGDRAPWPMAGGAGGEDGGAGGDADH GGAAGGPEDADLLDAVAAAET (SEQ ID NO: 855)
ORF2/2	MWTPPRNDQQYLNWQWYSSILSSHAAMCGCPDVVAHFNHLASVLRAPQN PPPPGPQRNLPLRRLPALPAAPEAPGDRAPWPMAGGAGGEDGGAGGDADH

[1572]

	GGAAGGPEDADLLDAVAAAETLLEIPAKTPTPRSIESLEAYKSLTRNTTHRN SHSIRGTSDVASLARKLLRECNNNQQLLTFFQQAARDPGGTPRCTTPAKKGS KKKAYFSPQSSSSDESPRGKTRSRRKAGRKAQRKRRRPSPSSSSSSCSNSESW ESNSYSCSTKSKKSNKIKISTLPCYQGGGI (SEQ ID NO: 856)
ORF2/3	MWTPPRNDQQYLNWQWYSSILSSHAAMCGCPDVVAHFNHLASVLRAPQN PPPPGPQRNLPLRRLPALPAAPEAPGDRAPWPMAGGAGGEDGGAGGDADH GGAAGGPEDADLLDAVAAAETPQETQEGHRGVPLQPRRGAKRKLTFPPSQ APQTSPPVGRLAAGGKRVAKLRGRDADRLPAAQAAAAATANPGSQTHTPV QPSPKNPTKSRYQPYLVTKGGGSSILISNSTINMFGDPKPYNPSSNDWKEEYE ACRIWDRPPRGNLRDTPFYPWAPKENQYRVNFKLGFQ (SEQ ID NO: 857)
ORF2t/3	MWTPPRNDQQYLNWQWYSSILSSHAAMPQETQEGHRGVPLQPRRGAKRK LTFPPSQAPQTSPPVGRLAAGGKRVAKLRGRDADRLPAAQAAAAATANPGS QTHTPVQPSPKNPTKSRYQPYLVTKGGGSSILISNSTINMFGDPKPYNPSSND WKEEYEACRIWDRPPRGNLRDTPFYPWAPKENQYRVNFKLGFQ (SEQ ID NO: 858)
ORF1	MAYGWWRRRRRRWRRWRRRPWRRRWRTRRRRPVRRRGRRRNVRRRRR GGRWRRRYRRWKRKGRRRKKAKIIIRQWQPNYRRRCNIVGYIPVLICGENT VSRNYATHSDDTNYPGPFGGGMTTDKFTLRILYDEYKRFMNYWTASNEDL DLCRYLGVNLYFFRHPEVDFIIKINTMPPFLDTELTAPSIHPGMLALDKRARW IPSLKSRPGKKHYIKIRVGAPKMFTDKWYPQTDLCDMVLLTVYATAADMQ YPFGSPLTDSVVVNFQVLQSMYDETISILPDQKEKRITLLTSIAFYNTTQTIAQ LKPFIDAGNMTSTTTATTWGSYINTTKFNTAATTTYTYPGSTTTTVTMLTCN DSWYRGTVYNDQIKNLPKEAAQLYLKATKTLLGNTFTNDDHTLEYHGGLY SSIWLSPGRSYFETPGAYTDIKYNPFTDRGEGNMLWIDWLSKKNMNYDKLQ SKCLISDLPLWAAAYGYLEFCAKSTGDQNIHMNARLLIRSPFTDPQLLVHTN PTKGFVPYSLNFGNGKMPGGSSNVPIRMRAKWYPTLFHQQEVLEALAQSGP FAYHSDIKKVSLGMKYRFKWIWGGNPVRQQVVRNPCKDSHSSVNRVPRSL QIVDPKYNSPELTFHTWDFRRGLFGQKAIERMQQQPTTTDIFSAGRKRPRRD TEVYHSSQEGEQKESLLFPPVKLLRRVPPWEDSQQEESGSQSSEEETQTVSQ QLKQQLQQQRILGVKLILLFNQVQKIQQNQDINPTLLPRGGDLASLFQIAP (SEQ ID NO: 859)

[1573]

[1574]

ORF1/1	MAYGWWRRRRRRWRRWRRRPWRRRWRTRRRRPVRRRGRRRNVVRNPC KDSHSSVNRVPRSLQIVDPKYNSPELTFHTWDFRRGLFGQKAIERMQQQPTT TDIFSAGRKRPRRDTEVYHSSQEGEQKESLLFPPVKLLRRVPPWEDSQQEES GSQSSEEETQTVSQQLKQQLQQQRILGVKLILLFNQVQKIQQNQDINPTLLPR GGDLASLFQIAP (SEQ ID NO: 860)
ORF1/2	MAYGWWRRRRRRWRRWRRRPWRRRWRTRRRRPVRRRGRRRNAARDPG GTPRCTTPAKKGSKKKAYFSPQSSSSDESPRGKTRSRRKAGRKAQRKRRRPS PSSSSSSCSNSESWESNSYSCSTKSKKSNKIKISTLPCYQGGGI (SEQ ID NO: 861)

[1575] 表A9.新颖指环病毒核酸序列(甲型细环病毒)

[1576] 名称 TTV-RTx5b

[1577] 属/进化分枝 甲型细环病毒,进化分枝5

[1578] 登录号 SRR3481639

[1579] 全序列:3553bp

```
1         10         20         30         40         50
|         |          |          |          |          |
ATACCTCATCATATAAAGCGGCGCACTTCCGAATGGCTGAGTTTTCCACG
CCCGTCCGCAGCGAGATCGCGACGGAGGAGCGATCGAGCGTCCCGAGGGC
GGGTGCCGGAGGTGAGTTTACACACCGCAGTCAAGGGGCAATTCGGGCTC
GGGACTGGCCGGGCTATGGGGCAAGACTCTTAAAAAAGCCATGTTTCTCG
GTAAACTTTACAGAAAGAAAAGGGCACTGTCACTGCTACGCGTGCGAGCT
CCAGAGGCGAAACCACCTGCTATGAGTTGGAGACCCCCGGTGCACAACCC
CAATGGGATCGAGAGAAACCTGTGGGAGGCATTCTTTCGCATGCATGCTT
CAGCTTGTGGTTGTGGCGATCTTGTTGGCCATCTTACTGTACTGGCTGGT
CGGTATGGTGCTCCTCCTCGTCCCCCGGCCCCCGGCGCTCCCAGACCACC
GCTGATACGCCAGCTGGCCCTTCCGGCGCCCCCCGCCGATCCTCAACAGG
CTAACCCACAATGGCCTGGTGGGGACGGTGGAGAAGATGGCGCTGGAGGC
CCCGCCGCTGGCGGCGCCGTCGCAGACGCCGAGTACCAAGAAGACGAGCT
CAACGCCCTGTTCGACGCCGTCGAGCAAGAAGAGTAAGGAGGAGGCGATG
GGGGAGGCGGAGGTGGAGACGGGGGTACAGACGCAGACTGAGACTAAGAC
GCAGACGCAGACGAAAGCGAAAGATAGTACTAACTCAGTGGAATCCCGCC
AAAGTGCGGAGGTGTACTATTAAGGGAGTTCTGCCCATGATCCTGTGCGG
GGCCGGGCGCTCGGGGTTTAACTACGGACTGCACAGCGACGACTACACTG
TACAGAAGCCCCTTGGCCAGAACCCCCACGGGGGCGGCATGAGTACAGTG
ACTTTTAGCCTACAGGTGCTCTATGACCAGTACCAGAGGTTTATGAACAA
GTGGTCGTACTCCAACGACCAGCTAGACCTCGCCAGGTACTTTGGCTGCA
CCTTCTGGTTCTACAGACACCCAGAGGTGGACTTTGTAGCTCAGTTTGAC
AACGTTCCCCCCATGAAAATGGACGAGAACACAGCCCCAAACACTCATCC
CTCTTTCTTACTACAGAACAAACACAAGGTTAAAATTCCCAGCTTTAAAA
CAAAGCCTTTTGGTAAAAAAAGAGTTAGAGTTACAGTAGGGCCCCCCAAA
CTGTTTGAAGATAAGTGGTACAGCCAACATGACTTGTGTAAGGTGCCCCT
```

[1580]

AGTCAGTTGGCGGTTAACCGCAGCTGACTTCAGGTTTCCGTTCTGCTCAC
CACAAACTGACAACCCTTGCTACACCTTCCAGGTATTGCATGAAGAGTAT
TACCCAGTAATAGGCACTTCTGCTTTAGAAAACGGCAGTAACTACAATAG
CTCAGCTATAACAGCCTTAGAAAAATTCTTATATGAAAAATGCACACACT
ATCAAACATTTGCCACAGACACCAGACTTAATCCTCAGCGACCAGTGTCA
TCTACAAATGCAAACAAAACATACACCCCCTCAGGCTCCCAAGAAACAAT
AGTGTGGGGGCAGTCAGATTTTAATTTATTTAAAAAGCACACAGACAGCA
ACTATGGCTACTGCACCTACTGTCCTACCAATGACTTAGCTACAAAAATT
AAAAAGTACAGAGACAAAAGATTCGACTGGCTAACAAACATGCCAGTAAC
AAACACCTGCCACATAAATGCCACCTTCGCCCGAGGCAAAATTAAAGAAT
GGGAGTACCACCTAGGGTGGTTCTCAAACATCTTTATAGGCAACCTGAGA
CACAACCTAGCATTCCGGGCCGCATACATAGACATCACCTANACAGACAA
GGGAGAAGGCAACATTATCTGGTTCCAGTACCTCACTAAACCCACCACAG
AGTACATAGAAGCCCAAGCAAAGTGCTCCATCACAAACATACCCCTGTAT
GCTGCTTTTTATGGCTACGAAGACTACCTCCAGAGAACACTAGGCCCCTA
CCAAGATGTAGAAACCCTAGGTATAATCTGTGTTAAATGTCCCTACACAG
ATCCCCCTCTAGTTCACAAGTCTACAGATAAAAAGAACTGGGGCTACGTG
TTCTACGACGTGCACTTTGGCAACGGAAAGACCCCAGAGGGACTGGGCCA
GGTGCACCCTTACTGGATGCAGAGGTGGAGACCCTACGTACAGTTTCAGA
AAGACACTATGAACAAAATAGCCAGGACGGGACCGTTCAGCTACAGAGAC
GAGACGCCTTCCATCACCCTGACCGCCGGGTACAAGTTTCATTTTAACTG
GGGGGGCGACTCTATATTTCCACAGATTATTAAAAACCCCTGCCCAGACA
GCGGGGTACGACCTTCATCCAGTAGAGAGCGTCGCTCAGTACAAGTCGTT
[1581] AGCCCGCTCACAATGGGGCCAGAGTACATATTCCACCGGTGGGACTGGCG
ACGGGGGTTCTTTAATCAAAAAGCTCTCAAAAGAATGCTTGAAAAATCAA
TTAATGATGGAGAGTATCCAACAGGCCCAAAGGTCCCTCGATGGTTTCCC
CCACTCGACAACCAAGAGCAAGAAGGCGCCTCAGGTTCAGAGGAGACAAG
GTCGCAGTCCTCGCAAGAAGAAGCCGCTCAAGAAGCCCTCCAAGAAGTCC
AAGAGGCGTCGCTACAGCAGCACCTCCTCCAGCAGTACCGAGAGCAGCGA
CGGATCGGAAAGCAACTCCAACTCGTCATGCTGCAGCTCACCAAGACGCA
GAGCAACCTGCACATAAACCCCCGTGTTCTTGGCCATGCATAAATAAAGT
CTACATGTTTCCCCCCGACAAGCCCATGCCCATACACGGGTACCACGGGT
GGGAGACGGAGTACCAGGCCTGCAAGGCCTTCAACAGGCCCCCCAGAAAC
TACCTTTCAGACAAACCCATCTACCCTTGGCTCCCTCGCCCCGAACCCGA
AATAATAGTGAGCTTTAGGTTCGGTTTCAAATAAACAAGGCCGCAAATAA
ACAAGGCCGTGGGAGTTTCACTGGTCGGTGTCTACCTCTTAAGGTCACTA
AGCACTCCGAGCGTTAGCGAGGAGTGCGACCCTTCCCCCTGGTGCCACGC
CCTCGGCGGCCGCGCGCTACGCCTNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNTGAATCAGTAACGAAAGTGAGTGG
GGCCAGACTTCGCCATAAGGCCTTTATCTTCTTGCCATTGGTCCGTGTGG
GGAGTCGCCATAGGCTTCGGGCTCGGTTTTAGGCCTTCCGGACTACAAAA
ACCGCCATTTTAGTGACGTCACGGCGGCCATTTTAAGTAAGCATGGCGGG
CGGTGACGTACAAGTTGAAAGGTCACCGCGCTTCCGTGTTTACTCAAAAT
GGTGGCCAACTGCTTCCGGGTCAAAGGTCGGCGGCCACGTCATAAGTCAC
GTGGGAGGGCTGCGTCACAAACACGGAAGTGGCTGTCCCACGTGACTTGT
CACGTGATTGCTACGTCACGGCCGCCATTTTAGTTCACAAAATGGCGGAC
TTC (SEQ ID NO: 862)

[1582] 注释：

[1583] *推定结构域* *碱基范围*

[1584]

TATA 盒	12 - 17
起始元件	30 - 45
转录起始位点	40
5' UTR 保守结构域	100 - 171
ORF2	272 - 637
ORF2/2	272 - 633; 2326 - 2767
ORF2/3	272 - 633; 2525 - 2984
ORF2t/3	272 - 633; 2525 - 2984
ORF1	511 - 2793
ORF1/1	511 - 711; 2326 - 2793
ORF1/2	511 - 711; 2525 - 2767
三开放阅读框区	2525 - 2767
聚(A)信号	2981 - 2985
未知序列	3125-3176

[1585] *注释:为保持阅读框所做的修改:

[1586] -"C"插入ORF2 430

[1587] -"N"插入ORF1 1842

[1588] 表A10.新颖指环病毒氨基酸序列(甲型细环病毒,进化分枝5)

[1589]

TTV-RTx5b (甲型细环病毒进化分枝 5)	
ORF2	MSWRPPVHNPNGIERNLWEAFFRMHASACGCGDLVGHLTVLAGRYGAPPR PPAPGAPRPPLIRQLALPAPPADPQQANPQWPGGDGGEDGAGGPAAGGAVA DAEYQEDELNALFDAVEQEE (SEQ ID NO: 863)
ORF2/2	MSWRPPVHNPNGIERNLWEAFFRMHASACGCGDLVGHLTVLAGRYGAPPR PPAPGAPRPPLIRQLALPAPPADPQQANPQWPGGDGGEDGAGGPAAGGAVA DAEYQEDELNALFDAVEQEELLKTPAQTAGYDLHPVESVAQYKSLARSQW GQSTYSTGGTGDGGSLIKKLSKECLKNQLMMESIQQAQRSLDGFPHSTTKSK KAPQVQRRQGRSPRKKKPLKKPSKKSKRRRYSSTSSSSTESSDGSESNSNSSC CSSPRRRATCT (SEQ ID NO: 864)
ORF2/3	MSWRPPVHNPNGIERNLWEAFFRMHASACGCGDLVGHLTVLAGRYGAPPR PPAPGAPRPPLIRQLALPAPPADPQQANPQWPGGDGGEDGAGGPAAGGAVA

[1590]

	DAEYQEDELNALFDAVEQEEPKGPSMVSPTRQPRARRRLRFRGDKVAVLAR RSRSRSPPRSPRGVATAAPPPAVPRAATDRKATPTRHAAAHQDAEQPAHKP PCSWPCINKVYMFPPDKPMPIHGYHGWETEYQACKAFNRPPRNYLSDKPIY PWLPRPEPEIIVSFRFGFK (SEQ ID NO:865)
ORF2t/3	MSWRPPVHNPNGIERNLWEAFFRMHASACGCGDLVGHLTVLAGRPKGPSM VSPTRQPRARRRLRFRGDKVAVLARRSRSRSPPRSPRGVATAAPPPAVPRAA TDRKATPTRHAAAHQDAEQPAHKPPCSWPCINKVYMFPPDKPMPIHGYHG WETEYQACKAFNRPPRNYLSDKPIYPWLPRPEPEIIVSFRFGFK (SEQ ID NO: 866)
ORF1	MAWWGRWRRWRWRPRRWRRRRRRRVPRRRAQRPVRRRRARRVRRRRW GRRRWRRGYRRRLRLRRRRRRKRKIVLTQWNPAKVRRCTIKGVLPMILCG AGRSGFNYGLHSDDYTVQKPLGQNPHGGGMSTVTFSLQVLYDQYQRFMN KWSYSNDQLDLARYFGCTFWFYRHPEVDFVAQFDNVPPMKMDENTAPNT HPSFLLQNKHKVKIPSFKTKPFGKKRVRVTVGPPKLFEDKWYSQHDLCKVP LVSWRLTAADFRFPFCSPQTDNPCYTFQVLHEEYYPVIGTSALENGSNYNSS AITALEKFLYEKCTHYQTFATDTRLNPQRPVSSTNANKTYTPSGSQETIVWG QSDFNLFKKHTDSNYGYCTYCPTNDLATKIKKYRDKRFDWLTNMPVTNTC HINATFARGKIKEWEYHLGWFSNIFIGNLRHNLAFRAAYIDITXTDKGEGNII WFQYLTKPTTEYIEAQAKCSITNIPLYAAFYGYEDYLQRTLGPYQDVETLGII CVKCPYTDPPLVHKSTDKKNWGYVFYDVHFGNGKTPEGLGQVHPYWMQR WRPYVQFQKDTMNKIARTGPFSYRDETPSITLTAGYKFHFNWGGDSIFPQIIK NPCPDSGVRPSSSRERRSVQVVSPLTMGPEYIFHRWDWRRGFFNQKALKRM LEKSINDGEYPTGPKVPRWFPPLDNQEQEGASGSEETRSQSSQEEAAQEALQ EVQEASLQQHLLQQYREQRRIGKQLQLVMLQLTKTQSNLHINPRVLGHA (SEQ ID NO: 867)
ORF1/1	MAWWGRWRRWRWRPRRWRRRRRRRVPRRRAQRPVRRRRARRIIKNPCPD SGVRPSSSRERRSVQVVSPLTMGPEYIFHRWDWRRGFFNQKALKRMLEKSI NDGEYPTGPKVPRWFPPLDNQEQEGASGSEETRSQSSQEEAAQEALQEVQE ASLQQHLLQQYREQRRIGKQLQLVMLQLTKTQSNLHINPRVLGHA (SEQ ID NO: 868)
ORF1/2	MAWWGRWRRWRWRPRRWRRRRRRRVPRRRAQRPVRRRRARRAQRSLDG
	FPHSTTKSKKAPQVQRRQGRSPRKKKPLKKPSKKSKRRRYSSTSSSSTESSD GSESNSNSSCCSSPRRRATCT (SEQ ID NO: 869)

[1591] 表A11.新颖指环病毒核酸序列(甲型细环病毒)

[1592] 名称 TTV-RTx6

[1593] 属/进化分枝 甲型细环病毒,进化分枝5

[1594] 登录号 SRR3438066

[1595] 全序列:3896bp

[1596]

```
1         10         20         30         40         50
|         |          |          |          |          |
TAAACTTCCTCTTTTAATAGGAAACCACAAAATTTGCATTGCCGACCACA
AACGCATATGCAAATTTACTTCCCCAAAAACTCAACCACAAAATTTGCAT
TGCCGCCCACAAACGTCTACTTTAACCACATCCTCTAACATGTTAGAAAC
TCCACCCAACTACTTCATTAGTATACAGCATCACAAGGGAGGAGCCAAAC
AACTATATAACCAAGTGTACTTCCGAATGGCTGAGTTTATGCCGCCAGAC
GGAGACGGGATCGCGACGGAGGAGCGATCGAGCGTCCCGAGGGCGGGTGC
CGGAGGTGAGTTTACACACCGCAGTCAAGGGGCAATTCGGGCTCGGGACT
GGCCGGGCTATGGGCAAGGCTCTTAAAAAAGCCATGTTTCTCGGTCGACC
TTACAGAAAGAAAAGGGCACTGTCACTGCTACGCGTGCGAGCTCCAGAGG
CGAAACCACCTGCTATGAGCTGGAGGCCCCCGGTGCACAACCCTAATGGG
ATCCAGAGAAACCTGTGGGAGGCATTCTTTCGCATGCATGCTGCAGCTTG
TGGTTGTGGCGATCTTGTTGGCCATATTACTGTACTGGCTGGTCGGTATG
GTGCTCCTCCTCGTCCCCCGGCCCCCGGGGCTCCCAGACCACCGCTGATA
CGCCAGCTGGCCCTTCCGGCGCCCCCCGCCGATCCTCAACAGGCTAACCC
ACAATGGCCTGGTGGGGACGGTGGAGAAGATGGCGCTGGAGGCCCCGCCG
CTGGCGGCGCCGTCGCAGACGCCGAGTACCAAGAAGACGAGCTCAACGCC
CTGTTCGACGCCGTCGAGCAAGAAGAGTAAGGAGGAGGCGATGGGGGAGG
CGGAGGTGGAGACGGGGGTACAGACGCAGACTAAGACTGAGACGCAGACG
CAGACGAAAGAAAATAAGACTGACTCAGTGGAACCCAGCCAAAGTCAGGA
GATGTACTATTAAGGGGGTGCTACCCATGATCTTATGCGGCGCCGGCCGC
TCGGGGTTTAACTATGGACTGCACAGCGACGACTACACGGTGCAGAAACC
CCTGGGGCAGAACCCCCACGGGGGCGGCATGAGCACAGTAACTTTTAGCC
TACAAGTACTATTTGACCAGTACCAGAGGTTTATGAACCGGTGGTCGTAC
TCCAACGACCAGCTAGACCTCGCCAGGTACTTTGGCTGCACCTTCTACTT
TTACAGACACCCTGAAATTGACTTTGTAGCTCAGTATGACAATGTACCCC
CAATGAAAATGGACGAGAACACGGCNCCTAACACTCACCCCTCTTTTCTA
CTACAAAACAAACGCAAAATTAAAATCCCCAGCTTTAAAACCAAGCCATT
TGGCAGAAAAAGAGTAAAAGTAACAGTGGGGCCCCCCAAACTGTTTGAAG
ATAAATGGTACAGCCAGCATGACTTGTGTAAGGTGCCCCTAGTCAGTTGG
CGGTTAACCGCATGTGACTTCAGGTTTCCGTTCTGCTCACCACTAACTGA
CAACCCTTGCTACACCTTCCAGGTATTGCATGAAAACTATTACCCAGTCA
TAGGCACTTCCTCTTTAGAAAACGGTACAAACTACAATAACACTGCTATA
ACTACCCTTGAGACATGGCTATATGGAAAATGCACACACTATCAAACATT
TGCCACAGACACCAGACTTAATCCACAGAGACCTGTATCTTCAAGTAATG
CAAATGAAACTTATACTCCTAGTGGTTCTAAAGAATCAATAATATGGGGA
CAGTCTGACTGGGCAAACTTTAAAAAGAACACAGACAGCAACTATGGCTA
CTGTTCCTACTGCCCCTCAAATGGCACTAACGGAACAGTAGATAAAATTA
AAAAATACAGAGACCAAAGATTTAGATGGCTTACAGAAATGCCAGTACCT
```

[1597]

AACACCTGTCACATACATGCCACCTTCGCCCGAGGCACTATTAAATACTG
GGAGTACCACCTAGGCTGGTACTCAAACATATTTATTGGCAACCTCAGAC
ACAACTTAGCCTTCAGACCAGCCTACATAGACATTACCTACAATCCCATC
ACTGACAAAGGAGAGGGCAACATTATCTGGTTCCAGTACCTCACTAAGCC
CACCACAGAATACATAGAAACCCAGGCAAAATGCACCATTACTAACATTC
CCCTTTATGCTGCTTTCTATGGCTACGAAGACTACCTCCAGAGAACACTA
GGCCCCTACCAAGATGTAGAAACCCTAGGCATAATCTGTGTTAAATGTCC
CTACACAGATCCCCCTCTAGTTCACAAAGACAAAAGTAAAACCAACTGGG
GCTACGTATTCTACGACGCCCACTTTGGCAACGGAAAGACCCCAGAGGGA
CTAGGCCAAGTACACCCTTACTGGATGCAGAGATGGAGACCCTATGTACA
GTTTCAAAAAGACACCATGCACAAAATATCCAGAACGGGACCCTTCAGCT
ACAGAGACGACACGCCTTCCATCACCCTCACTGCCGAATACAAGTTTCGT
TTTAACTGGGGGGGCGACTCTATATTTCCACAGATTATTAAAAACCCCTG
CCCAGACACCGGGGTTCGACCTTCAACCGGTAGAGACCGTCGCTCAGTAC
AAGTCGTTAGCCCGCTCACAATGGGACCCCAGTTTATATTCCACTCATGG
GACTGGAGACGGGGGTTCTTTAATCAAAAAACTCTCAAAAGAATGCTTGA
AAAACCAGTTAATGATGGAGAATATCCAACAGGCCCAAAGGTGCCTCGAT
GGTTTCCCCCACTCGACAACCAAGAGCAAGAAGGCGTCTCAGATACAGAG
ACGACAACCTCGCAGTCCTCGCAAGAAGAAGCCGCTCAAGAAGCCCTCCA
AGAAGTCCAAGAGGCGTCGCTACAGCAGCACCTCCTCCAGCAGTACCGAG
AGCAGCGAAGAATCGGAAAGCAACTCCAACTCGTCATGCTCCAACTCACC
AAGACGCAGAGCAACCTGCACATAAATCCCCGTGTCCTTGGCCATGCATA
AATAAAGTGTACATGTTTCCCCCCGAAAAGCCAATGCCCATACACGGCTA
CCACGGGTGGGAGACAGAGTATCAGGCCTGCAAGGCCTTTGACAGGCCCC
CTAGAAACTACCTATCAGACAAACCCATCTACCCCTGGCTTCCCCGCTCC
CAACCAGAATTTAAAGTGAGTTTTAAGCTTGGCTGTCAATAAACAAGNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNGTTTACACAAAATGGTGGCCAAGTCCTTCCGGGTGAAAGGTCGGC
GCCTACGTCATAAGTCACGTGGGGAGGGCTGCGTCACAACCAGGAAGCAA
TCCTCACCACGTGATTTGTCACGTGATCGCTACGTCACGGCCGCCATTTT
AGTTTACAAAATGGCGGACTTCCTTCCTCTTTTTCAAAAATAACGGCCCT
GCGGCGGCGCGCGCGCTGCGCGCGCGCGCCGGGGGCTGCCGCCCCA (SEQ ID NO: 870)

[1598] 注释：

[1599]

推定结构域	*碱基范围*
TATA 盒	206 - 210
起始元件	224 - 239
转录起始位点	234
5'UTR 保守结构域	294 - 364
ORF2	465 - 830
ORF2/2	465 - 826; 2534 - 2975
ORF2/3	465 - 826; 2721 - 3192

[1600]

ORF2t/3	465 - 595; 2721 - 3192
ORF1	704 - 3001
ORF1/1	704 - 826; 2534 - 3001
ORF1/2	704 - 826; 2721 - 2975
三开放阅读框区	2721 - 2975
聚(A)信号	3189 - 3193
未知序列	3198 - 3655
GC 富集区或其一部分**	3844 - 3895

[1601] 表A12.新颖指环病毒氨基酸序列(甲型细环病毒,进化分枝5)

[1602]

TTV-RTx6 (甲型细环病毒进化分枝 5)	
ORF2	MSWRPPVHNPNGIQRNLWEAFFRMHAAACGCGDLVGHITVLAGRYGAPPR PPAPGAPRPPLIRQLALPAPPADPQQANPQWPGGDGGEDGAGGPAAGGAVA DAEYQEDELNALFDAVEQEE (SEQ ID NO: 871)
ORF2/2	MSWRPPVHNPNGIQRNLWEAFFRMHAAACGCGDLVGHITVLAGRYGAPPR PPAPGAPRPPLIRQLALPAPPADPQQANPQWPGGDGGEDGAGGPAAGGAVA DAEYQEDELNALFDAVEQEELLKTPAQTPGFDLQPVETVAQYKSLARSQW DPSLYSTHGTGDGGSLIKKLSKECLKNQLMMENIQQAQRCLDGFPHSTTKS KKASQIQRRQPRSPRKKKPLKKPSKKSKRRRYSSTSSSSTESSEESESNSNSSC SNSPRRRATCT (SEQ ID NO: 872)
ORF2/3	MSWRPPVHNPNGIQRNLWEAFFRMHAAACGCGDLVGHITVLAGRYGAPPR PPAPGAPRPPLIRQLALPAPPADPQQANPQWPGGDGGEDGAGGPAAGGAVA DAEYQEDELNALFDAVEQEEISNRPKGASMVSPTRQPRARRRLRYRDDNLA VLARRSRSRSPPRSPRGVATAAPPPAVPRAAKNRKATPTRHAPTHQDAEQP AHKSPCPWPCINKVYMFPPEKPMPIHGYHGWETEYQACKAFDRPPRNYLSD KPIYPWLPRSQPEFKVSFKLGCQ (SEQ ID NO: 873)
ORF2t/3	MSWRPPVHNPNGIQRNLWEAFFRMHAAACGCGDLVGHITVLAGRISNRPK GASMVSPTRQPRARRRLRYRDDNLAVLARRSRSRSPPRSPRGVATAAPPPA VPRAAKNRKATPTRHAPTHQDAEQPAHKSPCPWPCINKVYMFPPEKPMPIH GYHGWETEYQACKAFDRPPRNYLSDKPIYPWLPRSQPEFKVSFKLGCQ (SEQ ID NO: 874)

[1603]

ORF1	MAWWGRWRRWRWRPRRWRRRRRRRVPRRRAQRPVRRRRARRVRRRRW GRRRWRRGYRRRLRLRRRRRRKKIRLTQWNPAKVRRCTIKGVLPMILCGA GRSGFNYGLHSDDYTVQKPLGQNPHGGGMSTVTFSLQVLFDQYQRFMNR WSYSNDQLDLARYFGCTFYFYRHPEIDFVAQYDNVPPMKMDENTAPNTHP SFLLQNKRKIKIPSFKTKPFGRKRVKVTVGPPKLFEDKWYSQHDLCKVPLVS WRLTACDFRFPFCSPLTDNPCYTFQVLHENYYPVIGTSSLENGTNYNNTAIT TLETWLYGKCTHYQTFATDTRLNPQRPVSSSNANETYTPSGSKESIIWGQSD WANFKKNTDSNYGYCSYCPSNGTNGTVDKIKKYRDQRFRWLTEMPVPNTC HIHATFARGTIKYWEYHLGWYSNIFIGNLRHNLAFRPAYIDITYNPITDKGEG NIIWFQYLTKPTTEYIETQAKCTITNIPLYAAFYGYEDYLQRTLGPYQDVETL GIICVKCPYTDPPLVHKDKSKTNWGYVFYDAHFGNGKTPEGLGQVHPYWM QRWRPYVQFQKDTMHKISRTGPFSYRDDTPSITLTAEYKFRFNWGGDSIFPQ IIKNPCPDTGVRPSTGRDRRSVQVVSPLTMGPQFIFHSWDWRRGFFNQKTLK RMLEKPVNDGEYPTGPKVPRWFPPLDNQEQEGVSDTETTTSQSSQEEAAQE ALQEVQEASLQQHLLQQYREQRRIGKQLQLVMLQLTKTQSNLHINPRVLGH A (SEQ ID NO: 875)
ORF1/1	MAWWGRWRRWRWRPRRWRRRRRRRVPRRRAQRPVRRRRARRIIKNPCPD TGVRPSTGRDRRSVQVVSPLTMGPQFIFHSWDWRRGFFNQKTLKRMLEKPV NDGEYPTGPKVPRWFPPLDNQEQEGVSDTETTTSQSSQEEAAQEALQEVQE ASLQQHLLQQYREQRRIGKQLQLVMLQLTKTQSNLHINPRVLGHA (SEQ ID NO: 876)
ORF1/2	MAWWGRWRRWRWRPRRWRRRRRRRVPRRRAQRPVRRRRARRNIQQAQR CLDGFPHSTTKSKKASQIQRRQPRSPRKKKPLKKPSKKSKRRRYSSTSSSSTE SSEESESNSNSSCSNSPRRRATCT (SEQ ID NO: 877)

[1604] 表1.示例性指环病毒核酸序列(甲型细环病毒,进化分枝1)

[1605] 名称 TTV-CT30F

[1606] 属/进化分枝 甲型细环病毒,进化分枝1

[1607] 登录号 AB064597.1

[1608] 全序列:3570bp

ATTTTGTGCAGCCCGCCAATTCTCGTTCAAACAGGCCAATCAGGAGGCTC
TACGTACACTTCCTGGGGTGTGTCTTCGAAGAGTATATAAGCAGAGGCGG
TGACGAATGGTAGAGTTTTTCCTGGCCCGTCCGCGGCGAGAGCGCGAGCG
GAGCGAGCGATCGAGCGTCCCGTGGGCGGGTGCCGTAGGTGAGTTTACAC
ACCGCAGTCAAGGGGCAATTCGGGCTCGGGACTGGCCGGGCTATGGGCAA
GATTCTTAAAAAATTCCCCCGATCCCTCTGTCGCCAGGACATAAAAACAT
GCCGTGGAGACCGCCGGTGCATAGTGTCCAGGGGCGAGAGGATCAGTGGT
TCGCGAGCTTTTTTCACGGCCACGCTTCATTTTGCGGTTGCGGTGACGCT
GTTGGCCATCTTAATAGCATTGCTCCTCGCTTTCCTCGCGCCGGTCCACC
AAGGCCCCCTCCGGGGCTAGAGCAGCCTAACCCCCCGCAGCAGGGCCCGG
CCGGGCCCGGAGGGCCGCCCGCCATCTTGGCGCTGCCGGCTCCGCCCGCG
GAGCCTGACGACCCGCAGCCACGGCGTGGTGGTGGGGACGGTGGCGCCGC
CGCTGGCGCCGCAGGCGACCGTGGAGACCGAGACTACGACGAAGAAGAGC
TAGACGAGCTTTTCCGCGCCGCCGCCGAAGACGATTTGTAAGTAGGAGAT
GGCGCCGGCCTTACAGGCGCAGGAGGAGACGCGGGCGACGCAGACGCAGA
CGCAGACGCAGACATAAGCCCACCCTAGTACTCAGACAGTGGCAACCTGA
CGTTATCAGACACTGTAAGATAACAGGACGGATGCCCCTCATTATCTGTG
GAAAGGGGTCCACCCAGTTCAACTACATCACCCACGCGGACGACATCACC
CCCAGGGGAGCCTCCTACGGGGGCAACTTCACAAACATGACTTTCTCCCT
GGAGGCAATATACGAACAGTTTCTGTACCACAGAAACAGGTGGTCAGCCT
CCAACCACGACCTCGAACTCTGCAGATACAAGGGTACCACCCTAAAACTG
TACAGGCACCCAGATGTAGACTACATAGTCACCTACAGCAGAACGGGACC
CTTTGAGATCAGCCACATGACCTACCTCAGCACTCACCCCCTTCTCATGC
TGCTAAACAAACACCACATAGTGGTGCCCAGCCTAAAGACTAAGCCCAGG
GGCAGAAAGGCCATAAAAGTCAGAATAAGACCCCCCAAACTCATGAACAA
CAAGTGGTACTTCACCAGAGACTTCTGTAACATAGGCCTCTTCCAGCTCT
GGGCCACAGGCTTAGAACTCAGAAACCCCTGGCTCAGAATGAGCACCCTG
[1609] AGCCCCTGCATAGGCTTCAATGTCCTTAAAAACAGCATTTACACAAACCT
CAGCAACCTACCTCAGCACAGAGAAGACAGACTTAACATTATTAACAACA
CATTACACCCACATGACATAACAGGACCAAACAATAAAAAATGGCAGTAC
ACATATACCAAACTCATGGCCCCCATTTACTATTCAGCAAACAGGGCCAG
CACCTATGACTTACTACGAGAGTATGGCCTCTACAGTCCATACTACCTAA
ACCCCACAAGGATAAACCTTGACTGGATGACCCCCTACACACACGTCAGG
TACAATCCACTAGTAGACAAGGGCTTCGGAAACAGAATATACATACAGTG
GTGCTCAGAGGCAGATGTAAGCTACAACAGGACTAAATCCAAGTGTCTCT
TACAAGACATGCCCCTGTTTTTCATGTGCTATGGCTACATAGACTGGGCA
ATTAAAAACACAGGGGTCTCCTCACTAGCGAGAGACGCCAGAATCTGCAT
CAGGTGTCCCTACACAGAGCCACAGCTGGTGGGCTCCACAGAAGACATAG
GGTTCGTACCCATCACAGAGACCTTCATGAGGGGCGACATGCCGGTACTT
GCACCATACATACCGTTGAGCTGGTTTTGCAAGTGGTATCCCAACATAGC
TCACCAGAAGGAAGTACTTGAGGCAATCATTTCCTGCAGCCCCTTCATGC
CCCGTGACCAGGGCATGAACGGTTGGGATATTACAATAGGTTACAAAATG
GACTTCTTATGGGGCGGTTCCCCTCTCCCCTCACAGCCAATCGACGACCC
CTGCCAGCAGGGAACCCACCCGATTCCCGACCCCGATAAGCACCCTCGCC
TCCTACAAGTGTCGAACCCGAAACTGCTCGGACCGAGGACAGTGTTCCAC
AAGTGGGACATCAGACGTGGGCAGTTTAGCAAAAGAAGTATTAAAAGAGT
GTCAGAATACTCATCGGATGATGAATCTCTTGCGCCAGGTCTCCCATCAA
AGCGAAACAAGCTCGACTCGGCCTTCAGAGGAGAAAACCCAGAGCAAAAA
GAATGCTATTCTCTCCTCAAAGCACTCGAGGAAGAAGAGACCCCAGAAGA
AGAAGAACCAGCACCCCAAGAAAAAGCCCAGAAAGAGGAGCTACTCCACC
AGCTCCAGCTCCAGAGACGCCACCAGCGAGTCCTCAGACGAGGGCTCAAG
CTCGTCTTTACAGACATCCTCCGACTCCGCCAGGGAGTCCACTGGAACCC
CGAGCTCACATAGAGCCCCCACCTTACATACCAGACCTACTTTTTCCCAA
TACTGGTAAAAAAAAAAAATTCTCTCCCTTCGACTGGGAAACGGAGGCCC
AGCTAGCAGGGATATTCAAGCGTCCTATGCGCTTCTATCCCTCAGACACC
CCTCACTACCCGTGGTTACCCCCCAAGCGCGATATCCCGAAAATATGTAA

[1610]
CATAAACTTCAAAATAAAGCTGCAAGAGTGAGTGATTCGAGGCCCTCCTC
TGTTCACTTAGCGGTGTCTACCTCTTAAAGTCACCAAGCACTCCGAGCGT
CAGCGAGGAGTGCGACCCTCCACCAAGGGGCAACTTCCTCGGGGTCCGGC
GCTACGCGCTTCGCGCTGCGCCGGACGCCTCGGACCCCCCCCCGACCCGA
ATCGCTCGCGCGATTCGGACCTGCGGCCTCGGGGGGGGTCGGGGGCTTTA
CTAAACAGACTCCGAGTTGCCACTGGACTCAGGAGCTGTGAATCAGTAAC
GAAAGTGAGTGGGGCCAGACTTCGCCATAGGGCCTTTAACTTGGGGTCGT
CTGTCGGTGGCTTCCGGGTCCGCCTGGGCGCCGCCATTTTAGCTTTAGAC
GCCATTTTAGGCCCTCGCGGGCACCCGTAGGCGCGTTTTAATGACGTCAC
GGCAGCCATTTTGTCGTGACGTTTGAGACACGTGATGGGGGCGTGCCTAA
ACCCGGAAGCATCCCTGGTCACGTGACTCTGACGTCACGGCGGCCATTTT
GTGCTGTCCGCCATCTTGTGACTTCCTTCCGCTTTTTCAAAAAAAAAGAG
GAAGTATGACAGTAGCGGCGGGGGGGCGGCCGCGTTCGCGCGCCGCCCAC
CAGGGGGTGCTGCGCGCCCCCCCCCGCGCATGCGCGGGGCCCCCCCCCGG
GGGGGCTCCGCCCCCCCGGCCCCCCCCCGTGCTAAACCCACCGCGCATGC
GCGACCACGCCCCCGCCGCC (SEQ ID NO: 1)

[1611] 注释：

推定结构域	*碱基范围*
TATA 盒	84 - 90
帽位点	107 - 114
转录起始位点	114
5' UTR 保守结构域	177 - 247
ORF2	299 - 691
ORF2/2	299 - 687; 2137 - 2659
ORF2/3	299 - 687; 2339 - 2831
ORF2t/3	299 - 348; 2339 - 2831
ORF1	571 - 2613
ORF1/1	571 - 687; 2137 - 2613
ORF1/2	571 - 687; 2339 - 2659
三开放阅读框区	2325 - 2610
聚(A)信号	2813 - 2818
GC 富集区	3415 - 3570

[1612]

[1613] 表2.示例性指环病毒氨基酸序列(甲型细环病毒,进化分枝1)

[1614]

TTV-CT30F (甲型细环病毒进化分枝 1)	
ORF2	MPWRPPVHSVQGREDQWFASFFHGHASFCGCGDAVGHLNSIAPRFPRAGPPRPPPG LEQPNPPQQGPAGPGGPPAILALPAPPAEPDDPQPRRGGGDGGAAAGAAGDRGDRD YDEEELDELFRAAAEDDL (SEQ ID NO: 2)
ORF2/2	MPWRPPVHSVQGREDQWFASFFHGHASFCGCGDAVGHLNSIAPRFPRAGPPRPPPG
	LEQPNPPQQGPAGPGGPPAILALPAPPAEPDDPQPRRGGGDGGAAAGAAGDRGDRD YDEEELDELFRAAAEDDFQSTTPASREPTRFPTPISTLASYKCRTRNCSDRGQCSTSG TSDVGSLAKEVLKECQNTHRMMNLLRQVSHQSETSSTRPSEEKTQSKKNAILSSKH SRKKRPQKKKNQHPKKKPRKRSYSTSSSSRDATSESSDEGSSSSLQTSSDSARESTGT PSSHRAPTLHTRPTFSQYW (SEQ ID NO: 3)
ORF2/3	MPWRPPVHSVQGREDQWFASFFHGHASFCGCGDAVGHLNSIAPRFPRAGPPRPPPG LEQPNPPQQGPAGPGGPPAILALPAPPAEPDDPQPRRGGGDGGAAAGAAGDRGDRD YDEEELDELFRAAAEDDLSPIKAKQARLGLQRRKPRAKRMLFSPQSTRGRRDPRRR RTSTPRKSPERGATPPAPAPETPPASPQTRAQARLYRHPPTPPGSPLEPRAHIEPPPYIP DLLFPNTGKKKKFSPFDWETEAQLAGIFKRPMRFYPSDTPHYPWLPPKRDIPKICNIN FKIKLQE (SEQ ID NO: 4)
ORF2t/3	MPWRPPVHSVQGREDQWSPIKAKQARLGLQRRKPRAKRMLFSPQSTRGRRDPRRR RTSTPRKSPERGATPPAPAPETPPASPQTRAQARLYRHPPTPPGSPLEPRAHIEPPPYIP DLLFPNTGKKKKFSPFDWETEAQLAGIFKRPMRFYPSDTPHYPWLPPKRDIPKICNIN FKIKLQE (SEQ ID NO: 5)
ORF1	TAWWWGRWRRRWRRRRPWRPRLRRRRARRAFPRRRRRRFVSRRWRRPYRRRRR RGRRRRRRRRRHKPTLVLRQWQPDVIRHCKITGRMPLIICGKGSTQFNYITHADDIT PRGASYGGNFTNMTFSLEAIYEQFLYHRNRWSASNHDLELCRYKGTTLKLYRHPD VDYIVTYSRTGPFEISHMTYLSTHPLLMLLNKHHIVVPSLKTKPRGRKAIKVRIRPPK LMNNKWYFTRDFCNIGLFQLWATGLELRNPWLRMSTLSPCIGFNVLKNSIYTNLSN LPQHREDRLNIINNTLHPHDITGPNNKKWQYTYTKLMAPIYYSANRASTYDLLREY GLYSPYYLNPTRINLDWMTPYTHVRYNPLVDKGFGNRIYIQWCSEADVSYNRTKSK CLLQDMPLFFMCYGYIDWAIKNTGVSSLARDARICIRCPYTEPQLVGSTEDIGFVPIT ETFMRGDMPVLAPYIPLSWFCKWYPNIAHQKEVLEAIISCSPFMPRDQGMNGWDITI GYKMDFLWGGSPLPSQPIDDPCQQGTHPIPDPDKHPRLLQVSNPKLLGPRTVFHKW DIRRGQFSKRSIKRVSEYSSDDESLAPGLPSKRNKLDSAFRGENPEQKECYSLLKALE EEETPEEEEPAPQEKAQKEELLHQLQLQRRHQRVLRRGLKLVFTDILRLRQGVHWN PELT (SEQ ID NO: 6)
ORF1/1	TAWWWGRWRRRWRRRRPWRPRLRRRRARRAFPRRRRRRFPIDDPCQQGTHPIPDP DKHPRLLQVSNPKLLGPRTVFHKWDIRRGQFSKRSIKRVSEYSSDDESLAPGLPSKR NKLDSAFRGENPEQKECYSLLKALEEEETPEEEEPAPQEKAQKEELLHQLQLQRRH QRVLRRGLKLVFTDILRLRQGVHWNPELT (SEQ ID NO: 7)
ORF1/2	TAWWWGRWRRRWRRRRPWRPRLRRRRARRAFPRRRRRRFVSHQSETSSTRPSEE

[1615]

	KTQSKKNAILSSKHSRKKRPQKKKNQHPKKKPRKRSYSTSSSSRDATSESSDEGSSS SLQTSSDSARESTGTPSSHRAPTLHTRPTFSQYW (SEQ ID NO: 8)

[1616] 表3.示例性指环病毒核酸序列(甲型细环病毒,进化分枝2)

[1617] 名称 TTV-P13-1

[1618] 属/进化分枝 甲型细环病毒,进化分枝2

[1619] 登录号 KT163896.1

[1620] 全序列:3451bp

[1621]

```
1         10         20         30         40         50
|         |          |          |          |          |
AATTTTGCTAAACAGACTCCGAGGTGCTCTTGGACACTGAGTGGGCGTAC
AGCAACGAAAGTGAGTGGGGCCAGACTTCGCCATAAGGCCTTTATCTTCG
GGTCTACATCATAATATAAAGATGTGCACTTCCGAATGGCTGAGTTTTTC
ACGCCATTCCGCAGCGGTGGAGCAGCGCAGCCACGACCCCCGCGTCCCGA
GGGCGGGTGCCGGAGGTGAGTTTACACACCGCAGTCAAGGGGCAATTCGG
GCTCGGGACTGGCCGGGCCCGGGCAAGGCTCTTAAAGCGAAACCATGTTC
CTCGGCAGGCCCTACCGCCACAGAAAGCGGCACCAGGCCGGCAAGAAAGG
GCCACTGCCACTGCCAAATCTGCAACCTGCACAGGAGAAACGGGCTGGTG
GTCCGTCCTTGATGGCCTCCGGACGCAGGGGATGGATGCCCCCGGACCTG
ACGGTCCAGGAGAGGGAGGATGCCTGGTGGACCAGCTTCTGCGCTAGCCA
CCGCAGCTTTTGTAGCTGCGACGATCCTGTGGGCCATATTAATACTCTCG
CCCGCGATAATAGTCCTCTGGCCCAGACTCCTACTACAACTTCAGGCCAG
GGGCCGCCGCCGCCGCCTACGCCTCCGCGGACGCCGGGGCCGCGCCCTGG
GTCTGCTCCGGACCAGGGGGGAAGGATCAGGGCCTCCTGGACCTACCCCC
TAGCCCCCGGAGGTCCCGGTAGCACGCCATGGCCTACTGGTGGGGCCGGA
GACGCCGGTGGCGCCGCTGGAGGAGGCGCCGGCGTCCTCTCCGCCGCCGC
CGGCGGTGGCGGAGAAGGCGACGCTGGCCCAGAAGGCGCCGGTGGAGGCG
AAGGAGACGACGTGCGAGACCTGCTCGCCGCTATCGAAGGAGACGTGGGC
GCAGACGGGTAAGGAGACGCCGTCGCCCCCAGAAACTAGTACTGACTCAG
TGGAATCCCCAGACTGTGAGAAAGTGTGTTATTAGGGGGTTTCTGCCCCT
GTTCTTCTGCGGACAGGGGGCCTACCACAGAAACTTTACAGACCACTATG
ACGATGTGTTCCCCAAGGGACCCAGCGGAGGTGGGCACGGGAGCATGGTG
TTCAACCTGTCCTTTCTGTACCAAGAGTTTAAGAAGCACCACAATAAGTG
GTCGCGCAGCAACCTGGACTTTGACTTAGTGAGATACAAGGGCACAGTGA
TAAAGCTGTACAGACACCAGGACTTTGACTACATAGTGTGGATAAGCAGG
ACCCCTCCCTTCCAGGAGAGCCTGCTCACAGTAATGACCCACCAGCCCAG
CGTCATGCTGCAGGCAAAAAAGTGCATAATAGTAAAGAGCTACAGGACCC
ACCCGGGGGGCAAACCCTATGTAACTGCAAAAGTTAGGCCCCCCAGACTC
CTAACTGACAAGTGGTACTTCCAGTCAGACTTCTGCAACGTTCCGCTTTT
TAGCCTACAGTTTGCCCTTGCGGAACTGCGGTTTCCGATCTGCTCACCAC
AAACTGACACCAATTGCATTAACTTCCTGGTGTTAGATGACATCTACTAC
AAGTTTCTAGATAATAAGCCTAAACAGAGTTCAGACCCTAATGACGAAAA
CAGAATAAAATTCTGGCACGGCCTATGGTCCACTATGAGATATTTAAACA
CCACCTACATAAACACACTGTTTCCAGGCACAGACAGTCTAGTGGCCGCC
AAAGATACTGACAATAGTGTAAATAAATACCCCAGCACAGCCACTAAACA
GCCCTACAAAGACAGTCAGTACATGCAAAATATATGGAATACATCAAAAA
TACATGCCTTATATACGTGGGTAGCAGAGACAAACTACAAAAGACTGCAG
GCCTACTACACACAGACCTACGGAGGCTACCAGAGACAATTTTTCACAGG
AAAACAGTACTGGGACTACAGAGTAGGCATGTTTAGTCCAGCCTTCCTGA
GTCCCAGCAGACTAAATCCCCAGAACCCAGGGGCATACACAGAGGTCTCC
```

[1622]
TACAACCCCTGGACAGACGAGGGCACGGGCAACGTAGTGTGCCTGCAGTA
TCTGACTAAAGAGACCTCAGACTACAAACCAGGTGGTGGGAGCAAGTTCT
GCATAGAAGGTGTGCCTCTATGGGCAGCGCTGGTGGGATACGTAGACATG
TGTAAAAAAGAGGGCAAGGACCCGGGCATCAGACTAAACTGTCTCCTGTT
AGTCAAGTGTCCCTATACAAAGCCTCAGCTGTATGACAAAAAAAACCCCG
AGAAACTGTTTGTACCTTACTCCTATAACTTTGGGCACGGCAAGATGCCG
GGGGGAGACAAATACATACCCATAGAGTTCAAAGACAGGTGGTACCCCTG
CCTGCTCCACCAAGAGGAGTGGATAGAGGACATTGTCAGGTCGGGACCCT
TCGTTCCAAAAGACATGCCCAGCAGCGTCACCTGCATGATGAGGTACAGC
TCTCTTTTTAACTGGGGCGGTAATATAATCCAAGAACAGGCCGTGGAAGA
CCCCTGTAAGAAAGGCACCTTCGTCGTTCCCGGAACCAGTGGCATCGCTC
GCATACTACAAGTCAGCAACCCGGCCAAGCAGACCCCCACGACAACCTGG
CACTCGTGGGACTGGAGACGATCCCTCTTTACAGAGACGGGTCTTAAAAG
AATGCGCGAACAACAACCATATGATGAACTGTCTTATACGGGCCCTAAAA
AGCCAAAACTGTCCCTTCCCGCAGGGCCCGCCGTCCCCGGTGCCGCCGTC
GCCTCCTCCTGGTGGGAAACAAAACAGGTCACCTCGCCAGACGTCAGCGA
GACGGAGACCGAAGCAGAAGCCCACCAAGAGGAAGAGACGGAGCCGGAGG
AGGGAGTCCAGCTCCAGCAGCTGTGGGAGCAGCAACTCCTGCAAAAGCGA
CAGCTGGGAGTCGTGTTCCAGCAACTCCTCCGACTCAGACAGGGGGCGGA
GATCCACCCGGGCCTCGTATAATTCCTGGGCCCCAGAACCCGTACCTGCT
TTTCCCGGAGCAGGCCCCTCCAAAAGTGCCTATTTTTGACCCCTTTGGTC
AGAAAACAGAGCTAGAGCTGTGCGGCTGCTTCGACAGGCCGCCCAGGAAC
AACCCCTACGACCACCCCTTCTACCCCTGGCTGCCCAAAGAGCCTCCCTC
CTACTACCAGGGCTACAAAGTGTCTTTCAAACTAGGGTTCCACCCAGACA
AGCATGTGTGAACCCCGCCAATAAACCACTGCTGCTACACTGATTCTTAG
GCCGTGGGAGTCTCACTGGTCGGTGTCTACCTCTTAAGGTCACTAAGCAC
TCCGAGCGTTAGCGAGGAGTGCGACCCTACCCCCTGGGCCCACTTCTTCG
GAGCCGCGCGCTACGCCTTCGGCTGCGCGCGGCACCTCAGACCCCCGCTC
GTGCTGACACGCTTGCGCGTGTCAGACCACTTCGGGCTCGCGGGGGTCGG
G (SEQ ID NO: 9)

[1623] 注释：

推定结构域	*碱基范围*
TATA 盒	112 - 119
起始元件	128-148
转录起始位点	148
5'UTR 保守结构域	204 - 273
ORF2	412-912
ORF2/2	412 - 908; 2490 - 3039
ORF2/3	412 - 908; 2725 - 3208
ORF1	729 - 2972
ORF1/1	729 - 908; 2490 - 2972
ORF1/2	729 - 908; 2725 - 3039
三开放阅读框区	2699 - 2969
聚(A)信号	3220 - 3225

[1625] GC 富集区 3302 - 3541

[1626] 表4.示例性指环病毒氨基酸序列(甲型细环病毒,进化分枝2)

TTV-P13-1 (甲型细环病毒进化分枝 2)	
ORF2	MASGRRGWMPPDLTVQEREDAWWTSFCASHRSFCSCDDPVGHINTLARDN SPLAQTPTTTSGQGPPPPPTPPRTPGPRPGSAPDQGGRIRASWTYPLAPGGPG STPWPTGGAGDAGGAAGGGAGVLSAAAGGGGEGDAGPEGAGGGEGDDV RDLLAAIEGDVGADG (SEQ ID NO: 10)
ORF2/2	MASGRRGWMPPDLTVQEREDAWWTSFCASHRSFCSCDDPVGHINTLARDN SPLAQTPTTTSGQGPPPPPTPPRTPGPRPGSAPDQGGRIRASWTYPLAPGGPG STPWPTGGAGDAGGAAGGGAGVLSAAAGGGGEGDAGPEGAGGGEGDDV RDLLAAIEGDVGADGPWKTPVRKAPSSFPEPVASLAYYKSATRPSRPPRQPG TRGTGDDPSLQRRVLKECANNNHMMNCLIRALKSQNCPFPQGPPSPVPPSPP PGGKQNRSPRQTSARRRPKQKPTKRKRRSRRRESSSSSCGSSNSCKSDSWES CSSNSSDSDRGRRSTRASYNSWAPEPVPAFPGAGPSKSAYF (SEQ ID NO: 11)
ORF2/3	MASGRRGWMPPDLTVQEREDAWWTSFCASHRSFCSCDDPVGHINTLARDN SPLAQTPTTTSGQGPPPPPTPPRTPGPRPGSAPDQGGRIRASWTYPLAPGGPG STPWPTGGAGDAGGAAGGGAGVLSAAAGGGGEGDAGPEGAGGGEGDDV RDLLAAIEGDVGADGARRPRCRRRLLLVGNKTGHLARRQRDGDRSRSPPRG RDGAGGGSPAPAAVGAATPAKATAGSRVPATPPTQTGGGDPPGPRIIPGPQN PYLLFPEQAPPKVPIFDPFGQKTELELCGCFDRPPRNNPYDHPFYPWLPKEPP SYYQGYKVSFKLGFHPDKHV (SEQ ID NO: 12)
ORF1	MAYWWGRRRRWRRWRRRRRPLRRRRRWRRRRRWPRRRRWRRRRRRARP ARRYRRRRGRRRVRRRRRPQKLVLTQWNPQTVRKCVIRGFLPLFFCGQGAY HRNFTDHYDDVFPKGPSGGGHGSMVFNLSFLYQEFKKHHNKWSRSNLDFD LVRYKGTVIKLYRHQDFDYIVWISRTPPFQESLLTVMTHQPSVMLQAKKCII VKSYRTHPGGKPYVTAKVRPPRLLTDKWYFQSDFCNVPLFSLQFALAELRF PICSPQTDTNCINFLVLDDIYYKFLDNKPKQSSDPNDENRIKFWHGLWSTMR YLNTTYINTLFPGTDSLVAAKDTDNSVNKYPSTATKQPYKDSQYMQNIWNT SKIHALYTWVAETNYKRLQAYYTQTYGGYQRQFFTGKQYWDYRVGMFSP AFLSPSRLNPQNPGAYTEVSYNPWTDEGTGNVVCLQYLTKETSDYKPGGGS KFCIEGVPLWAALVGYVDMCKKEGKDPGIRLNCLLLVKCPYTKPQLYDKK

[1627]

	NPEKLFVPYSYNFGHGKMPGGDKYIPIEFKDRWYPCLLHQEEWIEDIVRSGP FVPKDMPSSVTCMMRYSSLFNWGGNIIQEQAVEDPCKKGTFVVPGTSGIARI LQVSNPAKQTPTTTWHSWDWRRSLFTETGLKRMREQQPYDELSYTGPKKP KLSLPAGPAVPGAAVASSWWETKQVTSPDVSETETEAEAHQEEETEPEEGV QLQQLWEQQLLQKRQLGVVFQQLLRLRQGAEIHPGLV (SEQ ID NO: 13)
ORF1/1	MAYWWGRRRRWRRWRRRRRPLRRRRRWRRRRRWPRRRRWRRRRRRARP ARRYRRRRGRRRAVEDPCKKGTFVVPGTSGIARILQVSNPAKQTPTTTWHS WDWRRSLFTETGLKRMREQQPYDELSYTGPKKPKLSLPAGPAVPGAAVASS WWETKQVTSPDVSETETEAEAHQEEETEPEEGVQLQQLWEQQLLQKRQLG VVFQQLLRLRQGAEIHPGLV (SEQ ID NO: 14)
ORF1/2	MAYWWGRRRRWRRWRRRRRPLRRRRRWRRRRRWPRRRRWRRRRRRARP ARRYRRRRGRRRGPPSPVPPSPPPGGKQNRSPRQTSARRRPKQKPTKRKRRS RRRESSSSSCGSSNSCKSDSWESCSSNSSDSDRGRRSTRASYNSWAPEPVPAF PGAGPSKSAYF (SEQ ID NO: 15)

[1628]

[1629] 表5.示例性指环病毒核酸序列(甲型细环病毒,进化分枝3)

[1630] 名称 TTV-tth8

[1631] 属/进化分枝 甲型细环病毒,进化分枝3

[1632] 登录号 AJ620231.1

[1633] 全序列:3753bp

[1634]
```
1         10         20         30         40         50
|         |          |          |          |          |
TGCTACGTCACTAACCCACGTGTCCTCTACAGGCCAATCGCAGTCTATGT
CGTGCACTTCCTGGGCATGGTCTACATAATTATATAAATGCTTGCACTTC
CGAATGGCTGAGTTTTTGCTGCCCGTCCGCGGAGAGGAGCCACGGCAGGG
GATCCGAACGTCCTGAGGGCGGGTGCCGGAGGTGAGTTTACACACCGAAG
TCAAGGGGCAATTCGGGCTCAGGACTGGCCGGGCTTTGGGCAAGGCTCTT
AAAAATGCACTTTTCTCGAATAAGCAGAAAGAAAAGGAAAGTGCTACTGC
TTTGCGTGCCAGCAGCTAAGAAAAAACCAACTGCTATGAGCTTCTGGAAA
CCTCCGGTACACAATGTCACGGGGATCCAACGCATGTGGTATGAGTCCTT
TCACCGTGGCCACGCTTCTTTTTGTGGTTGTGGGAATCCTATACTTCACA
TTACTGCACTTGCTGAAACATATGGCCATCCAACAGGCCCGAGACCTTCT
GGGCCACCGGGAGTAGACCCCAACCCCCACATCCGTAGAGCCAGGCCTGC
CCCGGCCGCTCCGGAGCCCTCACAGGTTGATTCGAGACCAGCCCTGACAT
GGCATGGGGATGGTGGAAGCGACGGAGGCGCTGGTGGTTCCGGAAGCGGT
GGACCCGTGGCAGACTTCGCAGACGATGGCCTCGATCAGCTCGTCGCCGC
CCTAGACGACGAAGAGTAAGGAGGCGCAGACGGTGGAGGAGGGGGAGACG
AAAAACAAGGACTTACAGACGCAGGAGACGCTTTAGACGCAGGGGACGAA
AAGCAAAACTTATAATAAAACTGTGGCAACCTGCAGTAATTAAAAGATGC
```

AGAATAAAGGGATACATACCACTGATTATAAGTGGGAACGGTACCTTTGC
CACAAACTTTACCAGTCACATAAATGACAGAATAATGAAAGGCCCCTTCG
GGGGAGGACACAGCACTATGAGGTTCAGCCTCTACATTTTGTTTGAGGAG
CACCTCAGACACATGAACTTCTGGACCAGAAGCAACGATAACCTAGAGCT
AACCAGATACTTGGGGGCTTCAGTAAAAATATACAGGCACCCAGACCAAG
ACTTTATAGTAATATACAACAGAAGAACCCCTCTAGGAGGCAACATCTAC
ACAGCACCCTCTCTACACCCAGGCAATGCCATTTTAGCAAAACACAAAAT
ATTAGTACCAAGTTTACAGACAAGACCAAAGGGTAGAAAAGCAATTAGAC
TAAGAATAGCACCCCCCACACTCTTTACAGACAAGTGGTACTTTCAAAAG
GACATAGCCGACCTCACCCTTTTCAACATCATGGCAGTTGAGGCTGACTT
GCGGTTTCCGTTCTGCTCACCACAAACTGACAACACTTGCATCAGCTTCC
AGGTCCTTAGTTCCGTTTACAACAACTACCTCAGTATTAATACCTTTAAT
AATGACAACTCAGACTCAAAGTTAAAAGAATTTTTAAATAAAGCATTTCC
AACAACAGGCACAAAAGGAACAAGTTTAAATGCACTAAATACATTTAGAA
CAGAAGGATGCATAAGTCACCCACAACTAAAAAAACCAAACCCACAAATA
AACAAACCATTAGAGTCACAATACTTTGCACCTTTAGATGCCCTCTGGGG
AGACCCCATATACTATAATGATCTAAATGAAAACAAAAGTTTGAACGATA
TCATTGAGAAAATACTAATAAAAAACATGATTACATACCATGCAAAACTA
AGAGAATTTCCAAATTCATACCAAGGAAACAAGGCCTTTTGCCACCTAAC
AGGCATATACAGCCCACCATACCTAAACCAAGGCAGAATATCTCCAGAAA
TATTTGGACTGTACACAGAAATAATTTACAACCCTTACACAGACAAAGGA
ACTGGAAACAAAGTATGGATGGACCCACTAACTAAAGAGAACAACATATA
TAAAGAAGGACAGAGCAAATGCCTACTGACTGACATGCCCCTATGGACTT
TACTTTTTGGATATACAGACTGGTGTAAAAAGGACACTAATAACTGGGAC
TTACCACTAAACTACAGACTAGTACTAATATGCCCTTATACCTTTCCAAA
ATTGTACAATGAAAAAGTAAAAGACTATGGGTACATCCCGTACTCCTACA
AATTCGGAGCGGGTCAGATGCCAGACGGCAGCAACTACATACCCTTTCAG
TTTAGAGCAAAGTGGTACCCCACAGTACTACACCAGCAACAGGTAATGGA
GGACATAAGCAGGAGCGGGCCCTTTGCACCTAAGGTAGAAAAACCAAGCA
[1635] CTCAGCTGGTAATGAAGTACTGTTTTAACTTTAACTGGGGCGGTAACCCT
ATCATTGAACAGATTGTTAAAGACCCCAGCTTCCAGCCCACCTATGAAAT
ACCCGGTACCGGTAACATCCCTAGAAGAATACAAGTCATCGACCCGCGGG
TCCTGGGACCGCACTACTCGTTCCGGTCATGGGACATGCGCAGACACACA
TTTAGCAGAGCAAGTATTAAGAGAGTGTCAGAACAACAAGAAACTTCTGA
CCTTGTATTCTCAGGCCCAAAAAAGCCTCGGGTCGACATCCCAAAACAAG
AAACCCAAGAAGAAAGCTCACATTCACTCCAAAGAGAATCGAGACCGTGG
GAGACCGAGGAAGAAAGCGAGACAGAAGCCCTCTCGCAAGAGAGCCAAGA
GGTCCCCTTCCAACAGCAGTTGCAGCAGCAGTACCAAGAGCAGCTCAAGC
TCAGACAGGGAATCAAAGTCCTCTTCGAGCAGCTCATAAGGACCCAACAA
GGGGTCCATGTAAACCCATGCCTACGGTAGGTCCCAGGCAGTGGCTGTTT
CCAGAGAGAAAGCCAGCCCCAGCTCCTAGCAGTGGAGACTGGGCCATGGA
GTTTCTCGCAGCAAAAATATTTGATAGGCCAGTTAGAAGCAACCTTAAAG
ATACCCCTTACTACCCATATGTTAAAAACCAATACAATGTCTACTTTGAC
CTTAAATTTGAATAAACAGCAGCTTCAAACTTGCAAGGCCGTGGGAGTTT
CACTGGTCGGTGTCTACCTCTAAAGGTCACTAAGCACTCCGAGCGTAAGC
GAGGAGTGCGACCCTCCCCCCTGGAACAACTTCTTCGGAGTCCGGCGCTA
CGCCTTCGGCTGCGCCGGACACCTCAGACCCCCCCTCCACCCGAAACGCT
TGCGCGTTTCGGACCTTCGGCGTCGGGGGGGTCGGGAGCTTTATTAAACG
GACTCCGAAGTGCTCTTGGACACTGAGGGGGTGAACAGCAACGAAAGTGA
GTGGGGCCAGACTTCGCCATAAGGCCTTTATCTTCTTGCCATTTGTCAGT
GTCCGGGGTCGCCATAGGCTTCGGGCTCGTTTTTAGGCCTTCCGGACTAC
AAAAATCGCCATTTTGGTGACGTCACGGCCGCCATCTTAAGTAGTTGAGG
CGGACGGTGGCGTGAGTTCAAAGGTCACCATCAGCCACACCTACTCAAAA
TGGTGGACAATTTCTTCCGGGTCAAAGGTTACAGCCGCCATGTTAAAACA
CGTGACGTATGACGTCACGGCCGCCATTTTGTGACACAAGATGGCCGACT
TCCTTCCTCTTTTTCAAAAAAAAGCGGAAGTGCCGCCGCGGCGGCGGGGG
GCGGCGCGCTGCGCGCGCCGCCCAGTAGGGGGAGCCATGCGCCCCCCCCC
GCGCATGCGCGGGGCCCCCCCCCGCGGGGGGCTCCGCCCCCCGGCCCCCC
CCG (SEQ ID NO: 16)

[1636] 注释：

推定结构域	*碱基范围*
TATA 盒	83 - 88
帽位点	104 - 111
转录起始位点	111
5' UTR 保守结构域	170 - 240
ORF2	336 - 719
ORF2/2	336 - 715; 2363 - 2789
ORF2/3	336 - 715; 2565 - 3015
ORF2t/3	336 - 388; 2565 - 3015
ORF1	599 - 2830
ORF1/1	599 - 715; 2363 - 2830
ORF1/2	599 - 715; 2565 - 2789
三开放阅读框区	2551 - 2786
聚(A)信号	3011 - 3016
GC 富集区	3632 - 3753

[1637]

[1638] 表6.示例性指环病毒氨基酸序列(甲型细环病毒,进化分枝3)

TTV-tth8 (甲型细环病毒进化分枝 3)	
ORF2	MSFWKPPVHNVTGIQRMWYESFHRGHASFCGCGNPILHITALAETYGHPTGPRPSG PPGVDPNPHIRRARPAPAAPEPSQVDSRPALTWHGDGGSDGGAGGSGSGGPVADFA DDGLDQLVAALDDEE (SEQ ID NO: 17)
ORF2/2	MSFWKPPVHNVTGIQRMWYESFHRGHASFCGCGNPILHITALAETYGHPTGPRPSG PPGVDPNPHIRRARPAPAAPEPSQVDSRPALTWHGDGGSDGGAGGSGSGGPVADFA DDGLDQLVAALDDEELLKTPASSPPMKYPVPVTSLEEYKSSTRGSWDRTTRSGHGT CADTHLAEQVLRECQNNKKLLTLYSQAQKSLGSTSQNKKPKKKAHIHSKENRDRG RPRKKARQKPSRKRAKRSPSNSSCSSSTKSSSSSDRESKSSSSSS (SEQ ID NO: 18)
ORF2/3	MSFWKPPVHNVTGIQRMWYESFHRGHASFCGCGNPILHITALAETYGHPTGPRPSG PPGVDPNPHIRRARPAPAAPEPSQVDSRPALTWHGDGGSDGGAGGSGSGGPVADFA DDGLDQLVAALDDEEPKKASGRHPKTRNPRRKLTFTPKRIETVGDRGRKRDRSPLA REPRGPLPTAVAAAVPRAAQAQTGNQSPLRAAHKDPTRGPCKPMPTVGPRQWLFP ERKPAPAPSSGDWAMEFLAAKIFDRPVRSNLKDTPYYPYVKNQYNVYFDLKFE

[1639]

[1640]

	(SEQ ID NO: 19)
ORF2t/3	MSFWKPPVHNVTGIQRMWPKKASGRHPKTRNPRRKLTFTPKRIETVGDRGRKRDR SPLAREPRGPLPTAVAAAVPRAAQAQTGNQSPLRAAHKDPTRGPCKPMPTVGPRQ WLFPERKPAPAPSSGDWAMEFLAAKIFDRPVRSNLKDTPYYPYVKNQYNVYFDLK FE (SEQ ID NO: 20)
ORF1	MAWGWWKRRRRWWFRKRWTRGRLRRRWPRSARRRPRRRRVRRRRRWRRGRRK TRTYRRRRRFRRRGRKAKLIIKLWQPAVIKRCRIKGYIPLIISGNGTFATNFTSHINDR IMKGPFGGGHSTMRFSLYILFEEHLRHMNFWTRSNDNLELTRYLGASVKIYRHPDQ DFIVIYNRRTPLGGNIYTAPSLHPGNAILAKHKILVPSLQTRPKGRKAIRLRIAPPTLFT DKWYFQKDIADLTLFNIMAVEADLRFPFCSPQTDNTCISFQVLSSVYNNYLSINTFN NDNSDSKLKEFLNKAFPTTGTKGTSLNALNTFRTEGCISHPQLKKPNPQINKPLESQ YFAPLDALWGDPIYYNDLNENKSLNDIIEKILIKNMITYHAKLREFPNSYQGNKAFC HLTGIYSPPYLNQGRISPEIFGLYTEIIYNPYTDKGTGNKVWMDPLTKENNIYKEGQS KCLLTDMPLWTLLFGYTDWCKKDTNNWDLPLNYRLVLICPYTFPKLYNEKVKDY GYIPYSYKFGAGQMPDGSNYIPFQFRAKWYPTVLHQQQVMEDISRSGPFAPKVEKP STQLVMKYCFNFNWGGNPIIEQIVKDPSFQPTYEIPGTGNIPRRIQVIDPRVLGPHYSF RSWDMRRHTFSRASIKRVSEQQETSDLVFSGPKKPRVDIPKQETQEESSHSLQRESR PWETEEESETEALSQESQEVPFQQQLQQQYQEQLKLRQGIKVLFEQLIRTQQGVHV NPCLR (SEQ ID NO: 21)
ORF1/1	MAWGWWKRRRRWWFRKRWTRGRLRRRWPRSARRRPRRRRIVKDPSFQPTYEIPG TGNIPRRIQVIDPRVLGPHYSFRSWDMRRHTFSRASIKRVSEQQETSDLVFSGPKKPR VDIPKQETQEESSHSLQRESRPWETEEESETEALSQESQEVPFQQQLQQQYQEQLKL RQGIKVLFEQLIRTQQGVHVNPCLR (SEQ ID NO: 22)
ORF1/2	MAWGWWKRRRRWWFRKRWTRGRLRRRWPRSARRRPRRRRAQKSLGSTSQNKK PKKKAHIHSKENRDRGRPRKKARQKPSRKRAKRSPSNSSCSSSTKSSSSSDRESKSSS SSS (SEQ ID NO: 23)

[1641] 表7.示例性指环病毒核酸序列(甲型细环病毒,进化分枝4)

[1642] 名称 TTV-HD20a

[1643] 属/进化分枝 甲型细环病毒,进化分枝4

[1644] 登录号 FR751492.1

[1645] 全序列:3878bp

```
1         10         20         30         40         50
|         |          |          |          |          |
AAATACGTCACTAACCACGTGACTCCCACAGGCCAACCACAGTCTATGTC
GTGCACTTCCTGGGCATGGTCTACGTGATAATATAAAGCGGTGCACTTCC
GAATGGCTGAGTTTTCCACGCCCGTCCGCAGCGAGATCGCGACGTAGGAG
CGATCGAGCGTCCCGAGGGCGGGTGCCGGAGGTGAGTTTACACACCGCAG
TCAAGGGGCAATTCGGGCTCGGGAGGCCGGGCCATGGGCAAGGCTCTTAA
AAAGCTATGTTTCTCGGTAAAATCTACAGGAAGAAAAGGAAACTGCTTCT
GCAGGCTGTGCGTGCTCCGCAGACGCCATCTTCCATGAGCCGCTGCTGGT
GTCCCCCTCGGGGTGATGTCTCCTCCCGCGAGTCTCGATGGTACGAGGCG
GTTCGAGGAAGCCACGATGCTTTTTGTGGCTGTAGTGATCCTATTCTTCA
TCTTTCTCGTCTGGCTGCACGTTTTAACCATCAGGGACCTCCGACGCCCC
CCACGGACGACCGTGCGCCGCAGAATACCCCAGTGAGACGCCTGCTGCCT
CTCCCCAGCTACCCCGGCGAGGGTCCCCAGGCTAGATGGCCTGGTGGGGA
TGGAGGCGCCGCTGGTGGCGACCGAAGAGAAGGTGGAGATGGCGGCGCGC
GCGCCGCCGAAGACGAGTACCAGCCCGAAGACCTAGACGAGCTTTTCGGC
GCTATCGAACAAGAACAGTAAGGAGGAGGCGAAGGGGGAGGCGGAGGGGC
TACCGGCGCCGTTACAGACTGAGACGCTATGCCAGACGCAGGTTCCGACG
CAAAAAGATAGTACTGACTCAGTGGAACCCCCAGACTACCAGAAAATGTA
TAATAAGGGGCATGATGCCAGTACTGTGGGCCGGCATGGGTACGGGGGGC
AGAAACTATGCAGTGAGGTCAGATGACTATGTGGTGAACAAAGGGTTCGG
GGGCTCCTTCGCCACGGAGACCTTCTCCCTGAAGGTTCTCTATGACCAGT
TTCAAAGGGGCTTCAACAGGTGGTCCCACACTAACGAGGACCTAGACCTG
GCCCGCTACAGGGGCTGCAGGTGGACTTTTTACAGACATAAAGACACAGA
CTTTATAGTGTACTTTACAAACAATCCTCCCATGAAGACCAACCAGTTCT
CCGCGCCCCTGACGACCCCCGGCATGCTCATGCGCAGTAAATACAAAGTC
[1646] CTCATTCCCAGCTTCCAGACCAGACCCAAGGGTCGCAAAACAGTAACCGT
TAAAATAAGACCCCCCAAACTATTTCAAGACAAGTGGTACACCCAGCAGG
ACCTGTGTTCAGTTCCTCTTGTCCAACTGAACGTGACCGCAGCTGATTTC
ACACATCCGTTCGGCTCACCACTAACTGAAACTCCTTGCGTAGAGTTCCA
GGTGCTGGGTGACTTGTACAATACATGTCTCAATATCGACCTTCCGCAAT
TTAGTGAATTAGGAGAAATAACTAGTGCCTACTCAAAACCAAACTCAAAT
AACCTAAAAGAATTATACAAAGAATTGTTCACAAAAGCCACATCAGGACA
CTACTGGCAGACATTCATAACCAACAGCATGGTCAGAGCACACATAGATG
CAGACAAAGCTAAAGAAGCACAAAGAGCATCCACCACACCCTCATACAAC
AATGACCCCTTCCCCACAATACCTGTTAAATCAGAGTTTGCACAGTGGAA
AAAGAAATTCACAGACACTAGAGACAGCCCCTTTCTTTTTGCCACTTACC
ATCCCGAAGCTATAAAAGACACAATTATGAAAATGAGAGAGAACAACTTT
AAGCTAGAGACAGGACCCAATGACAAGTATGGAGACTACACAGCACAGTA
CCAAGGAAACACACACATGCTAGACTACTACCTTGGCTTTTACAGCCCCA
TATTCCTCTCAGATGGAAGGTCTAACGTAGAATTCTTCACTGCCTACAGA
GACATAGTATACAATCCCTTCTTAGACAAGGCCCAGGGCAACATGGTGTG
GTTTCAGTACCACACAAAGACAGACAACAAGTTTAAAAAACCAGAGTGCC
ACTGGGAAATCAAAGACATGCCCCTGTGGGCCCTCCTAAACGGATATGTA
GACTACTTAGAGACTCAAATACAGTATGGTGACCTCAGTAAAGAAGGGAA
AGTCCTCATCAGGTGTCCCTACACCAAGCCAGCACTAGTAGACCCCAGAG
ACGACACTGCAGGATATGTAGTCTACAACAGAAACTTTGGCAGAGGCAAG
TGGATAGACGGAGGGGGCTACATCCCCCTGCACGAGAGGACAAAATGGTA
CGTGATGCTCAGATACCAGACGGACGTCTTCCATGACATAGTGACCTGTG
GGCCCTGGCAGTACAGAGACGACAACAAAAACAGCCAGCTAGTGGCCAAA
TACCGCTTCAGCTTTATATGGGGAGGTAACACTGTCCACTCTCAGGTCAT
CAGAAACCCGTGCAAAGACAACCAAGTATCCGGTCCCCGTCGACAGCCTA
GGGATATACAAGTCGTTGACCCGCAACGCATCACGCCGCCGTGGGTCCTC
```

CACAGCTTCGACCAGCGAAGAGGCCTCTTTACTGAAACAGCTCTCAGGCG
CCTGCTCCAGGAACCACTACCTGGCGAGTATGCTGTTAGCACCCTCAGGA
CACCCCTCCTCTTTCTACCCTCAGAATACCAGCGAGAAGACGGCGCTGCA
GAAAGCGCCTCAGGTTCACCGGCCAAAAGACCCCGTATCTGGTCAGAAGA
GAGTCAGACGGAGACGATCTCCTCGGAGGAGAACCCGGCGGAGACGACGA
GGGAGCTCCTCCAGCGAAAGCTCCGAGAGCAGCGAGCACTCCAGTTCCAA
CTCCAGCACTTCGCGGTCCAACTCGCCAAGACCCAGGCGAATCTCCACGT
AAACCCCCTGTTATCTTTCCCGCAATGAATAAGGTCTTTCTGTTTCCCCC
AGAGGGTCCCAAGCCCATCCTGGGCAAAGAGGCCTGGCAGGACGAGTACG
AGACCTGCAGGGTCTGGAACAGACCTGCCAGAACCCACCACACAGACACC
CCCTTCTATCCCTGGGCCCCCCACAAGTTCCATGTAAGCTTCAAACTTGG
CTTCCAATAAAATTACTAGGCCGTGGAACTCTCACTGGTCGGTGTCTACC
TCTTAAGGTCACTAAGCACTCCGAGCGTCAGCGAGGAGTGCGACCCTCTA
[1647] CCCTGGTGCAACGCCCTCGGCGGCCGCGCGCTACGCCTTCGGCTGCGCGC
GGCACCTCGGACCCCCGCTCGTGCTGACGCGCTCGCGCGCGTCAGACCAC
TTCGGGCTCGCGGGGGTCGGGAATTTTGCTAAACAGACTCCGAGTTGCCA
TTGGACACTGTAGCTGTGAATCAGTAACGAAAGTGAGTGGGGCCAGACTT
CGCCATAGGGCCTTTATCTTCTTGCCATTGGTCCGTGTAGGGGGTCGCCA
TAGGCTTCGACCTCCCTTTTAGGCCTTCCGGACTACAAAAATGGCGGATT
CAGTGACGTCACGGCCGCCATTTTAAGTAGGTGCCGTCCAGGACTGCAGT
TCCGGGTCAGAGTGCATCCTCGGCGGAACCTGCACAAAATGGCGGTCAAT
ATCTTCCGGGTCAAAGGTCACACCTACGTCATAAGTCACGTGACTGGGTC
CTGCTACGTCATATGCGGAAGTAGGCCCCGCCACGTGACTCGTCACGTGG
GCGCTGCGTCACGGCGGCCATTTTGTATCACAAAATGGCGGACTTCCTTC
CTCTTTTTTAAAAATAACGGCCCAGCGGCGGCGCGCGCTTCGCGCGCG
CGCCGGGGGGCTCCGCCCCCCCCCGCGCATGCGCGGGGCCCCCCCCGCG
GGGGGCTCCGCCCCCCGGTCCCCCCCCG (SEQ ID NO: 24)

[1648] 注释：

推定结构域	*碱基范围*
TATA 盒	82 - 87
起始元件	95-115
转录起始位点	115
5’ UTR 保守结构域	170 - 238
ORF2	335 - 721
[1649] ORF2/2	335 - 717; 2446 - 2902
ORF2/3	335 - 717; 2675 - 3109
ORF1	586 - 2928
ORF1/1	586 - 717; 2446 - 2928
ORF1/2	586 - 717; 2675 - 2902
三开放阅读框区	2640 - 2899
聚(A)信号	3106 - 3114
[1650] GC 富集区	3768 - 3878

[1651] 表8.示例性指环病毒氨基酸序列(甲型细环病毒,进化分枝4)

[1652]

TTV-HD20a (甲型细环病毒进化分枝 4)	
ORF2	MSRCWCPPRGDVSSRESRWYEAVRGSHDAFCGCSDPILHLSRLAARFNHQG PPTPPTDDRAPQNTPVRRLLPLPSYPGEGPQARWPGGDGGAAGGDRREGGD GGARAAEDEYQPEDLDELFGAIEQEQ (SEQ ID NO: 25)
ORF2/2	MSRCWCPPRGDVSSRESRWYEAVRGSHDAFCGCSDPILHLSRLAARFNHQG PPTPPTDDRAPQNTPVRRLLPLPSYPGEGPQARWPGGDGGAAGGDRREGGD GGARAAEDEYQPEDLDELFGAIEQEQSSETRAKTTKYPVPVDSLGIYKSLTR NASRRRGSSTASTSEEASLLKQLSGACSRNHYLASMLLAPSGHPSSFYPQNT SEKTALQKAPQVHRPKDPVSGQKRVRRRRSPRRRTRRRRRGSSSSESSESSE HSSSNSSTSRSNSPRPRRIST (SEQ ID NO: 26)
ORF2/3	MSRCWCPPRGDVSSRESRWYEAVRGSHDAFCGCSDPILHLSRLAARFNHQG PPTPPTDDRAPQNTPVRRLLPLPSYPGEGPQARWPGGDGGAAGGDRREGGD GGARAAEDEYQPEDLDELFGAIEQEQIPARRRRCRKRLRFTGQKTPYLVRRE SDGDDLLGGEPGGDDEGAPPAKAPRAASTPVPTPALRGPTRQDPGESPRKPP VIFPAMNKVFLFPPEGPKPILGKEAWQDEYETCRVWNRPARTHHTDTPFYP WAPHKFHVSFKLGFQ (SEQ ID NO: 27)
ORF1	MAWWGWRRRWWRPKRRWRWRRARRRRRVPARRPRRAFRRYRTRTVRR RRRGRRRGYRRRYRLRRYARRRFRRKKIVLTQWNPQTTRKCIIRGMMPVL WAGMGTGGRNYAVRSDDYVVNKGFGGSFATETFSLKVLYDQFQRGFNRW SHTNEDLDLARYRGCRWTFYRHKDTDFIVYFTNNPPMKTNQFSAPLTTPGM LMRSKYKVLIPSFQTRPKGRKTVTVKIRPPKLFQDKWYTQQDLCSVPLVQL NVTAADFTHPFGSPLTETPCVEFQVLGDLYNTCLNIDLPQFSELGEITSAYSK PNSNNLKELYKELFTKATSGHYWQTFITNSMVRAHIDADKAKEAQRASTTP SYNNDPFPTIPVKSEFAQWKKKFTDTRDSPFLFATYHPEAIKDTIMKMRENN FKLETGPNDKYGDYTAQYQGNTHMLDYYLGFYSPIFLSDGRSNVEFFTAYR DIVYNPFLDKAQGNMVWFQYHTKTDNKFKKPECHWEIKDMPLWALLNGY VDYLETQIQYGDLSKEGKVLIRCPYTKPALVDPRDDTAGYVVYNRNFGRGK WIDGGGYIPLHERTKWYVMLRYQTDVFHDIVTCGPWQYRDDNKNSQLVA KYRFSFIWGGNTVHSQVIRNPCKDNQVSGPRRQPRDIQVVDPQRITPPWVLH

[1653]

	SFDQRRGLFTETALRRLLQEPLPGEYAVSTLRTPLLFLPSEYQREDGAAESAS GSPAKRPRIWSEESQTETISSEENPAETTRELLQRKLREQRALQFQLQHFAVQ LAKTQANLHVNPLLSFPQ (SEQ ID NO: 28)
ORF1/1	MAWWGWRRRWWRPKRRWRWRRARRRRRVPARRPRRAFRRYRTRTVIRN PCKDNQVSGPRRQPRDIQVVDPQRITPPWVLHSFDQRRGLFTETALRRLLQE PLPGEYAVSTLRTPLLFLPSEYQREDGAAESASGSPAKRPRIWSEESQTETISS EENPAETTRELLQRKLREQRALQFQLQHFAVQLAKTQANLHVNPLLSFPQ (SEQ ID NO: 29)
ORF1/2	MAWWGWRRRWWRPKRRWRWRRARRRRRVPARRPRRAFRRYRTRTNTSE KTALQKAPQVHRPKDPVSGQKRVRRRRSPRRRTRRRRRGSSSSESSESSEHS SSNSSTSRSNSPRPRRIST (SEQ ID NO: 30)

[1654] 表9.示例性指环病毒核酸序列(甲型细环病毒,进化分枝5)

[1655] 名称 TTV-16(TUS01)

[1656] 属/进化分枝 甲型细环病毒,进化分枝5

[1657] 登录号 AB017613.1

[1658] 全序列:3818bp

[1659]

```
1          10         20         30         40         50
|          |          |          |          |          |
AAGTCCGCCACTAACCACGTGACTCCCGCAGGCCAACCCAGTACTATGTC
GTCCACTTCCTGGGACGAGTCTACGTCCTGATATAAGTAAGTGCACTTCC
GAATGGCTGAGTTTTCCACGCCCGTCCGCAGCGAGAACGCCACGGAGGGG
AGTCCGCGCGTCCCGAGGGCGGGTGCCGGAGGTGAGTTTACACACCGCAG
TCAAGGGGCAATTCGGGCTCGGGACTGGCCGGGCCCCGGGCAAGGCTCTT
AAAAAATGCACTTTCGCAGAGTGCGAGCGAAAAGGAAACTGCTACTGCAA
GCTGTGCGAGCTCCACCGAAGGCACCTGCCATGAGCTTCACCACACCTAC
TATTAATGCCGGGATCCGAGAGCAGCAATGGTTCGAGTCCACCCTTAGAT
CCCACCACTCGTTCTGTGGCTGTGGTGATCCCGTGCTTCATTTTACTAAC
CTTGCTACTCGCTTTAACTATCTGCCTGCTACCTCTTCGCCTCTGGACCC
TCCCGGCCCAGCGCCGCGAGGCCGCCCGGCGCTCCGCCGCCTCCCGGCAC
TCCCTTCAGCCCCCGCGACCCCTTCTAGAGAACTAGCATGGCCTACTGGT
TCAGAAGGTGGGGCTGGAGGCCGAGGCGCCGGTGGAGAAGGTGGCGCCGC
CGTCGAAGGAGACTACCGAGAAGAAGAACTAGACGAGCTGTTCGCGGCCT
TGGAAGAAGACGCAAACCAAGGGTAAGGAGGCGCCGCAGAACTCGCAGAC
GTACCTACAGACGGGGGTGGAGACGCAGGAGGTACATAAGACGGGGGCGA
CGCAAAAAGAAACTCATACTGACTCAGTGGAACCCGGCAATAGTTAAGAG
GTGCAACATTAAGGGCGGACTTCCAATAATTATATGCGGAGAGCCCAGGG
CAGCCTTTAACTATGGCTACCACATGGAGGACTACACTCCTCAACCTTTC
```

CCCTTCGGAGGGGGAATGAGCACAGTGACTTTCTCTCTGAAAGCCTTGTA
TGACCAGTACCTAAAACACCAAAACAGGTGGACTTTCTCAAACGACCAGC
TAGACCTCGCCAGATACAGGGGCTGTAAACTAAGGTTCTACAGAAGCCCC
GTCTGTGACTTTATAGTACACTACAACCTAATACCTCCACTAAAAATGAA
CCAGTTCACAAGTCCCAACACGCACCCGGGACTACTCATGCTCAGCAAAC
ACAAGATAATAATTCCCAGCTTTCAAACAAGACCTGGGGGCAGACGCTTT
GTTAAAATAAGACTTAATCCCCCCAAACTATTTGAAGACAAGTGGTACAC
TCAGCAAGACCTGTGCAAGGTTCCGCTCGTTAGTATTACAGCAACTGCGG
CTGACTTGCGGTATCCGTTCTGCTCACCACAAACGAACAACCCTTGCACC
ACCTTCCAGGTACTGCGCAAGAACTACAATACAGTTATAGGAACTTCCGT
AAAAGACCAAGAGTCCACACAAGACTTTGAAAATTGGCTTTATAAAACAG
ACTCACACTATCAAACATTTGCCACAGAGGCTCAACTAGGCAGAATTCCT
GCATTTAATCCTGATGGCACTAAAAACACTAAACAGCAGTCGTGGCAAGA
TAACTGGAGCAAAAAAAATTCACCATGGACAGGTAACTCAGGTACATACC
CACAAACAACCAGTGAAATGTACAAAATTCCATATGACAGTAACTTCGGC
TTTCCCACATACAGAGCCCAAAAAGACTACATTTTAGAAAGAAGACAGTG
CAACTTTAACTATGAAGTTAATAATCCAGTTAGCAAAAAAGTATGGCCAC
AACCTAGTACAACAACACCCACAGTAGACTACTATGAATACCACTGTGGA
TGGTTCAGCAACATATTCATAGGCCCCAACAGATACAACCTACAGTTTCA
AACAGCATATGTAGACACCACATACAACCCACTAATGGACAAGGGCAAAG
GCAACAAAATATGGTTTCAATATCTGTCTAAAAAGGGCACAGACTACAAT
GAAAAACAATGCTACTGCACCCTAGAAGACATGCCCCTATGGGCAATATG
CTTTGGATACACTGACTATGTAGAGACTCAACTAGGACCCAATGTGGACC
ATGAAACAGCAGGCTTAATAATTATGATCTGTCCATACACTCAACCACCT
ATGTATGACAAAAACAGACCTAACTGGGGATACGTAGTCTATGACACAAA
CTTTGGCAATGGAAAAATGCCCTCAGGAAGTGGCCAAGTCCCAGTATACT
GGCAATGCCGATGGAGGCCCATGCTGTGGTTCCAACAACAAGTACTCAAT
GACATCTCAAAGACTGGACCGTACGCCTACAGAGACGAATATAAAAATGT
ACAACTGACTCTCTACTACAACTTTATTTTTAACTGGGGGGGCGACATGT
[1660] ATTACCCACAGGTCGTTAAAAACCCCTGTGGAGACTCCGGAATCGTTCCC
GGTTCCGGTAGATTCACTCGAGAAGTACAAGTCGTTAGCCCGCTTTCCAT
GGGACCGGCCTACATCTTCCACTACTTCGACTCCAGACGCGGGTTCTTTA
GTGAAAAAGCTCTTAAAAGAATGCAACAACAACAAGAATTTGATGAATCT
TTTACATTCAAACCTAAGAGACCCAAACTTTCTACAGCAGCCGCAGAAAT
CCTCCAGCTCGAAGAAGACTCGACTTCAGGGGAAGGAAAATCGCCACTAC
AGCAAGAAGAGAAAGAAGTCGAAGTCCTCCAAACGCCGACAGTACAGCTC
CAGCTCCAGCGAAACATCCAGGAGCAGCTCGCAATCAAGCAGCAGCTCCA
ATTCCTCTTGCTCCAACTCCTCAAAACCCAATCCAATTTGCATTTAAACC
CACAATTTTTAAGCCCTTCATAAAATATGACATGTTTGGGGACCCCCTTC
CTCACCCCCCAACAGCCGAAGAGTGGGAAACAGAGTACCAGTGCTGTAAG
GCCTTTAACAGACCACCTAGAACCAACCTAAAAGACACCCCCTTCTACCC
CTGGGTACCTAAACCTAAACCTCAATTCCGTGTATCTTTTAAACTTGGTT
TTCAATAAACAAGGCCGTGGGAGTTTCACTTGTCGGTGTCAACCTCTTAA
GGTCACTAAGCACTCCGAGCGTAAGCGAGGAGTGCGACCCTCCCCCCTGG
GGCAACTCCCTCGAAGTCCGGCGCTACGCGCTTCGCGCTGCGCCGGACAT
CTCGGACCCCCCCTCCACCCGAAACGCTTGCGCGTTTCGGACCTTCGGCG
TCGGGGGGGTCGGGGGCTTTACTAAACAGACTCCGAGGTGCCATTGGACA
CTGAGGGGATGAACAGCAACGAAAGTGAGTGGGGCCAGACTTCGCCATAA
GGCCTTTATCTTCTTGCCATTTGTCAGTATAGAGGGTCGCCATAGGCTTC
GGCCTCCATTTTAACCTCTAAAAACTACCAAAATGGCCGTTCCAGTGACG
TCACAGCCGCCATTTTAAGTAGCTGACGTCAAGGATTGACGTGAAGGTTA
AAGGTCATCCTCGGCGGAAGCTACACAAAATGGTGGACAACATCTTCCGG
GTCAAAGGTCGTGCACACGTCATAAGTCACGTGGTGGGGACCCGCTGTAA
CCCGGAAGTAGGCCCCGTCACGTGATTTGTCACGTGTGTACACGTCACAA
CCGCCATTTTGTTTTACAAAATGGCTGACTTCCTTCCTCTTTTTTAAAAA
AAACGGCCGTGCGGCGGCGCGCGCGCTTCGCGCGCGCGCCGGGGGCTGCC
GCCCCCCCCCGCGCATGCGCGCGGGGCCCCCCCCCGCGGGGGGCTCCGCC

CCCCGGCCCCCCCCCCCG (SEQ ID NO: 31)

[1661] 注释：

推定结构域	*碱基范围*
TATA 盒	82 - 86
起始元件	100-115
转录起始位点	115
5' UTR 保守结构域	170 - 240
ORF2	331- 726
ORF2/2	331 - 722; 2412 - 2847
ORF2/3	331 - 722; 2638 - 3058
ORF2t/3	331 - 380; 2638 - 3058
ORF1	588 - 2873
ORF1/1	588 - 722; 2412 - 2873
ORF1/2	588 - 722; 2638 - 2847
三开放阅读框区	2699 - 2969
聚(A)信号	3220 - 3225
GC 富集区	3302 - 3541

[1662]

[1663] 表10.示例性指环病毒氨基酸序列(甲型细环病毒,进化分枝5)

