202235622

(19)中華民國智慧財產局

(12)發明說明書公開本 (11)公開編號：TW 202235622 A

(43)公開日： 中華民國 111 (2022) 年 09 月 16 日

(21)申請案號：110148173 (22)申請日： 中華民國 110 (2021) 年 12 月 22 日

(51)Int. Cl.： *C12N15/86 (2006.01)* *A61K48/00 (2006.01)*

(30)優先權：2020/12/23 美國 63/130,360

2021/02/08 美國 63/147,064

(71)申請人：美商旗艦先鋒創新公司 (美國) FLAGSHIP PIONEERING INNOVATIONS V, INC. (US)

美國

(72)發明人：哈捷爾 羅傑 喬瑟夫 HAJJAR, ROGER JOSEPH (US)；德拉葛瑞夫 賽門 DELAGRAVE, SIMON (US)；斯旺森 寇特 亞當 SWANSON, KURT ADAM (US)

(74)代理人：陳長文

申請實體審查：無 申請專利範圍項數：9 項 圖式數：4 共 509 頁

(54)名稱

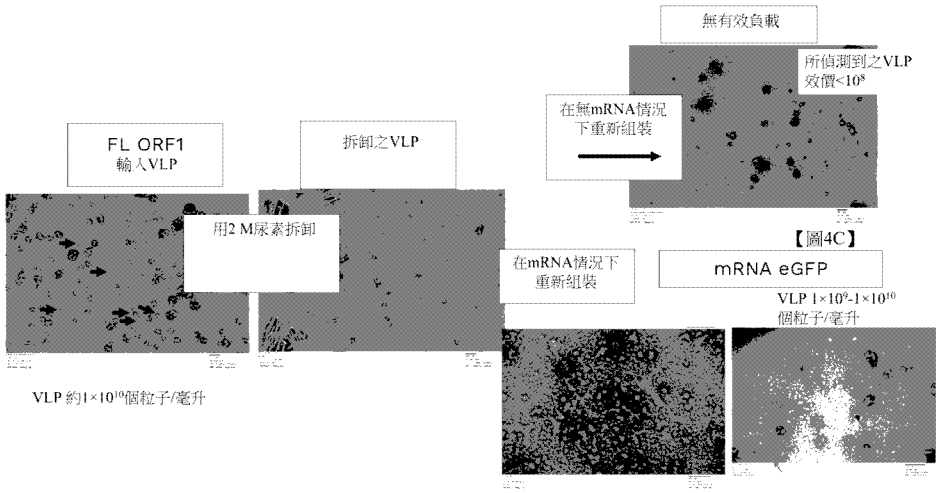
包裹 RNA 之指環病毒(ANELLOVIRUS)蛋白殼之活體外組裝

(57)摘要

本發明大體上係關於用於製造指環載體(anellovectors)之組合物及其用途。

This invention relates generally to compositions for making anellovectors and uses thereof.

指定代表圖：




【圖4A】 【圖4B】 【圖4C】 【圖4D】 【圖4E】

202233027
202259022
專利案號：110148173

# 【發明摘要】

【中文發明名稱】

包裹RNA之指環病毒(ANELLOVIRUS)蛋白殼之活體外組裝

【英文發明名稱】

IN VITRO ASSEMBLY OF ANELLOVIRUS CAPSIDS ENCLOSING RNA

【中文】

本發明大體上係關於用於製造指環載體(anellovectors)之組合物及其用途。

【英文】

This invention relates generally to compositions for making anellovectors and uses thereof.

【指定代表圖】

圖4A-4E

【代表圖之符號簡單說明】

無

# 【發明說明書】

【中文發明名稱】

包裹RNA之指環病毒(ANELLOVIRUS)蛋白殼之活體外組裝

【英文發明名稱】

IN VITRO ASSEMBLY OF ANELLOVIRUS CAPSIDS ENCLOSING RNA

【技術領域】

【先前技術】

【0001】 持續需要開發用於製成適合之載體以向患者遞送治療性效應子的組合物及方法。

【發明內容】

【0002】 本發明提供用於產生指環載體(例如合成指環載體)之組合物及方法，該指環載體可用作遞送媒劑，例如用於遞送遺傳物質、用於遞送效應子(例如有效負載)或用於遞送治療劑或治療效應子至真核細胞(例如人類細胞或人類組織)。雖然天然存在之指環病毒具有DNA基因體，但本發明提供具有包含RNA之遺傳元件的指環載體。

【0003】 指環載體(例如使用如本文所述之組合物或方法產生)一般包含囊封於蛋白質外部(例如包含指環病毒蛋白殼蛋白例如指環病毒ORF1蛋白或由指環病毒ORF1核酸編碼之多肽的蛋白質外部)之遺傳元件(例如包含或編碼效應子之遺傳元件，例如外源或內源性效應子，例如治療效應子)，該指環載體能夠將遺傳元件引入至細胞(例如哺乳動物細胞，例如人類細胞)中。遺傳元件可包含RNA。在一些實施例中，指環載體為包含蛋白質外部之感染性媒劑或粒子，該蛋白質外部包含由指環病毒ORF1核酸

(例如α細環病毒(*Alphatorquevirus*)、β細環病毒(*Betatorquevirus*)或γ細環病毒(*Gammatorquevirus*)之ORF1核酸，例如α細環病毒分枝系1、α細環病毒分枝系2、α細環病毒分枝系3、α細環病毒分枝系4、α細環病毒分枝系5、α細環病毒分枝系6或α細環病毒分枝系7之ORF1，例如如本文所述)編碼的多肽。本發明之指環載體之遺傳元件通常為環狀及/或單股RNA分子(例如，環狀及單股的)，其具有與包裹其之蛋白質外部結合的蛋白質結合序列或與其連接之多肽，此可促進遺傳元件包裹在蛋白質外部內及/或使遺傳元件相對於其他核酸在蛋白質外部內富集。在一些實施例中，使用桿狀病毒、核酸構築體(例如桿狀病毒質體及/或供體載體)、昆蟲細胞及/或例如如本文所描述之動物細胞株來產生指環載體之遺傳元件。

【0004】 在一些情況下，遺傳元件包含或編碼例如可在目標細胞中表現之效應子(例如包含核酸效應子，諸如非編碼RNA，或多肽效應子，例如蛋白質)。在一些實施例中，效應子為治療劑或治療性效應子，例如如本文所描述。在一些實施例中，效應子為內源性效應子或外源性效應子，例如針對野生型指環病毒或目標細胞為外源性的。在一些實施例中，效應子針對野生型指環病毒或目標細胞為外源性的。在一些實施例中，指環載體可藉由接觸細胞且將編碼效應子之遺傳元件引入細胞中而將效應子遞送至細胞中，使得效應子由細胞製成或表現。在某些情況下，效應子係內源性效應子(例如針對目標細胞為內源性的，但例如藉由指環載體以增加的量提供)。在其他情況下，效應子為外源性效應子。在一些情況下，效應子可調節細胞功能或調節細胞中目標分子之活性或水準。例如，效應子可降低細胞中目標蛋白之水準(例如，如PCT/US19/65995之實例3及4中所描述)。在另一實例中，指環載體可活體內遞送及表現效應子，例如外

源性蛋白質(例如，如PCT/US19/65995之實例19及28中所描述)。指環載體可用於例如將遺傳物質遞送至目標細胞、組織或個體；將效應子遞送至目標細胞、組織或個體；或用於治療疾病及病症，例如藉由遞送可作為治療劑操作之效應子至所需細胞、組織或個體。

【0005】 在一些實施例中，本文所述之組合物及方法可用於在例如宿主細胞中產生合成指環載體之遺傳元件。合成指環載體與野生型病毒(例如野生型指環病毒，例如本文所描述)相比具有至少一個結構差異，例如相對於野生型病毒的缺失、插入、取代、修飾(例如酶修飾)。一般而言，合成指環載體包括包裹於蛋白質外部內之外源性遺傳元件，其可用於將遺傳元件或其中所編碼之效應子(例如外源性效應子或內源性效應子)(例如多肽或核酸效應子)遞送至真核(例如人類)細胞中。在實施例中，指環載體不引起可偵測及/或非所需免疫或炎症反應，例如不會使炎症分子標記(例如TNF-α、IL-6、IL-12、IFN)以及B細胞反應(例如反應性或中和抗體)增加超過1%、5%、10%、15%，例如指環載體可實質上對目標細胞、組織或個體無免疫原性。

【0006】 在一些實施例中，本文所述之組合物及方法可用於產生指環載體之遺傳元件，該指環載體包含：(i)包含ORF1分子之蛋白質外部；及(ii)包含RNA之遺傳元件；其中該遺傳元件包裹於該蛋白質外部內。在一些實施例中，遺傳元件由至少10%、20%、25%、30%、40%、50%、60%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100% RNA組成。在一些實施例中，遺傳元件不包含DNA。在一些實施例中，遺傳元件不包含ssDNA。替代地，在一些實施例中，遺傳元件包含DNA區。在一些實施例中，本文所描述之DNA或RNA分子包含一或多個

經修飾之核苷酸(例如鹼基修飾、糖修飾或主鏈修飾)。在一些實施例中，遺傳元件為單股。在一些實施例中，遺傳元件包含雙股區。在一些實施例中，遺傳元件為線性多肽。替代地，在一些實施例中，遺傳元件為環狀聚核苷酸。在一些實施例中，核酸序列經密碼子最佳化，例如用於昆蟲細胞中之表現。在一些實施例中，核酸序列中至少50%、60%、70%、80%、90%、95%、96%、97%、98%、99%或100%之密碼子經密碼子最佳化，例如用於昆蟲細胞中之表現。在一些實施例中，核酸序列經密碼子最佳化，例如用於在哺乳動物(例如人類)細胞中表現。在一些實施例中，核酸序列中至少50%、60%、70%、80%、90%、95%、96%、97%、98%、99%或100%之密碼子經密碼子最佳化，例如用於在哺乳動物(例如人類)細胞中表現。在一些實施例中，遺傳元件之長度為約10-20、20-30、30-40、50-60、60-70、70-80、80-90、90-100、100-125、125-150、150-200、200-300、300-400、400-500、500-600、600-700、700-800、800-900、900-1000、1000-1500、1500-2000、2000-2500、2500-3000、3000-3500、3500-4000或4000-4500個核苷酸。在一些實施例中，遺傳元件之長度為至少約10、20、30、40、50、60、70、80、90、100、125、150、175、200、250、300、400、500、600、700、800、900、1000、1500、2000、2500、3000、3500、4000或4500個核苷酸。

【0007】 在一些實施例中，本文所述之組合物及方法可用於產生感染性(例如針對人類細胞)指環載體、媒劑或粒子之遺傳元件，該指環載體、媒劑或粒子包含蛋白殼(例如包含指環病毒ORF，例如ORF1，多肽之蛋白殼)，該蛋白殼囊封包含結合至蛋白殼之蛋白質結合序列及編碼治療性效應子之異源(針對指環病毒)序列的遺傳元件。在實施例中，指環載體

能夠將遺傳元件遞送至哺乳動物，例如人類細胞中。

【0008】 在一態樣中，本發明之特徵為一種藉由活體外組裝製成指環載體之方法。在一些實施例中，製成指環載體之方法包含：(a)提供包含以下之混合物：(i)包含RNA之遺傳元件及(ii) ORF1分子；及(b)在適用於將遺傳元件包裹於包含ORF1分子之蛋白質外部內的條件下培育該混合物，由此製成指環載體；視情況其中，該混合物不包含於細胞中。在一些實施例中，一種方法在提供(a)之前進一步包含例如在宿主細胞(例如昆蟲細胞或哺乳動物細胞)中表現ORF1分子。在一些實施例中，表現包含在適用於產生ORF1分子之條件下培育包含編碼ORF1分子之核酸分子(例如桿狀病毒表現載體)的宿主細胞(例如昆蟲細胞或哺乳動物細胞)。在一些實施例中，一種方法在提供(a)之前進一步包含純化由宿主細胞表現之ORF1分子。

【0009】 在一些實施例中，如本文所述之指環載體可用作有效遞送媒劑，用於將藥劑，諸如本文所描述之效應子引入至目標細胞，例如待治療性或預防性治療之個體中的目標細胞。

【0010】 在一態樣中，本發明之特徵在於一種醫藥組合物，其包含如本文所述之指環載體(例如合成指環載體)。在一些實施例中，醫藥組合物進一步包含醫藥學上可接受之載劑或賦形劑。在實施例中，醫藥組合物包含單位劑量，該單位劑量包含每公斤目標個體之指環載體約$10^5$-$10^{14}$(例如約$10^6$-$10^{13}$、$10^7$-$10^{12}$、$10^8$-$10^{11}$或$10^9$-$10^{10}$)基因體當量。在一些實施例中，包含製劑之醫藥組合物在可接受之時段及溫度內將為穩定的，及/或與所需投與途徑及/或此投與途徑將需要之任何裝置相容，例如針或注射器。在一些實施例中，醫藥組合物經調配而以單次劑量或多次劑量投

與。在一些實施例中，醫藥組合物係在投與地點，例如由醫療保健專業人員調配。在一些實施例中，醫藥組合物包含所需濃度之指環載體基因體或基因體當量(例如如每體積之基因體數目所定義)。

【0011】 在一態樣中，本發明提供一種治療個體之疾病或病症之方法，該方法包含向個體投與指環載體，例如合成指環載體，例如如本文所述。

【0012】 在一態樣中，本發明提供一種向細胞、組織或個體遞送效應子或有效負載(例如內源性或外源性效應子)之方法，該方法包含向個體投與指環載體，例如合成指環載體，例如如本文所述，其中該指環載體包含編碼效應子之核酸序列。在一些實施例中，有效負載為核酸。在一些實施例中，有效負載為多肽。

【0013】 在一態樣中，本發明提供一種將指環載體遞送至細胞之方法，其包含使指環載體，例如合成指環載體，例如如本文所述，與細胞，例如真核細胞，例如哺乳動物細胞，例如活體內或離體接觸。

【0014】 前述指環載體、組合物或方法中之任一者之額外特徵包括以下所列舉實施例中之一或多者。

【0015】 熟習此項技術者將認識到或能夠僅使用常規實驗即可確定本文所描述之本發明特定實施例的許多等效物。此類等效物意欲由以下所列舉之實施例涵蓋。

【0016】

**所列舉實施例**

1. 一種指環載體，其包含：

a)包含ORF1分子之蛋白質外部；

b)包含RNA之遺傳元件，

其中該遺傳元件包裹於該蛋白質外部內。

2. 如實施例1之指環載體，其中該遺傳元件由至少10%、20%、25%、30%、40%、50%、60%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100% RNA組成。

3. 如實施例1或2之指環載體，其中該RNA包含一或多個化學修飾。

4. 如前述實施例中任一項之指環載體，其中該遺傳元件由RNA組成或基本上由RNA組成。

5. 如前述實施例中任一項之指環載體，其中該指環載體不包含DNA。

6. 如前述實施例中任一項之指環載體，其中該指環載體不包含ssDNA。

7. 如前述實施例中任一項之指環載體，其中該遺傳元件包含DNA區。

8. 如前述實施例中任一項之指環載體，其中該DNA區之所有核苷酸經化學修飾。

9. 如實施例7或8之指環載體，其中該DNA區共價連接至該遺傳元件之RNA。

10. 如前述實施例中任一項之指環載體，其中該DNA區之至少一部分與該遺傳元件之該RNA的至少一部分雜交。

11. 如前述實施例中任一項之指環載體，其中該DNA區之長度為10-20、20-30、30-40、40-50、50-60、60-70、70-80、80-90、90-

100、100-150、150-200、200-300、300-400、400-500、500-600、600-700、700-800、800-900或900-1000個核苷酸。

12. 如前述實施例中任一項之指環載體，其中該遺傳元件為單股。

13. 如前述實施例中任一項之指環載體，其中該遺傳元件包含雙股區(例如RNA與RNA配對或DNA與RNA配對之區域)。

14. 如前述實施例中任一項之指環載體，其中該遺傳元件為線狀。

15. 如前述實施例中任一項之指環載體，其中該遺傳元件為環狀。

16. 如前述實施例中任一項之指環載體，其中該遺傳元件包含第一區域及可與該第一區域雜交之第二區域。

17. 如前述實施例中任一項之指環載體，其中該遺傳元件不包含5'端或3'端。

18. 如實施例15至17中任一項之指環載體，其中該遺傳元件不包含游離磷酸根及游離糖中之一者或兩者。

19. 如實施例15至18中任一項之指環載體，其中該遺傳元件中之每個磷酸根藉由該磷酸根所包含之第一氧原子共價連接至第一糖及藉由該磷酸根所包含之第二氧原子共價連接至第二糖。

20. 如實施例15至19中任一項之指環載體，其中該遺傳元件中之每個糖藉由該糖所包含之第一碳原子共價連接至第一磷酸根及藉由該糖所包含之第二碳原子共價連接至第二磷酸根。

21. 如實施例15至20中任一項之指環載體，其中該遺傳元件係藉由使線性RNA環化而產生。

22. 如前述實施例中任一項之指環載體，其中該遺傳元件之長度為約10-20、20-30、30-40、50-60、60-70、70-80、80-90、90-100、100-

125、125-150、150-200、200-300、300-400、400-500、500-600、600-700、700-800、800-900、900-1000、1000-1500、1500-2000、2000-2500、2500-3000、3000-3500、3500-4000或4000-4500個核苷酸。

23. 如前述實施例中任一項之指環載體，其中該遺傳元件之長度為至少約10、20、30、40、50、60、70、80、90、100、125、150、175、200、250、300、400、500、600、700、800、900、1000、1500、2000、2500、3000、3500、4000或4500個核苷酸。

24. 如前述實施例中任一項之指環載體，其中該遺傳元件結合ORF1分子。

25. 如前述實施例中任一項之指環載體，其中該遺傳元件結合ORF1分子所包含之果醬捲域。

26. 如前述實施例中任一項之指環載體，其中該遺傳元件結合ORF1分子所包含之富含精胺酸之結構域。

27. 如前述實施例中任一項之指環載體，其中該指環載體包含複數個遺傳元件，例如至少2、3、4、5、10、20、30、40、50或60個遺傳元件。

28. 如實施例27之指環載體，其中該複數個遺傳元件各自包含相同序列。

29. 如實施例27之指環載體，其中該複數個遺傳元件包含不同序列。

30. 如前述實施例中任一項之指環載體，其中該遺傳元件編碼外源性效應子。

31. 如實施例30之指環載體，其中該編碼外源性效應子之序列的長

度為至少約10、20、30、40、50、60、70、80、90、100、125、150、175、200、250、300、400、500、600、700、800、900、1000、1500、2000、2500或3000個核苷酸。

32. 如實施例30至31中任一項之指環載體，其中該外源性效應子包含治療性效應子(例如多肽或核酸分子)。

33. 如實施例30至32中任一項之指環載體，其中該外源性效應子包含有人類蛋白質或包含與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%序列一致性之胺基酸序列的多肽。

34. 如實施例30至33中任一項之指環載體，其中該外源性效應子包含核酸分子。

35. 如實施例30至34中任一項之指環載體，其中該外源性效應子包含非編碼核酸分子，例如功能性RNA，例如mRNA、miRNA或siRNA。

36. 如實施例30至35中任一項之指環載體，其中該遺傳元件為編碼外源性效應子(例如肽或多肽，例如治療肽或多肽)之mRNA分子。

37. 如實施例36之指環載體，其中非編碼核酸分子為人類非編碼核酸分子或包含與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%序列一致性之序列的核酸分子。

38. 如實施例30至37中任一項之指環載體，其中該外源性效應子包含細胞溶質多肽或細胞溶質肽(例如DPP-4抑制劑、GLP-1信號傳導之活化劑或嗜中性球彈性蛋白酶之抑制劑或其功能片段)。

39. 如實施例38之指環載體，其中該外源性效應子包含調控性胞內多肽。

40. 如實施例30至39中任一項之指環載體，其中外源性效應子包含

分泌多肽或肽(例如細胞介素、抗體分子、激素、生長因子或凝血相關因子或其功能片段)。

41. 如實施例30至40中任一項之指環載體，其中該外源性效應子包含蛋白質替代治療劑。

42. 如實施例30至41中任一項之指環載體，其中該外源性效應子包含酶。

43. 如實施例30至42中任一項之指環載體，其中外源性效應子包含紅血球生成素(EPO)或人類生長激素(hGH)或其功能片段。

44. 如實施例30至43中任一項之指環載體，其中該外源性效應子包含基因編輯系統之組分(例如CRISPR系統之組分，例如Cas9、Cpf1或其功能片段)。

45. 如前述實施例中任一項之指環載體，其中例如如本文所描述，RNA包含經化學修飾之RNA。

46. 如前述實施例中任一項之指環載體，其中RNA包含帽。

47. 如前述實施例中任一項之指環載體，其中RNA包含聚A尾，例如長度為至少約5、10、15、20、25、30、40、50、60、70、80、90或100個腺苷。

48. 如前述實施例中任一項之指環載體，其中RNA缺乏聚A尾，例如包含不超過5、10、15、20、25、30、40、50、60、70、80、90或100個連續腺苷。

49. 如前述實施例中任一項之指環載體，其中蛋白質外部包含ORF1分子之約60個(例如約40、50、60、70或80個)複本。

50. 如前述實施例中任一項之指環載體，其中ORF1分子之果醬捲

域面向蛋白質外部之內部。

51. 如前述實施例中任一項之指環載體，其中ORF1分子包含如表N-S及表37A-37C中之任一者中所列的胺基酸序列或與其具有至少30%、40%、50%、60%、70%、75%、80%、85%、90%、95%、96%、97%、98%或99%序列一致性的胺基酸序列。

52. 如前述實施例中任一項之指環載體，其中ORF1分子包含富含精胺酸之區域，例如包含如表N-S及表37A-37C中之任一者中所列的富含精胺酸之區域的胺基酸序列或與其具有至少30%、40%、50%、60%、70%、75%、80%、85%、90%、95%、96%、97%、98%或99%序列一致性的胺基酸序列。

53. 如實施例52之指環載體，其中富含精胺酸之區域包含至少約70% (例如至少約70%、75%、80%、85%、90%、95%或100%)鹼性殘基(例如精胺酸或離胺酸)。

54. 如前述實施例中任一項之指環載體，其中ORF1分子包含果醬捲域，其例如包含如表N-S及表37A-37C中之任一者中所列的果醬捲域之胺基酸序列或與其具有至少30%、40%、50%、60%、70%、75%、80%、85%、90%、95%、96%、97%、98%或99%序列一致性的胺基酸序列。

55. 如實施例54之指環載體，其中果醬捲域包含以下特徵中之一或多者(例如1、2、3或4者)：

(i)果醬捲域之胺基酸中之至少30% (例如至少30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、90%或更多)為一或多個β片之一部分；

(ii)果醬捲域之二級結構包含至少四條(例如至少4、5、6、7、8、9、10、11或12條) β股；及/或

(iii)果醬捲域之三級結構包含至少兩個(例如至少2、3或4個) β片；及/或

(iv)果醬捲域包含至少2:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1或10:1之β片與α螺旋之比。

56. 如實施例54之指環載體，其中果醬捲域包含兩個β片，例如相對於彼此以反平行定向佈置。

57. 如實施例54之指環載體，其中果醬捲域包含八條β股。

58. 如實施例52至57中任一項之指環載體，其中果醬捲域包含與D β股之序列具有至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的區域，例如如圖3中所示。

59. 如實施例52之指環載體，其中D β股包含1、2或3個或更多個鹼性殘基(例如精胺酸或離胺酸)。

60. 如實施例52至59中任一項之指環載體，其中果醬捲域包含與G β股之序列具有至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的區域，例如如圖3中所示。

61. 如實施例52之指環載體，其中G β股包含至少約1、2或3個或更多個鹼性殘基(例如精胺酸或離胺酸)。

62. 如前述實施例中任一項之指環載體，其中ORF1分子包含N22域，例如包含如表N-S及表37A-37C中之任一者中所列的N22域之胺基酸序列或與其具有至少30%、40%、50%、60%、70%、75%、80%、85%、90%、95%、96%、97%、98%或99%序列一致性的胺基酸序列。

63. 如前述實施例中任一項之指環載體，其中N22域包含胺基酸序列YNPX$^2$DXGX$^2$N，其中X$^n$各自獨立地為任何n個胺基酸之連續序列。

64. 如前述實施例中任一項之指環載體，其中N22域包含側接胺基酸序列YNPX$^2$DXGX$^2$N之第一β股及第二β股，例如其中第一β股包含胺基酸序列YNPX$^2$DXGX$^2$N之酪胺酸(Y)殘基，及/或其中第二β股包含胺基酸序列YNPX$^2$DXGX$^2$N之第二天冬醯胺(N)殘基(自N至C)。

65. 如前述實施例中任一項之指環載體，其中ORF1分子包含C端域，例如包含如表N-S及表37A-37C中之任一者中所列的C端域之胺基酸序列或與其具有至少30%、40%、50%、60%、70%、75%、80%、85%、90%、95%、96%、97%、98%、或99%序列一致性的胺基酸序列。

66. 如前述實施例中任一項之指環載體，其中遺傳元件缺乏編碼指環病毒ORF1蛋白之序列(例如如本文所述)。

67. 如前述實施例中任一項之指環載體，其中遺傳元件缺乏編碼指環病毒ORF2蛋白之序列(例如如本文所述)。

68. 如前述實施例中任一項之指環載體，其中遺傳元件缺乏編碼指環病毒ORF3蛋白之序列(例如如本文所述)。

69. 如前述實施例中任一項之指環載體，其中指環載體經組態以將遺傳元件遞送至細胞(例如真核細胞，例如哺乳動物細胞，例如人類細胞)。

70. 如實施例69之指環載體，其中至少1000個指環載體之群體能夠將遺傳元件之至少約100個複本(例如至少1、2、3、4、5、10、20、30、40、50、100、200、300、400、500、600、700、800、900、1000、

5000、10,000、50,000、100,000、500,000或1,000,000個複本)遞送至一或多個細胞中。

71. 如實施例69或70之指環載體，其中指環載體之群體(例如，每個細胞之遺傳元件至少1、2、3、4、5、10、20、30、40、50、100、200、300、400、500、600、700、800、900或1000基因體當量)能夠將遺傳元件遞送至至少10%、20%、30%、40%、50%、60%、70%、80%、90%、95%、99%或更多之細胞群體。

72. 如實施例69至71中任一項之指環載體，其中指環載體之群體(例如，每個細胞之遺傳元件為至少1、2、3、4、5、10、20、30、40、50、100、200、300、400、500、600、700、800、900或1000基因體當量)能夠將每個細胞至少1、2、3、4、5、6、7、8、9、10、20、50、100、200、500、1000、2000、5000、8,000、$1\times10^4$、$1\times10^5$、$1\times10^6$、$1\times10^7$個或更多個遺傳元件複本遞送至細胞群體。

73. 如實施例69至72中任一項之指環載體，其中指環載體之群體(例如，每個細胞之遺傳元件為至少1、2、3、4、5、10、20、30、40、50、100、200、300、400、500、600、700、800、900或1000基因體當量)能夠將每個細胞至少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-$10^4$、$1\times10^4$-$1\times10^5$、$1\times10^4$-$1\times10^6$、$1\times10^4$-$1\times10^7$、$1\times10^5$-$1\times10^6$、$1\times10^5$-$1\times10^7$或$1\times10^6$-$1\times10^7$個或更多個遺傳元件複本遞送至細胞群體。

74. 如前述實施例中任一項之指環載體，其中指環載體將效應子選擇性遞送至或以較高水準存在於(例如優先積聚於)所需細胞類型、組織或器官(例如骨髓、血液、心臟、GI、皮膚、視網膜中之感光體、上皮內層

或胰臟)中。

75. 如前述實施例中任一項之指環載體，其中遺傳元件經保護而不被RNA酶消化(例如由蛋白質外部消化)或對其具有抗性。

76. 如前述實施例中任一項之指環載體，其中包裹於蛋白質外部內之遺傳元件對核酸內切酶消化，例如對RNA酶消化具有抗性。

77. 如前述實施例中任一項之指環載體，其中遺傳元件包含啟動子元件。

78. 如前述實施例中任一項之指環載體，其中遺傳元件包含蛋白質結合序列。

79. 如實施例78之指環載體，其中蛋白質結合序列能夠結合至ORF1分子。

80. 一種組合物，其包含複數個前述實施例中任一項之指環載體。

81. 如實施例80之組合物，其中組合物中至少50%、60%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%的包含ORF1分子之蛋白質外部包含遺傳元件之至少一個複本。

82. 如實施例80之組合物，其中組合物中至少50%、60%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%的包含ORF1分子之蛋白質外部包含指環載體遺傳元件之至少一個複本。

83. 如實施例80至82中任一項之組合物，其中組合物包含至少$10^2$、$10^3$、$10^4$、$10^4$、$10^5$、$10^6$或$10^7$個相同指環載體。

84. 如實施例80至83中任一項之組合物，其中該複數個包含至少$10^3$、$10^4$、$10^5$、$10^6$、$10^7$、$10^8$、$10^9$、$10^{10}$、$10^{11}$、$10^{12}$、$10^{13}$、$10^{14}$或$10^{15}$個指環載體(例如指環載體之複本)；或其中組合物包含每毫升至少

$10^5$、$10^6$、$10^7$、$10^8$、$10^9$、$10^{10}$、$10^{11}$、$10^{12}$、$10^{13}$、$10^{14}$或$10^{15}$個指環載體基因體。

85. 如實施例80至84中任一項之組合物，其具有以下特徵中之一或多者(例如1、2、3、4、5或6者)：

a)該組合物滿足醫藥或良好作業規範(GMP)標準；

b)根據良好作業規範(GMP)製得組合物；

c)該組合物具有低於預定參考值之病原體含量，例如實質上不含病原體；

d)該組合物具有低於預定參考值之污染物含量，例如實質上不含污染物(例如變性劑，例如尿素)；

e)該組合物具有預定含量之非感染性粒子或預定比率之粒子:感染性單元(例如<300:1、<200:1、<100:1或<50:1)，或

f)該醫藥組合物具有較低免疫原性或實質上為非免疫原性的，例如如本文所描述；

視情況其中該組合物包含濃度小於0.1M、0.2M、0.3M、0.4M、0.5M、0.6M、0.7M、0.8M、0.9M、1M、1.1M、1.2M、1.3M、1.5M、1.5M、1.6M、1.7M、1.8M、1.9M或2M之尿素。

86. 如實施例80至85中任一項之組合物，其中該醫藥組合物之污染物含量低於預定參考值，例如實質上不含污染物。

87. 如實施例80至86中任一項之組合物，其中該組合物包含濃度為小 於0.1M、0.2M、0.3M、0.4M、0.5M、0.6M、0.7M、0.8M、0.9M、1M、1.1M、1.2M、1.3M、1.5M、1.5M、1.6M、1.7M、1.8M、1.9M或2M之尿素。

88. 如實施例86或87之組合物，其中該污染物包含以下中之一或多者：徽漿菌、內毒素、宿主細胞核酸(例如宿主細胞DNA及/或宿主細胞RNA)、生物衍生之過程雜質(例如血清白蛋白或胰蛋白酶)、複製勝任型試劑(RCA)，例如複製勝任型病毒或非所需指環載體(例如除所需指環載體以外的指環載體，例如如本文所述之合成指環載體)、游離病毒蛋白殼蛋白、偶然性物質及/或聚集體。

89. 如實施例80至88中任一項之組合物，其中該組合物包含小於10重量% (例如小於約10重量%、5重量%、4重量%、3重量%、2重量%、1重量%、0.5重量%或0.1重量%)之污染物。

90. 如實施例80至89中任一項之組合物，其中至少90%之蛋白質外部包含相同遺傳元件(例如，指環載體之遺傳元件)。

91. 如實施例80至90中任一項之組合物，其中至少90%之蛋白質外部包含相同ORF1分子。

92. 一種醫藥組合物，其包含如前述實施例中任一項之指環載體或組合物及醫藥學上可接受之載劑或賦形劑。

93. 一種製成指環載體之方法，該方法包含：

(a)提供包含以下之混合物：

(i)包含RNA之遺傳元件，及

(ii) ORF1分子；及

(b)在適用於將該遺傳元件包裹於包含該ORF1分子之蛋白質外部內的條件下培育該混合物，由此製成指環載體；

視情況其中，該混合物不包含於細胞中。

94. 如實施例93之方法，其在提供(a)之前，例如在宿主細胞(例如

昆蟲細胞或哺乳動物細胞)中，進一步包含表現ORF1分子。

95. 如實施例94之方法，其中該表現包含在適用於產生ORF1分子之條件下培育包含編碼ORF1分子之核酸分子(例如桿狀病毒表現載體)的宿主細胞(例如昆蟲細胞或哺乳動物細胞)。

96. 如實施例94或95之方法，其在提供(a)之前進一步包含純化由宿主細胞表現之ORF1分子。

97. 一種純化指環載體之方法，該方法包含：

(a)提供一種指環載體(例如如本文所描述)，其包含：

(i)遺傳元件，例如包含RNA之遺傳元件，及

(ii)包含ORF1分子之蛋白質外部，該蛋白質外部包裹該遺傳元件；及

(b)純化該指環載體。

98. 如實施例96或97之方法，其中純化(例如純化ORF1分子或純化指環載體)包含親和純化，例如肝素親和純化。

99. 如實施例96至98中任一項之方法，其中純化(例如純化ORF1分子或純化指環載體)包含尺寸排阻層析(例如使用Tris緩衝液移動相)。

100. 如實施例96至99中任一項之方法，其中純化(例如純化ORF1分子或純化指環載體)包含親和純化(例如肝素親和純化)，隨後為尺寸排阻層析。

101. 如實施例96至100中任一項之方法，其中純化(例如純化ORF1分子或純化指環載體)包含陰離子交換層析(例如Mustang Q膜層析)。

102. 如實施例96至101中任一項之方法，其中純化(例如純化ORF1分子或純化指環載體)包含混合模式層析(例如使用混合模式樹脂，例如

Cato700樹脂)。

103. 如實施例96至102中任一項之方法，其中純化(例如純化ORF1分子或純化指環載體)產生包含一或多個病毒樣粒子(VLP)的組合物，該一或多個病毒樣粒子包含ORF1分子之至少約20、30、40、50或60個複本或20-30、30-40、40-50或50-60個複本。

104. 如實施例103之方法，其中至少75%、80%、90%、95%、96%、97%、98%、99%或100%之病毒樣粒子包含蛋白質外部，該等蛋白質外部為60聚體或為直徑為至少30、31、32、33、34或35 nm之粒子。

105. 如實施例103之方法，其中該組合物包含至少$10^5$、$10^6$、$10^7$、$10^8$、$10^9$、$10^{10}$個粒子/毫升，或包含$10^5$ - $10^6$、$10^6$ - $10^7$、$10^7$ - $10^9$、$10^8$ - $10^9$、$10^9$ - $10^{10}$或$10^{10}$ - $10^{11}$個粒子/毫升(例如如藉由電子顯微法所量測)。

106. 如實施例93至105中任一項之方法，其在提供(a)之前進一步包含在適用於拆卸包含ORF1分子之蛋白質外部(例如，病毒樣粒子(VLP))之條件下培育ORF1分子。

107. 如實施例106之方法，其中適用於拆卸包含ORF1分子之蛋白質外部之條件包含在變性劑存在下培育。

108. 如實施例106至107中任一項之方法，其中變性劑包含離液劑(例如尿素)或清潔劑(例如SDS (例如0.1% SDS)、吐溫(Tween)或曲拉通(Triton))。

109. 如實施例106至108中任一項之方法，其中適用於拆卸包含ORF1分子之蛋白質外部之條件包含預定傳導性、高鹽溶液(例如，包含NaCl之溶液，例如濃度為至少約1M，例如至少約0.5、0.6、0.7、0.8、

0.9、1.0、1.1、1.2、1.3、1.4、1.5、2、3、4或5M)、熱量(例如，溫度高於約85、86、87、88、89、90、91、92、93、94或95℃)或pH (例如酸性pH或鹼性pH)。

110. 如實施例106至109中任一項之方法，其中適用於拆卸包含ORF1分子之蛋白質外部之條件包含在包含尿素之溶液中培育，該尿素之濃度為至少0.1M、0.2M、0.3M、0.4M、0.5M、0.6M、0.7M、0.8M、0.9M、1M、1.1M、1.2M、1.3M、1.5M、1.5M、1.6M、1.7M、1.8M、1.9M或2M。

111. 如實施例106至110中任一項之方法，其中(b)之培育引起至少50%、60%、70%、75%、80%、85%、90%、96%或100%的包含ORF1分子或其複本之粒子群體經拆卸。

112. 如實施例106至111中任一項之方法，其中適用於拆卸蛋白質外部之條件足以拆卸包含ORF1分子之至少約20、30、40、50或60個複本或20-30、30-40、40-50或50-60個複本的複合物(例如蛋白質外部)。

113. 如實施例106至112中任一項之方法，其中適用於拆卸蛋白質外部之條件產生每毫升少於$10^8$個剩餘完整粒子。

114. 如實施例106至113中任一項之方法，其在提供(a)之前進一步包含自適用於拆卸蛋白質外部之條件移除ORF1分子(例如使ORF1分子經受非變性條件)。

115. 如實施例114之方法，其中自適用於蛋白質外部之拆卸的條件移除ORF1分子包含降低變性劑之濃度，例如將變性劑(例如尿素)之濃度降低至低於0.1M、0.2M、0.3M、0.4M、0.5M、0.6M、0.7M、0.8M、0.9M、1M、1.1M、1.2M、1.3M、1.5M、1.5M、1.6M、1.7M、1.8M、

1.9M或2M。

116. 如實施例114或115之方法，其中該移除引起一或多個指環載體之形成，該一或多個指環載體各自包裹遺傳元件(例如mRNA)之至少一個複本(例如，至少1、2、3、4、5、6、7、8、9、10、15、20、25、30、40、50、60、70、80、90、100、150、200、250、300、400、500、600、700、800、900或1000個複本)。

117. 如實施例116之方法，其中所得溶液中包裹遺傳元件之至少一個複本的指環載體之數目為至少$10^5$、$10^6$、$10^7$、$10^8$、$10^9$、$10^{10}$或$10^{11}$；或介於$10^5$ - $10^6$、$10^6$ - $10^7$、$10^7$ - $10^9$、$10^8$ - $10^9$、$10^9$ - $10^{10}$或$10^{10}$ - $10^{11}$之間(例如如藉由電子顯微法所量測)。

118. 如實施例116之方法，其中所得溶液中包裹遺傳元件之至少一個複本的指環載體之數目為至少$10^5$、$10^6$、$10^7$、$10^8$、$10^9$、$10^{10}$或$10^{11}$個指環載體/毫升；或介於$10^5$ - $10^6$、$10^6$ - $10^7$、$10^7$ - $10^9$、$10^8$ - $10^9$、$10^9$ - $10^{10}$或$10^{10}$ - $10^{11}$個指環載體/毫升之間(例如如藉由電子顯微法所量測)。

119. 如實施例114-118中任一項之方法，其中移除使得溶液包含至少$10^5$、$10^6$、$10^7$、$10^8$、$10^9$、$10^{10}$或$10^{11}$個指環載體/毫升；或介於$10^5$ - $10^6$、$10^6$ - $10^7$、$10^7$ - $10^9$、$10^8$ - $10^9$、$10^9$ - $10^{10}$或$10^{10}$ - $10^{11}$個指環載體/毫升之間(例如如藉由電子顯微法所量測)，其中指環載體各自包裹遺傳元件(例如mRNA)之至少一個複本(例如至少1、2、3、4、5、6、7、8、9、10、15、20、25、30、40、50、60、70、80、90、100、150、200、250、300、400、500、600、700、800、900或1000個複本)。

120. 如實施例114至119中任一項之方法，其中至少75%、80%、90%、95%、96%、97%、98%、99%或100%之指環載體包含蛋白質外

部，該等蛋白質外部為60聚體或粒子或為直徑為至少30、31、32、33、34或35 nm之粒子。

121. 如實施例93至120中任一項之方法，其中指環載體之遺傳元件對核酸內切酶(例如，RNA酶)具有抗性。

122. 如實施例93至121中任一項之方法，其中(a)包含摻合(i)及(ii)。

123. 如實施例93至122中任一項之方法，其係在無細胞系統中進行。

124. 一種製造指環載體組合物之方法，其包含：

(a)提供複數個如前述實施例中任一項之指環載體或組合物；

(b)視情況評估該複數個以下中之一或多者：本文所描述之污染物、光學密度量測(例如，OD 260)、粒子數目(例如，藉由HPLC)、感染性(例如，粒子:感染性單位比率，例如，如藉由螢光及/或ELISA所測定)；及

(c)例如如果(b)之參數中之一或多者符合指定臨限值時，調配複數個指環載體，例如作為適用於向個體投與之醫藥組合物。

125. 一種治療有需要之個體之疾病或病症之方法，該方法包含向該個體投與如前述實施例中任一項之指環載體或組合物，由此治療個體之疾病或病症(例如如本文所描述)。

126. 一種調節，例如增強或抑制個體之生物功能(例如如本文所描述)的方法，該方法包含向個體投與如前述實施例中任一項之指環載體或組合物。

127. 一種向細胞遞送遺傳元件之方法，該方法包含使該如前述實施例中任一項之指環載體或組合物與細胞，例如真核細胞，例如哺乳動物細

胞，例如人類細胞接觸。

128. 一種如前述實施例中任一項之指環載體或組合物之用途，其係用於治療個體之疾病或病症(例如如本文所描述)。

129. 如前述實施例中任一項之指環載體或組合物，其用於治療個體之疾病或病症(例如如本文所描述)之方法中。

130. 如前述實施例中任一項之指環載體或組合物，其用於製造供治療個體之疾病或病症(例如如本文所描述)用之藥劑。

【0017】 本發明之其他特徵、目標及優點將自實施方式及圖式及自申請專利範圍顯而易知。

【0018】 除非另有定義，否則本文中所使用之所有技術及科學術語具有與一般熟習本發明所屬之技術者通常所理解相同之意義。所有公開案、專利申請案、專利及本文所提及之其他參考案均以全文引用之方式併入。另外，材料、方法及實例僅為說明性的且並不意欲為限制性的。

【圖式簡單說明】

【0019】

圖1為一系列電子顯微圖，展示來自指環病毒株之重組蛋白殼蛋白活體外形成病毒樣粒子(VLP)。如藉由陰性染色電子顯微法所觀測，不同細胞株中產生之蛋白殼蛋白活體外形成VLP。(A)自昆蟲細胞純化的環2 ORF1 VLP的觀測直徑為大約35 nm。(B)自昆蟲細胞純化的環10 ORF1 VLP的觀測直徑為大約35 nm。(C)自哺乳動物細胞純化的CAV VP1 VLP的觀測直徑為大約20 nm。

圖2為展示喙羽病病毒(Beak and Feather Disease Virus，BFDV)之結構中觀測到之八條β股果醬捲域(或果醬捲摺疊)的圖式。按照慣例，β股經

標記為B至I。股形成定向為B-I-D-G及C-H-E-F之四個反平行β片。B-I-D-G片形成病毒蛋白殼之內部。

圖3為描繪指環病毒ORF1蛋白之果醬捲序列相比於喙羽病病毒(BFDV)/E型肝炎蛋白殼蛋白(HepE)之果醬捲域的胺基酸序列比對。

圖4A-4E為展示產生包裹編碼eGFP之mRNA分子的基於指環病毒ORF1蛋白之病毒樣粒子(VLP)的例示性方法之一系列圖式。(A)細胞中產生ORF1蛋白且如本文所述分離。(B)由ORF1蛋白形成的VLP隨後在2 M尿素溶液中拆卸。(C)當在不存在mRNA之情況下移除尿素時，重新形成少量VLP (藉由電子顯微法偵測到效價小於$10^8$個粒子/毫升)。(D，E)當在編碼eGFP之mRNA存在下移除尿素時，藉由電子顯微法(EM)偵測到實質性量之VLP (效價為$1\times10^9$至$1\times10^{10}$個粒子/毫升)。

當結合附圖閱讀時，將更好地理解本發明之實施例的以下詳細描述。出於說明本發明之目的，在圖式中展示本發明例示之實施例。然而，應理解，本發明不限於圖式中所示實施例之精確佈置及工具。

【實施方式】

【0020】

**定義**

將關於特定實施例及參考某些圖式描述本發明，但本發明不限於此，而僅僅受申請專利範圍限制。除非另外指明，否則如下文所闡述之術語一般應按其常識來理解。

【0021】 當術語「包含」用於本發明說明書及申請專利範圍時，其不排除其他要素。出於本發明之目，認為術語「由……組成(consisting of)」為術語「包含(comprising of)」之較佳實施例。若下文中將群組定義

為包含至少某一數目之實施例，則此亦理解為揭示較佳僅由此等實施例組成之群組。

【0022】 除非另外具體規定，否則若在提及單數名詞時使用不定冠詞或定冠詞，例如「一(a/an)」或「該(the)」，則其包括複數個該名詞。

【0023】 措辭「用於治療、調節等的化合物、組合物、產物等」應理解為係指適用於治療、調節等指示目的之化合物、組合物、產物等本身。措辭「用於治療、調節等的化合物、組合物、產物等」另外揭示，作為一實施例，此類化合物、組合物、產物等用於治療、調節等。

【0024】 措辭「用於...之化合物、組合物、產物等」、「化合物、組合物、產物等在製造用於...之藥劑、醫藥組合物、獸醫組合物、診斷組合物等中的用途」或「用作藥劑...之化合物、組合物、產物等」指示此類化合物、組合物、產物等將待用於可在人類或動物身體上實踐之治療方法中。其被視為與關於治療方法等的實施例及技術方案等效之揭示內容。若實施例或技術方案因此係指「用於治療疑似患有疾病之人類或動物的化合物」，則此亦被視為揭示「化合物在製造供治療疑似患有疾病之人類或動物用之藥劑中的用途」或「藉由向疑似患有疾病之人類或動物投與化合物進行治療之方法」。措辭「用於治療、調節等的化合物、組合物、產物等」應理解為係指適用於治療、調節等指示目的之化合物、組合物、產物等本身。

【0025】 若在下文中將術語、值、數目等之實例提供於括號中，則此應理解為括號中提及之實例可構成實施例之指示。例如，若其陳述為「在實施例中，核酸分子包含與表1之編碼指環病毒ORF1的核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%

或100%序列一致性的核酸序列(例如，表1之核酸序列的核苷酸571-2613)」，則隨後一些實施例係關於包含與表1之核酸序列的核苷酸571-2613具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列的核酸分子。

【0026】 如本文所用，術語「擴增」係指複製核酸分子或其部分以產生核酸分子或其部分之一或多個額外複本(例如遺傳元件或遺傳元件區域)。在一些實施例中，擴增引起核酸序列部分複製。在一些實施例中，經由滾環複製進行擴增。

【0027】 如本文所用，術語「指環載體」係指包含遺傳元件，例如RNA，例如環狀RNA，包裹於蛋白質外部中之媒劑。在一些實施例中，遺傳元件實質上由蛋白質外部保護而不被RNA酶消化。如本文所用，「合成指環載體」通常係指非天然存在的指環載體，例如具有相對於野生型病毒(例如如本文所述之野生型指環病毒)不同的序列。在一些實施例中，合成指環載體經工程化或重組，例如包含相對於野生型病毒基因體(例如如本文所述之野生型指環病毒基因體)包含差異或修飾的遺傳元件。在一些實施例中，包裹於蛋白質外部內涵蓋由蛋白質外部進行之100%覆蓋以及小於100%覆蓋，例如95%、90%、85%、80%、70%、60%、50%或更小覆蓋。例如，只要遺傳元件保留在蛋白質外部中或經保護而不被RNA酶消化，例如在進入宿主細胞之前，則蛋白質外部中可存在間隙或不連續處(例如，使蛋白質外部對水、離子、肽或小分子可透)。在一些實施例中，指環載體經純化，例如其與其原始來源分離及/或實質上不含(>50%、>60%、>70%、>80%、>90%)其他組分。在一些實施例中，指環載體能夠將遺傳元件引入至目標細胞中(例如經由感染)。在一些實施例中，指環

載體為含有某些指環病毒元件，諸如指環病毒ORF1分子之感染性合成病毒粒子。

【0028】 如本文所用，術語「抗體分子」係指蛋白質，例如免疫球蛋白鏈或其片段，其包含至少一個免疫球蛋白可變域序列。術語「抗體分子」涵蓋全長抗體及抗體片段(例如scFv)。在一些實施例中，抗體分子為多特異性抗體分子，例如該抗體分子包含複數個免疫球蛋白可變域序列，其中該複數個中之第一免疫球蛋白可變域序列對第一抗原決定基具有結合特異性且該複數個中之第二免疫球蛋白可變域序列對第二抗原決定基具有結合特異性。在一些實施例中，多特異性抗體分子為雙特異性抗體分子。雙特異性抗體分子之一般由對於第一抗原決定基具有結合特異性之第一免疫球蛋白可變域序列及對於第二抗原決定基具有結合特異性之第二免疫球蛋白可變域序列特徵化。

【0029】 如本文所用，術語「主鏈」或「主鏈區」係指包含一或多個參與在宿主細胞中複製及/或維持核酸分子(例如對其為必需及/或足夠的)之元件的核酸分子內(例如在例如如本文所述之桿狀病毒質體或供體載體內)的區域。在一些實施例中，主鏈區，諸如「桿狀病毒主鏈區」包含一或多種桿狀病毒元件(例如桿狀病毒基因體或其功能片段)，例如適用於在昆蟲細胞(例如Sf9細胞)中複製核酸構築體。在一些實施例中，主鏈進一步包含可選標記。在一些實施例中，核酸分子包含遺傳元件區域及主鏈區(例如適用於在細菌細胞中進行複製的桿狀病毒主鏈區及/或主鏈區)。

【0030】 如本文所用，術語「桿狀病毒質體」係指包含足夠桿狀病毒主鏈元件，以使其適用於在昆蟲細胞中複製，且此外適用於在細菌細胞中複製的核酸分子。在一些實施例中，核酸分子適用於在細菌細胞(例如

大腸桿菌細胞，例如DH 10Bac細胞)中複製。

【0031】 如本文所用，「環狀」核酸係指形成無游離5'或3'端之結構的核酸。在一些實施例中，環狀核酸經由共價或非共價鍵閉合。舉例而言，環狀核酸可以藉由用例如磷酸鹽-糖鍵或合成連接子部分共價連接線性核酸之末端來製成。在其他實施例中，環狀核酸包含鄰近且非游離(與核酸外切酶實質上不可接近)之兩個末端。舉例而言，環狀核酸可藉由使線性核酸之末端直接或經由核酸夾板雜交來製成。

【0032】 如本文所用，「DNA區」係指包含複數個DNA核苷酸之聚核苷酸股的一部分。舉例而言，在一些實施例中，DNA區為併入至RNA股中之複數個DNA核苷酸。舉例而言，DNA區在聚核苷酸股內包含約5-10、10-20、20-30、30-40、40-50、50-60、60-70、70-80、80-90或90-100個DNA核苷酸。

【0033】 如本文所用，「編碼」之核酸係指編碼胺基酸序列或聚核苷酸，例如mRNA或功能性聚核苷酸(例如非編碼RNA，例如siRNA或miRNA)的核酸序列。

【0034】 如本文所用，「外源性」試劑(例如效應子、核酸(例如RNA)、基因、有效負載、蛋白質)係指不包含相應野生型病毒(例如如本文所述之指環病毒)或不由其編碼的試劑。在一些實施例中，外源性試劑並非天然存在，諸如具有相對於天然存在之蛋白質或核酸改變(例如藉由插入、缺失或取代)之序列的蛋白質或核酸。在一些實施例中，外源性試劑並不天然存在於宿主細胞中。在一些實施例中，外源性試劑天然存在於宿主細胞中但對於病毒為外源性的。在一些實施例中，外源性試劑天然存在於宿主細胞中，但不以所需水準或以所需時間存在。

【0035】 如本文所用之關於另一試劑或元件(例如效應子、核酸序列、胺基酸序列)的「異源性」試劑或元件(例如效應子、核酸序列、胺基酸序列)係指例如在野生型病毒(例如指環病毒)中未天然發現於一起的試劑或元件。在一些實施例中，異源核酸序列可與天然存在之核酸序列(例如指環病毒中天然存在之序列)存在於相同核酸中。在一些實施例中，異源性試劑或元件相對於指環病毒為外源性的，其中指環載體之其他元件(例如剩餘部分)係基於該指環病毒。

【0036】 如本文所用，術語「遺傳元件」係指包裹於或可包裹於蛋白質外部內(例如由蛋白質外部保護而不被RNA酶消化)，以例如形成如本文所述之指環載體的核酸分子。應理解，遺傳元件可以裸RNA形式產生且視情況進一步組裝成蛋白質外部。亦應理解，指環載體可將其遺傳元件插入細胞中，使得遺傳元件存在於細胞中且蛋白質外部不一定進入細胞。

【0037】 如本文所用，「遺傳元件構築體」係指包含遺傳元件序列或其片段之核酸構築體(例如質體、桿狀病毒質體、供體載體、黏質體或微型環)。在一些實施例中，如本文所述之桿狀病毒質體或供體載體為包含遺傳元件序列或其片段之遺傳元件構築體。

【0038】 如本文所用，術語「遺傳元件區域」係指包含遺傳元件之序列的構築體區域。在一些實施例中，遺傳元件區域包含與野生型指環病毒序列或其片段具有足夠一致性之序列，以由蛋白質外部包裹，藉此形成指環載體(例如與野生型指環病毒序列或其片段具有至少70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的序列)。在實施例中，遺傳元件區域包含蛋白質結合序列，例如如本文所述(例如如本文所述之5' UTR、3' UTR及/或富含GC之區域，或與其具有

至少70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的序列)。在一些實施例中，包含遺傳元件區域之構築體不包裹於蛋白質外部中，但由構築體產生之遺傳元件可包裹於蛋白質外部中。在一些實施例中，包含遺傳元件區域之構築體進一步包含載體主鏈(例如，桿狀病毒質體主鏈或供體載體主鏈)。在一些實施例中，構築體(例如桿狀病毒質體)包含一或多個桿狀病毒元件(例如桿狀病毒基因體，例如包含遺傳元件區域)。

【0039】 如本文所用，當相對於基因體(例如指環病毒基因體)或其片段使用時，術語「突變體」係指相對於對應野生型指環病毒序列具有至少一種變化的序列。在一些實施例中，相對於對應野生型指環病毒序列，突變體基因體或其片段包含至少一種單核苷酸多型性、添加、缺失或框移。在一些實施例中，相對於對應野生型指環病毒序列，突變體基因體或其片段包含至少一個指環病毒ORF (例如ORF1、ORF2、ORF2/2、ORF2/3、ORF1/1及/或ORF1/2中之一或多者)的缺失。在一些實施例中，相對於對應野生型指環病毒序列，突變體基因體或其片段包含所有指環病毒ORF (例如所有ORF1、ORF2、ORF2/2、ORF2/3、ORF1/1及ORF1/2)的缺失。在一些實施例中，相對於對應野生型指環病毒序列，突變體基因體或其片段包含至少一個指環病毒非編碼區域(例如5' UTR、3' UTR及/或富含GC之區域中之一或多者)的缺失。在一些實施例中，突變基因體或其片段包含或編碼外源性效應子。

【0040】 如本文所用，術語「ORF分子」係指具有指環病毒ORF蛋白(例如包含指環病毒ORF1、ORF2、ORF2/2、ORF2/3、ORF1/1及/或ORF1/2蛋白之多肽)之活性及/或結構特徵的多肽，或其功能片段。當一般

使用(亦即「ORF分子」)時，多肽可包含本文所述之指環病毒ORF (例如指環病毒ORF1、ORF2、ORF2/2、ORF2/3、ORF1/1及/或ORF1/2中之任一者)或其功能片段中之任一者的活性及/或結構特徵。當與修飾劑一起用於指示特定開放閱讀框架(例如「ORF1分子」、「ORF2分子」、「ORF2/2分子」、「ORF2/3分子」、「ORF1/1分子」或「ORF1/2分子」)時，通常意謂多肽包含對應指環病毒ORF蛋白或其功能片段(例如如下文所定義之「ORF1分子」)之活性及/或結構特徵。舉例而言，「ORF2分子」包含指環病毒ORF2蛋白或其功能片段的活性及/或結構特徵。

【0041】 如本文所用，術語「ORF1分子」係指具有指環病毒ORF1蛋白(例如如本文所述的指環病毒ORF1蛋白)或其功能片段之活性及/或結構特徵的多肽。在一些情況下，ORF1分子可以包含以下一或多者(例如1、2、3或4者)：包含至少60%鹼性殘基(例如至少60%精胺酸殘基)之第一區域、包含至少約六條β股之第二區域(例如至少4、5、6、7、8、9、10、11或12條β股)、包含指環病毒N22域(例如如本文所述，例如如本文所述之指環病毒ORF1蛋白的N22域)之結構或活性的第三區域、及/或包含指環病毒C端域(CTD) (例如如本文所述，例如如本文所述之指環病毒ORF1蛋白的CTD)之結構或活性的第四區域。在一些實施例中，ORF1分子以N端至C端之次序包含第一、第二、第三及第四區域。在一些實施例中，指環載體包含ORF1分子，該分子以N端至C端之次序包含第一、第二、第三及第四區域。在一些情況下，ORF1分子可包含由指環病毒ORF1核酸編碼之多肽。在一些實施例中，ORF1分子可進一步包含異源序列，例如高變區(HVR)，例如來自指環病毒ORF1蛋白之HVR，例如如本文所描述。如本文所用，「指環病毒ORF1蛋白」係指由指環病毒基因體(例如

野生型指環病毒基因體，例如如本文所述)編碼之ORF1蛋白。

【0042】 如本文所用，術語「ORF2分子」係指具有指環病毒ORF2蛋白(例如如本文所述的指環病毒ORF2蛋白)或其功能片段之活性及/或結構特徵的多肽。如本文所用，「指環病毒ORF2蛋白」係指由指環病毒基因體(例如野生型指環病毒基因體，例如如本文所述)編碼之ORF2蛋白。

【0043】 如本文所用，術語「蛋白質外部」係指主要為(例如>50%、>60%、>70%、>80%、>90%)蛋白質之外部組分。

【0044】 如本文所用，術語「調控核酸」係指修飾編碼表現產物之DNA序列之表現(例如轉錄及/或轉譯)的核酸序列。在一些實施例中，表現產物包含RNA或蛋白質。

【0045】 如本文所用，術語「調控序列」係指修飾目標基因產物之轉錄的核酸序列。在一些實施例中，調控序列為啟動子或強化子。

【0046】 如本文所用，「實質上非病原性」生物體、粒子或組分係指不會例如在宿主生物體(例如哺乳動物，例如人類)中引起或誘導不可接受之疾病或病原性病狀之生物體、粒子(例如病毒或指環載體，例如如本文所述)或其組分。在一些實施例中，向個體投與指環載體可引起作為標準照護之一部分可接受的輕微反應或副作用。

【0047】 如本文所用，術語「非病原性」係指例如在宿主生物體，例如哺乳動物，例如人類中，不會引起或誘導不希望之病狀(例如疾病或病原性病狀)的生物體或其組分。

【0048】 如本文所用，「實質上非整合型」遺傳元件係指一種遺傳元件，例如病毒或指環載體中之遺傳元件，例如如本文所述，其中進入宿主細胞(例如真核細胞)或生物體(例如哺乳動物，例如人類)中之小於約

0.01%、0.05%、0.1%、0.5%或1%之遺傳元件整合至基因體中。在一些實施例中，遺傳元件不會以可偵測方式整合至例如宿主細胞之基因體中。在一些實施例中，可使用如本文所述之技術，例如核酸定序、PCR偵測及/或核酸雜交來偵測遺傳元件至基因體中之整合。在一些實施例中，整合頻率藉由對自游離載體分離之基因體DNA的定量凝膠純化分析來測定，例如如Wang等人(2004, *Gene Therapy* 11: 711-721，以全文引用之方式併入本文中)中所述。

【0049】 如本文所用，「實質上非免疫原性」生物體、粒子或組分係指不會例如在宿主組織或生物體(例如哺乳動物，例如人類)中引起或誘導非所需或非靶向免疫反應的生物體、粒子(例如病毒或指環載體，例如如本文所述)或其組分。在一些實施例中，實質上非免疫原性生物體、粒子或組分不產生臨床上顯著的免疫反應。在一些實施例中，實質上非免疫原性指環載體不產生針對蛋白質的臨床顯著免疫反應，該蛋白質包含胺基酸序列或由指環病毒或指環載體遺傳元件之核酸序列編碼。在一些實施例中，免疫反應(例如非所需或非靶向免疫反應)係藉由分析個體中之抗體(例如中和抗體)存在或含量(例如抗指環載體抗體之存在或含量，例如針對如本文所述之指環載體之抗體的存在或含量)來偵測，例如根據Tsuda等人中所述之抗TTV抗體偵測方法(1999; *J. Virol. Methods* 77: 199-206；以引用之方式併入本文中)及/或Kakkola等人中所述之用於測定抗TTV IgG含量之方法(2008; *Virology* 382: 182-189；以引用之方式併入本文中)。針對指環病毒或基於指環病毒之指環載體的抗體(例如中和抗體)亦可藉由此項技術中用於偵測抗病毒抗體之方法來偵測，例如偵測抗AAV抗體之方法，例如如Calcedo等人(2013; *Front. Immunol.* 4(341): 1-7；以引用之方

式併入本文中)中所述。

【0050】 如本文所用之「子序列」係指分別包含於較大核酸序列或胺基酸序列中的核酸序列或胺基酸序列。在一些情況下，子序列可包含較大序列之結構域或功能片段。在一些情況下，子序列可包含當與較大序列分離時能夠形成二級及/或三級結構之較大序列的片段，類似於當與較大序列之其餘部分一起存在時由子序列形成之二級及/或三級結構。在一些情況下，子序列可經另一序列置換(例如，包含外源性序列之子序列或與較大序列之其餘部分異源的序列，例如來自不同指環病毒之對應子序列)。

【0051】 本發明大體上係關於指環載體，例如合成指環載體，及其用途。本發明提供指環載體、包含指環載體之組合物及製成或使用指環載體之方法。指環載體通常可用作遞送媒劑，例如用於將治療劑遞送至真核細胞。一般而言，本文所述之指環載體將包括包含RNA序列(例如編碼效應子，例如外源性效應子或內源性效應子之RNA序列)之遺傳元件，該RNA序列包裹於蛋白質外部內。指環載體可包括相對於指環病毒序列(例如如本文所述)之一或多個序列缺失(例如如本文所述之區域或域)。指環載體可用作實質上非免疫原性媒劑，用於將遺傳元件或其中編碼之效應子(例如多肽或核酸效應子，例如如本文所述)遞送至真核細胞中，例如用於治療包含該等細胞之個體的疾病或病症。

【0052】

**目錄**

I.藉由活體外組裝製成指環載體之組合物及方法

A.指環載體之組分及組裝

i.用於組裝指環載體之ORF1分子

ii.用於組裝指環載體之ORF2分子

B.遺傳元件

i.包含RNA之遺傳元件

a.僅RNA遺傳元件

b.雜交RNA-ssDNA遺傳元件

c.RNA/DNA結合物

C.遺傳元件構築體

D.基於RNA之遺傳元件之產生

E.蛋白質組分之產生

i.桿狀病毒表現系統

ii.昆蟲細胞系統

iii.哺乳動物細胞系統

F.效應子

G.活體外組裝方法

H.增濃及純化

II.指環載體

A.指環病毒

B.ORF1分子

C.ORF2分子

D.遺傳元件

E.蛋白質結合序列

F. 5' UTR區

G.富含GC之區域

H.效應子

I.調控序列

J.其他序列

K.蛋白質外部

III.使用方法

IV.投與/遞送

【0053】

**I.藉由活體外組裝製成指環載體之組合物及方法**

在一些態樣中，本發明提供可用於產生如本文所描述之指環載體，例如具有包含RNA之遺傳元件的指環載體的組合物及方法。在一些實施例中，本文所述之組合物及方法可用於產生遺傳元件或遺傳元件構築體。在一些實施例中，本文所描述之組合物及方法可用於藉由活體外組裝產生遺傳元件或遺傳元件構築體。在一些實施例中，本文所述之組合物及方法可以用於產生一或多個指環病毒ORF分子(例如ORF1、ORF2、ORF2/2、ORF2/3、ORF1/1或ORF1/2分子，或其功能片段或剪接變異體)。在一些實施例中，本文所述之組合物及方法可用於例如在宿主細胞(例如昆蟲細胞，例如Sf9細胞)中產生蛋白質外部或其組分(例如ORF1分子)。

【0054】

***指環載體之組分及組裝***

本文中之組合物及方法可用於產生指環載體。如本文所述，指環載體一般包含包裹於蛋白質外部(例如包含由指環病毒ORF1核酸編碼之多肽，例如如本文所述)內之遺傳元件(例如RNA分子)。在一些實施例中，

遺傳元件包含一或多個編碼指環病毒ORF (例如指環病毒ORF1、ORF2、ORF2/2、ORF2/3、ORF1/1或ORF1/2中之一或多者)的序列。如本文所用，指環病毒ORF或ORF分子(例如指環病毒ORF1、ORF2、ORF2/2、ORF2/3、ORF1/1或ORF1/2)包括多肽，該多肽包含與對應指環病毒ORF序列具有至少70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的胺基酸序列，例如如PCT/US2018/037379或PCT/US19/65995中所述(其各自以全文引用之方式併入本文中)。在實施例中，遺傳元件包含編碼指環病毒ORF1或其剪接變異體或功能片段的序列(例如果醬捲區，例如如本文所述)。在一些實施例中，蛋白質外部包含由指環病毒ORF1核酸(例如指環病毒ORF1分子或其剪接變異體或功能片段)編碼之多肽。

【0055】 在一些實施例中，藉由在蛋白質外部(例如如本文所述)內包裹遺傳元件(例如如本文所述)來組裝指環載體。在一些實施例中，遺傳元件包裹於宿主細胞(例如昆蟲細胞，例如Sf9細胞)中之蛋白質外部內。在一些實施例中，宿主細胞表現蛋白質外部中所包含之一或多種多肽(例如由指環病毒ORF1核酸編碼之多肽，例如ORF1分子)。舉例而言，在一些實施例中，宿主細胞包含編碼指環病毒ORF1分子之核酸序列，該分子例如指環病毒ORF1多肽之剪接變異體或功能片段(例如野生型指環病毒ORF1蛋白或由野生型指環病毒ORF1核酸編碼之多肽，例如如本文所述)。在實施例中，編碼指環病毒ORF1分子之核酸序列包含於核酸構築體(例如質體、病毒載體、病毒、微型環、桿狀病毒質體或人工染色體)中，該核酸構築體包含於宿主細胞中。在實施例中，將編碼指環病毒ORF1分子之核酸序列整合至宿主細胞之基因體中。

【0056】 在一些實施例中，宿主細胞包含遺傳元件及/或包含遺傳元件之序列的核酸構築體。在一些實施例中，核酸構築體係選自質體、病毒核酸、微型環、桿狀病毒質體或人工染色體。在一些實施例中，遺傳元件自核酸構築體(例如桿狀病毒質體)切除，且視情況自雙股形式轉化成單股形式(例如藉由變性)。在一些實施例中，遺傳元件由基於核酸構築體(例如桿狀病毒質體)中之模板序列的聚合酶產生。在一些實施例中，聚合酶產生遺傳元件序列之單股複本，其可視情況環化以形成如本文所述之遺傳元件。

【0057】 在一些實施例中，宿主細胞包含遺傳元件構築體(例如桿狀病毒質體、質體或微型環)及包含一或多個編碼指環病毒ORF分子(例如ORF1、ORF2、ORF2/2、ORF2/3、ORF1/1及/或ORF1/2 ORF分子)或其功能片段之序列的桿狀病毒質體。在一些實施例中，蛋白質外部蛋白質由桿狀病毒質體表現。在實施例中，蛋白質外部蛋白質由包裹遺傳元件之桿狀病毒質體表現，藉此形成指環載體。在一些實施例中，桿狀病毒質體包含適用於在昆蟲細胞(例如Sf9細胞)中複製核酸構築體之主鏈，例如桿狀病毒主鏈區。在一些實施例中，桿狀病毒質體包含適用於在細菌細胞(例如大腸桿菌細胞，例如DH 10Bac細胞)中複製遺傳元件構築體之主鏈區。在一些實施例中，遺傳元件構築體包含適用於在昆蟲細胞(例如Sf9細胞)中複製核酸構築體之主鏈，例如桿狀病毒主鏈區。在一些實施例中，遺傳元件構築體包含適用於在細菌細胞(例如大腸桿菌細胞，例如DH 10Bac細胞)中複製遺傳元件構築體之主鏈區。在一些實施例中，經由桿狀病毒粒子將桿狀病毒質體引入至宿主細胞中。在實施例中，桿狀病毒質體藉由生產細胞，例如昆蟲細胞(例如Sf9細胞)或細菌細胞(例如大腸桿菌細胞，例

如DH 10Bac細胞)產生。在實施例中，生產細胞包含桿狀病毒質體及/或供體載體，例如如本文所述。在實施例中，生產細胞進一步包含用於複製桿狀病毒質體及/或供體載體之充足細胞機構。

【0058】

***例如用於組裝指環載體之ORF1分子***

指環載體可例如藉由將遺傳元件包裹於蛋白質外部內而製成。在一些實施例中，包裹發生在無細胞系統或細胞中。指環載體之蛋白質外部一般包含由指環病毒ORF1核酸(例如指環病毒ORF1分子或其剪接變異體或功能片段，例如如本文所描述)編碼之多肽。在一些實施例中，ORF1分子可包含以下中之一或多者：包含富含精胺酸之區域的第一區域，例如具有至少60%鹼性殘基(例如至少60%、65%、70%、75%、80%、85%、90%、95%或100%鹼性殘基；例如60%-90%、60%-80%、70%-90%或70-80%之間的鹼性殘基)之區域，及包含果醬捲域之第二區域，例如至少六條β股(例如4、5、6、7、8、9、10、11或12條β股)。在實施例中，蛋白質外部包含指環病毒ORF1富含精胺酸之區域、果醬捲區、N22域、高變區及/或C端域中之一或多者(例如1、2、3、4或全部5者)。在一些實施例中，蛋白質外部包含指環病毒ORF1果醬捲區(例如如本文所描述)。在一些實施例中，蛋白質外部包含指環病毒ORF1富含精胺酸之區域(例如如本文所描述)。在一些實施例中，蛋白質外部包含指環病毒ORF1 N22域(例如如本文所描述)。在一些實施例中，蛋白質外部包含指環病毒高變區(例如如本文所述)。在一些實施例中，蛋白質外部包含指環病毒ORF1 C端域(例如如本文所描述)。

【0059】在一些實施例中，指環載體包含ORF1分子及/或編碼

ORF1分子之核酸。一般而言，ORF1分子包含具有指環病毒ORF1蛋白(例如如本文所述之指環病毒ORF1蛋白)或其功能片段之結構特徵及/或活性的多肽。在一些實施例中，ORF1分子包含相對於指環病毒ORF1蛋白(例如如本文所述之指環病毒ORF1蛋白)之截短。在一些實施例中，ORF1分子經截短指環病毒ORF1蛋白之至少10、20、30、40、50、60、70、80、90、100、150、200、250、300、350、400、450、500、550、600、650或700個胺基酸。在一些實施例中，ORF1分子包含與例如如本文所描述之α細環病毒屬、β細環病毒屬或γ細環病毒屬ORF1蛋白具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%序列一致性的胺基酸序列。ORF1分子通常可結合至核酸分子，諸如DNA (例如遺傳元件，例如如本文所描述)。在一些實施例中，ORF1分子定位至細胞核。在某些實施例中，ORF1分子定位至細胞之細胞核。

【0060】 不希望受理論所束縛，ORF1分子可能能夠結合至其他ORF1分子，例如以形成蛋白質外部(例如如本文所描述)。此類ORF1分子可描述為能夠形成蛋白殼。在一些實施例中，蛋白質外部可包裹核酸分子(例如，如本文所述之遺傳元件)。在一些實施例中，複數個ORF1分子可形成多聚體，例如以產生蛋白質外部。在一些實施例中，多聚體可為均多聚體。在其他實施例中，多聚體可為雜多聚體。

【0061】

**例如用於組裝指環載體之*ORF2分子***

使用本文所描述之組合物或方法產生指環載體可涉及表現指環病毒ORF2分子(例如如本文所描述)或其剪接變異體或功能片段。在一些實施例中，指環載體包含ORF2分子或其剪接變異體或功能片段，及/或編碼

ORF2分子或其剪接變異體或功能片段之核酸。在一些實施例中，指環載體不包含ORF2分子或其剪接變異體或功能片段，及/或編碼ORF2分子或其剪接變異體或功能片段之核酸。在一些實施例中，產生指環載體包含表現ORF2分子或其剪接變異體或功能片段，但ORF2分子不併入至指環載體中。

【0062】

***遺傳元件***

***包含RNA之遺傳元件***

在一些實施例中，遺傳元件為或包含核酸。在一些實施例中，遺傳元件為單股聚核苷酸。在一些實施例中，遺傳元件包含一或多個雙股區。在一些實施例中，遺傳元件包含RNA。在一些實施例中，遺傳元件包含RNA髮夾結構。在一些實施例中，遺傳元件為mRNA，例如經化學修飾之mRNA。在一些實施例中，遺傳元件由至少10%、20%、25%、30%、40%、50%、60%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100% RNA組成。在一些實施例中，遺傳元件包含DNA股及RNA股，例如其中DNA股之至少一部分與RNA股之至少一部分雜交。

【0063】在一些實施例中，遺傳元件不編碼指環病毒ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3或ORF2t/3中之任一者。

【0064】在一些實施例中，RNA遺傳元件編碼效應子，例如效應蛋白。

【0065】在一些實施例中，RNA遺傳元件為或包含效應子，例如功能性RNA。在一些實施例中，RNA係選自由以下組成之群：mRNA、rRNA、tRNA (例如TREM)、調控性RNA、非編碼RNA、長鏈非編碼

RNA (lncRNA)、環狀RNA (circRNA)、雙股RNA (dsRNA)、引導RNA (gRNA)、小干擾RNA (siRNA)、短髮夾RNA (shRNA)、piwi相互作用RNA (piRNA)、小核仁RNA (snoRNA)、小核RNA (snRNA)、細胞外RNA (exRNA)、小卡哈爾體特異性RNA (small Cajal body-specific RNA，scaRNA)、微小RNA (miRNA)及其他RNAi分子。

【0066】 在一些實施例中，遺傳元件包含RNA，例如經化學修飾之RNA。在一些實施例中，遺傳元件之RNA之一或多個核苷酸經化學修飾。在一些實施例中，RNA包含一或多個鹼基之一或多個化學修飾。在一些實施例中，RNA包含一或多種糖之一或多種化學修飾。在一些實施例中，遺傳元件之RNA之至少1、2、3、4、5、6、7、8、9或10個核苷酸經化學修飾。在一些實施例中，RNA包含一或多個主鏈修飾。在一些實施例中，修飾包含非天然存在之修飾，例如表5-9中之任一者中所描述之修飾。非天然存在之修飾可根據此項技術中已知之方法進行。

【0067】 在一些實施例中，本文所述之遺傳元件包含表5中所提供之非天然存在之修飾或其組合。

**表5：例示性非天然存在之修飾**

| 名稱 | 符號 | 鹼基 | 天然存在 |
|---|---|---|---|
| 7-去氮-腺苷 | -- | A | 否 |
| Nl-甲基-腺苷 | -- | A | 否 |
| N6, N6 (二甲基)腺嘌呤 | -- | A | 否 |
| N6-順-羥基-異戊烯基-腺苷 | -- | A | 否 |
| a-硫基-腺苷 | -- | A | 否 |
| 2 (胺基)腺嘌呤 | -- | A | 否 |
| 2 (胺基丙基)腺嘌呤 | -- | A | 否 |
| 2 (甲基硫基) N6 (異戊烯基)腺嘌呤 | -- | A | 否 |
| 2-(烷基)腺嘌呤 | -- | A | 否 |
| 2-(胺基烷基)腺嘌呤 | -- | A | 否 |
| 2-(胺基丙基)腺嘌呤 | -- | A | 否 |
| 2-(鹵基)腺嘌呤 | -- | A | 否 |
| 2-(鹵基)腺嘌呤 | -- | A | 否 |

| | | | |
|---|---|---|---|
| 2-(丙基)腺嘌呤 | -- | A | 否 |
| 2'-胺基-2'-去氧-ATP | -- | A | 否 |
| 2'-疊氮基-2'-去氧-ATP | -- | A | 否 |
| 2'-去氧-2'-a-胺基腺苷TP | -- | A | 否 |
| 2'-去氧-2'-a-疊氮基腺苷TP | -- | A | 否 |
| 6 (烷基)腺嘌呤 | -- | A | 否 |
| 6 (甲基)腺嘌呤 | -- | A | 否 |
| 6-(烷基)腺嘌呤 | -- | A | 否 |
| 6-(甲基)腺嘌呤 | -- | A | 否 |
| 7 (去氮)腺嘌呤 | -- | A | 否 |
| 8 (烯基)腺嘌呤 | -- | A | 否 |
| 8 (炔基)腺嘌呤 | -- | A | 否 |
| 8 (胺基)腺嘌呤 | -- | A | 否 |
| 8 (硫代烷基)腺嘌呤 | -- | A | 否 |
| 8-(烯基)腺嘌呤 | -- | A | 否 |
| 8-(烷基)腺嘌呤 | -- | A | 否 |
| 8-(炔基)腺嘌呤 | -- | A | 否 |
| 8-(胺基)腺嘌呤 | -- | A | 否 |
| 8-(鹵基)腺嘌呤 | -- | A | 否 |
| 8-(羥基)腺嘌呤 | -- | A | 否 |
| 8-(硫代烷基)腺嘌呤 | -- | A | 否 |
| 8-(硫醇)腺嘌呤 | -- | A | 否 |
| 8-疊氮基-腺苷 | -- | A | 否 |
| 氮雜腺嘌呤 | -- | A | 否 |
| 去氮腺嘌呤 | -- | A | 否 |
| N6 (甲基)腺嘌呤 | -- | A | 否 |
| N6-(異戊基)腺嘌呤 | -- | A | 否 |
| 7-去氮-8-氮雜-腺苷 | -- | A | 否 |
| 7-甲基腺嘌呤 | -- | A | 否 |
| 1-去氮腺苷TP | -- | A | 否 |
| 2'氟-N6-Bz-去氧腺苷TP | -- | A | 否 |
| 2'-OMe-2-胺基-ATP | -- | A | 否 |
| 2'O-甲基-N6-Bz-去氧腺苷TP | -- | A | 否 |
| 2'-a-乙炔基腺苷TP | -- | A | 否 |
| 2-胺基腺嘌呤 | -- | A | 否 |
| 2-胺基腺苷TP | -- | A | 否 |
| 2-胺基-ATP | -- | A | 否 |
| 2'-a-三氟甲基腺苷TP | -- | A | 否 |
| 2-疊氮基腺苷TP | -- | A | 否 |
| 2'-b-乙炔基腺苷TP | -- | A | 否 |
| 2-溴腺苷TP | -- | A | 否 |
| 2'-b-三氟甲基腺苷TP | -- | A | 否 |

| | | | |
|---|---|---|---|
| 2-氯腺苷TP | -- | A | 否 |
| 2'-去氧-2',2'-二氟腺苷TP | -- | A | 否 |
| 2'-去氧-2'-a-巰基腺苷TP | -- | A | 否 |
| 2'-去氧-2'-a-硫代甲氧基腺苷TP | -- | A | 否 |
| 2'-去氧-2'-b-胺基腺苷TP | -- | A | 否 |
| 2'-去氧-2'-b-疊氮基腺苷TP | -- | A | 否 |
| 2'-去氧-2'-b-溴腺苷TP | -- | A | 否 |
| 2'-去氧-2'-b-氯腺苷TP | -- | A | 否 |
| 2'-去氧-2'-b-氟腺苷TP | -- | A | 否 |
| 2'-去氧-2'-b-碘腺苷TP | -- | A | 否 |
| 2'-去氧-2'-b-巰基腺苷TP | -- | A | 否 |
| 2'-去氧-2'-b-硫代甲氧基腺苷TP | -- | A | 否 |
| 2-氟腺苷TP | -- | A | 否 |
| 2-碘腺苷TP | -- | A | 否 |
| 2-巰基腺苷TP | -- | A | 否 |
| 2-甲氧基-腺嘌呤 | -- | A | 否 |
| 2-甲基硫代-腺嘌呤 | -- | A | 否 |
| 2-三氟甲基腺苷TP | -- | A | 否 |
| 3-去氮-3-溴腺苷TP | -- | A | 否 |
| 3-去氮-3-氯腺苷TP | -- | A | 否 |
| 3-去氮-3-氟腺苷TP | -- | A | 否 |
| 3-去氮-3-碘腺苷TP | -- | A | 否 |
| 3-去氮腺苷TP | -- | A | 否 |
| 4'-疊氮基腺苷TP | -- | A | 否 |
| 4'-碳環腺苷TP | -- | A | 否 |
| 4'-乙炔基腺苷TP | -- | A | 否 |
| 5'-均-腺苷TP | -- | A | 否 |
| 8-氮雜-ATP | -- | A | 否 |
| 8-溴-腺苷TP | -- | A | 否 |
| 8-三氟甲基腺苷TP | -- | A | 否 |
| 9-去氮腺苷TP | -- | A | 否 |
| 2-胺基嘌呤 | -- | A/G | 否 |
| 7-去氮-2,6-二胺基嘌呤 | -- | A/G | 否 |
| 7-去氮-8-氮雜-2,6-二胺基嘌呤 | -- | A/G | 否 |
| 7-去氮-8-氮雜-2-胺基嘌呤 | -- | A/G | 否 |
| 2,6-二胺基嘌呤 | -- | A/G | 否 |
| 7-去氮-8-氮雜-腺嘌呤、7-去氮-2-胺基嘌呤 | -- | A/G | 否 |
| 4-甲基胞嘧啶核苷 | -- | C | 否 |
| 5-氮雜-胞嘧啶核苷 | -- | C | 否 |
| 假-異-胞嘧啶核苷 | -- | C | 否 |
| 吡咯并-胞嘧啶核苷 | -- | C | 否 |
| a-硫基-胞嘧啶核苷 | -- | C | 否 |

| | | | |
|---|---|---|---|
| 2-(硫基)胞嘧啶 | -- | C | 否 |
| 2'-胺基-2'-去氧-CTP | -- | C | 否 |
| 2'-疊氮基-2'-去氧-CTP | -- | C | 否 |
| 2'-去氧-2'-a-胺基胞嘧啶核苷TP | -- | C | 否 |
| 2'-去氧-2'-a-疊氮基胞嘧啶核苷TP | -- | C | 否 |
| 3 (去氮) 5 (氮雜)胞嘧啶 | -- | C | 否 |
| 3 (甲基)胞嘧啶 | -- | C | 否 |
| 3-(烷基)胞嘧啶 | -- | C | 否 |
| 3-(去氮) 5 (氮雜)胞嘧啶 | -- | C | 否 |
| 3-(甲基)胞嘧啶核苷 | -- | C | 否 |
| 4,2'-O-二甲基胞嘧啶核苷 | -- | C | 否 |
| 5 (鹵基)胞嘧啶 | -- | C | 否 |
| 5 (甲基)胞嘧啶 | -- | C | 否 |
| 5 (丙炔基)胞嘧啶 | -- | C | 否 |
| 5 (三氟甲基)胞嘧啶 | -- | C | 否 |
| 5-(烷基)胞嘧啶 | -- | C | 否 |
| 5-(炔基)胞嘧啶 | -- | C | 否 |
| 5-(鹵基)胞嘧啶 | -- | C | 否 |
| 5-(丙炔基)胞嘧啶 | -- | C | 否 |
| 5-(三氟甲基)胞嘧啶 | -- | C | 否 |
| 5-溴-胞嘧啶核苷 | -- | C | 否 |
| 5-碘-胞嘧啶核苷 | -- | C | 否 |
| 5-丙炔基胞嘧啶 | -- | C | 否 |
| 6-(偶氮基)胞嘧啶 | -- | C | 否 |
| 6-氮雜-胞嘧啶核苷 | -- | C | 否 |
| 氮雜胞嘧啶 | -- | C | 否 |
| 去氮胞嘧啶 | -- | C | 否 |
| N4 (乙醯基)胞嘧啶 | -- | C | 否 |
| l-甲基-1-去氮-假異胞嘧啶核苷 | -- | C | 否 |
| 1-甲基-假異胞嘧啶核苷 | -- | C | 否 |
| 2-甲氧基-5-甲基-胞嘧啶核苷 | -- | C | 否 |
| 2-甲氧基-胞嘧啶核苷 | -- | C | 否 |
| 2-硫基-5-甲基-胞嘧啶核苷 | -- | C | 否 |
| 4-甲氧基-1-甲基-假異胞嘧啶核苷 | -- | C | 否 |
| 4-甲氧基-假異胞嘧啶核苷 | -- | C | 否 |
| 4-硫基-l-甲基-1-去氮-<br>假異胞嘧啶核苷 | -- | C | 否 |
| 4-硫基-1-甲基-假異胞嘧啶核苷 | -- | C | 否 |
| 4-硫基-假異胞嘧啶核苷 | -- | C | 否 |
| 5-氮雜-澤布拉林(zebularine) | -- | C | 否 |
| 5-甲基-澤布拉林 | -- | C | 否 |
| 吡咯并-假異胞嘧啶核苷 | -- | C | 否 |

| | | | |
|---|---|---|---|
| 澤布拉林 | -- | C | 否 |
| (E)-5-(2-溴-乙烯基)胞嘧啶核苷TP | -- | C | 否 |
| 2,2'-脫水-胞嘧啶核苷TP鹽酸鹽 | -- | C | 否 |
| 2'氟-N4-Bz-胞嘧啶核苷TP | -- | C | 否 |
| 2'氟-N4-乙醯基-胞嘧啶核苷TP | -- | C | 否 |
| 2'-O-甲基-N4-乙醯基-胞嘧啶核苷TP | -- | C | 否 |
| 2'O-甲基-N4-Bz-胞嘧啶核苷TP | -- | C | 否 |
| 2'-a-乙炔基胞嘧啶核苷TP | -- | C | 否 |
| 2'-a-三氟甲基胞嘧啶核苷TP | -- | C | 否 |
| 2'-b-乙炔基胞嘧啶核苷TP | -- | C | 否 |
| 2'-b-三氟甲基胞嘧啶核苷TP | -- | C | 否 |
| 2'-去氧-2',2'-二氟胞嘧啶核苷TP | -- | C | 否 |
| 2'-去氧-2'-a-巰基胞嘧啶核苷TP | -- | C | 否 |
| 2'-去氧-2'-a-硫代甲氧基胞嘧啶核苷TP | -- | C | 否 |
| 2'-去氧-2'-b-胺基胞嘧啶核苷TP | -- | C | 否 |
| 2'-去氧-2'-b-疊氮基胞嘧啶核苷TP | -- | C | 否 |
| 2'-去氧-2'-b-溴胞嘧啶核苷TP | -- | C | 否 |
| 2'-去氧-2'-b-氯胞嘧啶核苷TP | -- | C | 否 |
| 2'-去氧-2'-b-氟胞嘧啶核苷TP | -- | C | 否 |
| 2'-去氧-2'-b-碘胞嘧啶核苷TP | -- | C | 否 |
| 2'-去氧-2'-b-巰基胞嘧啶核苷TP | -- | C | 否 |
| 2'-去氧-2'-b-硫代甲氧基胞嘧啶核苷TP | -- | C | 否 |
| 2'-O-甲基-5-(1-丙炔基)胞嘧啶核苷TP | -- | C | 否 |
| 3'-乙炔基胞嘧啶核苷TP | -- | C | 否 |
| 4'-疊氮基胞嘧啶核苷TP | -- | C | 否 |
| 4'-碳環胞嘧啶核苷TP | -- | C | 否 |
| 4'-乙炔基胞嘧啶核苷TP | -- | C | 否 |
| 5-(1-丙炔基)阿糖-胞嘧啶核苷TP | -- | C | 否 |
| 5-(2-氯-苯基)-2-硫代胞嘧啶核苷TP | -- | C | 否 |
| 5-(4-胺基-苯基)-2-硫代胞嘧啶核苷TP | -- | C | 否 |
| 5-胺基烯丙基-CTP | -- | C | 否 |
| 5-氰基胞嘧啶核苷TP | -- | C | 否 |
| 5-乙炔基阿糖-胞嘧啶核苷TP | -- | C | 否 |
| 5-乙炔基胞嘧啶核苷TP | -- | C | 否 |
| 5'-均-胞嘧啶核苷TP | -- | C | 否 |
| 5-甲氧基胞嘧啶核苷TP | -- | C | 否 |
| 5-三氟甲基-胞嘧啶核苷TP | -- | C | 否 |
| N4-胺基-胞嘧啶核苷TP | -- | C | 否 |
| N4-苯甲醯基-胞嘧啶核苷TP | -- | C | 否 |
| 假異胞嘧啶核苷 | -- | C | 否 |
| 6-硫基-鳥苷 | -- | G | 否 |
| 7-去氮-鳥苷 | -- | G | 否 |

| | | | |
|---|---|---|---|
| 8-側氧基-鳥苷 | -- | G | 否 |
| Nl-甲基-鳥苷 | -- | G | 否 |
| a-硫基-鳥苷 | -- | G | 否 |
| 2 (丙基)鳥嘌呤 | -- | G | 否 |
| 2-(烷基)鳥嘌呤 | -- | G | 否 |
| 2'-胺基-2'-去氧-GTP | -- | G | 否 |
| 2'-疊氮基-2'-去氧-GTP | -- | G | 否 |
| 2'-去氧-2'-a-胺基鳥苷TP | -- | G | 否 |
| 2'-去氧-2'-a-疊氮基鳥苷TP | -- | G | 否 |
| 6 (甲基)鳥嘌呤 | -- | G | 否 |
| 6-(烷基)鳥嘌呤 | -- | G | 否 |
| 6-(甲基)鳥嘌呤 | -- | G | 否 |
| 6-甲基-鳥苷 | -- | G | 否 |
| 7 (烷基)鳥嘌呤 | -- | G | 否 |
| 7 (去氮)鳥嘌呤 | -- | G | 否 |
| 7 (甲基)鳥嘌呤 | -- | G | 否 |
| 7-(烷基)鳥嘌呤 | -- | G | 否 |
| 7-(去氮)鳥嘌呤 | -- | G | 否 |
| 7-(甲基)鳥嘌呤 | -- | G | 否 |
| 8 (烷基)鳥嘌呤 | -- | G | 否 |
| 8 (炔基)鳥嘌呤 | -- | G | 否 |
| 8 (鹵基)鳥嘌呤 | -- | G | 否 |
| 8 (硫代烷基)鳥嘌呤 | -- | G | 否 |
| 8-(烯基)鳥嘌呤 | -- | G | 否 |
| 8-(烷基)鳥嘌呤 | -- | G | 否 |
| 8-(炔基)鳥嘌呤 | -- | G | 否 |
| 8-(胺基)鳥嘌呤 | -- | G | 否 |
| 8-(鹵基)鳥嘌呤 | -- | G | 否 |
| 8-(羥基)鳥嘌呤 | -- | G | 否 |
| 8-(硫代烷基)鳥嘌呤 | -- | G | 否 |
| 8-(硫醇)鳥嘌呤 | -- | G | 否 |
| 氮雜鳥嘌呤 | -- | G | 否 |
| 去氮鳥嘌呤 | -- | G | 否 |
| N (甲基)鳥嘌呤 | -- | G | 否 |
| N-(甲基)鳥嘌呤 | -- | G | 否 |
| l-甲基-6-硫基-鳥苷 | -- | G | 否 |
| 6-甲氧基-鳥苷 | -- | G | 否 |
| 6-硫基-7-去氮-8-氮雜-鳥苷 | -- | G | 否 |
| 6-硫基-7-去氮-鳥苷 | -- | G | 否 |
| 6-硫基-7-甲基-鳥苷 | -- | G | 否 |
| 7-去氮-8-氮雜-鳥苷 | -- | G | 否 |
| 7-甲基-8-側氧基-鳥苷 | -- | G | 否 |

| | | | |
|---|---|---|---|
| N2,N2-二甲基-6-硫基-鳥苷 | -- | G | 否 |
| N2-甲基-6-硫基-鳥苷 | -- | G | 否 |
| 1-Me-GTP | -- | G | 否 |
| 2'氟-N2-異丁基-鳥苷TP | -- | G | 否 |
| 2'O-甲基-N2-異丁基-鳥苷TP | -- | G | 否 |
| 2'-a-乙炔基鳥苷TP | -- | G | 否 |
| 2'-a-三氟甲基鳥苷TP | -- | G | 否 |
| 2'-b-乙炔基鳥苷TP | -- | G | 否 |
| 2'-b-三氟甲基鳥苷TP | -- | G | 否 |
| 2'-去氧-2',2'-二氟鳥苷TP | -- | G | 否 |
| 2'-去氧-2'-a-巰基鳥苷TP | -- | G | 否 |
| 2'-去氧-2'-a-硫代甲氧基鳥苷TP | -- | G | 否 |
| 2'-去氧-2'-b-胺基鳥苷TP | -- | G | 否 |
| 2'-去氧-2'-b-疊氮基鳥苷TP | -- | G | 否 |
| 2'-去氧-2'-b-溴鳥苷TP | -- | G | 否 |
| 2'-去氧-2'-b-氯鳥苷TP | -- | G | 否 |
| 2'-去氧-2'-b-氟鳥苷TP | -- | G | 否 |
| 2'-去氧-2'-b-碘鳥苷TP | -- | G | 否 |
| 2'-去氧-2'-b-巰基鳥苷TP | -- | G | 否 |
| 2'-去氧-2'-b-硫代甲氧基鳥苷TP | -- | G | 否 |
| 4'-疊氮基鳥苷TP | -- | G | 否 |
| 4'-碳環鳥苷TP | -- | G | 否 |
| 4'-乙炔基鳥苷TP | -- | G | 否 |
| 5'-均-鳥苷TP | -- | G | 否 |
| 8-溴-鳥苷TP | -- | G | 否 |
| 9-去氮鳥苷TP | -- | G | 否 |
| N2-異丁基-鳥苷TP | -- | G | 否 |
| 7-甲基肌苷 | | A | 否 |
| 烯丙基胺基-胸苷 | -- | T | 否 |
| 氮雜胸苷 | -- | T | 否 |
| 去氮胸苷 | -- | T | 否 |
| 去氧-胸苷 | -- | T | 否 |
| 5-丙炔基尿嘧啶 | -- | U | 否 |
| a-硫基-尿苷 | -- | U | 否 |
| 1 (胺基烷基胺基-羰基伸乙基)-2(硫基)-假尿嘧啶 | -- | U | 否 |
| 1 (胺基烷基胺基羰基伸乙基)-2,4-(二硫基)假尿嘧啶 | -- | U | 否 |
| 1 (胺基烷基胺基羰基伸乙基)-4 (硫基)假尿嘧啶 | -- | U | 否 |
| 1 (胺基烷基胺基羰基伸乙基)-假尿嘧啶 | -- | U | 否 |
| 1 (胺基羰基伸乙基)-2(硫基)-<br>假尿嘧啶 | -- | U | 否 |
| 1 (胺基羰基伸乙基)-2,4-<br>(二硫基)假尿嘧啶 | -- | U | 否 |
| 1 (胺基羰基伸乙基)-4 (硫基)假尿嘧啶 | -- | U | 否 |

