(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 112601454 B
(45)授权公告日 2023.10.27

(21)申请号 201980040365.8

(22)申请日 2019.04.16

(65)同一申请的已公布的文献号
申请公布号 CN 112601454 A

(43)申请公布日 2021.04.02

(30)优先权数据
62/658464 2018.04.16 US

(85)PCT国际申请进入国家阶段日
2020.12.16

(86)PCT国际申请的申请数据
PCT/US2019/027675 2019.04.16

(87)PCT国际申请的公布数据
WO2019/204303 EN 2019.10.24

(73)专利权人 宾夕法尼亚州大学信托人
地址 美国宾夕法尼亚州

(72)发明人 H·斯特德曼

(74)专利代理机构 中国专利代理(香港)有限公司 72001
专利代理师 权陆军 李志强

(51)Int.Cl.
C12N 15/861(2006.01)

(56)对比文件
WO 2018/053632 A1,2018.03.29
US 2017/0368198 A1,2017.12.28
US 7,771,993 B2,2010.08.10
CN 107250364 A,2017.10.13
CN 102203253 A,2011.09.28
Tuffery-Giraud S等.“dystrophin Dp427p1 isoform [Homo sapiens]”.《GenBank Database》.2017,Accession NO:NP_004000.1.
Tinsley,J.M.等.“utrophin (dystrophin related protein) [Homo sapiens]”.《GenBank Database》.2008,Accession NO:CAA48829.1.
黎红华等.“重组腺相关病毒载体介导的小基因SMCKA3999治疗Duchenne肌营养不良小鼠的研究”.《中华神经科杂志》.2005,(第9期),第27-30页.
VanBelzen, DJ等.“Mechanism of Deletion Removing All Dystrophin Exons in a Canine Model for DMD Implicates Concerted Evolution of X Chromosome Pseudogenes”.《MOLECULAR THERAPY-METHODS & CLINICAL DEVELOPMENT》.2017,第4卷第62-71页.

审查员 刘魏

权利要求书9页 说明书70页
序列表186页 附图41页

(54)发明名称

用于治疗杜兴肌营养不良的组合物和方法

(57)摘要

本文描述了肌养蛋白或肌养蛋白相关蛋白的三重剪接突变体及其用于治疗杜兴肌营养不良的方法。还提供了病毒载体,其包含编码在指导其表达的调节元件控制下的三重剪接突变肌养蛋白或肌养蛋白相关蛋白的核酸。还提供了组合物,其含有配制用于递送给人患者的这种病毒载体。

CN 112601454 B

1.具有AAV衣壳和载体基因组的重组腺伴随病毒(rAAV),其中所述载体基因组包含编码包含杂合三股螺旋结构域的肌养蛋白超家族突变蛋白的核酸序列,所述核酸序列在指导其表达的调节序列的控制下,所述杂合三股螺旋结构域包含

第一杂合螺旋、第二杂合螺旋、第三杂合螺旋和连接天然肌养蛋白超家族蛋白的第一螺旋结构域和第二不相邻的螺旋结构域的区段的突变剪接点,

其中所述天然肌养蛋白超家族蛋白的第一螺旋结构域具有平行螺旋A和C和反平行B螺旋,

其中所述天然肌养蛋白超家族蛋白的第二不相邻的螺旋结构域具有平行螺旋A'和C'和反平行B'螺旋,

所述杂合三股螺旋结构域的每个螺旋包含天然肌养蛋白超家族蛋白的三股螺旋重复序列中的螺旋的N-末端部分和C-末端部分,其中所述杂合三股螺旋结构域具有平行的第一和第三螺旋和反平行的第二螺旋,其中所述杂合三股螺旋结构域中第一和第三螺旋的N-末端部分来自所述天然肌养蛋白超家族蛋白的所述第一螺旋结构域,且所述第一和第三螺旋的C-末端部分来自所述天然肌养蛋白超家族蛋白的所述第二不相邻的螺旋结构域,

所述杂合三股螺旋结构域的第一杂合螺旋包含与螺旋A'的C-末端部分融合的螺旋A的N-末端部分;

所述杂合三股螺旋结构域的第二杂合螺旋包含与螺旋B的C-末端部分融合的螺旋B'的N-末端部分,其中所述螺旋B和螺旋B'在与所述杂合三股螺旋结构域中连接的所述第一螺旋结构域和所述第二不相邻的螺旋结构域的疏水核中心的中央保守色氨酸残基直接相邻的位置处剪接,所述疏水核对应于将所述色氨酸残基二等分的横截面平面,通过隐蔽马尔科夫模型和序列比对所确定的;以及

所述杂合三股螺旋结构域的第三杂合螺旋包含与螺旋C'的C-末端部分融合的螺旋C的N-末端部分;

其中所述第一杂合螺旋和所述第三杂合螺旋在所述杂合三股螺旋结构域中连接的所述第一螺旋结构域和所述第二不相邻的螺旋结构域的疏水核中心的保守色氨酸残基处剪接,所述疏水核对应于将所述色氨酸残基二等分的横截面平面,通过隐蔽马尔科夫模型和序列比对所确定的。

2.根据权利要求1所述的rAAV,其中所述肌养蛋白超家族突变蛋白是三重剪接突变肌养蛋白。

3.根据权利要求1或2所述的rAAV,其中所述肌养蛋白超家族突变蛋白包含全长肌养蛋白的至少螺旋重复序列3至螺旋重复序列21中的缺失。

4.根据权利要求1所述的rAAV,其中所述肌养蛋白超家族突变蛋白是三重剪接突变肌养蛋白,且所述突变蛋白的N-末端螺旋重复序列由全长肌养蛋白的螺旋重复序列1组成,且所述突变蛋白的C-末端螺旋重复序列由全长肌养蛋白的螺旋重复序列23和螺旋重复序列24组成,且形成杂合三股螺旋重复序列的两个螺旋重复序列的第一个是全长肌养蛋白中的螺旋重复序列2,且形成杂合三股螺旋重复序列的两个螺旋重复序列的第二个是全长肌养蛋白中的螺旋重复序列22。

5.根据权利要求1或2所述的rAAV,其中所述肌养蛋白超家族突变蛋白包含全长肌养蛋白的至少螺旋重复序列3至螺旋重复序列23中的缺失。

6.根据权利要求1或2所述的rAAV,其中所述肌养蛋白超家族突变蛋白包含SEQ ID NO:1的氨基酸序列。

7.根据权利要求1或2所述的rAAV,其中编码所述肌养蛋白超家族突变蛋白的核酸序列包含SEQ ID NO:2或是进行密码子优化用于在人中表达的编码SEQ ID NO:1的核酸序列。

8.根据权利要求1所述的rAAV,其中所述肌养蛋白超家族突变蛋白包含全长肌养蛋白的至少螺旋重复序列2至螺旋重复序列19中的缺失。

9.根据权利要求1或8所述的rAAV,其中所述肌养蛋白超家族突变蛋白包含SEQ ID NO:22的氨基酸序列。

10.根据权利要求1所述的rAAV,其中所述肌养蛋白超家族突变蛋白是突变肌养蛋白且具有5个血影蛋白样三股螺旋重复序列,包括通过剪接全长人肌养蛋白中的三股螺旋重复序列1和20形成的杂合三股螺旋结构域,任选地其中所述肌养蛋白超家族突变蛋白包含SEQ ID NO:22的氨基酸序列。

11.根据权利要求1所述的rAAV,其中所述肌养蛋白超家族突变蛋白是三重剪接突变肌养蛋白相关蛋白。

12.根据权利要求1或11所述的rAAV,其中所述肌养蛋白超家族突变蛋白是三重剪接突变肌养蛋白相关蛋白,并且包含全长肌养蛋白相关蛋白的至少螺旋重复序列3至螺旋重复序列19中的缺失。

13.根据权利要求12所述的rAAV,其中所述肌养蛋白超家族突变蛋白包含SEQ ID NO:3的氨基酸序列。

14.根据权利要求1、11或13中任一项所述的rAAV,其中编码所述肌养蛋白超家族突变蛋白的核酸序列包含SEQ ID NO:4。

15.根据权利要求1所述的rAAV,其中所述肌养蛋白超家族突变蛋白包含SEQ ID NO:20的氨基酸序列。

16.根据权利要求15所述的rAAV,其中编码所述肌养蛋白超家族突变蛋白的核酸序列包含SEQ ID NO:19,或与SEQ ID NO:19至少约95%相同的序列。

17.根据权利要求1或11所述的rAAV,其中所述肌养蛋白超家族突变蛋白是三重剪接突变肌养蛋白相关蛋白,并且包含全长肌养蛋白相关蛋白的至少螺旋重复序列2至螺旋重复序列17中的缺失,且具有5个血影蛋白样三股螺旋重复序列,包括通过剪接全长人肌养蛋白相关蛋白中的三股螺旋重复序列1和18形成的杂合三股螺旋结构域。

18.根据权利要求17所述的rAAV,其中所述肌养蛋白超家族突变蛋白包含SEQ ID NO:21的氨基酸序列。

19.根据权利要求11所述的rAAV,其中所述突变蛋白的N-末端螺旋重复序列由全长肌养蛋白相关蛋白中的螺旋重复序列1组成,且所述突变蛋白的C-末端螺旋重复序列由全长肌养蛋白相关蛋白中的螺旋重复序列21和螺旋重复序列22组成,且形成杂合三股螺旋重复序列的两个螺旋重复序列的第一个是全长肌养蛋白相关蛋白中的螺旋重复序列2,且形成杂合三股螺旋重复序列的两个螺旋重复序列的第二个是全长肌养蛋白相关蛋白中的螺旋重复序列20。

20.根据权利要求11所述的rAAV,其中所述突变蛋白的C-末端螺旋重复序列包括全长肌养蛋白相关蛋白中的螺旋重复序列19、20、21和22,且形成杂合三股螺旋重复序列的两个

螺旋重复序列的第一个是全长肌养蛋白相关蛋白中的螺旋重复序列1，且形成杂合三股螺旋重复序列的两个螺旋重复序列的第二个是全长肌养蛋白相关蛋白中的螺旋重复序列18。

21.具有AAV衣壳和载体基因组的rAAV，其中所述载体基因组包含编码肌养蛋白超家族三重剪接突变蛋白的核酸序列，所述核酸序列在指导其表达的调节序列的控制下，其中所述三重剪接突变蛋白包含一个或多个N-末端螺旋重复序列、杂合三股螺旋重复序列和一个或多个C-末端螺旋重复序列，其中所述三重剪接突变蛋白中包括杂合重复序列的螺旋重复序列的总数目选自1至比全长肌养蛋白超家族蛋白的螺旋重复序列数目小1的任何整数，并且其中所述杂合三股螺旋重复序列由如图2F所描绘的在垂直于其长轴将螺旋重复序列二等分的平面上剪接的两个螺旋重复序列形成。

22.根据权利要求21所述的rAAV，其中所述突变蛋白的一个或多个N-末端螺旋重复序列当在全长肌养蛋白超家族蛋白中时与形成杂合三股螺旋重复序列的两个重复序列中的螺旋重复序列1直接相邻；且其中所述突变蛋白的一个或多个C-末端螺旋重复序列当在全长肌养蛋白超家族蛋白中时与形成杂合三股螺旋重复序列的两个重复序列中的螺旋重复序列2直接相邻。

23.根据权利要求21或22所述的rAAV，其中所述全长肌养蛋白超家族蛋白是肌养蛋白。

24.根据权利要求21或22所述的rAAV，其中所述突变蛋白的一个或多个N-末端螺旋重复序列包含全长肌养蛋白中的螺旋重复序列1，和/或其中一个或多个C-末端螺旋重复序列包含全长肌养蛋白中的螺旋重复序列23和螺旋重复序列24。

25.根据权利要求21或22所述的rAAV，其中所述N-末端螺旋重复序列由全长肌养蛋白中的螺旋重复序列1组成，其中所述C-末端螺旋重复序列由全长肌养蛋白中的螺旋重复序列23和螺旋重复序列24组成，其中形成杂合三股螺旋重复序列的两个重复序列中的螺旋重复序列1是全长肌养蛋白中的螺旋重复序列2，并且其中形成杂合三股螺旋重复序列的两个重复序列中的螺旋重复序列2是全长肌养蛋白中的螺旋重复序列22。

26.根据权利要求21或22中任一项所述的rAAV，其中所述三重剪接突变蛋白包含SEQ ID NO:1的氨基酸序列。

27.根据权利要求26所述的rAAV，其中编码所述三重剪接突变蛋白的序列包含SEQ ID NO:2的核酸序列。

28.根据权利要求21或22所述的rAAV，其中所述全长肌养蛋白超家族蛋白是肌养蛋白相关蛋白。

29.根据权利要求21或22所述的rAAV，其中所述突变蛋白的一个或多个N-末端螺旋重复序列包含全长肌养蛋白相关蛋白中的螺旋重复序列1，和/或其中所述突变蛋白的C-末端螺旋重复序列包含全长肌养蛋白相关蛋白中的螺旋重复序列21和22。

30.根据权利要求21或22所述的rAAV，其中所述突变蛋白的N-末端螺旋重复序列由全长肌养蛋白相关蛋白中的螺旋重复序列1组成，其中所述突变蛋白的C-末端螺旋重复序列由全长肌养蛋白相关蛋白中的螺旋重复序列21和螺旋重复序列22组成，其中形成杂合三股螺旋重复序列的两个重复序列中的螺旋重复序列1是全长肌养蛋白相关蛋白中的螺旋重复序列2，且其中形成杂合三股螺旋重复序列的两个重复序列中的螺旋重复序列2是全长肌养蛋白相关蛋白中的螺旋重复序列20。

31.根据权利要求21或22所述的rAAV，其中所述三重剪接突变蛋白包含SEQ ID NO:3的

氨基酸序列。

32.根据权利要求31所述的rAAV,其中编码所述三重剪接突变蛋白的序列包含SEQ ID NO:4的核酸序列。

33.根据权利要求21或22所述的rAAV,其中所述三重剪接突变蛋白包含SEQ ID NO:20的氨基酸序列。

34.根据权利要求33所述的rAAV,其中编码所述三重剪接突变蛋白的核酸序列包含SEQ ID NO:19,或与SEQ ID NO:19至少约95%相同的序列。

35.具有AAV衣壳和载体基因组的rAAV,其中所述载体基因组包含编码肌养蛋白超家族三重剪接突变蛋白的核酸序列,所述核酸序列在指导其表达的调节序列的控制下,其中所述三重剪接突变蛋白包含杂合三股螺旋重复序列和C-末端螺旋重复序列,其中所述三重剪接突变蛋白中包括杂合重复序列的螺旋重复序列的总数目是5,并且其中所述杂合三股螺旋重复序列由如图2F所描绘的在垂直于其长轴将螺旋重复序列二等分的平面上剪接的两个螺旋重复序列形成。

36.根据权利要求35所述的rAAV,其中所述突变蛋白的C-末端螺旋重复序列由全长肌养蛋白中的螺旋重复序列21、22、23和24组成,其中形成杂合三股螺旋重复序列的两个重复序列中的螺旋重复序列1是全长肌养蛋白中的螺旋重复序列1,并且其中形成杂合三股螺旋重复序列的两个重复序列中的螺旋重复序列2是全长肌养蛋白中的螺旋重复序列20。

37.根据权利要求36所述的rAAV,其中所述三重剪接突变蛋白包含SEQ ID NO:22的氨基酸序列。

38.根据权利要求35所述的rAAV,其中所述突变蛋白的C-末端螺旋重复序列由全长肌养蛋白相关蛋白中的螺旋重复序列19、20、21和22组成,其中形成杂合三股螺旋重复序列的两个重复序列中的螺旋重复序列1是全长肌养蛋白相关蛋白中的螺旋重复序列1,并且其中形成杂合三股螺旋重复序列的两个重复序列中的螺旋重复序列2是全长肌养蛋白相关蛋白中的螺旋重复序列18。

39.根据权利要求38所述的rAAV,其中所述三重剪接突变蛋白包含SEQ ID NO:21的氨基酸序列。

40.根据权利要求1、2、8、11、13、16、18-22、27、32和34-39中任一项所述的rAAV,其中调节序列包含启动子,其中所述启动子是组成型启动子,或者其中所述启动子是肌肉特异性启动子。

41.根据权利要求40所述的rAAV,其中所述肌肉特异性启动子是肌肉肌酸激酶(MCK)启动子或SPc5-12启动子。

42.根据权利要求1、2、8、11、13、16、18-22、27、32、34-39和41中任一项所述的rAAV,其中所述调节序列包含嵌合CS-CRM4/SPc5-12启动子。

43.根据权利要求1、2、8、11、13、16、18-22、27、32、34-39和41中任一项所述的rAAV,其中所述调节序列包含增强子。

44.根据权利要求43所述的rAAV,其中所述增强子是MCK增强子的双重或三重串联。

45.根据权利要求1、2、8、11、13、16、18-22、27、32、34-39、41和44中任一项所述的rAAV,其中所述重组腺伴随病毒具有AAV1、AAV5、AAV6、AAV8、AAV8triple、AAV9、Anc80、Anc81或Anc82的衣壳。

46.重组肌养蛋白超家族三重剪接突变蛋白,其包含杂合三股螺旋结构域,所述杂合三股螺旋结构域包含

第一杂合螺旋、第二杂合螺旋、第三杂合螺旋和连接天然肌养蛋白超家族蛋白的第一螺旋结构域和第二不相邻的螺旋结构域的区段的突变剪接点,

其中所述天然肌养蛋白超家族蛋白的第一螺旋结构域具有平行螺旋A和C和反平行B螺旋,

其中所述天然肌养蛋白超家族蛋白的第二不相邻的螺旋结构域具有平行螺旋A'和C'和反平行B'螺旋,

所述杂合三股螺旋结构域的每个螺旋包含天然肌养蛋白超家族蛋白的三股螺旋重复序列中的螺旋的N-末端部分和C-末端部分,其中所述杂合三股螺旋结构域具有平行的第一和第三螺旋和反平行的第二螺旋,其中所述杂合三股螺旋结构域中第一和第三螺旋的N-末端部分来自所述天然肌养蛋白超家族蛋白的所述第一螺旋结构域,且所述第一和第三螺旋的C-末端部分来自所述天然肌养蛋白超家族蛋白的所述第二不相邻的螺旋结构域,

所述杂合三股螺旋结构域的第一杂合螺旋包含与螺旋A'的C-末端部分融合的螺旋A的N-末端部分;

所述杂合三股螺旋结构域的第二杂合螺旋包含与螺旋B的C-末端部分融合的螺旋B'的N-末端部分,其中螺旋B和螺旋B'在与所述杂合三股螺旋结构域中连接的所述第一螺旋结构域和所述第二不相邻的螺旋结构域的疏水核中心的中央保守色氨酸残基直接相邻的位置处剪接,所述疏水核对应于将所述色氨酸残基二等分的横截面平面,通过隐蔽马尔科夫模型和序列比对所确定的;以及

所述杂合螺旋结构域的第三杂合螺旋包含与螺旋C'的C-末端部分融合的螺旋C的N-末端部分;

其中所述第一杂合螺旋和所述第三杂合螺旋在所述杂合三股螺旋结构域中连接的所述第一螺旋结构域和所述第二不相邻的螺旋结构域的疏水核中心的保守色氨酸残基处剪接,所述疏水核对应于将所述色氨酸残基二等分的横截面平面,通过隐蔽马尔科夫模型和序列比对所确定的。

47.根据权利要求46所述的重组肌养蛋白超家族三重剪接突变蛋白,其中所述肌养蛋白超家族突变蛋白是三重剪接突变肌养蛋白。

48.根据权利要求46或47所述的重组肌养蛋白超家族三重剪接突变蛋白,其中所述三重剪接突变蛋白包含全长肌养蛋白的至少螺旋重复序列3至螺旋重复序列21中的缺失。

49.根据权利要求46所述的重组肌养蛋白超家族三重剪接突变蛋白,其中所述肌养蛋白超家族突变蛋白是三重剪接突变肌养蛋白,且所述突变蛋白的N-末端螺旋重复序列由全长肌养蛋白的螺旋重复序列1组成,且所述突变蛋白的C-末端螺旋重复序列由全长肌养蛋白的螺旋重复序列23和螺旋重复序列24组成,且形成杂合三股螺旋重复序列的两个螺旋重复序列的第一个是全长肌养蛋白中的螺旋重复序列2,且形成杂合三股螺旋重复序列的两个螺旋重复序列的第二个是全长肌养蛋白中的螺旋重复序列22。

50.根据权利要求46或47所述的重组肌养蛋白超家族三重剪接突变蛋白,其中所述三重剪接突变蛋白包含全长肌养蛋白的至少螺旋重复序列3至螺旋重复序列23中的缺失。

51.根据权利要求46或47中任一项所述的重组肌养蛋白超家族三重剪接突变蛋白,其

中所述三重剪接突变蛋白包含SEQ ID NO:1的氨基酸序列。

52.根据权利要求46或47中任一项所述的重组肌养蛋白超家族三重剪接突变蛋白,其中编码所述三重剪接突变蛋白的核酸序列包含SEQ ID NO:2的序列或是进行密码子优化用于在人中表达的编码SEQ ID NO:1的核酸序列。

53.根据权利要求46所述的重组肌养蛋白超家族三重剪接突变蛋白,其中所述三重剪接突变蛋白包含全长肌养蛋白的至少螺旋重复序列2至螺旋重复序列19中的缺失。

54.根据权利要求46或53所述的重组肌养蛋白超家族三重剪接突变蛋白,其中所述三重剪接突变蛋白包含SEQ ID NO:22的氨基酸序列。

55.根据权利要求46所述的重组肌养蛋白超家族三重剪接突变蛋白,其中所述三重剪接突变蛋白是突变肌养蛋白且具有5个血影蛋白样三股螺旋重复序列,包括通过剪接全长人肌养蛋白中的三股螺旋重复序列1和20形成的杂合三股螺旋结构域,任选地其中所述三重剪接突变蛋白包含SEQ ID NO:22的氨基酸序列。

56.根据权利要求46所述的重组肌养蛋白超家族三重剪接突变蛋白,其中所述肌养蛋白超家族突变蛋白是三重剪接突变肌养蛋白相关蛋白。

57.根据权利要求46或56所述的重组肌养蛋白超家族三重剪接突变蛋白,其中所述肌养蛋白超家族突变蛋白是三重剪接突变肌养蛋白相关蛋白,并且包含全长肌养蛋白相关蛋白的至少螺旋重复序列3至螺旋重复序列19中的缺失。

58.根据权利要求57所述的重组肌养蛋白超家族三重剪接突变蛋白,其中所述三重剪接突变蛋白包含SEQ ID NO:3的氨基酸序列。

59.根据权利要求46、56和58中任一项所述的重组肌养蛋白超家族三重剪接突变蛋白,其中编码所述三重剪接突变蛋白的核酸序列包含SEQ ID NO:4。

60.根据权利要求46所述的重组肌养蛋白超家族三重剪接突变蛋白,其中所述三重剪接突变蛋白包含SEQ ID NO:20的氨基酸序列。

61.根据权利要求60所述的重组肌养蛋白超家族三重剪接突变蛋白,其中编码所述三重剪接突变蛋白的核酸序列包含SEQ ID NO:19,或与SEQ ID NO:19至少约95%相同的序列。

62.根据权利要求46或56所述的重组肌养蛋白超家族三重剪接突变蛋白,其中所述肌养蛋白超家族突变蛋白是三重剪接突变肌养蛋白相关蛋白,并且包含全长肌养蛋白相关蛋白的至少螺旋重复序列2至螺旋重复序列17中的缺失,且具有5个血影蛋白样三股螺旋重复序列,包括通过剪接全长人肌养蛋白相关蛋白中的三股螺旋重复序列1和18形成的杂合三股螺旋结构域。

63.根据权利要求46或56所述的重组肌养蛋白超家族三重剪接突变蛋白,其中所述三重剪接突变蛋白包含SEQ ID NO:21的氨基酸序列。

64.根据权利要求46所述的重组肌养蛋白超家族三重剪接突变蛋白,其中所述突变蛋白的N-末端螺旋重复序列由全长肌养蛋白相关蛋白中的螺旋重复序列1组成,且所述突变蛋白的C-末端螺旋重复序列由全长肌养蛋白相关蛋白中的螺旋重复序列21和螺旋重复序列22组成,且形成杂合三股螺旋重复序列的两个螺旋重复序列的第一个是全长肌养蛋白相关蛋白中的螺旋重复序列2,且形成杂合三股螺旋重复序列的两个螺旋重复序列的第二个是全长肌养蛋白相关蛋白中的螺旋重复序列20。

65.根据权利要求46所述的重组肌养蛋白超家族三重剪接突变蛋白,其中所述突变蛋白的C-末端螺旋重复序列包括全长肌养蛋白相关蛋白中的螺旋重复序列19、20、21和22,且形成杂合三股螺旋重复序列的两个螺旋重复序列的第一个是全长肌养蛋白相关蛋白中的螺旋重复序列1,且形成杂合三股螺旋重复序列的两个螺旋重复序列的第二个是全长肌养蛋白相关蛋白中的螺旋重复序列18。

66.重组肌养蛋白超家族三重剪接突变蛋白,其包含一个或多个N-末端螺旋重复序列、杂合三股螺旋重复序列和一个或多个C-末端螺旋重复序列,其中所述三重剪接突变蛋白中包括杂合重复序列的一个或多个螺旋重复序列的总数目选自1至比全长肌养蛋白超家族蛋白的螺旋重复序列数目小1的任何整数,并且其中所述杂合三股螺旋重复序列由如图2F所描绘的在垂直于其长轴将螺旋重复序列二等分的平面上剪接的两个螺旋重复序列形成。

67.根据权利要求66所述的重组肌养蛋白超家族三重剪接突变蛋白,其中所述突变蛋白的一个或多个N-末端螺旋重复序列当在全长肌养蛋白超家族蛋白中时与形成杂合三股螺旋重复序列的两个重复序列中的螺旋重复序列1直接相邻;且其中所述突变蛋白的一个或多个C-末端螺旋重复序列当在全长肌养蛋白超家族蛋白中时与形成杂合三股螺旋重复序列的两个重复序列中的螺旋重复序列2直接相邻。

68.根据权利要求66或67所述的重组肌养蛋白超家族三重剪接突变蛋白,其中所述全长肌养蛋白超家族蛋白是肌养蛋白。

69.根据权利要求66或67所述的重组肌养蛋白超家族三重剪接突变蛋白,其中所述突变蛋白的一个或多个N-末端螺旋重复序列包含全长肌养蛋白中的螺旋重复序列1,和/或其中一个或多个C-末端螺旋重复序列包含全长肌养蛋白中的螺旋重复序列23和螺旋重复序列24。

70.根据权利要求66或67所述的重组肌养蛋白超家族三重剪接突变蛋白,其中所述N-末端螺旋重复序列由全长肌养蛋白中的螺旋重复序列1组成,其中所述C-末端螺旋重复序列由全长肌养蛋白中的螺旋重复序列23和螺旋重复序列24组成,其中形成杂合三股螺旋重复序列的两个重复序列中的螺旋重复序列1是全长肌养蛋白中的螺旋重复序列2,并且其中形成杂合三股螺旋重复序列的两个重复序列中的螺旋重复序列2是全长肌养蛋白中的螺旋重复序列22。

71.根据权利要求66或67所述的重组肌养蛋白超家族三重剪接突变蛋白,其中所述三重剪接突变蛋白包含SEQ ID NO:1的氨基酸序列。

72.根据权利要求71所述的重组肌养蛋白超家族三重剪接突变蛋白,其中编码所述三重剪接突变蛋白的序列包含SEQ ID NO:2的核酸序列。

73.根据权利要求66或67所述的重组肌养蛋白超家族三重剪接突变蛋白,其中所述全长肌养蛋白超家族蛋白是肌养蛋白相关蛋白。

74.根据权利要求66或67所述的重组肌养蛋白超家族三重剪接突变蛋白,其中所述突变蛋白的一个或多个N-末端螺旋重复序列包含全长肌养蛋白相关蛋白中的螺旋重复序列1,和/或其中所述突变蛋白的C-末端螺旋重复序列包含全长肌养蛋白相关蛋白中的螺旋重复序列21和22。

75.根据权利要求66或67所述的重组肌养蛋白超家族三重剪接突变蛋白,其中所述突变蛋白的N-末端螺旋重复序列由全长肌养蛋白相关蛋白中的螺旋重复序列1组成,其中所

述突变蛋白的C-末端螺旋重复序列由全长肌养蛋白相关蛋白中的螺旋重复序列21和螺旋重复序列22组成，其中形成杂合三股螺旋重复序列的两个重复序列中的螺旋重复序列1是全长肌养蛋白相关蛋白中的螺旋重复序列2，且其中形成杂合三股螺旋重复序列的两个重复序列中的螺旋重复序列2是全长肌养蛋白相关蛋白中的螺旋重复序列20。

76.根据权利要求66或67所述的重组肌养蛋白超家族三重剪接突变蛋白，其中所述三重剪接突变蛋白包含SEQ ID NO:3的氨基酸序列。

77.根据权利要求76所述的重组肌养蛋白超家族三重剪接突变蛋白，其中编码所述三重剪接突变蛋白的序列包含SEQ ID NO:4的核酸序列。

78.根据权利要求66或67所述的重组肌养蛋白超家族三重剪接突变蛋白，其中所述三重剪接突变蛋白包含SEQ ID NO:20的氨基酸序列。

79.根据权利要求78所述的重组肌养蛋白超家族三重剪接突变蛋白，其中编码所述三重剪接突变蛋白的核酸序列包含SEQ ID NO:19，或与SEQ ID NO:19至少约95％相同的序列。

80.重组肌养蛋白超家族三重剪接突变蛋白，其包含杂合三股螺旋重复序列和C-末端螺旋重复序列，其中所述三重剪接突变蛋白中包括杂合重复序列的螺旋重复序列的总数目是5，并且其中所述杂合三股螺旋重复序列由如图2F所描绘的在垂直于其长轴将螺旋重复序列二等分的平面上剪接的两个螺旋重复序列形成。

81.根据权利要求80所述的重组肌养蛋白超家族三重剪接突变蛋白，其中所述突变蛋白的C-末端螺旋重复序列由全长肌养蛋白中的螺旋重复序列21、22、23和24组成，其中形成杂合三股螺旋重复序列的两个重复序列中的螺旋重复序列1是全长肌养蛋白中的螺旋重复序列1，并且其中形成杂合三股螺旋重复序列的两个重复序列中的螺旋重复序列2是全长肌养蛋白中的螺旋重复序列20。

82.根据权利要求81所述的重组肌养蛋白超家族三重剪接突变蛋白，其中所述三重剪接突变蛋白包含SEQ ID NO:22的氨基酸序列。

83.根据权利要求80所述的重组肌养蛋白超家族三重剪接突变蛋白，其中所述突变蛋白的C-末端螺旋重复序列由全长肌养蛋白相关蛋白中的螺旋重复序列19、20、21和22组成，其中形成杂合三股螺旋重复序列的两个重复序列中的螺旋重复序列1是全长肌养蛋白相关蛋白中的螺旋重复序列1，并且其中形成杂合三股螺旋重复序列的两个重复序列中的螺旋重复序列2是全长肌养蛋白相关蛋白中的螺旋重复序列18。

84.根据权利要求83所述的重组肌养蛋白超家族三重剪接突变蛋白，其中所述三重剪接突变蛋白包含SEQ ID NO:21的氨基酸序列。

85.重组突变肌养蛋白，其包含SEQ ID NO:1、13、14、15、16、17、18或22的氨基酸序列。

86.重组突变肌养蛋白相关蛋白，其包含SEQ ID NO:3、20或21的氨基酸序列。

87.重组核酸分子，其编码根据权利要求46至86中任一项所述的重组肌养蛋白超家族三重剪接突变蛋白。

88.包含根据权利要求87的核酸分子的质粒，所述核酸分子在指导其表达的调节序列的控制下。

89.药物组合物，其包含在制剂缓冲液中的根据权利要求1-45中任一项所述的rAAV。

90.根据权利要求89所述的组合物在制备用于治疗杜兴肌营养不良(DMD)的药物中的

用途。

91.根据权利要求1-45中任一项所述的rAAV在制备用于治疗杜兴肌营养不良(DMD)的药物中的用途。

用于治疗杜兴肌营养不良的组合物和方法

[0001] 以电子形式提交的材料的引入作为参考

[0002] 申请人特此引入与此一道以电子形式提交的序列表材料作为参考。所述文件被称为“UPN_18-8438PCT_ST25.txt”,创建于2019年4月16日,且为269,817字节。

[0003] 联邦资助研究的声明

[0004] 本发明是借助在美国国立卫生研究院(National Institutes of Health)/美国国立神经病学与中风研究所(NationalInstitute of Neurological Disorders and Stroke)授予的R01NS094705项下的政府支持完成的。政府拥有本发明的某些权利。

[0005] 发明背景

[0006] 杜兴肌营养不良(DMD)是一种严重致残的全身性儿童期发病疾病,其从早期运动肌无力进展到末期呼吸和心肌病,其特征在于非常昂贵的护理者和技术依赖性(Landfeldt, E.等人, The burden of Duchenne musculardystrophy: an international, cross-sectional study. Neurology, 2014. 83(6): 第529-36页;和Schreiber-Katz, O.等人, Comparative cost of illness analysisand assessment of health care burden of Duchenne and Becker muscular dystrophies in Germany. Orphanet J Rare Dis, 2014. 9: 第210页)。DMD是人类最常见的单基因致死性疾病之一,其历史上全球发病率为约1:4000男性出生。分子基础是细胞骨架蛋白肌养蛋白的427 kd同工型(Dp427)的缺乏,其在大多数情况下是由DMD基因的多外显子移码缺失引起的。较轻的疾病贝克尔肌营养不良(BMD)是等位的,且大多数病例是由改变了编码的蛋白的杆结构域的长度的肌养蛋白基因的内部缺失或重复引起的。

[0007] 肌养蛋白是通过“定位克隆”发现的第一种蛋白,且所述发现提供了关于使用疾病的人遗传学图位置作为阐明其分子基础的主要基础的初始概念证明(proof of concept)。通过这种方法发现的许多蛋白在受影响的细胞中的丰度非常低,从而使得确定蛋白的生理功能变得复杂化。肌养蛋白发现于1987年,且在所述领域中的30年都面临着对蛋白在肌细胞中的精确功能的理解的较大空白。然而,已经提供了肌养蛋白在保护肌细胞膜免受肌肉收缩过程中产生的力中的作用的间接证据。肌养蛋白的机械负荷特性仍然是很差表征的。

[0008] 在美国,携带者检测和出生前咨询已多少降低了DMD的发病率(Pegoraro, E.等人, SPP1 genotype is a determinant ofdisease severity in Duchenne muscular dystrophy. Neurology, 2011. 76(3): 第219-26页)。目前结合糖皮质类固醇、ACE抑制剂和机械通气支持的疗法可能暂时减慢进展速度,但最终的临床过程是不可抗拒的(McDonald, C.M.等人, The cooperative international neuromuscular research group duchenne natural history study-a longitudinal investigation in the era ofglucocorticoid therapy: Design of protocol and the methods used. Muscle Nerve, 2013)。

[0009] 在基因治疗的时代,各种AAV载体已作为用于全身基因递送的毒性最小且最广泛传播的平台出现。那些基因治疗对全身生物分布具有巨大希望,但局限性包括(a)克隆能力限于对于全长Dp427所需的三分之一,(b)在持久治疗潜在所需剂量的载体免疫原性和毒性

尚未解决的问题,以及(c)在人疗法所需的规模上进行常规制造的特别成本。AAV载体在结构上与细小病毒亚科的野生型成员有关,其包壳了约5千碱基的单链DNA基因组。全长肌养蛋白的mRNA为14千碱基,具有约12 kb的可读框。

[0010] 引起DMD的大多数突变是在>2.5兆碱基的、X-连锁基因中的散发性多外显子、移码缺失(Kunkel, L.M.等人, Analysis of deletions in DNA frompatients with Becker and Duchenne muscular dystrophy. Nature, 1986. 322(6074): 第73-7页;Monaco, A.P.等人, Isolation of candidate cDNAs for portions of the Duchenne muscular dystrophy gene. Nature, 1986. 323(6089): 第646-50页;和Koenig, M.等人, The molecular basis for Duchenne versus Becker musculardystrophy: correlation of severity with type of deletion. Am J Hum Genet, 1989. 45(4): 第498-506页)。在没有对缺陷的全长蛋白的中枢性(胸腺)耐受的情况下,重组肌养蛋白具有诱导宿主对异种蛋白的免疫应答的能力(Mendell,J.R.等人, Dystrophin immunity in Duchenne's muscular dystrophy. N Engl J Med, 2010. 363(15): p. 1429-37)。新的载体和血管递送方法已在概念证明临床前研究中实现了有希望的区域和全身基因转移,从而提出了用于DMD的基因治疗的合理方法(Greelish, J.P.等人,Stable restoration of the sarcoglycan complex in dystrophic muscle perfused with histamine and a recombinant adeno-associated viral vector. Nat Med, 1999.5(4): 第439-43页;Su, L.T.等人, Uniform scale-independent genetransfer to striated muscle after transvenular extravasation of vector. Circulation, 2005. 112(12): 第1780-8页; Gao, G., L.H. Vandenberghe, 和J.M. Wilson, New recombinant serotypes of AAV vectors.Curr Gene Ther, 2005. 5(3): 第285-97页;和Katz, M.G.等人, Cardiac gene therapy: optimization of gene delivery techniques in vivo. HumGene Ther, 2010. 21(4): p. 371-80)。然而,这些进展也强调了这个领域中主要的载体发现挑战和患者安全忧虑(Mendell, J.R.等人, Dystrophin immunity in Duchenne'smuscular dystrophy. N Engl J Med, 2010. 363(15): 第1429-37页;Mendell, J.R.等人, Myoblast transfer in the treatment of Duchenne's muscular dystrophy. N Engl JMed, 1995. 333(13): 第832-8页;Mouly, V.等人, Myoblast transfer therapy: is there any light at the end of the tunnel?Acta Myol, 2005. 24(2): 第128-33页; 和Wang, Z.等人, Immunity toadeno-associated virus-mediated gene transfer in a random-bred canine model of Duchenne muscular dystrophy. Hum Gene Ther, 2007. 18(1): 第18-26页)。作为治疗DMD的AAV载体的另一个例子,美国专利No. 7,771,993提供了"微-肌养蛋白相关蛋白(micro-utrophin)"(也记录为"m-肌养蛋白相关蛋白(utrophin)"、"μ-肌养蛋白相关蛋白"或"μ-U"),其具有位于N-末端肌养蛋白相关蛋白区域内的相对于人肌养蛋白相关蛋白的约270个氨基酸的"辅肌动蛋白结合结构域"的功能部分、至少富含脯氨酸的铰链区1和4(H1)和(H4)的功能部分和 C-末端肌养蛋白相关蛋白蛋白的部分。微-肌养蛋白相关蛋白含有中央杆重复序列结构域的内部缺失和下游C-末端区域中的截短。

[0011] 仍然存在用于治疗杜兴肌营养不良和相关疾病的需要。

发明内容

[0012] 本发明提供了组合物和方法,其可用于治疗肌营养不良(MD),包括杜兴肌营养不良(DMD)和贝克尔肌营养不良(BMD)和其他疾病。本文提供了具有AAV衣壳和载体基因组的重组腺伴随病毒(rAAV)载体。载体基因组包含编码肌养蛋白超家族三重剪接(triple splice)突变蛋白的核酸序列,所述核酸序列在指导其表达的调节序列的控制下。

[0013] 在某些实施方案中,肌养蛋白超家族三重剪接突变蛋白包含杂合螺旋结构域,所述杂合螺旋结构域包含包含与螺旋A'的C-末端部分融合的螺旋A的N-末端部分的第一螺旋,包含与螺旋B的末端C-末端部分融合的螺旋B'的N-末端部分的第二螺旋,以及包含与螺旋C'的C-末端部分融合的螺旋C的N-末端部分的第三螺旋,其中螺旋A、B和C存在于与天然肌养蛋白超家族蛋白中具有螺旋A'、B'和C'的第二三股螺旋重复序列不相邻的第一三股螺旋重复序列中。在某些实施方案中,肌养蛋白超家族突变蛋白是三重剪接突变肌养蛋白或三重剪接突变肌养蛋白相关蛋白。

[0014] 在再进一步的实施方案中,肌养蛋白超家族三重剪接突变蛋白包含一个或多个N-末端螺旋重复序列、杂合三股螺旋重复序列和一个或多个C-末端螺旋重复序列,其中所述三重剪接突变蛋白中包括杂合重复序列的螺旋重复序列的总数目选自1至比全长肌养蛋白超家族蛋白的螺旋重复序列数目小1的任何整数,并且其中所述杂合三股螺旋重复序列由如图2F所描绘的在垂直于其长轴将螺旋重复序列二等分的平面上剪接的两个螺旋重复序列形成。在某些实施方案中,肌养蛋白超家族突变蛋白是三重剪接突变肌养蛋白或三重剪接突变肌养蛋白相关蛋白。

