



(12)发明专利申请

(10)申请公布号 CN 107864657 A

(43)申请公布日 2018.03.30

(21)申请号 201680033119.6

(22)申请日 2016.04.07

(30)优先权数据

62/144,862 2015.04.08 US

62/220,067 2015.09.17 US

(85)PCT国际申请进入国家阶段日

2017.12.06

(86)PCT国际申请的申请数据

PCT/US2016/026486 2016.04.07

(87)PCT国际申请的公布数据

W02016/164609 EN 2016.10.13

(71)申请人 建新公司

地址 美国马萨诸塞州

(72)发明人 S·库斯迪奥-穆尔 D·索萨

K·文森特

(74)专利代理机构 北京坤瑞律师事务所 11494

代理人 谢梦欣

(51)Int.Cl.

C12N 7/00(2006.01)

C12N 7/01(2006.01)

C07K 14/755(2006.01)

C12N 5/10(2006.01)

A61K 35/76(2015.01)

C12N 15/861(2006.01)

C12N 7/04(2006.01)

A61K 48/00(2006.01)

C12N 15/864(2006.01)

C12N 15/86(2006.01)

C12N 15/35(2006.01)

权利要求书8页 说明书49页

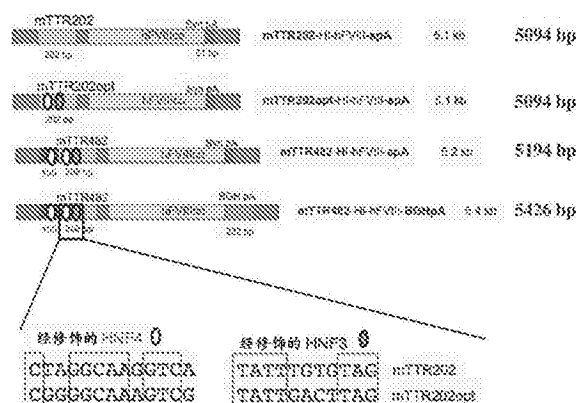
序列表14页 附图23页

(54)发明名称

超大腺相关载体的产生

(57)摘要

本文中提供了用于产生含有超大重组腺相关病毒(AAV)基因组(例如,大于4.7kb)的AAV颗粒的方法。在一些态样,本发明提供了包含超大rAAV基因组的AAV载体和AAV颗粒。还提供了生产细胞系以产生包含超大基因组的AAV颗粒。



1. 一种用于产生包含超大重组AAV基因组的腺相关病毒 (AAV) 颗粒的方法, 该方法包括:

a) 在产生rAAV颗粒的条件下培养AAV生产细胞系, 其中该AAV生产细胞系包含:

i) 编码AAV rep和cap基因的核酸, 和

ii) rAAV基因组, 其中该rAAV基因组大于约4.7kb;

b) 提供AAV辅助功能; 并

c) 收集包含超大rAAV基因组的rAAV颗粒。

2. 权利要求1的方法, 其中该编码AAV rep和cap基因的核酸和/或该rAAV基因组在该生产细胞系中稳定维持。

3. 权利要求1或2的方法, 其中该编码AAV rep和cap基因的核酸和/或该rAAV基因组稳定整合入该生产细胞系的基因组中。

4. 权利要求1-3中任一项的方法, 其中该rAAV基因组包含一个或多个AAV反向末端重复 (ITR) 和异源转基因。

5. 权利要求1-4中任一项的方法, 其中该rAAV基因组包含两个AAV ITR。

6. 权利要求1-5中任一项的方法, 其中该rAAV基因组是约4.7kb-约9.4kb, 任选地约4.7kb-6.7kb。

7. 权利要求1-6中任一项的方法, 其中步骤c) 中收集的该AAV颗粒包含大于约4.7kb的rAAV基因组。

8. 权利要求1-7中任一项的方法, 其中步骤c) 中收集的该AAV颗粒包含约4.7kb-约9.4kb的rAAV基因组。

9. 权利要求1-8中任一项的方法, 其中该rAAV基因组是约4.7kb-约5kb、约4.7kb-约6kb、约4.7kb-约7kb、约4.7kb-约8kb或约4.7kb-约9kb。

10. 权利要求1-8中任一项的方法, 其中该rAAV基因组是约4.7kb-6.7kb或约5.2kb-约8.7kb。

11. 权利要求1-10中任一项的方法, 其中该异源转基因编码治疗性转基因产物。

12. 权利要求1-11中任一项的方法, 其中该异源转基因编码因子VIII、抗肌萎缩蛋白、dysferlin或囊性纤维化跨膜传导调节蛋白 (CFTR)。

13. 权利要求1-12中任一项的方法, 其中该异源转基因是人转基因。

14. 权利要求1-13中任一项的方法, 其中该异源转基因与启动子可操作连接。

15. 权利要求14的方法, 其中该启动子是小鼠转甲状腺素 (mTTR) 启动子。

16. 权利要求1-15中任一项的方法, 其中该rAAV基因组包含内含子。

17. 权利要求16的方法, 其中该内含子是合成内含子。

18. 权利要求1-17中任一项的方法, 其中该rAAV基因组包含多聚腺苷酸化信号。

19. 权利要求18的方法, 其中该多聚腺苷酸化信号是合成多聚腺苷酸化信号或牛生长激素多聚腺苷酸化信号。

20. 权利要求1-19中任一项的方法, 其中该rAAV颗粒包含AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10、AAVrh10、AAV11、AAV12、AAV2R471A、AAV2/2-7m8、AAV DJ、AAV2N587A、AAV2E548A、AAV2N708A、AAV V708K、山羊AAV、AAV1/AAV2嵌合、牛AAV或小鼠AAV壳体rAAV2/HBoV1血清型壳体。

21. 权利要求20的方法, 其中该AAV血清型是AAV1、AAV2、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10或AAVrh10。

22. 权利要求4-21中任一项的方法, 其中该AAV ITR是AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10、AAVrh10、AAV11、AAV12、AAV2R471A、AAV DJ、山羊AAV、牛AAV或小鼠AAV血清型ITR。

23. 权利要求4-20中任一项的方法, 其中该AAV ITR是AAV2ITR。

24. 权利要求20-23中任一项的方法, 其中该rAAV颗粒的ITR和壳体源自相同AAV血清型。

25. 权利要求24的方法, 其中该ITR和该壳体源自AAV2。

26. 权利要求20-23中任一项的方法, 其中该rAAV颗粒的ITR和壳体源自不同AAV血清型。

27. 权利要求26的方法, 其中该AAV颗粒包含AAV2ITR和AAVrh8R壳体。

28. 权利要求26的方法, 其中该AAV颗粒包含AAV2ITR和AAV8壳体。

29. 权利要求1-28中任一项的方法, 其中该生产细胞系源自灵长类细胞。

30. 权利要求1-28中任一项的方法, 其中该生产细胞系源自HeLa、293、A549或Perc.6细胞。

31. 权利要求1-30中任一项的方法, 其中该生产细胞系适合于在悬浮液中生长。

32. 权利要求1-31中任一项的方法, 其中该AAV辅助功能由腺病毒、HSV或杆状病毒提供。

33. 权利要求32的方法, 其中在提供辅助功能后约48小时-约96小时收集该rAAV颗粒。

34. 权利要求1-33中任一项的方法, 其进一步包括纯化该rAAV颗粒。

35. 权利要求34的方法, 其中该纯化包括一个或多个层析步骤。

36. 通过权利要求1-35中任一项的方法产生的包含超大rAAV基因组的rAAV颗粒。

37. 一种包含rAAV颗粒的组合物, 其中至少约15%、至少约20%、至少约25%、至少约30%、至少约35%、至少约40%、至少约45%、至少约50%、至少约55%、至少约60%或至少约70%的该rAAV颗粒壳体化大于约4.7kb的rAAV基因组。

38. 权利要求37的组合物, 其中该rAAV基因组包含一个或多个AAV反向末端重复 (ITR) 和异源转基因。

39. 权利要求37或38的组合物, 其中该rAAV基因组包含两个AAV ITR。

40. 权利要求37-39中任一项的组合物, 其中该rAAV基因组是约4.7kb-约9.4kb。

41. 权利要求37-40中任一项的组合物, 其中该rAAV基因组是约4.7kb-约5kb、约4.7kb-约6kb、约4.7kb-约7kb、约4.7kb-约8kb或约4.7kb-约9kb。

42. 权利要求37-40中任一项的组合物, 其中该rAAV基因组是约4.7kb-6.7kb或约5.2kb-约8.7kb。

43. 权利要求37-42中任一项的组合物, 其中该异源转基因编码治疗性转基因产物。

44. 权利要求37-43中任一项的组合物, 其中该异源转基因编码因子VIII、抗肌萎缩蛋白、dysferlin或囊性纤维化跨膜传导调节蛋白 (CFTR)。

45. 权利要求37-44中任一项的组合物, 其中该异源转基因是人转基因。

46. 权利要求37-45中任一项的组合物, 其中该异源转基因与启动子可操作连接。

47. 权利要求46的组合物,其中该启动子是小鼠转甲状腺素(mTTR)启动子。
48. 权利要求37-47中任一项的组合物,其中该rAAV基因组包含内含子。
49. 权利要求48的组合物,其中该内含子是合成内含子。
50. 权利要求37-49中任一项的组合物,其中该rAAV基因组包含多聚腺苷酸化信号。
51. 权利要求50的组合物,其中该多聚腺苷酸化信号是合成多聚腺苷酸化信号或牛生长激素多聚腺苷酸化信号。
52. 权利要求37-51中任一项的组合物,其中该rAAV颗粒包含AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10、AAVrh10、AAV11、AAV12、AAV2R471A、AAV2/2-7m8、AAV DJ、AAV2N587A、AAV2E548A、AAV2N708A、AAV V708K、山羊AAV、AAV1/AAV2嵌合、牛AAV或小鼠AAV壳体rAAV2/HBoV1血清型壳体。
53. 权利要求52的组合物,其中该AAV血清型是AAV1、AAV2、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10或AAVrh10。
54. 权利要求38-51中任一项的组合物,其中该AAV ITR是AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10、AAVrh10、AAV11、AAV12、AAV2R471A、AAV DJ、山羊AAV、牛AAV或小鼠AAV血清型ITR。
55. 权利要求38-54中任一项的组合物,其中该AAV ITR是AAV2ITR。
56. 权利要求52-55中任一项的组合物,其中该rAAV颗粒的ITR和壳体源自相同AAV血清型。
57. 权利要求56的组合物,其中该ITR和该壳体源自AAV2。
58. 权利要求52-55中任一项的组合物,其中该rAAV颗粒的ITR和壳体源自不同AAV血清型。
59. 权利要求58的组合物,其中该AAV颗粒包含AAV2ITR和AAVrh8R壳体。
60. 权利要求58的组合物,其中该AAV颗粒包含AAV2ITR和AAV8壳体。
61. 权利要求37-60中任一项的组合物,其中在生产细胞系中产生该rAAV颗粒。
62. 权利要求61的组合物,其中编码AAV rep和cap基因的核酸和/或该rAAV基因组在该生产细胞系中稳定维持。
63. 权利要求61或62的组合物,其中该编码AAV rep和cap基因的核酸和/或该rAAV基因组稳定整合入该生产细胞系的基因组中。
64. 权利要求61-63中任一项的组合物,其中该生产细胞系源自灵长类细胞。
65. 权利要求61-64中任一项的组合物,其中该生产细胞系源自HeLa、293、A549或Perc.6细胞。
66. 权利要求61-65中任一项的组合物,其中该生产细胞系适合于在悬浮液中生长。
67. 权利要求61-66中任一项的组合物,其中通过对该生产细胞系提供AAV辅助功能而产生该rAAV颗粒。
68. 权利要求67的组合物,其中该AAV辅助功能由腺病毒、HSV或杆状病毒提供。
69. 权利要求68的组合物,其中在提供辅助功能后约48小时-约96小时收集该rAAV颗粒。
70. 一种用于增强超大rAAV基因组的表达的方法,该方法包括通过对生产细胞系提供AAV辅助功能而在该细胞系中生产rAAV颗粒,其中该生产细胞系包含:

- a) 编码AAV rep和cap基因的核酸,和
- b) rAAV基因组,其中该rAAV基因组大于约4.7kb。

71. 权利要求70的方法,其中该超大rAAV基因组的表达是来自通过瞬时转染产生的rAAV颗粒的该超大rAAV基因组的表达的大约1.25倍、约1.5倍、约1.75倍、约2.0倍、约2.5倍、约2.75倍、约3倍或约5倍。

72. 权利要求70或71的方法,其中与来自通过瞬时转染产生的rAAV颗粒的超大rAAV基因组的表达动力学相比,来自自由生产细胞系产生的rAAV颗粒的该超大rAAV基因组的表达动力学是更快的表达动力学。

73. 权利要求72的方法,其中来自自由生产细胞系产生的rAAV颗粒的该超大rAAV基因组的表达动力学比来自通过瞬时转染产生的rAAV颗粒的超大rAAV基因组的表达动力学快约5%、快约10%、快约25%、快约50%、快约75%或快约90%。

74. 权利要求70-73中任一项的方法,其中该编码AAV rep和cap基因的核酸和/或该rAAV基因组在该生产细胞系中稳定维持。

75. 权利要求70-74中任一项的方法,其中该编码AAV rep和cap基因的核酸和/或该rAAV基因组稳定整合入该生产细胞系的基因组中。

76. 权利要求70-75中任一项的方法,其中该rAAV基因组包含一个或多个AAV反向末端重复(ITR)和异源转基因。

77. 权利要求70-76中任一项的方法,其中该rAAV基因组包含两个AAV ITR。

78. 权利要求70-77中任一项的方法,其中该rAAV基因组是约4.7kb-约9.4kb。

79. 权利要求70-78中任一项的组合物,其中该rAAV基因组是约4.7kb-约5kb、约4.7kb-约6kb、约4.7kb-约7kb、约4.7kb-约8kb或约4.7kb-约9kb。

80. 权利要求70-78中任一项的组合物,其中该rAAV基因组是约4.7kb-6.7kb或约5.2kb-约8.7kb。

81. 权利要求70-78中任一项的方法,其中该异源转基因编码治疗性转基因产物。

82. 权利要求70-81中任一项的方法,其中该异源转基因编码因子VIII、抗肌萎缩蛋白、dysferlin或囊性纤维化跨膜传导调节蛋白(CFTR)。

83. 权利要求70-82中任一项的方法,其中该异源转基因是人转基因。

84. 权利要求70-83中任一项的方法,其中该异源转基因与启动子可操作连接。

85. 权利要求84的方法,其中该启动子是小鼠转甲状腺素(mTTR)启动子。

86. 权利要求70-85中任一项的方法,其中该rAAV基因组包含内含子。

87. 权利要求86的方法,其中该内含子是合成内含子。

88. 权利要求70-87中任一项的方法,其中该rAAV基因组包含多聚腺苷酸化信号。

89. 权利要求88的方法,其中该多聚腺苷酸化信号是合成多聚腺苷酸化信号或牛生长激素多聚腺苷酸化信号。

90. 权利要求70-89中任一项的方法,其中该rAAV颗粒包含AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10、AAVrh10、AAV11、AAV12、AAV2R471A、AAV2/2-7m8、AAV DJ、AAV2N587A、AAV2E548A、AAV2N708A、AAV V708K、山羊AAV、AAV1/AAV2嵌合、牛AAV或小鼠AAV壳体rAAV2/HBoV1血清型壳体。

91. 权利要求90的方法,其中该AAV血清型是AAV1、AAV2、AAV5、AAV6、AAV7、AAV8、

AAVrh8、AAVrh8R、AAV9、AAV10或AAVrh10。

92. 权利要求76-91中任一项的方法, 其中该AAV ITR是AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10、AAVrh10、AAV11、AAV12、AAV2R471A、AAV DJ、山羊AAV、牛AAV或小鼠AAV血清型ITR。

93. 权利要求76-92中任一项的方法, 其中该AAV ITR是AAV2ITR。

94. 权利要求90-93中任一项的方法, 其中该rAAV颗粒的ITR和壳体源自相同AAV血清型。

95. 权利要求94的方法, 其中该ITR和该壳体源自AAV2。

96. 权利要求90-93中任一项的方法, 其中该rAAV颗粒的ITR和壳体源自不同AAV血清型。

97. 权利要求96的方法, 其中该AAV颗粒包含AAV2ITR和AAVrh8R壳体。

98. 权利要求96的方法, 其中该AAV颗粒包含AAV2ITR和AAV8壳体。

99. 权利要求70-98中任一项的方法, 其中该生产细胞系源自灵长类细胞。

100. 权利要求70-99中任一项的方法, 其中该生产细胞系源自HeLa、293、A549或Perc.6细胞。

101. 权利要求70-100中任一项的方法, 其中该生产细胞系适合于在悬浮液中生长。

102. 权利要求70-101中任一项的方法, 其中该AAV辅助功能由腺病毒、HSV或杆状病毒提供。

103. 权利要求102的方法, 其中在提供辅助功能后约48小时-约96小时收集该rAAV颗粒。

104. 权利要求70-103中任一项的方法, 其进一步包括纯化该rAAV颗粒。

105. 权利要求104的方法, 其中该纯化包括一个或多个层析步骤。

106. 一种用于产生包含超大重组AAV基因组的腺相关病毒(AAV)颗粒的细胞系, 该细胞系包含:

a) 编码AAV rep和cap基因的核酸, 和

b) rAAV基因组, 其中该rAAV基因组大于约4.7kb。

107. 权利要求106的细胞系, 其中该编码AAV rep和cap基因的核酸和/或该rAAV基因组在该生产细胞系中稳定维持。

108. 权利要求106或107的细胞系, 其中该编码AAV rep和cap基因的核酸和/或该rAAV基因组稳定整合入该生产细胞系的基因组中。

109. 权利要求106-108中任一项的细胞系, 其中该rAAV基因组包含一个或多个AAV反向末端重复(ITR)和异源转基因。

110. 权利要求106-109中任一项的细胞系, 其中该rAAV基因组包含两个AAV ITR。

111. 权利要求106-110中任一项的细胞系, 其中该rAAV基因组是约4.7kb-约9.4kb。

112. 权利要求106-111中任一项的方法, 其中该rAAV基因组是约4.7kb-约5kb、约4.7kb-约6kb、约4.7kb-约7kb、约4.7kb-约8kb或约4.7kb-约9kb。

113. 权利要求106-111中任一项的方法, 其中该rAAV基因组是约4.7kb-6.7kb或约5.2kb-约8.7kb。

114. 权利要求106-113中任一项的细胞系, 其中该异源转基因编码治疗性转基因产物。

115. 权利要求106-114中任一项的细胞系,其中该异源转基因编码因子VIII、抗肌萎缩蛋白、dysferlin或囊性纤维化跨膜传导调节蛋白(CFTR)。

116. 权利要求106-115中任一项的细胞系,其中该异源转基因是人转基因。

117. 权利要求106-116中任一项的细胞系,其中该异源转基因与启动子可操作连接。

118. 权利要求117的细胞系,其中该启动子是小鼠转甲状腺素(mTTR)启动子。

119. 权利要求106-118中任一项的细胞系,其中该rAAV基因组包含内含子。

120. 权利要求119的细胞系,其中该内含子是合成内含子。

121. 权利要求106-120中任一项的细胞系,其中该rAAV基因组包含多聚腺苷酸化信号。

122. 权利要求121的细胞系,其中该多聚腺苷酸化信号是合成多聚腺苷酸化信号或牛生长激素多聚腺苷酸化信号。

123. 权利要求106-122中任一项的细胞系,其中该rAAV颗粒包含AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10、AAVrh10、AAV11、AAV12、AAV2R471A、AAV2/2-7m8、AAV DJ、AAV2N587A、AAV2E548A、AAV2N708A、AAV V708K、山羊AAV、AAV1/AAV2嵌合、牛AAV或小鼠AAV壳体rAAV2/HBoV1血清型壳体。

124. 权利要求123的细胞系,其中该AAV血清型是AAV1、AAV2、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10或AAVrh10。

125. 权利要求109-124中任一项的细胞系、其中该AAV ITR是AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10、AAVrh10、AAV11、AAV12、AAV2R471A、AAV DJ、山羊AAV、牛AAV或小鼠AAV血清型ITR。

126. 权利要求109-125中任一项的细胞系,其中该AAV ITR是AAV2ITR。

127. 权利要求123-126中任一项的细胞系,其中该rAAV颗粒的ITR和壳体源自相同AAV血清型。

128. 权利要求127的细胞系,其中该ITR和该壳体源自AAV2。

129. 权利要求123-126中任一项的细胞系,其中该rAAV颗粒的ITR和壳体源自不同AAV血清型。

130. 权利要求129的细胞系,其中该AAV颗粒包含AAV2ITR和AAVrh8R壳体。

131. 权利要求129的细胞系,其中该AAV颗粒包含AAV2ITR和AA8壳体。

132. 权利要求106-131中任一项的细胞系,其中该生产细胞系源自灵长类细胞。

133. 权利要求106-132中任一项的细胞系,其中该生产细胞系源自HeLa、293、A549或Perc.6细胞。

134. 权利要求106-133中任一项的细胞系,其中该生产细胞系适合于在悬浮液中生长。

135. 权利要求106-134中任一项的细胞系,其中该AAV辅助功能由腺病毒、HSV或杆状病毒提供。

136. 一种腺相关病毒(AAV)颗粒,其包含由AAV壳体壳体化的rAAV基因组,其中该rAAV基因组大于约4.7kb。

137. 权利要求136的AAV颗粒,其中该rAAV基因组包含一个或多个AAV反向末端重复(ITR)和异源转基因。

138. 权利要求136或137的AAV颗粒,其中该rAAV基因组包含两个AAV ITR。

139. 权利要求136-138中任一项的AAV颗粒,其中该rAAV基因组是约4.7kb-约9.4kb。

140. 权利要求136-139中任一项的AAV颗粒,其中该rAAV基因组是约4.7kb-约5kb、约4.7kb-约6kb、约4.7kb-约7kb、约4.7kb-约8kb或约4.7kb-约9kb。

141. 权利要求136-139中任一项的AAV颗粒,其中该rAAV基因组是约4.7kb-6.7kb或约5.2kb-约8.7kb。

142. 权利要求136-141中任一项的AAV颗粒,其中该异源转基因编码治疗性转基因产物。

143. 权利要求136-142中任一项的AAV颗粒,其中该异源转基因编码因子VIII、抗肌萎缩蛋白、dysferlin或囊性纤维化跨膜传导调节蛋白(CFTR)。

144. 权利要求136-143中任一项的AAV颗粒,其中该异源转基因是人转基因。

145. 权利要求136-144中任一项的AAV颗粒,其中该异源转基因与启动子可操作连接。

146. 权利要求145的AAV颗粒,其中该启动子是小鼠转甲状腺素(mTTR)启动子。

147. 权利要求136-146中任一项的AAV颗粒,其中该rAAV基因组包含内含子。

148. 权利要求147的AAV颗粒,其中该内含子是合成内含子。

149. 权利要求136-148中任一项的AAV颗粒,其中该rAAV基因组包含多聚腺苷酸化信号。

150. 权利要求149的AAV颗粒,其中该多聚腺苷酸化信号是合成多聚腺苷酸化信号或牛生长激素多聚腺苷酸化信号。

151. 权利要求136-150中任一项的AAV颗粒,其中该rAAV颗粒包含AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10、AAVrh10、AAV11、AAV12、AAV2R471A、AAV2/2-7m8、AAV DJ、AAV2N587A、AAV2E548A、AAV2N708A、AAV V708K、山羊AAV、AAV1/AAV2嵌合、牛AAV或小鼠AAV壳体rAAV2/HBoV1血清型壳体。

152. 权利要求151的AAV颗粒,其中该AAV血清型是AAV1、AAV2、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10或AAVrh10。

153. 权利要求137-152中任一项的AAV颗粒,其中该AAV ITR是AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10、AAVrh10、AAV11、AAV12、AAV2R471A、AAV DJ、山羊AAV、牛AAV或小鼠AAV血清型ITR。

154. 权利要求151-153中任一项的AAV颗粒,其中该AAV ITR是AAV2 ITR。

155. 权利要求137-154中任一项的AAV颗粒,其中该rAAV颗粒的ITR和壳体源自相同AAV血清型。

156. 权利要求155的AAV颗粒,其中该ITR和该壳体源自AAV2。

157. 权利要求151-154中任一项的AAV颗粒,其中该rAAV颗粒的ITR和壳体源自不同AAV血清型。

158. 权利要求157的AAV颗粒,其中该AAV颗粒包含AAV2ITR和AAVrh8R壳体。

159. 权利要求158的AAV颗粒,其中该AAV颗粒包含AAV2ITR和AAV8壳体。

160. 权利要求136-142中任一项的AAV颗粒,其中该rAAV基因组从5'至3'包含AAV2ITR、mTTR启动子、合成内含子、编码人FVIII的转基因、合成多聚腺苷酸化序列、和AAV2ITR。

161. 权利要求136-142中任一项的AAV颗粒,其中该rAAV基因组从5'至3'包含AAV2ITR、mTTR启动子、合成内含子、编码人FVIII的转基因、牛生长激素合成多聚腺苷酸化序列、和AAV2ITR。

162. 权利要求160或161的AAV颗粒, 其中该FVIII包含全部或部分的B域的缺失。

163. 权利要求160-162中任一项的AAV颗粒, 其中该AAV颗粒包含AAVrh8R壳体。

164. 权利要求160-162中任一项的AAV颗粒, 其中该AAV颗粒包含AAV8壳体。

165. 一种包含rAAV基因组的rAAV载体, 其中该rAAV基因组从5' 至3' 包含AAV2ITR、mTTR启动子、合成内含子、编码人FVIII的转基因、牛生长激素合成多聚腺苷酸化序列、和AAV2ITR。

166. 一种包含rAAV基因组的rAAV载体, 其中该rAAV基因组从5' 至3' 包含AAV2ITR、mTTR启动子、合成内含子、编码人FVIII的转基因、牛生长激素合成多聚腺苷酸化序列、和AAV2ITR。

167. 权利要求165或166的AAV载体, 其中该FVIII包含全部或部分的B域的缺失。

168. 一种用于治疗个体中的疾病或病症的方法, 其包括投予包含超大AAV基因组的AAV颗粒, 其中该超大AAV基因组包含适合于治疗该疾病或病症的转基因。

169. 权利要求168的方法, 其中该疾病或病症是血友病A, 并且该包含超大AAV基因组的AAV颗粒编码因子VIII转基因。

170. 权利要求169的方法, 其中该AAV颗粒是权利要求160-164中任一项的AAV颗粒。

171. 权利要求168的方法, 其中该疾病或病症是肌营养不良症, 并且该包含超大AAV基因组的AAV颗粒编码抗肌萎缩蛋白转基因。

172. 权利要求168的方法, 其中该疾病或病症是dysferlinopathy, 并且该包含超大AAV基因组的AAV颗粒编码dysferlin转基因。

173. 权利要求168的方法, 其中该疾病或病症是囊性纤维化, 并且该包含超大AAV基因组的AAV颗粒编码CFTR转基因。

174. 权利要求168-173中任一项的方法, 其中该个体是人。

超大腺相关载体的产生

[0001] 对相关申请的交叉引用

[0002] 本申请要求2015年4月8日提交的美国临时申请No.62/144,862和2015年9月17日提交的美国临时申请No.62/220,067的优先权,在此通过提及完整并入每篇的内容用于所有目的。

[0003] ASCII文本文件上的序列表的提交

[0004] ASCII文本文件上的下述提交的内容通过提述以其整体并入本文:序列表的计算可读形式(CRF)(文件名称:159792013240SEQLIST.txt,记录日期:2016年4月6日,大小:27KB)。

发明领域

[0005] 本发明涉及用于产生具有超大重组AAV基因组的腺相关病毒(AAV)颗粒的方法和细胞系。

[0006] 发明背景

[0007] 重组AAV(rAAV)载体已经变为用于针对遗传和慢性疾病的基因转移的有吸引力的投递媒介物。使用rAAV载体的限制之一是它们的小包装容量,其已经阻碍了用于许多需要大cDNA的临床应用的基因疗法,例如因子VIII(FVIII)、抗肌萎缩蛋白、dysferlin和囊性纤维化跨膜传导调节蛋白(CFTR)。早期研究限定包装限制为4.7至4.8kb(Dong, J-Y等(1996) Human Gene Therapy 7:2101-2112)。新近的研究已经对AAV2、AAV5或AAV8壳体确认了大致在5.0至5.2kb大小的包装载体基因组的限制。在这些研究中,两种极性的超大(或“片段化”)基因组在5'端通常是缺失的,并且大多数包装的基因组不超过约5.2kb(Lu, H.等(2008) Human Gene Therapy 19:648-654; Wu, Z.等(2010) Molecular Therapy 18:80-86; Grose, W.E.等(2012) PLoS One 7:e39233)。

[0008] 因而,需要用于超大载体的更好的生产平台,其容许稳健产率及足够质量的产生。

[0009] 发明概述

[0010] 本文中描述了通过生产细胞系(PCL)平台的超大载体产生的全面分析。如下文描述,此PCL平台产生较高产率的较好质量的超大重组腺相关病毒(rAAV)载体。相比于通过标准三重转染方法产生的载体中观察到的,产生的rAAV载体含有更高量的壳体化的较大基因组。另外,细胞系是稳定的,并且载体含有较少的污染性异常DNA。载体在基因转移后在体内产生完全的表达盒,并且导致功能性蛋白质的产生。

[0011] 本发明提供了用于产生包含超大重组AAV基因组的腺相关病毒(AAV)颗粒的方法,所述方法包括:a)在产生rAAV颗粒的条件下培养AAV生产细胞系,其中所述AAV生产细胞系包含:i)编码AAV rep和cap基因的核酸,和ii) rAAV基因组,其中所述rAAV基因组大于约4.7kb;b)提供AAV辅助功能;并c)收集包含超大rAAV基因组的rAAV颗粒。在一些实施方案中,所述编码AAV rep和cap基因的核酸和/或所述rAAV基因组在所述生产细胞系中稳定维持。在一些实施方案中,所述编码AAV rep和cap基因的核酸和/或所述rAAV基因组稳定整合入所述生产细胞系的基因组中。在一些实施方案中,所述rAAV基因组包含一个或多个AAV反

向末端重复 (ITR) 和异源转基因。在一些实施方案中,所述rAAV基因组包含两个AAV ITR。在一些实施方案中,rAAV基因组是约4.7kb-约9.4kb,任选地约4.7kb-6.7kb。在一些实施方案中,步骤c)中收集的AAV颗粒包含大于约4.7kb的rAAV基因组。在一些实施方案中,步骤c)中收集的AAV颗粒包含约4.7kb-约9.4kb,任选地约4.7kb-6.7kb的rAAV基因组。在一些实施方案中,rAAV基因组是约4.7kb-约5kb、约4.7kb-约6kb、约4.7kb-约7kb、约4.7kb-约8kb或约4.7kb-约9kb。在一些实施方案中,rAAV基因组是约4.7kb-6.7kb或约5.2kb-约8.7kb。在一些实施方案中,rAAV基因组的长度大于约5.0kb、5.1kb、5.2kb、5.3kb、5.4kb、5.5kb、5.6kb、5.7kb、5.8kb、5.9kb、6.0kb、6.1kb、6.2kb、6.3kb、6.4kb、6.5kb、6.6kb、6.7kb、6.8kb、6.9kb、7.0kb、8.0kb或9.0kb中任一者或其间的任何数值。

[0012] 在上述方法的一些实施方案中,异源转基因编码治疗性转基因产物。在一些实施方案中,异源转基因是人转基因。在一些实施方案中,异源转基因编码因子VIII、抗肌萎缩蛋白、dysferlin或囊性纤维化跨膜传导调节蛋白 (CFTR)。在一些实施方案中,异源转基因与启动子可操作连接。在别的实施方案中,启动子是小鼠转甲状腺素 (mTTR) 启动子。在一些实施方案中,rAAV基因组包含内含子。在别的实施方案中,内含子是合成内含子。在一些实施方案中,rAAV基因组包含多聚腺苷酸化信号。在别的实施方案中,多聚腺苷酸化信号是合成多聚腺苷酸化信号或牛生长激素多聚腺苷酸化信号。

[0013] 在上述方法的一些实施方案中,rAAV颗粒包含AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10、AAVrh10、AAV11、AAV12、AAV2R471A、AAV2/2-7m8、AAV DJ、AAV2N587A、AAV2E548A、AAV2N708A、AAV V708K、山羊AAV、AAV1/AAV2嵌合、牛AAV或小鼠AAV壳体rAAV2/HBoV1血清型壳体。在一些实施方案中,AAV血清型是AAV1、AAV2、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10或AAVrh10。在上述方法的一些实施方案中,AAV ITR是AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10、AAVrh10、AAV11、AAV12、AAV2R471A、AAV DJ、山羊AAV、牛AAV或小鼠AAV血清型ITR。在一些实施方案中,AAV ITR是AAV2ITR。在一些实施方案中,rAAV颗粒的ITR和壳体源自相同AAV血清型。在一些实施方案中,ITR和所述壳体源自AAV2。在其它实施方案中,rAAV颗粒的ITR和壳体源自不同AAV血清型。在一些实施方案中,AAV颗粒包含AAV2ITR和AAVrh8R壳体。在一些实施方案中,AAV颗粒包含AAV2ITR和AAV8壳体。

[0014] 在上述方法的一些实施方案中,生产细胞系源自灵长类细胞。在一些实施方案中,生产细胞系源自HeLa、293、A549或Perc.6细胞。在一些实施方案中,生产细胞系适合于在悬浮液中生长。在一些实施方案中,AAV辅助功能由腺病毒或HSV提供。在一些实施方案中,在提供辅助功能后约48小时-约96小时收集所述rAAV颗粒。在一些实施方案中,方法进一步包括纯化rAAV颗粒。在一些实施方案中,纯化包括一个或多个层析步骤。在一些态样中,本发明提供了通过本文中描述的方法产生的包含超大rAAV基因组的rAAV颗粒。

[0015] 在一些态样中,本发明提供了包含rAAV颗粒的组合物,其中至少约15%、至少约20%、至少约25%、至少约30%、至少约35%、至少约40%、至少约45%、至少约50%、至少约55%、至少约60%或至少约70%的所述rAAV颗粒壳体化大于约4.7kb的rAAV基因组。在一些实施方案中,所述rAAV基因组包含一个或多个AAV反向末端重复 (ITR) 和异源转基因。在一些实施方案中,所述rAAV基因组包含两个AAV ITR。在一些实施方案中,rAAV基因组是约4.7kb-约9.4kb。在一些实施方案中,rAAV基因组是约4.7kb-约5kb、约4.7kb-约6kb、约

4.7kb-约7kb、约4.7kb-约8kb或约4.7kb-约9kb。在一些实施方案中，rAAV基因组是约4.7kb-6.7kb或约5.2kb-约8.7kb。在一些实施方案中，rAAV基因组的长度大于约5.0kb、5.1kb、5.2kb、5.3kb、5.4kb、5.5kb、5.6kb、5.7kb、5.8kb、5.9kb、6.0kb、6.1kb、6.2kb、6.3kb、6.4kb、6.5kb、6.6kb、6.7kb、6.8kb、6.9kb、7.0kb、8.0kb或9.0kb中任一者或其间的任何数值。

[0016] 在上述组合物的一些实施方案中，异源转基因编码治疗性转基因产物。在一些实施方案中，异源转基因是人转基因。在一些实施方案中，异源转基因编码因子VIII、抗肌萎缩蛋白、dysferlin或囊性纤维化跨膜传导调节蛋白(CFTR)。在一些实施方案中，异源转基因与启动子可操作连接。在别的实施方案中，启动子是小鼠转甲状腺素(mTTR)启动子。在一些实施方案中，rAAV基因组包含内含子。在别的实施方案中，内含子是合成内含子。在一些实施方案中，rAAV基因组包含多聚腺苷酸化信号。在别的实施方案中，多聚腺苷酸化信号是合成多聚腺苷酸化信号或牛生长激素多聚腺苷酸化信号。

[0017] 在上述组合物的一些实施方案中，rAAV颗粒包含AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10、AAVrh10、AAV11、AAV12、AAV2R471A、AAV2/2-7m8、AAV DJ、AAV2N587A、AAV2E548A、AAV2N708A、AAV V708K、山羊AAV、AAV1/AAV2嵌合、牛AAV或小鼠AAV壳体rAAV2/HBoV1血清型壳体。在一些实施方案中，AAV血清型是AAV1、AAV2、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10或AAVrh10。在上述方法的一些实施方案中，AAV ITR是AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10、AAVrh10、AAV11、AAV12、AAV2R471A、AAV DJ、山羊AAV、牛AAV或小鼠AAV血清型ITR。在一些实施方案中，AAV ITR是AAV2ITR。在一些实施方案中，rAAV颗粒的ITR和壳体源自相同AAV血清型。在一些实施方案中，ITR和所述壳体源自AAV2。在其它实施方案中，rAAV颗粒的ITR和壳体源自不同AAV血清型。在一些实施方案中，AAV颗粒包含AAV2ITR和AAVrh8R壳体。在一些实施方案中，AAV颗粒包含AAV2ITR和AAV8壳体。

[0018] 在上述组合物的一些实施方案中，在生产细胞中产生包含超大AAV基因组的AAV颗粒。在一些实施方案中，生产细胞系源自灵长类细胞。在一些实施方案中，生产细胞系源自HeLa、293、A549或Perc.6细胞。在一些实施方案中，生产细胞系适合于在悬浮液中生长。在一些实施方案中，AAV辅助功能由腺病毒或HSV提供。在一些实施方案中，在提供辅助功能后约48小时-约96小时收集所述rAAV颗粒。

[0019] 在一些态样中，本发明提供了用于增强超大rAAV基因组的表达的方法，所述方法包括通过对生产细胞系提供AAV辅助功能而在所述细胞系中产生rAAV颗粒，其中所述生产细胞系包含：a) 编码AAV rep和cap基因的核酸，和b) rAAV基因组，其中所述rAAV基因组大于约4.7kb。在一些实施方案中，超大rAAV基因组的表达是来自通过瞬时转染产生的rAAV颗粒的所述超大rAAV基因组的表达的大约1.25倍、约1.5倍、约1.75倍、约2.0倍、约2.5倍、约2.75倍、约3倍或约5倍。在一些实施方案中，与来自通过瞬时转染产生的rAAV颗粒的超大rAAV基因组的表达动力学相比，来自生产细胞系产生的rAAV颗粒的所述超大rAAV基因组的表达动力学是更快的表达动力学。在一些实施方案中，由生产细胞系产生的超大rAAV基因组的表达动力学比来自通过瞬时转染产生的rAAV颗粒的超大rAAV基因组的表达动力学快约5%、快约10%、快约25%、快约50%、快约75%或快约90%。

[0020] 在超大rAAV基因组的增强表达的一些实施方案中，所述编码AAV rep和cap基因的

核酸和/或所述rAAV基因组在所述生产细胞系中稳定维持。在一些实施方案中,所述编码AAV rep和cap基因的核酸和/或所述rAAV基因组稳定整合入所述生产细胞系的基因组中。在一些实施方案中,所述rAAV基因组包含一个或多个AAV反向末端重复(ITR)和异源转基因。在一些实施方案中,所述rAAV基因组包含两个AAV ITR。在一些实施方案中,rAAV基因组是约4.7kb-约9.4kb。在一些实施方案中,rAAV基因组是约4.7kb-约5kb、约4.7kb-约6kb、约4.7kb-约7kb、约4.7kb-约8kb或约4.7kb-约9kb。在一些实施方案中,rAAV基因组是约4.7kb-6.7kb或约5.2kb-约8.7kb。

[0021] 在超大rAAV基因组的增强表达的一些实施方案中,异源转基因编码治疗性转基因产物。在一些实施方案中,异源转基因编码因子VIII、抗肌萎缩蛋白、dysferlin或囊性纤维化跨膜传导调节蛋白(CFTR)。在一些实施方案中,异源转基因是人转基因。在一些实施方案中,异源转基因与启动子可操作连接。在一些实施方案中,启动子是小鼠转甲状腺素(mTTR)启动子。在一些实施方案中,rAAV基因组包含内含子。在一些实施方案中,内含子是合成内含子。在一些实施方案中,rAAV基因组包含多聚腺苷酸化信号。在一些实施方案中,多聚腺苷酸化信号是合成多聚腺苷酸化信号或牛生长激素多聚腺苷酸化信号。