| | | | |
|---|---|---|---|
| 1 (胺基羰基伸乙基)-假尿嘧啶 | -- | U | 否 |
| 1 取代之2(硫基)-假尿嘧啶 | -- | U | 否 |
| 1 取代之2,4-(二硫基)假尿嘧啶 | -- | U | 否 |
| 1 取代之4 (硫基)假尿嘧啶 | -- | U | 否 |
| 1 取代之假尿嘧啶 | -- | U | 否 |
| 1-(胺基烷基胺基-羰基伸乙基)-2-(硫基)-假尿嘧啶 | -- | U | 否 |
| l-甲基-3-(3-胺基-3-羧基丙基)假尿苷TP | -- | U | 否 |
| l-甲基-3-(3-胺基-3-羧基普魯維(羧基丙基))假-UTP | -- | U | 否 |
| 1-甲基-假-UTP | -- | U | 否 |
| 2 (硫基)假尿嘧啶 | -- | U | 否 |
| 2' 去氧尿苷 | -- | U | 否 |
| 2'氟尿苷 | -- | U | 否 |
| 2-(硫基)尿嘧啶 | -- | U | 否 |
| 2,4-(二硫基)假尿嘧啶 | -- | U | 否 |
| 2' 甲基、2'胺基、2'疊氮基、2'氟-<br>鳥苷 | -- | U | 否 |
| 2'-胺基-2'-去氧-UTP | -- | U | 否 |
| 2'-疊氮基-2'-去氧-UTP | -- | U | 否 |
| 2'-疊氮基-去氧尿苷TP | -- | U | 否 |
| 2'-O-甲基假尿苷 | -- | U | 否 |
| 2' 去氧尿苷 | 2'dU | U | 否 |
| 2'氟尿苷 | -- | U | 否 |
| 2'-去氧-2'-a-胺基尿苷TP | -- | U | 否 |
| 2'-去氧-2'-a-疊氮基尿苷TP | -- | U | 否 |
| 2-甲基假尿苷 | m3'P | U | 否 |
| 3 (3胺基-3 羧基丙基)尿嘧啶 | -- | U | 否 |
| 4 (硫基)假尿嘧啶 | -- | U | 否 |
| 4-(硫基)假尿嘧啶 | -- | U | 否 |
| 4-(硫基)尿嘧啶 | -- | U | 否 |
| 4-硫代尿嘧啶 | -- | U | 否 |
| 5 (l ,3-二唑-1-烷基)尿嘧啶 | -- | U | 否 |
| 5 (2-胺基丙基)尿嘧啶 | -- | U | 否 |
| 5 (胺基烷基)尿嘧啶 | -- | U | 否 |
| 5 (二甲基胺基烷基)尿嘧啶 | -- | U | 否 |
| 5 (胍烷基)尿嘧啶 | -- | U | 否 |
| 5 (甲氧基羰基甲基)-2-(硫基)尿嘧啶 | -- | U | 否 |
| 5 (甲氧基羰基-甲基)尿嘧啶 | -- | U | 否 |
| 5 (甲基) 2 (硫基)尿嘧啶 | -- | U | 否 |
| 5 (甲基) 2,4 (二硫基)尿嘧啶 | -- | U | 否 |
| 5 (甲基) 4 (硫基)尿嘧啶 | -- | U | 否 |
| 5 (甲基胺基甲基)-2 (硫基)尿嘧啶 | -- | U | 否 |
| 5 (甲基胺基甲基)-2,4 (二硫基)尿嘧啶 | -- | U | 否 |

| | | | |
|---|---|---|---|
| 5 (甲基胺基甲基)-4 (硫基)尿嘧啶 | -- | U | 否 |
| 5 (丙炔基)尿嘧啶 | -- | U | 否 |
| 5 (三氟甲基)尿嘧啶 | -- | U | 否 |
| 5-(2-胺基丙基)尿嘧啶 | -- | U | 否 |
| 5-(烷基)-2-(硫基)假尿嘧啶 | -- | U | 否 |
| 5-(烷基)-2,4 (二硫基)假尿嘧啶 | -- | U | 否 |
| 5-(烷基)-4 (硫基)假尿嘧啶 | -- | U | 否 |
| 5-(烷基)假尿嘧啶 | -- | U | 否 |
| 5-(烷基)尿嘧啶 | -- | U | 否 |
| 5-(炔基)尿嘧啶 | -- | U | 否 |
| 5-(烯丙基胺基)尿嘧啶 | -- | U | 否 |
| 5-(氰基烷基)尿嘧啶 | -- | U | 否 |
| 5-(二烷基胺基烷基)尿嘧啶 | -- | U | 否 |
| 5-(二甲基胺基烷基)尿嘧啶 | -- | U | 否 |
| 5-(胍烷基)尿嘧啶 | -- | U | 否 |
| 5-(鹵基)尿嘧啶 | -- | U | 否 |
| 5-(1,3-二唑-l-烷基)尿嘧啶 | -- | U | 否 |
| 5-(甲氧基)尿嘧啶 | -- | U | 否 |
| 5-(甲氧基羰基甲基)-2-(硫基)尿嘧啶 | -- | U | 否 |
| 5-(甲氧基羰基-甲基)尿嘧啶 | -- | U | 否 |
| 5-(甲基) 2(硫基)尿嘧啶 | -- | U | 否 |
| 5-(甲基) 2,4 (二硫基)尿嘧啶 | -- | U | 否 |
| 5-(甲基) 4 (硫基)尿嘧啶 | -- | U | 否 |
| 5-(甲基)-2-(硫基)假尿嘧啶 | -- | U | 否 |
| 5-(甲基)-2,4 (二硫基)假尿嘧啶 | -- | U | 否 |
| 5-(甲基)-4 (硫基)假尿嘧啶 | -- | U | 否 |
| 5-(甲基)假尿嘧啶 | -- | U | 否 |
| 5-(甲基胺基甲基)-2 (硫基)尿嘧啶 | -- | U | 否 |
| 5-(甲基胺基甲基)-2,4(二硫基)尿嘧啶 | -- | U | 否 |
| 5-(甲基胺基甲基)-4-(硫基)尿嘧啶 | -- | U | 否 |
| 5-(丙炔基)尿嘧啶 | -- | U | 否 |
| 5-(三氟甲基)尿嘧啶 | -- | U | 否 |
| 5-胺基烯丙基-尿苷 | -- | U | 否 |
| 5-溴-尿苷 | -- | U | 否 |
| 5-碘-尿苷 | -- | U | 否 |
| 5-尿嘧啶 | -- | U | 否 |
| 6 (偶氮基)尿嘧啶 | -- | U | 否 |
| 6-(偶氮基)尿嘧啶 | -- | U | 否 |
| 6-氮雜-尿苷 | -- | U | 否 |
| 烯丙基胺基-尿嘧啶 | -- | U | 否 |
| 氮雜尿嘧啶 | -- | U | 否 |
| 去氮尿嘧啶 | -- | U | 否 |

| | | | |
|---|---|---|---|
| N3 (甲基)尿嘧啶 | -- | U | 否 |
| 假-UTP-1-2-乙酸 | -- | U | 否 |
| 假尿嘧啶 | -- | U | 否 |
| 4-硫基-假-UTP | -- | U | 否 |
| 1-羧基甲基-假尿苷 | -- | U | 否 |
| l-甲基-1-去氮-假尿苷 | -- | U | 否 |
| 1-丙炔基-尿苷 | -- | U | 否 |
| l-牛磺酸甲基-1-甲基-尿苷 | -- | U | 否 |
| l-牛磺酸甲基-4-硫基-尿苷 | -- | U | 否 |
| 1-牛磺酸甲基-假尿苷 | -- | U | 否 |
| 2-甲氧基-4-硫基-假尿苷 | -- | U | 否 |
| 2-硫基-l-甲基-1-去氮-假尿苷 | -- | U | 否 |
| 2-硫基-1-甲基-假尿苷 | -- | U | 否 |
| 2-硫基-5-氮雜-尿苷 | -- | U | 否 |
| 2-硫基-二氫假尿苷 | -- | U | 否 |
| 2-硫基-二氫尿苷 | -- | U | 否 |
| 2-硫基-假尿苷 | -- | U | 否 |
| 4-甲氧基-2-硫基-假尿苷 | -- | U | 否 |
| 4-甲氧基-假尿苷 | -- | U | 否 |
| 4-硫基-1-甲基-假尿苷 | -- | U | 否 |
| 4-硫基-假尿苷 | -- | U | 否 |
| 5-氮雜-尿苷 | -- | U | 否 |
| 二氫假尿苷 | -- | U | 否 |
| (±)1-(2-羥基丙基)假尿苷TP | -- | U | 否 |
| (2R)-l-(2-羥基丙基)假尿苷TP | -- | U | 否 |
| (2S)-l-(2-羥基丙基)假尿苷<br>TP | -- | U | 否 |
| (E)-5-(2-溴-乙烯基)阿糖-尿苷TP | -- | U | 否 |
| (E)-5-(2-溴-乙烯基)尿苷TP | -- | U | 否 |
| (Z)-5-(2-溴-乙烯基)阿糖-尿苷TP | -- | U | 否 |
| (Z)-5-(2-溴-乙烯基)尿苷TP | -- | U | 否 |
| 1-(2,2,2-三氟乙基)-假-UTP | -- | U | 否 |
| 1-(2,2,3,3,3-<br>五氟丙基)假尿苷TP | -- | U | 否 |
| 1-(2,2-二乙氧基乙基)假尿苷TP | -- | U | 否 |
| 1-(2,4,6-三甲基苯甲基)假尿苷TP | -- | U | 否 |
| 1-(2,4,6-三甲基-苯甲基)假-UTP | -- | U | 否 |
| 1-(2,4,6-三甲基-苯基)假-UTP | -- | U | 否 |
| 1-(2-胺基-2-羧基乙基)假-UTP | -- | U | 否 |
| 1-(2-胺基-乙基)假-UTP | -- | U | 否 |
| 1-(2-羥基乙基)假尿苷TP | -- | U | 否 |
| 1-(2-甲氧基乙基)假尿苷TP | -- | U | 否 |
| 1-(3,4-雙-三氟甲氧基苯甲基)假尿苷TP | -- | U | 否 |

| | | | |
|---|---|---|---|
| 1-(3,4-二甲氧基苯甲基)假尿苷TP | -- | U | 否 |
| 1-(3-胺基-3-羧基丙基)假-UTP | -- | U | 否 |
| 1-(3-胺基-丙基)假-UTP | -- | U | 否 |
| 1-(3-環丙基-丙-2-炔基)假尿苷TP | -- | U | 否 |
| 1-(4-胺基-4-羧基丁基)假-UTP | -- | U | 否 |
| 1-(4-胺基-苯甲基)假-UTP | -- | U | 否 |
| 1-(4-胺基-丁基)假-UTP | -- | U | 否 |
| 1-(4-胺基-苯基)假-UTP | -- | U | 否 |
| 1-(4-疊氮基苯甲基)假尿苷TP | -- | U | 否 |
| 1-(4-溴苯甲基)假尿苷TP | -- | U | 否 |
| 1-(4-氯苯甲基)假尿苷TP | -- | U | 否 |
| 1-(4-氟苯甲基)假尿苷TP | -- | U | 否 |
| 1-(4-碘苯甲基)假尿苷TP | -- | U | 否 |
| 1-(4-甲磺醯基苯甲基)假尿苷TP | -- | U | 否 |
| 1-(4-甲氧基苯甲基)假尿苷TP | -- | U | 否 |
| 1-(4-甲氧基-苯甲基)假-UTP | -- | U | 否 |
| 1-(4-甲氧基-苯基)假-UTP | -- | U | 否 |
| 1-(4-甲基苯甲基)假尿苷TP | -- | U | 否 |
| 1-(4-甲基-苯甲基)假-UTP | -- | U | 否 |
| 1-(4-硝基苯甲基)假尿苷TP | -- | U | 否 |
| 1-(4-硝基-苯甲基)假-UTP | -- | U | 否 |
| 1(4-硝基-苯基)假-UTP | -- | U | 否 |
| 1-(4-硫代甲氧基苯甲基)假尿苷TP | -- | U | 否 |
| 1-(4-<br>三氟甲氧基苯甲基)假尿苷TP | -- | U | 否 |
| 1-(4-三氟甲基苯甲基)假尿苷TP | -- | U | 否 |
| 1-(5-胺基-戊基)假-UTP | -- | U | 否 |
| 1-(6-胺基-己基)假-UTP | -- | U | 否 |
| 1,6-二甲基-假-UTP | -- | U | 否 |
| l-[3-(2-{2-[2-(2-胺基乙氧基)-乙氧基]-乙氧基}-乙氧基)-丙醯基]假尿苷TP | -- | U | 否 |
| 1-{3-[2-(2-胺基乙氧基)-乙氧基]-丙醯基}假尿苷TP | -- | U | 否 |
| 1-乙醯基假尿苷TP | -- | U | 否 |
| l-烷基-6-(1-丙炔基)-假-UTP | -- | U | 否 |
| l-烷基-6-(2-丙炔基)-假-UTP | -- | U | 否 |
| l-烷基-6-烯丙基-假-UTP | -- | U | 否 |
| l-烷基-6-乙炔基-假-UTP | -- | U | 否 |
| l-烷基-6-均烯丙基-假-UTP | -- | U | 否 |
| l-烷基-6-乙烯基-假-UTP | -- | U | 否 |
| 1-烯丙基假尿苷TP | -- | U | 否 |
| 1-胺基甲基-假-UTP | -- | U | 否 |
| 1-苯甲醯基假尿苷TP | -- | U | 否 |
| 1-苯甲基氧基甲基假尿苷TP | -- | U | 否 |

| | | | |
|---|---|---|---|
| 1-苯甲基-假-UTP | -- | U | 否 |
| l-生物素基-PEG2-假尿苷TP | -- | U | 否 |
| 1-生物素基假尿苷TP | -- | U | 否 |
| 1-丁基-假-UTP | -- | U | 否 |
| 1-氰基甲基假尿苷TP | -- | U | 否 |
| 1-環丁基甲基-假-UTP | -- | U | 否 |
| 1-環丁基-假-UTP | -- | U | 否 |
| 1-環庚基甲基-假-UTP | -- | U | 否 |
| 1-環庚基-假-UTP | -- | U | 否 |
| 1-環己基甲基-假-UTP | -- | U | 否 |
| 1-環己基-假-UTP | -- | U | 否 |
| 1-環辛基甲基-假-UTP | -- | U | 否 |
| 1-環辛基-假-UTP | -- | U | 否 |
| 1-環戊基甲基-假-UTP | -- | U | 否 |
| 1-環戊基-假-UTP | -- | U | 否 |
| 1-環丙基甲基-假-UTP | -- | U | 否 |
| 1-環丙基-假-UTP | -- | U | 否 |
| 1-乙基-假-UTP | -- | U | 否 |
| 1-己基-假-UTP | -- | U | 否 |
| 1-均烯丙基假尿苷TP | -- | U | 否 |
| 1-羥基甲基假尿苷TP | -- | U | 否 |
| 1-異-丙基-假-UTP | -- | U | 否 |
| 1-Me-2-硫基-假-UTP | -- | U | 否 |
| 1-Me-4-硫基-假-UTP | -- | U | 否 |
| 1-Me-α-硫基-假-UTP | -- | U | 否 |
| 1-甲磺醯基甲基假尿苷TP | -- | U | 否 |
| 1-甲氧基甲基假尿苷TP | -- | U | 否 |
| l-甲基-6-(2,2,2-三氟乙基)假-UTP | -- | U | 否 |
| l-甲基-6-(4-N-味啉基)-假-DTP | -- | U | 否 |
| l-甲基-6-(4-硫代N-味啉基)-假-UTP | -- | U | 否 |
| l-甲基-6-(取代之苯基)假-UTP | -- | U | 否 |
| 1-甲基-6-胺基-假-UTP | -- | U | 否 |
| l-甲基-6-疊氮基-假-UTP | -- | U | 否 |
| 1-甲基-6-溴-假-UTP | -- | U | 否 |
| l-甲基-6-丁基-假-UTP | -- | U | 否 |
| l-甲基-6-氯-假-UTP | -- | U | 否 |
| 1-甲基-6-氰基-假-UTP | -- | U | 否 |
| l-甲基-6-二甲基胺基-假-UTP | -- | U | 否 |
| l-甲基-6-乙氧基-假-UTP | -- | U | 否 |
| l-甲基-6-乙基甲酸酯-假-UTP | -- | U | 否 |
| l-甲基-6-乙基-假-UTP | -- | U | 否 |
| l-甲基-6-氟-假-UTP | -- | U | 否 |

| | | | |
|---|---|---|---|
| 1-甲基-6-甲醯基-假-UTP | -- | U | 否 |
| 1-甲基-6-羥基胺基-假-UTP | -- | U | 否 |
| 1-甲基-6-羥基-假-UTP | -- | U | 否 |
| 1-甲基-6-碘-假-UTP | -- | U | 否 |
| 1-甲基-6-異-丙基-假-UTP | -- | U | 否 |
| 1-甲基-6-甲氧基-假-UTP | -- | U | 否 |
| 1-甲基-6-甲基胺基-假-UTP | -- | U | 否 |
| 1-甲基-6-苯基-假-UTP | -- | U | 否 |
| 1-甲基-6-丙基-假-UTP | -- | U | 否 |
| 1-甲基-6-三級丁基-假-UTP | -- | U | 否 |
| 1-甲基-6-三氟甲氧基-假-UTP | -- | U | 否 |
| 1-甲基-6-三氟甲基-假-UTP | -- | U | 否 |
| 1-N-味啉基甲基假尿苷TP | -- | U | 否 |
| 1-戊基-假-UTP | -- | U | 否 |
| 1-苯基-假-UTP | -- | U | 否 |
| 1-特戊醯基假尿苷TP | -- | U | 否 |
| 1-炔丙基假尿苷TP | -- | U | 否 |
| 1-丙基-假-UTP | -- | U | 否 |
| 1-丙炔基-假尿苷 | -- | U | 否 |
| 1-對甲苯基-假-UTP | -- | U | 否 |
| 1-三級丁基-假-UTP | -- | U | 否 |
| 1-硫代甲氧基甲基假尿苷TP | -- | U | 否 |
| 1-硫代N-味啉基甲基假尿苷TP | -- | U | 否 |
| 1-三氟乙醯基假尿苷TP | -- | U | 否 |
| 1-三氟甲基-假-UTP | -- | U | 否 |
| 1-乙烯基假尿苷TP | -- | U | 否 |
| 2,2'-脫水-尿苷TP | -- | U | 否 |
| 2'-溴-去氧尿苷TP | -- | U | 否 |
| 2'-F-5-甲基-2'-去氧-UTP | -- | U | 否 |
| 2'-OMe-5-Me-UTP | -- | U | 否 |
| 2'-OMe-假-UTP | -- | U | 否 |
| 2'-a-乙炔基尿苷TP | -- | U | 否 |
| 2'-a-三氟甲基尿苷TP | -- | U | 否 |
| 2'-b-乙炔基尿苷TP | -- | U | 否 |
| 2'-b-三氟甲基尿苷TP | -- | U | 否 |
| 2'-去氧-2',2'-二氟尿苷TP | -- | U | 否 |
| 2'-去氧-2'-a-巰基尿苷TP | -- | U | 否 |
| 2'-去氧-2'-a-硫代甲氧基尿苷TP | -- | U | 否 |
| 2'-去氧-2'-b-胺基尿苷TP | -- | U | 否 |
| 2'-去氧-2'-b-疊氮基尿苷TP | -- | U | 否 |
| 2'-去氧-2'-b-溴尿苷TP | -- | U | 否 |
| 2'-去氧-2'-b-氯尿苷TP | -- | U | 否 |

| | | | |
|---|---|---|---|
| 2'-去氧-2'-b-氟尿苷TP | -- | U | 否 |
| 2'-去氧-2'-b-碘尿苷TP | -- | U | 否 |
| 2'-去氧-2'-b-甾基尿苷TP | -- | U | 否 |
| 2'-去氧-2'-b-硫代甲氧基尿苷TP | -- | U | 否 |
| 2-甲氧基-4-硫基-尿苷 | -- | U | 否 |
| 2-甲氧基尿苷 | -- | U | 否 |
| 2'-O-甲基-5-(1-丙炔基)尿苷TP | -- | U | 否 |
| 3-烷基-假-UTP | -- | U | 否 |
| 4'-疊氮基尿苷TP | -- | U | 否 |
| 4'-碳環尿苷TP | -- | U | 否 |
| 4'-乙炔基尿苷TP | -- | U | 否 |
| 5-(1-丙炔基)阿糖-尿苷TP | -- | U | 否 |
| 5-(2-呋喃基)尿苷TP | -- | U | 否 |
| 5-氰基尿苷TP | -- | U | 否 |
| 5-二甲基胺基尿苷TP | -- | U | 否 |
| 5'-均-尿苷TP | -- | U | 否 |
| 5-碘-2'-氟-去氧尿苷TP | -- | U | 否 |
| 5-苯基乙炔基尿苷TP | -- | U | 否 |
| 5-三氘甲基-6-氘尿苷TP | -- | U | 否 |
| 5-三氟甲基-尿苷TP | -- | U | 否 |
| 5-乙烯基阿糖尿苷TP | -- | U | 否 |
| 6-(2,2,2-三氟乙基)-假-UTP | -- | U | 否 |
| 6-(4-N-味啉基)-假-DTP | -- | U | 否 |
| 6-(4-硫代N-味啉基)-假-UTP | -- | U | 否 |
| 6-(取代之-苯基)-假-UTP | -- | U | 否 |
| 6-胺基-假-UTP | -- | U | 否 |
| 6-疊氮基-假-UTP | -- | U | 否 |
| 6-溴-假-UTP | -- | U | 否 |
| 6-丁基-假-UTP | -- | U | 否 |
| 6-氯-假-UTP | -- | U | 否 |
| 6-氰基-假-UTP | -- | U | 否 |
| 6-二甲基胺基-假-UTP | -- | U | 否 |
| 6-乙氧基-假-UTP | -- | U | 否 |
| 6-乙基甲酸酯-假-UTP | -- | U | 否 |
| 6-乙基-假-UTP | -- | U | 否 |
| 6-氟-假-UTP | -- | U | 否 |
| 6-甲醯基-假-UTP | -- | U | 否 |
| 6-羥基胺基-假-UTP | -- | U | 否 |
| 6-羥基-假-UTP | -- | U | 否 |
| 6-碘-假-UTP | -- | U | 否 |
| 6-異-丙基-假-UTP | -- | U | 否 |
| 6-甲氧基-假-UTP | -- | U | 否 |

| | | | |
|---|---|---|---|
| 6-甲基胺基-假-UTP | -- | U | 否 |
| 6-甲基-假-UTP | -- | U | 否 |
| 6-苯基-假-UTP | -- | U | 否 |
| 6-苯基-假-UTP | -- | U | 否 |
| 6-丙基-假-UTP | -- | U | 否 |
| 6-三級丁基-假-UTP | -- | U | 否 |
| 6-三氟甲氧基-假-UTP | -- | U | 否 |
| 6-三氟甲基-假-UTP | -- | U | 否 |
| α-硫基-假-UTP | -- | U | 否 |
| 假尿苷1-(4-甲基苯磺酸) TP | -- | U | 否 |
| 假尿苷1-(4-甲基苯甲酸) TP | -- | U | 否 |
| 假尿苷TP l-[3-(2-乙氧基)]丙酸 | -- | U | 否 |
| 假尿苷TP l-[3-{2-(2-[2-(2-乙氧基)-乙氧基]-乙氧基)-乙氧基}]丙酸 | -- | U | 否 |
| 假尿苷TP l-[3-{2-(2-[2-{2(2-乙氧基)-乙氧基}-乙氧基]-乙氧基)-乙氧基}]丙酸 | -- | U | 否 |
| 假尿苷TP l-[3-{2-(2-[2-乙氧基]-乙氧基)-乙氧基}]丙酸 | -- | U | 否 |
| 假尿苷TP l-[3-{2-(2-乙氧基)-乙氧基}]丙酸 | -- | U | 否 |
| 假尿苷TP 1-甲基膦酸 | -- | U | 否 |
| 假尿苷TP 1-甲基膦酸二乙基酯 | -- | U | 否 |
| 假-UT對N1-3-丙酸 | -- | U | 否 |
| 假-UT對N1-4-丁酸 | -- | U | 否 |
| 假-UT對N1-5-戊酸 | -- | U | 否 |
| 假-UT對N1-6-己酸 | -- | U | 否 |
| 假-UT對Nl-7-庚酸 | -- | U | 否 |
| 假-UT對N1-甲基-對苯甲酸 | -- | U | 否 |
| 假-UT對N1-對苯甲酸 | -- | U | 否 |

【0068】在一些實施例中，本文所述之遺傳元件包含表6中所提供之修飾或其組合。表6中所提供之修飾天然存在於RNA中，且可在本文中用於遺傳元件中在自然界中不存在之位置處。

**表6：額外例示性修飾**

| 名稱 | 符號 | 鹼基 | 天然存在 |
|---|---|---|---|
| 2-甲基硫代-N6-(順-羥基異戊烯基)腺苷 | ms2i6A | A | 是 |
| 2-甲基硫代-N6-甲基腺苷 | ms2m6A | A | 是 |
| 2-甲基硫代-N6-羥丁胺醯基胺甲醯基腺苷 | ms2t6A | A | 是 |
| N6-甘胺醯基胺甲醯基腺苷 | g6A | A | 是 |
| N6-異戊烯基腺苷 | i6A | A | 是 |

| | | | |
|---|---|---|---|
| N6-甲基腺苷 | m6A | A | 是 |
| N6-羥丁胺醯基胺甲醯基腺苷 | t6A | A | 是 |
| 1,2'-O-二甲基腺苷 | mlAm | A | 是 |
| 1-甲基腺苷 | mlA | A | 是 |
| 2'-O-甲基腺苷 | Am | A | 是 |
| 2'-O-核糖基腺苷(磷酸鹽) | Ar(p) | A | 是 |
| 2-甲基腺苷 | m2A | A | 是 |
| 2-甲基硫代-N6異戊烯基腺苷 | ms2i6A | A | 是 |
| 2-甲基硫代-N6-羥基去纈胺醯基胺甲醯基腺苷 | ms2hn6A | A | 是 |
| 2'-O-甲基腺苷 | m6A | A | 是 |
| 2'-O-核糖基腺苷(磷酸鹽) | Ar(p) | A | 是 |
| 異戊烯基腺苷 | Iga | A | 是 |
| N6-(順-羥基異戊烯基)腺苷 | io6A | A | 是 |
| N6,2'-O-二甲基腺苷 | m6Am | A | 是 |
| N6,2'-O-二甲基腺苷 | $m^6Am$ | A | 是 |
| N6,N6,2'-O-三甲基腺苷 | m62Am | A | 是 |
| N6,N6-二甲基腺苷 | m62A | A | 是 |
| N6-乙醯基腺苷 | ac6A | A | 是 |
| N6-羥基去纈胺醯基胺甲醯基腺苷 | hn6A | A | 是 |
| N6-甲基-N6-羥丁胺醯基胺甲醯基腺苷 | m6t6A | A | 是 |
| 2-甲基腺苷 | $m^2A$ | A | 是 |
| 2-甲基硫代-$N^6$-異戊烯基腺苷 | $ms^2i^6A$ | A | 是 |
| 2-硫代胞嘧啶核苷 | s2C | C | 是 |
| 3-甲基胞嘧啶核苷 | m3C | C | 是 |
| 5-甲醯基胞嘧啶核苷 | f5C | C | 是 |
| 5-羥基甲基胞嘧啶核苷 | hm5C | C | 是 |
| 5-甲基胞嘧啶核苷 | m5C | C | 是 |
| N4-乙醯基胞嘧啶核苷 | ac4C | C | 是 |
| 2'-O-甲基胞嘧啶核苷 | Cm | C | 是 |
| 2'-O-甲基胞嘧啶核苷 | Cm | C | 是 |
| 5,2'-O-二甲基胞嘧啶核苷 | m5Cm | C | 是 |
| 5-甲醯基-2'-O-甲基胞嘧啶核苷 | f5Cm | C | 是 |
| 立西啶(lysidine) | k2C | C | 是 |
| N4,2'-O-二甲基胞嘧啶核苷 | m4Cm | C | 是 |
| N4-乙醯基-2'-O-甲基胞嘧啶核苷 | ac4Cm | C | 是 |
| N4-甲基胞嘧啶核苷 | m4C | C | 是 |
| N4,N4-二甲基-2'-OMe-胞嘧啶核苷TP | -- | C | 是 |
| 7-甲基鳥苷 | m7G | G | 是 |
| N2,2'-O-二甲基鳥苷 | m2Gm | G | 是 |
| N2-甲基鳥苷 | m2G | G | 是 |
| 懷俄苷 | imG | G | 是 |
| 1,2'-O-二甲基鳥苷 | mlGm | G | 是 |

| | | | |
|---|---|---|---|
| 1-甲基鳥苷 | mlG | G | 是 |
| 2'-O-甲基鳥苷 | Gm | G | 是 |
| 2'-O-核糖基鳥苷(磷酸鹽) | Gr(p) | G | 是 |
| 2'-O-甲基鳥苷 | Gm | G | 是 |
| 2'-O-核糖基鳥苷(磷酸鹽) | Gr(p) | G | 是 |
| 7-胺基甲基-7-去氮鳥苷 | preQl | G | 是 |
| 7-氰基-7-去氮鳥苷 | preQ0 | G | 是 |
| 古嘌苷 | G+ | G | 是 |
| 甲基懷俄苷 | mimG | G | 是 |
| N2,7-二甲基鳥苷 | m2,7G | G | 是 |
| N2,N2,2'-O-三甲基鳥苷 | m22Gm | G | 是 |
| N2,N2,7-三甲基鳥苷 | m2,2,7G | G | 是 |
| N2,N2-二甲基鳥苷 | m22G | G | 是 |
| N2, 7,2 '-O-三甲基鳥苷 | m2 7<br>' Gm | G | 是 |
| 1-甲基肌苷 | mll | A | 是 |
| 肌苷 | I | A | 是 |
| 1,2'-O-二甲基肌苷 | mlim | A | 是 |
| 2'-O-甲基肌苷 | Im | A | 是 |
| 2'-O-甲基肌苷 | Im | A | 是 |
| 環氧基Q核苷 | oQ | G | 是 |
| 半乳糖苷-Q核苷 | galQ | G | 是 |
| 甘露糖基-Q核苷 | manQ | G | 是 |
| 2'-O-甲基尿苷 | -- | U | 是 |
| 2-硫代尿苷 | s2U | U | 是 |
| 3-甲基尿苷 | m3U | U | 是 |
| 5-羧基甲基尿苷 | cm5U | U | 是 |
| 5-羥基尿苷 | ho5U | U | 是 |
| 5-甲基尿苷 | m5U | U | 是 |
| 5-牛磺酸甲基-2-硫代尿苷 | rm5s2U | U | 是 |
| 5-牛磺酸甲基尿苷 | rm5U | U | 是 |
| 二氫尿苷 | D | U | 是 |
| 假尿苷 | Q | U | 是 |
| (3-(3-胺基-3-羧基丙基)尿苷 | acp3U | U | 是 |
| l-甲基-3-(3-胺基-5-<br>羧基丙基)假尿苷 | mlacp3'P | U | 是 |
| 1-甲基假尿苷 | ml'P | U | 是 |
| 1-甲基-假尿苷 | -- | U | 是 |
| 2'-0-甲基尿苷 | Um | U | 是 |
| 2'-0-甲基假尿苷 | 'Pm | U | 是 |
| 2'-0-甲基尿苷 | Um | U | 是 |
| 2-硫基-2'-0-甲基尿苷 | s2Um | U | 是 |
| 3-(3-胺基-3-羧基丙基)尿苷 | acp3U | U | 是 |

| | | | |
|---|---|---|---|
| 3,2'-0-二甲基尿苷 | m3Um | U | 是 |
| 3-甲基-假-尿苷TP | -- | U | 是 |
| 4-硫代尿苷 | s4U | U | 是 |
| 5-(羧基羥基甲基)尿苷 | chm5U | U | 是 |
| 5-(羧基羥基甲基)尿苷甲基酯 | mchm5U | U | 是 |
| 5,2'-0-二甲基尿苷 | m5Um | U | 是 |
| 5,6-二氫-尿苷 | -- | U | 是 |
| 5-胺基甲基-2-硫代尿苷 | nm5s2U | U | 是 |
| 5-胺甲醯基甲基-2'-0-甲基尿苷 | ncm5Um | U | 是 |
| 5-胺甲醯基甲基尿苷 | ncm5U | U | 是 |
| 5-羧基羥基甲基尿苷 | -- | U | 是 |
| 5-羧基羥基甲基尿苷甲基酯 | -- | U | 是 |
| 5-羧基甲基胺基甲基-2'-0-甲基尿苷 | cmnm5Um | U | 是 |
| 5-羧基甲基胺基甲基-2-硫代尿苷 | cmnm5s2U | U | 是 |
| 5-羧基甲基胺基甲基-2-硫代尿苷 | -- | U | 是 |
| 5-羧基甲基胺基甲基尿苷 | cmnm5U | U | 是 |
| 5-羧基甲基胺基甲基尿苷 | -- | U | 是 |
| 5-胺甲醯基甲基尿苷TP | -- | U | 是 |
| 5-甲氧基羰基甲基-2'-0-甲基尿苷 | mcm5Um | U | 是 |
| 5-甲氧基羰基甲基-2-硫代尿苷 | mcm5s2U | U | 是 |
| 5-甲氧基羰基甲基尿苷 | mcm5U | U | 是 |
| 5-甲氧基尿苷 | mo5U | U | 是 |
| 5-甲基-2-硫代尿苷 | m5s2U | U | 是 |
| 5-甲基胺基甲基-2-硒基尿苷 | mnm5se2U | U | 是 |
| 5-甲基胺基甲基-2-硫代尿苷 | mnm5s2U | U | 是 |
| 5-甲基胺基甲基尿苷 | mnm5U | U | 是 |
| 5-甲基二氫尿苷 | -- | U | 是 |
| 5-氧基乙酸-尿苷TP | -- | U | 是 |
| 5-氧基乙酸-甲基酯-尿苷TP | -- | U | 是 |
| Nl-甲基-假-尿苷 | -- | U | 是 |
| 尿苷5-氧基乙酸 | cmo5U | U | 是 |
| 尿苷5-氧基乙酸甲基酯 | mcmo5U | U | 是 |
| 3-(3-胺基-3-羧基丙基)-尿苷TP | -- | U | 是 |
| 5-(異-戊烯基胺基甲基)-2-硫代尿苷TP | -- | U | 是 |
| 5-(異-戊烯基胺基甲基)-2'-0-甲基尿苷TP | -- | U | 是 |
| 5-(異-戊烯基胺基甲基)尿苷TP | -- | U | 是 |
| 懷丁苷 | yW | A/T | 是 |
| 羥基懷丁苷 | OHyW | A/T | 是 |
| 異懷俄苷 | imG2 | A/T | 是 |
| 過氧基懷丁苷 | o2yW | A/T | 是 |

| | | | |
|---|---|---|---|
| 欠修飾之羥基懷丁苷 | OHyW* | A/T | 是 |
| 4-去甲基懷俄苷 | imG-14 | A/T | 是 |

【0069】在一個實施例中，本文所述之遺傳元件包含表7中所提供之非天然存在之修飾或其組合。

**表7：額外例示性非天然存在之修飾**

| 名稱 |
|---|
| 2,6-(二胺基)嘌呤 |
| 1-(氮雜)-2-(硫基)-3-(氮雜)-啡噁 吲-1-基 |
| 1,3-(二氮雜)-2-(側氧基)-吩噻吲-1-基 |
| 1,3-(二氮雜)-2-(側氧基)-啡噁 吲-1-基 |
| 1,3,5-(三氮雜)-2,6-(二氧雜)-萘 |
| 2 (胺基)嘌呤 |
| 2,4,5-(三甲基)苯基 |
| 2' 甲基、2'胺基、2'疊氮基、2'氟-胞嘧啶核苷 |
| 2' 甲基、2'胺基、2'疊氮基、2'氟-腺嘌呤 |
| 2'甲基、2'胺基、2'疊氮基、2'氟-尿苷 |
| 2'-胺基-2'-去氧核糖 |
| 2-胺基-6-氯-嘌呤 |
| 2-氮雜-肌苷基 |
| 2'-疊氮基-2'-去氧核糖 |
| 2'氟-2 '-去氧核糖 |
| 2'-氟-修飾之鹼基 |
| 2'-O-甲基-核糖 |
| 2-側氧基-7-胺基吡啶并嘧啶-3-基 |
| 2-側氧基-吡啶并嘧啶-3-基 |
| 2-吡啶酮 |
| 3 硝基吡咯 |
| 3-(甲基)-7-(丙炔基)異喹諾酮基 |
| 3-(甲基)異喹諾酮基 |
| 4-(氟)-6-(甲基)苯并咪唑 |
| 4-(甲基)苯并咪唑 |
| 4-(甲基)吲哚基 |
| 4,6-(二甲基)吲哚基 |
| 5 硝基吲哚 |
| 5 取代之嘧啶 |
| 5-(甲基)異喹諾酮基 |
| 5-硝基吲哚 |
| 6-(氮雜)嘧啶 |
| 6-(偶氮基)胸腺嘧啶 |
| 6-(甲基)-7-(氮雜)吲哚基 |

| |
|---|
| 6-氯-嘌呤 |
| 6-苯基-吡咯并-嘧啶-2-酮-3-基 |
| 7-(胺基烷基羥基)-1-(氮雜)-2-(硫基)-3-(氮雜)-吩噻吨-1-基 |
| 7-(胺基烷基羥基)-1-(氮雜)-2-(硫基)-3-(氮雜)-<br>啡噁 吨-1-基 |
| 7-(胺基烷基羥基)-1,3-(二氮雜)-2-(側氧基)-啡噁 吨-1-基 |
| 7-(胺基烷基羥基)-1,3-(二氮雜)-2-(側氧基)-吩噻吨-1-基 |
| 7-(胺基烷基羥基)-1,3-(二氮雜)-2-(側氧基)-啡噁 吨-l-基 |
| 7-(氮雜)吲哚基 |
| 7-(胍烷基羥基)-1-(氮雜)-2-(硫基)-3-(氮雜)-啡噁 吨-l-基 |
| 7-(胍烷基羥基)-1-(氮雜)-2-(硫基)-3-(氮雜)-<br>吩噻吨-1-基 |
| 7-(胍烷基羥基)-1-(氮雜)-2-(硫基)-3-(氮雜)-啡噁 吨-1-基 |
| 7-(胍烷基羥基)-1,3-(二氮雜)-2-(側氧基)-啡噁 吨-1-基 |
| 7-(胍烷基-羥基)-1,3-(二氮雜)-2-(側氧基)-<br>吩噻吨-1-基 |
| 7-(胍烷基羥基)-1,3-(二氮雜)-2-(側氧基)-啡噁 吨-1-基 |
| 7-(丙炔基)異喹諾酮基 |
| 7-(丙炔基)異喹諾酮基、丙炔基-7-(氮雜)吲哚基 |
| 7-去氮-肌苷基 |
| 7-取代之1-(氮雜)-2-(硫基)-3-(氮雜)-啡噁 吨-1-基 |
| 7-取代之1,3-(二氮雜)-2-(側氧基)-啡噁 吨-1-基 |
| 9-(甲基)-咪唑并吡啶基 |
| 胺基吲哚基 |
| 蒽基(anthracenyl) |
| 雙-正-(胺基烷基羥基)-6-苯基-吡咯并-嘧啶-2-酮-3-基 |
| 雙-正-取代之-6-苯基-吡咯并-嘧啶-2-酮-3-基 |
| 二氟甲苯基 |
| 次黃嘌呤 |
| 咪唑并吡啶基(imidizopyridinyl) |
| 肌苷基(inosinyl) |
| 異喹諾酮基 |
| 異鳥嘌呤 |
| N2-取代之嘌呤 |
| N6-甲基-2-胺基-嘌呤 |
| N6-取代之嘌呤 |
| N-烷基化衍生物 |
| 萘基 |
| 硝基苯并咪唑基 |
| 硝基咪唑基 |
| 硝基吲唑基 |
| 硝基吡唑基 |

| |
|---|
| 水粉蕈素 |
| O6-取代之嘌呤 |
| O-烷基化衍生物 |
| 正-(胺基烷基羥基)-6-苯基-吡咯并-嘧啶-2-酮-3-基 |
| 正-取代之-6-苯基-吡咯并-嘧啶-2-酮-3-基 |
| 側氧基間型黴素TP (Oxoformycin TP) |
| 對-(胺基烷基羥基)-6-苯基-吡咯并-嘧啶-2-酮-3-基 |
| 對-取代之-6-苯基-吡咯并-嘧啶-2-酮-3-基 |
| 稠五苯基 |
| 菲基(phenanthracenyl) |
| 苯基 |
| 丙炔基-7-(氮雜)吲哚基 |
| 芘基 |
| 吡啶并嘧啶-3-基 |
| 吡啶并嘧啶-3-基、2-側氧基-7-胺基-吡啶并嘧啶-3-基 |
| 吡咯并-嘧啶-2-酮-3-基 |
| 吡咯并嘧啶基 |
| 吡咯并吡咔基 |
| 芪基(stilbenzyl) |
| 取代之1,2,4-三唑 |
| 稠四苯基 |
| 殺結核菌素(tubercidine) |
| 黃嘌呤 |
| 黃嘌呤核苷-5 '-TP |
| 2-硫基-澤布拉林 |
| 5-氮雜-2-硫基-澤布拉林 |
| 7-去氮-2-胺基-嘌呤 |
| 吡啶-4-酮核糖核苷 |
| 2-胺基-核糖苷-TP |
| 間型黴素A TP |
| 間型黴素B TP |
| 吡咯離胺酸TP (Pyrrolosine TP) |
| 2'-OH-阿糖-腺苷TP |
| 2'-OH-阿糖-胞嘧啶核苷TP |
| 2'-OH-阿糖-尿苷TP |
| 2'-OH-阿糖-鳥苷TP |
| 5-(2-甲氧羰基乙烯基)尿苷TP |
| N6-(19-胺基-五氧雜十九烷基)腺苷TP |

【0070】在一個實施例中，本文所述之遺傳元件包含表8中所提供之非天然存在之修飾或其組合。

**表8：例示性主鏈修飾**

| 名稱 |
|---|
| 3'-伸烷基膦酸酯 |
| 3'-胺基胺基磷酸酯 |
| 含有烯烴之主鏈 |
| 胺基烷基胺基磷酸酯 |
| 胺基烷基磷酸三酯 |
| 硼烷磷酸鹽 |
| -CH2-0-N(CH3)-CH2- |
| -CH2-N(CH3)-N(CH3)-CH2- |
| -CH2-NH-CH2- |
| 對掌性膦酸酯 |
| 對掌性硫代磷酸酯 |
| 甲乙醯基及硫代甲乙醯基主鏈 |
| 亞甲基(甲基亞胺基) |
| 亞甲基甲乙醯基及硫代甲乙醯基主鏈 |
| 亞甲基亞胺基及亞甲基肼基主鏈 |
| N-味啉基鍵聯 |
| -N(CH3)-CH2-CH2- |
| 具有雜原子核苷間鍵聯之寡核苷 |
| 亞膦酸酯 |
| 胺基磷酸酯 |
| 二硫代磷酸酯 |
| 硫代磷酸酯核苷間鍵聯 |
| 硫代磷酸酯 |
| 磷酸三酯 |
| PNA |
| 矽氧烷主鏈 |
| 胺基磺酸酯主鏈 |
| 硫化亞碸及碸主鏈 |
| 磺酸酯及磺醯胺主鏈 |
| 硫羰基烷基膦酸酯 |
| 硫羰基烷基磷酸三酯 |
| 硫羰基胺基磷酸酯 |

【0071】在一個實施例中，本文所述之遺傳元件包含表9中所提供之非天然存在之修飾或其組合。

**表9：例示性非天然存在之主鏈修飾**

| 合成主鏈修飾之名稱 |
|---|
| 硫代磷酸酯 |
| 限制性核酸(CNA) |

| |
|---|
| 2' O'甲基化 |
| 2'-O-甲氧基乙基核糖(MOE) |
| 2'氟 |
| 鎖核酸(LNA) |
| (*S*)-限制性乙基(*cEt*) |
| 氟已醣醇核酸(FHNA) |
| 5'硫代磷酸酯 |
| 二胺基磷酸酯N-味啉基寡聚物(PMO) |
| 三環-DNA (tcDNA) |
| (*S*) 5'-C-甲基 |
| (*E*)-乙烯基膦酸酯 |
| 膦酸甲酯 |
| 具有磷酸鹽之(*S*) 5'-C-甲基 |

【0072】 在一些實施例中，遺傳元件包含帽。帽通常置放於mRNA之5'端，但帽亦可安置於RNA之3'端。在一些實施例中，帽保護遺傳元件不被核酸外切酶降解，且可以有助於在細胞內遞送及/或定位。帽可存在於5'-端(5'-端帽)處或3'-端(3'-端帽)處，或可存在於兩端上。5'-端帽之非限制性實例包括但不限於甘油基、反向去氧無鹼基殘基(部分)；4',5'-亞甲基核苷酸；1-(β-D-赤呋喃醣基)核苷酸、4'-硫代核苷酸；碳環核苷酸；1,5-無水己糖醇核苷酸；L-核苷酸；α-核苷酸；修飾鹼基核苷酸；二硫代磷酸酯鍵聯；蘇-呋喃戊醣基核苷酸；非環狀3',4'-斷核苷酸；非環狀3,4-二羥丁基核苷酸；非環狀3,5-二羥基戊基核苷酸、3'-3'-反向核苷酸部分；3'-3'-反向無鹼基部分；3'-2'-反向核苷酸部分；3'-2'-反向無鹼基部分；1,4-丁二醇磷酸酯；3'-胺基磷酸酯；己基磷酸酯；磷酸胺基己酯；3'-磷酸酯；3'-硫代磷酸酯；二硫代磷酸酯；或橋接或非橋接甲基膦酸酯部分。

【0073】 3'-端帽之非限制性實例包括但不限於甘油基、反向去氧無鹼基殘基(部分)、4',5'-亞甲基核苷酸；1-(β-D-赤呋喃醣基)核苷酸；4'-硫代核苷酸、碳環核苷酸；5'-胺基-磷酸烷基酯；磷酸1,3-二胺-2-丙酯；磷酸3-胺丙酯；磷酸6-胺己酯；磷酸1,2-胺基十二烷酯；磷酸羥丙酯；1,5-

無水己糖醇核苷酸；L-核苷酸；α-核苷酸；修飾鹼基核苷酸；二硫代磷酸酯；蘇-呋喃戊醣基核苷酸；非環狀3',4'-斷核苷酸；3,4-二羥丁基核苷酸；3,5-二羥基戊基核苷酸、5'-5'-反向核苷酸部分；5'-5'-反向無鹼基部分；5'-胺基磷酸酯；5'-硫代磷酸酯；1,4-丁二醇磷酸酯；5'-胺基；橋接及/或非橋接5'-胺基磷酸酯、硫代磷酸酯及/或二硫代磷酸酯、橋接或非橋接甲基膦酸酯及5'-巰基部分(更多細節，參見Beaucage及Iyer, 1993, Tetrahedron 49, 1925；以引用之方式併入本文中)。

【0074】在一些實施例中，遺傳元件包含聚A尾。在一些實施例中，聚A尾之長度包含至少約5、10、15、20、25、30、40、50、60、70、80、90或100個腺苷。在一些實施例中，RNA缺乏聚A尾。在一些實施例中，其中RNA缺乏聚A尾，RNA包含不超過約5、10、15、20、25、30、40、50、60、70、80、90或100個連續腺苷。

【0075】在一些實施例中，遺傳元件為線性的。在一些實施例中，遺傳元件為環狀。在一些實施例中，遺傳元件包含第一區域及可與第一區域雜交之第二區域。在一些實施例中，遺傳元件包含第一區域及第二區域，該第二區域可與該第一區域雜交以形成環。在一些實施例中，遺傳元件不包含5'或3'端。在一些實施例中，遺傳元件不包含游離磷酸根及游離糖中之一者或兩者。在一些實施例中，遺傳元件中之每個磷酸根藉由磷酸根所包含之第一氧原子共價連接至第一糖及藉由磷酸根所包含之第二氧原子共價連接至第二糖。在一些實施例中，遺傳元件中之每一糖藉由糖所包含之第一碳原子共價連接至第一磷酸根及藉由糖所包含之第二碳原子共價連接至第二磷酸根。在一些實施例中，遺傳元件係藉由使線性RNA環化而產生。環狀RNA描述於例如美國專利公開案20200306286中，其以全文

引用之方式併入本文中。

【0076】在一些實施例中，遺傳元件之長度為約10-20、20-30、30-40、50-60、60-70、70-80、80-90、90-100、100-125、125-150、150-200、200-300、300-400、400-500、500-600、600-700、700-800、800-900、900-1000、1000-1500、1500-2000、2000-2500、2500-3000、3000-3500、3500-4000或4000-4500個核苷酸。在一些實施例中，遺傳元件之長度為至少約10、20、30、40、50、60、70、80、90、100、125、150、175、200、250、300、400、500、600、700、800、900、1000、1500、2000、2500、3000、3500、4000或4500個核苷酸。

【0077】

***僅RNA遺傳元件***

在一些實施例中，遺傳元件由RNA組成或基本上由RNA組成。舉例而言，在一些實施例中，遺傳元件實質上不含DNA。在一些實施例中，遺傳元件包含單股RNA。在一些實施例中，遺傳元件包含至少一個雙股區。在一些實施例中，遺傳元件之雙股區包含RNA與RNA配對之區域。

【0078】

***雜交RNA-ssDNA遺傳元件***

在一些實施例中，遺傳元件包含DNA區。在一些實施例中，包含RNA之遺傳元件進一步包含DNA區。舉例而言，遺傳元件可為單股的，其中單股之第一部分包含核糖核苷酸且單股之第二部分包含去氧核糖核苷酸。在一些實施例中，包含DNA區之遺傳元件包含一或多個具有化學修飾之DNA核苷酸。在一些實施例中，遺傳元件包含DNA區，其中DNA區之所有核苷酸經化學修飾。

【0079】 在一些實施例中，遺傳元件之至少一部分為單股。在一些實施例中，遺傳元件為單股的。在一些實施例中，遺傳元件包含ssDNA。在一些實施例中，遺傳元件包含雙股區。在一些實施例中，遺傳元件之雙股區包含RNA與RNA配對之區域。在一些此類實施例中，遺傳元件之雙股區包含DNA與RNA配對之區域。在一些實施例中，DNA區之至少一部分與遺傳元件之RNA的至少一部分雜交。

【0080】 在一些實施例中，DNA區之長度為約10-20、20-30、30-40、40-50、50-60、60-70、70-80、80-90、90-100、100-150、150-200、200-300、300-400、400-500、500-600、600-700、700-800、800-900或900-1000個核苷酸。

【0081】

***RNA/DNA結合物***

在一些實施例中，遺傳元件包含DNA區。在一些實施例中，包含RNA之遺傳元件進一步包含DNA區。在一些實施例中，包含DNA區之遺傳元件包含一或多個具有化學修飾之DNA核苷酸。在一些實施例中，遺傳元件包含DNA區，其中，DNA區之所有核苷酸經化學修飾。

【0082】 在一些實施例中，遺傳元件之至少一部分為單股。在一些實施例中，遺傳元件為單股的。在一些實施例中，遺傳元件包含ssDNA。在一些實施例中，遺傳元件包含雙股區。在一些實施例中，遺傳元件之雙股區包含RNA與RNA配對之區域。在一些此類實施例中，遺傳元件之雙股區包含DNA與RNA配對之區域。在其中遺傳元件包含RNA之一些實施例中，DNA區共價連接至遺傳元件之RNA。

【0083】 在一些實施例中，DNA區之長度為約10-20、20-30、30-

40、40-50、50-60、60-70、70-80、80-90、90-100、100-150、150-200、200-300、300-400、400-500、500-600、600-700、700-800、800-900或900-1000個核苷酸。

【0084】

***遺傳元件構築體***

在一些實施例中，遺傳元件由遺傳元件構築體產生。舉例而言，在一些實施例中，遺傳元件構築體為DNA，例如雙股DNA，且遺傳元件可藉由轉錄產生，從而產生RNA遺傳元件。

【0085】 如本文所述之指環載體之遺傳元件可由遺傳元件構築體產生，該構築體包含遺傳元件區域及視情況存在之其他序列，諸如桿狀病毒質體(例如包含桿狀病毒基因體或其片段，例如一或多個桿狀病毒元件)或供體載體主鏈。在一些實施例中，遺傳元件構築體包含指環病毒5' UTR(例如如本文所述)。遺傳元件構築體可為適用於將遺傳元件之序列遞送至宿主細胞或無細胞系統中之任何核酸構築體，其中遺傳元件可包裹於蛋白質外部內。在一些實施例中，遺傳元件構築體包含啟動子。在一些實施例中，自遺傳元件構築體之轉錄產生RNA遺傳元件。

【0086】 在一些實施例中，遺傳元件構築體為線性核酸分子。在一些實施例中，遺傳元件構築體為環狀核酸分子(例如質體、桿狀病毒質體、供體載體或微型環，例如如本文所述)。在一些實施例中，遺傳元件構築體可為雙股的。在其他實施例中，遺傳元件為單股的。在一些實施例中，遺傳元件構築體包含DNA。在一些實施例中，遺傳元件構築體包含RNA。在一些實施例中，遺傳元件構築體包含一或多個經修飾之核苷酸。

【0087】

*質體*

在一些實施例中，遺傳元件構築體為質體。質體將一般包含如本文所描述之遺傳元件之序列以及適用於在宿主細胞中複製之複製起點(例如細菌細胞中複製之細菌複製起點)及可選標記(例如抗生素抗性基因)。在一些實施例中，遺傳元件之序列可自質體切除。在一些實施例中，質體能夠在細菌細胞中複製。在一些實施例中，質體能夠在哺乳動物細胞(例如人類細胞)中複製。在一些實施例中，質體之長度為至少300、400、500、600、700、800、900、1000、2000、3000、4000或5000 bp。在一些實施例中，質體之長度為小於600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000或10,000 bp。在一些實施例中，質體之長度在300-400、400-500、500-600、600-700、700-800、800-900、900-1000、1000-1500、1500-2000、2000-2500、2500-3000、3000-4000或4000-5000 bp之間。

【0088】

*小環狀核酸構築體*

在一些實施例中，遺傳元件構築體為環狀核酸構築體，例如缺乏載體主鏈(例如缺乏細菌複製起點及/或可選標記)。在實施例中，遺傳元件為單股或雙股環狀核酸構築體。在實施例中，環狀核酸構築體係藉由活體外環化(IVC)產生，例如如本文所述。在實施例中，雙股環狀核酸構築體可引入至宿主細胞中，在其中其可轉化為或用作用於產生單股環狀遺傳元件之模板，例如如本文所描述。在一些實施例中，環狀核酸構築體不包含質體主鏈或其功能片段。在一些實施例中，環狀核酸構築體之長度為至少

2000、2100、2200、2300、2400、2500、2600、2700、2800、2900、3000、3100、3200、3300、3400、3500、3600、3700、3800、3900、4000、4100、4200、4300、4400或4500 bp。在一些實施例中，環狀核酸構築體之長度小於2900、3000、3100、3200、3300、3400、3500、3600、3700、3800、3900、4000、4100、4200、4300、4400、4500、4600、4700、4800、4900、5000、5500或6000 bp。在一些實施例中，環狀核酸構築體之長度在2000-2100、2100-2200、2200-2300、2300-2400、2400-2500、2500-2600、2600-2700、2700-2800、2800-2900、2900-3000、3000-3100、3100-3200、3200-3300、3300-3400、3400-3500、3500-3600、3600-3700、3700-3800、3800-3900、3900-4000、4000-4100、4100-4200、4200-4300、4300-4400或4400-4500 bp之間。在一些實施例中，環狀核酸構築體為微型環。

【0089】

順式/反式構築體

在一些實施例中，如本文所述之遺傳元件構築體(例如桿狀病毒質體或供體載體)包含一或多個編碼一或多個指環病毒ORF之序列，例如蛋白質外部組分(例如由指環病毒ORF1核酸編碼之多肽，例如如本文所述)。例如，遺傳元件構築體可包含編碼指環病毒ORF1分子之核酸序列。*此類遺傳元件構築體可適於以順式將遺傳元件及指環病毒ORF引入至宿主細胞中*。在其他實施例中，如本文所描述之遺傳元件構築體不包含編碼一或多個指環病毒ORF之序列，例如蛋白質外部組分(例如由指環病毒ORF1核酸編碼之多肽，例如如本文所描述)。例如，遺傳元件構築體可不包含編碼指環病毒ORF1分子之核酸序列。此類遺傳元件構築體可適用於將遺傳元

件引入至宿主細胞中，其中一或多個指環病毒ORF以反式提供(例如經由引入編碼指環病毒ORF中之一或多者的第二核酸構築體，或經由整合至宿主細胞之基因體中的指環病毒ORF卡匣)。在一些實施例中，遺傳元件構築體包含適用於在昆蟲細胞(例如Sf9細胞)中複製核酸構築體之主鏈，例如桿狀病毒主鏈區。在一些實施例中，遺傳元件構築體包含適用於在細菌細胞(例如大腸桿菌細胞，例如DH 10Bac細胞)中複製遺傳元件構築體之主鏈區。

【0090】 在一些實施例中，遺傳元件構築體(例如桿狀病毒質體或供體載體)包含編碼指環病毒ORF1分子或其剪接變異體或功能片段(例如果醬捲區，例如如本文所述)的序列。在實施例中，不包含遺傳元件序列之遺傳元件部分包含編碼指環病毒ORF1分子之序列或其剪接變異體或功能片段(例如在包含啟動子及編碼指環病毒ORF1分子或其剪接變異體或功能片段之序列的卡匣中)。在其他實施例中，包含遺傳元件序列之構築體部分包含編碼指環病毒ORF1分子或其剪接變異體或功能片段之序列(例如果醬捲，例如如本文所描述)。在實施例中，將此類遺傳元件包裹於蛋白質外部(例如如本文所描述)中產生複製組分指環載體(例如在感染細胞後使得細胞能夠產生指環載體之額外複本而不將其他核酸構築體(例如編碼如本文所描述之一或多個指環病毒ORF)引入至細胞中的指環載體)。

【0091】 在其他實施例中，遺傳元件不包含編碼指環病毒ORF1分子或其剪接變異體或功能片段的序列(例如果醬捲區，例如如本文所描述)。在實施例中，將此類遺傳元件包裹於蛋白質外部(例如如本文所描述)中產生複製不勝任型指環載體(例如在感染細胞後不能使感染細胞產生額外指環載體之指環載體，例如在不存在一或多種額外構築體，例如編碼

如本文所描述之一或多個指環病毒ORF的情況下)。

【0092】

*表現卡匣*

在一些實施例中，遺傳元件構築體(例如桿狀病毒質體或供體載體)包含一或多個用於表現多肽或非編碼RNA (例如miRNA或siRNA)的卡匣。在一些實施例中，遺傳元件構築體包含用於表現效應子(例如外源性或內源性效應子)，例如如本文所描述之多肽或非編碼RNA的卡匣。在一些實施例中，遺傳元件構築體包含用於表現指環病毒蛋白質(例如指環病毒ORF1、ORF2、ORF2/2、ORF2/3、ORF1/1或ORF1/2或其功能片段)的卡匣。在一些實施例中，表現卡匣可位於遺傳元件序列內。在實施例中，效應子之表現卡匣位於遺傳元件序列內。在實施例中，指環病毒蛋白質之表現卡匣位於遺傳元件序列內。在其他實施例中，表現卡匣位於遺傳元件序列之外的遺傳元件構築體內的位置(例如主鏈中)。在實施例中，指環病毒蛋白質之表現卡匣位於遺傳元件序列之外的遺傳元件構築體內的位置(例如主鏈中)。

【0093】 多肽表現卡匣一般包含啟動子及編碼多肽之編碼序列，該多肽例如效應子(例如如本文所描述之外源性或內源性效應子)或指環病毒蛋白(例如編碼指環病毒ORF1、ORF2、ORF2/2、ORF2/3、ORF1/1或ORF1/2或其功能片段的序列)。可包括於多肽表現卡匣中(例如以驅動多肽表現)之例示性啟動子包括但不限於組成性啟動子(例如CMV、RSV、PGK、EF1a或SV40)、細胞或組織特異性啟動子(例如骨骼α-肌動蛋白啟動子、肌凝蛋白輕鏈2A啟動子、肌縮蛋白啟動子、肌肉肌酸激酶啟動子、肝白蛋白啟動子、B型肝炎病毒核心啟動子、骨鈣化素啟動子、骨唾

液蛋白啟動子、CD2啟動子、免疫球蛋白重鏈啟動子、T細胞受體a鏈啟動子、神經元特異性烯醇化酶(NSE)啟動子或神經絲輕鏈啟動子)及誘導性啟動子(例如鋅誘導性綿羊金屬硫蛋白(MT)啟動子；地塞米松(Dex)誘導性小鼠乳房腫瘤病毒(MMTV)啟動子；T7聚合酶啟動子系統、四環素抑制性系統、四環素誘導性系統、RU486誘導性系統、雷帕黴素誘導性系統)，例如如本文所描述。在一些實施例中，表現卡匣進一步包含強化子，例如如本文所描述。

【0094】

***基於RNA之遺傳元件之產生***

基於RNA之遺傳元件可藉由各種方法產生。舉例而言，包含DNA之遺傳元件構築體可經轉錄以產生包含RNA之遺傳元件，例如如上文所描述。轉錄可發生於例如細胞或無細胞系統中。RNA可在活體外合成，例如藉由固相合成。

【0095】

***蛋白質組分之產生***

指環載體，例如ORF1之蛋白質組分可以本文所述之多種方式產生。

【0096】

***桿狀病毒表現系統***

病毒表現系統，例如桿狀病毒表現系統，可用於表現蛋白質(例如用於產生指環載體)，例如如本文所述。桿狀病毒為桿形病毒，具有環狀、超螺旋雙股DNA基因體。桿狀病毒屬包括：α桿狀病毒(自鱗翅目昆蟲分離之核型多角體病毒(NPV))、β桿狀病毒(自鱗翅目昆蟲分離之顆粒體病毒(GV))、γ桿狀病毒(自膜翅目昆蟲分離之NPV)及λ桿狀病毒(自雙翅目昆

蟲分離之NPV)。雖然GV通常每個包膜僅含有一個核蛋白殼，但NPV通常每個包膜含有單個(SNPV)或多個(MNPV)核蛋白殼。包膜病毒粒子在GV中進一步封閉於顆粒體蛋白基質中且在NPV中進一步封閉於多角體蛋白中。桿狀病毒通常具有溶解及封閉的生命週期。在一些實施例中，溶解及包涵的生命週期貫穿三個病毒複製階段獨立地表現：早期、晚期及極晚期。在一些實施例中，在早期階段內，病毒DNA複製發生在病毒進入宿主細胞、早期病毒基因表現及宿主基因表現機構關閉之後。在一些實施例中，在表現編碼病毒DNA複製之晚期基因的晚期階段內，病毒粒子經組裝，且由宿主細胞產生細胞外病毒(EV)。在一些實施例中，在表現多角體蛋白及p10基因之極晚期內，包涵體病毒(OV)由宿主細胞產生，且宿主細胞經溶解。由於桿狀病毒感染昆蟲物種，因此其可用作生物製劑以在桿狀病毒容許之昆蟲細胞或幼蟲中產生外源性蛋白質。不同桿狀病毒分離株，諸如加洲苜蓿夜蛾(Autographa californica)多核多角體病毒(AcMNPV)及家蠶(蠶)核多角體病毒(BmNPV)可用於外源性蛋白質表現。各種桿狀病毒表現系統為市售的，例如獲自ThermoFisher。

【0097】 在一些實施例中，本文所述之蛋白質(例如指環病毒ORF分子，例如ORF1、ORF2、ORF2/2、ORF2/3、ORF1/1或ORF1/2或其功能片段或剪接變異體)可使用包含一或多種本文所述之組分的桿狀病毒表現載體(例如桿狀病毒質體)表現。舉例而言，桿狀病毒表現載體可包括以下中之一或多者(例如全部)：可選標記(例如kanR)、複製起點(例如細菌複製起點及昆蟲細胞複製起點中之一或兩者)、重組酶識別位點(例如att位點)及啟動子。在一些實施例中，桿狀病毒表現載體(例如如本文所述之桿狀病毒質體)可藉由用編碼本文所述之蛋白質的基因替換編碼桿狀病毒包涵

體的天然存在之野生型多角體蛋白基因產生。在一些實施例中，將編碼本文所述之蛋白質的基因選殖至含有桿狀病毒啟動子之桿狀病毒表現載體(例如如本文所述之桿狀病毒質體)中。在一些實施例中，桿狀病毒載體包含一或多種非桿狀病毒啟動子，例如哺乳動物啟動子或指環病毒啟動子。在一些實施例中，將編碼本文所述之蛋白質的基因選殖至供體載體(例如如本文所述)中，接著使該供體載體與空桿狀病毒表現載體(例如空桿狀病毒質體)接觸，使得編碼本文所述之蛋白質的基因自供體載體轉移(例如藉由同源重組或轉座酶活性)至桿狀病毒表現載體(例如桿狀病毒質體)中。在一些實施例中，桿狀病毒啟動子由來自非必需多角體蛋白基因座之桿狀病毒DNA側接。在一些實施例中，本文所述之蛋白質在處於病毒複製極後期之AcNPV多角體蛋白啟動子的轉錄控制下。在一些實施例中，適用於昆蟲細胞中之桿狀病毒表現之強啟動子包括但不限於桿狀病毒p10啟動子、多角體蛋白(polh)啟動子、p6.9啟動子及蛋白殼蛋白啟動子。適用於昆蟲細胞中之桿狀病毒表現的弱啟動子包括桿狀病毒之ie1、ie2、ie0、et1、39K (亦稱為pp31)及gp64啟動子。

【0098】 在一些實施例中，重組桿狀病毒由桿狀病毒基因體(例如，野生型或突變型桿狀病毒基因體)與轉移載體之間進行同源重組產生。在一些實施例中，將一或多個編碼本文所述之蛋白質的基因選殖至轉移載體中。在一些實施例中，轉移載體進一步含有由來自非必需基因座(例如多角體蛋白基因)之DNA側接的桿狀病毒啟動子。在一些實施例中，藉由桿狀病毒基因體與轉移載體之間的同源重組會將一或多個編碼本文所述之蛋白質的基因插入桿狀病毒基因體中。在一些實施例中，桿狀病毒基因體在一或多個獨特位點經線性化。在一些實施例中，線性化位點位於用於將編

碼本文所述之蛋白質的基因插入至桿狀病毒基因體中的目標位點附近。在一些實施例中，在基因(例如多角體蛋白基因)下游缺失桿狀病毒基因體之片段的線性化桿狀病毒基因體可用於同源重組。在一些實施例中，桿狀病毒基因體及轉移載體共轉染至昆蟲細胞中。在一些實施例中，產生重組桿狀病毒之方法包含以下步驟：製備桿狀病毒基因體，以與含有編碼一或多種本文所述之蛋白質之基因的轉移載體進行同源重組，且將轉移載體及桿狀病毒基因體DNA共轉染至昆蟲細胞中。在一些實施例中，桿狀病毒基因體包含與轉移載體之區域同源的區域。此等同源區域可增強桿狀病毒基因體與轉移載體之間的重組機率。在一些實施例中，轉移載體中之同源區域位於啟動子上游或下游。在一些實施例中，為誘導同源重組，將桿狀病毒基因體及轉移載體以約1:1至10:1之重量比混合。

【0099】 在一些實施例中，重組桿狀病毒由包含藉由Tn7進行位點特異性轉座(site-specific transposition)的方法產生，例如藉此將編碼本文所述之蛋白質的基因插入至桿狀病毒質體DNA中，例如在細菌，例如大腸桿菌(例如DH 10Bac細胞)中繁殖。在一些實施例中，將編碼本文所述之蛋白質的基因選殖至pFASTBAC®載體中且轉型至含有具有微型*att*Tn7目標位點之桿狀病毒質體DNA的勝任細胞，例如DH10BAC®勝任細胞中。在一些實施例中，桿狀病毒表現載體，例如pFASTBAC®載體可具有啟動子，例如雙啟動子(例如多角體蛋白啟動子、p10啟動子)。市售之pFASTBAC®供體質體包括：pFASTBAC 1、pFASTBAC HT及pFASTBAC DUAL。在一些實施例中，鑑別含有重組桿狀病毒質體DNA之菌落且分離桿狀病毒質體DNA以轉染昆蟲細胞。

【0100】 在一些實施例中，桿狀病毒載體連同輔助核酸一起引入至

昆蟲細胞中。引入可為同時的或依序的。在一些實施例中，輔助核酸提供一或多種桿狀病毒蛋白質，例如以促進桿狀病毒載體之封裝。

【0101】 在一些實施例中，昆蟲細胞中產生之重組桿狀病毒(例如藉由同源重組)經擴增且用於感染用於重組蛋白表現的昆蟲細胞(例如在對數生長中期內)。在一些實施例中，藉由細菌(例如大腸桿菌)中之位點特異性轉座產生的重組桿狀病毒質體DNA用於使用轉染劑(例如Cellfectin® II)轉染昆蟲細胞。關於桿狀病毒表現系統之額外資訊論述於美國專利申請案第14/447,341號、第14/277,892號及第12/278,916號中，該等申請案以引用之方式併入本文中。

【0102】

***昆蟲細胞系統***

本文所述之蛋白質可在經重組桿狀病毒或桿狀病毒質體DNA感染或轉染之昆蟲細胞中表現，例如如上文所述。在一些實施例中，昆蟲細胞包括：衍生自草地黏蟲之Sf9及Sf21細胞以及衍生自粉紋夜蛾(*Trichoplusia ni*)之Tn-368及High Five™ BTI-TN-5B1-4細胞(亦稱為Hi5細胞)。在一些實施例中，衍生自草地黏蟲蛹秋黏蟲之卵巢的昆蟲細胞株Sf21及Sf9可用於使用桿狀病毒表現系統來表現重組蛋白。在一些實施例中，Sf21及Sf9昆蟲細胞可在市售補充血清或無血清培養基中培養。適用於培養昆蟲細胞之培養基包括：格里斯氏補充(Grace's Supplemented) (TNM-FH)、IPL-41、TC-100、施奈德果蠅(Schneider's Drosophila)、SF-900 II SFM或及EXPRESS-FIVE™ SFM。在一些實施例中，一些無血清培養基調配物利用磷酸鹽緩衝液系統將培養物pH維持在6.0-6.4範圍內(Licari等人 Insect cell hosts for baculovirus expression vectors contain endogenous

exoglycosidase activity. Biotechnology Progress 9: 146-152 (1993)及Drugmand等人 Insect cells as factories for biomanufacturing. Biotechnology Advances 30:1140-1157 (2012))，以用於培養及重組蛋白生產。在一些實施例中，可使用6.0-6.8之pH培養各種昆蟲細胞株。在一些實施例中，昆蟲細胞在25℃至30℃之間的溫度下、在充氣下懸浮培養或以單層形式培養。關於昆蟲細胞之額外資訊論述於例如美國專利申請案第14/564,512號及第14/775,154號中，其各自以引用的方式併入本文中。

【0103】

***哺乳動物細胞系統***

在一些實施例中，本文所描述之蛋白質可在用例如如本文所描述之編碼蛋白質之載體感染或轉染的動物細胞株中活體外表現。在本發明之上下文中設想之動物細胞株包括豬細胞株，例如永生化豬細胞株，諸如但不限於豬腎上皮細胞株PK-15及SK、單骨髓細胞株3D4/31及睪丸細胞株ST。另外，包括其他哺乳動物細胞株，諸如CHO細胞(中國倉鼠卵巢)、MARC-145、MDBK、RK-13、EEL。另外或替代地，本發明之方法之特定實施例利用作為上皮細胞株，亦即上皮譜系細胞之細胞株的動物細胞株。適用於表現本文所描述之蛋白質的細胞株包括但不限於人類或靈長類動物來源之細胞株，諸如人類或靈長類動物腎臟癌細胞株。

【0104】

***效應子***

本文所述之組合物及方法可用於產生包含編碼效應子(例如外源性效應子或內源性效應子)之序列的指環載體之遺傳元件，例如如本文所述。在一些實施例中，遺傳元件為效應子，例如遺傳元件為功能性RNA。在

一些情況下，效應子可為內源性效應子或外源性效應子。在一些實施例中，效應子係治療性效應子。在一些實施例中，效應子包含多肽(例如治療性多肽或肽，例如如本文所描述)。在一些實施例中，效應子包含非編碼RNA (例 如miRNA、siRNA、shRNA、mRNA、lncRNA、RNA、DNA、反義RNA或gRNA)。在一些實施例中，效應子包含例如如本文所描述之調控核酸。

【0105】

*活體外組裝方法*

可例如藉由活體外組裝產生指環載體。在一些實施例中，遺傳元件在允許組裝之條件下活體外接觸ORF1。

【0106】 在一些實施例中，桿狀病毒構築體用於產生指環病毒蛋白質。此等蛋白質接著可用於例如活體外組裝以蛋白殼化遺傳元件，例如包含RNA之遺傳元件。在一些實施例中，編碼一或多種指環病毒蛋白質之聚核苷酸與啟動子融合以用於在宿主細胞，例如昆蟲或動物細胞中表現。在一些實施例中，將聚核苷酸選殖至桿狀病毒表現系統中。在一些實施例中，宿主細胞，例如昆蟲細胞經桿狀病毒表現系統感染且培育一段時間。在一些實施例中，將經感染細胞培育約1、2、3、4、5、10、15或20天。在一些實施例中，將經感染細胞溶解以回收指環病毒蛋白質。

【0107】 在一些實施例中，純化經分離之指環病毒蛋白質。在一些實施例中，指環病毒蛋白質使用純化技術純化，包括但不限於螯合純化、肝素純化、梯度沈降純化及/或SEC純化。在一些實施例中，將經純化之指環病毒蛋白質與遺傳元件混合以蛋白殼化遺傳元件，例如包含RNA之遺傳元件。在一些實施例中，使用ORF1蛋白、ORF2蛋白或其經修飾形式

來蛋白殼化遺傳元件。在一些實施例中，兩個核酸經蛋白殼化。舉例而言，第一核酸可為mRNA，例如經化學修飾之mRNA，且第二核酸可為DNA。

【0108】 在一些實施例中，編碼指環病毒(AV) ORF1之DNA (例如野生型ORF1蛋白、含有突變例如以改良組裝效率、產量或穩定性之ORF1蛋白、嵌合ORF1蛋白或其片段)表現於昆蟲細胞株(例如Sf9及/或HighFive)、動物細胞株(例如雞細胞株(MDCC))、細菌細胞(例如大腸桿菌)及/或哺乳動物細胞株(例如293expi及/或MOLT4)中。在一些實施例中，編碼AV ORF1之DNA可不經標記。在一些實施例中，編碼AV ORF1之DNA可含有N端及/或C端融合之標籤。在一些實施例中，編碼AV ORF1之DNA可在ORF1蛋白內具有突變、插入或缺失以引入標籤，例如以例如經由免疫染色分析(包括但不限於ELISA或西方墨點法)來輔助純化及/或鑑別確定。在一些實施例中，編碼AV ORF1之DNA可單獨或與任何數目之輔助蛋白質組合表現。在一些實施例中，編碼AV ORF1之DNA與AV ORF2及/或ORF3蛋白組合表現。

【0109】 在一些實施例中，具有改良組裝效率之突變的ORF1蛋白可包括但不限於具有引入至N端精胺酸臂(ARG臂)中以改變ARG臂之pI，從而容許pH敏感性核酸結合至觸發粒子組裝(SEQ ID 3-5)之突變的ORF1蛋白。在一些實施例中，具有改良穩定性之突變的ORF1蛋白可包括接觸典型果醬捲β桶形之β股F及G的內原聚體之突變，以改變原聚體表面之疏水性狀態且提高蛋白殼形成之熱力學有利性。

【0110】 在一些實施例中，嵌合ORF1蛋白可包括但不限於ORF1蛋白，其序列的一或多個部分經另一蛋白殼蛋白的可比部分置換，該蛋白殼

蛋白為例如喙羽病病毒(BFDV)蛋白殼蛋白，或E型肝炎蛋白殼蛋白，例如環9 ORF1的ARG臂或F及G β股經來自BFDV蛋白殼蛋白的可比組分置換。在一些實施例中，嵌合ORF1蛋白亦可包括其序列之一或多個部分經另一AV ORF1蛋白之可比部分置換(例如環2 ORF1之果醬捲片段或C端部分經環9 ORF1之可比部分置換)的ORF1蛋白。

【0111】 在一些實施例中，本發明描述製成指環載體之方法，該方法包含：(a)提供包含以下之混合物：(i)包含RNA之遺傳元件，及(ii) ORF1分子；及(b)在適用於將遺傳元件包裹於包含ORF1分子之蛋白質外部內的條件下培育該混合物，由此製成指環載體；視情況其中該混合物不包含於細胞中。在一些實施例中，該方法在提供(a)之前進一步包含例如在宿主細胞，例如昆蟲細胞或哺乳動物細胞中表現ORF1分子。在一些實施例中，表現包含在適用於產生ORF1分子之條件下培育包含編碼ORF1分子之核酸分子(例如桿狀病毒表現載體)的宿主細胞(例如昆蟲細胞或哺乳動物細胞)。在一些實施例中，該方法進一步包含在提供(a)之前純化由宿主細胞表現之ORF1分子。在一些實施例中，該方法在無細胞系統中進行。在一些實施例中，本發明描述一種製造指環載體組合物之方法，其包含：(a)提供複數個如前述實施例中任一項之指環載體或組合物；(b)視情況針對以下中之一或多者評估該複數個指環載體或組合物：本文所描述之污染物、光密度量測(例如，OD 260)、粒子數目(例如，藉由HPLC)、感染性(例如，粒子:感染性單位比率，例如，如藉由螢光及/或ELISA測定)；及(c)例如如何(b)之參數中之一或多者符合指定臨限值時，調配該複數個指環載體，例如作為適用於向個體投與之醫藥組合物。

【0112】

*富集及純化*

可純化及/或富集所收集之指環載體，例如以產生指環載體製劑。在一些實施例中，所收集之指環載體與存在於收集溶液中之其他成分或污染物分離，例如使用此項技術中已知用於純化病毒粒子之方法(例如藉由沈降、層析及/或超濾進行純化)。在一些實施例中，所收集之指環載體藉由親和純化(例如肝素親和純化)純化。在一些實施例中，收集之指環載體藉由尺寸排阻層析(例如使用Tris緩衝液移動相)純化。在一些實施例中，收集之指環載體藉由陰離子交換層析(例如Mustang Q膜層析)純化。在一些實施例中，收集之指環載體藉由混合模式層析(例如使用混合模式樹脂，例如Cato700樹脂)純化。在一些實施例中，純化步驟包含自製劑中移除血清、宿主細胞DNA、宿主細胞蛋白質、缺乏遺傳元件之粒子及/或酚紅中之一或多者。在一些實施例中，收集之指環載體相對於收集溶液中存在之其他成分或污染物富集，例如使用此項技術中已知用於富集病毒粒子之方法。

【0113】 在一些實施例中，所得製備或包含製劑之醫藥組合物在可接受之時段及溫度內將為穩定的，及/或與所需投與途徑及/或此投與途徑將需要之任何裝置，例如針或注射器相容。

【0114】

**II.指環載體**

在一些態樣中，本發明提供使用及製成指環載體、指環載體製劑之組合物及方法及使用及製成治療組合物之方法。在一些實施例中，指環載體包含一或多種核酸或多肽，其包含基於指環病毒(例如如本文所述之指環病毒)或其片段或部分，或其他實質上非致病性病毒，例如共生性病

毒、共生病毒、天然病毒之序列、結構及/或功能。在一些實施例中，基於指環病毒之指環載體包含至少一個外源性針對指環病毒為外源性之元件，例如外源性效應子或編碼安置於指環載體之遺傳元件內之外源性效應子的核酸序列。在一些實施例中，基於指環病毒之指環載體包含至少一個與來自指環病毒之另一元件異源的元件，例如與另一連接核酸序列，諸如啟動子元件異源之效應子編碼核酸序列。在一些實施例中，指環載體包含遺傳元件(例如環狀DNA，例如單股DNA)，其包含相對於遺傳元件及/或蛋白質外部之其餘部分異源的至少一個元件(例如編碼效應子之外源性元件，例如如本文所述)。指環載體可為用於有效負載至宿主(例如人類)中之遞送媒劑(例如實質上非致病性遞送媒劑)。在一些實施例中，指環載體能夠在真核細胞，例如哺乳動物細胞，例如人類細胞中複製。在一些實施例中，指環載體為實質上非致病性的及/或實質上非整合於哺乳動物(例如人類)細胞中。在一些實施例中，指環載體在哺乳動物(例如人類)中實質上為非免疫原性的。在一些實施例中，指環載體為複製缺陷型的。在一些實施例中，指環載體為複製勝任型的。

【0115】 在一些實施例中，指環載體包含curon或其組分(例如遺傳元件，例如包含編碼效應子之序列及/或蛋白質外部)，例如如PCT申請案第PCT/US2018/037379中所述，其以全文引用之方式併入本文中。在一些實施例中，指環載體包含指環載體或其組分(例如遺傳元件，例如包含編碼效應子之序列及/或蛋白質外部)，例如如PCT申請案第PCT/US19/65995中所述，其以全文引用之方式併入本文中。

【0116】 在一態樣中，本發明包括一種指環載體，其包含(i)包含啟動子元件之遺傳元件、編碼效應子之序列(例如內源性效應子或外源性效

應子，例如有效負載)及蛋白質結合序列(例如外部蛋白質結合序列，例如封裝信號)，其中遺傳元件為單股DNA，且具有以下特性中之一或兩者：為環狀及/或以小於進入細胞之遺傳元件之約0.001%、0.005%、0.01%、0.05%、0.1%、0.5%、1%、1.5%或2%的頻率整合至真核細胞之基因體中；及(ii)蛋白質外部；其中遺傳元件包裹於蛋白質外部內；且其中指環載體能夠將遺傳元件遞送至真核細胞中。

【0117】 在本文所描述之指環載體之一些實施例中，遺傳元件以小於進入細胞之遺傳元件的約0.001%、0.005%、0.01%、0.05%、0.1%、0.5%、1%、1.5%或2%之頻率整合。在一些實施例中，來自投與至個體之複數個指環載體之小於約0.01%、0.05%、0.1%、0.5%、1%、2%、3%、4%或5%的遺傳元件將整合至個體之一或多個宿主細胞的基因體中。在一些實施例中，例如如本文所描述之指環載體群體之遺傳元件以小於可比AAV病毒群體之頻率的頻率，例如以比可比AAV病毒群體低約50%、60%、70%、75%、80%、85%、90%、95%、100%或以上的頻率，整合至宿主細胞之基因體中。

【0118】 在一態樣中，本發明包括一種指環載體，其包含：(i)遺傳元件，該遺傳元件包含啟動子元件、編碼效應子(例如內源性效應子或外源性效應子，例如有效負載)之序列及蛋白質結合序列(例如外部蛋白質結合序列)，其中該遺傳元件與野生型指環病毒序列(例如野生型細環病毒屬(Torque Teno virus，TTV)、微型細環病毒屬(Torque Teno mini virus，TTMV)或TTMDV序列，例如如本文所述之野生型指環病毒序列)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%)的序列一致

性；及(ii)蛋白質外部；其中該遺傳元件包裹在該蛋白質外部內；且其中該指環載體能夠將該遺傳元件遞送至真核細胞中。

【0119】 在一個態樣中，本發明包括一種指環載體，其包含：

a)遺傳元件，其包含(i)編碼外部蛋白質(例如非病原性外部蛋白質)之序列、(ii)將遺傳元件結合至非病原性外部蛋白質之外部蛋白質結合序列及(iii)編碼效應子(例如內源性或外源性效應子)之序列；及

b)與遺傳元件相關聯之外部蛋白質，例如包封或包裹該遺傳元件。

【0120】 在一些實施例中，指環載體包括來自無包膜、環狀、單股DNA病毒(或與其具有>70%、75%、80%、85%、90%、95%、97%、98%、99%、100%同源性)之序列或表現產物。動物環狀單股DNA病毒一般係指單股DNA (ssDNA)病毒之子組，其感染真核非植物宿主且具有環狀基因體。因此，動物環狀ssDNA病毒可與感染原核生物之ssDNA病毒(亦即微小噬菌體科(Microviridae)及絲狀噬菌體科(Inoviridae))及感染植物之ssDNA病毒(亦即雙生病毒科(Geminiviridae)及矮化病毒科(Nanoviridae))區分。其亦可與感染非植物真核生物之線性ssDNA病毒(亦即小病毒科(Parvoviridiae))區分。

【0121】 在一些實施例中，指環載體例如短暫或長期地調節宿主細胞功能。在某些實施例中，細胞功能經穩定改變，諸如調節持續至少約1小時至約30天，或至少約2小時、6小時、12小時、18小時、24小時、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天或更長時間或其間的任何時間。在某些實施例中，細胞功能短暫改變，例如諸如調節持續不超

過約30分鐘至約7天，或不超過約1小時、2小時、3小時、4小時、5小時、6小時、7小時、8小時、9小時、10小時、11小時、12小時、13小時、14小時、15小時、16小時、17小時、18小時、19小時、20小時、21小時、22小時、24小時、36小時、48小時、60小時、72小時、4天、5天、6天、7天或其間的任何時間。

【0122】在一些實施例中，遺傳元件包含啟動子元件。在實施例中，啟動子元件係選自RNA聚合酶II依賴性啟動子、RNA聚合酶III依賴性啟動子、PGK啟動子、CMV啟動子、EF-1α啟動子、SV40啟動子、CAGG啟動子或UBC啟動子、TTV病毒啟動子、組織特異性U6(pollIII)、具有活化蛋白之上游DNA結合位點的最小CMV啟動子(TetR-VP16、Gal4-VP16、dCas9-VP16等)。在實施例中，啟動子元件包含TATA盒。在實施例中，啟動子元件對於野生型指環病毒為內源性的，例如如本文所述。

【0123】在一些實施例中，遺傳元件包含以下特徵中之一或多者：單股、環狀、負股及/或DNA。在實施例中，遺傳元件包含游離基因體。在一些實施例中，不包括效應子之遺傳元件部分具有約2.5-5 kb (例如約2.8-4 kb、約2.8-3.2 kb、約3.6-3.9 kb或約2.8-2.9 kb)、小於約5 kb (例如，小於約2.9 kb、3.2 kb、3.6 kb、3.9 kb或4 kb)或至少100個核苷酸(例如至少1 kb)之組合尺寸。

【0124】在一些實施例中，將指環載體或包含於指環載體中之遺傳元件引入至細胞(例如人類細胞)中。在一些實施例中，例如在指環載體或遺傳元件已引入至細胞中之後，例如由指環載體之遺傳元件編碼之效應子(例如RNA，例如miRNA)表現於細胞(例如人類細胞)中。在實施例中，將

指環載體或包含於其中之遺傳元件引入至細胞中以調節(例如增加或降低)細胞中之目標分子(例如目標核酸，例如RNA或目標多肽)之水準，例如藉由改變細胞之目標分子的表現量。在實施例中，引入指環載體或包含於其中之遺傳元件會降低細胞所產生之干擾素水準。在實施例中，將指環載體或包含於其中之遺傳元件引入至細胞中會調節(例如增加或降低)細胞之功能。在實施例中，將指環載體或包含於其中之遺傳元件引入至細胞中會調節(例如增加或降低)細胞之活力。在實施例中，將指環載體或包含於其中之遺傳元件引入至細胞中會降低細胞(例如癌細胞)之活力。

【0125】 在一些實施例中，本文所描述之指環載體(例如合成指環載體)誘導小於70%之抗體發病率(例如小於約60%、50%、40%、30%、20%或10%抗體發病率)。在實施例中，抗體發病率係根據此項技術中已知之方法測定。在實施例中，抗體流行率係例如根據Tsuda等人(1999; *J. Virol. Methods* 77: 199-206；以引用之方式併入本文中)所述之抗TTV抗體偵測方法，及/或Kakkola等人(2008; *Virology* 382: 182-189；以引用之方式併入本文中)所述之用於測定抗TTV IgG血清陽性率之方法，藉由偵測生物樣品中針對指環病毒(例如如本文所述)或基於指環病毒之指環載體的抗體來測定。針對指環病毒或基於指環病毒之指環載體的抗體亦可藉由此項技術中用於偵測抗病毒抗體之方法來偵測，例如偵測抗AAV抗體之方法，例如如Calcedo等人(2013; *Front. Immunol.* 4(341): 1-7；以引用之方式併入本文中)中所述。

【0126】 在一些實施例中，複製缺失型、複製缺陷型或複製不勝任型遺傳元件不編碼複製遺傳元件所需之所有必需機構或組分。在一些實施例中，複製缺陷型遺傳元件不編碼複製因子。在一些實施例中，複製缺陷

型遺傳元件不編碼一或多個ORF (例如ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3及/或ORF2t/3，例如如本文所描述)。在一些實施例中，可以反式提供(例如編碼於宿主細胞所包含之核酸中，例如整合至宿主細胞之基因體中)不由遺傳元件編碼之機構或組分，例如使得遺傳元件可在以反式提供之機構或組分存在下經歷複製。

【0127】 在一些實施例中，封裝缺失型、封裝缺陷型或封裝不勝任型遺傳元件無法封裝至蛋白質外部(例如其中蛋白質外部包含蛋白殼或其部分，例如包含由ORF1核酸編碼之多肽，例如如本文所描述)。在一些實施例中，相比於野生型指環病毒(例如如本文所描述)，封裝缺失型遺傳元件以小於10% (例如小於10%、9%、8%、7%、6%、5%、4%、3%、2%、1%、0.5%、0.1%、0.01%或0.001%)之效率封裝至蛋白質外部中。在一些實施例中，封裝缺失型遺傳元件即使在容許封裝野生型指環病毒(例如如本文所描述)之遺傳元件的因子(例如ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3或ORF2t/3)存在下亦無法封裝至蛋白質外部中。在一些實施例中，相比於野生型指環病毒(例如如本文所述)，封裝缺陷型遺傳元件以小於10% (例如小於10%、9%、8%、7%、6%、5%、4%、3%、2%、1%、0.5%、0.1%、0.01%或0.001%)之效率封裝至蛋白質外部中，即使在容許封裝野生型指環病毒(例如如本文所述)之遺傳元件的因子(例如ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3或ORF2t/3)存在下亦如此。

【0128】 在一些實施例中，封裝勝任型遺傳元件可封裝至蛋白質外部(例如其中蛋白質外部包含蛋白殼或其部分，例如包含由ORF1核酸編碼之多肽，例如如本文所描述)中。在一些實施例中，相比於野生型指環病

毒(例如如本文所述)，封裝勝任型遺傳元件以至少20% (例如至少20%、30%、40%、50%、60%、70%、80%、85%、90%、95%、96%、97%、98%、99%、100%或更高)的效率封裝至蛋白質外部中。在一些實施例中，封裝勝任型遺傳元件在容許封裝野生型指環病毒(例如如本文所描述)之遺傳元件的因子(例如ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3或ORF2t/3)存在下可封裝至蛋白質外部中。在一些實施例中，在容許封裝野生型指環病毒(例如如本文所述如)之遺傳元件的因子(例如ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3或ORF2t/3)存　在下，相比於野生型指環病毒(例如如本文所述)，封裝勝任型遺傳元件以至少20% (例如至少20%、30%、40%、50%、60%、70%、80%、85%、90%、95%、96%、97%、98%、99%、100%或更高)的效率封裝至蛋白質外部中。

【0129】

***指環病毒***

在一些實施例中，例如如本文所描述之指環載體包含衍生自指環病毒之序列或表現產物。在一些實施例中，指環載體包括相對於指環病毒為外源性的一或多種序列或表現產物。在一些實施例中，指環載體包括相對於指環病毒為內源性的一或多種序列或表現產物。在一些實施例中，指環載體包括相對於指環載體中之一或多種其他序列或表現產物為異源的一或多種序列或表現產物。指環病毒一般具有帶負極性之單股環狀DNA基因體。

【0130】 在一些實施例中，遺傳元件包含編碼胺基酸序列或其功能片段或與本文所描述之胺基酸序列(例如指環病毒胺基酸序列)中之任一者

具有至少約60%、70%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性之序列的核苷酸序列。

【0131】 在一些實施例中，如本文所述之指環載體包含一或多個核酸分子(例如如本文所述之遺傳元件)，該一或多個核酸分子包含與例如如本文所述之指環病毒序列或其片段具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的序列。

【0132】 在一些實施例中，如本文所述之指環載體包含一或多個核酸分子(例如如本文所述之遺傳元件)，其包含與以下中之一或多者具有至少 約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的序列：指環病毒之TATA盒、加帽位點、起始元件、轉錄起始位點、5' UTR保守域、ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3、ORF2t/3、三個開放閱讀框架區、聚(A)信號、富含GC之區域或其任何組合，例如如本文所述。在一些實施例中，核酸分子包含編碼蛋白殼蛋白之序列，該蛋白殼蛋白例如本文所描述之任一指環病毒 的ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3、ORF2t/3序列。在實施例中，核酸分子包含編碼蛋白殼蛋白之序列，該蛋白殼蛋白包含與指環病毒ORF1蛋白(或其剪接變異體或功能片段)具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的胺基酸序列或由指環病毒ORF1核酸編碼之多肽。

【0133】 在實施例中，核酸分子包含與表A1之指環病毒ORF1核酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表A1之指環病毒ORF1/1核苷酸序列具有至少約70%、75%、80%、

85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表A1之指環病毒ORF1/2核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表A1之指環病毒ORF2核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表A1之指環病毒ORF2/2核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表A1之指環病毒ORF2/3核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表A1之指環病毒ORF2t/3核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表A1之指環病毒TATA盒核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性之核酸序列。在實施例中，核酸分子包含與表A1之指環病毒起始元件核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性之核酸序列。在實施例中，核酸分子包含與表A1之指環病毒轉錄起始位點核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性之核酸序列。在實施例中，核酸分子包含與表A1之指環病毒5' UTR保守域核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一

致性之核酸序列。在實施例中，核酸分子包含與表A1之指環病毒三個開放閱讀框架區核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性之核酸序列。在實施例中，核酸分子包含與表A1之指環病毒聚(A)信號核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性之核酸序列。在實施例中，核酸分子包含與表A1之指環病毒富含GC之核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性之核酸序列。

【0134】在實施例中，核酸分子包含與表B1之指環病毒ORF1核酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性之核酸序列。在實施例中，核酸分子包含與表B1之指環病毒ORF1/1核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性之核酸序列。在實施例中，核酸分子包含與表B1之指環病毒ORF1/2核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表B1之指環病毒ORF2核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表B1之指環病毒ORF2/2核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表B1之指環病毒ORF2/3核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核

