

[19] 中华人民共和国国家知识产权局



## [12] 发明专利申请公布说明书

[21] 申请号 200680052999.8

[51] Int. Cl.

*G01N 33/53 (2006.01)*

*C12P 21/06 (2006.01)*

*C12N 1/00 (2006.01)*

*C12N 5/02 (2006.01)*

*C12N 1/20 (2006.01)*

*C07H 21/04 (2006.01)*

[43] 公开日 2009年3月11日

[11] 公开号 CN 101384902A

[51] Int. Cl. (续)

*A61K 31/70 (2006.01)*

*A01K 67/00 (2006.01)*

[22] 申请日 2006.12.18

[21] 申请号 200680052999.8

[30] 优先权

[32] 2005.12.16 [33] US [31] 60/751,407

[32] 2005.12.16 [33] US [31] 60/751,117

[86] 国际申请 PCT/US2006/048243 2006.12.18

[87] 国际公布 WO2007/102872 英 2007.9.13

[85] 进入国家阶段日期 2008.8.18

[71] 申请人 UAB 研究基金会

地址 美国阿拉巴马州

[72] 发明人 X·吴 J·卡比斯

[74] 专利代理机构 北京北翔知识产权代理有限公司

代理人 张广育 姜建成

权利要求书 34 页 说明书 86 页 序列表 38 页  
附图 18 页

[54] 发明名称

与使用病毒载体控制的基因表达相关的组合物和方法

[57] 摘要

本发明提供了与病毒载体相关的方法和组合物。本发明还提供了有效转染宿主（例如通过高效慢病毒送递载体）的方法和组合物，以及通过单纯地将调节物给予携带被转移的目的序列的宿主来灵敏控制所述被转移的目的序列的表达时机和水平。本发明还公开了制备转基因小鼠的方法以及使用与病毒载体相关的组合物和方法制备得到的转基因小鼠。

1. 一种核酸构建体，包括一种载体，其中所述载体包括第一核酸序列、第二核酸序列和第三核酸序列，其中所述第一核酸序列包括与第一转录调控元件有效连接的目的序列，其中所述第二核酸序列与第二转录调控元件有效连接，并且编码可调控所述第一核酸序列表达的多肽，其中所述第三核酸序列包括与第一转录调控元件有效连接的调节子靶序列，并且其中所述第一和第二转录调控元件取向相反。
2. 权利要求1的核酸构建体，其中所述载体为病毒载体。
3. 权利要求1的核酸构建体，其中所述病毒载体为自我失活型。
4. 权利要求1的核酸构建体，其中所述第一或第二转录调控元件影响所述其他转录调控元件的活性。
5. 权利要求1的核酸构建体，其中所述第一和第二转录调控元件并置。
6. 权利要求1的核酸构建体，其中所述第一和第二转录调控元件之间存在一个接头序列。
7. 权利要求6的核酸构建体，其中所述接头序列为一种染色体隔离子。
8. 一种含有权利要求1的核酸构建体的细胞系。
9. 权利要求1的核酸构建体，其中所述调节子靶序列包括至少一个 tet 操纵子序列。
10. 权利要求1的核酸构建体，其中所述调节子靶序列与一个 TATA 框有效连接。
11. 权利要求1的核酸构建体，其中所述调节子靶序列包括一个侧翼接有两个 tet 操纵子序列的 TATA 框。
12. 权利要求1的核酸构建体，其中所述调节子靶序列包括 SEQ ID NO: 6 的序列。
13. 权利要求1的核酸构建体，其中所述第一核酸序列的表达依赖于 Cre 的存在。
14. 权利要求13的核酸构建体，其中所述调节子靶序列包括一个侧翼接有 TATA 序列的可筛选标记，其中所述 TATA 序列与至少一个 tet 操纵子序列有效连接，其中所述可筛选标记的侧翼接有 lox P 位点，并且其中所述调节子序列位于所述第一转录调控元件和所述第一核酸序列之间。

15. 权利要求 1 的核酸构建体, 其中所述目的序列编码一种目的蛋白。
16. 权利要求 15 的核酸构建体, 其中所述被编码的蛋白为一种 DNA 聚合酶 III 的转录因子。
17. 权利要求 1 的核酸构建体, 其中所述目的序列为微小 RNA 序列、shRNA 序列、siRNA 序列或 mRNA 序列。
18. 权利要求 1 的核酸构建体, 其中所述第二核酸序列包括一种编码四环素抑制子的核酸序列。
19. 权利要求 18 的核酸构建体, 其中所述第二核酸序列还包括一种编码核定位信号的核酸序列。
20. 权利要求 19 的核酸构建体, 其中所述被编码的核定位信号为 SV40 核定位信号。
21. 权利要求 1 的核酸构建体, 其中所述第二核酸序列包括 SEQ ID NO: 1 的序列。
22. 权利要求 18 的核酸构建体, 其中所述第二核酸序列还包括 VP16。
23. 权利要求 22 的核酸构建体, 其中所述第二核酸序列包括 SEQ ID NO: 2 的序列。
24. 权利要求 1 的核酸构建体, 其中所述第二核酸序列包括一种编码四环素激活子的核酸序列。
25. 权利要求 24 的核酸构建体, 其中所述第二核酸序列还包括一种编码核定位信号的核酸序列。
26. 权利要求 25 的核酸构建体, 其中所述被编码的核定位信号为 SV40 核定位信号。
27. 权利要求 1 的核酸构建体, 其中所述第二核酸序列包括 SEQ ID NO: 2 的序列。
28. 权利要求 24 的核酸构建体, 其中所述第二核酸序列还包括 VP16。
29. 权利要求 28 的核酸构建体, 其中所述第二核酸序列包括 SEQ ID NO: 3 的序列。
30. 权利要求 1 的核酸构建体, 还包括与第三转录调控元件有效连接的第四核酸序列, 其中所述第四核酸序列包括第二目的序列。
31. 权利要求 30 的核酸构建体, 其中所述第三转录调控元件为启动子。
32. 权利要求 31 的核酸构建体, 其中所述第二目的序列编码一种目的蛋白。

33. 权利要求 31 的核酸构建体, 其中所述第二目的序列为微小 RNA、shRNA、siRNA 或 mRNA。
34. 权利要求 33 的核酸构建体, 其中所述第三转录调控元件选自小鼠 Hi 启动子、人 H1 启动子或 U6 启动子。
35. 权利要求 1 的核酸构建体, 还包括第四核酸序列, 其中所述第四核酸序列包括一种编码可筛选标记的序列。
36. 权利要求 35 的核酸构建体, 其中所述可筛选标记选自 GFP 或 RFP。
37. 权利要求 35 的核酸构建体, 还包括一个内部核糖体进入位点。
38. 权利要求 1 的核酸构建体, 其中所述第一核酸序列的表达是可调节的。
39. 权利要求 1 的核酸构建体, 其中所述转录调控元件选自启动子或增强子。
40. 权利要求 39 的核酸构建体, 其中所述启动子选自 mH1 启动子、hH1 启动子、CAG 启动子、CMV 启动子、基于 CMV 的启动子、鸡  $\beta$ -肌动蛋白启动子、泛素 C 启动子或 EF-1 $\alpha$  启动子。
41. 权利要求 40 的核酸构建体, 其中所述转录调控元件是可调节的。
42. 权利要求 16 的核酸构建体, 还包括一种与所述调节子靶序列有效连接的编码 GAL-4 的核酸序列。
43. 权利要求 42 的核酸构建体, 包括至少两个 tetO 序列。
44. 权利要求 42 的核酸构建体, 其中所述调节子靶序列包括至少一个 tet 调节子靶序列。
45. 权利要求 42 的核酸构建体, 其中所述调节子靶序列与一个 TATA 框有效连接。
46. 权利要求 42 的核酸构建体, 其中所述调节子靶序列包括一个侧翼接有两个 tet 操纵子序列的 TATA 框。
47. 权利要求 42 的核酸构建体, 其中所述调节子靶序列包括 SEQ ID NO: 6 的序列。
48. 权利要求 42 的核酸构建体, 还包括与第二调节子序列有效连接的编码 GAL-4 的核酸序列和第二目的序列, 其中所述第二目的序列与第三转录调控元件有效连接, 其中所述第二目的序列选自微小 RNA、shRNA 和 siRNA, 其中所述第二调节子序列位于所述编码 GAL-4 的核酸序列和所述第二目的序列之间。

49. 权利要求 48 的核酸构建体, 其中所述第二调节子靶序列包括至少一个调节子靶序列。
50. 权利要求 48 的核酸构建体, 其中所述第二调节子靶序列与一个 TATA 框有效连接。
51. 权利要求 48 的核酸构建体, 其中所述第二调节子靶序列包括一个侧翼接有两个 tet 操纵子序列的 TATA 框。
52. 权利要求 48 的核酸构建体, 其中所述第二调节子靶序列包括 SEQ ID NO: 6 的序列。
53. 权利要求 48 的核酸构建体, 其中所述第三转录调控元件选自启动子或增强子。
54. 权利要求 53 的核酸构建体, 其中所述启动子选自 mH1 启动子、hH1 启动子、泛素 C 启动子或 EF-1 $\alpha$  启动子。
55. 一种核酸构建体, 包括一种载体, 其中所述载体包括第一核酸序列、第二核酸序列和第三核酸序列, 其中所述第一核酸序列、第二核酸序列和第三核酸序列与一个转录调控元件有效连接, 其中所述第一核酸序列包括一个目的序列, 其中所述第二核酸序列编码一个可调控所述第一核酸序列表达的多肽, 其中所述第三核酸序列包括一个与所述第一转录调控元件有效连接的调节子靶序列, 并且其中所述第一转录调控元件能够驱动所述第一和第二核酸序列的表达。
56. 权利要求 55 的核酸构建体, 其中所述载体为病毒载体。
57. 权利要求 56 的核酸构建体, 其中所述病毒载体为自我失活型。
58. 一种含有权利要求 55 的核酸构建体的细胞系。
59. 权利要求 55 的核酸构建体, 其中所述调节子靶序列包括至少一个 tet 操纵子序列。
60. 权利要求 55 的核酸构建体, 其中所述调节子靶序列与一个 TATA 框有效连接。
61. 权利要求 55 的核酸构建体, 其中所述调节子靶序列包括一个侧翼接有两个 tet 操纵子序列的 TATA 框。
62. 权利要求 55 的核酸构建体, 其中所述调节子靶序列包括 SEQ ID NO: 6 的序列。
63. 权利要求 55 的核酸构建体, 其中所述第一核酸序列的表达依赖于 Cre 的存在。

64. 权利要求 63 的核酸构建体, 其中所述调节子序列包括一个侧翼接有 TATA 序列的可筛选标记, 其中所述 TATA 序列与至少一个 tet 操纵子序列连接, 其中所述调节子序列的侧翼还接有 lox P 位点, 并且其中所述调节子序列位于所述第一转录调控元件和所述第一核酸序列之间。
65. 权利要求 55 的核酸构建体, 其中所述目的序列编码一种目的蛋白。
66. 权利要求 65 的核酸构建体, 其中所述目的蛋白为一种 DNA 聚合酶 III 的转录因子。
67. 权利要求 55 的核酸构建体, 其中所述目的序列为微小 RNA、shRNA、siRNA 或 mRNA。
68. 权利要求 55 的核酸构建体, 还包括一种与所述调节子靶序列有效连接的编码 GAL-4 的核酸序列。
69. 权利要求 68 的核酸构建体, 包括 SEQ ID NO: 6 的序列。
70. 权利要求 68 的核酸构建体, 其中所述调节子靶序列包括至少一个 tet 调节子靶序列。
71. 权利要求 68 的核酸构建体, 其中所述调节子靶序列与一个 TATA 框有效连接。
72. 权利要求 68 的核酸构建体, 其中所述调节子靶序列包括一个侧翼接有两个 tet 操纵子序列的 TATA 框。
73. 权利要求 68 的核酸构建体, 其中所述调节子靶序列包括 SEQ ID NO: 6 的序列。
74. 权利要求 55 的核酸构建体, 还包括一种与第二调节子靶序列有效连接的编码 GAL-4 的核酸序列和第二目的序列, 其中所述第二目的序列与第三转录调控元件有效连接, 其中所述第二目的序列选自微小 RNA、shRNA 和 siRNA, 其中所述第二调节子靶序列位于所述编码 GAL-4 的核酸序列和所述第二目的序列之间。
75. 权利要求 74 的核酸构建体, 其中所述第二调节子靶序列包括至少一个 tet 调节子靶序列。
76. 权利要求 74 的核酸构建体, 其中所述第二调节子靶序列与一个 TATA 框有效连接。
77. 权利要求 74 的核酸构建体, 其中所述第二调节子靶序列包括一个侧翼接有两个 tet 操纵子序列的 TATA 框。

78. 权利要求 74 的核酸构建体,其中所述第二调节子靶序列包括 SEQ ID NO: 6 的序列。
79. 权利要求 74 的核酸构建体,其中所述第三转录调控元件选自启动子或增强子。
80. 权利要求 79 的核酸构建体,其中所述启动子选自 mH1 启动子或 hH1 启动子。
81. 权利要求 55 的核酸构建体,其中所述第二核酸序列包括一种编码四环素抑制子的核酸序列。
82. 权利要求 81 的核酸构建体,其中所述第二核酸序列还包括一种编码核定位信号的核酸序列。
83. 权利要求 82 的核酸构建体,其中所述被编码的核定位信号为 SV40 核定位信号。
84. 权利要求 55 的核酸构建体,其中所述第二核酸序列包括 SEQ ID NO: 1 的序列。
85. 权利要求 81 的核酸构建体,其中所述第二核酸序列还包括 VP16。
86. 权利要求 55 的核酸构建体,其中所述第二核酸序列包括 SEQ ID NO: 2 的序列。
87. 权利要求 55 的核酸构建体,其中所述第二核酸序列包括一种编码四环素激活子的核酸序列。
88. 权利要求 87 的核酸构建体,其中所述第二核酸序列还包括一种编码核定位信号的核酸序列。
89. 权利要求 88 的核酸构建体,其中所述被编码的核定位信号序列为 SV40 核定位信号。
90. 权利要求 55 的核酸构建体,其中所述第二核酸序列包括 SEQ ID NO: 2 的序列。
91. 权利要求 87 的核酸构建体,其中所述第二核酸序列还包括 VP16。
92. 权利要求 91 的核酸构建体,其中所述第二核酸序列包括 SEQ ID NO: 3 的序列。
93. 权利要求 55 的核酸构建体,还包括与第三转录调控元件有效连接的第四核酸序列,其中所述第四核酸序列包括第二目的序列。
94. 权利要求 93 的核酸构建体,其中所述第三转录调控元件为启动子。
95. 权利要求 93 的核酸构建体,其中所述第二目的序列编码一种目的

蛋白。

96. 权利要求 93 的核酸构建体, 其中所述第三转录调控元件为启动子。
97. 权利要求 93 的核酸构建体, 其中所述第二目的序列为微小 RNA、shRNA、siRNA 或 mRNA。
98. 权利要求 97 的核酸构建体, 其中所述第三转录调控元件选自小鼠 Hi 启动子、人 H1 启动子、U6 启动子、泛素 C 启动子或 EF-1 $\alpha$  启动子。
99. 权利要求 55 的核酸构建体, 还包括第四核酸序列, 其中所述第四核酸序列包括一种编码可筛选标记的序列。
100. 权利要求 99 的核酸构建体, 其中所述可筛选标记选自 GFP 或 RFP。
101. 权利要求 99 的核酸构建体, 还包括一个内部核糖体进入位点。
102. 权利要求 55 的核酸构建体, 其中所述第一核酸序列的表达是可调节的。
103. 权利要求 55 的核酸构建体, 其中所述转录调控元件选自启动子或增强子。
104. 权利要求 55 的核酸构建体, 其中所述转录调控元件包括 CMV 启动子的 3' 部分。
105. 权利要求 55 的核酸构建体, 其中所述转录调控元件包括 CMV 启动子的 5' 部分。
106. 权利要求 55 的核酸构建体, 其中所述转录调控元件包括  $\beta$ -肌动蛋白启动子的一部分。
107. 权利要求 55 的核酸构建体, 其中所述转录调控元件包括 CAG 启动子。
108. 权利要求 55 的核酸构建体, 其中所述转录调控元件包括与 CAG 启动子有效连接的 3'CMV 启动子序列。
109. 权利要求 55 的核酸构建体, 其中所述转录调控元件序列包括 SEQIDNO: 12 的序列。
110. 权利要求 67 的核酸构建体, 其中所述转录调控元件选自小鼠 Hi 启动子、人 H1 启动子、U6 启动子、泛素 C 启动子或 EF-1 $\alpha$  启动子。
111. 一种双向转录调控元件, 包括与 CAG 启动子的 5' 端融合的 CMV 启动子的 3' 端。
112. 权利要求 111 的双向转录调控元件, 包括 SEQIDNO: 12 中列出的序

- 列。
113. 权利要求 111 的双向转录调控元件, 其中所述双向启动子是可调节的。
114. 权利要求 111 的双向转录调控元件, 其中所述双向启动子为组成型的。
115. 一种双向转录调控元件, 包括与 EF-1 $\alpha$  启动子的 5'端融合的 CMV 启动子的 3'端。
116. 权利要求 115 的双向转录调控元件, 包括 SEQIDNO: 13 中列出的序列。
117. 权利要求 115 的双向转录调控元件, 其中所述双向启动子是可调节的。
118. 权利要求 115 的双向转录调控元件, 其中所述双向启动子为组成型的。
119. 一种选择性调节目的序列表达的方法, 包括在合适于调节目的序列的条件下将权利要求 1 所述的核酸构建体引入靶细胞。
120. 一种制备重组蛋白的方法, 包括在合适于表达重组蛋白的条件下将权利要求 1 所述的核酸构建体引入靶细胞, 其中所述目的序列为编码所述重组蛋白的核酸序列。
121. 权利要求 120 的方法, 其中所述靶细胞产生均一糖基化模式的糖蛋白。
122. 权利要求 120 的方法, 其中与对照细胞相比, 所述靶细胞具有降低的 GnTI 活性。
123. 权利要求 120 的方法, 其中通过在适合于调节所述目的序列的条件下将权利要求 1 所述的核酸构建体引入靶细胞中来制备所述靶细胞。
124. 权利要求 120 的方法, 其中在体外将权利要求 1 所述的核酸构建体引入所述靶细胞中。
125. 一种在受试者体内诱导免疫应答的方法, 包括将按照权利要求 120 所述方法制备的重组蛋白引入受试者体内, 所述重组蛋白的量足以诱导免疫应答。
126. 权利要求 120 的方法, 其中在体内将权利要求 1 所述的核酸构建体引入所述靶细胞中。

127. 一种选择性调节目的序列表达的方法, 包括在适合于调节所述目的序列的条件下将权利要求 55 所述的核酸构建体引入靶细胞。
128. 一种制备重组蛋白的方法, 包括在合适于表达重组蛋白的条件下将权利要求 55 所述的核酸构建体引入靶细胞, 其中所述目的序列为编码所述重组蛋白的核酸序列。
129. 权利要求 128 的方法, 其中所述靶细胞产生均一糖基化模式的糖蛋白。
130. 权利要求 128 的方法, 其中与对照细胞相比, 所述靶细胞具有降低的 GnTI 活性。
131. 权利要求 128 的方法, 其中所述靶细胞可使用权利要求 331 的方法来产生。
132. 权利要求 128 的方法, 其中在体外将所述核酸构建体引入所述靶细胞中。
133. 一种在受试者体内诱导免疫应答的方法, 包括将按照权利要求 128 所述方法制备的重组蛋白引入受试者体内, 所述重组蛋白的量足以诱导免疫应答。
134. 权利要求 128 的方法, 其中在体内将所述核酸构建体引入所述靶细胞中。
135. 一种制备转基因动物的方法, 包括:
- a) 将权利要求 1 所述的核酸构建体引入合子中;
  - b) 允许所述合子发育至足月;
  - c) 获得基因组中含有所述核酸构建体的动物;
  - d) 将所述动物与非转基因动物杂交以获得 F1 代; 和
  - e) 选择基因组中含有所述核酸构建体的动物。
136. 一种基因组中含有转基因的转基因动物, 其中所述转基因含有权利要求 1 所述的核酸构建体。
137. 一种制备转基因动物的方法, 包括:
- a) 将权利要求 55 所述的核酸构建体引入合子中;
  - b) 允许所述合子发育至足月;
  - c) 获得基因组中含有所述核酸构建体的动物;
  - d) 将所述动物与非转基因动物杂交以获得 F1 代; 和
  - e) 选择基因组中含有所述核酸构建体的动物。

138. 一种基因组中含有转基因的转基因动物，其中所述转基因含有权利要求 55 所述的核酸构建体。
139. 一种表达 Kiss-1 的转基因动物。
140. 权利要求 139 的转基因动物，其中所述 Kiss-1 的表达是可调节的。
141. 一种使用权利要求 135 所述的方法产生的表达 Kiss-1 的转基因动物。
142. 一种使用权利要求 137 所述的方法产生的表达 Kiss-1 的转基因动物。
143. 一种表达 Kiss-1 的转基因动物，其中所述动物包括权利要求 1 所述的核酸构建体。
144. 一种表达 Kiss-1 的转基因动物，其中所述动物包括权利要求 55 所述的核酸构建体。
145. 一种表达 FOXP3 的转基因动物。
146. 权利要求 145 的转基因动物，其中所述 FOXP3 的表达是可调节的。
147. 一种使用权利要求 135 所述的方法产生的表达 FOXP3 的转基因动物。
148. 一种使用权利要求 137 所述的方法产生的表达 FOXP3 的转基因动物。
149. 一种表达 FOXP3 的转基因动物，其中所述动物包括权利要求 1 所述的核酸构建体。
150. 一种表达 FOXP3 的转基因动物，其中所述动物包括权利要求 55 所述的核酸构建体。
151. 一种表达 NF  $\kappa$   $\beta$  的转基因动物。
152. 权利要求 151 的转基因动物，其中所述 NF  $\kappa$   $\beta$  的表达是可调节的。
153. 一种使用权利要求 135 所述的方法产生的表达 NF  $\kappa$   $\beta$  的转基因动物。
154. 一种使用权利要求 137 所述的方法产生的表达 NF  $\kappa$   $\beta$  的转基因动物。
155. 一种表达 NF  $\kappa$   $\beta$  的转基因动物，其中所述动物包括权利要求 1 所述的核酸构建体。
156. 一种表达 NF  $\kappa$   $\beta$  的转基因动物，其中所述动物包括权利要求 55 所述的核酸构建体。

157. 一种表达微小 RNA223 的转基因动物。
158. 权利要求 157 的转基因动物，其中所述微小 RNA223 的表达是可调节的。
159. 一种使用权利要求 135 所述的方法产生的表达微小 RNA223 的转基因动物。
160. 一种使用权利要求 137 所述的方法产生的表达微小 RNA223 的转基因动物。
161. 一种表达微小 RNA223 的转基因动物，其中所述动物包括权利要求 1 所述的核酸构建体。
162. 一种表达微小 RNA223 的转基因动物，其中所述动物包括权利要求 55 所述的核酸构建体。
163. 一种表达 Cre 的转基因动物。
164. 权利要求 163 的转基因动物，其中所述 Cre 的表达是可调节的。
165. 一种使用权利要求 447 所述的方法产生的权利要求 163 所述的转基因动物。
166. 一种使用权利要求 448 所述的方法产生的权利要求 163 所述的转基因动物。
167. 一种表达 Cre 的转基因动物，其中所述动物包括权利要求 1 所述的核酸构建体。
168. 一种表达 Cre 的转基因动物，其中所述动物包括权利要求 55 所述的核酸构建体。
169. 一种含有权利要求 1 所述的核酸构建体的疫苗。
170. 一种含有权利要求 55 所述的核酸构建体的疫苗。
171. 一种使受试者体内产生免疫应答的方法，包括给予所述受试者权利要求 169 所述的组合物。
172. 一种使受试者体内产生免疫应答的方法，包括给予所述受试者权利要求 170 所述的组合物。
173. 一种使受试者体内产生免疫应答的方法，其中所述免疫应答为针对 HIV 的免疫应答，包括给予所述受试者权利要求 169 所述的组合物。
174. 一种使受试者体内产生免疫应答的方法，其中所述免疫应答为针对 HIV 的免疫应答，包括给予所述受试者权利要求 170 所述的组

合物。

175. 一种制备重组蛋白的方法，包括：

- a. 将一种含有与调节子序列有效连接的启动子的第一核酸构建体引入靶细胞中，所述调节子序列与至少一个 VP16 序列有效连接；
- b. 使所述细胞处于可使所述待整合的第一核酸构建体整合至所述靶细胞基因组中的条件下；
- c. 将一种第二核酸构建体引入步骤 (b) 所述的细胞中，所述第二核酸构建体含有与目的序列有效连接的调节子靶序列；  
其中所述目的序列为编码重组蛋白的核酸序列；并
- d. 使步骤 (c) 的细胞处于可使重组蛋白表达的条件下。

176. 权利要求 175 的方法，其中所述第一核酸构建体包括 SEQIDNO: 44 的序列。

177. 权利要求 175 的方法，其中第二核酸构建体包括与转录调控元件有效连接的目的序列，所述转录调控元件与调节子靶序列有效连接。

178. 权利要求 175 的方法，其中所述目的序列与 IRES 序列有效连接，所述 IRES 序列与可筛选标记有效连接。

178. 权利要求 175 的方法，其中所述目的序列与 IRES 序列有效连接，所述 IRES 序列与可筛选标记有效连接。

179. 权利要求 175 的方法，其中所述目的序列与 IRES 样序列有效连接，所述 IRES 样序列与可筛选标记有效连接。

180. 一种制备重组蛋白的方法，包括：

- a. 将含有与调节子序列有效连接的启动子的第一核酸构建体引入靶细胞中，所述调节子序列与至少一个 VP16 序列有效连接；
- b. 将第二核酸构建体引入步骤 (a) 所述的同一靶细胞中，所述第二核酸构建体含有与目的序列有效连接的调节子靶序列；  
其中所述目的序列为能够编码重组蛋白的核酸序列；并
- c. 使所述靶细胞处于可使所述待整合的第一和第二核酸序列整合至所述靶细胞基因组中的条件下；并
- d. 使步骤 (c) 的细胞处于可使所述重组蛋白表达的条件下。

181. 权利要求 180 的方法，其中所述第一核酸构建体包括 SEQIDNO: 44 的序列。

182. 权利要求 180 的方法，其中第二核酸构建体包括与转录调控元件有

- 效连接的目的序列，所述转录调控元件与调节子靶序列有效连接。
183. 权利要求 180 的方法，其中所述目的序列与 IRES 序列有效连接，所述 IRES 序列与可筛选标记有效连接。
184. 权利要求 180 的方法，其中所述目的序列与 IRES 样序列有效连接，所述 IRES 样序列与可筛选标记有效连接。
185. 一种在受试者体内诱导免疫应答的方法，包括：
- 将权利要求 1 所述的核酸构建体引入靶细胞中；
  - 使所述细胞处于可使步骤 (a) 所述待整合的核酸构建体整合到所述靶细胞基因组中的条件下；并
  - 将步骤 (b) 所述的靶细胞引入所述受试者体内，所述靶细胞的数量足以诱导免疫应答。
186. 权利要求 185 的方法，其中所述权利要求 1 的核酸构建体的目的序列能够编码一种膜蛋白。
187. 一种在受试者体内诱导免疫应答的方法，包括：
- 将权利要求 55 所述的核酸构建体引入靶细胞中；
  - 使所述细胞处于可使步骤 (a) 所述待整合的核酸构建体整合到所述靶细胞基因组中的条件下；并
  - 将步骤 (b) 所述的靶细胞引入所述受试者体内，所述靶细胞的数量足以诱导免疫应答。
188. 权利要求 187 的方法，其中所述权利要求 55 的核酸构建体的目的序列能够编码一种膜蛋白。
189. 一种在受试者体内诱导免疫应答的方法，包括：
- 将权利要求 1 所述的核酸构建体引入靶细胞中，其中所述权利要求 1 的核酸构建体的转录调控元件为可调节的；
  - 使所述细胞处于可使步骤 (a) 所述待整合的核酸构建体整合到所述靶细胞基因组中的条件下；并
  - 将步骤 (b) 所述的靶细胞引入所述受试者体内；并
  - 给予所述受试者有效量的能够调节权利要求 1 的核酸构建体的转录调控元件的物质，所述物质的量足以诱导所述目的序列的表达，其中所述目的序列的表达量足以诱导免疫应答。
190. 权利要求 189 的方法，其中所述权利要求 1 的核酸构建体的目的序

列能够编码一种膜蛋白。

191. 一种在受试者体内诱导免疫应答的方法，包括：

- a. 将权利要求 55 所述的核酸构建体引入靶细胞中；
- b. 使所述细胞处于可使步骤 (a) 所述待整合的核酸构建体整合到所述靶细胞基因组中的条件下；并
- c. 将步骤 (b) 所述的靶细胞引入所述受试者体内；并
- d. 给予所述受试者有效量的能够调节权利要求 55 的核酸构建体的转录调控元件的物质，所述物质的量足以诱导所述目的序列的表达，

其中所述目的序列的表达量足以诱导免疫应答。

192. 权利要求 191 的方法，其中所述权利要求 55 的核酸构建体的目的序列能够编码一种膜蛋白。

193. 一种生产抗目的蛋白的抗体的方法，包括：

- a. 将权利要求 1 所述的核酸构建体引入靶细胞中，  
其中所述权利要求 1 的核酸构建体的转录调控元件为可调节的，  
其中所述目的序列能够编码一种目的蛋白；
- b. 使所述细胞处于可使步骤 (a) 所述待整合的核酸构建体整合到所述靶细胞基因组中的条件下；
- c. 将步骤 (b) 所述的细胞引入所述受试者体内；
- d. 给予所述受试者有效量的能够调节权利要求 1 的核酸构建体的转录调控元件的物质，所述物质的量足以诱导所述目的序列的表达，

其中所述目的序列的表达量足以诱导免疫应答，并且

其中所述免疫应答可产生抗所述目的蛋白的抗体。

194. 权利要求 193 的方法，还包括分离通过所述方法产生的抗体。

195. 权利要求 193 的方法，其中所述权利要求 1 的核酸构建体的目的序列能够编码一种膜蛋白。

196. 一种生产抗目的蛋白的抗体的方法，包括：

- a. 将权利要求 55 所述的核酸构建体引入靶细胞中，  
其中所述权利要求 55 的核酸构建体的转录调控元件为可调节的，  
其中所述目的序列能够编码一种目的蛋白；
- b. 使所述细胞处于可使步骤 (a) 所述待整合的核酸构建体整合到所

- 述靶细胞基因组中的条件下；
- c. 将步骤 (b) 所述的细胞引入所述受试者体内；
- d. 给予所述受试者有效量的一种能够调节权利要求 55 的核酸构建体的转录调控元件的物质，所述物质的量足以诱导所述目的序列的表达，
- 其中所述目的序列的表达量足以诱导免疫应答，并且
- 其中所述免疫应答可产生抗所述目的蛋白的抗体。
197. 权利要求 196 的方法，还包括分离通过所述方法产生的抗体。
198. 权利要求 196 的方法，其中所述权利要求 55 的核酸构建体的目的序列能够编码一种膜蛋白。
199. 一种生产抗目的蛋白的抗体的方法，包括：
- a. 将权利要求 1 所述的核酸构建体引入靶细胞中，其中所述目的序列能够编码一种目的蛋白；
- b. 使所述细胞处于可使步骤 (a) 所述待整合的核酸构建体整合到所述靶细胞基因组中的条件下；并
- c. 将步骤 (b) 所述的细胞引入所述受试者体内，
- 其中所述细胞可表达目的序列，所述目的序列的表达量足以诱导免疫应答，
- 其中所述免疫应答可产生抗所述目的蛋白的抗体。
200. 权利要求 199 的方法，还包括分离通过所述方法产生的抗体。
201. 权利要求 199 的方法，其中所述权利要求 1 的核酸构建体的目的序列能够编码一种膜蛋白。
202. 一种生产抗目的蛋白的抗体的方法，包括：
- a. 将权利要求 55 所述的核酸构建体引入靶细胞中，
- 其中所述权利要求 55 的核酸构建体的转录调控元件为可调节的，
- 其中所述目的序列能够编码一种目的蛋白；
- b. 使所述细胞处于可使步骤 (a) 所述待整合的核酸构建体整合到所述靶细胞基因组中的条件下；
- c. 将步骤 (b) 所述的细胞引入所述受试者体内，
- 其中所述细胞可表达目的序列，所述目的序列的表达量足以诱导免疫应答，
- 其中所述免疫应答可产生抗所述目的蛋白的抗体。

203. 权利要求 202 的方法，还包括分离通过所述方法产生的抗体。
204. 权利要求 202 的方法，其中所述权利要求 55 的核酸构建体的目的序列能够编码一种膜蛋白。
205. 一种双向转录调控元件，包括与 ef1 $\alpha$  启动子的 5'端融合的 CMV 启动子的 5'端。
206. 权利要求 205 的双向转录调控元件，还包括一种与所述 CMV 启动子的 3'端有效连接的调节子靶序列。
207. 权利要求 205 的双向转录调控元件，包括 SEQIDNO: 51 中列出的序列。
208. 一种包装构建体，包括第一和第二核酸序列，其中所述第一核酸序列编码 Gag 多聚蛋白，其中所述第二核酸序列编码 Gag-Pro-Pol 多聚蛋白，其中所述第一和第二核酸序列各包括一个或多个可减少移码或翻译连读的突变，其中所述第一和第二核酸序列可由同一核苷酸序列的不同编码区表达，并且其中所述第一和第二核酸序列与至少一个转录调控元件有效连接。
209. 权利要求 208 的包装构建体，还包括一种含有 rev 应答元件的第三核酸序列。
210. 一种含有权利要求 208 所述的包装构建体的细胞系。
211. 权利要求 208 的包装构建体，还包括一种位于所述第一和第二核酸序列之间的元件，其中所述元件可使所述第一和第二核酸序列产生差异表达。
212. 权利要求 211 的包装构建体，其中所述位于第一和第二核酸序列之间的元件为内部核糖体进入位点或内部核糖体进入位点样元件。
213. 权利要求 212 的包装构建体，其中所述内部核糖体进入位点为 EMC-病毒 IRES 或 HCV-病毒 IRES。
214. 权利要求 212 的包装构建体，其中所述内部核糖体进入位点样元件为天然存在的。
215. 权利要求 212 的包装构建体，其中所述内部核糖体进入位点或内部核糖体进入位点样元件为突变的。
216. 权利要求 208 的包装构建体，其中至少一个转录调控元件是可调节的。
217. 权利要求 208 的包装构建体，其中至少一个转录调控元件可由四环

素或多西环素调节。

218. 权利要求 208 的包装构建体, 其中至少一个转录调控元件含有至少一个 tet 操纵子序列。
219. 权利要求 208 的包装构建体, 其中至少一个 tet 操纵子序列与一个 TATA 框有效连接。
220. 权利要求 208 的包装构建体, 其中至少一个转录调控元件序列包括 SEQIDNO: 6 的序列。
221. 权利要求 208 的包装构建体, 其中至少一个转录调控元件为启动子。
222. 权利要求 221 的包装构建体, 其中所述启动子为基于 CMV 的启动子、CAG 启动子、基于 SV40 的启动子、基于热休克蛋白的启动子、基于 mH1 的启动子、基于 hH1 的启动子、基于鸡  $\beta$ -肌动蛋白的启动子、基于 U6 的启动子、基于泛素 C 的启动子或基于 EF-1 $\alpha$  启动子。
223. 权利要求 208 的包装构建体, 其中至少一个转录调控元件是组成型的。
224. 权利要求 223 的包装构建体, 其中至少一个所述组成型转录调控元件为启动子。
225. 权利要求 224 的包装构建体, 其中所述启动子为基于 CMV 的启动子、CAG 启动子、基于 SV40 的启动子、基于热休克蛋白的启动子、基于 mH1 的启动子、基于 hH1 的启动子、基于鸡  $\beta$ -肌动蛋白的启动子、基于 U6 的启动子、基于泛素 C 的启动子或基于 EF-1 $\alpha$  启动子。
226. 权利要求 208 的包装构建体, 其中所述第一核酸序列含有在 gag 序列内的移码所必需的至少一个点突变。
227. 权利要求 208 的包装构建体, 其中所述第一核酸序列包括 SEQIDNO: 19 的核苷酸序列。
228. 权利要求 208 的包装构建体, 其中所述第二核酸序列含有在 gag-pol 序列内的移码所必需的一个单核苷酸插入和至少一个点突变。
229. 权利要求 208 的包装构建体, 其中所述第二核酸序列包括 SEQIDNO: 11 的核苷酸序列。
230. 权利要求 208 的包装构建体, 其中一个或多个所述第一和第二核酸序列是密码子优化的。
231. 权利要求 208 的包装构建体, 还包括一个编码 Tat 的核酸序列。

232. 权利要求 208 的包装构建体, 还包括一个编码 Rev 的核酸序列。
233. 权利要求 208 的包装构建体, 其中所述构建体能够以约 99:1 至约 80:20 的比例表达 Gag 和 Gag-Pol。
234. 权利要求 208 的包装构建体, 其中所述构建体能够产生无复制能力的重组体。
235. 权利要求 208 的包装构建体, 其中所述第一核酸序列与第一转录调控元件有效连接, 并且所述第二核酸序列与第二转录调控元件有效连接。
236. 权利要求 235 的包装构建体, 其中所述第一转录调控元件强于所述第二转录调控元件。
237. 权利要求 208 的包装构建体, 其中所述第一和第二核酸序列的表达方向相反。
238. 权利要求 208 的包装构建体, 其中至少一个所述转录调控元件是可调节的。
239. 权利要求 208 的包装构建体, 其中所述第一和第二核酸序列与一单转录调控元件有效连接。
240. 权利要求 239 的包装构建体, 其中所述单转录调控元件是可调节的。
241. 权利要求 239 的包装构建体, 其中所述单转录调控元件为双向启动子。
242. 权利要求 241 的包装构建体, 其中所述双向启动子是可调节的。
243. 一种包装系统, 包括权利要求 208 所述的包装构建体和一种表达包膜糖蛋白的核酸构建体。
244. 一种包装系统, 包括权利要求 209 所述的包装构建体和一种表达包膜糖蛋白的核酸构建体。
245. 一种使用权利要求 243 所述的包装系统制备病毒样颗粒的方法。
246. 一种使用权利要求 244 所述的包装系统制备病毒样颗粒的方法。
247. 一种包装构建体, 包括第一和第二核酸序列, 其中所述第一核酸序列编码 Gag 多聚蛋白, 其中所述第二核酸序列编码 Gag-Pro 多聚蛋白, 其中所述第一和第二核酸序列包括一个或多个能减少移码或翻译连续的突变, 其中所述第一和第二核酸序列可由同一核苷酸序列的不同编码区表达, 并且其中所述第一和第二核酸序列与至少一个转录调控元件有效连接。

248. 权利要求 247 的包装构建体, 还包括一种含有 rev 应答元件的第三核酸序列。
249. 权利要求 247 的包装构建体, 还包括一种位于所述第一和第二核酸序列之间的元件, 其中所述元件可使所述两种核酸序列产生差异表达。
250. 权利要求 249 的包装构建体, 其中所述位于第一和第二核酸序列之间的元件为内部核糖体进入位点或内部核糖体进入位点样元件。
251. 权利要求 250 的包装构建体, 其中所述内部核糖体进入位点为 EMC-病毒 IRES 或 HCV-病毒 IRES。
252. 权利要求 250 的包装构建体, 其中所述内部核糖体进入位点样元件为天然存在的。
253. 权利要求 250 的包装构建体, 其中所述内部核糖体进入位点或内部核糖体进入位点样元件为突变的。
254. 权利要求 247 的包装构建体, 其中至少一个转录调控元件是可调节的。
255. 权利要求 247 的包装构建体, 其中至少一个转录调控元件可由四环素或多西环素调节。
256. 权利要求 247 的包装构建体, 其中至少一个转录调控元件含有至少一个 tet 操纵子序列。
257. 权利要求 256 的包装构建体, 其中至少一个 tet 操纵子序列与一个 TATA 框有效连接。
258. 权利要求 247 的包装构建体, 其中至少一个转录调控元件序列含有 SEQIDNO: 6 的序列。
259. 权利要求 247 的包装构建体, 其中至少一个转录调控元件为启动子。
260. 权利要求 259 的包装构建体, 其中所述启动子为基于 CMV 的启动子、CAG 启动子、基于 SV40 的启动子、基于热休克蛋白的启动子、基于 mH1 的启动子、基于 hH1 的启动子、基于鸡  $\beta$ -肌动蛋白的启动子、基于 U6 的启动子、基于泛素 C 的启动子或基于 EF-1 $\alpha$  启动子。
261. 权利要求 247 的包装构建体, 其中一种或多种所述转录调控元件是组成型的。
262. 权利要求 247 的包装构建体, 其中所述第一核酸序列含有在 gag 序列内的移码所必需的至少一个点突变。

263. 权利要求 247 的包装构建体，其中所述第一核酸序列包括 SEQIDNO: 19 的核苷酸序列。
264. 权利要求 247 的包装构建体，其中所述第二核酸序列含有在 gag-pol 序列内的移码所必需的一个单核苷酸插入和至少一个点突变。
265. 权利要求 247 的包装构建体，其中所述第二核酸序列包括 SEQIDNO: 20 的核苷酸序列。
266. 权利要求 247 的包装构建体，其中一个或多个所述第一和第二核酸序列是密码子优化的。
267. 权利要求 247 的包装构建体，还包括一个编码 Tat 的核酸序列。
268. 权利要求 247 的包装构建体，还包括一个编码 Rev 的核酸序列。
269. 权利要求 247 的包装构建体，其中所述构建体能够以约 99:1 至约 80:20 的比例表达 Gag 和 Gag-Pol。
270. 权利要求 247 的包装构建体，其中所述构建体能够产生无复制能力的重组体。
271. 权利要求 247 的包装构建体，其中所述第一核酸序列与第一转录调控元件有效连接，并且所述第二核酸序列与第二转录调控元件有效连接。
272. 权利要求 271 的包装构建体，其中所述第一转录调控元件强于所述第二转录调控元件。
273. 权利要求 271 的包装构建体，其中所述第一和第二核酸序列的表达方向相反。
274. 权利要求 271 的包装构建体，其中所述第一和第二转录调控元件是相同的。
275. 权利要求 271 的包装构建体，其中所述第一和第二转录调控元件是不同的。
276. 权利要求 271 的包装构建体，其中至少一个所述转录调控元件是可调节的。
277. 权利要求 247 的包装构建体，其中所述靶第一和第二核酸序列与一单转录调控元件有效连接。
278. 权利要求 277 的包装构建体，其中所述单转录调控元件是可调节的。
279. 权利要求 277 的包装构建体，其中所述单转录调控元件为双向启动

子。

280. 权利要求 279 的包装构建体, 其中所述双向启动子是可调节的。
281. 一种包装系统, 包括权利要求 247 所述的包装构建体和一种表达包膜糖蛋白的核酸构建体。
282. 一种包装系统, 包括权利要求 248 所述的包装构建体和一种表达包膜糖蛋白的核酸构建体。
283. 一种使用权利要求 281 所述的包装系统制备病毒样颗粒的方法。
284. 一种使用权利要求 282 所述的包装系统制备病毒样颗粒的方法。
285. 一种含有权利要求 247 所述的靶包装构建体的细胞系。
286. 一种包装构建体, 包括第一、第二和第三核酸序列, 其中所述第一核酸序列编码 Gag 多聚蛋白, 其中所述第二核酸序列编码 Gag-Pro 多聚蛋白, 其中所述第三核酸序列编码 Vpr-反转录酶-整合酶蛋白, 其中所述第一和第二核酸序列包括一个或多个能减少移码或翻译连续的突变, 其中所述第一、第二和第三核酸序列可由同一核苷酸序列的不同编码区表达, 其中所述第一、第二和第三核酸序列与至少一个转录调控元件有效连接。
287. 权利要求 286 的包装构建体, 还包括位于所述第一和第二核酸序列之间的元件, 其中所述元件可使所述第一和第二核酸序列产生差异表达。
288. 权利要求 287 的包装构建体, 其中所述位于第一和第二核酸序列之间的可使二者产生差异表达的元件为内部核糖体进入位点或内部核糖体进入位点样元件。
289. 权利要求 288 的包装构建体, 其中所述内部核糖体进入位点为 EMC-病毒 IRES 或 HCV-病毒 IRES。
290. 权利要求 288 的包装构建体, 其中所述内部核糖体进入位点样元件为天然存在的。
291. 权利要求 288 的包装构建体, 其中所述内部核糖体进入位点或内部核糖体进入位点样元件为突变的。
292. 权利要求 286 的包装构建体, 其中至少一个转录调控元件是可调节的。
293. 权利要求 286 的包装构建体, 其中至少一个转录调控元件可由四环素或多西环素调节。

294. 权利要求 286 的包装构建体, 其中至少一个转录调控元件含有至少一个 tet 操纵子序列。
295. 权利要求 294 的包装构建体, 其中至少一个 tet 操纵子序列与一个 TATA 框有效连接。
296. 权利要求 286 的包装构建体, 其中至少一个转录调控元件含有 SEQIDNO: 6 的序列。
297. 权利要求 286 的包装构建体, 其中至少一个转录调控元件为启动子。
298. 权利要求 297 的包装构建体, 其中所述启动子为基于 CMV 的启动子、CAG 启动子、基于 SV40 的启动子、基于热休克蛋白的启动子、基于 mH1 的启动子、基于 hH1 的启动子、基于鸡  $\beta$ -肌动蛋白的启动子、基于 U6 的启动子、基于泛素 C 的启动子或基于 EF-1  $\alpha$  启动子。
299. 权利要求 286 的包装构建体, 其中一种或多种所述转录调控元件是组成型的。
300. 权利要求 299 的包装构建体, 其中至少一个所述组成型转录调控元件为启动子。
301. 权利要求 300 的包装构建体, 其中所述启动子为基于 CMV 的启动子、CAG 启动子、基于 SV40 的启动子、基于热休克蛋白的启动子、基于 mH1 的启动子、基于 hH1 的启动子、基于鸡  $\beta$ -肌动蛋白的启动子、基于 U6 的启动子、基于泛素 C 的启动子或基于 EF-1  $\alpha$  启动子。
302. 权利要求 286 的包装构建体, 其中所述第一核酸序列含有在 gag 序列内的移码所必需的至少一个点突变。
303. 权利要求 286 的包装构建体, 其中所述第一核酸序列包括 SEQIDNO: 19 的核苷酸序列。
304. 权利要求 286 的包装构建体, 其中所述第二核酸序列含有在 gag-pol 序列内的移码所必需的一个单核苷酸插入和至少一个点突变。
305. 权利要求 286 的包装构建体, 其中所述第二核酸序列包括 SEQIDNO: 20 的核苷酸序列。
306. 权利要求 286 的包装构建体, 其中一个或多个所述第一和第二核酸序列是密码子优化的。
307. 权利要求 286 的包装构建体, 还包括一个编码 Tat 的核酸序列。
308. 权利要求 286 的包装构建体, 还包括一个编码 Rev 的核酸序列。

309. 权利要求 286 的包装构建体, 其中所述构建体能够以约 99:1 至约 80:20 的比例表达 Gag 和 Gag-Pro。
310. 权利要求 286 的包装构建体, 其中所述构建体能够产生无复制能力的重组体。
311. 权利要求 286 的包装构建体, 其中所述第一和第二核酸序列与第一转录调控元件有效连接, 并且所述第三核酸序列与第二转录调控元件有效连接。
312. 权利要求 311 的包装构建体, 其中所述第一转录调控元件强于所述第二转录调控元件。
313. 权利要求 286 的包装构建体, 其中所述第一和第二核酸序列或第三核酸序列的表达方向相反。
314. 权利要求 313 的包装构建体, 其中所述第一和第二核酸序列与第一转录调控元件有效连接, 并且所述第三核酸序列与第二转录调控元件有效连接。
315. 权利要求 314 的包装构建体, 其中至少一个所述转录调控元件是可调节的。
316. 权利要求 313 的包装构建体, 其中所述第一、第二和第三核酸序列与一单转录调控元件有效连接。
317. 权利要求 316 的包装构建体, 其中所述单转录调控元件是可调节的。
318. 权利要求 316 的包装构建体, 其中所述单转录调控元件为双向启动子。
319. 权利要求 318 的包装构建体, 其中所述双向启动子是可调节的。
320. 权利要求 286 的包装构建体, 还包括一个位于所述第一或第二核酸序列与所述第三核酸序列之间的元件, 其中所述第三核酸序列并不位于所述第一和第二核酸序列之间, 并且其中所述元件可使所述第一或第二核酸序列与所述第三核酸序列之间产生差异表达。
321. 权利要求 320 的包装构建体, 其中所述元件为内部核糖体进入位点或内部核糖体进入位点样元件。
322. 权利要求 321 的包装构建体, 其中所述内部核糖体进入位点为 EMC-病毒 IRES 或 HCV-病毒 IRES。
323. 权利要求 321 的包装构建体, 其中所述内部核糖体进入位点样元件为天然存在的。

324. 权利要求 321 的包装构建体, 其中内部核糖体进入位点或内部核糖体进入位点样元件为突变的。
325. 一种包装系统, 包括权利要求 286 所述的包装构建体和一种表达包膜糖蛋白的核酸构建体。
326. 一种使用权利要求 325 所述的包装系统制备病毒样颗粒的方法。
327. 一种表达系统, 包括:
- a. 权利要求 208、247 或 286 所述的包装构建体;
  - b. 一种包膜核酸构建体, 包括一种编码包膜糖蛋白的核酸序列, 其中所述编码包膜糖蛋白的核酸序列与至少一个转录调控元件有效连接; 和
  - c. 一种含有一种或多种目的基因的基因转移构建体。
328. 权利要求 327 的表达系统, 还包括一种编码核定位信号的构建体, 所述构建体含有与一种编码四环素反式作用子的核酸有效连接的编码核定位信号的核酸序列。
329. 权利要求 328 的表达系统, 其中所述编码核定位信号的构建体还包括一种转录调控元件。
330. 权利要求 329 的表达系统, 其中所述转录调控元件为启动子或增强子。
331. 权利要求 328 的表达系统, 其中所述编码核定位信号的核酸序列的侧翼接有至少一个接头序列。
332. 权利要求 331 的表达系统, 其中所述接头序列编码 SEQIDNO: 15。
333. 权利要求 328 的表达系统, 其中所述编码核定位信号的构建体从 5' 端至 3' 端依次包括巨细胞病毒启动子、第一接头编码序列、第二核定位信号、第二接头序列和四环素反式作用子编码序列, 其中所述被编码的接头为 SEQIDNO: 15。
334. 权利要求 327 的表达系统, 其中所述包膜糖蛋白有助于进入细胞。
335. 权利要求 327 的表达系统, 其中所述包膜糖蛋白为病毒包膜糖蛋白。
336. 权利要求 335 的表达系统, 其中所述病毒包膜糖蛋白为囊泡口炎病毒的 G 蛋白 (VSV-G)。
337. 权利要求 327 的表达系统, 其中所述基因转移构建体还包括一种标记编码序列, 并且其中所述目的基因和标记编码序列与至少一个转录调控元件有效连接。

338. 权利要求 327 的表达系统, 其中所述基因转移构建体包括两种目的基因。
339. 权利要求 327 的表达系统, 其中所述基因转移构建体还包括一个位于所述目的基因 3'端的土拨鼠肝炎病毒转录后调节元件。
340. 权利要求 327 的表达系统, 其中所述基因转移构建体包括一个或多个长末端重复序列。
341. 权利要求 340 的表达系统, 其中所述基因转移构建体包括一个突变, 所述突变位于 3'长末端重复序列内。
342. 权利要求 340 的表达系统, 其中用一种启动子序列代替 5'或 3'长末端重复序列。
343. 权利要求 337 的表达系统, 其中所述基因转移构建体的至少一个转录调控元件为启动子。
344. 权利要求 337 的表达系统, 其中所述基因转移构建体的转录调控元件是可调节的。
345. 权利要求 344 的表达系统, 其中所述基因转移构建体的转录调控元件还包括一种调节子构建体, 并且其中所述调节子构建体包括一种与至少一个转录调控元件有效连接的可调节元件。
346. 权利要求 337 的表达系统, 其中所述基因转移构建体还包括一种位于所述目的基因和所述标记编码序列之间的元件, 并且其中所述元件可使所述目的基因和所述标记编码序列产生差异表达。
347. 权利要求 346 的表达系统, 其中所述位于目的基因和标记编码序列之间的元件为内部核糖体进入位点或内部核糖体进入位点样元件。
348. 权利要求 347 的表达系统, 其中所述内部核糖体进入位点为 EMC-病毒 IRES 或 HCV-病毒 IRES。
349. 权利要求 347 的表达系统, 其中所述内部核糖体进入位点样元件为天然存在的。
350. 权利要求 347 的表达系统, 其中内部核糖体进入位点或内部核糖体进入位点样元件为突变的。
351. 权利要求 337 的表达系统, 其中所述基因转移构建体包括至少一个转录调控元件。
352. 权利要求 337 的表达系统, 其中所述转录调控元件是可调节的。
353. 权利要求 337 的表达系统, 其中所述转录调控元件可由四环素或多

四环素调节。

354. 权利要求 337 的表达系统, 其中至少一个转录调控元件含有至少一个 tet 操纵子序列。
355. 权利要求 337 的表达系统, 其中至少一个转录调控元件序列含有 SEQIDNO: 6 的序列。
356. 权利要求 351 的表达系统, 其中一种或多种所述转录调控元件为启动子。
357. 权利要求 356 的表达系统, 其中一种或多种所述启动子为基于 CMV 的启动子、CAG 启动子、基于 SV40 的启动子、基于热休克蛋白的启动子、基于 mH1 的启动子、基于 hH1 的启动子、基于鸡  $\beta$ -肌动蛋白的启动子、基于 U6 的启动子、基于泛素 C 的启动子或基于 EF-1  $\alpha$  启动子。
358. 权利要求 351 的表达系统, 其中所述基因转移构建体的至少一个转录调控元件为组成型的。
359. 权利要求 358 的表达系统, 其中一种或多种所述组成型转录调控元件为启动子。
360. 权利要求 359 的表达系统, 其中一种或多种所述启动子为基于 CMV 的启动子、CAG 启动子、基于 SV40 的启动子、基于热休克蛋白的启动子、基于 mH1 的启动子、基于 hH1 的启动子、基于鸡  $\beta$ -肌动蛋白的启动子、基于 U6 的启动子、基于泛素 C 的启动子或基于 EF-1  $\alpha$  启动子。
361. 权利要求 337 的表达系统, 其中所述标记编码序列和所述目的基因与不同的转录调控元件有效连接。
362. 权利要求 361 的表达系统, 其中所述与标记编码序列有效连接的转录调控元件强于所述与目的基因有效连接的转录调控元件。
363. 权利要求 337 的表达系统, 其中所述标记编码序列和所述目的基因的表达方向相反。
364. 权利要求 363 的表达系统, 其中至少一个所述转录调控元件是可调节的。
365. 权利要求 363 的表达系统, 其中所述标记编码序列和所述目的基因与一单转录调控元件有效连接。
366. 权利要求 365 的表达系统, 其中所述单转录调控元件是可调节的。

367. 权利要求 366 的表达系统, 其中所述单转录调控元件为双向启动子。
368. 一种含有权利要求 327 所述的表达系统的细胞系。
369. 一种表达系统, 包括:
- a. 第一包装构建体, 包括含有一种编码 Gag 多聚蛋白的核酸序列的第一核酸构建体, 其中所述编码 Gag 的核酸序列包括一个或多个可减少移码或翻译连读的突变, 并且与至少一个转录调控元件有效连接;
  - b. 第二包装构建体, 包括含有一种编码 Gag-Pol 多聚蛋白的核酸序列的第一核酸构建体, 其中所述编码 Gag-Pol 的核酸序列包括一个或多个可减少移码或翻译连读的突变, 并且与至少一个转录调控元件有效连接;
  - c. 第三核酸构建体, 包括编码包膜糖蛋白的第三核酸序列, 其中所述第三核酸序列与至少一个转录调控元件有效连接; 和
  - d. 一种含有至少一个目的基因的基因转移构建体。
370. 一种含有权利要求 369 所述的表达系统的细胞系。
371. 权利要求 369 的表达系统, 还包括含有编码核定位信号的第四核酸序列的第四核酸构建体, 所述核定位信号与四环素反式作用子编码序列有效连接。
372. 权利要求 371 的表达系统, 其中所述第四核酸构建体还包括一种转录调控元件。
373. 权利要求 371 的表达系统, 其中所述编码核定位信号的序列的侧翼接有至少一个接头序列。
374. 权利要求 373 的表达系统, 其中所述接头序列编码 SEQIDNO: 15。
375. 权利要求 371 的表达系统, 其中第四核酸序列从 5' 端至 3' 端依次包括巨细胞病毒启动子、编码 SEQIDNO: 15 的核酸序列、编码核定位信号的核酸序列、编码 SEQIDNO: 15 的核酸序列和编码四环素反式作用子的核酸序列。
376. 权利要求 371 的表达系统, 其中所述病毒包膜糖蛋白为囊泡口炎病毒的 G 蛋白 (VSV-G)。
377. 权利要求 369 的表达系统, 其中所述基因转移构建体还包括一种标记编码序列, 并且其中所述目的基因和标记编码序列与至少一个转录调控元件有效连接。

378. 权利要求 369 的表达系统, 其中所述基因转移构建体还包括第二目的基因。
379. 权利要求 369 的表达系统, 其中所述基因转移构建体还包括一种位于所述目的基因和所述标记编码序列之间的元件, 并且其中所述元件可使所述目的基因和所述标记编码序列产生差异表达。
380. 权利要求 379 的表达系统, 其中所述位于目的基因和标记编码序列之间的元件为内部核糖体进入位点或内部核糖体进入位点样元件。
381. 权利要求 380 的表达系统, 其中所述内部核糖体进入位点为 EMC-病毒 IRES 或 HCV-病毒 IRES。
382. 权利要求 380 的表达系统, 其中所述内部核糖体进入位点样元件为天然存在的。
383. 权利要求 380 的表达系统, 其中内部核糖体进入位点或内部核糖体进入位点样元件为突变的。
384. 权利要求 369 的表达系统, 其中所述基因转移构建体包括至少一个转录调控元件。
385. 权利要求 369 的表达系统, 其中所述转录调控元件是可调节的。
386. 权利要求 369 的表达系统, 其中所述转录调控元件可由四环素或多西环素调节。
387. 权利要求 369 的表达系统, 其中至少一个转录调控元件含有至少一个 tet 操纵子序列。
388. 权利要求 369 的表达系统, 其中至少一个转录调控元件序列含有 SEQIDNO: 6 的序列。
389. 权利要求 369 的表达系统, 其中一种或多种所述转录调控元件为启动子。
390. 权利要求 389 的表达系统, 其中一种或多种所述启动子为基于 CMV 的启动子、CAG 启动子、基于 SV40 的启动子、基于热休克蛋白的启动子、基于 mH1 的启动子、基于 hH1 的启动子、基于鸡  $\beta$ -肌动蛋白的启动子、基于 U6 的启动子、基于泛素 C 的启动子或基于 EF-1  $\alpha$  启动子。
391. 权利要求 369 的表达系统, 其中至少一个转录调控元件是组成型的。
392. 权利要求 391 的表达系统, 其中一个或多个所述组成型转录调控元件为启动子。

393. 权利要求 392 的表达系统, 其中一个或多个所述启动子为基于 CMV 的启动子、CAG 启动子、基于 SV40 的启动子、基于热休克蛋白的启动子、基于 mH1 的启动子、基于 hH1 的启动子、基于鸡  $\beta$ -肌动蛋白的启动子、基于 U6 的启动子、基于泛素 C 的启动子或基于 EF-1  $\alpha$  启动子。
394. 权利要求 377 的表达系统, 其中所述标记编码序列和所述目的基因与不同的转录调控元件有效连接。
395. 权利要求 394 的表达系统, 其中所述与所述标记编码序列有效连接的转录调控元件强于所述与所述目的基因有效连接的转录调控元件。
396. 权利要求 377 的表达系统, 其中所述标记编码序列和所述目的基因的表达方向相反。
397. 权利要求 396 的表达系统, 其中所述标记编码序列和所述目的基因与两种不同的转录调控元件有效连接。
398. 权利要求 397 的表达系统, 其中至少一个所述转录调控元件是可调节的。
399. 权利要求 396 的表达系统, 其中所述标记编码序列和所述目的基因与一单转录调控元件有效连接。
400. 权利要求 399 的表达系统, 其中所述单转录调控元件是可调节的。
401. 权利要求 399 的表达系统, 其中所述单转录调控元件为双向启动子。
402. 一种制备病毒样颗粒的方法, 包括:
- a. 将权利要求 208、247 或 286 所述的包装核酸构建体引入细胞中;  
并
  - b. 将一种包膜构建体引入所述细胞中, 所述包膜构建体含有一种编码包膜糖蛋白的核酸序列, 其中所述编码包膜糖蛋白的核酸序列与至少一个转录调控元件有效连接; 并
  - c. 使所述细胞处于可形成病毒样颗粒的条件下。
403. 一种基因转移方法, 包括:
- a. 将权利要求 208、247 或 286 所述的包装核酸构建体引入细胞中;  
并
  - b. 将一种包膜构建体引入所述细胞中, 所述包膜构建体含有一种编码包膜糖蛋白的核酸序列, 其中所述编码包膜糖蛋白的核酸序

- 列与至少一个转录调控元件有效连接;
- c. 将一种含有一种或多种目的基因的基因转移构建体引入细胞; 并
- d. 使所述细胞处于可形成病毒样颗粒的条件下。
404. 权利要求 403 的方法, 其中所述基因转移构建体还包括一种编码标记的核酸序列, 其中所述目的基因和编码标记的核酸序列与至少一个转录调控元件有效连接。
405. 权利要求 404 的方法, 其中所述基因转移构建体还包括一个位于所述目的基因 3'端的土拨鼠肝炎病毒转录后调节元件。
406. 权利要求 404 的方法, 其中所述基因转移构建体的转录调控元件为启动子。
407. 权利要求 404 的方法, 其中所述转录调控元件是可调节的。
408. 权利要求 404 的方法, 其中所述基因转移构建体还包括一种位于所述目的基因和所述标记编码序列之间的元件, 其中所述元件可使所述目的基因和所述标记编码序列产生差异表达。
409. 权利要求 408 的方法, 其中所述位于目的基因和标记编码序列之间的元件为内部核糖体进入位点或内部核糖体进入位点样元件。
410. 权利要求 409 的方法, 其中所述内部核糖体进入位点或内部核糖体进入位点样元件为天然存在的。
411. 权利要求 409 的方法, 其中所述内部核糖体进入位点或内部核糖体进入位点样元件为非天然存在的。
412. 权利要求 409 的方法, 其中所述内部核糖体进入位点为 EMC-病毒 IRES 或 HCV-病毒 IRES。
413. 权利要求 409 的方法, 其中所述内部核糖体进入位点样元件为天然存在的。
414. 权利要求 409 的基因转移构建体, 其中所述内部核糖体进入位点或内部核糖体进入位点样元件为突变的。
415. 权利要求 403 的方法, 其中所述基因转移构建体包括至少一个转录调控元件。
416. 权利要求 415 的方法, 其中所述转录调控元件是可调节的。
417. 权利要求 415 的方法, 其中至少一个转录调控元件可由四环素或多西环素调节。
418. 权利要求 415 的方法, 其中至少一个转录调控元件含有至少一个

- tet 操纵子序列。
419. 权利要求 415 的方法，其中至少一个转录调控元件序列含有 SEQIDNO: 6 的序列。
420. 权利要求 415 的方法，其中一个或多个所述转录调控元件为启动子。
421. 权利要求 420 的方法，其中一个或多个所述启动子为基于 CMV 的启动子、CAG 启动子、基于 SV40 的启动子、基于热休克蛋白的启动子、基于 mH1 的启动子、基于 hH1 的启动子、基于鸡  $\beta$ -肌动蛋白的启动子、基于 U6 的启动子、基于泛素 C 的启动子或基于 EF-1 $\alpha$  启动子。
422. 权利要求 404 的方法，其中至少一个转录调控元件是组成型的。
423. 权利要求 422 的方法，其中一个或多个所述组成型转录调控元件为启动子。
424. 权利要求 423 的方法，其中一个或多个所述启动子为基于 CMV 的启动子、CAG 启动子、基于 SV40 的启动子、基于热休克蛋白的启动子、基于 mH1 的启动子、基于 hH1 的启动子、基于鸡  $\beta$ -肌动蛋白的启动子、基于 U6 的启动子、基于泛素 C 的启动子或基于 EF-1 $\alpha$  启动子。
425. 权利要求 404 的方法，其中所述标记编码序列和所述目的基因与不同的转录调控元件有效连接。
426. 权利要求 425 的方法，其中所述与所述目的基因有效连接的转录调控元件强于所述与所述编码标记的核酸序列有效连接的转录调控元件。
427. 权利要求 404 的方法，其中所述标记编码序列和所述目的基因的表达方向相反。
428. 权利要求 427 的方法，其中所述标记编码序列和所述目的基因与两种不同的转录调控元件有效连接。
429. 权利要求 428 的方法，其中至少一个所述转录调控元件是可调节的。
430. 权利要求 427 的方法，其中所述标记编码序列和所述目的基因与一单转录调控元件有效连接。
431. 权利要求 430 的方法，其中所述单转录调控元件是可调节的。
432. 权利要求 427 的方法，其中所述单转录调控元件为双向启动子。
433. 权利要求 430 的方法，其中所述双向启动子是可调节的。

434. 权利要求 427 的方法, 还包括将一种编码核定位信号的构建体引入所述细胞中, 所述构建体包括一种编码核定位信号的核酸序列, 所述核定位信号与一种编码四环素反式作用子的核酸序列有效连接。
435. 权利要求 434 的方法, 其中所述编码核定位信号的构建体还包括一种转录调控元件。
436. 权利要求 435 的方法, 其中所述转录调控元件为启动子或增强子。
437. 权利要求 434 的方法, 其中所述编码核定位信号的序列的侧翼接有至少一个接头序列。
438. 权利要求 437 的方法, 其中所述接头序列编码 SEQIDNO: 15。
439. 权利要求 434 的方法, 其中所述编码核定位信号的构建体从 5'端至 3'端依次包括巨细胞病毒启动子、编码 SEQIDNO: 15 的核酸序列、编码核定位信号的核酸序列、编码 SEQIDNO: 15 的核酸序列和编码四环素反式作用子的核酸序列。
440. 一种含有外源目的基因的细胞, 其中按照权利要求 403 所述方法将所述目的基因转移到所述细胞内。
441. 一种制备重组蛋白的方法, 包括在适合于使靶细胞表达所述重组蛋白的条件下, 使所述靶细胞与按照权利要求 402 所述方法制备的病毒颗粒接触。
442. 一种在受试者体内诱导免疫应答的方法, 包括将按照权利要求 441 所述方法制备的重组蛋白引入所述受试者体内, 所述重组蛋白的量足以诱导免疫应答。
443. 权利要求 442 的方法, 其中所述靶细胞产生均一糖基化模式的糖蛋白。
444. 权利要求 443 的方法, 其中与对照细胞相比, 所述靶细胞具有降低的 GnTI 活性。
445. 一种使用选定的蛋白处理受试者的方法, 包括给予所述受试者按照权利要求 443 所述方法制备的蛋白。
446. 一种筛选可调整病毒颗粒形成的药剂的方法, 包括
- a. 将一种含有第一和第二核酸序列的包装核酸构建体引入细胞中, 其中所述第一核酸序列编码 Gag 多聚蛋白, 其中所述第二核酸序列编码 Gag-Pro-Pol 多聚蛋白, 其中所述第一和第二核酸序列包括一个或多个可减少移码或翻译连读的突变, 其中所述第