[0015] 提供了具有SEQ ID NO:1或22的氨基酸序列的新重组突变肌养蛋白,并且还提供了具有选自SEQ ID NOs:3、7和20的氨基酸序列的新重组突变肌养蛋白相关蛋白。在某些实施方案中,三重剪接突变蛋白是肌养蛋白相关蛋白,并由包含SEQID NO:19的核酸或与其约95%至约99%相同的序列编码。

[0016] 在进一步的实施方案中,三重剪接突变蛋白包含杂合三股螺旋重复序列和C-末端螺旋重复序列,其中所述三重剪接突变蛋白中包括杂合重复序列的螺旋重复序列的总数目为5,并且其中所述杂合三股螺旋重复序列由如图2F所描绘的在垂直于其长轴将螺旋重复序列二等分的平面上剪接的两个螺旋重复序列形成。在某些实施方案中,突变蛋白的C-末端螺旋重复序列由全长肌养蛋白中的螺旋重复序列21、22、23和24组成,其中形成杂合三股螺旋重复序列的两个重复序列中的螺旋重复序列1是全长肌养蛋白中的螺旋重复序列1,并且其中形成杂合三股螺旋重复序列的两个重复序列中的螺旋重复序列2是全长肌养蛋白中的螺旋重复序列20。在再另一个实施方案中,突变蛋白的C-末端螺旋重复序列由全长肌养蛋白相关蛋白中的螺旋重复序列19、20、21和22组成,其中形成杂合三股螺旋重复序列的两个重复序列中的螺旋重复序列1是全长肌养蛋白相关蛋白中的螺旋重复序列1,并且其中形成杂合三股螺旋重复序列的两个重复序列中的螺旋重复序列2是全长肌养蛋白相关蛋白中的螺旋重复序列18。

[0017] 在再进一步的实施方案中,提供了包含SEQ ID NO:1、13、14、15、16、17、18或22的氨基酸序列的新重组突变肌养蛋白蛋白。在某些实施方案中,提供了编码突变肌养蛋白超家族蛋白的核酸。在再进一步的实施方案中,提供了包含编码突变肌养蛋白超家族蛋白的核酸的质粒。

[0018] 在某些实施方案中，提供了包含包含载体基因组的rAAV的药物组合物，所述载体基因组包含编码三重剪接突变肌养蛋白超家族蛋白的核酸序列。

[0019] 在再进一步的实施方案中，提供了治疗被诊断患有杜兴肌营养不良的主体的方法，其包括施用包含rAAV的药物组合物，所述rAAV包含具有编码三重剪接突变肌养蛋白超家族蛋白的核酸序列的载体基因组。

[0020] 根据以下的发明详述，本发明的其他方面和优点将显而易见。

[0021] 附图简述

[0022] 图1A至图1D提供了形成杂合三股螺旋结构域的例示。图1A显示了例示的营养不良超家族蛋白的组分。从蛋白的N-末端(表示为NH_2-)到C-末端(表示为-COOH)，组分为1)两个钙调理蛋白同源结构域(CH1和CH2)，2)几个三股螺旋结构域(记录为TH再加上一个数字)，3)模块式多结构域球状区域(WW-EF-ZZ)和4)最后的C-末端区域。在所述例示中，有24个三股螺旋结构域。灰色箭头表示在以前形成短肌养蛋白超家族蛋白中使用的剪接点位点(例如在美国专利No. 7,771,993中所述的，其整体引入本文作为参考)，而黑色箭头表示用于形成杂合三股螺旋结构域的剪接点位点。图1B提供了对被剪接以形成如图1C中所示的杂合三股螺旋的三股螺旋结构域(以白色箭标形式的THn和以阴影箭标形式的THn')的更精细查看。从图1B到图1D中的每个白色箭标或阴影箭标表示结构域中的螺旋(对于THn的螺旋A、B和C;对于THn'的螺旋A'、B'和C';以及杂合三股螺旋结构域中的螺旋A-A'、B'-B和C-C')。黑色箭标线指示剪接点位点。图1D提供了杂合三股螺旋结构域的线性显示。条/箭标的白色部分表示它们最初来自THn，而阴影部分表示它们最初来自THn'。如图1D中所示的，杂合三股螺旋结构域的第一螺旋包含在N-末端的螺旋A和在C-末端的螺旋A';杂合三股螺旋结构域的第二螺旋包含在N-末端的螺旋B'和在C-末端的螺旋B;且杂合三股螺旋结构域的第三螺旋包含在N-末端螺旋C和在C-末端的螺旋C'。

[0023] 图2A至图2F举例说明了本文提供的肌养蛋白相关蛋白修饰。较浅着色的残基描绘色氨酸，其是对于α-辅肌动蛋白、β-和α-血影蛋白、肌养蛋白、肌养蛋白相关蛋白和spectroplakins共享的三股螺旋重复序列结构域的隐蔽马尔科夫模造(HMM)中最保守的残基。在具有广泛进化偶联(evolutionary coupling)的所有其他位置处的序列趋异证实各个重复序列的这些区域在不使三股螺旋不稳定的情况下不能在一个重复序列与另一个重复序列之间互换。单个剪接不可以解决这个问题，但是通过保留色氨酸残基的平面左侧和右侧的偶联氨基酸穿过所述色氨酸残基的平面的三重剪接可以。图2A举例说明了α-辅肌动蛋白的三股螺旋重复序列结构域(TH1)的一部分。图2B举例说明了α-辅肌动蛋白的TH2部分。图2C举例说明了α-辅肌动蛋白的TH3部分。图2D举例说明了α-辅肌动蛋白的TH4部分。图2E举例说明了β-和α-血影蛋白tet位点。图2F举例说明通过保留色氨酸残基的平面左侧和右侧的偶联氨基酸穿过所述色氨酸残基的平面的三重剪接(为了帮助定向，将色氨酸显示两次)。

[0024] 图3显示了血影蛋白家族(PF00435)的HMM标识(logo)的一部分。可以在pfam.xfam.org/family/PF00435#tabview=tab4中找到更多细节。所述HMM标识以图形形式提供了HMM性质的快速概述。两个锚定色氨酸各自的相对保守性，如对于在位置＃16的血影蛋白样三股螺旋重复序列的HMM标识所示的。本领域技术人员可以查明如何解释标识，如例如在Schuster-Böckler B等人, HMM Logos for visualization of protein families.

BMC Bioinformatics. 2004年1月21日;5:7中所述的。

[0025] 图4提供了如本文所述形成毫微-肌养蛋白(nano-dystrophin)的概述以及与形成微-肌养蛋白(micro-dystrophin)的比较。

[0026] 图5提供了显示肌养蛋白的螺旋B中可能的剪接点的网格。关于所述图的更详细的描述可以在发明详述的第一段,I. 肌养蛋白、肌养蛋白相关蛋白等,A. 螺旋中的剪接点中找到。所描述的变体对应于SEQ ID NOs:13-18的毫微-肌养蛋白氨基酸序列。

[0027] 图6A至图6C显示肌养蛋白杆的长度在肌节出现之前确立。图6A显示比对的人肌养蛋白基因、mRNA和蛋白结构域结构以显示内含子位置和定相(phasing)。人肌养蛋白的24个血影蛋白重复序列中的22个在HMM位置46(带有黑色轮廓的点)处含有相位0内含子。刺胞动物(Cnidarian)肌养蛋白共享杆结构域相位0内含子(为清楚起见,在刺胞动物mRNA中仅描绘了HMM位置46相位0内含子)。图6B提供了在杆结构域与三股螺旋结构域HMM的共有序列垂直比对的情况下描绘的人β重血影蛋白、肌养蛋白和MACF1。叠加的是相对于相应编码序列的内含子位置和相位,从而清楚地显示了肌养蛋白和MACF1基因的结构相似性。为了清楚起见,仅显示与来自最远的相关物种的直向同源基因共享的内含子,而血影蛋白基因显示来自13个重复序列的遥远部分基因重复的残迹内含子(箭标)(血影蛋白-大堡礁海绵(*A. queenslandica*);肌养蛋白-星状海葵(*N. vectensis*);MACF1-丝盘虫(*T. adhaerens*))。图6C提供了在挑选的真核谱系中α-辅肌动蛋白蛋白超家族成员的种系基因组学分布。杆结构域血影蛋白重复序列的数目在圆括号中指定。dystroplakin超家族成员在HMM位置46处携带相位0内含子,而α-辅肌动蛋白和血影蛋白家族蛋白在所述位置缺乏相位0内含子的保守性。在真菌和扁盘动物(Placozoa)之间观察到祖先MACF1直向同源物的HMM 46内含子驱动的扩充。祖先MACF1直向同源物经历了部分基因重复,其向祖先WW-EF-ZZ肌养蛋白直向同源物贡献出N-末端肌动蛋白结合结构域和全长杆结构域。

[0028] 图7A至图7F表明广泛的转导在AAV9-μ肌养蛋白相关蛋白治疗的mdx小鼠中恢复肌养蛋白结合蛋白复合体,预防肌纤维变性,使血清CK水平正常化并改善肌肉功能。如图7A中所示的,代表性肢肌肉对于天然和重组肌养蛋白相关蛋白共享的表位(Utro_N),对于天然肌养蛋白相关蛋白独特的表位(Utro_C),γ-肌膜蛋白聚糖,胚胎肌球蛋白重链(eMHC)和层粘连蛋白的免疫染色;比例尺条25 μm。图7B提供了代表性肢肌肉的H&E,其显示了肌肉坏死和单核的细胞浸润的抑制;比例尺条100 μm。图7C提供了检测重组肌养蛋白相关蛋白(AAV9-μUtro)和γ-肌膜蛋白聚糖(γ-sarc)的表达的蛋白印迹分析。图7D提供了中央成核肌纤维(CNF)的百分比,统计量度,如在方法中所定义的(颜色=不同的动物,形状=不同的肌肉,相同的颜色/形状=技术重复)。图7E提供了在治疗的小鼠(n=5)、未治疗的小鼠(n=12)(***p<0.0001)和N.S.来自野生型(n=7)中血清CK水平的测量。图7F提供了在治疗的小鼠(n=8)、未治疗的*mdx*小鼠(n=11)和野生型小鼠(n=6)中握力后一小时的垂直活动(vertical activity)的定量。误差棒表示SD,**p<0.001;N.S表示不显著;通过具有多组比较的Kruskal-Wallis检验评估统计显著性。

[0029] 图8A至图8G显示在7周龄的GRMD狗中全身递送AAV9-μ肌养蛋白相关蛋白预防肌肉坏死并导致血清CK水平快速降低。图8A提供了实验时间表。图8B提供了股外侧肌和颞肌肌肉的代表性H&E,其显示未处理的肌肉中丰富的肌肉坏死纤维和单核的细胞浸润,而处理的肌肉类似WT。图8C提供茜素红S染色,其显示具有相应的定量(图8E)的表明肌肉变性(左

图,红色)的钙化纤维。图10D提供了利用f1.652的免疫荧光染色,其显示了具有相应的定量(图8F)的eMHC-阳性纤维的簇(右图,红色)。图8G提供了在全身AAV9-μ肌养蛋白相关蛋白输注前/后的各个时间点的血清肌酸激酶(CK)水平。比例尺条100 μm。

[0030] 图9A至图9D显示,在7周龄全身递送后,μ肌养蛋白相关蛋白的广泛表达在治疗的GRMD狗中拯救了肌养蛋白结合蛋白复合体蛋白。图9A和图9B提供了代表性肢肌肉的免疫荧光染色。图9A显示了天然和重组肌养蛋白相关蛋白(Utro_N)、天然肌养蛋白相关蛋白(Utro_C)、层粘连蛋白。图9B显示了β-肌营养不良蛋白聚糖(绿色)、β-肌膜蛋白聚糖(绿色)和γ-肌膜蛋白聚糖(红色)。比例尺条100 μm。图9C显示了蛋白印迹分析,其显示尸体剖检时横纹肌中μ肌养蛋白相关蛋白(～135kD)的广泛生物分布。图9D显示了蛋白印迹分析,其显示了在股外侧肌(VL)和颅缝匠肌(cranial sartorius)(CS)的肌肉活组织检查中β-肌营养不良蛋白聚糖的表达。治疗的(H)/治疗的(B)表示来自治疗的狗Hann和Beetle的组织。

[0031] 图10A至图10D显示μ肌养蛋白而不是μ肌养蛋白相关蛋白的局部表达在肌养蛋白缺失-无效狗模型中引出可检测的外周和局部免疫应答。图10A提供了实验时间表。肌养蛋白缺失-无效狗(Grinch和Ned)各自接受相等剂量($1x10^{12}$vg/kg)的AAV9-μ肌养蛋白(右)和AAV9-μ肌养蛋白相关蛋白(左)到它们胫骨前肌区室中的IM注射。如图10B中所示的,在注射前、注射后2、4、6、8周收集PBMCs,并用跨越整个μ肌养蛋白(库(Pool)A-D)和μ肌养蛋白相关蛋白(库E-J)肽序列的合成肽进行培养,而疫苗肽和PMA/Ion充当阳性对照。阳性结果被解释为≥5斑点形成单位(SFU)/1E5 PBMCs(虚线)。图10C显示了在注射后4周收集的肌肉活组织检查针对Utro N(上排)和肌养蛋白(下排)的免疫荧光(绿色)染色。具有红色框的插图-供参考,对于肌养蛋白染色绿色的正常肌肉的外观。图10D显示了注射后4周收集的肌肉活组织检查的代表性H&E。

[0032] 图11提供了微-肌养蛋白相关蛋白的概述。显示了缺失连接(剪接点)的位点。铰链1、2和4标记为H1、H2和H4。SR1、SR2、SR3和SR22对应于全长肌养蛋白相关蛋白的TH1、TH3、TH3和TH22。

[0033] 图12A至图12D提供了杂合测试的设计:垂直活动监控和全肢力测试。图12A提供了显示实验时间表的示意图。12B显示在全肢力测试前c57(n=11)和*mdx*小鼠(n=12)之间的旷场(open field)笼中垂直活动的定量在c57和*mdx*小鼠之间未显示显著差异(P>0.05 Mann-Whitney检验)。图12C显示对于c57(n=11)和*mdx*小鼠(n=17)两者在一系列七次拉伸中进行了全肢力测试。C57小鼠用正方形表示,且*mdx*小鼠用圆表示。显示了对于每个系列拉伸的全肢力的分布。描绘了方程式(P<0.0001,双因素(2-way)ANOVA检验)。图12D提供了对于力测试后1小时的累积后垂直活动(post-vertical activity)的分析,显示在c57和*mdx*小鼠之间存在显著差异(P<0.0001,双因素ANOVA)。

[0034] 图13提供了对于全肢力测试之前5分钟(Pre),以及在力测试之后的最初5分钟和次5分钟的c57和*mdx*小鼠的动作的可视表示。线代表水平活动,且点代表垂直活动。

[0035] 图14A至图14G显示在有力的收缩后,BIO 14.6仓鼠骨骼肌的簇生肌细胞中的急性肌膜破坏。图14A提供了静脉内(*i.v.*)注射伊文思蓝(Evans blue)染料后72小时肌肉纤维中伊文思蓝染料和肌养蛋白复染的同时观察。在图14B中,如通过FITC滤器观察的,在来自图14A的切片中在所有伊文思蓝阳性纤维中都明显完全不存在肌养蛋白染色。图14C显示,在单次自发跑步的8小时内,肌细胞损伤导致大多数伊文思蓝阳性纤维中的肌养蛋白的丧

失。图14D显示了在加长速率为每秒0.75肌肉长度的强直性挛缩后3小时的胫骨前肌肌肉。通过普施安(procion)橙色染料吸收显示了急性损伤;肌养蛋白复染表明这些纤维中不存在完全的膜分解和非特异性蛋白酶解。图14E显示了在转轮(running wheel)运动1小时之后的肌肉冷冻切片的肌养蛋白复染。图14F显示了在与图14E中相同的切片中的伊文思蓝染料荧光。图14G提供连续冷冻切片相同区域的茜素红S染色。原始放大率:图14A、14B和14E至14G,100X;图14C和图14D,200X。

[0036] 图15提供了肌养蛋白结合蛋白复合体糖基化酶,配体,辅肌动蛋白血影蛋白超家族,发动蛋白和肌联蛋白-遮蔽蛋白(obsurin)超家族的分析。可以在实施例3中找到关于所述图的更多讨论。

[0037] 图16A和图16B提供了使用衍生自人β2-血影蛋白(3EDV,图16A)和人网蛋白(5J1G)的模板的人肌养蛋白相关蛋白的相邻三股螺旋重复序列的模型。可以在实施例3中找到更多细节。

[0038] 图17提供了微-肌养蛋白相关蛋白的概述。所述微-肌养蛋白相关蛋白将非结构的(unstructured)、富含脯氨酸的螺旋间“铰链-2”结构域(H2)相对于全长肌养蛋白相关蛋白的最后一个三股螺旋重复序列(第22个)并列。R1、R2、R3和R22对应于全长肌养蛋白相关蛋白的TH1、TH3、TH3和TH22。可以在实施例3的F部分中找到更多细节。

[0039] 图18提供了如实施例3中所述的最优化的微-肌养蛋白相关蛋白的表达水平。

[0040] 图19显示了如实施例3中所述的在GRMD狗中注射后六周的肌膜中稳健剂量依赖性μ肌养蛋白相关蛋白表达和野生型肌膜蛋白聚糖表达水平的稳定化。

[0041] 图20显示了如实施例2和3中所讨论的在注射的GRMD狗中的μ肌养蛋白相关蛋白特异性T细胞应答。将注射后第5和8周收集的外周血单核细胞(PBMCs)用三个对应于AAV 9衣壳的合成肽库(A、B和C)以及五个跨越整个μ肌养蛋白相关蛋白肽序列的合成肽库(A-B、C-D、E-F、G-H、I-J)进行培养。通过计数每百万PBMCs的斑点形成单位来评估γ-干扰素的产生,而在注射的狗中针对AAV9衣壳或肌养蛋白相关蛋白衍生的肽库均没有高于背景的应答。右下,腺病毒-CMV-lacZ注射后的对照测定,其基于对阳性结果的最保守解释(虚线),显示了对Had5(1-4)和lacZ(5-8)肽库两者的阳性应答(星号)。

[0042] 图21显示了如实施例3的G部分中所讨论的在AAV-μ-U注射的*mdx*小鼠的承重肌肉中的79 kd条带。泳道1、2和4:非承重(例如屈肌)肌肉;3、5和7:承重(例如伸肌)肌肉;6肝;8 PBS-注射的*mdx*肌肉;9,分子量标准参照物。

[0043] 图22显示了鉴定作为微-肌养蛋白相关蛋白的N-末端部分的79 kd片段的蛋白印迹、免疫亲和纯化和LC/MS-MS的组合。框包含序列PPPPP,紧接在缺失连接上游的“铰链2”的一部分,如所示的。

[0044] 图23显示SH3结构域与来自两侧的相邻三股螺旋的相容氨基酸侧链形成多个高亲和力接触,一种具有传递纵向力和抵抗伸展的潜力的构型。

[0045] 图24A至图24C显示,在这个意义上,Anc80在心肌和骨骼肌的强转导的情况下实现了与AAV9类似的微-肌养蛋白相关蛋白的整体生物分布(参见图24A和图24B中的显微照片,以及图24C中的蛋白印迹),如在实施例3中所讨论的。

[0046] 图25显示了对于在以2.5×10^{12}vg/小鼠的相等剂量全身施用这些载体后在mdx小鼠群组(cohort)中μ肌养蛋白相关蛋白的生物分布的AAV9和Anc80的定性比较。显示了来自

每种载体的两只小鼠的多种肌肉的代表性蛋白印迹,从而证明了用两种载体对横纹肌的广泛且有效的转导。最上面的条带是μ肌养蛋白相关蛋白,如用特异性识别对应于蛋白N-末端的表位的多克隆抗体所标记的。样品加样对照是用对于蛋白黏着斑蛋白的抗体标记的最低条带。代表的横纹肌:膈肌、三头肌、四头肌、腓肠肌、腹壁、胸肌、胫骨前肌、心。

[0047] 图26A和图26B显示AAV9-μ肌养蛋白相关蛋白消除MuRF-1(+)、TUNEL(+)和中央成核肌纤维以及在*mdx*小鼠肌肉中。在图26A中,如所标记的图像代表作为新生儿用AAV9μU或PBS注射的8-周龄*mdx*小鼠。用MuRF-1和TUNEL进行的组织学染色分别用作活性蛋白酶解和凋亡的生物标志物。图26B提供了总结在*mdx*PBS-治疗的、*mdx*AAV9-μ肌养蛋白相关蛋白治疗的和c57野生型PBS-治疗的组(n=3)中的中央成核肌纤维、MuRF-1、TUNEL和胚胎阳性肌纤维的平均值和标准差的表。*mdx*肌肉纤维的中央成核现象表示出至少有一个以前的坏死发作继之以随着*mdx*小鼠达到8周龄而再生。

[0048] 图27提供了随机化至AAV9-μ肌养蛋白相关蛋白的GRMD狗的正常生长,作为抗免疫介导的肌炎的证据。在没有免疫抑制的情况下随机化至最高剂量($1\times10^{13.5}$vg/kg)的AAV9-cU(犬μ-肌养蛋白相关蛋白)的狗以及相关对照(包括随机化至PBS的同窝出生的仔畜和其他同窝出生的仔畜携带者雌性和非同窝出生的仔畜GRMD雄性和雌性)的各个重量。为比较起见,还包括以前报道的接受AAV9-hD(人μ-肌养蛋白)的GRMD雌性的相关重量,其紧接在安死术前显示快速体重减轻,且尸体剖检显示全身肌炎的病征(Kornegay, J.N.等人, Widespread muscle expression of anAAV9 human mini-dystrophin vector after intravenous injection in neonatal dystrophin-deficient dogs. Mol Ther, 2010. 18(8): p. 1501-8.)。

[0049] 图28显示了来自比较体内微-肌养蛋白相关蛋白和毫微-肌养蛋白相关蛋白的稳定性的研究的蛋白印迹。将Mdx小鼠用编码微-肌养蛋白相关蛋白或毫微-肌养蛋白相关蛋白的AAV载体进行注射,继之以在肌肉组织中检测蛋白(包括79 kd N-末端亚片段)。

[0050] 图29A和图29B显示了通过剪接全长肌养蛋白的TH 1和TH 20形成的突变杂合螺旋重复序列的3-D再现的替代视点(viewpoints)。B反平行螺旋中的剪接定位于与平行A和C螺旋上的W残基(黄色)所成的平面中。

[0051] 图30A和图30B显示了通过剪接全长肌养蛋白相关蛋白的TH 1和TH 18形成的突变杂合螺旋重复序列的3-D再现的替代视点。B反平行螺旋中的剪接定位于与平行A和C螺旋上的W残基(黄色)所成的平面中。

[0052] 发明详述

[0053] 提供了组合物和方法,其用于通过肌养蛋白超家族三重剪接突变蛋白、编码其的核酸序列或包含这种核酸序列的载体治疗肌营养不良(MD),包括杜兴肌营养不良(DMD)和贝克尔肌营养不良(BMD)和其他相关疾病。

[0054] 进行了肌养蛋白及其超基因家族的不寻常的分子进化分析,从而导致范式转移(paradigm shift)。发明人的分析表明,肌养蛋白的主要作用是强的、牵着性(tethering)、不可伸展的“支柱”样杆的作用,而不是如以前所推理的减震体(shock absorber);其长度在重要性上次于其强度;其强度在总共占蛋白的一级结构的80%的整个杆结构域上取决于相邻的“三股螺旋”、“血影蛋白样”重复序列之间边界处的氨基酸的相互作用;肌养蛋白仅与其最弱的环节一样强;进化已相对于消弱蛋白的内部缺失进行了选择。因此,为了在不削

弱蛋白的情况下缩短,要穿过具有最大序列保守性的中央三股螺旋结构域在多个可以在的折叠的蛋白中对齐的位点切割多肽的编码序列(例如,在肌养蛋白中穿过将疏水核中的相互作用色氨酸残基二等分的2-维平面)。这是对于所有血影蛋白样三股螺旋重复序列的隐蔽马尔科夫模型中最强的要素,并且适用于肌养蛋白、肌养蛋白相关蛋白和spectroplakins(MACF、dystonin等)的大多数重复序列。

[0055] I. 肌养蛋白、肌养蛋白相关蛋白等。

[0056] 如本文所用的,肌养蛋白超家族蛋白是指包含由三个α-螺旋组成并且在蛋白中作为单拷贝或以多个重复序列的串联排列出现的"血影蛋白样"和"杆样"结构域的蛋白,例如肌养蛋白,肌养蛋白相关蛋白,α-辅肌动蛋白,α-血影蛋白,β-血影蛋白,或血影蛋白家族的其他成员,斑蛋白,spectraplakins(即,spectroplakins)。三股-α-螺旋结构域包含由非螺旋接头连接的两个类似地(螺旋A和C)和一个相反地(螺旋B)定向的α-螺旋。结构域的疏水核中的许多芳族残基一般是保守的。参见,例如,Parry DA等人,Analysis of the three-alpha-helix motifin the spectrin superfamily of proteins. Biophys J. 1992年4月;61(4):858-67;和Djinovic-Carugo K等人, The spectrin repeat: a structural platform for cytoskeletal protein assemblies.FEBS Lett. 2002年2月20日;513(1):119-23。这种重复序列的例子可以在图2A至2E中找到。在某些实施方案中,全长肌养蛋白超家族蛋白是指肌养蛋白超家族蛋白或其同工型,其可以存在于健康对照中。在某些实施方案中,全长肌养蛋白超家族蛋白是指被本领域技术人员认为是规范序列的肌养蛋白超家族蛋白或其同工型。这种规范序列是可获得的,例如,在www.uniprot.org。

[0057] 血影蛋白是嵌入真核细胞质膜的细胞内侧的细胞骨架蛋白。血影蛋白形成五角形或六角形排列,从而形成支架并在维持质膜完整性和细胞骨架结构中起重要作用。参见,例如,Huh GY等人, Calpain proteolysis of alphaII-spectrin in the normal adult human brain. Neurosci Lett. 2001年12月4日;316(1):41-4。所述超家族中的蛋白可以通过两个特征来定义:(1)N-末端肌动蛋白结合结构域;和(2)α-螺旋血影蛋白重复序列部分。参见,例如,Röper K等人,The 'spectraplakins': cytoskeletalgiants with characteristics of both spectrin and plakin families. J Cell Sci. 2002年11月15日;115(Pt 22):4215-25。血影蛋白超家族的成员包括但不限于,α-辅肌动蛋白(例如,α-肌动蛋白-1,参见,例如,UniProtKB - P12814和www.genecards.org/cgi-bin/carddisp.pl?gene=ACTN1;α-肌动蛋白-2,参见,例如,UniProtKB - P35609和www.genecards.org/cgi-bin/carddisp.pl?gene=ACTN2;α-肌动蛋白-3,参见,例如,UniProtKB - Q08043和www.genecards.org/cgi-bin/carddisp.pl?gene=ACTN3;和α-肌动蛋白-4,参见,例如,UniProtKB - O43707和www.genecards.org/cgi-bin/carddisp.pl?gene=ACTN4,其各自整体引入作为参考)、α-血影蛋白(例如,血影蛋白α链,红细胞1,参见,例如,UniProtKB - P02549和www.genecards.org/cgi-bin/carddisp.pl?gene=SPTA1;和血影蛋白α链,非红细胞1,参见,例如,UniProtKB - Q13813和www.genecards.org/cgi-bin/carddisp.pl?gene=SPTAN1,其各自整体引入作为参考)、β-血影蛋白(例如,血影蛋白β链,红细胞,参见,例如,UniProtKB - P11277和www.genecards.org/cgi-bin/carddisp.pl?gene=SPTB;和血影蛋白β链,非红细胞1,参见,例如,UniProtKB - Q01082和www.genecards.org/cgi-bin/carddisp.pl?gene=SPTBN1,其各自整体引入作为参考)、肌

养蛋白和肌养蛋白相关蛋白。

[0058] 斑蛋白是与细胞骨架元件和连接复合体结合的细胞连接(cytolinker)蛋白。参见,例如,Leung Cl等人,Plakins: a family of versatilecytolinker proteins. Trends Cell Biol. 2002年1月;12(1):37-45。已经鉴定了七个斑蛋白家族成员:桥粒斑蛋白(UniProtKB - P15924和www.genecards.org/cgi-bin/carddisp.pl?gene=DSP,其包括其中列出的序列,在此整体附上)、网蛋白(UniProtKB - P15924和www.genecards.org/cgi-bin/carddisp.pl?gene=PLEC,其包括其中列出的序列,在此整体附上)、大天疱疮抗原1(BPAG1,Dystonin)(UniProtKB - Q03001和www.genecards.org/cgi-bin/carddisp.pl?gene=DST,其包括其中列出的序列,在此整体附上)、微管-肌动蛋白交联因子(MACF)(UniProtKB - Q9UPN3和www.genecards.org/cgi-bin/carddisp.pl?gene=MACF1,其包括其中列出的序列,在此整体附上)、外被斑蛋白(UniProtKB - Q92817和www.genecards.org/cgi-bin/carddisp.pl?gene=EVPL,其包括其中列出的序列,在此整体附上)、斑周蛋白(UniProtKB - O60437和www.genecards.org/cgi-bin/carddisp.pl?gene=PPL,其包括其中列出的序列,在此整体附上)和epiplakin(UniProtKB - P58107和www.genecards.org/cgi-bin/carddisp.pl?gene=EPPK1,其包括其中列出的序列,在此整体附上)。所述蛋白家族由斑蛋白结构域和/或斑蛋白重复序列结构域(PRD)的存在定义。除这两个结构域外,斑蛋白还携带在一些但不是全部成员中共有的其他结构域:肌动蛋白结合结构域(ABD)、卷曲螺旋杆、含血影蛋白重复序列的杆和微管结合结构域。

[0059] spectraplakins属于血影蛋白和斑蛋白超家族两者,且是异常长的细胞内蛋白,其具有罕见的结合以下所有三种细胞骨架元件的能力:肌动蛋白、微管和中间丝。spectraplakins对于组织完整性和功能至关重要,与单个细胞骨架元件一起起作用以及使这些元件协调。参见,例如,如上文所引用的Röper K等人和Huelsmann S等人,Spectraplakins. Curr Biol. 2014年4月14日;24(8):R307-8. doi: 10.1016/j.cub.2014.02.003。spectraplakins的成员包括但不限于如上文所述的BPAG1和MACF。肌养蛋白通过F-肌动蛋白将细胞外基质锚定在细胞骨架上,并且是肌营养不良蛋白聚糖的配体。肌养蛋白结合性糖蛋白复合体的组分在神经肌肉接头(NMJ)以及周围和中枢神经系统的多种突触中累积,并在稳定肌膜中具有结构功能。它也与信号传导事件和突触传递牵连。参见,例如,www.uniprot.org/uniprot/P11532和www.genecards.org/cgi-bin/carddisp.pl?gene=DMD&keywords=dystrophin。存在肌养蛋白的10种同工型。同工型4被认为是规范序列,并且含有具有UniProtKB标识符P11532-1的氨基酸序列,其引入本文。其他同工型,包括具有UniProtKB标识符:P11532-2的同工型1、具有UniProtKB标识符:P11532-3的同工型2、具有UniProtKB标识符:P11532-4的同工型3、具有UniProtKB标识符:P11532-5的同工型5、具有UniProtKB标识符:P11532-6的同工型6、具有UniProtKB标识符:P11532-7的同工型7、具有UniProtKB标识符:P11532-8的同工型8、具有UniProtKB标识符:P11532-9的同工型9和具有UniProtKB标识符:P11532-10的同工型10,每种同工型的序列都引入本文。如本文所用的术语“全长肌养蛋白”可以指任何肌养蛋白同工型。在某些实施方案中,术语“全长肌养蛋白”是指同工型4(Dp427)。已在多种生物中鉴定了肌养蛋白的同源物(homolog),包括小鼠(UniProt P11531)、大鼠(UniProt P11530)和狗(UniProt O97592)。编码Dp427或肌养蛋白的任何其他同工型或同源物的可能核酸序列是公众可获得的,参见,

例如,NCBI参考序列:NM_000109.3、NM_004006.2、NM_004009.3、NM_004010.3、NM_004011.3、NM_004012.3、NM_004013.2、NM_004014.2、NM_004015.2、NM_004016.,2NM_004017.2、NM_004018.2、NM_004019.2、NM_004020.3、NM_004021.2、NM_004022.2、NM_004023.2、NM_004007.2、XM_006724468.2、XM_006724469.3、XM_006724470.3、XM_006724473.2、XM_006724474.3、XM_006724475.2、XM_011545467.1、XM_011545468.2、XM_011545469.1、XM_017029328.1、XM_017029329.1、XM_017029330.1和XM_017029331.1,其每一个都引入本文。可以通过用于反向翻译的工具,例如,www.ebi.ac.uk/Tools/st/、www.ebi.ac.uk/Tools/st/emboss_transeq/、www.ebi.ac.uk/Tools/st/emboss_sixpack/、www.ebi.ac.uk/Tools/st/emboss_backtranseq/和www.ebi.ac.uk/Tools/st/emboss_backtranambig/,产生编码Dp427或肌养蛋白的任何其他同工型或同源物的额外序列。此外,可以对编码序列进行密码子优化以在主体例如人、小鼠、大鼠或狗中表达。

[0060] Dp427的长度为3685个氨基酸。同样,参见,例如,US7892824B2和GenBank:AAA53189.1。Dp427的N-末端240个氨基酸折叠成两个钙调理蛋白同源结构域(CH1&2),它们具有与细胞骨架肌动蛋白纤丝高亲和力结合的能力。从约氨基酸340到3040的中央区域由24个串联链接的可通过隐蔽马尔科夫模造鉴定为三股螺旋的结构域组成(TH1-24),其与杆样蛋白血影蛋白和α-辅肌动蛋白的结晶重复序列具有可测量的结构同源性。从3057到3352的区域包含模块式多结构域球状区域(WW-EF-ZZ),其对以β-肌营养不良蛋白聚糖为中心的蛋白的跨膜复合体具有高亲和力。从3353到3685的最后的C-末端区域具有看上去消耗性的对于蛋白小肌营养蛋白和肌养蛋白结合蛋白的高亲和力结合结构域。Dp427已被模造为横纹肌细胞皮质中的杆状蛋白,通过N-末端钙调理蛋白同源结构域与细胞骨架F-肌动蛋白的最外缘结合,并通过C-末端WW-EF-ZZ结构域与DGC的跨膜成员结合。中央杆结构域由与三股螺旋重复序列铰链1-4具有低水平同源性的24个结构域组成。在等位疾病贝克尔MD(BMD)中,血影蛋白样重复序列的内部缺失通常与疾病进展的较慢速度有关。用于DMD的基因治疗的主要挑战是安全、有效且持久地取代大多数骨骼和心肌细胞中的Dp427。理想地,所述取代将与Dp427的功能性相适应;人们担心,显著小于427 kd的蛋白可能只会将DMD转化为严重的BMD表型。在图1A中可以找到Dp427的举例说明。可以通过常规方法执行隐蔽马尔科夫模造,而其参数可以由本领域技术人员调整。参见,例如,en.wikipedia.org/wiki/Hidden_Markov_model。在这些重组蛋白的合成编码序列在AAV载体的克隆能力内的意义上,19-20个邻接的TH结构域和3353-3685的区域的缺失产生了“AAV-大小的”小型化肌养蛋白。所有这种重组蛋白共享Dp427中杆长度的1/5至1/6的杆样结构域,从而引起了所述缩短逐渐损害重组蛋白“吸收”与24-重复序列全长蛋白一样多的“震动”的能力的担心。

[0061] 肌养蛋白相关蛋白与肌养蛋白基本同源,具有在杆结构域中中存在的显著趋异,在所述杆结构域处肌养蛋白相关蛋白缺乏重复序列15和19以及两个铰链区(参见,例如,Love等人, Nature 339:55 [1989];Winder等人, FEBS Lett., 369:27 [1995];www.uniprot.org/uniprot/P46939;和www.genecards.org/cgi-bin/carddisp.pl?gene=UTRN&keywords=Utrophin)。发现了四种肌养蛋白相关蛋白的同工型。同工型1被认为是规范序列,且含有具有UniProt标识符P46939-1的氨基酸序列,其引入本文。其他同工型包括具有UniProt标识符P46939-2的同工型2;具有UniProt标识符P46939-3的同工型UP71;和具有UniProt标识符P46939-4的同工型Up140。全长肌养蛋白相关蛋白可能是指任何肌养蛋白

相关蛋白同工型。在某些实施方案中，全长肌养蛋白相关蛋白是指肌养蛋白相关蛋白同工型1，其含有22个血影蛋白样重复序列(SR1-SR22，或TH1-TH22)和两个铰链区。已经在多种生物中鉴定了肌养蛋白相关蛋白的同源物，包括小鼠(Genbank检索号Y12229和UniProt E9Q6R7)、大鼠(Genbank检索号AJ002967和UniProt G3V7L1)和狗(GenBank检索号NW-139836)。可以使用任何合适的方法将这些或额外的同源物的核酸序列与人肌养蛋白相关蛋白的核酸序列进行比较。编码肌养蛋白相关蛋白同工型1或任何其他同工型或任何同源物的核酸序列是可获得的。参见，例如，NCBI参考序列:NM_007124.2、XM_005267127.4、XM_005267130.2、XM_005267133.2、XM_006715560.3、XM_011536101.2、XM_011536102.2、XM_011536106.2、XM_011536107.2、XM_011536109.2、XM_017011243.1、XM_017011244.1、XM_017011245.1;Genbank检索号X69086和GenBank检索号AL357149，其每一个都引入本文。可以通过用于反向翻译的工具，例如，www.ebi.ac.uk/Tools/st/、www.ebi.ac.uk/Tools/st/emboss_transeq/、www.ebi.ac.uk/Tools/st/emboss_sixpack/、www.ebi.ac.uk/Tools/st/emboss_backtranseq/和www.ebi.ac.uk/Tools/st/emboss_backtranambig/，产生编码肌养蛋白相关蛋白同工型1或任何其他同工型或任何同源物的额外序列。此外，可以对编码序列进行密码子优化以在主体例如人、小鼠、大鼠或狗中表达。

[0062] 在一个方面，本文提供了肌养蛋白超家族三重剪接突变蛋白。在一个实施方案中，三重剪接突变蛋白包含多个螺旋重复序列的内部缺失，以及通过连接全长肌养蛋白超家族蛋白的螺旋重复序列的部分形成的杂合螺旋结构域。在某些实施方案中，肌养蛋白超家族三重剪接突变蛋白含有具有与螺旋A'的C-末端部分融合的螺旋A的N-末端部分的第一螺旋，包含与螺旋B的末端C-末端部分融合的螺旋B'的N-末端部分的第二螺旋，以及包含与螺旋C'的C-末端部分融合的螺旋C的N-末端部分的第三螺旋，其中螺旋A、B和C存在于第一三股螺旋重复序列中，且螺旋A'、B'和C'存在于天然肌养蛋白超家族蛋白中的第二三股螺旋结构域中。在某些实施方案中，第一和第二三股螺旋结构域是不相邻的，且因此提供了突变肌养蛋白超家族蛋白，其具有杂合三股螺旋结构域和天然肌养蛋白超家族蛋白中存在的一个或多个三股螺旋结构域的缺失。因此，三重剪接突变蛋白中螺旋重复序列的总数目选自3至比全长肌养蛋白超家族蛋白的螺旋重复序列数目小1的任何整数，例如，4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22和23。在某些实施方案中，本文提供了具有在全长肌养蛋白的至少螺旋重复序列3至螺旋重复序列21中的缺失的突变肌养蛋白蛋白。在再另一个实施方案中，突变蛋白具有在全长肌养蛋白的至少螺旋重复序列3至23中的缺失。在再进一步的实施方案中，突变蛋白具有在全长肌养蛋白中的至少螺旋重复序列2至螺旋重复序列19中的缺失。在某些实施方案中，提供了突变肌养蛋白相关蛋白，其具有在全长肌养蛋白的至少螺旋重复序列3至螺旋重复序列10中的缺失。在进一步的实施方案中，突变肌养蛋白相关蛋白具有在全长肌养蛋白相关蛋白的至少螺旋重复序列2至螺旋重复序列17中的缺失。