酸分子包含與表B1之指環病毒TATA盒核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性之核酸序列。在實施例中，核酸分子包含與表B1之指環病毒起始元件核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性之核酸序列。在實施例中，核酸分子包含與表B1之指環病毒轉錄起始位點核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性之核酸序列。在實施例中，核酸分子包含與表B1之指環病毒5' UTR保守域核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性之核酸序列。在實施例中，核酸分子包含與表B1之指環病毒三個開放閱讀框架區核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性之核酸序列。在實施例中，核酸分子包含與表B1之指環病毒聚(A)信號核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性之核酸序列。在實施例中，核酸分子包含與表B1之指環病毒富含GC之核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性之核酸序列。

【0135】在實施例中，核酸分子包含與表C1之指環病毒ORF1核酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表C1之指環病毒ORF1/1核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性之核酸序

列。在實施例中，核酸分子包含與表C1之指環病毒ORF1/2核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表C1之指環病毒ORF2核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表C1之指環病毒ORF2/2核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表C1之指環病毒ORF2/3核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表C1之指環病毒TAIP核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表C1之指環病毒TATA盒核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性之核酸序列。在實施例中，核酸分子包含與表C1之指環病毒起始元件核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性之核酸序列。在實施例中，核酸分子包含與表C1之指環病毒轉錄起始位點核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性之核酸序列。在實施例中，核酸分子包含與表C1之指環病毒5' UTR保守域核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性之核酸序列。在實施例中，核酸分子包含與表C1之指環病毒三個開放閱讀

框架區核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性之核酸序列。在實施例中，核酸分子包含與表C1之指環病毒聚(A)信號核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性之核酸序列。在實施例中，核酸分子包含與表C1之指環病毒富含GC之核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性之核酸序列。

【0136】在實施例中，核酸分子包含與表D1之指環病毒ORF1核酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表D1之指環病毒ORF1/1核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表D1之指環病毒ORF1/2核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表D1之指環病毒ORF2核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表D1之指環病毒ORF2/2核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表D1之指環病毒ORF2/3核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表D1之指環病毒TAIP核苷酸序列具有至少約70%、75%、

80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表D1之指環病毒TATA盒核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表D1之指環病毒起始元件核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表D1之指環病毒轉錄起始位點核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表D1之指環病毒5' UTR保守域核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表D1之指環病毒三個開放閱讀框架區核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表D1之指環病毒聚(A)信號核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表D1之指環病毒富含GC之核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。

【0137】在實施例中，核酸分子包含與表E1之指環病毒ORF1核酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表E1之指環病毒ORF1/1核苷酸序列具有至少約70%、75%、80%、

85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表E1之指環病毒ORF1/2核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表E1之指環病毒ORF2核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表E1之指環病毒ORF2/2核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表E1之指環病毒ORF2/3核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表E1之指環病毒TAIP核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表E1之指環病毒TATA盒核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性之核酸序列。在實施例中，核酸分子包含與表E1之指環病毒起始元件核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性之核酸序列。在實施例中，核酸分子包含與表E1之指環病毒轉錄起始位點核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表E1之指環病毒5' UTR保守域核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的

核酸序列。在實施例中，核酸分子包含與表E1之指環病毒三個開放閱讀框架區核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表E1之指環病毒聚(A)信號核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。在實施例中，核酸分子包含與表E1之指環病毒富含GC之核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核酸序列。

【0138】 在一些實施例中，遺傳元件包含編碼胺基酸序列或其功能片段或與本文所描述之胺基酸序列(例如指環病毒胺基酸序列)中之任一者具有至少約60%、70%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性之序列的核苷酸序列。

【0139】 在一些實施例中，如本文所述之指環載體包含一或多個核酸分子(例如如本文所述之遺傳元件)，該一或多個核酸分子包含與例如如本文所述之指環病毒序列或其片段具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的序列。在實施例中，指環載體包含選自以下之核酸序列：表A1-M2中之任一者中所示之序列，或與其具有至少70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的序列。在實施例中，指環載體包含多肽，該多肽包含如表A2-M2中之任一者中所示之序列，或與其具有至少70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的序列。

【0140】 在一些實施例中，如本文所述之指環載體包含一或多個核

酸分子(例如如本文所述之遺傳元件)，其包含與以下中之一或多者具有至少　約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的序列：本文所述之任一指環病毒(例如如所註釋，或如由表A-M中之任一者中所列之序列編碼的指環病毒序列)之TATA盒、加帽位點、起始元件、轉錄起始位點、5' UTR保守域、ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3、ORF2t/3、三個開放閱讀框架區、聚(A)信號、富含GC之區域或其任何組合。在一些實施例中，核酸分子包含編碼蛋白殼蛋白之序列，例如本文所述之任一指環病毒(例如如表A-M中之任一者中所註釋或由所列之序列編碼的指環病毒序列)之ORF1、ORF1/1、ORF1/2，ORF2、ORF2/2、ORF2/3、ORF2t/3序列。在實施例中，核酸分子包含編碼蛋白殼蛋白之序列，該蛋白殼蛋白包含與指環病毒ORF1或ORF2蛋白(例如表A2-M2中之任一者中所示之ORF1或ORF2胺基酸序列，或由表A1-M1中之任一者中所示之核酸序列編碼的ORF1或ORF2胺基酸序列)具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的胺基酸序列。在實施例中，核酸分子包含編碼蛋白殼蛋白之序列，該蛋白殼蛋白包含與指環病毒ORF1蛋白(例如表A2-M2中之任一者中所示之ORF1胺基酸序列，或由表A1-M1中之任一者中所示之核酸序列編碼的ORF1胺基酸序列)具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的胺基酸序列。

【0141】 在一些實施例中，如本文所描述之指環載體為嵌合指環載體。在一些實施例中，嵌合指環載體進一步包含一或多種來自除指環病毒以外之病毒的元件、多肽或核酸。

【0142】 在實施例中，嵌合指環載體包含複數個多肽(例如指環病毒ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3及/或ORF2t/3)，該等多肽包含來自複數個不同指環病毒(例如如本文所描述)之序列。例如，嵌合指環載體可包含來自一種指環病毒之ORF1分子(例如環1 ORF1分子，或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%胺基酸序列一致性的ORF1分子)，及來自不同指環病毒之ORF2分子(例如環2 ORF2分子，或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%胺基酸序列一致性的ORF2分子)。在另一實例中，嵌合指環載體可包含來自一種指環病毒之第一ORF1分子(例如環1 ORF1分子，或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%胺基酸序列一致性的ORF1分子)，及來自不同指環病毒之第二ORF1分子(例如環2 ORF1分子，或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%胺基酸序列一致性的ORF1分子)。

【0143】 在一些實施例中，指環載體包含嵌合多肽(例如指環病毒ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3及/或ORF2t/3)，例如包含來自指環病毒(例如如本文所描述)之至少一部分及來自不同病毒(例如如本文所描述)之至少一部分。

【0144】 在一些實施例中，指環載體包含嵌合多肽(例如指環病毒ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3及/或ORF2t/3)，例如包含來自一種指環病毒(例如如本文所描述)之至少一部分及來自不同指環病毒(例如如本文所描述)之至少一部分。在實施例中，指環載體包含嵌合ORF1分子，該分子包含來自一種指環病毒(例如如本文所述)之ORF1

分子，或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%胺基酸序列一致性之ORF1分子的至少一部分，及來自不同指環病毒(例如如本文所述)之ORF1分子，或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%胺基酸序列一致性之ORF1分子的至少一部分。在實施例中，嵌合ORF1分子包含來自一種指環病毒之ORF1果醬捲域，或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%序列一致性之序列，及來自不同指環病毒之ORF1胺基酸子序列(例如如本文所述)，或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%序列一致性之序列。在實施例中，嵌合ORF1分子包含來自一種指環病毒之ORF1富含精胺酸之區域，或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%序列一致性的序列，及來自不同指環病毒之ORF1胺基酸子序列(例如如本文所述)，或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%序列一致性的序列。在實施例中，嵌合ORF1分子包含來自一種指環病毒之ORF1高變域，或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%序列一致性之序列，及來自不同指環病毒之ORF1胺基酸子序列(例如如本文所述)，或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%序列一致性之序列。在實施例中，嵌合ORF1分子包含來自一種指環病毒之ORF1 N22域，或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%序列一致性之序列，及來自不同指環病毒之ORF1胺基酸子序列(例如如本文所述)，或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%序列一致性之序列。在實施例中，嵌合ORF1分子包含來自一種指

環病毒之ORF1 C端域，或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%序列一致性之序列，及來自不同指環病毒之ORF1胺基酸子序列(例如如本文所述)，或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%序列一致性之序列。

【0145】在實施例中，指環載體包含嵌合ORF1/1分子，該分子包含來自一種指環病毒(例如如本文所述)之ORF1/1分子，或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%胺基酸序列一致性之ORF1/1分子的至少一部分，及來自不同指環病毒(例如如本文所述)之ORF1/1分子，或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%胺基酸序列一致性之ORF1/1分子的至少一部分。在實施例中，指環載體包含嵌合ORF1/2分子，該分子包含來自一種指環病毒(例如如本文所述)之ORF1/2分子，或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%胺基酸序列一致性之ORF1/2分子的至少一部分，及來自不同指環病毒(例如如本文所述)之ORF1/2分子，或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%胺基酸序列一致性之ORF1/2分子的至少一部分。在實施例中，指環載體包含嵌合ORF2分子，該分子包含來自一種指環病毒(例如如本文所述)之ORF2分子，或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%胺基酸序列一致性之ORF2分子的至少一部分，及來自不同指環病毒(例如如本文所述)之ORF2分子，或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%胺基酸序列一致性之ORF2分子的至少一部分。在實施例中，指環載體包含嵌合ORF2/2分子，該分子包含來自一種指環病毒(例如如本文所述)之ORF2/2分子，

或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%胺基酸序列一致性之ORF2/2分子的至少一部分，及來自不同指環病毒(例如如本文所述)之ORF2/2分子，或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%胺基酸序列一致性之ORF2/2分子的至少一部分。在實施例中，指環載體包含嵌合ORF2/3分子，該分子包含來自一種指環病毒(例如如本文所述)之ORF2/3分子，或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%胺基酸序列一致性之ORF2/3分子的至少一部分，及來自不同指環病毒(例如如本文所述)之ORF2/3分子，或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%胺基酸序列一致性之ORF2/3分子的至少一部分。在實施例中，指環載體包含嵌合ORF2T/3分子，該分子包含來自一種指環病毒(例如如本文所述)之ORF2T/3分子，或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%胺基酸序列一致性之ORF2T/3分子的至少一部分，及來自不同指環病毒(例如如本文所述)之ORF2T/3分子，或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%胺基酸序列一致性之ORF2T/3分子的至少一部分。

【0146】其中包含之序列或子序列可用於本文所述之組合物及方法中的額外例示性指環病毒基因體描述於例如PCT申請案第PCT/US2018/037379號及第PCT/US19/65995號(以全文引用之方式併入本文中)中。在一些實施例中，例示性指環病毒序列包含如以引用之方式併入本文中之PCT/US19/65995之表A1、A3、A5、A7、A9、A11、B1-B5、1、3、5、7、9、11、13、15或17中之任一者中所列的核酸序列。在一些實施例中，例示性指環病毒序列包含以引用之方式併入本文中之

PCT/US19/65995之　表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16或18中之任一者中所列的胺基酸序列。在一些實施例中，例示性指環病毒序列包含ORF1分子序列或編碼其之核酸序列，例如如以引用之方式併入本文中之PCT/US19/65995之表21、23、25、27、29、31、33、35、D2、D4、D6、D8、D10或37A-37C中之任一者中所列。

## 表A1.例示性指環病毒核酸序列(α細環病毒屬，分枝系3)

| | |
|---|---|
| 名稱 | 環1 |
| 屬/分枝系 | α細環病毒屬，分枝系3 |
| 寄存編號 | AJ620231.1 |

**全序列：**3753 bp

```
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
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)

**註釋：**

| *假定域* | *鹼基範圍* |
|---|---|
| TATA盒 | 83 - 88 |
| 加帽位點 | 104 - 111 |
| 轉錄起始位點 | 111 |
| 5' UTR保守域 | 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 |
| 三個開放閱讀框架區 | 2551 - 2786 |
| 聚(A)信號 | 3011 - 3016 |
| 富含GC之區域 | 3632 - 3753 |

**表A2.例示性指環病毒胺基酸序列(α細環病毒屬，分枝系3)**

| 環1 (α細環病毒屬分枝系3) | |
|---|---|
| ORF2 | MSFWKPPVHNVTGIQRMWYESFHRGHASFCGCGNPILHITALAETYGHPTG PRPSGPPGVDPNPHIRRARPAPAAPEPSQVDSRPALTWHGDGGSDGGAGGSG SGGPVADFADDGLDQLVAALDDEE (SEQ ID NO: 17) |
| ORF2/2 | MSFWKPPVHNVTGIQRMWYESFHRGHASFCGCGNPILHITALAETYGHPTG PRPSGPPGVDPNPHIRRARPAPAAPEPSQVDSRPALTWHGDGGSDGGAGGSG SGGPVADFADDGLDQLVAALDDEELLKTPASSPPMKYPVPVTSLEEYKSSTR GSWDRTTRSGHGTCADTHLAEQVLRECQNNKKLLTLYSQAQKSLGSTSQNK KPKKKAHIHSKENRDRGRPRKKARQKPSRKRAKRSPSNSSCSSSTKSSSSSD RESKSSSSSS (SEQ ID NO: 18) |
| ORF2/3 | MSFWKPPVHNVTGIQRMWYESFHRGHASFCGCGNPILHITALAETYGHPTG PRPSGPPGVDPNPHIRRARPAPAAPEPSQVDSRPALTWHGDGGSDGGAGGSG SGGPVADFADDGLDQLVAALDDEEPKKASGRHPKTRNPRRKLTFTPKRIETV GDRGRKRDRSPLAREPRGPLPTAVAAAVPRAAQAQTGNQSPLRAAHKDPTR GPCKPMPTVGPRQWLFPERKPAPAPSSGDWAMEFLAAKIFDRPVRSNLKDTP YYPYVKNQYNVYFDLKFE (SEQ ID NO: 19) |
| ORF2t/3 | MSFWKPPVHNVTGIQRMWPKKASGRHPKTRNPRRKLTFTPKRIETVGDRGR KRDRSPLAREPRGPLPTAVAAAVPRAAQAQTGNQSPLRAAHKDPTRGPCKP MPTVGPRQWLFPERKPAPAPSSGDWAMEFLAAKIFDRPVRSNLKDTPYYPY VKNQYNVYFDLKFE (SEQ ID NO: 20) |
| ORF1 | MAWGWWKRRRRWWFRKRWTRGRLRRRWPRSARRRPRRRRVRRRRRWRR GRRKTRTYRRRRRFRRRGRKAKLIIKLWQPAVIKRCRIKGYIPLIISGNGTFAT NFTSHINDRIMKGPFGGGHSTMRFSLYILFEEHLRHMNFWTRSNDNLELTRY LGASVKIYRHPDQDFIVIYNRRTPLGGNIYTAPSLHPGNAILAKHKILVPSLQT RPKGRKAIRLRIAPPTLFTDKWYFQKDIADLTLFNIMAVEADLRFPFCSPQTD NTCISFQVLSSVYNNYLSINTFNNDNSDSKLKEFLNKAFPTTGTKGTSLNALN TFRTEGCISHPQLKKPNPQINKPLESQYFAPLDALWGDPIYYNDLNENKSLND IIEKILIKNMITYHAKLREFPNSYQGNKAFCHLTGIYSPPYLNQGRISPEIFGLY TEIIYNPYTDKGTGNKVWMDPLTKENNIYKEGQSKCLLTDMPLWTLLFGYT DWCKKDTNNWDLPLNYRLVLICPYTFPKLYNEKVKDYGYIPYSYKFGAGQ MPDGSNYIPFQFRAKWYPTVLHQQQVMEDISRSGPFAPKVEKPSTQLVMKY CFNFNWGGNPIIEQIVKDPSFQPTYEIPGTGNIPRRIQVIDPRVLGPHYSFRSW DMRRHTFSRASIKRVSEQQETSDLVFSGPKKPRVDIPKQETQEESSHSLQRES RPWETEEESETEALSQESQEVPFQQQLQQQYQEQLKLRQGIKVLFEQLIRTQ QGVHVNPCLR (SEQ ID NO: 21) |
| ORF1/1 | MAWGWWKRRRRWWFRKRWTRGRLRRRWPRSARRRPRRRRIVKDPSFQPT YEIPGTGNIPRRIQVIDPRVLGPHYSFRSWDMRRHTFSRASIKRVSEQQETSDL VFSGPKKPRVDIPKQETQEESSHSLQRESRPWETEEESETEALSQESQEVPFQ QQLQQQYQEQLKLRQGIKVLFEQLIRTQQGVHVNPCLR (SEQ ID NO: 22) |
| ORF1/2 | MAWGWWKRRRRWWFRKRWTRGRLRRRWPRSARRRPRRRRAQKSLGSTS QNKKPKKKAHIHSKENRDRGRPRKKARQKPSRKRAKRSPSNSSCSSSTKSSS SSDRESKSSSSSS (SEQ ID NO: 23) |

## 表B1.例示性指環病毒核酸序列(β細環病毒屬)

| | |
|---|---|
| 名稱 | 環2 |
| 屬/分枝系 | β細環病毒屬 |
| 寄存編號 | JX134045.1 |

**全序列：**2797 bp

```
1         10         20         30         40         50
|         |          |          |          |          |
TAATAAATATTCAACAGGAAAACCACCTAATTTAAATTGCCGACCACAAA
CCGTCACTTAGTTCCCCTTTTTGCAACAACTTCTGCTTTTTTCCAACTGC
CGGAAAACCACATAATTTGCATGGCTAACCACAAACTGATATGCTAATTA
ACTTCCACAAAACAACTTCCCCTTTTAAAACCACACCTACAAATTAATTA
TTAAACACAGTCACATCCTGGGAGGTACTACCACACTATAATACCAAGTG
CACTTCCGAATGGCTGAGTTTATGCCGCTAGACGGAGAACGCATCAGTTA
CTGACTGCGGACTGAACTTGGGCGGGTGCCGAAGGTGAGTGAAACCACCG
AAGTCAAGGGGCAATTCGGGCTAGTTCAGTCTAGCGGAACGGGCAAGAAA
CTTAAAATTATTTTATTTTTCAGATGAGCGACTGCTTTAAACCAACATGC
TACAACAACAAAACAAAGCAAACTCACTGGATTAATAACCTGCATTTAAC
CCACGACCTGATCTGCTTCTGCCCAACACCAACTAGACACTTATTACTAG
CTTTAGCAGAACAACAAGAAACAATTGAAGTGTCTAAACAAGAAAAAGAA
AAAATAACAAGATGCCTTATTACTACAGAAGAAGACGGTACAACTACAGA
CGTCCTAGATGGTATGGACGAGGTTGGATTAGACGCCCTTTTCGCAGAAG
ATTTCGAAGAAAAAGAAGGGTAAGACCTACTTATACTACTATTCCTCTAA
AGCAATGGCAACCGCCATATAAAAGAACATGCTATATAAAAGGACAAGAC
TGTTTAATATACTATAGCAACTTAAGACTGGGAATGAATAGTACAATGTA
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
GAATTAGCCAAGACTGGGCACTTAAAGACACTGTTTCTAAAATTACAGAG
CCAGATCGACAGCAACTGCTCAAACAAGCCCTCGAATGCCTGCAAATCTC
GGAAGAAACGCAGGAGAAAAAAGAAAAAGAAGTACAGCAGCTCATCAGCA
ACCTCAGACAGCAGCAGCAGCTGTACAGAGAGCGAATAATATCATTATTA
AAGGACCAATAACTTTTAACTGTGTAAAAAAGGTGAAATTGTTTGATGAT
AAACCAAAAAACCGTAGATTTACACCTGAGGAATTTGAAACTGAGTTACA
AATAGCAAAATGGTTAAAGAGACCCCCAAGATCCTTTGTAAATGATCCTC
CCTTTTACCCATGGTTACCACCTGAACCTGTTGTAAACTTTAAGCTTAAT
TTTACTGAATAAAGGCCAGCATTAATTCACTTAAGGAGTCTGTTTATTTA
AGTTAAACCTTAATAAACGGTCACCGCCTCCCTAATACGCAGGCGCAGAA
AGGGGGCTCCGCCCCCTTTAACCCCCAGGGGGCTCCGCCCCCTGAAACCC
CCAAGGGGGCTACGCCCCCTTACACCCCC (SEQ ID NO: 54)

**註釋：**

| *假定域* | *鹼基範圍* |
|---|---|
| TATA盒 | 237- 243 |
| 加帽位點 | 260 - 267 |
| 轉錄起始位點 | 267 |
| 5' UTR保守域 | 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 |
| 三個開放閱讀框架區 | 2441 - 2586 |
| 聚(A)信號 | 2808 - 2813 |
| 富含GC之區域 | 2868 - 2929 |

**表B2.例示性指環病毒胺基酸序列(β細環病毒屬)**

| 環2 (β細環病毒屬) | |
|---|---|
| ORF2 | MSDCFKPTCYNNKTKQTHWINNLHLTHDLICFCPTPTRHLLLALAEQQETIE VSKQEKEKITRCLITTEEDGTTTDVLDGMDEVGLDALFAEDFEEKEG (SEQ ID NO: 55) |
| ORF2/2 | MSDCFKPTCYNNKTKQTHWINNLHLTHDLICFCPTPTRHLLLALAEQQETIE VSKQEKEKITRCLITTEEDGTTTDVLDGMDEVGLDALFAEDFEEKEGFNIPYP VTSMKQLRYRVQGKPQNPSYTPSTIDTGTTQQQLCHELAKTGHLKTLFLKL QSQIDSNCSNKPSNACKSRKKRRRKKKKKYSSSSATSDSSSSCTESE (SEQ ID NO: 56) |
| ORF2/3 | MSDCFKPTCYNNKTKQTHWINNLHLTHDLICFCPTPTRHLLLALAEQQETIE VSKQEKEKITRCLITTEEDGTTTDVLDGMDEVGLDALFAEDFEEKEGARSTA TAQTSPRMPANLGRNAGEKRKRSTAAHQQPQTAAAAVQRANNIIIKGPITFN CVKKVKLFDDKPKNRRFTPEEFETELQIAKWLKRPPRSFVNDPPFYPWLPPE PVVNFKLNFTE (SEQ ID NO: 57) |
| ORF1 | MPYYYRRRRYNYRRPRWYGRGWIRRPFRRRFRRKRRVRPTYTTIPLKQWQP PYKRTCYIKGQDCLIYYSNLRLGMNSTMYEKSIVPVHWPGGGSFSVSMLTL DALYDIHKLCRNWWTSTNQDLPLVRYKGCKITFYQSTFTDYIVRIHTELPANS NKLTYPNTHPLMMMMSKYKHIIPSRQTRRKKKPYTKIFVKPPPQFENKWYF ATDLYKIPLLQIHCTACNLQNPFVKPDKLSNNVTLWSLNTISIQNRNMSVDQG QSWPFKILGTQSFYFYFYTGANLPGDTTQIPVADLLPLTNPRINRPGQSLNEA KITDHITFTEYKNKFTNYWGNPFNKHIQEHLDMILYSLKSPEAIKNEWTTEN MKWNQLNNAGTMALTPFNEPIFTQIQYNPDRDTGEDTQLYLLSNATGTGWD PPGIPELILEGFPLWLIYWGFADFQKNLKKVTNIDTNYMLVAKTKFTQKPGTF YLVILNDTFVEGNSPYEKQPLPEDNIKWYPQVQYQLEAQNKLLQTGPFTPNI QGQLSDNISMFYKFYFKWGGSPPKAINVENPAHQIQYPIPRNEHETTSLQSPG EAPESILYSFDYRHGNYTTTALSRISQDWALKDTVSKITEPDRQQLLKQALEC LQISEETQEKKEKEVQQLISNLRQQQQLYRERIISLLKDQ (SEQ ID NO: 58) |
| ORF1/1 | MPYYYRRRRYNYRRPRWYGRGWIRRPFRRRFRRKRRIQYPIPRNEHETTSLQ SPGEAPESILYSFDYRHGNYTTTALSRISQDWALKDTVSKITEPDRQQLLKQA LECLQISEETQEKKEKEVQQLISNLRQQQQLYRERIISLLKDQ (SEQ ID NO: 59) |
| ORF1/2 | MPYYYRRRRYNYRRPRWYGRGWIRRPFRRRFRRKRRSQIDSNCSNKPSNAC KSRKKRRRKKKKKYSSSSATSDSSSSCTESE (SEQ ID NO: 60) |

## 表C1.例示性指環病毒核酸序列(γ細環病毒屬)

名稱　環4

屬/分枝系　γ細環病毒屬

寄存編號

**全序列：**3176 bp

```
1         10         20         30         40         50
|         |          |          |          |          |
TAAAATGGCGGGAGCCAATCATTTTATACTTTCACTTTCCAATTAAAAAT
GGCCACGTCACAAACAAGGGGTGGAGCCATTTAAACTATATAACTAAGTG
GGGTGGCGAATGGCTGAGTTTACCCCGCTAGACGGTGCAGGGACCGGATC
GAGCGCAGCGAGGAGGTCCCCGGCTGCCCATGGGCGGGAGCCGAGGTGAG
TGAAACCACCGAGGTCTAGGGGCAATTCGGGCTAGGGCAGTCTAGCGGAA
CGGGCAAGAAACTTAAAACAATATTTGTTTTACAGATGGTTAGTATATCC
TCAAGTGATTTTTTTAAGAAAACGAAATTTAATGAGGAGACGCAGAACCA
AGTATGGATGTCTCAAATTGCTGACTCTCATGATAATATCTGCAGTTGCT
GGCATCCATTTGCTCACCTTCTTGCTTCCATATTTCCTCCTGGCCACAAA
GATCGTGATCTTACTATTAACCAAATTCTTCTAAGAGATTATAAAGAAAA
ATGCCATTCTGGTGGAGAAGAAGGAGAAAATTCTGGACCAACAACAGGTT
TAATTACACCAAAAGAAGAAGATATAGAAAAAGATGGCCCAGAAGGCGCC
GCAGAAGAAGACCATACAGACGCCCTGTTCGCCGCCGCCGTAGAAAACTT
CGAAAGGTAAAGAGAAAAAAAAAATCTTTAATTGTTAGACAATGGCAACC
AGACAGTATAAGAACTTGTAAAATTATAGGACAGTCAGCTATAGTTGTTG
GGGCTGAAGGAAAGCAAATGTACTGTTATACTGTCAATAAGTTAATTAAT
GTGCCCCCAAAAACACCATATGGGGGAGGCTTTGGAGTAGACCAATACAC
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
CCTGGATTTGAAGAGCAAACAGAGAGAGAATTAGCAATTATATTTCATAG
GCCCCCTAGAACCTACAAAGAGGACCTTCCATTCTATCCCTGGCTACCAC
CTGCACCCCTTGTACAATTTAACCTTAACTTCAAAGGCTAGGCCAACAAT
GTACACTTAGTAAAGCATGTTTATTAAAGCACAACCCCCAAAATAAATGT
AAAAATAAAAAAAAAAAAAAAAAAATAAAAAATTGCAAAAATTCGGCGCT
CGCGCGCATGTGCGCCTCTGGCGCAAATCACGCAACGCTCGCGCGCCCGC
GTATGTCTCTTTACCACGCACCTAGATTGGGGTGCGCGCGCTAGCGCGCG
CACCCCAATGCGCCCCGCCCTCGTTCCGACCCGCTTGCGCGGGTCGGACC
ACTTCGGGCTCGGGGGGGCGCGCCTGCGGCGCTTTTTTACTAAACAGACT
CCGAGCCGCCATTTGGCCCCCTAAGCTCCGCCCCCCTCATGAATATTCAT
AAAGGAAACCACATAATTAGAATTGCCGACCACAAACTGCCATATGCTAA
TTAGTTCCCCTTTTACAAAGTAAAAGGGGAAGTGAACATAGCCCCACACC
CGCAGGGGCAAGGCCCCGCACCCCTACGTCACTAACCACGCCCCCGCCGC
CATCTTGGGTGCGGCAGGGCGGGGGC (SEQ ID NO: 886)

**註釋：**

| *假定域* | *鹼基範圍* |
|---|---|
| TATA盒 | 87- 93 |
| 加帽位點 | 110 - 117 |
| 轉錄起始位點 | 117 |
| 5' UTR保守域 | 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 |
| 三個開放閱讀框架區 | 2209 - 2439 |
| 聚(A)信號 | 2672 - 2678 |
| 富含GC之區域 | 3076 - 3176 |

## 表C2.例示性指環病毒胺基酸序列(γ細環病毒屬)

| 環4 (γ細環病毒屬) | |
|---|---|
| ORF2 | MVSISSSDFFKKTKFNEETQNQVWMSQIADSHDNICSCWHPFAHLLASIFPPG<br>HKDRDLTINQILLR<br>DYKEKCHSGGEEGENSGPTTGLITPKEEDIEKDGPEGAAEEDHTDALFAAAV<br>ENFER (SEQ ID NO: 887) |
| ORF2/2 | MVSISSSDFFKKTKFNEETQNQVWMSQIADSHDNICSCWHPFAHLLASIFPPG<br>HKDRDLTINQILLRDYKEKCHSGGEEGENSGPTTGLITPKEEDIEKDGPEGAA<br>EEDHTDALFAAAVENFESGVDHNSMNQKLLTLANKSSMMSPILSTKQYKLKI<br>QKDKTPDLSSMIGTTDEALLKKDLLKECQLTSQLIQISKQLQRKTFPKRKREL<br>DPNSQSHNKKKRRHCHVSSLSAKKIPSKKQRHKKTSSSSSSSSRSSSSSSRETS<br>SSSSTN (SEQ ID NO: 888) |
| ORF2/3 | MVSISSSDFFKKTKFNEETQNQVWMSQIADSHDNICSCWHPFAHLLASIFPPG<br>HKDRDLTINQILLRDYKEKCHSGGEEGENSGPTTGLITPKEEDIEKDGPEGAA<br>EEDHTDALFAAAVENFERSASNFRGRHSQKEKENWTPTHSPTTKRRGDTVM<br>SPLSLQKRYLPRNRDTRRPPAAHQAAAGAAAPPQEKHPPAHPQTKRESTNAS<br>ASHRHVTLTRFKPGFEEQTERELAIIFHRPPRTYKEDLPFYPWLPPAPLVQFNL<br>NFKG (SEQ ID NO: 889) |
| TAIP | MRRRRTKYGCLKLLTLMIISAVAGIHLLTFLLPYFLLATKIVILLLTKFF (SEQ<br>ID NO: 890) |
| ORF1 | MPFWWRRRRKFWTNNRFNYTKRRRYRKRWPRRRRRRRPYRRPVRRRRRK<br>LRKVKRKKKSLIVRQWQPDSIRTCKIIGQSAIVVGAEGKQMYCYTVNKLINV<br>PPKTPYGGGFGVDQYTLKYLYEEYRFAQNIWTQSNVLKDLCRYINVKLIFYR<br>DNKTDFVLSYDRNPPFQLTKFTYPGAHPQQIMLQKHHKFILSQMTKPNGRLT<br>KKLKIKPPKQMLSKWFFSKQFCKYPLLSLKASALDLRHSYLGCCNENPQVFF<br>YYLNHGYYTITNWGAQSSTAYRPNSKVTDTTYYRYKNDRKNINIKSHEYEK<br>SISYENGYFQSSFLQTQCIYTSERGEACIAEKPLGIAIYNPVKDNGDGNMIYL<br>VSTLANTWDQPPKDSAILIQGVPIWLGLFGYLDYCRQIKADKTWLDSHVLVI<br>QSPAIFTYPNPGAGKWYCPLSQSFINGNGPFNQPPTLLQKAKWFPQIQYQQEI<br>INSFVESGPFVPKYANQTESNWELKYKYVFTFKWGGPQFHEPEIADPSKQEQ<br>YDVPDTFYQTIQIEDPEGQDPRSLIHDWDYRRGFIKERSLKRMSTYFSTHTD<br>QQATSEEDIPKKKKRIGPQLTVPQQKEEETLSCLLSLCKKDTFQETETQEDLQ<br>QLIKQQQEQQLLLKRNILQLIHKLKENQQMLQLHTGMLP (SEQ ID NO: 891) |
| ORF1/1 | MPFWWRRRRKFWTNNRFNYTKRRRYRKRWPRRRRRRRPYRRPVRRRRRK<br>LRKWGGPQFHEPEIADPSKQEQYDVPDTFYQTIQIEDPEGQDPRSLIHDWDY<br>RRGFIKERSLKRMSTYFSTHTDQQATSEEDIPKKKKRIGPQLTVPQQKEEETL<br>SCLLSLCKKDTFQETETQEDLQQLIKQQQEQQLLLKRNILQLIHKLKENQQM<br>LQLHTGMLP (SEQ ID NO: 892) |
| ORF1/2 | MPFWWRRRRKFWTNNRFNYTKRRRYRKRWPRRRRRRRPYRRPVRRRRRK<br>LRKISKQLQRKTFPKRKR<br>ELDPNSQSHNKKKRRHCHVSSLSAKKIPSKKQRHKKTSSSSSSSSRSSSSSSR<br>ETSSSSTN (SEQ ID NO: 893) |

## 表D1.例示性指環病毒核酸序列(β細環病毒屬)

| | |
|---|---|
| 名稱 | 環9 |
| 屬/分枝系 | β細環病毒屬 |
| 寄存編號 | MH649263.1 |

**全序列：**2845 bp

```
1         10         20         30         40         50
|         |          |          |          |          |
TTATTAATATTCAACAGGAAAACCACCTAATTTAAATTGCCGACCACAAA
CCGTCACTAACTTCCTTATTTAACATTACTTCCCTTTTAACCAATGAATA
TTCATACAACACATCACACTTCCTGGGAGGAGACATAAAACTATATAACT
AACTACACAGACGAATGGCTGAGTTTATGCCGCTAGACGGAGGACGCACA
GCTACTGCTGCGACCTGAACTTGGGCGGGTGCCGAAGGTGAGTGTAACCA
CCGTAGTCAAGGGGCAATTCGGGCTAGTTCAGTCTAGCGGAACGGGCAAG
ATTATTAATACAAACTTATTTTTACAGATGAGCAAACAACTAAAACCAAC
TTTATACAAAGACAAATCATTGGAATTACAATGGCTAAACAACATTTTTA
GCTCTCACGACCTGTGCTGCGGCTGCAACGATCCAGTTTTACATTTACTG
ATTTTAATTAACAAAACCGGAGAAGCACCTAAACCAGAAGAAGACATTAA
AAATATAAAATGCCTCCTTACTGGCGCCAAAAATACTACCGAAGAAGATA
TAGACCTTTCTCCTGGAGAACTAGAAGAATTATTCAAAGAAGAAAAAGAT
GGAGATACCGCAAACCAAGAAAAACATACTGGAGAAGAAAACTGCGGGTA
AGAAAACGTTTTTATAAAAGAAAGTTAAAAAAAATTGTACTTAAACAGTT
TCAACCAAAAATTATTAGAAGATGTACAATATTTGGAACAATCTGCCTAT
TTCAAGGCTCTCCAGAAAGAGCCAACAATAATTATATTCAAACAATCTAC
TCCTACGTACCAGATAAAGAACCAGGAGGAGGGGGATGGACTTTAATAAC
TGAAAGCTTAAGTAGTTTATGGGAAGACTGGGAACATTTAAAAAATGTAT
GGACTCAAAGTAACGCTGGTTTACCACTTGTAAGATACGGGGGAGTAACA
TTATACTTTTATCAATCTGCCTATACTGACTATATTGCTCAAGTTTTCAA
CTGTTATCCTATGACAGACACAAAATACACACATGCAGACTCAGCACCAA
ACAGAATGTTATTAAAAAAACATGTAATAAGAGTACCTAGCAGAGAAACA
CGCAAAAAAAGAAAGCCATACAAAAGAGTTAGAGTAGGACCTCCTTCTCA
AATGCAAAACAAATGGTACTTTCAAAGAGACATATGTGAAATACCATTAA
TAATGATTGCAGCCACAGCCGTTGACTTTAGATATCCCTTTTGTGCAAGC
GACTGTGCTAGTAACAACTTAACTCTAACATGTTTAAACCCACTATTGTT
TCAAAACCAAGACTTTGACCACCCATCCGATACACAAGGCTACTTTCCAA
AACCTGGAGTATATCTATACTCAACACAAAGAAGTAACAAGCCAAGTTCT
TCAGACTGTATATACTTAGGAAACACAAAAGACAATCAAGAAGGTAAATC
TGCAAGTAGTCTAATGACTCTAAAAACACAAAAAATAACAGATTGGGGAA
ATCCATTTTGGCATTATTATATAGACGGTTCTAAAAAAATATTTTCTTAC
TTTAAACCCCCATCACAATTAGACAGCAGCGACTTTGAACACATGACAGA
ATTAGCAGAACCAATGTTTATACAAGTTAGATACAACCCAGAAAGAGACA
CAGGACAAGGAAACTTAATATACGTAACAGAAAACTTTAGAGGACAACAC
TGGGACCCTCCATCTAGTGACAACCTAAAATTAGATGGATTTCCCTTATA
TGACATGTGCTGGGGTTTCATAGACTGGATAGAAAAAGTTCATGAAACAG
AAAACTTACTTACCAACTACTGCTTCTGTATTAGAAGCAGCGCTTTCAAT
GAAAAAAAAACAGTTTTTATACCTGTAGATCATTCATTTTTAACAGGTTT
TAGCCCATATGAAACTCCAGTTAAATCATCAGACCAAGCTCACTGGCACC
CACAAATAAGATTTCAAACAAAATCAATAAATGACATTTGTTTAACAGGC
CCCGGTTGTGCTAGGTCCCCATATGGCAATTACATGCAGGCAAAAATGAG
TTATAAATTTCATGTAAAATGGGGAGGATGTCCAAAAACTTATGAAAAAC
```

CATATGATCCTTGTTCACAGCCCAATTGGACTATTCCCCATAACCTCAAT
GAAACAATACAAATCCAGAATCCAAACACATGCCCACAAACAGAACTCCA
AGAATGGGACTGGCGACGTGATATTGTTACAAAAAAAGCTATCGAAAGAA
TTAGACAACACACGGAACCTCATGAAACTTTGCAAATCTCTACAGGTTCC
AAACACAACCCACCAGTACACAGACAAACATCACCGTGGACGGACTCAGA
AACGGACTCGGAAGAGGAAAAAGACCAAACACAAGAGATCCAGATCCAGC
TCAACAAGCTCAGAAAGCATCAACAGCATCTCAAGCAGCAGCTCAAGCAG
TACCTGAAACCCCAAAATATAGAATAGTTGCAAGCAACATAAAAGTTGAA
CTTTTTCCTACTAAAAAACCTTTTAAAAACAGACGCTTTACTCCTTCTGA
AAGAGAAACAGAAAGACAATGTGCTAAAGCTTTTTGTAGACCAGAAAGAC
ATTTCTTTTATGATCCTCCTTTTTACCCTTACTGTGTACCTGAACCTATT
GTAAACTTTGCTTTGGGATATAAAATTTAAGGCCAACAAATTTCACTTAG
TGGTGTCTGTTTATTAAAGTTTAACCTTAATAAGCATACTCCGCCTCCCT
ACATTAAGGCGCCAAAAGGGGGCTCCGCCCCCTTAAACCCCAAGGGGGCT
CCGCCCCCTTAAACCCCCAAGGGGGCTCCGCCCCCTTACACCCCC (SEQ ID NO: 1001)

**註釋：**

| *假定域* | *鹼基範圍* |
|---|---|
| TATA盒 | 142 - 148 |
| 起始元件 | 162 - 177 |
| 轉錄起始位點 | 172 |
| 5' UTR保守域 | 226 - 296 |
| ORF2 | 328 - 651 |
| ORF2/2 | 328 - 647; 2121 - 2457 |
| ORF2/3 | 328 - 647; 2296 - 2680 |
| ORF1 | 510 - 2477 |
| ORF1/1 | 510 - 647; 2121 - 2477 |
| ORF1/2 | 510 - 647; 2296 - 2457 |
| 三個開放閱讀框架區 | 2296 - 2454 |
| 富含GC之區域 | 2734 - 2845 |

**表D2.例示性指環病毒胺基酸序列(β細環病毒屬)**

| 環9 (β細環病毒屬) | |
|---|---|
| ORF2 | MSKQLKPTLYKDKSLELQWLNNIFSSHDLCCGCNDPVLHLLILINKTGEAPK PEEDIKNIKCLLTGAKNTTEEDIDLSPGELEELFKEEKDGDTANQEKHTGEEN CG (SEQ ID NO: 1002) |
| ORF2/2 | MSKQLKPTLYKDKSLELQWLNNIFSSHDLCCGCNDPVLHLLILINKTGEAPK PEEDIKNIKCLLTGAKNTTEEDIDLSPGELEELFKEEKDGDTANQEKHTGEEN CGPIGLFPITSMKQYKSRIQTHAHKQNSKNGTGDVILLQKKLSKELDNTRNL MKLCKSLQVPNTTHQYTDKHHRGRTQKRTRKRKKTKHKRSRSSSTSSESINS ISSSSSSST (SEQ ID NO: 1003) |
| ORF2/3 | MSKQLKPTLYKDKSLELQWLNNIFSSHDLCCGCNDPVLHLLILINKTGEAPK PEEDIKNIKCLLTGAKNTTEEDIDLSPGELEELFKEEKDGDTANQEKHTGEEN CGFQTQPTSTQTNITVDGLRNGLGRGKRPNTRDPDPAQQAQKASTASQAAA QAVPETPKYRIVASNIKVELFPTKKPFKNRRFTPSERETERQCAKAFCRPERH FFYDPPFYPYCVPEPIVNFALGYKI (SEQ ID NO: 1004) |
| ORF1 | MPPYWRQKYYRRRYRPFSWRTRRIIQRRKRWRYRKPRKTYWRRKLRVRKR FYKRKLKKIVLKQFQPKIIRRCTIFGTICLFQGSPERANNNYIQTIYSYVPDKE PGGGGWTLITESLSSLWEDWEHLKNVWTQSNAGLPLVRYGGVTLYFYQSAY TDYIAQVFNCYPMTDTKYTHADSAPNRMLLKKHVIRVPSRETRKKRKPYKR VRVGPPSQMQNKWYFQRDICEIPLIMIAATAVDFRYPFCASDCASNNLTLTCL NPLLFQNQDFDHPSDTQGYFPKPGVYLYSTQRSNKPSSSDCIYLGNTKDNQE GKSASSLMTLKTQKITDWGNPFWHYYIDGSKKIFSYFKPPSQLDSSDFEHMT ELAEPMFIQVRYNPERDTGQGNLIYVTENFRGQHWDPPSSDNLKLDGFPLYD MCWGFIDWIEKVHETENLLTNYCFCIRSSAFNEKKTVFIPVDHSFLTGFSPYE TPVKSSDQAHWHPQIRFQTKSINDICLTGPGCARSPYGNYMQAKMSYKFHV KWGGCPKTYEKPYDPCSQPNWTIPHNLNETIQIQNPNTCPQTELQEWDWRR DIVTKKAIERIRQHTEPHETLQISTGSKHNPPVHRQTSPWTDSETDSEEEKDQ TQEIQIQLNKLRKHQQHLKQQLKQYLKPQNIE (SEQ ID NO: 1005) |
| ORF1/1 | MPPYWRQKYYRRRYRPFSWRTRRIIQRRKRWRYRKPRKTYWRRKLRPNWT IPHNLNETIQIQNPNTCPQTELQEWDWRRDIVTKKAIERIRQHTEPHETLQIST GSKHNPPVHRQTSPWTDSETDSEEEKDQTQEIQIQLNKLRKHQQHLKQQLK QYLKPQNIE (SEQ ID NO: 1006) |
| ORF1/2 | MPPYWRQKYYRRRYRPFSWRTRRIIQRRKRWRYRKPRKTYWRRKLRVPNT THQYTDKHHRGRTQKRTRKRKKTKHKRSRSSSTSSESINSISSSSSSST (SEQ ID NO: 1007) |

## 表E1.例示性指環病毒核酸序列(β細環病毒屬)

| | |
|---|---|
| 名稱 | 環10 |
| 屬/分枝系 | β細環病毒屬 |
| 寄存編號 | JX134044.1 |

**全序列：**2912 bp

```
1         10         20         30         40         50
|         |          |          |          |          |
TAATAAATATTCAACAGGAAAACCACCTAATTTAAATTGCCGACCACAAA
CCGTCACTTAGTTCCTCTTTTTCCACAACTTCCTCTTTTACTAATGAATA
TTCATGTAATTAATTAATAATCACCGTAATTCCGGGGAGGAGCCTTTAAA
CTATAAAACTAACTACACATTCGAATGGCTGAGTTTATGCCGCCAGACGG
AGACGGGATCACTTCAGTGACTCCAGGCTGATCAAGGGCGGGTGCCGAAG
GTGAGTGAAACCACCGTAGTCAAGGGGCAATTCGGGCTAGATCAGTCTGG
CGGAACGGGCAAGAAACTTAAAATGTACTTTATTTTACAGAAATGTTCAA
ATCTCCAACATACTTAACAACTAAAGGCAAAAACAATGCCTTAATCAACT
GCTTCGTTGGAGACCACGATCTTCTGTGCAGCTGTAACAATCCTGCCTAC
CATTGCCTCCAAATACTTGCAACTACCTTAGCACCTCAACTAAAACAAGA
AGAAAAACAACAAATAATACAATGCCTTGGTGGTACAGACGCCGTAGCTA
CAACCCGTGGAGACGAAGAAATTGGTTTAGAAGACCTAGAAAAACTATTT
ACAGAAGATACAGAAGAAGACGCCGCTGGGTAAGAAGAAAACCTTTTTAC
AAACGTAAAATTAAGAGACTAAATATAGTAGAATGGCAACCTAAATCAAT
TAGAAAATGTAGAATAAAAGGAATGCTATGCTTGTTTCAAACGACAGAAG
ACAGACTGTCATATAACTTTGATATGTATGAAGAGTCTATTATACCAGAA
AAACTGCCGGGAGGGGGGGGATTTAGCATTAAGAATATAAGCTTATATGC
CTTATACCAAGAACACATACATGCACACAACATATTTACACACACAAACA
CAGACAGACCACTAGCAAGATACACAGGCTGTTCTTTAAAATTCTACCAA
AGCAAAGACATAGACTACGTAGTAACATATTCTACATCACTCCCACTAAG
AAGCTCAATGGGAATGTACAACTCCATGCAACCATCCATACATCTAATGC
AACAAAACAAACTAATTGTACCAAGCAAACAAACACAAAAAAGAAGAAAA
CCATATATTAAAAAACATATATCACCACCAACACAAATGAAATCTCAATG
GTACTTTCAACATAACATTGCAAACATACCGCTACTAATGATAAGAACCA
CAGCATTAACATTAGATAATTACTATATAGGAAGCAGACAATTAAGTACA
AATGTCACTATACATACACTTAACACAACATACATCCAAAACAGAGACTG
GGGAGACAGAAATAAAACTTACTACTGCCAAACATTAGGAACACAAAGAT
ACTTCCTATATGGAACACATTCAACTGCACAAAATATTAATGACATAAAG
CTACAAGAACTAATACCTTTAACAAACACACAAGACTATGTACAAGGCTT
TGATTGGACAGAAAAAGACAAACATAACATAACAACCTACAAAGAATTCT
TAACTAAAGGAGCAGGAAATCCATTTCACGCAGAATGGATAACAGCACAA
AACCCAGTAATACACACAGCAAACAGTCCTACACAAATAGAACAAATATA
CACCGCTTCAACAACAACATTCCAAAACAAAAAACTAACAGACCTACCAA
CGCCAGGATATATATTTATAACTCCAACAGTAAGCTTAAGATACAACCCA
TACAAAGACCTAGCAGAAAGAAACAAATGCTACTTTGTAAGAAGCAAAAT
AAATGCACACGGGTGGGACCCAGAACAACACCAAGAATTAATAAACAGTG
ACCTACCACAATGGTTACTATTATTTGGCTACCCAGACTACATAAAAAGA
ACACAAAACTTTGCATTAGTAGACACAAATTACATACTAGTAGACCACTG
CCCATACACAAATCCAGAAAAAACACCATTTATACCTTTAAGCACATCAT
TTATAGAAGGTAGAAGCCCATACAGTCCTTCAGACACACATGAACCAGAT
GAAGAAGACCAAAACAGGTGGTACCCATGCTACCAATATCAACAAGAATC
AATAAATTCAATATGTCTTAGCGGTCCAGGCACACCAAAAATACCAAAAG
```

GAATAACAGCAGAAGCAAAAGTAAAATATTCCTTTAATTTTAAGTGGGGT
GGTGACCTACCACCAATGTCTACAATTACAAACCCGACAGACCAGCCAAC
ATATGTTGTTCCCAATAACTTCAATGAAACAACTTCGTTACAGAATCCAA
CCACCAGACCAGAGCACTTCTTGTACTCCTTTGACGAAAGGAGGGGACAA
CTTACAGAAAAAGCTACAAAACGCTTGCTTAAAGACTGGGAAACTAAAGA
AACTTCTTTATTGTCTACAGAATACAGATTCGCGGAGCCAACACAAACAC
AAGCCCCACAAGAGGACCCGTCCTCGGAAGAAGAAGAAGAGAGCAACCTC
TTCGAGCGACTCCTCCGACAGCGAACCAAGCAGCTCCAGCTCAAGCGCAG
AATAATACAAACATTGAAAGACCTACAAAAATTAGAATAACTAACAGCAA
AAACACCGTTTACCTATTTCCACCTGAACAAAAGAACAGAAGACTAACAC
CATGGGAAATACAAGAAGACAAAGAAATAGCCAATTTATTTGGCAGACCA
CATAGATACTTTTTAAAAGACATTCCTTTCTATTGGGATATACCCCCAGA
GCCTAAAGTAAACTTTGATTTAAATTTTCAATAAAGAAATAAAGGGCAAG
GCCCCATTAACTCAAAGTCGGTGTCTACCTCTTTAAGTTTAACTTTACTA
AACGGACTCCGCCTCCCTAAATTTGGGCGCCAAAAGGGGGCTCCGCCCCC
TTAAACCCCAGGGGGCTCCGCCCCCTAAAACCCCCAAGGGGGCTACGCCC
CCTTACACCCCC (SEQ ID NO: 1008)

**註釋：**

| *假定域* | *鹼基範圍* |
|---|---|
| TATA盒 | 152 - 158 |
| 起始元件 | 172 - 187 |
| 轉錄起始位點 | 182 |
| 5' UTR保守域 | 239 - 309 |
| ORF2 | 343 - 633 |
| ORF2/2 | 343 - 629; 2196 - 2505 |
| ORF2/3 | 343 - 629; 2371 - 2734 |
| ORF1 | 522 - 2540 |
| ORF1/1 | 522 - 629; 2196 - 2540 |
| ORF1/2 | 522 - 629; 2371 - 2505 |
| 三個開放閱讀框架區 | 2276 - 2502 |
| 富含GC之區域 | 2803 - 2912 |

**表E2.例示性指環病毒胺基酸序列(β細環病毒屬)**

| 環10 (β細環病毒屬) | |
|---|---|
| ORF2 | MFKSPTYLTTKGKNNALINCFVGDHDLLCSCNNPAYHCLQILATTLAPQLKQEEKQQIIQCLGGTDAVATTRGDEEIGLEDLEKLFTEDTEEDAAG (SEQ ID NO: 1009) |
| ORF2/2 | MFKSPTYLTTKGKNNALINCFVGDHDLLCSCNNPAYHCLQILATTLAPQLKQEEKQQIIQCLGGTDAVATTRGDEEIGLEDLEKLFTEDTEEDAAGQHMLFPITSMKQLRYRIQPPDQSTSCTPLTKGGDNLQKKLQNACLKTGKLKKLLYCLQNTDSRSQHKHKPHKRTRPRKKKKRATSSSDSSDSEPSSSSSSAE (SEQ ID NO: 1010) |
| ORF2/3 | MFKSPTYLTTKGKNNALINCFVGDHDLLCSCNNPAYHCLQILATTLAPQLKQEEKQQIIQCLGGTDAVATTRGDEEIGLEDLEKLFTEDTEEDAAGIQIRGANTNTSPTRGPVLGRRRREQPLRATPPTANQAAPAQAQNNTNIERPTKIRITNSKNTVYLFPPEQKNRRLTPWEIQEDKEIANLFGRPHRYFLKDIPFYWDIPPEPKVNFDLNFQ (SEQ ID NO: 1011) |
| ORF1 | MPWWYRRRSYNPWRRRNWFRRPRKTIYRRYRRRRRWVRRKPFYKRKIKRLNIVEWQPKSIRKCRIKGMLCLFQTTEDRLSYNFDMYEESIIPEKLPGGGGFSIKNISLYALYQEHIHAHNIFTHTNTDRPLARYTGCSLKFYQSKDIDYVVTYSTSLPLRSSMGMYNSMQPSIHLMQQNKLIVPSKQTQKRRKPYIKKHISPPTQMKSQWYFQHNIANIPLLMIRTTALTLDNYYIGSRQLSTNVTIHTLNTTYIQNRDWGDRNKTYYCQTLGTQRYFLYGTHSTAQNINDIKLQELIPLTNTQDYVQGFDWTEKDKHNITTYKEFLTKGAGNPFHAEWITAQNPVIHTANSPTQIEQIYTASTTTFQNKKLTDLPTPGYIFITPTVSLRYNPYKDLAERNKCYFVRSKINAHGWDPEQHQELINSDLPQWLLLFGYPDYIKRTQNFALVDTNYILVDHCPYTNPEKTPFIPLSTSFIEGRSPYSPSDTHEPDEEDQNRWYPCYQYQQESINSICLSGPGTPKIPKGITAEAKVKYSFNFKWGGDLPPMSTITNPTDQPTYVVPNNFNETTSLQNPTTRPEHFLYSFDERRGQLTEKATKRLLKDWETKETSLLSTEYRFAEPTQTQAPQEDPSSEEEEESNLFERLLRQRTKQLQLKRRIIQTLKDLQKLE (SEQ ID NO: 1012) |
| ORF1/1 | MPWWYRRRSYNPWRRRNWFRRPRKTIYRRYRRRRRWPTYVVPNNFNETTSLQNPTTRPEHFLYSFDERRGQLTEKATKRLLKDWETKETSLLSTEYRFAEPTQTQAPQEDPSSEEEEESNLFERLLRQRTKQLQLKRRIIQTLKDLQKLE |
| ORF1/2 | MPWWYRRRSYNPWRRRNWFRRPRKTIYRRYRRRRRWNTDSRSQHKHKPHKRTRPRKKKKRATSSSDSSDSEPSSSSSSAE (SEQ ID NO: 1013) |

【0147】在一些實施例中，指環載體包含有包含以全文引用之方式併入本文中之PCT申請案第PCT/US2018/037379號中所列之序列的核酸。在一些實施例中，指環載體包含有包含以全文引用之方式併入本文中之PCT申請案第PCT/US2018/037379號中所列之序列的多肽。在一些實施例中，指環載體包含有包含以全文引用之方式併入本文中之PCT申請案第PCT/US19/65995號中所列之序列的核酸。在一些實施例中，指環載體包含有包含以全文引用之方式併入本文中之PCT申請案第

PCT/US19/65995號中所列之序列的多肽。

【0148】

*ORF1分子*

在一些實施例中，指環載體包含ORF1分子及/或編碼ORF1分子之核酸。一般而言，ORF1分子包含具有指環病毒ORF1蛋白(例如如本文所述之指環病毒ORF1蛋白)之結構特徵及/或活性的多肽。在一些實施例中，ORF1分子包含相對於指環病毒ORF1蛋白(例如如本文所述之指環病毒ORF1蛋白)之截短。ORF1分子可能能夠結合至其他ORF1分子，例如以形成蛋白質外部(例如如本文所述)，例如蛋白殼。在一些實施例中，蛋白質外部可包裹核酸分子(例如，如本文所述之遺傳元件)。在一些實施例中，複數個ORF1分子可形成多聚體，例如以形成蛋白質外部。在一些實施例中，多聚體可為均多聚體。在其他實施例中，多聚體可為雜多聚體。

【0149】 在一些實施例中，ORF1分子可包含以下中之一或多者：包含富含精胺酸之區域的第一區域，例如具有至少60%鹼性殘基(例如至少60%、65%、70%、75%、80%、85%、90%、95%或100%鹼性殘基；例如60%-90%、60%-80%、70%-90%或70-80%之間的鹼性殘基)之區域，及包含果醬捲域之第二區域，例如至少六條β股(例如4、5、6、7、8、9、10、11或12條β股)。

【0150】 在一些實施例中，如本文所述的ORF1分子相對於野生型ORF1蛋白序列(例如如本文所述)包含一或多個離胺酸至組胺酸突變。在某些實施例中，ORF1分子在富含精胺酸之區域及/或第一β股中包含一或多個離胺酸至組胺酸突變。

【0151】

*富含精胺酸之區域*

富含精胺酸之區域與本文所述之富含精胺酸之區域序列或包含至少60%、70%或80%鹼性殘基(例如精胺酸、離胺酸或其組合)之至少約40個胺基酸的序列具有至少70% (例如至少約70%、80%、90%、95%、96%、97%、98%、99%或100%)序列一致性。

【0152】

*果醬捲域*

果醬捲域或區包含多肽(例如包含於較大多肽中之結構域或區域) (例如由其組成)，該多肽包含以下特徵中之一或多者(例如1、2或3者)：

(i)果醬捲域之胺基酸中之至少30% (例如至少30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、90%或更多)為一或多個β片之一部分；

(ii)果醬捲域之二級結構包含至少四條(例如至少4、5、6、7、8、9、10、11或12條) β股；及/或

(iii)果醬捲域之三級結構包含至少兩個(例如至少2、3或4個) β片；及/或

(iv)果醬捲域包含至少2:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1或10:1之β片與α螺旋之比。

【0153】 在某些實施例中，果醬捲域包含兩個β片。

【0154】 在某些實施例中，β片中之一或多者(例如1、2、3、4、5、6、7、8、9或10者)包含約八條(例如4、5、6、7、8、9、10、11或12條) β股。在某些實施例中，β片中之一或多者(例如1、2、3、4、5、6、7、8、9或10者)包含八條β股。在某些實施例中，β片中之一或多者(例如

1、2、3、4、5、6、7、8、9或10者)包含七條β股。在某些實施例中，β片中之一或多者(例如1、2、3、4、5、6、7、8、9或10者)包含六條β股。在某些實施例中，β片中之一或多者(例如1、2、3、4、5、6、7、8、9或10者)包含五條β股。在某些實施例中，β片中之一或多者(例如1、2、3、4、5、6、7、8、9或10者)包含四條β股。

【0155】 在一些實施例中，果醬捲域包含與第二β片呈反平行定向之第一β片。在某些實施例中，第一β片包含約四條(例如3、4、5或6條) β股。在某些實施例中，第二β片包含約四條(例如3、4、5或6條) β股。在實施例中，第一及第二β片總共包含約八條(例如6、7、8、9、10、11或12條) β股。

【0156】 在某些實施例中，果醬捲域為蛋白殼蛋白(例如如本文所述之ORF1分子)之組分。在某些實施例中，果醬捲域具有自組裝活性。在一些實施例中，包含果醬捲域之多肽結合至包含果醬捲域之多肽的另一複本。在一些實施例中，第一多肽之果醬捲域結合至多肽之第二複本之果醬捲域。

【0157】

*N22域*

ORF1分子亦可包括包含指環病毒N22域(例如如本文所述，例如來自如本文所述之指環病毒ORF1蛋白的N22域)之結構或活性的第三區域，及/或包含指環病毒C端域(CTD) (例如如本文所述，例如來自如本文所述之指環病毒ORF1蛋白的CTD)之結構或活性的第四區域。在一些實施例中，ORF1分子以N端至C端次序包含第一、第二、第三及第四區域。

【0158】

*高變區(HVR)*

在一些實施例中，ORF1分子可進一步包含高變區(HVR)，例如來自指環病毒ORF1蛋白之HVR，例如如本文所述。在一些實施例中，HVR位於第二區域與第三區域之間。在一些實施例中，HVR包含包含至少約55個(例如至少約45、50、51、52、53、54、55、56、57、58、59、60或65個)胺基酸(例如約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個胺基酸)。

【0159】

*例示性ORF1序列*

例示性指環病毒ORF1胺基酸序列及例示性ORF1域之序列提供於下表中。在一些實施例中，本文所述之多肽(例如ORF1分子)包含與一或多個指環病毒ORF1子序列(例如如表N-Z中之任一者中所述)具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的胺基酸序列。在一些實施例中，本文所述之指環載體包含ORF1分子，其包含與一或多個指環病毒ORF1子序列(例如如表N-Z中之任一者中所述)具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的胺基酸序列。在一些實施例中，本文所述之指環載體包含核酸分子(例如遺傳元件)，該核酸分子編碼包含與一或多個指環病毒ORF1子序列(例如如表N-Z中之任一者中所述)具有至少　約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的胺基酸序列的ORF1分子。

【0160】 在一些實施例中，一或多個指環病毒ORF1子序列包含以下中之一或多者：精胺酸(Arg)富集域、果醬捲域、高變區(HVR)、N22域

或C端域(CTD) (例如如表N-Z中之任一者中所列)或與其具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%之序列一致性的序列。在一些實施例中，ORF1分子包含來自不同指環病毒之複數個子序列(例如選自表N-Z中所列之α細環病毒屬分枝系1-7子序列的ORF1子序列之任何組合)。在實施例中，ORF1分子包含以下中之一或多者：Arg富集域、果醬捲域、N22域及來自一種指環病毒之CTD及來自另一種之HVR。在實施例中，ORF1分子包含以下中之一或多者：果醬捲域、HVR、N22域及來自一種指環病毒之CTD及來自另一種之Arg富集域。在實施例中，ORF1分子包含以下中之一或多者：Arg富集域、HVR、N22域及來自一種指環病毒之CTD及來自另一種之果醬捲域。在實施例中，ORF1分子包含以下中之一或多者：Arg富集域、果醬捲域、HVR及來自一種指環病毒之CTD及來自另一種之N22域。在實施例中，ORF1分子包含以下中之一或多者：Arg富集域、果醬捲域、HVR及來自一種指環病毒之N22域及來自另一種之CTD。

【0161】 ORF1分子或其剪接變異體或功能片段可用於本文所述之組合物及方法中(例如以形成指環載體之蛋白質外部，例如藉由包裹遺傳元件)的其他例示性指環病毒描述於例如PCT申請案第PCT/US2018/037379號及第PCT/US19/65995號(以全文引用之方式併入本文中)中。

## 表N.例示性指環病毒ORF1胺基酸子序列(α細環病毒屬，分枝系3)

| | |
|---|---|
| 名稱 | 環1 |
| 屬/分枝系 | α細環病毒屬，分枝系3 |
| 寄存編號 | AJ620231.1 |
| 蛋白質寄存編號 | CAF05750.1 |

**全序列：**743 AA

```
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)
```

**註釋：**

| *假定域* | *AA範圍* |
|---|---|
| 富含Arg之區域 | 1 - 68 |
| 果醬捲域 | 69 - 280 |
| 高變區 | 281 - 413 |
| N22 | 414 - 579 |
| C端域 | 580 - 743 |

**表O.例示性指環病毒ORF1胺基酸子序列(α細環病毒屬，分枝系3)**

| 環1 ORF1 (α細環病毒屬分枝系3) | |
|---|---|
| 富含Arg之區域 | MAWGWWKRRRRWWFRKRWTRGRLRRRWPRSARRRPRRRRVRRRRRWRRGRRKTRTYRRRRRFRRRGRK (SEQ ID NO: 186) |
| 果醬捲域 | AKLIIKLWQPAVIKRCRIKGYIPLIISGNGTFATNFTSHINDRIMKGPFGGGHSTMRFSLYILFEEHLRHMNFWTRSNDNLELTRYLGASVKIYRHPDQDFIVIYNRRTPLGGNIYTAPSLHPGNAILAKHKILVPSLQTRPKGRKAIRLRIAPPTLFTDKWYFQKDIADLTLFNIMAVEADLRFPFCSPQTDNTCISFQVLSSVYNNYLSI (SEQ ID NO: 187) |
| 高變域 | NTFNNDNSDSKLKEFLNKAFPTTGTKGTSLNALNTFRTEGCISHPQLKKPNPQINKPLESQYFAPLDALWGDPIYYNDLNENKSLNDIIEKILIKNMITYHAKLREFPNSYQGNKAFCHLTGIYSPPYLNQGR (SEQ ID NO: 188) |
| N22 | ISPEIFGLYTEIIYNPYTDKGTGNKVWMDPLTKENNIYKEGQSKCLLTDMPLWTLLFGYTDWCKKDTNNWDLPLNYRLVLICPYTFPKLYNEKVKDYGYIPYSYKFGAGQMPDGSNYIPFQFRAKWYPTVLHQQQVMEDISRSGPFAPKVEKPSTQLVMKYCFNFN (SEQ ID NO: 189) |
| C端域 | WGGNPIIEQIVKDPSFQPTYEIPGTGNIPRRIQVIDPRVLGPHYSFRSWDMRRHTFSRASIKRVSEQQETSDLVFSGPKKPRVDIPKQETQEESSHSLQRESRPWETEEESETEALSQESQEVPFQQQLQQQYQEQLKLRQGIKVLFEQLIRTQQGVHVNPCLR (SEQ ID NO: 190) |

## 表P.例示性指環病毒ORF1胺基酸子序列(β細環病毒屬)

| | |
|---|---|
| 名稱 | 環2 |
| 屬/分枝系 | β細環病毒屬 |
| 寄存編號 | JX134045.1 |
| 蛋白質寄存編號 | AGG91484.1 |

**全序列：**666 AA

```
1         10         20         30         40         50
|         |          |          |          |          |
MPYYYRRRRYNYRRPRWYGRGWIRRPFRRRFRRKRRVRPTYTTIPLKQWQ
PPYKRTCYIKGQDCLIYYSNLRLGMNSTMYEKSIVPVHWPGGGSFSVSML
TLDALYDIHKLCRNWWTSTNQDLPLVRYKGCKITFYQSTFTDYIVRIHTE
LPANSNKLTYPNTHPLMMMMSKYKHIIPSRQTRRKKKPYTKIFVKPPPQF
ENKWYFATDLYKIPLLQIHCTACNLQNPFVKPDKLSNNVTLWSLNTISIQ
NRNMSVDQGQSWPFKILGTQSFYFYFYTGANLPGDTTQIPVADLLPLTNP
RINRPGQSLNEAKITDHITFTEYKNKFTNYWGNPFNKHIQEHLDMILYSL
KSPEAIKNEWTTENMKWNQLNNAGTMALTPFNEPIFTQIQYNPDRDTGED
TQLYLLSNATGTGWDPPGIPELILEGFPLWLIYWGFADFQKNLKKVTNID
TNYMLVAKTKFTQKPGTFYLVILNDTFVEGNSPYEKQPLPEDNIKWYPQV
QYQLEAQNKLLQTGPFTPNIQGQLSDNISMFYKFYFKWGGSPPKAINVEN
PAHQIQYPIPRNEHETTSLQSPGEAPESILYSFDYRHGNYTTTALSRISQ
DWALKDTVSKITEPDRQQLLKQALECLQISEETQEKKEKEVQQLISNLRQ
    QQQLYRERIISLLKDQ (SEQ ID NO: 215)
```

**註釋：**

| *假定域* | *AA範圍* |
|---|---|
| 富含Arg之區域 | 1 - 38 |
| 果醬捲域 | 39 - 246 |
| 高變區 | 247 - 374 |
| N22 | 375 - 537 |
| C端域 | 538 - 666 |

## 表Q.例示性指環病毒ORF1胺基酸子序列(β細環病毒屬)

| 環2 ORF1 (β細環病毒屬) | |
|---|---|
| 富含Arg之區域 | MPYYYRRRRYNYRRPRWYGRGWIRRPFRRRFRRKRRVR (SEQ ID NO: 216) |
| 果醬捲域 | PTYTTIPLKQWQPPYKRTCYIKGQDCLIYYSNLRLGMNSTMYEKSIVPV HWPGGGSFSVSMLTLDALYDIHKLCRNWWTSTNQDLPLVRYKGCKITF YQSTFTDYIVRIHTELPANSNKLTYPNTHPLMMMMSKYKHIIPSRQTRR KKKPYTKIFVKPPPQFENKWYFATDLYKIPLLQIHCTACNLQNPFVKPD KLSNNVTLWSLNT (SEQ ID NO: 217) |
| 高變域 | ISIQNRNMSVDQGQSWPFKILGTQSFYFYFYTGANLPGDTTQIPVADLLP LTNPRINRPGQSLNEAKITDHITFTEYKNKFTNYWGNPFNKHIQEHLDM ILYSLKSPEAIKNEWTTENMKWNQLNNAG (SEQ ID NO: 218) |
| N22 | TMALTPFNEPIFTQIQYNPDRDTGEDTQLYLLSNATGTGWDPPGIPELIL EGFPLWLIYWGFADFQKNLKKVTNIDTNYMLVAKTKFTQKPGTFYLVIL NDTFVEGNSPYEKQPLPEDNIKWYPQVQYQLEAQNKLLQTGPFTPNIQ GQLSDNISMFYKFYFK (SEQ ID NO: 219) |
| C端域 | WGGSPPKAINVENPAHQIQYPIPRNEHETTSLQSPGEAPESILYSFDYRH GNYTTTALSRISQDWALKDTVSKITEPDRQQLLKQALECLQISEETQEK KEKEVQQLISNLRQQQQLYRERIISLLKDQ (SEQ ID NO: 220) |

## 表R.例示性指環病毒ORF1胺基酸子序列(γ細環病毒屬)

名稱 環4

屬/分枝系 γ細環病毒屬

寄存編號

蛋白質寄存編號

**全序列：**662 AA