TTV-16-TUS01 (甲型细环病毒进化分枝 5)	
ORF2	MSFTTPTINAGIREQQWFESTLRSHHSFCGCGDPVLHFTNLATRFNYLPATSS PLDPPGPAPRGRPALRRLPALPSAPATPSRELAWPTGSEGGAGGRGAGGEGG AAVEGDYREEELDELFAALEEDANQG (SEQ ID NO: 32)
ORF2/2	MSFTTPTINAGIREQQWFESTLRSHHSFCGCGDPVLHFTNLATRFNYLPATSS PLDPPGPAPRGRPALRRLPALPSAPATPSRELAWPTGSEGGAGGRGAGGEGG AAVEGDYREEELDELFAALEEDANQGSLKTPVETPESFPVPVDSLEKYKSLA RFPWDRPTSSTTSTPDAGSLVKKLLKECNNNKNLMNLLHSNLRDPNFLQQP QKSSSSKKTRLQGKENRHYSKKRKKSKSSKRRQYSSSSSETSRSSSQSSSSSN SSCSNSSKPNPICI (SEQ ID NO: 33)
ORF2/3	MSFTTPTINAGIREQQWFESTLRSHHSFCGCGDPVLHFTNLATRFNYLPATSS PLDPPGPAPRGRPALRRLPALPSAPATPSRELAWPTGSEGGAGGRGAGGEGG AAVEGDYREEELDELFAALEEDANQGSRRNPPARRRLDFRGRKIATTARRE

[1664]

[1665]

	RSRSPPNADSTAPAPAKHPGAARNQAAAPIPLAPTPQNPIQFAFKPTIFKPFIK YDMFGDPLPHPPTAEEWETEYQCCKAFNRPPRTNLKDTPFYPWVPKPKPQF RVSFKLGFQ (SEQ ID NO: 34)
ORF2t/3	MSFTTPTINAGIREQQCSRRNPPARRRLDFRGRKIATTARRERSRSPPNADST APAPAKHPGAARNQAAAPIPLAPTPQNPIQFAFKPTIFKPFIKYDMFGDPLPH PPTAEEWETEYQCCKAFNRPPRTNLKDTPFYPWVPKPKPQFRVSFKLGFQ (SEQ ID NO: 35)
ORF1	MAYWFRRWGWRPRRRWRRWRRRRRRLPRRRTRRAVRGLGRRRKPRVRR RRRTRRRTYRRGWRRRRYIRRGRRKKKLILTQWNPAIVKRCNIKGGLPIIICG EPRAAFNYGYHMEDYTPQPFPFGGGMSTVTFSLKALYDQYLKHQNRWTFS NDQLDLARYRGCKLRFYRSPVCDFIVHYNLIPPLKMNQFTSPNTHPGLLMLS KHKIIIPSFQTRPGGRRFVKIRLNPPKLFEDKWYTQQDLCKVPLVSITATAAD LRYPFCSPQTNNPCTTFQVLRKNYNTVIGTSVKDQESTQDFENWLYKTDSH YQTFATEAQLGRIPAFNPDGTKNTKQQSWQDNWSKKNSPWTGNSGTYPQT TSEMYKIPYDSNFGFPTYRAQKDYILERRQCNFNYEVNNPVSKKVWPQPST TTPTVDYYEYHCGWFSNIFIGPNRYNLQFQTAYVDTTYNPLMDKGKGNKIW FQYLSKKGTDYNEKQCYCTLEDMPLWAICFGYTDYVETQLGPNVDHETAG LIIMICPYTQPPMYDKNRPNWGYVVYDTNFGNGKMPSGSGQVPVYWQCR WRPMLWFQQQVLNDISKTGPYAYRDEYKNVQLTLYYNFIFNWGGDMYYP QVVKNPCGDSGIVPGSGRFTREVQVVSPLSMGPAYIFHYFDSRRGFFSEKAL KRMQQQQEFDESFTFKPKRPKLSTAAAEILQLEEDSTSGEGKSPLQQEEKEV EVLQTPTVQLQLQRNIQEQLAIKQQLQFLLLQLLKTQSNLHLNPQFLSPS (SEQ ID NO: 36)
ORF1/1	MAYWFRRWGWRPRRRWRRWRRRRRRLPRRRTRRAVRGLGRRRKPRVVK NPCGDSGIVPGSGRFTREVQVVSPLSMGPAYIFHYFDSRRGFFSEKALKRMQ QQQEFDESFTFKPKRPKLSTAAAEILQLEEDSTSGEGKSPLQQEEKEVEVLQT PTVQLQLQRNIQEQLAIKQQLQFLLLQLLKTQSNLHLNPQFLSPS (SEQ ID NO: 37)

ORF1/2	MAYWFRRWGWRPRRRWRRWRRRRRRLPRRRTRRAVRGLGRRRKPRQPQ KSSSSKKTRLQGKENRHYSKKRKKSKSSKRRQYSSSSSETSRSSSQSSSSSNS SCSNSSKPNPICI (SEQ ID NO: 38)

[1666] 表11.示例性指环病毒核酸序列(甲型细环病毒,进化分枝6)

[1667] 名称 TTV-TJN02

[1668] 属/进化分枝 甲型细环病毒,进化分枝6

[1669] 登录号 AB028669.1

[1670] 全序列:3794bp

[1671]

```
1         10         20         30         40         50
|         |          |          |          |          |
CCCGAAGTCCGTCACTAACCACGTGACTCCTGTCGCCCAATCAGAGTGTA
TGTCGTGCATTTCCTGGGCATGGTCTACATCCTGATATAACTAAGTGCAC
TTCCGAATGGCTGAGTTTTCCACGCCCGTCCGCAGCGAGGGAGCGACGGA
GGAGCTCCCGAGCGTCCCGAGGGCGGGTGCCGGAGGTGAGTTTACACACC
GCAGTCAAGGGGCAATTCGGGCTCGGGACTGGCCGGGCTATGGGCAAGGC
TCTTAGGGTCTTCATTCTTAATATGTTTCTTGGCAGAGTTTACCGCCACA
AGAAAAGGAAAGTGCTACTGTCCACACTGCGAGCTCCACAGGCGTCTCGC
AGGGCTATGAGTTGGCGACCCCCGGTACACGATGCACCCGGCATCGAGCG
CAATTGGTACGAGGCCTGTTTCAGAGCCCACGCTGGAGCTTGTGGCTGTG
GCAATTTTATTATGCACCTTAATCTTTTGGCTGGGCGTTATGGTTTTACT
CCGGGGTCAGCGCCGCCAGGTGGTCCTCCTCCGGGCACCCCGCAGATAAG
GAGAGCCAGGCCTAGTCCCGCCGCACCAGAGCAGCCCGCTGCCCTACCAT
GGCATGGGGATGGTGGAGATGGCGGCGCCGCTGGCCCGCCAGACGCTGGA
GGAGACGCCGTCGCCGGCGCCCCGTACGGAGAACAAGAGCTCGCCGACCT
GCTCGACGCTATAGAAGACGACGAACAGTAAGAACCAGGCGAAGGCGGTG
GGGGCGCAGACGGTACAGACGGGGCTGGAGACGCAGGACTTATGTGAGAA
AGGGGCGACACAGAAAAAAGAAAAAGAGACTGATACTGAGACAGTGGCAA
CCAGCCACAAGACGCAGATGTACCATAACTGGGTACCTGCCCATAGTGTT
CTGCGGCCACACTAGGGGCAATAAAAACTATGCACTACACTCTGACGACT
ACACCCCCCAAGGACAACCATTTGGAGGGGCTCTAAGCACTACCTCATTC
TCTTTAAAAGTACTATTTGACCAGCATCAGAGAGGACTAAACAAGTGGTC
TTTTCCAAACGACCAACTAGACCTCGCCAGATATAGAGGCTGCAAATTTA
TATTTTATAGAACAAAACAAACTGACTGGGTGGGCCAGTATGACATATCA
GAACCCTACAAGCTAGACAAATACAGCTGCCCCAACTATCACCCTGGAAA
CATGATTAAGGCAAAGCACAAATTTTTAATACCAAGCTATGACACTAATC
CTAGAGGCAGACAAAAAATTATAGTTAAAATTCCCCCCCCAGACCTCTTT
GTAGACAAGTGGTACACTCAAGAGGATCTGTGTTCCGTTAATCTTGTGTC
ACTTGCGGTTTCTGCGGCTTCCTTTCTCCACCCATTCGGCTCACCACAAA
CTGACAACCCTTGCTACACCTTCCAGGTGTTGAAAGAGTTCTACTATCAG
GCAATAGGCTTCTCTGCAAGCACACAAGCAATGACATCAGTATTAGACAC
GCTATACACACAAAACAGTTATTGGGAATCTAATCTAACTCAGTTTTATG
TACTTAATGCAAAAAAAGGCAGTGATACAACACAGCCTTTAACTAGCAAT
ATGCCAACTCGTGAAGAGTTTATGGCAAAAAAAAATACCAATTACAACTG
GTATACATACAAGGCCGCGTCAGTAAAAAATAAACTACATCAAATGAGAC
```

[1672]

AAACCTATTTTGAGGAGTTAACCTCTAAGGGGCCACAAACAACAAAAAGT
GAGGAAGGCTACAGTCAGCACTGGACCACCCCCTCCACAAACGCCTACGA
ATATCACTTAGGAATGTTTAGTGCAATATTTCTAGCCCCAGACAGGCCAG
TACCTAGATTTCCATGCGCCTACCAAGATGTAACTTACAACCCCTTAATG
GACAAAGGGGTGGGAAACCACATTTGGTTTCAGTACAACACAAAGGCAGA
CACTCAGCTAATAGTCACAGGAGGGTCCTGCAAAGCACACATACAAGACA
TACCACTGTGGGCGGCCTTCTATGGATACAGTGACTTTATAGAGTCAGAA
CTAGGCCCCTTTGTAGATGCAGAGACGGTAGGCTTAGTGTGTGTAATATG
CCCTTATACAAAACCCCCCATGTACAACAAGACAAACCCCGCCATGGGCT
ACGTGTTCTATGACAGAAACTTTGGTGACGGAAAATGGACTGACGGACGG
GGCAAAATAGAGCCCTACTGGCAAGTTAGGTGGAGGCCCGAAATGCTTTT
CCAAGAAACTGTAATGGCAGACCTAGTTCAGACTGGGCCCTTTAGCTACA
AAGACGAACTTAAAAACAGCACCCTAGTGTGCAAGTACAAATTCTATTTC
ACCTGGGGAGGTAACATGATGTTCCAACAGACGATCAAAAACCCGTGCAA
GACGGACGGACAACCCACCGACTCCAGTAGACACCCTAGAGGAATACAAG
TGGCGGACCCGGAACAAATGGGACCCCGCTGGGTGTTCCACTCCTTTGAC
TGGCGAAGGGGCTATCTTAGCGAGAAAGCTCTCAAACGCCTGCAAGAAAA
ACCTCTTGACTATGACGAATATTTTACACAACCAAAAAGACCTAGAATCT
TTCCTCCAACAGAATCAGCAGAGGGAGAGTTCCGAGAGCCCGAAAAAGGC
TCGTATTCAGAGGAAGAAAGGTCGCAAGCCTCTGCCGAAGAGCAGACGCA
GGAGGCGACAGTACTCCTCCTCAAGCGACGACTCAGAGAGCAACAGCAGC
TCCAGCAGCAGCTCCAATTCCTCACCCGAGAAATGTTCAAAACGCAAGCG
GGTCTCCACCTAAACCCTATGTTATTAAACCAGCGATAAACCAAGTGTAC
CTGTTTCCAGAGAGGGCCCCAAAACCCCCTCCTAGCAGCCAAGACTGGCA
GCAGGAGTACGAGGCCTGCGCAGCCTGGGACAGGCCCCCTAGATACAATC
TGTCCTCTCCTCCTTTCTACCCCAGCTGCCCTTCAAAATTCTGTGTAAAA
TTCAGCCTTGGCTTTAAATAAATGGCAACTTTACTGTGCAAGGCCGTGGG
AGTTTCACTGGTCGGTGTCTACCTCTAAAGGTCACTAAGCACTCCGAGCG
TTAGCGAGGAGTGCGACCCTTCCCCCTGACTCAACTTCTTCGGAGCCGCG
CGCTACGCCTTCGGCTGCGCGCGGCACCTCAGACCCCCGCTCGTGCTGAC
ACGCTCGCGCGTGTCAGACCACTTCGGGCTCGCGGGGGTCGGGAATTTTG
CTAAACAGACTCCGAGTTGCTCTTGGACACTGAGGGGGCATATCAGTAAC
GAAAGTGAGTGGGGCCAGACTTCGCCATAAGGCCTTTATCTTCTTGCCAT
TGGATAGTATCGAGGGTTGCCATAGGCTTCGACCTCCATTTTAGGCCTTC
CGGACTACAAAAATGGCCGTTTTAGTGACGTCACGGCCGCCATTTTAAGT
AAGGCGGAAGCAGCTCGGCGTACACAAAATGGCGGCGGAGCACTTCCGGC
TTGCCCAAAATGGTGGGCAACTTCTTCCGGGTCAAAGGTCACAGCTACGT
CACAAGTCACGTGGGGAGGGTTGGCGTTTAACCCGGAAGCCAATCCTCTT
ACGTGGCCTGTCACGTGACTTGTACGTCACGACCACCATTTTGTTTTACA
AAATGGCCGACTTCCTTCCTCTTTTTTAAAAATAACGGTTCGGCGGCGGC
GCGCGCGCTACGCGCGCGCGCCGGGGGGCTGCCGCCCCCCCCCCGCGCAT
GCGCGGGGCCCCCCCCCGCGGGGGGCTCCGCCCCCGGCCCCCC (SEQ ID NO: 39)

[1673] 注释：

推定结构域	*碱基范围*

[1674]

TATA 盒	89 - 90
帽位点	107 - 114

[1675]

转录起始位点	114
5' UTR 保守结构域	174 - 244
ORF2	357 - 731
ORF2/2	357 - 727; 2381 - 2813
ORF2/3	357 - 727; 2619 - 3021
ORF2t/3	357 - 406; 2619 - 3021
ORF1	599 - 2839
ORF1/1	599 - 727; 2381 - 2839
ORF1/2	599 - 727; 2619 - 2813
三开放阅读框区	2596 - 2810
聚(A)信号	3017 - 3022
GC 富集区	3691 - 3794

[1676] 表12.示例性指环病毒氨基酸序列(甲型细环病毒,进化分枝6)

[1677]

TTV-TJN02 (甲型细环病毒进化分枝 6)	
ORF2	MSWRPPVHDAPGIERNWYEACFRAHAGACGCGNFIMHLNLLAGRYGFTPGSAPPG GPPPGTPQIRRARPSPAAPEQPAALPWHGDGGDGGAAGPPDAGGDAVAGAPYGEQ ELADLLDAIEDDEQ (SEQ ID NO: 40)
ORF2/2	MSWRPPVHDAPGIERNWYEACFRAHAGACGCGNFIMHLNLLAGRYGFTPGSAPPG GPPPGTPQIRRARPSPAAPEQPAALPWHGDGGDGGAAGPPDAGGDAVAGAPYGEQ ELADLLDAIEDDEQRSKTRARRTDNPPTPVDTLEEYKWRTRNKWDPAGCSTPLTGE GAILARKLSNACKKNLLTMTNILHNQKDLESFLQQNQQRESSESPKKARIQRKKGR KPLPKSRRRRRQYSSSSDDSESNSSSSSSSNSSPEKCSKRKRVST (SEQ ID NO: 41)
ORF2/3	MSWRPPVHDAPGIERNWYEACFRAHAGACGCGNFIMHLNLLAGRYGFTPGSAPPG GPPPGTPQIRRARPSPAAPEQPAALPWHGDGGDGGAAGPPDAGGDAVAGAPYGEQ ELADLLDAIEDDEHRGRVPRARKRLVFRGRKVASLCRRADAGGDSTPPQATTQRAT AAPAAAPIPHPRNVQNASGSPPKPYVIKPAINQVYLFPERAPKPPPSSQDWQQEYEA CAAWDRPPRYNLSSPPFYPSCPSKFCVKFSLGFK (SEQ ID NO: 42)
ORF2t/3	MSWRPPVHDAPGIERNCRGRVPRARKRLVFRGRKVASLCRRADAGGDSTPPQATT QRATAAPAAAPIPHPRNVQNASGSPPKPYVIKPAINQVYLFPERAPKPPPSSQDWQQ EYEACAAWDRPPRYNLSSPPFYPSCPSKFCVKFSLGFK (SEQ ID NO: 43)
ORF1	MAWGWWRWRRRWPARRWRRRRRRPVRRTRARRPARRYRRRRTVRTRRRRWG RRRYRRGWRRRTYVRKGRHRKKKRLILRQWQPATRRRCTITGYLPIVFCGHTRG NKNYALHSDDYTPQGQPFGGALSTTSFSLKVLFDQHQRGLNKWSFPNDQLDLARY

[1678]

	RGCKFIFYRTKQTDWVGQYDISEPYKLDKYSCPNYHPGNMIKAKHKFLIPSYDTNP RGRQKIIVKIPPPDLFVDKWYTQEDLCSVNLVSLAVSAASFLHPFGSPQTDNPCYTF QVLKEFYYQAIGFSASTQAMTSVLDTLYTQNSYWESNLTQFYVLNAKKGSDTTQPL TSNMPTREEFMAKKNTNYNWYTYKAASVKNKLHQMRQTYFEELTSKGPQTTKSE EGYSQHWTTPSTNAYEYHLGMFSAIFLAPDRPVPRFPCAYQDVTYNPLMDKGVGN HIWFQYNTKADTQLIVTGGSCKAHIQDIPLWAAFYGYSDFIESELGPFVDAETVGLV CVICPYTKPPMYNKTNPAMGYVFYDRNFGDGKWTDGRGKIEPYWQVRWRPEMLF QETVMADLVQTGPFSYKDELKNSTLVCKYKFYFTWGGNMMFQQTIKNPCKTDGQ PTDSSRHPRGIQVADPEQMGPRWVFHSFDWRRGYLSEKALKRLQEKPLDYDEYFT QPKRPRIFPPTESAEGEFREPEKGSYSEEERSQASAEEQTQEATVLLLKRRLREQQQL QQQLQFLTREMFKTQAGLHLNPMLLNQR (SEQ ID NO: 44)
ORF1/1	MAWGWWRWRRRWPARRWRRRRRRRPVRRTRARRPARRYRRRRTTIKNPCKTDG QPTDSSRHPRGIQVADPEQMGPRWVFHSFDWRRGYLSEKALKRLQEKPLDYDEYF TQPKRPRIFPPTESAEGEFREPEKGSYSEEERSQASAEEQTQEATVLLLKRRLREQQ LQQQLQFLTREMFKTQAGLHLNPMLLNQR (SEQ ID NO: 45)
ORF1/2	MAWGWWRWRRRWPARRWRRRRRRRPVRRTRARRPARRYRRRRTQRESSESPKK ARIQRKKGRKPLPKSRRRRRQYSSSSDDSESNSSSSSSSNSSPEKCSKRKRVST (SEQ ID NO: 46)

[1679] 表13.示例性指环病毒核酸序列(甲型细环病毒,进化分枝7)

[1680] 名称 TTV-HD16d

[1681] 属/进化分枝 甲型细环病毒,进化分枝7

[1682] 登录号 FR751479.1

[1683] 全序列:3866bp

[1684]
```
1          10         20         30         40         50
|          |          |          |          |          |
AAGTCCGTCACTAACCACGTGACTCCCGCAGGCCAATCAGAGTCTATGTC
GTGCACTTCCTGGGCATGGTCTACGTTCTCATATAACTAACTGCACTTCC
GAATGGCTGAGTTTTCCACGCCCGTCCGCAGCGGCAGCACCACGGAGGGT
GATCCCCGCGTCCCGAGGGCGGGTGCCGAAGGTGAGTTTACACACCGCAG
TCAAGGGGCAATTCGGGCTCGGGACTGGCCGGGCTATGGGCAAGGCTCTT
AGGGCTTTCATTGTTAAAAATGTTTCTCGGCAGGCCTTACAGGAGAAAGA
AAAGGGCGCTGTCACTGCCTGGCGTGCGAGCTGCACAGGCGAAACAACCT
GGTGATATGAGCTGGAGCCGTCCAGTACATAATGCCGCCGGGATCGAAAG
GCAGTGGTTCGAATCCACCTTTAGATCCCACGCTAGTTGCTGTGGCTGCG
GCAATTTTGTTAATCATATTAATGTACTGGCTGCTCGCTACGGCTTTACT
GGGGGGCCGACGCCGCCAGGTGGTCCTGGGCCGCGTCCACAACTGAGGCC
```

CGCGCTTCCCGCGCCGGACCCCGACCCCCAGGCGCCCAACCGTGAGCCAT
GGCGTGGAGCTGGTGGTGGCAACGATGGAGAAGGCGCCGCTGGAAACCCA
GGAGGCGCCGCTGGAGACGTCTACGATGGAGAAGACCTAGACGCGCTGTT
CGCCGCCGTCGTCGAGGACGTAGAGTAAGGAGGCGGAGGTGGGCGCGTAG
ACGGGGGCGACGCAGACGGTACGCCACCAGACGAAAGAGACGTTATAGGG
GTCGCCGCTTTAAAAAGAAACTAGTACTGACTCAGTGGCACCCTAATACC
ATGAGACGCTGCTTAATCAAGGGCATAGTCCCCCTGGTAATATGCGGCCA
CACCAGGTGGAACTACAACTACGCCCTCCATAGCAAGGACTACACAGAGG
AGGGTCGCTACCCTCACGGGGGGGCCCTCAGCACCACTACGTGGTCCCTT
AAGGTGCTGTATGACGAGCACCTCAAACACCACGACTTCTGGGGCTATCC
CAACAACCAGCTAGACCTGGCCAGGTACAAGGGGGCCAAGTTCACCTTCT
ACAGACACAAAAAGACTGACTTTATAATATTCTTTAACAGAAAGCCTCCC
TTTAAGCTAAACAAGTACAGCTGTGCCTCCTATCACCCAGGCATGCTGAT
GCAGCAGAGACACAAGATCCTGCTACCCAGCTACGAAACTAAACCCAAGG
GCAGGCCAAAGATAACAGTTAGAATAAAGCCCCCCACTCTGTTAGAGGAC
AAGTGGTACACCCAGCAGGACCTGTGCGACGTTAACCTGTTGCAACTTGT
GGTCACTGCGGCTGACTTTCGACATCCACTCTGCTCACCACAAACGAACA
CTCCAACCACAACCTTCCAGGTGTTGAAAGACATCTATTATGACACTATG
AGCATATCTGAACCCACAGACTCCTACACTAGTGTTAACAATAAAAGTAC
AACACAAACTTTTACTAACTACTCAAACACCTTAGAAAACATTCTGTACA
CACGAGCCTCCTACTGGAACTCGTTCCACGCCACTGAATACCTAAACCCC
AACATCATATACAAAAACGGTGAAAAACTATTCAAAGAACATGAAGACTT
AATAACCTGGATGACCCAAACTAACAATACCGGGTTTCTAACTAAAAACA
ACACAGCTTTTGGCAACAACAGCTACAGGCCCAATGCAGACAAAATTAAA
AAAGCCAGAAAGACATACTGGAACGCCCTAATAGGCACCAACGACCTGGC
CACTAATATAGGCCAGGCCAGAGCAGAAAGGTTCGAGTACCACCTAGGCT
GGTACTCCCCCATATTTCTCAGCAGACACAGGAGCAACATGAACTTTGCC
AGGGCCTACCAAGACGTCACATACAACCCCAACTGTGACAGGGGAGTTAA
[1685] CAACAGGGTGTGGGTTCAGCCTCTAACTAAACCCACCACAGAGTTCGACG
AGAAAAGGTGTAAGTGCGTAGTGCAGCACCTGCCTCTGTGGGCGGCTCTG
TACTGCTACCAAGACTTTGTAGAGGAGGAGCTGGGGTCCTCCTCAGAGAT
ATTAAATTCATGCCTACTGGTATTACAGTGCCCTTACACCTTTCCCCCAA
TGTATGACAAAAAGCTACCAGACAAGGGATTCGTGTTTTATGACTCCCTT
TTTGGAGACGGCAAAATGTCTGACGGACGCGGACAGGTGGACATTTTCTG
GCAACAGCGATGGTACCCTCGCTTAGCCACTCAGATGCAAGTCATGCACG
ACATCACCATGACGGGCCCCTTCTCCTACCGAGACGAGCTAGTTAGCACC
CAACTGACTGCCAAGTACACCTTTGACTTTATGTGGGGCGGAAATATGAT
CTCCACACAGATCATCAAGAACCCCTGCAAAGACAGTGGACTGGAACCCG
CCTACCCCGGTAGACAGCGTCGCGACTTACAAATTGTTGACCCATACTCC
ATGGGCCCCCAATTCTCGTTCCACAACTGGGACTACAGACATGGCCTTTT
TGGCCAAGACGCTATCGACAGAGTGTCTAAACAACCAAAAGATGATGCAG
ACTATCCTAACCCATACAAAAGGCCTAGATATTTTCCACCCACAGACCAA
GCCGCCCAAGAGCAAGAAAAAGACTTCAGTTTCCTCAAAACAGCACCGTC
GAACTCAGAAGAGAGCGATCAAGAAGTCCTCCAAGAAACGCAAGTACTCC
GATTCCAGCCAGAGCAGCACAAGCAACTCCACCTGCAGCTCGCAGAGCGG
CAGCGAATCGGAGAGCAACTCCGATACCTACTCCAACAGATGTTCAAAAC
TCAGGCCAATCTCCACCTAAACCCATATACATTTACCCAGCTGTAAAGCA
GGTGTTTATGTTTGACCCCCCGGGCCCTAAGGCTATCTCGGGCGCCAAGG
CCTGGGAGGACGAGTTCCTCACCGCAAAAGTGTGGAACCGCCCGGTACGC
AAGTACTACTCAGACACCCCCTACTACCCCTGGGCCCCCAAACCCCAGTA
CTCTGTCAGTTTCAAACTCGGCTGGAAATAAAAAAAGCCTGCTCCACTGT
ACTAGGCCGTGGGAGTTTCACTCGTCGGTGTCTACCTCTTAAGGTCACCA
AGCACTCCGAGCGTCAGCGAGGAGTGCGACCCTTGGGGGTGGGTGCAACG
CCCTCGGCGGCCGCGCGCTACGCCTTCGGCTGCGCGCGGCACCTCGGACC
CCCGCTCGTGCTGACGCGCTTGCGCGCGTCAGACCACTTCGGGCTCGCGG
GGGTCGGAAATTTTGCTAAACAGACTCCGAGTTGCCATTGGACACTGGAG
CCGTGAATCAGTAACGAAAGTGAGTGGGGCCAGACTTCGCCATAAGGCCT

[1686]

TTATCTTTTTGCCATTTGTCCGTGGGGAAGGGTCGCTGCAAGCGCGGACC
CCGTTTTCACCCCTTCCGGACTACAAAAATAGCGCATTAGTGACGTCACG
GCCGCCATTTTAAGTAAGGCGGAAGCAACTCCACTTTCTCACAAAATGGC
GGCGGAGCACTTCCGGCTTGCCCAAAATGGCCGCCAAAAACATCCGGGTC
AAAGTTCGCCGCTACGTCATAAGTCACGTGACTGGGGAGGTACTTAAACA
CGGAAGTATCCTCAACCACGTAACTGGTCACGTGGTGCGCACGTCACGGC
AACCATTTTGTTTTACAAAATGGCGCATTTCCTTCCTCTTTTTTAAAAAT
TAACCGTTGGCGGCGGCGCGCGCGCTACGCGCGCGCGCCGGGGAGCTCTG
CCCCCCCCCGCGCATGCGCGCGGGTCCCCCCCCCGCGGGGGGCTCCGCCC
CCCGGTCCCCCCCCCG (SEQ ID NO: 47)

[1687] 注释：

[1688]

推定结构域	碱基范围
TATA 盒	82 - 86
起始元件	94-115
转录起始位点	115
5' UTR 保守结构域	170 - 240
ORF2	357 - 728
ORF2/2	357 - 724; 2411 - 2870
ORF2/3	357 - 724; 2646 - 3081
ORF1	599 - 2896
ORF1/1	599 - 724; 2411 - 2896
ORF1/2	599 - 724; 2646 - 2870
三开放阅读框区	2629 - 2867
聚(A)信号	3076 - 3086
GC 富集区	3759 - 3866

[1689] 表14.示例性指环病毒氨基酸序列(甲型细环病毒,进化分枝7)

[1690]

TTV-HD16d (甲型细环病毒进化分枝 7)	
ORF2	MSWSRPVHNAAGIERQWFESTFRSHASCCGCGNFVNHINVLAARYGFTGGP TPPGGPGPRPQLRPALPAPDPDPQAPNREPWRGAGGGNDGEGAAGNPGGA AGDVYDGEDLDALFAAVVEDVE (SEQ ID NO: 48)
ORF2/2	MSWSRPVHNAAGIERQWFESTFRSHASCCGCGNFVNHINVLAARYGFTGGP TPPGGPGPRPQLRPALPAPDPDPQAPNREPWRGAGGGNDGEGAAGNPGGA AGDVYDGEDLDALFAAVVEDVESSRTPAKTVDWNPPTPVDSVATYKLLTH

[1691]

	TPWAPNSRSTTGTTDMAFLAKTLSTECLNNQKMMQTILTHTKGLDIFHPQT KPPKSKKKTSVSSKQHRRTQKRAIKKSSKKRKYSDSSQSSTSNSTCSSQSGSE SESNSDTYSNRCSKLRPIST (SEQ ID NO: 49)
ORF2/3	MSWSRPVHNAAGIERQWFESTFRSHASCCGCGNFVNHINVLAARYGFTGGP TPPGGPGPRPQLRPALPAPDPDPQAPNREPWRGAGGGNDGEGAAGNPGGA AGDVYDGEDLDALFAAVVEDVEPSRPRARKRLQFPQNSTVELRRERSRSPP RNASTPIPARAAQATPPAARRAAANRRATPIPTPTDVQNSGQSPPKPIYIYPA VKQVFMFDPPGPKAISGAKAWEDEFLTAKVWNRPVRKYYSDTPYYPWAPK PQYSVSFKLGWK (SEQ ID NO: 50)
ORF1	MAWSWWWQRWRRRRWKPRRRRWRRLRWRRPRRAVRRRRRGRRVRRRR WARRRGRRRRYATRRKRRYRGRRFKKKLVLTQWHPNTMRRCLIKGIVPLV ICGHTRWNYNYALHSKDYTEEGRYPHGGALSTTTWSLKVLYDEHLKHHDF WGYPNNQLDLARYKGAKFTFYRHKKTDFIIFFNRKPPFKLNKYSCASYHPG MLMQQRHKILLPSYETKPKGRPKITVRIKPPTLLEDKWYTQQDLCDVNLLQ LVVTAADFRHPLCSPQTNTPTTTFQVLKDIYYDTMSISEPTDSYTSVNNKSTT QTFTNYSNTLENILYTRASYWNSFHATEYLNPNIIYKNGEKLFKEHEDLITW MTQTNNTGFLTKNNTAFGNNSYRPNADKIKKARKTYWNALIGTNDLATNI GQARAERFEYHLGWYSPIFLSRHRSNMNFARAYQDVTYNPNCDRGVNNRV WVQPLTKPTTEFDEKRCKCVVQHLPLWAALYCYQDFVEEELGSSSEILNSC LLVLQCPYTFPPMYDKKLPDKGFVFYDSLFGDGKMSDGRGQVDIFWQQRW YPRLATQMQVMHDITMTGPFSYRDELVSTQLTAKYTFDFMWGGNMISTQII KNPCKDSGLEPAYPGRQRRDLQIVDPYSMGPQFSFHNWDYRHGLFGQDAID RVSKQPKDDADYPNPYKRPRYFPPTDQAAQEQEKDFSFLKTAPSNSEESDQE VLQETQVLRFQPEQHKQLHLQLAERQRIGEQLRYLLQQMFKTQANLHLNPY TFTQL (SEQ ID NO: 51)
ORF1/1	MAWSWWWQRWRRRRWKPRRRRWRRLRWRRPRRAVRRRRRGRRIIKNPC KDSGLEPAYPGRQRRDLQIVDPYSMGPQFSFHNWDYRHGLFGQDAIDRVSK QPKDDADYPNPYKRPRYFPPTDQAAQEQEKDFSFLKTAPSNSEESDQEVLQE TQVLRFQPEQHKQLHLQLAERQRIGEQLRYLLQQMFKTQANLHLNPYTFTQ L (SEQ ID NO: 52)
ORF1/2	MAWSWWWQRWRRRRWKPRRRRWRRLRWRRPRRAVRRRRRGRRTKPPK SKKKTSVSSKQHRRTQKRAIKKSSKKRKYSDSSQSSTSNSTCSSQSGSESESN SDTYSNRCSKLRPIST (SEQ ID NO: 53)

[1692] 表15.示例性指环病毒核酸序列(乙型细环病毒)

[1693] 名称	TTMV-LY2
[1694] 属/进化分枝	乙型细环病毒
[1695] 登录号	JX134045.1

[1696] 全序列:2797bp

```
1         10         20         30         40         50
|         |          |          |          |          |
TAATAAATATTCAACAGGAAAACCACCTAATTTAAATTGCCGACCACAAA
CCGTCACTTAGTTCCCCTTTTTGCAACAACTTCTGCTTTTTTCCAACTGC
CGGAAAACCACATAATTTGCATGGCTAACCACAAACTGATATGCTAATTA
ACTTCCACAAAACAACTTCCCCTTTTAAAACCACACCTACAAATTAATTA
TTAAACACAGTCACATCCTGGGAGGTACTACCACACTATAATACCAAGTG
CACTTCCGAATGGCTGAGTTTATGCCGCTAGACGGAGAACGCATCAGTTA
CTGACTGCGGACTGAACTTGGGCGGGTGCCGAAGGTGAGTGAAACCACCG
AAGTCAAGGGGCAATTCGGGCTAGTTCAGTCTAGCGGAACGGGCAAGAAA
CTTAAAATTATTTTATTTTTCAGATGAGCGACTGCTTTAAACCAACATGC
TACAACAACAAAACAAAGCAAACTCACTGGATTAATAACCTGCATTTAAC
CCACGACCTGATCTGCTTCTGCCCAACACCAACTAGACACTTATTACTAG
CTTTAGCAGAACAACAAGAAACAATTGAAGTGTCTAAACAAGAAAAAGAA
AAAATAACAAGATGCCTTATTACTACAGAAGAAGACGGTACAACTACAGA
CGTCCTAGATGGTATGGACGAGGTTGGATTAGACGCCCTTTTCGCAGAAG
ATTTCGAAGAAAAAGAAGGGTAAGACCTACTTATACTACTATTCCTCTAA
AGCAATGGCAACCGCCATATAAAAGAACATGCTATATAAAAGGACAAGAC
TGTTTAATATACTATAGCAACTTAAGACTGGGAATGAATAGTACAATGTA
```

[1697]

```
TGAAAAAAGTATTGTACCTGTACATTGGCCGGGAGGGGGTTCTTTTTCTG
TAAGCATGTTAACTTTAGATGCCTTGTATGATATACATAAACTTTGTAGA
AACTGGTGGACATCCACAAACCAAGACTTACCACTAGTAAGATATAAAGG
ATGCAAAATAACATTTTATCAAAGCACATTTACAGACTACATAGTAAGAA
TACATACAGAACTACCAGCTAACAGTAACAAACTAACATACCCAAACACA
CATCCACTAATGATGATGATGTCTAAGTACAAACACATTATACCTAGTAG
ACAAACAAGAAGAAAAAAGAAACCATACACAAAAATATTTGTAAAACCAC
CTCCGCAATTTGAAAACAAATGGTACTTTGCTACAGACCTCTACAAAATT
CCATTACTACAAATACACTGCACAGCATGCAACTTACAAAACCCATTTGT
AAAACCAGACAAATTATCAAACAATGTTACATTATGGTCACTAAACACCA
TAAGCATACAAAATAGAAACATGTCAGTGGATCAAGGACAATCATGGCCA
TTTAAAATACTAGGAACACAAAGCTTTTATTTTTACTTTTACACCGGAGC
AAACCTACCAGGTGACACAACACAAATACCAGTAGCAGACCTATTACCAC
TAACAAACCCAAGAATAAACAGACCAGGACAATCACTAAATGAGGCAAAA
ATTACAGACCATATTACTTTCACAGAATACAAAAACAAATTTACAAATTA
TTGGGGTAACCCATTTAATAAACACATTCAAGAACACCTAGATATGATAC
TATACTCACTAAAAAGTCCAGAAGCAATAAAAAACGAATGGACAACAGAA
AACATGAAATGGAACCAATTAAACAATGCAGGAACAATGGCATTAACACC
ATTTAACGAGCCAATATTCACACAAATACAATATAACCCAGATAGAGACA
CAGGAGAAGACACTCAATTATACCTACTCTCTAACGCTACAGGAACAGGA
```

TGGGACCCACCAGGAATTCCAGAATTAATACTAGAAGGATTTCCACTATG
GTTAATATATTGGGGATTTGCAGACTTTCAAAAAAACCTAAAAAAAGTAA
CAAACATAGACACAAATTACATGTTAGTAGCAAAAACAAAATTTACACAA
AAACCTGGCACATTCTACTTAGTAATACTAAATGACACCTTTGTAGAAGG
CAATAGCCCATATGAAAAACAACCTTTACCTGAAGACAACATTAAATGGT
ACCCACAAGTACAATACCAATTAGAAGCACAAAACAAACTACTACAAACT
GGGCCATTTACACCAAACATACAAGGACAACTATCAGACAATATATCAAT
GTTTTATAAATTTTACTTTAAATGGGGAGGAAGCCCACCAAAAGCAATTA
ATGTTGAAAATCCTGCCCACCAGATTCAATATCCCATACCCCGTAACGAG
CATGAAACAACTTCGTTACAGAGTCCAGGGGAAGCCCCAGAATCCATCTT
ATACTCCTTCGACTATAGACACGGGAACTACACAACAACAGCTTTGTCAC

[1698] GAATTAGCCAAGACTGGGCACTTAAAGACACTGTTTCTAAAATTACAGAG
CCAGATCGACAGCAACTGCTCAAACAAGCCCTCGAATGCCTGCAAATCTC
GGAAGAAACGCAGGAGAAAAAAGAAAAAGAAGTACAGCAGCTCATCAGCA
ACCTCAGACAGCAGCAGCAGCTGTACAGAGAGCGAATAATATCATTATTA
AAGGACCAATAACTTTTAACTGTGTAAAAAAGGTGAAATTGTTTGATGAT
AAACCAAAAAACCGTAGATTTACACCTGAGGAATTTGAAACTGAGTTACA
AATAGCAAAATGGTTAAAGAGACCCCCAAGATCCTTTGTAAATGATCCTC
CCTTTTACCCATGGTTACCACCTGAACCTGTTGTAAACTTTAAGCTTAAT
TTTACTGAATAAAGGCCAGCATTAATTCACTTAAGGAGTCTGTTTATTTA
AGTTAAACCTTAATAAACGGTCACCGCCTCCCTAATACGCAGGCGCAGAA
AGGGGGCTCCGCCCCCTTTAACCCCCAGGGGGCTCCGCCCCCTGAAACCC
CCAAGGGGGCTACGCCCCCTTACACCCCC (SEQ ID NO: 54)

[1699] 注释：

推定结构域	*碱基范围*
TATA 盒	237- 243
帽位点	260 - 267
转录起始位点	267
5’ UTR 保守结构域	323 - 393
ORF2	424 - 723
ORF2/2	424 - 719; 2274 - 2589
ORF2/3	424 - 719; 2449 - 2812
ORF1	612 - 2612
ORF1/1	612 - 719; 2274 - 2612
ORF1/2	612 - 719; 2449 - 2589
三开放阅读框区	2441 - 2586
聚(A)信号	2808 - 2813
GC 富集区	2868 - 2929

[1700]

[1701] 表16.示例性指环病毒氨基酸序列(乙型细环病毒)

TTMV-LY2（乙型细环病毒）	
ORF2	MSDCFKPTCYNNKTKQTHWINNLHLTHDLICFCPTPTRHLLLALAEQQETIEVSKQE
	KEKITRCLITTEEDGTTTDVLDGMDEVGLDALFAEDFEEKEG (SEQ ID NO: 55)
ORF2/2	MSDCFKPTCYNNKTKQTHWINNLHLTHDLICFCPTPTRHLLLALAEQQETIEVSKQE KEKITRCLITTEEDGTTTDVLDGMDEVGLDALFAEDFEEKEGFNIPYPVTSMKQLRY RVQGKPQNPSYTPSTIDTGTTQQQLCHELAKTGHLKTLFLKLQSQIDSNCSNKPSNA CKSRKKRRRKKKKKYSSSSATSDSSSSCTESE (SEQ ID NO: 56)
ORF2/3	MSDCFKPTCYNNKTKQTHWINNLHLTHDLICFCPTPTRHLLLALAEQQETIEVSKQE KEKITRCLITTEEDGTTTDVLDGMDEVGLDALFAEDFEEKEGARSTATAQTSPRMP ANLGRNAGEKRKRSTAAHQQPQTAAAAVQRANNIIIKGPITFNCVKKVKLFDDKPK NRRFTPEEFETELQIAKWLKRPPRSFVNDPPFYPWLPPEPVVNFKLNFTE (SEQ ID NO: 57)
ORF1	MPYYYRRRRYNYRRPRWYGRGWIRRPFRRRFRRKRRVRPTYTTIPLKQWQPPYKR TCYIKGQDCLIYYSNLRLGMNSTMYEKSIVPVHWPGGGSFSVSMLTLDALYDIHKL CRNWWTSTNQDLPLVRYKGCKITFYQSTFTDYIVRIHTELPANSNKLTYPNTHPLM MMMSKYKHIIPSRQTRRKKKPYTKIFVKPPPQFENKWYFATDLYKIPLLQIHCTACN LQNPFVKPDKLSNNVTLWSLNTISIQNRNMSVDQGQSWPFKILGTQSFYFYFYTGA NLPGDTTQIPVADLLPLTNPRINRPGQSLNEAKITDHITFTEYKNKFTNYWGNPFNK HIQEHLDMILYSLKSPEAIKNEWTTENMKWNQLNNAGTMALTPFNEPIFTQIQYNP DRDTGEDTQLYLLSNATGTGWDPPGIPELILEGFPLWLIYWGFADFQKNLKKVTNID TNYMLVAKTKFTQKPGTFYLVILNDTFVEGNSPYEKQPLPEDNIKWYPQVQYQLEA QNKLLQTGPFTPNIQGQLSDNISMFYKFYFKWGGSPPKAINVENPAHQIQYPIPRNE HETTSLQSPGEAPESILYSFDYRHGNYTTTALSRISQDWALKDTVSKITEPDRQQLLK QALECLQISEETQEKKEKEVQQLISNLRQQQQLYRERIISLLKDQ (SEQ ID NO: 58)
ORF1/1	MPYYYRRRRYNYRRPRWYGRGWIRRPFRRRFRRKRRIQYPIPRNEHETTSLQSPGE APESILYSFDYRHGNYTTTALSRISQDWALKDTVSKITEPDRQQLLKQALECLQISEE TQEKKEKEVQQLISNLRQQQQLYRERIISLLKDQ (SEQ ID NO: 59)
ORF1/2	MPYYYRRRRYNYRRPRWYGRGWIRRPFRRRFRRKRRSQIDSNCSNKPSNACKSRK KRRRKKKKKYSSSSATSDSSSSCTESE (SEQ ID NO: 60)

[1702]

[1703] 表17.示例性指环病毒核酸序列(丙型细环病毒)

[1704] 名称 TTMDV-MD1-073

[1705] 属/进化分枝 丙型细环病毒

[1706] 登录号 AB290918.1

[1707] 全序列:3242bp

```
1         10         20         30         40         50
|         |          |          |          |          |
AGGTGGAGACTCTTAAGCTATATAACCAAGTGGGGTGGCGAATGGCTGAG
TTTACCCCGCTAGACGGTGCAGGGACCGGATCGAGCGCAGCGAGGAGGTC
CCCGGCTGCCCGTGGGCGGGAGCCCGAGGTGAGTGAAACCACCGAGGTCT
AGGGGCAATTCGGGCTAGGGCAGTCTAGCGGAACGGGCAAGAAACTTAAA
AATATTTCTTTTACAGATGCAAAACCTATCAGCCAAAGACTTCTACAAAC
CATGCAGATACAACTGTGAAACTAAAAACCAAATGTGGATGTCTGGCATT
GCTGACTCCCATGACAGTTGGTGTGACTGTGATACTCCTTTTGCTCACCT
CCTGGCTAGTATTTTTCCTCCTGGTCACACAGATCGCACACGAACCATCC
AAGAAATACTTACCAGAGATTTTAGGAAAACATGCCTTTCTGGTGGGGCC
GACGCAACAAATTCTGGTATGGCCGAAACTATAGAAGAAAAAAGAGAAGA
TTTCCAAAAAGAAGAAAAAGAAGATTTTACAGAAGAACAAAATATAGAAG
ACCTGCTCGCCGCCGTCGCAGACGCAGAAGGAAGGTAAGAAGAAAAAAAA
AAACTCTTATAGTAAGACAATGGCAGCCAGACTCTATTGTACTCTGTAAA
ATTAAAGGGTATGACTCTATAATATGGGGAGCTGAAGGCACACAGTTTCA
ATGTTCTACACATGAAATGTATGAATATACAAGACAAAAGTACCCTGGGG
GAGGAGGATTTGGTGTACAACTTTACAGCTTAGAGTATTTGTATGACCAA
TGGAAACTTAGAAATAATATATGGACTAAAACAAATCAACTCAAAGATTT
GTGTAGATACTTAAAATGTGTTATGACCTTTTACAGACACCAACACATAG
ATTTTGTAATTGTATATGAAAGACAACCCCCATTTGAAATAGATAAACTA
ACATACATGAAATATCATCCATATATGTTATTACAAAGAAAGCATAAAAT
AATTTTACCTAGTCAAACAACTAATCCTAGAGGTAAATTAAAAAAAAAGA
AAACTATTAAACCTCCCAAACAAATGCTCAGCAAATGGTTTTTTCAACAA
CAATTTGCTAAATATGATCTACTACTTATTGCTGCAGCAGCATGTAGTTT
AAGATACCCTAGAATAGGCTGCTGCAATGAAAATAGAATGATAACCTTAT
ACTGTTTAAATACTAAATTTTATCAAGATACAGAATGGGGAACTACAAAA
[1708] CAGGCCCCCCACTACTTTAAACCATATGCAACAATTAATAAATCCATGAT
ATTTGTCTCTAACTATGGAGGTAAAAAAACAGAATATAACATAGGCCAAT
GGATAGAAACAGATATACCTGGAGAAGGTAATCTAGCAAGATACTACAGA
TCAATAAGTAAAGAAGGAGGTTACTTTTCACCTAAAATACTGCAAGCATA
TCAAACAAAAGTAAAGTCTGTAGACTACAAACCTTTACCAATTGTTTTAG
GTAGATATAACCCAGCAATAGATGATGGAAAAGGCAACAAAATTTACTTA
CAAACTATAATGAATGGCCATTGGGGCCTACCTCAAAAAACACCAGATTA
TATAATAGAAGAGGTCCCTCTTTGGCTAGGCTTCTGGGGATACTATAACT
ACTTAAAACAAACAAGAACTGAAGCTATATTTCCACTACACATGTTTGTA
GTGCAAAGCAAATACATTCAAACACAACAAACAGAAACACCTAACAATTT
TTGGGCATTTATAGACAACAGCTTTATACAGGGCAAAAACCCATGGGACT
CAGTTATTACTTACTCAGAACAAAAGCTATGGTTTCCTACAGTTGCATGG
CAACTAAAAACCATAAATGCTATTTGTGAAAGTGGACCATATGTACCTAA
ACTAGACAATCAAACATATAGTACCTGGGAACTAGCAACTCATTACTCAT
TTCACTTTAAATGGGGTGGTCCACAGATATCAGACCAACCAGTTGAAGAC
CCAGGAAACAAAAACAAATATGATGTGCCCGATACAATCAAAGAAGCATT
ACAAATTGTTAACCCAGCAAAAAACATTGCTGCCACGATGTTCCATGACT
GGGACTACAGACGGGGTTGCATTACATCAACAGCTATTAAAAGAATGCAA
CAAAACCTCCCAACTGATTCATCTCTCGAATCTGATTCAGACTCAGAACC
AGCACCCAAGAAAAAAAGACTACTACCAGTCCTCCACGACCCACAAAAGA
AAACGGAAAAGATCAACCAATGTCTCCTCTCTCTCTGCGAAGAAAGTACA
TGCCAGGAGCAGGAAACGGAGGAAAACATCCTCAAGCTCATCCAGCAGCA
GCAGCAGCAGCAGCAGAAACTCAAGCACAACCTCTTAGTACTAATCAAGG
ACTTAAAAGTGAAACAAAGATTATTACAACTACAAACGGGGGTACTAGAA
TAACCCTTACCAGATTTAAACCAGGATTTGAGCAAGAAACTGAAAAAGAG
TTAGCACAAGCATTTAACAGACCCCCTAGACTGTTCAAAGAAGATAAACC
CTTTTACCCCTGGCTACCCAGATTTACACCCCTTGTAAACTTTCACCTTA
ATTTTAAAGGCTAGGCCTACACTGCTCACTTAGTGGTGTATGTTTATTAA
```

[1709]

AGTTTGCACCCCAGAAAAATTGTAAAATAAAAAAAAAAAAAAAAATAAA
AAATTGCAAAAATTCGGCGCTCGCGCGCGCTGCGCGCGCGAGCGCCGTCA
CGCGCCGGCGCTCGCGCGCCGCGCGTATGTGCTAACACACCACGCACCTA
GATTGGGGTGCGCGCGTAGCGCGCGCACCCCAATGCGCCCCGCCCTCGTT
CCGACCCGCTTGCGCGGGTCGGACCACTTCGGGCTCGGGGGGGCGCGCCT
GCGGCGCTTATTTACTAAACAGACTCCGAGTCGCCATTGGGCCCCCCCTA
AGCTCCGCCCCCCTCATGAATATTCATAAAGGAAACCACAAAATTAGAAT
TGCCGACCACAAACTGCCATATGCTAATTAGTTCCCCTTTTACACAGTAA
AAAGGGGAAGTGGGGGGGCAGAGCCCCCCCACACCCCCCGCGGGGGGGGC
AGAGCCCCCCCCGCACCCCCCCTACGTCACAGGCCACGCCCCCGCCGCCA
TCTTGGGTGCGGCAGGGCGGGGACTAAAATGGCGGGACCCAATCATTTTA
TACTTTCACTTTCCAATTAAAACCCGCCACGTCACACAAAAG (SEQ ID NO: 61)

[1710] 注释：

[1711]

推定结构域	*碱基范围*
TATA 盒	21- 25
帽位点	42 - 49
转录起始位点	49
5’UTR 保守结构域	117 - 187
ORF2	283 - 588
ORF2/2	283 - 584; 1977 - 2388
ORF2/3	283 - 584; 2197 - 2614
ORF1	432 - 2453
ORF1/1	432 - 584; 1977 - 2453
ORF1/2	432 - 584; 2197 - 2388
三开放阅读框区	2186 - 2385
聚(A)信号	2676 - 2681
GC 富集区	3054 - 3172

[1712] 表18.示例性指环病毒氨基酸序列(丙型细环病毒)

[1713]

TTMDV-MD1-073（丙型细环病毒）	
ORF2	MWMSGIADSHDSWCDCDTPFAHLLASIFPPGHTDRTRTIQEILTRDFRKTCLSGGAD ATNSGMAETIEEKREDFQKEEKEDFTEEQNIEDLLAAVADAEGR (SEQ ID NO: 62)
ORF2/2	MWMSGIADSHDSWCDCDTPFAHLLASIFPPGHTDRTRTIQEILTRDFRKTCLSGGAD ATNSGMAETIEEKREDFQKEEKEDFTEEQNIEDLLAAVADAEGRYQTNQLKTQETK TNMMCPIQSKKHYKLLTQQKTLLPRCSMTGTTDGVALHQQLLKECNKTSQLIHLSN LIQTQNQHPRKKDYYQSSTTHKRKRKRSTNVSSLSAKKVHARSRKRRKTSSSSSSSS SSSSRNSSTTS (SEQ ID NO: 63)
ORF2/3	MWMSGIADSHDSWCDCDTPFAHLLASIFPPGHTDRTRTIQEILTRDFRKTCLSGGAD

[1714]

	ATNSGMAETIEEKREDFQKEEKEDFTEEQNIEDLLAAVADAEGRTSTQEKKTTTSPP RPTKENGKDQPMSPLSLRRKYMPGAGNGGKHPQAHPAAAAAAAETQAQPLSTNQ GLKSETKIITTTNGGTRITLTRFKPGFEQETEKELAQAFNRPPRLFKEDKPFYPWLPRF TPLVNFHLNFKG (SEQ ID NO: 64)
ORF1	MPFWWGRRNKFWYGRNYRRKKRRFPKRRKRRFYRRTKYRRPARRRRRRRRKVR RKKKTLIVRQWQPDSIVLCKIKGYDSIIWGAEGTQFQCSTHEMYEYTRQKYPGGGG FGVQLYSLEYLYDQWKLRNNIWTKTNQLKDLCRYLKCVMTFYRHQHIDFVIVYER QPPFEIDKLTYMKYHPYMLLQRKHKIILPSQTTNPRGKLKKKKTIKPPKQMLSKWFF QQQFAKYDLLLIAAAACSLRYPRIGCCNENRMITLYCLNTKFYQDTEWGTTKQAPH YFKPYATINKSMIFVSNYGGKKTEYNIGQWIETDIPGEGNLARYYRSISKEGGYFSPK ILQAYQTKVKSVDYKPLPIVLGRYNPAIDDGKGNKIYLQTIMNGHWGLPQKTPDYII EEVPLWLGFWGYYNYLKQTRTEAIFPLHMFVVQSKYIQTQQTETPNNFWAFIDNSFI QGKNPWDSVITYSEQKLWFPTVAWQLKTINAICESGPYVPKLDNQTYSTWELATH YSFHFKWGGPQISDQPVEDPGNKNKYDVPDTIKEALQIVNPAKNIAATMFHDWDY RRGCITSTAIKRMQQNLPTDSSLESDSDSEPAPKKKRLLPVLHDPQKKTEKINQCLLS LCEESTCQEQETEENILKLIQQQQQQQQKLKHNLLVLIKDLKVKQRLLQLQTGVLE (SEQ ID NO: 65)
ORF1/1	MPFWWGRRNKFWYGRNYRRKKRRFPKRRKRRFYRRTKYRRPARRRRRRRRKISD QPVEDPGNKNKYDVPDTIKEALQIVNPAKNIAATMFHDWDYRRGCITSTAIKRMQQ NLPTDSSLESDSDSEPAPKKKRLLPVLHDPQKKTEKINQCLLSLCEESTCQEQETEEN ILKLIQQQQQQQQKLKHNLLVLIKDLKVKQRLLQLQTGVLE (SEQ ID NO: 66)
ORF1/2	MPFWWGRRNKFWYGRNYRRKKRRFPKRRKRRFYRRTKYRRPARRRRRRRRKISD QPVEDPGNKNKYDVPDTIKEALQIVNPAKNIAATMFHDWDYRRGCITSTAIKRMQQ NLPTDSSLESDSDSEPAPKKKRLLPVLHDPQKKTEKINQCLLSLCEESTCQEQETEEN ILKLIQQQQQQQQKLKHNLLVLIKDLKVKQRLLQLQTGVLE (SEQ ID NO: 67)

[1715] 表B1.示例性指环病毒核酸序列(丙型细环病毒)

[1716] 名称 环3.1

[1717] 属/进化分枝 丙型细环病毒

[1718] 登录号

[1719] 全序列:3264bp

[1720]

```
1         10         20         30         40         50
|         |          |          |          |          |
TAAAATGGCGGCAACCAATCATTTTATACTTTCACTTTCCAATTACAAGC
```

CGCCACGTCACAGAACAGGGGTGGAGACTTTAAAACTATATAACCAAGTG
ATGTGACGAATGGCTGAGTTTACCCCGCTAGACGGTGCAGGGACCGGATC
GAGCGCAGCGAGGAGGTCCCCGGCTGCCCGTGGGCGGGAGCCCGAGGTGA
GTGAAACCACCGAGGTCTAGGGGCAATTCGGGCTAGGGCAGTCTAGCGGA
ACGGGCAAGAAACTTAAAATATGTTTTGTTTCAGATGCAGACACCTGCTT
CACAGATAAGCTCAGACGACTTCTTTGTACACACTCCATTTAATGCAGTA
ACTAAACAGCAAATATGGATGTCTCAAATTGCTGATGGACATGACAACAT
TTGTCACTGCCACCGTCCTTTTGCTCACCTGCTTGCTAATATTTTTCCTC
CTGGTCATAAAGACAGGGATCTTACCATTAATCAAATACTTGCTAGAGAT
CTTACAGAAACATGCCATTCTGGTGGAGACGAAGGAACAAGCGGTGGTGG
GGTCGCCGCTTCCGCTACCGCCGCTACAACAAATATAAAACCAGAAGGAG
ACGCAGAATACCCAGAAGACGAAATAGAAGATTTACTAAGACACGCAGGA
GAAGAAAAAGAAAGAAGGTAAGAAGAAAACTTAAAAAAATTACTATTAAA
CAATGGCAGCCAGATTCAGTGAAAAAATGTAAAATTAAAGGATATAGTAC
TTTAGTTATGGGTGCACAAGGAAAACAATACAACTGTTACACAAACCAAG
CAAGTGACTATGTTCAGCCTAAAGCACCACAAGGTGGGGGCTTTGGCTGT
GAAGTATTTAATTTAAAATGGCTATACCAAGAATATACTGCACACAGAAA
TATTTGGACAAAAACAAATGAATATACAGACCTTTGTAGATACACTGGAG
CTCAAATAATTTTATACAGGCACCCAGATGTTGATTTTATAGTCAGCTGG
GACAATCAGCCACCTTTTTTACTTAACAAATATACATATCCAGAACTGCA
ACCACAAAACCTTTTACTAGCTAGAAGGAAAAGAATTATTCTTAGTCAAA
AATCAAACCCCAAAGGAAAACTAAGAATTAAACTAAGAATACCACCACCA
AAACAAATGATAACAAAATGGTTTTTTCAAAGAGACTTTTGTGATGTGAA
TCTGTTTAAACTATGTGCTTCTGCTGCTTCTTTCCGCTACCCAGGTATCA
GTCATGGAGCTCAAAGTACTATTTTTTCTGCATATGCTTTAAACACTGAC
TTTTATCAATGCAGTGACTGGTGCCAAACTAACACAGAAACTGGCTACCT
AAACATTAAAACACAACAAATGCCACTATGGTTTCATTACAGAGAGGGTG
GCAAAGAGAAATGGTATAAATACACCAACAAAGAACACAGACCATATACA
[1721] AATACATATCTTAAAAGTATTAGCTATAATGATGGATTGTTTTCTCCTAA
AGCCATGTTTGCATTTGAAGTAAAAGCGGGGGGTGAAGGAACAACAGAAC
CACCACAAGGCGCCCAATTAATTGCTAACCTTCCACTCATTGCACTAAGA
TATAATCCACATGAAGACACAGGCCATGGCAATGAAATTTACCTTACATC
AACTTTTAAAGGTACATATGACAAACCTAAAGTTACTGATGCTCTATACT
TTAACAATGTACCCCTGTGGATGGGATTTTATGGCTACTGGGACTTTATA
TTACAAGAAACAAAAAACAAAGGTGTCTTTGATCAACATATGTTTGTTGT
TAAATGTCCTGCCTTAAGGCCCATATCACAAGTCACAAAACAAGTATACT
ACCCACTTGTAGACATGGACTTTTGTTCAGGGAGACTGCCATTTGATGAA
TATTTATCCAAAGACATTAAAAGTCATTGGTATCCCACTGCAGAAAGACA
AACAGTTACAATAAATAATTTTGTTACAGCAGGTCCATACATGCCTAAAT
TTGAACCCACAGACAAAGACAGTACATGGCAATTAAACTATCACTATAAA
TTTTTTTTTAAGTGGGGTGGTCCACAAGTCACAGACCCAACTGTTGAAGA
CCCATGCAGCAGAAACAAATATCCTGTCCCCGATACAATGCAACAAACAA
TACAAATTAAAAACCCTGAAAAGCTGCACCCAGCAACCCTCTTCCATGAC
TGGGACCTTAGAAGGGGCTTCATTACACAAGCAGCTATTAAAAGAATGTC
AGAAAACCTCCAAATTGATTCATCTTTCGAATCTGATGGCACAGAATCAC
CCAAAAAAAAGAAAAGATGCACCAAAGAAATCCCAACACAAAACCAAAAG
CAAGAAGAGATCCAAGAATGTCTCCTCTCACTCTGCGAAGAGCCTACATG
CCAAGAAGAAACAGAGGACCTCCAGCTCTTCATCCAGCAGCAGCAGCAGC
AGCAGTACAAGCTCAGAAAAAACCTCTTCAAACTCCTCACTCACCTGAAA
AAAGGACAGAGAATAAGTCAACTACAAACGGGACTTTTAGAGTAATACCA
TTTAAACCAGGTTTTGAACAAGAAACAGAAAAAGAACTTGCCATAGCTTT
CTGCAGACCACCTAGAAAATATAAAAATGATCCCCCTTTTTATCCCTGGT
TACCATGGACACCCCTTGTACACTTTAACCTTAATTACAAAGGCTAGGCC
AACACTGTTCACTTAGTGGTGTATGTTTAATAAAGTTTCACCCCCAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAATTGCAAAAATTCG
GCGCTCGCGCGCGCTGCGCGCGCGCGAGCGCCGTCACGCGCCGGCGCTCG
CGCGCCGCGCGTATGTGCTAACACACCACGCACCTAGATTGGGGTGCGCG

[1722]
CGCTAGCGCGCGCACCCCAATGCGCCCCGCCCTCGTTCCGACCCGCTTGC
GCGGGTCGGACCACTTCGGGCTCGGGGGGGCGCGCCTGCGGCGCTTTTTT
ACTAAACAGACTCCGAGCCGCCATTTGGCCCCCCCTAAGCTCCGCCCCCC
TCATGAATATTCATAAAGGAAACCACATAATTAGAATTGCCGACCACAAA
CTGCCATATGCTAATTAGTTCCCCTTTTACACAGTAAAAAGGGGAAGTGG
GGGGGCATAGCCCCCCCCACACCCCCCGCGGGGGGGGCAGAGCCCCCCCCC
GCACCCCCCCCCCTACGTCACAATCCACGCCCCCGCCGCCATCTTGGGTGC
GGCAGGGCGGGGGC (SEQ ID NO: 878)

[1723] 注释：

推定结构域	碱基范围
TATA 盒	87- 93
帽位点	110 - 117
转录起始位点	117
5’ UTR 保守结构域	185 - 255
ORF2	285 - 671
ORF2/2	285 - 667; 2063 - 2498
ORF2/3	285 - 667; 2295 - 2697
TAIP	385 - 585
ORF1	512 - 2545
ORF1/1	512 - 667; 2063 - 2545
ORF1/2	512 - 667; 2295 - 2498
三开放阅读框区	2295 - 2495
聚(A)信号	2729 - 2734
GC 富集区	3141 - 3264

[1725] 表C1.示例性指环病毒氨基酸序列(丙型细环病毒)

[1726]

Ring 3.1 (丙型细环病毒)	
ORF2	MQTPASQISSDDFFVHTPFNAVTKQQIWMSQIADGHDNICHCHRPFAHLLAN IFPPGHKDRDLTINQ ILARDLTETCHSGGDEGTSGGGVAASATAATTNIKPEGDAEYPEDEIEDLLR HAGEEKERR (SEQ ID NO: 879)
ORF2/2	MQTPASQISSDDFFVHTPFNAVTKQQIWMSQIADGHDNICHCHRPFAHLLAN IFPPGHKDRDLTINQ ILARDLTETCHSGGDEGTSGGGVAASATAATTNIKPEGDAEYPEDEIEDLLR HAGEEKERSGVVHKSQTQLLKTHAAETNILSPIQCNKQYKLKTLKSCTQQPS

[1727]

	SMTGTLEGASLHKQLLKECQKTSKLIHLSNLMAQNHPKKRKDAPKKSQHK TKSKKRSKNVSSHSAKSLHAKKKQRTSSSSSSSSSSSTSSEKTSSNSSLT (SEQ ID NO: 880)
ORF2/3	MQTPASQISSDDFFVHTPFNAVTKQQIWMSQIADGHDNICHCHRPFAHLLAN IFPPGHKDRDLTINQ ILARDLTETCHSGGDEGTSGGGVAASATAATTNIKPEGDAEYPEDEIEDLLR HAGEEKERRITQKKEKMHQRNPNTKPKARRDPRMSPLTLRRAYMPRRNRG PPALHPAAAAAAVQAQKKPLQTPHSPEKRTENKSTTNGTFRVIPFKPGFEQE TEKELAIAFCRPPRKYKNDPPFYPWLPWTPLVHFNLNYKG (SEQ ID NO: 881)
TAIP	MDMTTFVTATVLLLTCLLIFFLLVIKTGILPLIKYLLEILQKHAILVETKEQAV VGSPLPLPPLQQI (SEQ ID NO: 882)
ORF1	MPFWWRRRNKRWWGRRFRYRRYNKYKTRRRRRIPRRRNRRFTKTRRRRK RKKVRRKLKKITIKQWQP DSVKKCKIKGYSTLVMGAQGKQYNCYTNQASDYVQPKAPQGGGFGCEVF NLKWLYQEYTAHRNIWTKTNEYTDLCRYTGAQIILYRHPDVDFIVSWDNQP PFLLNKYTYPELQPQNLLLARRKRIILSQKSNPKGKLRIKLRIPPPKQMITKWF FQRDFCDVNLFKLCASAASFRYPGISHGAQSTIFSAYALNTDFYQCSDWCQT NTETGYLNIKTQQMPLWFHYREGGKEKWYKYTNKEHRPYTNTYLKSISYN DGLFSPKAMFAFEVKAGGEGTTEPPQGAQLIANLPLIALRYNPHEDTGHGNE IYLTSTFKGTYDKPKVTDALYFNNVPLWMGFYGYWDFILQETKNKGVFDQ HMFVVKCPALRPISQVTKQVYYPLVDMDFCSGRLPFDEYLSKDIKSHWYPT AERQTVTINNFVTAGPYMPKFEPTDKDSTWQLNYHYKFFFKWGGPQVTDP TVEDPCSRNKYPVPDTMQQTIQIKNPEKLHPATLFHDWDLRRGFITQAAIKR MSENLQIDSSFESDGTESPKKKKRCTKEIPTQNQKQEEIQECLLSLCEEPTCQ EETEDLQLFIQQQQQQQYKLRKNLFKLLTHLKKGQRISQLQTGLLE (SEQ ID NO: 883)
ORF1/1	MPFWWRRRNKRWWGRRFRYRRYNKYKTRRRRRIPRRRNRRFTKTRRRRK RKKWGGPQVTDPTVEDPCSRNKYPVPDTMQQTIQIKNPEKLHPATLFHDWD LRRGFITQAAIKRMSENLQIDSSFESDGTESPKKKKRCTKEIPTQNQKQEEIQE CLLSLCEEPTCQEETEDLQLFIQQQQQQQYKLRKNLFKLLTHLKKGQRISQL QTGLLE (SEQ ID NO: 884)

ORF1/2	MPFWWRRRNKRWWGRRFRYRRYNKYKTRRRRRIPRRRNRRFTKTRRRRK RKKNHPKKRKDAPKKSQHKTKSKKRSKNVSSHSAKSLHAKKKQRTSSSSSS SSSSSSTSSEKTSSNSSLT (SEQ ID NO: 885)

[1728] 表B2.示例性指环病毒核酸序列(丙型细环病毒)

[1729] 名称 环4.0

[1730] 属/进化分枝 丙型细环病毒

[1731] 登录号

[1732] 全序列:3176bp

1 10 20 30 40 50

TAAAATGGCGGGAGCCAATCATTTTATACTTTCACTTTCCAATTAAAAAT
GGCCACGTCACAAACAAGGGGTGGAGCCATTTAAACTATATAACTAAGTG
GGGTGGCGAATGGCTGAGTTTACCCCGCTAGACGGTGCAGGGACCGGATC
GAGCGCAGCGAGGAGGTCCCCGGCTGCCCATGGGCGGGAGCCGAGGTGAG
TGAAACCACCGAGGTCTAGGGGCAATTCGGGCTAGGGCAGTCTAGCGGAA
CGGGCAAGAAACTTAAAACAATATTTGTTTTACAGATGGTTAGTATATCC
TCAAGTGATTTTTTTAAGAAAACGAAATTTAATGAGGAGACGCAGAACCA
AGTATGGATGTCTCAAATTGCTGACTCTCATGATAATATCTGCAGTTGCT
GGCATCCATTTGCTCACCTTCTTGCTTCCATATTTCCTCCTGGCCACAAA
GATCGTGATCTTACTATTAACCAAATTCTTCTAAGAGATTATAAAGAAAA
ATGCCATTCTGGTGGAGAAGAAGGAGAAAATTCTGGACCAACAACAGGTT
TAATTACACCAAAAGAAGAAGATATAGAAAAAGATGGCCCAGAAGGCGCC
GCAGAAGAAGACCATACAGACGCCCTGTTCGCCGCCGCCGTAGAAAACTT
CGAAAGGTAAAGAGAAAAAAAAAATCTTTAATTGTTAGACAATGGCAACC
AGACAGTATAAGAACTTGTAAAATTATAGGACAGTCAGCTATAGTTGTTG
GGGCTGAAGGAAAGCAAATGTACTGTTATACTGTCAATAAGTTAATTAAT
GTGCCCCCAAAAACACCATATGGGGGAGGCTTTGGAGTAGACCAATACAC

[1733] ACTGAAATACTTATATGAAGAATACAGATTTGCACAAAACATTTGGACAC
AATCTAATGTACTGAAAGACTTATGCAGATACATAAATGTTAAGCTAATA
TTCTACAGAGACAACAAAACAGACTTTGTCCTTTCCTATGACAGAAACCC
ACCTTTTCAACTAACAAAATTTACATACCCAGGAGCACACCCACAACAAA
TCATGCTTCAAAAACACCACAAATTCATACTATCACAAATGACAAAGCCT
AATGGAAGACTAACAAAAAAACTCAAAATTAAACCTCCTAAACAAATGCT
TTCTAAATGGTTCTTTTCAAAACAATTCTGTAAATACCCTTTACTATCTC
TTAAAGCTTCTGCACTAGACCTTAGGCACTCTTACCTAGGCTGCTGTAAT
GAAAATCCACAGGTATTTTTTTATTATTTAAACCATGGATACTACACAAT
AACAAACTGGGGAGCACAATCCTCAACAGCATACAGACCTAACTCCAAGG
TGACAGACACAACATACTACAGATACAAAAATGACAGAAAAAATATTAAC
ATTAAAAGCCATGAATACGAAAAAAGTATATCATATGAAAACGGTTATTT
TCAATCTAGTTTCTTACAAACACAGTGCATATATACCAGTGAGCGTGGTG
AAGCCTGTATAGCAGAAAAACCACTAGGAATAGCTATTTACAATCCAGTA
AAAGACAATGGAGATGGTAATATGATATACCTTGTAAGCACTCTAGCAAA
CACTTGGGACCAGCCTCCAAAAGACAGTGCTATTTTAATACAAGGAGTAC
CCATATGGCTAGGCTTATTTGGATATTTAGACTACTGTAGACAAATTAAA
GCTGACAAAACATGGCTAGACAGTCATGTACTAGTAATTCAAAGTCCTGC
TATTTTTACTTACCCAAATCCAGGAGCAGGCAAATGGTATTGTCCACTAT
CACAAAGTTTTATAAATGGCAATGGTCCGTTTAATCAACCACCTACACTG

CTACAAAAAGCAAAGTGGTTTCCACAAATACAATACCAACAAGAAATTAT
TAATAGCTTTGTAGAATCAGGACCATTTGTTCCCAAATATGCAAATCAAA
CTGAAAGCAACTGGGAACTAAAATATAAATATGTTTTTACATTTAAGTGG
GGTGGACCACAATTCCATGAACCAGAAATTGCTGACCCTAGCAAACAAGA
GCAGTATGATGTCCCCGATACTTTCTACCAAACAATACAAATTGAAGATC
CAGAAGGACAAGACCCCAGATCTCTCATCCATGATTGGGACTACAGACGA
GGCTTTATTAAAGAAAGATCTCTTAAAAGAATGTCAACTTACTTCTCAAC
TCATACAGATCAGCAAGCAACTTCAGAGGAAGACATTCCCAAAAAGAAAA
AGAGAATTGGACCCCAACTCACAGTCCCACAACAAAAAGAAGAGGAGACA
CTGTCATGTCTCCTCTCTCTCTGCAAAAAAGATACCTTCCAAGAAACAGA
GACACAAGAAGACCTCCAGCAGCTCATCAAGCAGCAGCAGGAGCAGCAGC
TCCTCCTCAAGAGAAACATCCTCCAGCTCATCCACAAACTAAAAGAGAAT
CAACAAATGCTTCAGCTTCACACAGGCATGTTACCTTAACCAGATTTAAA
[1734] CCTGGATTTGAAGAGCAAACAGAGAGAGAATTAGCAATTATATTTCATAG
GCCCCCTAGAACCTACAAAGAGGACCTTCCATTCTATCCCTGGCTACCAC
CTGCACCCCTTGTACAATTTAACCTTAACTTCAAAGGCTAGGCCAACAAT
GTACACTTAGTAAAGCATGTTTATTAAAGCACAACCCCCAAAATAAATGT
AAAAATAAAAAAAAAAAAAAAAAAATAAAAAATTGCAAAAATTCGGCGCT
CGCGCGCATGTGCGCCTCTGGCGCAAATCACGCAACGCTCGCGCGCCCGC
GTATGTCTCTTTACCACGCACCTAGATTGGGGTGCGCGCGCTAGCGCGCG
CACCCCAATGCGCCCCGCCCTCGTTCCGACCCGCTTGCGCGGGTCGGACC
ACTTCGGGCTCGGGGGGGCGCGCCTGCGGCGCTTTTTTACTAAACAGACT
CCGAGCCGCCATTTGGCCCCCTAAGCTCCGCCCCCCTCATGAATATTCAT
AAAGGAAACCACATAATTAGAATTGCCGACCACAAACTGCCATATGCTAA
TTAGTTCCCCTTTTACAAAGTAAAAGGGGAAGTGAACATAGCCCCACACC
CGCAGGGGCAAGGCCCCGCACCCCTACGTCACTAACCACGCCCCCGCCGC
CATCTTGGGTGCGGCAGGGCGGGGGC (SEQ ID NO: 886)

[1735] 注释：

[1736]

推定结构域	*碱基范围*
TATA 盒	87- 93
帽位点	110 - 117
转录起始位点	117
5’ UTR 保守结构域	185 - 254
ORF2	286 - 660
ORF2/2	286 - 656; 1998 - 2442
ORF2/3	286 - 656; 2209 - 2641
TAIP	385 - 484
ORF1	501 - 2489

	ORF1/1	501 - 656; 1998 - 2489
	ORF1/2	501 - 656; 2209 - 2442
[1737]	三开放阅读框区	2209 - 2439
	聚(A)信号	2672 - 2678
	GC 富集区	3076 - 3176

[1738] 表C2.示例性指环病毒氨基酸序列(丙型细环病毒)

Ring 4.0 (丙型细环病毒)	
ORF2	MVSISSSDFFKKTKFNEETQNQVWMSQIADSHDNICSCWHPFAHLLASIFPP GHKDRDLTINQILLR DYKEKCHSGGEEGENSGPTTGLITPKEEDIEKDGPEGAAEEDHTDALFAAAV ENFER (SEQ ID NO: 887)
ORF2/2	MVSISSSDFFKKTKFNEETQNQVWMSQIADSHDNICSCWHPFAHLLASIFPP GHKDRDLTINQILLRDYKEKCHSGGEEGENSGPTTGLITPKEEDIEKDGPEGA AEEDHTDALFAAAVENFESGVDHNSMNQKLLTLANKSSMMSPILSTKQYKL KIQKDKTPDLSSMIGTTDEALLKKDLLKECQLTSQLIQISKQLQRKTFPKRKR ELDPNSQSHNKKKRRHCHVSSLSAKKIPSKKQRHKKTSSSSSSSSRSSSSSSR ETSSSSSTN (SEQ ID NO: 888)
ORF2/3	MVSISSSDFFKKTKFNEETQNQVWMSQIADSHDNICSCWHPFAHLLASIFPP GHKDRDLTINQILLRDYKEKCHSGGEEGENSGPTTGLITPKEEDIEKDGPEGA AEEDHTDALFAAAVENFERSASNFRGRHSQKEKENWTPTHSPTTKRRGDTV MSPLSLQKRYLPRNRDTRRPPAAHQAAAGAAAPPQEKHPPAHPQTKRESTN ASASHRHVTLTRFKPGFEEQTERELAIIFHRPPRTYKEDLPFYPWLPPAPLVQ FNLNFKG (SEQ ID NO: 889)
TAIP	MRRRRTKYGCLKLLTLMIISAVAGIHLLTFLLPYFLLATKIVILLLTKFF (SEQ ID NO: 890)
ORF1	MPFWWRRRRKFWTNNRFNYTKRRRYRKRWPRRRRRRRPYRRPVRRRRRK LRKVKRKKKSLIVRQWQPDSIRTCKIIGQSAIVVGAEGKQMYCYTVNKLINV PPKTPYGGGFGVDQYTLKYLYEEYRFAQNIWTQSNVLKDLCRYINVKLIFY RDNKTDFVLSYDRNPPFQLTKFTYPGAHPQQIMLQKHHKFILSQMTKPNGR LTKKLKIKPPKQMLSKWFFSKQFCKYPLLSLKASALDLRHSYLGCCNENPQ VFFYYLNHGYYTITNWGAQSSTAYRPNSKVTDTTYYRYKNDRKNINIKSHE YEKSISYENGYFQSSFLQTQCIYTSERGEACIAEKPLGIAIYNPVKDNGDGNM IYLVSTLANTWDQPPKDSAILIQGVPIWLGLFGYLDYCRQIKADKTWLDSHV LVIQSPAIFTYPNPGAGKWYCPLSQSFINGNGPFNQPPTLLQKAKWFPQIQYQ QEIINSFVESGPFVPKYANQTESNWELKYKYVFTFKWGGPQFHEPEIADPSK

[1739]

[1740]

	QEQYDVPDTFYQTIQIEDPEGQDPRSLIHDWDYRRGFIKERSLKRMSTYFST HTDQQATSEEDIPKKKKRIGPQLTVPQQKEEETLSCLLSLCKKDTFQETETQE DLQQLIKQQQEQQLLLKRNILQLIHKLKENQQMLQLHTGMLP (SEQ ID NO: 891)
ORF1/1	MPFWWRRRRKFWTNNRFNYTKRRRYRKRWPRRRRRRRPYRRPVRRRRRK LRKWGGPQFHEPEIADPSKQEQYDVPDTFYQTIQIEDPEGQDPRSLIHDWDY RRGFIKERSLKRMSTYFSTHTDQQATSEEDIPKKKKRIGPQLTVPQQKEEETL SCLLSLCKKDTFQETETQEDLQQLIKQQQEQQLLLKRNILQLIHKLKENQQM LQLHTGMLP (SEQ ID NO: 892)
ORF1/2	MPFWWRRRRKFWTNNRFNYTKRRRYRKRWPRRRRRRRPYRRPVRRRRRK LRKISKQLQRKTFPKRKR ELDPNSQSHNKKKRRHCHVSSLSAKKIPSKKQRHKKTSSSSSSSSRSSSSSSR ETSSSSSTN (SEQ ID NO: 893)

[1741] 表B3.示例性指环病毒核酸序列(甲型细环病毒)-进化分枝1

[1742] 名称 环5.2

[1743] 属/进化分枝 甲型细环病毒进化分枝1

[1744] 登录号

[1745] 全序列:3696bp

[1746]

```
1         10         20         30         40         50
|         |          |          |          |          |
ATTTTGTTCAGCCCGCCAATTTCTCTTTCAAACAGGCCAATCAGCTACTA
CTTCGTGCACTTCCTGGGGCGTGTCCTGCCGCTCTATATAAGCAGAGGCG
GTGACGAATGGTAGAGTTTTTCTTGGCCCGTCCGCGGCGAGAGCGCGAGC
GAAGCGAGCGATCGAGCGTCCCGAGGGCGGGTGCCGGAGGTGAGTTTACA
CACCGCAGTCAAGGGGCAATTCGGGCTCGGGACTGGCCGGGCTATGGGCA
AGATTCTTAAAAAATTCCCCCGATCCCTTTGCCGCCAGGACATAAAAACA
TGCCGTGGAGACCGCCGGTCCATAGTGTCCAGGGGCGAGAGGATCAGTGG
TTCGCAAGCTTTTTTCACGGCCACGATTCGTTTTGCGGCTGCGGTGACCC
TCTTGGCCATATTAATAGCATTGCTCATCGCTTTCCTCGCGCCGGTCCAC
CAAGGCCCCCTCCGGGGCTAGATCAGCCTAACCCCCGGGAGCAGGGCCCG
GCCGGACCCGGAGGGCCGCCCGCCATCTTGGCCCTGCCGGCTCCGCCCGC
GGAGCCTGACGACCCGCAGCCACGGCGTGGTGGTGGGGACGGTGGCGCCG
CCGCTGGCGCCGCAGACGACCATACACAACGAGACTACGACGAAGAAGAG
CTAGACGAGCTTTTCCGCGCCGCCGCCGAAGACGATTTGTAAGTAGGAGA
TGGCGCCGGCCTTACAGGCGCAGGAGGAGACGCGGGCGACGCAGACGCAG
ACGCAGACGCAGACATAAGCCCACCCTAATACTCAGACAGTGGCAACCTG
ACTGTATCAGACACTGTAAAATAACAGGATGGATGCCCCTCATTATCTGT
```

GGAAAGGGGTCCACCCAGTTCAACTACATCACCCACGCGGACGATATCAC
CCCCAGGGGAGCCTCCTACGGAGGCAATTTCACAAACATGACTTTCTCCC
TGGAGGCCATATATGAACAGTTCCTATACCACAGAAACAGGTGGTCGGCC
TCTAACCACGACCTAGAACTGTGCAGATACAAGGGGACCACCTTAAAACT
CTACAGACACCCAGAAGTAGACTACATAGTTACCTACAGCAGAACAGGAC
CCTTTGAAATCAGCCACATGACCTACCTCAGCACTCACCCCATGCTAATG
CTGCTAAACAAGCACCACATTGTGGTGCCCAGCTTAAAGACTAAGCCCAG
AGGCAGAAAGGCCATAAAAGTCAGGATAAGGCCCCCAAAACTCATGAACA
ACAAGTGGTACTTCACCAGAGACTTCTGTAACATAGGCCTCTTCCAGCTC
TGGGCCACAGGCTTAGAACTCAGAAACCCCTGGCTCAGAATGAGCACCCT
GAGCCCCTGCATAGGCTTTAATGTCCTCAAAAACAGCATTTACACAAACC
TCAGCAACCTGCCACAATACAAAAACGAAAGACTAAACATCATTAACAAC
ATACTTCACCCACAAGAAATTACAGGTACAAACAACAAAAAGTGGCAGTA
CACATACACCAAACTCATGGCCCCTATTTACTATTCAGCAAACAGGGCCA
GCACCTATGACTGGGAAAATTACAGCAAAGAAACAAACTACAATAATACA
TATGTTAAATTTACCCAGAAAAGACAGGAAAAACTAACTAAAATTAGAAA
AGAGTGGCAGATGCTTTATCCACAACAACCCACAGCACTGCCAGACTCCT
ATGACCTCCTACAAGAGTATGGCCTCTACAGTCCATACTACCTAAACCCC
ACAAGAATAAACCTAGACTGGATGACCCCATACACACACGTCAGATACAA
TCCCCTAGTAGACAAGGGCTTTGGAAACAGAATATACATCCAGTGGTGCT
CAGAAGCAGATGTTAGCTACAACAGGACAAAATCCAAGTGTCTGCTACAA
GACATGCCCCTGTTTTTCATGTGCTATGGCTACATAGACTGGGCAATAAA
AAACACTGGAGTGTCATCTCTAGTGAAGGACGCCAGAATCTGCATCAGGT
GTCCCTACACAGAGCCACAACTAGTTGGCTCCACAGAAGACATAGGCTTT
GTACCCATCTCAGAAACCTTCATGAGGGGCGACATGCCGGTACTTGCACC
ATACATACCGTTAAGCTGGTTTTGCAAGTGGTATCCCAACATAGCTCACC
AAAAGGAAGTCCTTGAGTCAATCATTTCCTGCAGCCCCTTCATGCCCCGT
GACCAAGACATGAACGGTTGGGATATCACAATCGGTTACAAAATGGACTT
[1747] CTTATGGGGCGGTTCCCCTCTCCCCTCACAGCCAATCGACGACCCCTGCC
AGCAGGGAACCCACCCGATTCCCGACCCCGATAAACACCCTCGCCTCCTA
CAAGTCTCGAACCCGAAACTACTCGGACCGAGGACAGTGTTCCACAAGTG
GGACATCAGACGTGGGCAGTTTAGCAAAAGAAGTATTAAGAGAGTGTCAG
AATACTCAAGCGATGATGAATCTCTTGCGCCAGGTCTCCCATCAAAGCGA
AACAAGCTCGACTCGGCGTTCCGAGGAGAAAATCGAGAGCAAAAAGAATG
CTATTCTCTCCTCAAAGCGCTCGAGGAAGAAGAGACCCCAGAAGAAGAAG
AACCAGCACCCCAAGAAAAAGCCCAGAAAGAGGAGCTACTCCACCAGCTC
CAGCTCCAGAGACGCCACCAGCGAGTCCTCAGACGAGGGCTCAAGCTCGT
CTTTACAGACATCCTCCGACTCCGCCAGGGAGTCCACTGGAACCCGGAGC
TCACATAGCGCCCCCACCTTACATACCAGACCTGCTTTTTCCCAATACTG
GTAAAAAAAAAAAATTCTCTCCCTTCGATTGGGAGACAGAGGCGCAAATA
GCGGGGTGGATGCGGCGGCCCATGCGCTTCTATCCCTCAGACACCCCTCA
CTACCCGTGGCTACCCCCCGAGCGAGATATCCCGAAAATATGTAACATAA
ACTTCAAAATAAAGCTTCAAGAGTGAGTGATTCGAGGCCCTCCTCTGTTC
ACTTAGCGGTGTCTACCTCTTAAGGTCACTAAGCACTCCGAGCGTAAGCG
AGGAGTGCGACCCTCTACCAAGGGGCAACTTCCTCGGGGTCCGGCGCTAC
GCGCTTCGCGCTGCGCCGGACATCTCGGACCCCTCGACCCGAATCGCTTG
CGCGATTCGGACCTGCGGCCTCGGGGGGGTCGGGGGCTTTACTAAACAGA
CTCCGAGGTGCCATTGGACACTGTAGGGGGTGAACAGCAACGAAAGTGAG
TGGGGCCAGACTTCGCCATAAGGCCTTTATCTTCTTGCCATTGGATAGTG
ACTTCCGGGTCCGCCTGGGGGCCGCCATTTTAGCTTCGGCCGCCATTTTA
GGCCCTCGCGGGCCTCCGTAGGCGCGCTTTAGTGACGTCACGGCAGCCAT
TTTGTCGTGACGTTTGAGACACGTGATGGGGGCGTGCCTAAACCCGGAAG
CATCCCTGGTCACGTGACTCTGACGTCACGGCGGCCATCTTGTGCTGTCC
GCCATCTTGTAACTTCCTTCCGCTTTTTCAAAAAAAAAGAGGAAGTGTGA
CGTAGCGGCGGGGGGGCGGCGCGCTTCGCGCGCCGCCCACCAGGGGGCGC
TGCGCGCCCCCCGCGCATGCGCAGGGGCCTCTCGAGGGGCTCCGCCCCCC
CCCCGTGCTAAATTTACCGCGCATGCGCGACCACGCCCCCGCCGCC (SEQ ID NO: 894)

[1748] 注释：

[1749]

推定结构域	*碱基范围*
TATA 盒	85- 91
帽位点	108 - 115
转录起始位点	115
5' UTR 保守结构域	178 - 248
ORF2	300 - 692
ORF2/2	300 - 688; 2282 - 2804
ORF2/3	300 - 688; 2484 - 2976
ORF2t/3	300 - 349: 2484 - 2976
TAIP	322 - 471
ORF1	572 - 2758
ORF1/1	572 - 688; 2282 - 2758
ORF1/2	572 - 688; 2484 - 2804
三开放阅读框区	2484 - 2755
聚(A)信号	3018 -3023
GC 富集区	3555 - 3696

[1750] 表C3.示例性指环病毒氨基酸序列(甲型细环病毒)进化分枝1

[1751]

Ring 5.2 (甲型细环病毒)进化分枝1	
ORF2	MPWRPPVHSVQGREDQWFASFFHGHDSFCGCGDPLGHINSIAHRFPRAGPP RPPPGLDQPNPREQGPAGPGGPPAILALPAPPAEPDDPQPRRGGGDGGAAAG AADDHTQRDYDEEELDELFRAAAEDDL (SEQ ID NO: 895)
ORF2/2	MPWRPPVHSVQGREDQWFASFFHGHDSFCGCGDPLGHINSIAHRFPRAGPP RPPPGLDQPNPREQGPAGPGGPPAILALPAPPAEPDDPQPRRGGGDGGAAAG AADDHTQRDYDEEELDELFRAAAEDDFQSTTPASREPTRFPTPINTLASYKS RTRNYSDRGQCSTSGTSDVGSLAKEVLRECQNTQAMMNLLRQVSHQSETSS TRRSEEKIESKKNAILSSKRSRKKRPQKKKNQHPKKKPRKRSYSTSSSSRDAT SESSDEGSSSSLQTSSDSARESTGTRSSHSAPTLHTRPAFSQYW (SEQ ID NO: 896)
ORF2/3	MPWRPPVHSVQGREDQWFASFFHGHDSFCGCGDPLGHINSIAHRFPRAGPP RPPPGLDQPNPREQGPAGPGGPPAILALPAPPAEPDDPQPRRGGGDGGAAAG

[1752]

	AADDHTQRDYDEEELDELFRAAAEDDLSPIKAKQARLGVPRRKSRAKRMLF SPQSARGRRDPRRRRTSTPRKSPERGATPPAPAPETPPASPQTRAQARLYRHP PTPPGSPLEPGAHIAPPPYIPDLLFPNTGKKKKFSPFDWETEAQIAGWMRRPM RFYPSDTPHYPWLPPERDIPKICNINFKIKLQ (SEQ ID NO: 897)
ORF2t/3	MPWRPPVHSVQGREDQWSPIKAKQARLGVPRRKSRAKRMLFSPQSARGRR DPRRRRTSTPRKSPERGATPPAPAPETPPASPQTRAQARLYRHPPTPPGSPLEP GAHIAPPPYIPDLLFPNTGKKKKFSPFDWETEAQIAGWMRRPMRFYPSDTPH YPWLPPERDIPKICNINFKIKLQE (SEQ ID NO: 898)
TAIP	IVSRGERISGSQAFFTATIRFAAAVTLLAILIALLIAFLAPVHQGPLRG (SEQ ID NO: 899)
ORF1	TAWWWGRWRRRWRRRRPYTTRLRRRRARRAFPRRRRRRFVSRRWRRPYR RRRRRGRRRRRRRRRHKPTLILRQWQPDCIRHCKITGWMPLIICGKGSTQFN YITHADDITPRGASYGGNFTNMTFSLEAIYEQFLYHRNRWSASNHDLELCRY KGTTLKLYRHPEVDYIVTYSRTGPFEISHMTYLSTHPMLMLLNKHHIVVPSL KTKPRGRKAIKVRIRPPKLMNNKWYFTRDFCNIGLFQLWATGLELRNPWLR MSTLSPCIGFNVLKNSIYTNLSNLPQYKNERLNIINNILHPQEITGTNNKKWQ YTYTKLMAPIYYSANRASTYDWENYSKETNYNNTYVKFTQKRQEKLTKIR KEWQMLYPQQPTALPDSYDLLQEYGLYSPYYLNPTRINLDWMTPYTHVRY NPLVDKGFGNRIYIQWCSEADVSYNRTKSKCLLQDMPLFFMCYGYIDWAIK NTGVSSLVKDARICIRCPYTEPQLVGSTEDIGFVPISETFMRGDMPVLAPYIPL SWFCKWYPNIAHQKEVLESIISCSPFMPRDQDMNGWDITIGYKMDFLWGGS PLPSQPIDDPCQQGTHPIPDPDKHPRLLQVSNPKLLGPRTVFHKWDIRRGQFS KRSIKRVSEYSSDDESLAPGLPSKRNKLDSAFRGENREQKECYSLLKALEEE ETPEEEEPAPQEKAQKEELLHQLQLQRRHQRVLRRGLKLVFTDILRLRQGVH WNPELT (SEQ ID NO: 900)
ORF1/1	TAWWWGRWRRRWRRRRPYTTRLRRRRARRAFPRRRRRRFPIDDPCQQGT HPIPDPDKHPRLLQVSNPKLLGPRTVFHKWDIRRGQFSKRSIKRVSEYSSDDE SLAPGLPSKRNKLDSAFRGENREQKECYSLLKALEEEETPEEEEPAPQEKAQ KEELLHQLQLQRRHQRVLRRGLKLVFTDILRLRQGVHWNPELT (SEQ ID NO: 901)
ORF1/2	TAWWWGRWRRRWRRRRPYTTRLRRRRARRAFPRRRRRRFVSHQSETSSTR
	RSEEKIESKKNAILSSK RSRKKRPQKKKNQHPKKKPRKRSYSTSSSSRDATSESSDEGSSSSLQTSSDSA RESTGTRSSHSAPTLHTRPAFSQYW (SEQ ID NO: 902)

[1753] 表B4.示例性指环病毒核酸序列(甲型细环病毒)-进化分枝3

[1754] 名称 环6.0

[1755] 属/进化分枝 甲型细环病毒-进化分枝3

[1756] 登录号

[1757] 全序列:3828bp

1 10 20 30 40 50

GTGCTACGTCACTAACCTACGTGTCCGTCTCCCATAGGCCGGACACCGTA
TACGTCATACACTTCCTGGGCATGGTCTACGTGATAATATAAGTGGCTGC
ACTTCCGAATGGCTGAGTTTTCCACGCCCGTCCGCAGCGAGGACGCCACG
GAGGGGGATCCGCGCGTCCCGAGGGCGGGTGCCGGAGGTGAGTTTACACA
CCGCAGTCAAGGGGCAATTCGGGCTCGGGACTGGCCGGGCTATGGGCAAG
GCTCTTAAAAATGCACTTTTCTAGGTGCAGTAGAAAGAAAAGGACATTGT
CACTGCTACCACTGTACCATTCACAGAAAGCTAGGCCATCTGTGACAGGT
ATGTGGAGACCCCCGACTCGAAATGCGTTCAATATTCAACGTGACTGGTT
CTACAGTTGCTTTCACTCCCACGCTTCTATGTGCGGCTGTGCTGATTTTA
TTGGTCATTTCAATCATATCGCTGCTATGCTCGGCCGTCCGGAAGACCAG
AACCCTCCTCCGCCACCCGGGGCTCTGAGACCCCTACCCGCTCTCCCGGC
CTCTTCCGAGGCACCCGGTGATCGAGCGCCATGGCCTATGGGTGGTGGCG
GAGGCGACGGAGGCGCCCGTGGTGGAGGAGGAGATGGCGCCGCTGGAGAC
GCCGTCGGAGACCCCGCAGACGCCGACCTCGTCGCCGCTATCGACGCCGC
AGAACAGTAAGGAGGCGCGGCAGGGGGAGGTGGACTAGAGCACACAGGAG
[1758] ATGGCGCCGCAAGGGAAAACGCAGTCGCAAAAAAAAAGATTATTATAAGAC
AATGGCAGCCCAACTACACTCGCAGATGCAACATAGTGGGCTACATGCCT
CTACTAATATGTGGGGAAAATACTGTTGCTACAAACTATGCCACCCACTC
AGACGACAGCTACTACCCCGGACCCTTTGGGGGGGGAATGACTACAGACA
AATTTACTCTAAGAATACTGTATGATGAGTACAAAAGGTTCATGAACTAC
TGGACCTCTTCAAACGAGGACCTAGACCTATGTAGATACCTGGGATGCAC
TCTATATGTGTTTAGACACCCAGAAGTAGACTTTATAATCATTATAAATA
CCTCTCCTCCATTCCTAGACACAGAAATAACAGGGCCTAGCATACACCCA
GGTATGATGGCCCTTAACAAAAGAAGCAGATGGATACCTAGCATAAAAAA
CAGACCAGGCAGAAAGCACTATATAAAGATTAAAGTAGGAGCCCCCCGAA
TGTTCACAGATAAGTGGTACCCCCAAACAGACCTCTGTGACATGACACTC
CTAACGATCTTTGCCAGTGCGGCGGATATGCAATATCCGTTCGGCTCACC
ACTAACTGACACCATAGTTGTGTCATTCCAAGTTCTGCAATCCATGTACA
ACGACTGCCTGAGTGTACTTCCTGATAATTTTGCAGAGACATCAGGCAAA
GGCACCCAACTACATGAGAACATAATACAACATCTGCCCTACTACAACAC
CACACAAACACAAGCACAATTTAAAAGATTTATAGAAAACATGAATGCAA
CAAATGGAGACAATATATGGGCAAGCTACATAAACACAACCAAGTTCTCA
TCCGCAAACACTCCAAAGAATGACACAGGCATAGGAGGCCCTTACACTAC
ATATTCAGACTCATGGTACAAAGGCACAGTATACAATGACAAAATTAAAA