[0063] 如本文所用的，术语“三股螺旋结构域(triple helical domain)”、“三股螺旋结构域(triple helix domain)”、“三股螺旋重复序列(triple helical repeat)”、“三股螺旋重复序列(triple helix repeat)”和“TH”是可互换的，并且是指肌养蛋白超家族蛋白的杆样和血影蛋白样重复序列，其由三个α螺旋组成，即，由非螺旋接头连接的两个类似地(螺旋A和C)和一个相反地(螺旋B)定向的α-螺旋。那些重复序列可以通过隐蔽马尔科夫模造鉴

定。这种重复序列的例子可以在图2A至图2E中找到。杂合三股螺旋重复序列由在垂直于其长轴将螺旋重复序列二等分的平面上剪接的两个螺旋重复序列形成。这种平面在下文进一步讨论并且在实施例以及图2F中例示。在图1中提供了形成这种杂合三股螺旋重复序列的举例说明，而在图1C和图1D中显示了杂合三股螺旋结构域(本文也记录为杂合三股螺旋重复序列)的举例说明。如本文所用的，“剪接点”表示在核酸序列、氨基酸序列或具有二级、三级或四级结构的蛋白中的位置，在所述位置处内部缺失在全长肌养蛋白超家族蛋白或在形成杂合三股螺旋结构域的两个三股螺旋重复序列中的任何一个中开始或结束；或在所述位置处在全长蛋白中其对应序列彼此不直接相邻的序列但在杂合三股螺旋结构域或在肌养蛋白超家族三重剪接突变蛋白中连接。术语“直接相邻”意指两个序列、结构域或重复序列分别未被任何其他序列、结构域或重复序列分开。术语“连接”、“再连接(re-join)”或其任何语法变化形式表示两个序列、结构域或重复序列变得直接相邻。术语“形成”或其任何语法变化形式是指剪接或连接序列。序列或序列中位置的对应可以通过序列比对或隐蔽马尔科夫模型(HMM)确定。

[0064] 如本文所用的，“N-末端部分”是指对于所选择的螺旋的剪接点的氨基末端侧的氨基酸序列。在某些实施方案中，所选择的螺旋的“N-末端部分”是指从最初Met直到在剪接点之前(在N-末端侧)的最后一个氨基酸序列的全长氨基酸序列。在某些实施方案中，在N-末端部分中可以存在氨基酸取代、缺失、截短和/或插入。在某些实施方案中，这种取代是保守氨基酸改变。在某些实施方案中，缺失、截短或插入的长度为1-5个氨基酸，其不影响螺旋的折叠。

[0065] 术语“C-末端部分”是指对于所选择的螺旋的剪接点的羧基末端侧的氨基酸序列。在某些实施方案中，所选择的螺旋的“C-末端部分”是指从剪接点之后(在C-末端侧)的第一个氨基酸序列起的全长氨基酸序列。在某些实施方案中，在N-末端部分中可以存在氨基酸取代、缺失、截短和/或插入。在某些实施方案中，这种取代是保守氨基酸改变。在某些实施方案中，缺失、截短或插入的长度为1-5个氨基酸，其不影响螺旋的折叠。

[0066] 例如，在某些实施方案中，三重突变杂合螺旋结构域具有通过连接不相邻的螺旋结构域的区段而形成的三个剪接螺旋，其中每个螺旋包含天然肌养蛋白超家族蛋白的螺旋重复序列中螺旋的N-末端部分和C-末端部分。因为被连接以形成突变连接(mutant junction)的三股螺旋结构域各自具有平行的A和C螺旋以及反平行的B螺旋，所以三股螺旋突变体中的A和C螺旋的N-末端部分来自天然肌养蛋白超家族蛋白中的相同的三股螺旋重复序列，而这些螺旋的C-末端部分来自天然肌养蛋白超家族蛋白中的另一个三股螺旋重复序列。作为连接的定位以形成三股螺旋突变结构域的结果，N-末端部分和C-末端部分可以具有不同的长度，但是一起形成突变三股螺旋结构域的螺旋。

[0067] 在整个所述说明书中，使用序数，例如“第一”、“第二”、“第三”、“第四”或术语“额外的”作为参考术语，以区别组合物和方法的各种形式和组分。除非指定，否则如果使用序数表示TH、氨基酸序列或核酸序列，则所述数字在氨基酸序列或蛋白中从N-末端到C-末端或在核酸序列中从5’至3’计数。

[0068] 如本文所用的，除非特别指定，否则蛋白重复序列再加上数字是指从N-末端起计数的参考蛋白或参考氨基酸序列中所有重复序列中的重复序列序数。例如，三股螺旋结构域2是指图1A所举例说明的TH2。当重复序列和参考序列是多核苷酸时，除非特别指定，否则

重复序列的序数是从5'端到3'端计数的。

[0069] 在某些实施方案中,肌养蛋白超家族三重剪接突变蛋白包含一个或多个N-末端螺旋重复序列、杂合三股螺旋重复序列和一个或多个C-末端螺旋重复序列。在某些实施方案中,肌养蛋白超家族三重剪接突变蛋白包含杂合三股螺旋重复序列和C-末端螺旋重复序列。三重剪接突变蛋白中一个或多个螺旋重复序列的总数目选自3至比全长肌养蛋白超家族蛋白的螺旋重复序列数目小1的任何整数,例如,4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22和23。在某些实施方案中,与突变蛋白中的一个或多个N-末端螺旋重复序列相对应的全长肌养蛋白超家族蛋白中的一个或多个序列与形成杂合三股螺旋重复序列的两个螺旋重复序列中的第一个相对应的序列直接相邻。在某些实施方案中,与突变蛋白中的一个或多个C-末端螺旋重复序列相对应的全长肌养蛋白超家族蛋白中的一个或多个序列与形成杂合三股螺旋重复序列的两个螺旋重复序列中的第二个相对应的序列直接相邻。在进一步的实施方案中,一个或多个N-末端螺旋重复序列可包括TH1、TH2、TH3、TH4、TH5、TH6、TH7、TH8、TH9、TH10、TH11、TH12、TH13、TH14、TH15、TH16、TH17、TH18和TH19。在再进一步的实施方案中,一个或多个C-末端螺旋重复序列可以包括TH1、TH2、TH3、TH4、TH5、TH6、TH7、TH8、TH9、TH10、TH11、TH12、TH13、TH14、TH15、TH16、TH17、TH18和TH19,其每个TH序数均从C-末端起计数。另外,突变蛋白中可能只有一个螺旋重复序列,其中螺旋重复序列是全长肌养蛋白超家族蛋白的螺旋结构域1和最后一个螺旋结构域形成的杂合三股螺旋结构域。此外,突变蛋白中可能有两个螺旋重复序列。这种2-TH突变蛋白可以包含全长蛋白的螺旋重复序列1,以及由全长蛋白的螺旋重复序列2和最后一个螺旋重复序列形成的一个杂合三股螺旋结构域。在另一个实施方案中,2-TH突变蛋白可包含全长蛋白中的最后一个螺旋重复序列,以及由全长蛋白的螺旋重复序列1和倒数第二个螺旋重复序列形成的一个杂合三股螺旋结构域。如本文所用的,术语"肌养蛋白超家族三重剪接突变蛋白"、"三重剪接突变蛋白"和"突变蛋白"可互换使用。同样,除了在杂合三股螺旋重复序列中的一个或多个剪接点之外,突变蛋白中的重复序列和序列如它们在全长肌养蛋白超家族蛋白中那样分别和相同的重复序列和序列直接相邻。

[0070] 在某些实施方案中,肌养蛋白超家族三重剪接突变蛋白可在C-末端被截短,例如,达1至500个氨基酸的任何整数。这种截短不包含任何三股螺旋重复序列。在某些实施方案中,这种截短可发生在与全长肌养蛋白超家族蛋白的外显子的开始或结束相对应的位置。

[0071] 本文还提供了编码肌养蛋白超家族三重剪接突变蛋白的核酸序列。这种编码序列可以通过用于反向翻译的工具产生。此外,可以对编码序列进行密码子优化以在主体例如人、小鼠、大鼠或狗中表达。

[0072] 在一个实施方案中,肌养蛋白超家族突变蛋白是三重剪接突变肌养蛋白。在进一步的实施方案中,三重剪接突变肌养蛋白包含全长肌养蛋白的至少螺旋重复序列3至螺旋重复序列21中的缺失。在另一个实施方案中,三重剪接突变肌养蛋白包含全长肌养蛋白的至少螺旋重复序列3至螺旋重复序列23中的缺失。在再另一个实施方案中,突变肌养蛋白的一个或多个N-末端螺旋重复序列包含全长肌养蛋白中的螺旋重复序列1。突变蛋白的一个或多个C-末端螺旋重复序列包含全长肌养蛋白中的螺旋重复序列23和螺旋重复序列24。在进一步的实施方案中,突变蛋白的N-末端螺旋重复序列由全长肌养蛋白中的螺旋重复序列1组成。突变蛋白的C-末端螺旋重复序列由全长肌养蛋白中的螺旋重复序列23和螺旋重复

序列24组成。形成杂合三股螺旋重复序列的两个螺旋重复序列中的第一个是全长肌养蛋白中的螺旋重复序列2。形成杂合三股螺旋重复序列的两个螺旋重复序列中的第二个是全长肌养蛋白中的螺旋重复序列22。在再另一个实施方案中，三重剪接突变肌养蛋白包含全长肌养蛋白蛋白的至少螺旋重复序列2至螺旋重复序列19中的缺失。C-末端螺旋重复序列包含全长肌养蛋白中的螺旋重复序列21、22、23和24。形成杂合三股螺旋重复序列的两个螺旋重复序列中的第一个是全长肌养蛋白中的螺旋重复序列1，且形成杂合三股螺旋重复序列的两个螺旋重复序列中的第二个是全长肌养蛋白中的螺旋重复序列19。在某些实施方案中，三重剪接突变肌养蛋白可在N-末端进一步包含肌养蛋白同工型4的氨基酸(aa)约1至约aa 338。在某些实施方案中，三重剪接突变肌养蛋白可在C-末端进一步包含肌养蛋白同工型4的约aa 3041至约aa 3352、约aa 3041至约aa 3054、约aa 3041至约aa 3056、约aa 3041至约aa 3057、约aa 3041至约aa 3088、约aa 3041至约aa 3408或约aa 3041至约aa 3685。在某些实施方案中，可能存在三重剪接突变肌养蛋白的这些C-末端非-TH序列的进一步截短。如本文所用的，截短是指从C-末端开始的连续氨基酸的缺失。在某些实施方案中，这种截短可发生在与全长肌养蛋白的外显子的开始或结束相对应的位置。在某些实施方案中，截短的长度可以是1、2、3、4、5、约10、约15、约20、约30、约40、约50、约60、约70、约100、约150、约200、约250、约300、约400、约500或约600 aa。本文还提供了编码三重剪接突变肌养蛋白的核酸序列。这种编码序列可以通过用于反向翻译的工具产生。此外，可以对编码序列进行密码子优化以在主体例如人、小鼠、大鼠或狗中表达。

[0073] 在一个实施方案中，三重突变肌养蛋白是具有SEQ ID NO:1的氨基酸序列的毫微-肌养蛋白(也记录为n-肌养蛋白)。在再进一步的实施方案中，提供了具有SEQ ID NO:2的核酸序列的编码三重剪接突变肌养蛋白的序列。在一个实施方案中，对编码SEQ ID NO:1的核酸序列进行密码子优化以在主体中表达。在进一步的实施方案中，对编码SEQ ID NO:1的核酸序列进行密码子优化以在人中表达。用于密码子优化的常规工具是本领域技术人员可公开或商业获得的。参见，例如，FuglsangA(Codon optimizer: a freeware tool for codon optimization. Protein Expr Purif. 2003年10月;31(2):247-9,)www.genscript.com/codon-opt.html、www.thermofisher.com/us/en/home/life-science/cloning/gene-synthesis/geneart-gene-synthesis/geneoptimizer.html和www.idtdna.com/CodonOpt。

[0074] 在再另一个实施方案中，三重突变肌养蛋白是具有SEQ ID NO:13、14、15、16、17或18的氨基酸序列的毫微-肌养蛋白。在再进一步的实施方案中，本文提供了编码具有SEQ ID NO:13、14、15、16、17或18的氨基酸序列的三重剪接突变肌养蛋白的核酸序列。在某些实施方案中，将编码SEQ ID NO:SEQ ID NO:13、14、15、16、17或18的核酸序列密码子优化以在主体中表达。在进一步的实施方案中，对编码SEQID NO:SEQ ID NO:13、14、15、16、17或18的核酸序列进行密码子优化以在人中表达。

[0075] 在再另一个实施方案中，三重突变肌养蛋白是具有SEQ ID NO:22的氨基酸序列的毫微-肌养蛋白。

[0076] 如本文所用的术语“主体”意指雄性或雌性哺乳动物，包括人、兽医或农场动物、家畜或宠物以及通常用于临床研究的动物。在一个实施方案中，这些方法和组合物的主体是人。在一个实施方案中，这些方法和组合物的主体是出生前的、新生儿、婴儿、学步的幼儿、学龄前儿童、小学生、青少年、年轻成人或成人。新生儿人是指0到12个月龄的人；人学步的

幼儿具有1至3岁的年龄;人学龄前儿童具有3至5岁的年龄;人小学生具有5至12岁的年龄;人青少年具有12至18岁的年龄;人年轻成人具有18至21岁的年龄;而人成人具有超过18岁的年龄。“健康主体”是指没有疾病的主体。如本文所用的,术语“疾病”可以指DMD和/或BMD。在某些实施方案中,术语“疾病”可以指由异常肌养蛋白超家族蛋白引起的另一种疾病。在某些实施方案中,“疾病”是指冯维勒布兰德病。

[0077] 在一个实施方案中,肌养蛋白超家族突变蛋白是三重剪接突变肌养蛋白相关蛋白,且包含全长肌养蛋白相关蛋白的至少螺旋重复序列3和螺旋重复序列19中的缺失。在一个实施方案中,突变蛋白的一个或多个N-末端螺旋重复序列包含全长肌养蛋白相关蛋白中的螺旋重复序列1。突变蛋白的一个或多个C-末端螺旋重复序列包含全长肌养蛋白相关蛋白中的螺旋重复序列21和螺旋重复序列22。在进一步的实施方案中,突变蛋白的N-末端螺旋重复序列由全长肌养蛋白相关蛋白中的螺旋重复序列1组成。突变蛋白的C-末端螺旋重复序列由全长肌养蛋白相关蛋白中的螺旋重复序列21和螺旋重复序列22组成。形成杂合三股螺旋重复序列的两个螺旋重复序列中的第一个是全长肌养蛋白相关蛋白中的螺旋重复序列2。形成杂合三股螺旋重复序列的两个螺旋重复序列中的第二个是全长肌养蛋白相关蛋白中的螺旋重复序列20。在再另一个实施方案中,肌养蛋白超家族突变蛋白是三重剪接突变肌养蛋白相关蛋白且包含全长肌养蛋白相关蛋白的至少螺旋重复序列2至螺旋重复序列17中的缺失。突变蛋白的C-末端螺旋重复序列包含全长肌养蛋白相关蛋白中的螺旋重复序列19、20、21和22。形成杂合三股螺旋重复序列的两个螺旋重复序列中的第一个是全长肌养蛋白相关蛋白中的螺旋重复序列1,且形成杂合三股螺旋重复序列的两个螺旋重复序列中的第二个是全长肌养蛋白相关蛋白中的螺旋重复序列18。在某些实施方案中,三重剪接突变肌养蛋白相关蛋白可在N-末端进一步包含肌养蛋白相关蛋白同工型1的约aa 1至约aa 311。在某些实施方案中,三重剪接突变肌养蛋白相关蛋白可在C-末端进一步包含肌养蛋白相关蛋白同工型1的约aa 2797至约2811、约aa 2797至约2845、约aa 2797至约3124、约aa 2797至约3134、约aa 2797至约3165、约aa 2797至约3168或约aa 2797至约3433。在某些实施方案中,可能存在三重剪接突变肌养蛋白相关蛋白的这些C-末端非-TH序列的进一步截短。在某些实施方案中,这种截短可发生在与全长肌养蛋白相关蛋白的外显子的开始或结束相对应的位置。在某些实施方案中,截短的长度可以是1、2、3、4、5、约10、约15、约20、约30、约40、约50、约60、约70、约100、约150、约200、约250、约300、约400、约500或约600 aa。本文还提供了编码三重剪接突变肌养蛋白相关蛋白的核酸序列。这种编码序列可以通过用于反向翻译的工具产生。此外,可以对编码序列进行密码子优化以在主体例如人、小鼠、大鼠或狗中表达。

[0078] 在一个实施方案中,三重突变肌养蛋白相关蛋白是包含以下氨基酸序列的毫微-肌养蛋白相关蛋白(也记录为具有或不具有破折号的n-肌养蛋白相关蛋白或n-U):SEQ ID NO:3。在进一步的实施方案中,本文提供了编码三重剪接突变肌养蛋白相关蛋白并且包含SEQ ID NO:4的核酸序列的序列。在某些实施方案中,三重突变肌养蛋白相关蛋白是具有SEQ ID NO:5的氨基酸序列的毫微-肌养蛋白相关蛋白(也记录为具有或不具有破折号的n-肌养蛋白相关蛋白或n-U)。在进一步的实施方案中,本文提供了编码SEQ ID NO:5并具有SEQ ID NO:6的核酸序列或与SEQ ID NO:6约95%至约99%相同的核酸序列的序列。在某些实施方案中,三重突变肌养蛋白相关蛋白是具有SEQ ID NO:7的氨基酸序列的毫微-肌养蛋

白相关蛋白。在进一步的实施方案中，本文提供了编码SEQ ID NO:7并具有SEQ ID NO:8的核酸序列或与SEQ ID NO:8约95％至约99％相同的核酸序列的序列。在一个实施方案中，对编码SEQ ID NOs:3、5或7的核酸序列进行密码子优化以在主体中表达。在某些实施方案中，三重突变肌养蛋白相关蛋白是具有SEQ ID NO:20的氨基酸序列的毫微-肌养蛋白相关蛋白。在进一步的实施方案中，本文提供了编码SEQ ID NO:20并具有SEQ ID NO:19的核酸序列或与SEQ ID NO:19约95％至约99％相同的核酸序列的序列。在一个实施方案中，对编码SEQ ID NOs:3、5、7或20的核酸序列进行密码子优化以在主体中表达。在进一步的实施方案中，对编码SEQID NOs:3、5、7或20的核酸序列进行密码子优化以在人中表达。在再另一个实施方案中，三重突变肌养蛋白相关蛋白是包含SEQ ID NO:21的氨基酸序列的毫微-肌养蛋白相关蛋白。

[0079] 在某些实施方案中，所述公开内容包括以“三重剪接的”为特征的人毫微-肌养蛋白相关蛋白和人毫微-肌养蛋白两者的氨基酸和所有编码合成核酸序列。在某些实施方案中，杂合三股螺旋可以将三股螺旋重复序列2的中间与倒数第三个三股螺旋重复结构域(在肌养蛋白相关蛋白中的22个中的＃20，在肌养蛋白中的24个中的＃22)的中间连接，从而在两种重组蛋白中都给出总计四个重复结构域，如本文所述并如SEQ ID NOs:1和3的特征所举例说明的。这些氨基酸序列定义了可以在单个AAV载体基因组的编码能力内编码的最大强度的重组蛋白。参见图4和实施例2和3。

[0080] 肌养蛋白的力学生物学(mechanobiology)鲜为人知，但是对蛋白的生理作用的间接研究提示蛋白的杆结构域可能在肌肉收缩过程中被纵向负荷。本文提供的实施例为此提供了最早的直接证据。在表达微-肌养蛋白相关蛋白的幼年和骨骼成熟的*mdx*(肌养蛋白无效)小鼠之间的比较中，使用针对重组蛋白N-末端的抗体进行的蛋白印迹分析揭示了单个剪接点位置处杆破坏的证据，因为肌肉成熟和肌球蛋白同工型转换的过程增加穿过肌肉膜的机械负荷。这强有力地支持了以下假设:杆结构域的强度在单个剪接点的精确位置处受到损害。构成毫微-肌养蛋白相关蛋白和毫微-肌养蛋白开发的基础的设计原理通过消除杆的不相容亚结构域的并列补偿了所述以前的缺点。

[0081] A. 螺旋中的剪接点

[0082] 杂合三股螺旋重复序列由如图1A-图1D和图2F所示的在垂直于其长轴将螺旋重复序列二等分的平面上剪接的两个螺旋重复序列组成。反平行的“B”螺旋中剪接点的选择是通过间接方式得出的。仅已确定了对于肌养蛋白三股螺旋TH1的单个重复序列的一种X-射线晶体结构(即，迄今为止没有对于TH2-24的结构)。肌养蛋白的相邻三股螺旋比血影蛋白和α-辅肌动蛋白的那些更加重叠，从而在纵向承重过程中稳定了螺旋，可能部分解释了结构信息的缺乏，这是因为在结晶杆的亚区中可能存在困难。尽管如此，但疏水核中心的色氨酸残基的保守性为三个剪接中的两个(“A”和“C”螺旋中的那些)提供了“锚定点(anchor points)”。注意例如在图3中所示的HMM标识中的位置16处W的突出(Wheeler等人，BMCBioinformatics 2014)。所有晶体结构都对于包含两个相互作用色氨酸残基的三股螺旋重复序列进行分析，并使用在HMMer门户网站(web portal)上的HMMscan分析以定义“B”螺旋内各个位置将对应于将色氨酸二等分的横截面平面(即，垂直于其长轴将螺旋重复序列二等分的平面)的概率。

[0083] LQGEI*E*AHTDVY(N-末端至C-末端，全长肌养蛋白TH22中的氨基酸序列)

...QEDLEQEQV(N-末端至C-末端,全长肌养蛋白TH2中的氨基酸序列)是围绕形成毫微-肌养蛋白中杂合TH的全长肌养蛋白的两个TH的螺旋B(螺旋2,TH中三个螺旋中的第二个螺旋)的剪接点的序列。加下划线的字母E和Q表示螺旋B中的剪接点,而E和Q两者均保持在如SEQ ID NO:1所举例说明的作为结果而产生的杂合TH中。然而,本领域技术人员将理解,本文使用的三重剪接突变肌养蛋白可具有除如作为SEQ ID NO:1的特征所示的EQ之间以外的螺旋B中的剪接点。这种剪接点可以位于如LQGEIEAHTDVY…QEDLEQEQV以及图5和SEQ ID NOs:13-18的相应氨基酸序列中所示的螺旋B中任何氨基酸的N-末端或C-末端,或另一种TH中螺旋B的另一个相应位置。序列中位置的对应可以通过任何两个或更多个肌养蛋白TH之间的氨基酸序列比对或隐蔽马尔科夫模型(HMM)来确定。

[0084] 类似地,本文使用的三重剪接突变肌养蛋白相关蛋白可以具有螺旋B中的剪接点,其在如作为SEQ ID NO:3的特征所示的HQ之间,或者可以位于AEIDAHNDIFKS(N-末端至C-末端,全长肌养蛋白相关蛋白TH20中的氨基酸序列)… DL EAEQVKV(N-末端至C-末端,全长肌养蛋白相关蛋白TH2中的氨基酸序列)中螺旋B中任何氨基酸的N-末端或C-末端,或另一种肌养蛋白相关蛋白TH中螺旋B的另一个相应位置。序列中位置的对应可以通过任何两个或更多个肌养蛋白相关蛋白TH之间的氨基酸序列比对或隐蔽马尔科夫模型(HMM)来确定。

[0085] 在某些实施方案中,螺旋A或C(其是三股螺旋充分序列中三个螺旋中的螺旋1,即,第一螺旋,和螺旋3,即,第三螺旋)中的剪接点可位于所述超家族中蛋白的隐蔽马尔科夫模型(HMM)和所有晶体结构的核心处的色氨酸(W)处。在进一步的实施方案中,螺旋A或C中的剪接点可位于从W(s)至蛋白的C-末端侧或至N-末端侧1个氨基酸、2个氨基酸、3个氨基酸、4个氨基酸或5个氨基酸的位置。

[0086] B. 人毫微-肌养蛋白相关蛋白序列

[0087] 以下序列中的较大的大写字母表示与全长人肌养蛋白相关蛋白的N-末端区域相同的部分。较小的大写字母表示与全长人肌养蛋白相关蛋白的C-末端区域相同的区域。以斜体形式的W对应于位于所述超家族中蛋白的隐蔽马尔科夫模型(HMM)和所有晶体结构的核心处的“A”和“C”螺旋中的色氨酸残基,以斜体形式的HQ对应于侧接如图2F中所描绘的假设横切面平面的“B”螺旋的超家族HMM内的位置。折叠的蛋白的预期二级和三级结构对应于图4中描绘的杂合TH。

[0088] MAKYGEHEASPDNGQNEFSDIIKSRSDEHNDVQKKTFTKWIN
ARFSKSGKPPINDMFTDLKDGRKLLDLLEGLTGTSLPKERGSTRVHA
LNNVNRVLQVLHQNNVELVNIGGTDIVDGNHKLTLGLLWSIILHWQ
VKDVMKDVMSDLQQTNSEKILLSWVRQTTRPYSQVNVLNFTTSWT
DGLAFNAVLHRHKPDLFSWDKVVKMSPIERLEHAFSKAQTYLGIEK
LLDPEDVAVQLPDKKSIIMYLTSLFEVLPQQVTIDAIREVETLPRKYK
KECEEEAINIQSTAPEEEHESPRAETPSTVTEVDMDLDSYQIALEEVL
TWLLSAEDTFQEQDDISDDVEEVKDQFATHEAFMMELTAHQSSVGS
VLQAGNQLITQGTLSDEEEFEIQEQMTLLNARWEALRVESMDRQSR
LHDVLMELQKKQLQQLSA*W*IQEAETTVNVLVDASHRENALQDSILARELKQ
QMQDIQAEIDA*HQ*VKVNSLTHMVVIVDENSGESATAILEDQLQKLGER
*W*NDLKAKSASIRAHLEASAEKWNRLLMSLEELIKWLNMKDEELKKQMPIGGDVPA
LQLQYDHCKALRRELKEKEYSVLNAVDQARVFLADQPIEAPEEPRRNLQSKTELTPE
ERAQKIAKAMRKQSSEVKEKWESLNAVTSNWQKQVDKALEKLRDLQGAMDDLDA

[0089] DMKEAESVRNGWKPVGDLLIDSLQDHIEKIMAFREEIAPINFKVKTVNDLSSQLSPL
DLHPSLKMSRQLDDLNMRWKLLQVSVDDRLKQLQEAHRDFGPSSQHFLSTSVQLP
WQRSISHNKVPYYINHQTQTTCWDHPKMTELFQSLADLNNVRFSAYRTAIKIRRLQ
KALCLDLLELSTTNEIFKQHKLNQNDQLLSVPDVINCLTTTYDGLEQMHKDLVNVP
LCVDMCLNWLLNVYDTGRTGKIRVQSLKIGLMSLSKGLLEEKYRYLFKEVAGPTE
MCDQRQLGLLLHDAIQIPRQLGEVAAFGGSNIEPSVRSCFQQNNNKPEISVKEFIDW
MHLEPQSMVWLPVLHRVAAAETAKHQAKCNICKECPIVGFRYRSLKHFNYDVCQS
CFFSGRTAKGHKLHYPMVEYCIPTTSGE (SEQ ID NO: 3)

[0090] 在某些实施方案中，本文提供的毫微-肌养蛋白相关蛋白包含以下氨基酸序列，其包含全长肌养蛋白相关蛋白中连接重复序列2和20的三重剪接突变：

[0091] MAKYGEHEASPDNGQNEFSDIIKSRSDEHNDVQKKTFTKWINARFS
KSGKPPINDMFTDLKDGRKLLDLLEGLTGTSLPKERGSTRVHALNN
VNRVLQVLHQNNVELVNIGGTDIVDGNHKLTLGLLWSIILHWQVKD
VMKDVMSDLQQTNSEKILLSWVRQTTRPYSQVNVLNFTTSWTDGL
AFNAVLHRHKPDLFSWDKVVKMSPIERLEHAFSKAQTYLGIEKLLD
PEDVAVQLPDKKSIIMYLTSLFEVLPQQVTIDAIREVETLPRKYKKEC
EEEAINIQSTAPEEEHESPRAETPSTVTEVDMDLDSYQIALEEVLTWL
LSAEDTFQEQDDISDDVEEVKDQFATHEAFMMELTAHQSSVGSVLQ
AGNQLITQGTLSDEEEFEIQEQMTLLNARWEALRVESMDRQSRLHD
VLMELQKKQLQQLSA*W*IQEAETTVNVLVDASHRENALQDSILARELKQQMQD
IQAEIDA*HQ*VKVNSLTHMVVIVDENSGESATAILEDQLQKLGER*W*NDL
KAKSASIRAHLEASAEKWNRLLMSLEELIKWLNMKDEELKKQMPIGGDVPALQLQ
YDHCKALRRELKEKEYSVLNAVDQARVFLADQPIEAPEEPRRNLQSKTELTPEERAQ
KIAKAMRKQSSEVKEKWESLNAVTSNWQKQVDKALEKLRDLQGAMDDLDADMK
EAESVRNGWKPVGDLLIDSLQDHIEKIMAFREEIAPINFKVKTVNDLSSQLSPLDLHP
SLKMSRQLDDLNMRWKLLQVSVDDRLKQLQEAHRDFGPSSQHFLSTSVQLPWQRSI
SHNKVPYYINHQTQTTCWDHPKMTELFQSLADLNNVRFSAYRTAIKIRRLQKALCL
DLLELSTTNEIFKQHKLNQNDQLLSVPDVINCLTTTYDGLEQMHKDLVNVPLCVDM
CLNWLLNVYDTGRTGKIRVQSLKIGLMSLSKGLLEEKYRYLFKEVAGPTEMCDQRQ
LGLLLHDAIQIPRQLGEVAAFGGSNIEPSVRSCFQQNNNKPEISVKEFIDWMHLEPQS
[0092] MVWLPVLHRVAAAETAKHQAKCNICKECPIVGFRYRSLKHFNYDVCQSCFFSGRTA
KGHKLHYPMVEYCIPTTSGEDVRDFTKVLKNKFRSKKYFAKHPRLGYLPVQTVLEG
DNLET (SEQ ID NO: 20)

[0093] 本文描述了毫微-肌养蛋白相关蛋白的详细设计以处理结构约束。系统发生分析提示,祖先三股螺旋重复序列存在于与α-辅肌动蛋白直向同源的蛋白中,因为这是大多数单细胞真核生物蛋白质组中与血影蛋白共有序列匹配的唯一蛋白。包括α-辅肌动蛋白、α-和β-血影蛋白以及dystroplakins的训练组(training sets)产生了HMMs,其标识显示了色氨酸残基的异常保守性。在可用的高分辨率晶体结构中,侧链的位置和芳族相互作用是高度保守的,正如第三"B"α螺旋的结构。在上面的序列中,请注意加下划线的W和B螺旋氨基酸H和Q的位置。对应于这些亚结构域的多肽序列的重排或"剪接"用不同的字体大小描绘,而三个局部不连续(focaldiscontinuity)位点对应于图2F中所举例说明的截面平面,从而在肌养蛋白相关蛋白重复序列2和20之间产生3-D杂合体(hybrid)。

[0094] 本文还提供了具有五个血影蛋白样三股螺旋重复序列的毫微-肌养蛋白相关蛋白,其包括通过剪接全长人肌养蛋白相关蛋白蛋白中的TH 1和18而形成的杂合三股螺旋结构域。在某些实施方案中,毫微-肌养蛋白相关蛋白具有以下序列:

MAKYGEHEASPDNGQNEFSDIIKSRSDEHNDVQKKTFTKWINARFS
KSGKPPINDMFTDLKDGRKLLDLLEGLTGTSLPKERGSTRVHALNN
VNRVLQVLHQNNVELVNIGGTDIVDGNHKLTLGLLWSIILHWQVKD
VMKDVMSDLQQTNSEKILLSWVRQTTRPYSQVNVLNFTTSWTDGL
AFNAVLHRHKPDLFSWDKVVKMSPIERLEHAFSKAQTYLGIEKLLD
[0095] PEDVAVQLPDKKSIIMYLTSLFEVLPQQVTIDAIREVETLPRKYKKEC
EEEAINIQSTAPEEEHESPRAETPSTVTEVDMDLDSYQIALEEVLTWL
VLIDQMLKSNIVTVGDVEEINKTVSRMKITKADLEQRQSSVGSVLQAGNQLIT
QGTLSDEEEFEIQEQMTLLNARWDGTQHGVELRQQQLEDMIIDSLQWDDHR
EETEELMRKYEARLYILQQARRDPLTKQISDNQILLQELGPGDGIVMAFDNVLQKLL
EEYGSDDTRNVKETTEYLKTSWINLKQSIADRQNALEAEWRTVQASRRDLENFLKW
IQEAETTVNVLVDASHRENALQDSILARELKQQMQDIQAEIDAHNDIFKSIDGNRQK
MVKALGNSEEATMLQHRLDDMNQRWNDLKAKSASIRAHLEASAEKWNRLLMSLE
ELIKWLNMKDEELKKQMPIGGDVPALQLQYDHCKALRRELKEKEYSVLNAVDQAR
VFLADQPIEAPEEPRRNLQSKTELTPEERAQKIAKAMRKQSSEVKEKWESLNAVTSN
WQKQVDKALEKLRDLQGAMDDLDADMKEAESVRNGWKPVGDLLIDSLQDHIEKI
MAFREEIAPINFKVKTVNDLSSQLSPLDLHPSLKMSRQLDDLNMRWKLLQVSVDDR
[0096] LKQLQEAHRDFGPSSQHFLSTSVQLPWQRSISHNKVPYYINHQTQTTCWDHPKMTE
LFQSLADLNNVRFSAYRTAIKIRRLQKALCLDLLELSTTNEIFKQHKLNQNDQLLSVP
DVINCLTTTYDGLEQMHKDLVNVPLCVDMCLNWLLNVYDTGRTGKIRVQSLKIGL
MSLSKGLLEEKYRYLFKEVAGPTEMCDQRQLGLLLHDAIQIPRQLGEVAAFGGSNIE
PSVRSCFQQNNNKPEISVKEFIDWMHLEPQSMVWLPVLHRVAAAETAKHQAKCNIC
KECPIVGFRYRSLKHFNYDVCQSCFFSGRTAKGHKLHYPMVEYCIPTTSGE (SEQ ID
NO: 21)

[0097] C．人毫微-肌养蛋白序列：

[0098] 以下较大的大写字母表示与全长人肌养蛋白的N-末端区域相同的部分。较小的大写字母表示与全长人肌养蛋白的C-末端区域相同的区域。加下划线的W对应于位于所述超家族中蛋白的隐蔽马尔科夫模型(HMM)和所有晶体结构的核心处的“A”和“C”螺旋中的色氨酸残基。加下划线的EQ对应于侧接如图2F中所描绘的假设横切面平面的“B”螺旋的超家族HMM内的位置。折叠的蛋白的预期二级和三级结构对应于图4中描绘的杂合TH。

[0099] MLWWEEVEDCYEREDVQKKTFTKWVNAQFSKFGKQHIENLF
SDLQDGRRLLDLLEGLTGQKLPKEKGSTRVHALNNVNKALRVLQN
NNVDLVNIGSTDIVDGNHKLTLGLIWNIILHWQVKNVMKNIMAGLQ
QTNSEKILLSWVRQSTRNYPQVNVINFTTSWSDGLALNALIHSHRPD
LFDWNSVVCQQSATQRLEHAFNIARYQLGIEKLLDPEDVDTTYPDK
KSILMYITSLFQVLPQQVSIEAIQEVEMLPRPPKVTKEEHFQLHHQM
HYSQQITVSLAQGYERTSSPKPRFKSYAYTQAAYVTTSDPTRSPFPS
QHLEAPEDKSFGSSLMESEVNLDRYQTALEEVLSWLLSAEDTLQAQ
GEISNDVEVVKDQFHTHEGYMMDLTAHQGRVGNILQLGSKLIGTG
KLSEDEETEVQEQMNLLNSRWECLRVASMEKQSNLHRVLMDLQNQ
KLKELNDWLTEAETTANVLQDATRKERLLEDSKGVKELMKQWQDLQGEIEQE
QVRVNSLTHMVVVVDESSGDHATAALEEQLKVLGDRWANICRWSE
LRKKSLNIRSHLEASSDQWKRLHLSLQELLVWLQLKDDELSRQAPIGGDFPAVQKQ
NDVHRAFKRELKTKEPVIMSTLETVRIFLTEQPLEGLEKLYQEPRELPPEERAQNVTR
LLRKQAEEVNTEWEKLNLHSADWQRKIDETLERLQELQEATDELDLKLRQAEVIKG
[0100] SWQPVGDLLIDSLQDHLEKVKALRGEIAPLKENVSHVNDLARQLTTLGIQLSPYNLS
TLEDLNTRWKLLQVAVEDRVRQLHEAHRDFGPASQHFLSTSVQGPWERAISPNKVP
YYINHETQTTCWDHPKMTELYQSLADLNNVRFSAYRTAMKLRRLQKALCLDLLSL
SAACDALDQHNLKQNDQPMDILQIINCLTTIYDRLEQEHNNLVNVPLCVDMCLNWL
LNVYDTGRTGRIRVLSFKTGIISLCKAHLEDKYRYLFKQVASSTGFCDQRRLGLLLH
DSIQIPRQLGEVASFGGSNIEPSVRSCFQFANNKPEIEAALFLDWMRLEPQSMVWLPV
LHRVAAAETAKHQAKCNICKECPIIGFRYRSLKHFNYDICQSCFFSGRVAKGHKMH
YPMVEYCTPTTSGE (SEQ ID NO: 1)

[0101] 本文还提供了具有五个血影蛋白样三股螺旋重复序列的毫微-肌养蛋白，其包括通过剪接全长人肌养蛋白中的TH 1和20而形成的杂合三股螺旋结构域。在某些实施方案中，毫微-肌养蛋白相关蛋白具有以下序列：

MLWWEEVEDCYEREDVQKKTFTKWVNAQFSKFGKQHIENLFSDLQ
DGRRLLDLLEGLTGQKLPKEKGSTRVHALNNVNKALRVLQNNNVD
LVNIGSTDIVDGNHKLTLGLIWNIILHWQVKNVMKNIMAGLQQTNS
EKILLSWVRQSTRNYPQVNVINFTTSWSDGLALNALIHSHRPDLFDW
NSVVCQQSATQRLEHAFNIARYQLGIEKLLDPEDVDTTYPDKKSILM

[0102] YITSLFQVLPQQVSIEAIQEVEMLPRPPKVTKEEHFQLHHQMHYSQQI
TVSLAQGYERTSSPKPRFKSYAYTQAAYVTTSDPTRSPFPSQHLEAP
EDKSFGSSLMESEVNLDRYQTALEEVLSWLSLLDQVIKSQRVMVGDLEDI
NEMIIKQKATMQDLEQRQGRVGNILQLGSKLIGTGKLSEDEETEVQEQM
NLLNSRWDEVQEHLQNRRQQLNEMLKDSTQWLEAKEEAEQVLGQARAKLESW
KEGPYTVDAIQKKITETKQLAKDLRQWQTNVDVANDLALKLLRDYSADDTRKVH
MITENINASWRSIHKRVSEREAALEETHRLLQQFPLDLEKFLAWLTEAETTANVLQD
ATRKERLLEDSKGVKELMKQWQDLQGEIEAHTDVYHNLDENSQKILRSLEGSDDA
VLLQRRLDNMNFKWSELRKKSLNIRSHLEASSDQWKRLHLSLQELLVWLQLKDDE
LSRQAPIGGDFPAVQKQNDVHRAFKRELKTKEPVIMSTLETVRIFLTEQPLEGLEKLY
QEPRELPPEERAQNVTRLLRKQAEEVNTEWEKLNLHSADWQRKIDETLERLQELQE
ATDELDLKLRQAEVIKGSWQPVGDLLIDSLQDHLEKVKALRGEIAPLKENVSHVND

[0103] LARQLTTLGIQLSPYNLSTLEDLNTRWKLLQVAVEDRVRQLHEAHRDFGPASQHFL
STSVQGPWERAISPNKVPYYINHETQTTCWDHPKMTELYQSLADLNNVRFSAYRTA
MKLRRLQKALCLDLLSLSAACDALDQHNLKQNDQPMDILQIINCLTTIYDRLEQEHN
NLVNVPLCVDMCLNWLLNVYDTGRTGRIRVLSFKTGIISLCKAHLEDKYRYLFKQV
ASSTGFCDQRRLGLLLHDSIQIPRQLGEVASFGGSNIEPSVRSCFQFANNKPEIEAALF
LDWMRLEPQSMVWLPVLHRVAAAETAKHQAKCNICKECPIIGFRYRSLKHFNYDIC
QSCFFSGRVAKGHKMHYPMVEYCTPTTSGE (SEQ ID NO: 22)