```
1          10         20         30         40         50
|          |          |          |          |          |
MPFWWRRRRKFWTNNRFNYTKRRRYRKRWPRRRRRRRPYRRPVRRRRRKL
RKVKRKKKSLIVRQWQPDSIRTCKIIGQSAIVVGAEGKQMYCYTVNKLIN
VPPKTPYGGGFGVDQYTLKYLYEEYRFAQNIWTQSNVLKDLCRYINVKLI
FYRDNKTDFVLSYDRNPPFQLTKFTYPGAHPQQIMLQKHHKFILSQMTKP
NGRLTKKLKIKPPKQMLSKWFFSKQFCKYPLLSLKASALDLRHSYLGCCN
ENPQVFFYYLNHGYYTITNWGAQSSTAYRPNSKVTDTTYYRYKNDRKNIN
IKSHEYEKSISYENGYFQSSFLQTQCIYTSERGEACIAEKPLGIAIYNPV
KDNGDGNMIYLVSTLANTWDQPPKDSAILIQGVPIWLGLFGYLDYCRQIK
ADKTWLDSHVLVIQSPAIFTYPNPGAGKWYCPLSQSFINGNGPFNQPPTL
LQKAKWFPQIQYQQEIINSFVESGPFVPKYANQTESNWELKYKYVFTFKW
GGPQFHEPEIADPSKQEQYDVPDTFYQTIQIEDPEGQDPRSLIHDWDYRR
GFIKERSLKRMSTYFSTHTDQQATSEEDIPKKKKRIGPQLTVPQQKEEET
LSCLLSLCKKDTFQETETQEDLQQLIKQQQEQQLLLKRNILQLIHKLKEN
        QQMLQLHTGMLP (SEQ ID NO: 925)
```

**註釋：**

| *假定域* | *AA範圍* |
|---|---|
| 富含Arg之區域 | 1 - 58 |

| | |
|---|---|
| 果醬捲域 | 59 - 260 |
| 高變區 | 261 - 339 |
| N22 | 340 - 499 |
| C端域 | 500 - 662 |

**表S.例示性指環病毒ORF1胺基酸子序列(γ細環病毒屬)**

| 環4 (γ細環病毒屬) | |
|---|---|
| 富含Arg之區域 | MPFWWRRRRKFWTNNRFNYTKRRRYRKRWPRRRRRRRPYRRPVRRRRRKLRKVKRKKK (SEQ ID NO: 926) |
| 果醬捲域 | SLIVRQWQPDSIRTCKIIGQSAIVVGAEGKQMYCYTVNKLINVPPKTPYGGGFGVDQYTLKYLYEEYRFAQNIWTQSNVLKDLCRYINVKLIFYRDNKTDFVLSYDRNPPFQLTKFTYPGAHPQQIMLQKHHKFILSQMTKPNGRLTKKLKIKPPKQMLSKWFFSKQFCKYPLLSLKASALDLRHSYLGCCNENPQVFFYYL (SEQ ID NO: 927) |
| 高變域 | NHGYYTITNWGAQSSTAYRPNSKVTDTTYYRYKNDRKNINIKSHEYEKSISYENGYFQSSFLQTQCIYTSERGEACIAE (SEQ ID NO: 928) |
| N22 | KPLGIAIYNPVKDNGDGNMIYLVSTLANTWDQPPKDSAILIQGVPIWLGLFGYLDYCRQIKADKTWLDSHVLVIQSPAIFTYPNPGAGKWYCPLSQSFINGNGPFNQPPTLLQKAKWFPQIQYQQEIINSFVESGPFVPKYANQTESNWELKYKYVFTFK (SEQ ID NO: 929) |
| C端域 | WGGPQFHEPEIADPSKQEQYDVPDTFYQTIQIEDPEGQDPRSLIHDWDYRRGFIKERSLKRMSTYFSTHTDQQATSEEDIPKKKKRIGPQLTVPQQKEEETLSCLLSLCKKDTFQETETQEDLQQLIKQQQEQQLLLKRNILQLIHKLKENQQMLQLHTGMLP (SEQ ID NO: 930) |

【0162】在一些實施例中，第一區域可結合至核酸分子(例如DNA)。在一些實施例中，鹼性殘基係選自精胺酸、組胺酸或離胺酸或其組合。在一些實施例中，第一區域包含至少60%、65%、70%、75%、80%、85%、90%、95%或100%精胺酸殘基(例如60%-90%、60%-80%、70%-90%或70%-80%精胺酸殘基)。在一些實施例中，第一區域包含約30-120個胺基酸(例如約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個胺基酸)。在一些實施例中，第一區域包含病毒ORF1富含精胺酸之區域(例如來自指環病毒ORF1蛋白之富含精胺酸之區域，例如如本文所述)的結構或活性。在一些實施例中，第一區域包含核定位信號。

【0163】在一些實施例中，第二區域包含果醬捲域，例如病毒

ORF1果醬捲域(例如來自指環病毒ORF1蛋白之果醬捲域，例如如本文所述)的結構或活性。在一些實施例中，第二區域能夠結合至另一ORF1分子之第二區域，例如以形成蛋白質外部(例如蛋白殼)或其一部分。

【0164】 在一些實施例中，第四區域暴露於蛋白質外部(例如包含ORF1分子之多聚體的蛋白質外部，例如如本文所述)之表面上。

【0165】 在一些實施例中，第一區域、第二區域、第三區域、第四區域及/或HVR各自包含少於四個(例如0、1、2或3個) β片 。

【0166】 在一些實施例中，第一區域、第二區域、第三區域、第四區域及/或HVR中之一或多者可由異源性胺基酸序列(例如來自異源性ORF1分子之對應區域)替換。在一些實施例中，異源性胺基酸序列具有所需功能性，例如如本文所述。

【0167】 在一些實施例中，ORF1分子包含複數個保守模體(例如，包含約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或更多個胺基酸的模體) (例如，如PCT/US19/65995之圖34中所示)。在一些實施例中，保守模體可展示與一或多種野生型指環病毒分枝系(例如α細環病毒屬，分枝系1；α細環病毒屬，分枝系2；α細環病毒屬，分枝系3；α細環病毒屬，分枝系4；α細環病毒屬，分枝系5；α細環病毒屬，分枝系6；α細環病毒屬，分枝系7；β細環病毒屬；及/或γ細環病毒屬)之ORF1蛋白的60、70、80、85、90、95或100%序列一致性。在實施例中，保守模體各自之長度為在1-1000個(例如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個)之間的胺基酸。在某些實施例中，保守模體由ORF1分子之序列的約2-4% (例如

約1-8%、1-6%、1-5%、1-4%、2-8%、2-6%、2-5%或2-4%)組成，且各自展示與野生型指環病毒分枝系之ORF1蛋白中之對應模體100%序列一致性。在某些實施例中，保守模體由ORF1分子之序列的約5-10% (例如約1-20%、1-10%、5-20%或5-10%)組成，且各自展示與野生型指環病毒分枝系之ORF1蛋白中之對應模體80%序列一致性。在某些實施例中，保守模體由ORF1分子之序列的約10-50% (例如約10-20%、10-30%、10-40%、10-50%、20-40%、20-50%或30-50%)組成，且各自展示與野生型指環病毒分枝系之ORF1蛋白中之對應模體60%序列一致性。在一些實施例中，保守模體包含一或多個如表19中所列之胺基酸序列。

【0168】 在一些實施例中，相對於野生型ORF1蛋白，ORF1分子或編碼其之核酸分子包含至少一個差異(例如突變、化學修飾或表觀遺傳改變)，例如如本文所述。

【0169】

*N22域中之保守ORF1模體*

在一些實施例中，本文所述之多肽(例如ORF1分子)包含胺基酸序列 $YNPX^2DXGX^2N$ (SEQ ID NO: 829)，其中$X^n$為任何*n*個胺基酸之連續序列。舉例而言，$X^2$指示任何兩個胺基酸之連續序列。在一些實施例中，$YNPX^2DXGX^2N$ (SEQ ID NO: 829)包含於ORF1分子之N22域內，例如如本文所述。在一些實施例中，本文所述之遺傳元件包含編碼胺基酸序列 $YNPX^2DXGX^2N$ (SEQ ID NO: 829)之核酸序列(例如編碼ORF1分子之核酸序列，例如如本文所述)，其中$X^n$為任何*n*個胺基酸之連續序列。

【0170】 在一些實施例中，多肽(例如ORF1分子)包含保守二級結構，例如側接及/或包含$YNPX^2DXGX^2N$ (SEQ ID NO: 829)模體之一部

分，例如在N22域中。在一些實施例中，保守二級結構包含第一β股及/或第二β股。在一些實施例中，第一β股之長度為約5-6個(例如3、4、5、6、7或8個)胺基酸。在一些實施例中，第一β股包含在YNPX$^2$DXGX$^2$N (SEQ ID NO: 829)模體之N端的酪胺酸(Y)殘基。在一些實施例中，YNPX$^2$DXGX$^2$N (SEQ ID NO: 829)模體包含無規捲曲(例如約8-9個胺基酸之無規捲曲)。在一些實施例中，第二β股之長度為約7-8個(例如5、6、7、8、9或10個)胺基酸。在一些實施例中，第二β股包含在YNPX$^2$DXGX$^2$N (SEQ ID NO: 829)模體之C端的天冬醯胺(N)殘基。

【0171】 例示性YNPX$^2$DXGX$^2$N (SEQ ID NO: 829)模體側接二級結構描述於PCT/US19/65995之實例47及圖48中；該案以全文引用之方式併入本文中。在一些實施例中，ORF1分子包含有包含PCT/US19/65995之圖48中所示之一或多個(例如1、2、3、4、5、6、7、8、9、10個或全部)二級結構元件(例如β股)的區域。在一些實施例中，ORF1分子包含有包含PCT/US19/65995之圖48中所示之一或多個(例如1、2、3、4、5、6、7、8、9、10個或全部)二級結構元件(例如β股)的區域，側接YNPX$^2$DXGX$^2$N (SEQ ID NO: 829)模體(例如如本文所述)。

【0172】

*ORF1果醬捲域中之保守二級結構模體*

在一些實施例中，本文所述之多肽(例如ORF1分子)包含指環病毒ORF1蛋白(例如如本文所述)所包含之一或多個二級結構元件。在一些實施例中，ORF1分子包含指環病毒ORF1蛋白之果醬捲域所包含之一或多個二級結構元件(例如如本文所述)。通常，ORF1果醬捲域包含二級結構，該二級結構按N端至C端方向之順序包含第一β股、第二β股、第一α螺旋、

第三β股、第四β股、第五β股、第二α螺旋、第六β股、第七β股、第八β股及第九β股。在一些實施例中，ORF1分子包含二級結構，該二級結構按N端至C端方向之順序包含第一β股、第二β股、第一α螺旋、第三β股、第四β股、第五β股、第二α螺旋、第六β股、第七β股、第八β股及/或第九β股。

【0173】 在一些實施例中，一對保守二級結構元件(亦即β股及/或α螺旋)由間隙胺基酸序列隔開，例如包含無規捲曲序列、β股或α螺旋或其組合。保守二級結構元件之間的間隙胺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個或更多個胺基酸。在一些實施例中，ORF1分子可進一步包含一或多個額外β股及/或α螺旋(例如在果醬捲域中)。在一些實施例中，可組合連續β股或連續α螺旋。在一些實施例中，第一β股及第二β股包含於較大β股中。在一些實施例中，第三β股及第四β股包含於較大β股中。在一些實施例中，第四β股及第五β股包含於較大β股中。在一些實施例中，第六β股及第七β股包含於較大β股中。在一些實施例中，第七β股及第八β股包含於較大β股中。在一些實施例中，第八β股及第九β股包含於較大β股中。

【0174】 在一些實施例中，第一β股之長度為約5-7個(例如3、4、5、6、7、8、9或10個)胺基酸。在一些實施例中，第二β股之長度為約15-16個(例如13、14、15、16、17、18或19個)胺基酸。在一些實施例中，第一α螺旋之長度為約15-17個(例如13、14、15、16、17、18、19或20個)胺基酸。在一些實施例中，第三β股之長度為約3-4個(例如1、2、3、4、5或6個)胺基酸。在一些實施例中，第四β股之長度為約10-11個(例如

8、9、10、11、12或13個)胺基酸。在一些實施例中，第五β股之長度為約6-7個(例如4、5、6、7、8、9或10個)胺基酸。在一些實施例中，第二α螺旋之長度為約8-14個(例如5、6、7、8、9、10、11、12、13、14、15、16或17個)胺基酸。在一些實施例中，第二α螺旋可分解為兩個較小α螺旋(例如由無規捲曲序列隔開)。在一些實施例中，兩個較小α螺旋中之每一者之長度為約4-6個(例如2、3、4、5、6、7或8個)胺基酸。在一些實施例中，第六β股之長度為約4-5個(例如，2、3、4、5、6或7個)胺基酸。在一些實施例中，第七β股之長度為約5-6個(例如3、4、5、6、7、8或9個)胺基酸。在一些實施例中，第八β股之長度為約7-9個(例如5、6、7、8、9、10、11、12或13個)胺基酸。在一些實施例中，第九β股之長度為約5-7個(例如，3、4、5、6、7、8、9或10個)胺基酸。

【0175】 例示性果醬捲域二級結構描述於PCT/US19/65995之實例47及圖47中。在一些實施例中，ORF1分子含有包含PCT/US19/65995之圖47中所示之任一果醬捲域二級結構之一或多個(例如1、2、3、4、5、6、7、8、9、10個或全部)二級結構元件(例如β股及/或α螺旋)的區域。

【0176】

*共有ORF1域序列*

在一些實施例中，例如如本文所述之ORF1分子包含果醬捲域、N22域及/或C端域(CTD)中之一或多者。在一些實施例中，果醬捲域包含具有如本文所述(例如如表37A-37C中之任一者中所列)之果醬捲域共有序列的胺基酸序列。在一些實施例中，N22域包含具有如本文所述(例如如表37A-37C中之任一者中所列)之N22域共有序列的胺基酸序列。在一些實施例中，CTD域包含具有如本文所述(例如如表37A-37C中之任一者中所列)

之CTD域共有序列的胺基酸序列。在一些實施例中，呈型式「($X_{a-b}$)」的表37A-37C中之任一者中所列之胺基酸包含一系列連續胺基酸，其中該系列包含至少*a*個且至多*b*個胺基酸。在某些實施例中，該系列中之所有胺基酸均相同。在其他實施例中，該系列包含至少兩個(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個)不同胺基酸。

**表37A.α細環病毒ORF1域共有序列**

| 域 | 序列 | SEQ ID NO: |
|---|---|---|
| 果醬捲 | LVLTQWQPNTVRRCYIRGYLPLIICGEN($X_{0-3}$)TTSRNYATHSDDTIQKGPFGGGMSTTTFSLRVLYDEYQRFMNRWTYSNEDLDLARYLGCKFTFYRHPDXDFIVQYNTNPPFKDTKLTAPSIHP($X_{1-5}$)GMLMLSKRKILIPSLKTRPKGKHYVKVRIGPPKLFEDKWYTQSDLCDVPLVXLYATAADLQHPFGSPQTDNPCVTFQVLGSXYNKHLSISP;<br><br>其中X=任何胺基酸。 | 227 |
| N22 | SNFEFPGAYTDITYNPLTDKGVGNMVWIQYLTKPDTIXDKTQS($X_{0-3}$)KCLIEDLPLWAALYGYVDFCEKETGDSAIIXNXGRVLIRCPYTKPPLYDKT($X_{0-4}$)NKGFVPYSTNFGNGKMPGGSGYVPIYWRARWYPTLFHQKEVLEDIVQSGPFAYKDEKPSTQLVMKYCFNFN;<br><br>其中X=任何胺基酸。 | 228 |
| CTD | WGGNPISQQVVRNPCKDSG($X_{0-3}$)SGXGRQPRSVQVVDPKYMGPEYTFHSWDWRRGLFGEKAIKRMSEQPTDDEIFTGGXPKRPRRDPPTXQXPEE($X_{1-4}$)QKESSSFR($X_{2-14}$)PWESSSQEXESESQEEEE($X_{0-30}$)EQTVQQQLRQQLREQRRLRVQLQLLFQQLLKT($X_{0-4}$)QAGLHINPLLLSQA($X_{0-40}$)*;<br><br>其中X=任何胺基酸。 | 229 |

## 表37B.β細環病毒ORF1域共有序列

| 域 | 序列 | SEQ ID NO: |
|---|---|---|
| 果醬捲 | LKQWQPSTIRKCKIKGYLPLFQCGKGRISNNYTQYKESIVPHHEPGGGGWSIQQFTLGALYEEHLKLRNWWTKSNDGLPLVRYLGCTIKLYRSEDTDYIVTYQRCYPMTATKLTYLSTQPSRMLMNKHKIIVPSKXT($X_{1-4}$)NKKKKPYKKIFIKPPSQMQNKWYFQQDIANTPLLQLTXTACSLDRMYLSSDSISNNITFTSLNTNFFQNPNFQ;<br><br>其中X=任何胺基酸。 | 230 |
| N22 | ($X_{4-10}$)TPLYFECRYNPFKDKGTGNKVYLVSNN($X_{1-8}$)TGWDPPTDPDLIIEGFPLWLLLWGWLDWQKKLGKIQNIDTDYILVIQSXYYIPP($X_{1-3}$)KLPYYVPLDXD($X_{0-2}$)FLHGRSPY($X_{3-16}$)PSDKQHWHPKVRFQXETINNIALTGPGTPKLPNQKSIQAHMKYKFYFK;<br><br>其中X=任何胺基酸。 | 231 |
| CTD | WGGCPAPMETITDPCKQPKYPIPNNLLQTTSLQXPTTPIETYLYKFDERRGLLTKKAAKRIKKDXTTETTLFTDTGXXTSTTLPTXXQTETTQEEXTSEEE($X_{0-5}$)ETLLQQLQQLRRKQKQLRXRILQLLQLLXLL($X_{0-26}$)*;<br><br>其中X=任何胺基酸。 | 232 |

## 表37C.γ細環病毒ORF1域共有序列

| 域 | 序列 | SEQ ID NO: |
|---|---|---|
| 果醬捲 | TIPLKQWQPESIRKCKIKGYGTLVLGAEGRQFYCYTNEKDEYTPPKAPGGGGFGVELFSLEYLYEQWKARNNIWTKSNXYKDLCRYTGCKITFYRHPTTDFIVXYSRQPPFEIDKXTYMXXHPQXLLLRKHKKIILSKATNPKGKLKKKIKIKPPKQMLNKWFFQKQFAXYGLVQLQAAACBLRYPRLGCCNENRLITLYYLN;<br><br>其中X=任何胺基酸。 | 233 |
| N22 | LPIVVARYNPAXDTGKGNKXWLXSTLNGSXWAPPTTDKDLIIEGLPLWLALYGYWSYJKKVKKDKGILQSHMFVVKSPAIQPLXTATTQXTFYPXIDNSFIQGKXPYDEPJTXNQKKLWYPTLEHQQETINAIVESGPYVPKLDNQKNSTWELXYXYTFYFK;<br><br>其中X=任何胺基酸。 | 234 |
| CTD | WGGPQIPDQPVEDPKXQGTYPVPDTXQQTIQIXNPLKQKPETMFHDWDYRRGIITSTALKRMQENLETDSSFXSDSEETP($X_{0-2}$)KKKKRLTXELPXPQEETEEIQSCLLSLCEESTCQEE($X_{1-6}$)ENLQQLIHQQQQQQQQLKHNILKLLSDLKZKQRLLQLQTGILE($X_{1-10}$)*;<br><br>其中X=任何胺基酸。 | 235 |

【0177】在一些實施例中，果醬捲域包含如表21、23、25、27、29、31、33、35、D2、D4、D6、D8、D10或37A-37C中之任一者中所列之果醬捲域胺基酸序列，或與其具有至少70%、75%、80%、8%、90%、95%、96%、97%、98%、99%或100%序列一致性的胺基酸序列。在一些實施例中，N22域包含如表21、23、25、27、29、31、33、35、D2、D4、D6、D8、D10或37A-37C中之任一者中所列之N22域胺基酸序列，或與其具有至少70%、75%、80%、8%、90%、95%、96%、97%、98%、99%或100%序列一致性的胺基酸序列。在一些實施例中，CTD域包含如表21、23、25、27、29、31、33、35、D2、D4、D6、D8、D10或37A-37C中之任一者中所列之CTD域胺基酸序列，或與其具有至少70%、75%、80%、8%、90%、95%、96%、97%、98%、99%或100%序列一致性的胺基酸序列。

【0178】

*鑑別ORF1蛋白序列*

在一些實施例中，可自指環病毒之基因體(例如經鑑別之假定指環病毒基因體，例如藉由核酸定序技術，例如深度定序技術)鑑別指環病毒ORF1蛋白序列或編碼ORF1蛋白之核酸序列。在一些實施例中，ORF1蛋白序列係藉由以下選擇標準中之一或多者(例如1、2或全部3者)鑑別：

*(i)長度選擇：*蛋白質序列(例如符合下文(ii)或(iii)中所述之準則的假定指環病毒ORF1序列)可對於大於約600個胺基酸殘基之彼等進行尺寸選擇，以鑑別假定指環病毒ORF1蛋白。在一些實施例中，指環病毒ORF1蛋白序列之長度為至少約600、650、700、750、800、850、900、950或1000個胺基酸殘基。在一些實施例中，α細環病毒ORF1蛋白序列之長度

為至少約700、710、720、730、740、750、760、770、780、790、800、900或1000個胺基酸殘基。在一些實施例中，β細環病毒ORF1蛋白序列之長度為至少約650、660、670、680、690、700、750、800、900或1000個胺基酸殘基。在一些實施例中，γ細環病毒ORF1蛋白序列之長度為至少約650、660、670、680、690、700、750、800、900或1000個胺基酸殘基。在一些實施例中，編碼指環病毒ORF1蛋白之核酸序列的長度為至少約1800、1900、2000、2100、2200、2300、2400或2500個核苷酸。在一些實施例中，編碼α細環病毒ORF1蛋白序列之核酸序列的長度為至少約2100、2150、2200、2250、2300、2400或2500個核苷酸。在一些實施例中，編碼β細環病毒ORF1蛋白序列之核酸序列的長度為至少約1900、1950、2000、2500、2100、2150、2200、2250、2300、2400或2500或1000個核苷酸。在一些實施例中，編碼γ細環病毒ORF1蛋白序列之核酸序列的長度為至少約1900、1950、2000、2500、2100、2150、2200、2250、2300、2400或2500或1000個核苷酸。

*(ii) ORF1模體之存在：*蛋白質序列(例如符合以上(i)或以下(iii)中所述之準則的假定指環病毒ORF1序列)可經過濾以鑑別在上文所述之N22域中含有保守ORF1模體之彼等蛋白質序列。在一些實施例中，假定指環病毒ORF1序列包含序列YNPXXDXGXXN。在一些實施例中，假定指環病毒ORF1序列包含序列Y[NCS]PXXDX[GASKR]XX[NTSVAK]。

*(iii)富含精胺酸之區域之存在：*蛋白質序列(例如符合以上(i)及/或(iii)中所述之準則的假定指環病毒ORF1序列)可對於包括富含精胺酸之區域(例如如本文所述)之彼等進行過濾。在一些實施例中，假定指環病毒ORF1序列包含至少約30、35、40、45、50、55、60、65或70個胺基酸

之連續序列，其包含至少30% (例如至少約20%、25%、30%、35%、40%、45%或50%)精胺酸殘基。在一些實施例中，假定指環病毒ORF1序列包含約35-40、40-45、45-50、50-55、55-60、60-65或65-70個胺基酸之連續序列，其包含至少30% (例如至少約20%、25%、30%、35%、40%、45%或50%)精胺酸殘基。在一些實施例中，富含精胺酸之區域位於假定指環病毒ORF1蛋白之起始密碼子下游至少約30、40、50、60、70或80個胺基酸處。在一些實施例中，富含精胺酸之區域位於假定指環病毒ORF1蛋白之起始密碼子下游至少約50個胺基酸處。

【0179】

*ORF2分子*

在一些實施例中，指環載體包含ORF2分子及/或編碼ORF2分子之核酸。一般而言，ORF2分子包含具有指環病毒ORF2蛋白(例如如本文中所述之指環病毒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中之任一者中所列)或其功能片段之結構特徵及/或活性的多肽。在一些實施例中，ORF2分子包含與表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16或18中之任一者中所示之指環病毒ORF2蛋白序列具有至少70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的胺基酸序列。

【0180】 在一些實施例中，ORF2分子包含與α細環病毒、β細環病毒或γ細環病毒ORF2蛋白具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%序列一致性的胺基酸序列。在一些實施例中，ORF2分子(例如與α細環病毒ORF2蛋白具有至少75%、80%、85%、

90%、95%、96%、97%、98%或99%序列一致性之ORF2分子)的長度為250個或更少胺基酸(例如約150-200個胺基酸)。在一些實施例中，ORF2分子(例如與β細環病毒ORF2蛋白具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%序列一致性之ORF2分子)的長度為約50-150個胺基酸。在一些實施例中，ORF2分子(例如與γ細環病毒ORF2蛋白具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%序列一致性之ORF2分子)的長度為約100-200個胺基酸(例如約100-150個胺基酸)。在一些實施例中，ORF2分子包含螺旋-轉角-螺旋模體(例如包含兩個側接轉角區之α螺旋的螺旋-轉角-螺旋模體)。在一些實施例中，ORF2分子不包含TTV分離株TA278或TTV分離株SANBAN之ORF2蛋白的胺基酸序列。在一些實施例中，ORF2分子具有蛋白質磷酸酶活性。在一些實施例中，相對於野生型ORF2蛋白，例如如本文所述(例如如表A2、A4、A6、A8、A10、A12、C1-C5、2、4、6、8、10、12、14、16或18中 之任一者中所示)，ORF2分子或編碼其之核酸分子包含至少一個差異(例如突變、化學修飾或表觀遺傳改變)。

【0181】

*保守ORF2模體*

在一些實施例中，本文所述之多肽(例如ORF2分子)包含胺基酸序列[W/F]$X^7$H$X^3$C$X^1$C$X^5$H (SEQ ID NO: 949)，其中$X^n$為任何*n*個胺基酸之連續序列。在實施例中，$X^7$指示任何七個胺基酸之連續序列。在實施例中，$X^3$指示任何三個胺基酸之連續序列。在實施例中，$X^1$指示任何單一胺基酸。在實施例中，$X^5$指示任何五個胺基酸之連續序列。在一些實施例中，[W/F]可為色胺酸或苯丙胺酸。在一些實施例中，[W/F]$X^7$H$X^3$C$X^1$C$X^5$H

(SEQ ID NO: 949)包含於ORF2分子之N22域內，例如如本文所述。在一些實施例中，本文所述之遺傳元件包含編碼胺基酸序列[W/F]$X^7$HX$^3$CX$^1$CX$^5$H (SEQ ID NO: 949)之核酸序列(例如編碼ORF2分子之核酸序列，例如如本文所述)，其中$X^n$為任何*n*個胺基酸之連續序列。

【0182】

***遺傳元件***

在一些實施例中，指環載體包含遺傳元件。在一些實施例中，遺傳元件具有以下特徵中之一或多者：與宿主細胞之基因體實質上不整合，為游離型核酸，為單股RNA，為環狀，為約1至10 kb，存在於細胞核內，可與內源性蛋白結合，產生靶向宿主或目標細胞之基因、活性或功能的效應子，諸如多肽或核酸(例如RNA、iRNA、微小RNA)。在一個實施例中，遺傳元件為實質上非整合RNA。在一些實施例中，遺傳元件包含封裝信號，例如結合蛋白殼蛋白之序列。在一些實施例中，在封裝或蛋白殼結合序列之外，遺傳元件與野生型指環病毒核酸序列具有小於70%、60%、50%、40%、30%、20%、10%、5%序列一致性，例如與指環病毒核酸序列(例如如本文所述)具有小於70%、60%、50%、40%、30%、20%、10%、5%序列一致性。在一些實施例中，在封裝或蛋白殼結合序列之外，遺傳元件具有小於500 450、400、350、300、250、200、150或100個連續核苷酸，該等連續核苷酸與指環病毒核酸序列至少70%、75%、80%、8%、90%、95%、96%、97%、98%、99%或100%一致。

【0183】 在一些實施例中，遺傳元件之長度為小於20 kb (例如小於約19 kb、18 kb、17 kb、16 kb、15 kb、14 kb、13 kb、12 kb、11 kb、10 kb、9 kb、8 kb、7 kb、6 kb、5 kb、4 kb、3 kb、2 kb、1 kb或更

小)。在一些實施例中，遺傳元件之長度獨立地或另外為1000 kb (例如至少　　約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)。在一些實施例中，遺傳元件之長度為約2.5-4.6 kb、2.8-4.0 kb、3.0-3.8 kb或3.2-3.7 kb。在一些實施例中，遺傳元件之長度為約1.5-2.0 kb、1.5-2.5 kb、1.5-3.0 kb、1.5-3.5 kb、1.5-3.8 kb、1.5-3.9 kb、1.5-4.0 kb、1.5-4.5 kb或1.5-5.0 kb。在一些實施例中，遺傳元件之長度為約2.0-2.5 kb、2.0-3.0 kb、2.0-3.5 kb、2.0-3.8 kb、2.0-3.9 kb、2.0-4.0 kb、2.0-4.5 kb或2.0-5.0 kb。在一些實施例中，遺傳元件之長度為約2.5-3.0 kb、2.5-3.5 kb、2.5-3.8 kb、2.5-3.9 kb、2.5-4.0 kb、2.5-4.5 kb或2.5-5.0 kb。在一些實施例中，遺傳元件之長度為約3.0-5.0 kb、3.5-5.0 kb、4.0-5.0 kb或4.5-5.0 kb。在一些實施例中，遺傳元件之長度為約1.5-2.0 kb、2.0-2.5 kb、2.5-3.0 kb、3.0-3.5 kb、3.1-3.6 kb、3.2-3.7 kb、3.3-3.8 kb、3.4-3.9 kb、3.5-4.0 kb、4.0-4.5 kb或4.5-5.0 kb。在一些實施例中，遺傳元件之長度為約3.6-3.9 kb。在一些實施例中，遺傳元件之長度為約2.8-2.9 kb。在一些實施例中，遺傳元件之長度為約2.0-3.2 kb。

【0184】在一些實施例中，遺傳元件包含本文所述之特徵中之一或多者，例如編碼實質上非病原蛋白之序列、蛋白質結合序列、編碼調控核酸之一或多個序列、一或多個調控序列、編碼複製蛋白之一或多個序列及其他序列。

【0185】 在實施例中，遺傳元件由雙股環狀DNA (例如藉由轉錄產生)產生。

【0186】 在一些實施例中，遺傳元件不包含一或多種細菌質體元件(例如細菌複製起點或可選標記，例如細菌抗性基因)。在一些實施例中，遺傳元件不包含細菌質體主鏈。

【0187】 在一個實施例中，本發明提供一種包含編碼以下之核酸序列(例如RNA序列)的遺傳元件：(i)實質上非病原性外部蛋白質；(ii)將遺傳元件結合至實質上非病原性外部蛋白質之外部蛋白質結合序列；及(iii)調控核酸。在此類實施例中，遺傳元件可包含一或多個與天然病毒序列(例如天然指環病毒序列，例如如本文所描述)之核苷酸序列中之任一者具有至少約60%、70%、80%、85%、90%、95%、96%、97%、98%及99%核苷酸序列一致性的序列。

【0188】

*蛋白質結合序列*

【0189】 在一些實施例中，遺傳元件編碼結合至實質上非病原性蛋白質之蛋白質結合序列。在一些實施例中，蛋白質結合序列有助於將遺傳元件封裝至蛋白質外部中。在一些實施例中，蛋白質結合序列特異性結合實質上非病原性蛋白質之富含精胺酸之區域。在實施例中，遺傳元件包含如PCT/US19/65995之實例8中所述之蛋白質結合序列。

【0190】 在一些實施例中，遺傳元件包含與指環病毒序列之5' UTR保守域或富含GC之結構域具有至少70%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的蛋白質結合序列，例如如本文所述。

【0191】 在一些實施例中，蛋白質結合序列與指環病毒5' UTR保守域核苷酸序列具有至少約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性，例如如本文所述。

【0192】

*5' UTR區*

在一些實施例中，如本文所述之核酸分子(例如遺傳元件、遺傳元件構築體或遺傳元件區域)包含5' UTR序列，例如如本文(例如在表A1、B1或C1中的任一者中)所述之5' UTR保守域序列，或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%序列一致性的序列。

【0193】 在一些實施例中，5' UTR序列包含核酸序列AGGTGAGTGAAACCACCGAAGTCAAGGGGCAATTCGGGCTAGGGX$_1$CAGTCT，或與其具有至少85%、90%、95%、96%、97%、98%或99%序列一致性的核酸序列。在一些實施例中，5' UTR序列包含核酸序列AGGTGAGTGAAACCACCGAAGTCAAGGGGCAATTCGGGCTAGGGX$_1$CAGTCT，或相對於其具有不超過1、2、3、4、5、6、7、8、9或10個核苷酸差異(例如取代、缺失或添加)的核酸序列。在實施例中，$X_1$為A。在實施例中，$X_1$不存在。

【0194】 在一些實施例中，5' UTR序列包含α細環病毒(例如環1)之5' UTR之核酸序列，或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%序列一致性的序列。在實施例中，5' UTR序列包含表A1中所列之5' UTR保守域核酸序列或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%序列一致性的序列。在一些實施例中，核酸分子包含與表A1中所列之5' UTR保守域具有至少95%序列一致性的核酸序

列。在一些實施例中，核酸分子包含與表A1中所列之5' UTR保守域具有至少95.775%序列一致性的核酸序列。在一些實施例中，核酸分子包含與表A1中所列之5' UTR保守域具有至少97%序列一致性的核酸序列。在一些實施例中，核酸分子包含與表A1中所列之5' UTR保守域具有至少97.183%序列一致性的核酸序列。在一些實施例中，5' UTR序列包含核酸序列AGGTGAGTTTACACACCGCAGTCAAGGGGCAATTCGGGCTCGGGACTGGC，或與其具有至少85%、90%、95%、96%、97%、98%或99%序列一致性的核酸序列。在一些實施例中，5' UTR序列包含核酸序列AGGTGAGTTTACACACCGCAGTCAAGGGGCAATTCGGGCTCGGGACTGGC，或相對於其具有不超過1、2、3、4、5、6、7、8、9或10個核苷酸差異(例如取代、缺失或添加)的核酸序列。

【0195】 在一些實施例中，5' UTR序列包含β細環病毒(例如環2)之5' UTR的核酸序列，或與其具有至少75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列一致性的序列。在實施例中，5' UTR序列包含表B1中所列之5' UTR保守域核酸序列或與其具有至少75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列一致性的序列。在一些實施例中，核酸分子包含與表B1中所列之5' UTR保守域具有至少85%序列一致性的核酸序列。在一些實施例中，核酸分子包含與表B1中所列之5' UTR保守域具有至少87%序列一致性的核酸序列。在一些實施例中，核酸分子包含與表B1中所列之5' UTR保守域具有至少87.324%序列一致性的核酸序列。在一些實施例中，核酸分子包含與表B1中所列之5' UTR保守域具有至少88%序列一致性的核酸序

列。在一些實施例中，核酸分子包含與表B1中所列之5' UTR保守域具有至少88.732%序列一致性的核酸序列。在一些實施例中，核酸分子包含與表B1中所列之5' UTR保守域具有至少91%序列一致性的核酸序列。在一些實施例中，核酸分子包含與表B1中所列之5' UTR保守域具有至少91.549%序列一致性的核酸序列。在一些實施例中，核酸分子包含與表B1中所列之5' UTR保守域具有至少92%序列一致性的核酸序列。在一些實施例中，核酸分子包含與表B1中所列之5' UTR保守域具有至少92.958%序列一致性的核酸序列。在一些實施例中，核酸分子包含與表B1中所列之5' UTR保守域具有至少94%序列一致性的核酸序列。在一些實施例中，核酸分子包含與表B1中所列之5' UTR保守域具有至少94.366%序列一致性的核酸序列。在一些實施例中，核酸分子包含與表B1中所列之5' UTR保守域具有至少95%序列一致性的核酸序列。在一些實施例中，核酸分子包含與表B1中所列之5' UTR保守域具有至少95.775%序列一致性的核酸序列。在一些實施例中，核酸分子包含與表B1中所列之5' UTR保守域具有至少97%序列一致性的核酸序列。在一些實施例中，核酸分子包含與表B1中所列之5' UTR保守域具有至少97.183%序列一致性的核酸序列。在一些實施例中，5' UTR序列包含核酸序列AGGTGAGTGAAACCACCGAAGTCAAGGGGCAATTCGGGCTAGATCAGTCT，或與其具有至少85%、90%、95%、96%、97%、98%或99%序列一致性的核酸序列。在一些實施例中，5' UTR序列包含核酸序列AGGTGAGTGAAACCACCGAAGTCAAGGGGCAATTCGGGCTAGATCAGTCT，或相對於其具有不超過1、2、3、4、5、6、7、8、9或10個核苷酸差異(例如取代、缺失或添加)的核酸序列。

【0196】 在一些實施例中，5' UTR序列包含γ細環病毒(例如環4)之

5' UTR的核酸序列，或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%序列一致性的序列。在實施例中，5' UTR序列包含表C1中所列之5' UTR保守域核酸序列或與其具有至少75%、80%、85%、90%、95%、96%、97%、98%或99%序列一致性的序列。在一些實施例中，核酸分子包含與表C1中所列之5' UTR保守域具有至少97%序列一致性的核酸序列。在一些實施例中，核酸分子包含與表C1中所列之5' UTR保守域具有至少97.183%序列一致性的核酸序列。在一些實施例中，5' UTR序列包含核酸序列AGGTGAGTGAAACCACCGAGGTCTAGGGGCAATTCGGGCTAGGGCAGTCT，或與其具有至少85%、90%、95%、96%、97%、98%或99%序列一致性的核酸序列。在一些實施例中，5' UTR序列包含核酸序列AGGTGAGTGAAACCACCGAGGTCTAGGGGCAATTCGGGCTAGGGCAGTCT，或相對於其具有不超過1、2、3、4、5、6、7、8、9或10個核苷酸差異(例如取代、缺失或添加)的核酸序列。

【0197】 在一些實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含與指環病毒5' UTR序列(例如表38中所示之核酸序列)具有至少約75% (例如至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)一致性的核酸序列。在一些實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含具有表38中所示之共有5' UTR序列的核酸序列，其中$X_1$、$X_2$、$X_3$、$X_4$及$X_5$各自獨立地為任何核苷酸，例如其中$X_1$=G或T，$X_2$=C或A，$X_3$=G或A，$X_4$=T或C，且$X_5$=A、C或T)。在實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含與表38中所示之共有5' UTR序列具有至少約75% (例如至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)一致性的核酸序列。在實施例中，遺傳元件(例

如遺傳元件之蛋白質結合序列)包含與表38中所示之例示性TTV 5' UTR序列具有至少約75% (例如至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)一致性的核酸序列。在實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含與表38中所示之TTV-CT30F 5' UTR序列具有至少約75% (例如至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)一致性的核酸序列。在實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含與表38中所示之TTV-HD23a 5' UTR序列具有至少約75% (例如至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)一致性的核酸序列。在實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含與表38中所示之TTV-JA20 5' UTR序列具有至少約75% (例如至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)一致性的核酸序列。在實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含與表38中所示之TTV-TJN02 5' UTR序列具有至少約75% (例如至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)一致性的核酸序列。在實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含與表38中所示之TTV-tth8 5' UTR序列具有至少約75% (例如至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)一致性的核酸序列。

【0198】 在實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含與表38中所示之α細環病毒屬共有5' UTR序列具有至少約75% (例如至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)一致性的核酸序列。在實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含與表38中所示之α細環病毒屬分枝系1 5' UTR序列具有至少約75%

(例如至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)一致性的核酸序列。在實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含與表38中所示之α細環病毒屬分枝系2 5' UTR序列具有至少約75% (例如至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)一致性的核酸序列。在實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含與表38中所示之α細環病毒屬分枝系3 5' UTR序列具有至少約75% (例如至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)一致性的核酸序列。在實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含與表38中所示之α細環病毒屬分枝系4 5' UTR序列具有至少約75% (例如至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)一致性的核酸序列。在實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含與表38中所示之α細環病毒屬分枝系5 5' UTR序列具有至少約75% (例如至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)一致性的核酸序列。在實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含與表38中所示之α細環病毒屬分枝系6 5' UTR序列具有至少約75% (例如至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)一致性的核酸序列。在實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含與表38中所示之α細環病毒屬分枝系7 5' UTR序列具有至少約75% (例如至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)一致性的核酸序列。

**表38.來自指環病毒之例示性5' UTR序列**

| 來源 | 序列 | SEQ ID NO: |
|---|---|---|
| 共同 | CGGGTGCCGX$_1$AGGTGAGTTTACACACCGX$_2$ AGTCAAGGGGCAATTCGGGCTCX$_3$GGACTGG CCGGGCX$_4$X$_5$TGGG<br>X$_1$ = G或T<br>X$_2$ = C或A<br>X$_3$ = G或A<br>X$_4$ = T或C<br>X$_5$ = A、C或T | 105 |
| 例示性TTV序列 | CGGGTGCCGGAGGTGAGTTTACACACCGCA GTCAAGGGGCAATTCGGGCTCGGGACTGGC CGGGCTWTGGG | 106 |
| TTV-CT30F | CGGGTGCCGTAGGTGAGTTTACACACCGCAG TCAAGGGGCAATTCGGGCTCGGGACTGGCC GGGCTATGGG | 107 |
| TTV-HD23a | CGGGTGCCGGAGGTGAGTTTACACACCGCA GTCAAGGGGCAATTCGGGCTCGGGACTGGC CGGGCCCTGGG | 108 |
| TTV-JA20 | CGGGTGCCGGAGGTGAGTTTACACACCGCA GTCAAGGGGCAATTCGGGCTCGGGACTGGC CGGGCTTTGGG | 109 |
| TTV-TJN02 | CGGGTGCCGGAGGTGAGTTTACACACCGCA GTCAAGGGGCAATTCGGGCTCGGGACTGGC CGGGCTATGGG | 110 |
| TTV-tth8 | CGGGTGCCGGAGGTGAGTTTACACACCGAA GTCAAGGGGCAATTCGGGCTCAGGACTGGC CGGGCTTTGGG | 111 |
| α細環病毒屬共有5' UTR | CGGGTGCCGGAGGTGAGTTTACACACCGCA GTCAAGGGGCAATTCGGGCTCGGGACTGGC CGGGC<br>X$_1$X$_2$TGGG；其中X$_1$包含T或C，且其中X$_2$包含A、C或T。 | 112 |
| α細環病毒屬分枝系1 5' UTR (例如TTV-CT30F) | CGGGTGCCGTAGGTGAGTTTACACACCGCAG TCAAGGGGCAATTCGGGCTCGGGACTGGCC GGGCTATGGG | 113 |
| α細環病毒屬分枝系2 5' UTR (例如TTV-P13-1) | CGGGTGCCGGAGGTGAGTTTACACACCGCA GTCAAGGGGCAATTCGGGCTCGGGACTGGC CGGGCCCGGG | 114 |
| α細環病毒屬分枝系3 5' UTR (例如TTV-tth8) | CGGGTGCCGGAGGTGAGTTTACACACCGAA GTCAAGGGGCAATTCGGGCTCAGGACTGGC CGGGCTTTGGG | 115 |
| α細環病毒屬分枝系4 5' UTR (例如TTV-HD20a) | CGGGTGCCGGAGGTGAGTTTACACACCGCA GTCAAGGGGCAATTCGGGCTCGGGAGGCCG GGCCATGGG | 116 |
| α細環病毒屬分枝系5 5' UTR (例如TTV-16) | CGGGTGCCGGAGGTGAGTTTACACACCGCA GTCAAGGGGCAATTCGGGCTCGGGACTGGC CGGGCCCCGGG | 117 |

| | | |
|---|---|---|
| α細環病毒屬分枝系6 5' UTR (例如TTV-TJN02) | CGGGTGCCGGAGGTGAGTTTACACACCGCA GTCAAGGGGCAATTCGGGCTCGGGACTGGC CGGGCTATGGG | 118 |
| α細環病毒屬分枝系7 5' UTR (例如TTV-HD16d) | CGGGTGCCGAAGGTGAGTTTACACACCGCA GTCAAGGGGCAATTCGGGCTCGGGACTGGC CGGGCTATGGG | 119 |

【0199】

*鑑別5' UTR序列*

在一些實施例中，可在指環病毒之基因體(例如鑑別之假定指環病毒基因體，例如藉由核酸定序技術，例如深度定序技術)內鑑別指環病毒5' UTR序列。在一些實施例中，藉由以下步驟中之一或兩者鑑別指環病毒5' UTR序列：

*(i) 鑑別環化接合點：*在一些實施例中，5' UTR將位於全長環化指環病毒基因體之環化接合點附近。可例如藉由鑑別序列之重疊區域來鑑別環化接合點。在一些實施例中，序列之重疊區域可自序列修整以產生已環化之全長指環病毒基因體序列。在一些實施例中，基因體序列以此方式使用軟體環化。不希望受理論所束縛，計算環化基因體可使得序列之起始位置在非生物中定向。序列內之地標可用於在恰當方向上重新定向序列。舉例而言，地標序列可包括與如本文所述之指環病毒基因體內之一或多個元件(例如以下中之一或多者：TATA盒、加帽位點、起始元件、轉錄起始位點、5' UTR保守域、ORF1、ORF1/1、ORF1/2、ORF2、ORF2/2、ORF2/3、ORF2t/3、三個開放閱讀框架區、聚(A)信號或指環病毒之富含GC之區域，例如如本文所述)具有實質性同源性的序列。

*(ii) 鑑別5' UTR序列：*一旦已獲得假定指環病毒基因體序列，則可將序列(或其部分，例如長度為約40-50、50-60、60-70、70-80、80-90或90-100個核苷酸)與一或多個指環病毒5' UTR序列(例如如本文所述)相

比，以鑑別與其具有實質性同源性之序列。在一些實施例中，假定指環病毒5' UTR區與如本文所述之指環病毒5' UTR序列具有至少50%、60%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性。

【0200】

*富含GC之區域*

在一些實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含與表39中所示之核酸序列具有至少約75% (例如至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)一致性的核酸序列。在實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含與表39中所示之GC富集序列具有至少約75% (例如至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)一致性的核酸序列。

【0201】 在實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含與表39中所示之36個核苷酸之GC富集序列(例如36個核苷酸之共有富含GC之區域序列1、36個核苷酸之共有富含GC之區域序列2、TTV分枝系1 36個核苷酸之區域、TTV分枝系3 36個核苷酸之區域、TTV分枝系3分離株GH1 36個核苷酸之區域、TTV分枝系3 sle1932 36個核苷酸之區域、TTV分枝系4 ctdc002 36個核苷酸之區域、TTV分枝系5 36個核苷酸之區域、TTV分枝系6 36個核苷酸之區域或TTV分枝系7 36個核苷酸之區域)具有至少約75% (例如至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)一致性的核酸序列。在實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含一核酸序列，該核酸序列包含具有如表39中所示之36個核苷酸之GC富集序列(例如36個核苷酸之共有富含GC之區域

序列1、36個核苷酸之共有富含GC之區域序列2、TTV分枝系1 36個核苷酸之區域、TTV分枝系3 36個核苷酸之區域、TTV分枝系3分離株GH1 36個核苷酸之區域、TTV分枝系3 sle1932 36個核苷酸之區域、TTV分枝系4 ctdc002 36個核苷酸之區域、TTV分枝系5 36個核苷酸之區域、TTV分枝系6 36個核苷酸之區域或TTV分枝系7 36個核苷酸之區域)的至少10、15、20、25、30、31、32、33、34、35或36個連續核苷酸。

【0202】 在實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含與α細環病毒屬富含GC之區域序列具有至少約75% (例如至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)一致性的核酸序列，該α細環病毒屬富含GC之區域序列例如選自TTV-CT30F、TTV-P13-1、TTV-tth8、TTV-HD20a、TTV-16、TTV-TJN02或TTV-HD16d，例如如表39中所列。在實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含一核酸序列，該核酸序列包含具有α細環病毒屬富含GC之區域序列之至少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個連續核苷酸，該α細環病毒屬富含GC之區域序列例如選自TTV-CT30F、TTV-P13-1、TTV-tth8、TTV-HD20a、TTV-16、TTV-TJN02或TTV-HD16d，例如如表39中所列。

【0203】 在實施例中，36個核苷酸之GC富集序列係選自：

(i) CGCGCTGCGCGCGCCGCCCAGTAGGGGGAGCCATGC (SEQ ID NO: 160)，

(ii) GCGCTX$_1$CGCGCGCGCGCCGGGGGGCTGCGCCCCCCC (SEQ ID NO: 164)，其中X$_1$係選自T、G或A；

(iii) GCGCTTCGCGCGCCGCCCACTAGGGGGCGTTGCGCG (SEQ ID NO: 165)；

(iv) GCGCTGCGCGCGCCGCCCAGTAGGGGGCGCAATGCG (SEQ ID NO: 166)；

(v) GCGCTGCGCGCGCGGCCCCCGGGGGAGGCATTGCCT (SEQ ID NO: 167)；

(vi) GCGCTGCGCGCGCGCCGGGGGGGCGCCAGCGCCC (SEQ ID NO: 168)；

(vii) GCGCTTCGCGCGCGCCGGGGGGCTCCGCCCCCCC (SEQ ID NO: 169)；

(viii) GCGCTTCGCGCGCGCCGGGGGGCTGCGCCCCCCC (SEQ ID NO: 170)；

(ix) GCGCTACGCGCGCGCCGGGGGGCTGCGCCCCCCC (SEQ ID NO: 171)；或

(x) GCGCTACGCGCGCGCCGGGGGGCTCTGCCCCCCC (SEQ ID NO: 172)。

【0204】在實施例中，遺傳元件(例如，遺傳元件之蛋白質結合序列)包含核酸序列CGCGCTGCGCGCGCCGCCCAGTAGGGGGAGCCATGC (SEQ ID NO: 160)。

【0205】在實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含具有表39中所示之共有GC富集序列的核酸序列，其中$X_1$、$X_4$、$X_5$、$X_6$、$X_7$、$X_{12}$、$X_{13}$、$X_{14}$、$X_{15}$、$X_{20}$、$X_{21}$、$X_{22}$、$X_{26}$、$X_{29}$、$X_{30}$及$X_{33}$各自獨立地為任何核苷酸，且其中$X_2$、$X_3$、$X_8$、$X_9$、$X_{10}$、$X_{11}$、$X_{16}$、

$X_{17}$、$X_{18}$、$X_{19}$、$X_{23}$、$X_{24}$、$X_{25}$、$X_{27}$、$X_{28}$、$X_{31}$、$X_{32}$及$X_{34}$各自獨立地不存在或為任何核苷酸。在一些實施例中，$X_1$至$X_{34}$中之一或多者(例如全部)各自獨立地為表39中所指定之核苷酸(或不存在)。在實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含與表39中所示之TTV GC富集序列(例如全序列、片段1、片段2、片段3或其任何組合，例如依序片段1-3)具有至少約75% (例如至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)一致性的核酸序列。在實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含與表39中所示之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或其任何組合，例如依序片段1-7)具有至少約75% (例如至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)一致性的核酸序列。在實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含與表39中所示之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或其任何組合，例如依序片段1-6)具有至少約75% (例如至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)一致性的核酸序列。在實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含與表39中所示之TTV-JA20 GC富集序列(例如全序列、片段1、片段2或其任何組合，例如依序片段1及片段2)具有至少約75% (例如至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)一致性的核酸序列。在實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含與表39中所示之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或其任何組合，例如依序片段1-8)具有至少約75% (例如至少75%、80%、

85%、90%、95%、96%、97%、98%、99%或100%)一致性的核酸序列。在實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含與表39中所示之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或其任何組合，例如依序片段1-6)具有至少約75% (例如至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)一致性的核酸序列。在實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含與表39中所示之片段7具有至少約75% (例如至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)一致性的核酸序列。在實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含與表39中所示之片段8具有至少約75% (例如至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)一致性的核酸序列。在實施例中，遺傳元件(例如遺傳元件之蛋白質結合序列)包含與表39中所示之片段9具有至少約75% (例如至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)一致性的核酸序列。

**表39.來自指環病毒之例示性GC富集序列**

| 來源 | | 序列 | SEQ ID NO: |
|---|---|---|---|
| 共同 | | CGGCGGX$_1$GGX$_2$GX$_3$X$_4$X$_5$CGCGCTX$_6$CGCGCGCX$_7$X$_8$X$_9$X$_{10}$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$_{26}$CGCGCATX$_{27}$X$_{28}$GCX$_{29}$CGGGX$_{30}$CCCCCCCCCX$_{31}$X$_{32}$X$_{33}$GGGGGGCTCCGX$_{34}$CCCCCCGGCCCCCC<br><br>X$_1$ = G或C<br>X$_2$ = G、C或不存在<br>X$_3$ = C或不存在<br>X$_4$ = G或C<br>X$_5$ = G或C<br>X$_6$ = T、G或A<br>X$_7$ = G或C<br>X$_8$ = G或不存在<br>X$_9$ = C或不存在<br>X$_{10}$ = C或不存在<br>X$_{11}$ = G、A或不存在<br>X$_{12}$ = G或C<br>X$_{13}$ = C或T<br>X$_{14}$ = G或A<br>X$_{15}$ = G或A<br>X$_{16}$ = A、G、T或不存在<br>X$_{17}$ = G、C或不存在<br>X$_{18}$ = G、C或不存在<br>X$_{19}$ = C、A或不存在<br>X$_{20}$ = C或A<br>X$_{21}$ = T或A<br>X$_{22}$ = G或C<br>X$_{23}$ = G、T或不存在<br>X$_{24}$ = C或不存在<br>X$_{25}$ = G、C或不存在<br>X$_{26}$ = G或C<br>X$_{27}$ = G或不存在<br>X$_{28}$ = C或不存在<br>X$_{29}$ = G或A<br>X$_{30}$ = G或T<br>X$_{31}$ = C、T或不存在<br>X$_{32}$ = G、C、A或不存在<br>X$_{33}$ = G或C<br>X$_{34}$ = C或不存在 | 120 |
| 例示性TTV序 | 全序列 | GCCGCCGCGGCGGCGGSGGNGNSGCGC | 121 |

| | | | |
|---|---|---|---|
| 列 | | GCTDCGCGCGCSNNNCRCCRGGGGGNNNNCWGCSNCNCCCCCCCCCGCGCATGCGCGGGKCCCCCCCCCNNCGGGGGGCTCCGCCCCCCGGCCCCCCCCCGTGCTAAACCCACCGCGCATGCGCGACCACGCCCCCGCCGCC | |
| | 片段1 | GCCGCCGCGGCGGCGGSGGNGNSGCGCGCTDCGCGCGCSNNNCRCCRGGGGGNNNNCWGCSNCNCCCCCCCCCGCGCAT | 122 |
| | 片段2 | GCGCGGGKCCCCCCCCCNNCGGGGGGCTCCG | 123 |
| | 片段3 | CCCCCCGGCCCCCCCCCGTGCTAAACCCACCGCGCATGCGCGACCACGCCCCCGCCGCC | 124 |
| TTV-CT30F | 全序列 | GCGGCGG-GGGGGCG-GCCGCG-TTCGCGCGCCGCCCACCAGGGGGTG--CTGCG-CGCCCCCCCCCGCGCAT GCGCGGGGCCCCCCCCC--GGGGGGGCTCCGCCCCCCCGGCCCCCCCCCGTGCTAAACCCACCGCGCATGCGCGACCACGCCCCCGCCGCC | 125 |
| | 片段1 | GCGGCGG | 126 |
| | 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 | GGGGGGGCTCCGCCCCCCCGGCCCCCCCCCGTGCTAAACCCACCGCGCATGCGCGACCACGCCCCCGCCGCC | 133 |
| TTV-HD23a | 全序列 | CGGCGGCGGCGGCG-CGCGCGCTGCGCGCGCG---CGCCGGGGGGGGCGCCAGCG-CCCCCCCCCCCCGCGCAT GCACGGGTCCCCCCCCCCCACGGGGGGCTCCG CCCCCCGGCCCCCCCCC | 134 |
| | 片段1 | CGGCGGCGGCGGCG | 135 |
| | 片段2 | CGCGCGCTGCGCGCGCG | 136 |
| | 片段3 | CGCCGGGGGGGGCGCCAGCG | 137 |
| | 片段4 | CCCCCCCCCCCCGCGCAT | 138 |
| | 片段5 | GCACGGGTCCCCCCCCCCCACGGGGGGCTCCG | 139 |
| | 片段6 | CCCCCCGGCCCCCCCCC | 140 |
| TTV-JA20 | 全序列 | CCGTCGGCGGGGGGGCCGCGCGCTGCGCGCGCGGCCC-CCGGGGGAGGCACAGCCTCCCCCCCCCGCGCGCATGCGCGCGGGTCCCCCCCCCTCCGGGGGGCTCCGCCCCCCGGCCCCCC | 141 |

| | | | |
|---|---|---|---|
| | | CC | |
| | 片段1 | CCGTCGGCGGGGGGGCCGCGCGCTGCG CGCGCGGCCC | 142 |
| | 片段2 | CCGGGGGAGGCACAGCCTCCCCCCCCC GCGCGCATGCGCGCGGGTCCCCCCCCCT CCGGGGGGCTCCGCCCCCCGGCCCCCC CC | 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-GCGGCGCGCTGCGCGCGCCGCCCAGTA GGGGGAGCCATGCG---CCCCCCCCCGCGCAT GCGCGGGGCCCCCCCCC-GCGGGGGGCTCCG CCCCCCGGCCCCCCCCG | 153 |
| | 片段1 | GCCGCCGCGGCGGCGGGGG | 154 |
| | 片段2 | GCGGCGCGCTGCGCGCGCCGCCCAGTA GGGGGAGCCATGCG | 155 |
| | 片段3 | CCCCCCCCCGCGCAT | 156 |
| | 片段4 | GCGCGGGGCCCCCCCCC | 157 |
| | 片段5 | GCGGGGGGCTCCG | 158 |
| | 片段6 | CCCCCCGGCCCCCCCCG | 159 |
| | 片段7 | CGCGCTGCGCGCGCCGCCCAGTAGGGG GAGCCATGC | 160 |
| | 片段8 | CCGCCATCTTAAGTAGTTGAGGCGGACG GTGGCGTGAGTTCAAAGGTCACCATCA GCCACACCTACTCAAAATGGTGG | 161 |
| | 片段9 | CTTAAGTAGTTGAGGCGGACGGTGGCG TGAGTTCAAAGGTCACCATCAGCCACA CCTACTCAAAATGGTGGACAATTTCTTC CGGGTCAAAGGTTACAGCCGCCATGTTA AAACACGTGACGTATGACGTCACGGCC GCCATTTTGTGACACAAGATGGCCGACT TCCTTCC | 162 |
| 額外GC富集 | 36個核苷酸之共 | CGCGCTGCGCGCGCCGCCCAGTAGGGG GAGCCATGC | 163 |

| | | | |
|---|---|---|---|
| 序列 | 有富含GC之區域序列1 | | |
| | 36個核苷酸之共有富含GC之區域序列2 | GCGCTX$_1$CGCGCGCGCGCCGGGGGGCTGCGCCCCCCC，其中X$_1$係選自T、G或A | 164 |
| | TTV分枝系1 36個核苷酸之區域 | GCGCTTCGCGCGCCGCCCACTAGGGGGCGTTGCGCG | 165 |
| | TTV分枝系3 36個核苷酸之區域 | GCGCTGCGCGCGCCGCCCAGTAGGGGGCGCAATGCG | 166 |
| | TTV分枝系3分離株GH1 36個核苷酸之區域 | GCGCTGCGCGCGCGGCCCCCGGGGGAGGCATTGCCT | 167 |
| | TTV分枝系3 sle1932 36個核苷酸之區域 | GCGCTGCGCGCGCGCCGGGGGGGCGCCAGCGCCC | 168 |
| | TTV分枝系4 ctdc002 36個核苷酸之區域 | GCGCTTCGCGCGCGCCGGGGGGCTCCGCCCCCCC | 169 |
| | TTV分枝系5 36個核苷酸之區域 | GCGCTTCGCGCGCGCCGGGGGGCTGCGCCCCCCC | 170 |
| | TTV分枝系6 36個核苷酸之區域 | GCGCTACGCGCGCGCCGGGGGGCTGCGCCCCCCC | 171 |
| | TTV分枝系7 36個核苷酸之區域 | GCGCTACGCGCGCGCCGGGGGGCTCTGCCCCCCC | 172 |
| 額外α細環病毒屬富含GC之區域序列 | TTV-CT30F | GCGGCGGGGGGGCGGCCGCGTTCGCGCGCCGCCCACCAGGGGGTGCTGCGCGCCCCCCCCCGCGCATGCGCGGGGCCCCCCCCCCGGGGGGGCTCCGCCCCCCCGGCCCCCCCCGTGCTAAACCCACCGCGCATGCGCGACCACGCCCCCGCCGCC | 801 |
| | TTV-P13-1 | CCGAGCGTTAGCGAGGAGTGCGACCCTACCCCCTGGGCCCACTTCTTCGGAGCCGCGCGCTACGCCTTCGGCTGCGCGCGGCACCTCAGACCCCCGCTCGTGCTGACACGCTTGCGCGTGTCAGACCACTTCGGGCTCGCGGGGGTCGGG | 802 |
| | TTV-tth8 | GCCGCCGCGGCGGCGGGGGGCGGCGCGCTGCGCGCGCCGCCCAGTAGGGGGAGCCATGCGCCCCCCCCCGCGCATGCGCGGGGCCCCCCCCCCGCGGGGGGCTCCGCCCCCCGGCCCCCCCCG | 803 |
| | TTV-HD20a | CGGCCCAGCGGCGGCGCGCGCGCTTCGCGCGCGCCGGGGGGCTCCGCCCCCCCCCGCGCATGCGCGGGGCCCCCCCCCGCGGGGGGCTCCGCCCCCCGGTCCCCCCCCG | 804 |
| | TTV-16 | CGGCCGTGCGGCGGCGCGCGCGCTTCG | 805 |

| | | | |
|---|---|---|---|
| | | CGCGCGCGCCGGGGGCTGCCGCCCCCC<br>CCCGCGCATGCGCGCGGGGCCCCCCCC<br>CGCGGGGGGCTCCGCCCCCCGGCCCCC<br>CCCCCCG | |
| | TTV-TJN02 | CGGCGGCGGCGCGCGCGCTACGCGCGC<br>GCGCCGGGGGGCTGCCGCCCCCCCCC<br>GCGCATGCGCGGGGCCCCCCCCCGCGG<br>GGGGCTCCGCCCCCCGGCCCCCC | 806 |
| | TTV-HD16d | GGCGGCGGCGCGCGCGCTACGCGCGCG<br>CGCCGGGGAGCTCTGCCCCCCCCCGCG<br>CATGCGCGCGGGTCCCCCCCCCGCGGG<br>GGGCTCCGCCCCCCGGTCCCCCCCCCG | 807 |

【0206】

*效應子*

在一些實施例中，遺傳元件可包括一或多個作為效應子或編碼效應子，例如功能性效應子，例如內源性效應子或外源性效應子，例如治療性多肽或核酸，例如細胞毒性或細胞溶解RNA或蛋白質的序列。在一些實施例中，功能性核酸為非編碼RNA。在一些實施例中，功能性核酸為編碼RNA。效應子可調節生物活性，例如提高或降低酶活性、基因表現、細胞信號傳導及細胞或器官功能。效應子活性亦可包括結合調控蛋白以調節調節因子之活性，諸如轉錄或轉譯。效應子活性亦可包括活化子或抑制劑功能。舉例而言，效應子可藉由觸發酶中受質親和力的提高來誘導酶活性，例如果糖2,6-雙磷酸活化磷酸果糖激酶1且提高響應於胰島素之糖酵解速率。在另一實例中，效應子可抑制受質與受體結合且抑制其活化，例如納曲酮(naltrexone)及納洛酮(naloxone)在不活化類鴉片受體之情況下結合類鴉片受體且阻斷受體結合類鴉片之能力。效應子活性亦可包括調節蛋白質穩定性/降解及/或轉錄本穩定性/降解。舉例而言，蛋白質可經靶向以藉由多肽輔因子泛素降解至蛋白質上，以標記蛋白質用於降解。在另一實例中，效應子藉由阻斷酶活性位點來抑制酶活性，例如甲胺喋呤為四氫葉酸之結構類似物，該四氫葉酸為酶二氫葉酸還原酶之輔酶，與二氫葉酸還

原酶之結合比天然受質更緊密1000倍，且抑制核苷酸鹼基合成。

【0207】 在一些實施例中，編碼效應子之序列包含100-2000、100-1000、100-500、100-200、200-2000、200-1000、200-500、500-1000、500-2000或1000-2000個核苷酸。在一些實施例中，效應子為例如如本文所述之核酸或蛋白質有效負載。

【0208】

*調控核酸*

在一些實施例中，效應子為調控核酸。調控核酸修飾內源性基因及/或外源性基因之表現。在一個實施例中，調控核酸靶向宿主基因。調控核酸可包括但不限於與內源性基因雜交之核酸(例如如本文其他地方所述之miRNA、siRNA、mRNA、lncRNA、RNA、DNA、反義RNA、gRNA)、與諸如病毒DNA或RNA之外源性核酸雜交之核酸、與RNA雜交之核酸、干擾基因轉錄之核酸、干擾RNA轉譯之核酸、諸如經由靶向而使RNA穩定或使RNA不穩定以進行降解之核酸、及調節DNA或RNA結合因子之核酸。在實施例中，調控核酸編碼miRNA。在一些實施例中，調控核酸對於野生型指環病毒為內源性的。在一些實施例中，調控核酸對於野生型指環病毒為外源性的。

【0209】 在一些實施例中，調控核酸包含通常含有5-500個鹼基對之RNA或RNA樣結構(視特定RNA結構而定，例如miRNA 5-30 bp、lncRNA 200-500 bp)，且可具有與細胞內表現之目標基因中之編碼序列或編碼細胞內表現之目標基因之序列一致(或互補)或幾乎一致(或實質上互補)的核鹼基序列。

【0210】 在一些實施例中，調控核酸包含核酸序列，例如嚮導RNA

(gRNA)。在一些實施例中，DNA靶向部分包含嚮導RNA或編碼嚮導RNA之核酸。gRNA短合成RNA由與不完全效應部分結合所必需的「骨架」序列及使用者定義之基因體目標之約20個核苷酸靶向序列構成。實際上，嚮導RNA序列通常經設計以具有17-24個核苷酸(例如19、20或21個核苷酸)之間的長度且與所靶向核酸序列互補。定製gRNA產生器及演算法可商業購得，用於設計有效嚮導RNA。亦使用嵌合「單嚮導RNA」(「sgRNA」)實現基因編輯，該單嚮導RNA為模擬天然存在之crRNA-tracrRNA複合物且含有tracrRNA (用於結合核酸酶)及至少一種crRNA (以將核酸酶導引至經靶向以進行編輯之序列)的經工程改造(合成)之單一RNA分子。亦已證明經化學修飾之sgRNA在基因體編輯中有效；參見例如Hendel等人 (2015) Nature Biotechnol., 985 - 991。

【0211】 調控核酸包含識別特異性DNA序列(例如鄰近於基因之啟動子、強化子、緘默子或抑制子或在其內之序列)的gRNA。

【0212】 某些調控核酸可經由RNA干擾(RNAi)之生物過程抑制基因表現。RNAi分子包含RNA或RNA樣結構，該等結構通常含有15-50個鹼基對(諸如約18-25個鹼基對)且具有與細胞內所表現之目標基因中之編碼序列一致(互補)或幾乎一致(實質上互補)的核鹼基序列。RNAi分子包括但不限於：短干擾RNA (siRNA)、雙股RNA (dsRNA)、微小RNA (miRNA)、短髮夾RNA (shRNA)、部分雙螺旋體及切丁酶受質(美國專利第8,084,599號、第8,349,809號及第8,513,207號)。

【0213】 長的非編碼RNA (lncRNA)被定義為長於100個核苷酸之非蛋白質編碼轉錄本。此略微任意之限制將lncRNA與較小調控性RNA (諸如微小RNA (miRNA)、短干擾RNA (siRNA)及其他短RNA)區分開來。一

般而言，大部分(約78%)之lncRNA被描述為組織特異性的。在與附近蛋白質編碼基因相反之方向上轉錄之發散lncRNA (佔哺乳動物基因體中總lncRNA之顯著比例，約20%)可能調控鄰近基因之轉錄。

【0214】 遺傳元件可編碼具有與內源性基因或基因產物(例如mRNA)之全部或片段實質上互補或完全互補之序列的調控核酸。調控核酸可與內含子與外顯子之間的邊界處之序列互補，以防止特定基因之新產生的細胞核RNA轉錄本成熟為mRNA進行轉錄。與特定基因互補之調控核酸可與該基因之mRNA雜交且防止其轉譯。反義調控核酸可為DNA、RNA或其衍生物或雜合體。

【0215】 雜交至所關注轉錄本之調控核酸之長度可為5至30個核苷酸、約10至30個核苷酸或約11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30或更多個核苷酸。調控核酸與所靶向轉錄本之一致性程度應為至少75%、至少80%、至少85%、至少90%或至少95%。

【0216】 遺傳元件可編碼調控核酸，例如與目標基因之約5至約25個連續核苷酸一致的微小RNA (miRNA)分子。在一些實施例中，miRNA序列靶向mRNA且以二核苷酸AA開始，包含約30-70% (約30-60%、約40-60%或約45-55%)之GC含量，且與欲引入之哺乳動物基因體中除目標以外之任何核苷酸序列不具有高比例一致性，例如如藉由標準BLAST搜尋所確定。

【0217】 siRNA及shRNA類似內源性微小RNA (miRNA)基因之加工路徑中之中間物(Bartel, Cell 116:281-297, 2004)。在一些實施例中，siRNA可充當miRNA且反之亦然(Zeng等人, Mol Cell 9:1327-1333,

2002；Doench等人, Genes Dev 17:438-442, 2003)。微小RNA，如siRNA一般，使用RISC下調目標基因，但不同於siRNA，大部分動物miRNA不使mRNA裂解。實際上，miRNA經由轉譯遏制或多聚A移除及mRNA降解來減少蛋白質輸出(Wu等人, Proc Natl Acad Sci USA 103:4034-4039, 2006)。已知miRNA結合位點在mRNA 3' UTR內；miRNA似乎靶向與來自miRNA 5'端之核苷酸2-8具有接近完美的互補性之位點(Rajewsky, Nat Genet 38 增刊:S8-13, 2006；Lim等人, Nature 433:769-773, 2005)。此區域稱為種子區。由於siRNA及miRNA為可互換的，因此外源性siRNA下調與siRNA具有種子互補性之mRNA (Birmingham等人, Nat Methods 3:199-204, 2006。3' UTR內之多個目標位點提供較強下調(Doench等人, Genes Dev 17:438-442, 2003)。

【0218】已知miRNA序列之清單可見於諸如以下之研究組織所維護之資料庫中：Wellcome Trust Sanger Institute、Penn Center for Bioinformatics、Memorial Sloan Kettering Cancer Center及European Molecule Biology Laboratory。已知的有效siRNA序列及同源結合位點亦充分呈現於相關文獻中。RNAi分子容易藉由此項技術中已知之技術設計及產生。此外，存在增加發現有效及特定序列模體之機率的計算工具(Lagana等人, Methods Mol. Bio., 2015, 1269:393-412)。

【0219】調控核酸可調節由基因編碼之RNA的表現。因為多個基因可彼此共有一定程度之序列同源性，所以在一些實施例中，調控核酸可經設計以靶向具有足夠序列同源性之一類基因。在一些實施例中，調控核酸可含有與在不同基因目標中共有或特定基因目標所獨有之序列互補的序列。在一些實施例中，調控核酸可經設計以靶向在若干基因之間具有同源

性之RNA序列的保守區，藉此靶向基因家族中之若干基因(例如不同基因同功異型物、剪接變異體、突變基因等)。在一些實施例中，調控核酸可經設計以靶向為單個基因之特定RNA序列所獨有的序列。

【0220】 在一些實施例中，遺傳元件可包括一或多個編碼調節一或多個基因之表現之調控核酸的序列。

【0221】 在一個實施例中，本文他處所述之gRNA用作用於基因編輯之CRISPR系統的一部分。出於基因編輯之目的，指環載體可經設計以包括一或多個對應於所需目標DNA序列的嚮導RNA序列；參見例如Cong等人 (2013) Science, 339:819-823；Ran等人 (2013) Nature Protocols, 8:2281 - 2308。gRNA序列之至少約16或17個核苷酸通常允許發生Cas9介導之DNA裂解；對於Cpf1，需要gRNA序列之至少約16個核苷酸來實現可偵測DNA裂解。

【0222】

*治療性效應子(例如肽或多肽)*

在一些實施例中，遺傳元件包含治療性表現序列，例如編碼治療性肽或多肽之序列，例如胞內肽或胞內多肽、分泌型多肽或蛋白質替代療法。在一些實施例中，遺傳元件包括編碼蛋白質(例如治療性蛋白)之序列。治療性蛋白質之一些實例可包括但不限於激素、細胞介素、酶、抗體(例如編碼至少重鏈或輕鏈之一種或複數種多肽)、轉錄因子、受體(例如膜受體)、配位體、膜轉運體、分泌型蛋白質、肽、載體蛋白、結構蛋白、核酸酶或其組分。

【0223】 在一些實施例中，遺傳元件包括編碼肽(例如治療性肽)之序列。肽可為線性的或分支的。肽之長度為約5至約500個胺基酸、約15

至約400個胺基酸、約20至約325個胺基酸、約25至約250個胺基酸、約50至約200個胺基酸或其間的任何範圍。

【0224】 在一些實施例中，由治療性表現序列編碼之多肽可為以上中之任一者之功能性變異體或其片段，例如與參考其UniProt ID揭示於本文之表中之蛋白質序列具有至少80%、85%、90%、95%、967%、98%、99%之一致性的蛋白質。

【0225】 在一些實施例中，治療性表現序列可編碼結合以上中之任一者之抗體或抗體片段，例如針對與參考其UniProt ID揭示於本文之表中之蛋白質序列具有至少80%、85%、90%、95%、967%、98%、99%之一致性的蛋白質的抗體。術語「抗體」在本文中以最廣泛意義使用且涵蓋各種抗體結構，包括但不限於單株抗體、多株抗體、多特異性抗體(例如雙特異性抗體)及抗體片段，只要其呈現所需抗原結合活性即可。「抗體片段」係指包括至少一個重鏈或輕鏈且結合抗原之分子。抗體片段之實例包括但不限於Fv、Fab、Fab'、Fab'-SH、$F(ab')_2$；雙功能抗體；線性抗體；單鏈抗體分子(例如scFv)；及由抗體片段形成之多特異性抗體。

【0226】

**例示性胞內多肽效應子**

在一些實施例中，效應子包含細胞溶質多肽或細胞溶質肽。在一些實施例中，效應子包含細胞溶質肽，其為DPP-4抑制劑、GLP-1信號傳導之活化劑或嗜中性球彈性蛋白酶之抑制劑。在一些實施例中，效應子提高生長因子或其受體(例如FGF受體，例如FGFR3)之含量或活性。在一些實施例中，效應子包含n-myc相互作用蛋白質活性之抑制劑(例如，n-myc相互作用蛋白質抑制劑)；EGFR活性之抑制劑(例如，EGFR抑制劑)；IDH1

及/或IDH2活性之抑制劑(例如，IDH1抑制劑及/或IDH2抑制劑)；LRP5及/或DKK2活性之抑制劑(例如，LRP5及/或DKK2抑制劑)；KRAS活性之抑制劑；HTT活性之活化劑；或DPP-4活性之抑制劑(例如，DPP-4抑制劑)。

【0227】 在一些實施例中，效應子包含調控性胞內多肽。在一些實施例中，調控性胞內多肽結合對於目標細胞為內源性的一或多個分子(例如蛋白質或核酸)。在一些實施例中，調控性胞內多肽提高對於目標細胞為內源性的一或多個分子(例如蛋白質或核酸)的含量或活性。在一些實施例中，調控性胞內多肽減小對於目標細胞為內源性的一或多個分子(例如蛋白質或核酸)的含量或活性。

【0228】

**例示性分泌型多肽效應子**

例示性分泌型治療劑描述於本文中，例如下表中。

**表50.例示性細胞介素及細胞介素受體**

| 細胞介素 | 細胞介素受體 | Entrez基因ID | UniProt ID |
|---|---|---|---|
| IL-1α、IL-1β或其雜二聚體 | IL-1 1型受體、IL-1 2型受體 | 3552, 3553 | P01583, P01584 |
| IL-1Ra | IL-1 1型受體、IL-1 2型受體 | 3454, 3455 | P17181, P48551 |
| IL-2 | IL-2R | 3558 | P60568 |
| IL-3 | IL-3受體α + β c (CD131) | 3562 | P08700 |
| IL-4 | IL-4R I型，IL-4R II型 | 3565 | P05112 |
| IL-5 | IL-5R | 3567 | P05113 |
| IL-6 | IL-6R (sIL-6R) gp130 | 3569 | P05231 |
| IL-7 | IL-7R及sIL-7R | 3574 | P13232 |
| IL-8 | CXCR1及CXCR2 | 3576 | P10145 |
| IL-9 | IL-9R | 3578 | P15248 |
| IL-10 | IL-10R1/IL-10R2複合物 | 3586 | P22301 |
| IL-11 | IL-11Rα 1 gp130 | 3589 | P20809 |
| IL-12 (例如p19、p40或其雜二聚體) | IL-12Rβ1及IL-12Rβ2 | 3593, 3592 | P29459, P29460 |
| IL-13 | IL-13R1α1及IL-13R1α2 | 3596 | P35225 |

| | | | |
|---|---|---|---|
| IL-14 | IL-14R | 30685 | P40222 |
| IL-15 | IL-15R | 3600 | P40933 |
| IL-16 | CD4 | 3603 | Q14005 |
| IL-17A | IL-17RA | 3605 | Q16552 |
| IL-17B | IL-17RB | 27190 | Q9UHF5 |
| IL-17C | IL-17RA至IL-17RE | 27189 | Q9P0M4 |
| e | SEF | 53342 | Q8TAD2 |
| IL-17F | IL-17RA, IL-17RC | 112744 | Q96PD4 |
| IL-18 | IL-18受體 | 3606 | Q14116 |
| IL-19 | IL-20R1/IL-20R2 | 29949 | Q9UHD0 |
| IL-20 | L-20R1/IL-20R2及IL-22R1/ IL-20R2 | 50604 | Q9NYY1 |
| IL-21 | IL-21R | 59067 | Q9HBE4 |
| IL-22 | IL-22R | 50616 | Q9GZX6 |
| IL-23 (例如p19、p40或其雜二聚體) | IL-23R | 51561 | Q9NPF7 |
| IL-24 | IL-20R1/IL-20R2及IL-22R1/ IL-20R2 | 11009 | Q13007 |
| IL-25 | IL-17RA及IL-17RB | 64806 | Q9H293 |
| IL-26 | IL-10R2鏈及IL-20R1鏈 | 55801 | Q9NPH9 |
| IL-27 (例如p28、EBI3或其雜二聚體) | WSX-1及gp130 | 246778 | Q8NEV9 |
| IL-28A、IL-28B及IL29 | IL-28R1/IL-10R2 | 282617, 282618 | Q8IZI9, Q8IU54 |
| IL-30 | IL6R/gp130 | 246778 | Q8NEV9 |
| IL-31 | IL-31RA/OSMRβ | 386653 | Q6EBC2 |
| IL-32 | | 9235 | P24001 |
| IL-33 | ST2 | 90865 | O95760 |
| IL-34 | 群落刺激因子1受體 | 146433 | Q6ZMJ4 |
| IL-35 (例如p35、EBI3或其雜二聚體) | IL-12Rβ2/gp130; IL-12Rβ2/IL-12Rβ2; gp130/gp130 | 10148 | Q14213 |
| IL-36 | IL-36Ra | 27179 | Q9UHA7 |
| IL-37 | IL-18Rα及IL-18BP | 27178 | Q9NZH6 |
| IL-38 | IL-1R1, IL-36R | 84639 | Q8WWZ1 |
| IFN-α | IFNAR | 3454 | P17181 |
| IFN-β | IFNAR | 3454 | P17181 |
| IFN-γ | IFNGR1/IFNGR2 | 3459 | P15260 |
| TGF-β | TβR-I及TβR-II | 7046, 7048 | P36897, P37173 |
| TNF-α | TNFR1, TNFR2 | 7132, 7133 | P19438, P20333 |

【0229】在一些實施例中，本文所述之效應子包含表50之細胞介素或其功能性變異體，例如同源物(例如直系同源物或旁系同源物)或其片段。在一些實施例中，本文所述之效應子包含與參考其UniProt ID在表50中所列之胺基酸序列具有至少80%、85%、90%、95%、967%、98%、

99%序列一致性的蛋白質。在一些實施例中，功能性變異體以比在相同條件下對應野生型細胞介素對相同受體之Kd高或低不超過10%、20%、30%、40%或50%之Kd結合至對應細胞介素受體。在一些實施例中，效應子包含融合蛋白質，其包含第一區域(例如表50之細胞介素多肽或其功能性變異體或片段)及第二異源區域。在一些實施例中，第一區域為表50之第一細胞介素多肽。在一些實施例中，第二區域為表50之第二細胞介素多肽，其中第一及第二細胞介素多肽在野生型細胞中彼此形成細胞介素雜二聚體。在一些實施例中，表50之多肽或其功能性變異體包含信號序列，例如對於效應子為內源性的信號序列，或異源性信號序列。在一些實施例中，編碼表50之細胞介素的指環載體或其功能性變異體用於治療本文所述之疾病或病症。

【0230】 在一些實施例中，本文所述之效應子包含結合表50之細胞介素的抗體分子(例如scFv)。在一些實施例中，本文所述之效應子包含結合表50之細胞介素受體的抗體分子(例如scFv)。在一些實施例中，抗體分子包含信號序列。

【0231】 例示性細胞介素及細胞介素受體描述於例如Akdis等人，「Interleukins (from IL-1 to IL-38), interferons, transforming growth factor β, and TNF-α: Receptors, functions, and roles in diseases」2016年10月 第138卷, 第4期, 第984-1010頁中，其以全文引用之方式併入本文中，包括其中之表I。

**表51.例示性多肽激素及受體**

| 激素 | 受體 | Entrez基因ID | UniProt ID |
|---|---|---|---|
| 利尿鈉肽，例如心房利尿鈉肽(ANP) | NPRA, NPRB, NPRC | 4878 | P01160 |
| 腦尿鈉肽(BNP) | NPRA, NPRB | 4879 | P16860 |
| C型利尿鈉肽(CNP) | NPRB | 4880 | P23582 |
| 生長激素(GH) | GHR | 2690 | P10912 |
| 人類生長激素(hGH) | hGH受體(人類GHR) | 2690 | P10912 |
| 催乳素(PRL) | PRLR | 5617 | P01236 |
| 甲狀腺刺激激素(TSH) | TSH受體 | 7253 | P16473 |
| 促腎上腺皮質激素(ACTH) | ACTH受體 | 5443 | P01189 |
| 促卵泡激素(FSH) | FSHR | 2492 | P23945 |
| 促黃體生成激素(LH) | LHR | 3973 | P22888 |
| 抗利尿激素(ADH) | 血管加壓素受體，例如V2；AVPR1A；AVPR1B；AVPR3；AVPR2 | 554 | P30518 |
| 催產素 | OXTR | 5020 | P01178 |
| 降鈣素 | 降鈣素受體(CT) | 796 | P01258 |
| 副甲狀腺激素(PTH) | PTH1R及PTH2R | 5741 | P01270 |
| 胰島素 | 胰島素受體(IR) | 3630 | P01308 |
| 升糖素 | 升糖素受體 | 2641 | P01275 |

【0232】 在一些實施例中，本文所述之效應子包含表51之激素或其功能性變異體，例如同源物(例如直系同源物或旁系同源物)或其片段。在一些實施例中，本文所述之效應子包含與參考其UniProt ID在表51中所列之胺基酸序列具有至少80%、85%、90%、95%、967%、98%、99%序列一致性的蛋白質。在一些實施例中，功能性變異體以比在相同條件下對應野生型激素對相同受體之Kd高不超過10%、20%、30%、40%或50%之Kd結合至對應受體。在一些實施例中，表51之多肽或其功能性變異體包含信號序列，例如對於效應子為內源性的信號序列，或異源性信號序列。在一些實施例中，編碼表51之激素或其功能性變異體的指環載體用於治療本文所述之疾病或病症。

【0233】 在一些實施例中，本文所述之效應子包含結合表51之激素

的抗體分子(例如scFv)。在一些實施例中，本文所述之效應子包含結合表51之激素受體的抗體分子(例如scFv)。在一些實施例中，抗體分子包含信號序列。

**表52.例示性生長因子**

| 生長因子 | | Entrez基因ID | UniProt ID |
|---|---|---|---|
| PDGF家族 | | | |
| PDGF (例如，PDGF-1、PDGF-2或其雜二聚體) | PDGF受體，例如PDGFRα、PDGFRβ | 5156 | P16234 |
| CSF-1 | CSF1R | 1435 | P09603 |
| SCF | CD117 | 3815 | P10721 |
| VEGF家族 | | | |
| VEGF (例如同功異型物VEGF 121、VEGF 165、VEGF 189及VEGF 206) | VEGFR-1, VEGFR-2 | 2321 | P17948 |
| VEGF-B | VEGFR-1 | 2321 | P17949 |
| VEGF-C | VEGFR-2及VEGFR-3 | 2324 | P35916 |
| PlGF | VEGFR-1 | 5281 | Q07326 |
| EGF家族 | | | |
| EGF | EGFR | 1950 | P01133 |
| TGF-α | EGFR | 7039 | P01135 |
| 雙調蛋白 | EGFR | 374 | P15514 |
| HB-EGF | EGFR | 1839 | Q99075 |
| β細胞調節素 | EGFR, ErbB-4 | 685 | P35070 |
| 表皮調節素 | EGFR, ErbB-4 | 2069 | O14944 |
| 調蛋白 | EGFR, ErbB-4 | 3084 | Q02297 |
| FGF家族 | | | |
| FGF-1, FGF-2, FGF-3, FGF-4, FGF-5, FGF-6, FGF-7, FGF-8, FGF-9 | FGFR1、FGFR2、FGFR3及FGFR4 | 2246, 2247, 2248, 2249, 2250, 2251, 2252, 2253, 2254 | P05230, P09038, P11487, P08620, P12034, P10767, P21781, P55075, P31371 |
| 胰島素家族 | | | |
| 胰島素 | IR | 3630 | P01308 |
| IGF-I | IGF-I受體，IGF-II受體 | 3479 | P05019 |
| IGF-II | IGF-II受體 | 3481 | P01344 |
| HGF家族 | | | |
| HGF | MET受體 | 3082 | P14210 |
| MSP | RON | 4485 | P26927 |
| 神經滋養素家族 | | | |
| NGF | LNGFR, trkA | 4803 | P01138 |

| | | | |
|---|---|---|---|
| BDNF | trkB | 627 | P23560 |
| NT-3 | trkA, trkB, trkC | 4908 | P20783 |
| NT-4 | trkA, trkB | 4909 | P34130 |
| NT-5 | trkA, trkB | 4909 | P34130 |
| 血管生成素家族 | | | |
| ANGPT1 | HPK-6/TEK | 284 | Q15389 |
| ANGPT2 | HPK-6/TEK | 285 | O15123 |
| ANGPT3 | HPK-6/TEK | 9068 | O95841 |
| ANGPT4 | HPK-6/TEK | 51378 | Q9Y264 |

【0234】 在一些實施例中，本文所述之效應子包含表52之生長因子或其功能性變異體，例如同源物(例如直系同源物或旁系同源物)或其片段。在一些實施例中，本文所述之效應子包含與參考其UniProt ID在表52中所列之胺基酸序列具有至少80%、85%、90%、95%、967%、98%、99%序列一致性的蛋白質。在一些實施例中，功能性變異體以比在相同條件下對應野生型生長因子對相同受體之Kd高不超過10%、20%、30%、40%或50%之Kd結合至對應受體。在一些實施例中，表52之多肽或其功能性變異體包含信號序列，例如對於效應子為內源性的信號序列，或異源性信號序列。在一些實施例中，編碼表52之生長因子或其功能性變異體的指環載體用於治療本文所述之疾病或病症。

【0235】 在一些實施例中，本文所述之效應子包含結合表52之生長因子的抗體分子(例如scFv)。在一些實施例中，本文所述之效應子包含結合表52之生長因子受體的抗體分子(例如scFv)。在一些實施例中，抗體分子包含信號序列。

【0236】 例示性生長因子及生長因子受體描述於例如Bafico等人，「Classification of Growth Factors and Their Receptors」 Holland-Frei Cancer Medicine. 第6版中，其以全文引用之方式併入本文中。

**表53.凝結相關因子**

| 效應子 | 適應症 | Entrez基因ID | UniProt ID |
|---|---|---|---|
| 因子I (血纖維蛋白原) | 無纖維蛋白原血症 | 2243, 2266, 2244 | P02671, P02679, P02675 |
| 因子II | 因子II缺乏症 | 2147 | P00734 |
| 因子IX | B型血友病 | 2158 | P00740 |
| 因子V | 奧倫氏病(Owren's disease) | 2153 | P12259 |
| 因子VIII | A型血友病 | 2157 | P00451 |
| 因子X | 斯圖亞特因子缺乏症(Stuart-Prower Factor Deficiency) | 2159 | P00742 |
| 因子XI | C型血友病 | 2160 | P03951 |
| 因子XIII | 纖維蛋白穩定因子缺乏症 | 2162, 2165 | P00488, P05160 |
| vWF | 馮威里氏病(von Willebrand disease) | 7450 | P04275 |

【0237】 在一些實施例中，本文所述之效應子包含表53之多肽或其功能性變異體，例如同源物(例如直系同源物或旁系同源物)或其片段。在一些實施例中，本文所述之效應子包含與參考其UniProt ID在表53中所列之胺基酸序列具有至少80%、85%、90%、95%、967%、98%、99%序列一致性的蛋白質。在一些實施例中，功能性變異體催化與對應野生型蛋白質相同的反應，例如以比野生型蛋白質低不少於10%、20%、30%、40%或50%之速率。在一些實施例中，表53之多肽或其功能性變異體包含信號序列，例如對於效應子為內源性的信號序列，或異源性信號序列。在一些實施例中，編碼表53之多肽的指環載體或其功能性變異體用於治療表53之疾病或病症。

【0238】

**例示性蛋白質置換治療劑**

例示性蛋白質置換治療劑描述於本文中，例如下表中。

## 表54.例示性酶效應子及對應適應症

| 效應子 | 缺乏症 | Entrez基因ID | UniProt ID |
|---|---|---|---|
| 3-甲基巴豆醯基-CoA羧化酶 | 3-甲基巴豆醯基-CoA羧化酶缺乏症 | 56922, 64087 | Q96RQ3, Q9HCC0 |
| 乙醯基-CoA-胺基葡糖苷N-乙醯基轉移酶 | 黏多糖病MPS III (聖菲利柏氏症候群(Sanfilippo's syndrome)) III-C型 | 138050 | Q68CP4 |
| ADAMTS13 | 血栓性血小板減少性紫癜 | 11093 | Q76LX8 |
| 腺嘌呤磷酸核糖轉移酶 | 腺嘌呤磷酸核糖轉移酶缺乏症 | 353 | P07741 |
| 腺苷去胺酶 | 腺苷去胺酶缺乏症 | 100 | P00813 |
| ADP-核糖蛋白質水解酶 | 麩胺醯基核糖-5-磷酸貯積病 | 26119, 54936 | Q5SW96, Q9NX46 |
| α葡糖苷酶 | 2型肝糖貯積病(龐貝氏病(Pompe's disease)) | 2548 | P10253 |
| 精胺酸酶 | 家族性高精胺酸血症 | 383, 384 | P05089, P78540 |
| 芳基硫酸酯酶A | 異染性腦白質營養不良 | 410 | P15289 |
| 組織蛋白酶K | 密骨發育障礙 | 1513 | P43235 |
| 神經醯胺酶 | 法伯氏病(Farber's disease)(脂肪肉芽腫病) | 125981, 340485, 55331 | Q8TDN7, Q5QJU3, Q9NUN7 |
| 胱硫醚B合成酶 | 高胱胺酸尿 | 875 | P35520 |
| 多萜醇-P-甘露糖合成酶 | 先天性N-醣基化障礙CDG Ie | 8813, 54344 | O60762, Q9P2X0 |
| 多萜醇-P-Glc:Man9GlcNAc2-PP-多萜醇葡萄糖基轉移酶 | 先天性N-醣基化障礙CDG Ic | 84920 | Q5BKT4 |
| 多萜醇-P-Man:Man5GlcNAc2-PP-多萜醇甘露糖基轉移酶 | 先天性N-糖基化障礙CDG Id | 10195 | Q92685 |
| 多萜醇基-P-葡萄糖:Glc-1-Man-9-GlcNAc-2-PP-多萜醇基-α-3-葡萄糖基轉移酶 | 先天性N-醣基化障礙CDG Ih | 79053 | Q9BVK2 |
| 多萜醇基-P-甘露糖:Man-7-GlcNAc-2-PP-多萜醇基-α-6-甘露糖基轉移酶 | 先天性N-醣基化障礙CDG Ig | 79087 | Q9BV10 |
| 因子II | 因子II缺乏症 | 2147 | P00734 |
| 因子IX | B型血友病 | 2158 | P00740 |
| 因子V | 奧倫氏病 | 2153 | P12259 |

| | | | |
|---|---|---|---|
| 因子VIII | A型血友病 | 2157 | P00451 |
| 因子X | 斯圖亞特因子缺乏症 | 2159 | P00742 |
| 因子XI | C型血友病 | 2160 | P03951 |
| 因子XIII | 纖維蛋白穩定因子缺乏症 | 2162, 2165 | P00488, P05160 |
| 半乳胺糖-6-硫酸酯硫酸酯酶 | 黏多糖病MPS IV (莫奎歐氏症候群(Morquio's syndrome)) IV-A型 | 2588 | P34059 |
| 半乳糖基神經醯胺β-半乳糖苷酶 | 克拉伯氏病(Krabbe's disease) | 2581 | P54803 |
| 神經節苷脂β-半乳糖苷酶 | 全身性GM1神經節苷脂貯積病 | 2720 | P16278 |
| 神經節苷脂β-半乳糖苷酶 | GM2神經節苷脂貯積病 | 2720 | P16278 |
| 神經節苷脂β-半乳糖苷酶 | I型神經鞘脂貯積症 | 2720 | P16278 |
| 神經節苷脂β-半乳糖苷酶 | II型神經鞘脂貯積症(幼年型) | 2720 | P16278 |
| 神經節苷脂β-半乳糖苷酶 | III型神經鞘脂貯積症(成人型) | 2720 | P16278 |
| 葡糖苷酶I | 先天性N-糖基化障礙CDG IIb | 2548 | P10253 |
| 葡糖神經醯胺β-葡糖苷酶 | 高歇氏病(Gaucher's disease) | 2629 | P04062 |
| 乙醯肝素-S-硫酸酯硫酸醯胺酶 | 黏多糖病MPS III (聖菲利柏氏症候群) III-A型 | 6448 | P51688 |
| 尿黑酸氧化酶 | 黑尿症 | 3081 | Q93099 |
| 玻尿酸酶 | 黏多糖病MPS IX (玻尿酸酶缺乏症) | 3373, 8692, 8372, 23553 | Q12794, Q12891, O43820, Q2M3T9 |
| 艾杜糖醛硫酸酯硫酸酯酶 | 黏多糖病MPS II (亨特氏症候群(Hunter's syndrome)) | 3423 | P22304 |
| 卵磷脂-膽固醇醯基轉移酶(LCAT) | 完全LCAT缺乏症、魚眼病、動脈粥樣硬化、高膽固醇血症 | 3931 | 606967 |
| 離胺酸氧化酶 | 戊二酸血症I型 | 4015 | P28300 |
| 溶酶體酸性脂肪酶 | 膽甾醇酯貯積病(CESD) | 3988 | P38571 |
| 溶酶體酸性脂肪酶 | 溶酶體酸性脂肪酶缺乏症 | 3988 | P38571 |
| 溶酶體酸性脂肪酶 | 沃爾曼氏病(Wolman's disease) | 3988 | P38571 |
| 溶酶體胃酶抑素不敏感性肽酶 | 幼兒型蠟樣脂褐質沈積症(CLN2、詹-比二氏病(Jansky-Bielschowsky disease)) | 1200 | O14773 |
| 甘露糖(Man)磷酸酯(P)異構酶 | 先天性N-糖基化障礙CDG Ib | 4351 | P34949 |

| | | | |
|---|---|---|---|
| 甘露糖基-α-1,6-糖蛋白-β-1,2-N-乙醯葡糖胺基轉移酶 | 先天性N-糖基化障礙CDG IIa | 4247 | Q10469 |
| 金屬蛋白酶-2 | 溫徹斯特症候群 (Winchester syndrome) | 4313 | P08253 |
| 甲基丙二醯基-CoA變位酶 | 甲基丙二酸血症(維生素b12無反應型) | 4594 | P22033 |
| N-乙醯基半乳糖胺α-4-硫酸酯硫酸酯酶(芳基硫酸酯酶B) | 黏多糖病MPS VI (馬洛特-拉米症候群(Maroteaux-Lamy syndrome)) | 411 | P15848 |
| N-乙醯基-D-胺基葡糖苷酶 | 黏多糖病MPS III (聖菲利柏氏症候群) III-B型 | 4669 | P54802 |
| N-乙醯基-胺基半乳糖苷酶 | I型辛德勒氏病(Schindler's disease) (嬰兒嚴重型) | 4668 | P17050 |
| N-乙醯基-胺基半乳糖苷酶 | II型辛德勒氏病(神琦病(Kanzaki disease)，成年發作型) | 4668 | P17050 |
| N-乙醯基-胺基半乳糖苷酶 | III型辛德勒氏病(中間型) | 4668 | P17050 |
| N-乙醯基-葡萄糖胺-6-硫酸酯硫酸酯酶 | 黏多糖病MPS III (聖菲利柏氏症候群) III-D型 | 2799 | P15586 |
| N-乙醯基葡糖胺基-1-磷酸轉移酶 | 黏脂貯積症ML III (假賀勒氏多種營養不良(pseudo-Hurler's polydystrophy)) | 79158 | Q3T906 |
| N-乙醯基葡糖胺基-1-磷酸轉移酶催化次單元 | 黏脂貯積症ML II (I細胞病) | 79158 | Q3T906 |
| N-乙醯基葡糖胺基-1-磷酸轉移酶，受質識別次單元 | 黏脂貯積症ML III (假賀勒氏多種營養不良) III-C型 | 84572 | Q9UJJ9 |
| N-天冬胺醯胺基葡萄糖苷酶 | 天冬胺醯葡萄糖胺尿症 | 175 | P20933 |
| 神經胺糖酸苷酶1 (唾液酸酶) | 唾液腺病 | 4758 | Q99519 |
| 軟脂醯基-蛋白質硫酯酶-1 | 成人型蠟樣脂褐質沈積症 (CLN4、庫夫斯病 (Kufs'disease)) | 5538 | P50897 |
| 軟脂醯基-蛋白質硫酯酶-1 | 嬰兒型蠟樣脂褐質沈積症 (CLN1、桑塔沃里-哈爾蒂亞病(Santavuori-Haltia disease)) | 5538 | P50897 |
| 苯丙胺酸羥化酶 | 苯酮尿症 | 5053 | P00439 |
| 磷酸甘露糖酶-2 | 先天性N-糖基化障礙CDG Ia (僅神經及神經-多內臟型) | 5373 | O15305 |

| | | | |
|---|---|---|---|
| 膽色素原去胺酶 | 急性間歇卟啉症 | 3145 | P08397 |
| 嘌呤核苷磷酸化酶 | 嘌呤核苷磷酸化酶缺乏症 | 4860 | P00491 |
| 嘧啶5'核苷酸酶 | 溶血性貧血及/或嘧啶5'核苷酸酶缺乏症 | 51251 | Q9H0P0 |
| 神經磷脂酶 | 尼曼-匹克病(Niemann-Pick disease) A型 | 6609 | P17405 |
| 神經磷脂酶 | 尼曼-匹克病B型 | 6609 | P17405 |
| 固醇27-羥化酶 | 腦腱黃瘤病(膽甾醇脂沈積症) | 1593 | Q02318 |
| 胸苷磷酸化酶 | 粒線體神經胃腸道腦肌病(MNGIE) | 1890 | P19971 |
| 三己糖神經醯胺α-半乳糖苷酶 | 法布立氏病(Fabry's disease) | 2717 | P06280 |
| 酪胺酸酶，例如OCA1 | 白化病，例如眼白化病 | 7299 | P14679 |
| UDP-GlcNAc:多萜醇基-P NAcGlc磷酸轉移酶 | 先天性N-醣基化障礙CDG Ij | 1798 | Q9H3H5 |
| UDP-N-乙醯基葡糖胺-2-表異構酶/N-乙醯甘露糖胺激酶，唾液酸轉運蛋白(sialin) | 法國型涎尿(Sialuria French type) | 10020 | Q9Y223 |
| 尿酸酶 | 萊尼症候群(Lesch-Nyhan syndrome)、痛風 | 391051 | No protein |
| 二磷酸尿苷葡萄糖醛酸轉移酶(例如UGT1A1) | 克果-納傑氏症候群(Crigler-Najjar syndrome) | 54658 | P22309 |
| α-1,2-甘露糖基轉移酶 | 先天性N-醣基化障礙CDG Il (608776) | 79796 | Q9H6U8 |
| α-1,2-甘露糖基轉移酶 | 先天性N-醣基化障礙，I型(前高基體醣基化缺陷) | 79796 | Q9H6U8 |
| α-1,3-甘露糖基轉移酶 | 先天性N-醣基化障礙CDG Ii | 440138 | Q2TAA5 |
| α-D-甘露糖苷酶 | α-甘露糖苷貯積病I型(重度)或II型(輕度) | 10195 | Q92685 |
| α-L-岩藻糖苷酶 | 岩藻糖苷貯積症 | 4123 | Q9NTJ4 |
| α-l-艾杜糖苷 | 黏多糖病MPS I H/S (賀勒-施艾氏侯症群(Hurler-Scheie syndrome)) | 2517 | P04066 |
| α-l-艾杜糖苷 | 黏多糖病MPS I-H (賀勒氏症候群) | 3425 | P35475 |
| α-l-艾杜糖苷 | 黏多糖病MPS I-S (施艾氏症候群(Scheie's syndrome)) | 3425 | P35475 |
| β-1,4-半乳糖苷基轉移酶 | 先天性N-醣基化障礙CDG IId | 3425 | P35475 |

| | | | |
|---|---|---|---|
| β-1,4-甘露糖基轉移酶 | 先天性N-醣基化障礙CDG Ik | 2683 | P15291 |
| β-D-甘露糖苷酶 | β-甘露糖苷貯積病 | 56052 | Q9BT22 |
| β-半乳糖苷酶 | 黏多糖病MPS IV (莫奎歐氏症候群) IV-B型 | 4126 | O00462 |
| β-葡萄糖醛酸苷酶 | 黏多糖病MPS VII (斯利症候群(Sly's syndrome)) | 2720 | P16278 |
| β-己糖苷酶A | 泰-薩克斯病(Tay-Sachs disease) | 2990 | P08236 |
| β-己糖苷酶B | 桑德霍夫氏病(Sandhoff's disease) | 3073 | P06865 |

