



(12) 发明专利申请



(10) 申请公布号 CN 118488784 A

(43) 申请公布日 2024. 08. 13

(21) 申请号 202280073000.7

(22) 申请日 2022.11.04

(30) 优先权数据

63/275,582 2021.11.04 US

(85) PCT国际申请进入国家阶段日

2024.04.30

(86) PCT国际申请的申请数据

PCT/US2022/079336 2022.11.04

(87) PCT国际申请的公布数据

W02023/081847 EN 2023.05.11

(71) 申请人 瑞泽恩制药公司

地址 美国纽约州

(72) 发明人 T·斯蒂特 苏珊娜·布里奇斯

A·O·穆吉卡 洛葛仙妮·艾莉

(74) 专利代理机构 北京市君合律师事务所

11517

专利代理师 张怡 张梦倩

(51) Int.Cl.

A01K 67/0275 (2024.01)

A01K 67/0276 (2024.01)

C07K 14/47 (2006.01)

权利要求书10页 说明书27页

序列表(电子公布) 附图23页

(54) 发明名称

包含经修饰的CACNG1基因座的非人动物

(57) 摘要

提供了包含人源化Cacng1基因座的非人动物细胞和非人动物以及使用此类非人动物细胞和非人动物的方法。包含人源化Cacng1基因座的非人动物细胞或非人动物表达人CACNG1蛋白或其片段。

1. 一种非人动物细胞,其中所述非人动物细胞包含编码异源钙电压门控通道辅助亚基 $\gamma 1$ (CACNG1) 蛋白或其部分的核酸序列。

2. 如权利要求1所述的非人动物细胞,其中编码所述异源CACNG1蛋白或其部分的所述核酸序列包含:

- (i) 包含人CACNG1基因的外显子1或其部分的核酸序列;
- (ii) 包含人CACNG1基因的外显子2或其部分的核酸序列;
- (iii) 包含人CACNG1基因的外显子3或其部分的核酸序列;
- (iv) 包含人CACNG1基因的外显子4或其部分的核酸序列;或
- (v) (i) - (iv) 的任何组合。

3. 如权利要求1或权利要求2所述的非人动物细胞,其中编码所述异源CACNG1蛋白或其部分的所述核酸序列包含:

- (i) 包含人CACNG1基因的外显子1或其部分的核酸序列;
- (ii) 包含人CACNG1基因的内含子1或其部分的核酸序列;
- (iii) 包含人CACNG1基因的外显子2或其部分的核酸序列;
- (iv) 包含人CACNG1基因的内含子2或其部分的核酸序列;
- (v) 包含人CACNG1基因的外显子3或其部分的核酸序列;
- (vi) 包含人CACNG1基因的内含子3或其部分的核酸序列;
- (vii) 包含人CACNG1基因的外显子4或其部分的核酸序列;
- (viii) 包含人CACNG1基因的3' 非翻译区 (UTR) 的核酸序列;或者
- (ix) (i) - (viii) 的任何组合。

4. 如权利要求1-3中任一项所述的非人动物细胞,其中编码所述异源CACNG1蛋白或其部分的所述核酸序列包含选自以下的核酸序列、基本上由选自以下的核酸序列组成或由选自以下的核酸序列组成:如SEQ ID NO:5所示的核酸序列、如SEQ ID NO:27所示的核酸序列和如SEQ ID NO:28所示的核酸序列。

5. 如权利要求1-4中任一项所述的非人动物细胞,其中编码所述异源CACNG1蛋白或其部分的所述核酸序列处于内源Cacng1基因座处。

6. 如权利要求1-5中任一项所述的非人动物细胞,其中编码所述异源CACNG1蛋白或其部分的所述核酸序列替换编码内源CACNG1蛋白或其部分的直系同源内源核酸序列。

7. 如权利要求1-6中任一项所述的非人动物细胞,其中所述非人动物细胞包含内源Cacng1基因座,并且其中所述内源Cacng1基因座包含编码内源CACNG1蛋白或其部分的内源核酸序列被编码所述异源CACNG1蛋白或其部分的所述核酸序列的杂合或纯合替换,

其中编码所述内源CACNG1蛋白或其部分的所述内源核酸序列和编码所述异源CACNG1蛋白或其部分的所述核酸序列是直系同源的。

8. 如权利要求1-8中任一项所述的非人动物细胞,其中所述异源CACNG1蛋白或其部分包含人CACNG1蛋白或其部分的氨基酸序列。

9. 如权利要求1-8中任一项所述的非人动物细胞,其中所述异源CACNG1蛋白或其部分包含

- (i) 如SEQ ID NO:8所示的氨基酸序列;
- (ii) 如SEQ ID NO:10所示的氨基酸序列;

- (iii) 如SEQ ID NO:12所示的氨基酸序列;
- (iv) 如SEQ ID NO:14所示的氨基酸序列;
- (v) 如SEQ ID NO:16所示的氨基酸序列;
- (vi) 如SEQ ID NO:18所示的氨基酸序列;
- (vii) 如SEQ ID NO:20所示的氨基酸序列;
- (viii) 如SEQ ID NO:22所示的氨基酸序列;
- (ix) 如SEQ ID NO:24所示的氨基酸序列;或
- (x) (i) - (ix) 的任何组合。

10. 如权利要求1-3或5-9中任一项所述的非人动物细胞,其中所述异源CACNG1蛋白或其部分包含如SEQ ID NO:4所示的氨基酸序列。

11. 如权利要求1-10中任一项所述的非人动物细胞,其中所述非人动物细胞是哺乳动物细胞。

12. 如权利要求1-11中任一项所述的非人动物细胞,其中所述哺乳动物细胞是啮齿动物细胞。

13. 如权利要求1-12中任一项所述的非人动物细胞,其中所述啮齿动物细胞是大鼠细胞或小鼠细胞。

14. 如权利要求1-13中任一项所述的非人动物细胞,其中所述异源CACNG1蛋白是全长人CACNG1蛋白,并且

其中所述非人动物细胞在其细胞表面上表达全长人CACNG1蛋白。

15. 如权利要求1-14中任一项所述的非人动物细胞,其中所述非人动物细胞是在其细胞表面上表达所述异源CACNG1蛋白或其部分的非人动物骨骼肌细胞。

16. 如权利要求1-13中任一项所述的非人动物细胞,其中所述非人动物细胞是在其细胞表面上不表达所述异源CACNG1蛋白或其部分的非人动物细胞。

17. 如权利要求1-13和16中任一项所述的非人动物细胞,其中所述非人动物细胞在其细胞表面上不表达所述异源CACNG1蛋白或其部分,并且其中所述非人动物细胞不是骨骼细胞。

18. 如权利要求1-13和16-17中任一项所述的非人动物细胞,其中所述非人动物细胞在其细胞表面上不表达所述异源CACNG1蛋白或其部分,并且其中所述非人动物细胞是多能细胞。

19. 如权利要求1-13和16-18中任一项所述的非人动物细胞,其中所述非人动物细胞在其细胞表面上不表达所述异源CACNG1蛋白或其部分,并且其中所述非人动物细胞是胚胎干细胞。

20. 如权利要求1-13和16-18中任一项所述的非人动物细胞,其中所述非人动物细胞在其细胞表面上不表达所述异源CACNG1蛋白或其部分,并且其中所述非人动物细胞是生殖细胞。

21. 如权利要求1-21中任一项所述的非人动物细胞,其中所述非人动物细胞是小鼠细胞。

22. 如权利要求1-21中任一项所述的非人动物细胞,其中所述非人动物细胞是小鼠细胞,并且其中编码异源CACNG1蛋白或其部分的所述核酸序列包含如SEQ ID NO:6所示的核

酸序列、基本上由其组成或由其组成。

23. 一种非人动物, 所述非人动物包含编码异源钙电压门控通道辅助亚基 $\gamma 1$ (CACNG1) 蛋白或其部分的核酸序列。

24. 如权利要求23所述的非人动物, 其中编码所述异源CACNG1蛋白或其部分的所述核酸序列包含:

- (i) 包含人CACNG1基因的外显子1或其部分的核酸序列;
- (ii) 包含人CACNG1基因的外显子2或其部分的核酸序列;
- (iii) 包含人CACNG1基因的外显子3或其部分的核酸序列;
- (iv) 包含人CACNG1基因的外显子4或其部分的核酸序列; 或
- (v) (i) - (iv) 的任何组合。

25. 如权利要求23或权利要求24所述的非人动物, 其中编码所述异源CACNG1蛋白或其部分的所述核酸序列包含:

- (i) 包含人CACNG1基因的外显子1或其部分的核酸序列;
- (ii) 包含人CACNG1基因的内含子1或其部分的核酸序列;
- (iii) 包含人CACNG1基因的外显子2或其部分的核酸序列;
- (iv) 包含人CACNG1基因的内含子2或其部分的核酸序列;
- (v) 包含人CACNG1基因的外显子3或其部分的核酸序列;
- (vi) 包含人CACNG1基因的内含子3或其部分的核酸序列;
- (vii) 包含人CACNG1基因的外显子4或其部分的核酸序列;
- (viii) 包含人CACNG1基因的3' 非翻译区 (UTR) 的核酸序列; 或者
- (ix) (i) - (viii) 的任何组合。

26. 如权利要求23-25中任一项所述的非人动物, 其中编码所述异源CACNG1蛋白或其部分的所述核酸序列包含选自以下的核酸序列、基本上由选自以下的核酸序列组成或由选自以下的核酸序列组成: 如SEQ ID NO:5所示的核酸序列、如SEQ ID NO:27所示的核酸序列和如SEQ ID NO:28所示的核酸序列。

27. 如权利要求23-26中任一项所述的非人动物, 其中编码所述异源CACNG1蛋白或其部分的所述核酸序列处于内源Cacng1基因座处。

28. 如权利要求23-27中任一项所述的非人动物, 其中编码所述异源CACNG1蛋白或其部分的所述核酸序列替换编码内源CACNG1蛋白或其部分的直系同源内源核酸序列。

29. 如权利要求23-28中任一项所述的非人动物, 其中所述非人动物包含内源Cacng1基因座, 并且其中所述内源Cacng1基因座包含编码内源CACNG1蛋白或其部分的内源核酸序列被编码所述异源CACNG1蛋白或其部分的所述核酸序列的杂合或纯合替换,

其中编码所述内源CACNG1蛋白或其部分的所述内源核酸序列和编码所述异源CACNG1蛋白或其部分的所述核酸序列是直系同源的。

30. 如权利要求23-29中任一项所述的非人动物, 其中所述异源CACNG1蛋白或其部分包含人CACNG1蛋白或其部分的氨基酸序列。

31. 如权利要求23-30中任一项所述的非人动物, 其中所述异源CACNG1蛋白或其部分包含

- (i) 如SEQ ID NO:8所示的氨基酸序列;

- (ii) 如SEQ ID NO:10所示的氨基酸序列;
- (iii) 如SEQ ID NO:12所示的氨基酸序列;
- (iv) 如SEQ ID NO:14所示的氨基酸序列;
- (v) 如SEQ ID NO:16所示的氨基酸序列;
- (vi) 如SEQ ID NO:18所示的氨基酸序列;
- (vii) 如SEQ ID NO:20所示的氨基酸序列;
- (viii) 如SEQ ID NO:22所示的氨基酸序列;
- (ix) 如SEQ ID NO:24所示的氨基酸序列;或
- (x) (i) - (ix) 的任何组合。

32. 如权利要求23-31中任一项所述的非人动物,其中所述异源CACNG1蛋白或其部分包含如SEQ ID NO:4所示的氨基酸序列。

33. 如权利要求23-32中任一项所述的非人动物,其中所述非人动物是哺乳动物。

34. 如权利要求23-33中任一项所述的非人动物,其中所述哺乳动物为啮齿动物。

35. 如权利要求23-34中任一项所述的非人动物,其中所述啮齿动物是大鼠或小鼠。

36. 如权利要求23-35中任一项所述的非人动物,其中所述非人动物包含在其细胞表面上表达全长人CACNG1蛋白的非人肌细胞。

37. 如权利要求23-36中任一项所述的非人动物,其中所述非人动物包含在其细胞表面上表达所述异源CACNG1蛋白或其部分的非人动物骨骼肌细胞。

38. 如权利要求23-37中任一项所述的非人动物,其中所述异源CACNG1蛋白包含全长人CACNG1蛋白,并且

其中所述非人动物包含在细胞表面上表达所述全长人CACNG1蛋白的非人动物骨骼肌细胞。

39. 如权利要求23-38中任一项所述的非人动物,其中所述非人动物包含在细胞表面上不表达所述异源CACNG1蛋白或其部分的非人动物细胞。

40. 如权利要求23-35和39中任一项所述的非人动物,其中所述非人动物包含在其骨骼细胞的细胞表面上不表达所述异源CACNG1蛋白或其部分的非人动物生殖细胞。

41. 如权利要求23-40中任一项所述的非人动物,其中所述非人动物是大鼠或小鼠。

42. 一种非人动物基因组,其中所述非人动物基因组包含编码异源钙电压门控通道辅助亚基 $\gamma 1$ (CACNG1) 蛋白或其部分的核酸序列。

43. 如权利要求42所述的非人动物基因组,其中编码所述异源CACNG1蛋白或其部分的所述核酸序列包含:

- (i) 包含人CACNG1基因的外显子1或其部分的核酸序列;
- (ii) 包含人CACNG1基因的外显子2或其部分的核酸序列;
- (iii) 包含人CACNG1基因的外显子3或其部分的核酸序列;
- (iv) 包含人CACNG1基因的外显子4或其部分的核酸序列;或
- (v) (i) - (iv) 的任何组合。

44. 如权利要求42或权利要求43所述的非人动物基因组,其中编码所述异源CACNG1蛋白或其部分的所述核酸序列包含:

- (i) 包含人CACNG1基因的外显子1或其部分的核酸序列;

- (ii) 包含人CACNG1基因的内含子1或其部分的核酸序列;
- (iii) 包含人CACNG1基因的外显子2或其部分的核酸序列;
- (iv) 包含人CACNG1基因的内含子2或其部分的核酸序列;
- (v) 包含人CACNG1基因的外显子3或其部分的核酸序列;
- (vi) 包含人CACNG1基因的内含子3或其部分的核酸序列;
- (vii) 包含人CACNG1基因的外显子4或其部分的核酸序列;
- (viii) 包含人CACNG1基因的3' 非翻译区 (UTR) 的核酸序列;或者
- (ix) (i) - (viii) 的任何组合。

45. 如权利要求42-44中任一项所述的非人动物基因组,其中编码所述异源CACNG1蛋白或其部分的所述核酸序列包含选自以下的核酸序列、基本上由选自以下的核酸序列组成或由选自以下的核酸序列组成:如SEQ ID NO:5所示的核酸序列、如SEQ ID NO:27所示的核酸序列和如SEQ ID NO:28所示的核酸序列。

46. 如权利要求42-45中任一项所述的非人动物基因组,其中编码所述异源CACNG1蛋白或其部分的所述核酸序列处于内源Cacng1基因座处。

47. 如权利要求42-46中任一项所述的非人动物基因组,其中编码所述异源CACNG1蛋白或其部分的所述核酸序列替换编码内源CACNG1蛋白或其部分的直系同源内源核酸序列。

48. 如权利要求42-47中任一项所述的非人动物基因组,其中所述非人动物基因组包含内源Cacng1基因座,并且其中所述内源Cacng1基因座包含编码内源CACNG1蛋白或其部分的内源核酸序列被编码所述异源CACNG1蛋白或其部分的所述核酸序列的杂合或纯合替换,

其中编码所述内源CACNG1蛋白或其部分的所述内源核酸序列和编码所述异源CACNG1蛋白或其部分的所述核酸序列是直系同源的。

49. 如权利要求42-48中任一项所述的非人动物基因组,其中所述异源CACNG1蛋白或其部分包含人CACNG1蛋白或其部分的氨基酸序列。

50. 如权利要求42-49中任一项所述的非人动物基因组,其中所述异源CACNG1蛋白或其部分包含

- (i) 如SEQ ID NO:8所示的氨基酸序列;
- (ii) 如SEQ ID NO:10所示的氨基酸序列;
- (iii) 如SEQ ID NO:12所示的氨基酸序列;
- (iv) 如SEQ ID NO:14所示的氨基酸序列;
- (v) 如SEQ ID NO:16所示的氨基酸序列;
- (vi) 如SEQ ID NO:18所示的氨基酸序列;
- (vii) 如SEQ ID NO:20所示的氨基酸序列;
- (viii) 如SEQ ID NO:22所示的氨基酸序列;
- (ix) 如SEQ ID NO:24所示的氨基酸序列;或
- (x) (i) - (ix) 的任何组合。

51. 如权利要求42-50中任一项所述的非人动物基因组,其中所述异源CACNG1蛋白或其部分包含如SEQ ID NO:4所示的氨基酸序列。

52. 如权利要求42-51中任一项所述的非人动物基因组,其中所述非人基因组是哺乳动物核酸。

53. 如权利要求42-52中任一项所述的非人动物基因组,其中所述哺乳动物是啮齿动物核酸。

54. 如权利要求42-53中任一项所述的非人动物基因组,其中所述啮齿动物是大鼠基因组或小鼠核酸。

55. 如权利要求42-54中任一项所述的非人动物基因组,其中所述非人动物基因组是小鼠核酸。

56. 一种嵌合核酸分子,所述嵌合核酸分子包含非人动物Cacng1基因的核酸序列,所述核酸序列(a) 编码CACNG1蛋白,并且(b) 经修饰为包含编码所述CACNG1蛋白或其部分的序列被编码异源CACNG1蛋白或其部分的同源序列的替换,

其中所述嵌合核酸分子编码功能性CACNG1蛋白。

57. 如权利要求56所述的嵌合核酸分子,其中所述嵌合核酸序列还包含所述非人动物Cacng1基因的启动子和/或调节序列。

58. 如权利要求57所述的嵌合核酸分子,其中所述同源核酸序列包含:

- (i) 包含人CACNG1基因的外显子1或其部分的核酸序列;
- (ii) 人CACNG1基因的内含子1或其部分的核酸序列;
- (iii) 包含人CACNG1基因的外显子2或其部分的核酸序列;
- (iv) 人CACNG1基因的内含子2或其部分的核酸序列;
- (v) 包含人CACNG1基因的外显子3或其部分的核酸序列;
- (vi) 人CACNG1基因的内含子3或其部分的核酸序列;
- (vii) 包含人CACNG1基因的外显子4或其部分的核酸序列;
- (viii) 人CACNG1基因的3' 非翻译区(UTR)的核酸序列;或者
- (ix) (i) - (viii)的任何组合。

59. 如权利要求57或权利要求58所述的嵌合核酸分子,其中所述经修饰的非人动物Cacng1基因还包含药物选择盒。

60. 如权利要求57-59中任一项所述的嵌合核酸分子,所述嵌合核酸分子还包含:

- (i) 所述经修饰的非人动物Cacng1基因上游的5' 同源臂;和
- (ii) 所述经修饰的非人动物Cacng1基因下游的3' 同源臂。

61. 如权利要求60所述的嵌合核酸,其中所述5' 同源臂和3' 同源臂与感兴趣的非人动物Cacng1基因座进行同源重组,并且

其中在与所述感兴趣的非人动物Cacng1基因座同源重组后,所述经修饰的非人动物Cacng1基因替换所述感兴趣的非人动物Cacng1基因座处的所述非人动物Cacng1基因,并与驱动所述感兴趣的非人动物Cacng1基因座处的所述非人动物Cacng1基因表达的内源启动子可操作地连接。

62. 如权利要求60或权利要求61所述的嵌合核酸,其中:

- (i) 所述5' 同源臂包含如SEQ ID NO:25所示的核酸序列;或者
- (ii) 所述3' 同源臂包含如SEQ ID NO:26所示的核酸序列。

63. 如权利要求57-62中任一项所述的嵌合核酸分子,其中所述嵌合核酸的核酸序列包含如SEQ ID NO:6所示的核酸序列。

64. 一种制备权利要求1-22中任一项所述的非人动物细胞的方法,所述方法包括将编

码异源CACNG1蛋白或其部分的核酸序列插入所述非人动物细胞的基因组中。

65. 如权利要求64所述的方法,其中所述非人动物细胞是非人动物胚胎干(ES)细胞,并且

其中所述插入包括将编码所述异源CACNG1蛋白或其部分的所述核酸序列插入所述非人动物ES细胞的基因组中,以形成经修饰的非人动物ES细胞,在所述非人动物ES细胞的基因组中包含编码所述异源CACNG1蛋白或其部分的所述核酸序列。

66. 如权利要求64或权利要求65所述的方法,其中将编码所述异源CACNG1蛋白或其部分的所述核酸序列插入内源Cacng1基因座中。

67. 如权利要求64-66中任一项所述的方法,其中所述插入步骤包括用编码所述异源CACNG1蛋白或其部分的所述核酸序列替换编码内源CACNG1蛋白或其部分的内源核酸序列,

其中编码内源CACNG1蛋白或其部分的所述内源核酸序列和编码所述异源CACNG1蛋白或其部分的所述核酸序列是直系同源的。

68. 如权利要求64-67中任一项所述的方法,其中编码异源CACNG1蛋白或其部分的所述核酸序列包含:

- (i) 包含人CACNG1基因的外显子1或其部分的核酸序列;
- (ii) 包含人CACNG1基因的外显子2或其部分的核酸序列;
- (iii) 包含人CACNG1基因的外显子3或其部分的核酸序列;
- (iv) 包含人CACNG1基因的外显子4或其部分的核酸序列;或
- (v) (i) - (iv)的任何组合。

69. 如权利要求64-68中任一项所述的方法,其中编码异源CACNG1蛋白或其部分的所述核酸序列包含:

- (i) 包含人CACNG1基因的外显子1或其部分的核酸序列;
- (ii) 人CACNG1基因的内含子1或其部分的核酸序列;
- (iii) 包含人CACNG1基因的外显子2或其部分的核酸序列;
- (iv) 人CACNG1基因的内含子2或其部分的核酸序列;
- (v) 包含人CACNG1基因的外显子3或其部分的核酸序列;
- (vi) 人CACNG1基因的内含子3或其部分的核酸序列;
- (vii) 包含人CACNG1基因的外显子4或其部分的核酸序列;
- (viii) 人CACNG1基因的3' 非翻译区(UTR)的核酸序列;或者
- (ix) (i) - (viii)的任何组合。