- 一和第二核酸序列可由同一核苷酸序列的不同编码区表达，并且其中所述第一和第二核酸序列与所要筛选的药剂有效连接；
- b. 将一种包膜构建体引入所述细胞中，所述包膜构建体含有编码包膜糖蛋白的第三核酸序列，其中所述第三核酸序列与至少一个转录调控元件有效连接；
- c. 在适合于形成病毒颗粒的条件下培养所述细胞；
- d. 检测所述病毒颗粒；
- e. 与对照培养物相比，所要筛选的培养物中病毒颗粒数量的增加或减少表明所述药剂可调整病毒颗粒的形成。
447. 一种制备转基因动物的方法，包括：
- a. 将按照权利要求 402 所述方法制备的病毒颗粒引入合子中；
- b. 允许所述合子发育至足月；
- c. 获得一种其基因组中含有能够表达所述目的基因的核酸构建体的动物；
- d. 将所述动物与非转基因动物杂交以获得 F1 代；并
- e. 选择一种其基因组中含有所述能够表达目的基因的核酸构建体的动物，其中所述动物可表达选定的目的基因。
448. 一种按照权利要求 447 所述方法制备的转基因动物。
449. 一种包装系统，包括：
- a. 第一核酸构建体，包括编码 Gag 多聚蛋白的第一突变核酸，其中所述第一突变核酸与一种转录调控元件有效连接；和
- b. 第二核酸构建体，包括编码 Gag-Pol 多聚蛋白的第二突变核酸，其中所述第二突变核酸与一种转录调控元件有效连接；
- c. 其中所述第一和第二核酸构建体中的突变会使得所述 Gag 和 Gag-Pol 多聚蛋白的表达比例可形成病毒颗粒。
450. 权利要求 449 的包装系统，其中所述 Gag 和 Gag-Pol 多聚蛋白的表达比例为约 99:1 至约 80:20。
451. 权利要求 449 的包转系统，其中所述第一突变核酸与最小 CMV 启动子有效连接，并且所述第二突变核酸与所述热休克蛋白启动子有效连接。
452. 一种鉴别可结合目的抗原的抗体的方法，包括：
- a. 使疑似含有可结合目的抗原的抗体的样本与可表达所述目的抗原

- 的靶细胞接触，  
其中所述靶细胞含有一种或多种权利要求 1、55、208、247 和  
286 所述的核酸构建体；并
- b. 确定样本中的抗体是否与所述由靶细胞表达的目的抗原结合，由  
此将所述结合目的抗原的抗体鉴别为可结合目的抗原的抗体。

## 与使用病毒载体控制的基因表达相关的组合物和方法

### 相关专利申请的交互引用

本申请要求 2005 年 12 月 16 日提交的流水号为 60/751,407 的美国临时申请和 2005 年 12 月 16 日提交的流水号为 60/751,117 的美国临时申请的权益。2005 年 12 月 16 日提交的流水号为 60/751,407 的美国临时申请和 2005 年 12 月 16 日提交的流水号为 60/751,117 的美国临时申请的全文均通过引用的方式纳入本文。

### 致谢

本发明在美国国立变态反应和传染病研究所和国立卫生研究所的编号为 R01 AI47717 和 R01 AI48852 的政府拨款资助下完成。政府对本发明具有某些权利。

### 背景技术

经常需要在细胞或活体中转移和调控目的序列的表达，无论受试者是培养物中的细胞，还是活体例如动物模型或者是需要接受治疗基因的人类受试者。当使用基于 HIV 的慢病毒载体作为转移和/或表达目的序列的模式时，就会产生关于其使用安全性的忧虑，因为该病毒是 AIDS 的病原。其他的忧虑包括细胞的癌基因插入激活的可能性，该载体或构建体成功地和有效地与核糖体关联的能力，以及该载体或构建体成功地传导入核转运信号的能力。迄今为止，尚未产生可安全有效地用于在哺乳动物宿主或细胞内转移和/或表达目的序列、并且可以提供可诱导和可逆地表达所述被转移的目的基因或序列这两种重要能力的慢病毒载体。

慢病毒是复杂的反转录病毒，这些慢病毒——基于其较高的复杂度——可以整合到非增殖性细胞的基因组中并调整其生命周期，同感染

潜伏期中的情况一样。这些病毒包括 HIV-1、HIV-2 和 SIV。同其他的反转录病毒一样，慢病毒具有 gag、pol 和 env 基因，所述基因侧翼接有两个长末端重复 (LTR) 序列。每个上述基因均编码多种蛋白，初始时均表达为一种前体多聚蛋白。所述 gag 基因编码内部结构 (基质衣壳和核衣壳) 蛋白。pol 基因编码 RNA 指导的 DNA 聚合酶 (反转录酶、整合酶和蛋白酶)。env 基因编码病毒包膜糖蛋白并且还含有负责病毒 RNA 的出核转运的顺式作用元件 (RRE)。所述 5' 和 3' LTR 可起到促进病毒体 RNA 的转录和多腺苷酸化的作用。所述 LTR 含有病毒复制所必需的全部其他顺式作用序列。邻近所述 5' LTR 的是反转录所述基因组所必需的序列 (tRNA 引物结合位点) 和将病毒 RNA 有效地壳体化为颗粒所必需的序列 (Psi 位点)。如果将壳体化 (即将反转录病毒 RNA 包装为感染性病毒体) 所必需的序列从所述病毒基因组上缺失，则会产生可阻止基因组 RNA 壳体化的顺式缺陷。然而，所得突变体仍旧能够指导所有病毒体蛋白的合成。例如 Field's Virology (Raven Publishers), eds. B. N. Fields et al., (1996) 中提供了关于慢病毒 (例如 HIV) 的全面综述。

尽管慢病毒载体可用于许多用途，但是通过遗传重组产生有复制能力的反转录病毒 (RCR) 的可能性导致了关于安全性的忧虑。研究人员试图克服该问题的一个途径是构建基于 HIV 的包装系统 (反式-慢病毒)，所述包装系统将 gag/gag-pol 分为两部分：一部分表达 gag/gag-pro，而另一部分以病毒蛋白 R (Vpr) 的融合伴侣的形式表达反转录酶和整合酶。然而，该方法被发现具有许多缺陷，如生产感染性病毒载体颗粒的效率远低于理想的效率。

其他用于生产不产生重组反转录病毒的、有效的反转录病毒包装细胞系 (特别是慢病毒包装细胞系) 的方法和系统将具有很高价值。

## 发明内容

本发明通过将基因转移构建体或其他表达系统与基因调节系统结合起来，以便有效地将基因递送到细胞和活体中并控制其表达，提供了一种针对上述问题的解决方案。因此，本发明可 (例如通过高效慢病毒载体递送系统) 保证有效转染宿主，并通过单纯地将调节物 (例如，抗生素如四环素) 给予携带被转移的目的序列的宿主，来灵敏地控制所述被

转移的目的序列的表达时机和水平。本发明所提供的额外好处是可在数种物种(例如啮齿动物、灵长类动物和犬类)宿主的非分裂细胞中实现该有效的转染和调节。

本发明提供了基因转移构建体和表达系统。本发明的基因转移构建体和表达系统可以是慢病毒载体。上述构建体包括多种组件,所述组件可使得所述构建体可安全有效地将目的序列转移到哺乳动物宿主细胞中;本发明还通过给予细胞或受试者合适的调节物,提供了在所述哺乳动物宿主细胞内显著调控所述被转移目的序列的表达的极其重要的能力,所述细胞或受试者含有所述可诱导并且可逆转的基因转移构建体。本发明的基因转移构建体可包括下列组件的一种或多种:自我失活的5'LTR、调节子应答性启动子、入核转运信号、与编码调节子应答性受体的核酸有效连接的启动子、RNA稳定元件或自我失活的3'LTR。所公开的基因转移构建体可用于包装DNA并将其递送至分裂和非分裂细胞中。本发明公开的包装和转移构建体可以彼此结合使用,也可以与本领域已知的其他包装和基因转移构建体、系统和方法以及本文所公开的系统和方法结合使用。

本发明还提供了特异性基因转移构建体和使用所述构建体在靶细胞中诱导性地并且可逆转地表达目的序列的方法。本发明还提供了使用所述公开的基因转移构建体作为用于处理哺乳动物受试者的表达系统的活体外方法。本发明还提供了制备用于表达目的序列的动物模型的方法。此外,本发明提供了并入了或含有本文所公开的基因转移构建体或表达系统的细胞。因此,所公开的基因转移构建体有助于构建稳定的、可诱导/可逆转的细胞系,因为所述假型慢病毒载体可转导许多不适用标准DNA转染技术的细胞类型。

本发明还提供了双向启动子,所述双向启动子可在相反的两个方向上驱动至少两个单独的序列表达。所公开的双向启动子还可与本文公开的包装和基因转移构建体共同使用。

本发明还提供了含有本文所述各种基因转移构建体的细胞系。

本发明还公开了基因转移构建体,其中所述构建体能够产生无复制能力的重组体。

本发明还提供了含有本文所述的各种基因转移构建体的表达系统。本发明还提供了含有本文所述的基因转移构建体或表达系统的细胞系,

以及按照本文所述方法制备的细胞。

本发明还提供了选择性调节目基因表达的方法，包括将本文所公开的基因转移构建体引入靶细胞中。

本发明还提供了制备重组蛋白、抗体和转基因动物的方法。

本发明还提供了含有编码 Gag 和 Gag-Pro-Pol 多聚蛋白的核酸序列的包装构建体。上述构建体都是安全的，而且与在本发明之前可获得的构建体相比，本发明的构建体具有提高的包装效率。本发明还提供了含有编码 Gag 和 Gag-Pro-Pol 多聚蛋白的核酸序列的包装构建体，所述构建体还包括一个或多个位于所述编码 Gag 和 Gag-Pro-Pol 多聚蛋白的核酸序列内的突变，所述突变可减少合成 Gag-Pro 和 Gag-Pro-Pol 多聚蛋白所需的移码或翻译连读 (read-through)。

本发明还提供了一种包装构建体，包括第一和第二核酸序列，其中所述第一核酸序列编码 Gag 多聚蛋白，其中所述第二核酸序列编码 Gag-Pro 多聚蛋白，其中所述第一和第二核酸序列包括一个或多个可减少移码或翻译连读的突变，其中所述第一和第二核酸序列可由同一核苷酸序列的不同编码区表达，并且其中所述第一和第二核酸序列与至少一个转录调控元件有效连接。

本发明还提供了包括第一、第二和第三核酸序列的包装构建体，其中所述第一核酸序列编码 Gag 多聚蛋白，其中所述第二核酸序列编码 Gag-Pro 多聚蛋白，其中所述第三核酸序列编码 Vpr-反转录酶-整合酶蛋白。

本发明还提供了包装构建体，其中 Gag 和 Gag-Pol 为反式 (*in trans*)，其中所述编码 Gag 多聚蛋白的核酸序列和编码 Gag-Pro 多聚蛋白的核酸序列含有一个或多个可减少移码或翻译连读的突变，并且所述编码 Gag 多聚蛋白的核酸序列和编码 Gag-Pro 多聚蛋白的核酸序列与至少一个转录调控元件有效连接。

本发明还提供了含有各种本文所述包装构建体的细胞系、包装系统和表达系统。本发明还提供了含有本文所述表达系统的细胞系。

任选地，本文所述包装构建体能够产生无复制能力的重组体。

本发明还提供了制备病毒样颗粒的方法。

本发明还提供了制备和使用本文所述细胞系、包装构建体、基因转移构建体、包装系统和表达系统的方法。

本发明还提供了筛选可调节病毒颗粒形成的药剂 (agent) 的方法。

本发明还提供了含有本文所公开的基因转移构建体的疫苗, 以及在受试者体内诱导免疫应答的方法, 包括给予受试者本文所公开的疫苗。

因此, 本发明成功地将有效的目的序列送递系统与紧密调节的目的序列表达系统结合起来, 并且具有在目的序列的送递和表达技术上的显著进步。

## 附图说明

附图图示说明了下文所述的几个方面, 所述附图被编入本说明书中并构成其一部分。

图 1 示出了含有点突变的移码所需的突变 gag 序列与野生型 gag 序列之间的比较。

图 2 示出了含有点突变的移码所需的突变 gag-pol 序列与野生型 gag-pol 序列之间的比较。

图 3 示出了移码所需的 HIV gag 和 HIV gag-pol 中的环结构。

图 4 示出了移码所需的 HIV gag 和 HIV gag-pol 中环结构的改变的序列, 所述改变的序列导致移码所需的环结构被破坏。

图 5 示出了在给小鼠喂食 DOX 之前和 18 天后对转基因 CAG 建立者 (founder) 的血细胞中 GFP 表达的 FACS 分析的结果。

图 6 示出了转基因 CAG 建立者的血细胞中 GFP 表达的诱导动力学。

图 7 示出了人和小鼠 H1 启动子均能表达设计用于靶向 eGFP 的 shRNA, 这本身又可以有效地沉默 HeLa 细胞中的 eGFP 表达。

图 8 示出了人和小鼠 H1 启动子均能表达设计用于靶向 eGFP 的 shRNA, 这本身又可以有效地沉默人 T 细胞中的 eGFP 表达。

图 9 示出了含有靶向小鼠 CXCR4 的 shRNA 的诱导型单慢病毒载体可诱导性地降低小鼠内源 CXCR4 蛋白的表达。

图 10 示出了多个拷贝的整合的诱导型单慢病毒载体可诱发高水平的基因沉默, 所述慢病毒载体含有靶向小鼠 CXCR4 的 shRNA。

图 11 示出了利用 DOX 在转基因小鼠血细胞内诱导 siRNA 表达以减少 GFP。

图 12 示出了利用 DOX 在转基因小鼠血细胞内诱导 siRNA 表达以减

少 GFP。

图 13 示出了在给小鼠喂食 DOX 之前,非转基因小鼠以及转基因 CAG-建立者 F1-6#和 F1-9#的血细胞中 GFP 的表达水平。

图 14 示出了在给小鼠喂食 DOX 后第 10、17、27 天时,非转基因小鼠以及转基因 CAG-建立者 F1-4#和 F1-11#的血细胞中 GFP 的表达水平。

图 15 示出了本文所公开的基因转移构建体的实例。

图 16 示出了 A)基于 HIV 的含有 hCCR1-m 的慢病毒载体,所述慢病毒载体可用于产生诱导性地并可逆地表达人 CCR1 基因的细胞系。B)示出了 CCR1-m 的 C 末端氨基酸序列。所述 CCR1 的终止密码子被突变并且被置换为 TEV 蛋白酶位点 (ENLYFQG)。在 TEV 和 10 个组氨酸之间插入 M2 标签以便分析所述蛋白 (CCR1-m) 的表达和纯化。所述 10 个 His 标签起到使用 Ni-NTA 柱来纯化 CCR1-m 的作用。

图 17 示出了 A)基于 HIV 的含有 hEP2R 的慢病毒载体,所述慢病毒载体可用于产生诱导性地并可逆地表达人 EP2 基因的细胞系。B)示出了 hEP2R-m 的 C 末端氨基酸序列。其终止密码子被突变并且被置换为 TEV 蛋白酶位点 (ENLYFQG)。在 TEV 和 10 个组氨酸之间插入 M2 标签以便分析所述蛋白 (hEP2R-m) 的表达和纯化。所述 10 个 His 标签起到使用 Ni-NTA 柱来纯化 hEP2R-m 的作用。

## 具体实施方式

在公开和描述本发明的化合物、组合物、制品、设备和/或方法之前,应该理解的是下文所述的各方面不限于特定的合成方法或特定的给药方法,下文所述的各方面本身当然是可以变化的。还应该理解的是本文所用的术语仅是为了描述具体的方面而并非意在进行限制。

必须注意到除非上下文中另外明确指出,本说明书和所附权利要求中使用的单数形式“一”、“一个”和“该”包括复数指代对象。

本说明书全文中所用的“受试者”是指个体。因此,所述“受试者”可以包括驯养的动物(例如猫、狗等)、家畜(例如牛、马、猪、绵羊、山羊等)、实验用动物(例如小鼠、兔、大鼠、豚鼠等)和禽类。在一个方面中,所述受试者为哺乳动物,例如灵长类或人。

“任选的”或“任选地”意为随后描述的事件或情况可能发生,也

可能不发生,并且这样的描述包括所述事件或情况发生的情况和所述事件或情况不发生的情况。例如,短语“所述组合物任选地可以包括一种组合”是指该组合物可以包括不同分子的一种组合,也可以不包括这一组合,使得该描述同时包括存在所述组合和不存在所述组合的情况(即组合的各个成员)。

短语“包装细胞系”或“包装细胞”是指含有生产病毒颗粒或病毒样颗粒所需编码序列的细胞(通常是哺乳动物细胞系),所述细胞在包装病毒RNA和生产有复制能力的辅助病毒的能力上存在缺陷。当所述细胞内提供了包装功能时,该包装细胞系或包装细胞会产生重组的反转录病毒,因此成为“反转录病毒生产细胞系”或“反转录病毒生产细胞”。

术语“反转录病毒”是指任何已知的反转录病毒(例如c型反转录病毒,如莫洛尼鼠白血病毒(MoMuLV)、哈维鼠肉瘤病毒(HaMuSV)、鼠乳腺肿瘤病毒(MuMTV)、长臂猿白血病毒(GaLV)、猫白血病毒(FLV)和劳斯肉瘤病毒(RSV))。本发明的“反转录病毒”还包括人T细胞白血病毒(HTLV-1和HTLV-2)和反转录病毒的慢病毒家族,例如人免疫缺陷病毒(HIV-1和HIV-2)、猿免疫缺陷病毒(SIV)、猫免疫缺陷病毒(FIV)、马免疫缺陷病毒(EIV)和其他反转录病毒类别。

术语“Gag多聚蛋白”或“Gag蛋白”、“Pro多聚蛋白”或“Pro蛋白”以及“Pol多聚蛋白”或“Pol蛋白”是指由反转录病毒gag、pro和pol基因编码的多种蛋白,所述蛋白通常被表达为单一的前体“多聚蛋白”。例如,HIV gag编码p17、p24、p7和p6等蛋白。HIV pro编码病毒蛋白酶。HIV pol编码蛋白酶(PR)、反转录酶(RT)和整合酶(IN)等蛋白。本文所用术语“多聚蛋白”应包括gag、pro或pol多聚蛋白的全部或其任何部分。

术语“载体”或“构建体”是指能够将与所述载体序列相连接的另一核酸转运到细胞内的核酸序列。术语“表达载体”包括以适合于被细胞表达的形式(例如与转录调控元件相连)含有基因构建体的任何载体(例如质粒、粘粒或噬菌体染色体)。“质粒”和“载体”可互换使用,因为质粒是载体的一种常用形式。此外,本发明意在包括其他可发挥同样功能的载体。

术语“目的序列”或“目的基因”可以指这样一种核酸序列(例如治疗基因),所述核酸序列对于被引入该序列的细胞来说部分或全部异

源，即外来的。

术语“目的序列”或“目的基因”还可以指这样一种核酸序列，所述核酸序列对于被引入该序列的细胞的内源基因来说部分或全部同源，但是所述核酸序列被设计用于插入到该细胞的基因组中以改变该基因组（例如该序列被插入到与天然基因所处位置不同的位置上或者其插入导致“敲除”）。例如，目的序列可以是 cDNA、DNA 或 mRNA。

术语“目的序列”或“目的基因”还可以指这样一种核酸序列，所述核酸序列与被引入该序列的细胞的内源基因部分或完全互补。例如，所述目的序列可以是微小 RNA、shRNA 或 siRNA。

“目的序列”或“目的基因”还可以包括一种或多种转录调节序列以及优化表达选定核酸所需的任何其他核酸，例如内含子。“目的蛋白”是指由目的序列或目的基因表达的肽或多肽序列（例如治疗蛋白）。

“基因转移构建体”是指这样一种核酸序列，所述核酸序列通常与其他慢病毒或反式慢病毒载体系统的载体共同使用来生产病毒颗粒，例如，使得所述病毒颗粒可转导目的靶细胞。

术语“有效连接”是指一种核酸与另一种核酸序列之间的功能关系。启动子、增强子、转录和翻译终止位点以及其他信号序列是与其他序列有效连接的核酸序列的实例。例如，DNA 与转录调控元件之间的有效连接是指所述 DNA 与启动子之间的物理和功能关系，使得可利用 RNA 聚合酶、由所述启动子启动这类 DNA 的转录，所述 RNA 聚合酶特异性识别、结合并转录所述 DNA。

术语“转化”和“转染”是指将核酸（例如表达载体）引入接收者细胞，包括将核酸引入所述细胞的染色体 DNA 中。

术语“RNA 外排元件 (export element)”是指顺式作用转录后调节元件，所述元件可调节 RNA 转录物由细胞的细胞核至细胞质的转运。RNA 外排元件的实例包括但不限于，人免疫缺陷病毒 (HIV) rev 应答元件 (RRE) (参见例如，Cullen et al. (1991) *J. Virol.* 65:1053; 和 Cullen et al. (1991) *Cell* 58:423-426) 和乙型肝炎病毒转录后调节元件 (PRE) (参见例如，Huang et al. (1995) *Molec. and Cell. Biol.* 15(7):3864-3869; Huang et al. (1994) *J. Virol.* 68(5):3193-3199; Huang et al. (1993) *Molec. and Cell. Biol.* 13(12):7476-7486)，以及美国专利 NO. 5,744,326，上述所有文献关于 RNA 外排元件的内容通过引用的方

式全部纳入本文)。通常,所述RNA外排元件位于基因的3'UTR内,并且可以一个或多个拷贝插入。RNA外排元件可以插入至产生本发明的包装细胞系的任何或全部的独立载体中。

本文中范围可以表示为从“大约”一个具体的数值,和/或至“大约”另一个具体的数值。当表示为这种范围时,另一个实施方案包括从所述的一个具体的数值和/或至所述的另一个具体的数值。类似地,当通过使用前面的“大约”将数值表示为近似值时,应当理解的是,该具体的数值构成另一个实施方案。应当进一步理解的是,每个范围的端点在与另一个端点相关和独立于另一个端点时都是有意义的。还应当理解的是本文公开了许多数值,并且除所述数值本身外每一个数值在此还以“大约”该具体数值公开。例如,如果公开了数值“10”,那么也公开了“大约10”。还应当理解的是,当公开一个数值时,也公开了“小于或等于”该数值、“大于或等于该数值”以及数值间的可能范围,这正如本领域技术人员所恰当理解的。例如,如果公开了数值“10”,那么也公开了“小于或等于10”以及“大于或等于10”。还应当理解的是,在本申请全文中,数据以多种不同格式给出,并且该数据代表了端点和起点以及这些数据点任何组合的范围。例如,如果公开了具体数据点“10”和具体数据点“15”,那么应该理解的是,大于、大于或等于、小于、小于或等于以及等于10和15以及10和15之间的数值也被认为已经公开。

“病毒样颗粒”或“病毒颗粒”是指蛋白性壳体样病毒体,所述病毒体是通过在宿主细胞内表达至少一种下列病毒基因来生产的:gag、pro、rt、in和env。所产生的颗粒优选地含有所述基因转移构建体的mRNA等价物,并且对所转导的给定细胞类型具有感染性,或者可被加工成对所转导的给定细胞类型具有感染性。

#### 组合物

1型人免疫缺陷病毒(HIV-1)的gag和pol基因最初被表达为前体多聚蛋白Gag和Gag-Pro-Pol。在出芽过程中或之后,利用病毒蛋白酶(PR)将上述前体加工成为其成熟的产物。55 kDa的Gag前体生成基质(MA)、壳体(CA)、间隔肽p2、核壳体(NC)、间隔肽p1和p6。160 kDa的Gag-Pro-Pol多聚蛋白生成MA、CA、p2、NC、p6、PR、反转录酶(RT)和整合酶(IN)。Gag和Gag-Pro-Pol多聚蛋白由同一mRNA编码但合成速

率不同。罕见的核糖体移码事件会使 Gag 和 Gag-Pro-Pol 多聚蛋白的生成比例约为 20:1。维持该比例对于病毒颗粒的形成和感染性至关重要。

单独在胞内表达 Gag 已足以生产病毒样颗粒 (VLP)。此外,在装配过程中 Gag 和病毒基因组 RNA 相互作用起到重要作用,且基因组 RNA 的包装和二聚主要通过 RNA-Gag 相互作用发生。Gag 的 NC 结构域与病毒 RNA 结合并且已经被证明可促进所述 RNA 包装过程和二聚化过程。基因组 RNA 和 HIV-1 Gag 之间的初始相互作用似乎是通过 Gag 前体中的 NC 序列发生的,因为具有缺陷型病毒 PR 的 HIV-1 仍旧可以包装 RNA。此外,对野生型 (WT) 病毒体和 PR-缺陷型 (PR-) 病毒体的分析显示 HIV-1 中基因组 RNA 的二聚化开始于蛋白酶解加工之前,表明 Gag 和 Gag-Pro-Pol 前体蛋白在未进行蛋白加工的情况下亦可支持 RNA 二聚化。

与其他反转录病毒不同,慢病毒除 gag、pol 和 env 外还具有数种具有调节和结构性功能的“附属”基因。具体而言,HIV 具有至少六种这类基因,包括 Vif、Vpr、Tat、Rev、Vpu 和 Nef。密切相关的 HIV-2 不编码 Vpu,但可编码另一种不存在于 HIV-1 中的无关蛋白 Vpx。

HIV-1 Vpr 基因编码一 14 kD 的蛋白(96 个氨基酸)(Myers et al. (1993) *Human Retroviruses and AIDS*, Los Alamos National Laboratory, N.M.)。所述 Vpr 开放读码框还存在于大多数 HIV-2 和 SIV 病毒中。HIV-2 Vpr 和 Vpx 之间的氨基酸比较显示存在高度同源性区域,表明 Vpx 可能是由 Vpr 基因的复制产生的。成熟病毒颗粒中存在多个拷贝的 Vpr 和 Vpx,并且已经证明 Vpr 和 Vpx 可结合 p6 蛋白,所述 p6 蛋白是与病毒装配有关的 gag 编码的前体多聚蛋白的一部分(WO 96/07741; WO 96/32494)。因此,Vpr 和 Vpx 通过与 p6 之间的相互作用被引入到病毒颗粒中(Lavallee et al. (1994) *J. Virol.* 68: 1926-1934;和 Wu et al. (1994) *J. Virol.* 68: 6161)。还已经证明 Vpr 具体是与 p6 的羧基末端区域相关联。已经证明 Vpr 和 Vpx——以相对于 HIV 基因组的反式表达方式表达——可被用于将异源蛋白靶向 HIV 病毒(WO 96/07741; WO 96/32494)。下列文件给出了对 Vpr 和 Vpx 的结构和功能的描述,包括这些蛋白及其结合结构域的全长核苷酸和氨基酸序列,所述文件包括: WO 96/07741 和 Zhao et al. (1994) *J. Biol Chem.* 269(22): 1577 (Vpr); Mahalingham et al. (1995) *Virology* 207: 297 (Vpr); 和 Hu et al. (1989) *Virology* 173: 624 (Vpx)。其他与 Vpr 相关的参考文献包括,

例如 Kondo et al. (1995) *J. Virol* 69:2759; Lavallee et al. (1994) *J. Virol.* 68:1926; 和 Levy et al. (1993) *Cell* 72:541. 其他与 Vpx 相关的参考文献包括, 例如 Wu et al. (1994) *J. Virol.* 68:6161. 上述参考文献关于 Vpr 和 Vpx 结构和功能的教导全部通过引用的方式纳入本说明书。

这些反转录病毒整合酶 (IN) 蛋白可催化前病毒的整合并且是在体内维持所述感染状态所必需的。对这种关键性酶, 特别是它的蛋白质结构和所述催化整合反应的生物化学机理的认识已经取得了显著进步 (Brown, P. 1997. Integration, p. 161-204. *In* J. M. Coffin, S. H. Hughes 和 H. E. Varmus (ed.), *Retroviruses*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.; Dyda, F., A. B. Hickman, T. M. Jenkins, A. Engelman, R. Craigie 和 D. R. Davies. 1994. Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. *Science* 266:1981-1986; Katz, R. A., 和 A. M. Skalka. 1994. The retroviral enzymes. *Annu. Rev. Biochem.* 63:133-173. 所有上述参考文献关于 IN 蛋白结构和所述催化整合反应的生物化学机理的教导全部均通过引用的方式纳入本说明书)。HIV-1 IN 被表达为一个较大的 160-kDa Gag-Pol 前体多聚蛋白 (Pr160<sup>GagPol</sup>) 的一部分并且装配进所述病毒颗粒中, 所述 Gag-Pol 前体多聚蛋白含有其他 Gag (基质、壳体、核壳体和 p6) 和 Pol (蛋白酶、反转录酶 [RT] 和 IN) 组件。装配之后, 利用所述病毒蛋白酶对 Pr160<sup>GagPol</sup> 进行蛋白酶解加工以便释放所述各个 Gag 和 Pol 组件, 包括所述 32-kDa IN 蛋白。使用复制型病毒进行的 IN 功能研究已经证实, 除催化所述病毒 cDNA 的整合外, IN 还对病毒复制具有其他影响 (Gallay, P., S. Swingler, J. Song, F. Bushman 和 D. Trono. 1995. HIV nuclear import is governed by the phosphotyrosine-mediated binding of matrix to the core domain of integrase. *Cell* 83:569-576; Leavitt, A. D., G. Robles, N. Alesandro 和 H. E. Varmus. 1996. Human immunodeficiency virus type 1 integrase mutants retain in vitro integrase activity yet fail to integrate viral DNA efficiently during infection. *J. Virol.* 70:721-728; Masuda, T., V. Planelles, P. Krogstad 和 I. S. Y. Chen.

1995. Genetic analysis of human immunodeficiency virus type 1 integrase and the U3 *att* site: unusual phenotype of mutants in the zinc finger-like domain. *J. Virol.* 69:6687-6696。所有上述参考文献关于 IN 蛋白结构和所述催化整合反应的生物化学机理的教导全部均通过引用的方式纳入本说明书)。在使用前病毒克隆进行的研究中,已经证明 IN 基因突变可在多个水平上影响病毒复制。IN 基因中的突变可以影响 Gag-Pol 前体蛋白并且改变装配、成熟和其他后续病毒事件。IN 基因突变还可以影响成熟 IN 蛋白以及其在病毒颗粒和核蛋白整合前复合体内的组构。因此,这类突变是多效的并且可以通过多种机制在病毒生活周期的不同阶段改变病毒复制。

反转录可被 RT 催化,并且尽管反转录可以在体外使用重组 RT、模板和引物来进行,但是在体内该过程更加复杂。在病毒复制的过程中,病毒 cDNA 的完全合成并不只是简单地将不同的蛋白和核酸放在一起;相反,病毒复制是一个复杂的多步过程,包括许多过渡结构。在被感染的细胞中,反转录在核酸-蛋白(核蛋白)复合体中进行,所述复合体含有其他病毒因子和细胞因子。此外,所述病毒 cDNA 的合成在很大程度上取决于在反转录之前进行的多个分子事件的正确执行。

本发明公开了包装和基因转移构建体。用于生产重组反转录病毒载体和用于转化包装细胞系的方案在本领域中是已知的(*Current Protocols in Molecular Biology*, Ausubel, F. M. et al. (eds.) Greene Publishing Associates, (1989), Sections 9.10-9.14 和其他标准的实验室手册; Eglitis, et al. (1985) *Science* 230:1395-1398; Danos and Mulligan (1988) *Proc. Natl. Acad. Sci. USA* 85:6460-6464; Wilson et al. (1988) *Proc. Natl. Acad. Sci. USA* 85:3014-3018; Armentano et al. (1990) *Proc. Natl. Acad. Sci. USA* 87:6141-6145; Huber et al. (1991) *Proc. Natl. Acad. Sci. USA* 88:8039-8043; Ferry et al. (1991) *Proc. Natl. Acad. Sci. USA* 88:8377-8381; Chowdhury et al. (1991) *Science* 254:1802-1805; van Beusechem et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:7640-7644; Kay et al. (1992) *Human Gene Therapy* 3:641-647; Dai et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:10892-10895; Hwu et al. (1993) *J. Immunol.* 150:4104-4115; U. S. Pat. No. 4,868,116; U. S. Pat. No. 4,980,286; PCT 申请 WO 89/07136;

PCT 申请 WO 89/02468; PCT 申请 WO 89/05345; 和 PCT 申请 WO 92/07573。所有上述参考文献关于生产重组反转录病毒构建体和载体以及将其用于转化细胞系的方案的教导全部均通过引用的方式纳入本说明书)。此外, 适合用于本发明的反转录病毒序列可从商业来源获得。例如, 购得的这类序列可以是反转录病毒质粒形式, 例如 pLJ、pZIP、pWE 和 pEM。本发明载体中所采用的合适的包装序列也可以是市售的, 包括例如质粒  $\psi$ Crip、 $\psi$ Cre、 $\psi$ 2 和  $\psi$ Am。因此, 当使用具体实施方案(例如具体的慢病毒载体)描述本发明时, 其他可用于本发明的反转录病毒载体也可以按照本文所述的指南进行制备。此外, 本文所公开的基因转移载体可以与本文所公开的包装和表达系统共同使用。

具体而言, 本发明公开了含有编码 Gag 和 Gag-Pro-Pol 蛋白的核酸序列的包装构建体。任选地, 所述包装构建体包括第一和第二核酸序列, 其中所述第一核酸序列编码 Gag 蛋白, 其中所述第二核酸序列编码 Gag-Pro-Pol 蛋白, 其中所述第一和第二核酸序列各包括一个或多个可减少移码或翻译连续的突变, 其中所述第一和第二核酸序列由同一核苷酸序列的不同编码区表达, 并且其中所述第一和第二核酸序列与至少一个转录调控元件有效连接。

本发明还公开了一种含有第一和第二核酸序列的包装构建体, 其中所述第一核酸序列编码 Gag 蛋白, 其中所述第二核酸序列编码 Gag-Pro 蛋白, 其中所述第一和第二核酸序列各包括一个或多个可减少移码或翻译连续的突变, 其中所述第一和第二核酸序列由同一核苷酸序列的不同编码区表达, 并且其中所述第一和第二核酸序列与至少一个转录调控元件有效连接。在该构建体中, 所述通常编码多聚 RT 和 IN 的核酸序列被移除或被突变, 使得所述 Pol 或 RT-IN 蛋白无法表达。移除或突变所述编码 Pol 蛋白的核酸序列还会降低通过遗传重组生成有复制能力的反转录病毒(RCR)的可能性。然后, RT 和 IN 可由单独的构建体(反式)表达。例如, 反转录酶和整合酶可以以病毒蛋白 R(Vpr)的融合伴侣的形式被表达。

本发明还公开了含有第一、第二和第三核酸序列的包装构建体, 其中所述第一核酸序列编码 Gag 蛋白, 其中所述第二核酸序列编码 Gag-Pro 蛋白, 并且其中所述第三核酸序列编码 Vpr-反转录酶-整合酶蛋白。此外, 所述第一和第二核酸序列可以包括一个或多个可减少移码

或翻译连读的突变，其中所述第一、第二以及第三核酸序列由同一核苷酸序列的不同编码区表达，并且其中所述第一、第二和第三核酸序列与至少一个转录调控元件有效连接。在该构建体中，所述能够编码 Pol 蛋白的核酸序列可以被移除或被突变，使得所述 Pol 蛋白无法表达。所述反转录酶和整合酶可由编码 Vpr-反转录酶-整合酶蛋白的核酸序列以反式提供。

本发明还公开了位于更下游的 IRES 或 IRES 样元件以便调控 Vpr-RT-IN。下文描述了 IRES 和 IRES 样元件。例如，本发明公开了一种包装构建体，所述包装构建体还包括一个位于所述第一或第二核酸序列与所述第三核酸序列之间的元件，其中所述第三核酸序列并不位于所述第一和第二核酸序列之间，并且其中所述元件可使所述第一或第二核酸序列与所述第三核酸序列之间产生差异表达。上述的实例包括内部核糖体进入位点或内部核糖体进入位点样元件。本发明公开了可用于该方法的 IRES 和 IRES 样元件。所述 IRES 可以是，例如 EMC-病毒 IRES、HCV-病毒 IRES 或不同来源的 IRES。其他可以使用的 IRES 实例包括但不限于 <http://ifr31w3.toulouse.inserm.fr/IRESdatabase/> 的 IRES 数据库中存在的 IRES。

本发明还公开了包装构建体，所述包装构建体还包括含有 rev 应答元件的核酸序列。

实际上，Gag 和 Gag-Pol 蛋白可由部分重叠的开放读码框编码。Gag 具有其自身的起始和终止密码子，而 HIV-1 Gag-Pol 前体的合成会引起移码事件，所述移码事件的发生频率为 Gag 翻译时移码事件的发生频率的约 5-10%。其他反转录病毒也使用类似的移码机制或连读抑制机制来调节 Gag-Pol 或 Gag-Pro 蛋白的表达。因此，在所有的反转录病毒复制过程中，胞内 Gag/Gag-Pol 比例都受到调节。已发现 HIV 移码位点(富含 AU 的七核苷酸序列)位于核壳体(NC)编码序列的 3'末端。在合成 Gag 的过程中，该位点和下游紧邻的茎结构会停住所述核糖体，使得所述核糖体滑回一个核苷酸以使得罕见(相对于 Gag)的 Gag-Pol 融合蛋白合成得以进行。

Gag 蛋白的多聚化会产生病毒颗粒，同时 Gag-Pol 前体蛋白的表达会确保在病毒装配过程中病毒酶被整合到病毒颗粒内。在细胞释放病毒体的过程中以及释放之后，Gag 前体蛋白被病毒蛋白酶(PR)切割成为成

熟蛋白：基质、壳体 (CA)、NC、p6 和两种间隔肽 p2 和 p1。切割 Gag-Pol 融合体以产生基质、CA、p2 和 NC，以及跨读框蛋白、PR、反转录酶 (RT) 和整合酶 (IN)。

已经报道了单独合成 Gag 前体蛋白已足以进行病毒样颗粒的装配和释放。将 Gag-Pol 或其成熟产物并入到病毒体中是实现感染性所必需的，因为它们可介导被感染细胞内的病毒 cDNA 的合成和整合。此外，利用 PR 切割所述前体蛋白是所述病毒体核的形态成熟和感染性病毒颗粒的形成所必需的。在装配过程中，病毒基因组 RNA 也被包装进病毒体内，所述包装过程由位于所述基因组 5'端附近的基因组 RNA 包装序列以及与 Gag 的 NC 结构域之间的相互作用驱动。

与其他反转录病毒类似，例如 HIV-1 具有二聚的 RNA 基因组。对 HIV-1 病毒 RNA 进行的体外二聚分析已经定位了 (mapped) 一种含 50-60 个核苷酸的序列 (称为二聚体起始序列)，该序列对于所述二聚 RNA 复合体的形成很重要。二聚体起始序列中的突变会妨碍基因组 RNA 的二聚化以及病毒体 RNA 的包装，并且导致产生无感染性的病毒颗粒。一般认为 RNA 的二聚化是 HIV-1 中 RNA 包装的先决条件，并且在其他反转录病毒中，基因组 RNA 的病毒体包装和 RNA 的二聚化也是相关的。来源于 PR 缺陷型 HIV-1 病毒体的 RNA 二聚体的热稳定性低于来源于野生型成熟 HIV-1 的二聚体。在莫洛尼鼠白血病病毒中也观察到了类似的现象。

尽管单独表达 Gag-Pol 前体并不足以生产感染性反转录病毒颗粒，但是 Gag/Gag-Pol 比例对于所述病毒复制周期和 RNA 二聚化的影响是一个关键性因素。已经证明在生产病毒体的细胞内，Gag/Gag-Pol 比例对于感染性病毒颗粒的形成和所述病毒体 RNA 二聚体的稳定性都很重要 (Xhila et al. *Journal of Virology*, February 2001, p. 1834-1841, Vol. 75, No. 4)。

本发明公开了包装系统，其中所述 Gag 和 Gag-Pol 蛋白的比例为约 99:1、98:2、97:3、96:4、95:5、94:6、93:7、92:8、91:9、90:10、89:11、88:12、87:13、86:14、85:15、84:16、83:17、82:18、81:19、80:20、79:21、78:22、77:23、76:24、75:25、74:26、73:27、72:28、71:29、70:30、69:31、68:32、67:33、66:34、65:35、64:36、63:37、62:38、61:39、60:40 或任一介于其间的比例。

此外，基于慢病毒的包装构建体可以任选地包括能够表达 Vpr-反转

录病酶-整合酶蛋白(顺式)的核酸序列,所述包装构建体缺乏能够表达 Pol 蛋白的核酸序列。或者,基于慢病毒的包装系统还可以包括能够表达 Vpr-反转录病酶-整合酶蛋白(反式)的单独的核酸构建体,所述包装系统包括缺乏能够表达 Pol 蛋白的核酸序列的包装构建体。

本发明所公开的基因转移构建体可以包括目的序列。所述基因转移构建体还可以包括标记编码序列。例如,所述目的序列和标记编码序列可以与至少一个转录调控元件有效连接。任选地,所述基因转移构建体可以含有两、三、四、五个或更多个目的序列。所述目的序列可以是相同的或不同的,并且可以与单独的转录调控元件有效连接,或者可以与同另一目的序列、标记编码序列或调节序列有效连接的转录调控元件有效连接。例如,在本发明所公开的表达系统中,所述基因转移构建体可以包括第七、第八、第九或更高编号(ordered)的核酸序列,其中所述第七核酸序列编码第三、第四、第五或更高编号的选定的目的蛋白。

所述基因转移构建体还可以包括位于所述目的序列 3'端的土拨鼠肝炎病毒转录后调节元件。所述基因转移构建体还可以包括一个或多个长末端重复(LTR)序列,本文的其他部分对该序列进行了讨论。

所述基因转移构建体以及本文所公开的其他构建体还可以是自我失活型(SIN)构建体。SIN 载体是新一代的反转录病毒载体,所述载体利用所述病毒反转录酶的独特性质使整合的转移载体前病毒 DNA 的顺式作用序列失活。上述序列可以包括存在于所述 LTR 中的病毒启动子以及存在于整合的载体前病毒 DNA 中的任何包装序列。几种制备 SIN 载体的策略是可以获得的并且在本领域中是公知的例如,可以使用 Tahir A Rizvi 在 Non-Human Primate Lentiviral Vectors for Human Gene Therapy; Genetic Disorders in the Arab World United Arab Emirates (获自网址 <http://www.cags.org.ae/cbc101v.pdf>) 中所述的“断裂内含子”策略,所述文献关于断裂内含子策略的教导全部通过引用的方式纳入本文。所述“断裂内含子”策略利用有效真核剪接位点的并入来从整合载体前病毒 DNA 上缺失所述包装序列,使得它不能形成可以被所述病毒蛋白进一步包装和增殖的 RNA。这消除了所述载体 RNA 与任何内源或外源病毒的 RNA 发生潜在重组的可能性,所述病毒可以是偶然地或因其他原因存在于经反转录载体转导的细胞内。另外,所述基因转移构建体还可以任选地在 3'长末端重复序列中含有突变。还可以用启动子序列来取代

5'或 3'长末端重复序列。

此外，本发明所公开的表达系统可以包括含有目的序列和标记编码序列的基因转移构建体，且在所述目的序列和标记编码序列之间具有一元件，其中所述元件可使所述目的序列和标记编码序列产生差异表达。目的序列和标记编码序列之间的元件可以是内部核糖体进入位点 (IRES) 或内部核糖体进入位点样元件 (IRES 样)。本文的其他部分中讨论了 IRES 和 IRES 样元件。所述基因转移构建体还可以包括至少一个转录调控元件，本文的其他部分中对其进行了讨论。如本文其他部分所述，所述存在于基因转移构建体中的一个或多个转录调控元件还可以是可调节的。所述基因转移构建体还可以包含一个调节子序列 (regulator sequence)，或者所述调节子序列可由本文所述的单独的调节子构建体 (regulator construct) 提供。

此外，所述基因转移构建体可以包括标记编码序列和目的序列，其中所述标记编码序列和目的序列与相同或不同的转录调控元件 (TCE) 有效连接。例如，本发明公开了这样的基因转移构建体，其中所述目的序列与第一转录调控元件有效连接，并且所述标记编码序列与第二转录调控元件有效连接。在一个实例中，所述第一转录调控元件可以强于所述第二转录调控元件。在这种情况下，所述与第一 TCE 有效连接的标记编码序列的表达应高于所述与第二 TCE 有效连接的目的序列。例如，所述与第一 TCE 有效连接的标记编码序列与所述与第二 TCE 有效连接的目的序列的表达比例可以为 99:1、98:2、97:3、96:4、95:5、94:6、93:7、92:8、91:9、90:10、89:11、88:12、87:13、86:14、85:15、84:16、83:17、82:18、81:19、80:20、79:21、78:22、77:23、76:24、75:25、74:26、73:27、72:28、71:29、70:30、69:31、68:32、67:33、66:34、65:35、64:36、63:37、62:38、61:39 或 60:40。

此外，本发明公开了含有两个方向相反的启动子以及双向启动子的基因转移构建体。例如，所述目的序列和标记编码序列可以按相反方向表达。在另一个实例中，所述目的序列和标记编码序列可以按相反方向表达。此外，所述目的序列可以与第一转录调控元件有效连接，并且所述标记编码序列可以与第二转录调控元件有效连接。所述第一和第二转录调控元件可以是相同的或不同的。此外，至少一个所述转录调控元件是可调节的。同时，所述目的序列和标记编码序列还可以与一单转录调

控元件有效连接,所述转录调控元件是可调节的。所述单转录调控元件可以是可调节的双向启动子。

本发明具体地公开了包括载体的基因转移构建体,其中所述载体包括第一核酸序列、第二核酸序列和第三核酸序列,其中所述第一核酸序列包括与第一转录调控元件有效连接的目的序列,其中所述第二核酸序列与第二转录调控元件有效连接,并且编码调控所述第一核酸序列表达的多肽,其中所述第三核酸序列包括与第一转录调控元件有效连接的调节子靶序列,并且其中所述第一和第二转录调控元件取向相反。

本发明还公开了包括载体的基因转移构建体,其中所述载体包括第一核酸序列、第二核酸序列和第三核酸序列,其中所述第一核酸序列、第二核酸序列和第三核酸序列与一单转录调控元件有效连接,其中所述第一核酸序列包括目的序列,其中所述第二核酸序列编码可调控所述第一核酸序列表达的多肽,其中所述第三核酸序列包括与所述第一转录调控元件有效连接的调节子序列,并且其中所述转录调控元件能够驱动所述第一和第二核酸序列的表达。

所述基因转移构建体的载体可以是病毒载体,并且所述病毒载体可以任选地为自我失活型的。此外,所述基因转移载体的第一核酸序列的表达可以是可调节的。

本发明还公开了包含本文所公开的基因转移构建体的细胞和细胞系。

本发明还公开了任选地包含 RNA 外排元件的构建体。术语“RNA 外排元件”是指顺式作用转录后调节元件,所述元件可调节 RNA 转录物由细胞的细胞核至细胞质的转运。RNA 外排元件的实例包括但不限于人免疫缺陷病毒(HIV) rev 应答元件(RRE)(参见例如, Cullen et al. (1991) *J. Virol.* 65:1053; 和 Cullen et al. (1991) *Cell* 58:423-426)和乙型肝炎病毒转录后调节元件(PRE)(参见例如, Huang et al. (1995) *Molec. and Cell. Biol.* 15(7): 3864-3869; Huang et al. (1994) *J. Virol* 68(5): 3193-3199; Huang et al. (1993) *Molec. and Cell. Biol* 13(12): 7476-7486), 和美国专利 No. 5,744,326。这些参考文献关于 RNA 外排元件的教导全部通过引用的方式纳入本说明书)。通常,所述 RNA 外排元件位于基因的 3'UTR, 并且可以一个或多个拷贝插入。RNA 外排元件可以插入形成本发明的包装细胞系的任一或全部独立的载体中。

本发明公开的构建体可以任选地包括编码 Tat 的核酸序列。本发明还公开了可以任选地包括编码 Rev 的核酸序列的构建体。所述编码 Tat 和 Rev 的核酸序列可以为所述编码 Gag 或 Gag-Pol 的核酸序列的一部分或者是与所述编码 Gag 或 Gag-Pol 的核酸序列相分离。所述 Tat 和 Rev 蛋白分别在转录水平和转录后水平调节 HIV 基因的表达水平。例如，由于 HIV 长末端重复 (LTR) 的弱基础转录活性，前病毒的表达最初会产生少量的编码 Tat、Rev 和 Nef 蛋白的多剪接转录物。Tat 通过结合新生 RNA 中的茎环结构 (反式作用应答元件 [TAR]) 显著地增加了 HIV 转录，从而募集细胞周期蛋白-激酶复合体，所述复合体可以利用聚合酶 II 复合体激发转录延伸。

具体而言，Rev 是可直接与其 Rev 应答元件 (RRE) RNA 靶序列结合的核质穿梭蛋白，所述 RNA 靶序列是所有未剪接和不完全剪接的病毒 mRNA 的一部分。当发生多聚化并且随后与细胞辅因子相互作用时，Rev 促进上述 mRNA 跨核膜易位，导致产生晚期病毒蛋白。

Rev 通过起到 RNA 基序 (RRE，天然存在于 HIV 转录物的被膜编码区中) 和所述细胞核外排机制的组件之间的连接物作用来实现这一作用。Rev 结合序列是这样的核酸：它会在体外或体内特异性地与 Rev 结合 (通常是 RNA)，或者会与编码在体外或体内与 Rev 结合的核酸的核酸特异性地结合 (即 RNA 或 DNA)。几篇论文描述了用于监测 Rev 结合的体外结合测定法，包括 Wong-Staal et al. (1991) *Viral And Cellular Factors that Bind to the Rev Response Element in Genetic Structure and Regulation of HIV* (Haseltine and Wong-Staal eds.; Harvard AIDS Institute Series on Gene Regulation of Human Retroviruses 第一卷的一部分)，311-322 页和其中所引用的参考文献，所述参考文献描述了凝胶迁移率变换测定法和用于检测生物样本 (包括人血液) 中 Rev 的足迹测定法。这些参考文献关于监测 RNA 结合的结合分析的教导全部通过引用的方式纳入本说明书。

本发明所公开的构建体可以任选地包括含有 RRE 的核酸序列。RRE 通常存在于 HIV 转录物的被膜编码区以及细胞核外排机制的组件中。如上所述，当 RRE 与 Rev 发生多聚化，并且随后与多个细胞辅因子相互作用时，上述病毒 mRNA 就会跨核膜易位。

本发明还公开了内部核糖体进入位点 (IRES) 和内部核糖体进入位

点样元件。内部核糖体进入位点(IRES)为能够介导40S核糖体亚基进入内部的顺式作用RNA序列,所述40S核糖体亚基位于一些真核生物和病毒的信使RNA的翻译起始密码子的上游。虽然IRES序列极其多样,并且存在于越来越多的mRNA中,但是IRES元件含有保守的Y<sub>n</sub>-X<sub>m</sub>-AUG单元(Y:嘧啶;X:核苷酸),该单元似乎是IRES发挥功能所必需的。每年都有新的IRES序列不断被添加到公共数据库中,且未知IRES序列的数量肯定仍然很多。

IRES样元件也是能够介导40S核糖体亚基进入内部的顺式作用序列,所述核糖体亚基位于一些真核生物和病毒信使RNA的翻译起始密码子的上游。与IRES元件不同,在IRES样元件中,Y<sub>n</sub>-X<sub>m</sub>-AUG单元(Y:嘧啶;X:核苷酸)不是必需的,该单元似乎是IRES发挥功能所必需的。

本发明所公开的构建体可以任选地包括IRES或IRES样元件。例如,本发明公开的包装构建体还可以包括位于所述第一和第二核酸序列之间的元件,其中所述元件可使所述第一与第二核酸序列之间产生差异表达。在另一个实例中,所述位于第一和第二核酸序列之间的元件为内部核糖体进入位点或内部核糖体进入位点样元件。在另一个实例中,本发明所公开的包装构建体还可以包括位于所述第一或第二核酸序列与所述第三核酸序列之间的元件,其中所述第三核酸序列并不位于所述第一和第二核酸序列之间,并且其中所述元件可使所述第一或第二核酸序列与所述第三核酸序列之间产生差异表达。

所述IRES或IRES样元件可以是天然存在的或非天然存在的。IRES的实例包括但不限于<http://ifr31w3.toulouse.inserm.fr/IRESdatabase/>的IRES数据库中存在的IRES。IRES的实例还可以包括但不限于EMC-病毒IRES或HCV-病毒IRES。此外,所述IRES或IRES样元件可以是突变过的,其中所述IRES或IRES样元件的功能得到了保持。

本发明还公开了转录调控元件(TCE)。TCE是能够驱动与其有效连接的核酸序列表达的元件。本发明所公开的构建体包括至少一个TCE。TCE可以任选地为组成型的或可调节的。

本发明还公开了含有取向相反的第一和第二转录调控元件的构建体,其中一个所述转录调控元件的活性可影响另一个转录调控元件的活性。任选地,所述两个转录调控元件可以是并置的或者在所述第一和第

二转录调控元件之间存在一个接头序列。例如，所述接头序列可以是一种染色体隔离子。

可调节的 TCE 可以包含这样一种核酸序列，所述核酸序列能够结合由调节子构建体表达的融合蛋白的结合结构域，从而使得所述转录抑制结构域起到抑制所述可调节 TCE 中所含核酸序列转录的作用。

本发明还公开了调节子构建体和调节子序列。调节子构建体可以是含有调节子序列的构建体。调节子序列可以是能够调控与调节子靶序列有效连接的序列表达的序列。例如，调节子序列可以是能够编码多肽的序列，所述多肽可调控与本文其他部分所述的核酸构建体中调节子靶序列有效连接的核酸序列的表达。例如，调节子构建体可以是含有能够编码可药物调控的（例如药物诱导的）阻遏融合蛋白的构建体，所述融合蛋白含有 DNA 结合结构域和转录抑制结构域。或者，所述构建体除含有可调节 TCE 之外还含有能够编码可药物调控的（例如药物诱导的）阻遏融合蛋白的核酸序列，所述融合蛋白含有 DNA 结合结构域和转录抑制结构域。在这种情况下，能够编码可药物调控的（例如药物诱导的）阻遏融合蛋白的核酸序列与所述可调节 TCE 位于相同的构建体上，所述阻遏融合蛋白会与所述可调节 TCE 结合。例如，所述包装构建体和基因转移构建体可以包括能够编码可药物调控的（例如药物诱导的）阻遏融合蛋白和与所述阻遏融合蛋白结合的可调节 TCE 两者的核酸序列。

如本说明书通篇所述，本发明所公开的构建体可以含有调节子序列或含有调节子靶序列的可调节 TCE 或者同时含有两者。所述调控构建体可以含有能够编码四环素抑制子 (tetR) 或四环素激活子 (tetA) (或者被称为反向 tetR-VP16) 蛋白的调节子序列，所述蛋白可以结合 tetO 序列。所述 tetO 序列可以位于 TCE 中。所述调节子构建体可以任选地包括编码核定位信号的核酸序列，例如 SV40 核定位信号。例如，所述调节子构建体可以包括 SEQ ID NO: 1 的序列。此外，所述调节子构建体还可以任选地包括一个或多个 VP16 最小反式激活结构域。例如，所述调节子构建体可以包括 SEQ ID NO: 2 或 SEQ ID NO: 3 的序列。TetR-VP16 还可以被称为“tet-off”。反向 tetR-VP16 还可以被称为“tet-on”。

所述调节子构建体可以任选地包括改变形式的 tetR 和 tetA，以防止在所述 tetR 和 tetA 蛋白之间形成异源二聚体。所述改变形式的 tetR 和 tetA 可以包括彼此独立或相结合的 E 和 B tet 操纵子 DNA 结合结构

域。例如，所述调节子构建体可以包括 SEQ ID NO: 4 或 SEQ ID NO: 5 的序列。

可调节的 TCE 可以任选地包括调节子靶序列。可调节的 TCE 可以包含这样一种核酸序列，所述核酸序列能够结合由调节子构建体表达的融合蛋白的结合结构域，从而使得转录抑制结构域起到抑制所述可调节 TCE 中所含核酸序列转录的作用。调节子靶序列可以包括一个或多个 tet 操纵子序列 (tetO)。所述调节子靶序列可以与其他序列有效连接，所述其他序列包括但不限于 TATA 框或编码 GAL-4 的核酸序列。

本发明所描述的基因转移构建体可以任选地包括第二调节子序列。例如，本文所述的基因转移构建体可以任选地包括与第二调节子序列有效连接的第二 GAL-4 编码核酸序列和第二目的序列，其中所述第二目的序列与第三转录调控元件有效连接，其中所述第二目的序列选自微小 RNA、shRNA 和 siRNA，其中所述第二调节子序列位于所述第二 GAL-4 编码核酸序列和所述第二目的序列之间。

当存在可调节 TCE 和调节子序列时，无论它们是否位于相同或不同的构建体上，都使得可以诱导性且可逆转地表达所述与可调节 TCE 有效连接的序列。这样一来，所述可调节 TCE 就可以提供一种用于选择性地可诱导和可逆转地表达目的序列的方法。

例如，可调节 TCE 可以由四环素或多西环素调节。此外，所述 TCE 可以任选地包括至少一个 tet 操纵子序列。在一个实例中，至少一个 tet 操纵子序列可以与 TATA 框有效连接。

此外，如本文其他部分所述，所述 TCE 可以是启动子。本说明书的全文给出了可用于本文公开的包装构建体的启动子的实例。例如，启动子可以包括但不限于基于 CMV 的启动子、CAG 启动子、基于 SV40 的启动子、热休克蛋白启动子、mH1 启动子、hH1 启动子、鸡  $\beta$ -肌动蛋白启动子、U6 启动子、泛素 C 启动子或 EF-1  $\alpha$  的启动子。

此外，本文所公开的 TCE 还可以包括彼此有效连接的一个或多个启动子、启动子的一部分或者彼此有效连接的启动子的一部分。例如，转录调控元件可以包括但不限于 CMV 启动子的 3' 部分、CMV 启动子的 5' 部分、 $\beta$ -肌动蛋白启动子的一部分或者与 CAG 启动子有效连接的 3' CMV 启动子。

调控哺乳动物宿主细胞中载体转录的优选启动子可以获取自不同

的来源例如病毒基因组，或者获取自异源哺乳动物例如 $\beta$ -肌动蛋白启动子，所述病毒例如多瘤病毒、猿猴病毒 40 (SV40)、腺病毒、反转录病毒、乙型肝炎病毒并且最优选巨细胞病毒。可以很方便地以 SV40 限制性片段的形式获得 SV40 病毒的早期和晚期启动子，所述限制性片段还含有 SV40 病毒复制起点 (Fiers et al., Nature, 273:113 (1978)，所述文献关于病毒启动子的部分全部通过引用的方式纳入本文)。可以很方便地以 HindIII E 限制性片段的形式获得人巨细胞病毒的立即早期启动子 (Greenway, P.J. et al., Gene 18:355 360 (1982)，所述文献关于病毒启动子的部分全部通过引用的方式纳入本文)。当然，本发明还可以使用来自宿主细胞或相关物种的启动子，且所述启动子可以被用于进行组织特异性基因表达或组织特异性调节的基因表达。所引用的参考文献关于启动子的教导全部通过引用的方式纳入本说明书。

“增强子”通常是指在与所述转录起始位点的距离不固定的地方发挥作用的 DNA 序列，并且可以位于所述转录单元的 5' 方向 (Laimins, L. et al., Proc. Natl. Acad. Sci. 78:993 (1981)) 或 3' 方向 (Lusky, M.L., et al., Mol. Cell Bio. 3: 1108 (1983))。所引用的各参考文献关于增强子的教导全部通过引用的方式纳入本说明书。此外，增强子可以位于内含子中 (Banerji, J.L. et al., Cell 33:729 (1983)) 以及所述编码序列本身中 (Osborne, T.F., et al., Mol. Cell Bio. 4: 1293 (1984))。所引用的各参考文献关于增强子可能位置的教导全部通过引用的方式纳入本说明书。它们的长度通常在 10 至 300 bp 之间，并且它们以顺式发挥作用。增强子起到增加由邻近启动子起始的转录的作用。增强子还常常含有介导转录调节的应答元件。启动子也可以含有介导转录调节的应答元件。增强子常常可以决定对基因表达的调节。虽然目前已知许多来源于哺乳动物基因 (珠蛋白、弹性蛋白酶、胎球蛋白和胰岛素) 的增强子序列，但是研究人员通常使用来源于真核细胞病毒的增强子来进行一般性表达。优选的实例为位于复制起点 (bp 100 270) 晚期侧的 SV40 增强子、巨细胞病毒早期启动子增强子、位于复制起点晚期侧的多瘤病毒增强子和腺病毒增强子。

所述启动子和/或增强子可以通过可触发其功能的特定化学事件或光而被特异性地激活。可以使用例如四环素和地塞米松等试剂来调节系统。还存在通过暴露于辐照 (例如  $\gamma$  辐照) 或烷化化学治疗药物来增强病

毒载体基因表达的方式。

在某些实施方案中，所述启动子和/或增强子区域可以起到组成型启动子和/或增强子的作用，以使所要转录的转录单元区域的表达最大化。在某些构建体中，所述启动子和/或增强子区域在所有真核细胞类型中都具有活性，尽管所述启动子和/或增强子区域只能在特定的时间在特定的细胞类型中表达。一种优选的这类启动子为 CMV 启动子(650 个碱基)。其他优选的启动子为 SV40 启动子、巨细胞病毒(全长启动子)和反转录病毒载体 LTR。

本发明还公开了双向转录调控元件。例如，本发明公开了一种双向转录调控元件，它包含与 CAG 启动子的 5'端融合的 CMV 启动子的 3'端。本发明还公开了一种双向转录调控元件，它包含与人 EF-1 $\alpha$  启动子的 5'端融合的 CMV 启动子的 3'端。本发明还公开了一种双向转录调控元件，它包含与 CAG 启动子的 5'端融合的小鼠 H1 启动子的 5'端。本发明还公开了一种双向转录调控元件，它包含与 SV40 启动子的 5'端融合的 CMV 启动子的 3'端。所述双向转录调控元件如本文其他部分所公开的转录调控元件一样，可以是可调节的或者是组成型的。本发明还公开了一种双向转录调控元件，它包含与 ef1 $\alpha$  启动子的 5'端融合的 CMV 启动子的 5'端。

所述双向转录调控元件可以包括 SEQ ID NO: 12、SEQ ID NO: 13、SEQ ID NO: 14 或 SEQ ID NO: 51 中列出的序列。双向转录调控元件还可以包括调节子靶序列，并且可以利用抗生素(例如四环素或多西环素)来对其进行调节。

已经证明，所有特异性调节元件都可被克隆并用于构建表达载体，所述表达载体可在特定细胞类型(例如黑色素瘤细胞)中被选择性表达。胶质纤维酸性蛋白(GFAP)启动子已经被用于选择性表达胶质源性细胞内的基因。

真核宿主细胞(酵母、真菌、昆虫、植物、动物、人或有核细胞)中所用的表达载体还可以含有终止转录所必需的序列，所述序列可以影响 mRNA 的表达。上述区域被转录为编码组织因子蛋白的 mRNA 的非翻译部分中的多腺苷酸化区段。所述 3'非翻译区还包括转录终止位点。优选地，所述转录单元还含有多腺苷酸化区域。该区域的一个好处是它可增加了所述被转录单元如 mRNA 那样被加工和转运的可能性。多腺苷酸化信号

在表达构建体中的鉴别和用途是公知的。优选地，在所述转基因构建体中使用同源多腺苷酸化信号。在某些转录单元中，所述多腺苷酸化区域来源于 SV40 早期多腺苷酸化信号，并且由大约 400 个碱基组成。还优选地，所述被转录单元仅含有其他标准序列或含有其他标准序列以及上述序列，以便促进所述构建体的表达或增加其稳定性。

Cre 重组酶是一种来源于噬菌体 P1 的 I 型拓扑异构酶，它可催化位于 loxP 位点之间的 DNA 的位点特异性重组 (Abremski, K. and Hoess, R. (1984) *J. Biol. Chem.*, 259, 1509-1514, 所述文献关于 Cre 重组酶结构和功能的教导全部通过引用的方式纳入本说明书)。所述酶无需能量辅助因子，且 Cre 介导的重组可以迅速在底物和反应产物之间达到平衡 (Abremski, K. et al. (1983) *Cell*, 32, 1301-1311, 所述文献关于 Cre 重组酶作用机制的教导全部通过引用的方式纳入本说明书)。所述 loxP 识别元件是 34 碱基对 (bp) 的序列，包括位于一个起定向作用的 8 bp 的间隔区两侧的两个 13 bp 的反向重复序列 (Metzger, D. and Feil, R. (1999) *Curr. Opin. Biotechnol.*, 10, 470-476, 所述文献关于 loxP 识别元件以及它们在 Cre 重组酶作用中的作用的教导全部通过引用的方式纳入本说明书)。重组产物依赖于 loxP 位点的位置和相对取向。两种含有单 loxP 位点的 DNA 可以融合。位于同向 loxP 位点之间的 DNA 将以环状形式被切除，而位于反向 loxP 位点之间的 DNA 将相对于外部序列而被翻转。

在本文所述基因转移构建体中，与转录调控元件有效连接的核酸序列的表达也可以被 Cre 重组酶调节。例如，一种基因构建体可以包括一种载体，其中所述载体包括第一核酸序列、第二核酸序列、第三核酸序列和一种含有能够编码可筛选标记的核酸序列的调节子靶序列，其中所述第一核酸序列包括与第一转录调控元件有效连接的目的序列，其中所述第二核酸序列与第二转录调控元件有效连接，并且编码可调控所述第一核酸序列表达的多肽，其中所述第三核酸序列包括与所述第一转录调控元件有效连接的调节子序列，并且其中所述调节子靶序列也与所述第一转录调控元件有效连接，且位于所述第一转录调控元件和所述第一核酸序列之间。在这种情况下，所述调节子靶序列的侧翼可以接有 TATA 序列，所述 TATA 序列还与至少一个 tet 操纵子序列连接。所述带有伴随序列的调节子靶序列的侧翼还可以接有 loxP 位点，使得当 Cre 介导

重组时，所述调节子靶序列可被切除，并且所述目的序列可以与所述第一转录调控元件融合，从而表达所述目的序列。

本发明还公开了包装构建体，其中所述第一核酸序列与第一转录调控元件有效连接，并且所述第二核酸序列与第二转录调控元件有效连接。本发明还公开了包装构建体，其中所述第一和第二核酸序列与第一转录调控元件有效连接，并且所述第三核酸序列与第二转录调控元件有效连接。

任选地，所述第一转录调控元件可以强于所述第二转录调控元件。在这种情况下，所述与第一 TCE 有效连接的一个或多个序列的表达应高于所述与第二 TCE 有效连接的一个或多个序列。例如，所述与第一 TCE 有效连接的一个或多个序列与所述与第二 TCE 有效连接的一个或多个序列的表达比例可以为大约 99:1、98:2、97:3、96:4、95:5、94:6、93:7、92:8、91:9、90:10、89:11、88:12、87:13、86:14、85:15、84:16、83:17、82:18、81:19、80:20、79:21、78:22、77:23、76:24、75:25、74:26、73:27、72:28、71:29、70:30、69:31、68:32、67:33、66:34、65:35、64:36、63:37、62:38、61:39、60:40 或任一介于其间的比例。

本发明还公开了含有两种方向相反的启动子以及双向启动子的包装构建体。例如，所述第一和第二核酸序列可以按相反方向表达。在另一个实例中，所述第一和第二核酸序列相对于所述第三核酸序列可以按相反方向表达。任选地，所述标记编码序列和目的基因可以按相反方向表达。此外，所述第一核酸序列可以与第一转录调控元件有效连接，并且所述第二核酸序列可以与第二转录调控元件有效连接。此外，所述第一和第二核酸序列可以与第一转录调控元件有效连接，并且所述第三核酸序列可以与第二转录调控元件有效连接。所述第一和第二转录调控元件可以是相同的或不同的。此外，至少一个所述转录调控元件是可调节的。而且，所述第一和第二核酸序列还可以与一单转录调控元件有效连接，所述转录调控元件是可调节的。此外，所述第一、第二和第三核酸序列还可以与一单转录调控元件有效连接，所述单转录调控元件是可调节的。所述单转录调控元件是同样可调节的双向启动子。

典型的启动子由最小启动子和其他上游顺式元件组成。Lewin, B. *Gene VI* (Oxford University Press, Oxford, 1997)、Odell, J. T., Nagy, F. & Chua, N.-H. *Nature* 313, 810-812 (1990), 和 Benfey, P.