【0239】 在一些實施例中，本文所述之效應子包含表54之酶，或其功能性變異體，例如同源物(例如直系同源物或旁系同源物)或其片段。在一些實施例中，本文所述之效應子包含與參考其UniProt ID在表54中所列之胺基酸序列具有至少80%、85%、90%、95%、967%、98%、99%序列一致性的蛋白質。在一些實施例中，功能性變異體催化與對應野生型蛋白質相同的反應，例如以比野生型蛋白質低不少於10%、20%、30%、40%或50%之速率。在一些實施例中，編碼表54之酶的指環載體或其功能性變異體用於治療表54之疾病或病症。在一些實施例中，指環載體用於將葡萄糖醛酸苷轉移酶或其功能性變異體遞送至目標細胞，例如肝細胞。在一些實施例中，指環載體用於將OCA1或其功能性變異體遞送至目標細胞，例如視網膜細胞。

## 表55.例示性非酶效應子及對應適應症

| 效應子 | 適應症 | Entrez基因ID | UniProt ID |
|---|---|---|---|
| 存活運動神經元蛋白質(SMN) | 脊髓性肌萎縮 | 6606 | Q16637 |
| 肌縮蛋白或微肌縮蛋白 | 肌肉萎縮症(例如杜氏肌肉萎縮(Duchenne muscular dystrophy)或貝氏肌肉萎縮(Becker muscular dystrophy)) | 1756 | P11532 |
| 補體蛋白，例如補體因子C1 | 補體因子I缺乏症 | 3426 | P05156 |
| 補體因子H | 非典型性溶血性尿毒症症候群 | 3075 | P08603 |
| 胱胺酸素(溶酶體胱胺酸轉運體) | 胱胺酸症 | 1497 | O60931 |
| 附睾分泌蛋白1(HE1；NPC2蛋白) | 尼曼-匹克病C2型 | 10577 | P61916 |
| GDP-岩藻糖轉運體-1 | 先天性N-糖基化障礙CDG IIc (拉瑪-哈薩隆症候群(Rambam-Hasharon syndrome)) | 55343 | Q96A29 |
| GM2活化蛋白質 | GM2活化蛋白缺乏症(泰-薩克斯病AB變異體，GM2A) | 2760 | Q17900 |
| 溶酶體跨膜CLN3蛋白質 | 幼年型蠟樣脂褐質沈積症(CLN3、巴氏病(Batten disease)、沃格特-施皮耳麥耶病(Vogt-Spielmeyer disease)) | 1207 | Q13286 |
| 溶酶體跨膜CLN5蛋白 | 幼兒型蠟樣脂褐質沈積症變異體，芬蘭型(CLN5) | 1203 | O75503 |
| Na磷酸協同轉運蛋白，唾液酸轉運蛋白 | 嬰兒唾液酸貯積病 | 26503 | Q9NRA2 |
| Na磷酸協同轉運蛋白，唾液酸轉運蛋白 | 芬蘭型涎尿(薩拉病(Salla disease)) | 26503 | Q9NRA2 |
| NPC1蛋白質 | 尼曼-匹克病C1型/D型 | 4864 | O15118 |
| 寡聚高基體複合物-7 | 先天性N-醣基化障礙CDG IIe | 91949 | P83436 |
| 鞘脂激活蛋白原 | 鞘脂激活蛋白原缺乏 | 5660 | P07602 |

| | | | |
|---|---|---|---|
| | 症 | | |
| 保護蛋白/組織蛋白酶A (PPCA) | 半乳糖唾液酸貯積症(戈德堡症候群(Goldberg's syndrome)、神經胺糖酸苷酶與β-半乳糖苷酶組合缺乏症) | 5476 | P10619 |
| 與甘露糖-P-多萜醇利用有關之蛋白質 | 先天性N-醣基化障礙CDG If | 9526 | O75352 |
| 鞘脂激活蛋白B | 鞘脂激活蛋白B缺乏症(硫苷脂活化子缺乏症) | 5660 | P07602 |
| 鞘脂激活蛋白C | 鞘脂激活蛋白C缺乏症(高歇氏活化子缺乏症(Gaucher's activator deficiency)) | 5660 | P07602 |
| 硫酸酯酶修飾因子-1 | 黏硫脂病(多硫酸酯酶缺乏症) | 285362 | Q8NBK3 |
| 跨膜CLN6蛋白質 | 幼兒型蠟樣脂褐質沈積症變異體(CLN6) | 54982 | Q9NWW5 |
| 跨膜CLN8蛋白質 | 蠟樣脂褐質沈積症進展性癲癇症伴智力障礙 | 2055 | Q9UBY8 |
| vWF | 馮威里氏病(von Willebrand disease) | 7450 | P04275 |
| 因子I (血纖維蛋白原) | 無纖維蛋白原血症 | 2243, 2244, 2266 | P02671, P02675, P02679 |
| 紅血球生成素(hEPO) | | | |

【0240】 在一些實施例中，本文所述之效應子包含紅血球生成素(EPO)，例如人類紅血球生成素(hEPO)或其功能性變異體。在一些實施例中，編碼紅血球生成素之指環載體或其功能性變異體用於刺激紅血球生成。在一些實施例中，編碼紅血球生成素或其功能性變異體之指環載體用於治療疾病或病症，例如貧血。在一些實施例中，指環載體用於將EPO或其功能性變異體遞送至目標細胞，例如紅血球。

【0241】 在一些實施例中，本文所述之效應子包含表55之多肽或其功能性變異體，例如同源物(例如直系同源物或旁系同源物)或其片段。在

一些實施例中，本文所述之效應子包含與參考其UniProt ID在表55中所列之胺基酸序列具有至少80%、85%、90%、95%、967%、98%、99%序列一致性的蛋白質。在一些實施例中，編碼表55之多肽或其功能性變異體的指環載體用於治療表55之疾病或病症。在一些實施例中，指環載體用於將SMN或其功能性變異體遞送至目標細胞，例如脊髓及/或運動神經元之細胞。在一些實施例中，指環載體用於將微肌縮蛋白遞送至目標細胞，例如肌細胞。

【0242】 例示性微肌縮蛋白描述於Duan, 「Systemic AAV Micro-dystrophin Gene Therapy for Duchenne Muscular Dystrophy.」 Mol Ther. 2018年10月3日;26(10):2337-2356. doi: 10.1016/j.ymthe.2018.07.011. 電子版2018年7月17日中。

【0243】 在一些實施例中，本文所述之效應子包含凝血因子，例如本文之表54或表55中所列之凝血因子。在一些實施例中，本文所述之效應子包含當突變時引起溶酶體貯積病之蛋白質，例如本文之表54或表55中所列之蛋白質。在一些實施例中，本文所述之效應子包含轉運蛋白，例如本文之表55中所列之轉運蛋白。

【0244】 在一些實施例中，野生型蛋白質之功能性變異體包含具有野生型蛋白質之一或多種活性的蛋白質，例如功能性變異體催化與對應野生型蛋白質相同的反應，例如以比野生型蛋白質低不少於10%、20%、30%、40%或50%之速率。在一些實施例中，功能性變異體以例如比在相同條件下對應野生型蛋白質對相同結合搭配物之Kd高不超過10%、20%、30%、40%或50%之Kd結合至由野生型蛋白質結合之相同結合搭配物。在一些實施例中，功能性變異體之多肽序列與野生型多肽至少70%、75%、

80%、85%、90%、95%、96%、97%、98%或99%一致。在一些實施例中，功能性變異體包含對應野生型蛋白質之同源物(例如，直系同源物或旁系同源物)。在一些實施例中，功能性變異體為融合蛋白。在一些實施例中，融合包含與對應野生型蛋白質具有至少70%、75%、80%、85%、90%、95%、96%、97%、98%或99%之一致性的第一區域及第二異源區域。在一些實施例中，功能性變異體包含對應野生型蛋白質之片段或由其組成。

【0245】

**再生、修復及纖維化因子**

本文所述之治療性多肽包括例如如表56中所揭示之生長因子，或其功能性變異體，例如與藉由參考UniProt ID而揭示於表56中之蛋白質序列具有至少80%、85%、90%、95%、967%、98%、99%一致性的蛋白質。亦包括針對此類生長因子之抗體或其片段，或促進再生及修復之miRNA。

**表56.例示性再生、修復及纖維化因子**

| 目標 | 基因寄存編號 | 蛋白質寄存編號 |
|---|---|---|
| VEGF-A | NG_008732 | NP_001165094 |
| NRG-1 | NG_012005 | NP_001153471 |
| FGF2 | NG_029067 | NP_001348594 |
| FGF1 | 基因ID：2246 | NP_001341882 |
| miR-199-3p | MIMAT0000232 | |
| miR-590-3p | MIMAT0004801 | |
| mi-17-92 | MI0000071 | https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2732113/figure/F1/ |
| miR-222 | MI0000299 | |
| miR-302-367 | MIR302A及MIR367 | https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4400607/ |

【0246】

**轉型因子**

本文所述之治療性多肽亦包括轉型因子，例如將纖維母細胞轉型成分化細胞之蛋白質因子，例如如表57中所揭示之因子，或其功能性變異體，例如與藉由參考UniProt ID而揭示於表57中之蛋白質序列具有至少80%、85%、90%、95%、967%、98%、99%一致性的蛋白質。

**表57.例示性轉型因子**

| 目標 | 適應症 | 基因寄存編號 | 蛋白質寄存編號 |
|---|---|---|---|
| MESP1 | 藉由纖維母細胞轉型之器官修復 | 基因ID：55897 | EAX02066 |
| ETS2 | 藉由纖維母細胞轉型之器官修復 | 基因ID：2114 | NP_005230 |
| HAND2 | 藉由纖維母細胞轉型之器官修復 | 基因ID：9464 | NP_068808 |
| MYOCARDIN | 藉由纖維母細胞轉型之器官修復 | 基因ID：93649 | NP_001139784 |
| ESRRA | 藉由纖維母細胞轉型之器官修復 | 基因ID：2101 | AAH92470 |
| miR-1 | 藉由纖維母細胞轉型之器官修復 | MI0000651 | n/a |
| miR-133 | 藉由纖維母細胞轉型之器官修復 | MI0000450 | n/a |
| TGFb | 藉由纖維母細胞轉型之器官修復 | 基因ID：7040 | NP_000651.3 |
| WNT | 藉由纖維母細胞轉型之器官修復 | 基因ID：7471 | NP_005421 |
| JAK | 藉由纖維母細胞轉型之器官修復 | 基因ID：3716 | NP_001308784 |
| NOTCH | 藉由纖維母細胞轉型之器官修復 | 基因ID：4851 | XP_011517019 |

【0247】

**刺激細胞再生之蛋白質**

本文所述之治療性多肽亦包括刺激細胞再生之蛋白質，例如如表58中所揭示之蛋白質，或其功能性變異體，例如與藉由參考UniProt ID而揭示於表58中之蛋白質序列具有至少80%、85%、90%、95%、967%、98%、99%一致性的蛋白質。

**表58.刺激細胞再生之例示性蛋白質**

| 目標 | 基因寄存編號 | 蛋白質寄存編號 |
|---|---|---|
| MST1 | NG_016454 | NP_066278 |
| STK30 | 基因ID：26448 | NP_036103 |
| MST2 | 基因ID：6788 | NP_006272 |
| SAV1 | 基因ID：60485 | NP_068590 |
| LATS1 | 基因ID：9113 | NP_004681 |

| | | |
|---|---|---|
| LATS2 | 基因ID：26524 | NP_055387 |
| YAP1 | NG_029530 | NP_001123617 |
| CDKN2b | NG_023297 | NP_004927 |
| CDKN2a | NG_007485 | NP_478102 |

【0248】

**STING調節物效應子**

在一些實施例中，本文所述之分泌效應子調節STING/cGAS信號傳導。在一些實施例中，STING調節物為多肽，例如病毒多肽或其功能性變異體。舉例而言，效應子可包含描述於Maringer等人.「Message in a bottle: lessons learned from antagonism of STING signalling during RNA virus infection」Cytokine & Growth Factor Reviews 第25卷, 第6期, 2014年12月, 第669-679頁中之STING調節物(例如抑制劑)，該等內容以全文引用的方式併入本文中。額外STING調節物(例如活化劑)描述於例如Wang等人.「STING activator c-di-GMP enhances the anti-tumor effects of peptide vaccines in melanoma-bearing mice.」Cancer Immunol Immunother. 2015年8月;64(8):1057-66. doi: 10.1007/s00262-015-1713-5. Epub 2015年5月19日; Bose「cGAS/STING Pathway in Cancer: Jekyll and Hyde Story of Cancer Immune Response」Int J Mol Sci. 2017年11月; 18(11): 2456；及Fu等人.「STING agonist formulated cancer vaccines can cure established tumors resistant to PD-1 blockade」Sci Transl Med. 2015年4月15日; 7(283): 283ra52中，其各自以全文引用的方式併入本文中。

【0249】肽之一些實例包括但不限於螢光標籤或標記物、抗原、治療性肽、來自天然生物活性肽之合成或模擬肽、促效性或拮抗性肽、抗微生物肽、靶向或細胞毒性肽、降解或自毀肽及降解或自毀肽。本文所述之

適用於本發明之肽亦包括抗原結合肽，例如抗原結合抗體或抗體樣片段，諸如單鏈抗體、奈米抗體(參見例如Steeland等人. 2016. Nanobodies as therapeutics: big opportunities for small antibodies. Drug Discov Today: 21(7):1076-113)。此類抗原結合肽可結合細胞溶質抗原、細胞核抗原或細胞器內抗原。

【0250】 在一些實施例中，遺傳元件包含編碼小肽、肽模擬物(例如，類肽)、胺基酸及胺基酸類似物之序列。此類治療劑通常具有每莫耳低於約5,000公克之分子量、每莫耳低於約2,000公克之分子量、每莫耳低於約1,000公克之分子量、每莫耳低於約500公克之分子量及此類化合物之鹽、酯及其他醫藥學上可接受之形式。此類治療劑可包括但不限於神經傳遞素、激素、藥物、毒素、病毒或微生物粒子、合成分子及其促效劑或拮抗劑。

【0251】 在一些實施例中，本文所述之組合物或指環載體包括連接至能夠靶向特定位置、組織或細胞之配位體的多肽。

【0252】

*基因編輯組分*

指環載體之遺傳元件可包括編碼基因編輯系統之組分的一或多個基因。例示性基因編輯系統包括成簇規律間隔短回文重複序列(clustered regulatory interspaced short palindromic repeat，CRISPR)系統、鋅指核酸酶(zinc finger nucleases，ZFN)及基於類轉錄活化因子效應子之核酸酶(Transcription Activator-Like Effector-based Nucleases，TALEN)。基於ZFN、TALEN及CRISPR之方法描述於例如Gaj等人. Trends Biotechnol. 31.7(2013):397-405中；基因編輯之CRISPR方法描述於例如Guan等人,

Application of CRISPR-Cas system in gene therapy: Pre-clinical progress in animal model. DNA Repair 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. BioTechniques, 第57卷, 第3號, 2014年9月, 第115-124頁中。

【0253】CRISPR系統為最初在細菌及古菌中發現的適應性防禦系統。CRISPR系統使用稱為CRISPR相關性或「Cas」核酸內切酶(例如，Cas9或Cpf1)之RNA引導之核酸酶以裂解外來DNA。在典型CRISPR/Cas系統中，核酸內切酶係藉由靶向單股或雙股DNA序列之序列特定性非編碼「嚮導RNA」指向目標核苷酸序列(例如，待進行序列編輯之基因體中的位點)。已鑑別出三類(I-III) CRISPR系統。II類CRISPR系統使用單一Cas核酸內切酶(而非多個Cas蛋白質)。一個II類CRISPR系統包括II型Cas核酸內切酶，諸如Cas9、CRISPR RNA (「crRNA」)及反式活化crRNA (「tracrRNA」)。crRNA含有「嚮導RNA」，通常約20個核苷酸RNA序列對應於目標DNA序列。crRNA亦含有結合至tracrRNA以形成由RNA酶III裂解之部分雙股結構之區域，產生crRNA/tracrRNA雜合體。crRNA/tracrRNA雜合體隨後引導Cas9核酸內切酶識別及裂解目標DNA序列。目標DNA序列通常必須鄰近於對指定Cas核酸內切酶具有特異性之「原間隔序列相鄰模體」(「PAM」)；然而，PAM序列出現在整個給定基因體中。

【0254】在一些實施例中，指環載體包括CRISPR核酸內切酶之基因。舉例而言，自多種原核物種鑑別之一些CRISPR核酸內切酶具有不同PAM序列要求；PAM序列之實例包括5'-NGG (化膿性鏈球菌)、5'-

NNAGAA (嗜熱鏈球菌CRISPR1)、5'-NGGNG (嗜熱鏈球菌CRISPR3)及5'-NNNGATT (腦膜炎雙球菌)。一些核酸內切酶，例如Cas9核酸內切酶係與G富集PAM位點，例如5'-NGG相關，且在PAM位點上游(5')之位置3核苷酸處進行目標DNA之平端裂解。另一II類CRISPR系統包括V型核酸內切酶Cpf1，其小於Cas9；實例包括AsCpf1 (來自胺基酸球菌屬物種)及LbCpf1 (來自毛螺菌科物種)。Cpf1核酸內切酶與T富集PAM位點，例如5'-TTN相關。Cpf1亦可識別5'-CTA PAM模體。Cpf1如下地裂解目標DNA：藉由引入具有4-或5-核苷酸5'突出物之偏移或交錯雙股斷裂，例如使用位於編碼股上之PAM位點下游(3') 18個核苷酸及互補股上之PAM位點下游23個核苷酸的5-核苷酸偏移或交錯切口裂解目標DNA；由此類偏移裂解產生之5-核苷酸突出物使得藉由同源重組之DNA插入與藉由插入平端裂解的DNA相比更精確的基因體編輯。參見例如Zetsche等人 (2015) Cell, 163:759 - 771。

【0255】 多種CRISPR相關(Cas)基因可包括於指環載體中。基因之特定實例為編碼來自包括Cas1、Cas2、Cas3、Cas4、Cas5、Cas6、Cas7、Cas8、Cas9、Cas10、Cpf1、C2C1或C2C3之II類系統之Cas蛋白的彼等基因。在一些實施例中，指環載體包括編碼Cas蛋白，例如Cas9蛋白之基因，可來自多種原核物種中之任一者。在一些實施例中，指環載體包括編碼Cas蛋白，例如Cas9蛋白之基因，經選擇以識別特定前間隔子相鄰模體(PAM)序列。在一些實施例中，指環載體包括編碼兩個或更多個不同Cas蛋白或兩個或更多個Cas蛋白之核酸，可引入至細胞、受精卵、胚胎或動物中，例如以允許識別及修飾包含相同、類似或不同PAM模體之位點。在一些實施例中，指環載體包括編碼具有去活化核酸酶之經修飾

Cas蛋白，例如核酸酶缺陷型Cas9的基因。

【0256】 儘管野生型Cas9蛋白質在由gRNA靶向之特定DNA序列處產生雙股斷裂(DSB)，但已知具有經修飾功能之多種CRISPR核酸內切酶，例如：Cas核酸內切酶之「切口酶」型式(例如Cas9)僅產生單股斷裂；無催化活性Cas核酸內切酶，例如Cas9 (「dCas9」)不切割目標DNA。編碼dCas9之基因可與編碼效應子域之基因融合以抑制(CRISPRi)或活化(CRISPRa)目標基因之表現。舉例而言，基因可使用轉錄緘默子(例如，KRAB域)或轉錄活化因子(例如，dCas9-VP64融合)編碼Cas9融合。可包括編碼融合至FokI核酸酶之無催化活性Cas9 (dCas9)之基因(「dCas9-FokI」)，以在與兩個gRNA同源之目標序列處產生DSB。參見例如揭示於Addgene repository (Addgene, 75 Sidney St., Suite 550A, Cambridge, MA 02139; addgene.org/crispr/)且公開可購自其中之許多CRISPR/Cas9質體。引入兩個單獨雙股斷裂(各自藉由單獨嚮導RNA引導)之「雙切口酶」Cas9由Ran等人 (2013) Cell, 154:1380 - 1389描述為實現更精確的基因體編輯。

【0257】 用於編輯真核生物之基因的CRISPR技術揭示於美國專利申請公開案2016/0138008A1及US2015/0344912A1中，及美國專利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及8,906,616中。Cpf1核酸內切酶及對應嚮導RNA及PAM位點揭示於美國專利申請公開案2016/0208243 A1中。

【0258】 在一些實施例中，指環載體包含編碼本文所述之多肽，例如經靶向核酸酶，例如Cas9，例如野生型Cas9、切口酶Cas9 (例如，

Cas9 D10A)、死亡Cas9 (dCas9)、eSpCas9、Cpf1、C2C1或C2C3及gRNA的基因。編碼核酸酶及gRNA之基因的選擇係藉由所靶向突變是否為核苷酸之缺失、取代或添加，例如將核苷酸缺失、取代或添加至所靶向序列來確定。編碼無催化活性核酸內切酶之基因，例如與(一或多個)效應子域(例如VP64)之全部或一部分(例如，生物活性部分)繫留之死亡Cas9(dCas9，例如D10A；H840A)產生可調節一或多個目標核酸序列之活性及/或表現的嵌合蛋白質。

【0259】 在一些實施例中，指環載體包括編碼dCas9與一或多個效應子域(例如，全長野生型效應子域或其片段或變異體，例如其生物活性部分)之全部或一部分之融合物的基因以產生適用於本文所述之方法的嵌合蛋白。因此，在一些實施例中，指環載體包括編碼dCas9甲基化酶融合物之基因。在其他一些實施例中，指環載體包括編碼與位點特異性gRNA之dCas9-酶融合物的基因以靶向內源性基因。

【0260】 在其他態樣中，指環載體包括編碼與dCas9融合之1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20或更多個效應子域(全部或生物活性部分)的基因。

【0261】

*調控序列*

在一些實施例中，遺傳元件包含可操作地連接於編碼效應子之序列的調控序列，例如啟動子或強化子。在例如其中遺傳元件為mRNA之一些實施例中，遺傳元件中可能不存在啟動子。在一些實施例中，遺傳元件構築體包含用於驅動RNA遺傳元件之產生的啟動子。

【0262】 在一些實施例中，啟動子包括與編碼表現產物之DNA序列

相鄰定位的DNA序列。啟動子可以操作方式連接於相鄰DNA序列。相比於不存在啟動子時表現產物之量，啟動子通常增加自DNA序列表現之產物的量。來自一個生物體之啟動子可用於增強來自來源於另一生物體之DNA序列的產物表現。舉例而言，脊椎動物啟動子可用於在脊椎動物中表現水母GFP。因此，一個啟動子元件可增強一或多種產物之表現。多個啟動子元件為一般熟習此項技術者熟知的。

【0263】 在一個實施例中，需要高含量組成性表現。此類啟動子之實例包括但不限於反轉錄病毒勞氏肉瘤病毒(Rous sarcoma virus，RSV)長末端重複序列(LTR)啟動子/強化子、細胞巨大病毒(CMV)即刻早期啟動子/強化子(參見例如Boshart等人, Cell, 41:521-530 (1985))、SV40啟動子、二氫葉酸還原酶啟動子、細胞質.β.-肌動蛋白啟動子及磷酸甘油激酶(PGK)啟動子。

【0264】 在另一實施例中，可能需要誘導性啟動子。誘導性啟動子為由外源提供之化合物調控之彼等啟動子，例如以順式或反式提供，包括但不限於鋅誘導性綿羊金屬硫蛋白(MT)啟動子；地塞米松(Dex)誘導性小鼠乳房腫瘤病毒(MMTV)啟動子；T7聚合酶啟動子系統(WO 98/10088)；四環素抑制性系統(Gossen等人, Proc. Natl. Acad. Sci. USA, 89:5547-5551 (1992))；四環素誘導性系統(Gossen等人, Science, 268:1766-1769 (1995)；亦參見Harvey等人, Curr. Opin. Chem. Biol., 2:512-518 (1998))；RU486誘導性系統(Wang等人, Nat. Biotech., 15:239-243 (1997)及Wang等人, Gene Ther., 4:432-441 (1997)]；及雷帕黴素誘導性系統(Magari等人, J. Clin. Invest., 100:2865-2872 (1997)；Rivera等人, Nat. Medicine. 2:1028-1032 (1996))。在此情形下可適用之其他類型之誘導性

啟動子為藉由特定生理狀態，例如溫度、急性期或僅在複製細胞中調節之啟動子。

【0265】 在一些實施例中，使用所關注之基因或核酸序列之天然啟動子。當期望基因或核酸序列之表現應模擬天然表現時，可使用天然啟動子。當基因或其他核酸序列之表現必須在時間上或發育上，或以組織特異性方式，或回應於特定轉錄刺激進行調節時，可使用天然啟動子。在另一實施例中，其他天然表現控制元件，諸如強化子元件、聚腺苷酸化位點或Kozak共有序列，亦可用於模擬天然表現。

【0266】 在一些實施例中，遺傳元件包含可操作地連接至組織特異性啟動子的基因。舉例而言，若期望骨胳肌肉中之表現，則可使用在肌肉中有活性之啟動子。此等啟動子包括來自編碼骨骼α-肌動蛋白、肌凝蛋白輕鏈2A、肌縮蛋白、肌肉肌酸激酶之基因的啟動子，以及具有高於天然存在之啟動子之活性的合成肌肉啟動子。參見Li等人, Nat. Biotech., 17:241-245 (1999)。組織特異性啟動子之實例為吾人所知：肝白蛋白，Miyatake等人 J. Virol., 71:5124-32 (1997)；B型肝炎病毒核心啟動子，Sandig等人, Gene Ther. 3:1002-9 (1996)；α-胎蛋白(AFP)，Arbuthnot等人, Hum. Gene Ther., 7:1503-14 (1996)]；骨(骨鈣化素，Stein等人, Mol. Biol. Rep., 24:185-96 (1997)；骨唾液蛋白，Chen等人, J. Bone Miner. Res. 11:654-64 (1996))；淋巴球(CD2，Hansal等人, J. Immunol., 161:1063-8 (1998)；免疫球蛋白重鏈；T細胞受體a鏈)；神經元(神經元特異性烯醇化酶(NSE)啟動子，Andersen等人 Cell. Mol. Neurobiol., 13:503-15 (1993)；神經絲輕鏈基因，Piccioli等人, Proc. Natl. Acad. Sci. USA, 88:5611-5 (1991)；神經元特異性vgf基因，Piccioli等人, Neuron,

15:373-84 (1995)]；以及其他。

【0267】 遺傳元件構築體可包括強化子，例如與編碼基因之DNA序列相鄰定位的DNA序列。強化子元件通常位於啟動子元件上游或可位於編碼DNA序列(例如，轉錄或轉譯成一或多種產物之DNA序列)下游或其內。因此，強化子元件可位於編碼產物之DNA序列上游或下游100個鹼基對、200個鹼基對或300個或更多個鹼基對處。強化子元件可將自DNA序列表現之重組產物之量增加至高於由啟動子元件提供之增加的表現。多個強化子元件可易於供一般熟習此項技術者使用。

【0268】 在一些實施例中，遺傳元件包含一或多個側接編碼本文所述之表現產物之序列的反向末端重複序列(ITR)。在一些實施例中，遺傳元件包含一或多個側接編碼本文所述之表現產物之序列的長末端重複序列(LTR)。可使用之啟動子序列之實例包括但不限於猿猴病毒40 (SV40)早期啟動子、小鼠乳房腫瘤病毒(MMTV)、人類免疫缺乏病毒(HIV)長末端重複序列(LTR)啟動子、MoMuLV啟動子、禽類白血病病毒啟動子、埃-巴二氏病毒(Epstein-Barr virus)即刻早期啟動子及勞氏肉瘤病毒啟動子。

【0269】

*其他序列*

在一些實施例中，遺傳元件進一步包括編碼產物(例如核糖核酸酶、編碼蛋白質之治療性mRNA、外源性基因)之核酸。

【0270】 在一些實施例中，遺傳元件包括一或多個影響以下之序列：物種及/或組織及/或細胞向性(例如蛋白殼蛋白序列)、感染性(例如蛋白殼蛋白序列)、免疫抑制/活化(例如調控核酸)、病毒基因體結合及/或封裝、免疫逃避(非免疫原性及/或耐受性)、藥物動力學、內吞作用及/或細

胞附著、核進入、胞內調節及定位、胞外分泌調節、繁殖及宿主或宿主細胞中指環載體之核酸保護。

【0271】 在一些實施例中，遺傳元件可包含包括DNA、RNA或人工核酸之其他序列。其他序列可包括但不限於基因體DNA、cDNA或編碼tRNA、mRNA、rRNA、miRNA、gRNA、siRNA或其他RNAi分子之序列。在一個實施例中，遺傳元件包括編碼siRNA以靶向與調控核酸相同之基因表現產物之不同基因座的序列。在一個實施例中，遺傳元件包括編碼siRNA以靶向與調控核酸不同之基因表現產物的序列。

【0272】 在一些實施例中，遺傳元件進一步包含以下序列中之一或多者：編碼一或多個miRNA之序列、編碼一或多個複製蛋白之序列、編碼外源性基因之序列、編碼治療劑之序列、調控序列(例如啟動子、強化子)、編碼一或多個靶向內源性基因(siRNA、lncRNA、shRNA)之調控序列的序列以及編碼治療性mRNA或蛋白質之序列。

【0273】 其他序列之長度可為約2至約5000 nt、約10至約100 nt、約50至約150 nt、約100至約200 nt、約150至約250 nt、約200至約300 nt、約250至約350 nt、約300至約500 nt、約10至約1000 nt、約50至約1000 nt、約100至約1000 nt、約1000至約2000 nt、約2000至約3000 nt、約3000至約4000 nt、約4000至約5000 nt，或其間任何範圍。

【0274】

*經編碼之基因*

舉例而言，遺傳元件可包括與傳訊生化途徑相關之基因，例如傳訊生化途徑相關基因或聚核苷酸。實例包括疾病相關基因或聚核苷酸。「疾病相關」基因或聚核苷酸係指相比於非疾病對照之組織或細胞，在衍生自

受疾病影響之組織的細胞中以異常含量或以異常形式產生轉錄或轉譯產物的任何基因或聚核苷酸。其可為以異常高含量表現之基因；其可為以異常低含量表現之基因，其中改變之表現與疾病之出現及/或進展相關。疾病相關基因亦指具有直接負責或與負責疾病病因之基因處於連鎖不平衡之突變或遺傳變異的基因。

【0275】 疾病相關基因及聚核苷酸之實例係購自McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, Md.)及National Center for Biotechnology Information, National Library of Medicine (Bethesda, Md.)。疾病相關基因及聚核苷酸之實例列於美國專利第8,697,359號之表A及B中，該專利以全文引用之方式併入本文中。疾病特定資訊可獲自McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, Md.)及National Center for Biotechnology Information, National Library of Medicine (Bethesda, Md.)。信號傳導生化途徑相關基因及聚核苷酸之實例列於美國專利第8,697,359號之表A-C中，該專利以全文引用之方式併入本文中。

【0276】 此外，遺傳元件可編碼靶向部分，如本文其他處所述。此可例如藉由插入編碼糖、醣脂或蛋白質，諸如抗體之聚核苷酸來實現。熟習此項技術者已知用於產生靶向部分之額外方法。

【0277】

*病毒序列*

在一些實施例中，遺傳元件包含至少一個病毒序列。在一些實施例中，序列與一或多個來自單股DNA病毒，例如指環病毒(Anellovirus)、雙DNA病毒(Bidnavirus)、環狀病毒(Circovirus)、雙生病毒

(Geminivirus)、基因體病毒(Genomovirus)、絲狀病毒(Inovirus)、微小病毒(Microvirus)、矮化病毒(Nanovirus)、小病毒(Parvovirus)及螺旋病毒(Spiravirus)的序列具有同源性或一致性。在一些實施例中，序列與一或多個來自雙股DNA病毒，例如腺病毒(Adenovirus)、壺腹病毒(Ampullavirus)、囊泡病毒(Ascovirus)、非洲豬瘟病毒(Asfarvirus)、桿狀病毒(Baculovirus)、微小紡錘形噬菌體屬(Fusellovirus)、球狀病毒(Globulovirus)、滴狀病毒(Guttavirus)、肥大唾腺炎病毒(Hytrosavirus)、疱疹病毒(Herpesvirus)、虹彩病毒(Iridovirus)、脂毛病毒(Lipothrixvirus)、線極病毒(Nimavirus)及痘病毒(Poxvirus)的序列具有同源性或一致性。在一些實施例中，序列與一或多個來自RNA病毒，例如α病毒(Alphavirus)、真菌傳棒狀病毒(Furovirus)、肝炎病毒(Hepatitis virus)、大麥病毒(Hordeivirus)、菸草花葉病毒(Tobamovirus)、菸草脆裂病毒(Tobravirus)、三角病毒(Tricornavirus)、風疹病毒(Rubivirus)、雙RNA病毒(Birnavirus)、囊狀病毒(Cystovirus)、分病毒(Partitivirus)及里奧病毒(Reovirus)的序列具有同源性或一致性。

【0278】在一些實施例中，遺傳元件可以包含一或多個來自非致病性病毒，例如共生性病毒，例如共生病毒，例如天然病毒，例如指環病毒之序列。命名法之近期變化將能夠感染人類細胞之三種指環病毒分類為病毒之指環病毒科的α細環病毒(TT)、β細環病毒(TTM)及γ細環病毒(TTMD)屬。在一些實施例中，遺傳元件可包含與細環病毒(Torque Teno Virus，TT)，一種具有環狀、反義基因體之無包膜、單股DNA病毒具有同源性或一致性的序列。在一些實施例中，遺傳元件可包含與SEN病毒、哨兵病毒(Sentinel virus)、TTV樣微型病毒及TT病毒具有同源性或一致性

之序列。已描述不同類型之TT病毒，包括TT病毒基因型6、TT病毒組、TTV樣病毒DXL1及TTV樣病毒DXL2。在一些實施例中，遺傳元件可包含與以下具有同源性或一致性之序列：較小病毒細環樣微型病毒(TTM)，或基因體尺寸在TTV與TTMV之間的第三病毒，稱為細環樣中型病毒(TTMD)。在一些實施例中，遺傳元件可包含一或多個來自非致病性病毒之序列或序列片段，其與本文所述之核苷酸序列中之任一者具有至少約60%、70%、80%、85%、90%、95%、96%、97%、98%及99%核苷酸序列一致性。

【0279】在一些實施例中，遺傳元件可包含一或多個來自實質上非致病性病毒之序列或序列片段，其與本文，例如表41所述之核苷酸序列中之任一者具有至少約60%、70%、80%、85%、90%、95%、96%、97%、98%及99%核苷酸序列一致性。

**表41：指環病毒及其序列之實例.**如參考2018年12月11日，可在www.ncbi.nlm.nih.gov/genbank/寄存編號及相關序列資訊獲得。

| 寄存編號 | 描述 |
|---|---|
| AB017613.1 | 細環病毒16 DNA，完全基因體，分離株：TUS01 |
| AB026345.1 | ORF1及ORF2之TT病毒基因，完全cd，分離株：TRM1 |
| AB026346.1 | ORF1及ORF2之TT病毒基因，完全cd，分離株：TK16 |
| AB026347.1 | ORF1及ORF2之TT病毒基因，完全cd，分離株：TP1-3 |
| AB028669.1 | ORF1及ORF2之TT病毒基因，完全基因體，分離株：TJN02 |
| AB030487.1 | pORF2a、pORF2b、pORF1之TT病毒基因，完全cd，純系：JaCHCTC19 |
| AB030488.1 | pORF2a、pORF2b、pORF1之TT病毒基因，完全cd，純系：JaBD89 |
| AB030489.1 | pORF2a、pORF2b、pORF1，完全cd，純系：JaBD98 |
| AB038340.1 | ORF2s、ORF1、ORF3之TT病毒基因，完全cd |
| AB038622.1 | ORF2、ORF1、ORF3之TT病毒基因，完全cd，分離株：TTVyon-LC011 |
| AB038623.1 | ORF2、ORF1、ORF3之TT病毒基因，完全cd，分離株：TTVyon-KC186 |
| AB038624.1 | ORF2、ORF1、ORF3之TT病毒基因，完全cd，分離株：TTVyon-KC197 |
| AB041821.1 | VP1之TT病毒mRNA，完全cd |
| AB050448.1 | ORF1、ORF2、ORF3、ORF4之細環病毒基因，完全cd，分離株：TYM9 |
| AB060592.1 | ORF1、ORF2、ORF3、ORF4之細環病毒基因，純系：SAa-39 |
| AB060593.1 | ORF1、ORF2、ORF3、ORF4之細環病毒基因，完全cd，純系：SAa-38 |
| AB060595.1 | ORF1、ORF2、ORF3、ORF4之TT病毒基因，完全cd，純系：SAj-30 |

| | |
|---|---|
| AB060596.1 | ORF1、ORF2、ORF3、ORF4之TT病毒基因，完全cd，純系：SAf-09 |
| AB064596.1 | 細環病毒DNA，完全基因體，分離株：CT25F |
| AB064597.1 | 細環病毒DNA，完全基因體，分離株：CT30F |
| AB064599.1 | 細環病毒DNA，完全基因體，分離株：JT03F |
| AB064600.1 | 細環病毒DNA，完全基因體，分離株：JT05F |
| AB064601.1 | 細環病毒DNA，完全基因體，分離株：JT14F |
| AB064602.1 | 細環病毒DNA，完全基因體，分離株：JT19F |
| AB064603.1 | 細環病毒DNA，完全基因體，分離株：JT41F |
| AB064604.1 | 細環病毒DNA，完全基因體，分離株：CT39F |
| AB064606.1 | 細環病毒DNA，完全基因體，分離株：JT33F |
| AB290918.1 | 細環中型病毒1 DNA，完全基因體，分離株：MD1-073 |
| AF079173.1 | TT病毒株TTVCHN1，完全基因體 |
| AF116842.1 | TT病毒株BDH1，完全基因體 |
| AF122914.3 | TT病毒分離株JA20，完全基因體 |
| AF122917.1 | TT病毒分離株JA4，完全基因體 |
| AF122919.1 | TT病毒分離株JA10未知基因 |
| AF129887.1 | TT病毒TTVCHN2，完全基因體 |
| AF247137.1 | TT病毒分離株TUPB，完全基因體 |
| AF254410.1 | TT病毒ORF2蛋白及ORF1蛋白基因，完全cd |
| AF298585.1 | TT病毒拋光分離株P/1C1，完全基因體 |
| AF315076.1 | TTV樣病毒DXL1未知基因 |
| AF315077.1 | TTV樣病毒DXL2未知基因 |
| AF345521.1 | TT病毒分離株TCHN-G1 Orf2及Orf1基因，完全cd |
| AF345522.1 | TT病毒分離株TCHN-E Orf2及Orf1基因，完全cd |
| AF345525.1 | TT病毒分離株TCHN-D2 Orf2及Orf1基因，完全cd |
| AF345527.1 | TT病毒分離株TCHN-C2 Orf2及Orf1基因，完全cd |
| AF345528.1 | TT病毒分離株TCHN-F Orf2及Orf1基因，完全cd |
| AF345529.1 | TT病毒分離株TCHN-G2 Orf2及Orf1基因，完全cd |
| AF371370.1 | TT病毒ORF1、ORF3及ORF2基因，完全cd |
| AJ620212.1 | 細環病毒，分離株tth6，完全基因體 |
| AJ620213.1 | 細環病毒，分離株tth10，完全基因體 |
| AJ620214.1 | 細環病毒，分離株tth11g2，完全基因體 |
| AJ620215.1 | 細環病毒，分離株tth18，完全基因體 |
| AJ620216.1 | 細環病毒，分離株tth20，完全基因體 |
| AJ620217.1 | 細環病毒，分離株tth21，完全基因體 |
| AJ620218.1 | 細環病毒，分離株tth3，完全基因體 |
| AJ620219.1 | 細環病毒，分離株tth9，完全基因體 |
| AJ620220.1 | 細環病毒，分離株tth16，完全基因體 |
| AJ620221.1 | 細環病毒，分離株tth17，完全基因體 |
| AJ620222.1 | 細環病毒，分離株tth25，完全基因體 |
| AJ620223.1 | 細環病毒，分離株tth26，完全基因體 |
| AJ620224.1 | 細環病毒，分離株tth27，完全基因體 |
| AJ620225.1 | 細環病毒，分離株tth31，完全基因體 |
| AJ620226.1 | 細環病毒，分離株tth4，完全基因體 |

| | |
|---|---|
| AJ620227.1 | 細環病毒，分離株tth5，完全基因體 |
| AJ620228.1 | 細環病毒，分離株tth14，完全基因體 |
| AJ620229.1 | 細環病毒，分離株tth29，完全基因體 |
| AJ620230.1 | 細環病毒，分離株tth7，完全基因體 |
| AJ620231.1 | 細環病毒，分離株tth8，完全基因體 |
| AJ620232.1 | 細環病毒，分離株tth13，完全基因體 |
| AJ620233.1 | 細環病毒，分離株tth19，完全基因體 |
| AJ620234.1 | 細環病毒，分離株tth22g4，完全基因體 |
| AJ620235.1 | 細環病毒，分離株tth23，完全基因體 |
| AM711976.1 | TT病毒sle1957完全基因體 |
| AM712003.1 | TT病毒sle1931完全基因體 |
| AM712004.1 | TT病毒sle1932完全基因體 |
| AM712030.1 | TT病毒sle2057完全基因體 |
| AM712031.1 | TT病毒sle2058完全基因體 |
| AM712032.1 | TT病毒sle2072完全基因體 |
| AM712033.1 | TT病毒sle2061完全基因體 |
| AM712034.1 | TT病毒sle2065完全基因體 |
| AY026465.1 | TT病毒分離株L01 ORF2及ORF1基因，完全cd |
| AY026466.1 | TT病毒分離株L02 ORF2及ORF1基因，完全cd |
| DQ003341.1 | 細環病毒純系P2-9-02 ORF2 (ORF2)、ORF1A (ORF1A)及ORF1B (ORF1B)基因，完全cd |
| DQ003342.1 | 細環病毒純系P2-9-07 ORF2 (ORF2)、ORF1A (ORF1A)及ORF1B (ORF1B)基因，完全cd |
| DQ003343.1 | 細環病毒純系P2-9-08 ORF2 (ORF2)、ORF1A (ORF1A)及ORF1B (ORF1B)基因，完全cd |
| DQ003344.1 | 細環病毒純系P2-9-16 ORF2 (ORF2)、ORF1A (ORF1A)及ORF1B (ORF1B)基因，完全cd |
| DQ186994.1 | 細環病毒純系P601 ORF2 (ORF2)及ORF1 (ORF1)基因，完全cd |
| DQ186995.1 | 細環病毒純系P605 ORF2 (ORF2)及ORF1 (ORF1)基因，完全cd |
| DQ186996.1 | 細環病毒純系BM1A-02 ORF2 (ORF2)及ORF1 (ORF1)基因，完全cd |
| DQ186997.1 | 細環病毒純系BM1A-09 ORF2 (ORF2)及ORF1 (ORF1)基因，完全cd |
| DQ186998.1 | 細環病毒純系BM1A-13 ORF2 (ORF2)及ORF1 (ORF1)基因，完全cd |
| DQ186999.1 | 細環病毒純系BM1B-05 ORF2 (ORF2)及ORF1 (ORF1)基因，完全cd |
| DQ187000.1 | 細環病毒純系BM1B-07 ORF2 (ORF2)及ORF1 (ORF1)基因，完全cd |
| DQ187001.1 | 細環病毒純系BM1B-11 ORF2 (ORF2)及ORF1 (ORF1)基因，完全cd |
| DQ187002.1 | 細環病毒純系BM1B-14 ORF2 (ORF2)及ORF1 (ORF1)基因，完全cd |
| DQ187003.1 | 細環病毒純系BM1B-08 ORF2 (ORF2)基因，完全cd；及非功能性ORF1 (ORF1)基因，完全序列 |
| DQ187004.1 | 細環病毒純系BM1C-16 ORF2 (ORF2)及ORF1 (ORF1)基因，完全cd |
| DQ187005.1 | 細環病毒純系BM1C-10 ORF2 (ORF2)及ORF1 (ORF1)基因，完全cd |
| DQ187007.1 | 細環病毒純系BM2C-25 ORF2 (ORF2)基因，完全cd；及非功能性ORF1 (ORF1)基因，完全序列 |
| DQ361268.1 | 細環病毒分離株ViPi04 ORF1基因，完全cd |
| EF538879.1 | 細環病毒分離株CSC5 ORF2及ORF1基因，完全cd |

| | |
|---|---|
| EU305675.1 | 細環病毒分離株LTT7 ORF1基因，完全cd |
| EU305676.1 | 細環病毒分離株LTT10 ORF1基因，完全cd |
| EU889253.1 | 細環病毒分離株ViPi08非功能性ORF1基因，完全序列 |
| FJ392105.1 | 細環病毒分離株TW53A25 ORF2基因，部分cd；及ORF1基因，完全cd |
| FJ392107.1 | 細環病毒分離株TW53A27 ORF2基因，部分cd；及ORF1基因，完全cd |
| FJ392108.1 | 細環病毒分離株TW53A29 ORF2基因，部分cd；及ORF1基因，完全cd |
| FJ392111.1 | 細環病毒分離株TW53A35 ORF2基因，部分cd；及ORF1基因，完全cd |
| FJ392112.1 | 細環病毒分離株TW53A39 ORF2基因，部分cd；及ORF1基因，完全cd |
| FJ392113.1 | 細環病毒分離株TW53A26 ORF2基因，完全cd；及非功能性ORF1基因，完全序列 |
| FJ392114.1 | 細環病毒分離株TW53A30 ORF2及ORF1基因，完全cd |
| FJ392115.1 | 細環病毒分離株TW53A31 ORF2及ORF1基因，完全cd |
| FJ392117.1 | 細環病毒分離株TW53A37 ORF1基因，完全cd |
| FJ426280.1 | 細環病毒株SIA109，完全基因體 |
| FR751500.1 | 細環病毒完全基因體，病毒株TTV-HD23a (rheu215) |
| GU797360.1 | 細環病毒純系8-17，完全基因體 |
| HC742700.1 | 來自專利WO2010044889之序列7 |
| HC742710.1 | 來自專利WO2010044889之序列17 |
| JX134044.1 | TTV樣微型病毒分離株TTMV_LY1，完全基因體 |
| JX134045.1 | TTV樣微型病毒分離株TTMV_LY2，完全基因體 |
| KU243129.1 | TTV樣微型病毒分離株TTMV-204，完全基因體 |
| KY856742.1 | TTV樣微型病毒分離株zhenjiang，完全基因體 |
| LC381845.1 | 細環病毒Human/Japan/KS025/2016 DNA，完全基因體 |
| MH648892.1 | 指環病毒物種分離株ctdc048，完全基因體 |
| MH648893.1 | 指環病毒物種分離株ctdh007，完全基因體 |
| MH648897.1 | 指環病毒物種分離株ctcb038，完全基因體 |
| MH648900.1 | 指環病毒物種分離株ctfc019，完全基因體 |
| MH648901.1 | 指環病毒物種分離株ctbb022，完全基因體 |
| MH648907.1 | 指環病毒物種分離株ctcf040，完全基因體 |
| MH648911.1 | 指環病毒物種分離株cthi018，完全基因體 |
| MH648912.1 | 指環病毒物種分離株ctea38，完全基因體 |
| MH648913.1 | 指環病毒物種分離株ctbg006，完全基因體 |
| MH648916.1 | 指環病毒物種分離株ctbg020，完全基因體 |
| MH648925.1 | 指環病毒物種分離株ctci019，完全基因體 |
| MH648932.1 | 指環病毒物種分離株ctid031，完全基因體 |
| MH648946.1 | 指環病毒物種分離株ctdb017，完全基因體 |
| MH648957.1 | 指環病毒物種分離株ctch017，完全基因體 |
| MH648958.1 | 指環病毒物種分離株ctbh011，完全基因體 |
| MH648959.1 | 指環病毒物種分離株ctbc020，完全基因體 |
| MH648962.1 | 指環病毒物種分離株ctif015，完全基因體 |
| MH648966.1 | 指環病毒物種分離株ctei055，完全基因體 |
| MH648969.1 | 指環病毒物種分離株ctjg000，完全基因體 |
| MH648976.1 | 指環病毒物種分離株ctcj064，完全基因體 |
| MH648977.1 | 指環病毒物種分離株ctbj022，完全基因體 |

| | |
|---|---|
| MH648982.1 | 指環病毒物種分離株ctbf014，完全基因體 |
| MH648983.1 | 指環病毒物種分離株ctbd027，完全基因體 |
| MH648985.1 | 指環病毒物種分離株ctch016，完全基因體 |
| MH648986.1 | 指環病毒物種分離株ctbd020，完全基因體 |
| MH648989.1 | 指環病毒物種分離株ctga035，完全基因體 |
| MH648990.1 | 指環病毒物種分離株cthf001，完全基因體 |
| MH648995.1 | 指環病毒物種分離株ctbd067，完全基因體 |
| MH648997.1 | 指環病毒物種分離株ctce026，完全基因體 |
| MH648999.1 | 指環病毒物種分離株ctfb058，完全基因體 |
| MH649002.1 | 指環病毒物種分離株ctjj046，完全基因體 |
| MH649006.1 | 指環病毒物種分離株ctcf030，完全基因體 |
| MH649008.1 | 指環病毒物種分離株ctbg025，完全基因體 |
| MH649011.1 | 指環病毒物種分離株ctbh052，完全基因體 |
| MH649014.1 | 指環病毒物種分離株ctba003，完全基因體 |
| MH649017.1 | 指環病毒物種分離株ctbb016，完全基因體 |
| MH649022.1 | 指環病毒物種分離株ctch023，完全基因體 |
| MH649023.1 | 指環病毒物種分離株ctbd051，完全基因體 |
| MH649028.1 | 指環病毒物種分離株ctbf9，完全基因體 |
| MH649038.1 | 指環病毒物種分離株ctbi030，完全基因體 |
| MH649039.1 | 指環病毒物種分離株ctca057，完全基因體 |
| MH649040.1 | 指環病毒物種分離株ctch033，完全基因體 |
| MH649042.1 | 指環病毒物種分離株ctjd005，完全基因體 |
| MH649045.1 | 指環病毒物種分離株ctdc021，完全基因體 |
| MH649051.1 | 指環病毒物種分離株ctdg044，完全基因體 |
| MH649056.1 | 指環病毒物種分離株ctcc062，完全基因體 |
| MH649061.1 | 指環病毒物種分離株ctid009，完全基因體 |
| MH649062.1 | 指環病毒物種分離株ctdc018，完全基因體 |
| MH649063.1 | 指環病毒物種分離株ctbf012，完全基因體 |
| MH649068.1 | 指環病毒物種分離株ctcc066，完全基因體 |
| MH649070.1 | 指環病毒物種分離株ctda011，完全基因體 |
| MH649077.1 | 指環病毒物種分離株ctbh034，完全基因體 |
| MH649083.1 | 指環病毒物種分離株ctdg028，完全基因體 |
| MH649084.1 | 指環病毒物種分離株ctii061，完全基因體 |
| MH649085.1 | 指環病毒物種分離株cteh021，完全基因體 |
| MH649092.1 | 指環病毒物種分離株ctbg012，完全基因體 |
| MH649101.1 | 指環病毒物種分離株ctif053，完全基因體 |
| MH649104.1 | 指環病毒物種分離株ctei657，完全基因體 |
| MH649106.1 | 指環病毒物種分離株ctca015，完全基因體 |
| MH649114.1 | 指環病毒物種分離株ctbf050，完全基因體 |
| MH649122.1 | 指環病毒物種分離株ctdc002，完全基因體 |
| MH649125.1 | 指環病毒物種分離株ctbb15，完全基因體 |
| MH649127.1 | 指環病毒物種分離株ctba013，完全基因體 |
| MH649137.1 | 指環病毒物種分離株ctbb000，完全基因體 |
| MH649141.1 | 指環病毒物種分離株ctbc019，完全基因體 |

| | |
|---|---|
| MH649142.1 | 指環病毒物種分離株ctid026，完全基因體 |
| MH649144.1 | 指環病毒物種分離株ctfj004，完全基因體 |
| MH649152.1 | 指環病毒物種分離株ctcj13，完全基因體 |
| MH649156.1 | 指環病毒物種分離株ctci006，完全基因體 |
| MH649157.1 | 指環病毒物種分離株ctbd025，完全基因體 |
| MH649158.1 | 指環病毒物種分離株ctbf005，完全基因體 |
| MH649161.1 | 指環病毒物種分離株ctcf045，完全基因體 |
| MH649165.1 | 指環病毒物種分離株ctcc29，完全基因體 |
| MH649169.1 | 指環病毒物種分離株ctib021，完全基因體 |
| MH649172.1 | 指環病毒物種分離株ctbh857，完全基因體 |
| MH649174.1 | 指環病毒物種分離株ctbj049，完全基因體 |
| MH649178.1 | 指環病毒物種分離株ctfc006，完全基因體 |
| MH649179.1 | 指環病毒物種分離株ctbe000，完全基因體 |
| MH649183.1 | 指環病毒物種分離株ctbb031，完全基因體 |
| MH649186.1 | 指環病毒物種分離株ctcb33，完全基因體 |
| MH649189.1 | 指環病毒物種分離株ctcc12，完全基因體 |
| MH649196.1 | 指環病毒物種分離株ctci060，完全基因體 |
| MH649199.1 | 指環病毒物種分離株ctbb017，完全基因體 |
| MH649203.1 | 指環病毒物種分離株cthc018，完全基因體 |
| MH649204.1 | 指環病毒物種分離株ctbj003，完全基因體 |
| MH649206.1 | 指環病毒物種分離株ctbg010，完全基因體 |
| MH649208.1 | 指環病毒物種分離株ctid008，完全基因體 |
| MH649209.1 | 指環病毒物種分離株ctbg056，完全基因體 |
| MH649210.1 | 指環病毒物種分離株ctda001，完全基因體 |
| MH649212.1 | 指環病毒物種分離株ctcf004，完全基因體 |
| MH649217.1 | 指環病毒物種分離株ctbe029，完全基因體 |
| MH649223.1 | 指環病毒物種分離株ctci016，完全基因體 |
| MH649224.1 | 指環病毒物種分離株ctce11，完全基因體 |
| MH649228.1 | 指環病毒物種分離株ctcf013，完全基因體 |
| MH649229.1 | 指環病毒物種分離株ctcb036，完全基因體 |
| MH649241.1 | 指環病毒物種分離株ctda027，完全基因體 |
| MH649242.1 | 指環病毒物種分離株ctbf003，完全基因體 |
| MH649254.1 | 指環病毒物種分離株ctjb007，完全基因體 |
| MH649255.1 | 指環病毒物種分離株ctbb023，完全基因體 |
| MH649256.1 | 指環病毒物種分離株ctca002，完全基因體 |
| MH649258.1 | 指環病毒物種分離株ctcg010，完全基因體 |
| MH649263.1 | 指環病毒物種分離株ctgh3，完全基因體 |
| MK012439.1 | 指環病毒物種分離株cthe000，完全基因體 |
| MK012440.1 | 指環病毒物種分離株ctjd008，完全基因體 |
| MK012448.1 | 指環病毒物種分離株ctch012，完全基因體 |
| MK012457.1 | 指環病毒物種分離株ctda009，完全基因體 |
| MK012458.1 | 指環病毒物種分離株ctcd015，完全基因體 |
| MK012485.1 | 指環病毒物種分離株ctfd011，完全基因體 |
| MK012489.1 | 指環病毒物種分離株ctba003，完全基因體 |

| | |
|---|---|
| MK012492.1 | 指環病毒物種分離株ctbb005，完全基因體 |
| MK012493.1 | 指環病毒物種分離株ctcj014，完全基因體 |
| MK012500.1 | 指環病毒物種分離株ctcb001，完全基因體 |
| MK012504.1 | 指環病毒物種分離株ctcj010，完全基因體 |
| MK012516.1 | 指環病毒物種分離株ctcf003，完全基因體 |
| NC_038336.1 | 細環病毒5分離株TCHN-C1 Orf2及Orf1基因，完全cd |
| NC_038338.1 | 細環病毒11分離株TCHN-D1 Orf2及Orf1基因，完全cd |
| NC_038339.1 | 細環病毒13分離株TCHN-A Orf2及Orf1基因，完全cd |
| NC_038340.1 | 細環病毒20 ORF4、ORF3、ORF2、ORF1基因，完全cd，純系：SAa-10 |
| NC_038341.1 | 細環病毒21分離株TCHN-B ORF2及ORF1基因，完全cd |
| NC_038342.1 | 細環病毒23 ORF2、ORF1基因，完全cd，分離株：s-TTV CH65-2 |
| NC_038343.1 | 細環病毒24 ORF4、ORF3、ORF2、ORF1基因，完全cd，純系：SAa-01 |
| NC_038344.1 | 細環病毒29 ORF2、ORF1、ORF3基因，完全cd，分離株：TTVyon-KC009 |
| NC_038345.1 | 微型細環病毒10分離株LIL-y1 ORF2、ORF1、ORF3及ORF4基因，完全cd |
| NC_038346.1 | 微型細環病毒11分離株LIL-y2 ORF2、ORF1及ORF3基因，完全cd |
| NC_038347.1 | 微型細環病毒12分離株LIL-y3 ORF2、ORF1、ORF3及ORF4基因，完全cd |
| NC_038350.1 | 細環中型病毒3分離株2PoSMA ORF2及ORF1基因，完全cd |
| NC_038351.1 | 細環中型病毒4分離株6PoSMA ORF2、ORF1及ORF3基因，完全cd |
| NC_038352.1 | 細環中型病毒5 DNA，完全基因體，分離株：MDJHem2 |
| NC_038353.1 | 細環中型病毒6 DNA，完全基因體，分離株：MDJHem3-1 |
| NC_038354.1 | 細環中型病毒7 DNA，完全基因體，分離株：MDJHem3-2 |
| NC_038355.1 | 細環中型病毒8 DNA，完全基因體，分離株：MDJN1 |
| NC_038356.1 | 細環中型病毒9 DNA，完全基因體，分離株：MDJN2 |
| NC_038357.1 | 細環中型病毒10 DNA，完全基因體，分離株：MDJN14 |
| NC_038358.1 | 細環中型病毒11 DNA，完全基因體，分離株：MDJN47 |
| NC_038359.1 | 細環中型病毒12 DNA，完全基因體，分離株：MDJN51 |
| NC_038360.1 | 細環中型病毒13 DNA，完全基因體，分離株：MDJN69 |
| NC_038361.1 | 細環中型病毒14 DNA，完全基因體，分離株：MDJN97 |
| NC_038362.1 | 細環中型病毒15 DNA，完全基因體，分離株：Pt-TTMDV210 |

【0280】在一些實施例中，遺傳元件包含與來自以下之一或多個序列具有同源性或一致性的一或多個序列：一或多個非指環病毒，例如腺病毒、疱疹病毒、痘病毒、痘瘡病毒、SV40、乳頭狀瘤病毒；RNA病毒，諸如反轉錄病毒，例如慢病毒；單股RNA病毒，例如肝炎病毒；或雙股RNA病毒，例如輪狀病毒。在一些實施例中，由於缺乏重組反轉錄病毒，因此可提供輔助以產生感染性粒子。此類輔助可例如藉由使用輔助細胞株來提供，該等輔助細胞株含有在LTR內之調控序列控制下編碼反轉錄

病毒之所有結構基因的質體。適用於複製本文所述之指環載體的細胞株包括此項技術中已知之細胞株，例如A549細胞，其可如本文所述地修飾。該遺傳元件可另外含有編碼可選標記之基因，以使得可鑑別所需遺傳元件。

【0281】 在一些實施例中，遺傳元件包括非靜默突變，例如在編碼多肽中產生胺基酸差異之鹼基取代、缺失或添加，只要序列保持與由第一核苷酸序列編碼之多肽至少約70%、75%、80%、85%、90%、95%、96%、97%、98%或99%一致或以其他方式適用於實踐本發明。就此而言，可進行通常認為不使總體蛋白質功能不活化的某些保守胺基酸取代，諸如關於帶正電荷的胺基酸(且反之亦然)：離胺酸、精胺酸及組胺酸；關於帶負電荷的胺基酸(且反之亦然)：天冬胺酸及麩胺酸；以及關於帶中性電荷之某些胺基酸群組(且在所有情況下，亦反之亦然)：(1)丙胺酸及絲胺酸，(2)天冬醯胺、麩醯胺酸及組胺酸，(3)半胱胺酸及絲胺酸，(4)甘胺酸及脯胺酸，(5)異白胺酸、白胺酸及纈胺酸，(6)甲硫胺酸、白胺酸及異白胺酸，(7)苯丙胺酸、甲硫胺酸、白胺酸及酪胺酸，(8)絲胺酸及蘇胺酸，(9)色胺酸及酪胺酸，以及(10)例如酪胺酸、色胺酸及苯丙胺酸。可根據物理特性及對二級與三級蛋白質結構的影響來對胺基酸進行分類。保守取代在此項技術中公認為一個胺基酸取代具有類似特性之另一胺基酸。

【0282】 具有相同或指定百分比之相同核苷酸或胺基酸殘基的兩個或更多個核酸或多肽序列之一致性(例如在指定區域上之約60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高一致性，當在比較窗口或指示區域上比較及比對最大一致性時)可使用BLAST或BLAST 2.0序列比較演算法在下文所述之

預設參數下，或藉由手動比對及目視檢查來量測(參見例如NCBI網站www.ncbi.nlm.nih.gov/BLAST/或其類似者)。一致性亦可指或可應用於測試序列之互補序列。一致性亦包括具有缺失及/或添加之序列以及具有取代之彼等者。如本文所述，演算法考慮空隙及其類似物。一致性可存在於以下區域上：其為至少約10個胺基酸或核苷酸長、約15個胺基酸或核苷酸長、約20個胺基酸或核苷酸長、約25個胺基酸或核苷酸長、約30個胺基酸或核苷酸長、約35個胺基酸或核苷酸長、約40個胺基酸或核苷酸長、約45個胺基酸或核苷酸長、約50個胺基酸或核苷酸長或更多個胺基酸或核苷酸長。由於遺傳密碼簡併，因此同源核苷酸序列可包括任意數目個靜默鹼基變化，亦即仍然編碼相同胺基酸的核苷酸取代。

【0283】

*蛋白質外部*

在一些實施例中，指環載體，例如合成指環載體，包含包裹遺傳元件之蛋白質外部。蛋白質外部可包含未能在哺乳動物中引發非所需免疫反應之實質上非致病性外部蛋白質。指環載體之蛋白質外部通常包含可自裝配成構成蛋白質外部之二十面體結構的實質上非致病性蛋白質。

【0284】 在一些實施例中，蛋白質外部蛋白質由指環載體之遺傳元件之序列(例如，與遺傳元件呈順式)編碼。在其他實施例中，蛋白質外部蛋白質由與指環載體之遺傳元件分離之核酸(例如，與遺傳元件呈反式)編碼。

【0285】 在一些實施例中，蛋白質，例如實質上非致病性蛋白質及/或蛋白質外部蛋白質，包含一或多個醣基化胺基酸，例如2、3、4、5、6、7、8、9、10或更多個。

【0286】 在一些實施例中，蛋白質，例如實質上非致病性蛋白質及/或蛋白質外部蛋白質，包含至少一個親水性DNA結合區、富含精胺酸之區域、富含蘇胺酸之區域、富含麩醯胺之區域、N端聚精胺酸序列、可變區、C端聚麩醯胺酸/麩胺酸序列及一或多個二硫橋鍵。

【0287】 在一些實施例中，蛋白質為蛋白殼蛋白，例如具有以下之序列：與由編碼本文所述之蛋白殼蛋白，例如指環病毒ORF1分子及/或蛋白殼蛋白序列，例如如本文所述，的核苷酸序列中之任一者編碼之蛋白質具有至少約60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性。在一些實施例中，蛋白質或蛋白殼蛋白之功能片段由與指環病毒ORF1核酸，例如如本文所述，具有至少約60%、70%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核苷酸序列編碼。

【0288】 在一些實施例中，指環載體包含編碼蛋白殼蛋白或蛋白殼蛋白之功能片段之核苷酸序列，或與如本文所述之指環病毒ORF1分子具有至少約60%、70%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的序列。

【0289】 在一些實施例中，具有較低序列一致性之胺基酸的範圍可提供本文所述之特性及細胞/組織/物種特異性(例如向性)之差異中之一或多者。

【0290】 在一些實施例中，指環載體缺乏蛋白質外部之脂質。在一些實施例中，指環載體缺乏脂質雙層，例如病毒包封。在一些實施例中，指環載體之內部完全藉由蛋白質外部覆蓋(例如，100%覆蓋)。在一些實施例中，指環載體之內部低於100%由蛋白質外部覆蓋，例如95%、

90%、85%、80%、70%、60%、50%或更低覆蓋度。在一些實施例中，蛋白質外部包含間隔或不連續處，例如容許對水、離子、肽或小分子之滲透性，只要遺傳元件保留於指環載體中。

【0291】在一些實施例中，蛋白質外部包含一或多種蛋白質或多肽，其特異性識別及/或結合宿主細胞，例如互補蛋白質或多肽，以介導遺傳元件進入宿主細胞中。

【0292】在一些實施例中，蛋白質外部包含以下中之一或多者：富含精胺酸之區域、果醬捲(jelly-roll)區、N22域、高變區及/或C端域，例如ORF1分子，例如如本文所述。在一些實施例中，蛋白質外部包含以下中之一或多者：一或多個醣基化蛋白質、親水性DNA結合區、富含精胺酸之區域、富含蘇胺酸之區域、富含麩醯胺之區域、N端聚精胺酸序列、可變區、C端聚麩醯胺酸/麩胺酸序列及一或多個二硫橋鍵。舉例而言，蛋白質外部包含由指環病毒ORF1核酸，例如如本文所述，編碼之蛋白質。

【0293】在一些實施例中，蛋白質外部包含以下特徵中之一或多者：二十面體對稱性，識別及/或結合與一或多個宿主細胞分子相互作用以介導進入宿主細胞中之分子，缺乏脂質分子，缺乏碳水化合物，pH及溫度穩定性、耐清潔劑，及在宿主中為實質上非免疫原性或非致病性的。

【0294】

**III.使用方法**

本文所述之指環載體及包含指環載體之組合物可用於治療例如有需要之個體(例如哺乳動物個體，例如人類個體)之病症、疾病或病狀的方法中。投與本文所述之醫藥組合物可例如藉助於非經腸(包括靜脈內、瘤內、腹膜內、肌肉內、腔內及皮下)投與。指環載體可單獨投與或調配為

醫藥組合物。

【0295】 指環載體可以單位劑量組合物，諸如單位劑量非經腸組合物形式投與。此類組合物通常藉由摻合來製備，且可適合地經調適用於非經腸投與。此類組合物可例如呈可注射及可輸注溶液或懸浮液或栓劑或氣溶膠形式。

【0296】 在一些實施例中，投與指環載體或包含該指環載體(例如如本文所述)之組合物可使得將包含指環載體之遺傳元件遞送至目標細胞，例如個體之目標細胞。

【0297】 本文所述之指環載體或其組合物，例如包含效應子(例如，內源性或外源性效應子)，可用於將效應子遞送至細胞、組織或個體。在一些實施例中，指環載體或其組合物用於將效應子遞送至骨髓、血液、心臟、GI或皮膚。藉由投與本文所述之指環載體組合物遞送效應子可調節(例如增加或減少)細胞、組織或個體之非編碼RNA或多肽的表現量。以此方式調節表現量可引起使效應子遞送至細胞中之功能活性改變。在一些實施例中，經調節功能活性在本質上可為酶、結構或調控性的。

【0298】 在一些實施例中，指環載體或其複本可在遞送至細胞中之後24小時(例如，1天、2天、3天、4天、5天、6天、1週、2週、3週、4週、30天或1個月)在細胞中偵測到。在實施例中，指環載體或其組合物介導對目標細胞之作用，且該作用持續至少1、2、3、4、5、6或7天，2、3或4週，或1、2、3、6或12個月。在一些實施例中(例如，其中該指環載體或其組合物包含編碼外源性蛋白質之遺傳元件)，該作用持續低於1、2、3、4、5、6或7天，2、3或4週，或1、2、3、6或12個月。

【0299】 可用本文所述之指環載體或包含指環載體之組合物處理的

疾病、病症及病狀的實例包括但不限於：免疫病症、干擾素病變(例如，I型干擾素病變)、傳染病、發炎性病症、自體免疫病狀、癌症(例如，實體腫瘤，例如肺癌；非小細胞肺癌，例如表現對mIR-625反應之基因，例如凋亡蛋白酶-3的腫瘤)及腸胃疾病。在一些實施例中，指環載體調節(例如提高或降低)與指環載體接觸之細胞中之活性或功能。在一些實施例中，指環載體調節(例如提高或降低)與指環載體接觸之細胞中之分子(例如核酸或蛋白質)之含量或活性。在一些實施例中，指環載體減小與指環載體接觸之細胞(例如癌細胞)的存活率，例如至少約10%、20%、30%、40%、50%、60%、70%、80%、90%、95%、99%或更高。在一些實施例中，指環載體包含與指環載體接觸之細胞(例如癌細胞)的存活率，例如至少約10%、20%、30%、40%、50%、60%、70%、80%、90%、95%、99%或更高的效應子，例如miRNA，例如miR-625。在一些實施例中，指環載體例如藉由提高凋亡蛋白酶-3活性來提高與指環載體接觸之細胞(例如癌細胞)的細胞凋亡，例如提高至少約10%、20%、30%、40%、50%、60%、70%、80%、90%、95%、99%或更多。在一些實施例中，指環載體包含例如藉由提高凋亡蛋白酶-3活性來提高與指環載體接觸之細胞(例如癌細胞)的細胞凋亡，例如至少約10%、20%、30%、40%、50%、60%、70%、80%、90%、95%、99%或更高的效應子，例如miRNA，例如miR-625。

【0300】

**IV.投與/遞送**

該組合物(例如，包含如本文所述之指環載體的醫藥組合物)可經調配以包括醫藥學上可接受之賦形劑。醫藥組合物可視情況包含一或多種額外

活性物質，例如治療性及/或預防性活性物質。本發明之醫藥組合物可無菌及/或不含熱原質。在調配及/或製造藥劑中之一般考慮因素可見於例如Remington: The Science and Practice of Pharmacy 第21版, Lippincott Williams & Wilkins, 2005 (以引用之方式併入本文中)中。

【0301】 儘管描述本文所提供之醫藥組合物主要針對適用於向人類投與之醫藥組合物，但熟習此項技術者應理解，此類組合物一般適用於向任何其他動物投與，例如向非人類動物，例如非人類哺乳動物投與。應充分理解，為使組合物適用於向各種動物投與，對適用於向人類投與之醫藥組合物進行修改，且一般熟練的獸醫藥理學家可僅用一般實驗(若存在)設計及/或進行此類修改。醫藥組合物之投藥所涵蓋的個體包括但不限於人類及/或其他靈長類動物；哺乳動物，包括商業相關之哺乳動物，諸如牛、豬、馬、綿羊、貓、狗、小鼠及/或大鼠；及/或鳥類，包括商業相關之鳥類，諸如家禽、雞、鴨、鵝及/或火雞。

【0302】 本文所述之醫藥組合物的調配物可藉由藥理學技術中已知或此後研發之任何方法來進行製備。一般而言，此類製備方法包括使活性成分與賦形劑及/或一或多種其他附屬成分結合，且隨後必要時及/或需要時將產物分割、成型及/或封裝之步驟。

【0303】 在一個態樣中，本發明提供一種向個體遞送指環載體之方法。該方法包括向個體投與包含如本文所述之指環載體的醫藥組合物。在一些實施例中，所投與之指環載體在個體中複製(例如，成為個體之病毒體的一部分)。

【0304】 醫藥組合物可包括野生型或天然病毒元件及/或經修飾病毒元件。指環載體可包括一或多個指環病毒序列(例如，核酸序列或編碼其

胺基酸序列之核酸序列)或與其具有至少約60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%及99%核苷酸序列一致性的序列。指環載體可包含有包含與一或多個指環病毒序列(例如，指環病毒ORF1核酸序列)具有至少約60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%及99%序列一致性之核酸序列的核酸分子。指環載體可包含編碼與指環病毒胺基酸序列(例如，指環病毒ORF1分子之胺基酸序列)具有至少約60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%及99%序列一致性之胺基酸序列的核酸分子。指環載體可包含有包含與指環病毒胺基酸序列(例如，指環病毒ORF1分子之胺基酸序列)具有至少約60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%及99%序列一致性之胺基酸序列的多肽。

【0305】 在一些實施例中，指環載體足以提高(刺激)內源性基因及蛋白質表現，例如與參考，例如健康對照相比，提高至少約5%、10%、15%、20%、25%、30%、35%、40%、45%、50%或更多。在某些實施例中，指環載體足以降低(抑制)內源性基因及蛋白質表現，例如與參考，例如健康對照相比，降低至少約5%、10%、15%、20%、25%、30%、35%、40%、45%、50%或更多。

【0306】 在一些實施例中，指環載體抑制/增強宿主或宿主細胞中之一或多種病毒特性，例如向性、感染性、免疫抑制/活化，例如與參考，例如健康對照相比，至少約5%、10%、15%、20%、25%、30%、35%、40%、45%、50%或更多。

【0307】 在一些實施例中，向個體投與進一步包含未在病毒遺傳資訊中表示之一或多個病毒株的醫藥組合物。

【0308】在一些實施例中，包含本文所述之指環載體的醫藥組合物以足以調節病毒感染之劑量及時間投與。病毒感染之一些非限制性實例包括腺相關病毒、愛知病毒(Aichi virus)、澳大利亞蝙蝠狂犬病毒、BK多瘤病毒、班納病毒(Banna virus)、巴馬森林病毒(Barmah forest virus)、布尼安維拉病毒(Bunyamwera virus)、拉克羅斯本揚病毒(Bunyavirus La Crosse)、雪足野兔布尼亞病毒(Bunyavirus snowshoe hare)、獼猴皰疹病毒(Cercopithecine herpesvirus)、金迪普拉病毒(Chandipura virus)、屈公病毒(Chikungunya virus)、科薩病毒A (Cosavirus A)、牛痘病毒(Cowpox virus)、柯薩奇病毒(Coxsackievirus)、克里米亞-岡果出血熱病毒(Crimean-Congo hemorrhagic fever virus)、登革熱病毒(Dengue virus)、多理病毒(Dhori virus)、達格畢病毒(Dugbe virus)、杜文海格病毒(Duvenhage virus)、東部馬腦炎病毒、伊波拉病毒(Ebolavirus)、腸道細胞病變性人類孤兒病毒(Echovirus)、腦心肌炎病毒、埃-巴二氏病毒(Epstein-Barr virus)歐洲蝙蝠狂犬病毒(European bat lyssavirus)、GB病毒C/G型肝炎病毒、漢坦病毒(Hantaan virus)、亨德拉病毒(Hendra virus) A型肝炎病毒、B型肝炎病毒、C型肝炎病毒、E型肝炎病毒、D型肝炎病毒、馬痘病毒、人類腺病毒、人類星狀病毒、人類冠狀病毒、人類巨細胞病毒、人類腸病毒68、人類腸病毒70、人類皰疹病毒1、人類皰疹病毒2、人類皰疹病毒6、人類皰疹病毒7、人類皰疹病毒8、人類免疫缺乏病毒、人類乳頭狀瘤病毒1、人類乳頭狀瘤病毒2、人類乳頭狀瘤病毒16、人類乳頭狀瘤病毒18、人類副流感、人類小病毒B19、人類呼吸道合胞病毒、人類鼻病毒、人類SARS冠狀病毒、人類泡沫病毒、人類嗜T淋巴球病毒、人類環曲病毒、A型流感病毒、B型流感病毒、C型流感病毒、伊斯

法罕病毒(Isfahan virus)、JC多瘤病毒、日本腦炎病毒、胡寧沙粒病毒(Junin arenavirus)、KI多瘤病毒、庫京病毒(Kunjin virus)、拉各斯蝙蝠病毒(Lagos bat virus)維多利亞湖馬堡病毒(Lake Victoria marburgvirus)、蘭加特病毒(Langat virus)、賴薩熱病毒(Lassa virus)、洛茲達雷病毒(Lordsdale virus)、跳躍病病毒(Louping ill virus)、淋巴球性脈絡叢腦膜炎病毒、馬丘波病毒(Machupo virus)、馬雅羅病毒(Mayaro virus)、MERS冠狀病毒、麻疹病毒、門戈腦心肌炎病毒(Mengo encephalomyocarditis virus)、梅克爾細胞多瘤病毒(Merkel cell polyomavirus)、莫科拉病毒(Mokola virus)、傳染性軟疣病毒、猴痘病毒(Monkeypox virus)、腮腺炎病毒、莫雷谷線腦炎病毒(Murray valley encephalitis virus)、紐約病毒(New York virus)、立百病毒(Nipah virus)、諾沃克病毒(Norwalk virus)、奧尼永-尼永病毒(O'nyong-nyong virus)、口瘡病毒(Orf virus)、奧羅普切病毒(Oropouche virus)、皮欽德病毒(Pichinde virus)、脊髓灰白質炎病毒、龐塔托魯白蛉熱病毒(Punta toro phlebovirus)、撲嗎拉病毒(Puumala virus)、狂犬病病毒、東非瑞夫特河谷羊熱病病毒(Rift valley fever virus)、羅沙病毒A (Rosavirus A)、羅斯河病毒(Ross river virus)、輪狀病毒A (Rotavirus A)、輪狀病毒B、輪狀病毒C、德國麻疹病毒(Rubella virus)、鷺山病毒(Sagiyama virus)、薩利病毒A (Salivirus A)、西西里糠蚊熱病毒(Sandfly fever sicilian virus)、箚幌病毒(Sapporo virus)、勝利基森林病毒(Semliki forest virus)、漢城病毒(Seoul virus)、猴泡沫病毒(Simian foamy virus)、猴病毒5、辛得比斯病毒(Sindbis virus)、南安普頓病毒(Southampton virus)、聖路易腦炎病毒(St. louis encephalitis virus)、蜱傳播波瓦生病毒(Tick-borne powassan

virus)、細環病毒、托斯卡納病毒(Toscana virus)、尤庫尼米病毒(Uukuniemi virus)、痘瘡病毒(Vaccinia virus)、水痘帶狀皰狀病毒、天花病毒、委內瑞拉馬腦炎病毒(Venezuelan equine encephalitis virus)、水泡性口炎病毒、西部馬腦炎病毒、WU多瘤病毒、西尼羅河病毒(West Nile virus)、亞巴猴腫瘤病毒(Yaba monkey tumor virus)、亞巴樣病病毒、黃熱病病毒(Yellow fever virus)及茲卡病毒(Zika Virus)。在某些實施例中，指環載體足以勝過及/或取代已存在於個體中之病毒，例如與參考相比至少約5%、10%、15%、20%、25%、30%、35%、40%、45%、50%或更多。在某些實施例中，指環載體足以與慢性或急性病毒感染競爭。在某些實施例中，可預防性投與指環載體以保護免於病毒感染(例如前病毒)。在一些實施例中，指環載體之量足以調節(例如，表型、病毒含量、基因表現、與其他病毒競爭、疾病病狀等，至少約5%、10%、15%、20%、25%、30%、35%、40%、45%、50%或更多)。在一些實施例中，治療(treatment)、治療(treating)及其同源詞包含個體之醫療管理(例如，藉由投與指環載體，例如如本文所述製成之指環載體)，例如意欲改善、改進、穩定、預防或治癒疾病、病理性病狀或病症。在一些實施例中，治療包含主動性治療(針對改善疾病、病理性病狀或病症的治療)、病因性治療(針對相關疾病、病理性病狀或病症之病因的治療)、緩解性治療(針對症狀緩解而設計之治療)、預防性治療(針對預防、最小化或部分或完全地抑制相關疾病、病理性病狀或病症之發展的治療)及/或支持性治療(用於補充另一療法之治療)。

【0309】 本文中引用之所有參考文獻及公開案均以引用之方式併入本文中。

【0310】 提供以下實例以進一步說明本發明之一些實施例，但不意欲限制本發明之範疇；藉由其示例性之性質，應瞭解，可替代地使用熟習此項技術者已知之其他程序、方法或技術。

【0311】

**實例**

**目錄**

實例1：使用利用桿狀病毒產生之ORF1的指環載體之活體外組裝

實例2：蛋白殼化mRNA之基於環2 ORF1之指環載體的活體外組裝

實例3：使用經修飾之ORF1蛋白的mRNA蛋白殼化指環載體之活體外組裝

實例4：使用經修飾之mRNA的mRNA蛋白殼化指環載體之活體外組裝

實例5：指環病毒ORF1蛋白殼蛋白之結構分析

【0312】

**實例1：使用利用桿狀病毒產生之ORF1的指環載體之活體外組裝**

在此實例中，藉由活體外組裝產生適用於表現指環病毒蛋白質(例如ORF1)之桿狀病毒構築體。

【0313】 在第一個實例中，編碼與N端$HIS_6$-標籤融合之環2 ORF1(HIS-ORF1)的DNA經密碼子最佳化以用於昆蟲表現，且根據製造商說明書(ThermoFisher Scientific)選殖至桿狀病毒表現載體pFASTbac系統中。用環2 HIS-ORF1桿狀病毒感染10公升昆蟲細胞培養物(Sf9)且在感染後3天藉由離心收集細胞。溶解細胞且使用本領域中之標準方法，使用螯合樹脂管柱純化溶解物。對含有HIS-ORF1之溶離份進行滲析且用DNA酶處理

以消化宿主細胞DNA。所得物質再次使用螯合樹脂管柱純化且保留含有ORF1之溶離份用於核酸蛋白殼化及病毒載體純化。亦藉由陰性染色電子顯微法分析含ORF1溶離份。

【0314】 在第二個實例中，編碼與N端$HIS_6$-標籤融合之環10 ORF1 (HIS-ORF1)的DNA經密碼子最佳化以用於昆蟲表現，且根據製造商說明書(ThermoFisher Scientific)選殖至桿狀病毒表現載體pFASTbac系統中。用環HIS-ORF1桿狀病毒感染昆蟲細胞(Sf9)且在感染後3天藉由離心收集細胞。溶解細胞且使用本領域之標準方法，使用螯合樹脂親和管柱(HisTrap，GE Healthcare)純化蛋白質。所得物質再次使用肝素親和管柱(Heparin HiTrap, GE Healthcare)純化且含有ORF1之溶離份藉由陰性染色電子顯微法分析。

【0315】 在第三個實例中，編碼與N端$HIS_6$-Flag-標籤融合之雞貧血病毒(CAV)蛋白殼蛋白(Vp1) (HIS-Flag-Vp1)及輔助蛋白質(Vp2)的DNA經密碼子最佳化以用於哺乳動物表現，且使用CMV啟動子選殖至哺乳動物表現載體中。哺乳動物細胞(293expi)用CAV Vp1及Vp2表現載體轉染。在感染後3天藉由離心收集細胞。溶解細胞且使用螯合及肝素純化製程純化溶解物。藉由陰性染色電子顯微法分析含有雞貧血病毒(CAV) Vp1之溶離份。

【0316】 如圖1中所示，環2 ORF1及環10 ORF1均展示出形成約35 nm病毒樣粒子的傾向。

【0317】

**核酸蛋白殼化及病毒載體純化：**環ORF1 (野生型蛋白質、嵌合蛋白質或其片段)將用足以解離VLP或病毒蛋白殼之條件處理，以使得能夠與

核酸承載物重新組裝。核酸承載物可定義為例如編碼希望作為治療劑遞送之相關基因的RNA。界定濃度之核酸承載物將與界定濃度之環ORF1組合，且用足以容許核酸蛋白殼化之條件處理，且定義為病毒載體之所得粒子隨後將使用標準病毒純化程序純化。

【0318】

**實例2：蛋白殼化mRNA之基於環2 ORF1之指環載體的活體外組裝**

環2 ORF1藉由尺寸排阻層析(SEC)用移動相純化，該等移動相包括具有500 mM NaCl之Tris pH 8.0、具有500 mM NaCl及0.1% SDS之Tris pH 8.0、具有150 mM NaCl之CAPS緩衝液pH 10.5、具有500 mM NaCl之CAPS緩衝液pH 10.5或具有500 mM NaCl及0.1% SDS之CAPS緩衝液pH 10.5，以將病毒粒子或VLP解離成分散的蛋白質或次蛋白衣。

【0319】在第一個實例中，ORF1與mRNA，亦即經螢光標記之mRNA或以化學方式結合至實例1中所示之ssDNA之區段的mRNA轉殖基因混合，以能夠誘導載體形成。病毒載體經由滲析形成，且使用Tris pH 8.0緩衝液進行SEC純化，以分離蛋白殼化RNA之指環載體(例如如藉由保留之螢光吸收所量測)。藉由諸如DLS或電子顯微法之生物物理學評定進一步評估指環載體組裝。

【0320】在第二個實例中，用含0.1% SDS之1 M NaCl處理經純化ORF1，將寡聚物或VLP解離成分散的蛋白質或次蛋白衣。ORF1隨後與mRNA，諸如轉譯相關基因(例如報導基因，例如GFP、mCherry；或相關效應子，例如EPO)之mRNA混合，且針對具有150 mM NaCl之Tris pH 8.0經滲析以容許VLP形成。後續複合物藉由SEC使用Tris pH 8.0緩衝液純化以分離蛋白殼化mRNA之AV載體。指環載體的載體組裝可進一步藉由活

體外或活體內讀數，例如藉由轉導細胞及觀測報導基因(例如mCherry或GFP)之表現或經由相關效應子之表現(例如使用ELISA偵測基因，諸如EPO之表現)進行評估。

【0321】

*蛋白殼化GFP mRNA之基於環2 ORF1之指環載體的活體外組裝*

在另一實例中，環2 ORF1蛋白表現為昆蟲細胞中之全長蛋白，且經組裝VLP藉由肝素親和管柱隨後尺寸排阻層析(SEC)，使用Tris緩衝液移動相純化。由經分離環2 ORF1蛋白形成的VLP用陰性染色電子顯微法(EM)來觀測，且估計粒子效價為$10^{10}$個粒子/毫升(pts/ml；圖4A)。用2莫耳(2 M)尿素處理VLP以拆卸VLP。藉由EM再成像展示未觀測到VLP (圖4B)。經尿素處理之VLP隨後在不存在mRNA之情況下(圖4C)或在約10×過量的編碼GFP之mRNA存在下(圖4D及圖4E)經滲析以移除尿素。對於用尿素處理且在不存在mRNA之情況下經滲析之VLP樣品，藉由EM觀測到極少粒子(小於$10^8$個粒子/毫升；圖4C)。相比之下，藉由EM，在存在過量mRNA之情況下滲析產生粒子效價實質上較高之觀測結果(約$10^9$至$10^{10}$個粒子/毫升；圖4D及圖4E)。此等資料表明VLP之拆卸及重新組裝在mRNA存在下更有效，且指示指環病毒ORF1蛋白可用於活體外蛋白殼化mRNA以形成指環載體。

【0322】

**實例3：使用經修飾之ORF1蛋白的mRNA蛋白殼化指環載體之活體外組裝**

在此實例中，mRNA遺傳元件之封裝藉由修飾ORF1蛋白以具有結合mRNA之接觸殘基來改良。在此實例中，ssDNA接觸殘基及/或接觸

ssDNA之果醬捲β股及/或N端富含精胺酸之模體(ARM)可經mRNA結合病毒蛋白或其他mRNA結合蛋白之組分置換以容許mRNA之有效結合及封裝。接著用含0.1% SDS之1 M NaCl處理此結合mRNA的嵌合ORF1以將寡聚物或VLP解離成分散的蛋白質或次蛋白衣。隨後嵌合ORF1與mRNA，諸如編碼相關基因(例如報導基因，例如GFP、mCherry；或相關效應子，例如EPO)之mRNA混合，且針對具有150 mM NaCl之Tris pH 8.0經滲析以容許VLP形成。後續複合物藉由SEC使用Tris pH 8.0緩衝液純化以分離蛋白殼化mRNA之指環載體。指環載體組裝可進一步藉由活體外或活體內讀數，例如藉由轉導細胞及觀測報導基因(例如mCherry或GFP)之表現或經由相關效應子之表現(諸如使用ELISA偵測基因，諸如EPO之表現)進行評估。

【0323】 **對ORF1分子之例示性修飾：**可用於本文所描述之方法中的環ORF1分子包括例如若干野生型指環病毒ORF1蛋白；CAV蛋白殼蛋白(VP1)變異體；具有改良組裝效率、產量或穩定性之突變的指環病毒ORF1蛋白；及嵌合ORF1株或其功能片段。在一些情況下，親和標籤連接至ORF1分子，例如在N端(SEQ ID NO: 561-562)。在一些情況下，ORF1分子係未標記蛋白質。環ORF1分子可單獨或與任何數目之輔助蛋白質組合表現，該等輔助蛋白質包括但不限於指環病毒ORF2及/或ORF3蛋白。

【0324】 具有改良組裝效率之突變的環ORF1蛋白可包括但不限於具有引入至N端富含精胺酸之模體(ARM)中之突變的ORF1蛋白，該等突變例如改變ARG臂之pI，其可容許pH敏感性核酸結合以觸發粒子組裝(SEQ ID NO: 563-565)。改良穩定性之ORF1突變可以包括例如接觸典型果醬捲β桶形之β股F及G (F及G β股)的內原聚體之突變，例如以改變原聚

體表面之疏水性狀態及/或使蛋白殼形成在熱力學上更為有利。

【0325】 嵌合ORF1蛋白可包括但不限於其序列之一或多個部分經另一蛋白殼蛋白，諸如BFDV蛋白殼蛋白、E型肝炎蛋白殼蛋白(例如ARG臂及/或F及G β股，或其可比組分)之可比部分置換的ORF1蛋白。嵌合ORF1蛋白亦可包括其序列之一或多個部分經另一指環病毒ORF1蛋白之可比部分置換(諸如環2 ORF1之果醬捲片段或C端部分經環9 ORF1之可比部分置換；參見例如SEQ ID NO: 568-575)的ORF1蛋白。

【0326】 一般而言，ORF1分子可使用純化技術純化，包括但不限於螯合純化、肝素純化、梯度沈降純化及/或SEC純化。

【0327】

**實例4：使用經修飾之mRNA的mRNA蛋白殼化指環載體之活體外組裝**

在此實例中，基於mRNA之遺傳元件之蛋白殼化係藉由使mRNA分子結合至ssDNA，或藉由修飾mRNA轉殖基因，以使得主鏈之一段容許與野生型ORF1之ssDNA接觸殘基結合的方式而最佳化。mRNA一般編碼相關基因，諸如報導基因(例如GFP或mCherry)及/或效應基因(例如EPO)。

【0328】 在一個實例中，將可藉助於其糖鏈主鏈結合ORF1但亦可與mRNA非共價配對之經修飾ssDNA與相關mRNA混合以產生mRNA/DNA複合物。此mRNA/DNA複合物接著可使用環ORF1蛋白殼化以形成例如如下所述之指環載體。

【0329】 在另一實例中，mRNA分子用具有容許藉由ORF1結合及蛋白殼化之DNA主鏈的mRNA分子之一或多段合成，同時保留編碼待遞送之基因(例如，報導基因或效應基因)的mRNA部分。此mRNA/DNA雜合分子接著可使用環ORF1蛋白殼化以形成例如如下所述之指環載體。

【0330】 **藉由活體外組裝之蛋白殼化：**上文所描述之mRNA/DNA遺傳元件隨後藉由活體外組裝進行蛋白殼化。簡言之，接著用含0.1% SDS之1 M NaCl處理指環載體ORF1以將寡聚物或VLP解離成分散的蛋白質或次蛋白衣。隨後ORF1與合成mRNA複合物或雜合分子混合，且針對具有150 mM NaCl之Tris pH 8.0經滲析以容許VLP形成。後續粒子藉由SEC使用Tris pH 8.0緩衝液純化以分離蛋白殼化mRNA之指環載體。指環載體組裝可藉由活體外或活體內讀數，藉由轉導細胞及觀測報導基因或效應基因之表現來進一步評估，例如如本文所描述。

【0331】

**實例5：指環病毒ORF1蛋白殼蛋白之結構分析**

指環病毒與其他充分表徵之病毒，諸如禽類病原體喙羽病病毒(BFDV)或雞貧血病毒(CAV)共有預測的結構特徵。

【0332】 指環病毒ORF1蛋白殼蛋白含有與BFDV的類似之N端ARM序列。二級結構預測展示ORF1之前約250個殘基視病毒株而定包括8個預測β股(圖3)。當遵循果醬捲域命名定則命名為B至I的ORF1之8個預測β股藉由二級結構預測引導對準至BFDV及Hep E之蛋白殼蛋白時，ORF1中之保守離胺酸及精胺酸殘基與BFDV及Hep E蛋白殼蛋白之已知ssDNA接觸殘基對準(圖3，由星號表示)。

【0333】

**序列**

*下文列出之序列如下註釋。加粗且帶下劃線的文字指示包含用於螯合純化之His6標籤(HHHHHH)及例如用於對低表現蛋白質進行西方墨點法偵測之Flag標籤(DYKDDDDK)（強抗原決定基)的序列。加粗及斜體序*

列指示環9 ORF1序列或其部分。加粗、非加下劃線序列為環2序列或其部分。未加粗、帶下劃線的序列來自豬肘病病毒(BFDV)。灰色高光顯示指示離胺酸至組胺酸突變之位置，例如在環9 ORF1之富含精胺酸之區域及第一β股中。

SEQ ID NO: 561：

環2 N端HIS-FLAG-3C蛋白酶-ORF1：

MGSSHHHHHHGSDYKDDDDKSGSLEVLFQGPSGMPYYYRRRKYNYRRPRWYGRGWIRRPLRRRLRRKRKRVRPLYY
TIPLKQWQPPYKRTCYIKGQDCLIYYSNLRLGMNSTMYLKSIVPVLIWPGGGSLSVSMLTLDALYDLHKLCRNWWTSTN
QDLPLVRYKGCKLTLYQSTLTDYIVRLLTLLPANSNKLTYPNTLIPLMMMMSKYKLIIPSRQTRRKKKPYTKLVKPPPQLF
NKWYLATDLYKIPLLQLICTACNLQNPLVKPDKLSNNVLTWSLNTISQNRNMSVDQGQSWPLKLGTQSLYLYLYGAN
LPGDTTQIPVADLLPLTNPRINRPGQSLNLAKTDLHTLTLYKNKLTNYWGNPLNKHIQELLDMILYSLKSPLAIKNLWT
TLNMKWNQLNNAGTMALTPLNLPLTQIQYNPDRDTGLDTQLYTLSNATGTGWDPPGIPLTLEGLPLWLIYWGLAD
LQKNLKKVTNIDTNYMLVAKTKLTQKPGTLYLVLLNDTLVLGNSPYLKQPLPLDNIKWYPQVQYQLLAQNKLLQTGPFT
PNIQGQLSDNISMLYKLYLKWGGSPPKAINVLNPALIQIQYPIPRNLLLETTSLQSPGLAPLSILYSLDYRLIGNYTTTAISRL
SQDWALKDTVSKTLPDRQQLLKQALLCLQSLLTQLKKLKVQQLISNLRQQQQLYRLRLISLLKDQ

SEQ ID NO: 562：

環9 N端HIS-FLAG-3C蛋白酶-ORF1：

MGSSHHHHHHGSDYKDDDDKSGSL*EVLFQGPSG*MPYYWRQKYYRRRYRPLSWRTRRIIQRRKRWRYRKPRKTY
WRRKTRVRKRLYKGRKTKKIVLKQLQPKHRRGTLGTICLLQGSPLRANINNYIQTIYSYVPDKLPGGGGWTLTLSLSSL
WLDWFLILKNVWTQSNAGLPLVRYGGVTLYLYQSAYTDYIAQVLNCYPMTDTKYTLADSAPNRMILIKKLVIRVPS
RETRKKRKKPYKRWRVGGPSQMQNKWYLQRDICLPLIMIAATAVDLRYPLGASDCASNNLTLTCLNPTLLQNQQPLDL
PSDTQGYLPKPGVYLYSTQRSNKLPSSSDCIYLGNTKDNQLGKSASSLMTLKTQKLTDWGNPLWLIYYIDGSKKLSYLK
PPSQLDSSDLLELMLLALPMLIQVRYNPLRDTGQGNLIYWTLNLAGQLAVDPPSSDNLKLDGLPLYDMCWGLDWL
LKVHLNLMLLNYCLQKSSALNLKCVLLPVDLSLLTGLSPYLTPVKSSDQAALWLLPQRLQLKSLNDKLTGPGCARSP
YGNYMQAKMSYKLLHVKWGGCPKCYLKPYDPCSQPNWVLPLINLNLLTQIQNLPNLCPQTLLQLWDWRRDLVLKKAL
ERLRQHLLPHLLTQLSLGSKLNLPVLRQLSPWLDSLLDSLLLKDQLQLQLQLNKLLKHQQHLKQQLKQYLKPQNLL

SEQ ID NO: 563：

具有環9之ARG替代環2 ORF1 (環291)

MGSSHHHHHHGSDYKDDDDKSGSLEVLFQGPSG*MPYYWRQKYYRRRYRPLSWRTRRIIQRRKRWRYRKPRKTY
WRRKTRVRKRRVHPLYTTIPLKQWQPPYKRTCYIKGQDCLIYYSNLRLGMNSTMYLKSIVPVLIWPGGGSLSVSMLTLD
ALYDLHKLCRNWWTSTNQDLPLVRYKGCKLTLYQSTLTDYIVRLLTLLPANSNKLTYPNTLIPLMMMMSKYKLIIPSRQT
RRKKKPYTKLVKPPPQLFNKWYLATDLYKIPLLQLICTACNLQNPLVKPDKLSNNVLTWSLNTISQNRNMSVDQGQS
WPLKLGTQSLYLYLYGANLPGDTTQIPVADLLPLTNPRINRPGQSLNLAKTDLHTLTLYKNKLTNYWGNPLNKHIQE
LLDMILYSLKSPLAIKNLWTTLNMKWNQLNNAGTMALTPLNLPLTQIQYNPDRDTGLDTQLYTLSNATGTGWDPPG
IPLTLEGLPLWLIYWGLADLQKNLKKVTNIDTNYMLVAKTKLTQKPGTLYLVLLNDTLVLGNSPYLKQPLPLDNIKWYP
QVQYQLLAQNKLLQTGPFTPNIQGQLSDNISMLYKLYLKWGGSPPKAINVLNPALIQIQYPIPRNLLLETTSLQSPGLAPL
SILYSLDYRLIGNYTTTAISRLSQDWALKDTVSKTLPDRQQLLKQALLCLQSLLTQLKKLKVQQLISNLRQQQQLYRLRLI
SLLKDQ

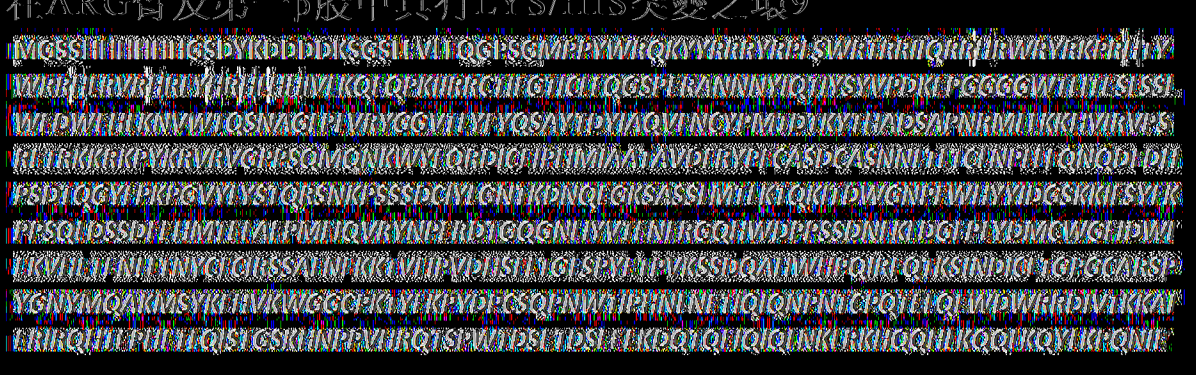
SEQ ID NO: 564：

具有環9之ARG腎及β股1 + 2 722抗原決定基之環2 ORF1 (環292)

MGSSHHHHHHGSDYKDDDDKSGSLEVLFQGPSGMPPYWHQKYYRRRYRPLSWRTRAHQRHKRWRYRKPRKTY
WHRKKLAVRKRLYKHKKKKWLKQLQPKHRRCLHGTCLLQGSNLRLGMNSTMYEKSIVPVHWPGGGSLSVSMETLDA
LYDLHKLCRNWWWTSTNQDLPLVRYKGCKTLYQSTFTDYIVRHLTLLPANSNKLTYPNTLHPLMMMMSKYKHHPSRQTR
RKKKPYTKLLVKPPPQLLENKWYLATDLYKPLLQLHCLACNLQNPLVKPDKLSNNVTLWSLNTLSIQNRNMSVDQGQSW
PLKLGTQSLYLYLYLGANLPGDTLQIPVADLLPLTNPRINRPGQSLNLAKLTDLHTLTLYKNKLTNYWGNPLNKLHQLHL
DMDLYSLKSPLAIKNLWTLTLNMKWNQLNNAGLMALTPLNLPLTQIQYNPDRDTGLDLQLYLLSNATGLGWDPPGIP
LHLLGLPLWLLYWGLADLQKNLKKVLNIDLTNYMLVAKLKLTQKPGTLYLVLLNDTLVLGNSPYLKQPLPLDNIKWYPQV
QYQLLAQNKLLQTGPLTPNLQGQLSDNISMLYKLYLKWGGSPPKAINVLNPALHQIQYPIPRNLLLETTSLQSPGLAPLSLL
YSLDYRLIGNYTLTALSRISQDWALKDTVSKTLPPDRQQLLKQALLCLQSLFTQLKKLKLVQQLISNLRQQQQLYRLRLISL
LKDQ

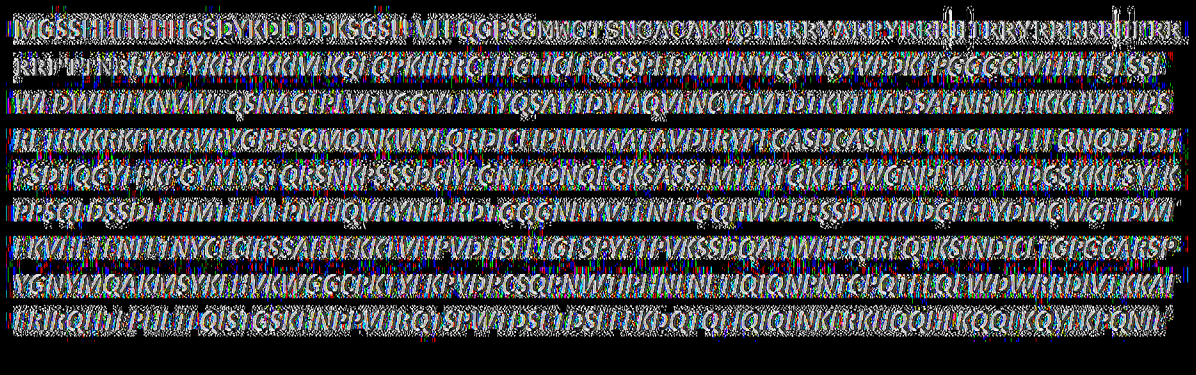
SEQ ID NO: 565：

在ARG腎及第一β股中具有LYS/HIS突變之環9



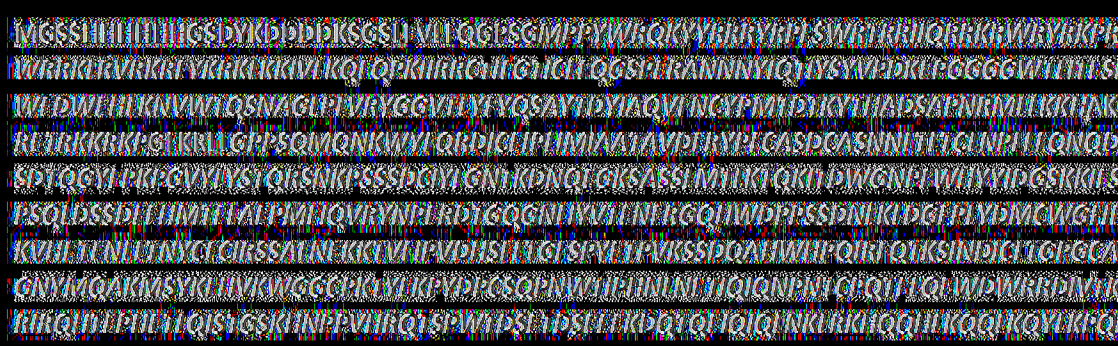
SEQ ID NO: 566：

具有BFDV之ARG腎的環9：



SEQ ID NO: 567：

具有BFDV蛋白殼蛋白之β股F及G的環9：



SEQ ID NO: 568 ：

具有環9之β C的環2：

MGSSHHHHHHGSDYKDDDDKSGSLVLQGPSGMPYYYRRRYNYRRPRWYGRGWIRRPLRRRLRRKRRVRPTYT
TPLKQWQPPYKRHGILGTGIIQGSNLRLGMNSTMYLKSIVPVLWPGGGSLSVSMLLTDALYDLLKLCRNWWLSLN
QDLPLVRYKGCKLTYQSTLFDYIVRLLTLLPANSNKLTYPNTLPLMMMMSKYKLIIPSRQTRRKKKPYTKLVKPPPQL
NKWYLATDLYKPLLQLHGTACNLQNPLVKPDKLSNNVTLWSLNTLSIQNRNMSVDQGQSWPLKLGTQSLYLYLYTGA
NLPGDTTQLPVADLLPLTNPRINRPGQSLNLAKTDLHTLFTLYKNKLTNYWGNPLNKLHQLFLDMLLYSLKSPLAIKNLWL
TLNMKWNQLNNAGTMALTPLNLPLTQLQYNPDRDTGLDTQLYLLSNATGTGWDPPGIPLLLGLPLWLIYWGLAD
LQKNLKKVTNLDTNYMLVAKTKLTQKPGTLYLVLLNDTLVLGNSPYLKQPLPLDNLKWYPQVQYQLLAQNKLLQTGPFT
PNLQGQLSDNLSMLYKLYLKWGGSPPKAINVLNPALQLQYPLPRNLLLLTLSLQSPGLAPLSLLYSLDYRLHGNYTLTALSRL
SQDWALKDLVSKLLPDRQQLLKQALLCLQSLLLQLKKLKLKLQQLLKQQLKQYLKLQNLL

SEQ ID NO: 569 ：

具有環9之連接子1的環2：

MGSSHHHHHHGSDYKDDDDKSGSLVLQGPSGMPYYYRRRYNYRRPRWYGRGWIRRPLRRRLRRKRRVRPTYT
TPLKQWQPPYKRTCYIKGQDCLIYYSPLRANNWYQLIYSYVVPDKLPGGGGWILLLTLSLSSLWLDWLLNLKNVWLQSN
AGLPLVRYKGCKLTYQSTLFDYIVRLLTLLPANSNKLTYPNTLPLMMMMSKYKLIIPSRQTRRKKKPYTKLVKPPPQL
NKWYLATDLYKPLLQLHGTACNLQNPLVKPDKLSNNVTLWSLNTLSIQNRNMSVDQGQSWPLKLGTQSLYLYLYTGA
NLPGDTTQLPVADLLPLTNPRINRPGQSLNLAKTDLHTLFTLYKNKLTNYWGNPLNKLHQLFLDMLLYSLKSPLAIKNLWL
TLNMKWNQLNNAGTMALTPLNLPLTQLQYNPDRDTGLDTQLYLLSNATGTGWDPPGIPLLLGLPLWLIYWGLAD
LQKNLKKVTNLDTNYMLVAKTKLTQKPGTLYLVLLNDTLVLGNSPYLKQPLPLDNLKWYPQVQYQLLAQNKLLQTGPFT
PNLQGQLSDNLSMLYKLYLKWGGSPPKAINVLNPALQLQYPLPRNLLLLTLSLQSPGLAPLSLLYSLDYRLHGNYTLTALSRL
SQDWALKDLVSKLLPDRQQLLKQALLCLQSLLLQLKKLKLKLQQLLKQQLKQYLKLQNLL

SEQ ID NO: 570 ：

具有環9之β股D的環2：

MGSSHHHHHHGSDYKDDDDKSGSLVLQGPSGMPYYYRRRYNYRRPRWYGRGWIRRPLRRRLRRKRRVRPTYT
TPLKQWQPPYKRTCYIKGQDCLIYYSNLRLGMNSTMYLKSIVPVLWPGGGSLSVSMLLTDALYDLLKLCRNWWLSLN
QDLPLVRYGGVILYLYQSTLFDYIVRLLTLLPANSNKLTYPNTLPLMMMMSKYKLIIPSRQTRRKKKPYTKLVKPPPQL
NKWYLATDLYKPLLQLHGTACNLQNPLVKPDKLSNNVTLWSLNTLSIQNRNMSVDQGQSWPLKLGTQSLYLYLYTGA
NLPGDTTQLPVADLLPLTNPRINRPGQSLNLAKTDLHTLFTLYKNKLTNYWGNPLNKLHQLFLDMLLYSLKSPLAIKNL
WLTLNMKWNQLNNAGTMALTPLNLPLTQLQYNPDRDTGLDTQLYLLSNATGTGWDPPGIPLLLGLPLWLIYWGLA
DLQKNLKKVTNLDTNYMLVAKTKLTQKPGTLYLVLLNDTLVLGNSPYLKQPLPLDNLKWYPQVQYQLLAQNKLLQTG
PFTPNLQGQLSDNLSMLYKLYLKWGGSPPKAINVLNPALQLQYPLPRNLLLLTLSLQSPGLAPLSLLYSLDYRLHGNYTLTALA
SRLSQDWALKDLVSKLLPDRQQLLKQALLCLQSLLLQLKKLKLKLQQLLKQQLKQYLKLQNLL

SEQ ID NO: 571 ：

具有環9之連接子2的環2：

MGSSHHHHHHGSDYKDDDDKSGSLVLFQGPSGMPYYYRRRYNYRRPRWYGRGWIRRPFRRRFRRKRRVRPTYT
TPFKQWQPPYKRTCYIKGQDCFIYYSNFRIGMNSTMYFKSIVPVFIWPGGGSFSVSMFTFDAFYDFHKFCRNWWFSFN
QDFPFVRYKGCKFTYQSFFFDYIVRFFFFPANSNKFIYPNTFPFMMMMSKYKHFIPSRQTRRKKKPYFKFVKPPPQF
FNKWYFATDFYKIPFFQFHCFACNFQNPFVKPDKFSNNVFTWSFNTFSIQNRNMSVDQGQSWPFKFIGTQSFYFYFYTG
ANFPGDTFQIPVADFFPFTNPRINRPGQSFNFAKTDHFTFFYKNKFTNYWGNPFNKFHQFHFDMFFYSFKSPFAIKNF
WTFFNMKWNQFNNAGTMAFTPFNFPFFTQIQYNPDRDTGFDTQFYFFSNATGTGWDPPGIPFFFFGFPFWFFYWGF
ADFQKNFKKVFNIDTNYMFVAKTKFTQKPGTFYFVFFNDTFVFGNSPYFKQPFPFDNIKWYPQVQYQFFAQNKFFQTG
PFTPNIQGQFSDNISMFYKFYFKWGGSPPKAINVFNPAFHQIQYPIPRNFFFTSFQSPGFAPFSFFYSFDYRFHGNYTTTAF
SRFSQDWAFKDTVSKTFPDRQQFFKQAFFCFQSFFTQFKKFKFVQQFISNFRQQQQQFYRFRFISFFKDQ

SEQ ID NO: 572 ：

具有環9之β股G DNA結合的環2：

MGSSHHHHHHGSDYKDDDDKSGSLVLFQGPSGMPYYYRRRHYNYRRPRWYGRGWIRRPFRRRFRRKRRVRPTYT
TPFKQWQPPYKRTCYIKGQDCFIYYSNFRIGMNSTMYFKSIVPVFIWPGGGSFSVSMFTFDAFYDFHKFCRNWWFSFN
QDFPFVRYKGCKFTYQSFFFDYIVRFFFFPANSNKFIYPNTFPFMMMMSKYKHFIPSRQTRRKKKPYKFVMVKPPPQF
FNKWYFATDFYKIPFFQFHCFACNFQNPFVKPDKFSNNVFTWSFNTFSIQNRNMSVDQGQSWPFKFIGTQSFYFYFYTG
ANFPGDTFQIPVADFFPFTNPRINRPGQSFNFAKTDHFTFFYKNKFTNYWGNPFNKFHQFHFDMFFYSFKSPFAIKNF
WTFFNMKWNQFNNAGTMAFTPFNFPFFTQIQYNPDRDTGFDTQFYFFSNATGTGWDPPGIPFFFFGFPFWFFYWGF
ADFQKNFKKVFNIDTNYMFVAKTKFTQKPGTFYFVFFNDTFVFGNSPYFKQPFPFDNIKWYPQVQYQFFAQNKFFQTG
PFTPNIQGQFSDNISMFYKFYFKWGGSPPKAINVFNPAFHQIQYPIPRNFFFTSFQSPGFAPFSFFYSFDYRFHGNYTTTAF
SRFSQDWAFKDTVSKTFPDRQQFFKQAFFCFQSFFTQFKKFKFVQQFISNFRQQQQQFYRFRFISFFKDQ

SEQ ID NO: 573 ：

具有環9之β股F內原紫黴接觸的環2：

MGSSHHHHHHGSDYKDDDDKSGSLVLFQGPSGMPYYYRRRYNYRRPRWYGRGWIRRPFRRRFRRKRRVRPTYT
TPFKQWQPPYKRTCYIKGQDCFIYYSNFRIGMNSTMYFKSIVPVFIWPGGGSFSVSMFTFDAFYDFHKFCRNWWFSFN
QDFPFVRYKGCKFTYQSFFFDYIVRFFFFPANSNKFIYPNTFPFMMMMSKYKHFIPSRQTRRKKKPYKFVKPPPQF
FNKWYFATDFYKIPFFQFHCFACNFQNPFVKPDKFSNNVFTWSFNTFSIQNRNMSVDQGQSWPFKFIGTQSFYFYFYTG
ANFPGDTFQIPVADFFPFTNPRINRPGQSFNFAKTDHFTFFYKNKFTNYWGNPFNKFHQFHFDMFFYSFKSPFAIKNF
WTFFNMKWNQFNNAGTMAFTPFNFPFFTQIQYNPDRDTGFDTQFYFFSNATGTGWDPPGIPFFFFGFPFWFFYWGF
ADFQKNFKKVFNIDTNYMFVAKTKFTQKPGTFYFVFFNDTFVFGNSPYFKQPFPFDNIKWYPQVQYQFFAQNKFFQTG
PFTPNIQGQFSDNISMFYKFYFKWGGSPPKAINVFNPAFHQIQYPIPRNFFFTSFQSPGFAPFSFFYSFDYRFHGNYTTTAF
SRFSQDWAFKDTVSKTFPDRQQFFKQAFFCFQSFFTQFKKFKFVQQFISNFRQQQQQFYRFRFISFFKDQ

SEQ ID NO: 574 ：

具有環9之股H及I以及C端片段的環2 ORF1：

MGSSHHHHHHGSDYKDDDDKSGSLVLFQGPSGMPYYYRRRYNYRRPRWYGRGWIRRPFRRRFRRKRRVRPTYT
TPFKQWQPPYKRTCYIKGQDCFIYYSNFRIGMNSTMYFKSIVPVFIWPGGGSFSVSMFTFDAFYDFHKFCRNWWFSFN
QDFPFVRYKGCKFTYQSFFFDYIVRFFFFPANSNKFIYPNTFPFMMMMSKYKHFIPSRQTRRKKKPYKFVKPPPQF
FNKWYFATDFYKIPFFQFHCFACNFQNPFVKPDKFSNNVFTWSFNTFSIQNRNMSVDQGQSWPFKFIGTQSFYFYFYTG
ANFPGDTFQIPVADFFPFTNPRINRPGQSFNFAKTDHFTFFYKNKFTNYWGNPFNKFHQFHFDMFFYSFKSPFAIKNF
WTFFNMKWNQFNNAGTMAFTPFNFPFFTQIQYNPDRDTGFDTQFYFFSNATGTGWDPPGIPFFFFGFPFWFFYWGF
ADFQKNFKKVFNIDTNYMFVAKTKFTQKPGTFYFVFFNDTFVFGNSPYFKQPFPFDNIKWYPQVQYQFFAQNKFFQTG
PFTPNIQGQFSDNISMFYKFYFKWGGSPPKAINVFNPAFHQIQYPIPRNFFFTSFQSPGFAPFSFFYSFDYRFHGNYTTTAF
SRFSQDWAFKDTVSKTFPDRQQFFKQAFFCFQSFFTQFKKFKFVQQFISNFRQQQQQFYRFRFISFFKDQ

SEQ ID NO: 575 ：

具有環9之股I及C端片段的環2 ORF1 ：

MGSSHHHHHHGSDYKDDDDKSGSLVPRGSMPYYYRRRYNYRRPRWYGRGWIRRPFRRRFRRKRRVRPTYT
TPIKQWQPPYKRTCYIKGQDCIIYYSNIRIGMNSTMYEKSIVPVFWPGGGSFSVSMFTTDAIYDFHKICRNWWTSTN
QDIPFVRYKGCKITEYQSTFTDYTVRHTTTPANSNKTYPNTHPTMMMMSKYKHHPSRQTRRKKKCPYTKFVKPPQDF
NKWYFATDIYKPIIQHICTACNFQNPFVKPDKISNWTETTGINPFIFQNQDEDHPSDTQGVFHKPGVWTYSTQRSNKP
SSSDGIYTGNTKDNQKGKSASSFMTTKTQKTTDWGNHFWHTYYIDGSKKFSYTKPPSQIDSSDFTHMTTATFPMFIQV
RYNPFRDTGQGNIIYWTFNFEGQFHWDPSSDNFKIDGFPLYDMCWGFIDWFKVFFTTNFITNYCTGRSSALNFKK
TMTPVDFSTTGFSPVTTPVKSSDQAFTMTFQHFQTKSNDICTGPGCARSPYGNYMQAKMSYKTTVKTWGGCPK
TYTKPYDFCSQFNWTHFTHWNFTTQIQNFNTCPQITQFWDWRRDIVTKKAFRRQFTTFFFTTQISTGSKTNFPV
FHQTSFWTDSFTDSFTFKDQTQFIQIQINKTRKHQQFHKQQIKQYTKPQNFF

【序列表】

<110> 美商旗艦先鋒創新公司 (FLAGSHIP PIONEERING INNOVATIONS V, INC.)

<120> 包裹RNA之指環病毒(ANELLOVIRUS)蛋白殼之活體外組裝

<130> V2057-7017TW

<140> TW 110148173
<141> 2021-12-22

<150> 63/147,064
<151> 2021-02-08

<150> 63/130,360
<151> 2020-12-23

<160> 968

<170> PatentIn version 3.5

<210> 1
<211> 699
<212> PRT
<213> 人工序列

<220>
<223> 人工序列描述：合成多肽

<400> 1

