

[19] 中华人民共和国国家知识产权局

[51] Int. Cl<sup>7</sup>

C12N 15/86

C12N 15/41 A61K 48/00

A61K 39/00



# [12] 发明专利申请公开说明书

[21] 申请号 02810488.9

[43] 公开日 2005 年 2 月 2 日

[11] 公开号 CN 1575339A

[22] 申请日 2002.5.23 [21] 申请号 02810488.9

[30] 优先权

[32] 2001.5.23 [33] US [31] 60/292,515

[86] 国际申请 PCT/IB2002/002810 2002.5.23

[87] 国际公布 WO2002/095023 英 2002.11.28

[85] 进入国家阶段日期 2003.11.24

[71] 申请人 巴斯德研究院

地址 法国巴黎

[72] 发明人 尼古拉斯·埃斯克瑞欧

西尔维·范-德尔-韦夫

马尔科·维纽齐 西尔维·热尔博

[74] 专利代理机构 永新专利商标代理有限公司

代理人 林晓红

权利要求书 8 页 说明书 44 页 序列表 13 页  
附图 14 页

[54] 发明名称 用于生产异源性蛋白的、衍生自正链 RNA 病毒基因组的复制子

[57] 摘要

本发明涉及由心肌炎病毒及口疮病毒基因组衍生的复制子或自我复制 RNA 分子，其能用于在动物细胞内表达异源性蛋白。当以裸 RNA 形式注射入动物宿主体内时，这些复制子能使所编码异源性蛋白得以翻译。如果所编码异源性蛋白为外源性抗原，则这些复制子能诱导针对所编码异源性蛋白的免疫应答。本发明使用心肌炎病毒及口疮病毒基因组构建这些复制子。本发明证实了这些复制子在以裸 RNA 形式注射时，不需要任何载体或佐剂即能在动物宿主体内诱导针对复制子编码异源性蛋白的免疫应答。

ISSN 1008-4274

1. 一种 RNA 病毒的病毒基因组的自我复制重组正链 RNA 分子，其中所述 RNA 分子包含：

- a) 编码该 RNA 病毒非结构蛋白的 RNA 序列；
- b) 病毒复制所必需的病毒非编码 RNA 序列；以及
- c) 编码一异源性蛋白或异源性蛋白片段的 RNA 序列。

2. 一种 RNA 病毒的病毒基因组的自我复制重组正链 RNA 分子，其中所述 RNA 分子包含：

- (a) 编码该 RNA 病毒非结构蛋白的 RNA 序列；
- (b) 病毒复制所必需的病毒非编码 RNA 序列，其中 a) 中的 RNA 序列和/或 b) 中的病毒非编码 RNA 序列为突变或截断的形式，以及
- (c) 编码一异源性蛋白或异源性蛋白片段的 RNA 序列。

3. 权利要求 1 或 2 的自我复制重组正链 RNA 分子，其中的 RNA 病毒属于心肌炎病毒属或口疮病毒属。

4. 权利要求 3 的自我复制重组正链 RNA 分子，其中的 RNA 病毒为门戈病毒。

5. 权利要求 4 的自我复制重组正链 RNA 分子，其进一步包含门戈病毒 VP2 基因的顺式作用复制元件（CRE）。

6. 权利要求 4 的自我复制重组正链 RNA 分子，其进一步包含泰累尔病毒 VP2 基因的顺式作用复制元件（CRE）。

7. 权利要求 1 至 6 中任一项的自我复制重组正链 RNA 分子，其中所述的异源性蛋白选自生物活性蛋白、报告蛋白、细胞毒性蛋白、病原体蛋白或肿瘤蛋白。

8. 权利要求 7 的自我复制重组正链 RNA 分子，其中所述的报告蛋白为绿色荧光蛋白。

9. 权利要求 7 的自我复制重组正链 RNA 分子，其中所述的病原体蛋白为流感病毒核蛋白或流感病毒红细胞凝集素。

10. 权利要求 1 至 6 中任一项的自我复制重组正链 RNA 分子，其中所述的异源性蛋白片段为一抗原或所述异源性蛋白的表位。

11. 一种疫苗，其包含至少一种权利要求 1 至 7 及 9 至 10 中任一项的自我复制重组正链 RNA 分子及一种药用可接受的载体。

12. 权利要求 11 的疫苗，其中所述的自我复制重组正链 RNA 分子为裸 RNA。

13. 权利要求 11 的疫苗，其中所述的自我复制重组正链 RNA 分子为衣壳化的。

14. 权利要求 11 至 13 中任一项的疫苗，其中所述的药用可接受的载体选自水、石油、动物油、植物油、花生油、豆油、矿物油、芝麻油、生理盐水溶液、葡萄糖水溶液、甘油溶液、聚合阳离子颗粒、蛋白质颗粒、鱼精蛋白颗粒、脂质体和金颗粒。

15. 一种在宿主体内诱导保护性免疫应答的方法，其包含：

(a) 在药用可接受的载体中制备至少一种权利要求 1 至 7 及 9 至 10 中任一项的自我复制重组正链 RNA 分子；以及

(b) 用步骤(a)的制剂免疫宿主。

16. 权利要求 15 的在宿主体内诱导免疫应答的方法，其中(a)中的权利要求 1 至 7 及 9 至 10 中任一项的自我复制重组正链 RNA 分子制备为裸 RNA 形式。

17. 权利要求 15 的在宿主体内诱导免疫应答的方法，其中(a)中的权利要求 1 至 7 及 9 至 10 中任一项的自我复制重组正链 RNA 分子为衣壳化的 RNA。

18. 权利要求 15 至 17 中任一项的方法，其中所述的药用可接受的载体选自水、石油、动物油、植物油、花生油、豆油、矿物油、芝麻油，生理盐水溶液、葡萄糖水溶液、甘油溶液、聚合阳离子颗粒、蛋白质颗粒、鱼精蛋白颗粒、脂质体和金颗粒。

19. 权利要求 15 至 18 中任一项的方法，其中所述的宿主为人、猪、狗、猫、牛、鸡、小鼠或马。

20. 一种 DNA 分子，其编码一种 RNA 病毒的病毒基因组的自我复制重组正链 RNA 分子，其中所述的 RNA 分子包含：

- (a) 编码该 RNA 病毒非结构蛋白的 RNA 序列；
- (b) 病毒复制所必需的病毒非编码 RNA 序列；以及
- (c) 编码一异源性蛋白或异源性蛋白片段的 RNA 序列。

21. 一种 DNA 分子，其编码一种 RNA 病毒的病毒基因组的自我复制重组正链 RNA 分子，其中所述的 RNA 分子包含：

- (a) 编码该 RNA 病毒非结构蛋白的 RNA 序列；

(b) 病毒复制所必需的病毒非编码 RNA 序列；其中 a)中的 RNA 序列和/或 b)中的病毒非编码 RNA 序列为突变或截断的形式；以及

(c) 编码一异源性蛋白或异源性蛋白片段的 RNA 序列。

22. 权利要求 20 或 21 的 DNA 分子，其中所述的 RNA 病毒为心肌炎病毒属或口疮病毒属。

23. 权利要求 22 的 DNA 分子，其中所述的 RNA 病毒为门戈病毒。

24. 权利要求 23 的 DNA 分子，其进一步编码包含门戈病毒 VP2 基因的顺式作用复制元件 (CRE) 的 RNA。

25. 权利要求 23 的 DNA 分子，其进一步编码包含泰累尔病毒 VP2 基因的顺式作用复制元件 (CRE) 的 RNA。

26. 权利要求 20 至 25 中任一项的 DNA 分子，其中所述的异源性蛋白选自生物活性蛋白、报告蛋白、细胞毒性蛋白、病原体蛋白或肿瘤蛋白。

27. 权利要求 26 的 DNA 分子，其中所述的报告蛋白为绿色荧光蛋白。

28. 权利要求 26 的 DNA 分子，其中所述的病原体蛋白为流感病毒核蛋白或流感病毒红细胞凝集素。

29. 权利要求 26 的 DNA 分子，其中所述的异源性蛋白片段为一抗原或所述异源性蛋白的表位。

30. 权利要求 26 的 DNA 分子，其进一步包含一合适的克隆载体。

31. 一种 DNA 分子, 其包含 SEQ ID NO: 26 的序列 (于 2001 年 5 月 21 日保藏于 CNCM, 巴斯德研究院, 28 rue du Docteur Roux, 75724 Paris Cedex 15, 法国, 保藏号为 I-2668) 或其片段, 以及可表达形式的、编码异源性蛋白或异源性蛋白片段的 DNA 序列。

32. 一种 DNA 分子, 其包含突变或截断的 SEQ ID NO: 26 序列 (于 2001 年 5 月 21 日保藏于 CNCM, 巴斯德研究院, 28 rue du Docteur Roux, 75724 Paris Cedex 15, 法国, 保藏号为 I-2668) 或其片段, 以及可表达形式的、编码异源性蛋白或异源性蛋白片段的 DNA 序列。

33. 权利要求 31 或 32 的 DNA 分子, 其中所述的异源性蛋白选自生物活性蛋白、报告蛋白、细胞毒性蛋白、病原体蛋白或肿瘤蛋白。

34. 权利要求 33 的 DNA 分子, 其中所述的报告蛋白为绿色荧光蛋白。

35. 权利要求 33 的 DNA 分子, 其中所述的病原体蛋白是流感病毒核蛋白或流感病毒红细胞凝集素。

36. 权利要求 31 或 32 的 DNA 分子, 其中所述的异源性蛋白片段为一抗原或所述异源性蛋白的表位。

37. 一种 DNA 分子, 其包含 SEQ ID NO: 27 序列 (于 2001 年 5 月 21 日保藏于 CNCM, 巴斯德研究院, 28 rue du Docteur Roux, 75724 Paris Cedex 15, 法国, 保藏号为 I-2669) 或其片段, 以及编码异源性蛋白或异源性蛋白片段的 DNA 序列。

38. 一种 DNA 分子, 其包含突变或截断的 SEQ ID NO: 27 序列 (于 2001 年 5 月 21 日保藏于 CNCM, 巴斯德研究院, 28 rue du Docteur

Roux, 75724 Paris Cedex 15, 法国, 保藏号为 I-2669) 或其片段, 以及可表达形式的、编码异源性蛋白或异源性蛋白片段的 DNA 序列。

39. 权利要求 37 或 38 的 DNA 分子, 其中所述的异源性蛋白选自生物活性蛋白、报告蛋白、细胞毒性蛋白、病原体蛋白或肿瘤蛋白。

40. 权利要求 39 的 DNA 分子, 其中所述的病原体蛋白为流感病毒核蛋白或流感病毒红细胞凝集素。

41. 权利要求 37 或 38 的 DNA 分子, 其中所述的异源性蛋白片段为一抗原或所述异源性蛋白的表位。

42. 一种在宿主体内诱导保护性免疫应答的方法, 其包含:

(a) 在药用可接受的载体中制备至少一种权利要求 20 至 41 中任一项的 DNA 分子; 以及

(b) 用步骤(a)的制剂免疫宿主。

43. 权利要求 42 的在宿主体内诱导保护性免疫应答的方法, 其中所述的 DNA 分子为裸 DNA。

44. 权利要求 42 的在宿主体内诱导保护性免疫应答的方法, 其中所述的 DNA 分子为衣壳化的 DNA。

45. 一种治疗组合物, 其包含在一可接受的介质内的、至少一种权利要求 20 至 41 中任一项的 DNA 分子或一种权利要求 1 至 7 及 9 至 10 中任一项的自我复制重组正链 RNA 分子。

46. 一种治疗试剂盒, 其包含在一可接受的介质内的、至少一种权利要求 20 至 41 中任一项的 DNA 分子或一种权利要求 1 至 7 及 9 至 10 中任一项的自我复制重组正链 RNA 分子。

47. 一种在宿主体内调节免疫应答的方法，其包含：

(a) 在一药用可接受的载体内制备至少一种选自权利要求 20 至 41 中任一项的 DNA 分子和一种权利要求 1 至 7 及 9 至 10 中任一项的自我复制重组正链 RNA 分子的分子；以及

(b) 用步骤(a)的制剂免疫宿主。

48. 权利要求 42 的方法，其中所述的药用可接受的载体选自水、石油、动物油、植物油、花生油、豆油、矿物油、芝麻油、生理盐水溶液、葡萄糖水溶液、甘油溶液、聚合阳离子颗粒、蛋白质颗粒、鱼精蛋白颗粒、脂质体和金颗粒、。

49. 权利要求 42 的方法，其中所述的宿主为人、猪、狗、猫、牛、鸡、小鼠或马。

50. 用于通过生产一衣壳化的 RNA 病毒的病毒基因组的自我复制重组正链 RNA 分子而提高 RNA 病毒的病毒基因组的自我复制重组正链 RNA 分子的免疫原性的方法，其包含：

(a) 将权利要求 1 至 7 及 9 至 10 中任一项的自我复制重组正链 RNA 分子或权利要求 20 至 41 中任一项的 DNA 分子转染入表达衣壳蛋白 P1 前体的细胞内；

(b) 由该转染的细胞制备衣壳化的自我复制重组正链 RNA 分子；以及

(c) 用步骤(b)的制剂免疫宿主。

51. 用于提高 RNA 病毒的病毒基因组的自我复制重组正链 RNA 分子的免疫原性的方法，其包含：

(a) 浓缩权利要求 1 至 7 及 9 至 10 中任一项的自我复制重组正链 RNA 分子；以及

(b) 用步骤(a)的浓缩的RNA分子免疫宿主。

52. 权利要求31的DNA分子，其包含SEQ ID NO: 28的序列（于2002年5月16日保藏于CNCM, 巴斯德研究院, 28 rue du Docteur Roux, 75724 Paris Cedex 15, 法国，保藏号为I-2879）。

53. 权利要求 36 或 41 的 DNA 分子，其中所述异源性蛋白的表位为淋巴细胞性脉络丛脑膜炎病毒核蛋白的 NP118-126 表位。

用于生产异源性蛋白的、衍生自正链 RNA 病毒基因组的复制子

本发明涉及由心肌炎病毒 (cardiovirus) 及口疮病毒 (aphtovirus) 基因组衍生的复制子或自我复制 RNA 分子, 其可用于在动物细胞内表达异源性蛋白。例如当将其以裸 RNA 的形式注入宿主动物时, 这些复制子能允许所编码异源性蛋白的翻译。如果编码异源性蛋白为外源性抗原的话, 这些复制子会诱导出抗编码异源性蛋白的免疫反应。本发明使用心肌炎病毒及口疮病毒基因组来构建这些复制子。本发明证明这些复制子以裸 RNA 的形式注入时, 能在没有任何形式载体或佐剂参与的情况下在受体动物体内诱导出抗复制子编码的异源性蛋白的免疫反应。

遗传免疫是疫苗发展过程中的一个强有力的替代工具。其基本原理为接种含有编码外来蛋白的基因序列的 DNA 表达载体。例如, 接种编码流感病毒核蛋白 (NP) 的裸 DNA 载体已被显示能诱导抗体及细胞反应的发生, 因而能保护动物宿主免受流感病毒 A 变体的同源及交叉毒株 (2,27,28) 的感染。DNA 免疫的显性包括生产、纯化和使用的便利性, 及其产生的持久的免疫效果。

DNA 免疫相关的长期免疫可能与注射 DNA 的长期存在及表达有关。事实上, 注射的 DNA 分子在小鼠模型体内 (31) 能持续存在 1 年以上。但是从临床观点来看, 某些问题正是由于该特点而依然存在, 例如 DNA 序列与宿主基因组整合的潜在风险性。尽管在动物体内 (19) 进行的初步研究尚未证实有基因组整合事件的发生, 但是这种整合能引

起插入诱变、原癌基因的活化、或染色体的不稳定性，并从而导致疾病如癌症的发生（35）。

为了避免这个潜在问题的发生，本发明人制造了由正链 RNA 病毒基因组衍生的裸露自我复制 RNA 分子或复制子。将 RNA 作为 DNA 遗传免疫的替代物的提议由来已久，但是这种方法面临着由 RNA 短暂的细胞内半衰期及其被普遍存在的 RNA 酶降解而引起新问题。最初用 mRNA 诱导免疫反应的尝试（5）是以肌肉注射的形式进行的，其施用形式为金颗粒包被基因枪给药（25）或脂质体胶囊注射以在施用过程中保护 RNA（17）。为了进一步改善这些分子的输送及其编码的异源性蛋白的表达，已经开发了可将异源性序列运载进入细胞的、由衣壳包裹的正链 RNA 病毒基因组衍生的自我复制 RNA 或复制子。在这些复制子中，基因组结构基因被异源性序列所取代，而其非结构基因则依然保留以进行一轮复制。该分子设计允许外源蛋白的表达。

甲病毒（alphavirus）、塞姆利基森林病毒（Semliki forest virus, SFV）、辛德毕斯病毒（Sindbis virus）及委内瑞拉马脑炎病毒（Venezuelan equine encephalitis virus）基因组都已用该方法操作使其可表达外源蛋白（11,24）。RNA 复制子经蛋白包装能使其稳定，使所获病毒样颗粒能进行注射并诱导一系列抗异源性蛋白免疫反应的发生。与此相类似的是，脊髓灰质炎病毒的正链 RNA 去掉了其衣壳编码序列以允许外源蛋白的表达（3, 21），在将其包装入病毒样颗粒并给予脊髓灰质炎病毒受体转基因小鼠注射后（18,23），可诱导免疫反应的发生。

与有关包装 RNA 分子的研究相反的是，本发明的本发明人研究了裸 RNA 复制子诱导免疫反应的能力，并提出对这些载体进行包装是不必

要的，这是因为这些复制子的复制本性减少了输入大量 RNA 的需要。在编码红细胞凝集素（HA）及流感病毒 NP 分子的重组 SFV 的案例中，注射裸 RNA 已被发现能诱导特异性抗体的产生（6,34）。近来，一些出版物报告了衍生自 SFV 的重组复制子能诱导抗流感病毒 A、抗呼吸道合胞病毒及抗 Looping Ill 病毒的保护性抗体的产生(10)，以及对作为模型抗原的 lacZ 产生细胞毒 T 淋巴细胞（CTL）（33）。

最近本发明人报道（30），将 RNA 疫苗以裸剂形式注射后，重组编码内在流感病毒 A NP 蛋白（rSFV-NP）的 SFV 复制子能诱导出抗流感病毒 A 的体液和细胞免疫反应，这些反应与质粒 DNA 所诱导的相差无几。此外，本发明人还进一步证明，裸 rSFV-NP 复制子在注射后，能诱导出流感病毒 NP 免疫显性表位特异性的 CTL 反应，并能减少感染适于小鼠的流感病毒的小鼠肺内的病毒负荷量，其程度与已被广泛描述的 DNA 免疫技术所能诱导的程度相同。

本发明人还报道，编码内在流感病毒 A NP 蛋白（r $\Delta$ P1-E-NP）的脊髓灰质炎病毒复制子在以裸 RNA 形式注射后，其在小鼠体内所诱导的体液免疫反应与塞姆利基 rSFV-NP 所诱导的相比，要弱得多。此外，注射 r $\Delta$ P1-E-NP 复制子 RNA 的小鼠体内（30）检测不出抗流感病毒 NP 的 CTL 反应。因此，本发明人决定探索使用细小病毒家族其他成员病毒基因组的可能性，以构建在以裸 RNA 形式注射后，能在动物细胞内及动物受体内表达异源性蛋白的新复制子。口疮病毒和心肌炎病毒属的成员有着相同的基因结构，因而能用于该实验。本发明人将门戈病毒（Mengo virus）作为心肌炎病毒原型的工作实施例。

为了构建基于门戈病毒基因组的复制子，本发明人测出了部分基因组序列可在不影响分子复制的前提下加以去除。接下来，在门戈病毒基因组中进行了一系列包含全部或部分 L-P1-2A 前体蛋白编码区框架内缺失的遗传工程化操作。并对相应亚基因组 RNA 分子的复制特征加以分析。本发明人证明，除了编码 VP2 的一小段核酸序列外，从门戈病毒基因组中去除所有 L-P1-2A 前体蛋白的编码区不会影响该基因组的复制能力。事实上，本发明人证明，包含门戈病毒基因组第 1137 位至 1267 位核苷酸的区域（在此指 vMC24 减毒株的编号方式）含有一个顺式作用复制元件（CRE），该元件对于亚基因组门戈病毒 RNA 分子在转染细胞中进行复制是绝对必需的（15）。这种情形与在脊髓灰质炎病毒和口疮病毒基因组中观察到的是截然不同的，在后两者中，去除整个衣壳蛋白前体（P1）不会影响相应亚基因组 RNA 分子的复制（1,12）。

在构建了门戈病毒衍生的复制子之后，本发明人证明了亚基因组门戈病毒复制子能够表达异源性序列。复制子的免疫原性能通过不同的方法加以改进。例如，本发明人证明门戈病毒复制子可以反式衣壳化的形式转染表达衣壳蛋白 P1 前体蛋白的细胞。同样，Hoerr 等人也描述了用阳离子多肽鱼精蛋白浓缩复制子 RNA 的方法（37）。

本发明描述了由心肌炎病毒属病毒基因组构建的复制子的构建方法及其用途。由于口疮病毒同样是细小病毒家族的成员，且和心肌炎病毒有着同样的遗传学结构，因此类似的复制子还可由口疮病毒属病毒的基因组得以构建。

在本文中使用的“复制子”一词包括但不限于自我复制的重组正链 RNA 分子。

在本文中使用的“正链”一词包括但不限于直接编码蛋白的 RNA 链。

在本文中使用的“表达”一词包括但不限于引起或编码一种蛋白或部分蛋白的生产。

本发明提供了一种来自 RNA 病毒的病毒基因组的自我复制的重组正链 RNA 分子（复制子），其中该 RNA 分子包含：

- (a) 编码该 RNA 病毒非结构蛋白的 RNA 序列；
- (b) 病毒复制所必需的病毒非编码 RNA 序列；以及
- (c) 编码一种异源性蛋白或异源性蛋白片段的 RNA 序列；

根据所述复制子的一个有利的实施方案，a)中编码非结构蛋白的 RNA 序列，和/或 b)中病毒复制所必需的病毒非编码 RNA 序列是以突变或截断的形式存在的。

根据所述复制子的另一个有利的实施方案，该 RNA 病毒是心肌炎病毒属或口疮病毒属的，优选地为门戈病毒；最优选地，所述复制子还包含门戈病毒或泰累尔病毒（Theiler's virus）的 VP2 基因中的顺式作用复制元件（CRE）。

根据所述复制子的另一个有利的实施方案，c)中所定义的异源性蛋白选自生物活性蛋白、报告蛋白、细胞毒性蛋白、病原体蛋白、或是肿瘤蛋白，优选地，报告蛋白是绿色荧光蛋白，而病原体蛋白则为流感病毒核蛋白或流感病毒红细胞凝集素（influenza hemagglutinin）。