[0022] 在超大rAAV基因组的增强表达的一些实施方案中,rAAV颗粒包含AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10、AAVrh10、AAV11、AAV12、AAV2R471A、AAV2/2-7m8、AAV DJ、AAV2N587A、AAV2E548A、AAV2N708A、AAV V708K、山羊AAV、AAV1/AAV2嵌合、牛AAV或小鼠AAV壳体rAAV2/HBoV1血清型壳体。在一些实施方案中,AAV血清型是AAV1、AAV2、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10或AAVrh10。在一些实施方案中,AAV ITR是AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10、AAVrh10、AAV11、AAV12、AAV2R471A、AAV DJ、山羊AAV、牛AAV或小鼠AAV血清型ITR。在一些实施方案中,AAV ITR是AAV2ITR。在一些实施方案中,rAAV颗粒的ITR和壳体源自相同AAV血清型。在一些实施方案中,ITR和所述壳体源自AAV2。在一些实施方案中,rAAV颗粒的ITR和壳体源自不同AAV血清型。在一些实施方案中,AAV颗粒包含AAV2ITR和AAVrh8R壳体。在一些实施方案中,AAV颗粒包含AAV2ITR和AAV8壳体。

[0023] 在超大rAAV基因组的增强表达的一些实施方案中,生产细胞系源自灵长类细胞。在一些实施方案中,生产细胞系源自HeLa、293、A549或Perc.6细胞。在一些实施方案中,生产细胞系适合于在悬浮液中生长。在一些实施方案中,AAV辅助功能由腺病毒、HSV或杆状病毒提供。在一些实施方案中,在提供辅助功能后约48小时-约96小时收集所述rAAV颗粒。在一些实施方案中,方法进一步包括纯化rAAV颗粒。在一些实施方案中,纯化包括一个或多个层析步骤。

[0024] 在一些态样中,本发明提供了用于产生包含超大重组AAV基因组的腺相关病毒(AAV)颗粒的细胞系,所述细胞系包含:a)编码AAV rep和cap基因的核酸,和b) rAAV基因组,其中所述rAAV基因组大于约4.7kb。在一些实施方案中,所述编码AAV rep和cap基因的核酸和/或所述rAAV基因组在所述生产细胞系中稳定维持。在一些实施方案中,所述编码AAV rep和cap基因的核酸和/或所述rAAV基因组稳定整合入所述生产细胞系的基因组中。在一些实施方案中,所述rAAV基因组包含一个或多个AAV反向末端重复(ITR)和异源转基因。在一些实施方案中,rAAV基因组是约4.7kb-约9.4kb。在一些实施方案中,rAAV基因组是约4.7kb-约5kb、约4.7kb-约6kb、约4.7kb-约7kb、约4.7kb-约8kb或约4.7kb-约9kb。在一些实

施方案中,rAAV基因组是约4.7kb-6.7kb或约5.2kb-约8.7kb。在一些实施方案中,rAAV基因组的长度大于约5.0kb、5.1kb、5.2kb、5.3kb、5.4kb、5.5kb、5.6kb、5.7kb、5.8kb、5.9kb、6.0kb、6.1kb、6.2kb、6.3kb、6.4kb、6.5kb、6.6kb、6.7kb、6.8kb、6.9kb、7.0kb、8.0kb、8.7kb或9.0kb中任一者或其间的任何数值。

[0025] 在细胞系的一些实施方案中,异源转基因编码治疗性转基因产物。在一些实施方案中,异源转基因是人转基因。在一些实施方案中,异源转基因编码因子VIII、抗肌萎缩蛋白、dysferlin或囊性纤维化跨膜传导调节蛋白(CFTR)。在一些实施方案中,异源转基因与启动子可操作连接。在别的实施方案中,启动子是小鼠转甲状腺素(mTTR)启动子。在一些实施方案中,rAAV基因组包含内含子。在别的实施方案中,内含子是合成内含子。在一些实施方案中,rAAV基因组包含多聚腺苷酸化信号。在别的实施方案中,多聚腺苷酸化信号是合成多聚腺苷酸化信号或牛生长激素多聚腺苷酸化信号。

[0026] 在上述细胞系的一些实施方案中,rAAV颗粒包含AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10、AAVrh10、AAV11、AAV12、AAV2R471A、AAV2/2-7m8、AAV DJ、AAV2N587A、AAV2E548A、AAV2N708A、AAV V708K、山羊AAV、AAV1/AAV2嵌合、牛AAV或小鼠AAV壳体rAAV2/HBoV1血清型壳体。在一些实施方案中,AAV血清型是AAV1、AAV2、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10或AAVrh10。在上述方法的一些实施方案中,AAV ITR是AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10、AAVrh10、AAV11、AAV12、AAV2R471A、AAV DJ、山羊AAV、牛AAV或小鼠AAV血清型ITR。在一些实施方案中,AAV ITR是AAV2ITR。在一些实施方案中,rAAV颗粒的ITR和壳体源自相同AAV血清型。在一些实施方案中,ITR和所述壳体源自AAV2。在其它实施方案中,rAAV颗粒的ITR和壳体源自不同AAV血清型。在一些实施方案中,AAV颗粒包含AAV2ITR和AAVrh8R壳体。在一些实施方案中,AAV颗粒包含AAV2ITR和AAV8壳体。

[0027] 在上述细胞系的一些实施方案中,生产细胞系源自灵长类细胞。在一些实施方案中,生产细胞系源自HeLa、293、A549或Perc.6细胞。在一些实施方案中,生产细胞系适合于在悬浮液中生长。在一些实施方案中,通过提供AAV辅助功能而在细胞系中生产AAV颗粒。在一些实施方案中,AAV辅助功能由腺病毒或HSV提供。在一些实施方案中,在提供辅助功能后约48小时-约96小时收集所述rAAV颗粒。

[0028] 在一些态样中,本发明提供了包含由AAV壳体壳体化的rAAV基因组的腺相关病毒(AAV)颗粒,其中rAAV基因组大于约4.7kb。在一些实施方案中,所述rAAV基因组包含一个或多个AAV反向末端重复(ITR)和异源转基因。在一些实施方案中,所述rAAV基因组包含两个AAV ITR。在一些实施方案中,rAAV基因组是约4.7kb-约9.4kb。在一些实施方案中,rAAV基因组是约4.7kb-约5kb、约4.7kb-约6kb、约4.7kb-约7kb、约4.7kb-约8kb或约4.7kb-约9kb。在一些实施方案中,rAAV基因组是约4.7kb-6.7kb或约5.2kb-约8.7kb。在一些实施方案中,rAAV基因组的长度大于约5.0kb、5.1kb、5.2kb、5.3kb、5.4kb、5.5kb、5.6kb、5.7kb、5.8kb、5.9kb、6.0kb、6.1kb、6.2kb、6.3kb、6.4kb、6.5kb、6.6kb、6.7kb、6.8kb、6.9kb、7.0kb、8.0kb或9.0kb中任一者或其间的任何数值。

[0029] 在一些实施方案中,本发明提供了包含超大rAAV基因组的AAV颗粒,其中异源转基因编码治疗性转基因产物。在一些实施方案中,异源转基因是人转基因。在一些实施方案中,异源转基因编码因子VIII、抗肌萎缩蛋白、dysferlin或囊性纤维化跨膜传导调节蛋白

(CFTR)。在一些实施方案中,异源转基因与启动子可操作连接。在别的实施方案中,启动子是小鼠转甲状腺素(mTTR)启动子。在一些实施方案中,rAAV基因组包含内含子。在别的实施方案中,内含子是合成内含子。在一些实施方案中,rAAV基因组包含多聚腺苷酸化信号。在别的实施方案中,多聚腺苷酸化信号是合成多聚腺苷酸化信号或牛生长激素多聚腺苷酸化信号。

[0030] 在上述细胞系的一些实施方案中,rAAV颗粒包含AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10、AAVrh10、AAV11、AAV12、AAV2R471A、AAV2/2-7m8、AAV DJ、AAV2N587A、AAV2E548A、AAV2N708A、AAV V708K、山羊AAV、AAV1/AAV2嵌合、牛AAV或小鼠AAV壳体rAAV2/HBoV1血清型壳体。在一些实施方案中,AAV血清型是AAV1、AAV2、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10或AAVrh10。在上述方法的一些实施方案中,AAV ITR是AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10、AAVrh10、AAV11、AAV12、AAV2R471A、AAV DJ、山羊AAV、牛AAV或小鼠AAV血清型ITR。在一些实施方案中,AAV ITR是AAV2ITR。在一些实施方案中,rAAV颗粒的ITR和壳体源自相同AAV血清型。在一些实施方案中,ITR和所述壳体源自AAV2。在其它实施方案中,rAAV颗粒的ITR和壳体源自不同AAV血清型。在一些实施方案中,AAV颗粒包含AAV2ITR和AAVrh8R壳体。在一些实施方案中,AAV颗粒包含AAV2ITR和AAV8壳体。

[0031] 在本发明的一些实施方案中,在生产细胞中产生包含超大AAV基因组的AAV颗粒。在一些实施方案中,生产细胞系源自灵长类细胞。在一些实施方案中,生产细胞系源自HeLa、293、A549或Perc.6细胞。在一些实施方案中,生产细胞系适合于在悬浮液中生长。在一些实施方案中,AAV辅助功能由腺病毒或HSV提供。在一些实施方案中,在提供辅助功能后约48小时-约96小时收集所述rAAV颗粒。

[0032] 在一些实施方案中,本发明提供了包含超大rAAV基因组的AAV颗粒,其中rAAV基因组从5'至3'包含AAV2ITR、mTTR启动子、合成内含子、编码人FVIII的转基因、合成多聚腺苷酸化序列、和AAV2ITR。在一些实施方案中,rAAV基因组从5'至3'包含AAV2ITR、mTTR启动子、合成内含子、编码人FVIII的转基因、牛生长激素合成多聚腺苷酸化序列、和AAV2ITR。在一些实施方案中,FVIII包含全部或部分的B域的缺失。在一些实施方案中,AAV颗粒包含AAVrh8R壳体。在一些实施方案中,AAV颗粒包含AAV8壳体。

[0033] 在一些态样中,本发明提供了包含rAAV基因组的rAAV载体,其中rAAV基因组从5'至3'包含AAV2ITR、mTTR启动子、合成内含子、编码人FVIII的转基因、合成多聚腺苷酸化序列、和AAV2ITR。在一些实施方案中,rAAV基因组从5'至3'包含AAV2ITR、mTTR启动子、合成内含子、编码人FVIII的转基因、牛生长激素合成多聚腺苷酸化序列、和AAV2ITR。在一些实施方案中,FVIII包含全部或部分的B域的缺失。

[0034] 在一些实施方案中,本发明提供了治疗具有疾病或病症的个体的方法,其包括对个体授予包含超大rAAV基因组的AAV颗粒,所述超大rAAV基因组编码治疗性转基因,其中治疗性转基因适合于治疗疾病或病症。在一些实施方案中,个体是哺乳动物(例如人)。在一些实施方案中,疾病或病症是血友病A。在一些实施方案中,治疗性转基因编码因子VIII;例如人因子VIII,包括B域缺失的人因子VIII。

[0035] 在一些实施方案中,本发明提供了包含如本文中描述的AAV颗粒的套组,所述AAV颗粒包含超大rAAV基因组。

[0036] 本文中引用的所有参考文献,包括专利申请和公开文本通过提述以其整体并入。

[0037] 附图简述

[0038] 图1A显示了基于小鼠转甲状腺素(mTTR)启动子、范围为5.1至5.4kb载体基因组大小(如指示)的hFVIII表达盒的图。显示了HNF4结合位点(空心圆圈)和HNF3结合位点(实心圆圈)中的序列修饰及其位置。图1A、1B和1C的缩写:ITR,rAAV反向末端重复;mTTR,小鼠转甲状腺素启动子(202或482bp);HI,杂合内含子;FVIII,B域缺失的人FVIII cDNA;syn pA,合成的(syn pA);BGH或牛生长激素(BGH)多聚A(pA)。

[0039] 图1B显示了本文中描述的实验中使用的mTTR启动子序列的比对。

[0040] 图1C:在体内来自mTTR-FVIII表达盒的FVIII水平。通过高体积注射将质粒载体静脉内注射入C56BL/6小鼠中,并且通过ELISA测定法测量血浆中的因子VIII水平。

[0041] 图1D显示了5.1kb FVIII表达盒的结构。此盒包含rAAV反向末端重复(ITR)、小鼠转甲状腺素(mTTR)启动子、杂合内含子(HI)、B域缺失的人FVIII cDNA、和合成的多聚A序列。

[0042] 图1E显示了TriplePlay质粒,其含有FVIII载体基因组、AAV rep和cap基因、以及负责嘌呤霉素和卡那霉素药物抗性的基因。

[0043] 图2A和2B显示了来自选定的主孔克隆(masterwell clone,MW)的基因组DNA的Southern印迹分析。(图2A)用SpeI切割FVIII TriplePlay质粒,这产生13kb线性片段。这用作单位长度TriplePlay质粒和整合拷贝数目标标准品的大小对照。(图2B)通过用BglII和HincII消化分析整合的载体基因组的完整性。这些酶在FVIII表达盒内切割,产生1.8和2.8kb片段。在这两幅图中,提供了显示载体和限制性位点的图。

[0044] 图3A和3B显示了对AAVrh8R/5.1kb mTTR-FVIII载体生产产率和稳定性的分析。(图3A)AAVrh8R/5.1kb载体产生的时间过程。用野生型腺病毒(wt Ad)感染摇动培养物,在第2天、第3天和第4天收集样品,并且以每ml的载体基因组(VG/ml)通过qPCR定量载体产率。(图3B)选定的高生产主孔的稳定性。对MW#287(AAV8/5.1kb)、MW#35(AAVrh8R/5.1kb)和MW#163(AAVrh8R/5.4kb)显示了rAAV载体产生水平。传代主孔直至第20代或第26代,并且通过qPCR定量rAAV生产率(productivity)(DRP/ml)。

[0045] 图4A和4B显示了对超大5.1kb rAAVrh8R/FVIII载体质量的分析。通过AUC分析比较PCL和TXN产生的5.1kb载体批次。使用MW#35将AAVrh8R/5.1kb FVIII产生三次(图4A),并且与通过TXN方法产生的相同载体比较(图4B)。通过测量病毒质量差异的分析超速离心分析(AUC)评估载体的质量。插页指示具有不同沉积(S)数值的壳体%。空壳体通常具有63至66的S值,而具有野生型大小载体基因组的壳体通常为100至103的S。

[0046] 图5A和5B显示了对超大5.4kb rAAVrh8R/FVIII载体质量的分析。通过AUC分析比较PCL(图5A)和TXN(图5B)产生的5.4kb载体批次。插页指示不同沉积(S)数值的壳体%。圈出空壳体(64S/63S)和具有较大基因组的颗粒(101S/99S)的百分比。

[0047] 图6A和6B显示了通过Southern印迹表征PCL或TXN产生的rAAVrh8R/5.1kb载体中的包装的载体基因组。从纯化的病毒体中分离载体基因组,并且通过碱性凝胶电泳,接着使用对载体特异性的探针进行Southern印迹分析大小。(图6A)使用4.0kb FVIII探针的Southern分析(FVIII域A1、A2、A3和C1)。以 1.1 和 6.0×10^9 VG/道上样VG,并且在1%碱性凝胶上分开。将通过PCL(MW#35)或三重转染产生的5.1kb FVIII载体与4.6kb大小载体(与

rh8R/5.1kb载体相同,只是缺失C1域以创建正常大小的载体)比较。(图6B)将每个独特VG大小的信号强度定量,并且以每道中的总信号的%绘图。

[0048] 图7A、7B和7C显示了通过DNA斑点印迹分析表征PCL或TXN产生的rAAVrh8R/5.1kb载体中的包装的载体基因组的5'端。通过将每种载体的2倍连续稀释物应用到膜上(以 2.4×10^9 开始;总共8个降低的载体浓度加上作为阴性对照应用的无基因组)分析图5中使用的载体批次。使每个印迹与对载体基因组(正或负极性)的中间或5'末端特异性的3'端标记的寡核苷酸探针杂交。定量信号强度,并且相对于4.6kb载体(完全包装)标准化。使用三个浓度来产生标准误差。(图7A)显示了使用的寡核苷酸探针的位置的图。数值指示相应的3'端的核苷酸距离。(图7B)负链的分析。(图7C)正链的分析。

[0049] 图8A、8B、8C和8D显示了PCL或TXN产生的rAAVrh8R/5.1kb载体中的包装载体基因组的5'和3'端的表征。(图8A)图显示了针对使用的载体基因组的正和负链的5'和3'端的寡核苷酸探针的位置。(图8B)每个批次中的5.1kb载体基因组的正和负链的定量。分析的载体包含4.6kb或5.1kb mTTR-FVIII基因组。指示载体产生方法(PCL或TXN)。以相似的方式纯化所有载体。通过将每种载体的2倍连续稀释物应用到膜上(以 3.0×10^9 开始;总共8个降低的载体浓度)进行分析,如图5中描述。使用含有FIX(阴性对照)或FVIII(阳性对照)cDNA的质粒作为信号特异性的对照。(图8C)使用针对5.1kb载体的3'和5'端寡核苷酸探针的Southern分析。以 1.5 和 7.5×10^9 VG/道上样VG,并且在1%碱性凝胶上分开。将通过PCL(MW#35)或三重转染(TXN)产生的5.1kb FVIII载体与4.6kb大小载体比较。显示了大小标志物(2.7、4.7和5.1kb)。顶部小图,正链分析;底部小图,负链分析。显示了用于每道的寡核苷酸。白色箭指示了丢失的信号。(图8D)每个载体中的基因组大小的定量。通过ImageJ定量用3'端寡核苷酸探针(检测所有包装的基因组)探查的小图中的信号强度。将每个独特VG大小(>4.7kb、4.7kb和<4.7kb)的强度定量,并且以每道中的总信号的%绘图。

[0050] 图9A、9B和9C显示了表征PCL或TXN产生的rAAVrh8R/5.4kb载体中的包装载体基因组的5'和3'端。(图9A)显示了针对使用的载体基因组的正和负链的5'和3'端的寡核苷酸探针的位置的图。(图9B)通过斑点印迹分析定量每个批次中的5.4kb载体基因组的正和负链。如图8中描述进行分析。(图9C)使用3'和5'端寡核苷酸探针的5.4kb载体的Southern分析。如图8中描述进行实验。

[0051] 图10A和10B显示了PCL产生的rAAVrh8R/5.1kb载体在血友病A KO小鼠中的体内效力。通过尾静脉以 3×10^{11} 和 4×10^{10} DRP/小鼠对小鼠投予载体,并且分析血浆FVIII水平直至第56天。(图10A)血浆FVIII蛋白活性。通过Coatest测定法在第7天、第14天、第28天、第42天和第56天血浆样品中测量活性。(图10B)第28天和第56天的凝固时间。通过激活部分促凝血酶原激酶时间(aPTT)分析凝固时间。每个处理组含有n=7-10只小鼠/组。通过Student t检验指示统计学显著性如下:*, $p < 0.05$;**, $p < 0.01$ ***, $p < 0.001$ 。

[0052] 图11A、11B和11C显示了使用血友病A KO小鼠在体内比较PCL和TXN产生的5.1kb AAVrh8R/FVIII载体。通过尾静脉以 4×10^{10} DRP/小鼠对小鼠投予载体。(图11A)血浆FVIII蛋白活性。通过Coatest测定法在第21天、第35天、第56天、第70天和第84天样品中测量活性。(图11B)第21天的血浆凝固时间。(图11C)第56天的血浆凝固时间。通过aPTT测定法测量血浆凝固时间。显示了小鼠品系(129S和BALB/c)的凝固时间以进行比较。(图11D)第84天的肝中的载体基因组(VG)拷贝。通过qPCR定量VG拷贝,并且以拷贝/500ng总肝DNA显示。每个处

理组含有 $n=8$ 只小鼠/组。通过Student t检验指示统计学显著性如下：*, $p<0.05$ ；**, $p<0.01$ ，***, $p<0.001$ 。在每幅小图中指示病毒产生的方法(PCL或TXN)。

[0053] 图12A、12B和12C显示了使用血友病A KO小鼠在体内比较PCL和TXN产生的5.4kb AAVrh8R/FVIII载体。通过尾静脉以 4×10^{10} DRP/小鼠对小鼠投予载体,并且在载体投予后第24天和第43天收集血浆样品。(图12A) 血浆FVIII活性。通过Coatest测定法在第24天和第43天血浆样品中测量活性。(图12B) 通过aPTT测定法得到的第24天血浆凝固时间。(图12C) 第3天和第43天在肝中的载体基因组(VG)拷贝。在载体投予后3和43天处死动物,并且通过qPCR定量VG拷贝,并且以拷贝/500ng总肝DNA显示。每个处理组含有 $n=6-8$ 只小鼠/组。通过Student t检验指示统计学显著性如下：*, $p<0.05$ ；**, $p<0.01$ ，***, $p<0.001$ 。

[0054] 图13A显示了5.1、5.9和6.7kb AAV2/SEAP载体的图。图13B显示了就相对和比生产(specific production) ($n=2$) 测定法中的载体产率而言来自各个主孔(MW)的数据。载体产率以DRP/ml指示。

[0055] 发明详述

[0056] 如本文中详细讨论,发明人已经开发出能够产生较高产率的较高质量的超大重组腺相关病毒(rAAV)载体的生产细胞系平台。已经使用含有人因子VIII cDNA的rAAV载体作为例示性构建体表征了此平台。与使用标准三重转染方法的产量相比,此平台产生具有更高量的壳体化的较大基因组的rAAV载体。使用此平台产生的rAAV载体在体内基因转移方面也有能力,并且导致功能性因子VIII的产生。

[0057] 因而,本发明提供了用于产生含有超大重组AAV基因组的腺相关病毒(AAV)颗粒的方法。在一些实施方案中,方法包括在产生rAAV颗粒的条件下培养AAV生产细胞系,其中所述AAV生产细胞系含有:i) 编码AAV rep和cap基因的核酸,和ii) rAAV基因组,其中所述rAAV基因组是约4.7kb-约9.4kb,任选地约4.7kb-6.7kb;b) 提供AAV辅助功能;并c) 收集含有超大rAAV基因组的rAAV颗粒。在一些实施方案中,rAAV基因组大于约5kb。在一些实施方案中,rAAV基因组的长度大于约5.0kb、5.1kb、5.2kb、5.3kb、5.4kb、5.5kb、5.6kb、5.7kb、5.8kb、5.9kb、6.0kb、6.1kb、6.2kb、6.3kb、6.4kb、6.5kb、6.6kb、6.7kb、6.8kb、6.9kb、7.0kb、8.0kb或9.0kb中任一者或其间的任何数值。本文中进一步提供了通过本公开内容的方法产生的含有超大重组AAV基因组的rAAV颗粒。

[0058] 本文中还提供进一步提供了包含rAAV颗粒的组合物,其中至少约15%、至少约20%、至少约25%、至少约30%、至少约35%、至少约40%、至少约45%、至少约50%、至少约55%、至少约60%或至少约70% rAAV颗粒壳体化大于约5kb的rAAV基因组。

[0059] 本文中还提供进一步提供了用于产生含有超大重组AAV基因组的腺相关病毒(AAV)颗粒的细胞系,所述细胞系包含a) 编码AAV rep和cap基因的核酸,和b) rAAV基因组,其中rAAV基因组是约4.7kb-约9.4kb,任选地约4.7kb-6.7kb。在一些实施方案中,rAAV基因组大于约5kb。

[0060] 本文中还提供进一步提供了含有通过AAV壳体壳体化的rAAV基因组的腺相关病毒(AAV)颗粒,其中rAAV基因组是约4.7kb-约9.4kb,任选地约4.7kb-6.7kb。在一些实施方案中,rAAV基因组大于约5kb。

[0061] I.通用技术

[0062] 本领域技术人员使用常规方法,如例如下列各项中描述的广泛利用的方法一般完

全理解并且通常采用本文中描述或引用的技术和规程:Molecular Cloning:A Laboratory Manual (Sambrook等,第4版,Cold Spring Harbor Laboratory Press,Cold Spring Harbor,N.Y.,2012);Current Protocols in Molecular Biology (F.M.Ausubel,等编,2003);系列Methods in Enzymology (Academic Press,Inc.);PCR 2:A Practical Approach (M.J.MacPherson,B.D.Hames和G.R.Taylor编,1995);Antibodies,A Laboratory Manual (Harlow和Lane,编,1988);Culture of Animal Cells:A Manual of Basic Technique and Specialized Applications (R.I.Freshney,第6版,J.Wiley and Sons,2010);Oligonucleotide Synthesis (M.J.Gait编,1984);Methods in Molecular Biology, Humana Press;Cell Biology:A Laboratory Notebook (J.E.Cellis编,Academic Press,1998);Introduction to Cell and Tissue Culture (J.P.Mather和P.E.Roberts, Plenum Press,1998);Cell and Tissue Culture:Laboratory Procedures (A.Doyle, J.B.Griffiths,和D.G.Newell,编,J.Wiley and Sons,1993-8);Handbook of Experimental Immunology (D.M.Weir和C.C.Blackwell,编,1996);Gene Transfer Vectors for Mammalian Cells (J.M.Miller和M.P.Calos,编,1987);PCR:The Polymerase Chain Reaction, (Mullis等编,1994);Current Protocols in Immunology (J.E.Coligan等编,1991);Short Protocols in Molecular Biology (Ausubel等编,J.Wiley and Sons,2002);Immunobiology (C.A.Janeway等,2004);Antibodies (P.Finch,1997);Antibodies:A Practical Approach (D.Catty编,IRL Press,1988-1989);Monoclonal Antibodies:A Practical Approach (P.Shepherd和C.Dean,编,Oxford University Press,2000);Using Antibodies:A Laboratory Manual (E.Harlow和D.Lane,Cold Spring Harbor Laboratory Press,1999);The Antibodies (M.Zanetti和J.D.Capra,编,Harwood Academic Publishers,1995);及Cancer:Principles and Practice of Oncology (V.T.DeVita等编,J.B.Lippincott Company,2011)。

[0063] II.定义

[0064] 如本文中使用的,“载体”指包含要在体外或在体内投递到宿主细胞中的核酸的重组质粒或病毒。

[0065] 如本文中使用的,术语“多核苷酸”或“核酸”指任何长度的核苷酸(核糖核苷酸或脱氧核糖核苷酸)的聚合形式。如此,此术语包括但不限于单链、双链或多链DNA或RNA、基因组DNA、cDNA、DNA-RNA杂合物或包含嘌呤和嘧啶碱基或其它天然的、化学或生物化学修饰的、非天然的或衍生化的核苷酸碱基的聚合物。多核苷酸的主链可以包含糖和磷酸基团(其通常可以在RNA或DNA中找到),或者经修饰的或取代的糖或磷酸基团。或者,多核苷酸的主链可以包含合成亚基如氨基磷酸酯的聚合物,并且如此可以是寡脱氧核苷氨基磷酸酯(P-NH₂)或混合的氨基磷酸酯-磷酸二酯寡聚物。另外,可以通过合成互补链并且在合适的条件下使链退火,或者通过用合适的引物使用DNA聚合酶从头合成互补链从化学合成的单链多核苷酸产物获得双链多核苷酸。

[0066] 术语“多肽”和“蛋白质”可互换使用,指氨基酸残基的聚合物,并且不限于最小长度。氨基酸残基的此类聚合物可以含有天然的或非天然的氨基酸残基,并且包括但不限于肽、寡肽、氨基酸残基的二聚体、三聚体、和多聚体。该定义涵盖全长蛋白质和其片段两者。术语还包括多肽的表达后修饰,例如糖基化、唾液酸化、乙酰化、磷酸化等。此外,出于本发

明的目的,“多肽”指包含对天然序列的修饰,如缺失、添加和取代(一般实质上保守)的蛋白质,只要蛋白质维持期望的活性。这些修饰可以是有意的,如经由定点诱变,或者可以是意外的,如经由生成蛋白质的宿主的突变或者由于PCR扩增所致的错误。

[0067] “重组病毒载体”指包含一种或多种异源序列(即不是病毒起源的核酸序列)的重组多核苷酸载体。在重组AAV载体的情况下,重组核酸侧翼有至少一个反向末端重复序列(ITR)。在一些实施方案中,重组核酸侧翼有两个ITR。

[0068] “重组AAV载体(rAAV载体)”指包含一种或多种异源序列(即不是AAV起源的核酸序列)的多核苷酸载体,所述异源序列侧翼有至少一个或两个AAV反向末端重复序列(ITR)。当存在于宿主细胞中时,此类rAAV载体可以被复制并包装成感染性病毒颗粒,所述宿主细胞已经感染有合适的辅助病毒(或者在表达合适的辅助功能),并且表达AAV rep和cap基因产物(即AAV Rep和Cap蛋白)。当rAAV载体并入较大的多核苷酸中(例如在染色体中或者在另一种载体如用于克隆或转染的质粒中)时,则rAAV载体可以称为“原载体(pro-vector)”,其可以通过在存在AAV包装功能和合适的辅助功能的情况下复制和壳体化来“挽救”。rAAV载体可以为许多形式之任一种,包括但不限于质粒、线性人工染色体、与脂质复合、在脂质体内包囊、和在病毒颗粒(例如AAV颗粒)中壳体化。可以将rAAV载体包装成AAV病毒壳体以产生“重组腺相关病毒颗粒(rAAV颗粒)”。

[0069] 如本文中使用的,“生产细胞系”是能够产生AAV颗粒的稳定细胞系。在一些实施方案中,AAV复制和/或壳体基因在宿主细胞系中稳定维持。在一些实施方案中,包含一种或多种AAV ITR和异源核酸(例如异源转基因)的AAV载体基因组在宿主细胞系中稳定维持。在一些实施方案中,AAV复制和/或壳体基因和包含一种或多种AAV ITR和异源核酸(例如异源转基因)的AAV载体基因组在宿主细胞系中稳定维持。在一些实施方案中,AAV复制基因、壳体基因或包含一种或多种AAV ITR的AAV载体基因组中的一种或多种稳定整合到宿主细胞系的基因组中。本领域技术人员会理解稳定维持的核酸在多次传代(例如5、10、15、25或更多次传代)后在宿主细胞系中维持。

[0070] “异源”意指源自与同它比较并且将其导入或整合的实体的剩余部分在基因型上截然不同的实体。例如,通过遗传工程技术导入不同细胞类型中的多核苷酸是异源多核苷酸(并且在表达时可以编码异源多肽)。类似地,并入病毒载体中的细胞序列(例如基因或其部分)相对于载体而言是异源核苷酸序列。

[0071] 术语“转基因”指导入细胞中并且能够被转录成RNA,并且任选地在合适的条件下翻译和/或表达的多核苷酸。在各态样中,它对接受其导入的细胞赋予期望的特性,或者在其它情况下导致期望的治疗或诊断结果。在另一个态样中,它可以被转录成介导RNA干扰的分子,如miRNA、siRNA或shRNA。

[0072] 术语“转甲状腺素(TTR)启动子”指能够驱动源自转甲状腺素基因的基因表达的多核苷酸序列。在一些实施方案中,转甲状腺素蛋白启动子可以来自小鼠转甲状腺素(mTTR)基因(例如小家鼠(*Mus musculus*)转甲状腺素,如以GenBank Entrez Gene ID 22139表示)。图1B中呈现了TTR启动子的例子。

[0073] 如提及病毒滴度使用,术语“基因组颗粒(gp)”、“基因组等同物”或“基因组拷贝”指含有重组AAV DNA基因组的病毒体的数目,不管感染性或功能性如何。特定载体制备物中的基因组颗粒的数目可以通过方法,如记载于本文中的实施例,或者例如Clark等(1999)

Hum. Gene Ther., 10:1031-1039; Veldwijk等 (2002) Mol. Ther., 6:272-278中的方法测量。

[0074] 如本文中使用的术语“载体基因组 (vg)”可以指构成载体 (例如病毒载体) 的多核苷酸序列组的一种或多种多核苷酸。可以在病毒颗粒中壳体化载体基因组。取决于具体的病毒载体, 载体基因组可以包含单链DNA、双链DNA或单链RNA或双链RNA。载体基因组可以包含与特定病毒载体有关的内源序列和/或经由重组技术插入特定病毒载体中的任何异源序列。例如, 重组AAV载体基因组可以包含至少一个在启动子侧翼的ITR序列、感兴趣的序列 (例如异源转基因)、任选地内含子、和多聚腺苷酸化序列。完整的载体基因组可以包含载体的多核苷酸序列的完整套组。在一些实施方案中, 病毒载体的核酸滴度可以按照vg/mL测量。适合于测量此滴度的方法是本领域中已知的 (例如定量PCR)。

[0075] 术语“超大重组AAV基因组”可以指具有大于AAV基因组的常规包装限度的大小 (如以核苷酸碱基对测量) 的重组AAV基因组, 所述常规包装限度在本领域中已经限定为4.7至4.8kb (参见例如Dong, J-Y等 (1996) Human Gene Therapy 7:2101-2112)。在一些实施方案中, 超大重组AAV基因组大于约4.7kb。在一些实施方案中, 超大重组AAV基因组大于约5kb。在一些实施方案中, 超大重组AAV基因组是约4.7kb-约9.4kb, 任选地约4.7kb-6.7kb。

[0076] 如提及病毒滴度使用, 术语“感染单位 (iu)”、“感染性颗粒”或“复制单位”指感染性并且有复制能力的重组AAV载体颗粒的数目, 如通过如记载于例如McLaughlin等 (1988) J. Virol., 62:1963-1973的感染性中心测定法 (又称为复制中心测定法) 测量。

[0077] 如提及病毒滴度使用, 术语“转导单位 (tu)”指导致功能性转基因产物的生成的感染性重组AAV载体颗粒的数目, 如在功能性测定法, 如记载于本文中的实施例, 或例如Xiao等 (1997) Exp. Neurobiol., 144:113-124; 或Fisher等 (1996) J. Virol., 70:520-532 (LFU测定法) 中的功能性测定法中测量。

[0078] “反向末端重复”或“ITR”序列是本领域中完全理解的术语, 并且指以相反取向在病毒基因组的末端找到的相对较短的序列。

[0079] “AAV反向末端重复 (ITR)”序列, 本领域中完全理解的术语, 是存在于天然单链AAV基因组的两个末端的约145个核苷酸的序列。ITR的最外面的125个核苷酸可以以两个备选取向之任一存在, 导致不同AAV基因组之间和单一AAV基因组的两个末端之间的异质性。最外面的125个核苷酸也含有自身互补的几个较短的区域 (称为A、A'、B、B'、C、C' 和D区), 容许在ITR的此部分内发生链内碱基配对。

[0080] “末端解析序列 (terminal resolution sequence)”或“trs”是AAV ITR的D区中在病毒DNA复制过程中被AAV rep蛋白切割的序列。突变体末端解析序列难以通过AAV rep蛋白切割。

[0081] “AAV辅助功能”指让AAV被宿主细胞复制并包装的功能。AAV辅助功能可以以许多形式之任一种提供, 包括但不限于辅助病毒或辅助病毒基因, 所述辅助病毒基因帮助AAV复制和包装。其它AAV辅助功能是本领域中已知的, 如遗传毒性剂 (genotoxic agent)。

[0082] 用于AAV的“辅助病毒”指让AAV (其是缺陷型细小病毒 (parvovirus)) 被宿主细胞复制并包装的病毒。已经鉴定出许多此类辅助病毒, 包括腺病毒、疱疹病毒、细小病毒如痘病毒和杆状病毒。腺病毒涵盖许多不同亚组, 尽管最常使用C亚组的腺病毒5型 (Ad5)。人、非人哺乳动物和禽起源的许多腺病毒是已知的, 并且从保藏所如ATCC可获得。疱疹家族的病毒 (其也从保藏所如ATCC可获得) 包括例如单纯疱疹病毒 (HSV)、埃巴病毒 (EBV)、巨细胞病

毒 (CMV) 和假性狂犬病病毒 (PRV)。用于复制AAV的腺病毒辅助功能的例子包括E1A功能、E1B功能、E2A功能、VA功能和E4orf6功能。从保藏所可获得的杆状病毒包括苜蓿银纹夜蛾 (*Autographa californica*) 核型多角体病毒。

[0083] 若感染性AAV颗粒与感染性辅助病毒颗粒的比率是至少约 $10^2:1$;至少约 $10^4:1$;至少约 $10^6:1$;或至少约 $10^8:1$ 或更多,则rAAV的制备物被说成是“基本上没有”辅助病毒。在一些实施方案中,制备物也没有等同量的辅助病毒蛋白(即,若上文记录的辅助病毒颗粒杂质以破坏形式存在,则会由于辅助病毒的此类水平而存在的蛋白质)。病毒和/或细胞蛋白质污染一般可以在SDS凝胶上以考马斯染色条带的存在观察到(例如出现除了与AAV壳体蛋白VP1、VP2和VP3对应的条带外的条带)。

[0084] “百分比(%)序列同一性”就参照多肽或核酸序列而言定义为在比对序列并且在必要时引入缺口以实现最大百分比序列同一性后,并且不将任何保守取代看作是序列同一性的一部分,与参照多肽或核酸序列中的氨基酸残基或核苷酸相同的候选序列中的氨基酸残基或核苷酸的百分比。为了测定百分比氨基酸或核酸序列同一性的比对可以在本领域技术内的各种方式实现,例如使用公开可得到的计算机软件程序,例如那些记载于Current Protocols in Molecular Biology (Ausubel等编,1987), Supp.30,第7.7.18节,表7.7.1的,并且包括BLAST、BLAST-2、ALIGN或Megalign (DNASTAR) 软件。比对程序的一个例子是ALIGN Plus (Scientific and Educational Software, Pennsylvania)。本领域技术人员可以确定适合于测量比对的参数,包括为了在所比较的序列的全长里实现最大比对需要的任何算法。出于本文中的目的,如下计算给定氨基酸序列A对、与或相对于给定氨基酸序列B的%氨基酸序列同一性(或者,其可以表述为对、与或相对于给定的氨基酸序列B具有或包含一定%氨基酸序列同一性的给定氨基酸序列A):分数 X/Y 的100倍,其中X是通过序列比对程序在该程序的A和B的比对中评分为相同匹配的氨基酸残基的数目,并且其中Y是B中的氨基酸残基的总数。应当领会,在氨基酸序列A的长度不等于氨基酸序列B的长度的情况下,A对B的%氨基酸序列同一性不会等于B对A的%氨基酸序列同一性。出于本文中的目的,如下计算给定核酸序列C对、与或相对于给定核酸序列D的%核酸序列同一性(或者,可以表述为对、与或相对于给定的核酸序列D具有或包含某个%核酸序列同一性的给定核酸序列C):分数 W/Z 的100倍,其中W是通过序列比对程序在该程序的C和D的比对中评分为相同匹配的核苷酸的数目,并且其中Z是D中的核苷酸的总数。应当领会,在核酸序列C的长度不等于核酸序列D的长度的情况下,C对D的%核酸序列同一性不会等于D对C的%核酸序列同一性。

[0085] “分离的”分子(例如核酸或蛋白质)或细胞意味着它已经得到鉴定并且从其天然环境的组分中分离和/或回收。

[0086] “有效量”指有效实现有益或期望的结果,包括临床结果(例如改善症状、实现临床终点等)的量。可以在一次或多次投予中投予有效量。就疾病状态而言,有效量是足以改善疾病、稳定疾病或延迟疾病发生/发展的量。

[0087] “个体”或“受试者”是哺乳动物。哺乳动物包括但不限于家养动物(例如牛、羊、猫、犬和马)、灵长类(例如人和非人灵长类,如猴)、兔、和啮齿类(例如小鼠和大鼠)。在某些实施方案中,个体或受试者是人。

[0088] 如本文中使用的,“治疗/处理”是用于获得有益或期望的临床结果的办法。出于本发明的目的,有益的或期望的临床结果包括但不限于改善症状、降低疾病程度、稳定的(例如

不恶化)疾病状态、防止疾病扩散(例如转移)、延迟或减缓疾病进展、改善或减轻疾病状态、和消退(无论部分还是完全),无论可检出还是检测不到。“治疗/处理”也可以意指与若不接受治疗的预期存活相比延长存活。

[0089] 如本文中使用的术语“预防性治疗/处理”指下述处理,其中已知或怀疑个体具有或有风险具有病症,但是尚未展现症状或者已经展现出最小限度的疾病症状。可以在症状发作前治疗在经历预防性治疗/处理的个体。

[0090] 如本文中使用的,“治疗”剂(例如治疗性多肽、核酸或转基因)是提供有益或期望的临床结果,如上文描述的例示性的临床结果的药剂。因而,治疗剂可以在如上文描述的治疗/处理中使用。