70. 如权利要求64-69中任一项所述的方法,其中编码异源CACNG1蛋白或其部分的所述核酸序列包含选自以下的核酸序列、基本上由选自以下的核酸序列组成或由选自以下的核酸序列组成:如SEQ ID NO:5所示的核酸序列、如SEQ ID NO:27所示的核酸序列和如SEQ ID NO:28所示的核酸序列。

71. 如权利要求64-70中任一项所述的方法,其中所述异源CACNG1蛋白或其部分包含人CACNG1蛋白或其部分的氨基酸序列。

72. 如权利要求64-71中任一项所述的方法,其中所述异源CACNG1蛋白或其部分包含

- (i) 如SEQ ID NO:8所示的氨基酸序列;
- (ii) 如SEQ ID NO:10所示的氨基酸序列;

- (iii) 如SEQ ID NO:12所示的氨基酸序列;
- (iv) 如SEQ ID NO:14所示的氨基酸序列;
- (v) 如SEQ ID NO:16所示的氨基酸序列;
- (vi) 如SEQ ID NO:18所示的氨基酸序列;
- (vii) 如SEQ ID NO:20所示的氨基酸序列;
- (viii) 如SEQ ID NO:22所示的氨基酸序列;
- (ix) 如SEQ ID NO:24所示的氨基酸序列;或
- (x) (i) - (ix) 的任何组合。

73. 如权利要求64-72中任一项所述的方法,其中所述异源CACNG1蛋白包含如SEQ ID NO:4所示的氨基酸序列。

74. 如权利要求64-73中任一项所述的方法,其中所述非人动物细胞是哺乳动物细胞。

75. 如权利要求64-74中任一项所述的方法,其中所述哺乳动物细胞是啮齿动物细胞。

76. 如权利要求64-75中任一项所述的方法,所述啮齿动物细胞是小鼠细胞或大鼠细胞。

77. 如权利要求64-76中任一项所述的方法,其中所述非人动物细胞是小鼠细胞,并且其中编码异源CACNG1蛋白或其部分的所述核酸序列包含如SEQ ID NO:6所示的核酸序列、基本上由其组成或由其组成。

78. 如权利要求64-77中任一项所述的方法,其中插入包括使所述非人动物细胞的基因组与权利要求55-63中任一项所述的嵌合核酸分子接触。

79. 如权利要求64-78中任一项所述的方法,其中插入包括使所述非人动物细胞的基因组与权利要求56-63中任一项所述的嵌合核酸分子接触。

80. 一种制备权利要求23-41中任一项所述的非人动物的方法,所述方法包括将编码异源CACNG1蛋白或其部分的核酸序列插入非人动物胚胎干(ES)细胞的基因组中以形成经修饰的非人动物ES细胞,在所述非人动物ES细胞的基因组中包含编码所述异源CACNG1蛋白或其部分的所述核酸序列;

将所述非人动物ES细胞引入体外宿主胚胎细胞中;以及

在非人代孕母体动物中孕育包含所述经修饰的非人动物ES细胞的所述宿主胚胎细胞,并且其中在所述孕育后,所述非人代孕母体动物分娩包含生殖细胞的非人动物后代,所述生殖细胞包含编码所述异源CACNG1蛋白或其部分的所述核酸序列。

81. 如权利要求80所述的方法,其中将编码所述异源CACNG1蛋白或其部分的所述核酸序列插入内源Cacng1基因座中。

82. 如权利要求80或权利要求81所述的方法,其中所述插入步骤包括用编码所述异源CACNG1蛋白或其部分的所述核酸序列替换编码内源CACNG1蛋白或其部分的内源核酸序列,

其中编码内源CACNG1蛋白或其部分的所述内源核酸序列和编码所述异源CACNG1蛋白或其部分的所述核酸序列是直系同源的。

83. 如权利要求80-82中任一项所述的方法,其中编码异源CACNG1蛋白或其部分的所述核酸序列包含:

- (i) 包含人CACNG1基因的外显子1或其部分的核酸序列;
- (ii) 包含人CACNG1基因的外显子2或其部分的核酸序列;

- (iii) 包含人CACNG1基因的外显子3或其部分的核酸序列;
- (iv) 包含人CACNG1基因的外显子4或其部分的核酸序列;或
- (v) (i) - (iv)的任何组合。

84. 如权利要求80-83中任一项所述的方法,其中编码异源CACNG1蛋白或其部分的所述核酸序列包含:

- (i) 包含人CACNG1基因的外显子1或其部分的核酸序列;
- (ii) 人CACNG1基因的内含子1或其部分的核酸序列;
- (iii) 包含人CACNG1基因的外显子2或其部分的核酸序列;
- (iv) 人CACNG1基因的内含子2或其部分的核酸序列;
- (v) 包含人CACNG1基因的外显子3或其部分的核酸序列;
- (vi) 人CACNG1基因的内含子3或其部分的核酸序列;
- (vii) 包含人CACNG1基因的外显子4或其部分的核酸序列;
- (viii) 人CACNG1基因的3' 非翻译区 (UTR) 的核酸序列;或者
- (ix) (i) - (viii)的任何组合。

85. 如权利要求80-84中任一项所述的方法,其中编码异源CACNG1蛋白或其部分的所述核酸序列包含选自以下的核酸序列、基本上由选自以下的核酸序列组成或由选自以下的核酸序列组成:如SEQ ID NO:5所示的核酸序列、如SEQ ID NO:27所示的核酸序列和如SEQ ID NO:28所示的核酸序列。

86. 如权利要求80-85中任一项所述的方法,其中所述异源CACNG1蛋白或其部分包含人CACNG1蛋白或其部分的氨基酸序列。

87. 如权利要求80-86中任一项所述的方法,其中所述异源CACNG1蛋白或其部分包含

- (i) 如SEQ ID NO:8所示的氨基酸序列;
- (ii) 如SEQ ID NO:10所示的氨基酸序列;
- (iii) 如SEQ ID NO:12所示的氨基酸序列;
- (iv) 如SEQ ID NO:14所示的氨基酸序列;
- (v) 如SEQ ID NO:16所示的氨基酸序列;
- (vi) 如SEQ ID NO:18所示的氨基酸序列;
- (vii) 如SEQ ID NO:20所示的氨基酸序列;
- (viii) 如SEQ ID NO:22所示的氨基酸序列;
- (ix) 如SEQ ID NO:24所示的氨基酸序列;或
- (x) (i) - (ix)的任何组合。

88. 如权利要求80-87中任一项所述的方法,其中所述异源CACNG1蛋白包含如SEQ ID NO:4所示的氨基酸序列。

89. 如权利要求80-88中任一项所述的方法,其中所述非人动物是哺乳动物。

90. 如权利要求80-89中任一项所述的方法,其中所述哺乳动物是啮齿动物。

91. 如权利要求80-90中任一项所述的方法,所述啮齿动物细胞是小鼠细胞或大鼠细胞。

92. 如权利要求80-91中任一项所述的方法,其中所述非人动物是小鼠,并且其中编码异源CACNG1蛋白或其部分的所述核酸序列包含如SEQ ID NO:6所示的核酸序列、基本上由

其组成或由其组成。

93. 如权利要求23-40中任一项所述的非人动物,或根据权利要求80-92中任一项所述的方法制备的非人动物,所述非人动物包含结合所述异源CACNG1蛋白的抗原结合蛋白,其中所述非人动物在骨骼肌细胞的表面表达所述异源CACNG1蛋白或其胞外结构域。

94. 如权利要求93所述的非人动物,其中所述异源CACNG1蛋白是人CACNG1蛋白。

95. 如权利要求93或权利要求94所述的非人动物,其中所述非人动物是小鼠。

96. 一种非人动物、非人动物细胞或非人动物基因组,所述非人动物、非人动物细胞或非人动物基因组包含内源Cacng1基因的敲除突变。

97. 如权利要求96所述的非人动物、非人动物细胞或非人动物基因组,其中所述敲除突变包含所述Cacng1基因或其部分的缺失。

98. 如权利要求96或权利要求97所述的非人动物、非人动物细胞或非人动物基因组,其中所述敲除突变包含所述Cacng1基因的整个编码序列的缺失。

99. 如权利要求96-98中任一项所述的非人动物、非人动物细胞或非人动物基因组,其中所述非人动物、非人动物细胞或非人动物基因组不表达任何CACNG1蛋白。

100. 如权利要求96-99中任一项所述的非人动物、非人动物细胞或非人动物基因组,其中所述非人动物、非人动物细胞或非人动物基因组不表现出任何重大突变表型。

101. 如权利要求96-100中任一项所述的非人动物、非人动物细胞或非人动物基因组,其中所述非人动物、非人动物细胞或非人动物基因组包含含有如SEQ ID NO:29所示的序列的内源Cacng1基因座。

102. 一种制备CACNG1敲除的非人动物的方法,所述方法包括将所述非人动物的内源Cacng1基因座修饰为包含敲除突变。

103. 一种靶向载体,所述靶向载体包含:

(i) 5' 同源臂;以及

(ii) 3' 同源臂,

其中所述5' 同源臂和3' 同源臂与感兴趣的非人动物Cacng1基因座进行同源重组,并且其中在与所述感兴趣的非人动物Cacng1基因座同源重组后,所述靶向载体在所述感兴趣的非人动物Cacng1基因座处的所述非人动物Cacng1基因中插入敲除突变。

104. 一种根据本文描述的任何方法制备的非人动物、非人动物细胞或非人动物基因组。

包含经修饰的CACNG1基因座的非人动物

[0001] 相关申请的交叉引用

[0002] 本申请要求2021年11月4日提交的美国临时申请63/275,582号的优先权权益,该临时申请以引用的方式整体并入。

[0003] 序列表

[0004] 书写于文件11102W001_ST26.txt中的序列表是79千字节,创建于2022年11月4日,并且特此以引用的方式整体并入。

技术领域

[0005] 描述了经遗传修饰的非人动物(例如啮齿动物,例如小鼠或大鼠),其基因组中包含编码人钙电压门控通道辅助亚基 $\gamma 1$ (CACNG1) 蛋白或其部分的核酸。因此,还描述了例如在骨骼肌细胞的表面表达人CACNG1蛋白或其部分的经遗传修饰的非人动物。例如在骨骼肌细胞的表面表达人CACNG1蛋白或其部分的此类经遗传修饰的非人动物可用作基于CACNG1的治疗剂例如基于CACNG1的抗体的临床前测试模型。

背景技术

[0006] 骨骼肌是人体三大重要肌肉组织之一。每块骨骼肌含有数千条由结缔组织鞘包裹在一起的肌纤维。骨骼肌使人类能够移动和进行日常活动。骨骼肌在呼吸力学中起着至关重要的作用,并有助于保持姿势和平衡。骨骼肌还保护身体的重要器官。

[0007] 由于骨骼肌功能异常,可导致无数医学病状,一种能够测试针对治疗异常肌肉功能的疗法的合适动物模型(例如通过靶向肌肉特异性表面蛋白)可有助于研究与骨骼肌相关的医学病状。

发明内容

[0008] 本文提供了具有编码人钙电压门控通道辅助亚基 $\gamma 1$ (CACNG1) 蛋白的重组遗传基因座的经遗传修饰的非人动物。本文还提供了用于产生和使用此类经修饰的非人动物的组合和方法。

[0009] 本文描述了包含异源(例如,人)Cacng1基因或其部分的基因工程化的非人动物基因组、工程化细胞和非人动物。在一些实施方案中,本文所述的基因工程化动物从所需基因座(例如,从内源Cacng1区段)表达异源(例如,人)CACNG1蛋白。该非人动物可以是哺乳动物,诸如啮齿动物(例如,小鼠或大鼠)。该非人动物细胞可以是哺乳动物细胞,诸如啮齿动物细胞(例如,小鼠细胞或大鼠细胞)。该非人动物基因组可以是哺乳动物核酸,诸如啮齿动物核酸(例如,小鼠核酸或大鼠核酸)。

[0010] 在一些实施方案中,非人动物、非人动物细胞或非人动物基因组包含编码异源(例如,人)CACNG1蛋白或其部分的核酸序列。

[0011] 在一些实施方案中,编码异源(例如,人)CACNG1蛋白或其部分的核酸序列包含:
(i) 包含人CACNG1基因的外显子1或其部分的核酸序列; (ii) 包含人CACNG1基因的外显子2

或其部分的核酸序列; (iii) 包含人CACNG1基因的外显子3或其部分的核酸序列; (iv) 包含人CACNG1基因的外显子4或其部分的核酸序列; 或 (v) (i) - (iv) 的任何组合。在一些实施方案中, 编码异源 (例如, 人) CACNG1蛋白或其部分的核酸序列包含: (i) 包含人CACNG1基因的外显子1或其部分的核酸序列; (ii) 人CACNG1基因的内含子1或其部分的核酸序列; (iii) 包含人CACNG1基因的外显子2或其部分的核酸序列; (iv) 人CACNG1基因的内含子2或其部分的核酸序列; (v) 包含人CACNG1基因的外显子3或其部分的核酸序列; (vi) 人CACNG1基因的内含子3或其部分的核酸序列; (v) 包含人CACNG1基因的外显子4或其部分的核酸序列; (vii) 人CACNG1基因的3' 非翻译区 (UTR) 的核酸序列; 或 (v) (i) - (iv) 的任何组合。在一些实施方案中, 编码异源CACNG1蛋白或其部分的核酸序列包含选自以下的核酸序列、基本上由选自以下的核酸序列组成或由选自以下的核酸序列组成: 如SEQ ID NO:5所示的核酸序列、如SEQ ID NO:27所示的核酸序列和如SEQ ID NO:28所示的核酸序列。

[0012] 在一些实施方案中, 编码异源 (例如, 人) CACNG1蛋白或其部分的核酸序列掺入 (基因组、细胞或非人类动物的) 内源Cacng1基因座中。在一些实施方案中, 编码异源CACNG1蛋白或其部分的核酸序列 (在非人动物基因组、非人动物细胞或非人动物的内源基因座处) 替换编码内源CACNG1蛋白或其部分的直系同源内源核酸序列。在一些实施方案中, 内源Cacng1基因座包含编码内源CACNG1蛋白或其部分的内源核酸序列被编码异源CACNG1蛋白或其部分的核酸序列的杂合或纯合替换, 其中编码内源CACNG1蛋白或其部分的内源核酸序列和编码异源CACNG1蛋白或其部分的核酸序列是直系同源的。

[0013] 在一些实施方案中, 异源CACNG1蛋白或其部分包含人CACNG1蛋白或其部分的氨基酸序列。在一些实施方案中, 异源CACNG1蛋白或其部分包含 (i) 如SEQ ID NO:8所示的氨基酸序列; (ii) 如SEQ ID NO:10所示的氨基酸序列; (iii) 如SEQ ID NO:12所示的氨基酸序列; (iv) 如SEQ ID NO:14所示的氨基酸序列; (v) 如SEQ ID NO:16所示的氨基酸序列; (vi) 如SEQ ID NO:18所示的氨基酸序列; (vii) 如SEQ ID NO:20所示的氨基酸序列; (viii) 如SEQ ID NO:22所示的氨基酸序列; (ix) 如SEQ ID NO:24所示的氨基酸序列; 或 (x) (i) - (ii) 的任何组合。在一些实施方案中, 异源CACNG1蛋白包含如SEQ ID NO:4所示的氨基酸序列。

[0014] 在一些实施方案中, 如本文所述的非人动物细胞在其细胞表面上表达异源CACNG1蛋白或其部分, 所述蛋白可以是全长人CACNG1蛋白。在一些实施方案中, 如本文所述的非人动物细胞是在其细胞表面上表达异源 (例如, 人) CACNG1蛋白或其部分的非人动物骨骼肌细胞。

[0015] 在一些实施方案中, 如本文所述的非人动物细胞是在其细胞表面上不表达异源 (例如, 人) CACNG1蛋白或其部分的非人动物细胞, 例如其中非人动物细胞不是骨骼细胞和/或例如其中非人动物细胞是多能细胞、胚胎干细胞、生殖细胞等。

[0016] 在一些实施方案中, 非人动物细胞是小鼠细胞并且编码异源CACNG1蛋白或其部分的核酸序列包含如SEQ ID NO:6所示的核酸序列、基本上由其组成或由其组成。

[0017] 在一些实施方案中, 如本文所述的非人动物包含在其细胞表面上表达异源CACNG1蛋白或其部分的骨骼肌细胞。在一些实施方案中, 非人动物包含在其细胞表面上表达异源CACNG1蛋白或其部分 (例如, 全长人CACNG1蛋白) 的非人骨骼肌细胞。在一些实施方案中, 非人动物包含含有异源 (例如, 人) Cacng1基因或其部分并且在其细胞表面上不表达由异源 (例如, 人) Cacng1基因或其部分编码的异源 (例如, 人) CACNG1蛋白或其部分的非人动物细

胞,例如其中非人动物细胞不是骨骼细胞和/或其中非人动物细胞是多能细胞、胚胎干细胞、生殖细胞等。

[0018] 在一些实施方案中,本文所述的非人动物是小鼠,本文所述的非人动物细胞是小鼠细胞,或者本文所述的非人动物基因组是小鼠核酸,并且编码异源CACNG1蛋白或其部分的核酸序列包含如SEQ ID NO:6所示的核酸序列、基本上由其组成或由其组成。

[0019] 本文还描述了编码功能性CACNG1蛋白的嵌合核酸分子,该嵌合核酸分子包含编码非人CACNG1蛋白或其部分的经修饰的非人动物Cacng1基因的核酸序列,其中经修饰的非人动物Cacng1基因包含编码非人动物CACNG1蛋白的一部分的核酸序列被编码异源CACNG1蛋白或其部分的同源核酸序列的替换。在一些实施方案中,如本文所述的嵌合核酸分子包含非人动物Cacng1基因的核酸序列,该核酸序列(a) 编码CACNG1蛋白,并且(b) 经修饰为包含编码CACNG1蛋白或其部分的序列被编码异源CACNG1蛋白或其部分的同源序列的替换,其中该嵌合核酸分子编码功能性CACNG1蛋白,并且任选地,其中该嵌合核酸序列还包含非人动物Cacng1基因的启动子和/或调节序列。在一些实施方案中,同源核酸序列包含:(i) 包含人CACNG1基因的外显子1或其部分的核酸序列;(ii) 人CACNG1基因的内含子1或其部分的核酸序列;(iii) 包含人CACNG1基因的外显子2或其部分的核酸序列;(iv) 人CACNG1基因的内含子2或其部分的核酸序列;(v) 包含人CACNG1基因的外显子3或其部分的核酸序列;(vi) 人CACNG1基因的内含子3或其部分的核酸序列;(v) 包含人CACNG1基因的外显子4或其部分的核酸序列;(vii) 人CACNG1基因的3' 非翻译区(UTR) 的核酸序列;或(v) (i) - (iv) 的任何组合。在一些实施方案中,经修饰的Cacng1基因还包含药物选择盒。在一些实施方案中,本文所述的嵌合核酸分子还包含(i) 经修饰的非人动物Cacng1基因上游的5' 同源臂和(ii) 经修饰的非人动物Cacng1基因下游的3' 同源臂。在一些情况下,5' 同源臂和3' 同源臂可以与感兴趣的非人动物Cacng1基因座进行同源重组,并且其中在与感兴趣的非人动物Cacng1基因座同源重组后,经修饰的Cacng1基因可以替换感兴趣的非人动物Cacng1基因座处的非人动物Cacng1基因,并与驱动感兴趣的非人动物Cacng1基因座处的非人动物Cacng1基因表达的内源启动子可操作地连接。在具体的实施方案中,本文所述的嵌合核酸具有(i) 包含如SEQ ID NO:25所示的核酸序列的5' 同源臂和/或(ii) 包含如SEQ ID NO:26所示的核酸序列的3' 同源臂。在一些实施方案中,该核酸序列包含如SEQ ID NO:6所示的核酸序列。

[0020] 还描述了通过将编码异源CACNG1蛋白或其部分的核酸序列插入非人动物的基因组、非人动物细胞的基因组或非人动物基因组中来制备本文所述的非人动物、非人动物细胞或非人动物基因组的方法。在一些实施方案中,非人动物细胞是非人动物胚胎干(ES) 细胞,并且其中所述插入包括将编码异源CACNG1蛋白或其部分的核酸序列插入非人动物ES细胞的基因组中,以形成经修饰的非人动物ES细胞,在其基因组中包含编码异源CACNG1蛋白或其部分的核酸序列。在一些实施方案中,该方法包括将经修饰的非人动物ES细胞引入体外宿主胚胎细胞中。在一些实施方案中,该方法包括在合适的非人代孕母体动物中孕育包含经修饰的非人动物ES细胞的宿主胚胎细胞,并且使非人代孕母体动物分娩包含生殖细胞的非人动物后代,该生殖细胞包含编码异源CACNG1蛋白或其部分的核酸序列。在一些实施方案中,将编码异源CACNG1蛋白或其部分的核酸序列插入内源Cacng1基因座中。在此类实施方案中,插入步骤包括用编码异源CACNG1蛋白或其部分的核酸序列替换编码内源CACNG1蛋白或其部分的内源核酸序列,其中编码内源CACNG1蛋白或其部分的内源核酸序列和编码