N. & Chua, N.-H. *Science* 250, 959-966 (1990)。所述最小启动子主要是其中 RNA 聚合酶与起始转录区结合的 TATA 框区，但是它本身不具有转录活性。Benfey, P. N. & Chua, N.-H. *Science* 250, 959-966 (1990)。当所述顺式元件一旦与特异性转录因子结合时，它单独地或共同地决定启动子的时空表达模式 (Benfey, P. N. & Chua, N.-H. *Science* 250, 959-966 (1990))。美国专利 No. 5, 814, 618 公开了一种含有多个 tet 操纵子序列 (在说明书中被定义为增强子或抑制子) 并且侧翼接有最小启动子的双向启动子。美国专利 No. 5, 955, 646 公开了双向异源构建体。美国专利 No. 5, 368, 855 公开了一种天然存在的双向启动子。美国专利 No. 5, 359, 142 公开了经过处理以使得在增强基因表达方面存在变化的构建体。美国专利 No. 5, 627, 046 公开了一种天然存在的双向启动子。美国专利 No. 5, 827, 693 公开了经修饰的血红蛋白启动子。所有上述参考文献关于双向启动子的教导全部通过引用的方式纳入本说明书。

本发明还公开了含有一种或多种突变的包装构建体，所述突变位于编码 Gag 和 Gag-Pro-Pol 蛋白的核酸序列中，所述突变可减少合成 Gag-Pro 和 Gag-Pro-Pol 蛋白所必需的移码或翻译连续。本发明还公开了含有一种或多种突变的包装构建体，所述突变可减少合成 Gag-Pro 和 Gag-Pro 蛋白所需的移码或翻译连续。

在 I 型 HIV (HIV-1) 和 SIV-感染的细胞内，Gag 和 Gag-Pol 天然地由相同的 mRNA 转录物制备，摩尔比例为约 20:1。该比例可由 *gag* 和 *pol* 或 *gag* 和 *pro* 读码框之间重叠区中的核糖体移码或连续来实现 (Swanstrom, R., and J. W. Wills. 1997. Synthesis, assembly, and processing of viral proteins, p. 263-334. *In* J. M. Coffin, S. H. Hughes, and H. E. Varmus (ed.), *Retroviruses*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 所述文献关于移码的教导全部通过引用的方式纳入本说明书)。作为成熟病毒体的催化亚基的前体，Pol 是病毒体成熟和具备感染性所必需的，并且 Pol 是否能并入装配中的病毒颗粒中依赖于它与 Gag 的关联 (Id.)。所述 *gag-pol* 移码位点由保守性的七核苷酸滑动序列 (UUUUUA) SEQ ID NO: 7 组成，所述序列下游紧邻一个 RNA 二级结构区域 (Swanstrom, R., and J. W. Wills. 1997. Synthesis, assembly, and processing of viral proteins, p. 263-334. *In* J. M. Coffin, S. H. Hughes, and H. E.

Varmus (ed.), *Retroviruses*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.). 当苯丙氨酸和亮氨酸的 tRNA (密码子 UUU UUA; SEQ ID NO: 8) 相对于所述 *gag* 框滑回一个核苷酸 (-1) (UUU UUA; SEQ ID NO: 8 → UUU UUU; SEQ ID NO: 9) 并且在所述 *pol* 读码框内继续翻译时, 就会在所述滑动序列内发生核糖体的物理移码 (Jacks, T., M. D. Power, F. R. Masiarz, P. A. Luciw, P. J. Barr, and H. E. Varmus. 1988. Characterization of ribosomal frameshifting in HIV-1 *gag-pol* expression. *Nature* 331:280-283, 所述参考文献关于 HIV-1 中的核糖体移码的教导全部通过引用的方式纳入本说明书)。例如, 所述突变可以破坏移码所必需的环结构。这可以通过改变或移除各个核苷酸以破坏环结构来实现。

例如, 图 3 示出了 HIV *gag* 和 HIV *gag-pol* 中移码所必需的环结构。图 4 示出了 HIV *gag* 和 HIV *gag-pol* 中移码所必需的环结构的改变的序列, 所述改变的序列导致移码所必需的环结构被破坏。本发明公开了包装构建体, 其中所述第一核酸序列包括移码所必需的 *gag* 和 *gag-pol* 序列中的突变。任选地, 移码所必需的 *gag* 序列可以包括点突变。例如, 移码所需的所述 *gag* 序列可以包括图 1 中所示的点突变。任选地, 所述第一核酸序列可以包括 SEQ ID NO: 10 的核苷酸序列。例如, 所述第二核酸序列可以包括单核苷酸插入以及图 2 中所示的几个点突变。任选地, 所述第二核酸序列可以包括 SEQ ID NO: 11 的核苷酸序列。

不同物种之间的密码子偏好可能明显不同。为增强异种蛋白在特定表达系统 (大肠杆菌 (*E. coli*)、酵母、昆虫或哺乳动物细胞) 中的表达水平, 调整所述异种蛋白的密码子频率以便与所述宿主表达系统的密码子频率匹配是十分重要的。这一过程被称为密码子优化。密码子优化是指改变基因序列以使得密码子选择与目的细胞/物种中的有效 tRNA 池相匹配。密码子优化是一种利用来自小 RNA 和 DNA 病毒的基因来增加蛋白表达的有力工具, 所述病毒通常含有重叠的读码框以及嵌入编码区内的结构元件; 这些特征在大 DNA 病毒中并不普遍。

使用密码子优化的 I 型人免疫缺陷病毒 (HIV-1) 的 *env* 基因 (Andre, S., B. Seed, J. Eberle, W. Schraut, A. Bultmann, and J. Haas. 1998. Increased immune response elicited by DNA vaccination with a synthetic gp120 sequence with optimized codon usage. *J. Virol.*

72:1497-1503) 和 *gag* 基因 (Deml, L., A. Bojak, S. Steck, M. Graf, J. Wild, R. Schirmbeck, H. Wolf, and R. Wagner. 2001. Multiple effects of codon usage optimization on expression and immunogenicity of DNA candidate vaccines encoding the human immunodeficiency virus type 1 Gag protein. *J. Virol.* 75:10991-11001; zur Megede, J., M. C. Chen, B. Doe, M. Schaefer, C. E. Greer, M. Selby, G. R. Otten, and S. W. Bamett. 2000. Increased expression and immunogenicity of sequence-modified human immunodeficiency virus type 1 *gag* gene. *J. Virol.* 74:2628-2635) 进行免疫导致所述基因表达增强和针对所述抗原的免疫应答加强。使用许多其他病原性生物进行的类似研究确认了密码子优化会增强 DNA 疫苗的效率的潜能, 所述病原性生物例如李斯特菌属 (*Listeria*) (Nagata, T., M. Uchijima, A. Yoshida, M. Kawashima, and Y. Koide. 1999. Codon optimization effect on translational efficiency of DNA vaccine in mammalian cells: analysis of plasmid DNA encoding a CTL epitope derived from microorganisms. *Biochem. Biophys. Res. Commun.* 261:445-451), 产破伤风毒素细菌 (Stratford, R., G. Douce, L. Zhang-Barber, N. Fairweather, J. Eskola, and G. Dougan. 2000. Influence of codon usage on the immunogenicity of a DNA vaccine against tetanus. *Vaccine* 19:810-815), 原质团 (*Plasmodium*) (Nagata, T., M. Uchijima, A. Yoshida, M. Kawashima, and Y. Koide. 1999. Codon optimization effect on translational efficiency of DNA vaccine in mammalian cells: analysis of plasmid DNA encoding a CTL epitope derived from microorganisms. *Biochem. Biophys. Res. Commun.* 261:445-451), 人乳头状瘤病毒 (human papillomavirus) (Cid-Arregui, A., V. Juarez, and H. zur Hausen. 2003. A synthetic E7 gene of human papillomavirus type 16 that yields enhanced expression of the protein in mammalian cells and is useful for DNA immunization studies. *J. Virol.* 77:4928-4937; Liu, W., F. Gao, K. Zhao, W. Zhao, G. Fernando, R. Thomas, and I. Frazer. 2002. Codon modified human papillomavirus type 16 E7 DNA vaccine enhances cytotoxic T-lymphocyte induction

and anti-tumour activity. *Virology* 301: 43-52), 以及其他病原性生物 (Gurunathan, S., D. M. Klinman, and R. A. Seder. 2000. DNA vaccines: immunology, application, and optimization. *Annu. Rev. Immunol.* 18: 927-974)。可以使用各种本领域技术人员已知的技术进行密码子优化。例如, 可以使用 Ramakrishna L, Anand KK, Mohankumar KM, Ranga U. *J Virol.* 2004 Sep; 78(17): 9174-89 中所述的方法。所有引用的参考文献关于密码子优化的教导全部通过引用的方式纳入本说明书。

本文还公开了已经进行了密码子优化的包装构建体。例如, 可以修饰本文所述的包装构建体以使得所述第一核酸序列是经过密码子优化的。在另一个实施方案中, 所述第二核酸序列可以经过密码子优化的。此外, 所述第一和第二核酸都可以是经过密码子优化的。

本发明还公开了包装构建体, 其中所述构建体能够产生无复制能力的重组体。本发明还公开了包装构建体, 其中所述构建体不能产生有复制能力的重组体。如上所述, 考虑到反转录载体 (特别是能够感染不分裂细胞的慢病毒) 的优势, 许多研究的目标是获得用于生产纯净的重组病毒储液 (不含有复制能力的辅助病毒) 的改进方法。通常可以通过将合适的前病毒 DNA 载体引入哺乳动物细胞 (“包装细胞”) 中来生产重组反转录病毒, 所述细胞可生产将所需重组 RNA 壳体化所必须的病毒蛋白, 但是缺乏包装病毒 RNA 的信号 ( $\psi$  序列)。因此, 虽然所需的反转录病毒的 gag、pol 和 env 基因是完整的, 但是这些包装系不会释放野生型辅助病毒。然而, 当使用含有包装所必需的  $\psi$  序列的单独载体转染所述细胞时, 就可以通过重组产生野生型反转录病毒 (Mann et al. (1983) *Cell* 33: 153)。对于所述载体的某些应用 (例如基因疗法) 来说, 上述情况具有明显的安全隐患, 特别是使用慢病毒 (例如 HIV) 时。

目前避免由于重组导致产生有复制能力辅助病毒的危险的方法包括在所述用于产生包装系的病毒构建体中进行额外的突变 (例如 LTR 缺失) 和将产生病毒体所必需的病毒基因分开放到不同的质粒上。例如, 最近已经证实, 可以通过在包装细胞中将 gag 和 pol 基因与 env 基因分离开, 来生产重组莫洛尼鼠白血病毒 (MuLV) (不含可检测的辅助病毒) (Markowitz et al. (1998) *J. Virol.* 62(4): 1120)。在使用第三种含有  $\psi$  包装信号的载体共转染上述包装细胞时, 会获得含有两种共同编

码病毒体生产所需病毒蛋白的不同质粒的包装细胞，这样降低了产生完整反转录病毒所必需的重组事件的发生机率。

本发明公开的构建体可以任选地包括编码核定位信号的核酸序列。此外，本发明所公开的构建体可以任选地包括与编码四环素反式作用子的核酸有效连接的编码核定位信号的核酸序列。例如，本发明所公开的构建体可以包括与编码四环素反式作用子(例如 tet-on)的核酸有效连接的编码核定位信号的核酸序列。核定位序列是指导多肽从胞质穿过核膜进入细胞核的序列。所述编码核定位信号的核酸还可以包括转录调控元件。本文的其他部分公开了转录调控元件。所述编码核定位信号的核酸序列的侧翼还可以接有至少一个接头序列，所述接头序列可以编码例如 SEQ ID NO: 15 (GGGGS)，它包括后面接有一个丝氨酸残基的四个甘氨酸残基。接头序列可以是染色体隔离子并且还可以是通用序列。通常，所述接头序列的作用是减少融合蛋白各功能结构域之间的干扰。例如，所述接头序列可以减少与 tetR 或 tetA 蛋白、SV40 NLS、VP16 的干扰或与 ZNF10 沉默蛋白的干扰。作为染色体隔离子的接头可以减少本发明所公开的构建体中的可诱导型启动子与组成型启动子之间的干扰，从而减少所述诱导型启动子的渗漏。

本发明还公开了含有本文所公开的包装构建体的细胞系。本文的其他部分还描述了用于生产细胞系的方法。

上文和下文所述的实施方案适用于本文所公开的任何组合和方法。

## 系统

本文还公开了使用上文所述的包装构建体的包装系统。例如，包装系统可以包括本发明的包装构建体和本文其他部分所述的表达包膜糖蛋白的核酸构建体。本发明还公开了包装细胞系。用于生产病毒样颗粒的包装细胞系包括靶细胞和本文所公开的包装构建体。包装细胞系还可以包括本文其他部分所述的表达包膜糖蛋白的核酸构建体。本文所用的包膜糖蛋白将可以使得由所述包装构建体产生的颗粒假型化。本文描述了含有能够表达包膜糖蛋白的核酸序列的构建体。例如，所述包膜构建体可以包括水泡性口膜炎病毒的 G 糖蛋白 (VSV G) 和莫洛尼氏白血病毒 (MLV) 的包膜。

本文还公开了包装和表达系统，其中所述包装构建体含有反式编码 Gag 和 Gag-Pro-Pol 的核酸。例如，本发明公开了含有第一、第二和第三包装构建体的表达系统。所述第一包装构建体包括含有编码 Gag 多聚蛋白的核酸序列的第一核酸构建体，其中所述编码 Gag 的核酸序列包括一个或多个可减少移码或翻译连续的突变，并且与至少一个转录调控元件有效连接。所述第二包装构建体包括含有编码 Gag-Pro-Pol 蛋白的核酸序列的第二核酸构建体，其中所述编码 Gag-Pro-Pol 的核酸序列包括一个或多个可减少移码或翻译连续的突变，并且与至少一个转录调控元件有效连接。所述第三核酸构建体包括编码包膜糖蛋白的第三核酸序列，其中所述第三核酸序列与至少一个转录调控元件有效连接。所述含有上述构建体的包装和表达系统还可以包括含有至少一个目的基因的基因转移构建体。

本发明还公开了一种含有第一核酸构建体和第二核酸构建体的包装系统，所述第一核酸构建体包括编码 Gag 多聚蛋白的第一突变核酸，其中所述第一突变核酸与一个转录调控元件有效连接，所述第二核酸构建体包括编码 Gag-Pol 多聚蛋白的第二突变核酸，其中所述第二突变核酸与一个转录调控元件有效连接。所述第一和第二核酸构建体中的突变会使得所述 Gag 和 Gag-Pol 多聚蛋白的表达比例可形成病毒颗粒。任选地，所述包装系统的第一突变核酸可以与最小 CMV 启动子有效连接，并且所述第二突变核酸可以与所述热休克蛋白启动子有效连接。本文其他部分描述了适用于所述包装系统的构建体的其他启动子。

可以在体外、在体内和活体外 (*ex vivo*) 条件下使用本发明的构建体和病毒颗粒，以将目的序列引入靶细胞中 (例如真核细胞) 或哺乳动物体内 (例如人或其他哺乳动物或脊椎动物)。所述细胞可通过商业途径获得或获取自保藏单位，或例如通过活组织检查而直接获自哺乳动物。所述细胞可以获取自其将返回的哺乳动物；或者所述细胞可以获取自相同或不同物种的另一/不同哺乳动物。例如，使用本发明的包装细胞系或病毒颗粒可以将目的 DNA 引入到非人细胞 (例如猪细胞) 中，然后再将所述细胞引入到人体内。或者，例如在需要将本发明的病毒颗粒递送到进行基因疗法的哺乳动物体内的情况下，可不必将所述细胞从哺乳动物体内分离出来。

例如在 Kasid et al., Proc. Natl. Acad. Sci. USA, 87: 473 (1990);

Rosenberg et al., N. Engl. J. Med., 323:570 (1990); Williams et al., Nature, 310:476 (1984); Dick et al., Cell, 42:71 (1985); Keller et al., Nature, 318:149 (1985); 和 Anderson et al., U.S. Pat. No. 5,399,346 中描述了活体外疗法, 上述文献关于活体外疗法的教导全部通过引用的方式纳入本说明书。

用于将病毒颗粒直接给予(引入)哺乳动物的方法通常是本领域技术人员公知的。例如, 给药模式包括肠胃外、注射、粘膜、全身、植入、腹膜内、口腔、皮内、经皮肤(例如以缓释聚合物形式)、肌内、静脉内(包括输注和/或快速浓注)、皮下、局部、硬膜外等。优选地, 本发明的病毒颗粒可以通过可药用载体来给药, 所述可药用载体例如盐水、无菌水、林格氏溶液和等渗氯化钠溶液。

给予哺乳动物的本发明的病毒颗粒的剂量(包括给药频率)随各种因素而变, 包括给药模式和给药途径; 受体哺乳动物的体型、年龄、性别、健康状况、体重和饮食; 所要治疗的疾病或障碍症状的性质和程度; 同期治疗的种类、治疗频率和所期望的效果。

本发明公开了含有本文所述的包装构建体的表达系统, 其中所述表达系统还包括含有编码包膜糖蛋白的核酸序列的包膜核酸构建体, 其中所述编码包膜糖蛋白的核酸序列与至少一个转录调控元件有效连接; 并且所述表达系统还包括含有一个或多个目的序列的基因转移构建体。本发明还公开了表达系统, 其中包膜糖蛋白有助于进入细胞。任选地, 所述包膜糖蛋白可以是病毒包膜糖蛋白(例如水泡性口膜炎病毒的 G 蛋白(VSV-G))或者是本领域已知可介导进入细胞的几种其它病毒糖蛋白之一。

任选地, 所述表达系统还包括编码核定位信号的构建体, 所述构建体含有与上文所述编码四环素反式作用子的核酸有效连接的核定位信号编码核酸序列。上文公开了核定位序列。例如, 所述编码核定位信号的构建体从 5'端至 3'端还依次包括巨细胞病毒启动子、第一接头编码序列、第二核定位信号、第二接头序列和四环素反式作用子编码序列, 其中所述被编码的接头为 GGGGS (SEQ ID NO: 15)。

本发明还公开了含有第一包装构建体并同时含有第二包装构建体的表达系统, 其中所述第一包装构建体包括含有编码 Gag 多聚蛋白的核酸序列的第一核酸构建体, 其中所述编码 Gag 的核酸序列含有一个或多

个可减少移码或翻译连续的突变，并且与至少一个转录调控元件有效连接；所述第二包装构建体包括含有编码 Gag-Pol 多聚蛋白的核酸序列的第二核酸构建体，其中所述编码 Gag-Pol 的核酸序列含有一个或多个可减少移码或翻译连续的突变，并且与至少一个转录调控元件有效连接。所述表达系统还包括第三核酸构建体，所述第三核酸构建体包括编码包膜糖蛋白的第三核酸序列，其中所述第三核酸序列与至少一个转录调控元件有效连接。所述表达系统还包括含有一个或多个目的序列的基因转移构建体。任选地，所述表达系统还可以包括编码核定位信号的构建体，所述构建体含有与编码四环素反式作用子的核酸有效连接的核定位信号编码核酸序列。核定位序列是指导多肽从胞质穿过核膜进入细胞核的序列。

上文公开的表达系统还可以包括含有编码核定位信号的第四核酸序列的第四核酸构建体，所述核定位信号与四环素反式作用子有效连接。所述第四核酸构建体还可以包括转录调控元件，例如启动子。所述核定位信号编码序列的侧翼还可以接有至少一个接头序列。所述第四核酸序列从 5'端至 3'端还可以依次包括巨细胞病毒启动子、编码 SEQ ID NO: 15 (GGGGS) 的核酸序列、编码核定位信号的核酸序列、编码 SEQ ID NO: 15 (GGGGS) 的核酸序列和编码四环素反式作用子的核酸序列。

本发明还公开了含有本文其他部分所公开的表达系统的细胞系。

本发明还公开了包膜核酸构建体，所述包膜核酸构建体包括编码包膜糖蛋白的核酸序列，其中所述编码包膜糖蛋白的核酸序列与至少一个转录调控元件有效连接。所述包膜糖蛋白可以有助于进入细胞。任选地，所述包膜糖蛋白可以是病毒的包膜糖蛋白。在一个实例中，所述包膜糖蛋白可以是水泡性口膜炎病毒的 G 蛋白 (VSV-G)。

本发明还公开了下述实施方案，其中顺式作用元件是进行壳体化、反转录和整合所必需的。所提供的顺式作用元件与本文所述构建体可以是反式或顺式 (*in cis*) 的。例如，缺乏能够表达 Pol 蛋白的核酸序列的包装构建体可以任选地包括能够表达 Vpr-反转录病酶-整合酶蛋白的核酸序列 (顺式)。或者，包装系统还可以包括能够表达 Vpr-反转录病酶-整合酶蛋白的单独的核酸构建体 (反式)，所述包装系统包括缺乏能够表达 Pol 蛋白的核酸序列的包装构建体。

本发明还公开了一种基因转移方法，包括将本文其他部分所述的包

装核酸构建体引入细胞中，并且将含有编码包膜糖蛋白的核酸序列的包膜构建体引入所述细胞中；其中所述编码包膜糖蛋白的核酸序列与至少一个转录调控元件有效连接，并且将本文其他部分所述的含有一个或多个目的基因的基因转移构建体引入细胞；使所述细胞处于可形成病毒样颗粒的条件下，所述病毒样颗粒含有一个或多个目的基因或目的序列。

本发明还公开了含有外源目的序列的细胞，其中使用上文所述的基因转移方法将所述目的序列转移到细胞内。

本文所述构建体可以任选地包括编码标记产物的核酸序列。该标记产物被用于确定所述基因是否已经被送递到细胞内，并且所述基因在被送递后是否立即表达。优选的标记基因为大肠杆菌 *lacZ* 基因(编码  $\beta$  半乳糖苷酶)和绿色荧光蛋白基因。

在一些实施方案中，所述标记可以是可筛选标记。适合用于哺乳动物细胞的可筛选标记的实例为二氢叶酸还原酶(DHFR)、胸腺嘧啶核苷激酶、新霉素、新霉素类似物 G418、潮霉素和嘌呤霉素。当这类可筛选标记被成功地转移到哺乳动物宿主细胞内时，如果将细胞置于筛选压力下则被转化的哺乳动物宿主细胞可以存活下来。存在两类广泛应用的不同的筛选方法。第一类是基于细胞代谢并且使用突变细胞系，所述细胞系缺少不依赖于补充培养基独立生长的能力。两个实例是 CHO DHFR 细胞和小鼠 LTK 细胞。上述细胞缺少在不添加养分(例如胸腺嘧啶脱氧核苷或次黄嘌呤)的情况下生长的能力。因为上述细胞缺少完整的核苷酸合成途径中所必需的某些基因，所以它们在补充培养基中未提供所缺少的核苷酸的情况下无法存活。补充所述培养基的一种替代方法是将完整的 DHFR 或 TK 基因引入缺少对应基因的细胞内，由此改变其生长需求。未使用 DHFR 或 TK 基因转化的细胞个体将不能在非补充培养基中存活。

第二类方法是优势选择，这是指可用于任何细胞类型中且不需要使用突变细胞系的筛选方案。上述方案通常使用药物来阻碍宿主细胞的生长。那些具有新基因的细胞将会表达可赋予细胞抗药性的蛋白并且将会在筛选中存活下来。这种优势选择的实例使用药物新霉素(Southern P. and Berg, P., J. Molec. Appl. Genet. 1:327 (1982))、麦考酚酸(Mulligan, R. C. and Berg, P. Science 209:1422 (1980))或潮霉素(Sugden, B. et al., Mol. Cell. Biol. 5:410-413 (1985. These))。所引用的参考文献关于优势选择实例的教导全部通过引用的方式纳入

本说明书。上述三个实例采用受真核细胞调控的细菌基因来分别赋予细胞对适当的药物 G418 或新霉素(遗传霉素)、xgpt(麦考酚酸)或潮霉素的抗性。其它的药物包括新霉素类似物 G418 和嘌呤霉素。

本发明还公开了包膜核酸构建体, 所述包膜核酸构建体包括编码包膜糖蛋白的核酸序列, 其中所述编码包膜糖蛋白的核酸序列与至少一个转录调控元件有效连接。所述包膜糖蛋白可以有助于进入细胞。任选地, 所述包膜糖蛋白可以是病毒的包膜糖蛋白。在一个实例中, 所述包膜糖蛋白可以是水泡性口膜炎病毒的 G 蛋白(VSV-G)。

### 方法

本文公开了制备病毒样颗粒的方法。例如, 一种制备病毒样颗粒的方法包括使用本发明的包装构建体。本发明还公开了制备病毒样颗粒的方法, 包括将任一种上文所述的包装核酸构建体引入细胞中; 并且将包膜构建体引入所述细胞中, 所述包膜构建体含有编码包膜糖蛋白的核酸序列, 其中所述编码包膜糖蛋白的核酸序列与至少一个转录调控元件有效连接; 并且使所述细胞处于可形成病毒样颗粒的条件下。

本发明还公开了制备病毒样颗粒的方法, 包括将任一种上文所述的包装核酸构建体引入细胞中; 并且将包膜构建体引入所述细胞中, 所述包膜构建体含有编码包膜糖蛋白的核酸序列, 其中所述编码包膜糖蛋白的核酸序列与至少一个转录调控元件有效连接; 并且使所述细胞处于可形成病毒样颗粒的条件下。

可以通过将选定的慢病毒序列插入到合适的载体(例如, 含有适合的调节元件(例如启动子和增强子)、限制性克隆位点、标记基因等的市售表达质粒)中来制备病毒样颗粒。这可以使用本领域公知的标准克隆技术来实现, 包括 PCR。所要克隆到这类载体中的慢病毒序列可以获取自任何已知来源, 包括慢病毒基因组 RNA 或与病毒 RNA 对应的 cDNA。适合的与慢病毒基因组 RNA 对应的 cDNA 是市售的, 并且包括例如含有 HIV 基因组序列的 pNLENV-1 (Maldarelli et al. (1991) *J. Virol.* 65: 5732), 该文献关于适合的与慢病毒基因组 RNA 对应的 cDNA 的教导全部通过引用的方式纳入本说明书。其他反转录病毒(例如慢病毒) cDNA 克隆的来源包括位于马里兰州洛克威尔的美国典型培养物保藏中心(ATCC)。上述参考文献关于目前可获得的与慢病毒基因组 RNA 对应的 cDNA 实例的教导全部通过引用的方式纳入本说明书, 这些克隆全部通过

引用的方式纳入本说明书，例如可用于本文所公开的组合物和方法中的反转录病毒 cDNA 克隆。

在反转录病毒序列(例如 gag、pol、env、LTR 和顺式作用序列)被克隆到适合的载体(例如表达载体)中之后，则可以按照本文所述对其进行修饰。在一个实施方案中，由质粒扩增的慢病毒序列(例如 pNLENV-1)可以被克隆到适合的骨架载体中(例如 pUC 载体(例如 pUC19)(University of California, San Francisco)、pBR322 或 pcDNA1(Invitrogen, Inc., Carlsbad, California))，然后通过缺失(使用限制性内切酶)、置换(例如使用定向诱变)或其他(例如化学)修饰方法修饰来防止选定的慢病毒序列表达或发挥功能。如上所述，可以移除或突变 gag、pol 和 env 基因的一部分以及选定的附加基因。例如，在一个实施方案中，对编码 Gag 和 Gag-Pro-Pol 多聚蛋白的核酸序列进行突变以减少合成 Gag-Pro 和 Gag-Pro-Pol 多聚蛋白所必需的移码或翻译连续。

本发明的每种载体均可以含有编码所需慢病毒蛋白(例如 gag、pol 和 env)或指导所需慢病毒功能(例如包装 RNA)所必需的最小慢病毒序列。即优选地，所述载体其余部分为非病毒来源的或者来自于非慢病毒(例如 HIV)的其他病毒。在一个实施方案中，对本发明的慢病毒载体中所含的慢病毒 LTR 进行修饰，所采用的方法是通过用来自另一种病毒的功能相似的序列置换所述 LTR 的一部分，从而产生杂合 LTR。例如，起启动子作用的所述慢病毒 5'LTR 可以被 CMV 启动子或来自不同反转录病毒(例如 MuLV 或 MuSV)的 LTR 部分地置换。替代地或额外地，所述慢病毒 3' LTR 可以被来自另一种基因或反转录病毒的多腺苷酸化序列部分地置换。任选地，所述 HIV-1 的 3'LTR 的一部分被兔 $\beta$ -珠蛋白基因的多腺苷酸化序列置换。通过以这种方式使本发明载体内的慢病毒序列总体上最少，载体之间发生重组的机率显著减少，所述重组会导致产生有复制能力的辅助慢病毒。

任何适合的表达载体均可用于本发明。如上所述，适合的表达构建体可以包括人巨细胞病毒(CMV)立即早期启动子构建体。所述巨细胞病毒启动子可以获自任何适合的来源。例如，完整的巨细胞病毒增强子-启动子可以来源于人巨细胞病毒(hCMV)。其他用于获得 CMV 启动子的适合来源包括商业来源，例如 Clontech (Mountain View, California)、

Invitrogen (Carlsbad, California) 和 Stratagene (La Jolla, California)。本发明可以使用部分或全部的 CMV 启动子。其他可用于实施本发明的构建体的实例包括使用 MuLV、SV40、劳斯氏肉瘤病毒 (RSV)、疫苗 P7.5、热休克蛋白和大鼠  $\beta$ -肌动蛋白启动子的构建体。在某些情况下,例如 RSV 和 MuLV,上述启动子-增强子元件位于所述 LTR 序列内或者邻近所述 LTR 序列。

基因转录、翻译、加工和分泌所必需的适合的调节序列在本领域中是已知的,并且可被选择用于在合适的细胞内指导所需蛋白的表达。因此,本文所用的术语“调节序列”包括任何位于基因翻译区 5'端(上游)或 3'端(下游)并可调控或影响所述基因的表达的遗传元件,例如增强子和启动子序列。例如,Goeddel, *Gene expression Technology: Methods in Enzymology*, 185 页, Academic Press, San Diego, Calif. (1990) 中讨论了这类调节序列,所述文献关于调节序列的教导全部通过引用的方式纳入本说明书。本领域普通技术人员能够选择可用于本发明的调节序列。

在一个实施方案中,本发明在本文公开的构建体内使用了诱导型启动子,使得选定基因的转录可以被启动或关闭。这将由细胞毒性病毒蛋白的表达导致的细胞毒性减到最少,增加了含有所述载体的包装细胞的稳定性。例如,高水平表达的 VSV-G(包膜蛋白)和 Vpr 可能是有细胞毒性的(Yee, J. -K., et al., *Proc. Natl. Acad. Sci*, 91:9654-9568 (1994), 并且因此可以利用诱导型启动子系统(例如诱导型 Tet 操纵子系统(GIBCO BRL, Carlsbad, California))来调控上述蛋白在本发明的包装细胞内的表达,使得可以利用所述培养基中的四环素浓度来精确调节基因的表达(即生成反转录病毒颗粒)。即在存在四环素的情况下使用 Tet 操纵子系统,四环素与 Tet 反式作用子融合蛋白(tTA)结合,防止 tTA 与所述 Tet 操纵子序列结合并且使得所述基因在 Tet 操纵子序列的调控下表达(Gossen et al. (1992) *PNAS* 89: 5547-5551),所述文献关于 tTA 和使得所述基因在 Tet 操纵子序列的调控下表达的教导全部通过引用的方式纳入本说明书。在不存在四环素的情况下,tTA 与 Tet 操纵子序列结合,防止所述基因在 Tet 操纵子的调控下表达。

其他可用于受控地表达所述蛋白的诱导型操纵子系统的实例——其中所述蛋白提供了假型包膜——为: 1) 响应于金属离子或糖皮质激素

的诱导型真核启动子(例如金属硫蛋白启动子); 和 2) 大肠杆菌的 LacSwitch™ 诱导型哺乳动物表达系统(Stratagene) (La Jolla, California)。简言之,在所述大肠杆菌的乳糖操纵子中,所述 Lac 抑制子以同源四聚体的形式与 lac 操纵子结合,阻断所述 lac2 基因的转录。诱导剂(例如异乳糖(生理诱导剂)或异丙基- $\beta$ -D-硫代半乳糖苷(IPTG, 合成诱导剂))与 Lac 抑制子结合,导致发生构象变化并且有效地降低所述抑制子对操纵子的亲和力。当将所述抑制子从操纵子中移出时,由所述乳糖操纵子调控的转录会继续进行。

本发明还公开了选择性调节目的序列表达的方法,包括在合适于调节目的序列的条件下将本文所述的基因转移构建体引入靶细胞内。本文公开的方法还可以被用于指导目的序列以组织特异性方式表达。例如,基因转移构建体可以包括组织特异性 TCE,所述组织特异性 TCE 可以被用于驱动目的序列在特定组织中表达。这种基因转移载体可与包装构建体联合使用来制备本文所述的病毒颗粒。然后所述病毒颗粒可以被引入合子中。任选地,可以使用本文所公开的用于形成转基因动物的方法来实现组织特异性表达,其中所述目的序列的表达处于可诱导/可逆转的 TCE 的调控下。在这种动物中,所述目的序列的表达可以被限制于例如 DOX 的给药位点上。这样,目的序列的表达将仅发生在 DOX 给药位点上。

本发明还公开了将所述使用本发明方法形成的病毒颗粒给予受试者的方法。可以在体外、在体内和活体外使用本发明的构建体和病毒颗粒将目的序列引入靶细胞(例如哺乳动物细胞)中或哺乳动物(例如人)体内。所述细胞可通过商业途径获得或获取自保藏单位,或者例如通过活组织检查而直接获自哺乳动物。所述细胞可以获取自其将返回的哺乳动物;或者所述细胞可以获取自相同或不同物种的另一/不同哺乳动物。例如,使用本发明的包装细胞系或病毒颗粒可以将目的 DNA 引入到非人细胞(例如猪细胞)中,然后再将所述细胞引入到人体内。或者,例如在需要将本发明的病毒颗粒送递到进行基因疗法的哺乳动物体内的情况下,可不必将所述细胞从哺乳动物体内分离出来。

例如在 Kasid et al., *Proc. Natl. Acad. Sci. USA*, 87:473 (1990); Rosenberg et al., *N. Engl. J. Med.*, 323:570 (1990); Williams et al., *Nature*, 310:476 (1984); Dick et al., *Cell*, 42:71 (1985); Keller et al., *Nature*, 318:149 (1985); 和 Anderson et al., U.S.

Pat. No. 5,399,346 中描述了活体外疗法,上述文献关于活体外疗法的教导全部通过引用的方式纳入本说明书。

本发明还公开了将所述使用本发明方法形成的病毒颗粒给予受试者的方法。通常,成功的抗病毒疫苗大多数依赖于减毒活病毒。候选 HIV 减毒活疫苗并不理想,因为它们具有逆转、重组或突变的危险。其他现有的候选 HIV 疫苗在形成针对原代 HIV 分离物的细胞毒性 T 细胞免疫应答和普遍有效的中和抗体方面存在困难。已经证明病毒样颗粒 (VLP) 可以被安全地给予动物和人类患者,并且是细胞和体液免疫应答的强有力的和有效的刺激物。因此,VLP 可用作 HIV 疫苗。嵌合 HIV-1 VLP 对早期 VLP 设计进一步作出改进,产生了得到增强的免疫刺激,所述嵌合 HIV-1 VLP 是使用 HIV 或 SIV 衣壳蛋白加上 HIV 免疫表位和免疫刺激分子构建的。经由粘膜表面给予 VLP 疫苗也是一种极具前景的策略,使用该方法可诱发粘膜及全身的体液和细胞免疫应答。此外,关于抗原加工以及通过树突细胞 (DC) 呈递特定 (particulate) 抗原的新信息已形成用于改进的候选 VLP 疫苗的新策略。

用于将病毒颗粒直接给予(引入)哺乳动物的方法通常是本领域技术人员已知的。例如,给药模式包括肠胃外、注射、粘膜、全身、植入、腹膜内、口腔、皮内、经皮肤(例如以缓释聚合物形式)、肌内、静脉内(包括输注和/或快速浓注)、皮下、局部、硬膜外等。优选地,本发明的病毒颗粒可以通过可药用载体来给药,所述可药用载体例如盐水、无菌水、林格氏溶液和等渗氯化钠溶液。

给予哺乳动物的本发明的病毒颗粒的剂量(包括给药频率)随各种因素而变,包括给药模式和给药途径;受体哺乳动物的体型、年龄、性别、健康状况、体重和饮食;所要治疗的疾病或障碍症状的性质和程度;同期治疗的种类、治疗频率和所期望的效果。

本发明公开了含有本文所述的包装构建体的表达系统,其中所述表达系统还包括含有编码包膜糖蛋白的核酸序列的包膜核酸构建体,其中所述编码包膜糖蛋白的核酸序列与至少一个转录调控元件有效连接;并且所述表达系统还包括含有一个或多个目的序列的基因转移构建体。本发明还公开了表达系统,其中包膜糖蛋白有助于进入细胞。任选地,所述包膜糖蛋白可以是病毒包膜糖蛋白,例如水泡性口膜炎病毒的 G 蛋白 (VSV-G)。

任选地,所述表达系统还包括编码核定位信号的构建体,所述构建体含有与编码四环素反式作用子(例如 tet-on)的核酸有效连接的核定位信号编码核酸序列。核定位序列是指导多肽从胞质穿过核膜进入细胞核的序列。所述编码核定位信号的核酸还可以包括转录调控元件。本文的其他部分公开了转录调控元件。所述编码核定位信号的核酸序列的侧翼还可以接有至少一个接头序列,所述接头序列可以编码例如 SEQ ID NO: 15 (GGGGS)。接头序列还可以是通用序列。所述编码核定位信号的构建体从 5'端至 3'端还可以依次包括巨细胞病毒启动子、第一接头编码序列、第二核定位信号、第二接头序列和四环素反式作用子编码序列,其中所述被编码的接头为 GGGGS (SEQ ID NO: 15)。

本发明还公开了含有第一和第二包装构建体、第三核酸构建体和基因转移构建体的表达系统。所述第一包装构建体包括含有编码 Gag 蛋白的核酸序列的第一核酸构建体,其中所述编码 Gag 的核酸序列包括一个或多个可减少移码或翻译连读的突变,并且与至少一个转录调控元件有效连接。所述第二包装构建体包括含有编码 Gag-Pol 蛋白的核酸序列的第二核酸构建体,其中所述编码 Gag-Pol 的核酸序列包括一个或多个可减少移码或翻译连读的突变,并且与至少一个转录调控元件有效连接。所述表达系统还包括第三核酸构建体,所述第三核酸构建体包括编码包膜糖蛋白的第三核酸序列,其中所述第三核酸序列与至少一个转录调控元件有效连接。所述表达系统还包括含有一个或多个目的序列的基因转移构建体。任选地,所述表达系统还可以包括编码核定位信号的构建体,所述构建体含有与编码四环素反式作用子的核酸有效连接的上述核定位信号编码核酸序列。核定位序列是指导多肽从胞质穿过核膜进入细胞核的序列。

此外,所述编码核定位信号的核酸还可以包括转录调控元件。本文的其他部分公开了转录调控元件。所述编码核定位信号的核酸序列的侧翼还可以接有至少一个接头序列,所述接头序列可以编码例如 SEQ ID NO: 15 (GGGGS)。所述编码核定位信号的构建体从 5'端至 3'端还依次包括巨细胞病毒启动子、第一接头编码序列、第二核定位信号、第二接头序列和四环素反式作用子编码序列,其中所述被编码的接头为 GGGGS (SEQ ID NO: 15)。

上文公开的表达系统还可以包括含有编码核定位信号的第四核酸

序列的第四核酸构建体，所述核定位信号与四环素反式作用子有效连接。所述第四核酸构建体还可以包括转录调控元件，例如启动子。如上所述，所述核定位信号编码序列的侧翼还可以接有至少一个接头序列。所述第四核酸序列从5'端至3'端还可以依次包括巨细胞病毒启动子、编码 SEQ ID NO: 15 (GGGGS) 的核酸序列、编码核定位信号的核酸序列、编码 SEQ ID NO: 15 (GGGGS) 的核酸序列和编码四环素反式作用子的核酸序列。

本发明还公开了含有本文其他部分所公开的表达系统的细胞系。

本发明还公开了一种基因转移方法，包括将本文其他部分所述的包装核酸构建体引入细胞中，并且将包膜构建体引入所述细胞中，所述包膜构建体含有编码包膜糖蛋白的核酸序列，其中所述编码包膜糖蛋白的核酸序列与至少一个转录调控元件有效连接，并且将种本文其他部分所述的含有一个或多个目的基因的基因转移构建体引入细胞；使所述细胞处于可形成病毒样颗粒的条件下。所述病毒样颗粒含有一个或多个目的基因或目的序列。

本发明还公开了含有外源目的序列的细胞，其中使用上文所述的基因转移方法将所述目的序列转移到细胞内。

本发明还公开了由目的基因制备重组蛋白的方法，包括在合适于使细胞表达重组蛋白的条件下将靶细胞与病毒颗粒接触，所述病毒颗粒含有本文其他部分所公开的目的基因。例如，可以使所述靶细胞在体外或在体内与所述病毒颗粒接触。

本发明还公开了由目的基因或目的序列制备重组蛋白的方法，包括在合适于表达重组蛋白的条件下将本文所公开的基因转移构建体引入靶细胞中，其中所述目的序列为编码所述重组蛋白的核酸序列。如本文其他部分所公开的，所述重组蛋白的表达可以是可调节的。例如，所述重组蛋白的表达可以是可诱导的和可逆转的。

本发明还公开了制备重组蛋白的方法，包括在合适于使细胞表达所述重组蛋白的条件下将靶细胞与病毒颗粒接触，所述病毒颗粒含有一个或多个本文其他部分公开的编码所述重组蛋白的目的基因或序列。例如，可以使所述靶细胞在体外或在体内与所述病毒颗粒接触。

本发明还公开了制备重组蛋白的方法，包括将含有与调节子序列有效连接的启动子的第一核酸构建体引入靶细胞中，所述调节子序列与至

少一个 VP16 序列有效连接；使所述靶细胞处于可使所述待整合的第一核酸序列整合至所述靶细胞基因组中并且形成经修饰的靶细胞的条件下；将第二核酸构建体引入步骤 (b) 的经修饰的靶细胞中，所述第二核酸构建体含有与目的序列有效连接的调节子靶序列；其中所述目的序列为编码重组蛋白的核酸序列；使所述经修饰的靶细胞处于可使所述重组蛋白表达的条件下。

本发明还公开了制备重组蛋白的方法，包括将含有与调节子序列有效连接的启动子的第一核酸构建体引入靶细胞中，所述调节子序列与至少一个 VP16 序列有效连接；将第二核酸构建体引入步骤 (a) 的同一靶细胞中，所述第二核酸构建体含有与目的序列有效连接的调节子靶序列，其中所述目的序列为能够编码重组蛋白的核酸序列；并且使所述靶细胞处于可使所述待整合的第一和第二核酸序列整合至所述靶细胞基因组中并且形成经修饰的靶细胞的条件下；并且使所述经修饰的靶细胞处于可使所述重组蛋白表达的条件下。

例如，所述第一核酸构建体可以包括 SEQ ID NO: 44 的序列。所述第二核酸序列可以是任何本文其他部分所述的任一基因转移载体。例如，所述第二核酸可以包括与转录调控元件有效连接的目的序列，所述转录调控元件与调节子靶序列有效连接。所述第一核酸构建体还可以包括 IRES 或 IRES 样序列。例如，所述目的序列可以与 IRES 或 IRES 样序列有效连接，所述 IRES 或 IRES 样序列与可筛选标记有效连接。

所述制备重组蛋白的方法可以使用任何已知的细胞转染技术。本文其他部分公开了用于将细胞与病毒颗粒接触的其他方法。通常对于体外方法而言，在适合的转染条件下，将细胞与所述构建体或载体一起在适合的培养基中孵育（即培养），这在本领域中是公知的。例如，可以使用例如电穿孔和磷酸钙沉淀的方法 (O'Mahoney et al. (1994) *DNA & Cell Biol.* 13(12): 1227-1232)。

本发明还公开了含有本文公开的基因转移构建体的疫苗。本发明还公开了在受试者体内产生免疫应答的方法，包括给予所述受试者本文公开的基因转移构建体。

此外，本发明公开了在受试者体内产生免疫应答的方法，其中所述免疫应答为针对 HIV 的免疫应答，包括给予所述受试者本文公开的基因转移构建体，其中所述目的序列为能够表达 HIV 抗原的序列。

本文所用的特异性针对特定病原的“疫苗”或“用于接种受试者的组合物”是指这样一种制品，当将所述制品给予受试者时，会在受试者体内产生免疫原性应答。本文所用的“免疫原性”应答是指可赋予所述受试者针对所述病原的保护性免疫力的免疫应答。不希望受限于理论，人们认为免疫原性应答是由于产生中和抗体(即体液免疫应答)或由于免疫系统的细胞毒性细胞(即细胞免疫应答)或二者同时所导致的。本文所用的“免疫原性抗原”是指这样一种抗原，当将该抗原引入受试者体内时或者当在宿主或受试者细胞内合成该抗原时，它可以诱导免疫原性应答。本文所用的疫苗或免疫接种组合物的“有效量”是指这样一个量，当按照该量给予受试者所述疫苗或免疫接种组合物时，其足以赋予所述受试者保护性免疫力。历史上，人们已经了解到疫苗含有一种或多种特定分子组件或结构作为有效成分，所述特定分子组件或结构包含病原体，特别是病原体表面。这类结构可以包括表面组件例如通常存在于病原性生物中的蛋白、复合糖和/或复脂。

然而，应该强调的是本文所用的术语“疫苗”或“用于接种受试者的组合物”并不限于上述段落中所总结的一般含义。本文所用的这些术语还指本发明的目的序列或含有所述目的序列的组合物。所述目的序列促使在受试者细胞内生物合成一种或多种由所述目的序列编码的特定基因产物，其中所述基因产物是某种病原体的特异性抗原。然后所述生物合成的抗原起免疫原作用。已经注意到，所述目的序列以及由此而来的疫苗可以是任何编码所述特定免疫原性抗原的核酸。在本发明的一个优选实施方案中，所述疫苗的目的序列为DNA。为了分子生物学、细胞生物学和病毒免疫学领域技术人员的方便起见，所述目的序列可以包括并入有另外的基因或特定序列的质粒或载体(参见 Molecular Cloning: A Laboratory Manual, 2nd Ed., Sambrook, Fritsch and Maniatis, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989; 和 Current Protocols in Molecular Biology. Ausubel et al., John Wiley and Sons, New York 1987 (季刊)，上述文献关于质粒或载体的实例和用途的教导全部通过引用的方式纳入本说明书)。

近年来已经设计了数种重组亚基和病毒疫苗。美国专利 No. 4, 810, 492 描述了用作疫苗中的抗原的日本脑炎病毒(JEV)的E糖蛋白的生产，所述专利关于重组亚基和病毒疫苗的教导全部通过引用的方

式纳入本说明书。将对应的 DNA 克隆到表达系统中以便在适合的宿主细胞(例如大肠杆菌、酵母或高等动物细胞培养物)中表达所述抗原蛋白。美国专利 5,229,293 公开了含有 JEV E 蛋白基因的重组杆状病毒,所述专利关于将 DNA 克隆到表达系统中以表达抗原蛋白的方法的教导全部通过引用的方式纳入本说明书。所述病毒被用于感染培养物中的昆虫细胞使得可以生产并回收所述 E 蛋白用作疫苗。

美国专利 5,021,347 公开了一种重组痘苗病毒基因组,其中已经整合有 JEV E 蛋白基因。所述活重组痘苗病毒可被用作进行抗 JEV 免疫的疫苗。美国专利 5,494,671 公开了重组痘苗病毒和杆状病毒,其中所述病毒整合有编码登革血清型 2(dengue serotype 2)、登革血清型 4 和 JEV E 蛋白 C-末端截短的基因。美国专利 5,514,375 公开了各种表达一部分 JEV 开放读码框(从 prM 至 NS2B)的重组痘苗病毒。这些 pox 病毒诱导形成含有所述经加工的 M 蛋白和所述 E 蛋白的胞外颗粒。两种编码上述 JEV 蛋白的重组病毒可在小鼠体内产生高效价的中和抗体和血凝素抑制抗体,以及保护性免疫力。两次免疫处理后这些效应的程度要大于一次免疫处理后这些效应的程度。将含有 JEV prM/M 和 E 蛋白编码基因的重组痘苗病毒给予小鼠时,其可赋予小鼠保护性免疫力(Konishi et al., *Virology* 180:401-410 (1991))。已经证明使用含有 JEV prM 和 E 基因的重组痘苗病毒感染的 Hela 细胞可产生亚病毒颗粒(Konishi et al., *Virology* 188:714-720 (1992))。Dmitriev 等报导使用编码蜱传脑炎病毒的结构蛋白和某些非结构蛋白的重组痘苗病毒免疫小鼠(*J. Biotechnology* 44:97-103 (1996))。上述每篇参考文献关于重组痘苗病毒的教导全部通过引用的方式纳入本说明书。

已经制备了重组病毒载体来作为登革热的病毒疫苗。Zhao 等(*J. Virol.* 61:4019-4022 (1987))制备了含有登革血清型 4 的结构蛋白和 NS1 的重组痘苗病毒,并且在使用所述重组病毒感染哺乳动物细胞之后实现了表达。使用重组杆状病毒感染靶昆虫细胞获得了类似的表达(Zhang et al., *J. Virol.* 62:3027-3031(1988))。Bray 等(*J. Virol.* 63:2853-2856 (1989))还报道了基于 E 蛋白基因的重组痘苗登革疫苗,当使用所述疫苗进行攻击时其可赋予小鼠抗登革脑炎的保护性免疫力。Falgout 等(*J. Virol.* 63:1852-1860 (1989))和 Falgout 等(*J. Virol.* 64:4356-4363 (1990))报道了类似的结果。Zhang 等(*J. Virol.*

62: 3027-3031 (1988))证明当使用编码登革 E 和 NS1 蛋白的重组杆状病毒进行攻击时,其同样会保护小鼠免遭登革脑炎的伤害。其中将结构基因和非结构基因并入重组病毒疫苗中的其他组合未能产生明显的免疫力(Bray et al., *J. Virol.* 63:2853-2856 (1989))。同时,当使用表达所述 E 蛋白的重组杆状病毒进行免疫时,猴未能产生出针对登革病毒攻击的完全的保护性免疫力(Lai et al. (1990) pp 119-124 in F. Brown, R. M. Chancock, H. S. Ginsberg and R. Lerner (eds.) *Vaccines: 90 Modern approaches to new vaccines including prevention of AIDS*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY)。上述每篇参考文献关于将基因并入重组病毒疫苗中的方法以及将结构基因和非结构基因并入重组病毒疫苗中的实例的教导全部通过引用的方式纳入本说明书。

已经报道了使用断奶小鼠作为模型,使用重组 DNA 制品进行免疫来抵御 SLEV 和登革 2 病毒(Phillpotts et al., *Arch. Virol.* 141: 743-749 (1996); Kochel et al., *Vaccine* 15: 547-552 (1997))。编码 SLEV 的 prM 和 E 基因的质粒 DNA 可以通过单剂或双剂的 DNA 免疫来针对 SLEV 攻击提供部分保护。在这些实验中,对照小鼠的存活率为约 25%,并且在所述 DNA 免疫小鼠体内未检测到保护性抗体(Phillpotts et al., *Arch. Virol.* 141:743-749(1996))。在接受了三次皮内注射含有 prM 的重组登革-2 质粒 DNA 的小鼠中,全部小鼠均产生了抗登革-2 中和抗体,并且接受了对应 E 基因的小鼠中有 92%同样产生了中和抗体(Kochel et al., *Vaccine* 15: 547-552 (1997))。然而,使用双剂方案进行的攻击实验未能保护小鼠免遭致命的登革 2 病毒攻击。仅使用某些虫媒病毒(例如 JEV)的蛋白的重组疫苗不会导致高抗体滴度,所述重组疫苗是通过在细胞培养物中生物合成表达然后纯化或抗原处理来产生的。同时,与所述全病毒制品类似,这些疫苗也具有针对宿主抗原或所述载体产生有害变态反应的危险。抗登革病毒和 WNV 疫苗的开发水平比较落后,并且这类基于病毒或基于重组蛋白的疫苗面临着与上文提及的疫苗类似的问题。上述每篇参考文献关于将基因并入重组病毒疫苗中的方法、将结构基因和非结构基因并入重组病毒疫苗中的实例以及使用重组 DNA 制品进行免疫的方法的教导全部通过引用的方式纳入本说明书。

本发明还公开了制备抗体的方法。例如，本发明公开了一种通过在受试者体内诱发免疫应答来诱发生产抗体的体内方法。所述体外方法包括将按照本文其他部分公开的方法制备的重组蛋白引入受试者体内，其量足以诱发免疫应答。例如，可以在体外或体内将所述靶细胞与本文公开的基因转移构建体或病毒颗粒接触。

本发明还公开了生成抗目的蛋白的抗体的方法，包括：(a)将本文其他部分公开的基因转移构建体引入靶细胞中，其中所述基因转移构建体的转录调控元件为可调节或组成型的，其中所述目的序列能够编码目的蛋白；(b)使所述细胞处于可使步骤(a)中的核酸构建体整合到所述靶细胞的基因组中并且形成经修饰的靶细胞的条件下；(c)将步骤(b)的经修饰的靶细胞引入所述受试者体内；(d)给予所述受试者有效量的能够调节所述基因转移构建的转录调控元件的物质，所述物质的量足以诱导所述目的序列的表达，其中所述目的序列的表达量足以诱发免疫应答，并且其中所述免疫应答会生成针对所述目的蛋白的抗体。此外，可以分离由本文所述方法产生的抗体。

本发明还公开了鉴别可结合目的抗原的抗体的方法，所述方法包括：使疑似含有可结合目的抗原的抗体的样本与可表达所述目的抗原的靶细胞接触，并且确定样本中的抗体是否与所述由靶细胞表达的目的抗原结合，由此将所述结合目的抗原的抗体鉴别为可结合目的抗原的抗体。表达所述目的抗原的靶细胞可以是按照本文所述方法制成的靶细胞。所公开的鉴别可结合目的抗原的抗体的方法中所用的靶细胞可以是含有本文其他部分所述的基因转移构建体的靶细胞。例如，本发明还公开了鉴别可结合目的抗原的抗体的方法，所述方法包括：使疑似含有可结合目的抗原的抗体的样本与可表达所述目的抗原的靶细胞接触，其中所述靶细胞含有一种或多种权利要求1、55、208、247和286所述的核酸构建体；并且确定样本中的抗体是否与所述由靶细胞表达的目的抗原结合，由此将所述结合目的抗原的抗体鉴别为可结合目的抗原的抗体。

本发明还公开了生成抗目的蛋白的抗体的方法，包括：(a)将本文其他部分公开的基因转移构建体引入靶细胞中，其中所述基因转移构建体的转录调控元件为可调节或组成型的，其中所述目的序列能够编码目的蛋白；(b)使所述细胞处于可使步骤(a)中的核酸构建体整合到所述靶细胞的基因组中并且形成经修饰的靶细胞的条件下；(c)将步骤(b)的经

修饰的靶细胞引入所述受试者体内；(d) 给予所述受试者有效量的能够调节所述基因转移构建的转录调控元件的物质，所述物质的量足以诱发所述目的序列的表达，其中所述目的序列的表达量足以诱发免疫应答，并且其中所述免疫应答会生成针对所述目的蛋白的抗体。

此外，在该方法中可以使用不表达所述目的抗原的对照，使得疑似含有可结合目的抗原的抗体但实际不含可结合目的抗原的抗体的样本不会被鉴别为可结合所述目的抗原的抗体。所公开的鉴别可结合目的抗原的抗体的方法还可以被用于鉴别中和抗体。

本文所用术语“抗体”包括但不限于任何类的全免疫球蛋白(即完整的抗体)。天然抗体通常是异四聚体糖蛋白，由两条相同的轻(L)链和两条相同的重(H)链组成。通常，每条轻链均通过一个共价二硫键与一条重链相连，而不同免疫球蛋白同种型的重链间所述二硫键的数量是不相同的。每条重链和轻链还具有均匀分布的链内二硫键。每条重链均在一端具有一个可变区(V(H))，其后接有几个恒定区。每条轻链均在一端具有一个可变区(V(L))，并且在其另一端具有一个恒定区；所述轻链恒定区与所述重链第一恒定区并排，并且所述轻链可变区与所述重链可变区并排。人们认为特定的氨基酸残基在所述轻链和重链可变区之间形成了界面。根据其恒定区的氨基酸序列，来自任何脊椎动物的抗体的轻链均可被划为两种明显不同的类型，称为 kappa( $\kappa$ )和 lambda( $\lambda$ )。根据其重链恒定区的氨基酸序列，可以将免疫球蛋白划分为不同的类型。存在五大类人免疫球蛋白：IgA、IgD、IgE、IgG 和 IgM，并且其中几类还可被分为亚类(同种型)，例如 IgG-1、IgG-2、IgG-3 和 IgG-4；IgA-1 和 IgA-2。本领域技术人员将可识别小鼠中的对应类别。与不同类免疫球蛋白对应的重链恒定区分别被称为  $\alpha$ 、 $\delta$ 、 $\epsilon$ 、 $\gamma$  和  $\mu$ 。

本文使用术语“可变”来描述抗体之间序列不同的可变区的某些部分，并且决定了每种特定抗体针对其特定抗原的结合和特异性。然而，在抗体的整个可变区中变异性通常并非均匀地分布。在轻链和重链可变区中，变异性通常集中在三个被称为互补决定区(CDR)或高变区的区段中。可变区中保守度较高的部分被称为骨架(FR)。天然重链和轻链的可变区均包括四个 FR 区，主要采取  $\beta$  折叠构型、由三个 CDR 连接，这会形成连接所述  $\beta$  折叠结构的环(在某些情况下形成所述  $\beta$  折叠结构的一部分)。每条链中的 CDR 均被所述 FR 区紧密地结合在一起，并且与其它

链的 CDR 一起形成抗体的抗原结合位点(参见 Kabat E. A. et al., "Sequences of Proteins of Immunological Interest," National Institutes of Health, Bethesda, Md. (1987))。恒定区并不直接参与抗体与抗原的结合,但表现出各种效应子功能,例如抗体依赖性细胞毒性中抗体的参与。

本文所述术语“抗体或其片段”包括具有两种或多种抗原或表位特异性的嵌合抗体和杂交抗体和片段,例如 F(ab')<sub>2</sub>、Fab'和 Fab 等,包括杂交片段。因此,提供保持了与其特异性抗原结合的能力的抗体片段。例如,保持了目的蛋白结合活性的抗体片段均属于术语“抗体或其片段”的含义范围。这类抗体和片段可以利用本领域中已知的技术来制备,并且可以就特异性和活性按照实施例给出的方法以及用于生产抗体并就特异性和活性筛选抗体的常规方法对其进行筛选(参见 Harlow and Lane. Antibodies, A Laboratory Manual. Cold Spring Harbor Publications, New York, (1988), 所述文献涉及用于生产抗体并就特异性和活性对其进行筛选的常规方法的教导全部通过引用的方式纳入本说明书)。

例如,美国专利 4,704,692 中所述的抗体片段与抗原结合蛋白(单链抗体)的缀合物也属于“抗体或其片段”的含义范围,所述文献关于抗体片段与抗原结合蛋白(单链抗体)的缀合物的教导全部通过引用的方式纳入本说明书。

任选地,在其他物种中制备所述抗体并且对其进行“人源化”以用于在人体内给药。人源化形式的非人(例如鼠类)抗体为嵌合免疫球蛋白、免疫球蛋白链或其片段(例如 Fv、Fab、Fab'、F(ab')<sub>2</sub> 或抗体的其他抗原结合子序列),所述片段含有来源于非人免疫球蛋白的最小序列。人源化抗体包括人免疫球蛋白(接受者抗体),其中来源于接受者的互补决定区(CDR)的残基被来源于具有所需特异性、亲和力和能力的非人物种(例如小鼠、大鼠或兔)的 CDR 的残基(提供者抗体)所置换。在某些情况下,人免疫球蛋白的 Fv 骨架残基被对应的非人残基所置换。人源化抗体还可以包括既不存在于所述接受者抗体也不存在于输入的 CDR 或骨架序列中的残基。通常,所述人源化抗体将基本上包括至少一个——通常两个——可变区的全部,其中全部或基本上全部的 CDR 区与非人免疫球蛋白的 CDR 区对应,并且全部或基本上全部的 FR 区为人免疫球蛋白

白的 FR 区。最佳地，所述人源化抗体还将包括至少一部分免疫球蛋白恒定区 (Fc)，通常为人免疫球蛋白的恒定区的一部分 (Jones et al., *Nature*, 321: 522-525 (1986); Riechmann et al., *Nature*, 332: 323-327 (1988); 和 Presta, *Curr. Op. Struct. Biol.*, 2: 593-596 (1992)，所述文献关于人源化抗体的教导全部通过引用的方式纳入本说明书)。

用于人源化非人抗体的方法在本领域中是公知的。通常，人源化抗体具有一个或多个由非人来源引入其中的氨基酸残基。这些非人氨基酸残基通常是被称为“输入”残基，这通常取自“输入”可变区。

基本上可以按照 Winter 及其同事的方法 (Jones et al., *Nature*, 321: 522-525 (1986); Riechmann et al., *Nature*, 332: 323-327 (1988); Verhoeyen et al., *Science*, 239: 1534-1536 (1988)，上述文献关于抗体人源化的教导全部通过引用的方式纳入本说明书)，用一个或多个啮齿动物的 CDR 取代人抗体的对应序列来进行人源化。因此，这类“人源化”抗体为嵌合抗体 (美国专利 4, 816, 567，上述文献关于人源化抗体和嵌合抗体的教导全部通过引用的方式纳入本说明书)，其中一个基本上完整的人可变区已经被非人物种的对应序列所取代。事实上，人源化抗体通常是人抗体，其中一些 CDR 残基并且可能一些 FR 残基被啮齿动物抗体中类似位点的残基所取代。

为降低抗原性，用于制备人源化抗体的人可变区 (轻链和重链) 的选择非常重要。根据“最佳适配”法，用所述啮齿动物抗体的可变区序列在整个已知的人可变区序列文库中进行筛选。然后将与所述啮齿动物序列最接近的人序列作为人源化抗体的人骨架区 (FR) (Sims et al., *J. Immunol.*, 151: 2296 (1993) 和 Chothia et al., *J. Mol. Biol.*, 196: 901 (1987)，上述文献关于使用与所述啮齿动物序列最接近的人序列作为人源化抗体的人骨架区 (FR) 的教导全部通过引用的方式纳入本说明书)。另一种方法使用了一种来源于所有人抗体的共有序列的特定骨架，所述人抗体具有特定亚型的轻链或重链。同一种骨架可以被用于数种不同的人源化抗体 (Carter et al., *Proc. Natl. Acad. Sci. USA*, 89: 4285 (1992); Presta et al., *J. Immunol.*, 151: 2623 (1993)，上述文献关于使用与所述啮齿动物序列最接近的人序列作为人源化抗体的人骨架区 (FR) 的教导也全部通过引用的方式纳入本说明书)。

更加重要的是，将抗体人源化，并保持针对所述抗原的高亲和力以及其他有利的生物性质。为实现这一目标，根据一个优选的方法，利用一种使用亲代序列和人源化序列的三维模型来分析亲代序列和各种概念上的人源化产物的方法来制备人源化抗体。三维免疫球蛋白模型通常是可获得的并且是本领域技术人员所熟知的。计算机程序是可获得的，所述程序可举例说明和显示所选候选免疫球蛋白序列的可能三维构象结构。对上述展示的检查使得可以分析所述残基在候选免疫球蛋白序列发挥功能的过程中可能的作用，即对可影响候选免疫球蛋白结合其抗原的能力的残基进行分析。通过这种方式，可以从所述共有序列和输入序列中选择 FR 残基并加以组合以使得可以实现所需的抗体性质，例如对所述一种或多种靶抗原的亲和力增加。通常，所述 CDR 残基大多数会直接影响抗原结合(参见 1994 年 3 月 3 日公开的 WO 94/04679，所述专利关于 CDR 残基以及其对抗原结合的影响的教导全部通过引用的方式纳入本说明书)。

可以使用在进行免疫时能够生产人抗体的所有组成成分的转基因动物(例如小鼠)，所述转基因动物不能进行内源免疫球蛋白生产。例如，已经描述了在嵌合和种系突变小鼠体内进行的抗体重链结合区(J(H))基因的纯合缺失会导致对内源抗体生产的完全抑制。将人种系免疫球蛋白基因列转移到这类种系突变小鼠体内将会导致小鼠在受到抗原攻击时产生人抗体(例如参见 Jakobovits et al., Proc. Natl. Acad. Sci. USA, 90:2551-255 (1993); Jakobovits et al., Nature, 362:255-258 (1993); Bruggemann et al., Year in Immuno., 7:33 (1993)，上述文献关于在受到抗原攻击时产生人抗体的教导全部通过引用的方式纳入本说明书)。还可以使用噬菌体展示文库生产人抗体(Hoogenboom et al., J. Mol. Biol., 227:381 (1991); Marks et al., J. Mol. Biol., 222:581 (1991)，上述文献关于使用噬菌体展示文库生产人抗体的教导全部通过引用的方式纳入本说明书)。Cote 等和 Boerner 等的用于制备人单克隆抗体的技术也是可以获得的(Cole et al., Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, p. 77 (1985); Boerner et al., J. Immunol., 147(1):86-95 (1991)，上述文献关于制备人单克隆抗体的教导全部通过引用的方式纳入本说明书)。