```
Met Gly Ser Ser His His His His His His Gly Ser Asp Tyr Lys Asp
1               5                  10                  15


Asp Asp Asp Lys Ser Gly Ser Leu Glu Val Leu Phe Gln Gly Pro Ser
            20                  25                  30


Gly Met Pro Tyr Tyr Tyr Arg Arg Arg Arg Tyr Asn Tyr Arg Arg Pro
        35                  40                  45


Arg Trp Tyr Gly Arg Gly Trp Ile Arg Arg Pro Phe Arg Arg Arg Phe
    50                  55                  60


Arg Arg Lys Arg Arg Val Arg Pro Thr Tyr Thr Thr Ile Pro Leu Lys
65                  70                  75                  80
```

Gln Trp Gln Pro Pro Tyr Lys Arg Thr Cys Tyr Ile Lys Gly Gln Asp
85 90 95

Cys Leu Ile Tyr Tyr Ser Asn Leu Arg Leu Gly Met Asn Ser Thr Met
100 105 110

Tyr Glu Lys Ser Ile Val Pro Val His Trp Pro Gly Gly Gly Ser Phe
115 120 125

Ser Val Ser Met Leu Thr Leu Asp Ala Leu Tyr Asp Ile His Lys Leu
130 135 140

Cys Arg Asn Trp Trp Thr Ser Thr Asn Gln Asp Leu Pro Leu Val Arg
145 150 155 160

Tyr Lys Gly Cys Lys Ile Thr Phe Tyr Gln Ser Thr Phe Thr Asp Tyr
165 170 175

Ile Val Arg Ile His Thr Glu Leu Pro Ala Asn Ser Asn Lys Leu Thr
180 185 190

Tyr Pro Asn Thr His Pro Leu Met Met Met Met Ser Lys Tyr Lys His
195 200 205

Ile Ile Pro Ser Arg Gln Thr Arg Arg Lys Lys Lys Pro Tyr Thr Lys
210 215 220

Ile Phe Val Lys Pro Pro Pro Gln Phe Glu Asn Lys Trp Tyr Phe Ala
225 230 235 240

Thr Asp Leu Tyr Lys Ile Pro Leu Leu Gln Ile His Cys Thr Ala Cys
245 250 255

Asn Leu Gln Asn Pro Phe Val Lys Pro Asp Lys Leu Ser Asn Asn Val
260 265 270

Thr Leu Trp Ser Leu Asn Thr Ile Ser Ile Gln Asn Arg Asn Met Ser

275 280 285

Val Asp Gln Gly Gln Ser Trp Pro Phe Lys Ile Leu Gly Thr Gln Ser
290 295 300

Phe Tyr Phe Tyr Phe Tyr Thr Gly Ala Asn Leu Pro Gly Asp Thr Thr
305 310 315 320

Gln Ile Pro Val Ala Asp Leu Leu Pro Leu Thr Asn Pro Arg Ile Asn
325 330 335

Arg Pro Gly Gln Ser Leu Asn Glu Ala Lys Ile Thr Asp His Ile Thr
340 345 350

Phe Thr Glu Tyr Lys Asn Lys Phe Thr Asn Tyr Trp Gly Asn Pro Phe
355 360 365

Asn Lys His Ile Gln Glu His Leu Asp Met Ile Leu Tyr Ser Leu Lys
370 375 380

Ser Pro Glu Ala Ile Lys Asn Glu Trp Thr Thr Glu Asn Met Lys Trp
385 390 395 400

Asn Gln Leu Asn Asn Ala Gly Thr Met Ala Leu Thr Pro Phe Asn Glu
405 410 415

Pro Ile Phe Thr Gln Ile Gln Tyr Asn Pro Asp Arg Asp Thr Gly Glu
420 425 430

Asp Thr Gln Leu Tyr Leu Leu Ser Asn Ala Thr Gly Thr Gly Trp Asp
435 440 445

Pro Pro Gly Ile Pro Glu Leu Ile Leu Glu Gly Phe Pro Leu Trp Leu
450 455 460

Ile Tyr Trp Gly Phe Ala Asp Phe Gln Lys Asn Leu Lys Lys Val Thr
465 470 475 480

Asn Ile Asp Thr Asn Tyr Met Leu Val Ala Lys Thr Lys Phe Thr Gln
485 490 495

Lys Pro Gly Thr Phe Tyr Leu Val Ile Leu Asn Asp Thr Phe Val Glu
500 505 510

Gly Asn Ser Pro Tyr Glu Lys Gln Pro Leu Pro Glu Asp Asn Ile Lys
515 520 525

Trp Tyr Pro Gln Val Gln Tyr Gln Leu Glu Ala Gln Asn Lys Leu Leu
530 535 540

Gln Thr Gly Pro Phe Thr Pro Asn Ile Gln Gly Gln Leu Ser Asp Asn
545 550 555 560

Ile Ser Met Phe Tyr Lys Phe Tyr Phe Lys Trp Gly Gly Ser Pro Pro
565 570 575

Lys Ala Ile Asn Val Glu Asn Pro Ala His Gln Ile Gln Tyr Pro Ile
580 585 590

Pro Arg Asn Glu His Glu Thr Thr Ser Leu Gln Ser Pro Gly Glu Ala
595 600 605

Pro Glu Ser Ile Leu Tyr Ser Phe Asp Tyr Arg His Gly Asn Tyr Thr
610 615 620

Thr Thr Ala Leu Ser Arg Ile Ser Gln Asp Trp Ala Leu Lys Asp Thr
625 630 635 640

Val Ser Lys Ile Thr Glu Pro Asp Arg Gln Gln Leu Leu Lys Gln Ala
645 650 655

Leu Glu Cys Leu Gln Ile Ser Glu Glu Thr Gln Glu Lys Lys Glu Lys
660 665 670

Glu Val Gln Gln Leu Ile Ser Asn Leu Arg Gln Gln Gln Gln Leu Tyr

675 680 685

Arg Glu Arg Ile Ile Ser Leu Leu Lys Asp Gln
690 695

<210> 2
<211> 688
<212> PRT
<213> 人工序列

<220>
<223> 人工序列描述：合成多肽

<400> 2
Met Gly Ser Ser His His His His His His Gly Ser Asp Tyr Lys Asp
1 5 10 15

Asp Asp Asp Lys Ser Gly Ser Leu Glu Val Leu Phe Gln Gly Pro Ser
20 25 30

Gly Met Pro Pro Tyr Trp Arg Gln Lys Tyr Tyr Arg Arg Arg Tyr Arg
35 40 45

Pro Phe Ser Trp Arg Thr Arg Arg Ile Ile Gln Arg Arg Lys Arg Trp
50 55 60

Arg Tyr Arg Lys Pro Arg Lys Thr Tyr Trp Arg Arg Lys Leu Arg Val
65 70 75 80

Arg Lys Arg Phe Tyr Lys Arg Lys Leu Lys Lys Ile Val Leu Lys Gln
85 90 95

Phe Gln Pro Lys Ile Ile Arg Arg Cys Thr Ile Phe Gly Thr Ile Cys
100 105 110

Leu Phe Gln Gly Ser Pro Glu Arg Ala Asn Asn Asn Tyr Ile Gln Thr
115 120 125

Ile Tyr Ser Tyr Val Pro Asp Lys Glu Pro Gly Gly Gly Gly Trp Thr
130 135 140

Leu Ile Thr Glu Ser Leu Ser Ser Leu Trp Glu Asp Trp Glu His Leu
145 150 155 160

Lys Asn Val Trp Thr Gln Ser Asn Ala Gly Leu Pro Leu Val Arg Tyr
165 170 175

Gly Gly Val Thr Leu Tyr Phe Tyr Gln Ser Ala Tyr Thr Asp Tyr Ile
180 185 190

Ala Gln Val Phe Asn Cys Tyr Pro Met Thr Asp Thr Lys Tyr Thr His
195 200 205

Ala Asp Ser Ala Pro Asn Arg Met Leu Leu Lys Lys His Val Ile Arg
210 215 220

Val Pro Ser Arg Glu Thr Arg Lys Lys Arg Lys Pro Tyr Lys Arg Val
225 230 235 240

Arg Val Gly Pro Pro Ser Gln Met Gln Asn Lys Trp Tyr Phe Gln Arg
245 250 255

Asp Ile Cys Glu Ile Pro Leu Ile Met Ile Ala Ala Thr Ala Val Asp
260 265 270

Phe Arg Tyr Pro Phe Cys Ala Ser Asp Cys Ala Ser Asn Asn Leu Thr
275 280 285

Leu Thr Cys Leu Asn Pro Leu Leu Phe Gln Asn Gln Asp Phe Asp His
290 295 300

Pro Ser Asp Thr Gln Gly Tyr Phe Pro Lys Pro Gly Val Tyr Leu Tyr
305 310 315 320

Ser Thr Gln Arg Ser Asn Lys Pro Ser Ser Ser Asp Cys Ile Tyr Leu
325 330 335

Gly Asn Thr Lys Asp Asn Gln Glu Gly Lys Ser Ala Ser Ser Leu Met
340 345 350

Thr Leu Lys Thr Gln Lys Ile Thr Asp Trp Gly Asn Pro Phe Trp His
355 360 365

Tyr Tyr Ile Asp Gly Ser Lys Lys Ile Phe Ser Tyr Phe Lys Pro Pro
370 375 380

Ser Gln Leu Asp Ser Ser Asp Phe Glu His Met Thr Glu Leu Ala Glu
385 390 395 400

Pro Met Phe Ile Gln Val Arg Tyr Asn Pro Glu Arg Asp Thr Gly Gln
405 410 415

Gly Asn Leu Ile Tyr Val Thr Glu Asn Phe Arg Gly Gln His Trp Asp
420 425 430

Pro Pro Ser Ser Asp Asn Leu Lys Leu Asp Gly Phe Pro Leu Tyr Asp
435 440 445

Met Cys Trp Gly Phe Ile Asp Trp Ile Glu Lys Val His Glu Thr Glu
450 455 460

Asn Leu Leu Thr Asn Tyr Cys Phe Cys Ile Arg Ser Ser Ala Phe Asn
465 470 475 480

Glu Lys Lys Thr Val Phe Ile Pro Val Asp His Ser Phe Leu Thr Gly
485 490 495

Phe Ser Pro Tyr Glu Thr Pro Val Lys Ser Ser Asp Gln Ala His Trp
500 505 510

His Pro Gln Ile Arg Phe Gln Thr Lys Ser Ile Asn Asp Ile Cys Leu
515 520 525

Thr Gly Pro Gly Cys Ala Arg Ser Pro Tyr Gly Asn Tyr Met Gln Ala
530 535 540

Lys Met Ser Tyr Lys Phe His Val Lys Trp Gly Gly Cys Pro Lys Thr
545 550 555 560

Tyr Glu Lys Pro Tyr Asp Pro Cys Ser Gln Pro Asn Trp Thr Ile Pro
565 570 575

His Asn Leu Asn Glu Thr Ile Gln Ile Gln Asn Pro Asn Thr Cys Pro
580 585 590

Gln Thr Glu Leu Gln Glu Trp Asp Trp Arg Arg Asp Ile Val Thr Lys
595 600 605

Lys Ala Ile Glu Arg Ile Arg Gln His Thr Glu Pro His Glu Thr Leu
610 615 620

Gln Ile Ser Thr Gly Ser Lys His Asn Pro Pro Val His Arg Gln Thr
625 630 635 640

Ser Pro Trp Thr Asp Ser Glu Thr Asp Ser Glu Glu Glu Lys Asp Gln
645 650 655

Thr Gln Glu Ile Gln Ile Gln Leu Asn Lys Leu Arg Lys His Gln Gln
660 665 670

His Leu Lys Gln Gln Leu Lys Gln Tyr Leu Lys Pro Gln Asn Ile Glu
675 680 685

<210> 3
<211> 714
<212> PRT
<213> 人工序列

<220>
<223> 人工序列描述：合成多肽

<400> 3
Met Gly Ser Ser His His His His His His Gly Ser Asp Tyr Lys Asp
1 5 10 15

```
Asp Asp Asp Lys Ser Gly Ser Leu Glu Val Leu Phe Gln Gly Pro Ser
            20                  25                  30

Gly Met Pro Pro Tyr Trp Arg Gln Lys Tyr Tyr Arg Arg Arg Tyr Arg
        35                  40                  45

Pro Phe Ser Trp Arg Thr Arg Arg Ile Ile Gln Arg Arg Lys Arg Trp
    50                  55                  60

Arg Tyr Arg Lys Pro Arg Lys Thr Tyr Trp Arg Arg Lys Leu Arg Val
65                  70                  75                  80

Arg Lys Arg Arg Val Arg Pro Thr Tyr Thr Thr Ile Pro Leu Lys Gln
                85                  90                  95

Trp Gln Pro Pro Tyr Lys Arg Thr Cys Tyr Ile Lys Gly Gln Asp Cys
            100                 105                 110

Leu Ile Tyr Tyr Ser Asn Leu Arg Leu Gly Met Asn Ser Thr Met Tyr
        115                 120                 125

Glu Lys Ser Ile Val Pro Val His Trp Pro Gly Gly Gly Ser Phe Ser
    130                 135                 140

Val Ser Met Leu Thr Leu Asp Ala Leu Tyr Asp Ile His Lys Leu Cys
145                 150                 155                 160

Arg Asn Trp Trp Thr Ser Thr Asn Gln Asp Leu Pro Leu Val Arg Tyr
                165                 170                 175

Lys Gly Cys Lys Ile Thr Phe Tyr Gln Ser Thr Phe Thr Asp Tyr Ile
            180                 185                 190

Val Arg Ile His Thr Glu Leu Pro Ala Asn Ser Asn Lys Leu Thr Tyr
        195                 200                 205

Pro Asn Thr His Pro Leu Met Met Met Met Ser Lys Tyr Lys His Ile
```

210 215 220

Ile Pro Ser Arg Gln Thr Arg Arg Lys Lys Lys Pro Tyr Thr Lys Ile
225 230 235 240

Phe Val Lys Pro Pro Pro Gln Phe Glu Asn Lys Trp Tyr Phe Ala Thr
245 250 255

Asp Leu Tyr Lys Ile Pro Leu Leu Gln Ile His Cys Thr Ala Cys Asn
260 265 270

Leu Gln Asn Pro Phe Val Lys Pro Asp Lys Leu Ser Asn Asn Val Thr
275 280 285

Leu Trp Ser Leu Asn Thr Ile Ser Ile Gln Asn Arg Asn Met Ser Val
290 295 300

Asp Gln Gly Gln Ser Trp Pro Phe Lys Ile Leu Gly Thr Gln Ser Phe
305 310 315 320

Tyr Phe Tyr Phe Tyr Thr Gly Ala Asn Leu Pro Gly Asp Thr Thr Gln
325 330 335

Ile Pro Val Ala Asp Leu Leu Pro Leu Thr Asn Pro Arg Ile Asn Arg
340 345 350

Pro Gly Gln Ser Leu Asn Glu Ala Lys Ile Thr Asp His Ile Thr Phe
355 360 365

Thr Glu Tyr Lys Asn Lys Phe Thr Asn Tyr Trp Gly Asn Pro Phe Asn
370 375 380

Lys His Ile Gln Glu His Leu Asp Met Ile Leu Tyr Ser Leu Lys Ser
385 390 395 400

Pro Glu Ala Ile Lys Asn Glu Trp Thr Thr Glu Asn Met Lys Trp Asn
405 410 415

Gln Leu Asn Asn Ala Gly Thr Met Ala Leu Thr Pro Phe Asn Glu Pro
420 425 430

Ile Phe Thr Gln Ile Gln Tyr Asn Pro Asp Arg Asp Thr Gly Glu Asp
435 440 445

Thr Gln Leu Tyr Leu Leu Ser Asn Ala Thr Gly Thr Gly Trp Asp Pro
450 455 460

Pro Gly Ile Pro Glu Leu Ile Leu Glu Gly Phe Pro Leu Trp Leu Ile
465 470 475 480

Tyr Trp Gly Phe Ala Asp Phe Gln Lys Asn Leu Lys Lys Val Thr Asn
485 490 495

Ile Asp Thr Asn Tyr Met Leu Val Ala Lys Thr Lys Phe Thr Gln Lys
500 505 510

Pro Gly Thr Phe Tyr Leu Val Ile Leu Asn Asp Thr Phe Val Glu Gly
515 520 525

Asn Ser Pro Tyr Glu Lys Gln Pro Leu Pro Glu Asp Asn Ile Lys Trp
530 535 540

Tyr Pro Gln Val Gln Tyr Gln Leu Glu Ala Gln Asn Lys Leu Leu Gln
545 550 555 560

Thr Gly Pro Phe Thr Pro Asn Ile Gln Gly Gln Leu Ser Asp Asn Ile
565 570 575

Ser Met Phe Tyr Lys Phe Tyr Phe Lys Trp Gly Gly Ser Pro Pro Lys
580 585 590

Ala Ile Asn Val Glu Asn Pro Ala His Gln Ile Gln Tyr Pro Ile Pro
595 600 605

Arg Asn Glu His Glu Thr Thr Ser Leu Gln Ser Pro Gly Glu Ala Pro

610 615 620

Glu Ser Ile Leu Tyr Ser Phe Asp Tyr Arg His Gly Asn Tyr Thr Thr
625 630 635 640

Thr Ala Leu Ser Arg Ile Ser Gln Asp Trp Ala Leu Lys Asp Thr Val
645 650 655

Ser Lys Ile Thr Glu Pro Asp Arg Gln Gln Leu Leu Lys Gln Ala Leu
660 665 670

Glu Cys Leu Gln Ile Ser Glu Glu Thr Gln Glu Lys Lys Glu Lys Glu
675 680 685

Val Gln Gln Leu Ile Ser Asn Leu Arg Gln Gln Gln Gln Leu Tyr Arg
690 695 700

Glu Arg Ile Ile Ser Leu Leu Lys Asp Gln
705 710

<210> 4
<211> 714
<212> PRT
<213> 人工序列

<220>
<223> 人工序列描述：合成多肽

<400> 4
Met Gly Ser Ser His His His His His His Gly Ser Asp Tyr Lys Asp
1 5 10 15

Asp Asp Asp Lys Ser Gly Ser Leu Glu Val Leu Phe Gln Gly Pro Ser
20 25 30

Gly Met Pro Pro Tyr Trp Arg Gln Lys Tyr Tyr Arg Arg Arg Tyr Arg
35 40 45

Pro Phe Ser Trp Arg Thr Arg Arg Ile Ile Gln Arg Arg Lys Arg Trp
50 55 60

Arg Tyr Arg Lys Pro Arg Lys Thr Tyr Trp Arg Arg Lys Leu Arg Val
65 70 75 80

Arg Lys Arg Phe Tyr Lys Arg Lys Leu Lys Lys Ile Val Leu Lys Gln
85 90 95

Phe Gln Pro Lys Ile Ile Arg Arg Cys Thr Ile Phe Gly Thr Ile Cys
100 105 110

Leu Phe Gln Gly Ser Asn Leu Arg Leu Gly Met Asn Ser Thr Met Tyr
115 120 125

Glu Lys Ser Ile Val Pro Val His Trp Pro Gly Gly Gly Ser Phe Ser
130 135 140

Val Ser Met Leu Thr Leu Asp Ala Leu Tyr Asp Ile His Lys Leu Cys
145 150 155 160

Arg Asn Trp Trp Thr Ser Thr Asn Gln Asp Leu Pro Leu Val Arg Tyr
165 170 175

Lys Gly Cys Lys Ile Thr Phe Tyr Gln Ser Thr Phe Thr Asp Tyr Ile
180 185 190

Val Arg Ile His Thr Glu Leu Pro Ala Asn Ser Asn Lys Leu Thr Tyr
195 200 205

Pro Asn Thr His Pro Leu Met Met Met Met Ser Lys Tyr Lys His Ile
210 215 220

Ile Pro Ser Arg Gln Thr Arg Arg Lys Lys Lys Pro Tyr Thr Lys Ile
225 230 235 240

Phe Val Lys Pro Pro Pro Gln Phe Glu Asn Lys Trp Tyr Phe Ala Thr
245 250 255

Asp Leu Tyr Lys Ile Pro Leu Leu Gln Ile His Cys Thr Ala Cys Asn
260 265 270

Leu Gln Asn Pro Phe Val Lys Pro Asp Lys Leu Ser Asn Asn Val Thr
275 280 285

Leu Trp Ser Leu Asn Thr Ile Ser Ile Gln Asn Arg Asn Met Ser Val
290 295 300

Asp Gln Gly Gln Ser Trp Pro Phe Lys Ile Leu Gly Thr Gln Ser Phe
305 310 315 320

Tyr Phe Tyr Phe Tyr Thr Gly Ala Asn Leu Pro Gly Asp Thr Thr Gln
325 330 335

Ile Pro Val Ala Asp Leu Leu Pro Leu Thr Asn Pro Arg Ile Asn Arg
340 345 350

Pro Gly Gln Ser Leu Asn Glu Ala Lys Ile Thr Asp His Ile Thr Phe
355 360 365

Thr Glu Tyr Lys Asn Lys Phe Thr Asn Tyr Trp Gly Asn Pro Phe Asn
370 375 380

Lys His Ile Gln Glu His Leu Asp Met Ile Leu Tyr Ser Leu Lys Ser
385 390 395 400

Pro Glu Ala Ile Lys Asn Glu Trp Thr Thr Glu Asn Met Lys Trp Asn
405 410 415

Gln Leu Asn Asn Ala Gly Thr Met Ala Leu Thr Pro Phe Asn Glu Pro
420 425 430

Ile Phe Thr Gln Ile Gln Tyr Asn Pro Asp Arg Asp Thr Gly Glu Asp
435 440 445

Thr Gln Leu Tyr Leu Leu Ser Asn Ala Thr Gly Thr Gly Trp Asp Pro
450 455 460

Pro Gly Ile Pro Glu Leu Ile Leu Glu Gly Phe Pro Leu Trp Leu Ile
465 470 475 480

Tyr Trp Gly Phe Ala Asp Phe Gln Lys Asn Leu Lys Lys Val Thr Asn
485 490 495

Ile Asp Thr Asn Tyr Met Leu Val Ala Lys Thr Lys Phe Thr Gln Lys
500 505 510

Pro Gly Thr Phe Tyr Leu Val Ile Leu Asn Asp Thr Phe Val Glu Gly
515 520 525

Asn Ser Pro Tyr Glu Lys Gln Pro Leu Pro Glu Asp Asn Ile Lys Trp
530 535 540

Tyr Pro Gln Val Gln Tyr Gln Leu Glu Ala Gln Asn Lys Leu Leu Gln
545 550 555 560

Thr Gly Pro Phe Thr Pro Asn Ile Gln Gly Gln Leu Ser Asp Asn Ile
565 570 575

Ser Met Phe Tyr Lys Phe Tyr Phe Lys Trp Gly Gly Ser Pro Pro Lys
580 585 590

Ala Ile Asn Val Glu Asn Pro Ala His Gln Ile Gln Tyr Pro Ile Pro
595 600 605

Arg Asn Glu His Glu Thr Thr Ser Leu Gln Ser Pro Gly Glu Ala Pro
610 615 620

Glu Ser Ile Leu Tyr Ser Phe Asp Tyr Arg His Gly Asn Tyr Thr Thr
625 630 635 640

Thr Ala Leu Ser Arg Ile Ser Gln Asp Trp Ala Leu Lys Asp Thr Val
645 650 655

Ser Lys Ile Thr Glu Pro Asp Arg Gln Gln Leu Leu Lys Gln Ala Leu
660 665 670

Glu Cys Leu Gln Ile Ser Glu Glu Thr Gln Glu Lys Lys Glu Lys Glu
675 680 685

Val Gln Gln Leu Ile Ser Asn Leu Arg Gln Gln Gln Gln Leu Tyr Arg
690 695 700

Glu Arg Ile Ile Ser Leu Leu Lys Asp Gln
705 710

<210> 5
<211> 688
<212> PRT
<213> 人工序列

<220>
<223> 人工序列描述：合成多肽

<400> 5
Met Gly Ser Ser His His His His His His Gly Ser Asp Tyr Lys Asp
1 5 10 15

Asp Asp Asp Lys Ser Gly Ser Leu Glu Val Leu Phe Gln Gly Pro Ser
20 25 30

Gly Met Pro Pro Tyr Trp Arg Gln Lys Tyr Tyr Arg Arg Arg Tyr Arg
35 40 45

Pro Phe Ser Trp Arg Thr Arg Arg Ile Ile Gln Arg Arg His Arg Trp
50 55 60

Arg Tyr Arg Lys Pro Arg His Thr Tyr Trp Arg Arg His Leu Arg Val
65 70 75 80

Arg His Arg Phe Tyr His Arg His Leu His His Ile Val Leu Lys Gln
85 90 95

Phe Gln Pro Lys Ile Ile Arg Arg Cys Thr Ile Phe Gly Thr Ile Cys

100 105 110

Leu Phe Gln Gly Ser Pro Glu Arg Ala Asn Asn Asn Tyr Ile Gln Thr
115 120 125

Ile Tyr Ser Tyr Val Pro Asp Lys Glu Pro Gly Gly Gly Gly Trp Thr
130 135 140

Leu Ile Thr Glu Ser Leu Ser Ser Leu Trp Glu Asp Trp Glu His Leu
145 150 155 160

Lys Asn Val Trp Thr Gln Ser Asn Ala Gly Leu Pro Leu Val Arg Tyr
165 170 175

Gly Gly Val Thr Leu Tyr Phe Tyr Gln Ser Ala Tyr Thr Asp Tyr Ile
180 185 190

Ala Gln Val Phe Asn Cys Tyr Pro Met Thr Asp Thr Lys Tyr Thr His
195 200 205

Ala Asp Ser Ala Pro Asn Arg Met Leu Leu Lys Lys His Val Ile Arg
210 215 220

Val Pro Ser Arg Glu Thr Arg Lys Lys Arg Lys Pro Tyr Lys Arg Val
225 230 235 240

Arg Val Gly Pro Pro Ser Gln Met Gln Asn Lys Trp Tyr Phe Gln Arg
245 250 255

Asp Ile Cys Glu Ile Pro Leu Ile Met Ile Ala Ala Thr Ala Val Asp
260 265 270

Phe Arg Tyr Pro Phe Cys Ala Ser Asp Cys Ala Ser Asn Asn Leu Thr
275 280 285

Leu Thr Cys Leu Asn Pro Leu Leu Phe Gln Asn Gln Asp Phe Asp His
290 295 300

Pro Ser Asp Thr Gln Gly Tyr Phe Pro Lys Pro Gly Val Tyr Leu Tyr
305 310 315 320

Ser Thr Gln Arg Ser Asn Lys Pro Ser Ser Ser Asp Cys Ile Tyr Leu
325 330 335

Gly Asn Thr Lys Asp Asn Gln Glu Gly Lys Ser Ala Ser Ser Leu Met
340 345 350

Thr Leu Lys Thr Gln Lys Ile Thr Asp Trp Gly Asn Pro Phe Trp His
355 360 365

Tyr Tyr Ile Asp Gly Ser Lys Lys Ile Phe Ser Tyr Phe Lys Pro Pro
370 375 380

Ser Gln Leu Asp Ser Ser Asp Phe Glu His Met Thr Glu Leu Ala Glu
385 390 395 400

Pro Met Phe Ile Gln Val Arg Tyr Asn Pro Glu Arg Asp Thr Gly Gln
405 410 415

Gly Asn Leu Ile Tyr Val Thr Glu Asn Phe Arg Gly Gln His Trp Asp
420 425 430

Pro Pro Ser Ser Asp Asn Leu Lys Leu Asp Gly Phe Pro Leu Tyr Asp
435 440 445

Met Cys Trp Gly Phe Ile Asp Trp Ile Glu Lys Val His Glu Thr Glu
450 455 460

Asn Leu Leu Thr Asn Tyr Cys Phe Cys Ile Arg Ser Ser Ala Phe Asn
465 470 475 480

Glu Lys Lys Thr Val Phe Ile Pro Val Asp His Ser Phe Leu Thr Gly
485 490 495

Phe Ser Pro Tyr Glu Thr Pro Val Lys Ser Ser Asp Gln Ala His Trp

500 505 510

His Pro Gln Ile Arg Phe Gln Thr Lys Ser Ile Asn Asp Ile Cys Leu
515 520 525

Thr Gly Pro Gly Cys Ala Arg Ser Pro Tyr Gly Asn Tyr Met Gln Ala
530 535 540

Lys Met Ser Tyr Lys Phe His Val Lys Trp Gly Gly Cys Pro Lys Thr
545 550 555 560

Tyr Glu Lys Pro Tyr Asp Pro Cys Ser Gln Pro Asn Trp Thr Ile Pro
565 570 575

His Asn Leu Asn Glu Thr Ile Gln Ile Gln Asn Pro Asn Thr Cys Pro
580 585 590

Gln Thr Glu Leu Gln Glu Trp Asp Trp Arg Arg Asp Ile Val Thr Lys
595 600 605

Lys Ala Ile Glu Arg Ile Arg Gln His Thr Glu Pro His Glu Thr Leu
610 615 620

Gln Ile Ser Thr Gly Ser Lys His Asn Pro Pro Val His Arg Gln Thr
625 630 635 640

Ser Pro Trp Thr Asp Ser Glu Thr Asp Ser Glu Glu Glu Lys Asp Gln
645 650 655

Thr Gln Glu Ile Gln Ile Gln Leu Asn Lys Leu Arg Lys His Gln Gln
660 665 670

His Leu Lys Gln Gln Leu Lys Gln Tyr Leu Lys Pro Gln Asn Ile Glu
675 680 685

<210> 6
<211> 687
<212> PRT

<213> 人工序列

<220>
<223> 人工序列描述：合成多肽

<400> 6

```
Met Gly Ser Ser His His His His His His Gly Ser Asp Tyr Lys Asp
1               5                  10                  15

Asp Asp Asp Lys Ser Gly Ser Leu Glu Val Leu Phe Gln Gly Pro Ser
            20                  25                  30

Gly Met Trp Gly Thr Ser Asn Cys Ala Cys Ala Lys Phe Gln Ile Arg
        35                  40                  45

Arg Arg Tyr Ala Arg Pro Tyr Arg Arg Arg His Ile Arg Arg Tyr Arg
    50                  55                  60

Arg Arg Arg Arg His Phe Arg Arg Arg Arg Phe Thr Thr Asn Arg Arg
65                  70                  75                  80

Lys Arg Phe Tyr Lys Arg Lys Leu Lys Lys Ile Val Leu Lys Gln Phe
                85                  90                  95

Gln Pro Lys Ile Ile Arg Arg Cys Thr Ile Phe Gly Thr Ile Cys Leu
            100                 105                 110

Phe Gln Gly Ser Pro Glu Arg Ala Asn Asn Asn Tyr Ile Gln Thr Ile
        115                 120                 125

Tyr Ser Tyr Val Pro Asp Lys Glu Pro Gly Gly Gly Gly Trp Thr Leu
    130                 135                 140

Ile Thr Glu Ser Leu Ser Ser Leu Trp Glu Asp Trp Glu His Leu Lys
145                 150                 155                 160

Asn Val Trp Thr Gln Ser Asn Ala Gly Leu Pro Leu Val Arg Tyr Gly
                165                 170                 175
```

```
Gly Val Thr Leu Tyr Phe Tyr Gln Ser Ala Tyr Thr Asp Tyr Ile Ala
            180                 185                 190

Gln Val Phe Asn Cys Tyr Pro Met Thr Asp Thr Lys Tyr Thr His Ala
        195                 200                 205

Asp Ser Ala Pro Asn Arg Met Leu Leu Lys Lys His Val Ile Arg Val
    210                 215                 220

Pro Ser Arg Glu Thr Arg Lys Lys Arg Lys Pro Tyr Lys Arg Val Arg
225                 230                 235                 240

Val Gly Pro Pro Ser Gln Met Gln Asn Lys Trp Tyr Phe Gln Arg Asp
                245                 250                 255

Ile Cys Glu Ile Pro Leu Ile Met Ile Ala Ala Thr Ala Val Asp Phe
            260                 265                 270

Arg Tyr Pro Phe Cys Ala Ser Asp Cys Ala Ser Asn Asn Leu Thr Leu
        275                 280                 285

Thr Cys Leu Asn Pro Leu Leu Phe Gln Asn Gln Asp Phe Asp His Pro
    290                 295                 300

Ser Asp Thr Gln Gly Tyr Phe Pro Lys Pro Gly Val Tyr Leu Tyr Ser
305                 310                 315                 320

Thr Gln Arg Ser Asn Lys Pro Ser Ser Ser Asp Cys Ile Tyr Leu Gly
                325                 330                 335

Asn Thr Lys Asp Asn Gln Glu Gly Lys Ser Ala Ser Ser Leu Met Thr
            340                 345                 350

Leu Lys Thr Gln Lys Ile Thr Asp Trp Gly Asn Pro Phe Trp His Tyr
        355                 360                 365

Tyr Ile Asp Gly Ser Lys Lys Ile Phe Ser Tyr Phe Lys Pro Pro Ser
    370                 375                 380
```

Gln Leu Asp Ser Ser Asp Phe Glu His Met Thr Glu Leu Ala Glu Pro
385 390 395 400

Met Phe Ile Gln Val Arg Tyr Asn Pro Glu Arg Asp Thr Gly Gln Gly
405 410 415

Asn Leu Ile Tyr Val Thr Glu Asn Phe Arg Gly Gln His Trp Asp Pro
420 425 430

Pro Ser Ser Asp Asn Leu Lys Leu Asp Gly Phe Pro Leu Tyr Asp Met
435 440 445

Cys Trp Gly Phe Ile Asp Trp Ile Glu Lys Val His Glu Thr Glu Asn
450 455 460

Leu Leu Thr Asn Tyr Cys Phe Cys Ile Arg Ser Ser Ala Phe Asn Glu
465 470 475 480

Lys Lys Thr Val Phe Ile Pro Val Asp His Ser Phe Leu Thr Gly Phe
485 490 495

Ser Pro Tyr Glu Thr Pro Val Lys Ser Ser Asp Gln Ala His Trp His
500 505 510

Pro Gln Ile Arg Phe Gln Thr Lys Ser Ile Asn Asp Ile Cys Leu Thr
515 520 525

Gly Pro Gly Cys Ala Arg Ser Pro Tyr Gly Asn Tyr Met Gln Ala Lys
530 535 540

Met Ser Tyr Lys Phe His Val Lys Trp Gly Gly Cys Pro Lys Thr Tyr
545 550 555 560

Glu Lys Pro Tyr Asp Pro Cys Ser Gln Pro Asn Trp Thr Ile Pro His
565 570 575

Asn Leu Asn Glu Thr Ile Gln Ile Gln Asn Pro Asn Thr Cys Pro Gln
580 585 590

Thr Glu Leu Gln Glu Trp Asp Trp Arg Arg Asp Ile Val Thr Lys Lys
595 600 605

Ala Ile Glu Arg Ile Arg Gln His Thr Glu Pro His Glu Thr Leu Gln
610 615 620

Ile Ser Thr Gly Ser Lys His Asn Pro Pro Val His Arg Gln Thr Ser
625 630 635 640

Pro Trp Thr Asp Ser Glu Thr Asp Ser Glu Glu Glu Lys Asp Gln Thr
645 650 655

Gln Glu Ile Gln Ile Gln Leu Asn Lys Leu Arg Lys His Gln Gln His
660 665 670

Leu Lys Gln Gln Leu Lys Gln Tyr Leu Lys Pro Gln Asn Ile Glu
675 680 685

<210> 7
<211> 688
<212> PRT
<213> 人工序列

<220>
<223> 人工序列描述：合成多肽

<400> 7
Met Gly Ser Ser His His His His His His Gly Ser Asp Tyr Lys Asp
1 5 10 15

Asp Asp Asp Lys Ser Gly Ser Leu Glu Val Leu Phe Gln Gly Pro Ser
20 25 30

Gly Met Pro Pro Tyr Trp Arg Gln Lys Tyr Tyr Arg Arg Arg Tyr Arg
35 40 45

Pro Phe Ser Trp Arg Thr Arg Arg Ile Ile Gln Arg Arg Lys Arg Trp

50 55 60

Arg Tyr Arg Lys Pro Arg Lys Thr Tyr Trp Arg Arg Lys Leu Arg Val
65 70 75 80

Arg Lys Arg Phe Tyr Lys Arg Lys Leu Lys Lys Ile Val Leu Lys Gln
85 90 95

Phe Gln Pro Lys Ile Ile Arg Arg Cys Thr Ile Phe Gly Thr Ile Cys
100 105 110

Leu Phe Gln Gly Ser Pro Glu Arg Ala Asn Asn Asn Tyr Ile Gln Thr
115 120 125

Ile Tyr Ser Tyr Val Pro Asp Lys Glu Pro Gly Gly Gly Gly Trp Thr
130 135 140

Leu Ile Thr Glu Ser Leu Ser Ser Leu Trp Glu Asp Trp Glu His Leu
145 150 155 160

Lys Asn Val Trp Thr Gln Ser Asn Ala Gly Leu Pro Leu Val Arg Tyr
165 170 175

Gly Gly Val Thr Leu Tyr Phe Tyr Gln Ser Ala Tyr Thr Asp Tyr Ile
180 185 190

Ala Gln Val Phe Asn Cys Tyr Pro Met Thr Asp Thr Lys Tyr Thr His
195 200 205

Ala Asp Ser Ala Pro Asn Arg Met Leu Leu Lys Lys His Ala Lys Lys
210 215 220

Trp Phe Ser Arg Glu Thr Arg Lys Lys Arg Lys Pro Gly Phe Lys Arg
225 230 235 240

Leu Leu Gly Pro Pro Ser Gln Met Gln Asn Lys Trp Tyr Phe Gln Arg
245 250 255

Asp Ile Cys Glu Ile Pro Leu Ile Met Ile Ala Ala Thr Ala Val Asp
260 265 270

Phe Arg Tyr Pro Phe Cys Ala Ser Asp Cys Ala Ser Asn Asn Leu Thr
275 280 285

Leu Thr Cys Leu Asn Pro Leu Leu Phe Gln Asn Gln Asp Phe Asp His
290 295 300

Pro Ser Asp Thr Gln Gly Tyr Phe Pro Lys Pro Gly Val Tyr Leu Tyr
305 310 315 320

Ser Thr Gln Arg Ser Asn Lys Pro Ser Ser Ser Asp Cys Ile Tyr Leu
325 330 335

Gly Asn Thr Lys Asp Asn Gln Glu Gly Lys Ser Ala Ser Ser Leu Met
340 345 350

Thr Leu Lys Thr Gln Lys Ile Thr Asp Trp Gly Asn Pro Phe Trp His
355 360 365

Tyr Tyr Ile Asp Gly Ser Lys Lys Ile Phe Ser Tyr Phe Lys Pro Pro
370 375 380

Ser Gln Leu Asp Ser Ser Asp Phe Glu His Met Thr Glu Leu Ala Glu
385 390 395 400

Pro Met Phe Ile Gln Val Arg Tyr Asn Pro Glu Arg Asp Thr Gly Gln
405 410 415

Gly Asn Leu Ile Tyr Val Thr Glu Asn Phe Arg Gly Gln His Trp Asp
420 425 430

Pro Pro Ser Ser Asp Asn Leu Lys Leu Asp Gly Phe Pro Leu Tyr Asp
435 440 445

Met Cys Trp Gly Phe Ile Asp Trp Ile Glu Lys Val His Glu Thr Glu

450 455 460

Asn Leu Leu Thr Asn Tyr Cys Phe Cys Ile Arg Ser Ser Ala Phe Asn
465 470 475 480

Glu Lys Lys Thr Val Phe Ile Pro Val Asp His Ser Phe Leu Thr Gly
485 490 495

Phe Ser Pro Tyr Glu Thr Pro Val Lys Ser Ser Asp Gln Ala His Trp
500 505 510

His Pro Gln Ile Arg Phe Gln Thr Lys Ser Ile Asn Asp Ile Cys Leu
515 520 525

Thr Gly Pro Gly Cys Ala Arg Ser Pro Tyr Gly Asn Tyr Met Gln Ala
530 535 540

Lys Met Ser Tyr Lys Phe His Val Lys Trp Gly Gly Cys Pro Lys Thr
545 550 555 560

Tyr Glu Lys Pro Tyr Asp Pro Cys Ser Gln Pro Asn Trp Thr Ile Pro
565 570 575

His Asn Leu Asn Glu Thr Ile Gln Ile Gln Asn Pro Asn Thr Cys Pro
580 585 590

Gln Thr Glu Leu Gln Glu Trp Asp Trp Arg Arg Asp Ile Val Thr Lys
595 600 605

Lys Ala Ile Glu Arg Ile Arg Gln His Thr Glu Pro His Glu Thr Leu
610 615 620

Gln Ile Ser Thr Gly Ser Lys His Asn Pro Pro Val His Arg Gln Thr
625 630 635 640

Ser Pro Trp Thr Asp Ser Glu Thr Asp Ser Glu Glu Glu Lys Asp Gln
645 650 655

Thr Gln Glu Ile Gln Ile Gln Leu Asn Lys Leu Arg Lys His Gln Gln
660 665 670

His Leu Lys Gln Gln Leu Lys Gln Tyr Leu Lys Pro Gln Asn Ile Glu
675 680 685

<210> 8
<211> 699
<212> PRT
<213> 人工序列

<220>
<223> 人工序列描述：合成多肽

<400> 8
Met Gly Ser Ser His His His His His His Gly Ser Asp Tyr Lys Asp
1 5 10 15

Asp Asp Asp Lys Ser Gly Ser Leu Glu Val Leu Phe Gln Gly Pro Ser
20 25 30

Gly Met Pro Tyr Tyr Tyr Arg Arg Arg Arg Tyr Asn Tyr Arg Arg Pro
35 40 45

Arg Trp Tyr Gly Arg Gly Trp Ile Arg Arg Pro Phe Arg Arg Arg Phe
50 55 60

Arg Arg Lys Arg Arg Val Arg Pro Thr Tyr Thr Thr Ile Pro Leu Lys
65 70 75 80

Gln Trp Gln Pro Pro Tyr Lys Arg Arg Cys Thr Ile Phe Gly Thr Ile
85 90 95

Cys Leu Phe Gln Gly Ser Asn Leu Arg Leu Gly Met Asn Ser Thr Met
100 105 110

Tyr Glu Lys Ser Ile Val Pro Val His Trp Pro Gly Gly Gly Ser Phe
115 120 125

Ser Val Ser Met Leu Thr Leu Asp Ala Leu Tyr Asp Ile His Lys Leu
130 135 140

Cys Arg Asn Trp Trp Thr Ser Thr Asn Gln Asp Leu Pro Leu Val Arg
145 150 155 160

Tyr Lys Gly Cys Lys Ile Thr Phe Tyr Gln Ser Thr Phe Thr Asp Tyr
165 170 175

Ile Val Arg Ile His Thr Glu Leu Pro Ala Asn Ser Asn Lys Leu Thr
180 185 190

Tyr Pro Asn Thr His Pro Leu Met Met Met Met Ser Lys Tyr Lys His
195 200 205

Ile Ile Pro Ser Arg Gln Thr Arg Arg Lys Lys Lys Pro Tyr Thr Lys
210 215 220

Ile Phe Val Lys Pro Pro Pro Gln Phe Glu Asn Lys Trp Tyr Phe Ala
225 230 235 240

Thr Asp Leu Tyr Lys Ile Pro Leu Leu Gln Ile His Cys Thr Ala Cys
245 250 255

Asn Leu Gln Asn Pro Phe Val Lys Pro Asp Lys Leu Ser Asn Asn Val
260 265 270

Thr Leu Trp Ser Leu Asn Thr Ile Ser Ile Gln Asn Arg Asn Met Ser
275 280 285

Val Asp Gln Gly Gln Ser Trp Pro Phe Lys Ile Leu Gly Thr Gln Ser
290 295 300

Phe Tyr Phe Tyr Phe Tyr Thr Gly Ala Asn Leu Pro Gly Asp Thr Thr
305 310 315 320

Gln Ile Pro Val Ala Asp Leu Leu Pro Leu Thr Asn Pro Arg Ile Asn
325 330 335

Arg Pro Gly Gln Ser Leu Asn Glu Ala Lys Ile Thr Asp His Ile Thr
340 345 350

Phe Thr Glu Tyr Lys Asn Lys Phe Thr Asn Tyr Trp Gly Asn Pro Phe
355 360 365

Asn Lys His Ile Gln Glu His Leu Asp Met Ile Leu Tyr Ser Leu Lys
370 375 380

Ser Pro Glu Ala Ile Lys Asn Glu Trp Thr Thr Glu Asn Met Lys Trp
385 390 395 400

Asn Gln Leu Asn Asn Ala Gly Thr Met Ala Leu Thr Pro Phe Asn Glu
405 410 415

Pro Ile Phe Thr Gln Ile Gln Tyr Asn Pro Asp Arg Asp Thr Gly Glu
420 425 430

Asp Thr Gln Leu Tyr Leu Leu Ser Asn Ala Thr Gly Thr Gly Trp Asp
435 440 445

Pro Pro Gly Ile Pro Glu Leu Ile Leu Glu Gly Phe Pro Leu Trp Leu
450 455 460

Ile Tyr Trp Gly Phe Ala Asp Phe Gln Lys Asn Leu Lys Lys Val Thr
465 470 475 480

Asn Ile Asp Thr Asn Tyr Met Leu Val Ala Lys Thr Lys Phe Thr Gln
485 490 495

Lys Pro Gly Thr Phe Tyr Leu Val Ile Leu Asn Asp Thr Phe Val Glu
500 505 510

Gly Asn Ser Pro Tyr Glu Lys Gln Pro Leu Pro Glu Asp Asn Ile Lys
515 520 525

Trp Tyr Pro Gln Val Gln Tyr Gln Leu Glu Ala Gln Asn Lys Leu Leu
530 535 540

Gln Thr Gly Pro Phe Thr Pro Asn Ile Gln Gly Gln Leu Ser Asp Asn
545 550 555 560

Ile Ser Met Phe Tyr Lys Phe Tyr Phe Lys Trp Gly Gly Ser Pro Pro
565 570 575

Lys Ala Ile Asn Val Glu Asn Pro Ala His Gln Ile Gln Tyr Pro Ile
580 585 590

Pro Arg Asn Glu His Glu Thr Thr Ser Leu Gln Ser Pro Gly Glu Ala
595 600 605

Pro Glu Ser Ile Leu Tyr Ser Phe Asp Tyr Arg His Gly Asn Tyr Thr
610 615 620

Thr Thr Ala Leu Ser Arg Ile Ser Gln Asp Trp Ala Leu Lys Asp Thr
625 630 635 640

Val Ser Lys Ile Thr Glu Pro Asp Arg Gln Gln Leu Leu Lys Gln Ala
645 650 655

Leu Glu Cys Leu Gln Ile Ser Glu Glu Thr Gln Glu Lys Lys Glu Lys
660 665 670

Glu Val Gln Gln Leu Ile Ser Asn Leu Arg Gln Gln Gln Gln Leu Tyr
675 680 685

Arg Glu Arg Ile Ile Ser Leu Leu Lys Asp Gln
690 695

<210> 9
<211> 699
<212> PRT
<213> 人工序列

<220>

<223> 人工序列描述：合成多肽

<400> 9

```
Met Gly Ser Ser His His His His His His Gly Ser Asp Tyr Lys Asp
1               5                  10                  15

Asp Asp Asp Lys Ser Gly Ser Leu Glu Val Leu Phe Gln Gly Pro Ser
            20                  25                  30

Gly Met Pro Tyr Tyr Tyr Arg Arg Arg Arg Tyr Asn Tyr Arg Arg Pro
        35                  40                  45

Arg Trp Tyr Gly Arg Gly Trp Ile Arg Arg Pro Phe Arg Arg Arg Phe
    50                  55                  60

Arg Arg Lys Arg Arg Val Arg Pro Thr Tyr Thr Thr Ile Pro Leu Lys
65                  70                  75                  80

Gln Trp Gln Pro Pro Tyr Lys Arg Thr Cys Tyr Ile Lys Gly Gln Asp
                85                  90                  95

Cys Leu Ile Tyr Tyr Ser Pro Glu Arg Ala Asn Asn Asn Tyr Ile Gln
            100                 105                 110

Thr Ile Tyr Ser Tyr Val Pro Asp Lys Glu Pro Gly Gly Gly Gly Trp
        115                 120                 125

Thr Leu Ile Thr Glu Ser Leu Ser Ser Leu Trp Glu Asp Trp Glu His
    130                 135                 140

Leu Lys Asn Val Trp Thr Gln Ser Asn Ala Gly Leu Pro Leu Val Arg
145                 150                 155                 160

Tyr Lys Gly Cys Lys Ile Thr Phe Tyr Gln Ser Thr Phe Thr Asp Tyr
                165                 170                 175

Ile Val Arg Ile His Thr Glu Leu Pro Ala Asn Ser Asn Lys Leu Thr
            180                 185                 190
```

Tyr Pro Asn Thr His Pro Leu Met Met Met Met Ser Lys Tyr Lys His
195 200 205

Ile Ile Pro Ser Arg Gln Thr Arg Arg Lys Lys Lys Pro Tyr Thr Lys
210 215 220

Ile Phe Val Lys Pro Pro Pro Gln Phe Glu Asn Lys Trp Tyr Phe Ala
225 230 235 240

Thr Asp Leu Tyr Lys Ile Pro Leu Leu Gln Ile His Cys Thr Ala Cys
245 250 255

Asn Leu Gln Asn Pro Phe Val Lys Pro Asp Lys Leu Ser Asn Asn Val
260 265 270

Thr Leu Trp Ser Leu Asn Thr Ile Ser Ile Gln Asn Arg Asn Met Ser
275 280 285

Val Asp Gln Gly Gln Ser Trp Pro Phe Lys Ile Leu Gly Thr Gln Ser
290 295 300

Phe Tyr Phe Tyr Phe Tyr Thr Gly Ala Asn Leu Pro Gly Asp Thr Thr
305 310 315 320

Gln Ile Pro Val Ala Asp Leu Leu Pro Leu Thr Asn Pro Arg Ile Asn
325 330 335

Arg Pro Gly Gln Ser Leu Asn Glu Ala Lys Ile Thr Asp His Ile Thr
340 345 350

Phe Thr Glu Tyr Lys Asn Lys Phe Thr Asn Tyr Trp Gly Asn Pro Phe
355 360 365

Asn Lys His Ile Gln Glu His Leu Asp Met Ile Leu Tyr Ser Leu Lys
370 375 380

Ser Pro Glu Ala Ile Lys Asn Glu Trp Thr Thr Glu Asn Met Lys Trp

385 390 395 400

Asn Gln Leu Asn Asn Ala Gly Thr Met Ala Leu Thr Pro Phe Asn Glu
405 410 415

Pro Ile Phe Thr Gln Ile Gln Tyr Asn Pro Asp Arg Asp Thr Gly Glu
420 425 430

Asp Thr Gln Leu Tyr Leu Leu Ser Asn Ala Thr Gly Thr Gly Trp Asp
435 440 445

Pro Pro Gly Ile Pro Glu Leu Ile Leu Glu Gly Phe Pro Leu Trp Leu
450 455 460

Ile Tyr Trp Gly Phe Ala Asp Phe Gln Lys Asn Leu Lys Lys Val Thr
465 470 475 480

Asn Ile Asp Thr Asn Tyr Met Leu Val Ala Lys Thr Lys Phe Thr Gln
485 490 495

Lys Pro Gly Thr Phe Tyr Leu Val Ile Leu Asn Asp Thr Phe Val Glu
500 505 510

Gly Asn Ser Pro Tyr Glu Lys Gln Pro Leu Pro Glu Asp Asn Ile Lys
515 520 525

Trp Tyr Pro Gln Val Gln Tyr Gln Leu Glu Ala Gln Asn Lys Leu Leu
530 535 540

Gln Thr Gly Pro Phe Thr Pro Asn Ile Gln Gly Gln Leu Ser Asp Asn
545 550 555 560

Ile Ser Met Phe Tyr Lys Phe Tyr Phe Lys Trp Gly Gly Ser Pro Pro
565 570 575

Lys Ala Ile Asn Val Glu Asn Pro Ala His Gln Ile Gln Tyr Pro Ile
580 585 590

Pro Arg Asn Glu His Glu Thr Thr Ser Leu Gln Ser Pro Gly Glu Ala
595 600 605

Pro Glu Ser Ile Leu Tyr Ser Phe Asp Tyr Arg His Gly Asn Tyr Thr
610 615 620

Thr Thr Ala Leu Ser Arg Ile Ser Gln Asp Trp Ala Leu Lys Asp Thr
625 630 635 640

Val Ser Lys Ile Thr Glu Pro Asp Arg Gln Gln Leu Leu Lys Gln Ala
645 650 655

Leu Glu Cys Leu Gln Ile Ser Glu Glu Thr Gln Glu Lys Lys Glu Lys
660 665 670

Glu Val Gln Gln Leu Ile Ser Asn Leu Arg Gln Gln Gln Gln Leu Tyr
675 680 685

Arg Glu Arg Ile Ile Ser Leu Leu Lys Asp Gln
690 695

<210> 10
<211> 699
<212> PRT
<213> 人工序列

<220>
<223> 人工序列描述：合成多肽

<400> 10
Met Gly Ser Ser His His His His His His Gly Ser Asp Tyr Lys Asp
1 5 10 15

Asp Asp Asp Lys Ser Gly Ser Leu Glu Val Leu Phe Gln Gly Pro Ser
20 25 30

Gly Met Pro Tyr Tyr Tyr Arg Arg Arg Arg Tyr Asn Tyr Arg Arg Pro
35 40 45

Arg Trp Tyr Gly Arg Gly Trp Ile Arg Arg Pro Phe Arg Arg Arg Phe
50 55 60

Arg Arg Lys Arg Arg Val Arg Pro Thr Tyr Thr Thr Ile Pro Leu Lys
65 70 75 80

Gln Trp Gln Pro Pro Tyr Lys Arg Thr Cys Tyr Ile Lys Gly Gln Asp
85 90 95

Cys Leu Ile Tyr Tyr Ser Asn Leu Arg Leu Gly Met Asn Ser Thr Met
100 105 110

Tyr Glu Lys Ser Ile Val Pro Val His Trp Pro Gly Gly Gly Ser Phe
115 120 125

Ser Val Ser Met Leu Thr Leu Asp Ala Leu Tyr Asp Ile His Lys Leu
130 135 140

Cys Arg Asn Trp Trp Thr Ser Thr Asn Gln Asp Leu Pro Leu Val Arg
145 150 155 160

Tyr Gly Gly Val Thr Leu Tyr Phe Tyr Gln Ser Thr Phe Thr Asp Tyr
165 170 175

Ile Val Arg Ile His Thr Glu Leu Pro Ala Asn Ser Asn Lys Leu Thr
180 185 190

Tyr Pro Asn Thr His Pro Leu Met Met Met Met Ser Lys Tyr Lys His
195 200 205

Ile Ile Pro Ser Arg Gln Thr Arg Arg Lys Lys Lys Pro Tyr Thr Lys
210 215 220

Ile Phe Val Lys Pro Pro Pro Gln Phe Glu Asn Lys Trp Tyr Phe Ala
225 230 235 240

Thr Asp Leu Tyr Lys Ile Pro Leu Leu Gln Ile His Cys Thr Ala Cys
245 250 255

Asn Leu Gln Asn Pro Phe Val Lys Pro Asp Lys Leu Ser Asn Asn Val
260 265 270

Thr Leu Trp Ser Leu Asn Thr Ile Ser Ile Gln Asn Arg Asn Met Ser
275 280 285

Val Asp Gln Gly Gln Ser Trp Pro Phe Lys Ile Leu Gly Thr Gln Ser
290 295 300

Phe Tyr Phe Tyr Phe Tyr Thr Gly Ala Asn Leu Pro Gly Asp Thr Thr
305 310 315 320

Gln Ile Pro Val Ala Asp Leu Leu Pro Leu Thr Asn Pro Arg Ile Asn
325 330 335

Arg Pro Gly Gln Ser Leu Asn Glu Ala Lys Ile Thr Asp His Ile Thr
340 345 350

Phe Thr Glu Tyr Lys Asn Lys Phe Thr Asn Tyr Trp Gly Asn Pro Phe
355 360 365

Asn Lys His Ile Gln Glu His Leu Asp Met Ile Leu Tyr Ser Leu Lys
370 375 380

Ser Pro Glu Ala Ile Lys Asn Glu Trp Thr Thr Glu Asn Met Lys Trp
385 390 395 400

Asn Gln Leu Asn Asn Ala Gly Thr Met Ala Leu Thr Pro Phe Asn Glu
405 410 415

Pro Ile Phe Thr Gln Ile Gln Tyr Asn Pro Asp Arg Asp Thr Gly Glu
420 425 430

Asp Thr Gln Leu Tyr Leu Leu Ser Asn Ala Thr Gly Thr Gly Trp Asp
435 440 445

Pro Pro Gly Ile Pro Glu Leu Ile Leu Glu Gly Phe Pro Leu Trp Leu
450 455 460

Ile Tyr Trp Gly Phe Ala Asp Phe Gln Lys Asn Leu Lys Lys Val Thr
465 470 475 480

Asn Ile Asp Thr Asn Tyr Met Leu Val Ala Lys Thr Lys Phe Thr Gln
485 490 495

Lys Pro Gly Thr Phe Tyr Leu Val Ile Leu Asn Asp Thr Phe Val Glu
500 505 510

Gly Asn Ser Pro Tyr Glu Lys Gln Pro Leu Pro Glu Asp Asn Ile Lys
515 520 525

Trp Tyr Pro Gln Val Gln Tyr Gln Leu Glu Ala Gln Asn Lys Leu Leu
530 535 540

Gln Thr Gly Pro Phe Thr Pro Asn Ile Gln Gly Gln Leu Ser Asp Asn
545 550 555 560

Ile Ser Met Phe Tyr Lys Phe Tyr Phe Lys Trp Gly Gly Ser Pro Pro
565 570 575

Lys Ala Ile Asn Val Glu Asn Pro Ala His Gln Ile Gln Tyr Pro Ile
580 585 590

Pro Arg Asn Glu His Glu Thr Thr Ser Leu Gln Ser Pro Gly Glu Ala
595 600 605

Pro Glu Ser Ile Leu Tyr Ser Phe Asp Tyr Arg His Gly Asn Tyr Thr
610 615 620

Thr Thr Ala Leu Ser Arg Ile Ser Gln Asp Trp Ala Leu Lys Asp Thr
625 630 635 640

Val Ser Lys Ile Thr Glu Pro Asp Arg Gln Gln Leu Leu Lys Gln Ala
645 650 655

Leu Glu Cys Leu Gln Ile Ser Glu Glu Thr Gln Glu Lys Lys Glu Lys
660 665 670

Glu Val Gln Gln Leu Ile Ser Asn Leu Arg Gln Gln Gln Gln Leu Tyr
675 680 685

Arg Glu Arg Ile Ile Ser Leu Leu Lys Asp Gln
690 695

<210> 11
<211> 699
<212> PRT
<213> 人工序列

<220>
<223> 人工序列描述：合成多肽

<400> 11
Met Gly Ser Ser His His His His His His Gly Ser Asp Tyr Lys Asp
1 5 10 15

Asp Asp Asp Lys Ser Gly Ser Leu Glu Val Leu Phe Gln Gly Pro Ser
20 25 30

Gly Met Pro Tyr Tyr Tyr Arg Arg Arg Arg Tyr Asn Tyr Arg Arg Pro
35 40 45

Arg Trp Tyr Gly Arg Gly Trp Ile Arg Arg Pro Phe Arg Arg Arg Phe
50 55 60

Arg Arg Lys Arg Arg Val Arg Pro Thr Tyr Thr Thr Ile Pro Leu Lys
65 70 75 80

Gln Trp Gln Pro Pro Tyr Lys Arg Thr Cys Tyr Ile Lys Gly Gln Asp
85 90 95

Cys Leu Ile Tyr Tyr Ser Asn Leu Arg Leu Gly Met Asn Ser Thr Met
100 105 110

Tyr Glu Lys Ser Ile Val Pro Val His Trp Pro Gly Gly Gly Ser Phe
115 120 125

Ser Val Ser Met Leu Thr Leu Asp Ala Leu Tyr Asp Ile His Lys Leu
130 135 140

Cys Arg Asn Trp Trp Thr Ser Thr Asn Gln Asp Leu Pro Leu Val Arg
145 150 155 160

Tyr Lys Gly Cys Lys Ile Thr Phe Tyr Gln Ser Thr Phe Thr Asp Tyr
165 170 175

Ile Val Arg Ile His Asn Cys Tyr Pro Met Thr Asp Thr Lys Tyr Thr
180 185 190

His Ala Asp Ser Ala Pro Asn Arg Met Leu Leu Lys Lys His Lys His
195 200 205

Ile Ile Pro Ser Arg Gln Thr Arg Arg Lys Lys Lys Pro Tyr Thr Lys
210 215 220

Ile Phe Val Lys Pro Pro Pro Gln Phe Glu Asn Lys Trp Tyr Phe Ala
225 230 235 240

Thr Asp Leu Tyr Lys Ile Pro Leu Leu Gln Ile His Cys Thr Ala Cys
245 250 255

Asn Leu Gln Asn Pro Phe Val Lys Pro Asp Lys Leu Ser Asn Asn Val
260 265 270

Thr Leu Trp Ser Leu Asn Thr Ile Ser Ile Gln Asn Arg Asn Met Ser
275 280 285

Val Asp Gln Gly Gln Ser Trp Pro Phe Lys Ile Leu Gly Thr Gln Ser
290 295 300

Phe Tyr Phe Tyr Phe Tyr Thr Gly Ala Asn Leu Pro Gly Asp Thr Thr

305 310 315 320

Gln Ile Pro Val Ala Asp Leu Leu Pro Leu Thr Asn Pro Arg Ile Asn
325 330 335

Arg Pro Gly Gln Ser Leu Asn Glu Ala Lys Ile Thr Asp His Ile Thr
340 345 350

Phe Thr Glu Tyr Lys Asn Lys Phe Thr Asn Tyr Trp Gly Asn Pro Phe
355 360 365

Asn Lys His Ile Gln Glu His Leu Asp Met Ile Leu Tyr Ser Leu Lys
370 375 380

Ser Pro Glu Ala Ile Lys Asn Glu Trp Thr Thr Glu Asn Met Lys Trp
385 390 395 400

Asn Gln Leu Asn Asn Ala Gly Thr Met Ala Leu Thr Pro Phe Asn Glu
405 410 415

Pro Ile Phe Thr Gln Ile Gln Tyr Asn Pro Asp Arg Asp Thr Gly Glu
420 425 430

Asp Thr Gln Leu Tyr Leu Leu Ser Asn Ala Thr Gly Thr Gly Trp Asp
435 440 445

Pro Pro Gly Ile Pro Glu Leu Ile Leu Glu Gly Phe Pro Leu Trp Leu
450 455 460

Ile Tyr Trp Gly Phe Ala Asp Phe Gln Lys Asn Leu Lys Lys Val Thr
465 470 475 480

Asn Ile Asp Thr Asn Tyr Met Leu Val Ala Lys Thr Lys Phe Thr Gln
485 490 495

Lys Pro Gly Thr Phe Tyr Leu Val Ile Leu Asn Asp Thr Phe Val Glu
500 505 510

Gly Asn Ser Pro Tyr Glu Lys Gln Pro Leu Pro Glu Asp Asn Ile Lys
515 520 525

Trp Tyr Pro Gln Val Gln Tyr Gln Leu Glu Ala Gln Asn Lys Leu Leu
530 535 540

Gln Thr Gly Pro Phe Thr Pro Asn Ile Gln Gly Gln Leu Ser Asp Asn
545 550 555 560

Ile Ser Met Phe Tyr Lys Phe Tyr Phe Lys Trp Gly Gly Ser Pro Pro
565 570 575

Lys Ala Ile Asn Val Glu Asn Pro Ala His Gln Ile Gln Tyr Pro Ile
580 585 590

Pro Arg Asn Glu His Glu Thr Thr Ser Leu Gln Ser Pro Gly Glu Ala
595 600 605

Pro Glu Ser Ile Leu Tyr Ser Phe Asp Tyr Arg His Gly Asn Tyr Thr
610 615 620

Thr Thr Ala Leu Ser Arg Ile Ser Gln Asp Trp Ala Leu Lys Asp Thr
625 630 635 640

Val Ser Lys Ile Thr Glu Pro Asp Arg Gln Gln Leu Leu Lys Gln Ala
645 650 655

Leu Glu Cys Leu Gln Ile Ser Glu Glu Thr Gln Glu Lys Lys Glu Lys
660 665 670

Glu Val Gln Gln Leu Ile Ser Asn Leu Arg Gln Gln Gln Gln Leu Tyr
675 680 685

Arg Glu Arg Ile Ile Ser Leu Leu Lys Asp Gln
690 695

<210> 12

<211> 699
<212> PRT
<213> 人工序列

<220>
<223> 人工序列描述：合成多肽

<400> 12

```
Met Gly Ser Ser His His His His His His Gly Ser Asp Tyr Lys Asp
1               5                  10                  15

Asp Asp Asp Lys Ser Gly Ser Leu Glu Val Leu Phe Gln Gly Pro Ser
            20                  25                  30

Gly Met Pro Tyr Tyr Tyr Arg Arg Arg Arg Tyr Asn Tyr Arg Arg Pro
        35                  40                  45

Arg Trp Tyr Gly Arg Gly Trp Ile Arg Arg Pro Phe Arg Arg Arg Phe
    50                  55                  60

Arg Arg Lys Arg Arg Val Arg Pro Thr Tyr Thr Thr Ile Pro Leu Lys
65                  70                  75                  80

Gln Trp Gln Pro Pro Tyr Lys Arg Thr Cys Tyr Ile Lys Gly Gln Asp
                85                  90                  95

Cys Leu Ile Tyr Tyr Ser Asn Leu Arg Leu Gly Met Asn Ser Thr Met
            100                 105                 110

Tyr Glu Lys Ser Ile Val Pro Val His Trp Pro Gly Gly Gly Ser Phe
        115                 120                 125

Ser Val Ser Met Leu Thr Leu Asp Ala Leu Tyr Asp Ile His Lys Leu
    130                 135                 140

Cys Arg Asn Trp Trp Thr Ser Thr Asn Gln Asp Leu Pro Leu Val Arg
145                 150                 155                 160

Tyr Lys Gly Cys Lys Ile Thr Phe Tyr Gln Ser Thr Phe Thr Asp Tyr
                165                 170                 175
```

Ile Val Arg Ile His Thr Glu Leu Pro Ala Asn Ser Asn Lys Leu Thr
180 185 190

Tyr Pro Asn Thr His Pro Leu Met Met Met Met Ser Lys Tyr Lys His
195 200 205

Ile Ile Pro Ser Arg Gln Thr Arg Arg Lys Lys Lys Pro Tyr Lys Arg
210 215 220

Val Arg Val Lys Pro Pro Pro Gln Phe Glu Asn Lys Trp Tyr Phe Ala
225 230 235 240

Thr Asp Leu Tyr Lys Ile Pro Leu Leu Gln Ile His Cys Thr Ala Cys
245 250 255

Asn Leu Gln Asn Pro Phe Val Lys Pro Asp Lys Leu Ser Asn Asn Val
260 265 270

Thr Leu Trp Ser Leu Asn Thr Ile Ser Ile Gln Asn Arg Asn Met Ser
275 280 285

Val Asp Gln Gly Gln Ser Trp Pro Phe Lys Ile Leu Gly Thr Gln Ser
290 295 300

Phe Tyr Phe Tyr Phe Tyr Thr Gly Ala Asn Leu Pro Gly Asp Thr Thr
305 310 315 320

Gln Ile Pro Val Ala Asp Leu Leu Pro Leu Thr Asn Pro Arg Ile Asn
325 330 335

Arg Pro Gly Gln Ser Leu Asn Glu Ala Lys Ile Thr Asp His Ile Thr
340 345 350

Phe Thr Glu Tyr Lys Asn Lys Phe Thr Asn Tyr Trp Gly Asn Pro Phe
355 360 365

```
Asn Lys His Ile Gln Glu His Leu Asp Met Ile Leu Tyr Ser Leu Lys
    370                 375                 380

Ser Pro Glu Ala Ile Lys Asn Glu Trp Thr Thr Glu Asn Met Lys Trp
385                 390                 395                 400

Asn Gln Leu Asn Asn Ala Gly Thr Met Ala Leu Thr Pro Phe Asn Glu
                405                 410                 415

Pro Ile Phe Thr Gln Ile Gln Tyr Asn Pro Asp Arg Asp Thr Gly Glu
            420                 425                 430

Asp Thr Gln Leu Tyr Leu Leu Ser Asn Ala Thr Gly Thr Gly Trp Asp
        435                 440                 445

Pro Pro Gly Ile Pro Glu Leu Ile Leu Glu Gly Phe Pro Leu Trp Leu
    450                 455                 460

Ile Tyr Trp Gly Phe Ala Asp Phe Gln Lys Asn Leu Lys Lys Val Thr
465                 470                 475                 480

Asn Ile Asp Thr Asn Tyr Met Leu Val Ala Lys Thr Lys Phe Thr Gln
                485                 490                 495

Lys Pro Gly Thr Phe Tyr Leu Val Ile Leu Asn Asp Thr Phe Val Glu
            500                 505                 510

Gly Asn Ser Pro Tyr Glu Lys Gln Pro Leu Pro Glu Asp Asn Ile Lys
        515                 520                 525

Trp Tyr Pro Gln Val Gln Tyr Gln Leu Glu Ala Gln Asn Lys Leu Leu
    530                 535                 540

Gln Thr Gly Pro Phe Thr Pro Asn Ile Gln Gly Gln Leu Ser Asp Asn
545                 550                 555                 560

Ile Ser Met Phe Tyr Lys Phe Tyr Phe Lys Trp Gly Gly Ser Pro Pro
                565                 570                 575
```

Lys Ala Ile Asn Val Glu Asn Pro Ala His Gln Ile Gln Tyr Pro Ile
580 585 590

Pro Arg Asn Glu His Glu Thr Thr Ser Leu Gln Ser Pro Gly Glu Ala
595 600 605

Pro Glu Ser Ile Leu Tyr Ser Phe Asp Tyr Arg His Gly Asn Tyr Thr
610 615 620

Thr Thr Ala Leu Ser Arg Ile Ser Gln Asp Trp Ala Leu Lys Asp Thr
625 630 635 640

Val Ser Lys Ile Thr Glu Pro Asp Arg Gln Gln Leu Leu Lys Gln Ala
645 650 655

Leu Glu Cys Leu Gln Ile Ser Glu Glu Thr Gln Glu Lys Lys Glu Lys
660 665 670

Glu Val Gln Gln Leu Ile Ser Asn Leu Arg Gln Gln Gln Gln Leu Tyr
675 680 685

Arg Glu Arg Ile Ile Ser Leu Leu Lys Asp Gln
690 695

<210> 13
<211> 699
<212> PRT
<213> 人工序列

<220>
<223> 人工序列描述：合成多肽

<400> 13
Met Gly Ser Ser His His His His His His Gly Ser Asp Tyr Lys Asp
1 5 10 15

Asp Asp Asp Lys Ser Gly Ser Leu Glu Val Leu Phe Gln Gly Pro Ser
20 25 30

Gly Met Pro Tyr Tyr Tyr Arg Arg Arg Arg Tyr Asn Tyr Arg Arg Pro
35 40 45

Arg Trp Tyr Gly Arg Gly Trp Ile Arg Arg Pro Phe Arg Arg Arg Phe
50 55 60

Arg Arg Lys Arg Arg Val Arg Pro Thr Tyr Thr Thr Ile Pro Leu Lys
65 70 75 80

Gln Trp Gln Pro Pro Tyr Lys Arg Thr Cys Tyr Ile Lys Gly Gln Asp
85 90 95

Cys Leu Ile Tyr Tyr Ser Asn Leu Arg Leu Gly Met Asn Ser Thr Met
100 105 110

Tyr Glu Lys Ser Ile Val Pro Val His Trp Pro Gly Gly Gly Ser Phe
115 120 125

Ser Val Ser Met Leu Thr Leu Asp Ala Leu Tyr Asp Ile His Lys Leu
130 135 140

Cys Arg Asn Trp Trp Thr Ser Thr Asn Gln Asp Leu Pro Leu Val Arg
145 150 155 160

Tyr Lys Gly Cys Lys Ile Thr Phe Tyr Gln Ser Thr Phe Thr Asp Tyr
165 170 175

Ile Val Arg Ile His Thr Glu Leu Pro Ala Asn Ser Asn Lys Leu Thr
180 185 190

Tyr Pro Asn Thr His Pro Leu Met Met Met Met Ser Lys His Val Ile
195 200 205

Arg Val Pro Ser Arg Gln Thr Arg Arg Lys Lys Lys Pro Tyr Thr Lys
210 215 220

Ile Phe Val Lys Pro Pro Pro Gln Phe Glu Asn Lys Trp Tyr Phe Ala

225 230 235 240

Thr Asp Leu Tyr Lys Ile Pro Leu Leu Gln Ile His Cys Thr Ala Cys
245 250 255

Asn Leu Gln Asn Pro Phe Val Lys Pro Asp Lys Leu Ser Asn Asn Val
260 265 270

Thr Leu Trp Ser Leu Asn Thr Ile Ser Ile Gln Asn Arg Asn Met Ser
275 280 285

Val Asp Gln Gly Gln Ser Trp Pro Phe Lys Ile Leu Gly Thr Gln Ser
290 295 300

Phe Tyr Phe Tyr Phe Tyr Thr Gly Ala Asn Leu Pro Gly Asp Thr Thr
305 310 315 320

Gln Ile Pro Val Ala Asp Leu Leu Pro Leu Thr Asn Pro Arg Ile Asn
325 330 335

Arg Pro Gly Gln Ser Leu Asn Glu Ala Lys Ile Thr Asp His Ile Thr
340 345 350

Phe Thr Glu Tyr Lys Asn Lys Phe Thr Asn Tyr Trp Gly Asn Pro Phe
355 360 365

Asn Lys His Ile Gln Glu His Leu Asp Met Ile Leu Tyr Ser Leu Lys
370 375 380

Ser Pro Glu Ala Ile Lys Asn Glu Trp Thr Thr Glu Asn Met Lys Trp
385 390 395 400

Asn Gln Leu Asn Asn Ala Gly Thr Met Ala Leu Thr Pro Phe Asn Glu
405 410 415

Pro Ile Phe Thr Gln Ile Gln Tyr Asn Pro Asp Arg Asp Thr Gly Glu
420 425 430

Asp Thr Gln Leu Tyr Leu Leu Ser Asn Ala Thr Gly Thr Gly Trp Asp
435 440 445

Pro Pro Gly Ile Pro Glu Leu Ile Leu Glu Gly Phe Pro Leu Trp Leu
450 455 460

Ile Tyr Trp Gly Phe Ala Asp Phe Gln Lys Asn Leu Lys Lys Val Thr
465 470 475 480

Asn Ile Asp Thr Asn Tyr Met Leu Val Ala Lys Thr Lys Phe Thr Gln
485 490 495

Lys Pro Gly Thr Phe Tyr Leu Val Ile Leu Asn Asp Thr Phe Val Glu
500 505 510

Gly Asn Ser Pro Tyr Glu Lys Gln Pro Leu Pro Glu Asp Asn Ile Lys
515 520 525

Trp Tyr Pro Gln Val Gln Tyr Gln Leu Glu Ala Gln Asn Lys Leu Leu
530 535 540

Gln Thr Gly Pro Phe Thr Pro Asn Ile Gln Gly Gln Leu Ser Asp Asn
545 550 555 560

Ile Ser Met Phe Tyr Lys Phe Tyr Phe Lys Trp Gly Gly Ser Pro Pro
565 570 575

Lys Ala Ile Asn Val Glu Asn Pro Ala His Gln Ile Gln Tyr Pro Ile
580 585 590

Pro Arg Asn Glu His Glu Thr Thr Ser Leu Gln Ser Pro Gly Glu Ala
595 600 605

Pro Glu Ser Ile Leu Tyr Ser Phe Asp Tyr Arg His Gly Asn Tyr Thr
610 615 620

Thr Thr Ala Leu Ser Arg Ile Ser Gln Asp Trp Ala Leu Lys Asp Thr

625 630 635 640

Val Ser Lys Ile Thr Glu Pro Asp Arg Gln Gln Leu Leu Lys Gln Ala
645 650 655

Leu Glu Cys Leu Gln Ile Ser Glu Glu Thr Gln Glu Lys Lys Glu Lys
660 665 670

Glu Val Gln Gln Leu Ile Ser Asn Leu Arg Gln Gln Gln Gln Leu Tyr
675 680 685

Arg Glu Arg Ile Ile Ser Leu Leu Lys Asp Gln
690 695

<210> 14
<211> 673
<212> PRT
<213> 人工序列

<220>
<223> 人工序列描述：合成多肽

<400> 14
Met Gly Ser Ser His His His His His His Gly Ser Asp Tyr Lys Asp
1 5 10 15

Asp Asp Asp Lys Ser Gly Ser Leu Glu Val Leu Phe Gln Gly Pro Ser
20 25 30

Gly Met Pro Tyr Tyr Tyr Arg Arg Arg Arg Tyr Asn Tyr Arg Arg Pro
35 40 45

Arg Trp Tyr Gly Arg Gly Trp Ile Arg Arg Pro Phe Arg Arg Arg Phe
50 55 60

Arg Arg Lys Arg Arg Val Arg Pro Thr Tyr Thr Thr Ile Pro Leu Lys
65 70 75 80

Gln Trp Gln Pro Pro Tyr Lys Arg Thr Cys Tyr Ile Lys Gly Gln Asp
85 90 95

Cys Leu Ile Tyr Tyr Ser Asn Leu Arg Leu Gly Met Asn Ser Thr Met
100 105 110

Tyr Glu Lys Ser Ile Val Pro Val His Trp Pro Gly Gly Gly Ser Phe
115 120 125

Ser Val Ser Met Leu Thr Leu Asp Ala Leu Tyr Asp Ile His Lys Leu
130 135 140

Cys Arg Asn Trp Trp Thr Ser Thr Asn Gln Asp Leu Pro Leu Val Arg
145 150 155 160

Tyr Lys Gly Cys Lys Ile Thr Phe Tyr Gln Ser Thr Phe Thr Asp Tyr
165 170 175

Ile Val Arg Ile His Thr Glu Leu Pro Ala Asn Ser Asn Lys Leu Thr
180 185 190

Tyr Pro Asn Thr His Pro Leu Met Met Met Met Ser Lys Tyr Lys His
195 200 205

Ile Ile Pro Ser Arg Gln Thr Arg Arg Lys Lys Lys Pro Tyr Thr Lys
210 215 220

Ile Phe Val Lys Pro Pro Pro Gln Phe Glu Asn Lys Trp Tyr Phe Ala
225 230 235 240

Thr Asp Leu Tyr Lys Ile Pro Leu Ile Met Ile Ala Ala Thr Ala Val
245 250 255

Asp Phe Arg Tyr Pro Phe Cys Ala Ser Asp Cys Ala Ser Asn Asn Leu
260 265 270

Thr Leu Thr Cys Leu Asn Pro Leu Leu Phe Gln Asn Gln Asp Phe Asp
275 280 285

His Pro Ser Asp Thr Gln Gly Tyr Phe Pro Lys Pro Gly Val Tyr Leu
290 295 300

Tyr Ser Thr Gln Arg Ser Asn Lys Pro Ser Ser Ser Asp Cys Ile Tyr
305 310 315 320

Leu Gly Asn Thr Lys Asp Asn Gln Glu Gly Lys Ser Ala Ser Ser Leu
325 330 335

Met Thr Leu Lys Thr Gln Lys Ile Thr Asp Trp Gly Asn Pro Phe Trp
340 345 350

His Tyr Tyr Ile Asp Gly Ser Lys Lys Ile Phe Ser Tyr Phe Lys Pro
355 360 365

Pro Ser Gln Leu Asp Ser Ser Asp Phe Glu His Met Thr Glu Leu Ala
370 375 380

Glu Pro Met Phe Ile Gln Val Arg Tyr Asn Pro Glu Arg Asp Thr Gly
385 390 395 400

Gln Gly Asn Leu Ile Tyr Val Thr Glu Asn Phe Arg Gly Gln His Trp
405 410 415

Asp Pro Pro Ser Ser Asp Asn Leu Lys Leu Asp Gly Phe Pro Leu Tyr
420 425 430

Asp Met Cys Trp Gly Phe Ile Asp Trp Ile Glu Lys Val His Glu Thr
435 440 445

Glu Asn Leu Leu Thr Asn Tyr Cys Phe Cys Ile Arg Ser Ser Ala Phe
450 455 460

Asn Glu Lys Lys Thr Val Phe Ile Pro Val Asp His Ser Phe Leu Thr
465 470 475 480

Gly Phe Ser Pro Tyr Glu Thr Pro Val Lys Ser Ser Asp Gln Ala His
485 490 495

Trp His Pro Gln Ile Arg Phe Gln Thr Lys Ser Ile Asn Asp Ile Cys
500 505 510

Leu Thr Gly Pro Gly Cys Ala Arg Ser Pro Tyr Gly Asn Tyr Met Gln
515 520 525

Ala Lys Met Ser Tyr Lys Phe His Val Lys Trp Gly Gly Cys Pro Lys
530 535 540

Thr Tyr Glu Lys Pro Tyr Asp Pro Cys Ser Gln Pro Asn Trp Thr Ile
545 550 555 560

Pro His Asn Leu Asn Glu Thr Ile Gln Ile Gln Asn Pro Asn Thr Cys
565 570 575

Pro Gln Thr Glu Leu Gln Glu Trp Asp Trp Arg Arg Asp Ile Val Thr
580 585 590

Lys Lys Ala Ile Glu Arg Ile Arg Gln His Thr Glu Pro His Glu Thr
595 600 605

Leu Gln Ile Ser Thr Gly Ser Lys His Asn Pro Pro Val His Arg Gln
610 615 620

Thr Ser Pro Trp Thr Asp Ser Glu Thr Asp Ser Glu Glu Glu Lys Asp
625 630 635 640

Gln Thr Gln Glu Ile Gln Ile Gln Leu Asn Lys Leu Arg Lys His Gln
645 650 655

Gln His Leu Lys Gln Gln Leu Lys Gln Tyr Leu Lys Pro Gln Asn Ile
660 665 670

Glu

<210> 15
<211> 673
<212> PRT
<213> 人工序列

<220>
<223> 人工序列描述：合成多肽

<400> 15

```
Met Gly Ser Ser His His His His His His Gly Ser Asp Tyr Lys Asp
1               5                  10                  15

Asp Asp Asp Lys Ser Gly Ser Leu Glu Val Leu Phe Gln Gly Pro Ser
            20                  25                  30

Gly Met Pro Tyr Tyr Tyr Arg Arg Arg Arg Tyr Asn Tyr Arg Arg Pro
        35                  40                  45

Arg Trp Tyr Gly Arg Gly Trp Ile Arg Arg Pro Phe Arg Arg Arg Phe
    50                  55                  60

Arg Arg Lys Arg Arg Val Arg Pro Thr Tyr Thr Thr Ile Pro Leu Lys
65                  70                  75                  80

Gln Trp Gln Pro Pro Tyr Lys Arg Thr Cys Tyr Ile Lys Gly Gln Asp
                85                  90                  95

Cys Leu Ile Tyr Tyr Ser Asn Leu Arg Leu Gly Met Asn Ser Thr Met
            100                 105                 110

Tyr Glu Lys Ser Ile Val Pro Val His Trp Pro Gly Gly Gly Ser Phe
        115                 120                 125

Ser Val Ser Met Leu Thr Leu Asp Ala Leu Tyr Asp Ile His Lys Leu
    130                 135                 140

Cys Arg Asn Trp Trp Thr Ser Thr Asn Gln Asp Leu Pro Leu Val Arg
145                 150                 155                 160