异源CACNG1蛋白或其部分的核酸序列是直系同源的。在一些实施方案中,编码异源CACNG1蛋白或其部分的核酸序列包含:(i)包含人CACNG1基因的外显子1或其部分的核酸序列,(ii)包含人CACNG1基因的外显子2或其部分的核酸序列,(iii)包含人CACNG1基因的外显子3或其部分的核酸序列,(iv)包含人CACNG1基因的外显子4或其部分的核酸序列,或(v)(i)-(iv)的任何组合。在一些示例中,编码异源CACNG1蛋白或其部分的核酸序列包含:(i)包含人CACNG1基因的外显子1或其部分的核酸序列,(ii)人CACNG1基因的内含子1或其部分的核酸序列,(iii)包含人CACNG1基因的外显子2或其部分的核酸序列,(iv)人CACNG1基因的内含子2或其部分的核酸序列,(v)包含人CACNG1基因的外显子3或其部分的核酸序列,(vi)人CACNG1基因的内含子3或其部分的核酸序列,(v)包含人CACNG1基因的外显子4或其部分的核酸序列,(vii)人CACNG1基因的3'非翻译区(UTR)的核酸序列,或(v)(i)-(iv)的任何组合。在一些实施方案中,编码异源CACNG1蛋白或其部分的核酸序列包含选自以下的核酸序列、基本上由选自以下的核酸序列组成或由选自以下的核酸序列组成:如SEQ ID NO:5所示的核酸序列、如SEQ ID NO:27所示的核酸序列和如SEQ ID NO:28所示的核酸序列。在一些实施方案中,异源CACNG1蛋白或其部分包含人CACNG1蛋白或其部分的氨基酸序列。在一些实施方案中,异源CACNG1蛋白或其部分包含(i)如SEQ ID NO:8所示的氨基酸序列(人胞质结构域1);(ii)如SEQ ID NO:10所示的氨基酸序列(人跨膜结构域1);(iii)如SEQ ID NO:12所示的氨基酸序列(人胞外结构域1);(iv)如SEQ ID NO:14所示的氨基酸序列(人跨膜结构域2);(v)如SEQ ID NO:16所示的氨基酸序列(人胞质结构域2);(vi)如SEQ ID NO:18所示的氨基酸序列(人跨膜结构域3);(vii)如SEQ ID NO:20所示的氨基酸序列(人胞外结构域3);(viii)如SEQ ID NO:22所示的氨基酸序列(人跨膜结构域4);(ix)如SEQ ID NO:24所示的氨基酸序列(人胞质结构域4);或(x)(i)-(ii)的任何组合。在一些实施方案中,异源CACNG1蛋白包含如SEQ ID NO:4所示的氨基酸序列。在一些实施方案中;(i)非人动物是哺乳动物,诸如啮齿动物;(ii)非人动物细胞是哺乳动物细胞,诸如啮齿动物细胞;或(iii)非人动物基因组是哺乳动物核酸,诸如啮齿动物核酸。在一些实施方案中,(i)非人动物是大鼠或小鼠;(ii)非人动物细胞是大鼠细胞或小鼠细胞;或(iii)非人动物基因组是大鼠核酸或小鼠核酸。在一些实施方案中,非人动物是小鼠,非人动物细胞是小鼠细胞,或者非人动物基因组是小鼠核酸,并且编码异源CACNG1蛋白或其部分的核酸序列包含如SEQ ID NO:6所示的核酸序列、基本上由其组成或由其组成。

[0021] 在一些实施方案中,核酸的插入包括使非人动物的基因组、非人动物细胞的基因组或非人动物基因组与本公开的任何嵌合核酸分子接触。

[0022] 非人动物、非人动物细胞或非人动物基因组可以根据本公开的任何方法制备。

[0023] 在一些实施方案中,如本文所述的非人动物包含结合异源CACNG1蛋白的抗原结合蛋白,其中该非人动物在骨骼肌细胞的表面表达异源CACNG1蛋白或其胞外结构域。在一些实施方案中,异源CACNG1蛋白是人CACNG1蛋白。在一些实施方案中,所述非人动物是小鼠。

[0024] 在一些实施方案中,非人动物、非人动物细胞或非人动物基因组包含内源Cacng1基因的敲除突变。在一些情况下,敲除突变包括Cacng1基因或其部分的缺失。在具体的实施方案中,敲除突变包括Cacng1基因的全部编码序列的缺失。

[0025] 本文还提供了不表达任何CACNG1蛋白的非人动物、非人动物细胞或非人动物基因组。

[0026] 本文还提供了不表达骨骼肌特异性蛋白的非人动物、非人动物细胞或非人动物基因组,然而该非人动物、非人动物细胞或非人动物基因组不表现出任何重大突变表型。

[0027] 在一些实施方案中,本公开提供了一种靶向载体,该靶向载体包含:(i) 5' 同源臂和(ii) 3' 同源臂,其中5' 同源臂和3' 同源臂与感兴趣的非人动物Cacng1基因座进行同源重组,并且其中在与感兴趣的非人动物Cacng1基因座同源重组后,该靶向载体在感兴趣的非人动物Cacng1基因座处的非人动物Cacng1基因中插入敲除突变。

[0028] 在一些实施方案中,本公开提供了一种制备CACNG1敲除的非人动物的方法,该方法包括将非人动物的内源Cacng1基因座修饰为包含敲除突变。

附图说明

[0029] 图1是显示CACNG1表达的图表(GTEx Portal)。

[0030] 图2描绘产生CACNG1敲除的策略(未按比例)。

[0031] 图3A是说明与野生型(WT)对照动物相比,小鼠中CACNG1基因缺失(CACNG1^{-/-})不改变骨骼肌重量的图表。

[0032] 图3B是说明与野生型(WT)对照动物相比,小鼠中CACNG1基因缺失(CACNG1^{-/-})不改变颤搐(1Hz)或强直(125Hz)收缩力的图表。

[0033] 图4A示出CACNG1^{hu/hu}小鼠人源化的示意图(未按比例)。星号指示用于等位基因获得性测定的上游(7450hTU)和下游(7450hTD)引物的位置。图的上部分例示用于非动物基因组人源化的源自人CACNG1的12,484bp序列。图的下部分例示鼠类12,795bp基因组序列,其中Cacng1基因座靶向缺失。

[0034] 图4B详述7450等位基因(包括Neo自缺失盒)的CACNG1人源化策略(未按比例)。将小鼠Cacng1的编码外显子1、内含子1、编码外显子2-4(和插入内含子)和3' 非翻译区(UTR)的82bp的部分替换为相应的部分编码外显子1序列、内含子1、编码外显子2-4(和插入内含子)、完整的3' UTR和人CACNG1的3' UTR后另外的158bp。编码序列开始处的15bp仍然是小鼠序列。loxP-mPrm1-Crei-pA-hUb1-em7-Neo-pA-loxP盒(4,805bp)示于人序列的下游,随后是小鼠3' UTR的剩余部分。

[0035] 图4C详述7451等位基因的CACNG1人源化策略(未按比例),其中盒缺失,而LoxP位点保留。将小鼠Cacng1的编码外显子1、内含子1、编码外显子2-4(和插入内含子)和3' 非翻译区(UTR)的82bp的部分替换为相应的部分编码外显子1序列、内含子1、编码外显子2-4(和插入内含子)、完整的3' UTR和人CACNG1的3' UTR后另外的158bp。编码序列开始处的15bp仍然是小鼠序列。loxP-mPrm1-Crei-pA-hUb1-em7-Neo-pA-loxP盒(4,805bp)示于人序列的下游,随后是小鼠3' UTR的剩余部分。盒缺失后,LoxP和克隆位点(77bp)保留在人3' UTR之后。

[0036] 图5示出了小鼠CACNG1蛋白(mCACNG1; SEQ ID NO:2)与人hCACNG1蛋白(hCACNG1; SEQ ID NO:4)以及由7451等位基因编码的CACNG1蛋白(7451; SEQ ID NO:4)的比对。星号表示保持不变的残基。粗实线表示跨膜结构域。加下划线的残基是由引入的人外显子编码的残基。胞质结构域和胞外结构域被标记并示出。

[0037] 图6A是证明通过qPCR在CACNG1^{hu/hu}小鼠肌肉中无法检测到小鼠CACNG1(mCACNG1)表达(左图),而人CACNG1(hCACNG1)在CACNG1^{hu/hu}中表达,但在WT小鼠肌肉中不表达(右图)的图表。

[0038] 图6B是例示用100nM的Alexa 647缀合的人特异性 α -CACNG1 Ab对单个骨骼肌纤维进行活染色的图像,显示出与从CACNG1^{hu/hu}小鼠分离的肌纤维结合,但不与从WT小鼠分离的肌纤维结合。

[0039] 图6C是例示注射10mg/kg Alexa 647缀合的人特异性 α -CACNG1 Ab的CACNG1^{hu/hu}小鼠的低温荧光断层扫描(CryoFT)图像的图像,显示与同型对照Ab相比,在注射后6天对骨骼肌具有高特异性。

[0040] 图7是例示来自10mg/kg Alexa 647缀合的人特异性 α -CACNG1 Ab给药的CACNG1^{hu/hu}小鼠的组织学切片的图像,显示注射后6天与骨骼肌结合。上图展示来自体内注射的Ab的内源性Alexa 647信号,下图展示Alexa647-Ab与层粘连蛋白结合和DAPI共染色的叠加,以使肌肉形态可视化。

[0041] 图8A提供NP_031608.1所指的小鼠Cacng1蛋白的胞质结构域(氨基酸1-10、131-135和206-223)、跨膜结构域(氨基酸11-29、110-130、136-156和181-205)和胞外结构域(氨基酸30-109和157-180)的注释。

[0042] 图8B提供了编码小鼠编码DNA序列(CDS)的胞质结构域(核酸1-30、391-405和616-669)、跨膜结构域(核酸31-87、328-390、406-468和541-615)和胞外结构域(核酸88-327和469-540)的核酸序列的注释。

[0043] 图9A提供NP_000718.1所指的人CACNG1蛋白的胞质结构域(氨基酸1-10、130-134和205-222)、跨膜结构域(氨基酸11-29、109-129、135-155和180-204)和胞外结构域(氨基酸30-108和156-179)的注释。

[0044] 图9B提供编码人编码DNA序列(CDS)的胞质结构域、跨膜结构域和胞外结构域的核酸序列的注释。

[0045] 图10A提供由7451等位基因编码的CACNG1蛋白的例示性序列。

[0046] 图10B提供小鼠/人CACNG1核酸编码序列(CDS)的例示性序列,包括3'非翻译序列。

[0047] 图11提供7450等位基因的核酸序列。带有Neo自缺失盒的CACNG1人源化区域=小鼠(小写)_人_XhoI_LoxP_Prm_Crei_sv40_polya(小写)-hUbi-em7(小写)-NEO-PGK_polyA_LoxP_ICeUI_小鼠(小写)。

[0048] 图12提供7450等位基因的核酸序列。带有Neo自缺失盒的CACNG1人源化区域=小鼠(小写)_人_XhoI_LoxP_Prm_Crei_sv40_polya(小写)-hUbi-em7(小写)-NEO-PGK_polyA_LoxP_ICeUI_小鼠(小写)。

[0049] 定义

[0050] 术语“蛋白质”、“多肽”和“肽”在本文中可互换使用,并且包括任何长度的氨基酸的聚合形式,这些聚合形式包括编码和非编码的氨基酸以及经化学或生物化学修饰或衍生的氨基酸。该术语还包括已进行修饰的聚合物,例如具有修饰的肽主链的多肽。术语结构域可以指具有特定功能或结构的蛋白质或多肽的任何部分。

[0051] 蛋白质被说成具有“N末端”和“C末端”。术语“N末端”指蛋白质或多肽的起始,由具有游离胺基(-NH₂)的氨基酸。术语“C末端”指氨基酸链(蛋白质或多肽)的末端,由游离羧基(-COOH)终止。

[0052] 本文可互换使用的术语“核酸”和“多核苷酸”包括任何长度的核苷酸的聚合形式,包括核糖核苷酸、脱氧核糖核苷酸、或其类似物或修饰形式。核酸和多核苷酸可以包括单

链、双链和多链DNA或RNA、基因组DNA、cDNA、DNA-RNA杂合体以及包含嘌呤碱基、嘧啶碱基或其他天然的、化学修饰的、生物化学修饰的、非天然的或衍生的核苷酸碱基的聚合物。

[0053] 核酸被说成具有“5' 末端”和“3' 末端”，因为单核苷酸以这样的方式反应以制备寡核苷酸，使得一个单核苷酸戊糖环的5' 磷酸盐经由磷酸二酯键合附着至其在一个方向上的邻居的3' 氧。如果寡核苷酸的5' 磷酸盐未与单核苷酸戊糖环的3' 氧连接，则将寡核苷酸的末端称为“5' 末端”。如果寡核苷酸的3' 氧未与另一个单核苷酸戊糖环的5' 磷酸盐连接，则将寡核苷酸的末端称为“3' 末端”。即使是对于较大的寡核苷酸内部的核酸序列，也可以被说成具有5' 末端和3' 末端。在线性或环状DNA分子中，不连续元件称为5' 元件的“下游”或3' 元件的“下游”。

[0054] 术语“基因组整合的”指这样的核酸，其已被引入细胞内，使得核苷酸序列整合到细胞的基因组内，并且能够被其子代遗传。任何方案都可以用于将核酸稳定整合到细胞的基因组内。

[0055] 术语“靶向载体”指这样的重组核酸，其可以通过同源重组、非同源末端连接介导的连接、或以任何其它重组手段引入细胞的基因组中的靶位置。

[0056] 术语“病毒载体”指重组核酸，其包括至少一个病毒起源的元件，并且包括足以或允许包装到病毒载体颗粒内的元件。该载体和/或颗粒可以用于将DNA、RNA或其它核酸离体或体内转移到细胞内的目的。众多形式的病毒载体是已知的。

[0057] 术语“野生型”包括具有如在正常（与突变、患病、改变等等形成对比）状态或背景下发现的结构和/或活性的实体。野生型基因和多肽经常以多重不同形式（例如等位基因）存在。

[0058] 表述“重大突变表型”是指本公开的工程化的非人小鼠与“野生型”之间表型的显著差异或变异。

[0059] 术语“内源”指天然存在于细胞或非人动物内的核酸序列。例如，非人动物的内源Cacng1序列是指在非人动物的Cacng1基因座处天然存在的天然Cacng1序列。

[0060] “外源”分子或序列包括通常不以该形式存在于细胞中的分子或序列。正常存在包括关于细胞的特定发育阶段和环境条件的存在。例如，外源分子或序列可以包括细胞内的相应内源序列的突变形式，例如内源序列的人源化形式，或者可以包括与细胞内的内源序列相对应，但以不同形式（即，不在染色体内）的序列。相比之下，内源分子或序列包括在特定环境条件下、在特定发育阶段、在特定细胞中通常以该形式存在的分子或序列。

[0061] 当在核酸或蛋白质的上下文中使用时，术语“异源的”指示该核酸或蛋白质包含通常不在相同分子中一起存在的至少两个部分。例如，当关于核酸的一部分或蛋白质的一部分使用时，术语“异源的”指示该核酸或蛋白质包含在自然界中未发现处于彼此的相同关系中（例如，连接在一起）的两个或更多个子序列。作为一个例子，核酸载体的“异源”区域是在自然界中未发现与另一种分子结合的、在另一种核酸分子内或附着至另一种核酸分子的核酸的区段。例如，核酸载体的异源区域可以包括侧翼为在自然界中未发现与编码序列结合的序列的编码序列。同样地，蛋白质的“异源”区域是在自然界中未发现与另一种肽分子结合的、在另一种肽分子内或附着至另一种肽分子的氨基酸的区段（例如，融合蛋白或具有标签的蛋白质）。类似地，核酸或蛋白质可以包含异源标记或者异源分泌序列或定位序列。

[0062] “密码子优化”利用密码子的简并性，如指定氨基酸的三碱基对密码子组合的多样

性所表现的,并且通常包括通过用在宿主细胞基因中使用频率更高或使用频率最高的密码子替换天然序列的至少一个密码子,同时保持天然氨基酸序列来修饰核酸序列以增强在特定宿主细胞中的表达的过程。例如,可以修饰编码Cas9蛋白的核酸,以取代与天然存在的核酸序列相比、在给定的原核或真核细胞中具有更高使用频率的密码子,所述原核或真核细胞包括细菌细胞、酵母细胞、人细胞、非人细胞、哺乳动物细胞、啮齿类动物细胞、小鼠细胞、大鼠细胞、仓鼠细胞或任何其它宿主细胞。密码子使用表例如在“密码子使用数据库(Codon Usage Database)”处可容易获得。这些表可以通过多种方式进行修改。参见Nakamura等人(2000)Nucleic Acids Research 28:292,其以引用的方式整体并入本文用于所有目的。用于在特定宿主中表达的特定序列的密码子优化的计算机算法也是可获得的(参见例如, Gene Forge)。

[0063] 术语“基因座”指基因(或重要序列)、DNA序列、多肽编码序列的特定定位、或在生物体的基因组的染色体上的位置。例如,“Cacng1基因座”可以指Cacng1基因、Cacng1 DNA序列、编码Cacng1的序列的特定位置,或已鉴定为此类序列所在的地方的生物体基因组的染色体上的Cacng1位置。“Cacng1基因座”可以包含Cacng1基因的调控元件,包括例如增强子、启动子、5' 和/或3' 非翻译区(UTR)或它们的组合。

[0064] 术语“基因”指染色体中的DNA序列,其编码产物(例如,RNA产物和/或多肽产物),并且包括被非编码内含子中断的编码区、以及在5'和3'末端两者上与编码区相邻的序列,使得该基因对应于全长mRNA(包括5'和3'非翻译序列)。术语“基因”还包括其它非编码序列,包括调节序列(例如,启动子、增强子和转录因子结合位点)、聚腺苷酸化信号、内部核糖体进入位点、沉默子、绝缘序列和基质附着区。这些序列可能靠近基因的编码区(例如,在10kb内)或在遥远位点处,并且它们影响基因的转录和翻译水平或速率。

[0065] 术语“等位基因”指基因的变体形式。一些基因具有各种不同形式,其定位于染色体上的相同位置或遗传基因座处。二倍体生物在每个基因基因座处上具有两个等位基因。每对等位基因代表特定基因基因座的基因型。如果在特定基因座处存在两个相同的等位基因,则基因型被描述为纯合的;如果两个等位基因不同,则被描述为杂合的。

[0066] “启动子”是DNA的调节区,其通常包含能够指导RNA聚合酶II在特定多核苷酸序列的适当转录起始位点处起始RNA合成的TATA盒。启动子可以另外包含影响转录起始速率的其它区域。本文公开的启动子序列调节可操作地连接的多核苷酸的转录。启动子可以在本文公开的一种或多种细胞类型(例如,真核细胞、非人哺乳动物细胞、人细胞、啮齿类动物细胞、多能细胞、单细胞期胚胎、分化细胞或其组合)中是活跃的。启动子可以是例如组成型活性启动子、条件启动子、诱导型启动子、时间受限的启动子(例如,发育调节的启动子)或空间受限的启动子(例如,细胞特异性或组织特异性启动子)。启动子的例子可以例如在WO 2013/176772中发现,所述专利以引用的方式整体并入本文用于所有目的。

[0067] “可操作的连接”或“可操作地连接”包括两个或更多个组分(例如启动子和另一个序列元件)这样的并置,使得两个组分均正常起作用,并且允许至少一个组分可以介导对其它组分中的至少一个发挥的功能的可能性。例如,如果启动子响应一种或多种转录调节因子的存在或不存在而控制编码序列的转录水平,则该启动子可以与编码序列可操作地连接。可操作的连接可以包括彼此邻接或反式起作用的此类序列(例如,调节序列可在一定距离处起作用,以控制编码序列的转录)。

[0068] 术语“变体”指不同于群体中最普遍的序列(例如,相差一个核苷酸)的核苷酸序列,或不同于群体中最普遍的序列(例如,相差一个氨基酸)的蛋白质序列。

[0069] 当涉及蛋白质时,术语“片段”意指比全长蛋白质更短或具有更少氨基酸的蛋白质。当涉及核酸时,术语“片段”意指比全长核酸更短或具有更少核苷酸的核酸。片段可以是例如N末端片段(即去除蛋白质的C末端的一部分)、C末端片段(即,去除蛋白质的N末端的一部分)、或内部片段。

[0070] 在两个多核苷酸或多肽序列的上下文中,“序列同一性”或“同一性”指当在指定的比较窗上就最大对应性进行比对时,两个序列中相同的残基。当序列同一性的百分比提及蛋白质使用时,并不相同的残基位置经常因保守的氨基酸取代而不同,其中氨基酸残基被具有相似化学特性(例如电荷或疏水性)的其它氨基酸残基取代,且因此不改变分子的功能特性。当序列在保守取代中不同时,可以向上调整序列同一性百分比,以校正取代的保守性质。通过此类保守取代而不同的序列被称为具有“序列相似性”或“相似性”。用于进行这种调整的方法是众所周知的。通常,这涉及将保守取代评分为部分错配而不是完全错配,从而增加序列同一性百分比。因此,例如,当相同氨基酸给予得分1,而非保守取代给予得分零时,保守取代给予在零至1之间的得分。例如,如程序PC/GENE(Intelligenetics,Mountain View,California)中实现的计算保守取代的得分。

[0071] “序列同一性百分比”包括通过比较窗中比较两个最佳比对的序列(完全匹配残基的最大数目)而确定的值,其中在比较窗中的多核苷酸序列的部分可以包含与参考序列(其不包含添加或缺失)相比的添加或缺失(即,缺口),用于两个序列的最佳比对。该百分比通过以下来计算:测定在其处相同的核酸碱基或氨基酸残基在两个序列中出现的位置数目,以得出匹配位置数目,将匹配位置数目除以比较窗中的位置总数目,并且将结果乘以100,以得出序列同一性的百分比。除非另有说明(例如,较短的序列包括连接的异源序列),否则比较窗是待比较的两个序列中的较短者的全长。

[0072] 除非另有说明,否则序列同一性/相似性值包括使用GAP版本10使用下述参数获得的值:关于核苷酸序列的%同一性和%相似性,使用GAP权重50和长度权重3以及nwsgapdna.cmp评分矩阵;关于氨基酸序列的%同一性和%相似性,使用GAP权重8和长度权重2以及BLOSUM62评分矩阵;或其任何等价程序。“等价程序”包括任何序列比较程序,当与通过GAP版本10生成的相应比对相比较时,对于所讨论的任何两个序列,所述序列比较程序生成具有相同的核苷酸或氨基酸残基匹配和相同的序列百分比同一性的比对。

[0073] 术语“保守氨基酸取代”指序列中通常存在的氨基酸被具有相似大小、电荷或极性的不同氨基酸取代。保守取代的例子包括非极性(疏水性)残基,例如异亮氨酸、缬氨酸或亮氨酸被另一个非极性残基的取代。同样,保守替换的实例包括对另一种极性(亲水性)残基(诸如精氨酸与赖氨酸)之间的替代、谷氨酰胺与天冬酰胺之间或甘氨酸与丝氨酸之间的替代。另外,对另一种酸性残基,诸如赖氨酸、精氨酸或组氨酸,或者一种酸性残基(诸如天冬氨酸或谷氨酸)的取代是保守替代的另外实例。非保守替代的实例包括对非极性(疏水性)氨基酸残基诸如异亮氨酸、缬氨酸、亮氨酸、丙氨酸或甲硫氨酸进行取代,所述极性(亲水性)残基诸如半胱氨酸、谷氨酰胺、谷氨酸或赖氨酸和/或极性残基用于非极性残基。典型的氨基酸分类概括于下文。