根据所述复制子的另一个有利的实施方案，c)中所定义的异源性蛋白片段是一种抗原或所述异源性蛋白的表位。

复制子的构建可通过去除所有或部分衣壳编码序列并保留所有复制必需的编码及非编码序列而完成。保留基因组复制序列使得病毒及异源性基因产物能得以在合适的细胞中进行表达。例如，在门戈病毒 VP2 基因中发现的 CRE 对于复制就是非常重要的。

复制子可用多种方法进行制备。在一个实施方案中，用一种 DNA 依赖的 RNA 聚合酶，如噬菌体 T7、T3 或 SP6 聚合酶对合适的 DNA 序列进行体外转录。在另一个实施方案中，复制子可通过用包含合适 DNA 序列的质粒转染动物细胞，再由转染细胞中分离复制子 RNA 而制备。例如，编码一个复制子的互补 DNA (cDNA) 可置于转录控制之下，位于聚合酶 I 启动子的下游和肝炎病毒 $\delta$ 核酶的上游。在本文中使用的“转染”包括但不限于通过电穿孔、DEAE-葡聚糖处理、磷酸钙沉淀、脂质体（如 Lipofectin）、蛋白质包装（如在假病毒颗粒中）、鱼精蛋白浓缩或任何其他导入方法将 DNA 或 RNA 导入细胞。

本发明还提供了下列 DNA 分子，这些 DNA 分子在根据本发明进行的自我复制重组正链 RNA 分子的生产非常有用：

-一种编码本发明的 RNA 病毒的病毒基因组的自我复制重组正链 RNA 分子的 DNA 分子。在一个优选实施方案中，所述 DNA 分子还含有一个适当的克隆载体；

-一种 DNA 分子，其包含选自 SEQ ID NO: 26 和 SEQ ID NO: 27（于 2001 年 5 月 21 日保藏于 CNCM, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France 的质粒，保藏号分别为 I-2668 和 2669）或其片段的序列，以及可表达形式的、编码一种异源性蛋白或异源性蛋白片段的 DNA 序列；所述 DNA 分子的优选包含 SEQ ID NO: 28（于 2002

年5月16日保藏于 CNCM, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France 的质粒, 保藏号为 I-2879), 以及

-一种 DNA 分子, 其包含选自突变或截断形式的 SEQ ID NO: 26 和 SEQ ID NO: 27 (于 2001 年 5 月 21 日保藏于 CNCM, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France 的质粒, 保藏号分别为 I-2668 和 2669) 或其片段的序列, 以及可表达形式的、编码一种异源性蛋白或异源性蛋白片段的 DNA 序列。

根据所述 DNA 分子的一个优选实施方案, 该异源性蛋白选自生物活性蛋白、报告蛋白、细胞毒性蛋白、病原体蛋白或肿瘤蛋白。优选地, 报告蛋白为绿色荧光蛋白, 病原体蛋白为流感病毒核蛋白, 流感病毒红细胞凝集素, 或淋巴细胞性脉络丛脑膜炎病毒核蛋白, 而异源性蛋白片段为一种抗原或所述异源性蛋白的表位, 优选为淋巴细胞性脉络丛脑膜炎病毒核蛋白的 NP118-126 表位。

本发明的复制子有多种潜在用途。在第一个实施方案中, 可通过在复制子中插入编码异源性多肽的序列, 并将复制子导入动物细胞或动物宿主体内, 从而在动物细胞内或在动物宿主体内表达异源性蛋白。在一个实施方案中, 所述动物宿主是狗、猫、猪、母牛、鸡、小鼠或马。在优选实施方案中, 所述动物宿主为人。复制子可通过多种方法导入宿主体内, 包括肌肉注射、金颗粒包被基因枪注射、蛋白包装注射(如包装在假病毒颗粒中)、鱼精蛋白浓缩注射、或脂质体胶囊注射。例如, 一种由门戈病毒衍生的复制子能在注射部位或其附近瞬时一种治疗蛋白, 或者在直接注射后在实体肿瘤中表达一种毒性蛋白或促凋亡蛋白, 从而提供一种抗肿瘤基因治疗的形式。此外, 重组复制子可用于体外或体内

表达便于检测的报告蛋白。这些复制子可用于监测 RNA 复制和 RNA 输送，从而使动物细胞转染或 RNA 输送入动物宿主体内的过程得以优化。最后，复制子还可用于表达任何令人感兴趣的蛋白，以进一步研究有关蛋白质特性、蛋白质生产或蛋白质定位。

在另一个实施方案中，复制子可用于在动物宿主体内诱导抗编码异源性蛋白的免疫反应。因此，本发明的复制子与一种药用可接受的载体能一起组成一种疫苗。药用可接受的载体包括但不限于无菌液体如水、油，包括石油（petroleum oil）、动物油、植物油、花生油、豆油、矿物油、芝麻油，以及生理盐水溶液、葡萄糖水溶液、甘油溶液、聚合阳离子颗粒、蛋白质颗粒、鱼精蛋白颗粒、脂质体、金颗粒或任何其他能浓缩 RNA 的蛋白或分子。例如，复制子可以“裸露”RNA 或衣壳化 RNA 的形式进行注射。

在一个实施例中，复制子能表达任何病原体的表位，包括细菌、真菌、病毒或寄生虫。复制子还能表达肿瘤抗原或肿瘤抗原和病原体抗原的组合。这种复制子能诱导抗病原体或肿瘤的免疫反应，从而组成抗相应疾病的疫苗。在这点上，门戈病毒衍生的复制子能诱导强烈细胞免疫的能力是一个有利的特性。

在第二个实施例中，复制子还能作为免疫治疗剂用于治疗已经患病的个体。明确地说，复制子能增强已经存在的免疫反应或诱导抗已经存在于个体体内的病原体或肿瘤的新的免疫反应，从而组成了抗相应疾病的治疗方法。例如，乙型肝炎可通过给予表达乙型肝炎病毒表面抗原的复制子，用该方法进行治疗。

在第三个实施例中，可构建复制子用于表达包含一串由相同抗原或不同抗原衍生的 T 细胞表位的合成多肽。这些表位能特异性刺激 CD4<sup>+</sup>T 细胞（辅助 T 细胞）或 CD8<sup>+</sup>T 细胞（CTL）。这些复制子能（1）在考虑 HLA 可变性时诱导多特异性免疫反应，以及（2）限制该病原体或肿瘤细胞通过抗原逃逸对免疫反应的逃避。

在第四个实施例中，能用复制子表达任何生物活性蛋白。在一个实施方案中，生物活性蛋白是一种免疫调节蛋白，如细胞因子或趋化因子，它们能调节宿主的免疫反应。如果在复制子表达外来抗原的相同时间和相同部位进行注射的话，则细胞因子复制子能调节抗该外来抗原的免疫反应。这些复制子还能单独用于调节抗任何病原体抗原或癌抗原的免疫反应。如果合理使用的话，这些复制子还能用于调节自身免疫的病理过程。

因此，本发明提供了一种由至少一种根据本发明制备的自我复制重组正链 RNA 分子和一种药用可接受的载体组成的疫苗。

在所述疫苗的一个有利的实施方案中，其中的自我复制重组正链 RNA 分子为裸 RNA。

在所述疫苗的另一有利的实施方案中，其中的自我复制重组正链 RNA 分子为衣壳化 RNA。

本发明还提供了一种在宿主体内诱导保护性免疫反应的方法，其包含：

(a)在一种药用可接受的载体中制备至少一种选自本发明的自我复制重组正链 RNA 分子和 DNA 分子的分子；以及

(b)用步骤(a)制备的产物免疫宿主。

在所述方法的一个有利的实施方案中，步骤(a)中的自我复制重组正链 RNA 分子和 DNA 分子是裸露的。I

在另一个所述方法的有利的实施方案中，步骤(a)中的自我复制重组正链 RNA 分子是衣壳化的。

本发明还提供了一种治疗组合物，其包含在可接受的介质中的至少一种选自本发明的自我复制重组正链 RNA 分子和 DNA 分子的分子。

本发明还提供了一种治疗试剂盒，其包含在可接受的介质中的至少一种选自本发明的自我复制重组正链 RNA 分子和 DNA 分子的分子。

本发明还提供了一种在宿主体内调节免疫反应的方法，其包含：

(a)在一种药用可接受的载体中制备至少一种选自本发明的自我复制重组正链 RNA 分子和 DNA 分子的分子；以及

(b)用步骤 (a) 制备的产物免疫宿主。

在所述方法的另一个有利的实施方案中，药用可接受的载体选自水、石油、动物油、植物油、花生油、豆油、矿物油、芝麻油、生理盐水溶液、葡萄糖水溶液、甘油溶液、聚合阳离子颗粒、蛋白质颗粒、鱼精蛋白颗粒、脂质体、和金颗粒。

在所述方法的另一个有利的实施方案中，宿主选自人、猪、狗、猫、母牛、鸡、小鼠或马。

本发明还提供了一种通过制备衣壳化 RNA 病毒的病毒基因组 RNA 的自我复制重组正链 RNA 分子以提高其免疫原性的方法，其包含：

(a)用本发明的 DNA 或自我复制重组正链 RNA 分子转染表达衣壳蛋白 P1 前体蛋白的细胞；

(b)由该经转染的细胞制备衣壳化自我复制重组正链 RNA 分子；以及

(c)用步骤 (b) 制备的产物免疫宿主。

本发明还提供了一种提高 RNA 病毒的病毒基因组的自我复制重组正链 RNA 分子的免疫原性的方法，其包含：

(a)浓缩本发明的自我复制重组正链 RNA 分子；以及

(b)用步骤 (a) 制备的产物免疫宿主。

在下文中还将通过图解及工作实施例的形式对本发明进行进一步的证明，其中的复制子由门戈病毒基因组经遗传工程化而获得。然而必需了解的是，这些实施例仅仅用于对本发明进行说明，而并不以任何形式对本发明构成限制。

## 附图说明

图 1 为编码门戈病毒病毒基因组衍生的亚基因组重组复制子的质粒示意图。绿色荧光蛋白 (GFP)、HA 和 NP 基因分别用阴影框表示。CRE 用点彩框表示。HA 蛋白信号肽 (SP) 及 HA 跨膜区 (TM) 用黑色条带表示。

图 2 为证明重组门戈病毒聚合蛋白在兔网织红细胞裂解液内进行体外翻译及处理的 SDS-PAGE 分析。分子量标记物的位置在图右显示。门

戈病毒蛋白前体及其部分主要裂解产物在图左显示。重组复制子编码的 GFP-NP、GFP 多肽及流感病毒 NP 由实心箭头表示。

图 3 为证明亚基因组门戈病毒衍生的复制子复制的狭线杂交 (slot blot)。在所示转染后时间点上, 细胞质 RNA 被提取以供分析。

图 4 为用荧光细胞计数仪阅读重组复制子 rM $\Delta$ BB、rM $\Delta$ BB-GFP 或 rMAXBB-GFP 转染的 HeLa 细胞内的 GFP 的表达。

图 5 为重组复制子 rM $\Delta$ BB-NP 转染的用 [<sup>35</sup>S]甲硫氨酸标记的 HeLa 细胞内表达的流感病毒 NP 蛋白免疫沉淀的 SDS-PAGE 分析。上样情况如下: 假转染 HeLa 细胞 (第 1 泳道); 复制子 rM $\Delta$ BB 转染的 HeLa 细胞 (第 2 泳道); rM $\Delta$ BB-NP 转染的 HeLa 细胞 (第 3 泳道); rM $\Delta$ BB-GFP-NP 转染的 HeLa 细胞 (第 4 泳道), 其中第 1~4 泳道均为转染后 10 小时收获; 假转染 HeLa 细胞 (第 5 泳道); A/PR/8/34 病毒转染的 HeLa 细胞 (第 6 泳道), 其中第 5、第 6 泳道均为转染后 20 小时收获。分子量及病毒 HA 蛋白、病毒 NP 蛋白、病毒 M1 蛋白的位置在图右表示。

图 6 为证明用 rM $\Delta$ BB-NP 免疫的 C57BL/6 小鼠体内诱导的 NP 特异性 CTL 活性的 CTL 实验。每组 4 只 C57BL/6 小鼠以 3 周的间隔分别用下列疫苗方案进行免疫: 1、注射 50  $\mu$ g 的 pCI(O) 或 pCI-NP(●) DNA; 2、注射 25  $\mu$ g 的 rM $\Delta$ BB(□)或 rM $\Delta$ BB-NP(■) RNA。在最末一次注射 3 周后收获小鼠脾细胞, 在体外刺激并进行抗加载 NP366 肽 (a) 或未加载的 NP366 肽 (b) 的同基因 EL4 靶细胞铬释放实验。图中表示了在不同效应剂:靶细胞比例下的特异性溶解的比例。所示数据为两次实验中的一组数据。在最末一次注射 3 周后, 用 IFN $\gamma$  ELISPOT 实验

在有免疫显性 NP366 肽存在 (c) 的条件下对流感病毒特异性 CD8+ T 细胞的频率进行测试, 具体描述详见材料与amp;方法。数据用每  $10^5$  个脾细胞中的 SFC 数目表示。

图 7 为证明用 rM $\Delta$ BB-NP 免疫的 C57BL/6 小鼠体内诱导的 NP 特异性抗体的 ELISA 分析, 其疫苗注射方案同图 6 方案。在直接 ELISA 实验中使用纯化 A/PR/8/34 病毒体作为抗原, 采用 450nm 的光密度值使其两倍于背景值, 取 5 或 6 只小鼠的血清, 抗体效价值用混合血清 (pooled serum) 最高稀释倍数的倒数表示。

图 8 为用 rMBB $\Delta$ -NP 免疫后再以流感病毒感染的小鼠肺内病毒负荷的绘图。空心圆圈代表每组的平均值, 条柱代表标准差。所示数据为两次实验中的一组数据。

图 9A 为证明天然 HA 在兔网织红细胞裂解液内进行体外翻译的 SDS-PAGE 分析。rM $\Delta$ FM-HA 重组复制子编码的流感病毒 HA 多肽以实心箭头表示, 未裂解的前体以空心箭头表示。

图 9B 为证明单顺反子门戈病毒复制子不能在转染真核细胞内表达外源糖基化蛋白的狭线杂交。在所示转染后时间点, 提取细胞质 RNA 并将其狭线杂交至一张尼龙膜上以供分析。

图 10 为转染重组门戈病毒复制子的 [ $^{35}$ S]甲硫氨酸标记 HeLa 细胞内表达 GFP 融合多肽免疫沉淀的 SDS-PAGE 分析。上样情况如下: 假转染的 HeLa 细胞或转染复制子 RNA rM $\Delta$ BB-GFP、rM $\Delta$ BB-GFP-NP118 (2 个克隆), 或转染复制子 rM $\Delta$ BB-GFP-lcmvNP 的 HeLa 细胞。分子量在图左显示。

图 11 为证明在以 rM $\Delta$ BB-GFP-NP118 和 rM $\Delta$ BB-GFP-lcmvNP 复制子 RNA 免疫的 BALB/c 小鼠体内, 以及在作为对照的以 pCMV-NP 和 pCMV-MG34 质粒 DNA 免疫的小鼠体内 LCMV 特异性 T 细胞的诱导的 ELISPOT 实验。在最末一次注射 3 周后, 用 IFN $\gamma$  ELISPOT 实验在有免疫显性 NP118-126 肽存在的条件下测定 LCMV 特异性 CD8<sup>+</sup> T 细胞的频率, 具体方法详见材料与方法。所获数据用每 10<sup>5</sup> 个脾细胞中的 SFC 数目表示。

图 12 为用荧光细胞计数仪阅读重组门戈病毒复制子 rM $\Delta$ BB-GFP 或 rM $\Delta$ XBB-GFP-NP118 或 rM $\Delta$ BB-GFP-lcmvNP 转染的 HeLa 细胞内 GFP 的表达。

## 实施例

将门戈病毒基因组衍生的复制子 cDNA 以正有义方向 (positive sense orientation) 克隆入一个细菌质粒, 位于 T7 RNA 聚合酶 I 启动子的下游及一个独特的 *Bam*H I 裂解位点的上游。用 *Bam*H I 将细菌质粒线性化后, 用 T7 RNA 聚合酶去合成一个病毒 RNA 样转录子, 该转录子可用于转染动物细胞或用于动物宿主的体内注射。

第一个复制子系列, 即 rM $\Delta$ BB 系列的构建方法详见材料与方法及例 1。几乎所有的 L-P1-2A 前体的编码序列都被缺失, 只留下了 CRE。这些复制子确实在转染的 HeLa 细胞内复制, 并随后与载体衍生的残余部分一起以融合蛋白的形式表达 GFP 或流感病毒 NP。rM $\Delta$ BB-NP 复制子以裸 RNA 的形式注射后, 在小鼠体内诱导出抗 NP 的免疫反应。在此基础上, 又构建了其他复制子, 它们也确实如在例 9 中所述的进行复

制，并随后表达淋巴细胞性脉络丛脑膜炎病毒（LCMV）的 NP 以及 H2<sup>d</sup> 小鼠 LCMV 免疫显性 NP118-26 表位相应的合成多肽。

第二个复制子系列，即 rMΔFM 系列的构建的目的在于以一种更天然的方式表达外源蛋白序列，其方法为将与外源蛋白序列融合的载体序列量最小化。这些 rMΔFM 复制子同样也在转染的 HeLa 细胞中复制。与此相反的是，包含流感病毒 HA 完整序列，包括其 SP 和 TM 区在内的 rMΔFM-HA 重组复制子并不具有复制能力。

小核糖核酸病毒（Picornavirus）基因组在正常情况下不编码糖蛋白。本发明人注意到，与其过去显示的脊髓灰质炎病毒基因组衍生的复制子一样，单顺反子门戈病毒衍生的复制子不能表达外源糖基化蛋白。但是，本发明人过去曾证明双顺反子脊髓灰质炎病毒（PV）复制子能表达糖蛋白。明确地说，本发明人构建了双顺反子复制子 rΔPV-IR-HA，该复制子对于 HA 和 PV 序列的翻译被 EMCV 内源性核糖体插入位点（IRES）的插入所解耦联了。rΔPV-IR-HA 复制子在转染后进行复制，并在细胞表面表达正确糖基化的 HA（29）。同样，双顺反子门戈病毒复制子的构建也可通过外源性、病毒或哺乳动物 IRES 的插入而完成，并可测试其复制及指挥糖基化蛋白如病毒或肿瘤抗原，或生物活性多肽表达的能力。

## 材料与方法

### 细胞、病毒及质粒

HeLa 细胞（ATCC 保藏号 CCL-2）在 37°C、5% CO<sub>2</sub> 的条件下在加入 5% 热灭活胎牛血清（FCS）（TechGen # 8010050）的 DMEM 完全培

培养基（Dulbecco 改进的 Eagle 培养基加入 1mM 丙酮酸钠、4.5mg/ml L-葡萄糖、100U/ml 青霉素及 100 $\mu$ g/ml 链霉素）中生长。

EL4（小鼠淋巴瘤，H-2<sup>b</sup>）（ATCC 保藏号 TIB-39）及 P815（小鼠肥大细胞瘤，H-2<sup>d</sup>）（ATCC 保藏号 TIB-64）在加入 10% FCS 的 RPMI 完全培养基（PRMI 1640，10mM HEPES，50 $\mu$ M  $\beta$ -巯基乙醇，100U/ml 青霉素，100 $\mu$ g/ml 链霉素）内保存。

适于小鼠的流感病毒 A/PR/8/34(ma) (H1N1)由感染裸鼠的肺组织匀浆液系列传代获得，具体方法前文已有描述（20）。以后的病毒原料通过单个尿囊途径在 11 日龄的胚鸡蛋内生产，该方法生产的病毒不影响其对小鼠的致病性。

质粒 pCI-NP 的构建是通过在表达质粒 pCI（Promega # E1731）的 *Sal* I 和 *Sma* I 位点之间插入流感病毒 NP 的编码序列而完成的，其位置在 CMV 直接-早期增强子/启动子的下游，具体方法在其他文献中已有描述（30）。质粒 pCI-NP 包含 A/PR/8/34(ma) NP cDNA 的共有序列，需要者可通过向本发明人索取而获得，同时该质粒还包含一位于密码子 107 的沉默突变（E: GAG $\rightarrow$ GAA），以及另外一个位于密码子 277 的突变（Pro $\rightarrow$ Ser）。密码子 277 的突变不直接影响 I 类主要组织相容性复合物（MHC-I）限制性免疫显性表位 NP366-374。

#### 用于重组复制子体外翻译的质粒的构建

包含去除及取代 L-P1-2A 部分的门戈病毒 cDNA 质粒是由质粒 pMC24（又叫 pM16.1；由威斯康辛大学，Madison, WI 的 Ann Palmenberg 提供）衍生得到的，该质粒包含位于噬菌体 T7 启动子下游的减毒门戈病毒毒株的全长感染性 cDNA（8）。

质粒 pM $\Delta$ BB (SEQ ID NO: 26) 包含一个亚基因组门戈病毒 cDNA, 其中第 737 至 3787 位的核苷酸被一个 *Sac* I/*Xho* I 多接头 (GAGCTCGAG) (SEQ ID NO: 1) 所取代, vMC24 cDNA 的第 1137 至 1267 位核苷酸包含门戈病毒 CRE (图 1)。质粒 pM $\Delta$ BB 的构建是通过用 *Bst* B I 消化质粒 pMN34 (15), 再进行自我拼接而完成的。含有 pM $\Delta$ BB 的细菌于 2001 年 5 月 21 日保藏于国立微生物培养物保藏中心 (Collection Nationale de Cultures de Microorganismes, CNCM), 巴黎, 法国, 保藏号为 I-2668。质粒 pM $\Delta$ N34 在设计上与 pM $\Delta$ BB 相似, 但去除的门戈病毒基因组部分 (第 737 至 3680 位核苷酸) 较少。

质粒 pM $\Delta$ XBB 的构建是为了去除包含 pM $\Delta$ BB 质粒序列的 CRE。简而言之, 通过寡核苷酸 5'-TCGAGGCTAGCTT-3' (SEQ ID NO: 2) 及 5'-CGAAGCTAGCC-3' (SEQ ID NO: 3) 的退火获得了一个 *Xho* I-*Bst* B I 连接子, 并将其克隆入质粒 pMN $\Delta$ 34 的 *Xho* I 及 *Bst* B I 位点之间。用 Big Dye 终止子测序试剂盒 (Perkin Elmer # P/N 4303150) 及 ABI377 自动测序仪 (Perkin-Elmer) 测定阳性克隆的序列。