本发明还公开了可产生所述单克隆抗体的细胞。本文所用术语“单

克隆抗体”是指获取自基本上同源的抗体群的抗体，即构成所述群的各个抗体除了可能存在少量的天然存在的突变之外都是相同的。本发明的单克隆抗体特别地包括“嵌合”抗体，其中所述重链和/或轻链的一部分与来源于特定的物种或者属于特定的抗体类或亚类的抗体中对应的序列相同或同源，同时所述链的其余部分与来源于另一物种或者属于另一抗体类或亚类的抗体中对应的序列相同或同源，所述术语“单克隆抗体”还包括这类抗体的具有所需活性的片段(参见美国专利 4,816,567 和 Morrison et al., Proc. Natl. Acad. Sci. USA, 81:6851-6855 (1984), 上述参考文献关于具体地包括嵌合抗体的单克隆抗体的教导全部通过引用的方式纳入本说明书)。

还可以使用杂交瘤法制备单克隆抗体，例如 Kohler and Milstein, Nature, 256:495 (1975) or Harlow and Lane. Antibodies, A Laboratory Manual. Cold Spring Harbor Publications, New York, (1988) 所述的那些方法。在一种杂交瘤方法中，通常使用免疫剂免疫小鼠或其他适合的宿主动物，以诱发产生或能够产生特异性结合所述免疫剂的抗体的淋巴细胞。或者，可以在体外对所述淋巴细胞进行免疫。优选地，所述免疫剂包括一种或多种存在于所述基因转移构建体中的目的序列。通常，单克隆抗体的生成依赖于是否可以获得纯化的蛋白或肽来用作免疫原。因此，本文所公开的方法提供了一种通过提供大量位于病毒颗粒内的目的蛋白来诱发强免疫应答并产生单克隆抗体的方法，所述病毒颗粒可以被注射到宿主动物体内。

该系统的优势包括容易生成、高水平表达和翻译后修饰与哺乳动物系统中所观察到的情况高度类似。该系统的用途包括以融合蛋白形式表达目的蛋白结构域的抗体。还可以通过将在信号序列和目的蛋白的成熟蛋白结构域的抗体核苷酸序列之间插入读框内基因片段来产生所述抗原。这会导致所述外源蛋白在病毒体表面上展示。该方法使得可以使用全病毒来进行免疫，不需要纯化靶抗原。

通常，当制备单克隆抗体时，如果需要人源的细胞则可以在生产单克隆抗体的方法中使用外周血淋巴细胞(“PBL”)，如果需要非人哺乳动物来源的细胞则使用脾细胞或淋巴结细胞。然后，使用适合的融合剂(例如聚乙二醇)将所述淋巴细胞与永生化细胞系融合以形成杂交瘤细胞(Goding, "Monoclonal Antibodies: Principles and Practice"

Academic Press, (1986) pp 59-103)。永生化细胞系通常是经转化的哺乳动物细胞，包括啮齿动物、牛、马和人来源的骨髓瘤细胞。通常，使用大鼠或小鼠的骨髓瘤细胞系。所述杂交瘤细胞可以在适合的培养基中培养，所述培养基优选地含有一种或多种可抑制非融合的永生化细胞的生长或存活物质。例如，如果亲代细胞缺少次黄嘌呤鸟嘌呤磷酸核糖转移酶(HGPRT 或 HPRT)，则所述用于培养杂交瘤的培养基将通常含有次黄嘌呤、氨基蝶呤和胸腺嘧啶脱氧核苷("HAT"培养基)，这些物质可防止 HGPRT 缺陷型细胞的生长。优选的永生化细胞系为具有下列特征的细胞：可有效融合、支持所选的产抗体细胞高水平的稳定表达抗体，以及对培养基(例如 HAT 培养基)敏感。更优选的永生化细胞系为鼠类骨髓瘤细胞系，所述细胞系可以获自，例如位于加利福尼亚州圣地亚哥的沙克研究所细胞分配中心(Salk Institute Cell Distribution Center, San Diego, Calif)和位于马里兰州洛克威尔的美国典型培养物保藏中心(American Type Culture Collection, Rockville, Md)。还已经描述了用于生产人单克隆抗体的人骨髓瘤和小鼠-人异种骨髓瘤细胞系(Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., "Monoclonal Antibody Production Techniques and Applications" Marcel Dekker, Inc., New York, (1987) pp 51-63)。然后可以测定所述培养杂交瘤细胞的培养基中是否存在抗所述目的蛋白的单克隆抗体。优选地，由所述杂交瘤细胞产生的单克隆抗体的结合特异性可以通过免疫沉淀或通过体外结合测定(例如放射免疫测定(RIA)或酶联免疫吸附测定(ELISA))来确定。这类技术和测定在本领域中是已知的，并且本文和 Harlow and Lane "Antibodies, A Laboratory Manual" Cold Spring Harbor Publications, New York, (1988)中有进一步表述。

在鉴别了所需的杂交瘤细胞之后，所述克隆可以通过有限稀释或 FACS 分选方法来亚克隆并且按照标准方法来培养。适合于该目的的培养基包括例如达尔伯克氏改良伊格尔氏培养基和 RPMI-1640 培养基。或者，所述杂交瘤细胞可以在体内培养，例如在哺乳动物腹水中。

可以按照常规的免疫球蛋白纯化方法来从所述培养基或腹水液中分离或纯化由所述亚克隆分泌的单克隆抗体，所述免疫球蛋白纯化方法例如蛋白 A-琼脂糖、蛋白 G、羟基磷灰石色谱、凝胶电泳、透析或亲和色谱。

体外方法也适用于制备单价抗体。可以使用本领域中已知的常规技术来消化抗体以产生其片段，特别是 Fab 片段。例如，可以使用木瓜蛋白酶来进行消化。1994 年 12 月 22 日公布的 WO 94/29348、美国专利 4,342,566 和 Harlow and Lane, *Antibodies, A Laboratory Manual*, Cold Spring Harbor Publications, New York, (1988) 中描述了木瓜蛋白酶消化的实例。抗体的木瓜蛋白酶消化通常会产生两个相同的被称为 Fab 片段的抗原结合片段——每个片段均带有一个抗原结合位点——和一个残留的 Fc 片段。胃蛋白酶处理会产生一个被称为 F(ab')<sub>2</sub> 片段的片段，所述片段具有两个抗原结合位点并且仍然能够交联抗原。

所述通过抗体消化产生的 Fab 片段还含有轻链恒定区和重链的第一恒定区。Fab' 片段与 Fab 片段的不同之处在于 Fab' 片段在重链结构域的羧基末端多了几个残基，包括来自于抗体铰链区的一个或多个半胱氨酸。F(ab')<sub>2</sub> 片段是含有两个 Fab' 片段的二价片段，所述两个 Fab' 片段由一个二硫键在铰链区相连接。本文中所命名的 Fab'-SH 是指其中恒定区的一个或多个半胱氨酸残基带有一个自由巯基基团的 Fab'。抗体片段最初是以 Fab' 片段对形式产生的，所述 Fab' 片段对之间带有铰链半胱氨酸。抗体片段的其他化学偶联形式也是已知的。

本发明还提供了一种分离的免疫原特异性的抗体互补位或片段。可以通过对分子进行化学性或机械性破坏来从整个抗体中分离出抗体特定的免疫原性表位。因此，可以按照本文中教导的方法对所获得的纯化片段进行测试以确定它们的免疫原性和特异性。任选地，所述抗体的免疫反应互补位是直接合成的。免疫反应片段被定义为至少含有大约 2-5 个来源于所述抗体氨基酸序列的连续氨基酸的氨基酸序列。

本发明还公开了具有生物活性的抗体片段。所述多肽片段可以通过将编码该多肽的核酸克隆到能够产生该多肽片段的表达系统(例如本文所公开的表达系统)中获得的。例如，技术人员可以确定获自特定杂交瘤的抗体的活性域，所述活性域可以导致与抗体和目的蛋白之间的相互作用相关的生物效应。例如，可以将对所述抗体的活性或结合特异性或亲和力没有贡献的氨基酸缺失而不会导致各自活性的丧失。例如，可以将氨基或羧基末端的氨基酸从天然的或经修饰的非免疫球蛋白分子或免疫球蛋白分子上连续的移除，并且用许多可实现的测定方法之一来测定各自的活性。在另一个实例中，抗体片段包括经修饰的抗体，其中

已经用至少一个氨基酸取代了位于特定位置的天然存在的氨基酸，并且氨基或羧基末端氨基酸的一部分甚至抗体的内部区域已经被多肽片段或其他部分(例如生物素)所置换，这有助于纯化所述经修饰的抗体。例如，可以下列方式将经修饰的抗体与麦芽糖结合蛋白融合：通过肽化学或将各种编码所述两个多肽片段的核酸克隆到表达载体中以使得所述编码区的表达可产生杂交多肽。可以通过将所述杂交多肽上样到淀粉酶亲和层析柱上来对其进行亲和纯化，然后通过用特异性蛋白酶因子 Xa 切割所述杂交多肽来从麦芽糖结合区分离所述经修饰的抗体受体(参见，例如 New England Biolabs Product Catalog, 1996, pg 164.)。类似的用于从真核细胞中分离杂交蛋白的纯化方法同样是可实现的。

所述片段——无论是否连接到其他序列上——包括对特定区域或特定氨基酸残基的插入、缺失、取代或其他所选的修饰，只要与未修饰的抗体或抗体片段相比，所述片段的活性没有被明显改变或削弱。这些修饰可以产生一些额外的性质，例如移除或增加能够形成二硫键的氨基酸、增加其生物寿命、改变其分泌特征等。在任何情况下，所述片段必须具有生物活性，例如结合活性、调节在结合结构域的结合等。可以通过诱变蛋白的特定区域然后表达并且测试所表达的多肽，从而鉴别所述抗体的功能或活性区域。这些方法对于本领域技术人员来说是显而易见的，并且可以包括对编码所述抗原的核酸进行定点诱变(Zoller MJ et al. Nucl. Acids Res. 10: 6487-500 (1982))。

可以使用各种免疫测定方法来选择可与特定蛋白、变体或片段选择性结合的抗体。例如，通常使用固相 ELISA 免疫测定来选择可选择性地与蛋白、蛋白变体或其片段发生免疫反应的抗体。关于可用于测定选择性结合的免疫测定方法和条件的描述请参见 Harlow and Lane. Antibodies, A Laboratory Manual. Cold Spring Harbor Publications, New York, (1988)。例如，可以通过 Munson et al., Anal Biochem., 107: 220 (1980) 的 Scatchard 分析法来确定单克隆抗体的结合亲和力。

本发明还提供了一种抗体试剂盒，包括装有单克隆抗体或其片段的容器以及一种或多种用于检测所述抗体或其片段与目的蛋白的结合的试剂。所述试剂可以包括，例如荧光标记、酶标记或其他标记。所述试剂还可以包括二抗或三抗或用于酶促反应的试剂，其中所述酶促反应会产生可视产物。

本发明还公开了在受试者体内诱导免疫应答的方法，包括将按照本文其他部分公开的方法制备的重组蛋白引入受试者体内，所述重组蛋白的量足以诱导免疫应答。例如，可以使所述靶细胞在体外或在体内与所述病毒颗粒接触。

本发明还公开了在受试者体内诱导免疫应答的方法，包括：(a)将基因转移构建体引入靶细胞中；(b)使所述细胞处于可使待整合的步骤(a)所述的核酸构建体整合到所述靶细胞基因组中的条件下；和(c)将步骤(b)所述的靶细胞引入所述受试者体内，所述靶细胞的数量足以诱导免疫应答。例如，所述目的序列能够编码膜蛋白(例如HIV膜蛋白)。此外，所述目的序列的表达可以是可诱导的、可逆转的或可诱导且可逆转的。

本文所用的“免疫应答”是指身体作为整体对所存在的抗原的反应，所述反应包括产生抗体、产生免疫力、产生对所述抗原的超敏感性和产生耐受性。因此，对抗原的免疫应答还包括在受试者体内对所述目的抗原产生体液和/或细胞免疫应答。“体液免疫应答”通过由浆细胞产生的抗体来介导。“细胞免疫应答”通过T淋巴细胞和/或其他白细胞介导。

本文所用的术语“抗原”是指任何可被抗体识别和/或可在个体体内诱发免疫应答的药剂(例如，任何物质、化合物、分子、蛋白或其他部分)。

本文所公开的方法可以用于任何细胞类型。换言之，任何细胞类型均可用作本文所公开的方法的靶细胞。真核宿主细胞可以包括但不限于酵母、真菌、昆虫、植物、动物、人和有核细胞。哺乳动物细胞也可被用于本文所述的方法中。靶细胞还可以含有任何本文所述的核酸构建体。例如，靶细胞可以含有一种或多种本文所述的基因转移构建体和/或一种或多种本文所述的包装构建体。

本文所用术语“哺乳动物”和“哺乳动物的”是指任何脊椎动物，包括单孔类动物、有袋类动物和胎盘类动物，所述脊椎动物会给其后代哺乳并且生育幼体(真哺乳亚纲哺乳动物或胎盘哺乳动物)或产卵(后兽亚纲哺乳动物或无胎盘哺乳动物)。哺乳动物物种的实例包括人和其他灵长类(例如猴、猩猩)、啮齿动物(例如大鼠、小鼠、豚鼠)和反刍动物(例如牛、猪、马)。

哺乳动物细胞的实例包括人(例如HeLa细胞、293T细胞、NIH 3T3

细胞)、牛、羊、猪、鼠(例如胚胎干细胞)、兔和猴(例如 COS1 细胞)的细胞。所述细胞可以是非分裂细胞(包括肝细胞、肌纤维、造血干细胞、神经元)或分裂细胞。所述细胞可以是胚胎细胞、骨髓干细胞或其他祖细胞。当所述细胞为体细胞时,所述细胞可以是例如上皮细胞、成纤维细胞、平滑肌细胞、血细胞(包括造血细胞、红细胞、T-细胞、B-细胞等)、肿瘤细胞、心肌细胞、巨噬细胞、树突细胞、神经元细胞(例如神经胶质细胞或星形胶质细胞)或病原体感染的细胞(例如那些被细菌、病毒、拟病毒、寄生虫或朊病毒感染的细胞)。

通常,分离自特定组织的细胞(例如上皮细胞、成纤维细胞或造血细胞)可被分类为一种“细胞型”。所述细胞可通过商业途径获得或获取自保藏单位或例如通过活组织检查而直接获自动物。或者,例如在需要将所述病毒送递到进行基因疗法的动物体内时,根本不必将所述细胞从动物身上分离出来。

尽管可以使用任何细胞类型来制备本文其他部分所述的重组蛋白或抗体,但是细胞表面存在的寡糖会给结晶和抗体的产生带来困难。本文公开了使用靶细胞制备重组蛋白和抗体的方法,所述靶细胞的一个或多个与蛋白糖基化有关的酶存在缺陷,这可以被用于刺激抗体生产。一种与蛋白糖基化有关的酶为 UDP-GlcNAc: -D-甘露糖苷-1,2-N-乙酰葡萄糖氨基转移酶 I (GnTI)。

许多分泌的蛋白以及所述分泌系统的膜整合蛋白为糖蛋白,即它们被聚糖(寡糖)所修饰,所述聚糖与天冬氨酸 N-连接或与丝氨酸、苏氨酸或羟脯氨酸 O-连接。N-糖基化决定着蛋白的正确折叠和稳定性、防止蛋白降解、蛋白构象和识别、蛋白的可溶性、蛋白向胞外空间的分泌以及蛋白的生物活性。

GnTI 是一种 II 型膜整合蛋白,定位到高尔基体中间膜囊,所述 GnTI 可催化由高麦芽糖 N-聚糖转化为复合且混合的结构的第一步。复合 N-聚糖对于发育中的胚胎的存活来说至关重要,因为缺少功能性 GnTI 基因的小鼠会在出生前死亡。然而,复合-N聚糖对于体外培养的细胞的存活来说并非至关重要,因为已经分离到许多缺少 GnTI 活性的突变体。

一个处理纯化的糖蛋白上的异源 N-聚糖的实例是使用衣霉素处理来消除所有的糖基化。因此,已经使用衣霉素处理和四环素诱导的表达来纯化毫克量级的非糖基化视紫红质。然而,该方法并不理想,因为移

除 N-聚糖并不能了解使其在糖蛋白的结构和功能中所起的作用。例如，尽管没有完全了解糖基化在视紫红质结构和功能中的准确作用，但是它明显具有重要作用。先前已经报道了由于未对光感受器进行糖基化导致的信号转导性质的明显缺陷。此外，当在存在衣霉素的情况下进行表达时，具有三个氨基酸变化 (E113Q/E134Q/M257Y) 的视紫红质突变体不能被纯化。Puthalakath et al., *Glycosylation Defect in Lee!Chinese Hamster Ovary Mutant is Due to a Point Mutation in N-Acetylglucosaminyltransferase I Gene*, J. B. C., 271, 27818-27822 (1996) 中描述了其他已经被突变的细胞系，所述文献关于缺少 GnTI 活性的细胞系的教导全部通过引用的方式纳入本说明书。

另一个处理异源 N-聚糖的实例为在细胞内生产所述蛋白，所述细胞中各种与 N-聚糖合成相关的酶 (例如 GlcNAc 转移酶 I) 中的一种存在缺陷。该方法之前已经被用于分离中国仓鼠卵巢 (CHO) 细胞系的不同收集物，所述细胞系由于各种与 N-聚糖合成有关的酶存在缺陷而导致对各种凝集素具有抗性。美国专利申请 2004/0029229 描述了已经经过突变以产生均一糖基化模式的细胞系，所述文献关于已经经过突变以确保 N-聚糖均一的细胞系的教导全部通过引用的方式纳入本说明书。Reeves 等也描述了已经经过突变以产生均一糖基化模式的细胞系 (*Structure and function in rhodopsin: high-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible N-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line*; PNAS 2002 Oct 15; 99 (21): 13419-24. Epub 2002 Oct 7)。如 Koprivova et al., *N-Glycosylation in the Moss Physcomitrella patens is Organized Similarly to that in Higher Plants*, Plant Biology 5 (2003): 582-591 所述，还已经破坏了植物中的 GnTI 基因，所述文献关于已经经过突变以破坏所述 gntI 基因的细胞系的教导全部通过引用的方式纳入本说明书。

例如，本文所述靶细胞可以在糖蛋白上产生均一的糖基化模式。任选地，与对照细胞相比，所述靶细胞具有降低的 GnTI 活性。反义寡核苷酸、RNAi 分子、核酶和 siRNA 分子可用于中断表达。反义寡核苷酸、RNAi 分子、核酶和 siRNA 分子可以单独使用或者可以与其他治疗剂 (例

如抗病毒化合物)组合使用。这类方法还可以与本文公开的构建体和方  
法共同使用。例如,所述靶细胞还可以含有能够表达 GnTI siRNA 的基  
因转移构建体,其中 GnTI siRNA 的表达可以是组成型的或可调节的。

本发明还公开了一种使用选定的蛋白处理受试者的方法,包括给予  
所述受试者按照本文所公开的方法制备的蛋白。所选蛋白的给药方法包  
括但不限于,注射(皮下、表皮、皮内)、粘膜内(例如鼻、直肠和阴道);  
腹膜内、静脉内、口腔或肌内。其他给药模式包括口腔和肺给药、栓剂  
和经皮施用。剂量处理可以是单次给药方案或多次给药方案。

在本文所述的方法中,所述方法包括将外源 DNA 给予到受试者细胞  
内并且被吸收(即基因转导或转染),所公开的核酸可以是用于将所述核  
酸送递到细胞内的载体的形式,由此所述编码抗体的 DNA 片段受到启动  
子的转录调节,这是本领域普通技术人员所公知的。所述载体可以是本  
文所公开的任何载体。可以经由各种机制将所述核酸或载体送递到细胞  
内。作为一个实例,可以经由脂质体进行送递,使用市售的脂质体制品  
例如 LIPOFECTIN、LIPOFECTAMINE (GIBCO-BRL, Inc., Gaithersburg,  
MD)、SUPERFECT (Qiagen, Inc. Hilden, Germany)和 TRANSFECTAM  
(Promega Biotec, Inc., Madison, WI),以及其他按照本领域中的标  
准方法制备的脂质体。此外,可以利用电穿孔将所公开的核酸或载体送  
递到体内,用于进行电穿孔的技术可获自 Genetronics, Inc. (San  
Diego, CA)和利用 SONOPORATION 仪器 (ImaRx Pharmaceutical Corp.,  
Tucson, AZ)。

在一个实例中,本发明公开的重组反转录病毒可被用于进行感染并  
由此将编码广义中和抗体(或其活性片段)的核酸送递到被感染的细胞  
内。

如果需要的话,所述核酸或载体的肠胃外给药通常是通过注射进行  
的。可以将注射剂制成常规形式,以液体溶液或悬液形式、适合于在注  
射前溶于液体形成悬液的固体形式、或者以乳液形式。一种新近改良的  
用于肠胃外给药的方法包括使用缓释或持续释放系统以使得可维持恒  
定剂量。例如参见美国专利 3,610,795,所述文献关于肠胃外给药的方法  
的教导全部通过引用的方式纳入本说明书。关于治疗化合物的适合制  
剂和各种给药途径的其他讨论,例如参见 Remington: The Science and  
Practice of Pharmacy (19th ed.) ed. A. R. Gennaro, Mack Publishing

Company, Easton, PA (1995), 所述文献关于治疗化合物的适合制剂和各种给药途径的教导全部通过引用的方式纳入本说明书。

本文还公开了筛选可调整病毒颗粒形成的药剂的方法。例如, 本发明公开了一种筛选可调整病毒颗粒形成的药剂的方法, 包括将含有第一和第二核酸序列的包装核酸构建体引入细胞中, 其中所述第一核酸序列编码 Gag 蛋白, 并且其中所述第二核酸序列编码 Gag-Pro-Pol 蛋白, 并且其中所述第一和第二核酸序列包括一个或多个可减少移码或翻译连续的突变。此外, 所述第一和第二核酸序列可由同一核苷酸序列的不同编码区表达, 并且所述第一和第二核酸序列可以与所要筛选的药剂有效连接。接下来, 可以将包膜构建体引入所述细胞中, 并且所述包膜构建体可以含有编码包膜糖蛋白的第三核酸序列, 其中所述第三核酸序列与至少一个转录调控元件有效连接。然后在适合于形成病毒颗粒的条件下培养所述细胞。然后可以对所述病毒颗粒进行检测; 并且与对照相比, 在存在所要筛选的药剂情况下, 病毒颗粒数量的增加或减少表明所述药剂可调整病毒颗粒的形成。所述对照培养物可以是单独的培养物或者可以与给予所述药剂之前或之后的同一培养物。还可以将含有调节序列的调节子构建体引入所述细胞中, 其中所述可调节元件与至少一个转录调控元件有效连接。本说明书通篇对各种可调节的转录调控元件和调节子序列进行了讨论。例如, 所述转录调控元件可以是 CMV 启动子, 并且所述可调节元件可以是 tetR 或 tetA。

可以使用各种本领域中公知的筛选标记来筛选阳性包装细胞转化体(即已经接纳并整合了所述反转录病毒载体的细胞)。例如, 所述构建体中可以包括下列标记基因: 绿色荧光蛋白(GFP)基因、潮霉素抗性(Hyg)基因、新霉素抗性(Neo)基因和 $\beta$ -半乳糖苷酶( $\beta$ -gal)基因; 并且使用例如酶活性或药物抗性测定法来对上述基因进行测定。或者, 可以就 Goff et al. (1981) *J. Virol.* 38:239 所述的、作为病毒蛋白生产的量度的反转录酶(RT)活性对细胞进行测量。

可以使用类似的测定法来测试包装细胞所产生的不期望的、有复制能力的辅助病毒。例如, 本文所述的构建体可以含有标记基因, 例如本文所述的那些标记基因。在使用本文所述的包装构建体对靶细胞进行瞬时转染之后, 将包装细胞(含有至少一种本文所公开的包装构建体的细胞)和其他非包装细胞共同传代培养。可以使用携带所述标记基因的本

发明的重组的复制缺陷型构建体来感染这些非包装细胞。然而，由于上述非包装细胞不含产生病毒颗粒所必需的基因(例如 *gag*、*pol* 和 *env* 基因)，所以当与其他细胞共同传代培养时，上述非包装细胞应该不能感染所述其他细胞。当将上述其他细胞与所述非包装细胞共同传代培养时，如果所述其他细胞中的标记基因呈现阳性反应，则说明已经产生了不期望的、有复制能力的辅助病毒。

因此，为测试不期望的辅助病毒的产生情况，可以将本发明的包装细胞与第一细胞系(例如 NIH3T3 细胞)共同传代培养，再将所述第一细胞系与第二细胞系共同传代培养，测试所述第二细胞系中标记基因或 RT 活性的存在情况，所述标记基因或 RT 活性的存在情况表明了有复制能力的辅助反转录病毒的存在情况。可以使用例如 FACS 染色和酶活性测定法来测定标记基因，所述技术在本领域中是公知的。

本发明还公开了制备转基因动物的方法。具体而言，本发明公开了制备转基因动物的方法，包括：将按照本文所公开的方法制备的病毒颗粒引入合子中；允许所述合子发育至足月；获得一种其基因组中含有能够表达所述目的基因的核酸构建体的动物；将所述动物与非转基因动物杂交以获得 F1 子代，并且选择其基因组中含有所述能够表达目的序列或含有目的序列的核酸构建体的动物，其中所述动物可表达选定的目的序列或含有选定的目的序列。本发明还公开了按照本文所公开的方法制备的转基因动物。

可以将本发明的病毒颗粒引入到动物的基因组内，以产生转基因的非人动物来用于实施本发明所述的方法。可筛选标记还可以被用作鉴别那些含有目的序列的动物的报道分子。例如，发光蛋白可以被用作报道分子，通常使用完整的、活的、非人转基因动物例如活的转基因啮齿动物(例如小鼠或大鼠)来进行显像。可以使用任何可被用于将核酸引入所选的动物细胞内的技术(例如，"Transgenic Animal Technology: A Laboratory Handbook," by Carl A. Pinkert, (Editor) First Edition, Academic Press; ISBN: 0125571658; "Manipulating the Mouse Embryo: A Laboratory Manual," Brigid Hogan, et al., ISBN: 0879693843, Publisher: Cold Spring Harbor Laboratory Press, Pub. Date: September 1999, Second Edition, 上述参考文献关于可被用于将核酸引入动物细胞内的技术的教导全部通过引用的方式纳入本

说明书)。各种转化技术在本领域中是公知的。可被用于将核酸引入所选的动物细胞内的方法包括但不限于以下方法：

(i) 直接显微注射到细胞核中：可以使用微量吸管将病毒颗粒直接显微注射到动物细胞核内以机械地转移所述重组 DNA。该方法的优势在于不会使所述 DNA 与除细胞核之外的细胞区室接触并且可以高频率地产生稳定的重组体。参见 Capecchi, M., *Cell* 22: 479-488 (1980), 所述参考文献关于直接显微注射到细胞核中的教导全部通过引用的方式纳入本说明书。

例如，在合子形成雄原核膜之后并且在雄原核膜被合子雌原核加工之前，尽快将所述病毒颗粒显微注射到合子的早期雄原核内。因此，应当在所述雄原核和雌原核被完全分开并且二者都位于距离细胞膜很近的位置上时，采用本方法的显微注射。参见，例如 Wagner 等的美国专利 4,873,191(于 1989 年 10 月 10 日授权)；和 Richa, J., (2001) "Production of Transgenic Mice," *Mol. Biotech.*, 17: 261-8, 上述参考文献关于直接显微注射到合子的早期雄原核内的教导全部通过引用的方式纳入本说明书。

(ii) ES 细胞转染：本发明的病毒颗粒还可以被引入到胚胎干细胞 ("ES") 内。鉴别与靶载体发生同源重组的 ES 细胞克隆，然后生产 ES 细胞-小鼠嵌合体。通过半合子嵌合动物交配来产生纯合动物。所述方法在例如 Koller, B. H. and Smithies, O., (1992) "Altering genes in animals by gene targeting", *Ann. R. Imm* 10: 705-30 中有所描述。

(iii) 电穿孔：还可以利用电穿孔将本发明的病毒颗粒引入到动物细胞内。在该技术中，在存在病毒颗粒的情况下对动物细胞进行电穿孔。高场强的电脉冲可逆地透化生物膜，使得可以引入所述核酸。在电穿孔过程中产生的孔使得可以摄取大分子，例如核酸。所述方法在例如 Potter, H., et al., *Proc. Nat'l. Acad. Sci. U. S. A.* 81: 7161-7165 (1984)；和 Sambrook, ch. 16 中有所描述，所述参考文献关于利用电穿孔将核酸或病毒颗粒引入到动物细胞内的教导全部通过引用的方式纳入本说明书。

(iv) 磷酸钙沉淀：还可以利用其它直接摄取方法(例如使用磷酸钙)将所述病毒颗粒转移到细胞内。参见，例如 Graham, F., and A. Van der

Eb, *Virology* 52:456-467 (1973); 和 Sambrook, ch.16, 所述文献关于利用磷酸钙沉淀将核酸或病毒颗粒引入到动物细胞内的教导全部通过引用的方式纳入本说明书。

(v) 脂质体: 还可以使用下述方法将核酸引入到动物细胞内: 将核酸包装到人工膜囊(脂质体)中, 然后将所述脂质体与所述靶细胞膜融合。参见 Mannino, R. and S. Gould-Fogerite, *BioTechniques*, 6:682 (1988), 所述参考文献关于使用脂质体将核酸引入到动物细胞内的教导全部通过引用的方式纳入本说明书。

(vi) 使用聚凝胺(Polybrene)或 DEAE-葡聚糖进行转染: 这些技术在 Sambrook, ch.16 中有所描述, 所述参考文献关于使用聚凝胺或 DEAE-葡聚糖将核酸引入到动物细胞内的教导全部通过引用的方式纳入本说明书。

(vii) 原生质体融合: 原生质体融合通常包括将携带有大量目的质粒的细菌原生质体与所培养的动物细胞融合, 通常由使用聚乙二醇的处理来介导(Rassoulzadegan, M., et al., *Nature*, 295:257 (1982), 所述参考文献关于使用原生质体融合将核酸引入到动物细胞内的教导全部通过引用的方式纳入本说明书)。

(iix) 弹道侵入(Ballistic Penetration): 另一种引入核酸区段的方法是利用小颗粒进行高速弹道侵入, 所述核酸位于小珠或小颗粒的基质内或位于其表面上, 参见 Klein, et al., *Nature*, 327, 70-73, 1987, 所述参考文献关于利用弹道侵入将核酸引入到动物细胞内的教导全部通过引用的方式纳入本说明书。

电穿孔的优势在于容易实施, 并且已经得到了广泛的应用, 但是在电穿孔过程中大部分靶细胞可能会被杀死。因此, 对于敏感细胞或只能少量获得的细胞来说, 优选直接显微注射到细胞内。此外, 当高效率的核酸引入特别重要时, 例如不使用可筛选标记的转化(如上所述), 直接显微注射到细胞核内就是一种有利的方法, 因为通常 5-25%的靶细胞将会具有稳定整合的经显微注射的核酸。

此外, 本发明公开了含有目的序列的转基因动物。例如, 本发明公开了表达 KISS-1、FOX P3、NF  $\kappa$   $\beta$ 、微小 RNA 223 或 Cre 重组酶的转基因动物。

本发明还公开了含有本文所述的基因转移构建体的转基因动物。本

发明还公开了按照本文所公开的方法制备的转基因动物。

在本申请通篇中，引用了各种出版物。这些出版物的公开内容全部通过引用的方式纳入本申请，以便更全面地描述本文所述的化合物、组合物和方法。

可以对本文所述的化合物、组合物和方法进行各种修改和变化。考虑到对本文所公开的化合物、组合物和方法的说明和实践，本文所述的化合物、组合物和方法的其他方面将是显而易见的。所述说明和实例意在举例说明。

### 实施例

给出下列实施例以便为本领域普通技术人员提供关于如何制备并评估本文所述和所要求保护的化合物、组合物、制品、设备和/或方法的完整公开和描述，并且意在纯粹地进行举例说明而并非意在对发明人所认为的本发明的范围进行限制。已经努力确保数字(例如数量、温度等)的准确性，但是应该考虑到会存在一些错误和偏差。除非另有说明，份是指重量份、温度以为℃单位或指室温，压力是指大气压或接近大气压。存在许多的反应条件(例如组分浓度、所需溶剂、溶剂混合物、温度、压力以及其他反应范围和条件)的变化和组合，所述反应条件可被用于优化通过所述工艺获得的产物的纯度和产率。为优化这类工艺条件，仅需要进行合理的和常规的实验。

#### 实施例 1

##### 构建基于四环素的可诱导和可逆转的单慢病毒载体

构建了基于四环素的可诱导和可逆转的单基因转移载体，以驱动目的序列 eGFP 的表达。首先，利用 PCR 由 pEF4/His (Invitrogen) 扩增 1.2kb 的人 EF1-a 启动子，并且使用 EcoRI 和 BamHI 限制性内切酶将其克隆到 pHRCMVeGFP/bias 中。将所得的载体命名为 pHREFeGFP/bias。接下来，对一能够编码四环素抑制子的序列进行密码子优化并且将其与 SV40 核定位信号连接。然后使用 NcoI 和 XhoI 限制性内切酶将所述与 SV40 核定位信号连接的经过密码子优化的四环素抑制子的基因克隆到 pHREFeGFP/bias 中，用所述基因置换 eGFP。将所得的载体命名为 pHREFtet/bias。然后，利用 PCR 扩增 500 bp 的人 CMV 启动子，将两条

tet 启动子序列引入到 3'CMV 启动子中。使用 ClaI 和 EcoRI 限制性内切酶将所述 PCR 片段克隆到 pHREtet/blas 中。将所得的载体命名为 pHRCMV02(R)EFtet/blas。这样所述 CMV 启动子的方向会被逆转。然后将含有牛生长激素多腺苷酸化信号的 EGFP 片段克隆到 pHRCMV02(R)EFtet/blas 中, 所述片段就会被 CMV 启动子所调控。将所得的载体命名为 pHReGFP02/EFtet/blas。接下来, 利用 PCR 由 pUB6/V5-His (Invitrogen) 扩增 1.2 kb 的人泛素 6 启动子并且使用 EcoRI 和 NcoI 限制性内切酶将其克隆到 pHReGFP02/EFtet/blas 中。将所得的载体命名为 pHReGFP02/UB6tet/blas。该步骤之后, 从 pDRIVE-CAG (Invivogen) 获得 1.6 kb 的含有 300 bp 的 5'人 CMV 启动子序列的 CAG 启动子和 1.2 kb 的鸡  $\beta$ -肌动蛋白启动子, 并且将其克隆到 pHReGFP02/EFtet/blas 中。利用 SnaBI 和 NcoI 限制性内切酶切割 pHReGFP02/EFtet/blas, 将 5' CMV 序列和 EF 启动子移除并且替换为 CAG 启动子。将所得的载体命名为 pHReGFP02/CAGtet/blas。

## 实施例 2

生成高效价的基于四环素的可诱导和可逆转的单病毒颗粒

用包装构建体、包膜构建体和不同基因转移构建体(包括 pHReGFP02/EFtet/blas、pHReGFP02/CAGtet/blas; pHReGFP02/UB6tet/blas 和 pHReGFP02/CAGtet/blas)共转染 293Y 细胞以生产不同类型的诱导型病毒颗粒。使用荧光显微术测量由共转染得到的所述病毒颗粒的效价, 以测定 HeLa 细胞内的 eGFP 表达。来自被转染的细胞的上清液的效价为  $1-4 \times 10^6$ /ml, 而浓缩的上清液的效价为  $2-10 \times 10^8$ /ml(高了 400 倍)。

## 实施例 3

紧密调节的诱导型单慢病毒载体

使用 100  $\mu$ l 源自 pHReGFP02/EFtet/blas(效价为  $2.5 \times 10^6$ /ml)的病毒颗粒上清液感染小鼠 T 细胞系 ( $4 \times 10^4$ )。第二天, 将感染的细胞分为以下两组: 第一组, 在含有 0.1  $\mu$ g DOX/ml 的培养基中培养; 第二组, 在不含 DOX 的培养基中培养。感染三天后, 利用 FACS 分析对所述细胞进行分析以确定 GFP 的表达水平。对第一组的分析显示 GFP 表达信号的

平均强度为 16,195, 与第二组的相比增加了 44.2 倍。

#### 实施例 4

单慢病毒载体对 DOX 高度敏感并且可迅速诱导基因表达

为测定诱导所述单载体系统内基因表达所需的 DOX 浓度, 将不同浓度的 DOX 加入到被病毒颗粒感染的细胞内, 所述病毒颗粒源自 pHReGFP02/EFtet/blas (效价为  $2.5 \times 10^6$ /ml)。利用荧光显微术检测 GFP 表达。图 4A 示出了 15 ng 的 DOX 足以在 48 小时内诱导 GFP 表达。

#### 实施例 5

组成型启动子活性会显著影响诱导型单慢病毒载体的诱导型启动子活性

为确定四环素抑制子的表达水平是否会影响基因转移载体的可诱导性, 将不同启动子克隆到基因转移载体中以驱动四环素抑制子的表达。所用的启动子为人 EF-1 $\alpha$  启动子 (pHReGFP02/EF-1 $\alpha$ /blas)、CAG 启动子 (pHReGFP02/CAGtet/blas) 和人泛素 6 启动子 (pHReGFP02/UB6tet/blas)。EF-1 $\alpha$  为三种启动子中最强的启动子, 而人泛素 6 启动子为最弱的。用来自所述 293T 细胞的病毒颗粒感染小鼠 T 细胞系。使用抗生素杀稻瘟素来筛选阳性感染细胞。经过 3 天的筛选之后, 将被感染的细胞划分为以下两组: 第一组, 在含有 DOX 的培养基中培养; 第二组, 在不含 DOX 的培养基中培养。利用 FACS 分析对被感染的细胞进行分析以测量在存在或不存在 DOX 的情况下培养三天后的 GFP 表达。下表示出了使用不同载体诱导的 GFP 的表达水平。

表 1

不同启动子	DOX (-)	DOX (+)	诱导
EF-1 $\alpha$	133	15,071	113 倍
CAG	157	7,071	45 倍
UB6	263	4,550	17 倍

含有 EF-1 $\alpha$  启动子的构建体可产生最低的 eGFP 基础表达水平, 然而, 它也产生了最高的 eGFP 诱导表达水平。对于 EF-1 $\alpha$  启动子构建体

来说, eGFP 表达的诱导水平超过 100 倍。人泛素 6 启动子产生最高的 eGFP 基础表达水平和最低的 eGFP 诱导表达水平。对于人泛素 6 启动子构建体来说, eGFP 表达的诱导水平约为 17 倍。

可以在两个水平上观察到组成型启动子的作用, 首先该启动子会影响基础泄漏水平, 其次所述组成型启动子会影响所述目的基因(本文中为 eGFP)的最高表达水平。所述强的组成型启动子可以驱动高水平的四环素抑制子表达, 在不存在 DOX 的情况下, 所述高水平的四环素抑制子表达可以促进和调控基础泄漏水平。与其他类型细胞相比(例如 HeLa 细胞), 在 T 细胞中所述基于 CMV 启动子的诱导型启动子的活性较弱。

当将一种与调节子构建体相连的强组成型启动子(例如与 tetR 相连的 EF-1 $\alpha$ )应用于所述诱导型系统中时, 这类强组成型启动子可以刺激基于 CMV 的诱导型启动子的活性。当将所述与目的基因有效连接的诱导型启动子再与一种可驱动调节子构建体表达的强组成型启动子有效连接时, 所述目的基因的表达在 T 细胞中可变得非常活跃。

### 实施例 6

使用诱导型单慢病毒载体来产生 eGFP 转基因小鼠

如前所述, 使用孕马血清和人体绒毛膜促性腺激素(HCG)的组合来使年龄在 22-24 天之间的雌性小鼠(B6 株)超量排卵(B. Hogan, R. Beddington, F. Costantini, E. Lacy, *Manipulating the Mouse Embryo* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1994)。然后, 按照 B. Hogan, R. Beddington, F. Costantini, E. Lacy, *Manipulating the Mouse Embryo* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1994 来所述收集提供者胚胎。在使用显微注射系统(CellTram, Eppendorf GmbH, Hamburg, Germany)收集的同一天, 将使用上文所述方法制备的浓缩病毒颗粒(效价为大约  $2 \times 10^8$ /ml)送递到单细胞期胚胎中。使用显微操作器来引导所述吸管, 推动微量吸管穿过透明带进入卵黄周隙, 并且将 10 p1-100 p1 的病毒储液送递到所述胚胎内。将所述被感染的胚胎在 KSOM-AA 中培养过夜并且按照 B. Hogan, R. Beddington, F. Costantini, E. Lacy, *Manipulating the Mouse Embryo* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, (1994)所述将这些二细胞期胚胎转

移到假孕雌鼠(10 周龄 CD1)体内, 所述参考文献关于制备转基因动物的方法的教导全部通过引用的方式纳入本说明书。鉴别了源自 pHReGFP02/EFtet 的 11 个建立者(本文中称为 EF-建立者)和源自 pHReGFP02/CAGtet 的 8 个建立者(本文中称为 CAG-建立者)。

通过移除杀稻瘟素基因来由 pHReGFP02/Eftet/blas 和 pHReGFP02/CAGtet/blas 制备两种形式的 pHReGFP02/EFtet 和 pHReGFP02/CAGtet, 以防止对所述转基因小鼠产生有害作用的可能性。从三周龄的建立者中提取基因组 DNA 并且利用 PCR 和 RNA 印迹进行分析, 以确定是否存在阳性转基因小鼠和被整合的构建体的拷贝数量。表 2 示出了阳性转基因小鼠的数量。

表 2

不同启动子	建立者编号	PCR 阳性的比例	单拷贝的比例
EF-1 $\alpha$	11	63.6% (7/11)	42.8% (3/7)
CAG	8	61.5% (5/8)	40% (2/5)

通过使用一对靶向四环素抑制子基因的引物—SEQ ID NO: 16 和 SEQ ID NO: 17 的 PCR 分析来鉴别阳性转基因小鼠。这两种构建体可以产生相似阳性比例的转基因小鼠, 因为这两种构建体的效价都大约为  $2 \times 10^8$ /ml。RNA 印迹显示在 EF-建立者组中存在三种单拷贝建立者, 而在 CAG-建立者组中鉴别了两种单拷贝建立者。在这两组中, 半数以上的阳性转基因小鼠具有两个或多个拷贝(范围为 1-4)。与之前的报道相比, 其他研究人员使用 5 倍高的效价( $10 \times 10^8$  /ml)来产生建立者小鼠, 所述建立者小鼠在这两组中具有两个或多个拷贝(范围为 1-20)。因此, 本方法提供了一种更加有效的方法。

### 实施例 7

使用含有 DOX 的饮用水在所述转基因小鼠体内诱导 eGFP 表达

为确定所述诱导型构建体是否可以在体内诱导 eGFP 表达, 给所述转基因小鼠喂食含有  $100 \mu\text{g}/\text{ml}$  DOX 的饮用水。利用荧光显微术和 FACS 分析法来分析喂食 DOX 前后所述转基因小鼠体内(脚爪)和 PBMC 中的 GFP 表达。所有 12 只阳性小鼠均能在 PBMC 和体内(脚爪)诱导 eGFP 表达,

但是这些小鼠的诱导水平之间存在差异。在喂食 DOX 之前, 在所有转基因小鼠体内均检测到 eGFP 表达, 但是这些小鼠的表达水平之间存在差异。在给所述转基因小鼠喂食 DOX 之后, eGFP 表达显著升高。与使用源自含有 EF-1 $\alpha$  启动子的构建体的病毒颗粒感染的转基因小鼠相比, 使用源自含有 CAG 启动子的构建体的病毒颗粒感染的转基因动物可产生最高水平的诱导。这些数据与上文所述的体外实验的结果不同。

### 实施例 8

转基因小鼠体内 (指) eGFP 的可视表达是可诱导的和可逆转的

为确定所述基因转移构建体中含有的转基因是否会在全身表达, 使用上述的含有由 CAG 启动子驱动的 eGFP 基因的基因转移构建体, 如上所述地产生转基因动物。当 eGFP 处于 CAG 启动子的调控之下时, eGFP 可能会在全身表达。为确定 eGFP 是否会在含有上述基因转移构建体的转基因动物全身表达, 利用荧光显微术分析所述转基因动物的指。在本研究中, 选择四只上文所述的 CAG-建立者 (分别命名为 CAG-建立者 1#、2#、6#和 7#) 来进行分析。在添加 DOX 之后, 在 CAG-建立者-2#和 6#体内观察到 eGFP 表达, 表明这两只小鼠体内的 GFP 表达可以受到 DOX 的紧密调控。添加 DOX 之后, 在所测试的转基因建立者中, CAG-建立者-1#小鼠体内的 eGFP 表达的显像最亮, 这说明该小鼠的指中表达的 eGFP 水平最低, 并未受到 DOX 的诱导。CAG-建立者-7#在响应于 DOX 添加而表达 eGFP 时表现出某种延迟, 并且与其他 CAG-建立者的 eGFP 表达强度相比, CAG-建立者-7#的总表达强度较弱。

不再给予 CAG-建立者 DOX 12 天后, 再次利用荧光显微术来分析所述小鼠的指。除 CAG-建立者-1#外, 所述指内 GFP 表达强度急剧下降到类似于诱导之前的表达水平。结果表明这些转基因小鼠体内的 GFP 表达是可诱导的和可逆转的, 这取决于是否存在 DOX。

### 实施例 9

转基因小鼠血细胞内的 GFP 表达可被 DOX 诱导并且是可逆的

在下述 4 个不同时间点检测 CAG-建立者小鼠血细胞内的 GFP 表达:  
(1) 在给小鼠喂食含 DOX (0.1 mg/ml DOX) 饮用水之前; (2) 给小鼠喂食含 DOX (0.1 mg/ml DOX) 饮用水 12 天后; (3) 从时间点 (2) 开始从饮用水中

移除 DOX 12 天后,然后在再次给时间点 (3) 的小鼠喂食含 DOX (0.1 mg/ml DOX) 饮用水 1 和 2 天后。CAG-建立者 1#和 CAG-建立者 2#小鼠血细胞内的 GFP 表达受到 DOX 的紧密调控。此外,当停止给予 DOX 时,所述 GFP 表达可被逆转。此外,所测试的血细胞内的 GFP 表达水平可以回到本底水平(添加 DOX 之前的水平)。该数据表明单慢病毒载体系统可以诱导所述基因转移构建体内序列的表达并且可以使所述表达逆转。

### 实施例 10

#### 利用 DOX 诱导多种器官内的 GFP 表达

为确定上文所述慢病毒系统是否能够在动物全身表达目的序列,检测了 GFP 的表达。在生成 CAG-建立者-2#后,将所述动物解剖并且使用荧光显微术逐个分析各器官。在所述 Tg 小鼠(CAG 建立者 2#)的骨骼和肌肉中观察到高 GFP 表达,但是在正常小鼠体内未观察到 GFP 表达。在所述 Tg 小鼠的心、肺、肝、肾、脾和肠内观察到高 GFP 表达,而所述 Tg 小鼠的脑内 GFP 表达较弱。该数据表明所述转基因小鼠全身的表达均可由 DOX 诱导,但是在对不同器官的诱导水平之间存在差异。

### 实施例 11

#### 确定饮用水中诱导 GFP 表达所需的 DOX 浓度

一项之前的研究报道称用于在含有 tet 调节系统的转基因小鼠体内进行诱导型基因表达所需的该小鼠饮用水中的 DOX 浓度为 0.1-10 mg/ml。为确定用于在上文所述的转基因小鼠体内进行 GFP 诱导型表达所需的 DOX 浓度,将源自 CAG-建立者 6#的 F1 小鼠分为四组,所述四组分别被喂食含有不同浓度 DOX 的饮用水,包括 0  $\mu$ g/ml(第一组)、4  $\mu$ g/ml(第二组)、20  $\mu$ g/ml(第三组)和 100  $\mu$ g/ml(第四组)。在喂食所述小鼠 DOX 0、1、2、3、5 和 18 天后,通过在 UV 光下使在所述转基因小鼠的指内表达的 GFP 显影来监测 GFP 表达。观察整个实验过程中所测试小鼠的指内的荧光信号强度。在 DOX 喂养 1 天后,第三组和第四组小鼠开始表达 GFP,表明 DOX 会通过饮用水迅速诱导所述基因表达。在 DOX 喂养 5 天后,第三组和第四组的荧光信号强度表明其 GFP 表达水平达到最高。此外,给第二组的小鼠喂食 4  $\mu$ g/ml DOX 足以诱导 GFP 表达,然而,与第三组和第四组小鼠相比,诱导被延迟且强度较弱。还应该注意

到, 荧光信号强度表现为剂量依赖性方式。

使用 FACS 分离并且量化表达 GFP 的阳性血细胞。图 5 示出了在给小鼠喂食 DOX 之前和 18 天后对血细胞中 GFP 表达的 FACS 分析的结果。对第二组、第三组和第四组的所有小鼠而言, 在给所述小鼠喂食 DOX 后, 表达 GFP 的细胞的数量和强度均有所增加。

为确定 DOX 的药物代谢动力学, 第四组小鼠 43% 的血细胞表达 GFP (DOX 喂养 18 天后), 使用该水平作为 100% 诱导的阈值。图 6 示出了第 2、3 和 4 组小鼠的血细胞中 GFP 表达的诱导动力学。该数据表明 DOX 可以诱导所公开转基因小鼠的血液和指内的 GFP 表达, 并且所述诱导水平为剂量依赖性的。

### 实施例 12

构建基于四环素的可诱导和可逆转的单慢病毒载体以表达 shRNA

使用一条含有 NotI 的正义引物 (5'-GCGGCCGCAATTCATATTGCATGTCGCTATGT-3') (SEQ ID NO: 18) 和一条反义引物 (5'-GAATTCGCGGATCCTCTCTATCACTGATAGGGACTTATAAGTCTCTATCACTGATAGGGATTTACGTTTATGGTGA-3') (SEQ ID NO: 19), 利用 PCR 从 Hela 细胞中扩增人 H1 启动子, 其中所述反义引物含有位于 TATA 框上游的一个最小 19bp 的 tetO 序列和位于 TATA 框下游的另一个 tetO 序列。然后将含有人 H1 启动子的 PCR 片段克隆到 pHREFtet/blas 中。所得的载体被命名为 pHRhH1tetOEFtet/blas。

使用一条含有 NotI 的正义引物 (5'-GGCGGCCGCATATGACTAGTCATGCAAATTACGCGCT-3') (SEQ ID NO: 20) 和一条反义引物 (5'-GAATTCTGGATCCTCTCTATCACTGATAGGGATTATAAGTCTCTATCACTGATAGGGATTTTACGTTTAGGGTGATTT-3') (SEQ ID NO: 21) 利用 PCR 从 3T3 细胞系中扩增小鼠 H1 启动子, 其中所述反义引物含有位于 TATA 框上游的一个最小 19bp 的 tetO 序列和位于 TATA 框下游的另一个 tetO 序列。然后将含有小鼠 H1 启动子的 PCR 片段克隆到 pHREFtet/blas 中。所得的载体被命名为 pHRmH1tetOEFtet/blas。

用于上述实验的目的序列为设计用于靶向 eGFP 编码区 (从 nt126 至 144) 的 shRNA。使用正义引物 (5'-GATCCAGCTGACCCTGAAGTTCATCTTCAAGAGAGATGAACTTCAGGGTCAGCTTTTTGG-3') (SEQ ID NO: 22) 和反

义引物 (5'-AATCCAAAAAGCTGACCCTGAAGTTCATCTCTCTTGAAGATGAACTTCA GGGTCAGCTG-3') (SEQ ID NO: 23) 来生成 shRNA, 所述 shRNA 彼此退火并被克隆至 pHRhH1tet0EFtet/blas 和 pHRmH1tet0EFtet/blas 中, 所述载体之前已经用 BamHI 和 EcoRI 限制性内切酶消化过。将所得载体分别命名为 pHRhH1GFPi (126) EFtet/blas 和 pHRmH1GFPi (126) EFtet/blas。

### 实施例 13

利用小鼠 H1 诱导型启动子有效沉默基因表达

使用源自 pHRhH1GFPi (126) EFtet/blas 和 pHRmH1GFPi (126) EFtet/blas 的病毒颗粒感染能够表达 eGFP 的不同细胞系 [HeLa 细胞、CEM-SS 细胞 (人 T 细胞系) 和小鼠 T 细胞系]。感染 2 天后, 通过使细胞与杀稻瘟素接触 3 天来使用抗生素 (10  $\mu$ g/ml 杀稻瘟素) 筛选含有慢病毒载体的细胞。将阳性细胞分为两组: 第一组, 在含有 0.5  $\mu$ g/ml DOX 的培养基中培养; 第二组, 在不含 DOX 的培养基中培养。在 DOX 诱导 7 天后, 使用 FACS 分析第一组和第二组的细胞。图 7 示出了人和小鼠 H1 启动子均能表达 shRNA, 这本身又可以有效地沉默 HeLa 细胞中的 eGFP 表达。eGFP 表达的抑制最高可达 50 倍。

还应该注意到的是, 人 H1 启动子在沉默人 T 细胞内的 eGFP 表达方面效率较低 (1-2 倍), 而小鼠 H1 启动子可将 eGFP 表达降低最高达 10 倍 (图 8)。在小鼠 T 细胞系内, 利用小鼠 H1 启动子将 eGFP 表达降低至本底水平, 而人 H1 启动子可将 eGFP 表达降低 4 倍。该数据表明使用上文所述病毒颗粒感染的细胞内的 eGFP 表达水平会受到 DOX 的紧密调控。

### 实施例 14

利用单慢病毒载体可诱导地沉默内源蛋白 CXCR4

为确定所述诱导型单慢病毒载体是否可以降低内源蛋白表达, 构建一种含有靶向于小鼠 CXCR4 mRNA 的 shRNA 的单慢病毒载体。使用正义引物 (5'-GATCCAGGATGGTGGTGTTC AATTCCTTCAAGAGAGGA ATTGAAACACCACCATCCTTTTTGG-3') (SEQ ID NO: 24) 和反义引物 (5'-AATCCAAAAAGGATGGTGGTGTTC AATTCCTTCTTGAAGGAATTGAAACACCAC CATCCTG-3') (SEQ ID NO: 25) 来生成设计用于靶向 CXCR4 编码区 (从

nt652 至 702) 的 shRNA, 所述 shRNA 彼此退火并被克隆至 pHRmH1tet0EFtet/blas 中, 所述载体之前已经用 BamHI 和 EcoRI 限制性内切酶消化过, 并且利用 eGFP 置换所述杀稻瘟素抗性标记。将所得载体命名为 pHRmH1GFPi (682) EFtet/GFP。

使用源自 pHRmH1GFPi (682) EFtet/GFP 的病毒颗粒感染两组小鼠 T 细胞系。第一组使用效价为  $5 \times 10^6$  IU/ml 的病毒颗粒感染; 第二组使用效价为  $5 \times 10^7$  IU/ml 的病毒颗粒感染。感染 3 天后, 将每组细胞再分为两个亚组。将所述亚组分别在含有  $0.5 \mu\text{g/ml}$  DOX 的培养基 (1a 组和 2a 组) 或者在不含 DOX 的培养基 (1b 组和 2b 组) 中培养。培养 5 天后, 使用与 PE 缀合的抗-CXCR4 抗体 (BD Pharmgen) 将所有的细胞染色。然后, 利用 FACS 分析上述被染色的细胞。使用  $5 \times 10^6$  IU/ml 的源自 pHRmH1GFPi (682) EFtet/GFP 的病毒颗粒感染的上述细胞中有 85% 会表达 GFP, 而使用  $5 \times 10^7$  IU/ml 的所述病毒颗粒感染的细胞中有 98% 会表达 GFP (图 9)。在存在 DOX 的情况下, 1a 组的细胞使 CXCR4 的强度降低 60%, 而 2a 组的细胞使 CXCR4 的强度降低 85%。这些数据表明所述慢病毒载体可以诱导 shRNA 活性, 所述 shRNA 又会降低内源蛋白表达。此外, 上述数据表明多拷贝的所述整合载体可以诱发高水平的基因沉默。

### 实施例 15

单慢病毒载体可以诱导性地表达 shRNA 以沉默转基因小鼠体内的基因表达。

为确定所述单慢病毒载体是否可以表达 shRNA 以降低动物体内的蛋白表达, 选择来自 Jackson 实验室的 eGFP 转基因小鼠的纯合品系作为靶标。当使用 eGFP 转基因小鼠的纯合品系时, 可以测量所述 shRNA 对 eGFP 蛋白表达的影响。为产生用于本实验的慢病毒载体, 使用 CAG 启动子置换所述 pHRmH1GFPi (126) EFtet/blas 质粒中的 EF-1a 启动子, 以提高所述转基因小鼠体内单慢病毒载体的基因表达能力。将所得载体命名为 pHRmH1GFPi (126) CAGtet/blas。在细胞培养物中, 源自 pHRmH1GFPi (126) CAGtet/blas 的载体与来源于 pHRmH1GFPi (126) EFtet/blas 的载体相似, 均可表达能够诱导性地沉默 GFP 表达的 shRNA。

使用微注射系统 (上文所述) 将  $2 \times 10^8$  IU/ml 的源自

pHRmH1GFPi(126)EFtet/blas 构建体或 pHRmH1GFPi(126)CAGtet/blas 构建体的病毒颗粒送递到 GFP 小鼠纯合品系的单细胞期胚胎中。将这样获得的转基因小鼠命名为 GFP/CAG-建立者小鼠。第二天,将双细胞期胚胎植入到 CD1 代孕母鼠 (foster mother) 体内。在 11 只使用源自 pHRmH1GFPi(126)EFtet/blas 的病毒颗粒感染的小鼠中,有 5 只为转基因阳性,这一点得到了 PCR 分析的确认;而在 9 只使用源自 pHRmH1GFPi(126)CAGtet/blas 的慢病毒载体感染的小鼠中,有 4 只为转基因阳性,这一点得到了 PCR 分析的确认。因此,对于所述两种慢病毒载体的每一种来说,利用 PCR 分析推断出的转基因阳性小鼠的比例为大约 40%。

将两只被鉴别为转基因表达阳性的小鼠——GFP/CAG-建立者 6#和 GFP/CAG-建立者 9#饲养 4 周。然后收集取自 4 周龄转基因小鼠尾静脉的血液,并且利用 FACS 分析法分析所述血细胞中的 GFP 表达水平。然后通过饮用水给该小鼠喂食 DOX 以诱导所述 shRNA 的表达。在喂食 DOX 5 日和 10 日后,再次收集取自 4 周龄转基因小鼠尾静脉的血液。如上文所述地利用 FACS 分析法分析所述血细胞中的 GFP 表达水平。在被喂食 DOX 之后,四只源自 pHRmH1GFPi(126)CAGtet/blas 的阳性转基因小鼠中有两只降低了 GFP 表达水平(参见图 12 中的 GFP/CAG-建立者 6#和 GFP/CAG-建立者 9#)。所述血细胞中 GFP 表达水平的降低并不一致,一些细胞中 GFP 表达水平降低最高达 10 倍,而一些细胞没有变化。此外,五只使用源自 pHRmH1GFPi(126)EFtet/blas 构建体的病毒颗粒感染的阳性转基因小鼠的 GFP 表达水平全部没有发生变化。然而,诱导前 GFP/CAG-建立者 6#和 GFP/CAG-建立者 9#的 GFP 表达水平与非转基因小鼠(不含 shRNA 载体)相同,表明所述转基因动物体内的 H1 诱导型启动子会受到 DOX 的紧密调控(图 11)。

### 实施例 16

单慢病毒载体可以诱导性地表达 shRNA 以沉默转基因小鼠体内的基因表达。

为确定经过种系传递之后所述诱导型单慢病毒载体是否会保持功能,使两只阳性转基因小鼠——GFP/CAG-建立者 6#(雌性)和 GFP/CAG-建立者 9#(雄性)交配。利用 DNA 印迹分析法分析所有 F1 小鼠以测定慢

病毒载体的整合拷贝的数量。在 11 只 F1 小鼠中，有 2 只含有两个整合拷贝的慢病毒载体。其他 F1 小鼠含有一个整合拷贝 (5 只小鼠) 或者为整合阴性小鼠 (4 只小鼠)。

为确定所述含有慢病毒载体的 F1 小鼠是否可以通过调节 DOX 来诱导性地降低 GFP 表达水平，分析所述含有两个整合拷贝的慢病毒载体的 F1 小鼠 (F1-6# 和 F1-9#)。在给小鼠喂食 DOX 之前以及给小鼠喂食 DOX 10、17、27 天后，收集取自 4 周龄 F1-6# 和 F1-9# 转基因小鼠的血液。利用 FACS 分析法分析所述血细胞中的 GFP 表达水平。图 13 示出了在给所述小鼠喂食 DOX 之前所述血细胞中的 GFP 表达水平。两只转基因小鼠 (F1-6# 和 F1-9#) 体内的 GFP 表达水平均类似于非转基因小鼠体内的 GFP 表达水平。图 14 示出了在给所述小鼠喂食 DOX 10、17、27 天后所述转基因小鼠体内的 GFP 表达水平。在给所述小鼠喂食 DOX 之后，所述血细胞中的 GFP 表达水平下降。所述血细胞中 GFP 表达水平的降低并不一致，一些细胞中 GFP 表达水平降低最高达 30 倍，而一些细胞中 GFP 表达水平维持不变。喂食 DOX 17 天后，75% 的血细胞的 GFP 表达水平降低了 20 倍。喂食 DOX 27 天后，85% 的血细胞的 GFP 表达水平降低了 30 倍。这些数据表明所述诱导型慢病毒载体可表达其数量足以沉默 F1 小鼠体内 GFP 表达的 shRNA，表明经过种系传递之后所述诱导型单慢病毒载体为功能性的。

### 实施例 17

经过种系传递之后表达 shRNA 的诱导型单慢病毒载体为功能性的

为确定经过种系传递之后所述诱导型单慢病毒载体是否为功能性的，使两只阳性转基因小鼠 (GFP/CAG-建立者 6# 为雌性，GFP/CAG-建立者 9# 为雄性) 交配。当使用两只建造者来彼此交配时，我们希望可增加 shRNA 表达以明显地抑制 GFP 水平。利用 DNA 印迹分析法分析所有 F1 小鼠以测定所述慢病毒载体整合拷贝的数量。在 11 只 F1 小鼠中，有 2 只含有两个整合拷贝的慢病毒载体。其他小鼠含有一个整合拷贝的慢病毒载体 (5 只小鼠) 或者为整合阴性的 (4 只小鼠)。

为确定含有所述慢病毒载体的 F1 小鼠是否可以利用 DOX 降低 GFP，分析了含有两个整合拷贝的载体的 F1 小鼠 (F1-6# 和 F1-9#)。在给小鼠喂食 DOX 之前以及给小鼠喂食 DOX 10、17、27 天后，收集 4 周龄转基因

因小鼠的血液。利用 FACS 分析法分析所述血细胞中的 GFP 水平。图 13 示出了在喂食 DOX 前所述血细胞中的 GFP 水平。两只转基因小鼠 (F1-6# 和 F1-9#) 的 GFP 表达水平均类似于非转基因小鼠体内的 GFP 表达水平。图 14 示出了喂食 DOX 10、17、27 天后所述血细胞内的 GFP 水平。喂食 DOX 后, 所述血细胞内的 GFP 水平降低。所述血细胞中 GFP 水平的降低并不一致, 一些细胞中 GFP 表达降低最高达 30 倍, 而其他细胞中 GFP 表达水平无变化。喂食 DOX 17 天后, 75% 的血细胞的 GFP 表达水平降低了 20 倍, 而喂食 DOX 27 天后, 85% 的血细胞的 GFP 表达水平降低了 30 倍。这些数据表明所述诱导型慢病毒载体可表达其数量足以沉默 F1 小鼠体内 GFP 蛋白的 shRNA, 表明经过种系传递之后所述诱导型单慢病毒载体为功能性的。

### 实施例 18

诱导型单慢病毒载体可利用 II 型聚合酶表达基于微小 RNA 的 shRNA, 以便沉默基因表达

之前, 其他研究人员已经报道了使用由 H. Bujard 及其同事开发的使用四环素 (Tet)-调节系统的诱导型单慢病毒载体。在单一载体中的四环素诱导型启动子和人 CMV 启动子的调控下, 这种载体可表达 GFP 报道基因和四环素反式作用子。所述诱导型和组成型启动子排列反向相同, 均从 5'-LTR 至 3'-LTR。这类单一载体表达微小 RNA 或 shRNA, 所述微小 RNA 或 shRNA 可能会与非特异性 RNA 序列杂交。上述非特异性 RNA 序列可以降低微小 RNA 或 shRNA 的效率和功能。

为克服这类问题, 制备含有双顺反子 (bistronic) 诱导型启动子和组成型启动子的诱导型单慢病毒载体, 上述两种启动子取向相反。为减少所述诱导型启动子的启动子干扰和基础水平渗漏, 将 1.2 kb 的鸡隔离子插入到所述诱导型启动子和组成型启动子之间。选择 CAG 启动子来驱动所述四环素抑制子基因 (tetR-VP16 融合蛋白) 的表达和提高所述可诱导的体内基因表达。此外, 该载体使用改进的 tet-on 系统, 包括被称为 M2 的突变体形式 tet-on, 并且用四拷贝的最小 Vp16 反式作用子结构域置换所述单一全长 Vp16 结构域。还将 DsRed-exp 基因插入到所述四环素激活子基因的下游, 所述四环素激活子基因的表达由 CAG 启动子驱动并且由 IRES 表达。所述最终构建体被命名为

pHRpATRE/CAGM2Red。

当使用 Invitrogen miRNA 试剂盒时, 鉴别了一个 21 bp 的 miRNA 靶向序列(从 480 至 500 为 5'-CGGCATCAAGGTGAACTTCAA-3') (SEQ ID NO: 26) 可以有效地沉默 GFP 蛋白表达。利用 PCR 扩增 157 个碱基对的 miRNA-GFP(480) 并且将其克隆到 pHRpATRE/CAGM2Red 中。此外, 将 DsRed-exp 基因插入到所述四环素激活子基因的下游, 所述四环素激活子基因的表达由 CAG 启动子驱动并且由 IRES 表达。为促进所述诱导型启动子中转录的终止, 引入了两个 pA 信号元件(pA-BGH 和 pA-TK)。将所得的构建体命名为 pHR miRNA-GFP(480)/CAGM2Red (SEQ ID NO: 37)。使用源自构建体 pHRmiRNA-GFP(480)/CAGM2Red 的病毒颗粒来感染表达 GFP 的 HeLa 细胞。然后将被感染的表达 GFP 的 HeLa 细胞分为两组, 第一组在含有 0.5  $\mu$ g/ml DOX 的培养基中培养, 第二组在不含 DOX 的培养基中培养。感染 7 天后, 利用荧光显微术分析所述细胞。该数据表明基于 Pol II 的单慢病毒载体可以表达功能性 shRNA, 所述 shRNA 能够以可诱导和可逆转的方式降低其靶蛋白的表达。