[0074]	丙氨酸	Ala	A	非极性	中性	1.8
	精氨酸	Arg	R	极性	阳性	-4.5
	天冬酰胺	Asn	N	极性	中性	-3.5
	天冬氨酸	Asp	D	极性	阴性	-3.5
	半胱氨酸	Cys	C	非极性	中性	2.5
[0075]	谷氨酸	Glu	E	极性	阴性	-3.5
	谷氨酰胺	Gln	Q	极性	中性	-3.5
	甘氨酸	Gly	G	非极性	中性	-0.4
	组氨酸	His	H	极性	阳性	-3.2
	异亮氨酸	Ile	I	非极性	中性	4.5
	亮氨酸	Leu	L	非极性	中性	3.8
	赖氨酸	Lys	K	极性	阳性	-3.9
	甲硫氨酸	Met	M	非极性	中性	1.9
	苯丙氨酸	Phe	F	非极性	中性	2.8
	脯氨酸	Pro	P	非极性	中性	-1.6
	丝氨酸	Ser	S	极性	中性	-0.8
	苏氨酸	Thr	T	极性	中性	-0.7
	色氨酸	Trp	W	非极性	中性	-0.9
	酪氨酸	Tyr	Y	极性	中性	-1.3
	缬氨酸	Val	V	非极性	中性	4.2

[0076] “同源”序列(例如,核酸序列)包括与已知参考序列相同或基本上相似的序列,使得它与已知参考序列例如至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或100%相同。同源序列可以包括例如直系同源序列和旁系同源序列。同源基因例如通常通过物种形成事件(直系同源基因)或遗传复制事件(旁系同源基因),遗传自共同的祖先DNA序列。“直系同源”基因包括通过物种形成由共同的祖先基因进化的在不同物种中的基因。直系同源物通常在进化过程中保留相同的功能。“旁系同源”基因包括通过基因组内的复制相关的基因。旁系同源物可以在进化过程中进化出新功能。

[0077] 术语“体外”包括人工环境,以及在人工环境(例如,试管)内发生的过程或反应。术语“体内”包括自然环境(例如,细胞或生物或机体),以及在自然环境中发生的过程或反应。术语“离体”包括已从个体体内取出的细胞,以及在此类细胞内发生的过程或反应。

[0078] 术语“报道基因”指具有编码基因产物(通常是酶)的序列的核酸,当包含与异源启动子和/或增强子元件可操作地连接的报道基因序列的构建体,被引入含有(或其可以制备为含有)启动子和/或增强子元件活化所必需的因子的细胞内时,所述核酸很容易且可定量地测定。报道基因的例子包括但不限于编码 β -半乳糖苷酶(lacZ)的基因,细菌氯霉素乙酰转移酶(cat)基因,萤火虫萤光素酶基因,编码 β -葡萄糖醛酸酶(GUS)的基因和编码荧光蛋白的基因。“报道蛋白”指由报道基因编码的蛋白质。

[0079] 如本文使用的,术语“荧光报道蛋白”意指基于荧光可检测的报道蛋白,其中荧光可以直接来自报道蛋白、报道蛋白对荧光底物的活性、或对于与加上荧光标签的化合物结合具有亲和力的蛋白质。荧光蛋白的例子包括绿色荧光蛋白(例如GFP、GFP-2、tagGFP、turboGFP、eGFP、Emerald、Azami Green、Monomeric Azami Green、CopGFP、AceGFP和ZsGreen1),黄色荧光蛋白质(例如YFP、eYFP、Citrine、Venus、YPet、PhiYFP和ZsYellow1),蓝色荧光蛋白(例如BFP、eBFP、eBFP2、Azurite、mKalamal、GFPuv、Sapphire和T-sapphire),

青色荧光蛋白(例如CFP、eCFP、Cerulean、CyPet、AmCyan1和Midoriishi-Cyan),红色荧光蛋白(例如RFP、mKate、mKate2、mPlum、DsRed monomer、mCherry、mRFP1、DsRed-Express、DsRed2、DsRed-Monomer、HcRed-Tandem、HcRed1、AsRed2、eqFP611、mRaspberry、mStrawberry和Jred),橙色荧光蛋白(例如mOrange、mKO、Kusabira-Orange、Monomeric Kusabira-Orange、mTangerine和tdTomato),以及其在细胞中的存在可以通过流式细胞术方法检测的任何其它合适的荧光蛋白。

[0080] 术语“重组”包括在两个多核苷酸之间交换遗传信息的任何过程,并且可以通过任何机制发生。响应双链断裂(DSB)的重组主要通过两个保守的DNA修复途径发生:非同源末端连接(NHEJ)和同源重组(HR)。参见Kasperek&Humphrey(2011) *Seminars in Cell& Dev. Biol.* 22:886-897,其以引用的方式整体并入本文用于所有目的。同样地,由外源供体核酸介导的靶核酸的修复可以包括在两个多核苷酸之间交换遗传信息的任何过程。

[0081] NHEJ包括通过断裂末端彼此的直接连接或无需同源模板与外源序列的直接连接,来修复核酸中的双链断裂。NHEJ包括通过断裂末端彼此的直接连接或无需同源模板与外源序列的直接连接,来修复核酸中的双链断裂。非邻接序列通过NHEJ的连接经常可以导致接近双链断裂位点的缺失、掺入或易位。例如,NHEJ还可以通过断裂末端与外源供体核酸的末端的直接连接,导致外源供体核酸的靶向整合(即,基于NHEJ的捕获)。当同源性指导的修复(HDR)途径不易使用时(例如,在非分裂细胞、原代细胞、以及弱执行基于同源性的DNA修复的细胞),此类NHEJ介导的靶向整合对于外源供体核酸的插入可以是优选的。另外,与同源性指导的修复形成对比,不需要关于侧接切割位点的序列同一性的大区域的了解,当尝试靶向插入具有关于其存在基因组序列的有限了解的基因组的生物内时,这可以是有益的。整合可以经由在外源供体核酸与切割的基因组序列之间平端连接、或经由使用侧翼为突出端(所述突出端与在切割的基因组序列中由核酸酶试剂生成的那些相容)的外源供体核酸的粘性末端(即具有5'或3'突出端)连接来进行。参见例如,US2011/020722、WO 2014/033644、WO 2014/089290和Maresca等人(2013) *Genome Res.* 23(3):539-546,其各自以引用的方式整体并入本文用于所有目的。如果连接平端,则可能需要靶和/或供体切除,以生成片段连接所需的微同源区域,其可能在靶序列中产生不需要的改变。

[0082] 重组也可以经由同源性指导的修复(HDR)或同源重组(HR)发生。HDR或HR包括核酸修复形式,其可能需要核苷酸序列同源性,使用“供体”分子作为用于修复“靶”分子(即,经历双链断裂的分子)的模板,并且导致遗传信息从供体转移到靶。不希望受到任何特定理论的束缚,此类转移可以涉及在断裂的靶和供体之间形成的异源双链体DNA的错配校正和/或合成依赖性链退火,其中供体用于重新合成将成为靶和/或相关过程的部分的遗传信息。在一些情况下,供体多核苷酸、供体多核苷酸的一部分、供体多核苷酸的拷贝、或供体多核苷酸的拷贝的一部分整合到靶DNA内。参见Wang等人(2013) *Cell* 153:910-918; Mandalos等人(2012) *PLOS ONE* 7:e45768:1-9;以及Wang等人(2013) *Nat Biotechnol.* 31:530-532,其各自以引用的方式整体并入本文用于所有目的。

[0083] 术语“抗原结合蛋白”包括与抗原结合的任何蛋白质。抗原结合蛋白的例子包括抗体,抗体的抗原结合片段,多特异性抗体(例如,双特异性抗体),scFV,双scFV,双抗体,三抗体,四抗体,V-NAR,VHH,VL,F(ab),F(ab)₂,DVD(双可变结构域抗原结合蛋白),SVD(单可变结构域抗原结合蛋白),双特异性T细胞衔接子(BiTE)或Davisbody(美国专利号8,586,713,

其以引用的方式整体并入本文用于所有目的)。

[0084] 提及抗原结合蛋白的术语“多特异性”或“双特异性”意指该蛋白识别相同抗原或不同抗原上的不同表位。多特异性抗原结合蛋白可以是单个多功能多肽,或者它也可以是彼此共价或非共价结合的两个或更多个多肽的多聚体复合物。例如,抗体或其片段可与一个或多个其它分子实体例如蛋白质或其片段功能性连接(例如通过化学偶联、基因融合、非共价结合或其它方式),以产生具有第二结合特异性的双特异性或多特异性抗原结合分子。

[0085] 术语“抗原”指物质,无论是整个分子还是分子内的结构域,其能够引发对该物质具有结合特异性的抗体产生。术语抗原还包括这样的物质,其在野生型宿主生物中不由于自身识别引发抗体产生,但可以用适当的基因工程来破坏免疫耐受性而在宿主动物中引发此类应答。

[0086] 术语“表位”指抗原上抗原结合蛋白(例如抗体)与之结合的位点。表位可以由邻接氨基酸、或者通过一种或多种蛋白质的三级折叠并置的非邻接氨基酸形成。由邻接氨基酸形成的表位(也称为线性表位)通常在暴露于变性溶剂时保留,而通过三级折叠形成的表位(也称为构象表位)通常在用变性溶剂处理后丧失。表位通常包含处于独特的空间构象中的至少3个,本文更通常为至少5个或8-10个氨基酸。确定表位的空间构象的方法包括例如x射线晶体学和2维核磁共振。参见例如,Epitope Mapping Protocols,于Methods in Molecular Biology,第66卷,Glenn E.Morris,编辑(1996),其以引用的方式整体并入本文用于所有目的。

[0087] 如本文所述的抗体互补位一般至少包含特异性识别异源表位的互补决定区(CDR)(例如,重链和/或轻链可变结构域的CDR3区)。

[0088] 术语“抗体”包括免疫球蛋白分子,其包含通过二硫键互连的四条多肽链,两条重(H)链和两条轻(L)链。每条重链包含重链可变域和重链恒定区(C_H)。重链恒定区包含三个结构域: C_{H1} 、 C_{H2} 和 C_{H3} 。每条轻链包含轻链可变域和轻链恒定区(C_L)。重链和轻链可变域可以进一步细分成高变区,称为互补决定区(CDR),其间穿插有保守性更高的区域,称为构架区(FR)。每个重链和轻链可变域包含三个CDR和四个FR,其从氨基端到羧基端按以下顺序排列:FR1、CDR1、FR2、CDR2、FR3、CDR3、FR4(重链CDR可以简称为HCDR1、HCDR2和HCDR3;轻链CDR可以简称为LCDR1、LCDR2和LCDR3)。术语“高亲和力”抗体是指相对于其靶表位的 K_D 约为 10^{-9} M或更低(例如,约 1×10^{-9} M、 1×10^{-10} M、 1×10^{-11} M或约 1×10^{-12} M)的抗体。在一个实施方案中, K_D 通过表面等离子共振例如BIAcore™进行测量;在另一个实施方案中, K_D 通过ELISA进行测量。

[0089] 术语“双特异性抗体”包括能够选择性结合两个或更多个表位的抗体。双特异性抗体一般包含两条不同的重链,其中每条重链特异性结合不同的表位—在两种不同的分子上(例如,在两种不同的抗原上)或在相同的分子上(例如在相同的抗原上)。如果双特异性抗体能够选择性结合两个不同的表位(第一表位和第二表位),则第一重链对于第一表位的亲和力一般比第一重链对于第二表位的亲和力低至少一个到两个或三个或四个数量级,并且反之亦然。由双特异性抗体识别的表位可在相同或不同的靶上(例如在相同或不同的蛋白质上)。双特异性抗体可例如通过组合识别相同抗原的不同表位的重链来制备。例如,编码识别相同抗原的不同表位的重链可变序列的核酸序列可与编码不同重链恒定区的核酸序列融合,并且这样的序列可在表达免疫球蛋白轻链的细胞中表达。通常的双特异性抗体具

有两条重链以及免疫球蛋白轻链,所述两条重链各自具有三个重链CDR,随后为(N末端至C末端)CH1结构域、铰链、CH2结构域和CH3结构域,所述免疫球蛋白轻链不赋予抗原结合特异性,但可与每条重链结合,或者可与每条重链结合并且可结合由重链抗原结合区结合的一个或多个表位,或者可与每条重链结合并且使得重链之一或两者能够与一个或两个表位结合。

[0090] 术语“重链”或“免疫球蛋白重链”包括来自任何生物的免疫球蛋白重链序列,包括免疫球蛋白重链恒定区序列。除非另外说明,否则重链可变结构域包括三个重链CDR和四个FR区。重链的片段包括CDR、CDR和FR以及其组合。典型重链在可变域之后具有(从N端到C端) C_H1 结构域、铰链、 C_H2 结构域和 C_H3 结构域。重链的功能性片段包括能够特异性识别表位(例如,以微摩尔、纳摩尔或皮摩尔范围内的 K_D 识别表位)、能够由细胞表达和分泌以及包含至少一个CDR的片段。重链可变结构域由可变区核苷酸序列编码,所述可变区核苷酸序列通常包括源自存在于种系中的 V_H 、 D_H 和 J_H 区段库的 V_H 、 D_H 和 J_H 区段。关于各种生物的V、D和J重链区段的序列、定位和命名法可以在IMGT数据库中找到,所述IMGT数据库可经由因特网在万维网(www)上的URL“imgt.org”处访问。

[0091] 术语“轻链”包括来自任何生物的免疫球蛋白轻链序列,并且除非另有说明,否则包括人kappa(κ)和lambda(λ)轻链和VpreB以及替代轻链。除非另外说明,否则轻链可变域通常包括三个轻链CDR和四个构架(FR)区。一般地,全长轻链从氨基末端到羧基末端包括可变结构域和轻链恒定区,所述可变结构域包含FR1-CDR1-FR2-CDR2-FR3-CDR3-FR4和轻链恒定区氨基酸序列。轻链可变结构域由轻链可变区核苷酸序列编码,所述轻链可变区核苷酸序列一般包含源自种系中存在的轻链V和J基因区段储库的轻链 V_L 和轻链 J_L 基因区段。关于各种生物的轻链V和J基因区段的序列、定位和命名法可以在IMGT数据库中找到,所述IMGT数据库可经由因特网在万维网(www)上的URL“imgt.org”处访问。轻链包括例如不选择性结合第一表位或第二表位的那些,所述第一表位或第二表位被它们在其中出现的表位结合蛋白选择性结合。轻链还包括这样的轻链,其结合且识别或帮助重链结合且识别,被它们在其中出现的表位结合蛋白选择性结合的一个或多个表位。

[0092] 如本文使用的,术语“互补决定区”或“CDR”包括由生物的免疫球蛋白基因的核酸序列编码的氨基酸序列,其通常(即,在野生型动物中)出现在免疫球蛋白分子(例如抗体或T细胞受体)的轻链或重链可变区中的两个构架区之间。CDR可以由例如种系序列或重排序列编码,并且例如由幼稚或成熟B细胞或T细胞编码。CDR可以进行体细胞突变(例如,不同于动物的种系中编码的序列),人源化和/或用氨基酸取代、添加或缺失进行修饰。在一些情况下(例如,对于CDR3),CDR可以由两个或更多个序列(例如,种系序列)编码,所述两个或更多个序列并非邻接的(例如,在未重排的核酸序列中),但在B细胞核酸序列中是邻接的,例如,由于剪接或连接序列(例如V-D-J重组以形成重链CDR3)。

[0093] 抗原结合蛋白与其靶抗原的特异性结合包括以至少 10^6 、 10^7 、 10^8 、 10^9 或 $10^{10}M^{-1}$ 的亲合力的结合。特异性结合在量级中可检测地更高,并且可与对至少一种无关靶发生的非特异性结合区分开。特异性结合可以是特定官能团之间或特定空间配合(例如锁和钥匙类型)之间的键形成的结果,而非特异性结合通常是范德华力的结果。然而,特异性结合不一定暗示抗原结合蛋白结合一个靶且仅结合一个靶。

[0094] “任选的”或“任选地”意指随后描述的事件或情形可能发生或可能不发生,并且该

描述包括其中事件或情形发生的情况、以及其中事件或情形不发生的情况。

[0095] 值范围的指定包括在该范围内或定义该范围的所有整数,以及由在该范围内的整数定义的所有子范围。

[0096] 除非由上下文显而易见的,否则术语“约”涵盖所述值的标准测量误差界限内的值(例如,SEM)。

[0097] 术语“和/或”指并且涵盖一个或多个相关联的所列项目的任何和所有可能的组合,以及当以替代方式(“或”)解释时组合的缺乏。

[0098] 术语“或”指特定列表的任何一个成员,并且还包括该列表的成員的任何组合。

[0099] 除非上下文另有明确说明,否则冠词“一个”、“一种”和“该/所述”的单数形式包括复数所指物。例如,术语“一种蛋白质”或“至少一种蛋白质”可以包括多种蛋白质,包括其混合物。

[0100] 统计学显著的意指 $p \leq 0.05$ 。

具体实施方式

[0101] I. 概述

[0102] 本文公开了包含发现在骨骼肌中特异性表达的外源序列的非人动物细胞、非人动物和非人基因组及用于制备这些非人动物细胞、非人动物和非人基因组的试剂。在一些实施方案中,外源序列掺入基因的内源基因座中。

[0103] 骨骼肌是人体三大重要肌肉组织之一。每块骨骼肌由数千条由结缔组织鞘包裹在一起的肌纤维组成。骨骼肌中单独的肌纤维束称为纤维束。围绕整个肌肉的最外层结缔组织鞘称为肌外膜。覆盖每个纤维束的结缔组织鞘称为肌束膜,并且围绕单条肌纤维的最内层鞘称为肌内膜。每条肌纤维都包含含多条肌丝的肌原纤维。

[0104] 当捆绑在一起时,所有肌原纤维以独特的条纹图案排列形成肌节,肌节是骨骼肌的基本收缩单位。两种最重要的肌丝是肌动蛋白丝和肌球蛋白丝,它们独特排列以在骨骼肌上形成各种条带。分化成成熟肌纤维的干细胞称为卫星细胞,可以发现于基底膜和肌膜(横纹肌纤维细胞周围的细胞膜)之间。当受到生长因子刺激时,干细胞分化并繁殖形成新的肌纤维细胞。

[0105] 骨骼肌的主要功能是经由骨骼肌的内在兴奋-收缩耦联过程进行的。当肌肉附着在骨腱上时,肌肉收缩引起骨骼移动,骨骼移动允许特定移动的进行。骨骼肌还提供结构支撑并保持保持身体姿势。骨骼肌还充当氨基酸的储存来源,氨基酸可以被身体不同器官用于合成器官特异性蛋白质。骨骼肌在保持体温恒定方面也起着重要作用并在饥饿时充当能量来源。已经发现CACNG1蛋白在骨骼肌中特异性表达。

[0106] 在一些实施方案中,本文提供了非人动物细胞和非人动物,在本文提供的非人动物细胞或非人动物的基因组中具有异源Cacng1序列。异源Cacng1序列可以插入内源Cacng1基因座中,从而提供具有经遗传修饰的内源Cacng1基因座的非人动物细胞和非人动物。

[0107] 在一些实施方案中,本文提供了编码异源序列的核酸,该异源序列编码Cacng1序列的至少一部分,并且本文提供了用此类核酸制备非人动物细胞和非人动物的方法。在一些实施方案中,此类核酸有利于非人动物编辑的序列(例如,loxP位点),该序列位于编码Cacng1基因的序列的侧翼。

[0108] 在一些实施方案中,本文提供了针对由本公开的非人动物细胞和/或非人动物产生的嵌合CACNG1蛋白的抗体。

[0109] 在一些实施方案中,本公开提供了可用于制备此类非人动物(例如,啮齿动物,例如大鼠或小鼠)、源自此类非人动物的细胞和/或组织以及核苷酸(例如,靶向载体、基因组等)的方法。

[0110] 在一些实施方案中,本公开还提供了非人动物基因组,该非人动物基因组包含具有异源Cacng1序列的经遗传修饰的内源CACNG1基因座。在一些实施方案中,异源Cacng1序列编码CACNG1人蛋白序列。在一些情况下,CACNG1结构域的全部或部分由内源Cacng1基因座的区段编码,该区段已缺失并用异源Cacng1序列替换。

[0111] 在一些实施方案中,提供了包含人源化Cacng1基因座并由人源化Cacng1基因座表达人源化或嵌合CACNG1蛋白的非人动物,以及使用此类非人动物(例如,啮齿动物,例如大鼠或小鼠)、源自此类非人动物的细胞和/或组织以及可用于制备此类动物的核苷酸(例如,靶向载体、基因组等)的方法。

[0112] 在一些实施方案中,本文描述了包含编码经修饰的CACNG1蛋白的经遗传修饰的Cacng1基因座的非人动物,其中经修饰的CACNG1蛋白包含人CACNG1序列的结构域,并且该结构域的全部或部分由内源Cacng1基因座的区段编码,该区段已经缺失并用直系同源的人CACNG1序列替换,并且其中该非人动物表达经修饰的Cacng1蛋白。

[0113] 在一些实施方案中,人CACNG1序列的结构域由内源Cacng1基因座的区段编码,该区段已经缺失并用异源序列替换。此类结构域可以是人Cacng1胞外结构域。本公开所考虑的编码胞外结构域的合适序列在细胞内翻译时包括对应于CACNG1蛋白的氨基酸30-108 (SEQ ID NO:12)、氨基酸156-179 (SEQ ID NO:20) 或两者的人胞外结构域。