为了克隆, 将质粒 pEGFP-N1 (Clontech # 6085-1) 作为模板、各自包含一个 *Sac* I 限制性内切酶位点 (划线部分) 的寡核苷酸 5'-GCTGAGCTCATGGTGAGCAAGGGCGAGGAGC-3' (SEQ ID NO: 4) 及 5'-GCAGAGCTCCTTGTACAGCTCGTCCATGCCG-3' (SEQ ID NO: 5) 作为引物, 用校阅 *PWO* 聚合酶 (Roche # 1644947)、通过 PCR 法扩增 GFP 编码序列。在质粒 pM $\Delta$ BB 及 pM $\Delta$ XBB 框内的 *Sac* I 位点插入 GFP 序列, 分别得到质粒 pM $\Delta$ BB-GFP 及 pM $\Delta$ XBB-GFP。用上文的方法测定阳性克隆的序列。

pM $\Delta$ BB-NP 质粒分两步构建。根据重叠扩展 PCR 方法，用 *PWO* 聚合酶产生包含流感病毒 A/PR/8/34 (ma) NP 突变 cDNA 的重组 cDNA 片段。诱变的目的在于将第 277 位密码子还原为正确的 Pro277，并在第 160 位密码子引入一沉默突变 (D: GAT→GAC)，从而破坏一个 *Bam*HI 位点以便进行下面的实验。简而言之，用 PCR 将质粒 pCI-NP 与寡核苷酸 5'- TCTCCACAGGTGTCCACTCC- 3' (SEQ ID NO: 6) 及 5'- CACATCCTGGGGTCCATTCCGGTGC GAAC -3' (SEQ ID NO: 7)、质粒 pTG-NP24 (与参考文献 30 中的 pTG-NP82 类似，但不包含 P277S 突变) 与寡核苷酸 5' - ACCGGAATGGACCCCAGGATGTGCTCTCTG - 3' (SEQ ID NO: 8) 及 5' - GTCCCATCGAGTGC GGCTAC -3' (SEQ ID NO: 9) 一起扩增，获得两个重叠 DNA 片段。用各自包含一个 *Xho*I 限制性内切酶位点 ( 划线部分 ) 的寡核苷酸 5'- CGGAATTCTCGAGATGGCGTCTCAAGGCACCAAACG-3' (SEQ ID NO: 10) 及 5'-GCGAATTCTCGAGATTGTCGTACTCCTCTGCATTGTC-3' (SEQ ID NO: 11) 获得的融合 PCR 产物再克隆入质粒 pTG186 的 *Eco*R I 位点 (13)，获得质粒 pTG-R4。用上文的方法测定阳性克隆的序列。其次，用 *Xho*I 消化 pTG-R4 获得 NP 编码序列，将该序列插入 pM $\Delta$ BB 的 *Xho*I 位点，使得 NP 序列和残余门戈病毒多蛋白序列都在框内。GFP 编码序列插入 pM $\Delta$ BB-NP 质粒的方法与 pM $\Delta$ BB 质粒的相同，使用的位点为一个独特的 *Sac*I 位点 (见上文)，从而获得质粒 pM $\Delta$ BB-GFP-NP。为了构建 pM $\Delta$ BB-GFP-lcmvNP 质粒，将寡核苷酸 5'- CGGAATTCTCGAGATGTCCTTGTCTAAGGAAGTTAAG-3' (SEQ. ID. NO 12) 及 5'-GCGAATTCTCGAGTGT CACAACATTTGGGCCTC-3' (SEQ. ID NO. 13) 作为引物，质粒 pCMV-NP 作为模板，用 PCR 法扩增 LCMV

病毒的 NP 编码序列。所获 DNA 片段克隆入质粒 pM $\Delta$ BB-GFP 的 *Xho* I 位点。用上文的方法测定阳性克隆的序列。

为了重建 NP118-126 H2<sup>d</sup> 的 LCMV 限制性免疫显性表位的编码序列，将浓度为 100  $\mu$ M 的寡核苷酸 5'TCGAAGCTAGCGAAAGACCCCAAGCTTCAGGTGTGTATATGGGTAATTTGACAC-3' (SEQ ID NO: 14) 及 5'TCGAGTGTCAAATTACCCATATACACACCTGAAGCTTGGGGTCTTTCGCTAGCT-3' (SEQ ID NO: 15) 在 750 mM Tris-HCl (pH 7.7) 内，在 100°C 的条件下处理 5 分钟，再在 20°C 条件下处理 1 小时进行退火而获得一个合成连接子。将该连接子插入 pM $\Delta$ BB-GFP 质粒的 *Xho* I 位点，获得质粒 pM $\Delta$ BB-GFP-NP118。用上文的方法测定阳性克隆的序列。

为了构建 pM $\Delta$ FM 质粒 (SEQ ID NO: 27)，将浓度为 100  $\mu$ M 的寡核苷酸 5'TCGAGGCTAGCCAGCTTTGAATTTTGACCTTCTTAAGCTTGC GGAGACGTCGAGTCCAACCCTGGGCCCT-3' (SEQ ID NO: 16) 及 5'TCGAAGGGCCCAGGGTTGGACTCGACGTCTCCCGCAAGCTTAAGAAGGTCAA AATTCAACAGCTGGCTAGCC-3' (SEQ ID NO: 17) 在 750 mM Tris-HCl (pH 7.7) 内，在 100°C 的条件下处理 5 分钟，再在 20°C 条件下处理 1 小时进行退火而获得一个合成连接子。将该连接子插入 pM $\Delta$ BB 质粒的 *Xho* I 位点，获得质粒 p2AB。接下来，将浓度寡核苷酸 5'-CGAGCATG-3' (SEQ ID NO: 18) 及 5'-CTAGCATGCTCGAGCT-3' (SEQ ID NO: 19) 退火获得第二个连接子。将该连接子插入 p $\Delta$ 2AB 的 *Sac* I 和 *Nhe* I 位点，获得质粒 pM $\Delta$ FM。用上文的方法测定阳性克隆的序列。包含 pM $\Delta$ FM 质粒的细菌于 2001 年 5 月 21 日保藏于 CNCM，保藏号为 I-2669。

为了克隆流感病毒 HA 序列，采用标准 RNA 提取步骤用 5M 硫氰酸胍硫氰酸胍和苯酚从感染 A/PR/8/34(ma)的小鼠肺组织匀浆中提取病毒基因组 RNA。将所获病毒 RNA 逆转录为 cDNA。接下来，用 *PWO* 聚合酶、以及寡核苷酸 5'-CTGGATCCAAAATGAAGGCAAACCT-3' (SEQ ID NO: 20)和 5'-CAGGATCCTAGATGCATATTCTGCACTG-3' (SEQ ID NO: 21)、通过 PCR 法扩增在起始密码子之前和终止密码子之后都包含 *Bam* HI 位点的 HA 编码序列。将所获 DNA 片段克隆入质粒 pTG186 的 *Bam* HI 位点，获得质粒 pTG-HA8。

再用寡核苷酸 5'-GAAAGGCAAACCTACTGGTCCTGTT-3' (SEQ ID NO: 22)及 5'-CGTGCAGTCGACAGGATGCATATTCTGCACTGCAAAG-3' (SEQ ID NO: 23)作为引物，质粒 pTG-HA8 作为模板，用 PCR 扩增 A/PR/8/34(ma)病毒的 HA 编码序列。寡核苷酸引物的设计要使所获 DNA 片段能被 *Sal* I 消化，并在 *klenow* 破坏的 *Sal* I 位点和 *Nhe* I 位点之间克隆入质粒 p2ΔAB 的框内，获得质粒 pMΔFM-HA。用上文的方法测定阳性克隆的序列。该质粒包含一个重组复制子 cDNA，其中翻译起始密码子 AUG 后跟随的 HA 序列与口蹄疫病毒 (FMDV) 2A/2B 自我催化裂解位点在框内融合，其后跟随 CRE，原始门戈病毒 2A/2B 裂解位点以及病毒多蛋白的残余部分 (见图 1)。

### 质粒 DNA 的体外转录

用 *Bam*HI 将门戈病毒衍生的质粒线性化，再根据生产厂商的使用说明用 Promega RiboMAX-T7 大规模 RNA 生产系统 (Promega # P1300) 对其进行转录。在体内实验中，采用标准分子生物学技术，用 RQ1 DNA 酶 (1.5 U/μg DNA, Promega # M6101) 在 37°C 处理反应混合

物 20 分钟，用苯酚-氯仿抽提一次，先在乙酸铵-异丙醇溶液中沉淀，再在乙酸钠-异丙醇内沉淀，将沉淀在无内毒素 PBS 中复悬（生命科学）。在体外翻译实验中，用相同步骤处理反应混合物，但仅用乙酸铵-异丙醇沉淀一次，并将沉淀复悬于无 RNA 酶的水中。

### 兔网织红细胞裂解液的体外翻译

用 Flexi™兔网织红细胞裂解液系统（Promega # L4540）加入 0.8 mCi/ml [<sup>35</sup>S]-甲硫氨酸(Amersham # SJ1515; 1000 Ci/mmol)、0.5 mM MgCl<sub>2</sub> 及 100 mM KCl，将体外合成的 RNA（10μg/ml）进行体外翻译。在 30℃ 孵育反应混合物 3 小时，在 30℃用 100 μg/ml RNA 酶 A 在 10mM 的 EDTA 内处理 15 分钟，在 12% SDS 聚丙烯酰胺凝胶内进行电泳分析，穿孔将凝胶在柯达 X-OMAT 胶片上放射自显影。

### RNA 转染

用 Easyject plus 电穿孔装置(Equibio)通过电穿孔法将 RNA 转染入 HeLa 细胞。简而言之，将 16x10<sup>6</sup> 个细胞用胰蛋白酶消化，用 PBS 洗两次，将沉淀复悬于 800 μl 以冰预冷的 PBS 中，并在 32 μg RNA 或 DNA 存在的条件下以单个脉冲（240 V, 1800 μF, 最大阻抗）在 0.4cm 电极间隙的电穿孔槽中电穿孔。细胞立即转移入加入 2%FCS 的 DMEM 完全培养基内，分装入 8 个 35mm 直径的组织培养皿中，在 37℃、5% CO<sub>2</sub> 条件下孵育。

### RNA 复制的分析

在转染后的不同时间间隔，用标准步骤制备细胞质 RNA（26）。在 1×SSC、50%甲酰胺、7%甲醛中于 65℃温育变性 15 分钟后，将 RNA 样品转移至一张尼龙膜（Hybond N, Amersham # RPN203N）上，并用

$^{32}\text{P}$  标记的 RNA 探针杂交, 该探针与门戈病毒 RNA 的第 6022 至 7606 位核苷酸互补。杂交在  $65^\circ\text{C}$  条件下, 在含  $6\times\text{SSC}$ 、 $5\times\text{Denhardt}$  溶液及  $0.1\%$  SDS 的溶液内进行 18 小时。在室温下用  $2\times\text{SSC}$ 、 $0.1\%$  SDS 溶液将膜洗 3 遍, 再在  $65^\circ\text{C}$  下用  $0.1\times\text{SSC}$ 、 $0.1\%$  SDS 溶液洗 3 遍。最后将膜曝光于 STORM™ 820 磷屏成像仪 (分子动力学), 并用 Image Quant 程序分析 (分子动力学)

### RNA 转染细胞内 GFP 表达的分析

用所述方法转染 HeLa 细胞。在转染后 8 至 12 小时, 用胰蛋白酶消化细胞, 用 PBS 洗涤, 并在  $100\mu\text{l}$  PBS、 $1\%$  多聚甲醛内于  $4^\circ\text{C}$  孵育 60 分钟进行固定。接下来在 FACScalibur 荧光细胞计数仪 (Becton-Dickinson) 上分析样品的荧光强度。

### RNA 转染细胞内流感病毒 NP 表达的分析

在峰表达 (peak expression) 时刻, 用  $[\text{}^{35}\text{S}]$ -甲硫氨酸 ( $50\ \mu\text{Ci/ml}$ ; Amersham;  $1000\ \text{Ci/mmol}$ ) 对流感病毒 A/PR/8/34 感染或 RNA/DNA 转染细胞进行代谢标记 2 小时。峰表达时刻是通过通过对 rM $\Delta$ BB-GFP 复制子 RNA 或 pCI-GFP 质粒 DNA 转染的 HeLa 细胞内 GFP 的表达进行研究而确定的。对于 RNA 转染细胞, 在转染后 6 至 9 小时可观察到峰表达。对于 DNA 转染细胞, 在转染后 20 小时可观察到峰表达。流感病毒 A/PR/8/34 转染的 HeLa 细胞在转染后 20 小时观察到峰表达。接下来用 PBS 洗涤细胞, 并用  $50\ \text{mM}$  Tris-HCl ( $\text{pH } 7.5$ )、 $150\ \text{mM}$  NaCl、 $1\ \text{mM}$  EDTA、 $1\%$  NP40 及  $0.5\%$  蛋白酶抑制剂 Cocktail (Sigma) 裂解细胞。细胞提取物在有吸附兔抗流感病毒 A/PR/8/34 抗体的蛋白 A 琼脂糖凝胶珠 (Amersham Pharmacia Biotech # 17-0780-01) 的 RIPA 缓冲液 ( $50\ \text{mM}$

Tris-HCl, 150 mM NaCl, 1 mM EDTA, 0.1% 脱氧胆酸盐, 0.1% 十二烷基硫酸钠, 0.5% NP40 和 0.5% 蛋白酶抑制剂 Protease Inhibitor Cocktail) 中于 4°C 免疫沉淀过夜。用 RIPA 缓冲液洗涤免疫沉淀物, 用 Laemmli 样品缓冲液 (50 mM Tris-HCl (pH 6.8), 2% SDS, 5%  $\beta$ -巯基乙醇, 20% 甘油) 于 65°C 进行萃取, 用 SDS-PAGE 分析, 并在柯达 X-OMAT 胶片上放射自显影使其显像。

### RNA 转染细胞内 GFP 融合蛋白表达的分析

用上文中描述的 NP 表达分析方法对 RNA/DNA 转染 HeLa 细胞的提取物进行免疫沉淀和分析, 但使用的抗体为兔抗 GFP 抗体 (Invitrogen #46-0092)。

### 免疫

用包含 50  $\mu$ g 质粒 DNA 或 25  $\mu$ g 门戈病毒复制子 RNA 的 100  $\mu$ l PBS 对周龄为 7 至 8 周的 C57BL/6 雄性小鼠 (IFFA CREDO) 进行肌肉注射 (i.m.) (双侧胫前肌各 50  $\mu$ l)。间隔 3 周再以 i.m. 的形式行加强注射。用于注射的 DNA 是用 Nucleobond PC2000 试剂盒 (Nucleobond # 740576) 制备的, 随后再用 triton X 114 进行提取步骤, 最后用苯酚-氯仿进行抽提。用 QCL-1000 内毒素试剂盒 (BioWhittaker # 50-647U) 检测样品是否确实不含有内毒素 (<100 U/mg)。RNA 的制备在前文中已进行了分析, 在注射后用琼脂糖凝胶电泳证实其未被降解。

### 抗体效价

在最末一次注射的 3 周后取小鼠的血液。将混合血清连续稀释以用 ELISA 法测定 NP 特异性抗体效价, 每孔用 0.5  $\mu$ g 被去污剂破坏的 A/PR/8/34 病毒作为抗原。简而言之, 在 0.2 M 碳酸钠、0.2 M 碳酸氢钠

(pH 9.6) 溶液内, 将 96 孔 ELISA 板 (NUNC Maxisorp, # 439454) 于 4°C 用 0.5  $\mu\text{g}$  被去污剂破坏的 A/PR/8/34 病毒包被过夜。用与辣根过氧化物酶(HRP) (Biosystems # BI2413C) 偶联的 1/2000 稀释抗小鼠 IgG(H+L) 抗体检测结合抗体, 并根据厂商的说明加入 TMB 过氧化物酶底物 (KPL # 50-76-00) 使其显像。

在 450nm 波长光密度值为背景值两倍的混合血清稀释倍数的倒数即为抗体的效价。每组混合血清由 4 或 5 只小鼠制备。

### 细胞毒性实验

在最末一次免疫后 3 周收集小鼠脾细胞, 并将其以  $2 \times 10^6$  个细胞/ml 的密度接种于直立 T75 瓶内, 培养基为加入 10% FCS、1.0mM 非必需氨基酸、1mM 丙酮酸钠及 2.5% 伴刀豆球蛋白 A 上清的 RPMI 完全培养基。用与 10  $\mu\text{M}$  NP366 肽(ASNENMETM, Neosystem; SEQ ID NO: 24) 一起在加入 5%FCS 的 PRMI 完全培养基内于 37°C 脉冲处理 3 小时、洗涤及照射 (2500 拉德) 的同基因脾细胞 ( $10^6/\text{ml}$ ) 对脾细胞再刺激 7 天。用标准 4 小时  $^{51}\text{Cr}$  释放细胞毒性实验测定再刺激效应细胞的细胞毒活性, 基本原理已有前文论述 (9)。在  $^{51}\text{Cr}$  标记过程中用或不用脉冲将 EL4 和 P815 靶细胞与 NP366 肽一起处理。将细胞在培养基中单独孵育或在 1% triton X-100 中孵育, 分别测定放射活性的自发值和最大值。特异性  $^{51}\text{Cr}$  释放的百分比用如下公式计算: (实验释放值-自发释放值) / (最大释放值-自发释放值)  $\times 100$ 。

### IFN $\gamma$ ELISPOT 实验

在最末一次接种后 3 周收集小鼠脾细胞, 并用标准 IFN $\gamma$  ELISPOT 实验系统检测流感病毒或 LCMV 病毒特异性 CD8+ T 细胞的存在。简而

言之，用 1 $\mu$ M 流感病毒 NP366 合成肽 (ASNENMETM, Neosystem; SEQ ID NO: 24)、LCMV NP118-26 肽(RPQASGVYM, Neosystem, SEQ ID NO: 25)及 IL-2 (10 U/ml)刺激脾细胞 20 小时，并在已用大鼠抗小鼠 IFN $\gamma$  抗体 (R4-6A2, Becton-Dickinson)包被的 96 孔 Multiscreen HA 硝酸纤维素板 (Millipore) 中每孔加入 5X10<sup>5</sup> 经照射 (2000 拉德) 同基因脾细胞作为饲养细胞。将细胞与结合了生物素的大鼠抗小鼠 IFN $\gamma$  抗体 (XMG1.2, Becton-Dickinson)、碱性磷酸酶偶联的链霉抗生物素蛋白 (Becton-Dickinson) 及 BCIP/NBT 底物 (Sigma) 一起连续孵育以显示斑点。通过计数每孔中斑点产生细胞 (SFC) 的数目确定产 IFN $\gamma$  细胞的频率。所获结果用每 10<sup>5</sup> 个脾细胞中的 SFC 数目表示。

#### 用 A/PR/8/34(ma)病毒感染小鼠进行免疫试验

在第三次免疫后 1 周或 3 周，用 100 mg/kg 氯胺酮 (Merial) 将 C57BL/6 小鼠轻度麻醉，并用溶解于 40  $\mu$ l PBS 内的 100 pfu (0.1 LD<sub>50</sub>) A/PR/8/34(ma) 病毒经鼻感染小鼠。在感染后 7 天杀死小鼠。制备肺组织匀浆，并用标准噬斑实验 (36) 在单层 MDCK 细胞上测定病毒的效价。用小样本假设 (变量常态分布，进行比较的总体偏差相同) 以 Student's 独立 t 检验对个体小鼠病毒效价的对数 (log<sub>10</sub>) 进行统计学分析。

包含质粒 pM $\Delta$ BB 和 pM $\Delta$ FM 的菌种存放于 CNCM，巴斯德研究院，28, rue du Docteur Roux, 75724 Paris Cedex 15, 法国，如下表所示：

质粒	保藏号	存放日期
pM $\Delta$ BB (SEQ ID NO: 26)	I-2668	May 21, 2001
pM $\Delta$ FM (SEQ ID NO: 27)	I-2669	May 21, 2001
pM $\Delta$ BB-GFP-lcmvNP (SEQ ID NO: 28)	I-2879	May 16, 2002

### 实施例1: 制备衍生自门戈病毒基因组的重组复制子

为了制备门戈病毒基因组衍生的复制子,首先要构建质粒载体 pM $\Delta$ BB,其中衣壳蛋白的 L-P1-2A 前体蛋白编码序列被一个 *Sac* I/*Xho* I 多接头及原先位于 VP2 衣壳蛋白编码序列内的门戈病毒 CRE (15) 所取代。该取代的操作仍保留了最佳 2A/2B 自我催化裂解位点相应的序列,该位点的组成为 2A 的 C 端 19 个氨基酸和 2B 的第一个氨基酸 (7) (图 1)。明确地说,去除了质粒 pMC24 (其包含位于 T7 噬菌体 $\phi$ 10 启动子的下游的门戈病毒减毒株完整感染性 cDNA) 的第 737-3787 位核苷酸,即编码结构蛋白、L 蛋白和 2A 蛋白的 L-P1-2A 区。去除的序列被一个 *Sac* I, *Xho* I 多接头及包含门戈病毒 CRE 的序列所取代。如上所述,包含在 2A/2B 位点顺式自我催化裂解最佳序列的、编码 2A 的 C 端 22 个氨基酸的序列被保留下来。所获质粒 pM $\Delta$ BB (SEQ ID NO: 26, 8017 个碱基对)能与 T7 RNA 聚合酶一起体外转录合成 rM $\Delta$ BB 复制子 RNA。SEQ ID NO: 26 的第一个碱基对应于复制子 RNA 的第一个碱基,用于在转录前将质粒线性化的 *Bam*H I 位点位于 4837 位, T7 启动子为第 7999 位至第 8017 位核苷酸, 2 个 G 残基 (第 8016 和第 8017 位核苷酸) 实际上是由 T7 RNA 聚合酶得到的该质粒合成转录子的组成部分。

将编码 GFP、流感病毒 NP 或 GFP-NP 融合蛋白的序列插入位于 CRE 及重建 2A/2B 裂解位点上游的 pM $\Delta$ BB 多接头,与编码门戈病毒多蛋白的剩余序列一起位于框内,从而获得质粒 pM $\Delta$ BB-GFP、pM $\Delta$ BB-NP 及 pM $\Delta$ BB-GFP-NP (图 1)。

作为阴性对照的质粒 pM $\Delta$ XBB 和 pM $\Delta$ XBB-GFP 分别与 pM $\Delta$ BB 和 pM $\Delta$ BB-GFP 类似,但其中的 $\Delta$ X 构建部分不含门戈病毒 CRE (图 1)。

在此应用的所有质粒都是通过业内已知技术在实验室中获得的。它们的核苷酸序列都是已知且可获得的。用 PCR 对这些序列进行扩增时，已对插入部分进行了完全测序从而对其进行了检验。