CCATACCAATAAAAGCAAGCAAGTTATACTACGAGCAAACCAAAAACCTC
ATTGGCATTACATTCACTGGATCCACACACAGACTCCATTACTGTGGAGG
CCTATACTCCTCCGTATGGCTATCAGCAGGTAGATCCTACTTTGAAACCA
AAGGCCCATACACAGACATAACTTACAACCCCTTTTCAGACAGAGGAGAG
GGTAACATGCTATGGATAGACTGGCTAACTAAAAATGACTCAGTGTACTC
AAAAACAAGTAGCAAGTGTCTTATAGAAAACCTGCCCCTGTGGGCCTCAG
TATACGGATATAAAGAATACTGCAGCAAGGTAACAGGAGACACAAACATA
GAACACAACTGTAGATGTGTTATCAGAAGCCCCTACACAGTACCACAACT
GTTAGACCACAACAATCCCTTCAGAGGATACGTGCCTTATAGCTTCAACT
TTGGAAATGGTAAAATGCCAGGCGGTAGCAGCCTAGTGCCCATTAGAATG
AGAGCCAAGTGGTACCCCACTCTGTTCCACCAAAAAGAAGTTCTAGAAGC
CATAGCACAGGCGGGCCCCTTCGCATACCACTCAGATATTAAAAAAGTGT
CCCTGGGCATAAAGTACAGATTTAAGTGGGTGTGGGGTGGCAACCCCGTG
TCCCAACAGGTTGTTAGAAACCCCTGCAAGACCACCCAAGGTTCCTCGGG
CAATAGAGTGCCTCGATCAATACAAGTCGTTGACCCGCGGTACAACACGC
CAGAACTCACCATACACGCGTGGGACTTCAGACATGGGTTCTTTGGCAGA
AAAGCTATTAAGAGAATGCAAGAACAACCAATACCTCATGACACTTTTTC
AGCAGGGTTCAAGCGCAGTCGCCGAGATACAGAAGCACTCCAATGCAGCC
AAGAAGAGCAACAAAAAGAAAACTTACTTTTCCCAGTCCAGCAGCTCAAG
CGAGTCCCCCCGTGGGAGACCTCGCAAGAGAGCCAAAGCGAGGAAGAAAA
CTCGCAAAAACAGGAGACCCTCTCCCAGCAACTCAGAGACCAGCTGCACA
[1759] AGCAGCGGCTCATGGGAGAGCAACTCCGATCGCTCCTCTACCAAATGCAG
AGGGTCCAACAAAATCAACACATAAACCCTATGTTATTGCCAAAGGGTCT
GGCATTAACTTCTATTTCTCACAATGTAATATAGATATGTTTGGTGACCC
CAAACCCTACAAGCCCTCCTCCAATGACTGGAAGGAGGAGTACGAGGCCG
CAAAGTACTGGGACAGACCCCCCAGACGCGACCTGAGGAGCACCCCCTTC
TACCCCTGGGCCCCCACCCCCAAACCATACAATGTCAACTTTGCCCTCAA
CTACAAATAAACGGTGGCCGTGGGAGTTTCACTTGTCGGTGTCTACCTCT
TAAGGTCACTAAGCACTCCGAGCGTAAGCGAGGAGTGCGACCCTTCACCA
AGGGCAACTCCCTCGAAGTCCGGCGCTACGCGCTTCGCGCTGCGCCGGAC
ATCTCGGACCCCCCCTCGACCCGAATCGCTTGCGCGATTCGGACCTGCGG
CCTCGGGGGGGTCGGGGGCTTTACTAAACAGACTCCGAGGTGCCATTGGA
CACTGAGGGGGTGAACAGCAACGAAAGTGAGTGGGGCCAGACTTCGCCAT
AAGGCCTTTATCTTCTTGCCATTTGTCCGCGACCGGGGGTCGCTCCTAGG
CGCGGACCCCGTTTCGGGGTCCTTCCGGGTTCATCGGCGCCGTTCCAGTG
ACGTCACGGGCGCCATGTTAAGTGGCTGTCGCCGAGGATTGACGTCACAG
TTCAAAGGTCATCCTCGGCGGTAACCGCAAACATGGCGGTCAATCTCTTC
CGGGTCAAAGGTCGTGCATACGTCATAAGTCACATGACAGGGGTCCACTT
AAACACGGAAGTAGGCCCCGACATGTGACTCGTCACGTGTGTACACGTCA
CGGCCGCCATTTTGTTTTACAAAATGGCCGACTTCCTTCCTGTTTTTTAA
AAAAAGGCGCGAAAAAACCGTCGGCGGGGGCCGCGCGCTGCGCGCGGGG
AGGCAATGCCTCCCCCCCCCCGCGCGCATGCGCGCGGGTCCCCCCCCCTC
CGGGGGGCTCCGCCCCCCGGCCCCCCCC (SEQ ID NO: 903)

[1760] 注释：

推定结构域	*碱基范围*
TATA 盒	85- 92
帽位点	109 - 116
转录起始位点	116
5' UTR 保守结构域	176 - 246
ORF2	351 - 710
ORF2/2	351 - 706; 2360 - 2825
ORF2/3	351 - 706; 2556 - 3060
TAIP	373 - 528
ORF1	581 - 2884
ORF1/1	581 - 706; 2360 - 2884
ORF1/2	581 - 706; 2556 - 2825
三开放阅读框区	2556 - 2821
聚(A)信号	3055 - 3061
GC 富集区	3720 - 3828

[1761]

[1762] 表C4.示例性指环病毒氨基酸序列(甲型细环病毒)进化分枝3

[1763]

环6.0（甲型细环病毒）	
ORF2	MWRPPTRNAFNIQRDWFYSCFHSHASMCGCADFIGHFNHIAAMLGRPEDQ NPPPPPGALRPLPALPASSEAPGDRAPWPMGGGGGDGGARGGGGDGAAGD AVGDPADADLVAAIDAAEQ (SEQ ID NO: 904)
ORF2/2	MWRPPTRNAFNIQRDWFYSCFHSHASMCGCADFIGHFNHIAAMLGRPEDQ NPPPPPGALRPLPALPASSEAPGDRAPWPMGGGGGDGGARGGGGDGAAGD AVGDPADADLVAAIDAAEQLLETPARPPKVPRAIECLDQYKSLTRGTTRQNS PYTRGTSDMGSLAEKLLRECKNNQYLMTLFQQGSSAVAEIQKHSNAAKKS NKKKTYFSQSSSSSESPRGRPRKRAKARKKTRKNRRPSPSNSETSCTSSGSW ESNSDRSSTKCRGSNKINT (SEQ ID NO: 905)
ORF2/3	MWRPPTRNAFNIQRDWFYSCFHSHASMCGCADFIGHFNHIAAMLGRPEDQ NPPPPPGALRPLPALPASSEAPGDRAPWPMGGGGGDGGARGGGGDGAAGD AVGDPADADLVAAIDAAEQVQAQSPRYRSTPMQPRRATKRKLTFPSPAAQA SPPVGDLAREPKRGRKLAKTGDPLPATQRPAAQAAAHGRATPIAPLPNAEG PTKSTHKPYVIAKGSGINFYFSQCNIDMFGDPKPYKPSSNDWKEEYEAAKY WDRPPRRDLRSTPFYPWAPTPKPYNVNFALNYK (SEQ ID NO: 906)
TAIP	MRSIFNVTGSTVAFTPTLLCAAVLILLVISIISLLCSAVRKTRTLLRHPGL (SEQ ID NO: 907)
ORF1	MAYGWWRRRRRRPWWRRRWRRWRRRRRPRRRRPRRRYRRRRTVRRRGR GRWTRAHRRWRRKGKRSRKKKIIIRQWQPNYTRRCNIVGYMPLLICGENTV

[1764]

	ATNYATHSDDSYYPGPFGGGMTTDKFTLRILYDEYKRFMNYWTSSNEDLD LCRYLGCTLYVFRHPEVDFIIIINTSPPFLDTEITGPSIHPGMMALNKRSRWIPS IKNRPGRKHYIKIKVGAPRMFTDKWYPQTDLCDMTLLTIFASAADMQYPFG SPLTDTIVVSFQVLQSMYNDCLSVLPDNFAETSGKGTQLHENIIQHLPYYNTT QTQAQFKRFIENMNATNGDNIWASYINTTKFSSANTPKNDTGIGGPYTTYSD SWYKGTVYNDKIKTIPIKASKLYYEQTKNLIGITFTGSTHRLHYCGGLYSSV WLSAGRSYFETKGPYTDITYNPFSDRGEGNMLWIDWLTKNDSVYSKTSSKC LIENLPLWASVYGYKEYCSKVTGDTNIEHNCRCVIRSPYTVPQLLDHNNPFR GYVPYSFNFGNGKMPGGSSLVPIRMRAKWYPTLFHQKEVLEAIAQAGPFAY HSDIKKVSLGIKYRFKWVVWGGNPVSQQVVRNPCKTTQGSSGNRVPRSIQVV DPRYNTPELTIHAWDFRHGFFGRKAIKRMQEQPIPHDTFSAGFKRSRRDTEA LQCSQEEQQKENLLFPVQQLKRVPPWETSQESQSEEENSQKQETLSQQLRD QLHKQRLMGEQLRSLLYQMQRVQQNQHINPMLLPKGLALTSISHNVI (SEQ ID NO: 908)
ORF1/1	MAYGWWRRRRRRPWWRRRWRRWRRRRRPRRRRPRRRYRRRRTVVRNPC KTTQGSSGNRVPRSIQVVDPRYNTPELTIHAWDFRHGFFGRKAIKRMQEQPI PHDTFSAGFKRSRRDTEALQCSQEEQQKENLLFPVQQLKRVPPWETSQESQS EEENSQKQETLSQQLRDQLHKQRLMGEQLRSLLYQMQRVQQNQHINPMLL PKGLALTSISHNVI (SEQ ID NO: 909)
ORF1/2	MAYGWWRRRRRRPWWRRRWRRWRRRRRPRRRRPRRRYRRRRTGSSAVA EIQKHSNAAKKSNKKKTYFSQSSSSSESPRGRPRKRAKARKKTRKNRRPSPS NSETSCTSSGSWESNSDRSSTKCRGSNKINT (SEQ ID NO: 910)

[1765] 表B5.示例性指环病毒核酸序列(甲型细环病毒)-进化分枝7

[1766] 名称 环7

[1767] 属/进化分枝 甲型细环病毒-进化分枝7

[1768] 登录号

[1769] 全序列:3815bp

[1770]
```
1          10         20         30         40         50
|          |          |          |          |          |
AAGATCGTCACTAACCACGTGACTCCTCTCGCCCAATCAGTGTCTACGTC
```

GTCCATTTCCTGGGCATGGTCTACATCCTGATATAAAGCGATGCACTTCC
GAATGGCTGAGTTTTCCACGCCCGTCCGCGGCGAGATCGCGACGGAGGAG
CGATCGAGCGTCCCGAGGGCGGGTGCCGGAGGTGAGTTTACACACCGCAG
TCAAGGGGCAATTCGGGCTCGGGACTGGCCGGGCTATGGGCAAGGCTCTT
AAAGCGTACGTCCCCCGCTATGTTTCTCGGCAGGGTGTGGAGGAAACAGA
AAAGGAAAGTGCTTCTGCTGGCTGTGCGAGCTACACAGAAAACATCTTCC
ATGAGTATCTGGCGTCCCCCCCTTGGGAATGTCTCCTACAGGGAGAGAAA
TTGGCTTCAGGCCGTCGAAACATCCCACAGTTCTTTTTGTGGCTGTGGTG
ATTTTATTCTTCATCTTACTAATTTGGCTGCACGCTTTGCTCTCCAGGGG
CCCCCGCCAGAGGGTGGTCCACCTCGGCCGAGGCCGCCGCTCCTGAGAGC
GCTGCCGGCCCCCGAGGTCCGCAGGGAGACGCGCACAGAGAACCGGGGCG
CCTCCGGTGAGCCATGGCCTGGCGATGGTGGTGGCAGAGACGATGGCGCC
GCCGCCGGTGGCCCCGCAGACGGTGGAGACGCCTACGACGCCGGAGACCT
AGACGACCTGTTCGCCGCCGTCGAAGAAGAACAACAGTAAGGAGGCGGAG
GTGGAGGGGCAGACGTGGGCGACGCACATACACCCGACGCGCGGTCAGAC
GCAGACGCAGACCCAGAAAGAGACTTGTACTGACTCAGTGGAGCCCCCAG
ACAGTCAGAAACTGCTCAATAAGGGGCATAGTGCCCATGGTAATATGCGG
ACACACAAAAGCAGGTAGAAACTATGCTATTCATAGCGAGGACTTCACCA
CACAGATACAACCCTTCGGGGGCAGTTTCAGCACGACCACCTGGTCCCTA
AAAGTGCTGTGGGACGAGCACCAGAAATTCCAGAACAGATGGTCCTACCC
AAACACACAACTAGACCTGGCCAGATACAGAGGGGTCACCTTCTGGTTCT
ACAGAGACCAGAAAACAGACTATATAGTACAGTGGAGTAGGAATCCCCCT
TTTAAACTCAATAAATACAGCAGTGCCATGTACCACCCGGGCATGATGAT
GCAGGCCAAAAGGAAACTAGTTGTACCTAGTTTCCAGACCAGACCCAAAG
GCAAGAAGAGATACAGAGTCACAATAAAACCCCCTAACATGTTTGCTGAC
AAGTGGTACACTCAAGAGGACCTGTGTCCGGTACCTCTTGTGCAAATTGT
GGTTTCTGCGGCGAGCCTGCTACATCCGTTCTGCCCACCACAAACGAACA
ACCCTTGCATCACCTTCCAGGTTTTGAAAGACATATATGATGAATGCATA
[1771] GGAGTTAACGAAACTATGAAAGATAAGTATAAGAAATTACAAACAACACT
ATACACCACTTGCACATACTATCAAACAACACAAGTACTGGCACAGCTAT
CTCCTGCCTTTCAACCTGCTATGAAACCTACTACTACACAATCAGCAGCT
ACAGCGACAACACTAGGAAACTATGTACCAGAGTTAAAGTACAACAATGG
CTCTTTTCACACAGGACAAAACGCAGTATTCGGCATGTGCTCATACAAAC
CAACAGACAGCATAATGACAAAAGCTAATGGCTGGTTTTGGCAAAACCTA
ATGGTAGACAACAACCTACATAGTTCTTATGGCAAGGCAACATTAGAATG
CATGGAGTATCACACAGGCATATACAGCTCTATATTTCTAAGTCCACAAA
GATCTTTAGAATTCCCAGCAGCATACCAAGACGTTACATACAACCCTAAC
TGTGATAGAGCAGTTGGAAACGTAGTTTGGTTTCAGTACAGCACTAAAAT
GGATACAAATTTTGATGAAACAAAATGTAAATGTGTCCTTAAAAACATTC
CACTGTGGGCGGCCTTCAATGGCTACTCAGACTTTATAATGCAAGAACTC
AGCATAAGTACAGAAATCCACAACTTTGGCATAGTGTGCTTTCAGTGCCC
GTACACTTTTCCCCCCTGTTTCAATAAAAACAAACCCCTAAAGGGGTACG
TGTTCTATGACACCACCTTTGGTAATGGAAAAATGCCAGACGGATCGGGG
CACGTACCCATCTACTGGCAGCAGAGATGGTGGATCAGACTAGCCTTCCA
GGTCCAGGTCATGCATGACTTTGTACTAACAGGCCCCTTTAGCTACAAAG
ATGACCTAGCAAACACCACACTCACAGCCAGATACAAATTTAAATTCAAA
TGGGGCGGCAATATCATCCCTGAACAGATTATCAAGAACCCGTGTCACAG
AGAGCAGTCCCTCGCTTCCTATCCCGATAGACAACGTCGCGACCTACAAG
TTGTTGACCCATCAACCATGGGCCCGATCTACACCTTCCACACATGGGAC
TGGCGACGGGGGCTTTTTGGTGCAGATGCTATCCAGAGAGTGTCACAAAA
ACCGGGAGATGCTCTCCGCTTTACAAACCCTTTCAAGAGACCCAGATATC
TTCCCCCGACAGACAGAGAAGACTACCGACAAGAAGAAGACTTCGCTTTA
CAGGAAAAAAGACGGCGCACATCCACAGAAGAAGCCCAGGACGAGGAGAG
CCCCCCGGAAAGCGCGCCGCTCCTACAGCAGCAGCAGCAGCAGCGGCAGC
TCTCAGTCCACCTCGCGGAGCAGCAGCGACTCGGAGTCCAACTCCGATAC
ATCCTCCAAGAAGTCCTCAAAACGCAAGCGGGTCTCCACCTAAACCCCCT
ATTATTAGGCCCGCCACAAACAAGGTCTATCTCTTTGAGCCCTCCAAAGG

CCTACTCCCCATAGTAGGAAAAGAGGCCTGGGAGGACGAGTACTGCACCT
GCAAGTACTGGGATCGCCCTCCCAGAACCAACCACCTAGACATCCCCACT
TATCCCTGGATGCCCACAAACTTCAAAGTCAGCTTCAAACTTGGATTTAA
ACCCTAAATAAAAATACAAGGCCGTACACTGTTCACTTGTCGGTGTCTAC
CTCTATAAGTCACTAAGCACTCCGAGCGCAGCGAGGAGTGCGACCCTCAG
CGGTGGGTGCAACGCCCTCGGCGGCCGCGCGCTACGCCTTCGGCTGCGCG
CGGCACCTCGGACCCCCGCTCGTGCTGACACGCTCGCGCGTGTCAGACCA
CTTCGGGCTCGCGGGGGTCGGGAATTTTGCTAAACAGACTCCGAGTTGCT
CTTGGACACTGTAGCTGTGAATCAGTAACGAAAGTGAGTGGGGCCAGACT
[1772] TCGCCATAAGGCCTTTATCTTCTTGCCATTGGTCCGTCTCGGGGGTCGCC
ATAGGCTTCGGGCTCGGTTTTAGGCCTTCCGGACTACCAAAATGGCGGAT
TCCGTGACGTCATGGCCGCCATTTTAAGTAAGGCGGAACAGGCTGTCACC
CCGTGTCAAAGTTCAGGGGTCAGCCTTCCGCTTTACACAAAATGGAGGTC
AATATCTTCCGGGTCAAAGGTCGCTACCGCGTCATAAGTCACGTGGGGAA
GGCTGCTGTGAATCCGGAAGTAGCTGACCCACGTGACTTGTCACGTGACT
AGCACGTCACGGCAGCCATTTTGAATCACAAAATGGCCGACTTCCTTCCT
CTTTTTTAAAAATAACGGCCCGGCGGCGGCGCGCGCTTCGCGCCGCTC
CGCCCCCCCCGCGCATGCGCGGGACCCCCCCCCGCGGGGGGCTCCGCCCC
CCGGTCCCCCCCCCG (SEQ ID NO: 911)

[1773] 注释：

推定结构域	*碱基范围*
TATA 盒	82- 87
帽位点	103 - 110
转录起始位点	110
5' UTR 保守结构域	170 - 240
ORF2	351 - 740
ORF2/2	351 - 737; 2378 - 2843
[1774] ORF2/3	351 - 737; 2526 - 3057
TAIP	379 - 543
ORF1	614 - 2911
ORF1/1	614 - 737; 2378 - 2911
ORF1/2	614 - 737; 2526 - 2843
三开放阅读框区	2526 - 2840
聚(A)信号	3056 - 3062
GC 富集区	3716 - 3815

[1775] 表C5.示例性指环病毒氨基酸序列(甲型细环病毒)-进化分枝7

[1776]

环7.0（甲型细环病毒）	
ORF2	MSIWRPPLGNVSYRERNWLQAVETSHSSFCGCGDFILHLTNLAARFALQGPP

[1777]

	PEGGPPRPRPPLLRA LPAPEVRRETRTENRGASGEPWPGDGGGRDDGAAAGGPADGGDAYDAGD LDDLFAAVEEEQQ (SEQ ID NO: 912)
ORF2/2	MSIWRPPLGNVSYRERNWLQAVETSHSSFCGCGDFILHLTNLAARFALQGPP PEGGPPRPRPPLLRA LPAPEVRRETRTENRGASGEPWPGDGGGRDDGAAAGGPADGGDAYDAGD LDDLFAAVEEEQQLSRTRVTESSPSLPIPIDNVATYKLLTHQPWARSTPSTHG TGDGGFLVQMLSRECHKNREMLSALQTLSRDPDIFPRQTEKTTDKKKTSLY RKKDGAHPQKKPRTRRAPRKARRSYSSSSSSGSSQSTSRSSSDSESNSDTSSK KSSKRKRVST (SEQ ID NO: 913)
ORF2/3	MSIWRPPLGNVSYRERNWLQAVETSHSSFCGCGDFILHLTNLAARFALQGPP PEGGPPRPRPPLLRA LPAPEVRRETRTENRGASGEPWPGDGGGRDDGAAAGGPADGGDAYDAGD LDDLFAAVEEEQQCYPESVTKTGRCSPLYKPFQETQISSPDRQRRLPTRRRLR FTGKKTAHIHRRSPGRGEPPGKRAAPTAAAAAAAALSPPRGAAATRSPTPIH PPRSPQNASGSPPKPPIIRPATNKVYLFEPSKGLLPIVGKEAWEDEYCTCKYW DRPPRTNHLDIPTYPWMPTNFKVSFKLGFKP (SEQ ID NO: 914)
TAIP	MSPTGREIGFRPSKHPTVLFVAVVILFFILLIWLHALLSRGPRQRVVHLGRGR RS (SEQ ID NO: 915)
ORF1	MAWRWWWQRRWRRRRWPRRRWRRLRRRRPRRPVRRRRRRTTVRRRRW RGRRGRRTYTRRAVRRRRRPRKRLVLTQWSPQTVRNCSIRGIVPMVICGHT KAGRNYAIHSEDFTTQIQPFGGSFSTTTWSLKVLWDEHQKFQNRWSYPNTQ LDLARYRGVTFWFYRDQKTDYIVQWSRNPPFKLNKYSSAMYHPGMMMQA KRKLVVPSFQTRPKGKKRYRVTIKPPNMFADKWYTQEDLCPVPLVQIVVSA ASLLHPFCPPQTNNPCITFQVLKDIYDECIGVNETMKDKYKKLQTTLYTTCT YYQTTQVLAQLSPAFQPAMKPTTTQSAATATTLGNYVPELKYNNGSFHTGQ NAVFGMCSYKPTDSIMTKANGWFWQNLMVDNNLHSSYGKATLECMEYHT GIYSSIFLSPQRSLEFPAAYQDVTYNPNCDRAVGNVVWFQYSTKMDTNFDE TKCKCVLKNIPLWAAFNGYSDFIMQELSISTEIHNFGIVCFQCPYTFPPCFNK NKPLKGYVFYDTTFGNGKMPDGSGHVPIYWQQRWWIRLAFQVQVMHDFV LTGPFSYKDDLANTTLTARYKFKFKWGGNIIPEQIIKNPCHREQSLASYPDRQ

[1778]

	RRDLQVVDPSTMGPIYTFHTWDWRRGLFGADAIQRVSQKPGDALRFTNPFK RPRYLPPTDREDYRQEEDFALQEKRRRTSTEEAQDEESPPESAPLLQQQQQQ RQLSVHLAEQQRLGVQLRYILQEVLKTQAGLHLNPLLLGPPQTRSISLSPPKA YSP (SEQ ID NO: 916)
ORF1/1	MAWRWWWQRRWRRRRWPRRRWRRLRRRRPRRPVRRRRRRRTTIIKNPCHR EQSLASYPDRQRRDLQVVDPSTMGPIYTFHTWDWRRGLFGADAIQRVSQKP GDALRFTNPFKRPRYLPPTDREDYRQEEDFALQEKRRRTSTEEAQDEESPPE SAPLLQQQQQQRQLSVHLAEQQRLGVQLRYILQEVLKTQAGLHLNPLLLGP PQTRSISLSPPKAYSP (SEQ ID NO: 917)
ORF1/2	MAWRWWWQRRWRRRRWPRRRWRRLRRRRPRRPVRRRRRRRTTMLSRECH KNREMLSALQTLSRDPDIFPRQTEKTTDKKKTSLYRKKDGAHPQKKPRTRR APRKARRSYSSSSSSGSSQSTSRSSSDSESNSDTSSKKSSKRKRVST (SEQ ID NO: 918)

[1779] 在一些实施例中，指环体包含核酸，所述核酸包含在PCT申请号PCT/US2018/037379(通过引用整体并入本文)中列出的序列。在一些实施例中，指环体包含多肽，所述多肽包含在PCT申请号PCT/US 2018/037379(通过引用整体并入本文)中列出的序列。

[1780] 在一些实施例中，指环体包含指环病毒基因组，例如，根据实例9中所述的方法所鉴定的。在一些实施例中，指环体包含指环病毒序列或其一部分，如实例13中所述。

[1781] 在一些实施例中，指环体包含遗传元件，所述遗传元件包含共有的指环病毒基序，例如，如表19所示。在一些实施例中，指环体包含遗传元件，所述遗传元件包含共有的指环病毒ORF1基序，例如，如表19所示。在一些实施例中，指环体包含遗传元件，所述遗传元件包含共有的指环病毒ORF1/1基序，例如，如表19所示。在一些实施例中，指环体包含遗传元件，所述遗传元件包含共有的指环病毒ORF1/2基序，例如，如表19所示。在一些实施例中，指环体包含遗传元件，所述遗传元件包含共有的指环病毒ORF2/2基序，例如，如表19所示。在一些实施例中，指环体包含遗传元件，所述遗传元件包含共有的指环病毒ORF2/3基序，例如，如表19所示。在一些实施例中，指环体包含遗传元件，所述遗传元件包含共有的指环病毒ORF2t/3基序，例如，如表19所示。在一些实施例中，如表19中所示，X表示任何氨基酸。在一些实施例中，如表19中所示，Z表示谷氨酸或谷氨酰胺。在一些实施例中，如表19所示，B表示天冬氨酸或天冬酰胺。在一些实施例中，如表19中所示，J表示亮氨酸或异亮氨酸。

[1782] 表19.指环病毒的开放阅读框(ORF)中的共有基序

[1783]

共有阈值	开放阅读框	位置	基序	SEQ ID NO:
50	ORF1	79	LIJRQWQPXXIRRCXIXGYXPLIXC	68
50	ORF1	111	NYXXHXD	69
50	ORF1	135	FSLXXLYDZ	70
50	ORF1	149	NXWTXSNXDLDLCRYXGC	71
50	ORF1	194	TXPSXHPGXMXLXKHK	72
50	ORF1	212	IPSLXTRPXG	73
50	ORF1	228	RIXPPXLFXDKWYFQXDL	74
50	ORF1	250	LLXIXATA	75
50	ORF1	260	LXXPFXSPXTD	76
50	ORF1	448	YNPXXDKGXGNXIW	77
50	ORF1	519	CPYTZPXL	78
50	ORF1	542	XFGXGXMP	79
50	ORF1	569	HQXEVXEX	80
50	ORF1	600	KYXFXFXWGGNP	81
50	ORF1	653	HSWDXRRG	82
50	ORF1	666	AIKRXQQ	83
50	ORF1	750	XQZQXXLR	84
50	ORF1/1	73	PRXJQXXDP	85
50	ORF1/1	91	HSWDXRRG	86
50	ORF1/1	105	AIKRXQQ	87
50	ORF1/1	187	QZQXXLR	88
50	ORF1/2	97	KXKRRRR	89
50	ORF2/2	158	PIXSLXXYKXXTR	90
50	ORF2/2	189	LAXQLLKECXKN	91
50	ORF2/3	39	HLNXLA	92
50	ORF2/3	272	DRPPR	93
50	ORF2/3	281	DXPFYPWXP	94
50	ORF2/3	300	VXFKLXF	95
50	ORF2t/3	4	WXPPVHBVXGIERXW	96
50	ORF2t/3	37	AKRKLX	97
50	ORF2t/3	140	PSSXDWXXEY	98
50	ORF2t/3	156	DRPPR	99
50	ORF2t/3	167	PFYPW	100
50	ORF2t/3	183	NVXFKLXF	101
50	ORF1	84	JXXXXWQPXXXXXCXIXGXXXJWQP	102
50	ORF1	149	NXWXXXNXXXXLXRY	103
50	ORF1	448	YNPXXDXG	104

[1784] ORF1分子

[1785] 在一些实施例中，指环体包含ORF1分子和/或编码ORF1分子的核酸。通常，ORF1分子包含具有指环病毒ORF1蛋白(例如，如本文所述的指环病毒ORF1蛋白，例如，如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所列)的结构特征和/或活性的多肽，或其功能片段。在一些实施例中，ORF1分子包含相对于指环病毒

ORF1蛋白(例如,如本文所述的指环病毒ORF1蛋白,例如,如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所列)的截短。在一些实施例中,ORF1分子被截短了指环病毒ORF1蛋白的至少10、20、30、40、50、60、70、80、90、100、150、200、250、300、350、400、450、500、550、600、650或700个氨基酸。在一些实施例中,ORF1分子包含与如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所示的指环病毒ORF1蛋白序列具有至少70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在一些实施例中,ORF1分子包含与例如如本文所述的甲型细环病毒、乙型细环病毒或丙型细环病毒ORF1蛋白具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%序列同一性的氨基酸序列。ORF1分子通常可以结合核酸分子,例如DNA(例如,遗传元件,例如,如本文所述)。在一些实施例中,ORF1分子定位于细胞核。在某些实施例中,ORF1分子定位于细胞的核仁。

[1786] 不希望受理论束缚,ORF1分子可能能够与其他ORF1分子结合,例如,形成蛋白质外部(例如,如本文所述)。这种ORF1分子可被描述为具有形成衣壳的能力。在一些实施例中,蛋白质外部可以包壳核酸分子(例如,如本文所述的遗传元件)。在一些实施例中,多个ORF1分子可形成多聚体,例如以产生蛋白质外部。在一些实施例中,多聚体可以是同多聚体。在其他实施例中,多聚体可以是异多聚体(例如,包含多个不同的ORF1分子)。还预期ORF1分子可具有复制酶活性。

[1787] 在一些实施例中,ORF1分子可以包含以下中的一种或多种:包含精氨酸富集区的第一区域,例如具有至少60%的碱性残基(例如,至少60%、65%、70%、75%、80%、85%、90%、95%或100%碱性残基;例如,60%-90%、60%-80%、70%-90%或70-80%的碱性残基)的区域,以及包含胶冻卷结构域的第二区域,例如至少六个β链(例如,4、5、6、7、8、9、10、11或12个β链)。

[1788] 精氨酸富集区

[1789] 精氨酸富集区与本文描述的精氨酸富集区序列具有至少70%(例如,至少约70%、80%、90%、95%、96%、97%、98%、99%或100%)序列同一性,或者包含至少60%、70%或80%碱性残基(例如,精氨酸、赖氨酸或其组合)的至少约40个氨基酸的序列。

[1790] 胶冻卷结构域

[1791] 胶冻卷结构域或区域包含以下(例如,由以下组成):包含以下特征中的一个或多个(例如,1、2或3个)的多肽(例如,包含在较大多肽中的结构域或区域):

[1792] (i)胶冻卷结构域的至少30%(例如,至少30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、90%或更多)的氨基酸是一个或多个β-折叠的部分;

[1793] (ii)胶冻卷结构域的二级结构包含至少四个(例如,至少4、5、6、7、8、9、10、11或12)个β-链;和/或

[1794] (iii)胶冻卷结构域的三级结构包含至少两个(例如,至少2、3或4个)β-折叠;和/或

[1795] (iv)胶冻卷结构域包含β-折叠与α-螺旋的比例为至少2:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1或10:1。

[1796] 在某些实施例中,胶冻卷结构域包含两个β-折叠。

[1797] 在某些实施例中,一个或多个(例如,1、2、3、4、5、6、7、8、9或10个)β-折叠包含约八

个(例如,4、5、6、7、8、9、10、11或12个)β-链。在某些实施例中,一个或多个(例如,1、2、3、4、5、6、7、8、9或10个)β-折叠包含八个β-链。在某些实施例中,一个或多个(例如,1、2、3、4、5、6、7、8、9或10个)β-折叠包含七个β-链。在某些实施例中,一个或多个(例如,1、2、3、4、5、6、7、8、9或10个)β-折叠包含六个β-链。在某些实施例中,一个或多个(例如,1、2、3、4、5、6、7、8、9或10个)β-折叠包含五个β-链。在某些实施例中,一个或多个(例如,1、2、3、4、5、6、7、8、9或10个)β-折叠包含四个β-链。

[1798] 在一些实施例中,胶冻卷结构域包含与第二β-折叠反平行取向的第一β-折叠。在某些实施例中,第一β-折叠包含约四个(例如,3、4、5或6个)β-链。在某些实施例中,第二β-折叠包含约四个(例如,3、4、5或6个)β-链。在实施例中,第一和第二β-折叠总共包含约八个(例如,6、7、8、9、10、11或12个)β-链。

[1799] 在某些实施例中,胶冻卷结构域是衣壳蛋白(例如,如本文所述的ORF1分子)的组分。在某些实施例中,胶冻卷结构域具有自组装活性。在一些实施例中,包含胶冻卷结构域的多肽结合包含胶卷结构域的多肽的另一拷贝。在一些实施例中,第一多肽的胶冻卷结构域结合多肽的第二拷贝的胶冻卷结构域。

[1800] ORF1分子还可包括包含指环病毒N22结构域(例如,如本文所述,例如来自如本文所述的指环病毒ORF1蛋白的N22结构域)的结构或活性的第三区域,和/或包含指环病毒C末端结构域(CTD)(例如,如本文所述,例如来自如本文所述的指环病毒ORF1蛋白的CTD)的结构或活性的第四区域。在一些实施例中,ORF1分子以N末端到C末端的顺序包含所述第一、第二、第三和第四区域。

[1801] 在一些实施例中,ORF1分子可以进一步包含高变区(HVR),例如来自例如本文所述的指环病毒ORF1蛋白的HVR。在一些实施例中,HVR位于所述第二区域和所述第三区域之间。在一些实施例中,HVR包含至少约55(例如,至少约45、50、51、52、53、54、55、56、57、58、59、60或65)个氨基酸(例如,约45-160、50-160、55-160、60-160、45-150、50-150、55-150、60-150、45-140、50-140、55-140或60-140个氨基酸)。

[1802] 在一些实施例中,第一区域可以结合核酸分子(例如,DNA)。在一些实施例中,碱性残基选自精氨酸、组氨酸或赖氨酸,或其组合。在一些实施例中,第一区域包含至少60%、65%、70%、75%、80%、85%、90%、95%或100%的精氨酸残基(例如,60%-90%、60%-80%、70%-90%或70-80%精氨酸残基)。在一些实施例中,第一区域包含约30-120个氨基酸(例如,约40-120、40-100、40-90、40-80、40-70、50-100、50-90、50-80、50-70、60-100、60-90或60-80个氨基酸)。在一些实施例中,第一区域包含病毒ORF1精氨酸富集区(例如,来自例如本文所述的指环病毒ORF1蛋白的精氨酸富集区)的结构或活性。在一些实施例中,第一区域包含核定位信号。

[1803] 在一些实施例中,第二区域包含胶冻卷结构域,例如病毒ORF1胶冻卷结构域的结构或活性(例如,来自例如如本文所述的指环病毒ORF1蛋白的胶冻卷结构域)。在一些实施例中,第二区域能够与另一个ORF1分子的第二区域结合,例如,形成蛋白质外部(例如衣壳)或其一部分。

[1804] 在一些实施例中,第四区域暴露在蛋白质外部(例如,包含例如如本文所述的ORF1分子的多聚体的蛋白质外部)的表面上。

[1805] 在一些实施例中,第一区域、第二区域、第三区域、第四区域和/或HVR各自包含少

于四个(例如,0、1、2或3)个β折叠。

[1806] 在一些实施例中,第一区域、第二区域、第三区域、第四区域和/或HVR中的一个或多个可以被异源氨基酸序列(例如来自异源ORF1分子的相应区域)替换。在一些实施例中,异源氨基酸序列具有所期望的功能性,例如,如本文所述。

[1807] 在一些实施例中,ORF1分子包含多个保守基序(例如,包含约5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45、50、60、70、80、90、100或更多个氨基酸的基序)(例如,如图34所示)。在一些实施例中,保守基序可显示与一个或多个野生型指环病毒进化分枝(例如,甲型细环病毒、进化分枝1;甲型细环病毒,进化分枝2;甲型细环病毒,进化分枝3;甲型细环病毒,进化分枝4;甲型细环病毒,进化分枝5;甲型细环病毒,进化分枝6;甲型细环病毒,进化分枝7;乙型细环病毒;和/或丙型细环病毒)的ORF1蛋白的60%、70%、80%、85%、90%、95%或100%序列同一性。在实施例中,保守基序各自具有1-1000(例如,介于5-10、5-15、5-20、10-15、10-20、15-20、5-50、5-100、10-50、10-100、10-1000、50-100、50-1000或100-1000)个氨基酸的长度。在某些实施例中,保守基序由ORF1分子序列的约2%-4%(例如,约1%-8%、1%-6%、1%-5%、1%-4%、2%-8%、2%-6%、2%-5%或2%-4%)组成,并且每个都显示与野生型指环病毒进化分枝的ORF1蛋白中相应基序的100%序列同一性。在某些实施例中,保守基序由ORF1分子序列的约5%-10%(例如,约1%-20%、1%-10%、5%-20%或5%-10%)组成,并且每个都显示与野生型指环病毒进化分枝的ORF1蛋白中相应基序的80%序列同一性。在某些实施例中,保守基序由ORF1分子序列的约10%-50%(例如,约10%-20%、10%-30%、10%-40%、10%-50%、20%-40%、20%-50%或30%-50%)组成,并且每个都显示与野生型指环病毒进化分枝的ORF1蛋白中相应基序的60%序列同一性。在一些实施例中,保守基序包含一个或多个如表19中所列的氨基酸序列。

[1808] 在一些实施例中,ORF1分子包含相对于例如如本文所述的野生型ORF1蛋白(例如,如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16、18、20-37或D1-D10中的任一个中所示)的至少一个差异(例如,突变、化学修饰或表观遗传改变)。

[1809] N22结构域中的保守ORF1基序

[1810] 在一些实施例中,本文所述的多肽(例如,ORF1分子)包含氨基酸序列YNPX^{2}DXGX^{2}N(SEQ ID NO:829),其中X^{n}是任何n个氨基酸的连续序列。例如,X^{2}表示任何两个氨基酸的连续序列。在一些实施例中,YNPX^{2}DXGX^{2}N(SEQ ID NO:829)包含在例如如本文所述的ORF1分子的N22结构域内。在一些实施例中,本文所述的遗传元件包含编码氨基酸序列YNPX^{2}DXGX^{2}N(SEQ ID NO:829)的核酸序列(例如,编码例如如本文所述的ORF1分子的核酸序列),其中X^{n}是任何n个氨基酸的连续序列。

[1811] 在一些实施例中,多肽(例如,ORF1分子)包含保守的二级结构,例如,侧接和/或包含YNPX^{2}DXGX^{2}N(SEQ ID NO:829)部分,例如在N22结构域中。在一些实施例中,保守二级结构包含第一β链和/或第二β链。在一些实施例中,第一β链的长度为约5-6个(例如,3、4、5、6、7或8个)氨基酸。在一些实施例中,第一β链包含位于YNPX^{2}DXGX^{2}N(SEQ ID NO:829)基序的N末端的酪氨酸(Y)残基。在一些实施例中,YNPX^{2}DXGX^{2}N(SEQ ID NO:829)基序包含无规卷曲(例如,约8-9个氨基酸的无规卷曲)。在一些实施例中,第二β链的长度为约7-8个(例如,5、6、7、8、9或10个)氨基酸。在一些实施例中,第二β链包含位于YNPX^{2}DXGX^{2}N(SEQ ID NO:829)基序的C末端的天冬酰胺(N)残基。

[1812] 示例性YNPX^{2}DXGX^{2}N(SEQ ID NO:829)基序侧翼二级结构在实例47和图48中描述。在一些实施例中,ORF1分子包含区域,所述区域包含图48中所示的二级结构元件(例如,β链)中的一个或多个(例如,1、2、3、4、5、6、7、8、9、10个或全部)。在一些实施例中,ORF1分子包含区域,所述区域包含图48中所示的侧翼于YNPX^{2}DXGX^{2}N(SEQ ID NO:829)基序(例如,如本文所述)的二级结构元件(例如,β链)中的一个或多个(例如,1、2、3、4、5、6、7、8、9、10个或全部)。

[1813] ORF1胶冻卷结构域中的保守二级结构基序

[1814] 在一些实施例中,本文所述的多肽(例如,ORF1分子)包含一个或多个由指环病毒ORF1蛋白(例如,如本文所述)所包含的二级结构元件。在一些实施方案中,ORF1分子包含一个或多个由指环病毒ORF1蛋白(例如,如本文所述)的胶冻卷结构域所包含的二级结构元件。通常,ORF1胶冻卷结构域包含二级结构,所述二级结构在N末端到C末端的方向上依次包含第一β链、第二β链、第一α螺旋、第三β链、第四β链、第五β链、第二α螺旋、第六β链、第七β链、第八β链和第九β链。在一些实施例中,ORF1分子包含二级结构,所述二级结构按从N末端到C末端的方向依次包含第一β链、第二β链、第一α螺旋、第三β链、第四β链、第五β链、第二α螺旋、第六β链、第七β链、第八β链和/或第九β链。

[1815] 在一些实施例中,一对保守的二级结构元件(即β链和/或α螺旋)被间隙氨基酸序列隔开,例如包含无规卷曲序列、β链或α螺旋,或其组合。保守二级结构元件之间的间隙氨基酸序列可包含例如1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30或更多个氨基酸。在一些实施例中,ORF1分子可进一步包含一个或多个另外的β链和/或α螺旋(例如,在胶冻卷域中)。在一些实施例中,可以组合连续的β链或连续的α螺旋。在一些实施例中,第一β链和第二β链包含在更大的β链中。在一些实施例中,第三β链和第四β链包含在更大的β链中。在一些实施例中,第四β链和第五β链包含在更大的β链中。在一些实施例中,第六β链和第七β链包含在更大的β链中。在一些实施例中,第七条β链和第八条β链包含在更大的β链中。在一些实施例中,第八β链和第九β链包含在更大的β链中。

[1816] 在一些实施例中,第一β链的长度为约5-7个(例如,3、4、5、6、7、8、9或10个)氨基酸。在一些实施例中,第二β链的长度为约15-16个(例如,13、14、15、16、17、18或19个)氨基酸。在一些实施例中,第一α螺旋的长度为约15-17个(例如,13、14、15、16、17、18、19或20个)氨基酸。在一些实施例中,第三β链的长度为约3-4个(例如,1、2、3、4、5或6个)氨基酸。在一些实施例中,第四β链的长度为约10-11个(例如,8、9、10、11、12或13个)氨基酸。在一些实施例中,第五β链的长度为约6-7个(例如,4、5、6、7、8、9或10个)氨基酸。在一些实施例中,第二α螺旋的长度为约8-14个(例如,5、6、7、8、9、10、11、12、13、14、15、16或17个)氨基酸。在一些实施例中,第二α螺旋可以分解成两个较小的α螺旋(例如,由无规卷曲序列隔开)。在一些实施例中,两个较小的α螺旋中的每一个的长度为约4-6个(例如,2、3、4、5、6、7或8个)氨基酸。在一些实施例中,第六β链的长度为约4-5个(例如,2、3、4、5、6或7个)氨基酸。在一些实施例中,第七β链的长度为约5-6个(例如,3、4、5、6、7、8或9个)氨基酸。在一些实施例中,第八β链的长度为约7-9个(例如,5、6、7、8、9、10、11、12或13个)氨基酸。在一些实施例中,第九β链的长度为约5-7个(例如,3、4、5、6、7、8、9或10个)氨基酸。

[1817] 示例性胶冻卷结构域二级结构在实例47和图47中描述。在一些实施例中,ORF1分

子包含区域，所述区域包含图47中所示的任何胶冻卷结构域二级结构的二级结构元件(例如，β链和/或α螺旋)中的一个或多个(例如，1、2、3、4、5、6、7、8、9、10或全部)。

[1818] 示例性ORF1序列

[1819] 在一些实施例中，本文所述的多肽(例如，ORF1分子)包含与例如如表20-37或D1-D10中的任一个中所述的一个或多个指环病毒ORF1子序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在一些实施例中，本文所述的指环体包含ORF1分子，所述ORF1分子包含与例如如表20-37或D1-D10中的任一个中所述的一个或多个指环病毒ORF1子序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在一些实施例中，本文所述的指环体包含编码ORF1分子的核酸分子(例如，遗传元件)，所述ORF1分子包含与例如如表20-37或D1-D10中的任一个中所述的一个或多个指环病毒ORF1子序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。

[1820] 在一些实施例中，一个或多个指环病毒ORF1子序列包含精氨酸(Arg)富集结构域、胶冻卷结构域、高变区(HVR)、N22结构域或C末端结构域(CTD)中的一个或多个(例如，如表20-37或D1-D10中的任一个中所列)，或与其具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的序列。在一些实施例中，ORF1分子包含来自不同指环病毒的多个子序列(例如，选自表20-37或D1-D10中所列的甲型细环病毒进化分枝1-7子序列的ORF1子序列的任何组合)。在实施例中，ORF1分子包含来自一种指环病毒的Arg富集结构域、胶冻卷结构域、N22结构域和CTD中的一个或多个，以及来自另一种的HVR。在实施例中，ORF1分子包含来自一种指环病毒的胶冻卷结构域、HVR、N22结构域和CTD中的一个或多个，以及来自另一种的Arg富集结构域。在实施例中，ORF1分子包含来自一种指环病毒的Arg富集结构域、HVR、N22结构域和CTD中的一个或多个，以及来自另一种的胶冻卷结构域。在实施例中，ORF1分子包含来自一种指环病毒的Arg富集结构域、胶冻卷结构域、HVR和CTD中的一个或多个，以及来自另一种的N22结构域。在实施例中，ORF1分子包含来自一种指环病毒的Arg富集结构域、胶冻卷结构域、HVR和N22结构域中的一个或多个，以及来自另一种的CTD。

[1821] 在实施例中，一个或多个指环病毒ORF1子序列包含与表20的Arg富集区氨基酸序列(例如，表20的氨基酸1-66)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表21的Arg富集区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表20的胶冻卷区氨基酸序列(例如，表20的氨基酸67-277)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表21的胶冻卷区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表20的HVR氨基酸序列(例如，表20的氨基酸278-347)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表21的HVR氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、

98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表20的N22结构域氨基酸序列(例如，表20的氨基酸348-513)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表21的N22结构域氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表20的CTD氨基酸序列(例如，表20的氨基酸513-680)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表21的CTD区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。

[1822] 在实施例中，一个或多个指环病毒ORF1子序列包含与表22的Arg富集区氨基酸序列(例如，表22的氨基酸1-69)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表23的Arg富集区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表22的胶冻卷区氨基酸序列(例如，表22的氨基酸70-279)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表23的胶冻卷区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表22的HVR氨基酸序列(例如，表22的氨基酸280-411)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表23的HVR氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表22的N22结构域氨基酸序列(例如，表22的氨基酸412-578)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表23的N22结构域氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表22的CTD氨基酸序列(例如，表22的氨基酸579-747)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表23的CTD区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。

[1823] 在实施例中，一个或多个指环病毒ORF1子序列包含与表24的Arg富集区氨基酸序列(例如，表24的氨基酸1-68)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表25的Arg富集区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表24的胶冻卷区氨基酸序列(例如，表24的氨基酸69-280)具有至少约

70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表25的胶冻卷区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表24的HVR氨基酸序列(例如，表24的氨基酸281-413)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表25的HVR氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表24的N22结构域氨基酸序列(例如，表24的氨基酸414-479)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表25的N22结构域氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表24的CTD氨基酸序列(例如，表24的氨基酸580-743)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表25的CTD区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。

[1824] 在实施例中，一个或多个指环病毒ORF1子序列包含与表26的Arg富集区氨基酸序列(例如，表26的氨基酸1-74)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表27的Arg富集区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表26的胶冻卷区氨基酸序列(例如，表26的氨基酸75-284)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表27的胶冻卷区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表26的HVR氨基酸序列(例如，表26的氨基酸285-445)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表27的HVR氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表26的N22结构域氨基酸序列(例如，表26的氨基酸446-611)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表27的N22结构域氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表26的CTD氨基酸序列(例如，表26的氨基酸612-780)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表27的CTD区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、

99％或100％序列同一性的氨基酸序列。

[1825] 在实施例中，一个或多个指环病毒ORF1子序列包含与表28的Arg富集区氨基酸序列(例如，表28的氨基酸1-75)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表29的Arg富集区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表28的胶冻卷区氨基酸序列(例如，表28的氨基酸75-284)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表29的胶冻卷区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表28的HVR氨基酸序列(例如，表28的氨基酸285-432)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表29的HVR氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表28的N22结构域氨基酸序列(例如，表28的氨基酸433-599)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表29的N22结构域氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表28的CTD氨基酸序列(例如，表28的氨基酸600-780)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表29的CTD区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。

[1826] 在实施例中，一个或多个指环病毒ORF1子序列包含与表30的Arg富集区氨基酸序列(例如，表30的氨基酸1-77)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表31的Arg富集区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表30的胶冻卷区氨基酸序列(例如，表30的氨基酸78-286)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表31的胶冻卷区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表30的HVR氨基酸序列(例如，表30的氨基酸287-416)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表31的HVR氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表30的N22结构域氨基酸序列(例如，表30的氨基酸417-585)具有至少约70％、

75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表31的N22结构域氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表30的CTD氨基酸序列(例如，表30的氨基酸586-746)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表31的CTD区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。

[1827] 在实施例中，一个或多个指环病毒ORF1子序列包含与表32的Arg富集区氨基酸序列(例如，表32的氨基酸1-74)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表33的Arg富集区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表32的胶冻卷区氨基酸序列(例如，表32的氨基酸75-286)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表33的胶冻卷区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表32的HVR氨基酸序列(例如，表32的氨基酸287-428)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表33的HVR氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表32的N22结构域氨基酸序列(例如，表32的氨基酸429-595)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表33的N22结构域氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表32的CTD氨基酸序列(例如，表32的氨基酸596-765)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表33的CTD区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。

[1828] 在实施例中，一个或多个指环病毒ORF1子序列包含与表34的Arg富集区氨基酸序列(例如，表34的氨基酸1-38)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表35的Arg富集区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表34的胶冻卷区氨基酸序列(例如，表34的氨基酸39-246)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表35的胶冻卷区氨基酸序列具有

至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表34的HVR氨基酸序列(例如，表34的氨基酸247-374)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表35的HVR氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表34的N22结构域氨基酸序列(例如，表34的氨基酸375-537)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表35的N22结构域氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表34的CTD氨基酸序列(例如，表34的氨基酸538-666)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表35的CTD区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。

[1829] 在实施例中，一个或多个指环病毒ORF1子序列包含与表36的Arg富集区氨基酸序列(例如，表36的氨基酸1-57)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表37的Arg富集区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表36的胶冻卷区氨基酸序列(例如，表36的氨基酸58-259)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表37的胶冻卷区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表36的HVR氨基酸序列(例如，表36的氨基酸260-351)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表37的HVR氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表36的N22结构域氨基酸序列(例如，表36的氨基酸352-510)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表37的N22结构域氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表36的CTD氨基酸序列(例如，表36的氨基酸511-673)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表37的CTD区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。

[1830] 在实施例中，一个或多个指环病毒ORF1子序列包含与表D1的Arg富集区氨基酸序

列(例如,表D1的氨基酸1-66)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在实施例中,一个或多个指环病毒ORF1子序列包含与表D2的Arg富集区氨基酸序列具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在实施例中,一个或多个指环病毒ORF1子序列包含与表D1的胶冻卷区氨基酸序列(例如,表D1的氨基酸67-277)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在实施例中,一个或多个指环病毒ORF1子序列包含与表D2的胶冻卷区氨基酸序列具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在实施例中,一个或多个指环病毒ORF1子序列包含与表D1的HVR氨基酸序列(例如,表D1的氨基酸278-347)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在实施例中,一个或多个指环病毒ORF1子序列包含与表D2的HVR氨基酸序列具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在实施例中,一个或多个指环病毒ORF1子序列包含与表D1的N22结构域氨基酸序列(例如,表D1的氨基酸348-513)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在实施例中,一个或多个指环病毒ORF1子序列包含与表D2的N22结构域氨基酸序列具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在实施例中,一个或多个指环病毒ORF1子序列包含与表D1的CTD氨基酸序列(例如,表D1的氨基酸513-680)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在实施例中,一个或多个指环病毒ORF1子序列包含与表D2的CTD区氨基酸序列具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。

[1831] 在实施例中,一个或多个指环病毒ORF1子序列包含与表D3的Arg富集区氨基酸序列(例如,表D3的氨基酸1-66)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在实施例中,一个或多个指环病毒ORF1子序列包含与表D4的Arg富集区氨基酸序列具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在实施例中,一个或多个指环病毒ORF1子序列包含与表D3的胶冻卷区氨基酸序列(例如,表D3的氨基酸67-277)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在实施例中,一个或多个指环病毒ORF1子序列包含与表D4的胶冻卷区氨基酸序列具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在实施例中,一个或多个指环病毒ORF1子序列包含与表D3的HVR氨基酸序列(例如,表D3的氨基酸278-347)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在实施例中,一个或多个指环病毒ORF1子序列包含与表D4的HVR氨基酸序列具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在实施例中,一个或多个指环病毒ORF1子序列包含与表D3的N22结构域氨基酸序列(例如,表D3的氨基酸348-513)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在实施例中,一个或多个指环病毒ORF1子序列包含与表D4的N22结构域氨基酸序列具有至少

约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表D3的CTD氨基酸序列(例如，表D3的氨基酸513-680)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表D4的CTD区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。

[1832] 在实施例中，一个或多个指环病毒ORF1子序列包含与表D5的Arg富集区氨基酸序列(例如，表D5的氨基酸1-66)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表D6的Arg富集区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表D5的胶冻卷区氨基酸序列(例如，表D5的氨基酸67-277)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表D6的胶冻卷区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表D5的HVR氨基酸序列(例如，表D5的氨基酸278-347)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表D6的HVR氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表D5的N22结构域氨基酸序列(例如，表D5的氨基酸348-513)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表D6的N22结构域氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表D5的CTD氨基酸序列(例如，表D5的氨基酸513-680)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表D6的CTD区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。

[1833] 在实施例中，一个或多个指环病毒ORF1子序列包含与表D7的Arg富集区氨基酸序列(例如，表D7的氨基酸1-57)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表D8的Arg富集区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表D7的胶冻卷区氨基酸序列(例如，表D7的氨基酸58-259)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表D8的胶冻卷区氨基酸序列具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中，一个或多个指环病毒ORF1子序列包含与表D7的HVR氨基酸序列(例

如,表D7的氨基酸260-351)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在实施例中,一个或多个指环病毒ORF1子序列包含与表D8的HVR氨基酸序列具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在实施例中,一个或多个指环病毒ORF1子序列包含与表D7的N22结构域氨基酸序列(例如,表D7的氨基酸352-510)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在实施例中,一个或多个指环病毒ORF1子序列包含与表D8的N22结构域氨基酸序列具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在实施例中,一个或多个指环病毒ORF1子序列包含与表D7的CTD氨基酸序列(例如,表D7的氨基酸511-673)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在实施例中,一个或多个指环病毒ORF1子序列包含与表D8的CTD区氨基酸序列具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。

[1834] 在实施例中,一个或多个指环病毒ORF1子序列包含与表D9的Arg富集区氨基酸序列(例如,表D9的氨基酸1-57)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在实施例中,一个或多个指环病毒ORF1子序列包含与表D10的Arg富集区氨基酸序列具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在实施例中,一个或多个指环病毒ORF1子序列包含与表D9的胶冻卷区氨基酸序列(例如,表D9的氨基酸58-259)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在实施例中,一个或多个指环病毒ORF1子序列包含与表D10的胶冻卷区氨基酸序列具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在实施例中,一个或多个指环病毒ORF1子序列包含与表D9的HVR氨基酸序列(例如,表D9的氨基酸260-351)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在实施例中,一个或多个指环病毒ORF1子序列包含与表D10的HVR氨基酸序列具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在实施例中,一个或多个指环病毒ORF1子序列包含与表D9的N22结构域氨基酸序列(例如,表D9的氨基酸352-510)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在实施例中,一个或多个指环病毒ORF1子序列包含与表D10的N22结构域氨基酸序列具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在实施例中,一个或多个指环病毒ORF1子序列包含与表D9的CTD氨基酸序列(例如,表D9的氨基酸511-673)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在实施例中,一个或多个指环病毒ORF1子序列包含与表D10的CTD区氨基酸序列具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。

[1835] 表20.示例性指环病毒ORF1氨基酸子序列(甲型细环病毒,进化分枝1)

[1836]

名称	CT30F
属/进化分枝	甲型细环病毒，进化分枝 1
株登录号	AB064597.1
蛋白质登录号	ANQ39351.1

[1837] 全序列:680AA

[1838]

```
1         10         20         30         40         50
|         |          |          |          |          |
TAWWWGRWRRRWRRRRPWRPRLRRRRARRAFPRRRRRRFVSRRWRRPYRR
RRRRGRRRRRRRRRHKPTLVLRQWQPDVIRHCKITGRMPLIICGKGSTQF
NYITHADDITPRGASYGGNFTNMTFSLEAIYEQFLYHRNRWSASNHDLEL
CRYKGTTLKLYRHPDVDYIVTYSRTGPFEISHMTYLSTHPLLMLLNKHHI
VVPSLKTKPRGRKAIKVRIRPPKLMNNKWYFTRDFCNIGLFQLWATGLEL
RNPWLRMSTLSPCIGFNVLKNSIYTNLSNLPQHREDRLNIINNTLHPHDI
TGPNNKKWQYTYTKLMAPIYYSANRASTYDLLREYGLYSPYYLNPTRINL
DWMTPYTHVRYNPLVDKGFGNRIYIQWCSEADVSYNRTKSKCLLQDMPLF
FMCYGYIDWAIKNTGVSSLARDARICIRCPYTEPQLVGSTEDIGFVPITE
TFMRGDMPVLAPYIPLSWFCKWYPNIAHQKEVLEAIISCSPFMPRDQGMN
GWDITIGYKMDFLWGGSPLPSQPIDDPCQQGTHPIPDPDKHPRLLQVSNP
KLLGPRTVFHKWDIRRGQFSKRSIKRVSEYSSDDESLAPGLPSKRNKLDS
AFRGENPEQKECYSLLKALEEEETPEEEEPAPQEKAQKEELLHQLQLQRR
HQRVLRRGLKLVFTDILRLRQGVHWNPELT (SEQ ID NO: 173)
```

[1839] 注释：

[1840]

推定结构域	*AA 范围*
Arg 富集区	1 - 66
胶冻卷结构域	67 - 277
高变区	278 - 347
N22	348 - 513
C 末端结构域	513 - 680

[1841]

[1842] 表21.示例性指环病毒ORF1氨基酸子序列(甲型细环病毒,进化分枝1)

[1843]

TTV-CT30F-ORF1（甲型细环病毒进化分枝 1）	
Arg 富集区	TAWWWGRWRRRWRRRRPWRPRLRRRRARRAFPRRRRRRFVSRRWR RPYRRRRRRGRRRRRRRRRHK (SEQ ID NO: 174)
胶冻卷结构域	PTLVLRQWQPDVIRHCKITGRMPLIICGKGSTQFNYITHADDITPRGASY GGNFTNMTFSLEAIYEQFLYHRNRWSASNHDLELCRYKGTTLKLYRHP DVDYIVTYSRTGPFEISHMTYLSTHPLLMLLNKHHIVVPSLKTKPRGRK AIKVRIRPPKLMNNKWYFTRDFCNIGLFQLWATGLELRNPWLRMSTLS PCIGFNVLKNSIYTNL (SEQ ID NO: 175)
高变结构域	SNLPQHREDRLNIINNTLHPHDITGPNNKKWQYTYTKLMAPIYYSANR ASTYDLLREYGLYSPYYLNPTR (SEQ ID NO: 176)
N22	INLDWMTPYTHVRYNPLVDKGFGNRIYIQWCSEADVSYNRTKSKCLL QDMPLFFMCYGYIDWAIKNTGVSSLARDARICIRCPYTEPQLVGSTEDI GFVPITETFMRGDMPVLAPYIPLSWFCKWYPNIAHQKEVLEAIISCSPFM PRDQGMNGWDITIGYKMDFL (SEQ ID NO: 177)
C 末端结构域	WGGSPLPSQPIDDPCQQGTHPIPDPDKHPRLLQVSNPKLLGPRTVFHKW DIRRGQFSKRSIKRVSEYSSDDESLAPGLPSKRNKLDSAFRGENPEQKEC YSLLKALEEEETPEEEEPAPQEKAQKEELLHQLQLQRRHQRVLRRGLK LVFTDILRLRQGVHWNPELT (SEQ ID NO: 178)

[1844] 表22.示例性指环病毒ORF1氨基酸子序列(甲型细环病毒,进化分枝2)

[1845]

名称	TTV-P13-1
属/进化分枝	甲型细环病毒，进化分枝 2
登录号	KT163896.1
蛋白质登录号	ANQ39351.1

[1846] 全序列:747AA

[1847]

```
1         10         20         30         40         50
|         |          |          |          |          |
MAYWWGRRRRWRRWRRRRRPLRRRRRWRRRRRWPRRRRWRRRRRRARPAR
RYRRRRGRRRVRRRRRPQKLVLTQWNPQTVRKCVIRGFLPLFFCGQGAYH
RNFTDHYDDVFPKGPSGGGHGSMVFNLSFLYQEFKKHHNKWSRSNLDFDL
VRYKGTVIKLYRHQDFDYIVWISRTPPFQESLLTVMTHQPSVMLQAKKCI
IVKSYRTHPGGKPYVTAKVRPPRLLTDKWYFQSDFCNVPLFSLQFALAEL
RFPICSPQTDTNCINFLVLDDIYYKFLDNKPKQSSDPNDENRIKFWHGLW
STMRYLNTTYINTLFPGTDSLVAAKDTDNSVNKYPSTATKQPYKDSQYMQ
NIWNTSKIHALYTWVAETNYKRLQAYYTQTYGGYQRQFFTGKQYWDYRVG
MFSPAFLSPSRLNPQNPGAYTEVSYNPWTDEGTGNVVCLQYLTKETSDYK
PGGGSKFCIEGVPLWAALVGYVDMCKKEGKDPGIRLNCLLLVKCPYTKPQ
LYDKKNPEKLFVPYSYNFGHGKMPGGDKYIPIEFKDRWYPCLLHQEEWIE
DIVRSGPFVPKDMPSSVTCMMRYSSLFNWGGNIIQEQAVEDPCKKGTFVV
PGTSGIARILQVSNPAKQTPTTTWHSWDWRRSLFTETGLKRMREQQPYDE
LSYTGPKKPKLSLPAGPAVPGAAVASSWWETKQVTSPDVSETETEAEAHQ
EEETEPEEGVQLQQLWEQQLLQKRQLGVVFQQLLRLRQGAEIHPGLV
(SEQ ID NO: 179)
```

[1848] 注释：

[1849]

推定结构域	*AA 范围*
Arg 富集区	1 - 69
胶冻卷结构域	70 - 279
高变区	280 - 411
N22	412 - 578
C 末端结构域	579 - 747

[1850] 表23.示例性指环病毒ORF1氨基酸子序列(甲型细环病毒,进化分枝2)

[1851]

TTV-P13-1-ORF1（甲型细环病毒进化分枝 2）	
Arg 富集区	MAYWWGRRRRWRRWRRRRRPLRRRRRWRRRRRWPRRRRWRRRRR RARPARRYRRRRGRRRVRRRRRPQK (SEQ ID NO: 180)
胶冻卷结构域	LVLTQWNPQTVRKCVIRGFLPLFFCGQGAYHRNFTDHYDDVFPKGPSG GGHGSMVFNLSFLYQEFKKHHNKWSRSNLDFDLVRYKGTVIKLYRHQ DFDYIVWISRTPPFQESLLTVMTHQPSVMLQAKKCIIVKSYRTHPGGKP YVTAKVRPPRLLTDKWYFQSDFCNVPLFSLQFALAELRFPICSPQTDTN CINFLVLDDIYYKFLDN (SEQ ID NO: 181)
高变结构域	KPKQSSDPNDENRIKFWHGLWSTMRYLNTTYINTLFPGTDSLVAAKDT DNSVNKYPSTATKQPYKDSQYMQNIWNTSKIHALYTWVAETNYKRLQ AYYTQTYGGYQRQFFTGKQYWDYRVGMFSPAFLSPSR (SEQ ID NO: 182)
N22	LNPQNPGAYTEVSYNPWTDEGTGNVVCLQYLTKETSDYKPGGGSKFCI EGVPLWAALVGYVDMCKKEGKDPGIRLNCLLLVKCPYTKPQLYDKK NPEKLFVPYSYNFGHGKMPGGDKYIPIEFKDRWYPCLLHQEEWIEDIVR SGPFVPKDMPSSVTCMMRYSSLFN (SEQ ID NO: 183)
C 末端结构域	WGGNIIQEQAVEDPCKKGTFVVPGTSGIARILQVSNPAKQTPTTTWHS

[1852]

	WDWRRSLFTETGLKRMREQQPYDELSYTGPKKPKLSLPAGPAVPGAA VASSWWETKQVTSPDVSETETEAEAHQEEETEPEEGVQLQQLWEQQL LQKRQLGVVFQQLLRLRQGAEIHPGLV (SEQ ID NO: 184)

[1853] 表24.示例性指环病毒ORF1氨基酸子序列(甲型细环病毒,进化分枝3)

[1854]

名称	TTV-tth8
属/进化分枝	*甲型细环病毒，进化分枝 3*
登录号	AJ620231.1
蛋白质登录号	CAF05750.1

[1855] 全序列:743AA

[1856]

```
1         10         20         30         40         50
|         |          |          |          |          |
MAWGWWKRRRRWWFRKRWTRGRLRRRWPRSARRRPRRRRVRRRRRWRRGR
RKTRTYRRRRRFRRRGRKAKLIIKLWQPAVIKRCRIKGYIPLIISGNGTF
ATNFTSHINDRIMKGPFGGGHSTMRFSLYILFEEHLRHMNFWTRSNDNLE
LTRYLGASVKIYRHPDQDFIVIYNRRTPLGGNIYTAPSLHPGNAILAKHK
ILVPSLQTRPKGRKAIRLRIAPPTLFTDKWYFQKDIADLTLFNIMAVEAD
LRFPFCSPQTDNTCISFQVLSSVYNNYLSINTFNNDNSDSKLKEFLNKAF
PTTGTKGTSLNALNTFRTEGCISHPQLKKPNPQINKPLESQYFAPLDALW
GDPIYYNDLNENKSLNDIIEKILIKNMITYHAKLREFPNSYQGNKAFCHL
TGIYSPPYLNQGRISPEIFGLYTEIIYNPYTDKGTGNKVWMDPLTKENNI
YKEGQSKCLLTDMPLWTLLFGYTDWCKKDTNNWDLPLNYRLVLICPYTFP
KLYNEKVKDYGYIPYSYKFGAGQMPDGSNYIPFQFRAKWYPTVLHQQQVM
EDISRSGPFAPKVEKPSTQLVMKYCFNFNWGGNPIIEQIVKDPSFQPTYE
IPGTGNIPRRIQVIDPRVLGPHYSFRSWDMRRHTFSRASIKRVSEQQETS
DLVFSGPKKPRVDIPKQETQEESSHSLQRESRPWETEEESETEALSQESQ
EVPFQQQLQQQYQEQLKLRQGIKVLFEQLIRTQQGVHVNPCLR
(SEQ ID NO: 185)
```

[1857] 注释：

[1858]

推定结构域	*AA 范围*
Arg 富集区	1 - 68
胶冻卷结构域	69 - 280
高变区	281 - 413
N22	414 - 579
C 末端结构域	580 - 743

[1859] 表25.示例性指环病毒ORF1氨基酸子序列(甲型细环病毒,进化分枝3)

[1860]

TTV-tth8-ORF1（甲型细环病毒进化分枝 3）	
Arg 富集区	MAWGWWKRRRRWWFRKRWTRGRLRRRWPRSARRRPRRRRVRRRRRWRRGRRKTRTYRRRRRFRRRGRK (SEQ ID NO: 186)
胶冻卷结构域	AKLIIKLWQPAVIKRCRIKGYIPLIISGNGTFATNFTSHINDRIMKGPFGGGHSTMRFSLYILFEEHLRHMNFWTRSNDNLELTRYLGASVKIYRHPDQDFIVIYNRRTPLGGNIYTAPSLHPGNAILAKHKILVPSLQTRPKGRKAIRLRIAPPTLFTDKWYFQKDIADLTLFNIMAVEADLRFPFCSPQTDNTCISFQVLSSVYNNYLSI (SEQ ID NO: 187)
高变结构域	NTFNNDNSDSKLKEFLNKAFPTTGTKGTSLNALNTFRTEGCISHPQLKKPNPQINKPLESQYFAPLDALWGDPIYYNDLNENKSLNDIIEKILIKNMITYHAKLREFPNSYQGNKAFCHLTGIYSPPYLNQGR (SEQ ID NO: 188)
N22	ISPEIFGLYTEIIYNPYTDKGTGNKVWMDPLTKENNIYKEGQSKCLLTDMPLWTLLFGYTDWCKKDTNNWDLPLNYRLVLICPYTFPKLYNEKVKDYGYIPYSYKFGAGQMPDGSNYIPFQFRAKWYPTVLHQQQVMEDISRSGPFAPKVEKPSTQLVMKYCFNFN (SEQ ID NO: 189)
C 末端结构域	WGGNPIIEQIVKDPSFQPTYEIPGTGNIPRRIQVIDPRVLGPHYSFRSWDMRRHTFSRASIKRVSEQQETSDLVFSGPKKPRVDIPKQETQEESSHSLQRESRPWETEEESETEALSQESQEVPFQQQLQQQYQEQLKLRQGIKVLFEQLIRTQQGVHVNPCLR (SEQ ID NO: 190)

[1861] 表26.示例性指环病毒ORF1氨基酸子序列(甲型细环病毒,进化分枝4)

[1862]

名称	TTV-HD20a
属/进化分枝	甲型细环病毒，进化分枝 4
登录号	FR751492.1
蛋白质登录号	NA

[1863] 全序列:780AA

[1864]

```
1         10         20         30         40         50
|         |          |          |          |          |
MAWWGWRRRWWRPKRRWRWRRARRRRRVPARRPRRAFRRYRTRTVRRRRR
GRRRGYRRRYRLRRYARRRFRRKKIVLTQWNPQTTRKCIIRGMMPVLWAG
MGTGGRNYAVRSDDYVVNKGFGGSFATETFSLKVLYDQFQRGFNRWSHTN
EDLDLARYRGCRWTFYRHKDTDFIVYFTNNPPMKTNQFSAPLTTPGMLMR
SKYKVLIPSFQTRPKGRKTVTVKIRPPKLFQDKWYTQQDLCSVPLVQLNV
TAADFTHPFGSPLTETPCVEFQVLGDLYNTCLNIDLPQFSELGEITSAYS
```

[1865]

```
KPNSNNLKELYKELFTKATSGHYWQTFITNSMVRAHIDADKAKEAQRAST
TPSYNNDPFPTIPVKSEFAQWKKKFTDTRDSPFLFATYHPEAIKDTIMKM
RENNFKLETGPNDKYGDYTAQYQGNTHMLDYYLGFYSPIFLSDGRSNVEF
FTAYRDIVYNPFLDKAQGNMVWFQYHTKTDNKFKKPECHWEIKDMPLWAL
LNGYVDYLETQIQYGDLSKEGKVLIRCPYTKPALVDPRDDTAGYVVYNRN
FGRGKWIDGGGYIPLHERTKWYVMLRYQTDVFHDIVTCGPWQYRDDNKNS
QLVAKYRFSFIWGGNTVHSQVIRNPCKDNQVSGPRRQPRDIQVVDPQRIT
PPWVLHSFDQRRGLFTETALRRLLQEPLPGEYAVSTLRTPLLFLPSEYQR
EDGAAESASGSPAKRPRIWSEESQTETISSEENPAETTRELLQRKLREQR
ALQFQLQHFAVQLAKTQANLHVNPLLSFPQ (SEQ ID NO: 191)
```

[1866] 注释：

[1867]

推定结构域	*AA 范围*
Arg 富集区	1 - 74
胶冻卷结构域	75 - 284
高变区	285 - 445
N22	446 - 611
C 末端结构域	612 - 780

[1868] 表27.示例性指环病毒ORF1氨基酸子序列(甲型细环病毒,进化分枝4)

[1869]

TTV-HD20a-ORF1（甲型细环病毒进化分枝 4）	
Arg 富集区	MAWWGWRRRWWRPKRRWRWRRARRRRRVPARRPRRAFRRYRTRT VRRRRRGRRRGYRRRYRLRRYARRRFRRKK (SEQ ID NO: 192)
胶冻卷结构域	IVLTQWNPQTTRKCIIRGMMPVLWAGMGTGGRNYAVRSDDYVVNKG FGGSFATETFSLKVLYDQFQRGFNRWSHTNEDLDLARYRGCRWTFYR HKDTDFIVYFTNNPPMKTNQFSAPLTTPGMLMRSKYKVLIPSFQTRPKG RKTVTVKIRPPKLFQDKWYTQQDLCSVPLVQLNVTAADFTHPFGSPLT ETPCVEFQVLGDLYNTCLNI (SEQ ID NO: 193)
高变结构域	DLPQFSELGEITSAYSKPNSNNLKELYKELFTKATSGHYWQTFITNSMV RAHIDADKAKEAQRASTTPSYNNDPFPTIPVKSEFAQWKKKFTDTRDSP FLFATYHPEAIKDTIMKMRENNFKLETGPNDKYGDYTAQYQGNTHML DYYLGFYSPIFLSDGR (SEQ ID NO: 194)
N22	SNVEFFTAYRDIVYNPFLDKAQGNMVWFQYHTKTDNKFKKPECHWEI KDMPLWALLNGYVDYLETQIQYGDLSKEGKVLIRCPYTKPALVDPRD DTAGYVVYNRNFGRGKWIDGGGYIPLHERTKWYVMLRYQTDVFHDI VTCGPWQYRDDNKNSQLVAKYRFSFI (SEQ ID NO: 195)
C 末端结构域	WGGNTVHSQVIRNPCKDNQVSGPRRQPRDIQVVDPQRITPPWVLHSFD QRRGLFTETALRRLLQEPLPGEYAVSTLRTPLLFLPSEYQREDGAAESA SGSPAKRPRIWSEESQTETISSEENPAETTRELLQRKLREQRALQFQLQH FAVQLAKTQANLHVNPLLSFPQ (SEQ ID NO: 196)

[1870] 表28.示例性指环病毒ORF1氨基酸子序列(甲型细环病毒,进化分枝5)

[1871]

名称	TTV-16 (TUS01)
属/进化分枝	甲型细环病毒，进化分枝 5

[1872]

登录号	AB017613.1
蛋白质登录号	BAA82454.1

[1873] 全序列:761AA

[1874]

```
1         10         20         30         40         50
|         |          |          |          |          |
MAYWFRRWGWRPRRRWRRWRRRRRRLPRRRTRRAVRGLGRRRKPRVRRRR
RTRRRTYRRGWRRRRYIRRGRRKKKLILTQWNPAIVKRCNIKGGLPIIIC
GEPRAAFNYGYHMEDYTPQPFPFGGGMSTVTFSLKALYDQYLKHQNRWTF
SNDQLDLARYRGCKLRFYRSPVCDFIVHYNLIPPLKMNQFTSPNTHPGLL
MLSKHKIIIPSFQTRPGGRRFVKIRLNPPKLFEDKWYTQQDLCKVPLVSI
TATAADLRYPFCSPQTNNPCTTFQVLRKNYNTVIGTSVKDQESTQDFENW
LYKTDSHYQTFATEAQLGRIPAFNPDGTKNTKQQSWQDNWSKKNSPWTGN
SGTYPQTTSEMYKIPYDSNFGFPTYRAQKDYILERRQCNFNYEVNNPVSK
KVWPQPSTTTPTVDYYEYHCGWFSNIFIGPNRYNLQFQTAYVDTTYNPLM
DKGKGNKIWFQYLSKKGTDYNEKQCYCTLEDMPLWAICFGYTDYVETQLG
PNVDHETAGLIIMICPYTQPPMYDKNRPNWGYVVYDTNFGNGKMPSGSGQ
VPVYWQCRWRPMLWFQQQVLNDISKTGPYAYRDEYKNVQLTLYYNFIFNW
GGDMYYPQVVKNPCGDSGIVPGSGRFTREVQVVSPLSMGPAYIFHYFDSR
RGFFSEKALKRMQQQQEFDESFTFKPKRPKLSTAAAEILQLEEDSTSGEG
KSPLQQEEKEVEVLQTPTVQLQLQRNIQEQLAIKQQLQFLLLQLLKTQSN
LHLNPQFLSPS (SEQ ID NO: 197)
```

[1875] 注释：

[1876]

推定结构域	*AA 范围*
Arg 富集区	1 - 75
胶冻卷结构域	75 - 284
高变区	285 - 432
N22	433 - 599
C 末端结构域	600 - 780

[1877] 表29.示例性指环病毒ORF1氨基酸子序列(甲型细环病毒,进化分枝5)

[1878]

TTV-16(TUS01)-ORF1（甲型细环病毒进化分枝 5）	
Arg 富集区	MAYWFRRWGWRPRRRWRRWRRRRRRLPRRRTRRAVRGLGRRRKPR VRRRRRTRRRTYRRGWRRRRYIRRGRRKKK (SEQ ID NO: 198)
胶冻卷结构域	LILTQWNPAIVKRCNIKGGLPIIICGEPRAAFNYGYHMEDYTPQPFPFGG GMSTVTFSLKALYDQYLKHQNRWTFSNDQLDLARYRGCKLRFYRSPV CDFIVHYNLIPPLKMNQFTSPNTHPGLLMLSKHKIIIPSFQTRPGGRRFV KIRLNPPKLFEDKWYTQQDLCKVPLVSITATAADLRYPFCSPQTNNPCT TFQVLRKNYNTVI (SEQ ID NO: 199)

[1879]

高变结构域	GTSVKDQESTQDFENWLYKTDSHYQTFATEAQLGRIPAFNPDGTKNTK QQSWQDNWSKKNSPWTGNSGTYPQTTSEMYKIPYDSNFGFPTYRAQK DYILERRQCNFNYEVNNPVSKKVWPQPSTTTPTVDYYEYHCGWFSNIFI GPNR (SEQ ID NO: 200)
N22	YNLQFQTAYVDTTYNPLMDKGKGNKIWFQYLSKKGTDYNEKQCYCT LEDMPLWAICFGYTDYVETQLGPNVDHETAGLIIMICPYTQPPMYDKN RPNWGYVVYDTNFGNGKMPSGSGQVPVYWQCRWRPMLWFQQQVLN DISKTGPYAYRDEYKNVQLTLYYNFIFN (SEQ ID NO: 201)
C 末端结构域	WGGDMYYPQVVKNPCGDSGIVPGSGRFTREVQVVSPLSMGPAYIFHY FDSRRGFFSEKALKRMQQQQEFDESFTFKPKRPKLSTAAAEILQLEEDS TSGEGKSPLQQEEKEVEVLQTPTVQLQLQRNIQEQLAIKQQLQFLLLQL LKTQSNLHLNPQFLSPS (SEQ ID NO: 202)

[1880] 表30.示例性指环病毒ORF1氨基酸子序列(甲型细环病毒,进化分枝6)

[1881]

名称	TTV-TJN02
属/进化分枝	*甲型细环病毒，进化分枝 6*
登录号	AB028669.1
蛋白质登录号	BAA94878.1

[1882] 全序列:746AA

[1883]

```
1         10         20         30         40         50
|         |          |          |          |          |
MAWGWWRWRRRWPARRWRRRRRRRPVRRTRARRPARRYRRRRTVRTRRRR
WGRRRYRRGWRRRTYVRKGRHRKKKKRLILRQWQPATRRRCTITGYLPIV
FCGHTRGNKNYALHSDDYTPQGQPFGGALSTTSFSLKVLFDQHQRGLNKW
SFPNDQLDLARYRGCKFIFYRTKQTDWVGQYDISEPYKLDKYSCPNYHPG
NMIKAKHKFLIPSYDTNPRGRQKIIVKIPPPDLFVDKWYTQEDLCSVNLV
SLAVSAASFLHPFGSPQTDNPCYTFQVLKEFYYQAIGFSASTQAMTSVLD
TLYTQNSYWESNLTQFYVLNAKKGSDTTQPLTSNMPTREEFMAKKNTNYN
WYTYKAASVKNKLHQMRQTYFEELTSKGPQTTKSEEGYSQHWTTPSTNAY
EYHLGMFSAIFLAPDRPVPRFPCAYQDVTYNPLMDKGVGNHIWFQYNTKA
DTQLIVTGGSCKAHIQDIPLWAAFYGYSDFIESELGPFVDAETVGLVCVI
CPYTKPPMYNKTNPAMGYVFYDRNFGDGKWTDGRGKIEPYWQVRWRPEML
FQETVMADLVQTGPFSYKDELKNSTLVCKYKFYFTWGGNMMFQQTIKNPC
KTDGQPTDSSRHPRGIQVADPEQMGPRWVFHSFDWRRGYLSEKALKRLQE
KPLDYDEYFTQPKRPRIFPPTESAEGEFREPEKGSYSEEERSQASAEEQT
QEATVLLLKRRLREQQQLQQQLQFLTREMFKTQAGLHLNPMLLNQR
(SEQ ID NO: 203)
```

[1884] 注释：

[1885]

推定结构域	*AA 范围*
Arg 富集区	1 - 77
胶冻卷结构域	78 - 286
高变区	287 - 416
N22	417 - 585
C 末端结构域	586 - 746

[1886]

[1887] 表31.示例性指环病毒ORF1氨基酸子序列(甲型细环病毒,进化分枝6)

[1888]

TTV-TJN02-ORF1（甲型细环病毒进化分枝 6）	
Arg 富集区	MAWGWWRWRRRWPARRWRRRRRRRPVRRTRARRPARRYRRRRTVR TRRRRWGRRRYRRGWRRRTYVRKGRHRKKKKR (SEQ ID NO: 204)
胶冻卷结构域	LILRQWQPATRRRCTITGYLPIVFCGHTRGNKNYALHSDDYTPQGQPFG GALSTTSFSLKVLFDQHQRGLNKWSFPNDQLDLARYRGCKFIFYRTKQ TDWVGQYDISEPYKLDKYSCPNYHPGNMIKAKHKFLIPSYDTNPRGRQ KIIVKIPPPDLFVDKWYTQEDLCSVNLVSLAVSAASFLHPFGSPQTDNPC YTFQVLKEFYYQAI (SEQ ID NO: 205)
高变结构域	GFSASTQAMTSVLDTLYTQNSYWESNLTQFYVLNAKKGSDTTQPLTSN MPTREEFMAKKNTNYNWYTYKAASVKNKLHQMRQTYFEELTSKGPQ TTKSEEGYSQHWTTPSTNAYEYHLGMFSAIFLAPDR (SEQ ID NO: 206)
N22	PVPRFPCAYQDVTYNPLMDKGVGNHIWFQYNTKADTQLIVTGGSCKA HIQDIPLWAAFYGYSDFIESELGPFVDAETVGLVCVICPYTKPPMYNKT NPAMGYVFYDRNFGDGKWTDGRGKIEPYWQVRWRPEMLFQETVMA DLVQTGPFSYKDELKNSTLVCKYKFYFT (SEQ ID NO: 207)
C 末端结构域	WGGNMMFQQTIKNPCKTDGQPTDSSRHPRGIQVADPEQMGPRWVFHS FDWRRGYLSEKALKRLQEKPLDYDEYFTQPKRPRIFPPTESAEGEFREP EKGSYSEEERSQASAEEQTQEATVLLLKRRLREQQQLQQQLQFLTREM FKTQAGLHLNPMLLNQR (SEQ ID NO: 208)

[1889] 表32.示例性指环病毒ORF1氨基酸子序列(甲型细环病毒,进化分枝7)

[1890]

名称	TTV-HD16d
属/进化分枝	甲型细环病毒，进化分枝 7
登录号	FR751479.1
蛋白质登录号	NA

[1891] 全序列:765AA

[1892]

```
1         10         20         30         40         50
|         |          |          |          |          |
MAWSWWWQRWRRRRWKPRRRRWRRLRWRRPRRAVRRRRRGRRVRRRRWAR
RRGRRRRYATRRKRRYRGRRFKKKLVLTQWHPNTMRRCLIKGIVPLVICG
HTRWNYNYALHSKDYTEEGRYPHGGALSTTTWSLKVLYDEHLKHHDFWGY
PNNQLDLARYKGAKFTFYRHKKTDFIIFFNRKPPFKLNKYSCASYHPGML
MQQRHKILLPSYETKPKGRPKITVRIKPPTLLEDKWYTQQDLCDVNLLQL
VVTAADFRHPLCSPQTNTPTTTFQVLKDIYYDTMSISEPTDSYTSVNNKS
TTQTFTNYSNTLENILYTRASYWNSFHATEYLNPNIIYKNGEKLFKEHED
LITWMTQTNNTGFLTKNNTAFGNNSYRPNADKIKKARKTYWNALIGTNDL
ATNIGQARAERFEYHLGWYSPIFLSRHRSNMNFARAYQDVTYNPNCDRGV
NNRVWVQPLTKPTTEFDEKRCKCVVQHLPLWAALYCYQDFVEEELGSSSE
ILNSCLLVLQCPYTFPPMYDKKLPDKGFVFYDSLFGDGKMSDGRGQVDIF
WQQRWYPRLATQMQVMHDITMTGPFSYRDELVSTQLTAKYTFDFMWGGNM
ISTQIIKNPCKDSGLEPAYPGRQRRDLQIVDPYSMGPQFSFHNWDYRHGL
FGQDAIDRVSKQPKDDADYPNPYKRPRYFPPTDQAAQEQEKDFSFLKTAP
SNSEESDQEVLQETQVLRFQPEQHKQLHLQLAERQRIGEQLRYLLQQMFK
TQANLHLNPYTFTQL (SEQ ID NO: 209)
```

[1893] 注释：

[1894]

推定结构域	*AA 范围*
Arg 富集区	1 - 74
胶冻卷结构域	75 - 286
高变区	287 - 428
N22	429 - 595
C 末端结构域	596 - 765

[1895] 表33.示例性指环病毒ORF1氨基酸子序列(甲型细环病毒,进化分枝7)

[1896]

TTV-HD16d-ORF1（甲型细环病毒进化分枝 7）	
Arg 富集区	MAWSWWWQRWRRRRWKPRRRRWRRLRWRRPRRAVRRRRRGRRVR RRRWARRRGRRRRYATRRKRRYRGRRFKKK (SEQ ID NO: 210)
胶冻卷结构域	LVLTQWHPNTMRRCLIKGIVPLVICGHTRWNYNYALHSKDYTEEGRYP HGGALSTTTWSLKVLYDEHLKHHDFWGYPNNQLDLARYKGAKFTFY RHKKTDFIIFFNRKPPFKLNKYSCASYHPGMLMQQRHKILLPSYETKPK GRPKITVRIKPPTLLEDKWYTQQDLCDVNLLQLVVTAADFRHPLCSPQ TNTPTTTFQVLKDIYYDTMSI (SEQ ID NO: 211)
高变结构域	SEPTDSYTSVNNKSTTQTFTNYSNTLENILYTRASYWNSFHATEYLNPN IIYKNGEKLFKEHEDLITWMTQTNNTGFLTKNNTAFGNNSYRPNADKI KKARKTYWNALIGTNDLATNIGQARAERFEYHLGWYSPIFLSRHR (SEQ ID NO: 212)
N22	SNMNFARAYQDVTYNPNCDRGVNNRVWVQPLTKPTTEFDEKRCKCV VQHLPLWAALYCYQDFVEEELGSSSEILNSCLLVLQCPYTFPPMYDKK LPDKGFVFYDSLFGDGKMSDGRGQVDIFWQQRWYPRLATQMQVMHD ITMTGPFSYRDELVSTQLTAKYTFDFM (SEQ ID NO: 213)

[1897]

C 末端结构域	WGGNMISTQIIKNPCKDSGLEPAYPGRQRRDLQIVDPYSMGPQFSFHN WDYRHGLFGQDAIDRVSKQPKDDADYPNPYKRPRYFPPTDQAAQEQE KDFSFLKTAPSNSEESDQEVLQETQVLRFQPEQHKQLHLQLAERQRIGE QLRYLLQQMFKTQANLHLNPYTFTQL (SEQ ID NO: 214)

[1898] 表34.示例性指环病毒ORF1氨基酸子序列(乙型细环病毒)

[1899]

名称	TTMV-LY2
属/进化分枝	乙型细环病毒
登录号	JX134045.1
蛋白质登录号	AGG91484.1

[1900] 全序列:666AA

[1901]

```
1         10         20         30         40         50
|         |          |          |          |          |
MPYYYRRRRYNYRRPRWYGRGWIRRPFRRRFRRKRRVRPTYTTIPLKQWQ
PPYKRTCYIKGQDCLIYYSNLRLGMNSTMYEKSIVPVHWPGGGSFSVSML
TLDALYDIHKLCRNWWTSTNQDLPLVRYKGCKITFYQSTFTDYIVRIHTE
LPANSNKLTYPNTHPLMMMMSKYKHIIPSRQTRRKKKPYTKIFVKPPPQF
ENKWYFATDLYKIPLLQIHCTACNLQNPFVKPDKLSNNVTLWSLNTISIQ
NRNMSVDQGQSWPFKILGTQSFYFYFYTGANLPGDTTQIPVADLLPLTNP
RINRPGQSLNEAKITDHITFTEYKNKFTNYWGNPFNKHIQEHLDMILYSL
KSPEAIKNEWTTENMKWNQLNNAGTMALTPFNEPIFTQIQYNPDRDTGED
TQLYLLSNATGTGWDPPGIPELILEGFPLWLIYWGFADFQKNLKKVTNID
TNYMLVAKTKFTQKPGTFYLVILNDTFVEGNSPYEKQPLPEDNIKWYPQV
QYQLEAQNKLLQTGPFTPNIQGQLSDNISMFYKFYFKWGGSPPKAINVEN
PAHQIQYPIPRNEHETTSLQSPGEAPESILYSFDYRHGNYTTTALSRISQ
DWALKDTVSKITEPDRQQLLKQALECLQISEETQEKKEKEVQQLISNLRQ
QQQLYRERIISLLKDQ (SEQ ID NO: 215)
```

[1902] 注释:

[1903]

推定结构域	AA 范围
Arg 富集区	1 - 38
胶冻卷结构域	39 - 246
高变区	247 - 374
N22	375 - 537
C 末端结构域	538 - 666

[1904] 表35.示例性指环病毒ORF1氨基酸子序列(乙型细环病毒)

[1905]

TTV-HD16d-ORF1（乙型细环病毒）	
Arg 富集区	MPYYYRRRRYNYRRPRWYGRGWIRRPFRRRFRRKRRVR (SEQ ID NO: 216)
胶冻卷结构域	PTYTTIPLKQWQPPYKRTCYIKGQDCLIYYSNLRLGMNSTMYEKSIVPV HWPGGGSFSVSMLTLDALYDIHKLCRNWWTSTNQDLPLVRYKGCKIT FYQSTFTDYIVRIHTELPANSNKLTYPNTHPLMMMMSKYKHIIPSRQTR RKKKPYTKIFVKPPPQFENKWYFATDLYKIPLLQIHCTACNLQNPFVKP DKLSNNVTLWSLNT (SEQ ID NO: 217)
高变结构域	ISIQNRNMSVDQGQSWPFKILGTQSFYFYFYTGANLPGDTTQIPVADLL PLTNPRINRPGQSLNEAKITDHITFTEYKNKFTNYWGNPFNKHIQEHLD MILYSLKSPEAIKNEWTTENMKWNQLNNAG (SEQ ID NO: 218)
N22	TMALTPFNEPIFTQIQYNPDRDTGEDTQLYLLSNATGTGWDPPGIPELIL EGFPLWLIYWGFADFQKNLKKVTNIDTNYMLVAKTKFTQKPGTFYLVI LNDTFVEGNSPYEKQPLPEDNIKWYPQVQYQLEAQNKLLQTGPFTPNI QGQLSDNISMFYKFYFK (SEQ ID NO: 219)
C 末端结构域	WGGSPPKAINVENPAHQIQYPIPRNEHETTSLQSPGEAPESILYSFDYRH GNYTTTALSRISQDWALKDTVSKITEPDRQQLLKQALECLQISEETQEK KEKEVQQLISNLRQQQQLYRERIISLLKDQ (SEQ ID NO: 220)

[1906] 表36.示例性指环病毒ORF1氨基酸子序列(丙型细环病毒)

[1907]

名称	TTMDV-MD1-073
属/进化分枝	丙型细环病毒
登录号	AB290918.1
蛋白质登录号	BAG49427.1

[1908] 全序列:673AA

[1909]

```
1         10         20         30         40         50
|         |          |          |          |          |
MPFWWGRRNKFWYGRNYRRKKRRFPKRRKRRFYRRTKYRRPARRRRRRRR
KVRRKKKTLIVRQWQPDSIVLCKIKGYDSIIWGAEGTQFQCSTHEMYEYT
RQKYPGGGGFGVQLYSLEYLYDQWKLRNNIWTKTNQLKDLCRYLKCVMTF
YRHQHIDFVIVYERQPPFEIDKLTYMKYHPYMLLQRKHKIILPSQTTNPR
GKLKKKKTIKPPKQMLSKWFFQQQFAKYDLLLIAAAACSLRYPRIGCCNE
NRMITLYCLNTKFYQDTEWGTTKQAPHYFKPYATINKSMIFVSNYGGKKT
EYNIGQWIETDIPGEGNLARYYRSISKEGGYFSPKILQAYQTKVKSVDYK
PLPIVLGRYNPAIDDGKGNKIYLQTIMNGHWGLPQKTPDYIIEEVPLWLG
FWGYYNYLKQTRTEAIFPLHMFVVQSKYIQTQQTETPNNFWAFIDNSFIQ
GKNPWDSVITYSEQKLWFPTVAWQLKTINAICESGPYVPKLDNQTYSTWE
LATHYSFHFKWGGPQISDQPVEDPGNKNKYDVPDTIKEALQIVNPAKNIA
ATMFHDWDYRRGCITSTAIKRMQQNLPTDSSLESDSDSEPAPKKKRLLPV
LHDPQKKTEKINQCLLSLCEESTCQEQETEENILKLIQQQQQQQQQKLKHN
LLVLIKDLKVKQRLLQLQTGVLE(SEQ ID NO: 221)
```

[1910] 注释：

[1911]

推定结构域	*AA 范围*
Arg 富集区	1 - 57
胶冻卷结构域	58 - 259
高变区	260 - 351
N22	352 - 510
C 末端结构域	511 - 673

[1912] 表37.示例性指环病毒ORF1氨基酸子序列(丙型细环病毒)

[1913]

TTV- HD16d -ORF1 (*Gammatorquevirus*)	
Arg 富集区	MPFWWGRRNKFWYGRNYRRKKRRFPKRRKRRFYRRTKYRRPARRRRRRRRRKVRRKKK (SEQ ID NO: 222)
胶冻卷结构域	TLIVRQWQPDSIVLCKIKGYDSIIWGAEGTQFQCSTHEMYEYTRQKYPGGGGFGVQLYSLEYLYDQWKLRNNIWTKTNQLKDLCRYLKCVMTFYRHQHIDFVIVYERQPPFEIDKLTYMKYHPYMLLQRKHKIILPSQTTNPRGKLKKKKTIKPPKQMLSKWFFQQQFAKYDLLLIAAAACSLRYPRIGCCNENRMITLYCL (SEQ ID NO: 223)
高变结构域	NTKFYQDTEWGTTKQAPHYFKPYATINKSMIFVSNYGGKKTEYNIGQWIETDIPGEGNLARYYRSISKEGGYFSPKILQAYQTKVKSVDYKP (SEQ ID NO: 224)
N22	LPIVLGRYNPAIDDGKGNKIYLQTIMNGHWGLPQKTPDYIIEEVPLWLGFWGYYNYLKQTRTEAIFPLHMFVVQSKYIQTQQTETPNNFWAFIDNSFIQGKNPWDSVITYSEQKLWFPTVAWQLKTINAICESGPYVPKLDNQTYSTWELATHYSFHFK (SEQ ID NO: 225)
C 末端结构域	WGGPQISDQPVEDPGNKNKYDVPDTIKEALQIVNPAKNIAATMFHDWDYRRGCITSTAIKRMQQNLPTDSSLESDSDSEPAPKKKRLLPVLHDPQKKTEKINQCLLSLCEESTCQEQETEENILKLIQQQQQQQQQKLKHNLLVLIKDLKVKQRLLQLQTGVLE (SEQ ID NO: 226)

[1914] 表D1.示例性指环病毒ORF1氨基酸子序列(丙型细环病毒)

[1915]