[0091] 如本文中使用的,“差别系数分布数值”或“C(S)”是为了描述沉积颗粒分布(例如超速离心期间)的Lamm方程解的分布变体。

[0092] 如本文中使用的,“斯韦伯单位(Svedberg unit)”指沉积速率的单位。给定大小和形状的颗粒的沉积速率测量颗粒多么快速地沉积。1个斯韦伯单位等同于 10^{-13} 秒。例如,斯韦伯单位经常用于反映分子在离心的离心力下运动的速率。

[0093] 如本文中使用的,“沉积速度条件”或“边界沉积速度条件”可以指样品溶液进行沉积速度分析的任何实验条件。沉积速度容许在较宽范围的pH和离子强度条件并且在温度4至40°C的颗粒研究。沉积边界移动的速率是沉积种类的沉积系数的量度。

[0094] 沉积系数取决于分子重量(颗粒越大,沉积越快)并且也取决于分子形状。沉积边界的最小宽度与分子的扩散系数相关;具有相似沉积系数的多个种类的存在会使边界宽于仅基于扩散预期的边界。沉积速度条件可以包括但不限于与转子速度、样品和转子中心之间的距离、温度、溶剂、样品、缓冲液、超速离心时间、用于检测的时间间隔、扇区和光学窗口特征、AUC仪器(包括超速离心机和检测设备)、参照溶剂的平衡透析、和数据分析算法相关的任何条件。

[0095] 如本文中使用的,术语“分析密度梯度沉积平衡”指用于测量颗粒的浮力密度,或者使用浮力密度的差异来分离不同颗粒种类的方法。这些方法可以使用例如AUC沉积平衡技术。在这些方法中,颗粒溶液(例如不限于多肽、多核苷酸或病毒壳体的溶液)可以进行梯度溶剂合物,如氯化铯或硫酸铯梯度中的超速离心,直至达到与溶剂合物的平衡。在平衡时,颗粒溶液会在颗粒密度等于溶剂合物密度的梯度中的位置处浓缩或成带。可以使用条带的位置来计算颗粒密度,或者可以提取条带以分离单一种类的颗粒。

[0096] 如本文中使用的,“SEDFIT算法”是容许分析水动力学数据,如沉积速度的算法(Schuck(2000)Biophys.J.,78:1606-19)。在SEDFIT算法中,创建跨越预期范围的沉积系数的栅格。假设恒定的颗粒形状和溶剂摩擦比,使用每个沉积系数的Lamm方程的解模拟沉积边界。

[0097] 如本文中使用的,术语“F统计量”或“F比率”指置信水平。此参数控制使用的正则化(regularization)的量。它对于不同范围具有不同意义:从0至0.5,不使用正则化。从0.5至0.999的数值对应于概率P(置信水平)。从这些P值,用F统计量计算正则化的简约性约束容许的期望的卡方(chi-square)增加。0.51的数值会引起非常少的正则化;0.68至0.90的数值会对应于通常使用的置信水平(通常,通过50次以上的扫描,与0.7的概率对应的卡方增加是0.1%量级的),而接近0.99的数值会引起非常高的正则化。可以使用F统计量计算器检

查这些数值与概率的关系。若输入数目>1,则它们直接取为卡方比率(因为没有概率>1)。例如,1.1的数值会导致以10%卡方增加的正则化。

[0098] 在本文中提及“约”数值或参数包括(并且描述)涉及该数值或参数本身的实施方案。例如,提及“约X”的描述包括描述“X”。

[0099] 如本文中使用的冠词“一个”、“一种”和“所述/该”的单数形式包括复数提及物,除非另有指示。

[0100] 应当理解,本文中描述的本发明的态样和实施方案包括“包括”、“由…组成”和/或“基本上由…组成”的态样和实施方案。

[0101] III. 病毒颗粒

[0102] 本公开内容的某些态样涉及含有超大重组AAV(rAAV)基因组的腺相关病毒(AAV)颗粒(例如通过本文中公开的方法和/或细胞系生产)。本公开内容的某些态样涉及腺相关病毒(AAV)颗粒,其含有约4.7kb-约9.4kb,任选地约4.7kb-6.7kb的rAAV基因组。在一些实施方案中,通过AAV壳体壳体制化的rAAV基因组大于约5kb。在一些实施方案中,rAAV颗粒包含rAAV载体。在一些实施方案中,rAAV载体含有约4.7kb-约9.4kb,任选地约4.7kb-6.7kb的rAAV基因组。在一些实施方案中,rAAV基因组大于约5kb。在一些实施方案中,rAAV基因组是约5kb-约7.0kb、约4.7kb-约9.4kb或约4.7kb-约6.7kb。在一些实施方案中,rAAV基因组的长度大于约5.0kb、5.1kb、5.2kb、5.3kb、5.4kb、5.5kb、5.6kb、5.7kb、5.8kb、5.9kb、6.0kb、6.1kb、6.2kb、6.3kb、6.4kb、6.5kb、6.6kb、6.7kb、6.8kb、6.9kb、7.0kb、7.1kb、7.2kb、7.3kb、7.4kb、7.5kb、7.6kb、7.7kb、7.8kb、7.9kb、8.0kb、8.1kb、8.2kb、8.3kb、8.4kb、8.5kb、8.6kb、8.7kb、8.8kb、8.9kb、9.0kb、9.2kb、9.3kb或9.4kb中的任一者或其间的任何数值。

[0103] 在一些实施方案中,病毒颗粒是包含核酸的重组AAV颗粒,所述核酸包含侧翼为一个或两个AAV反向末端重复(ITR)的异源核酸(例如异源转基因)。在AAV颗粒中壳体化核酸。AAV颗粒也包含壳体蛋白。在一些实施方案中,核酸包含感兴趣的编码序列(例如异源转基因)、在转录方向上的可操作连接的组分、控制序列,包括转录启动和终止序列,由此形成表达盒。表达盒在5'和3'末端侧翼有至少一个功能性AAV ITR序列。“功能性AAV ITR序列”意指ITR序列与意图用于挽救、复制和包装AAV病毒体一样发挥功能。参见Davidson等,PNAS,2000,97(7)3428-32;Passini等,J.Virol.,2003,77(12):7034-40;及Pechan等,Gene Ther.,2009,16:10-16,它们全部通过提述以其整体并入本文。为了实施本发明的一些态样,重组载体包含至少所有对于壳体化必需的AAV序列和用于rAAV感染的物理结构。本发明的载体中使用的AAV ITR不需要具有野生型核苷酸序列(例如如记载于Kotin,Hum.Gene Ther.,1994,5:793-801),并且可以通过插入、缺失或取代核苷酸而改变或者AAV ITR可以源自几种AAV血清型之任一者。目前已知超过40种AAV血清型,并且继续鉴定出新的血清型和现存血清型的变体。参见Gao等,PNAS,2002,99(18):11854-6;Gao等,PNAS,2003,100(10):6081-6;及Bossis等,J.Virol.,2003,77(12):6799-810。认为任何AAV血清型的使用在本发明的范围内。在一些实施方案中,rAAV载体是源自AAV血清型,包括但不限于AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10、AAVrh10、AAV11、AAV12、AAV2R471A、AAV DJ、山羊AAV、牛AAV或小鼠AAV等的载体。例如,在一些实施方案中,AAV血清型是AAV1、AAV2、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10或AAVrh10。

在一些实施方案中, AAV ITR中的核酸是AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10、AAVrh10、AAV11、AAV12、AAV2R471A、AAV DJ、山羊AAV、牛AAV或小鼠AAV血清型ITR等。在某些实施方案中, AAV中的核酸包含AAV2ITR。

[0104] 在别的实施方案中, rAAV颗粒包含AAV1壳体、AAV2壳体、AAV3壳体、AAV4壳体、AAV5壳体、AAV6壳体(例如野生型AAV6壳体或变体AAV6壳体, 如ShH10, 如记载于U.S. PG Pub. 2012/0164106)、AAV7壳体、AAV8壳体、AAVrh8壳体、AAVrh8R壳体、AAV9壳体(例如野生型AAV9壳体, 或经修饰的AAV9壳体, 如记载于U.S. PG Pub. 2013/0323226)、AAV10壳体、AAVrh10壳体、AAV11壳体、AAV12壳体、酪氨酸壳体突变体、肝素结合壳体突变体、AAV2R471A壳体、AAVAAV2/2-7m8壳体、AAV DJ壳体(例如AAV-DJ/8壳体、AAV-DJ/9壳体或U.S. PG Pub. 2012/0066783中描述的任何其它壳体)、AAV2N587A壳体、AAV2E548A壳体、AAV2N708A壳体、AAV V708K壳体、山羊AAV壳体、AAV1/AAV2嵌合壳体、牛AAV壳体、小鼠AAV壳体、rAAV2/HBoV1壳体或美国专利No. 8, 283, 151或国际公开文本No. WO/2003/042397中描述的AAV壳体。在一些实施方案中, 突变体壳体蛋白维持形成AAV壳体的能力。在一些实施方案中, rAAV颗粒包含AAV5酪氨酸突变体壳体(Zhong L. 等, (2008) Proc Natl Acad Sci U S A 105 (22): 7827-7832。在别的实施方案中, rAAV颗粒包含来自A-F进化枝的AAV血清型的壳体蛋白(Gao, 等, J. Virol. 2004, 78 (12): 6381)。在一些实施方案中, rAAV颗粒包含AAV1壳体蛋白或其突变体。在其它实施方案中, rAAV颗粒包含AAV2壳体蛋白或其突变体。在一些实施方案中, AAV血清型是AAV1、AAV2、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10或AAVrh10。在一些实施方案中, rAAV颗粒包含AAV血清型1 (AAV1) 壳体。在一些实施方案中, rAAV颗粒包含AAV血清型2 (AAV2) 壳体。在一些实施方案中, rAAV颗粒包含AAVrh8R壳体或其突变体。

[0105] 使用不同AAV血清型来优化特定靶细胞的转化或者靶向特定靶组织(例如肝或CNS组织)内的特定细胞类型。rAAV颗粒可以包含源自相同血清型或不同血清型(例如混合血清型)的病毒蛋白和病毒核酸。例如, 在一些实施方案中, rAAV颗粒可以包含AAV1壳体蛋白和至少一个AAV2ITR或者它可以包含AAV2壳体蛋白和至少一个AAV1ITR。本文中提供了用于产生rAAV颗粒的AAV血清型的任何组合, 就像本文中已经明确叙述每种组合一样。在一些实施方案中, 本发明提供了rAAV颗粒, 其包含AAV1壳体和本公开内容的rAAV载体(例如包含异源核酸的表达盒), 侧翼有至少一个AAV2ITR。在一些实施方案中, 本发明提供了包含AAV2壳体的rAAV颗粒。在一些实施方案中, ITR和壳体源自AAV2。在其它实施方案中, ITR源自AAV2, 并且壳体源自AAVrh8R。

[0106] 本公开内容的其它态样涉及包含rAAV颗粒的组合物, 其中至少约15%、至少约20%、至少约25%、至少约30%、至少约35%、至少约40%、至少约45%、至少约50%、至少约55%、至少约60%或至少约70%、至少约80%、至少约90%或至少约95%的rAAV颗粒壳体化约4.7kb-约9.4kb, 任选地约4.7kb-6.7kb的rAAV基因组。在一些实施方案中, rAAV颗粒壳体化大于约5kb的基因组。在一些实施方案中, rAAV颗粒壳体化长度大于约5.0kb、5.1kb、5.2kb、5.3kb、5.4kb、5.5kb、5.6kb、5.7kb、5.8kb、5.9kb、6.0kb、6.1kb、6.2kb、6.3kb、6.4kb、6.5kb、6.6kb、6.7kb、6.8kb、6.9kb、7.0kb、8.0kb或9.0kb中任一者或其间的任何数值的基因组。在一些实施方案中, 包装的AAV基因组不含5'端的截短。在一些实施方案中, 包装的AAV基因组不含3'端的截短。用于测定rAAV基因组的大小的方法是本领域中已知的, 并

且包括但不限于Southern印迹和分析超速离心,如下文描述。

[0107] 在一些实施方案中,本公开内容的组合物含有rAAV颗粒,其中至少约15%、至少约20%、至少约25%、至少约30%、至少约35%、至少约40%、至少约45%、至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%或至少约95%的rAAV颗粒壳体化大于约4.7kb、大于约5.0kb、大于约5.1kb、大于约5.2kb、大于约5.3kb、大于约5.4kb、大于约5.5kb、大于约5.6kb、大于约5.7kb、大于约5.8kb、大于约5.9kb、大于约6.0kb、大于约6.5kb、大于约7.0kb、大于约7.5kb、大于约8.0kb、大于约8.5kb、大于约9.0kb或大于约9.4kb的rAAV基因组。在一些实施方案中,包装的AAV基因组不含5'端的截短。在一些实施方案中,包装的AAV基因组不含3'端的截短。

[0108] 在本发明的一些实施方案中,组合物中的重组病毒颗粒是高度纯化的,适当缓冲的,并浓缩的。在一些实施方案中,将病毒颗粒浓缩到至少约 1×10^7 vg/mL至约 9×10^{13} vg/mL或其间的任何浓度。

[0109] 如本文中描述,一种用于表征病毒颗粒的制备物(例如一种或多种与载体基因组大小和/或完整性相关的特性)的技术是经由使用Southern印迹法。例如,如下文实施例中更为详细描述,可以用DNA酶处理rAAV颗粒的制备物(任选如本文中描述的那样纯化)以除去任何非壳体化的核酸,用停止DNA酶消化的试剂(例如EDTA)处理,用蛋白酶消化,然后进行DNA提取以除去包装的载体基因组。然后,可以使用电泳分离载体基因组,交联到膜上,并且用一种或多种标记的探针探查,所述探针与载体基因组特异性杂交。通过与标记的探针杂交而标记的DNA片段的大小(例如与一种或多种特定的大小标志物比较)指示载体基因组大小。另外,可以使用一种或多种探针,其与载体基因组的已知区段(例如5'或3'端)杂交。若这些探针中的一种或多种未能与载体基因组杂交,则这指示可以截短或以其它方式缺失制备物的载体基因组,使得它们短于其预测的完整大小。由于已知AAV基因组的包装从3'端开始发生(King, J.A.等(2001)EMBO J. 20:3282-3291),当基因组大小超过4.7kb时,超大载体可以缺乏负和正链的5'端中的序列。在一些实施方案中,病毒颗粒包含大于约5.0kb的超大rAAV基因组,其中rAAV颗粒中壳体化的病毒基因组包含相对完整的5'和3'端;例如如通过与对5'和/或3'端特异性的探针杂交来测量。可通过本领域中已知的方法测量杂交,所述方法如但不限于Southern印迹分析或PCR。在一些实施方案中,包装的AAV基因组不含5'端的截短。在一些实施方案中,包装的AAV基因组不含3'端的截短。

[0110] 分析超速离心

[0111] 如本文中描述,一种用于表征病毒颗粒的制备物的技术(例如一种或多种与载体基因组大小和/或完整性相关的特性)是经由使用分析超速离心(AUC)。例如,在一些实施方案中,使用AUC评估rAAV颗粒制备物中重组腺相关病毒(rAAV)颗粒的载体基因组完整性以区分具有完全、完整基因组的病毒颗粒、空病毒壳体和具有变体(例如截短、聚集体、杂质等)病毒基因组的病毒颗粒。使用分析超速离心表征病毒(例如AAV)颗粒的更多描述可以参见2015年1月20日提交的美国临时专利申请系列号62/105,714,“Analytical Ultracentrifugation for Characterization of Recombinant Viral Particles”,在此通过提述以其整体并入。

[0112] 分析超速离心是评估蛋白质或其它大分子的分子量和水动力学和热力学特征的手段。在一系列条件(包括浓度、温度、离子强度和pH)里,通过沉积速度得出蛋白质或大分

子的异质性。例如,可以在临床相关配制剂中分析蛋白质。分析超速离心表征腺病毒制备物的用途由Berkowitz, SA&Philo JS, (2007) Anal. Biochem., 362:16-37提供。

[0113] AUC分析指用于表征颗粒(例如多肽、多核苷酸和病毒壳体)的生物物理特征的定量方法,该方法通过测量其迁移通过离心场中的溶剂进行。在几十年里已经完善表征AUC分析,并且是多用途的。由于AUC分析依赖于第一原理水动力学和热力学信息,可以应用AUC以在颗粒浓度和大小的较宽范围中测定许多类型的颗粒的生物物理特性。AUC分析通常涵盖两种基本类型的实验:沉积速度和沉积平衡。沉积平衡分析产生颗粒的热力学特性,其可以用于测量特征,如化学计量学和结合常数。沉积速度产生颗粒的水动力学特性,其可以用于测量特征,如大小、形状和浓度。病毒制备物的AUC分析的特征是可以使用相同测定条件来分析病毒颗粒的不同制备物,而不论病毒基因组的核苷酸序列或壳体的血清型如何。

[0114] 本公开内容的某些态样涉及沉积速度分析表征病毒壳体特性的用途。在一些实施方案中,沉积速度分析使用超速离心机速度池(ultracentrifuge velocity cell),其具有透析平衡中的两个扇区(一个用于实验样品和一个用于仅溶剂的参照样品),每个含有让光通过区室的两个光学窗口。超速离心对细胞施加角速度,并且导致溶质颗粒向扇区底部的快速沉积。由于发生沉积,溶质在池的顶部接近弯月面(meniscus)被消减,创建消减区域和沉积溶质之间的沉积边界。通过采取以特定时间间隔(对于沉积速度,这些间隔通常为分钟量级)比较样品和参照扇区的特性的测量来测量沉积边界的运动或迁移速率。若存在多个种类的溶质,则这可以导致多个沉积边界的形成,每个对应于可解析的种类。

[0115] 用于光学检测沉积边界并且测量其运动或迁移速率的几种方法是本领域中已知的(关于参考文献,参见Cole等(2008) Methods Cell Biol., 84:143-79)。在一些实施方案中,可以使用吸光度检测测定参照和样品扇区。在此检测方法中,可以在每个扇区内的不同径向位置处对样品和参照扇区测量特定波长的吸光度。或者,可以测量单一径向位置处的吸光度的时间过程。比尔定律(Beer's Law)提供吸光度和溶质的消光系数之间的数学关系。

[0116] 在一些实施方案中,可以使用干涉检测(例如Rayleigh干涉检测)测定参照和样品扇区。在Rayleigh干涉检测法中,干涉光学系统含有两个平行狭缝。单一相干光束被分开,使得它通过两个窗口,然后再合并两个束。当合并这两个光波时,它们形成交替的亮条纹和暗条纹的干涉模式。若样品和参照样品要具有相同的折射率,则所得的干涉条纹会是完全直的。增加溶质浓度增加了溶液的折射率,由此延迟样品光束,并且引起垂直条纹位移(shift)。通过测量此条纹位移,可以测量样品中的溶质浓度。与吸光度测量(其测量样品和参照的绝对值)不一样,干涉检测测量样品和参照之间的相对差异。然而,干涉检测产生与浓度成正比的积分峰,并且它可以用于不显著吸收的样品类型。关于使用具有AUC的Rayleigh干涉光学的参考文献,参见Furst(1997) Eur. Biophys. J. 35:307-10。

[0117] 可以使用沉积边界移动的速率的测量来得出溶质颗粒的许多物理特性。边界移动的速率决定沉积系数,其基于颗粒的质量和形状(摩擦系数)。颗粒的沉积系数指其速度与通过离心场对它施加的加速度的比率。沉积系数以斯韦伯单位S(1个斯韦伯单位等同于 10^{-13} 秒)表示。颗粒或颗粒溶液的沉积系数取决于其特性,例如分子量(针对浮力校准),和溶剂的特性。

[0118] 可以使用Lamm方程(Schuck(2000) Biophys. J., 78:1606-19)测定超速离心期间溶

质的浓度边界随时间的变化。简言之,考虑扇形形状池和由转子产生的离心场,Lamm方程计算响应于沉积(其使溶质浓缩)和扩散(其使溶质分散)的竞争力,溶质的浓度边界随时间的变化。Lamm方程可以表示为:

[0119] 方程1: $\partial c / \partial t = D [(\partial^2 c / \partial r^2) + 1/r(\partial c / \partial r)] - s\omega^2 [r(\partial c / \partial r) + 2c]$

[0120] 其中c是溶质浓度,D代表溶质扩散常数,s代表沉积系数, ω 代表转子的角速度,r是半径,并且t是时间。

[0121] 通过拟合原始AUC数据到Lamm方程的解,有可能测定溶质特征,如沉积系数和浓度分布的变化。例如,可以使用Lamm方程对沉积边界的变化速率的实验测定的数值建模来得出形成边界的溶质的沉积系数、分子量或浓度。可以使用本领域中已知的几种程序,如SEDFIT (Schuck (2000) Biophys. J., 78:1606-19) 将Lamm方程建模到AUC数据。这些程序也能够将Lamm方程应用到含有多种溶质或多种沉积边界的溶液。

[0122] 适合于测定溶质特征的程序的一个例子是SEDFIT算法。在一些实施方案中,SEDFIT算法可以用于计算差别系数分布数值或C(S),其使用来自含有颗粒种类混合物的溶液的AUC数据进行(为了参考,参见Schuck (2000) Biophys. J., 78:1606-19)。在SEDFIT算法中,创建跨越预期范围的沉积系数的栅格。假设恒定的颗粒形状和溶剂摩擦比,使用每个沉积系数的Lamm方程的解模拟沉积边界。然后,将实际AUC数据拟合到这些Lamm解,以得出差别系数分布数值或C(S)。可用于分析AUC数据的许多其它程序可以参见Cole和Hansen (1999) J. Biomol. Tech. 10:163-76。

[0123] 在一些实施方案中,病毒颗粒在合适的宿主细胞中产生并纯化。在一些实施方案中,通过亲和层析纯化病毒颗粒。纯化AAV颗粒的方法是本领域中已知的。例如,通过使用层析介质上固定化的病毒壳体蛋白的抗体或病毒壳体蛋白的结合配体。

[0124] 在一些实施方案中,使用分析超速离心来进行沉积速度分析超速离心(SV-AUC)分析,所述分析超速离心能够在生物学相关溶液条件下表征天然状态的样品(例如ProteomeLab™XL-I (Beckman Coulter))。当使用ProteomeLab™XL-1时,将样品上样到两个扇区速度池的样品扇区中,将媒介物对照(例如没有重组病毒的PBS)上样到相应的参照扇区中。在四孔转子中放置样品,并且容许该样品在仪器中平衡,直至将约20℃的温度和完全真空维持约1小时。在一个例示性的实施方案中,以约20,000RPM、约20℃和约0.003cm径向步骤设置进行沉积速度离心,没有延迟且没有重复。如下文所述,不同参数可以用于离心。在一些实施方案中,使用吸光度(260nm)和/或干涉光学(例如Rayleigh干涉光学)以同时记录径向浓度作为时间的函数,直至最小的沉积组分使光学窗口澄清。在一些实施方案中,记录径向浓度,直至具有最低密度的沉积种类使扇区澄清。在一些实施方案中,监测沉积,直至具有最低密度的重组病毒颗粒沉积到超速离心机的扇区的底部。扇区可以是超速离心机的部分;例如超速离心机速度池。在一些实施方案中,扇区可以是检测样品的超速离心机的部分。在一些实施方案中,超速离心利用包含超速离心机速度池的超速离心机。在一些实施方案中,进行监测,直至重组病毒颗粒沉积到超速离心机速度池的底部。在一些实施方案中,监测沉积,直至具有最低密度的重组病毒颗粒沉积,并且使光学窗口澄清。在一些实施方案中,记录径向浓度至少约0.5小时、0.75小时、1.0小时、1.5小时、2.0小时、3.0小时、4.0小时或5.0小时中任一者。在一些实施方案中,记录径向浓度约1.2小时。优化运行条件可以包括例如继续运行,直至所有沉积种类以保持恒定于20℃的温度和18,000rpm-20,000rpm

的速度完全沉积到扇区的底部。如下文所述,可以使用其它温度和速度。

[0125] 通过使用SEDFIT连续大小C(S)分布模型从每种检测方法分析多次扫描(例如75)测定百分比完全壳体。对拟合应用二阶(第2)导数正则化。在一些实施方案中,F统计量的置信水平是约0.68。在一些实施方案中,F统计量的置信水平是超过约0.68、0.70、0.75、0.80、0.85、0.90、0.95或0.99中任一者,或其间的任何数值。在一些实施方案中,将下列C(S)参数保持恒定:约200S至约5000S的分辨率,S_{min}是约1S至约100S,S_{max}是约100S至约5000S,并且摩擦比是约1.0或保持浮动于由离心软件确定的数值。在一些实施方案中,分辨率是约200S、300S、400S、500S、600S、700S、800S、900S或1000S中任一者或其间的任何数值。在一些实施方案中,分辨率是约200S。在一些实施方案中,S_{max}是约100S、200S、300S、400S、500S、600S、700S、800S、900S或1000S中任一者或其间的任何数值。在一些实施方案中,其中S_{max}是约200S。在一些实施方案中,使摩擦比浮动于由离心软件确定的数值。在一些实施方案中,摩擦比是约1.0。在一些实施方案中,应用径向不变量(RI)和时间不变量(TI)噪音扣除。在一些实施方案中,容许弯月面位置浮动,让软件选择最佳的位置。在一些实施方案中,容许摩擦比浮动,让软件选择最佳的位置。模型将数据拟合到Lamm方程,并且所得的大小分布是“沉积系数分布”,其看起来像具有与以条纹(Fringes)单位或OD260单位计的浓度成比例的每个峰下的面积的层析图。对分布中的每个组分测定沉积系数(以斯韦伯单位计)和相对浓度(以OD单位计)。在一些实施方案中,多个AUC运行是独立的测定法,并且每次分析监测下列属性以确保结果的质量:拟合优度(rmsd)、每个峰的OD260nm/以条纹计的干涉信号的比率(A260/IF比率)、运行间的每个种类的沉积系数的一致性、和扫描的总体质量。

[0126] 在本发明的一些实施方案中,使用消光系数从吸光度数据计算完整载体峰的实际百分比数值和摩尔浓度。可以基于发表的公式(Sommer等(2003) *Mol Ther.*, 7:122-8)计算用于空壳体($\epsilon_{260}/\text{壳体}=3.72\text{e}6$)和完整载体($\epsilon_{260}/\text{载体}=3.00\text{e}7$)两者的摩尔吸光度消光系数。消光系数可用于空壳体和完整载体峰。可以使用由Schuck(2000) *Biophys. J.*, 78:1606-19描述的SEDFIT算法确定C(S)值。可以使用比尔定律计算完整载体和空壳体两者的摩尔浓度,并且从这些数值中计算完整壳体的百分比。在一些实施方案中,按照完整壳体的百分比报告数值。

[0127] 在一些实施方案中,不可能凭经验确定特定种类的重组病毒颗粒(例如具有未知大小和序列的片段化基因组的病毒颗粒)的消光系数。可以通过分析具有壳体化的已知核苷酸大小的病毒基因组的重组病毒载体制备物建立S值和基因组大小之间的关系,并且测定相应的S值,如本文中描述的。可以将计算的S值绘图以产生标准曲线,可以与所述标准曲线比较未知分子量或基因组大小的重组病毒种类以测定未知种类的分子量。

[0128] 在一些态样中,如下表征重组病毒颗粒(例如rAAV颗粒)的制备物:a)使制备物在边界沉积速度条件下进行分析超速离心,其中以时间间隔监测重组病毒颗粒的沉积(例如一次或多次),b)将差别沉积系数分布数值(C(s))相对于以斯韦伯单位(S)计的沉积系数绘图,c)对C(s)分布中的每个峰下的面积进行积分以测定每个峰的相对浓度,其中每个峰代表重组病毒颗粒的种类。在一些实施方案中,鉴定的重组病毒颗粒的种类包括但不限于:包含完整重组病毒基因组的完全重组病毒颗粒、空重组病毒壳体颗粒、和包含变体重组病毒基因组的重组病毒颗粒。在一些实施方案中,变体基因组小于完整的重组病毒基因组(例如截短的基因组)。在一些实施方案中,变体基因组大于完整的重组病毒基因组(例如聚集体、

重组体等)。在一些实施方案中,如下表征重组病毒颗粒(例如rAAV颗粒)的制备物:a)使制备物在边界沉积速度条件下进行分析超速离心,其中以时间间隔监测重组病毒颗粒的沉积(例如一次或多次),b)将差别沉积系数分布数值($C(s)$)相对于以斯韦伯单位(S)计的沉积系数绘图,c)通过与S值对应的图上的峰的存在来鉴定制备物中的重组病毒颗粒种类,其中通过比较种类的S值与标准曲线来计算特定种类的重组病毒颗粒的基因组大小,所述标准曲线由包含不同已知大小的壳体化的病毒基因组的重组病毒颗粒的S值产生。在一些实施方案中,方法进一步包括对 $C(s)$ 分布中的每个峰下的面积进行积分以确定每个种类的重组病毒颗粒的相对浓度。在一些实施方案中,以一个时间间隔监测重组病毒颗粒的沉积。在一些实施方案中,以超过一个时间间隔监测重组病毒颗粒的沉积。

[0129] 在一些实施方案中,通过测量在约260nm的光密度或吸光度监测重组病毒颗粒(例如rAAV颗粒)的沉积。测量吸光度的手段是本领域中已知的。在一些实施方案中,用于AUC的超速离心机装备有用于测量吸光度的手段/装置。在其它实施方案中,通过干涉监测重组病毒颗粒的沉积。在一些实施方案中,通过Rayleigh干涉监测重组病毒颗粒的沉积。测量干涉的手段是本领域中已知的(Furst (1997) Eur. Biophys. J. 35:307-10)。在一些实施方案中,用于AUC的超速离心机装备有用于测量干涉的手段/装置。在一些实施方案中,通过吸光度和干涉两者监测重组病毒颗粒的沉积。在一些实施方案中,使用参照标准品测量吸光度和/或干涉。在一些实施方案中,参照标准品匹配重组病毒制备物的溶液,只是不存在重组病毒。例如,重组病毒制备物可以包含缓冲液(如磷酸盐缓冲盐水)中的重组病毒。在此例子中,参照标准品可以是没有重组病毒颗粒的磷酸盐缓冲盐水。

[0130] 在一些实施方案中,通过在超速离心期间连续监测病毒颗粒的沉积来测定超速离心期间病毒颗粒的沉积速度。在熟练技术人员范围内的是对不同类型的病毒颗粒优化AUC参数。在一些实施方案中,以约3,000-约20,000rpm的AUC速度进行用于rAAV颗粒的数据采集。在一些实施方案中,以约1S的 S_{min} 和约1000S的 S_{max} 进行用于rAAV颗粒的数据分析。在一些实施方案中,以约200S至约1,000S的分辨率进行用于rAAV颗粒的数据分析。在一些实施方案中,分辨率是约200S、300S、400S、500S、600S、700S、800S、900S或1000S中任一者或其间的任何数值。在一些实施方案中,分辨率是约200S。在一些实施方案中,以约100S、200S、300S、400S、500S、600S、700S、800S、900S或1000S中任一者或其间的任何数值的 S_{max} 进行用于rAAV颗粒的数据分析。在一些实施方案中, S_{max} 是约200S至约5000S。在一些实施方案中,其中 S_{max} 是约200S。在一些实施方案中,应用径向不变量(RI)和时间不变量(TI)噪音扣除。在一些实施方案中,容许弯月面位置浮动,让软件选择最佳的位置。在一些实施方案中,容许摩擦比浮动,让软件选择最佳的位置。在一些实施方案中,用于rAAV颗粒的数据分析在1处保持恒定。在一些实施方案中,通过使用具有使用非线性回归优化的数值的FIT命令容许用于rAAV颗粒的数据分析浮动。

[0131] 就重组病毒颗粒(例如rAAV颗粒)而言,在一些实施方案中,通过以超过约每15秒、30秒、45秒、1分钟(60秒)、2分钟、3分钟、4分钟、5分钟、6分钟、7分钟、8分钟、9分钟、10分钟、15分钟、20分钟、25分钟一次监测(例如扫描)重组病毒颗粒的沉积来测定超速离心期间重组病毒的沉积速度。可以在没有延迟的情况下与光学系统容许一样快地连续采集扫描。干涉扫描是快速的,并且单次扫描在~10-15秒中完成,而吸光度扫描需要~60秒。当使用双重检测时,用于这两者的扫描采集的速度由吸光度系统决定。在本发明的一些实施方案中,

在超速离心期间使用超过约5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95或100次扫描以监测重组病毒颗粒的沉积。在一些实施方案中,需要最少30次扫描进行分析,并且收集扫描,直至沉积过程是完全的。在一些实施方案中,通常可以通过40-75次扫描描述沉积过程。在一些实施方案中,基于约75次扫描测定重组病毒颗粒的沉积速度。在一些实施方案中,基于约55次扫描至约75次扫描测定重组病毒颗粒的沉积速度。在一些实施方案中,基于约55次扫描至约60次扫描测定重组病毒颗粒的沉积速度。在一些实施方案中,基于约60次扫描至约75次扫描测定重组病毒颗粒的沉积速度。在一些实施方案中,基于约60次扫描至约70次扫描测定重组病毒颗粒的沉积速度。在一些实施方案中,基于多次超速离心(运行)测定重组病毒颗粒的沉积速度。在一些实施方案中,基于1、2、3、4、5、6、7、8、9、10或更多次中任一者的超速离心运行测定重组病毒颗粒的沉积速度。在一些实施方案中,使用SEDFIT算法,使用沉积速度测定C(S)值。在一些实施方案中,以F统计量的约0.68的置信水平对拟合水平应用二阶导数正则化。在一些实施方案中,下列C(S)参数保持恒定:分辨率100S至约200S, S_{min} 是约1, S_{max} 是约200S至300S,并且摩擦比是约1.0至1.2S。在一些实施方案中,应用径向不变量(RI)和时间不变量(TI)噪音扣除。

[0132] 在一些实施方案中,通过以超过约5,000rpm;10,000rpm;15,000rpm;20,000rpm;25,000rpm;30,000rpm;35,000rpm;40,000rpm;45,000rpm;或50,000rpm中任一者或其间的任何数值超速离心重组病毒颗粒的制备物来测定重组病毒颗粒制备物中的重组病毒颗粒(例如rAAV颗粒)的边界沉积速度。在本发明的一些实施方案中,通过以约20,000rpm超速离心重组病毒颗粒的制备物来测定重组病毒颗粒制备物中的重组病毒颗粒的边界沉积速度。在本发明的一些实施方案中,通过以约15,000rpm至约20,000rpm超速离心重组病毒颗粒的制备物来测定重组病毒颗粒制备物中的重组病毒颗粒的边界沉积速度。

[0133] 在一些实施方案中,通过于约或超过4℃、10℃、15℃、20℃、25℃或30℃或其间的任何温度超速离心重组病毒颗粒制备物来测定重组病毒颗粒(例如rAAV颗粒)制备物中的重组病毒颗粒的边界沉积速度。在一些实施方案中,通过于约20℃超速离心重组病毒颗粒制备物来测定重组病毒颗粒制备物中的重组病毒颗粒的边界沉积速度。在一些实施方案中,通过于约15℃至约20℃超速离心重组病毒颗粒制备物来测定重组病毒颗粒制备物中的重组病毒颗粒的边界沉积速度。

[0134] 具有增强的表达的病毒颗粒

[0135] 在一些态样中,本发明提供了具有增强的表达的包含超大载体基因组的病毒颗粒。在一些实施方案中,与通过细胞的瞬时转染制备的AAV颗粒相比,超大rAAV基因组在使用生产细胞系在AAV颗粒中包装时展现增强的表达。在一些实施方案中,本发明提供了用于增强超大rAAV基因组的表达的方法,所述方法包括通过对细胞系提供AAV辅助功能在生产细胞系中产生rAAV颗粒,其中生产细胞系包含a)编码AAV rep和cap基因的核酸,和b) rAAV基因组,其中rAAV基因组大于约4.7kb。在一些实施方案中,超大rAAV基因组的表达是当通过瞬时转染产生时超大rAAV基因组的表达的大约1.25倍、约1.5倍、约1.75倍、约2.0倍、约2.25倍、约2.5倍、约2.75倍、约3倍、约3.25倍、约3.5倍、约3.75倍、约4倍、约4.25倍、约4.5倍、约4.75倍或约5倍。在一些实施方案中,与来自通过瞬时转染产生的rAAV颗粒的超大rAAV基因组的表达动力学相比,超大rAAV基因组的增强的表达是更快的表达动力学。在一些实施方案中,更快的表达动力学是在对细胞投递包含超大rAAV基因组的AAV颗粒后超大

rAAV基因组的表达随时间的更快增加。在一些实施方案中,与投递包含来自通过瞬时转染制备的rAAV颗粒的超大rAAV基因组的AAV颗粒后超大rAAV基因组的表达水平相比,更快的表达动力学是在对细胞投递包含超大rAAV基因组的AAV颗粒后达到超大rAAV基因组的最大或稳定状态表达水平的时间更快。在一些实施方案中,由生产细胞系产生的超大rAAV基因组的表达动力学比来自通过瞬时转染产生的rAAV颗粒的超大rAAV基因组的表达动力学快约5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%或100%中任一者。在一些实施方案中,超大载体基因组的长度大于约5.0kb、5.1kb、5.2kb、5.3kb、5.4kb、5.5kb、5.6kb、5.7kb、5.8kb、5.9kb、6.0kb、6.1kb、6.2kb、6.3kb、6.4kb、6.5kb、6.6kb、6.7kb、6.8kb、6.9kb、7.0kb、7.1kb、7.2kb、7.3kb、7.4kb、7.5kb、7.6kb、7.7kb、7.8kb、7.9kb、8.0kb、8.1kb、8.2kb、8.3kb、8.4kb、8.5kb、8.6kb、8.7kb、8.8kb、8.9kb、9.0kb、9.2kb、9.3kb或9.4kb中任一者或其间的任何数值。

[0136] 异源转基因

[0137] 在一些实施方案中,病毒颗粒是包含超大载体基因组的重组AAV颗粒,所述超大载体基因组包含侧翼有一个或两个AAV反向末端重复(ITR)的异源核酸(例如异源转基因)。在AAV颗粒中壳体化核酸。在一些实施方案中,本公开内容的rAAV基因组含有一个或多个AAV反向末端重复(ITR)和异源转基因。例如,在一些实施方案中,本公开内容的rAAV基因组含有两个AAV反向末端重复(ITR)。在某些实施方案中,本公开内容的rAAV基因组含有两个AAV反向末端重复(ITR)和异源转基因。在一些实施方案中,载体基因组是约4.7kb-约9.4kb,任选地约4.7kb-6.7kb。在一些实施方案中,载体基因组是大于约5kb。在一些实施方案中,载体基因组是约5kb-约7kb、约4.7kb-约9.4kb或约4.7kb-6.7kb,或其间的任何数值。在一些实施方案中,载体基因组的长度大于约5.0kb、5.1kb、5.2kb、5.3kb、5.4kb、5.5kb、5.6kb、5.7kb、5.8kb、5.9kb、6.0kb、6.1kb、6.2kb、6.3kb、6.4kb、6.5kb、6.6kb、6.7kb、6.8kb、6.9kb、7.0kb、7.1kb、7.2kb、7.3kb、7.4kb、7.5kb、7.6kb、7.7kb、7.8kb、7.9kb、8.0kb、8.1kb、8.2kb、8.3kb、8.4kb、8.5kb、8.6kb、8.7kb、8.8kb、8.9kb、9.0kb、9.2kb、9.3kb或9.4kb中任一者或其间的任何数值。

[0138] 在一些实施方案中,异源转基因编码治疗性转基因产物。在一些实施方案中,治疗性转基因产物是治疗性多肽。例如,治疗性多肽可以供应细胞或生物体中缺乏或以降低的水平存在的多肽和/或酶促活性。或者,治疗性多肽可以供应间接抵消细胞或生物体中的不平衡的多肽和/或酶促活性。例如,用于与由代谢酶或活性缺乏引起的代谢物积累相关的病症的治疗性多肽可以供应缺少的代谢酶或活性,或者它可以供应导致代谢物减少的替换代谢酶或活性。也可以使用治疗性多肽通过例如充当显性负性多肽来降低多肽(例如过表达、通过功能获得突变活化或其活性以其它方式误调节的多肽)的活性。

[0139] 在一些实施方案中,异源转基因编码因子VIII。在一些实施方案中,因子VIII是人因子VIII编码序列,包括但不限于由人因子VIII基因表达的任何编码序列。人因子VIII基因(例如GenBank Entrez Gene ID 2157)又称为AHF、F8、F8B、F8C、HEMA、FVIII和DXS1253E。在一些实施方案中,因子VIII具有人因子VIII的氨基酸序列(例如如以GenBank登录号AAA52484代表)。可以使用编码因子VIII的异源转基因,例如以在患有血友病A(一种与因子VIII缺乏有关的隐性X连锁的凝固病症)的个体中表达因子VIII。已知因子VIII作为固有血液凝固途径的一部分参与血液凝固,并且通常由遍及整个身体的肝血窦细胞和内皮细胞表