### 实施例 19

基于 Cre-loxP 的有条件的、可诱导的、可逆转的慢病毒载体系统的开发

制备了基于 Cre-loxP 的有条件的、可诱导系统, 并且将其应用到转基因动物内。为制备该系统, 将 Cre-loxP 系统与一种四环素诱导的系统结合, 以便以组织特异性、可诱导、可逆转的方式表达基因。按照下列方式制备构建体: 将 850 bp 的 loxp-DsRed-loxp 插入到 pHRpATRE/CAGM2Red 中所述 M2 基因的上游, 并且将 M2 下游的 IRES-DsRed 片段删除。所得构建体为基于 Cre-loxp 的有条件的、可诱导、可逆转的慢病毒载体, 命名为 pHRpATRE/CAGloxRedM2。接下来, 将获自 pHR miRNA-GFP(480)/CAGM2Red 的 miRNA-GFP(480) 片段克隆到 pHRpATRE/CAGloxRedM2 内, 由此制备被命名为 pHRmiRNA-GFP(480)/CAGloxRedM2 的构建体。DsRed 荧光蛋白提供了一种监测 Cre-loxp 功能的方法。

使用含有 Bgl II 限制性内切酶和 SV40 NLS(用下划线标出)位点的正义引物(5'- GGA AGA TCT GAA TTC ACC ATG GAT CCC AAA AAG AAA AGA

AAG GTA GCA TCC AAT TTA CTA ACC GTA CAC -3') (SEQ ID NO: 39) 和含有 XhoI I 限制性内切酶位点的反义引物 (5'- ATG CCG CTC GAG CTA ATC GCC ATC TTC CAG CAG GCG -3') (SEQ ID NO: 40), 利用 PCR 扩增所述 Cre 基因。利用 Bgl II 和 XhoI I 限制性内切酶消化所述 PCR 产物, 并且使用 BamHI 和 XhoI 限制性内切酶将其克隆到 pHREFGFPblas 中, 命名为 pHREFla/CreNLS/blas。使用源自 pHREFla/CreNLS/blas 构建体的慢病毒载体颗粒感染表达 GFP 的 HeLa 细胞, 以组成性地表达所述 Cre 酶。利用杀稻瘟素筛选所述被感染的细胞。筛选所得的细胞在本文中命名为 GFP/Cre HeLa 细胞。使用所述源自 pHR miRNA-GFP(480)/CAGloxRedM2 的构建体病毒颗粒感染所述 GFP 或 GFP/Cre HeLa 细胞或者使所述病毒颗粒进入所述 GFP 或 GFP/Cre HeLa 细胞。感染三天后, 将所述细胞分为两组。使第一组暴露于 DOX (0.5  $\mu$ g/ml), 第二组不与 DOX 接触。感染 7 天后, 利用荧光显微术分析所述细胞。与在不存在 DOX 的情况下培养的细胞相比, 在所述暴露于 DOX 的 GFP/Cre HeLa 细胞内 GFP 表达水平显著下降。在所述 GFP/Cre HeLa 细胞内未检测到 DsRed 表达, 因此, 所述 Cre 酶可以移除所述 Loxp-DsRed-Loxp 片段, 并且可以利用 Cre 酶有条件地表达 M2。此外, 上述结果表明 M2 可以诱导表达功能性 shRNA 以便以四环素调控方式减低所述靶蛋白表达。

## 实施例 20

### 构建 pTREGag-HCV-Gag-Pol 包装构建体

如上文所述, 利用 PCR 由 pTRE 质粒(购自 Clontech)扩增所述四环素诱导型启动子片段。然后将所述 PCR 产物克隆到 pcDNA 3.1 中以置换 CMV 启动子, 从而制备所述四环素诱导型质粒, 所得的质粒在本文中被命名为 pTRE-neo。

接下来, 使用含有 EcoRI 限制性位点的正义引物 (5'-CGAATTCGAGCTCGGTACCCGGGATCGCGTGAAGCGCGCACGGCAAGAGGCGAG-3') (SEQ ID NO: 27) 与含有 MscI 限制性位点和 7 碱基突变的反义引物 (5'-CATGTTGGCCAAATTTTGGCCAGGAAATTAGCCTGTCTCTCAG-3') (SEQ ID NO: 28) 利用 PCR 扩增一个 1357 bp 的含有 MA(这是什么)、CA 和 NC 编码序列的 HIV-1 gag 片段。将所述 7 点突变引入到反义引物中是为了破坏所

述 PCR 产物中移码所必需的二级结构(环结构)。所述突变不会改变 gag 氨基酸序列。

接下来,使用含有 MscI 限制性位点的正义引物(5'-TTTGGCCAAGTCACAAGGGAAGGCCAG-3')(SEQ ID NO: 29)与含有 XhoI 和 MluI 限制性位点和 3 个点突变的反义引物(5'-CTCGACATGACGCGTTATTGTGACGAGGGGTCGCTGCCAAA-3')(SEQ ID NO: 30)利用 PCR 扩增一个 194 bp 的含有 P2 和 P6 编码序列的 HIV-1 gag 片段。将所述 3 个点突变引入到正义引物中是为了破坏所述 PCR 产物中移码所必需的二级结构(环结构)。所述突变不会改变 gag 氨基酸序列。使用 EcoRI 和 MscI 限制性内切酶消化所述 1357 bp 的 HIV-1 gag 片段。此外,使用 MscI 和 XhoI 限制性内切酶消化所述 194 bp 的 HIV-1 gag 片段。

还使用 EcoRI 和 XhoI 限制性内切酶消化所述 pTRE-neo 载体。然后将所述 PCR 产物的两个片段克隆到 pTRE-neo 内。将所得质粒命名为 pTRE-Gag 质粒。

接下来,使用含有 EcoRI、MluI 和 BssHII 限制性内切酶位点的正义引物(5'-GAATTCACGCGTATGGGCGCGCGTGCCTCAGTATTGAGCGGGGG-3')(SEQ ID NO: 31)与含有 BglII 限制性位点以及点突变和额外的碱基对插入的反义引物(5'-CGCAGATCTTCCCTGAAGAAGTTAGCCTGTCTCTCAGTACAATC-3')(SEQ ID NO: 32),利用 PCR 扩增一个 1313 bp 的含有 MA、CA 和 NC 编码序列的 HIV-1 gag 片段。将所述点突变和碱基对插入引入到正义引物中以破坏所述 PCR 产物的二级结构(环结构),并且生成所述 Gag-pol 融合蛋白。

然后,使用含有 BglII 限制性位点的正义引物(5'-AGATCTGGCATTTCGCGAGGGTAAAGCGCGTGAATTTTCCTCAGAGCAGACCAGAGCCAACA-3')(SEQ ID NO: 33)和含有 XhoI 和 SalI 限制性位点的反义引物(5'-GCCTCGAGCGATGTCGACACCCAATTCTGAAAAGAGTAAACAGCAG-3')(SEQ ID NO: 34)利用 PCR 扩增一个 3695 bp 的含有 P2、TF、蛋白酶、反转录酶、整合酶、vif 和 vpr 的 HIV-1 gag 和 Pol 片段。使用 EcoRI 和 BglII 限制性内切酶消化所述 1313 bp 的 PCR 产物,同时使用 BglII 和 XhoI 限制性内切酶消化所述 3695 bp 的 PCR 产物。

然后使用 EcoRI 和 XhoI 限制性内切酶消化 pTRE-neo。然后将两个 PCR 产物片段克隆到 pTRE-neo 内。将所得质粒命名为

pTRE-Gag-Pol/dTat/dRev。

利用 XhoI 和 SalI 限制性内切酶消化 pCMV-Gag-Pol, 获得一个 1710 bp 的含有 vpr、Tat、Rev 和 RRE 的片段。然后使用 XhoI 和 Sal I 限制性内切酶将该片段克隆到 pTRE-Gag-Pol 质粒中以制备被命名为 pTRE-Gag-Pol 的质粒。

使用含有 MluI 限制性内切酶位点的正义引物 (5'-CTGACGACGCGTGCCAGCCCCCTGATGGGGCGAC-3') (SEQ ID NO: 35) 和含有 BssH II 位点的反义引物 (5'-CGCACGCGCGCCCATGGTGCGCTGTGTACGAGACCTCCCGGGGCA-3') (SEQ ID NO: 36), 利用 PCR 扩增一个 340 bp 的 HCV IRES 片段。然后使用 MluI 和 BssH II 限制性内切酶消化所述 PCR 产物, 并且使用 MluI 和 BssH II 限制性内切酶将其克隆到 pTRE-Gag-Pol 质粒内。将所得质粒命名为 pTRE-HCV-Gag-Pol。

使用 MluI 和 XhoI 限制性内切酶消化 pTRE Gag, 获得了 1646 bp 的 Gag 片段, 并且使用 MluI 和 XhoI 限制性内切酶将其亚克隆到 pTRE-HCV-Gag-Pol 中。将最终的质粒命名为 pTREGag-HCV-Gag-Pol。所得质粒缺少所述保守性移码环结构。其次, Gag-Pol 融合蛋白的表达受到 HCV IRES 的调节。

### 实施例 21

#### 制备并分析 KiSS-1 转基因小鼠

肿瘤迁移抑制素 (Metastin) 是一种由癌细胞中 KiSS-1 基因编码的抗迁移肽。最近的研究发现肿瘤迁移抑制素是孤儿 G 蛋白偶联受体 GPR54 的配体, 所述 GPR54 在特定脑区域中 (例如下丘脑和部分海马体) 高水平表达。kisspeptins 在调节促性腺激素释放激素 (GnRH) 分泌过程中起到了重要作用。新的证据证实了 kisspeptin 通过作用于 GPR54 来刺激 GnRH 分泌。Kisspeptins 和 GPR54 对灵长类的青春期成熟至关重要。然而, 目前尚未有关于 KiSS-1 转基因小鼠的报道。下文所述实验描述了基于诱导型单慢病毒载体的转基因小鼠的制备, 所述转基因小鼠能够可诱导地和可逆转地表达人 KiSS-1 基因。

使用含有 BamHI 限制性内切酶位点的正义引物 (5'-ATCGCGGATCCCTGCCTCTTCTCACCAAGATGAACTCACTGGT-3') (SEQ ID NO:

41) 和含有 XhoI 限制性内切酶位点的反义引物 (5'-TTTCTCGAGTCACTGCCCCGCACCTGCGCC-3') (SEQ ID NO: 42), 利用 PCR 扩增人 KiSS-1 基因。使用 BamHI 和 XhoI 限制性内切酶消化所述 PCR 产物, 并且使用 BamHI 和 XhoI 限制性内切酶将其克隆到 pHRpAtet0CMVCAGtetGFP 中。将最终的构建体命名为 pHRKiSS02CAGtetGFP (SEQ ID NO: 43)。

如上所述, 使用源自 pHRKiSS02CAGtetGFP 的高效价慢病毒载体的感染性颗粒感染单细胞期胚胎。使用荧光显微术通过 GFP 阳性细胞测定感染性颗粒的效价。在感染后的第二天, 将所述双细胞期胚胎转移到代孕母鼠 (CD1) 体内。根据 GFP 表达情况筛选阳性转基因小鼠 (具有绿色身体的小鼠)。所测试的 11 只小鼠中有 7 只阳性转基因小鼠。使用含有 500  $\mu\text{g}/\text{ml}$  DOX 的水喂养 4 只 (2 只雌性和 2 只雄性) 4 周龄的建立者转基因小鼠以诱导 KiSS 基因表达。为测量所述 KiSS 转基因小鼠的表型, 监测雌性小鼠的阴道开口和雄性小鼠的阴茎。DOX 诱导 5 天后, 5 周龄的雌性小鼠的阴道张开。此外, 5 周龄雄性小鼠的阴茎颜色和大小均有变化。与对照小鼠的阴茎相比, KiSS Tg 雄性小鼠的阴茎较大并且发育更快。

### 实施例 22

使用慢病毒载体制备表达 M2 反式活化蛋白的 GNT1 细胞系

使用 BamHI 和 XhoI 将经修饰的 M2 基因 (含有与 VP16 有效连接的 tetON) 克隆到 pHREF-1ablas 载体中 (SEQ ID NO: 52)。将所得载体命名为 PS839pHREFM2blas (SEQ ID NO: 44)。通过使用 PS839pHREFM2blas (一种包装构建体 (p8.91, Trono lab, Lausanne, Switzerland)) 和 pCMV-VSV-G (pMD-G, Trono lab, Lausanne, Switzerland) 共转染来制备含有 PS839pHREFM2blas 的感染性颗粒。使用所述感染性颗粒感染 HEK293S 细胞 (GnT1<sup>+</sup>) 和 GnT1<sup>-</sup>HEK293S 细胞 (所述 HEK293S 细胞 (GnT1<sup>+</sup>) 和 GnT1<sup>-</sup>HEK293S 细胞由美国麻省理工学院 (Massachusetts Institute of Technology) 提供)。将所述被转导的细胞用杀稻瘟素 (20  $\mu\text{g}/\text{ml}$ ) 持续筛选一周。将抗性细胞系命名为 GnT1<sup>+</sup>HEK293S W2 细胞和 GnT1<sup>-</sup>HEK293S W2 细胞。上述细胞系含有上文所述的 M2 构建体。

制备表达 CCR1 的四环素诱导型细胞系

迄今为止, 已经描述了至少 10 种 CC 趋化因子受体家族成员。根据

IUIS/WHO 小组委员会关于趋化因子命名的规定,将所述成员命名为 CCR1 至 CCR10。CCR1 是首个被鉴别出来的 CC 趋化因子受体,并且可结合多种炎症/诱导型 CC 趋化因子(例如 CCL4-6 和 CCL14-16)。在人体内,该受体存在于外周血淋巴细胞和单核细胞上。该受体也被称为分化抗原簇标记 CD191。

#### 构建含有人 CCR1 的慢病毒载体

人 CCR1 cDNA (GENBANK 编号:BC051306) 获自 Open Biosystems (Huntsville, AL), 并且利用 PCR 对其进行扩增, 并使用 Invitrogen TA 克隆试剂盒 (Carlsbad, CA) 将其克隆到 pCR-2.1 载体中。将所得的载体命名为 pCR-hCCR1。对 CCR1 基因的终止密码子进行突变以便与 Tag 基因融合 (TEV-Flag-10His) (参见图 16B)。使用限制性内切酶消化 hEP2R 基因并且将其克隆到 pHTRE-puro 内 (也被称为 L494pHRTREpuro; SEQ ID NO: 45)。将所得载体命名为 pHTRE-hCCR1-TEV-Flag-10His (也被称为 PT834pHRTRE-hCCR1TEVpur; SEQ ID NO: 46)。图 16A 示出了载体的示意图。如图 16 所示, 驱动 hCCR1 表达的启动子为 tet-调节元件 (TRE), 其后紧邻 hCCR1 编码区。整合的载体转录双顺反子 mRNA, 使所述嘌呤霉素抗性基因与 hCCR1 为顺式。

#### 构建表达 hCCR1 的四环素诱导型细胞系

为产生源自 pHTRE-hCCR1-TEV-Flag-10His 的感染性颗粒, 将所述 pHTRE-hCCR1-TEV-Flag-10His 质粒与 p8.91 包装构建体 (Trono lab, Lausanne, Switzerland) 和 pCMV-VSV-G (pMD-G; Trono lab, Lausanne, Switzerland) 共转染到 293T 细胞内。使用所述含有 pHTRE-hCCR1-TEV-Flag-10His 的病毒颗粒感染 G<sub>n</sub>T1<sup>-</sup> HEK293S W2 细胞, 所述细胞具有降低的 G<sub>n</sub>T1 活性并且还表达四环素反式作用子。发现被感染的细胞可产生高水平的 CFTR 蛋白 (3-5mg/10<sup>9</sup> 细胞), 这比其他已知被用于表达膜蛋白的系统高 1 个对数值。利用嘌呤霉素 (从 1.0 至 4.0 μg/ml) 筛选被转导的细胞, 以建立被命名为 hCCR1-m 细胞的稳定细胞系。

#### 分析四环素诱导型细胞系内的 hCCR1 表达

使 hCCR1-m 细胞 (使用 0、1、2 和 4 μg 嘌呤霉素/ml 进行筛选) 在 6 孔板中生长并且将 1 μg/ml DOX 加入到培养基中。第二天, 收集诱导的 hCCR1-m 细胞并且利用蛋白质印迹对其进行分析, 所述蛋白质印迹方法

使用了一抗(M2 flag 抗体)和二抗(HRP 缀合的抗-小鼠抗体)。还对所述印迹进行抗微管蛋白共染色作为对照。利用 M2 flag 抗体在 hCCR1-m 细胞内检测到一条 52 kDa 的条带,而未在 HEK 细胞内检测到该条带,同时所有细胞中均检测到一条 55 kDa 的条带。

#### CCR1 的高水平表面表达

为确定所述抗 CCR1 特异性抗体是否可以在 hCCR1-m 细胞中检测到 CCR1 的细胞表面表达,使用与 Alexa Fluor® 647 缀合的小鼠抗人 CCR1 单克隆抗体(CAT#557914, BD Bioscience, San Jose, CA)对使用或未使用 DOX 诱导的 hCCR1-m 细胞进行染色(24 小时)。使用未转导的 293 细胞作为对照。结果表明 DOX 诱导的细胞系所表达的 CCR1 的细胞表面水平非常高。与未诱导的 hCCR1-m 细胞相比,在诱导的 hCCR1-m 细胞内诱导水平为大约 10 倍。

#### 诱导 CCR1 以阻止细胞生长和/或导致细胞凋亡

将所述 hCCR1-m 细胞在含有或不含有 1  $\mu$ g/ml DOX 的 6 孔板中培养。利用显微镜来监测细胞生长。在使用 DOX 诱导 24 小时之后,所述 hCCR1-m 细胞停止生长,同时大多数被诱导的 hCCR1-m 细胞会从板上脱附。数据表明高水平的 hCCR1 表达会导致 hCCR1 信号的激活,而不需要配体。

### 实施例 23

#### 制备表达 CFTR (Cl<sup>-</sup>阴离子穿膜通道)的四环素诱导型细胞系

#### 构建 CFTR His (10x)慢病毒载体

使用 BstXI 和 XhoI 消化 CFTR His (6x) 慢病毒载体(又称为 PT764pHRTRECFTR-His6puro; SEQ ID NO: 47)DNA 以移除含有 6xHis 的 C 末端 DNA 片段。使用野生型 CFTR 质粒 DNA 作为模板,利用 PCR 扩增含有 10xHis 的 C 末端 CFTR BstXI/XhoI DNA 片段。在使用 BstXI 和 XhoI 消化之后,克隆所述片段并且使用内切酶限制性分析和核苷酸序列分析进行确认。将所得载体命名为 CFTR His (10x)(又称为 PT823pHRTRECFTR-His10pur; SEQ ID NO: 48)。

使用 CFTR His (6x)和 CFTR His (10x)慢病毒载体转导 GnT1<sup>+</sup>和 GnT1<sup>-</sup> HEK293S 细胞。

如本文其他部分所述,包装每个慢病毒载体并且用其转导 GnT1<sup>+</sup> HEK293S W2 细胞和 GnT1<sup>-</sup> HEK293S W2 细胞。在给所述培养基添加 25  $\mu$

g/ml 嘌呤霉素两天后，筛选高度表达 CFTR 的细胞系。筛选四天后，在不含嘌呤霉素的培养基中放大培养存活的细胞。

### 免疫印迹分析

收集三百万个每种类型的细胞(转导的  $GnT1^+$  HEK293S W2 细胞和  $GnT1^-$  HEK293S W2 细胞)，并且制备膜组分以进行免疫印迹分析。此外还分析了 CFTR<sup>+</sup>对照 (CFTR-FLAG)。使用 R1104 抗-CFTR MAb 检测被印迹的蛋白。结果表明转导的  $GnT1^-$  HEK293S W2 细胞表达了 6x 和 10x 标记的 CFTR 蛋白。观察到与转导的  $GnT1^+$  HEK293S W2 细胞中所表达的同一种蛋白的条带相比，转导的  $GnT1^-$  HEK293S W2 细胞中所表达的蛋白的条带在聚丙烯酰胺凝胶中迁移的更快。

分析表达 CFTR-His (10x) 的  $GnT1^-$  HEK293S 细胞中的离子通道功能

为确定所表达的 CFTR-His-tag (10x) 蛋白是否具有活性，使用卤化物淬灭染料 6-甲氧基-N-3-(磺丙基)喹啉铵 (SPQ, Molecular Probes) 测量  $GnT1^+$  和  $GnT1^-$  细胞系中的卤化物外流。为进行比较，平行分析了表达野生型 CFTR 的  $GnT1^+$  细胞。简言之，将转导的 HEK293S wt-CFTR、HEK293S CFTR-His (10x)、HEK293S  $GnT1^-$  CFTR-His (10x) 细胞系在盖玻片上接种并且培养至约 50% 汇合度。然后将所述细胞用 10 mM SPQ 低渗负载 10 分钟，并且置于淬灭 NaI 缓冲液中。使用 Zeiss 倒置显微镜、PTI 成像系统和 Hamamatsu 照相机测量单细胞荧光。以 340 nm 的波长激发，测量 410 nm 波长的发射光。在实验开始时将细胞置于淬灭缓冲液中 (NaI)，在建立了稳定的基线之后，在 200 秒时将细胞换到不含卤化物的去淬灭缓冲液中。在 620 秒时使用激动剂 (20  $\mu$ M 福斯高林) 刺激细胞，然后将细胞放回到所述淬灭 NaI 缓冲液中。用其基线值将每个细胞的荧光标准化，并且将荧光变化表示为与基线荧光相比增加的百分数。将每个时间点所分析的细胞总数 (至少 30 个) 的平均值作图。所得结果证实每种细胞系的卤化物外流均被显著活化。与 HEK293S wt-CFTR 相比，所述 HEK293S CFTR-His (10x) 和  $GnT1^-$  HEK293 CFTR-His (10x) 细胞系的荧光均产生更大变化。

### 实施例 24

制备表达人 EP2R 的四环素诱导型细胞系

构建含有 EP2 的慢病毒载体

hEP2 cDNA 获自 Schering Ag (Berlin, Germany), 并且利用 PCR 对其进行扩增, 并使用 Invitrogen TA 克隆试剂盒 (Carlsbad, CA) 将其克隆到 pCR-2.1 载体中。将所得构建体命名为 pCR-hEP2R (SEQ ID NO: 49)。对 EP2 基因的终止密码子进行突变以便与 Tag 基因融合 (TEV-Flag-10His) (图 17)。使用限制性内切酶消化 hEP2R 基因, 并且将其克隆到 pHTRE-puro (SEQ ID NO: 45) 内。将所得构建体命名为 pHTRE-hEP2R-TEV-Flag-10His (SEQ ID NO: 50)。图 17 示出了载体的示意图。如图 17 所示, 驱动 hEP2R 表达的启动子为 tet-调节元件 (TRE), 其后紧邻 hEP2R 编码区。所述整合的载体转录一种双顺反子 mRNA, 使所述嘌呤霉素抗性基因与 hEP2R 同向。

#### 构建表达 hEP2R 的四环素诱导型细胞系

为产生源自 pHTRE-hEP2R-TEV-Flag-10His (SEQ ID NO: 50) 的感染性颗粒, 将所述 pHTRE-hEP2R-TEV-Flag-10His 质粒 (SEQ ID NO: 50) 与 p8.91 包装构建体 (Trono lab, Lausanne, Switzerland) 和 pCMV-VSV-G (pMD-G; Trono lab, Lausanne, Switzerland) 共转染到 293T 细胞内。使用所述含有 pHTRE-hEP2R-TEV-Flag-10His 的病毒颗粒感染经过基因工程改造的细胞系, 所述细胞系具有降低的 GnTI 活性 (GnTI<sup>-</sup> HEK293S W2 细胞) 并且还表达四环素反式作用子。发现被感染的细胞可产生高水平的 hEP2R 蛋白 (3-5mg/10<sup>9</sup> 细胞), 这比其他已知被用于表达膜蛋白的系统高 1 个对数值。利用嘌呤霉素 (从 1.0 至 4.0 μg/ml) 筛选被转导的细胞以建立被命名为 hEP2R-m 细胞的稳定细胞系。

#### 分析四环素诱导型细胞系内的 hEP2R 表达

使 hEP2R-m 细胞 (使用 0 μg 或 2 μg /ml 嘌呤霉素进行筛选) 在 6 孔板中生长并且将 1 μg/ml DOX 加入到培养基中。第二天, 收集诱导的 hEP2R-m 细胞并且利用蛋白质印迹对其进行分析, 所述蛋白质印迹方法使用了一抗 (M2 flag 抗体) 和二抗 (HRP 缀合的抗-小鼠抗体) 的。对所述印迹利用抗-微管蛋白共染色作为对照。利用 M2 flag 抗体在 hEP2R-m 细胞内检测到三条带, 而在 HEK293S 细胞内未检测到所述条带。三条带的大小为 45-53 kDa, 所述条带均小于微管蛋白的大小 (55 kDa)。hEP2R 的期望大小为 53 kDa。

#### 诱导 hEP2R 以阻止细胞生长和/或导致细胞凋亡

将所述 hEP2R-m 细胞在含有或不含有 1 μg/ml DOX 的 6 孔板中培养。

---

利用显微镜来监测细胞生长。在诱导 24 小时之后，所述 hEP2R-m 细胞停止生长，同时大多数被诱导的 hEP2R-m 细胞会从板上脱附。数据表明高水平的 hEP2R 表达会导致 hEP2R 信号的活化，而不需要配体。

<110> UAB研究基金会

<120> 与使用病毒载体控制的基因表达相关的方法和组合物

<130> 21085.0140P1

<150> 60/751,407

<151> 2005-12-16

<150> 60/751,117

<151> 2005-12-16

<160> 52

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 654

<212> DNA

<213> 人工序列

<220>

<223> 对人工序列的描述: 合成构建体

<400> 1

```

tgcttaatg aggtcggaaat cgaaggttta acaaccogta aactogccca gaagctaggt      60
gtagagcagc ctacattgta ttggcatgta aaaaataage gggctttgct cgacgcotta      120
gccattgaga tgttagatag gcaccatact cacttttgcc ctttagaagg ggaaagctgg      180
caagattttt tacgtaataa cgctaaaagt tttagatgtg ctttactaag tcatcgcgat      240
ggagcaaaag tacatttagg tacacggcct acagaaaaac agtatgaaac tctcgaaaat      300
caattagcct ttttatgcoa acaaggtttt tcactagaga atgcattgta cgccctgtcc      360
gccgtcggcc acttcaccct gggctgtgtg ctggaggacc aagagcatca agtcgctaaa      420
gaagaaaggg aaacacctac tactgatagt atgcgcgcat tattacgaca agctatcgaa      480
ttatttgatc accaaggtgc agagccagcc ttcttattcg gccttgaatt gatcatatgc      540
ggattagaaa aacaacttaa atgtgaaagt gggtcgcgct acagccgctg cggaggcgga      600
ggcagtcgcg gcgcgcatcc caaaaagaaa agaaaggtag cagccatggc ctaa      654

```

<210> 2

<211> 891

<212> DNA

<213> 人工序列

<220>

<223> 对人工序列的描述: 合成构建体

<400> 2

```

atggccagc cgccctggaca agtccaaggt catcaattcc gcattagagc tgcttaatga      60
ggtcggaaatc gaaggtttaa caaccogtaa actcgcgccag aagctaggtg tagagcagcc      120
tacattgtat tggcatgtaa aaaataagcg ggcctttgctc gacgccttag ccattgagat      180

```

tttatgccaa	caagggttttt	cactagagaa	tgcattgtac	gccctgtccg	ccgtcggcca	420
cttcaccctg	ggctgtgtgc	tggaggacca	agagcatcaa	gtcgctaaag	aagaaagggg	480
aacacctact	actgatagta	tgccgccatt	attacgacaa	gctatcgaat	tatttgatca	540
ccaaggtgca	gagccagcct	tcttattcgg	ccttgaattg	atcatatgcg	gattagaaaa	600
acaacttaaa	tgtgaaagtg	ggtccgcgta	cagccgcggc	ggagggcgag	gcagtccgcg	660
cgccgatccc	aaaaagaaaa	gaaaggtagc	acgcgtcggc	ggagggcgaa	gtgggtcccc	720
ggccgacgcc	ctggacgact	tgcacctgga	catgctgccc	gccgacgccc	tggacgactt	780
cgacctggac	atgctgcccg	ccgacgccct	ggacgacttc	gacctggaca	tgctgcccgg	840
cgacgccctg	gacgacttcg	acctggacat	gtgcccgggg	taactaagta	a	891

&lt;210&gt; 3

&lt;211&gt; 891

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;223&gt; 对人工序列的描述: 合成构建体

&lt;400&gt; 3

atggccagc	cgccctggaca	agtccaaggt	catcaatggc	gccctggagc	tgctgaacgg	60
cgctgggaatc	gaagggtttaa	caacccgtaa	actcgcccag	aagctaggtg	tagagcagcc	120
tacattgtat	tggcatgtaa	aaaataagcg	ggctttgctc	gacgccttac	ccatcgagat	180
gctggaccgc	caccacaccc	acttetgccc	cctggagggc	gagagctggc	aggacttctt	240
acgtaataac	gctaaaagtt	ttagatgtgc	tttactaagt	catcgcgatg	gagcaaaagt	300
acatttaggt	acacggccta	cagaaaaaca	gtatgaaact	ctcgaaaatc	aattagcctt	360
tttatgccaa	caagggttttt	cactagagaa	tgcattgtac	gccctgtccg	ccgtcggcca	420
cttcaccctg	ggctgtgtgc	tggaggagca	ggagcatcaa	gtcgctaaag	aagaaagggg	480
aacacctact	actgatagta	tgccgccatt	attacgacaa	gctatcgaat	tatttgatcg	540
ccaagggcgc	gagccgcctt	tcctgttcgg	cctggagctg	atcatctgcg	gcctggagaa	600
gcagctgaag	tgcgagagcg	gcagcgccta	cagccgcggc	ggagggcgag	gcagtcocog	660
cgccgatccc	aaaaagaaaa	gaaaggtagc	acgcgtcggc	ggagggcgaa	gtgggtcccc	720
ggccgacgcc	ctggacgact	tgcacctgga	catgctgccc	gccgacgccc	tggacgactt	780
cgacctggac	atgctgcccg	ccgacgccct	ggacgacttc	gacctggaca	tgctgcccgg	840
cgacgccctg	gacgacttcg	acctggacat	gtgcccgggg	taactaagta	a	891

&lt;210&gt; 4

&lt;211&gt; 901

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;223&gt; 对人工序列的描述: 合成构建体

&lt;400&gt; 4

atggcctcc	agattagata	aaagtaaagt	gattaacagc	gcattagagc	tgcttaatga	60
ggctgggaatc	gaagggtttaa	caacccgtaa	actcgcccag	aagctaggtg	tagagcagcc	120
tacattgtat	tggcagctgc	gcaacaagca	gactcctatg	aacatgcttt	cagaggcaat	180
actggcgaag	catcacaccc	gttcagcacc	gttaccgact	gagagttggc	agcagtttct	240
ccaggaaaaat	gctctgagtt	tccgtaaagc	attactggtc	catcgtgatg	gagcccattt	300
gcatataggg	acctctceta	gccccccca	gtttgaacaa	gcagagggcg	aactacgctg	360
tctatgcatg	gcagggtttt	cggtcgagga	ggctcctttc	attctgcaat	ctataagcca	420
ttttagcttg	gggtgcagtat	tagaggagca	agcaacaac	cagatagaaa	ataatcatgt	480
gatagacgct	gcaccacat	tattacaaga	ggcatttaat	attcaggcga	gaacctctgc	540
tgaaatggcc	ttccatttctg	ggctgaaatc	attaataat	ggattttctg	cacagttaga	600
tgaaaaaaag	catacaccca	ttgaggatgg	taataaaggc	ggagggcgag	ggcgcgccga	660
tcccaaaaag	aaaagaaagg	tagcacgcgc	cgggggaggg	ggcctggcag	tgtcagtgac	720
atgtgaagat	gtggctgtgc	tctttactcg	ggacgagtgg	aagaagctgg	atctgtctca	780
gagaagcctg	taccgtgagg	tgatgctgga	gaattacagc	aacctggcct	ccatggcagg	840
atctctgttt	accaaacc	aggtgatctc	cctgttgcag	caaggagagg	acctctggt	900

**a** **901**

<210> 5  
 <211> 1000  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 对人工序列的描述: 合成构建体

<400> 5  
 atggcctcc agattagata aaagtaaagt gattaacagc gcattagagc tgcttaatga 60  
 ggtcggaatc gaaggtttaa caaccgtaa actegcccag aagctagggtg tagagcagcc 120  
 tacattgtat tggcacgtgc gcaacaagca gactcttatg aacatgcttt cagaggcaat 180  
 actggcgaag catcacaccc gttcagcacc gttaccgact gagagttggc agcagtttct 240  
 ccaggaaaaat gctctgagtt tccgtaaagc attactggtc catcgtgatg gagcccgatt 300  
 gcataataggg acctotccta gccccccca gtttgaacaa gcagagggcg aactacgctg 360  
 tctatgcatg gcagggtttt cggtcgagga ggctcttttc attctgcaat ctataagcca 420  
 ttttagcttg ggtgcagtat tagaggagca agcaacaac cagatagaaa ataatcatgt 480  
 gatagacgct gcaccaccat tattacaaga ggcatttaat attcaggcga gaacctctgc 540  
 tgaaatggcc ttccatttog ggtgaaatc attaatatat ggattttctg cacagttaga 600  
 tgaaaaaaaaag catacaccca ttgaggatgg taataaaggc ggaggcggag ggcgcgccga 660  
 tccccaaaaag aaaagaaagg tagcacgcgc cgggggaggc ggctgatgg atgctaagtc 720  
 actaactgcc tggcccgga cactggtgac cttcaaggat gtatttgtgg acttaccag 780  
 ggaggagtgg aagctgctgg aactgctca gcagatcgtg tacagaaatg tgatgctgga 840  
 gaactataag aacctgggtt ccttgggtta tcagcttact aagccagatg tgatcctccg 900  
 gttggagaag ggagaagagc cctggctggt ggagagagaa attcaccaag agaccatcc 960  
 tgattcagag actgcatttg aatcaaatc atcagtttaa 1000

<210> 6  
 <211> 107  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 对人工序列的描述: 合成构建体

<400> 6  
 actagtcat gcaaattacg cgctgtgctt tgtgggaaat caccctaac gtaaaatccc 60  
 tatcagtgat agagacttat aatccctatc agtgatagag aggatcc 107

<210> 7  
 <211> 8  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 对人工序列的描述: 合成构建体

<400> .  
 uuuuuua 8

<210> 8  
 <211> 7  
 <212> DNA  
 <213> 人工序列

<220>

<223> 对人工序列的描述: 合成构建体

<400> 8  
uuuuua

7

<210> 9  
<211> 7  
<212> DNA  
<213> 人工序列

<220>  
<223> 对人工序列的描述: 合成构建体

<400> 9  
uuuuuu

7

<210> 10  
<211> 104  
<212> DNA  
<213> 人工序列

<220>  
<223> 对人工序列的描述: 合成构建体

<400> 10  
ggaaaggaa ggacaccaa tgaaagattg tactgagaga caggctaatt tcttgggcaa 60  
aatttggcca agtcacaagg gaaggccagg gaattttctt caga 104

<210> 11  
<211> 105  
<212> DNA  
<213> 人工序列

<220>  
<223> 对人工序列的描述: 合成构建体

<400> 11  
aggctaact tcttcaggga agatctggca tttcgcagg gtaaagcgcg tgaattttcc 60  
tcagagcaga ccagagccaa cagccccacc agaagagagc ttcag 105

<210> 12  
<211> 1878  
<212> DNA  
<213> 人工序列

<220>  
<223> 对人工序列的描述: 合成构建体

<400> 12  
atctctatc actgatagg agatctctat aactgatagg gagagctctg cttatataga 60  
cctcccaccg tacacgccta ccgccattt gcgtcaatgg ggcggagtg ttacgacatt 120  
ttggaagt cegtgattt tggttccaaa acaaactccc attgacgtca atgggggtgga 180  
gacttggaat tccccgtgag tcaaacgcgt atccacgccc attgatgtac tgccaaaacc 240  
gcatcaccat ggtaatagcg atgactaata caattctaaa tggcccgcct ggctgaccgc 300  
ccaacgacc ccgccattg acgtcaataa tgacgtatgt tcccatagta acgccaatag 360  
ggactttcca ttgacgtcaa tgggtggagt atttacggtg aactgcccac ttggcagtac 420

atcaagtgta	toatatgcca	agtacgeccc	ctattgacgt	caatgacggg	aaatggcccc	480
cctggcatta	tgcccagtac	atgaccttat	gggactttcc	tacttggcag	tacatctacg	540
tattagtcac	cgctattaac	atgggtcgagg	tgagccccac	gttctgcttc	actctcccca	600
tctccccccc	ctccccacco	ccaattttgt	atattattat	tttttaatta	ttttgtgcag	660
cgatgggggg	gggggggggg	ggggggcgcg	cgccaggcgg	ggcggggcgg	ggcgaagggg	720
ggggcggggg	gaggcgagga	ggtgcggcgg	cagccaatca	gagcggcgcg	ctccgaaaag	780
ttccttttat	ggcgaggcgg	cgccggcgcg	ggcctataaa	aaagcgaagc	gcgcggcggg	840
cggggagtcg	ctgcgacgct	gccttcgccc	cgtgccccgc	tccgcgcgcg	cctcgcgcgc	900
cccgccccgg	ctctgactga	cogcgttact	cccacagggt	agcggggcgg	acggcccttc	960
tectccgggc	tgtaattagc	gcttggttta	atgacggctt	gtttcttttc	tgtggctgcg	1020
tgaaagcctt	gaggggctcc	gggaggggccc	tttgtgcggg	gggagcggct	cggggggtgc	1080
gtgcgtgtgt	gtgtgcgtgg	ggagcgcccg	gtgcgcctcc	gcctgcgccg	gcggctgtga	1140
cgctgcgggg	cgccgcggcg	ggctttgtgc	gctccgcagt	gtgcgcgagg	ggagcggcgc	1200
cgggggcggt	gccccgcggg	gcgggggggg	ctgcgagggg	aacaaaggct	gcgtgcgggg	1260
tgtgtgcgtg	ggggggtgag	cagggggtgt	gggcgcgtcg	gtcgggctgc	aacccccctt	1320
gcacccccct	ccccgagttg	ctgagcacgg	cccggcttcg	gggtgcgggg	tccgtacggg	1380
gcgtggcgcg	gggctcgccg	tgccggggcg	ggggtggcgg	caggtggggg	tgccgggcgg	1440
ggcgggggcc	octcgggocg	gggaggggctc	gggggagggg	cgcgggcgcc	cccggagcgc	1500
ggcgggcgct	cgaggcgcgg	cgagccgcag	ccattgcctt	ttatggtaat	cgtgcgagag	1560
ggcgagggga	cttcttttgt	cccaaatctg	tgccgagccg	aaatctggga	ggcgccgcgc	1620
cacccccctt	agcgggcggg	gggcgaagcg	gtgcggcgcc	ggcaggaagg	aaatgggcgg	1680
ggagggcctt	cgtgcgtcgc	cgcgcccgcc	tccccttctc	cctctccagc	ctcggggctg	1740
tccgcggggg	gacggctgcc	ttcggggggg	acggggcagg	gcgggggttc	gcttctggcg	1800
tgtgaccggc	ggctctagac	aattgtacta	accttctctc	cttctctctc	ctgacaggtt	1860
ggtgtacagt	agcttcca					1878

&lt;210&gt; 13

&lt;211&gt; 1732

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;223&gt; 对人工序列的描述: 合成构建体

&lt;400&gt; 13

ggatccgat	ctctatcact	gatagggaga	tctctatcac	tgatagggag	agctctgctt	60
atatagacct	cccaccgtac	acgcctaccg	cccattttgg	tcaatggggc	ggagttgtta	120
cgacattttg	gaaagtcccg	ttgattttgg	ttccaaaaca	aactcccatt	gacgtcaatg	180
gggtggagac	ttggaaatcc	ccgtgagtca	aaccgcctatc	caagccattt	gatgtactgc	240
caaaaccyca	tcaccatggt	aatagcgatg	actaatacgt	agatgtactg	ccaagtagga	300
aagtcccata	aggtcatgta	ctgggcataa	tgccaggcgg	gccatttacc	gtcattgacg	360
tcaatagggg	catatgatac	acttgatgta	ctgccaaagtg	ggcagtttac	ggcagtttac	420
cgtaaatact	ccacccattg	acgtcaatgg	aaagtcccta	ttggcgttac	tatgggaaca	480
tacgtcatta	ttgacgtcaa	tgggcggggg	tcgttggggc	gtcagccagg	cgggccattt	540
agaattcaag	cttcgtgagg	ctccgggtgcc	cgtcagtggg	cagagcgcac	atcgccaca	600
gtocccgaga	agttgggggg	aggggtcggc	aattgaaccg	gtgcctagag	aagggtggcg	660
ggggtaaaact	gggaaagtga	tgtcgtgtac	tggctccgcc	tttttccoga	gggtggggga	720
gaaccgtata	taagtgcagt	agtcgcccgtg	aacgttcttt	ttcgcaacgg	gtttgcggcc	780
agaacacagg	taagtgcctg	gtgtggttcc	cgggggcttg	gcctctttac	gggttatggc	840
octtgcgtgc	cttgaattac	ttccacctgg	ctccagtacg	tgattcttga	tcccagctg	900
gagccagggg	cgggccttgc	gctttaggag	ccccttcgcc	togtgcttga	gttgaggcct	960
ggcctggggc	ctggggcgcg	cgcgtgcgaa	tctgggtggca	cettcgcgoc	tgtctcgetg	1020
ctttcgataa	gtctctagcc	atttaaattt	tttgatgac	tgctgcgacg	ctttttttct	1080
ggcaagatag	tcttgtaaat	gcgggcccagg	atctgcacac	tggatatttc	gtttttgggc	1140
ccgcggccgg	cgacggggcc	cgtgcgtccc	agcgcacatg	ttcggcgagg	cggggcctgc	1200
gagcgcggcc	accgagaatc	ggacgggggt	agtctcaage	tggccggcct	gctctgggtg	1260
ctggcctcgc	gcccgcctgt	atcgcgccgc	cctgggcggc	aaggctggcc	cggctcggcac	1320
cagttgcgtg	agcggaaaga	tggccgcttc	ccggccctgc	tccagggggc	tcaaaatgga	1380
ggacgcggcg	ctcgggagag	cgggggggtg	agtcacccac	acaaaggaaa	agggcctttc	1440
cgtcctcagc	cgtcgcctca	tgtgactcca	cggagtaacc	ggcgcgctcc	aggcacctcg	1500

```

attagttctg gagcttttgg agtacgtcgt ctttaggttg gggggagggg ttttatgcga 1560
tggagtttcc ccacactgag tgggtggaga ctgaagttag gdcagcttgg cacttgatgt 1620
aattctcctt ggaatttggc ctttttgagt ttggatcttg gttcattctc aagcctcaga 1680
cagtggttca aagttttttt cttccatttc aggtgtcgtg aggatctact ag 1732

```

<210> 14  
 <211> 1715  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 对人工序列的描述: 合成构建体

```

<400> 14
ggatcctct ctaactga tagggattat aagtctctat cactgatagg gattttacgt 60
ttagggatg ttcccacaaa gcacagcgcg taatttgcac gactagtcaa ttctaaatgg 120
cccgcctggc tgaccgcca acgacccccg ccattgacg tcaataatga cgtatgttcc 180
catagtaacg ccaatagggc ctttecaty acgtcaatgg gtggagtatt tacggtaaac 240
tgcccacttg gcagtacatc aagtgtatca tatgccagt acgcccccta ttgacgtcaa 300
tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg actttcctac 360
ttggcagtac atctacgtat tagtcatcgc tattaacatg gtccaggtga gccccacgtt 420
ctgcttcaat ctccccatct cccccccctc cccacccccca attttgtatt tatttatttt 480
ttaattattt tgtgcagcga tgggggcggg gggggggggg ggggcgcgcgc caggcggggc 540
ggggcggggc ggggggcggg ggggggcgag ggggagaggt ggggcggcag ccaatcagag 600
cggcgcgcgc cgaaagtctc cttttatggc gaggcggcgg cggcggcggc cctataaaaa 660
gcgaagcgcg cggcggcggg ggagtcgctg cgacgctgcc ttccgcccgt gccccgctcc 720
gccgcgcgct cgcgcgcgcc gccccggctc tgactgaccg cgttactccc acaggtgagc 780
gggcgggacg gcccttctcc tccgggctgt aattagcgtc tggtttaatg acggcttgtt 840
tctttctgt ggcctgcgtg aagccttgag gggctccggg agggcccttt gtgcgggggg 900
agcggctcgg ggggtgcgtg cgtgtgtgtg tgcgtgggga ggcgcgcgtg cggctccgcg 960
ctgcccggcg gctgtgagcg ctgcccgcgc ggcgcggggc tttgtgcgct ccgcagtgtg 1020
cgcgagggga gcgcggccgg gggcgggtgcc ccgcggtgcg gggggggctg cgaggggaac 1080
aaaggtgcg tgcggggtgt gtcgctgggg ggggtgagcag ggggtgtggg cgcgtcggtc 1140
gggctgcaac cccccctgca cccccctccc cgagtgtctg agcacggccc ggettegggt 1200
gcggggctcc gtacggggcg tggcgcgggg ctccgctgct cgggcggggg gtggcggcag 1260
gtgggggtgc cgggcggggc ggggcgcgct cgggcggggg agggctcggg ggaggggcgc 1320
ggcggccccc ggagcgcggg cggctgtcga ggcgcggoga gccgcagcca ttgcctttta 1380
tggtaatcgt gcgagagggc gcagggactt cctttgtccc aaactctgtc ggagccgaaa 1440
tctgggaggg gcgcgcgcac cccctctagc gggcgcgggg cgaagcgggt cggcgcgggc 1500
aggaaggaaa tgggcgggga gggccttctg gcgtgcgcgc gccgcgctcc ccttctccct 1560
ctccagcctc ggggtgtctc ggggggggac ggcctgcttc gggggggacg gggcagggcg 1620
gggttcggct tctggcgtgt gaccggcggc totagacaat tgtactaacc ttcttctctt 1680
tcctctcctg acaggttggg gtacagtagc ttcca 1715

```

<210> 15  
 <211> 6  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 对人工序列的描述: 合成构建体

```

<400>
ggggg 6

<210> 16
<211> 21
<212> DNA
<213> 人工序列

```

<220>  
 <223> 对人工序列的描述: 合成构建体

<400> 16  
 tcgaagggt taacaaccog t 21

<210> 17  
 <211> 21  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 对人工序列的描述: 合成构建体

<400> 17  
 ttgtcgtaa taatggggc a 21

<210> 18  
 <211> 34  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 对人工序列的描述: 合成构建体

<400> 18  
 gcggcgca attcatattt gcatgctgct atgt 34

<210> 19  
 <211> 78  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 对人工序列的描述: 合成构建体

<400> 19  
 gaattcgcg gatcctctct atcaactgata gggacttata agtctctatc actgataggg 60  
 atttcacggt tatgggtga 78

<210> 20  
 <211> 38  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 对人工序列的描述: 合成构建体

<400> 20  
 ggcggcgcg atatgactag tcatgcaaat tacgcgct 38

<210> 21  
 <211> 79  
 <212> DNA  
 <213> 人工序列

<220>

<223> 对人工序列的描述：合成构建体

<400> 21

gaattctgg atcctctcta tcaactgatag ggattataag tctctatcac tgatagggat 60  
tttacgttta gggtgattt 79

<210> 22

<211> 61

<212> DNA

<213> 人工序列

<220>

<223> 对人工序列的描述：合成构建体

<400> 22

gatccagct gaccctgaag ttcactctca agagagatga acttcagggt cagctttttg 60  
g 61

<210> 23

<211> 61

<212> DNA

<213> 人工序列

<220>

<223> 对人工序列的描述：合成构建体

<400> 23

aattccaaa aagctgacct tgaagttcat ctctcttgaa gatgaacttc agggctcagct 60  
g 61

<210> 24

<211> 65

<212> DNA

<213> 人工序列

<220>

<223> 对人工序列的描述：合成构建体

<400> 24

gatccagga tgggtggtt tcaattcctt caagagagga attgaaacac caccatcctt 60  
tttgg 65

<210> 25

<211> 65

<212> DNA

<213> 人工序列

<220>

<223> 对人工序列的描述：合成构建体

<400> 25

aattccaaa aaggatggtg gtgtttcaat tcctctcttg aaggattga aacaccacca 60  
tcctg 65

<210> 26  
 <211> 22  
 <212> DNA  
 <213> 人工序列  
  
 <220>  
 <223> 对人工序列的描述: 合成构建体  
  
 <400> 26  
 cggcatcaa ggtgaacttc aa 22  
  
 <210> 27  
 <211> 55  
 <212> DNA  
 <213> 人工序列  
  
 <220>  
 <223> 对人工序列的描述: 合成构建体  
  
 <400> 27  
 cgaattcga gctcgggtacc cgggatcgcg tgaagcgcgc acggcaagag gcgag 55  
  
 <210> 28  
 <211> 44  
 <212> DNA  
 <213> 人工序列  
  
 <220>  
 <223> 对人工序列的描述: 合成构建体  
  
 <400> 28  
 catgttggc caaattttgc ccaggaaatt agcctgtctc tcag 44  
  
 <210> 29  
 <211> 28  
 <212> DNA  
 <213> 人工序列  
  
 <220>  
 <223> 对人工序列的描述: 合成构建体  
  
 <400> 29  
 tttggccaa gtcacaaggg aaggccag 28  
  
 <210> 30  
 <211> 42  
 <212> DNA  
 <213> 人工序列  
  
 <220>  
 <223> 对人工序列的描述: 合成构建体  
  
 <400> 30  
 ctcgacatg acgcgttatt gtgacgaggg stcgcctgcc aa 42  
  
 <210> 31

<211> 45  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 对人工序列的描述: 合成构建体

<400> 31 45  
 gaattcaag cgtatgggag cgcgtgcgac agtattgagc gggggg

<210> 32  
 <211> 45  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 对人工序列的描述: 合成构建体

<400> 32 45  
 cgcagatct tccctgaaga agttagcctg tctctcagta caatc

<210> 33  
 <211> 63  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 对人工序列的描述: 合成构建体

<400> 33 60  
 agatctggc atttccgcag ggtaaagcgc gtgaattttc ctccagagcag accagagcca  
 aca 63

<210> 34  
 <211> 47  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 对人工序列的描述: 合成构建体

<400> 34 47  
 gcctcagac gatgtcgaca cccaattctg aaaagagtaa acagcag

<210> 35  
 <211> 35  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 对人工序列的描述: 合成构建体

<400> 35 35  
 ctgacgacg cgtgccagcc cctgatggg gcgac

<210> 36

<211> 46  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 对人工序列的描述: 合成构建体

<400> 36  
 cgcaogcgc gcccatggtg cgctgtgtac gagacctccc ggggca 46

<210> 37  
 <211> 9416  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 对人工序列的描述: 合成构建体

<400> 37  
 ttggaaggg ctaattcact cccaagaag acaagatata cttgatctgt ggatctacca 60  
 cacacaaggc tacttccctg attagcagaa ctacacacca gggccagggg tcagatatcc 120  
 actgaccttt gyatggtgct acaagctagt accagttgag ccagataagg tagaagaggc 180  
 caataaagga gagaacacca gcttgttaca ccoctgtgagc ctgcatggga tggatgacco 240  
 ggagagagaa gtgttagagt ggaggtttga cagccgccta gcatttcac acgtggcccg 300  
 agagctgcat ccggagtact tcaagaactg ctgatataca gcttgctaca agggactttc 360  
 cgctggggac tttccaggga ggcgtggcct gggcgggact ggggagtggc gagccctcag 420  
 atcctgcata taagcagctg ctttttgect gtactgggaa gcttttagaca agatagagga 480  
 agagcaaaac aaaagtaaga ccaccgcaca gcaggtctct ctggttagac cagatctgag 540  
 cctgggagct ctctggctaa ctagggaacc cactgcttaa gcctcaataa agcttgccct 600  
 gagtgcttca agtagtggtg gcccgctgtg tgtgtgactc tggtaactag agatocctca 660  
 gaccctttta gtcagtgtgg aaaatctcta gcagtggcgc ccgaacaggg acttgaaagc 720  
 gaaaggggaa ccagaggagc tctctcgacg caggactcgg cttgctgaag cgcgcacggc 780  
 aagagggcgag gggcgggcag tggtagtacc gccaaaaatt ttgactagcg gaggctagaa 840  
 ggagagagat ggggtgcgaga gcgtcagtat taagcggggg agaattagat cgcgatggga 900  
 aaaaattcgg ttaaggccag ggggaaagaa aaaatataaa ttaaaacata tagtatgggc 960  
 aagcagggag ctagaacgat tcgcagttaa tcctggcctg ttagaaacat cagaaggctg 1020  
 tagacaaata ctgggacagc tacaaccatc ccttcagaca ggatcagaag aacttagatc 1080  
 attatataat acagtagcaa cctctatttg tgtgcatcaa aggatagaga taaaagcac 1140  
 caaggaagct ttagacaaga tagaggaaga gcaaaaacaa agtaagacca ccgcacagca 1200  
 agcggccgct gatcttcaga cctggaggag gagatatgag ggacaattgg agaagtgaat 1260  
 tatataaata taaagtagta aaaattgaac cattaggagt agcaccacc aaggcaaga 1320  
 gaagagtggt gcagagagaa aaaagagcag tgggaatagg agctttgttc cttgggttct 1380  
 tgggagcagc aggaagcact atgggcgcag cgtcaatgac gctgacggtc caggccagac 1440  
 aattattgtc tggatatagt cagcagcaga acaatttgct gagggtctatt gaggcgcaac 1500  
 agcatctggt gcaactcaca gtctggggca tcaagcagct ccaggcaaga atcctggctg 1560  
 tggaaagata cctaaaggat caacagctcc tggggatttg gggttgctct ggaaaactca 1620  
 tttgcaccac tgctgtgcoct tggaatgcta gttggagtaa taaatctctg gaacagattt 1680  
 ggaatcacac gacctggatg gagtgggaca gagaaattaa caattacaca agcttaatac 1740  
 actocctaat tgaagaatcg caaaaccagc aagaaaagaa tgaacaagaa ttattggaat 1800  
 tagataaatg ggcaagtttg tggaaattggt ttaacataac aaattggctg tggatatata 1860  
 aattattcat aatgatagta ggaggcttgg taggtttaag aatagttttt gotgtacttt 1920  
 ctatagttaa tagagttagg cagggatatt caccattatc gtttcagacc cacctcccaa 1980  
 ccccgagggg acccgacagg cccgaaggaa tagaagaaga aggtggagag agagacagag 2040  
 acagatccat tcgattagtg aacggatctc gacggtatcg attttaaaag aaaagggggg 2100  
 attggggggg acagtgcagg ggaaagaata gtagacataa tagcaacaga catacaaact 2160  
 aaagaactac aaaaaacaaat taaaaaatt caaaattttc gggtttatta cagggacagc 2220  
 agagatccag tttggaattg cgcgttacag ggcgcgtggg gatccccct agagccccag 2280  
 ctggttcttt ccgcctcaga agccatagag cccaccgat cccagcatg cctgctattg 2340  
 tcttcccaat cctccccctt gctgtcctgc cccacccac ccccagaat agaatgacac 2400

ctactcagac	aatgcgatgc	aatttcoctca	ttttattagg	aaaggacagt	gggagtggca	2450
ccttccaggg	tcaaggaagg	cacgggggag	gggcaacaa	cagatggctg	gcaactagaa	2520
ggcacagtcg	aggctgatca	gcgggtttgg	ttctctgacg	ctagcggtag	cacgcgttac	2580
agggcgcgtg	gggatacccc	ctagagcccc	agctggttct	ttccgcotca	gaagccatag	2640
agcccaccgc	atccccagca	tgectgctat	tgtcttccca	atcctccccc	ttgctgtcct	2700
gccccacccc	acccccaga	atagaatgac	acctactcag	acaatgegat	gcaatttcoct	2760
cattttatta	ggaaaggaca	gtgggagtgg	caccttccag	ggccaaggaa	ggcacggggg	2820
aggggcaaac	aacagatggc	tggcaactag	aaggcacagt	cgaggctgat	cagtgcggcc	2880
agatctgggg	catttgttcc	atgtgagtgc	tagtaacagg	ccttgtgtcc	tgttgaagtt	2940
cactgatgcc	ggtcagtcag	tggccaaaac	cggcatcaag	gtgaacttca	acagcataca	3000
gccttcagca	agcctccagg	atccggatcc	ggatggcgtc	tccaggcgat	ctgaaggttc	3060
actaaacgag	ctctgcttat	ataggcctcc	cacogtacac	gctactcga	cccggtacc	3120
gagctcggag	tggtaaactc	gactttcact	ttctctatc	actgataggg	agtggtaaac	3180
tcgaacttca	cttttctcta	tcaactgatag	ggagtggtaa	actcgacttt	cacttttctc	3240
tatcaactgat	agggagtggg	aaactcgact	ttcacttttc	tctatcactg	atagggagtg	3300
gtaaactcga	ctttcacttt	tctctatcac	tgatagggag	tggtaaactc	gacgtcaggg	3360
tcgataatca	agaattogaa	ttccggcggc	cgctctcaa	gggcatcggg	cgactctaga	3420
gggacagccc	cccccaagg	ccccaggga	tgtaattacg	tccctcccc	gctaggggca	3480
gcagcgagcc	gcccgggct	ccgctccggg	ccggcgtccc	ccccgcatcc	ccgagccggc	3540
agcgtgcggg	gacagcccgg	gcacggggaa	ggtggcacgg	gatcgctttc	ctctgaacgc	3600
ttctcgctgc	tctttgagcc	tgccagaccc	tggggggata	cggygaaaaa	gctttaggct	3660
gaaagagaga	tttagaatga	cagaatcata	gaacggcctg	ggttgcaaaag	gagcacagtg	3720
ctcatccaga	tccaaccccc	tgctatgtgc	agggcatca	accagcagcc	caggctgccc	3780
agagccacat	ccagcctggc	cttgaatgcc	tgccagggatg	gggcatccac	agcctccttg	3840
ggcaacctgt	tcagtgctgc	accacctct	gggggaaaaa	ctgctcctc	atatccaacc	3900
caaacctccc	ctgtctcagt	gtaaagccat	tcccccttgt	cctatcaagg	gggagtttgc	3960
tgtgacattg	ttggtctggg	gtgacacatg	tttgccaatt	cagtgcacac	cgagagggca	4020
gatcttgggg	ataaggaagt	gcaggacagc	atggacgtgg	gacatgcagg	tgttgagggc	4080
tctgggacac	tctccaagtc	acagcgttca	gaacagcctt	aaggataaga	agataggata	4140
gaaggacaaa	gagcaagtta	aaacccagca	tggagagggg	cacaaaaagg	ccacagacac	4200
tgatggtccc	tggtcttgag	cctgcattgt	tgatggtctc	tggatgcaag	cagaaagggg	4260
ggaaagcctt	gcctggagag	atacagctgg	gtcagtagga	ctgggacagg	cagctggaga	4320
attgccatgt	agatggtcat	acaatcgtca	aatcatgaag	gctggaaaag	ctccaagatc	4380
cccagacca	acccccaccc	acccaaccgtg	cccactggcc	atgtccctca	gtgccacatc	4440
cccacagttc	ttcatcaoct	ccagggacgg	tgaccccccc	acctccgtgg	gcagctgtgc	4500
cactgcagca	ccgctctttg	gagaaggtaa	atcttgceta	atccagcccc	acctccctct	4560
ggcacaacgt	aaggccatta	tctctcatcc	aaactccaga	cggagtcaat	gaggtggggg	4620
cactagtoat	atgaagccga	attcaattct	aaatggcccc	cctggctgac	cgccccacga	4680
ccccgcocca	ttgacgtcaa	taatgacgta	tgttcccata	gtaacgocaa	tagggacttt	4740
ccattgacgt	caatgggtgg	agtattttacg	gtaaaactgcc	cacttggcag	tacatcaagt	4800
gtatcatatg	ccaagtacgc	ccccatttga	cgtcaatgac	ggtaaatggc	ccgcttgga	4860
ttatgcccag	tacatgacct	tatgggactt	tctactttgg	cagtacatct	acgtattagt	4920
catcgctatt	aacatgctcg	aggtgagccc	ccagttctgc	ttcactctcc	ccatctcccc	4980
ccccccocca	cccccaattt	tgtattttatt	tatttttttaa	ttattttgtg	cagcagatggg	5040
ggcggggggg	gggggggggc	gcgcgccagg	cggggcgggg	cggggcgagg	ggcggggcgg	5100
ggcgagggcg	agaggtgctg	cggcagccaa	tcagagcggc	gcgctccgaa	agtttccctt	5160
tatggcgagg	cggcggcggc	ggcggccocta	taaaaagcga	agcgcgcggc	ggcgggggag	5220
tcgctgcgac	gctgccttgc	ccccgtgccc	cgctccggcg	ccgcctcgcg	ccgcocggcc	5280
cgctctgac	tgaccgcggt	actcccacag	gtgagcgggc	gggacggccc	ttctcctcgg	5340
ggctgtaatt	agcgccttgg	ttaatgacgg	cctgtttctt	ttctgtggct	gcgtgaaagc	5400
cttgaggggc	tccgggaggg	ccctttgtgc	ggggggagcg	gctcgggggg	tgctgctgtg	5460
tgtgtgtgcg	tggggagcgc	cgctgctggc	tccgcgctgc	ccggcggctg	tgagcgcgtc	5520
gggcgcggcg	cggggctttg	tgcgctccgc	agtgctgcgc	aggggagcgc	ggcggggggc	5580
ggtgccccgc	ggtgcggggg	gggctgagag	gggaacaaag	gctgcgtgct	gggtgtgtgc	5640
ctgggggggt	gagcaggggg	tgtgggcggc	tccgtcgggc	tgcaaccccc	cctgcacccc	5700
cctccccgag	ttgctgagca	cgcccggcct	tcgggtcggg	ggctccgtac	ggggcgtggc	5760
gcggggctcg	ccgtgcgggg	cggggggtgg	cggcaggtgg	gggtgcccgg	cggggcgggg	5820
ccgctcgggg	ccggggaggg	ctcgggggag	gggcgcggcg	gccccgggag	cgccggcggc	5880
tgtcgagggc	cggcgagccg	cagccattgc	cttttatggg	aatcgtgcga	gagggcgcag	5940
ggacttccct	tgtcccaaat	ctgtgoggag	ccgaaatctg	ggaggcgcgg	ccgcaccccc	6000
tctagcgggc	gcggggcga	gcgggtgoggc	gcccggcagga	aggaaatggg	cggggagggc	6060

cttcgtggt	ogccgcgccc	cogtcccctt	ctccctctcc	agcctcgggg	ctgtccgccc	6120
ggggacggct	gccttcgggg	gggacggggc	agggcggggg	tccgcttctg	gcgtgtgacc	6180
ggcggctcta	gacaattgta	ctaaccctct	tctctttcct	ctcctgacag	gttggtgtac	6240
agtagcttcc	aatggccagc	cgctgggaca	agtccaaggt	catcaatggc	gocctggagc	6300
tgctgaacgg	cgtcggaatc	gaaggtttaa	caaccogtaa	actogcccag	aagctaggtg	6360
tagagcagcc	tacattgtat	tggcatgtaa	aaaataagcg	ggctttgctc	gacgccttac	6420
ccatcgagat	gctggaccgc	caccacacc	acttctgccc	cctggagggc	gagagctggc	6480
aggacttctt	acgtaataac	gctaaaagtt	ttagatgtgc	tttactaagt	catcgcgatg	6540
gagcaaaagt	acatttaggt	acacggccta	cagaaaaaca	gtatgaaact	ctcgaaaatc	6600
aattagcctt	tttatgcoaa	caaggttttt	cactagagaa	tgcatgttac	gocctgtccg	6660
ccgtcggcca	cttcaccctg	ggctgtgtgc	tggaggagca	ggagcatcaa	gtcgcataaag	6720
aagaagggga	aacacctaact	actgatagta	tgcgcccaat	attacgaoaa	gctatcgaat	6780
tatttgatcg	ccaaggcggc	gagcccgctt	tcctgttcgg	cctggagctg	atcatctgcg	6840
gcctggagaa	gcagctgaag	tgcgagagcg	gcagcgccta	cagcgcgggc	ggaggcggag	6900
gcagtcggcg	cgccgatccc	aaaaagaaaa	gaaaggtagc	acgcgtcggc	ggaggcggaa	6960
gtgggtcccc	ggccgacggc	ctggacgact	tgcacctgga	catgctgccc	gcccagcccc	7020
tggacgactt	cgacctggac	atgctgcccg	cgcagcccc	ggacgacttc	gacctggaca	7080
tgctgcggcg	ogacggctcg	gacgacttcg	acctggacat	gctgcggggg	taactaagta	7140
atttccctct	agcgggatca	attccgcccc	ccccctctcc	ctcccccccc	ctaactgtac	7200
tggccgaagc	cgcttggaat	aaggccgggtg	tgcgtttgtc	tatatgttat	tttccaccat	7260
attgccgctct	tttggcaatg	tgagggcocg	gaaacctggc	cctgtcttct	tgacgagcat	7320
tcctaggggt	ctttccccct	tgcocaaagg	aatgcaaggt	ctggtgaatg	tcgtgaaggga	7380
agcagttcct	ctggaagctt	cttgaagaca	aacaactgct	gtagcgaacc	tttgcaggca	7440
gcggaacccc	acagctggcg	acaggtgcct	ctgcccgcac	aagccactgt	tataagatac	7500
acctgcaaaag	gcggcacaac	cccagtgcca	cgttgtgagt	tggatagttg	tggaaagagt	7560
caaatggctc	tctcaagcg	tattcaacaa	ggggctgaag	gatgcccaga	aggtaccca	7620
ttgtatggga	tctgatctgg	ggcctcgggtg	cacatgcttt	acatgtgttt	agtcgaggtt	7680
aaaaaaaagt	ctaggccccc	cgaaccacgg	ggacgtggtt	ttcctttgaa	aaacacgatg	7740
ataatggcca	caacctgggc	ctcctccgag	gacgtcatca	aggagttcat	goccttcaag	7800
gtgcgcattg	agggctccgt	gaaccggcac	gagttcagga	tcgagggcga	gggctgaggc	7860
cgccctacg	ctggattaca	gaccgccaag	ctgaagggtga	ccaagggcgg	ccccctgcc	7920
ttcgcctggg	acatcctgtc	ccccagttc	cagtaaggct	ccaaggtgta	cgtgaagcac	7980
cccgcgcaca	tcccogacta	caagaagctg	tccttcccgc	agggcttcaa	gtgggagcgc	8040
gtgatgaaact	tcgaggacgg	cggcgtgggtg	accgtgaccc	aggactcctc	cctgcaggac	8100
ggctcactca	tctacaaggt	gaagttcatc	ggcgtgaaat	tcccctccga	cgcccccgtc	8160
atgcagaaga	agactatggg	ctgggagggc	tcaccagagc	gcctgtaccc	ccgcagcggc	8220
gtgctgaagg	gcgagatcca	caaggccctg	aagctgaagg	acggcggcca	ctacctggtg	8280
gagttcaagt	tatctatatg	gccaagaagc	cogtgcagct	gcccggctac	tactacgtgg	8340
actccaagct	ggacatcacc	tcccacaacg	aggactacac	catcgtggag	cagtaacgagc	8400
gcgcccaggg	ccgcccaccac	ctgttccctg	agtgcagctc	gacgtcaccg	ccgacgtcga	8460
ggtgcocgaa	ggaccgcgca	cctggtgcat	gaccocgaag	cccgggtgct	gacgcctcga	8520
caatcaacct	ctggattaca	aaatttgtga	aagattgact	ggatattctta	actatgtttg	8580
tccttttaag	ctatgtggat	acgtgtcttt	aatgcctttg	tatcatgcta	ttgcttcccg	8640
tatggcttbc	atthctctct	ccttgtataa	atcctgggtg	ctgtctcttt	atgaggagtt	8700
gtggcccgct	gtcaggcaac	gtggcgtggg	gtgcaactgtg	tttgctgacg	caacccccac	8760
tggttggggc	attgccacca	cctgtcagct	cctttccggg	actttcgcct	tccccctccc	8820
tattgccacg	gcggaactca	tgcggcctg	ccttgccccg	tgctggacag	gggctcggct	8880
gttgggcaact	gacaattccg	tggtgttgtc	ggggaagctg	acgtcctttc	catggctgct	8940
cgccctgtgt	gccacctgga	ttctgcggcg	gacgtccttc	tgctacgtcc	cttcggccct	9000
caatccagcg	gaccttccct	cccggggcct	gctgcggctc	ctgcggcctc	ttccgcgtct	9060
tgccttccgc	cctcagaoga	gtcggatctc	cctttggggc	gcctccccgc	ctgggtacct	9120
ttaagaccaa	tgacttacia	ggcagctgta	gatcttagcc	actttttaaa	agaaaagggg	9180
ggactggaag	ggctaattca	ctcccaacga	agacaagato	tgotttttgc	ttgtacggtc	9240
tctctggtta	gaccagatct	gagcctggga	tctctctggc	taactagggg	accocctgct	9300
taagcctcaa	taagccttgc	cttgagtgct	tcagtagtg	tgtgcccgtc	tgttgtgtga	9360
ctctggtaac	tagagatccc	tcagaccctt	ttagtcagtg	tggaaaatct	ctagca	9416