名称	环 3.1
属/进化分枝	丙型细环病毒
登录号	
蛋白质登录号	

[1916] 全序列:677AA

[1917]

```
1         10         20         30         40         50
|         |          |          |          |          |
MPFWWRRRNKRWWGRRFRYRRYNKYKTRRRRRIPRRRNRRFTKTRRRRKR
KKVRRKLKKITIKQWQPDSVKKCKIKGYSTLVMGAQGKQYNCYTNQASDY
VQPKAPQGGGFGCEVFNLKWLYQEYTAHRNIWTKTNEYTDLCRYTGAQII
LYRHPDVDFIVSWDNQPPFLLNKYTYPELQPQNLLLARRKRIILSQKSNP
KGKLRIKLRIPPPKQMITKWFFQRDFCDVNLFKLCASAASFRYPGISHGA
QSTIFSAYALNTDFYQCSDWCQTNTETGYLNIKTQQMPLWFHYREGGKEK
WYKYTNKEHRPYTNTYLKSISYNDGLFSPKAMFAFEVKAGGEGTTEPPQG
AQLIANLPLIALRYNPHEDTGHGNEIYLTSTFKGTYDKPKVTDALYFNNV
PLWMGFYGYWDFILQETKNKGVFDQHMFVVKCPALRPISQVTKQVYYPLV
DMDFCSGRLPFDEYLSKDIKSHWYPTAERQTVTINNFVTAGPYMPKFEPT
DKDSTWQLNYHYKFFFKWGGPQVTDPTVEDPCSRNKYPVPDTMQQTIQIK
NPEKLHPATLFHDWDLRRGFITQAAIKRMSENLQIDSSFESDGTESPKKK
KRCTKEIPTQNQKQEEIQECLLSLCEEPTCQEETEDLQLFIQQQQQQQYK
   LRKNLFKLLTHLKKGQRISQLQTGLLE (SEQ ID NO: 919)
```

[1918] 注释:

[1919]

推定结构域	*AA 范围*
Arg 富集区	1 - 59
胶冻卷结构域	60 - 260
高变区	261 - 356
N22	357 - 517
C 末端结构域	518 - 677

[1920] 表D2.示例性指环病毒ORF1氨基酸子序列(丙型细环病毒)

[1921]

环 3.1（丙型细环病毒）	
Arg 富集区	MPFWWRRRNKRWWGRRFRYRRYNKYKTRRRRRIPRRRNRRFTKTRR RRKRKKVRRKLKK (SEQ ID NO: 920)
胶冻卷结构域	ITIKQWQPDSVKKCKIKGYSTLVMGAQGKQYNCYTNQASDYVQPKAP QGGGFGCEVFNLKWLYQEYTAHRNIWTKTNEYTDLCRYTGAQIILYR HPDVDFIVSWDNQPPFLLNKYTYPELQPQNLLLARRKRIILSQKSNPKG KLRIKLRIPPPKQMITKWFFQRDFCDVNLFKLCASAASFRYPGISHGAQS TIFSAYAL (SEQ ID NO: 921)
高变结构域	NTDFYQCSDWCQTNTETGYLNIKTQQMPLWFHYREGGKEKWYKYTN KEHRPYTNTYLKSISYNDGLFSPKAMFAFEVKAGGEGTTEPPQGAQLIA N (SEQ ID NO: 922)
N22	LPLIALRYNPHEDTGHGNEIYLTSTFKGTYDKPKVTDALYFNNVPLWM GFYGYWDFILQETKNKGVFDQHMFVVKCPALRPISQVTKQVYYPLVD MDFCSGRLPFDEYLSKDIKSHWYPTAERQTVTINNFVTAGPYMPKFEPT DKDSTWQLNYHYKFFFK (SEQ ID NO: 923)
C 末端结构域	WGGPQVTDPTVEDPCSRNKYPVPDTMQQTIQIKNPEKLHPATLFHDWD LRRGFITQAAIKRMSENLQIDSSFESDGTESPKKKKRCTKEIPTQNQKQE EIQECLLSLCEEPTCQEETEDLQLFIQQQQQQQYKLRKNLFKLLTHLKK GQRISQLQTGLLE (SEQ ID NO: 924)

[1922] 表D3.示例性指环病毒ORF1氨基酸子序列(丙型细环病毒)

[1923]

名称	环 4.0
属/进化分枝	丙型细环病毒
登录号	
蛋白质登录号	

[1924] 全序列:662AA

[1925]

```
1         10         20         30         40         50
|         |          |          |          |          |
MPFWWRRRRKFWTNNRFNYTKRRRYRKRWPRRRRRRRPYRRPVRRRRRKL
RKVKRKKKSLIVRQWQPDSIRTCKIIGQSAIVVGAEGKQMYCYTVNKLIN
VPPKTPYGGGFGVDQYTLKYLYEEYRFAQNIWTQSNVLKDLCRYINVKLI
FYRDNKTDFVLSYDRNPPFQLTKFTYPGAHPQQIMLQKHHKFILSQMTKP
NGRLTKKLKIKPPKQMLSKWFFSKQFCKYPLLSLKASALDLRHSYLGCCN
ENPQVFFYYLNHGYYTITNWGAQSSTAYRPNSKVTDTTYYRYKNDRKNIN
IKSHEYEKSISYENGYFQSSFLQTQCIYTSERGEACIAEKPLGIAIYNPV
KDNGDGNMIYLVSTLANTWDQPPKDSAILIQGVPIWLGLFGYLDYCRQIK
ADKTWLDSHVLVIQSPAIFTYPNPGAGKWYCPLSQSFINGNGPFNQPPTL
LQKAKWFPQIQYQQEIINSFVESGPFVPKYANQTESNWELKYKYVFTFKW
GGPQFHEPEIADPSKQEQYDVPDTFYQTIQIEDPEGQDPRSLIHDWDYRR
GFIKERSLKRMSTYFSTHTDQQATSEEDIPKKKKRIGPQLTVPQQKEEET
LSCLLSLCKKDTFQETETQEDLQQLIKQQQEQQLLLKRNILQLIHKLKEN
          QQMLQLHTGMLP (SEQ ID NO: 925)
```

[1926] 注释:

[1927]

推定结构域	*AA 范围*
Arg 富集区	1 - 58
胶冻卷结构域	59 - 260
高变区	261 - 339
N22	340 - 499
C 末端结构域	500 - 662

[1928] 表D4.示例性指环病毒ORF1氨基酸子序列(丙型细环病毒)

[1929]

环 4.0（丙型细环病毒）	
Arg 富集区	MPFWWRRRRKFWTNNRFNYTKRRRYRKRWPRRRRRRRPYRRPVRRRRRKLRKVKRKKK (SEQ ID NO: 926)
胶冻卷结构域	SLIVRQWQPDSIRTCKIIGQSAIVVGAEGKQMYCYTVNKLINVPPKTPY

[1930]

	GGGFGVDQYTLKYLYEEYRFAQNIWTQSNVLKDLCRYINVKLIFYRDNKTDFVLSYDRNPPFQLTKFTYPGAHPQQIMLQKHHKFILSQMTKPNGRLTKKLKIKPPKQMLSKWFFSKQFCKYPLLSLKASALDLRHSYLGCCNENPQVFFYYL (SEQ ID NO: 927)
高变结构域	NHGYYTITNWGAQSSTAYRPNSKVTDTTYYRYKNDRKNINIKSHEYEKSISYENGYFQSSFLQTQCIYTSERGEACIAE (SEQ ID NO: 928)
N22	KPLGIAIYNPVKDNGDGNMIYLVSTLANTWDQPPKDSAILIQGVPIWLGLFGYLDYCRQIKADKTWLDSHVLVIQSPAIFTYPNPGAGKWYCPLSQSFINGNGPFNQPPTLLQKAKWFPQIQYQQEIINSFVESGPFVPKYANQTESNWELKYKYVFTFK (SEQ ID NO: 929)
C 末端结构域	WGGPQFHEPEIADPSKQEQYDVPDTFYQTIQIEDPEGQDPRSLIHDWDYRRGFIKERSLKRMSTYFSTHTDQQATSEEDIPKKKKRIGPQLTVPQQKEEETLSCLLSLCKKDTFQETETQEDLQQLIKQQQEQQLLLKRNILQLIHKLKENQQMLQLHTGMLP (SEQ ID NO: 930)

[1931] 表D5.示例性指环病毒ORF1氨基酸子序列(甲型细环病毒)进化分枝1

[1932]

名称	*环 5.2*
属/进化分枝	*甲型细环病毒进化分枝 1*
登录号	
蛋白质登录号	

[1933] 全序列:728AA

[1934]

```
1         10         20         30         40         50
|         |          |          |          |          |
TAWWWGRWRRRWRRRRPYTTRLRRRRARRAFPRRRRRRFVSRRWRRPYRR
RRRRGRRRRRRRRRHKPTLILRQWQPDCIRHCKITGWMPLIICGKGSTQF
NYITHADDITPRGASYGGNFTNMTFSLEAIYEQFLYHRNRWSASNHDLEL
CRYKGTTLKLYRHPEVDYIVTYSRTGPFEISHMTYLSTHPMLMLLNKHHI
VVPSLKTKPRGRKAIKVRIRPPKLMNNKWYFTRDFCNIGLFQLWATGLEL
RNPWLRMSTLSPCIGFNVLKNSIYTNLSNLPQYKNERLNIINNILHPQEI
TGTNNKKWQYTYTKLMAPIYYSANRASTYDWENYSKETNYNNTYVKFTQK
```

[1935]

```
RQEKLTKIRKEWQMLYPQQPTALPDSYDLLQEYGLYSPYYLNPTRINLDW
MTPYTHVRYNPLVDKGFGNRIYIQWCSEADVSYNRTKSKCLLQDMPLFFM
CYGYIDWAIKNTGVSSLVKDARICIRCPYTEPQLVGSTEDIGFVPISETF
MRGDMPVLAPYIPLSWFCKWYPNIAHQKEVLESIISCSPFMPRDQDMNGW
DITIGYKMDFLWGGSPLPSQPIDDPCQQGTHPIPDPDKHPRLLQVSNPKL
LGPRTVFHKWDIRRGQFSKRSIKRVSEYSSDDESLAPGLPSKRNKLDSAF
RGENREQKECYSLLKALEEEETPEEEEPAPQEKAQKEELLHQLQLQRRHQ
   RVLRRGLKLVFTDILRLRQGVHWNPELT (SEQ ID NO: 931)
```

[1936] 注释：

[1937]

推定结构域	*AA 范围*
Arg 富集区	1 - 66
胶冻卷结构域	67 - 277
高变区	278 - 395
N22	396 - 561
C 末端结构域	562 - 728

[1938] 表D6.示例性指环病毒ORF1氨基酸子序列(甲型细环病毒)进化分枝1

环 5.2（甲型细环病毒）进化分枝 1	
Arg 富集区	TAWWWGRWRRRWRRRRPYTTRLRRRRARRAFPRRRRRRFVSRRWRR PYRRRRRRGRRRRRRRRRHK (SEQ ID NO: 932)
胶冻卷结构域	PTLILRQWQPDCIRHCKITGWMPLIICGKGSTQFNYITHADDITPRGASY GGNFTNMTFSLEAIYEQFLYHRNRWSASNHDLELCRYKGTTLKLYRHP EVDYIVTYSRTGPFEISHMTYLSTHPMLMLLNKHHIVVPSLKTKPRGRK AIKVRIRPPKLMNNKWYFTRDFCNIGLFQLWATGLELRNPWLRMSTLS PCIGFNVLKNSIYTNL (SEQ ID NO: 933)
高变结构域	SNLPQYKNERLNIINNILHPQEITGTNNKKWQYTYTKLMAPIYYSANRA STYDWENYSKETNYNNTYVKFTQKRQEKLTKIRKEWQMLYPQQPTAL PDSYDLLQEYGLYSPYYLNPTR (SEQ ID NO: 934)
N22	INLDWMTPYTHVRYNPLVDKGFGNRIYIQWCSEADVSYNRTKSKCLL QDMPLFFMCYGYIDWAIKNTGVSSLVKDARICIRCPYTEPQLVGSTEDI GFVPISETFMRGDMPVLAPYIPLSWFCKWYPNIAHQKEVLESIISCSPFM PRDQDMNGWDITIGYKMDFL (SEQ ID NO: 935)
C 末端结构域	WGGSPLPSQPIDDPCQQGTHPIPDPDKHPRLLQVSNPKLLGPRTVFHKW DIRRGQFSKRSIKRVSEYSSDDESLAPGLPSKRNKLDSAFRGENREQKE CYSLLKALEEEETPEEEEPAPQEKAQKEELLHQLQLQRRHQRVLRRGL KLVFTDILRLRQGVHWNPELT (SEQ ID NO: 936)

[1939]

[1940] 表D7.示例性指环病毒ORF1氨基酸子序列(甲型细环病毒)-进化分枝3

[1941]

名称	环 6.0
属/进化分枝	甲型细环病毒进化分枝 3

[1942]

登录号
蛋白质登录号

[1943] 全序列:767AA

[1944]

```
1         10         20         30         40         50
|         |          |          |          |          |
MAYGWWRRRRRRPWWRRRWRRWRRRRRPRRRRPRRRYRRRRTVRRRGRGR
WTRAHRRWRRKGKRSRKKKIIIRQWQPNYTRRCNIVGYMPLLICGENTVA
TNYATHSDDSYYPGPFGGGMTTDKFTLRILYDEYKRFMNYWTSSNEDLDL
CRYLGCTLYVFRHPEVDFIIIINTSPPFLDTEITGPSIHPGMMALNKRSR
WIPSIKNRPGRKHYIKIKVGAPRMFTDKWYPQTDLCDMTLLTIFASAADM
QYPFGSPLTDTIVVSFQVLQSMYNDCLSVLPDNFAETSGKGTQLHENIIQ
HLPYYNTTQTQAQFKRFIENMNATNGDNIWASYINTTKFSSANTPKNDTG
IGGPYTTYSDSWYKGTVYNDKIKTIPIKASKLYYEQTKNLIGITFTGSTH
RLHYCGGLYSSVWLSAGRSYFETKGPYTDITYNPFSDRGEGNMLWIDWLT
KNDSVYSKTSSKCLIENLPLWASVYGYKEYCSKVTGDTNIEHNCRCVIRS
PYTVPQLLDHNNPFRGYVPYSFNFGNGKMPGGSSLVPIRMRAKWYPTLFH
QKEVLEAIAQAGPFAYHSDIKKVSLGIKYRFKWVWGGNPVSQQVVRNPCK
TTQGSSGNRVPRSIQVVDPRYNTPELTIHAWDFRHGFFGRKAIKRMQEQP
IPHDTFSAGFKRSRRDTEALQCSQEEQQKENLLFPVQQLKRVPPWETSQE
SQSEEENSQKQETLSQQLRDQLHKQRLMGEQLRSLLYQMQRVQQNQHINP
        MLLPKGLALTSISHNVI (SEQ ID NO: 937)
```

[1945] 注释:

[1946]

推定结构域	*AA 范围*
Arg 富集区	1 - 69
胶冻卷结构域	70 - 269
高变区	270 - 424
N22	425 - 584
C 末端结构域	585 - 767

[1947] 表D8.示例性指环病毒ORF1氨基酸子序列(甲型细环病毒)-进化分枝3

[1948]

环 6.0（甲型细环病毒）	
Arg 富集区	MAYGWWRRRRRRPWWRRRWRRWRRRRRPRRRRPRRRYRRRRTVRRRGRGRWTRAHRRWRRKGKRSRKKK (SEQ ID NO: 938)
胶冻卷结构域	IIIRQWQPNYTRRCNIVGYMPLLICGENTVATNYATHSDDSYYPGPFGGGMTTDKFTLRILYDEYKRFMNYWTSSNEDLDLCRYLGCTLYVFRHPEVDFIIIINTSPPFLDTEITGPSIHPGMMALNKRSRWIPSIKNRPGRKHYIKIKVGAPRMFTDKWYPQTDLCDMTLLTIFASAADMQYPFGSPLTDTIVVSFQVL (SEQ ID NO: 939)
高变结构域	QSMYNDCLSVLPDNFAETSGKGTQLHENIIQHLPYYNTTQTQAQFKRFIENMNATNGDNIWASYINTTKFSSANTPKNDTGIGGPYTTYSDSWYKGTVYNDKIKTIPIKASKLYYEQTKNLIGITFTGSTHRLHYCGGLYSSVWLSAGRSYFETK (SEQ ID NO: 940)

N22	GPYTDITYNPFSDRGEGNMLWIDWLTKNDSVYSKTSSKCLIENLPLWA SVYGYKEYCSKVTGDTNIEHNCRCVIRSPYTVPQLLDHNNPFRGYVPY SFNFGNGKMPGGSSLVPIRMRAKWYPTLFHQKEVLEAIAQAGPFAYHS DIKKVSLGIKYRFKWV (SEQ ID NO: 941)
C 末端结构域	WGGNPVSQQVVRNPCKTTQGSSGNRVPRSIQVVDPRYNTPELTIHAWD FRHGFFGRKAIKRMQEQPIPHDTFSAGFKRSRRDTEALQCSQEEQQKEN LLFPVQQLKRVPPWETSQESQSEEENSQKQETLSQQLRDQLHKQRLMG EQLRSLLYQMQRVQQNQHINPMLLPKGLALTSISHNVI (SEQ ID NO: 942)

[1949]

[1950] 表D9.示例性指环病毒ORF1氨基酸子序列(甲型细环病毒)-进化分枝7

名称	环 7.0
属/进化分枝	甲型细环病毒-进化分枝 7
登录号	
蛋白质登录号	

[1951]

[1952] 全序列:766AA

[1953]

```
1          10         20         30         40         50
|          |          |          |          |          |
MAWRWWWQRRWRRRRWPRRRWRRLRRRRPRRPVRRRRRRTTVRRRRWRGR
RGRRTYTRRAVRRRRRPRKRLVLTQWSPQTVRNCSIRGIVPMVICGHTKA
GRNYAIHSEDFTTQIQPFGGSFSTTTWSLKVLWDEHQKFQNRWSYPNTQL
DLARYRGVTFWFYRDQKTDYIVQWSRNPPFKLNKYSSAMYHPGMMMQAKR
KLVVPSFQTRPKGKKRYRVTIKPPNMFADKWYTQEDLCPVPLVQIVVSAA
SLLHPFCPPQTNNPCITFQVLKDIYDECIGVNETMKDKYKKLQTTLYTTC
TYYQTTQVLAQLSPAFQPAMKPTTTQSAATATTLGNYVPELKYNNGSFHT
GQNAVFGMCSYKPTDSIMTKANGWFWQNLMVDNNLHSSYGKATLECMEYH
TGIYSSIFLSPQRSLEFPAAYQDVTYNPNCDRAVGNVVWFQYSTKMDTNF
DETKCKCVLKNIPLWAAFNGYSDFIMQELSISTEIHNFGIVCFQCPYTFP
PCFNKNKPLKGYVFYDTTFGNGKMPDGSGHVPIYWQQRWWIRLAFQVQVM
HDFVLTGPFSYKDDLANTTLTARYKFKFKWGGNIIPEQIIKNPCHREQSL
ASYPDRQRRDLQVVDPSTMGPIYTFHTWDWRRGLFGADAIQRVSQKPGDA
LRFTNPFKRPRYLPPTDREDYRQEEDFALQEKRRRTSTEEAQDEESPPES
APLLQQQQQQRQLSVHLAEQQRLGVQLRYILQEVLKTQAGLHLNPLLLGP
          PQTRSISLSPPKAYSP (SEQ ID NO: 943)
```

[1954] 注释:

推定结构域	*AA 范围*
Arg 富集区	1 - 70
胶冻卷结构域	71 - 271
高变区	272 - 418
N22	419 - 579
C 末端结构域	580 - 766

[1955]

[1956]

[1957] 表D10.示例性指环病毒ORF1氨基酸子序列(甲型细环病毒)-进化分枝7

[1958]

环 7.0（甲型细环病毒）	
Arg 富集区	MAWRWWWQRRWRRRRWPRRRWRRLRRRRPRRPVRRRRRRRTTVRRRRWRGRRGRRTYTRRAVRRRRRPRKR (SEQ ID NO: 944)
胶冻卷结构域	LVLTQWSPQTVRNCSIRGIVPMVICGHTKAGRNYAIHSEDFTTQIQPFGGSFSTTTWSLKVLWDEHQKFQNRWSYPNTQLDLARYRGVTFWFYRDQKTDYIVQWSRNPPFKLNKYSSAMYHPGMMMQAKRKLVVPSFQTRPKGKKRYRVTIKPPNMFADKWYTQEDLCPVPLVQIVVSAASLLHPFCPPQTNNPCITFQVL (SEQ ID NO: 945)
高变结构域	KDIYDECIGVNETMKDKYKKLQTTLYTTCTYYQTTQVLAQLSPAFQPAMKPTTTQSAATATTLGNYVPELKYNNGSFHTGQNAVFGMCSYKPTDSIMTKANGWFWQNLMVDNNLHSSYGKATLECMEYHTGIYSSIFLSPQRSLEFP (SEQ ID NO: 946)
N22	AAYQDVTYNPNCDRAVGNVVWFQYSTKMDTNFDETKCKCVLKNIPLWAAFNGYSDFIMQELSISTEIHNFGIVCFQCPYTFPPCFNKNKPLKGYVFYDTTFGNGKMPDGSGHVPIYWQQRWWIRLAFQVQVMHDFVLTGPFSYKDDLANTTLTARYKFKFK (SEQ ID NO: 947)
C 末端结构域	WGGNIIPEQIIKNPCHREQSLASYPDRQRRDLQVVDPSTMGPIYTFHTWDWRRGLFGADAIQRVSQKPGDALRFTNPFKRPRYLPPTDREDYRQEEDFALQEKRRRTSTEEAQDEESPPESAPLLQQQQQQRQLSVHLAEQQRLGVQLRYILQEVLKTQAGLHLNPLLLGPPQTRSISLSPPKAYSP (SEQ ID NO: 948)

[1959] 共有ORF1结构域序列

[1960] 在一些实施例中，ORF1分子，例如，如本文所述，包含胶冻卷结构域、N22结构域和/或C末端结构域(CTD)中的一个或多个。在一些实施例中，胶冻卷结构域包含具有如本文所述的胶冻卷结构域共有序列的氨基酸序列(例如，如表37A-37C中的任一个中所列)。在一些实施例中，N22结构域包含具有如本文所述的N22结构域共有序列的氨基酸序列(例如，如表37A-37C中的任一个中所列)。在一些实施例中，CTD结构域包含具有如本文所述的CTD结构域共有序列的氨基酸序列(例如，如表37A-37C中的任一个中所列)。在一些实施例中，以“(X_{a-b})”格式在表37A-37C中的任一个中列出的氨基酸包含连续系列的氨基酸，其中该系列包含至少a个和至多b个氨基酸。在某些实施例中，该系列中的所有氨基酸都是相同的。在其他实施例中，该系列包括至少两个(例如，至少2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20或21)个不同的氨基酸。

[1961] 表37A.甲型细环病毒ORF1结构域共有序列

[1962]

结构域	序列	SEQ ID NO:
胶冻卷	LVLTQWQPNTVRRCYIRGYLPLIICGEN(X_{0-3})TTSRNYATHSDDTIQKGPFGGGMSTTTFSLRVLYDEYQRFMNRWTYSNEDLDLARYLGCKFTFYRHPDXDFIVQYNTNPPFKDTKLTAPSIHP(X_{1-5})GMLMLSKRKILIPSLKTRPKGKHYVKVRIGPPKLFEDKWYTQSDLCDVPLVXLYATAADLQHPFGSPQTDNPCVTFQVLGSXYNKHLSISP; 其中 X= 任何氨基酸。	227
N22	SNFEFPGAYTDITYNPLTDKGVGNMVWIQYLTKPDTIXDKTQS(X_{0-3})KCLIEDLPLWAALYGYVDFCEKETGDSAIIXNXGRVLIRCPYTKPPLYDKT(X_{0-4})NKGFVPYSTNFGNGKMPGGSGYVPIYWRARWYPTLFHQKEVLEDIVQSGPFAYKDEKPSTQLVMKYCFNFN; 其中 X= 任何氨基酸。	228
CTD	WGGNPISQQVVRNPCKDSG(X_{0-3})SGXGRQPRSVQVVDPKYMGPEYTFHSWDWRRGLFGEKAIKRMSEQPTDDEIFTGGXPKRPRRDPPTXQXPEE(X_{1-4})QKESSSFR(X_{2-14})PWESSSQEXESESQEEEE(X_{0-30})EQTVQQQLRQQLREQRRLRVQLQLLFQQLLKT(X_{0-4})QAGLHINPLLLSQA(X_{0-40})*; 其中 X= 任何氨基酸。	229

[1963] 表37B.乙型细环病毒ORF1结构域共有序列

[1964]

结构域	序列	SEQ ID NO:
胶冻卷	LKQWQPSTIRKCKIKGYLPLFQCGKGRISNNYTQYKESIVPHHEPGGGGWSIQQFTLGALYEEHLKLRNWWTKSNDGLPLVRYLGCTIKLYRSEDTDYIVTYQRCYPMTATKLTYLSTQPSRMLMNKHKIIVPSKXT(X_{1-4})NKKKKPYKKIFIKPPSQMQNKWYFQQDIANTPLLQLTXTACSLDRMYLSSDSISNNITFTSLNTNFFQNPNFQ; 其中 X= 任何氨基酸。	230
N22	(X_{4-10})TPLYFECRYNPFKDKGTGNKVYLVSNN(X_{1-8})TGWDPPTDPDLIIEGFPLWLLLWGWLDWQKKLGKIQNIDTDYILVIQSXYYIPP(X_{1-3})KLPYYVPLDXD(X_{0-2})FLHGRSPY(X_{3-16})PSDKQHWHPKVRFQXETINNIALTGPGTPKLPNQKSIQAHMKYKFYFK; 其中 X= 任何氨基酸。	231
CTD	WGGCPAPMETITDPCKQPKYPIPNNLLQTTSLQXPTTPIETYLYKFDERRGLLTKKAAKRIKKDXTTETTLFTDTGXXTSTTLPTXXQTETTQEEXTSEEE(X_{0-5})ETLLQQLQQLRRKQKQLRXRILQLLQLLXLL(X_{0-26})*; 其中 X= 任何氨基酸。	232

[1965] 表37C.丙型细环病毒ORF1结构域共有序列

[1966]

结构域	序列	SEQ ID NO:
胶冻卷	TIPLKQWQPESIRKCKIKGYGTLVLGAEGRQFYCYTNE KDEYTPPKAPGGGGFGVELFSLEYLYEQWKARNNIWT KSNXYKDLCRYTGCKITFYRHPTTDFIVXYSRQPPFEID KXTYMXXHPQXLLLRKHKKIILSKATNPKGKLKKKIKI KPPKQMLNKWFFQKQFAXYGLVQLQAAACBLRYPRL GCCNENRLITLYYLN; 其中 X= 任何氨基酸。	233
N22	LPIVVARYNPAXDTGKGNKXWLXSTLNGSXWAPPTTD KDLIIEGLPLWLALYGYWSYJKKVKKDKGILQSHMFV VKSPAIQPLXTATTQXTFYPXIDNSFIQGKXPYDEPJTX NQKKLWYPTLEHQQETINAIVESGPYVPKLDNQKNST WELXYXYTFYFK; 其中 X= 任何氨基酸。	234
CTD	WGGPQIPDQPVEDPKXQGTYPVPDTXQQTIQIXNPLKQ KPETMFHDWDYRRGIITSTALKRMQENLETDSSFXSDS EETP(X_{0-2})KKKKRLTXELPXPQEETEEIQSCLLSLCEEST CQEE(X_{1-6})ENLQQLIHQQQQQQQQLKHNILKLLSDLKZ KQRLLQLQTGILE(X_{1-10})*; 其中 X= 任何氨基酸。	235

[1967] 在一些实施例中,胶冻卷结构域包含如表21、23、25、27、29、31、33、35、D2、D4、D6、D8、D10、或37A-37C中的任一个中所列的胶冻卷结构域氨基酸加,或与其具有至少70%、75%、80%、8%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在一些实施例中,N22结构域包含如表21、23、25、27、29、31、33、35、D2、D4、D6、D8、D10、或37A-37C中的任一个中所列的N22结构域氨基酸加,或与其具有至少70%、75%、80%、8%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。在一些实施例中,CTD结构域包含如表21、23、25、27、29、31、33、35、D2、D4、D6、D8、D10、或37A-37C中的任一个中所列的CTD结构域氨基酸加,或与其具有至少70%、75%、80%、8%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。

[1968] ORF2分子

[1969] 在一些实施例中,指环体包含ORF2分子和/或编码ORF2分子的核酸。通常,ORF2分子包含具有指环病毒ORF2蛋白(例如,如本文所述的指环病毒ORF2蛋白,例如,如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16或18中的任一个中所列)的结构特征和/或活性的多肽,或其功能片段。在一些实施例中,ORF2分子包含与如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16或18中的任一个中所示的指环病毒ORF2蛋白序列具有至少70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的氨基酸序列。

[1970] 在一些实施例中,ORF2分子包含与甲型细环病毒、乙型细环病毒或丙型细环病毒ORF2蛋白具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%序列同一性的氨基酸序列。在一些实施例中,ORF2分子(例如,与甲型细环病毒ORF2蛋白具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%序列同一性的ORF2分子)的长度为250个或更少个氨基酸(例如,约150-200个氨基酸)。在一些实施例中,ORF2分子(例如,与乙型细环病毒

ORF2蛋白具有至少75％、80％、85％、90％、95％、96％、97％、98％或99％序列同一性的ORF2分子)的长度约为50-150个氨基酸。在一些实施例中,ORF2分子(例如,与丙型细环病毒ORF2蛋白具有至少75％、80％、85％、90％、95％、96％、97％、98％或99％序列同一性的ORF2分子)的长度为约100-200个氨基酸(例如,约100-150个氨基酸)。在一些实施例中,ORF2分子包含螺旋-转角-螺旋基序(例如,包含两个位于转角区侧翼的α螺旋的螺旋-转角-螺旋基序)。在一些实施例中,ORF2分子不包含TTV分离株TA278或TTV分离株SANBAN的ORF2蛋白的氨基酸序列。在一些实施例中,ORF2分子具有蛋白磷酸酶活性。在一些实施例中,ORF2分子包含相对于例如如本文所述的野生型ORF2蛋白(例如,如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16或18中的任一个中所示)的至少一个差异(例如,突变、化学修饰或表观遗传改变)。

[1971] 保守的ORF2基序

[1972] 在一些实施例中,本文所述的多肽(例如,ORF2分子)包含氨基酸序列[W/F]$X^7HX^3CX^1CX^5H$(SEQ ID NO:949),其中X^n是任何n个氨基酸的连续序列。在实施例中,X^7表示任何七个氨基酸的连续序列。在实施例中,X^3表示任何三个氨基酸的连续序列。在实施例中,X^1表示任何单个氨基酸。在实施例中,X^5表示任何五个氨基酸的连续序列。在一些实施例中,[W/F]可以是色氨酸或苯丙氨酸。在一些实施例中,[W/F]$X^7HX^3CX^1CX^5H$(SEQ ID NO:949)包含在例如如本文所述的ORF2分子的N22结构域内。在一些实施例中,本文所述的遗传元件包含编码氨基酸序列[W/F]$X^7HX^3CX^1CX^5H$(SEQ ID NO:949)的核酸序列(例如,编码例如如本文所述的ORF2分子的核酸序列),其中X^n是任何n个氨基酸的连续序列。

[1973] 遗传元件

[1974] 在一些实施例中,指环体包含遗传元件。在一些实施例中,遗传元件具有以下一种或多种特征:基本上不与宿主细胞的基因组整合,是游离核酸,是单链DNA,是环状的,大约1至10kb,存在于细胞核内,可以与内源性蛋白质结合,产生效应子,例如靶向宿主或靶细胞的基因、活性或功能的多肽或核酸(例如,RNA、iRNA、微小RNA)。在一个实施例中,遗传元件是基本上非整合的DNA。在一些实施例中,遗传元件包含包装信号,例如结合衣壳蛋白的序列。在一些实施例中,在包装或衣壳结合序列之外,遗传元件与野生型指环病毒核酸序列具有小于70％、60％、50％、40％、30％、20％、10％、5％序列同一性,例如,与例如如本文所述的指环病毒核酸序列具有小于70％、60％、50％、40％、30％、20％、10％、5％序列同一性。在一些实施例中,在包装或衣壳结合序列之外,遗传元件具有与指环病毒核酸序列至少70％、75％、80％、8％、90％、95％、96％、97％、98％、99％或100％相同的少于500、450、400、350、300、250、200、150或100个连续核苷酸。在某些实施例中,遗传元件是包含启动子序列、编码治疗效应子的序列和衣壳结合蛋白的环状单链DNA。

[1975] 在一实施例中,遗传元件与例如本文所述(例如,表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中任一个所述)的指环病毒核酸序列或其片段具有至少约70％、75％、80％、8％、90％、95％、96％、97％、98％、99％或100％序列同一性,或编码与指环病毒氨基酸序列(例如,如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16或18中任一个所述)或其片段具有至少约70％、75％、80％、8％、90％、95％、96％、97％、98％、99％或100％序列同一性的氨基酸序列。在实施例中,遗传元件包含编码效应子(例如内源性效应子或外源性效应子,例如有效负载),例如多肽效应子(例如蛋白质)或核酸效应子(例如非

编码RNA,例如miRNA、siRNA、mRNA、lncRNA、RNA、DNA、反义RNA、gRNA)的序列。

[1976] 在一些实施例中,遗传元件的长度小于20kb(例如,小于约19kb、18kb、17kb、16kb、15kb、14kb、13kb、12kb、11kb、10kb、9kb、8kb、7kb、6kb、5kb、4kb、3kb、2kb、1kb或更小)。在一些实施例中,遗传元件独立地或额外地具有大于1000b的长度(例如,至少约1.1kb、1.2kb、1.3kb、1.4kb、1.5kb、1.6kb、1.7kb、1.8kb、1.9kb、2kb、2.1kb、2.2kb、2.3kb、2.4kb、2.5kb、2.6kb、2.7kb、2.8kb、2.9kb、3kb、3.1kb、3.2kb、3.3kb、3.4kb、3.5kb、3.6kb、3.7kb、3.8kb、3.9kb、4kb、4.1kb、4.2kb、4.3kb、4.4kb、4.5kb、4.6kb、4.7kb、4.8kb、4.9kb、5kb或更大)。在一些实施例中,遗传元件的长度为约2.5kb-4.6kb、2.8kb-4.0kb、3.0kb-3.8kb或3.2kb-3.7kb。在一些实施例中,遗传元件具有约1.5-2.0、1.5-2.5、1.5-3.0、1.5-3.5、1.5-3.8、1.5-3.9、1.5-4.0、1.5-4.5或1.5-5.0kb的长度。在一些实施例中,遗传元件具有约2.0-2.5、2.0-3.0、2.0-3.5、2.0-3.8、2.0-3.9、2.0-4.0、2.0-4.5或2.0-5.0kb的长度。在一些实施例中,遗传元件具有约2.5-3.0、2.5-3.5、2.5-3.8、2.5-3.9、2.5-4.0、2.5-4.5或2.5-5.0kb的长度。在一些实施例中,遗传元件具有约3.0-5.0、3.5-5.0、4.0-5.0或4.5-5.0kb的长度。在一些实施例中,遗传元件具有约1.5-2.0、2.0-2.5、2.5-3.0、3.0-3.5、3.1-3.6、3.2-3.7、3.3-3.8、3.4-3.9、3.5-4.0、4.0-4.5或4.5-5.0kb的长度。

[1977] 在一些实施例中,遗传元件包含本文描述的一个或多个特征,例如,编码基本上非致病性蛋白的序列,蛋白质结合序列,编码调节性核酸的一个或多个序列,一个或多个调节性序列,编码复制蛋白的一个或多个序列,以及其他序列。在一些实施例中,基本上非致病性的蛋白包含氨基酸序列或其功能片段或与本文所述的氨基酸序列中任一个(指环病毒氨基酸序列,例如,表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16或18中任一个所列)具有至少约60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的序列。

[1978] 在实施例中,遗传元件由双链环状DNA产生(例如,通过体外环化产生)。在一些实施例中,遗传元件通过双链环状DNA的滚环复制产生。在实施例中,滚环复制发生在细胞(例如宿主细胞,例如哺乳动物细胞,例如人细胞,例如HEK293T细胞、A549细胞或Jurkat细胞)中。在实施例中,遗传元件可以通过细胞中的滚环复制以指数扩增。在实施例中,遗传元件可以通过细胞中的滚环复制以线性扩增。在实施例中,双链环状DNA或遗传元件能够通过细胞中的滚环复制产生原始量的至少2、4、8、16、32、64、128、256、518、1024倍或更多倍。在实施例中,将例如如本文所述的双链环状DNA引入细胞中。

[1979] 在一些实施例中,双链环状DNA和/或遗传元件不包含一种或多种细菌质粒元件(例如细菌复制起点或可选择标记,例如细菌抗性基因)。在一些实施例中,双链环状DNA和/或遗传元件不包含细菌质粒骨架。

[1980] 在一个实施例中,本发明包括一种遗传元件,所述遗传元件包括编码(i)基本上非致病性的外部蛋白质,(ii)将所述遗传元件结合至所述基本上非致病性的外部蛋白质的外部蛋白质结合序列,以及(iii)调节性核酸的核酸序列(例如DNA序列)。在这样的实施例中,遗传元件可包含一个或多个序列,所述序列与天然病毒序列(例如,天然指环病毒序列,例如,如本文所述)的核苷酸序列中任一个具有至少约60%、70%、80%、85%、90%、95%、96%、97%、98%和99%核苷酸序列同一性。

[1981] 蛋白质结合序列

[1982] 许多病毒采用的策略是病毒衣壳蛋白识别其基因组中的特定蛋白质结合序列。例如,在具有未分段基因组的病毒(例如酵母的L-A病毒)中,在基因组的5'端有二级结构(茎环)和特定序列,它们都用于结合病毒衣壳蛋白。但是,具有分段基因组的病毒(如呼肠孤病毒科、正粘病毒科(流行性感冒)、布尼亚病毒科和沙粒病毒)需要包装每个基因组片段。一些病毒利用片段的互补区来帮助病毒包括每个基因组分子中的一个。其他病毒对于每个不同的片段具有特定的结合位点。例如,参见Curr Opin Struct Biol.[结构生物学最新观点]2010年2月;20(1):114-120;和Journal of Virology[病毒学杂志](2003),77(24),13036-13041。

[1983] 在一些实施例中,遗传元件编码与基本上非致病性的蛋白质结合的蛋白质结合序列。在一些实施例中,蛋白质结合序列有助于将遗传元件包装到蛋白质外部中。在一些实施例中,蛋白质结合序列特异性结合基本上非致病性的蛋白质的精氨酸富集区。在一些实施例中,遗传元件包含如实例8中所述的蛋白质结合序列。在一些实施例中,遗传元件包含蛋白质结合序列,其与指环病毒序列的5'UTR保守结构域或GC富集结构域(例如,如表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中任一个所示)具有至少70%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性。

[1984] 在实施例中,蛋白质结合序列与表A1的指环病毒5'UTR保守结构域核苷酸序列(例如,表A1的核酸序列的核苷酸165-235)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性。在实施例中,蛋白质结合序列与表A1的指环病毒GC富集核苷酸序列(例如,表A1的核酸序列的核苷酸3620-3648)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性。在实施例中,蛋白质结合序列与表A3的指环病毒5'UTR保守结构域核苷酸序列(例如,表A3的核酸序列的核苷酸175-245)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性。在实施例中,蛋白质结合序列与表A5的指环病毒5'UTR保守结构域核苷酸序列(例如,表A5的核酸序列的核苷酸138-208)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性。在实施例中,蛋白质结合序列与表A7的指环病毒5'UTR保守结构域核苷酸序列(例如,表A7的核酸序列的核苷酸174-244)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性。在实施例中,蛋白质结合序列与表A7的指环病毒GC富集核苷酸序列(例如,表A7的核酸序列的核苷酸3720-3742)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性。在实施例中,蛋白质结合序列与表A9的指环病毒5'UTR保守结构域核苷酸序列(例如,表A9的核酸序列的核苷酸100-171)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性。在实施例中,蛋白质结合序列与表A11的指环病毒5'UTR保守结构域核苷酸序列(例如,表A11的核酸序列的核苷酸294-364)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性。在实施例中,蛋白质结合序列与表A1的指环病毒GC富集核苷酸序列(例如,表A11的核酸序列的核苷酸3844-3895)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性。

[1985] 在实施例中,蛋白质结合序列与表1的指环病毒5'UTR保守结构域核苷酸序列(例如,表1的核酸序列的核苷酸177-247)具有至少约70%、75%、80%、85%、90%、95%、96%、

97％、98％、99％或100％序列同一性。在实施例中，蛋白质结合序列与表1的指环病毒GC富集核苷酸序列(例如，表1的核酸序列的核苷酸3415-3570)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性。在实施例中，蛋白质结合序列与表3的指环病毒5'UTR保守结构域核苷酸序列(例如，表3的核酸序列的核苷酸204-273)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性。在实施例中，蛋白质结合序列与表3的指环病毒GC富集核苷酸序列(例如，表3的核酸序列的核苷酸3302-3541)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性。在实施例中，蛋白质结合序列与表5的指环病毒5'UTR保守结构域核苷酸序列(例如，表5的核酸序列的核苷酸170-240)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性。在实施例中，蛋白质结合序列与表5的指环病毒GC富集核苷酸序列(例如，表5的核酸序列的核苷酸3632-3753)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性。在实施例中，蛋白质结合序列与表7的指环病毒5'UTR保守结构域核苷酸序列(例如，表7的核酸序列的核苷酸170-238)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性。在实施例中，蛋白质结合序列与表7的指环病毒GC富集核苷酸序列(例如，表7的核酸序列的核苷酸3768-3878)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性。在实施例中，蛋白质结合序列与表9的指环病毒5'UTR保守结构域核苷酸序列(例如，表9的核酸序列的核苷酸170-240)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性。在实施例中，蛋白质结合序列与表9的指环病毒GC富集核苷酸序列(例如，表9的核酸序列的核苷酸3302-3541)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性。在实施例中，蛋白质结合序列与表11的指环病毒5'UTR保守结构域核苷酸序列(例如，表11的核酸序列的核苷酸174-244)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性。在实施例中，蛋白质结合序列与表11的指环病毒GC富集核苷酸序列(例如，表11的核酸序列的核苷酸3691-3794)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性。在实施例中，蛋白质结合序列与表13的指环病毒5'UTR保守结构域核苷酸序列(例如，表13的核酸序列的核苷酸170-240)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性。在实施例中，蛋白质结合序列与表13的指环病毒GC富集核苷酸序列(例如，表13的核酸序列的核苷酸3759-3866)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性。在实施例中，蛋白质结合序列与表15的指环病毒5'UTR保守结构域核苷酸序列(例如，表15的核酸序列的核苷酸323-393)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性。在实施例中，蛋白质结合序列与表15的指环病毒GC富集核苷酸序列(例如，表15的核酸序列的核苷酸2868-2929)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性。在实施例中，蛋白质结合序列与表17的指环病毒5'UTR保守结构域核苷酸序列(例如，表17的核酸序列的核苷酸117-187)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性。在实施例中，蛋白质结合序列与表17的指环病毒GC富集核苷酸序列(例如，表17的核酸序列的核苷酸3054-3172)具有至少约70％、75％、80％、85％、90％、95％、96％、

97％、98％、99％或100％序列同一性。

[1986] 在实施例中，蛋白质结合序列与表B1的指环病毒5'UTR保守结构域核苷酸序列(例如，表B1的核酸序列的核苷酸185-255)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性。在实施例中，蛋白质结合序列与表B1的指环病毒GC富集核苷酸序列(例如，表B1的核酸序列的核苷酸3141-3264)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性。在实施例中，蛋白质结合序列与表B2的指环病毒5'UTR保守结构域核苷酸序列(例如，表B2的核酸序列的核苷酸185-254)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性。在实施例中，蛋白质结合序列与表B2的指环病毒GC富集核苷酸序列(例如，表B2的核酸序列的核苷酸3076-3176)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性。在实施例中，蛋白质结合序列与表B3的指环病毒5'UTR保守结构域核苷酸序列(例如，表B3的核酸序列的核苷酸178-248)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性。在实施例中，蛋白质结合序列与表B3的指环病毒GC富集核苷酸序列(例如，表B3的核酸序列的核苷酸3555-3696)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性。在实施例中，蛋白质结合序列与表B4的指环病毒5'UTR保守结构域核苷酸序列(例如，表B4的核酸序列的核苷酸176-246)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性。在实施例中，蛋白质结合序列与表B4的指环病毒GC富集核苷酸序列(例如，表B4的核酸序列的核苷酸3720-3828)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性。在实施例中，蛋白质结合序列与表B5的指环病毒5'UTR保守结构域核苷酸序列(例如，表B5的核酸序列的核苷酸170-240)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性。在实施例中，蛋白质结合序列与表B5的指环病毒GC富集核苷酸序列(例如，表B5的核酸序列的核苷酸3716-3815)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性。

[1987] 5'UTR区域

[1988] 在一些实施例中，遗传元件(例如，遗传元件的蛋白质结合序列)包括与表38和/或图20中所示的核酸序列具有至少约75％(例如，至少75％、80％、85％、90％、95％、96％、97％、98％、99％或100％)同一性的核酸序列。在一些实施例中，遗传元件(例如，遗传元件的蛋白质结合序列)包括表38所示的共有5'UTR序列的核酸序列，其中X_1、X_2、X_3、X_4和X_5各自独立地是任何核苷酸，例如其中X_1＝G或T，X_2＝C或A，X_3＝G或A，X_4＝T或C，并且X_5＝A、C或T)。在实施例中，遗传元件(例如，遗传元件的蛋白质结合序列)包括与表38中所示的共有5'UTR序列具有至少约75％(例如，至少75％、80％、85％、90％、95％、96％、97％、98％、99％或100％)同一性的核酸序列。在实施例中，遗传元件(例如，遗传元件的蛋白质结合序列)包括与表38中所示的示例性TTV 5'UTR序列具有至少约75％(例如，至少75％、80％、85％、90％、95％、96％、97％、98％、99％或100％)同一性的核酸序列。在实施例中，遗传元件(例如，遗传元件的蛋白质结合序列)包括与表38中所示的TTV-CT30F 5'UTR序列具有至少约75％(例如，至少75％、80％、85％、90％、95％、96％、97％、98％、99％或100％)同一性的核酸序列。在实施例中，遗传元件(例如，遗传元件的蛋白质结合序列)包括与表38中所示的TTV-HD23a

5'UTR序列具有至少约75%(例如,至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)同一性的核酸序列。在实施例中,遗传元件(例如,遗传元件的蛋白质结合序列)包括与表38中所示的TTV-JA20 5'UTR序列具有至少约75%(例如,至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)同一性的核酸序列。在实施例中,遗传元件(例如,遗传元件的蛋白质结合序列)包括与表38中所示的TTV-TJN02 5'UTR序列具有至少约75%(例如,至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)同一性的核酸序列。在实施例中,遗传元件(例如,遗传元件的蛋白质结合序列)包括与表38中所示的TTV-tth8 5'UTR序列具有至少约75%(例如,至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)同一性的核酸序列。

[1989] 在实施例中,遗传元件(例如,遗传元件的蛋白质结合序列)包括与表38中所示的甲型细环病毒共有5'UTR序列具有至少约75%(例如,至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)同一性的核酸序列。在实施例中,遗传元件(例如,遗传元件的蛋白质结合序列)包括与表38中所示的甲型细环病毒进化分枝1 5'UTR序列具有至少约75%(例如,至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)同一性的核酸序列。在实施例中,遗传元件(例如,遗传元件的蛋白质结合序列)包括与表38中所示的甲型细环病毒进化分枝2 5'UTR序列具有至少约75%(例如,至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)同一性的核酸序列。在实施例中,遗传元件(例如,遗传元件的蛋白质结合序列)包括与表38中所示的甲型细环病毒进化分枝3 5'UTR序列具有至少约75%(例如,至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)同一性的核酸序列。在实施例中,遗传元件(例如,遗传元件的蛋白质结合序列)包括与表38中所示的甲型细环病毒进化分枝4 5'UTR序列具有至少约75%(例如,至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)同一性的核酸序列。在实施例中,遗传元件(例如,遗传元件的蛋白质结合序列)包括与表38中所示的甲型细环病毒进化分枝5 5'UTR序列具有至少约75%(例如,至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)同一性的核酸序列。在实施例中,遗传元件(例如,遗传元件的蛋白质结合序列)包括与表38中所示的甲型细环病毒进化分枝6 5'UTR序列具有至少约75%(例如,至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)同一性的核酸序列。在实施例中,遗传元件(例如,遗传元件的蛋白质结合序列)包括与表38中所示的甲型细环病毒进化分枝7 5'UTR序列具有至少约75%(例如,至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)同一性的核酸序列。

[1990] 在实施例中,遗传元件包含与表A1的指环病毒5'UTR保守结构域核苷酸序列(例如,表A1的核酸序列的核苷酸165-235)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,遗传元件包含与表A3的指环病毒5'UTR保守结构域核苷酸序列(例如,表A3的核酸序列的核苷酸175-245)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,遗传元件包含与表A5的指环病毒5'UTR保守结构域核苷酸序列(例如,表A5的核酸序列的核苷酸138-208)具有至少约70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在实施例中,遗传元件包含与表A7的指环病毒5'UTR保守结构域核苷酸序列(例如,表A7的核酸序列的核苷酸174-244)具有至少约

70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，遗传元件包含与表A9的指环病毒5'UTR保守结构域核苷酸序列(例如，表A9的核酸序列的核苷酸100-171)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，遗传元件包含与表A11的指环病毒5'UTR保守结构域核苷酸序列(例如，表A11的核酸序列的核苷酸294-364)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。

[1991] 在实施例中，遗传元件包含与表1的指环病毒5'UTR保守结构域核苷酸序列(例如，表1的核酸序列的核苷酸177-247)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，遗传元件包含与表3的指环病毒5'UTR保守结构域核苷酸序列(例如，表3的核酸序列的核苷酸204-273)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，遗传元件包含与表5的指环病毒5'UTR保守结构域核苷酸序列(例如，表5的核酸序列的核苷酸170-240)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，遗传元件包含与表7的指环病毒5'UTR保守结构域核苷酸序列(例如，表7的核酸序列的核苷酸170-238)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，遗传元件包含与表9的指环病毒5'UTR保守结构域核苷酸序列(例如，表9的核酸序列的核苷酸170-240)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，遗传元件包含与表11的指环病毒5'UTR保守结构域核苷酸序列(例如，表11的核酸序列的核苷酸174-244)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，遗传元件包含与表13的指环病毒5'UTR保守结构域核苷酸序列(例如，表13的核酸序列的核苷酸170-240)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，遗传元件包含与表15的指环病毒5'UTR保守结构域核苷酸序列(例如，表15的核酸序列的核苷酸323-393)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，遗传元件包含与表17的指环病毒5'UTR保守结构域核苷酸序列(例如，表17的核酸序列的核苷酸117-187)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。

[1992] 在实施例中，遗传元件包含与表B1的指环病毒5'UTR保守结构域核苷酸序列(例如，表B1的核酸序列的核苷酸185-255)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，遗传元件包含与表B2的指环病毒5'UTR保守结构域核苷酸序列(例如，表B2的核酸序列的核苷酸185-254)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，遗传元件包含与表B3的指环病毒5'UTR保守结构域核苷酸序列(例如，表B3的核酸序列的核苷酸178-248)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，遗传元件包含与表B4的指环病毒5'UTR保守结构域核苷酸序列(例如，表B4的核酸序列的核苷酸176-246)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。

在实施例中，遗传元件包含与表B5的指环病毒5'UTR保守结构域核苷酸序列(例如，表B5的核酸序列的核苷酸170-240)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。

[1993] 表38.来自指环病毒的示例性5'UTR序列

[1994]

来源	序列	SEQ ID NO:
共有	CGGGTGCCGX_1AGGTGAGTTTACACACCGX_2AGT CAAGGGGCAATTCGGGCTCX_3GGACTGGCCGGG CX_4X_5TGGG X_1 = G 或 T X_2 = C 或 A X_3 = G 或 A X_4 = T 或 C X_5 = A、C 或 T	105
示例性 TTV 序列	CGGGTGCCGGAGGTGAGTTTACACACCGCAGTC AAGGGGCAATTCGGGCTCGGGACTGGCCGGGCT WTGGG	106
TTV-CT30F	CGGGTGCCGTAGGTGAGTTTACACACCGCAGTC AAGGGGCAATTCGGGCTCGGGACTGGCCGGGCT ATGGG	107
TTV-HD23a	CGGGTGCCGGAGGTGAGTTTACACACCGCAGTC AAGGGGCAATTCGGGCTCGGGACTGGCCGGGCC CTGGG	108
TTV-JA20	CGGGTGCCGGAGGTGAGTTTACACACCGCAGTC AAGGGGCAATTCGGGCTCGGGACTGGCCGGGCT TTGGG	109
TTV-TJN02	CGGGTGCCGGAGGTGAGTTTACACACCGCAGTC AAGGGGCAATTCGGGCTCGGGACTGGCCGGGCT ATGGG	110
TTV-tth8	CGGGTGCCGGAGGTGAGTTTACACACCGAAGTC AAGGGGCAATTCGGGCTCAGGACTGGCCGGGCT TTGGG	111
甲型细环病毒	CGGGTGCCGGAGGTGAGTTTACACACCGCAGTC	112

[1995]

共有 5' UTR	AAGGGGCAATTCGGGCTCGGGACTGGCCGGGC X_1X_2TGGG; 其中 X_1 包括 T 或 C,并且其中 X_2 包括 A、C 或 T。	
甲型细环病毒进化分枝 1 5' UTR（例如，TTV-CT30F）	CGGGTGCCGTAGGTGAGTTTACACACCGCAGTC AAGGGGCAATTCGGGCTCGGGACTGGCCGGGCT ATGGG	113
甲型细环病毒进化分枝 2 5' UTR（例如，TTV-P13-1）	CGGGTGCCGGAGGTGAGTTTACACACCGCAGTC AAGGGGCAATTCGGGCTCGGGACTGGCCGGGCC CGGG	114
甲型细环病毒进化分枝 3 5' UTR（例如，TTV-tth8）	CGGGTGCCGGAGGTGAGTTTACACACCGAAGTC AAGGGGCAATTCGGGCTCAGGACTGGCCGGGCT TTGGG	115
甲型细环病毒进化分枝 4 5' UTR（例如，TTV-HD20a）	CGGGTGCCGGAGGTGAGTTTACACACCGCAGTC AAGGGGCAATTCGGGCTCGGGAGGCCGGGCCAT GGG	116
甲型细环病毒进化分枝 5 5' UTR（例如，TTV-16）	CGGGTGCCGGAGGTGAGTTTACACACCGCAGTC AAGGGGCAATTCGGGCTCGGGACTGGCCGGGCC CCGGG	117
甲型细环病毒进化分枝 6 5' UTR（例如，TTV-TJN02）	CGGGTGCCGGAGGTGAGTTTACACACCGCAGTC AAGGGGCAATTCGGGCTCGGGACTGGCCGGGCT ATGGG	118
甲型细环病毒进化分枝 7 5' UTR（例如，TTV-HD16d）	CGGGTGCCGAAGGTGAGTTTACACACCGCAGTC AAGGGGCAATTCGGGCTCGGGACTGGCCGGGCT ATGGG	119

[1996] GC富集区

[1997] 在一些实施例中,遗传元件(例如,遗传元件的蛋白质结合序列)包括与表39和/或图20和32中的任一个中所示的核酸序列具有至少约75%(例如,至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)同一性的核酸序列。在实施例中,遗传元件(例如,遗传元件的蛋白质结合序列)包括与表39中显示的GC富集序列具有至少约75%(例如,至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)同一性的核酸序列。

[1998] 在实施例中,遗传元件(例如,遗传元件的蛋白质结合序列)包括与表39中所示的36个核苷酸的GC富集序列(例如,36个核苷酸的共有GC富集区序列1、36个核苷酸的共有GC富集区序列2、TTV进化分枝1 36个核苷酸的区域、TTV进化分枝3 36个核苷酸的区域、TTV进化分枝3分离株GH1 36个核苷酸的区域、TTV进化分枝3sle1932 36个核苷酸的区域、TTV进化分枝4ctdc002 36个核苷酸的区域、TTV进化分枝5 36个核苷酸区域、TTV进化分枝6 36个核苷酸的区域或TTV进化分枝7 36个核苷酸的区域)具有至少约75%(例如,至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)同一性的核酸序列。在实施例中,遗传元件(例如,遗传元件的蛋白质结合序列)包含核酸序列,所述核酸序列包含表39中所示的36个核苷酸的GC富集序列(例如,36个核苷酸的共有GC富集区序列1、36个核苷酸的共有GC富集区序列2、TTV进化分枝1 36个核苷酸的区域、TTV进化分枝3 36个核苷酸的区域、TTV进化分枝3分离株GH1 36个核苷酸的区域、TTV进化分枝3sle1932 36个核苷酸的区域、TTV进化分枝4ctdc002 36个核苷酸的区域、TTV进化分枝5 36个核苷酸区域、TTV进化分枝6 36个核苷酸的区域或TTV进化分枝7 36个核苷酸的区域)的至少10、15、20、25、30、31、32、33、34、35或36个连续核苷酸。

[1999] 在实施例中,遗传元件(例如,遗传元件的蛋白质结合序列)包含与甲型细环病毒

GC富集区序列(例如选自TTV-CT30F、TTV-P13-1、TTV-tth8、TTV-HD20a、TTV-16、TTV-TJN02或TTV-HD16d,例如,如表39中所列)具有至少约75%(例如,至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)同一性的核酸序列。在实施例中,遗传元件(例如,遗传元件的蛋白质结合序列)包含核酸序列,所述核酸序列包含甲型细环病毒GC富集区序列(例如选自TTV-CT30F、TTV-P13-1、TTV-tth8、TTV-HD20a、TTV-16、TTV-TJN02或TTV-HD16d,例如,如表39中所列)的至少10、15、20、25、30、35、40、45、50、60、70、80、90、100、104、105、108、110、111、115、120、122、130、140、145、150、155或156个连续核苷酸。

[2000] 在实施例中,36个核苷酸的GC富集序列选自:

[2001] (i)CGCGCTGCGCGCGCCGCCCAGTAGGGGGAGCCATGC(SEQ ID NO:160),

[2002] (ii)GCGCTX$_1$CGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:164),其中X$_1$选自T、G或A;

[2003] (iii)GCGCTTCGCGCGCCGCCCACTAGGGGGCGTTGCGCG(SEQ ID NO:165);

[2004] (iv)GCGCTGCGCGCGCCGCCCAGTAGGGGGCGCAATGCG(SEQ ID NO:166);

[2005] (v)GCGCTGCGCGCGCGGCCCCCGGGGGAGGCATTGCCT(SEQ ID NO:167);

[2006] (vi)GCGCTGCGCGCGCGCGCCGGGGGGGCGCCAGCGCCC(SEQ ID NO:168);

[2007] (vii)GCGCTTCGCGCGCGCGCCGGGGGGCTCCGCCCCCCC(SEQ ID NO:169);

[2008] (viii)GCGCTTCGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:170);

[2009] (ix)GCGCTACGCGCGCGCGCCGGGGGGCTGCGCCCCCCC(SEQ ID NO:171);或

[2010] (x)GCGCTACGCGCGCGCGCCGGGGGGCTCTGCCCCCCC(SEQ ID NO:172)。

[2011] 在实施例中,遗传元件(例如,遗传元件的蛋白质结合序列)包含核酸序列CGCGCTGCGCGCGCCGCCCAGTAGGGGGAGCCATGC(SEQ ID NO:160)。

[2012] 在实施例中,遗传元件(例如,遗传元件的蛋白质结合序列)包含表39中所示的共有GC富集序列的核酸序列,其中X$_1$、X$_4$、X$_5$、X$_6$、X$_7$、X$_{12}$、X$_{13}$、X$_{14}$、X$_{15}$、X$_{20}$、X$_{21}$、X$_{22}$、X$_{26}$、X$_{29}$、X$_{30}$和X$_{33}$各自独立地是任何核苷酸,并且其中X$_2$、X$_3$、X$_8$、X$_9$、X$_{10}$、X$_{11}$、X$_{16}$、X$_{17}$、X$_{18}$、X$_{19}$、X$_{23}$、X$_{24}$、X$_{25}$、X$_{27}$、X$_{28}$、X$_{31}$、X$_{32}$和X$_{34}$各自独立地是不存在的或任何核苷酸。在一些实施例中,X$_1$至X$_{34}$中的一个或多个(例如全部)各自独立地是表39中指定的核苷酸(或不存在)。在实施例中,遗传元件(例如,遗传元件的蛋白质结合序列)包括与表39中显示的示例性TTV GC富集序列(例如,全序列、片段1、片段2、片段3或其任何组合,例如按顺序排列的片段1-3)具有至少约75%(例如,至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)同一性的核酸序列。在实施例中,遗传元件(例如,遗传元件的蛋白质结合序列)包括与表39中显示的TTV-CT30F GC富集序列(例如,全序列、片段1、片段2、片段3、片段4、片段5、片段6、片段7、片段8或其任何组合,例如按顺序排列的片段1-7)具有至少约75%(例如,至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)同一性的核酸序列。在实施例中,遗传元件(例如,遗传元件的蛋白质结合序列)包括与表39中显示的TTV-HD23a GC富集序列(例如,全序列、片段1、片段2、片段3、片段4、片段5、片段6、或其任何组合,例如按顺序排列的片段1-6)具有至少约75%(例如,至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)同一性的核酸序列。在实施例中,遗传元件(例如,遗传元件的蛋白质结合序列)包括与表39中显示的TTV-JA20 GC富集序列(例如,全序列、片段1、片段2或其任何组合,例如按顺序排列的片段1和2)具有至少约75%(例如,至少75%、80%、85%、90%、95%、96%、97%、98%、99%或

100%)同一性的核酸序列。在实施例中,遗传元件(例如,遗传元件的蛋白质结合序列)包括与表39中显示的TTV-TJN02 GC富集序列(例如,全序列、片段1、片段2、片段3、片段4、片段5、片段6、片段7、片段8或其任何组合,例如按顺序排列的片段1-8)具有至少约75%(例如,至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)同一性的核酸序列。在实施例中,遗传元件(例如,遗传元件的蛋白质结合序列)包括与表39中显示的TTV-tth8 GC富集序列(例如,全序列、片段1、片段2、片段3、片段4、片段5、片段6、片段7、片段8、片段9或其任何组合,例如按顺序排列的片段1-6)具有至少约75%(例如,至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)同一性的核酸序列。在实施例中,遗传元件(例如,遗传元件的蛋白质结合序列)包括与表39中显示的片段7具有至少约75%(例如,至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)同一性的核酸序列。在实施例中,遗传元件(例如,遗传元件的蛋白质结合序列)包括与表39中显示的片段8具有至少约75%(例如,至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)同一性的核酸序列。在实施例中,遗传元件(例如,遗传元件的蛋白质结合序列)包括与表39中显示的片段9具有至少约75%(例如,至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)同一性的核酸序列。

[2013] 表39.来自指环病毒的示例性GC富集序列

[2014]

来源	序列	SEQ ID NO:
共有	$CGGCGGX_1GGX_2GX_3X_4X_5CGCGCTX_6CGCGC$ $GCX_7X_8X_9X_{10}CX_{11}X_{12}X_{13}X_{14}GGGGX_{15}X_{16}X_{17}X_{18}$ $X_{19}X_{20}X_{21}GCX_{22}X_{23}X_{24}X_{25}CCCCCCCX_{26}CGCGC$ $ATX_{27}X_{28}GCX_{29}CGGGX_{30}CCCCCCCCCX_{31}X_{32}X_{33}GGGGGGCTCCGX_{34}CCCCCCGGCCCCCC$ X_1 = G 或 C X_2 = G、C 或不存在 X_3 = C 或不存在	120

[2015]

		X_4 = G 或 C X_5 = G 或 C X_6 = T、G 或 A X_7 = G 或 C X_8 = G 或不存在 X_9 = C 或不存在 X_{10} = C 或不存在 X_{11} = G、A 或不存在 X_{12} = G 或 C X_{13} = C 或 T X_{14} = G 或 A X_{15} = G 或 A X_{16} = A、G、T 或不存在 X_{17} = G、C 或不存在 X_{18} = G、C 或不存在 X_{19} = C、A 或不存在 X_{20} = C 或 A X_{21} = T 或 A X_{22} = G 或 C X_{23} = G、T 或不存在 X_{24} = C 或不存在 X_{25} = G、C 或不存在 X_{26} = G 或 C X_{27} = G 或不存在 X_{28} = C 或不存在 X_{29} = G 或 A X_{30} = G 或 T X_{31} = C、T 或不存在 X_{32} = G、C、A 或不存在 X_{33} = G 或 C X_{34} = C 或不存在	
示例性 TTV 序列	全序列	GCCGCCGCGGCGGCGGSGGNGNSGCGCGCT DCGCGCGCSNNNCRCCRGGGGGNNNNCWG CSNCNCCCCCCCCCGCGCATGCGCGGGKCC CCCCCCCNNCGGGGGGCTCCGCCCCCCGGC CCCCCCCCGTGCTAAACCCACCGCGCATGC GCGACCACGCCCCCGCCGCC	121
	片段 1	GCCGCCGCGGCGGCGGSGGNGNSGCGCGCT DCGCGCGCSNNNCRCCRGGGGGNNNNCWG CSNCNCCCCCCCCCGCGCAT	122
	片段 2	GCGCGGGKCCCCCCCCCNNCGGGGGGCTC CG	123
	片段 3	CCCCCCGGCCCCCCCCCGTGCTAAACCCAC CGCGCATGCGCGACCACGCCCCCGCCGCC	124
TTV-CT30F	全序列	GCGGCGG-GGGGGCG-GCCGCG- TTCGCGCGCCGCCCACCAGGGGGTG-- CTGCG-CGCCCCCCCCCGCGCAT GCGCGGGGCCCCCCCCC-- GGGGGGGCTCCGCCCCCCCGGCCCCCCCCC GTGCTAAACCCACCGCGCATGCGCGACCAC GCCCCCGCCGCC	125
	片段 1	GCGGCGG	126

[2016]

	片段 2	GGGGGCG	127
	片段 3	GCCGCG	128
	片段 4	TTCGCGCGCCGCCCACCAGGGGGTG	129
	片段 5	CTGCG	130
	片段 6	CGCCCCCCCCCGCGCAT	131
	片段 7	GCGCGGGGCCCCCCCCC	132
	片段 8	GGGGGGGCTCCGCCCCCCCGGCCCCCCCCC GTGCTAAACCCACCGCGCATGCGCGACCAC GCCCCCGCCGCC	133
TTV-HD23a	全序列	CGGCGGCGGCGGCG-CGCGCGCTGCGCGCGCG---CGCCGGGGGGGCGCCAGCG-CCCCCCCCCCCGCGCAT GCACGGGTCCCCCCCCCCACGGGGGGCTCC G CCCCCCGGCCCCCCCCC	134
	片段 1	CGGCGGCGGCGGCG	135
	片段 2	CGCGCGCTGCGCGCGCG	136
	片段 3	CGCCGGGGGGGCGCCAGCG	137
	片段 4	CCCCCCCCCCCGCGCAT	138
	片段 5	GCACGGGTCCCCCCCCCCACGGGGGGCTCC G	139
	片段 6	CCCCCCGGCCCCCCCCC	140
TTV-JA20	全序列	CCGTCGGCGGGGGGGCCGCGCGCTGCGCG CGCGGCCC-CCGGGGGAGGCACAGCCTCCCCCCCCCGCG CGCATGCGCGCGGGTCCCCCCCCCCTCCGGG GGGCTCCGCCCCCCGGCCCCCCCC	141
	片段 1	CCGTCGGCGGGGGGGCCGCGCGCTGCGCG CGCGGCCC	142
	片段 2	CCGGGGGAGGCACAGCCTCCCCCCCCCGCG CGCATGCGCGCGGGTCCCCCCCCCCTCCGGG GGGCTCCGCCCCCCGGCCCCCCCC	143
TTV-TJN02	全序列	CGGCGGCGGCG-CGCGCGCTACGCGCGCG---CGCCGGGGGG----CTGCCGC-CCCCCCCCCGCGCAT GCGCGGGGCCCCCCCCC-GCGGGGGGCTCCG CCCCCCGGCCCCCC	144
	片段 1	CGGCGGCGGCG	145
	片段 2	CGCGCGCTACGCGCGCG	146
	片段 3	CGCCGGGGGG	147
	片段 4	CTGCCGC	148
	片段 5	CCCCCCCCCGCGCAT	149
	片段 6	GCGCGGGGCCCCCCCCC	150
	片段 7	GCGGGGGGCTCCG	151
	片段 8	CCCCCCGGCCCCCC	152
TTV-tth8	全序列	GCCGCCGCGGCGGCGGGGG-GCGGCGCGCTGCGCGCGCCGCCCAGTAGG GGGAGCCATGCG---CCCCCCCCCGCGCAT GCGCGGGGCCCCCCCCC-GCGGGGGGCTCCG CCCCCCGGCCCCCCCCG	153
	片段 1	GCCGCCGCGGCGGCGGGGG	154
	片段 2	GCGGCGCGCTGCGCGCGCCGCCCAGTAGG	155

[2017]

		GGGAGCCATGCG	
	片段 3	CCCCCCCCCGCGCAT	156
	片段 4	GCGCGGGGCCCCCCCCC	157
	片段 5	GCGGGGGGCTCCG	158
	片段 6	CCCCCCGGCCCCCCCCG	159
	片段 7	CGCGCTGCGCGCGCCGCCCAGTAGGGGGA GCCATGC	160
	片段 8	CCGCCATCTTAAGTAGTTGAGGCGGACGGT GGCGTGAGTTCAAAGGTCACCATCAGCCAC ACCTACTCAAAATGGTGG	161
	片段 9	CTTAAGTAGTTGAGGCGGACGGTGGCGTGA GTTCAAAGGTCACCATCAGCCACACCTACT CAAAATGGTGGACAATTTCTTCCGGGTCAA AGGTTACAGCCGCCATGTTAAAACACGTGA CGTATGACGTCACGGCCGCCATTTTGTGAC ACAAGATGGCCGACTTCCTTCC	162
另外的 GC 富集序列（如图 32 所示）	36 个核苷酸的共有 GC 富集区序列 1	CGCGCTGCGCGCGCCGCCCAGTAGGGGGA GCCATGC	163
	36 个核苷酸的区域共有序列 2	GCGCTX$_1$CGCGCGCGCCGGGGGGCTGCG CCCCCCC，其中 X$_1$ 选自 T、G 或 A	164
	TTV 进化分枝 1 36 个核苷酸的区域	GCGCTTCGCGCGCCGCCCACTAGGGGGCGT TGCGCG	165
	TTV 进化分枝 3 36 个核苷酸的区域	GCGCTGCGCGCGCCGCCCAGTAGGGGGCG CAATGCG	166
	TTV 进化分枝 3 分离株 GH1 36 个核苷酸的区域	GCGCTGCGCGCGCGGCCCCCGGGGGAGGC ATTGCCT	167
	TTV 进化分枝 3 sle1932 36 个核苷酸的区域	GCGCTGCGCGCGCGCCGGGGGGGCGCC AGCGCCC	168
	TTV 进化分枝 4 ctdc002 36 个核苷酸的区域	GCGCTTCGCGCGCGCCGGGGGGCTCCGC CCCCCC	169
	TTV 进化分枝 5 36 个核苷酸的区域	GCGCTTCGCGCGCGCCGGGGGGCTGCGC CCCCCC	170
	TTV 进化分枝 6 36 个核苷酸的区域	GCGCTACGCGCGCGCCGGGGGGCTGCG CCCCCCC	171
	TTV 进化分枝 7 36 个核苷酸的区域	GCGCTACGCGCGCGCCGGGGGGCTCTGC CCCCCC	172
其他甲型细环病毒 GC 富集区序列	TTV-CT30F	GCGGCGGGGGGGCGGCCGCGTTCGCGCGC CGCCCACCAGGGGGTGCTGCGCGCCCCCCC CCGCGCATGCGCGGGGCCCCCCCCCGGGG GGGCTCCGCCCCCCCGGCCCCCCCCCGTGC TAAACCCACCGCGCATGCGCGACCACGCCC	801

[2018]

		CCGCCGCC	
	TTV-P13-1	CCGAGCGTTAGCGAGGAGTGCGACCCTACC CCCTGGGCCCACTTCTTCGGAGCCGCGCGC TACGCCTTCGGCTGCGCGCGGCACCTCAGA CCCCCGCTCGTGCTGACACGCTTGCGCGTG TCAGACCACTTCGGGCTCGCGGGGGTCGGG	802
	TTV-tth8	GCCGCCGCGGCGGCGGGGGGCGGCGCGCT GCGCGCGCCGCCCAGTAGGGGGAGCCATG CGCCCCCCCCCGCGCATGCGCGGGGCCCCC CCCCGCGGGGGGCTCCGCCCCCCGGCCCCC CCCG	803
	TTV-HD20a	CGGCCCAGCGGCGGCGCGCGCTTCGCGC GCGCGCCGGGGGGCTCCGCCCCCCCCCGCG CATGCGCGGGGCCCCCCCCCGCGGGGGGCT CCGCCCCCCGGTCCCCCCCCG	804
	TTV-16	CGGCCGTGCGGCGGCGCGCGCTTCGCGC GCGCGCCGGGGGCTGCCGCCCCCCCCCGCG CATGCGCGCGGGGCCCCCCCCCGCGGGGG GCTCCGCCCCCCGGCCCCCCCCCCG	805
	TTV-TJN02	CGGCGGCGGCGCGCGCGCTACGCGCGCGC GCCGGGGGGCTGCCGCCCCCCCCCCGCGCA TGCGCGGGGCCCCCCCCCGCGGGGGGCTCC GCCCCCCGGCCCCCC	806
	TTV-HD16d	GGCGGCGGCGCGCGCGCTACGCGCGCGCG CCGGGGAGCTCTGCCCCCCCCCGCGCATGC GCGCGGGTCCCCCCCCCGCGGGGGGCTCCG CCCCCCGGTCCCCCCCCCG	807

[2019] 在实施例中，遗传元件包含与表A1的指环病毒GC富集核苷酸序列(例如，表A1的核酸序列的核苷酸3620-3648)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，遗传元件包含与表A7的指环病毒GC富集核苷酸序列(例如，表A7的核酸序列的核苷酸3720-3742)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，遗传元件包含与表A1的指环病毒GC富集核苷酸序列(例如，表A11的核酸序列的核苷酸3844-3895)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。

[2020] 在实施例中，遗传元件包含与表1的指环病毒GC富集核苷酸序列(例如，表1的核酸序列的核苷酸3415-3570)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，遗传元件包含与表3的指环病毒GC富集核苷酸序列(例如，表3的核酸序列的核苷酸3302-3541)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，遗传元件包含与表5的指环病毒GC富集核苷酸序列(例如，表5的核酸序列的核苷酸3632-3753)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，遗传元件包含与表7的指环病毒GC富集核苷酸序列(例如，表7的核酸序列的核苷酸3768-3878)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，遗传元件包含与表9的指环病毒GC富集核苷酸序列(例如，表9的核酸序列的核苷酸3302-3541)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，遗传元件包含与表11的指环病毒GC富集核苷酸序列(例如，表11的核酸序列的核苷酸3691-3794)具

有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，遗传元件包含与表13的指环病毒GC富集核苷酸序列(例如，表13的核酸序列的核苷酸3759-3866)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，遗传元件包含与表15的指环病毒GC富集核苷酸序列(例如，表15的核酸序列的核苷酸2868-2929)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，遗传元件包含与表17的指环病毒GC富集核苷酸序列(例如，表17的核酸序列的核苷酸3054-3172)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。

[2021] 在实施例中，遗传元件包含与表B1的指环病毒GC富集核苷酸序列(例如，表B1的核酸序列的核苷酸3141-3264)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，遗传元件包含与表B2的指环病毒GC富集核苷酸序列(例如，表B2的核酸序列的核苷酸3076-3176)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，遗传元件包含与表B3的指环病毒GC富集核苷酸序列(例如，表B3的核酸序列的核苷酸3555-3696)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，遗传元件包含与表B4的指环病毒GC富集核苷酸序列(例如，表B4的核酸序列的核苷酸3720-3828)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。在实施例中，遗传元件包含与表B5的指环病毒GC富集核苷酸序列(例如，表B5的核酸序列的核苷酸3716-3815)具有至少约70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的核酸序列。

[2022] 效应子

[2023] 在一些实施例中，遗传元件可包括一种或多种编码功能性效应子的序列，例如内源性效应子或外源性效应子，例如治疗性多肽或核酸，例如细胞毒性或细胞溶解性RNA或蛋白质。在一些实施例中，功能性核酸是非编码RNA。在一些实施例中，功能性核酸是编码RNA。效应子可以调节生物活性，例如增加或降低酶活性、基因表达、细胞信号传导以及细胞或器官功能。效应子活性还可包括结合调节性蛋白以调节调节子的活性，例如转录或翻译。效应子活性还可以包括激活剂或抑制剂功能。例如，效应子可通过触发酶中底物亲和力的增加来诱导酶活性，例如，果糖2,6-二磷酸激活磷酸果糖激酶1并增加糖酵解响应于胰岛素的速率。在另一个实例中，效应子可以抑制底物与受体的结合并抑制其激活，例如，纳曲酮和纳洛酮结合阿片样物质受体而不激活它们，并阻断受体结合阿片样物质的能力。效应子活性还可以包括调节蛋白质的稳定性/降解和/或转录物的稳定性/降解。例如，可以通过多肽辅因子泛素在蛋白质上以将它们标记为降解，从而将蛋白质靶向降解。在另一个实例中，效应子通过阻断酶的活性位点来抑制酶活性，例如，甲氨蝶呤是四氢叶酸(一种二氢叶酸还原酶的辅酶)的结构类似物，其与二氢叶酸还原酶的结合力比天然底物高1000倍，并抑制核苷酸碱基合成。

[2024] 在一些实施例中，编码效应子的序列是遗传元件的一部分，例如，它可以插入在如实例10、12或22中所述的插入位点。在实施例中，在非编码区处将编码效应子的序列插入遗

传元件中，例如，位于开放阅读框的3’和遗传元件的GC富集区的5’的非编码区，在TATA盒上游的5’非编码区中，在5’UTR中，在聚A信号下游或GC富集区上游的3’非编码区中。在实施例中，在例如本文所述的TTV-tth8质粒的约核苷酸3588处或在例如本文所述的TTMV-LY2质粒的约核苷酸2843处将编码效应子的序列插入遗传元件中。在实施例中，在例如本文所述的TTV-tth8质粒的核苷酸336-3015处或之内或在例如本文所述的TTV-LY2质粒的核苷酸242-2812处或之内将编码效应子的序列插入遗传元件中。在一些实施例中，编码效应子的序列替代了开放阅读框的一部分或全部(例如，本文所述的ORF，例如，表A1-A12、B1-B5、C1-C5或1-18中任一个中所示的ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3和/或ORF2t/3)。

[2025] 在一些实施例中，编码效应子的序列包括100-2000、100-1000、100-500、100-200、200-2000、200-1000、200-500、500-1000、500-2000或1000-2000个核苷酸。在一些实施例中，效应子是核酸或蛋白质有效负载，例如，如实例11中所述。

[2026] 调节性核酸

[2027] 在一些实施例中，效应子是调节性核酸。调节性核酸修饰内源基因和/或外源基因的表达。在一实施例中，调控核酸靶向宿主基因。调节性核酸可包括但不限于与内源基因杂交的核酸(例如，如本文其他地方所述的miRNA、siRNA、mRNA、lncRNA、RNA、DNA、反义RNA、gRNA)、与外源核酸(例如病毒DNA或RNA)杂交的核酸、与RNA杂交的核酸、干扰基因转录的核酸、干扰RNA翻译的核酸、稳定RNA或去稳定RNA的核酸(例如通过靶向降解)，以及调节DNA或RNA结合因子的核酸。在实施例中，调节性核酸编码miRNA。

[2028] 在一些实施例中，调节性核酸包含通常包含5-500个碱基对的RNA或RNA样结构(取决于特定的RNA结构，例如miRNA 5-30bp、lncRNA 200-500bp)并且可以具有与细胞内表达的靶基因中的编码序列或编码细胞内表达的靶基因的序列相同(或互补)或几乎相同(或基本上互补)的核碱基序列。

[2029] 在一些实施例中，调节性核酸包含核酸序列，例如，指导RNA(gRNA)。在一些实施例中，DNA靶向部分包含指导RNA或编码指导RNA的核酸。gRNA短合成RNA可以由与不完整的效应子部分结合所必需的“支架”序列和用于基因组靶标的用户定义的约20个核苷酸靶向序列组成。在实践中，指导RNA序列通常被设计为具有17-24个核苷酸(例如19、20或21个核苷酸)的长度，并且与靶核酸序列互补。定制的gRNA生成器和算法可从商业上获得，用于设计有效的指导RNA。还使用嵌合的“单指导RNA”(“sgRNA”)，一种模拟天然存在的crRNA-tracrRNA复合物并包含tracrRNA(用于结合核酸酶)和至少一个crRNA(以将核酸酶指导至编辑目标序列)的工程化(合成)单RNA分子，实现了基因编辑。也已经证明，在基因组编辑中，化学修饰的sgRNA是有效的；参见，例如，Hendel等人(2015)Nature Biotechnol.[自然生物技术]，985-991。

[2030] 调节性核酸包含识别特定DNA序列(例如，与基因的启动子、增强子、沉默子或阻遏子相邻或在其内的序列)的gRNA。

[2031] 某些调节性核酸可以通过RNA干扰(RNAi)的生物过程抑制基因表达。RNAi分子包含RNA或RNA样结构，其通常包含15-50个碱基对(例如约18-25个碱基对)并且具有与细胞内表达的靶基因中的编码序列相同(互补)或几乎相同(基本上互补)的核碱基序列。RNAi分子包括但不限于：短干扰RNA(siRNA)、双链RNA(dsRNA)、微小RNA(miRNA)、短发夹RNA(shRNA)、部分双螺旋和dicer底物(美国专利号8,084,599、8,349,809和8,513,207)。

[2032] 长非编码RNA(lncRNA)定义为长于100个核苷酸的非蛋白质编码转录本。这个有点随意的限制将lncRNA与小型调控RNA(例如微小RNA(miRNA)、短干扰RNA(siRNA)和其他短RNA)区分开来。通常,大多数(约78%)lncRNA的特征为组织特异性的。以与附近蛋白质编码基因相反的方向转录的发散lncRNA(占哺乳动物基因组中总lncRNA的约20%大比例)可能会调控附近基因的转录。

[2033] 遗传元件可以编码具有与内源基因或基因产物(例如,mRNA)的全部或片段基本上互补或完全互补的序列的调节性核酸。调节性核酸可以与内含子和外显子之间的边界处的序列互补,从而防止特异性基因的新生成的核RNA转录物成熟为用于转录的mRNA。与特定基因互补的调控核酸可与该基因的mRNA杂交并阻止其翻译。反义调控核酸可以是DNA、RNA或其衍生物或杂交体。

[2034] 与目的转录物杂交的调节性核酸的长度可以在5至30个核苷酸之间,在约10至30个核苷酸之间,或约11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30个或更多个核苷酸。调节性核酸与靶向转录物的同一性程度应为至少75%、至少80%、至少85%、至少90%或至少95%。

[2035] 遗传元件可以编码调节性核酸,例如与靶基因的约5至约25个连续核苷酸相同的微小RNA(miRNA)分子。在一些实施例中,miRNA序列靶向mRNA并从二核苷酸AA开始,其GC含量为约30%-70%(约30%-60%、约40%-60%或约45%-55%),并且例如通过标准BLAST搜索测定,与要引入其中的哺乳动物基因组中的靶标以外的任何核苷酸序列不具有高百分比同一性。

[2036] 在一些实施例中,调节性核酸是至少一种miRNA,例如2、3、4、5、6种或更多种。在一些实施例中,遗传元件包含编码miRNA的序列,所述miRNA与本文所述的核苷酸序列或与序列互补的序列中任一个具有至少约75%、80%、85%、90%、95%、96%、97%、98%、99%或100%核苷酸序列同一性。

[2037] siRNA和shRNA类似于内源性微小RNA(miRNA)基因的加工途径中的中间体(Bartel,Cell[细胞]116:281-297,2004)。在一些实施例中,siRNA可以用作miRNA,反之亦然(Zeng等人,Mol Cell[分子细胞学]9:1327-1333,2002;Doench等人,Genes Dev[基因与发育]17:438-442,2003)。像siRNA一样,微小RNA使用RISC来下调靶基因,但与siRNA不同,大多数动物miRNA都不切割mRNA。相反,miRNA通过翻译抑制或聚A去除和mRNA降解降低蛋白质输出(Wu等人,Proc Natl Acad Sci USA[美国科学院院报]103:4034-4039,2006)。已知的miRNA结合位点位于mRNA 3'UTR内;miRNA似乎靶向与miRNA 5'端的2-8个核苷酸几乎完全互补的位点(Rajewsky,Nat Genet[自然遗传学]38增刊:S8-13,2006;Lim等人,Nature[自然]433:769-773,2005)。该区域称为种子区域。由于siRNA和miRNA是可互换的,因此外源siRNA下调与siRNA具有种子互补性的mRNA(Birmingham等人,Nat Methods[自然方法]3:199-204,2006)。3'UTR内的多个靶位点会产生更强的下调(Doench等人,Genes Dev[基因与发育]17:438-442,2003)。

[2038] 已知miRNA序列的列表可在研究组织维护的数据库中找到,这些组织例如维康信托基金会桑格研究院(Wellcome Trust Sanger Institute)、宾夕法尼亚生物信息学中心(Penn Center for Bioinformatics)、斯隆凯特灵癌症中心(Memorial Sloan Kettering Cancer Center)和欧洲分子生物学实验室(European Molecule Biology Laboratory)等。

已知的有效的siRNA序列和同源结合位点也很好地呈现在相关文献中。通过本领域已知的技术,RNAi分子易于设计和生产的。此外,存在增加发现有效并且特异性的基序的机会的计算工具(Lagana等人,Methods Mol.Bio.[分子生物学方法],2015,1269:393-412)。

[2039] 调节性核酸可以调节基因编码的RNA的表达。因为多个基因可以彼此共享一定程度的序列同源性,所以在一些实施例中,可以将调节性核酸设计为靶向具有足够序列同源性的一类基因。在一些实施例中,调节性核酸可包含与在不同基因靶之间共享的序列互补的或特定基因靶所特有的序列。在一些实施例中,可以将调节性核酸设计为靶向在几个基因之间具有同源性的RNA序列的保守区,从而靶向基因家族中的几个基因(例如,不同的基因同工型、剪接变体、突变基因等)。在一些实施例中,可以将调节性核酸设计为靶向单一基因的特定RNA序列所特有的序列。

[2040] 在一些实施例中,遗传元件可以包括一个或多个编码调节一个或多个基因表达的调节性核酸的序列。

[2041] 在一个实施例中,本文其他地方描述的gRNA被用作用于基因编辑的CRISPR系统的一部分。出于基因编辑的目的,可将指环体设计为包括一个或多个与所需的靶DNA序列相对应的指导RNA序列;参见,例如,Cong等人(2013)Science[科学],339:819-823;Ran等人(2013)Nature Protocols[自然实验手册],8:2281-2308。gRNA序列的至少约16或17个核苷酸通常允许Cas9介导的DNA切割发生;对于Cpf1,需要gRNA序列的至少约16个核苷酸,从而实现可检测的DNA切割。

[2042] 治疗性肽或多肽

[2043] 在一些实施例中,遗传元件包含编码治疗性肽或多肽,例如蛋白质替代治疗剂的序列。这样的治疗剂包括但不限于小肽、拟肽(例如类肽)、氨基酸和氨基酸类似物。这样的治疗剂通常具有小于约5,000克/摩尔的分子量、小于约2,000克/摩尔的分子量、小于约1,000克/摩尔的分子量、小于约500克/摩尔的分子量和此类化合物的盐、酯和其他药学上可接受的形式。这样的治疗剂可以包括但不限于神经递质、激素、药物、毒素、病毒或微生物颗粒、合成分子及其激动剂或拮抗剂。

[2044] 在一些实施例中,遗传元件包括编码肽例如治疗性肽的序列。肽可以是直链或支链的。所述肽的长度为约5至约500个氨基酸、约15至约400个氨基酸、约20至约325个氨基酸、约25至约250个氨基酸、约50至约150个氨基酸或其间的任何范围。

[2045] 示例性效应子在本文中描述,例如在下表中。

[2046] 表A.示例性酶促效应子和相应的适应症

效应子	适应症
3-甲基巴豆酰-辅酶 A 羧化酶	3-甲基巴豆酰-辅酶 A 羧化酶缺乏症
乙酰辅酶 A-氨基葡萄糖 N-乙酰基转移酶	粘多糖贮积症 MPS III（沙费利波（Sanfilippo）综合征）III-C 型
ADAMTS13	血栓性血小板减少性紫癜
腺嘌呤磷酸核糖基转移酶	腺嘌呤磷酸核糖基转移酶缺乏症
腺苷脱氨酶	腺苷脱氨酶缺乏症
ADP-核糖蛋白水解酶	谷酰基核糖-5-磷酸贮积病
α-葡萄糖苷酶	糖原贮积病 2 型（庞贝病）
氨单加氧酶	精氨酸血症
氨单加氧酶	精氨酸琥珀酸尿症
氨单加氧酶	瓜氨酸血症 I 型
氨单加氧酶	瓜氨酸血症 II 型
氨单加氧酶	高血氨症/鸟氨酸血症/瓜氨酸血症（鸟氨酸转运蛋白缺陷）
精氨酸酶	家族性高精氨酸血症
芳基硫酸酯酶 A	异染性脑白质营养不良
组织蛋白酶 K	致密成骨不全症
神经酰胺酶	法伯氏病（脂肪肉芽肿病）
胱硫醚 B 合酶	高胱氨酸尿症
多萜醇-P-甘露糖合酶	N-糖基化先天性疾病 CDG Ie
多萜醇-P-Glc：Man9GlcNAc2-PP-多萜醇葡萄糖基转移酶	N-糖基化先天性疾病 CDG Ic
多萜醇-P-Man：Man5GlcNAc2-PP-多萜醇甘露糖基转移酶	N-糖基化先天性疾病 CDG Id
多萜基-P-葡萄糖：Glc-1-Man-9-GlcNAc-2-PP-多萜基-α-3-葡萄糖基转移酶	N-糖基化先天性疾病 CDG Ih
多萜基-P-甘露糖：Man-7-GlcNAc-2-PP-多萜基-α-6-甘露糖基转移酶	N-糖基化先天性疾病 CDG Ig

[2047]

[2048]

因子 II	因子 II 缺乏症
因子 IX	血友病 B
因子 V	欧伦氏病（Owren's disease）
因子 VIII	血友病 A
因子 X	Stuart-Prower 因子缺乏症
因子 XI	血友病 C
因子 XIII	纤维蛋白稳定因子缺乏症
半乳糖胺-6-硫酸酯硫酸酯酶	粘多糖贮积症 MPS IV（莫奎奥综合征（Morquio's syndrome））IV-A 型
半乳糖脱氢酶	转移酶缺陷型半乳糖血症（半乳糖血症 1 型）
半乳糖基神经酰胺 β-半乳糖苷酶	克拉布病（Krabbe's disease）
神经节苷脂 β-半乳糖苷酶	GM1 神经节苷脂沉积症，全身性
神经节苷脂 β-半乳糖苷酶	GM2 神经节苷脂沉积症
神经节苷脂 β-半乳糖苷酶	神经鞘脂病 I 型
神经节苷脂 β-半乳糖苷酶	神经鞘脂病 II 型（幼年型）
神经节苷脂 β-半乳糖苷酶	神经鞘脂病 III 型（成人型）
葡萄糖苷酶 I	N-糖基化先天性疾病 CDG IIb
葡萄糖神经酰胺 β-葡萄糖苷酶	戈谢病（Gaucher's disease）
乙酰肝素-S-硫酸酯磺酰胺酶	粘多糖贮积症 MPS III（沙费利波综合征）III-A 型
黑尿酸氧化酶	黑尿酸尿症
透明质酸酶	粘多糖贮积症 MPS IX(透明质酸酶缺乏症)
硫酸艾杜糖酸硫酸酯酶	粘多糖贮积症 MPS II(亨特综合征(Hunter's syndrome))
卵磷脂-胆固醇酰基转移酶（LCAT）	完全 LCAT 缺乏、鱼眼病、动脉粥样硬化、高胆固醇血症
亮氨酸代谢酶	异戊酸血症
亮氨酸代谢酶	枫糖尿病
赖氨酸氧化酶	戊二酸血症 I 型
溶酶体酸性脂肪酶	胆固醇酯贮积病（CESD）
溶酶体酸性脂肪酶	溶酶体酸性脂肪酶缺乏症
溶酶体酸性脂肪酶	沃尔曼病（Wolman's disease）
溶酶体胃酶抑素不敏感肽酶	蜡样脂褐质沉积症晚期婴儿型（CLN2，比-詹二氏病（Jansky-Bielschowsky disease））
甘露糖（Man）磷酸（P）异构酶	N-糖基化先天性疾病 CDG Ib

[2049]

甘露糖基-α-1,6-糖蛋白-β-1,2-N-乙酰基氨基葡萄糖转移酶	N-糖基化先天性疾病 CDG IIa
金属蛋白酶-2	温彻斯特综合征（Winchester syndrome）
甲基丙二酰辅酶 A 变位酶	甲基丙二酸血症（维生素 b12 无应答）
N-乙酰半乳糖胺 α-4-硫酸酯硫酸酯酶（芳基硫酸酯酶 B）	粘多糖贮积症 MPS VI（拉-米氏综合征（Maroteaux-Lamy syndrome））
N-乙酰-D-氨基葡萄糖苷酶	粘多糖贮积症 MPS III（沙费利波综合征）III-B 型
N-乙酰基-氨基半乳糖苷酶	辛德勒病（Schindler's disease）I 型（婴儿重症型）
N-乙酰基-氨基半乳糖苷酶	辛德勒病 II 型（神崎病，成人发病型）
N-乙酰基-氨基半乳糖苷酶	辛德勒病 III 型（中间型）
N-乙酰基-氨基葡萄糖-6-硫酸酯硫酸酯酶	粘多糖贮积症 MPS III（沙费利波综合征）III-D 型
N-乙酰基氨基葡萄糖基-1-磷酸转移酶	粘脂贮积症 ML III(假性贺勒氏多营养不良（pseudo-Hurler's polydystrophy）)
N-乙酰基氨基葡萄糖基-1-磷酸转移酶催化亚基	黏脂贮积症 ML II（I 细胞病）
N-乙酰基氨基葡萄糖基-1-磷酸转移酶，底物识别亚基	黏脂贮积症 ML III（假性贺勒氏多营养不良）III-C 型
N-天冬氨酰基氨基葡萄糖苷酶	天冬氨酰基葡萄糖胺尿
神经氨酸酶 1（唾液酸酶）	唾液酸沉积症
草酸氧化酶	原发性高草酸尿症
棕榈酰基-蛋白硫酯酶-1	蜡样脂褐质沉积症成人型（CLN4，库夫斯病（Kufs' disease））
棕榈酰基-蛋白硫酯酶-1	蜡样脂褐质沉积症婴儿型（CLN1，桑塔沃里-哈尔蒂亚病（Santavuori-Haltia disease））
苯丙氨酸羟化酶	苯丙酮尿症
磷酸甘露糖变位酶-2	N-糖基化先天性疾病 CDG Ia（仅神经系统和神经系统-多内脏形式）
胆色素原脱氨酶	急性间歇性卟啉症
嘌呤核苷磷酸化酶	嘌呤核苷磷酸化酶缺乏症
嘧啶 5'核苷酸酶	溶血性贫血和/或嘧啶 5'核苷酸酶缺乏症
鞘磷脂酶	尼曼-匹克病（Niemann-Pick disease）A 型
鞘磷脂酶	尼曼-匹克病 B 型
甾醇 27-羟化酶	脑腱黄瘤病（胆甾烷醇脂质沉积症）
胸苷磷酸化酶	线粒体神经胃肠脑肌病（MNGIE）

[2050]

三己糖基神经酰胺 α-半乳糖苷酶	法布里病（Fabry's disease）
酪氨酸酚裂解酶	酪氨酸血症 1 型
酪氨酸酶，例如 OCA1	白化病，例如眼白化病
UDP-GlcNAc：多萜基-P NAcGlc 磷酸转移酶	N-糖基化先天性疾病 CDG Ij
UDP-N-乙酰基氨基葡萄糖-2-差向异构酶/N-乙酰基甘露糖胺激酶，唾液酸	Sialuria 法国型
尿酸酶	莱施-奈恩综合征(Lesch-Nyhan syndrome)，痛风
尿苷二磷酸葡萄糖醛酸基转移酶（例如，UGT1A1）	克里格勒-纳贾尔综合征
α-1,2-甘露糖基转移酶	N-糖基化先天性疾病 CDG Il（608776）
α-1,2-甘露糖基转移酶	N-糖基化先天性疾病 I 型(前高尔基体糖基化缺陷)
α-1,3-甘露糖基转移酶	N-糖基化先天性疾病 CDG Ii
α-D-甘露糖苷酶	α-甘露糖苷贮积症 I 型（严重）或 II 型（轻度）
α-L-岩藻糖苷酶	岩藻糖苷贮积症
α-l-艾杜糖苷酶	粘多糖贮积症 MPS I H/S（贺勒-施艾综合征（Hurler-Scheie syndrome)）
α-l-艾杜糖苷酶	粘多糖贮积症 MPS I-H（贺勒综合征）
α-l-艾杜糖苷酶	粘多糖贮积症 MPS I-S（施艾综合征（Scheie's syndrome)）
β-1,4-半乳糖基转移酶	N-糖基化先天性疾病 CDG IId
β-1,4-甘露糖基转移酶	N-糖基化先天性疾病 CDG Ik
β-D-甘露糖苷酶	β-甘露糖苷贮积症
β-半乳糖苷酶	粘多糖贮积症 MPS IV(莫奎奥综合征)IV-B 型
β-葡萄糖醛酸酶	粘多糖贮积症 MPS VII（斯莱综合征（Sly's syndrome)）
β-己糖胺酶 A	泰-萨氏病（Tay-Sachs disease）
β-己糖胺酶 B	桑德霍夫病（Sandhoff's disease）