Tyr Lys Gly Cys Lys Ile Thr Phe Tyr Gln Ser Thr Phe Thr Asp Tyr
```

165 170 175

Ile Val Arg Ile His Thr Glu Leu Pro Ala Asn Ser Asn Lys Leu Thr
180 185 190

Tyr Pro Asn Thr His Pro Leu Met Met Met Met Ser Lys Tyr Lys His
195 200 205

Ile Ile Pro Ser Arg Gln Thr Arg Arg Lys Lys Lys Pro Tyr Thr Lys
210 215 220

Ile Phe Val Lys Pro Pro Pro Gln Phe Glu Asn Lys Trp Tyr Phe Ala
225 230 235 240

Thr Asp Leu Tyr Lys Ile Pro Leu Leu Gln Ile His Cys Thr Ala Cys
245 250 255

Asn Leu Gln Asn Pro Phe Val Lys Pro Asp Lys Leu Ser Asn Asn Leu
260 265 270

Thr Leu Thr Cys Leu Asn Pro Leu Leu Phe Gln Asn Gln Asp Phe Asp
275 280 285

His Pro Ser Asp Thr Gln Gly Tyr Phe Pro Lys Pro Gly Val Tyr Leu
290 295 300

Tyr Ser Thr Gln Arg Ser Asn Lys Pro Ser Ser Ser Asp Cys Ile Tyr
305 310 315 320

Leu Gly Asn Thr Lys Asp Asn Gln Glu Gly Lys Ser Ala Ser Ser Leu
325 330 335

Met Thr Leu Lys Thr Gln Lys Ile Thr Asp Trp Gly Asn Pro Phe Trp
340 345 350

His Tyr Tyr Ile Asp Gly Ser Lys Lys Ile Phe Ser Tyr Phe Lys Pro
355 360 365

Pro Ser Gln Leu Asp Ser Ser Asp Phe Glu His Met Thr Glu Leu Ala
370 375 380

Glu Pro Met Phe Ile Gln Val Arg Tyr Asn Pro Glu Arg Asp Thr Gly
385 390 395 400

Gln Gly Asn Leu Ile Tyr Val Thr Glu Asn Phe Arg Gly Gln His Trp
405 410 415

Asp Pro Pro Ser Ser Asp Asn Leu Lys Leu Asp Gly Phe Pro Leu Tyr
420 425 430

Asp Met Cys Trp Gly Phe Ile Asp Trp Ile Glu Lys Val His Glu Thr
435 440 445

Glu Asn Leu Leu Thr Asn Tyr Cys Phe Cys Ile Arg Ser Ser Ala Phe
450 455 460

Asn Glu Lys Lys Thr Val Phe Ile Pro Val Asp His Ser Phe Leu Thr
465 470 475 480

Gly Phe Ser Pro Tyr Glu Thr Pro Val Lys Ser Ser Asp Gln Ala His
485 490 495

Trp His Pro Gln Ile Arg Phe Gln Thr Lys Ser Ile Asn Asp Ile Cys
500 505 510

Leu Thr Gly Pro Gly Cys Ala Arg Ser Pro Tyr Gly Asn Tyr Met Gln
515 520 525

Ala Lys Met Ser Tyr Lys Phe His Val Lys Trp Gly Gly Cys Pro Lys
530 535 540

Thr Tyr Glu Lys Pro Tyr Asp Pro Cys Ser Gln Pro Asn Trp Thr Ile
545 550 555 560

Pro His Asn Leu Asn Glu Thr Ile Gln Ile Gln Asn Pro Asn Thr Cys

565 570 575

Pro Gln Thr Glu Leu Gln Glu Trp Asp Trp Arg Arg Asp Ile Val Thr
580 585 590

Lys Lys Ala Ile Glu Arg Ile Arg Gln His Thr Glu Pro His Glu Thr
595 600 605

Leu Gln Ile Ser Thr Gly Ser Lys His Asn Pro Pro Val His Arg Gln
610 615 620

Thr Ser Pro Trp Thr Asp Ser Glu Thr Asp Ser Glu Glu Glu Lys Asp
625 630 635 640

Gln Thr Gln Glu Ile Gln Ile Gln Leu Asn Lys Leu Arg Lys His Gln
645 650 655

Gln His Leu Lys Gln Gln Leu Lys Gln Tyr Leu Lys Pro Gln Asn Ile
660 665 670

Glu

<210> 16
<211> 3753
<212> DNA
<213> α細環病毒屬

<400> 16
tgctacgtca ctaacccacg tgtcctctac aggccaatcg cagtctatgt cgtgcacttc 60

ctgggcatgg tctacataat tatataaatg cttgcacttc cgaatggctg agtttttgct 120

gcccgtccgc ggagaggagc cacggcaggg gatccgaacg tcctgagggc gggtgccgga 180

ggtgagttta cacaccgaag tcaaggggca attcgggctc aggactggcc gggctttggg 240

caaggctctt aaaaatgcac ttttctcgaa taagcagaaa gaaaaggaaa gtgctactgc 300

tttgcgtgcc agcagctaag aaaaaaccaa ctgctatgag cttctggaaa cctccggtac 360

acaatgtcac ggggatccaa cgcatgtggt atgagtcctt tcaccgtggc cacgcttctt 420

tttgtggttg tgggaatcct atacttcaca ttactgcact tgctgaaaca tatggccatc 480

caacaggccc gagaccttct gggccaccgg gagtagaccc caacccccac atccgtagag 540

ccaggcctgc cccggccgct ccggagccct cacaggttga ttcgagacca gccctgacat 600

ggcatgggga tggtggaagc gacggaggcg ctggtggttc cggaagcggt ggacccgtgg 660

cagacttcgc agacgatggc ctcgatcagc tcgtcgccgc cctagacgac gaagagtaag 720

gaggcgcaga cggtggagga gggggagacg aaaaacaagg acttacagac gcaggagacg 780

ctttagacgc aggggacgaa aagcaaaact tataataaaa ctgtggcaac ctgcagtaat 840

taaaagatgc agaataaagg gatacatacc actgattata agtgggaacg gtacctttgc 900

cacaaacttt accagtcaca taaatgacag aataatgaaa ggccccttcg ggggaggaca 960

cagcactatg aggttcagcc tctacatttt gtttgaggag cacctcagac acatgaactt 1020

ctggaccaga agcaacgata acctagagct aaccagatac ttgggggctt cagtaaaaat 1080

atacaggcac ccagaccaag actttatagt aatatacaac agaagaaccc ctctaggagg 1140

caacatctac acagcaccct ctctacaccc aggcaatgcc attttagcaa aacacaaaat 1200

attagtacca agtttacaga caagaccaaa gggtagaaaa gcaattagac taagaatagc 1260

acccccccaca ctctttacag acaagtggta ctttcaaaag gacatagccg acctcaccct 1320

tttcaacatc atggcagttg aggctgactt gcggtttccg ttctgctcac cacaaactga 1380

caacacttgc atcagcttcc aggtccttag ttccgtttac aacaactacc tcagtattaa 1440

tacctttaat aatgacaact cagactcaaa gttaaaagaa tttttaaata aagcatttcc 1500

aacaacaggc acaaaaggaa caagtttaaa tgcactaaat acatttagaa cagaaggatg 1560

cataagtcac ccacaactaa aaaaaccaaa cccacaaata aacaaaccat tagagtcaca 1620

atactttgca cctttagatg ccctctgggg agaccccata tactataatg atctaaatga 1680

aaacaaaagt ttgaacgata tcattgagaa aatactaata aaaaacatga ttacatacca 1740

tgcaaaacta agagaatttc caaattcata ccaaggaaac aaggcctttt gccacctaac 1800

aggcatatac agcccaccat acctaaacca aggcagaata tctccagaaa tatttggact 1860

gtacacagaa ataatttaca acccttacac agacaaagga actggaaaca aagtatggat 1920

ggacccacta actaaagaga acaacatata taaagaagga cagagcaaat gcctactgac 1980

tgacatgccc ctatggactt tacttttgg atatacagac tggtgtaaaa aggacactaa 2040

taactgggac ttaccactaa actacagact agtactaata tgcccttata cctttccaaa 2100

attgtacaat gaaaaagtaa aagactatgg gtacatcccg tactcctaca aattcggagc 2160

gggtcagatg ccagacggca gcaactacat acccctttcag tttagagcaa agtggtaccc 2220

cacagtacta caccagcaac aggtaatgga ggacataagc aggagcgggc cctttgcacc 2280

taaggtagaa aaaccaagca ctcagctggt aatgaagtac tgttttaact ttaactgggg 2340

cggtaaccct atcattgaac agattgttaa agaccccagc ttccagccca cctatgaaat 2400

acccggtacc ggtaacatcc ctagaagaat acaagtcatc gacccgcggg tcctgggacc 2460

gcactactcg ttccggtcat gggacatgcg cagacacaca tttagcagag caagtattaa 2520

gagagtgtca gaacaacaag aaacttctga ccttgtattc tcaggcccaa aaaagcctcg 2580

ggtcgacatc ccaaaacaag aaacccaaga agaaagctca cattcactcc aaagagaatc 2640

gagaccgtgg gagaccgagg aagaaagcga gacagaagcc ctctcgcaag agagccaaga 2700

ggtccccttc caacagcagt tgcagcagca gtaccaagag cagctcaagc tcagacaggg 2760

aatcaaagtc ctcttcgagc agctcataag gacccaacaa ggggtccatg taaacccatg 2820

cctacggtag gtcccaggca gtggctgttt ccagagagaa agccagcccc agctcctagc 2880

agtggagact gggccatgga gtttctcgca gcaaaaatat ttgataggcc agttagaagc 2940

aaccttaaag atacccctta ctacccatat gttaaaaacc aatacaatgt ctactttgac 3000

cttaaatttg aataaacagc agcttcaaac ttgcaaggcc gtgggagttt cactggtcgg 3060

tgtctacctc taaaggtcac taagcactcc gagcgtaagc gaggagtgcg accctccccc 3120

ctggaacaac ttcttcggag tccggcgcta cgccttcggc tgcgccggac acctcagacc 3180

ccccctccac ccgaaacgct tgcgcgtttc ggaccttcgg cgtcgggggg gtcgggagct 3240

ttattaaacg gactccgaag tgctcttgga cactgagggg gtgaacagca acgaaagtga 3300

gtggggccag acttcgccat aaggccttta tcttcttgcc atttgtcagt gtccggggtc 3360

gccataggct tcgggctcgt ttttaggcct tccggactac aaaaatcgcc attttggtga 3420

cgtcacggcc gccatcttaa gtagttgagg cggacggtgg cgtgagttca aaggtcacca 3480

tcagccacac ctactcaaaa tggtggacaa tttcttccgg gtcaaaggtt acagccgcca 3540

tgttaaaaca cgtgacgtat gacgtcacgg ccgccatttt gtgacacaag atggccgact 3600

tccttcctct ttttcaaaaa aaagcggaag tgccgccgcg gcggcggggg gcggcgcgct 3660

gcgcgcgccg cccagtaggg ggagccatgc gccccccccc gcgcatgcgc ggggcccccc 3720

cccgcggggg gctccgcccc ccggcccccc ccg 3753

<210> 17
<211> 127
<212> PRT
<213> α細環病毒屬

<400> 17
Met Ser Phe Trp Lys Pro Pro Val His Asn Val Thr Gly Ile Gln Arg
1 5 10 15

Met Trp Tyr Glu Ser Phe His Arg Gly His Ala Ser Phe Cys Gly Cys
20 25 30

Gly Asn Pro Ile Leu His Ile Thr Ala Leu Ala Glu Thr Tyr Gly His
35 40 45

Pro Thr Gly Pro Arg Pro Ser Gly Pro Pro Gly Val Asp Pro Asn Pro
50 55 60

His Ile Arg Arg Ala Arg Pro Ala Pro Ala Ala Pro Glu Pro Ser Gln
65 70 75 80

Val Asp Ser Arg Pro Ala Leu Thr Trp His Gly Asp Gly Gly Ser Asp
85 90 95

Gly Gly Ala Gly Gly Ser Gly Ser Gly Gly Pro Val Ala Asp Phe Ala
100 105 110

Asp Asp Gly Leu Asp Gln Leu Val Ala Ala Leu Asp Asp Glu Glu
115 120 125

<210> 18
<211> 268
<212> PRT
<213> α細環病毒屬

<400> 18

```
Met Ser Phe Trp Lys Pro Pro Val His Asn Val Thr Gly Ile Gln Arg
1               5                  10                  15

Met Trp Tyr Glu Ser Phe His Arg Gly His Ala Ser Phe Cys Gly Cys
            20                  25                  30

Gly Asn Pro Ile Leu His Ile Thr Ala Leu Ala Glu Thr Tyr Gly His
        35                  40                  45

Pro Thr Gly Pro Arg Pro Ser Gly Pro Pro Gly Val Asp Pro Asn Pro
    50                  55                  60

His Ile Arg Arg Ala Arg Pro Ala Pro Ala Ala Pro Glu Pro Ser Gln
65                  70                  75                  80

Val Asp Ser Arg Pro Ala Leu Thr Trp His Gly Asp Gly Gly Ser Asp
                85                  90                  95

Gly Gly Ala Gly Gly Ser Gly Ser Gly Gly Pro Val Ala Asp Phe Ala
            100                 105                 110

Asp Asp Gly Leu Asp Gln Leu Val Ala Ala Leu Asp Asp Glu Glu Leu
        115                 120                 125

Leu Lys Thr Pro Ala Ser Ser Pro Pro Met Lys Tyr Pro Val Pro Val
    130                 135                 140

Thr Ser Leu Glu Glu Tyr Lys Ser Ser Thr Arg Gly Ser Trp Asp Arg
145                 150                 155                 160

Thr Thr Arg Ser Gly His Gly Thr Cys Ala Asp Thr His Leu Ala Glu
                165                 170                 175
```

Gln Val Leu Arg Glu Cys Gln Asn Asn Lys Lys Leu Leu Thr Leu Tyr
180 185 190

Ser Gln Ala Gln Lys Ser Leu Gly Ser Thr Ser Gln Asn Lys Lys Pro
195 200 205

Lys Lys Lys Ala His Ile His Ser Lys Glu Asn Arg Asp Arg Gly Arg
210 215 220

Pro Arg Lys Lys Ala Arg Gln Lys Pro Ser Arg Lys Arg Ala Lys Arg
225 230 235 240

Ser Pro Ser Asn Ser Ser Cys Ser Ser Ser Thr Lys Ser Ser Ser Ser
245 250 255

Ser Asp Arg Glu Ser Lys Ser Ser Ser Ser Ser Ser
260 265

<210> 19
<211> 276
<212> PRT
<213> α細環病毒屬

<400> 19
Met Ser Phe Trp Lys Pro Pro Val His Asn Val Thr Gly Ile Gln Arg
1 5 10 15

Met Trp Tyr Glu Ser Phe His Arg Gly His Ala Ser Phe Cys Gly Cys
20 25 30

Gly Asn Pro Ile Leu His Ile Thr Ala Leu Ala Glu Thr Tyr Gly His
35 40 45

Pro Thr Gly Pro Arg Pro Ser Gly Pro Pro Gly Val Asp Pro Asn Pro
50 55 60

His Ile Arg Arg Ala Arg Pro Ala Pro Ala Ala Pro Glu Pro Ser Gln
65 70 75 80

Val Asp Ser Arg Pro Ala Leu Thr Trp His Gly Asp Gly Gly Ser Asp
85 90 95

Gly Gly Ala Gly Gly Ser Gly Ser Gly Gly Pro Val Ala Asp Phe Ala
100 105 110

Asp Asp Gly Leu Asp Gln Leu Val Ala Ala Leu Asp Asp Glu Glu Pro
115 120 125

Lys Lys Ala Ser Gly Arg His Pro Lys Thr Arg Asn Pro Arg Arg Lys
130 135 140

Leu Thr Phe Thr Pro Lys Arg Ile Glu Thr Val Gly Asp Arg Gly Arg
145 150 155 160

Lys Arg Asp Arg Ser Pro Leu Ala Arg Glu Pro Arg Gly Pro Leu Pro
165 170 175

Thr Ala Val Ala Ala Ala Val Pro Arg Ala Ala Gln Ala Gln Thr Gly
180 185 190

Asn Gln Ser Pro Leu Arg Ala Ala His Lys Asp Pro Thr Arg Gly Pro
195 200 205

Cys Lys Pro Met Pro Thr Val Gly Pro Arg Gln Trp Leu Phe Pro Glu
210 215 220

Arg Lys Pro Ala Pro Ala Pro Ser Ser Gly Asp Trp Ala Met Glu Phe
225 230 235 240

Leu Ala Ala Lys Ile Phe Asp Arg Pro Val Arg Ser Asn Leu Lys Asp
245 250 255

Thr Pro Tyr Tyr Pro Tyr Val Lys Asn Gln Tyr Asn Val Tyr Phe Asp
260 265 270

Leu Lys Phe Glu
275

<210> 20
<211> 167
<212> PRT
<213> α細環病毒屬

<400> 20
Met Ser Phe Trp Lys Pro Pro Val His Asn Val Thr Gly Ile Gln Arg
1 5 10 15

Met Trp Pro Lys Lys Ala Ser Gly Arg His Pro Lys Thr Arg Asn Pro
20 25 30

Arg Arg Lys Leu Thr Phe Thr Pro Lys Arg Ile Glu Thr Val Gly Asp
35 40 45

Arg Gly Arg Lys Arg Asp Arg Ser Pro Leu Ala Arg Glu Pro Arg Gly
50 55 60

Pro Leu Pro Thr Ala Val Ala Ala Ala Val Pro Arg Ala Ala Gln Ala
65 70 75 80

Gln Thr Gly Asn Gln Ser Pro Leu Arg Ala Ala His Lys Asp Pro Thr
85 90 95

Arg Gly Pro Cys Lys Pro Met Pro Thr Val Gly Pro Arg Gln Trp Leu
100 105 110

Phe Pro Glu Arg Lys Pro Ala Pro Ala Pro Ser Ser Gly Asp Trp Ala
115 120 125

Met Glu Phe Leu Ala Ala Lys Ile Phe Asp Arg Pro Val Arg Ser Asn
130 135 140

Leu Lys Asp Thr Pro Tyr Tyr Pro Tyr Val Lys Asn Gln Tyr Asn Val
145 150 155 160

Tyr Phe Asp Leu Lys Phe Glu
165

<210> 21
<211> 743
<212> PRT
<213> α細環病毒屬

<400> 21
Met Ala Trp Gly Trp Trp Lys Arg Arg Arg Arg Trp Trp Phe Arg Lys
1 5 10 15

Arg Trp Thr Arg Gly Arg Leu Arg Arg Arg Trp Pro Arg Ser Ala Arg
20 25 30

Arg Arg Pro Arg Arg Arg Arg Val Arg Arg Arg Arg Arg Trp Arg Arg
35 40 45

Gly Arg Arg Lys Thr Arg Thr Tyr Arg Arg Arg Arg Arg Phe Arg Arg
50 55 60

Arg Gly Arg Lys Ala Lys Leu Ile Ile Lys Leu Trp Gln Pro Ala Val
65 70 75 80

Ile Lys Arg Cys Arg Ile Lys Gly Tyr Ile Pro Leu Ile Ile Ser Gly
85 90 95

Asn Gly Thr Phe Ala Thr Asn Phe Thr Ser His Ile Asn Asp Arg Ile
100 105 110

Met Lys Gly Pro Phe Gly Gly Gly His Ser Thr Met Arg Phe Ser Leu
115 120 125

Tyr Ile Leu Phe Glu Glu His Leu Arg His Met Asn Phe Trp Thr Arg
130 135 140

Ser Asn Asp Asn Leu Glu Leu Thr Arg Tyr Leu Gly Ala Ser Val Lys
145 150 155 160

Ile Tyr Arg His Pro Asp Gln Asp Phe Ile Val Ile Tyr Asn Arg Arg
165 170 175

Thr Pro Leu Gly Gly Asn Ile Tyr Thr Ala Pro Ser Leu His Pro Gly
180 185 190

Asn Ala Ile Leu Ala Lys His Lys Ile Leu Val Pro Ser Leu Gln Thr
195 200 205

Arg Pro Lys Gly Arg Lys Ala Ile Arg Leu Arg Ile Ala Pro Pro Thr
210 215 220

Leu Phe Thr Asp Lys Trp Tyr Phe Gln Lys Asp Ile Ala Asp Leu Thr
225 230 235 240

Leu Phe Asn Ile Met Ala Val Glu Ala Asp Leu Arg Phe Pro Phe Cys
245 250 255

Ser Pro Gln Thr Asp Asn Thr Cys Ile Ser Phe Gln Val Leu Ser Ser
260 265 270

Val Tyr Asn Asn Tyr Leu Ser Ile Asn Thr Phe Asn Asn Asp Asn Ser
275 280 285

Asp Ser Lys Leu Lys Glu Phe Leu Asn Lys Ala Phe Pro Thr Thr Gly
290 295 300

Thr Lys Gly Thr Ser Leu Asn Ala Leu Asn Thr Phe Arg Thr Glu Gly
305 310 315 320

Cys Ile Ser His Pro Gln Leu Lys Lys Pro Asn Pro Gln Ile Asn Lys
325 330 335

Pro Leu Glu Ser Gln Tyr Phe Ala Pro Leu Asp Ala Leu Trp Gly Asp
340 345 350

Pro Ile Tyr Tyr Asn Asp Leu Asn Glu Asn Lys Ser Leu Asn Asp Ile
355 360 365

Ile Glu Lys Ile Leu Ile Lys Asn Met Ile Thr Tyr His Ala Lys Leu
370 375 380

Arg Glu Phe Pro Asn Ser Tyr Gln Gly Asn Lys Ala Phe Cys His Leu
385 390 395 400

Thr Gly Ile Tyr Ser Pro Pro Tyr Leu Asn Gln Gly Arg Ile Ser Pro
405 410 415

Glu Ile Phe Gly Leu Tyr Thr Glu Ile Ile Tyr Asn Pro Tyr Thr Asp
420 425 430

Lys Gly Thr Gly Asn Lys Val Trp Met Asp Pro Leu Thr Lys Glu Asn
435 440 445

Asn Ile Tyr Lys Glu Gly Gln Ser Lys Cys Leu Leu Thr Asp Met Pro
450 455 460

Leu Trp Thr Leu Leu Phe Gly Tyr Thr Asp Trp Cys Lys Lys Asp Thr
465 470 475 480

Asn Asn Trp Asp Leu Pro Leu Asn Tyr Arg Leu Val Leu Ile Cys Pro
485 490 495

Tyr Thr Phe Pro Lys Leu Tyr Asn Glu Lys Val Lys Asp Tyr Gly Tyr
500 505 510

Ile Pro Tyr Ser Tyr Lys Phe Gly Ala Gly Gln Met Pro Asp Gly Ser
515 520 525

Asn Tyr Ile Pro Phe Gln Phe Arg Ala Lys Trp Tyr Pro Thr Val Leu
530 535 540

His Gln Gln Gln Val Met Glu Asp Ile Ser Arg Ser Gly Pro Phe Ala
545 550 555 560

Pro Lys Val Glu Lys Pro Ser Thr Gln Leu Val Met Lys Tyr Cys Phe
565 570 575

Asn Phe Asn Trp Gly Gly Asn Pro Ile Ile Glu Gln Ile Val Lys Asp
580 585 590

Pro Ser Phe Gln Pro Thr Tyr Glu Ile Pro Gly Thr Gly Asn Ile Pro
595 600 605

Arg Arg Ile Gln Val Ile Asp Pro Arg Val Leu Gly Pro His Tyr Ser
610 615 620

Phe Arg Ser Trp Asp Met Arg Arg His Thr Phe Ser Arg Ala Ser Ile
625 630 635 640

Lys Arg Val Ser Glu Gln Gln Glu Thr Ser Asp Leu Val Phe Ser Gly
645 650 655

Pro Lys Lys Pro Arg Val Asp Ile Pro Lys Gln Glu Thr Gln Glu Glu
660 665 670

Ser Ser His Ser Leu Gln Arg Glu Ser Arg Pro Trp Glu Thr Glu Glu
675 680 685

Glu Ser Glu Thr Glu Ala Leu Ser Gln Glu Ser Gln Glu Val Pro Phe
690 695 700

Gln Gln Gln Leu Gln Gln Gln Tyr Gln Glu Gln Leu Lys Leu Arg Gln
705 710 715 720

Gly Ile Lys Val Leu Phe Glu Gln Leu Ile Arg Thr Gln Gln Gly Val
725 730 735

His Val Asn Pro Cys Leu Arg
740

<210> 22
<211> 194

<212> PRT
<213> α細環病毒屬

<400> 22

```
Met Ala Trp Gly Trp Trp Lys Arg Arg Arg Arg Trp Trp Phe Arg Lys
1               5                  10                  15

Arg Trp Thr Arg Gly Arg Leu Arg Arg Arg Trp Pro Arg Ser Ala Arg
            20                  25                  30

Arg Arg Pro Arg Arg Arg Arg Ile Val Lys Asp Pro Ser Phe Gln Pro
        35                  40                  45

Thr Tyr Glu Ile Pro Gly Thr Gly Asn Ile Pro Arg Arg Ile Gln Val
    50                  55                  60

Ile Asp Pro Arg Val Leu Gly Pro His Tyr Ser Phe Arg Ser Trp Asp
65                  70                  75                  80

Met Arg Arg His Thr Phe Ser Arg Ala Ser Ile Lys Arg Val Ser Glu
                85                  90                  95

Gln Gln Glu Thr Ser Asp Leu Val Phe Ser Gly Pro Lys Lys Pro Arg
            100                 105                 110

Val Asp Ile Pro Lys Gln Glu Thr Gln Glu Glu Ser Ser His Ser Leu
        115                 120                 125

Gln Arg Glu Ser Arg Pro Trp Glu Thr Glu Glu Glu Ser Glu Thr Glu
    130                 135                 140

Ala Leu Ser Gln Glu Ser Gln Glu Val Pro Phe Gln Gln Gln Leu Gln
145                 150                 155                 160

Gln Gln Tyr Gln Glu Gln Leu Lys Leu Arg Gln Gly Ile Lys Val Leu
                165                 170                 175

Phe Glu Gln Leu Ile Arg Thr Gln Gln Gly Val His Val Asn Pro Cys
            180                 185                 190
```

Leu Arg

<210> 23
<211> 113
<212> PRT
<213> α細環病毒屬

<400> 23
Met Ala Trp Gly Trp Trp Lys Arg Arg Arg Arg Trp Trp Phe Arg Lys
1 5 10 15

Arg Trp Thr Arg Gly Arg Leu Arg Arg Arg Trp Pro Arg Ser Ala Arg
20 25 30

Arg Arg Pro Arg Arg Arg Arg Ala Gln Lys Ser Leu Gly Ser Thr Ser
35 40 45

Gln Asn Lys Lys Pro Lys Lys Lys Ala His Ile His Ser Lys Glu Asn
50 55 60

Arg Asp Arg Gly Arg Pro Arg Lys Lys Ala Arg Gln Lys Pro Ser Arg
65 70 75 80

Lys Arg Ala Lys Arg Ser Pro Ser Asn Ser Ser Cys Ser Ser Ser Thr
85 90 95

Lys Ser Ser Ser Ser Ser Asp Arg Glu Ser Lys Ser Ser Ser Ser Ser
100 105 110

Ser

<210> 24

<400> 24
000

<210> 25

<400> 25
000

<210> 26

<400> 26
000

<210> 27

<400> 27
000

<210> 28

<400> 28
000

<210> 29

<400> 29
000

<210> 30

<400> 30
000

<210> 31

<400> 31
000

<210> 32

<400> 32
000

<210> 33

<400> 33
000

<210> 34

<400> 34
000

<210> 35

<400> 35
000

<210> 36

<400> 36
000

<210> 37

<400> 37
000

<210> 38

<400> 38
000

<210> 39

<400> 39
000

<210> 40

<400> 40
000

<210> 41

<400> 41
000

<210> 42

<400> 42
000

<210> 43

<400> 43
000

<210> 44

<400> 44
000

<210> 45

<400> 45
000

<210> 46

<400> 46
000

<210> 47

<400> 47
000

<210> 48

<400> 48
000

<210> 49

<400> 49
000

<210> 50

<400> 50
000

<210> 51

<400> 51
000

<210> 52

<400> 52
000

<210> 53

<400> 53
000

<210> 54
<211> 2979
<212> DNA
<213> β細環病毒屬

<400> 54
taataaatat tcaacaggaa aaccacctaa tttaaattgc cgaccacaaa ccgtcactta 60

gttccccttt ttgcaacaac ttctgctttt ttccaactgc cggaaaacca cataatttgc 120

atggctaacc acaaactgat atgctaatta acttccacaa aacaacttcc ccttttaaaa 180

ccacacctac aaattaatta ttaaacacag tcacatcctg ggaggtacta ccacactata 240

ataccaagtg cacttccgaa tggctgagtt tatgccgcta gacggagaac gcatcagtta 300

ctgactgcgg actgaacttg ggcgggtgcc gaaggtgagt gaaaccaccg aagtcaaggg 360

gcaattcggg ctagttcagt ctagcggaac gggcaagaaa cttaaaatta ttttatttt 420

cagatgagcg actgctttaa accaacatgc tacaacaaca aaacaaagca aactcactgg 480

attaataacc tgcatttaac ccacgacctg atctgcttct gcccaacacc aactagacac 540

ttattactag ctttagcaga acaacaagaa acaattgaag tgtctaaaca agaaaaagaa 600

aaaataacaa gatgccttat tactacagaa gaagacggta caactacaga cgtcctagat 660

ggtatggacg aggttggatt agacgccctt ttcgcagaag atttcgaaga aaaagaaggg 720

taagacctac ttatactact attcctctaa agcaatggca accgccatat aaaagaacat 780

gctatataaa aggacaagac tgtttaatat actatagcaa cttaagactg ggaatgaata 840

gtacaatgta tgaaaaaagt attgtacctg tacattggcc gggaggggggt tctttttctg 900

taagcatgtt aactttagat gccttgtatg atatacataa actttgtaga aactggtgga 960

catccacaaa ccaagactta ccactagtaa gatataaagg atgcaaaata acattttatc 1020

aaagcacatt tacagactac atagtaagaa tacatacaga actaccagct aacagtaaca 1080

aactaacata cccaaacaca catccactaa tgatgatgat gtctaagtac aaacacatta 1140

tacctagtag acaaacaaga agaaaaaaga aaccatacac aaaaatattt gtaaaaccac 1200

ctccgcaatt tgaaaacaaa tggtactttg ctacagacct ctacaaaatt ccattactac 1260

aaatacactg cacagcatgc aacttacaaa acccatttgt aaaaccagac aaattatcaa 1320

acaatgttac attatggtca ctaaacacca taagcataca aaatagaaac atgtcagtgg 1380

atcaaggaca atcatggcca tttaaaatac taggaacaca aagcttttat ttttactttt 1440

acaccggagc aaacctacca ggtgacacaa cacaaatacc agtagcagac ctattaccac 1500

taacaaaccc aagaataaac agaccaggac aatcactaaa tgaggcaaaa attacagacc 1560

atattacttt cacagaatac aaaaacaaat ttacaaatta ttggggtaac ccatttaata 1620

aacacattca agaacaccta gatatgatac tatactcact aaaaagtcca gaagcaataa 1680

aaaacgaatg gacaacagaa aacatgaaat ggaaccaatt aaacaatgca ggaacaatgg 1740

cattaacacc atttaacgag ccaatattca cacaaataca atataaccca gatagagaca 1800

caggagaaga cactcaatta tacctactct ctaacgctac aggaacagga tgggacccac 1860

caggaattcc agaattaata ctagaaggat ttccactatg gttaatatat tggggatttg 1920

cagactttca aaaaaaccta aaaaaagtaa caaacataga cacaaattac atgttagtag 1980

caaaaacaaa atttacacaa aaacctggca cattctactt agtaatacta aatgacacct 2040

ttgtagaagg caatagccca tatgaaaaac aacctttacc tgaagacaac attaaatggt 2100

acccacaagt acaataccaa ttagaagcac aaaacaaact actacaaact gggccattta 2160

caccaaacat acaaggacaa ctatcagaca atatatcaat gttttataaa ttttacttta 2220

aatggggagg aagcccacca aaagcaatta atgttgaaaa tcctgcccac cagattcaat 2280

atcccatacc ccgtaacgag catgaaacaa cttcgttaca gagtccaggg gaagccccag 2340

aatccatctt atactccttc gactatagac acgggaacta cacaacaaca gctttgtcac 2400

gaattagcca agactgggca cttaaagaca ctgtttctaa aattacagag ccagatcgac 2460

agcaactgct caaacaagcc ctcgaatgcc tgcaaatctc ggaagaaacg caggagaaaa 2520

aagaaaaaga agtacagcag ctcatcagca acctcagaca gcagcagcag ctgtacagag 2580

agcgaataat atcattatta aaggaccaat aacttttaac tgtgtaaaaa aggtgaaatt 2640

gtttgatgat aaaccaaaaa accgtagatt tacacctgag gaatttgaaa ctgagttaca 2700

aatagcaaaa tggttaaaga gacccccaag atcctttgta aatgatcctc ccttttaccc 2760

atggttacca cctgaacctg ttgtaaactt taagcttaat tttactgaat aaaggccagc 2820

attaattcac ttaaggagtc tgtttattta agttaaacct taataaacgg tcaccgcctc 2880

cctaatacgc aggcgcagaa agggggctcc gccccttta acccccaggg ggctccgccc 2940

cctgaaaccc ccaaggggggc tacgcccccct tacaccccc 2979

<210> 55
<211> 99
<212> PRT
<213> β細環病毒屬

<400> 55

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Met | Ser | Asp | Cys | Phe | Lys | Pro | Thr | Cys | Tyr | Asn | Asn | Lys | Thr | Lys | Gln |
| 1 | | | | | 5 | | | | 10 | | | | | 15 | |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Thr | His | Trp | Ile | Asn | Asn | Leu | His | Leu | Thr | His | Asp | Leu | Ile | Cys | Phe |
| | | | 20 | | | | | 25 | | | | | 30 | | |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Cys | Pro | Thr | Pro | Thr | Arg | His | Leu | Leu | Leu | Ala | Leu | Ala | Glu | Gln | Gln |
| | | 35 | | | | | 40 | | | | | 45 | | | |

Glu Thr Ile Glu Val Ser Lys Gln Glu Lys Glu Lys Ile Thr Arg Cys

50 55 60

Leu Ile Thr Thr Glu Glu Asp Gly Thr Thr Thr Asp Val Leu Asp Gly
65 70 75 80

Met Asp Glu Val Gly Leu Asp Ala Leu Phe Ala Glu Asp Phe Glu Glu
85 90 95

Lys Glu Gly

<210> 56
<211> 203
<212> PRT
<213> β細環病毒屬

<400> 56
Met Ser Asp Cys Phe Lys Pro Thr Cys Tyr Asn Asn Lys Thr Lys Gln
1 5 10 15

Thr His Trp Ile Asn Asn Leu His Leu Thr His Asp Leu Ile Cys Phe
20 25 30

Cys Pro Thr Pro Thr Arg His Leu Leu Leu Ala Leu Ala Glu Gln Gln
35 40 45

Glu Thr Ile Glu Val Ser Lys Gln Glu Lys Glu Lys Ile Thr Arg Cys
50 55 60

Leu Ile Thr Thr Glu Glu Asp Gly Thr Thr Thr Asp Val Leu Asp Gly
65 70 75 80

Met Asp Glu Val Gly Leu Asp Ala Leu Phe Ala Glu Asp Phe Glu Glu
85 90 95

Lys Glu Gly Phe Asn Ile Pro Tyr Pro Val Thr Ser Met Lys Gln Leu
100 105 110

Arg Tyr Arg Val Gln Gly Lys Pro Gln Asn Pro Ser Tyr Thr Pro Ser

115 120 125

Thr Ile Asp Thr Gly Thr Thr Gln Gln Gln Leu Cys His Glu Leu Ala
130 135 140

Lys Thr Gly His Leu Lys Thr Leu Phe Leu Lys Leu Gln Ser Gln Ile
145 150 155 160

Asp Ser Asn Cys Ser Asn Lys Pro Ser Asn Ala Cys Lys Ser Arg Lys
165 170 175

Lys Arg Arg Arg Lys Lys Lys Lys Lys Tyr Ser Ser Ser Ser Ala Thr
180 185 190

Ser Asp Ser Ser Ser Ser Cys Thr Glu Ser Glu
195 200

<210> 57
<211> 219
<212> PRT
<213> β細環病毒屬

<400> 57
Met Ser Asp Cys Phe Lys Pro Thr Cys Tyr Asn Asn Lys Thr Lys Gln
1 5 10 15

Thr His Trp Ile Asn Asn Leu His Leu Thr His Asp Leu Ile Cys Phe
20 25 30

Cys Pro Thr Pro Thr Arg His Leu Leu Leu Ala Leu Ala Glu Gln Gln
35 40 45

Glu Thr Ile Glu Val Ser Lys Gln Glu Lys Glu Lys Ile Thr Arg Cys
50 55 60

Leu Ile Thr Thr Glu Glu Asp Gly Thr Thr Thr Asp Val Leu Asp Gly
65 70 75 80

Met Asp Glu Val Gly Leu Asp Ala Leu Phe Ala Glu Asp Phe Glu Glu

85 90 95

Lys Glu Gly Ala Arg Ser Thr Ala Thr Ala Gln Thr Ser Pro Arg Met
100 105 110

Pro Ala Asn Leu Gly Arg Asn Ala Gly Glu Lys Arg Lys Arg Ser Thr
115 120 125

Ala Ala His Gln Gln Pro Gln Thr Ala Ala Ala Ala Val Gln Arg Ala
130 135 140

Asn Asn Ile Ile Ile Lys Gly Pro Ile Thr Phe Asn Cys Val Lys Lys
145 150 155 160

Val Lys Leu Phe Asp Asp Lys Pro Lys Asn Arg Arg Phe Thr Pro Glu
165 170 175

Glu Phe Glu Thr Glu Leu Gln Ile Ala Lys Trp Leu Lys Arg Pro Pro
180 185 190

Arg Ser Phe Val Asn Asp Pro Pro Phe Tyr Pro Trp Leu Pro Pro Glu
195 200 205

Pro Val Val Asn Phe Lys Leu Asn Phe Thr Glu
210 215

<210> 58
<211> 666
<212> PRT
<213> β細環病毒屬

<400> 58
Met Pro Tyr Tyr Tyr Arg Arg Arg Arg Tyr Asn Tyr Arg Arg Pro Arg
1 5 10 15

Trp Tyr Gly Arg Gly Trp Ile Arg Arg Pro Phe Arg Arg Arg Phe Arg
20 25 30

Arg Lys Arg Arg Val Arg Pro Thr Tyr Thr Thr Ile Pro Leu Lys Gln

35 40 45

Trp Gln Pro Pro Tyr Lys Arg Thr Cys Tyr Ile Lys Gly Gln Asp Cys
50 55 60

Leu Ile Tyr Tyr Ser Asn Leu Arg Leu Gly Met Asn Ser Thr Met Tyr
65 70 75 80

Glu Lys Ser Ile Val Pro Val His Trp Pro Gly Gly Gly Ser Phe Ser
85 90 95

Val Ser Met Leu Thr Leu Asp Ala Leu Tyr Asp Ile His Lys Leu Cys
100 105 110

Arg Asn Trp Trp Thr Ser Thr Asn Gln Asp Leu Pro Leu Val Arg Tyr
115 120 125

Lys Gly Cys Lys Ile Thr Phe Tyr Gln Ser Thr Phe Thr Asp Tyr Ile
130 135 140

Val Arg Ile His Thr Glu Leu Pro Ala Asn Ser Asn Lys Leu Thr Tyr
145 150 155 160

Pro Asn Thr His Pro Leu Met Met Met Met Ser Lys Tyr Lys His Ile
165 170 175

Ile Pro Ser Arg Gln Thr Arg Arg Lys Lys Lys Pro Tyr Thr Lys Ile
180 185 190

Phe Val Lys Pro Pro Pro Gln Phe Glu Asn Lys Trp Tyr Phe Ala Thr
195 200 205

Asp Leu Tyr Lys Ile Pro Leu Leu Gln Ile His Cys Thr Ala Cys Asn
210 215 220

Leu Gln Asn Pro Phe Val Lys Pro Asp Lys Leu Ser Asn Asn Val Thr
225 230 235 240

Leu Trp Ser Leu Asn Thr Ile Ser Ile Gln Asn Arg Asn Met Ser Val
245 250 255

Asp Gln Gly Gln Ser Trp Pro Phe Lys Ile Leu Gly Thr Gln Ser Phe
260 265 270

Tyr Phe Tyr Phe Tyr Thr Gly Ala Asn Leu Pro Gly Asp Thr Thr Gln
275 280 285

Ile Pro Val Ala Asp Leu Leu Pro Leu Thr Asn Pro Arg Ile Asn Arg
290 295 300

Pro Gly Gln Ser Leu Asn Glu Ala Lys Ile Thr Asp His Ile Thr Phe
305 310 315 320

Thr Glu Tyr Lys Asn Lys Phe Thr Asn Tyr Trp Gly Asn Pro Phe Asn
325 330 335

Lys His Ile Gln Glu His Leu Asp Met Ile Leu Tyr Ser Leu Lys Ser
340 345 350

Pro Glu Ala Ile Lys Asn Glu Trp Thr Thr Glu Asn Met Lys Trp Asn
355 360 365

Gln Leu Asn Asn Ala Gly Thr Met Ala Leu Thr Pro Phe Asn Glu Pro
370 375 380

Ile Phe Thr Gln Ile Gln Tyr Asn Pro Asp Arg Asp Thr Gly Glu Asp
385 390 395 400

Thr Gln Leu Tyr Leu Leu Ser Asn Ala Thr Gly Thr Gly Trp Asp Pro
405 410 415

Pro Gly Ile Pro Glu Leu Ile Leu Glu Gly Phe Pro Leu Trp Leu Ile
420 425 430

Tyr Trp Gly Phe Ala Asp Phe Gln Lys Asn Leu Lys Lys Val Thr Asn

435 440 445

Ile Asp Thr Asn Tyr Met Leu Val Ala Lys Thr Lys Phe Thr Gln Lys
450 455 460

Pro Gly Thr Phe Tyr Leu Val Ile Leu Asn Asp Thr Phe Val Glu Gly
465 470 475 480

Asn Ser Pro Tyr Glu Lys Gln Pro Leu Pro Glu Asp Asn Ile Lys Trp
485 490 495

Tyr Pro Gln Val Gln Tyr Gln Leu Glu Ala Gln Asn Lys Leu Leu Gln
500 505 510

Thr Gly Pro Phe Thr Pro Asn Ile Gln Gly Gln Leu Ser Asp Asn Ile
515 520 525

Ser Met Phe Tyr Lys Phe Tyr Phe Lys Trp Gly Gly Ser Pro Pro Lys
530 535 540

Ala Ile Asn Val Glu Asn Pro Ala His Gln Ile Gln Tyr Pro Ile Pro
545 550 555 560

Arg Asn Glu His Glu Thr Thr Ser Leu Gln Ser Pro Gly Glu Ala Pro
565 570 575

Glu Ser Ile Leu Tyr Ser Phe Asp Tyr Arg His Gly Asn Tyr Thr Thr
580 585 590

Thr Ala Leu Ser Arg Ile Ser Gln Asp Trp Ala Leu Lys Asp Thr Val
595 600 605

Ser Lys Ile Thr Glu Pro Asp Arg Gln Gln Leu Leu Lys Gln Ala Leu
610 615 620

Glu Cys Leu Gln Ile Ser Glu Glu Thr Gln Glu Lys Lys Glu Lys Glu
625 630 635 640

Val Gln Gln Leu Ile Ser Asn Leu Arg Gln Gln Gln Leu Tyr Arg
645 650 655

Glu Arg Ile Ile Ser Leu Leu Lys Asp Gln
660 665

<210> 59
<211> 148
<212> PRT
<213> β細環病毒屬

<400> 59
Met Pro Tyr Tyr Tyr Arg Arg Arg Arg Tyr Asn Tyr Arg Arg Pro Arg
1 5 10 15

Trp Tyr Gly Arg Gly Trp Ile Arg Arg Pro Phe Arg Arg Arg Phe Arg
20 25 30

Arg Lys Arg Arg Ile Gln Tyr Pro Ile Pro Arg Asn Glu His Glu Thr
35 40 45

Thr Ser Leu Gln Ser Pro Gly Glu Ala Pro Glu Ser Ile Leu Tyr Ser
50 55 60

Phe Asp Tyr Arg His Gly Asn Tyr Thr Thr Thr Ala Leu Ser Arg Ile
65 70 75 80

Ser Gln Asp Trp Ala Leu Lys Asp Thr Val Ser Lys Ile Thr Glu Pro
85 90 95

Asp Arg Gln Gln Leu Leu Lys Gln Ala Leu Glu Cys Leu Gln Ile Ser
100 105 110

Glu Glu Thr Gln Glu Lys Lys Glu Lys Glu Val Gln Gln Leu Ile Ser
115 120 125

Asn Leu Arg Gln Gln Gln Gln Leu Tyr Arg Glu Arg Ile Ile Ser Leu
130 135 140

Leu Lys Asp Gln
145

<210> 60
<211> 82
<212> PRT
<213> β細環病毒屬

<400> 60
Met Pro Tyr Tyr Tyr Arg Arg Arg Arg Tyr Asn Tyr Arg Arg Pro Arg
1 5 10 15

Trp Tyr Gly Arg Gly Trp Ile Arg Arg Pro Phe Arg Arg Arg Phe Arg
20 25 30

Arg Lys Arg Arg Ser Gln Ile Asp Ser Asn Cys Ser Asn Lys Pro Ser
35 40 45

Asn Ala Cys Lys Ser Arg Lys Lys Arg Arg Arg Lys Lys Lys Lys Lys
50 55 60

Tyr Ser Ser Ser Ser Ala Thr Ser Asp Ser Ser Ser Ser Cys Thr Glu
65 70 75 80

Ser Glu

<210> 61

<400> 61
000

<210> 62

<400> 62
000

<210> 63

<400> 63

000

<210> 64

<400> 64
000

<210> 65

<400> 65
000

<210> 66

<400> 66
000

<210> 67

<400> 67
000

<210> 68

<400> 68
000

<210> 69

<400> 69
000

<210> 70

<400> 70
000

<210> 71

<400> 71
000

<210> 72

<400> 72
000

<210> 73

<400> 73
000

<210> 74

<400> 74
000

<210> 75

<400> 75
000

<210> 76

<400> 76
000

<210> 77

<400> 77
000

<210> 78

<400> 78
000

<210> 79

<400> 79
000

<210> 80

<400> 80
000

<210> 81

<400> 81
000

<210> 82

<400> 82
000

<210> 83

<400> 83
000

<210> 84

<400> 84
000

<210> 85

<400> 85
000

<210> 86

<400> 86
000

<210> 87

<400> 87
000

<210> 88

<400> 88

000

<210> 89

<400> 89
000

<210> 90

<400> 90
000

<210> 91

<400> 91
000

<210> 92

<400> 92
000

<210> 93

<400> 93
000

<210> 94

<400> 94
000

<210> 95

<400> 95
000

<210> 96

<400> 96
000

<210> 97

<400> 97
000

<210> 98

<400> 98
000

<210> 99

<400> 99
000

<210> 100

<400> 100
000

<210> 101

<400> 101
000

<210> 102

<400> 102
000

<210> 103

<400> 103
000

<210> 104

<400> 104
000

<210> 105

<211> 71
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 105
cgggtgccgk aggtgagttt acacaccgma gtcaaggggc aattcgggct crggactggc 60

cgggcyhtgg g 71

<210> 106
<211> 71
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 106
cgggtgccgg aggtgagttt acacaccgca gtcaaggggc aattcgggct cgggactggc 60

cgggctwtgg g 71

<210> 107
<211> 71
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 107
cgggtgccgt aggtgagttt acacaccgca gtcaaggggc aattcgggct cgggactggc 60

cgggctatgg g 71

<210> 108
<211> 71
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 108

cgggtgccgg aggtgagttt acacaccgca gtcaaggggc aattcgggct cgggactggc 60

cgggcgctgg g 71

<210> 109
<211> 71
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 109
cgggtgccgg aggtgagttt acacaccgca gtcaaggggc aattcgggct cgggactggc 60

cgggctttgg g 71

<210> 110
<211> 71
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 110
cgggtgccgg aggtgagttt acacaccgca gtcaaggggc aattcgggct cgggactggc 60

cgggctatgg g 71

<210> 111
<211> 71
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 111
cgggtgccgg aggtgagttt acacaccgaa gtcaaggggc aattcgggct caggactggc 60

cgggctttgg g 71

<210> 112
<211> 71
<212> DNA

<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 112
cgggtgccgg aggtgagttt acacaccgca gtcaaggggc aattcgggct cgggactggc 60

cgggcyhtgg g 71

<210> 113
<211> 71
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 113
cgggtgccgt aggtgagttt acacaccgca gtcaaggggc aattcgggct cgggactggc 60

cgggctatgg g 71

<210> 114
<211> 70
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 114
cgggtgccgg aggtgagttt acacaccgca gtcaaggggc aattcgggct cgggactggc 60

cgggcccggg 70

<210> 115
<211> 71
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 115
cgggtgccgg aggtgagttt acacaccgaa gtcaaggggc aattcgggct caggactggc 60

cgggctttgg g 71

<210> 116
<211> 69
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 116
cgggtgccgg aggtgagttt acacaccgca gtcaaggggc aattcgggct cgggaggccg 60

ggccatggg 69

<210> 117
<211> 71
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 117
cgggtgccgg aggtgagttt acacaccgca gtcaaggggc aattcgggct cgggactggc 60

cgggccccgg g 71

<210> 118
<211> 71
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 118
cgggtgccgg aggtgagttt acacaccgca gtcaaggggc aattcgggct cgggactggc 60

cgggctatgg g 71

<210> 119
<211> 71
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 119
cgggtgccga aggtgagttt acacaccgca gtcaaggggc aattcgggct cgggactggc 60

cgggctatgg g 71

<210> 120
<211> 117
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成多核苷酸

<220>
<221> misc_feature
<222> (10)..(10)
<223> 可能存在或可能不存在

<220>
<221> misc_feature
<222> (12)..(12)
<223> 可能存在或可能不存在

<220>
<221> misc_feature
<222> (30)..(32)
<223> 可能存在或可能不存在

<220>
<221> misc_feature
<222> (34)..(34)
<223> 可能存在或可能不存在

<220>
<221> misc_feature
<222> (43)..(46)
<223> 可能存在或可能不存在

<220>
<221> misc_feature
<222> (52)..(54)
<223> 可能存在或可能不存在

<220>
<221> misc_feature

<222> (70)..(71)
<223> 可能存在或可能不存在

<220>
<221> misc_feature
<222> (89)..(90)
<223> 可能存在或可能不存在

<220>
<221> misc_feature
<222> (103)..(103)
<223> 可能存在或可能不存在

<400> 120
cggcggsggs gcsscgcgct dcgcgcgcsg cccrsyrggg grdssmmwgc skcsccccccc 60

cscgcgcatg cgcrcgggkc ccccccccyv sgggggggctc cgcccccccg gccccc 117

<210> 121
<211> 169
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成多核苷酸

<220>
<221> modified_base
<222> (20)..(20)
<223> a、c、t、g、未知或其他

<220>
<221> modified_base
<222> (22)..(22)
<223> a、c、t、g、未知或其他

<220>
<221> modified_base
<222> (40)..(42)
<223> a、c、t、g、未知或其他

<220>
<221> modified_base
<222> (53)..(56)
<223> a、c、t、g、未知或其他

<220>
<221> modified_base

<222> (62)..(62)
<223> a、c、t、g、未知或其他

<220>
<221> modified_base
<222> (64)..(64)
<223> a、c、t、g、未知或其他

<220>
<221> modified_base
<222> (97)..(98)
<223> a、c、t、g、未知或其他

<400> 121
gccgccgcgg cggcggsggn gnsgcgcgct dcgcgcgcsn nncrccrggg ggnnnncwgc 60

sncncccccc cccgcgcatg cgcgggkccc ccccccnncg gggggctccg ccccccggcc 120

cccccccgtg ctaaacccac cgcgcatgcg cgaccacgcc cccgccgcc 169

<210> 122
<211> 79
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<220>
<221> modified_base
<222> (20)..(20)
<223> a、c、t、g、未知或其他

<220>
<221> modified_base
<222> (22)..(22)
<223> a、c、t、g、未知或其他

<220>
<221> modified_base
<222> (40)..(42)
<223> a、c、t、g、未知或其他

<220>
<221> modified_base
<222> (53)..(56)
<223> a、c、t、g、未知或其他

<220>
<221> modified_base
<222> (62)..(62)
<223> a、c、t、g、未知或其他

<220>
<221> modified_base
<222> (64)..(64)
<223> a、c、t、g、未知或其他

<400> 122
gccgccgcgg cggcggsggn gnsgcgcgct dcgcgcgcsn nncrccrggg ggnnnncwgc 60

sncncccccc cccgcgcat 79

<210> 123
<211> 31
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<220>
<221> modified_base
<222> (18)..(19)
<223> a、c、t、g、未知或其他

<400> 123
gcgcgggkcc ccccccnnc gggggggctcc g 31

<210> 124
<211> 59
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 124
ccccccggcc cccccccgtg ctaaacccac cgcgcatgcg cgaccacgcc cccgccgcc 59

<210> 125
<211> 156
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成多核苷酸

<400> 125
gcggcgggggg ggcggccgcg ttcgcgcgcc gcccaccagg gggtgctgcg cgccccccccc 60

cgcgcatgcg cggggccccc ccccgggggg gctccgcccc cccggccccc ccccgtgcta 120

aacccaccgc gcatgcgcga ccacgccccc gccgcc 156

<210> 126
<211> 7
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 126
gcggcgg 7

<210> 127
<211> 7
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 127
ggggggcg 7

<210> 128
<211> 6
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 128
gccgcg 6

<210> 129
<211> 25

<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 129
ttcgcgcgcc gcccaccagg gggtg 25

<210> 130
<211> 5
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 130
ctgcg 5

<210> 131
<211> 17
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 131
cgccccccccc cgcgcat 17

<210> 132
<211> 17
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 132
gcgcggggcc ccccccc 17

<210> 133
<211> 72
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 133
ggggggggctc cgccccccccg gccccccccc gtgctaaacc caccgcgcat gcgcgaccac 60

gcccccgccg cc 72

<210> 134
<211> 115
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成多核苷酸

<400> 134
cggcggcggc ggcgcgcgcg ctgcgcgcgc gcgccgggggg ggcgccagcg ccccccccccc 60

cgcgcatgca cgggtccccc cccccacggg gggctccgcc ccccggcccc ccccc 115

<210> 135
<211> 14
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 135
cggcggcggc ggcg 14

<210> 136
<211> 17
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 136
cgcgcgctgc gcgcgcg 17

<210> 137
<211> 19

<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 137
cgccgggggg gcgccagcg 19

<210> 138
<211> 17
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 138
ccccccccccc cgcgcat 17

<210> 139
<211> 31
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 139
gcacgggtcc ccccccccac gggggggctcc g 31

<210> 140
<211> 17
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 140
ccccccggcc ccccccc 17

<210> 141
<211> 121
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成多核苷酸

<400> 141
ccgtcggcgg gggggccgcg cgctgcgcgc gcggcccccg ggggaggcac agcctccccc 60

ccccgcgcgc atgcgcgcgg gtcccccccc ctccgggggg ctccgccccc cggccccccc 120

c 121

<210> 142
<211> 37
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 142
ccgtcggcgg gggggccgcg cgctgcgcgc gcggccc 37

<210> 143
<211> 84
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 143
ccgggggagg cacagcctcc ccccccccgcg cgcatgcgcg cgggtccccc cccctccggg 60

gggctccgcc ccccggcccc cccc 84

<210> 144
<211> 104
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成多核苷酸

<400> 144
cggcggcggc gcgcgcgcta cgcgcgcgcg ccgggggggct gccgcccccc ccccgcgcat 60

gcgcggggcc ccccccccgcg gggggctccg ccccccggcc cccc 104

<210> 145
<211> 11
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 145
cggcggcggc g 11

<210> 146
<211> 17
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 146
cgcgcgctac gcgcgcg 17

<210> 147
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 147
cgccgggggg 10

<210> 148
<211> 7
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 148
ctgccgc 7

<210> 149
<211> 15
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 149
ccccccccg cgcat 15

<210> 150
<211> 17
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 150
gcgcggggcc ccccccc 17

<210> 151
<211> 13
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 151
gcgggggct ccg 13

<210> 152
<211> 14
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 152
ccccccggcc cccc 14

<210> 153
<211> 122

<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成多核苷酸

<400> 153
gccgccgcgg cggcgggggg cggcgcgctg cgcgcgccgc ccagtagggg gagccatgcg 60

cccccccccg cgcatgcgcg gggccccccc ccgcgggggg ctccgccccc cggccccccc 120

cg 122

<210> 154
<211> 19
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 154
gccgccgcgg cggcgggggg 19

<210> 155
<211> 41
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 155
gcggcgcgct gcgcgcgccg cccagtaggg ggagccatgc g 41

<210> 156
<211> 15
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 156
cccccccccg cgcat 15

<210> 157
<211> 17
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 157
gcgcggggcc ccccccc 17

<210> 158
<211> 13
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 158
gcgggggggct ccg 13

<210> 159
<211> 17
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 159
ccccccggcc ccccccg 17

<210> 160
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 160
cgcgctgcgc gcgccgccca gtaggggggag ccatgc 36

<210> 161
<211> 78

<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 161
ccgccatctt aagtagttga ggcggacggt ggcgtgagtt caaaggtcac catcagccac 60

acctactcaa aatggtgg 78

<210> 162
<211> 172
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成多核苷酸

<400> 162
cttaagtagt tgaggcggac ggtggcgtga gttcaaaggt caccatcagc cacacctact 60

caaaatggtg gacaatttct tccgggtcaa aggttacagc cgccatgtta aaacacgtga 120

cgtatgacgt cacggccgcc attttgtgac acaagatggc cgacttcctt cc 172

<210> 163
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 163
cgcgctgcgc gcgccgccca gtaggggggag ccatgc 36

<210> 164
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 164
gcgctdcgcg cgcgcgccgg ggggctgcgc cccccc 36

<210> 165
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 165
gcgcttcgcg cgccgcccac taggggcgt tgcgcg 36

<210> 166
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 166
gcgctgcgcg cgccgcccag taggggcgc aatgcg 36

<210> 167
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 167
gcgctgcgcg cgcggccccc gggggaggca ttgcct 36

<210> 168
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 168
gcgctgcgcg cgcgcgccgg gggggcgcca gcgccc 36

<210> 169
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 169
gcgcttcgcg cgcgcgccgg ggggctccgc cccccc 36

<210> 170
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 170
gcgcttcgcg cgcgcgccgg ggggctgcgc cccccc 36

<210> 171
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 171
gcgctacgcg cgcgcgccgg ggggctgcgc cccccc 36

<210> 172
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<400> 172
gcgctacgcg cgcgcgccgg ggggctctgc cccccc 36

<210> 173

<400> 173
000

<210> 174

<400> 174
000

<210> 175

<400> 175
000

<210> 176

<400> 176
000

<210> 177

<400> 177
000

<210> 178

<400> 178
000

<210> 179

<400> 179
000

<210> 180

<400> 180
000

<210> 181

<400> 181
000

<210> 182

<400> 182
000

<210> 183

<400> 183
000

<210> 184

<400> 184
000

<210> 185
<211> 743
<212> PRT
<213> α細環病毒屬

<400> 185
Met Ala Trp Gly Trp Trp Lys Arg Arg Arg Arg Trp Trp Phe Arg Lys
1 5 10 15

Arg Trp Thr Arg Gly Arg Leu Arg Arg Arg Trp Pro Arg Ser Ala Arg
20 25 30

Arg Arg Pro Arg Arg Arg Arg Val Arg Arg Arg Arg Arg Trp Arg Arg
35 40 45

Gly Arg Arg Lys Thr Arg Thr Tyr Arg Arg Arg Arg Arg Phe Arg Arg
50 55 60

Arg Gly Arg Lys Ala Lys Leu Ile Ile Lys Leu Trp Gln Pro Ala Val
65 70 75 80

Ile Lys Arg Cys Arg Ile Lys Gly Tyr Ile Pro Leu Ile Ile Ser Gly
85 90 95

Asn Gly Thr Phe Ala Thr Asn Phe Thr Ser His Ile Asn Asp Arg Ile
100 105 110

Met Lys Gly Pro Phe Gly Gly Gly His Ser Thr Met Arg Phe Ser Leu
115 120 125

Tyr Ile Leu Phe Glu Glu His Leu Arg His Met Asn Phe Trp Thr Arg
130 135 140

Ser Asn Asp Asn Leu Glu Leu Thr Arg Tyr Leu Gly Ala Ser Val Lys
145 150 155 160

Ile Tyr Arg His Pro Asp Gln Asp Phe Ile Val Ile Tyr Asn Arg Arg
165 170 175

Thr Pro Leu Gly Gly Asn Ile Tyr Thr Ala Pro Ser Leu His Pro Gly
180 185 190

Asn Ala Ile Leu Ala Lys His Lys Ile Leu Val Pro Ser Leu Gln Thr
195 200 205

Arg Pro Lys Gly Arg Lys Ala Ile Arg Leu Arg Ile Ala Pro Pro Thr
210 215 220

Leu Phe Thr Asp Lys Trp Tyr Phe Gln Lys Asp Ile Ala Asp Leu Thr
225 230 235 240

Leu Phe Asn Ile Met Ala Val Glu Ala Asp Leu Arg Phe Pro Phe Cys
245 250 255

Ser Pro Gln Thr Asp Asn Thr Cys Ile Ser Phe Gln Val Leu Ser Ser
260 265 270

Val Tyr Asn Asn Tyr Leu Ser Ile Asn Thr Phe Asn Asn Asp Asn Ser
275 280 285

Asp Ser Lys Leu Lys Glu Phe Leu Asn Lys Ala Phe Pro Thr Thr Gly
290 295 300

Thr Lys Gly Thr Ser Leu Asn Ala Leu Asn Thr Phe Arg Thr Glu Gly
305 310 315 320

Cys Ile Ser His Pro Gln Leu Lys Lys Pro Asn Pro Gln Ile Asn Lys
325 330 335

Pro Leu Glu Ser Gln Tyr Phe Ala Pro Leu Asp Ala Leu Trp Gly Asp
340 345 350

Pro Ile Tyr Tyr Asn Asp Leu Asn Glu Asn Lys Ser Leu Asn Asp Ile
355 360 365

Ile Glu Lys Ile Leu Ile Lys Asn Met Ile Thr Tyr His Ala Lys Leu
370 375 380

Arg Glu Phe Pro Asn Ser Tyr Gln Gly Asn Lys Ala Phe Cys His Leu
385 390 395 400

Thr Gly Ile Tyr Ser Pro Pro Tyr Leu Asn Gln Gly Arg Ile Ser Pro
405 410 415

Glu Ile Phe Gly Leu Tyr Thr Glu Ile Ile Tyr Asn Pro Tyr Thr Asp
420 425 430

Lys Gly Thr Gly Asn Lys Val Trp Met Asp Pro Leu Thr Lys Glu Asn
435 440 445

Asn Ile Tyr Lys Glu Gly Gln Ser Lys Cys Leu Leu Thr Asp Met Pro
450 455 460

Leu Trp Thr Leu Leu Phe Gly Tyr Thr Asp Trp Cys Lys Lys Asp Thr
465 470 475 480

Asn Asn Trp Asp Leu Pro Leu Asn Tyr Arg Leu Val Leu Ile Cys Pro
485 490 495

Tyr Thr Phe Pro Lys Leu Tyr Asn Glu Lys Val Lys Asp Tyr Gly Tyr
500 505 510

Ile Pro Tyr Ser Tyr Lys Phe Gly Ala Gly Gln Met Pro Asp Gly Ser
515 520 525

Asn Tyr Ile Pro Phe Gln Phe Arg Ala Lys Trp Tyr Pro Thr Val Leu
530 535 540

His Gln Gln Gln Val Met Glu Asp Ile Ser Arg Ser Gly Pro Phe Ala
545 550 555 560

Pro Lys Val Glu Lys Pro Ser Thr Gln Leu Val Met Lys Tyr Cys Phe
565 570 575

Asn Phe Asn Trp Gly Gly Asn Pro Ile Ile Glu Gln Ile Val Lys Asp
580 585 590

Pro Ser Phe Gln Pro Thr Tyr Glu Ile Pro Gly Thr Gly Asn Ile Pro
595 600 605

Arg Arg Ile Gln Val Ile Asp Pro Arg Val Leu Gly Pro His Tyr Ser
610 615 620

Phe Arg Ser Trp Asp Met Arg Arg His Thr Phe Ser Arg Ala Ser Ile
625 630 635 640

Lys Arg Val Ser Glu Gln Gln Glu Thr Ser Asp Leu Val Phe Ser Gly
645 650 655

Pro Lys Lys Pro Arg Val Asp Ile Pro Lys Gln Glu Thr Gln Glu Glu
660 665 670

Ser Ser His Ser Leu Gln Arg Glu Ser Arg Pro Trp Glu Thr Glu Glu
675 680 685

Glu Ser Glu Thr Glu Ala Leu Ser Gln Glu Ser Gln Glu Val Pro Phe
690 695 700

Gln Gln Gln Leu Gln Gln Gln Tyr Gln Glu Gln Leu Lys Leu Arg Gln
705 710 715 720

Gly Ile Lys Val Leu Phe Glu Gln Leu Ile Arg Thr Gln Gln Gly Val
725 730 735

His Val Asn Pro Cys Leu Arg
740

<210> 186
<211> 68
<212> PRT
<213> α細環病毒屬

<400> 186
Met Ala Trp Gly Trp Trp Lys Arg Arg Arg Arg Trp Trp Phe Arg Lys
1 5 10 15

Arg Trp Thr Arg Gly Arg Leu Arg Arg Arg Trp Pro Arg Ser Ala Arg
20 25 30

Arg Arg Pro Arg Arg Arg Arg Val Arg Arg Arg Arg Arg Trp Arg Arg
35 40 45

Gly Arg Arg Lys Thr Arg Thr Tyr Arg Arg Arg Arg Arg Phe Arg Arg
50 55 60

Arg Gly Arg Lys
65

<210> 187
<211> 212
<212> PRT
<213> α細環病毒屬

<400> 187
Ala Lys Leu Ile Ile Lys Leu Trp Gln Pro Ala Val Ile Lys Arg Cys
1 5 10 15

Arg Ile Lys Gly Tyr Ile Pro Leu Ile Ile Ser Gly Asn Gly Thr Phe
20 25 30

Ala Thr Asn Phe Thr Ser His Ile Asn Asp Arg Ile Met Lys Gly Pro
35 40 45

Phe Gly Gly Gly His Ser Thr Met Arg Phe Ser Leu Tyr Ile Leu Phe
50 55 60

Glu Glu His Leu Arg His Met Asn Phe Trp Thr Arg Ser Asn Asp Asn
65 70 75 80

Leu Glu Leu Thr Arg Tyr Leu Gly Ala Ser Val Lys Ile Tyr Arg His
85 90 95

Pro Asp Gln Asp Phe Ile Val Ile Tyr Asn Arg Arg Thr Pro Leu Gly
100 105 110

Gly Asn Ile Tyr Thr Ala Pro Ser Leu His Pro Gly Asn Ala Ile Leu
115 120 125

Ala Lys His Lys Ile Leu Val Pro Ser Leu Gln Thr Arg Pro Lys Gly
130 135 140

Arg Lys Ala Ile Arg Leu Arg Ile Ala Pro Pro Thr Leu Phe Thr Asp
145 150 155 160

Lys Trp Tyr Phe Gln Lys Asp Ile Ala Asp Leu Thr Leu Phe Asn Ile
165 170 175

Met Ala Val Glu Ala Asp Leu Arg Phe Pro Phe Cys Ser Pro Gln Thr
180 185 190

Asp Asn Thr Cys Ile Ser Phe Gln Val Leu Ser Ser Val Tyr Asn Asn
195 200 205

Tyr Leu Ser Ile
210

<210> 188
<211> 133
<212> PRT
<213> α細環病毒屬

<400> 188
Asn Thr Phe Asn Asn Asp Asn Ser Asp Ser Lys Leu Lys Glu Phe Leu
1 5 10 15

Asn Lys Ala Phe Pro Thr Thr Gly Thr Lys Gly Thr Ser Leu Asn Ala
20 25 30

Leu Asn Thr Phe Arg Thr Glu Gly Cys Ile Ser His Pro Gln Leu Lys
35 40 45

Lys Pro Asn Pro Gln Ile Asn Lys Pro Leu Glu Ser Gln Tyr Phe Ala
50 55 60

Pro Leu Asp Ala Leu Trp Gly Asp Pro Ile Tyr Tyr Asn Asp Leu Asn
65 70 75 80

Glu Asn Lys Ser Leu Asn Asp Ile Ile Glu Lys Ile Leu Ile Lys Asn
85 90 95

Met Ile Thr Tyr His Ala Lys Leu Arg Glu Phe Pro Asn Ser Tyr Gln
100 105 110

Gly Asn Lys Ala Phe Cys His Leu Thr Gly Ile Tyr Ser Pro Pro Tyr
115 120 125

Leu Asn Gln Gly Arg
130

<210> 189
<211> 166
<212> PRT
<213> α細環病毒屬

<400> 189

Ile Ser Pro Glu Ile Phe Gly Leu Tyr Thr Glu Ile Ile Tyr Asn Pro
1 5 10 15

Tyr Thr Asp Lys Gly Thr Gly Asn Lys Val Trp Met Asp Pro Leu Thr
20 25 30

Lys Glu Asn Asn Ile Tyr Lys Glu Gly Gln Ser Lys Cys Leu Leu Thr
35 40 45

Asp Met Pro Leu Trp Thr Leu Leu Phe Gly Tyr Thr Asp Trp Cys Lys
50 55 60

Lys Asp Thr Asn Asn Trp Asp Leu Pro Leu Asn Tyr Arg Leu Val Leu
65 70 75 80

Ile Cys Pro Tyr Thr Phe Pro Lys Leu Tyr Asn Glu Lys Val Lys Asp
85 90 95

Tyr Gly Tyr Ile Pro Tyr Ser Tyr Lys Phe Gly Ala Gly Gln Met Pro
100 105 110

Asp Gly Ser Asn Tyr Ile Pro Phe Gln Phe Arg Ala Lys Trp Tyr Pro
115 120 125

Thr Val Leu His Gln Gln Gln Val Met Glu Asp Ile Ser Arg Ser Gly
130 135 140

Pro Phe Ala Pro Lys Val Glu Lys Pro Ser Thr Gln Leu Val Met Lys
145 150 155 160

Tyr Cys Phe Asn Phe Asn
165

<210> 190
<211> 164
<212> PRT
<213> α細環病毒屬

<400> 190

```
Trp Gly Gly Asn Pro Ile Ile Glu Gln Ile Val Lys Asp Pro Ser Phe
1               5                  10                  15

Gln Pro Thr Tyr Glu Ile Pro Gly Thr Gly Asn Ile Pro Arg Arg Ile
            20                  25                  30

Gln Val Ile Asp Pro Arg Val Leu Gly Pro His Tyr Ser Phe Arg Ser
        35                  40                  45

Trp Asp Met Arg Arg His Thr Phe Ser Arg Ala Ser Ile Lys Arg Val
    50                  55                  60

Ser Glu Gln Gln Glu Thr Ser Asp Leu Val Phe Ser Gly Pro Lys Lys
65                  70                  75                  80

Pro Arg Val Asp Ile Pro Lys Gln Glu Thr Gln Glu Glu Ser Ser His
                85                  90                  95

Ser Leu Gln Arg Glu Ser Arg Pro Trp Glu Thr Glu Glu Glu Ser Glu
            100                 105                 110

Thr Glu Ala Leu Ser Gln Glu Ser Gln Glu Val Pro Phe Gln Gln Gln
        115                 120                 125

Leu Gln Gln Gln Tyr Gln Glu Gln Leu Lys Leu Arg Gln Gly Ile Lys
    130                 135                 140

Val Leu Phe Glu Gln Leu Ile Arg Thr Gln Gln Gly Val His Val Asn
145                 150                 155                 160

Pro Cys Leu Arg
```

<210> 191

<400> 191
000

<210> 192

<400> 192
000

<210> 193

<400> 193
000

<210> 194

<400> 194
000

<210> 195

<400> 195
000

<210> 196

<400> 196
000

<210> 197

<400> 197
000

<210> 198

<400> 198
000

<210> 199

<400> 199
000

<210> 200

<400> 200
000

<210> 201

<400> 201
000

<210> 202

<400> 202
000

<210> 203

<400> 203
000

<210> 204

<400> 204
000

<210> 205

<400> 205
000

<210> 206

<400> 206
000

<210> 207

<400> 207
000

<210> 208

<400> 208
000

<210> 209

<400> 209
000

<210> 210

<400> 210
000

<210> 211

<400> 211
000

<210> 212

<400> 212
000

<210> 213

<400> 213
000

<210> 214

<400> 214
000

<210> 215
<211> 666
<212> PRT
<213> β細環病毒屬

<400> 215

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Met | Pro | Tyr | Tyr | Tyr | Arg | Arg | Arg | Arg | Tyr | Asn | Tyr | Arg | Arg | Pro | Arg |
| 1 | | | | 5 | | | | 10 | | | | | 15 | | |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Trp | Tyr | Gly | Arg | Gly | Trp | Ile | Arg | Arg | Pro | Phe | Arg | Arg | Arg | Phe | Arg |
| | | 20 | | | | | 25 | | | | | 30 | | | |

```
Arg Lys Arg Arg Val Arg Pro Thr Tyr Thr Thr Ile Pro Leu Lys Gln
        35                  40                  45

Trp Gln Pro Pro Tyr Lys Arg Thr Cys Tyr Ile Lys Gly Gln Asp Cys
    50                  55                  60

Leu Ile Tyr Tyr Ser Asn Leu Arg Leu Gly Met Asn Ser Thr Met Tyr
65                  70                  75                  80

Glu Lys Ser Ile Val Pro Val His Trp Pro Gly Gly Gly Ser Phe Ser
                85                  90                  95

Val Ser Met Leu Thr Leu Asp Ala Leu Tyr Asp Ile His Lys Leu Cys
            100                 105                 110

Arg Asn Trp Trp Thr Ser Thr Asn Gln Asp Leu Pro Leu Val Arg Tyr
        115                 120                 125

Lys Gly Cys Lys Ile Thr Phe Tyr Gln Ser Thr Phe Thr Asp Tyr Ile
    130                 135                 140

Val Arg Ile His Thr Glu Leu Pro Ala Asn Ser Asn Lys Leu Thr Tyr
145                 150                 155                 160

Pro Asn Thr His Pro Leu Met Met Met Met Ser Lys Tyr Lys His Ile
                165                 170                 175

Ile Pro Ser Arg Gln Thr Arg Arg Lys Lys Lys Pro Tyr Thr Lys Ile
            180                 185                 190

Phe Val Lys Pro Pro Pro Gln Phe Glu Asn Lys Trp Tyr Phe Ala Thr
        195                 200                 205

Asp Leu Tyr Lys Ile Pro Leu Leu Gln Ile His Cys Thr Ala Cys Asn
    210                 215                 220
```

```
Leu Gln Asn Pro Phe Val Lys Pro Asp Lys Leu Ser Asn Asn Val Thr
225                 230                 235                 240

Leu Trp Ser Leu Asn Thr Ile Ser Ile Gln Asn Arg Asn Met Ser Val
                245                 250                 255

Asp Gln Gly Gln Ser Trp Pro Phe Lys Ile Leu Gly Thr Gln Ser Phe
            260                 265                 270

Tyr Phe Tyr Phe Tyr Thr Gly Ala Asn Leu Pro Gly Asp Thr Thr Gln
        275                 280                 285

Ile Pro Val Ala Asp Leu Leu Pro Leu Thr Asn Pro Arg Ile Asn Arg
    290                 295                 300

Pro Gly Gln Ser Leu Asn Glu Ala Lys Ile Thr Asp His Ile Thr Phe
305                 310                 315                 320

Thr Glu Tyr Lys Asn Lys Phe Thr Asn Tyr Trp Gly Asn Pro Phe Asn
                325                 330                 335

Lys His Ile Gln Glu His Leu Asp Met Ile Leu Tyr Ser Leu Lys Ser
            340                 345                 350

Pro Glu Ala Ile Lys Asn Glu Trp Thr Thr Glu Asn Met Lys Trp Asn
        355                 360                 365

Gln Leu Asn Asn Ala Gly Thr Met Ala Leu Thr Pro Phe Asn Glu Pro
    370                 375                 380

Ile Phe Thr Gln Ile Gln Tyr Asn Pro Asp Arg Asp Thr Gly Glu Asp
385                 390                 395                 400

Thr Gln Leu Tyr Leu Leu Ser Asn Ala Thr Gly Thr Gly Trp Asp Pro
                405                 410                 415

Pro Gly Ile Pro Glu Leu Ile Leu Glu Gly Phe Pro Leu Trp Leu Ile
            420                 425                 430
```

Tyr Trp Gly Phe Ala Asp Phe Gln Lys Asn Leu Lys Lys Val Thr Asn
435 440 445

Ile Asp Thr Asn Tyr Met Leu Val Ala Lys Thr Lys Phe Thr Gln Lys
450 455 460

Pro Gly Thr Phe Tyr Leu Val Ile Leu Asn Asp Thr Phe Val Glu Gly
465 470 475 480

Asn Ser Pro Tyr Glu Lys Gln Pro Leu Pro Glu Asp Asn Ile Lys Trp
485 490 495

Tyr Pro Gln Val Gln Tyr Gln Leu Glu Ala Gln Asn Lys Leu Leu Gln
500 505 510

Thr Gly Pro Phe Thr Pro Asn Ile Gln Gly Gln Leu Ser Asp Asn Ile
515 520 525

Ser Met Phe Tyr Lys Phe Tyr Phe Lys Trp Gly Gly Ser Pro Pro Lys
530 535 540

Ala Ile Asn Val Glu Asn Pro Ala His Gln Ile Gln Tyr Pro Ile Pro
545 550 555 560

Arg Asn Glu His Glu Thr Thr Ser Leu Gln Ser Pro Gly Glu Ala Pro
565 570 575

Glu Ser Ile Leu Tyr Ser Phe Asp Tyr Arg His Gly Asn Tyr Thr Thr
580 585 590

Thr Ala Leu Ser Arg Ile Ser Gln Asp Trp Ala Leu Lys Asp Thr Val
595 600 605

Ser Lys Ile Thr Glu Pro Asp Arg Gln Gln Leu Leu Lys Gln Ala Leu
610 615 620

Glu Cys Leu Gln Ile Ser Glu Glu Thr Gln Glu Lys Lys Glu Lys Glu
625 630 635 640

Val Gln Gln Leu Ile Ser Asn Leu Arg Gln Gln Gln Gln Leu Tyr Arg
645 650 655

Glu Arg Ile Ile Ser Leu Leu Lys Asp Gln
660 665

<210> 216
<211> 38
<212> PRT
<213> β細環病毒屬

<400> 216
Met Pro Tyr Tyr Tyr Arg Arg Arg Arg Tyr Asn Tyr Arg Arg Pro Arg
1 5 10 15

Trp Tyr Gly Arg Gly Trp Ile Arg Arg Pro Phe Arg Arg Arg Phe Arg
20 25 30

Arg Lys Arg Arg Val Arg
35

<210> 217
<211> 208
<212> PRT
<213> β細環病毒屬

<400> 217
Pro Thr Tyr Thr Thr Ile Pro Leu Lys Gln Trp Gln Pro Pro Tyr Lys
1 5 10 15

Arg Thr Cys Tyr Ile Lys Gly Gln Asp Cys Leu Ile Tyr Tyr Ser Asn
20 25 30

Leu Arg Leu Gly Met Asn Ser Thr Met Tyr Glu Lys Ser Ile Val Pro
35 40 45

Val His Trp Pro Gly Gly Gly Ser Phe Ser Val Ser Met Leu Thr Leu
50 55 60

Asp Ala Leu Tyr Asp Ile His Lys Leu Cys Arg Asn Trp Trp Thr Ser
65 70 75 80

Thr Asn Gln Asp Leu Pro Leu Val Arg Tyr Lys Gly Cys Lys Ile Thr
85 90 95

Phe Tyr Gln Ser Thr Phe Thr Asp Tyr Ile Val Arg Ile His Thr Glu
100 105 110

Leu Pro Ala Asn Ser Asn Lys Leu Thr Tyr Pro Asn Thr His Pro Leu
115 120 125

Met Met Met Met Ser Lys Tyr Lys His Ile Ile Pro Ser Arg Gln Thr
130 135 140

Arg Arg Lys Lys Lys Pro Tyr Thr Lys Ile Phe Val Lys Pro Pro Pro
145 150 155 160

Gln Phe Glu Asn Lys Trp Tyr Phe Ala Thr Asp Leu Tyr Lys Ile Pro
165 170 175

Leu Leu Gln Ile His Cys Thr Ala Cys Asn Leu Gln Asn Pro Phe Val
180 185 190

Lys Pro Asp Lys Leu Ser Asn Asn Val Thr Leu Trp Ser Leu Asn Thr
195 200 205

<210> 218
<211> 128
<212> PRT
<213> β細環病毒屬

<400> 218
Ile Ser Ile Gln Asn Arg Asn Met Ser Val Asp Gln Gly Gln Ser Trp
1 5 10 15

Pro Phe Lys Ile Leu Gly Thr Gln Ser Phe Tyr Phe Tyr Phe Tyr Thr
20 25 30

Gly Ala Asn Leu Pro Gly Asp Thr Thr Gln Ile Pro Val Ala Asp Leu
35 40 45

Leu Pro Leu Thr Asn Pro Arg Ile Asn Arg Pro Gly Gln Ser Leu Asn
50 55 60

Glu Ala Lys Ile Thr Asp His Ile Thr Phe Thr Glu Tyr Lys Asn Lys
65 70 75 80

Phe Thr Asn Tyr Trp Gly Asn Pro Phe Asn Lys His Ile Gln Glu His
85 90 95

Leu Asp Met Ile Leu Tyr Ser Leu Lys Ser Pro Glu Ala Ile Lys Asn
100 105 110

Glu Trp Thr Thr Glu Asn Met Lys Trp Asn Gln Leu Asn Asn Ala Gly
115 120 125

<210> 219
<211> 163
<212> PRT
<213> β細環病毒屬

<400> 219
Thr Met Ala Leu Thr Pro Phe Asn Glu Pro Ile Phe Thr Gln Ile Gln
1 5 10 15

Tyr Asn Pro Asp Arg Asp Thr Gly Glu Asp Thr Gln Leu Tyr Leu Leu
20 25 30

Ser Asn Ala Thr Gly Thr Gly Trp Asp Pro Pro Gly Ile Pro Glu Leu
35 40 45

Ile Leu Glu Gly Phe Pro Leu Trp Leu Ile Tyr Trp Gly Phe Ala Asp
50 55 60

Phe Gln Lys Asn Leu Lys Lys Val Thr Asn Ile Asp Thr Asn Tyr Met
65 70 75 80

Leu Val Ala Lys Thr Lys Phe Thr Gln Lys Pro Gly Thr Phe Tyr Leu
85 90 95

Val Ile Leu Asn Asp Thr Phe Val Glu Gly Asn Ser Pro Tyr Glu Lys
100 105 110

Gln Pro Leu Pro Glu Asp Asn Ile Lys Trp Tyr Pro Gln Val Gln Tyr
115 120 125

Gln Leu Glu Ala Gln Asn Lys Leu Leu Gln Thr Gly Pro Phe Thr Pro
130 135 140

Asn Ile Gln Gly Gln Leu Ser Asp Asn Ile Ser Met Phe Tyr Lys Phe
145 150 155 160

Tyr Phe Lys

<210> 220
<211> 129
<212> PRT
<213> β細環病毒屬

<400> 220
Trp Gly Gly Ser Pro Pro Lys Ala Ile Asn Val Glu Asn Pro Ala His
1 5 10 15

Gln Ile Gln Tyr Pro Ile Pro Arg Asn Glu His Glu Thr Thr Ser Leu
20 25 30

Gln Ser Pro Gly Glu Ala Pro Glu Ser Ile Leu Tyr Ser Phe Asp Tyr
35 40 45

Arg His Gly Asn Tyr Thr Thr Thr Ala Leu Ser Arg Ile Ser Gln Asp
50 55 60

Trp Ala Leu Lys Asp Thr Val Ser Lys Ile Thr Glu Pro Asp Arg Gln
65 70 75 80

Gln Leu Leu Lys Gln Ala Leu Glu Cys Leu Gln Ile Ser Glu Glu Thr
85 90 95

Gln Glu Lys Lys Glu Lys Glu Val Gln Gln Leu Ile Ser Asn Leu Arg
100 105 110

Gln Gln Gln Gln Leu Tyr Arg Glu Arg Ile Ile Ser Leu Leu Lys Asp
115 120 125

Gln

<210> 221

<400> 221
000

<210> 222

<400> 222
000

<210> 223

<400> 223
000

<210> 224

<400> 224
000

<210> 225

<400> 225
000

<210> 226

<400> 226
000

<210> 227
<211> 220
<212> PRT
<213> 人工序列

<220>
<223> 人工序列描述：合成多肽

<220>
<221> MOD_RES
<222> (29)..(31)
<223> 任何胺基酸

<220>
<221> SITE
<222> (29)..(31)
<223> 此區域可涵蓋0-3個殘基

<220>
<221> MOD_RES
<222> (100)..(100)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (125)..(129)
<223> 任何胺基酸

<220>
<221> SITE
<222> (125)..(129)
<223> 此區域可涵蓋1-5個殘基

<220>
<221> MOD_RES
<222> (181)..(181)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (211)..(211)
<223> 任何胺基酸

<400> 227
Leu Val Leu Thr Gln Trp Gln Pro Asn Thr Val Arg Arg Cys Tyr Ile

1 5 10 15

Arg Gly Tyr Leu Pro Leu Ile Ile Cys Gly Glu Asn Xaa Xaa Xaa Thr
20 25 30

Thr Ser Arg Asn Tyr Ala Thr His Ser Asp Asp Thr Ile Gln Lys Gly
35 40 45

Pro Phe Gly Gly Gly Met Ser Thr Thr Thr Phe Ser Leu Arg Val Leu
50 55 60

Tyr Asp Glu Tyr Gln Arg Phe Met Asn Arg Trp Thr Tyr Ser Asn Glu
65 70 75 80

Asp Leu Asp Leu Ala Arg Tyr Leu Gly Cys Lys Phe Thr Phe Tyr Arg
85 90 95

His Pro Asp Xaa Asp Phe Ile Val Gln Tyr Asn Thr Asn Pro Pro Phe
100 105 110

Lys Asp Thr Lys Leu Thr Ala Pro Ser Ile His Pro Xaa Xaa Xaa Xaa
115 120 125

Xaa Gly Met Leu Met Leu Ser Lys Arg Lys Ile Leu Ile Pro Ser Leu
130 135 140

Lys Thr Arg Pro Lys Gly Lys His Tyr Val Lys Val Arg Ile Gly Pro
145 150 155 160

Pro Lys Leu Phe Glu Asp Lys Trp Tyr Thr Gln Ser Asp Leu Cys Asp
165 170 175

Val Pro Leu Val Xaa Leu Tyr Ala Thr Ala Ala Asp Leu Gln His Pro
180 185 190

Phe Gly Ser Pro Gln Thr Asp Asn Pro Cys Val Thr Phe Gln Val Leu
195 200 205

Gly Ser Xaa Tyr Asn Lys His Leu Ser Ile Ser Pro
210 215 220

<210> 228
<211> 172
<212> PRT
<213> 人工序列

<220>
<223> 人工序列描述：合成多肽

<220>
<221> MOD_RES
<222> (38)..(38)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (44)..(46)
<223> 任何胺基酸

<220>
<221> SITE
<222> (44)..(46)
<223> 此區域可涵蓋0-3個殘基

<220>
<221> MOD_RES
<222> (77)..(77)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (79)..(79)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (98)..(101)
<223> 任何胺基酸

<220>
<221> SITE
<222> (98)..(101)
<223> 此區域可涵蓋0-4個殘基

<400> 228

Ser Asn Phe Glu Phe Pro Gly Ala Tyr Thr Asp Ile Thr Tyr Asn Pro
1 5 10 15

Leu Thr Asp Lys Gly Val Gly Asn Met Val Trp Ile Gln Tyr Leu Thr
20 25 30

Lys Pro Asp Thr Ile Xaa Asp Lys Thr Gln Ser Xaa Xaa Xaa Lys Cys
35 40 45

Leu Ile Glu Asp Leu Pro Leu Trp Ala Ala Leu Tyr Gly Tyr Val Asp
50 55 60

Phe Cys Glu Lys Glu Thr Gly Asp Ser Ala Ile Ile Xaa Asn Xaa Gly
65 70 75 80

Arg Val Leu Ile Arg Cys Pro Tyr Thr Lys Pro Pro Leu Tyr Asp Lys
85 90 95

Thr Xaa Xaa Xaa Xaa Asn Lys Gly Phe Val Pro Tyr Ser Thr Asn Phe
100 105 110

Gly Asn Gly Lys Met Pro Gly Gly Ser Gly Tyr Val Pro Ile Tyr Trp
115 120 125

Arg Ala Arg Trp Tyr Pro Thr Leu Phe His Gln Lys Glu Val Leu Glu
130 135 140

Asp Ile Val Gln Ser Gly Pro Phe Ala Tyr Lys Asp Glu Lys Pro Ser
145 150 155 160

Thr Gln Leu Val Met Lys Tyr Cys Phe Asn Phe Asn
165 170

<210> 229
<211> 258
<212> PRT
<213> 人工序列

<220>

<223> 人工序列描述：合成多肽

<220>
<221> MOD_RES
<222> (20)..(22)
<223> 任何胺基酸

<220>
<221> SITE
<222> (20)..(22)
<223> 此區域可涵蓋0-3個殘基

<220>
<221> MOD_RES
<222> (25)..(25)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (78)..(78)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (89)..(89)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (91)..(91)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (95)..(98)
<223> 任何胺基酸

<220>
<221> SITE
<222> (95)..(98)
<223> 此區域可涵蓋1-4個殘基

<220>
<221> MOD_RES
<222> (107)..(120)
<223> 任何胺基酸

<220>
<221> SITE

<222> (107)..(120)
<223> 此區域可涵蓋2-14個殘基

<220>
<221> MOD_RES
<222> (129)..(129)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (139)..(168)
<223> 任何胺基酸

<220>
<221> SITE
<222> (139)..(168)
<223> 此區域可涵蓋0-30個殘基

<220>
<221> MOD_RES
<222> (201)..(204)
<223> 任何胺基酸

<220>
<221> SITE
<222> (201)..(204)
<223> 此區域可涵蓋0-4個殘基

<220>
<221> MOD_RES
<222> (219)..(258)
<223> 任何胺基酸

<220>
<221> SITE
<222> (219)..(258)
<223> 此區域可涵蓋0-40個殘基

<400> 229

```
Trp Gly Gly Asn Pro Ile Ser Gln Gln Val Val Arg Asn Pro Cys Lys
1               5                  10                  15


Asp Ser Gly Xaa Xaa Xaa Ser Gly Xaa Gly Arg Gln Pro Arg Ser Val
            20                  25                  30


Gln Val Val Asp Pro Lys Tyr Met Gly Pro Glu Tyr Thr Phe His Ser
        35                  40                  45
```

Trp Asp Trp Arg Arg Gly Leu Phe Gly Glu Lys Ala Ile Lys Arg Met
50 55 60

Ser Glu Gln Pro Thr Asp Asp Glu Ile Phe Thr Gly Gly Xaa Pro Lys
65 70 75 80

Arg Pro Arg Arg Asp Pro Pro Thr Xaa Gln Xaa Pro Glu Glu Xaa Xaa
85 90 95

Xaa Xaa Gln Lys Glu Ser Ser Ser Phe Arg Xaa Xaa Xaa Xaa Xaa Xaa
100 105 110

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Pro Trp Glu Ser Ser Ser Gln Glu
115 120 125

Xaa Glu Ser Glu Ser Gln Glu Glu Glu Glu Xaa Xaa Xaa Xaa Xaa Xaa
130 135 140

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
145 150 155 160

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Glu Gln Thr Val Gln Gln Gln Leu
165 170 175

Arg Gln Gln Leu Arg Glu Gln Arg Arg Leu Arg Val Gln Leu Gln Leu
180 185 190

Leu Phe Gln Gln Leu Leu Lys Thr Xaa Xaa Xaa Xaa Gln Ala Gly Leu
195 200 205

His Ile Asn Pro Leu Leu Leu Ser Gln Ala Xaa Xaa Xaa Xaa Xaa Xaa
210 215 220

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
225 230 235 240

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

245 250 255

Xaa Xaa

<210> 230
<211> 214
<212> PRT
<213> 人工序列

<220>
<223> 人工序列描述：合成多肽

<220>
<221> MOD_RES
<222> (136)..(136)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (138)..(141)
<223> 任何胺基酸

<220>
<221> SITE
<222> (138)..(141)
<223> 此區域可涵蓋1-4個殘基

<220>
<221> MOD_RES
<222> (179)..(179)
<223> 任何胺基酸

<400> 230
Leu Lys Gln Trp Gln Pro Ser Thr Ile Arg Lys Cys Lys Ile Lys Gly
1 5 10 15

Tyr Leu Pro Leu Phe Gln Cys Gly Lys Gly Arg Ile Ser Asn Asn Tyr
20 25 30

Thr Gln Tyr Lys Glu Ser Ile Val Pro His His Glu Pro Gly Gly Gly
35 40 45

Gly Trp Ser Ile Gln Gln Phe Thr Leu Gly Ala Leu Tyr Glu Glu His

50 55 60

Leu Lys Leu Arg Asn Trp Trp Thr Lys Ser Asn Asp Gly Leu Pro Leu
65 70 75 80

Val Arg Tyr Leu Gly Cys Thr Ile Lys Leu Tyr Arg Ser Glu Asp Thr
85 90 95

Asp Tyr Ile Val Thr Tyr Gln Arg Cys Tyr Pro Met Thr Ala Thr Lys
100 105 110

Leu Thr Tyr Leu Ser Thr Gln Pro Ser Arg Met Leu Met Asn Lys His
115 120 125

Lys Ile Ile Val Pro Ser Lys Xaa Thr Xaa Xaa Xaa Xaa Asn Lys Lys
130 135 140

Lys Lys Pro Tyr Lys Lys Ile Phe Ile Lys Pro Pro Ser Gln Met Gln
145 150 155 160

Asn Lys Trp Tyr Phe Gln Gln Asp Ile Ala Asn Thr Pro Leu Leu Gln
165 170 175

Leu Thr Xaa Thr Ala Cys Ser Leu Asp Arg Met Tyr Leu Ser Ser Asp
180 185 190

Ser Ile Ser Asn Asn Ile Thr Phe Thr Ser Leu Asn Thr Asn Phe Phe
195 200 205

Gln Asn Pro Asn Phe Gln
210

<210> 231
<211> 187
<212> PRT
<213> 人工序列

<220>
<223> 人工序列描述：合成多肽

<220>
<221> MOD_RES
<222> (1)..(10)
<223> 任何胺基酸

<220>
<221> SITE
<222> (1)..(10)
<223> 此區域可涵蓋4-10個殘基

<220>
<221> MOD_RES
<222> (38)..(45)
<223> 任何胺基酸

<220>
<221> SITE
<222> (38)..(45)
<223> 此區域可涵蓋1-8個殘基

<220>
<221> MOD_RES
<222> (94)..(94)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (100)..(102)
<223> 任何胺基酸

<220>
<221> SITE
<222> (100)..(102)
<223> 此區域可涵蓋1-3個殘基

<220>
<221> MOD_RES
<222> (112)..(112)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (114)..(115)
<223> 任何胺基酸

<220>
<221> SITE
<222> (114)..(115)

<223> 此區域可涵蓋0-2個殘基

<220>
<221> MOD_RES
<222> (124)..(139)
<223> 任何胺基酸

<220>
<221> SITE
<222> (124)..(139)
<223> 此區域可涵蓋3-16個殘基

<220>
<221> MOD_RES
<222> (154)..(154)
<223> 任何胺基酸

<400> 231

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Thr Pro Leu Tyr Phe Glu
1 5 10 15

Cys Arg Tyr Asn Pro Phe Lys Asp Lys Gly Thr Gly Asn Lys Val Tyr
20 25 30

Leu Val Ser Asn Asn Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Thr Gly Trp
35 40 45

Asp Pro Pro Thr Asp Pro Asp Leu Ile Ile Glu Gly Phe Pro Leu Trp
50 55 60

Leu Leu Leu Trp Gly Trp Leu Asp Trp Gln Lys Lys Leu Gly Lys Ile
65 70 75 80

Gln Asn Ile Asp Thr Asp Tyr Ile Leu Val Ile Gln Ser Xaa Tyr Tyr
85 90 95

Ile Pro Pro Xaa Xaa Xaa Lys Leu Pro Tyr Tyr Val Pro Leu Asp Xaa
100 105 110

Asp Xaa Xaa Phe Leu His Gly Arg Ser Pro Tyr Xaa Xaa Xaa Xaa Xaa
115 120 125

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Pro Ser Asp Lys Gln
130 135 140

His Trp His Pro Lys Val Arg Phe Gln Xaa Glu Thr Ile Asn Asn Ile
145 150 155 160

Ala Leu Thr Gly Pro Gly Thr Pro Lys Leu Pro Asn Gln Lys Ser Ile
165 170 175

Gln Ala His Met Lys Tyr Lys Phe Tyr Phe Lys
180 185

<210> 232
<211> 163
<212> PRT
<213> 人工序列

<220>
<223> 人工序列描述：合成多肽

<220>
<221> MOD_RES
<222> (34)..(34)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (65)..(65)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (77)..(78)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (86)..(87)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (96)..(96)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (102)..(106)
<223> 任何胺基酸

<220>
<221> SITE
<222> (102)..(106)
<223> 此區域可涵蓋0-5個殘基

<220>
<221> MOD_RES
<222> (125)..(125)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (135)..(135)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (138)..(163)
<223> 任何胺基酸

<220>
<221> SITE
<222> (138)..(163)
<223> 此區域可涵蓋0-26個殘基

<400> 232

```
Trp Gly Gly Cys Pro Ala Pro Met Glu Thr Ile Thr Asp Pro Cys Lys
1               5                  10                  15

Gln Pro Lys Tyr Pro Ile Pro Asn Asn Leu Leu Gln Thr Thr Ser Leu
            20                  25                  30

Gln Xaa Pro Thr Thr Pro Ile Glu Thr Tyr Leu Tyr Lys Phe Asp Glu
        35                  40                  45

Arg Arg Gly Leu Leu Thr Lys Lys Ala Ala Lys Arg Ile Lys Lys Asp
    50                  55                  60