[0104] D. 其他

[0105] 许多高分子量蛋白具有重复的内部结构域。肌养蛋白作为一大类蛋白的例子，对于所述一大类蛋白，对于重复的内部结构域有限的三维结构信息是可获得的。概念上相同的方法可能适用于其他遗传病。例如，常见的遗传性凝血紊乱冯维勒布兰德病是由具有多个重复序列的蛋白(“vWF”)的8千碱基编码序列中的突变引起的。最近发表的研究已揭示，重组蛋白在肝中的转基因表达足以治疗所述疾病，但是编码序列对于单个AAV基因组来说太大，并且两个AAV基因组之间的反式剪接太低效以致于无法达到治疗水平的重组蛋白表达。没有整个蛋白的晶体结构，但是通过本文概述的方法进行的分析立即提示了产生可取代全长vWF的小型化毫微-vWF蛋白的机会。

[0106] 冯维勒布兰德病一般是由*VWF*基因中的变异(突变)引起的遗传病。*VWF*基因提供了制备称为冯维勒布兰德因子(von Willebrand factor)的凝血蛋白的指令，所述蛋白对于

形成血凝块并防止损伤后进一步的失血是重要的。如果冯维勒布兰德因子不能正常起作用或太少的蛋白是可获得的，则血凝块将不能正确形成。减少冯维勒布兰德因子的量或导致所述蛋白异常起作用(或根本不起作用)的*VWF*基因突变是与所述状况有关的病征和症状的原因。这些变异可能以常染色体显性或常染色体隐性方式遗传，或可能在家族中没有任何其他病例的情况下首次在患病的人中发生(称为从头突变)。参见，例如，ghr.nlm.nih.gov/condition/von-willebrand-disease。本领域技术人员将理解，本文在整个说明书中描述的任何组合物、方案(regiment)、方面、实施方案和方法都意图应用于冯维勒布兰德病、突变冯维勒布兰德因子、编码突变冯维勒布兰德因子的核酸序列或包含这种编码序列的载体。

[0107] 冯维勒布兰德因子(vWF)在维持止血中是重要的。它通过在内皮下胶原基质和血小板表面受体复合体GPIb-IX-V之间形成分子桥，促进血小板黏附到血管损伤部位。它也可充当凝固因子VIII的蛋白伴侣，从而将其递送至损伤部位，稳定其异二聚体结构并保护其免于从血浆过早清除。有2个冯维勒布兰德因子的同工型。同工型1被认为是规范序列，并且含有具有UniProtKB标识符:P04275-1的氨基酸序列，其引入本文。另一个同工型是具有UniProtKB标识符:P04275-2的同工型2，所述同工型的序列引入本文。“全长”vWF可指同工型1。在某些实施方案中，“全长”vWF可以指vWF的同工型2或其他vWF的同源物。已在多种生物中鉴定了冯维勒布兰德因子的同源物，包括小鼠(UniProt Q8CIZ8)、大鼠(UniProt Q62935)、猪(UniProt Q28833)和狗(UniProt Q28295)。编码冯维勒布兰德因子或其任何其他同工型或同源物的可能核酸序列是可公开获得的，参见，例如，NCBI参考序列:NM_000552.4、X04385.1、M10321.1、X04146.1、AK128487.1、AK297600.1、AK292122.1、BC069030.1、BC022258.1、U81237.1、K03028.1、M17588.1、AF086470.1和X02672.1，其每一个都引入本文。

[0108] 如本文所用的，“突变冯维勒布兰德因子”或“突变vWF”是指具有一个或多个重复序列的内部缺失和连接两个重复序列的剪接点的冯维勒布兰德因子，其中所述剪接点位点在重复序列内而不是重复序列之间。在某些实施方案中，通过类似于A. 螺旋中的剪接点、I. 肌养蛋白、肌养蛋白相关蛋白等或任何实施例的方法确定剪接点。vWF中一个或多个重复序列的鉴定或表征可以通过隐蔽马尔科夫模型(HMM)或任何其他常规方法确定。参见，例如，Zhou YF等人，Sequence and structure relationshipswithin von Willebrand factor. Blood. 2012年7月12日;120(2):449-58. doi: 10.1182/blood-2012-01-405134. Epub 2012年4月6日;Sadler JE.Biochemistry and genetics of von Willebrand factor. Annu Rev Biochem. 1998;67:395-424;和Perkins SJ等人，The secondary structure of the von Willebrand factor typeA domain in factor B of human complement by Fourier transform infrared spectroscopy. Its occurrence in collagen types VI, VII, XII and XIV, theintegrins and other proteins by averaged structure predictions. J Mol Biol. 1994年4月22日;238(1):104-19。编码突变vWF的核酸序列可以通过用于反向翻译的工具产生。此外，可以对编码序列进行密码子优化以在主体例如人、小鼠、大鼠或狗中表达。

[0109] 应当理解，本文描述的任何组合物意图应用于整个说明书中描述的其他组合物、方案、方面、实施方案和方法。

[0110] II. 表达盒

[0111] 本文提供了表达盒,其包含编码肌养蛋白超家族三重剪接突变蛋白的核酸序列,所述核酸序列在指导其表达的调节序列控制下。

[0112] 如本文所用的,术语“表达”或“基因表达”是指将来自基因的信息用于功能性基因产物的合成的过程。基因产物可以是蛋白、肽或核酸聚合物(例如RNA、DNA或PNA)。在某些实施方案中,功能基因产物是肌养蛋白超家族三重剪接突变蛋白。在某些实施方案中,术语“基因”、“小基因”和“转基因”是指编码肌养蛋白超家族三重剪接突变蛋白例如毫微-肌养蛋白或毫微-肌养蛋白相关蛋白的序列。

[0113] 如本文所用的,“表达盒”是指核酸聚合物,其包含编码序列,启动子,并且可以包括对于其的其他调节序列,所述盒可以包装到载体中。

[0114] 如本文所用的,术语“调节序列”或“表达控制序列”是指诱导、阻抑或以其他方式控制与其可操作地连接的编码蛋白的核酸序列转录的核酸序列,例如起始子序列、增强子序列和启动子序列。

[0115] 如本文所用的,术语“可操作地连接”是指与编码序列邻接的表达控制序列和反式或远距离起作用以控制编码序列的表达控制序列两者。在某些实施方案中,编码序列编码肌养蛋白超家族三重剪接突变蛋白。

[0116] 当关于蛋白或核酸使用时,术语“异源的”表示蛋白或核酸包含在自然界中未以相同的彼此关系发现的两个或更多个序列或子序列(subsequences)。例如,核酸一般是重组产生的,具有被安排以产生新的功能核酸的来自不相关基因的两个或更多个序列。例如,在一个实施方案中,核酸具有来自一个基因的启动子,所述启动子被安排以导来自不同基因的编码序列的表达。因此,关于编码序列,启动子是异源的。

[0117] 关于序列的同一性或相似性在本文中定义为,在比对序列并引入缺口(如果需要的话)以达到最大百分比序列同一性后,候选序列中与本文提供的肽和多肽区域相同(即,相同残基)或相似(即,来自基于共同的侧链性质的相同组的氨基酸残基,参见下文)的氨基酸残基的百分比。百分比(%)同一性是两个多核苷酸或两个多肽之间的关系的量度,如分别通过比较它们的核苷酸或氨基酸序列所确定的。通常,将要比较的两个序列进行比对以给出序列之间的最大相关性。检查两个序列的比对,并且确定给出两个序列之间精确的氨基酸或核苷酸对应的位置数目,除以比对的总长度并乘以100,以给出%同一性数值。可以在要比较的序列的全长上确定所述%同一性数值,这特别适合于具有相同或非常相似的长度且高度同源的序列,或者在较短的确定长度上,这更适合于具有不相等的长度或具有较低水平的同源性的序列。存在许多算法和基于其的计算机程序,其是可获得的以被用于文献和/或公开或商业可获得的以用于执行比对和同一性百分比。算法或程序的选择不是本发明的限制。

[0118] 合适的比对程序的例子包括,例如,在Unix下并然后被输入到Bioedit程序中的软件CLUSTALW(Hall, T. A. 1999, BioEdit: a user-friendly biological sequence alignment editor andanalysis program for Windows 95/98/NT. Nucl. Acids. Symp. Ser. 41:95-98);可从EMBL-EBI获得的Clustal Omega(Sievers, Fabian等人,″Fast, scalable generation of high-quality proteinmultiple sequence alignments using Clustal Omega.″ Molecular systems biology 7.1 (2011): 539和Goujon, Mickael等

人,"A new bioinformatics analysis tools framework atEMBL-EBI." Nucleic acids research 38.suppl 2 (2010): W695-W699);威斯康星序列分析程序包(WisconsinSequence Analysis Package),版本9.1(Devereux J等人., Nucleic Acids Res., 12:387-395, 1984,可从美国的GeneticsComputer Group, Madison, Wis.获得)。程序BESTFIT和GAP可用于确定两个多核苷酸之间的%同一性和两个多肽序列之间的%同一性。

[0119] 用于确定序列之间的同一性和/或相似性的其他程序包括,例如,可从美国Bethesda, Md.的美国国家生物技术信息中心(National Center for BiotechnologyInformation)(NCB)获得且可通过位于www.ncbi.nlm.nih.gov)的NCBI主页进行访问的BLAST程序系列、作为GCG序列比对软件包的一部分的ALIGN程序(2.0版)。当利用ALIGN程序比较氨基酸序列时,可以使用PAM120权重残基表(weightresidue table)、缺口长度罚分12和缺口罚分4;以及FASTA(Pearson W. R.和Lipman D. J., Proc. Natl. Acad. Sci. USA, 85:2444-2448,可作为威斯康星序列分析程序包的一部分获得)。SeqWeb软件(GCG威斯康星程序包:Gap程序(GCG Wisconsin Package: Gap program)的基于万维网的界面)。

[0120] 在一个实施方案中,表达盒被进行设计用于在主体例如人、大鼠、小鼠或狗中表达和分泌。在一个实施方案中,表达盒被进行设计用于在包括心肌、骨骼肌和平滑肌的肌肉中表达。

[0121] 在某些实施方案中,调节控制元件包括作为表达控制序列一部分的启动子序列,例如,位于选择的5'ITR序列和编码序列之间。在本文所述的载体中可以使用可以利用组成型启动子、可调节的启动子[参见,例如,WO 2011/126808和WO 2013/04943]、组织特异性启动子(参见,例如,www.invivogen.com/tissue-specific-promoter)或对生理线索有反应的启动子。在某些实施方案中,可以使用肌肉特异性启动子,例如,肌肉肌酸激酶(MCK)启动子,结蛋白启动子,Mb启动子,或用于肌球蛋白-重多肽2,肌球蛋白,肌钙蛋白T类型3,肌钙蛋白C类型2,肌球蛋白结合蛋白C,快骨骼肌球蛋白(fast skeletal myosin)轻链2,辅肌动蛋白α2,突触小泡相关膜蛋白5,甲状腺激素受体相互作用蛋白(interactor)10,原肌球蛋白3,肌膜蛋白聚糖γ,生肌分化1,生肌因子6(力蛋白)或钙通道,电压依赖性γ1的启动子。另一种有用的启动子是合成的SPc5-12启动子,其允许在骨骼肌和心肌中稳健表达(参见例如,Rasowo等人, European Scientific Journal, 2014年6月版次第10卷第18期,和美国专利申请公开Nos. 20040192593和2017/0275649,其全部引入本文作为参考)。

[0122] 在某些实施方案中,调节控制元件包括心特异性顺式调节模块(CS-CRM),其包括CS-CRM元件1-8中的任何一个。在某些实施方案中,调节序列包括CS-CRM4元件或与SPc5-12启动子组合的CS-CRM4元件,例如以前描述的嵌合合成CS-CRM4/SPc5-12启动子(Rincon等人,Genome-wide computational analysis reveals cardiomyocyte-specific transcriptional Cis-regulatory motifs that enableefficient cardiac gene therapy. Mol Ther. 2015年1月;23(1):43-52),其引入本文作为参考)。

[0123] 适合于控制治疗用产物表达的组成型启动子的例子包括,但不限于鸡β-肌动蛋白(CB)启动子、人巨细胞病毒(CMV)启动子、泛素C启动子(UbC)、猿猴病毒40(SV40)的早期和晚期启动子、U6启动子、金属硫蛋白启动子、EFlα启动子、泛素启动子、次黄嘌呤磷酸核糖转

移酶(HPRT)启动子、二氢叶酸还原酶(DHFR)启动子(Scharfmann等人, Proc. Natl. Acad. Sci. USA 88:4626-4630 (1991)、腺苷脱氨酶启动子、磷酸甘油激酶(PGK)启动子、丙酮酸激酶启动子磷酸甘油酸变位酶启动子、β-肌动蛋白启动子(Lai等人, Proc. Natl. Acad. Sci. USA 86: 10006-10010 (1989)、莫洛尼白血病病毒和其他反转录病毒的长末端重复序列(LTR)、单纯疱疹病毒的胸苷激酶启动子和本领域技术人员已知的其他组成型启动子。适合用于本发明的组织或细胞特异性启动子的实例包括,但不限于,内皮缩血管肽-I(ET -I)和Flt-I,其对内皮细胞是特异性的,FoxJ1(其靶向纤毛细胞)。

[0124] 适合于控制治疗用产物表达的诱导型启动子包括对外源试剂(例如,药理学试剂)或对生理线索有反应的启动子。这些应答元件包括,但不限于结合HIF-Iα和β的低氧应答元件(HRE),金属离子应答元件,如由Mayo等人(1982, Cell 29:99-108);Brinster等人(1982, Nature 296:39-42)和Searle等人(1985, Mol. Cell. Biol. 5:1480-1489)所述的;或热激应答元件,如由Nouer等人(在Heat Shock Response, 编辑Nouer, L., CRC, Boca Raton, Fla., 第167-220页, 1991中)所述的。

[0125] 在一个实施方案中,编码序列的表达由可调节的启动子控制,所述可调节的启动子提供对于编码序列的转录的严格控制,例如,药理学试剂,或被药理学试剂激活的转录因子,或在备择实施方案中,生理线索。优选非渗漏且可被严格控制的启动子系统。可用于本发明中的作为配体依赖性转录因子复合体的可调节的启动子的实例包括,但不限于,被其各自配体(例如,糖皮质激素、雌激素、黄体酮、类视黄醇、蜕皮素及其类似物和模拟物(mimetics))激活的核受体超家族成员和被四环素激活的rTTA。在本发明的一个方面,基因开关(gene switch)是基于EcR的基因开关。这种系统的例子包括,但不限于,美国专利Nos.6,258,603、7,045,315、美国公开专利申请Nos. 2006/0014711、2007/0161086和国际公开申请No. WO 01/70816中描述的系统。嵌合蜕皮素受体系统的实例描述于美国专利No.7,091,038、美国公开专利申请Nos. 2002/0110861、2004/0033600、2004/0096942、2005/0266457和2006/0100416以及国际公开申请Nos. WO 01/70816、WO 02/066612、WO 02/066613、WO 02/066614、WO02/066615、WO 02/29075和WO 2005/108617中,其各自整体引入本文作为参考。非类固醇蜕皮素激动剂调节的系统的例子是RheoSwitch哺乳动物诱导型表达系统(RheoSwitch Mammalian Inducible Expression System)(New England Biolabs, Ipswich, MA)。

[0126] 又其他的启动子系统可包括应答元件,包括但不限于四环素(tet)应答元件(例如由Gossen&Bujard (1992, Proc. Natl. Acad. Sci. USA 89:5547-551)所述的;或激素应答元件,例如由Lee等人(1981, Nature 294:228-232);Hynes等人(1981,Proc. Natl. Acad. Sci. USA 78:2038-2042);Klock等人(1987, Nature 329:734-736);和Israel&Kaufman(1989, Nucl. Acids Res. 17:2589-2604)所述的和本领域已知的其他诱导型启动子。使用这种启动子,可以例如通过Tet-on/off系统(Gossen等人, 1995, Science 268:1766-9;Gossen等人,1992, Proc. Natl. Acad. Sci. USA., 89(12):5547-51);TetR-KRAB系统(Urrutia R., 2003, Genome Biol., 4(10):231;Deuschle U等人, 1995, Mol Cell Biol. (4):1907-14);米非司酮(RU486)可调节的系统(Geneswitch; Wang Y等人, 1994, Proc. Natl. Acad. Sci. USA., 91(17):8180-4;Schillinger等人, 2005, Proc. Natl. Acad. Sci. U S A.102(39):13789-94);人源化的他莫西芬-dep可调节的系统

(Roscilli等人, 2002, Mol. Ther. 6(5):653-63)控制转基因的表达。基因开关可基于FK506结合蛋白(FKBP)与FKBP雷帕霉素相关蛋白(FRAP)的异源二聚化,并通过雷帕霉素或其非免疫抑制类似物进行调节。这种系统的例子包括,但不限于,ARGENT™转录技术(ARGENT™ Transcriptional Technology)(ARIAD Pharmaceuticals, Cambridge, Mass.)和在美国专利Nos. 6,015,709、6,117,680、6,479,653、6,187,757和6,649,595、美国公开No. 2002/0173474、美国公开No. 200910100535、美国专利No.5,834,266、美国专利No. 7,109,317、美国专利No. 7,485,441、美国专利No. 5,830,462、美国专利No. 5,869,337、美国专利No. 5,871,753、美国专利No.6,011,018、美国专利No. 6,043,082、美国专利No. 6,046,047、美国专利No. 6,063,625、美国专利No. 6,140,120、美国专利No. 6,165,787、美国专利No.6,972,193、美国专利No. 6,326,166、美国专利No. 7,008,780、美国专利No. 6,133,456、美国专利No. 6,150,527、美国专利No. 6,506,379、美国专利No.6,258,823、美国专利No. 6,693,189、美国专利No. 6,127,521、美国专利No. 6,150,137、美国专利No. 6,464,974、美国专利No. 6,509,152、美国专利No.6,015,709、美国专利No. 6,117,680、美国专利No. 6,479,653、美国专利No.6,187,757、美国专利No. 6,649,595、美国专利No. 6,984,635、美国专利No. 7,067,526、美国专利No. 7,196,192、美国专利No. 6,476,200、美国专利No.6,492,106、WO 94/18347、WO 96/20951、WO 96/06097、WO 97/31898、WO 96/41865、WO98/02441、WO 95/33052、WO 99110508、WO 99110510、WO 99/36553、WO 99/41258、WO 01114387中描述的系统,ARGENT™调节的转录反转录病毒试剂盒(ARGENT™ Regulated Transcription Retrovirus Kit),版本2.0(9109102)和ARGENT™调节的转录质粒试剂盒(ARGENT™ Regulated Transcription Plasmid Kit),版本2.0(9109/02),其各自整体引入本文作为参考。将Ariad系统进行设计以由雷帕霉素及其称为“rapalogs”的类似物诱导。合适的雷帕霉素的实例在以上关于ARGENT系统的描述列出的文件中提供。在一个实施方案中,分子是雷帕霉素[例如,由Pfizer以雷帕鸣(Rapamune)市售的]。在另一个实施方案中,使用了被称为AP21967 [ARIAD]的rapalog。可用于本发明的这些二聚化剂(dimerizer)分子的实例包括,但不限于,雷帕霉素,FK506,FK1012(FK506的同型二聚体),雷帕霉素类似物(“rapalogs”),其可以通过对天然产物进行化学修饰以添加减少或消除对于内源FKBP和/或FRAP的亲和力的“凸起(bump)”而容易地制备。rapalogs的实例包括,但不限于例如AP26113(Ariad)、AP1510(Amara, J.F.等人,1997, Proc Natl Acad Sci USA, 94(20): 10618-23) AP22660、AP22594、AP21370、AP22594、AP23054、AP1855、AP1856、AP1701、AP1861、AP1692和AP1889,具有使与内源FKBP的相互作用减至最小的设计的“凸起”。还可以选择又其他的rapalogs,例如,AP23573 [Merck]。

[0127] 其他合适的增强子包括适于期望的靶组织指征的那些。在一个实施方案中,表达盒包含一种或多种表达增强子。在一个实施方案中,表达盒含有两种或更多种表达增强子。这些增强子可以相同或可以彼此不同。例如,增强子可以包括CMV立即早期增强子。所述增强子可以以彼此相邻定位的两个拷贝存在。另一方面,增强子的双重拷贝可以被一个或多个序列分开。在又另一个实施方案中,表达盒进一步含有内含子,例如,鸡β-肌动蛋白内含子。其他合适的内含子包括本领域中已知的那些,例如,如WO 2011/126808中所述的。合适的polyA序列的实例包括,例如,兔结合球蛋白(rBG),SV40,SV50,牛生长素(bGH),人生长激素和合成的polyAs。任选地,可以选择一种或多种序列以稳定mRNA。这种序列的实例是修饰

的WPRE序列，其可以在polyA的上游和编码序列的下游进行人工改造[参见，例如，MA Zanta-Boussif等人, Gene Therapy (2009) 16: 605-619。在一个实施方案中，增强子是双重或三重串联MCK增强子。

[0128] 在一个实施方案中，调节序列进一步包含多腺苷酸化信号(polyA)。在进一步的实施方案中，polyA是兔珠蛋白polyA。参见，例如，WO 2014/151341。另一方面，另一种polyA，例如人生长激素(hGH)多腺苷酸化序列，SV40 polyA或合成polyA可包括在表达盒中。

[0129] 应当理解，本文所述的表达盒中的组合物意图应用于整个说明书中描述的其他组合物、方案、方面、实施方案和方法。

[0130] III. 载体

[0131] 在某些实施方案中，将编码肌养蛋白超家族三重剪接突变蛋白的核酸序列在载体中人工改造，包括病毒载体和非病毒载体。

[0132] 如本文所用的“载体”是包含核酸序列的生物学或化学部分，其可以被引入适当的靶细胞中用于复制或表达所述核酸序列。载体的实例包括但不限于重组病毒、质粒、Lipoplexes、聚合物囊泡(Polymersome)、Polyplexes、树状聚体、细胞穿透肽(CPP)缀合物、磁性颗粒或纳米颗粒。这种载体优选具有一个或多个复制起点，以及一个或多个可向其中插入编码序列或表达盒的位点。载体经常具有这样的手段，通过所述手段可以从不具有载体的那些细胞选择具有载体的细胞，例如，它们编码抗药性基因。常见的载体包括质粒、病毒基因组和“人工染色体”。载体的产生、生产、表征或定量的常规方法对本领域技术人员而言是可获得的。

[0133] 如本文所用的，术语“宿主细胞”可以指其中生产载体(例如，重组AAV)的包装细胞系。宿主细胞可以是原核或真核细胞(例如，人、昆虫或酵母)，其含有已通过任何方式引入细胞中的外源或异源DNA，例如，电穿孔、磷酸钙沉淀、显微注射、转化、病毒感染、转染、脂质体递送、膜融合技术、高速DNA-包被的丸粒、病毒感染和原生质体融合。宿主细胞的实例可包括，但不限于分离的细胞、细胞培养物、大肠杆菌(Escherichia coli)细胞、酵母细胞、人细胞、非人细胞、哺乳动物细胞、非哺乳动物细胞、昆虫细胞、HEK-293细胞、肝细胞、肾细胞、肌细胞、平滑肌细胞、心肌细胞或骨骼肌细胞。

[0134] 如用于描述核酸序列或蛋白的术语“外源”意指核酸或蛋白不天然出现于其在染色体或宿主/靶细胞中存在的位置。外源核酸序列还指衍生自并插入相同宿主细胞或主体中但以非天然状态(例如不同的拷贝数)存在或在不同的调节元件控制下的序列。

[0135] 如本文所用的，术语“靶细胞”是指其中期望肌养蛋白超家族三重剪接突变蛋白表达的任何靶细胞。在某些实施方案中，术语“靶细胞”意图涉及正对于MD(包括DMD和BMD)进行治疗的主体的细胞。靶细胞的实例可包括，但不限于，肝细胞、肾细胞、肌细胞、平滑肌细胞、心肌细胞或骨骼肌细胞。在某些实施方案中，将载体先体外后体内递送至靶细胞。在某些实施方案中，将载体体内递送至靶细胞。

[0136] 非病毒载体可以是携带表达盒的质粒，所述表达盒至少包括编码肌养蛋白超家族三重剪接突变蛋白的核酸序列和任选地启动子或其他调节元件，其被递送至心。核酸分子到平滑肌和心肌系统的非病毒递送可包括化学或物理方法。化学方法包括使用阳离子脂质体(“lipoplex”)、聚合物(“polyplex”)、两者的组合(“lipopolyplex”)、磷酸钙和DEAE葡聚糖。另外地或任选地，这种核酸分子可以用于进一步包含一种或多种试剂的组合物中，所述

试剂包括，例如，脂质体试剂，如例如DOTAP/DOPE、Lipofectin、Lipofectamine等，以及阳离子聚合物，例如PEI、Effectene和树状聚体。这种试剂对于转染平滑肌细胞是有效的。除化学方法以外，存在许多促进未复合的DNA直接进入细胞的物理方法。这些方法可以包括个别细胞的显微注射、hydroporation、电穿孔、超声和生物射弹递送(即，基因枪)。

[0137] 在某些实施方案中，包含编码肌养蛋白超家族三重剪接突变蛋白的核酸序列的表达盒由病毒载体携带，例如，重组腺病毒、慢病毒、博卡病毒(bocavirus)、杂种AAV/博卡病毒(参见，例如Yan Z等人，A novel chimeric adenoassociated virus 2/human bocavirus 1 parvovirus vector efficiently transduces human airway epithelia.Mol Ther. 2013年12月;21(12):2181-94. doi: 10.1038/mt.2013.92. Epub 2013年7月30日)、单纯疱疹病毒或腺伴随病毒。在这种实施方案中，病毒载体可以是复制缺陷型病毒。

[0138] “复制缺陷型病毒”或“病毒载体”是指其中包含表达盒的载体基因组包装在病毒衣壳或包膜中的合成或人工病毒颗粒，其中也包装在病毒衣壳或包膜中的任何病毒基因组序列是复制缺陷型的；即，它们不能产生子代病毒体，但保留感染靶细胞的能力。在一个实施方案中，病毒载体的基因组不包括编码复制所需的酶的基因(基因组可以被人工改造为“无内容的”-仅含有由人工基因组的扩增和包装所需的信号侧接的感兴趣的转基因)，但这些基因可以在生产过程中提供。因此，因为除非在有复制所需的病毒酶的情况下，否则子代病毒体不能发生复制和感染，所以其被认为对于在基因治疗中使用是安全的。

[0139] 载体可以是本领域已知或上文公开的任何载体，包括裸DNA、质粒、噬菌体、转座子、黏粒、附加体、病毒等。载体向宿主细胞中的引入可以通过本领域已知或如上文所公开的任何方式实现，包括转染和感染。可以将一种或多种腺病毒基因稳定整合到宿主细胞的基因组中、作为附加体稳定表达或瞬时表达。基因产物可以全部瞬时表达、在附加体上或稳定整合，或者一些基因产物可以稳定表达，而其他则瞬时表达。此外，每种腺病毒基因的启动子可以独立地选自组成型启动子、诱导型启动子或天然腺病毒启动子。启动子可以通过生物或细胞的特定生理状态(即，通过分化状态或在复制或休眠细胞中)或通过例如外源添加的因子调节。

[0140] 也可以使用技术人员已知的且如在整个说明书中讨论的技术完成将分子(作为质粒或病毒)引入宿主细胞。在优选的实施方案中，使用标准转染技术，例如，$CaPO_4$转染或电穿孔。将腺病毒的选择的DNA序列(以及转基因和其他载体元件装配入各种中间质粒，以及使用质粒和载体以产生重组病毒颗粒都使用常规技术实现。这种技术包括常规的cDNA克隆技术，例如教科书[Sambrook等人，Molecular Cloning: A Laboratory Manual]中所述的那些，腺病毒基因组的重叠寡核苷酸序列的使用，聚合酶链反应以及提供所期望的核苷酸序列的任何合适方法。使用标准转染和共转染技术，例如，$CaPO_4$沉淀技术。所使用的其他常规方法包括病毒基因组的同源重组、琼脂覆层中病毒的噬斑、测量信号产生的方法等。

[0141] 病毒载体的剂量将主要取决于诸如正被治疗的状况，患者的年龄、重量和健康状况的因素，且因此在患者之间可能不同。例如，病毒载体的治疗有效的成人或兽医剂量通常在约100 μL至约100 mL载体的范围内，所述载体含有约1×10^{6}至约1×10^{15}个颗粒、约1×10^{11}至1×10^{13}个颗粒或约1×10^{9}至1×10^{12}个颗粒病毒的浓度。剂量将取决于动物的大小和施用途径而变动。例如，对于单个部位，用于肌内注射的合适的人或兽医剂量(对于约80 kg

的动物)在每mL约1×10^{9}至约5×10^{12}个颗粒的范围内。任选地,可以递送多个施用部位。在另一个实例中,对于制剂,合适的人或兽医剂量可以在约1×10^{11}至约1×10^{15}个颗粒的范围内。本领域技术人员可以取决于施用途径以及采用重组载体的治疗或接种应用来调节这些剂量。可以监控转基因的表达水平,以确定剂量施用的频率。确定施用频率的时间控制的再其他方法对于本领域技术人员将是显而易见的。

[0142] 如本文所用的,"载体基因组"是指包装在载体内的核酸序列。

[0143] A. 复制缺陷型腺病毒载体

[0144] 在一个实施方案中,使用复制缺陷型腺病毒载体。许多合适的腺病毒中的任何一种都可以用作腺病毒衣壳序列的来源和/或用于生产。参见,例如,美国专利Nos. 9,617,561;9,592,284;9,133,483;8,846,031;8,603,459;8,394,386;8,105,574;7,838,277;7,344,872;8,387,368;6,365,394;6,287,571;6,281,010;6,270,996;6,261,551;6,251,677;6,203,975;6,083,716;6,019,978;6,001,557;5,872,154;5,871,982;5,856,152;5,698,202。又其他的腺病毒可从美国典型培养物保藏中心(American Type Culture Collection)获得。在一个实施方案中,通过E1a和/或E1b基因中的缺失使腺病毒颗粒成为复制缺陷型的。另一方面,通过另一种方式使腺病毒成为复制缺陷型的,任选地同时保留E1a和/或E1b基因。腺病毒载体还可以含有对腺病毒基因组的其他突变,例如,温度敏感突变或其他基因中的缺失。在其他实施方案中,期望在腺病毒载体中保留完整的E1a和/或E1b区域。这种完整的E1区域可以位于其在腺病毒基因组中的天然位置,或者放置于天然腺病毒基因组中的缺失位点(例如,在E3区域中)。

[0145] 在构建有用的腺病毒载体用于将基因递送给人(或其他哺乳动物)细胞中,可以在载体中使用一系列腺病毒核酸序列。例如,可以从形成重组病毒一部分的腺病毒序列中消除全部或部分腺病毒延迟早期基因E3。据信E3的功能与重组病毒颗粒的功能和产生不相干。还可以构建具有至少E4基因的ORF6区域缺失的腺病毒载体,并且由于所述区域整个E4区域的功能冗余而是更期望的。又另一个腺病毒载体含有延迟早期基因E2a中的缺失。也可以在腺病毒基因组的晚期基因L1至L5中的任一种中进行缺失。类似地,中期基因IX和IVa_2中的缺失对于一些目的可能是有用的。可以在其他结构或非结构腺病毒基因中进行其他缺失。上文讨论的缺失可以个别地使用,即,如本文所述供使用的腺病毒序列可以含有仅单个区域中的缺失。另一方面,可以以任何组合使用有效破坏其生物活性的整个基因或其部分的缺失。例如,在一种示例性载体中,腺病毒序列可具有E1基因和E4基因,或E1、E2a和E3基因,或E1和E3基因,或在有或没有E3缺失的情况下E1、E2a和E4基因等等的缺失。如上文所讨论的,这种缺失可以与其他突变(例如温度敏感突变)组合使用,以实现所期望的结果。

[0146] 缺乏任何必需腺病毒序列(例如,E1a、E1b、E2a、E2b、E4 ORF6、L1、L2、L3、L4和L5)的腺病毒载体可以在有对于腺病毒颗粒的病毒感染性和繁殖所需要的缺少的腺病毒基因产物的情况下培养。这些辅助功能可以通过在有一种或多种辅助构建体(例如,质粒或病毒)或包装宿主细胞的情况下培养腺病毒载体来提供。参见,例如,1996年5月9日公开且引入本文作为参考的国际专利申请WO96/13597中对于制备"最小"人Ad载体描述的技术。

[0147] a. 辅助病毒

[0148] 因此,取决于用于携带表达盒的病毒载体的腺病毒基因内容物,辅助腺病毒或非复制性病毒片段可能是必需的,以提供产生含有表达盒的感染性重组病毒颗粒所必需的足

够的腺病毒基因序列。有用的辅助病毒含有腺病毒载体构建体中不存在和/或在其中转染了载体的包装细胞系中不表达的选择的腺病毒基因序列。在一个实施方案中，辅助病毒是复制缺陷型的，且除上述序列外还含有多种腺病毒基因。期望将这种辅助病毒与表达E1的细胞系组合使用。

[0149] 辅助病毒也可以形成聚阳离子缀合物，如Wu等人，J. Biol. Chem.，264:16985-16987(1989)；K. J. Fisher和J. M.Wilson，Biochem. J.，299:49(1994年4月1日)中所述的。辅助病毒可以任选含有第二报道小基因。许多这种报道基因是本领域已知的。与腺病毒载体上的转基因不同的辅助病毒上报道基因的存在允许Ad载体和辅助病毒被独立监控。所述第二报道基因用于使得能够在纯化时分离所得到的重组病毒和辅助病毒。

[0150] b. 互补细胞系

[0151] 为了产生在上述任何基因中缺失的重组腺病毒(Ad)，必须通过辅助病毒或细胞系，即，互补或包装细胞系，将缺失的基因区域的功能(如果对于病毒的复制和感染性必需的话)提供给重组病毒。在许多情况中，表达人E1的细胞系可用于反式互补(trans-complement)Ad载体。然而，在某些情况中，将期望利用表达E1基因产物的细胞系，可用于产生E1-缺失的腺病毒。已经描述了这种细胞系。参见，例如，美国专利6,083,716。

[0152] 如果期望，那么人们可以利用本文提供的序列以产生包装细胞或细胞系，其至少表达在用于在选择的亲本细胞系中表达的启动子的转录控制下的腺病毒E1基因。诱导型或组成型启动子可用于所述目的。这种启动子的实例在本说明书的别处有详细描述。选择亲本细胞以用于产生表达任何所期望的腺病毒基因的新细胞系。在没有限制的情况下，这种亲本细胞系除了其他的之外可以是HeLa [ATCC检索号CCL 2]、A549 [ATCC检索号CCL 185]、HEK 293、KB [CCL 17]、Detroit [例如，Detroit 510, CCL 72]和WI-38 [CCL 75]细胞。这些细胞系均可从美国典型培养物保藏中心，10801 University Boulevard, Manassas,Virginia 20110-2209获得。其他合适的亲本细胞系可以从其他来源获得。

[0153] 这种表达E1的细胞系可用于产生重组腺病毒E1缺失的载体。此外或另一方面，表达一种或多种腺病毒基因产物例如E1a、E1b、E2a和/或E4 ORF6的细胞系可以使用与用于产生重组病毒载体基本上相同的程序来构建。这种细胞系可用于反式互补在编码那些产物的必需基因中缺失的腺病毒载体，或提供对于包装依赖于辅助病毒的病毒(例如，腺伴随病毒)所必需的辅助功能。宿主细胞的制备包括诸如装配所选择的DNA序列的技术。所述装配可能利用常规技术完成。这种技术包括cDNA和基因组克隆，其是众所周知的且在上文引用的Sambrook等人中进行了描述，腺病毒基因组的重叠寡核苷酸序列的使用，结合聚合酶链反应，合成方法和提供所期望的核苷酸序列的任何其他合适的方法。

[0154] 在又另一个备择方案中，必需的腺病毒基因产物由腺病毒载体和/或辅助病毒反式提供。在这种情况中，合适的宿主细胞可以选自任何生物学生物，包括原核(例如，细菌)细胞，以及真核细胞，包括，昆虫细胞、酵母细胞和哺乳动物细胞。特别期望的宿主细胞选自任何哺乳动物物种，包括，但不限于，细胞，例如A549、WEHI、3T3、10T1/2、HEK 293细胞或PERC6(两者均表达功能性腺病毒E1)[Fallaux, FJ等人, (1998), Hum Gene Ther, 9:1909-1917]、Saos、C2C12、L细胞、HT1080、HepG2和衍生自哺乳动物包括人、猴、小鼠、大鼠、兔和仓鼠的原代成纤维细胞、肝细胞和成肌细胞。提供所述细胞的哺乳动物物种的选择不是本发明的限制；哺乳动物细胞的类型，即，成纤维细胞、肝细胞、肿瘤细胞等也不是。

[0155] c. 病毒颗粒的装配和细胞系的转染

[0156] 通常,当通过转染递送包含小基因的载体时,载体以约5 μg至约100 μg DNA,且优选约10至约50 μg DNA的量递送至约1×10^{4}个细胞至约1×10^{13}个细胞,且优选约10^{5}个细胞。然而,可以考虑诸如选择的载体、递送方法和选择的宿主细胞的因素来调节载体DNA与宿主细胞的相对量。

[0157] B. 慢病毒系统

[0158] 多种不同的慢病毒系统是本领域已知的。参见,例如,关于用于利用慢病毒系统获得稳定的心血管转导的方法的WO2001089580 A1。参见,例如,美国专利6,521,457。也参见NBWasala等人,The evolution of heart gene delivery vectors, J Gen Med., 2011年10月; 13(10): 557-565中的讨论,其引入本文作为参考。

[0159] C. 重组AAV

[0160] 在某些实施方案中,载体基因组是指包装在载体例如rAAV内的核酸序列。对于rAAV,这种核酸序列可以含有AAV反向末端重复序列(ITRs)和表达盒。在一个实例中,载体基因组从5'至3'至少含有AAV 5'ITR、编码肌养蛋白超家族三重剪接突变蛋白的核酸序列和AAV 3'ITR。在一个实例中,载体基因组从5'至3'至少含有AAV 5'ITR、表达盒和AAV 3'ITR。ITRs可能来自AAV2或来自除AAV2以外的不同来源AAV。在其他实施方案中,载体基因组可以含有自身互补的AAV载体所需的末端重复序列(TRs)。

[0161] 在一个实施方案中,本文提供了具有AAV衣壳和载体基因组的重组腺伴随病毒(rAAV),其中所述载体基因组包含如本文所述的表达盒或编码肌养蛋白超家族三重剪接突变蛋白的核酸序列(即,如本文所用的cDNA),所述核酸序列在指导其表达的调节序列的控制下。

[0162] 在一些实施方案中,肌养蛋白超家族三重剪接突变蛋白进行设计以从重组腺伴随病毒表达,并且载体基因组还含有AAV反向末端重复序列(ITRs)。在一个实施方案中,rAAV是假型的,即,AAV衣壳来自与提供ITRs的AAV不同的来源AAV。在一个实施方案中,使用了AAV血清型2的ITRs。然而,可以选择来自其他合适来源的ITRs。任选地,AAV可以是自身互补的AAV。

[0163] 缩写"sc"是指自身互补的。"自身互补的AAV"是指其中由重组AAV核酸序列携带的编码区已被进行设计以形成分子内双链DNA模板的构建体。感染时,scAAV的两个互补的一半将结合以形成随时可以立即复制和转录的一个双链DNA(dsDNA)单位,而不是等待细胞介导的第二链的合成。参见,例如,D M McCarty等人, Self-complementary recombinantadeno-associated virus (scAAV) vectors promote efficient transduction independently of DNA synthesis, Gene Therapy, (2001年8月), 第8卷, 第16期, 第1248-1254页。自身互补的AAVs描述于例如美国专利Nos. 6,596,535;7,125,717;和7,456,683,其每一个都整体引入本文作为参考。

[0164] 在要从AAV表达基因的场合,本文所述的表达盒包括AAV 5'反向末端重复序列(ITR)和AAV 3'ITR。然而,这些元件的其他构型可能是合适的。已经描述了称为ΔITR的5'ITR的缩短的形式,其中缺失了D-序列和末端解离位点(terminal resolution site)(trs)。在其他实施方案中,使用了全长AAV 5'和/或3'ITRs。在要产生假型AAV的场合,表达中的ITRs选自与衣壳的AAV来源不同的来源。例如,可以选择AAV2 ITRs供与具有用于靶向

肌肉的特定效率的AAV衣壳一起使用。在一个实施方案中，为方便起见并用于加快监管批准，使用来自AAV2的ITR序列或其缺失的形式（ΔITR）。然而，可以选择来自其他AAV来源的ITRs。在ITRs的来源来自AAV2且AAV衣壳来自另一种AAV来源的场合，所得到的载体可以被称为假型的。然而，可以利用AAV ITRs的其他来源。