[2051] 在一些实施例中，本文所述的效应子包含表A的酶或其功能变体，例如同源物(例如直向同源物或旁系同源物)或片段。在一些实施例中，功能变体催化与相应野生型蛋白质相同的反应，例如，催化速率比野生型蛋白低不少于10％、20％、30％、40％或50％。在一些实施例中，编码表A的酶或其功能变体的指环体用于治疗表A的疾病或障碍。在一些实施例中，指环体用于递送尿苷二磷酸葡糖醛酸转移酶或其功能变体至靶细胞，例如肝细胞。在一些实施例中，使用指环体将OCA1或其功能变体递送至靶细胞，例如视网膜细胞。

[2052] 表B.示例性非酶促效应子和相应的适应症

[2053]

效应子	适应症
存活运动神经元蛋白（SMN）	脊髓性肌萎缩
肌营养不良蛋白或微肌营养不良蛋白	肌营养不良症（例如，杜氏肌营养不良症或贝克尔肌营养不良症）
补体蛋白质，例如补体因子 C1	补体因子 I 缺乏症
补体因子 H	非典型溶血性尿毒症综合征
胱氨酸转运蛋白（溶酶体胱氨酸转运蛋白）	胱氨酸病
附睾分泌蛋白 1（HE1；NPC2 蛋白）	尼曼-匹克病 C2 型
GDP-岩藻糖转运蛋白-1	N-糖基化先天性疾病 CDG IIc（Rambam-Hasharon 综合征）
GM2 激活蛋白	GM2 激活蛋白缺乏症（泰-萨氏病 AB 变体，GM2A）
溶酶体跨膜 CLN3 蛋白	蜡样脂褐质沉积症幼年型（CLN3、贝敦氏症（Batten disease）、伏-施-斯三氏病（Vogt-Spielmeyer disease））
溶酶体跨膜 CLN5 蛋白	蜡样脂褐质沉积症变体晚期婴儿型，芬兰型（CLN5）
磷酸钠协同转运蛋白，唾液酸转运蛋白	婴儿唾液酸贮积症
磷酸钠协同转运蛋白，唾液酸转运蛋白	Sialuria 芬兰型（唾液酸贮积病）
NPC1 蛋白	尼曼-匹克病 C1 型/D 型
寡聚高尔基复合体-7	N-糖基化先天性疾病 CDG IIe
鞘脂激活蛋白原（Prosaposin）	鞘脂激活蛋白原缺乏症
保护性蛋白/组织蛋白酶 A（PPCA）	半乳糖唾液酸贮积症（戈德堡综合征（Goldberg's syndrome），神经氨酸酶和 β-半乳糖苷酶缺乏症

	联合）
参与甘露糖-P-多萜醇利用的蛋白质	N-糖基化先天性疾病 CDG If
皂素 B	皂素 B 缺乏症（硫苷脂激活剂缺乏症）
皂素 C	皂素 C 缺乏症（戈谢激活剂（Gaucher's activator）缺乏症）
硫酸酯酶修饰因子-1	黏脑硫脂贮积症（多种硫酸酯酶缺乏症）
跨膜 CLN6 蛋白	蜡样脂褐质沉积症变异晚期婴儿型（CLN6）
跨膜 CLN8 蛋白	蜡样脂褐质沉积症进行性癫痫伴智力障碍
vWF	血管性血友病
因子 I（纤维蛋白原）	无纤维蛋白原血症

[2054]

[2055] 在一些实施例中，本文所述的效应子包含促红细胞生成素(EPO)，例如人促红细胞生成素(hEPO)或其功能变体。在一些实施例中，编码促红细胞生成素或其功能变体的指环体用于刺激红细胞生成。在一些实施例中，编码促红细胞生成素或其功能变体的指环体用于治疗疾病或障碍，例如贫血。在一些实施例中，使用指环体将EPO或其功能变体递送至靶细胞，例如红细胞。

[2056] 在一些实施例中，本文所述的效应子包含表B的多肽或其功能变体，例如同源物(例如，直向同源物或旁系同源物)或片段。在一些实施例中，编码表B的多肽或其功能变体的指环体用于治疗表B的疾病或障碍。在一些实施例中，指环体用于将SMN或其功能变体递送至靶细胞，例如脊髓和/或运动神经元的细胞。在一些实施例中，使用指环体将微肌营养不良蛋白递送至靶细胞，例如，肌细胞。

[2057] 示例性的微肌营养不良蛋白描述于Duan，“Systemic AAV Micro-dystrophin Gene Therapy for Duchenne Muscular Dystrophy.[杜氏肌营养不良症的全身AAV微肌营养不良蛋白基因治疗]”Mol Ther.[分子疗法]2018年10月3日;26(10):2337-2356.doi:10.1016/j.ymthe.2018.07.011.2018年7月17日电子公开。

[2058] 在一些实施例中，本文所述的效应子包含凝血因子，例如本文表A或表B中所列的凝血因子。在一些实施例中，本文所述的效应子包含在突变时引起溶酶体贮积症的蛋白质，例如本文表A或表B中列出的蛋白质。在一些实施例中，本文所述的效应子包含转运蛋白，例如本文表B中所列的转运蛋白。

[2059] 在一些实施例中，野生型蛋白质的功能变体包含具有野生型蛋白质的一种或多种活性的蛋白质，例如，功能变体催化与相应野生型蛋白质相同的反应，例如，催化速率比野生型蛋白低不少于10％、20％、30％、40％或50％。在一些实施例中，功能变体结合与野生型蛋白结合的相同结合配偶体，例如，Kd比相应野生型蛋白质在相同条件下对于相同结合配偶体的Kd高不超过10％、20％、30％、40％或50％。在一些实施例中，功能变体具有与野生多

肽的多肽序列至少70％、75％、80％、85％、90％、95％、96％、97％、98％或99％相同的多肽序列。在一些实施例中，功能变体包含相应野生型蛋白质的同源物(例如直向同源物或旁系同源物)。在一些实施例中，功能变体是融合蛋白。在一些实施例中，融合包含与相应野生型蛋白质具有至少70％、75％、80％、85％、90％、95％、96％、97％、98％或99％同一性的第一区域，和第二异源区域。在一些实施例中，功能变体包含相应野生型蛋白质的片段或由其组成。

[2060] 肽的一些实例包括但不限于荧光标签或标记、抗原、治疗性肽、天然生物活性肽的合成或类似物肽、激动剂或拮抗剂肽、抗微生物肽、靶向或细胞毒性肽、降解或自毁肽、以及多种降解或自毁肽。本文所述的可用于本发明的肽还包括抗原结合肽，例如抗原结合抗体或抗体样片段，例如单链抗体、纳米抗体(参见，例如，Steeland等人2016.Nanobodies as therapeutics:big opportunities for small antibodies.[纳米抗体作为治疗剂:小抗体的巨大机会]Drug Discov Today[今日药物发现]:21(7):1076-113)。这样的抗原结合肽可以结合细胞质抗原、核抗原或细胞内抗原。

[2061] 在一些实施例中，遗传元件包括编码蛋白质例如治疗性蛋白质的序列。治疗性蛋白质的一些实例可以包括但不限于激素、细胞因子、酶、抗体、转录因子、受体(例如膜受体)、配体、膜转运蛋白、分泌的蛋白质、肽、载剂蛋白、结构蛋白、核酸酶或其组分。

[2062] 在一些实施例中，本文所述的组合物或指环体包括与能够靶向特定位置、组织或细胞的配体连接的多肽。

[2063] 调节性序列

[2064] 在一些实施例中，遗传元件包含与编码效应子的序列可操作地连接的调控序列，例如启动子或增强子。

[2065] 在一些实施例中，启动子包括与编码表达产物的DNA序列相邻的DNA序列。启动子可以可操作地连接至相邻的DNA序列。与不存在启动子时表达的产物的量相比，启动子通常增加DNA序列表达的产物的量。来自一种生物体的启动子可用于增强来自另一生物体的DNA序列的产物表达。例如，脊椎动物启动子可用于在脊椎动物中表达水母GFP。另外，一个启动子元件可以增加串联连接的多个DNA序列表达的产物的量。因此，一种启动子元件可以增强一种或多种产物的表达。多个启动子元件是本领域普通技术人员众所周知的。

[2066] 在一个实施例中，期望高水平的组成型表达。此类启动子的实例包括但不限于逆转录病毒劳斯肉瘤病毒(RSV)长末端重复(LTR)启动子/增强子、巨细胞病毒(CMV)立即早期启动子/增强子(参见，例如，Boshart等人，Cell[细胞]，41:521-530(1985))、SV40启动子、二氢叶酸还原酶启动子、细胞质β-肌动蛋白启动子和磷酸甘油激酶(PGK)启动子。

[2067] 在另一个实施例中，可能期望诱导型启动子。诱导型启动子是由外源供应的化合物以顺式或反式调节的那些启动子，包括但不限于锌诱导的绵羊金属硫蛋白(MT)启动子；地塞米松(Dex)诱导型小鼠乳腺肿瘤病毒(MMTV)启动子；T7聚合酶启动子系统(WO 98/10088)；四环素阻抑型系统(Gossen等人，Proc.Natl.Acad.Sci.USA[美国科学院院报]，89:5547-5551(1992))；四环素诱导型系统(Gossen等人，Science[科学]，268:1766-1769(1995)；还参见Harvey等人，Curr.Opin.Chem.Biol.[化学生物学新见]，2:512-518(1998))；RU486诱导型系统(Wang等人，Nat.Biotech.[自然生物技术]，15:239-243(1997)和Wang等人，Gene Ther.[基因治疗]，4:432-441(1997))；和雷帕霉素诱导型系统(Magari

等人,J.Clin.Invest.[临床投资杂志],100:2865-2872(1997);Rivera等人,Nat.Medicine.[自然医学]2:1028-1032(1996))。可在本文中使用的其他类型的诱导型启动子是那些由特定的生理状态例如温度、急性期或仅在复制细胞中调节的启动子。

[2068] 在一些实施例中,使用目的基因或核酸序列的天然启动子。当期望基因或核酸序列的表达应模拟天然表达时,可以使用天然启动子。当基因或其他核酸序列的表达必须在时间或发育上,或以组织特异性方式或响应于特定转录刺激进行调节时,可以使用天然启动子。在另一个实施例中,也可以使用其他天然表达控制元件,例如增强子元件、聚腺苷酸化位点或Kozak共有序列来模拟天然表达。

[2069] 在一些实施例中,遗传元件包含可操作地连接至组织特异性启动子的基因。例如,如果需要在骨骼肌中表达,则可以使用在肌肉中有活性的启动子。这些包括来自编码骨骼肌α-肌动蛋白、肌球蛋白轻链2A、肌营养不良蛋白、肌肉肌酸激酶、以及具有比天然启动子更高活性的合成肌肉启动子的启动子。参见Li等人,Nat.Biotech.[自然生物技术],17:241-245(1999)。对于以下的组织特异性的启动子的实例是已知的:肝白蛋白,Miyatake等人J.Virol.[病毒学杂志],71:5124-32(1997);乙型肝炎病毒核心启动子,Sandig等人,Gene Ther.[基因治疗]3:1002-9(1996);甲胎蛋白(AFP),Arbuthnot等人,Hum.Gene Ther.[人类基因治疗],7:1503-14(1996),骨骼(骨钙素,Stein等人,Mol.Biol.Rep.[分子生物学报告],24:185-96(1997);骨骼唾液蛋白,Chen等人,J.Bone Miner.Res.[骨与矿物研究杂志]11:654-64(1996)),lymphocytes[淋巴细胞](CD2,Hansal等人,J.Immunol.[免疫学杂志],161:1063-8(1998);免疫球蛋白重链;T细胞受体a链),神经元(神经元特异性烯醇化酶(NSE)启动子,Andersen等人Cell.Mol.Neurobiol.[细胞与分子神经生物学],13:503-15(1993);神经丝轻链基因,Piccioli等人,Proc.Natl.Acad.Sci.USA[美国科学院院报],88:5611-5(1991);神经元特异性vgf基因,Piccioli等人,Neuron[神经元],15:373-84(1995)];等。

[2070] 遗传元件可包括增强子,例如与编码基因的DNA序列相邻的DNA序列。增强子元件通常位于启动子元件的上游,或者可以位于编码DNA序列(例如,转录或翻译成一种或多种产物的DNA序列)的下游或之内。因此,增强子元件可位于编码产物的DNA序列上游或下游的100个碱基对、200个碱基对或300个或更多个碱基对。增强子元件可以增加DNA序列表达的重组产物的量,超出由启动子元件提供的增加的表达。对于本领域普通技术人员而言,多个增强子元件是容易获得的。

[2071] 在一些实施例中,遗传元件包括在编码本文所述的表达产物的序列侧翼的一个或多个反向末端重复序列(ITR)。在一些实施例中,遗传元件包含在编码本文所述的表达产物的序列侧翼的一个或多个长末端重复序列(LTR)。可以使用的启动子序列的实例包括但不限于猿猴病毒40(SV40)早期启动子、小鼠乳腺肿瘤病毒(MMTV)、人免疫缺陷病毒(HIV)长末端重复序列(LTR)启动子、MoMuLV启动子、禽白血病病毒启动子、Epstein-Barr病毒立即早期启动子和劳斯肉瘤病毒启动子。

[2072] 复制蛋白

[2073] 在一些实施例中,指环体例如合成的指环体的遗传元件可包括编码一种或多种复制蛋白的序列。在一些实施例中,指环体可以通过滚环复制法复制,例如,前导链和滞后链的合成是不偶联的。在这样的实施例中,指环体包含三个附加元件:i)编码起始蛋白的基

因,ii)双链起点,和iii)单链起点。包含复制蛋白的滚环复制(RCR)蛋白复合物与前导链结合并使复制起点去稳定。RCR复合物切割基因组以产生游离的3'OH末端。细胞DNA聚合酶从游离的3'OH末端开始病毒DNA复制。复制基因组后,RCR复合物共价地闭合环。这导致释放出正的环状单链亲本DNA分子和由负的亲本链和新合成的正链组成的环状双链DNA分子。单链DNA分子可以被包壳或涉及第二轮复制。参见例如Virology Journal[病毒学杂志]2009,6:60doi:10.1186/1743-422X-6-60。

[2074] 遗传元件可包含编码聚合酶,例如RNA聚合酶或DNA聚合酶的序列。

[2075] 其他序列

[2076] 在一些实施例中,遗传元件还包括编码产物(例如核酶、编码蛋白质的治疗性mRNA、外源基因)的核酸。

[2077] 在一些实施例中,遗传元件包括一个或多个影响指环体在宿主或宿主细胞中的物种和/或组织和/或细胞嗜性(例如衣壳蛋白序列)、感染性(例如衣壳蛋白序列)、免疫抑制/激活(例如调节核酸)、病毒基因组结合和/或包装、免疫逃逸(非免疫原性和/或耐受性)、药代动力学、内吞作用和/或细胞附着、核进入、细胞内调节和定位、胞吐作用调节、繁殖和核酸保护的序列。

[2078] 在一些实施例中,遗传元件可以包含其他序列,其包括DNA、RNA或人工核酸。其他序列可以包括但不限于基因组DNA,cDNA或编码tRNA、mRNA、rRNA、miRNA、gRNA、siRNA或其他RNAi分子的序列。在一个实施例中,遗传元件包括编码siRNA的序列,以靶向与调节性核酸相同的基因表达产物的不同基因座。在一个实施例中,遗传元件包括编码siRNA的序列,以靶向与调节性核酸不同的基因表达产物。

[2079] 在一些实施例中,遗传元件进一步包含以下序列中的一个或多个:编码一个或多个miRNA的序列、编码一个或多个复制蛋白的序列、编码外源基因的序列、编码治疗剂的序列、调节性序列(例如启动子、增强子)、编码一个或多个靶向内源基因(siRNA、lncRNA、shRNA)的调节性序列的序列和编码治疗性mRNA或蛋白质的序列。

[2080] 其他序列的长度可为约2nt至约5000nt、约10nt至约100nt、约50nt至约150nt、约100nt至约200nt、约150nt至约250nt、约200至约300nt、约250nt至约350nt、约300nt至约500nt、约10nt至约1000nt、约50nt至约1000nt、约100nt至约1000nt、约1000nt至约2000nt、约2000nt至约3000nt、约3000nt至约4000nt、约4000nt至约5000nt或其间的任何范围。

[2081] 编码的基因

[2082] 例如,遗传元件可以包括与信号传导生化途径相关的基因,例如与信号传导生化途径相关的基因或多核苷酸。实例包括与疾病相关的基因或多核苷酸。"与疾病相关的"基因或多核苷酸是指与非疾病对照的组织或细胞相比,在患病组织衍生的细胞中以异常水平或异常形式产生转录或翻译产物的任何基因或多核苷酸。它可能是会以异常高水平表达的基因;它可能是会以异常低水平表达的基因,其中表达的改变与疾病的发生和/或进展相关。与疾病相关的基因还指具有直接引起病因或与引起病因的一个或多个基因连锁不平衡的一个或多个突变或基因变异的基因。

[2083] 与疾病相关的基因和多核苷酸的实例可得自约翰·霍普金斯大学的麦考斯克-纳森斯遗传医学研究所(马里兰州巴尔的摩)(McKusick-Nathans Institute of Genetic Medicine,Johns Hopkins University(Baltimore,Md.))和国家医学图书馆的国家生物技

术信息中心(马里兰州贝塞斯达)(National Center for Biotechnology Information,National Library of Medicine(Bethesda,Md.))。与疾病相关的基因和多核苷酸的实例列于美国专利号:8,697,359的表A-C中,所述专利通过引用整体并入本文。特定疾病信息可获自约翰·霍普金斯大学的麦考斯克-纳森斯遗传医学研究所(马里兰州巴尔的摩)(McKusick-Nathans Institute of Genetic Medicine,Johns Hopkins University(Baltimore,Md.))和国家医学图书馆的国家生物技术信息中心(马里兰州贝塞斯达)(National Center for Biotechnology Information,National Library of Medicine(Bethesda,Md.))。与信号传导生化途径相关的基因和多核苷酸的实例列于美国专利号:8,697,359的表A-C中,所述专利通过引用整体并入本文。

[2084] 此外,如本文其他地方所述,遗传元件可以编码靶向部分。这可以例如通过插入编码糖、糖脂或蛋白质例如抗体的多核苷酸来实现。本领域技术人员知道用于生成靶向部分的其他方法。

[2085] 病毒序列

[2086] 在一些实施例中,遗传元件包含至少一个病毒序列。在一些实施例中,所述序列与来自单链DNA病毒,例如指环病毒、比那病毒、圆环病毒、双生病毒、基诺病毒、丝状病毒、微小病毒、纳米病毒、细小病毒和斯派拉病毒的一个或多个序列具有同源性或同一性。在一些实施例中,所述序列与来自双链DNA病毒,例如腺病毒、瓶状病毒、囊泡病毒、非洲猪瘟病毒、杆状病毒、福斯罗病毒、球状病毒、滴状病毒、腺炎病毒、疱疹病毒、虹彩病毒、脂毛病毒、尼马病毒和痘病毒的一个或多个序列具有同源性或同一性。在一些实施例中,所述序列与来自RNA病毒,例如甲病毒、真菌传杆状病毒、肝炎病毒、大麦病毒、烟草花叶病毒、烟草脆裂病毒、三角病毒、风疹病毒、双RNA病毒、囊状病毒、分体病毒和呼肠病毒的一个或多个序列具有同源性或同一性。

[2087] 在一些实施例中,遗传元件可包含来自非致病性病毒,例如共生病毒(symbiotic virus),例如共栖病毒(commensal virus),例如天然病毒,例如指环病毒的一个或多个序列。命名法的最新变化已将三种能够感染人细胞的指环病毒划分为病毒的指环病毒科的甲型细环病毒(TT)、乙型细环病毒(TTM)和丙型细环病毒(TTMD)属。迄今为止,指环病毒尚未与任何人疾病相关联。在一些实施例中,遗传元件可包含与细环病毒(TT)具有同源性或同一性的序列,所述细环病毒(TT)是一种具有环状负义基因组的非包膜单链DNA病毒。在一些实施例中,遗传元件可包含与SEN病毒、前哨病毒、TTV样微小病毒和TT病毒具有同源性或同一性的序列。已经描述了不同类型的TT病毒,包括TT病毒基因型6、TT病毒群、TTV样病毒DXL1和TTV样病毒DXL2。在一些实施例中,遗传元件可包含与较小病毒、细环样微小病毒(TTM)或基因组大小在TTV和TTMV之间的第三种病毒(称为细环样中型病毒(TTMD))具有同源性或同一性的序列。在一些实施例中,遗传元件可包含与本文所述核苷酸序列中任一个具有至少约60%、70%、80%、85%、90%、95%、96%、97%、98%和99%核苷酸序列同一性的来自非致病性病毒的一个或多个序列或序列的片段。

[2088] 在一些实施例中,遗传元件可包含与本文所述例如表41中核苷酸序列中任一个具有至少约60%、70%、80%、85%、90%、95%、96%、97%、98%和99%核苷酸序列同一性的来自基本上非致病性病毒的一个或多个序列或序列的片段。

[2089] 表41:指环病毒及其序列的实例。登录号和相关序列信息可在

www.ncbi.nlm.nih.gov/genbank/上获得,以2018年12月11日作为参考。

[2090]

登录号	描述
AB017613.1	细环病毒 16 DNA，全基因组，分离株：TUS01
AB026345.1	ORF1 和 ORF2 的 TT 病毒基因，全 cds，分离株：TRM1
AB026346.1	ORF1 和 ORF2 的 TT 病毒基因，全 cds，分离株：TK16
AB026347.1	ORF1 和 ORF2 的 TT 病毒基因，全 cds，分离株：TP1-3
AB028669.1	ORF1 和 ORF2 的 TT 病毒基因，全基因组，分离株：TJN02
AB030487.1	pORF2a、pORF2b、pORF1 的 TT 病毒基因，全 cds，克隆：JaCHCTC19
AB030488.1	pORF2a、pORF2b、pORF1 的 TT 病毒基因，全 cds，克隆：JaBD89
AB030489.1	pORF2a、pORF2b、pORF1 的 TT 病毒基因，全 cds，克隆：JaBD98
AB038340.1	ORF2s、ORF1、ORF3 的 TT 病毒基因，全 cds
AB038622.1	ORF2、ORF1、ORF3 的 TT 病毒基因，全 cds，分离株：TTVyon-LC011
AB038623.1	ORF2、ORF1、ORF3 的 TT 病毒基因，全 cds，分离株：TTVyon-KC186
AB038624.1	ORF2、ORF1、ORF3 的 TT 病毒基因，全 cds，分离株：TTVyon-KC197
AB041821.1	VP1 的 TT 病毒 mRNA，全 cds
AB050448.1	ORF1、ORF2、ORF3、ORF4 的细环病毒基因，全 cds，分

[2091]

	离株：TYM9
AB060592.1	ORF1、ORF2、ORF3、ORF4 的细环病毒基因，克隆：SAa-39
AB060593.1	ORF1、ORF2、ORF3、ORF4 的细环病毒基因，全 cds，克隆：SAa-38
AB060595.1	ORF1、ORF2、ORF3、ORF4 的 TT 病毒基因，全 cds，克隆：SAj-30
AB060596.1	ORF1、ORF2、ORF3、ORF4 的 TT 病毒基因，全 cds，克隆：SAf-09
AB064596.1	细环病毒 DNA，全基因组，分离株：CT25F
AB064597.1	细环病毒 DNA，全基因组，分离株：CT30F
AB064599.1	细环病毒 DNA，全基因组，分离株：JT03F
AB064600.1	细环病毒 DNA，全基因组，分离株：JT05F
AB064601.1	细环病毒 DNA，全基因组，分离株：JT14F
AB064602.1	细环病毒 DNA，全基因组，分离株：JT19F
AB064603.1	细环病毒 DNA，全基因组，分离株：JT41F
AB064604.1	细环病毒 DNA，全基因组，分离株：CT39F
AB064606.1	细环病毒 DNA，全基因组，分离株：JT33F
AB290918.1	细环中型病毒 1 DNA，全基因组，分离株：MD1-073
AF079173.1	TT 病毒菌株 TTVCHN1，全基因组
AF116842.1	TT 病毒菌株 BDH1，全基因组
AF122914.3	TT 病毒分离株 JA20，全基因组
AF122917.1	TT 病毒分离株 JA4，全基因组
AF122919.1	TT 病毒分离株 JA10 未知基因
AF129887.1	TT 病毒 TTVCHN2，全基因组
AF247137.1	TT 病毒分离株 TUPB，全基因组
AF254410.1	TT 病毒 ORF2 蛋白质和 ORF1 蛋白质基因，全 cds
AF298585.1	TT 病毒 Polish 分离株 P/1C1，全基因组
AF315076.1	TTV 样病毒 DXL1 未知基因
AF315077.1	TTV 样病毒 DXL2 未知基因
AF345521.1	TT 病毒分离株 TCHN-G1 Orf2 和 Orf1 基因，全 cds
AF345522.1	TT 病毒分离株 TCHN-E Orf2 和 Orf1 基因，全 cds
AF345525.1	TT 病毒分离株 TCHN-D2 Orf2 和 Orf1 基因，全 cds
AF345527.1	TT 病毒分离株 TCHN-C2 Orf2 和 Orf1 基因，全 cds
AF345528.1	TT 病毒分离株 TCHN-F Orf2 和 Orf1 基因，全 cds

[2092]

AF345529.1	TT 病毒分离株 TCHN-G2 Orf2 和 Orf1 基因，全 cds
AF371370.1	TT 病毒 ORF1、ORF3、和 ORF2 基因，全 cds
AJ620212.1	细环病毒，分离株 tth6，全基因组
AJ620213.1	细环病毒，分离株 tth10，全基因组
AJ620214.1	细环病毒，分离株 tth11g2，全基因组
AJ620215.1	细环病毒，分离株 tth18，全基因组
AJ620216.1	细环病毒，分离株 tth20，全基因组
AJ620217.1	细环病毒，分离株 tth21，全基因组
AJ620218.1	细环病毒，分离株 tth3，全基因组
AJ620219.1	细环病毒，分离株 tth9，全基因组
AJ620220.1	细环病毒，分离株 tth16，全基因组
AJ620221.1	细环病毒，分离株 tth17，全基因组
AJ620222.1	细环病毒，分离株 tth25，全基因组
AJ620223.1	细环病毒，分离株 tth26，全基因组
AJ620224.1	细环病毒，分离株 tth27，全基因组
AJ620225.1	细环病毒，分离株 tth31，全基因组
AJ620226.1	细环病毒，分离株 tth4，全基因组
AJ620227.1	细环病毒，分离株 tth5，全基因组
AJ620228.1	细环病毒，分离株 tth14，全基因组
AJ620229.1	细环病毒，分离株 tth29，全基因组
AJ620230.1	细环病毒，分离株 tth7，全基因组
AJ620231.1	细环病毒，分离株 tth8，全基因组
AJ620232.1	细环病毒，分离株 tth13，全基因组
AJ620233.1	细环病毒，分离株 tth19，全基因组
AJ620234.1	细环病毒，分离株 tth22g4，全基因组
AJ620235.1	细环病毒，分离株 tth23，全基因组
AM711976.1	TT 病毒 sle1957 全基因组
AM712003.1	TT 病毒 sle1931 全基因组
AM712004.1	TT 病毒 sle1932 全基因组
AM712030.1	TT 病毒 sle2057 全基因组
AM712031.1	TT 病毒 sle2058 全基因组
AM712032.1	TT 病毒 sle2072 全基因组
AM712033.1	TT 病毒 sle2061 全基因组
AM712034.1	TT 病毒 sle2065 全基因组

[2093]

AY026465.1	TT 病毒分离株 L01 ORF2 和 ORF1 基因，全 cds
AY026466.1	TT 病毒分离株 L02 ORF2 和 ORF1 基因，全 cds
DQ003341.1	细环病毒克隆 P2-9-02 ORF2（ORF2）、ORF1A（ORF1A）、和 ORF1B（ORF1B）基因，全 cds
DQ003342.1	细环病毒克隆 P2-9-07 ORF2（ORF2）、ORF1A（ORF1A）、和 ORF1B（ORF1B）基因，全 cds
DQ003343.1	细环病毒克隆 P2-9-08 ORF2（ORF2）、ORF1A（ORF1A）、和 ORF1B（ORF1B）基因，全 cds
DQ003344.1	细环病毒克隆 P2-9-16 ORF2（ORF2）、ORF1A（ORF1A）、和 ORF1B（ORF1B）基因，全 cds
DQ186994.1	细环病毒克隆 P601 ORF2（ORF2）和 ORF1（ORF1）基因，全 cds
DQ186995.1	细环病毒克隆 P605 ORF2（ORF2）和 ORF1（ORF1）基因，全 cds
DQ186996.1	细环病毒克隆 BM1A-02 ORF2（ORF2）和 ORF1（ORF1）基因，全 cds
DQ186997.1	细环病毒克隆 BM1A-09 ORF2（ORF2）和 ORF1（ORF1）基因，全 cds
DQ186998.1	细环病毒克隆 BM1A-13 ORF2（ORF2）和 ORF1（ORF1）基因，全 cds
DQ186999.1	细环病毒克隆 BM1B-05 ORF2（ORF2）和 ORF1（ORF1）基因，全 cds
DQ187000.1	细环病毒克隆 BM1B-07 ORF2（ORF2）和 ORF1（ORF1）基因，全 cds
DQ187001.1	细环病毒克隆 BM1B-11 ORF2（ORF2）和 ORF1（ORF1）基因，全 cds
DQ187002.1	细环病毒克隆 BM1B-14 ORF2（ORF2）和 ORF1（ORF1）基因，全 cds
DQ187003.1	细环病毒克隆 BM1B-08 ORF2（ORF2）基因，全 cds; 和非功能性 ORF1（ORF1）基因，全序列
DQ187004.1	细环病毒克隆 BM1C-16 ORF2（ORF2）和 ORF1（ORF1）基因，全 cds
DQ187005.1	细环病毒克隆 BM1C-10 ORF2（ORF2）和 ORF1（ORF1）基因，全 cds
DQ187007.1	细环病毒克隆 BM2C-25 ORF2（ORF2）基因，全 cds; 和非功能性 ORF1（ORF1）基因，全序列

[2094]

DQ361268.1	细环病毒分离株 ViPi04 ORF1 基因，全 cds
EF538879.1	细环病毒分离株 CSC5 ORF2 和 ORF1 基因，全 cds
EU305675.1	细环病毒分离株 LTT7 ORF1 基因，全 cds
EU305676.1	细环病毒分离株 LTT10 ORF1 基因，全 cds
EU889253.1	细环病毒分离株 ViPi08 非功能性 ORF1 基因，全序列
FJ392105.1	细环病毒分离株 TW53A25 ORF2 基因，部分 cds；和 ORF1 基因，全 cds
FJ392107.1	细环病毒分离株 TW53A27 ORF2 基因，部分 cds；和 ORF1 基因，全 cds
FJ392108.1	细环病毒分离株 TW53A29 ORF2 基因，部分 cds；和 ORF1 基因，全 cds
FJ392111.1	细环病毒分离株 TW53A35 ORF2 基因，部分 cds；和 ORF1 基因，全 cds
FJ392112.1	细环病毒分离株 TW53A39 ORF2 基因，部分 cds；和 ORF1 基因，全 cds
FJ392113.1	细环病毒分离株 TW53A26 ORF2 基因，全 cds；和非功能性 ORF1 基因，全序列
FJ392114.1	细环病毒分离株 TW53A30 ORF2 和 ORF1 基因，全 cds
FJ392115.1	细环病毒分离株 TW53A31 ORF2 和 ORF1 基因，全 cds
FJ392117.1	细环病毒分离株 TW53A37 ORF1 基因，全 cds
FJ426280.1	细环病毒菌株 SIA109，全基因组
FR751500.1	细环病毒全基因组，分离株 TTV-HD23a（rheu215）
GU797360.1	细环病毒克隆 8-17，全基因组
HC742700.1	来自专利 WO 2010044889 的序列 7
HC742710.1	来自专利 WO 2010044889 的序列 17
JX134044.1	TTV 样微小病毒分离株 TTMV_LY1，全基因组
JX134045.1	TTV 样微小病毒分离株 TTMV_LY2，全基因组
KU243129.1	TTV 样微小病毒分离株 TTMV-204，全基因组
KY856742.1	TTV 样微小病毒分离株镇江（zhenjiang），全基因组
LC381845.1	细环病毒人/日本/KS025/2016 DNA，全基因组
MH648892.1	指环病毒科物种分离株 ctdc048，全基因组
MH648893.1	指环病毒科物种分离株 ctdh007，全基因组
MH648897.1	指环病毒科物种分离株 ctcb038，全基因组
MH648900.1	指环病毒科物种分离株 ctfc019，全基因组

[2095]

MH648901.1	指环病毒科物种分离株 ctbb022，全基因组
MH648907.1	指环病毒科物种分离株 ctcf040，全基因组
MH648911.1	指环病毒科物种分离株 cthi018，全基因组
MH648912.1	指环病毒科物种分离株 ctea38，全基因组
MH648913.1	指环病毒科物种分离株 ctbg006，全基因组
MH648916.1	指环病毒科物种分离株 ctbg020，全基因组
MH648925.1	指环病毒科物种分离株 ctci019，全基因组
MH648932.1	指环病毒科物种分离株 ctid031，全基因组
MH648946.1	指环病毒科物种分离株 ctdb017，全基因组
MH648957.1	指环病毒科物种分离株 ctch017，全基因组
MH648958.1	指环病毒科物种分离株 ctbh011，全基因组
MH648959.1	指环病毒科物种分离株 ctbc020，全基因组
MH648962.1	指环病毒科物种分离株 ctif015，全基因组
MH648966.1	指环病毒科物种分离株 ctei055，全基因组
MH648969.1	指环病毒科物种分离株 ctjg000，全基因组
MH648976.1	指环病毒科物种分离株 ctcj064，全基因组
MH648977.1	指环病毒科物种分离株 ctbj022，全基因组
MH648982.1	指环病毒科物种分离株 ctbf014，全基因组
MH648983.1	指环病毒科物种分离株 ctbd027，全基因组
MH648985.1	指环病毒科物种分离株 ctch016，全基因组
MH648986.1	指环病毒科物种分离株 ctbd020，全基因组
MH648989.1	指环病毒科物种分离株 ctga035，全基因组
MH648990.1	指环病毒科物种分离株 cthf001，全基因组
MH648995.1	指环病毒科物种分离株 ctbd067，全基因组
MH648997.1	指环病毒科物种分离株 ctce026，全基因组
MH648999.1	指环病毒科物种分离株 ctfb058，全基因组
MH649002.1	指环病毒科物种分离株 ctjj046，全基因组
MH649006.1	指环病毒科物种分离株 ctcf030，全基因组
MH649008.1	指环病毒科物种分离株 ctbg025，全基因组
MH649011.1	指环病毒科物种分离株 ctbh052，全基因组
MH649014.1	指环病毒科物种分离株 ctba003，全基因组
MH649017.1	指环病毒科物种分离株 ctbb016，全基因组
MH649022.1	指环病毒科物种分离株 ctch023，全基因组
MH649023.1	指环病毒科物种分离株 ctbd051，全基因组

[2096]

MH649028.1	指环病毒科物种分离株 ctbf9，全基因组
MH649038.1	指环病毒科物种分离株 ctbi030，全基因组
MH649039.1	指环病毒科物种分离株 ctca057，全基因组
MH649040.1	指环病毒科物种分离株 ctch033，全基因组
MH649042.1	指环病毒科物种分离株 ctjd005，全基因组
MH649045.1	指环病毒科物种分离株 ctdc021，全基因组
MH649051.1	指环病毒科物种分离株 ctdg044，全基因组
MH649056.1	指环病毒科物种分离株 ctcc062，全基因组
MH649061.1	指环病毒科物种分离株 ctid009，全基因组
MH649062.1	指环病毒科物种分离株 ctdc018，全基因组
MH649063.1	指环病毒科物种分离株 ctbf012，全基因组
MH649068.1	指环病毒科物种分离株 ctcc066，全基因组
MH649070.1	指环病毒科物种分离株 ctda011，全基因组
MH649077.1	指环病毒科物种分离株 ctbh034，全基因组
MH649083.1	指环病毒科物种分离株 ctdg028，全基因组
MH649084.1	指环病毒科物种分离株 ctii061，全基因组
MH649085.1	指环病毒科物种分离株 cteh021，全基因组
MH649092.1	指环病毒科物种分离株 ctbg012，全基因组
MH649101.1	指环病毒科物种分离株 ctif053，全基因组
MH649104.1	指环病毒科物种分离株 ctei657，全基因组
MH649106.1	指环病毒科物种分离株 ctca015，全基因组
MH649114.1	指环病毒科物种分离株 ctbf050，全基因组
MH649122.1	指环病毒科物种分离株 ctdc002，全基因组
MH649125.1	指环病毒科物种分离株 ctbb15，全基因组
MH649127.1	指环病毒科物种分离株 ctba013，全基因组
MH649137.1	指环病毒科物种分离株 ctbb000，全基因组
MH649141.1	指环病毒科物种分离株 ctbc019，全基因组
MH649142.1	指环病毒科物种分离株 ctid026，全基因组
MH649144.1	指环病毒科物种分离株 ctfj004，全基因组
MH649152.1	指环病毒科物种分离株 ctcj13，全基因组
MH649156.1	指环病毒科物种分离株 ctci006，全基因组
MH649157.1	指环病毒科物种分离株 ctbd025，全基因组
MH649158.1	指环病毒科物种分离株 ctbf005，全基因组
MH649161.1	指环病毒科物种分离株 ctcf045，全基因组

[2097]

MH649165.1	指环病毒科物种分离株 ctcc29，全基因组
MH649169.1	指环病毒科物种分离株 ctib021，全基因组
MH649172.1	指环病毒科物种分离株 ctbh857，全基因组
MH649174.1	指环病毒科物种分离株 ctbj049，全基因组
MH649178.1	指环病毒科物种分离株 ctfc006，全基因组
MH649179.1	指环病毒科物种分离株 ctbe000，全基因组
MH649183.1	指环病毒科物种分离株 ctbb031，全基因组
MH649186.1	指环病毒科物种分离株 ctcb33，全基因组
MH649189.1	指环病毒科物种分离株 ctcc12，全基因组
MH649196.1	指环病毒科物种分离株 ctci060，全基因组
MH649199.1	指环病毒科物种分离株 ctbb017，全基因组
MH649203.1	指环病毒科物种分离株 cthc018，全基因组
MH649204.1	指环病毒科物种分离株 ctbj003，全基因组
MH649206.1	指环病毒科物种分离株 ctbg010，全基因组
MH649208.1	指环病毒科物种分离株 ctid008，全基因组
MH649209.1	指环病毒科物种分离株 ctbg056，全基因组
MH649210.1	指环病毒科物种分离株 ctda001，全基因组
MH649212.1	指环病毒科物种分离株 ctcf004，全基因组
MH649217.1	指环病毒科物种分离株 ctbe029，全基因组
MH649223.1	指环病毒科物种分离株 ctci016，全基因组
MH649224.1	指环病毒科物种分离株 ctce11，全基因组
MH649228.1	指环病毒科物种分离株 ctcf013，全基因组
MH649229.1	指环病毒科物种分离株 ctcb036，全基因组
MH649241.1	指环病毒科物种分离株 ctda027，全基因组
MH649242.1	指环病毒科物种分离株 ctbf003，全基因组
MH649254.1	指环病毒科物种分离株 ctjb007，全基因组
MH649255.1	指环病毒科物种分离株 ctbb023，全基因组
MH649256.1	指环病毒科物种分离株 ctca002，全基因组
MH649258.1	指环病毒科物种分离株 ctcg010，全基因组
MH649263.1	指环病毒科物种分离株 ctgh3，全基因组
MK012439.1	指环病毒科物种分离株 cthe000，全基因组
MK012440.1	指环病毒科物种分离株 ctjd008，全基因组
MK012448.1	指环病毒科物种分离株 ctch012，全基因组
MK012457.1	指环病毒科物种分离株 ctda009，全基因组

[2098]

MK012458.1	指环病毒科物种分离株 ctcd015，全基因组
MK012485.1	指环病毒科物种分离株 ctfd011，全基因组
MK012489.1	指环病毒科物种分离株 ctba003，全基因组
MK012492.1	指环病毒科物种分离株 ctbb005，全基因组
MK012493.1	指环病毒科物种分离株 ctcj014，全基因组
MK012500.1	指环病毒科物种分离株 ctcb001，全基因组
MK012504.1	指环病毒科物种分离株 ctcj010，全基因组
MK012516.1	指环病毒科物种分离株 ctcf003，全基因组
NC_038336.1	细环病毒 5 分离株 TCHN-C1 Orf2 和 Orf1 基因，全 cds
NC_038338.1	细环病毒 11 分离株 TCHN-D1 Orf2 和 Orf1 基因，全 cds
NC_038339.1	细环病毒 13 分离株 TCHN-A Orf2 和 Orf1 基因，全 cds
NC_038340.1	细环病毒 20 ORF4、ORF3、ORF2、ORF1 基因，全 cds，克隆：SAa-10
NC_038341.1	细环病毒 21 分离株 TCHN-B ORF2 和 ORF1 基因，全 cds
NC_038342.1	细环病毒 23 ORF2，ORF1 基因，全 cds，分离株：s-TTV CH65-2
NC_038343.1	细环病毒 24 ORF4、ORF3、ORF2、ORF1 基因，全 cds，克隆：SAa-01
NC_038344.1	细环病毒 29 ORF2、ORF1、ORF3 基因，全 cds，分离株：TTVyon-KC009
NC_038345.1	细环微小病毒 10 分离株 LIL-y1 ORF2、ORF1、ORF3 和 ORF4 基因，全 cds
NC_038346.1	细环微小病毒 11 分离株 LIL-y2 ORF2、ORF1 和 ORF3 基因，全 cds
NC_038347.1	细环微小病毒 12 分离株 LIL-y3 ORF2、ORF1、ORF3 和 ORF4 基因，全 cds
NC_038350.1	细环中型病毒 3 分离株 2PoSMA ORF2 和 ORF1 基因，全 cds
NC_038351.1	细环中型病毒 4 分离株 6PoSMA ORF2、ORF1 和 ORF3 基因，全 cds
NC_038352.1	细环中型病毒 5 DNA，全基因组，分离株：MDJHem2
NC_038353.1	细环中型病毒 6 DNA，全基因组，分离株：MDJHem3-1
NC_038354.1	细环中型病毒 7 DNA，全基因组，分离株：MDJHem3-2
NC_038355.1	细环中型病毒 8 DNA，全基因组，分离株：MDJN1
NC_038356.1	细环中型病毒 9 DNA，全基因组，分离株：MDJN2
NC_038357.1	细环中型病毒 10 DNA，全基因组，分离株：MDJN14

[2099]

NC_038358.1	细环中型病毒 11 DNA，全基因组，分离株：MDJN47
NC_038359.1	细环中型病毒 12 DNA，全基因组，分离株：MDJN51
NC_038360.1	细环中型病毒 13 DNA，全基因组，分离株：MDJN69
NC_038361.1	细环中型病毒 14 DNA，全基因组，分离株：MDJN97
NC_038362.1	细环中型病毒 15 DNA，全基因组，分离株：Pt-TTMDV210

[2100] 在一些实施例中，遗传元件包含与来自一种或多种非指环病毒，例如腺病毒、疱疹病毒、pox病毒、痘苗病毒、SV40、乳头瘤病毒、RNA病毒(例如逆转录病毒、例如慢病毒)、单链RNA病毒(例如肝炎病毒)或双链RNA病毒(例如轮状病毒)的一个或多个序列具有同源性或同一性的一个或多个序列。由于在一些实施例中，重组逆转录病毒是有缺陷的，因此可以提供协助以产生感染性颗粒。可以例如通过使用辅助细胞系来提供这种协助，所述辅助细胞系包含在LTR内的调节性序列的控制下编码逆转录病毒的所有结构基因的质粒。用于复制本文所述的指环体的合适细胞系包括本领域已知的细胞系，例如A549细胞，其可以如本文所述进行修饰。所述遗传元件可以另外包含编码可选择标记的基因，从而可以鉴定期望的遗传元件。

[2101] 在一些实施例中，遗传元件包括非沉默突变，例如导致编码的多肽中氨基酸差异的碱基置换、缺失或添加，只要序列与由第一核苷酸序列编码的多肽保持至少约70％、75％、80％、85％、90％、95％、96％、97％、98％或99％同一性或以其他方式可用于实施本发明。在这方面，可以进行某些保守的氨基酸置换，这些置换通常被认为不会使蛋白质的整体功能失活：例如对于带正电的氨基酸(反之亦然)，赖氨酸、精氨酸和组氨酸；对于带负电的氨基酸(反之亦然)，天冬氨酸和谷氨酸；对于某些电中性氨基酸的组(在所有情况下，反之亦然)，(1)丙氨酸和丝氨酸，(2)天冬酰胺、谷氨酰胺和组氨酸，(3)半胱氨酸和丝氨酸，(4)甘氨酸和脯氨酸，(5)异亮氨酸、亮氨酸和缬氨酸，(6)蛋氨酸、亮氨酸和异亮氨酸，(7)苯丙氨酸、蛋氨酸、亮氨酸和酪氨酸，(8)丝氨酸和苏氨酸，(9)色氨酸和酪氨酸，(10)以及例如酪氨酸、色氨酸和苯丙氨酸。氨基酸可以根据物理性质以及对二级和三级蛋白质结构的贡献进行分类。保守置换在本领域中被认为是一个氨基酸被具有相似特性的另一氨基酸置换。

[2102] 具有相同或指定百分比的相同核苷酸或氨基酸残基的两个或更多个核酸或多肽序列的同一性(例如，当在比较窗口或指定区域内进行最大对应性的比较和对齐时，在具体区域内约60％、65％、70％、75％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或更高同一性)可使用具有以下所述默认参数的BLAST或BLAST 2.0序列比较算法，或通过手动比对和目视检查(例如，参见NCBI网站www.ncbi.nlm.nih.gov/BLAST/等)进行测量。同一性也可以指或可以用于对测试序列的补充。同一性还包括具有缺失和/或添加的序列以及具有置换的序列。如本文所述，算法考虑了缺口等。同一性可以在长度为至少约10个氨基酸或核苷酸、长度为约15个氨基酸或核苷酸、长度为约20个氨基酸或核苷酸、长度为约25个氨基酸或核苷酸、长度为约30个氨基酸或核苷酸、长度为约35个氨基酸或核苷酸、长度为约40个氨基酸或核苷酸、长度为约45个氨基酸或核苷酸、长度为约50个氨基酸或核苷酸、或更多的区域中存在。

[2103] 在一些实施例中，遗传元件包含与本文所述例如如表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15、17或41中的任一个中所列的核苷酸序列中任一个具有至少约

75％核苷酸序列同一性，至少约80％、85％、90％、95％、96％、97％、98％、99％或100％核苷酸序列同一性的核苷酸序列。由于遗传密码是简并的，因此同源核苷酸序列可包括任何数目的“沉默”碱基改变，即仍然编码相同氨基酸的核苷酸置换。

[2104] 基因编辑组分

[2105] 指环体的遗传元件可以包括一个或多个编码基因编辑系统的组分的基因。示例性的基因编辑系统包括成簇的调节性间隔短回文重复序列(CRISPR)系统、锌指核酸酶(ZFN)和基于转录激活子样效应子的核酸酶(TALEN)。基于ZFN、TALEN和CRISPR的方法已在例如在Gaj等人Trends Biotechnol.[生物技术动向]31.7(2013):397-405中描述；CRISPR基因编辑方法例如在Guan等人，Application of CRISPR-Cas system in gene therapy:Pre-clinical progress in animal model.[CRISPR-Cas系统在基因治疗中的应用:动物模型中的临床前进展]DNARepair[DNA修复]2016年10月；46:1-8.doi:10.1016/J.dnarep.2016.07.004；Zheng等人，Precise gene deletion and replacement using the CRISPR/Cas9 system in human cells[使用CRISPR/Cas9系统在人细胞中进行精确基因缺失和替换].BioTechniques[生物技术]，第57卷，第3号，2014年9月，第115-124页中描述。

[2106] CRISPR系统是最初在细菌和古细菌中发现的自适应防御系统。CRISPR系统使用称为CRISPR相关或“Cas”内切核酸酶的RNA指导的核酸酶(例如，Cas9或Cpf1)来切割外源DNA。在典型的CRISPR/Cas系统中，内切核酸酶通过靶向单链或双链DNA序列的序列特异性的非编码“指导RNA”定向到靶核苷酸序列(例如，基因组中待序列编辑的位点)。已经鉴定了三类(I-III)CRISPR系统。II类CRISPR系统使用单个Cas内切核酸酶(而不是多个Cas蛋白)。一种II类CRISPR系统包括II类Cas内切核酸酶，例如Cas9、CRISPR RNA(“crRNA”)和反式激活crRNA(“tracrRNA”)。crRNA包含“指导RNA”，即通常对应于靶DNA序列的约20个核苷酸RNA序列。crRNA还包含与tracrRNA结合的区域，以形成被RNase III切割的部分双链结构，产生crRNA/tracrRNA杂交体。然后，crRNA/tracrRNA杂交体指导Cas9内切核酸酶识别并切割靶DNA序列。靶DNA序列必须总体上邻近针对给定Cas内切核酸酶来说是特异性的“原间隔子邻近基序”(“PAM”)；然而，PAM序列似乎遍布整个给定基因组。

[2107] 在一些实施例中，指环体包括CRISPR内切核酸酶的基因。例如，从不同原核物种鉴定的一些CRISPR内切核酸酶具有独特的PAM序列要求；PAM序列的实例包括5’-NGG(酿脓链球菌(Streptococcus pyogenes))、5’-NNAGAA(嗜热链球菌(Streptococcus thermophilus)CRISPR1)、5’-NGGNG(嗜热链球菌CRISPR3)、和5’-NNNGATT(奈瑟氏脑膜炎双球菌(Neisseria meningiditis))。一些内切核酸酶例如Cas9内切核酸酶与富含G的PAM位点例如5’-NGG有关，并在距PAM位点上游(5’)3个核苷酸处对靶DNA进行平末端切割。另一个II类CRISPR系统包括小于Cas9的V型内切核酸酶Cpf1；实例包括AsCpf1(来自氨基酸球菌属物种(Acidaminococcus sp.))和LbCpf1(来自毛螺旋菌属物种(Lachnospiraceae sp.))。Cpf1内切核酸酶与富含T的PAM位点例如5’-TTN相关。Cpf1也可以识别5’-CTAPAM基序。Cpf1通过引入具有4或5个核苷酸的5’突出端的错位或交错的双链断裂来切割靶DNA，例如切割如下靶DNA，该靶DNA中的5个核苷酸的错位或交错的切割位于距离编码链上的PAM位点下游(3’)18个核苷酸的位置处和距离互补链上的PAM位点下游23个核苷酸的位置处；由这种错位切割产生的5个核苷酸突出端使得通过同源重组的DNA插入比在平末端切割的DNA的插入

更精确地进行基因组编辑。参见例如,Zetsche等人(2015)Cell[细胞],163:759-771。

[2108] 指环体中可以包括多种CRISPR相关(Cas)基因。基因的具体实例是那些编码来自II类系统的Cas蛋白(包括Cas1、Cas2、Cas3、Cas4、Cas5、Cas6、Cas7、Cas8、Cas9、Cas10、Cpf1、C2C1或C2C3)的基因。在一些实施例中,指环体包括编码Cas蛋白例如Cas9蛋白的基因,所述Cas蛋白可以来自多种原核物种中的任一种。在一些实施例中,指环体包括编码特定Cas蛋白例如特定Cas9蛋白的基因,所述Cas蛋白被选择以识别特定的原间隔子邻近基序(PAM)序列。在一些实施例中,指环体包括编码两种或更多种不同Cas蛋白或两种或更多种Cas蛋白的核酸,所述Cas蛋白可以被引入细胞、合子、胚胎或动物中,例如,以允许识别和修饰包含相同、相似或不同PAM基序的位点。在一些实施例中,指环体包括编码具有失活的核酸酶的修饰的Cas蛋白(例如,核酸酶缺陷的Cas9)的基因。

[2109] 野生型Cas9蛋白在被gRNA靶向的特定DNA序列上产生双链断裂(DSB),而许多具有经修饰的功能的CRISPR内切核酸酶是已知的,例如:Cas9的“切口酶”版本仅产生单链断裂;不具有催化活性的Cas9(“dCas9”)不会切割靶DNA。可以将编码dCas9的基因与编码效应子结构域的基因融合,以抑制(CRISPRi)或激活(CRISPRa)靶基因的表达。例如,所述基因可以编码Cas9与转录沉默子(例如KRAB结构域)或转录激活子(例如dCas9-VP64融合体)的融合体。可以包括编码与FokI核酸酶(“dCas9-FokI”)融合的不具有催化活性的Cas9(dCas9)的基因,以在与两个gRNA同源的靶序列处产生DSB。参见例如,许多CRISPR/Cas9质粒披露于并且可公开获得自阿德基因质粒库(Addgene repository)(阿德基因公司(Addgene),西德尼大街75号(Sidney St.),550A单元,剑桥郡,马萨诸塞州02139;addgene.org/crispr/)。Ran等人(2013)Cell[细胞],154:1380-1389将引入两个独立的双链断裂(每个断裂被独立的指导RNA指向)的“双切口酶”Cas9描述为实现了更准确的基因组编辑。

[2110] 在美国专利申请公开2016/0138008A1和US 2015/0344912 A1以及美国专利8,697,359、8,771,945、8,945,839、8,999,641、8,993,233、8,895,308、8,865,406、8,889,418、8,871,445、8,889,356、8,932,814、8,795,965和8,906,616中披露了用于编辑真核生物的基因的CRISPR技术。Cpf1内切核酸酶和相应的指导RNA和PAM位点在美国专利申请公开2016/0208243A1中披露。

[2111] 在一些实施例中,指环体包含编码本文所述多肽(例如靶向核酸酶,例如Cas9,例如野生型Cas9、切口酶Cas9(例如Cas9 D10A)、失活Cas9(dCas9)、eSpCas9、Cpf1、C2C1或C2C3)以及gRNA的基因。编码核酸酶和一个或多个gRNA的基因的选择取决于靶向突变是核苷酸的缺失、置换或添加,例如靶序列的核苷酸缺失、置换或添加。编码与(一个或多个)效应子结构域(例如VP64)的全部或部分(例如,其生物活性部分)连接在一起的不具有催化活性的内切核酸酶(例如,失活Cas9(dCas9,例如D10A;H840A))的基因产生可以调节一个或多个靶核酸序列的活性和/或表达的嵌合蛋白。

[2112] 如本文所用,“效应子结构域的生物活性部分”是维持效应子结构域(例如,“最小”或“核心”结构域)的功能(例如,完全、部分或最小功能)的部分。在一些实施例中,指环体包括编码dCas9与一个或多个效应子结构域的全部或一部分的融合体的基因,以产生可用于本文所述方法的嵌合蛋白。因此,在一些实施例中,指环体包括编码dCas9-甲基化酶融合体的基因。在其他一些实施例中,指环体包括编码dCas9酶与位点特异性gRNA的融合体的基因,以靶向内源基因。

[2113] 在其他方面，指环体包括编码1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、或更多个与dCas9融合的效应子结构域(全部或生物学活性部分)的基因。

[2114] 蛋白质外部

[2115] 在一些实施例中，指环体例如合成的指环体包含封闭遗传元件的蛋白质外部。蛋白质外部可以包含不能引起不想要的哺乳动物免疫应答的基本上非致病性的外部蛋白质。指环体的蛋白质外部通常包含可自组装成构成蛋白质外部的二十面体结构的基本上非致病的蛋白质。

[2116] 在一些实施例中，蛋白质外部蛋白由指环体的遗传元件的序列编码(例如，与遗传元件顺式)。在其他实施例中，蛋白质外部蛋白由与指环体的遗传元件分离的核酸编码(例如，与遗传元件反式)。

[2117] 在一些实施例中，蛋白质，例如基本上非致病性的蛋白质和/或蛋白质外部蛋白质，包含一个或多个糖基化氨基酸，例如2、3、4、5、6、7、8、9、10或更多个。

[2118] 在一些实施例中，蛋白质，例如基本上非致病性的蛋白质和/或蛋白质外部蛋白质，包含至少一个亲水性DNA结合区、精氨酸富集区、苏氨酸富集区、谷氨酰胺富集区、N-末端聚精氨酸序列、可变区、C-末端聚谷氨酰胺/谷氨酸序列和一个或多个二硫键。

[2119] 在一些实施例中，蛋白质是衣壳蛋白，例如具有与由编码本文所述衣壳蛋白的任一核苷酸序列编码的蛋白(例如表1-18，A1-A12、B1-B5、C1-C5、D1-D10或20-37中的任一个中所列的指环病毒ORF1序列或衣壳蛋白序列)具有至少约60％、65％、70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的序列。在一些实施例中，蛋白质或衣壳蛋白的功能片段由核苷酸序列编码，所述核苷酸序列与本文所述的任一核苷酸序列具有至少约60％、70％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性，例如表A1-A12、B1-B5、C1-C5、D1-D10或20-37中的任一个中所列的指环病毒衣壳序列或衣壳蛋白序列。在一些实施例中，蛋白包含衣壳蛋白或衣壳蛋白的功能片段，其由衣壳核苷酸序列或与本文所述的核苷酸序列中任一个具有至少约60％、65％、70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％核苷酸序列同一性的序列编码，例如表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中的任一个中所列的指环病毒衣壳序列或衣壳蛋白序列。

[2120] 在一些实施例中，指环体包含编码以下的核苷酸序列：衣壳蛋白或衣壳蛋白的功能片段或与本文所述的氨基酸序列中任一个(例如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16或18中的任一个中所列的指环病毒衣壳序列或衣壳蛋白序列)具有至少约60％、70％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的序列。在一些实施例中，指环体包含编码以下的核苷酸序列：衣壳蛋白或衣壳蛋白的功能片段或与本文所述的氨基酸序列中任一个(例如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16或18中的任一个中所列的指环病毒衣壳序列或衣壳蛋白序列)具有至少约60％、65％、70％、75％、80％、85％、90％、95％、96％、97％、98％、99％或100％序列同一性的序列。

[2121] 在一些实施例中，指环体包含核苷酸序列，所述核苷酸序列编码具有本文所述的氨基酸序列(指环病毒氨基酸序列，例如，如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16或18中的任一个中所列或如图1所示)或其功能片段的约1位至约150位(例如，或每个范围内的氨基酸的任何子集，例如约20位至约35位、约25位至约30位、约26位至约30)、

约150位至约390位(例如,或每个范围内的氨基酸的任何子集,例如约200位至约380位、约205位至约375位、约205位至约371)、约390至约525位、约525位至约850位(例如,或每个范围内的氨基酸的任何子集,例如约530位至约840位、约545位至约830位、约550位至约820)、约850至约950位(例如,或每个范围内的氨基酸的任何子集,例如约860位至约940位、约870位至约930位、约880位至约923)的氨基酸序列。在一些实施例中,蛋白质包含氨基酸序列或其功能片段或与本文所述的氨基酸序列(指环病毒氨基酸序列,例如,如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16或18中的任一个中所列或如图1所示)的约1位至约150位(例如,或本文所述的每个范围内的氨基酸的任何子集)、约150位至约390位、约390位至约525位、约525位至约850位、约850位至约950位具有至少约60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的序列。

[2122] 在一些实施例中,蛋白包含氨基酸序列或其功能片段或与本文所述的氨基酸序列或氨基酸范围中任一个(指环病毒氨基酸序列,例如,表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16或18中任一个所列或如图1所示)具有至少约60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的序列。在一些实施例中,具有较少序列同一性的氨基酸范围可以提供本文所述的一种或多种性质以及细胞/组织/物种特异性(例如嗜性)的差异。

[2123] 在一些实施例中,指环体在蛋白质外部缺少脂质。在一些实施例中,指环体缺少脂质双层,例如病毒包膜。在一些实施例中,指环体的内部被蛋白质外部完全覆盖(例如100%覆盖)。在一些实施例中,指环体内部被蛋白质外部覆盖少于100%,例如95%、90%、85%、80%、70%、60%、50%或更少的覆盖。在一些实施例中,蛋白质外部包括间隙或不连续,例如,使得可透过水、离子、肽或小分子,只要遗传元件保留在指环体中即可。

[2124] 在一些实施例中,蛋白质外部包含特异性识别和/或结合宿主细胞例如互补蛋白或多肽以介导遗传元件进入宿主细胞的一种或多种蛋白质或多肽。

[2125] 在一些实施例中,蛋白质外部包含以下一种或多种:一个或多个糖基化蛋白、亲水性DNA结合区、精氨酸富集区、苏氨酸富集区、谷氨酰胺富集区、N末端聚精氨酸序列、可变区、C末端聚谷氨酰胺/谷氨酸序列、以及一个或多个二硫键。例如,蛋白质外部包含由本文所述的指环病毒ORF1编码的蛋白质。

[2126] 在一些实施例中,蛋白质外部包括以下一种或多种特征:二十面体对称、识别和/或结合与一个或多个宿主细胞分子相互作用以介导进入宿主细胞的分子、缺乏脂质分子、缺乏碳水化合物、pH和温度稳定、耐洗涤剂、以及在宿主中基本上是非免疫原性的或是非致病性的。

[2127] II.载体

[2128] 本文所述的遗传元件可以包含在载体中。合适的载体及其制造和使用方法在现有技术中是众所周知的。

[2129] 一方面,本发明包括一种载体,所述载体包含遗传元件,所述遗传元件包含(i)编码非致病性外部蛋白质的序列,(ii)使所述遗传元件与所述非致病性外部蛋白质结合的外部蛋白质结合序列,和(iii)编码调节性核酸的序列。

[2130] 遗传元件或遗传元件内的任何序列可以使用任何合适的方法获得。多种重组方法在本领域中是已知的,例如,使用标准技术,从具有病毒序列的细胞中筛选文库,从已知包

含序列的载体中获得所述序列，或直接从包含序列的细胞和组织中分离所述序列。可替代地或组合地，遗传元件的部分或全部可以合成产生，而不是克隆。

[2131] 在一些实施例中，载体包括调节性元件，与靶基因同源的核酸序列，以及用于在活细胞内和/或当靶细胞内存在细胞内分子时引起报道分子表达的各种报道构建体。

[2132] 报告基因用于鉴定潜在转染的细胞和评估调节性序列的功能。通常，报告基因是如下基因，所述基因不存在于受体生物体或组织中或由其表达并且编码多肽，所述多肽的表达通过一些可易于检测的特性(例如，酶活性)表现出来。在将DNA引入受体细胞后，在适当的时间测定报告基因的表达。合适的报告基因可包括编码萤光素酶、β-半乳糖苷酶、氯霉素乙酰转移酶、分泌的碱性磷酸酶、或绿色荧光蛋白基因的基因(例如，Ui-Tei等人，2000FEBS Letters[欧洲生化学会联合会快报]479:79-82)。合适的表达系统是熟知的，可以使用已知技术制备或商业获得。通常，具有显示报告基因最高表达水平的最小5'侧翼区的构建体被鉴定为启动子。此类启动子区可以连接至报告基因，并且用于评估药剂调节启动子驱动的转录的能力。

[2133] 在一些实施例中，载体在宿主细胞中是基本上非致病性的和/或基本上非整合的，或者在宿主中是基本上非免疫原性的。

[2134] 在一些实施例中，载体的量足以调节表型、病毒水平、基因表达、与其他病毒竞争、疾病状态等中的一种或多种至少约5％、10％、15％、20％、25％、30％、35％、40％、45％、50％或更高。

[2135] 用于量化指环体活性的模型系统

[2136] 可以在许多动物模型和体外模型中测定本文所述的指环体。例如，脊髓性肌萎缩症的小鼠模型可用于量化编码SMN或其功能变体的指环体的活性。在一些实施例中，肌营养不良的小鼠模型可用于量化编码微肌营养不良蛋白或其功能变体的指环体的活性。在一些实施例中，UGT1A1缺陷型小鼠模型可用于量化编码UGT1A1或其功能变体的指环体的活性。

[2137] III.组合物

[2138] 本文所述的指环体或载体也可以包括在具有例如本文所述的药物赋形剂的药物组合物中。在一些实施例中，药物组合物包含至少10^5、10^6、10^7、10^8、10^9、10^{10}、10^{11}、10^{12}、10^{13}、10^{14}或10^{15}个指环体。在一些实施例中，药物组合物包含约10^5-10^{15}、10^5-10^{10}、或10^{10}-10^{15}个指环体。在一些实施例中，药物组合物包含约10^8(例如约10^5、10^6、10^7、10^8、10^9、或10^{10})个基因组当量/mL的指环体。在一些实施例中，药物组合物包含10^5-10^{10}、10^6-10^{10}、10^7-10^{10}、10^8-10^{10}、10^9-10^{10}、10^5-10^6、10^5-10^7、10^5-10^8、10^5-10^9、10^5-10^{11}、10^5-10^{12}、10^5-10^{13}、10^5-10^{14}、10^5-10^{15}或10^{10}-10^{15}个基因组当量/mL的指环体，例如，根据实例18的方法所测定。在一些实施例中，药物组合物包含足够的指环体以将指环体中包含的遗传元件的至少1、2、5或10、100、500、1000、2000、5000、8,000、$1x10^4$、$1x10^5$、$1x10^6$、$1x10^7$个或更大数量的拷贝/细胞递送到真核细胞群体中。在一些实施例中，药物组合物包含足够的指环体以将指环体中包含的遗传元件的至少约$1x10^4$、$1x10^5$、$1x10^6$、$1x10^7$、或约$1x10^4$-$1x10^5$、$1x10^4$-$1x10^6$、$1x10^4$-$1x10^7$、$1x10^5$-$1x10^6$、$1x10^5$-$1x10^7$、或$1x10^6$-$1x10^7$个拷贝/细胞递送到真核细胞群体中。

[2139] 在一些实施例中，药物组合物具有以下一种或多种特征：所述药物组合物满足药物或良好生产规范(GMP)标准；所述药物组合物是根据良好生产规范(GMP)制成的；所述药物组合物具有低于预定参考值的病原体水平，例如基本上不含病原体；所述药物组合物具

有低于预定参考值的污染物水平,例如,基本上不含污染物;或所述药物组合物具有低免疫原性或基本上是非免疫原性的,例如,如本文所述。

[2140] 在一些实施例中,药物组合物包含低于阈值量的一种或多种污染物。在药物组合物中理想地不包括或最少包括的示例性污染物包括但不限于宿主细胞核酸(例如,宿主细胞DNA和/或宿主细胞RNA)、动物来源的组分(例如,血清白蛋白或胰蛋白酶)、具有复制能力的病毒、非感染性颗粒、游离病毒衣壳蛋白、外源因子和聚集体。在实施例中,污染物是宿主细胞DNA。在实施例中,组合物包含少于约10ng的宿主细胞DNA/剂量。在实施例中,组合物中宿主细胞DNA的水平通过宿主细胞DNA的过滤和/或酶促降解而降低。在实施例中,药物组合物由按重量计小于10%(例如小于约10%、5%、4%、3%、2%、1%、0.5%或0.1%)的污染物组成。

[2141] 一方面,本文所述的发明包括药物组合物,其包含:

[2142] a)指环体,其包含遗传元件,所述遗传元件包含(i)编码非致病性外部蛋白质的序列,(ii)使所述遗传元件与所述非致病性外部蛋白质结合的外部蛋白质结合序列,和(iii)编码调节性核酸的序列;和与所述遗传元件相关联(例如包封或封闭所述遗传元件)的蛋白质外部;以及

[2143] b)药物赋形剂。

[2144] 囊泡

[2145] 在一些实施例中,组合物进一步包含载剂组分,例如微粒、脂质体、囊泡或外来体。在一些实施例中,脂质体包含球形囊泡结构,所述球形囊泡结构由围绕内部水性隔室的单层或多层的脂质双层和相对不可渗透的外部亲脂性磷脂双层构成。脂质体可以是阴离子的、中性的或阳离子的。脂质体通常具有生物相容性,无毒,可以递送亲水性和亲脂性药物分子,保护其货物免受血浆酶的降解,并将其负载运输穿过生物膜(关于综述,参见,例如,Spuch和Navarro,Journal of Drug Delivery[药物递送杂志],第2011卷,文章ID 469679,第12页,2011.doi:10.1155/2011/469679)。

[2146] 囊泡可由几种不同类型的脂质制成;然而,磷脂最常用于产生脂质体作为药物载剂。囊泡可包括但不限于DOTMA、DOTAP、DOTIM、DDAB,单独使用或与胆固醇一起产生DOTMA和胆固醇、DOTAP和胆固醇、DOTIM和胆固醇以及DDAB和胆固醇。制备多层囊泡脂质的方法是本领域已知的(参见例如美国专利号6,693,086,其关于多层囊泡脂质制备的教导通过引用并入本文)。尽管当脂质膜与水溶液混合时,囊泡的形成是自发的,但也可以通过使用均质器、超声仪或挤压装置以振荡的形式施加力来加快囊泡的形成(关于综述,参见例如,Spuch和Navarro,Journal of Drug Delivery[药物递送杂志],第2011卷,文章ID 469679,第12页,2011.doi:10.1155/2011/469679)。可通过挤出通过具有减小尺寸的过滤器来制备挤出的脂质,如Templeton等人,Nature Biotech[自然生物技术],15:647-652,1997中所述,其关于挤出脂质制备的教导通过引用并入文中。

[2147] 如本文所述,可将添加剂添加至囊泡以改变其结构和/或性质。例如,可以将胆固醇或鞘磷脂加入混合物中,以帮助稳定结构并防止内部货物泄漏。此外,可以由氢化的卵磷脂酰胆碱或卵磷脂酰胆碱、胆固醇和磷酸二鲸蜡酯制备囊泡。(关于综述,参见,例如,Spuch和Navarro,Journal of Drug Delivery[药物递送杂志],第2011卷,文章ID 469679,第12页,2011.doi:10.1155/2011/469679)。同样,囊泡可以在合成期间或之后进行表面修饰以

包括与受体细胞上的反应性基团互补的反应性基团。这样的反应性基团包括但不限于马来酰亚胺基团。例如，可以合成囊泡以包括马来酰亚胺缀合的磷脂，例如但不限于DSPE-MaL-PEG2000。

[2148] 囊泡制剂可以主要由天然磷脂和脂质(例如1,2-二硬脂酰基-sn-甘油-3-磷脂酰胆碱(DSPC)、鞘磷脂、卵磷脂酰胆碱和单唾液酸神经节苷脂)构成。仅由磷脂组成的制剂在血浆中不太稳定。然而，用胆固醇操纵脂质膜降低了包封的货物的快速释放，或者1,2-二油酰基-sn-甘油-3-磷酸乙醇胺(DOPE)增加了稳定性(关于综述，参见，例如，Spuch和Navarro,Journal of Drug Delivery[药物递送杂志],第2011卷,文章ID 469679,第12页,2011.doi:10.1155/2011/469679)。

[2149] 在实施例中，脂质可用于形成脂质微粒。脂质包括但不限于DLin-KC2-DMA4、C12-200和可使用自发囊泡形成程序配制的共脂质二硬脂酰磷脂酰胆碱、胆固醇和PEG-DMG(参见例如Novobrantseva,Molecular Therapy-Nucleic Acids[分子治疗-核酸](2012)1,e4;doi:10.1038/mtna.2011.3)。组分摩尔比可以为约50/10/38.5/1.5(DLin-KC2-DMA或C12-200/二硬脂酰磷脂酰胆碱/胆固醇/PEG-DMG)。Tekmira在美国和国外拥有约95个专利同族组合，涉及脂质微粒和脂质微粒制剂的各个方面(参见，例如，美国专利7,982,027;7,799,565;8,058,069;8,283,333;7,901,708;7,745,651;7,803,397;8,101,741;8,188,263;7,915,399;8,236,943和7,838,658以及欧洲专利号1766035;1519714;1781593和1664316)，所有这些都可以用于和/或适用于本发明。

[2150] 在一些实施例中，微粒包含以随机方式排列的一种或多种固化聚合物。微粒可以是可生物降解的。可生物降解的微粒可以例如使用本领域已知的方法合成，包括但不限于溶剂蒸发、热熔微囊化、溶剂去除和喷雾干燥。Bershteyn等人,Soft Matter[软物质]4:1787-1787,2008和US 2008/0014144A1中描述了用于合成微粒的示例性方法，其关于微粒合成的具体教导通过引用并入本文。

[2151] 可用于形成可生物降解的微粒的示例性合成聚合物包括但不限于脂肪族聚酯、聚(乳酸)(PLA)、聚(乙醇酸)(PGA)、乳酸和乙醇酸的共聚物(PLGA)、聚己内酯(PCL)、聚酐、聚(原)酸酯、聚氨酯、聚(丁酸)、聚(丙酸)和聚(丙交酯-己内酯)以及天然聚合物，例如白蛋白、藻酸盐和其他多糖，包括葡聚糖和纤维素、胶原蛋白、其化学衍生物，包括化学基团的取代、添加，例如烷基、亚烷基、羟基化、氧化和本领域技术人员常规进行的其他修饰)、白蛋白和其他亲水蛋白、玉米蛋白和其他醇溶蛋白和疏水性蛋白质、共聚物及其混合物。通常，这些材料通过酶水解或暴露于水、通过表面或整体腐蚀而降解。

[2152] 微粒的直径范围为0.1-1000微米(μm)。在一些实施例中，它们的直径的尺寸范围为1μm-750μm、或50μm-500μm、或100μm-250μm。在一些实施例中，它们的直径的尺寸范围为50μm-1000μm、50μm-750μm、50μm-500μm或50μm-250μm。在一些实施例中，它们的直径的尺寸范围是.05μm-1000μm、10μm-1000μm、100μm-1000μm或500μm-1000μm。在一些实施例中，它们的直径为约0.5μm、约10μm、约50μm、约100μm、约200μm、约300μm、约350μm、约400μm、约450μm、约500μm、约550μm、约600μm、约650μm、约700μm、约750μm、约800μm、约850μm、约900μm、约950μm或约1000μm。如微粒直径的上下文中所用，术语“约”是指所述绝对值的+/-5％。

[2153] 在一些实施例中，配体通过存在于颗粒表面上并存在于待连接的配体上的官能化学基团(羧酸、醛、胺、巯基和羟基)与微粒表面缀合。可以通过例如在微粒的乳液制备期间，

并入具有官能化学基团的稳定剂来将官能度引入微粒中。

[2154] 将官能团引入微粒的另一个实例是在微粒制备后,通过用均双官能或异双官能交联剂直接交联颗粒和配体。该程序可以使用合适的化学物质和一类交联剂(CDI、EDAC、戊二醛等,如下文更详细讨论)或通过制备后对颗粒表面进行化学修饰将配体偶联至颗粒表面的任何其他交联剂。这还包括这样的程序,通过所述程序可以将两亲性分子(例如脂肪酸、脂质或功能稳定剂)被动吸附并粘附到颗粒表面,从而引入官能性端基用以连接到配体上。

[2155] 在一些实施例中,可以合成微粒以在其外表面上包含一个或多个靶向基团以靶向特定的细胞或组织类型(例如,心肌细胞)。这些靶向基团包括但不限于受体、配体、抗体等。这些靶向基团将其配偶体结合在细胞表面上。在一些实施例中,微粒将整合到构成细胞表面的脂质双层中,并且将线粒体递送至细胞中。

[2156] 微粒还可在其最外表面上包含脂质双层。该双层可以由一个或多个相同或不同类型的脂质组成。实例包括但不限于磷脂,例如磷酸胆碱和磷酸肌醇。具体实例包括但不限于DMPC、DOPC、DSPC和各种其他脂质,例如本文针对脂质体所述的那些。

[2157] 在一些实施例中,载剂包含例如本文所述的纳米颗粒。

[2158] 在一些实施例中,本文所述的囊泡或微粒用诊断剂官能化。诊断剂的实例包括但不限于用于正电子发射断层扫描(PET)、计算机辅助断层扫描(CAT)、单光子发射计算机断层扫描、x射线、荧光检查和磁共振成像(MRI)的市售成像剂;和造影剂。在MRI中用作造影剂的合适材料的实例包括钆螯合物,以及铁、镁、锰、铜和铬。

[2159] 载剂

[2160] 本文所述的组合物(例如,药物组合物)可包含载剂、与载剂一起配制和/或在载剂中递送。在一个方面,本发明包括组合物,例如药物组合物,其包含载剂(例如囊泡、脂质体、脂质纳米颗粒、外泌体、红细胞、外泌体(例如哺乳动物或植物外泌体)),融合体),所述载剂包含(例如,包封)本文所述的组合物(例如,指环体、指环病毒、指环载体或本文所述的遗传元件)。

[2161] 在一些实施例中,本文所述的组合物和系统可以配制在脂质体或其他类似的囊泡中。通常,脂质体是球形囊泡结构,所述球形囊泡结构由围绕内部水性隔室的单层或多层的脂质双层和相对不可渗透的外部亲脂性磷脂双层构成。脂质体可以是阴离子的、中性的或阳离子的。脂质体通常具有以下特征中的一种或多种(例如,所有):生物相容性,无毒,可以递送亲水性和亲脂性药物分子,可以保护其货物免受血浆酶的降解,并可以将其负载运输穿过生物膜和血脑屏障(BBB)(参见,例如,Spuch和Navarro,Journal of Drug Delivery[药物递送杂志],第2011卷,文章ID 469679,第12页,2011.doi:10.1155/2011/469679;和Zylberberg&Matosevic.2016.Drug Delivery[药物递送],23:9,3319-3329,doi:10.1080/10717544.2016.1177136)。

[2162] 囊泡可由几种不同类型的脂质制成;然而,磷脂最常用于产生脂质体作为药物载剂。制备多层囊泡脂质的方法是已知的(参见例如美国专利号6,693,086,其关于多层囊泡脂质制备的教导通过引用并入文中)。尽管当脂质膜与水溶液混合时,囊泡的形成是自发的,但也可以通过使用均质器、超声仪或挤压装置以振荡的形式施加力来加快囊泡的形成(关于综述,参见,例如,Spuch和Navarro,Journal of Drug Delivery[药物递送杂志],第2011卷,文章ID 469679,第12页,2011.doi:10.1155/2011/469679)。如empleton等人,

Nature Biotech[自然生物技术],15:647-652,1997中所述,挤出的脂质可以通过例如通过用于减小尺寸的过滤器挤出来制备。

[2163] 脂质纳米颗粒(LNP)是为本文所述的药物组合物提供生物相容性和可生物降解的递送系统的载剂的另一个实例。参见,例如,Gordillo-Galeano等人European Journal of Pharmaceutics and Biopharmaceutics[欧洲药剂学和生物药剂学杂志].第133卷,2018年12月,第285-308页。纳米结构化的脂质载剂(NLC)是经修饰的固体脂质纳米颗粒(SLN),其保留了SLN的特性,提高了药物的稳定性和载药量,并防止了药物泄漏。聚合物纳米颗粒(PNP)是药物递送的重要组成部分。这些纳米颗粒可以有效地将药物递送引导至特定靶并改善药物稳定性和受控的药物释放。也可以使用脂质聚合物纳米颗粒(PLN),其是一种组合了脂质体和聚合物的新型载剂。这些纳米颗粒具有PNP和脂质体的互补优势。PLN由核-壳结构构成;聚合物核提供了稳定的结构,磷脂壳提供了良好的生物相容性。这样,这两种组分提高了药物包封有效率,促进了表面修饰,并防止了水溶性药物的泄漏。对于综述,参见例如Li等人2017,Nanomaterials[纳米材料]7,122;doi:10.3390/nano7060122。

[2164] 外来体也可用作本文所述的组合物和系统的药物递送媒剂。对于综述,参见Ha等人2016年7月.Acta Pharmaceutica Sinica B.[药学学报]第6卷,第4期,第287-296页;doi.org/10.1016/j.apsb.2016.02.001。

[2165] 离体分化的红细胞也可用作本文所述组合物的载剂。参见,例如,WO 2015073587;WO 2017123646;WO 2017123644;WO 2018102740;WO 2016183482;WO 2015153102;WO 2018151829;WO 2018009838;Shi等人2014.Proc Natl Acad Sci USA.[美国国家科学院院刊]111(28):10131-10136;美国专利9,644,180;Huang等人2017.Nature Communications[自然通讯]8:423;Shi等人2014.Proc Natl Acad Sci USA.[美国国家科学院院刊]111(28):10131-10136。

[2166] 融合体组合物,例如,如WO 2018208728中所述,也可用作载剂以递送本文所述的组合物。

[2167] 膜穿透多肽

[2168] 在一些实施例中,组合物进一步包含膜穿透多肽(MPP),以将组分携带到细胞中或穿过膜例如细胞或核膜。能够促进物质跨膜运输的膜穿透多肽包括但不限于细胞穿透肽(CPP)(参见,例如,美国专利号:8,603,966)、用于植物细胞内递送的融合肽(参见,例如,Ng等人,PLoS One,2016,11:e0154081)、蛋白质转导结构域、木马肽和膜移位信号(MTS)(参见,例如,Tung等人,Advanced Drug Delivery Reviews[先进药物递送论述]55:281-294(2003))。一些MPP富含氨基酸,例如精氨酸,带有带正电的侧链。

[2169] 膜穿透多肽具有诱导组分的膜穿透的能力,并且在全身性施用后可在体内在多个组织的细胞内进行大分子移位。膜穿透多肽还可以指在适当条件下与细胞接触时,从外部环境通过进入细胞内环境(包括细胞质,细胞器如线粒体或细胞核)中的肽,其量明显超过被动扩散所能达到的量。

[2170] 跨膜运输的组分可以可逆地或不可逆地连接至膜穿透多肽。接头可以是化学键,例如一个或多个共价键或非共价键。在一些实施例中,接头是肽接头。这样的接头可以在2-30个氨基酸之间,或更长。接头包括柔性、刚性或可切割的接头。

[2171] 组合

[2172] 在一方面，本文所述的指环体或包含指环体的组合物还可包含一个或多个异源部分。在一方面，本文所述的指环体或包含指环体的组合物还可在融合体中包含一个或多个异源部分。在一些实施例中，异源部分可以与遗传元件连接。在一些实施例中，异源部分可以被封闭在蛋白质外部中作为指环体的一部分。在一些实施例中，异源部分可以与指环体一起施用。

[2173] 在一方面，本发明包括一种细胞或组织，其包含本文所述的指环体和异源部分中的任一种。

[2174] 在另一方面，本发明包括一种药物组合物，其包含本文所述的指环体和异源部分。

[2175] 在一些实施例中，异源部分可以是病毒(例如，效应子(例如，药物、小分子)、靶向剂(例如，DNA靶向剂、抗体、受体配体)、标签(例如，荧光团、光敏剂例如KillerRed)或本文所述的编辑或靶向部分。在一些实施例中，本文所述的膜移位多肽与一个或多个异源部分连接。在一个实施例中，异源部分是小分子(例如，肽模拟物或分子量小于2000道尔顿的有机小分子)、肽或多肽(例如，抗体或其抗原结合片段)、纳米颗粒、适体或药剂。

[2176] 病毒

[2177] 在一些实施例中，组合物可进一步包含病毒作为异源部分，例如单链DNA病毒，例如指环病毒、比那病毒(Bidnavirus)、圆环病毒、双生病毒、基诺病毒(Genomovirus)、丝状病毒、微小病毒、纳米病毒、细小病毒和斯派拉病毒(Spiravirus)。在一些实施例中，组合物可进一步包含双链DNA病毒，例如腺病毒、瓶状病毒、囊泡病毒、非洲猪瘟病毒、杆状病毒、福斯罗病毒(Fusellovirus)、球状病毒、滴状病毒、腺炎病毒、疱疹病毒、虹彩病毒、脂毛病毒、尼马病毒(Nimavirus)和痘病毒。在一些实施例中，组合物可以进一步包含RNA病毒，例如甲病毒、真菌传杆状病毒、肝炎病毒、大麦病毒、烟草花叶病毒、烟草脆裂病毒、三角病毒(Tricornavirus)、风疹病毒、双RNA病毒、囊状病毒、分体病毒和呼肠病毒。在一些实施例中，将指环体与作为异源部分的病毒一起施用。

[2178] 在一些实施例中，异源部分可以包含非致病性病毒，例如，共生病毒、共栖的病毒、天然病毒。在一些实施例中，非致病性病毒是一种或多种指环病毒，例如，甲型细环病毒(TT)、乙型细环病毒(TTM)和丙型细环病毒(TTMD)。在一些实施例中，指环病毒可以包括细环病毒(TT)、SEN病毒、前哨病毒、TTV样微小病毒、TT病毒、TT病毒基因型6、TT病毒群、TTV样病毒DXL1、TTV样病毒DXL2、细环样微小病毒(TTM)或细环样中型病毒(TTMD)。在一些实施例中，非致病性病毒包含一个或多个序列，所述序列与本文所述例如如表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15、17或41中的任一个中所列的核苷酸序列中任一个具有至少约60％、70％、80％、85％、90％、95％、96％、97％、98％和99％核苷酸序列同一性。

[2179] 在一些实施例中，异源部分可包含一种或多种被鉴定为受试者缺乏的病毒。例如，可以向被鉴定为患有病毒缺乏症的受试者施用组合物，所述组合物包含指环体和在所述受试者中失衡或具有不同于参考值例如健康受试者的比率的一种或多种病毒组分或病毒。

[2180] 在一些实施例中，异源部分可以包含一种或多种非指环病毒，例如腺病毒、疱疹病毒、pox病毒、痘苗病毒、SV40、乳头瘤病毒、RNA病毒(例如逆转录病毒、例如慢病毒)、单链RNA病毒(例如肝炎病毒)或双链RNA病毒(例如轮状病毒)。在一些实施例中，指环体或病毒是有缺陷的，或需要协助以产生感染性颗粒。可以例如通过使用辅助细胞系来提供这种协助，所述辅助细胞系包含在LTR内的调节性序列的控制下编码复制缺陷型指环体或病毒的

(例如所有)结构基因中一个或多个的核酸,例如整合到基因组中的质粒或DNA。用于复制本文所述的指环体的合适细胞系包括本领域已知的细胞系,例如A549细胞,其可以如本文所述进行修饰。

[2181] 靶向部分

[2182] 在一些实施例中,本文所述的组合物或指环体可进一步包含靶向部分,例如与存在于靶细胞上的目的分子特异性结合的靶向部分。靶向部分可调节目的分子或细胞的特定功能,调节特定分子(例如酶、蛋白质或核酸),例如通路中目的分子下游的特定分子,或特异性结合靶标以定位指环体或遗传元件。例如,靶向部分可以包括与目的特定分子相互作用以增加,减少或以其他方式调节其功能的治疗剂。

[2183] 标记部分或监测部分

[2184] 在一些实施例中,本文所述的组合物或指环体可以进一步包含标签以标记或监测本文所述的指环体或遗传元件。标记部分或监测部分可以通过化学试剂或酶促裂解,例如蛋白水解或内含肽剪接而去除。亲和标签可用于使用亲和技术纯化标记的多肽。一些实例包括几丁质结合蛋白(CBP)、麦芽糖结合蛋白(MBP)、谷胱甘肽-S-转移酶(GST)和聚(His)标签。增溶标签可用于帮助在分子伴侣缺乏的物种(例如大肠杆菌)中表达的重组蛋白,以帮助蛋白质正确折叠并防止其沉淀。一些实例包括硫氧还蛋白(TRX)和聚(NANP)。标记部分或监测部分可以包括光敏标签,例如荧光。荧光标签对于可视化是有用的。GFP及其变体是通常用作荧光标签的一些实例。蛋白质标签可允许发生特定的酶促修饰(例如通过生物素连接酶进行生物素化)或化学修饰(例如与FlAsH-EDT2反应进行荧光成像)。通常将标记部分或监测部分组合,以便将蛋白质连接到多个其他组分。标记部分或监测部分也可以通过特异性蛋白水解或酶促裂解(例如通过TEV蛋白酶、凝血酶、因子Xa或肠肽酶)去除。

[2185] 纳米颗粒

[2186] 在一些实施例中,本文所述的组合物或指环体可以进一步包含纳米颗粒。纳米颗粒包括无机材料,其尺寸在约1至约1000纳米之间、在约1至约500纳米之间、在约1至约100nm之间、在约50nm至约300nm之间、在约75nm至约200nm之间、在约100nm至约200nm之间、以及其间的任何范围。纳米颗粒通常具有纳米尺度的复合结构。在一些实施例中,纳米颗粒通常是球形的,尽管取决于纳米颗粒的组成,不同的形态是可能的。纳米颗粒与纳米颗粒外部的环境接触的部分通常被鉴定为纳米颗粒的表面。在本文所述的纳米颗粒中,尺寸限制可被限制为二维,从而纳米颗粒包括直径为约1至约1000nm的复合结构,其中特定直径取决于纳米颗粒的组成和根据实验设计的纳米颗粒的预期用途。例如,用于治疗应用的纳米颗粒通常具有约200nm或更小的尺寸。

[2187] 鉴于感兴趣的具体应用,纳米颗粒的其他期望的性质(例如表面电荷和空间稳定性)也可以变化。在Davis等人,Nature[自然]2008第7卷,第771-782页;Duncan,Nature[自然]2006第6卷,第688-701页;和Allen,Nature[自然]2002第2卷第750-763页中描述了在临床应用(例如癌症治疗)中可能需要的示例性特性,每个文章均通过引用整体并入本文。技术人员在阅读本发明后可以鉴定出其他性质。纳米颗粒的尺寸和性质可以通过本领域已知的技术来检测。检测颗粒尺寸的示例性技术包括但不限于动态光散射(DLS)和各种显微技术,例如透射电子显微镜(TEM)和原子力显微镜(AFM)。检测颗粒形态的示例性技术包括但不限于TEM和AFM。检测纳米颗粒的表面电荷的示例性技术包括但不限于ζ电势方法。适用于

检测其他化学性质的其他技术包括通过^{1}H、^{11}B、和^{13}C和^{19}F NMR、UV/Vis和红外/拉曼光谱和荧光光谱(当纳米颗粒与荧光标记组合使用时)和技术人员可以鉴定的其他技术。

[2188] 小分子

[2189] 在一些实施例中,本文所述的组合物或指环体可以进一步包含小分子。小分子部分包括但不限于小肽、拟肽(例如类肽)、氨基酸、氨基酸类似物、合成多核苷酸、多核苷酸类似物、核苷酸、核苷酸类似物、通常具有小于约5,000克/摩尔的分子量的有机和无机化合物(包括杂有机和有机金属化合物),例如,具有小于约2,000克/摩尔的分子量的有机或无机化合物,例如,具有小于约1,000克/摩尔的分子量的有机或无机化合物,例如,具有小于约500克/摩尔的分子量的有机或无机化合物,以及此类化合物的盐、酯和其他医药可接受形式。小分子可以包括但不限于神经递质、激素、药物、毒素、病毒或微生物颗粒、合成分子以及激动剂或拮抗剂。

[2190] 合适的小分子的实例包括在以下中描述的那些:“The Pharmacological Basis of Therapeutics[治疗学的药理基础],”Goodman和Gilman,McGraw-Hill,纽约州纽约,(1996),第九版,在以下章节:Drugs Acting at Synaptic and Neuroeffector Junctional Sites[药物作用于突触和神经效应连接点];Drugs Acting on the Central Nervous System[药物作用于中枢神经系统];Autacoids:Drug Therapy of Inflammation[自体活性物质:炎症的药物治疗];Water,Salts and Ions[水,盐和离子];Drugs Affecting Renal Function and Electrolyte Metabolism[药物影响肾功能和电解质代谢];Cardiovascular Drugs[心血管药物];Drugs Affecting Gastrointestinal Function[药物影响胃肠功能];Drugs Affecting Uterine Motility[药物影响子宫动力];Chemotherapy of Parasitic Infections[寄生虫感染的化学疗法];Chemotherapy of Microbial Diseases[微生物疾病的化学疗法];Chemotherapy of Neoplastic Diseases[肿瘤性疾病的化学疗法];Drugs Used for Immunosuppression[药物用于免疫抑制];Drugs Acting on Blood-Forming organs[药物作用于造血器官];Hormones and Hormone Antagonists;[激素和激素拮抗剂];Vitamins,Dermatology[维生素,皮肤病学];和Toxicology[毒理学],均以引用方式并入本文。小分子的一些实例包括但不限于朊病毒药物,例如他克莫司、泛素连接酶或HECT连接酶抑制剂例如heclin、组蛋白修饰药物例如丁酸钠、酶抑制剂例如5-氮杂胞苷、蒽环类药物例如阿霉素、β-内酰胺类例如青霉素、抗菌剂、化学治疗剂、抗病毒剂、来自其他生物体的调节剂例如VP64、以及生物利用度不足的药物、例如药代动力学不足的化学治疗药。

[2191] 在一些实施例中,小分子是一种表观遗传修饰剂,例如,如de Groote等人Nuc.Acids Res.[核酸研究](2012):1-18中所述的那些。示例性的小分子表观遗传修饰剂描述于例如Lu等人J.Biomolecular Screening[生物分子筛选杂志]17.5(2012):555-71(例如表1或2)中,其通过引用并入本文。在一些实施例中,表观遗传修饰剂包括伏立诺他或罗米地辛。在一些实施例中,表观遗传修饰剂包含I、II、III和/或IV类组蛋白脱乙酰酶(HDAC)的抑制剂。在一些实施例中,表观遗传修饰剂包含SirTI的激活剂。在一些实施例中,表观遗传修饰剂包括山竹子素,Lys-CoA、C646、(+)-JQI、I-BET、BICI、MS120、DZNep、UNC0321、EPZ004777、AZ505、AMI-I、吡唑酰胺7b、苯并[d]咪唑17b、酰化氨苯砜衍生物(例如PRMTI)、methylstat、4,4’-二羧基-2,2’-联吡啶、SID 85736331、异羟肟酸酯类似物8、

tanylcypromie、双胍和双胍多胺类似物、UNC669、维达扎、地西他滨、苯丁酸钠(SDB)、硫辛酸(LA)、槲皮素、丙戊酸、肼苯哒嗪、新诺明、绿茶提取物(例如表没食子儿茶素没食子酸酯(EGCG))、姜黄素、萝卜硫素和/或大蒜素/二烯丙基二硫化物。在一些实施例中,表观遗传修饰剂抑制DNA甲基化,例如DNA甲基转移酶的抑制剂(例如5-氮杂胞苷和/或地西他滨)。在一些实施例中,表观遗传修饰剂修饰组蛋白修饰,例如组蛋白乙酰化、组蛋白甲基化、组蛋白类泛素化和/或组蛋白磷酸化。在一些实施例中,表观遗传修饰剂是组蛋白脱乙酰基酶的抑制剂(例如伏立诺他和/或曲古抑菌素A)。

[2192] 在一些实施例中,小分子是药物活性剂。在一个实施例中,小分子是代谢活性或组分的抑制剂。有用的药物活性剂类别包括但不限于抗生素、抗炎药、血管生成剂或血管活性剂、生长因子和化学治疗(抗肿瘤)剂(例如肿瘤抑制剂)。可以使用来自本文描述的类别和实例或来自(Orme-Johnson 2007,Methods Cell Biol.[细胞生物学方法]2007;80:813-26)的分子的一种或组合。在一个实施例中,本发明包括一种组合物,其包含抗生素、抗炎药、血管生成剂或血管活性剂、生长因子或化学治疗剂。

[2193] 肽或蛋白质

[2194] 在一些实施例中,本文所述的组合物或指环体可以进一步包含肽或蛋白质。肽部分可包括但不限于肽配体或抗体片段(例如,结合受体如细胞外受体的抗体片段)、神经肽、激素肽、肽药物、毒性肽、病毒或微生物肽、合成肽、以及激动剂或拮抗剂肽。

[2195] 肽部分可以是直链或支链的。肽的长度为约5至约200个氨基酸、约15至约150个氨基酸、约20至约125个氨基酸、约25至约100个氨基酸或其间的任何范围。

[2196] 肽的一些实例包括但不限于荧光标签或标记、抗原、抗体、抗体片段例如单域抗体、配体和受体例如胰高血糖素样肽-1(GLP-1)、GLP-2受体2、胆囊收缩素B(CCKB)和生长抑素受体、肽治疗剂例如与特定细胞表面受体例如G蛋白偶联受体(GPCR)或离子通道结合的那些、天然生物活性肽的合成或类似物肽、抗微生物肽、成孔肽、肿瘤靶向或细胞毒性肽、以及降解或自毁肽例如凋亡诱导肽信号或光敏肽。

[2197] 本文所述的可用于本发明的肽还包括小抗原结合肽,例如抗原结合抗体或抗体样片段,例如单链抗体、纳米抗体(参见,例如,Steeland等人2016.Nanobodies as therapeutics:big opportunities for small antibodies.[纳米抗体作为治疗剂:小抗体的巨大机会]Drug Discov Today[今日药物发现]:21(7):1076-113)。这样的小抗原结合肽可以结合细胞质抗原、核抗原、细胞内抗原。