达。

[0140] 在一些实施方案中,异源转基因编码抗肌萎缩蛋白。在一些实施方案中,抗肌萎缩蛋白是人抗肌萎缩蛋白编码序列,包括但不限于由人抗肌萎缩蛋白基因表达的任何编码序列。人抗肌萎缩蛋白基因(例如GenBank Entrez Gene ID 1756)又称为DMD、BMD、CMD3B、MRX85、DXS142、DXS164、DXS206、DXS230、DXS239、DXS268、DXS269、DXS270和DXS272。在一些实施方案中,抗肌萎缩蛋白具有人抗肌萎缩蛋白的氨基酸序列(例如以GenBank登录号AAA53189代表)。可以使用编码抗肌萎缩蛋白的异源转基因,例如以在患有迪谢内(Duchenne)或贝克(Becker)肌营养不良(与抗肌萎缩蛋白中的突变有关的隐性X连锁的肌营养不良)的个体中表达抗肌萎缩蛋白。贝克肌营养不良症是一种由抗肌萎缩蛋白中的功能缺失突变引起的不太严重的病症,而迪谢内肌营养不良症与抗肌萎缩蛋白中的更严重的功能缺失或无效突变(例如无义或移码突变)有关。已知抗肌萎缩蛋白在抗肌萎缩蛋白-糖蛋白复合物(DGC)中发挥功能,需要所述抗肌萎缩蛋白-糖蛋白复合物(DGC)以将肌细胞的F-肌动蛋白连接于胞外基质,由此在肌肉收缩和松弛期间稳定肌膜。

[0141] 在一些实施方案中,异源转基因编码囊性纤维化跨膜传导调节蛋白(CFTR),又称为ATP结合盒亚家族C,成员7。在一些实施方案中,CFTR是人CFTR编码序列,包括但不限于由人CFTR基因表达的任何编码序列。人CFTR基因(例如GenBank Entrez Gene ID 1080)又称为CF、MRP7、ABC35、ABCC7、CFTR/MRP、TNR-CFTR和dj760C5.1。在一些实施方案中,CFTR具有人CFTR的氨基酸序列(例如以GenBank登录号NP_000483代表)。可以使用编码CFTR的异源转基因,例如以在患有囊性纤维化(一种与CFTR中的突变有关的常染色体隐性病症,其影响肺、胰腺、肠和许多其它器官)的个体中表达CFTR。已知CFTR作为参与Cl⁻离子转运的ATP门控离子通道发挥功能。足够的CFTR功能的缺乏导致多种病理学;一个例子是穿过上皮细胞的离子转运被破坏,导致增加的细胞性水吸收和肺和其它组织中粘液的病理性增稠和积累。

[0142] 在一些实施方案中,治疗性转基因产物是治疗性核酸。在一些实施方案中,治疗性核酸可以包括但不限于siRNA、shRNA、RNAi、miRNA、反义RNA、核酶或DNAzyme。因而,治疗性核酸可以编码RNA,其在从载体的核酸转录时能通过干扰与病症有关的异常或过多蛋白质的翻译或转录来治疗病症。例如,异源转基因可以编码RNA,其通过高度特异性消除或减少编码异常和/或过多蛋白质的mRNA治疗病症。治疗性RNA序列包括RNAi、小抑制性RNA(siRNA)、微小RNA(miRNA)、和/或核酶(如锤头和发夹核酶),其能通过高度特异性消除或减少编码异常和/或过多蛋白质的mRNA治疗病症。

[0143] 在一些实施方案中,异源转基因是人转基因。在一些实施方案中,异源转基因与启动子连接。在一些实施方案中,转基因(例如本文中描述的异源核酸)与启动子可操作连接。例示性的启动子包括但不限于巨细胞病毒(CMV)立即早期启动子、GUSB启动子、RSV LTR、MoMLV LTR、磷酸甘油酸激酶-1(PGK)启动子、猿病毒40(SV40)启动子和CK6启动子、转甲状腺素启动子(TTR)、TK启动子、四环素响应性启动子(TRE)、HBV启动子、hAAT启动子、LSP启动子、嵌合肝特异性启动子(LSP)、E2F启动子、端粒酶(hTERT)启动子;巨细胞病毒增强子/鸡β-肌动蛋白/兔β-珠蛋白启动子(CAG启动子;Niwa等,Gene,1991,108(2):193-9)和延长因子1-α启动子(EF1-α)启动子(Kim等,Gene,1990,91(2):217-23和Guo等,Gene Ther.,1996,3(9):802-10)。在一些实施方案中,启动子包含人β-葡糖醛酸糖苷酶启动子或与鸡β-

肌动蛋白 (CBA) 启动子连接的巨细胞病毒增强子。启动子可以是组成性、诱导型或阻遏性启动子。在一些实施方案中,启动子是小鼠转甲状腺素启动子。

[0144] 组成性启动子的例子包括但不限于逆转录病毒劳斯肉瘤病毒 (RSV) LTR启动子(任选地与RSV增强子一起)、巨细胞病毒 (CMV) 启动子(任选地与CMV增强子一起)[参见例如 Boshart等,Cell,41:521-530 (1985)]、SV40启动子、二氢叶酸还原酶启动子、13-肌动蛋白启动子、磷酸甘油激酶 (PGK) 启动子和EF1a启动子 (Invitrogen)。

[0145] 诱导型启动子容许调节基因表达,并且可以通过外源供应的化合物、环境因素如温度或特定生理状态的存在,例如急性阶段、细胞的特定分化状态,或者仅在复制细胞中调节。诱导型启动子和诱导型系统从多种商业来源可得到,包括但不限于Invitrogen、Clontech和Ariad。许多其它系统已经得到描述,并且可以是本领域技术人员容易选择的。通过外源供应的启动子调节的诱导型启动子的例子包括锌诱导型绵羊金属硫蛋白 (MT) 启动子、地塞米松 (dexamethasone) (Dex) 诱导型小鼠乳腺肿瘤病毒 (MMTV) 启动子、T7聚合酶启动子系统 (WO 98/10088);蜕皮激素昆虫启动子 (No等,Proc.Natl.Acad.Sci.USA,93:3346-3351 (1996))、四环素阻遏性系统 (Gossen等,Proc.Natl.Acad.Sci.USA,89:5547-5551 (1992))、四环素诱导型系统 (Gossen等,Science,268:1766-1769 (1995),还参见 Harvey等,Curr.Opin.Chem.Biol.,2:512-518 (1998))、RU486诱导型系统 (Wang等,Nat.Biotech.,15:239-243 (1997) 和Wang等,Gene Ther.,4:432-441 (1997)) 和雷帕霉素 (rapamycin) 诱导型系统 (Magari等,J.Clin.Invest.,100:2865-2872 (1997))。在此上下文中可为有用的其它类型的诱导型启动子是那些通过特定生理状态(例如温度、急性阶段、细胞的特定分化状态)或者仅在复制细胞中调节的。

[0146] 在另一个实施方案中,会使用转基因的天然启动子或其片段。当期望转基因的表达应当模拟天然表达时,天然启动子可以是优选的。当转基因的表达必须在时间上或发育上,或者以组织特异性方式,或者响应特定的转录刺激物而受调节时,可以使用天然启动子。在别的实施方案中,也可以使用其它天然表达控制元件,如增强子元件、多聚腺苷酸化位点或Kozak共有序列来模拟天然表达。

[0147] 在一些实施方案中,调节序列赋予组织特异性基因表达能力。在一些情况中,组织特异性调节序列结合以组织特异性方式诱导转录的组织特异性转录因子。例如,肝、肺、肌肉、肠、胰腺和/或其它组织中的组织特异性表达可为合意的。合适的组织特异性调节序列(例如启动子、增强子等)是本领域中公知的。例如,在一些实施方案中,启动子是小鼠转甲状腺素 (mTTR) 启动子,已知其驱动肝中的基因表达。

[0148] 在一些实施方案中,rAAV基因组包含内含子。在一些实施方案中,内含子是杂合内含子。

[0149] 在一些实施方案中,rAAV基因组包含多聚腺苷酸化信号。许多多聚腺苷酸化信号是本领域中已知的。在一些实施方案中,多聚腺苷酸化信号是合成多聚腺苷酸化信号。在其它实施方案中,多聚腺苷酸化信号是牛生长激素 (BGH) 多聚腺苷酸化信号。关于BGH多聚腺苷酸化信号的更为详细的描述,参见例如Goodwin,E.C.和Rottman,F.M. (1992) J.Biol.Chem.267:16330-16334。

[0150] 在一些实施方案中,本发明提供了包含超大AAV基因组的AAV颗粒,其中AAV基因组从5'至3'包含AAV2ITR、mTTR202启动子、杂合内含子、B域缺失的因子VIII转基因、合成多聚

腺苷酸化信号和AAV2ITR。在一些实施方案中,超大AAV基因组从5'至3'包含AAV2ITR、mTTR202opt启动子、杂合内含子、B域缺失的因子VIII转基因、合成多聚腺苷酸化信号和AAV2ITR。在一些实施方案中,超大AAV基因组从5'至3'包含AAV2ITR、mTTR482启动子、杂合内含子、B域缺失的因子VIII转基因、合成多聚腺苷酸化信号和AAV2ITR。在一些实施方案中,超大AAV基因组从5'至3'包含AAV2ITR、mTTR482启动子、杂合内含子、B域缺失的因子VIII转基因、牛生长激素多聚腺苷酸化信号和AAV2ITR。

[0151] 上文描述的rAAV基因组元件(例如启动子、内含子和多聚腺苷酸化信号)可以单独或在与本公开内容的异源转基因的任何组合中存在。rAAV基因组可以包含建立异源转基因表达的任何元件,例如启动子、异源核酸、ITR、核糖体结合元件、终止子、增强子、选择标志物、内含子、多聚腺苷酸化(polyA)信号、和/或复制起点。例如,在一些实施方案中,rAAV基因组含有异源转基因和一个或多个选自由下列组成之群组的元件:本公开内容的启动子、本公开内容的内含子、和本公开内容的多聚腺苷酸化信号。在一些实施方案中,rAAV基因组可以包含在异源转基因侧翼的至少一个ITR序列和一个或多个选自由下列组成之群组的元件:本公开内容的启动子、本公开内容的内含子、和本公开内容的多聚腺苷酸化信号。

[0152] 在一些实施方案中,超大rAAV载体是自身互补的rAAV载体,例如包含重组自身互补性(术语“自身互补”在本文中可互换使用)基因组的rAAV载体。具有自身互补性基因组的AAV病毒颗粒和使用自身互补性AAV基因组的方法记载于美国专利No.6,596,535;7,125,717;7,465,583;7,785,888;7,790,154;7,846,729;8,093,054;和8,361,457;及Wang Z.,等,(2003)Gene Ther10:2105-2111,每篇通过提述以其整体并入本文。包含自身互补性基因组的rAAV会凭借其部分互补序列(例如转基因的互补编码和非编码链)而快速形成双链DNA分子。在一些实施方案中,载体包含编码异源核酸的第一核酸序列和编码核酸的互补物的第二核酸序列,其中第一核酸序列可以沿着大部分或整个其长度与第二核酸序列形成链内碱基对。

[0153] 在一些实施方案中,通过突变ITR(例如正确的ITR)连接第一异源核酸序列和第二异源核酸序列。在一些实施方案中,ITR包含多核苷酸序列5'-CACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCACGCCCGGGCTTTGCCCGGGCG-3'(SEQ ID NO:24)。突变ITR包含含有末端解析序列的D区的缺失。结果,在复制AAV病毒基因组时,rep蛋白不会在突变ITR处切割病毒基因组,并且因而,会在病毒壳体中包装以5'至3'次序包含如下的重组病毒基因组:AAV ITR、包含调节序列的第一异源多核苷酸序列、突变AAV ITR、与第一异源多核苷酸为相反取向的第二异源多核苷酸和第三AAV ITR。在一些实施方案中,scAAV载体基因组大于约5.0kb、5.1kb、5.2kb、5.3kb、5.4kb、5.5kb、5.6kb、5.7kb、5.8kb、5.9kb、6.0kb、6.1kb、6.2kb、6.3kb、6.4kb、6.5kb、6.6kb、6.7kb、6.8kb、6.9kb或7.0kb中任一者或其间的任何数值。

[0154] IV.产生病毒颗粒的方法

[0155] 本公开内容的某些态样涉及用于产生含有超大重组AAV基因组的腺相关病毒(AAV)颗粒的方法。在一些实施方案中,方法包括在产生rAAV颗粒的条件下培养AAV生产细胞系,其中AAV生产细胞系包含i)编码AAV rep和cap基因的核酸,和ii) rAAV基因组,其中rAAV基因组是约4.7kb-约9.4kb,任选地约4.7kb-6.7kb;b)提供AAV辅助功能;并且c)收集含有超大rAAV基因组的rAAV颗粒。在一些实施方案中,AAV生产细胞系包含稳定维持的编码

AAV rep和cap基因的核酸。在一些实施方案中,AAV生产细胞系包含稳定维持的rAAV基因组,其中rAAV基因组是约4.7kb-约9.4kb,任选地约4.7kb-6.7kb。在一些实施方案中,AAV生产细胞系包含稳定维持的编码AAV rep和cap基因的核酸和稳定维持的rAAV基因组,其中rAAV基因组是约4.7kb-约9.4kb,任选地约4.7kb-6.7kb。在一些实施方案中,AAV生产细胞系包含稳定整合到细胞系基因组中的编码AAV rep和cap基因的核酸。在一些实施方案中,AAV生产细胞系包含稳定整合到细胞系基因组中的rAAV基因组,其中rAAV基因组是约4.7kb-约9.4kb,任选地约4.7kb-6.7kb。在上述实施方案的一些实施方案中,rAAV基因组的长度大于约5.0kb、5.1kb、5.2kb、5.3kb、5.4kb、5.5kb、5.6kb、5.7kb、5.8kb、5.9kb、6.0kb、6.1kb、6.2kb、6.3kb、6.4kb、6.5kb、6.6kb、6.7kb、6.8kb、6.9kb、7.0kb、7.1kb、7.2kb、7.3kb、7.4kb、7.5kb、7.6kb、7.7kb、7.8kb、7.9kb、8.0kb、8.1kb、8.2kb、8.3kb、8.4kb、8.5kb、8.6kb、8.7kb、8.8kb、8.9kb、9.0kb、9.2kb、9.3kb或9.4kb中任一者或其间的任何数值。在一些实施方案中,包装的AAV基因组不含5'端的截短。在一些实施方案中,包装的AAV基因组不含3'端的截短。

[0156] 本公开内容的其它态样涉及用于产生包含超大重组AAV基因组的腺相关病毒(AAV)颗粒的细胞系,所述细胞系包含a)编码AAV rep和cap基因的核酸,和b) rAAV基因组,其中rAAV基因组是约4.7kb-约9.4kb,任选地约4.7kb-6.7kb。在一些实施方案中,AAV生产细胞系包含稳定维持的编码AAV rep和cap基因的核酸。在一些实施方案中,AAV生产细胞系包含稳定维持的rAAV基因组,其中rAAV基因组是约4.7kb-约9.4kb,任选地约4.7kb-约6.7kb或约5.2kb-约8.7kb。在一些实施方案中,AAV生产细胞系包含稳定维持的编码AAV rep和cap基因的核酸和稳定维持的rAAV基因组,其中rAAV基因组是约4.7kb-约9.4kb,任选地约4.7kb-约6.7kb或约5.2kb-约8.7kb。在一些实施方案中,AAV生产细胞系包含稳定整合到细胞系基因组中的编码AAV rep和cap基因的核酸。在一些实施方案中,AAV生产细胞系包含稳定整合到细胞系基因组中的rAAV基因组,其中rAAV基因组是约4.7kb-约9.4kb,任选地约4.7kb-约6.7kb或约5.2kb-约8.7kb。在一些实施方案中,AAV生产细胞系包含稳定整合到细胞系基因组中的编码AAV rep和cap基因的核酸和rAAV基因组,其中rAAV基因组是约4.7kb-约9.4kb,任选地约4.7kb-约6.7kb或约5.2kb-约8.7kb。在一些实施方案中,rAAV基因组的长度大于约5.0kb、5.1kb、5.2kb、5.3kb、5.4kb、5.5kb、5.6kb、5.7kb、5.8kb、5.9kb、6.0kb、6.1kb、6.2kb、6.3kb、6.4kb、6.5kb、6.6kb、6.7kb、6.8kb、6.9kb、7.0kb、7.1kb、7.2kb、7.3kb、7.4kb、7.5kb、7.6kb、7.7kb、7.8kb、7.9kb、8.0kb、8.1kb、8.2kb、8.3kb、8.4kb、8.5kb、8.6kb、8.7kb、8.8kb、8.9kb、9.0kb、9.2kb、9.3kb或9.4kb中任一者或其间的任何数值。

[0157] 本领域中已知许多用于产生rAAV载体的方法,包括转染、稳定的细胞系生产、和感染性杂合病毒生产系统,其包括腺病毒-AAV杂合物、疱疹病毒-AAV杂合物(Conway, JE等, (1997) J. Virology 71 (11):8780-8789)和杆状病毒-AAV杂合物。用于产生rAAV病毒颗粒的rAAV产生培养都需要:1)合适的宿主细胞,2)合适的辅助病毒功能,3) AAV rep和cap基因和基因产物;4)侧翼有至少一个AAV ITR序列(例如超大rAAV载体基因组)的核酸(如治疗性核酸);和5)合适的培养基和培养基组分以支持rAAV产生。在一些实施方案中,合适的宿主细

胞是灵长类宿主细胞。在一些实施方案中,合适的宿主细胞是人衍生细胞系,如HeLa、A549、293或Perc.6细胞。在一些实施方案中,合适的辅助病毒功能由野生型或突变体腺病毒(如温度敏感性腺病毒)、疱疹病毒(HSV)、杆状病毒或提供辅助功能的质粒构建体提供。在一些实施方案中,AAV rep和cap基因产物可以来自任何AAV血清型。一般地,但不强制,AAV rep基因产物与rAAV载体基因组的ITR是相同血清型的,只要rep基因产物可以发挥功能以复制并包装rAAV基因组。本领域中已知的合适的培养基可以用于产生rAAV载体。这些培养基包括但不限于由Hyclone Laboratories和JRH生产的培养基,包括改良的伊格尔(Eagle)培养基(MEM)、Dulbecco氏改良的伊格尔培养基(DMEM)、定制配制物,如记载于美国专利No.6,566,118的那些,和如记载于美国专利No.6,723,551的Sf-900II SFM培养基,每篇通过提述以其整体并入本文,特别就在产生重组AAV载体中使用的定制培养基配制物而言。在一些实施方案中,AAV辅助功能由腺病毒或HSV提供。在一些实施方案中,AAV辅助功能由杆状病毒提供,并且宿主细胞是昆虫细胞(例如草地贪夜蛾(*Spodoptera frugiperda*) (Sf9)细胞)。

[0158] 一种用于产生rAAV颗粒的方法是三重转染方法。简言之,可以将含有rep基因和壳体基因的质粒以及辅助腺病毒质粒转染(例如使用磷酸钙方法)到细胞系(例如HEK-293细胞)中,并且可以将病毒收集,并且任选地纯化。如此,在一些实施方案中,通过将编码rAAV载体的核酸、编码AAV rep和cap的核酸、和编码AAV辅助病毒功能的核酸三重转染到宿主细胞中产生rAAV颗粒,其中将核酸转染到宿主细胞生成能够产生rAAV颗粒的宿主细胞。

[0159] 在一些实施方案中,可以通过生产细胞系方法,如下文提供的例示性生产细胞系方法(还可参见Martin等,(2013) Human Gene Therapy Methods 24:253-269; U.S. PG Pub.No.US2004/0224411; 及Liu,X.L.等(1999) Gene Ther. 6:293-299)产生rAAV颗粒。简言之,可以用质粒稳定转染细胞系(例如HeLa、293、A549或Perc.6细胞系),所述质粒含有rep基因、壳体基因、和包含启动子-异源核酸序列的超大载体基因组。可以筛选细胞系以选择用于rAAV产生的前导克隆(lead clone),然后将所述前导克隆扩充到生产生物反应器,并且用辅助病毒(例如腺病毒或HSV)感染以启动rAAV产生。随后,可收获病毒,可灭活(例如通过加热)和/或除去腺病毒,并且可纯化rAAV颗粒。如此,在一些实施方案中,通过生产细胞系产生rAAV颗粒,所述生产细胞系包含下列一者或多者:编码rAAV载体的核酸、编码AAV rep和cap的核酸、和编码AAV辅助病毒功能的核酸。如本文中描述,与三重转染方法相比,生产细胞系方法对于产生具有超大基因组的rAAV颗粒可以是有利的。

[0160] 在一些实施方案中,编码AAV rep和cap基因的核酸和/或rAAV基因组在生产细胞系中稳定维持。在一些实施方案中,在一个或多个质粒上将编码AAV rep和cap基因的核酸和/或rAAV基因组导入细胞系中以生成生产细胞系。在一些实施方案中,在同一质粒上将AAV rep、AAV cap和rAAV基因组导入细胞中。在其它实施方案中,在不同质粒上将AAV rep、AAV cap和rAAV基因组导入细胞中。在一些实施方案中,用质粒稳定转染的细胞系将质粒维持多次细胞系传代(例如5、10、20、30、40、50或超过50次细胞传代)。例如,质粒可以随细胞复制而复制,或者质粒可以整合到细胞基因组中。已经鉴定出使质粒能够在细胞(例如人细胞)中自主复制的多种序列(参见例如Krysan,P.J.等(1989) Mol.Cell Biol. 9:1026-1033)。在一些实施方案中,质粒可以含有选择标志物(例如抗生素抗性标志物),其容许选择维持质粒的细胞。哺乳动物细胞中通常使用的选择标志物包括但不限于杀稻瘟素(blasticidin)、G418、潮霉素B、zeocin、嘌呤霉素及它们的衍生物。用于将核酸导入细胞中

的方法是本领域中已知的,并且包括但不限于病毒转导、阳离子转染(例如使用阳离子聚合物,如DEAE-右旋糖苷或阳离子脂质如lipofectamine)、磷酸钙转染、显微注射、颗粒轰击、电穿孔、和纳米颗粒转染(关于更多详情,参见例如Kim,T.K.和Eberwine,J.H. (2010) Anal.Bioanal.Chem.397:3173-3178)。

[0161] 在一些实施方案中,编码AAV rep和cap基因的核酸和/或rAAV基因组稳定整合入生产细胞系的基因组中。在一些实施方案中,在一个或多个质粒上将编码AAV rep和cap基因的核酸和/或rAAV基因组导入细胞系中以生成生产细胞系。在一些实施方案中,在同一质粒上将AAV rep、AAV cap和rAAV基因组导入细胞中。在其它实施方案中,在不同质粒上将AAV rep、AAV cap和rAAV基因组导入细胞中。在一些实施方案中,质粒可以含有选择标志物(例如抗生素抗性标志物),其容许选择维持质粒的细胞。用于将核酸稳定整合到多种宿主细胞系中的方法是本领域中已知的(关于通过核酸的稳定整合创建的例示性生产细胞系的更为详细的描述,参见下文实施例)。例如,可以使用重复选择(例如经由使用选择性标志物)来选择已经整合含有选择性标志物(和AAV cap和rep基因和/或rAAV基因组)的核酸的细胞。在其它实施方案中,可以以位点特异性方式将核酸整合到细胞系中以生成生产细胞系。几种位点特异性重组系统是本领域中已知的,如FLP/FRT(参见例如O' Gorman,S.等(1991) Science 251:1351-1355)、Cre/loxP(参见例如Sauer,B.和Henderson,N. (1988) Proc.Natl.Acad.Sci.85:5166-5170)、和phi C31-att(参见例如Groth,A.C.等(2000) Proc.Natl.Acad.Sci.97:5995-6000)。

[0162] 在一些实施方案中,生产细胞系源自灵长类细胞系(例如非人灵长类细胞系,如Vero或FRhL-2细胞系)。在一些实施方案中,细胞系源自人细胞系。在一些实施方案中,生产细胞系源自HeLa、293、A549或PERC.6® (CruCell) 细胞。例如,在将编码AAV rep和cap基因的核酸和/或超大rAAV基因组导入和/或稳定维持/整合到细胞系中以生成生产细胞系前,细胞系是HeLa、293、A549或PERC.6® (CruCell) 细胞系或其衍生物。

[0163] 在一些实施方案中,生产细胞系适合于在悬浮液中生长。如本领域中已知的,锚定(anchorage)依赖性细胞通常不能在没有基底(如微载体珠)的情况下在悬浮液中生长。使细胞系适合于悬浮生长可以包括例如在具有搅拌桨的旋动培养中培养细胞系,使用缺乏钙和镁离子的培养基以防止结块(和任选地消泡剂),使用用硅化合物包被的培养容器,并且在每次传代时选择培养物中(而不是大块中或容器的侧面上)的细胞。关于进一步的描述,参见例如ATCC常问问题文件(在www.atcc.org/Global/FAQs/9/1/Adapting%20a%20monolayer%20cell%20line%20to%20suspension-40.aspx可得到) 及其中引用的参考文献。

[0164] 在一些态样中,提供了用于产生如本文中公开的任何rAAV颗粒的方法,其包括(a) 在产生rAAV颗粒的条件下培养宿主细胞,其中宿主细胞包含(i) 一种或多种AAV包装基因,其中每种所述AAV包装基因编码AAV复制和/或壳体化蛋白;(ii) rAAV原载体,其包含编码侧翼有至少一个AAV ITR的如本文中描述的异源核酸的核酸,和(iii) AAV辅助功能;和(b) 回收由宿主细胞产生的rAAV颗粒。在一些实施方案中,所述至少一种AAV ITR选自由下列组成之群组:AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10、AAVrh10、AAV11、AAV12、AAV2R471A、AAV DJ、山羊AAV、牛AAV或小鼠AAV血清型ITR等。例如,在一些实施方案中,AAV血清型是AAV1、AAV2、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、

AAV9、AAV10或AAVrh10。在某些实施方案中, AAV中的核酸包含AAV2ITR。在一些实施方案中, 所述壳体化蛋白选自由下列组成之群组: AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10、AAVrh10、AAV11、AAV12、AAV2R471A、AAV2/2-7m8、AAV DJ、AAV2N587A、AAV2E548A、AAV2N708A、AAV V708K、山羊AAV、AAV1/AAV2嵌合、牛AAV或小鼠AAV壳体rAAV2/HBoV1血清型壳体蛋白或其突变体。在一些实施方案中, 壳体化蛋白是AAV5壳体蛋白, 包括具有酪氨酸壳体突变的AAV5壳体蛋白。在一些实施方案中, 壳体化蛋白是AAV5壳体蛋白, 包括具有酪氨酸壳体突变的AAV5壳体蛋白, 并且ITR是AAV2ITR。在别的实施方案中, rAAV颗粒包含来自A-F进化枝的AAV血清型的壳体蛋白。在一些实施方案中, AAV血清型是AAV1、AAV2、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10或AAVrh10。在一些实施方案中, rAAV颗粒包含AAV血清型1 (AAV1) 壳体。在一些实施方案中, rAAV颗粒包含AAV血清型2 (AAV2) 壳体。在一些实施方案中, rAAV颗粒包含AAVrh8R壳体或其突变体。在一些实施方案中, rAAV颗粒包含AAV1壳体和重组基因组, 其包含AAV2ITR、突变体AAV2ITR和编码治疗性转基因/核酸的核酸。在一些实施方案中, AAV ITR是AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAVrh8、AAVrh8R、AAV9、AAV10、AAVrh10、AAV11、AAV12、AAV2R471A、AAV DJ、山羊AAV、牛AAV或小鼠AAV血清型ITR。在某些实施方案中, AAV ITR是AAV2ITR。在一些实施方案中, ITR源自AAV2, 并且壳体源自AAV2。在一些实施方案中, ITR源自AAV2, 并且壳体源自AAVrh8R。

[0165] 本发明的合适的rAAV生产培养基可以补充有0.5%–20% (v/v或w/v) 水平的血清或血清衍生的重组蛋白。或者, 如本领域中已知的, 可在无血清的条件中产生rAAV载体, 所述无血清的条件也可以称为没有动物衍生产物的培养基。本领域普通技术人员可以理解, 设计为支持rAAV载体产生的商业或定制培养基也可以补充有本领域中已知的一种或多种细胞培养组分, 包括但不限于葡萄糖、维生素、氨基酸和或生长因子, 以增加生产培养物中的rAAV滴度。

[0166] 可以在多种适合于利用的特定宿主细胞的条件下(在较宽的温度范围里, 持续可变的时间长度, 等等) 下培养rAAV生产培养物。如本领域中已知的, rAAV生产培养物包括附着依赖性培养物, 其可以在合适的附着依赖性容器(如例如滚瓶、中空纤维滤器、微载体、和填充床或流化床生物反应器) 中培养。rAAV载体生产培养物也可以包括悬浮适应性宿主细胞, 如HeLa、293和SF-9细胞, 所述宿主细胞可以以多种方式培养, 包括例如旋转瓶、搅拌釜生物反应器、和一次性系统, 如Wave袋系统。

[0167] 本公开内容的某些态样涉及收集含有超大rAAV基因组的rAAV颗粒。可以如下从rAAV生产培养物中收获本发明的rAAV载体颗粒, 即裂解生产培养物的宿主细胞或者从生产培养物中回收用完的培养基, 只要在本领域中已知的条件下培养细胞以引起rAAV颗粒从完整细胞中释放到培养基中, 如更完整记载于美国专利No. 6, 566, 118)。裂解细胞的合适方法也是本领域中已知的, 并且包括例如多个冷冻/融化循环、声处理、微流化和用化学品如去污剂和/或蛋白酶处理。

[0168] 在一些实施方案中, 收集的AAV颗粒含有大于约5.0kb的rAAV基因组。在一些实施方案中, 收集的rAAV颗粒含有长度大于约5.0kb、5.1kb、5.2kb、5.3kb、5.4kb、5.5kb、5.6kb、5.7kb、5.8kb、5.9kb、6.0kb、6.1kb、6.2kb、6.3kb、6.4kb、6.5kb、6.6kb、6.7kb、6.8kb、6.9kb或7.0kb、8.0kb或9.0kb中任一者或长度为其间的任何数值的rAAV基因组。在一些实施方案

中,收集的rAAV颗粒含有约5.0kb-约9.0kb、约5.0kb-约8.5kb、约5.0kb-约8.0kb、约5.0kb-约7.5kb、约5.0kb-约7.0kb、约5.0kb-约6.5kb、约5.0kb-约6.0kb、约5.0kb-约5.5kb、约5.2kb-约9.0kb、约5.2kb-约8.5kb、约5.2kb-约8.0kb、约5.2kb-约7.5kb、约5.2kb-约7.0kb、约5.2kb-约6.5kb、约5.2kb-约6.0kb、约5.2kb-约5.5kb、约5.5kb-约9.0kb、约5.5kb-约8.5kb、约5.5kb-约8.0kb、约5.5kb-约7.5kb、约5.5kb-约7.0kb、约5.5kb-约6.5kb、约5.5kb-约6.0kb、约6.0kb-约9.0kb、约6.0kb-约8.5kb、约6.0kb-约8.0kb、约6.0kb-约7.5kb、约6.0kb-约7.0kb、约6.0kb-约6.5kb、约6.5kb-约9.0kb、约6.5kb-约8.5kb、约6.5kb-约7.5kb、约6.5kb-约7.0kb、约7.0kb-约9.0kb、约7.0kb-约8.5kb、约7.0kb-约8.0kb、约7.0kb-约7.5kb、约7.5kb-约9.0kb、约7.5kb-约8.5kb、约7.5kb-约8.0kb、约8.0kb-约9.0kb、约8.0kb-约8.5kb或约8.5kb-约9.0kb中任一项的rAAV基因组。在一些实施方案中,收集的rAAV颗粒含有约4.7kb-约9.4kb,任选地约4.7kb-约6.7kb或约5.2kb-约8.7kb的rAAV基因组。

[0169] 在一些实施方案中,在提供辅助功能后约48小时-约96小时收集rAAV颗粒。例如,在一些实施方案中,在提供辅助功能后约48小时、约60小时、约72小时、约84小时或约96小时收集rAAV颗粒。在一些实施方案中,在提供辅助功能后约48小时-约96小时、约48小时-约84小时、约48小时-约72小时、约48小时-约60小时、约60小时-约96小时、约60小时-约84小时、约60小时-约72小时、约72小时-约96小时、约72小时-约84小时或约84小时-约96小时收集rAAV颗粒。

[0170] 在别的实施方案中,纯化rAAV颗粒。如本文中使用的,术语“纯化的”包括缺乏至少一些其它组分的rAAV颗粒的制备物,所述其它组分在天然存在或最初制备rAAV颗粒的地方也可以存在。因此,例如,分离的rAAV颗粒可以使用纯化技术从来源混合物,如培养物裂解物或生产培养物上清液中富集它来制备。可以以多种方式,如例如通过DNA酶抗性颗粒(DRP)的比例或溶液中存在的基因组拷贝(gc),或者通过感染性来测量富集,或者它可以相对于来源混合物中存在的第二潜在干扰性物质,如污染物,包括生产培养物污染物或过程中污染物,包括辅助病毒、培养基组分等来测量。

[0171] 在一些实施方案中,使rAAV生产培养收获物澄清以除去宿主细胞碎片。在一些实施方案中,通过过滤通过一系列深度滤器,包括例如DOHC Millipore Millistak+HC Pod级滤器、AlHC Millipore Millistak+HC Pod级滤器、和0.2 μ m Filter Opticap XL10 Millipore Express SHC Hydrophilic Membrane滤器使生产培养收获物澄清。也可以通过本领域中已知的多种其它标准技术,如离心或过滤通过本领域中已知的0.2 μ m以上孔径的任何乙酸纤维素滤器实现澄清。

[0172] 在一些实施方案中,用Benzonase[®]进一步处理rAAV生产培养收获物以消化生产培养物中存在的任何高分子量DNA。在一些实施方案中,在本领域中已知的标准条件下进行Benzonase[®]消化,所述标准条件包括例如于范围为环境温度至37℃的温度,终浓度1-2.5个单位/ml Benzonase[®]进行30分钟至几小时的时间。

[0173] 可以使用一个或多个下述纯化步骤分离或纯化rAAV颗粒:平衡离心;流过阴离子交换过滤(flow-through anionic exchange filtration);用于浓缩rAAV颗粒的切向流过滤(TFF);通过磷灰石层析的rAAV捕捉;辅助病毒的热灭活;通过疏水性相互作用层析的

rAAV捕捉;通过大小排阻层析(SEC)的缓冲液交换;纳米过滤;和通过阴离子交换层析、阳离子交换层析或亲和层析的rAAV捕捉。在一些实施方案中,纯化包括一个或多个层析步骤(例如上文描述的一个或多个层析步骤)。可以单独、以各种组合或以不同次序使用这些步骤。在一些实施方案中,方法包括以如下文描述的次序的所有步骤。纯化rAAV颗粒的方法参见例如Xiao等,(1998) *Journal of Virology* 72:2224-2232;美国专利号6,989,264和8,137,948;及WO 2010/148143。

[0174] 在一些实施方案中,rAAV颗粒在医药组合物中。在一些实施方案中,rAAV颗粒在包含医药上可接受之赋形剂的医药组合物中。如本领域中公知的,医药上可接受之赋形剂是相对惰性的物质,其促进药理学有效物质的授予,并且可以作为液体溶液或悬浮液,作为乳剂,或者作为适合于在使用前在液体中溶解或悬浮的固体形式供应。例如,赋形剂可以给出形式或一致性,或者充当稀释剂。合适的赋形剂包括但不限于稳定剂、润湿和乳化剂、用于改变渗量(osmolality)的盐、成胶囊剂(encapsulating agent)、pH缓冲物质、和缓冲剂。此类赋形剂包括可以在没有过度毒性的情况下授予的适合于对靶组织投递的任何药用剂。医药上可接受之赋形剂包括但不限于山梨糖醇、各种TWEEN化合物之任一者、和液体,如水、盐水、甘油和乙醇。医药上可接受之盐可以包含于其中,例如矿物酸盐,如氢氯酸盐、氢溴酸盐、磷酸盐、硫酸盐等;和有机酸盐,如乙酸盐、丙酸盐、丙二酸盐、苯甲酸盐等。医药上可接受之赋形剂的彻底讨论在REMINGTON'S PHARMACEUTICAL SCIENCES (Mack Pub.Co., N.J.1991)中可得到。

[0175] 此类医药上可接受之载体可以是无菌液体,如水和油,包括石油、动物、植物或合成起源的那些油,如花生油、大豆油、矿物油等。也可以采用盐水溶液和水性右旋糖、聚乙二醇(PEG)和甘油溶液作为液体载体,特别用于可注射溶液。医药组合物可以进一步包含其它成分,例如防腐剂、缓冲剂、张度剂、抗氧化剂和稳定剂、非离子型湿润或澄清剂、增粘剂等。本文中描述的医药组合物可以以单一单位剂量或以多剂量形式包装。组合物一般配制为无菌且基本上等张的溶液。

[0176] V. 治疗方法

[0177] 在一些态样中,本发明提供了治疗有此需要的个体中的疾病或病症的治疗方法,其包括对个体授予AAV颗粒。可以对感兴趣的特定组织授予AAV颗粒,或者它可以系统性授予。在一些实施方案中,可以胃肠外授予有效量的AAV颗粒。胃肠外授予路径可以包括但不限于静脉内、骨内、动脉内、脑内、肌肉内、鞘内、皮下、脑室内等。在一些实施方案中,可以由一种授予路径授予有效量的AAV颗粒。在一些实施方案中,可以由超过一种授予路径的组合授予有效量的AAV颗粒。在一些实施方案中,个体是哺乳动物。在一些实施方案中,个体是人。

[0178] 根据治疗的目的,授予有效量的包含超大AAV基因组的AAV颗粒。例如,在低百分比的转导可以实现期望的治疗效果的情况下,则治疗的目的一般满足或超过此转导水平。在一些情况中,可以通过转导期望组织类型的仅约1至5%靶细胞,在一些实施方案中期望组织类型的至少约20%细胞,在一些实施方案中期望组织类型的至少约50%,在一些实施方案中至少约80%,在一些实施方案中至少约95%,在一些实施方案中至少约99%细胞来实现此转导水平。作为指导,每次注射授予的颗粒的数目一般是约 1×10^6 -约 1×10^{14} 个颗粒、约 1×10^7 - 1×10^{13} 个颗粒、约 1×10^9 - 1×10^{12} 个颗粒或约 1×10^9 个颗粒、约 1×10^{10} 个颗粒或

约 1×10^{11} 个颗粒。可以在相同的方法期间或者间隔几天、几周、几个月或几年通过一次或多次授予来授予rAAV组合物。可以使用本文中描述的任何一种或多种授予路径。在一些实施方案中,可以使用多种载体来治疗人。

[0179] 鉴定通过AAV病毒颗粒转导的细胞的方法是本领域中已知的;例如免疫组织化学或使用标志物如增强型绿色荧光蛋白可以用于检测病毒颗粒的转导;例如,包含具有一个或多个氨基酸取代的rAAV壳体的病毒颗粒。

[0180] 在一些实施方案中,同时或序贯对超过一个位置授予包含超大AAV基因组的AAV病毒颗粒。在一些实施方案中,rAAV病毒颗粒的多次注射相隔不超过1小时、2小时、3小时、4小时、5小时、6小时、9小时、12小时或24小时。

[0181] 在一些实施方案中,本发明提供了治疗个体中的疾病或病症的方法,其包括授予包含超大AAV基因组的AAV颗粒,其中超大AAV基因组包含适合于治疗疾病或病症的转基因。在一些实施方案中,本发明提供了用包含超大AAV基因组的AAV颗粒治疗血友病A的方法,所述超大AAV基因组编码因子VIII转基因(例如人因子VIII转基因)。在一些实施方案中,本发明提供了用于用包含超大AAV基因组的AAV颗粒治疗肌营养不良症的方法,所述超大AAV基因组编码抗肌萎缩蛋白转基因(例如人抗肌萎缩蛋白转基因)。在一些实施方案中,本发明提供了用于用包含超大AAV基因组的AAV颗粒治疗dysferlinopathy的方法,所述超大AAV基因组编码dysferlin转基因(例如人dysferlin转基因)。在一些实施方案中,本发明提供了用于用包含超大AAV基因组的AAV颗粒治疗囊性纤维化的方法,所述超大AAV基因组编码CFTR转基因(例如人CFTR转基因)。然而,本发明不限于需要表达比4.8kb AAV载体基因组中所适合的更大的转基因的疾病或病症。例如,在一些实施方案中,本发明提供了包含AAV基因组的AAV颗粒,所述AAV基因组包含一种或多种异源转基因,其中异源转基因和调节因子(启动子、增强子、内含子等)的组合导致大于约5.0kb的AAV基因组。