用 T7 RNA 聚合酶对 pM $\Delta$ BB、pM $\Delta$ BB-GFP、pM $\Delta$ BB-NP 和 pM $\Delta$ BB-GFP-NP 质粒 DNA 进行体外转录，用 *Bam* HI 线性化所获的 rM $\Delta$ BB、rM $\Delta$ BB-GFP、rM $\Delta$ BB-NP 和 rM $\Delta$ BB-GFP-NP 重组 RNA 在兔网织红细胞裂解液内进行体外翻译。用 SDS-PAGE 分析翻译产物并通过放射自显影使其显像。如图 2 所示，复制子编码的多蛋白被 3C 蛋白酶正确地裂解，以表达 RNA 扩增所必需的非结构蛋白，证据为裂解的最终产物如 2C、3C、3D 和 3CD 蛋白。与此相反的是，没有在 rM $\Delta$ BB 衍生的复制子中观察到重建 2A/2B 位点的正确顺式裂解。本发明人预测，外源序列会与 7 个连接子编码残基、CRE 编码多肽（CREP，44 个氨基酸）及 2A 蛋白的最后 22 个残基一起表达为融合蛋白，从而将外源多肽扩大为 8 kD。对于重组 rM $\Delta$ BB-NP 复制子而言，正确裂解的 NP-CREP-2A\* 融合蛋白的表达会显示为在预测分子量为 63kDa 的位置出现一条条带，而实际上在分子量为 70kDa 或稍大一些的位置上观察到一条条带（图 2）。与此相反的是，GFP-CREP-2A\* 和 GFP-NP-CREP-2A\* 融合蛋白移行的分子量与预测的相近（分别为 35 kDa 和 89 kDa）。

本发明人解释由 rM $\Delta$ BB-NP 复制子编码的 NP 蛋白预测大小与实际大小之间的明显不一致是由于 2A/2B 裂解没有发生，而是在 2B 多肽内出现了一个代替裂解，其位置大约距其 N 端三分之一（假设 2B 蛋白大小为 151 个氨基酸）。在此例中，由 rM $\Delta$ BB-NP 载体编码的 NP 相关性异源性序列以 NP-CREP-2A\* $\Delta$ 2B 融合多肽的形式表达。可能是位于裂解位点之前由 NP 序列及 CRE 编码的氨基酸的延展（stretch）迫使 2A 序

列的残余部分折叠，从而使裂解不能进行。本发明人目前对于 2B 多肽内出现的异常裂解尚没有解释，但是对其他细小病毒的代替处理方法已进行了描述，尤其是当处理串联步骤中的一个裂解被阻断时的情况（4）。

#### 实施例 2：门戈病毒基因组衍生复制子 rMΔBB、rMΔBB-GFP、rMΔBB-NP 及 rMΔBB-GFP-NP 的复制特点

本发明人接下来对外源序列是否能在不影响 RNA 复制的前提下插入门戈病毒基因组进行了确定。此外，由于流感病毒 NP 在以往已显示出能与 RNA 发生非特异性联合的性质（14, 32），其与门戈病毒 RNA 之间发生的反应在理论上会影响整体复制效率。因此，将 rMΔBB、rMΔBB-GFP、rMΔBB-NP 及 rMΔBB-GFP-NP 的合成 RNA 转录子转染入 HeLa 细胞，并在转染后不同时间点提取全部细胞质 RNA。狭线杂交后用与门戈病毒 RNA 第 6022-7606 位核苷酸互补的<sup>[32P]</sup>放射性标记 riboprobe（体外转录）与 RNA 杂交发现，所有的 RNA 都进行了高效复制（图 3）。在本发明人的研究中用电穿孔法转染细胞，该法比经典的 DEAE-葡聚糖技术的转染效率高（能转染 50%以上的细胞）。在这些条件下，不论有没有衣壳蛋白存在，所有的 4 种 RNA 都诱导出了一种细胞病理效应（CPE），并导致单层细胞在转染后 24 小时全面破坏（数据未在本文中显示）。综上所述，这些结果说明插入外源序列如 GFP 或 NP 编码序列对于 RNA 的复制没有负效应。

#### 实施例 3：重组门戈病毒衍生的复制子对绿色荧光蛋白的表达

通过监测门戈病毒衍生复制子转染细胞内的 530nm 荧光，用荧光细胞计数仪分析 GFP 的表达。用电穿孔将 HeLa 细胞假转染或用

rM $\Delta$ BB、rM $\Delta$ BB-GFP、rM $\Delta$ XBB-GFP 复制子 RNA 转染。在转染后 9 小时，用胰蛋白酶消化细胞，并用 FACScalibur 荧光细胞计数器分析其荧光强度，根据预试验结果，这时所有的检测复制子都处于 GFP 表达高峰期内（7 至 12 小时）。如图 4 所示，在 rM $\Delta$ BB-GFP 转染细胞内能检测到 GFP 的表达，而在假转染细胞内或空载体 rM $\Delta$ BB 转染细胞内则无 GFP 的表达。有趣的是，复制子 rM $\Delta$ XBB-GFP RNA 转染的细胞内没有显示出任何荧光，该结果证实了门戈病毒 CRE 对于 RNA 复制是必需的结论，而且证实了外源序列的显著表达必需在 RNA 复制的前提下进行。因此，门戈病毒衍生重组复制子显示出能在转染细胞内指导 GFP 高效表达的能力。

#### 例4: 重组门戈病毒衍生复制子对流感病毒核蛋白的表达

如方法中所述，通过加入抗 A/PR/8/34 病毒抗体，用免疫沉淀法处理门戈病毒衍生的复制子转染的细胞或 A/PR/8/34 病毒感染的细胞的细胞质提取物。根据预试验结果，通过电穿孔法用复制子 RNA 转染 HeLa 细胞，并在表达峰值时用 [<sup>35</sup>S]-甲硫氨酸代谢标记 2 小时。制备细胞质提取物，用抗流感病毒 A/PR/8/34 多克隆抗体免疫沉淀蛋白质，用 SDS-PAGE 对其进行分析并通过放射自显影使其显像。如图 5 所示，在 rM $\Delta$ BB-NP 转染细胞的提取物（第 3 泳道）内，明显有一分子量为 70 kDa 的蛋白质被特异性免疫沉淀。正如预期的那样，在假转染细胞或空载体 rM $\Delta$ BB 复制子 RNA 转染的细胞细胞质内没有观察到任何免疫反应蛋白。

门戈病毒衍生复制子表达的 NP 融合多肽明显移动至分子量 70 kD 处（图 5，第 3 泳道），其分子量显著大于 A/PR/8/34 病毒感染细胞所表

达的天然 NP 的分子量 (55 kD) (第 6 泳道)。如实施例 1 中所讨论的, 这个分子量之间的差异是由 NP-CREP-2A\*融合蛋白的额外氨基酸残基及 2B 蛋白的额外残基造成的, 该结果与体外实验所观察到的一致。同样, 该结果与门戈病毒多肽 2A/2B 位点没有发生蛋白裂解, 而是在 2B 序列内部的一个替代裂解位点发生了蛋白裂解的假说相一致。有趣的是, 这并不影响复制子 RNA 的整体复制效率, 提示该替代裂解途径可能为门戈病毒多蛋白连锁反应的一部分。

重组复制子 rM $\Delta$ BB-GFP-NP 转染的 HeLa 细胞 (图 5, 第 4 泳道) 同样引起 NP 相关蛋白的高水平表达。由于 NP 相关物质移动至大于预期值的分子量处 (实际值约为 97 kD, 而非预期的 89 kD), 因此估计在 2A/2B 位点同样也没有发生裂解。

因此, 门戈病毒衍生重组复制子显示出能在转染细胞内指导大小至少可多达 2200 个核苷酸的异源性序列高效表达的能力。

#### 实施例 5: 以裸 RNA 形式注射重组门戈病毒衍生的复制子后诱导 NP 特异性 CTL 应答

为了确定使用裸门戈病毒衍生复制子注射以引出一异源性特异性免疫应答的可行性, 本发明人确定了将重组 rM $\Delta$ BB-NP 以裸 RNA 形式注射能否诱导出 NP 特异性 CTL 应答, 明确地说是抗 NP 显性 H-2D<sup>b</sup> 限制性表位 NP366 的免疫应答。

为了实现这个目的, 用 25  $\mu$ g rM $\Delta$ BB-NP 裸 RNA 给 C57BL/6 肌肉注射两次, 中间间隔一月, 或用 50  $\mu$ g pCI-NP 裸 DNA 给小鼠肌肉注射一次作为阳性对照。该免疫程序根据过去的实验确定, 并基于对以下现象的观察, 即注射一次质粒 DNA 足已诱导一可检测的 NP 特异性 CTL

应答，其水平仅次于由亚致死剂量流感病毒 A/PR/8/34 (ma) 感染中恢复小鼠所获水平（数据未在本文中显示）。在最末一次注射 3 周后收获小鼠脾细胞，在体外以 NP366 肽刺激并在 7 天后用经典铬释放实验检测其溶细胞活性，详见材料与方法。注射 rM $\Delta$ BB-NP RNA 或 pCI-NP DNA 的小鼠脾细胞在负荷 NP366 肽的条件下特异性溶解同基因 EL4 细胞（图 6a）。由 r $\Delta$ BB-NP 复制子 RNA 诱导的 CTL 活性与 pCI-NP DNA 诱导的极为相似，且水平很高（即在效应细胞对靶细胞比为 6.7:1 时特异性溶解 60% 至 70%）。所有未免疫对照小鼠或用不包含 NP 序列的对照空载体免疫的小鼠脾细胞在刺激后都没有观察到任何溶细胞反应（图 6，空心标志）；在没有负荷肽的靶细胞中亦未观察到任何溶细胞反应（图 6b）。最后，无论是否用肽孵育，所有效应细胞对同源异体 P815 靶细胞（H-2<sup>d</sup>）的溶细胞反应都保持为背景水平（数据未在本文中显示），提示该溶细胞活性为 H-2 限制性，因而可能由 I 类限制性 CD8<sup>+</sup> T 细胞效应产生。

最后，用 IFN  $\gamma$  ELISPOT 实验对肌肉注射 rM $\Delta$ BB-NP RNA 及 pCI-NP DNA 诱导的两种特异性 T 细胞应答进行了定量测定。用流感病毒免疫显性 NP366 肽体外刺激免疫小鼠脾细胞，根据其应答水平测定 IFN $\gamma$  产生细胞的频率，详见材料与方法。如图 6c 所示，复制子 RNA 和质粒 DNA 免疫小鼠的 T 细胞频率都非常高且在同一范围内（100/10<sup>5</sup> 脾细胞）。正如预期的那样，两个空对照即在无 NP366 肽存在的条件下或空载体免疫小鼠的每 10<sup>5</sup> 个脾细胞所获 SFC 数目都小于 1。

这些发现显示门戈病毒复制子在以裸 RNA 形式注射时具有免疫原性，并且能诱导出抗注射外源序列如流感病毒 NP 的异源性特异性免疫性。

### 实施例 6: 重组复制子 rM $\Delta$ BB-NP 免疫后诱导 NP 特异性抗体

为了评估重组复制子 rM $\Delta$ BB-NP 在以裸 RNA 形式注射后能否诱导抗流感病毒抗原特异性抗体的产生, 用 25  $\mu$ g rM $\Delta$ BB-NP RNA 给 C57BL/6 小鼠肌肉注射 3 次, 每次间隔 3 周, 或用 50  $\mu$ g pCI-NP DNA 注射作为阳性对照。在最末一次注射 3 周后收集小鼠血清 (DNA 注射 1 或 2 次, RNA 注射 2 次)。用 ELISA 法测定特异性抗 NP 抗体。

如图 7 所示, 25  $\mu$ g 裸 rM $\Delta$ BB-NP RNA 注射两次可诱导血清抗流感病毒 NP 抗体的产生。其 NP 特异性 ELISA 效价稍高于注射一次 50  $\mu$ g 质粒 pCI-NP DNA 所诱导的, 但明显低于注射两次 pCI-NP DNA 所诱导的抗体水平。

如同实施例 5 一样, 这些发现显示门戈病毒复制子以裸 RNA 注射时具有免疫原性, 并能诱导出抗插入流感病毒 NP 外源序列的异源性特异性免疫应答。将实施例 5 和实施例 6 综合考虑, 两者都证明了门戈病毒复制子既能诱导抗编码异源性蛋白的体液免疫 (抗体), 又能诱导细胞免疫 (CTL)。

### 实施例 7: 体外保护性免疫

为了显示 rM $\Delta$ BB-NP 产生体外保护性免疫的能力, 用 25  $\mu$ g rM $\Delta$ BB 或 rM $\Delta$ BB-NP 复制子 RNA, 或 50  $\mu$ g pCI 或 pCI-NP 质粒 DNA 免疫 C57BL/6 小鼠 (每组 6 只), 共注射 3 次, 每次间隔 3 周。最末一次注射 3 周后, 用  $10^2$  pfu (0.1 LD<sub>50</sub>) 适于小鼠的 A/PR/8/34 病毒刺激小鼠, 并在刺激 7 天后测定小鼠肺内的病毒效价。如图 8 所示, 用每一种 NP 编码载体注射的小鼠体内的病毒负荷明显低于用相应空载体注射的小鼠病毒负荷 ( $p < 0.001$ ; student's t 检验)。

值得注意的是，尽管由于本发明人使用的适于小鼠病毒株的毒力很高（C57BL/6 小鼠的 LD<sub>50</sub> 为 10<sup>3</sup> pfu），病毒负荷仅中度下降，但裸 RNA 免疫所引起的下降与裸 DNA 免疫所引起的同样有效。该结果证实由门戈病毒衍生复制子裸 RNA 引起免疫应答（很可能是 CTL）能通过降低肺内病毒效价起到抗流感病毒的保护作用。

#### 实施例 8：门戈病毒基因组衍生重组 rMΔFM 复制子的制备

为了以一种更为天然的形式表达外源序列，本发明人探索了将与外源序列融合的载体序列最小化的可能性。为了达到这个目的，通过在 pMΔBB 编码复制子的多接头和 CRE 序列之间插入 FMDV 的 2A/2B 自我催化裂解位点序列构建质粒 pMΔFM（图 1）。在这个最佳方案中，该裂解位点包含 2A 蛋白的 C 端 19 个残基和 2B 蛋白的第一个脯氨酸（7）。

所获质粒 pMΔFM (8092 个碱基对)对应于 SEQ ID NO: 27：第一个碱基对应于复制子 RNA 的第一个碱基，用于在转录前将质粒线性化的 *Bam*HI 位点位于 4912 位，T7 启动子为第 8074 位至第 8092 位核苷酸，2 个 G 残基（第 8091 和第 8092 位核苷酸）实际上是由 T7 RNA 聚合酶得到的该质粒合成转录子的组成部分。

接下来，将流感病毒 A/PR/8/34(ma)的 HA 基因插入于 pMΔFM 的 *Sac* I 和 *Nhe* I 位点之间，紧邻 FMDV 2A 序列的上游，且与剩余多蛋白序列一起位于框内，获得质粒 pMΔFM-HA。

为了验证这些构建的质粒能翻译为多蛋白并如预期在终末产物内裂解，如上文所述在体外对相应线性化质粒进行了转录，并将合成的 RNA 在兔网织红细胞裂解液内翻译。所有的复制子都显示出正确裂解终末产

物的类似翻译轮廓，证据为 2C、3C、3D 及 3CD 病毒多肽的存在（图 9A）。

特别是在重组复制子 rM $\Delta$ BB-HA 内观察到重建 FMDV 2A/2B 位点的正确顺式裂解；在预期的 65 kDa 分子量处出现的条带（图 9A）表明了包含 FMDV 2A 蛋白（21 个氨基酸）和多接头（5 个氨基酸）的 26 个额外氨基酸残基的正确裂解 HA-2A\*融合蛋白的表达。有趣的是，在更大分子量处出现的条带提示该裂解在此体外翻译实验中并非 100%有效。

相应母代复制子 rM $\Delta$ FM 的该裂解产物以 3.4 kDa 大小的 MCS-2A 融合蛋白形式出现，由于其分子量小因而不可见，但有一分子量为 16 kDa 的多肽条带存在；该多肽可能与跨越门戈病毒 CRE、门戈病毒 2A 蛋白的最后 22 个残基及 2B 的 N 端的序列相对应，提示在本例中 FMDV 2A/2B 位点亦发生了裂解，而原始的门戈病毒 2A/2B 仍未裂解，该现象与上文所述的 rM $\Delta$ BB 及 rM $\Delta$ BB-NP 实例中的现象一致。

为了测试这些第二代复制子的复制有效性，通过电穿孔法用合成 RNA 转染 HeLa 细胞，在转染后不同时间点提取细胞质 RNA 并用 Northern 杂交分析，使用的探针为门戈病毒特异性<sup>[32P]</sup>标记的 riboprobe（体外转录），其序列与门戈病毒基因组第 6022-7606 位核苷酸互补。如图 9B 所示，rM $\Delta$ FM 复制子的确如同其母代 rM $\Delta$ BB 一样高效复制，提示该新操作的 2A/2B 裂解对于 RNA 合成没有负效应。另一方面，rM $\Delta$ FM-HA 重组复制子没有复制能力。

由于 rM $\Delta$ FM-HA 复制子中的 HA 包含一 SP 区和 TM 区，这一发现可能与另一细小病毒基因组即脊髓灰质炎病毒构建的复制子类似。确实

发现在紧邻脊髓灰质炎病毒复制子多蛋白 N 端处 SP 区的存在废除了相应 RNA 的复制 (1, 16)。通过显示在 rM $\Delta$ FM-HA 脊髓灰质炎病毒复制子内插入实质为糖基化跨膜蛋白的流感病毒 HA 的完整序列, 该复制子的复制被废除 (29), 本发明人近来证实了这一现象。此外, 本发明人还演示了用去除其 P1 区的脊髓灰质炎病毒基因组衍生复制子表达 HA 糖基化序列是可能的, 前提为通过在外源序列和剩余 P2P3 多蛋白序列之间插入一异源性 IRES 如 EMCV IRES 将这些复制子双顺反子化 (29)。

因此可构建双顺反子门戈病毒复制子。可通过在 *Sac* I/*Xho* I 多接头和 pM $\Delta$ BB 质粒剩余多蛋白序列之间插入一外源病毒或哺乳动物 IRES 完成该构建。例如通过插入 A 型或 B 型马鼻炎病毒外源 IRES 可构建所述双顺反子门戈病毒复制子, 这是因为这两者的 IRES 有效地竞争心肌炎病毒属典型代表 EMCV 病毒 IRES 的翻译因子 (38)。正如本发明人过去用双顺反子脊髓灰质炎病毒演示的那样, 这些双顺反子门戈病毒复制子能够进行复制并表达糖基化外源多肽。例如可将流感病毒 HA 序列插入所述新构建的双顺反子门戈病毒复制子中的一个内。

当糖基化是多肽抗原性或生物活性的关键参数时, 这些新构建的双顺反子门戈病毒复制子允许研究人员感兴趣的外源抗原或蛋白的表达。例如门戈病毒双顺反子复制子可用于表达病毒抗原如乙肝病毒 HBs 抗原或人免疫缺陷病毒的包装糖蛋白, 或用于表达肿瘤抗原如人肿瘤细胞的表面抗原。

如在兔网织红细胞裂解液内观察到的, 门戈病毒 rM $\Delta$ FM 复制子载体还能用于指导非糖基化外源蛋白在转染细胞内的天然表达。

### 实施例 9: 门戈病毒复制子对其他抗原 LCMV 核蛋白 (NP) 或 LCMV NP118-126 表位的表达

为了显示门戈病毒衍生复制子在以裸 RNA 形式接种后能诱导抗其他抗原的异源性特异性免疫应答, 本发明人构建了 rM $\Delta$ BB-GFP-lcmvNP 和 rM $\Delta$ BB-GFP-NP118 复制子。这些复制子以 GFP 融合蛋白的形式分别编码 LCMV NP 及 LCMV NP118-126 H2<sup>d</sup> 限制性免疫显性表位。

为了达到该目的, 本发明人构建了质粒 pM $\Delta$ BB-GFP-lcmvNP (SEQ ID NO: 28, 10417 个碱基对), 详见材料与方法。SEQ ID NO: 28 的第一个碱基对应于复制子 RNA 的第一个碱基。用于在转录前将质粒线性化的 *Bam*HI 位点位于 7237 位, T7 启动子为第 10399 位至第 10417 位核苷酸, 2 个 G 残基 (第 10416 和第 10417 位核苷酸) 实际上是由 T7 RNA 聚合酶得到的该质粒合成转录子的组成部分。

接下来, rM $\Delta$ BB-GFP-lcmvNP 复制子 RNA 转染 HeLa 细胞细胞质提取物在预期的 97 kDa 分子量处出现的条带表明 LCMV NP 与 GFP 以融合多肽的形式得到了表达 (图 10)。用荧光细胞计数仪在 530nm 波长监测复制子转染 HeLa 细胞内荧光的存在也同样证明了 GFP 的表达 (图 12)。与此相类似的是, NP118-126 LCMV 表位的前体大小为 15 个氨基酸 (NP 116-130, 大约为 1.7 kDa), 其表达也作为融合蛋白被检测到了, 分子量略大于 GFP (35 kDa)。这表明重组 rM $\Delta$ BB-GFP-lcmvNP 和 rM $\Delta$ BB-GFP-NP118 RNA 如同其母代 rM $\Delta$ BB-GFP 复制子一样, 确实进行了复制, 并能使插入序列得到合成。此外, 本例及例 3 都显示了 GFP 的表达作为门戈病毒衍生复制子 RNA 复制标记物的适用性。

最后，用 25  $\mu\text{g}$  rMBB-GFP、r $\Delta$ BB-GFP-lcmvNP 或 r $\Delta$ BB-GFP-NP118 裸 RNA 给 BALB/c 小鼠肌肉注射两次，或用 50  $\mu\text{g}$  pCMV-NP 或 pCMV-MG34 (40)裸 DNA 给小鼠注射作为阳性对照。用 LCMV 免疫显性 NP118-126 肽对经免疫小鼠的脾细胞进行体外刺激，并用 IFN $\gamma$  ELISPOT 实验测定产 IFN $\gamma$ 细胞的频率，详见材料与方法。如图 11 所示，r $\Delta$ BB-GFP-lcmvNP 和 r $\Delta$ BB-GFP-NP118 复制子都诱导出了高频率的 LCMV 特异性 T 细胞（70~200/10<sup>5</sup>脾细胞）。有趣的是，这些频率都略高于用质粒 DNA 进行遗传免疫后所测到的频率。

总而言之，这些发现表明由于门戈病毒复制子能以全长外源抗原或相应外源表位短相关肽的免疫原性形式进行表达，因此该复制子对于诱导源性特异性免疫应答具有多种用途。

以上对整个发明都已进行了详细的论述，经验丰富的业内人员会发现本发明可在一定等效范围内进行操作，而不破坏其整体性，也不需要进行任何过度实验。此外，尽管在本文中用了一些实施方案及实施例对本发明进行了描述，但本发明人相信本发明还能进行进一步的调整。本申请书意在涵盖所有根据上文中总体原则对本发明进行的变动、使用或适应性调整。

所有在本文中引用的参考文献、手册、专利及专利申请书记的内容都作为整体并入作为参考。

**参考文献:**

1. Ansardi, D. C., Z. Moldoveanu, D. C. Porter, D. E. Walker, R. M. Conry, A. F. LoBuglio, S. McPherson, and C. D. Morrow 1994. Characterization of poliovirus replicons encoding carcinoembryonic antigen. *Cancer Res.* 54:6359-64.
2. Bot, A., S. Bot, A. Garcia-Sastre, and C. Bona 1996. DNA immunization of newborn mice with a plasmid-expressing nucleoprotein of influenza virus. *Viral Immunol.* 9:207-10.
3. Choi, W. S., R. Pal-Ghosh, and C. D. Morrow 1991. Expression of human immunodeficiency virus type 1 (HIV-1) gag, pol, and env proteins from chimeric HIV-1-poliovirus minireplicons. *J. Virol.* 65:2875-83.
4. Cohen, L., K. M. Kean, M. Girard, and S. Van der Werf 1996. Effects of P2 cleavage site mutations on poliovirus polyprotein processing. *Virology.* 224:34-42.
5. Conry, R. M., A. F. LoBuglio, M. Wright, L. Sumerel, M. J. Pike, F. Johanning, R. Benjamin, D. Lu, and D. T. Curiel 1995. Characterization of a messenger RNA polynucleotide vaccine vector. *Cancer Res.* 55:1397-400.
6. Dalemans, W., A. Delers, C. Delmelle, F. Denamur, R. Meykens, C. Thiriart, S. Veenstra, M. Francotte, C. Bruck, and J. Cohen 1995. Protection against homologous influenza challenge by genetic immunization with SFV-RNA encoding Flu-HA. *Annals of the New York Academy of Sciences.* 772:255-6.