[0165] 如本文所用的，“重组AAV病毒颗粒”或“AAV病毒颗粒”是指具有衣壳并在其中包装了异源核酸分子（载体基因组）的抗核酸酶的颗粒（nuclease-resistant particle）（NRP），所述异源核酸分子包含肌养蛋白超家族三重剪接突变蛋白的表达盒。这种表达盒一般含有侧接基因序列的AAV 5’和/或3’反向末端重复序列，其中所述基因序列可操作地连接至表达控制序列。在其中包装了载体基因组的这种衣壳也可以被称为“实心”AAV衣壳。当这种rAAV病毒颗粒将转基因递送至能够表达由表达盒携带的所期望的基因产物的宿主细胞时，其被称为“药理学活性的”。

[0166] 在许多情况中，rAAV颗粒被称为“抗DNA酶的”。然而，除了所述内切核酸酶（DNA酶）外，其他内切核酸酶和外切核酸酶也可用于本文所述的纯化步骤中，以去除污染核酸。可以选择这种核酸酶以降解单链DNA和/或双链DNA以及RNA。这种步骤可以含有单种核酸酶，或针对不同靶的核酸酶的混合物，并且可以是内切核酸酶或外切核酸酶。

[0167] 术语“抗核酸酶的”表示AAV衣壳已在被进行设计以将转基因递送至宿主细胞的表达盒周围完全装配，并在核酸酶温育步骤过程中保护这些包装的基因组序列免于降解（消化），所述核酸酶温育步骤被设计用于去除可能自生产过程中存在的污染核酸。

[0168] 如本文所用的，“AAV9衣壳”是由多种AAV9 vp蛋白组成的自动装配的AAV衣壳。AAV9 vp蛋白一般作为来自编码SEQ ID NO:9的vp1氨基酸序列（GenBank检索:AAS99264）的SEQ ID NO:10的核酸序列或与其至少70％、至少75％、至少80％、至少85％、至少90％、至少95％、至少97％、至少99％的序列的可变剪接变体产生的。这些剪接变体结果产生SEQ ID NO:9的不同长度的蛋白。在某些实施方案中，“AAV9衣壳”包括具有99％相同于SEQ ID NO:9（即，与参考序列小于约1％的变异）的氨基酸序列的AAV。还参见US7906111和WO 2005/033321。这种AAV可以包括，例如，天然分离物（例如，hu31，其vp1由SEQ ID NO:11编码；或hu32，其vp1由SEQ ID NO:12编码），或具有氨基酸取代、缺失或添加的AAV9的变体，例如，包括但不限于选自从与AAV9衣壳比对的任何其他AAV衣壳中的相应位置“募集”的替代残基的氨基酸取代；例如，如在US9,102,949、US 8,927,514、US 8,734,809和WO 2016/049230A1中描述的。然而，在其他实施方案中，可以选择与上述序列具有至少约95％同一性的AAV9或AAV9衣壳的其他变体。参见，例如，美国公开专利申请No. 2015/0079038。已经描述了产生衣壳的方法、用于其的编码序列以及产生rAAV病毒载体的方法。参见，例如，Gao等人，Proc. Natl. Acad. Sci. U.S.A. 100 (10), 6081-6086(2003)和US 2013/0045186A1。

[0169] 除了AAV9之外，可以使用其他AAV载体，例如AAV1、AAV5、AAV6、AAV8、AAV8 triple、AAV9、Anc80、Anc81和Anc82。参见，例如，Santiago-Ortiz等人, Gene Ther., 22(12):934-46 (2015)；US20170051257A1；和Zinn等人, Cell Rep., 12(6): 1056-1068 (2015)。

[0170] 任一种AAV衣壳的序列都可以容易地通过合成或使用多种分子生物学和基因工程技术产生。合适的生产技术是本领域技术人员众所周知的。参见，例如，Sambrook等人，Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press (Cold Spring Harbor, NY)。另一方面，编码肽（例如，CDRs）的寡核苷酸或肽本身可以例如通过众所周知

的固相肽合成方法合成产生(Merrifield, (1962) J.*Am. Chem. Soc.*, 85:2149;Stewart和Young, Solid Phase Peptide Synthesis (Freeman, San Francisco, 1969) 第27-62页)。这些和其他合适的生产方法在本领域技术人员的知识范围内,且不是对本发明的限制。

[0171] 制备基于AAV的载体的方法是已知的。参见,例如,美国公开专利申请No. 2007/0036760(2007年2月15日),其引入本文作为参考。具有对于肌细胞和/或心细胞的向性的AAV衣壳的使用特别好地适合于本文所述的组合物和方法。然而,可以选择其他靶。AAV9的序列和产生基于AAV9衣壳的载体的方法描述于US 7,906,111;US2015/0315612;WO2012/112832;和WO2017160360A3,其引入本文作为参考。在某些实施方案中,AAV1、AAV5、AAV6、AAV9、AAV8triple、Anc80、Anc81和Anc82的序列是已知的,并可用于产生AAV载体。参见,例如,US 7186552、WO 2017/180854、US 7,282,199 B2、US 7,790,449和US 8,318,480,其引入本文作为参考。许多这种AAV的序列提供于上面引用的美国专利7,282,199 B2、US7,790,449、US 8,318,480、US专利7,906,111、WO/2003/042397、WO/2005/033321、WO/2006/110689、US 8,927,514、US 8,734,809;WO2015054653A3、WO-2016065001-A1、WO-2016172008-A1、WO-2015164786-A1、US-2010186103-A1、WO-2010138263-A2和WO 2016/049230A1,和/或可从GenBank获得。已经描述了对于AAV1、AAV8和AAVrh10-样载体的相应方法。参见WO2017100676 A1;WO2017100674A1;和WO2017100704A1。

[0172] 本文所述的重组腺伴随病毒(AAV)可使用已知技术产生。参见,例如,WO 2003/042397;WO 2005/033321、WO 2006/110689;US7588772 B2。这种方法包括培养含有编码AAV衣壳的核酸序列;功能性rep基因;至少由AAV反向末端重复序列(ITRs)和转基因组成的表达盒;以及足够的辅助功能的宿主细胞以允许将表达盒包装到AAV衣壳蛋白中。宿主细胞可以是293细胞或悬浮293细胞。参见,例如,如本文所引用的Zinn, E等人;Joshua C Grieger等人,Production of RecombinantAdeno-associated Virus Vectors Using Suspension HEK293 Cells and Continuous Harvest of Vector From the Culture Media for GMP FIX and FLT1 Clinical Vector.Mol Ther. 2016年2月; 24(2): 287-297. 在线公布于2015年11月3日. 在线重新公布于2015年10月6日. doi: 10.1038/mt.2015.187;Laura Adamson-Small等人,SodiumChloride Enhances Recombinant Adeno-Associated Virus Production in a Serum-Free Suspension Manufacturing Platform Using the Herpes Simplex Virus System. HumGene Ther Methods. 2017年2月1日; 28(1): 1-14. 在线公布于2017年2月1日. doi: 10.1089/hgtb.2016.151;US20160222356A1;和Chahal PS等人,Production of adeno-associated virus(AAV) serotypes by transient transfection of HEK293 cell suspension cultures for gene delivery. J Virol Methods. 2014年2月;196:163-73. doi: 10.1016/j.jviromet.2013.10.038. Epub2013年11月13日。

[0173] 可以利用本领域技术人员可获得的产生rAAV的其他方法。合适的方法可包括但不限于杆状病毒表达系统(例如,杆状病毒感染的昆虫细胞系统)或通过酵母生产。参见,例如,WO2005072364A2;WO2007084773A2;WO2007148971A8;WO2017184879A1;WO2014125101A1;US6723551B2;Bryant, L.M.等人, Lessons Learned from the ClinicalDevelopment and Market Authorization of Glybera. Hum Gene Ther Clin Dev, 2013;Robert M. Kotin, Large-scale recombinant adeno-associated virus

production. Hum Mol Genet. 2011年4月15日; 20(R1): R2-R6. 在线公布于2011年4月29日. doi: 10.1093/hmg/ddr141;Aucoin MG等人, Production of adeno-associatedviral vectors in insect cells using triple infection: optimization of baculovirus concentration ratios. Biotechnol Bioeng. 2006年12月20日;95(6):1081-92;SAMIS. THAKUR, Production of Recombinant Adeno-associated viral vectors in yeast. 提供给佛罗里达大学研究生院(the Graduate School of the University of Florida)的论文, 2012;Kondratov O等人,Direct Head-to-Head Evaluation of Recombinant Adeno-associated Viral Vectors Manufactured in Human versus Insect Cells, MolTher. 2017年8月10日. pii: S1525-0016(17)30362-3. doi: 10.1016/j.ymthe.2017.08.003. [Epub领先于印刷];Mietzsch M等人, OneBac 2.0: Sf9 Cell Lines forProduction of AAV1, AAV2, and AAV8 Vectors with Minimal Encapsidation of Foreign DNA. Hum Gene Ther Methods. 2017年2月;28(1):15-22. doi: 10.1089/hgtb.2016.164.;Li L等人,Production and characterization of novel recombinant adeno-associated virus replicative-formgenomes: a eukaryotic source of DNA for gene transfer. PLoS One. 2013年8月1日;8(8):e69879. doi: 10.1371/journal.pone.0069879. 2013年印刷;Galibert L等人, Latest developments in the large-scale production of adeno-associated virusvectors in insect cells toward the treatment of neuromuscular diseases. J Invertebr Pathol. 2011年7月;107 Suppl:S80-93. doi: 10.1016/j.jip.2011.05.008;和Kotin RM, Large-scale recombinant adeno-associated virusproduction. Hum Mol Genet. 2011年4月15日;20(R1):R2-6. doi: 10.1093/hmg/ddr141. Epub 2011年4月29日。

[0174] 为了计算空颗粒和满颗粒的含量,将选择的样品(例如,在本文的实施例中,碘克沙醇梯度纯化的制剂,其中GC的#＝颗粒的#)的VP3带体积(band volumes)对负载的GC颗粒作图。所得到的线性方程式(y = mx + c)用于计算测试物品峰的带体积中的颗粒数目。然后将负载的每20 μL的颗粒(pt)数目乘以50,以给出颗粒(pt)/mL。Pt/mL除以GC/mL给出颗粒与基因组拷贝的比例(pt/GC)。Pt/mL- GC/mL给出空pt/mL。空pt/mL除以pt/mL并x 100给出空颗粒的百分比。

[0175] 通常,用于测定空衣壳和具有包装的基因组的AAV载体颗粒的方法是本领域已知的。参见,例如,Grimm等人, *Gene Therapy* (1999) 6:1322-1330;Sommer等人, *Molec. Ther.* (2003) 7:122-128。为了对于变性的衣壳进行测试,所述方法包括使处理的AAV原液经历由能够分离三种衣壳蛋白的任何凝胶(例如,含有在缓冲液中的3-8%Tris-乙酸盐的梯度凝胶)组成的SDS-聚丙烯酰胺凝胶电泳,然后运行所述凝胶直到分离了样品材料,并将凝胶印迹到尼龙或硝化纤维素膜上,优选尼龙。然后将抗-AAV衣壳抗体用作结合变性的衣壳蛋白的第一抗体,优选抗-AAV衣壳单克隆抗体,最优选B1抗-AAV-2单克隆抗体(Wobus等人,*J. Virol.* (2000) 74:9281-9293)。然后使用第二抗体,所述第二抗体与第一抗体结合并且含有用于检测与第一抗体结合的手段,更优选含有与其共价结合的检测分子的抗-IgG抗体,最优选与辣根过氧化物酶共价连接的绵羊抗小鼠IgG抗体。使用检测结合的方法以半定量地确定第一抗体和第二抗体之间的结合,优选能够检测放射性同位素发射、电磁辐射或比色变化的检测方法,最优选化学发光检测试剂盒。例如,对于SDS-PAGE,可以获取来自

柱级分的样品并在含有还原剂(例如,DTT)的SDS-PAGE加样缓冲液中加热,且将衣壳蛋白在预制梯度聚丙烯酰胺凝胶(例如,Novex)上分离。可以根据制造商的说明书使用SilverXpress(Invitrogen, CA)进行银染色,或其他合适的染色方法,即SYPRO ruby或考马斯染色剂。在一个实施方案中,柱级分中的AAV载体基因组(vg)的浓度可以通过定量实时PCR(Q-PCR)进行测量。将样品稀释并用DNA酶I(或其他合适的核酸酶)消化以去除外源DNA。核酸酶失活后,将样品进一步稀释并使用引物和对引物之间的DNA序列特异的TaqMan™荧光探针进行扩增。在Applied Biosystems Prism 7700序列检测系统上对于每个样品测量达到定义的荧光水平所需的循环数(阈值循环,Ct)。使用含有与AAV载体中所含的序列相同的序列的质粒DNA,以在Q-PCR反应中产生标准曲线。从样品获得的循环阈值(Ct)值通过将其相对于质粒标准曲线的Ct值归一化用于确定载体基因组滴度。也可以使用基于数字PCR的终点测定。

[0176] 在一个方面,使用了最优化的q-PCR方法,其利用广谱丝氨酸蛋白酶,例如,蛋白酶K(如可从Qiagen商购获得的)。更具体地,最优化的qPCR基因组滴度测定与标准测定相似,只是在DNA酶I消化后,将样品用蛋白酶K缓冲液稀释并用蛋白酶K处理,后面是热灭活。适当地,用等于样品大小的量的蛋白酶K缓冲液稀释样品。蛋白酶K缓冲液可以浓缩至2倍或更高。一般,蛋白酶K处理为约0.2 mg/mL,但是可以从0.1 mg/mL到约1 mg/mL不等。所述处理步骤通常于约55℃进行约15分钟,但是可以于较低温度(例如,约37℃至约50℃)在较长时间段(例如,约20分钟至约30分钟)内进行,或于较高温度(例如,最多到约60℃)进行达较短时间段(例如,约5至10分钟)。类似地,热灭活通常于约95℃达约15分钟,但是温度可以降低(例如,约70至约90℃)并且时间延长(例如,约20分钟至约30分钟)。然后将样品稀释(例如,1000倍)并如标准测定中所述地经历TaqMan分析。

[0177] 此外或另一方面,可以使用液滴数字PCR(ddPCR)。例如,已经描述了用于通过ddPCR测定单链和自身互补的AAV载体基因组滴度的方法。参见,例如,M. Lock等人, Hu Gene Therapy Methods, Hum Gene Ther Methods. 2014年4月;25(2):115-25. doi: 10.1089/hgtb.2013.131. Epub 2014年2月14日。

[0178] 简而言之,用于从基因组缺陷的AAV中间体中分离具有包装的基因组序列的rAAV颗粒的方法包括使包含重组AAV病毒颗粒和AAV衣壳中间体的悬浮液经历快速液相层析(fast performance liquid chromatography),其中AAV病毒颗粒和AAV中间体与在高(例如,对于AAV9 pH 10.2)平衡的强离子交换树脂结合,并在监控洗脱物在约260和约280的紫外线吸收的同时经历盐梯度。尽管对于rAAV9较不最佳,但是pH可以在约10.0至10.4的范围内。在所述方法中,从当A260/A280的比率达到回折点时洗脱的级分收集AAV实心衣壳。在一个实例中,对于亲和层析步骤,可以将渗滤的产物应用于有效捕获AAV2/9血清型的Capture Select Poros-AAV2/9亲和树脂(Life Technologies)。在这些离子条件下,显著百分比的残留细胞DNA和蛋白流过柱,而AAV颗粒被有效捕获。

[0179] 如本文所用的,术语“治疗(treatment)”或“治疗(treating)”是指为了改善包括DMD和BMD的MD的一种或多种症状、恢复期望的全长肌养蛋白功能或改善疾病的生物标志物的一种或多种组合物和/或一种或多种方法。在一些实施方案中,定义术语“治疗(treatment)”或“治疗(treating)”包括为了本文所示的目的向主体施用本文所述的一种或多种组合物。因此,“治疗”可以包括在给定主体中预防疾病、降低疾病症状的严重性、阻

滞其进展、消除疾病症状、延缓疾病的进展或提高治疗功效中的一种或多种。如本文所用的，术语疾病是指MD，包括DMD和BMD，或任何其他肌养蛋白相关疾病。

[0180] 应当理解，本文所述载体中的组合物意图应用于整个说明书中描述的其他组合物、方案、方面、实施方案和方法。

[0181] IV. 方法和试剂盒

[0182] 在其他实施方案中，期望用于靶向肌肉的方法，包括骨骼肌、心肌和/或平滑肌。这可能包括静脉内注射或肌内注射。然而，可以选择其他递送途径。

[0183] 在某些实施方案中，本发明的组合物被特异性地靶向(例如，通过直接注射)至心。在某些实施方案中，所述组合物或在心(例如，心肌细胞)中特异性表达。已经描述了用于优先靶向心细胞和/或使脱靶(off-target)非心基因转移减至最小的方法。参见，例如，Matkar PN等人, Cardiac gene therapy: are we there yet?Gene Ther. 2016年8月;23(8-9):635-48. doi: 10.1038/gt.2016.43. Epub 2016年4月29日；专利公开US20030148968A1、US20070054871A1、WO2000038518A1、US7078387B1、US6162796A和WO1994011506A1。

[0184] 在某些实施方案中，将诸如美国专利7,399,750中的那种的方法用于通过诱导低体温、心与循环的隔离以及接近或完全心脏停搏来增加携带感兴趣的基因的载体在心中的停留时间。透化剂是所述方法的必需组分，并在病毒施用期间用于增加心细胞对病毒的摄取。所述方法特别好地适合于病毒载体，其中基因表达可能对心肌高度特异，且特别在rAAV载体的情况下，可以在没有心肌炎症的病征的情况下长期维持表达。可以使用又另一种系统和技术，包括，但不限于，例如，“生物起搏器(bio-pacemaker)”，例如在美国专利8,642,747、US-2011-0112510中所描述的。

[0185] 在一个实施方案中，通过国际公开号WO2005027995A2中描述的总体心肌灌注方法来完成递送。在另一个实施方案中，通过于2004年9月24日提交的国际专利申请No. PCT/US2004/031322中描述的基因转移方法来完成递送。简而言之，所述方法包括通过将主体的微血管区域放血，并在高流体静力压下使用如在灌注过程中保护心和肺以蛋白器官所需的灌流套管和气囊的构型将复合体递送至所述区域，来将本发明的微肌养蛋白相关蛋白转移至肌细胞。提供了供载体的全身递送之用的气囊式导管，其具有基本上在插入主体内的主动脉或血管的整个长度上延伸的气囊。在又另一个实施方案中，本发明提供了经由灌注回路的递送，并且提供了用于在心肺旁路手术期间将物质原位递送至主体的心的外科手术方法。灌注回路限定了用于在外科手术程序过程中通过冠状循环回路使含有大分子复合体的溶液通过主体的心再循环的途径，其中防止物质被递送至主体的其他器官。

[0186] 在一个方面，本文提供了一种药物组合物，其包含在制剂缓冲液(即，媒介物)中的肌养蛋白超家族三重剪接突变蛋白、编码肌养蛋白超家族三重剪接突变蛋白的核酸序列、表达盒或包含这种核酸序列的载体。在一个实施方案中，制剂进一步包含溶解在水性悬浮液体中的表面活性剂、防腐剂、赋形剂和/或缓冲液。在一个实施方案中，缓冲液是PBS。各种合适的溶液是已知的，包括那些包括以下一种或多种的溶液：缓冲盐水，表面活性剂，和调节至相当于约100 mM氯化钠(NaCl)至约250mM氯化钠的离子强度的生理上相容的盐或盐的混合物，或调整为相等离子浓度的生理上相容的盐。适当地，将制剂调节至生理上可接受的pH，例如，在pH 6至8，或pH 6.5至7.5，pH 7.0至7.7或pH 7.2至7.8的范围内。

[0187] 合适的表面活性剂或表面活性剂的组合可以选自无毒性的非离子型表面活性剂。在一个实施方案中，选择在伯羟基封端的双官能嵌段共聚物表面活性剂，例如，如Pluronic® F68 [BASF]，也称为泊洛沙姆(Poloxamer)188，其具有中性pH，具有8400的平均分子量。可以选择其他表面活性剂和其他泊洛沙姆，即，由被两个聚氧乙烯(聚环氧乙烷)的亲水链侧接的中心聚氧丙烯(聚环氧丙烷)疏水链组成的非离子型三嵌段共聚物，SOLUTOLHS 15(聚乙二醇-15羟基硬脂酸酯(Macrogol-15 Hydroxystearate))，LABRASOL(聚羟辛酸甘油酯)，聚氧乙烯10油基醚(polyoxy 10 oleylether)，吐温(聚氧乙烯山梨糖醇酐脂肪酸酯)，乙醇和聚乙二醇。在一个实施方案中，制剂含有泊洛沙姆。这些共聚物通常用字母“P”(对于泊洛沙姆)再加上三个数字命名：前两个数字x 100给出聚氧丙烯核心的近似分子量，且最后一个数字x 10给出百分比聚氧乙烯含量。在一个实施方案中，选择泊洛沙姆188。表面活性剂可以以最多到悬浮液的约0.0005％至约0.001％的量存在。

[0188] 在一个实例中，制剂可含有例如缓冲盐溶液，其包含在水中的氯化钠、碳酸氢钠、葡萄糖、硫酸镁(例如，硫酸镁·7H2O)、氯化钾、氯化钙(例如，氯化钙·2H2O)、磷酸氢二钠及其混合物中的一种或多种。适当地，对于鞘内递送，摩尔渗透压浓度在与脑脊液相容的范围内(例如，约275至约290)；参见，例如，emedicine.medscape.com/article/2093316-overview。任选地，对于鞘内递送，可以将可商购获得的稀释剂用作悬浮剂，或与另一种悬浮剂和其他任选的赋形剂组合。参见，例如，Elliotts B®溶液[Lukare Medical]。

[0189] 在其他实施方案中，制剂可含有一种或多种渗透促进剂(permeation enhancer)。合适的渗透促进剂的实例可包括例如甘露糖醇、甘油胆酸钠、牛磺胆酸钠、脱氧胆酸钠、水杨酸钠、辛酸钠、癸酸钠、月桂基硫酸钠、聚氧乙烯-9-十二烷基醚或EDTA。

[0190] 如本文所用的，“载体”包括任何和所有溶剂、分散介质、媒介物、包衣、稀释剂、抗细菌和抗真菌剂、等渗和吸收延迟剂、缓冲液、载体溶液、悬浮液、胶体等。这种介质和试剂用于药物活性物质的用途是本领域众所周知的。补充活性成分也可以并入组合物中。递送载体例如脂质体、纳米型胶囊、微粒、小球体、脂质颗粒、小泡等可以用于将本发明的组合物引入合适的宿主细胞中。特别地，rAAV载体可以被配制用于包囊在脂质颗粒、脂质体、小泡、纳米球或纳米颗粒等中递送。在一个实施方案中，药物组合物中包括治疗有效量的所述载体。载体的选择不是本发明的限制。其他常规的药学上可接受的载体，例如防腐剂或化学稳定剂。合适的示例性防腐剂包括氯丁醇、山梨酸钾、山梨酸、二氧化硫、没食子酸丙酯、对羟基苯甲酸酯类、乙基香兰素、甘油、酚和对氯苯酚。合适的化学稳定剂包括明胶和清蛋白。

[0191] 短语“药学上可接受的”是指当施用于宿主时不产生变应性或相似的不适当反应的分子实体和组合物。

[0192] 如本文所用的，术语“剂量”或“量”可以指在治疗过程中递送至主体的总剂量或量，或在单个单位(或多个单位或分开的剂量(split dosage))施用中递送的剂量或量。

[0193] 同样，可以以剂量单位配制载体组合物以含有在约1.0 x 10^9个颗粒至约1.0 x 10^{18}个颗粒的范围内的载体量(以治疗一个主体)，包括在所述范围内的所有整数或分数量，并且对于人患者优选1.0 x 10^{12}个颗粒至1.0 x 10^{14}个颗粒。在一个实施方案中，将组合物进行配制以含有每剂至少1x10^9、2x10^9、3x10^9、4x10^9、5x10^9、6x10^9、7x10^9、8x10^9或9x10^9个颗粒，包括在所述范围内的所有整数或分数量。在另一个实施方案中，将组合物进行配制以含有每剂至少1x10^{10}、2x10^{10}、3x10^{10}、4x10^{10}、5x10^{10}、6x10^{10}、7x10^{10}、8x10^{10}或9x10^{10}个颗粒，包

括在所述范围内的所有整数或分数量。在另一个实施方案中，将组合物进行配制以含有每剂至少$1x10^{11}$、$2x10^{11}$、$3x10^{11}$、$4x10^{11}$、$5x10^{11}$、$6x10^{11}$、$7x10^{11}$、$8x10^{11}$或$9x10^{11}$个颗粒，包括在所述范围内的所有整数或分数量。在另一个实施方案中，将组合物进行配制以含有每剂至少$1x10^{12}$、$2x10^{12}$、$3x10^{12}$、$4x10^{12}$、$5x10^{12}$、$6x10^{12}$、$7x10^{12}$、$8x10^{12}$或$9x10^{12}$个颗粒，包括在所述范围内的所有整数或分数量。在另一个实施方案中，将组合物进行配制以含有每剂至少$1x10^{13}$、$2x10^{13}$、$3x10^{13}$、$4x10^{13}$、$5x10^{13}$、$6x10^{13}$、$7x10^{13}$、$8x10^{13}$或$9x10^{13}$个颗粒，包括在所述范围内的所有整数或分数量。在另一个实施方案中，将组合物进行配制以含有每剂至少$1x10^{14}$、$2x10^{14}$、$3x10^{14}$、$4x10^{14}$、$5x10^{14}$、$6x10^{14}$、$7x10^{14}$、$8x10^{14}$或$9x10^{14}$个颗粒，包括在所述范围内的所有整数或分数量。在另一个实施方案中，将组合物进行配制以含有每剂至少$1x10^{15}$、$2x10^{15}$、$3x10^{15}$、$4x10^{15}$、$5x10^{15}$、$6x10^{15}$、$7x10^{15}$、$8x10^{15}$或$9x10^{15}$个颗粒，包括在所述范围内的所有整数或分数量。在另一个实施方案中，将组合物进行配制以含有每剂至少$1x10^{16}$、$2x10^{16}$、$3x10^{16}$、$4x10^{16}$、$5x10^{16}$、$6x10^{16}$、$7x10^{16}$、$8x10^{16}$或$9x10^{16}$个颗粒，包括在所述范围内的所有整数或分数量。在另一个实施方案中，将组合物进行配制以含有每剂至少$1x10^{17}$、$2x10^{17}$、$3x10^{17}$、$4x10^{17}$、$5x10^{17}$、$6x10^{17}$、$7x10^{17}$、$8x10^{17}$或$9x10^{17}$个颗粒，包括在所述范围内的所有整数或分数量。在一个实施方案中，对于人应用，剂量可以从每剂$1x10^{10}$到约$1x10^{12}$个颗粒不等，包括在所述范围内的所有整数或分数量。在一个实施方案中，向主体递送治疗有效量的本文所述的载体。如本文所用的，“治疗有效量”是指包含编码肌养蛋白超家族三重剪接突变蛋白的核酸序列的组合物的量，其在靶细胞中递送和表达足以实现功效的量的酶。或“治疗有效量”是指递送至主体的包含肌养蛋白超家族三重剪接突变蛋白的组合物的量。在一个实施方案中，载体的剂量为约每kg体重1×10^{9}个颗粒(例如，基因组拷贝，GC)至约每kg体重1×10^{16}个颗粒，包括在所述范围内的所有整数或分数量和端点。在另一个实施方案中，剂量是每kg体重1×10^{10}个颗粒至约每kg体重1×10^{13}个颗粒。

[0194] 载体的剂量将主要取决于诸如正被治疗的状况，患者的年龄、重量和健康状况的因素，且因此在患者之间可能不同。例如，载体的治疗有效的人剂量通常在约1 ml至约100 ml溶液的范围内，所述溶液含有约1×10^{7}至1×10^{16}个基因组或颗粒载体的浓度。将调节剂量以相对于任何副作用平衡治疗益处，并且这种剂量可以取决于使用重组载体的治疗应用而变化。可以监控转基因的表达水平以确定结果产生载体的剂量的频率，优选含有小基因的AAV载体。任选地，可以利用与用于治疗目的所描述的那些相似的剂量方案以用于使用本发明的组合物进行免疫。

[0195] 任选地，用肌养蛋白超家族三重剪接突变蛋白(例如，毫微-肌养蛋白相关蛋白或毫微-肌养蛋白)或表达肌养蛋白超家族三重剪接突变蛋白的载体进行的疗法可以与其他疗法组合。

[0196] 肌养蛋白超家族三重剪接突变蛋白(例如，毫微-肌养蛋白相关蛋白或毫微-肌养蛋白)的表达可以通过免疫荧光染色和免疫印迹(蛋白印迹)来检测。肌养蛋白超家族三重剪接突变蛋白(例如，毫微-肌养蛋白相关蛋白或毫微-肌养蛋白)疗法可通过测量肌纤维质膜上缺少的DAP复合体(包括肌膜蛋白聚糖复合体)来监控，所述复合体一般由于原发性肌养蛋白缺乏而在未治疗的营养不良性肌肉中找不到。另一方面，肌养蛋白超家族三重剪接突变蛋白(例如，毫微-肌养蛋白相关蛋白或毫微-肌养蛋白)疗法可通过评估被保护以免于病理表型的肌肉来监控。

[0197] 在一个方面，本发明提供了供临床医生或其他人员使用的试剂盒。一般，这种试剂盒将包括本发明的突变蛋白或载体，以及任选地用于其重构和/或递送的说明书。在另一个实施方案中，试剂盒将包括在生理上相容的盐溶液中的突变蛋白或载体，以及任选地用于稀释并执行本文所述方法的说明书。

[0198] 本发明的试剂盒还可包括气囊式导管，以促进体细胞基因转移，如所述的(国际专利申请No. PCT/US2004/030463，或通过2004年9月24日提交的国际专利申请No. PCT/US2004/031322中描述的基因转移方法)，氧输送剂和/或至少一种体外循环支持和氧合系统的一次性元件。例如，至少一种一次性元件可以是充氧器，其具有中空主体、与主体内部流体连通的液体入口、与主体内部流体连通的液体出口、用于向气室内部提供气体的气体入口、至少一个将气室与主体内部隔开的透气膜以及允许气体从气室退出的气体出口，由此使得能够在主体内部的流体和气室中的气体之间进行气体交换。所述充氧器可以如美国专利No.6,177,403中所述构造，其中透气膜包括在管的至少一部分内伸展的PTFE管，并且其中气室包括PTFE管的内部。

[0199] 应当理解，本文描述的方法和试剂盒中的组合物意图应用于整个说明书中描述的其他组合物、方案、方面、实施方案和方法。

[0200] 术语“一个(a)”或“一种(an)”是指一个或多个。照这样，术语“一个(a)”(或“一种(an)”)、“一个或多个”和“至少一个”在本文中可互换使用。

[0201] 单词“包含(comprise)”、“包含(comprises)”和“包含(comprising)”将被非遍举地而不是排他地解释。单词“由……组成(consist)”、“由……组成(consisting)”及其变形要排他地解释，而不是非遍举地。虽然说明书中的各种实施方案是使用“包含”措辞来呈现的，但是在其他情况下，也意图使用“由……组成”或“基本上由……组成”的措辞来解释和描述相关的实施方案。

[0202] 除非另有规定，否则术语“约”包括在±10％之内并且包括±10％的变化。

[0203] 除非在本说明书中另有定义，否则本文所用的技术和科学术语具有与本领域普通技术人员和参考公开的文本所通常理解的相同的含义，所述公开的文本向本领域技术人员提供了对本申请中使用的许多术语的一般指导。

[0204] 实施例

[0205] 肌养蛋白的一级结构和轻于DMD的疾病BMD的分子基础提示了构建用于DMD中疗法的较小的、部分功能的蛋白的概念上简单的方法。肌养蛋白长重复结构域的单个邻接部分的内部缺失被用于实现BMD-样部分长度肌养蛋白变体，其定位于全长肌养蛋白正常占据的细胞地址，且从而部分取代了肌养蛋白的一种或多种关键生理功能。基于肌养蛋白充当分子“减震体”的观点，广泛假定对于所述作用中的正常功能将需要蛋白的全长。预期的结果是，在适当的测试下，所述整类重组蛋白将为DMD患者赋予BMD-样表型。体细胞递送的部分长度重组肌养蛋白均未完全使临床前研究中最灵敏的病理学测定正常化，从而提示由多个团队对于临床开发做好准备的载体将最好也不过暂时使DMD中疾病进展的速度“贝克尔化(Beckerize)”。短于和长于野生型肌养蛋白都可与贝克尔肌营养不良(BMD)中的严重疾病有关，从而表明肌养蛋白的机械作用并不像依赖于长度的“减震体”那样简单。杜兴和贝克尔肌营养不良中的这种基因型/表型相关性充当开发Dp427的低分子量取代物(包括潜在非免疫原性的肌养蛋白旁系同源物(paralog)肌养蛋白相关蛋白的衍生物)的起点。基于对肌

养蛋白和肌养蛋白相关蛋白的力学生物学的新理解,我们已开发了肌养蛋白相关蛋白或肌养蛋白变体,其可以递送至患者,以取代在处理上文指出的DMD-特异性局限性中先前研究的基因治疗的功效和安全性。

[0206] 实施例1-肌联蛋白而不是肌养蛋白的进化与运动动力的可调节性(scalability)相关

[0207] A. 结果与讨论

[0208] 在大型动物中,总是由肌节肌球蛋白给快速运动以动力,而最快运动的单细胞真核生物和最早分支的动物谱系使用纤毛动力蛋白作为主要的运动动力源(Colin, S. P.等人,Stealth predation and the predatorysuccess of the invasive ctenophore Mnemiopsis leidyi. Proc Natl Acad Sci U S A 107, 17223-17227 (2010);Srivastava, M.等人,The Trichoplax genome and the nature of placozoans.Nature 454, 955-960 (2008);Srivastava, M.等人,The Amphimedon queenslandica genome and the evolution of animal complexity. Nature 466, 720-726 (2010);Ryan,J. F.等人,The genome of the ctenophore Mnemiopsis leidyi and its implications for cell type evolution. Science 342, 1242592,doi:10.1126/science.1242592 (2013);和Moroz, L. L.等人,The ctenophore genome and theevolutionary origins of neural systems. Nature 510, 109-114, doi:10.1038/nature13400 (2014))。驱动从动力蛋白到肌球蛋白的进化转换的选择压力必须反映由这些发动蛋白在其中实现最大能量密度(powerdensity)的细胞器施加的几何约束,而肌节但不是纤毛适合于三维调节。这种关键性转换的分子基础很少理解。在这里,我们显示肌节的出现和与脊索动物肌联蛋白直向同源的大量、含聚-IgG-重复序列的蛋白的种系基因组学重构的出现有关,而肌养蛋白及其相关的膜结合糖蛋白复合体则在较早分支谱系趋异之前就逐渐出现。我们已经鉴定了保留推断的祖先肌联蛋白超基因结构的无脊椎动物物种,从而提供了以前使基因直向同源性和具有辐射和两侧对称性的动物肌节的共同起源难理解的基因重排的统一观点。令人惊讶地,基因结构提供了令人信服的证据证明肌养蛋白的杆结构域的异常尺寸反映了微管结合蛋白的旁系同源类的历史遗产,其中对逐渐增加的长度的选择发生在肌节萌芽之前。这些发现对于肌养蛋白的力学生物学和用于肌营养不良的治疗用途的小型化蛋白的设计具有关键意义(实施例2和3)。我们的重构提示,对细胞形态学和机体发育布局的几何约束要求在可将肌球蛋白以给快速、与尺度无关的运动以动力所需要的密度安全地排列到肌节中之前皮质细胞骨架和胞外基质之间强而柔韧的连接。

[0209] 肌联蛋白是人蛋白质组中最大的蛋白,以单体形式充当用于肌节形成的主要支架(Zoghbi, M. E., Woodhead, J. L., Moss, R. L.&Craig, R. Three-dimensional structure of vertebrate cardiac muscle myosin filaments. Proc Natl Acad Sci U S A 105,2386-2390, doi:10.1073/pnas.0708912105 (2008);和Kontrogianni-Konstantopoulos, A., Ackermann, M. A., Bowman, A. L., Yap, S. V.&Bloch, R. J. Muscle giants: molecular scaffolds in sarcomerogenesis. Physiol Rev 89, 1217-1267,doi:10.1152/physrev.00017.2009 (2009))。在脊椎动物中,肌联蛋白主要由免疫球蛋白(IgG)和纤连蛋白III型(Fn3)结构域组成,所述结构域组织成形成跨越半肌节的极化丝的“超重复序列(super-repeats)”,而独特的N-和C-末端分别定位于Z盘和M线内。然而,

以前在无脊椎动物物种中鉴定的肌联蛋白样蛋白在数目、一级结构、结构域组成和长度方面广泛趋异，从而使功能直向同源性的描述变得复杂化(Tskhovrebova,L.&Trinick, J. Titin: properties and family relationships. Nat Rev Mol Cell Biol 4, 679-689 (2003))。我们的发现表明，所述一般结构的“祖先肌联蛋白超基因”已经经历了广泛的谱系特异性基因组重排和模块重复序列扩充。

[0210] 在脊椎动物中，横纹肌纤维在工作负荷下的生存力取决于最外层肌节与胞外基质之间的肌养蛋白依赖性机械连接所赋予的膜保护(Hoffman, E. P., Brown, R. H., Jr.& Kunkel, L. M. Dystrophin: the protein product of the Duchenne muscular dystrophy locus. Cell 51, 919-928,doi:0092-8674(87)90579-4 [pii] (1987))。肌养蛋白分子量的约75％是由大的由24个血影蛋白重复序列结构域组成的中央“杆结构域”贡献的，而侧翼结构域在两端确立了黏着接触(图6A;和未显示的数据)。患有贝克尔肌营养不良(BMD)的患者可以具有局限于编码杆结构域的外显子的截短缺失或加长重复，从而引起了肌养蛋白的生理功能是否已在24个重复序列中得到最优化的问题。我们探询了在整个后生动物系统发生中，肌养蛋白直向同源物中血影蛋白样重复序列的数目是否在选择压力下增加，这可能与等级捕食食物链进化过程中所选择的分类单元的逐渐增加的能量输出有关。如所示的(图6C)，预测在刺胞动物-Bilatarian分开之前存在的祖先肌养蛋白已具有与人中相同长度的杆结构域，但是我们不能找到在较早直向同源物中显著较短的杆的证据。有趣的是，我们的系统发生分析提供了强有力的证据证明跨膜肌养蛋白结合蛋白复合体的出现远早于后生动物的多细胞性，而几乎所有疾病牵连的组分的直向同源物都存在于后生动物的单细胞姐妹群中(数据未显示)。最早的祖先肌养蛋白直向同源物既缺乏N-末端肌动蛋白结合结构域(ABD)又缺乏整个杆结构域，且仅由推定的肌营养不良蛋白聚糖结合性C-末端“WW-EF-ZZ”结构域组成(图6A，图6E;和未显示的数据)。最早的具有“现代”肌养蛋白直向同源物(即N-末端ABD和延长的杆结构域)的分支谱系是扁盘动物物种丝盘虫(*T. adherens*)，其中杆结构域的大小与人的相似(数据未显示)。因此，肌养蛋白在肌联蛋白的IgG扩充和肌节出现之前处于“全长”;然而，杆结构域的进化谱系迄今仍未解决，这是因为同源蛋白之间的显著序列趋异引起了“长分支吸引”假象。

[0211] 我们通过鉴定比序列中个别氨基酸或核苷酸进化更慢的祖先特征状态处理了这个问题:相对于编码的蛋白结构域的隐蔽马尔科夫模型的内含子的位置和相位。我们发现肌养蛋白和MACF1的血影蛋白重复序列在HMM共同位置46处以视觉明显的模式共享保守的相位0内含子，与血影蛋白基因的随机分布的内含子形成鲜明对比(图6A和图6B)。通过我们仅描绘遥远相关物种的直向同源基因中共享的那些着重指出了相关内含子位置的进化停滞的例证(例如，注意使β重血影蛋白ORF延伸13个重复序列的祖先部分基因重复的证据)(图6B)。所述杆结构域特征状态分析将MACF直向同源物而不是β-血影蛋白鉴定为肌养蛋白CH和杆结构域的最接近供体，从而暗示了所述分支分类群的名称“dystroplakin”(图6C)。基因结构构成了强有力的证据证明当编码祖先MACF1-样spectroplakin的N-末端部分的基因的部分重复变得与编码祖先Dp71-样(WW-EF-ZZ)肌养蛋白直向同源物的基因顺式连接时，进化地出现肌养蛋白。在选择的谱系的巨型MACF直向同源物中，存在编码血影蛋白重复序列的外显子的最近串联重复的证据。这支持了重构，其中在微管-肌动蛋白交联“支柱”的细胞背景中而不是在肌养蛋白本身的背景中出现了用于杆结构域的逐步加长的选择压力