[0182] VI. 套组

[0183] 在一些实施方案中,本发明包括包含AAV颗粒的套组,所述AAV颗粒包含本发明的超大基因组。在一些实施方案中,套组进一步包括用于投递(例如胃肠外授予)rAAV颗粒组合物的装置。在一些实施方案中,使用说明书包括根据本文中描述的方法之一的说明书。在一些实施方案中,在给容器提供(例如附于容器)的标签上印刷说明书。在一些实施方案中,使用说明书包括用于治疗疾病或病症的说明书。

[0184] 在一些实施方案中,套组包含单一流体(例如包含有效量的载体的医药上可接受之流体)。在一些实施方案中,套组包含2种以上流体。流体可以包括稀释剂、缓冲剂、赋形剂或本文中描述或本领域中已知的任何其它液体,其适合于投递、稀释、稳定、缓冲或以其它方式转运本公开内容的AAV颗粒。在一些实施方案中,系统包含一种或多种缓冲剂,例如水性pH缓冲溶液。缓冲剂的例子可以包括但不限于磷酸盐、柠檬酸盐、Tris、HEPES、和其它有机酸缓冲剂。

[0185] 在一些实施方案中,套组包含容器。合适的容器可以包括例如小瓶、袋、注射器和瓶。容器可以由一种或多种材料,如玻璃、金属或塑料制成。在一些实施方案中,容器用于容纳本公开内容的rAAV组合物。在一些实施方案中,容器也可以容纳流体和/或其它治疗剂。

实施例

[0186] 通过参考以下实施例会更充分理解本发明。然而,它们不应解释为限制本发明的范围。应当理解,本文中描述的实施例和实施方案是仅为了说明目的,并且鉴于其的各种修改或变化会对本领域技术人员有提示,并且应当包括在本申请的精神和范围及所附权利要求要求的范围内。

[0187] 实施例1:具有超大5.1kb和5.4kb FVIII载体的生产细胞系的生成

[0188] 如上文讨论,需要能够以高产率、一致产物、和高质量基因组产生具有超大基因组的rAAV载体的平台。本实施例描述了生成生产细胞系(PCL)平台,其对于产生具有含有大构建体(例如超过5kb)的基因组的rAAV载体是特别有利的。

[0189] 方法

[0190] 用于超大5.1和5.4kb FVIII载体的pTP质粒的构建

[0191] 在基于pUC57的质粒中生成FVIII表达盒,并且该FVIII表达盒由小鼠转甲状腺素(mTTR)启动子(Costa,RH等,Mol Cell Biol 1986,6:4697-4708.) (具有和没有100bp增强子序列的202bp核心序列)、杂合内含子(Jiang,H.等,Blood 2006 108:107-115)、经密码子优化的人B域缺失的FVIII cDNA、合成或BGH多聚A和rAAV2反向末端重复序列组成。这些生成具有范围为5.1和5.4kb的载体基因组大小的rAAV载体(图1A)。

[0192] 通过对正常C57BL/6小鼠的高体积注射在体内对具有FVIII表达盒的质粒载体测试FVIII产生。为了生成用于AAVrh8R/5.1kb FVIII载体的生产细胞系(PCL),用BglII消化TriplePlay质粒pAFTGEN-SEAP-caprh8R,并且进行平端处理。使用PvuI和SapI位点从pUC57-mTTR-hFVIIIco (pITR-mTTR-hFVIIIISQco-SpA) 中切出具有侧翼5'和3' AAV2ITR的FVIII载体基因组。将5.5kb PvuI/SapI平端化片段连接到TriplePlay质粒以生成具有5.1kb FVIII载体和AAVrh8R cap基因的质粒。生成含有AAV8cap基因的相似构建体。通过用牛生长激素(BGH)多聚A替换合成的多聚A区而生成具有AAVrh8R cap基因和5.4kb载体的TriplePlay质粒。将所得的卡那霉素抗性克隆转染到Huh7细胞中以测试FVIII蛋白产生。

[0193] 通过标准ELISA测定法定量培养基中的FVIII水平确认来自选定的TriplePlay质粒的FVIII产生。通过将pAdhelper共转染到293细胞中测试来自选定的TriplePlay质粒(pTGEN/AAVrh8R/mTTRhFVIII或pTGEN/AAV8/mTTRhFVIII)的rAAV载体生成。收获细胞裂解物,并且进行用FVIII引物/探针的qPCR以量化包装基因组的量。

[0194] 本文中呈现的实施例中使用的引物和探针见于表1。

[0195] 用于5.1kb和5.4kb FVIII载体的生产细胞系的生成

[0196] 用Lipofectamine和Plus试剂将质粒pTGEN/AAVrh8R/mTTR-hFVIII (具有5.1或5.4kb载体)或质粒pTGEN/AAV8/mTTR-hFVIII转染到HeLaS3细胞中。将细胞分配到60×96孔板上,并且将板清洗并且每周补料。在选择后,就集落生长对板评分。收获主孔(MW)并转移到24孔皿,并基于大小收获到24孔皿中。

[0197] 接着,将主孔分配到96孔板上以进行相对生产(RP)筛选,然后对来自RP筛选的用于载体产生的阳性MW测试比生产(SP)水平;例如通过qPCR经由载体产生(Martin,J.等,2013Hum.Gene Ther.Meth.24:253-269)。

[0198] 具有5.1kb和5.4kb FVIII载体的生产细胞系的基因组DNA的表征

[0199] 通过使用对每种序列特异性的引物和探针的qPCR对基因组DNA分析载体、rep和嘌呤霉素序列的拷贝数。或者,通过Southern印迹分析整合的“TriplePlay”质粒的大小和完

整性。对此,用SpeI (Tripleplay质粒中的单一切割剂) 消化基因组DNA以确定整合的TriplePlay质粒的大小,并且用BglIII/HincII消化基因组DNA以查看载体表达盒的完整性。在mTTR启动子、FVIII cDNA和合成多聚A内的BglIII/HincII消化切割产生1.8和2.8kb片段。在0.8%琼脂糖凝胶上运行经消化的基因组DNA和TriplePlay质粒(掺入基因组DNA中并且用作拷贝数和大小标志物)。将DNA转移到尼龙膜上,并且用经DIG标记的FVIII NcoI片段探查。

[0200] 从生产细胞系中的AAV/mTTR-hFVIII载体产生的表征

[0201] 对选定的MW分析rAAV载体产生。为了比较,通过三重转染产生方法通过将质粒pUC57-mTTR-hFVIIIco转染到293细胞中生成5.1和5.4kb FVIII载体。与生产细胞系方法相当,将细胞收获,裂解,并进行纯化。通过qPCR对来自两种方法的样品定量载体基因组拷贝,并且计算病毒回收率和产率。通过壳体的SDS-PAGE分析、AUC分析、及对于包装的基因组大小表征载体批次(参见下文)。

[0202] 从生产细胞系中生成的rAAV/mTTR-hFVIII载体基因组的表征

[0203] 如下从纯化的壳体中提取包装的载体基因组(VG)。在37℃将病毒与110U DNA酶(Promega)一起温育1小时。添加EDTA以停止消化,接着在存在N-月桂酰十二烷基肌氨酸钠(N-lauryl sarcosyl)的情况下于50℃用蛋白酶K消化温育45分钟。用酚:氯仿:异戊醇(25:24:1)提取DNA两次,并且以14,000rpm于4℃离心10分钟。T-80℃,用100%乙醇和3M乙酸钠沉淀DNA达1小时,离心1小时。在TE中重悬DNA和团粒。

[0204] 对于Southern分析,通过在由30mM NaOH和1mM EDTA组成的运行缓冲液中的1%碱性凝胶电泳分离基因组。将样品转移并交联到Hybond膜(Amersham)上,用对FVIII表达盒特异性的各种片段探查。这些包括含有除了C2外的所有FVIII域的4.0kb NheI-XcmI片段。另外,使用多个25至30聚体链特异性寡核苷酸探针。用AlkPhos直接标记系统(Amersham)标记较大的探针。使用DIG Oligo 3'端标记套组(Roche)根据制造商的说明书对寡核苷酸探针进行3'端标记。

[0205] 对于DNA斑点印迹分析,通过于100℃加热5分钟在TE缓冲液pH 7.0中变性VG,接着在冰上冷却5分钟,并且使用多通道移液管手动施加于尼龙膜。通过UV交联将DNA固定于膜。对于每种经DIG标记的寡核苷酸探针于50℃在Easy Hyb缓冲液中实施杂交6小时,接着进行高严格性清洗,封闭步骤(30分钟),用缀合有碱性磷酸酶的抗DIG Fab片段检测(30分钟),进一步清洗,与CDP-Star底物反应(5分钟),并且根据3'-端标记套组说明书(Roche)暴露于X-射线胶片。使用ImageJ软件(在rsb.info.nih.gov/ij可得到)定量Southern和斑点印迹中信号的密度(density)。

[0206] 表1:引物和探针

[0207]

描述	序列	SEQ ID NO:
FVIII载体基因组定量		
FVIII A1-正向引物	GACGTGGTGCCTTCGA	5
FVIII A1-反向	GGGCGTAATCCCAGTCCTCT	6

[0208]

引物		
FVIII A1探针	AAGCGTGGCCAAGAAGCACCCC	7
氨卡青霉素 ^R 基因定量		
Amp-正向引物	GTTGCCATTGCTACAGGCATC	8
Amp-反向引物	ACTCGCCTTGATCGTTGGG	9
Amp-探针	FAM-ACGCTCGTCGTTTGGTATGGCTTCATTC-TAMRA	10
嘌呤霉素 ^R 基因定量		
嘌呤霉素-正向引物	GGACCGCCACATCGAGC	11
嘌呤霉素-反向引物	CCCCGCTTCGACGCT	12
嘌呤霉素-探针	FAM-TCACCGAGCTGCAAGAACTCTTCCTCAC-TAMRA	13
Rep基因定量		
Rep-正向引物	GACCAGGCCTCATACTCTCCTT	14
Rep-反向引物	GGCAGCCTTGATTTGGGA	15
Rep-探针	FAM-AATGCGGCCTCCAACCTCGCG-TAMRA	16
E6基因定量		
E6-正向引物	CAACACGGCGACCCTACAA	17
E6-反向引物	TCCAATACTGTCTTGCAATATACACAGG	18
E6-探针	FAM -TGCACGGAACCTGAACACTTCACTGCAAG-TAMRA	19
载体基因组分析		
Oligo#4768 (+)	CCGTCGTGAATAGCCTGGACCCTC	20
Oligo#4924 (+)	ATCTGTGTGTTGGTTTTTTGTGTGCGGC	21
Oligo#3342 (-)	AATCCCAGTCCTCTTCCTCGGCGGCGATA	22
Oligo#4900 (-)	AGTATCGGAACACTCGCTCTACGAAATGT	23

[0209] 在体内评估从生产细胞系生成的AAV/mTTR-hFVIII载体

[0210] 在8-12周龄的雄性血友病A KO小鼠 (C56BL/6, 129S-F8^{tm1Kaz} [外显子16中的neo基因]) (Jackson Laboratories) 中评估rAAV载体。经由尾静脉通过静脉内路径投予载体 (4、

10和 30×10^{10} DRP/小鼠)。经由眶后窦将血液收集到柠檬酸钠管中,并且血浆冷冻贮存直至分析。使用Coatest测定法(Diapharma)根据制造商的方案(针对96孔形式修改)对血浆样品分析FVIII活性水平。数值以正常血浆中存在的%FVIII活性测量,并且转化成ng/ml ($100\% \text{ FVIII} = 150 \text{ ng FVIII/ml}$)。还对一些样品测试部分促凝血酶原激酶时间(PTT, IDEXX)。通过标准ELISA(Enzyme Research Laboratories)使用汇集的正常人血浆(Innovative Research)作为标准品定量FVIII蛋白水平。

[0211] 在每项研究结束时收集肝样品。用珠磨器(bead beater)-16匀浆化具有10 μ l β -巯基乙醇和1/4英寸氧化锆1mm珠的1mL RLTplus中的肝(50-400mg)。将匀浆的一部分放入Trizol(对于RNA)或DNA Stat-60(对于DNA)中。使用Trimega方案纯化RNA,接着用旋转柱(Promega Z3100)根据制造商纯化。用无核酸酶的水洗脱RNA,并且以15,000x g离心1分钟。使用RNA生成cDNA(Invitrogen)。通过DNA提取Purelink柱(Invitrogen)根据制造商的说明书纯化DNA。随后,通过使用对FVIII A2区特异性的引物和探针(表1)的qPCR,分别使用cDNA和DNA两者定量FVIII mRNA和载体基因组拷贝。

[0212] 结果

[0213] 为了生成用于超大rAAV载体产生的PCL平台,构建用于表达FVIII的新型盒。这些盒侧翼有AAV ITR,并且范围为5.1至5.4kb载体基因组(图1A)。每个盒包含源自mTTR启动子的启动子,并且构建不同mTTR变体以检查其对表达的影响(参见图1B中提供的变体的比对和解释)。

[0214] 质粒语境中的所有这些表达盒在小鼠中体内测试时生成FVIII(图1C)。图1A和1B中显示的HNF3和HNF4结合位点中的修饰相对于核心mTTR启动子(“202”)增加FVIII生成,但是其它修饰,如mTTR增强子和BGH多聚A则不然(图1C)。图1D显示FVIII表达盒的图。图1E显示TriplePlay质粒的设计。

[0215] 使用具有核心mTTR(5.1kb)的表达盒和具有增强子、mTTR和BGH多聚A(5.4kb)的表达盒,以在生成对于每种的TriplePlay质粒后随后测试对于超大FVIII载体的PCL生产。FVIII ELISA结果确认了当转染到Huh7细胞中时,经转染的TriplePlay/FVIII质粒在体外生成FVIII。FVIII质粒也能够在小规模包装实验中生成rAAV。

[0216] 总之,生成mTTR启动子修饰,其在体内增加自核心mTTR启动子的表达。所有TriplePlay质粒在体外表达FVIII,并且能够在小规模包装实验中生成病毒。

[0217] 为了生成具有超大5.1kb mTTR-FVIII载体的生产细胞系,对MW分析rAAVrh8R/FVIII产生水平。这些之中,鉴定出高生产者、中等生产者和低生产者。如此,显示可生成PCL用于超大rAAV/mTTR-FVIII载体。

[0218] 实施例2:对具有mTTR-FVIII载体的生产细胞系表征基因组DNA

[0219] 为了评估实施例1中描述的FVIII表达盒的完整性和TriplePlay质粒的整合拷贝,选择含有AAVrh8R/5.1kb、AAVrh8R/5.4kb或AAV8/5.1kb FVIII载体的MW以分析基因组DNA。

[0220] 在高生产性MW(MW#35)中,Southern分析揭示了细胞系基因组中约50个拷贝的载体,而根据Southern,中等生产性MW#272具有每种的小于10个拷贝(图2A)。

[0221] 还通过qPCR分析用于产生具有mTTR-FVIII的AAVrh8R/5.1kb、AAVrh8R/5.4kb或AAV8/5.1kb的高和中等生产性MW(表2)。FVIII、rep和puroR基因的拷贝数使用针对每种的特异性引物/探针测定,并将拷贝数相对于HeLaS3细胞中存在的E6基因(每个HeLaS3基因组

为11个拷贝)标准化。在高生产性MW (MW#35) 中,qPCR分析揭示了约59–67个拷贝的载体、rep和嘌呤霉素(表2),而根据qPCR,中等生产性MW#272具有每种的15至18个拷贝(表2)。这些数值与通过Southern获得的结果相比是略高的,但是具有相似的排序次序(图2A)。为了比较,将表达SEAP的正常大小载体(4.3kb)包装到AAV2或AAV8壳体中,并分析。

[0222] 表2:基因组分析选定MW的整合的TriplePlay质粒的拷贝。

[0223]

细胞系	主孔	拷贝/细胞			生产水平
		FVIII (stdev)	REP (stdev)	嘌呤霉素 (stdev)	
AAVrh8R/FVIII 5.1kb	MW#35	67 (2)	65 (0)	59 (1)	H
AAVrh8R/FVIII 5.1kb	MW#272	15	16	18	M
AAVrh8R/FVIII 5.1kb	MW#418	229	195	260	H
AAVrh8R/FVIII 5.4kb	MW#61	235 (13)	256 (1)	196 (12)	M
AAVrh8R/FVIII 5.4kb	MW#163	253 (40)	265 (1)	237 (5)	H
AAV8/FVIII 5.1kb	MW#287	270 (38)	294 (6)	266 (13)	H
AAV8/FVIII 5.1kb	MW#342	101 (5)	126 (2)	108 (4)	H
AAVrh8R/FVIIIopt 5.1KB	MW#14	1362 (48)	1499 (22)	1343 (42)	H
AAVrh8R/FVIIIopt 5.1KB	MW#27	77 (4)	82 (3)	73 (73)	M
AAV2/SEAP	对照(MW#156 SEAP)	0	73 (1)	70 (2)	H
	HeLaS3	0	0	0	

[0224] 用限制酶(SpeI)消化的基因组DNA的Southern印迹分析显示了与掺入基因组DNA中的线性化Tripleplay/FVIII质粒相似迁移的~13kb条带的产生,预测所述限制酶仅在Tripleplay/FVIII质粒中切割一次。因此,所有克隆含有整合到HeLaS3基因组中的整个质粒。

[0225] 如图2A中显示,所有MW还具有不同大小的低拷贝(1个拷贝)条带,其代表在整合位点侧翼的基因组DNA。虽然272(中等生产者)和35(高生产者)具有指示单一整合位点(仅观察到两个侧翼片段)的模式,但是MW418具有多个侧翼片段及较大的(2x 14=24kb左右)片段,其潜在代表整合质粒的正向和反向取向的串联。此事实连同多个整合模式,提示MW418是克隆的混合物。

[0226] 如图2B中显示,通过用在FVIII表达盒内切割的酶(HincII,BglII)消化来分析FVIII载体基因组完整性。与对照相比,基于通过消化初始TriplePlay质粒获得的相似结果,观察到正确的大小片段。这些结果证明了整合后不发生载体的重排。对具有5.4kb载体的生产细胞系及对具有AAV8壳体的5.1kb载体完成相似的分析,并且获得了相当的结果。总之,在从生产细胞系分离的基因组DNA中没有观察到整合的5.1或5.4kb载体序列中的重排或缺失,这指示含有超大AAV载体基因组的细胞系的生成是可行的。

[0227] 实施例3:使用生产细胞系的超大载体产生和载体分析

[0228] 接着,检查使用上文描述的MW35细胞系的rAAV载体产生。在小规模培养物中对

AAVrh8R/5.1kb FVIII载体的高生产性克隆(MW35)测试rAAV载体产生。在第3天和第4天看到峰rAAV产生,并在培养放大期间维持高生产水平(图3A)。

[0229] 这通过下文表3中显示的结果进一步证明。放大通过MW#35的生产以比较载体产生水平。另外,定量细胞团粒(细胞)和培养基(CM)中的载体水平。显示了正常大小AAV2/SEAP载体用于比较。

[0230] 表3:通过MW#35的AAVrh8R/5.1kb载体生产。

[0231]

FVIII MW#35	SP:VG/ml	细胞(%)	CM(%)
20ml (4x 10 ⁶ 个细胞)	2.41 × 10 ¹⁰	39	61
250ml (5x 10 ⁷ 个细胞)	2.64 × 10 ¹⁰	45	55
1000ml (2x 10 ⁸ 个细胞)	3.17 × 10 ¹⁰	35	64
MW156 (SEAP):20ml	4.15 × 10 ¹⁰	NT	NT

[0232] 对于使用的载体血清型,在细胞团粒和培养基中同等地检测到载体。MW 35和418两者在几次传代里都是稳定的。从第5次传代到第20次传代,MW35维持高水平产生($\geq 1 \times 10^{10}$ DRP/ml)(图3B)。类似地,从第5次传代到第21次传代,MW418(分类为中等生产者)维持稳定的中等水平产生($\geq 1 \times 10^9$ drp/ml)。还对MW287证明了稳定的载体产生,所述MW287生成具有不同壳体血清型(AAV8)的5.1kb载体(图3B)。

[0233] 这些数据证明了与先前已经对正常大小载体所显示的(Martin,J.等(2013)Molecular Therapy 21:2205-2216)相似,可独立于壳体血清型从超大载体获得高且稳定的产生。用具有5.4kb mTTR-FVIII载体的PCL以及对具有含AAV8壳体的5.1kb载体的PCL也获得稳固的载体产生,产生相当的结果。

[0234] 接着,放大使用高生产性主孔(对于AAVrh8R/5.1kb FVIII为MW#35,对于AAV8/5.1kb FVIII为MW#287,以及对于AAVrh8R/5.4kb FVIII为MW#163)的载体产生以评估载体产率和质量(表4和图4A)。与通过三重转染方法生成的载体比较载体产生(表4和图4B)。在AAVrh8R/5.1kb FVIII载体的比较中包括3个PCL和2个TXN批次。例如,显示了用具有5.1kb基因组的AAVrh8R壳体得到的数据。

[0235] 表4:通过PCL或三重转染方法产生的AAVrh8R/5.1kb FVIII载体的比较。

[0236]

分析	生产细胞系	三重转染
细胞/生产	2 × 10 ⁹ 个细胞	3 × 10 ⁹ 个细胞
总载体产率	2 × 10 ¹⁴ DRP	6 × 10 ¹³ DRP
产率/细胞	1 × 10 ⁵ DRP/细胞	2 × 10 ⁴ DRP/细胞
%含有VG的病毒	44-50%	24-30%
%具有>4.7kb VG的病毒	59-61%	43-50%

[0237] 用MW35的PCL生产运行导致一致的产物概况,如通过AUC分析评估(图4A)。如上文表4中汇总,对于通过PCL生成的病毒,通过此分析测定的含有载体基因组的壳体的百分比是44-50%,而三重转染材料具有更低的水平(30%;对于HLP19,即大主链载体质粒仅为1-5%)。此外,PCL生成的材料中存在较高部分的具有较大基因组(≥ 4.7 kb)的病毒。载体产率/细胞是1 × 10⁵DRP/细胞(通过高生产性PCL)和1-3 × 10⁴DRP/细胞(通过TXN)(用HLP19和

大载体主链更低, 1×10^3 DRP/细胞)。

[0238] 表5:通过PCL或三重转染方法生成的AAVrh8R/5.4kb FVIII载体的比较。

[0239]

分析	生产细胞系	三重转染
细胞/生产	2×10^9 个细胞	2×10^9 个细胞
总载体产率(粗制)	2×10^{13} DRP	2×10^{14} DRP
产率/细胞	2×10^4 DRP/细胞	7×10^4 DRP/细胞
%含有VG的病毒	24%	23%
%具有>4.7kb VG的病毒	68%	39%

[0240] 使用MW163对5.4kb载体的类似分析显示了与三重转染相比通过PCL得到的略低的DRP/细胞水平(上文表5)。虽然含有VG的壳体的总百分比是相当的,但是PCL生成的病毒具有较高水平的具有较大基因组的病毒(图5B)。

[0241] 当通过AUC分析表征PCL材料和三重转染材料时,与三重转染材料(图5B)相比,较大基因组的比例在PCL材料(图5A)中为2倍高。

[0242] 表6:通过PCL和三重转染产生的载体中的异常包装的分析。

[0243]

通过PCL得到的载体	载体滴度(DRP/ml)		
	FVIII	嘌呤霉素	%嘌呤霉素
AAVrh8R/5.1-FVIII	$1.2\text{E}+12$	$5.40\text{E}+09$	0.44
AAVrh8R/5.4-FVIII	$7.6\text{E}+12$	$1.54\text{E}+10$	0.20
AAV8/5.1-FVIII	$5.1\text{E}+12$	$1.21\text{E}+10$	0.24

[0244]

通过TXN得到的载体	FVIII	氨卡青霉素	%氨卡青霉素
AAVrh8R/5.1-FVIII	$2.3\text{E}+13$	$4.76\text{E}+11$	2.10
AAVrh8R/5.4-FVIII	$1.6\text{E}+13$	$5.38\text{E}+11$	3.40
AAVrh8R/FVIII 4.6kb	$1.2\text{E}+13$	$2.47\text{E}+11$	2.00

[0245] 如通过包装病毒中质粒衍生抗生素抗性基因(用于PCL的嘌呤霉素,用于TXN的氨苄青霉素)的存在测量的异常的、不想要的DNA包装的水平在通过PCL生成的病毒中较低(<1%) (表6)。比较而言, TXN生成的病毒具有约10倍高的异常包装水平。

[0246] 总之,数据显示了选定的生产细胞系能够生成高水平的超大载体(>100,000DRP/细胞)。此外,这些细胞系在几次传代(>20)里维持载体产生能力,所述几次传代会是用于制造的大型放大需要的。PCL生成的载体与通过标准三重转染方法生成的载体的比较显示了PCL材料含有更多的含有载体基因组的病毒和更高比例的野生型大小或较大载体基因组以及含有更少的不想要的非载体相关DNA。

[0247] 实施例4:PCL产生的超大载体中的包装的载体基因组的表征

[0248] 接着,分析由PCL平台产生的超大rAAV载体中的壳体化的载体基因组。从纯化的病毒体中分离载体基因组,并且通过碱性凝胶电泳,接着使用对载体特异性的探针的Southern印迹分析来分析单链基因组大小(图6A)。通过FVIII表达盒的4.0kb片段探查的

Southern印迹显示了大多数VG大小在通过任一种方法生成的载体中大约为4.6kb或更大(图6B)。使用ImageJ软件(<http://rsb.info.nih.gov/ij>)定量Southern印迹中的信号密度。

[0249] 还使用链特异性寡核苷酸探针分析VG以定量缺失的5' 末端的比例。由于已知AAV基因组的包装从3' 端开始发生(King, J.A.等(2001)EMBO J.20:3282-3291),当基因组大小超过4.7kb时,超大载体可以缺乏负和正链的5' 端中的序列。通过将每种载体的2倍连续稀释物施加于膜上(以 2.4×10^9 开始;总共8个降低的载体浓度加上作为阴性对照的无施加基因组)分析图4A和4B中使用的载体批次。使每个印迹与对载体基因组(正或负极性)的3' 或5' 末端特异性的3' 端标记的寡核苷酸探针杂交。通过将每种载体的2倍连续稀释物施加于膜上(以 2.4×10^9 开始;总共8个降低的载体浓度加上作为阴性对照的无施加基因组)分析图5中使用的载体批次。使每个印迹与对载体基因组(正或负极性)的中间或5' 末端特异性的3' 端标记的寡核苷酸探针杂交。定量信号强度,并且相对于4.6kb载体(完全包装)标准化。使用三个浓度来产生标准误差。

[0250] 数据显示了相比于完全包装的4.6kb载体(除了编码FVIII的C1域的区域外具有与5.1kb载体相当的序列),当对单链基因组两种极性使用与超出4.7kb的区域互补的寡核苷酸时,5.1kb载体具有更低的信号强度(图7A和7B和7C)。在使用的大多数5' 探针的情况下,与PCL载体相比,此差异在三重转染的材料中更高。

[0251] 当用与+或-链互补的寡核苷酸探针探查Southern印迹时也观察到PCL和三重转染生成的载体中的5' 端中的差异。显示了每种寡核苷酸探针与相应的3' 端的距离(图8A)。PCL和三重转染(TXN)生成的病毒的相等检测首先通过用链特异性寡核苷酸探针的DNA斑点印迹分析比较,并显示了每条链的每种病毒的相当量(使用4.6kb病毒作为完全包装病毒的对照;使用含有FIX和FVIII的质粒作为检测特异性的阴性和阳性对照)(图8B)。当用针对3' 端的寡核苷酸探针探查Southern印迹时,PCL和TXN生成的病毒两者证明了+和-链的存在(图8C,左小图)。与图6A中显示的结果相似,在PCL病毒中检出更高水平的大于4.6kb的载体基因组。当用对5' 端特异性的寡核苷酸探查Southern印迹时,PCL载体显示了大于4.6kb的包装的基因组的存在,而三重转染生成的病毒显示了这些较大基因组的信号的明确缺乏(使用的探针检测距3' 端为4.9kb的区域)(图8C)。还通过定量Southern印迹中的各种大小的包装的基因组的信号强度(使用针对3' 端的探针)来确认较高部分的大于4.7kb的基因组(图8D)。定量还显示了与TXN载体相比PCL载体中更少的片段化/较小基因组(<4.7kb)。

[0252] 接着,我们通过相似的方法评估了由PCL和TXN生成的5.4kb FVIII载体的载体基因组包装。显示了每个寡核苷酸探针的互补序列在5.4kb基因组中距其相应的3' 端的位置(图9A)。与5.1kb载体的结果相似,用针对基因组的3' 端的寡核苷酸探针以相当的水平检出5.4kb载体的+和-链(图9B)。然而,对距3' 端远于4.7kb定位的区域特异性的探针未能检测通过TXN方法生成的载体中的基因组(图9C)。比较而言,通过PCL方法生成的载体显示了大于4.7kb(尽管不与5.4kb一样大)的基因组的存在。

[0253] 总之,数据证明了与TXN方法相比通过PCL方法生成的载体中更高水平的较大载体基因组包装。

[0254] 实施例5:超大rAAV/mTTR-FVIII载体的体内效力

[0255] 接着,对由PCL平台产生的超大rAAV载体检查体内效力。通过尾静脉以 3×10^{11} 和4

$\times 10^{10}$ DRP/小鼠对小鼠投予载体,并且分析直至第56天。PCL产生的AAVrh8R/5.1kb mTTR-FVIII载体以剂量响应性方式产生在经处理的血友病A KO小鼠的血浆中可检出的活性FVIII蛋白(图10A)。除Coatest活性测定法外,还使用激活部分促凝血酶原激酶时间(aPTT)测定法通过凝固时间对FVIII活性评估功能性。此测定法对测试的低剂量和高剂量显示了相当的凝固时间,因此指示临床相关的低剂量足以使血友病A KO小鼠中的凝固时间正常化(图10B)。

[0256] 然后,使用临床相关的载体剂量(2×10^{12} DRP/kg, 4×10^{10} DRP/小鼠)比较PCL生成的载体与三重转染产生的载体。通过尾静脉对血友病A KO小鼠投予载体。PCL生成的病毒比TXN产生的病毒生成更有活性的FVIII蛋白,如通过Coatest活性测定法测定的(图11A)。这也与由PCL生成的载体通过aPTT测定法在第21天得到的显著更短的凝固时间相关联(图11B)。在第56天在PCL和TXN材料之间观察到很小的差异,这提示了PCL材料产生较快表达动力学(图11C)。肝载体基因组拷贝的定量显示了通过PCL生成的载体处理的动物中比用TXN材料观察到的情况更持久的载体基因组(图11D)。

[0257] 还对血友病A敲除小鼠测试PCL和TXN生成的较大的5.4kb FVIII载体。与5.1kb载体类似,由PCL生成的5.4kb载体显示通过Coatest测定法得到的更高的FVIII活性和第24天时更短的凝固时间(图12A,B)。为了分析肝中的载体基因组水平的动力学,在载体投予后3天和43天定量载体基因组拷贝。数据证明了在这两天,通过PCL生成的载体处理的动物中比用TXN材料的情况存在约2倍多的载体基因组(图12C)。

[0258] 总之,当在血友病A KO疾病模型中体内测试时,通过PCL方法生成的超大rAAV/mTTR-FVIII载体比TXN方法的情况产生2倍高的FVIII活性和更短的凝固时间。这些结果与经处理的动物的肝中存在的2倍高水平的持久载体基因组相关联。因此,这些结果证明了用于超大rAAV载体的两种载体产生方法间包装的基因组的质量中观察到的差异转化为通过PCL生成的载体得到的更高的体内效力,并且基于靶器官中转录活性载体基因组的增加的生成效率。

[0259] 实施例6:具有超大5.1、5.9和6.7kb SEAP载体的生产细胞系的生成

[0260] 方法:通过范围为5.1kb至5.9和6.7kb的载体大小逐步增加生成超大AAV2-SEAP载体基因组(参见图13A)。使用AAVsp70质粒作为模板经由PCR扩增3种长度(0.8、1.6和2.4kb)的AAT填充DNA片段,并且将每种填充片段克隆到具有AAV2cap和rep基因和每种SEAP载体基因组的TriplePlay质粒中以生成一系列pAF-SEAP质粒。

[0261] 为了生成用于超大SEAP载体的细胞系,对含有载体基因组和AAV2rep和cap基因的相应质粒并行比较在24孔高通量分析转染中生成高生产性细胞系的能力。按照标准方案,用每种构建体转染HeLaS3细胞的一式两份T75烧瓶。转染后1天,用75,000个细胞/孔接种每次转染8块 $\times$ 24孔板,并且启动药物选择。培养样品,并且评估集落大小和汇合,准备好用于相对生产率筛选。

[0262] 结果:首先确认5.1、5.9和6.7kb超大载体质粒经由瞬时转染包装到HeLaS3细胞(+wtAd5)中(图13A)。结果显示了每种质粒促进载体包装(数据未显示)。

[0263] 在相对生产筛选中为每个构建体筛选大致 ~ 100 – 170 个主孔。在相对生产筛选中的百分比阳性主孔(产生大于 1×10^7 DRP/ml的那些)全体都较高($>80\%$),其中仅6.7kb构建体显示减少的量(65.7%)。此外,虽然仅5.1kb构建体产生在高($>1 \times 10^{10}$ DRP/ml)范围中生

产的主孔(总共三个),但是在所有情况中鉴定出在中等-高范围($>1 \times 10^9$ DRP/ml)中生产的主孔。百分比中等-高遵循预期的模式,5.1kb为20%,5.9kb为15.4%并且6.7kb为10.7%。

[0264] 随后,所有较高生产性主孔进行比生产率分析。两个比生产率筛选的结果显示于图13B(灰色和白色柱形),并且与每个主孔的相对生产率值(黑色柱形)比较。结果显示在许多情况中,载体产率在比生产率筛选中保持稳定。

[0265] 总之,数据证明了可以对大小至少6.7kb的载体产生生产细胞系。

[0266] 序列

[0267] 以N端至C端呈现所有多肽序列,除非另有记录。

[0268] 以5'至3'呈现所有核酸序列,除非另有记录。

[0269] mTTR202-HI-hFVIIIco-spA (5097 bp)

[0270]

GAGCTCTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCCGACGCCCGGG
CTTTGCCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACCTCCATCACTAGGGGTTCTCTACG
CGTGTCTGTCTGCACATTTTCGTAGAGCGAGTGTTCGATACTCTAATCTCCCTAGGCAAGGTTTCATATTTGTGTAGG
TTACTTATTCTCCTTTTGTGACTAAGTCAATAATCAGAATCAGCAGGTTTGGAGTCAGCTTGGCAGGGATCAGCAG
CCTGGGTTGGAAGGAGGGGGTATAAAAGCCCCCTTACCAGGAGAAGCCGTCACACAGATCCACAAGCTCCTGCTAGC
AGGTAAGTGCCGTGTGTGGTTCCCGCGGGCCTGGCCTCTTTACGGGTTATGGCCCTTGCGTGCCTTGAATTACTGAC
ACTGACATCCACTTTTTCTTTTTCTCCACAGGTATCGATTCTCTAGAGCCACCATGCAGATCGAGCTGTCTACCTGC
TTCTTCCTGTGCCTGCTGCGGTTCTGCTTCAGCGCCACCAGACGGTACTATCTGGGCGCCGTGGAAGTCTGAGCTGGGA
CTACATGCAGAGCGACCTGGGCGAGCTGCCCGTGGATGCCAGATTCCCTCCAAGAGTGCCCAAGAGCTTCCCCTTCA
ACACCTCCGTGGTGTACAAGAAAACCTGTTCGTGGAATTCACCGACCACCTGTTCAATATCGCCAAGCCCAGACCC
CCCTGGATGGGCTGCTGGGACCTACAATTCAGGCCGAGGTGTACGACACCGTCGTGATCACCTGAAGAACATGGC
CAGCCACCCCGTGTCTCTGCATGCCGTGGGAGTGTCTACTGGAAGGCCTCTGAGGGCGCCGAGTACGACGATCAGA
CCAGCCAGCGCGAGAAAGAGGACGACAAGGTGTTCCCTGGCGGCAGCCACACCTACGTGTGGCAGGTGCTGAAAGAA
AACGGCCCCATGGCCTCCGACCCCTCTGTGCCTGACATACAGCTACCTGAGCCACGTGGACCTCGTGAAGGACCTGAA
CAGCGGCCTGATCGGAGCCCTGCTCGTGTGTAGAGAGGGCAGCCTGGCCAAAGAGAAAACCCAGACCCTGCACAAGT
TCATCCTGCTGTTTCGCCGTGTTTCGACGAGGGCAAGAGCTGGCACAGCGAGACAAAGAACAGCCTGATGCAGGACCGG
GACGCGCCTCTGCTAGAGCCTGGCCCAAAATGCACACCGTGAACGGCTACGTGAACAGAAGCCTGCCCGGACTGAT
CGGCTGCCACCGGAAGTCTGTGTACTGGCACGTGATCGGCATGGGCACCACCCCTGAGGTGCACAGCATCTTTCTGG
AAGGACACACCTTTCTCGTGCGAACACCGGCAGGCCAGCCTGGAAATCAGCCCTATCACCTTCCTGACCGCCCAG
ACACTGCTGATGGACCTGGGCCAGTTTCTGCTGTTCTGCCACATCAGCTCCCACCAGCACGACGGCATGGAAGCCTA
CGTGAAGGTGGACAGCTGCCCCGAGGAACCCAGCTGCGGATGAAGAACAACGAGGAAGCCGAGGACTACGACGACG
ACCTGACCGACAGCGAGATGGACGTGGTGCGCTTCGACGACGATAACAGCCCCAGCTTCATCCAGATCAGAAGCGTG
GCCAAGAAGCACCCCAAGACCTGGGTGCACTATATCGCCCGGAGGAAGAGGACTGGGATTACGCCCCTCTGGTGCT
GGCCCCCGACGACAGAAGCTACAAGAGCCAGTACCTGAACAATGGCCCCCAGCGGATCGGCCGGAAGTATAAGAAAG
TGCGGTTTCATGGCCTACACCGACGAGACATTCAAGACCAGAGAGGCCATCCAGCACGAGAGCGGCATCCTGGGCCCT
CTGCTGTATGGCGAAGTGGGCGACACCCTGCTGATCATCTTCAAGAACCAGGCCAGCAGACCCTACAACATCTACCC
TCACGGCATCACCGACGTGCGGCCCTGTACTCCAGAAGGCTGCCAAGGGCGTGAAACACCTGAAGGACTTCCCCA
TCCTGCCCGGCGAGATCTTCAAGTACAAGTGGACCGTGACCGTGGAAGATGGCCCCACCAAGAGCGACCCAGATGC