7. Donnelly, M. L., D. Gani, M. Flint, S. Monaghan, and M. D. Ryan 1997. The cleavage activities of aphthovirus and cardiovirus 2A proteins. *J. Gen. Virol.* 78:13-21.
8. Duke, G. M., and A. C. Palmenberg 1989. Cloning and synthesis of infectious cardiovirus RNA containing short, discrete Poly(C) tracts. *J. Virol.* 63:1822-1826.
9. Escriou, N., C. Leclerc, S. Gerbaud, M. Girard, and S. van der Werf 1995. Cytotoxic T cell response to Mengo virus in mice: effector cell phenotype and target proteins. *J. Gen. Virol.* 76:1999-2007.
10. Fleton, M. N., M. Chen, P. Berglund, G. Rhodes, S. E. Parker, M. Murphy, G. J. Atkins, and P. Liljeström 2001. Self-replicating RNA vaccines elicit protection against Influenza A Virus, Respiratory Syncytial Virus, and a Tickborne Encephalitis Virus. *J. Infect. Dis.* 183:1395-8.
11. Frolov, I., T. A. Hoffman, B. M. Pragai, S. A. Dryga, H. V. Huang, S. Schlesinger, and C. M. Rice 1996. Alphavirus-based expression vectors: strategies and applications. *Proc. Natl. Acad. Sci. USA.* 93:11371-7.
12. Kaplan, G., and V. R. Racaniello 1988. Construction and characterization of poliovirus subgenomic replicons. *J. Virol.* 62:1687-96.
13. Kieny, M. P., G. Rautmann, D. Schmitt, K. Dott, S. Wain-Hobson, M. Alizon, M. Girard, S. Chamaret, A. Laurent, L. Montagnier, and J. P. Lecocq 1986. AIDS virus env protein expressed from a recombinant vaccinia virus. *Biotechnology.* 4:790-795.

14. Kingsbury, D. W., I. M. Jones, and K. G. Murti 1987. Assembly of influenza ribonucleoprotein in vitro using recombinant nucleoprotein. *Virology*. 156:396-403.
15. Lobert, P. E., N. Escriou, J. Ruelle, and T. Michiels 1999. A coding RNA sequence acts as a replication signal in cardioviruses. *Proc. Natl. Acad. Sci. USA*. 96:11560-5.
16. Lu, H. H., L. Alexander, and E. Wimmer 1995. Construction and genetic analysis of dicistronic polioviruses containing open reading frames for epitopes of human immunodeficiency virus type 1 gp120. *J. Virol.* 69:4797-806.
17. Martinon, F., S. Krishnan, G. Lenzen, R. Magno, E. Gomard, J.-G. Guillet, J.-P. Levy, and P. Meulien 1993. Induction of virus-specific cytotoxic T lymphocytes in vivo by liposome-entrapped mRNA. *Eur. J. Immunol.* 23:1719-1722.
18. Moldoveanu, Z., D. C. Porter, A. Lu, S. McPherson, and C. D. Morrow 1995. Immune responses induced by administration of encapsidated poliovirus replicons which express HIV-1 gag and envelope proteins. *Vaccine*. 13:1013-22.
19. Nichols, W. W., B. J. Ledwith, S. V. Manam, and P. J. Troilo 1995. Potential DNA vaccine integration into host cell genome. *Annals of the New York Academy of Sciences*. 772:30-9.
20. Oukka, M., J. C. Manuguerra, N. Livaditis, S. Tourdot, N. Riche, I. Vergnon, P. Cordopatis, and K. Kosmatopoulos 1996. Protection against

lethal viral infection by vaccination with nonimmunodominant peptides. *J. Immunol.* 157:3039-45.

21. Percy, N., W. S. Barclay, M. Sullivan, and J. W. Almond 1992. A poliovirus replicon containing the chloramphenicol acetyltransferase gene can be used to study the replication and encapsidation of poliovirus RNA. *J. Virol.* 66:5040-6.

22. Pogulis, R. J., A. N. Vallejo, and L. R. Pease 1996. In vitro recombination and mutagenesis by overlap extension PCR. *Methods Mol Biol.* 57:167-76.

23. Porter, D. C., J. Wang, Z. Moldoveanu, S. McPherson, and C. D. Morrow 1997. Immunization of mice with poliovirus replicons expressing the C-fragment of tetanus toxin protects against lethal challenge with tetanus toxin. *Vaccine.* 15:257-64.

24. Pushko, P., M. Parker, G. V. Ludwig, N. L. Davis, R. E. Johnston, and J. F. Smith 1997. Replicon-helper systems from attenuated Venezuelan equine encephalitis virus: expression of heterologous genes in vitro and immunization against heterologous pathogens in vivo. *Virology.* 239:389-401.

25. Qiu, P., P. Ziegelhoffer, J. Sun, and N. S. Yang 1996. Gene gun delivery of mRNA in situ results in efficient transgene expression and genetic immunization. *Gene Therapy.* 3:262-8.

26. Sambrook, J., E. F. Fritsch, and T. Maniatis 1989. Molecular cloning: a laboratory manual, 2nd ed, vol. 1. Cold Spring Harbor Laboratory Press, New York.

27. Ulmer, J. B., J. J. Donnelly, S. E. Parker, G. H. Rhodes, P. L. Felgner, V. J. Dwarki, S. H. Gromkowski, R. R. Deck, C. M. DeWitt, A. Friedman, L. A. Hawe, K. R. Leander, D. Martinez, H. C. Perry, J. W. Shiver, D. L. Montgomery, and M. A. Liu 1993. Heterologous protection against influenza by injection of DNA encoding a viral protein. *Science*. 259:1745-1749.

28. Ulmer, J. B., T. M. Fu, R. R. Deck, A. Friedman, L. Guan, C. DeWitt, X. Liu, S. Wang, M. A. Liu, J. J. Donnelly, and M. J. Caulfield 1998. Protective CD4<sup>+</sup> and CD8<sup>+</sup> T cells against influenza virus induced by vaccination with nucleoprotein DNA. *J. Virol.* 72:5648-53.

29. Vignuzzi, M., S. Gerbaud, S. van der Werf, and N. Escriou 2002. Expression of a membrane-anchored glycoprotein, the Influenza hemagglutinin, by dicistronic replicons derived from the poliovirus genome. *J. Virol.* 76:5285-90.

30. Vignuzzi, M., S. Gerbaud, S. van der Werf, and N. Escriou 2001. Naked RNA immunization with replicons derived from the poliovirus and Semliki Forest virus genomes for the generation of a cytotoxic T cell (CTL) response against the Influenza A virus nucleoprotein. *J. Gen. Virol.* 82:1737-47.

31. Wolff, J. A., J. J. Ludtke, G. Acsadi, P. Williams, and A. Jani 1992. Long-term persistence of plasmid DNA and foreign gene expression in mouse muscle. *Human Molecular Genetics*. 1:363-9.

32. Yamanaka, K., A. Ishihama, and K. Nagata 1990. Reconstitution of influenza virus RNA-nucleoprotein complexes structurally resembling native viral ribonucleoprotein cores. *J. Biol. Chem.* 265:11151-5.
33. Ying, H., T. Z. Zaks, R. F. Wang, K. R. Irvine, U. S. Kammula, F. M. Marincola, W. W. Leitner, and N. P. Restifo 1999. Cancer therapy using a self-replicating RNA vaccine. *Nature Medicine.* 5:823-7.
34. Zhou, X., P. Berglund, G. Rhodes, S. E. Parker, M. Jondal, and P. Liljeström 1994. Self-replicating Semliki Forest virus RNA as recombinant vaccine. *Vaccine.* 12:1510-1514.
35. Kurth, R. 1995. Risk potential of the chromosomal insertion of foreign DNA. *Ann. N.Y. Acad. of Sci.* 772:140-51.
36. Manuguerra, J.C. and C. Hannoun 1999. Influenza and other Viral Respiratory Diseases, surveillance and laboratory diagnosis. Edited by the Institut Pasteur. ISBN 2-901320-28-7.
37. Hoerr, I., R. Obst, H.G. Rammensee, and G. Jung 2000. *In vivo* application of RNA leads to induction of specific cytotoxic T lymphocytes and antibodies. *Eur. J. Immunol.* 30:1-7.
38. Hinton. T.M. and B.S. Crabb 2001. The novel picornavirus Equine rhinitis B virus contains a strong type II internal ribosomal entry site which functions similarly to that of Encephalomyocarditis virus. *J. Gen. Virol.* 82:2257-69.

---

39. Yokoyama, M. J. Zhang, and J.L. Whitton 1995. DNA immunization confers protection against lethal lymphocytic choriomeningitis virus infection. *J. Virol.* 69:2684-88.

40. Rodriguez, F. L.L. An, S. Harkins, J. Zhang, M. Yokouama, G. Widera, J.T. Fuller, C. Kincaid, I.L. Campbell, and J.L. Whitton 1998. DNA immunization with minigenes: low frequency of memory cytotoxic T lymphocytes and inefficient antiviral protection are rectified by ubiquitination. *J. Virol.* 72:5174-81.

<110> 巴斯德研究院

<120> 用于生产异源性蛋白的、衍生自正链RNA病毒基因组的复制子

<130> SF226PCT102

<140>

<141>

<150> 美国临时申请 N?0/292,515

<151> 2001-05-23

<160> 28

<170> PatentIn Ver. 2.1

<210> 1

<211> 9

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述:引物

<400> 1

gagctcgag

9

<210> 2

<211> 13

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述:引物

<400> 2

tcgaggctag ctt

13

<210> 3

<211> 11

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述:引物

<400> 3

cgaagctagc c

11

<210> 4

<211> 31

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述:引物

<400> 4

gctgagctca tggtagcaa gggcgaggag c

31

---

<210> 5  
 <211> 31  
 <212> DNA  
 <213> 人工序列  
  
 <220>  
 <223> 人工序列描述:引物  
  
 <400> 5  
 gcagagctcc ttgtacagct cgtccatgcc g 31  
  
 <210> 6  
 <211> 20  
 <212> DNA  
 <213> 人工序列  
  
 <220>  
 <223> 人工序列描述:引物  
  
 <400> 6  
 tctccacagg tgtccactcc 20  
  
 <210> 7  
 <211> 29  
 <212> DNA  
 <213> 人工序列  
  
 <220>  
 <223> 人工序列描述:引物  
  
 <400> 7  
 cacatcctgg ggtccattcc ggtgcgaac 29  
  
 <210> 8  
 <211> 30  
 <212> DNA  
 <213> 人工序列  
  
 <220>  
 <223> 人工序列描述:引物  
  
 <400> 8  
 accggaatgg accccaggat gtgctctctg 30  
  
 <210> 9  
 <211> 20  
 <212> DNA  
 <213> 人工序列  
  
 <220>  
 <223> 人工序列描述:引物  
  
 <400> 9  
 gtcccatcga gtgcggctac 20  
  
 <210> 10  
 <211> 36  
 <212> DNA

<213> 人工序列	
<220>	
<223> 人工序列描述:引物	
<400> 10	
cggaattctc gagatggcgt ctcaaggcac caaacg	36
<210> 11	
<211> 37	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列描述:引物	
<400> 11	
gcgaattctc gagattgtcg tactcctctg cattgtc	37
<210> 12	
<211> 37	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列描述:引物	
<400> 12	
cggaattctc gagatgtcct tgtctaagga agttaag	37
<210> 13	
<211> 33	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列描述:引物	
<400> 13	
gcgaattctc gagtgtcaca acatttgggc etc	33
<210> 14	
<211> 54	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列描述:引物	
<400> 14	
tcgaagctag cgaaagaccc caagcttcag gtgtgtatat gggttaattg acac	54
<210> 15	
<211> 54	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列描述:质粒	

<400> 15  
 tcgagtgtca aattacccat atacacacct gaagcttggg gtctttcgct agct 54

<210> 16  
 <211> 71  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 人工序列描述:引物

<400> 16  
 tcgaggctag ccagctttga attttgacct tottaagctt gcgggagacg tcgagtccaa 60  
 ccctgggcc t 71

<210> 17  
 <211> 72  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 人工序列描述:引物

<400> 17  
 tcgaaggcc cagggttga ctcgacgtct cccgcaagct taagaaggtc aaaattcaac 60  
 agctggctag cc 72

<210> 18  
 <211> 8  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 人工序列描述:引物

<400> 18  
 cgagcatg 8

<210> 19  
 <211> 16  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 人工序列描述:引物

<400> 19  
 ctagcatgct cgagct 16

<210> 20  
 <211> 25  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 人工序列描述:引物

<400> 20  
 ctggatccaa aatgaaggca aacct 25

<210> 21  
 <211> 28  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 人工序列描述:引物

<400> 21  
 caggatccta gatgcatatt ctgcactg 28

<210> 22  
 <211> 25  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 人工序列描述:引物

<400> 22  
 gaaaggcaaa cctactggtc ctggt 25

<210> 23  
 <211> 37  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 人工序列描述:引物

<400> 23  
 cgtgcagtcg acaggatgca tattctgcac tgcaaag 37

<210> 24  
 <211> 9  
 <212> PRT  
 <213> 人工序列

<220>  
 <223> 人工序列描述:肽

<400> 24  
 Ala Ser Asn Glu Asn Met Glu Thr Met  
 1 5

<210> 25  
 <211> 9  
 <212> PRT  
 <213> 人工序列

<220>  
 <223> 人工序列描述:肽

<400> 25  
 Arg Pro Gln Ala Ser Gly Val Tyr Met  
 1 5

<210> 26  
 <211> 8017  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 人工序列描述: 质粒

<400> 26  
 tttgaaagcc gggggtggga gatccggatt gccgggtccgc tcgatatcgc gggccggggtc 60  
 cgtgactacc cactccccct ttcaacgtga aggctacgat agtgccaggg cgggtcctgc 120  
 cgaaagtgcc aacccaaac cacataaacc ccccccccc tcccccccc ctcacattac 180  
 tggccgaagc cgcttggaaat aaggccgggtg tgcgtctgtc tataatgttac ttctactaca 240  
 ttgtcgtctg tgacgatgta ggggcccggga acctggtcct gtcttcttga cgagtattcc 300  
 taggggtctt tcccctctcg acaaaggaat acaaggtctg ttgaatgtcg tgaaggaagc 360  
 agttcctctg gacgcttctt gaagacaagc aacgtctgta gcgaccctt gcaggcagcg 420  
 gaatccccca cctggtgaca ggtgcctctg cggccgaaag ccacgtgtgt aagacacacc 480  
 tgcaaaggcg gcacaacccc agtgccacgt tgtgcgttgg atagtgtgg aaagagtcaa 540  
 atggctctcc tcaagcgtat tcaacaaggg gctgaaggat gccagaagg taccctactg 600  
 gctgggatct gatctggggc ctcgggtcgc gtgctttaca cgcgttgagt cgaggttaaa 660  
 aaacgtctag gcccccgaa ccacggggac gtggtttcc tttgaaacc acgacaataa 720  
 tatggctaca accatggagc tcgagaatac agaggagatg gagaatttat cagaccgagt 780  
 gtctcaagac actgccggca acacggtcac aaacacccaa tcaaccgttg gtcgtcttgt 840  
 cggatacggg acagttcatg atggggaaca tccattcgaa acacattatg caggatactt 900  
 ttcagatctt ttgatccacg atgtcgagac caatcccggg cctttcacgt ttaaaccaag 960  
 acaacggccg gtttttcaga ctcaaggagc ggcagtgtca tcaatggctc aaaccctact 1020  
 gccgaacgac ttggccagca aagctatggg atcagcctt acggcttgc tcgatgcaa 1080  
 cgaggacgcc caaaaagcaa tgaagattat aaagacgta agttctctat cggatgcatg 1140  
 ggaaaatgta aaaggaacat tgaacaacc ggagttctgg aaacaactct taagcagatg 1200  
 tgtgcaactg attgccggga tgacgatagc agtgatgcat ccggaccctc tgacgctgct 1260  
 ttgcttggga gtcttgacag cagcagagat cacaagccag acaagcctgt gcgaagaaa 1320  
 agcagctaaa ttcaaaaaca tcttctactc tccccctc cgtttctctg tgatctcact 1380  
 tttccaacag cagtcccccc ttaaacaggt caatgatgtt ttctctctgg caaagaacct 1440  
 agactgggca gtgaagacag ttgaaaaagt ggttgattgg ttggaactt gggttgcaca 1500  
 agaagagaga gagcagaccc tggatcagct gctccagcga tccccgagc acgcaagag 1560  
 gatttcagac cttcgtaatg gaatggctgc ctatgttgaa tgcaaggaga gcttcgattt 1620  
 ctttgagaaa ctttacaatc aagcagttaa ggagaagaga actggaattg ctgccgtttg 1680  
 tgaaaagttc agacaaaaac atgacatgc cacggcacga tgtgaaccag ttgtgatcgt 1740  
 gttgcgcggt gatgctggtc agggaaagtc attgtcaagt caaatcattg cccaggctgt 1800  
 ttctaaaact atttttgggc gccagtcagt ctattctctt cctcctgatt cabagttctt 1860  
 tgatggctat gagaaccagt ttgccgcaat aatggatgat ttgggacaaa atcccgatgg 1920  
 ttcagatttt accaccttct gccagatggt gtccacgaca aacttactcc caaacatggc 1980  
 tagtctggag agaaaaggaa cccccttcac atctcagctc gtagtggcta cgacaaatct 2040  
 cccggagttt agacctgtta caattgccca ttatcctgct gttagcggcc gcattacttt 2100  
 cgactactcg gtgtctgcag gtccagtttg tcaaaagacc gaagctggtt gcaaagtgtt 2160  
 ggatgttgaa agagccttta ggccaacagg tgatgccctt ctccatggt tccaaaataa 2220  
 ttgectatct ttggaagagg ctggcctgca gttcagagat aataggtcca aggagatttt 2280  
 atctttggtt gatgtgatg agagagctgt gactagaata gagaggaaga agaagtcct 2340  
 cacagcggtg cagacccttg tggccaagg gcctgttgat gaagttagct ttactcgggt 2400  
 tgtccagcag ctcaaggcta gacaggaagc tacagatgag cagttggagg aactccagga 2460  
 agcctttgcc cgggttcagg agcggagttc agtgttctca gactggatga agatttccgc 2520  
 catgctttgt gccgccacc tagctctcac acaagtggtg aagatggcta aggctgtcaa 2580  
 acagatggtg agaccagact tgggtcgggt ccagctggt gagcaagaac agggctcctta 2640  
 taacgaaacc acccgtataa agcccaaac tcttcaattg ctagatgtcc agggtcctaaa 2700  
 tccgactatg gactttgaaa agtttgttgc taagttgtt acagcccca ttggtttgtg 2760  
 gtacccaca ggtgttagca ctcagatg ctaacttgtg aagggacgta ccctggcgggt 2820  
 gaatcggcac atggcagagt ctgactggac ctccattgta gtgcgtggtg ttagccacac 2880  
 ccgctcctca gtgaaaatta tcgccatagc caaagctggg aaggagactg atgtgtcgtt 2940  
 cattcgcctt tcatctggtc ccttgtttag agataaact agcaagtttg tgaaggccag 3000  
 tgacgtattg ccccatagct cttccccct tattgggatc atgaatgtgg acattccaat 3060  
 gatgtataca gggacatttc tgaaggctgg cgtctcggtg ccggttgaga cagggcagac 3120  
 tttcaaccac tgcattccact acaaagcaaa tacacggaaa ggctggtgtg ggtctgcaat 3180  
 cctggccgat cttggtggga gcaagaagt tctgggctc cattcagccg gctccatggg 3240  
 cgttgacgcc gcgtcgataa tttcacaaga aatgatcgt gcggtggtgc aggccttcga 3300  
 gccccagggt gcacttgagc gggtgccaga tggctccgcg atccatgtac cccgaaagac 3360