[0114] 在一些实施方案中,人CACNG1序列的至少两个结构域由人源化小鼠模型中内源Cacng1基因座的区段编码。人CACNG1序列的非限制性结构域的例示性实例包含胞质结构域、跨膜结构域和胞外结构域。在一些情况下,每个结构域的全部或部分由内源Cacng1基因座的区段编码,该区段已缺失并用直系同源的人CACNG1序列替换。在其他情况下,胞质结构域和胞外结构域可以任选地由内源基因组编码。在一些实施方案中,胞质结构域和跨膜结构域两者的全部或部分由内源Cacng1基因座的区段编码,该区段已经缺失并用直系同源的人CACNG1序列替换。在一些实施方案中,所有胞质结构域、跨膜结构域和胞外结构域都由内源Cacng1基因座的区段编码,该区段已经缺失并用直系同源的人CACNG1序列替换。后者将人CACNG1基因的多个个人源化结构域掺入非人基因组中;前者允许胞外膜的人源化,同时保留结构域中已知位于膜内和细胞内的内源结构域。编码本公开的胞质结构域的合适序列在细胞内翻译时产生对应于CACNG1蛋白的氨基酸1-10 (SEQ ID NO:8)、氨基酸130-134 (SEQ ID NO:16)、氨基酸205-222 (SEQ ID NO:24) 或它们的任何组合的人胞质结构域。编码本公开的跨膜结构域的合适序列在细胞内翻译时产生对应于CACNG1蛋白的氨基酸11-29 (SEQ ID NO:10)、氨基酸109-129 (SEQ ID NO:14)、氨基酸135-155 (SEQ ID NO:18)、氨基酸180-204 (SEQ ID NO:22) 或它们的任何组合的人跨膜结构域。因此,在一些替代实施方案中,胞质结构域或跨膜结构域的全部或部分由内源性非人动物Cacng1序列编码。

[0115] 在一些实施方案中,本文所述的非人动物或非人动物基因组编码代替内源性小鼠Cacng1序列的直系同源的人CACNG1序列。在一些实施方案中,非人动物或非人动物基因组

包含选自如SEQ ID NO:5所示的核酸序列、如SEQ ID NO:27所示的核酸序列和如SEQ ID NO:28所示的核酸序列。

[0116] 在一些实施方案中,人CACNG1序列由内源Cacng1基因座的区段编码,该区段已经缺失并用编码电压依赖性钙通道的全长 γ 1 结构域的人Cacng1序列替换。

[0117] 在一些实施方案中,本文所述的非人动物或非人动物基因组对于经遗传修饰的内源Cacng1基因座是杂合的。在一些实施方案中,非人动物或非人动物基因组对于经遗传修饰的内源Cacng1基因座是纯合的。

[0118] 在一些实施方案中,内源Cacng1基因座的区段缺失并用外源Cacng1序列替换。在这些情况中的一些情况下,已经缺失的内源Cacng1基因座可包含内源Cacng1基因座的3' 非翻译区区段、编码外显子1区段、内含子1区段、编码外显子2区段、内含子2区段、编码外显子3区段、内含子3区段、编码外显子4区段或上述区段的组合。

[0119] 在一些实施方案中,人CACNG1序列可用于替换非人动物或非人细胞内的基因座。在此类实施方案中,替换内源基因座区段的直系同源的人CACNG1序列可以包含人CACNG1序列的3' 非翻译区、人CACNG1序列的外显子1、人CACNG1序列的内含子1、人CACNG1序列的外显子2、人CACNG1序列内含子2、人CACNG1序列的外显子2、人CACNG1序列的内含子3、人CACNG1序列的外显子3、人CACNG1序列的内含子4、人CACNG1序列的外显子4或它们的任何组合的区段。

[0120] 在一些实施方案中,非人动物是哺乳动物,或者非人动物基因组是哺乳动物基因组。在一些实施方案中,非人动物可以是啮齿动物,或者非人动物基因组可以是啮齿动物基因组。在一些实施方案中,非人动物可以是大鼠或小鼠,或者非人动物基因组可以是大鼠基因组或小鼠基因组。

[0121] 在一些实施方案中,掺入非动物的基因组或非动物基因组上的异源序列编码人Cacng1胞外结构域、人Cacng1跨膜结构域和人Cacng1结构域。

[0122] 在一些实施方案中,掺入非动物的基因组或非动物基因组上的异源序列编码人CACNG1序列的至少两个结构域。两个或更多个结构域的非限制性实例包括第一胞质结构域、第一跨膜结构域、第一胞外结构域、第二跨膜结构域、第二胞质结构域、第三跨膜结构域、第二胞外结构域、第四跨膜结构域、第三胞质结构域。

[0123] 在一些实施方案中,掺入非动物的基因组或非动物基因组上的异源序列包含用于编码氨基酸1-10(胞质结构域)、氨基酸11-29(跨膜结构域)、氨基酸30-108(胞外结构域)、氨基酸109-129(跨膜结构域)、氨基酸130-134(胞质结构域)、氨基酸135-155(跨膜结构域)、氨基酸156-179(胞外结构域)、氨基酸180-204(跨膜结构域)、氨基酸205-222(胞质结构域)或它们的任何合适组合的合适序列。

[0124] 在一些实施方案中,本文提供了包含编码经修饰的CACNG1蛋白的经遗传修饰的内源Cacng1基因座的非人动物细胞,其中经修饰的Cacng1蛋白包含人CACNG1序列的结构域,并且该结构域的全部或部分由内源Cacng1基因座的区段编码,该区段已经缺失并用直系同源的人CACNG1序列替换。该非人动物细胞可以是骨骼肌细胞、多能细胞、ES细胞或生殖细胞。

[0125] 在一些实施方案中,本公开还提供了用于制备任何非人动物的方法,或制备如本文所述的非人动物所需的试剂。

[0126] A. 钙电压门控通道辅助亚基 γ 1 (CACNG1)

[0127] 本文所述的细胞和非人动物通常含有编码CACNG1蛋白区段(例如,人CACNG1蛋白结构域)的外源序列。电压依赖性钙通道通常由五个亚基组成。由CACNG1基因编码的蛋白质代表这些亚基之一。此外,由CACNG1基因编码的蛋白质 γ 是两种已知的 γ 亚基蛋白质之一。这种特殊的 γ 亚基是骨骼肌1,4-二氢吡啶敏感性钙通道的一部分并且是在兴奋-收缩耦联中起作用的整合膜蛋白。该基因是PMP-22/EMP/MP20家族功能多样的八成员蛋白亚家族的一部分,与两个起到跨膜AMPA受体调节蛋白(TARP)作用的家族成员成簇定位。

[0128] 编码人CACNG1的基因(CACNG1)位于17号染色体的长臂上。CACNG1包含4个外显子并且大约12,244个碱基长。

[0129] 人CACNG1的示例性序列分配NCBI登录号NM_0007582.2(参见图4A)。小鼠Cacng1的示例序列分配NCBI登录号NM_000727.4(参见图4A)。示例人CACNG1蛋白分配UniProt登录号070578(参见图4A和图5)。示例小鼠CACNG1蛋白分配UniProt登录号Q06432(参见图4A和图5)。图5示出了经修饰的非人Cacng1基因座(例如,7451)编码的示例人或人源化CACNG1蛋白。示例大鼠CACNG1蛋白分配NCBI参考序列:NP_062128.1。示例猩猩Cacng1蛋白分配NCBI参考序列:XP_002827789.2。

[0130] 在一些实施方案中,本公开提供了包含编码异源CACNG1蛋白或其部分的核酸序列的非人动物、非人动物细胞或非人动物基因组。编码异源CACNG1蛋白或其部分的此类核酸序列包含:(i)包含人CACNG1基因的外显子1或其部分的核酸序列;(ii)包含人CACNG1基因的外显子2或其部分的核酸序列;(iii)包含人CACNG1基因的外显子3或其部分的核酸序列;(iv)包含人CACNG1基因的外显子4或其部分的核酸序列;或(v)(i)-(iv)的任何组合。

[0131] 此外,掺入本文所述的非人动物的基因组、非人动物细胞或非人动物基因组的核酸序列可以包含内含子。在一些实施方案中,编码异源CACNG1蛋白或其部分的核酸序列包含:(i)包含人CACNG1基因的外显子1或其部分的核酸序列;(ii)人CACNG1基因的内含子1或其部分的核酸序列;(iii)包含人CACNG1基因的外显子2或其部分的核酸序列;(iv)人CACNG1基因的内含子2或其部分的核酸序列;(v)包含人CACNG1基因的外显子3或其部分的核酸序列;(vi)人CACNG1基因的内含子3或其部分的核酸序列;(v)包含人CACNG1基因的外显子4或其部分的核酸序列;(vii)人CACNG1基因的3'非翻译区(UTR)的核酸序列;或(v)(i)-(iv)的任何组合。

[0132] 在一些实施方案中,本文所述的非人动物、非人动物细胞或非人动物基因组编码CACNG1蛋白的人源化编码区(即,一些小鼠调节区和选定的人非编码/编码区)。在一些实施方案中,编码异源CACNG1蛋白或其部分的核酸序列可以包含编码人源化小鼠/人CACNG1蛋白的核酸序列、基本上由该核酸序列组成或由该核酸序列组成,该核酸序列诸如选自以下的核酸序列:如SEQ ID NO:5所示的核酸序列、如SEQ ID NO:27所示的核酸序列和如SEQ ID NO:28所示的核酸序列。任何此类核酸都可以掺入内源Cacng1基因座中。在一些实施方案中,编码异源CACNG1蛋白或其部分的核酸序列可以替换编码内源CACNG1蛋白或其部分的直系同源内源核酸序列。

[0133] 在一些实施方案中,本公开提供了非人动物、非人动物细胞或非人动物基因组,其中异源CACNG1蛋白或其部分包含(i)如SEQ ID NO:8所示的氨基酸序列;(ii)如SEQ ID NO:10所示的氨基酸序列;(iii)如SEQ ID NO:12所示的氨基酸序列;(iv)如SEQ ID NO:14所示

的氨基酸序列; (v) 如SEQ ID NO:16所示的氨基酸序列; (vi) 如SEQ ID NO:18所示的氨基酸序列; (vii) 如SEQ ID NO:20所示的氨基酸序列; (viii) 如SEQ ID NO:22所示的氨基酸序列; (ix) 如SEQ ID NO:24所示的氨基酸序列; 或 (x) (i) - (ii) 的任何组合。

[0134] B. 经修饰的Cacng1非人动物

[0135] 本公开提供了CACNG1蛋白功能丧失的非人动物和嵌合动物(例如, 表达人源化CACNG1蛋白的转基因啮齿动物)。人源化Cacng1基因座可以是其中整个Cacng1基因被相应直系同源的人CACNG1序列替换的Cacng1基因座, 或者可以是其中Cacng1基因的仅一部分被相应直系同源的人CACNG1序列替换(即人源化)的Cacng1基因座。任选地, 基于非人动物中的密码子使用, 相应直系同源的人CACNG1序列经修饰为密码子优化的。替换的(即, 人源化的)区域可以包括编码区例如外显子、非编码区例如内含子、非翻译区或调节区(例如启动子、增强子或转录阻遏物结合元件)、或其任何组合。作为一个实例, 对应于人CACNG1基因的1、2、3、4个或所有4个外显子的外显子可以是人源化的。例如, 对应于人CACNG1基因的外显子1-4的外显子可以是人源化的。可替代地, 编码被抗人CACNG1抗原结合蛋白识别的表位的Cacng1区域可以进行人源化。作为另一个实例, N-末端胞质结构域、跨膜结构域或胞内结构域中的一个或多个或全部可以是人源化的。例如, Cacng1基因座的编码胞外结构域的全部或部分区域可以是人源化的, Cacng1基因座的编码胞质结构域的全部或部分区域可以是人源化的, 和/或Cacng1基因座的编码跨膜结构域的全部或部分区域可以是人源化的。在一个实例中, 只有Cacng1基因座的编码跨膜结构域的全部或部分区域是人源化的, 只有Cacng1基因座的编码胞质结构域的全部或部分区域是人源化的, 或者只有Cacng1基因座的编码胞外区的全部或部分区域(即, 可用作表位的区域)是人源化的。例如, Cacng1基因座的编码胞外结构域的区域可以人源化, 使得产生具有内源性N-末端胞质结构域、内源跨膜结构域和人源化跨膜结构域(表位)的嵌合Cacng1蛋白。同样, 对应于人CACNG1基因的1、2、3个或所有4个内含子的内含子可以是人源化的。侧翼非翻译区包括调节序列也可以进行人源化。例如, 5'非翻译区(UTR)、3'UTR、或5'UTR和3'UTR两者可以进行人源化, 或5'UTR、3'UTR、或5'UTR和3'UTR两者可以保持内源的。在一个具体例子中, 3'UTR是人源化的, 但5'UTR保持内源的。取决于由直系同源序列替换的程度, 调节序列例如启动子可以是内源的, 或可以由替换的人直系同源序列提供。例如, 人源化Cacng1基因座可以包括内源非人动物Cacng1启动子。

[0136] 由人源化Cacng1基因座编码的Cacng1蛋白可以包含一个或多个来自哺乳动物CACNG1蛋白(例如, 人)的结构域。例如, Cacng1蛋白可以包含人胞外结构域、人CACNG1跨膜结构域和人CACNG1胞质结构域中的一个或多个或全部。作为一个实例, Cacng1蛋白可以仅包含人CACNG1胞外结构域。任选地, 由人源化Cacng1基因座编码的Cacng1蛋白也可以包含一个或多个来自内源(即, 天然)的非人动物Cacng1蛋白的结构域。

[0137] 来自人CACNG1蛋白的结构域可以由完全人源化的序列编码(即, 编码该结构域的整个序列被直系同源的人CACNG1序列替换)或者可以由部分人源化的序列编码(即, 编码该结构域的一些序列被直系同源的人CACNG1序列替换, 并且编码该结构域的剩余内源(即天然)序列编码与直系同源的人CACNG1序列相同的氨基酸, 使得编码的结构域与人CACNG1蛋白中的该结构域相同)。

[0138] 作为一个实例, 由人源化Cacng1基因座编码的Cacng1蛋白可以包含人CACNG1胞外结构域(例如: 人表位)。任选地, 人CACNG1跨膜结构域包含与图9A所示的序列至少85%、

90%、95%、96%、97%、98%、99%或100%相同的序列,基本上由该序列组成或由该序列组成,并且CACNG1蛋白保持天然CACNG1的活性(即,保持其在骨骼肌中的功能)。

[0139] 作为另一个实例,由人源化Cacng1基因座编码的CACNG1蛋白可以包含人跨膜或胞质CACNG1结构域。任选地,人CACNG1胞外结构域包含与图9A至少85%、90%、95%、96%、97%、98%、99%或100%相同的序列,基本上由该序列组成或由该序列组成,并且Cacng1蛋白保持天然CACNG1的活性。

[0140] 任选地,人源化Cacng1基因座可以包含其他元件。此类元件的实例可以包括选择盒、报道基因、重组酶识别位点或其他元件。可替代地,人源化Cacng1基因座可以缺少其他元件(例如,可以缺少选择标记或选择盒)。合适的报道基因和报道蛋白的例子在本文其它地方公开。合适的选择标记物的例子包括新霉素磷酸转移酶(neo_r)、潮霉素B磷酸转移酶(hyg_r)、嘌呤霉素-N-乙酰基转移酶(puro_r)、杀稻瘟菌素S脱氨酶(bsr_r)、黄嘌呤/鸟嘌呤磷酸核糖基转移酶(gpt)和单纯疱疹病毒胸苷激酶(HSV-k)。重组酶的例子包括Cre、Flp和Dre重组酶。Cre重组酶基因的一个例子是Crei,其中编码Cre重组酶的两个外显子被内含子分开,以防止其在原核细胞中表达。此类重组酶还可以包含核定位信号,以促进对核的定位(例如NLS-Crei)。重组酶识别位点包括这样的核苷酸序列,其被位点特异性重组酶识别,并且可以充当重组事件的底物。重组酶识别位点的例子包括FRT、FRT11、FRT71、attP、att、rox和lox位点,例如loxP、lox511、lox2272、lox66、lox71、loxM2和lox5171。

[0141] 其它元件,例如报道基因或选择盒可以是侧翼为重组酶识别位点的自缺失盒。参见例如,US 8,697,851和US2013/0312129,其各自以引用的方式整体并入本文用于所有目的。作为一个例子,自缺失盒可以包含与小鼠Prm1启动子可操作连接的Crei基因(包含编码Cre重组酶的两个外显子,其由内含子分开)、以及与人泛素启动子可操作连接的新霉素抗性基因。通过采用Prm1启动子,自缺失盒可以在F0动物的雄性生殖细胞中特异地缺失。可以将编码选择标记物的多核苷酸可操作地连接到在待靶性的细胞中活跃的启动子。启动子的例子在本文其它地方描述。作为另一个具体例子,自缺失选择盒可以包含与一个或多个启动子(例如,人泛素和EM7启动子两者)可操作连接的潮霉素抗性基因编码序列,随后为聚腺苷酸化信号,随后为与一个或多个启动子(例如mPrm1启动子)可操作连接的Crei编码序列,随后为另一个聚腺苷酸化信号,其中整个盒侧翼为loxP位点。

[0142] 一个示例人源化Cacng1基因座(例如,人源化的小鼠Cacng1基因座)是其中编码外显子1-4被侧翼是Neo自缺失盒的相应人序列替换的基因座。这些外显子编码Cacng1的编码结构域。将小鼠Cacng1的编码外显子1、内含子1、编码外显子2-4(和插入内含子)和3'非翻译区(UTR)的82bp的部分替换为相应的部分编码外显子1序列、内含子1、编码外显子2-4(和插入内含子)、完整的3'UTR和人CACNG1的3'UTR后另外的158bp(编码序列开始处的15bp仍然是小鼠序列)提供此类非人动物。参见图4B和图4C。

[0143] Cacng1^{hu/hu}小鼠

[0144] 本公开考虑了包含外源Cacng1基因座的细胞和非人动物。在一些实施方案中,包含异源Cacng1基因座的细胞或非人类动物可以表达异源CACNG1蛋白或嵌合CACNG1蛋白,其中天然Cacng1蛋白的一个或多个片段已被来自异源CACNG1序列的相应片段(例如,胞外结构域的全部或部分;CACNG1编码区的全部)替换。

[0145] 在一些实施方案中,本文公开的细胞和非人动物包含编码人CACNG1蛋白的氨基酸

1-10、氨基酸11-29、氨基酸30-108、氨基酸109-129、氨基酸130-134、氨基酸135-155、氨基酸156-179、氨基酸180-204、氨基酸205-222和/或它们的组合的外源核酸序列。

[0146] 功能丧失型Cacng1^{-/-}小鼠

[0147] 用基因编辑技术产生CACNG1^{-/-}小鼠,以确定CACNG1的缺失是否影响骨骼肌质量或功能。因为本文所述的一些非人动物缺少Cacng1基因座,所以此类非人动物可以整体方式提供对功能丧失对Cacng1蛋白的影响的了解。

[0148] 在一些实施方案中,本公开提供了包含内源Cacng1基因的敲除突变的非人动物、非人动物细胞或非人动物基因组。在一些实施方案中,此类敲除突变可以包括Cacng1基因或其部分的缺失。在一些情况下,敲除突变可以包括Cacng1基因的整个编码序列的缺失。在一些实施方案中,Cacng1^{-/-}人类动物基因组不表达任何CACNG1蛋白。

[0149] 在一些实施方案中,非人动物、非人动物细胞或非人动物基因组不表现出任何重大突变表型(即,不呈现出任何可测量的性状,特别是与野生型对应物相比具有统计学显著性的肌肉强度、结构或功能性状)。

[0150] C. 包含异源Cacng1基因座的非人细胞和非人动物

[0151] 提供了包含如本文所述的人源化Cacng1基因座的非人动物细胞和非人动物。细胞或非人动物对于人源化Cacng1基因座可以是杂合的或纯合的。二倍体生物在每个基因座处上具有两个等位基因。每对等位基因代表特定基因座的基因型。如果在特定基因座处存在两个相同的等位基因,则基因型被描述为纯合的;如果两个等位基因不同,则被描述为杂合的。

[0152] 本文提供的非人动物细胞可以是例如包含Cacng1基因座或与人CACNG1基因座同源或直系同源的基因组基因座的任何非人细胞。细胞可以是真核细胞,其包括例如真菌细胞(例如酵母)、植物细胞、动物细胞、哺乳动物细胞、非人哺乳动物细胞和人细胞。动物可以是例如哺乳动物、鱼或鸟。哺乳动物细胞可以是例如非人哺乳动物细胞、啮齿类动物细胞、大鼠细胞、小鼠细胞或仓鼠细胞。其它非人哺乳动物包括例如非人灵长类动物、猴、猿、猩猩、猫、犬、兔、马、公牛、鹿、野牛、家畜(例如牛族物种,例如母牛、阉牛等等;绵羊科物种,例如绵羊、山羊等等;以及猪科物种,例如猪、公猪)。鸟类包括例如鸡、火鸡、鸵鸟、鹅、鸭等等。还包括驯养动物和农业动物。术语“非人”排除人。

[0153] 细胞也可以是任何类型的未分化或分化状态。例如,细胞可以是全能细胞、多能细胞(例如人多能细胞或非人多能细胞,例如小鼠胚胎干(ES)细胞或大鼠ES细胞)、或非多能细胞。全能细胞包括可以产生任何细胞类型的未分化细胞,并且多能细胞包括具有发展成多于一个分化细胞类型的能力的未分化细胞。此类多能和/或全能细胞可以是例如ES细胞或ES样细胞,例如诱导性多能干(iPS)细胞。ES细胞包括胚胎来源的全能或多能细胞,它们在引入胚胎后可有助于发育中胚胎的任何组织。ES细胞可以源自胚泡的内细胞团,并且可以分化为脊椎动物三个胚层(内胚层、外胚层和中胚层)中任何胚层的细胞。

[0154] 本文提供的细胞也可以是生殖细胞(例如,精子或卵母细胞)。细胞可以是有丝分裂感受态细胞或无有丝分裂活性细胞,减数分裂感受态细胞或无减数分裂活性细胞。类似地,本文公开的细胞也可以是原代体细胞或非原代体细胞的细胞。体细胞包括并非配子、生殖细胞、配子体或未分化干细胞的任何细胞。例如,本文公开的细胞可以是肌细胞,诸如骨骼肌细胞。

[0155] 本文提供的合适的细胞还包括原代细胞。原代细胞包括已直接从生物、器官或组织中分离的细胞或细胞培养物。原代细胞包括既不转化也不是永生的细胞。原代细胞包括从生物体、器官或组织中获得的所有细胞,所述细胞先前未在组织培养中传代或先前已经在组织培养中传代但不能在组织培养中无限传代。此类细胞可以通过常规技术分离,包括例如肌细胞(例如,骨骼肌细胞)。