CTGACACGGTACTACAGCAGCTTCGTGAACATGGAACGGGACCTGGCCTCCGGCCTGATTGGCCCACTGCTGATCTG
CTACAAAGAAAGCGTGGACCAGCGGGGCAACCAGATCATGAGCGACAAGCGGAACGTGATCCTGTTTAGCGTGTTCCG
ATGAGAACCGGTCCTGGTATCTGACCGAGAATATCCAGCGGTTCCCTGCCAACCCCTGCCGGCGTGACGCTGGAAGAT
CCTGAGTTCCAGGCCTCCAACATCATGCACTCCATCAATGGCTATGTGTTTCGACAGCCTGCAGCTGAGCGTGTGCCT
GCACGAGGTGGCCTACTGGTACATCCTGAGCATCGGGGCCAGACCGACTTCCTGTCCGTGTTCTTCTCCGGCTACA
CCTTCAAGCACAAGATGGTGTACGAGGATACCCTGACCCTGTTCCCCCTTAGCGGCGAAACCGTGTTTCATGAGCATG
GAAAACCCCGGCCTGTGGATCCTGGGCTGCCACAACAGCGACTTCGGGAACAGAGGCATGACCGCCCTGCTGAAGGT
GTCCAGCTGCGACAAGAACACCGGCGACTACTACGAGGACAGCTATGAGGACATCAGCGCCTACCTGCTGAGCAAGA
ACAATGCCATCGAGCCCAGAAGCTTCAGCCAGAACCCCCCGTGCTGAAGCGGCACCAGAGAGAGATCACCCGGACC
ACCCTGCAGTCCGACCAGGAAGAGATCGATTACGACGACACCATCAGCGTGGAATGAAGAAAGAAGATTTTCGACAT
CTACGACGAGGACGAGAACCAGAGCCCCCGTCTTTCAGAAAAAGACCCGGCACTACTTCATTGCCGCTGTGGAAC
GGCTGTGGGACTACGGCATGAGCAGCAGCCCTCACGTGCTGAGAAACAGGGCCCAGAGCGGCAGCGTGCCCCAGTTC
AAGAAAGTGGTGTTCCAGGAATTCACAGACGGCAGCTTCACCCAGCCTCTGTACCGCGCGAGCTGAATGAGCACCT
GGGACTGCTGGGCCCCTATATCAGAGCCGAAGTGGAAGATAATATCATGGTCACCTTCGGGAATCAGGCCTCCCGGC
CCTACAGCTTCTACAGCTCCCTGATCAGCTACGAAGAGGACCAGAGACAGGGCGCTGAGCCCCGGAAGAACTTCGTG
AAGCCCAACGAGACTAAGACCTACTTTTGGAAGGTGCAGCACCACATGGCCCCTACAAAGGACGAGTTCGACTGCAA
GGCCTGGGCCTACTTCTCCGATGTGGACCTGAAAAAGGACGTGCACTCTGGGCTGATCGGGCCCCCTGCTCGTGTGCC
ACACCAACACCCTGAATCCCGCCACGGCAGACAAGTGACAGTGCAGGAATTCGCCCTGTTCTTCACCATCTTCGAC
GAAACAAAGAGCTGGTACTTCACCGAAAAACATGAAAAGAACTGCCGGGCTCCCTGCAACATCCAGATGGAAGATCC
CACCTTCAAAGAGAACTACCGGTTCCACGCCATCAACGGCTACATCATGGACACACTGCCCCGGCCTCGTGATGGCTC
AGGATCAGCGGATCCGGTGGTATCTGCTGTCCATGGGCTCCAACGAGAACATCCACAGCATCCACTTCAGCGGCCAC
GTGTTACCGTGCGGAAAAAAGAAGAGTACAAAATGGCCCTGTACAACCTGTACCCTGGGGTGTTTCGAGACAGTGGA
AATGCTGCCCAGCAAGGCCGGCATCTGGCGGGTGGAATGTCTGATCGGCGAGCATCTGCACGCTGGGATGAGCACAC
TGTTTCTGGTGACAGCAACAAGTGCCAGACACCTCTGGGCATGGCCTCTGGCCACATCCGGGACTTTCAGATCACA
GCCAGCGGCCAGTATGGCCAGTGGGCCCCAAAACCTGGCCAGACTGCACTACAGCGGCAGCATCAACGCCTGGTCCAC
CAAAGAGCCCTTCAGCTGGATCAAGGTGGACCTGCTGGCTCCCATGATCATCCACGGAATCAAGACCCAGGGCGCCA
GACAGAAGTTCAGCAGCCTGTACATCAGCCAGTTCATCATCATGTACAGCCTGGACGGCAAGAAGTGGCAGACCTAC
CGGGGCAATAGCACCGGCACCCCTGATGGTGTCTTCGGCAACGTGGACTCCAGCGGCATTAAGCACAACATCTTCAA
CCCCCCCATCATTGCCCCGTACATCCGGCTGCACCCCAACCCACTACAGCATCCGGTCCACCCTGAGAATGGAAGTGA
TGGGCTGCGACCTGAACTCCTGCAGCATGCCCTGGGGATGGAAGCAAGGCCATCTCCGACGCCAGATCACCGCC
TCCAGCTACTTCACCAACATGTTCCGCACCTGGTCCCCATCCAAGGCCCGGCTGCATCTGCAGGGCAGAAGCAATGC
TTGGAGGCCCCAAGTGAACAACCCCAAGAAATGGCTGCAGGTGGACTTCAGAAAACCATGAAAGTGACCGGCGTG
CCACCCAGGGCGTGAAGTCTCTGCTGACCTCTATGTACGTGAAAGAGTTCCTGATCTCCAGCAGCCAGGACGGCCAC
CAGTGGAACCTGTTTTTCCAGAACGGCAAAGTGAAAGTGTTCAGGGGAACAGGACTCCTTACCCCCGTGCTGAA
TAGCCTGGACCCTCCACTGCTGACCAGATACCTGCGGATCCACCCTCAGAGTTGGGTGCACCAGATTGCTCTGCGGA
TGGAAGTGCTGGGATGCGAGGCCCAGGACCTGTACTGACACTAGTAATAAAAGATCAGAGCTGTAGAGATCTGTGTG
TTGGTTTTTTGTGTGCGGCCGTACCAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGC
TCACTGAGGCCGGGCGACCAAAGGTCGCCGACGCCCGGGCTTGGCCGGGCGGCTCAGTGAGCGAGCGAGCGCGC
AGAGAGGGAGTGGCC (SEQ ID NO:1)

[0271] mTTR202opt

[0272]

TGTCTGTCTGCACATTTTCGTAGAGCGAGTGTTCCGATACTCTAATCTCCCGGGGCAAAGGTCGTATTGACTTAGGTT
ACTTATTCTCCTTTTGTGACTAAGTCAATAATCAGAATCAGCAGGTTTGGAGTCAGCTTGGCAGGGATCAGCAGCC
TGGGTTGGAAGGAGGGGTATAAAAGCCCCTTCACCAGGAGAAGCCGTCACACAGATCCACAAGCTCCTGCTAGC
(SEQ ID NO:2)

[0273] mTTR482opt

[0274]

CTACCTGCTGATCGCCCGGCCCTGTTCAAACATGTCCTAATACTCTGTGCGGGCAAAGGTCGGCAGTAGTTTTCCA
TCTTACTCAACATCCTCCCAGTGACGTAGGATCCTGTCTGTCTGCACATTTTCGTAGAGCGAGTGTTCCGATACTCT
AATCTCCCGGGCAAAGGTCGTATTGACTTAGGTTACTTATTCTCCTTTTGTGACTAAGTCAATAATCAGAATCAG
CAGGTTTGGAGTCAGCTTGGCAGGGATCAGCAGCCTGGGTTGGAAGGAGGGGTATAAAAGCCCCTTCACCAGGAGA
AGCCGTCACACAGATCCACAAGCTCCTGCTAGC (SEQ ID NO:3)

[0275] pTGENcaprh8R-ITRmTTRFVIII (具有AAVrh8R壳体和5.1kbITR-mTTR-hFVIIIco的
TriplePlay质粒) 13524bp

[0276]

GAATGCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCA
CAAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTTGCCAACTCATCAATGTATCTTATCATGTCTGGATC
CGCTAGAACTAGGAATTCGCTAGCGGTACCGATATCCTAGTGGATCCCCGTACACAGGAAGTGACAATTTTCGCGC
GGTTTTAGGCGGATGTTGTAGTAAATTTGGGCGTAACCGAGTAAGATTTGGGTGGTCACGCTGGGTATTTAAGCCCG
AGTGAGCACGCAGGGTCTCCATTTTGAAGCGGGAGGTTTGAACGCGCAGCCGCATGCCGGGGTTTTACGAGATTGT
GATTAAGGTCCCCAGCGACCTTGACGAGCATCTGCCCGGCATTTCTGACAGCTTTGTGAACTGGGTGGCCGAGAAGG
AATGGGAGTTGCCGCCAGATTCTGACATGGATCTGAATCTGATTGAGCAGGCACCCCTGACCGTGGCCGAGAAGCTG
CAGCGCGACTTTCTGACGGAATGGCGCGGTGTGAGTAAGGCCCGGAGGCCCTTTCTTTGTGCAATTTGAGAAGGG
AGAGAGCTACTTCCACATGCACGTGCTCGTGAAACCACCGGGGTGAAATCCATGGTTTTGGGACGTTTCCTGAGTC
AGATTCGCGAAAACTGATTTCAGAGAATTTACCGCGGATCGAGCCGACTTGGCCAACTGGTTCGCGGTCACAAAG
ACCAGAAATGGCGCCGAGGCGGGAACAAGGTGGTGGATGAGTGCTACATCCCCAATTACTTGCTCCCCAAAACCCA
GCCTGAGCTCCAGTGGGCGTGACTAATATGGAACAGTATTTAAGCGCCTGTTGAATCTCAGGAGCGTAAACGGT
TGGTGGCGCAGCATCTGACGCACGTGTCGCAGACGCAGGAGCAGAACAAAGAGAATCAGAATCCCAATTCTGATGCG
CCGGTGATCAGATCAAAAACCTTCAGCCAGGTACATGGAGCTGGTCGGGTGGCTCGTGGACAAGGGGATTACCTCGGA
GAAGCAGTGGATCCAGGAGGACCAGGCCTCATACATCTCCTTCAATGCGGCCTCCAACCTCGCGGTCCCAAATCAAGG
CTGCCTTGGACAATGCGGGAAAGATTATGAGCCTGACTAAAACCGCCCCGACTACCTGGTGGGCCAGCAGCCCGTG
GAGGACATTTCCAGCAATCGGATTTATAAAATTTTGGAACTAAACGGGTACGATCCCCAATATGCGGCTTCCGTCTT
TCTGGGATGGGCCACGAAAAAGTTTCGGCAAGAGGAACACCATCTGGCTGTTTGGGCCTGCAACTACCGGAAGACCA
ACATCGCGGAGGCCATAGCCACACTGTGCCCTTCTACGGGTGCGTAACTGGACCAATGAGAACTTTCCCTTCAAC
GACTGTGTCGACAAGATGGTGATCTGGTGGGAGGAGGGGAAGATGACCGCCAAGGTCGTGGAGTCGGCCAAAGCCAT
TCTCGGAGGAAGCAAGGTGCGCGTGGAACAGAAATGCAAGTCCTCGGCCAGATAGACCCGACTCCCGTGATCGTCA
CCTCCAACACCAACATGTGCGCCGTGATTGACGGGAACTCAACGACCTTCGAACACCAGCAGCCGTTGCAAGACCGG
ATGTTCAAATTTGAACTCACCCGCCGTCTGGATCATGACTTTGGGAAGGTCACCAAGCAGGAAGTCAAAGACTTTTT

CCGGTGGGCAAAGGATCACGTGGTTGAGGTGGAGCATGAATTCTACGTCAAAAAGGGTGGAGCCAAGAAAAGACCCG
CCCCAGTGACGCAGATATAAGTGAGCCCAACGGGTGCGCGAGTCAGTTGCGCAGCCATCGACGTCAGACGCGGAA
GCTTCGATCAACTACGCAGACAGGTACCAAAAACAAATGTTCTCGTCACGTGGGCATGAATCTGATGCTGTTTCCCTG
CAGACAATGCGAGAGAATGAATCAGAATTCAAATATCTGCTTCACTCACGGACAGAAAGACTGTTTAGAGTGCTTTC
CCGTGTCAGAATCTCAACCCGTTTCTGTCGTCAAAAAGGCGTATCAGAACTGTGCTACATTCATCATATCATGGGA
AAGGTGCCAGACGCTTGCACTGCCTGCGATCTGGTCAATGTGGATTTGGATGACTGCATCTTTGAACAATAAATGAT
TTAAATCAGGTATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACCTCTCTGAGGGCATTGCGAGTG
TGGGACCTGAAACCTGGAGCCCCGAAACCCAAAGCCAACCAGCAAAAGCAGGACGACGGCCGGGGTCTGGTGCTTCC
TGGCTACAAGTACCTCGGACCCTTCAACGGACTCGACAAGGGGGAGCCCGTCAACGCGGCGGACGCAGCGGCCCTCG
AGCACGACAAGGCCTACGACCAGCAGCTCAAAGCGGGTGACAATCCGTACCTGCGGTATAACCACGCCGACGCCGAG
TTTCAGGAGCGTCTGCAAGAAGATACGTCTTTTGGGGGCAACCTCGGGCGAGCAGTCTTCCAGGCCAAGAAGCGGGT
TCTCGAACCTCTCGGTCTGGTTGAGGAAGGCGTAAGACGGCTCCTGGAAGAAGAGACCGGTAGAGCAGTCACCCC
AAGAACCAGACTCATCTCGGGCATCGGCAAATCAGGCCAGCAGCCCGCTAAAAAGAGACTCAATTTTGGTCAGACT
GGCGACTCAGAGTCAGTCCCCGACCCACAACCTCTCGGAGAACCTCCAGAAGCCCCCTCAGGTCTGGGACCTAATAC
AATGGCTTCAGGCGGTGGCGCTCCAATGGCAGACAATAACGAAGGCGCCGACGGAGTGGGTAATTCCTCGGGAAATT
GGCATTGCGATTCCACATGGCTGGGGGACAGAGTCATCACCACCAGCACCCGAACCTGGGCATTGCCACCTACAAC
AACCACCTCTACAAGCAAATCTCCAATGGAACATCGGGAGGAAGCACCAACGACAACACCTACTTTGGCTACAGCAC
CCCCTGGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTCTCACCACGTGACTGGCAGCGACTCATCAACAACA
ACTGGGGATTCCGGCCAAAGAGACTCAACTTCAAGCTGTTCAACATCCAGGTCAAGGAGGTTACGACGAACGAAGGC
ACCAAGACCATCGCCAATAACCTTACCAGCACCGTCCAGGTCTTTACGGACTCGGAGTACCAGCTACCGTACGTCTT
AGGCTCTGCCCACCAAGGATGCCTGCCACCGTTTCTGTCAGACGTCTTCATGGTTCCTCAGTACGGCTACCTGACGC
TCAACAATGGAAGTCAAGCGTTAGGACGTTCTTCTTTCTACTGTCTGGAATACTTCCCTTCTCAGATGCTGAGAACC
GGCAACAACCTTTCAGTTCAGCTACACTTTCGAGGACGTGCCTTTCACACAGCAGCTACGCACACAGCCAGAGTCTAGA
TCGACTGATGAACCCCTCATCGACCAGTACCTATACTACCTGGTCAGAACACAGACAACCTGGAACCTGGGGGAACCTC
AACTTTGGCATTTCAGCCAAGCAGGCCCTAGCTCAATGGCCAATCAGGCTAGAACTGGGTACCCGGGCCTTGCTAC
CGTCAGCAGCGCGTCTCCACAACCACCAACCAAAATAACAACAGCAACTTTGCGTGAGCGGGAGCTGCTAAATTCAA
GCTGAACGGGAGAGACTCGCTAATGAATCCTGGCGTGCTATGGCATCGCACAAAGACGACGAGGACCGCTTCTTTC
CATCAAGTGGCGTTCTCATATTTGGCAAGCAAGGAGCCGGGAACGATGGAGTCGACTACAGCCAGGTGCTGATTACA
GATGAGGAAGAAATTAAGCCACCAACCTGTAGCCACAGAGGAATACGGAGCAGTGGCCATCAACAACCAGGCCGC
TAACACGCAGGCGCAAACCTGGACTTGTGCATAACCAGGGAGTTATTCCTGGTATGGTCTGGCAGAACCAGGACGTGT
ACCTGCAGGGCCCTATTTGGGCTAAAAACCTCACACAGATGGCAACTTTCACCCGTCTCCTCTGATGGGTGGATTT
GGACTGAAACACCCACCTCCACAGATTCTAATTAATAAATACACCAGTGCCGGCAGATCCTCCTTACCTTCAATCA
AGCCAAGCTGAACTCTTTCATCACGCAGTACAGCACGGGACAAGTCAGCGTGGAATCGAGTGGGAGCTGCAGAAAG
AAAACAGCAAGCGCTGGAATCCAGAGATCCAGTATACTTCAAACCTACTACAAATCTACAAATGTGGACTTTGCTGTC
AATACCGAAGGTGTTTACTCTGAGCCTCGCCCCATTGGTACTCGTTACCTCACCCGTAATTTGTAATTGCCTGTTAA
TCAATAAACCGGTTAATTTCGTTTCAGTTGAACTTTGGTCTCTGCGGGCCGGCCTTAATTAACATGTGAGCAAAAGGC
CAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCA
CAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCT
CCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTCTCCCTTCGGGAAGCGTGCGC

CTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACC
CCCCGTTCAGCCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGC
CACTGGCAGTAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGG
CCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAAGAGT
TGGTAGCTCTTGATCCGGCAAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCA
GAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACCTCACGTAA
GGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTCACGTAGAAAGCCAGTCCGCAGAAACG
GTGCTGACCCCGGATGAATGTCAGCTACTGGGCTATCTGGACAAGGGAAAACGCAAGCGCAAAGAGAAAAGCAGGTAG
CTTGACAGTGGGCTTACATGGCGATAGCTAGACTGGGCGGTTTTATGGACAGCAAGCGAACCAGGAATTGCCAGCTGGG
GCGCCCTCTGGTAAGGTTGGGAAGCCCTGCAAAGTAAACTGGATGGCTTTCTTGCCGCCAAGGATCTGATGGCGCAG
GGGATCAAGATCCGATCAAGAGACAGGATGAGGATCGTTTCGCATGATTGAACAAGATGGATTGCACGCAGGTTCTC
CGGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTT
CGGCTGTACGCGAGGGGCGCCTGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAAGACGA
GGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTACTGAAGCGGGAA
GGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCC
ATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTTCGACCACCAAGCGAAACATCG
CATCGAGCGAGCACGTAACGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCG
CGCCAGCCGAACGTTCGCCAGGCTCAAGGCGAGCATGCCCCACGGCGAGGATCTCGTCGTGACCCATGGCGATGCC
TGCTTGCCGAATATCATGGTGGAATAATGGCCGCTTTTCTGGATTTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCG
CTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGC
TTTACGGTATCGCCGCTCCCGATTTCGACGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAATTAATTA
GCGGCCGCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCGCGCACATTTCCCC
GAAAAGTGCCACCTGACGTCAGATCCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGC
AAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTTCGCGAGCT
CTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCTTTG
CCCCGGGCGGCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCTACGCGTGT
CTGTCTGCACATTTTCGTAGAGCGAGTGTTCGGATACTCTAATCTCCCTAGGCAAGGTTTCATATTTGTGTAGGTTACT
TATTCTCCTTTTGTGACTAAGTCAATAATCAGAATCAGCAGGTTTGGAGTCAGCTTGGCAGGGATCAGCAGCCTGG
GTTGGAAGGAGGGGGTATAAAAGCCCCCTTCACCAGGAGAAGCCGTCACACAGATCCACAAGCTCCTGCTAGCAGGTA
AGTGCCGTGTGTGGTTCCCGCGGGCCTGGCCTCTTTACGGGTATGGCCCTTGCGTGCCTTGAATTACTGACACTGA
CATCCACTTTTTCTTTTTCTCCACAGGTATCGATTCTCTAGAGCCACCATGCAGATCGAGCTGTCTACCTGCTTCTT
CCTGTGCCTGCTGCGGTTCTGCTTCAGCGCCACCAGACGGTACTATCTGGGCGCCGTGGAAGTGAAGTGGGACTACA
TGCAGAGCGACCTGGGCGAGCTGCCCCGTGGATGCCAGATTCCCTCCAAGAGTGCCCAAGAGCTTCCCTTCAACACC
TCCGTGGTGTACAAGAAAACCTGTTCTGTTGAATTCACCGACCACCTGTTCAATATCGCCAAGCCCAGACCCCCCTG
GATGGGCCTGCTGGGACCTACAATTCAGGCCGAGGTGTACGACACCGTCGTGATCACCTGAAGAACATGGCCAGCC
ACCCCGTGTCTCTGCATGCCGTGGGAGTGTCTACTGGAAGGCCTCTGAGGGCGCCGAGTACGACGATCAGACCAGC
CAGCGCGAGAAAGAGGACGACAAGGTGTTCCCTGGCGGCAGCCACACCTACGTGTGGCAGGTGCTGAAAGAAAACGG
CCCCATGGCCTCCGACCCTCTGTGCCTGACATACAGTACCTGAGCCACGTGGACCTCGTGAAGGACCTGAACAGCG
GCCTGATCGGAGCCCTGCTCGTGTGTAGAGAGGGCAGCCTGGCCAAAGAGAAAACCCAGACCCTGCACAAGTTTCATC

CTGCTGTTGCGCGTGTTCGACGAGGGCAAGAGCTGGCACAGCGAGACAAAGAACAGCCTGATGCAGGACCGGGACGC
CGCCTCTGCTAGAGCCTGGCCCAAAATGCACACCGTGAACGGGTACGTGAACAGAAGCCTGCCCGGACTGATCGGCT
GCCACCGGAAGTCTGTGTACTGGCACGTGATCGGCATGGGCACCACCCCTGAGGTGCACAGCATCTTTCTGGAAGGA
CACACCTTTCTCGTGCAGAACCCCGGCAGGCCAGCCTGGAATCAGCCCTATCACCTTCCTGACCGCCCAGACACT
GCTGATGGACCTGGGCCAGTTTCTGCTGTTCTGCCACATCAGCTCCCACCAGCACGACGGCATGGAAGCCTACGTGA
AGGTGGACAGCTGCCCCGAGGAACCCAGCTGCGGATGAAGAACAACGAGGAAGCCGAGGACTACGACGACGACCTG
ACCGACAGCGAGATGGACGTGGTGCCTTCGACGACGATAACAGCCCCAGCTTCATCCAGATCAGAAGCGTGGCCAA
GAAGCACCCCAAGACCTGGGTGCACTATATCGCCGCCGAGGAAGAGGACTGGGATTACGCCCTCTGGTGCTGGCCC
CCGACGACAGAAGCTACAAGAGCCAGTACCTGAACAATGGCCCCAGCGGATCGGCCGGAAGTATAAGAAAGTGC GG
TTCATGGCCTACACCGACGAGACATTCAAGACCAGAGAGGCCATCCAGCACGAGAGCGGCATCCTGGGCCCTCTGCT
GTATGGCGAAGTGGGCGACACCTGCTGATCATCTTCAAGAACCAGGCCAGCAGACCCTACAACATCTACCCTCACG
GCATCACCGACGTGCGGCCCTGTACTCCAGAAGGCTGCCCAAGGGCGTGAAACACCTGAAGGACTTCCCCATCCTG
CCCGGCGAGATCTTCAAGTACAAGTGGACCGTGACCGTGGAAGATGGCCCCACCAAGAGCGACCCAGATGCCTGAC
ACGGTACTACAGCAGCTTCGTGAACATGGAACGGGACCTGGCCTCCGGCCTGATTGGCCCACTGCTGATCTGCTACA
AAGAAAGCGTGGACCAGCGGGCAACCAGATCATGAGCGACAAGCGGAACGTGATCCTGTTTAGCGTGTTGATGAG
AACCGGTCTGGTATCTGACCGAGAATATCCAGCGGTTCTGCCCAACCCTGCCGGCGTGACGTGGAAGATCCTGA
GTTCCAGGCCTCCAACATCATGCACTCCATCAATGGCTATGTGTTGACAGCCTGCAGCTGAGCGTGTGCCTGCACG
AGGTGGCCTACTGGTACATCCTGAGCATCGGGGCCAGACCGACTTCCTGTCCGTGTTCTTCTCCGGCTACACCTTC
AAGCACAAGATGGTGTACGAGGATACCCTGACCCTGTTCCTTTAGCGGCGAAACCGTGTTTCATGAGCATGGAAAA
CCCCGGCCTGTGGATCCTGGGCTGCCACAACAGCGACTTCCGGAACAGAGGCATGACCGCCCTGCTGAAGGTGTCCA
GCTGCGACAAGAACACCGGCGACTACTACGAGGACAGCTATGAGGACATCAGCGCCTACCTGCTGAGCAAGAACAAT
GCCATCGAGCCCAGAAGCTTCAGCCAGAACCCCCCGTGCTGAAGCGGCACCAGAGAGAGATCACCCGACCAACCT
GCAGTCCGACCAGGAAGAGATCGATTACGACGACACCATCAGCGTGGAATGAAGAAAGAAGATTCGACATCTACG
ACGAGGACGAGAACCAGAGCCCCCGGTCTTTCAGAAAAAGACCCGGCACTACTTCATTGCCGCTGTGGAACGGCTG
TGGGACTACGGCATGAGCAGCAGCCCTCACGTGCTGAGAAACAGGGCCAGAGCGGCAGCGTGCCCCAGTTCAAGAA
AGTGGTGTTCAGGAATTCACAGACGGCAGCTTCACCCAGCCTCTGTACCGCGGCGAGCTGAATGAGCACCTGGGAC
TGCTGGGCCCCTATATCAGAGCCGAAGTGAAGATAATATCATGGTCACCTTCGGAATCAGGCCTCCCGGCCCTAC
AGCTTCTACAGCTCCCTGATCAGCTACGAAGAGGACCAGAGACAGGGCGCTGAGCCCCGGAAGAACTTCGTGAAGCC
CAACGAGACTAAGACCTACTTTTGAAGGTGCAGCACCACATGGCCCCACAAAGGACGAGTTCGACTGCAAGGCCT
GGGCCTACTTCTCCGATGTGGACCTGGAAAAAGGACGTGCACTCTGGGCTGATCGGCCCCCTGCTCGTGTGCCACACC
AACACCCTGAATCCCGCCACGGCAGACAAGTGACAGTGACGGAATTCGCCCTGTTCTTCACCATCTTCGACGAAAC
AAAGAGCTGGTACTTCACCGAAAAACATGGAAGAAAGTGCCTGGGCTCCCTGCAACATCCAGATGGAAGATCCACCT
TCAAAGAGAACTACCGGTTCCACGCCATCAACGGCTACATCATGGACACACTGCCCGGCCTCGTGATGGCTCAGGAT
CAGCGGATCCGGTGGTATCTGCTGTCCATGGGCTCCAACGAGAACATCCACAGCATCCACTTCAGCGGCCACGTGTT
CACCGTGCGAAAAAAGAAGAGTACAAAATGGCCCTGTACAACCTGTACCCTGGGGTGTTGAGACAGTGGAATGC
TGCCCAGCAAGGCCGGCATCTGGCGGGTGAATGTCTGATCGGCGAGCATCTGCACGCTGGGATGAGCACACTGTTT
CTGGTGTACAGCAACAAGTGCCAGACACCTCTGGGCATGGCCTCTGGCCACATCCGGGACTTTCAGATCACAGCCAG
CGGCCAGTATGGCCAGTGGGCCCCAAAAGTGGCCAGACTGCACTACAGCGGCAGCATCAACGCCTGGTCCACCAAAG
AGCCCTTCAGCTGGATCAAGGTGGACCTGCTGGCTCCCATGATCATCCACGGAATCAAGACCCAGGGCGCCAGACAG

AAGTTCAGCAGCCTGTACATCAGCCAGTTCATCATCATGTACAGCCTGGACGGCAAGAAGTGGCAGACCTACCGGGG
CAATAGCACC GG CACCCTGATGGTGTTCTTCGGCAACGTGGACTCCAGCGGCATTAAGCACAACATCTTCAACCCCC
CCATCATTGCCCGGTACATCCGGCTGCACCCCACCCACTACAGCATCCGGTCCACCCTGAGAATGGAAGTGTATGGG
TGCGACCTGAACTCCTGCAGCATGCCCCCTGGGGATGGAAGCAAGGCCATCTCCGACGCCCAGATCACC GG CCTCCAG
CTACTTCACCAACATGTTCCGCCACCTGGTCCCCATCCAAGGCCCGGCTGCATCTGCAGGGCAGAAGCAATGCTTGGA
GGCCCCAAGTGAACAACCCCAAGAATGGCTGCAGGTGGACTTCCAGAAAACCATGAAAGTGACCGGGCTGACCACC
CAGGGCGTGAAGTCTCTGCTGACCTCTATGTACGTGAAAGAGTTCCTGATCTCCAGCAGCCAGGACGGCCACCAGTG
GACCCTGTTTTTCCAGAACGGCAAAGTGAAAGTGTTTCAGGGGAACCAGGACTCCTTCACCCCCGTCGTGAATAGCC
TGGACCCTCCACTGCTGACCAGATACCTGCGGATCCACCCTCAGAGTTGGGTGCACCAGATTGCTCTGCGGATGGAA
GTGCTGGGATGCGAGGCCCAGGACCTGTAAGTGAACAAGTAAATAAAGATCAGAGCTGTAGAGATCTGTGTGTTGG
TTTTTTGTGTGCGGCCGGTACCCAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCA
CTGAGGCCGGGCGACCAAAGGTGCGCCGACGCCCCGGGCTTTGGTTCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGA
GAGGGAGTGGCCGGAAGCTTGGCGTAATCATGGTCATAGCTGTTTCTGTGTGAAATTGTTATCCGCTCACAATTCC
ACACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGT
TGCGCTCACTGCCCCGCTTTCCAGTCGGGAAACCTGTCTGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGA
GGCGGTTTGCGTATTGGGCGCTCTTCCGCTGATCTCATACTAGCGAACGCCAGCAAGACGTAGCCCAGCGCGTCGGC
CCCAGATGCGCCGCGTGGCGCTGCTGGAGATGGCGGACGCGATGGATATGTTCTGCCAAGGGTTGGTTTGCGCATT
CACAGTTCTCCGCAAGAATTGATTGGCTCCAATTCTTGGAGTGGTGAATCCGTTAGCGAGGTGCCGCCCTGCTTCAT
CCCCGTGGCCCCGTTGCTCGCGTTTGCTGGCGGTGTCCCCGGAAGAAATATATTTGCATGTCTTTAGTTCTATGATGA
CACAAACCCCGCCAGCGTCTTGTCAATTGGCGAATTCCGGCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAGTCCC
CAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAGTCCCCAGGC
TCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCCTAACTCCGCCCAT
CCCCCCCCCTAACTCCGCCAGTTCGCCCATTTCTCCGCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCG
AGGCCGCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGAGAGCCTAGGCTTTTGCAAAAAGCTT
GCATGCCTGCAGGTGGCCGCCACGACCGGTGCCGCCACCATCCCCTGACCCACGCCCCCTGACCCCTCACAAGGAGA
CGACCTTCCATGACCGAGTACAAGCCCACGGTGCGCCTCGCCACCCGCGACGACGTCCCCCGGGCCGTACGCACCCT
CGCCGCCGCGTTCCGCCACTACCCCGCCACGCGCCACACCGTCGACCCGGACCGCCACATCGAGCGGGTCACCGAGC
TGCAAGAAGTCTTCTCACGCGCTCGGGCTCGACATCGGCAAGGTGTGGGTGCGGACGACGGCGCCGCGGTGGCG
GTCTGGACCACGCCGGAGAGCGTCGAAGCGGGGGCGGTGTTCCGCCGAGATCGGCCCCGCGCATGGCCGAGTTGAGCGG
TTCCCGGCTGGCCGCGCAGCAACAGATGGAAGGCCCTCTGGCGCCGACCGGCCCAAGGAGCCCGCGTGGTTCTTG
CCACCGTCGGCGTCTCGCCGACCAACAGGGCAAGGTCTGGGCAGCGCCGTCGTGCTCCCCGAGTGAGAGCGGCC
GAGCGCGCCGGGTGCCCGCTTCTGGAGACCTCCGCGCCCCGCAACCTCCCCTTCTACGAGCGGCTCGGCTTCAC
CGTCACCGCCGACGTGAGGTGCCGAAGGACCGCGCACCTGGTGCATGACCCGCAAGCCCGGTGCCTGACGCCCGC
CCCACGACCCGACGCGCCCGACCGAAAGGAGCGCACGACCCCATGGCTCCGACCGAAGCCACCCGGGGCGGCCCGC
CGACCCCGCACCCGCCCCCGAGGCCACCGACTCTAGAGGATCATAATCAGCCATACCACATTTGTAGAGGTTTTAC
TTGCTTTAAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAAT (SEQ ID NO:4)

序列表

<110> S.库斯迪奥-穆尔

D.索萨

K.文森特

<120>超大腺相关载体的产生

<130> 159792013240

<140> 尚未指定

<141> 同时附上

<150> US 62/144,862

<151> 2015-04-08

<150> US 62/220,067

<151> 2015-09-17

<160> 24

<170> 用于Windows的FastSEQ 4.0版

<210> 1

<211> 5097

<212> DNA

<213> 人工序列

<220>

<223> 合成构建体

<400> 1

```

gagctcttgg ccactccctc tctgcgcgct cgctcgetca ctgaggccgg gcgaccaaag 60
gtcgcgccgac gcccgggctt tgcccgggcg gcctcagtga gcgagcgagc gcgcagagag 120
ggagtggcca actccatcac taggggttcc tacgcgtgtc tgtctgcaca tttcgtagag 180
cgagtgttcc gatactctaa tctccctagg caaggttcat atttgtgtag gttacttatt 240
ctccttttgt tgactaagtc aataatcaga atcagcaggt ttggagtcag cttggcaggg 300
atcagcagcc tgggttgga ggagggggta taaaagcccc ttcaccagga gaagccgtca 360
cacagatcca caagctcctg ctagcaggta agtgccgtgt gtggttcccg cgggcctggc 420
ctcttttacgg gttatggccc ttgcgtgcct tgaattactg aactgacat ccactttttc 480
tttttctcca caggtatcga ttctctagag ccaccatgca gatcgagctg tctacctgct 540
tcttcctgtg cctgctgcgg ttctgcttca gcgccaccag acggtactat ctgggcgccg 600
tggaactgag ctgggactac atgcagagcg acctgggcga gctgcccgtg gatgccagat 660
tcctccaag agtgcccaag agettcccet tcaacacctc cgtggtgtac aaaaaaccc 720
tgttcgtgga attcaccgac cacctgttca atatcgcaa gccagaccc ccctggatgg 780
gcctgctggg acctacaatt caggccgagg tgtacgacac cgtcgtgatc accctgaaga 840
acatggccag ccaccccggt tctctgcatg ccgtgggagt gtcctactgg aaggectctg 900
agggcgccga gtacgacgat cagaccagcc agcgcgagaa agaggacgac aaggtgttcc 960
ctggcggcag ccacacctac gtgtggcagg tgctgaaaga aaacggcccc atggcctccg 1020

```

accctctgtg cctgacatac agctacctga gccacgtgga cctcgtgaag gacctgaaca 1080
 gcggcctgat cggagccctg ctctgtgtga gagagggcag cctggccaaa gagaaaaccc 1140
 agaccctgca caagttcatc ctgctgttcg ccgtgttcga cgagggcaag agctggcaca 1200
 gcgagacaaa gaacagcctg atgcaggacc gggacgccgc ctctgctaga gcctggccca 1260
 aaatgcacac cgtgaacggc tacgtgaaca gaagcctgcc cggactgata ggctgccacc 1320
 ggaagtctgt gtactggcac gtgatcggca tgggcaccac ccctgaggtg cacagcatct 1380
 ttctggaagg acacaccttt ctctgctgga accaccggca ggccagcctg gaaatcagcc 1440
 ctatcacctt cctgaccgcc cagacactgc tgatggacct gggccagttt ctgctgttct 1500
 gccacatcag ctcccaccag cacgacggca tggaagccta cgtgaaggtg gacagctgcc 1560
 ccgaggaacc ccagctgcgg atgaagaaca acgaggaagc cgaggactac gacgacgacc 1620
 tgaccgacag cgagatggac gtggtgcgct tcgacgacga taacagcccc agcttcatcc 1680
 agatcagaag cgtggccaag aagcacccca agacctgggt gcactatata gccgccgagg 1740
 aagaggactg ggattacgcc cctctggtgc tggccccga cgacagaagc tacaagagcc 1800
 agtacctgaa caatggcccc cagcggatcg gccggaagta taagaaagtg cggttcatgg 1860
 cctacaccga cgagacattc aagaccagag aggccatcca gcacgagagc ggcatectgg 1920
 gccctctgct gtatggcgaa gtgggcgaca ccctgctgat catcttcaag aaccaggcca 1980
 gcagacccta caacatctac cctcacggca tcaccgacgt gcggccccctg tactccagaa 2040
 ggctgcccga gggcgtgaaa cacctgaagg acttccccat cctgccccgc gagatcttca 2100
 agtacaagtg gaccgtgacc gtggaagatg gccccacaa gagcgacccc agatgcctga 2160
 cacggtacta cagcagcttc gtgaacatgg aacgggacct ggcctccggc ctgattggcc 2220
 cactgctgat ctgctacaaa gaaagcgtgg accagcgggg caaccagatc atgagcgaca 2280
 agcggaaagt gatcctgttt agcgtgttcg atgagaaccg gtcttggtat ctgaccgaga 2340
 atatccagcg gttcctgccc aacctgccc gcgtgcagct ggaagatcct gaggttccagg 2400
 cctccaacat catgcactcc atcaatggct atgtgttcga cagcctgcag ctgagcgtgt 2460
 gcctgcacga ggtggcctac tggtagatcc tgagcatcgg ggcccagacc gacttccctgt 2520
 ccgtgttctt ctccggctac accttcaagc acaagatggt gtacaggat accctgacct 2580
 tgttccccct tagcggcgaa accgtgttca tgagcatgga aaaccccggc ctgtggatcc 2640
 tgggctgcca caacagcgac ttccggaaca gaggcacgac cgccctgctg aaggtgtcca 2700
 gctgcgacaa gaacaccggc gactactacg aggacagcta tgaggacatc agcgccctacc 2760
 tgctgagcaa gaacaatgcc atcgagccca gaagcttcag ccagaacccc cccgtgctga 2820
 agcggcacca gagagagatc acccggacca ccctgcagtc cgaccaggaa gagatcgatt 2880
 acgacgacac catcagcgtg gaaatgaaga aagaagattt cgacatctac gacgaggacg 2940
 agaaccagag cccccggtcc ttccagaaaa agaccggca ctacttcatt gccgtgtgg 3000
 aacggctgtg ggactacggc atgagcagca gccctcacgt gctgagaaac agggcccaga 3060
 gcggcagcgt gccccagttc aagaaagtgg tgttccagga attcacagac ggcagcttca 3120
 cccagcctct gtaccgcggc gagctgaatg agcacctggg actgctgggc ccctatatca 3180
 gagccgaagt ggaagataat atcatggtca cttccggaa tcaggcctcc cggccctaca 3240
 gcttctacag ctccctgatc agctacgaag aggaccagag acagggcgct gagccccgga 3300
 agaacttcgt gaagcccaac gagactaaga cctacttttg gaaggtgcag caccacatgg 3360

cccctacaaa ggacgagttc gactgcaagg cctgggccta cttctccgat gtggacctgg 3420
 aaaaggacgt gcactctggg ctgatcggcc ccctgctcgt gtgccacacc aacaccctga 3480
 atccccccca cggcagacaa gtgacagtgc aggaattcgc cctgttcttc accatcttcg 3540
 acgaaacaaa gagctggtac ttcaccgaaa acatggaaag aaactgccgg gctccctgca 3600
 acatccagat ggaagatccc accttcaaag agaactaccg gttccacgcc atcaacggct 3660
 acatcatgga cacactgccc ggcctcgtga tggctcagga tcagcggatc cgggtggtatc 3720
 tgctgtccat gggctccaac gagaacatcc acagcatcca cttcagcggc cacgtgttca 3780
 ccgtgcggaa aaaagaagag taaaaaatgg ccctgtacaa cctgtaccct ggggtgttcg 3840
 agacagtgga aatgctgccc agcaaggccg gcactctggcg ggtggaatgt ctgatcggcg 3900
 agcatctgca cgctgggatg agcacactgt ttctggtgta cagcaacaag tgccagacac 3960
 ctctgggcat ggcctctggc cacatccggg actttcagat cacagccagc ggccagtatg 4020
 gccagtgggc cccaaaactg gccagaactgc actacagcgg cagcatcaac gcctggtcca 4080
 ccaagagacc cttcagctgg atcaaggtgg acctgctggc tcccatgatc atccacggaa 4140
 tcaagacca gggcgccaga cagaagtcca gcagcctgta catcagccag ttcatcatca 4200
 tgtacagcct ggacggcaag aagtggcaga cctaccgggg caatagcacc ggcaccctga 4260
 tgggtgttctt cggcaacgtg gactccagcg gcattaagca caacatcttc aaaaaaaaa 4320
 tcattgcccc gtacatccgg ctgcaccca cccactacag catccggtcc accctgagaa 4380
 tggaactgat gggctgcgac ctgaactcct gcagcatgcc cctggggatg gaaagcaagg 4440
 ccattctcga cgcccagatc accgcctcca gctacttcac caacatgttc gccacctggt 4500
 ccccatccaa ggccccgctg catctgcagg gcagaagcaa tgcttgaggg cccaagtga 4560
 acaaccccaa agaatggctg caggtggact tccagaaaac catgaaagtg accggcgatga 4620
 ccaccaggg cgtgaagtct ctgctgacct ctatgtacgt gaaagagttc ctgatctcca 4680
 gcagccagga cggccaccag tggacctgt tttccagaa cggcaaagtg aaagtgtttc 4740
 aggggaacca ggactccttc acccccgtcg tgaatagcct ggacctcca ctgctgacca 4800
 gatacctgcg gatccaccct cagagttggg tgcaccagat tgctctgcgg atggaagtgc 4860
 tgggatgcga ggcccaggac ctgtactgac actagtaata aaagatcaga gctgtagaga 4920
 tctgtgtgtt ggTTTTTgt gtgcggccgg taccaggaac ccctagtgat ggagttggcc 4980
 actccctctc tgcgcgctcg ctcgtcact gaggcggggc gaccaaaggt cgcccacgc 5040
 ccgggctttg cccgggcggc ctcagtgagc gagcgagcgc gcagagaggg agtggcc 5097