tgctttgocg	ccgactgttg	ccagacaggt	cttccaaccc	gcttttgccc	cagctgttct	3420
ttctaaat	gacccacgca	cggatgctga	tggtgacgaa	gtagcttttt	caaaacatac	3480
atccaatcag	gaaaccctcc	ccccagtgtt	tagaatgggt	gctagggat	atgcaaacag	3540
agtattcgca	ctgttgggca	gagacaatgg	aaggctgtca	gtcaagcaag	ccttggatgg	3600
acttgagggg	atggaccocct	tggacaagaa	cacttcccca	ggccttccat	atactacgct	3660
aggaatgcgt	agaacagatg	ttgtagattg	ggaaaccgcc	actcttatcc	cctttgcagc	3720
agagagacta	gaaaaaatga	ataacaaaga	cttttccgac	attgtctatc	agacattcct	3780
caagagcgag	cttagaccta	tagagaaggt	acaagccgcc	aagacacgga	ttgtggatgt	3840
tccaccattt	gagcactgca	ttctgggtag	acaactgtct	gggaagtctg	cttccaaatt	3900
ccagacccaa	ccgggtctgg	aattgggctc	tgcaattggg	tgtgaccag	acgtgcattg	3960
gacagccttt	ggtgtggcaa	tgcaaggctt	tgaagggtg	tatgatgtgg	attattccaa	4020
ttttgattct	accattcag	tagctatatt	taggttattg	gcagaggaat	tcttttctga	4080
agagaatggc	ttcgacccat	tggttaagga	ttatcttgag	tccttagcca	tttcaaaaca	4140
tgcgatgag	gaaaagcgct	atctcataac	cgggtgctt	ccgtctggtt	gtgcagcgac	4200
ctcaatgta	aatacaata	tgaataatat	tattattagg	gccggttgt	atcttacata	4260
taaaaat	gagttgatg	acgtgaaggt	cttgtcttat	ggtgatgatc	ttctagtggc	4320
aactaattac	caattgaact	ttgatagagt	gagaacaagc	ctggcaaaga	caggatataa	4380
gattacacc	gctaacaaaa	cttctacctt	tcccctggaa	tcaactcttg	aggatgtagt	4440
attcctgaag	agaaaattta	agaaagagg	ccctctatat	cgacctgtca	tgaatagaga	4500
ggcgttagaa	gcaatggtgt	catattatcg	tccagggact	ctatctgaga	aactcacttc	4560
aatcactatg	cttgccgtgc	attctggcaa	acaggagtac	gatcgactct	ttgccccgtt	4620
tcgcgaggtt	ggagtgatcg	taccaacttt	tgagagtgtg	gagtacagat	ggaggagcct	4680
gttctgta	tagcgcggtc	actggcacia	cgcgttacc	ggtaagccaa	cocgggtgac	4740
acggtcgtca	taccgcagac	agggttcttc	tactttgcaa	gataaactag	agtagtaaaa	4800
taaatagttt	taaaaaaaa	aaaaaaaaaa	aaaacgggat	cctctagagt	cgacctgcag	4860
gcatgcaagc	ttttgttccc	tttagtgagg	gttaattccg	agcttggcgt	aatcatggtc	4920
atagctgttt	cctgtgtgaa	attgttatcc	gctcacaatt	ccacacaaca	tacgagccgg	4980
aagcataaag	tgtaaagcct	ggggtgccta	atgagtgagc	taactcacat	taattgcggt	5040
gcgctcactg	cccgttttcc	agtcgggaaa	cctgtcgtgc	cagctgcatt	aatgaactcg	5100
ccaacgcgcg	gggagaggcg	gtttgcgtat	ttggcgctct	tccgcttcc	cgctactgca	5160
ctcgttgcgc	tcggtcgttc	ggctgcgcg	agcggatca	gctcactcaa	aggcggtaat	5220
acggttatcc	acagaatcag	gggataacgc	aggaaagaac	atgtgagcaa	aaggccagca	5280
aaaggccag	aaccgtaaaa	aggccgcggt	gctggcgttt	ttccataggc	tccgcccccc	5340
tgacgagcat	cacaaaaatc	gacgctcaag	tcagaggtgg	cgaaaccoga	caggactata	5400
aagataccag	gcgtttcccc	ctggaagctc	cctcgtgcgc	tctcctgttc	cgaccctgcc	5460
gcttaccgga	tacctgtccg	cctttctccc	ttcgggaagc	gtggcgcttt	ctcatagctc	5520
acgctgtagg	tatctcagtt	cgggtgtaggt	cgctcgtccc	aagctgggct	gtgtgacaga	5580
acccccgctt	catcccgacc	gctgcgcctt	atccggtaac	tatcgtcttg	agtccaacc	5640
ggtaagacac	gacttatcgc	cactggcagc	agccactggt	aacaggatta	gcagagcgag	5700
gtatgtaggc	ggtgctacag	agttcttgaa	gtggtggcct	aactacggct	acaactagaag	5760
gacagtattt	ggtatctgcg	ctctgctgaa	gccagttacc	ttcggaaaaa	gagttggtag	5820
ctcttgatcc	ggcaaacaaa	ccaccgctgg	tagcgggtgt	ttttttgttt	gcaagcagca	5880
gattacgcgc	agaaaaaaag	gatctcaaga	agatcctttg	atcttttcta	cggggtctga	5940
cgctcagtg	aacgaaaact	cacgttaagg	gattttggtc	atgagattat	caaaaaggat	6000
cttcacctag	atccttttaa	attaaaaatg	aagttttaa	tcaatctaaa	gtatatatga	6060
gtaaaactgg	tctgacagtt	accaatgctt	aatcagtgag	gcacctatct	cagcagctcg	6120
tctatttctg	tcatccatag	ttgcctgact	ccccgtcgtg	tagataacta	cgatacggga	6180
gggcttacca	tctggcccca	gtgctgcaat	gataccgcga	gacccacgct	caocggctcc	6240
agatttatca	gcaataaacc	agccagccgg	aaggcccgag	cgcagaagtg	gtcctgcaac	6300
ttatccgcc	tccatccagt	ctattaattg	ttgccgggaa	gctagagtaa	gtagtctgcc	6360
agttaaatagt	ttgcgcaacg	ttgttgccat	tgctacaggc	atcgtgggtg	cacgctcgtc	6420
gtttggtatg	gcttcattca	gctccgggtc	ccaacgatca	aggcaggtta	catgatcccc	6480
catgtgtg	aaaaaagcgg	ttagctcctt	cggctcctcg	atcgttgtca	gaagtaagtt	6540
ggccgagtg	ttatcactca	tggttatggc	agcactgcat	aattctctta	ctgtcatgcc	6600
atccgtaaga	tgcttttctg	tgactgggtga	gtactcaacc	aagtcattct	gagaatagtg	6660
tatgcggcga	ccgagttgct	cttgcccggc	gtcaatacgg	gataataccg	cgccacatag	6720
cagaacttta	aaagtgtca	tcattggaaa	acgttcttcg	gggcgaaaac	tctcaaggat	6780
cttaccgctg	ttgagatcca	gttcgatgta	accactcgt	gcacccaact	gatcttcagc	6840
atcttttact	ttcaccagcg	tttctgggtg	agcaaaaaca	ggaaggcaaa	atgccgcaaa	6900
aaaggaata	agggcgacac	ggaaatggtg	aatactcata	ctcttctttt	ttcaatata	6960
ttgaagcatt	attgaggggtt	attgtctcat	gacgggatc	atatttgaat	gtatttagaa	7020
aaataaacia	ataggggttc	cgcgcacatt	tccccgaaaa	gtgccacctg	acgtctaaga	7080
aaccattatt	atcatgacat	taacctataa	aaataggcgt	atcacgaggc	cctttcgtct	7140
cgcgctttc	ggtgatgacg	gtgaaaacct	ctgacacatg	cagctcccg	agacggtcac	7200
agcttgtctg	taagcggatg	ccgggagcag	acaagcccgt	cagggcgcgt	cagcgggtgt	7260

tggcgggtgt	cggggctggc	ttaactatgc	ggcatcagag	cagattgtac	tgagagtgca	7320
ccatatgagg	tgtgaaatac	cgcacagatg	cgtaaggaga	aaataccgca	tcaggaaatt	7380
gtaaactgta	atattttgtt	aaaattcgcg	ttaaattttt	gttaaatacag	ctcatttttt	7440
aaccaatagg	ccgaaatcgg	caaaatccct	tataaatcaa	aagaatagac	cgagataggg	7500
ttgagtgttg	ttccagtttg	gaacaagagt	ccactattaa	agaacgtgga	ctccaacgtc	7560
aaagggcgaa	aaaccgtcta	tcagggcgat	ggcccactac	gtgaaccatc	accctaatac	7620
agttttttgg	ggtcggagg	ccgtaaagca	ctaaatcggg	accctaagag	gagccccgca	7680
tttagagctt	gacggggaaa	gcccggcgaac	gtggcgagaa	aggaagggaa	gaaagcgaaa	7740
ggagcgggag	ctagggcgct	ggcaagtgtg	gcggtcacgc	tgcgagtaac	caccacaccc	7800
gcccgcgcta	atgcccgcct	acagggcgcg	tcgcccatt	cgccattcag	gctgagcaac	7860
tgttgggaag	ggcgatcggt	gcccggcctct	tcgctattac	gccagctggc	gaaaggggga	7920
tgtgctgcaa	ggcgattaag	ttgggtaacg	ccagggtttt	cccagtcacg	acgttgtaaa	7980
acgacggcca	gtgaattgta	atacgactca	ctatagg			8017

&lt;210&gt; 27

&lt;211&gt; 8092

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;223&gt; 人工序列描述:质粒

&lt;400&gt; 27

tttgaaagcc	gggggtggga	gatccggatt	gcccgtccgc	tcgatatcgc	gggccgggtc	60
cgtgactacc	cactccccct	ttcaacgtga	aggctacgat	agtgccaggg	cgggtcctgc	120
cgaaagtgcc	aaccctaaac	cacataacc	cccccccccc	cccccccccc	ctcacattac	180
tggccgaagc	cgcttgggat	aaggccgggtg	tgcgtctgtc	tatatgttac	ttctactaca	240
ttgtcgtctg	tgacgatgta	ggggcccggg	acctggtcct	gtcttcttga	cgagtattcc	300
taggggtcct	tcccctctcg	acaaaggaat	acaaggctctg	ttgaatgtcg	tgaaggaagc	360
agttcctctg	gacgcttctt	gaagacaagc	aacgtctgta	gcgacccttt	gcaggcagcg	420
gaatccccca	cctggtgaca	ggtgectctg	cgccgaaag	ccacgtgtgt	aagacacacc	480
tgcaaaggcg	gcacaacccc	agtgccacgt	tgtgcggttg	atagttgtgg	aaagagtcaa	540
atggctctcc	tcaagcgtat	tcaacaagg	gctgaaggat	gccagaagg	taccccactg	600
gctgggatct	gatctggggc	ctcgggtgcg	gtgctttaca	cgcggtgagt	cgagggtaaa	660
aaacgtctag	gccccccgaa	ccacggggac	gtggttttcc	tttgaaaacc	acgacaataa	720
tatggctaca	accatggagc	tcgagcatgc	tagccagctg	ttgaattttg	accttcttaa	780
gcttgcggga	gacgtcgagt	ccaaccctgg	gcccttcgag	aatacagagg	agatgggaaa	840
tttatcagac	cgagtgtctc	aagacactgc	cggcaacacg	gtcacaaca	cccaatcaac	900
cgttggctcg	cttgtcggat	acggaacagt	tcgatgagg	gaacatccat	tcgaaacaca	960
ttatgcagga	tacttttcag	atcttttgat	ccacgatgtc	gagaccaatc	ccgggccttt	1020
cacgtttaaa	ccaagacaac	ggccggtttt	tcagactcaa	ggagcggcag	tgtcatcaat	1080
ggctcaaacc	ctactgccga	acgacttgcc	cagcaaagct	atgggatcag	cctttacggc	1140
tttgcctcag	gccaacgagg	acgccccaaa	agcaatgaag	attataaaga	cgttaagttc	1200
tctatcggat	atgtaaaagg	atgtaaaagg	aacattgaac	aaccggaggt	tctggaaaca	1260
actcttaagc	agatgtgtgc	aactgattgc	cgggatgacg	atagcagtga	tgcatccgga	1320
ccccttgacg	ctgctttgct	tgggagtctt	gacagcagca	gagatcaca	gccagacaag	1380
cctgtgcgaa	gaaatagcag	ctaaattcaa	aacaatcttc	actactccc	cccctcgttt	1440
tcctgtgatc	tcacttttcc	aacagcagtc	cccccttaa	caggtcaatg	atgttttctc	1500
tctggcaaaag	aacctagact	gggcagtgaa	gacagttgaa	aaagtgggtg	attggtttgg	1560
aacttgggtt	gcacaagaag	agagagagca	gaccctggat	cagctgctcc	agcgattccc	1620
cgagcacgag	aagaggatth	cagaccttcg	taatggaatg	gctgcctatg	ttgaatgcaa	1680
ggagagcttc	gatttctttg	agaaacttta	caatcaagca	gttaaggaga	agagaactgg	1740
aattgctgcc	gtttgtgaaa	agttcagaca	aaaacatgac	catgccacgg	cacgatgtga	1800
accagttgtg	atcgtgttgc	gcggtgatgc	tggtcagggg	aagtcattgt	caagtcaaat	1860
cattgccag	gctgtttcta	aaactattht	tgggcgcccag	tcagctctatt	ctcttctctc	1920
tgattcagat	ttctttgatg	gctatgagaa	ccagtttgcc	gcaataatgg	atgatttggg	1980
acaaaatccc	gatggttcag	atthttaccac	cttctgccag	atggtgtcca	cgacaaaact	2040
actoccaaac	atggctagtc	tgagagaaaa	aggaaccccc	ttcacatctc	agctcgtagt	2100
ggctaccgaca	aatctcccgg	agtttagacc	tgttacaatt	gccattatc	ctgctgttga	2160
gcccgcgact	actttcgaact	actcgggtgc	tcagagttcca	gtttgttcaa	agaccgaagc	2220
tggttgcaaa	gtgttgatg	ttgaaagagc	ctttaggcca	acaggtgatg	cccctcttcc	2280
atgtttccaa	aataattgcc	tattcttggg	aaaggctggc	ctgcagttca	gagataatag	2340
gtccaaggag	atthttatctt	tggttgatgt	gatcgagaga	gctgtgacta	gaatagagag	2400
gaagaagaaa	gtcctcacag	cgggtcagac	ccttgtggcc	caagggcctg	ttgatgaagt	2460

tagcttttac	tcggttgtcc	agcagctcaa	ggctagacag	gaagctacag	atgagcagtt	2520
ggaggaactc	caggaagcct	ttgcccggt	tcaggagcgg	agttcagtg	tctcagactg	2580
gatgaagatt	tccgccatgc	tttgtgccgc	cacctagct	ctcacacaag	tggatgaagat	2640
ggctaaggct	gtcaaacaga	tggtagacc	agacttggg	cgggtccagc	tggatgagca	2700
agaacagggt	ccttataacg	aaaccaccg	tataaagccc	aaaactcttc	aattgctaga	2760
tgtccagggt	ccaaatccga	ctatggactt	tgaaaagt	gttgctaagt	ttgttacagc	2820
ccccattggt	tttgtgtacc	ccacaggtgt	tagcactcag	acatgcctac	ttgtgaagg	2880
acgtaccctg	gcggtgaatc	ggcacatggc	agagtctgac	tggacctcca	ttgtagtgcg	2940
tgggttagc	cacaccgct	cctcagtgaa	aattatcgcc	atagccaaag	ctgggaagga	3000
gactgatgtg	tcgttcattc	gcctttcatc	tggcccttg	tttagagata	atactagcaa	3060
gtttgtgaag	gccagtgcg	tattgcccc	tagctcttc	ccccttattg	ggatcatgaa	3120
tgtggacatt	ccaatgatgt	atacaggac	atctctgaag	gctggcgtct	cgggtccgg	3180
tgagacagg	cagactttca	accactgcat	ccactacaaa	gcaaatacac	ggaaaggctg	3240
gtgtgggtct	gcaatcctgg	ccgatcttgg	tgggagcaag	aagattctgg	gcttccattc	3300
agccggctcc	atgggcgttg	cagccgcgtc	gataatttca	caagaaatga	tcgatgcggt	3360
tgtgcaggcc	ttcagagccc	agggcgact	tgagcggctg	ccagatggtc	cgccatcca	3420
tgtaccocga	aagactgctt	tgcgcccgc	tgctgccaga	caggtcttcc	aaccgcttt	3480
tgccccagct	gttctttcta	aattgacc	acgcacggat	gctgatgttg	acgaagtagc	3540
ttttcaaaa	catacatcca	atcaggaac	cctccccca	gtgtttagaa	tggttgctag	3600
ggaatatg	aacagagt	tcgactgtt	ggcagagac	aatggaaggc	tgtagtcaa	3660
gcaagccttg	gatggacttg	aggggatgga	ccctatggac	aagaacactt	ccccaggcct	3720
tccatatact	acgctagaa	tgcgtagaac	agatgttga	gattgggaaa	ccgccactct	3780
tatccocctt	gcagcagaga	gactagaaa	aatgaataac	aaagactttt	ccgacattgt	3840
ctatcagaca	ttcctcaagg	acgagcttag	ccctatagag	aaggtacaag	ccgccaagc	3900
acggattgtg	gatgttccac	catttgagca	ctgcattctg	ggtagacaac	tgctcgggaa	3960
gttcgcttcc	aaattccaga	cccaaccggg	tctggaattg	ggctctgcaa	ttgggtgtga	4020
cccagacgtg	cattggacag	cctttgggtg	ggcaatgcaa	ggctttgaaa	gggtgtatga	4080
tgtggattat	tccaattttg	attctacca	ttcagtagct	atatttaggt	tattggcaga	4140
ggaattcttt	tctgaagaga	atggcttcca	cccattgggt	aaggattatc	ttgagtcctt	4200
agccatttca	aaacatgcgt	atgaggaaaa	gcgctatctc	ataaccgggtg	gtcttccgtc	4260
tggttgtgca	gcgacctcaa	tgtaataac	aataatgaat	aatattatta	ttagggccgg	4320
tttgtatctt	acatataaaa	attttgagtt	tgatgacgtg	aaggtcttgt	cttatgggta	4380
tgatcttcta	gtggcaacta	attaccaatt	gaactttgat	agagtgagaa	caagcctggc	4440
aaagacagga	tataagatta	caccgctaa	caaaacttct	acctttcccc	tggaatcaac	4500
tcttgaggat	gtagtattcc	tgaagagaaa	atttaagaaa	gagggccctc	tatatcgacc	4560
tgcatgaat	agagaggcgt	tagaagcaat	gttgcatat	tatcgtccag	ggactctatc	4620
tgagaaaactc	acttcaatca	ctatgcttgc	cgtgcattct	ggcaaacagg	agtacgatcg	4680
actctttg	ccgtttcg	aggtggagt	gatcgtacca	acttttgaga	gtgtggagta	4740
cagatggagg	ggtaatagc	ggtaatagc	cggtcactgg	cacaacgcgt	taccgggtaa	4800
gccaacggg	tgtacacggt	cgcataccg	cagacaggg	tcttctactt	tgcaagataa	4860
actagagtag	taaaataaat	agtttataaa	aaaaaaaaaa	aaaaaaaaaac	gggatcctct	4920
agagtcgacc	tgaggcatg	caagctttg	ttcccttag	tgagggttaa	ttccgagctt	4980
ggcgtaatca	tggatcatgc	tgtttcctgt	gtgaaattgt	tatccgctca	caattccaca	5040
caacatacga	gccggaagca	taaagtgtaa	agcctgggg	gcctaagtag	tgagctaact	5100
cacattaatt	gcgttgcgct	cactgccgc	tttccagctg	ggaaacctgt	cgtgccagct	5160
gcattaatga	atcgccaac	gcgcggggag	agcggtttg	cgtattgggc	gctcttccgc	5220
ttctcgtc	actgactcgc	tgcgctcgt	cgttcggctg	cggcgagcgg	tatcagctca	5280
ctcaaaggcg	gtaatcgg	tatccacaga	atcagggat	aacgcaggaa	agaacatgtg	5340
agcaaaaggc	cagcaaaagg	ccaggaaccg	taaaaaggcc	gcgttgctgg	cgtttttcca	5400
taggctccgc	ccccctgacg	agcatcacia	aaatcgacgc	tcaagtcaga	ggtggcgaaa	5460
cccagacagga	ctataaagat	accaggcgtt	tccccctgga	agctccctcg	tgcgctctcc	5520
tgttccgacc	ctgccgctta	ccggatacct	gtccgcctt	ctcccttcgg	gaagcgtggc	5580
gctttctcat	agctcacgct	gtaggtatct	cagttcgggtg	taggtcgttc	gctccaagct	5640
gggctgtgtg	cacgaacccc	ccgttcagcc	cgaccgctgc	gccttatccg	gtaactatcg	5700
tcttgagtc	aaccggtaa	gacacgactt	atcgccactg	gcagcagcca	ctggtaacag	5760
gattagcaga	gcgaggtatg	taggcgggtc	tacagagttc	ttgaagtgg	ggcctaacta	5820
cggctacact	agaaggacag	tatttggat	ctgcgctctg	ctgaagccag	ttacctcgg	5880
aaaaagagtt	ggtagctctt	gatccggcaa	acaaaccacc	gctggtagcg	gtggttttt	5940
tgtttgcaag	cagcagatta	cgcgagaaa	aaaaggatct	caagaagatc	ctttgatctt	6000
ttctacggg	tctgacgctc	agtggaacga	aaactcacgt	taagggattt	tggatcatgag	6060
attatcaaaa	aggatcttca	cctagatcct	tttaatttaa	aaatgaagtt	ttaaatcaat	6120
ctaaagtata	tatgagtaaa	cttggctgga	cagttaccaa	tgcttaatca	gtgagccacc	6180
tatctcagcg	atctgtctat	ttcgttcatc	catagttgcc	tgactccccg	tcgtgtagat	6240
aactacgata	cgggagggtc	taccatctgg	ccccagtgct	gcaatgatac	cgcgagacc	6300
acgctcaccg	gctccagatt	tatcagcaat	aaaccagcca	gccggaagg	ccgagcgcag	6360

```

aagtggctcct gcaactttat ccgcctccat ccagctctatt aattggtgcc ggggaagctag 6420
agtaagtagt tgcgcagtta atagtttgcg caacggtggt gccattgcta caggcatcgt 6480
ggtgtcacgc tcgctgtttg gtatggcttc attcagctcc ggttcccaac gatcaaggcg 6540
agttacatga tccccatgt tgtgcaaaaa agcggtagc tccttcggtc ctccgatcgt 6600
tgtcagaagt aagttggccg cagtgttatc actcatgggt atggcagcac tgcataattc 6660
tcttactgtc atgccatccg taagatgctt ttctgtgact ggtgagtact caaccaagtc 6720
attctgagaa tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtcaa tacgggataa 6780
taccgcgcca catagcagaa ctttaaaagt gctcatcatt ggaaaacggt cttcggggcg 6840
aaaactctca aggatcttac cgctgttgag atccagttcg atgtaacca ctcgtgcacc 6900
caactgatct tcagcatctt ttactttcac cagcgtttct ggggtgagcaa aaacaggaag 6960
gcaaaatgcc gcaaaaaagg gaataagggc gacacggaaa tgttgaatac tcatactctt 7020
cctttttcaa tattattgaa gcatttatca ggttatttgt ctcatgagcg gatacatatt 7080
tgaatgtatt tagaaaaata acaaataggt gtttccgctc acatttcccc gaaaagtgcc 7140
acctgacgtc taagaaacca ttattatcat gacattaacc tataaaaaata ggcgtatcac 7200
gaggcccttt cgtctcgcgc gtttcgggta tgacggtgaa aacctctgac acatgcagct 7260
cccggagacg gtcacagctt gtctgtaagc ggatgccggg agcagacaag cccgtcaggg 7320
cgctgcagcg ggtgttggcg ggtgtcgggg ctggcttaac tatgcggcat cagagcagat 7380
tgtactgaga gtgcaccata tgcggtgtga aataccgcac agatgcgtaa ggagaaaata 7440
ccgcatcagg aaattgtaaa cgtaaatatt ttgttaaaat tcgcgttaaa tttttgttaa 7500
atcagctcat tttttaacca ataggccgaa atcggcaaaa tcccttataa atcaaaagaa 7560
tagaccgaga tagggtttag tgttgttcca gtttggaaaca agagtccact attaaagaac 7620
ctggactcca acgtcaaagg gcgaaaaacc gtctatcagg gcgatggccc actacgtgaa 7680
ccatcacctt aatcaagttt ttggggctcg aggtgccgta aagcactaaa tcggaacctt 7740
aaagggagcc cccgatttag agcttgacgg ggaaaagccgg cgaacgtggc gagaaaggaa 7800
gggaagaaag cgaaaggagc gggcgctagg gcgctggcaa gtgtagcggc cacgctgcgc 7860
gtaaccacca cccccgccgc gcttaatgcg ccgctacagg gcgctcgcgc ccattcgcca 7920
ttcaggctgc gcaactgttg ggaagggcga tcggtgcggg cctcttcgct attacgccag 7980
ctggcgaaag ggggatgtgc tgcaaggcga ttaagttggg taacgccagg gttttccag 8040
tcacgacggt gtaaaacgac ggccagttaa ttgtaatacg actcactata gg 8092