[2198] 在一些实施例中,本文所述的组合物或指环体包括与能够靶向特定位置、组织或细胞的配体连接的多肽。

[2199] 寡核苷酸适体

[2200] 在一些实施例中,本文所述的组合物或指环体可以进一步包含寡核苷酸适体。适体部分是寡核苷酸或肽适体。寡核苷酸适体是单链DNA或RNA(ssDNA或ssRNA)分子,其可以以高亲和力和特异性结合预先选择的靶标(包括蛋白质和肽)。

[2201] 寡核苷酸适体是可以通过重复多轮的体外选择或等效地SELEX(通过指数富集的配体系统进化)而工程化以结合到各种分子靶标(例如小分子、蛋白质、核酸甚至细胞、组织和生物体)的核酸种类。适体提供区分的分子识别,并且可以通过化学合成来产生。另外,适体可能具有理想的储存特性,并且在治疗应用中几乎没有引起免疫原性。

[2202] DNA和RNA适体均可显示出对各种靶标的稳健结合亲和力。例如,已经选择了DNA和RNA适体用于溶菌酶、凝血酶、人免疫缺陷病毒反式反应响应元件(HIV TAR)(参见en.wikipedia.org/wiki/Aptamer-cite_note-10)、氯化血红素、干扰素γ、血管内皮生长因子(VEGF)、前列腺特异性抗原(PSA)、多巴胺和非经典癌基因、热休克因子1(HSF1)。

[2203] 肽适体

[2204] 在一些实施例中,本文所述的组合物或指环体可以进一步包含肽适体。肽适体具有一个(或多个)短可变肽结构域,包括具有低分子量12kDa-14kDa的肽。肽适体可以被设计为特异性结合并干扰细胞内部的蛋白质-蛋白质相互作用。

[2205] 肽适体是经选择或经工程化以结合特定靶分子的人工蛋白质。这些蛋白质包括一个或多个可变序列的肽环。它们通常是从组合文库中分离出来的,并且常常随后通过定向突变或多轮的可变区诱变和选择而得到改进。在体内,肽适体可以结合细胞蛋白靶标并发挥生物学作用,包括干扰其靶向分子与其他蛋白的正常蛋白相互作用。特别地,针对附着于转录因子激活结构域的靶蛋白筛选附着于转录因子结合结构域的可变肽适体环。将肽适体通过该选择策略与其靶标的体内结合检测为下游酵母标记基因的表达。这样的实验鉴定了与适体结合的特定蛋白质,以及适体破坏的蛋白质相互作用以引起表型。另外,用适当的官能部分衍生的肽适体可引起其靶蛋白的特异性翻译后修饰,或改变靶标的亚细胞定位。

[2206] 肽适体还可以在体外识别靶标。已发现它们可代替生物传感器中的抗体,并用于从包含无活性和活性蛋白形式的群体中检测活性蛋白同工型。其中肽适体“头部”与独特的序列双链DNA“尾部”共价连接的称为蝌蚪的衍生物可通过其DNA尾部的PCR(例如,使用定量实时聚合酶链反应)定量混合物中的稀缺靶分子。

[2207] 可以使用不同的系统进行肽适体选择,但目前使用最多的是酵母双杂交系统。肽适体还可以选自通过噬菌体展示和其他表面展示技术(例如mRNA展示、核糖体展示、细菌展示和酵母展示)构建的组合肽文库。这些实验程序也称为生物淘选。在从生物淘选获得的肽中,模拟表位可被认为是一种肽适体。从组合肽文库淘选的所有肽已存储在名为MimoDB的特殊数据库中。

[2208] IV.宿主

[2209] 本发明进一步涉及包含本文所述的指环体的宿主或宿主细胞。在一些实施例中,宿主或宿主细胞是植物、昆虫、细菌、真菌、脊椎动物、哺乳动物(例如人)或其他生物体或细胞。在某些实施例中,如本文所证实的,所提供的指环体感染一系列不同的宿主细胞。靶宿主细胞包括中胚层、内胚层或外胚层来源的细胞。靶宿主细胞包括例如上皮细胞、肌肉细胞、白细胞(例如淋巴细胞)、肾组织细胞、肺组织细胞。

[2210] 在一些实施例中,指环体在宿主中是基本上非免疫原性的。指环体或遗传元件无法通过宿主的免疫系统产生不希望的实质性反应。一些免疫应答包括但不限于体液免疫应答(例如,抗原特异性抗体的产生)和细胞介导的免疫应答(例如,淋巴细胞增殖)。

[2211] 在一些实施例中,宿主或宿主细胞与指环体接触(例如,感染)。在一些实施例中,宿主是哺乳动物,例如人。施用后任何时候都可以测量宿主体内的指环体的量。在某些实施例中,测定培养物中指环体生长的时程。

[2212] 在一些实施例中,指环体例如本文所述的指环体是可遗传的。在一些实施例中,指环体在流体和/或细胞中从母亲线性传输给孩子。在一些实施例中,来自原始宿主细胞的子

细胞包含指环体。在一些实施例中，母亲以至少25％、50％、60％、70％、80％、85％、90％、95％或99％的效率将指环体传输给孩子或从宿主细胞到子细胞的传输效率为至少25％、50％、60％、70％、80％、85％、90％、95％或99％。在一些实施例中，宿主细胞中的指环体在减数分裂期间具有至少25％、50％、60％、70％、80％、85％、90％、95％或99％的传输效率。在一些实施例中，宿主细胞中的指环体在有丝分裂期间具有至少25％、50％、60％、70％、80％、85％、90％、95％或99％的传输效率。在一些实施例中，细胞中的指环体具有约10％-20％、20％-30％、30％-40％、40％-50％、50％-60％、60％-70％、70％-75％、75％-80％、80％-85％、85％-90％、90％-95％、95％-99％或其间任何百分比的传输效率。

[2213] 在一些实施例中，指环体例如指环体在宿主细胞内复制。在一实施例中，指环体能够在哺乳动物细胞例如人细胞中复制。在其他实施例中，指环体复制缺陷或复制无能力。

[2214] 尽管在一些实施例中，指环体在宿主细胞中复制，但是指环体未整合到宿主的基因组中，例如未与宿主的染色体整合。在一些实施例中，指环体具有例如与宿主的染色体的可忽略的重组频率。在一些实施例中，指环体与例如宿主的染色体的重组频率例如小于约1.0cM/Mb、0.9cM/Mb、0.8cM/Mb、0.7cM/Mb、0.6cM/Mb、0.5cM/Mb、0.4cM/Mb、0.3cM/Mb、0.2cM/Mb、0.1cM/Mb或更低。

[2215] V.使用方法

[2216] 本文所述的指环体和包含指环体的组合物可用于例如在需要其的受试者(例如哺乳动物受试者，例如人受试者)中治疗疾病、障碍或病症的方法中。本文描述的药物组合物的施用可以例如通过肠胃外(包括静脉内、肿瘤内、腹膜内、肌内、腔内和皮下)施用。指环体可以单独施用或配制成药物组合物。

[2217] 指环体可以以单位剂量组合物的形式，例如单位剂量肠胃外组合物的形式来施用。这样的组合物通常通过混合制备，并且可以适合于肠胃外施用。这样的组合物可以是例如可注射和可输注的溶液或悬浮液或栓剂或气雾剂的形式。

[2218] 在一些实施例中，例如本文所述的指环体或包含其的组合物的施用可导致将指环体所包含的遗传元件递送至例如受试者的靶细胞中。

[2219] 本文所述的指环体或其组合物，例如，包含效应子(例如，内源性或外源性效应子)，可用于将效应子递送至细胞、组织或受试者。在一些实施例中，指环体或其组合物用于将效应子递送至骨髓、血液、心脏、GI或皮肤。通过施用本文所述的指环体组合物来递送效应子可以调节(例如，增加或减少)细胞、组织或受试者中非编码RNA或多肽的表达水平。以这种方式调节表达水平可导致效应子被递送其中的细胞中功能活性的改变。在一些实施例中，调节的功能活性本质上可以是酶促的、结构的或调节的。

[2220] 在一些实施例中，指环体或其拷贝在递送至细胞中后24小时(例如1天、2天、3天、4天、5天、6天、1周、2周、3周、4周、30天或1个月)可在细胞中检测到。在实施例中，指环体或其组合物介导对靶细胞的作用，并且所述作用持续至少1、2、3、4、5、6或7天，2、3或4周，或1、2、3、6或12个月。在一些实施例中(例如，其中指环体或其组合物包含编码外源蛋白质的遗传元件)，所述作用持续少于1、2、3、4、5、6或7天，2、3或4周，或1、2、3、6或12个月。

[2221] 可以用本文所述的指环体或包含该指环体的组合物治疗的疾病、障碍和病症的实例包括但不限于：免疫障碍、干扰素病(例如I型干扰素病)、传染病、炎性障碍、自身免疫性病症、癌症(例如实体瘤，例如肺癌、非小细胞肺癌，例如，表达对mIR-625有响应的基因(例

如,胱天蛋白酶3)的肿瘤)和胃肠道障碍。在一些实施例中,指环体在与指环体接触的细胞中调节(例如增加或减少)活性或功能。在一些实施例中,指环体在与指环体接触的细胞中调节(例如增加或减少)分子(例如核酸或蛋白质)的水平或活性。在一些实施例中,指环体将与指环体接触的细胞例如癌细胞的存活力降低例如至少约10%、20%、30%、40%、50%、60%、70%、80%、90%、95%、99%或更高。在一些实施例中,指环体包含效应子,例如miRNA,例如miR-625,其将与指环体接触的细胞例如癌细胞的存活力降低例如至少约10%、20%、30%、40%、50%、60%、70%、80%、90%、95%、99%或更高。在一些实施例中,指环体例如通过增加胱天蛋白酶-3活性将与指环体接触的细胞例如癌细胞的细胞凋亡增加例如至少约10%、20%、30%、40%、50%、60%、70%、80%、90%、95%、99%或更高。在一些实施例中,指环体包含效应子,例如miRNA,例如miR-625,其例如通过增加胱天蛋白酶-3活性将与指环体接触的细胞例如癌细胞的细胞凋亡增加例如至少约10%、20%、30%、40%、50%、60%、70%、80%、90%、95%、99%或更高。

[2222] VI.生产方法

[2223] 产生遗传元件

[2224] 制备指环体的遗传元件的方法描述于以下中,例如,Khudyakov&Fields,Artificial DNA:Methods and Applications[人工DNA:方法与应用],CRC出版社(2002);Zhao,Synthetic Biology:Tools and Applications[合成生物学:工具和应用],(第一版),学术出版社(Academic Press)(2013);和Egli&Herdewijn,Chemistry and Biology of Artificial Nucleic Acids[人工核酸的化学与生物学],(第一版),Wiley-VCH(2012)。

[2225] 在一些实施例中,可以使用计算机辅助设计工具来设计遗传元件。可将指环体分为更容易合成的较小的重叠块(例如,在约100bp至约10kb片段的范围内或单独的ORF)。这些DNA片段由一组重叠的单链寡核苷酸合成。然后将所得的重叠合成子装配成较大的DNA块,例如指环体。可以将片段或ORF装配成指环体,例如,体外重组或在5'和3'末端的独特的限制性位点以实现连接。

[2226] 替代地,可以使用设计算法来合成遗传元件,所述设计算法将指环体解析为寡核苷酸长度的片段,考虑到序列空间的复杂性,为合成创造最佳的设计条件。然后在基于半导体的高密度芯片上化学合成寡核苷酸,每个芯片可合成超过200,000个单独的寡核苷酸。用诸如BioFab®的装配技术装配寡核苷酸,以由较小的寡核苷酸构建更长的DNA片段。这是通过并行方式完成的,因此一次可以构建成百上千个合成的DNA片段。

[2227] 可以对每个遗传元件或遗传元件的片段进行序列验证。在一些实施例中,可以使用AnyDot.chips(Genovoxx,德国)进行RNA或DNA的高通量测序,其允许监测生物学过程(例如,miRNA表达或等位基因变异性(SNP检测)。具体而言,AnyDot.chips使核苷酸荧光信号检测提高10x-50x。AnyDot.chips及其使用方法在以下中有部分地描述:国际公开申请号WO 02088382、WO 03020968、WO 0303 1947、WO 2005044836、PCTEP 05105657、PCMEP 05105655;以及德国专利申请号DE 101 49 786、DE 102 14 395、DE 103 56 837、DE 10 2004 009 704、DE 10 2004 025 696、DE 10 2004 025 746、DE 10 2004 025 694、DE 10 2004 025 695、DE 10 2004 025 744、DE 10 2004 025 745和DE 10 2005 012 301。

[2228] 其他高通量测序系统包括在Venter,J.,等人Science[科学]2001年2月16日;Adams,M.等人Science[科学]2000年3月24日;和M.J,Levene,等人Science[科学]299:682-

686,2003年1月;以及US公布申请号20030044781和2006/0078937中披露的那些。总体上,这样的系统包括通过经由在核酸分子上测量的聚合反应在时间上添加碱基来对具有多个碱基的靶核酸分子进行测序,即实时跟踪核酸聚合酶对要测序的模板核酸分子的活性。然后可以通过在碱基添加顺序的每个步骤中通过核酸聚合酶的催化活性来鉴定哪个碱基被掺入到靶核酸的生长互补链中来推导该序列。在适合于沿着靶核酸分子移动并在活性位点延伸寡核苷酸引物的位置提供在靶核酸分子复合物上的聚合酶。在活性位点附近提供多个经标记的类型的核苷酸类似物,每个可区分类型的核苷酸类似物与靶核酸序列中的不同核苷酸互补。通过使用聚合酶以在活性位点向核酸链添加核苷酸类似物来延伸正在生长的核酸链,其中所添加的核苷酸类似物在活性位点与靶核酸的核苷酸互补。鉴定作为聚合步骤结果的添加至寡核苷酸引物的核苷酸类似物。重复提供经标记的核苷酸类似物,聚合生长的核酸链和鉴定添加的核苷酸类似物的步骤,从而使核酸链进一步延伸并鉴定靶核酸的序列。

[2229] 在一些实施例中,执行鸟枪测序。在鸟枪测序中,将DNA随机分为许多小片段,使用链终止法对其进行测序以获得读数。通过执行几轮这种片段化和测序,可以获得靶DNA的多个重叠读数。然后,计算机程序使用不同读数的重叠末端以将它们组装成连续序列。

[2230] 在一些实施例中,相对于遗传元件,可以以顺式或以反式提供用于复制或包装的因子。例如,当以顺式提供时,遗传元件可以包含一个或多个编码例如如本文所述的指环病毒ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3或ORF2t/3的基因。在一些实施例中,复制和/或包装信号可以掺入遗传元件中,例如,以诱导扩增和/或包封。在一些实施例中,这在指环体基因组的较大区域的背景下进行(例如,将效应子插入基因组中的特定位点,或用效应子替换病毒ORF)。

[2231] 在另一个实例中,当以反式提供时,遗传元件可能缺乏编码例如如本文所述的指环病毒ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3,或ORF2t/3中的一个或多个的基因;该一种或多种蛋白质可以由例如另一核酸例如辅助核酸提供。在一些实施例中,遗传元件中存在最小的顺式信号(例如,5'UTR和/或GC富集区)。在一些实施例中,遗传元件不编码复制或包装因子(例如,复制酶和/或衣壳蛋白)。在一些实施例中,此类因子可由一种或多种辅助核酸(例如,辅助病毒核酸、辅助质粒或整合到宿主细胞基因组中的辅助核酸)提供。在一些实施例中,辅助核酸表达足以诱导扩增和/或包装的蛋白质和/或RNA,但可能缺乏它们自己的包装信号。在一些实施例中,将遗传元件和辅助核酸引入宿主细胞(例如,同时地或分开地),导致遗传元件而不是辅助核酸的扩增和/或包装。

[2232] 体外环化

[2233] 在某些情况下,要包装到蛋白质外部内的遗传元件是单链环状DNA。在一些情况下,遗传元件可以以不同于单链环状DNA的形式引入宿主细胞。例如,遗传元件可以作为双链环状DNA引入宿主细胞。双链环状DNA然后可以在宿主细胞中转化为单链环状DNA(例如,包含用于滚环复制的合适酶的宿主细胞,例如指环病毒Rep蛋白,例如,Rep68/78、Rep60、RepA、RepB、Pre、MobM、TraX、TrwC、Mob02281、Mob02282、NikB、ORF50240、NikK、TecH、OrfJ或TraI,例如,如Wawrzyniak等人2017,Front.Microbiol.[微生物学前沿]8:2353所述;就列出的酶而言通过引用并入本文)。在一些实施例中,双链环状DNA通过体外环化产生,例如,如实例35中所述。通常,体外环化的DNA构建体可以通过消化包含待包装的遗传元件序列的

质粒来产生，使得遗传元件序列被切除为线性DNA分子。然后可以连接所得线性DNA，例如，使用DNA连接酶，以形成双链环状DNA。在一些情况下，由体外环化产生的双链环状DNA可以进行滚环复制，例如，如本文所述。不希望受理论束缚，预期体外环化产生双链DNA构建体，所述构建体无需进一步修饰即可进行滚环复制，从而能够产生合适大小的单链环状DNA以包装进入指环体，例如，如本文所述。在一些实施例中，双链DNA构建体小于质粒(例如，细菌质粒)。在一些实施例中，双链DNA构建体从质粒(例如细菌质粒)中切除，然后例如通过体外环化进行环化。

[2234] 生产指环体

[2235] 可以多种方式使用如本文所述制备的遗传元件和包含遗传元件的载体，以在适当的宿主细胞中表达指环体。在一些实施例中，遗传元件和包含遗传元件的载体被转染到适当的宿主细胞中，并且所得的RNA可以指导指环体基因产物例如非致病性蛋白质和蛋白质结合序列的高水平表达。提供高水平表达的宿主细胞系统包括提供病毒功能的连续细胞系，例如分别重叠感染APV或MPV的细胞系，经工程化以补充APV或MPV功能的细胞系等。

[2236] 在一些实施例中，如实例1、2、5、6或15-17中任一项所述地产生指环体。

[2237] 在一些实施例中，将指环体在体外在连续的动物细胞系中培养。根据本发明的一个实施例，细胞系可以包括猪细胞系。在本发明的上下文中设想的细胞系包括永生的猪细胞系，例如但不限于猪肾上皮细胞系PK-15和SK、单髓细胞系3D4/31和睾丸细胞系ST。此外，还包括其他哺乳动物细胞系，如CHO细胞(中国仓鼠卵巢)、MARC-145、MDBK、RK-13、EEL。另外地或替代地，本发明方法的特定实施例利用了动物细胞系，其是上皮细胞系，即上皮谱系细胞的细胞系。易受指环体感染的细胞系包括但不限于人或灵长类动物来源的细胞系，例如人或灵长类动物肾癌细胞系。

[2238] 在一些实施例中，遗传元件和包含遗传元件的载体被转染到表达病毒聚合酶蛋白的细胞系中，以实现指环体的表达。为此，可以将表达指环体聚合酶蛋白的转化细胞系用作合适的宿主细胞。宿主细胞可以类似地被工程化以提供其他病毒功能或额外功能。

[2239] 为了制备本文披露的指环体，可以使用本文披露的遗传元件或包含遗传元件的载体来转染细胞，所述细胞提供指环体蛋白以及复制和产生所需的功能。可替代地，可以在被本文披露的遗传元件或包含遗传元件的载体转染之前、期间或之后用辅助病毒来转染细胞。在一些实施例中，辅助病毒可用于补充不完整病毒颗粒的产生。辅助病毒可能具有条件性生长缺陷，例如宿主范围限制或温度敏感性，从而允许随后选择转染病毒。在一些实施例中，辅助病毒可以提供宿主细胞利用的一种或多种复制蛋白以实现指环体的表达。在一些实施例中，可以用编码病毒蛋白例如一种或多种复制蛋白的载体转染宿主细胞。在一些实施例中，辅助病毒包括抗病毒敏感性。

[2240] 可以通过本领域已知的多种技术将本文披露的遗传元件或包含遗传元件的载体复制和产生成指环体颗粒，例如，美国专利号4,650,764;美国专利号5,166,057;美国专利号5,854,037;欧洲专利公开EP 0702085 A1;美国专利申请系列号09/152,845;国际专利公开PCT WO97/12032;WO 96/34625;欧洲专利公开EP-A780475;WO 99/02657;WO 98/53078;WO 98/02530;WO 99/15672;WO 98/13501;WO 97/06270;和EPO 780 47SA1中所述，其中每一个均通过引用整体并入本文。

[2241] 根据本发明的含指环体的细胞培养物的产生可以以不同的规模进行，例如在烧

瓶、滚瓶或生物反应器中。用于培养要感染的细胞的培养基是技术人员已知的，并且可以通常包含细胞活力所需的标准营养物，但是还可以包含取决于细胞类型的其他营养物。任选地，培养基可以不含蛋白质和/或不含血清。根据细胞类型，可以将细胞悬浮培养或在底物上培养。在一些实施例中，不同的培养基用于宿主细胞的生长和指环体的生产。

[2242] 可以根据病毒生产的技术人员已知的方法进行指环体的纯化和分离，例如Rinaldi等人,DNA Vaccines:Methods and Protocols(Methods in Molecular Biology)[DNA疫苗:方法和方案(分子生物学方法)],第3版.2014,Humana出版社所描述。

[2243] 在一方面，本发明包括一种本文所述的指环体的体外复制和繁殖的方法，所述方法可以包括以下步骤：(a)将线性化的遗传元件转染到对指环体感染敏感的细胞系中；(b)收获所述细胞并分离出显示所述遗传元件存在的细胞；(c)根据实验条件和基因表达，将在步骤(b)中获得的细胞培养至少三天，例如至少一周或更长时间；以及(d)收获步骤(c)的细胞。

[2244] 在一些实施例中，可以将指环体引入到生长至高细胞密度的宿主细胞系中。在一些实施例中，在用药物赋形剂配制之前，可以通过基于生物物理特性分离溶质，例如离子交换色谱或切向流过滤来收获和/或纯化指环体。

[2245] VII.施用/递送

[2246] 可以将组合物(例如，包含本文所述的指环体的药物组合物)配制成包括药学上可接受的赋形剂。药物组合物可任选地包含一种或多种另外的活性物质，例如治疗和/或预防活性物质。本发明的药物组合物可以是无菌的和/或无热原的。可以在以下中找到药学制剂的配制和/或生产中的一般考虑：例如，Remington:The Science and Practice of Pharmacy 21st ed.[雷明顿:药学科学与实践第21版],Lippincott Williams&Wilkins,2005(通过引用并入本文)。

[2247] 尽管本文提供的药物组合物的描述主要针对适合于施用给人的药物组合物，但是本领域技术人员应理解，此类组合物通常适合于施用给任何其他动物，例如非人动物，例如非人哺乳动物。为了使组合物适合于施用给各种动物而对适合于施用给人的药物组合物的修饰是众所周知的，并且普通兽医药理师可以仅通过普通的实验(如果有的话)来设计和/或进行这种修饰。预期施用药物组合物的受试者包括但不限于人和/或其他灵长类;哺乳动物，包括与商业有关的哺乳动物，例如牛、猪、马、绵羊、猫、狗、小鼠和/或大鼠；和/或鸟类，包括与商业相关的鸟类，例如家禽、鸡、鸭、鹅和/或火鸡。

[2248] 本文描述的药物组合物的制剂可以通过药理学领域中已知的或以后开发的任何方法来制备。通常，这样的制备方法包括以下步骤：使活性成分与赋形剂和/或一种或多种其他辅助成分结合，然后，如果必要和/或期望的话，将产品分开、成形和/或包装。

[2249] 在一方面，本发明的特征在于一种向受试者递送指环体的方法。所述方法包括向受试者施用包含本文所述的指环体的药物组合物。在一些实施例中，所施用的指环体在受试者中复制(例如，成为受试者的病毒组的一部分)。

[2250] 药物组合物可以包含野生型或天然病毒元件和/或修饰的病毒元件。指环体可以包括表A1-A12、B1-B5、C1-C5或1-18中任一个中的一个或多个序列(例如，核酸序列或编码其氨基酸序列的核酸序列)或与这些核苷酸序列中的任何一个具有至少约60％、65％、70％、75％、80％、85％、90％、95％、96％、97％、98％和99％核苷酸序列同一性的序列或与

表A1-A12、B1-B5、C1-C5或1-18中任一个中的序列互补的序列。指环体可包含核酸分子，所述核酸分子包含与表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15、17或41中的任一个中的一个或多个序列具有至少约60％、65％、70％、75％、80％、85％、90％、95％、96％、97％、98％和99％序列同一性的核酸序列。指环体可包含核酸分子，所述核酸分子编码与表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16或18中的任一个中的氨基酸序列的任何一个具有至少约60％、65％、70％、75％、80％、85％、90％、95％、96％、97％、98％和99％序列同一性的氨基酸序列。指环体可包含多肽，所述多肽包含与表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16或18中的任一个中的氨基酸序列的任何一个具有至少约60％、65％、70％、75％、80％、85％、90％、95％、96％、97％、98％和99％序列同一性的氨基酸序列。指环体可以包括表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15、17或41中的任一个中的一个或多个序列，或与表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15、17或41中的任一个中的核苷酸序列的任一个或跟其互补的序列具有至少约60％、65％、70％、75％、80％、85％、90％、95％、96％、97％、98％和99％核苷酸序列同一性的序列。

[2251] 在一些实施例中，指环体足以与参考例如健康对照相比，增加(刺激)内源基因和蛋白质表达例如至少约5％、10％、15％、20％、25％、30％、35％、40％、45％、50％或更多。在某些实施例中，与参考例如健康对照相比，指环体足以降低(抑制)内源基因和蛋白质表达例如至少约5％、10％、15％、20％、25％、30％、35％、40％、45％、50％或更多。

[2252] 在一些实施例中，与参考例如健康对照相比，指环体在宿主或宿主细胞中抑制/增强一种或多种病毒特性，例如嗜性、感染性、免疫抑制/激活例如至少约5％、10％、15％、20％、25％、30％、35％、40％、45％、50％或更多。

[2253] 在一些实施例中，向受试者施用药物组合物，所述药物组合物进一步包含一种或多种在病毒遗传信息中未表示的病毒株系。

[2254] 在一些实施例中，以足以调节病毒感染的剂量和时间施用包含本文所述的指环体的药物组合物。病毒感染的一些非限制性实例包括腺相关病毒、爱知病毒、澳大利亚蝙蝠狂犬病病毒、BK多瘤病毒、班纳病毒、巴马森林病毒、布尼亚韦拉病毒、布尼亚病毒拉克罗斯、布尼亚病毒雪兔、鹿角皮膜疱疹疱疹病毒、钱迪普拉病毒、基孔肯雅病毒、科萨病毒(Cosavirus)A、牛痘病毒、柯萨奇病毒、克里米亚-刚果出血热病毒、登革热病毒、多里病毒、杜拜病毒、杜温哈格病毒、东部马脑炎病毒、埃博拉病毒、回声病毒、脑心肌炎病毒、爱泼斯坦-巴尔病毒、欧洲蝙蝠狂犬病病毒、GB病毒C型/G型肝炎病毒、汉坦病毒、亨德拉病毒、甲型肝炎病毒、乙型肝炎病毒、丙型肝炎病毒、戊型肝炎病毒、丁型肝炎病毒、马痘病毒、人腺病毒、人星形病毒、人冠状病毒、人巨细胞病毒、人肠病毒68、人肠病毒70、人疱疹病毒1、人疱疹病毒2、人疱疹病毒6、人疱疹病毒7、人疱疹病毒8、人免疫缺陷病毒、人乳头瘤病毒1、人乳头瘤病毒2、人乳头瘤病毒16、人乳头瘤病毒18、人副流感、人细小病毒B19、人呼吸道合胞病毒、人鼻病毒、人SARS冠状病毒、人膜细小病毒、人T淋巴病毒、人环状病毒、甲型流感病毒、乙型流感病毒、丙型流感病毒、伊斯法罕病毒、JC多瘤病毒、日本脑炎病毒、胡宁球菌病毒、KI多瘤病毒、昆津病毒、拉各斯蝙蝠病毒、维多利亚湖马尔堡病毒、兰加特病毒、拉沙病毒、洛兹代尔病毒、跳跃病病毒、淋巴细胞性脉络膜脑膜炎病毒、马库波病毒、马亚罗病毒、MERS冠状病毒、麻疹病毒、芒果脑心肌炎病毒、默克尔细胞多瘤病毒、莫可拉病毒、软体动物接触性疱疹病毒、猴痘病毒、腮腺炎病毒、默里谷脑炎病毒、尼帕病毒、诺沃克病毒、奥-奈氏病

毒、口疮病毒、奥罗泼希病毒、皮钦德病毒、脊髓灰质炎病毒、庞塔托鲁静脉病毒、普马拉病毒、狂犬病病毒、裂谷热病毒、罗莎病毒A、罗斯河病毒、轮状病毒A、轮状病毒B、轮状病毒C、风疹病毒、鹭山病毒、萨里病毒(Salivirus)A、桑蝇热西西里病毒、塞姆利基森林病毒、猿猴泡沫病毒、猿猴病毒5、辛德比斯病毒、圣路易斯脑炎病毒、蜱传波瓦桑病毒、细环病毒、尤库尼米病毒、牛痘病毒、水痘带状疱疹病毒、天花病毒、委内瑞拉马脑炎病毒、水泡口腔炎病毒、西部马脑炎病毒、WU多瘤病毒、西尼罗河病毒、亚巴猴肿瘤病毒、亚巴样病病毒、黄热病病毒和寨卡病毒。在某些实施例中，相对于参考，指环体足以胜出和/或替换已经存在于受试者中的病毒例如至少约5％、10％、15％、20％、25％、30％、35％、40％、45％、50％或更高。在某些实施例中，指环体足以与慢性或急性病毒感染竞争。在某些实施例中，可预防性地施用指环体以免遭病毒感染(例如预防病毒)。在一些实施例中，指环体的量足以调节(例如表型、病毒水平、基因表达、与其他病毒竞争、疾病状态等至少约5％、10％、15％、20％、25％、30％、35％、40％、45％、50％或更高)。

[2255] 本文引用的所有参考和出版物均通过引用并入本文。

[2256] 提供以下实例以进一步说明本发明的一些实施例，但无意于限制本发明的范围；通过它们的示例性性质将理解，可以替换地使用本领域技术人员已知的其他程序、方法或技术。

[2257] 实例

[2258] 内容列表

[2259] 实例1：指环体的制备：抑制干扰素(IFN)表达的合成的指环体的设计和合成

[2260] 实例2：大规模产生指环体(指环体A和/或B)：指环体的产生和繁殖

[2261] 实例3：指环体的体外作用(指环体A)：细胞感染后指环体的表达和效应子功能例如miRNA的表达的体外评估

[2262] 实例4：指环体的免疫学作用(指环体)：施用后指环体的体内效应子功能，例如miRNA的表达

[2263] 实例5：制备合成的指环体：合成的指环体的体外产生

[2264] 实例6：指环体的装配和感染：使用实例5中所述的合成DNA序列在体外产生感染性指环体

[2265] 实例7：指环体的选择性：体外产生的合成的指环体感染多种组织来源的细胞系的能力

[2266] 实例8：蛋白质结合序列的鉴定和使用：指环病毒基因组中推定的蛋白质结合位点

[2267] 实例9：指环病毒基因组

[2268] 实例10：将各种长度的核苷酸插入指环病毒基因组中：将各种长度的DNA序列添加到指环病毒基因组中

[2269] 实例11：要递送的示例性货物：指环体中核酸和蛋白质有效负载的示例性类别

[2270] 实例12：示例性有效负载整合位点

[2271] 实例13：定义的类别的指环病毒和其保守区

[2272] 实例14：复制缺陷型指环体和辅助病毒

[2273] 实例15：具有复制能力的指环体的制造方法

[2274] 实例16：复制缺陷型指环体的制造方法：复制缺陷型指环体的回收和扩大生产

[2275] 实例17:使用悬浮细胞生产指环体:在悬浮细胞中生产指环体。

[2276] 实例18:通过qPCR对指环体基因组当量进行定量:用以定量指环体的基于水解探针的定量PCR测定的开发

[2277] 实例19:利用指环体在小鼠中表达外源蛋白:使用指环体在体内表达功能性模型蛋白

[2278] 实例20:基因组比对,以测定指环体DNA是否整合到宿主基因组中

[2279] 实例21:指环体整合到宿主基因组中的评估

[2280] 实例22:表达外源性微小RNA序列的指环体的功能作用:使用指环体表达功能性核酸效应子

[2281] 实例23:表达外源性非编码RNA的指环体的制备和产生:使用指环体表达外源性小非编码RNA

[2282] 实例24:指环病毒进化分枝中的保守性:甲型细环病毒属中七个进化分枝的鉴定

[2283] 实例25:从指环体表达内源性miRNA和内源性miRNA的缺失

[2284] 实例26:指环病毒ORF的定位

[2285] 实例27:指环体开发所需区域的表征

[2286] 实例28:外源蛋白质的指环体体内递送:本实例展示了指环体在施用后的体内效应子功能(例如蛋白质的表达)

[2287] 实例29:指环病毒中前体miRNA(前mIR)的鉴定:用以鉴定各种指环病毒编码的新颖前体miRNA的计算和实验方法

[2288] 实例30:确定指环病毒前miR的内源性靶:分析以确定由各种指环病毒株编码的前miR的内源性靶和潜在的治疗相关靶途径

[2289] 实例31:制造编码天然指环病毒前miR的指环体:包装表达天然指环病毒前miR的指环体的复制或非复制形式的过程

[2290] 实例32:在体外细胞培养模型中利用指环病毒前miR作为肿瘤抑制因子:从例如如实例29中所述的分析中鉴定为肿瘤抑制性的候选前miR的表型作用

[2291] 实例33:利用指环病毒前miR作为体内肿瘤抑制因子:体内实验,以确认来自如实例32中所述的体外分析的肿瘤抑制性指环病毒前miR和癌细胞系的肿瘤抑制作用

[2292] 实例34:指环病毒基因组的串联拷贝

[2293] 实例35:体外环化的指环病毒基因组:包含环状双链环病毒基因组DNA和最少非病毒DNA的构建体

[2294] 实例36:ORF1建模和保守残基和结构域的鉴定:乙型细环病毒ORF1蛋白的建模和定义推定的结构域

[2295] 实例37:生产含有来自不同细环病毒株的具有高变结构域的嵌合ORF1的指环体

[2296] 实例38:生产含有非TTV蛋白/肽代替高变结构域的嵌合ORF1

[2297] 实例39:分泌酶的指环体体内传递

[2298] 实例40:设计包含DNA有效负载的指环体

[2299] 实例41:基于tth8和LY2的指环体分别成功地将EPO基因转导到肺癌细胞中

[2300] 实例42:静脉内(i.v.)施用后,可以在体内检测到带有治疗性转基因的指环体

[2301] 实例43:指环病毒中的编码序列大小分布

[2302] 实例44:用以表征ORF2的高度保守基序

[2303] 实例45:人全长指环病毒ORF1 mRNA的证据

[2304] 实例46:体外环化的基因组作为体外生产指环体的输入材料

[2305] 实例47:指环病毒ORF1中保守二级结构基序的鉴定

[2306] 实例1:指环体的制备

[2307] 该实例描述了抑制干扰素(IFN)表达的合成的指环体的设计和合成。

[2308] 设计了一种指环体(指环体A),其起始于1)编码非致病性包装外壳的衣壳基因的DNA序列(Arch Virol(2007)152:1961-1975),登录号:A7XCE8.1(ORF11_TTW3);2)编码靶向宿主基因(例如IFN)的微小RNA的DNA序列(PLOS Pathogen[公共科学图书馆—病原体](2013),9(12),e1003818),登录号:AJ620231.1;和3)与衣壳蛋白中的特定区域(例如,具有登录号:Q99153.1的衣壳特定区域)结合的DNA序列(Journal of Virology[病毒学杂志](2003),77(24),13036-13041)。

[2309] 向该序列添加1kb非编码DNA序列(指环体B)。将所设计的指环体(图2)化学合成为3kb(总大小),并进行序列验证。

[2310] 按照制造商的建议,用JetPEI试剂(PolyPlus-transfection,法国伊尔基希(Illkirch,France))将指环体序列转染到人胚肾293T细胞中(12孔平板上每10^5个细胞1mg)。对照转染仅包含载体或单独用JetPEI转染的细胞,并用编码GFP的报告质粒优化了转染效率。测量对照转染的荧光以确保适当转染细胞。将转染培养物在37℃和5%二氧化碳中孵育过夜。

[2311] 18小时后,在加入新鲜培养基之前,将细胞用PBS洗涤3次。如下收集上清液,用于超速离心和收获指环体。通过以4,000xg离心30分钟然后以8,000xg离心15分钟以除去细胞和细胞碎片来清除培养基。然后将上清液通过0.45μm孔径的过滤器过滤。通过5%蔗糖垫(5ml)以27,000rpm将指环体沉淀1小时,然后将其重悬于1x磷酸盐缓冲盐水(PBS)和0.1%杆菌肽(以1/100原始体积)中。将浓缩的指环体以24,000rpm通过20%至35%蔗糖梯度离心2小时。收集梯度结处的指环体带。然后将指环体用1xPBS稀释并以27,000rpm沉淀1小时。将指环体沉淀重悬于1xPBS中,并通过20%至35%连续蔗糖梯度进一步纯化。

[2312] 实例2:大规模产生指环体(指环体A和/或B)

[2313] 此实例描述了指环体的生产和繁殖。

[2314] 制备实例1中所述的纯化的指环体,用于在具有悬浮生长的生产A549细胞的旋转瓶中大规模扩增。将A549细胞维持在F12K培养基、10%胎牛血清、2mM谷氨酰胺和抗生素中。在37℃和5%二氧化碳中孵育24小时后,以10^6个指环体的指环体负载用指环体感染A549细胞,从而产生约$1x10^7$个指环体颗粒。然后将细胞用PBS洗涤3次,并与新鲜培养基孵育6小时。

[2315] 对于指环体的纯化,如下进行基于氯化铯梯度的两个超速离心步骤,然后进行透析(Bio-Protocol[生物学方案](2012)Bio101:e201)。离心(6000xg,10分钟)除去细胞,然后将上清液先通过0.8μm过滤器过滤,然后再通过0.2μm过滤器过滤。滤液通过滤膜(100,000mw)浓缩至8ml体积。将渗余物负载到硫酸铯溶液中,并以247,000xg离心20h。去除指环体带,放入14,000mw截止分子量的透析管中,并进行透析。如果需要,可以进行进一步的浓缩。

[2316] 实例3:指环体的体外作用(指环体A)

[2317] 该实例描述了细胞感染后指环体的表达和效应子功能例如miRNA的表达的体外评估。

[2318] 通过内源基因调节(例如IFN信号传导)在体外评估实例1中所述的纯化的指环体的作用。将HEK293T细胞用双萤光素酶质粒(具有基于干扰素刺激的响应元件(ISRE)的启动子的萤火虫萤光素酶和具有组成型启动子的转染对照海肾萤光素酶):萤光素酶报告混合物(pcDNA3.1dsRluc与pISRE-Luc,1:4比率(Clonetech))共转染(J Virol[病毒学杂志](2008),82:9823-9828)。

[2319] 向接种在6孔板(2组一式三份的3个对照孔和3个使用指环体A的实验孔)中的HEK293T细胞,以10^7次感染复数施用指环体。

[2320] 48小时后,用含或不含100u/ml通用I型干扰素(PBL,新泽西州皮斯卡塔韦(Piscataway,NJ))的新培养基替换培养基。IFN治疗后十六小时,进行了双萤光素酶测定(J Virol[病毒学杂志](2008),82:9823-9828)以测定IFN信号传导。将萤火虫萤光素酶归一化为海肾萤光素酶表达,以控制转染差异。通过将相当的实验孔除以对照孔,计算出ISRE ffLuc报告基因的诱导倍数,并相对于阴性对照比较每种条件的诱导。

[2321] 在一个实施例中,与对照相比,在指环体治疗组中萤光素酶信号降低将表明该指环体降低了细胞中IFN的产生。

[2322] 实例4:指环体的免疫学作用(指环体)

[2323] 该实例描述了施用后指环体的体内效应子功能,例如miRNA的表达。

[2324] 使用的百倍稀释度,从10^{14}个基因组当量/千克开始,降到0个基因组当量/千克,将按照实例1和2中所述制备的纯化的指环体以各种剂量静脉内施用给健康的猪。为了评估对免疫耐受性的影响,每天向猪注射以上指定剂量的指环体或媒介物对照PBS,持续3天,并在3天后处死。

[2325] 收获脾脏、骨髓和淋巴结。从每个组织制备单细胞悬浮液,并用MHC-II、CD11c和细胞内IFN的细胞外标记物染色。通过流式细胞术分析来自每个组织的MHC+、CD11c+、IFN+抗原呈递细胞,例如,其中对于上述标记物中给定一种呈阳性的细胞是荧光表现高于在相同条件下缺少标记物表达但在其他方面与细胞分析群体相似的阴性对照群体中99%细胞的细胞。

[2326] 在一个实施例中,与对照相比,指环体治疗组中IFN+细胞数量的减少将表明指环体在施用后降低了细胞中IFN的产生。

[2327] 实例5:制备合成的指环体。

[2328] 该实例说明了合成的指环体的体外产生。

[2329] 将EcoRV限制性酶切位点之间来自TTMiniV的LY1和LY2株系的DNA序列(Eur Respir J.[欧洲呼吸杂志]2013年8月;42(2):470-9)克隆到卡那霉素载体(集成DNA技术公司(Integrated DNA Technologies))中。在实例6和7以及图6A-10B中,将包括来自TTMiniV的LY1和LY2株系的DNA序列的指环体分别称为愈指环体1(指环(Anello)1)和指环体2(指环2)。按照制造商的规程,将克隆的构建体转化到10-β感受态大肠杆菌中(新英格兰生物实验室公司(New England Biolabs Inc.)),然后纯化质粒(凯杰公司(Qiagen))。

[2330] 根据制造商的规程,将DNA构建物(图3和图4)在37℃下用EcoRV限制性酶切消化

(新英格兰生物实验室公司(New England Biolabs Inc.))线性化6小时,产生包含TTMiniV基因组的双链线性DNA片段,但不包括细菌骨架元件(例如复制起点和可选择标记)。然后进行琼脂糖凝胶电泳,切除TTMiniV基因组片段(2.9千碱基对)的正确大小的DNA带,并根据制造商的方案使用凝胶提取试剂盒(凯杰公司(Qiagen))从切除的琼脂糖带中凝胶纯化DNA。

[2331] 实例6:指环体的装配和感染

[2332] 该实例说明了使用实例5中所述的合成DNA序列成功地在体外产生感染性指环体。

[2333] 用脂质转染试剂(赛默飞世尔科技公司(Thermo Fisher Scientific))以完整质粒或线性化形式将双链线性化的经凝胶纯化的指环病毒基因组DNA(在实施例5中获得)转染到HEK293T细胞(人胚胎肾细胞系)或A549细胞(人肺癌细胞系)中。将6ug质粒或1.5ug线性化的指环病毒基因组DNA用于在T25烧瓶中转染70%汇合细胞。将缺少指环体中包含的病毒序列的空载体主链用作阴性对照。转染后6小时,将细胞用PBS洗涤两次,并使其在37℃和5%二氧化碳下在新鲜生长培养基中生长。从IDT合成了编码人Ef1α启动子后跟YFP基因的DNA序列。将该DNA序列平末端连接到克隆载体(赛默飞世尔科技公司(Thermo Fisher Scientific))中。将所得的载体用作评估转染效率的对照。转染后72小时,使用细胞成像系统(赛默飞世尔科技公司(Thermo Fisher Scientific))检测YFP。HEK293T和A549细胞的转染效率分别计算为85%和40%(图5)。

[2334] 转染后96小时,收获用指环体转染的293T和A549细胞的上清液。将收获的上清液在4℃下以2000rpm离心10分钟,以去除任何细胞碎片。将每种收获的上清液分别用于感染新的293T和A549细胞,它们在24孔板的孔中70%汇合。在37℃和5%二氧化碳中孵育24小时后,将上清液洗掉,然后两次PBS洗涤,并用新鲜的生长培养基替换。将这些细胞在37度和5%二氧化碳中再孵育48小时后,分别收获细胞用于基因组DNA提取。根据制造商的方案,使用基因组DNA提取试剂盒(赛默飞世尔科技公司(Thermo Fisher Scientific))收集每个样品的基因组DNA。

[2335] 为了证实体外产生的指环体成功感染了293T和A549细胞,使用针对乙型细环病毒或LY2特异性序列具有特异性的引物,将100ng如本文所述收获的基因组DNA用以进行定量聚合酶链反应(qPCR)。根据制造商的方案,使用SYBR绿色试剂(赛默飞世尔科技公司(Thermo Fisher Scientific))进行qPCR。使用针对GAPDH的基因组DNA序列具有特异性的引物的qPCR进行归一化。表42列出了所有使用的引物的序列。

[2336] 表42:

[2337]

	引物序列(5'>3')	
靶标	正向	反向
乙型细环病毒	ATTCGAATGGCTGAGTTTATGC (SEQ ID NO: 690)	CCTTGACTACGGTGGTTTCAC (SEQ ID NO: 693)
LY2TTMiniV株系	CACGAATTAGCCAAGACTGGGCAC (SEQ ID NO: 691)	TGCAGGCATTCGAGGGCTTGTT (SEQ ID NO: 694)
GAPDH	GCTCCCACTCCTGATTTCTG (SEQ ID NO: 692)	TTTAACCCCCTAGTCCCAGG (SEQ ID NO: 695)

[2338] 如图6A、6B、7A和7B所示的qPCR结果所示,体外产生且如本实例所述的指环体具有传染性。

[2339] 实例7:指环体的选择性

[2340] 该实例说明了体外产生的合成的指环体感染多种组织来源的细胞系的能力。

[2341] 将具有感染性TTMiniV指环体(如实例5中所述)的上清液与70％汇合293T、A549、Jurkat(急性T细胞白血病细胞系)、Raji(伯基特氏淋巴瘤B细胞系)和Chang细胞系在37度和5％二氧化碳下在24孔板的孔中一起孵育。感染后24小时,用PBS洗涤细胞两次,然后用新鲜的生长培养基替换。然后将细胞在37度和5％二氧化碳中再次孵育48小时,然后收获进行基因组DNA提取。根据制造商的方案,使用基因组DNA提取试剂盒(赛默飞世尔科技公司(Thermo Fisher Scientific))收集每个样品的基因组DNA。

[2342] 为了证实在先前实例中产生的指环体成功感染了这些细胞系,使用针对乙型细环病毒或LY2特异性序列具有特异性的引物,将100ng如本文所述收获的基因组DNA用以进行定量聚合酶链反应(qPCR)。根据制造商的方案,使用SYBR绿色试剂(赛默飞世尔科技公司(Thermo Fisher Scientific))进行qPCR。使用针对GAPDH的基因组DNA序列具有特异性的引物的qPCR进行归一化。表42列出了所有使用的引物的序列。

[2343] 如图6A-10B所示的qPCR结果所示,在体外产生的指环体不仅具有传染性,而且它们也能够感染多种细胞系,包括上皮细胞、肺组织细胞、肝细胞、癌细胞、淋巴细胞、淋巴母细胞、T细胞、B细胞和肾细胞的实例。还观察到合成的指环体能够感染HepG2细胞(肝细胞系),相对于对照增加超过100倍。

[2344] 实例8:蛋白质结合序列的鉴定和使用

[2345] 该实例描述了指环病毒基因组中推定的蛋白质结合位点,其可用于例如在本文所述的指环体中扩增和包装效应子。在一些情况下,蛋白质结合位点可能能够结合外部蛋白质,例如衣壳蛋白。

[2346] 指环病毒基因组中的两个保守结构域是假定的复制起点:5'UTR保守结构域(5CD)和GC富集结构域(GCR)(de Villiers等人,Journal of Virology[病毒学杂志]2011;Okamoto等人,Virology[病毒学]1999)。在一个实例中,为了证实这些序列是作为DNA复制位点还是作为衣壳包装信号,在包括TTMV-LY2的质粒中对每个区域进行缺失。用pTTMV-LY2Δ5CD或pTTMV-LY2ΔGCR转染A539细胞。将转染的细胞孵育四天,然后从上清液和细胞沉淀中分离出病毒。用病毒感染A549细胞,四天后,从上清液和感染的细胞沉淀中分离出病毒。进行qPCR以定量来自样品的病毒基因组。复制起点的破坏阻止了病毒复制酶扩增病毒DNA,与野生型病毒相比,导致从转染的细胞沉淀中分离出的病毒基因组减少。仍包装了少量病毒,并且可以在转染的上清液和感染的细胞沉淀中找到。在一些实施例中,包装信号的破坏将防止病毒DNA被衣壳蛋白包封。因此,在实施例中,在转染的细胞中仍然存在病毒基因组的扩增,但是在上清液或感染的细胞沉淀中未发现病毒基因组。

[2347] 在另一个实例中,为了表征DNA中的额外复制或包装信号,使用了整个TTMV-LY2基因组上的一系列缺失。在序列的长度上逐步缺失100bp。将包括TTMV-LY2缺失的质粒转染到A549中并如上所述进行测试。在一些实施例中,破坏病毒扩增或包装的缺失将包含潜在的顺式调节结构域。

[2348] 可以将复制和包装信号掺入编码效应子的DNA序列中(例如在指环体的遗传元件中)以诱导扩增和包封。这是在指环体基因组的较大区域的背景下完成的(即将效应子插入基因组的特定位点,或用效应子代替病毒ORF等),或者通过将最小的顺式信号掺入效应子DNA中来完成。在指环体缺乏反式复制或包装因子(例如,复制酶和衣壳蛋白等)的情况下,反式因子由辅助基因提供。辅助基因表达足以诱导扩增和包装的所有蛋白质和RNA,但缺乏

其自身的包装信号。指环体DNA与辅助基因共转染，导致效应子而非辅助基因的扩增和包装。

[2349] 实例9：指环病毒基因组

[2350] 该实例描述了指环病毒基因组中的缺失。

[2351] 在ORF下游但GC富集区上游的TTV-tth8的非编码区(NCR)中缺失了172个核苷酸(nt)(nt 3436至3607)。将56-nt随机序列TTTGTGACACAAGATGGCCGACTTCCTTCCTCTTTAGTCTTCCCCAAAG AAGACAA(SEQ ID NO:696))插入缺失中。产生了pTTV-tth8(3436-3707::56nt)，一个含有改变的TTV-tth8的DNA质粒。使用lipofectamine 2000将2μg双链环状质粒或双链SmaI线性化DNA(产生与细菌骨架元件分离的TTV-tth8基因组片段)以转染到6cm盘中60％汇合度的HEK293或A549细胞中，一式两份。转染后96小时，通过冻融从细胞沉淀和上清液中分离出病毒，在液氮与37℃水浴之间交替进行三次。将来自上清液的病毒用于再感染细胞(HEK293细胞被从HEK293分离出的病毒感染，并且A549细胞被从A549分离出的病毒感染)。感染后72小时，通过冻融从细胞沉淀和上清液中分离出病毒。对所有样品进行qPCR。如下表43所示，在感染的细胞的细胞沉淀和上清液中均观察到TTV-tth8，表明pTTV-tth8(3436-3707::56nt)成功产生了病毒。因此，TTV-tth8能够耐受nt 3436至3707的缺失。

[2352] 表43：HEK293和A549中的TTV-tth8(3436-3707::56nt)感染导致病毒扩增。与不添加质粒或病毒的阴性对照细胞相比，重复实验得到的平均基因组当量。

[2353]

基因组当量/Rx		HEK293 P0	HEK293 P1	A549 P0	A549 P1
TTH8 线性	上清液	2.45E+06	1.02E+03	1.87E+07	1.00E+04
	细胞	2.52E+08	3.92E+05	2.89E+08	7.57E+05
TTH8 环状	上清液	1.69E+06	6.83E+02	5.07E+02	1.05E+04
	细胞	2.00E+08	3.75E+05	2.61E+08	8.36E+05

阴性	
293 空白	1.42E+02
293 阴性	5.08E+02
549 空	1.73E+01
549 阴性	2.08E+01

[2354] 装配了TTMV-LY2的工程化版本，缺失了核苷酸574至1371和1432至2210(1577bp缺失)，并在ORF1的C末端(野生型TTMV-LY2的nt 2609之后)插入了513bp NanoLuc(nLuc)报告基因ORF。如实例17所述，将包括用于工程化TTMV-LY2的DNA序列的质粒(pVL46-015B)转染到A549细胞中，然后分离病毒并用于感染新的A549细胞。在感染的细胞的细胞沉淀和上清液中检测到nLuc发光，表明病毒复制(图11A-11B)。这说明TTMV-LY2可以耐受ORF区域中缺失至少1577bp。

[2355] 为了进一步表征病毒基因组，在TTMV-LY2 DNA中进行了一系列缺失。如前所述，制备了具有nt 574-1371和1432-2210缺失但没有nLuc插入的TTMV-LY2，并测试其病毒复制。对TTMV-LY2Δ574-1371，Δ1432-2210进行了进一步缺失。缺失nt 1372-1431，以产生TTMV-LY2Δ574-2210。此外，将ORF1下游的ORF3序列缺失(Δ2610-2809)。最后，为了测试非编码区的缺失，在整个NCR上依次进行了一系列100bp的缺失。如前所述，测试所有缺失突变体的病毒复制。将导致成功病毒产生的缺失(表明缺失的区域对于病毒复制不是必不可少的)组合在一起，以制成具有更多缺失核苷酸的TTMV-LY2变体。为了鉴定可以用辅助基因扩增的病毒基因组，将破坏病毒复制的每个缺失突变体与携带反式复制和包装元件的辅助基因一起进行测试。通过复制元件的反式表达挽救的缺失表明当从单独的来源提供辅助基因时可以在不阻止病毒形成的情况下缺失的病毒基因组区域。

[2356] 实例10：将各种长度的核苷酸插入指环病毒基因组中

[2357] 该实例描述了将各种长度的DNA序列添加到指环病毒基因组中，其在某些情况下

可用于产生本文所述的指环体。

[2358] 将DNA序列克隆到含有TTV-tth8(GenBank登录号AJ620231.1)和TTMV-LY2(GenBank登录号JX134045.1)的质粒中。在开放阅读框的非编码区(NCR)3'和GC富集区的5'中进行插入:在TTV-tth8的核苷酸3588或TTMV-LY2的核苷酸2843之后。

[2359] 将以下长度的随机DNA序列插入TTV-tth8和TTMV-LY2的NCR中:100个碱基对(bp)、200bp、500bp、1000bp和2000bp。设计这些序列以匹配每个病毒基因组的相对GC含量:约50%GC用于插入TTV-tth8,而约38%GC用于TTMV-LY2。另外,将几个转基因插入到NCR中。这些包括由U6启动子(351bp)驱动的miRNA(例如FF4 miRNA)和由组成型hEF1a启动子(2509bp)驱动的EGFP。

[2360] 如前所述,将包括各种大小的DNA插入物的TTV-tth8和TTMV-LY2变体转染到哺乳动物细胞系中,包括HEK293和A549。从上清液或细胞沉淀中分离出病毒。将分离的病毒用于感染其他细胞。通过定量PCR监测从感染细胞产生病毒。在一些实施例中,病毒的成功产生将指示插入的耐受性。实例11:要递送的示例性货物

[2361] 该实例描述了可以与例如本文所述的指环体例如基于指环病毒的指环体一起递送的示例性类别的核酸和蛋白质有效负载。

[2362] 有效负载的一个实例是用于蛋白质表达的mRNA。目的编码序列是从源病毒(例如指环病毒)天然的病毒启动子转录的,或者是从与有效负载一起作为转基因一部分而引入的启动子转录的。可替代地,在病毒mRNA的开放阅读框内编码mRNA,导致病毒蛋白与目的蛋白之间的融合。当需要时,切割结构域例如2A肽或蛋白酶靶位点可用于将目的蛋白与病毒蛋白分离。

[2363] 非编码RNA(ncRNA)是有效负载的另一个实例。通常使用RNA聚合酶III启动子(例如U6或VA)转录这些RNA。可替代地,使用RNA聚合酶II(例如天然病毒启动子或可调节的合成启动子)转录ncRNA。当从RNA聚合酶II启动子表达时,ncRNA被编码为mRNA外显子、内含子或在聚A信号下游转录的额外RNA的一部分。ncRNA通常被编码为较大RNA分子的一部分,或者使用核酶或核糖内切核酸酶切割。可在指环体基因组中编码为货物的ncRNA包括微小RNA(miRNA)、小干扰RNA(siRNA)、短发夹RNA(shRNA)、反义RNA、miRNA海绵、长非编码RNA(lncRNA)和指导RNA(gRNA)。

[2364] DNA可以用作功能元件,而无需RNA转录。例如,DNA可用作同源重组的模板。在另一个实例中,可以使用蛋白质结合DNA序列来驱动目的蛋白质包装到衣壳中(例如指环体的蛋白质外部)。为了同源重组,将与人基因组DNA同源的区域编码到载体DNA中以充当同源臂。重组可以由靶向内切核酸酶(例如具有gRNA的Cas9,或锌指核酸酶)驱动,所述靶向内切核酸酶可以从载体或从单独的来源表达。在细胞内部,单链DNA基因组被转换为双链DNA,然后作为基因组DNA断裂位点的同源重组的模板。为了募集目的蛋白质,可以将蛋白质结合序列编码到指环体DNA中。目的DNA结合蛋白或与DNA结合蛋白(例如Ga14)融合的目的蛋白与指环体DNA结合。当指环体DNA被衣壳蛋白包封时,DNA结合蛋白也被包封,并且可以与指环体一起递送到细胞中。

[2365] 实例12:示例性有效负载整合位点

[2366] 该实例描述了核酸有效负载可插入其中的TTV-tth8(GenBank登录号AJ620231.1)和TTMV-LY2(GenBank登录号JX134045)的基因组中的示例性位点。

[2367] 可以采用几种策略以插入TTV-tth8(核苷酸336至3015)和TTMV-LY2(核苷酸424至2812)的开放阅读框(ORF)区。在一个实例中,为了标记病毒蛋白或产生融合蛋白,将有效负载按读框插入目的特定ORF内。替代地,缺失部分或全部ORF区,这可能会或可能不会破坏病毒蛋白功能。然后将有效负载插入缺失区域。另外,TTV-tth8(核苷酸716和2362之间)或TTMV-LY2(核苷酸724和2273之间)的ORF1中的高变结构域(HVD)可以用作插入位点。在一些情况下,HVD中的插入或核苷酸替换可能会被更好地耐受和/或在较小程度上破坏病毒功能。

[2368] 替代地,将有效负载插入与TTV-tth8或TTMV-LY2的非编码区(NCR)相当的载体区中。特别是,在TATA盒上游的5'NCR、5'非翻译区(UTR)、聚A信号下游和GC富集区上游的3'NCR中进行插入。另外,在TTV-tth8的miRNA区(核苷酸3429至3506)中插入。对于5'NCR区,在TATA盒上游进行插入(在TTV-tth8中的核苷酸1和82之间,在TTMV-LY2中的核苷酸1和236之间)。在一些实施例中,以相反方向插入转基因以减少启动子干扰。对于5'UTR,在转录起始位点的下游(TTV-tth8中的核苷酸111,TTMV-LY2中的核苷酸267)和ORF2起始密码子的上游(TTV-tth8中的核苷酸336,TTMV-LY2中的核苷酸421)进行插入。5'UTR插入会在5'UTR中添加或替换核苷酸。如实例10中所述,在GC富集区上游进行3'NCR插入,特别是在TTV-tth8中的核苷酸3588或TTMV-LY2中的核苷酸2843之后。将TTV-tth8的miRNA替换为可替代的天然或合成miRNA发夹。

[2369] 实例13:定义的类别的指环病毒和其保守区

[2370] 人中存在三种指环病毒属:甲型细环病毒(细环病毒,TTV)、乙型细环病毒(细环中型病毒,TTMDV)和丙型细环病毒(细环小型病毒,TTMV)。甲型细环病毒包括至少五个(例如七个)得到充分支持的系统发育进化分枝(图11C)。可以预期的是,这些指环病毒中的任一种都可以用作来源病毒(例如,病毒DNA序列的来源),用于产生本文所述的指环体。

[2371] 在这些序列中,在5'UTR结构域(约75%保守性)和GC富集结构域(大于100个碱基对,GC含量大于70%,约70%保守性)中发现了最高保守性。另外,序列中的高变结构域(HVD)具有非常低的保守性(约30%保守性)。所有指环病毒也都包含其中所有三个阅读框均开放的区域。在一些情况下,5'UTR或GC富集区可能充当复制起点。

[2372] 本文还提供了来自TTV进化分枝以及TTMDV和TTMV中每一个的代表性病毒的示例性序列,其用保守区注释(参见例如表A1-A12、B1-B5、C1-C5或1-18)。

[2373] 实例14:复制缺陷型指环体和辅助病毒

[2374] 为了复制和包装指环体,可以反式提供一些元件。这些包括指导或支持DNA复制或包装的蛋白质或非编码RNA。在某些情况下,反式元件可以从替代指环体的来源中获得,例如辅助病毒、质粒或从细胞基因组中获得。

[2375] 通常以顺式提供其他元件。这些元件可以是例如指环体DNA中的序列或结构,其充当复制起点(例如,允许指环体DNA的扩增)或包装信号(例如,结合蛋白以将基因组负载到衣壳中)。通常,复制缺陷型病毒或指环体将丢失这些元件中的一个或多个,使得即使反式提供其他元件,DNA也不能包装成传染性病毒粒子或指环体。

[2376] 复制缺陷型病毒可用作辅助病毒,例如,用于控制指环体(例如复制缺陷型或包装缺陷型指环体)在同一细胞中的复制。在某些情况下,辅助病毒将缺少顺式复制或包装元件,但会表达反式元件,例如蛋白质和非编码RNA。通常,治疗性指环体将缺少这些反式元件

中的一些或所有,因此将无法单独复制,但会保留顺式元件。当共转染/感染到细胞中时,复制缺陷型辅助病毒将驱动指环体的扩增和包装。因此,收集的包装颗粒将仅由治疗性指环体组成,而不会感染辅助病毒。

[2377] 为了开发复制缺陷型指环体,将去除指环病毒的非编码区域中的保守元件。特别是,将分别或一起测试保守的5'UTR结构域和GC富集结构域的缺失。两种元件被认为对于病毒复制或包装是重要的。此外,将在整个非编码区中进行一系列缺失,以鉴定以前未知的目的区。

[2378] 复制元件的成功缺失将导致细胞内的指环体DNA扩增减少,例如,通过qPCR所测量,但将支持某些传染性指环体产生,例如,通过对受感染的细胞的测定所监测,所述测定可包括qPCR、蛋白质印迹、荧光测定或发光测定中任一种或全部。包装元件的成功缺失不会破坏指环体DNA扩增,因此通过qPCR在转染的细胞中可观察到指环体DNA的增加。但是,指环体基因组不会被包封,因此不会观察到传染性指环体产生。

[2379] 实例15:具有复制能力的指环体的制造方法

[2380] 此实例描述了一种用于回收和扩大具有复制能力的指环体的产生的方法。当它们在基因组中编码在细胞中复制所需的所有必需遗传元件和ORF时,指环体具有复制能力。由于这些指环体的复制没有缺陷,因此它们不需要反式提供的互补活性。但是,它们可能需要辅助活性,例如转录增强子(例如丁酸钠)或病毒转录因子(例如腺病毒E1、E2、E4、VA;HSV Vp16和立即早期蛋白)。

[2381] 在该实例中,将编码线性或环状形式的合成指环体的全序列的双链DNA通过化学转染引入T75烧瓶中的5E+05粘附哺乳动物细胞中,或通过电穿孔引入5E+05悬浮细胞中。在最佳时间段后(例如,转染后3-7天),通过将细胞刮入上清液培养基中来收集细胞和上清液。加入温和洗涤剂(如胆盐)至最终浓度为0.5%,并在37℃下孵育30分钟。加入氯化钙和氯化镁至最终浓度分别为0.5mM和2.5mM。加入内切核酸酶(例如DNAse I,Benzonase),并在25℃-37℃下孵育0.5-4小时。将指环体悬浮液在4℃下以1000xg离心10分钟。将澄清的上清液转移至新试管中,并用冷冻保护剂缓冲液(也称为稳定化缓冲液)以1:1稀释,并根据需要在-80℃下保存。这将产生第0代的指环体(P0)。为了使洗涤剂的浓度低于用于培养细胞的安全极限,根据指环体效价,将这种接种物在无血清培养基(SFM)中稀释至少100倍或更多倍。

[2382] 将T225烧瓶中新鲜的哺乳动物细胞单层用足以覆盖培养物表面的最小体积覆盖,并在37℃和5%二氧化碳下以温和摇动孵育90分钟。用于该步骤的哺乳动物细胞可以与用于P0回收的细胞类型相同或不同。孵育后,将接种物替换为40ml无血清、无动物源的培养基。将细胞在37℃和5%二氧化碳下孵育3-7天。加入4ml先前使用的相同温和洗涤剂的10X溶液,以使最终洗涤剂浓度达到0.5%,然后将混合物在37℃下以温和搅拌孵育30分钟。加入内切核酸酶并在25℃-37℃下孵育0.5-4小时。然后收集培养基,并在4℃下以1000xg离心10分钟。将澄清的上清液与40ml稳定缓冲液混合,并在-80℃下保存。这将产生种子储备物或第一代的指环体(P1)。

[2383] 根据储备物的效价,将其在SFM中稀释不少于100倍,并添加到在所需大小的多层烧瓶中生长的细胞中。感染复数(MOI)和孵育时间已在较小的规模上进行了优化,以确保最大的指环体生产。收获后,可根据需要纯化和浓缩指环体。图12中提供了显示工作流程的示

意图,例如,本实例中所述。

[2384] 实例16:复制缺陷型指环体的制造方法

[2385] 此实例描述了一种用于回收和扩大复制缺陷型指环体的产生的方法。

[2386] 通过缺失复制中涉及的一个或多个ORF(例如,ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3和/或ORF2t/3),可以使指环体成为复制缺陷型。复制缺陷型指环体可以在互补细胞系中生长。这种细胞系组成型表达促进指环体生长但在指环体的基因组中丢失或无功能的组分。

[2387] 在一个实例中,将参与指环体繁殖的任何一个或多个ORF的一个或多个序列克隆到适合于产生编码选择标记的稳定细胞系的慢病毒表达系统中,并如本文所述产生慢病毒载体。使能够支持指环体繁殖的哺乳动物细胞系感染该慢病毒载体,并经受选择标记(例如嘌呤霉素或任何其他抗生素)的选择压力,以选择已稳定整合了克隆的ORF的细胞群。一旦对这种细胞系进行了表征并证明可以互补工程化的指环体中的缺陷,从而支持此类指环体的生长和繁殖,就可以将其扩增并储存在低温储存库中。在这些细胞的扩增和维持过程中,将选择抗生素添加到培养基中以维持选择压力。一旦将指环体引入这些细胞中,就可以抑制选择抗生素。

[2388] 一旦建立了该细胞系,就进行复制缺陷型指环体的生长和产生,例如,实例15中所述。

[2389] 实例17:使用悬浮细胞生产指环体

[2390] 该实例描述了悬浮细胞中产生指环体。

[2391] 在该实例中,使适于在悬浮条件下生长的A549或293T生产细胞系在无动物组分和无抗生素的悬浮培养基(赛默飞世尔科技公司(Thermo Fisher Scientific))中于WAVE生物反应器袋中于37度和5%二氧化碳下生长。使用lipofectamine 2000(赛默飞世尔科技公司(Thermo Fisher Scientific))在目前的良好生产规范(cGMP)下,将以$1x10^6$个活细胞/mL接种的这些细胞用含有指环体序列的质粒以及适合或需要包装指环体的任何互补质粒(例如在复制缺陷型指环体的情况下,例如,如实例16中所述)进行转染。在一些情况下,互补质粒可以编码已经从指环体基因组(例如,基于例如本文所述例如指环病毒基因组的病毒基因组的指环体基因组)中缺失但是用于或需要用于复制和包装指环体的病毒蛋白。转染的细胞在WAVE生物反应器袋中生长,并在以下时间点收获上清液:转染后48小时、72小时和96小时。使用离心将上清液与每个样品的细胞沉淀分离。然后,使用离子交换色谱法从收获的上清液和裂解的细胞沉淀中纯化包装的指环体颗粒。

[2392] 可以例如通过使用纯化的制备物的少量等分试样以使用病毒基因组提取试剂盒(凯杰公司(Qiagen))来收获指环体基因组,然后使用靶向指环体DNA序列的引物和探针进行qPCR,例如实例18中所述,来测定指环体的纯化制备物中的基因组当量。

[2393] 可通过对纯化的制备物进行连续稀释以感染新的A549细胞来定量纯化的制备物中的指环体的感染性。转染后72小时,收获这些细胞,然后使用对指环体DNA序列具有特异性的引物和探针对基因组DNA进行qPCR测定。

[2394] 实例18:通过qPCR对指环体基因组当量进行定量

[2395] 该实例说明了用以定量指环体的基于水解探针的定量PCR测定的开发。使用Geneious软件,基于选择的TTV(登录号AJ620231.1)和TTMV(登录号JX134045.1)的基因组

序列设计引物和探针组，并进行最终用户优化。引物序列显示在下表44中。

[2396] 表44：正向和反向引物和水解探针的序列，用于通过定量PCR定量TTMV和TTV基因组当量。

[2397]

TTMV		SEQ ID NO:
正向引物	5'-GAAGCCCACCAAAAGCAATT-3'	697
反向引物	5'-AGTTCCCGTGTCTATAGTCGA-3'	698
探针	5'-ACTTCGTTACAGAGTCCAGGGG-3'	699

[2398]

TTV		
正向引物	5′-AGCAACAGGTAATGGAGGAC-3′	700
反向引物	5′-TGGAAGCTGGGGTCTTTAAC-3′	701
探针	5′-TCTACCTTAGGTGCAAAGGGCC-3′	702

[2399] 作为开发方法的第一步，使用TTV和TTMV引物以及SYBR绿色化学进行qPCR，以检查引物特异性。图13显示了每个引物对的一个不同的扩增峰。

[2400] 订购的水解探针在5’末端标记有荧光团6FAM，在3’末端标记有小沟结合非荧光猝灭剂(MGBNFQ)。然后使用两种不同的商业预混物，使用纯化的质粒DNA作为标准曲线的组成并增加引物的浓度来评估新引物和探针的PCR效率。通过使用纯化的质粒建立标准曲线，所述纯化的质粒包含用于不同引物-探针组的靶序列。进行了七次十倍系列稀释，以实现7个对数内的线性范围和每20ul反应15个拷贝的定量下限。2号预混物能够产生90％-110％的PCR效率，这是定量PCR可接受的值(图14)。qPCR的所有引物均购自IDT。与荧光团6FAM和小沟结合非荧光猝灭剂(MGBNFQ)缀合的水解探针以及所有qPCR预混物均购自赛默飞世尔(Thermo Fisher)。示例性扩增图如图15所示。

[2401] 使用这些引物-探针组和试剂，对基因组当量(GEq)/ml指环体储备物进行定量。线性范围为每20ul反应1.5E+07-15GEq，然后用于计算GEq/ml，如图16A-16B所示。浓度高于线性范围的样品可以根据需要进行稀释。

[2402] 实例19：利用指环体在小鼠中表达外源蛋白

[2403] 此实例描述了指环体的用法，其中将细小环病毒(TTMV)基因组工程化以在小鼠中表达萤火虫萤光素酶蛋白。

[2404] 将编码工程化TTMV的DNA序列的质粒通过化学转染引入A549细胞(人肺癌细胞系)中，所述工程化TTMV编码萤火虫萤光素酶基因。将18ug质粒DNA用于转染10cm组织培养板中70％汇合的细胞。将缺少TTMV序列的空载体主链用作阴性对照。转染后5小时，将细胞用PBS洗涤两次，并使其在37℃和5％二氧化碳下于新鲜生长培养基中生长。

[2405] 转染后96小时，收获转染的A549细胞及其上清液。将收获的物质在37℃下用0.5％脱氧胆酸盐(体积重量)处理1小时，然后进行内切核酸酶处理。使用离子交换色谱从该裂解物中纯化指环体颗粒。为了测定指环体浓度，将指环体储备物的样品运行通过病毒DNA纯化试剂盒，并使用靶向指环体DNA序列的引物和探针通过qPCR测定每ml的基因组当量。

[2406] 通过多种注射途径(例如静脉内、腹膜内、皮下、肌肉内)在8-10周龄的小鼠中进行1x磷酸盐缓冲液中的指环体的基因组当量的剂量范围。在注射后3、7、10和15天，对每只动

物进行腹侧和背侧生物发光成像。根据制造商的规程，在指定的时间点通过向每只动物腹膜内添加萤光素酶底物(Perkin-Elmer)，然后进行活体成像来进行成像。

[2407] 实例20:基因组比对，以测定指环体DNA是否整合到宿主基因组中

[2408] 此实例描述了通过检查细环病毒(TTV)是否已整合到人基因组中来测定指环体DNA是否可以整合到宿主基因组中而进行的计算分析。

[2409] 使用基本局部比对检索工具(BLAST)将来自五个示例性甲型细环病毒进化分枝中每一个的代表性TTV序列的全基因组与人基因组序列进行比对，所述工具会发现序列之间的局部相似性区域。分析了表45中所示的代表性TTV序列：

[2410] 表45:代表性TTV序列

[2411]

TTV 进化分枝	NCBI 登录号
进化分枝 A	AB064597.1
进化分枝 B	AB028669.1
进化分枝 C	AJ20231.1
进化分枝 D	AF122914.3

[2412]

进化分枝 E	AF298585.1

[2413] 没有发现来自对齐的TTV的序列与人基因组具有任何显著相似性，这表明TTV尚未整合到人基因组中。

[2414] 实例21:指环体整合到宿主基因组中的评估

[2415] 在此实例中，将A549细胞(人肺癌细胞系)和HEK293T细胞(人胚胎肾细胞系)以5、10、30或50的MOI感染指环体颗粒或AAV颗粒。感染后5小时，将细胞用PBS洗涤，更换为新鲜的生长培养基。然后使细胞在37度和5％二氧化碳下生长。感染后五天，收获细胞，并使用基因组DNA提取试剂盒(凯杰公司(Qiagen))将其处理以收获基因组DNA。还从未感染的细胞中收获基因组DNA(阴性对照)。根据制造商规程，使用Nextera DNA文库制备试剂盒(Illumina)为这些收获的DNA制备全基因组测序文库。根据制造商规程，使用NextSeq 550系统(Illumina)对DNA文库进行测序。将测序数据装配到参考基因组中，并进行分析以寻找指环体或AAV基因组与宿主基因组之间的连接。在检测到连接的情况下，将在原始基因组DNA样品中对其进行验证，然后再通过PCR进行测序文库制备。设计引物来扩增包含连接和连接周围的区域。通过qPCR定量样品中的连接数(代表整合事件)和指环体拷贝总数，来测定指环体整合到宿主基因组中的频率。可以将该比率与AAV的比率进行比较。

[2416] 实例22:表达外源性微小RNA序列的指环体的功能作用

[2417] 该实例展示了使用天然启动子从指环体基因组中成功表达外源miRNA(miR-625)。

[2418] 将500ng以下质粒DNA转染到24孔板中HEK293T细胞的60％汇合孔中：

[2419] i)空质粒骨架

[2420] ii)含有TTV-tth8基因组的质粒，其中内源性miRNA被敲除(KO)

[2421] iii)TTV-tth8，其中内源性miRNA被非靶向乱序miRNA取代

[2422] iv)TTV-tth8，其中内源性miRNA序列被编码miR-625的miRNA替换

[2423] 转染后72小时，使用Qiagen miRNeasy试剂盒从转染细胞中收集总miRNA，然后使用miRNA Script RT II试剂盒进行逆转录。使用应特异性检测miRNA-625或RNU6小RNA的引

物对逆转录的DNA进行定量PCR。RNU6小RNA用作管家基因，数据绘制在图17中，作为相对于空载体的倍数变化。如图17所示，miR-625指环体导致miR-625表达增加约100倍，而对于空载体、miR-敲除(KO)和乱序miR，未检测到信号。

[2424] 实例23：表达外源性非编码RNA的指环体的制备和产生

[2425] 该实例描述了表达外源性小非编码RNA的指环体的合成和产生。

[2426] 合成了来自TTV的tth8菌株的DNA序列(Jelcic等人，Journal of Virology[病毒学杂志]，2004)，并将其克隆到含有细菌复制起点和细菌抗生素抗性基因的载体中。在该载体中，编码TTV miRNA发夹的DNA序列被替换为编码外源性小非编码RNA(例如miRNA或shRNA)的DNA序列。然后，将工程化的构建体转化到电感受态细菌中，然后根据制造商的规程使用质粒纯化试剂盒进行质粒分离。

[2427] 将编码外源性小非编码RNA的指环体DNA转染到真核生产细胞系中，以生产指环体颗粒。在转染后的不同时间点，收获含有指环体颗粒的转染细胞的上清液。将来自过滤的上清液或纯化后的指环体颗粒用于下游应用，例如本文所述。

[2428] 实例24：指环病毒进化分枝中的保守性

[2429] 本实例描述了甲型细环病毒属中七个进化分枝的鉴定。这些进化分枝之间的代表性序列在整个序列中显示54.7％成对同一性(图18)。在开放阅读框中，成对同一性最低(约48.8％)，在非编码区则较高(在5'NCR中为69.1％，在3'NCR中为74.6％)(图18)。这表明非编码区的DNA序列或结构在病毒复制中起重要作用。

[2430] 还比较了甲型细环病毒中推定蛋白质的氨基酸序列。DNA序列显示约47％至50％成对同一性，而氨基酸序列显示约32％至38％成对同一性(图19)。有趣的是，来自甲型细环病毒进化分枝的代表性序列能够在体内成功复制并在人群中观察到。这表明指环病毒蛋白的氨基酸序列可以广泛变化，同时保留诸如复制和包装的功能。

[2431] 发现指环病毒在非编码区中具有局部高度保守的区域。在启动子下游的区域是71-bp 5'UTR保守结构域，其在七个甲型细环病毒进化分枝中具有95.2％成对同一性(图20)。在甲型细环病毒的3'非编码区的开放阅读框的下游，存在其中代表性序列之间具有大量成对同一性的区域。在这个3'保守非编码区的3'末端附近是高度保守的序列。指环病毒还包括GC含量超过70％的GC富集区，其在它们对齐的区域中表现出75.4％的成对同一性(图21)。

[2432] 实例25：从指环体表达内源性miRNA和内源性miRNA的缺失

[2433] 在一个实例中，使用包含经修饰的TTV-tth8基因组的指环体来感染培养中的Raji B细胞，其中TTV-tth8基因组被修饰为缺失GC富集区，如实例27所述。这些指环体包含编码TTV-tth8指环体病毒的内源有效负载的序列，其是一种靶向编码n-myc相互作用蛋白(NMI)的mRNA的miRNA，并通过引入包含指环病毒基因组的质粒进入宿主细胞。NMI在JAK/STAT途径的下游运行，以调节各种细胞内信号的转录，包括干扰素刺激的基因、增殖和生长基因以及炎症反应的介质。如图22所示，在靶Raji B细胞中检测到病毒基因组。与对照细胞相比，在靶Raji B细胞中也观察到了NMI的成功敲低(图23)。与对照细胞相比，包含针对NMI的miRNA的指环体诱导NMI蛋白水平降低75％以上。该实例说明具有天然指环病毒miRNA的指环体可以敲低宿主细胞中的靶分子。

[2434] 在另一个实例中，缺失了基于指环病毒的指环体的内源性miRNA。然后将所得的指

环体(ΔmiR)与宿主细胞一起孵育。然后将ΔmiR指环体遗传元件的基因组等效物与保留内源性miRNA的相应指环体的基因组等效物进行比较。如图24所示,在细胞中以与仍然存在内源性miRNA的指环体基因组所观察到的水平相当的水平检测到缺失了内源性miRNA的指环体基因组。该实例表明,基于指环病毒的指环体的内源性miRNA可以发生突变或完全缺失,并且仍然可以在靶细胞中检测到该指环体基因组。

[2435] 实例26:指环病毒ORF的定位

[2436] 该实例描述了指环病毒的各种推定ORF的新功能。在该实例中,推定的开放阅读框(ORF)序列被设计在每个ORF的N末端的标记蛋白(即纳米萤光素酶)的下游。通过化学转染将每个ORF-nLuc质粒引入12孔板中的5E+05贴壁细胞(Vero或HEK293T)或通过电穿孔引入悬浮的5E+05细胞。在最佳时间段后(例如,转染后3-7天),用PBS中的4%多聚甲醛(赛默飞世尔公司cat#28908)固定细胞,并用0.5%Triton X-100进行透化,并用兔多克隆抗nLuc抗体(普洛麦格公司(Promega)的赠品),然后是山羊抗兔Alexa488(赛默飞世尔公司cat#A-11008)偶联二抗对nLuc染色。细胞核用DAPI(赛默飞世尔公司Cat#D3571)染色。染色的细胞在带有20X物镜和单色Axiocam 506相机的Zeiss AxioVert A1上可视化,用于标记的蛋白的细胞定位。

[2437] 如图25A-25B所示,在Vero细胞和HEK293细胞中观察到ORF2定位于细胞质,观察到ORF1/1定位于细胞核。图25C显示了ORF1/2和ORF2/2的定位。

[2438] 实例27:指环体开发所需区域的表征

[2439] 该实例描述了指环病毒基因组中的缺失,以帮助表征足以复制病毒和指环病毒生产的基因组部分。在ORF(nt 3016至3753)下游的TTV-tth8的非编码区(NCR)中进行了一系列缺失。36个核苷酸(nt)的序列(CGCGCTGCGCGCGCCGCCCAGTAGGGGGAGCCATGC(SEQ ID NO:160))从GC区域中缺失(标记为Δ36nt(GC))。此外,从3'NCR中缺失了78-nt前微小RNA序列(CCGCCATCTTAAGTAGTTGAGGCGGACGGTGGCGTGAGTTCAAAGGTCACCATCAGCCACACCTACTCAAAATGGTGG(SEQ ID NO:161))(标记为Δ36nt(GC)ΔmiR)。最后,在Δ36nt(GC)的3'NCR缺失额外的171个nt(CTTAAGTAGTTGAGGCGGACGGTGGCGTGAGTTCAAAGGTCACCATCAGCCACACCTACTCAAAATGGTGGACAATTTCTTCCGGGTCAAAGGTTACAGCCGCCATGTTAAAACACGTGACGTATGACGTCACGGCCGCCATTTTGTGACACAAGATGGCCGACTTCCTTCC(SEQ ID NO:162))并且标记为Δ3'NCR(图26)。将上述分别描述的2μg环状pTTV-tth8(WT)、pTTV-tth8(Δ36nt(GC))、pTTV-tth8(Δ36nt(GC)ΔmiR)、包含改变的3'NCR TTV-tth8的pTTV-tth8(Δ3'NCR)DNA质粒使用lipofectamine 2000转染到12孔板中的60%汇合度的HEK293中,一式三份。转染后48小时,收获细胞沉淀并裂解以分离mRNA转录物(RNeasy,凯杰公司cat#74104)并转化为cDNA(高容量cDNA逆转录试剂盒,赛默飞世尔公司,cat#4368814)。对所有样品进行qPCR,测量每个缺失情况下的病毒转录物表达,并标准化为GAPDH的内部对照mRNA。

[2440] 如图27A-27D所示,所有三种缺失突变体在体外都显著抑制病毒转录物表达。因此,TTV-tth8的3'NCR是生产用于表达转基因的指环体所必需的。

[2441] TTV株tth8,GeneBank登录号AJ620231.1,以全基因组序列保藏。然而,在GC富集区,有一段36个核苷酸被注释为通用N。该区域在TTV株中高度保守,因此可能对这些病毒的生物学很重要。将数百个TTV菌株的DNA序列进行计算比对,并用于生成这36个核苷酸的强共有序列(CGCGCTGCGCGCGCCGCCCAGTAGGGGGAGCCATGC(SEQ ID NO:160))。TTV-tth8基因组

序列在本文中称为“野生型”序列，因此该共有序列被插入以替代公开可用的TTV-tth8序列中列出的36个N的区段。

[2442] 实例28:外源蛋白质的指环体体内递送

[2443] 本实例展示了指环体在施用后的体内效应子功能(例如蛋白质的表达)。

[2444] 制备包含编码纳米萤光素酶(nLuc)的转基因的指环体(图28A-28B)。简而言之，将含有TTMV-LY2非编码区和nLuc表达盒的双链DNA质粒与编码完整TTMV-LY2基因组的双链DNA质粒一起转染到HEK293T细胞中，作为反式复制和包装因子。转染后，培养细胞以允许产生指环体，并通过核酸酶处理、超滤/渗滤和无菌过滤收获和富集指环体材料。使用以下转染另外的HEK293T细胞:包含nLuc表达盒和TTMV-LY2 ORF转染盒的非复制DNA质粒，但缺乏复制和包装所必需的非编码结构域，以充当“非病毒”阴性对照。非病毒样品是按照与指环体材料相同的方案制备的。

[2445] 对三只健康小鼠进行肌内注射，并在九天内通过IVIS Lumina成像(布鲁克公司(Bruker))进行监测(图29A)。作为非病毒对照，将非复制型制剂施用于另外三只小鼠(图29B)。在第0天向左后腿注射25μL的指环体或非病毒制剂，并在第4天再注射右后腿(参见图29A和B中的箭头)。在IVIS成像9天后，与非病毒制剂(图29B)相比，在注射了指环体制剂(图29A)的小鼠中观察到更多的nLuc发光信号，这与体内指环体转导后的反式基因表达一致。

[2446] 实例29:指环病毒中前体miRNA(前mIR)的鉴定

[2447] 该实例描述了用以鉴定各种指环病毒编码的新颖前体miRNA的各种计算和实验方法。

[2448] 计算方法

[2449] 指环病毒株在核苷酸序列水平上彼此非常不同。然而，指环病毒株，尤其是同一进化分枝内的那些，在启动子、GC富集区、非编码区和编码区等各种成分的基因组组织方面可以表现出显著的相似性(参见例如图29D)。本文描述了一种方法，其中通过与已通过实验验证的前miR序列的指环病毒株进行比对来预测各种指环病毒株(其前miR序列未知)的前miR序列。

[2450] 简而言之，挖掘来自细胞系和各种人样品的小RNA的各种公开可用的小RNA测序数据集，以发现由各种指环病毒株编码的新颖前miR序列。基于结构预测或机器学习分类的公开可用的计算工具和算法，例如mFold程序、miRANDA算法、miRScan、miRanalyzer、miRDeep(https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1559940/,https://www.frontiersin.org/articles/10.3389/fbioe.2015.00007/full)用于预测由各种指环病毒编码的新颖miRNA。然后使用针对特异性miRNA序列设计的探针的RNA印迹和/或使用针对miRNA的特异性引物的RT-qPCR来确认、验证和定量新颖miRNA的表达。

[2451] 实验方法

[2452] 在一个实例中，对感染了指环病毒的人组织或血液样品进行高通量小RNA测序，以发现新颖的指环病毒编码的前miRNA。为此，需要从匀浆的人组织样品或人血液样品中采集RNA。使用Illumina试剂盒和测序平台制备和测序小RNA文库。测序读段在BaseSpace Sequence Hub(依诺米那公司(Illumina))上存储、比对和分析。

[2453] 在第二实例中，对用以下条件处理的各种细胞系进行高通量小RNA测序，以发现由指环病毒编码的新颖前miRNA:(a)感染了天然存在的指环病毒的细胞系，转染了体外合成

的指环病毒基因组的细胞系，以及(c)感染了使用合成的基因组在体外包装的指环病毒的细胞系。使用针对特异性miRNA序列设计的探针的RNA印迹和/或使用针对miRNA的特异性引物的RT-qPCR来确认、验证和定量新颖miRNA的表达。

[2454] 实例30：确定指环病毒前miR的内源性靶

[2455] 该实例描述了分析以确定由各种指环病毒株编码的前miR的内源性靶和潜在的治疗相关靶途径。由各种指环病毒编码的经计算预测和/或实验验证的单个前miRNA序列被克隆到慢病毒载体中，由U6启动子驱动。由U6启动子驱动的非靶向乱序miRNA序列也以类似的方式克隆，用作对照。慢病毒质粒被克隆，使得当包装时，其基因组将包含(i)由U6启动子驱动的前miRNA序列，(ii)由SV40启动子驱动的嘌呤霉素抗性基因，以及(iii)由CMV启动子驱动的绿色荧光蛋白(GFP)基因。这些慢病毒质粒中的每一个都与慢病毒辅助质粒一起单独共转染到HEK-293T细胞中以包装病毒。转染后六小时，吸出转染细胞的培养基，用PBS洗涤一次并更换为新鲜培养基。转染后72小时收集含有慢病毒的培养基。培养基通过0.4um过滤器过滤以去除任何细胞，然后用于感染目的细胞类型，如HeLa、Raji和THP1，一式三份。通过在感染后3天开始用嘌呤霉素处理来选择含有整合的慢病毒基因组的细胞。使用RNA提取试剂盒(凯杰公司)从稳定选择的细胞系中收获RNA，然后使用逆转录酶试剂盒(赛默飞世尔科技公司)逆转录为cDNA。处理cDNA样品以生成带索引的短读段文库。使用Illumina测序平台，经独特索引的短读段文库被多路复用测序，每个样品产生约2000万个读段。测序读段使用BaseSpace Sequence Hub(依诺米那公司(Illumina))进行存储、比对和分析。通过比较表达候选前miR的细胞系中的基因表达与表达乱序前miR的细胞系中的基因表达来确定每个单独的候选前miR的靶。进行Ingenuity Pathway分析以测试前miRNA是否靶向特定通路，尤其是治疗相关通路。该实例中描述的工作流程示意图如图30所示。

[2456] 实例31：制造编码天然指环病毒前miR的指环体

[2457] 该实例描述了包装表达天然指环病毒前miR的指环体的复制或非复制形式的过程。

[2458] 合成的非复制形式的指环体的基因组包含以下组分：(i)复制起点，(ii)编码指环病毒前miRNA的序列，(iii)RNA聚合酶III，例如U6或H1，驱动前miRNA的表达，以及(iv)包装信号。该基因组通过转染到辅助细胞系中进行包装，所述辅助细胞系稳定表达病毒包装所需的所有蛋白质。如本文所述，转染后7天收获转染的细胞并进行加工以制备指环体制剂。如本文所述，通过进行qPCR确定指环体制剂的基因组等效效价。然后将适当剂量的指环体制剂用于下游应用。

[2459] 可以合成指环体的复制形式的基因组，例如，通过产生天然指环病毒，除了使用外源启动子如U6或组织特异性启动子操纵前miRNA序列的表达。基因组通过转染到HEK-293T细胞中进行包装。如本文所述，转染后7天收获转染的细胞并进行加工以制备指环体制剂。如本文所述，通过进行qPCR确定指环体制剂的基因组等效效价。将适当剂量的指环体制剂用于下游应用。

[2460] 实例32：在体外细胞培养模型中利用指环病毒前miR作为肿瘤抑制因子

[2461] 该实例描述了研究以确认从例如如实例29中所述的分析中鉴定为肿瘤抑制性的候选前miR的表型作用。

[2462] 基于如实例29中所述的分析鉴定具有肿瘤抑制作用的候选前miRNA。如实例31中

所述制备编码这些候选前miRNA以及乱序前miRNA的指环体的复制形式的指环体制剂。来自NCI-60癌细胞系组的癌细胞系接种在96孔板中。当30％汇合时，这些细胞系用包含候选前miR或乱序前miR的指环体以每个细胞五个基因组当量的剂量处理。感染后五小时吸出含有指环体的培养基，然后用PBS洗涤两次，并更换为新鲜培养基。处理后三天对处理过的细胞进行Alamar blue测定，以确定哪些前miR抑制癌细胞系的增殖。

[2463] 实例33：利用指环病毒前miR作为体内肿瘤抑制因子

[2464] 该实例描述了体内实验，以确认来自如实例32中所述的体外分析的筛选的候选肿瘤抑制性指环病毒前miR和癌细胞系的肿瘤抑制作用。

[2465] 异种移植物是通过将来自实例32中描述的分析的经筛选的癌细胞系连同基质胶一起皮下注射到无胸腺小鼠的侧腹中来产生的。一旦异种移植肿瘤变得可触及，就会进行$3x10^6$基因组当量的编码肿瘤抑制性前miRNA或乱序前miRNA的指环体的局部肿瘤注射。通过三周的常规肿瘤生长测量，在实验结束时异种移植肿瘤的肿瘤重量测量以及通过BrdU掺入测定确定指环体注射对肿瘤生长的影响。

[2466] 实例34：指环病毒基因组的串联拷贝

[2467] 该实例描述了基于质粒的表达载体，其包含单个环病毒基因组的两个拷贝，串联排列，使得上游基因组的GC富集区靠近下游基因组的5’区域(图31A)。

[2468] 指环病毒通过滚环复制，其中复制酶(Rep)蛋白在复制起点与基因组结合并启动围绕该环的DNA合成。对于包含在质粒骨架中的指环病毒基因组，这需要复制完整的质粒长度，其比天然病毒基因组长，或者质粒的重组导致包含具有最少骨架的基因组的较小环。因此，从质粒复制病毒可能效率低下。为了提高病毒基因组复制效率，质粒被工程化为具有TTV-tth8和TTMV-LY2的串联拷贝。这些质粒呈现了指环病毒基因组的每一种可能的环状排列：无论Rep蛋白结合在哪里，它都能驱动病毒基因组从复制起点上游到起点下游的复制。类似的策略已用于生产猪指环病毒(Huang等人，2012，Journal of Virology[病毒学杂志]86(11)6042-6054)。

[2469] 串联TTV-tth8是通过将基因组的拷贝依次克隆到质粒骨架中来组装的，在两个序列之间留下12bp的非病毒DNA。几个TTV-tth8变体被组装成串联质粒，包括野生型和TTV-tth8(Δ36GC)(即，TTV-tth8基因组被工程化为包括本文所述的36个核苷酸的GC富集序列)，它缺失了来自GC富集区的36个碱基对。串联TTMV-LY2是通过金门(Golden-gate)组装进行组装的，同时将基因组的两个拷贝合并到骨架中，并且在基因组之间不留下额外的核苷酸。

[2470] 将携带TTV-tth8(Δ36GC)的串联拷贝的质粒转染到HEK239T细胞中。将细胞培养五天，然后使用0.1％Triton X-100裂解并用核酸酶处理以消化不受病毒衣壳保护的DNA。然后使用Taqman探针对TTV-tth8基因组序列和质粒骨架进行qPCR。TTV-tth8基因组拷贝被标准化为骨架拷贝。如图31B所示，串联TTV-tth8产生的病毒基因组数量是携带单拷贝的质粒的多于四倍。当考虑TTV-tth8基因组序列数量翻倍时，串联质粒每个转染拷贝产生的新合成基因组拷贝数是其两倍多。这些数据表明，工程化串联指环病毒基因组可以增加病毒基因组复制，并可用作增加指环病毒产量的策略。

[2471] 实例35：体外环化的指环病毒基因组

[2472] 该实例描述了包含环状双链环病毒基因组DNA和最少非病毒DNA的构建体。这些环

状病毒基因组与野生型指环病毒复制过程中发现的双链DNA中间体更加匹配。当引入细胞时,这种具有最少非病毒DNA的环状双链指环病毒基因组DNA可以进行滚环复制以产生例如本文所述的遗传元件。

[2473] 在一个实例中,携带TTV-tth8变体和TTMV-LY2的质粒用限制性内切核酸酶消化,所述限制性内切核酸酶识别基因组DNA侧翼的位点。然后将得到的线性化基因组连接起来形成环状DNA。这些连接反应是用不同的DNA浓度进行的,以优化分子内连接。连接的环要么直接转染到哺乳动物细胞中,要么通过用限制性内切核酸酶消化以切割质粒骨架和用外切核酸酶以降解线性DNA来进一步加工以去除非环状基因组DNA。对于TTV-tth8,XmaI核酸内切酶用于使DNA线性化;连接的环在GC富集区和5'非编码区域之间含有53bp的非病毒DNA。对于TTMV-LY2,使用IIS型限制性内切酶Esp3I,产生不含非病毒DNA的病毒基因组DNA环。该方案改编自先前公开的TTV-tth8环化(Kincaid等人,2013,PLoS Pathogens[公共科学图书馆—病原体]9(12):e1003818)。为了证明指环病毒产量的改进,将环化的TTV-tth8和TTMV-LY2转染到HEK293T细胞中。孵育7天后,裂解细胞,并进行qPCR以比较环化的和基于质粒的指环病毒基因组之间的指环病毒基因组水平。指环病毒基因组水平的增加表明病毒DNA的环化是增加指环病毒产量的有用策略。

[2474] 在另一个实例中,TTMV-LY2质粒(pVL46-240)和TTMV-LY2-nLuc分别用Esp3I或EcoRV-HF线性化。消化后的质粒在1%琼脂糖凝胶上纯化,然后电洗脱或Qiagen柱纯化,并用T4 DNA连接酶连接。转染前将环化的DNA在100kDa UF/DF膜上浓缩。环化通过凝胶电泳确认,如图31C所示。在用Lipofectamine 2000进行脂质转染前一天,将HEK293T以$3x10^4$个细胞/cm^2接种到T-225烧瓶中。在烧瓶接种后一天共转染九微克环化的TTMV-LY2 DNA和50μg环化的TTMV-LY2-nLuc。作为比较,另外的T-225烧瓶用50μg线性化的TTMV-LY2和50μg线性化的TTMV-LY2-nLuc共转染。

[2475] 在Triton X-100收获缓冲液中收获细胞前,指环体生产进行了八天。通常,可以通过例如宿主细胞的裂解、裂解物的澄清、过滤和色谱来富集指环体。在该实例中,收获的细胞在氯化钠调整和1.2μm/0.45μm正常流动过滤之前进行了核酸酶处理。在750kDa MWCO mPES中空纤维膜上浓缩澄清的收获物并将缓冲液交换到PBS中。将TFF渗余物用0.45μm过滤器过滤,然后装入在PBS中预平衡的Sephacryl S-500HR SEC柱。指环体以30cm/hr通过SEC柱进行处理。如图31D所示,将各个级分通过qPCR收集并且分析其病毒基因组拷贝数和转基因拷贝数。从SEC色谱图的空隙体积,级分7,开始观察到病毒基因组和转基因拷贝。在级分15处观察到残留质粒峰。TTMV-LY2基因组和TTMV-LY2-nLuc转基因的拷贝数与使用环化的输入DNA在级分7-级分10处产生的指环体非常一致,表明包装的指环体包含nLuc转基因。将SEC级分合并和使用100kDa MWCO PVDF膜浓缩,然后在体内施用之前0.2μm过滤。

[2476] 与线性化的指环体DNA相比,输入指环体DNA的环化导致在整个纯化过程中核酸酶保护的基因组的回收百分比增加了三倍,表明使用环化的输入指环体DNA提高了制造效率,如表46所示。

[2477] 表46.纯化过程产率

[2478]

步骤	线性化的 TTMV-LY2		环状 TTMV-LY2	
	病毒基因组总拷贝数	nLuc 转基因基因组总拷贝数	病毒基因组总拷贝数	nLuc 转基因基因组总拷贝数
核酸酶前收获	2.78E + 12	2.17E + 12	1.04E + 11	4.39E + 11
澄清的收获	9.96E + 09	5.48E + 09	6.55E + 08	9.81E + 08
TFF	1.01E + 10	7.66E + 09	2.58E + 08	3.56E + 08
SEC	3.18E + 07	8.73E + 06	9.16E + 06	7.75E + 06
UF/DF	8.82E + 06	3.25E + 06	1.78E + 06	2.73E + 06
无菌过滤	5.60E + 06	2.64E + 06	8.66E + 05	1.63E + 06
纯化过程产率 (%)	0.0002%	0.0001%	0.0006%	0.0004%

[2479] 实例36:ORF1建模和保守残基和结构域的鉴定

[2480] 该实例描述了乙型细环病毒的ORF1蛋白的计算机模拟建模,并根据结构基序和氨基酸保守性/相似性定义推定结构域。

[2481] ORF1蛋白被预测为指环病毒的主要衣壳蛋白,这是基于精氨酸富集区的存在和使用PSIpred的二级结构预测中β-折叠的高度存在(http://bioinf.cs.ucl.ac.uk/psipred/)。RaptorX(http://raptorx.uchicago.edu/)被用于对八种乙型细环病毒的序列进行结构预测和接触预测。使用乙型细环病毒ORF1序列是因为它们比甲型细环病毒(约750个氨基酸)更短(约650个氨基酸),后者预测为非结构化的区域更少。预测的结构中的五个包含用于鉴定ORF1推定结构域的具有相似性的元件(图33)。ORF1分为五个区域-精氨酸富集区、推定的核心(胶冻卷结构域)、高变区、N22区域和C末端结构域。

[2482] 乙型细环病毒株CBS203的结构模型用于展示在乙型细环病毒家族中具有一定保守性的残基/结构区域。为了分析保守残基,使用ClustalW比对算法在Geneious中比对了110个乙型细环病毒ORF1序列。然后使用阈值为1的BLOSUM62矩阵通过同一性和相似性百分比评估残基的保守性。结构模型上突出显示了比对中所有株的具有大于60%相似性的残基(图34)。总共有26个残基(约4%)与100%的比对序列具有氨基酸相似性。80%和60%截断值分别包含总残基的23.7%和36.7%。

[2483] 对甲型细环病毒的258个株进行了类似的比对算法和相似性测定。相似性和同一性展示在来自比对的共有序列中,并且基于与乙型细环病毒的一级序列比对来分配推定结构域(图35)。甲型细环病毒具有29个残基(3.9%),它们100%相似,与乙型细环病毒的观察结果非常一致。有趣的是,与具有至少80%(30.9%的残基)或60%(42.9%的残基)相似性的乙型细环病毒相比,甲型细环病毒具有更高百分比的残基。

[2484] 实例37:生产含有来自不同细环病毒株的具有高变结构域的嵌合ORF1的指环体

[2485] 该实例描述了ORF1的高变区的结构域交换,以产生嵌合指环体,所述嵌合指环体包含一个TTV株的ORF1精氨酸富集区、胶冻卷结构域、N22和C末端结构域,以及来自不同TTV

株的ORF1蛋白的高变结构域。

[2486] 乙型细环病毒的全长基因组LY2株已被克隆到表达载体中，用于在哺乳动物细胞中表达。对该基因组进行突变以去除LY2的高变结构域并将其替换为远亲乙型细环病毒的高变结构域(图36)。然后使用先前公布的方法(Kincaid等人,PLoS Pathogens[公共科学图书馆—病原体]2013)将含有具有交换的高变价格域的LY2基因组的质粒(pTTMV-LY2-HVRa-z)线性化和环化。HEK293T细胞用环化的基因组转染并孵育5-7天，以允许指环体生产。孵育期后，通过梯度超速离心从转染的细胞的上清液和细胞沉淀中纯化出指环体。

[2487] 为了确定嵌合指环体是否仍然具有感染性，将分离的病毒颗粒添加到未感染的细胞中。将细胞孵育5-7天以允许病毒复制。孵育后，嵌合指环体建立感染的能力将通过免疫荧光、蛋白质印迹和qPCR进行监测。嵌合病毒的结构完整性通过负染色和冷冻电子显微镜进行评估。可以进一步测试嵌合指环体在体内感染细胞的能力。通过高变结构域交换产生功能性嵌合指环体的能力的建立可以允许病毒工程化以改变向性并可能逃避免疫检测。

[2488] 实例38:生产含有非TTV蛋白/肽代替高变结构域的嵌合ORF1

[2489] 该实例描述了用其他目的蛋白质或肽替换ORF1的高变区，以产生嵌合ORF1蛋白，其包含一个TTV株的精氨酸富集区、胶冻卷结构域、N22和C末端结构域，以及替代高变结构域的非TTV蛋白/肽。

[2490] 如实例B所示，LY2的高变结构域从基因组中缺失，目的蛋白质或肽可以插入该区域(图37)。可以引入该区域的序列类型的实例包括但不限于亲和标签、抗体的单链可变区(scFv)和抗原肽。质粒(pTTMV-LY2-ΔHVR-POI)中的突变基因组如实例B中所述进行线性化和环化。将环化的基因组转染到HEK293T细胞中并孵育5-7天。孵育后，在合适的情况下，通过超速离心和/或亲和色谱从上清液和细胞沉淀中纯化含有POI的嵌合指环体。

[2491] 使用各种技术评估产生含有POI的功能性嵌合指环体的能力。首先，将纯化的病毒添加到未感染的细胞中，以确定嵌合指环体是否可以复制和/或将有效负载递送给原初细胞。此外，使用电子显微镜评估嵌合指环体的结构完整性。对于体外功能性的嵌合指环体，还评估了体内复制/递送有效负载的能力。

[2492] 实例39:分泌酶的指环体体内传递

[2493] 该实例说明了施用后由指环体递送的分泌酶的体内效应子功能。

[2494] 制备了包含编码ADAMTS13(具有血小板反应蛋白1型基序的解聚素和金属蛋白酶，成员13)的转基因的指环体。简而言之，产生了五个构建体：构建体A-TTMV-LY2载体±ADAMTS13;构建体B-ADAMTS13蛋白和TTMV-LY2 ORF;构建体C-用于生产TTMV-LY2载体的质粒;构建体D-用于生产ADAMTS13蛋白和TTMV-LY2 ORF的质粒;和构建体E-无菌PBS。构建体A和构建体B在HEK-293T细胞中产生，并通过核酸酶处理、超滤/渗滤和无菌过滤进行纯化。构建体C和构建体D在大肠杆菌中产生，通过MaxiPrep纯化，然后在PBS中稀释至目标拷贝数，然后进行无菌过滤。构建体E是通过PBS的无菌过滤产生。HEK-293T细胞在DMEM+10％FBS中以三天和四天的传代时间表从解冻开始扩展到第4代。在第5代，细胞以大约$5x10^4$个细胞/cm^2接种，用于第二天的转染。接种后一天，使用Lipofectamine 2000用构建体共转染细胞。转染后，孵育细胞以允许生产指环体并收获指环体。

[2495] 将25uL指环体制剂或适当的对照物静脉内施用给基因工程化的VWD型2B小鼠。每天对每只动物进行抽血以确定溶血和血小板减少症，以及测量血液中ADAMTS13的分泌

(SensoLyte®520ADAMTS13活性测地,*荧光测定*AnaSpec公司(AnaSpec,Inc.))