(并且在一些谱系中继续存在)。

[0212] 比较肌联蛋白和肌养蛋白中重复结构域的分子进化揭示了重要的差别。点矩阵揭示了肌联蛋白IgG和Fn3重复序列区域的谱系特异性“踏车(treadmilling)”的证据(数据未显示),估计可能在可互换的区域性串联倍增的基础上,只要总蛋白足够长以促进肌节发生(sarcomerogenesis)。在肌养蛋白直向同源物中,个别血影蛋白重复序列的类似转换看来似乎已几乎不存在,从而提示针对其达至少6亿年的强负选择(数据未显示)。基于这些结果,我们提出了肌养蛋白的血影蛋白重复序列的不能互换性的模型,从而反映了蛋白在纵向力传递中的作用,由此相邻的血影蛋白重复序列之间的氨基酸相互作用必须通过进化偶联来保持(Hopf, T. A.等人,Mutation effects predicted from sequenceco-variation. Nat Biotechnol 35, 128- 135, doi:10.1038/nbt.3769 (2017))以维持抗张强度(数据未显示)。在所述模型中,纯化选择已抵消了以前已与其祖先的相邻配偶体经历了偶联氨基酸进化的趋异血影蛋白样重复序列的新的并列(通过内部基因缺失或重复)(如可在BMD发病机理中观察到的)。通过在血影蛋白和斑蛋白提供的趋异模板的基础上对比肌养蛋白杆结构域的相邻三股螺旋的备择同源性模型进一步支持了所述祖先事件的重构(数据未显示)。换句话说,肌养蛋白的分子进化与杆结构域的抗张强度比其长度更重要的提议是一致的,其中后者是其历史遗产的副产物。由于蛋白定位于细胞骨架皮质的超薄边缘,所以维持对跨膜力传递问题的结构上冗余的祖先解决方案的代谢成本是无关紧要地小的。这个概念对于用于疾病治疗的转基因设计具有关键意义,如在实施例2和3中详细证实的。

[0213] B. 材料和方法:

[0214] RNA-Seq:星状海葵(*Nematostella vectensis*)的参考转录物组是由JGI基因组装配公开的原始簇集ESTs(Putnam, N. H.等人,Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomicorganization. Science 317, 86-94 (2007))以及由Finnerty Lab产生的起源于美国新泽西州(New Jersey, USA)(株系NJ3)的克隆谱系产生的转录物组[Lubinski等人,修改中]装配的。使用具有100%序列同一性的截断值的CD-HIT从合并的装配中去除冗余的重叠群。

[0215] 基本局部比对搜索工具(BLAST)搜索:使用具有预定参数(BLOSUM62矩阵,期望的E值阈值:10,缺口成本存在:11,缺口成本延伸:1)的BLASTp和/或tBLASTn算法对基因组进行blast。在未鉴定出全长同源物的场合,则下载评分最高的部分长度命中周围的基因组区域,并进行从头基因模造(参见基因模造方法部分)。使用tBLASTn算法,从NCBI BLAST服务器针对转录物组鸟枪法装配(TSA)数据库再次使用预定参数blast转录物组。

[0216] 基因模造:在没有相应的RNAseq数据的情况下,使用与在研究中的物种最紧密相关的所列生物的生物特异性基因发现参数,从FGENESH套件(www.softberry.com)的程序获得了蛋白编码基因模型。

[0217] 蛋白结构域分析:通过使用HMMER软件包的欧洲生物信息学研究所(European Bioinformatics Institute)的HMMscan功能(www.ebi.ac.uk/Tools/hmmer/-search/hmmscan)针对Pfam、TIGRFAM、CATH-Gene3D、超家族(Superfamily)和PIRSF蛋白家族HMM数据库运行一级氨基酸序列来分析蛋白结构域(Finn, R. D.等人,HMMER web server: 2015 update. Nucleic Acids Res 43,W30-38, doi:10.1093/nar/gkv397 (2015))。

[0218] 用于内含子位置/相位鉴定的血影蛋白重复序列比对:将所有Pfam序型-HMM可鉴定的血影蛋白重复序列结构域与HMMscan中的Pfam血影蛋白重复序列共有序列进行比对。根据它们相对于共有序列的比对,将这些血影蛋白重复序列序贯比对到多序列比对中。

[0219] 内含子位置/相位鉴定:使用具有94%序列同一性截断值的MacVector(v15.1)(macvector.com)中的点矩阵功能,将ORF注释的cDNA序列与其编码基因组支架进行比对。内含子的位置和相位被鉴定为比对中的断点。每个内含子位置和相位均通过分别紧接在每个100%同一性比对的区组之后和之前的基因组DNA序列中的共有剪接位点供体(-GT)和受体位点(AG-)的存在来确认。

[0220] 推断的祖先内含子鉴定:推断的祖先内含子是在智人(H. sapiens)和大堡礁海绵或星状海葵的直向同源蛋白之间共享的那些。

[0221] 点矩阵:cDNA/DNA、蛋白/基因组DNA和蛋白/蛋白点矩阵是在MacVector(v15.1)(macvector.com)中生成的。

[0222] 肌养蛋白血影蛋白重复序列的同源性模造:Phyre2用于模造来自人肌养蛋白的相邻血影蛋白重复序列(www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index)(Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N.&Sternberg, M. J. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc 10, 845-858, doi:10.1038/nprot.2015.053 (2015))。使用β2-血影蛋白(PDB-ID = 3EDV)(Davis, L.等人,Localization andstructure of the ankyrin-binding site on beta2-spectrin. J Biol Chem 284, 6982-6987, doi:10.1074/jbc.M809245200 (2009))或网蛋白(PDBID= 5J1G)(Ortega, E.等人,The Structure of the Plakin Domain of Plectin Reveals an Extended Rod-like Shape. JBiol Chem 291, 18643-18662, doi:10.1074/jbc.M116.732909 (2016))的晶体结构作为同源性模型的模板产生性质不同的同源性模型。

[0223] 数据可获得性:在补充信息中提供了用于支持本文章中研究结果的序列数据。于请求时可从相应的作者处获得所有其他数据。

[0224] 实施例2-使用AAV-介导的微-肌养蛋白相关蛋白递送的有效的用于肌营养不良的基因治疗

[0225] A. 结果和讨论

[0226] 杜兴肌营养不良(DMD)基因的必需蛋白产物是肌养蛋白(Hoffman, E. P., Brown, R. H., Jr.&Kunkel, L. M. Dystrophin: the protein product of the Duchenne muscular dystrophy locus. Cell 51, 919-928, doi:0092-8674(87)90579-4 [pii](1987)),杆样的427kd蛋白(Koenig, M., Monaco, A. P.&Kunkel, L. M. The complete sequence of dystrophin predicts a rod-shaped cytoskeletalprotein. Cell 53, 219-226(1988)),其通过将皮质细胞骨架与胞外基质连接(Ibraghimov-Beskrovnaya, O.等人,Primary structure of dystrophin-associated glycoproteins linking dystrophin to the extracellular matrix. Nature 355, 696-702,doi:10.1038/355696a0 (1992))保护横纹肌细胞免受收缩诱导的损伤(Petrof, B. J., Shrager, J. B., Stedman, H. H., Kelly, A. M.&Sweeney, H. L. Dystrophin protects the sarcolemma from stresses developed during musclecontraction.

Proceedings of the National Academy of Sciences of the United States of America 90, 3710-3714 (1993))。大多数DMD患者具有多外显子移码缺失,而许多患有较轻等位疾病贝克尔MD的患者具有改变肌养蛋白的150nm杆结构域的长度的保持读框的(frame-preserving)突变(Monaco, A. P., Bertelson, C. J., Liechti-Gallati, S., Moser, H.&Kunkel, L. M. An explanation for the phenotypic differences between patients bearing partial deletions of the DMD locus. Genomics 2, 90-95(1988);和Koenig, M.等人,The molecular basis for Duchenne versus Becker muscular dystrophy: correlation ofseverity with type of deletion. American journal of human genetics 45, 498-506(1989))。我们对肌养蛋白的深入进化史的分析提示,杆结构域是从更长的细胞骨架蛋白补选的,并在强有力的横纹肌出现之前出现(实施例1)。在这里,我们显示了密码子优化的合成转基因,其编码从旁系同源蛋白肌养蛋白相关蛋白合理设计的非免疫原性的肌养蛋白的25 nm取代物(Tinsley, J. M.等人,Primary structure of dystrophin-related protein. Nature 360, 591-593,doi:10.1038/360591a0 (1992)),以保持小型化杆结构域的抗张强度,预防动物模型中肌营养不良的最有害的组织学和生理学方面。在AAV载体向新生儿肌养蛋白缺陷的mdx小鼠全身性施用后,在到成体重量的整个生长期间完全抑制肌肉坏死和再生的所有组织学和生物化学标记。在于最多到4 kg体重类似治疗的肌养蛋白缺陷的狗中,转基因的全身分布和表达在没有细胞介导的蛋白产物免疫识别的情况下预防了肌肉坏死,从而提示由对全长肌养蛋白相关蛋白的中枢免疫耐受的保护。这些发现支持这样的模型,其中抗张强度是肌养蛋白和肌养蛋白相关蛋白杆的基本特征,而在大多数谱系中它们的150nm长度通过针对以强度为代价减少长度的突变的纯化选择来保持。

[0227] 尽管已使用衍生自人腺病毒的内部缺失载体实现了编码全长肌养蛋白的12 kb cDNAs的体细胞转移,但由于复合载体衣壳的免疫原性和有限的生物分布,所述方法已被放弃(Clemens, P. R.等人,In vivo muscle gene transfer of full-length dystrophin with an adenoviral vector that lacks all viral genes. Gene therapy 3, 965-972 (1996))。已显示衍生自人腺伴随病毒(AAVs)的多种载体促进全身基因转移(Wang, B.等人,Adeno-associated virus vector carrying human minidystrophin genes effectively amelioratesmuscular dystrophy in mdx mouse model. Proc Natl Acad Sci U S A 97, 13714-13719. (2000);Harper, S. Q.等人,Modular flexibility of dystrophin: implications for genetherapy of Duchenne muscular dystrophy. Nat Med 8, 253-261. (2002);Gregorevic, P.等人,Systemic delivery of genes to striated muscles using adeno-associated viral vectors. NatMed 10, 828-834 (2004);和Gregorevic, P.等人,rAAV6-microdystrophin preserves muscle function and extends lifespan in severely dystrophic mice. Nat Med 12, 787-789 (2006)),但其克隆能力限于野生型病毒的,大约5 kb。对用于DMD的基因治疗的同样重要的第二约束是大多数患者蛋白缺乏的缺失特性,而重组肌养蛋白具有作为“非自身”蛋白(Mendell, J. R.等人,Dystrophin immunity in Duchenne's muscular dystrophy. NEngl J Med 363, 1429-1437 (2010))以引发慢性自身免疫性肌炎的潜力。我们假设对肌养蛋白分子进化的详细分析可能通过揭示所述蛋白的历史遗产的以前未理解的方面而报

告对这两种约束的合成生物学方法。我们对肌养蛋白遥远历史的重构提示,在蛋白开端时,其杆结构域含有从另一个大得多的支柱样细胞骨架蛋白的N-末端部分补选的24个“血影蛋白样”三股螺旋结构域的重复序列(实施例1)。来自肌养蛋白、肌养蛋白相关蛋白和紧密相关的spectroplakin的三股螺旋重复序列的晶体结构提示,相邻重复序列之间的氨基酸侧链相互作用产生了对杆的强度关键的互锁界面。所述原理可以解释BMD患者中符合读框的缺失和重复引起的表型和脊索动物旁系同源物(例如七鳃鳗(Lamprey))缺失的罕见性,这是因为三股螺旋重复序列的天然序列的大多数破坏都具有局部地削弱杆结构域的潜力。为了将产生“最弱的环节”的危险减至最低,我们集中于在一侧被传统地标记为“铰链2”的无序结构域侧接的缺失,并且还缺失了超出大约ZZ结构域末端的C-末端序列(Ishikawa-Sakurai, M., Yoshida, M., Imamura, M., Davies,K. E.&Ozawa, E. ZZ domain is essentially required for the physiological binding of dystrophin and utrophin to beta-dystroglycan. Hum Mol Genet 13,693-702, (2004);Hnia, K.等人,ZZ domain of dystrophin and utrophin: topology and mapping of a beta-dystroglycan interaction site. Biochem J 401, 667-677,(2007))。为了利用通过在胸腺中早期发育表达获得的中枢免疫耐受(Mesnard-Rouiller,L.等人,Thymic myoid cells express high levels of muscle genes. J Neuroimmunol 148, 97-105, 2003),我们将肌养蛋白中的这些缺失定位到旁系同源蛋白肌养蛋白相关蛋白上,其在脊椎动物进化的早期与肌养蛋白趋异。基于这些考虑,我们合成了基于野生型肌养蛋白相关蛋白mRNA序列的转基因,且然后使用所述序列的人工改造形式改进了表达。在这里,我们就使用基于AAV9的载体和衍生的祖先衣壳“Anc80”以向所有横纹肌全身递送3.5kb合成转基因(AAV9-μU、AAV9-μ肌养蛋白相关蛋白)的盲法(blinded)临床前研究中获得的结果提出了报告。

[0228] 我们最初向重约5 gm的新生儿mdx小鼠中进行了最多到2.5×10^{12}vg AAV-μ肌养蛋白相关蛋白剂量的腹膜内注射,并研究了整个肌肉发育过程中的肌保护(myoprotection)程度。在这些随机化的盲法研究中,我们观察到用AAV9和Anc80两者到肌肉的相等的整体生物分布,并且两者在没有任何毒性病征的情况下在小鼠中均良好耐受(图7A-7C)。在2.5×10^{12}vg/小鼠剂量,重组μ肌养蛋白相关蛋白以足以定性地完全抑制肌营养不良的所有测试的组织学病征的水平表达,所述组织学病征包括肌纤维中央成核现象(centronucleation)(图7A和图7D)、胚胎肌球蛋白重链表达(图7A和7B)、天然肌养蛋白相关蛋白上调(图7A和图9A)、作为蛋白降解标记的MURF1表达、肌核(myonuclear)凋亡(数据未显示)、正在进行的肌肉坏死和单核的细胞浸润(图7B)。在这些病征中,中央成核现象是mdx小鼠中肌保护的定量地最灵敏指示物,这是因为它反映了以前的再生循环。我们首次显示了中央成核现象到与野生型在生物学上不能区别的水平的正常化(或预防)(图7D)。所述观察到的肌保护与心肌和骨骼肌的肌膜中的肌养蛋白结合性糖蛋白复合体(DGC)的持续正常化有关(图7A)。蛋白印迹分析进一步证实表达的μ肌养蛋白相关蛋白足以稳定DGC(图7C)。在整个载体递送后4个月时间段期间(研究的终点)观察到μ肌养蛋白相关蛋白在骨骼肌和心肌中的持续表达,从而表明由单剂治疗赋予的持久的肌保护水平。引人注目地,肌酸激酶(一种反映肌膜透性的生物标志物)与野生型小鼠的在统计学上不能区别(图7E),从而提示发育中早期的密码子优化和重组蛋白超表达相对于在肌肉病理学(myopathology)发病后通过尾静脉注射施用选择性转基因改善了反应(Gregorevic, P. 等人,Systemic delivery of genes to

striated muscles usingadeno-associated viral vectors. Nat Med 10, 828-834, 2004;Odom, G. L.等人,Microutrophin delivery through rAAV6 increases lifespan and improves muscle function indystrophic dystrophin/utrophin-deficient mice. Mol Ther 16, 1539-1545, 2008;Kennedy, T. L.等人,Micro-utrophin Improves Cardiac and Skeletal Muscle Function of Severely Affected D2/mdx Mice.Mol Ther Methods Clin Dev 11, 92-105, 2018)。

[0229] 为了测试AAV9-μ肌养蛋白相关蛋白是否赋予*mdx*小鼠中功能的改善,我们使用了确立的混合式测定,其将握力测试(force grip testing)与意志成分结合,其中我们无损伤地测量了动物的握力后垂直活动(Song, Y.等人,Suite of clinically relevant functional assays to address therapeutic efficacy and disease mechanism in the dystrophic mdx mouse. J ApplPhysiol 122, 593-602, 2017)。我们以前的研究表明,所述混合测试提供了区别*mdx*和野生型小鼠的最灵敏、临床相关的参数之一,从而捕获了与肌养蛋白缺陷的肌肉的过大疲劳反应有因果联系的行为(Kobayashi, Y. M.等人,Sarcolemma-localized nNOS is required tomaintain activity after mild exercise. Nature 456, 511-515, 2008)。所述测试证明了未治疗的和AAV9-μ肌养蛋白相关蛋白治疗的mdx小鼠之间的客观的、引人注目的且统计学上显著的差异,而后一组与野生型小鼠不能区别(图7F)。同样,相对于未治疗的小鼠,治疗的*mdx*小鼠不仅显示出增加的自主跑轮(voluntary wheel)(8周)和下坡踏车(downhilltreadmill running)(16周)距离,而且它们的先体外后体内分离的EDL肌肉还表现出增强的对偏心收缩诱导的损伤的抗性以及增强的通过握力测试的体内肌肉性能。这些发现提示,尽管逆向人工改造的蛋白与肌动蛋白细胞骨架的杆样连接长度比较短,并且缺乏R16-17 nNOS-结合基序,但μ肌养蛋白相关蛋白的早期超表达能够在没有全长肌养蛋白的情况下完全改善表型。

[0230] 这些结果引起了通过全身肌肉转导实现DMD病理生理学的完全而非BMD-样部分逆转的希望;然而,尚不清楚小的和大的营养不良动物之间的大小依赖性差异是否将揭示所述方法的局限性。在一项盲法研究中进一步研究了μ肌养蛋白相关蛋白基因转移的组织学和免疫学后果,其中将5只4-7天龄的金毛猎犬肌营养不良(Golden Retriever Muscular Dystrophy)(GRMD)狗随机化至在没有免疫抑制的情况下以注射时1×10^{13}和3.2×10^{13}vg/kg的剂量静脉内施用AAV9-μ肌养蛋白相关蛋白。注射后六周,我们观察到肌膜中稳健的μ肌养蛋白相关蛋白表达和野生型肌膜蛋白聚糖表达水平的稳定化(数据未显示)。此外,与以前报道的在相同GRMD模型中全身施用异种人肌养蛋白后与免疫肌炎有关的重量减少相反,这些治疗的狗实现了类似于携带者雌性的重量的四倍增加Kornegay, J. N.等人,Widespread muscle expression of an AAV9human mini-dystrophin vector after intravenous injection in neonatal dystrophin-deficient dogs. Mol Ther 18, 1501-1508, 2010)。所述持续的μ肌养蛋白相关蛋白表达与明显降低的肌肉坏死水平,肌纤维最小费雷特直径(Feretdiameter)的单核浸润正常化有关(数据未显示)。在载体施用后5和8周,犬干扰素-γ ELISpot测定揭示在我们的非免疫抑制治疗的GRMD狗中,没有针对AAV衣壳或μ肌养蛋白相关蛋白转基因产物的细胞介导的免疫(数据未显示)。所述概念证明研究的主要局限性归因于分别为25g和25kg的mdx小鼠和GRMD狗的成体重量之间1000-倍的差异,从而基于预期的成体重量将我们在该狗中可达到的AAV9剂量限于2.0×10^{12}vg/kg。在所

述剂量,狗将不可避免地“长得比”载体“快”,正如作为5 gm新生儿用2.15×10^{11}vg治疗的mdx小鼠一样(数据未显示)。我们因此集中于相对早期的组织学分析,以检测重组μ肌养蛋白相关蛋白表达、肌细胞保护以及对全身载体施用的免疫应答。

[0231] 我们的新生儿方法既提供不可逆的肌肉损伤发病之前的早期预防治疗的可能性,又将针对载体衣壳抗原的免疫反应的危险降至最低,这是因为识别野生型AAV的记忆T细胞通过连续环境暴露进行发育Nichols, T.等人,Translational Data from AAV-Mediated Gene Therapy of HemophiliaB in Dogs. Hum Gene Ther Clin Dev, 2014;Calcedo, R.等人,Adeno-associated virus antibody profiles in newborns, children, and adolescents. Clin VaccineImmunol 18, 1586-1588, 2011)。然而,大多数DMD患者一般在两岁后被诊断出来,在那时以前,已经发生了大量肌肉纤维变性、具有单核的细胞侵袭的坏死以及增加的纤维大小变异性(Yiu, E. M.&Kornberg, A. J. Duchenne muscular dystrophy. Journal of paediatrics and child health 51, 759-764, 2015)。为了探索我们的方法在患有DMD的年轻男孩中的可行性,在短暂使用每日1 mg/kg的抗炎剂量的强的松期间,以高达注射时1.25×10^{14}vg/kg的剂量向两只7.5周龄的幼年GRMD狗(Hann和Beetle)静脉内注射AAV9-μ肌养蛋白相关蛋白(Liu, J. M.等人,Effects of prednisone in canine muscular dystrophy. Muscle Nerve 30, 767-773, 2004)(图8A)。注射后四周获取的肌肉活组织检查的免疫染色显示μ肌养蛋白相关蛋白的均匀肌膜表达(图9A和数据未显示)、天然肌养蛋白相关蛋白的抑制(图9A)以及DGC的拯救(图9B)。通过蛋白印迹进一步证实了这一点(图9D)。

[0232] 来自治疗的GRMD狗的肢肌肉的组织病理学表征显示了正在进行的肌肉损伤的几乎完全抑制,如在未治疗的年龄匹配对照中通过高比例的肌肉坏死纤维、过多的钙累积(图8C和图8E)、簇生的再生肌肉纤维(图8D和图8F)、丰富的炎性细胞浸润和脂肪浸润(图10B和数据未显示)所证明的。令人印像深刻的是,AAV9-μ肌养蛋白相关蛋白治疗的狗还显示了几乎完全预防咀嚼肌中的肌肉变性和再生(图8B和数据未显示),其在未治疗的狗中受到了严重影响,这是因为它们表达了独特地强有力的MYH16肌球蛋白同工型(Stedman, H. H.等人,Myosin gene mutation correlates with anatomical changes in the human lineage. Nature428, 415-418 (2004);Toniolo, L.等人,Masticatory myosin unveiled: first determination of contractile parameters of muscle fibers from carnivore jaw muscles. Am JPhysiol Cell Physiol 295, C1535-1542 (2008))。于尸体剖检时(3.5个月龄)的进一步蛋白印迹分析显示了μ肌养蛋白相关蛋白在骨骼肌和心肌中的持久广泛表达(图9C)。与我们以前的GRMD新生儿狗研究一致,干扰素-γ ELISpot测定未揭示针对μ肌养蛋白相关蛋白的高于背景的信号(数据未显示)。此外,与以前在GRMD狗和非人灵长类动物中的研究相对比,未见严重急性毒性的病征(Kornegay, J. N.等人,Widespread muscle expression of an AAV9 human mini-dystrophin vector after intravenous injection in neonataldystrophin-deficient dogs. Mol Ther 18, 1501-1508 (2010);Hinderer, C.等人,Severe Toxicity in Nonhuman Primates and Piglets Following High-Dose IntravenousAdministration of an Adeno-Associated Virus Vector Expressing Human SMN. Hum Gene Ther, (2018);Hordeaux, J.等人,The Neurotropic Properties of AAV-PHP.B Are Limited toC57BL/6J Mice. Mol Ther,

(2018))。重要的是,在两只狗中,在用AAV9-μ肌养蛋白相关蛋白输注后1周测量到血清CK水平的80％下降(图8G),这是与观察到的组织学改善一致的发现。为了在从婴儿期到骨骼成熟的整个肌肉生长过程中实现持久的肌保护,患有DMD的营养不良的狗和男孩可能需要以与mdx小鼠幼崽所需剂量成比例的剂量全身性施用AAV载体,以在最严重受影响的肌肉膈中维持稳健的、均匀的表达(Stedman, H. H.等人,The mdx mouse diaphragm reproduces the degenerative changes of Duchenne muscular dystrophy. Nature 352, 536-539, (1991)),例如1×10^{15}vg/kg新生儿体重(数据未显示)。

[0233] 为了对载体和/或转基因免疫毒性(immunotoxicity)进行严格的临床前评估,我们利用了独特的德国短毛波音达犬(German Shorthaired Pointer)(GSHPMD)缺失无效犬模型(Schatzberg, S. J.等人,Molecular analysis of a spontaneousdystrophin 'knockout' dog. Neuromuscul Disord 9, 289-295. (1999);VanBelzen, D. J.等人,Mechanism of Deletion Removing All Dystrophin Exons in a Canine Model for DMD ImplicatesConcerted Evolution of X Chromosome Pseudogenes. Mol Ther Methods Clin Dev 4, 62-71, (2017))。GSHPMD为中枢性耐受研究提供了优越的平台,这是因为GRMD模型中的可变剪接允许潜在耐受水平的可检测的接近全长的肌养蛋白的连读,如促进以前显示的AAV-编码的犬微肌养蛋白的治疗性、长期全身表达所可能预料的(Schatzberg, S. J.等人,Alternative dystrophin gene transcripts in goldenretriever muscular dystrophy. Muscle Nerve 21, 991-998. (1998);Yue, Y.等人,Safe and bodywide muscle transduction in young adult Duchenne muscular dystrophy dogswith adeno-associated virus. Hum Mol Genet 24, 5880-5890, (2015);Le Guiner, C.等人,Long-term microdystrophin gene therapy is effective in a canine model of Duchennemuscular dystrophy. Nat Commun 8, 16105, (2017))。成体GSHPMD狗(Ned和Grinch)各自经由到对侧胫骨区室中的肌内注射接受了相等剂量(2×10^{12}vg/kg)的AAV9-μ肌养蛋白和AAV9-μ肌养蛋白相关蛋白(图10A)。经由ELISPOT进行的干扰素-γ检测揭示,早在注射后2周,就存在强的针对μ肌养蛋白的全身细胞介导的免疫应答,但尽管从组成型CMV启动子表达对于μ肌养蛋白相关蛋白也不存在所述免疫应答,从而表明中枢免疫耐受的强度(图10B)。注射后4周收集的肌肉活组织检查的免疫染色揭示了μ肌养蛋白相关蛋白的持久表达,但是仅有稀少量的μ肌养蛋白(图10C)。与其在μ肌养蛋白相关蛋白注射的一侧实际上不存在相比,H&E显示在μ肌养蛋白注射的一侧上严重的炎症和单核的细胞浸润(图10D和数据未显示)。这些发现表明,观察到的免疫应答是由μ肌养蛋白驱动的,而不是由载体衣壳驱动的,这是因为相等剂量的载体被注射到两肢。

[0234] 总之,在利用比较种系基因组学方法以鉴定进化约束之后,我们已逆向人工改造了用于安全地全身递送至DMD的鼠和犬模型的肌肉的高度治疗性的3.5 kb合成转基因。我们的盲法研究揭示了令人惊讶地完全的肌保护,只要基因递送的初始水平足以适应后来的肌肉生长。合起来看,这些发现可能使所述领域再聚焦于使用功能上最优化的、非免疫原性的基于肌养蛋白相关蛋白的基因治疗方法作为对于杜兴肌营养不良的治疗。

[0235] B. 材料与方法

[0236] a. 生物信息学和系统发生分析

[0237] 通过多种blast算法,特别是tBLASTn26,27,查询了在图例和原稿文本中列出的物

种的可公开获得的基因组DNA序列，以鉴定与全长人肌养蛋白和肌养蛋白相关蛋白同源的编码序列。对于大多数物种，来自mRNA序列的支持性证据是可获得的以定义内含子/外显子边界。在缺乏这种证据的场合，将FGENESH +(Softberry)与生物特异性基因发现参数和隐蔽马尔科夫模型加类似的基于蛋白的基因预测一起使用，以从组装的重叠群中鉴定推定的编码序列。作为这种方法的内部测试，实际上所有转录物组定义的编码序列均通过FGENESH+程序28正确地鉴定。我们认识到这些mRNA和蛋白序列文件遗漏了可公开获得的基因组DNA的不确定区域中的零星外显子。HMMER(hmmer.janelia.org/search/hmmscan)与E-值定义的截断值一起使用，以定义与钙调理蛋白同源、血影蛋白样重复序列、WW、EF手和ZZ结构域的隐蔽马尔科夫模型匹配的蛋白编码结构域。通过MacVector版本13.5.1中使用默认设置的ClustalW对所有推导的肽序列文件进行比对：具有用于成对比对开放缺口罚分(open gap penalty)10、延伸缺口罚分0.1的参数和用于多重比对开放缺口罚分10、延伸缺口罚分0.2和延迟趋异(delay divergence)30％的参数的Gonnet系列矩阵(Gonnet SeriesMatrix)。使用邻接树构建方法用全长和截短的序列两者产生系统发生重构，而在缺口按比例分布或被忽略的情况下树中的联系(ties)被随机解析且距离被泊松修正，以确定选择是否影响树的拓扑结构。在所有这种情况中，当选择“最佳树(best tree)”模式时，树的拓扑结构对缺口的管理不敏感。具有10000重复的自引模式的备择使用确认了距离系统发生图中的所有分支点。蛋白和DNA矩阵分析基于pam250评分矩阵。使用的缩写：紫海胆，紫色球海胆(Strongylocentrotus purpuratus)，S.pur；文昌鱼，佛罗里达文昌鱼(Branchiostomafloridae)，B.flo；姥鲨(Elephant shark)，米氏叶吻银鲛(Callorhinchusmilii)，C.mil；扬子鳄(Chinese alligator)，扬子鳄(Alligator sinensis)，A.sin；小鼠，小鼠(Mus musculus)，M.mus；狗，家犬(Canis familiaris)，C.fam；人，智人(Homo sapiens)，H.sap；卡罗莱纳安乐蜥(Carolina anole)，北美绿色安乐蜥(Anolis carolinensis)，A.car；普通黑猩猩(common chimpanzee)，黑猩猩(Pantroglodytes)，P.tro；鸭嘴兽，鸭嘴兽(Ornithorhynchidae anatinus)，O.ana.；日本河豚(Japanesepufferfish)，红鳍东方鲀(Takifugu rubripes)，T.rub；热带爪蟾，热带爪蟾(Xenopus tropicalis)，X.tro；D-肌养蛋白；U-肌养蛋白相关蛋白。

[0238] b. 微肌养蛋白相关蛋白转基因设计和载体生产

[0239] 基于BMD/DMD患者之间的序列保守性和基因型-表型相关性的种系基因组学分析，我们在计算机芯片上模造了AAV可编码的小型化肌养蛋白相关蛋白的谱，其保持钙调理蛋白同源结构域、最先三个和最后三个血影蛋白样重复序列以及WW、EF手和ZZ结构域的组合。为了保持相邻的血影蛋白样重复序列的螺旋间环之间的氨基酸侧链相互作用，我们仅集中于最先或最后三个重复序列保持完整的μ肌养蛋白相关蛋白的亚群。为了将免疫原性减至最低，我们仅考虑了可以通过单个内部缺失和C-末端截短组合而产生的那些μ肌养蛋白相关蛋白。尽管血影蛋白样重复序列与共有序列同源，但是趋异是这样的，以致于任何哺乳动物肌养蛋白相关蛋白中相同的十肽之间不能找到剪接。因此，我们使用如在hmmer.janelia.org/search/hmmscan上在线执行的序型隐蔽马尔科夫模型，以定义和注释全长犬肌养蛋白相关蛋白序列(3456 aa，XP_005615306)中的血影蛋白样三股螺旋重复序列边界。我们在新生儿小鼠中使用了编码与犬和人蛋白相匹配的蛋白的转基因，其中通过腹膜内注射AAV45容易地诱导了抗原特异性耐受，但是只有新生儿和较年长狗中的犬形式，

其中预期耐受需要在免疫个体发育过程中更早的出生前暴露于等基因的天然蛋白(Davey, M. G.等人,Induction of Immune Tolerance to ForeignProtein via Adeno-Associated Viral Vector Gene Transfer in Mid-Gestation Fetal Sheep. PLoS One 12, e0171132, (2017))。设计μ肌养蛋白相关蛋白转基因以含有肌动蛋白结合结构域、三股螺旋重复序列1-3和22、近似以前鉴定为"铰链"2的无序的富含脯氨酸的区域以及C-末端WW、EF手和ZZ结构域,从而产生了重组蛋白,其被进行设计以除了在缺失连接处的单个剪接位点之外其余都与犬和人肌养蛋白相关蛋白序列匹配,从而相对于以前报道的转基因将在肌养蛋白缺陷的狗和最终的人中潜在的免疫原性减至最低(Wang,B.等人,Adeno-associated virus vector carrying human minidystrophin genes effectively ameliorates muscular dystrophy in mdxmouse model. Proc Natl Acad Sci U S A 97, 13714-13719. (2000);Harper, S. Q.等人,Modular flexibility of dystrophin: implications for gene therapy of Duchenne musculardystrophy. Nat Med 8, 253-261. (2002);Gregorevic, P.等人, Systemic delivery of genes to striatedmuscles using adeno-associated viral vectors. Nat Med 10, 828-834 (2004);Gregorevic, P.等人,rAAV6-microdystrophin preserves muscle function and extends lifespan in severely dystrophic mice. NatMed 12, 787-789 (2006);和Odom, G. L.等人, Microutrophin delivery through rAAV6 increases lifespan and improves muscle function in dystrophic dystrophin/utrophin-deficient mice.Mol Ther 16, 1539-1545 (2008))。选择供在我们的研究之用的编码序列被选择为由竞争性生物技术公司(GeneArt和DNA 2.0)最优化和合成的cDNAs库中的最高表达候选物。通过在*mdx*小鼠胫骨前肌肌肉中电穿孔50 μg DNA后的免疫荧光染色和蛋白印迹来测定表达。找到选择供进一步使用的合成编码序列以在体外和体内测定中驱动约为编码相同一级结构的重组蛋白的野生型犬cDNA序列30倍的表达。最佳合成的cDNA与野生型之间的显著差异是密码子偏倚的水平,而只有最优化的合成cDNA与哺乳动物肌球蛋白重链的极端偏倚(例如 154 CTG亮氨酸, 0 TTA亮氨酸)紧密匹配。将合成的犬μ肌养蛋白相关蛋白cDNA亚克隆到由CMV立即早期增强子/启动子的833 bp片段或合成启动子spc5-12(分别为CMV和SP)驱动的AAV2表达载体盒中。如以前所述的,AAV9载体由宾夕法尼亚大学临床前载体中心(University of Pennsylvania preclinicalvector core)使用三重转染方法在HEK 293细胞中产生并纯化(Vandenberghe, L. H.等人,Efficient serotype-dependent release offunctional vector into the culture medium during adeno-associated virus manufacturing. Hum Gene Ther 21, 1251-1257, doi:10.1089/hum.2010.107(2010))。在合并用于将2×10^{11}AAV9μ肌养蛋白相关蛋白vg注射到*mdx*小鼠胫骨前肌肌肉中之前,测定了载体制剂的质量、纯度和内毒素水平(Lock, M.等人,Analysis of particle content ofrecombinant adenoassociated virus serotype 8 vectors by ion-exchange chromatography. Hum Gene Ther Methods, 23, 56-64, doi:10.1089/hgtb.2011.217[pii]。

[0240] c. 动物-一般

[0241] A&M大学(A&M University)和宾夕法尼亚大学的动物护理和使用委员会(Animal Care andUse Committee)批准了所有在小鼠和狗中进行的动物实验规程。

[0242] d. 鼠模型载体施用

[0243] 小鼠品系C57BL/10SnJ和*mdx*购自Jackson laboratory(Bar Harbor, ME)。所述研究包括23只C57BL/10SnJ小鼠和30只*mdx*小鼠,均于9±2天龄注射。在接受腹膜内注射AAV9 μ肌养蛋白相关蛋白或磷酸缓冲盐水(PBS)之前,将各个幼崽用Aramis Micro文身试剂盒(Ketchum Manufacturing Inc, 加拿大)进行趾文身,并随机分配到不同的剂量组。在所有注射和组织收获过程中,研究人员均不知情(blinded)。基于所述规程,经由32-规格胰岛素注射器将C57BL/10SnJ和*mdx*幼崽用50-250 μl的作为阴性对照的PBS或在PBS中稀释的AAV9 μ肌养蛋白相关蛋白进行注射。在注射之前,将每只小鼠称重。载体施用后,将所有小鼠放回到其褥草(litter),并在断奶后分离。

[0244] e. 鼠模型组织获得和贮藏

[0245] 根据制度性政策(institutional policy),在约8周龄,*mdx*和C57BL/10SnJ小鼠经历CO_2安死术。基于显示在小鼠中AAV9中<100-倍较低脱靶基因表达的研究,收获心、胫骨前肌、腓肠肌、四头肌、三头肌、腹、膈肌、吸肌肌肉和肝,并进一步处理;其他的被贮藏但不利用Zincarelli, C.等人,Analysis of AAV serotypes 1-9 mediated gene expression and tropism in mice after systemic injection. Mol Ther 16,1073-1080(2008))。将指定的组织学组织样品放置于含有包埋模子(Richard-Allan Scientific)的OCT(Tissue-Tek)中,并在液氮冷却的异戊烷中快速冷冻。将额外的指定的生物组织样品放置于组织容器中,并在液氮中快速冷冻。所有样品都贮藏于-80℃。在-25℃的恒冷切片机(Microm HM550, Thermo Scientific, 美国)上切割5-7 μm厚度的冷冻切片,并将其封固在载玻片(Superfrost Plus, Fisher Scientific, 美国)上。

[0246] f. 基本组织学-苏木精和伊红染色(H&E)和茜素红染色(ARS)

[0247] 将7 μm厚的横切片于室温风干15分钟。然后将载玻片用Harris苏木精染料染色2.5分钟,在蒸馏水中漂洗,浸入0.1%乙酸中达15秒,继之以在自来水中重复漂洗达4分钟,并用1%伊红复染1分钟。作为最后的步骤,将载玻片在乙醇中脱水3次,每次2分钟。拍摄有代表性的非重叠高倍视野(HPF)用于评分(现在显示数据)。还在7 μm厚的横切片上进行了茜素红染色。由10%缓冲的福尔马林磷酸盐固定10分钟后,将切片用PBS洗涤3×5分钟,并与茜素红染料于室温温育15分钟。在所述程序之后,将载玻片在乙醇中洗涤3次并通过cytoseal 60(Thermo Scientific)封固。

[0248] g. 形态度量分析

[0249] 由对样品鉴定不知情的研究人员研究了三组小鼠,C57BL/Sn10J(对照)和*mdx*,它们随机化至用PBS或AAV9 μ肌养蛋白相关蛋白注射。用H&E对来自每种胫骨前肌、腓肠肌、四头肌、三头肌、吸肌和腹肌肌肉的4至5个随机选择的区域进行染色,并用光学显微镜对于中央成核的肌纤维的定量进行筛选。由于肌腱连接(myotendonous junctions)处的区域在mdx和对照两者中的中央成核的肌纤维中均丰富,所以将它们从测量中排除。总计评价了11,649根纤维。

[0250] h. 免疫荧光染色程序:N-末端和C-末端肌养蛋白相关蛋白双重染色

[0251] 通过使用N-末端多克隆(抗全长和重组肌养蛋白相关蛋白)以及C-末端单克隆(抗全长肌养蛋白相关蛋白)抗体两者,对来自所有肌肉样品的切片进行处理用于肌养蛋白相关蛋白免疫染色。在于0.01M PBS(Roche Diagnostics GmbH, Mannheim, 德国)中稀释的Triton X-100(RocheDiagnostics GmbH, Mannheim, 德国)的1%溶液中初始温育20分钟

后,将样品在PBS中漂洗3次,达每次5分钟(3×5分钟)。然后将切片在5%正常驴血清中温育15分钟,继之以与N-末端肌养蛋白相关蛋白抗体(N-19,sc-7460,山羊多克隆IgG,Santa Cruz, CA,美国,稀度1:50)于37℃温育60分钟。在第二循环的3×5分钟PBS洗涤后,将载玻片与5%正常驴血清于室温温育15分钟。然后将制备的切片在驴抗山羊IgG-FITC(sc-2024, Santa Cruz, CA, 美国,稀度1:300)中于37℃温育30分钟。在第三PBS洗涤达3×5分钟后,首先将切片与10%正常山羊血清(Invitrogen, Scotland, 英国)温育15分钟,且然后与C-末端肌养蛋白相关蛋白抗体(MANCHO7, 小鼠单克隆IgG2a, Santa Cruz, CA, 美国,稀度1:25)于37℃温育60分钟。在PBS洗涤达3×5分钟并与10%正常山羊血清温育后,将切片在山羊抗小鼠IgG2a-Alexa Fluor® 594(A-21140, Life Technologies, 美国,稀度1:300)中于37℃温育30分钟。将切片再次在PBS中洗涤3×5分钟,并封固在Vectashield封固介质(H-1000)(Vector Laboratories, CA, 美国)或具有DAPI的封固介质(H-1500)(Vector Laboratories)中。用Leica DM6000B显微镜(Leica,德国)拍摄照片。