<210> 38  
<211> 9396

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;223&gt; 对人工序列的描述: 合成构建体

&lt;400&gt; 38

```

gtcagagttt accactccct atcagtgata gagaaaagtg aaagtcgagt ttaccactcc      60
ctatcagtgat tagagaaaag tgaaagtctga gtttaccact ccctatcagt gatagagaaa      120
agtgaaagtc gagtttacca ctccctatca gtgatagaga aaagtgaaaag tcgagtttac      180
cactccctat cagtgataga gaaaagtgaa agtcgagttt accactccct atcagtgata      240
gagaaaagtg aaagtcgagt ttaccactcc ctatcagtgat tagagaaaag tgaaagtctga      300
gctcgggtacc cgggtcgagt aggcgtgtac ggtgggaggc ctatataagc agagctcgtt      360
tagtgaaccg tcagatcgcc tggagacgcc atccaogctg ttttgacctc catagaagac      420
accgggaccg atccagcctc cgcggccccc aattcgagct cggtaaccgg gatcgcgtga      480
agcgcgcacg gcaagaggcg aggggcccgc actggtgaga gatgggtgag agagcgtcag      540
tattgagcgg gggaaaattg gataagtggg agaaaattcg gtttaaggcca gggggaaaga      600
aaaaatataa attaaaacat ctagtatggg caagcaggga gctagaacga ttcgcagtta      660
atcccgccct gttagaaca gcagaaggct gtagacaaat actgggacag ctacaaccgt      720
cccttcagac aggatcagaa gaacttaaat cattatataa tacaatagca gtccctctatt      780
gtgtgcatca aatgatagat gtaaaagaca ccaaggaagc ttagagaga atagaggaag      840
agcaaaaca cagtaagaaa aaagcacagc aagcagcagc tgacacagga aacagcagcc      900
aggtcagccg aattaccct atagtgcaga acatccaggg gcaaatggta catcaggcca      960
tatcaccag aactttaaat gcatgggtaa atagtagaga agagaaggct ttcagcccag      1020
aagtaatacc catgttttca gcattatcag aaggagccac cccacaagat ttaaacacca      1080
tgctaaacac agtgggggga catcaagcag ctatgcaaat gttaaaagag accatcaatg      1140
aggaagctgc agaatgggat agattgcatc cagtgcagc agggcctgtt gcaccaggcc      1200
agatgagaga accaagggga agtgacatag cacctatcc cagtaggaga aatctataaa      1260
taggtggat cacacataa gtaagaatgt atagccctac cagcattctg gacataagac      1320
tcctgggatt aaataaaaata gtaagaatgt atagccctac cagcattctg gacataagac      1380
aaggacaaa ggaacccttt agagactatg tagaccgatt ctataaaaact ctaagagccg      1440
agcaagcttc caaagaggta aaaaatttga tgacagaaac cttggttggto caaatgcca      1500
accagattg taagactatt ttaaaagcat tgggaccagg agcgcacta gaagaaatga      1560
tgacagcatg tcagggagtg gggggaccgg gccataaagc aagagttttg gctgaagcaa      1620
tgagccaagt aacaaatcca gctaccataa tgatacagaa aggcaatttt aggaacaaa      1680
gaaagactgt taagtgtttc aattgtggca aagtaggca catagccaaa aattgctgta      1740
ccctaggaa aaagggctgt tggaaatgtg gaaaggaagg acaccaaatg aaagattgta      1800
ctgagagaca ggctaatttc ctgggcaaaa tttggccaag tcacaagggga aggccaggga      1860
atcttcttca gagcagacca gagccaacag ccccaccaga agagagcttc aggtttgggg      1920
aagagacaac aactccctct cagaagcagg agccgataga caaggaactg tatcctttag      1980
cttccctcag atcactcttt gcagcagacc cctcgtcaca ataaacgcgt gccagcccc      2040
tgatggggcg aactccacc atagatcacc atagatcact cccctgtgag gaactctgt      2100
cttcacgcag aaagcgtcta gccatggcgt gtcgtgcagc ctccaggacc ccccctcccg      2160
ggagagccat agtggctctg ggaaccgggt agtacaccgg aattgccagg acgaccgggt      2220
ccttcttgg atcaaccgcg tcaatgcctg gagatttggg cgtgcccccg cgagactgct      2280
agccgagtag tgttgggtcg cgaaaggcct tgtggtagct cctgataggg tgcttgcgag      2340
tgccccggga ggtctcgtac acagcgcacc atgggcgccc gtgctcagc attgagcggg      2400
ggaaaattgg ataagtggga gaaaattcgg ttaaggccag ggggaaagaa aaaatataaa      2460
ttaaaccatc tagtatgggc aagcagggag ctagaacgat tcgcagttaa tccggcctg      2520
ttagaaacag cagaaggctg tagacaaata ctgggacagc tacaaccgtc ccttcagaca      2580
ggatcagaag aacttaaatc attatataat acaatagcag tcctctattg tgtgcatcaa      2640
atgatagatg taaaagacac caaggaagct ttagagaaga tagaggaaga gcaaaacaac      2700
agtaagaaaa aagcacagca agcagcagct gacacagga acagcagcca ggtcagccga      2760
aattacccta tagtgagaa catccagggg caaatggtag atcaggccat atcaccaga      2820
actttaaatg catgggtaaa agtagtagaa gagaaggctt tcagcccaga agtaataccc      2880
atgttttcag cattatcaga aggagccacc ccacaagatt taaacaccat gctaaacaca      2940
gtggggggac atcaagcagc tatgcaaatg ttaaaagaga ccatcaatga ggaagctgca      3000
gaatgggata gattgcatcc agtgcaagca gggcctgttg caccaggcca gatgagagaa      3060
ccaaggggaa gtgacatagc aggaactact agtacccttc aggaacaaat aggatggatg      3120
acacataatc cacctatccc agtaggagaa atctataaaa gatggataat cctgggatta      3180

```

aataaaatag	taagaatgta	tagcoctacc	agcattctgg	acataagaca	aggaccaaa	3240
gaacccttta	gagactatgt	agaccgattc	tataaaactc	taagagccga	gcaagcttca	3300
caagaggtaa	aaaattggat	gacagaaacc	ttgttgggtc	aaaatgcgaa	cccagattgt	3360
aagactatft	taaaagcatt	gggaccagga	gcgacactag	aagaaatgat	gacagcatgt	3420
cagggagttg	ggggaccocg	ccataaagca	agagttttgg	ctgaagcaat	gagccaagta	3480
acaaatccag	ctaccataat	gatacagaaa	ggcaatftta	ggaaaccaag	aaagactgtt	3540
aagtgtttca	attgtggcaa	agaagggcac	atagccaaaa	attgcagggc	ccctaggaaa	3600
aagggctgtt	ggaaatgtgg	aaaggaagga	caccaaataga	aagattgtac	tgagagacag	3660
gctaacttct	tcagggaaaga	tctggcattt	ccgcagggta	aagcgcgtga	atfttccctca	3720
gagcagacca	gagccaacag	ccccaccaga	agagagcttc	aggtttgggg	aagagacaac	3780
aactccctct	cagaagcag	agccgataga	caaggaactg	tatccttttag	cttccctcag	3840
atcactcttt	ggcagcgacc	cctcgtcaca	ataaagatcg	gggggcaatt	aaaggaaagt	3900
ctattagata	caggagcaga	tgatacagta	ttagaagaaa	tgaatttgcc	aggaagatgg	3960
aaaccaaaaa	tgataggggg	aattggaggt	tttatcaaa	taggacagta	tgatcagata	4020
cccatagaaa	tctgtggaca	taaagctata	ggtacagtat	tagtaggacc	tacacctgtc	4080
aacataattg	gaagaaatct	gttgactcag	attggttgca	ctctaaattt	tccgattagt	4140
cctattgaaa	ctgtaccagt	aaaattaaag	cccgggatgg	atggtccgaa	agttaaacaa	4200
tggccattga	cagaagcaaa	aataaaagca	ttagttagaa	tttgtacaga	aatggaaaag	4260
gaagggaaaga	tttcaaaaat	tgggcctgaa	aatccataca	atactccagt	atfttctata	4320
aagaaaaaag	acagtactaa	atggagaaaa	ttagttagatt	tcagagaact	taataagagg	4380
actcaagact	tctgggaagt	tcaattagga	ataccacatc	ccgctggatt	aaaaaagaaa	4440
aatcagtaa	cagtactaga	tgtgggtgat	cgctatttct	cagttccctt	agataaagac	4500
ttcaggaat	atactgcatt	taccatacct	agtataaaca	atgagacacc	agggattaga	4560
tatcagtaca	atgtgtctcc	acagggatgg	aaaggtcac	cagcaatatt	ccaaaagtgc	4620
atgacaaaaa	tcttagagcc	ttttagaaag	caaatccag	acatagttat	ctatcagtac	4680
atggatgatt	tgtatgtagg	atctgactta	gaaatagggc	agcatagaac	aaaaatagag	4740
gaactgagac	aacatctgtt	aaggtgggja	tttaccacac	cagacaaaaa	acatcagaaa	4800
gaacctccat	tcctttggat	gggttatgaa	ctccatcctg	ataaatggac	agtacagcct	4860
atagtgtctg	cagaaaaaga	cagctggact	gtcaatgaca	tacagaagtt	agtgggaaaa	4920
ttgaattggg	caagtcagat	ttactcaggg	atcaaagtga	agcagttatg	taaacctctt	4980
aggggaacca	aagcactaac	agaagttagta	acactaacag	aagaagcaga	gctagaactg	5040
gcagaaaaaca	gggaaattct	aaaagaacca	gtacatggag	tgtattatga	cccatcaaaa	5100
gacttaatag	cagaaataca	gaaacagggg	caagcccaat	ggacatatca	aatttatcaa	5160
gagccattta	aaaatctgaa	aacagcaaaa	tatgcaagaa	cgaggggtgc	ccacaotaat	5220
gatgtaaaaa	aattaacaga	ggcagtgcaa	aaaaataacca	cagaatgcat	aataatatgg	5280
ggaaaaactc	ctaaatttag	actgcccata	caaaaagaaa	catgggaaac	atggtggaca	5340
gagtaattgg	aagccacctg	gattcctgaa	tgggagttyg	tcaatacccc	tcccttagtg	5400
aaattatggt	accagttaga	gaaagagccc	atagaaggcg	cagaaacttt	ctatgtagat	5460
ggagcagcta	acagggagac	taaattagga	aaagcaggat	atgttactaa	caaaggaaga	5520
caaaaagttg	tcaccctaac	tgacacaaca	aatcagaaga	ctgagttaga	agcaattcat	5580
ctagctttgc	aggatctctg	attagaagta	aacatagtaa	cagactcaca	atatgcatta	5640
ggaatcattc	agcacaacc	agataaaagt	gaatcagta	tagtcagtca	aataatagag	5700
cagtttaata	aaaaggaaaa	ggtctacctg	gcatgggtac	cagcacacaa	aggaattgga	5760
ggaaatgaac	aagtagataa	attagtcagt	gotggatcca	ggaaagtact	atfttttagat	5820
ggaatagata	aggccaaga	agaacatgag	aaatatcaca	gtaattggag	agccatggct	5880
agtgatttta	acttaccacc	tgtagttagca	aaagaaatag	tagccagctg	tgataaatgt	5940
cagctaaaa	gagaagccat	gcatggacaa	gtagactgta	gtccaggaat	atggcaacta	6000
gattgcacac	atctagaagg	aaaaattatc	ctggtggcgg	ttcatgtagc	cagtggtat	6060
atagaagcag	aagttattcc	agcagagaca	gggcaggaaa	cagcatactt	tctcttaaaa	6120
ttagcaggaa	gatggccagt	aaaaacaata	catacagaca	atggcagcaa	tttcaccagt	6180
accacggtta	aggccgcctg	ttggtgggca	gggatcaagc	aggaatttgg	cattccctac	6240
aatccccaaa	gtcaaggagt	agtagaatct	atgaataaag	aattaaagaa	aattatagga	6300
caggttaagag	atcaggctga	acatcttaaa	acagcagtac	aaatggcagt	atfttatccac	6360
zattttaaaa	gaaaaggggg	gattgggggg	tacagtgca	gggaaagaat	agtagacata	6420
atagcaacag	acatacaaac	taaagaacta	caaaaacaaa	ttacaaaaat	tcaaaatftt	6480
cgggtttatt	acagggacaa	caaagatcca	ctttggaaag	gaccagcaaa	gcttctctgg	6540
aaaggtgaag	gggcagtagt	aatacaagat	aatagtgaca	taaaagtagt	gccaagaaga	6600
aaagcaaaaga	tcattagaga	ttatggaaaa	cagatggcag	gtgatgattg	tgtggcaagt	6660
agacaggatg	aggattagaa	catggataag	tttagtaaaa	cacpatatgt	atatttcaag	6720
gaaagcaaa	gatggtttta	tagacatcac	tatgaaagca	ctcacccaaa	aataagttca	6780
gaagtacaca	tcccactagg	ggatgctaga	ttggtaataa	caacatattg	gggtctgcat	6840

```

acaggagaaa gagattggca tttgggtcat ggagtctccg tagaatggag gaaaaagaga 6900
tatagcacac aagtagacc tgacctagca gaccaactaa ttcactctgta ttactttgat 6960
tgtttttcag aatctgccat aagaaatgcc atattaggac atatagttag toctaggtgt 7020
gaatatcaag caggacataa caaggtagga tctctacagt acctagcact agcagcatta 7080
ataacaccaa aaaggataaa gccacctttg cctagtgtta caaaactaac agggataga 7140
tgyaaccaagc cccagaagac caagggccac agagggagcc atacaatgaa tggacataga 7200
gcttttagaa gaacttaaga atgaagctgt tagacatttt cctaggatat ggctccatgg 7260
cttagggcaa tatatctatg aaacttatgg ggatacttgg gcaggagtgg aagccctagt 7320
aagaactctg caacaactgc tgtttactct tttagaattg ggtgtcgaca tagcagaata 7380
ggcattactc aaogaagaag agcaagaaat ggagccagta gatcctagac tagagccctg 7440
gaagcatcca ggaagccagc ctaaaactgc ttgtaocaaa tgctattgta aaaagtgtt 7500
cttaccattgc caagtttgtt tcatgacaaa aggcttaggc atctcctatg gcaggaagaa 7560
gcygagacag cagcgaagag ctctcaaga cagtcagact catcaagctt ctctatcaa 7620
gcagtaagta gtgcatgtaa tgcaacctat acaaatagca gcaatagtag cattagtagt 7680
ggtaggaata atagcaatag ttgtgtggta aaatattaag acaaagaaaa atagacaggt 7740
taattaaaag aataagtaaa agagcagaag acagtggcaa tgagagtgaa ggagatcagg 7800
aagaattatc agcacttgtg gagatggggc acctgctct tgggatatt gatgatctat 7860
agctacagag aaatataaag tagtaaaaa tagtaaaaa agcagtggga ataggagctc 7920
ccaccacggc aaagagaaga gtggtgcaaa gagaaaaaag agcagtggga ataggagctc 7980
tgttccttgg gttcttggga gcagcaggaa gcactatggg cgcagcgtca atgacgttga 8040
cggtagaggc cagacaatta ttgtctggta tagtgcaaca gcagaacaat ttgctgaggg 8100
ctattgaggc gcaacagcat ctgttgcaac tcacagtctg gggcatcaag cagctccagg 8160
caagagtctc ggctgtggaa agatacctaa aggatcaaca gctcctgggg atttggggtt 8220
gctctgaaa actcatttgc accactgctg tgcttggaa tgctagtgg agtaataaat 8280
ctctgaatca gatttgggat aacatgactt ggtgcagtg ggaaagagaa attgaaaatt 8340
acacagactt aatatacaac ttaattgaag aatcgcagaa ccagcaagaa aagaatgaac 8400
aagaattatt ggaattagat aaatgggcaa gtttgtggaa ttggtttaca ataacaaact 8460
ggctgtggta tataaaaata ttcataatga tagtaggagg cttgataggt ttaagaatag 8520
tttttactgt actttctata gtgaatagag ttaggcaggg atactacca ttgtcgtttc 8580
agaccacct cccaacccc aggggaccg acagggccga aggaatcgaa gaagaaggtg 8640
gagagagaga cagagacaga tccggtcgat tagtgaacgg attcttagca cttttctggg 8700
acgatctgag gagcctgtgc ctcttcagct accaccgctt gagagaacta atcttggttg 8760
taacgaggat tgtggaactt ctgggacgca ggggtggga agccctcaag tattgggtga 8820
gtctcctaca gtattggagc caggaactaa agaatagtgc tgtaacttg cttaatgtca 8880
cagccatagc agtagctgag ggaacagata gggttataga agtagtaca agaacttata 8940
gagctattct ccacatacct agaagaataa gacagggctt ggaaaggctt ttgctataag 9000
atgggtggca agtggcaca acgtatggag ggtggatggc atgctgtaag ggaaagaatg 9060
actcgagtct agagggccc tttaaacccg ctgatcagcc tcgactgtgc cttctagttg 9120
ccagccatct gttgtttgcc cctccccgt gcctccttg accctggaag gtgccactcc 9180
cactgtcctt tcttaataaa atgaggaaat tgcacgcat tgtctgagta ggtgtcattc 9240
tattctgggg ggtgggggtg ggcaggacag caagggggag gattgggaag acaatagcag 9300
gcatgctggg gatgcgggtg gctctatggc ttctgaggcg gaaagaacca gctggggctc 9360
taggggggtat ccccacgcgc cctgtagcgg egcatt

```

<210> 39

<211> 70

<212> DNA

<213> 人工序列

<220>

<223> 对人工序列的描述: 合成构建体

<400> 39

ggaagatct gaattcacca tggatcccaa aaagaaaaga aaggtagcat ccaatttact  
aacgtaaac

60

70

<210> 40

<211> 37

<212> DNA  
<213> 人工序列

<220>  
<223> 对人工序列的描述: 合成构建体

<400> 40  
atgccgctc gagctaatcg ccatcttcca gcaggcg 37

<210> 41  
<211> 44  
<212> DNA  
<213> 人工序列

<220>  
<223> 对人工序列的描述: 合成构建体

<400> 41  
atcgcggat ccctgcctct tctcaccaag atgaactcac tggt 44

<210> 42  
<211> 31  
<212> DNA  
<213> 人工序列

<220>  
<223> 对人工序列的描述: 合成构建体

<400> 42  
tttctcgag tcaactgcccc gcacctgcgc c 31

<210> 43  
<211> 7956  
<212> DNA  
<213> 人工序列

<220>  
<223> 对人工序列的描述: 合成构建体

<400> 43  
ttggaaggg ctaattcact cccaaagaag acaagatata cttgatctgt ggatctacca 60  
cacacaaggo tacttccctg attagcagaa ctacacacca gggccagggg tcagatatcc 120  
actgaccttt ggatgggtgct acaagctagt accagttgag ccagataaag tagaagagggc 180  
caataaagga gagaacacca gcttgttaca ccctgtgagc ctgcatggga tggatgaccc 240  
ggagagagaa gtgttagagt ggaggtttga cagccgccta gcatttoatc acgtggcccg 300  
agagctgcat ccggagtact tcaagaactg ctgatatcga gcttgctaca agggactttc 360  
cgctggggac tttccagga ggcgtggcct gggcgggact ggggagtggc gagccctcag 420  
atcctgcata taagcagctg ctttttgctt gtactgggaa gcttttagaca agatagagga 480  
agagcaaaac aaaagtaaga ccaccgcaca gcaggtctct ctgggttagac cagatctgag 540  
cctgggagct ctctggctaa ctagggaacc cactgcttaa gcctcaataa agcttgctt 600  
gagtgtctca agtagtgtgt gccggtctgt tgtgtgactc tggtaactag agatccctca 660  
gaccttttta gtcagtgtgg aaaatctcta gcagtggcgc ccgaacaggg acttgaaagc 720  
gaaagggaaa ccagaggagc tctctcgacg caggactcgg cttgctgaag cgcgcaoggc 780  
aagagggcag gggcggcgac tgggtgagtac gccaaaaatt ttgactagcg gaggctagaa 840  
ggagagagat ggggtcgaga gcgtcagat taagcggggg agaattagat ccgatggga 900  
aaaaattcgg ttaaggccag ggggaaagaa aaaatataaa ttaaaccata tagtatgggc 960  
aagcagggag ctagaacgat togcagttaa tctctggcctg ttagaaacat cagaaggctg 1020

tagacaaata	ctggggacagc	tacaaccatc	ccttcagaca	ggatcagaag	aacttagatc	1080
attatataat	acagtagcaa	ccctctattg	tgtgcatcaa	aggatagaga	taaaagacac	1140
caaggaagct	ttagacaaga	tagaggaaga	gcaaaaacaa	agtaagacca	ccgcacagca	1200
agcggccgct	gatcttcaga	cctggaggag	gagatatgag	ggacaattgg	agaagtgaat	1260
tatataaata	taaagtagta	aaaattgaac	cattaggagt	agcaccacc	aaggcaaaaga	1320
gaagagtgg	gcagagagaa	aaaagagcag	tgggaatagg	agctttgttc	cttgggttct	1380
tgggagcagc	aggaagcact	atgggcgcag	cgtcaatgac	gctgacggta	caggccagac	1440
aattattgtc	tggtatagt	cagcagcaga	acaatttgct	gagggctatt	gagggcgaac	1500
agcatctgtt	gcaactcaca	gtctggggca	tcaagcagct	ccaggcaaga	atcctggctg	1560
tggaaagata	cctaaaggat	caacagctcc	tggggatttg	gggttgctct	ggaaaactca	1620
tttgcaccac	tgctgtgct	tggaatgcta	gttggagtaa	taaattctctg	gaacagattt	1680
ggaatcacac	gacctggatg	gagtgggaca	gagaaattaa	caattacaca	agcttaatac	1740
actcottaat	tgaagaatcg	caaaaccagc	aagaaaagaa	tgaacaagaa	ttattggaat	1800
tagataaaatg	ggcaagtttg	tggaattgg	ttaacataac	aaattggctg	tggtatataa	1860
aattattcat	aatgatagta	ggaggottgg	taggtttaag	aatagttttt	gctgtacttt	1920
ctatagttaa	tagagttagg	cagggatatt	caccattatc	gtttcagacc	cacctcccaa	1980
ccccgagggg	accogacagg	ccogaaggaa	tagaagaaga	aggtggagag	agagacagag	2040
acagatccat	tcgattagt	aacggatctc	gacggtatcg	attttaaaag	aaaagggggg	2100
attggggggg	acagtgcagg	ggaagaata	gtagacataa	tagcaacaga	catacaaaact	2160
aaagaactac	aaaaacaaat	tacaaaaaatt	caaaattttc	gggtttatta	cagggacagc	2220
agagatccag	tttggaaattg	cgcgttacag	ggcgcgtggg	gataccccct	agagccccag	2280
ctgggtcttt	ccgcctcaga	agccatagag	cccaccgcat	ccccagcatg	cctgctattg	2340
tcttcccaat	cctccccctt	gctgtcctgc	cccaccccac	cccccagaat	agaatgacac	2400
ctactcagac	aatgogatgc	aatttctca	ttttattagg	aaaggacagt	gggagtggca	2460
ccttcacagg	tcaaggaagg	caogggggag	gggcaaacaa	cagatggctg	gcaactagaa	2520
ggccacagtg	aggctgatca	goggtttct	cgagtcactg	ccccgcacct	gogccccagc	2580
ccgcgccagc	gcttctgccc	tggttccctg	gtgccgcctc	ccgcttgccg	aagocgaggg	2640
cgaaggagtt	ccagttgtag	ttcggcaggt	ccttctcccg	ctgcaaccagc	accgcgccct	2700
ggggtgcggg	gatctggcgg	ctgtgggggg	cggacagggc	cggtgctgg	cggtcccgg	2760
agctctcggg	gggcggggac	agcaggtcc	ccttgtcgtc	atcgtctttg	tagtcccagc	2820
ggctcagcct	ggcagtagca	gctggcttcc	tctcgtgca	eggcaggctc	tgctccccgg	2880
ggcccaggag	gcccagggat	tctagctgct	ggcctgtggg	tctagaattc	cccacagagg	2940
ccaccttttc	taatggctcc	ccaaagtggg	tggcacagag	gaaaagcagt	agctgccaa	3000
aaaccagtga	gttcatcttg	gtgagaagag	gcagggatcc	atctctatca	ctgatagggg	3060
gatctctatc	actgataggg	agagctctgc	ttatatagac	ctcccaccgt	acacgcctac	3120
cgcccatttg	cgtcaatggg	goggtttgt	taogacattt	tggaaagtcc	cgttgatttt	3180
ggttccaaaa	caaacctcca	ttgaectcaa	tggggtggag	acttggaaat	ccccgtgagt	3240
caaaccgcta	tccacgcca	ttgatgtact	gccaaaaccg	catcaccatg	gtaatagcga	3300
tgaactaac	aattctaaat	ggcccgcctg	getgaccgcc	caacgacccc	cgcccattga	3360
cgccaataat	gacgtatggt	cccatagtaa	cgccaatagg	gactttccat	tgacgtcaat	3420
gggtggagta	tttacggtaa	actgcccact	tggcagtaga	tcaagtgtat	catatgccaa	3480
gtacgcccc	tattgacgtc	aatgacggta	aatggcccgc	ctggcattat	gcccagtaca	3540
tgaccttatg	ggaacttctc	acttggcagt	acatctacgt	attagtcatc	gctattaaca	3600
tggtcaggtg	gagccccacg	ttctgcttca	ctctcccact	ctcccccccc	tccccacccc	3660
caattttgta	tttattttat	ttttaattat	tttgtgcagc	gatggggggg	gggggggggg	3720
gggggcccgc	gccagggcgg	gcggggcggg	gogagggggg	gggcggggcg	aggcggagag	3780
gtgcggcggc	agccaatcag	agcggcgcgc	tccgaaagtt	tcccttttatg	gogagggcgg	3840
ggcggggcgg	gcccataaaa	aagcgaagcg	cgcgggcggc	ggggagtccg	tgcgacgctg	3900
ccttcgcccc	gtgccccgct	ccgcccgcgc	ctcgcgcgcg	ccgccccggc	tctgactgac	3960
ccggttactc	ccacaggtga	gogggcggga	cgcccttct	ctccgggct	gtaattagcg	4020
cttgggttaa	tgaaggcttg	tttcttttct	gtggttgcgt	gaaagccttg	aggggctccg	4080
ggagggccct	ttgtgcgggg	ggagcggctc	ggggggtgog	tgcgtgtgtg	tgtgcgtggg	4140
gagcgcggcg	tgcggctccg	cgctgcccgg	cggtgtgag	cgctgcgggc	gcggcgcggg	4200
gctttgtgog	ctccgcagtg	tgcgcgaggg	gagcgcggcc	gggggcgggtg	ccccgcggtg	4260
oggggggggg	tgcgagggga	acaaaggctg	cgtgcggggt	gtgtgcgtgg	gggggtgagc	4320
aggggggtgtg	tcggcgtcgg	acccccctg	cacccccctc	cacccccctc	cccaggttgc	4380
tgagcaoggc	ccggttogg	gtgoggggct	ccgtacgggg	cgtggcgcgg	ggctcgcctg	4440
gcogggcggg	gggtggcggc	aggtgggggt	gcccggggcg	gocgggcccg	ctogggccgg	4500
ggagggctcg	ggggaggggc	gocggggccc	ccggagcggc	ggcggctgtc	gagggcgggc	4560
gagccgcagc	cattgcottt	tatggtaatc	gtgcgagagg	gcgcagggac	ttcctttgtc	4620
ccaaatctgt	gogggagcga	aatctgggag	gocccgcgcg	acccccctca	gcgggcgcgg	4680

ggcgaagcgg	tgccggcgcg	gcaggaagga	aatgggcggg	gagggccttc	gtgogtgcce	4740
ggcgcgcggt	ccccttctcc	ctctccagcc	tggggctgt	ccgcgggggg	acgggtgcct	4800
tccgggggga	cggggcaggg	cggggttcgg	ctctggcgt	gtgaccggcg	gctctagaca	4860
attgtactaa	ccttctcttc	tttctctctc	tgacaggttg	gtgtacagta	gcttccacca	4920
tggccagccg	cctggacaag	tccaaggtca	tcaattccgc	attagagctg	cttaatgagg	4980
tccgaatcga	aggtttaaca	acccgtaaac	tgcgccagaa	gctaggtgta	gagcagccta	5040
cattgtattg	gcatgtaaaa	aataagcggg	ctttgctcga	cgcttagcc	attgagatgt	5100
tagataggca	ccatactcac	ttttgccctt	tagaagggga	aagctggcaa	gattttttac	5160
gtaataacgc	taaaagtttt	agatgtgctt	tactaagtca	tccgcatgga	gcaaaaagta	5220
atthaggtac	acggcctaca	gaaaaacagt	atgaaactct	cgaaaatcaa	ttagcctttt	5280
tatgccaaaca	aggtttttca	ctagagaatg	cattgtacgc	cctgtccgcc	gtcggccact	5340
tcaccctggg	ctgtgtgctg	gaggaccaag	agcatcaagt	cgctaaagaa	gaaagggaaa	5400
cacctactac	tgatagtatg	ccgccattat	tacgacaagc	tatcgaatta	tttgatcacc	5460
aaggtgcaga	gcoagccttc	ttattcggcc	ttgaattgat	catatgcgga	ttagaaaaac	5520
aaacttaaatg	tgaaagtggg	tccgcgtaca	gccgcggcgg	aggcggaggc	agtccgcgcg	5580
ccgatcccaa	aaagaaaaga	aaggtagcag	ccatggccta	actcgagttt	ccctctagcg	5640
ggatcaattc	cgcccccccc	ctctccctcc	ccccccctaa	egttactggc	cgaagccgct	5700
tggaaataagg	ccggtgtgcg	tttgtctata	tgttattttc	caccatattg	ccgtcttttg	5760
gcaatgtgag	ggcccggaaa	cctggccctg	tcttcttgac	gagcattcct	aggggtcttt	5820
cccctctcgc	caaaggaatg	caaggtctgt	tgaatgtcgt	gaaggaagca	gttcctctgy	5880
aagcttcttg	aagacaaaaca	acgtctgtag	cgaccctttg	caggcagcgg	aacccccccac	5940
ctggcgcacag	gtgcctctgc	ggccaaaagc	cacgtgtata	agatacacct	gcaaaggcgy	6000
cacaacccca	gtgccacggt	gtgagttgga	tagttgtgga	aagagtcaaa	tggctctcct	6060
caagcgtatt	caacaagggg	ctgaaggatg	cccagaaggt	accccattgt	atgggatctg	6120
atctggggcc	tccgtgcaca	tgcctttacat	gtgttttagtc	gaggttaaaa	aaacgtctag	6180
gocccccgag	ccacggggac	gtggttttcc	tttgaaaaac	acgatgataa	tggccacaac	6240
catggtgagc	aagggcaggg	agctgttcac	cggggtggtg	ccatcctcgg	tccagctgga	6300
cggcgacgta	aacggccaca	agttcagcgt	gtccggcgag	ggcgagggcg	atgccacctc	6360
cggaagcctg	acctggaagt	tcatctgcac	caccggcaag	ctgcccgctg	cctggcccac	6420
cctcgtgacc	acctgacct	acggcgtgca	gtgcttcagc	cgctaccccg	accacatgaa	6480
gcagcacgac	ttcttcaagt	cogccatgoc	cgaaggctac	gtccaggagc	gcaccactct	6540
cttcaagggc	acaggaacct	acaagaccoc	cgccaggtg	aagttccagg	gcgacacct	6600
ggtgaaccgc	atcgagctga	agggcatcga	cttcaaggag	gacggcaaca	tcctggggca	6660
caagctggag	tacaactaca	acagccacaa	cgtctatata	atggccgaca	agcagaagaa	6720
cggcatcaag	gtgaaactca	agatccgcca	caacatcgag	gacggcagcg	tgcagctcgc	6780
cgaccactac	cagcagaaca	ccccatcgg	cgacggcccc	gtgctgctgc	ccgacaacca	6840
ctacctgagc	accagtcgg	ccctgagcaa	agacccccac	gagaagcgcg	atcacatggt	6900
cctgctggag	ttcgtgaccg	ccgcccggat	cactctcggc	atggacgagc	tgtacaagtc	6960
eggactcaga	tctcgacgtc	gacgtcaccg	ccgacgtcga	ggtgcccgaa	ggaccgcgca	7020
cctggtgcat	gaccgcgaag	cccgtgctct	gacgcctcga	caatcaacct	ctggattaca	7080
aaatttbtga	aagattgact	ggtattctta	actatggtgc	tccttttacg	ctatgtggat	7140
acgctgcttt	aatgcctttg	tatcatgcta	ttgcttcccg	tatggctttc	attttctcct	7200
ccttgtataa	atcctggttg	ctgtctcttt	atgaggagtt	gtggcccgtt	gtcaggcaac	7260
gtggcgtggg	gtgcactgtg	tttgcctgacg	caaccccacc	tgggtggggc	attgccacca	7320
cctgtcagct	ccttccgggg	actttcgcctt	tccccctccc	tattgccacg	gcggaactca	7380
tccgcgcctg	ccttgcgccg	tgctggacag	gggctcggct	gttgggcact	gacaattccg	7440
tgggtgtgtc	ggggaagctg	acgtcctttc	catggctgct	cgctgtgttt	gccacctgga	7500
ttctgcgggg	gaegtccctc	tgctacgtcc	cttcggccct	caatccagcg	gaccttctct	7560
ccgcgggctc	gctgcccggc	ctgcggcctc	tccgcctctc	tccgcctcgc	cctcagacga	7620
gtcggatctc	cctttggggc	gcctccccgc	ctgggtacct	ttaagaccaa	tgacttataa	7680
ggcagctgta	gatcttagcc	actttttaa	agaaaagggg	ggactggaag	ggctaattca	7740
ctcccaacga	agacaagatc	tgctttttgc	ttgtaocggtc	tctctggtta	gaccagatct	7800
gagcctggga	gctctctggc	taactagggg	acccactgct	taagcctcaa	taaagcttgc	7860
ccttagtgct	tcaagtagtg	tgtgcccgct	tgttgtgtga	ctctggtaac	tagagatccc	7920
tcagaccctt	ttagtcagtg	tggaaaatct	ctagca			7956

<210> 44  
 <211> 6290  
 <212> DNA

<213> 人工序列

<220>

<223> 对人工序列的描述: 合成构建体

<400> 44

gggctaatt	cactoccaa	gaagacaaga	tatccttgat	ctgtggatct	accacacaca	60
aggctacttc	cctgattagc	agaactacac	accagggcca	ggggtcagat	atccactgac	120
ctttggatgg	tgctacaagc	tagtaccagt	tgagccagat	aaggtagaag	aggccaataa	180
aggagagaac	accagottgt	taacacctgt	gagcctgcat	gggatggatg	accgggagag	240
agaagtgtta	gagtggaggt	ttgacagccg	cctagcattt	catcacgtgg	cccgagagct	300
gcatccggag	tacttcaaga	actgctgata	tcgagcttgc	tacaagggac	tttccgctgg	360
ggactttcca	gggaggcgtg	gcctgggogg	gactggggag	tggcgagccc	tcagatcctg	420
catataagca	gctgcttttt	gcctgtaactg	ggtctctctg	gttagaccag	atctgagcct	480
gggagctctc	tggctaacta	gggaaccac	tgcttaagcc	tcaataaagc	ttgccttgag	540
tgcttcaagt	agtgtgtgcc	cgctgttgt	gtgactctgg	taactagaga	tcctcagac	600
ccttttagtc	agtggtgaaa	atctctagca	gtggcgcccg	aacagggact	tgaaagcgaa	660
agggaacca	gaggagctct	ctcgacgcag	gactcggctt	gctgaagcgc	gcacggcaag	720
aggcgagggg	cggcgactgg	tgagtacgcc	aaaaattttg	actagcggag	gctagaagga	780
gagagatggg	tgcgagagcg	tcagtattaa	gccccggaga	attagatcgc	gatgggaaaa	840
aattcggtta	aggccagggg	gaaagaaaa	atataaatta	aaacatatag	tatgggcaag	900
cagggagcta	gaacgattcg	cagttaatcc	tggcctgta	gaaacatcag	aaggctgtag	960
acaataactg	ggacagctac	aacctccct	tcagacagga	tcagaagaac	ttagatcatt	1020
atataataca	gtagcaaccc	tctatttgtt	gcatcaaagg	atagagataa	aagacaccaa	1080
ggaagcttta	gacaagatag	aggaagagca	aaacaaaagt	aagaccaccg	cacagcaagc	1140
ggccgctgat	cttcagacct	ggaggaggag	atatgaggga	caattggaga	agtgaattat	1200
ataaatataa	agtagtaaaa	attgaaacct	taggagtagc	accaccaag	gcaaagagaa	1260
gagtggtgca	gagagaaaa	agagcagtgg	gaataggagc	tttgttccct	gggttcttgg	1320
gagcagcagg	aagcactatg	ggcgacogt	caatgacgct	gacggtagac	gocagacaat	1380
tattgtctgg	tatagtgcag	cagcagaaca	atttgcctgag	ggctattgag	gcgcaacagc	1440
atctgtttca	actcacagtc	tggggcatca	agcagctcca	ggcaagaatc	ctggctgtgg	1500
aaagatacct	aaaggatcaa	cagctcctgg	ggatttgggg	ttgctctgga	aaactcattt	1560
gcaccactgc	tgtgccttgg	aatgctagt	ggagtaataa	atctctggaa	cagattttgga	1620
atcacacgac	ctggatggag	tgggacagag	aaattaacaa	ttacacaagc	ttaatacact	1680
ccttaattga	agaatcgaa	aaccagcaag	aaaagaatga	acaagaatta	ttggaattag	1740
ataaatgggc	aagtttgtgg	aattggttta	acataacaaa	ttggctgtgg	tatataaaat	1800
tattcataat	gatagtagga	ggcttggtag	gtttaagaat	agtttttgc	gtactttcta	1860
tagtgaatag	agttaggcag	ggatattcac	cattatcggt	tcagacccac	ctccaacct	1920
cgaggggacc	cgacaggccc	gaaggaatag	aagaagaagg	tggagagaga	gacagagaca	1980
gtccattcag	attagtgaac	ggatctcgac	ggtatogatt	ttaaaagaaa	aggggggatt	2040
gggggggtaca	gtgcagggga	aagaatagta	gacataatag	caacagacat	acaaactaaa	2100
gaactacaaa	aacaaattac	aaaaattcaa	aattttcggg	tttattacag	ggacagcaga	2160
gatccagttt	ggaattggaa	ttcaagcttc	gtgaggctcc	ggtgcccgtc	agtgggcaga	2220
gcgcacatcg	cccacagtc	ccgagaagtt	gggggggagg	gtcggcaatt	gaaccgggtc	2280
ctagagaagg	tggcgcgggg	taaactggga	aagtgatgtc	gtgtaactgg	tccgcctttt	2340
tcccaggggt	gggggagaa	cgtatataag	tgcagtagtc	gcccgtgaac	ttctttttcg	2400
caacgggttt	gocgacagaa	cacaggtaa	tgcctgtgtg	ggttcccgg	ggcctggcct	2460
ctttacgggt	tatggccctt	gogtgccttg	aattacttcc	acctggctcc	agtacgtgat	2520
tcttgatccc	gagctggagc	cagggcgggg	ccttgogctt	taggagcccc	ttogcctcgt	2580
gcttgagttg	aggcctggcc	tggcgctgg	ggccgcccgg	tgcgaatctg	gtggcacctt	2640
cgcgccctgc	tcgctgcttt	cgataagtet	ctagccattt	aaaatttttg	atgacctgct	2700
gcgacgcttt	ttttctggca	agatagcttt	gtaaatcggt	gocaggatct	gcacactggt	2760
atltcggttt	ttggcccgc	ggccggcgac	ggggcccgtg	cgtcccagcg	cacatgttcg	2820
gogaggcggg	gcctgcgagc	gogggccacc	agaatcggac	gggggtagtc	tcaagctggc	2880
cggcctgctc	tgggtgctgg	cctcgcccg	ccgtgtatcg	ccccgccttg	ggcggcaagg	2940
ctggcccgg	cgccaccagt	tgctgagcgc	gaaagatggc	cgcttcccgg	ccctgctcca	3000
gggggctcaa	aatggaggac	gogggcctcg	ggagagcggg	cggttgagtc	acccacacaa	3060
aggaaaagg	cctttccgtc	ctcagccgtc	gcttcatgtg	actccacgga	gtaccggggc	3120
ccgtccaggc	acctcgatta	gttctggagc	ttttggagta	cgtcgtcttt	aggttggggg	3180
gaggggtttt	atgcyatgga	gtttcccac	actgagtggg	tggagactga	agttaggcca	3240

gcttggcact	tgatgtaatt	ctccttggaa	tttggccttt	ttgagtttgg	atcttggttc	3300
attotcaage	ctcagacagt	ggttcaaagt	ttttttcttc	catttcaggt	gtcgtgagga	3360
tccaccatgg	ccagccgcct	ggacaagtcc	aaggatcatca	atggcgccct	ggagctgctg	3420
aacggcgtcg	gaatogaagg	tttaacaacc	cgtaaaactcg	cccagaagct	aggtgtagag	3480
cagcctacat	tgtattggca	tgtaaaaaat	aagegggctt	tgctcgacgc	cttaccctac	3540
gagatgctgg	accgccacca	caccctcttc	tgccccctgg	agggcgagag	ctggcaggac	3600
ttcttacgta	ataacgctaa	aagttttaga	tgtgctttac	taagtcatcg	cgatggagca	3660
aaagtacatt	taggtacacg	gcttacagaa	aaacagtatg	aaaotctoga	aaatcaatta	3720
gcctttttat	gccaacaagg	tttttccacta	gagaatgcat	tgtacgccct	gtccgccgtc	3780
ggccacttca	ccctgggctg	tgtgctggag	gagcaggagc	atcaagtctc	taaagaagaa	3840
agggaaacac	ctactactga	tagtatgccc	ccattattac	gacaagctat	cgaattattt	3900
gatcgccaag	tgcccgagcc	cgcttctctg	ttccgcttg	agctgatcat	ctgcygcttg	3960
gagaagcagc	gagcggcagc	gagcggcagc	gcctacagcc	gcggcggagg	cggagggcagt	4020
ccgctgcgcc	atcccaaaaa	gaaaagaaag	gtagcacgcg	tcggcggagg	cggaagtggg	4080
tcccgggccc	acgccctgga	cgacttccgac	ctggacatgc	tgccggccga	cgccctggac	4140
gacttcgacc	tggacatgct	gcccggccgac	gcccctggacg	acttcgacct	ggacatgctg	4200
ccggccgagc	ccctggacga	cttcgacctg	gacatgctgc	cggggtaact	aagtaaggat	4260
ctcgagtttc	caactcagcgg	gatcaatttc	gcccccccc	tctcccctccc	ccccctaac	4320
gttactggcc	gaagccgctt	ggaataaggc	cggtgtgctg	ttgtctatat	gttattttcc	4380
accatattgc	cgtcttttgg	caatgtgagg	gcccggaaac	ctggccctgt	cttcttgacg	4440
agcattccta	ggggtctttc	ccctctcgcc	aaaggaatgc	aaggctctgt	gaatgtcgtg	4500
aaggaagcag	ttctcttggg	agcttcttga	agacaaacaa	cgtctgtagc	gaccctttgc	4560
aggcagcggg	acccccacc	tggcgcaggg	tgctctgctg	gccaaaagcc	acgtgtataa	4620
gatacaactg	caaagggcgg	acaacccccg	tgccacgctg	tgagttggat	agttgtggaa	4680
agagtcaaat	ggctctctcc	aagcgtattc	aacaaggggc	tgaaaggtgc	ccagaaggta	4740
ccccattgta	tgggatctga	tctggggcct	cggtgcacat	gctttacatg	tgttttagctg	4800
aggttaaaaa	aacgtctagg	ccccccgaac	cacggggacg	tggttttcct	ttgaaaaaca	4860
cgatgataat	ggccacaacc	atggccaagc	ctttgtctca	agaagaatcc	accctcattg	4920
aaagagcaac	ggctacaatc	aacagcatcc	ccatctctga	agactacagc	gtcgcacagc	4980
caactctctc	tagcgcagcc	cgcatcttca	ctggtgtcaa	tgtatatcat	tttactgggg	5040
gaccttctgc	agaactcgtg	gtgctgggca	ctgctgctgc	tgccgdcagc	ggcaacctga	5100
cttgtatcgt	cgcgatcgga	aatgagaaca	ggggcatctt	gagccctg	ggacgggtgc	5160
gacaggtgct	tctcgatctg	catcctggga	tcaaagccat	agtgaaggac	agtgatggac	5220
agccgacggc	agttgggatt	cgtgaattgc	tgccctctgg	ttatgtgtgg	gagggctaag	5280
tcgacgtcac	cgccgacgct	gaggtgccc	aaggaccg	caactggtgc	atgacccgca	5340
agcccggtgc	ctgacgcctc	gacaatcaac	ctctggatta	caaaatttgt	gaaagattga	5400
ctggatctct	taactatggt	gctcctttta	cgctatgtgg	atacgtgct	ttaatgcctt	5460
tgtatcatgc	tattgcttcc	cgatggcctt	tcattttctc	ctccttgtat	aaatcctggg	5520
tgctgtctct	ttatgaggag	ttgtggccc	ttgtcaggca	acgtggcgtg	gtgtgcaactg	5580
tgtttgctga	cgcaaccccc	actggttggg	geattgccac	caactgtcag	ctcctttccg	5640
ggactttcgc	tttccccctc	cctattgcca	cggcggaact	catcgccgcc	tgccctgccc	5700
tgcctctggac	aggggctcgg	ctgttgggca	ctgacaatc	cggtggtgtg	tcggggaagc	5760
tgacgtcctt	tccatggctg	ctcgcctgtg	ttgccacctg	gattctgccc	gggacgtcct	5820
tctgctacgt	cccttcggcc	ctcaatccag	cggaccttcc	ttcccggggc	ctgctgccgg	5880
ctctgcccgc	tcttccggct	cttcgccttc	gcccctcagac	gagtcggatc	tccctttggg	5940
ccgcctcccc	gcctgggtao	ctttaagacc	aatgacttac	aaggcagctg	tagatcttag	6000
ccacttttta	aaagaaaagg	ggggactgga	agggctaatt	cactcccaac	gaagacaaga	6060
tctgcttttt	gcttgtaactg	ggtctctctg	gttagaccag	atctgagcct	gggagctctc	6120
tggctaacta	gggaacccac	tgtcttaagcc	tcaataaagc	ttgccttgag	tgcttcaagt	6180
agtgtgtgcc	cgtctgttgt	gtgactctgg	taactagaga	tcctcagac	ccttttagtc	6240
agtgtggaaa	atctctagca	gtagtagttc	atgtcatctt	attattcagt		6290

&lt;210&gt; 45

&lt;211&gt; 4891

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;223&gt; 对人工序列的描述: 合成构建体

&lt;400&gt; 45

gggctaatt	cactcocaaa	gaagacaaga	tatccttgat	ctgtggatct	accacacaca	60
aggctacttc	cctgattagc	agaactacac	accagggcca	ggggtcagat	atccactgac	120
ctttggatgg	tgctacaagc	tagtaccagt	tgagccagat	aaggtagaag	aggccaataa	180
aggagagAAC	accagcttgt	tacaccctgt	gagcctgcat	gggatggatg	accgggagag	240
agaagtgtta	gagtggaggt	ttgacagccg	cctagcattt	catcacgtgg	cccagagagct	300
gcatccggag	tacttcaaga	actgctgata	tcgagcttgc	tacaagggac	tttccgctgg	360
ggactttcca	gggagggcgtg	gcctgggccc	gactggggag	tggcgagccc	tcagatcctg	420
catataagca	gctgcttttt	gcctgtactg	ggtctctctg	gttagaccag	atctgagcct	480
gggagctctc	tggttaacta	gggaaccac	tgcttaagcc	tcataaagc	ttgccttgag	540
tgottcaagt	agtgtgtgcc	cgtctgttgt	gtgactctgg	taactagaga	tcctcagac	600
ccttttagtc	agtgtggaaa	atctctagca	gtggcgcccg	aacagggact	tgaagcgaa	660
agggaacca	gaggagctct	ctcgacgcag	gactcggctt	gctgaagcgc	gcacggcaag	720
aggcgagggg	cggcgactgg	tgagtacgcc	aaaaattttg	actagcggag	gctagaagga	780
gagagatggg	tgcgagagcg	tcagtattaa	gcgggggaga	attagatcgc	gatgggaaaa	840
aattcggtta	aggcaagggg	gaaagaaaa	atataaatta	aaacatatag	tatgggcaag	900
cagggagcta	gaaogattcg	cagttaatcc	tggcctgtta	gaaacatcag	aaggctgtag	960
acaaatactg	ggacagctac	aaccatccct	tcagacagga	tcagaagaac	ttagatcatt	1020
atataataca	gtagcaacc	tctattgtgt	gcatcaaagg	atagagataa	aagacacca	1080
ggaaagcttta	gacaagatag	aggaagagca	aaacaaaagt	aagaccaccg	cacagcaagc	1140
ggccgctgat	cctcagacct	ggaggaggag	atatgagggg	caattggaga	agtgaattat	1200
ataaatataa	attgaacct	attgaacct	taggagttag	accaccaag	gcaagagaa	1260
gagtgggtgca	gagagaaaa	agagcagtg	gaaagagagc	tttgtctctt	gggttcttgg	1320
gagcagcagg	aagcactatg	ggcgcagcgt	caatgacgct	gacggtaacg	gccagacaat	1380
tattgtctgg	tatagtgcag	cagcagaaca	atltgctgag	ggctattgag	gcgcaacagc	1440
atctgttgca	actcacagtc	tggggcatca	agcagctcca	ggcaagaatc	ctggctgtgg	1500
aaagataoct	aaaggatcaa	cagctcctgg	ggatttgggg	ttgctctgga	aaactcattt	1560
gcaccactgc	tgtgccttgg	aatgctagtt	ggagtaataa	atctctggaa	cagatttggg	1620
atcacacgac	ctggatggag	tgggacagag	aaattaaCAA	ttacacaagc	ttaatcaact	1680
cottaattga	agaatcgcaa	aaccagcaag	aaaagaatga	acaagaatta	ttggaattag	1740
ataaatgggc	aagtttgtgg	aatttggttta	acataacaaa	ttggctgtgg	tatataaaat	1800
tattcataat	gatagtagga	ggcttggtag	gtttaagaat	agtttttgc	gtactttcta	1860
tagtgaatag	agttaggcag	ggatattcac	cattatcgtt	tcagaccac	ctcccaacc	1920
cgaggggacc	cgacagggcc	gaaggaatag	aagaagaagg	tggagagaga	gacagagaca	1980
gatccattcg	attagtgaac	ggatctcgac	ggtatcgatt	ttaaaagaaa	aggggggatt	2040
ggggggtaca	gtgcagggga	aagaatagta	gacataatag	caacagacat	acaaactaaa	2100
gaactacaaa	aacaaattac	aaaaattcaa	aattttcggg	tttattacag	ggacagcaga	2160
gatccagttt	ggaatttcca	gtttaccact	ccctatcagt	gatagagaaa	agtgaaagtc	2220
cagtttacca	ctccctatca	gtgatagaga	aaagtgaag	tcgagtttac	cactccctat	2280
gagttagaga	gtcactgaa	agtcagttt	accactccct	atcagtgata	gagaaaagtg	2340
aaagtcaggt	ttaccactcc	ctatcagtga	tagagaaaag	tgaaagtcca	gtttaccact	2400
ccctatcagt	gatagagaaa	agtgaagtc	gagttacca	ctccctatca	gtgatagaga	2460
aaagtgaag	tcgagctcgg	taaccgggct	gagtagcgt	gtacgggtgg	aggootatat	2520
aagcagagct	cgttttagtga	accgtcagat	cgctgggaga	cgccatccac	gctgttttga	2580
cctccataga	agacaccggg	accgatccag	cctccgcggc	cccgaattcg	aattcggatc	2640
caogcgact	agtctcgagc	gagtttccct	ctagcgggat	caattccgcc	ccccccctct	2700
ccctccccc	ccctaacggt	actggccgaa	gocgcttggg	ataaggccgg	tgtgcgtttg	2760
tcctatagtt	atlttccacc	atattgccc	cttttggcaa	tgtgagggcc	cggaaaacctg	2820
gcctgtctt	cttgacgagc	attcctaggg	gtctttcccc	tctcgcaaaa	ggaatgcaag	2880
gtctgttgaa	tgtcgtgaag	gaagcagttc	ctctgggaagc	ttcttgaaga	caaaCaacgt	2940
ctgtagcgac	cctttgcagg	cagcgggaacc	ccccacctgg	cgacaggtgc	ctctgcggcc	3000
aaaagccacg	tgtataagat	acacctgcaa	aggcggcaca	accccagtg	caagtgtgga	3060
gttggatagt	tgtggaaaga	gtcaaatggc	tctcctcaag	ogtattcaac	aaggggctga	3120
aggatgccc	gaaggtacc	cattgtatgg	gatctgatct	ggggcctcgg	tgcacatgct	3180
ttacatgtgt	ttagtcgagg	ttaaaaaac	gtctaggccc	ccogaaccac	ggggacgtgg	3240
tttcccttg	aaaaacacga	tgataatggc	cacaaccatg	gtgactgaat	acaaaccaac	3300
tgttcgctg	gcaactcgtg	atgatgtcc	acgtgcagtt	cgcaacctgg	ctgctgcatt	3360
tgttgactac	cctgcaacc	gtcacactgt	ggaccagac	ogcaacattg	aacgtgtgac	3420
tgaactgcag	gagctgttcc	tgaccctgt	ggccctggac	attggcaag	tgtgggtggc	3480

agatgatggg	gctgctgtgg	cagtgtggac	cacccctgaa	tctgttgaag	ctgggtgcagt	3540
gtttgctgag	attggcccac	gbatggcaga	actgtctggc	agccgcctgg	cagcacaaca	3600
gcagatggaa	ggtctgctgg	caccacaccg	cccaaaagaa	cctgcttggg	tcctggcaac	3660
tgtgggtgtg	agccctgacc	accagggtaa	gggcctgggc	tctgcagtgg	tgctgcctgg	3720
tgtggaagca	gctgaacgtg	caggtgtgcc	tgctttcctg	gagacctcag	ctccacgcaa	3780
cctgcctttc	tatgaacgcc	tgggcttcac	tgtgactgct	gatgtggaag	tgccagaagg	3840
ccacgcact	tgggtcatga	ctcgcaaacc	agggtgctta	gtcgacgtca	ccgccgacgt	3900
cgaggtgccc	gaaggaccgc	gcacctgggt	catgacccgc	aagcccggtg	cctgacgcct	3960
cgacaatcaa	cctctggatt	acaaaatttg	tgaaagattg	actggatttc	ttactatgt	4020
tgctcctttt	acgctatgtg	gatacgctgc	tttaatgcct	ttgtatcatg	ctattgcttc	4080
ccgtatggct	ttcattttct	cctccttgta	taaactcctg	ttgctgtctc	tttatgagga	4140
gttgtggccc	tttgtcaggc	aacgtggcgt	gggtgtgca	gtggttgcctg	acgcaacccc	4200
caactggttg	ggcattgcca	ccacctgtca	gctcctttcc	gggactttcg	ctttccccct	4260
ccctattgcc	acggcggaac	tcctcgccgc	ctgccttgcc	cgctgctgga	caggggctcg	4320
gctgttgggc	actgacaatt	ccgtgggtgt	gtcggggaag	ctgaogtct	ttccatggct	4380
gctcgctgt	gttgccacct	ggattctgcg	cgggacgtcc	ttctgctacg	tccttctggc	4440
cctcaatcca	geggaccttc	ctcccgcg	ctgctgccc	gctctgccc	ctcttcccg	4500
tcttcgctt	cgccctcaga	cgagtcggat	ctccctttgg	gccgcctccc	cgccctgggta	4560
cctttaagac	caatgactta	caaggcagct	gtagatctta	gccacttttt	aaaagaaaag	4620
gggggactgg	aagggtcaat	tcactcccaa	ogaagacaag	atctgctttt	tgcttgtact	4680
gggtctctct	ggttagacca	gatctgagcc	tgggagctct	ctggctaact	agggaaacca	4740
ctgcttaagc	ctcaataaag	cttgccttga	gtgcttcaag	tagtgtgtgc	ccgtctgttg	4800
tgtgactctg	gtaactagag	atccctcaga	cccttttagt	cagtgtggaa	aatctctagc	4860
agtagtagtt	catgtcatct	tattattcag	t			4891

<210> 46

<211> 6031

<212> DNA

<213> 人工序列

<220>

<223> 对人工序列的描述: 合成构建体

<400> 46

gggctaatt	cactccccaaa	gaagacaaga	tatccttgat	ctgtggatct	accacacaca	60
aggctacttc	cctgatttagc	agaactacac	accagggcca	ggggtcagat	atccactgac	120
ctttggatgg	tgctacaagc	tagtaccagt	tgagccagat	aaggtagaag	aggccaataa	180
aggagagaac	accagcttgt	tacaccctgt	gagcctgcat	gggatggatg	accgggagag	240
agaagtgtta	gagtgagggt	ttgacagccg	cctagcattt	catcacgtgg	cccgagagct	300
gcatccggag	tacttcaaga	actgctgata	tcgagcttgc	tacaagggac	tttccgctgg	360
ggactttcca	gggagggcgtg	gcctggggcg	gactggggag	tggcgagccc	tcagatcctg	420
catataagca	gctgcttttt	gcctgtactg	ggctctctct	gttagaccag	atctgagcct	480
gggagctctc	tggctaacta	gggaaccacc	tgcttaagcc	tcataaagc	ttgccttgag	540
tgcttcaagt	agtgtgtgcc	cgtctgttgt	gtgactctgg	taactagaga	tcctctagac	600
ccttttagtc	agtgtggaaa	atctctagca	gtggcgccc	aacagggact	tgaaagcgaa	660
agggaaaacca	gagtagctct	ctcgacgcag	gactcggctt	gctgaagcgc	gcacggcaag	720
agggcagggg	cggcgactgg	tgagtacgcc	aaaaattttg	actagcggag	gctagaagga	780
gagagatggg	tgcgagagcg	tcagtattaa	gcgggggaga	attagatcgc	gatgggaaaa	840
aattcgggta	aggccagggg	gaaagaaaaa	atataaatta	aaacatatag	tatgggcaag	900
cagggagcta	gaaogattcg	cagttaatcc	tggcctgtta	gaaacatcag	aaggctgtag	960
acaaatactg	ggacagctac	aaccatccct	tcagacagga	tcagaagaac	ttagatcatt	1020
atataataca	gtagcaacc	tctattgtgt	gcatcaaagg	atagagataa	aagacaccaa	1080
ggaagcttta	gacaagatag	aggaagagca	aaacaaaagt	aagaccaccg	cacagcaagc	1140
ggccgctgat	cttcagacct	ggaggaggag	atatgagggg	caattggaga	agtgaattat	1200
ataaatataa	agtagtaaaa	attgaaccat	taggagtagc	accaccaag	gcaagagaaa	1260
gagtggtgca	gagagaaaaa	agagcagtyg	gaataggagc	tttgttccct	gggttcttgg	1320
gagcagcagg	aagcactatg	ggcgcagcgt	caatgacgt	gacggtagag	gccagacaat	1380
tattgtctgg	tatagtgag	cagcagaaca	atttgcctgag	ggctattgag	gcgcaacagc	1440
atctgttgca	actcacgtc	tggggcatca	agcagctcca	ggcaagaatc	ctggctgtgg	1500
aaagatacct	aaaggatcaa	cagctcctgg	ggatttgggg	ttgctctgga	aaactcattt	1560

gcaccactgc	tgtgcecttg	aatgctagtt	ggagtaataa	atctctggaa	cagatttgya	1620
atcacacgac	ctggatggag	tgggacagag	aaattaacaa	ttacacaagc	ttaatacact	1680
ccttaattga	agaatcgcaa	aaccagcaag	aaaagaatga	acaagaatta	ttggaattag	1740
ataaatgggc	aagtttgtgg	aattggttta	acataacaaa	ttggctgtgg	tatataaaat	1800
tattcataat	gatagtagga	ggcttggtag	gtttaagaat	agtttttgc	gtacttteta	1860
tagtgaatag	agttaggcag	ggatattcac	cattatcggt	tcagaccac	ctcccaacc	1920
cyaggggacc	cgacaggccc	gaaggaatag	sagaagaagg	tggagagaga	gacagagaca	1980
gatccattcg	attagtgaac	ggatctcgac	ggtatcgatt	ttaaaagaaa	aggggggatt	2040
gggggggtaca	gtgcagggga	aagaatagta	gacataatag	caacagacat	acaaactaaa	2100
gaactacaaa	aacaatttac	aaaaattcaa	aattttcggg	tttattacag	ggacagcaga	2160
gatccagttt	ggaatttoga	gtttaccact	ccctatcagt	gatagagaaa	agtgaaagtc	2220
gagtttacc	ctccctatca	gtgatagaga	aaagtgaag	togagtttac	cactccctat	2280
cagtgataga	gaaaagtga	agtcyagttt	accactccct	atcagtgata	gagaaaagt	2340
aaagtcgagt	ttaccactcc	ctatcagtg	tagagaaaag	tgaaagtcca	gtttaccact	2400
ccctatcagt	gatagagaaa	agtgaagtc	gagtttacc	ctccctatca	gtgatagaga	2460
aaagtgaag	tcgagctcgg	taccgggct	gagtaggct	gtacgggtgg	aggcctatat	2520
aagcagagct	cgtttagtga	accgtcagat	cgctcggaga	cgccatccac	gctgttttga	2580
cctccataga	agacaccggg	accgatccag	cctccggcga	cccgaattcg	gatccaccat	2640
ggaaactoca	aacaccacag	aggactatga	cacgaccaca	gagtttgact	atggggatgc	2700
aactccgtgc	cagaaggtga	acgagagggc	ctttggggcc	caactgctgc	cccctctgta	2760
ctccttggta	tttgtcattg	gcctggttgg	aaacatcctg	gtggctcctg	tccttgtgca	2820
atacaagagg	ctaaaaaca	tgaccagcat	ctacctcctg	aacctggcca	tttctgacct	2880
gctcttctgt	ttcacgcttc	ccttctggat	cgactacaag	ttgaaggatg	actgggtttt	2940
tggtagatgc	atgtgtaaga	tcctctctgg	gttttattac	acaggcttgt	acagcgagat	3000
cttttctatc	atcctgctga	cgattgacag	gtacctggcc	atcgtccacy	ccgtgtttgc	3060
cttgccggca	cggaaccgtca	cttttgggtg	catcaccage	atcatcattt	gggcccctgg	3120
catcttggct	tcocatgccag	gcttatactt	ttccaagacc	caatgggaat	tcactacca	3180
cacctgcagc	cttcaacttc	ctcacgaaa	cctacgagag	tggagctgt	ttcaggctct	3240
gaaactgaac	ctctttgggg	tggatttgcc	tttgttggta	atgatcatct	gctacacagg	3300
tattaraaag	attctgctaa	gacgaccaa	tgagaagaaa	tcocaaagctg	tcogtttgat	3360
ttttgtcatt	atgatcatct	ttttctctct	ttggaccccc	tacaatttga	ctatacttat	3420
ttctgttttc	caagacttcc	tgttcaccca	tgagtgtgag	cagagcagac	atttggacct	3480
ggctgtgcaa	gtgacggagg	tgatcgccta	cacgcaactgc	tgtgtcaacc	cagtgatcta	3540
cgcttctgtt	ggtgagagg	tccggaagta	cctcgggcag	ttgttccaca	ggcgtgtggc	3600
tgtgcacctg	gttaaattgg	tccccttcc	ctcogtggac	aggctggaga	gggtcagctc	3660
cacatctccc	tcacacgggg	agcatgaact	ctctgtggg	ttcgaaaacc	tgtattttca	3720
gggcgctcga	ggagattaca	aagatgacga	cgataagcgc	aacggccatc	atcacctca	3780
ccatcaccac	catcactaac	gagtttccct	ctagcgggat	caattccgoc	ccccccctct	3840
ccctcccccc	ccctaaegtt	actggccgaa	gccgcttgg	ataaggccgg	tgtgcgtttg	3900
tctatatggt	attttccacc	atattgccc	cttttggcaa	tgtgagggcc	cggaaccctg	3960
gcccctgtct	cttgaogagc	attcctagg	gtcttccccc	tctcgocaaa	ggaatgcaag	4020
gtctgttgaa	tgtcgtgaag	gaagcagttc	ctctggaagc	ttcttgaaga	caaacaacgt	4080
ctgtagcgac	cctttcgagg	cagcggaaac	cccacactgg	ogaagggtgc	ctctcgggoc	4140
aaaagccacg	tgtataagat	acacctgcaa	aggcggcaca	accccagtg	caogttgtga	4200
gttgatagat	tgtggaagaa	gtcaaatggc	tctcctcaag	cgtattcaac	aaggggctga	4260
aggatgccca	gaaggtaccc	cattgtatgg	gatctgatct	ggggcctcgg	tgccatgct	4320
ttacatgtgt	ttagtcgagg	ttaaaaaac	gtctaggccc	ccggaaccac	ggggaogtgg	4380
tttctccttg	aaaaacagca	tgataatggc	cacaaccatg	gtgactgaat	acaaaccaac	4440
tgttcgctg	gcaactcgtg	atgatgttcc	acgtgcagtt	cgcacccctg	ctgctgcatt	4500
tgetgactac	cctgcaacc	gtcacactgt	ggaccagac	cgccacattg	aacgtgtgac	4560
tgaactgcag	gagctgttcc	tgaccctgtg	ggcctggac	attggcaag	tgtgggtggc	4620
agatgatggt	gctgctgtgg	cagtgtggac	cacccctgaa	tctgttgaag	ctggtgcagt	4680
gtttgctgag	attggcccac	gcattggcaga	actgtctggc	agccgcctgg	cagcacaaca	4740
gcagatggaa	ggtctgctgg	caccacacog	cccaaaagaa	cctgcttgg	tccttggcaac	4800
tgtgggtgtg	agccctgacc	accaggttaa	gggcctgggc	tctgcagtg	tgctgctgg	4860
tgtggaagca	gctgaacgtg	caggtgtgcc	tgtttcctg	gagacctcag	ctccacgcaa	4920
cctgcctttc	tatgaacgcc	tgggcttcac	tgtgactgct	gatgtggaag	tgccagaagg	4980
cccacgcact	tggtgcatga	ctcgcaaac	agggtgctta	gtcgacgtca	ccgcccagct	5040
cgaggtgccc	gaaggaccgc	gcacctgggt	catgacccgc	aagcccgggt	cctgacgcct	5100
cgacaatcaa	cctctggatt	acaaaatttg	tgaagattg	actggtattc	ttaactatgt	5160
tgctcctttt	acgetatgtg	gatacgctgc	tttaatgccc	ttgtatcatg	ctattgcttc	5220

```

cogtatggct ttcattttct cctcottgta taaatcctgg ttgctgtctc tttatgagga 5280
gttgtggccc gttgtcaggc aacgtggcgt ggtgtgcact gtgtttgctg acgcaacccc 5340
cactggttgg ggcattgcca ccacctgtca gctcctttcc gggactttcg ctttcccct 5400
ccctattgcc acggcggaac tcatcgccgc ctgccttgcc cgtgctgga caggggctcg 5460
gctgttgggc actgacaatt ccgtgggtgtt gtcgggggaag ctgacgtcct ttccatggct 5520
gctcgcctgt gttgccacct ggattctgoc cgggacgtcc ttctgctacg tcccttggc 5580
cctcaatcca gcggaacctt cttcccggg cctgctgocg gctctgoggc ctcttccggc 5640
tcttcgcctt cgcctcaga cgagtggat ctcctttgg gccgcctccc cgcttggtta 5700
cctttaagac caatgactta caaggcagct gtagatctta gccacttttt aaaagaaaag 5760
gggggactgg aagggtaat tcactcccaa cgaagacaag atctgctttt tgcttgtact 5820
gggtctctct ggttagacca gatctgagcc tgggagctct ctggctaact agggaaacca 5880
ctgcttaagc tcaataaag cttgccttga gtgcttcaag tagtgtgtgc ccgtctgttg 5940
tgtgactctg gtaactagag atccctcaga cccttttagt cagtgtggaa aatctctagc 6000
agtagtagtt catgtcatct tattatcag t 6031