[0156] 本文提供的其它合适的细胞包括永生化细胞。永生化细胞包括来自多细胞生物的细胞,其通常不无限期地增殖,但由于突变或改变,已逃避正常的细胞衰老,而是可以保持经历分裂。此类突变或改变可以天然发生或被有意诱导。永生化细胞系的实例是肌纤维细胞系。永生化细胞或原代细胞包括可用于培养或表达重组基因或蛋白质的细胞。

[0157] 本文提供的细胞还包括单细胞期胚胎(即,受精卵母细胞或受精卵)。此类单细胞期胚胎可以来自任何遗传背景(例如来自小鼠的BALB/c、C57BL/6、129或其组合),可以是新鲜的或冷冻的,并且可以源自自然繁殖或体外受精。

[0158] 本文提供的细胞可以是正常的健康细胞,或者可以是患病细胞或携带突变的细胞。

[0159] 包含如本文所述的人源化Cacng1基因座的非人动物可通过本文别处所述的方法制备。动物可以是例如哺乳动物、鱼或鸟。非人哺乳动物包括例如非人灵长类动物、猴、猿、猩猩、猫、犬、兔、马、公牛、鹿、野牛、绵羊、兔、啮齿类动物(例如小鼠、大鼠、仓鼠和豚鼠)、以及家畜(例如牛族物种,例如母牛和公牛;绵羊科物种,例如绵羊和山羊;以及猪科物种,例如猪和公猪)。鸟类包括例如鸡、火鸡、鸵鸟、鹅和鸭。还包括驯养动物和农业动物。术语“非人动物”排除人。优选的非人动物包括例如啮齿类动物,例如小鼠和大鼠。

[0160] 非人动物可以来自任何遗传背景。例如,合适的小鼠可以来自129品系、C57BL/6品系、129和C57BL/6的混合物、BALB/c品系或Swiss Webster品系。129品系的例子包括129P1、129P2、129P3、129X1、129S1(例如129S1/SV、129S1/Sv1m)、129S2、129S4、129S5、129S9/SvEvH、129S6(129/SvEvTac)、129S7、129S8、129T1和129T2。参见例如,Festing等人(1999) Mammalian Genome 10:836,其以引用的方式整体并入本文用于所有目的。C57BL品系的例子包括C57BL/A、C57BL/An、C57BL/GrFa、C57BL/Kal_wN、C57BL/6、C57BL/6J、C57BL/6ByJ、C57BL/6NJ、C57BL/10、C57BL/10ScSn、C57BL/10Cr和C57BL/01a。合适的小鼠也可以来自上述129品系和上述C57BL/6品系的混合物(例如50% 129和50% C57BL/6)。同样地,合适的小鼠可以来自上述129品系的混合物或上述BL/6品系的混合物(例如129S6(129/SvEvTac)品系)。

[0161] 类似地,大鼠可以来自任何大鼠品系,包括例如ACI大鼠品系、Dark Agouti (DA) 大鼠品系、Wistar大鼠品系、LEA大鼠品系、Sprague Dawley (SD) 大鼠品系、或Fischer大鼠品系例如Fisher F344或Fisher F6。大鼠也可以得自源自上述两种或更多种品系的混合物的品系。例如,合适的大鼠可以来自DA品系或ACI品系。ACI大鼠品系表征为具有黑灰色,具有白色腹部和足部以及RT1^{av1}单体型。此类品系可得自各种来源包括Harlan Laboratories。Dark Agouti (DA) 大鼠品系表征为具有野灰色皮毛和RT1^{av1}单倍型。此类大鼠可得自各种来源包括Charles River和Harlan Laboratories。一些合适的大鼠可来自近交大鼠品系。参见例如,US2014/0235933,其以引用的方式整体并入本文用于所有目的。

[0162] III. 制备包含异源Cacng1基因座的非人动物的方法

[0163] 提供了各种方法用于制备包含如本文别处公开的异源Cacng1基因座的非人动物。用于产生遗传修饰的生物的任何方便的方法或方案都适合于产生此类遗传修饰的非人动物。参见例如,Cho等人(2009)Current Protocols in Cell Biology 42:19.11:19.11.1-19.11.22,以及Gama Sosa等人(2010)Brain Struct.Funct.214(2-3):91-109,其各自以引用的方式整体并入本文用于所有目的。此类经遗传修饰的非人动物可以例如通过在靶向Cacng1基因座处的基因敲入来产生。

[0164] 例如,产生包含人源化Cacng1基因座的非人动物的方法可以包括:(1)将多能细胞的基因组修饰为包含人源化Cacng1基因座;(2)鉴定或选择包含人源化Cacng1基因座的经遗传修饰的多能细胞;(3)在将经遗传修饰的多能细胞引入体外非人动物宿主胚胎细胞中;并且(4)在代孕母体中植入并孕育宿主胚胎细胞。任选地,可以将包含修饰的多能细胞(例如,非人ES细胞)的宿主胚胎孵育直至胚泡期,然后在代孕母体中植入并孕育,以产生F0非人动物。然后代孕母体可以产生包含人源化Cacng1基因座的F0代非人动物。

[0165] 该方法还可以包括鉴定具有修饰的靶基因组基因座的细胞或动物。各种方法可以用于鉴定具有靶向遗传修饰的细胞和动物。

[0166] 筛选步骤可以包括例如用于评价亲本染色体的等位基因修饰(MOA)的定量测定。例如,可以经由定量PCR,例如实时PCR(qPCR)来进行定量测定。实时PCR可以利用识别靶基因座的第一引物组,以及识别非靶向参考基因座的第二引物组。引物组可以包含识别扩增序列的荧光探针。

[0167] 合适的定量测定的其它例子包括荧光介导的原位杂交(FISH),比较基因组杂交,等温DNA扩增,与固定探针的定量杂交,INVADER[®]探针,TAQMAN[®]Molecular Beacon探针或ECLIPSE[™]探针技术(参见例如,US 2005/0144655,其以引用的方式整体并入本文用于所有目的)。

[0168] 合适的多能细胞的一个例子是胚胎干(ES)细胞(例如,小鼠ES细胞或大鼠ES细胞)。经修饰的多能细胞可以例如通过以下方式通过重组产生:(a)向细胞中引入一种或多种靶向载体,这些靶向载体包含侧翼是对应于5'和3'靶位点的5'和3'同源臂的插入核酸,其中插入核酸包含异源Cacng1基因座;并且(b)鉴定至少一种在其基因组中包含整合在靶基因组基因座处的插入核酸的细胞。可替代地,经修饰的多能细胞可以通过以下方式产生:(a)向细胞中引入:(i)核酸酶剂,其中该核酸酶剂在靶基因组基因座内的识别位点处产生缺口或双链断裂;和(ii)一种或多种靶向载体,这些靶向载体包含侧翼是对应于与识别位点足够接近地定位的5'和3'靶位点的5'和3'同源臂,其中插入核酸包含异源Cacng1基因座;并且(c)鉴定至少一种在靶基因组基因座处包含修饰(例如,插入核酸的整合)的细胞。可以使用在所需识别位点内诱导切口或双链断裂的任何核酸酶试剂。合适的核酸酶的例子包括转录活化因子样效应物核酸酶(TALEN),锌指核酸酶(ZFN),大范围核酸酶和成簇的规律间隔的短回文重复(CRISPR)/CRISPR相关的(Cas)系统或此类系统的组分(例如,CRISPR/Cas9)。参见例如,US 2013/0309670和US 2015/0159175,其各自以引用的方式整体并入本文用于所有目的。

[0169] 可以将供体细胞引入任何阶段(例如囊胚期或桑椹体前期(即,4细胞期或8细胞期))的宿主胚中。产生能够通过所述生殖系传递所述基因修饰的后代。参见例如,美国专利号7,294,754,其以引用的方式整体并入本文用于所有目的。

[0170] 可替代地,产生本文别处所述的非人动物的方法可包括:(1)使用上述用于修饰多能细胞的方法将单细胞期胚胎的基因组修饰为包含异源Cacng1基因座;(2)选择经遗传修饰的胚胎;并且(3)在代孕母体中植入并孕育经遗传修饰的胚胎。产生能够通过所述生殖系传递所述基因修饰的后代。

[0171] 也可以利用核转移技术产生非人哺乳动物。简言之,用于核转移的方法可以包括以下步骤:(1)将卵母细胞去核或提供去核的卵母细胞;(2)分离或提供待与去核的卵母细胞组合的供体细胞或核;(3)将细胞或核插入去核的卵母细胞内,以形成重建细胞;(4)将重建细胞植入动物的子宫内,以形成胚胎;并且(5)允许胚胎发育。在此类方法中,卵母细胞通常从已死亡的动物中回收,但它们也可以从活动物的输卵管和/或卵巢中分离出来。将供体细胞或核插入去核的卵母细胞中以形成重建细胞可以是将供体细胞显微注射于透明带下再融合来实现。融合可以通过跨越接触/融合平面施加DC电脉冲(电融合)、通过使细胞暴露于促进融合的化学物质(例如聚乙二醇)或借助于灭活病毒(例如仙台病毒)来诱导。重建的细胞可以在核供体和接受者卵母细胞融合之前、期间和/或之后,通过电学和/或非电学方式活化。活化方法包括电脉冲、化学上诱导的休克、利用精子穿透、增加卵母细胞中的二价阳离子含量,和减少卵母细胞中的细胞蛋白质磷酸化(如通过激酶抑制剂)。活化的重建细胞或胚胎可以在培养基中培养,然后转移到动物的子宫中。参见例如,US 2008/0092249、WO 1999/005266、US 2004/0177390、WO 2008/017234和美国专利号7,612,250,其各自以引用的方式整体并入本文用于所有目的。

[0172] 本文提供的各种方法允许产生经遗传修饰的非人F0动物,其中经遗传修饰的F0动物的细胞包含人源化Cacng1基因座。人们认识到,根据用于产生F0动物的方法,F0动物中具有异源Cacng1基因座的细胞数量会有所不同。经由例如VELOCIMOUSE®方法,将供体ES细胞从相应的生物(例如8细胞期的小鼠胚胎)引入桑椹体前期胚胎,允许F0动物的更大百分比的细胞群体包含具有目的核苷酸序列的细胞,所述目的核苷酸序列包含靶向遗传修饰。例如,非人F0动物的至少50%、60%、65%、70%、75%、85%、86%、87%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的细胞贡献可以包含具有靶向修饰的细胞群体。

[0173] 经遗传修饰的F0动物的细胞对于异源Cacng1基因座可以是杂合的,或者对于异源Cacng1基因座可以是纯合的。

[0174] 上文或下文引用的所有专利文件、网站、其它出版物、登录号等等都以引用的方式整体并入用于所有目的,其程度与每个个别项目具体地且个别地指出以引用的方式如此并入相同。如果序列的不同版本在不同时间与一个登录号相关联,则意指在本专利申请的有效提交日期与该登录号相关联的版本。有效提交日期意指实际提交日期,或适用的提及登录号的优先权申请的提交日期中的较早者。同样地,如果出版物的不同版本、网站等等在不同时间公布,则除非另有说明,否则意指在专利申请的有效提交日期最近公布的版本。除非另外明确说明,否则本发明的任何特征、步骤、要素、实施方案或实施方案可与任何其他特征、步骤、元素、实施方案或实施方案组合使用。尽管本发明已通过说明和实例略微详细地进行描述用于清楚和理解的目的,但显而易见的是可以在所附权利要求的范围内实践某些改变和修改。

[0175] 在一些实施方案中,本公开提供了一种制备本文所述的非人动物、非人动物细胞

或非人动物基因组的方法,该方法包括将编码异源CACNG1蛋白或其部分的核酸序列插入非人动物的基因组、非人动物细胞的基因组或非人动物基因组中。

[0176] 多种嵌合核酸可专门用于此类目的。在一些实施方案中,在本文所述的细胞或基因组的基因编辑中可以使用编码功能性CACNG1蛋白的,包含经修饰的非人动物Cacng1基因的核酸序列的嵌合核酸分子,其中经修饰的非人动物Cacng1基因包含编码非人动物CACNG1蛋白的一部分的核酸序列被编码异源CACNG1蛋白或其部分的同源核酸序列的替换。在一些情况下,此类嵌合核酸分子包含:(i) 包含人CACNG1基因的外显子1或其部分的核酸序列;(ii) 人CACNG1基因的内含子1或其部分的核酸序列;(iii) 包含人CACNG1基因的外显子2或其部分的核酸序列;(iv) 人CACNG1基因的内含子2或其部分的核酸序列;(v) 包含人CACNG1基因的外显子3或其部分的核酸序列;(vi) 人CACNG1基因的内含子3或其部分的核酸序列;(v) 包含人CACNG1基因的外显子4或其部分的核酸序列;(vii) 人CACNG1基因的3' 非翻译区(UTR)的核酸序列;或(v) (i) - (iv)的任何组合。在一些情况下,该嵌合分子提供药物选择盒。

[0177] 在一些实施方案中,本文所述的嵌合核酸分子包含(i) 经修饰的非人动物Cacng1基因上游的5' 同源臂和(ii) 经修饰的非人动物Cacng1基因下游的3' 同源臂。在一些实施方案中,5' 同源臂和3' 同源臂构造成与感兴趣的非人动物Cacng1基因座进行同源重组,并且在与感兴趣的非人动物Cacng1基因座同源重组后,经修饰的Cacng1基因替换感兴趣的非人动物Cacng1基因座处的非人动物Cacng1基因,并与驱动感兴趣的非人动物Cacng1基因座处的非人动物Cacng1基因表达的内源启动子可操作地连接。在一些实施方案中,嵌合核酸分子包含:(i) 包含如SEQ ID NO:25所示的核酸序列的5' 同源臂和/或(ii) 包含如SEQ ID NO:26所示的核酸序列的3' 同源臂。在一些实施方案中,嵌合核酸分子包含的核酸序列包含如SEQ ID NO:6所示的核酸序列。

[0178] 序列简述

[0179] 所附序列列表中列出的核苷酸和氨基酸序列使用用于核苷酸碱基的标准字母缩写和用于氨基酸的三字母代码显示。核苷酸序列遵循从序列的5' 末端开始并继续前进(即,每行从左到右)到3' 末端的标准惯例。每个核苷酸序列仅显示一条链,但互补链应理解为通过对所展示链的任何引用包括。氨基酸序列遵循从序列的氨基末端开始并继续前进(即,每行从左到右)到羧基末端的标准惯例。

[0180] 表1. 序列描述.

[0181]

SEQ ID NO	类型	描述
1	DNA	m <i>Cacng1</i> -编码序列-NCBI 基因 ID 12299
2	蛋白	mCACNG1 – UniProt ID O70578
3	DNA	h <i>CACNG1</i> 编码序列-NCBI 基因 ID 786
4	蛋白	hCACNG1 – UniProt ID O70578
5	DNA	小鼠/人 CACNG1 cds 的人类部分
6	DNA	小鼠/人 CACNG1 cds
7	蛋白	人 CYT1 的氨基酸序列
8	DNA	人 CYT1 的核酸序列
9	蛋白	人 TM1 的氨基酸序列
10	DNA	人 TM1 的核酸序列
11	蛋白	人 EX1 的氨基酸序列
12	DNA	人 EX1 的核酸序列
13	蛋白	人 TM2 的氨基酸序列
14	DNA	人 TM2 的核酸序列
15	蛋白	人 CYT2 的氨基酸序列
16	DNA	人 CYT2 的核酸序列
17	蛋白	人 TM3 的氨基酸序列
18	DNA	人 TM3 的核酸序列
19	蛋白	人 EX2 的氨基酸序列
20	DNA	人 EX2 的核酸序列
21	蛋白	人 TM4 的氨基酸序列
22	DNA	人 TM4 的核酸序列
23	蛋白	人 CYT3 的氨基酸序列
24	DNA	人 CYT3 的核酸序列

[0182]

SEQ ID NO	类型	描述
25	DNA	小鼠 5'臂
26	DNA	小鼠 3'臂
27	DNA	7450 等位基因
28	DNA	7451 等位基因
29	DNA	m <i>Cacng1</i> 基因座敲除等位基因 (6866)
30	DNA	hCACNG1: 转发-GGCGAGAGCTCGGAGATC
31	DNA	hCACNG1: rev-GGCTGCCAGGATGATGAAG
32	DNA	hCACNG1: 探针-TCGAATTCACCACTCAGAAGGAGTACA
33	DNA	mCACNG1: 正向-CCGTGCACAACAAAGACAAGAG
34	DNA	mCACNG1: 反向-GGCTGCCAGGATGATGAAG
35	DNA	mCACNG1: 探针-TGTGAGCACGTCACACCATCAGG

[0183] 实施例

[0184] 实施例1: CACNG1在骨骼肌中特异性表达

[0185] 骨骼肌二氢吡啶受体 (DHPR) 是一种参与兴奋-收缩耦联的L型钙通道。骨骼肌DHPR由5个亚基组成,其中 $\alpha 1s$ 亚基经由与兰尼碱受体 (ryanodine receptor) 的物理相互作用调节肌浆网的钙释放,在肌肉收缩中发挥关键作用。发现 $\gamma 1$ 亚基 (CACNG1) 在骨骼肌中高度特异性表达 (图1A)。因此,利用多种不同的方法产生人源化*Cacng1*小鼠,用于验证不同治疗剂的肝脏特异性递送。

[0186] 实施例2. CACNG1敲除小鼠 (CACNG1^{-/-}) 的产生和分析

[0187] 在内部产生并繁育CACNG1^{-/-}小鼠,以确定CACNG1的缺失是否影响骨骼肌质量或功能。Cacng1消融结构设计如下。对含有Cacng1基因组序列的细菌人工染色体进行修饰,使得处于人UBC (泛素) 启动子的控制下含有新霉素抗性基因的floxed lacZ报道基因盒替换Cacng1编码外显子1正好从起始ATG之后开始的224bp。克隆该盒,使得lacZ编码序列与起始ATG同框,并且Cacng1编码外显子1的3' 5bp保留在该盒之后。(参见,图2) 该构建体电穿孔到100% C57B1/6NTac胚胎干细胞中。通过TaqMan分析鉴定成功靶向的克隆。使用VelociGene©方法产生Cacng1^{-/+}小鼠 (Valenzuela 2003 Nat Biotech PMID:12730667; Poueymirou 2007 Nat Biotech PMID:17187059) 并根据需要繁育至纯合 (CACNG1^{-/-})。在F0种系中使用自缺失技术去除抗性盒。

[0188] 将来自成年 (5-7个月龄) 雄性WT和CACNG1^{-/-}小鼠的肌肉组织仔细解剖并称重以评估肌肉质量,并对分离的趾长伸肌 (EDL) 进行离体收缩性测量以评估肌肉功能。使用Aurora Scientific 1300A设备进行肌肉收缩性测量。小心地切除EDL肌肉,并通过缝线使其附着到肌肉生理装置上。然后在用95%O₂/5%CO₂充氧的Krebs-Henseleit缓冲液中以最佳长度平衡肌肉,随后用超最大双相电流刺激以引发颤搐反应。颤搐刺激后,通过40Hz、60Hz、80Hz、100Hz和125Hz的刺激生成力-频率强直曲线,刺激之间休息2分钟。在100Hz下记录到最大力的产生。CACNG1似乎不会在调节肌肉功能中发挥主要作用,因为WT与CACNG^{-/-}型小鼠的颤搐力和强直力相似。参见图3A和图3B。

[0189] 实施例3.CACNG1人源化小鼠 (CACNG1^{hu/hu}) 的产生和分析

[0190] Cacng1靶向构建体设计如下。含有完整小鼠Cacng1基因组序列的细菌人工染色体经修饰以使Cacng1基因座人源化。小鼠Cacng1的编码外显子1、内含子1、编码外显子2-4 (和插入内含子) 和3' 非翻译区 (UTR) 的82bp的部分被人CACNG1序列替换,该人CACNG1序列由编码外显子1序列减去前15bp (该起始序列仍为小鼠的)、内含子1、编码外显子2-4 (和插入内含子)、完整的3' UTR和在入CACNG1的3' UTR后的另外158bp组成。参见图2A-2C。将自缺失新霉素抗性盒插入入序列的下游,随后是小鼠3' UTR的剩余部分。参见图2A-2C,例示了自缺失新霉素抗性盒缺失后的靶位点。然后将该靶向载体电穿孔到50% C57B1/6NTac/50% 129SvEvTac胚胎干细胞系中。通过TaqMan分析鉴定成功靶向的克隆。使用VelociGene©方法产生Cacng1^{+/+}小鼠 (Valenzuela 2003 Nat Biotech PMID:12730667; Poueymirou 2007 Nat Biotech PMID:17187059) 并根据需要与C57B1/6NTac回交。在F0雄性种系中使用自缺失技术去除抗生素抗性盒。

[0191] 基因表达分析

[0192] 经由TRIzol匀化和氯仿相分离从组织中分离总RNA,随后用用于微阵列总RNA分离试剂盒的MagMAX-96进行纯化。使用无RNA酶的DNA酶套组去除基因组DNA,并使用SuperScript VILO主要混合物将mRNA逆转录成cDNA。使用12K Flex系统用SensiFAST探针Lo-Rox扩增cDNA。使用Taqman基因表达测定法确定相对于m18S (内源性对照) 的人 (h) 和小鼠 (m) CACNG1表达,并使用比较CT方法 ($\Delta\Delta Ct$) 分析数据。Taqman引物/探针序列如下: hCACNG1: 正向-GGCGAGAGCTCGGAGATC (SEQ ID NO:30), 反向-GGCTGCCAGGATGATGAAG (SEQ ID NO:31), 探针-TCGAATTCACCACTCAGAAGGAGTACA (SEQ ID NO:32); mCACNG1: 正向-CCGTGCACAACAAAGACAAGAG (SEQ ID NO:33), 反向-GCTCTCCCCTGGGTTGAAG (SEQ ID NO:34), 探针-TGTGAGCACGTCACACCATCAGG (SEQ ID NO:35)。图3A是证明通过qPCR在CACNG1^{hu/hu}小鼠

肌肉中无法检测到小鼠CACNG1 (mCACNG1) 表达(左图), 而人CACNG1 (hCACNG1) 在CACNG1^{hu/hu}中表达, 但在WT小鼠肌肉中不表达(右图)的图表。