[0252] i. 免疫荧光染色程序:对于γ-肌膜蛋白聚糖/层粘连蛋白、MuRF-1/层粘连蛋白和MyHC-胚胎/层粘连蛋白的双重免疫荧光染色

[0253] 对于这些蛋白的染色程序遵循与以前所述相同的规程(Song, Y.等人,Effects on contralateral muscles after unilateral electrical muscle stimulation and exercise. PloS one 7, e52230,doi:10.1371/journal.pone.0052230(2012))。兔抗γ-肌膜蛋白聚糖(NBP1-59744, Novus Biologicals, Littleton, CO)和MURF1(NBP1-31207, Novus Biologicals, Littleton, CO)多克隆抗体以在具有牛血清清蛋白(BSA)的PBS中1:50的稀度使用。MyHC-胚胎单克隆抗体(F1.652)(Developmentalstudies, Hybridoma Bank, Iowa, 美国)以在PBS中1:50-1:100的稀度使用。应用1:500-1:1000稀度的层粘连蛋白鸡多克隆抗体(ab14055-50, Abcam, Cambridge, MA, 美国)连同第二山羊抗鸡IgY(TR)抗体(ab7116, Abcam, Cambridge, MA, 美国,稀度1:300)以鉴定肌肉纤维。MYH16兔多克隆抗体肽序列使用MYH16的“环2”区域的人犬序列产生。肽序列:

[0254] LLALLFKEEEAPAGS

[0255] j. TUNEL测定

[0256] 切片最初在10%缓冲的福尔马林磷酸盐(Fisher Scientific,美国)中固定20分钟。然后使用TACS2 TdT荧光素凋亡检测试剂盒(Trevigen, Gaithersburg, MD, 美国)对片段化的DNA进行原位切口末端标记,如由制造商的说明书所述的。

[0257] k. 血清肌酸激酶(CK)测定

[0258] 使用5mm的动物刺血针(Goldenrod Animal Lancet, Braintree Scientific, Inc, Braintree, MA)经由下颌下静脉的静脉穿刺术收集血清。在肝素化的血液收集管(Terumo,目录号:TMLH)中收集总共150 μL。抽血后30分钟,仔细监控小鼠以观察潜在的不良应激病征。CK水平由宾夕法尼亚大学的Matthew J. Ryan Veterinary Hospital的临床病理学实验室(Clinical Pathology Laboratory)测定。

[0259] l. EDL肌肉收缩性质的先体外后体内评价

[0260] 先体外后体内评估由宾夕法尼亚大学的肌肉生理学评估中心(Muscle Physiology Assessment Core)进行。使用装备有Dynamic MuscleControl v.5.3软件的Aurora Mouse 1200A系统,在来自2个月龄*mdx*小鼠的新鲜分离的EDL肌肉上定量了生理性

质,包括等长颤搐力、等长强直力(isometric tetanic force)和ECCs后的力下降。在安死术先体外后体内测试之前24小时,所有这些小鼠都已经历了体内握力测试。EDL肌肉于24℃维持在不断加氧的林格液(100 mM NaCl, 4.7 mM KCl, 3.4 mM $CaCl_2$, 1.2 mM KH_2PO_4, 1.2 mM $MgSO_4$, 25 mM HEPES和5.5 mM D-葡萄糖)中。所应用的颤搐刺激规程是持续时间为0.2 ms的单个刺激。对于测量强直最大力产生,以120 Hz的频率重复相同的刺激达500 ms。在两次强直性收缩之间允许五分钟,以确保肌肉恢复。调节肌肉长度以获得最大颤搐应答,并在肌腱连接的最外可见尖端之间测量所述长度,并记录为最佳长度(L0)。通过将肌肉质量除以肌肉密度系数(1.06 g/cm^3)、肌肉L0和纤维长度系数(对于EDL为0.45)的乘积,计算EDL肌肉的肌肉截面积(CSA)。通过将力相对于CSA归一化来确定比力。

[0261] 在测试EDL的等长性质后,以从重复的500 ms等长收缩开始继之以在施加最大强直性刺激的同时将肌肉拉伸L_0的10%的循环应用一系列五次偏心收缩(ECCs,每五分钟一次)。每次ECC的报道的绝对力对应于ECC等长阶段期间的峰值力。

[0262] m. 垂直活动和握力测试

[0263] 将小鼠小心地放置在旷场笼中,并测定其基线垂直活动达五分钟。然后将小鼠放回它们原来的笼中,并允许其休息三分钟。使用轴向力传感器以测量力(VernierLabPro& Vernier Dual-Range Force Sensor ±10N, Beaverton Oregon),同时使用附带的软件(Logger Lite版本1.8.1)收集数据。所有实验均由相同实验者以盲法方式进行。为了减少偏倚的机会并确保盲法实验的稳健性,我们使用了如所描述的方法(Song,Y.等人,Suite of clinically relevant functional assays to address therapeutic efficacy and disease mechanism in the dystrophicmdx mouse. J Appl Physiol 122, 593-602, doi:10.1152/japplphysiol.00776.2016(2017))。

[0264] n. 犬模型

[0265] 两组狗用于我们的实验。第一组在A&M大学的群体中饲养,并在宾夕法尼亚大学下崽。所述研究包括五只受患GRMD狗和四只年龄匹配的同窝出生的仔畜,包括一只野生型和三只携带者雌性,其充当对照组。通过升高的血清肌酸磷酸激酶(CPK)水平鉴定所有营养不良的狗,并通过PCR测定进行基因型分型。将所有幼崽随机分配至治疗组,并且在临床和组织学评估期间使研究人员保持不知情。幼崽于6-10日龄以1.0×10^{13}vg/kg和$1.0\times10^{13.5}$vg/kg的剂量经由颈外静脉途径注射AAV9 μ肌养蛋白相关蛋白。向两只幼崽注射低剂量的AAV9 μ肌养蛋白相关蛋白,两只用高剂量进行注射,且剩余的一只仅用盐水进行注射。对于最先6周每天称重每只狗,且此后每周一次。

[0266] 第二个犬的组包括约7.5周龄的两只GRMD狗。在载体施用前三天,将狗安排口服强的松龙1mg/kg方案达25天。使用相同的方法以分别1.25×10^{14}vg/kg和5.0×10^{13}vg/kg的两种不同单剂注射AAV9 μ肌养蛋白相关蛋白。将狗随机分配至剂量,且在临床和组织学评估期间使研究人员保持不知情。

[0267] 缺失无效德国短毛波音达犬(German Short Hairpointer)(GSHPMD)狗被饲养并安置在德克萨斯A&M大学(Texas A&M University)。两只七岁龄的受患GSHPMD狗Ned和Grinch分别重21 kg和24.2 kg。每只狗接受了5次到它们的胫骨前肌区室中的总等效剂量为$1.0x10^{12}$vg/kg的AAV9-μ肌养蛋白(右)和AAV9-μ肌养蛋白相关蛋白(左)的肌内注射。所有五个注射部位均进行了文身,从而允许使我们对于注射后4周的肌肉活组织检查正确定出

注射部位。在注射前、注射后2、4、6、8周收集外周血，以收集外周血单核细胞(PBMCs)。

[0268] o. 犬组织获得和贮藏

[0269] 第一组GRMD狗在载体注射后约6周经历了颅缝匠肌、股外侧肌和和肱三头肌肌肉的针吸活组织检查。样品与一组盲法代码(blinding codes)一起贮藏，以防止在分析和解释过程中的偏倚。通过弹簧加载的14-规格针套针获得活组织检查，从而显著地将动物的术后疼痛减至最低。然后将肌肉活组织检查在液氮冷却的异戊烷中骤冻，包埋在OCT介质(Sakuru，美国)中，并于-80℃贮藏。组织分析后，由与所述研究的原作者无关的个人破坏了盲法代码。载体暴露后一个月，第二组注射的狗经历相同肌肉的直视(open)肌肉活组织检查。载体施用后七周，对这些狗实施安死术，并以相同方式将所收获的组织深低温保藏。

[0270] p. 犬组织学分析

[0271] 横向切割的7 μm连续切片用于明视野显微术分析和免疫荧光(IF)染色，以检查微肌养蛋白相关蛋白表达和肌膜蛋白聚糖的拯救。肌肉切片用H&E染色用于明视野显微术，并用permount封固。对于IF染色，将切片在PBS中的5％驴血清中封闭45分钟，继之以使用1:350稀度的多克隆山羊抗肌养蛋白相关蛋白抗体(N-19, sc-7460, Santa Cruz, 美国)和1:250稀度的单克隆γ-肌膜蛋白聚糖抗体(ab55683, Abcam, 美国)于37℃温育60分钟。然后将切片在PBS中漂洗3次，并在1:1000稀度的Alexa488驴抗山羊第二抗体或Alexa540驴抗小鼠第二抗体中温育45分钟。将载玻片用PBS洗涤两次达5分钟，继之以用水的单次洗涤，并用含有DAPI的抗褪色封固介质(H-1500,Vector Labs)封固。使用Olympus B-65荧光显微镜在相同的设置下捕获图像，并经由通过全部设置限制和增益以避免IF图像中的任何不一致的相同的方式进行处理。根据如2014年1月28日所更新的TREAT_NMD规程DMD_M.1.2.001计算最小费雷特直径和变异系数。

[0272] q. 免疫印迹分析

[0273] 免疫印迹分析是通过将20-40 μg/泳道的全细胞或全肌肉裂解物加样到10％十二烷基硫酸钠-聚丙烯酰胺凝胶上进行的。将蛋白转移到聚偏二氟乙烯膜上。通过1:500稀度的抗N-末端表位的山羊多克隆抗体(N-19, sc-7460, Santa Cruz, 美国)以及1:5000稀度的第二抗体与辣根过氧化物酶缀合的驴抗山羊抗体(Sigma-Aldrich)检测微肌养蛋白相关蛋白。使用Odyssey红外成像系统(LI-COR)进行蛋白检测和定量。γ-肌膜蛋白聚糖通过小鼠单克隆抗体(VectorLabs VP-G803)和驴，抗小鼠HRP缀合的第二抗体(Santa Cruz Biotechnology)进行检测。

[0274] r. 中和抗-AAV抗体的检测

[0275] 为了评估针对AAV 9衣壳蛋白的体液免疫应答，在出生日，然后在第4周和第8周经由外周静脉收集血清。将HEK 293细胞在200 μL含10％胎牛血清的DMEM中以10^5个细胞/孔的密度接种在48-孔平板中。将细胞于37℃培养3-4小时，并允许其黏附至孔。

[0276] 将AAV9-绿色荧光蛋白(GFP)载体(1 x 10^8个颗粒)与用PBS连续稀释的小鼠血清在25 μL总体积中于4℃温育2小时。然后将混合物以200 μl的终体积添加至细胞，其含有4×10^6个AAV9颗粒，并于37℃温育24或48小时。在荧光显微镜下计数表达GFP的细胞。使用最高稀度计算中和抗体滴度，其中GFP-阳性细胞的百分比比无血清的对照少50％。

[0277] s. T-细胞对衣壳衍生的肽的反应性的评价

[0278] 通过IFN-γ ELISpot测定定量了对新的AAV衣壳抗原的外周血T细胞应答

(Mingozzi, F.等人,AAV-1-mediatedgene transfer to skeletal muscle in humans results in dosedependent activation of capsid-specific T cells. Blood 114, 2077-2086 (2009))。简而言之,外周血单核细胞(PBMC)在菲可海帕克(Ficoll hypaque)梯度上分离并用跨越VP1衣壳蛋白的合成肽(长度为20个氨基酸,重叠10个残基)培养。为了鉴定库中引出IFN-γ活性的个别肽,每种肽均存在于两个交叉映射的亚库(sub-pools)中。于37℃温育36小时后,计数IFN-γ SFU。用来自对照库的肽观察到少于10 SFU/孔(增强的GFP)。当在重复孔中SFU超过50/10^6PBMC时,应答被认为是阳性的。

[0279] t. 统计分析

[0280] 基于可用的最灵敏和最广泛使用的组织学测定之一,估计*mdx*小鼠中AAV载体注射的提供信息的组大小:来自于8周龄尸体剖检的小鼠的中央成核肌肉纤维的比例。为了使使用的动物数目的统计功效(statistical power)达到最大,我们采用了Aarts等人所述的方法(Aarts,E.等人, A solution to dependency: using multilevel analysis to accommodate nested data. Nat Neurosci, 2014. 17(4): p.491-6),以适应小鼠中多个高倍视野(HPF)之间的依赖性。使用可交换相关结构计及小鼠中分群的混合效应模型用于比较由基因型和治疗定义的三个组。来自这些模型的群内相关性(ICC)的估计显示低的值(<10%),从而提示相对高度有效的样本量,尽管每组小鼠必定小的数目(至少四只)。在所述分析中计算的其他随机效应参数包括:群之间的方差,σ^2_u;群内的方差,σ^2_e;和有效样本量,n_{eff}。

[0281] 为了表征在野生型、治疗的和未治疗的营养不良狗中的最小费雷特直径的分布,绘制了代表来自代表性HPFs的各个测量结果的点。由于狗的小数目,所以所述分析完全是描述性的。

[0282] 除中央成核的肌肉纤维的比例之外的所有分析均作为平均值±S.D.提供。使用Prism 7软件(GraphPad)处理统计分析。p值的统计显著性在图中显示:*p<0.05;**p<0.001;***p<0.0001;n.s.不显著。

[0283] 实施例3-使用AAV-介导的毫微-肌养蛋白相关蛋白和毫微-肌养蛋白递送用于DMD的全身基因治疗

[0284] A. 人毫微-肌养蛋白相关蛋白设计

[0285] 杆结构域的主要作用是力的纵向传递,且肌养蛋白的异常长度反映了蛋白的进化遗产。作为所述理解的结果,我们已经鉴定了处理其他方法的重要局限性的方法,并设计新的转基因以实现与野生型不能区别的水平的力转导和肌保护。

[0286] 我们对来自广泛取样的后生动物分类单元的肌养蛋白基因和转录物的直向同源物和旁系同源物的序列分析使得我们假设,在具有保持读框的缺失或重复的患者中的BMD由机械负荷过程中杆结构域的局部去稳定引起。所述假设得到了来自肌养蛋白和其他含有三股螺旋重复序列结构域的蛋白的晶体学数据、相邻三股螺旋重复序列中氨基酸之间进化偶联的生物信息学分析以及来自我们对机械负荷的肌肉中的小型化重组蛋白的研究的未发表的数据的支持。迄今报道的重组小-肌养蛋白(mini-dystrophins)都没有避免在内部缺失位点处最弱的环节的潜力,这是因为使用的所有设计方法都将三股螺旋重复序列模造为可互换的模块单位。尽管在动物模型中其给人深刻印像的表型改善,但即使微-肌养蛋白相关蛋白(micro-Utrophin)也冒着这种局限性的危险。为了处理这种可能性并进一步缩短

转基因,我们已经开发了三重剪接的“毫微-肌养蛋白相关蛋白”的模型(图4)。我们横过两个不相邻的三股螺旋重复序列的结构上最保守的三维平面“切割”和“结合”肌养蛋白和肌养蛋白相关蛋白的杆,以在相邻重复序列之间的保守性较低的、承重环对环界面处保持所有进化上偶联的氨基酸侧链相互作用。这在图4中使用多米诺骨牌比喻进行模造。在下一节中,这将使用来自分子模造的图解法进行更详细的描述。通过在肌养蛋白缺陷的小鼠(中间的)和大鼠(更好的DMD纤维化、变性和虚弱的病理遗传模型,但需要10x更多载体)中的序贯测试,关于三重剪接的毫微-肌养蛋白相关蛋白评价了比起微-肌养蛋白相关蛋白来的潜在改善,以鉴定治疗用转基因。

[0287] B. 用于DMD中的全身递送的AAV的可调节的生产

[0288] 迄今在我们的在营养不良小鼠和狗中的临床前研究中利用的AAV9和Anc80载体已经通过应用三重质粒转染方法在人HEK 293细胞中进行了制备。所述技术依赖于依赖贴壁细胞在丰富培养基中的生长,而输入质粒在转染前在细菌菌株中大规模繁殖,从而终产物保持潜在免疫原性的细菌DNA甲基化模式。进行了使用大量衣壳和平台的AAV载体发现和生产,包括依赖贴壁和悬浮的HEK293细胞以及杆状病毒感染的Sf9昆虫细胞。

[0289] C. 动物模型

[0290] 使用报告将来的临床试验的设计和执行的初级、次级和探索终点研究了AAVμ-对n-肌养蛋白相关蛋白全身疗法的严格、盲法测试。选择了两个初级终点:在我们发表的在mdx小鼠中综合生理测定套件中最灵敏的中的性能(混合式肢力垂直活动测试,参见来自(Song, Y.等人, Suite of Clinically Relevant Functional Assays to Address Therapeutic Efficacy and Disease Mechanism in the Dystrophic mdx Mouse.J Appl Physiol, 2016: p. jap 00776 2016)的图,和肌联蛋白N-末端25 kDa片段的定量免疫检测(Robertson, A.S.等人., Dramatic elevation in urinary amino terminal titin fragment excretion quantified by immunoassay in Duchenne muscular dystrophypatients and in dystrophin deficient rodents. Neuromuscul Disord, 2017. 27(7): p. 635-645)。参见,例如,图12A-12D和图13。后者是无损伤的、基于尿的生物测定,其具有非凡的动态范围,独特地反映了肌肉坏死的全身水平,并基于肌原纤维形成中肌联蛋白同工型的作用(实施例1)。所述计划宗旨的主要目的是在重现了DMD的主要病理特征的有成本效益的模型中定量表征表型改善的程度和持久性。mdx小鼠中的研究已系统地扩展至肌养蛋白缺陷的大鼠,这是因为所述后者模型更贴近地类似DMD中所见的标志病理肌变性和纤维化,具有容易定量的进行性肌无力和心肌病指数Robertson, A.S.等人, Dramatic elevation in urinary amino terminal titinfragment excretion quantified by immunoassay in Duchenne muscular dystrophy patients and in dystrophin deficient rodents。参见Petrof,B.J.等人, Dystrophin protects the sarcolemma from stresses developed during muscle contraction. Proc Natl Acad Sci, 1993. 90:p. 3710-14;Neuromuscul Disord, 2017. 27(7): p. 635-645;Larcher, T.等人, Characterization of dystrophin deficient rats: a newmodel for Duchenne muscular dystrophy. PLoS One, 2014. 9(10): p. e110371;Nakamura, K.等人, Generation of muscular dystrophy model rats with a CRISPR/Cas system. Sci Rep,2014. 4: p. 5635;Stedman, H.H.等人, The mdx mouse diaphragm reproduces

the degenerative changes of Duchenne muscular dystrophy. Nature, 1991. 352 (6335): p. 536-9;Shrager, J.B.等人, Themdx mouse and mdx diaphragm implications for the pathogenesis of Duchenne Muscular Dystrophy., 在 NeuromuscularDevelopment and Disease, A.M. Kelly和H.M. Blau, 编辑. 1992, Raven Press, Ltd.: New York. p.317-328中;Krupnick, A.S.等人, Inspiratory loading does not accelerate dystrophy in mdx mouse diaphragm: implications for regenerative therapy. J Appl Physiol,2003. 94(2): p. 411-9;和Song, Y.等人, Suite of clinically relevant functional assays to address therapeutic efficacy and disease mechanism in the dystrophic mdx mouse.J Appl Physiol, 2017. 122(3): p. 593-602。几种犬疾病模型(Cooper, B.J.等人, The homologue of the Duchenne locus isdefective in X-linked muscular dystrophy of dogs. Nature, 1988. 334(6178): p. 154-6;Smith, B.F.等人, Molecular basis of canine muscle typephosphofructokinase deficiency. J Biol Chem, 1996. 271(33): p. 20070-4;Bridges, C.R.等人, Global cardiac-specific transgene expression using cardiopulmonary bypass withcardiac isolation. Ann Thorac Surg, 2002. 73(6): p. 1939-46;Arruda, V.R.等人, Regional intravascular delivery of AAV-2-F.IX to skeletal muscle achieveslong-term correction of hemophilia B in a large animal model. Blood, 2004. Epub领先于印刷;Arruda, V.R.等人, Peripheral transvenular delivery of adeno-associated viral vectors to skeletalmuscle as a novel therapy for hemophilia B. Blood, 2010. 115(23): p. 4678-88;Mead, A.F.等人, Diaphragm remodeling and compensatory respiratory mechanics in a canine modelof Duchenne muscular dystrophy. J Appl Physiol (1985), 2014. 116(7): p. 807-15;和Su, L.T.等人, Uniform scale-independent gene transfer to striated muscle after transvenularextravasation of vector. Circulation, 2005. 112(12): p. 1780-8)可以使用,以及LGMD的仓鼠模型(例如图14A-图14G中的组织学测定,来自Greelish等人, Nature Medicine,(Greelish,J.P.等人, Stable restoration of the sarcoglycan complex in dystrophic muscle perfused with histamine and a recombinantadeno-associated viral vector. Nat Med, 1999. 5(4): p. 439-43),其形成了在人肌营养不良中使用AAV载体的基因治疗的首次I期临床试验背景的基本部分:Stedman等人, Human Gene Therapy. Stedman, H.等人, Phase Iclinical trial utilizing gene therapy for limb girdle muscular dystrophy: alpha- , beta-, gamma-, or delta-sarcoglycan gene delivered with intramuscularinstillations of adeno-associated vectors. Hum Gene Ther, 2000. 11(5): p. 777-90)。

[0291] D. 产品简况

[0292] AAVμ肌养蛋白相关蛋白或AAVn肌养蛋白相关蛋白的靶产品简况。

[0293]

产品靶	最低可接受的结果	理想结果
主要产品适应症	由任何肌养蛋白基因突变引起的杜兴和贝克尔肌营养不良中肢肌肉功能障碍的症状和病征	由任何肌养蛋白基因突变引起的杜兴和贝克尔肌营养不良中运动、呼吸和心肌肉功能障碍的症状和病征
患者群体	具有进行性DMD的症状和病征的不卧床或不卧床后（post-ambulatory）患者	具有DMD/BMD的症状发生前的婴儿、有症状的男孩和年轻成人
治疗持续时间	单剂	单剂
递送模式	IV	IV
剂量形式	重组AAV载体的冷冻悬浮液，用于解冻后的单次使用	重组AAV载体的冷冻悬浮液，用于解冻后的单次使用
方案	输注-用于全身输注的肘前静脉中的长导管；用于肢的止血带远侧的头静脉和/或隐静脉注射的肢中功能丧失进展速度的多年变慢	输注-用于全身输注的肘前静脉中的长导管；用于肢的止血带远侧的头静脉和/或隐静脉基于证实的早期症状发生前诊断，与DMD/BMD有关的运动、呼吸和心脏疾病症状的*完全预防*
功效		
危险/副作用	没有来自编码的重组蛋白的长期免疫毒性。可能需要短暂免疫抑制以将对载体衣壳的免疫应答的危险减至最低。抗载体的蛋白衣壳的抗体的发展可能阻止再施用。可应用的最大剂量可能受涉及先天免疫的毒性和/或肝脏毒性的危险限制。	没有来自编码的重组蛋白的长期免疫毒性。可能需要短暂免疫抑制以将对载体衣壳的免疫应答的危险减至最低。抗载体的蛋白衣壳的抗体的发展可能阻止再施用。
治疗模式	编码人工剪接的肌养蛋白相关蛋白同工型的促肌腺伴随病毒载体	编码人工剪接的肌养蛋白相关蛋白同工型的促肌腺伴随病毒载体

[0294] 上面的表格描绘了AAVμ肌养蛋白相关蛋白或AAVn肌养蛋白相关蛋白的靶产品简况。我们目前对DMD中横纹肌细胞的进行性丧失和纤维脂肪置换(fibro-fatty replacement)的理解提示,在较年长主体中疾病的显著逆转可能是有限的,但是在重组μ-或n-肌养蛋白相关蛋白表达开始后的任何阶段都可能预防进一步的肌细胞丧失。主要针对肌养蛋白缺陷的小鼠和大鼠进行了实验,以报告对于在婴儿期中治疗的DMD患者中理想参

数的期望值。在婴儿期中用充足的AAV μ-或n-肌养蛋白相关蛋白进行转导以使肌膜蛋白聚糖表达在到骨骼成熟的整个生长过程中正常化可允许相对正常的肌肉生长以及因而强度中正常成熟的增加。至少四个因素可能限制较年长患者中的治疗益处:1)治疗前不可逆的肌细胞丧失的程度,2)肌肉的纤维脂肪置换损害载体递送和肌细胞转导的程度,3)对载体的内皮透性中的成熟减少,可能需要从较年长患者中止血带远侧的血管腔强制外渗,以及4)随着逐渐增长的患者年龄,自然暴露于AAV病毒的预期的增加,从而增加了具有对多种AAV血清型的高滴度抗体和对保守的AAV衣壳衍生肽的记忆T细胞两者的比例。剂量限制性毒性可能涉及先天免疫系统和/或肝。

[0295] E. 肌养蛋白缺乏症的μ-或n-肌养蛋白相关蛋白取代的实验和理论基础

[0296] 肌养蛋白相关蛋白最初是根据其与肌养蛋白的编码序列同源性发现的(Love, D.R.等人, An autosomal transcript in skeletal muscle with homology to dystrophin. Nature, 1989. 339(6219): p. 55-8)。我们最近根据来自大量分类单元的可公开获得的全基因组序列重构了肌养蛋白和肌养蛋白相关蛋白的进化史。两个观察结果是相关的:1)在横纹肌出现之前,两种蛋白的杆样结构域的"供体"基因远在其通过部分基因重复与Dp71-样结构域连接以前就具有至少21个串联的血影蛋白样重复序列结构域,2)肌养蛋白相关蛋白和肌养蛋白的单独基因在头索动物沿着通向无颌类和有颌类脊椎动物的共同祖先的谱系趋异之后以及少突胶质细胞的进化出现之前被固定(Putnam,N.H.等人, The amphioxus genome and the evolution of the chordate karyotype. Nature, 2008. 453(7198): p. 1064-71;和Smith, J.J.等人,Sequencing of the sea lamprey (Petromyzon marinus) genome provides insights into vertebrate evolution. Nat Genet, 2013. 45(4): p. 415-21, 421e1-2)。两种蛋白都保留了对于细胞骨架肌动蛋白和肌养蛋白(/肌养蛋白相关蛋白)结合蛋白复合体(D/UAPC)的跨膜糖蛋白的结合界面。充分确定的是,肌养蛋白相关蛋白的全长重组衍生物甚至可以逆转小鼠中严重的肌营养不良表型(Tinsley,J.等人, Expression of full-length utrophin prevents muscular dystrophy in mdx mice. Nat Med, 1998. 4(12): p. 1441-4;Gilbert, R.等人, Adenovirus-mediated utrophin gene transfer mitigates the dystrophic phenotype of mdx mouse muscles. Hum Gene Ther, 1999. 10(8): p. 1299-310;和Odom, G.L.等人,Microutrophin delivery through rAAV6 increases lifespan and improves muscle function in dystrophic dystrophin/utrophin-deficient mice. Mol Ther, 2008.16(9): p. 1539-45)。我们知识中未解决的空白是导致了祖先蛋白的三股螺旋重复最初串联重复的选择性压力的特性,以及肌养蛋白当前的一种或多种生理作用是否"需要"天然蛋白Dp427的估计的6nm/重复序列×24个重复序列或144nm长度。一些具有小内部缺失的BMD患者的严重临床表型提示,Dp427的长度对于其作为减震体的作用是必需的,但是重复-BMD向这一解释提出了疑问,这是因为这些患者具有比野生型更长的肌养蛋白。从我们对肌养蛋白的遥远进化史的重构引出了另一种解释(实施例1)。

[0297] 我们问了一个看似简单的问题"肌养蛋白和肌节哪个先出现?"目的是确定来自现存物种的基因组和推断的蛋白质组的可获得的证据是否与这样的模型相一致,在所述模型中,肌养蛋白杆的长度在从肌节出现到对肌养蛋白缺陷的哺乳动物中的急性损伤最易感的特化快缩肌肌肉的最近进化的时间段期间增加(Webster, C.等人, Fast muscle fibers

are preferentiallyaffected in Duchenne muscular dystrophy. Cell, 1988. 52(4): p. 503-13;和Petrof, B.J.等人, Adaptations in myosin heavy chain expression and contractile function indystrophic mouse diaphragm. Am. J. Physiol., 1993. 265: p. C834-C841)。如图15中所示的,证据支持肌养蛋白在时间上早于肌节的重构。详细的分析表明,肌养蛋白在现存脊椎动物和刺胞动物(水母)物种中的长度相同,具有高度保守的基因结构,如图6A至图6C中所示的。肌养蛋白的长度很可能在肌节出现之前很久就达到,甚至在扁盘动物(Placozoa)(由最简单的自由生活的动物物种所代表的门)丝盘虫(Trichoplax adhaerens)趋异之前。所述物种使用纤毛动力蛋白(不是肌球蛋白)作为其主要的运动动力源,且其机体发育布局以仅四种细胞类型为特色,没有一种显示可鉴定的肌节(Srivastava, M.等人, The Trichoplax genome and the natureof placozoans. Nature, 2008. 454(7207): p. 955-60)。基因结构的创新分析提供了令人信服的证据证明占蛋白长度的80%的肌养蛋白的杆状结构域看来似乎是作为一个整体从较大的祖先蛋白(不是血影蛋白,注意图6B中的性质不同的模式)补选的,具有在交联肌动蛋白纤丝和微管中的特殊作用(MACF)。所述发现的基本意义是MACF-样蛋白而不是肌养蛋白本身是驱动两种蛋白的共同祖先杆结构域最初加长的选择性压力的对象。在纵向力传递的潜力方面,MACF的晶体结构与血影蛋白的不同,从而为在大量分类单元中肌养蛋白和肌养蛋白相关蛋白长度的保持提供了解释。如图16A和16B中所示的,血影蛋白和dystroplakins具有在重叠程度和稳定氨基酸侧链相互作用的数目方面不同折叠的三股螺旋。在图16A和图16B中,我们使用衍生自(图16A)人β2-血影蛋白(3EDV)和(图16B)人网蛋白(5J1G)的模板对人肌养蛋白相关蛋白的相邻三股螺旋重复序列进行了模造。我们的模型预测,力在体内是穿过宽结构域间界面纵向传递的,如在不破坏肌膜的情况下将机械能从细胞内传递到细胞外基质所需要的。在相邻的和可能重叠的三股螺旋重复序列之间的侧链键合中的任何较大破坏都将因此在力传递过程中使肌养蛋白和肌养蛋白相关蛋白去稳定。

[0298] 所述模型与杆结构域最重要的特征是强度而不是长度的主张完全一致,而蛋白的异常大小可归因于其作为通过长得多的MACF同源物(homolog)的部分基因重复的衍生物的历史遗产。因此,专门设计以最优化在所有重复序列间界面上的杆结构域的结构完整性的短重组蛋白应完全补充全长蛋白例如Dp427的机械功能。在概念上接近这一点的最简单的方法是避免直接将不相容的三股螺旋重复序列并列的内部重排。问题是对于肌养蛋白的24个三股螺旋重复序列的仅仅一个存在关于结构的晶体学信息,并且除了螺旋A和C中心中保守的色氨酸残基之外重复序列间的一级结构同源性低。我们选择最初集中于相对于全长肌养蛋白相关蛋白具有一个内部和一个C-末端缺失的重组蛋白,并将其命名为微-或μ-肌养蛋白相关蛋白。如图17中所示的,所述构建体将非结构的、富含脯氨酸的螺旋间“铰链-2”结构域相对于全长肌养蛋白相关蛋白的最后一个三股螺旋重复序列(第22个)并列。在最合适的情况中,“铰链”可以充当在没有与三股螺旋重复序列22的序列精确匹配的情况下具有传递纵向力的能力的螺旋间间隔区(spacer)。这是我们对于μ-肌养蛋白相关蛋白的集中评价的基本原理,如下文和实施例2中所详细描述的。

[0299] 用于μ-肌养蛋白相关蛋白转基因最优化的系统方法确定了肌节肌球蛋白重链的密码子偏倚的使用在体外将可免疫检测的蛋白的表达比起野生型来增加到30-倍(泳道5对泳道8,图18)。实际上,表达水平如此高,以致于其与共转染的e-gfp对照的表达有效竞争。

为了分析包装到AAV9中后最优化的μU构建体的治疗功效，我们在DMD的mdx小鼠模型中进行了一系列随机的、盲法研究，如实施例2中所述的。

[0300] 根据对最小T细胞肽表位的MHC结合的计算机芯片上的分析，与肌养蛋白缺陷的宿主中所有假设的μ肌养蛋白相比，μ肌养蛋白相关蛋白具有小于1/100的数目的潜在免疫优势外源肽。这将DMD中的治疗后自身免疫的危险以及对长期免疫抑制的潜在需要减至最低。我们对非免疫抑制的GRMD狗中全身μ肌养蛋白相关蛋白表达的盲法研究通过证明由干扰素γ ELISpot测定完全不存在外周T细胞反应性而提供了进一步的再保证。这些研究必定以mdx小鼠中使用的最大剂量的1/10(相对于成体体重归一化)的剂量进行，并且只能在将来的日期进行扩展，以利用可调节的生产技术评估大的动物DMD模型中的最大耐受剂量。

[0301] F. 肌养蛋白的基本原理和结构的力学生物学-毫微-肌养蛋白相关蛋白

[0302] 我们的发现对整个领域都具有重要意义，这是因为其与所有正在开发用于转化入临床研究的AAV-大小的肌养蛋白候选物同等有关。

[0303] 我们开始了一系列盲法实验以处理AAV-注射的mdx小鼠中μ-肌养蛋白相关蛋白的稳定性。这些定量研究已经关于μ-肌养蛋白相关蛋白表达的水平在中间时间点提供了非常有利的数据。然而，在进一步的分析中，我们在用对肌养蛋白相关蛋白的N-末端特异的抗体染色后，在蛋白印迹上注意到不曾预料到的“额外条带”。如图21中所总结的，分配给所述新奇条带的分子量与μ-肌养蛋白相关蛋白的79 kd N-末端部分的分子量完全匹配。换句话说，所述发现提示135 kd的μ-肌养蛋白相关蛋白在紧接接近于侧接对应于全长-肌养蛋白相关蛋白的编码序列缺失的“铰链2”的部分和血影蛋白样重复序列22之间的连接处被破坏，从而释放了79 kd N-末端“亚结构域”作为片段。在我们的研究中，看来似乎在79 kd条带的出现和强度与由肌肉进行的力转导的最大最近水平之间存在相关性。我们接下来通过液相层析和串联质谱法的组合使额外的凝胶的79 kd区域经历了蛋白质组学分析。如图22中所示的，这证实了我们的假设。所述发现迫使我们重新考虑我们关于血影蛋白重复序列的互换性的假设，特别是当在直接一级结构接近于铰链2之处放置时的4和22。在再访问我们对肌养蛋白、肌养蛋白相关蛋白和肌联蛋白的生物信息学分析的发现时，显然肌联蛋白的聚IgG结构域显示出在进化时间内互换性的强有力的证据，但是在相同的物种比较中，我们未看到关于其他两种蛋白的互换性的证据(实施例1)。这为选择的贝克尔MD病例中的基因型-表型相关性提供了令人信服的解释，尤其是与基因重复有关的严重病例，其中据推断存在比野生型更长的肌养蛋白，其具有穿过杆的两个祖先不相邻的部分的新连接轴向力传递的最弱的环节。

[0304] 来自MACF/斑蛋白家族的一种晶体结构可能是有益的，这是因为肌养蛋白和肌养蛋白相关蛋白的结构数据限于第一个三股螺旋重复序列(即杆的最N-末端)，且铰链结构域被预测为“非结构的”。在确定所述结构之前，存在三股螺旋重复序列之间的SH3结构域也充当铰链从而使得出现杆的两个强有力部分之间的不稳定界面图像的推测。相反，结构揭示SH3结构域与来自两侧的相邻三股螺旋的相容氨基酸侧链形成多个高亲和力接触，一种具有传递纵向力和抵抗伸展的潜力的构型(图23显示了斑蛋白结构域的SR4&5的3PEo结构，以空间充满显示SH3:SR4/5结合界面)。对来自刺胞动物物种(水母)的几种肌养蛋白/肌养蛋白相关蛋白直向同源物的一级结构的分析表明存在插入的但HMM可识别的结构域，但没有非结构的“铰链”，从而进一步支持了我们关于肌养蛋白杆结构域从MACF-样蛋白进化起源

以及穿过结构域间界面的宽表面积的长轴力传递的幅度两者的假设。

[0305] 考虑到这些结构和功能约束,我们再访问了对于Dp427即全长肌养蛋白的AAV-相容的的小型化取代物的设计。我们设计了毫微-肌养蛋白相关蛋白,它利用独特的机会以在对结构域间界面的破坏最少的情况下在内部对杆结构域进行重排。这是通过穿过保守的轴横截面合并成对的不同的三股螺旋在计算机芯片上实现的,在使三股螺旋的核心稳定的相互作用的、高度保守的色氨酸残基的位点处举例说明的一种(图3)。我们已经合成了所述构建体的密码子优化的cDNAs,产生并大量制备了用于293细胞的三重转染的ITR-侧接的转录盒载体。进一步地,我们进行了盲法实验,其中*mdx*小鼠被随机化以接受编码"微-"或"毫微-"肌养蛋白相关蛋白的AAVs。结果表明,在79kd N-末端亚片段的末端精确地切割了微-肌养蛋白相关蛋白,而在相似的生理负荷下没有可检测的毫微-肌养蛋白相关蛋白的切割(图28)。两种蛋白都正确定位于*mdx*小鼠肌肉中的肌膜。来自所述实验的发现表明,毫微-肌养蛋白相关蛋白与微-肌养蛋白相关蛋白相比的优越强度更紧密地接近肌养蛋白同工型Dp427的力学生物学。利用以上在这些实施例中使用的方法进行额外的研究,以评价各种疾病模型中的表型改善。同时,在减少的载体基因组大小的基础上确立了转基因包装入AAV的效率的任何改善的程度。

[0306] G. 用于DMD治疗的重构的祖先AAV衣壳:Anc80、81、82

[0307] 基于天然存在的衣壳血清型9的AAV载体实现了到狗和非人灵长类动物的横纹肌的惊人的整体生物分布。尽管对于其他血清型很好地定义了与选择的膜受体结合所涉及的衣壳残基,但对此的结构基础很少理解。AAV8是在狗中进行有效心基因转移的较好选择,从而反映了物种差异,这是因为AAV9在灵长类动物中提供稳健的心肌转导。AAV8和9两者均与成年人群体的显著比例中的中和抗体有关。在努力在维持到横纹肌的整体生物分布的同时回避这种局限性的过程中,Zinn等人(Zinn, E.等人, In SilicoReconstruction of the Viral Evolutionary Lineage Yields a Potent Gene Therapy Vector. Cell Rep, 2015. 12(6): p. 1056-68)使用在计算机芯片上的祖先序列重构和体外基因合成的组合以制备具有"新"载体衣壳的AAVs(即很可能重现了天然存在但很久以前已灭绝的衣壳变体)。在这些中,标记为Anc80、81和82的那些代表了最直接的机会以在从在过去暴露于天然存在的现存AAVs后具有高滴度中和抗体的潜力的患者群体扩大合格库的同时重复或扩展AAV8和9的有利生物分布,如详细描述的(Zinn等人,如上文所引用的)。

[0308] 我们最近将我们的μ肌养蛋白相关蛋白基因组包装在Anc80衣壳中,并使用所得到的载体以评价mdx小鼠中相关肌肉的转导。这些研究表明,在这个意义上,Anc80在心肌和骨骼肌的强转导的情况下实现了与AAV9类似的整体生物分布(图24A至24C)。

[0309] 基于治疗用转基因的表达水平,来自随机化、盲法实验的数据证实Anc80和AAV9具有在营养不良的肌肉中类似的体内感染性的能力。

[0310] 对于在以2.5×10^{12}vg/小鼠的相等剂量全身施用这些载体后mdx小鼠群组中μ肌养蛋白相关蛋白的生物分布,进行了进行设计以定性比较AAV9和Anc80的实验。对于每种载体,来自两只小鼠的多种肌肉的代表性蛋白印迹显示在图25中,从而证明了用两种载体对横纹肌的广泛且有效的转导。