```

<210> 28

<211> 10417

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述:质粒

<400> 28

```

tttgaaagcc ggggggtgga gatccggatt gccggtccgc tcgatatcgc gggccggggtc 60
cgtgactacc cactccccct ttcaacgtga aggctacgat agtgccaggg cgggtcctgc 120
cgaaagtgcc aacccaaaac cacataaccc ccccccccc tcccccccc ctcacattac 180
tggccgaagc cgcttggaaat aaggccggtg tgcgtctgtc tataatgttac ttctactaca 240
ttgtcgtctg tgacgatgta ggggcccgga acctggctct gtcttcttga cgagtattcc 300
taggggtctt tcccctctcg acaaaggaaat acaaggtctg ttgaaatgctg tgaaggaaagc 360
agttcctctg gacgcttctt gaagacaagc aacgtctgta gcgacccttt gcaggcagcg 420
gaatccccca cctggtgaca ggtgcctctg cgccgaaag ccacgtgtgt aagacacacc 480
tgcaaaggcg gcacaacccc agtgccacgt tgtgcgttgg atagttgtgg aaagagtcaa 540
atggctctcc tcaagcgtat tcaacaaggg gctgaaggat gccagaagg taccctactg 600
gctgggatct gatctggggc ctcggtgcgc gtgctttaca cgcgttgagt cgaggttaaa 660
aaacgtctag gcccccgaa ccacggggac gtggttttcc tttgaaaacc acgacaataa 720
tatggctaca accatggagc tcatggtgag caagggcgag gagctgttca ccgggtggt 780
gcccacctcg gtcgagctgg acggcgacgt aaacggccac aagttcagcg tgtccggcga 840
gggaggggac gatgccacct acggcaagct gacctgaag ttcattctgca ccaccggcaa 900
gctgcccgtg ccctggccca ccctcgtgac caccctgacc tacggcgtgc agtgctttag 960
ccgctacccc gaccacatga agcagcacga cttcttcaag tccgccatgc ccgaaggcta 1020
cgtccaggag cgcacatct tcttcaagga cgacggcaac tacaagacc gcgcccagg 1080
gaagttcgag ggcgacaccc tgggtgaacc catcgagctg aagggcatcg acttcaagga 1140
ggacggcaac atcctggggc acaagctgga gtacaactac aacagccaca acgtctatat 1200
gatggccgac aagcagaaga acggcatcaa ggtgaacttc aagatccgcc acaacatcga 1260
cgcagctcagc gtgcagctcg ccgaccacta ccagcagaac accccatcg gcgacggccc 1320
cgtgtgctg cccgacaacc actacctgag caccagtcg gccctgagca aagaccccaa 1380
cgagaagcgc gatcacatgg tctgtctgga gttcgtgacc gccgcccggg tcaactctcg 1440
catggacgag ctgtacaagg agctcgagat gtccttgtct aaggaagtta agagcttcca 1500

```

atggacgcaa	gcattgagaa	gagaattgca	gagcttcaca	tcagatgtga	aggctgctgt	1560
cattaagga	gcaaccaacc	ttctgaatgg	gttgacttc	tctgaggtca	gcaatgctca	1620
gaggatcatg	aggaaggaaa	agagagatga	caaagaccta	cagagactca	gaagtctcaa	1680
ccagactgta	catttctctg	tggattttaa	gtcaacatca	aagaagaatg	ttttgaaagt	1740
ggggaggctc	agtgcagaag	aactgatgtc	tcttgcggt	gaccttgaga	agctgaaggc	1800
caagatcatg	aggtctgaaa	ggccccaggc	ttcaggggta	tatatgggga	acttaacaac	1860
acagcaacta	gaccaaagat	ctcagatcct	acagatagtt	gggatgagaa	agcctcagca	1920
gggtgcaagt	ggtgtggtaa	gagtttggga	tgtgaaagac	tcatcacttt	tgaacaatca	1980
atttggcaca	atgccaagtc	taactatggc	ttgtatggcc	aaacagtcac	agactccgct	2040
caatgacggt	gtacaagcgc	tcacagacct	tggcttgctt	tacacagtca	agtatccaaa	2100
tcttaatgat	cttgaaaggc	tgaagacaa	gcaccagtt	ctgggggtca	tactgaaca	2160
gcagtcacag	atcaacattt	ctggctataa	ctttagtctt	ggtgctgccg	tgaaggcagg	2220
ggcagccctg	ttggatgggg	gtaacatggt	agagtcaatt	ttgatcaagc	caagcaacag	2280
cgaggacctc	ttgaaaggcag	ttctcggggc	caagagaaaa	ctcaacatgt	ttgtttcaga	2340
ccaagtggg	gacaggaacc	cttatgaaaa	catcctctat	aaagtttgcc	tttcaggtga	2400
aggatggcca	tacatagctt	gtagaacatc	gattgtgggg	agagcatggg	aaaacacaac	2460
aattgactct	acaagcgaga	aacctgcagt	caactcacc	aggccagcgc	ctggagcagc	2520
aggtccacct	caggtgggct	taagctacag	ccagacaatg	cttttaaaag	acctcatggg	2580
aggaattgac	cccaacgctc	ctacatggat	tgacattgag	ggtagattta	atgatccagt	2640
ggaaatagca	attttccaac	cacagaacgg	gcagttcata	cacttttaca	gggaaccgct	2700
tgatcaaaaa	caattcaagc	aagattccaa	gtactcacac	ggcatggatc	ttgccgacct	2760
cttcaatgcy	caaccgggt	tgacctcgtc	agttataggt	gctcttccgc	aggggatggt	2820
tctaagctgt	caaggctccg	atgacatcag	aaagcttctg	gactcacaaa	ataggaagga	2880
cattaagctt	atcgatgttg	aatgaccag	ggaagcttcg	agggagtatg	aagacaaagt	2940
gtgggacaaa	tatggctggg	tgtgtaagat	gcatactgga	atagtaaggg	acaaaaagaa	3000
gaaagagatc	accccgact	gtgcactcat	ggactgcatc	atttttgaaa	gcgctccaa	3060
agcaaggctc	ccagatctga	aaactgttca	caacattctg	ccacatgacc	taatttttag	3120
aggcccaaat	gttgtgacac	tcgagaatac	agaggagatg	gagaatttat	cagaccgagt	3180
gtctcaagac	actgccggca	acacggtcac	aaacacccaa	tcaaccgttg	gtcgtcttgt	3240
cggatacggg	acagttcatg	atggggaaca	tccattcgaa	acacattatg	caggatactt	3300
ttcagatcct	ttgatccacg	atgtcgagac	caatcccggg	cctttcacgt	ttaaaccaag	3360
acaacggccg	gtttttcaga	ctcaaggagc	ggcaggtca	tcaatggctc	aaacctact	3420
gccgaacgac	ttggccagca	aagctatggg	atcagccttt	acggctttgc	tcgatgocaa	3480
cgaggacgcc	caaaaagcaa	tgaagattat	aaagacgtta	agttctctat	cggatgcatg	3540
ggaaaatgta	aaaggaacat	tgaacaaccc	ggagttctgg	aaacaactct	taagcagatg	3600
tgtgcaactg	attgccggga	tgacgatagc	agtgatgcat	ccggaccctt	tgacgctgct	3660
ttgcttggga	gtcttgacag	cagcagagat	cacaagccag	acaagcctgt	gccaagaaat	3720
agcagctaaa	ttcaaaaaca	tcttactac	tccccccct	cgttttctctg	tgatctcact	3780
ttccaacag	cagtcccccc	ttaaaccaggt	caatgatggt	ttctctctgg	caaagaacct	3840
agactgggca	gtgaagacag	ttgaaaaagt	ggttgattgg	tttggaactt	gggttgaca	3900
agaagagaga	gagcagaccc	tggatcagct	gctccagcga	ttccccgagc	acgcgaagag	3960
gatttcagac	cttcgtaatg	gaatggctgc	ctatgttgaa	tgcaaggaga	gcttcgattt	4020
ctttgagaaa	ctttacaatc	aagcagttaa	ggagaagaga	actggaattg	ctgccgtttg	4080
tgaaaagttc	agacaaaaac	atgaccatgc	cacggcacga	tgtgaaccag	ttgtgatcgt	4140
gttgcgcggt	gatgctggtc	agggaaagtc	attgtcaagt	caaatcattg	cccaggctgt	4200
ttctaaaact	atttttgggc	gccagtcagt	ctattctctt	cctcctgatt	cagatttctt	4260
tgatggctat	gagaaccagt	ttgccgcaat	aatggatgat	ttgggacaaa	atcccgatgg	4320
ttcagatttt	accaccttct	gccagatggg	gtcccagaca	aaactactcc	caaacatggc	4380
tagtctggag	agaaaaggaa	cccccttcac	atctcagctc	gtagtggcta	cgacaaatct	4440
cccggagttt	agacctgtta	caattgcca	ttatcctgct	gttgagcgcc	gcattacttt	4500
cgactactcg	gtgtctgcag	gtccagtttg	ttcaaagacc	gaagctgggt	gcaaagtgtt	4560
ggatgttgaa	agagccttta	ggccaacagg	tgatgccctt	cttccatggt	tccaaaataa	4620
ttgcctattc	ttggaaaagg	ctggcctgca	gttcagagat	aataggtcca	aggagatttt	4680
atctttggtt	gatgtgatcg	agagagctgt	gactagaata	gagaggaaga	agaaagtcct	4740
cacagcgggt	cgagaccttg	tggcccaagg	gcctgttgat	gaagttagct	tttactcggg	4800
tgtccagcag	ctcaaggcta	gacaggaagc	tacagatgag	cagttggagg	aactccagga	4860
agcctttgcc	cgggttcagg	agcggagttc	agtgttctca	gactggatga	agatttccgc	4920
catgctttgt	gccgccacc	tagctctcac	acaagtgggt	aagatggcta	aggctgtcaa	4980
acagatggtg	agaccagact	tggtgctggg	ccagctggat	gagcaagaac	agggctcctt	5040
taacgaaacc	accctgataa	agcccaaac	tcttcaattg	ctagatgtcc	agggccaaa	5100
tccgactatg	gactttgaaa	agtttgttgc	taagttgtt	acagcccca	ttggttttgt	5160
gtacccaca	ggtgttagca	ctcagacatg	cctacttggt	aagggcagta	ccctggcggg	5220
gaatcggcac	atgctcagag	ctgactggac	ctccattgta	gtgctggtg	ttagccacac	5280
ccgctcctca	gtgaaaatta	tcgccatagc	caaagctggg	aaggagactg	atgtgtcgtt	5340
cattcgcctt	tcatctggtc	ccttgttttag	agataatact	agcaagtttg	tgaaggccag	5400

tgacgtattg	ccccatagct	cttccccctt	tattgggac	atgaatgtgg	acattccaat	5460
gatgtataca	gggacatttc	tgaaggctgg	cgtctcgggtg	ccggttgaga	cagggcagac	5520
tttcaaccac	tgcatccact	acaaagcaaa	tacaocggaaa	ggctgggtg	ggtctgcaat	5580
cctggccgat	cttgggtggga	gcaagaagat	tctgggcttc	cattcagccg	gctccatggg	5640
cgttgcagcc	gcgtcgataa	tttcacaaga	aatgatcgat	gcgggtgggtc	aggccttcga	5700
gccccaggg	gcacttgagc	ggctgccaga	tggtccgcgc	atccatgtac	cccgaagac	5760
tgctttgcgc	ccgactgttg	ccagacaggt	cttccaacce	gcttttgccc	cagctgttct	5820
ttctaaat	gacccacgca	cggatgctga	tggtgacgaa	gtagcttttt	caaaacatac	5880
atccaatcag	gaaaccctcc	ccccagtggt	tagaatgggt	gctagggat	atgcgaacag	5940
agtattcgca	ctggtgggca	gagacaatgg	aaggctgtca	gtcaagcaag	ccttggatgg	6000
acttgagggg	atggacccta	tgacaagaa	cacttcccca	ggccttccat	atactacgct	6060
aggaatgcgt	agaacagatg	ttgtagattg	ggaaaccgcc	actcttatcc	cctttgcagc	6120
agagagacta	gaaaaaatga	ataacaaaga	cttttccgac	attgtctatc	agacattcct	6180
caaggacgag	cttagaccta	tagagaaggt	acaagccgcc	aagacacgga	ttgtggatgt	6240
tccaccat	gagcactgca	ttctgggtag	acaactgctc	gggaagtccg	cttccaaatt	6300
ccagaccct	ccgggtctgg	aattgggctc	tgcaatgggt	tgtagcccg	acgtgcatg	6360
gacagcct	ggtgtggcaa	tgcaaggctt	tgaaaggggtg	tatgatgtgg	attattccaa	6420
ttttgattct	accattcag	tagctatatt	taggttattg	gcagaggaat	tcttttctga	6480
agagaatggc	ttcgacccat	tggttaagga	ttatcttgag	tccttagcca	tttcaaaaca	6540
tgcgatagag	gaaaagcgct	atctcataac	cggtggctct	ccgtctgggt	gtgcagcgac	6600
ctcaatgta	aatacaataa	tgaataatat	tattattagg	gccggttgt	atcttacata	6660
taaaaaat	gagttgatg	acgtgaaggt	cttgtcttat	ggtgatgatc	ttctagtggc	6720
aactaattac	caattgaact	ttgatagagt	gagacaagc	ctggcaaaga	caggatataa	6780
gattacacc	gtaacaaaa	cttctacctt	tcccctggaa	tcaactcttg	aggatgtag	6840
attcctgaag	agaaaattta	agaaagagg	ccctctatat	cgacctgtca	tgaatagaga	6900
ggcgtagaa	gcaatgttgt	catattatcg	tccagggact	ctatctgaga	aactcacttc	6960
aatcactatg	cttgccgtgc	attctggcaa	acaggagtac	gatcgactct	ttgccccgtt	7020
tcgagaggt	ggagtgatcg	taccaacttt	tgagagtgtg	gagtacagat	ggaggagcct	7080
gttctggtaa	tagcgcggtc	actggcacia	cgcgttacc	ggtaagccaa	ccgggtgtac	7140
acggctcgta	taccgcagac	agggttcttc	tactttgcaa	gataaactag	agtagtaaaa	7200
taaatagttt	taaaaaaaaa	aaaaaaaaaa	aaaacgggat	cctctagagt	cgacctgacg	7260
gcatgcaagc	ttttgtccc	tttagtgagg	gttaattccg	agcttggcgt	aatcatggtc	7320
atagctgttt	cctgtgtgaa	attgttatcc	gctcacaatt	ccacacaaca	tacgagccgg	7380
aagcataaag	tgtaaagcct	gggtgccta	atgagtgagc	taactacat	taattgcgtt	7440
gcgctcactg	cccgtttcc	agtcgggaaa	cctgtcgtgc	cagctgcatt	aatgaatcgg	7500
ccaacgcgcg	gggagagcg	gtttgcgtat	tgggcgctct	tccgcttcc	cgctcactga	7560
ctcgtcgcgc	tcggtcgttc	ggctgcggcg	agcggtatca	gctcactcaa	aggcggtaat	7620
acggttatcc	acagaatcag	gggataacgc	aggaaagaac	atgtgagcaa	aaggccacca	7680
aaaggcagc	aggcgtaaaa	aggccgcggt	tggtgcggtt	ttccataggc	tccgcccccc	7740
tgacgagcat	cacaaaaatc	gacgctcaag	tcagaggtgg	cgaaaccgca	caggactata	7800
aagataccag	gcgtttcccc	ctggaagctc	cctcgtgcgc	tctcctgttc	cgacctgcc	7860
gcttaccgga	tacctgtccg	cctttctccc	ttcgggaagc	gtggcgcttt	ctcatagctc	7920
acgctgtagg	tatctcagtt	cggtgtaggt	cgttcgcctc	aagctgggct	gtgtgcacga	7980
accccccg	cagcccgacc	gctgcgcctt	atccggtaac	tatcgtcttg	agtccaacc	8040
ggtaagacac	gacttatcgc	cactggcagc	agccactggt	aacaggatta	gcagagcgag	8100
gtatgtaggg	ggtgctacag	agttcttgaa	gtgggtggcct	aactacggct	acactagaag	8160
gacagtattt	ggtatctcgc	ctctgctgaa	gccagttacc	ttcggaaaaa	gagttggtag	8220
ctcttgatcc	ggcaaacaaa	ccaccgctgg	tagcgggtgt	ttttttgttt	gcaagcagca	8280
gattacgcgc	agaaaaaaag	gatctcaaga	agatcctttg	atcttttcta	cggggtctga	8340
cgctcagtg	aacgaaaact	cacgttaagg	gattttggtc	atgagattat	caaaaaggat	8400
cttcacctag	atccttttaa	attaaaaatg	aagttttaaa	tcaatctaaa	gtatatatga	8460
gtaaacttgg	tctgacagtt	accaatgctt	aatcagtgag	gcacctatct	cagcgatctg	8520
tctatttcgt	tcatccatag	ttgcctgact	ccccgtcgtg	tagataacta	cgatacggga	8580
gggcttacca	tctggcccca	gtgctgcaat	gataccgcga	gacctacgct	caccggctcc	8640
agatttatca	gcaataaac	agccagccgg	aagggccgag	cgcagaagtg	gtcctgcaac	8700
tttatccgcc	tccatccagt	ctattaattg	ttgcccggaa	gctagagtaa	gtagttcgcc	8760
agttaatagt	ttgcgcaacg	ttgttgccat	tgctacaggc	atcgtgggtg	cacgctcgtc	8820
gtttggtag	gcttcattca	gctccggttc	ccaacgatca	aggcgagtta	catgatcccc	8880
catgttgtgc	aaaaaagcgg	ttagctcctt	cggtcctccg	atcgttgtca	gaagtaagtt	8940
ggccgcagtg	ttatcactca	tggttatggc	agcactgcat	aattctctta	ctgtcatgcc	9000
atccgtaaga	tgettttctg	tgactgggtga	gtactcaacc	aagtcttct	gagaatagtg	9060
tatcggcgca	ccgagttgct	cttggccggc	gtcaatacgg	gataataccg	cgccacatag	9120
cagaacttta	aaagtgctca	tcaattggaaa	acgttcttcg	ggcgaaaaac	tctcaaggat	9180
cttaccgctg	ttgagatcca	gttcgatgta	accactcgt	gcacccaact	gatcttcagc	9240
atcttttact	ttcaccagcg	tttctgggtg	agcaaaaaa	ggaaggcaaa	atgccgcaaa	9300

---

```

aaaggaata agggcgacac ggaaatgttg aatactcata ctcttccttt ttcaatatta 9360
ttgaagcatt tadcagggtt attgtctcat gagcggatac atatttgaat gtatttagaa 9420
aaataaacia ataggggttc cgcgcacatt tccccgaaaa gtgccacctg acgtctaaga 9480
aaccattatt atcatgacat taacctataa aaataggcgt atcacgaggc cctttcgtct 9540
cgcgcgtttc ggtgatgacg gtgaaaacct ctgacacatg cagctcccgg agacggtcac 9600
agcttgtctg taagcggatg cggggagcag acaagcccgt cagggcgcgt cagcgggtgt 9660
tggcgggtgt cggggctggc ttaactatgc ggcatcagag cagattgtac tgagagtgca 9720
ccatatgagg tgtgaaatac cgcacagatg cgtaaggaga aaataccgca tcaggaaatt 9780
gtaaaccgta atattttggt aaaattcgcg ttaaattttt gttaaatcag ctcatttttt 9840
aaccaatagg ccgaaatcgg caaaatccct tataaatcaa aagaatagac cgagataggg 9900
ttgagtgttg ttccagtttg gaacaagagt ccactattaa agaacgtgga ctccaacgtc 9960
aaagggcgaa aaaccgtcta tcagggcgat ggcccactac gtgaaccatc accctaatca 10020
agttttttgg ggtcagagggt ccgtaaagca ctaaatcgga accctaaagg gagccccga 10080
tttagagctt gacggggaaa gccggcgaac gtggcgagaa aggaagggaa gaaagcgaaa 10140
ggagcgggag ctagggcgct ggcaagtgtg gcggtcacgc tgcgcgtaac caccacacc 10200
gccgcgctta atgcgcgctt acagggcgcg tcgcgccatt cgccattcag gctgcgcaac 10260
tgttgggaag ggcgatcggg gcgggcctct tcgctattac gccagctggc gaaaggggga 10320
tgtgctgcaa ggcgattaag ttgggtaacg ccagggtttt cccagtcacg acgttgtaaa 10380
acgacggcca gtgaattgta atacgactca ctatagg 10417