[0193] 单肌纤维分离和活染色

[0194] 从成年雄性CACNG1^{hu/hu}小鼠的腓肠肌中分离出单肌纤维。小心地切除肌肉并在DMEM中用700U/mL胶原酶消化60分钟。用火焰抛光的玻璃巴斯德移液管(Pasteur pipette)分离单肌纤维, 并且在几轮消化和洗涤后, 将肌纤维接种在低粘附性组织培养板中过夜。第二天早上, 在30分钟或4小时内, 以100nM的浓度将人特异性、Alexa 647缀合的CACNG1抗体添加到活的单肌纤维中。然后用DMEM洗涤肌纤维, 在4% PFA中固定, 并进行DAPI染色。然后将单纤维转移到显微镜载玻片上, 用Fluoromount封片并用LSM880共聚焦显微镜成像。参见图3B。该实验证明用100nM的Alexa 647缀合的人特异性 α -CACNG1 Ab对单个骨骼肌纤维的活染色, 显示出与从CACNG1^{hu/hu}小鼠分离的肌纤维结合, 但不与从WT小鼠分离的肌纤维结合。

[0195] 抗体分布的低温荧光断层扫描(CryoFT)

[0196] 成年雄性CACNG1^{hu/hu}小鼠尾静脉注射10mg/kg人特异性、Alexa 647缀合的CACNG1抗体、Alexa 647缀合的同型对照抗体或盐水。注射后六天, 经由CO₂对小鼠实施安乐死, 整体冷冻, 并使用CryoFT处理和成像进行评估。参见图3C。注射10mg/kg Alexa 647缀合的人特异性 α -CACNG1 Ab的CACNG1^{hu/hu}小鼠的图像, 显示与同型对照Ab相比, 在注射后6天对骨骼肌具有高特异性。

[0197] 抗体分布的免疫荧光成像

[0198] 成年雄性CACNG1^{hu/hu}小鼠皮下注射10mg/kg人特异性、Alexa 647缀合的CACNG1抗体、Alexa 647缀合的同型对照抗体或盐水。注射后六天, 小鼠用PBS经心脏灌注, 并将腓肠肌/跖肌/比目鱼肌复合体浸没在OCT包埋介质中, 并在液氮冷却的异戊烷中冷冻。将组织冷冻切片至12 μ m厚, 随后用4% PFA固定并进行层粘连蛋白和DAPI染色。将载玻片用Fluoromount封片并用Axioscan载玻片扫描仪成像。参见图3D。上图展示来自体内注射的Ab的内源性Alexa 647信号, 下图展示Alexa647-Ab与层粘连蛋白结合和DAPI共染色的叠加, 以使肌肉形态可视化。

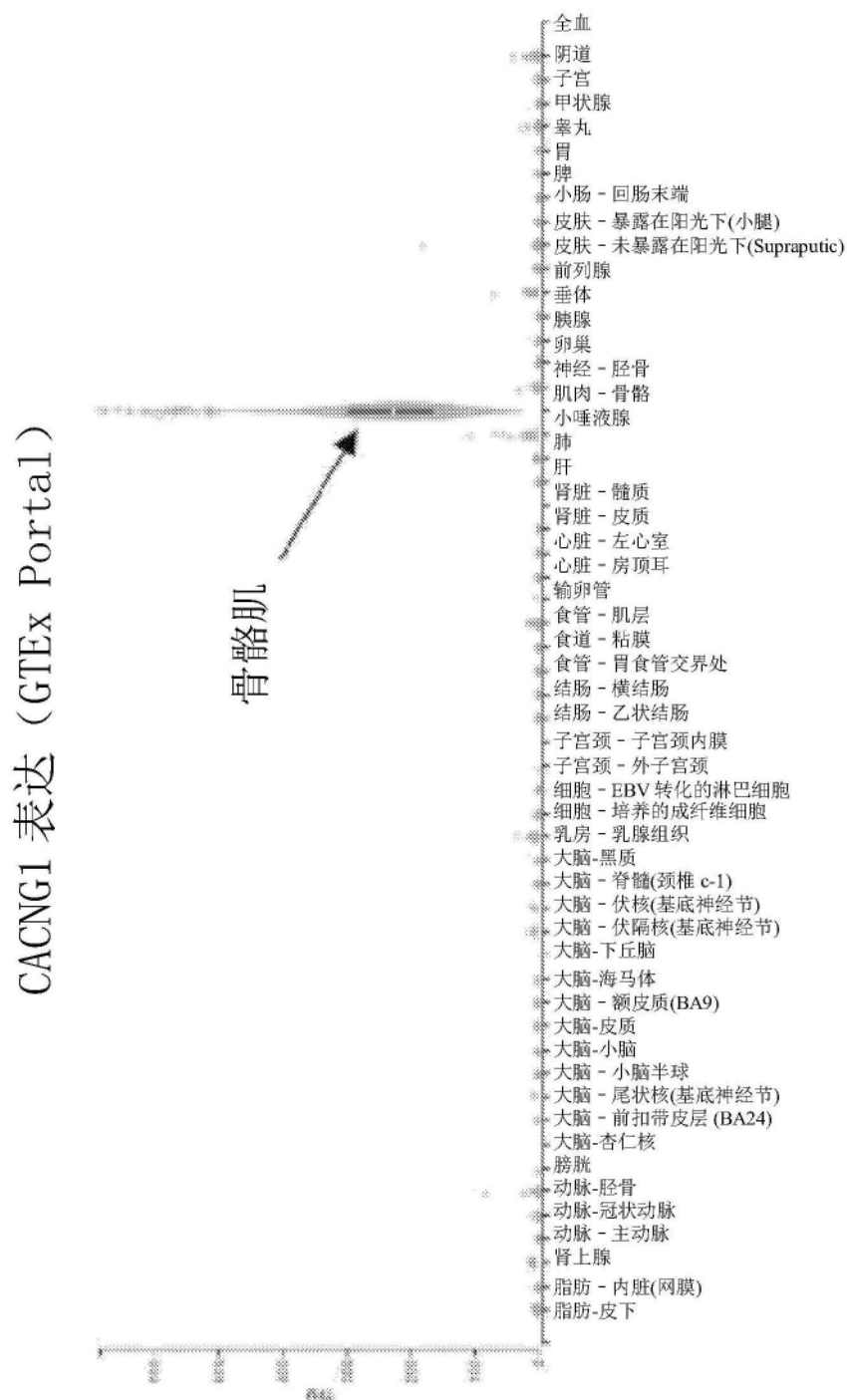


图1

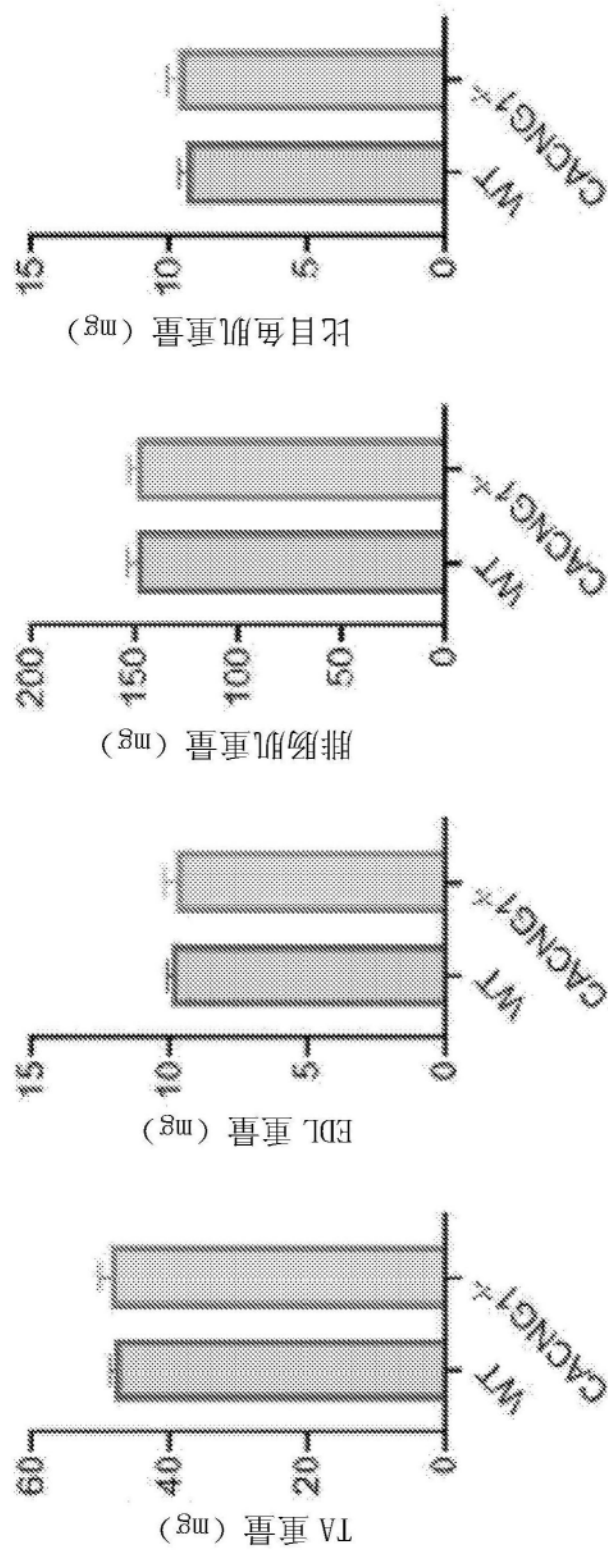


图3A

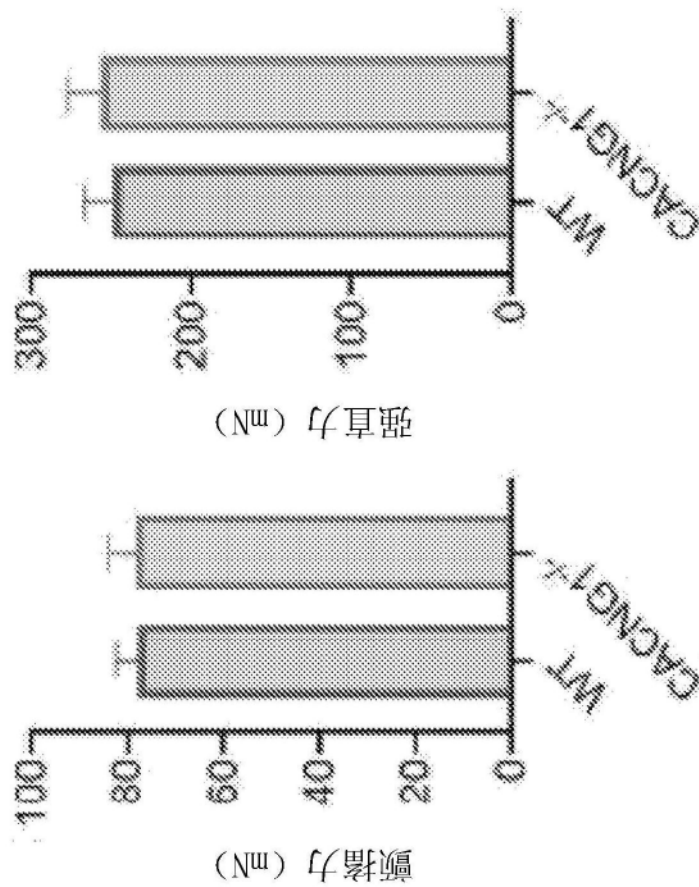


图3B

	正式符号	NCBI 基因 ID	主要来源	参考序列 mRNA ID	UniProt ID	基因组组装	位置
小鼠	<i>Cacng1</i>	12299	MG1:1206582	NM_007582.2	Q70578	GRC38/mm10	chr11:107,703,218-107,716,522 (-)
人	<i>CACNG1</i>	786	HGNC:1405	NM_000727.4	Q06432	GRCh38/hg38	chr17:67,044,554-67,056,797 (+)

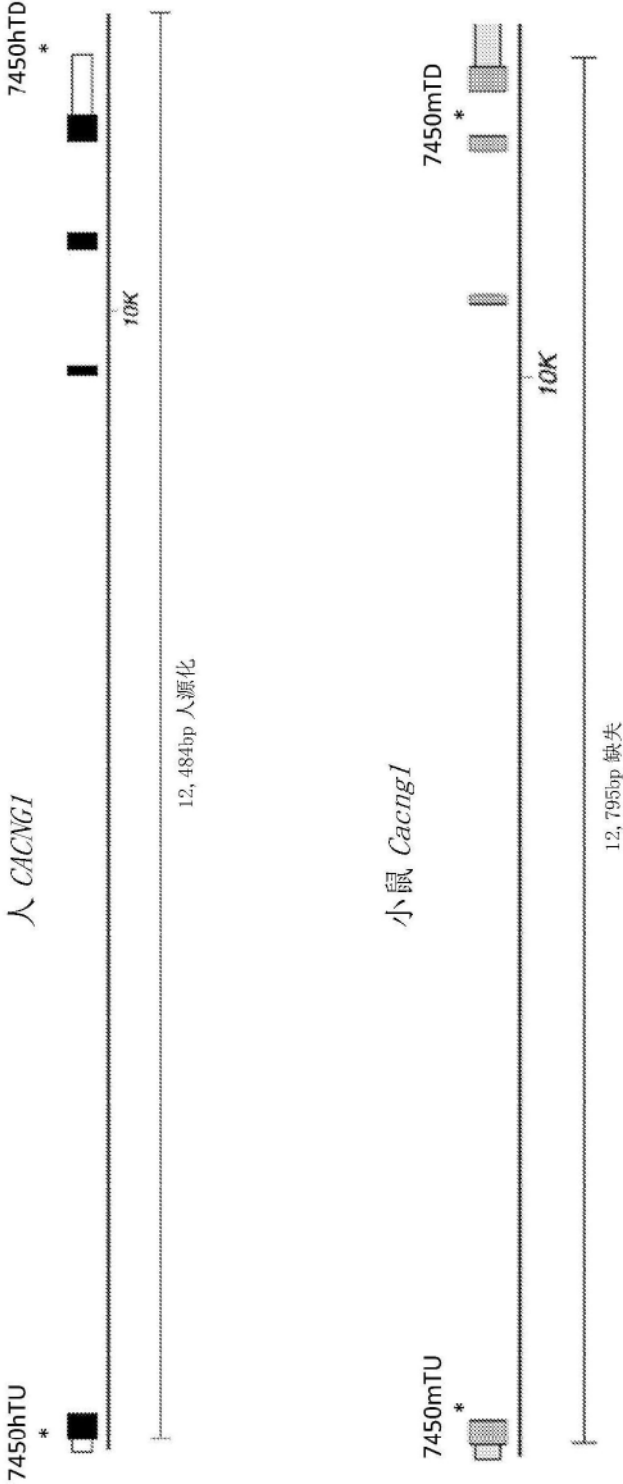


图4A

7450 等位基因 *CACNG1* 人源化，含有 Neo 自缺失盒

描述	将小鼠 <i>Cacng1</i> 的编码外显子 1、内含子 1、编码外显子 2-4（和插入内含子）和 3' 非翻译区（UTR）的 82bp 的部分替换为相应的部分编码外显子 1 序列、内含子 1、编码外显子 2-4（和插入内含子）、完整的 3' UTR 和人 <i>CACNG1</i> 的 3' UTR 后另外的 158 bp。编码序列开始处的 15bp 仍然是小鼠序列。 loxP-mPrm1-Crei-pA-hUbl-em7-Neo-pA-loxP 盒（4,805 bp）在人序列的下游，随后是小鼠 3' UTR 的剩余部分。
大小	12,795 bp 小鼠序列被 12,484 bp 的人序列替换

保留小鼠编码序列
的 15bp 起点。

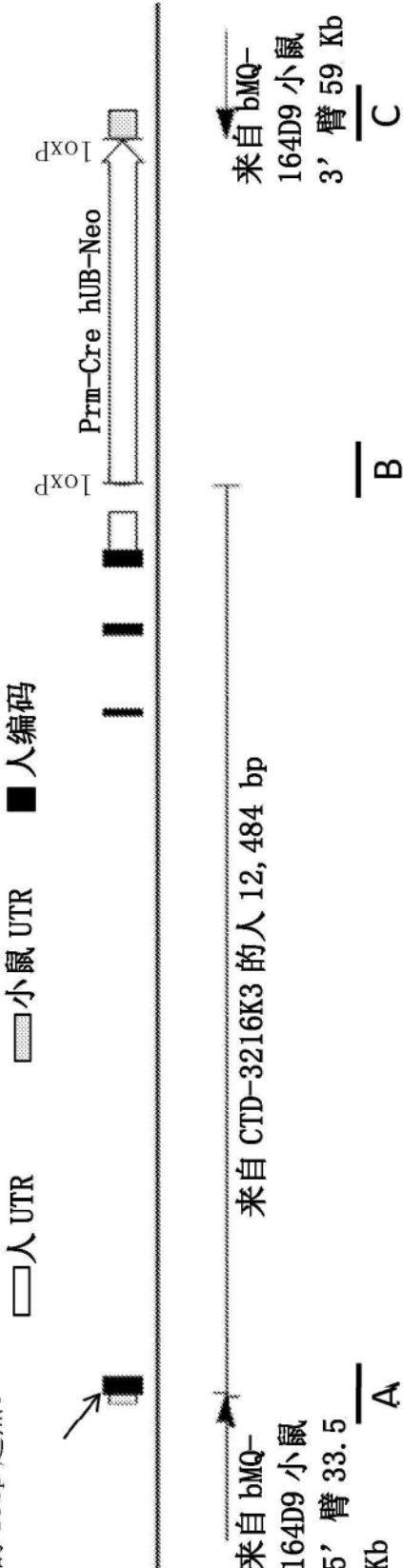


图4B

7451 等位基因 *CACNG1* 人源化，盒缺失，LoxP 位点保留

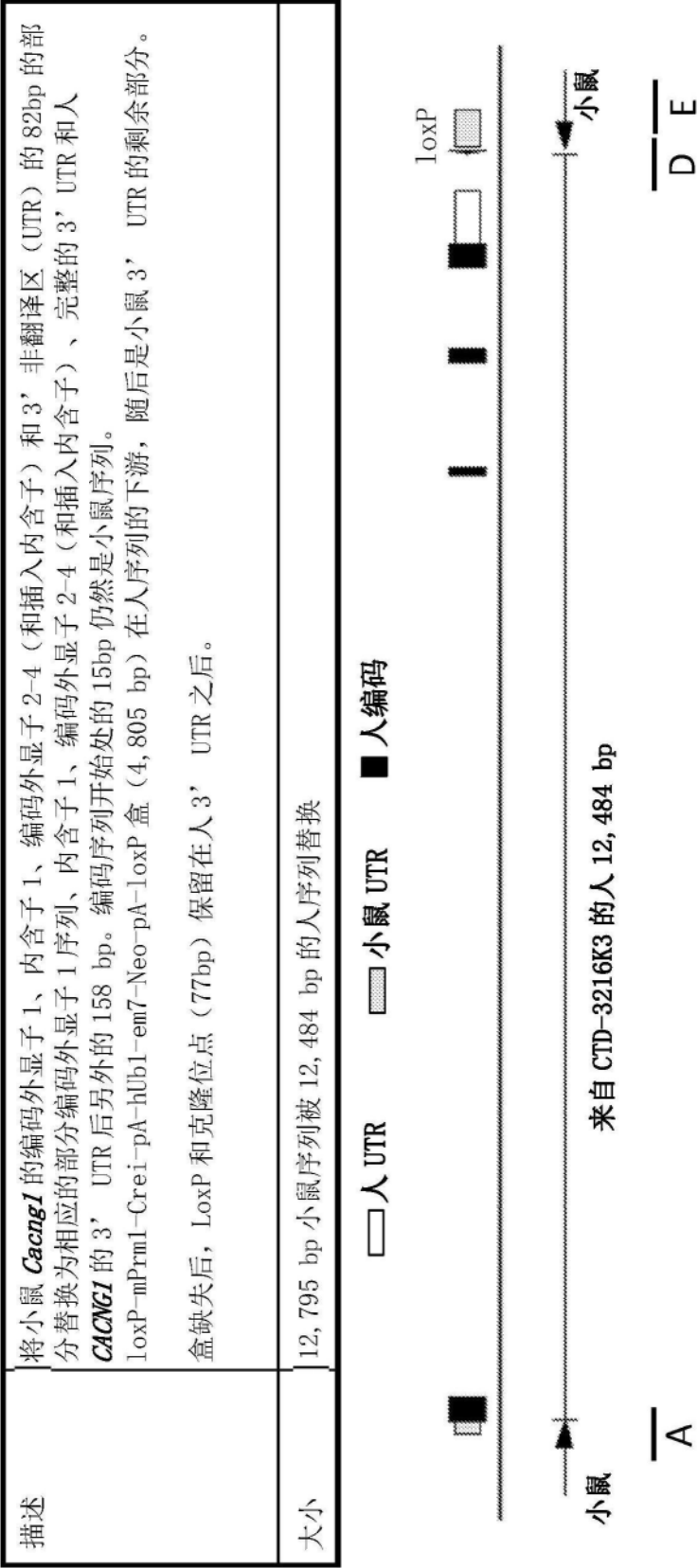


图4C

[illegible]