[2496] 在第0、1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20和21天的每一天测量ADAMTS13信号的存在。此外,在同一时间点,进行了体外溶血测定方案。简而言之,将血液离心并测量上清液的吸光度,所述上清液包括血浆和裂解的红细胞。从裂解的红细胞(Triton X-100)的标准曲线计算裂解百分比。ADAMTS13的存在增加和溶血减少将证明经由指环体递送的ADAMTS13的体内表达和活性。

[2497] 实例40:设计包含有效负载的指环体

[2498] 该实例描述了包含反式基因的示例性指环体遗传元件的设计。遗传元件由来自指环病毒科成员的必需顺式复制和包装结构域以及非指环病毒有效负载组成,所述非指环病毒有效负载可能包括例如蛋白质或非编码RNA表达基因。指环体缺乏复制和包装所必需的反式蛋白质元件,并且需要其他来源(例如,辅助病毒,例如复制病毒、表达质粒或基因组整合)提供的蛋白质来进行滚环复制和衣壳化。

[2499] 在一组实例中,缺失了整个蛋白质编码DNA序列,从第一个起始密码子到最后一个终止密码子(图38)。对于TTV-tth8,从ORF2起始密码子到ORF3终止密码子的第336至3015位核苷酸被缺失。对于TTMV-LY2,从ORF2起始密码子到ORF3终止密码子的424到2813被缺失。所得DNA保留了病毒非编码区(NCR),包括病毒启动子、5'UTR保守结构域、3'UTR(在某些指环病毒株中编码miRNA,例如TTV-tth8)和GC富集区。指环体NCR包含必需的顺式结构域,包括病毒复制起点和衣壳结合结构域。然而,由于缺乏编码指环病毒蛋白的开放阅读框,指环体无法表达DNA复制和衣壳化所需的必需蛋白质因子,因此除非以反式提供这些元件,否则无法扩增或包装。

[2500] 有效负载DNA,包括但不限于蛋白质编码序列、完整的反式基因(包括非指环病毒启动子序列)和非编码RNA基因,通过插入到缺失的指环病毒开放阅读框的位点中来掺入指环体遗传元件(图38)。例如,可以通过天然病毒启动子或作为反式基因掺入的合成启动子驱动蛋白质编码序列的表达。

[2501] 复制缺陷或无复制能力的指环体遗传元件(例如,如本文所述)可能缺乏用于病毒复制和/或衣壳因子的蛋白质编码序列。因此,包装的指环体是通过用该实例中描述的指环体DNA和病毒蛋白编码DNA共转染细胞来产生的。病毒蛋白由具有复制能力的野生型病毒基因组、包含受病毒启动子控制的病毒蛋白的非复制型质粒或包含受强组成型启动子控制的病毒蛋白的质粒表达。

[2502] 实例41:基于tth8和LY2的指环体分别成功地将EPO基因转导到肺癌细胞中

[2503] 在该实例中,非小细胞肺癌细胞系(EKVX)使用两种不同的携带促红细胞生成素基因(EPO)的指环体转导。如本文所述,指环体通过体外环化产生,并且包括基于LY2或tth8骨架的两种类型的指环体(例如,分别如表15和16或表5和6中所述)。LY2-EPO和tth8-EPO指环体中的每一个都包括遗传元件,所述遗传元件分别包括EPO编码盒和LY2或tth8基因组的非编码区(5'UTR,GC富集区),但不包括指环病毒ORF,例如,如实例40中所述。用纯化的指环体或阳性对照(高剂量或与指环体相同剂量的AAV2-EPO)接种细胞并孵育7天。指环病毒ORF以反式提供,位于单独的体外环化DNA中。在接种后3、5.5和7天对培养物上清液取样,并使用商业ELISA试剂盒进行测定以检测EPO。LY2-EPO和tth8-EPO指环体均成功转导细胞,与未处理(阴性)对照细胞相比,EPO效价显著更高(在所有时间点$P<0.013$)(图40)。

[2504] 实例42:静脉内(i.v.)施用后,可以在体内检测到带有治疗性转基因的指环体

[2505] 在该实例中,静脉内(i.v.)施用后,在体内检测到编码人生长激素(hGH)的指环体。基于LY2骨架并编码外源hGH(LY2-hGH)的复制缺陷型指环体通过如本文所述的体外环化产生。LY2-hGH指环体的遗传元件包括LY2非编码区(5'UTR,GC富集区)和hGH编码盒,但不包括指环体病毒ORF,例如,如实例40中所述。LY2-hGH指环体静脉内施用给小鼠。指环病毒ORF以反式提供,位于单独的体外环化DNA中。简而言之,在第0天静脉内注射指环体(LY2-hGH)或PBS(n=4只小鼠/组)。指环体以每只小鼠4.66E+07个指环体基因组施用于独立的动物组。

[2506] 在第一实例中,检测到指环体病毒基因组DNA拷贝。在第7天,收集血液和血浆并通过qPCR分析hGH DNA扩增子。在体内感染后7天后,LY2-hGH指环体存在于全血的细胞级分中(图41A)。此外,血浆中不存在指环体证明这些指环体无法在体内复制(图41B)。

[2507] 在第二实例中,在体内转导后检测到hGH mRNA转录物。在第7天,收集血液并通过qRT-PCR分析hGH mRNA转录扩增子。GAPDH用作对照看家基因。在全血的细胞级分中测量hGH mRNA转录物。在体内检测到来自指环体编码的转基因的mRNA(图42)。

[2508] 实例43:指环病毒中的编码序列大小分布

[2509] 所有指环病毒的编码序列(CDS)长度是利用内部鉴定的野生型株的广泛目录进行评估的。绘制指环病毒的CDS长度,比较三个人指环病毒属(甲型细环病毒,α;乙型细环病毒,β;和丙型细环病毒,γ)的病毒株,并将公开可用的基因组序列长度与发明人内部(内部)组装的那些进行比较。所有指环病毒的平均CDS长度为约2100个核苷酸。甲型细环病毒属中的TTV大于来自乙型细环病毒和丙型细环病毒属的指环病毒(分别为TTV小型和TTV中型)。具体而言,在甲型细环病毒TTV中观察到的平均CDS为2237个核苷酸,范围为1800-2541个核苷酸。观察到乙型细环病毒的平均CDS长度为2011个核苷酸,范围为1803-2229个核苷酸。观察到丙型细环病毒的平均CDS长度为2012个核苷酸,范围为1812-2379个核苷酸。

[2510] 实例44:用以表征ORF2的高度保守基序

[2511] 如图43A中的示例性基因组所示,指环病毒ORF2可能编码具有可能的磷酸酶活性和在病毒复制和宿主免疫调节中的作用的非结构蛋白。检查了广泛的病毒序列库中是否存在保守的ORF2氨基酸基序(图43B)。然后使用该基序在内部和公共序列中鉴定超过1,000个指环病毒ORF2序列。发现该ORF2基序在大量人指环病毒株以及所有检查的非人指环病毒(啮齿动物、猪和灵长类动物指环病毒,以及鸡贫血病毒)中保持保守,使其成为迄今为止鉴定的最高度保守的指环病毒基序。还进行了ORF2结构建模,这表明ORF2基序中的保守残基保持在螺旋-转角-螺旋结构中,其取向表明可能存在金属结合结构域(图43C)。有趣的是,ORF1与ORF2相比的系统发育树(图43D)显示出甲型细环病毒、乙型细环病毒和丙型细环病毒的类似属级分解,表明ORF2是属特异性的。

[2512] 实例45:人全长指环病毒ORF1 mRNA的证据

[2513] 指环病毒在体外表达至少三种可变剪接的mRNA,其中最长的(约2.2kb)预计编码全长ORF1。在该实例中,在体内评估ORF1 mRNA转录。

[2514] 为此,检查了来自GTEx(基因型-组织表达)项目的公开可用的RNA Seq组织数据。目标是鉴定包含足够的指环病毒RNA读段的人组织样品,以对病毒转录物进行分类。鉴定了104个具有指环病毒RNA读段的组织样品(占所有组织的2.4%,占血液样品的19%);其中7

个样品具有超过20个指环病毒RNA读段，允许进行病毒转录组分析。这7个指环病毒阳性样品中的3个还具有匹配的WGS数据，从中可以组装相应的指环病毒DNA基因组以进行精确读段映射(图44A)。由于缺乏相应的病毒参考基因组，指环病毒多样性禁止提供信息性的RNA读段映射。在三个供体(两个血液样品和一个肺组织样品)中检测到映射到ORF1区域的RNA读段。在一个供体血液样品中，鉴定了涵盖全长ORF1区域的指环病毒RNA读段(图44B，灰色条表示读段对)。这是使用RNA Seq数据首次在体内确认全长指环病毒转录物。

[2515] 实例46:体外环化的基因组作为体外生产指环体的输入材料

[2516] 该实例证明体外环化(IVC)的双链指环病毒DNA，作为本文所述指环病毒遗传元件的源材料，比质粒中的指环病毒基因组DNA更稳健，以产生预期密度的包装的指环病毒基因组。

[2517] T75烧瓶中的1.2E+07HEK293T细胞(人胚胎肾细胞系)用11.25ug的以下任一种转染:(i)体外环化的双链TTV-tth8基因组(IVC TTV-tth8)，(ii)在质粒骨架中的TTV-tth8基因组，或(iii)仅包含TTV-tth8(非复制型TTV-tth8)的ORF1序列的质粒。转染后7天收获细胞，用0.1％Triton裂解，并用100单位/ml的Benzonase处理。裂解物用于氯化铯密度分析；测量密度，并对氯化铯线性梯度的每个级分进行TTV-tth8拷贝定量。如图45所示，与TTV-tth8质粒相比，IVC TTV-tth8在1.33的预期密度下产生了显著更多的病毒基因组拷贝。

[2518] 1E+07Jurkat细胞(人T淋巴细胞细胞系)用体外环化LY2基因组(LY2IVC)或质粒中的LY2基因组进行核转染。转染后4天收获细胞并使用含有0.5％氚和300mM氯化钠的缓冲液裂解，然后进行两轮即时冻融。裂解物用100单位/ml的benzonase处理，然后进行氯化铯密度分析。对氯化铯线性梯度的每个级分进行密度测量和LY2基因组定量。如图46所示，与包含LY2基因组的质粒转染(其在图46中未显示可检测的峰)相比，Jurkat细胞中体外环化的LY2基因组的转染导致在预期密度处的尖峰。

[2519] 实例47:指环病毒ORF1中保守二级结构基序的鉴定

[2520] 在该实例中，计算模型被用来鉴定指环病毒ORF1蛋白的二级结构中的保守基序。使用JPred程序对单个菌株进行二级结构预测。

[2521] 通常，人TTV的胶冻卷结构域的长度为约200个氨基酸(AA)±3AA。示例性胶冻卷结构域的二级结构以5-7AA的β链开始，然后是3-5AA无规卷曲、15-16AAβ链、26-28AA无规卷曲、15-17AAα螺旋、2AA无规卷曲、3-4AAβ链、8AA无规卷曲、10-11AAβ链、5-6AA无规卷曲、6-7AAβ链、8-14AA无规卷曲、8-14AAα-螺旋(在一些情况下可能会分成2个较小的螺旋)、3-4AA无规卷曲、4-5AAβ链、10AA无规卷曲、5-6AAβ链、20-21AA无规卷曲、7-9AAβ链、14-16AA无规卷曲、5-7AAβ链。图47显示了来自甲型细环病毒、乙型细环病毒和丙型细环病毒进化分枝的示例性指环病毒ORF1二级结构的比对。

[2522] ORF1的N22结构域中$YNPX^2DXGX^2N$(SEQ ID NO:829)基序的二级结构也具有围绕它的保守二级结构。从在基序的位置1处的酪氨酸(Y)之后断裂的5-6AAβ链开始，大多数基序排队在8-9AA无规卷曲中，直到末端天冬酰胺(N)，此处另一个7-8AAβ链起始。示例性指环病毒ORF1 N22基序序列的比对显示在图48中。基序中的酪氨酸破坏了β链，第二β链从基序的末端天冬酰胺开始。

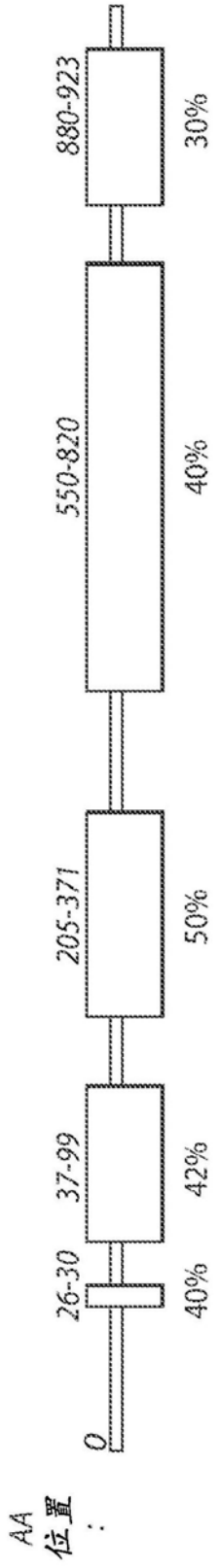


图1A

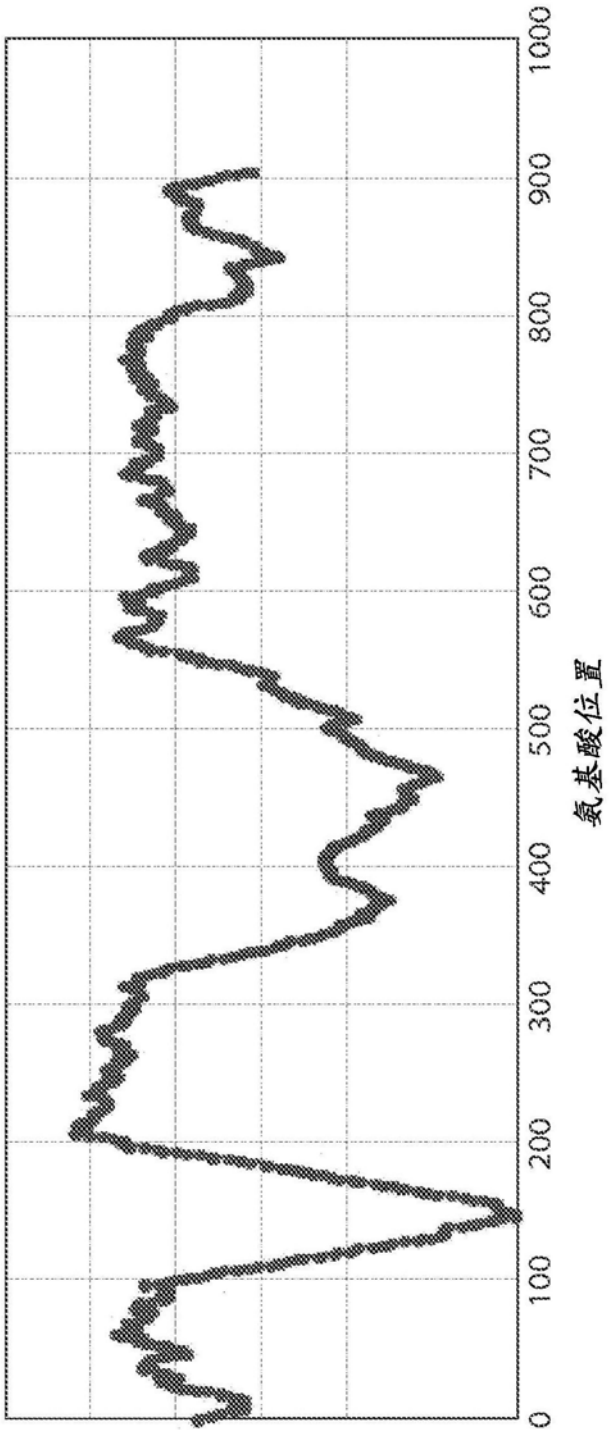


图1B

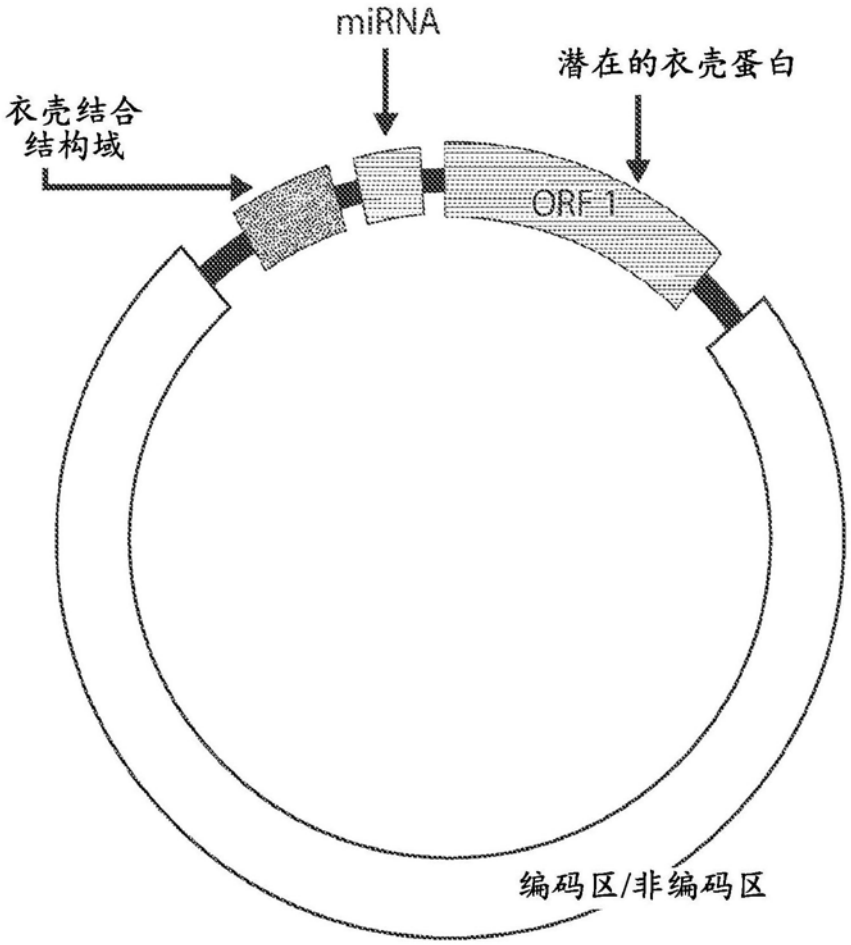
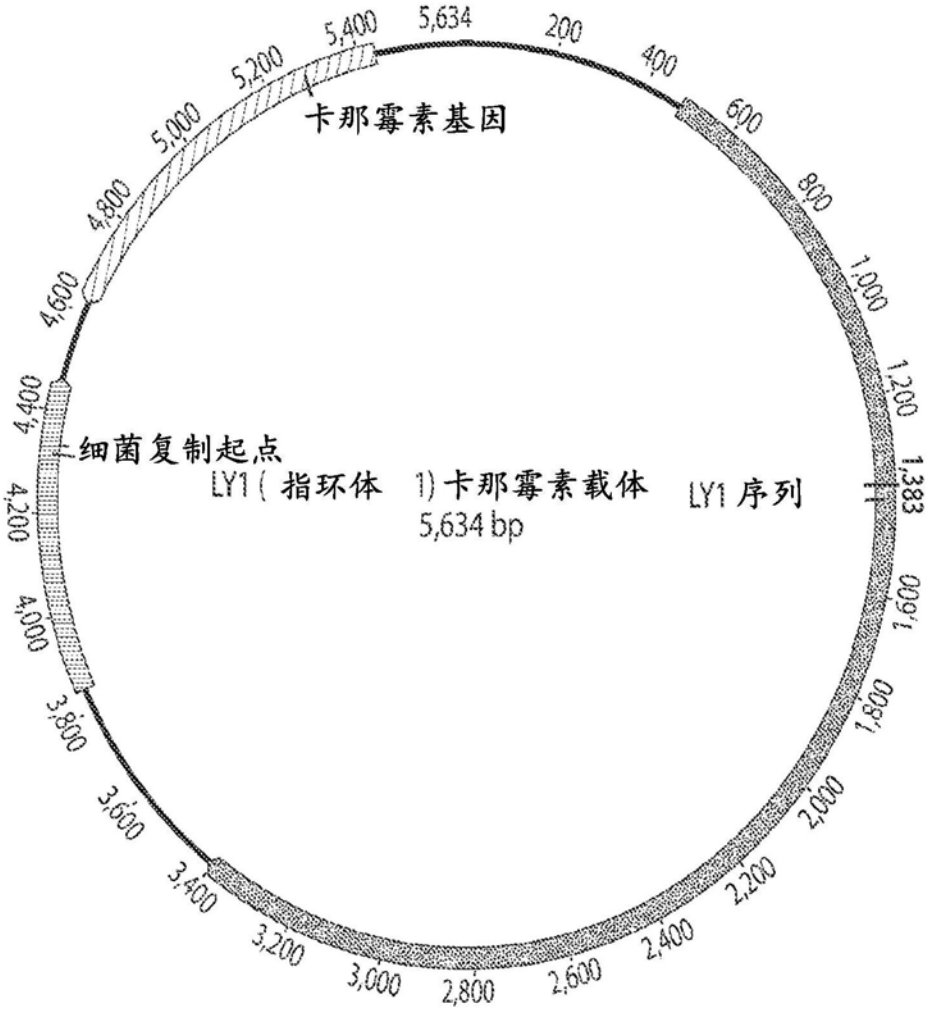
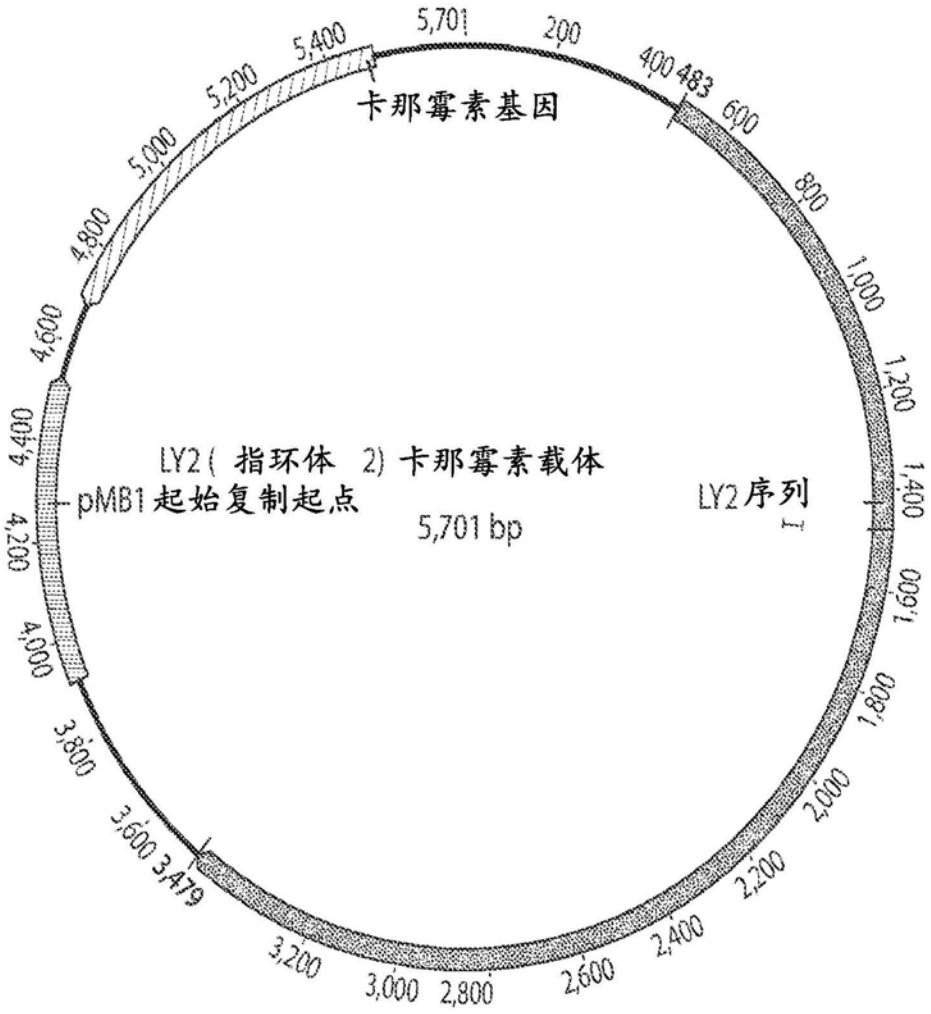


图2



编码TTMiniV的LY1株（指环体1）的卡那霉素载体的示意图

图3



编码TTMiniV的LY2株（指环体2）的卡那霉素载体的示意图

图4

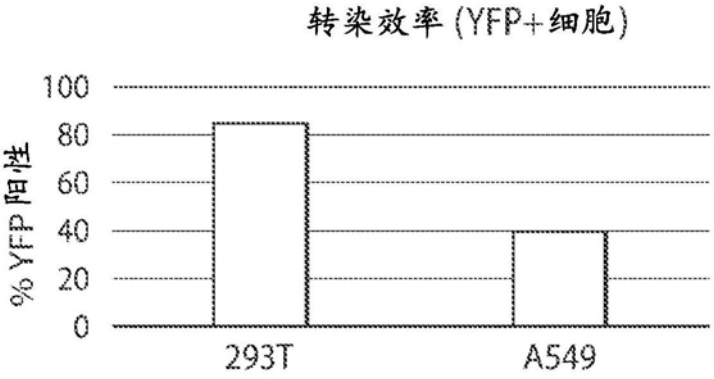


图5

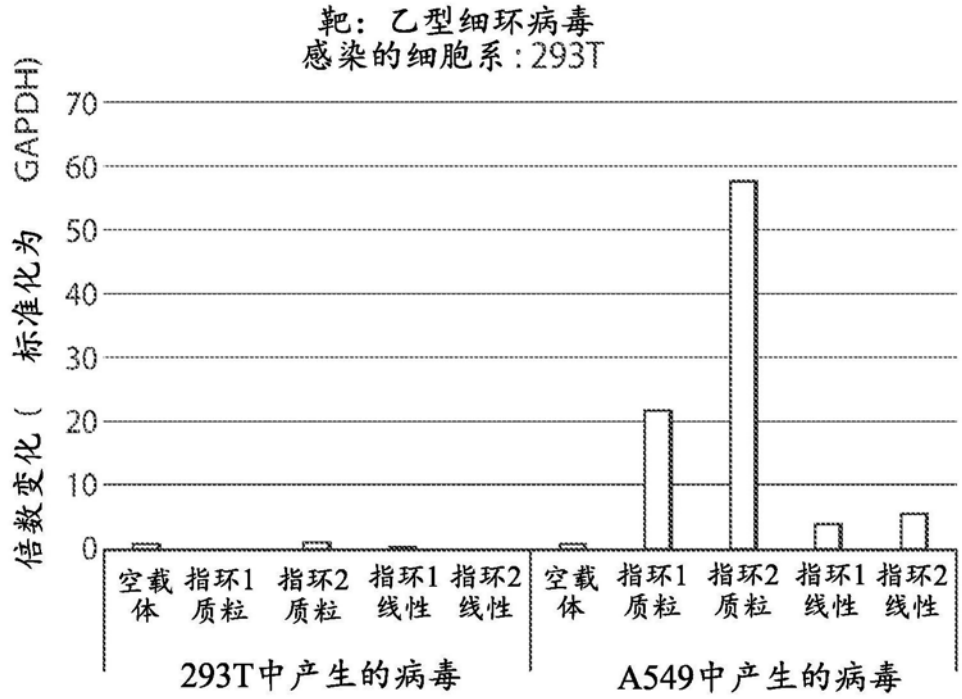

靶：乙型细环病毒
感染的细胞系：293T
倍数变化（标准化为GAPDH）
70
60
50
40
30
20
10
0
空载体
指环1质粒
指环2质粒
指环1线性
指环2线性
空载体
指环1质粒
指环2质粒
指环1线性
指环2线性
293T中产生的病毒
A549中产生的病毒

图6A

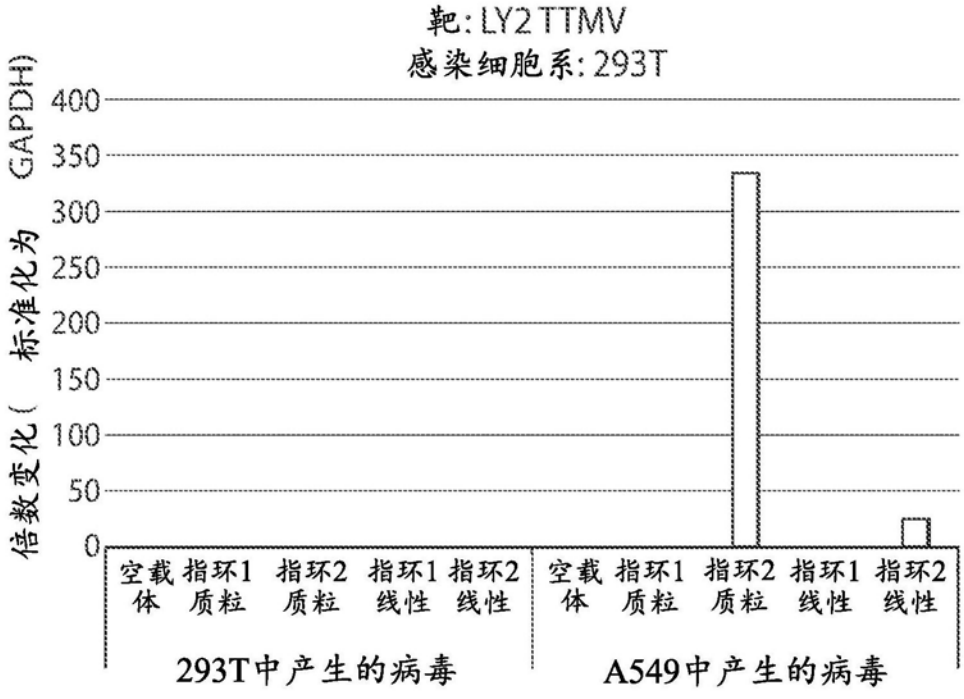

靶：LY2 TTMV
感染细胞系：293T
倍数变化（标准化为GAPDH）
400
350
300
250
200
150
100
50
0
空载体
指环1质粒
指环2质粒
指环1线性
指环2线性
空载体
指环1质粒
指环2质粒
指环1线性
指环2线性
293T中产生的病毒
A549中产生的病毒

图6B

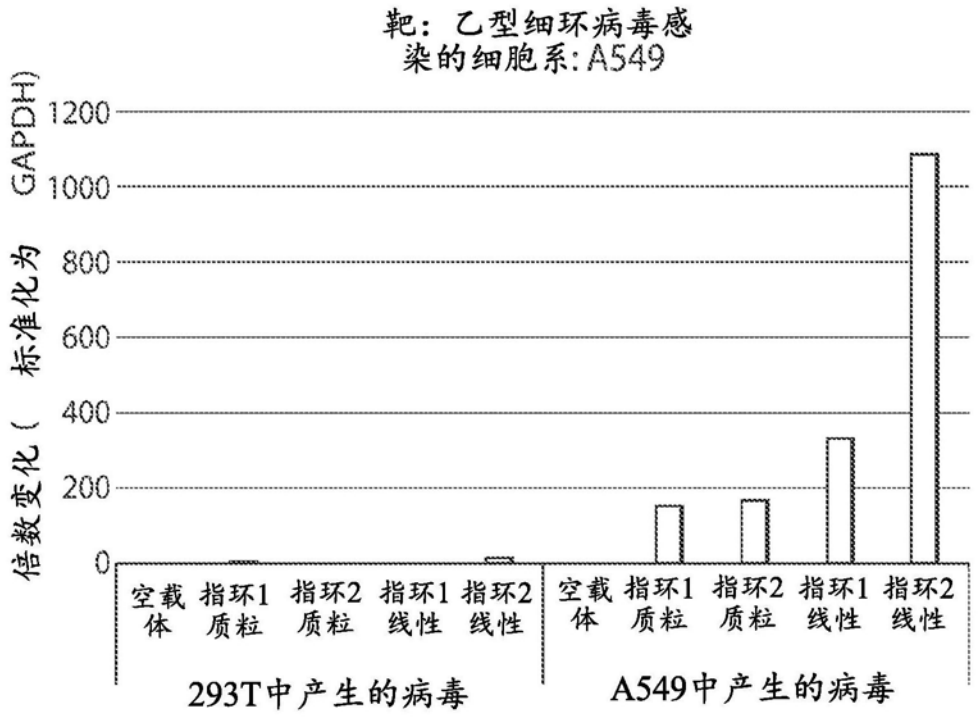

靶：乙型细环病毒感
染的细胞系:A549
倍数变化（ 标准化为 GAPDH)
1200
1000
800
600
400
200
0
空载体
指环1质粒
指环2质粒
指环1线性
指环2线性
空载体
指环1质粒
指环2质粒
指环1线性
指环2线性
293T中产生的病毒
A549中产生的病毒

图7A

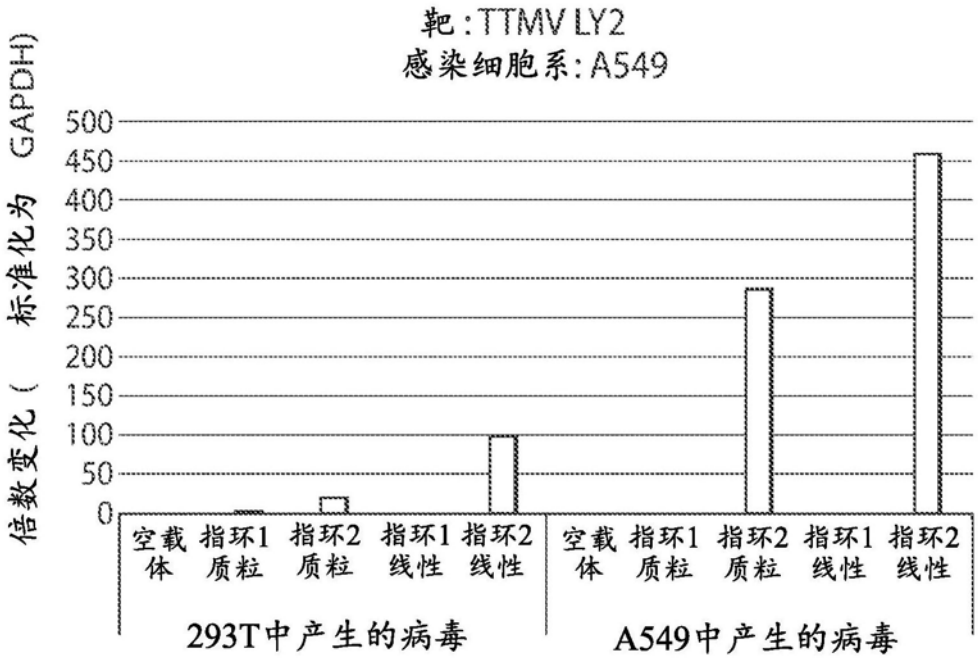

靶:TTMV LY2
感染细胞系:A549
倍数变化（ 标准化为 GAPDH)
500
450
400
350
300
250
200
150
100
50
0
空载体
指环1质粒
指环2质粒
指环1线性
指环2线性
空载体
指环1质粒
指环2质粒
指环1线性
指环2线性
293T中产生的病毒
A549中产生的病毒

图7B

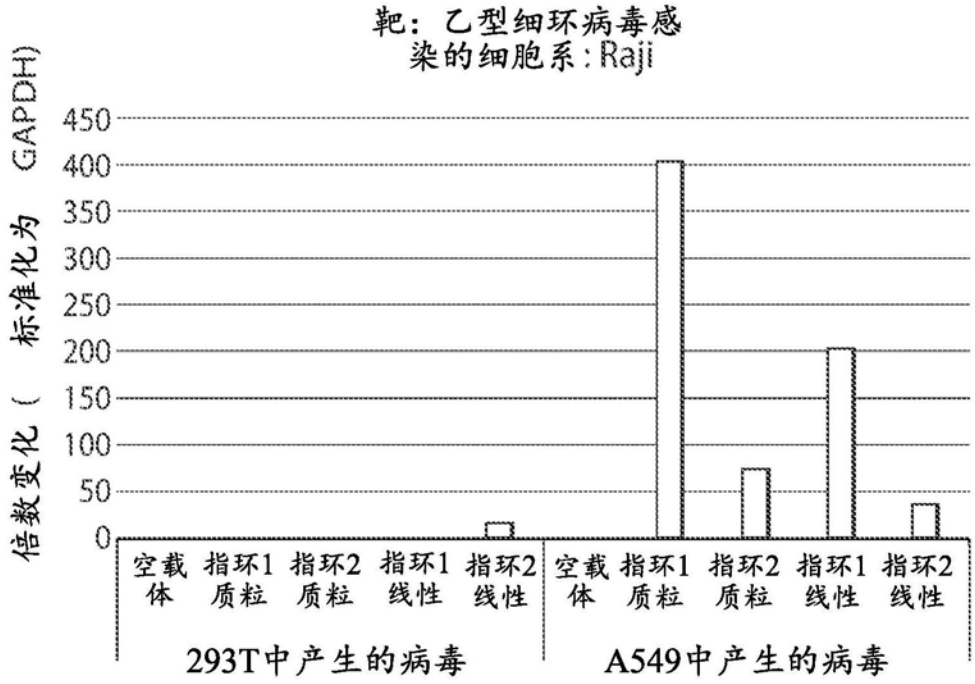

靶：乙型细环病毒感染的细胞系：Raji
倍数变化（标准化为GAPDH）
450
400
350
300
250
200
150
100
50
0
空载体
指环1质粒
指环2质粒
指环1线性
指环2线性
空载体
指环1质粒
指环2质粒
指环1线性
指环2线性
293T中产生的病毒
A549中产生的病毒

图8A

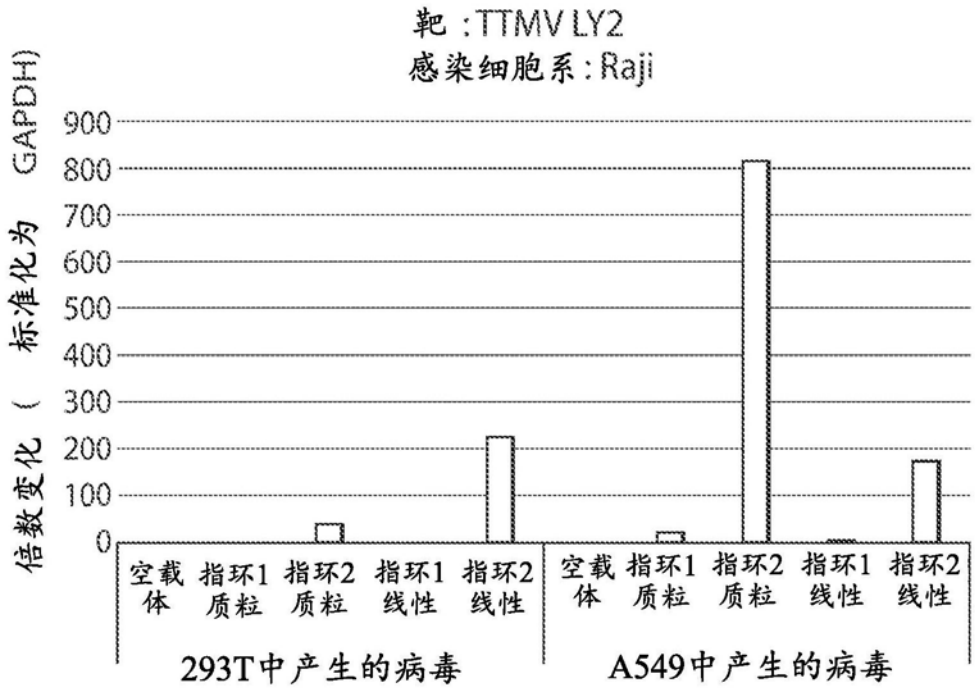

靶：TTMV LY2
感染细胞系：Raji
倍数变化（标准化为GAPDH）
900
800
700
600
500
400
300
200
100
0
空载体
指环1质粒
指环2质粒
指环1线性
指环2线性
空载体
指环1质粒
指环2质粒
指环1线性
指环2线性
293T中产生的病毒
A549中产生的病毒

图8B

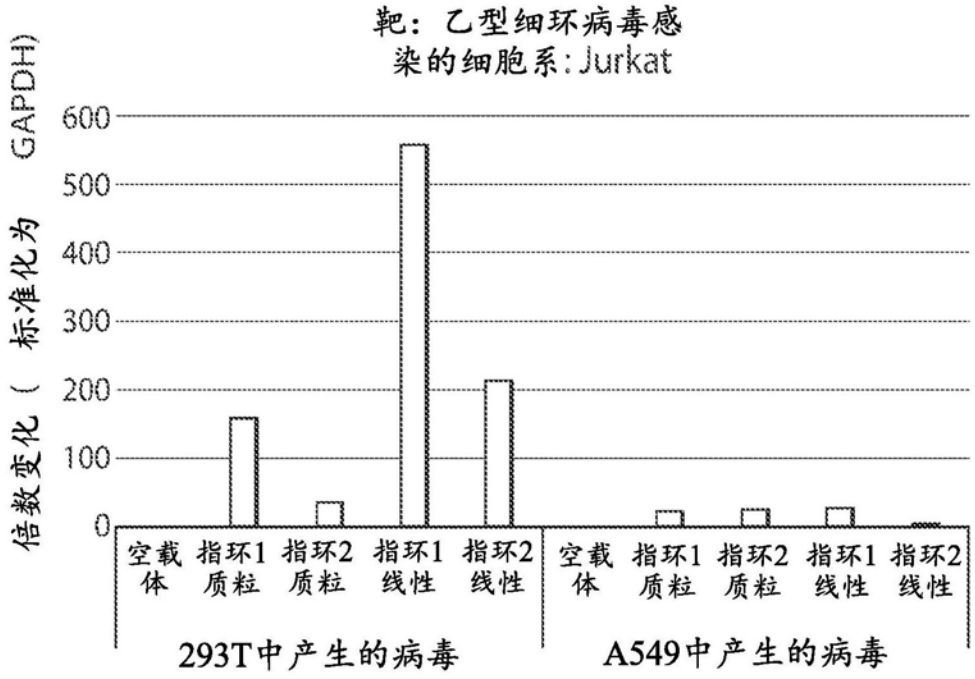

靶：乙型细环病毒感
染的细胞系：Jurkat
倍数变化（标准化为GAPDH）
600
500
400
300
200
100
0
空载体
指环1质粒
指环2质粒
指环1线性
指环2线性
空载体
指环1质粒
指环2质粒
指环1线性
指环2线性
293T中产生的病毒
A549中产生的病毒

图9A

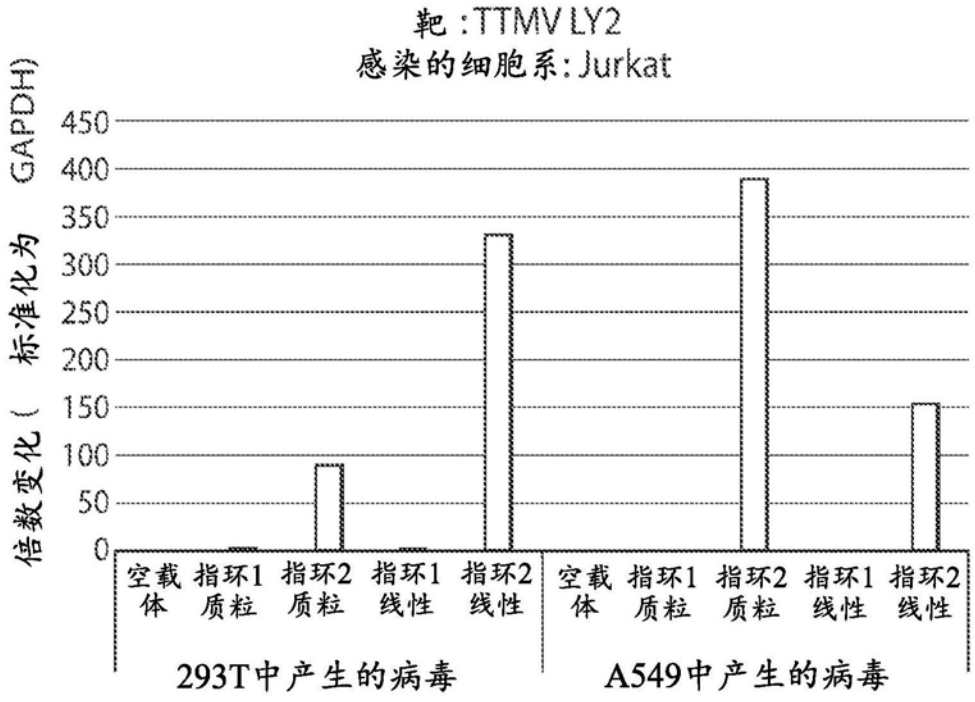

靶：TTMV LY2
感染的细胞系：Jurkat
倍数变化（标准化为GAPDH）
450
400
350
300
250
200
150
100
50
0
空载体
指环1质粒
指环2质粒
指环1线性
指环2线性
空载体
指环1质粒
指环2质粒
指环1线性
指环2线性
293T中产生的病毒
A549中产生的病毒

图9B

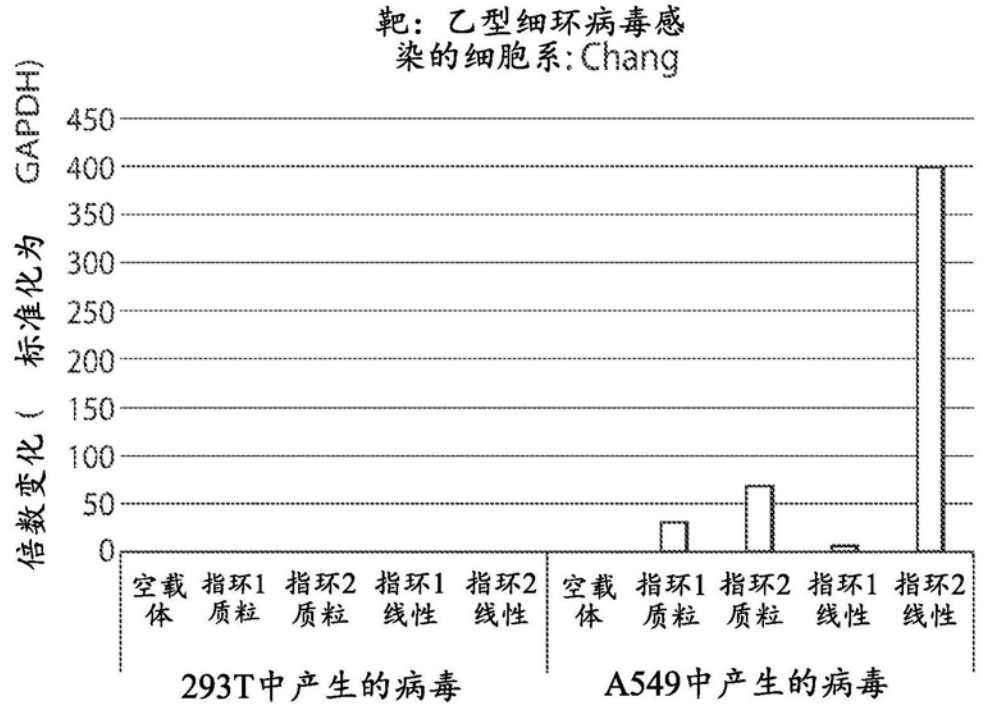

靶：乙型细环病毒感染的细胞系:Chang
倍数变化（ 标准化为 GAPDH）
450
400
350
300
250
200
150
100
50
0
空载体
指环1质粒
指环2质粒
指环1线性
指环2线性
空载体
指环1质粒
指环2质粒
指环1线性
指环2线性
293T中产生的病毒
A549中产生的病毒

图10A

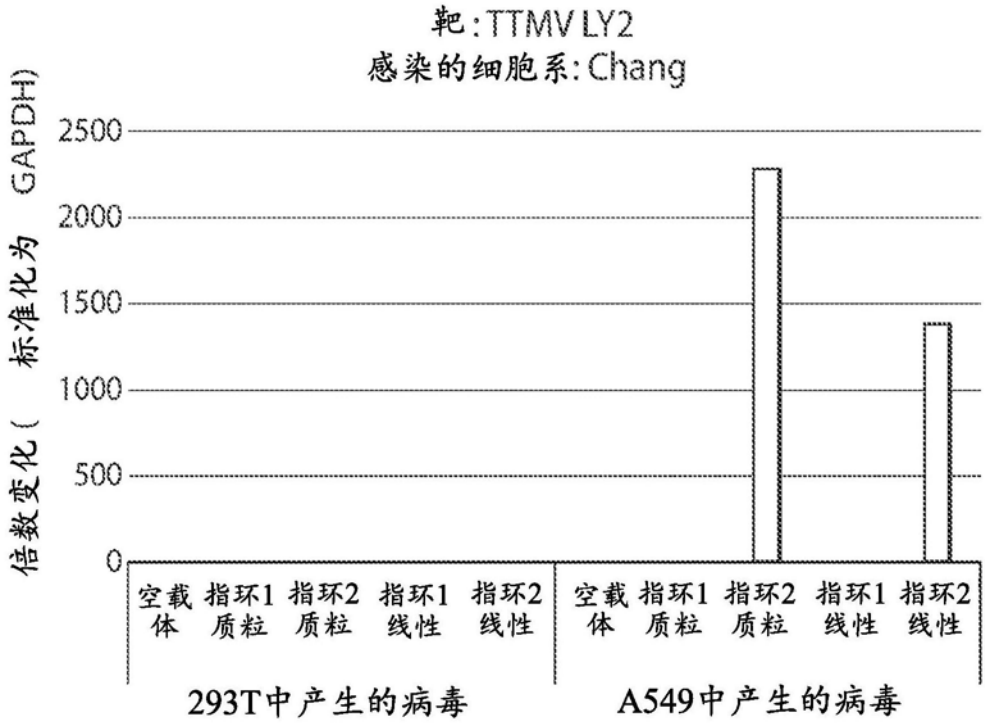

靶:TTMV LY2
感染的细胞系:Chang
倍数变化（ 标准化为 GAPDH）
2500
2000
1500
1000
500
0
空载体
指环1质粒
指环2质粒
指环1线性
指环2线性
空载体
指环1质粒
指环2质粒
指环1线性
指环2线性
293T中产生的病毒
A549中产生的病毒

图10B

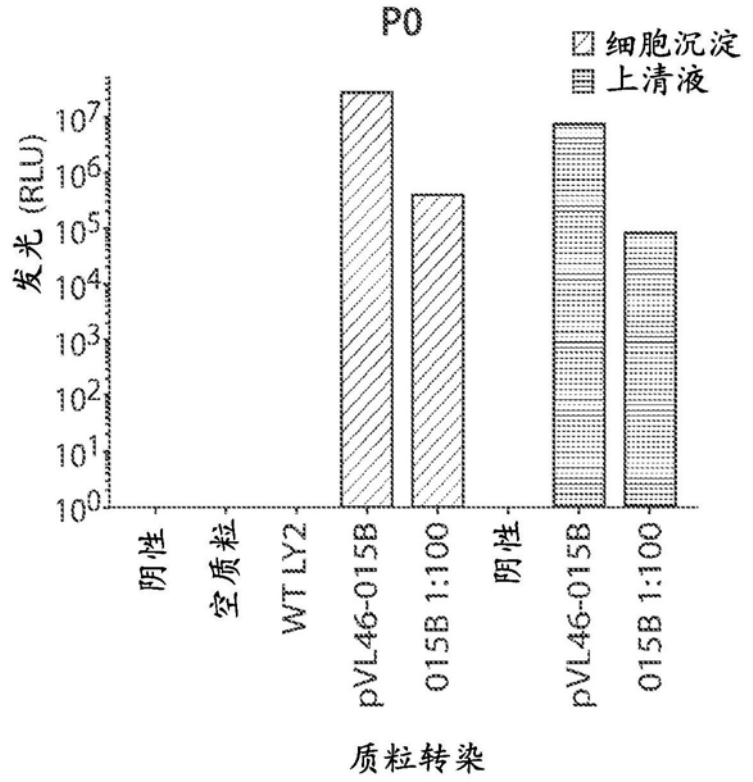

P0
细胞沉淀
上清液
发光 (RLU)
阴性
空质粒
WT LY2
pVL46-015B
015B 1:100
阴性
pVL46-015B
015B 1:100
质粒转染

图11A

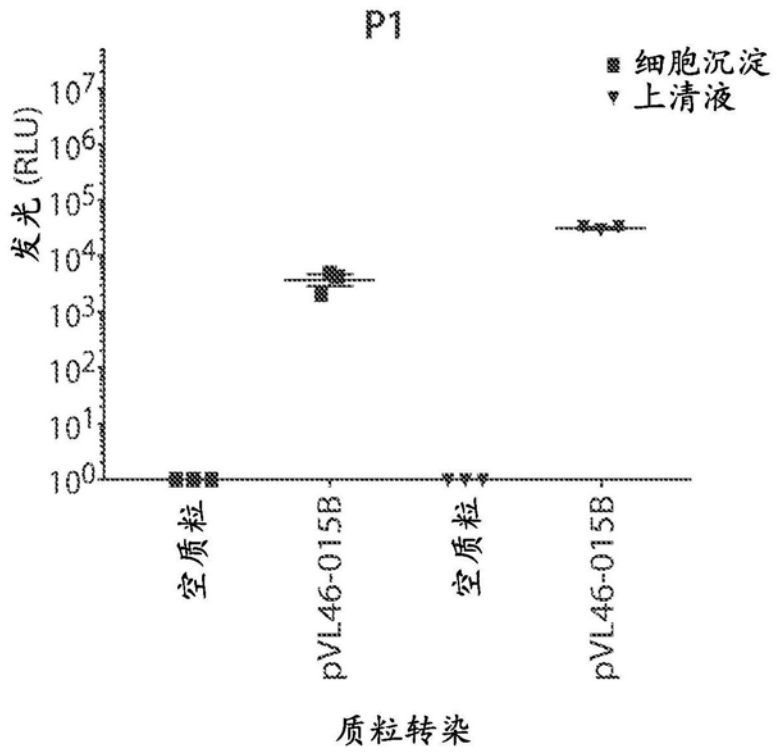

P1
细胞沉淀
上清液
发光 (RLU)
空质粒
pVL46-015B
空质粒
pVL46-015B
质粒转染

图11B

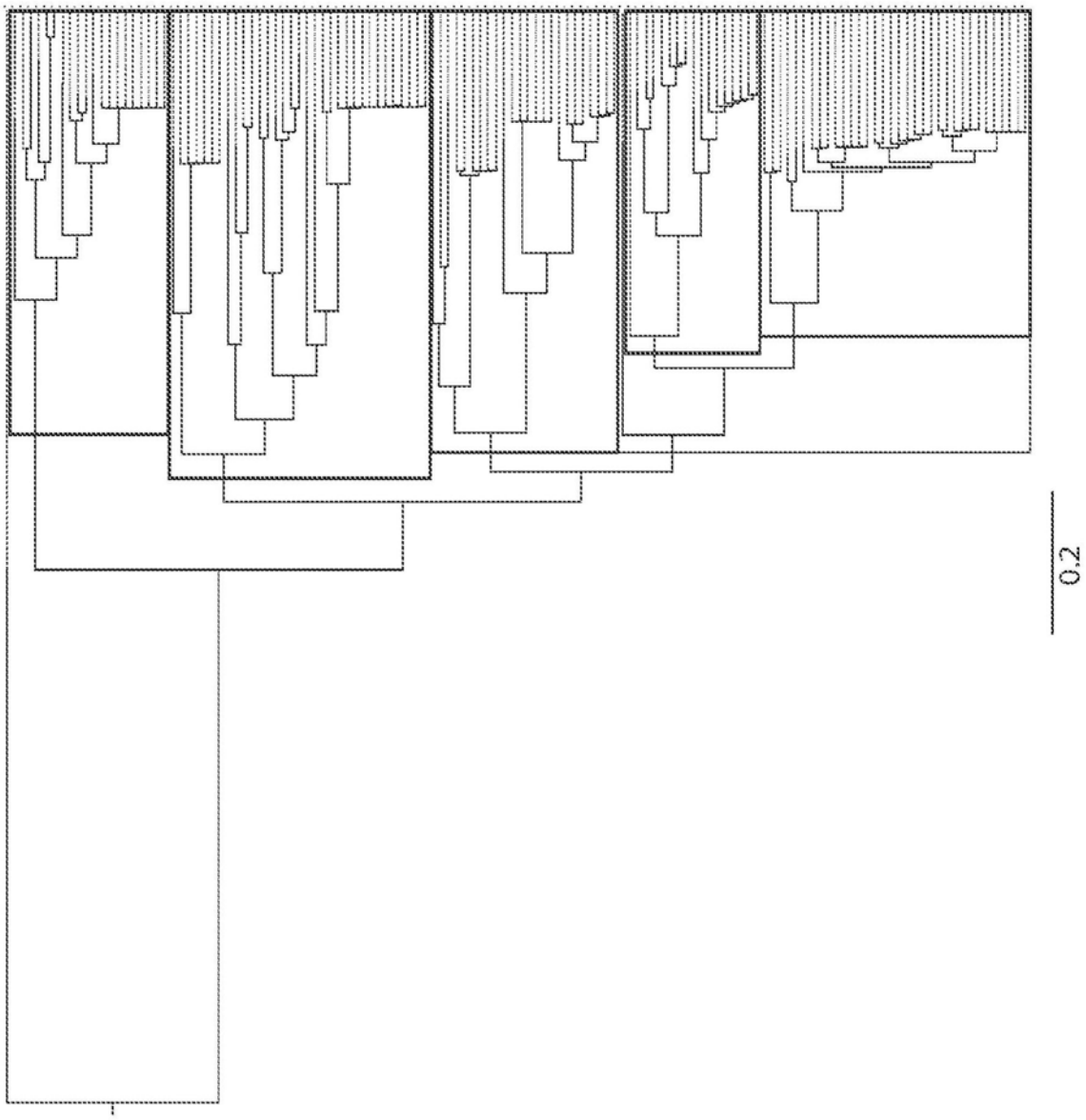


图11C

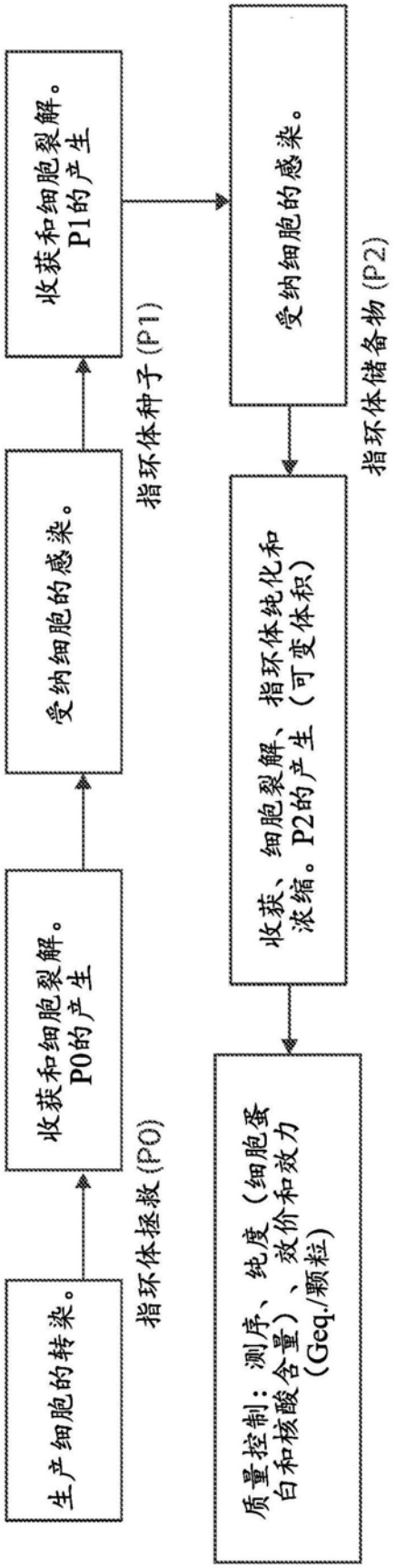

生产细胞的转染。
指环体拯救(P0)
收获和细胞裂解。
P0的产生
受纳细胞的感染。
指环体种子(P1)
收获和细胞裂解。
P1的产生
受纳细胞的感染。
指环体储备物(P2)
收获、细胞裂解、指环体纯化和
浓缩。P2的产生（可变体积）
质量控制：测序、纯度（细胞蛋
白和核酸含量）、效价和效力
(Geq./颗粒)

图12

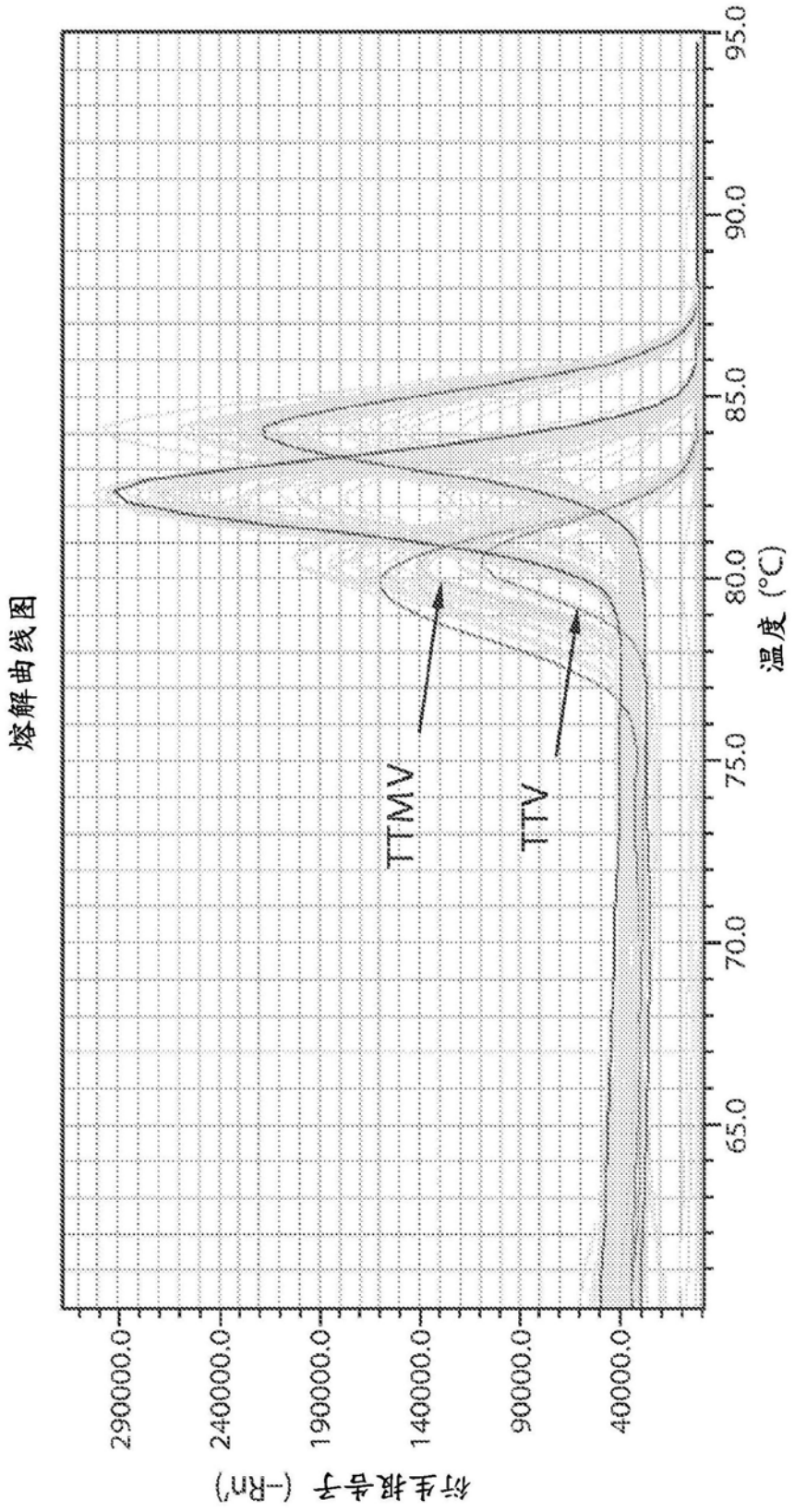

熔解曲线图
衍生报告子 (-RFU)
290000.0
240000.0
190000.0
140000.0
90000.0
40000.0
65.0
70.0
75.0
80.0
85.0
90.0
95.0
温度 (°C)
TTMV
TTV

图13

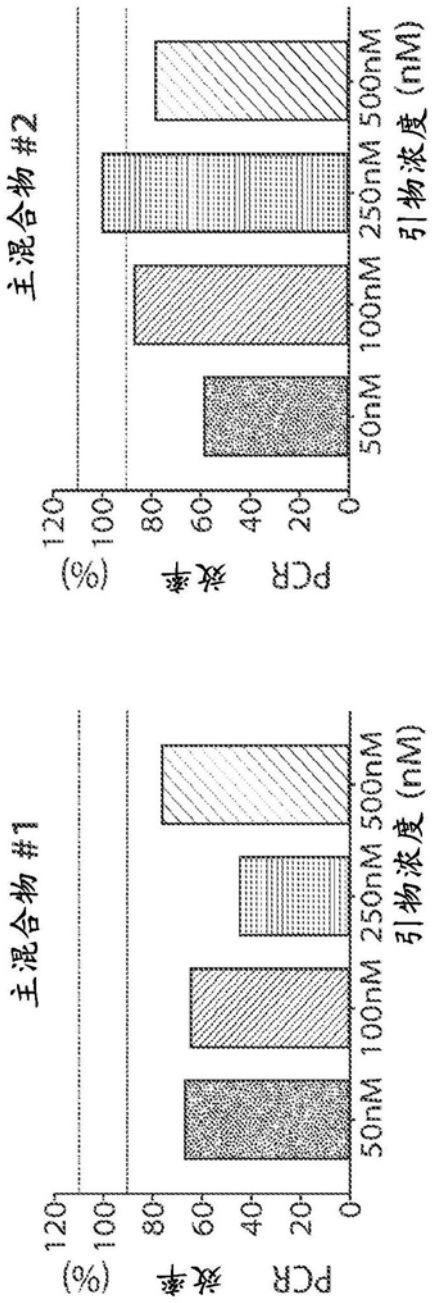

主混合物 #2
主混合物 #1
PCR 效率 (%)
0
20
40
60
80
100
120
50nM
100nM
250nM
500nM
引物浓度 (nM)

图14

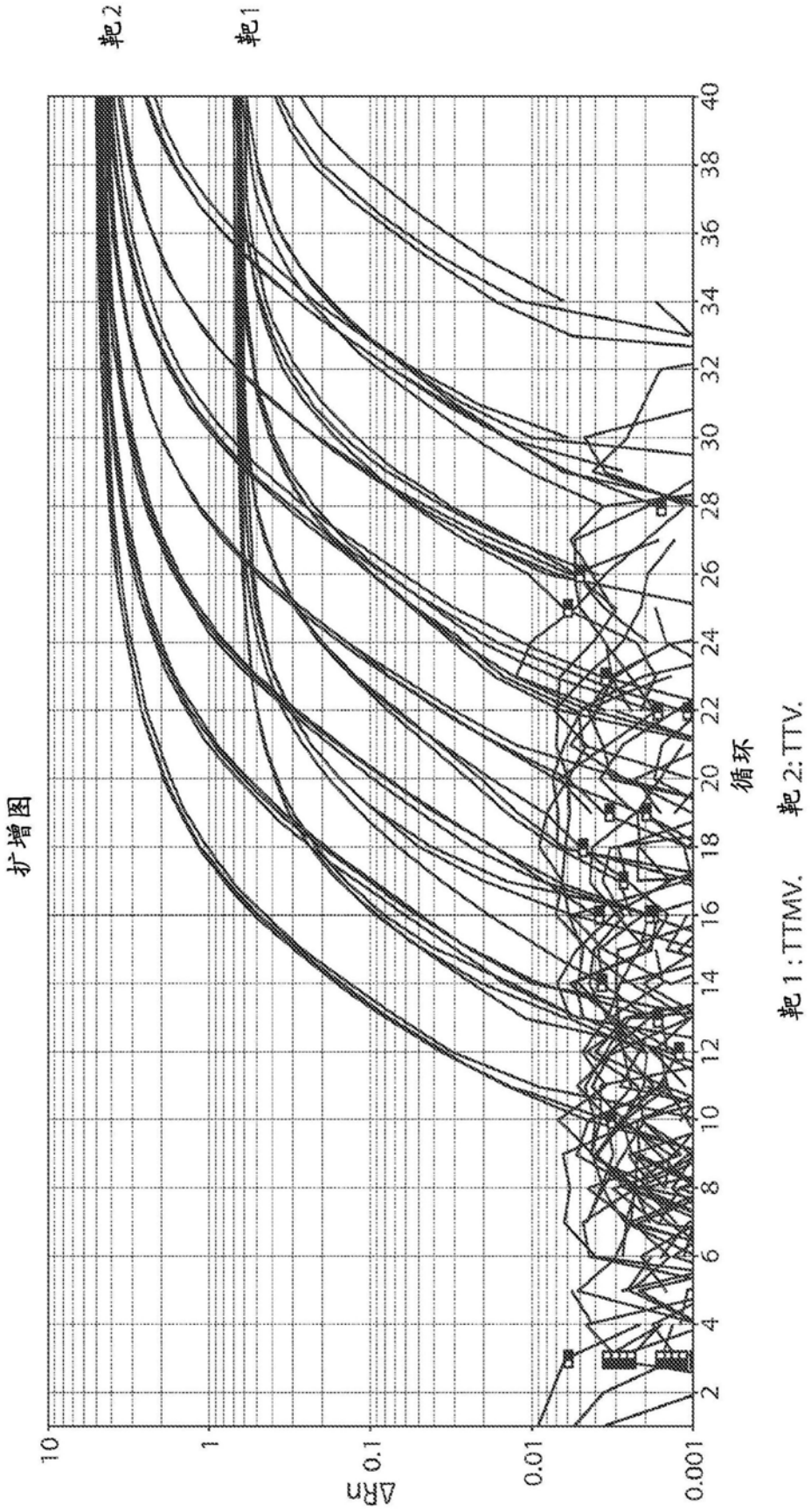


图15

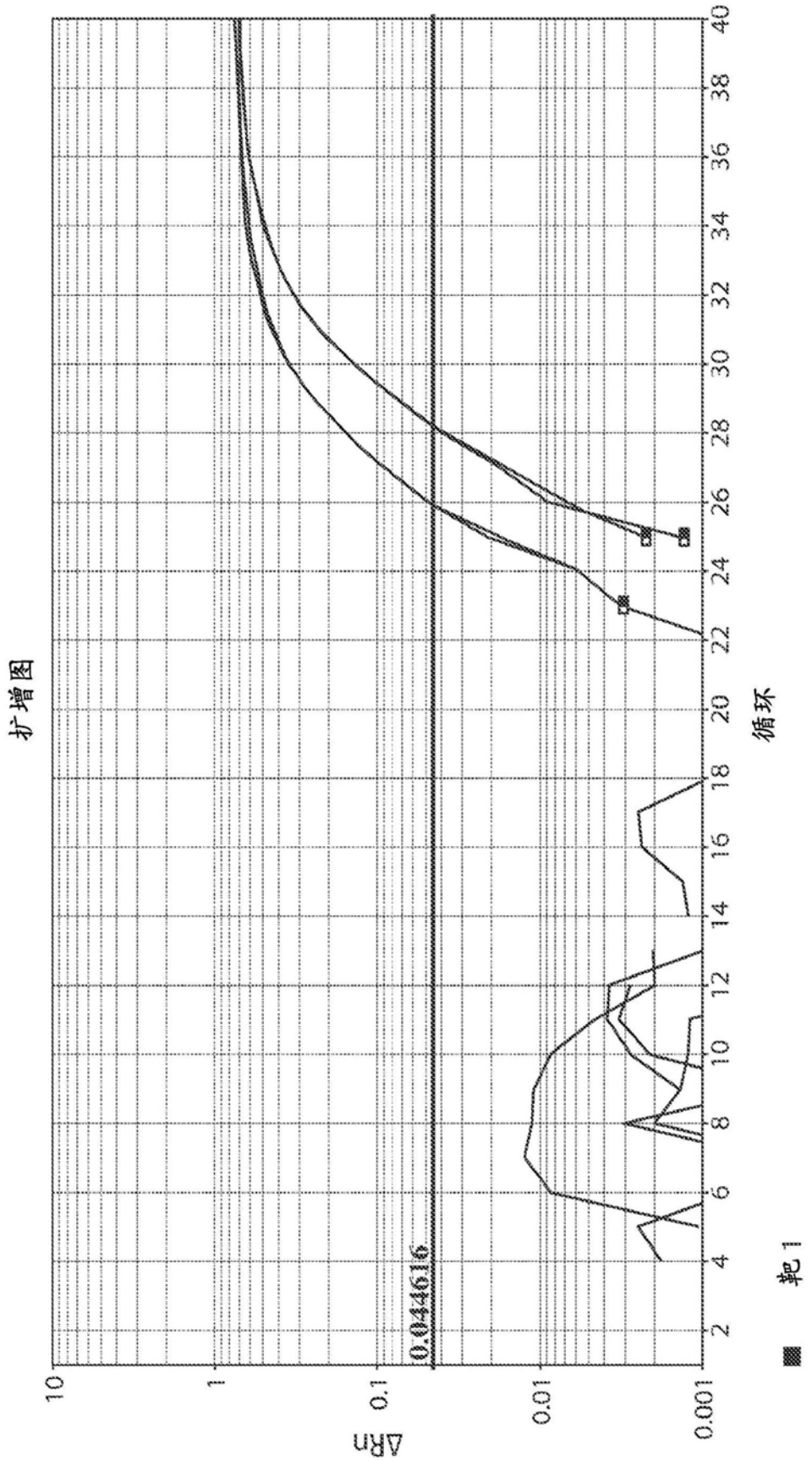

扩增图
ΔRn
循环
0.0446116
靶 1
10
1
0.1
0.01
0.001
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
38
40

图16A

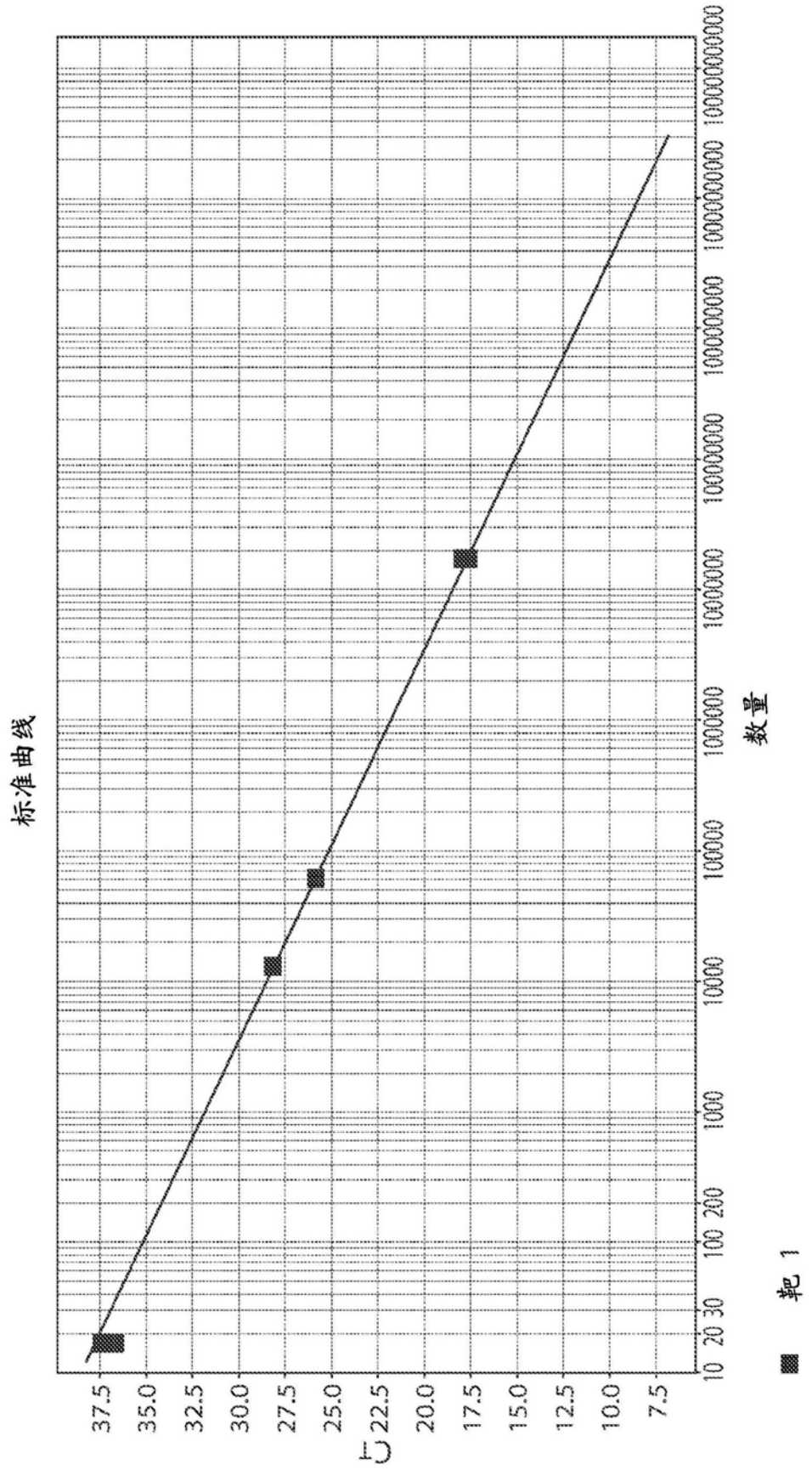

标准曲线
Ct
37.5
35.0
32.5
30.0
27.5
25.0
22.5
20.0
17.5
15.0
12.5
10.0
7.5
10 20 30 100 200 1000 10000 100000 1000000 10000000 100000000 1000000000 10000000000
数量
靶 1

图16B

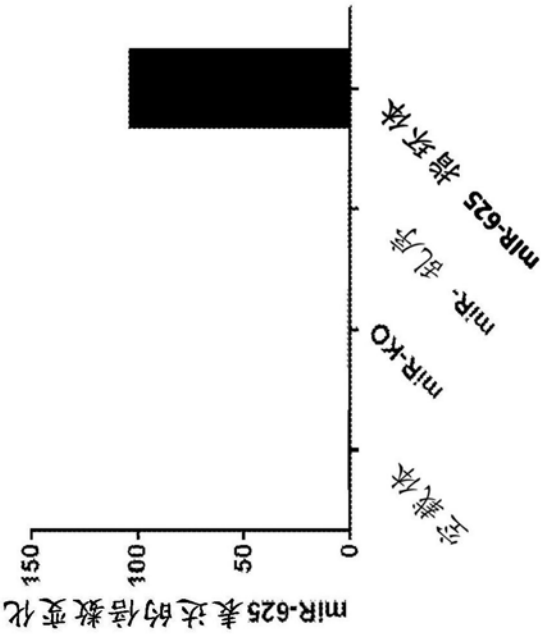


图17

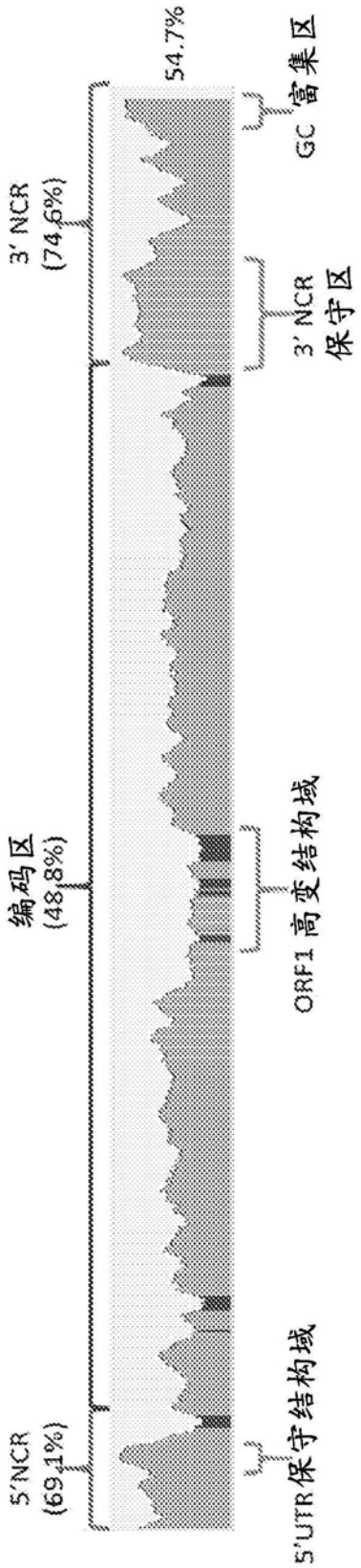

54.7%
GC 富集区
3' NCR
(74.6%)
3' NCR
保守区
编码区
(48.8%)
ORF1 高变结构域
5'NCR
(69.1%)
5'UTR保守结构域

图18

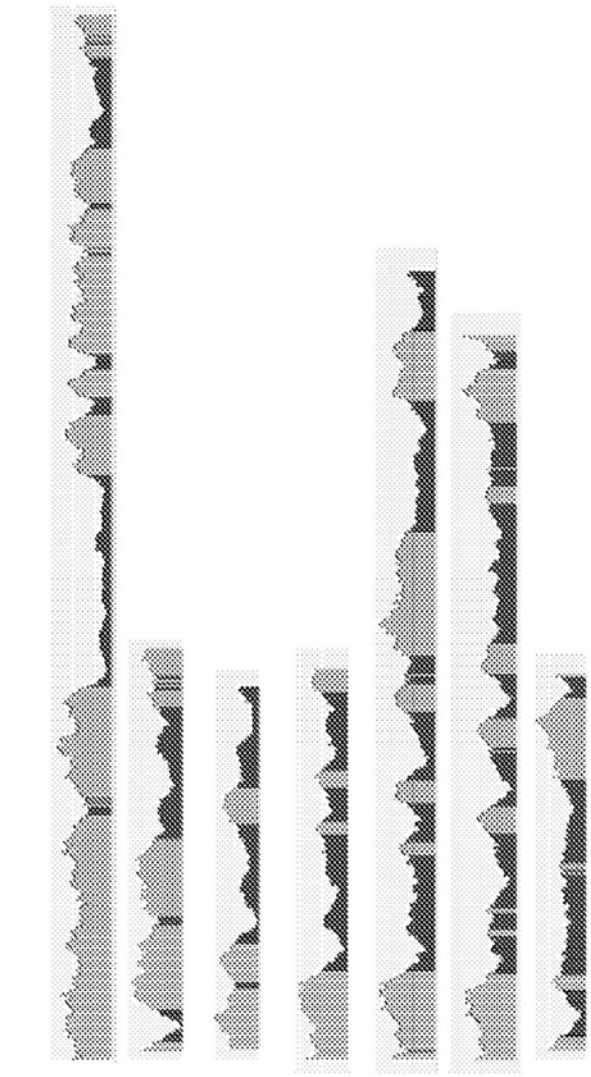

DNA AA
ORF 1 49.6% 34.8%
ORF 1/1 49.6% 34.2%
ORF 1/2 49.4% 30.2%
ORF 2 47.8% 34.2%
ORF 2/2 48.9% 32.0%
ORF 2/3 48.8% 32.9%
ORF 2t/3* 48.9% 37.5%

图19

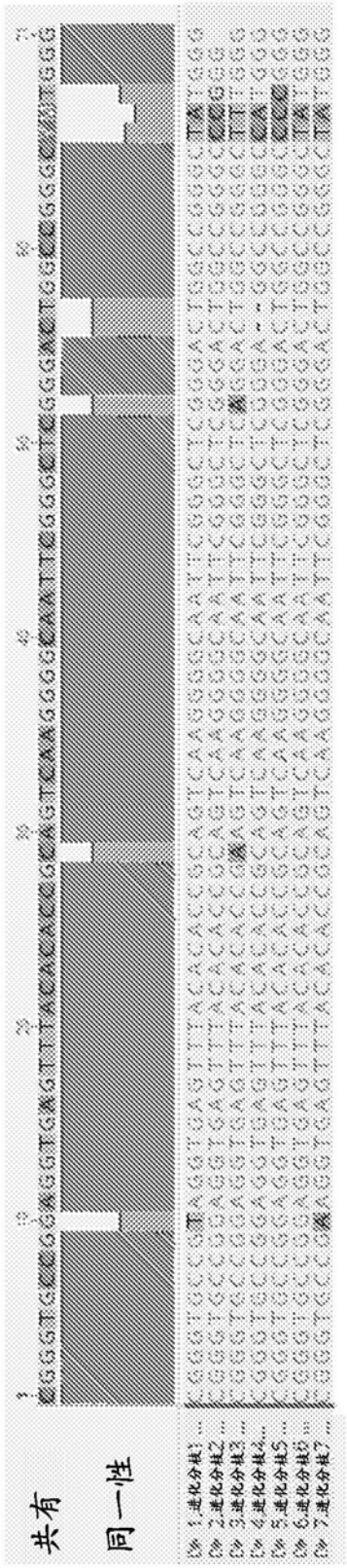


图20

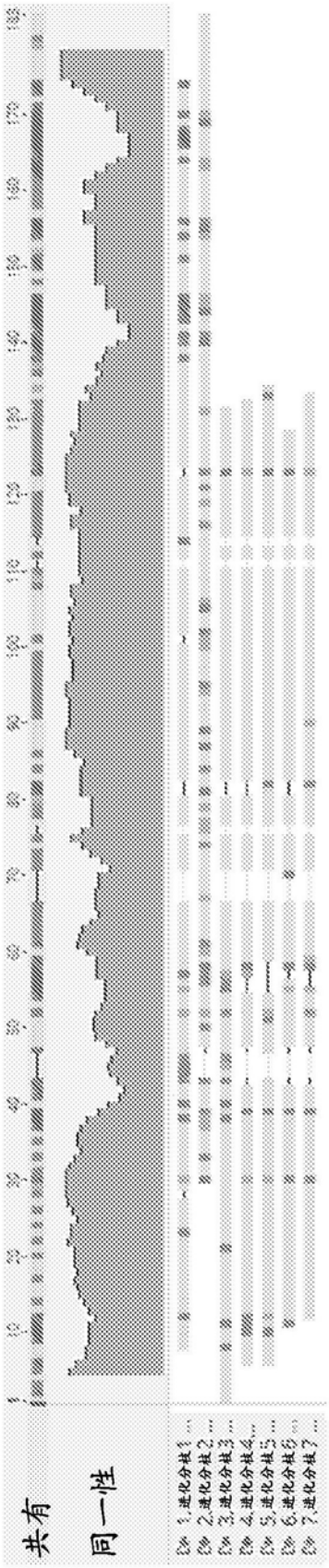


图21

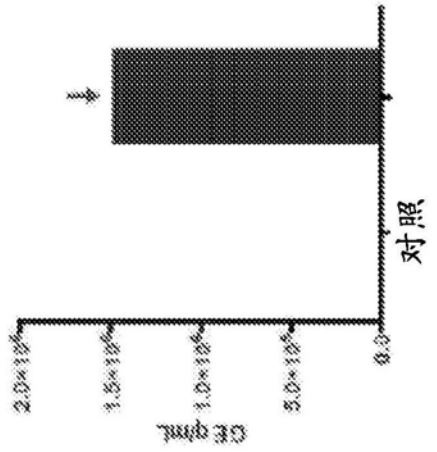

对照
0.0
5.0×10⁴
1.0×10⁵
1.5×10⁵
2.0×10⁵

图22

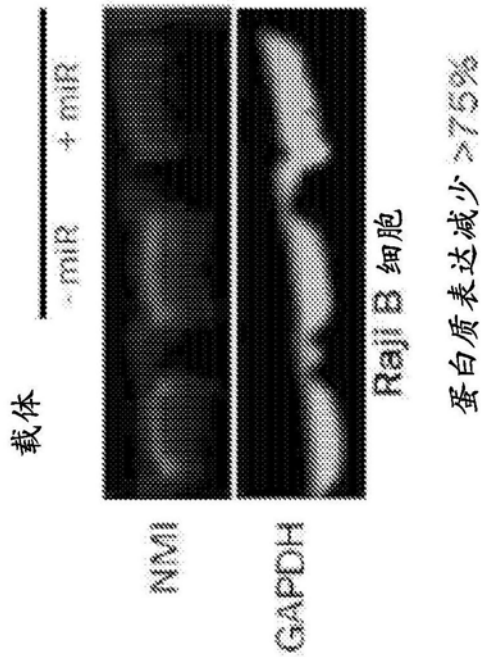

载体
- miR
+ miR
NMI
GAPDH
Raji B 细胞
蛋白质表达减少 >75%

图23

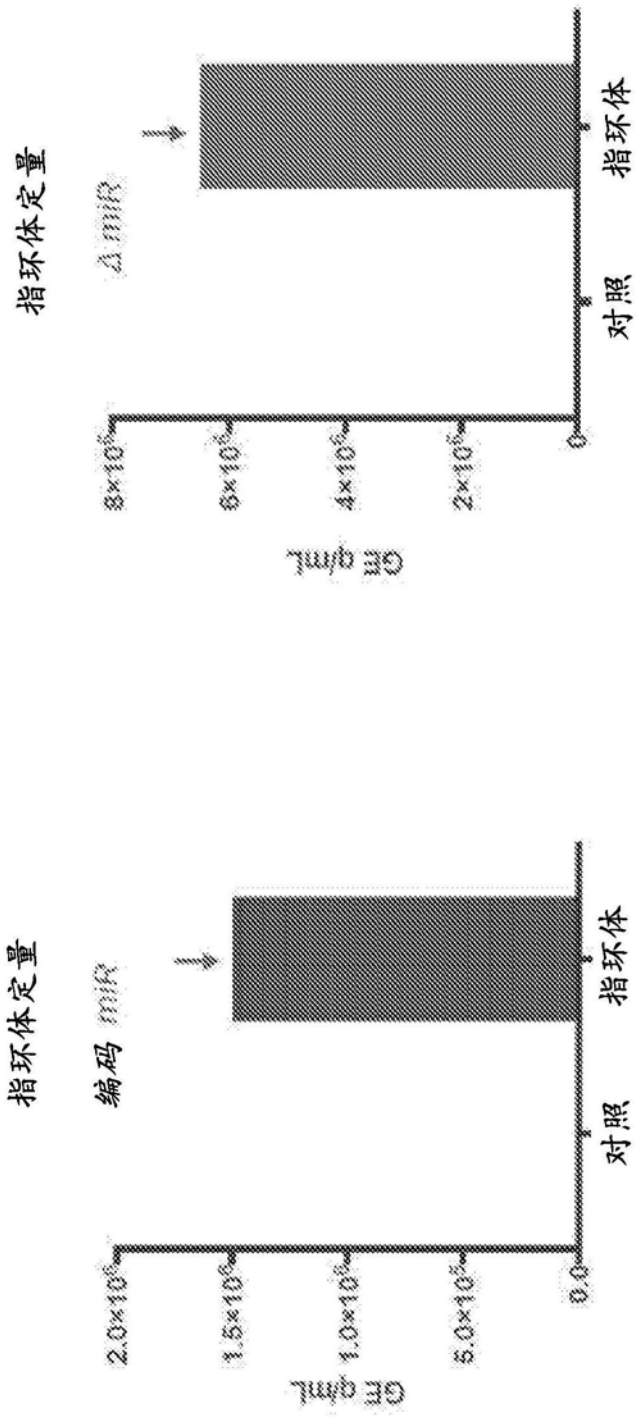

指环体定量
Δ miR
8×10⁶
6×10⁶
4×10⁶
2×10⁶
0
对照
指环体
指环体定量
编码 miR
2.0×10⁶
1.5×10⁶
1.0×10⁶
5.0×10⁵
0.0
对照
指环体