```

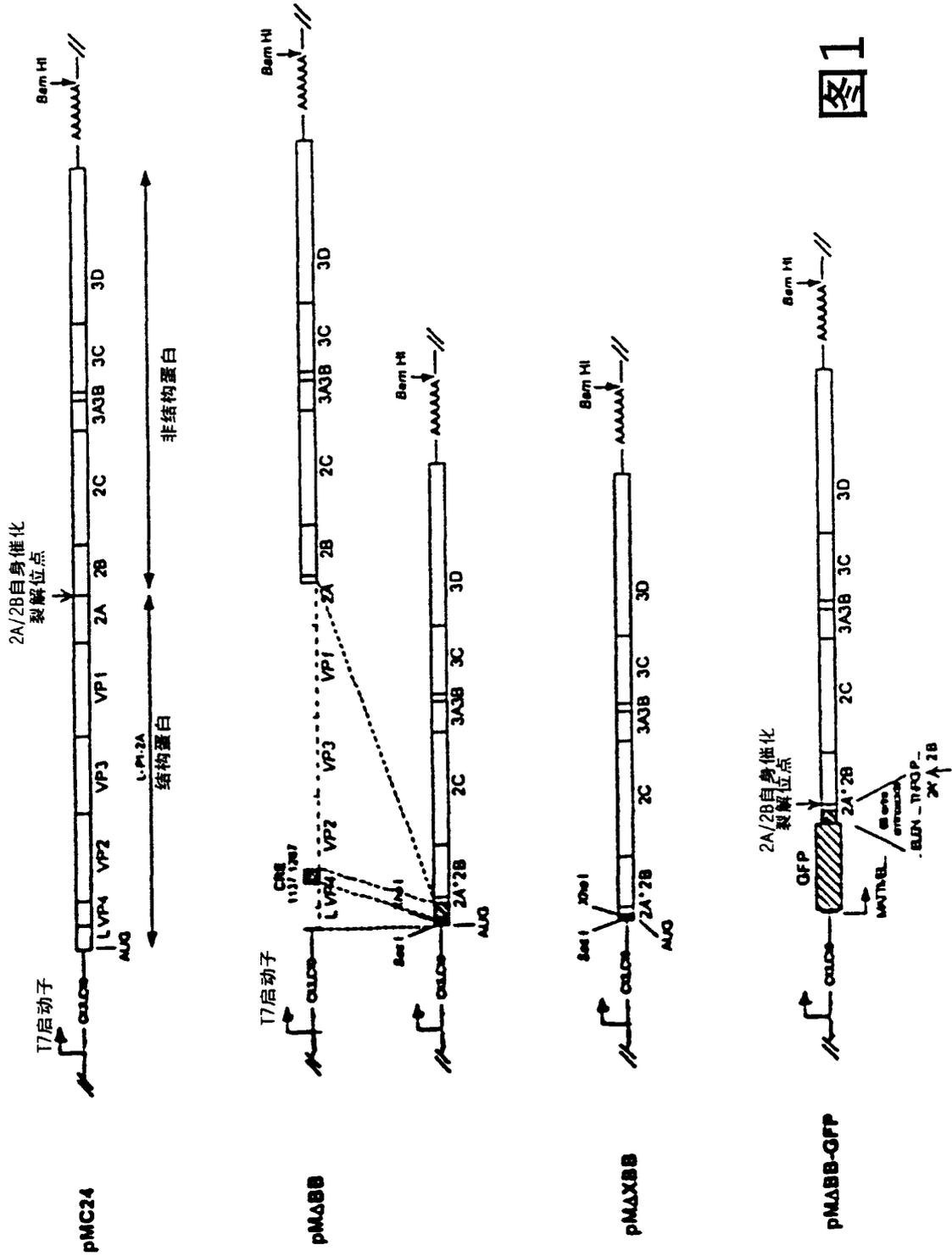


图1

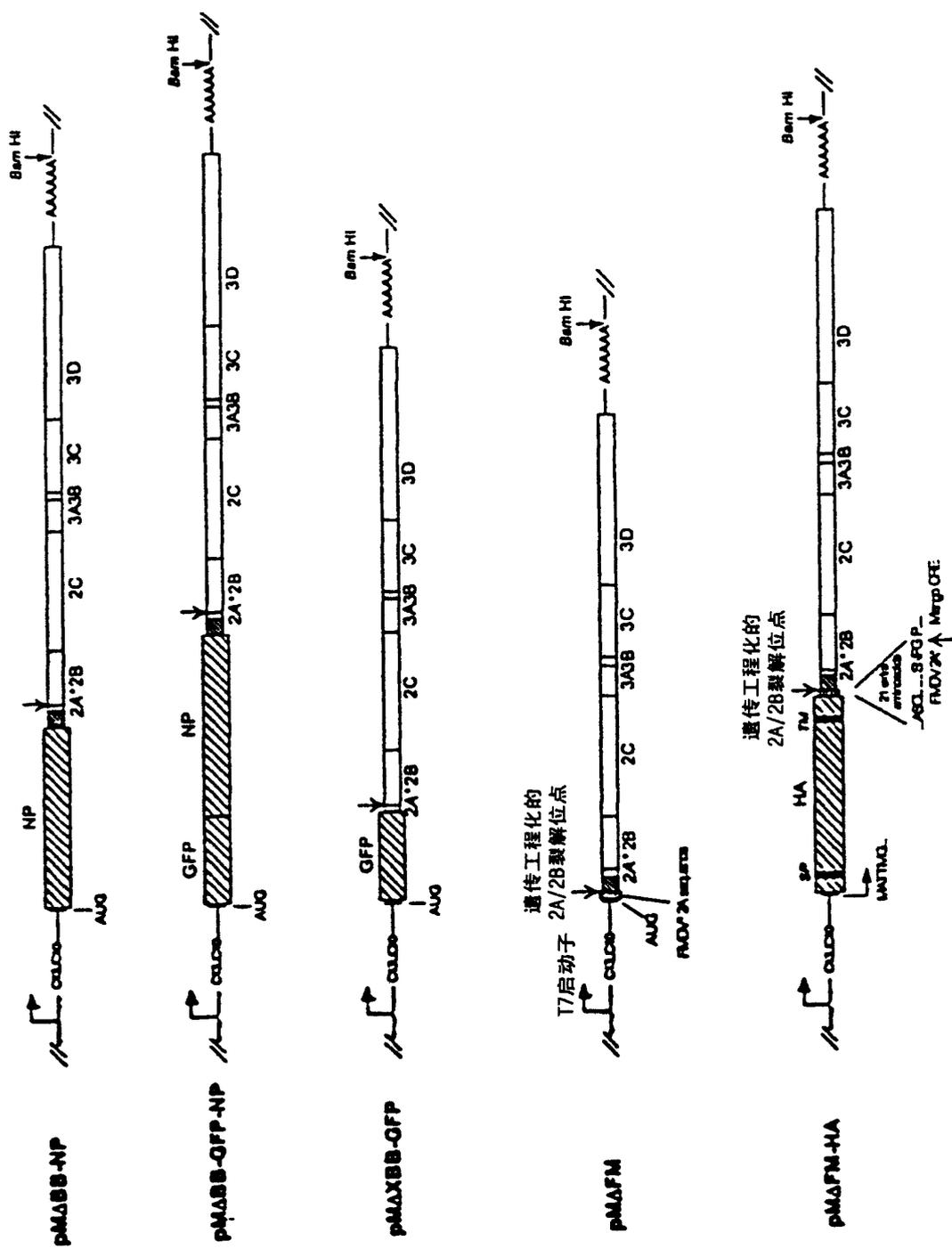


图1(续)

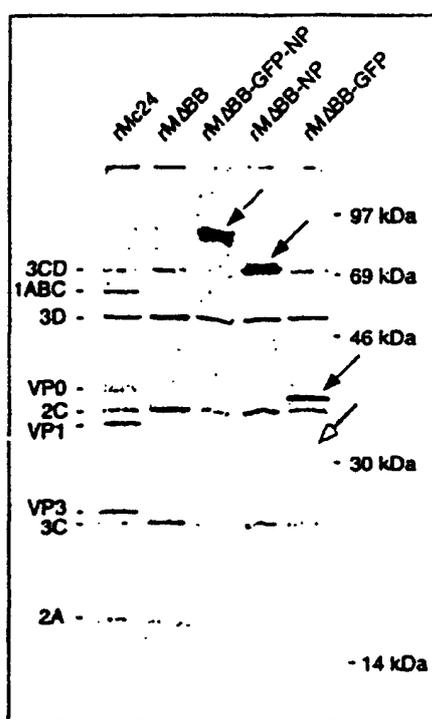


图2

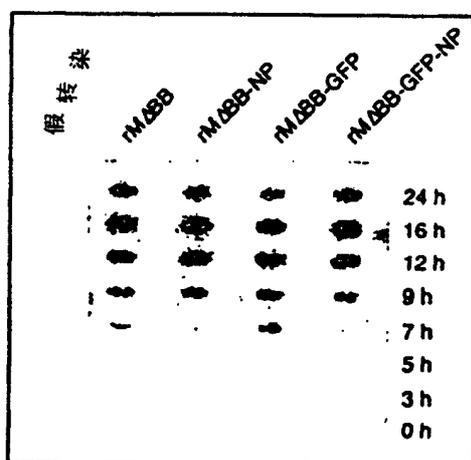


图3

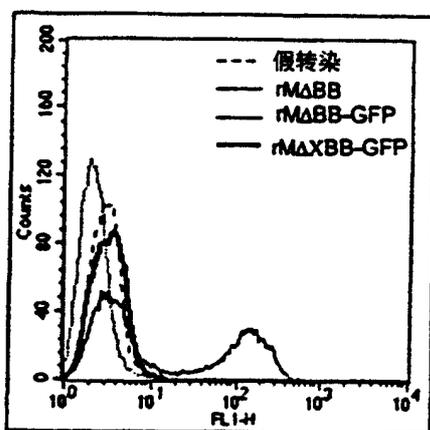


图4

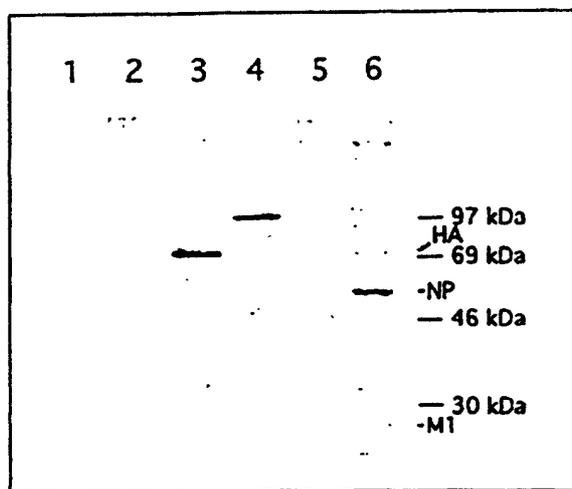


图5

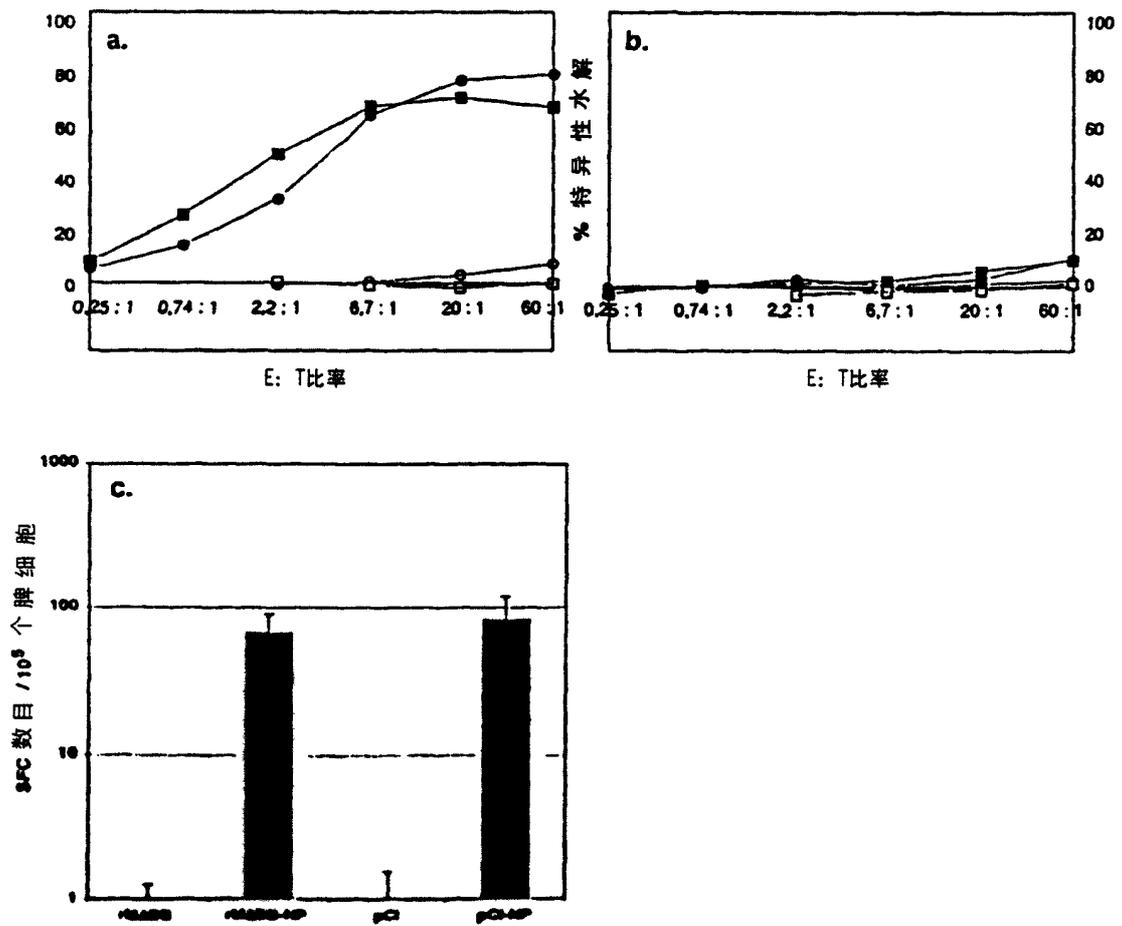


图6

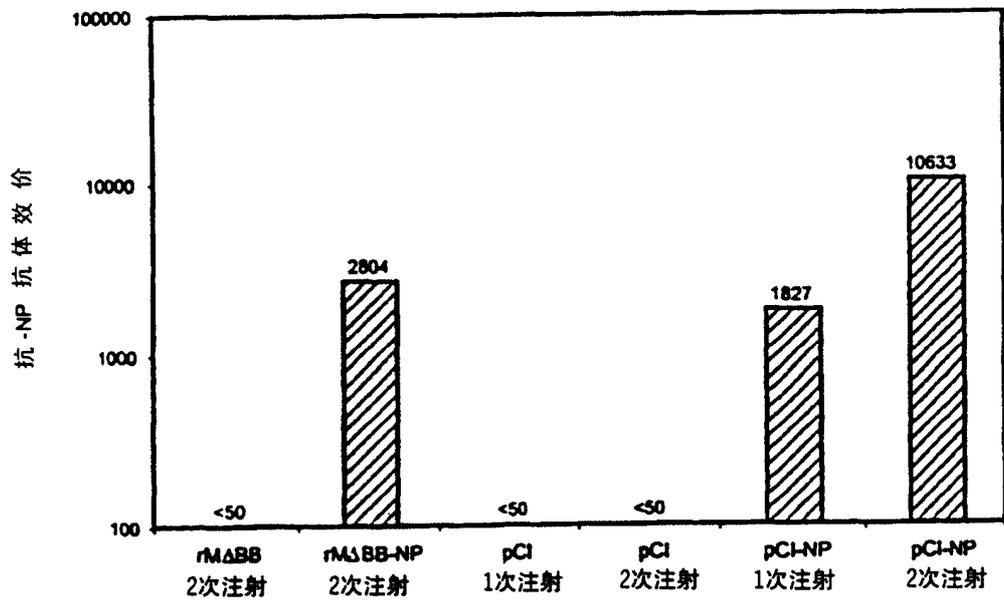


图7

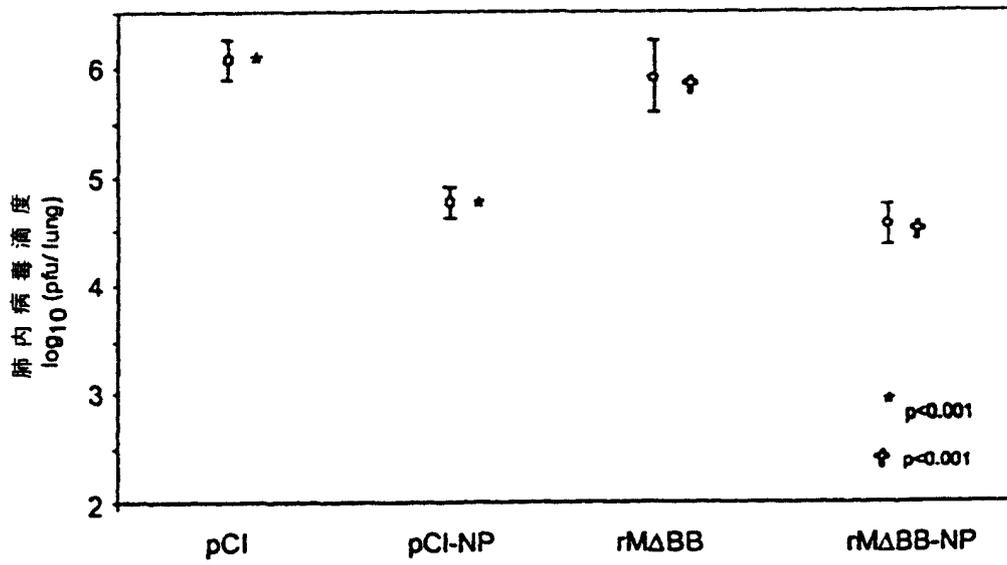


图8

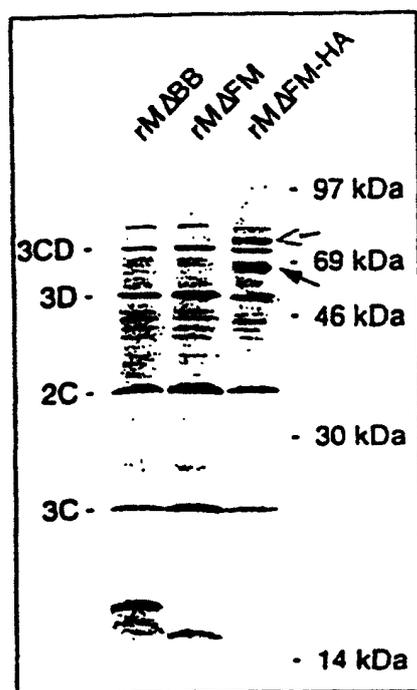


图9A

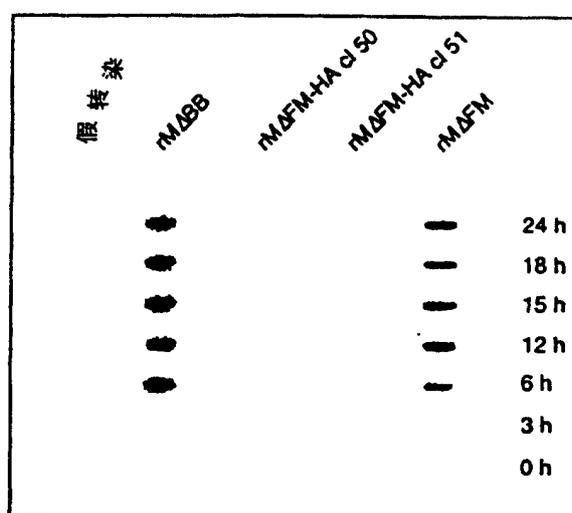


图9B

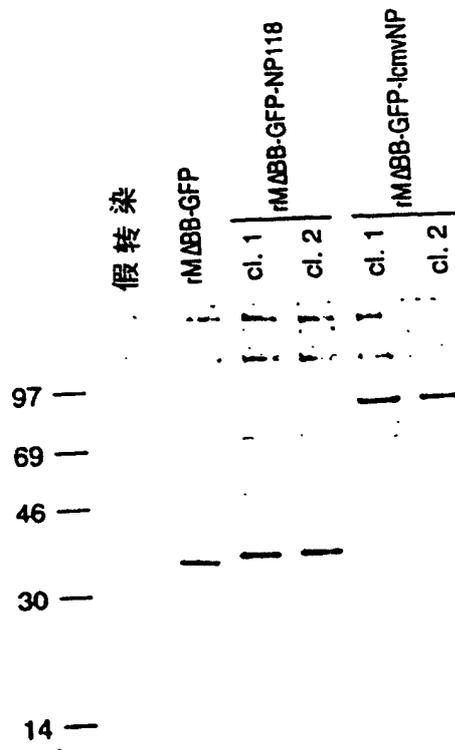


图10

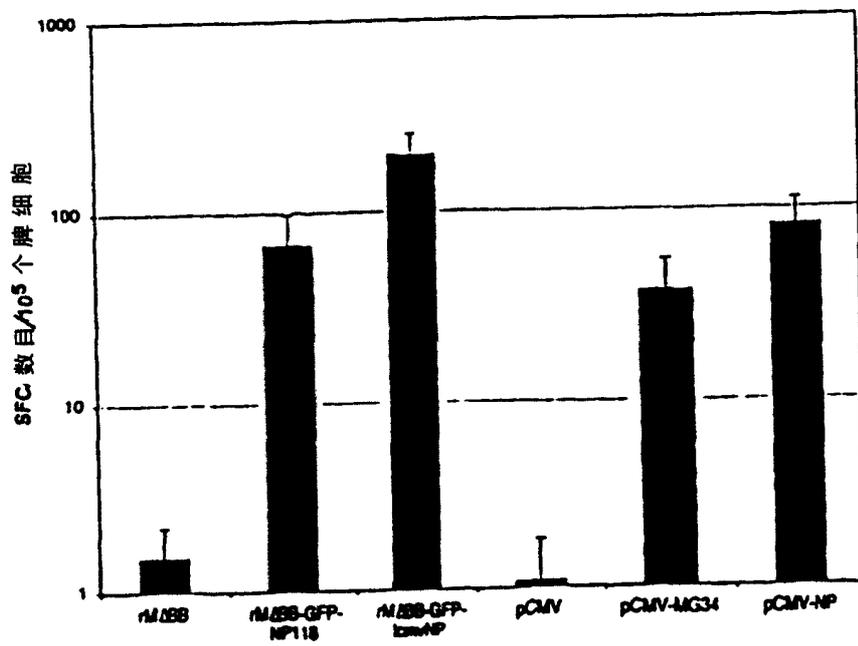


图11

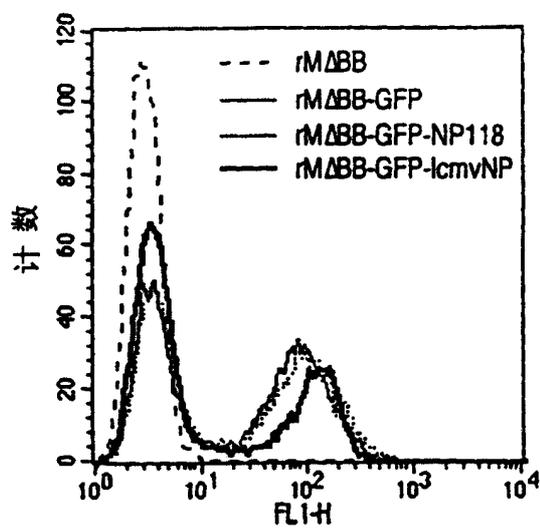


图12