图5

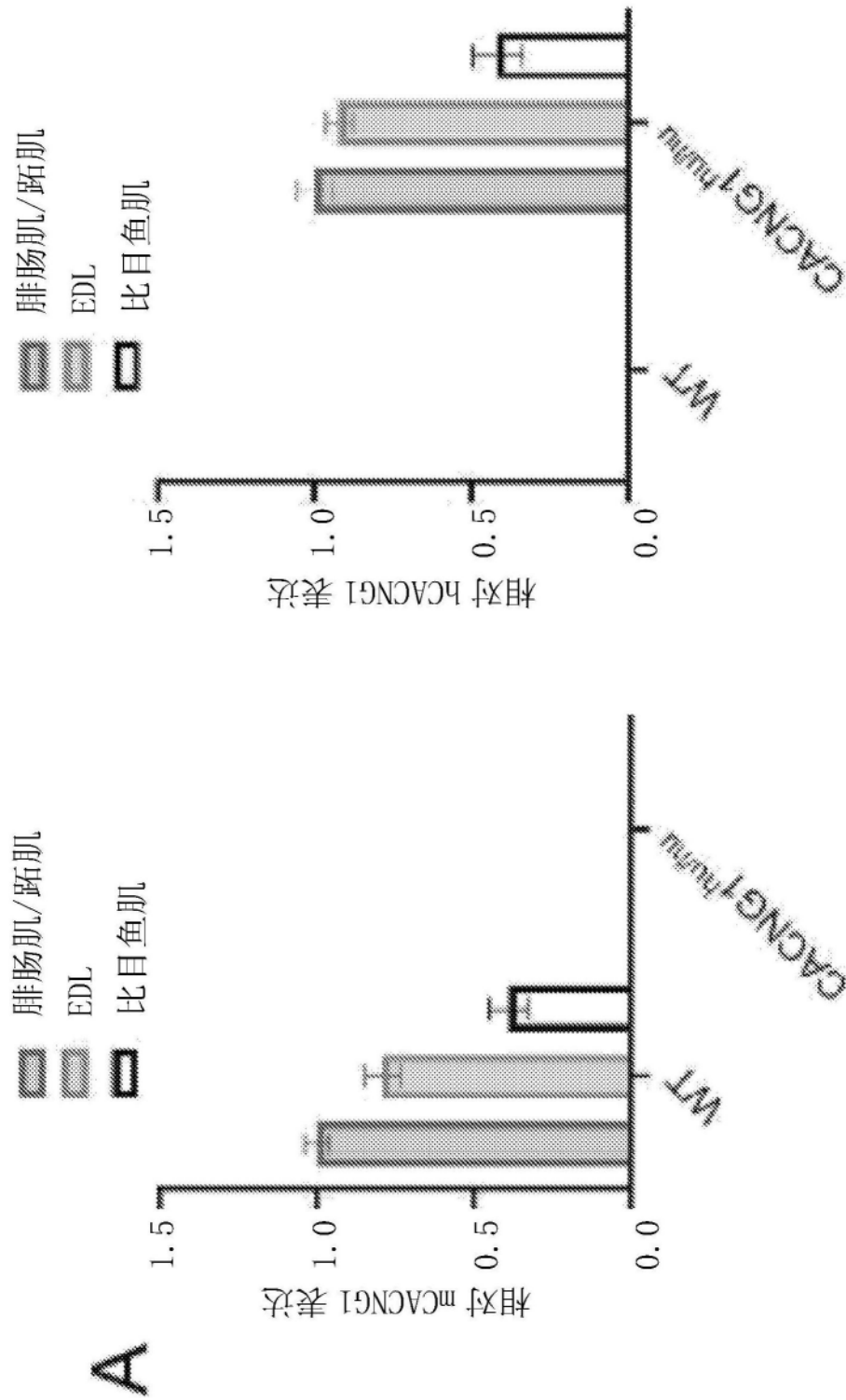


图6A

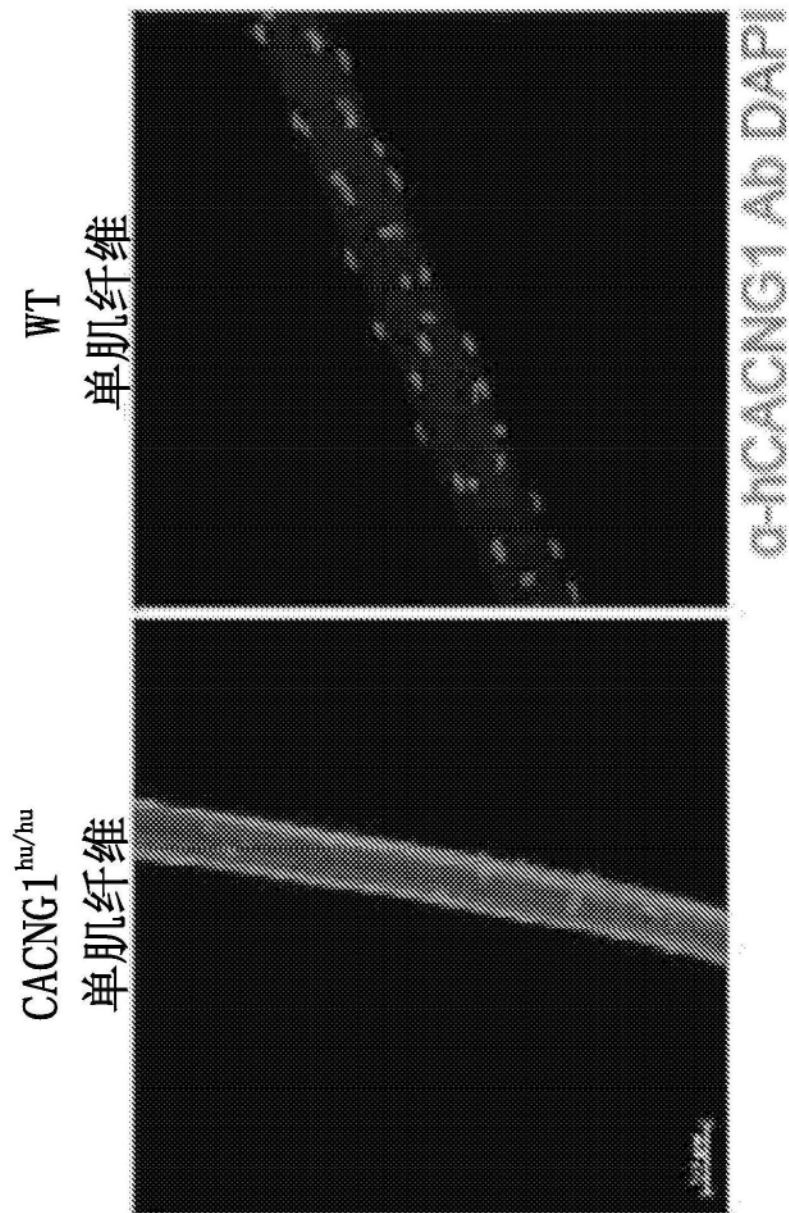


图6B

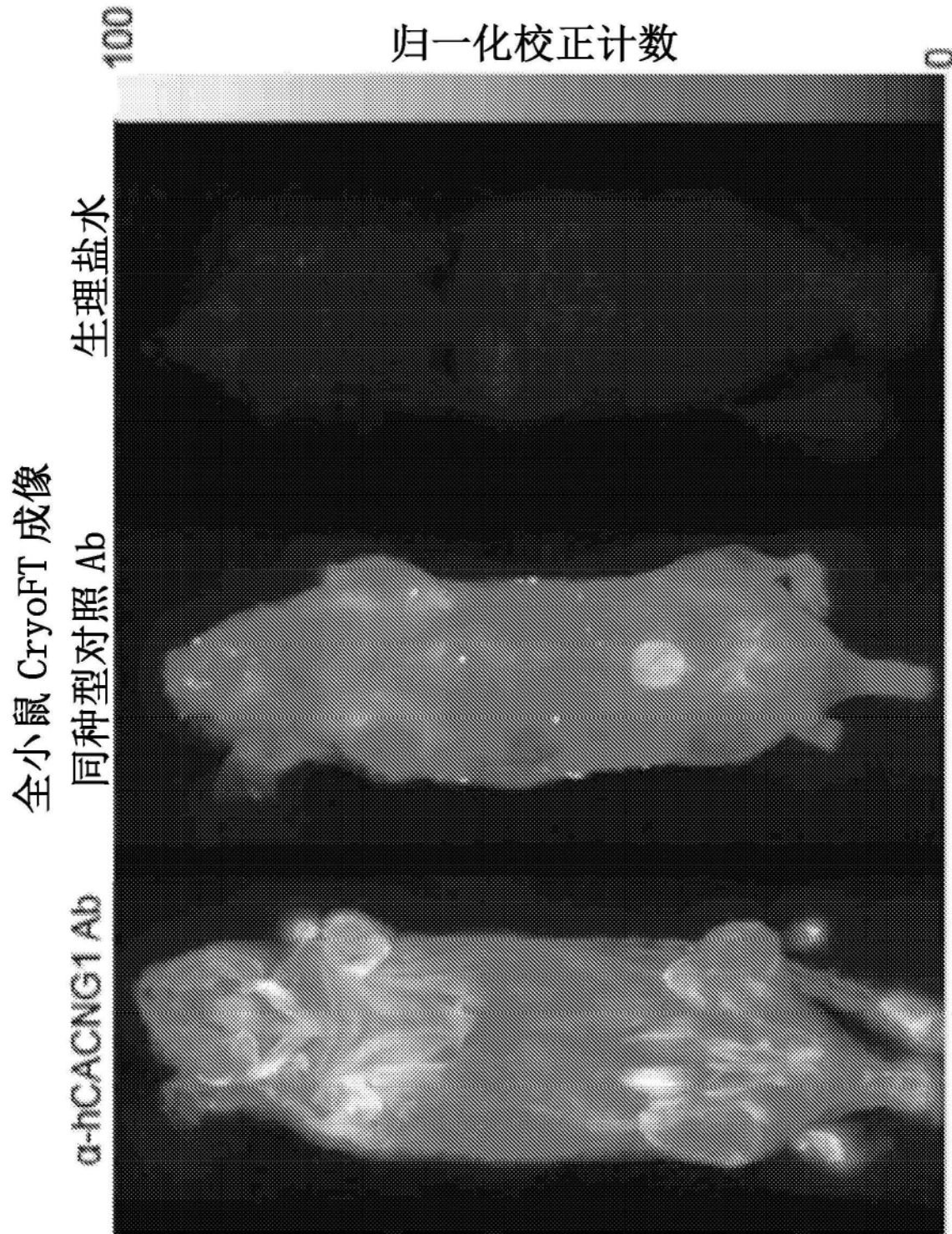


图6C

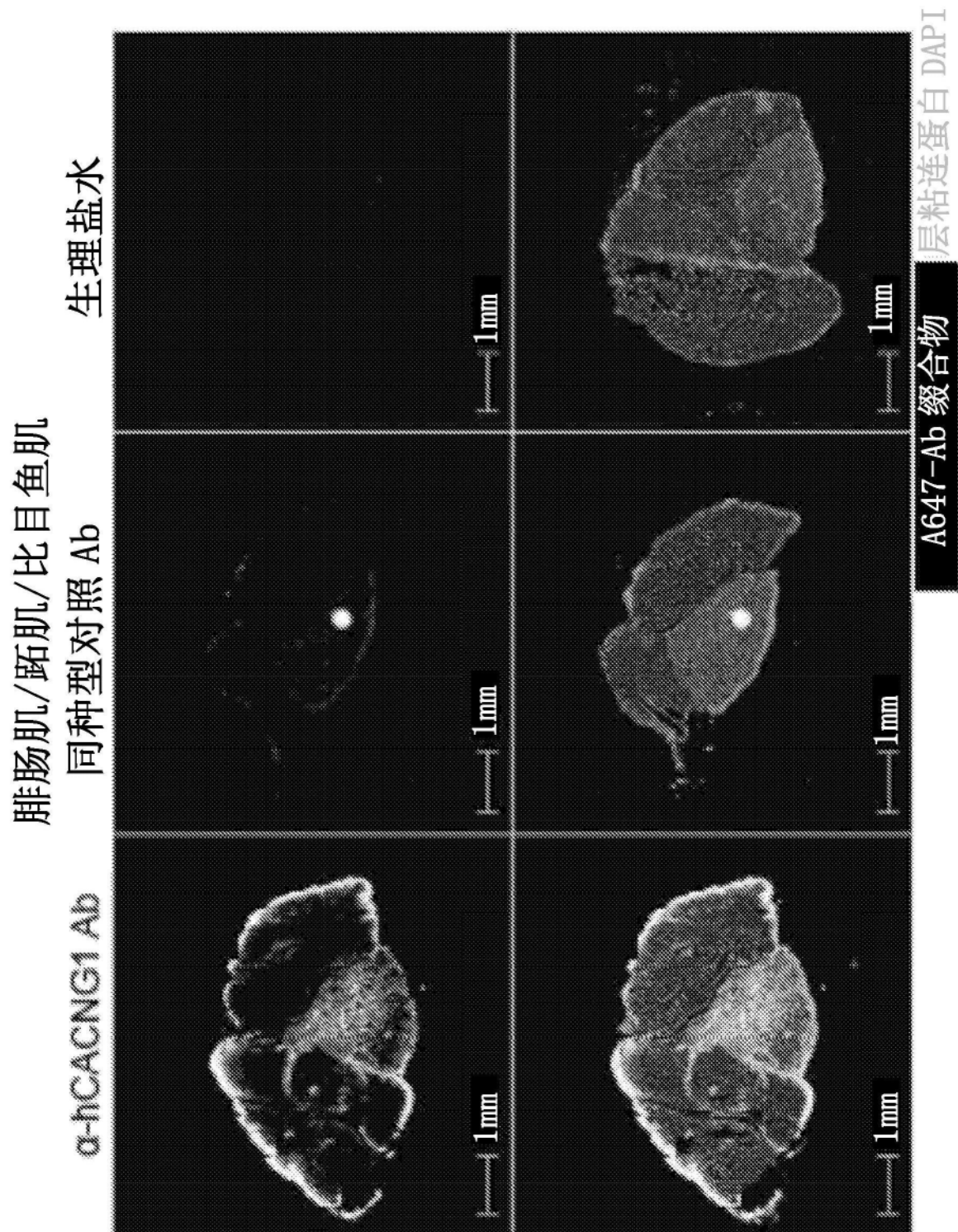


图7

人 Cacng1 蛋白 (070578; NP_031608.1) :

- 胞质结构域: 1-10
- 跨膜结构域: 11-29
- 胞外结构域: 30-109
- 跨膜结构域 2: 110-130
- 胞质结构域: 131-135
- 跨膜结构域 3: 136-156
- 胞外结构域: 157-180
- 跨膜结构域 4: 181-205
- 胞质结构域: 206-223

MSQTKTAKVRVTLFFILVGGVLAMVAVVTDHWAVLSPHLEHHNETCEAAHFGLWRICTARVAVHNKDKSCEHVT
P
SGEKNCSYFRHFNPGESSEIFEFTTQKEYSISAAAIATFSLGFIIVGSICAFISFGNKRDIYLLRPASMFYAFAGL
CLIVSVEVMRQSVKRMIDSEDTWIEHYYSWSFACACAAFIILFLGGLFLLFSLPRMPQNPWESCMDAEEPEH

图8A

小鼠 Cacng1 CDS:

- 胞质结构域: 1-30
- 跨膜结构域: 31-87
- 胞外结构域: 88-327
- 跨膜结构域 2: 328-390
- 胞质结构域: 391-405
- 跨膜结构域 3: 406-468
- 胞外结构域: 469-540
- 跨膜结构域 4: 541-615
- 胞质结构域: 616-669

ATGTCACAGACCAAAACAGCGAAGGTTTCGTGTGACCCCTCTTCTTCATCCTGGTGGGGGGGTGCTCGCCATGGTGGCCGGTGGTGAACCACT
GGGCCGTGCTGAGTCCACACCTGGAGCACCAACAATGAAACGTGCGAGGCAGCCCACTTTGGCCCTCTGGAGGATCTGCACCGCTCGGGTTGCCGT
GCACAAACAAGACAAGAGTTGTGAGCACGTCACACCATCAGGGGAAAAGAACTGCTCCTACTTCAGGCACCTTCAACCCAGGGGAGAGCTCGGAA
ATCTTTGAATTCAACCACTCAAAAGGAGTACAGCATCTCAGCAGCGGCCATTGCCATCTTCAGCCCTCGGCTTCATCATTTGTTGGTTCCATCTGCG
CATTTCTGTCCTTCGGGAATAAGCGTGATTACCTGCTGAGGCCAGCATCCATGTTTATGCTTCGAGGGCTCTGCCCTCATCGTCTCCGTGGA
GGTCATGAGGCAGTCSGTGAAGCGTATGATTGACAGCGAGGACACGGTCTGGATAGAGCACTACTATTCTGTGGTCTTTTCGCCCTGTGCATGTGCC
GCGTTCAATCTTGCTCTTCCTCGGTGGGCTGTTCCCTCCCTGCTCTTCTCCCTGCCCTCGGATGCCTCAGAACCCCTGGGAATCCTGCATGGACGCTG
AGCCAGAGCAC

图8B

人 CACNG1 蛋白 (Q06432; NP_000718.1) :

- 胞质结构域: 1-10
- 跨膜结构域: 11-29
- 胞外结构域: 30-108
- 跨膜结构域 2: 109-129
- 胞质结构域: 130-134
- 跨膜结构域 3: 135-155
- 胞外结构域: 156-179
- 跨膜结构域 4: 180-204
- 胞质结构域: 205-222

MSQTKMLKVRVTLFCILAGIVLAMTAVVTDHWAVLSPHMEHHNTTCEAAHFGLWRICTKRI PMDDSKTCGPTTLPGEKNC
SYFRHFNPGESSEIFEFTTQKEYSISAAAIAIFSLGFIIILGSLCVLISLGKKRDYLLRPASMFYAFAGLCILVSV EVMRQ
SVKRMIDSEDTVWIEYYYSWSEFACACAAFI LLFLGGLALLFSLPRMPRNPWESCMDAEPEH

图9A

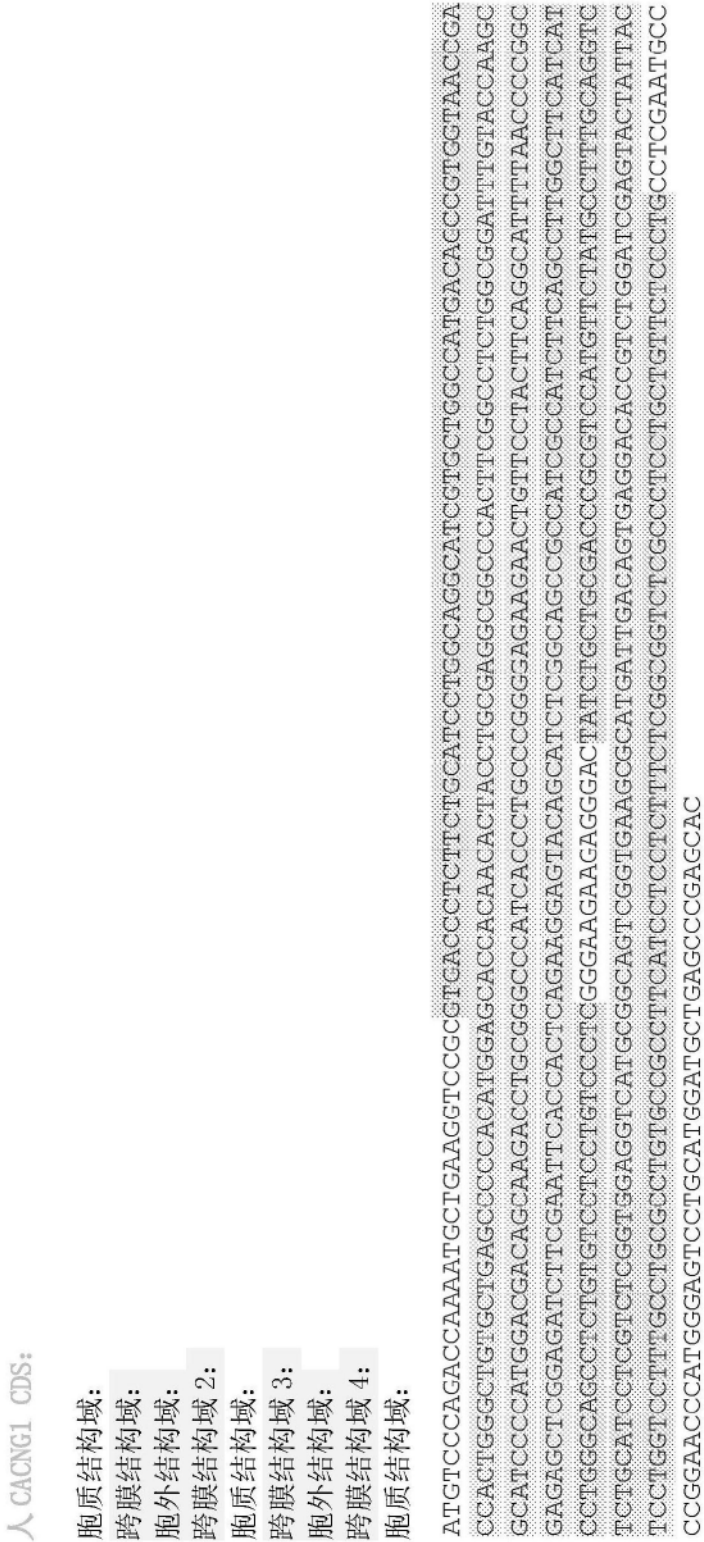


图9B

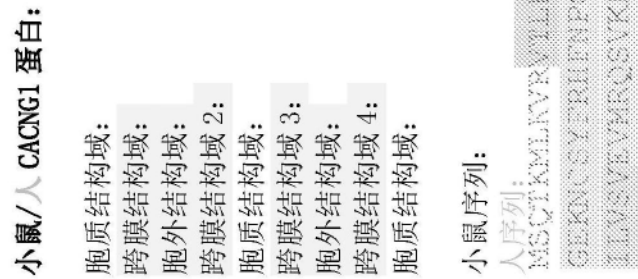


图10A

[illegible]

7450 等位基因 带有 Neo 自缺失盒的 *CACNG1* 人源化区域=小鼠 (小写) 人 Xho1 LoxP Prrm Crei sv40 polyA (小写) hUbi-em7 (小写) -
NEO-PGK polyA LoxP ICeU1 小鼠 (小写)

图11

7450 等位基因 带有 Neo 自缺失盒的 *CACNG1* 人源化区域=小鼠 (小写) 人 XhoI LoxP Prrm Crei sv40 polyA (小写) hUbi-em7 (小写) -
NEO-PGK polyA LoxP ICeU1 小鼠 (小写)

58

7450 等位基因 带有 Neo 自缺失盒的 *CACNG1* 人源化区域=小鼠 (小写) 人 Xho1 LoxP Prrm Crei sv40 polyA (小写) hUbi-em7 (小写) - NEO-PGK polyA LoxP ICeU1 小鼠 (小写)

gcccaggagctcagagg

TAACCTATAAGCGGTCCTCAGGTAGGGATgctcctctacaaacagtgcacatggacacaaacgagacc

7451 等位基因盒缺失的 *CACNG1* 人源化区域=小鼠 (小写) 人 *Xho1* *LoxP* ICEUI 小鼠 (小写)

图12 (续)