```

- <210> 47
- <211> 9372
- <212> DNA
- <213> 人工序列

<220>  
 <223> 对人工序列的描述: 合成构建体

```

<400> 47
gggctaatt cactccaaa gaagacaaga tacccttgat ctgtggatct accacacaca 60
aggctacttc cctgattagc agaactacac accagggcca ggggtcagat atccactgac 120
ctttggatgg tgctacaagc tagtaccagt tgagccagat aaggtagaag aggccataaa 180
aggagagaac accagcttgt tacacctgt gagcctgcat gggatggatg acccggagag 240
agaagtgtta gagtggaggt ttgacagccg cctagcattt catcacgtgg cccgagagct 300
gcatccggag tacttcaaga actgctgata tcgagcttgc tacaagggac tttccgctgg 360
ggactttcca gggaggcgtg gctggggcgg gactggggag tggcgagccc tcagatcctg 420
catataagca gctgcttttt gcctgtactg ggtctctctg gtttagaccag atctgagcct 480
gggagctctc tggctaacta gggaaaccac tgcttaagcc tcaataaagc ttgccttag 540
tgcttcaagt agtgtgtgcc cgtctgttgt gtgactctgg taactagaga tccctcagac 600
ccttttagtc agtgtggaaa atctctagca gtggcgccc aacagggact tgaagcga 660
agggaaacca gaggagctct ctgcagcag gactcggctt gctgaagcgc gcacggcaag 720
aggcgagggg cggcagctgg ttagtacgcc aaaaatttt actagcggag gctagaagga 780
gagagatggg tgcgagagcg tcagtattaa gggggggaga attagatcgc gatgggaaaa 840
aattcggtta agccagggg gaaagaaaa atataaatta aaacatatag tatgggcaag 900
cagggagcta gaacgattcg cagttaatcc tggcctgtta gaaacatcag aaggctgtag 960
acaaatactg ggacagctac aacctccct tcagacagga tcagaagaac ttagatcatt 1020
atataataca gtagcaacc tctatttgtt gcatcaaagg atagagataa aagacaccaa 1080
ggaagcttta gacaagatag aggaagagca aaacaaaagt aagaccaccg cacagcaagc 1140
ggcgcctgat cttcagacct ggaggaggag atatgagggc caattggaga agtgaattat 1200
ataaatataa agtagtaaaa attgaaccat taggagtagc accaccaag gcaaagagaa 1260
gagtggtgca gagagaaaa agagcagtg gaaataggagc tttgttcctt gggttcttgg 1320
gagcagcagg aagcactatg ggccagcgt caatgacgct gacggtaqag gccagacaat 1380
tattgtctgg tatagtgcag cagcagaaca atttgctgag ggtatattgag gcgcaacagc 1440
atctgttgca actcacagtc tggggcatca agcagctcca ggcaagaatc ctggctgtgg 1500
aaagatacct aaaggatcaa cagctcctgg ggatttgggg ttgctctgga aaactcattt 1560
gcaccactgc tgtgccttgg aatgctagtt ggagtaataa atctctggaa cagatctgga 1620
atcacacgac ctggatggag tgggacagag aaattaacaa ttacacaagc ttaatacact 1680
ccttaattga agaatcgcaa aaccagcaag aaaagaatga acaagaatta ttggaattag 1740
ataaatgggc aagtttgttg aattggttta acataacaaa ttggctgtgg tatataaaat 1800
tattcataat gatagttaga ggttggtag gtttaagaat agtttttgcgt gtactttcta 1860
tagtgaatag agttaggcag ggatattcac cattatcgtt tcagaccacc ctcccaacc 1920
cgaggggacc cgacagggcc gaaggaatag aagaagaagg tggagagaga gacagagaca 1980
gatccattcg attagtgaac ggtatctgac ggtatcgatt ttaaaagaaa aggggggatt 2040

```

gggggggtaca	gtgcagggga	aagaatagta	gacataatag	caacagacat	acaaactaaa	2100
gaactacaaa	aacaaattac	aaaaattcaa	aattttcggg	tttattacag	ggacagcaga	2160
gatccagttt	ggaatttcga	gtttaccaot	ccctatcagt	gatagagaaa	agtgaaagtc	2220
gagtttacca	ctccctatca	gtgatagaga	aaagtgaag	tcgagtttac	cactccctat	2280
cagtgataga	gaaaagtga	agtcgagttt	accactccct	atcagtgata	gagaaaagtg	2340
aaagtcgagt	ttaccactcc	ctatcagtga	tagagaaaag	tgaaagtcga	gtttaccact	2400
ccctatcagt	gatagagaaa	agtgaagtc	gagtttacca	ctccctatca	gtgatagaga	2460
aaagtgaag	tcgagctcgg	taccocgggtc	gagtaggcgt	gtacgggtggg	aggcctatat	2520
aagcagagct	cgtttagtga	accgtcagat	cgctggaga	cgccatccac	gctgttttga	2580
cctccataga	agacaccggg	accgatccag	cctccggggc	cccgaattcg	aattcatgca	2640
gaggtcgctt	ctggaaaagg	ccagcgttgt	ctccaaactt	tttttcagct	ggaccagacc	2700
aatcttgagg	aaaggatata	gacagcctct	ggaattgtca	gacatatacc	aaatcccttc	2760
tgttgattct	gctgacaatc	tatctgaaaa	attggaaaaga	gaatgggata	gagagctggc	2820
ttcaaagaaa	aatcctaaac	tcattaatgc	ccttcggcga	tgttttttct	ggagatttat	2880
gttctatgga	atctttttat	atthagggga	agtcaccaa	gcagtacagc	ctctcttact	2940
gggaagaatc	atagcttctt	atgacccgga	taacaaggag	gaacgctcta	tcogogattta	3000
tetaggcata	ggcttatgct	ttctctttat	tgtaggaca	ctgctcctac	accagccat	3060
ttttggcctt	actcacattg	gaatgcagat	gagaatagct	atgtttagtt	tgatttataa	3120
gaagacttta	aagctgtcaa	gcctgttctt	agataaaata	agtattggac	aaactgttag	3180
tctcctttcc	aacaacctga	acaaatttga	tgaaggactt	gcattggcac	atctcgtgtg	3240
gatcgctcct	ttgcaagtgg	cactcctcat	ggggctaact	tgaggagtgt	tacaggcgctc	3300
tgcttctgt	ggacttggtt	tcctgatagt	ccttgccctt	tttcaggctg	ggctagggag	3360
aatgatgatg	aagtacagag	atcagagagc	tggaagatc	agtgaagac	ttgtgattac	3420
ctcagaaatg	attgaaaata	tccaatctgt	taaggctaac	tgctgggaag	aagcaatgga	3480
aaaaatgatt	gaaaacttaa	gacaaacaga	actgaaactg	actcgggaagg	cagocctatgt	3540
gagatacttc	aatagctcag	ccttcttctt	ctcagggttc	tttggtgtgt	ttttatctgt	3600
gcttccctat	gcactaatca	aaggaatcat	cctccggaaa	atattcacca	ccatctcatt	3660
ctgcattggt	ctgoccatgg	cggtcactcg	gcaatttccc	tgggctgtac	aaacatggta	3720
tgactctott	ggagcaataa	acaaaataca	ggatttctta	caaaagcaag	aatataagac	3780
attggaatat	acttaacga	ctacagaagt	agtatggag	aatgtaacag	cctctggga	3840
ggagggattt	ggggaattat	ttgagaagc	aaaacaaaac	aataacaata	gaaaaacttc	3900
taatggtgat	gacagcctct	tcttcagtaa	tttctcactt	cttggtactc	ctgtcctgaa	3960
agatattaat	ttcaagatag	aaagaggaca	gttgttggcg	gttgctggat	ccactggagc	4020
aggcaagact	tcacttctaa	tgatgattat	gggagaactg	gagccttcag	agggtaaaat	4080
taagcacagt	ggaagaattt	cattctgttc	tcagtttccc	tggtattatgc	ctggcaccat	4140
taaagaaaaa	atcacttttg	gtgtttccta	tgatgaatat	agatacagaa	gcgtctcaa	4200
agcagccaa	ctagaagagg	acatctccaa	gtttgcagag	aaagacaata	tagtctctgg	4260
agaaggtgga	atcacactga	gtggagggtca	acgagcaaga	atctctttag	caagagcagt	4320
atacaaagat	gctgatattg	atattattaga	ctctcctttt	ggatacctag	atgttttaac	4380
agaaaaagaa	atatttgaaa	gctgtgtctg	taaactgatg	gctaacaaaa	ctaggatttt	4440
ggtcacttct	aaaatggaac	atlttaagaa	agctgacaaa	atattaattt	tgaatgaagg	4500
tagcactcat	ttttatggga	cattttcaga	actccaaaat	ctacagccag	actttagctc	4560
aaaactcatg	ggatgtgatt	ctttcgacca	atthagtgca	gaaagaagaa	attcaatcct	4620
aactgagacc	ttacaccggt	tctcattaga	aggagatgct	cctgtctcct	ggacagaaac	4680
aaaaaaacaa	tcttttaaac	agactggaga	gtttggggaa	aaaaggaaga	attctattct	4740
caatccaatc	aactctatac	gaaaattttc	cattgtgcaa	aagactccct	tacaaatgaa	4800
tggcatcgaa	gaggattctg	atgagccttt	agagagaagg	ctgtccttag	taccagattc	4860
tgagcagggg	gaggcgatac	tgctctcgcat	cagcgtgatc	agcactggcc	ccaogcttca	4920
ggcaocgaagg	aggcagctctg	tcctgaaacct	gatgacacac	tcagttaaac	aaggctcagaa	4980
cattcacccga	aagacaacag	catccacacg	aaaagtgtca	ctggcccctc	aggcaaacctt	5040
gactgaactg	gatatatatt	caagaagggt	atctcaagaa	actggcttgg	aaataagtgga	5100
agaaatatac	gaagaagact	taaaggagtg	cctttttgat	gatatggaga	gcataaccagc	5160
agtgactaca	tggaacacat	accttcgata	tattactgtc	cacaagagct	taatttttgt	5220
gctaatttgg	tgcttagtaa	ttttctggc	agaggtggct	gcttcttggg	ttgtgctgtg	5280
gctccttggg	aacactcctc	ttcaagacaa	agggatagtg	actcatagta	gaaataacag	5340
ctatgcagtg	attatcacca	gcaccagttc	gtattatgtg	ttttacattt	acgtgggagt	5400
agccgacact	ttgcttgcta	tgggattctt	cagaggtcta	ccactgggtc	atactctaata	5460
cacagtgctg	aaaattttac	accacaaaat	gttacattct	gttcttcaag	cacctatgtc	5520
aacctcaac	acgttgaaag	caggtgggat	tcttaataga	ttctccaaag	atatagcaat	5580
tttgatgac	cttctgcttc	ttaccatatt	tgacttcac	cagttgttat	taattgtgat	5640
tggagctata	gcagttgtcg	cagttttaca	accctacac	tttgttgcaa	cagtgccag	5700

gatagtggct	tttattatgt	tgagagcata	tttctcccaa	acctcacagc	aactcaaaca	5760
actggaatct	gaaggcagga	gtccaatttt	cactcatctt	gttacaagct	taaaaggact	5820
atggacactt	cgtgccttcc	gacggcagcc	ttactttgaa	actctgttcc	acaaagctct	5880
gaatttacat	actgccaaact	ggttcttgta	cctgtcaaca	ctgcgctggg	tccaaatgag	5940
aatagaaatg	atttttgtca	tcttcttcat	tgtgttacc	ttcatttcca	ttttaacaac	6000
aggagaagga	gaaggaagag	ttggtattat	cctgacttta	gccatgaata	tcatgagtac	6060
attgcagtg	gctgtaaact	ccagcataga	tgtggatagc	ttgatgcat	ctgtgagccg	6120
agtctttaag	ttcattgaca	tgccaacaga	aggtaaacct	accaagtcaa	ccaaaccata	6180
caagaatggc	caactctcga	aagttatgat	tattgagaat	tcacacgtga	agaaagatga	6240
catctggccc	tcagggggcc	aaatgactgt	caaagatctc	acagcaaaat	acacagaagg	6300
tggaatgccc	atattagaga	acatttcctt	ctcaataagt	ctcggccaga	gggtgggccc	6360
cttgggaaga	actggatcag	ggaagagtac	tttgttatca	gcttttttga	gactactgaa	6420
cactgaagga	gaaatccaga	tcgatgggtg	gtcttgggat	tcataaactt	tgcaacagtg	6480
gaggaaagcc	tttggagtga	taccacagaa	agtattttat	ttttctggaa	catttagaaa	6540
aaacttggat	ccctatgaac	agtggagtga	tcaagaaata	tggaaagtgt	cagatgaggt	6600
tgggctcaga	tctgtgatag	aacagtttcc	tgggaaget	gactttgtcc	ttgtggatgg	6660
gggctgtgtc	ataagcctg	gccacaagca	gttgatgtgc	ttggctagat	ctgttctcag	6720
taaggcgaag	atcttgcctg	ttgatgaaac	cagtgtcat	ttggatccag	taacatcaca	6780
aataattaga	agaactctaa	aacaagcatt	tgtgtattgc	acagtaattc	tctgtgaaca	6840
caggatagaa	gcaatgctgg	aatgccaaaca	atttttggtc	atagaagaga	acaaagtgcg	6900
gcagtacgat	tccatccaga	aactgctgaa	cgagaggagc	ctcttccggc	aagccatcag	6960
cccctccgac	agggtgaagc	tctttcccca	ccggaactca	agcaagtgca	agtctaagcc	7020
ccagattgct	gctctgaaag	aggagacaga	agaagagtg	caagatacaa	ggctttagct	7080
cgaggagatt	acaagatga	cgacgataag	cgcaaccggc	atcatcacca	taccacttaa	7140
cgagtctccc	tctagcggga	tcaattccgc	ccccccctc	tccttcccc	ccctaactaa	7200
tactggccga	agccgcttgg	aataaggccg	gtgtgctgtt	gtctatatgt	tattttccac	7260
catattgccc	tcttttggca	atgtgagggc	ccggaaacct	ggccctgtct	tcttgacgag	7320
catcctaggg	ggtctttccc	ctctcgccaa	aggaatgcaa	ggtctgttga	atgtcgtgaa	7380
ggaagcagtt	ctctggaag	cttcttgaag	acaaacaacg	tctgtagoga	ccctttgcag	7440
gcagcggaac	ccccacatg	gcgacaggtg	cctctgcggc	caaaagccac	gtgtataaga	7500
tacacctgca	aaggcggcac	aaccccagtg	ccactgtgtg	agttggatag	ttgtggaaaag	7560
agtcaaatgg	ctctctcaa	gcgtattcaa	caaggggctg	aaggatgccc	agaaggatcc	7620
ccattgtatg	ggatctgatc	tggggcctcg	gtgcacatgc	tttacatgtg	tttagtccag	7680
gttaaaaaaa	cgtctaggcc	ccccgaacca	cggggacgtg	gttttctctt	gaaaaacacc	7740
atgataatgg	ccacaacctg	ggtgactgaa	tacaaaccaa	ctgttcgccc	ggcaactcgt	7800
gatgatgttc	caagtgcagt	tcgcacctg	gctgctgcat	ttgctgacta	ccctgcaacc	7860
cgtcacactg	tggaccaga	ccgccacatt	gaactgttga	ctgaaactgca	ggagctgttc	7920
ctgaccctgt	tgggctgga	cattggcaaa	gtgtgggtgg	cagatgatgg	tgtctgtgtg	7980
gcagtgtgga	ccacccctga	atctgttgaa	gctgggtgcag	tgtttgctga	gatttggccc	8040
gcagatggcag	aactgtctgg	cagccgcctg	gcagcacaac	agcagatgga	aggtctgctg	8100
gcaccacacc	gcccgaaga	acctgcttgg	ttcctggcaa	ctgtgggtgt	gagccctgac	8160
caccagggtg	agggcctggg	ctctgcagtg	gtgctgctg	gtgtggaagc	agctgaaact	8220
gcagggtgtg	ctgctttcct	ggagacctca	gctccacgca	acctgccttt	ctatgaaagc	8280
ctgggcttca	ctgtgactgc	tgatgtggaa	gtgccagaag	gcccacgcac	ttggtgcatg	8340
actcgcaaac	caggtgotta	agtcgacgtc	accgcccagc	tcgaggtgoc	cgaaggaccg	8400
gcacactggt	goatgaacct	caagcccggg	gacctgacgc	tcgacaatca	acctctggat	8460
tacaaaatgt	gtgaaagatt	gactggatatt	cttaactatg	ttgctccttt	taagctatgt	8520
ggatagcctg	ctttaaagcc	tttgtatcat	gctattgctt	cccgtatggc	tttcattttc	8580
tcctccttgt	ataaatcctg	gttgcctgtc	ctttatgagg	agttgtggcc	cgttgtcagg	8640
caaactggcg	tgggtgtcac	tgtgtttgct	gacgcaacce	ccactgggtg	gggcatgccc	8700
accacotgtc	agctccttcc	cgggactttc	gctttcccc	tccttattgc	cacggcggaa	8760
ctcatcgccg	cctgccttgc	ccgctgctgg	acaggggctc	ggctgttggg	cactgacaat	8820
tccgtgggtg	tgtccgggaa	gctgacgtcc	tttccatggc	tgtctgctg	tggtgcccac	8880
tggattctgc	gcggaagctc	cttctgctac	gtcccttccg	ccctcaatcc	agcggacctt	8940
ccttcccggc	gcctgctgoc	ggctctgocg	cctcttccgc	gtcttgcctt	tcgcccctcag	9000
acgagtccga	tctccccttg	ggccgcctcc	ccgctgggtg	acctttaaga	ccaatgactt	9060
acaaggcagc	tgtagatctt	agccactttt	taaaagaaaa	ggggggactg	gaagggctaa	9120
ttcactccca	acgaagacaa	gatctgcttt	ttgcttgtac	tgggtctctc	tgggttagacc	9180
agactcagc	ctgggagctc	tctggctaac	tagggaacct	actgottaag	cctcaataaa	9240
gottgccttg	agtgactcaa	gtagtgtgtg	ccgctgtgtt	gtgtgactct	ggtaactaga	9300
gatccctcag	acccttttag	tcagtggtgga	aaactctctag	cagtagtagt	tcatgtcctc	9360

ttattattca gt

9372

&lt;210&gt; 48

&lt;211&gt; 9384

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;223&gt; 对人工序列的描述: 合成构建体

&lt;400&gt; 48

gggctaatt	cactcccaaa	gaagacaaga	tatccttgat	ctgtggatct	accacacaca	60
aggctacttc	cctgattagc	agaactacac	accagggcca	ggggtcagat	atcccactgac	120
ctttggatgg	tgctacaagc	tagtaccagt	tgagccagat	aaggtagaag	aggccaataa	180
aggagagaac	accagottgt	tacaccctgt	gagcctgcat	gggatggatg	accgggagag	240
agaagtgtta	gagtgagggt	ttgacagccg	cctagcattt	catcacgtgg	cccagagagct	300
gcatccggag	tacttcaaga	actgctgata	togagcttgc	tacaagggac	tttccgctgg	360
ggactttcca	gggagggcgtg	gcctgggocg	gactggggag	tygcgagccc	tcagatcctg	420
catataagca	gctgcttttt	gcctgtactg	ggtctctctg	gttagaccag	atctgagcct	480
gggagctctc	tggctaacta	gggaaccac	tgcttaagcc	tcaataaagc	ttgccttgag	540
tgcttcaagt	agtgtgtgcc	cgtctgttgt	gtgactctgg	taactagaga	tcctcagac	600
ccttttagtc	agtgtggaaa	atctctagca	gtggcgcccg	aacagggact	tgaaagcgaa	660
agggaaacca	gaggagctct	ctcgacgcag	gactcggctt	gctgaagcgc	gcacggcaag	720
aggcgagggg	cggcgactgg	tgagtaagcc	aaaaattttg	actagcggag	gctagaagga	780
gagagatggg	tgccgagagcg	tcagtattaa	gcgggggaga	attagatcgc	gatgggaaaa	840
aattcgggta	aggccagggg	gaaagaaaaa	atataaatta	aaacatatag	tatgggcaag	900
cagggagcta	gaacgattcg	cagttaatcc	tggcctgtta	gaaacatcag	aaggctgtag	960
acaaatactg	ggacagctac	aaccatccct	tcagacagga	tcagaagaac	ttagatcatt	1020
atataataca	gtagcaacc	tctatttgtt	gcatcaaaag	atagagataa	aagacaccaa	1080
ggaagcttta	gacaagatag	aggaagagca	aaacaaaagt	aagaccaccg	cacagcaagc	1140
ggccgctgat	cttcagacct	ggaggaggag	atbtgagggg	caattggaga	agtgaattat	1200
ataaataataa	agttagtaaaa	attgaaccat	taggagttagc	accacacaaag	gcaaagagaa	1260
gagtggtgca	gagagaaaaa	agagcagtg	gaataggagc	tttgttcoct	gggttcttgg	1320
gagcagcagg	aagcactatg	ggcgcagcgt	caatgacgct	gacggtacag	gccagacaat	1380
tattgtctgg	tatagtgagc	cagcagaaca	atttgctgag	ggctattgag	gcgcaacagc	1440
atctgttgca	actcacagtc	tggggcatca	agcagctcca	ggcaagaatc	ctggctgtgg	1500
aaagatacct	aaaggatcaa	cagctcctgg	ggatttgggg	ttgctctgga	aaactcattt	1560
gcaocactgc	tgtgccttgg	aatgctagtt	ggagttaataa	atctctggaa	cagatttggg	1620
atcacacgac	ctggatggag	tgggacagag	aaattaacaa	ttacacaagc	ttaatacact	1680
ccttaattga	agaatcgcaa	aaccagcaag	aaaagaatga	acaagaatta	ttggaattag	1740
ataaatgggc	aagtttgtgg	aattggttta	acataacaaa	ttggtgtggy	tatataaaat	1800
tattcataat	gatagtagga	ggcttggtag	gtttaagaat	agtttttgct	gtactttcta	1860
tagtgaatag	agttaggcag	ggatattcac	cattatcgtt	tcagaccac	ctcccaaccc	1920
cgaggggacc	cgacaggccc	gaaggaatag	aagaagaag	tggagagaga	gacagagaca	1980
gatccattcg	attagtgaac	ggatctcgac	ggtatcgatt	ttaaaagaaa	aggggggatt	2040
gggggggtaca	gtgcagggga	aagaatagta	gacataatag	caacagacat	acaaactaaa	2100
gaactacaaa	aacaaattac	aaaaattcaa	aattttcggg	tttattacag	ggacagcaga	2160
gatccagttt	ggaatttcca	gtttaccact	cactatcagt	gatagagaaa	agtgaaagtc	2220
gagtttacca	ctccctatca	gtgatagaga	aaagtgaag	tcgagttaac	cactccctat	2280
cagtgataga	gaaaagtga	agtcgagttt	accactcct	atcagtgata	gagaaaagtg	2340
aaagtcgagt	ttaccactcc	ctatcagtga	tagagaaaag	tgaaagtoga	gtttaccact	2400
ccctatcagt	gatagagaaa	agtgaaagtc	gagtttacca	ctccctatca	gtgatagaga	2460
aaagtgaag	tcgagctcgg	taccggggtc	gagtaggogt	gtacggtggy	aggcctatat	2520
aagcagagct	cgttttagtga	accgtcagag	cgcttgagga	cgccatccac	gctgttttga	2580
cctccataga	agacaccggg	accgatccag	cctccggcgc	cccgaattcg	aattcattga	2640
gaggtcgcct	ctggaaaagg	ccagogttgt	ctccaaaactt	tttttcagct	ggaccagacc	2700
aattttgagg	aaaggataca	gacagcgcct	ggaattgtca	gacatatacc	aaatcccttc	2760
tgttgattct	gctgacaatc	tatctgaaaa	attggaaaga	gaatgggata	gagagctggc	2820

ttcaagaaa	aatcctaaac	tcattaatgc	ccttogggca	tgtttttct	ggagatttat	2880
gttctatgga	atctttttat	atttagggga	agtcacccaaa	gcagtacagc	ctctcttact	2940
gggaagaatc	atagcttctc	atgaccogga	taacaaggag	gaacgctcta	togcgattta	3000
tctaggcata	ggcttatgcc	ttctotttat	tgtgaggaca	ctgctcctac	accagccat	3060
ttttggcctt	catcacattg	gaatgcagat	gagaatagct	atgtttagtt	tgatttataa	3120
gaagaottta	aagctgtcaa	gccgtgttct	agataaaaata	agtattggac	aacttgttag	3180
tctcctttcc	aacaacctga	acaaatttga	tgaaggactt	gcattggcac	atctctgtgt	3240
gatcctcct	ttcaagtg	cactcctcat	ggggctaate	tgggagttgt	tacaggogtc	3300
tgccttctgt	ggacttgggt	tcctgatagt	ccttgccctt	tttcaggctg	ggctaggagg	3360
aatgatgatg	aagtacagag	atcagagagc	tgggaagatc	agtgaaagac	ttgtgattac	3420
ctcagaaatg	attgaaaata	tccaatctgt	taaggcatac	tgcctgggaag	aagcaatgga	3480
aaaaatgatt	gaaaacttaa	gacaaacaga	actgaaactg	actcgggaagg	cagcctatgt	3540
gagatacttc	aatagctcag	cettctctct	ctcagggttc	tttgtgggtg	ttttatctgt	3600
gcttccctat	gcactaatca	aaggaatcat	cctccggaaa	atattcacca	ccatctcatt	3660
ctgcattggt	ctgcgcattg	eggctactcg	gcaatttccc	tgggctgtac	aaacatggta	3720
tgactctctt	ggagcaataa	acaaaataca	ggatttctta	caaaagcaag	aatataagac	3780
attggaatat	aacttaacga	ctacagaagt	agtgatggag	aatgtaacag	ccttctggga	3840
ggagggattt	ggggaattat	ttgagaaagc	aaaacaaaac	aataacaata	gaaaaacttc	3900
taatgggtgat	gacagcctct	tcttcagtaa	tttctcactt	cctggctactc	ctgtcctgaa	3960
agatattaat	ttcaagatag	aaagaggaca	gttgttggcg	gttgcctggat	ccactggagc	4020
aggcaagact	tcacttctaa	tgatgattat	gggagaactg	gagccttcag	agggtaaaat	4080
taagcacagt	ggaagaattt	cattctgttc	tcagtttccc	tggattatgc	ctggcaccat	4140
taaagaaaat	atcatctttg	gtgtttccta	tgatgaatat	agatacagaa	gcgtcatcaa	4200
agcatgccaa	ctagaagagg	acatctccaa	gtttgcagag	aaagacaata	tagttcttgg	4260
agaaggtgga	ctcacactga	gtggagggtca	acgagcaaga	atctcttttag	caagagcagt	4320
atacaaaagat	gtcgatttbt	atctattaga	ctctcctttt	ggatacctag	atgttttaac	4380
agaaaaagaa	atatttga	gctgtgtctg	taactgatg	gctaaca	ctaggatttt	4440
ggtcacttct	aaaatggaac	atttaagaa	agctgacaaa	atattaattt	tgaatgaagg	4500
tagcagctat	ttttatggga	cattttcaga	actccaaaat	ctacagccag	acttttagctc	4560
aaaactcatg	ggatgtgatt	ctttcgacca	atctagtgca	gaagaagaa	attcaatcct	4620
aactgagacc	ttacaccgtt	tctcattaga	aggagatgct	cctgtctcct	ggacagaaac	4680
aaaaaaacaa	tcttttaaac	agactggaga	gtttggggaa	aaaaggaaga	attctattct	4740
caatccaato	aactctatac	gaaaattttc	cattgtgcaa	aagactccct	tacaaatgaa	4800
tggcatogaa	gaggattctg	atgagccttt	agagagaagg	ctgtccttag	taccagattc	4860
tgagcagggg	gagcgatac	tgcctcgcat	cagcgtgatc	agcactggcc	ccacgcttca	4920
ggcacgaagg	aggcagtctg	tectgaacct	gatgacacac	tcagttaacc	aaggtcagaa	4980
cattcacoga	aagacaacag	catccacacg	aaaagtgtca	ctggccctc	aggcaactt	5040
gactgaaactg	gatataat	caagaagggt	atctcaagaa	actggcttgg	aaataagtga	5100
agaaattaac	gaagaagact	taaaggagtg	cctttttgat	gatatggaga	gcataccage	5160
agtgactaca	tggaacacat	acottcgata	tattactgtc	cacaagagct	taatttttgt	5220
gctaatttgg	tgcttagtaa	ttttctggc	agaggtggct	gcttctttgg	ttgtgctgtg	5280
gctccttggg	aacactcctc	ttcaagacaa	agggaaatag	actcatagta	gaaataacag	5340
ctatgcagtg	attatcacca	gcaccagttc	gtattatgtg	ttttacattt	acgtgggagt	5400
agccgacact	ttgcttgeta	tgggatttct	cagaggtcta	ccactygtgc	atactcta	5460
caagtgctog	aaaattttac	accacaaaat	gttacattct	gttcttcaag	cacctatgtc	5520
aacctcaac	acgttgaag	caggtgggat	tcttaataga	ttctccaaag	atatagcaat	5580
tttggatgac	cttctgctc	ttaccatatt	tgacttcate	cagttgttat	taattgtgat	5640
tggagctata	gcagttgtcg	cagttttaca	accctacatc	tttgttgcaa	cagtgccagt	5700
gatagtggct	tttattatgt	tgagagcata	tttctccaa	acctcacagc	aactcaaca	5760
actggaatct	gaaggcagga	gtcctaattt	cactcatctt	gttacaagct	taaaaggact	5820
atggacactt	cgtgccttcg	gacggcagcc	ttactttgaa	actctgttcc	acaaagctct	5880
gaatttacat	actgccaaact	ggttcttgta	cctgtcaaca	ctgcgctgg	tccaaatgag	5940
aatagaaatg	atttttgtca	tcttcttcat	tgctgttacc	ttcatttcca	ttttaacaac	6000
aggagaagga	gaaggaagag	ttggtattat	cctgacttta	gccatgaata	tcatgagtac	6060
attgcagtg	gctgtaaact	ccagcataga	tgtggatagc	ttgatgcat	ctgtgagccg	6120
agtctttaag	ttcattgaca	tgccaacaga	aggtaaacct	accaagtcaa	ccaaaccata	6180
caagaatggc	caactctoga	aagttatgat	tattgagaat	tcacacgtga	agaaagatga	6240
catctggccc	tcagggggcc	aaatgactgt	caagatctc	acagcaaaat	acacagaagg	6300
tggaaatgcc	atattagaga	acatttccct	ctcaataagt	cctggccaga	gggtgggctt	6360
cttgggaaga	actggatcag	ggaagagtac	tttgttatca	gcttttttga	gactactgaa	6420
cactgaagga	gaaatccaga	tcgatggtgt	gtcttgggat	tcaataactt	tgcaacagtg	6480

gaggaaagcc	tttggagtga	taccacagaa	agtatrratt	ttttctggaa	catttagaaa	6540
aaacttggat	ccctatgaac	agtggagtga	tcaagaaata	tggaaagtgtg	cagatgaggt	6600
tgggtcaga	tctgtgatag	aacagtttcc	tgggaagcct	gactttgtcc	ttgtggatgg	6660
gggtgtgtc	ctaagccatg	gccacaagca	gttgatgtgc	ttggctagat	ctgttctcag	6720
taaggcgaag	atcttgctgc	ttgatgaacc	cagtgetcat	ttggatccag	taacatacca	6780
aataattaga	agaactctaa	aacaagcatt	tgctgattgc	acagtaattc	tctgtgaaca	6840
caggatagaa	gcaatgctgg	aatgccaaaca	atTTTTgtgc	atagaagaga	acaaagtgcg	6900
gcagtacgat	tccatccaga	aactgctgaa	cgagaggagc	ctcttcoggc	aagccatcag	6960
ccctccgac	aggtggaagc	tctttcccca	ccggaactca	agcaagtgca	agtctaagcc	7020
ccagattgct	gctctgaaag	aggagacaga	agaagaggtg	caagatacaa	ggctttagct	7080
cgaggagatt	acaagatga	cgacgataag	cgcaacggcc	atcatcacca	tcacccatcac	7140
caoccatcact	aacgagtttc	cctctagcgg	gatcaattcc	gccccccccc	tctccctccc	7200
ccccoctaac	gttactggcc	gaagccgctt	ggaataaggc	cggtgtgctg	ttgtctatat	7260
gttattttcc	accatattgc	ogtcttttgg	caatgtgagg	gcccggaaac	ctggccctgt	7320
cttcttgacg	agcattocta	ggggtctttc	ccctctcgcc	aaaggaatgc	aaggtctgtt	7380
gaatgtogtg	aaggaagcag	ttcctctgga	agcttcttga	agacaaacaa	cgctctgtagc	7440
gaccctttgc	aggcagcgga	accccccaac	tggogacagg	tgccctctgcg	gccaaaagcc	7500
acgtgtataa	gatacacctg	caaaggcggc	acaaccccag	tgccacgttg	tgagttggat	7560
agttgtggaa	agagtcaaat	ggctctctcc	aagcgtatcc	aacaaggggc	tgaaggatgc	7620
ccagaaggta	ccccattgta	tgggatctga	tctggggcct	cggtgcacat	gctttacatg	7680
tgttttagtgc	aggttaaaaa	aacgtctagg	ccccccgaac	cacggggacg	tggttttccct	7740
ttgaaaaaca	cgatgataat	ggccacaacc	atggtgactg	aatacaaac	aactgttccg	7800
ctggcacttc	gtgatgatgt	gttcgcaccc	tggctgctgc	atctgtgac	tttccgtgac	7860
taccctgcaa	ccogtcacac	tgtggacceca	gaccgccaca	ttgaacgtgt	gactgaaactg	7920
caggagctgt	tcctgaccocg	tgtgggcctg	gacattggca	aagtgtgggt	ggcagatgat	7980
ggtgctgctg	tggcagtgty	gaccacccct	gaatctgtty	aagctggtgc	agtgtttgct	8040
gagattggcc	caogcatggc	agaactgtct	ggcagccgcc	tggcagcaca	acagcagatg	8100
gaaggtctgc	tggcaccaca	ccgccccaaa	gaacctgctt	ggttccctgce	aactgtgggt	8160
gtggcacttc	accaccaggg	taagggcctg	ggctctgcag	tggtgctgcc	tggtgtggaa	8220
gcagctgaac	gtgcaggtgt	gcctgcttcc	ctggagacct	cagctccacg	caacctgccc	8280
ttctatgaac	gootgggctt	cactgtgact	gctgatgtgg	aagtgccaga	aggcccacgc	8340
acttgggtgca	tgactcgcaa	accaggtgct	taagtcgacg	tcaccgcccga	cgctgaggtg	8400
ccccgaaggac	cgcgcacctg	gtgcatgacc	cgcaagcccg	gtgcccgaag	cctcgacaat	8460
caacctctgg	attacaaaaat	ttgtgaaaga	ttgactggta	ttcttaacta	tggtgtctcct	8520
tttaogctat	ctggatacgc	tgctttaatg	cttttgatc	atgctattgc	ttcccgatg	8580
gotttccattt	tctcctcctt	gtataaatcc	tggttgctgt	ctctttatga	ggagttgtgg	8640
cccgttgfca	ggcaacgtgg	cggtgtgtgc	actgtgtttg	ctgacgcaac	ccccactggg	8700
tggggcattg	ccaccacctg	tcagctcctt	tccgggactt	tcgctttccc	cctccctatt	8760
gccacggcgg	aaactaatcgc	cgccctgctt	gcccgtgct	ggacaggggc	tcggctggtg	8820
ggcactgaca	attccgttgg	gttgtcgggg	aagctgacgt	cctttccatg	gctgctcgcc	8880
tgtgttgcca	cctggattct	gcgcgggacg	tccttctgct	acgtcccttc	ggccctcaat	8940
ccagcggacc	ttccttcccg	cgccctgctg	ccgctctgc	ggcctcttcc	ggctcttccg	9000
cttcgcccctc	agacgagctg	gatctccctt	tgggcccctt	ccccgcttgg	gtacctttaa	9060
gaccaatgac	ttacaaggca	gctgtagatc	ttagccactt	tttaaaagaa	aaggggggac	9120
tggaaagggct	aattcactcc	caacgaagac	aagatctgct	ttttgcttgt	actgggtctc	9180
tctggttaga	ccagatctga	gcctgggagc	tctctggcta	actagggaac	ccactgctta	9240
agcctcaata	aagcttgcct	tgagtgtctc	aagtagtgtg	tgcccgtctg	ttgtgtgact	9300
ctggtaacta	gagatccctc	agaccctttt	agt.cagtgtg	gaaaatctct	agcagtagta	9360
gttccatgtca	tcttattate	cagt				9384

&lt;210&gt; 49

&lt;211&gt; 5015

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;223&gt; 对人工序列的描述: 合成构建体

<400> 49								
agegccc	aa	tacgcaa	acc	gcctctccc	gcgcgttggc	cgattcatta	atgcagctgg	60
cacgacaggt	ttccc	gactg	gaaagcgggc	agtgagcgca	acgcaattaa	tgtgagttag	120	
otcaactcatt	aggcacc	ccca	ggctttacac	tttatgcttc	cggtcgtat	gttgtgtgga	180	
attgtgagcg	gataaca	aatt	tcacacagga	aacagctatg	accatgatta	cgccaagctt	240	
ggtaocgagc	tcggatcc	ac	tagtaaggat	ccaccatggg	caatgcctcc	aatgactccc	300	
agtctgagga	ctgagag	acg	cgacagtggc	ttccccagg	cgaaagccca	gccatcagct	360	
ccgtcatggt	ctcggc	cg	gtgctgggga	acctcatagc	actggcgctg	ctggcgcgcc	420	
gctggcgggg	ggacgt	gg	tgcagcgccg	gcccagggag	ctocctctcc	ttgttccacg	480	
tgctgggtgac	cgagct	gg	ttcacogacc	tgctcgggac	ctgcctcatc	agcccagtg	540	
tactggcttc	gtacgc	cg	aaccagacc	tggtggcact	ggcgcccgag	agccgcgct	600	
gcacctactt	cgcttt	cg	atgacctct	tcagcctggc	cacgatgctc	atgctcttg	660	
ccatggccct	ggagcg	ct	ctctcgatcg	ggcacccta	cttctaccag	cgccgctct	720	
cgcgctccgg	gggect	gg	gtgctgcctg	tcctctatgc	agtctccctg	ctcttctgct	780	
cgctgcccgt	gctgg	act	gggcagtacg	tcagtaactg	ccccgggacc	tggtgcttca	840	
tccggcacgg	gcccag	cc	tacctgcagc	tgtaagccac	cctgctgctg	cttctcattg	900	
tctcgggtgct	cgctg	ca	ttcagtgcca	ttctcaacct	catccgcatg	caccgcgaa	960	
gcgggagaag	ccgct	cg	ccttccctgg	gcagtggccg	ggcgggcccc	ggggccccca	1020	
ggagaggggga	aagggt	gt	atggcggagg	agacggacca	cctcattctc	ctggctatca	1080	
tgaccatcac	cttcg	cg	tgctccttgc	ctttcaogat	ttttgcatat	atgaatgaaa	1140	
cctcttcccc	aaagg	aaaa	tgggacctcc	aagctcttag	gtttttatca	attaattcaa	1200	
taattgacct	ttgggt	ct	gtccatccta	ggctcctgt	tctgagacta	atgcgttcag	1260	
tccctctgtg	tcggat	tt	ttaagaacac	aagatgcaac	acaaacttcc	tgttctaac	1320	
agtccagatgc	cagtaa	ac	gctgaccttg	aaaacctgta	ttttcagggc	gctcgaggag	1380	
attacaaaaa	gcoga	att	gcagatatcc	atcacactgg	cgcccgctcg	agcatgcac	1440	
tagagggccc	aattc	gc	atagtgagtc	gtattacaat	tcactggccg	togttttaca	1500	
acgtcgtgac	tggg	aaa	ctggcgttac	ccaacttaat	cgccctgcag	cacatcccc	1560	
tttcgcccagc	tggcg	ta	gcgaagaggc	ccgcaccgat	cgcccttccc	aacagttgcg	1620	
cagctcgaat	ggcga	at	cgccctgt	agcggcgcat	taagcgggc	gggtgtggtg	1680	
gttacgcgca	gcgtg	ac	tcacttgc	agcgccctag	cgcccgctcc	tttcgcttcc	1740	
ttcccttcc	ttctc	gc	gttcgcggc	tttccccg	aagctctaaa	tcgggggctc	1800	
cctttaggg	tcog	at	tgctttacgg	caactcgacc	ccaaaaact	tgattaggg	1860	
gatggttcac	gtagt	gg	atcgccctga	tagacggttt	ttcgccctt	gacgttggag	1920	
tccacgttct	ttaat	ag	actcttgttc	caaactggaa	caacactcaa	ccctatctcg	1980	
gtctattctt	ttgatt	ta	agggattttg	ccgatttcgg	cctattgggt	aaaaaatgag	2040	
ctgatttaac	aaaa	at	cgcaatttt	aacaaaatc	agggcgcaag	ggctctgaaa	2100	
ggaagcggaa	cacgt	ag	gccagtccgc	agaaaagg	ctgaccccg	atgaatgtca	2160	
gctactgggc	tatct	gg	agggaaaacg	caagcgcaaa	gagaaaagcag	gtagcttgca	2220	
gtgggcttac	atggc	gat	ctagactggg	cggttttatg	gacagcaagc	gaaccggaat	2280	
tgccagctgg	ggcgc	ct	ggtaaaggtg	ggaagccctg	caaagttaac	tggtatggct	2340	
tcttcgccc	aagg	at	tgggcaggg	gatcaagatc	tgatcaagag	acaggatgag	2400	
gategtttcg	ctgatt	ga	caagatggat	tgcaogcagg	ttctccggcc	gcttgggtgg	2460	
agaggctatt	cggt	at	tgggcacaac	agacaatcgg	ctgctctgat	gcccgcgtgt	2520	
tccggctgtc	agcgc	ag	ogcccggttc	tttttgtcaa	gacogacctg	tcoggtgccc	2580	
tgaatgaact	gcagg	ac	gcagcgggc	tatcgtggct	ggccaagcag	ggcgttccct	2640	
gcgagctgt	gctcg	ag	gtcactgaag	cggaagggga	ctgggtgcta	ttggggcgaag	2700	
tgccggggca	ggat	ct	tcctccacc	ttgctcctg	cgagaaagta	tcctcatgg	2760	
ctgatgcaat	gggg	cg	catacgcttg	atccggctac	ctgcccattc	gaccaccaag	2820	
cgaaacatcg	catcg	ag	gcacgtactc	ggatggaagc	cggtcttgtc	gatcaggatg	2880	
atctggacga	agag	at	gggctcgcgc	cagccgaact	gttcgcccagg	ctcaaggcgc	2940	
gcatgcccga	cgcg	ag	ctcgtcgtga	cccatggcga	tgccctgctt	ccgaatatca	3000	
tggtggaaaa	tggcc	g	tctggattca	tgcactgttg	ccggctgggt	gtggcggacc	3060	
gctatcagga	catag	cg	gtaccocgtg	atattgctga	agagcttggc	ggcgaattgg	3120	
ctgaccgctt	cctcg	g	tacggtatcg	ccgctcccga	ttcgcgccgc	atcgacttct	3180	
atcgccctct	tgac	g	ttctgaattg	aaaaaggaag	agtatgagta	ttcaacattt	3240	
ccgtgtcggc	cttatt	cc	tttttgcggc	atthtgcctt	cctgtttttg	ctcaccaga	3300	
aacgctggtg	aaagt	aaa	atgctgaaga	tcagttgggt	gcacgagtg	gttacatcga	3360	
actggatctc	aacag	cg	agatccttga	gagttttcgc	cccgaagaac	gttttccaat	3420	
gatgagcact	tttaa	agt	tgctatgtgg	ogggatatta	tcccgtattg	acgcccggca	3480	
agagcaactc	ggtc	gc	tacactattc	tcagaatgac	ttgggtgagt	actcaccagt	3540	
cacagaaaag	catct	ta	atggcatgac	agtaagagaa	ttatgcagtg	ctgcataac	3600	

catgagtgat	aacactgogg	ccaacttact	tctgacaacg	atcggaggac	cgaaggagct	3660
aacgcgtttt	ttgcacaaca	tgggggatca	tgtaactcgc	cttgatcgtt	gggaaccgga	3720
gctgaatgaa	gccataccaa	acgacgagcg	tgacaccacg	atgcctgtag	caatggcaac	3780
aacgttgccg	aaactattaa	ctggcgaact	acttactcta	gcttcccggc	aacaattaat	3840
agactggatg	gagggggata	aagttgcagg	accacttctg	cgctcggccc	ttccggctgg	3900
ctggtttatt	gctgataaat	ctggagccgg	tgagcgtggg	tctcgcggta	tcattgcagc	3960
actggggcca	gatggtaagc	cctcccgtat	cgtagttatc	tacacgacgg	ggagtcaggc	4020
aactatggat	gaaacgaaata	gacagatcgc	tgagataggt	gcctcactga	ttaagcattg	4080
gtaactgtca	gaccaagttt	actcatatat	actttagatt	gatttaaaac	ttcattttta	4140
atltaaaagg	atctaggtga	agatcctttt	tgataatctc	atgaccaaaa	tcctttaacg	4200
tgagttttcg	ttccactgag	cgtagaaaag	atcaaaggat	cttcttgaga	cttcttgaga	4260
tccttttttt	ctgcgcgtaa	tctgctgctt	gcaaacaaaa	aaaccacccg	taccagcggc	4320
ggtttgtttg	cgggatcaag	agctaccaac	tctttttccg	aaggtaactg	gcttcagcag	4380
agcgcagata	ccaaatactg	ttcttctagt	gtagccgtag	ttaggccacc	acttcaagaa	4440
ctctgtagca	ccgcotacat	acctcgcctc	gctaatcctg	ttaccagtg	ctgctgccag	4500
tgggogataag	togtgotta	ccgggttggg	ctcaagacga	tagttaccgg	ataaggcgca	4560
gcggtcgggg	tgaaocgggg	gttcgtgcac	acagcccagc	ttggagcgaa	cgacctacac	4620
cgaactgaga	tacctacagc	gtgagctatg	agaaagcggc	acgcttcccg	aaggagagaaa	4680
ggcggacagg	tatccggtaa	gcccaggggt	cggaacagga	gagcgcacga	gggagcttcc	4740
agggggaaac	gcctgggtatc	tttatagttc	tgtegggttt	cgccacctct	gacttgagcg	4800
togatttttg	tgatgctcgt	cagggggggc	gagcctatgg	aaaaacggca	gcaacggggc	4860
cttttttaagg	ttcotggcot	tttgctggcc	ttttgctcac	atgttctttc	ctgcggtatc	4920
ccctgattct	gtggataaac	gtattaccgc	ctttgagtga	gctgataccg	ctcgcggcag	4980
ccgaacgacc	gagcgcagcg	agtcagtgag	cgagg			5015

&lt;210&gt; 50

&lt;211&gt; 6040

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;223&gt; 对人工序列的描述: 合成构建体

&lt;400&gt; 50

gggctaatt	cactcccaaa	gaagacaaga	tatccttgat	ctgtggatct	accacacaca	60
aggtaacttc	cctgattagc	agaactacac	accagggcca	ggggtcagat	atccactgac	120
ctttggatgg	tgctacaagc	tagtaccagt	tgagccagat	aaggtagaag	aggccaataa	180
aggagagaac	accagcttgt	tacaccctgt	gagcctgcat	gggatggatg	accggagag	240
agaagtgtta	gagtggaggt	ttgacagccg	cctagcattt	catcacgtgg	cccagagcct	300
gcatccggag	tacttcaaga	actgctgata	tcgagcttgc	tacaagggac	tttccgctgg	360
ggactttcca	gggaggcgtg	gcctgggccc	gactggggag	tggcggagccc	tcagatcctg	420
catataagca	gctgcttttt	gcctgtactg	ggtctctctg	gttagaccag	atctgagcct	480
gggagctctc	tggctaacta	gggaaccac	tgcttaagcc	tcaataaagc	ttgccttgag	540
tgcttcaagt	agtgtgtgcc	cgctctgtgt	gtgactctgg	taactagaga	tcctcagac	600
ccttttagtc	agtgtggaaa	atctctagca	gtggcggccc	aacagggact	tgaaagcgaa	660
agggaaacca	gaggagctct	ctcgacgcag	gactcggctt	gctgaagcgc	gcacggcaag	720
aggcgagggg	cggcgactgg	tgagtacgcc	aaaaattttg	actagcggag	gctagaagga	780
gagagatggg	tgcgagagcg	tcagtattaa	gcgggggaga	attagatcgc	gatgggaaaa	840
aattcgggta	aggccagggg	gaaagaaaaa	atataaatta	aaacatatag	tatgggcaag	900
cagggagcta	gaacgattcg	cagttaatcc	tggcctgcta	gaaacatcag	aaggctgtag	960
acaaatactg	ggacagctac	aacctccct	tcagacagga	tcagaagaac	ttagatcatt	1020
atataataca	gtagcaaccc	tctattgtgt	gcatcaaagg	atagagataa	aagacaccaa	1080
ggaagcttta	gacaagatag	aggaaagcca	aaacaaaagt	aagaccaccg	cacagcaagc	1140
ggccgctgat	cttcagacct	ggagagagg	atatgagggg	caattggaga	agtgaattat	1200
ataaataaaa	agttagtaaaa	attgaacct	taggagtagc	aoccaccaag	gcaagagaaa	1260
gagtgtgca	gagagaaaaa	agagcagtg	gaataggagc	tttgttctct	gggttcttgg	1320
gagcagcagg	aagcactatg	ggcgcagct	caatgacgct	gacggtagac	gccagacaat	1380
tattgtctgg	tatagtgcag	cagcagaaca	atttgcctgag	ggctattgag	ggcgaacagc	1440

atctgttgca	actcacagtc	tggggcatca	agcagctcca	ggcaagaatc	ctggctgtgg	1500
aaagatacct	aaaggatcaa	cagctcctgg	ggatttgggg	ttgctctgga	aaactcattt	1560
gcaccaotgc	tgtgccttgg	aatgctagtt	ggagtaataa	atctctggaa	cagatttggga	1620
atcacacgac	ctggatggag	tgggacagag	aaattaacaa	ttacacaagc	ttaatacact	1680
ccttaattga	agaatcgcaa	aaccagcaag	aaaagaatga	acaagaatta	ttggaattag	1740
ataaatgggc	aagtttgtgg	aattggttta	acataacaaa	ttggctgtgg	tatataaaat	1800
tattcataat	gatagtagga	ggcttggtag	gtttaagaat	agtttttgcct	gtactttcta	1860
tagtgaatag	agttaggcag	ggatattcac	cattatcgtt	tcagacccac	ctcccaaccc	1920
cgaggggacc	cgacaggccc	gaaggaatag	aagaagaagg	tggagagaga	gacagagaca	1980
gatccattcg	attagtgaac	ggatctcgac	ggtatcgatt	ttaaagaaa	aggggggatt	2040
gggggggtaca	gtgcagggga	aagaatagta	gacataatag	caacagacat	acaaactaaa	2100
gaactacaaa	aaacaaattac	aaaaattcaa	aattttcggg	tttattacag	ggacagcaga	2160
gatccagttt	ggaatttcga	gtttaccact	ccctatcagt	gatagagaaa	agtgaaagtc	2220
yagtttacca	ctccctatca	gtgatagaga	aaagtgaag	tcaggtttac	cactccctat	2280
cagtgataga	gaaaagtga	agtcgagttt	accactccct	atcagtgata	gagaaaagtg	2340
aaagtcgagt	ttaccactcc	ctatcagtga	tagagaaaag	tgaaagtoga	gtttaccact	2400
ccctatcagt	gatagagaaa	agtgaagtc	gagtttacca	ctccctatca	gtgatagaga	2460
aaagtgaag	tcgagctcgg	taccgggtc	gagtaggcgt	gtacgggtggg	aggcctatat	2520
aagcagagct	cgtttagtga	accgtcagat	cgctcggaga	cgccatccac	gctgttttga	2580
cctccataga	agacaccggg	accgatccag	cctccgcgag	ccogaattcg	gatccaccat	2640
ggcfaatgcc	tccaatgact	cccagtctga	ggactgogag	acggcagcagt	ggcttcccc	2700
aggcgaagc	ccagccatca	gctccgtcat	gttctcggcc	gggtgctgg	ggaacctcat	2760
agcactggcg	ctgctggcgc	gocgctggcg	gggggacgtg	gggtgcagcg	cggccgcag	2820
gagctccctc	tccttgttcc	acgtgctggg	gaccgagctg	gtgttcaccg	acctgctcgg	2880
gacctgcctc	atcagcccag	tggtactggc	ttcgtacgcg	cggaaccaga	ccctgggtggc	2940
actggcgccc	gagagccgcg	cgtgcaocta	cttcgcttcc	gccatgacct	tcttcagcct	3000
ggccaacgatg	ctcatgctct	tcgcatggc	cctggagcgc	tacctctcga	tcgggcaccc	3060
ctacttctac	cagocgcgog	tctcgcgctc	cgggggcctg	gocgtgctgc	ctgtcatcta	3120
tgcagtctcc	ctgctcttct	gctcgcgctc	gctgctggac	tatgggcagt	acgtccagta	3180
ctgccccggg	acctggtgct	tcatccggca	cggggcgacc	gcttacctgc	agctgtacgc	3240
cacctgctg	ctgcttctca	ttgtctcggg	gctcgcctgc	aacttcagtg	tcattctcaa	3300
cctcatccgc	atgcaccgcc	gaagccggag	aagccctctg	ggaccttccc	tgggcagtgg	3360
ccggggcgctc	ccggggcgctc	gcaggagagg	ggaaaagggtg	tccatggcgg	aggagacgga	3420
ccacctcatt	ctcctggcta	tcatgacct	caccttcgcc	gtctgctcct	tgcccttcac	3480
gatttttgca	tatatgaatg	aaacctcttc	ccgaaaggaa	aaatgggacc	tccaagctct	3540
taggttttta	tcaattaatt	caataattga	cccttgggtc	tttgccatcc	ttaggcctcc	3600
tgttctgaga	ctaattgcgtt	cagtcctctg	ttgtcggatt	tcattaagaa	cacaagatgc	3660
aacacaaact	tccgtttcta	cacagtcaga	tgccagtaaa	caggctgacc	ttgaaaacct	3720
gtattttcag	ggcgtctcag	gagattacaa	agatgaogac	gataagcgca	acggccatca	3780
tcaccatcac	oatcaaaccc	atcactaacg	agtttccctc	tagcgggatc	aattccgccc	3840
ccccctctc	cctccccccc	cctaocgtta	ctggccgaag	ccgcttggaa	taagcccggt	3900
gtgogtttgt	ctatatgtta	ttttccacca	tattgocgtc	ttttggcaat	gtgagggccc	3960
ggaaacctgg	ccctgtcttc	ttgacgagca	ttcctagggg	tctttcccct	ctcgccaaag	4020
gaatgcaagg	tctgttgaat	gtcgtgaagg	aagcagttcc	tctggaagct	tcttgaagac	4080
aaacaacgto	tgtagcgacc	ctttgcaggc	agcggaaacc	cccacctggc	gacaggtgcc	4140
tctgcgcca	aaagccacgt	gtataagata	cacctgcaaa	ggcggcacia	cccagtgcc	4200
acgttgtgag	ttggatagtt	gtggaaagag	tcaaatggct	ctcctcaage	gtattcaaca	4260
aggggctgaa	ggatgcccag	aaggtacccc	attgtatggg	atctgatctg	gggcctcggt	4320
gcacatgctt	tacatgtgtt	tagtcgaggt	taaaaaaacg	tctaggcccc	cogaaccacg	4380
gggacgtggt	tttcccttga	aaaacacgat	gataatggcc	acaacccatg	tgactgaata	4440
caaaccaact	gttcgctgg	caactcgtga	tgatgttcca	cgtgcagttc	gcaacctggc	4500
tgctgcattt	gctgactacc	ctgcaaccgg	tcacactgtg	gacccagacc	gccacattga	4560
acgtgtgact	gaactgcagg	agctgttccct	gaccogtgtg	ggcctggaca	ttggcaaaag	4620
gtgggtggca	gatgatggtg	ctgctgtggc	agtgtggacc	accctggaat	ctgttgaagc	4680
tggtgcagtg	tttgcctgaga	ttggcccacg	catggcagaa	ctgtctggca	gocgcctggc	4740
agcaacaacg	cagatggaa	gtctgctggc	accacacocg	ccaaaagaac	ctgcttgggt	4800
cctggcaact	gtgggtgaga	gccctgacca	ccagggtaag	ggcctgggct	ctgcagtggt	4860
gctgcctggt	gtggaagcag	ctgaaactgc	aggtgtgctt	gctttcctgg	agacctcage	4920
tccacgcaac	ctgcctttct	atgaaactgc	ggccttcaact	gtgactgctg	atgtggaagt	4980
gccagaaggc	ccacgcactt	ggtgcatgac	tcgcaaacca	ggtgcttaag	tcgacgtcac	5040
cgccgacgctc	gaggtgcccc	aaggaccgog	caactggtgc	atgacccgca	agcccgggtgc	5100

ctgacgcctc	gacaatcaac	ctctggatta	caaaatthgt	gaaagattga	ctggattct	5160
taactatgtt	gctcctttta	cgctatgtgg	atacgtctgt	ttaatgcctt	tgtatcatgc	5220
tattgcttcc	cgatggcctt	tcattttctc	ctccttgtat	aaatcctggg	tgctgtctct	5280
ttatgaggag	ttgtggcccg	ttgtcaggca	acgtggcgtg	gtgtgcactg	tgtttgctga	5340
cgcaaccccc	actggttggg	gcattgccc	cacctgtcag	ctcctttccg	ggactttcgc	5400
tttccccctc	cctattgcca	cggcggaaact	catcgcgcc	tgccttgccc	gctgctggac	5460
aggggctcgg	ctgttgggca	ctgacaattc	cgtgggtgtg	tccggggaagc	tgacgtcctt	5520
tccatggctg	ctcgctgtg	ttgccacctg	gattctgcgc	gggacgtcct	tctgctacgt	5580
cccttoggcc	ctcaatccag	cggaaccttc	ttcccgccgc	ctgctgcggg	ctctgcggcc	5640
tcttccgctg	cttcgccttc	gcccctcagac	gagtcggatc	tccctttggg	ccgcccctccc	5700
gcctgggtac	ctttaagacc	aatgacttac	aaggcagctg	tagatcttag	ccacttttta	5760
aaagaaaagg	ggggactgga	agggctaatt	cactcccaac	gaagacaaga	tctgcttttt	5820
gcttgtactg	ggtctctctg	gttagaccag	atctgagcct	gggagctctc	tggctaacta	5880
gggaacccac	tgcttaagcc	tcaataaagc	ttgccttgag	tgcttcaagt	agtgtgtgcc	5940
cgtctgttgt	gtgactctgg	taactagaga	tcctcagac	ccttttagtc	agtgtggaaa	6000
atctctagca	gtagtagttc	atgtcatctt	attattcagt			6040