<210> 2

<211> 229

<212> DNA

<213> 人工序列

<220>

<223> 合成构建体

<400> 2

tgtctgtctg cacatttcgt agagcgagtg ttccgatact ctaatctccc ggggcaaagg 60
 tcgtattgac ttaggttact tattctcctt ttgttgacta agtcaataat cagaatcagc 120
 aggtttggag tcagcttggc agggatcagc agcctgggtt ggaaggaggg ggtataaaag 180

cccccttcacc aggagaagcc gtcacacaga tccacaagct cctgctagc 229

<210> 3

<211> 341

<212> DNA

<213> 人工序列

<220>

<223> 合成构建体

<400> 3

ctacctgctg atcgccccgc cctgtttcaa acatgtccta atactctgtc ggggcaaagg 60
tcggcagtag ttttccatct tactcaacat cctcccagtg tacgtaggat cctgtctgtc 120
tgcacatttc gtagagcgag tgttccgata ctctaatttc ccggggcaaa ggtcgtattg 180
acttaggtta cttattctcc tttgttgac taagtcaata atcagaatca gcaggtttgg 240
agtcagcttg gcagggatca gcagcctggg ttggaaggag ggggtataaa agccccctca 300
ccaggagaag ccgtcacaca gatccacaag ctctgctag c 341

<210> 4

<211> 13524

<212> DNA

<213> 人工序列

<220>

<223> 合成构建体

<400> 4

gaatgcaatt gttgttggtta acttgtttat tgcagcttat aatggttaca aataaagcaa 60
tagcatcaca aatttcacaa ataaagcatt ttttctactg cattctagtt gtggtttgtc 120
caaactcatc aatgtatctt atcatgtctg gatccgctag aactaggaat tcgctagcgg 180
taccgatatc ctagtggatc cccgtacac aggaagtgc aattttcgcg cggtttttagg 240
cggatgttgt agtaaatttg ggcgtaaccg agtaagattt ggggtgtcac gctgggtatt 300
taagcccagag tgagcacgca gggctctccat ttgaagcgg gaggtttgaa cgcgagccc 360
ccatgccggg gttttacgag attgtgatta aggtccccag cgacctgac gagcatctgc 420
ccggcatttc tgacagcttt gtgaactggg tggccgagaa ggaatgggag ttgccgccag 480
attctgacat ggatctgaat ctgattgagc aggcacccct gaccgtggcc gagaagctgc 540
agcgcgactt tctgacggaa tggcgccgtg tgagtaaggc cccggaggcc cttttctttg 600
tgcaatttga gaaggagag agctacttcc acatgcacgt gctcgtggaa accaccgggg 660
tgaaatccat ggttttgga cgtttctga gtcagattcg cgaaaaactg attcagagaa 720
tttaccgcgg gatcgagccg actttgccaa actggttcgc ggtcacaaag accagaaatg 780
gcgccggagg cgggaacaag gtggtggatg agtctacat cccaattac ttgctcccca 840
aaaccagcc tgagctccag tgggcgtgga ctaatatgga acagtattta agcgctgtt 900
tgaatctcac ggagcgtaaa cggttggtg gcgagcatct gacgcacgtg tcgcagacgc 960
aggagcagaa caaagagaat cagaatccca attctgatgc gccggtgatc agatcaaaaa 1020
cttcagccag gtacatggag ctggtcgggt ggctcgtgga caaggggatt acctcgaga 1080

```

agcagtggat ccaggaggac caggcctcat acatctcctt caatgcggcc tccaactcgc 1140
gggtcccaaat caaggctgcc ttggacaatg cgggaaagat tatgagcctg actaaaaccg 1200
cccccgacta cctgggtgggc cagcagccccg tggaggacat ttccagcaat cggatttata 1260
aaatttttga actaaacggg tacgatcccc aatatgcggc ttccgtcttt ctgggatggg 1320
ccacgaaaaa gttcggcaag aggaacacca tctggctgtt tgggcctgca actaccggga 1380
agaccaacat cgcgagggcc atagcccaca ctgtgccctt ctacgggtgc gtaaactgga 1440
ccaatgagaa ctttcccttc aacgactgtg tcgacaagat ggtgatctgg tgggaggagg 1500
ggaagatgac cgccaaggtc gtggagtcgg ccaaagccat tctcggagga agcaaggcgc 1560
gcgtggacca gaaatgcaag tcctcggccc agatagacce gactcccgtg atcgtcacct 1620
ccaacaccaa catgtgcgcc gtgattgacg ggaactcaac gaccttcgaa caccagcagc 1680
cgttgcaaga ccgatgttc aaatttgaac tcaccgcgcg tctggatcat gactttggga 1740
aggtcaccaa gcagggaagtc aaagactttt tccggtgggc aaaggatcac gtggttgagg 1800
tgagcatga attctacgtc aaaaagggtg gagccaagaa aagaccgcgc ccagtgacg 1860
cagatataag tgagcccaaa cgggtgcgcg agtcagttgc gcagccatcg acgtcagacg 1920
cggaagcttc gatcaactac gcagacaggt accaaaacaa atgttctcgt cacgtgggca 1980
tgaatctgat gctgtttccc tgcagacaat gcgagagaat gaatcagaat tcaaatatct 2040
gcttactca cggacagaaa gactgtttag agtgtttcc cgtgtcagaa tctcaaccgc 2100
tttctgtcgt caaaaaggcg tatcagaaac tgtgtacat tcatcatatc atgggaaagg 2160
tgccagacgc ttgactgcc tgcgatctgg tcaatgtgga tttgatgac tgcatctttg 2220
aacaataaat gatttaaate aggtatggct gccgatggtt atcttccaga ttggctcgag 2280
gacaacctct ctgagggcac tcgcgagtg tgggacctga aacctggagc cccgaaaccc 2340
aaagccaacc agcaaaagca ggacgacggc cggggtctgg tgcttcttgg ctacaagtac 2400
ctcggaccct tcaacggact cgacaagggg gagcccgtca acgcggcgga cgcagcggcc 2460
ctcgagcacg acaaggccta cgaccagcag ctcaaagcgg gtgacaatcc gtacctgcgg 2520
tataaccacg ccgacgccga gtttcaggag cgtctgcaag aagatactc ttttgggggc 2580
aacctcgggc gagcagtcct ccaggccaag aagcgggttc tcgaacctct cggctctggtt 2640
gaggaaggcg ctaagacggc tcctggaag aagagaccgg tagagcagtc accccaagaa 2700
ccagactcat cctcgggcac cggcaaatca ggccagcagc ccgctaaaaa gagactcaat 2760
tttggtcaga ctggcgactc agagtcagtc cccgaccac aacctctcgg agaacctcca 2820
gaagccccct caggctctgg acctaataca atggttcag gcggtggcgc tccaatggca 2880
gacaataacg aaggcgccga cggagtgggt aattctcgg gaaattggca ttgcgattcc 2940
acatggctgg gggacagagt catcaccacc agcaccgaa cctgggcatt gccacctac 3000
aacaaccacc tctacaagca aatctccaat ggaacatcgg gaggaagcac caacgacaac 3060
acctactttg gctacagcac ccctggggg tattttgact tcaacagatt ccaactgccac 3120
ttctcaccac gtgactggca gcgactcatc aacaacaact ggggattccg gccaaagaga 3180
ctcaacttca agctgttcaa catccaggtc aaggaggtta cgacgaacga aggcaccaag 3240
accatcgcca ataaccttac cagcaccgtc caggtcttta cggactcgga gtaccagcta 3300
ccgtacgtcc taggtctcgc ccaccaagga tgctgccac cgtttctcgc agacgtcttc 3360
atggttcctc agtacggcta cctgacgtc aacaatggaa gtcaagcgtt aggacgttct 3420

```

```

tctttctact gtctggaata cttcccttct cagatgctga gaaccggcaa caactttcag 3480
ttcagctaca ctttcgagga cgtgcctttc cacagcagct acgcacacag ccagagtcta 3540
gatcgactga tgaacccccct catcgaccag tacctatact acctggtcag aacacagaca 3600
actggaactg ggggaactca aacttttgga ttcagccaag caggccctag ctcaatggcc 3660
aatcaggcta gaaactgggt acccgggcct tgctaccgtc agcagcgcgt ctccacaacc 3720
accaaccaaa ataacaacag caactttgcg tggacgggag ctgctaaatt caagctgaac 3780
gggagagact cgctaataaa tcctggcgtg gctatggcat cgcacaaaaga cgacaggagc 3840
cgcttctttc catcaagtgg cgttctcata tttggcaagc aaggagccgg gaacgatgga 3900
gtcgactaca gccagggtgt gattacagat gaggaagaaa tttaaagccac caaccctgta 3960
gccacagagg aatacggagc agtggccatc aacaaccagg ccgctaacac gcaggcgcaa 4020
actggacttg tgcataacca gggagttatt cctggtatgg tctggcagaa ccgggacgtg 4080
tacctgcagg gccctatttg ggetaaaaata cctcacacag atggcaactt tcaccctgt 4140
cctctgatgg gtggatttgg actgaaacac ccacctccac agattctaatt taaaaataca 4200
ccagtgccgg cagatcctcc tcttacctc aatcaagcca agctgaactc tttcatcacg 4260
cagtacagca cgggacaagt cagcgtggaa atcgagtggg agctgcagaa agaaaacagc 4320
aagcgtgga atccagagat ccagtatact tcaaactact acaaactctac aaatgtggac 4380
tttgcgtgca ataccgaagg tgtttactct gagcctcgcc ccatttgtac tcgttacctc 4440
acccgtaatt tgtaattgcc tgtaaatcaa taaaccggtt aattcgtttc agttgaactt 4500
tggtctctgc gggccggcct taattaacat gtgagcaaaa ggccagcaaa aggccaggaa 4560
ccgtaaaaag gccgcgttgc tggcgttttt ccataggtc cgccccctg acgagcatca 4620
caaaaatcga cgctcaagtc agagggtggc aaacccgaca ggactataaa gataccaggc 4680
gtttccccct ggaagctccc tcgtgcgtc tcctgttccg accctgccgc ttaccggata 4740
cctgtccgcc tttctccctt cgggaagcgt ggcgttttct catagctcac gctgtaggta 4800
tctcagttcg gtgtaggtcg ttcgtccaa gctgggctgt gtgcacgaac cccccgttca 4860
gcccgaaccg tcgccttat ccggttaacta tcgtcttgag tccaaccgg taagacacga 4920
cttatcgcca ctggcagtag ccactggtaa caggattagc agagcgaggt atgtaggcgg 4980
tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagaa cagtatttgg 5040
tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg 5100
caaacaaacc accgctggta gcggtggtt ttttgttgc aagcagcaga ttacgcgcag 5160
aaaaaaagga tctcaagaag atcctttgat cttttctac gggctctgac ctcaagtggaa 5220
cgaaaactca cgtaaggga ttttggtcat gagattatca aaaaggatct tcacctagat 5280
ccttttcacg tagaaaagcca gtccgcagaa acggtgctga ccccgatga atgtcagcta 5340
ctgggctatc tggacaaggg aaaacgcaag cgcaaagaga aagcaggtag cttgcagtgg 5400
gcttacatgg cgatagctag actgggcggt tttatggaca gcaagcgaac cggaattgcc 5460
agctggggcg ccctctggta aggttgggaa gccctgcaaa gttaaactgga tggttttctt 5520
gccccaagg atctgatggc gcaggggatc aagatccgat caagagacag gatgaggatc 5580
gtttcgcag attgaacaag atggattgca cgcaggttct ccggccgctt ggggtggagag 5640
gctattcggc tatgactggg cacaacagac aatcggtgc tctgatgcc ccgtgttccg 5700
gctgtcagcg caggggcgcc tggttctttt tgtcaagacc gacctgtccg gtgccctgaa 5760

```

tgaactgcaa gacgaggcag cgcggctatc gtggctggcc acgacgggcg ttccttgccg 5820
 agctgtgctc gacgttgtca ctgaagcggg aagggaactgg ctgctattgg gcgaagtgcc 5880
 ggggcaggat ctctgtcat ctacacctgc tcctgccgag aaagtatcca tcatggctga 5940
 tgcaatgcgg cggctgcata cgcttgatcc ggctacctgc ccattcgacc accaagcgaa 6000
 acatcgcatc gagcgagcac gtactcggat ggaagccggt cttgtcgatc aggatgatct 6060
 ggacgaagag catcaggggc tcgcgccagc cgaactgttc gccaggctca aggcgagcat 6120
 gcccgcggc gaggatctcg tcgtgaccca tggcgatgcc tgcttgccga atatcatggt 6180
 ggaaaatggc cgcttttctg gattcatcga ctgtggccgg ctgggtgtgg cggaccgcta 6240
 tcaggacata gcgttggtta cccgtgatat tgctgaagag cttggcggcg aatgggctga 6300
 ccgcttcctc gtgctttacg gtatcgccgc tcccgattcg cagcgcatcg ctttctatcg 6360
 ctttcttgac gaggttcttct gaattaatta agcggccgct catgagcgga tacatatttg 6420
 aatgtattta gaaaaataaa caaatagggg ttccgcgcac atttccccga aaagtgccac 6480
 ctgacgtcag atccggtgcg ggctctctcg ctattacgcc agctggcgaa agggggatgt 6540
 gctgcaaggc gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaacg 6600
 acggccagtg aattcgcgag ctcttgccca ctccctctct gcgcgctcgc tcgctcactg 6660
 aggcggggcg accaaaggtc gcccgcgcc cgggctttgc ccgggcggcc tcagtgcg 6720
 agcgagcgcg cagagaggga gtggccaact ccatcactag gggttcctac gcgtgtctgt 6780
 ctgcacattt cgtagagcga gtgttccgat actctaactt ccctaggcaa gttcatatt 6840
 tgtgtagggt acttattctc cttttgttga ctaagtcaat aatcagaatc agcaggtttg 6900
 gagtcagctt ggcagggatc agcagcctgg gttggaagga gggggtataa aagcccttc 6960
 accaggagaa gccgtcacac agatccacaa gctcctgcta gcagtaagt gccgtgtgtg 7020
 gttcccgcg gcctggcctc ttacgggtt atggcccttg cgtgccttga attactgaca 7080
 ctgacatcca ctttttcttt ttctccacag gtatcgatc tctagagcca ccatgcagat 7140
 cgagctgtct acctgcttct tcctgtgcct gctgcggttc tgcttcagcg ccaccagacg 7200
 gtactatctg ggcgccgtgg aactgagctg ggactacatg cagagcgacc tgggcgagct 7260
 gcccgtggat gccagattcc ctccaagagt gccaagagc ttcccttca acacctccgt 7320
 ggtgtacaag aaaaccctgt tcgtggaatt caccgaccac ctgttcaata tcgccaagcc 7380
 cagaccccc tggtggggc tgctgggacc tacaattcag gccaggtgt acgacaccgt 7440
 cgtgatcacc ctgaagaaca tggccagcca cccgtgtct ctgcatgccg tgggagtgtc 7500
 ctactggaag gcctctgagg gcgccgagta cgacgatcag accagccagc gcgagaaaga 7560
 ggacgacaag gtgttccctg gcggcagcca cacctacgtg tggcaggtgc tgaaagaaaa 7620
 cggcccatg gcctccgacc ctctgtgcct gacatacagc tacctgagcc acgtggacct 7680
 cgtgaaggac ctgaacagcg gcctgatcgg agccctgctc gtgtgtagag agggcagcct 7740
 ggccaaagag aaaaccaga ccctgcacaa gttcactctg ctgttcgccc gtgttcagca 7800
 gggcaagagc tggcacagcg agacaaagaa cagcctgatg caggaccggg acgccgctc 7860
 tgctagagcc tggcccaaaa tgcacaccgt gaacggctac gtgaacagaa gcctgcccgg 7920
 actgatcggc tgccaccgga agtctgtgta ctggcacgtg atcggcattg gcaccaccc 7980
 tgaggatgcac agcatcttct tggaaggaca cactttctc gtgcggaacc accggcagcg 8040
 cagcctggaa atcagcccta tcaccttct gaccgccag aactgctga tggacctggg 8100

ccagtttctg ctgttctgcc acatcagctc ccaccagcac gacggcatgg aagcctacgt 8160
 gaaggtggac agctgccccg aggaacccca gctgcggatg aagaacaacg aggaagccga 8220
 ggactacgac gacgacctga ccgacagcga gatggacgtg gtgcgcttcg acgacgataa 8280
 cagccccagc ttcatccaga tcagaagcgt ggccaagaag caccccaaga cctgggtgca 8340
 ctatatcgcc gccgaggaag aggactggga ttacgcccct ctggtgctgg ccccgacga 8400
 cagaagctac aagagccagt acctgaacaa tggccccag cggatcggcc ggaagtataa 8460
 gaaagtgcgg ttcatggcct acaccgacga gacattcaag accagagagg ccatccagca 8520
 cgagagcggc atcctgggccc ctctgctgta tggcgaagtg ggcgacaccc tgctgatcat 8580
 cttcaagaac caggccagca gacctacaa catctacct caccgcatca ccgacgtgcg 8640
 gcccctgtac tccagaaggc tgcccaaggc cgtgaaacac ctgaaggact tccccatcct 8700
 gcccggcgag atcttcaagt acaagtggac cgtgaccgtg gaagatggcc ccaccaagag 8760
 cgaccccaga tgctgacac ggtactacag cagcttcgtg aacatggaac gggacctggc 8820
 ctccggcctg attggccac tgctgatctg ctacaaagaa agcgtggacc agcggggcaa 8880
 ccagatcatg agcgacaagc ggaacgtgat cctgtttagc gtgttcgatg agaaccggtc 8940
 ctggtatctg accgagaata tccagcgggt cctgcccac cctgccggcg tgcagctgga 9000
 agatcctgag ttccaggcct ccaacatcat gcactccatc aatggctatg tgttcgacag 9060
 cctgcagctg agcgtgtgcc tgcacgaggt ggcctactgg tacatcctga gcatcggggc 9120
 ccagaccgac ttctgtccg tgttcttctc cggtacacc ttcaagcaca agatggtgta 9180
 cgaggatacc ctgacctgt tccccttag cggcgaaacc gtgttcatga gcatggaaaa 9240
 ccccggcctg tggatcctgg gctgccacaa cagcgacttc cggaacagag gcatgaccgc 9300
 cctgctgaag gtgtccagct gcgacaagaa caccggcgac tactacgagg acagctatga 9360
 ggacatcagc gcctacctgc tgagcaagaa caatgccatc gagcccagaa gcttcagcca 9420
 gaaccccccc gtgtgaagc ggcaccagag agagatcacc cggaccaccc tgcagtccga 9480
 ccaggaagag atcgattacg acgacaccat cagcgtggaa atgaagaaag aagatttcga 9540
 catctacgac gaggacgaga accagagccc ccggtccttt cagaaaaaga cccggcacta 9600
 cttcattgcc gctgtggaac ggctgtggga ctacggcatg agcagcagcc ctcacgtgct 9660
 gaaaaacagg gccagagcgc gcagcgtgcc ccagttcaag aaagtgggtg tccaggaatt 9720
 cacagacggc agcttcaccc agcctctgta ccgcggcgag ctgaatgagc acctgggact 9780
 gctgggcccc tatatcagag ccgaagtgga agataatata atggtcacct tccggaatca 9840
 ggccctcccg ccctacagct tctacagctc cctgatcagc tacgaagagg accagagaca 9900
 gggcgctgag ccccggaaga acttcgtgaa gcccacagag actaagacct acttttggaa 9960
 ggtgcagcac cacatggccc ctacaaagga cgagttcgac tgcaaggcct gggcctactt 10020
 ctccgatgtg gacctggaaa aggacgtgca ctctgggctg atcgggcccc tgctcgtgtg 10080
 ccacaccaac accctgaatc ccgcccacgg cagacaagtg acagtgcagg aattcgccct 10140
 gttcttcacc atcttcgacg aaacaaagag ctggtacttc accgaaaaca tggaaagaaa 10200
 ctgccgggct ccctgcaaca tccagatgga agatcccacc ttcaaagaga actaccggtt 10260
 ccacgccatc aacggctaca tcatggacac actgcccggc ctcgtgatgg ctcaggatca 10320
 gcggatccgg tggatctgc tgtccatggg ctccaacgag aacatccaca gcatccactt 10380
 cagcggccac gtgttcaccg tgcggaaaaa agaagagtac aaaatggccc tgtacaacct 10440

gtaccctggg gtgttcgaga cagtggaaat gctgcccagc aaggccggca tctggcgggt 10500
 ggaatgtctg atcggcgagc atctgcacgc tgggatgagc aactgtttc tgggtgtacag 10560
 caacaagtgc cagacacctc tgggcatggc ctctggccac atccgggact ttcagatcac 10620
 agccagcggc cagtatggcc agtgggcccc aaaactggcc agactgcact acagcggcag 10680
 catcaacgcc tgggtccacca aagagccctt cagctggatc aaggtggacc tgctggctcc 10740
 catgatcatc cacggaatca agaccaggc cgccagacag aagttcagca gcctgtacat 10800
 cagccagttc atcatcatgt acagcctgga cggcaagaag tggcagacct accggggcaa 10860
 tagcaccggc accctgatgg tgttcttcgg caacgtggac tccagcggca ttaagcacia 10920
 catcttcaac cccccatca ttgcccggta catecggtg caccceccc actacagcat 10980
 ccggtccacc ctgagaatgg aactgatggg ctgcgacctg aactcctgca gcatgcccct 11040
 ggggatggaa agcaaggcca tctccgacgc ccagatcacc gctccagct acttcacaa 11100
 catgttcgcc acctggtccc catccaaggc ccggtgcat ctgcagggca gaagcaatgc 11160
 ttggaggccc caagtgaaca accccaaaga atggtgcag gtggacttcc agaaaacat 11220
 gaaagtgacc ggcgtgacca cccaggcgct gaagtctctg ctgaccteta tgtactgaa 11280
 agagttcctg atctccagca gccaggacgg ccaccagtgg accctgtttt tccagaacgg 11340
 caaagtgaia gtgtttcagg ggaaccagga ctcttcacc cccgtcgtga atagcctgga 11400
 cctccactg ctgaccagat acctgcggat ccaccctcag agttgggtgc accagattgc 11460
 tctgcggatg gaagtgcctg gatgcgaggc ccaggacctg tactgacaac tagtaataaa 11520
 agatcagagc ttagagatc tgtgtgttg tttttgtgt gcggccggta cccaggaacc 11580
 cctagtgatg gattggcca ctccctctct gcgcgctcgc tcgctcactg aggcggggcg 11640
 accaaaggct gcccagccc cgggctttgg tcgggcggcc tcagtgcgcg agcgcgcgcg 11700
 cagagaggga gtggccggaa gcttggcgta atcatggtca tagctgtttc ctgtgtgaaa 11760
 ttgttatccg ctcaaatc cacacaacat acgagccgga agcataaagt gtaaagcctg 11820
 gggtcctaa tgagtgcct aactcacatt aattgcgtt gcctcactgc ccgctttcca 11880
 gtcgggaaac ctgtcgtgcc agctgcatta atgaatcggc caacgcgcgg ggagaggcgg 11940
 tttgcgtatt gggcgctctt ccgctgatct catactagcg aacgccagca agacgtagcc 12000
 cagcgcgtcg gccccgagat gcgcccgtg cggctgctgg agatggcgga cgcgatggat 12060
 atgttctgcc aagggttgg ttgcgcattc acagttctcc gcaagaattg attggctcca 12120
 attcttggag tggatgaatc gttagcgagg tgccgcccgt cttcatcccc gtggcccgtt 12180
 gctcgcgttt gctggcgggt tccccggaag aaatatattt gcattgtttt agttctatga 12240
 tgacacaaac cccgcccagc gtcttgtcat tggcgaatc cggctgtgga atgtgtgtca 12300
 gttagggtgt ggaaagtccc caggtcctcc agcaggcaga agtatgaaa gcatgcatct 12360
 caattagtca gcaaccagg gtggaaagtc cccaggtcc ccagcaggca gaagtatgca 12420
 aagcatgcat ctcaattagt cagcaaccat agtcccgc ctaactccgc ccatcccgc 12480
 cctaactccg cccagttccg cccattctcc gcccattggc tgactaattt tttttattta 12540
 tgcagaggcc gaggcgcct cggcctctga gtattccag aagtagtgag gaggttttt 12600
 tggaggccta ggcttttgca aaaagcttgc atgcctgcag gtcggccgcc acgaccggtg 12660
 ccgccaccat cccctgacct acgcccctga cccctcacia ggagacgacc ttccatgacc 12720
 gagtacaagc ccacggtgcg cctcgccacc cgcgacgacg tccccgggc cgtacgcacc 12780

ctcgccgccc cggtcgccga ctaccccgcc acgcgccaca ccgtcgaccc ggaccgccac 12840
atcgagcggg tcaccgagct gcaagaactc ttctcagc gcgtcgggct cgacatcggc 12900
aaggtgtggg tcgcggacga cggcgccgcg gtggcggtct ggaccacgcc ggagagcgtc 12960
gaagcggggg cgggtgttcgc cgagatcggc ccgcgcatgg ccgagttgag cggttcccgg 13020
ctggccgcgc agcaacagat ggaaggcctc ctggcgccgc accggcccaa ggagcccgcg 13080
tggttcctgg ccaccgtcgg cgtctcgccc gaccaccagg gcaagggtct gggcagcgcc 13140
gtcgtgctcc ccggagtgga ggcggccgag cgcgcggggg tgcccgcctt cctggagacc 13200
tccgcgcccc gcaacctccc cttctacgag cggtcgggt tcaccgtcac cgccgacgtc 13260
gaggtgcccc aaggaccgcg cacctggtgc atgaccgca agcccgggtgc ctgacgcccc 13320
ccccacgacc cgcagcgcgc gaccgaaagg agcgcacgac cccatggctc cgaccgaagc 13380
cacccggggc ggccccgcgc accccgcacc cgccccgag gccaccgac tctagaggat 13440
cataatcagc cataccacat ttgtagaggt tttacttget ttaaaaaacc tcccacacct 13500
ccccctgaac ctgaaacata aaat 13524

<210> 5

<211> 17

<212> DNA

<213> 人工序列

<220>

<223> 合成构建体

<400> 5

gacgtggtgc gcttcga 17

<210> 6

<211> 20

<212> DNA

<213> 人工序列

<220>

<223> 合成构建体

<400> 6

gggcgtaatc ccagtcctct 20

<210> 7

<211> 22

<212> DNA

<213> 人工序列

<220>

<223> 合成构建体

<400> 7

aagcgtggcc aagaagcacc cc 22

<210> 8

<211> 21

<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 8
gttgccattg ctacaggcat c 21
<210> 9
<211> 19
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 9
actcgccttg atcgttggg 19
<210> 10
<211> 28
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 10
acgctcgtcg ttggtatgg cttcattc 28
<210> 11
<211> 17
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 11
ggaccgccac atcgagc 17
<210> 12
<211> 15
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 12
ccccgcttcg acgct 15
<210> 13

<211> 28
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 13
tcaccgagct gcaagaactc ttcctcac 28
<210> 14
<211> 23
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 14
gaccaggcct catacatctc ctt 23
<210> 15
<211> 18
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 15
ggcagccttg atttggga 18
<210> 16
<211> 20
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 16
aatgcggcct ccaactcgcg 20
<210> 17
<211> 19
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 17
caacacggcg accctacaa 19

<210> 18
<211> 28
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 18
tccaataactg tcttgcaata tacacagg 28
<210> 19
<211> 28
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 19
tgcacggaac tgaacacttc actgcaag 28
<210> 20
<211> 24
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 20
ccgtcgtgaa tagcctggac cctc 24
<210> 21
<211> 28
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 21
atctgtgtgt tggttttttg tgtgcggc 28
<210> 22
<211> 29
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 22

aatcccagtc ctcttcctcg gcggcgata 29

<210> 23

<211> 29

<212> DNA

<213> 人工序列

<220>

<223> 合成构建体

<400> 23

agtatcggaa cactcgctct acgaaatgt 29

<210> 24

<211> 78

<212> DNA

<213> 人工序列

<220>

<223> 合成构建体

<400> 24

cactccctct ctgcgcgtc gctcgctcac tgaggccggg cgaccaaagg tcgcccacgc 60

ccgggctttg cccgggcg 78

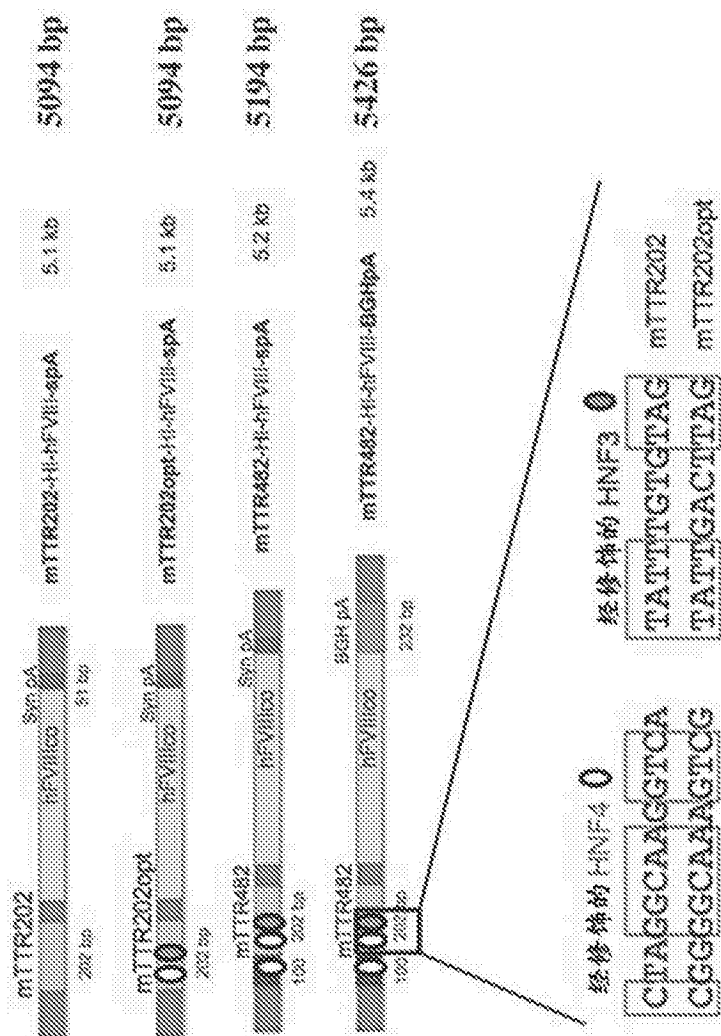


图 1A

```

mTTR202(ab)+ 1 CTACCTGCTGATCGCCCGGCCCCCTGTTCAAAACATGTCTTAATATCTCTCTCCGGCAAGG 60
mTTR202(ab) 1 ..... 60
mTTR202(a) 1 ..... 60
mTTR202(b) 1 ..... 60
mTTR202(wt) 1 ..... 60

mTTR202(ab)+ 81 TCGGCAGTASGTTTTCCATCTTACTCAGATCTCTCCAGTGTACGTACGATCTCTCTCTCT 120
mTTR202(ab) 1 ..... TCTCTCTCT 8
mTTR202(a) 1 ..... TCTCTCTCT 8
mTTR202(b) 1 ..... TCTCTCTCT 8
mTTR202(wt) 1 ..... TCTCTCTCT 8
                      xxxxxxxx

mTTR202(ab)+ 131 TGCACATTTCTAGAGCGAGTGTTCGGATATCTTAATCTCTCCGGGCAAGGCTCTATTTG 180
mTTR202(ab) 9 TGCACATTTCTAGAGCGAGTGTTCGGATATCTTAATCTCTCCGGGCAAGGCTCTATTTG 60
mTTR202(a) 9 TGCACATTTCTAGAGCGAGTGTTCGGATATCTTAATCTCTCCGGGCAAGGCTCTATTTG 60
mTTR202(b) 9 TGCACATTTCTAGAGCGAGTGTTCGGATATCTTAATCTCTCCGGGCAAGGCTCTATTTG 60
mTTR202(wt) 9 TGCACATTTCTAGAGCGAGTGTTCGGATATCTTAATCTCTCCGGGCAAGGCTCTATTTG 60
                      xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx

mTTR202(ab)+ 191 ACTTAGCTTACTTATTTCTCTTTTCTTCACTAAGTCAATAATCAGAAATCAGCAGCTTTTC 240
mTTR202(ab) 69 ACTTAGCTTACTTATTTCTCTTTTCTTCACTAAGTCAATAATCAGAAATCAGCAGCTTTTC 120
mTTR202(a) 69 ACTTAGCTTACTTATTTCTCTTTTCTTCACTAAGTCAATAATCAGAAATCAGCAGCTTTTC 120
mTTR202(b) 69 ACTTAGCTTACTTATTTCTCTTTTCTTCACTAAGTCAATAATCAGAAATCAGCAGCTTTTC 120
mTTR202(wt) 69 ACTTAGCTTACTTATTTCTCTTTTCTTCACTAAGTCAATAATCAGAAATCAGCAGCTTTTC 120
                      xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx

mTTR202(ab)+ 241 ACTCAGCTTTCGAGGATCAGCAGCCTGGCTTGGAAAGAGGGGCTATAAAGGCTCTTCA 300
mTTR202(ab) 129 ACTCAGCTTTCGAGGATCAGCAGCCTGGCTTGGAAAGAGGGGCTATAAAGGCTCTTCA 180
mTTR202(a) 129 ACTCAGCTTTCGAGGATCAGCAGCCTGGCTTGGAAAGAGGGGCTATAAAGGCTCTTCA 180
mTTR202(b) 129 ACTCAGCTTTCGAGGATCAGCAGCCTGGCTTGGAAAGAGGGGCTATAAAGGCTCTTCA 180
mTTR202(wt) 129 ACTCAGCTTTCGAGGATCAGCAGCCTGGCTTGGAAAGAGGGGCTATAAAGGCTCTTCA 180
                      xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx

mTTR202(ab)+ 301 CCAGGAGAGGCTCTCAGACAGATCCAGAGCTCTCTAGCAGCTAAGTCTCTCTCTCT 360
mTTR202(ab) 189 CCAGGAGAGGCTCTCAGACAGATCCAGAGCTCTCTAGCAGCTAAGTCTCTCTCTCT 240
mTTR202(a) 189 CCAGGAGAGGCTCTCAGACAGATCCAGAGCTCTCTAGCAGCTAAGTCTCTCTCTCT 240
mTTR202(b) 189 CCAGGAGAGGCTCTCAGACAGATCCAGAGCTCTCTAGCAGCTAAGTCTCTCTCTCT 240
mTTR202(wt) 189 CCAGGAGAGGCTCTCAGACAGATCCAGAGCTCTCTAGCAGCTAAGTCTCTCTCTCT 240
                      xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx

mTTR202(ab)+ 361 TTCCCGCGGGGCTGGCTCTTTACGGGTTATGGCCCTTGGCTGCTTCAATTACTGACAC 420
mTTR202(ab) 249 TTCCCGCGGGGCTGGCTCTTTACGGGTTATGGCCCTTGGCTGCTTCAATTACTGACAC 300
mTTR202(a) 249 TTCCCGCGGGGCTGGCTCTTTACGGGTTATGGCCCTTGGCTGCTTCAATTACTGACAC 300
mTTR202(b) 249 TTCCCGCGGGGCTGGCTCTTTACGGGTTATGGCCCTTGGCTGCTTCAATTACTGACAC 300
mTTR202(wt) 249 TTCCCGCGGGGCTGGCTCTTTACGGGTTATGGCCCTTGGCTGCTTCAATTACTGACAC 300
                      xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx

mTTR202(ab)+ 421 TGACATCCACTTTTTCCTTTTCTCT 445
mTTR202(ab) 369 TGACATCCACTTTTTCCTTTTCTCT 333
mTTR202(a) 369 TGACATCCACTTTTTCCTTTTCTCT 333
mTTR202(b) 369 TGACATCCACTTTTTCCTTTTCTCT 333
mTTR202(wt) 369 TGACATCCACTTTTTCCTTTTCTCT 333
                      xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx

```

mTTR202(ab)+ = "482": HNF-3和HNF-4域变化加100bp远端mTTR增强子

mTTR202(ab) = HNF-3和HNF-4域变化

mTTR202(a) = HNF-4域变化

mTTR202(b) = HNF-3域变化

mTTR202(wt) = 野生型小鼠转甲扶腺素蛋白启动子序列

图1B

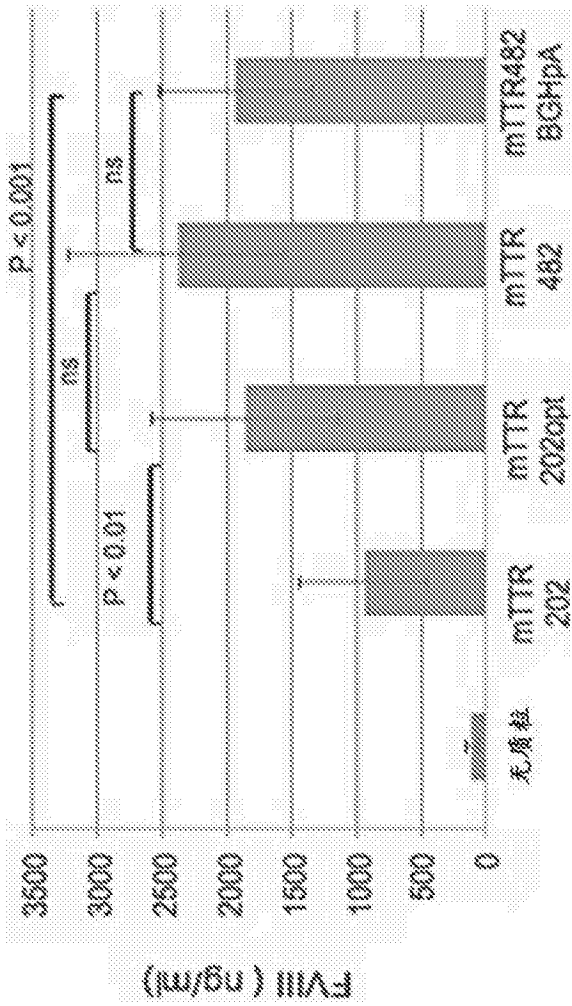


图1C

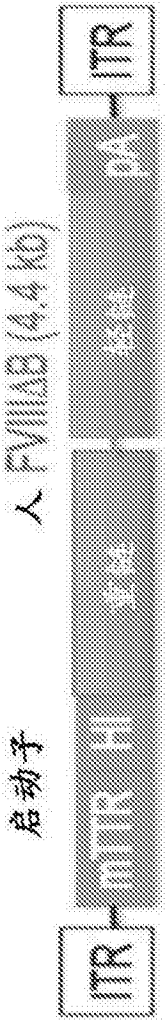


图1D

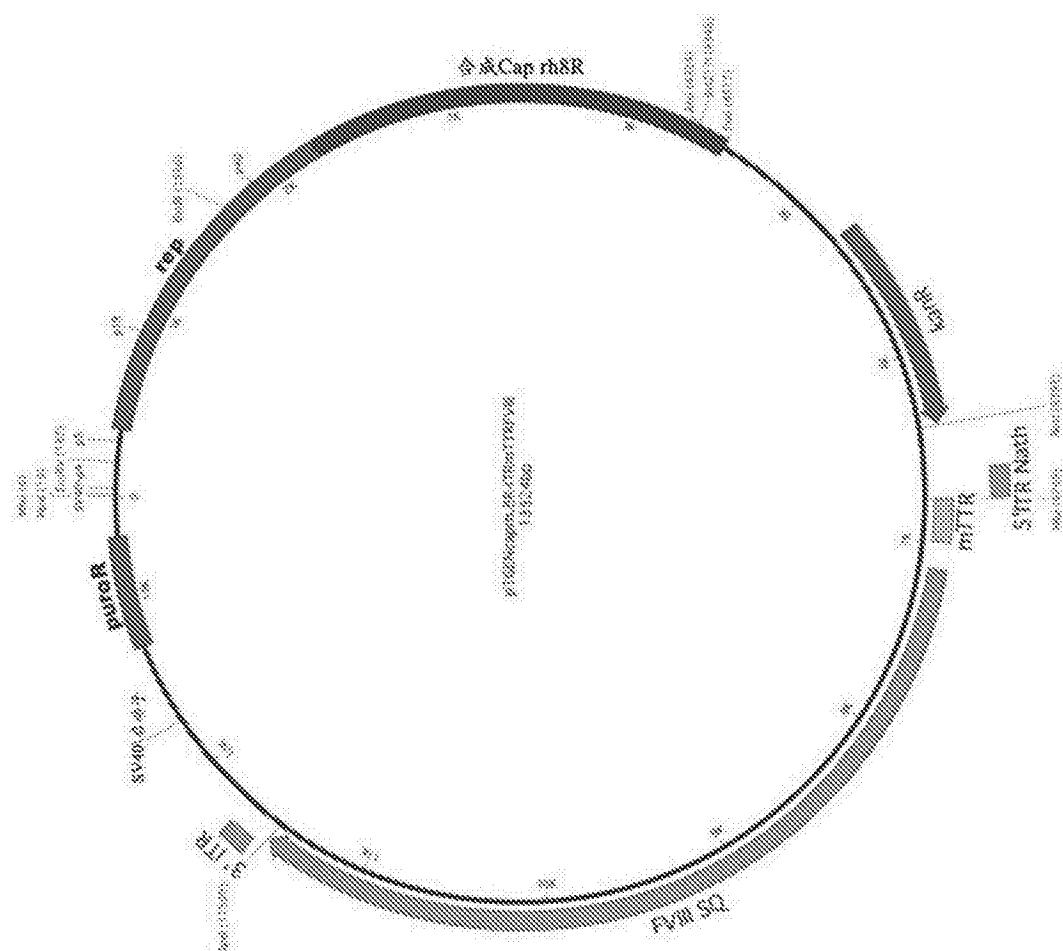


图1E

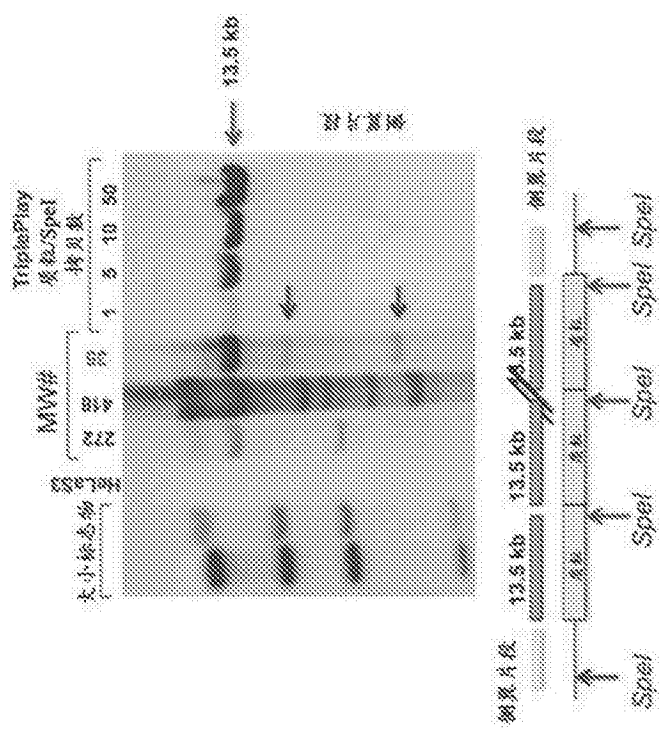


图2A

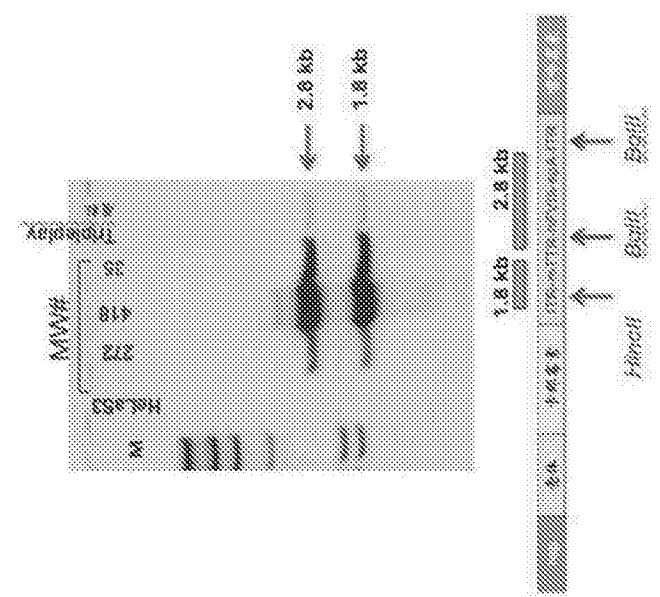


图2B

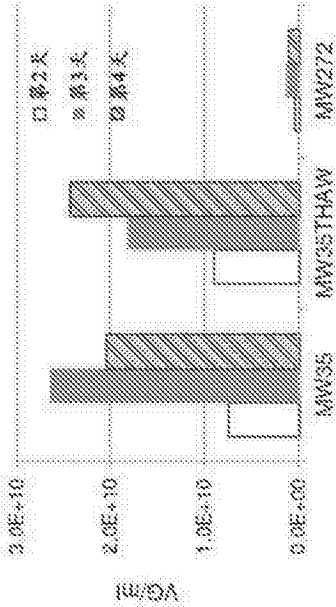


图3A

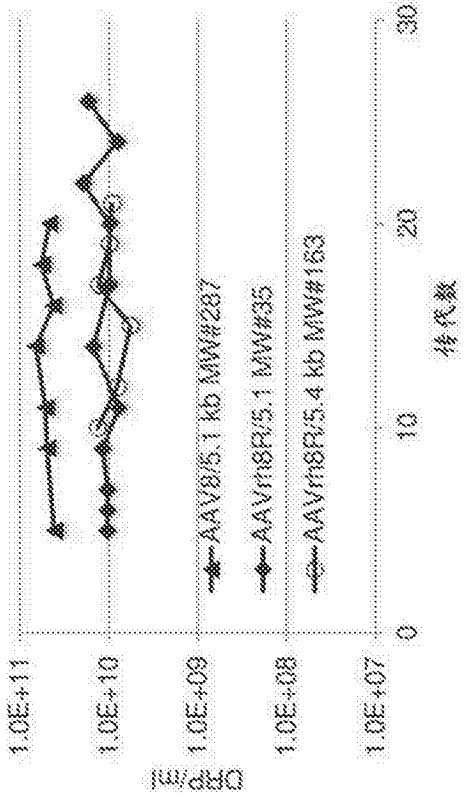


图3B

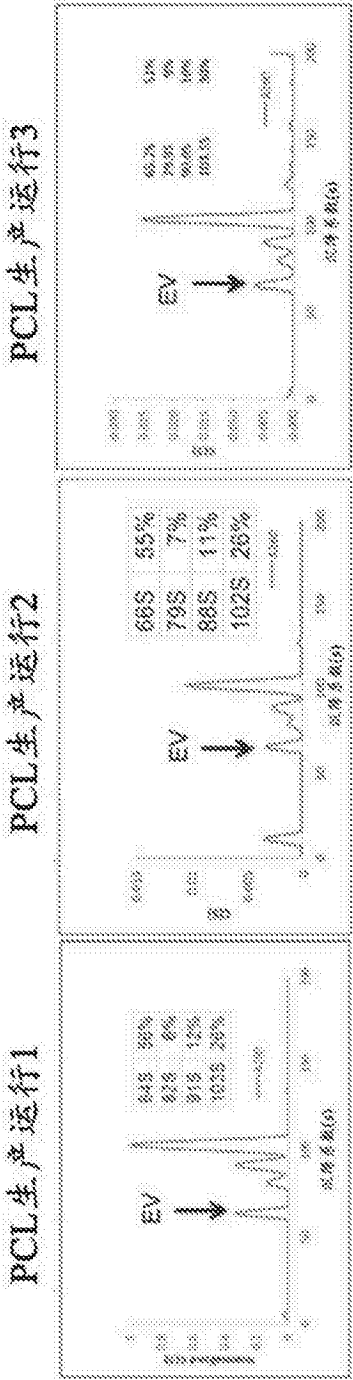


图4A

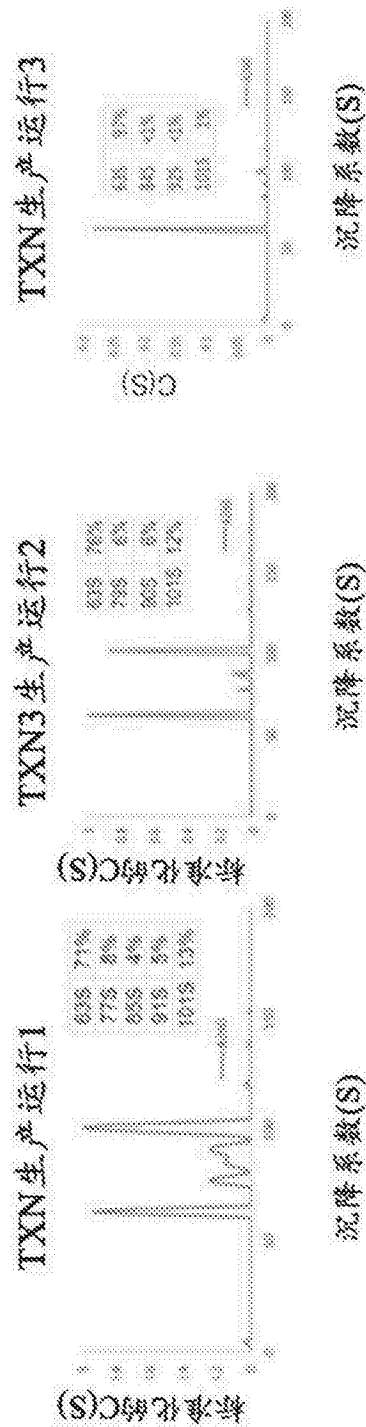


图4B

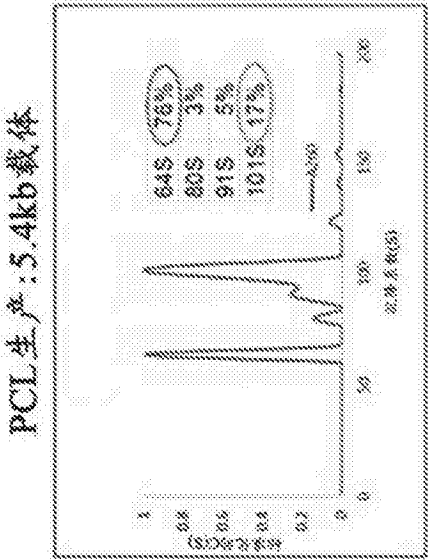


图5A

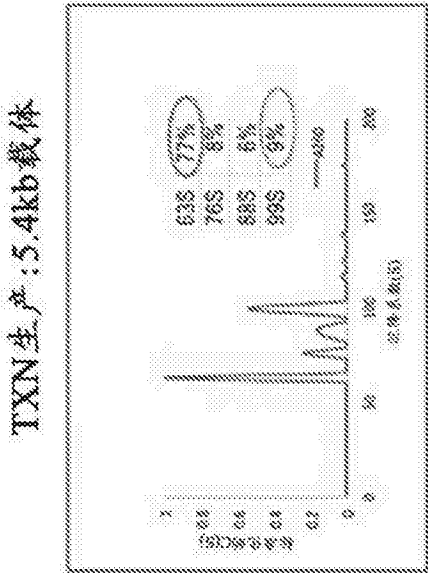


图5B

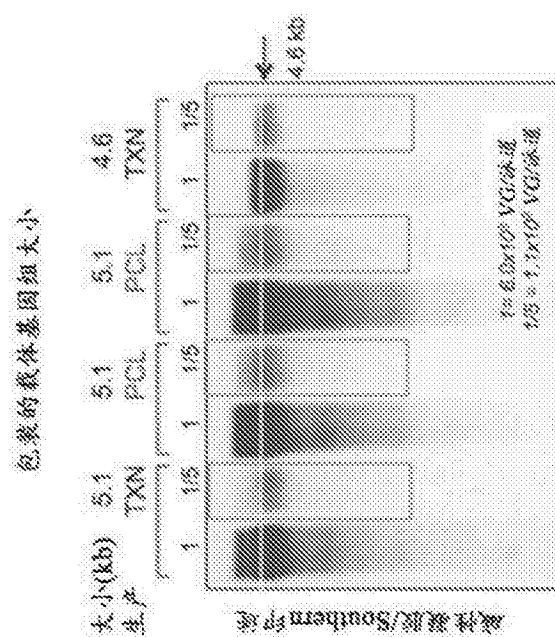


图6A

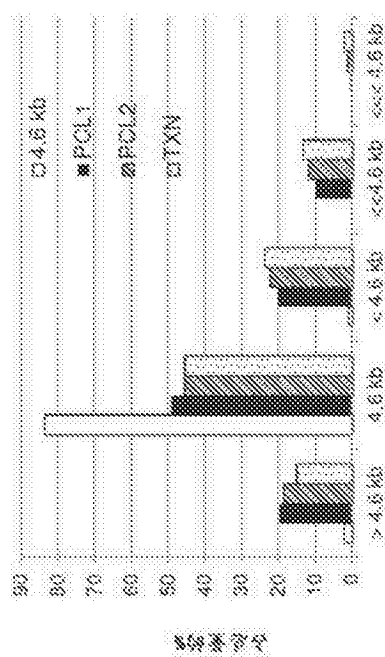


图6B

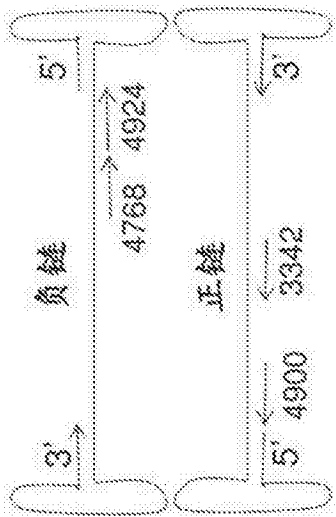


图7A

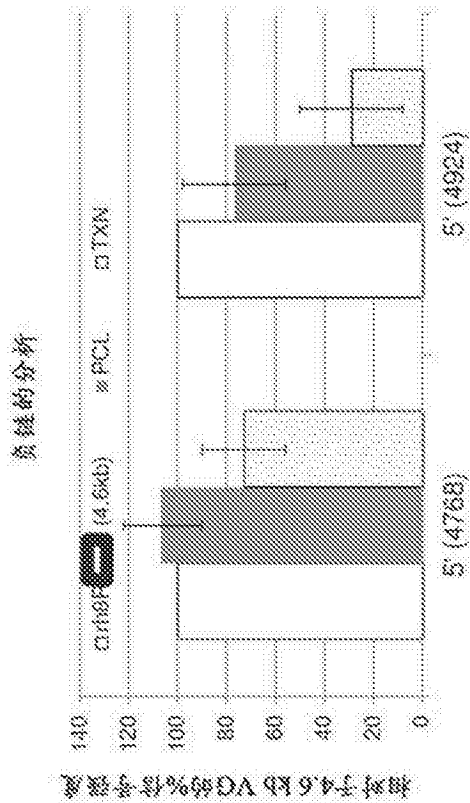


图7B

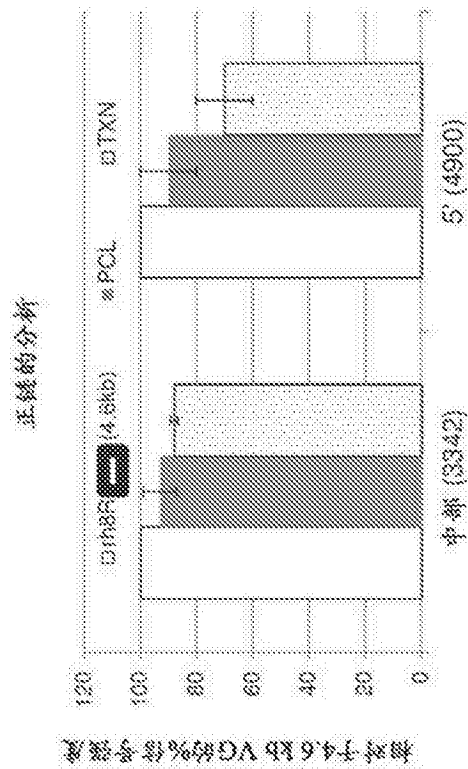


图7C

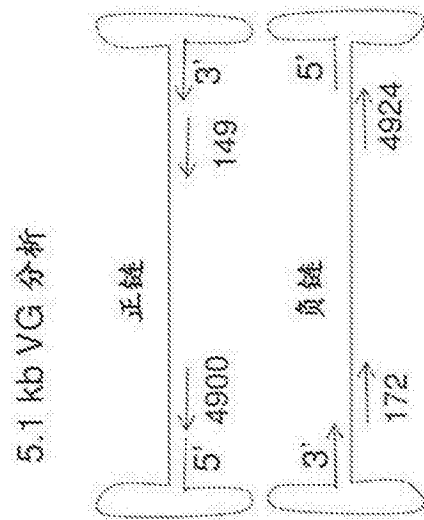


图8A

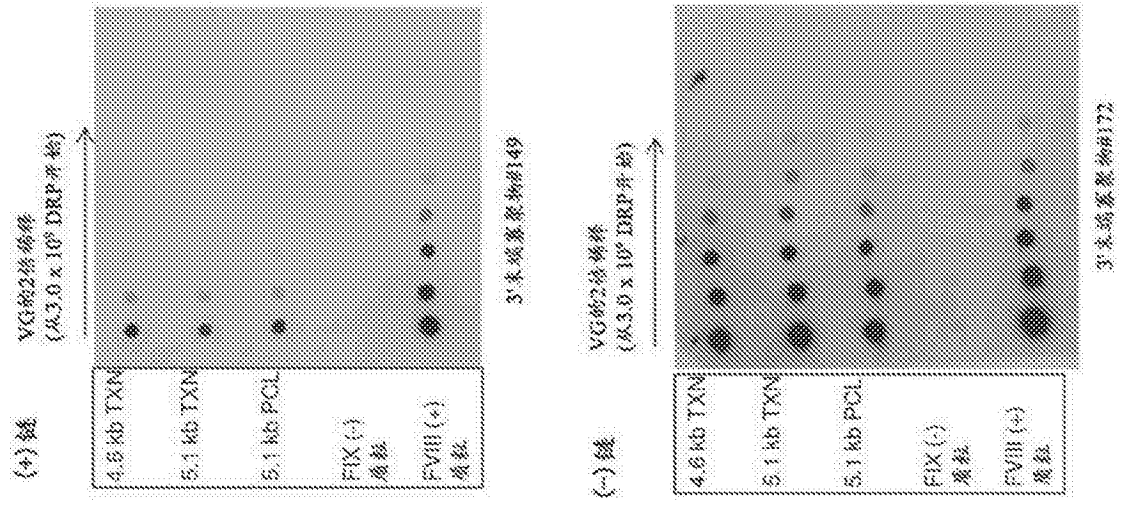


图8B

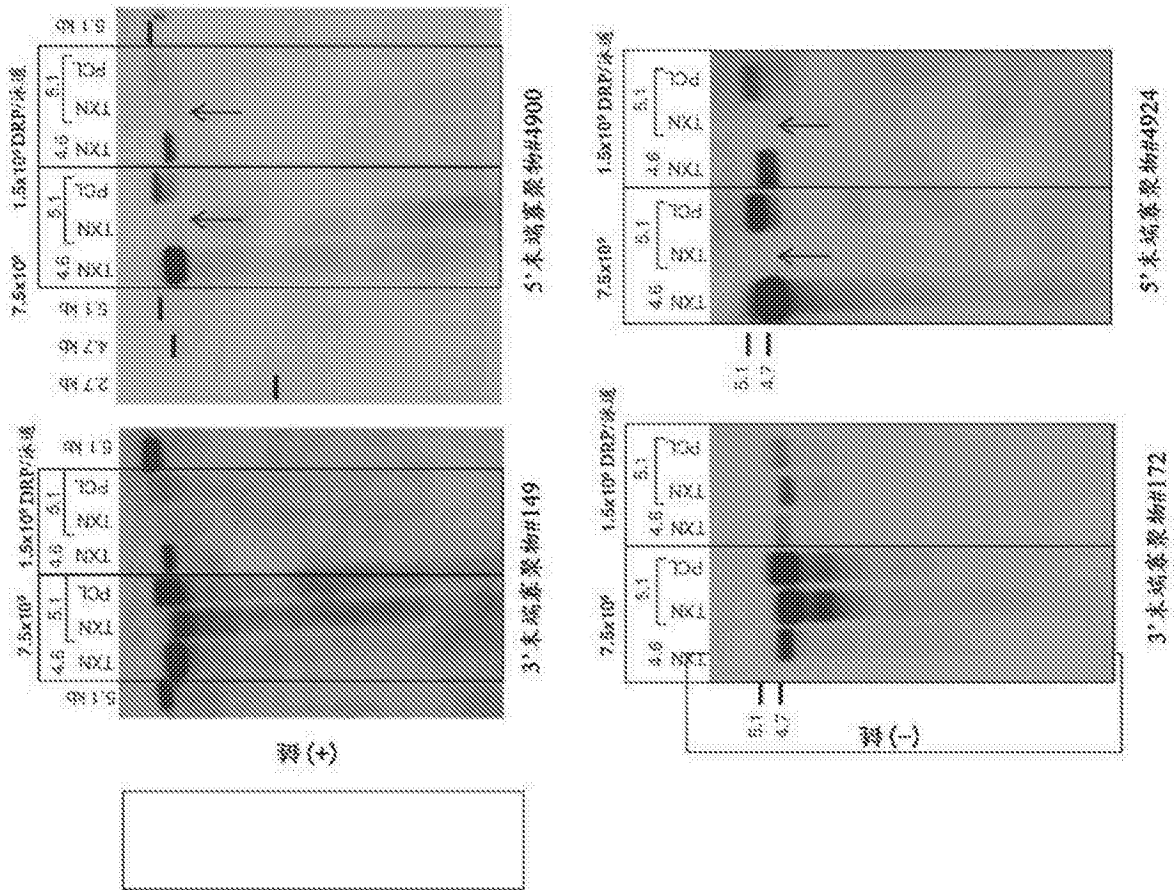


图8C

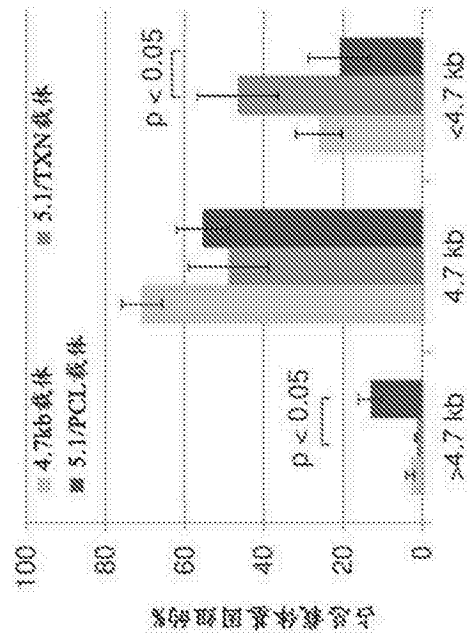


图8D

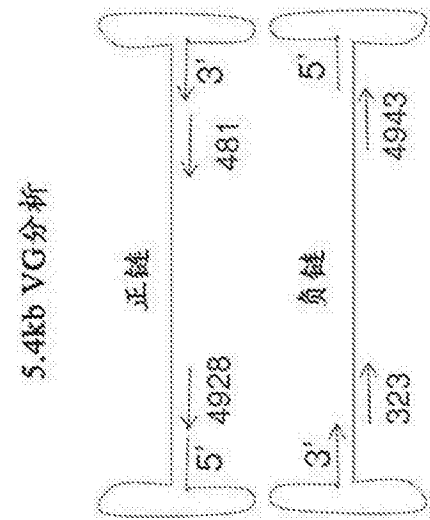


图9A

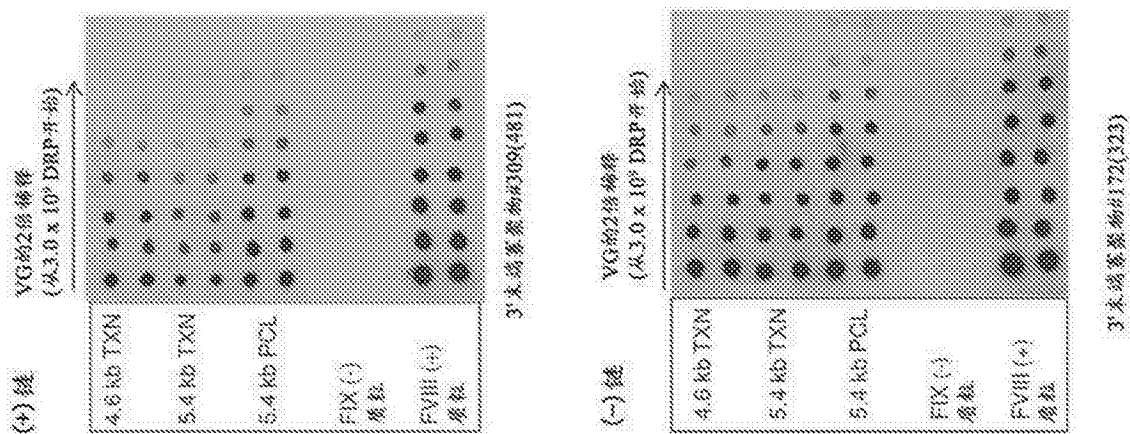


图9B

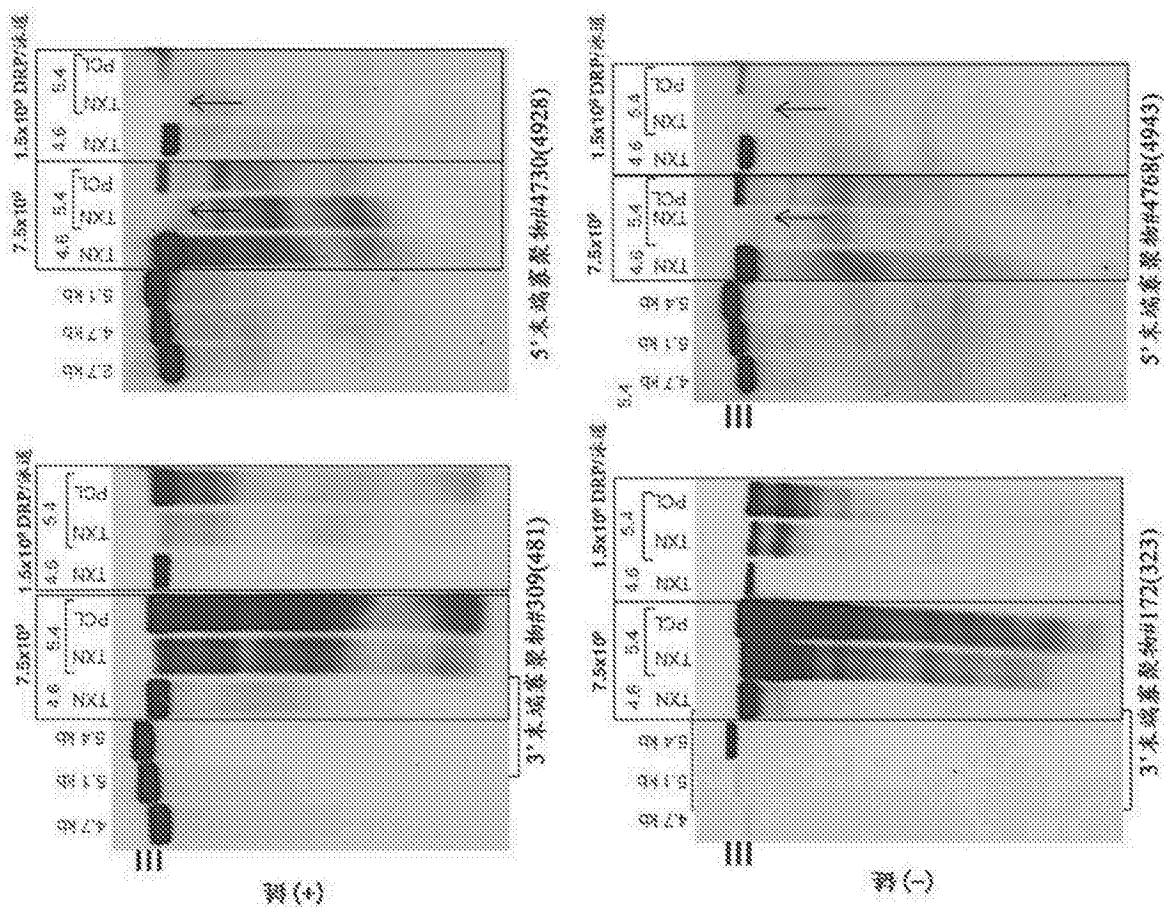


图9C

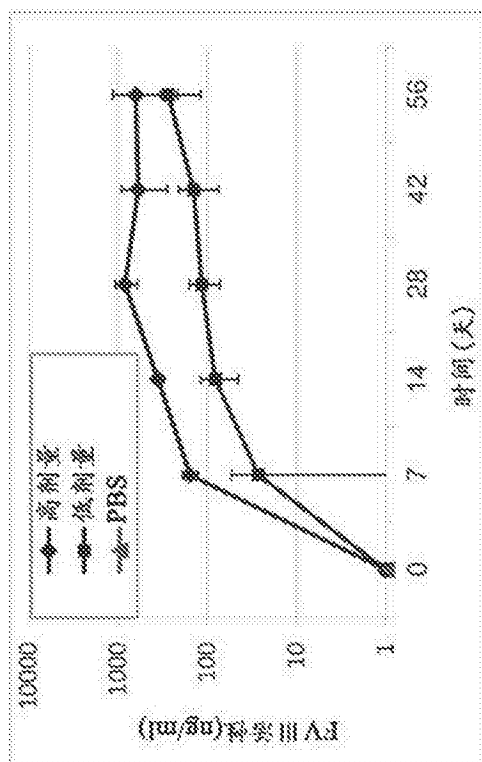


图10A

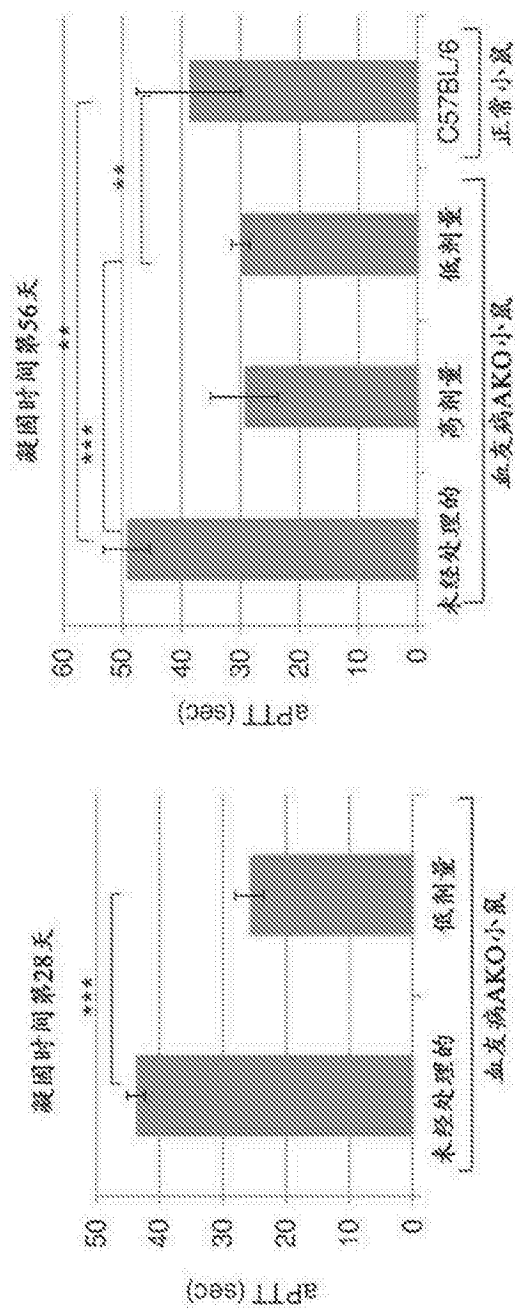


图10B

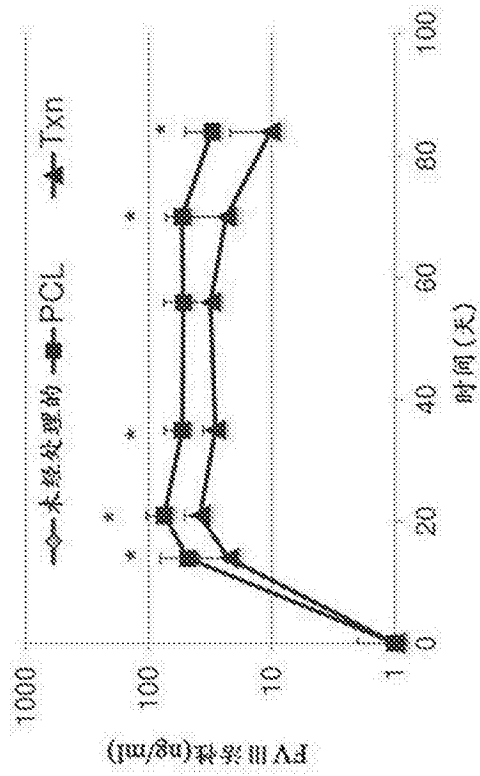


图11A

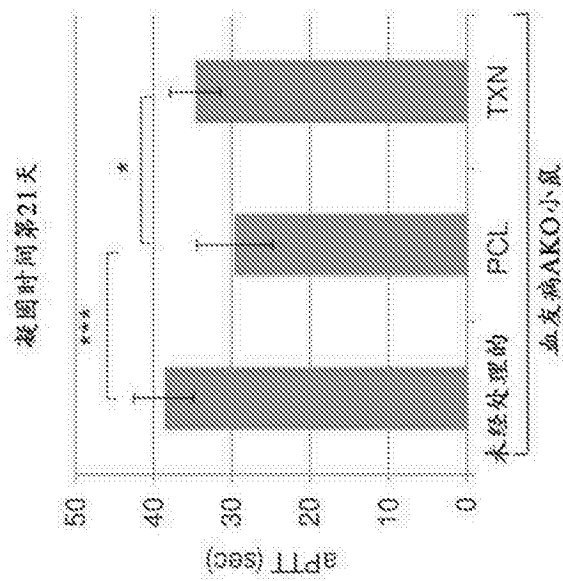


图11B

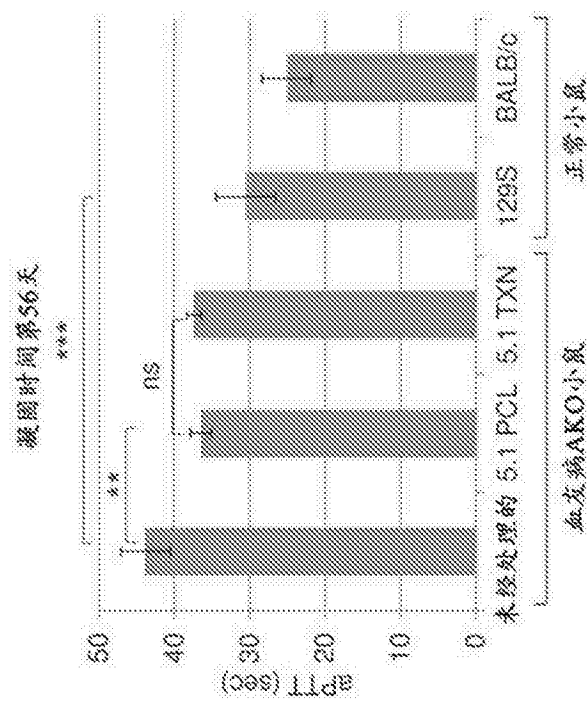


图11C

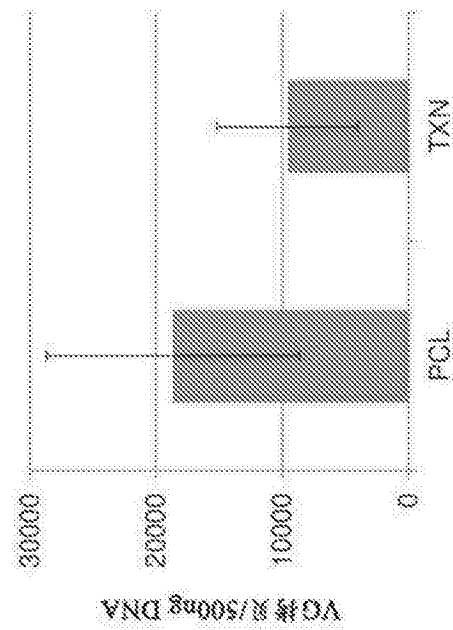


图11D

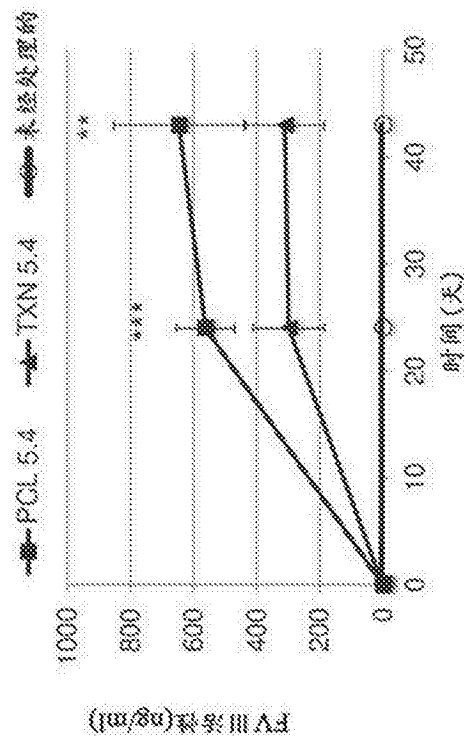


图12A

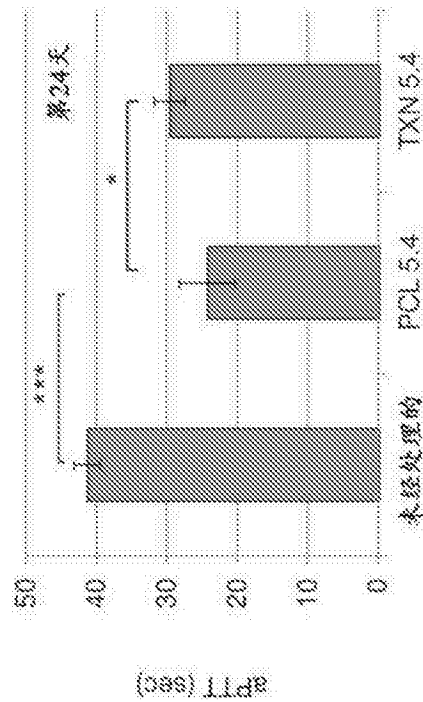


图12B

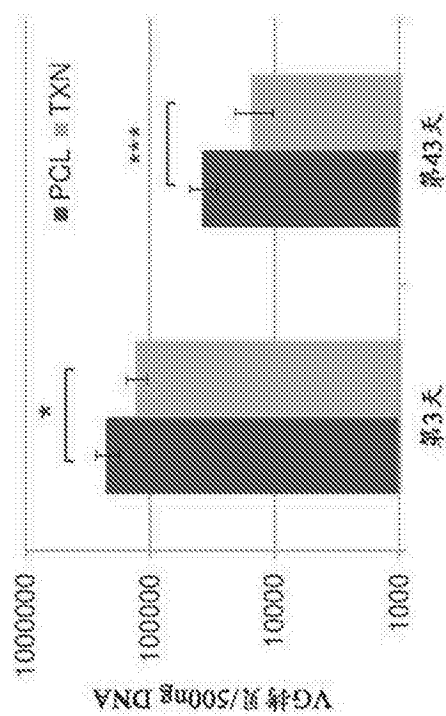


图12C

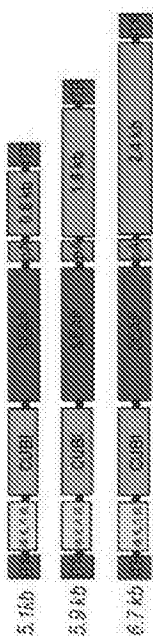


图13A