Xaa Thr Thr Glu Thr Thr Leu Phe Thr Asp Thr Gly Xaa Xaa Thr Ser
65                  70                  75                  80
```

Thr Thr Leu Pro Thr Xaa Xaa Gln Thr Glu Thr Thr Gln Glu Glu Xaa
85 90 95

Thr Ser Glu Glu Glu Xaa Xaa Xaa Xaa Xaa Glu Thr Leu Leu Gln Gln
100 105 110

Leu Gln Gln Leu Arg Arg Lys Gln Lys Gln Leu Arg Xaa Arg Ile Leu
115 120 125

Gln Leu Leu Gln Leu Leu Xaa Leu Leu Xaa Xaa Xaa Xaa Xaa Xaa Xaa
130 135 140

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
145 150 155 160

Xaa Xaa Xaa

<210> 233
<211> 203
<212> PRT
<213> 人工序列

<220>
<223> 人工序列描述：合成多肽

<220>
<221> MOD_RES
<222> (79)..(79)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (104)..(104)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (116)..(116)
<223> 任何胺基酸

<220>

<221> MOD_RES
<222> (120)..(121)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (125)..(125)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (170)..(170)
<223> 任何胺基酸

<400> 233

```
Thr Ile Pro Leu Lys Gln Trp Gln Pro Glu Ser Ile Arg Lys Cys Lys
1               5                  10                  15

Ile Lys Gly Tyr Gly Thr Leu Val Leu Gly Ala Glu Gly Arg Gln Phe
            20                  25                  30

Tyr Cys Tyr Thr Asn Glu Lys Asp Glu Tyr Thr Pro Pro Lys Ala Pro
        35                  40                  45

Gly Gly Gly Gly Phe Gly Val Glu Leu Phe Ser Leu Glu Tyr Leu Tyr
    50                  55                  60

Glu Gln Trp Lys Ala Arg Asn Asn Ile Trp Thr Lys Ser Asn Xaa Tyr
65                  70                  75                  80

Lys Asp Leu Cys Arg Tyr Thr Gly Cys Lys Ile Thr Phe Tyr Arg His
                85                  90                  95

Pro Thr Thr Asp Phe Ile Val Xaa Tyr Ser Arg Gln Pro Pro Phe Glu
            100                 105                 110

Ile Asp Lys Xaa Thr Tyr Met Xaa Xaa His Pro Gln Xaa Leu Leu Leu
        115                 120                 125

Arg Lys His Lys Lys Ile Ile Leu Ser Lys Ala Thr Asn Pro Lys Gly
    130                 135                 140
```

Lys Leu Lys Lys Lys Ile Lys Ile Lys Pro Pro Lys Gln Met Leu Asn
145 150 155 160

Lys Trp Phe Phe Gln Lys Gln Phe Ala Xaa Tyr Gly Leu Val Gln Leu
165 170 175

Gln Ala Ala Ala Cys Asx Leu Arg Tyr Pro Arg Leu Gly Cys Cys Asn
180 185 190

Glu Asn Arg Leu Ile Thr Leu Tyr Tyr Leu Asn
195 200

<210> 234
<211> 162
<212> PRT
<213> 人工序列

<220>
<223> 人工序列描述：合成多肽

<220>
<221> MOD_RES
<222> (12)..(12)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (20)..(20)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (23)..(23)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (30)..(30)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (58)..(58)
<223> I或L

<220>
<221> MOD_RES
<222> (84)..(84)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (90)..(90)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (95)..(95)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (105)..(105)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (111)..(111)
<223> I或L

<220>
<221> MOD_RES
<222> (113)..(113)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (154)..(154)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (156)..(156)
<223> 任何胺基酸

<400> 234
Leu Pro Ile Val Val Ala Arg Tyr Asn Pro Ala Xaa Asp Thr Gly Lys
1 5 10 15

Gly Asn Lys Xaa Trp Leu Xaa Ser Thr Leu Asn Gly Ser Xaa Trp Ala
20 25 30

Pro Pro Thr Thr Asp Lys Asp Leu Ile Ile Glu Gly Leu Pro Leu Trp
35 40 45

Leu Ala Leu Tyr Gly Tyr Trp Ser Tyr Xaa Lys Lys Val Lys Lys Asp
50 55 60

Lys Gly Ile Leu Gln Ser His Met Phe Val Val Lys Ser Pro Ala Ile
65 70 75 80

Gln Pro Leu Xaa Thr Ala Thr Thr Gln Xaa Thr Phe Tyr Pro Xaa Ile
85 90 95

Asp Asn Ser Phe Ile Gln Gly Lys Xaa Pro Tyr Asp Glu Pro Xaa Thr
100 105 110

Xaa Asn Gln Lys Lys Leu Trp Tyr Pro Thr Leu Glu His Gln Gln Glu
115 120 125

Thr Ile Asn Ala Ile Val Glu Ser Gly Pro Tyr Val Pro Lys Leu Asp
130 135 140

Asn Gln Lys Asn Ser Thr Trp Glu Leu Xaa Tyr Xaa Tyr Thr Phe Tyr
145 150 155 160

Phe Lys

<210> 235
<211> 177
<212> PRT
<213> 人工序列

<220>
<223> 人工序列描述：合成多肽

<220>
<221> MOD_RES
<222> (16)..(16)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (26)..(26)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (33)..(33)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (73)..(73)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (81)..(82)
<223> 任何胺基酸

<220>
<221> SITE
<222> (81)..(82)
<223> 此區域可涵蓋0-2個殘基

<220>
<221> MOD_RES
<222> (90)..(90)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (94)..(94)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (119)..(124)
<223> 任何胺基酸

<220>
<221> SITE
<222> (119)..(124)
<223> 此區域可涵蓋1-6個殘基

<220>
<221> MOD_RES
<222> (168)..(177)
<223> 任何胺基酸

<220>
<221> SITE
<222> (168)..(177)
<223> 此區域可涵蓋1-10個殘基

<400> 235

```
Trp Gly Gly Pro Gln Ile Pro Asp Gln Pro Val Glu Asp Pro Lys Xaa
1               5                  10                  15

Gln Gly Thr Tyr Pro Val Pro Asp Thr Xaa Gln Gln Thr Ile Gln Ile
            20                  25                  30

Xaa Asn Pro Leu Lys Gln Lys Pro Glu Thr Met Phe His Asp Trp Asp
        35                  40                  45

Tyr Arg Arg Gly Ile Ile Thr Ser Thr Ala Leu Lys Arg Met Gln Glu
    50                  55                  60

Asn Leu Glu Thr Asp Ser Ser Phe Xaa Ser Asp Ser Glu Glu Thr Pro
65                  70                  75                  80

Xaa Xaa Lys Lys Lys Lys Arg Leu Thr Xaa Glu Leu Pro Xaa Pro Gln
                85                  90                  95

Glu Glu Thr Glu Glu Ile Gln Ser Cys Leu Leu Ser Leu Cys Glu Glu
            100                 105                 110

Ser Thr Cys Gln Glu Glu Xaa Xaa Xaa Xaa Xaa Xaa Glu Asn Leu Gln
        115                 120                 125

Gln Leu Ile His Gln Gln Gln Gln Gln Gln Gln Gln Leu Lys His Asn
    130                 135                 140

Ile Leu Lys Leu Leu Ser Asp Leu Lys Glx Lys Gln Arg Leu Leu Gln
145                 150                 155                 160

Leu Gln Thr Gly Ile Leu Glu Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
                165                 170                 175
```

Xaa

<210> 236

<400> 236
000

<210> 237

<400> 237
000

<210> 238

<400> 238
000

<210> 239

<400> 239
000

<210> 240

<400> 240
000

<210> 241

<400> 241
000

<210> 242

<400> 242
000

<210> 243

<400> 243
000

<210> 244

<400> 244
000

<210> 245

<400> 245
000

<210> 246

<400> 246
000

<210> 247

<400> 247
000

<210> 248

<400> 248
000

<210> 249

<400> 249
000

<210> 250

<400> 250
000

<210> 251

<400> 251
000

<210> 252

<400> 252
000

<210> 253

<400> 253
000

<210> 254

<400> 254
000

<210> 255

<400> 255
000

<210> 256

<400> 256
000

<210> 257

<400> 257
000

<210> 258

<400> 258
000

<210> 259

<400> 259
000

<210> 260

<400> 260
000

<210> 261

<400> 261
000

<210> 262

<400> 262
000

<210> 263

<400> 263
000

<210> 264

<400> 264
000

<210> 265

<400> 265
000

<210> 266

<400> 266
000

<210> 267

<400> 267
000

<210> 268

<400> 268
000

<210> 269

<400> 269
000

<210> 270

<400> 270
000

<210> 271

<400> 271
000

<210> 272

<400> 272
000

<210> 273

<400> 273
000

<210> 274

<400> 274
000

<210> 275

<400> 275
000

<210> 276

<400> 276
000

<210> 277

<400> 277
000

<210> 278

<400> 278
000

<210> 279

<400> 279
000

<210> 280

<400> 280
000

<210> 281

<400> 281
000

<210> 282

<400> 282
000

<210> 283

<400> 283
000

<210> 284

<400> 284
000

<210> 285

<400> 285
000

<210> 286

<400> 286
000

<210> 287

<400> 287
000

<210> 288

<400> 288
000

<210> 289

<400> 289
000

<210> 290

<400> 290
000

<210> 291

<400> 291
000

<210> 292

<400> 292
000

<210> 293

<400> 293
000

<210> 294

<400> 294
000

<210> 295

<400> 295
000

<210> 296

<400> 296
000

<210> 297

<400> 297
000

<210> 298

<400> 298
000

<210> 299

<400> 299
000

<210> 300

<400> 300
000

<210> 301

<400> 301
000

<210> 302

<400> 302
000

<210> 303

<400> 303
000

<210> 304

<400> 304
000

<210> 305

<400> 305
000

<210> 306

<400> 306
000

<210> 307

<400> 307
000

<210> 308

<400> 308
000

<210> 309

<400> 309
000

<210> 310

<400> 310
000

<210> 311

<400> 311
000

<210> 312

<400> 312
000

<210> 313

<400> 313
000

<210> 314

<400> 314
000

<210> 315

<400> 315
000

<210> 316

<400> 316
000

<210> 317

<400> 317
000

<210> 318

<400> 318
000

<210> 319

<400> 319
000

<210> 320

<400> 320
000

<210> 321

<400> 321
000

<210> 322

<400> 322
000

<210> 323

<400> 323
000

<210> 324

<400> 324
000

<210> 325

<400> 325
000

<210> 326

<400> 326
000

<210> 327

<400> 327
000

<210> 328

<400> 328
000

<210> 329

<400> 329
000

<210> 330

<400> 330
000

<210> 331

<400> 331
000

<210> 332

<400> 332
000

<210> 333

<400> 333
000

<210> 334

<400> 334
000

<210> 335

<400> 335
000

<210> 336

<400> 336
000

<210> 337

<400> 337
000

<210> 338

<400> 338
000

<210> 339

<400> 339
000

<210> 340

<400> 340
000

<210> 341

<400> 341
000

<210> 342

<400> 342
000

<210> 343

<400> 343
000

<210> 344

<400> 344
000

<210> 345

<400> 345
000

<210> 346

<400> 346
000

<210> 347

<400> 347
000

<210> 348

<400> 348
000

<210> 349

<400> 349
000

<210> 350

<400> 350
000

<210> 351

<400> 351
000

<210> 352

<400> 352
000

<210> 353

<400> 353
000

<210> 354

<400> 354
000

<210> 355

<400> 355
000

<210> 356

<400> 356
000

<210> 357

<400> 357
000

<210> 358

<400> 358
000

<210> 359

<400> 359
000

<210> 360

<400> 360
000

<210> 361

<400> 361
000

<210> 362

<400> 362
000

<210> 363

<400> 363
000

<210> 364

<400> 364
000

<210> 365

<400> 365
000

<210> 366

<400> 366
000

<210> 367

<400> 367
000

<210> 368

<400> 368
000

<210> 369

<400> 369
000

<210> 370

<400> 370
000

<210> 371

<400> 371
000

<210> 372

<400> 372
000

<210> 373

<400> 373
000

<210> 374

<400> 374
000

<210> 375

<400> 375
000

<210> 376

<400> 376
000

<210> 377

<400> 377
000

<210> 378

<400> 378
000

<210> 379

<400> 379
000

<210> 380

<400> 380
000

<210> 381

<400> 381
000

<210> 382

<400> 382
000

<210> 383

<400> 383
000

<210> 384

<400> 384
000

<210> 385

<400> 385
000

<210> 386

<400> 386
000

<210> 387

<400> 387
000

<210> 388

<400> 388
000

<210> 389

<400> 389
000

<210> 390

<400> 390
000

<210> 391

<400> 391
000

<210> 392

<400> 392
000

<210> 393

<400> 393
000

<210> 394

<400> 394
000

<210> 395

<400> 395
000

<210> 396

<400> 396
000

<210> 397

<400> 397
000

<210> 398

<400> 398
000

<210> 399

<400> 399
000

<210> 400

<400> 400
000

<210> 401

<400> 401
000

<210> 402

<400> 402
000

<210> 403

<400> 403
000

<210> 404

<400> 404
000

<210> 405

<400> 405
000

<210> 406

<400> 406
000

<210> 407

<400> 407
000

<210> 408

<400> 408
000

<210> 409

<400> 409
000

<210> 410

<400> 410
000

<210> 411

<400> 411
000

<210> 412

<400> 412
000

<210> 413

<400> 413
000

<210> 414

<400> 414
000

<210> 415

<400> 415
000

<210> 416

<400> 416
000

<210> 417

<400> 417
000

<210> 418

<400> 418
000

<210> 419

<400> 419
000

<210> 420

<400> 420
000

<210> 421

<400> 421
000

<210> 422

<400> 422
000

<210> 423

<400> 423
000

<210> 424

<400> 424
000

<210> 425

<400> 425
000

<210> 426

<400> 426
000

<210> 427

<400> 427
000

<210> 428

<400> 428
000

<210> 429

<400> 429
000

<210> 430

<400> 430
000

<210> 431

<400> 431
000

<210> 432

<400> 432
000

<210> 433

<400> 433
000

<210> 434

<400> 434
000

<210> 435

<400> 435
000

<210> 436

<400> 436
000

<210> 437

<400> 437
000

<210> 438

<400> 438
000

<210> 439

<400> 439
000

<210> 440

<400> 440
000

<210> 441

<400> 441
000

<210> 442

<400> 442
000

<210> 443

<400> 443
000

<210> 444

<400> 444
000

<210> 445

<400> 445
000

<210> 446

<400> 446
000

<210> 447

<400> 447
000

<210> 448

<400> 448
000

<210> 449

<400> 449
000

<210> 450

<400> 450
000

<210> 451

<400> 451
000

<210> 452

<400> 452
000

<210> 453

<400> 453
000

<210> 454

<400> 454
000

<210> 455

<400> 455
000

<210> 456

<400> 456
000

<210> 457

<400> 457
000

<210> 458

<400> 458
000

<210> 459

<400> 459
000

<210> 460

<400> 460
000

<210> 461

<400> 461
000

<210> 462

<400> 462
000

<210> 463

<400> 463
000

<210> 464

<400> 464
000

<210> 465

<400> 465
000

<210> 466

<400> 466
000

<210> 467

<400> 467
000

<210> 468

<400> 468
000

<210> 469

<400> 469
000

<210> 470

<400> 470
000

<210> 471

<400> 471
000

<210> 472

<400> 472
000

<210> 473

<400> 473
000

<210> 474

<400> 474
000

<210> 475

<400> 475
000

<210> 476

<400> 476
000

<210> 477

<400> 477
000

<210> 478

<400> 478
000

<210> 479

<400> 479
000

<210> 480

<400> 480
000

<210> 481

<400> 481
000

<210> 482

<400> 482
000

<210> 483

<400> 483
000

<210> 484

<400> 484
000

<210> 485

<400> 485
000

<210> 486

<400> 486
000

<210> 487

<400> 487
000

<210> 488

<400> 488
000

<210> 489

<400> 489
000

<210> 490

<400> 490
000

<210> 491

<400> 491
000

<210> 492

<400> 492
000

<210> 493

<400> 493
000

<210> 494

<400> 494
000

<210> 495

<400> 495
000

<210> 496

<400> 496
000

<210> 497

<400> 497
000

<210> 498

<400> 498
000

<210> 499

<400> 499
000

<210> 500

<400> 500
000

<210> 501

<400> 501
000

<210> 502

<400> 502
000

<210> 503

<400> 503
000

<210> 504

<400> 504
000

<210> 505

<400> 505
000

<210> 506

<400> 506
000

<210> 507

<400> 507
000

<210> 508

<400> 508
000

<210> 509

<400> 509
000

<210> 510

<400> 510
000

<210> 511

<400> 511
000

<210> 512

<400> 512
000

<210> 513

<400> 513
000

<210> 514

<400> 514
000

<210> 515

<400> 515
000

<210> 516

<400> 516
000

<210> 517

<400> 517
000

<210> 518

<400> 518
000

<210> 519

<400> 519
000

<210> 520

<400> 520
000

<210> 521

<400> 521
000

<210> 522

<400> 522
000

<210> 523

<400> 523
000

<210> 524

<400> 524
000

<210> 525

<400> 525
000

<210> 526

<400> 526
000

<210> 527

<400> 527
000

<210> 528

<400> 528
000

<210> 529

<400> 529
000

<210> 530

<400> 530
000

<210> 531

<400> 531
000

<210> 532

<400> 532
000

<210> 533

<400> 533
000

<210> 534

<400> 534
000

<210> 535

<400> 535
000

<210> 536

<400> 536
000

<210> 537

<400> 537
000

<210> 538

<400> 538
000

<210> 539

<400> 539
000

<210> 540

<400> 540
000

<210> 541

<400> 541
000

<210> 542

<400> 542
000

<210> 543

<400> 543
000

<210> 544

<400> 544
000

<210> 545

<400> 545
000

<210> 546

<400> 546
000

<210> 547

<400> 547
000

<210> 548

<400> 548
000

<210> 549

<400> 549
000

<210> 550

<400> 550
000

<210> 551

<400> 551
000

<210> 552

<400> 552
000

<210> 553

<400> 553
000

<210> 554

<400> 554
000

<210> 555

<400> 555
000

<210> 556

<400> 556
000

<210> 557

<400> 557
000

<210> 558

<400> 558
000

<210> 559

<400> 559
000

<210> 560

<400> 560
000

<210> 561

<400> 561
000

<210> 562

<400> 562
000

<210> 563

<400> 563
000

<210> 564

<400> 564
000

<210> 565

<400> 565
000

<210> 566

<400> 566
000

<210> 567

<400> 567
000

<210> 568

<400> 568
000

<210> 569

<400> 569
000

<210> 570

<400> 570
000

<210> 571

<400> 571
000

<210> 572

<400> 572
000

<210> 573

<400> 573
000

<210> 574

<400> 574
000

<210> 575

<400> 575
000

<210> 576

<400> 576
000

<210> 577

<400> 577
000

<210> 578

<400> 578
000

<210> 579

<400> 579
000

<210> 580

<400> 580
000

<210> 581

<400> 581
000

<210> 582

<400> 582
000

<210> 583

<400> 583
000

<210> 584

<400> 584
000

<210> 585

<400> 585
000

<210> 586

<400> 586
000

<210> 587

<400> 587
000

<210> 588

<400> 588
000

<210> 589

<400> 589
000

<210> 590

<400> 590
000

<210> 591

<400> 591
000

<210> 592

<400> 592
000

<210> 593

<400> 593
000

<210> 594

<400> 594
000

<210> 595

<400> 595
000

<210> 596

<400> 596
000

<210> 597

<400> 597
000

<210> 598

<400> 598
000

<210> 599

<400> 599
000

<210> 600

<400> 600
000

<210> 601

<400> 601
000

<210> 602

<400> 602
000

<210> 603

<400> 603
000

<210> 604

<400> 604
000

<210> 605

<400> 605
000

<210> 606

<400> 606
000

<210> 607

<400> 607
000

<210> 608

<400> 608
000

<210> 609

<400> 609
000

<210> 610

<400> 610
000

<210> 611

<400> 611
000

<210> 612

<400> 612
000

<210> 613

<400> 613
000

<210> 614

<400> 614
000

<210> 615

<400> 615
000

<210> 616

<400> 616
000

<210> 617

<400> 617
000

<210> 618

<400> 618
000

<210> 619

<400> 619
000

<210> 620

<400> 620
000

<210> 621

<400> 621
000

<210> 622

<400> 622
000

<210> 623

<400> 623
000

<210> 624

<400> 624
000

<210> 625

<400> 625
000

<210> 626

<400> 626
000

<210> 627

<400> 627
000

<210> 628

<400> 628
000

<210> 629

<400> 629
000

<210> 630

<400> 630
000

<210> 631

<400> 631
000

<210> 632

<400> 632
000

<210> 633

<400> 633
000

<210> 634

<400> 634
000

<210> 635

<400> 635
000

<210> 636

<400> 636
000

<210> 637

<400> 637
000

<210> 638

<400> 638
000

<210> 639

<400> 639
000

<210> 640

<400> 640
000

<210> 641

<400> 641
000

<210> 642

<400> 642
000

<210> 643

<400> 643
000

<210> 644

<400> 644
000

<210> 645

<400> 645
000

<210> 646

<400> 646
000

<210> 647

<400> 647
000

<210> 648

<400> 648
000

<210> 649

<400> 649
000

<210> 650

<400> 650
000

<210> 651

<400> 651
000

<210> 652

<400> 652
000

<210> 653

<400> 653
000

<210> 654

<400> 654
000

<210> 655

<400> 655
000

<210> 656

<400> 656
000

<210> 657

<400> 657
000

<210> 658

<400> 658
000

<210> 659

<400> 659
000

<210> 660

<400> 660
000

<210> 661

<400> 661
000

<210> 662

<400> 662
000

<210> 663

<400> 663
000

<210> 664

<400> 664
000

<210> 665

<400> 665
000

<210> 666

<400> 666
000

<210> 667

<400> 667
000

<210> 668

<400> 668
000

<210> 669

<400> 669
000

<210> 670

<400> 670
000

<210> 671

<400> 671
000

<210> 672

<400> 672
000

<210> 673

<400> 673
000

<210> 674

<400> 674
000

<210> 675

<400> 675
000

<210> 676

<400> 676
000

<210> 677

<400> 677
000

<210> 678

<400> 678
000

<210> 679

<400> 679
000

<210> 680

<400> 680
000

<210> 681

<400> 681
000

<210> 682

<400> 682
000

<210> 683

<400> 683
000

<210> 684

<400> 684
000

<210> 685

<400> 685
000

<210> 686

<400> 686
000

<210> 687

<400> 687
000

<210> 688

<400> 688
000

<210> 689

<400> 689
000

<210> 690

<400> 690
000

<210> 691

<400> 691
000

<210> 692

<400> 692
000

<210> 693

<400> 693
000

<210> 694

<400> 694
000

<210> 695

<400> 695
000

<210> 696

<400> 696
000

<210> 697

<400> 697
000

<210> 698

<400> 698
000

<210> 699

<400> 699
000

<210> 700

<400> 700
000

<210> 701

<400> 701
000

<210> 702

<400> 702
000

<210> 703

<400> 703
000

<210> 704

<400> 704
000

<210> 705

<400> 705
000

<210> 706

<400> 706
000

<210> 707

<400> 707
000

<210> 708

<400> 708
000

<210> 709

<400> 709
000

<210> 710

<400> 710
000

<210> 711

<400> 711
000

<210> 712

<400> 712
000

<210> 713

<400> 713
000

<210> 714

<400> 714
000

<210> 715

<400> 715
000

<210> 716

<400> 716
000

<210> 717

<400> 717
000

<210> 718

<400> 718
000

<210> 719

<400> 719
000

<210> 720

<400> 720
000

<210> 721

<400> 721
000

<210> 722

<400> 722
000

<210> 723

<400> 723
000

<210> 724

<400> 724
000

<210> 725

<400> 725
000

<210> 726

<400> 726
000

<210> 727

<400> 727
000

<210> 728

<400> 728
000

<210> 729

<400> 729
000

<210> 730

<400> 730
000

<210> 731

<400> 731
000

<210> 732

<400> 732
000

<210> 733

<400> 733
000

<210> 734

<400> 734
000

<210> 735

<400> 735
000

<210> 736

<400> 736
000

<210> 737

<400> 737
000

<210> 738

<400> 738
000

<210> 739

<400> 739
000

<210> 740

<400> 740
000

<210> 741

<400> 741
000

<210> 742

<400> 742
000

<210> 743

<400> 743
000

<210> 744

<400> 744
000

<210> 745

<400> 745
000

<210> 746

<400> 746
000

<210> 747

<400> 747
000

<210> 748

<400> 748
000

<210> 749

<400> 749
000

<210> 750

<400> 750
000

<210> 751

<400> 751
000

<210> 752

<400> 752
000

<210> 753

<400> 753
000

<210> 754

<400> 754
000

<210> 755

<400> 755
000

<210> 756

<400> 756
000

<210> 757

<400> 757
000

<210> 758

<400> 758
000

<210> 759

<400> 759
000

<210> 760

<400> 760
000

<210> 761

<400> 761
000

<210> 762

<400> 762
000

<210> 763

<400> 763
000

<210> 764

<400> 764
000

<210> 765

<400> 765
000

<210> 766

<400> 766
000

<210> 767

<400> 767
000

<210> 768

<400> 768
000

<210> 769

<400> 769
000

<210> 770

<400> 770
000

<210> 771

<400> 771
000

<210> 772

<400> 772
000

<210> 773

<400> 773
000

<210> 774

<400> 774
000

<210> 775

<400> 775
000

<210> 776

<400> 776
000

<210> 777

<400> 777
000

<210> 778

<400> 778
000

<210> 779

<400> 779
000

<210> 780

<400> 780
000

<210> 781

<400> 781
000

<210> 782

<400> 782
000

<210> 783

<400> 783
000

<210> 784

<400> 784
000

<210> 785

<400> 785
000

<210> 786

<400> 786
000

<210> 787

<400> 787
000

<210> 788

<400> 788
000

<210> 789

<400> 789
000

<210> 790

<400> 790
000

<210> 791

<400> 791
000

<210> 792

<400> 792
000

<210> 793

<400> 793
000

<210> 794

<400> 794
000

<210> 795

<400> 795
000

<210> 796

<400> 796
000

<210> 797

<400> 797
000

<210> 798

<400> 798
000

<210> 799

<400> 799
000

<210> 800

<400> 800
000

<210> 801
<211> 156
<212> DNA
<213> 人工序列

<220>

<223> 人工序列描述：合成多核苷酸

<400> 801
gcggcggggg ggcggccgcg ttcgcgcgcc gcccaccagg gggtgctgcg cgccccccc 60

cgcgcatgcg cggggccccc ccccgggggg gctccgcccc cccggccccc ccccgtgcta 120

aacccaccgc gcatgcgcga ccacgccccc gccgcc 156

<210> 802
<211> 150
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成多核苷酸

<400> 802
ccgagcgtta gcgaggagtg cgaccctacc ccctgggccc acttcttcgg agccgcgcgc 60

tacgccttcg gctgcgcgcg gcacctcaga cccccgctcg tgctgacacg cttgcgcgtg 120

tcagaccact tcgggctcgc gggggtcggg 150

<210> 803
<211> 122
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成多核苷酸

<400> 803
gccgccgcgg cggcgggggg cggcgcgctg cgcgcgccgc ccagtagggg gagccatgcg 60

ccccccccg cgcatgcgcg gggcccccc ccgcgggggg ctccgccccc cggccccccc 120

cg 122

<210> 804
<211> 111
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成多核苷酸

<400> 804
cggcccagcg gcggcgcgcg cgcttcgcgc gcgcgccggg gggctccgcc ccccccgcg 60

catgcgcggg gccccccccc gcgggggggct ccgccccccg gtcccccccc g 111

<210> 805
<211> 115
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成多核苷酸

<400> 805
cggccgtgcg gcggcgcgcg cgcttcgcgc gcgcgccggg ggctgccgcc ccccccgcg 60

catgcgcgcg gggcccccccc ccgcgggggg ctccgccccc cggcccccccc ccccg 115

<210> 806
<211> 104
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成多核苷酸

<400> 806
cggcggcggc gcgcgcgcta cgcgcgcgcg ccgggggggct gccgcccccc ccccgcgcat 60

gcgcggggcc ccccccccgcg gggggctccg ccccccggcc cccc 104

<210> 807
<211> 108
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成多核苷酸

<400> 807
ggcggcggcg cgcgcgctac gcgcgcgcgc cggggagctc tgcccccccc cgcgcatgcg 60

cgcgggtccc ccccccgcgg ggggctccgc ccccccggtcc cccccccg 108

<210> 808

<400> 808
000

<210> 809

<400> 809
000

<210> 810

<400> 810
000

<210> 811

<400> 811
000

<210> 812

<400> 812
000

<210> 813

<400> 813
000

<210> 814

<400> 814
000

<210> 815

<400> 815
000

<210> 816

<400> 816
000

<210> 817

<400> 817
000

<210> 818

<400> 818
000

<210> 819

<400> 819
000

<210> 820

<400> 820
000

<210> 821

<400> 821
000

<210> 822

<400> 822
000

<210> 823

<400> 823
000

<210> 824

<400> 824
000

<210> 825

<400> 825
000

<210> 826

<400> 826
000

<210> 827

<400> 827
000

<210> 828

<400> 828
000

<210> 829
<211> 11
<212> PRT
<213> 人工序列

<220>
<223> 人工序列描述：合成肽

<220>
<221> MOD_RES
<222> (4)..(5)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (7)..(7)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (9)..(10)
<223> 任何胺基酸

<400> 829
Tyr Asn Pro Xaa Xaa Asp Xaa Gly Xaa Xaa Asn

1 5 10

<210> 830

<400> 830
000

<210> 831

<400> 831
000

<210> 832

<400> 832
000

<210> 833

<400> 833
000

<210> 834

<400> 834
000

<210> 835

<400> 835
000

<210> 836

<400> 836
000

<210> 837

<400> 837
000

<210> 838

<400> 838
000

<210> 839

<400> 839
000

<210> 840

<400> 840
000

<210> 841

<400> 841
000

<210> 842

<400> 842
000

<210> 843

<400> 843
000

<210> 844

<400> 844
000

<210> 845

<400> 845
000

<210> 846

<400> 846
000

<210> 847

<400> 847
000

<210> 848

<400> 848
000

<210> 849

<400> 849
000

<210> 850

<400> 850
000

<210> 851

<400> 851
000

<210> 852

<400> 852
000

<210> 853

<400> 853
000

<210> 854

<400> 854

000

<210> 855

<400> 855
000

<210> 856

<400> 856
000

<210> 857

<400> 857
000

<210> 858

<400> 858
000

<210> 859

<400> 859
000

<210> 860

<400> 860
000

<210> 861

<400> 861
000

<210> 862

<400> 862
000

<210> 863

<400> 863
000

<210> 864

<400> 864
000

<210> 865

<400> 865
000

<210> 866

<400> 866
000

<210> 867

<400> 867
000

<210> 868

<400> 868
000

<210> 869

<400> 869
000

<210> 870

<400> 870
000

<210> 871

<400> 871
000

<210> 872

<400> 872
000

<210> 873

<400> 873
000

<210> 874

<400> 874
000

<210> 875

<400> 875
000

<210> 876

<400> 876
000

<210> 877

<400> 877
000

<210> 878

<400> 878
000

<210> 879

<400> 879

000

<210> 880

<400> 880
000

<210> 881

<400> 881
000

<210> 882

<400> 882
000

<210> 883

<400> 883
000

<210> 884

<400> 884
000

<210> 885

<400> 885
000

<210> 886
<211> 3176
<212> DNA
<213> γ細環病毒屬

<400> 886
taaaatggcg ggagccaatc attttatact ttcactttcc aattaaaaat ggccacgtca 60

caaacaaggg gtggagccat ttaaactata taactaagtg gggtggcgaa tggctgagtt 120

taccccgcta gacggtgcag ggaccggatc gagcgcagcg aggaggtccc cggctgccca 180

tgggcgggag ccgaggtgag tgaaaccacc gaggtctagg ggcaattcgg gctagggcag 240

tctagcggaa cgggcaagaa acttaaaaca atatttgttt tacagatggt tagtatatcc 300

tcaagtgatt tttttaagaa aacgaaattt aatgaggaga cgcagaacca agtatggatg 360

tctcaaattg ctgactctca tgataatatc tgcagttgct ggcatccatt tgctcacctt 420

cttgcttcca tatttcctcc tggccacaaa gatcgtgatc ttactattaa ccaaattctt 480

ctaagagatt ataaagaaaa atgccattct ggtggagaag aaggagaaaa ttctggacca 540

acaacaggtt taattacacc aaaagaagaa gatatagaaa aagatggccc agaaggcgcc 600

gcagaagaag accatacaga cgccctgttc gccgccgccg tagaaaactt cgaaaggtaa 660

agagaaaaaa aaaatcttta attgttagac aatggcaacc agacagtata agaacttgta 720

aaattatagg acagtcagct atagttgttg gggctgaagg aaagcaaatg tactgttata 780

ctgtcaataa gttaattaat gtgcccccaa aaacaccata tgggggaggc tttggagtag 840

accaatacac actgaaatac ttatatgaag aatacagatt tgcacaaaac atttggacac 900

aatctaatgt actgaaagac ttatgcagat acataaatgt taagctaata ttctacagag 960

acaacaaaac agactttgtc ctttcctatg acagaaaccc acctttcaa ctaacaaaat 1020

ttacataccc aggagcacac ccacaacaaa tcatgcttca aaaacaccac aaattcatac 1080

tatcacaaat gacaaagcct aatggaagac taacaaaaaa actcaaaatt aaacctccta 1140

aacaaatgct ttctaaatgg ttcttttcaa aacaattctg taaataccct ttactatctc 1200

ttaaagcttc tgcactagac cttaggcact cttacctagg ctgctgtaat gaaaatccac 1260

aggtattttt ttattattta aaccatggat actacacaat aacaaactgg ggagcacaat 1320

cctcaacagc atacagacct aactccaagg tgacagacac aacatactac agatacaaaa 1380

atgacagaaa aaatattaac attaaaagcc atgaatacga aaaaagtata tcatatgaaa 1440

acggttattt tcaatctagt ttcttacaaa cacagtgcat atataccagt gagcgtggtg 1500

aagcctgtat agcagaaaaa ccactaggaa tagctattta caatccagta aaagacaatg 1560

gagatggtaa tatgatatac cttgtaagca ctctagcaaa cacttgggac cagcctccaa 1620

aagacagtgc tattttaata caaggagtac ccatatggct aggcttattt ggatatttag 1680

actactgtag acaaattaaa gctgacaaaa catggctaga cagtcatgta ctagtaattc 1740

aaagtcctgc tatttttact tacccaaatc caggagcagg caaatggtat tgtccactat 1800

cacaaagttt tataaatggc aatggtccgt ttaatcaacc acctacactg ctacaaaaag 1860

caaagtggtt tccacaaata caataccaac aagaaattat taatagcttt gtagaatcag 1920

gaccatttgt tcccaaatat gcaaatcaaa ctgaaagcaa ctgggaacta aaatataaat 1980

atgtttttac atttaagtgg ggtggaccac aattccatga accagaaatt gctgacccta 2040

gcaaacaaga gcagtatgat gtccccgata ctttctacca aacaatacaa attgaagatc 2100

cagaaggaca agaccccaga tctctcatcc atgattggga ctacagacga ggctttatta 2160

aagaaagatc tcttaaaaga atgtcaactt acttctcaac tcatacagat cagcaagcaa 2220

cttcagagga agacattccc aaaaagaaaa agagaattgg accccaactc acagtcccac 2280

aacaaaaaga agaggagaca ctgtcatgtc tcctctctct ctgcaaaaaa gataccttcc 2340

aagaaacaga gacacaagaa gacctccagc agctcatcaa gcagcagcag gagcagcagc 2400

tcctcctcaa gagaaacatc ctccagctca tccacaaact aaaagagaat caacaaatgc 2460

ttcagcttca cacaggcatg ttaccttaac cagatttaaa cctggatttg aagagcaaac 2520

agagagagaa ttagcaatta tatttcatag gccccctaga acctacaaag aggaccttcc 2580

attctatccc tggctaccac ctgcacccct tgtacaattt aaccttaact tcaaaggcta 2640

ggccaacaat gtacacttag taaagcatgt ttattaaagc acaacccccа aaataaatgt 2700

aaaaataaaa aaaaaaaaaa aaaaataaaa aattgcaaaa attcggcgct cgcgcgcatg 2760

tgcgcctctg gcgcaaatca cgcaacgctc gcgcgcccgc gtatgtctct ttaccacgca 2820

cctagattgg ggtgcgcgcg ctagcgcgcg caccccaatg cgccccgccc tcgttccgac 2880

ccgcttgcgc gggtcggacc acttcgggct cgggggggcg cgcctgcggc gctttttac 2940

taaacagact ccgagccgcc atttggcccc ctaagctccg ccccccctcat gaatattcat 3000

aaaggaaacc acataattag aattgccgac cacaaactgc catatgctaa ttagttcccc 3060

ttttacaaag taaaagggga agtgaacata gccccacacc cgcaggggca aggccccgca 3120

cccctacgtc actaaccacg cccccgccgc catcttgggt gcggcagggc gggggc 3176

<210> 887
<211> 124
<212> PRT
<213> γ細環病毒屬

<400> 887
Met Val Ser Ile Ser Ser Ser Asp Phe Phe Lys Lys Thr Lys Phe Asn
1 5 10 15

Glu Glu Thr Gln Asn Gln Val Trp Met Ser Gln Ile Ala Asp Ser His
20 25 30

Asp Asn Ile Cys Ser Cys Trp His Pro Phe Ala His Leu Leu Ala Ser
35 40 45

Ile Phe Pro Pro Gly His Lys Asp Arg Asp Leu Thr Ile Asn Gln Ile
50 55 60

Leu Leu Arg Asp Tyr Lys Glu Lys Cys His Ser Gly Gly Glu Glu Gly
65 70 75 80

Glu Asn Ser Gly Pro Thr Thr Gly Leu Ile Thr Pro Lys Glu Glu Asp
85 90 95

Ile Glu Lys Asp Gly Pro Glu Gly Ala Ala Glu Glu Asp His Thr Asp
100 105 110

Ala Leu Phe Ala Ala Ala Val Glu Asn Phe Glu Arg
115 120

<210> 888
<211> 271
<212> PRT
<213> γ細環病毒屬

<400> 888
Met Val Ser Ile Ser Ser Ser Asp Phe Phe Lys Lys Thr Lys Phe Asn
1 5 10 15

Glu Glu Thr Gln Asn Gln Val Trp Met Ser Gln Ile Ala Asp Ser His
20 25 30

Asp Asn Ile Cys Ser Cys Trp His Pro Phe Ala His Leu Leu Ala Ser
35 40 45

Ile Phe Pro Pro Gly His Lys Asp Arg Asp Leu Thr Ile Asn Gln Ile
50 55 60

Leu Leu Arg Asp Tyr Lys Glu Lys Cys His Ser Gly Gly Glu Glu Gly
65 70 75 80

Glu Asn Ser Gly Pro Thr Thr Gly Leu Ile Thr Pro Lys Glu Glu Asp
85 90 95

Ile Glu Lys Asp Gly Pro Glu Gly Ala Ala Glu Glu Asp His Thr Asp
100 105 110

Ala Leu Phe Ala Ala Ala Val Glu Asn Phe Glu Ser Gly Val Asp His
115 120 125

Asn Ser Met Asn Gln Lys Leu Leu Thr Leu Ala Asn Lys Ser Ser Met
130 135 140

Met Ser Pro Ile Leu Ser Thr Lys Gln Tyr Lys Leu Lys Ile Gln Lys
145 150 155 160

Asp Lys Thr Pro Asp Leu Ser Ser Met Ile Gly Thr Thr Asp Glu Ala
165 170 175

Leu Leu Lys Lys Asp Leu Leu Lys Glu Cys Gln Leu Thr Ser Gln Leu
180 185 190

Ile Gln Ile Ser Lys Gln Leu Gln Arg Lys Thr Phe Pro Lys Arg Lys
195 200 205

Arg Glu Leu Asp Pro Asn Ser Gln Ser His Asn Lys Lys Lys Arg Arg
210 215 220

His Cys His Val Ser Ser Leu Ser Ala Lys Lys Ile Pro Ser Lys Lys
225 230 235 240

Gln Arg His Lys Lys Thr Ser Ser Ser Ser Ser Ser Ser Ser Arg Ser
245 250 255

Ser Ser Ser Ser Ser Arg Glu Thr Ser Ser Ser Ser Ser Thr Asn
260 265 270

<210> 889
<211> 267
<212> PRT
<213> γ細環病毒屬

<400> 889
Met Val Ser Ile Ser Ser Ser Asp Phe Phe Lys Lys Thr Lys Phe Asn
1 5 10 15

Glu Glu Thr Gln Asn Gln Val Trp Met Ser Gln Ile Ala Asp Ser His
20 25 30

Asp Asn Ile Cys Ser Cys Trp His Pro Phe Ala His Leu Leu Ala Ser
35 40 45

Ile Phe Pro Pro Gly His Lys Asp Arg Asp Leu Thr Ile Asn Gln Ile
50 55 60

Leu Leu Arg Asp Tyr Lys Glu Lys Cys His Ser Gly Gly Glu Glu Gly
65 70 75 80

Glu Asn Ser Gly Pro Thr Thr Gly Leu Ile Thr Pro Lys Glu Glu Asp
85 90 95

Ile Glu Lys Asp Gly Pro Glu Gly Ala Ala Glu Glu Asp His Thr Asp
100 105 110

Ala Leu Phe Ala Ala Ala Val Glu Asn Phe Glu Arg Ser Ala Ser Asn
115 120 125

Phe Arg Gly Arg His Ser Gln Lys Glu Lys Glu Asn Trp Thr Pro Thr
130 135 140

His Ser Pro Thr Thr Lys Arg Arg Gly Asp Thr Val Met Ser Pro Leu
145 150 155 160

Ser Leu Gln Lys Arg Tyr Leu Pro Arg Asn Arg Asp Thr Arg Arg Pro
165 170 175

Pro Ala Ala His Gln Ala Ala Ala Gly Ala Ala Ala Pro Pro Gln Glu
180 185 190

Lys His Pro Pro Ala His Pro Gln Thr Lys Arg Glu Ser Thr Asn Ala
195 200 205

Ser Ala Ser His Arg His Val Thr Leu Thr Arg Phe Lys Pro Gly Phe
210 215 220

Glu Glu Gln Thr Glu Arg Glu Leu Ala Ile Ile Phe His Arg Pro Pro
225 230 235 240

Arg Thr Tyr Lys Glu Asp Leu Pro Phe Tyr Pro Trp Leu Pro Pro Ala
245 250 255

Pro Leu Val Gln Phe Asn Leu Asn Phe Lys Gly
260 265

<210> 890
<211> 50
<212> PRT
<213> γ細環病毒屬

<400> 890
Met Arg Arg Arg Arg Thr Lys Tyr Gly Cys Leu Lys Leu Leu Thr Leu
1 5 10 15

Met Ile Ile Ser Ala Val Ala Gly Ile His Leu Leu Thr Phe Leu Leu
20 25 30

Pro Tyr Phe Leu Leu Ala Thr Lys Ile Val Ile Leu Leu Leu Thr Lys
35 40 45

Phe Phe
50

<210> 891
<211> 662
<212> PRT
<213> γ細環病毒屬

<400> 891
Met Pro Phe Trp Trp Arg Arg Arg Arg Lys Phe Trp Thr Asn Asn Arg
1 5 10 15

Phe Asn Tyr Thr Lys Arg Arg Arg Tyr Arg Lys Arg Trp Pro Arg Arg
20 25 30

Arg Arg Arg Arg Arg Pro Tyr Arg Arg Pro Val Arg Arg Arg Arg Arg
35 40 45

Lys Leu Arg Lys Val Lys Arg Lys Lys Lys Ser Leu Ile Val Arg Gln
50 55 60

Trp Gln Pro Asp Ser Ile Arg Thr Cys Lys Ile Ile Gly Gln Ser Ala
65 70 75 80

Ile Val Val Gly Ala Glu Gly Lys Gln Met Tyr Cys Tyr Thr Val Asn
85 90 95

Lys Leu Ile Asn Val Pro Pro Lys Thr Pro Tyr Gly Gly Gly Phe Gly
100 105 110

Val Asp Gln Tyr Thr Leu Lys Tyr Leu Tyr Glu Glu Tyr Arg Phe Ala
115 120 125

Gln Asn Ile Trp Thr Gln Ser Asn Val Leu Lys Asp Leu Cys Arg Tyr
130 135 140

Ile Asn Val Lys Leu Ile Phe Tyr Arg Asp Asn Lys Thr Asp Phe Val
145 150 155 160

Leu Ser Tyr Asp Arg Asn Pro Pro Phe Gln Leu Thr Lys Phe Thr Tyr
165 170 175

Pro Gly Ala His Pro Gln Gln Ile Met Leu Gln Lys His His Lys Phe
180 185 190

Ile Leu Ser Gln Met Thr Lys Pro Asn Gly Arg Leu Thr Lys Lys Leu
195 200 205

Lys Ile Lys Pro Pro Lys Gln Met Leu Ser Lys Trp Phe Phe Ser Lys
210 215 220

Gln Phe Cys Lys Tyr Pro Leu Leu Ser Leu Lys Ala Ser Ala Leu Asp
225 230 235 240

Leu Arg His Ser Tyr Leu Gly Cys Cys Asn Glu Asn Pro Gln Val Phe
245 250 255

Phe Tyr Tyr Leu Asn His Gly Tyr Tyr Thr Ile Thr Asn Trp Gly Ala
260 265 270

Gln Ser Ser Thr Ala Tyr Arg Pro Asn Ser Lys Val Thr Asp Thr Thr
275 280 285

Tyr Tyr Arg Tyr Lys Asn Asp Arg Lys Asn Ile Asn Ile Lys Ser His
290 295 300

Glu Tyr Glu Lys Ser Ile Ser Tyr Glu Asn Gly Tyr Phe Gln Ser Ser
305 310 315 320

Phe Leu Gln Thr Gln Cys Ile Tyr Thr Ser Glu Arg Gly Glu Ala Cys
325 330 335

Ile Ala Glu Lys Pro Leu Gly Ile Ala Ile Tyr Asn Pro Val Lys Asp
340 345 350

Asn Gly Asp Gly Asn Met Ile Tyr Leu Val Ser Thr Leu Ala Asn Thr
355 360 365

Trp Asp Gln Pro Pro Lys Asp Ser Ala Ile Leu Ile Gln Gly Val Pro
370 375 380

Ile Trp Leu Gly Leu Phe Gly Tyr Leu Asp Tyr Cys Arg Gln Ile Lys
385 390 395 400

Ala Asp Lys Thr Trp Leu Asp Ser His Val Leu Val Ile Gln Ser Pro
405 410 415

Ala Ile Phe Thr Tyr Pro Asn Pro Gly Ala Gly Lys Trp Tyr Cys Pro
420 425 430

Leu Ser Gln Ser Phe Ile Asn Gly Asn Gly Pro Phe Asn Gln Pro Pro
435 440 445

Thr Leu Leu Gln Lys Ala Lys Trp Phe Pro Gln Ile Gln Tyr Gln Gln
450 455 460

Glu Ile Ile Asn Ser Phe Val Glu Ser Gly Pro Phe Val Pro Lys Tyr
465 470 475 480

Ala Asn Gln Thr Glu Ser Asn Trp Glu Leu Lys Tyr Lys Tyr Val Phe
485 490 495

Thr Phe Lys Trp Gly Gly Pro Gln Phe His Glu Pro Glu Ile Ala Asp
500 505 510

Pro Ser Lys Gln Glu Gln Tyr Asp Val Pro Asp Thr Phe Tyr Gln Thr
515 520 525

Ile Gln Ile Glu Asp Pro Glu Gly Gln Asp Pro Arg Ser Leu Ile His
530 535 540

```
Asp Trp Asp Tyr Arg Arg Gly Phe Ile Lys Glu Arg Ser Leu Lys Arg
545                 550                 555                 560

Met Ser Thr Tyr Phe Ser Thr His Thr Asp Gln Gln Ala Thr Ser Glu
                565                 570                 575

Glu Asp Ile Pro Lys Lys Lys Lys Arg Ile Gly Pro Gln Leu Thr Val
            580                 585                 590

Pro Gln Gln Lys Glu Glu Glu Thr Leu Ser Cys Leu Leu Ser Leu Cys
        595                 600                 605

Lys Lys Asp Thr Phe Gln Glu Thr Glu Thr Gln Glu Asp Leu Gln Gln
    610                 615                 620

Leu Ile Lys Gln Gln Gln Glu Gln Gln Leu Leu Leu Lys Arg Asn Ile
625                 630                 635                 640

Leu Gln Leu Ile His Lys Leu Lys Glu Asn Gln Gln Met Leu Gln Leu
                645                 650                 655

His Thr Gly Met Leu Pro
            660
```

<210> 892
<211> 215
<212> PRT
<213> γ細環病毒屬

<400> 892

```
Met Pro Phe Trp Trp Arg Arg Arg Arg Lys Phe Trp Thr Asn Asn Arg
1               5                   10                  15

Phe Asn Tyr Thr Lys Arg Arg Arg Tyr Arg Lys Arg Trp Pro Arg Arg
            20                  25                  30

Arg Arg Arg Arg Arg Pro Tyr Arg Arg Pro Val Arg Arg Arg Arg Arg
        35                  40                  45
```

Lys Leu Arg Lys Trp Gly Gly Pro Gln Phe His Glu Pro Glu Ile Ala
50 55 60

Asp Pro Ser Lys Gln Glu Gln Tyr Asp Val Pro Asp Thr Phe Tyr Gln
65 70 75 80

Thr Ile Gln Ile Glu Asp Pro Glu Gly Gln Asp Pro Arg Ser Leu Ile
85 90 95

His Asp Trp Asp Tyr Arg Arg Gly Phe Ile Lys Glu Arg Ser Leu Lys
100 105 110

Arg Met Ser Thr Tyr Phe Ser Thr His Thr Asp Gln Gln Ala Thr Ser
115 120 125

Glu Glu Asp Ile Pro Lys Lys Lys Lys Arg Ile Gly Pro Gln Leu Thr
130 135 140

Val Pro Gln Gln Lys Glu Glu Glu Thr Leu Ser Cys Leu Leu Ser Leu
145 150 155 160

Cys Lys Lys Asp Thr Phe Gln Glu Thr Glu Thr Gln Glu Asp Leu Gln
165 170 175

Gln Leu Ile Lys Gln Gln Gln Glu Gln Gln Leu Leu Leu Lys Arg Asn
180 185 190

Ile Leu Gln Leu Ile His Lys Leu Lys Glu Asn Gln Gln Met Leu Gln
195 200 205

Leu His Thr Gly Met Leu Pro
210 215

<210> 893
<211> 129
<212> PRT
<213> γ細環病毒屬

<400> 893
Met Pro Phe Trp Trp Arg Arg Arg Arg Lys Phe Trp Thr Asn Asn Arg
1 5 10 15

Phe Asn Tyr Thr Lys Arg Arg Arg Tyr Arg Lys Arg Trp Pro Arg Arg
20 25 30

Arg Arg Arg Arg Arg Pro Tyr Arg Arg Pro Val Arg Arg Arg Arg Arg
35 40 45

Lys Leu Arg Lys Ile Ser Lys Gln Leu Gln Arg Lys Thr Phe Pro Lys
50 55 60

Arg Lys Arg Glu Leu Asp Pro Asn Ser Gln Ser His Asn Lys Lys Lys
65 70 75 80

Arg Arg His Cys His Val Ser Ser Leu Ser Ala Lys Lys Ile Pro Ser
85 90 95

Lys Lys Gln Arg His Lys Lys Thr Ser Ser Ser Ser Ser Ser Ser Ser
100 105 110

Arg Ser Ser Ser Ser Ser Ser Arg Glu Thr Ser Ser Ser Ser Ser Thr
115 120 125

Asn

<210> 894

<400> 894
000

<210> 895

<400> 895
000

<210> 896

<400> 896
000

<210> 897

<400> 897
000

<210> 898

<400> 898
000

<210> 899

<400> 899
000

<210> 900

<400> 900
000

<210> 901

<400> 901
000

<210> 902

<400> 902
000

<210> 903

<400> 903
000

<210> 904

<400> 904
000

<210> 905

<400> 905
000

<210> 906

<400> 906
000

<210> 907

<400> 907
000

<210> 908

<400> 908
000

<210> 909

<400> 909
000

<210> 910

<400> 910
000

<210> 911

<400> 911
000

<210> 912

<400> 912
000

<210> 913

<400> 913
000

<210> 914

<400> 914
000

<210> 915

<400> 915
000

<210> 916

<400> 916
000

<210> 917

<400> 917
000

<210> 918

<400> 918
000

<210> 919

<400> 919
000

<210> 920

<400> 920
000

<210> 921

<400> 921
000

<210> 922

<400> 922
000

<210> 923

<400> 923
000

<210> 924

<400> 924
000

<210> 925
<211> 662
<212> PRT
<213> γ細環病毒屬

<400> 925

```
Met Pro Phe Trp Trp Arg Arg Arg Arg Lys Phe Trp Thr Asn Asn Arg
1               5                  10                  15

Phe Asn Tyr Thr Lys Arg Arg Arg Tyr Arg Lys Arg Trp Pro Arg Arg
            20                  25                  30

Arg Arg Arg Arg Arg Pro Tyr Arg Arg Pro Val Arg Arg Arg Arg Arg
        35                  40                  45

Lys Leu Arg Lys Val Lys Arg Lys Lys Lys Ser Leu Ile Val Arg Gln
    50                  55                  60

Trp Gln Pro Asp Ser Ile Arg Thr Cys Lys Ile Ile Gly Gln Ser Ala
65                  70                  75                  80
```

Ile Val Val Gly Ala Glu Gly Lys Gln Met Tyr Cys Tyr Thr Val Asn
85 90 95

Lys Leu Ile Asn Val Pro Pro Lys Thr Pro Tyr Gly Gly Gly Phe Gly
100 105 110

Val Asp Gln Tyr Thr Leu Lys Tyr Leu Tyr Glu Glu Tyr Arg Phe Ala
115 120 125

Gln Asn Ile Trp Thr Gln Ser Asn Val Leu Lys Asp Leu Cys Arg Tyr
130 135 140

Ile Asn Val Lys Leu Ile Phe Tyr Arg Asp Asn Lys Thr Asp Phe Val
145 150 155 160

Leu Ser Tyr Asp Arg Asn Pro Pro Phe Gln Leu Thr Lys Phe Thr Tyr
165 170 175

Pro Gly Ala His Pro Gln Gln Ile Met Leu Gln Lys His His Lys Phe
180 185 190

Ile Leu Ser Gln Met Thr Lys Pro Asn Gly Arg Leu Thr Lys Lys Leu
195 200 205

Lys Ile Lys Pro Pro Lys Gln Met Leu Ser Lys Trp Phe Phe Ser Lys
210 215 220

Gln Phe Cys Lys Tyr Pro Leu Leu Ser Leu Lys Ala Ser Ala Leu Asp
225 230 235 240

Leu Arg His Ser Tyr Leu Gly Cys Cys Asn Glu Asn Pro Gln Val Phe
245 250 255

Phe Tyr Tyr Leu Asn His Gly Tyr Tyr Thr Ile Thr Asn Trp Gly Ala
260 265 270

Gln Ser Ser Thr Ala Tyr Arg Pro Asn Ser Lys Val Thr Asp Thr Thr
275 280 285

Tyr Tyr Arg Tyr Lys Asn Asp Arg Lys Asn Ile Asn Ile Lys Ser His
290 295 300

Glu Tyr Glu Lys Ser Ile Ser Tyr Glu Asn Gly Tyr Phe Gln Ser Ser
305 310 315 320

Phe Leu Gln Thr Gln Cys Ile Tyr Thr Ser Glu Arg Gly Glu Ala Cys
325 330 335

Ile Ala Glu Lys Pro Leu Gly Ile Ala Ile Tyr Asn Pro Val Lys Asp
340 345 350

Asn Gly Asp Gly Asn Met Ile Tyr Leu Val Ser Thr Leu Ala Asn Thr
355 360 365

Trp Asp Gln Pro Pro Lys Asp Ser Ala Ile Leu Ile Gln Gly Val Pro
370 375 380

Ile Trp Leu Gly Leu Phe Gly Tyr Leu Asp Tyr Cys Arg Gln Ile Lys
385 390 395 400

Ala Asp Lys Thr Trp Leu Asp Ser His Val Leu Val Ile Gln Ser Pro
405 410 415

Ala Ile Phe Thr Tyr Pro Asn Pro Gly Ala Gly Lys Trp Tyr Cys Pro
420 425 430

Leu Ser Gln Ser Phe Ile Asn Gly Asn Gly Pro Phe Asn Gln Pro Pro
435 440 445

Thr Leu Leu Gln Lys Ala Lys Trp Phe Pro Gln Ile Gln Tyr Gln Gln
450 455 460

Glu Ile Ile Asn Ser Phe Val Glu Ser Gly Pro Phe Val Pro Lys Tyr
465 470 475 480

Ala Asn Gln Thr Glu Ser Asn Trp Glu Leu Lys Tyr Lys Tyr Val Phe
485 490 495

Thr Phe Lys Trp Gly Gly Pro Gln Phe His Glu Pro Glu Ile Ala Asp
500 505 510

Pro Ser Lys Gln Glu Gln Tyr Asp Val Pro Asp Thr Phe Tyr Gln Thr
515 520 525

Ile Gln Ile Glu Asp Pro Glu Gly Gln Asp Pro Arg Ser Leu Ile His
530 535 540

Asp Trp Asp Tyr Arg Arg Gly Phe Ile Lys Glu Arg Ser Leu Lys Arg
545 550 555 560

Met Ser Thr Tyr Phe Ser Thr His Thr Asp Gln Gln Ala Thr Ser Glu
565 570 575

Glu Asp Ile Pro Lys Lys Lys Lys Arg Ile Gly Pro Gln Leu Thr Val
580 585 590

Pro Gln Gln Lys Glu Glu Glu Thr Leu Ser Cys Leu Leu Ser Leu Cys
595 600 605

Lys Lys Asp Thr Phe Gln Glu Thr Glu Thr Gln Glu Asp Leu Gln Gln
610 615 620

Leu Ile Lys Gln Gln Gln Glu Gln Gln Leu Leu Leu Lys Arg Asn Ile
625 630 635 640

Leu Gln Leu Ile His Lys Leu Lys Glu Asn Gln Gln Met Leu Gln Leu
645 650 655

His Thr Gly Met Leu Pro
660

<210> 926
<211> 58

<212> PRT
<213> γ細環病毒屬

<400> 926
Met Pro Phe Trp Trp Arg Arg Arg Arg Lys Phe Trp Thr Asn Asn Arg
1 5 10 15

Phe Asn Tyr Thr Lys Arg Arg Arg Tyr Arg Lys Arg Trp Pro Arg Arg
20 25 30

Arg Arg Arg Arg Arg Pro Tyr Arg Arg Pro Val Arg Arg Arg Arg Arg
35 40 45

Lys Leu Arg Lys Val Lys Arg Lys Lys Lys
50 55

<210> 927
<211> 202
<212> PRT
<213> γ細環病毒屬

<400> 927
Ser Leu Ile Val Arg Gln Trp Gln Pro Asp Ser Ile Arg Thr Cys Lys
1 5 10 15

Ile Ile Gly Gln Ser Ala Ile Val Val Gly Ala Glu Gly Lys Gln Met
20 25 30

Tyr Cys Tyr Thr Val Asn Lys Leu Ile Asn Val Pro Pro Lys Thr Pro
35 40 45

Tyr Gly Gly Gly Phe Gly Val Asp Gln Tyr Thr Leu Lys Tyr Leu Tyr
50 55 60

Glu Glu Tyr Arg Phe Ala Gln Asn Ile Trp Thr Gln Ser Asn Val Leu
65 70 75 80

Lys Asp Leu Cys Arg Tyr Ile Asn Val Lys Leu Ile Phe Tyr Arg Asp
85 90 95

Asn Lys Thr Asp Phe Val Leu Ser Tyr Asp Arg Asn Pro Pro Phe Gln
100 105 110

Leu Thr Lys Phe Thr Tyr Pro Gly Ala His Pro Gln Gln Ile Met Leu
115 120 125

Gln Lys His His Lys Phe Ile Leu Ser Gln Met Thr Lys Pro Asn Gly
130 135 140

Arg Leu Thr Lys Lys Leu Lys Ile Lys Pro Pro Lys Gln Met Leu Ser
145 150 155 160

Lys Trp Phe Phe Ser Lys Gln Phe Cys Lys Tyr Pro Leu Leu Ser Leu
165 170 175

Lys Ala Ser Ala Leu Asp Leu Arg His Ser Tyr Leu Gly Cys Cys Asn
180 185 190

Glu Asn Pro Gln Val Phe Phe Tyr Tyr Leu
195 200

<210> 928
<211> 79
<212> PRT
<213> γ細環病毒屬

<400> 928
Asn His Gly Tyr Tyr Thr Ile Thr Asn Trp Gly Ala Gln Ser Ser Thr
1 5 10 15

Ala Tyr Arg Pro Asn Ser Lys Val Thr Asp Thr Thr Tyr Tyr Arg Tyr
20 25 30

Lys Asn Asp Arg Lys Asn Ile Asn Ile Lys Ser His Glu Tyr Glu Lys
35 40 45

Ser Ile Ser Tyr Glu Asn Gly Tyr Phe Gln Ser Ser Phe Leu Gln Thr
50 55 60

Gln Cys Ile Tyr Thr Ser Glu Arg Gly Glu Ala Cys Ile Ala Glu
65 70 75

<210> 929
<211> 160
<212> PRT
<213> γ細環病毒屬

<400> 929
Lys Pro Leu Gly Ile Ala Ile Tyr Asn Pro Val Lys Asp Asn Gly Asp
1 5 10 15

Gly Asn Met Ile Tyr Leu Val Ser Thr Leu Ala Asn Thr Trp Asp Gln
20 25 30

Pro Pro Lys Asp Ser Ala Ile Leu Ile Gln Gly Val Pro Ile Trp Leu
35 40 45

Gly Leu Phe Gly Tyr Leu Asp Tyr Cys Arg Gln Ile Lys Ala Asp Lys
50 55 60

Thr Trp Leu Asp Ser His Val Leu Val Ile Gln Ser Pro Ala Ile Phe
65 70 75 80

Thr Tyr Pro Asn Pro Gly Ala Gly Lys Trp Tyr Cys Pro Leu Ser Gln
85 90 95

Ser Phe Ile Asn Gly Asn Gly Pro Phe Asn Gln Pro Pro Thr Leu Leu
100 105 110

Gln Lys Ala Lys Trp Phe Pro Gln Ile Gln Tyr Gln Gln Glu Ile Ile
115 120 125

Asn Ser Phe Val Glu Ser Gly Pro Phe Val Pro Lys Tyr Ala Asn Gln
130 135 140

Thr Glu Ser Asn Trp Glu Leu Lys Tyr Lys Tyr Val Phe Thr Phe Lys
145 150 155 160

<210> 930
<211> 163
<212> PRT
<213> γ細環病毒屬

<400> 930
Trp Gly Gly Pro Gln Phe His Glu Pro Glu Ile Ala Asp Pro Ser Lys
1 5 10 15

Gln Glu Gln Tyr Asp Val Pro Asp Thr Phe Tyr Gln Thr Ile Gln Ile
20 25 30

Glu Asp Pro Glu Gly Gln Asp Pro Arg Ser Leu Ile His Asp Trp Asp
35 40 45

Tyr Arg Arg Gly Phe Ile Lys Glu Arg Ser Leu Lys Arg Met Ser Thr
50 55 60

Tyr Phe Ser Thr His Thr Asp Gln Gln Ala Thr Ser Glu Glu Asp Ile
65 70 75 80

Pro Lys Lys Lys Lys Arg Ile Gly Pro Gln Leu Thr Val Pro Gln Gln
85 90 95

Lys Glu Glu Glu Thr Leu Ser Cys Leu Leu Ser Leu Cys Lys Lys Asp
100 105 110

Thr Phe Gln Glu Thr Glu Thr Gln Glu Asp Leu Gln Gln Leu Ile Lys
115 120 125

Gln Gln Gln Glu Gln Gln Leu Leu Leu Lys Arg Asn Ile Leu Gln Leu
130 135 140

Ile His Lys Leu Lys Glu Asn Gln Gln Met Leu Gln Leu His Thr Gly
145 150 155 160

Met Leu Pro

<210> 931

<400> 931
000

<210> 932

<400> 932
000

<210> 933

<400> 933
000

<210> 934

<400> 934
000

<210> 935

<400> 935
000

<210> 936

<400> 936
000

<210> 937

<400> 937
000

<210> 938

<400> 938
000

<210> 939

<400> 939
000

<210> 940

<400> 940
000

<210> 941

<400> 941
000

<210> 942

<400> 942
000

<210> 943

<400> 943
000

<210> 944

<400> 944
000

<210> 945

<400> 945
000

<210> 946

<400> 946
000

<210> 947

<400> 947
000

<210> 948

<400> 948
000

<210> 949
<211> 21
<212> PRT
<213> 人工序列

<220>
<223> 人工序列描述：合成肽

<220>
<221> MOD_RES
<222> (1)..(1)
<223> W或F

<220>
<221> MOD_RES
<222> (2)..(8)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (10)..(12)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (14)..(14)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (16)..(20)
<223> 任何胺基酸

<400> 949
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa His Xaa Xaa Xaa Cys Xaa Cys Xaa
1 5 10 15

Xaa Xaa Xaa Xaa His
20

<210> 950
<211> 22
<212> PRT
<213> 人工序列

<220>
<223> 人工序列描述：合成肽

<220>
<221> MOD_RES
<222> (6)..(7)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (9)..(9)
<223> 任何胺基酸

<220>
<221> MOD_RES
<222> (15)..(16)
<223> 任何胺基酸

<400> 950
Tyr Asn Cys Ser Pro Xaa Xaa Asp Xaa Gly Ala Ser Lys Arg Xaa Xaa
1 5 10 15

Asn Thr Ser Val Ala Lys
20

<210> 951
<211> 51
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：合成寡核苷酸

<220>
<221> modified_base
<222> (45)..(45)
<223> a、c、t、g、未知或其他

<400> 951
aggtgagtga aaccaccgaa gtcaaggggc aattcgggct agggncagtc t 51

<210> 952
<211> 50
<212> DNA
<213> α細環病毒屬

<400> 952
aggtgagttt acacaccgca gtcaaggggc aattcgggct cgggactggc 50

<210> 953
<211> 50
<212> DNA
<213> β細環病毒屬

<400> 953
aggtgagtga aaccaccgaa gtcaaggggc aattcgggct agatcagtct 50

<210> 954
<211> 50
<212> DNA
<213> γ細環病毒屬

<400> 954
aggtgagtga aaccaccgag gtctaggggc aattcgggct agggcagtct 50

<210> 955
<211> 6
<212> PRT
<213> 人工序列

<220>
<223> 人工序列描述：合成6xHis標籤

<400> 955
His His His His His His
1 5

<210> 956
<211> 237
<212> PRT
<213> 未知

<220>
<223> 未知描述：指環病毒科序列

<400> 956

```
Arg Arg Lys Leu Arg Val Arg Lys Arg Phe Tyr Lys Arg Lys Leu Lys
1               5                  10                  15

Lys Ile Val Leu Lys Gln Phe Gln Pro Lys Ile Ile Arg Arg Cys Thr
            20                  25                  30

Ile Phe Gly Thr Ile Cys Leu Phe Gln Gly Ser Pro Glu Arg Ala Asn
        35                  40                  45

Asn Asn Tyr Ile Gln Thr Ile Tyr Ser Tyr Val Pro Asp Lys Glu Pro
    50                  55                  60

Gly Gly Gly Gly Trp Thr Leu Ile Thr Glu Ser Leu Ser Ser Leu Trp
65                  70                  75                  80

Glu Asp Trp Glu His Leu Lys Asn Val Trp Thr Gln Ser Asn Ala Gly
                85                  90                  95

Leu Pro Leu Val Arg Tyr Gly Gly Val Thr Leu Tyr Phe Tyr Gln Ser
            100                 105                 110

Ala Tyr Thr Asp Tyr Ile Ala Gln Val Phe Asn Cys Tyr Pro Met Thr
        115                 120                 125

Asp Thr Lys Tyr Thr His Ala Asp Ser Ala Pro Asn Arg Met Leu Leu
    130                 135                 140

Lys Lys His Val Ile Arg Val Pro Ser Arg Glu Thr Arg Lys Lys Arg
145                 150                 155                 160

Lys Pro Tyr Lys Arg Val Arg Val Gly Pro Pro Ser Gln Met Gln Asn
                165                 170                 175

Lys Trp Tyr Phe Gln Arg Asp Ile Cys Glu Ile Pro Leu Ile Met Ile
            180                 185                 190

Ala Ala Thr Ala Val Asp Phe Arg Tyr Pro Phe Cys Ala Ser Asp Cys
```

195 200 205

Ala Ser Asn Asn Leu Thr Leu Thr Cys Leu Asn Pro Leu Leu Phe Gln
210 215 220

Asn Gln Asp Phe Asp His Pro Ser Asp Thr Gln Gly Tyr
225 230 235

<210> 957
<211> 232
<212> PRT
<213> 未知

<220>
<223> 未知描述：指環病毒科序列

<400> 957
Arg Arg Lys Arg Arg Val Arg Pro Thr Tyr Thr Thr Ile Pro Leu Lys
1 5 10 15

Gln Trp Gln Pro Pro Tyr Lys Arg Thr Cys Tyr Ile Lys Gly Gln Asp
20 25 30

Cys Leu Ile Tyr Tyr Ser Asn Leu Arg Leu Gly Met Asn Ser Thr Met
35 40 45

Tyr Glu Lys Ser Ile Val Pro Val His Trp Pro Gly Gly Gly Ser Phe
50 55 60

Ser Val Ser Met Leu Thr Leu Asp Ala Leu Tyr Asp Ile His Lys Leu
65 70 75 80

Cys Arg Asn Trp Trp Thr Ser Thr Asn Gln Asp Leu Pro Leu Val Arg
85 90 95

Tyr Lys Gly Cys Lys Ile Thr Phe Tyr Gln Ser Thr Phe Thr Asp Tyr
100 105 110

Ile Val Arg Ile His Thr Glu Leu Pro Ala Asn Ser Asn Lys Leu Thr
115 120 125

Tyr Pro Asn Thr His Pro Leu Met Met Met Met Ser Lys Tyr Lys His
130 135 140

Ile Ile Pro Ser Arg Gln Thr Arg Arg Lys Lys Lys Pro Tyr Thr Lys
145 150 155 160

Ile Phe Val Lys Pro Pro Pro Gln Phe Glu Asn Lys Trp Tyr Phe Ala
165 170 175

Thr Asp Leu Tyr Lys Ile Pro Leu Leu Gln Ile His Cys Thr Ala Cys
180 185 190

Asn Leu Gln Asn Pro Phe Val Lys Pro Asp Lys Leu Ser Asn Asn Val
195 200 205

Thr Leu Trp Ser Leu Asn Thr Ile Ser Ile Gln Asn Arg Asn Met Ser
210 215 220

Val Asp Gln Gly Gln Ser Trp Pro
225 230

<210> 958
<211> 238
<212> PRT
<213> 未知

<220>
<223> 未知描述：指環病毒科序列

<400> 958
Arg Arg Arg Arg Arg Trp Arg Lys Arg Pro Thr Val Arg Arg Lys Leu
1 5 10 15

Lys Lys Leu Thr Ile Gln Gln Trp Gln Pro Lys Thr Ile Arg Lys Cys
20 25 30

Cys Ile Gln Gly Leu His Cys Leu Phe Leu Val Thr Glu Asp Thr Ile
35 40 45

Ser Arg Asn Tyr Arg Met Tyr Glu His Ser Tyr Thr Gly Glu Tyr Tyr
50 55 60

Pro Gly Gly Gly Gly Phe Ser Ile Thr Arg Tyr Ser Leu Asp Gly Leu
65 70 75 80

Tyr Glu Gln His Gln Leu Asp Arg Asn Trp Trp Thr Asn Ser Asn Thr
85 90 95

Asn Leu Pro Leu Val Arg Tyr Thr Gly Cys Lys Ile Lys Phe Tyr Gln
100 105 110

Ser Trp Ser Val Asp Tyr Ile Cys Asn Tyr Ser Leu Thr Trp Pro Met
115 120 125

Val Ala Thr Gln Leu Leu Tyr Gln Ser Cys Gln Pro Ser Phe Met Met
130 135 140

Met Asn Lys Asn Ser Ile Met Ile Pro Ser Lys Leu Thr Lys Pro Ile
145 150 155 160

Lys Lys Gly Tyr Lys Thr Ile Lys Glu Lys Pro Pro His Glu Met Leu
165 170 175

Asn Arg Trp Tyr Phe Ala Lys Asp Leu Ser Lys Val Gly Leu Leu Met
180 185 190

Leu Thr Ala Ala Ser Ala Ser Phe Asp His Tyr Tyr Gln Ala Thr Asp
195 200 205

Ser Leu Ser Asn Asn Cys Thr Phe Glu Ser Leu Asn Pro Tyr Phe Tyr
210 215 220

Met Arg His Asp Phe Ile Leu Phe Pro Val Thr Gly Tyr Val
225 230 235

<210> 959

<211> 238
<212> PRT
<213> 未知

<220>
<223> 未知描述：指環病毒科序列

<400> 959

```
Arg Arg Arg Arg Arg Trp Val Arg Arg Lys Pro Phe Tyr Lys Arg Lys
1               5                   10                  15

Ile Lys Arg Leu Asn Ile Val Glu Trp Gln Pro Lys Ser Ile Arg Lys
            20                  25                  30

Cys Arg Ile Lys Gly Met Leu Cys Leu Phe Gln Thr Thr Glu Asp Arg
        35                  40                  45

Leu Ser Tyr Asn Phe Asp Met Tyr Glu Glu Ser Ile Ile Pro Glu Lys
    50                  55                  60

Leu Pro Gly Gly Gly Gly Phe Ser Ile Lys Asn Ile Ser Leu Tyr Ala
65                  70                  75                  80

Leu Tyr Gln Glu His Ile His Ala His Asn Ile Phe Thr His Thr Asn
                85                  90                  95

Thr Asp Arg Pro Leu Ala Arg Tyr Thr Gly Cys Ser Leu Lys Phe Tyr
            100                 105                 110

Gln Ser Lys Asp Ile Asp Tyr Val Val Thr Tyr Ser Thr Ser Leu Pro
        115                 120                 125

Leu Arg Ser Ser Met Gly Met Tyr Asn Ser Met Gln Pro Ser Ile His
    130                 135                 140

Leu Met Gln Gln Asn Lys Leu Ile Val Pro Ser Lys Gln Thr Gln Lys
145                 150                 155                 160

Arg Arg Lys Pro Tyr Ile Lys Lys His Ile Ser Pro Pro Thr Gln Met
                165                 170                 175
```

Lys Ser Gln Trp Tyr Phe Gln His Asn Ile Ala Asn Ile Pro Leu Leu
180 185 190

Met Ile Arg Thr Thr Ala Leu Thr Leu Asp Asn Tyr Tyr Ile Gly Ser
195 200 205

Arg Gln Leu Ser Thr Asn Val Thr Ile His Thr Leu Asn Thr Thr Tyr
210 215 220

Ile Gln Asn Arg Asp Trp Gly Asp Arg Asn Lys Thr Tyr Tyr
225 230 235

<210> 960
<211> 240
<212> PRT
<213> 未知

<220>
<223> 未知描述：指環病毒科序列

<400> 960
Arg Arg Pro Val Arg Arg Arg Arg Arg Lys Leu Arg Lys Val Lys Arg
1 5 10 15

Lys Lys Lys Ser Leu Ile Val Arg Gln Trp Gln Pro Asp Ser Ile Arg
20 25 30

Thr Cys Lys Ile Ile Gly Gln Ser Ala Ile Val Val Gly Ala Glu Gly
35 40 45

Lys Gln Met Tyr Cys Tyr Thr Val Asn Lys Leu Ile Asn Val Pro Pro
50 55 60

Lys Thr Pro Tyr Gly Gly Gly Phe Gly Val Asp Gln Tyr Thr Leu Lys
65 70 75 80

Tyr Leu Tyr Glu Glu Tyr Arg Phe Ala Gln Asn Ile Trp Thr Gln Ser
85 90 95

Asn Val Leu Lys Asp Leu Cys Arg Tyr Ile Asn Val Lys Leu Ile Phe
100 105 110

Tyr Arg Asp Asn Lys Thr Asp Phe Val Leu Ser Tyr Asp Arg Asn Pro
115 120 125

Pro Phe Gln Leu Thr Lys Phe Thr Tyr Pro Gly Ala His Pro Gln Gln
130 135 140

Ile Met Leu Gln Lys His His Lys Phe Ile Leu Ser Gln Met Thr Lys
145 150 155 160

Pro Asn Gly Arg Leu Thr Lys Lys Leu Lys Ile Lys Pro Pro Lys Gln
165 170 175

Met Leu Ser Lys Trp Phe Phe Ser Lys Gln Phe Cys Lys Tyr Pro Leu
180 185 190

Leu Ser Leu Lys Ala Ser Ala Leu Asp Leu Arg His Ser Tyr Leu Gly
195 200 205

Cys Cys Asn Glu Asn Pro Gln Val Phe Phe Tyr Tyr Leu Asn His Gly
210 215 220

Tyr Tyr Thr Ile Thr Asn Trp Gly Ala Gln Ser Ser Thr Ala Tyr Arg
225 230 235 240

<210> 961
<211> 240
<212> PRT
<213> 未知

<220>
<223> 未知描述：指環病毒科序列

<400> 961
Arg Phe Thr Lys Thr Arg Arg Arg Arg Lys Arg Lys Lys Val Arg Arg
1 5 10 15

```
Lys Leu Lys Lys Ile Thr Ile Lys Gln Trp Gln Pro Asp Ser Val Lys
            20                  25                  30

Lys Cys Lys Ile Lys Gly Tyr Ser Thr Leu Val Met Gly Ala Gln Gly
        35                  40                  45

Lys Gln Tyr Asn Cys Tyr Thr Asn Gln Ala Ser Asp Tyr Val Gln Pro
    50                  55                  60

Lys Ala Pro Gln Gly Gly Gly Phe Gly Cys Glu Val Phe Asn Leu Lys
65                  70                  75                  80

Trp Leu Tyr Gln Glu Tyr Thr Ala His Arg Asn Ile Trp Thr Lys Thr
                85                  90                  95

Asn Glu Tyr Thr Asp Leu Cys Arg Tyr Thr Gly Ala Gln Ile Ile Leu
            100                 105                 110

Tyr Arg His Pro Asp Val Asp Phe Ile Val Ser Trp Asp Asn Gln Pro
        115                 120                 125

Pro Phe Leu Leu Asn Lys Tyr Thr Tyr Pro Glu Leu Gln Pro Gln Asn
    130                 135                 140

Leu Leu Leu Ala Arg Arg Lys Arg Ile Ile Leu Ser Gln Lys Ser Asn
145                 150                 155                 160

Pro Lys Gly Lys Leu Arg Ile Lys Leu Arg Ile Pro Pro Pro Lys Gln
                165                 170                 175

Met Ile Thr Lys Trp Phe Phe Gln Arg Asp Phe Cys Asp Val Asn Leu
            180                 185                 190

Phe Lys Leu Cys Ala Ser Ala Ala Ser Phe Arg Tyr Pro Gly Ile Ser
        195                 200                 205

His Gly Ala Gln Ser Thr Ile Phe Ser Ala Tyr Ala Leu Asn Thr Asp
    210                 215                 220
```

Phe Tyr Gln Cys Ser Asp Trp Cys Gln Thr Asn Thr Glu Thr Gly Tyr
225 230 235 240

<210> 962
<211> 239
<212> PRT
<213> 未知

<220>
<223> 未知描述：指環病毒科序列

<400> 962
Gly Arg Arg Thr Tyr Thr Arg Arg Ala Val Arg Arg Arg Arg Arg Pro
1 5 10 15

Arg Lys Arg Leu Val Leu Thr Gln Trp Ser Pro Gln Thr Val Arg Asn
20 25 30

Cys Ser Ile Arg Gly Ile Val Pro Met Val Ile Cys Gly His Thr Lys
35 40 45

Ala Gly Arg Asn Tyr Ala Ile His Ser Glu Asp Phe Thr Thr Gln Ile
50 55 60

Gln Pro Phe Gly Gly Ser Phe Ser Thr Thr Thr Trp Ser Leu Lys Val
65 70 75 80

Leu Trp Asp Glu His Gln Lys Phe Gln Asn Arg Trp Ser Tyr Pro Asn
85 90 95

Thr Gln Leu Asp Leu Ala Arg Tyr Arg Gly Val Thr Phe Trp Phe Tyr
100 105 110

Arg Asp Gln Lys Thr Asp Tyr Ile Val Gln Trp Ser Arg Asn Pro Pro
115 120 125

Phe Lys Leu Asn Lys Tyr Ser Ser Ala Met Tyr His Pro Gly Met Met
130 135 140

Met Gln Ala Lys Arg Lys Leu Val Val Pro Ser Phe Gln Thr Arg Pro
145 150 155 160

Lys Gly Lys Lys Arg Tyr Arg Val Thr Ile Lys Pro Pro Asn Met Phe
165 170 175

Ala Asp Lys Trp Tyr Thr Gln Glu Asp Leu Cys Pro Val Pro Leu Val
180 185 190

Gln Ile Val Val Ser Ala Ala Ser Leu Leu His Pro Phe Cys Pro Pro
195 200 205

Gln Thr Asn Asn Pro Cys Ile Thr Phe Gln Val Leu Lys Asp Ile Tyr
210 215 220

Asp Glu Cys Ile Gly Val Asn Glu Thr Met Lys Asp Lys Tyr Lys
225 230 235

<210> 963
<211> 240
<212> PRT
<213> 未知

<220>
<223> 未知描述：指環病毒科序列

<400> 963
Arg Arg Arg Arg Arg Arg Gly Arg Arg Arg Arg Arg Arg Arg Arg Arg
1 5 10 15

His Lys Pro Thr Leu Ile Leu Arg Gln Trp Gln Pro Asp Cys Ile Arg
20 25 30

His Cys Lys Ile Thr Gly Trp Met Pro Leu Ile Ile Cys Gly Lys Gly
35 40 45

Ser Thr Gln Phe Asn Tyr Ile Thr His Ala Asp Asp Ile Thr Pro Arg
50 55 60

Gly Ala Ser Tyr Gly Gly Asn Phe Thr Asn Met Thr Phe Ser Leu Glu
65 70 75 80

Ala Ile Tyr Glu Gln Phe Leu Tyr His Arg Asn Arg Trp Ser Ala Ser
85 90 95

Asn His Asp Leu Glu Leu Cys Arg Tyr Lys Gly Thr Thr Leu Lys Leu
100 105 110

Tyr Arg His Pro Glu Val Asp Tyr Ile Val Thr Tyr Ser Arg Thr Gly
115 120 125

Pro Phe Glu Ile Ser His Met Thr Tyr Leu Ser Thr His Pro Met Leu
130 135 140

Met Leu Leu Asn Lys His His Ile Val Val Pro Ser Leu Lys Thr Lys
145 150 155 160

Pro Arg Gly Arg Lys Ala Ile Lys Val Arg Ile Arg Pro Pro Lys Leu
165 170 175

Met Asn Asn Lys Trp Tyr Phe Thr Arg Asp Phe Cys Asn Ile Gly Leu
180 185 190

Phe Gln Leu Trp Ala Thr Gly Leu Glu Leu Arg Asn Pro Trp Leu Arg
195 200 205

Met Ser Thr Leu Ser Pro Cys Ile Gly Phe Asn Val Leu Lys Asn Ser
210 215 220

Ile Tyr Thr Asn Leu Ser Asn Leu Pro Gln Tyr Lys Asn Glu Arg Leu
225 230 235 240

<210> 964
<211> 233
<212> PRT
<213> 未知

<220>

<223> 未知描述：指環病毒科序列

<400> 964

```
Arg Arg Arg Arg Arg Phe Arg Arg Arg Gly Arg Lys Ala Lys Leu Ile
1               5                  10                  15

Ile Lys Leu Trp Gln Pro Ala Val Ile Lys Arg Cys Arg Ile Lys Gly
            20                  25                  30

Tyr Ile Pro Leu Ile Ile Ser Gly Asn Gly Thr Phe Ala Thr Asn Phe
        35                  40                  45

Thr Ser His Ile Asn Asp Arg Ile Met Lys Gly Pro Phe Gly Gly Gly
    50                  55                  60

His Ser Thr Met Arg Phe Ser Leu Tyr Ile Leu Phe Glu Glu His Leu
65                  70                  75                  80

Arg His Met Asn Phe Trp Thr Arg Ser Asn Asp Asn Leu Glu Leu Thr
                85                  90                  95

Arg Tyr Leu Gly Ala Ser Val Lys Ile Tyr Arg His Pro Asp Gln Asp
            100                 105                 110

Phe Ile Val Ile Tyr Asn Arg Arg Thr Pro Leu Gly Gly Asn Ile Tyr
        115                 120                 125

Thr Ala Pro Ser Leu His Pro Gly Asn Ala Ile Leu Ala Lys His Lys
    130                 135                 140

Ile Leu Val Pro Ser Leu Gln Thr Arg Pro Lys Gly Arg Lys Ala Ile
145                 150                 155                 160

Arg Leu Arg Ile Ala Pro Pro Thr Leu Phe Thr Asp Lys Trp Tyr Phe
                165                 170                 175

Gln Lys Asp Ile Ala Asp Leu Thr Leu Phe Asn Ile Met Ala Val Glu
            180                 185                 190
```

Ala Asp Leu Arg Phe Pro Phe Cys Ser Pro Gln Thr Asp Asn Thr Cys
195 200 205

Ile Ser Phe Gln Val Leu Ser Ser Val Tyr Asn Asn Tyr Leu Ser Ile
210 215 220

Asn Thr Phe Asn Asn Asp Asn Ser Asp
225 230

<210> 965
<211> 233
<212> PRT
<213> 未知

<220>
<223> 未知描述：指環病毒科序列

<400> 965
Arg Arg Trp Arg Arg Lys Gly Lys Arg Ser Arg Lys Lys Lys Ile Ile
1 5 10 15

Ile Arg Gln Trp Gln Pro Asn Tyr Thr Arg Arg Cys Asn Ile Val Gly
20 25 30

Tyr Met Pro Leu Leu Ile Cys Gly Glu Asn Thr Val Ala Thr Asn Tyr
35 40 45

Ala Thr His Ser Asp Asp Ser Tyr Tyr Pro Gly Pro Phe Gly Gly Gly
50 55 60

Met Thr Thr Asp Lys Phe Thr Leu Arg Ile Leu Tyr Asp Glu Tyr Lys
65 70 75 80

Arg Phe Met Asn Tyr Trp Thr Ser Ser Asn Glu Asp Leu Asp Leu Cys
85 90 95

Arg Tyr Leu Gly Cys Thr Leu Tyr Val Phe Arg His Pro Glu Val Asp
100 105 110

Phe Ile Ile Ile Ile Asn Thr Ser Pro Pro Phe Leu Asp Thr Glu Ile
115 120 125

Thr Gly Pro Ser Ile His Pro Gly Met Met Ala Leu Asn Lys Arg Ser
130 135 140

Arg Trp Ile Pro Ser Ile Lys Asn Arg Pro Gly Arg Lys His Tyr Ile
145 150 155 160

Lys Ile Lys Val Gly Ala Pro Arg Met Phe Thr Asp Lys Trp Tyr Pro
165 170 175

Gln Thr Asp Leu Cys Asp Met Thr Leu Leu Thr Ile Phe Ala Ser Ala
180 185 190

Ala Asp Met Gln Tyr Pro Phe Gly Ser Pro Leu Thr Asp Thr Ile Val
195 200 205

Val Ser Phe Gln Val Leu Gln Ser Met Tyr Asn Asp Cys Leu Ser Val
210 215 220

Leu Pro Asp Asn Phe Ala Glu Thr Ser
225 230

<210> 966
<211> 232
<212> PRT
<213> 未知

<220>
<223> 未知描述：指環病毒科序列

<400> 966
Arg Arg Trp Lys Arg Lys Gly Arg Arg Arg Arg Lys Ala Lys Ile Ile
1 5 10 15

Ile Arg Gln Trp Gln Pro Asn Tyr Arg Arg Arg Cys Asn Ile Val Gly
20 25 30

Tyr Leu Pro Ile Leu Ile Cys Gly Gly Asn Thr Val Ser Arg Asn Tyr

35 40 45

Ala Thr His Ser Asp Asp Thr Asn Tyr Pro Gly Pro Phe Gly Gly Gly
50 55 60

Met Thr Thr Asp Lys Phe Ser Leu Arg Ile Leu Tyr Asp Glu Tyr Lys
65 70 75 80

Arg Phe Met Asn Tyr Trp Thr Ala Ser Asn Glu Asp Leu Asp Leu Cys
85 90 95

Arg Tyr Leu Gly Cys Thr Phe Tyr Phe Phe Arg His Pro Glu Val Asp
100 105 110

Phe Ile Ile Lys Ile Asn Thr Met Pro Pro Phe Leu Asp Thr Thr Ile
115 120 125

Thr Ala Pro Ser Ile His Pro Gly Leu Met Ala Leu Asp Lys Arg Ala
130 135 140

Arg Trp Ile Pro Ser Leu Lys Asn Arg Pro Gly Lys Lys His Tyr Ile
145 150 155 160

Lys Ile Arg Val Gly Ala Pro Lys Met Phe Thr Asp Lys Trp Tyr Pro
165 170 175

Gln Thr Asp Leu Cys Asp Met Thr Leu Leu Thr Ile Tyr Ala Thr Ala
180 185 190

Ala Asp Met Gln Tyr Pro Phe Gly Ser Pro Leu Thr Asp Thr Val Val
195 200 205

Val Asn Ser Gln Val Leu Gln Ser Met Tyr Asp Glu Thr Ile Ser Ile
210 215 220

Leu Pro Asp Glu Lys Thr Lys Arg
225 230

<210> 967
<211> 170
<212> PRT
<213> 喙羽病病毒

<400> 967

```
Arg Arg Arg Arg Arg Arg Phe Ser Thr Asn Arg Ile Tyr Thr Leu Arg
1               5                  10                  15

Leu Thr Arg Gln Phe Gln Phe Lys Ile Asn Lys Gln Leu Ile Phe Asn
            20                  25                  30

Ala Asp Tyr Ile Thr Phe Ala Leu Asp Asp Phe Leu Gln Ala Val Pro
        35                  40                  45

Asn Pro His Thr Leu Asn Phe Glu Asp Tyr Arg Ile Lys Leu Ala Lys
    50                  55                  60

Met Glu Met Arg Pro Thr Gly Gly His Tyr Thr Gly Phe Gly His Thr
65                  70                  75                  80

Ala Val Ile Gln Asp Ser Arg Ile Thr Arg Phe Lys Thr Thr Ala Asp
                85                  90                  95

Pro Leu Ala Pro Phe Asp Gly Ala Lys Lys Trp Phe Val Ser Arg Gly
            100                 105                 110

Phe Lys Arg Leu Leu Arg Pro Lys Pro Gln Ile Thr Ile Glu Gly Pro
        115                 120                 125

Asn Ser Ala Gly Thr Lys Val Arg His Tyr Gly Ile Ala Phe Ser Phe
    130                 135                 140

Pro Gln Pro Glu Gln Thr Val Thr Lys Leu Thr Leu Tyr Val Gln Phe
145                 150                 155                 160

Arg Gln Phe Ala Pro Asn Asn Pro Ser Thr
                165                 170
```

<210> 968
<211> 179
<212> PRT
<213> 正肝病毒屬A物種(Orthohepevirus A)

<400> 968

```
Gly Ala Ile Leu Arg Arg Gln Tyr Asn Leu Ser Thr Ser Pro Leu Thr
1               5                  10                  15

Ser Ser Val Ala Ser Gly Thr Asn Leu Val Leu Tyr Ala Ala Pro Leu
            20                  25                  30

Asn Pro Leu Leu Pro Leu Gln Asp Gly Thr Asn Thr His Ile Met Ala
        35                  40                  45

Thr Glu Ala Ser Asn Tyr Ala Gln Tyr Arg Val Val Arg Ala Thr Ile
    50                  55                  60

Arg Tyr Arg Pro Leu Val Pro Asn Ala Val Ser Ile Ser Phe Trp Pro
65                  70                  75                  80

Gln Thr Thr Thr Thr Pro Thr Ser Val Asp Met Asn Ser Ile Thr Ser
                85                  90                  95

Thr Asp Val Arg Ile Leu Val Gln Pro Gly Ile Ala Ser Glu Leu Val
            100                 105                 110

Ile Pro Ser Glu Arg Leu His Tyr Arg Thr Arg Thr Gly Val Ala Glu
        115                 120                 125

Glu Glu Ala Thr Ser Gly Leu Val Met Leu Cys Ile His Gly Ser Pro
    130                 135                 140

Val Asn Ser Ala Leu Gly Leu Leu Asp Phe Ala Leu Glu Leu Glu Phe
145                 150                 155                 160

Arg Asn Leu Thr Pro Gly Asn Thr Asn Thr Arg Val Ser Arg Tyr Thr
                165                 170                 175
```

Ser Thr Ala

# 【發明申請專利範圍】

【請求項1】

一種指環載體(anellovector)，其包含：

a)包含ORF1分子之蛋白質外部；

b)包含RNA之遺傳元件，

其中該遺傳元件包裹於該蛋白質外部內。

【請求項2】

如請求項1之指環載體，其中該遺傳元件由至少10%、20%、25%、30%、40%、50%、60%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100% RNA組成。

【請求項3】

如請求項1或2之指環載體，其中該RNA包含一或多個化學修飾。

【請求項4】

如前述請求項中任一項之指環載體，其中該遺傳元件由RNA組成或基本上由RNA組成。

【請求項5】

如前述請求項中任一項之指環載體，其中該遺傳元件包含DNA區。

【請求項6】

如前述請求項中任一項之指環載體，其中該DNA區之至少一部分與該遺傳元件之該RNA的至少一部分雜交。

【請求項7】

如前述請求項中任一項之指環載體，其中該遺傳元件為環狀。

【請求項8】

一種製造指環載體之方法，該方法包含：

(a)提供包含以下之混合物：

(i)包含RNA之遺傳元件，及

(ii) ORF1分子；及

(b)在適用於將該遺傳元件包裹於包含該ORF1分子之蛋白質外部內的條件下培育該混合物，由此製造指環載體；

視情況其中，該混合物不包含於細胞中。

【請求項9】

一種向細胞遞送遺傳元件之方法，該方法包含使如請求項1至7中任一項之指環載體與細胞，例如真核細胞，例如哺乳動物細胞，例如人類細胞接觸。

# [(發明圖式)]

圖式