- <210> 51
- <211> 7647
- <212> DNA
- <213> 人工序列

- <220>
- <223> 对人工序列的描述: 合成构建体

<400> 51						
gggctaatt	cactcccaaa	gaagacaaga	tatccttgat	ctgtggatct	accacacaca	60
aggctacttc	cctgattagc	agaactacac	accagggcca	ggggtcagat	atccactgac	120
ctttggatgg	tgctacaagc	tagtaccagt	tgagccagat	aaggtagaag	aggccaataa	180
aggagagaac	accagcttgt	tacacccctgt	gagcctgcat	gggatggatg	acccggagag	240
agaagtgtta	gagtggaggt	ttgacagccg	cctagcattt	catcacgtgg	cccagagctc	300
gcatccggag	tacttcaaga	actgctgata	tcgagcttgc	tacaagggac	tttccgctgg	360
ggactttcca	gggagggcgtg	gcctggggcgg	gactggggag	tggcgagccc	tcagatcctg	420
catataagca	gctgcttttt	gcctgtactg	ggtctctctg	gttagaccag	atctgagcct	480
gggagctctc	tggctaacta	gggaaaccac	tgcttaagcc	tcaataaagc	ttgccttgag	540
tgcttcaagt	agtgtgtgcc	cgtctgttgt	gtgactctgg	taactagaga	tcctcagac	600
ccttttagtc	agtgtggaaa	atctctagca	gtggcgcccg	aacagggact	tgaagcaga	660
agggaacca	gaggagctct	ctcgcgcag	gactcggcct	gctgaagcgc	gcacggcaag	720
aggcgagggg	cggcgactgg	tgagtacgcc	aaaaatthtg	actagcggag	gctagaagga	780
gagagatggg	tgcgagagcg	tcagatttaa	gcgggggaga	attagatcgc	gatgggaaaa	840
aattcgggta	aggccagggg	gaaagaaaaa	atataaatta	aaacataatg	tatgggcaag	900
caggagagcta	gaacgatccg	cagttaatcc	tggcctgtta	gaaacatcag	aaggctgtag	960
acaaatactg	ggacagctac	aacctccct	tcagacagga	tcagaagaac	ttagatcatt	1020
atataataca	gtagcaaccc	tcatttctgt	gcatcaaagg	atagagataa	aagcaccaaa	1080
ggaagcttta	gacaagatag	aggaagagca	aaacaaaagt	aagaccaccg	cacagcaagc	1140
ggccgctgat	ctcagacct	ggaggaggag	atattgaggga	caattggaga	agtgaattat	1200
ataaatataa	agttagtaaaa	attgaacct	taggagtagc	accaccaag	gcaagagaaa	1260
gagtgggtgca	gagagaaaaa	agagcagtg	gaataggagc	tttgttctct	gggttcttgg	1320
gagcagcagg	aagcactatg	ggcgcagcgt	caatgacgct	gacggtagag	gccagacaat	1380
tattgtctgg	tatagtgagc	cagcagaaca	atttctctgag	ggctattgag	gcgcaacagc	1440
atctgttgca	actcacagtc	tggggcatca	agcagctcca	ggcaagaatc	ctggctgtgg	1500
aaagatacct	aaaggatcaa	cagctcctgg	ggatttgggg	ttgctctgga	aaactcattt	1560
gcaccactgc	tgtgccttgg	aatgctagtt	ggagtaataa	atctctggaa	cagatttggga	1620
ctcacacgac	ctggatggag	tgggacagag	aaattaacaa	ttacacaagc	ttaatacact	1680
ccttaattga	agaattcgcaa	aaccagcaag	aaagaatga	acaagaatta	ttggaattag	1740
ataaatgggc	aagtttctgg	aattggttta	acataacaaa	ttggctgtgg	tataaaaaat	1800
tattcataat	gatagtagga	ggcttggtag	gtttaagaat	agtttttctg	gtactttcta	1860
tagtgaatag	agttaggcag	ggatattcac	cattatcgtt	tcagaccac	ctcccaaccc	1920

cgaggggacc	cgacaggccc	gaaggaatag	aagaagaagg	tggagagaga	gacagagaca	1980
gatccattcg	attagtgaac	ggatctcgac	ggtatcgatt	ttaaaagaaa	aggggggatt	2040
gggggggtaca	gtgcagggga	aagaatagta	gacataatag	caacagacat	acaaactaaa	2100
gaactacaaa	aacaaattac	aaaaattcaa	aattttcggg	tttattacag	ggacagcaga	2160
gatccagttt	ggaattaatt	gcgcgttaca	gggcgcgtgg	ggatacccc	tagagcccca	2220
gctggttctt	tccgcctcag	aagccataga	gcccaccgca	tccccagcat	gcctgctatt	2280
gtcttcccaa	tcttccccct	tgctgtectg	ccccaccca	ccccccagaa	tagaatgaca	2340
cctactcaga	caatgcgatg	caatttctct	atthatttag	gaaaggacag	tgggagtggc	2400
accttccagg	gtcaaggaag	gcacggggga	ggggcaaaoc	acagatggct	ggcaactaga	2460
aggcacagtc	gaggctgac	agcgggttcc	togagatctg	agtcgggact	tgtacagctc	2520
gtccatgccc	agagtgatcc	cggcggcggt	caogaactcc	agcaggacca	tgtgatcgcg	2580
cttctcgttg	gggtctttgc	tcagggcgga	ctgggtgctc	aggtagtgg	tgtcgggcag	2640
cagcacgggg	ccgtcgccga	tgggggtgtt	ctgctggtag	tggtcggcga	gctgcacgct	2700
gcogtctog	atggtgtggc	ggatcttgaa	gtcaccttg	atgccgttct	tctgcttgtc	2760
ggccatgata	tagacgttgt	ggctgttcta	gttgaactcc	agcttgtgcc	ccaggatggt	2820
gcogtctctc	ttgaagtcca	tgcccttcag	ctogatcgcg	ttcaccaggg	tgtcgccctc	2880
gaacttcacc	tggcgcgggg	tcttgtagtt	gccgtcgctc	ttgaagaaga	tgggtcgctc	2940
ctggacgtag	ccttcgggca	tggcggactt	gaagaagtcc	tgctgcttca	tgtggtcggg	3000
gtagcggctg	aagcactgca	cgccgtaggt	caggggtggtc	acgagggtag	gccagggcac	3060
gggcagcttg	ccggatggtc	agatgaactt	cagggtcagc	ttgccgtagg	tggcatcgcc	3120
ctcgcctctg	ccggacagcc	tgaacttctg	gccgtttagc	tgccttcca	gctcgaccag	3180
gatgggcacc	accocggtga	acagctctct	gcccttctct	accatggtag	cgaccggtag	3240
cgctaggatc	catctctatc	actgataggg	agatctctat	cactgatagg	gagactctgc	3300
ttatatagac	ctcccaccgt	acacgcctac	cgccatttgg	cgccaatggg	gccggagtgt	3360
tacgacattt	tggaaagtcc	cgttgatttt	ggttccaaaa	caaactccca	ttgacgtcaa	3420
tgggggtggag	acttggaaat	ccccgtgagt	caaaccgcta	tccacgcccc	ttgatgtact	3480
gcoaaaaccg	catcaccatg	gtaatagcga	tgaactaac	gtagatgtac	tgccaaagtga	3540
gaaagtccca	taaggtcatg	tactgggcat	aatgccaggc	gggccattta	ccgtcattga	3600
cgtcaatagg	gggcgtactt	ggcatatgat	acacttgatg	tactgccaaag	tgggcagttt	3660
accgtaaata	ctccacccat	tgacgtcaat	ggaaagtccc	tattggcggt	actatgggaa	3720
catacgtcat	tattgacgtc	aatgggcggg	ggtcgttggg	eggtcagcca	ggcggggccat	3780
ttaggaattc	aagcttctgt	aggctccggt	gcccgctcagt	gggcagagcg	cacatogccc	3840
acagtcctcc	agaagttygg	gggaggggtc	ggcaattgaa	ccgggtgcta	gagaaggtgg	3900
cgcggggtaa	actgggaaag	tgatgtcgtg	tactggctcc	gccttttcc	cgagggtgg	3960
ggagaaccgt	atataagtgc	agtagtcgcc	gtgaacgctc	tttttcgcaa	ccgggttggc	4020
gccagaacac	aggtaaagtgc	cgtgtgtggt	tcccgcgggc	ctggcctctt	taogggttat	4080
ggcccttgcg	tgccctgaa	tacttccacc	tggctccagt	acgtgattct	tgatcccag	4140
ctggagccag	gggcgggccc	tgccctttag	gagccccttc	gcctcgtgct	tgagttaggg	4200
ccctggcctgg	gcgctggggc	gcgcgctg	gaatctgggtg	gcacctcgc	gctctctctg	4260
ctgctttega	taagttctca	gcoatttaaa	atthttgatg	acctgctg	acgctttttt	4320
tctggcaaga	tagtcttcta	aatgcgggccc	aggatctgca	cactgggtatt	tccggtttttg	4380
ggcccgcggc	cggcgacggg	gcccgtgctg	cccagcgcac	atgctcggcg	aggcggggccc	4440
tgcgagcgcg	gccaccgaga	atcgacgggg	ggtagtctca	agctggccgg	cctgctctgg	4500
tgccctggcc	cgcccgcccg	tgtatcgccc	cgccctgggc	ggcaaggctg	gcccggctcg	4560
caccagttgc	gtgagcggaa	agatggccgc	ttcccggccc	tgctccaggg	ggctcaaaat	4620
ggagyaccgc	gogctcggga	gagcggggcg	gtgagtcacc	cacacaaagg	aaaagggcct	4680
ttccgtcctc	agccgtcgc	tcatgtgact	ccaoggagta	ccgggcgcgc	tccaggcacc	4740
togattagtt	ctggagcttt	tggagtacgt	cgtctttagg	ttggggggag	gggtttttatg	4800
cgatggagtt	tcccacact	gagtgggtgg	agactgaagt	taggccagct	tggcacttga	4860
tgtaatctct	cttggaaatt	ggcctttttg	agtttggatc	ttggttcatt	ctcaagcctc	4920
agacagtgg	tcaaagttht	tttcttccat	ttcaggtgct	gtgaccatgg	ccagcgcgct	4980
ggacaagtcc	aaggtoatca	attccgcatt	agaggtcctt	aatgaggtcg	gaatcgagg	5040
tttaacaacc	cgtaaacctg	cccagaagct	aggtgtagag	cagcctacat	tgtattggca	5100
tgtaaaaaat	aagcgggctt	tgctcgacgc	cttagccatt	gagatgtag	ataggcacca	5160
tactcacttt	tgcccttttag	aaggggaaag	ctggcaagat	tttttacgta	ataacgctaa	5220
aagttttaga	tgtgcttttac	taagtcacog	cgatggagca	aaagtacatt	taggtacacg	5280
gctacagaa	aaacagtatg	aaactctcga	aaactcaatta	gcctttttat	gccaacaagg	5340
tttttacta	gagaatgcat	tgtacgccct	tgccgcctg	ggccacttca	ccctggctg	5400
tgtgctggag	gaccaagagc	atcaagtcgc	taaagaagaa	agggaaacac	ctactactga	5460
tagtatgccg	ccattattac	gacaagctat	cgaattattt	gatcaccac	gtgcagagcc	5520
agccttctta	ttcggccttg	aattgatcat	atgocgatta	gaaaaacaac	ttaaatgtga	5580

aagtgggtcc	gcgtacagcc	gaggcgccat	ggcctaactc	gagtttccct	ctagcgggat	5640
caattccgcc	ccccccctct	ccctcccccc	coctaacggt	actggcogaa	gccgcttgga	5700
ataaggccgg	tgtgogtttg	tctatatggt	atthttccacc	atattgcogt	cttttgccaa	5760
tgtgagggcc	cggaaacctg	gccctgtctt	cttgacgagc	atccttaggg	gtctttcccc	5820
tctcgccaaa	ggaatgcaag	gtctgttgaa	tgtcgtgaag	gaagcagttc	ctctggaagc	5880
ttcttgaaga	caaacaacgt	ctgtagcgac	cctttgcagg	cagcggaaacc	ccccacctgg	5940
cgacaggtgc	ctctgcggcc	aaaagccacg	tgtataagat	acacctgcaa	aggcggcaca	6000
accccagtg	cacgttgtga	gttggatagt	tgtggaaaga	gtcaaatggc	tctcctcaag	6060
cgtattcaac	aaggggctga	aggatgcccc	gaaggtacce	cattgtatgg	gatctgatct	6120
gggacctcgg	tgcacatget	ttacatgtgt	ttagtogagg	ttaaaaaac	gtctagggcc	6180
cccgaaccac	ggggacgtgg	ttttcctttg	aaaaacacga	tgataatggc	cacaaccatg	6240
gccaagcctt	tgtctcaaga	agaatccacc	ctcattgaaa	gagcaacggc	tacaatcaac	6300
agcatcccca	tctctgaaga	ctacagcgtc	gccagcgcag	ctctctctag	cgacggccgc	6360
atcttcaactg	gtgtcaatgt	atatcatttt	actgggggac	cttgtgcaga	actcgtggtg	6420
ctgggcaactg	ctgetgctgc	ggcagctggc	aacctgactt	gtatcgtcgc	gatcggaaat	6480
gagaacaggg	gcaccttgag	ccctgcgga	cggtgccgac	agggtcctct	cgatctgcat	6540
cctgggatca	aagccatagt	gaaggacagt	gatggacagc	cgacggcagt	tgggatctgt	6600
gaattgctgc	cctctgggta	tgtgtgggag	ggctaagtgc	acgtcacccg	cgacgtcgag	6660
gtgcccgaag	gaccgcgcac	ctggtgcatg	accgcgaagc	ccgggtccctg	acgcctcgac	6720
aatcaacctc	tggattacaa	aatttgtgaa	agattgactg	gtattcttaa	ctatgttgct	6780
ccttttaacg	tatgtggata	cgctgcttta	atgcctttgt	atcatgctat	tgettccogt	6840
atggctttca	ttttctcctc	cttgtataaaa	tctgtgttgc	tgtctcttta	tgaggagtgg	6900
tggcccgttg	tcaggcaacg	tggcgtggtg	tgcactgtgt	ttgctgacgc	aaacccact	6960
ggttggggca	ttgccaccac	ctgtcagctc	ctttccggga	ctttcgcttt	ccccctccct	7020
attgccacgg	cggaaactcat	cgccgcctgc	cttgcccgtc	gctggacagg	ggctcggctg	7080
ttgggcaactg	acaattccgt	ggtgttgtcg	gggaagctga	cgctccttcc	atggctgctc	7140
gcctgtgttg	ccacctggat	tctgcgcggg	acgtccttct	gctaogtccc	ttcggccctc	7200
aatcccaggg	accttccctc	ccgcggcctg	ctgcccgtct	tgcggcctct	tcocgctctt	7260
cgccttcgcc	ctcagacgag	tcggatctcc	ctttgggccc	ccctccccgc	tgggtacctt	7320
taagaccaat	gacttacaag	gcagctgtag	atcttagoca	cttttttaaaa	gaaaaggggg	7380
gactggaagg	gctaattcac	tcccacgaa	gacaagatct	gctttttgct	tgtactgggt	7440
ctctctggtt	agaccagatc	tgagcctggg	agctctctgg	ctaactaggg	aaacccactgc	7500
ttaagcctca	ataaagcttg	ccttgagtgc	ttcaagtagt	gtgtgcccgt	ctgttgtgtg	7560
actctggtaa	ctagagatcc	ctcagacctc	tttagtcagt	gtggaaaatc	tctagcagta	7620
gtagttcatg	tcattottatt	attcagt				7647

- <210> 52
- <211> 4987
- <212> DNA
- <213> 人工序列

- <220>
- <223> 对人工序列的描述: 合成构建体

<400> 52	gggctaatt	cactocccaaa	gaagacaaga	tatccttgat	ctgtggatct	accacacaca	60
aggctacttc	cctgattagc	agaactadac	accagggcca	ggggtcagat	atccactgac		120
ctttggatgg	tgctacaagc	tagtaccagt	tgagccagat	aaggtagaag	aggccaataa		180
aggagagaac	accagcttgt	tacaccctgt	gagcctgcat	gggatggatg	accoggagag		240
agaagtgtta	gagtggaggt	ttgacagccg	cctagcattt	catcacgtgg	cccagagact		300
gcattccggag	tacttcaaga	actgctgata	tcgagcttgc	tacaagggac	tttccgctgg		360
ggactttcca	gggaggcgtg	gcctggggcg	gactggggag	tggcgagccc	tcagatcctg		420
catataagca	gctgcttttt	gcctgtaactg	ggtctctctg	gtagaccag	atctgagcct		480
gggagctctc	tggctaacta	gggaacccac	tgettaagcc	tcaataaagc	ttgccttgag		540
tgcttcaagt	agtgtgtgcc	ogtctgttgt	gtgactctgg	taactagaga	tccttcagac		600
ccttttagtc	agtgtggaaa	atctctagca	gtggcgcggc	aacagggact	tgaagcgaa		660
agggaaaacca	gaggagctct	ctcgaagcag	gactcggctt	gctgaagcgc	gcacggcaag		720
aggcaggggg	cggcgactgg	tgagtaagcc	aaaaattttg	actagcggag	gctagaagga		780

gagagatggg	tgcgagagcg	tcagtattaa	gcgggggaga	attagatcgc	gatgggaaaa	840
aattcgggta	aggccagggg	gaaagaaaa	atataaatta	aaacatatag	tatgggcaag	900
cagggagcta	gaacgattcg	cagttaatcc	tggcctgtta	gaaacatcag	aaggctgtag	960
acaaatactg	ggacagctac	aaccatccct	tcagacagga	tcagaagaac	ttagatcatt	1020
atataataca	gtagcaaccc	tctattgtgt	gcatoaaagg	atagagataa	aagacaccaa	1080
ggaagcttta	gacaagatag	aggaagagca	aaacaaaagt	aagaccaccg	cacagcaagc	1140
ggccgctgat	cttcagacct	ggaggaggag	atatgaggga	caattggaga	agtgaattat	1200
ataaatataa	agtagtaaaa	attgaaccat	taggagtagc	accaccaag	gcaaagagaa	1260
gagtggtgca	gagagaaaa	agagcagtg	gaataggagc	tttgttcctt	gggttcttgg	1320
gagcagcagg	aagcactatg	ggcgagcgt	caatgacgct	gacggtacag	gccagacaat	1380
tattgtctgg	tatagtgcag	cagcagaaca	atttgctgag	ggctattgag	gcgcaacagc	1440
atctgttgca	actocagctc	tggggcatca	agcagctcca	ggcaagaatc	ctggctgtgg	1500
aaagatacct	aaaggatcaa	cagctectgg	ggatttgggg	ttgctotgga	aaactcattt	1560
gcaccactgc	tgtgccttgg	aatgctagtt	ggagtaataa	atctctggaa	cagatttggg	1620
atcacacgac	ctggatggag	tgggacagag	aaattaacaa	ttacacaagc	ttaatacact	1680
ccttaattga	agaatcgcaa	aaccagcaag	aaaagaatga	acaagaatta	ttggaattag	1740
ataaatgggc	aagtttgtgg	aattggttta	acataacaaa	ttggctgtgg	tatataaat	1800
tattcataat	gatagtagga	ggcttggtag	gtttaagaat	agttttgtct	gtactttcta	1860
tagtgaatag	agttaggcag	ggatattcac	cattatcggt	tcagaccac	ctccaaccc	1920
cgaggggacc	cgacaggccc	gaaggaaatg	aagaagaagg	tggagagaga	gacagagaca	1980
gatccattcg	attagtgaac	ggatctcgac	ggtatcgatt	ttaaaagaaa	aggggggatt	2040
ggggggtaca	gtgcagggga	agaatagta	gacataatag	caacagacat	acaaactaaa	2100
gaactacaaa	accaaattac	aaaaattcaa	aattttcggg	tttattacag	ggacagcaga	2160
gatccagttt	ggaatttcga	gtttaccact	ccctatcagt	gatagagaaa	agtgaaagtc	2220
gagtttacca	ctccctatca	gtgatagaga	aaagtgaag	tcgagtttac	cactccctat	2280
cagtgataga	gaaaagtga	agtcgagttt	accactccct	atcagtgata	gagaaaagtg	2340
aaagtcgagt	ttaccactcc	ctatcagtga	tagagaaaag	tgaaagtcga	gtttaccact	2400
ccctatcagt	gatagagaaa	agtgaagtc	gagtttacca	ctccctatca	gtgatagaga	2460
aaagtgaag	tcgagctcgg	tacccggtc	gagtagggct	gtacgggtgg	aggcctatat	2520
aagcagagct	cgtttagtga	accgtcagat	cgctgggaga	cgccatccac	gctgttttga	2580
cctccataga	agacaccggg	accgatccag	cctccgcggc	ccogaattcg	aattcgggatc	2640
cacgcgtact	agtctcgagg	aaaacctgta	ttttcagggc	gctcgaggag	attacaaaga	2700
tgacgacgat	aagcgcaacg	gccatcatca	ccatcaccat	caccaccatc	actaacgagt	2760
ttccctctag	cyggatcaat	tccgcccc	ccctctccct	ccccccct	aacgttactg	2820
gccgaagccg	cttggataaa	ggccgggtg	cggttctcta	tatgttattt	tcaccatat	2880
tgccgtcttt	tggcaatgtg	agggcccggg	aacctggccc	tgtcttcttg	acgagcattc	2940
ctaggggtct	ttccctcttc	gccaaaggaa	tgcaaggctc	gttgaatgtc	gtgaagggaag	3000
cagttcctct	ggaagcttct	tgaagacaaa	caacgtctgt	agcgaccctt	tgcaaggcagc	3060
ggaaccccc	acctggcgac	aggtgectct	ggggccaaaa	gccacgtgta	taagatacac	3120
ctgcaaaagc	gycacaaccc	cagtgccacg	ttgtgagttg	gatagttgtg	gaaagagtca	3180
aatggctctc	ctcaagcgta	ttcaacaagg	ggctgaagga	tgcccagaag	gtaccccatt	3240
gtatgggate	tgatctgggg	cctcggtgca	catgctttac	atgtgtttag	tcgagtttaa	3300
aaaaacgtct	aggccccccg	aaccaogggg	acgtggtttt	ccittgaaaa	acacgatgat	3360
aatggccaca	accatgggtg	ctgaatacaa	accaactggt	cgccctggcaa	ctcgtgatga	3420
tgttccacgt	gcagttcgca	ccctgggtgc	tgcatttgc	gactaccctg	caaccctca	3480
cactgtggac	ccagaccgcc	acattgaacg	tgtgactgaa	ctgcaggagc	tgttccctgac	3540
ccgtgtgggc	ctggacattg	gcaaagtgtg	gggtggcagat	gatggtgctg	ctgtggcagt	3600
gtggaccacc	cctgaatctg	ttgaagctgg	tgcaagtttt	gctgagattg	gccacgcat	3660
ggcagaactg	tctggcagcc	gcttggcagc	acaacagcag	atggaaggtc	tgctggcacc	3720
acaccgccca	aaagaacctg	cttggttcct	ggcaactgtg	ggtgtgagcc	ctgaccacca	3780
gggtaagggc	ctgggctctg	cagtgggtct	gcctgggtgtg	gaagcagctg	aacgtgcagg	3840
tgtgocctgct	ttcctggaga	cctcagctcc	acgcaacctc	ccittctatg	aacgcctggg	3900
cttcaactgtg	actgctgatg	tggaaagtgc	agaaggccca	cgacttgggt	gcatgactgc	3960
caaaccaggt	gcttaagtgc	acgtcacgcg	cgacgtcgag	gtgcccgaag	gaccgycac	4020
ctgggtgatg	accgcgaagc	ccgggtgctg	acgcctcgac	aatcaacctc	tggattacaa	4080
aatttgtgaa	agattgactg	gtattcttaa	ctatgttgct	ccitttaacg	tatgtggata	4140
cgctgcttta	atgcctttgt	atcatgctat	tgcttccgct	atggctttca	ttttctcttc	4200
cttgtataaa	tcctggttgc	tgtctcttta	tgaggagttg	tggcccgttg	tcaggcaacg	4260
tggcgtgggtg	tgcaactgtg	ttgctgacgc	aaaccccat	ggttggggca	ttgccaccac	4320
ctgtcagctc	ctttccggga	ctttcgcttt	ccccctccct	attgccacgg	oggaactcat	4380
cgccgcctgc	cttgcocgct	gctggacagg	ggctcggtcg	ttgggcaactg	acaattccgt	4440

ggtgttgteg	gggaagctga	egtceetted	atggetgctc	gcctgtgttg	ccacctggat	4500
tctgcgcggg	acgtceetted	getacgtccc	ttcggccctc	aatccagcgg	accttccttc	4560
ccgcggcctg	ctgcgcgctc	tgccgctctc	tcgggtatt	cgcttcgcc	ctcagacgag	4620
tcggatctcc	ctttgggcgg	cctcccggcc	tgggtacctt	taagaccaat	gacttacaag	4680
gcagctgtag	atcttagcca	ctttttaaaa	gaaaaggggg	gactggaagg	gctaattcac	4740
tcccaacgaa	gacaagatct	gctttttgct	tgtactgggt	ctctctgggt	agaccagatc	4800
tgagcctggg	agctctctgg	ctaactaggg	aacccactgc	ttaagcctca	ataaagcttg	4860
ccttgagtgc	ttcaagtagt	gtgtgcccg	ctgttgtgtg	actctggtaa	ctagagatcc	4920
ctcagacct	tttagtcagt	gtggaaaatc	tctagcagta	gtagttcatg	tcctcttatt	4980
attcagt						4987

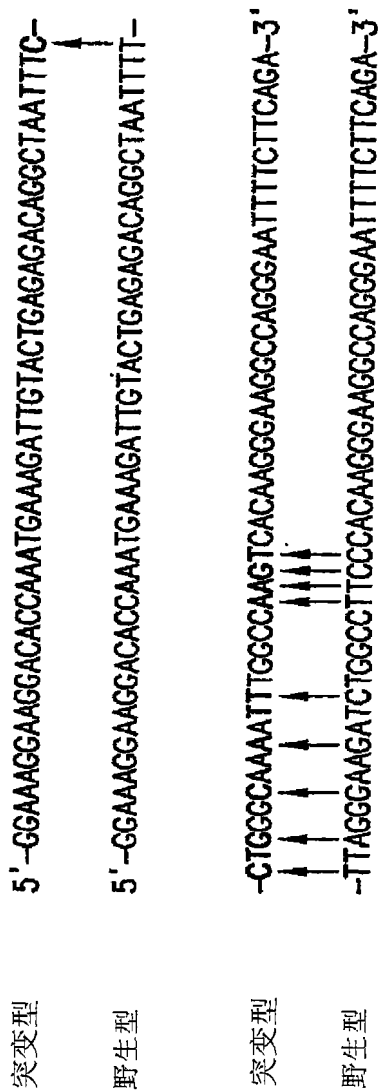


图 1

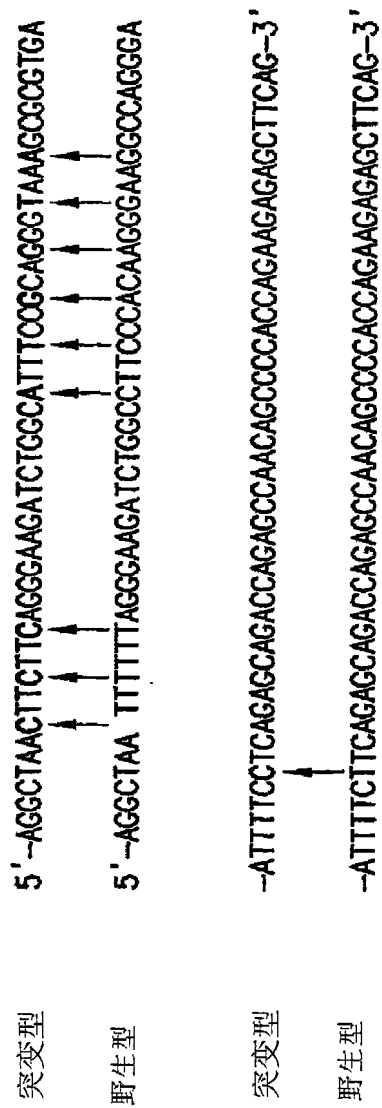


图 2

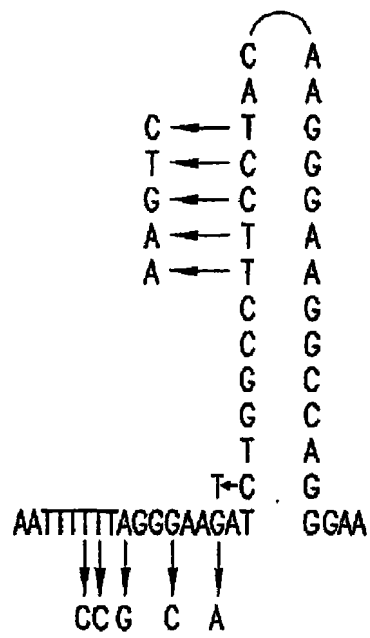


图 3

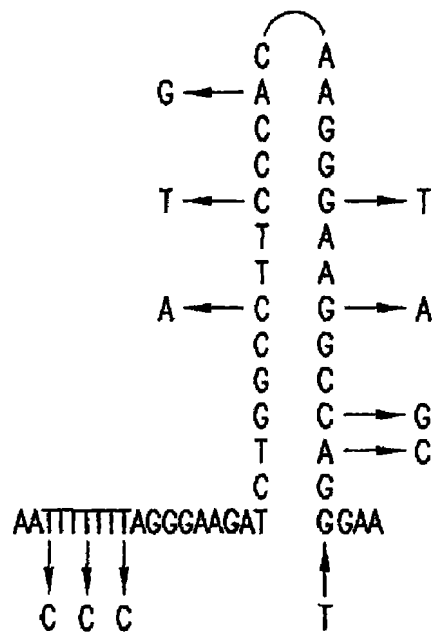


图 4

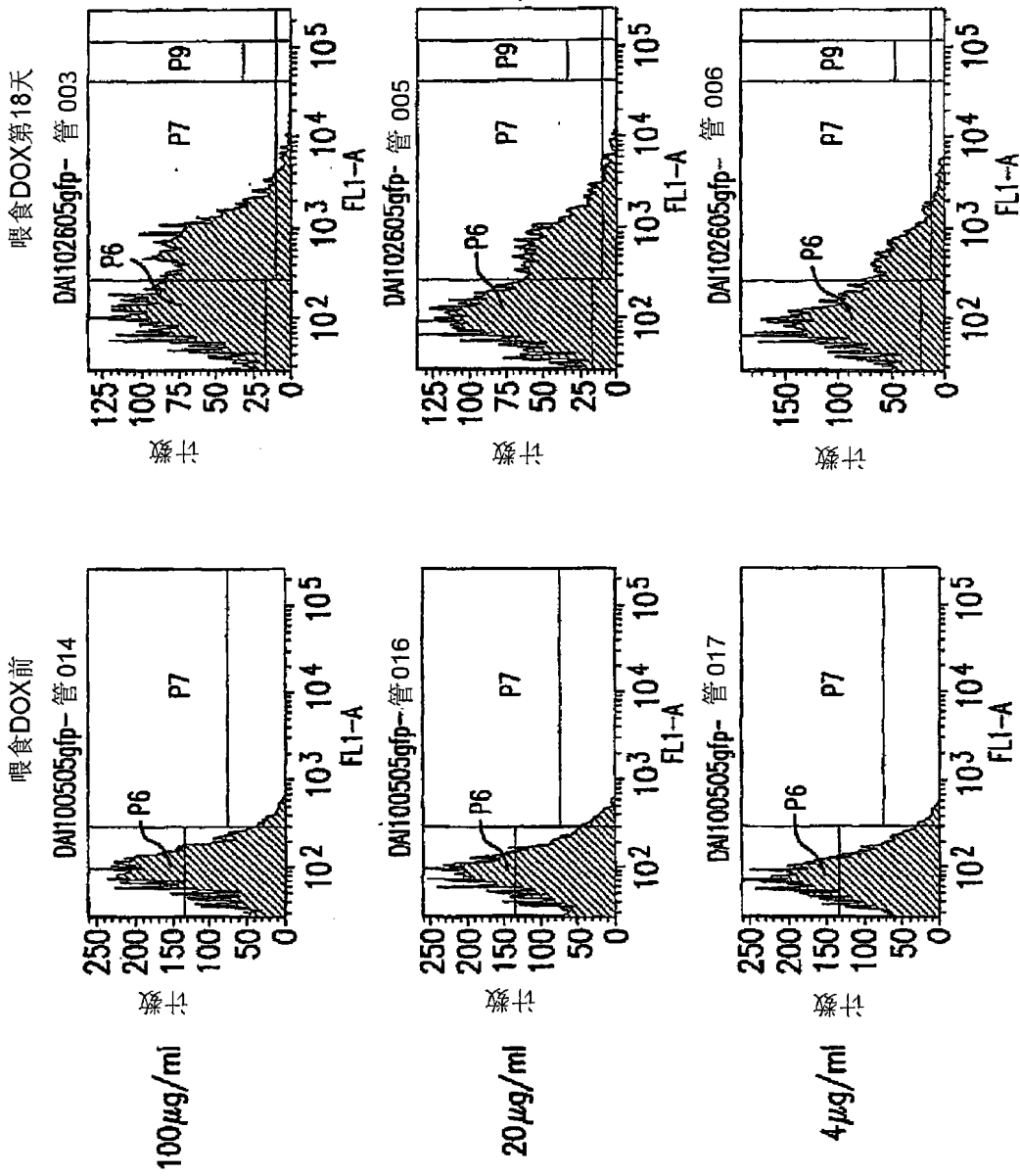


图 5

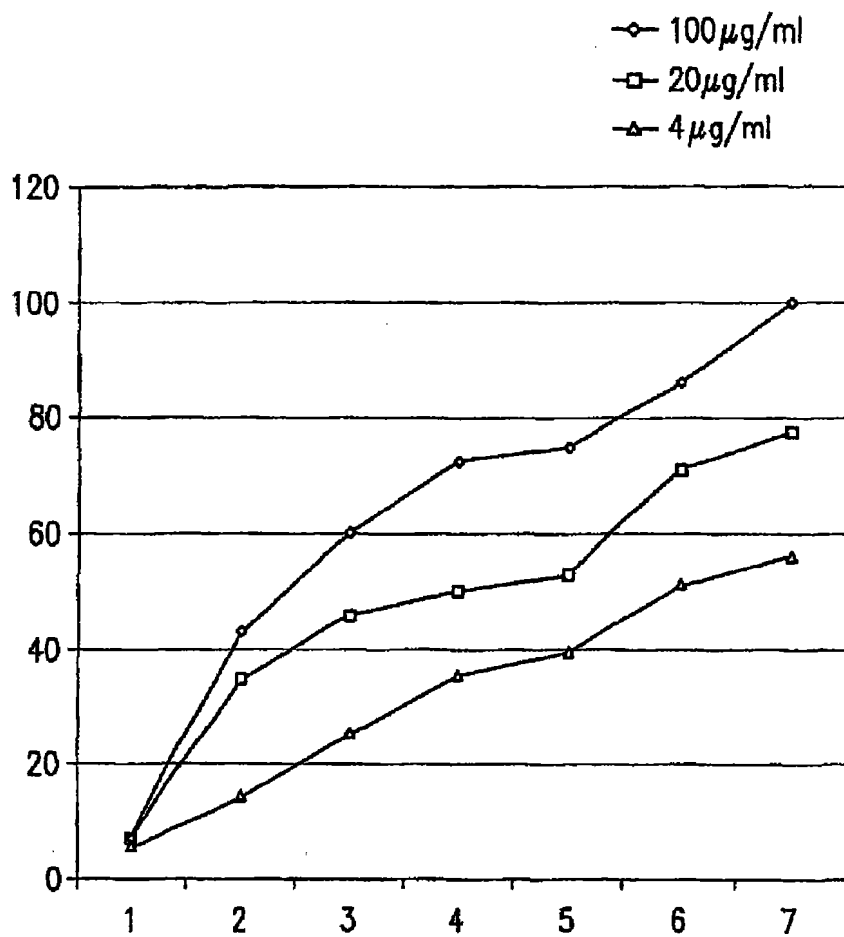


图 6

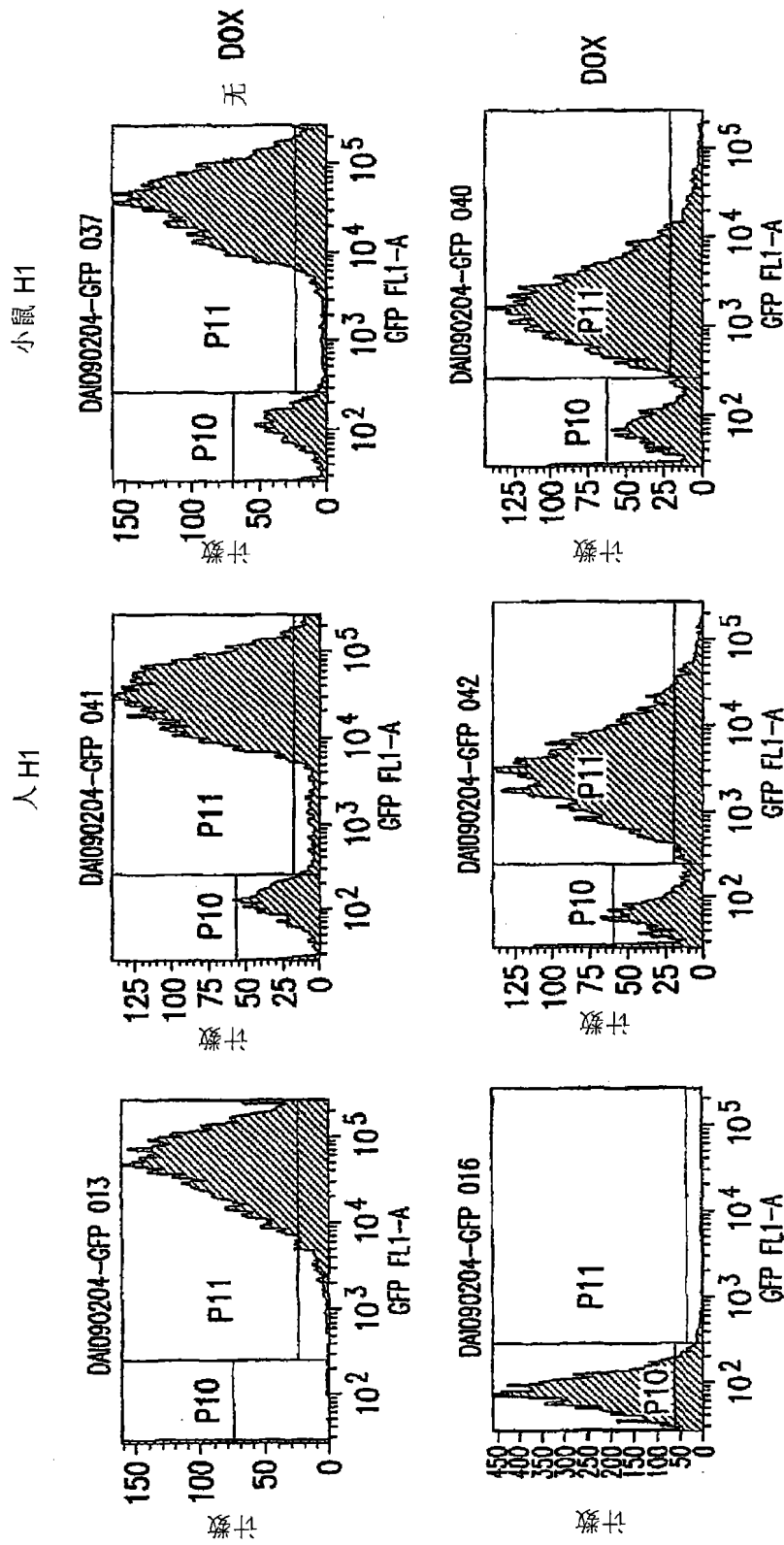


图 7

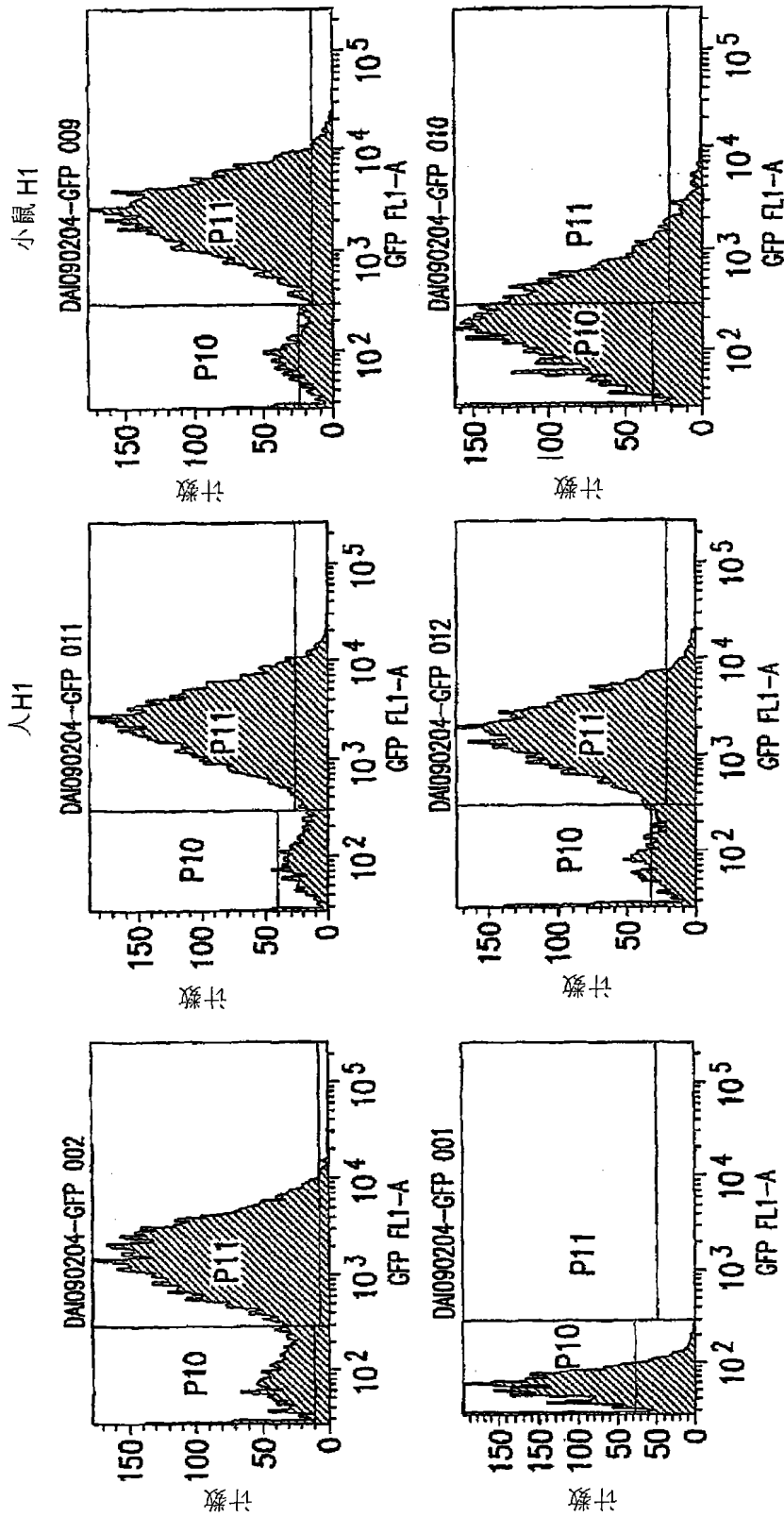


图 8

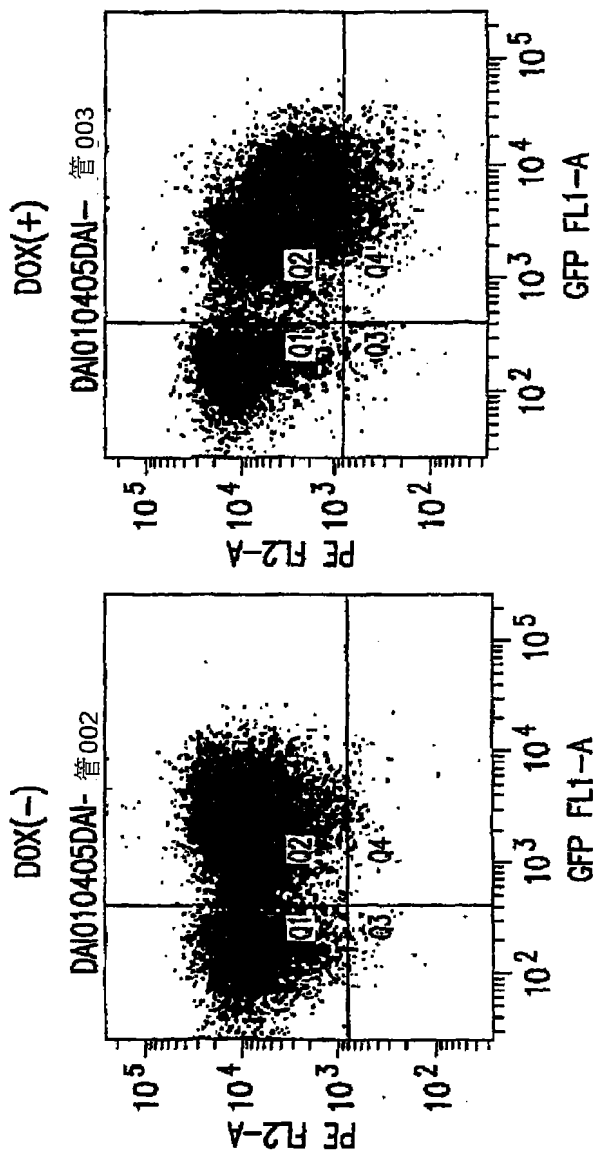


图 9

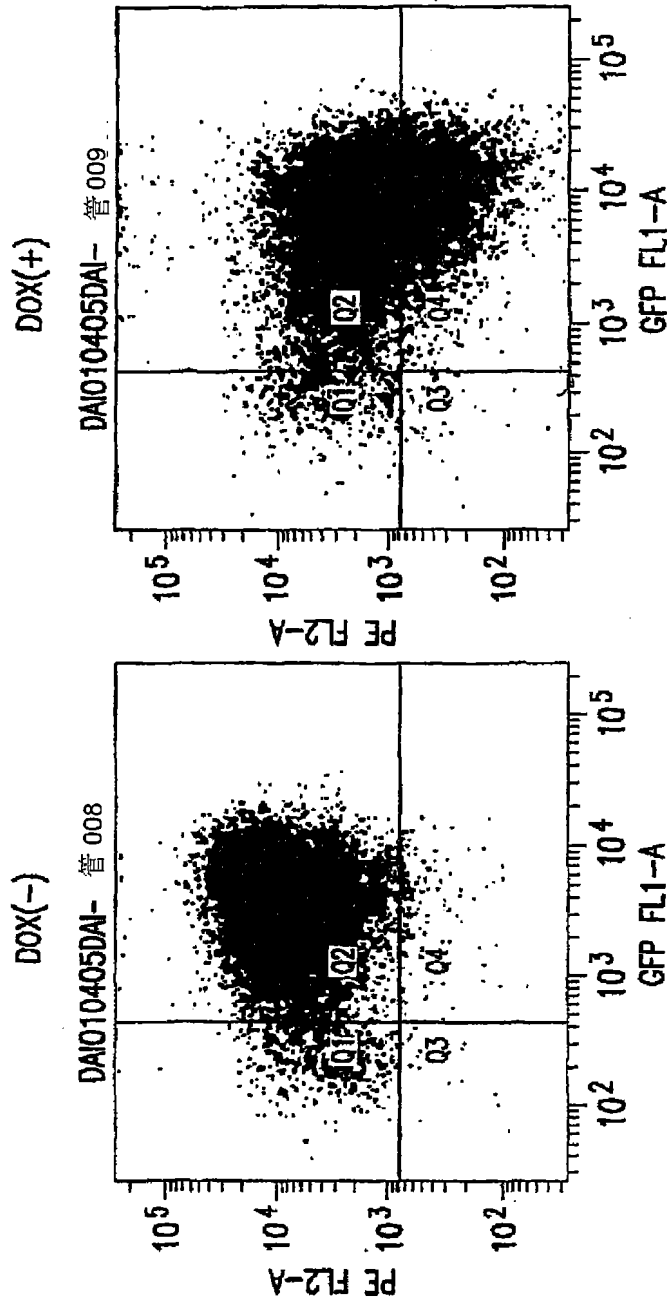


图 10

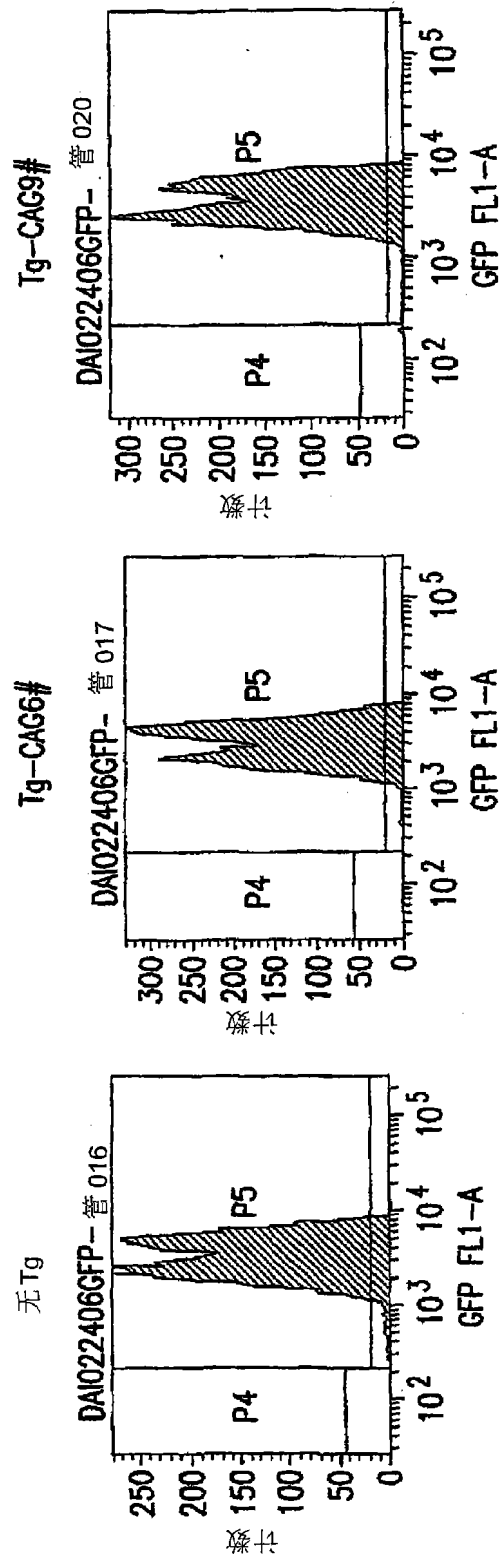


图 11

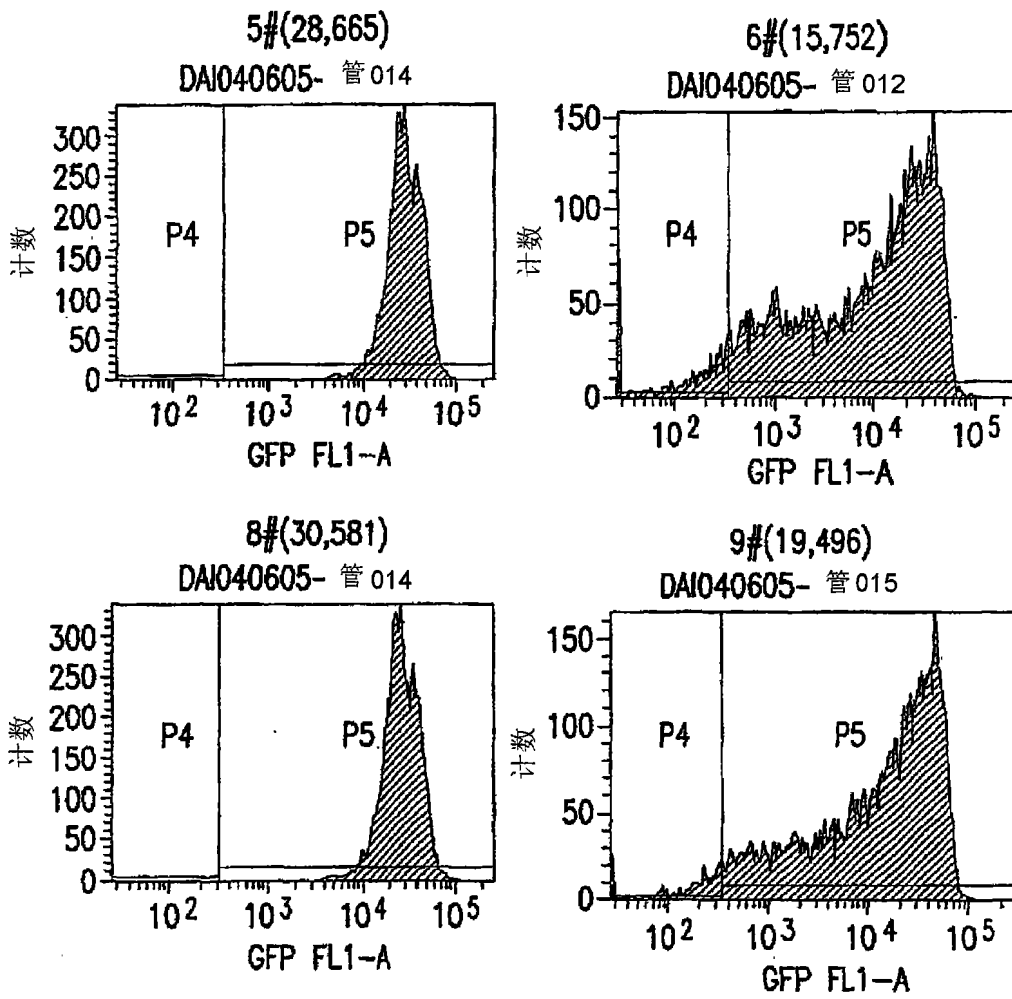


图 12

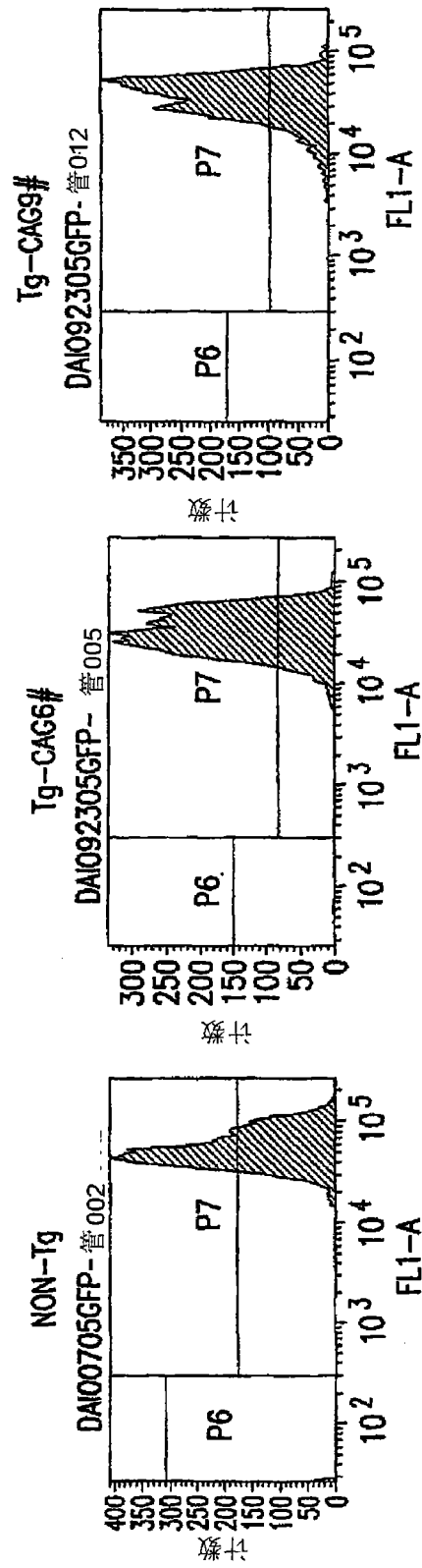


图 13

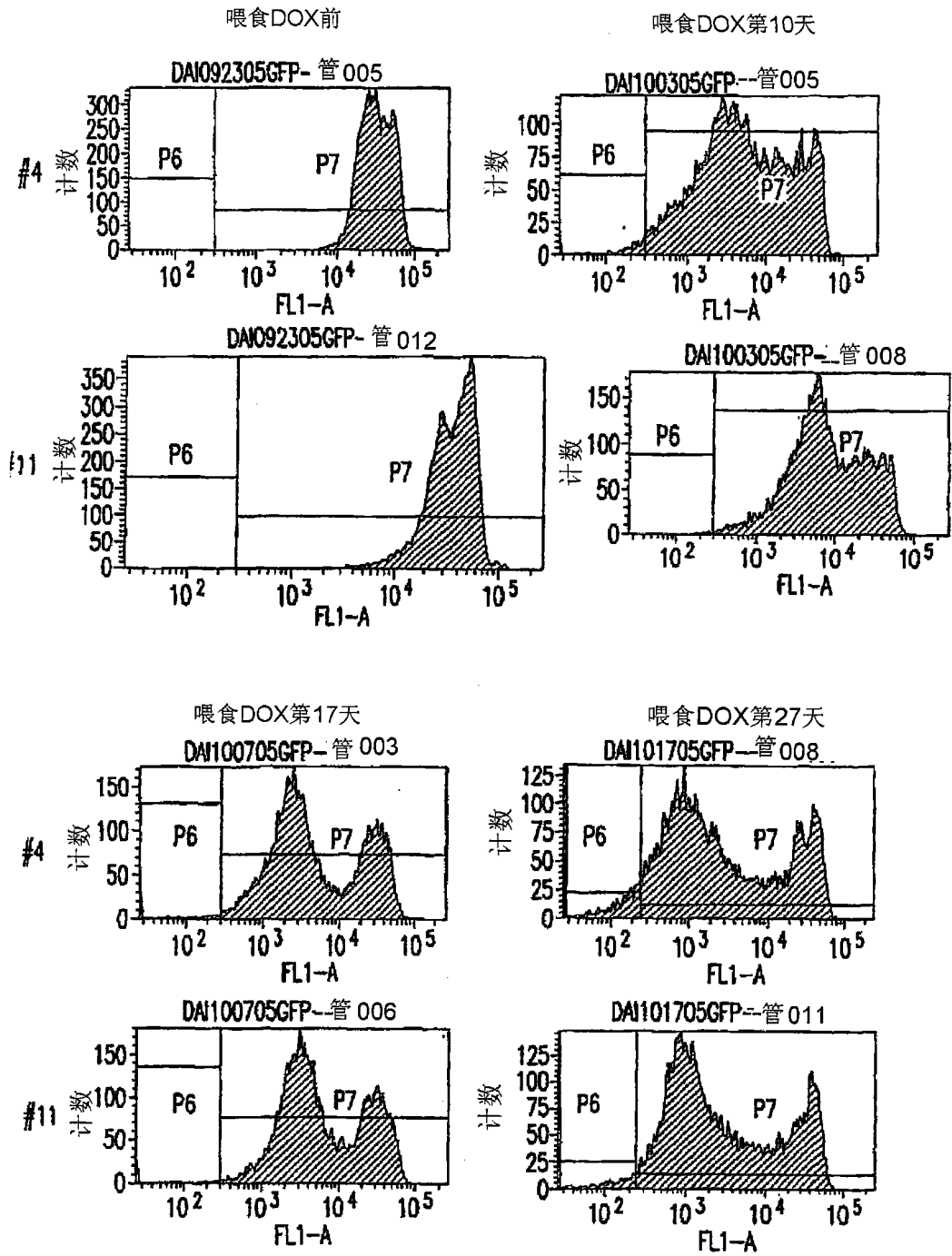
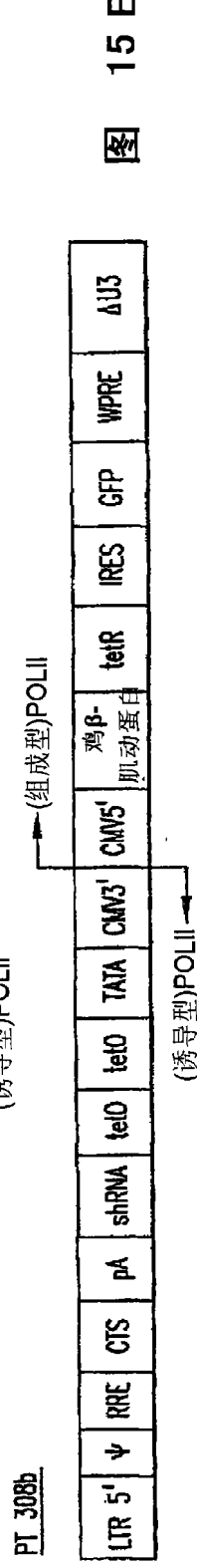
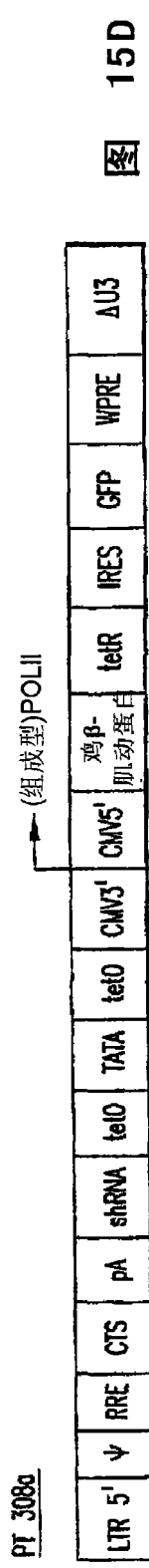
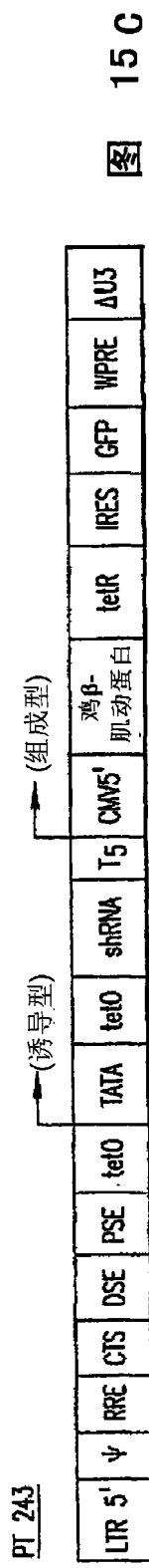
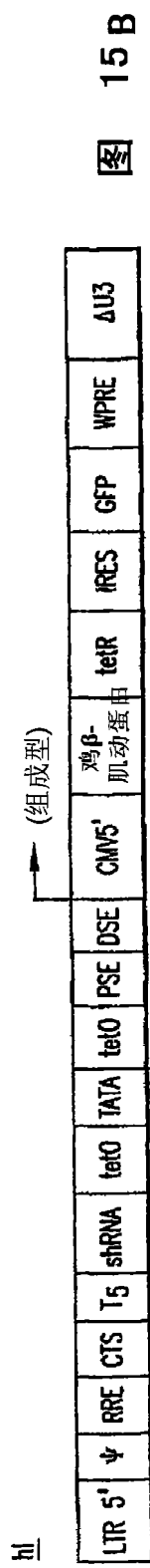
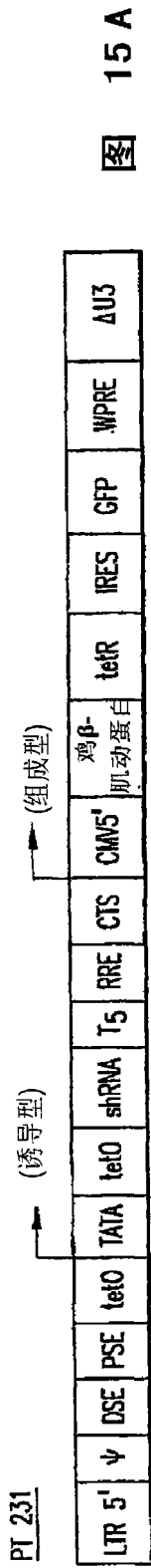
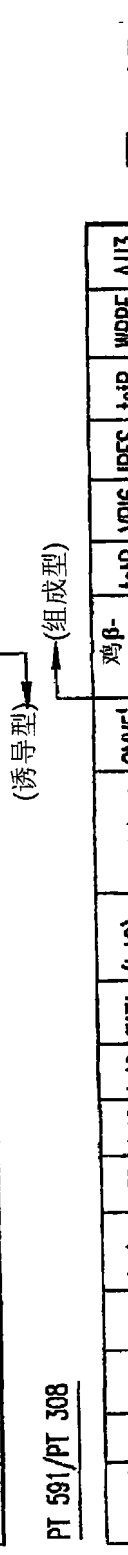
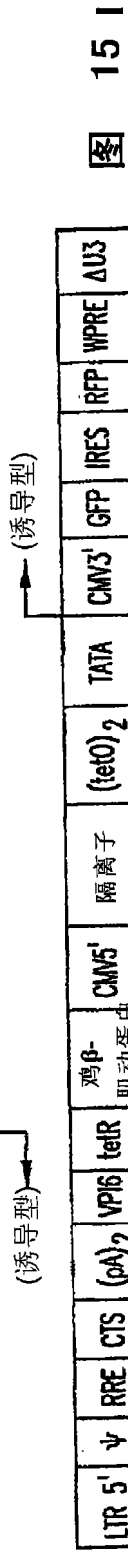
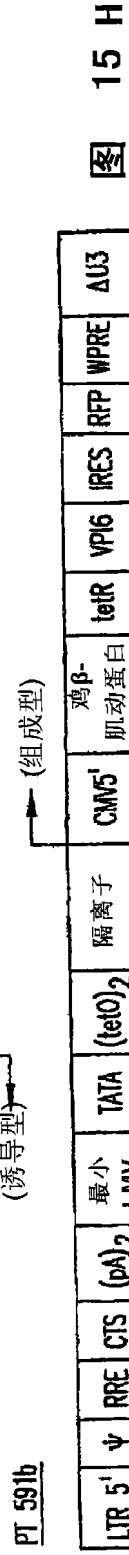
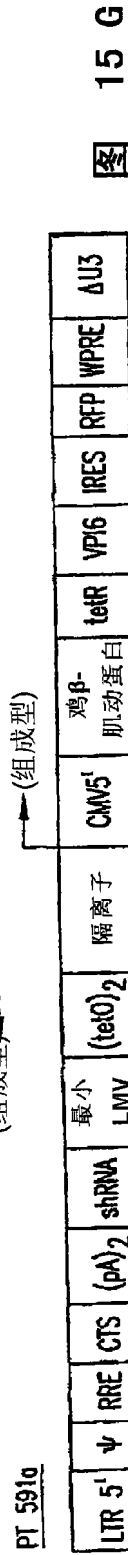
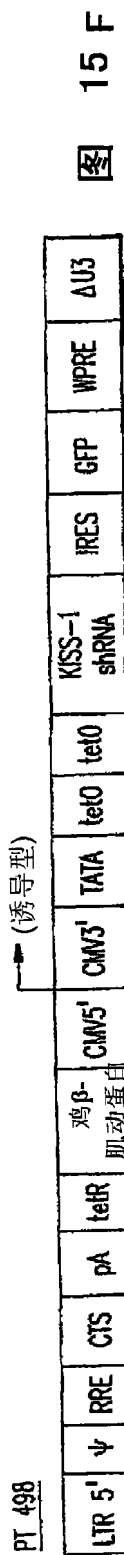


图 14





基于慢病毒载体的hCCR1-m表达

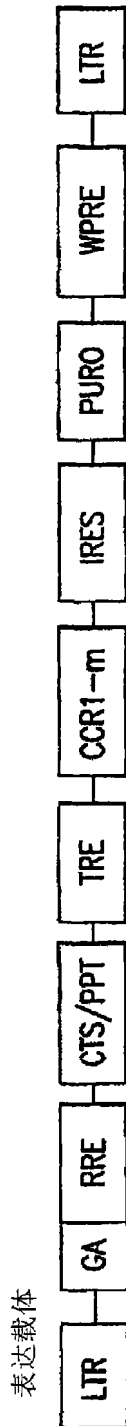


图 16 A

CCR1-m的C末端AA序列

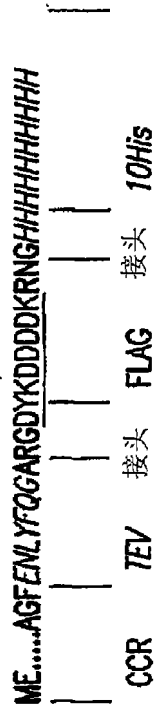


图 16 B

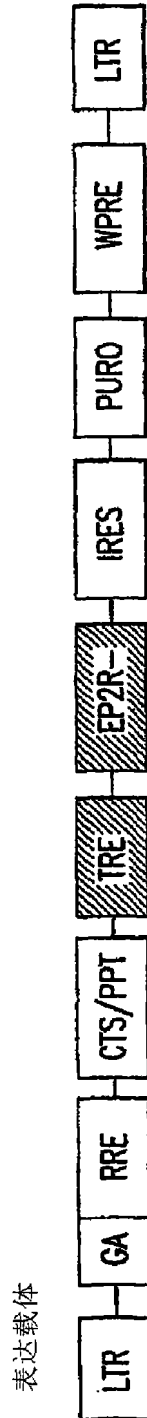


图 17 A

hEP2P的C末端AA序列

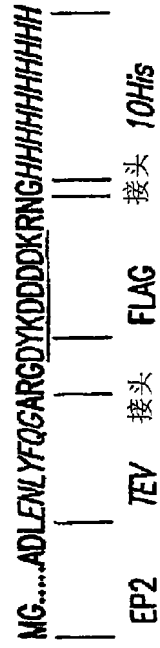


图 17 B

专利名称(译)	与使用病毒载体控制的基因表达相关的组合物和方法		
公开(公告)号	<a href="#">CN101384902A</a>	公开(公告)日	2009-03-11
申请号	CN200680052999.8	申请日	2006-12-18
[标]申请(专利权)人(译)	UAB研究基金会		
申请(专利权)人(译)	UAB研究基金会		
当前申请(专利权)人(译)	UAB研究基金会		
[标]发明人	X吴 J卡比斯		
发明人	X·吴 J·卡比斯		
IPC分类号	G01N33/53 C12P21/06 C12N1/00 C12N5/02 C12N1/20 C07H21/04 A61K31/70 A01K67/00		
CPC分类号	A01K67/0275 A01K2217/052 A01K2217/203 A61P31/18 C12N15/63 C12N15/635 C12N2740/16043 C12N2800/30 C12N2830/003		
代理人(译)	姜建成		
优先权	60/751407 2005-12-16 US 60/751117 2005-12-16 US		
外部链接	<a href="#">Espacenet</a> <a href="#">SIPO</a>		

摘要(译)

本发明提供了与病毒载体相关的方法和组合物。本发明还提供了有效转染宿主(例如通过高效慢病毒送递载体)的方法和组合物,以及通过单纯地将调节物给予携带被转移的目的序列的宿主来灵敏控制所述被转移的目的序列的表达时机和水平。本发明还公开了制备转基因小鼠的方法以及使用与病毒载体相关的组合物和方法制备得到的转基因小鼠。

不同启动子	DOX(-)	DOX(+)	诱导
EF-1 $\alpha$	133	15,071	113倍
CAG	157	7,071	45倍
UB6	263	4,550	17倍