图24

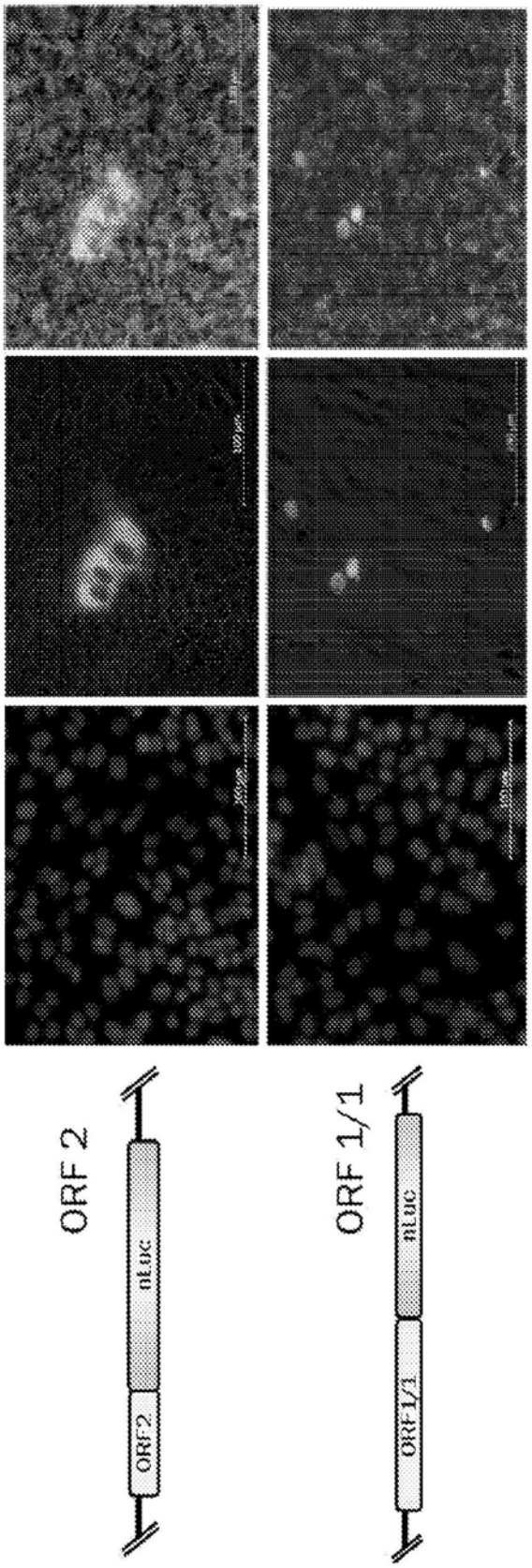


图25A

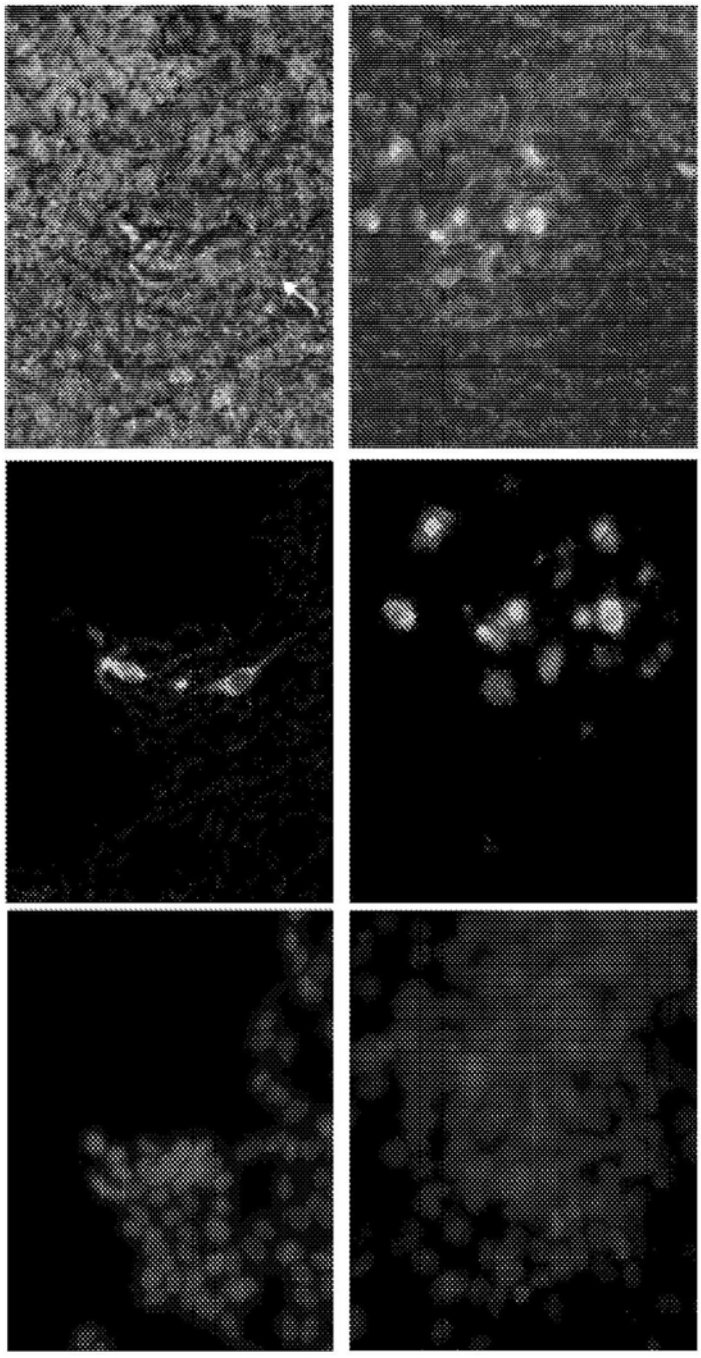


图25B

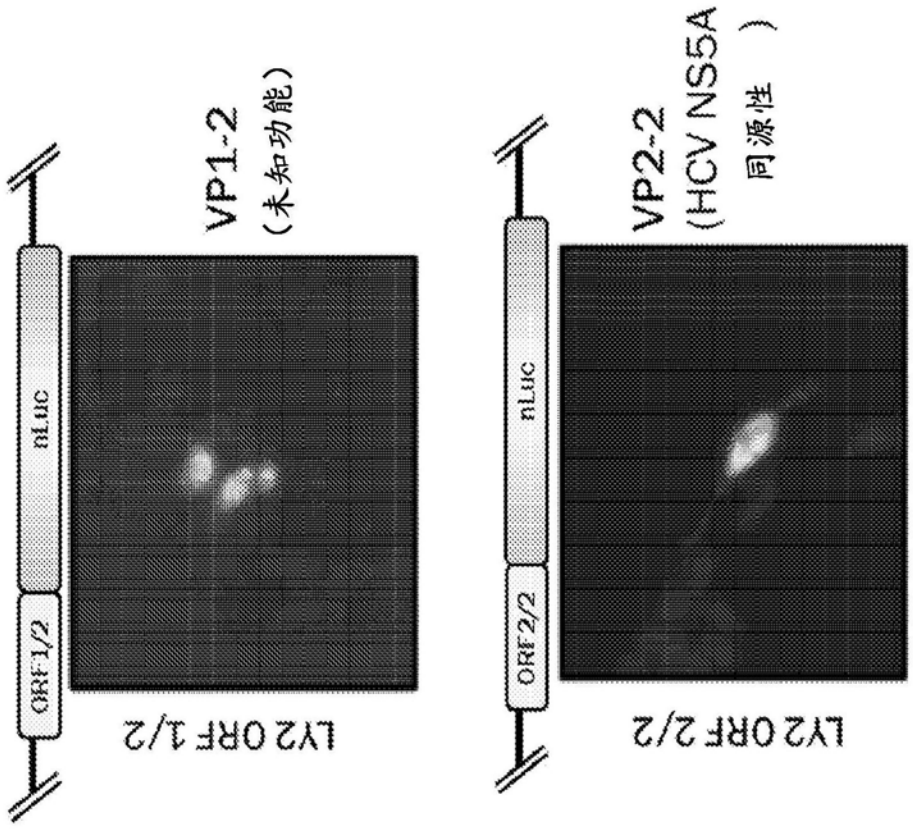

VP1-2
（未知功能）
nLuc
ORF1/2
LY2 ORF 1/2
VP2-2
（HCV NS5A
同源性 ）
nLuc
ORF2/2
LY2 ORF 2/2

图25C

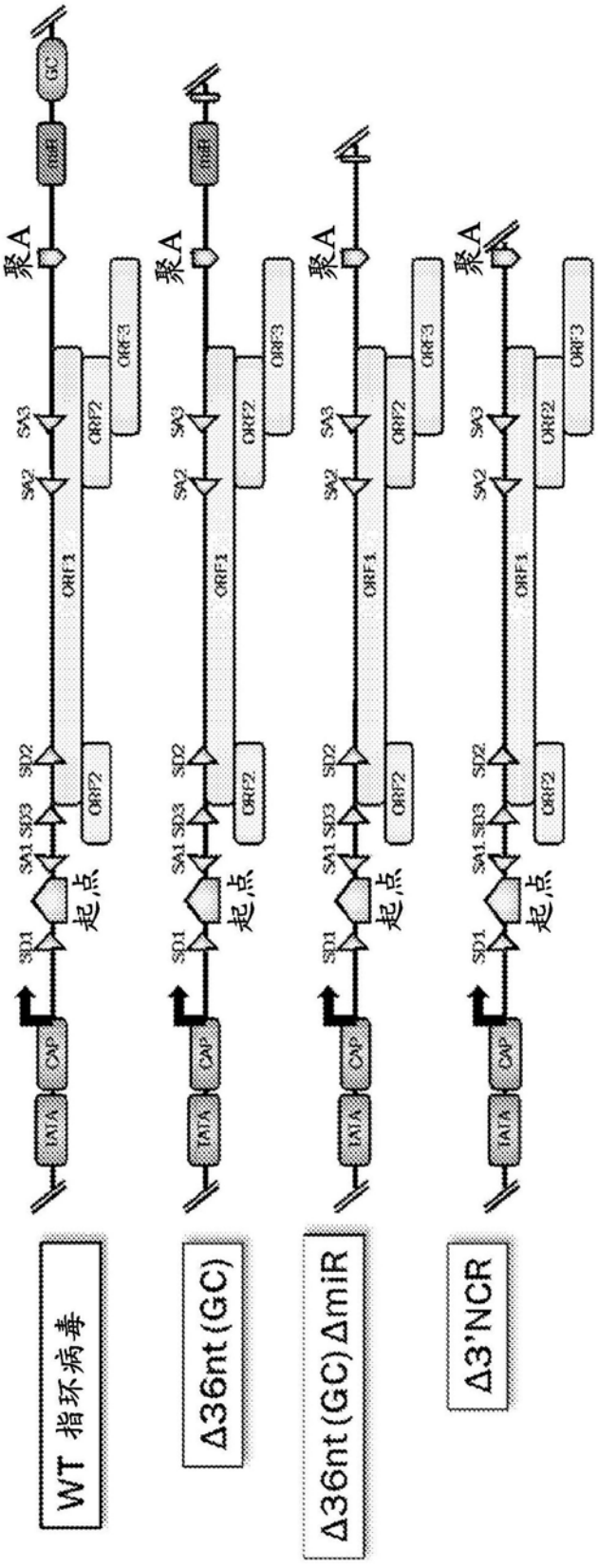


图26

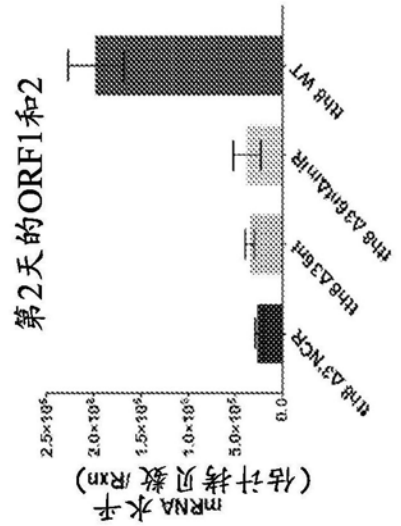

第2天的ORF1和2
mRNA水平
(估计拷贝数/Rxn)
2.5×10⁶
2.0×10⁶
1.5×10⁶
1.0×10⁶
5.0×10⁵
0.0
Ehb Δ3'NCR
Ehb Δ3Ent
Ehb Δ3EntΔnirR
Ehb WT

图27A

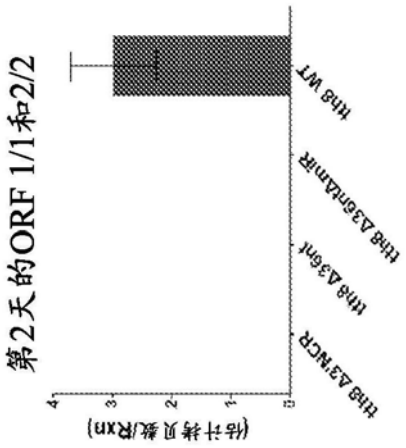

第2天的ORF 1/1和2/2
(估计拷贝数/Rxn)
4
3
2
1
0
Ehb Δ3'NCR
Ehb Δ3Ent
Ehb Δ3EntΔnirR
Ehb WT

图27B

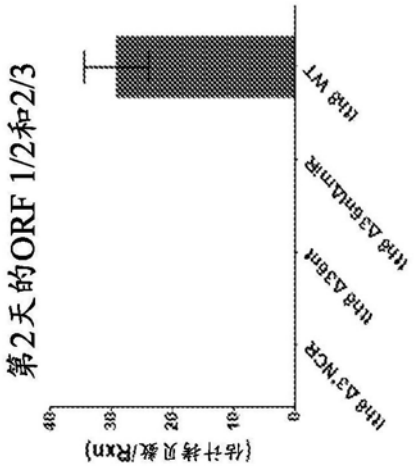

第2天的ORF 1/2和2/3
(估计拷贝数/Rxn)
0
10
20
30
40
Ithh8 A3'NCR
Ithh8 A3Sm
Ithh8 A3SmLmR
Ithh8 WT

图27C

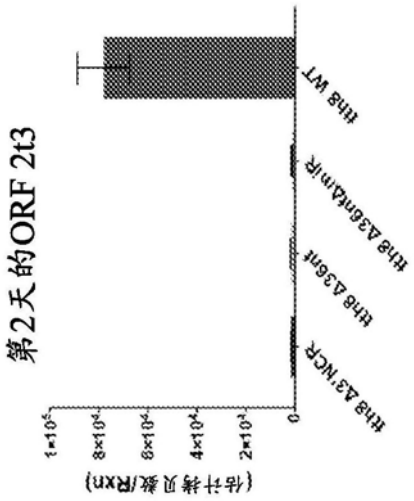

第2天的ORF 2t3
(估计拷贝数/Rxn)
0
2×10⁴
4×10⁴
6×10⁴
8×10⁴
1×10⁵
Ithh8 A3'NCR
Ithh8 A3Sm
Ithh8 A3SmLmR
Ithh8 WT

图27D

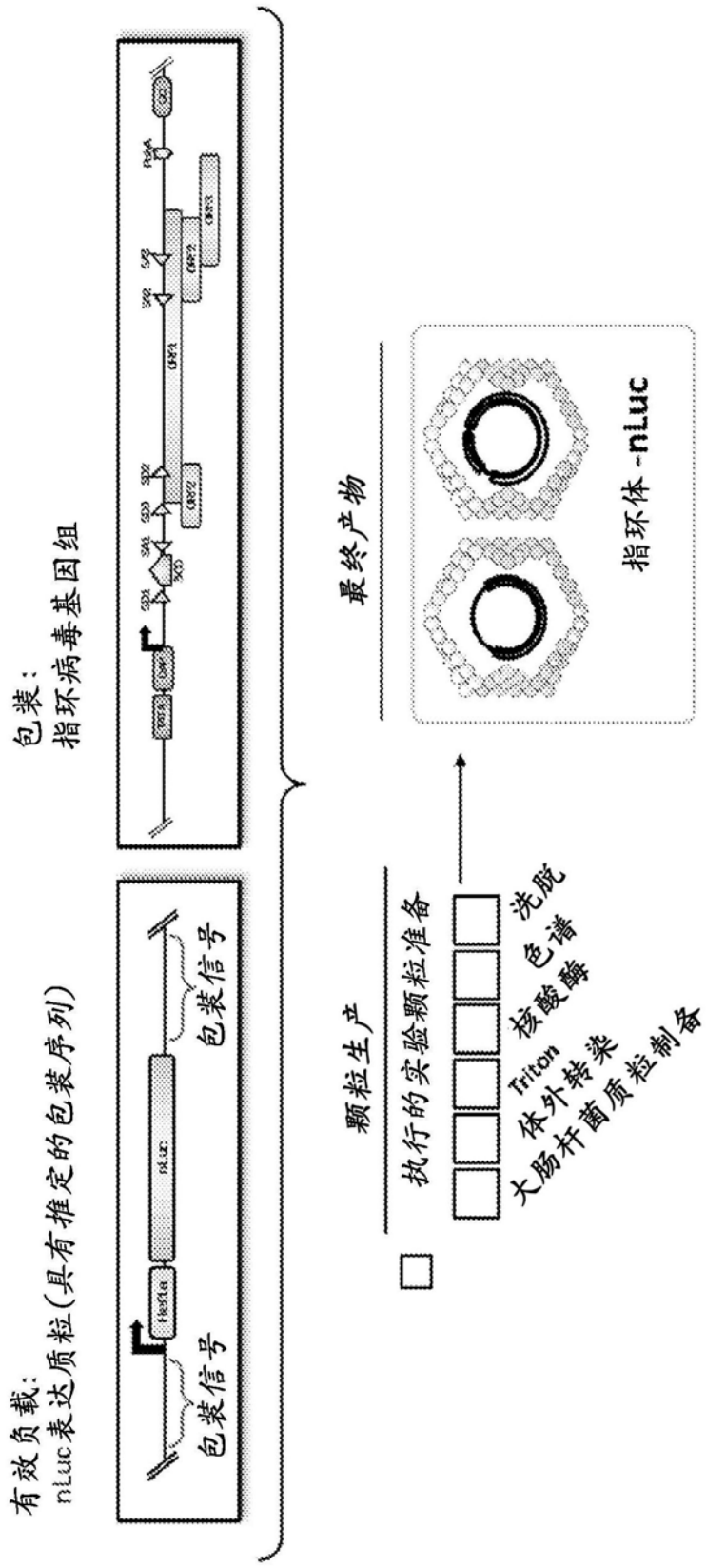

有效负载:
nLuc表达质粒(具有推定的包装序列)
包装信号
nLuc
包装信号
包装:
指环病毒基因组
颗粒生产
执行的实验颗粒准备
洗脱
色谱
核酸酶
Triton
体外转录
大肠杆菌质粒制备
最终产物
指环体-nLuc

图28A

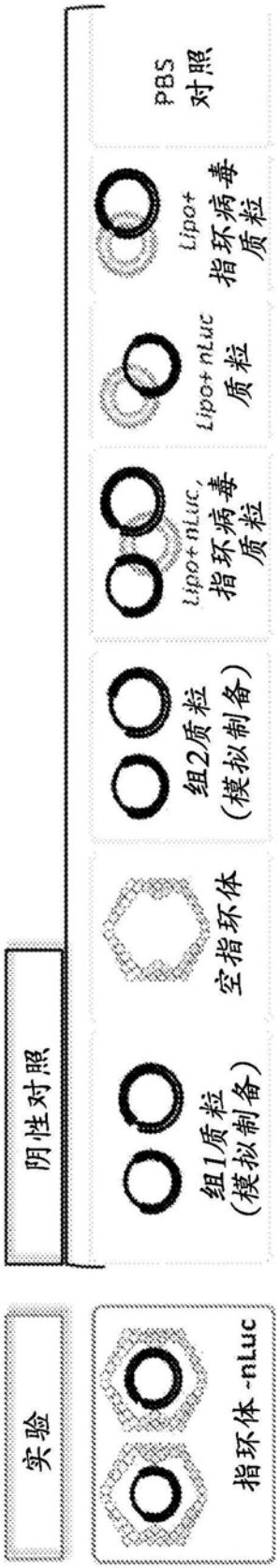

实验
指环体-nluc
阴性对照
组1质粒
(模拟制备)
空指环体
组2质粒
(模拟制备)
lipo+nluc,
指环病毒
质粒
lipo+nluc
质粒
lipo+
指环病毒
质粒
PBS
对照

图28B

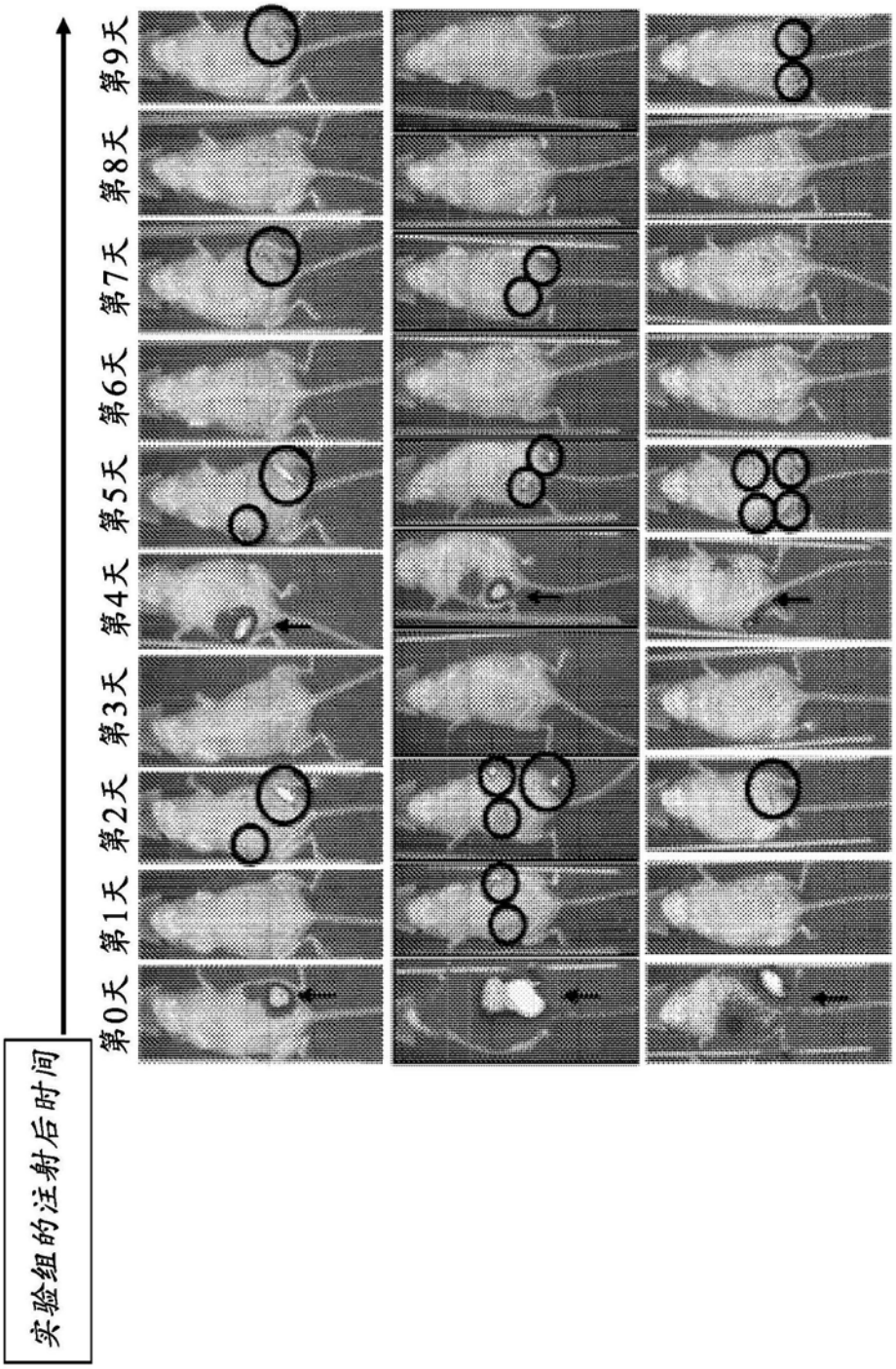


图29A

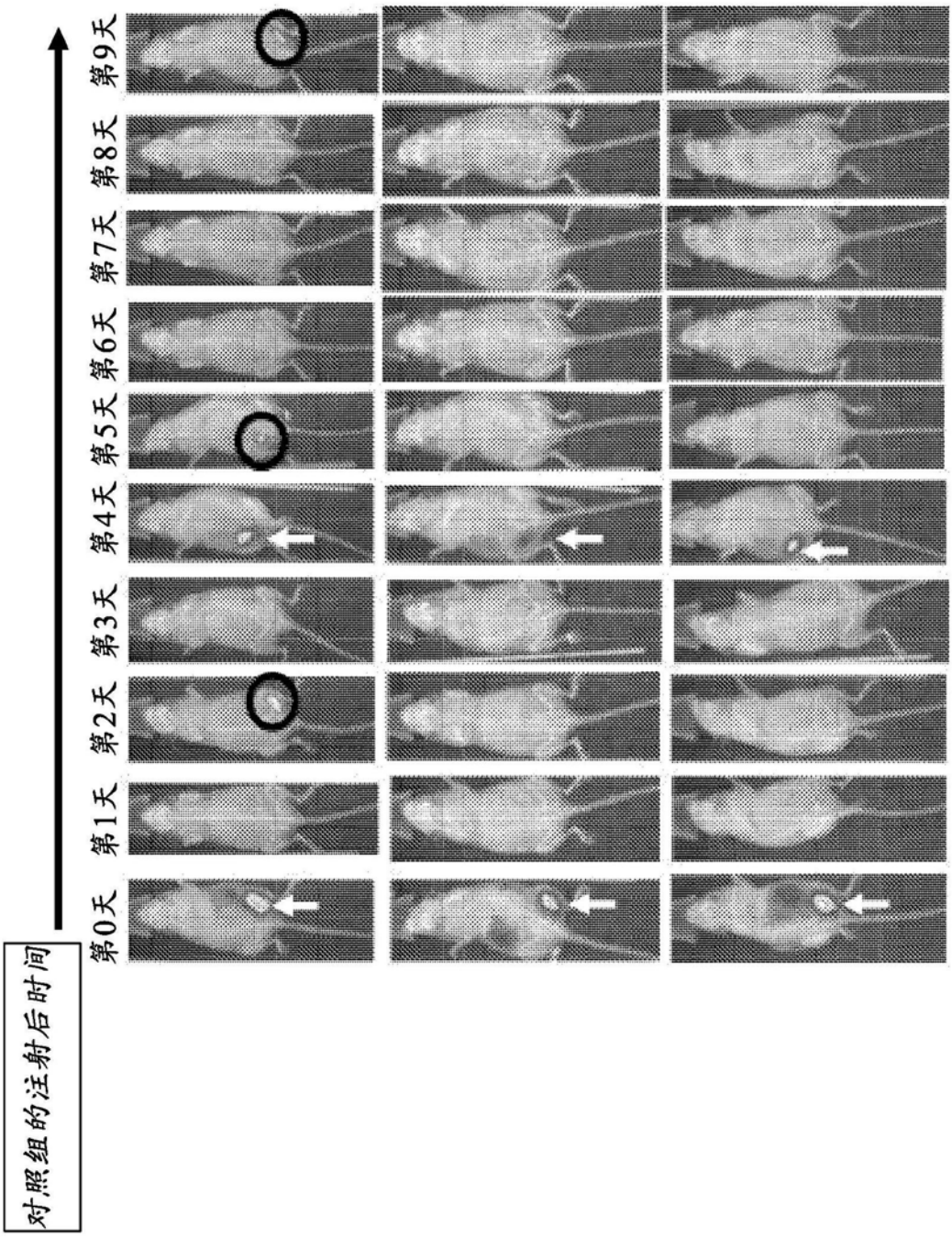


图29B

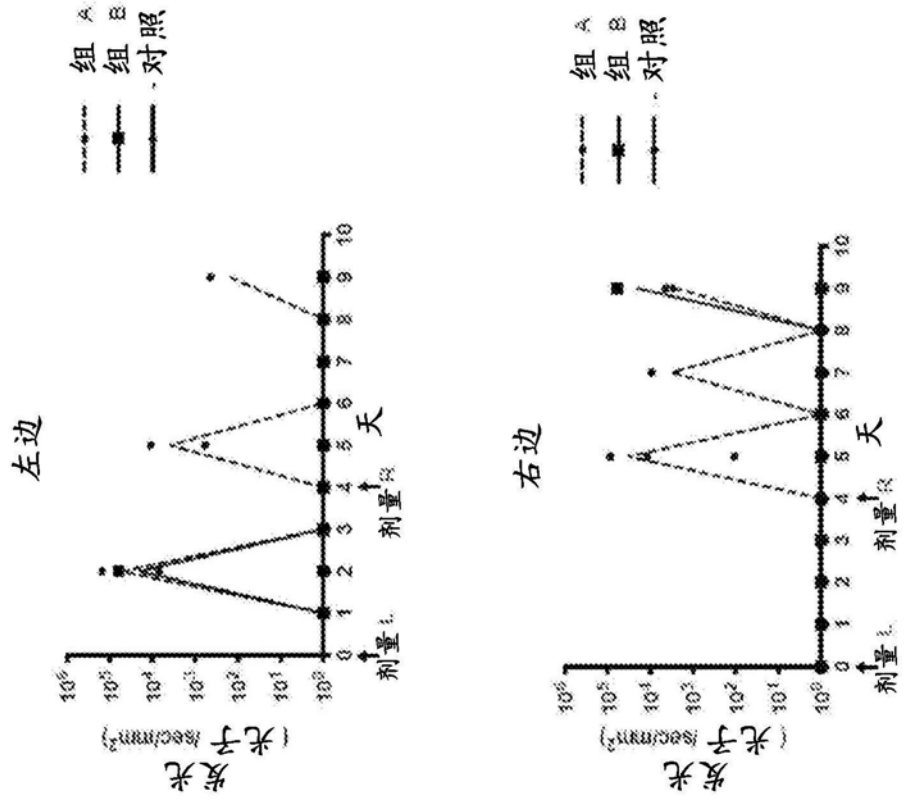


图29C

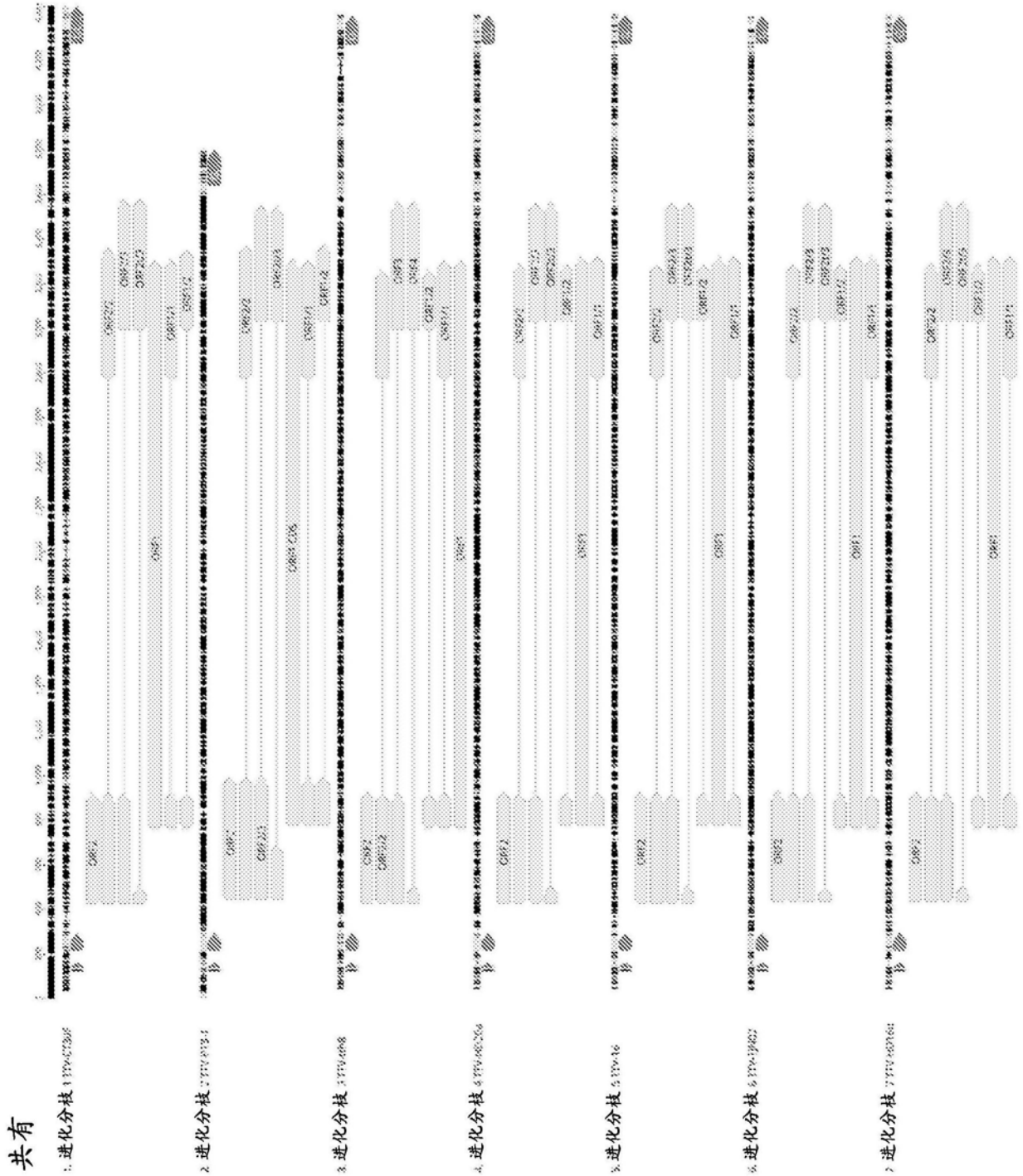


图29D

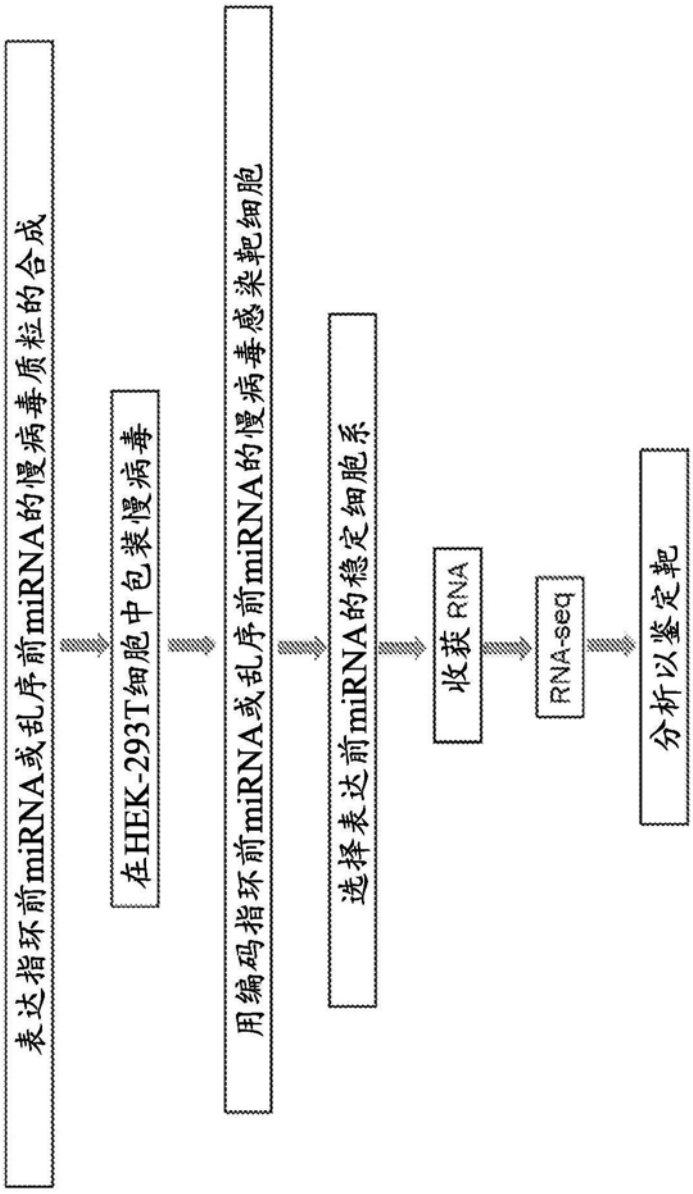

表达指环前miRNA或乱序前miRNA的慢病毒质粒的合成
在HEK-293T细胞中包装慢病毒
用编码指环前miRNA或乱序前miRNA的慢病毒感染靶细胞
选择表达前miRNA的稳定细胞系
收获RNA
RNA-seq
分析以鉴定靶

图30

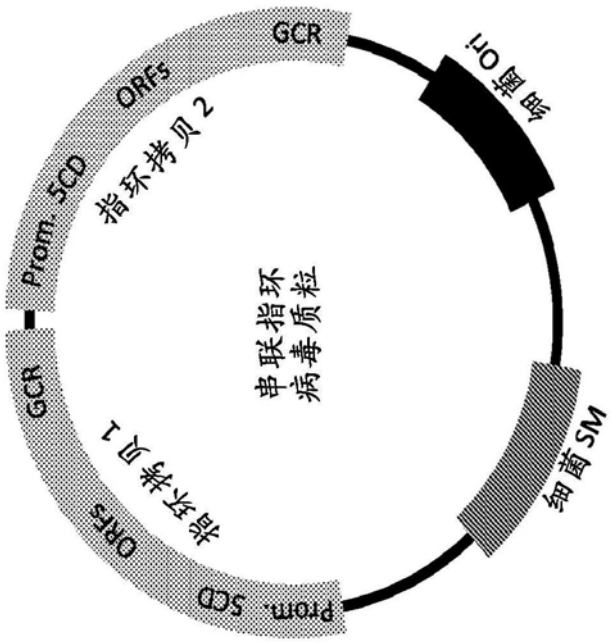

GCR
ORFs
指环拷贝 2
Prom. SCD
串联指环
病毒质粒
复制 Ori
细菌 SM
Prom. SCD
ORFs
指环拷贝 2
GCR

图31A

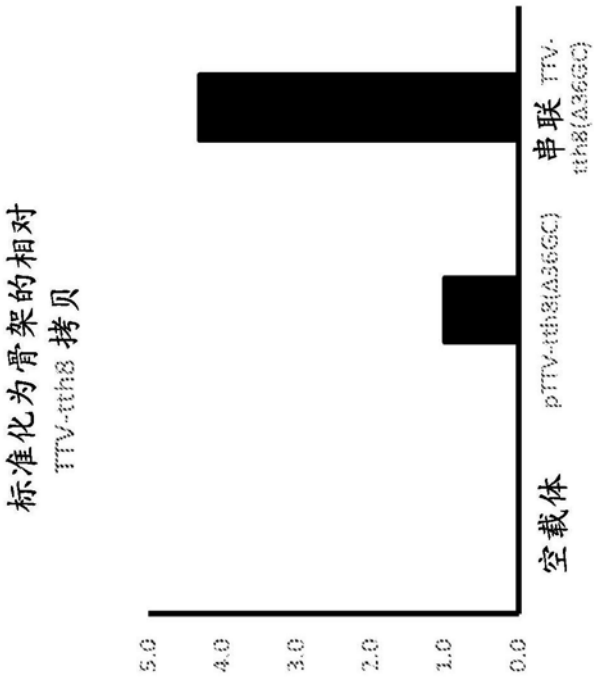

标准化为骨架的相对
TTV-tth8 拷贝
空载体
pTTV-tth8(A36GC)
串联 TTV-tth8(A36GC)
0.0
1.0
2.0
3.0
4.0
5.0

图31B

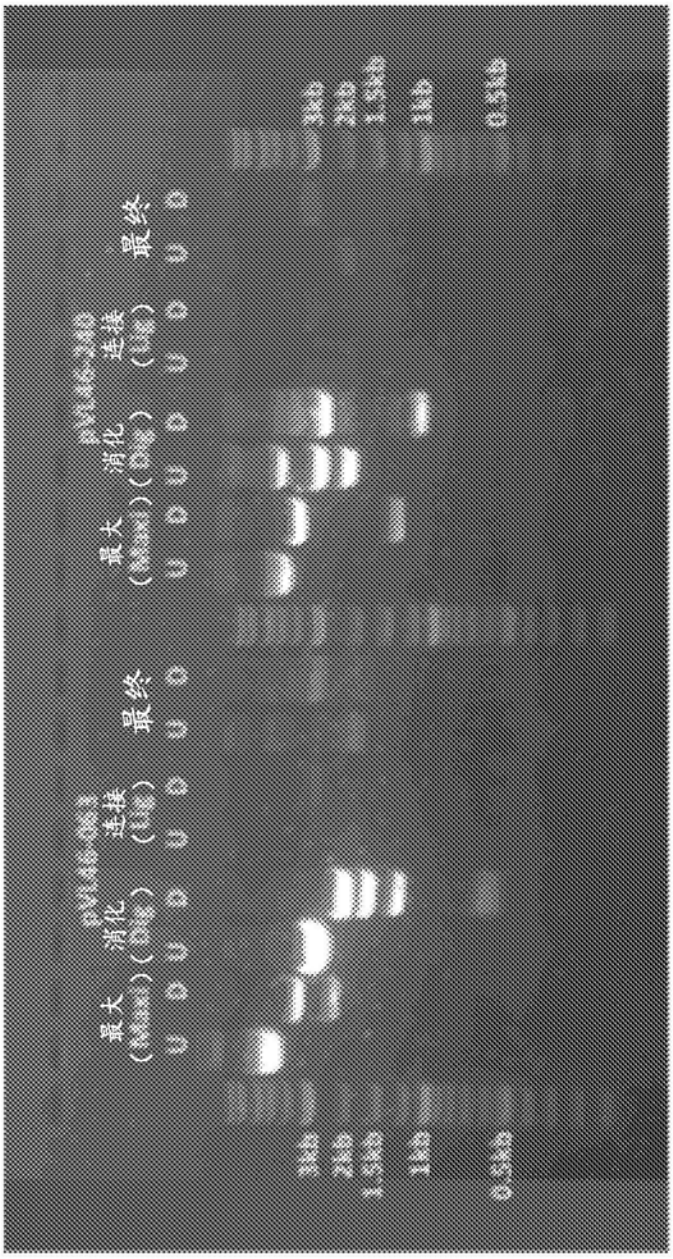

最大
(Mins)
消化
(Dec)
走捷
(Inc)
最终
pVU46-261
pVU46-240
0.5569
1.5569
219
149
219
149

图31C

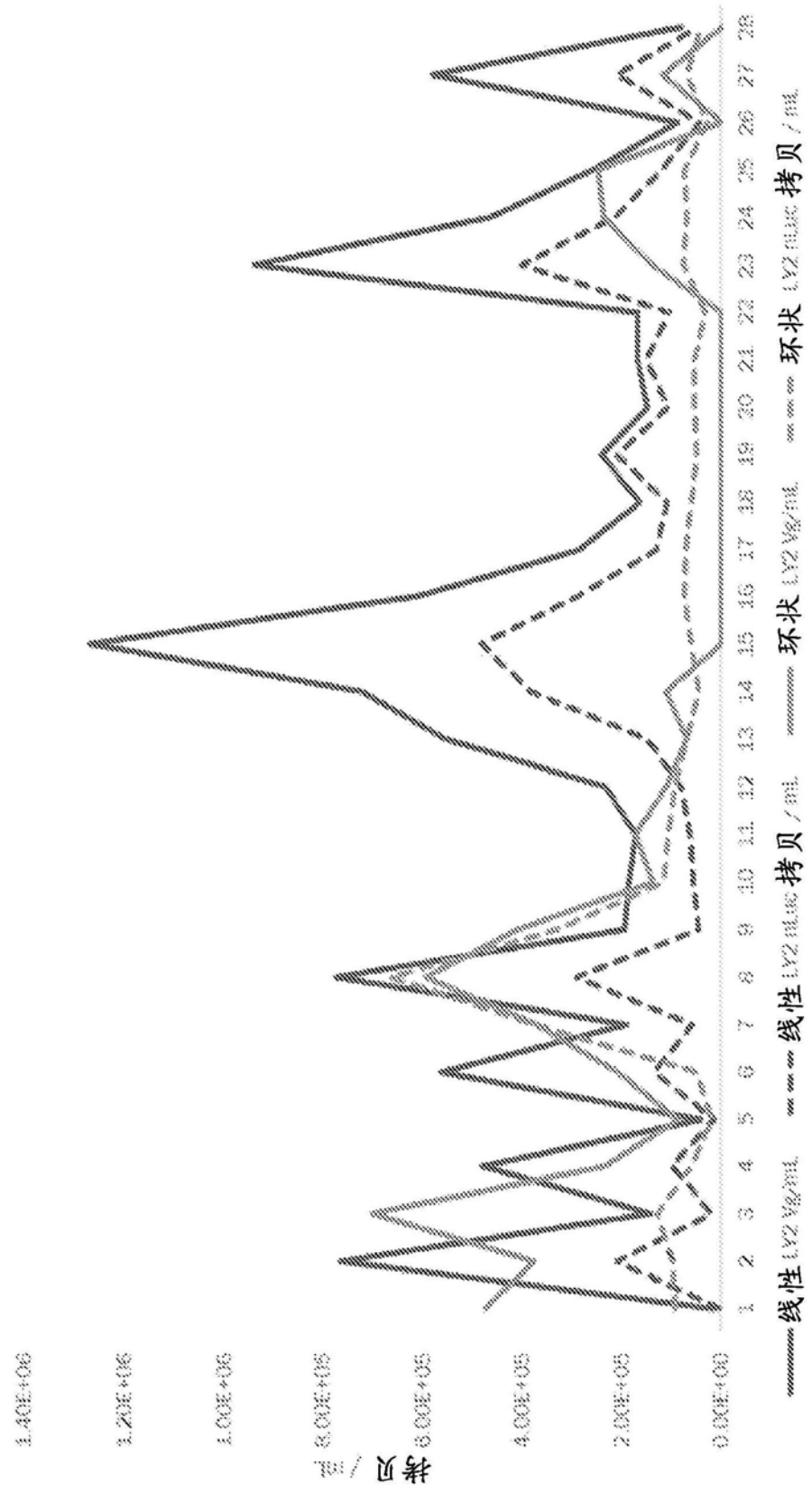

拷贝 / mL
1.40E+06
1.20E+06
1.00E+06
8.00E+05
6.00E+05
4.00E+05
2.00E+05
0.00E+00
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28
线性 LY2 Vg/mL
线性 LY2 mAb 拷贝 / mL
环状 LY2 Vg/mL
环状 LY2 mAb 拷贝 / mL

图31D

共有
同一性
1
10
20
30
36bp区域
1.AB064600.1 未分类 |TTV_ 进化分枝1 (细环病毒 DNA 全基因组,分离株 ...
2.AB064606.1 未分类 |TTV_ 进化分枝3
3.AF122913.1 未分类 |TTV_ 进化分枝3(TT病毒分离株 GH1, 全基因组)
4.AM712004.1 未分类 |TTV_ 进化分枝3(TT病毒 sa19932 全基因组)
5.FR751497.1 未分类 |TTV_ 进化分枝4(细环病毒全基因组,分离株 TT...
6.MH643122.1 未分类 |TTV_ 进化分枝4(指环病毒科物种分离株 cmc902,全基因组...
7.FR751498.1 未分类 |TTV_ 进化分枝5(细环病毒全基因组,分离株 TT...
8.FR648329.1 未分类 |TTV_ 进化分枝6(细环病毒全基因组,分离株 H...
9.FR751481.1 未分类 |TTV_ 进化分枝7(细环病毒全基因组,分离株 TT...

图32

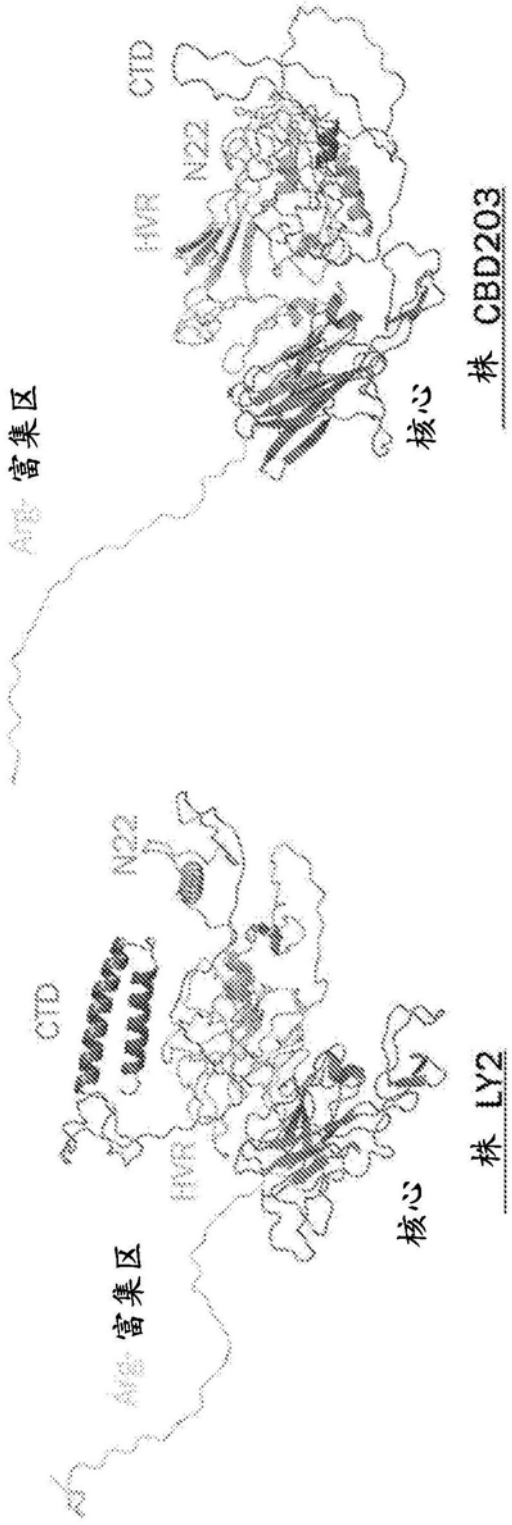


图33

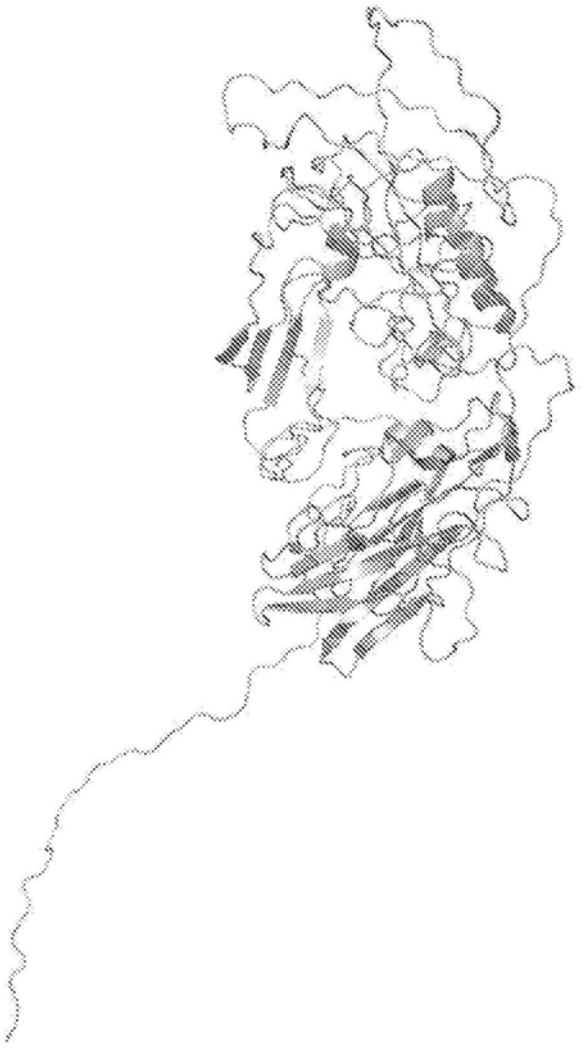


图34

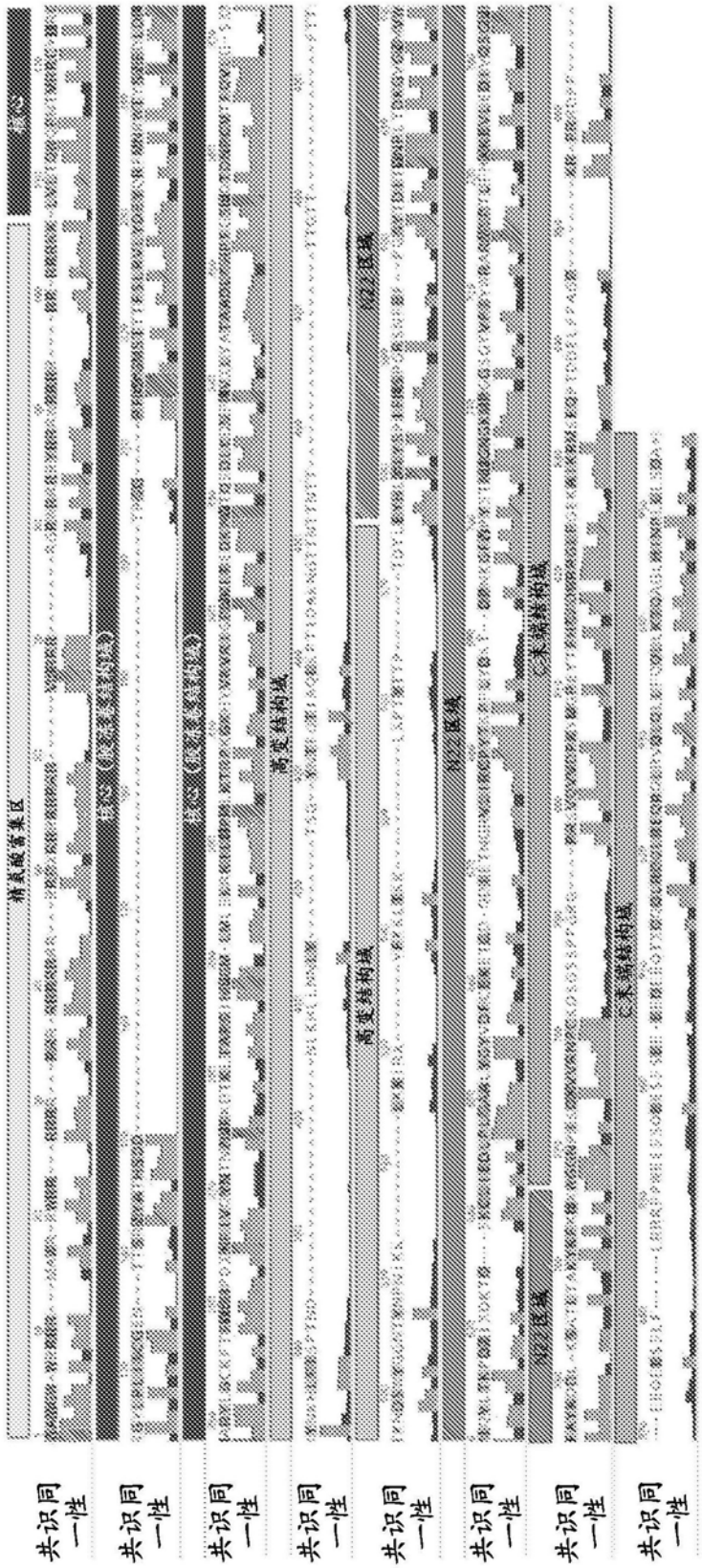

共识同一性
共识同一性
共识同一性
共识同一性
共识同一性
共识同一性
共识同一性
共识同一性
高变结构域
高变结构域
N2区域
N2区域
C末端结构域
C末端结构域

图35

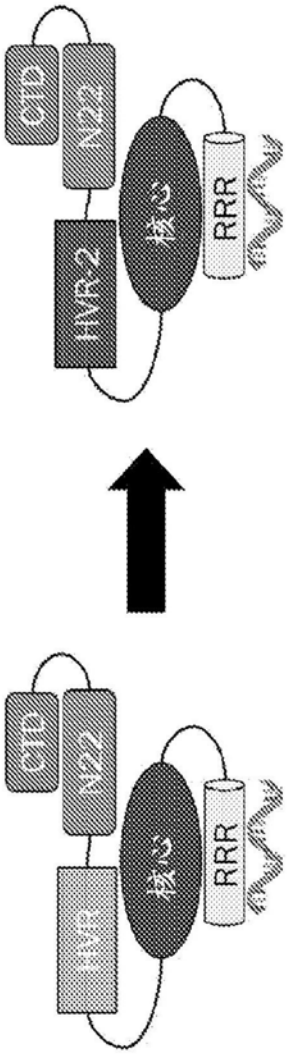


图36

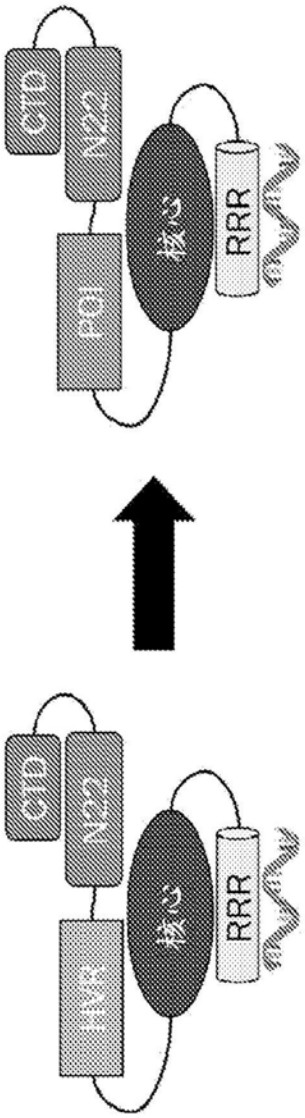


图37

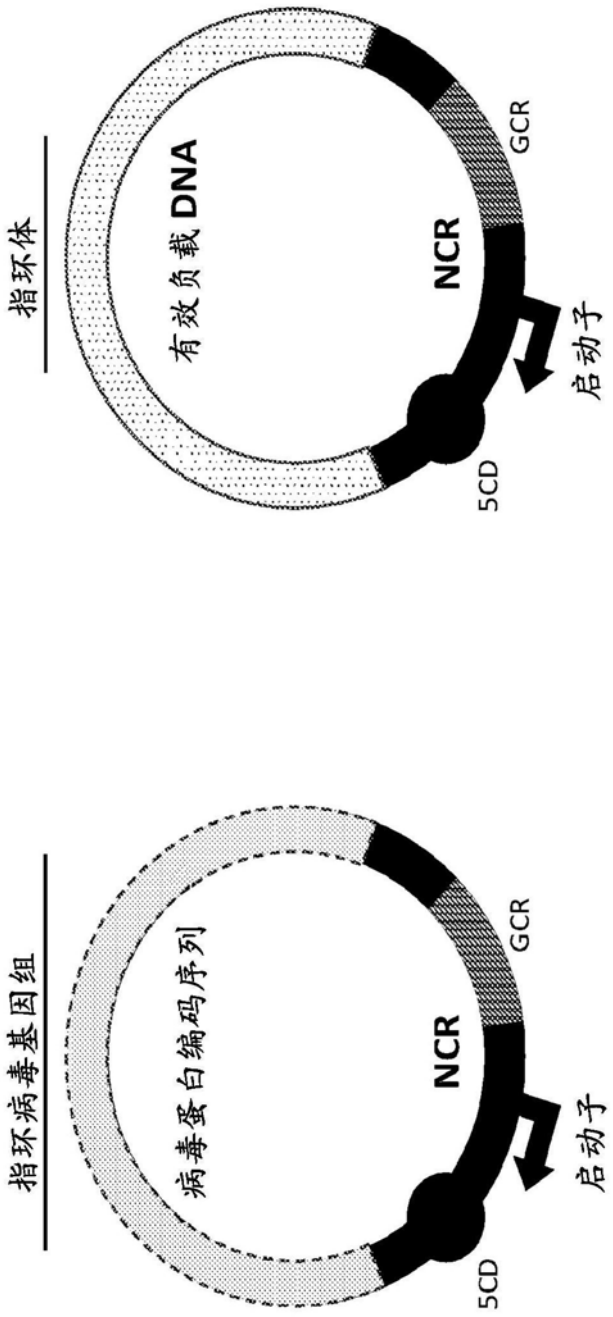


图38

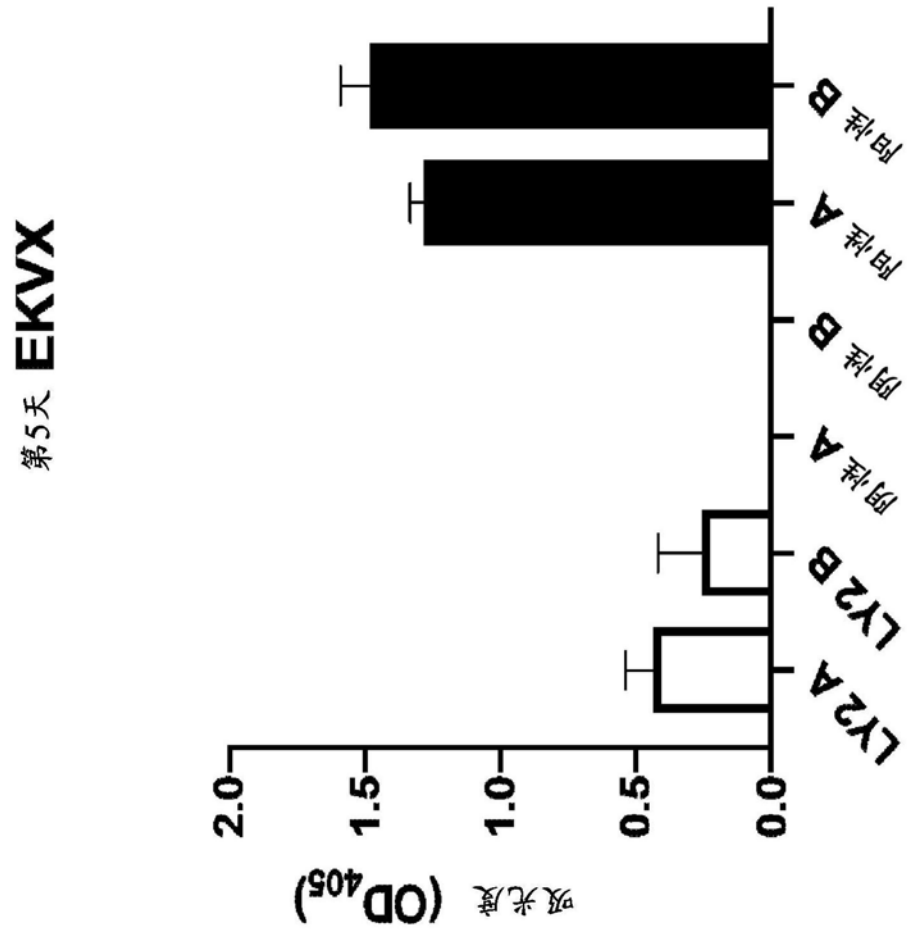


图39

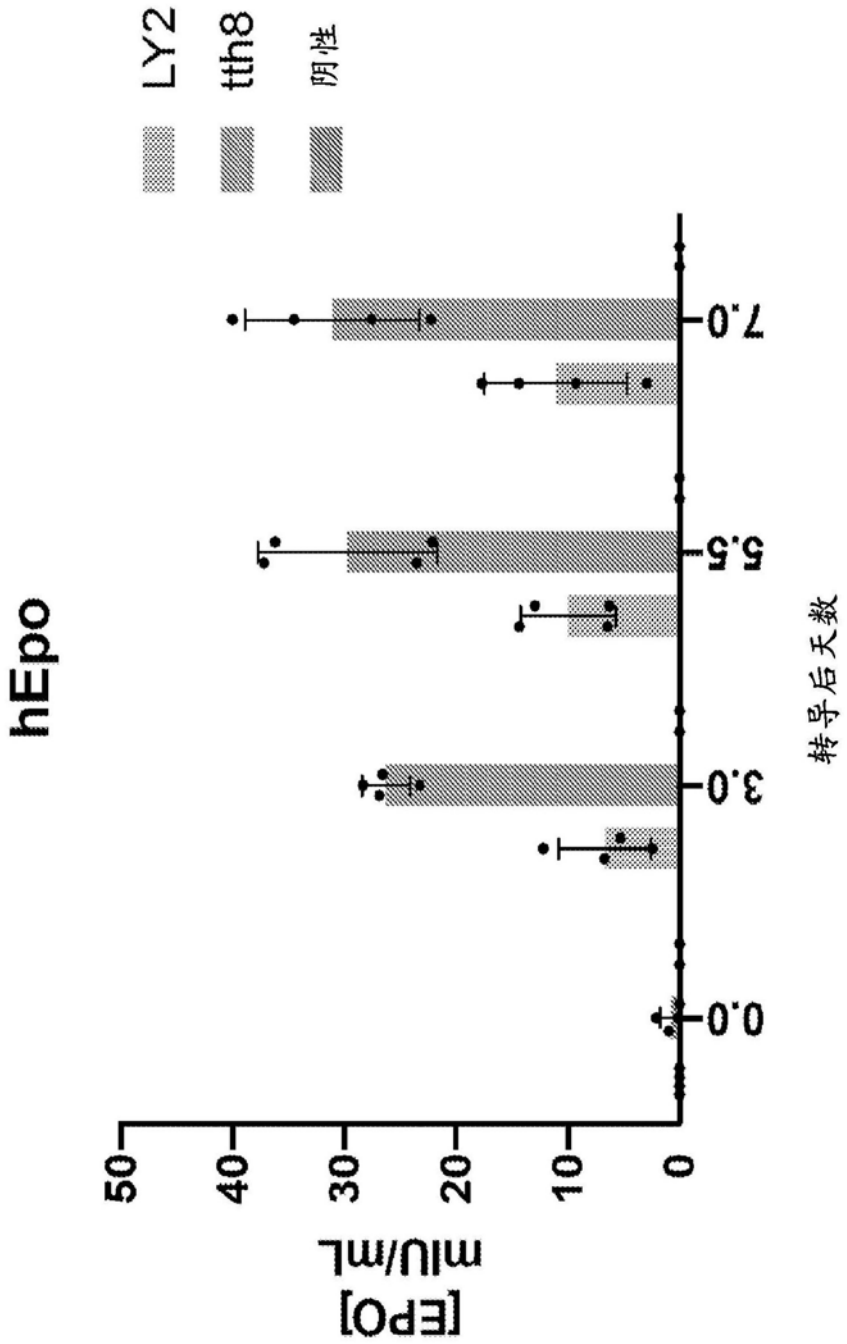


图40

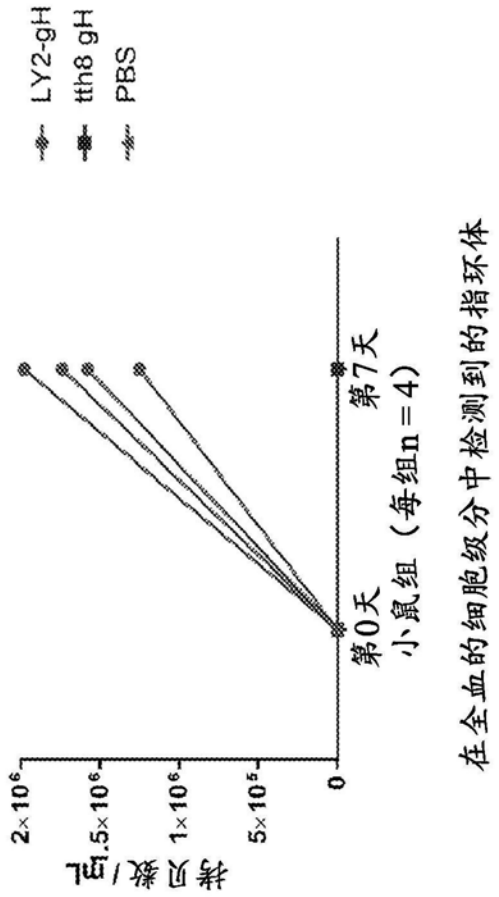

LY2-gH
tth8 gH
PBS
拷贝数/ml
2×10^6
1.5×10^6
1×10^6
5×10^5
0
第0天
第7天
小鼠组（每组n=4）
在全血的细胞级分中检测到的指环体

图41A

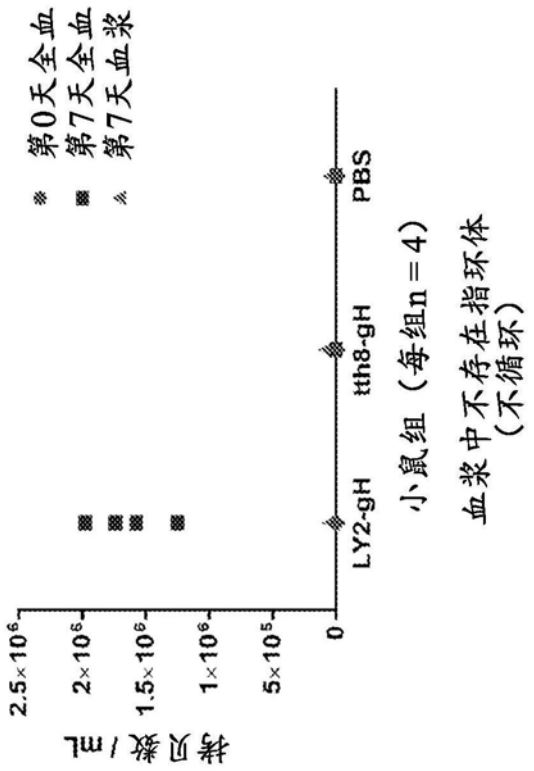

第0天全血
第7天全血
第7天血浆
拷贝数/mL
2.5×10⁶
2×10⁶
1.5×10⁶
1×10⁶
5×10⁵
0
LY2-gH
tth8-gH
PBS
小鼠组（每组n = 4）
血浆中不存在指环体
（不循环）

图41B

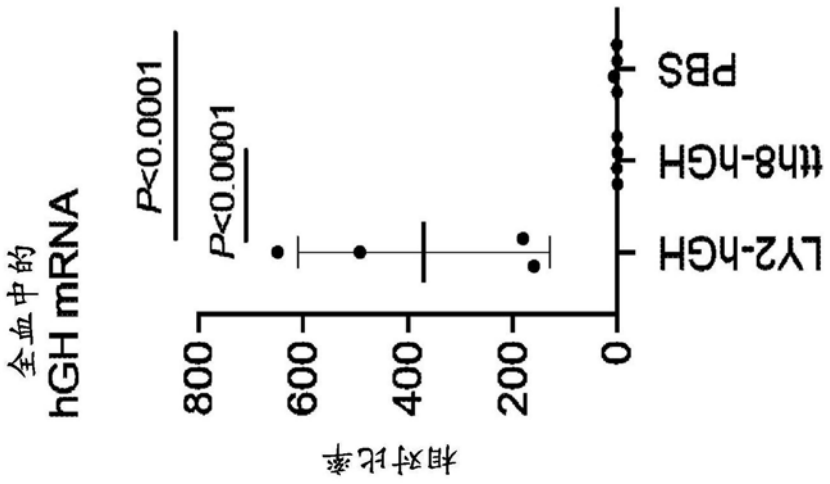

全血中的
hGH mRNA
P<0.0001
P<0.0001
相对比率
800
600
400
200
0
LY2-hGH
tth8-hGH
PBS

图42

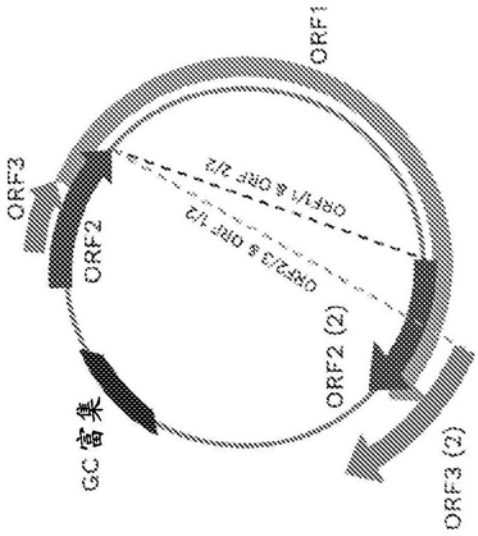

ORF1
ORF3
ORF2
ORF1/3 & ORF 2/2
ORF2/3 & ORF 1/2
ORF2 (2)
GC 富集
ORF3 (2)

图43A

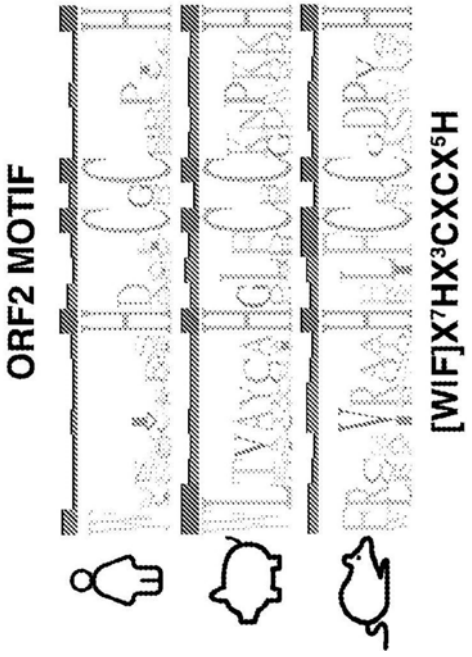

ORF2 MOTIF
[WIF]X7HX3CXCX5H

图43B

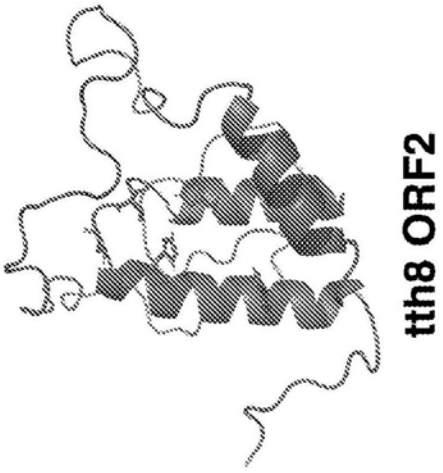


图43C

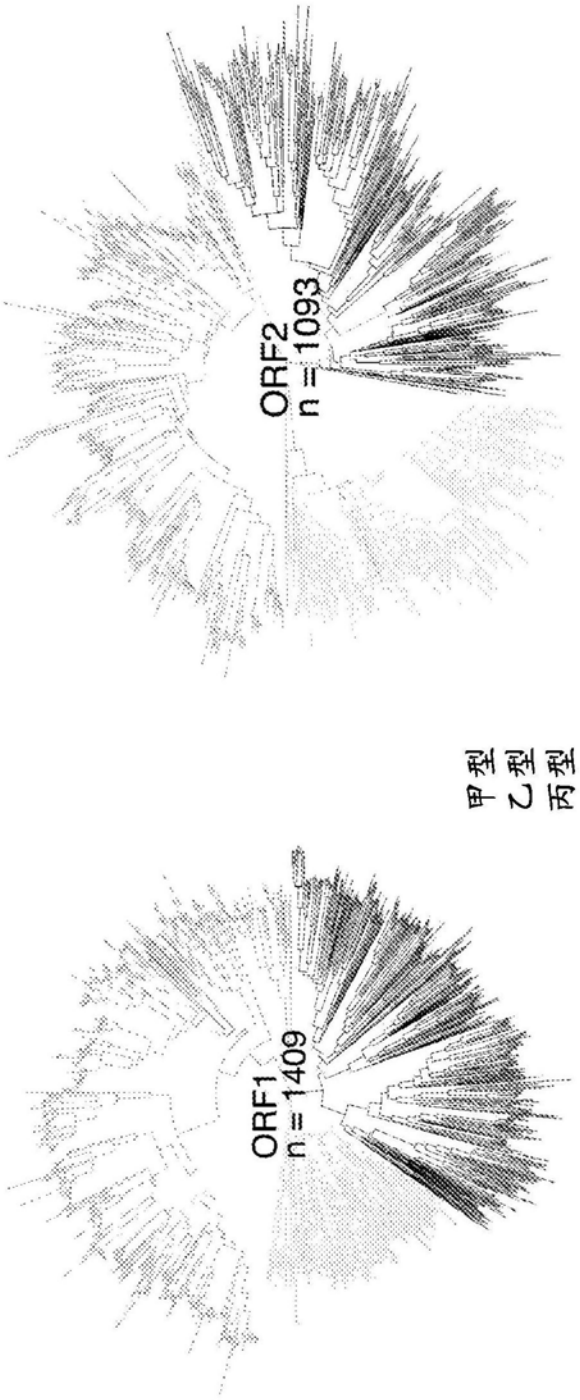


图43D

指环读段计数

供体	性别	组织	指环RNA读段	指环DNA读段
1	雌性	血液	660	109
2	雄性	血液	68	171
3	雄性	肺	33	113

图44A

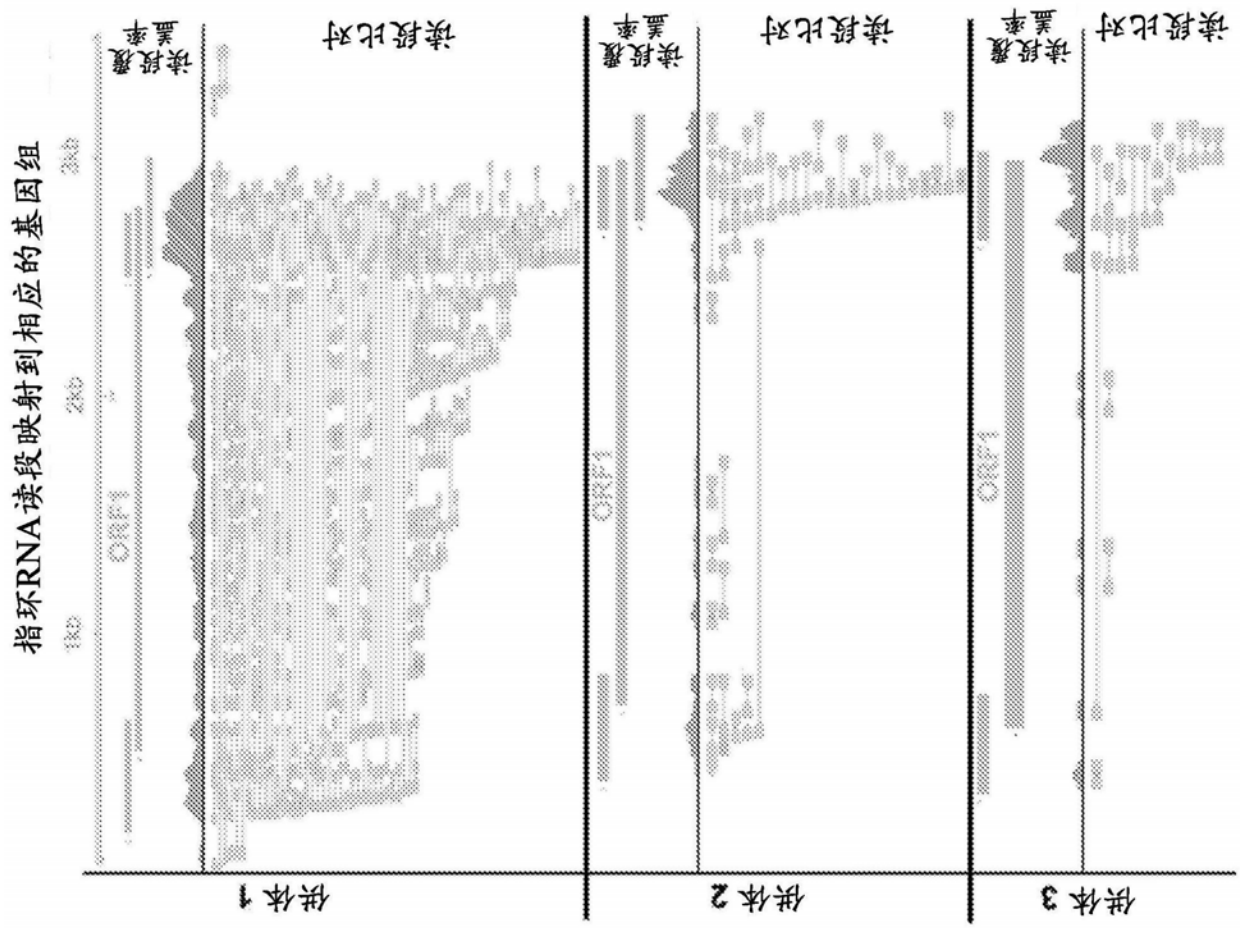


图44B

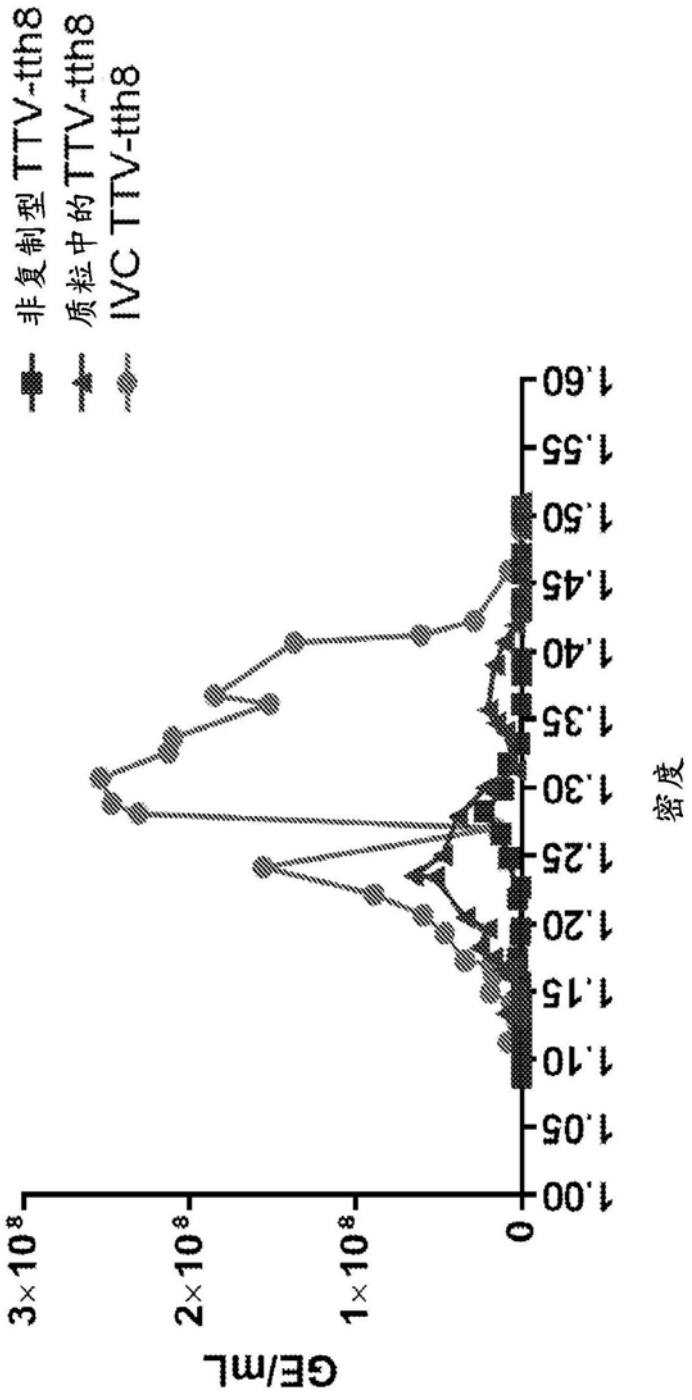

非复制型 TTV-tth8
质粒中的TTV-tth8
IVC TTV-tth8
GE/mL
3×10⁸
2×10⁸
1×10⁸
0
1.00
1.05
1.10
1.15
1.20
1.25
1.30
1.35
1.40
1.45
1.50
1.55
1.60
密度

图45

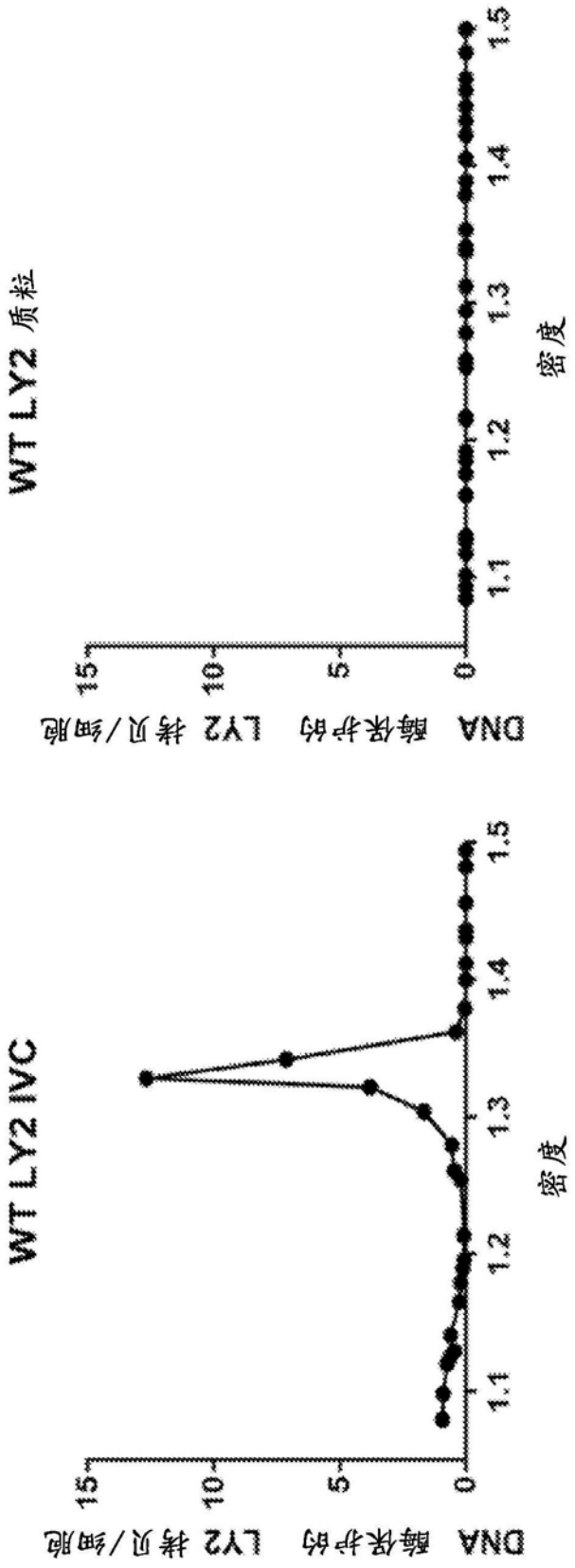

WT LY2 质粒
DNA 酶保护的 LY2 拷贝/细胞
密度
0
5
10
15
1.1
1.2
1.3
1.4
1.5
WT LY2 IVC
DNA 酶保护的 LY2 拷贝/细胞
密度

图46

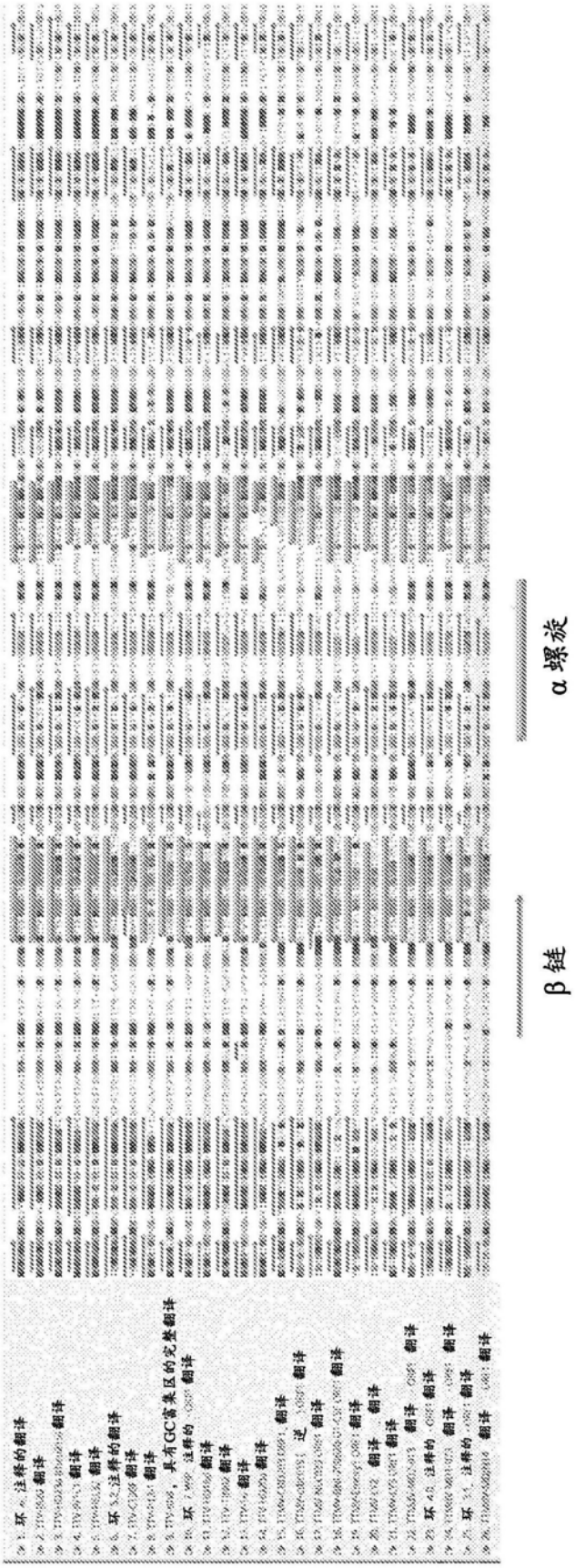
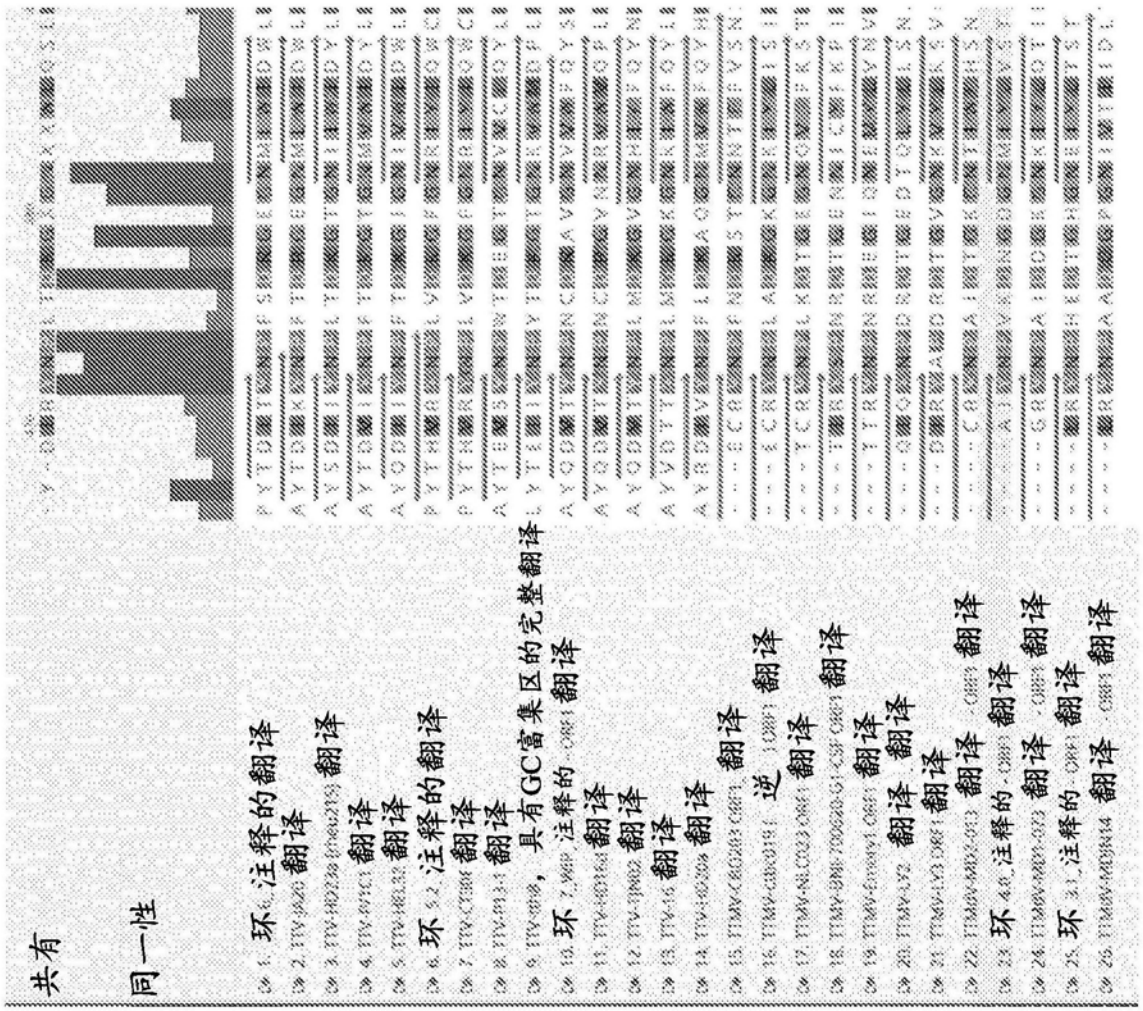


图47


共有
同一性
环 6_注释的 翻译
翻译
翻译
翻译
翻译
环 52_注释的 翻译
翻译
翻译
翻译
翻译