[0311] 实施例4-五重复序列突变体肌养蛋白相关蛋白和肌养蛋白

[0312] 我们已经设计了相对于上述最初的"毫微"突变体含有额外的三股螺旋的肌养蛋

白相关蛋白和肌养蛋白重组蛋白。这些“五重复序列”突变体具有在全长蛋白中存在的N-末端钙调理蛋白同源结构域和C-末端“WW-EF-ZZ”结构域之间的5个血影蛋白样三股螺旋重复序列。这些变体由于存在三重剪接突变而还具有改善的稳定性，并且进一地能够被包装到AAV载体中。重要的是，这些突变体的开发举例说明了设计四重复序列的毫微-肌养蛋白和毫微-肌养蛋白相关蛋白所依赖的原理，包括本文所述的那些，可以扩展到具有五个螺旋重复序列的变体。SEQ ID NO:22中提供了对应于五重复序列肌养蛋白的氨基酸序列，其中通过连接全长肌养蛋白蛋白的螺旋重复序列1和20形成了剪接突变(图29A和图29B)。SEQ ID NO:21中提供了对应于五重复序列肌养蛋白相关蛋白的氨基酸序列，其中通过连接全长肌养蛋白相关蛋白蛋白的螺旋重复序列1和18形成了剪接突变(图30A和图30B)。

[0313] (序列表自由文字)

[0314] 对于在数字标识符<223>下含有自由文字的序列提供了以下信息。

[0315]

SEQ ID NO: (含有自由文字)	<223>下的自由文字
1	<223> 毫微-肌养蛋白的氨基酸序列 <220> <221> MISC_FEATURE <222> (1)..(463) <223> 与全长人肌养蛋白的 N-末端区域相同的氨基酸序列。 <220> <221> MISC_FEATURE <222> (463)..(463) <223> 位于所述超家族中蛋白的隐蔽马尔科夫模型（HMM）和所有晶体结构的核心处的“A”螺旋中的色氨酸残基。

[0316]

SEQ ID NO: (含有自由文字)	<223>下的自由文字
	<220> <221> MISC_FEATURE <222> (464)..(505) <223> 与全长人肌养蛋白的 C-末端区域相同的氨基酸序列。 <220> <221> MISC_FEATURE <222> (505)..(506) <223> 侧接如图 2F 中所描绘的假设横切面平面的“B”螺旋的超家族 HMM 内的位置 <220> <221> MISC_FEATURE <222> (506)..(550) <223> 与全长人肌养蛋白的 N-末端区域相同的氨基酸序列。 <220> <221> MISC_FEATURE <222> (550)..(550) <223> 位于所述超家族中蛋白的隐蔽马尔科夫模型（HMM）和所有晶体结构的核心处的“C”螺旋中的色氨酸残基。 <220> <221> MISC_FEATURE <222> (551)..(1131) <223> 与全长人肌养蛋白的 C-末端区域相同的氨基酸序列。
2	<223> 编码毫微-肌养蛋白的核酸序列
3	<223> 毫微-肌养蛋白相关蛋白 1 的氨基酸序列 <220> <221> MISC_FEATURE <222> (1)..(433) <223> 与全长人肌养蛋白相关蛋白的 N-末端区域相同的氨基酸序列 <220> <221> MISC_FEATURE <222> (433)..(433) <223> 位于所述超家族中蛋白的隐蔽马尔科夫模型（HMM）和所

[0317]

SEQ ID NO: (含有自由文字)	<223>下的自由文字
	有晶体结构的核心处的“A”螺旋中的色氨酸残基。 <220> <221> MISC_FEATURE <222> (434)..(477) <223> 与全长人肌养蛋白相关蛋白的 C-末端区域相同的氨基酸序列 <220> <221> MISC_FEATURE <222> (477)..(478) <223> 侧接如图 2F 中所描绘的假设横切面平面的“B”螺旋的超家族 HMM 内的位置 <220> <221> MISC_FEATURE <222> (478)..(514) <223> 与全长人肌养蛋白相关蛋白的 N-末端区域相同的氨基酸序列 <220> <221> MISC_FEATURE <222> (514)..(514) <223> 位于所述超家族中蛋白的隐蔽马尔科夫模型（HMM）和所有晶体结构的核心处的“C”螺旋中的色氨酸残基 <220> <221> MISC_FEATURE <222> (515)..(1098) <223> 与全长人肌养蛋白相关蛋白的 C-末端区域相同的氨基酸序

[0318]

SEQ ID NO: (含有自由文字)	<223>下的自由文字
	列
4	<223> 编码毫微-肌养蛋白相关蛋白-1 蛋白的核酸序列
5	<223> 毫微-肌养蛋白相关蛋白-2 的氨基酸序列
6	<223> 编码毫微-肌养蛋白相关蛋白-2 蛋白的核酸序列
7	<223> 毫微-肌养蛋白相关蛋白-3 的氨基酸序列
8	<223> 编码毫微-肌养蛋白相关蛋白-3 蛋白的核酸序列
9	<223> 具有 GenBank 检索: AAS99264 的氨基酸序列
10	<223> AAV9 衣壳 vp1 的编码序列
11	<223> hu31 vp1 的编码序列
12	<223> hu32 vp1 的编码序列
13	<223> 毫微-肌养蛋白的氨基酸序列
14	<223> 毫微-肌养蛋白的氨基酸序列
15	<223> 毫微-肌养蛋白的氨基酸序列
16	<223> 毫微-肌养蛋白的氨基酸序列
17	<223> 毫微-肌养蛋白的氨基酸序列
18	<223> 毫微-肌养蛋白的氨基酸序列
19	<223> 编码毫微-肌养蛋白相关蛋白的人工改造的序列
20	<223> 合成的构建体
21	<223> 人肌养蛋白相关蛋白突变体 <220> <221> MISC_FEATURE <222> (1)..(324) <223> 与全长人肌养蛋白相关蛋白的 N-末端区域相同的氨基酸序列 <220>

[0319]

SEQ ID NO: (含有自由文字)	<223>下的自由文字
	<221> MISC_FEATURE <222> (311)..(417) <223> 杂合三股螺旋 <220> <221> MISC_FEATURE <222> (325)..(362) <223> 与全长人肌养蛋白相关蛋白的 C-末端区域相同的氨基酸序列 <220> <221> MISC_FEATURE <222> (363)..(401) <223> 与全长人肌养蛋白相关蛋白的 N-末端区域相同的氨基酸序列 <220> <221> MISC_FEATURE <222> (402)..(1208) <223> 与全长人肌养蛋白相关蛋白的 C-末端区域相同的氨基酸序列
22	<223> 突变人肌养蛋白蛋白 <220> <221> MISC_FEATURE <222> (1)..(354) <223> 与全长人肌养蛋白的 N-末端区域相同的氨基酸序列

[0320]

SEQ ID NO: (含有自由文字)	<223>下的自由文字
	<220> <221> MISC_FEATURE <222> (340)..(447) <223> 杂合三股螺旋 <220> <221> MISC_FEATURE <222> (355)..(392) <223> 与全长人肌养蛋白的 C-末端区域相同的氨基酸序列 <220> <221> MISC_FEATURE <222> (393)..(431) <223> 与全长人肌养蛋白的 N-末端区域相同的氨基酸序列 <220> <221> MISC_FEATURE <222> (432)..(1237) <223> 与全长人肌养蛋白的 C-末端区域相同的氨基酸序列

[0321] 本申请中引用的所有出版物、专利、专利申请和本文参考的序列表，以及于2018年4月16日提交的美国临时专利申请No. 62/658,464特此整体引入本文作为参考，好像每个个别的出版物或专利申请均被明确地并单独地指出引入作为参考。许多修改和变化形式都包括在上述确定的说明书的范围内，并且预期对本领域技术人员是显而易见的。对组合物和方法的这种修改和改变，例如不同的编码序列的选择或载体或免疫调谐剂的选择或剂量，被认为在所附权利要求的范围内。

序列表

<110> The Trustees of the University of Pennsylvania

<120> 用于治疗杜兴肌营养不良的组合物和方法

<130> UPN-18-8438PCT

<150> US 62/658,464
<151> 2018-04-16

<160> 22

<170> PatentIn version 3.5

<210> 1
<211> 1131
<212> PRT
<213> 人工序列

<220>
<223> 毫微-肌养蛋白的氨基酸序列

<220>
<221> MISC_FEATURE
<222> (1)..(463)
<223> 与全长人肌养蛋白的 N-末端区域相同的氨基酸序列。

[0001]

<220>
<221> MISC_FEATURE
<222> (463)..(463)
<223> 位于所述超家族中蛋白的隐蔽马尔科夫模型（HMM）和所有晶体结构的核心处的“A”螺旋中的色氨酸残基。

<220>
<221> MISC_FEATURE
<222> (464)..(505)
<223> 与全长人肌养蛋白的 C-末端区域相同的氨基酸序列。

<220>
<221> MISC_FEATURE
<222> (505)..(506)
<223> 侧接如图 2F 中所描绘的假设横切面平面的“B”螺旋的超家族 HMM 内的位置

<220>
<221> MISC_FEATURE
<222> (506)..(550)
<223> 与全长人肌养蛋白的 N-末端区域相同的氨基酸序列。

<220>
<221> MISC_FEATURE
<222> (550)..(550)
<223> 位于所述超家族中蛋白的隐蔽马尔科夫模型（HMM）和所有晶体结构的核心处的“C”螺旋中的色氨酸残基。

<220>
<221> MISC_FEATURE

<222> (551)..(1131)
<223> 与全长人肌养蛋白的C-末端区域相同的氨基酸序列。

<400> 1

Met Leu Trp Trp Glu Glu Val Glu Asp Cys Tyr Glu Arg Glu Asp Val
1 5 10 15

Gln Lys Lys Thr Phe Thr Lys Trp Val Asn Ala Gln Phe Ser Lys Phe
20 25 30

Gly Lys Gln His Ile Glu Asn Leu Phe Ser Asp Leu Gln Asp Gly Arg
35 40 45

Arg Leu Leu Asp Leu Leu Glu Gly Leu Thr Gly Gln Lys Leu Pro Lys
50 55 60

Glu Lys Gly Ser Thr Arg Val His Ala Leu Asn Asn Val Asn Lys Ala
65 70 75 80

Leu Arg Val Leu Gln Asn Asn Asn Val Asp Leu Val Asn Ile Gly Ser
85 90 95

[0002]

Thr Asp Ile Val Asp Gly Asn His Lys Leu Thr Leu Gly Leu Ile Trp
100 105 110

Asn Ile Ile Leu His Trp Gln Val Lys Asn Val Met Lys Asn Ile Met
115 120 125

Ala Gly Leu Gln Gln Thr Asn Ser Glu Lys Ile Leu Leu Ser Trp Val
130 135 140

Arg Gln Ser Thr Arg Asn Tyr Pro Gln Val Asn Val Ile Asn Phe Thr
145 150 155 160

Thr Ser Trp Ser Asp Gly Leu Ala Leu Asn Ala Leu Ile His Ser His
165 170 175

Arg Pro Asp Leu Phe Asp Trp Asn Ser Val Val Cys Gln Gln Ser Ala
180 185 190

Thr Gln Arg Leu Glu His Ala Phe Asn Ile Ala Arg Tyr Gln Leu Gly
195 200 205

Ile Glu Lys Leu Leu Asp Pro Glu Asp Val Asp Thr Thr Tyr Pro Asp

210 215 220

Lys Lys Ser Ile Leu Met Tyr Ile Thr Ser Leu Phe Gln Val Leu Pro
225 230 235 240

Gln Gln Val Ser Ile Glu Ala Ile Gln Glu Val Glu Met Leu Pro Arg
245 250 255

Pro Pro Lys Val Thr Lys Glu Glu His Phe Gln Leu His His Gln Met
260 265 270

His Tyr Ser Gln Gln Ile Thr Val Ser Leu Ala Gln Gly Tyr Glu Arg
275 280 285

Thr Ser Ser Pro Lys Pro Arg Phe Lys Ser Tyr Ala Tyr Thr Gln Ala
290 295 300

Ala Tyr Val Thr Thr Ser Asp Pro Thr Arg Ser Pro Phe Pro Ser Gln
305 310 315 320

His Leu Glu Ala Pro Glu Asp Lys Ser Phe Gly Ser Ser Leu Met Glu
[0003] 325 330 335

Ser Glu Val Asn Leu Asp Arg Tyr Gln Thr Ala Leu Glu Glu Val Leu
340 345 350

Ser Trp Leu Leu Ser Ala Glu Asp Thr Leu Gln Ala Gln Gly Glu Ile
355 360 365

Ser Asn Asp Val Glu Val Val Lys Asp Gln Phe His Thr His Glu Gly
370 375 380

Tyr Met Met Asp Leu Thr Ala His Gln Gly Arg Val Gly Asn Ile Leu
385 390 395 400

Gln Leu Gly Ser Lys Leu Ile Gly Thr Gly Lys Leu Ser Glu Asp Glu
405 410 415

Glu Thr Glu Val Gln Glu Gln Met Asn Leu Leu Asn Ser Arg Trp Glu
420 425 430

Cys Leu Arg Val Ala Ser Met Glu Lys Gln Ser Asn Leu His Arg Val
435 440 445

Leu Met Asp Leu Gln Asn Gln Lys Leu Lys Glu Leu Asn Asp Trp Leu
450 455 460

Thr Glu Ala Glu Thr Thr Ala Asn Val Leu Gln Asp Ala Thr Arg Lys
465 470 475 480

Glu Arg Leu Leu Glu Asp Ser Lys Gly Val Lys Glu Leu Met Lys Gln
485 490 495

Trp Gln Asp Leu Gln Gly Glu Ile Glu Gln Glu Gln Val Arg Val Asn
500 505 510

Ser Leu Thr His Met Val Val Val Val Asp Glu Ser Ser Gly Asp His
515 520 525

Ala Thr Ala Ala Leu Glu Glu Gln Leu Lys Val Leu Gly Asp Arg Trp
530 535 540

Ala Asn Ile Cys Arg Trp Ser Glu Leu Arg Lys Lys Ser Leu Asn Ile
545 550 555 560

[0004] Arg Ser His Leu Glu Ala Ser Ser Asp Gln Trp Lys Arg Leu His Leu
565 570 575

Ser Leu Gln Glu Leu Leu Val Trp Leu Gln Leu Lys Asp Asp Glu Leu
580 585 590

Ser Arg Gln Ala Pro Ile Gly Gly Asp Phe Pro Ala Val Gln Lys Gln
595 600 605

Asn Asp Val His Arg Ala Phe Lys Arg Glu Leu Lys Thr Lys Glu Pro
610 615 620

Val Ile Met Ser Thr Leu Glu Thr Val Arg Ile Phe Leu Thr Glu Gln
625 630 635 640

Pro Leu Glu Gly Leu Glu Lys Leu Tyr Gln Glu Pro Arg Glu Leu Pro
645 650 655

Pro Glu Glu Arg Ala Gln Asn Val Thr Arg Leu Leu Arg Lys Gln Ala
660 665 670

Glu Glu Val Asn Thr Glu Trp Glu Lys Leu Asn Leu His Ser Ala Asp

675 680 685

Trp Gln Arg Lys Ile Asp Glu Thr Leu Glu Arg Leu Gln Glu Leu Gln
690 695 700

Glu Ala Thr Asp Glu Leu Asp Leu Lys Leu Arg Gln Ala Glu Val Ile
705 710 715 720

Lys Gly Ser Trp Gln Pro Val Gly Asp Leu Leu Ile Asp Ser Leu Gln
725 730 735

Asp His Leu Glu Lys Val Lys Ala Leu Arg Gly Glu Ile Ala Pro Leu
740 745 750

Lys Glu Asn Val Ser His Val Asn Asp Leu Ala Arg Gln Leu Thr Thr
755 760 765

Leu Gly Ile Gln Leu Ser Pro Tyr Asn Leu Ser Thr Leu Glu Asp Leu
770 775 780

Asn Thr Arg Trp Lys Leu Leu Gln Val Ala Val Glu Asp Arg Val Arg
[0005] 785 790 795 800

Gln Leu His Glu Ala His Arg Asp Phe Gly Pro Ala Ser Gln His Phe
805 810 815

Leu Ser Thr Ser Val Gln Gly Pro Trp Glu Arg Ala Ile Ser Pro Asn
820 825 830

Lys Val Pro Tyr Tyr Ile Asn His Glu Thr Gln Thr Thr Cys Trp Asp
835 840 845

His Pro Lys Met Thr Glu Leu Tyr Gln Ser Leu Ala Asp Leu Asn Asn
850 855 860

Val Arg Phe Ser Ala Tyr Arg Thr Ala Met Lys Leu Arg Arg Leu Gln
865 870 875 880

Lys Ala Leu Cys Leu Asp Leu Leu Ser Leu Ser Ala Ala Cys Asp Ala
885 890 895

Leu Asp Gln His Asn Leu Lys Gln Asn Asp Gln Pro Met Asp Ile Leu
900 905 910

```
Gln Ile Ile Asn Cys Leu Thr Thr Ile Tyr Asp Arg Leu Glu Gln Glu
        915                 920                 925

His Asn Asn Leu Val Asn Val Pro Leu Cys Val Asp Met Cys Leu Asn
    930                 935                 940

Trp Leu Leu Asn Val Tyr Asp Thr Gly Arg Thr Gly Arg Ile Arg Val
945                 950                 955                 960

Leu Ser Phe Lys Thr Gly Ile Ile Ser Leu Cys Lys Ala His Leu Glu
                965                 970                 975

Asp Lys Tyr Arg Tyr Leu Phe Lys Gln Val Ala Ser Ser Thr Gly Phe
            980                 985                 990

Cys Asp Gln Arg Arg Leu Gly Leu  Leu Leu His Asp Ser  Ile Gln Ile
        995                 1000                1005

Pro Arg  Gln Leu Gly Glu Val  Ala Ser Phe Gly Gly  Ser Asn Ile
    1010                1015                1020
```

[0006]

```
Glu Pro  Ser Val Arg Ser Cys  Phe Gln Phe Ala Asn  Asn Lys Pro
    1025                1030                1035

Glu Ile  Glu Ala Ala Leu Phe  Leu Asp Trp Met Arg  Leu Glu Pro
    1040                1045                1050

Gln Ser  Met Val Trp Leu Pro  Val Leu His Arg Val  Ala Ala Ala
    1055                1060                1065

Glu Thr  Ala Lys His Gln Ala  Lys Cys Asn Ile Cys  Lys Glu Cys
    1070                1075                1080

Pro Ile  Ile Gly Phe Arg Tyr  Arg Ser Leu Lys His  Phe Asn Tyr
    1085                1090                1095

Asp Ile  Cys Gln Ser Cys Phe  Phe Ser Gly Arg Val  Ala Lys Gly
    1100                1105                1110

His Lys  Met His Tyr Pro Met  Val Glu Tyr Cys Thr  Pro Thr Thr
    1115                1120                1125

Ser Gly  Glu
```

1130

<210> 2
<211> 3393
<212> DNA
<213> 人工序列

<220>
<223> 编码毫微-肌养蛋白的核酸序列

<220>
<221> misc_feature
<222> (6)..(6)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (21)..(21)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (39)..(39)
<223> n是a、c、g或t

[0007]

<220>
<221> misc_feature
<222> (48)..(48)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (60)..(60)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (66)..(66)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (75)..(75)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (81)..(81)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (90)..(90)
<223> n是a、c、g或t

<220>

<220>
<221> misc_feature
<222> (99)..(99)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (120)..(120)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (126)..(126)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (132)..(132)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (141)..(141)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (144)..(144)
<223> n是a、c、g或t

[0008]

<220>
<221> misc_feature
<222> (147)..(147)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (150)..(150)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (153)..(153)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (159)..(159)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (162)..(162)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (168)..(168)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (171)..(171)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (174)..(174)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (177)..(177)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (186)..(186)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (189)..(189)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (201)..(201)
[0009] <223> n是a、c、g或t

<220>
<221> misc_feature
<222> (204)..(204)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (207)..(207)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (210)..(210)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (213)..(213)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (219)..(219)
<223> n是a、c、g或t

<220>
<221> misc_feature

<222> (222)..(222)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (231)..(231)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (240)..(240)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (243)..(243)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (246)..(246)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (249)..(249)
<223> n是a、c、g或t

[0010] <220>
<221> misc_feature
<222> (252)..(252)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (267)..(267)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (273)..(273)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (276)..(276)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (285)..(285)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (288)..(288)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (291)..(291)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (300)..(300)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (306)..(306)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (318)..(318)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (321)..(321)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (324)..(324)
<223> n是a、c、g或t

[0011]

<220>
<221> misc_feature
<222> (327)..(327)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (330)..(330)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (348)..(348)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (360)..(360)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (369)..(369)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (387)..(387)

<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (390)..(390)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (393)..(393)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (402)..(402)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (408)..(408)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (420)..(420)
<223> n是a、c、g或t

[0012]

<220>
<221> misc_feature
<222> (423)..(423)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (426)..(426)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (432)..(432)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (435)..(435)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (441)..(441)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (444)..(444)
<223> n是a、c、g或t

<220>

<220>
<221> misc_feature
<222> (447)..(447)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (456)..(456)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (462)..(462)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (468)..(468)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (480)..(480)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (483)..(483)
<223> n是a、c、g或t

[0013]

<220>
<221> misc_feature
<222> (486)..(486)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (492)..(492)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (498)..(498)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (501)..(501)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (504)..(504)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (507)..(507)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (513)..(513)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (516)..(516)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (525)..(525)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (531)..(531)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (534)..(534)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (540)..(540)
[0014] <223> n是a、c、g或t

<220>
<221> misc_feature
<222> (555)..(555)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (558)..(558)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (561)..(561)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (573)..(573)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (576)..(576)
<223> n是a、c、g或t

<220>
<221> misc_feature

<220>
<221> misc_feature
<222> (579)..(579)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (585)..(585)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (588)..(588)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (597)..(597)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (609)..(609)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (612)..(612)
<223> n是a、c、g或t

[0015]
<220>
<221> misc_feature
<222> (621)..(621)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (624)..(624)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (636)..(636)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (639)..(639)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (645)..(645)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (654)..(654)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (660)..(660)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (663)..(663)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (669)..(669)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (681)..(681)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (687)..(687)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (699)..(699)
<223> n是a、c、g或t

[0016]

<220>
<221> misc_feature
<222> (702)..(702)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (705)..(705)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (714)..(714)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (717)..(717)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (720)..(720)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (729)..(729)

<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (732)..(732)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (741)..(741)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (753)..(753)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (762)..(762)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (765)..(765)
<223> n是a、c、g或t

[0017]

<220>
<221> misc_feature
<222> (768)..(768)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (771)..(771)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (774)..(774)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (780)..(780)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (783)..(783)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (804)..(804)
<223> n是a、c、g或t

<220>

<220>
<221> misc_feature
<222> (825)..(825)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (837)..(837)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (840)..(840)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (843)..(843)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (846)..(846)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (849)..(849)
<223> n是a、c、g或t

[0018]

<220>
<221> misc_feature
<222> (855)..(855)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (864)..(864)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (867)..(867)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (870)..(870)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (873)..(873)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (876)..(876)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (882)..(882)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (885)..(885)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (894)..(894)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (900)..(900)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (906)..(906)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (912)..(912)
[0019] <223> n是a、c、g或t

<220>
<221> misc_feature
<222> (915)..(915)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (921)..(921)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (924)..(924)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (927)..(927)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (930)..(930)
<223> n是a、c、g或t

<220>
<221> misc_feature

<220>
<221> misc_feature
<222> (936)..(936)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (939)..(939)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (942)..(942)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (945)..(945)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (948)..(948)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (954)..(954)
<223> n是a、c、g或t

[0020]

<220>
<221> misc_feature
<222> (957)..(957)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (966)..(966)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (972)..(972)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (975)..(975)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (987)..(987)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (993)..(993)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (996)..(996)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (999)..(999)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1002)..(1002)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1011)..(1011)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1017)..(1017)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1023)..(1023)
<223> n是a、c、g或t

[0021]

<220>
<221> misc_feature
<222> (1029)..(1029)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1038)..(1038)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1041)..(1041)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1044)..(1044)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1053)..(1053)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1056)..(1056)

<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1059)..(1059)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1065)..(1065)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1068)..(1068)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1071)..(1071)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1074)..(1074)
<223> n是a、c、g或t

[0022]

<220>
<221> misc_feature
<222> (1083)..(1083)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1086)..(1086)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1092)..(1092)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1098)..(1098)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1107)..(1107)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1116)..(1116)
<223> n是a、c、g或t

<220>

<220>
<221> misc_feature
<222> (1122)..(1122)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1125)..(1125)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1143)..(1143)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1152)..(1152)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1167)..(1167)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1170)..(1170)
<223> n是a、c、g或t

[0023]

<220>
<221> misc_feature
<222> (1173)..(1173)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1182)..(1182)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1185)..(1185)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1188)..(1188)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1191)..(1191)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1200)..(1200)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1206)..(1206)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1209)..(1209)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1212)..(1212)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1218)..(1218)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1224)..(1224)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1227)..(1227)
[0024] <223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1230)..(1230)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1236)..(1236)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1239)..(1239)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1254)..(1254)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1260)..(1260)
<223> n是a、c、g或t

<220>
<221> misc_feature

<222> (1278)..(1278)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1281)..(1281)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1287)..(1287)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1290)..(1290)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1302)..(1302)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1305)..(1305)
<223> n是a、c、g或t

[0025] <220>
<221> misc_feature
<222> (1308)..(1308)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1311)..(1311)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1314)..(1314)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1329)..(1329)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1335)..(1335)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1341)..(1341)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1344)..(1344)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1347)..(1347)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1356)..(1356)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1371)..(1371)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1380)..(1380)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1392)..(1392)
<223> n是a、c、g或t

[0026]

<220>
<221> misc_feature
<222> (1395)..(1395)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1401)..(1401)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1407)..(1407)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1410)..(1410)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1413)..(1413)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1419)..(1419)

<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1422)..(1422)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1431)..(1431)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1434)..(1434)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1437)..(1437)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1446)..(1446)
<223> n是a、c、g或t

[0027]

<220>
<221> misc_feature
<222> (1449)..(1449)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1452)..(1452)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1461)..(1461)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1467)..(1467)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1470)..(1470)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1479)..(1479)
<223> n是a、c、g或t

<220>

<220>
<221> misc_feature
<222> (1500)..(1500)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1506)..(1506)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1527)..(1527)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1530)..(1530)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1533)..(1533)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1539)..(1539)
<223> n是a、c、g或t

[0028]

<220>
<221> misc_feature
<222> (1542)..(1542)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1545)..(1545)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1554)..(1554)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1557)..(1557)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1560)..(1560)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1563)..(1563)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1572)..(1572)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1575)..(1575)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1578)..(1578)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1587)..(1587)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1590)..(1590)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1593)..(1593)
[0029] <223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1596)..(1596)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1599)..(1599)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1611)..(1611)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1617)..(1617)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1620)..(1620)
<223> n是a、c、g或t

<220>
<221> misc_feature

<222> (1623)..(1623)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1629)..(1629)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1635)..(1635)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1647)..(1647)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1653)..(1653)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1659)..(1659)
<223> n是a、c、g或t

[0030] <220>
<221> misc_feature
<222> (1662)..(1662)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1671)..(1671)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1674)..(1674)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1683)..(1683)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1686)..(1686)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1692)..(1692)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1698)..(1698)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1701)..(1701)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1704)..(1704)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1719)..(1719)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1722)..(1722)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1728)..(1728)
<223> n是a、c、g或t

[0031]

<220>
<221> misc_feature
<222> (1731)..(1731)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1734)..(1734)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1743)..(1743)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1746)..(1746)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1749)..(1749)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1755)..(1755)

<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1761)..(1761)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1776)..(1776)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1779)..(1779)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1782)..(1782)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1788)..(1788)
<223> n是a、c、g或t

[0032]

<220>
<221> misc_feature
<222> (1791)..(1791)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1797)..(1797)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1800)..(1800)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1809)..(1809)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1812)..(1812)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1815)..(1815)
<223> n是a、c、g或t

<220>

<220>
<221> misc_feature
<222> (1833)..(1833)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1839)..(1839)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1842)..(1842)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1851)..(1851)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1857)..(1857)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1863)..(1863)
<223> n是a、c、g或t

[0033]

<220>
<221> misc_feature
<222> (1872)..(1872)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1875)..(1875)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1884)..(1884)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1887)..(1887)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1890)..(1890)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1896)..(1896)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1899)..(1899)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1902)..(1902)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1911)..(1911)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1914)..(1914)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1923)..(1923)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1926)..(1926)
[0034] <223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1932)..(1932)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1935)..(1935)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1944)..(1944)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1956)..(1956)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1959)..(1959)
<223> n是a、c、g或t

<220>
<221> misc_feature

<220>
<221> misc_feature
<222> (1965)..(1965)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1968)..(1968)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1971)..(1971)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1980)..(1980)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1983)..(1983)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1992)..(1992)
<223> n是a、c、g或t

[0035] <220>
<221> misc_feature
<222> (1995)..(1995)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1998)..(1998)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2001)..(2001)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2004)..(2004)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2007)..(2007)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2016)..(2016)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2025)..(2025)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2031)..(2031)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2046)..(2046)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2052)..(2052)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2058)..(2058)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2061)..(2061)
<223> n是a、c、g或t

[0036]

<220>
<221> misc_feature
<222> (2073)..(2073)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2088)..(2088)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2091)..(2091)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2097)..(2097)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2100)..(2100)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2109)..(2109)

<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2118)..(2118)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2121)..(2121)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2130)..(2130)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2136)..(2136)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2142)..(2142)
<223> n是a、c、g或t

[0037]

<220>
<221> misc_feature
<222> (2145)..(2145)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2151)..(2151)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2157)..(2157)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2166)..(2166)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2169)..(2169)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2178)..(2178)
<223> n是a、c、g或t

<220>

<220>
<221> misc_feature
<222> (2181)..(2181)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2184)..(2184)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2190)..(2190)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2193)..(2193)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2202)..(2202)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2205)..(2205)
<223> n是a、c、g或t

[0038]

<220>
<221> misc_feature
<222> (2217)..(2217)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2226)..(2226)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2232)..(2232)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2235)..(2235)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2238)..(2238)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2241)..(2241)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2250)..(2250)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2253)..(2253)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2256)..(2256)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2268)..(2268)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2271)..(2271)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2277)..(2277)
[0039] <223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2286)..(2286)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2289)..(2289)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2292)..(2292)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2298)..(2298)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2301)..(2301)
<223> n是a、c、g或t

<220>
<221> misc_feature

<220>
<221> misc_feature
<222> (2304)..(2304)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2307)..(2307)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2310)..(2310)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2319)..(2319)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2322)..(2322)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2325)..(2325)
<223> n是a、c、g或t

[0040] <220>
<221> misc_feature
<222> (2334)..(2334)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2337)..(2337)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2340)..(2340)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2343)..(2343)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2352)..(2352)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2358)..(2358)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2361)..(2361)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2370)..(2370)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2373)..(2373)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2379)..(2379)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2382)..(2382)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2385)..(2385)
<223> n是a、c、g或t

[0041]

<220>
<221> misc_feature
<222> (2394)..(2394)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2397)..(2397)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2400)..(2400)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2406)..(2406)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2415)..(2415)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2421)..(2421)

<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2430)..(2430)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2433)..(2433)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2436)..(2436)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2439)..(2439)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2451)..(2451)
<223> n是a、c、g或t

[0042]

<220>
<221> misc_feature
<222> (2454)..(2454)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2457)..(2457)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2460)..(2460)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2463)..(2463)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2469)..(2469)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2472)..(2472)
<223> n是a、c、g或t

<220>

<220>
<221> misc_feature
<222> (2481)..(2481)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2484)..(2484)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2490)..(2490)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2493)..(2493)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2502)..(2502)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2505)..(2505)
<223> n是a、c、g或t

[0043]

<220>
<221> misc_feature
<222> (2526)..(2526)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2532)..(2532)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2535)..(2535)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2550)..(2550)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2559)..(2559)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2565)..(2565)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2574)..(2574)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2577)..(2577)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2580)..(2580)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2586)..(2586)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2595)..(2595)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2598)..(2598)
[0044] <223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2604)..(2604)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2607)..(2607)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2613)..(2613)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2616)..(2616)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2619)..(2619)
<223> n是a、c、g或t

<220>
<221> misc_feature

<220>
<221> misc_feature
<222> (2628)..(2628)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2631)..(2631)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2634)..(2634)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2637)..(2637)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2646)..(2646)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2649)..(2649)
<223> n是a、c、g或t

[0045] <220>
<221> misc_feature
<222> (2655)..(2655)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2661)..(2661)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2664)..(2664)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2667)..(2667)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2670)..(2670)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2673)..(2673)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2676)..(2676)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2679)..(2679)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2688)..(2688)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2691)..(2691)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2706)..(2706)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2724)..(2724)
<223> n是a、c、g或t

[0046]

<220>
<221> misc_feature
<222> (2736)..(2736)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2754)..(2754)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2757)..(2757)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2760)..(2760)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2772)..(2772)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2775)..(2775)

<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2796)..(2796)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2799)..(2799)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2805)..(2805)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2808)..(2808)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2811)..(2811)
<223> n是a、c、g或t

[0047]

<220>
<221> misc_feature
<222> (2817)..(2817)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2829)..(2829)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2838)..(2838)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2841)..(2841)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2847)..(2847)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2856)..(2856)
<223> n是a、c、g或t

<220>

<220>
<221> misc_feature
<222> (2859)..(2859)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2862)..(2862)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2865)..(2865)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2868)..(2868)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2871)..(2871)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2877)..(2877)
<223> n是a、c、g或t

[0048]

<220>
<221> misc_feature
<222> (2880)..(2880)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2883)..(2883)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2886)..(2886)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2895)..(2895)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2898)..(2898)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2907)..(2907)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2910)..(2910)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2919)..(2919)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2925)..(2925)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2940)..(2940)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2946)..(2946)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2958)..(2958)
[0049] <223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2961)..(2961)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2964)..(2964)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2967)..(2967)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2970)..(2970)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2973)..(2973)
<223> n是a、c、g或t

<220>
<221> misc_feature

<220>
<221> misc_feature
<222> (2988)..(2988)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2991)..(2991)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2994)..(2994)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2997)..(2997)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3000)..(3000)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3003)..(3003)
<223> n是a、c、g或t

[0050]

<220>
<221> misc_feature
<222> (3006)..(3006)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3015)..(3015)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3027)..(3027)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3030)..(3030)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3036)..(3036)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3039)..(3039)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3045)..(3045)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3048)..(3048)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3051)..(3051)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3057)..(3057)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3060)..(3060)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3063)..(3063)
<223> n是a、c、g或t

[0051]

<220>
<221> misc_feature
<222> (3075)..(3075)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3078)..(3078)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3081)..(3081)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3084)..(3084)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3087)..(3087)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3102)..(3102)

<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3114)..(3114)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3126)..(3126)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3129)..(3129)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3132)..(3132)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3138)..(3138)
<223> n是a、c、g或t

[0052]

<220>
<221> misc_feature
<222> (3150)..(3150)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3153)..(3153)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3159)..(3159)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3165)..(3165)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3171)..(3171)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3177)..(3177)
<223> n是a、c、g或t

<220>

<220>
<221> misc_feature
<222> (3180)..(3180)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3183)..(3183)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3186)..(3186)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3192)..(3192)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3195)..(3195)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3198)..(3198)
<223> n是a、c、g或t

[0053]

<220>
<221> misc_feature
<222> (3201)..(3201)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3204)..(3204)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3210)..(3210)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3213)..(3213)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3225)..(3225)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3252)..(3252)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3261)..(3261)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3267)..(3267)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3273)..(3273)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3276)..(3276)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3279)..(3279)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3309)..(3309)
[0054] <223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3321)..(3321)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3324)..(3324)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3327)..(3327)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3330)..(3330)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3333)..(3333)
<223> n是a、c、g或t

<220>
<221> misc_feature

<220>
<221> misc_feature
<222> (3339)..(3339)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3357)..(3357)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3363)..(3363)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3375)..(3375)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3378)..(3378)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3381)..(3381)
<223> n是a、c、g或t

[0055]

<220>
<221> misc_feature
<222> (3384)..(3384)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3387)..(3387)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3390)..(3390)
<223> n是a、c、g或t

<400> 2

atgytntggt gggargargt ngargaytgy taygarmgng argaygtnca raaraaracn 60

ttyacnaart gggtnaaygc ncarttywsn aarttyggna arcarcayat hgaraayytn 120

ttywsngayy tncargaygg nmgnmgnytn ytngayytny tngarggnyt nacnggncar 180

aarytnccna argaraargg nwsnacnmgn gtncaygcny tnaayaaygt naayaargcn 240

ytnmgngtny tncaraayaa yaaygtngay ytngtnaaya thggnwsnac ngayathgtn 300

gayggnaayc ayaarytnac nytnggnytn athtggaaya thathytnca ytggcargtn 360

aaraaygtna tgaaraayat hatggcnggn ytncarcara cnaaywsnga raarathytn 420

ytnwsntggg tnmgncarws nacnmgnaay tayccncarg tnaaygtnat haayttyacn 480

acnwsntggw sngayggnyt ngcnytnaay gcnytnathc aywsncaymg nccngayytn 540

ttygaytgga aywsngtngt ntgycarcar wsngcnacnc armgnytnga rcaygcntty 600

aayathgcnm gntaycaryt nggnathgar aarytnytng ayccngarga ygtngayacn 660

acntayccng ayaaraarws nathytnatg tayathacnw snytnttyca rgtnytnccn 720

carcargtnw snathgargc nathcargar gtngaratgy tnccnmgncc nccnaargtn 780

acnaargarg arcayttyca rytncaycay caratgcayt aywsncarca rathacngtn 840

wsnytngcnc arggntayga rmgnacnwsn wsnccnaarc cnmgnttyaa rwsntaygcn 900

tayacncarg cngcntaygt nacnacnwsn gayccnacnm gnwsnccntt yccnwsncar 960

cayytngarg cnccngarga yaarwsntty ggnwsnwsny tnatggarws ngargtnaay 1020

ytngaymgnt aycaracngc nytngargar gtnytnwsnt ggytnytnws ngcngargay 1080

acnytncarg cncarggnga rathwsnaay gaygtngarg tngtnaarga ycarttycay 1140

acncaygarg gntayatgat ggayytnacn gcncaycarg gnmgngtngg naayathytn 1200

carytnggnw snaarytnat hggnacnggn aarytnwsng argaygarga racngargtn 1260

[0056] cargarcara tgaayytnyt naaywsnmgn tgggartgyy tnmgngtngc nwsnatggar 1320

aarcarwsna ayytncaymg ngtnytnatg gayytncara aycaraaryt naargarytn 1380

aaygaytggy tnacngargc ngaracnacn gcnaaygtny tncargaygc nacnmgnaar 1440

garmgnytny tngargayws naarggngtn aargarytna tgaarcartg gcargayytn 1500

carggngara thgarcarga rcargtnmgn gtnaaywsny tnacncayat ggtngtngtn 1560

gtngaygarw snwsnggnga ycaygcnacn gcngcnytng argarcaryt naargtnytn 1620

ggngaymgnt gggcnaayat htgymgntgg wsngarytnm gnaaraarws nytnaayath 1680

mgnwsncayy tngargcnws nwsngaycar tggaarmgny tncayytnws nytncargar 1740

ytnytngtnt ggytncaryt naargaygay garytnwsnm gncargcncc nathggnggn 1800

gayttyccng cngtncaraa rcaraaygay gtncaymgng cnttyaarmg ngarytnaar 1860

acnaargarc cngtnathat gwsnacnytn garacngtnm gnathttyyt nacngarcar 1920

ccnytngarg gnytngaraa rytntaycar garccnmgng arytnccncc ngargarmgn 1980

gcncaraayg tnacnmgnyt nytnmgnaar cargcngarg argtnaayac ngartgggar 2040

aarytnaayy tncaywsngc ngaytggcar mgnaarathg aygaracnyt ngarmgnytn 2100

cargarytnc argargcnac ngaygarytn gayytnaary tnmgncargc ngargtnath 2160

aarggnwsnt ggcarccngt nggngayytn ytnathgayw snytncarga ycayytngar 2220

aargtnaarg cnytnmgngg ngarathgcn ccnytnaarg araaygtnws ncaygtnaay 2280

gayytngcnm gncarytnac nacnytnggn athcarytnw snccntayaa yytnwsnacn 2340

ytngargayy tnaayacnmg ntggaarytn ytncargtng cngtngarga ymgngtnmgn 2400

carytncayg argcncaymg ngayttyggn ccngcnwsnc arcayttyyt nwsnacnwsn 2460

gtncarggnc cntgggarmg ngcnathwsn ccnaayaarg tnccntayta yathaaycay 2520

garacncara cnacntgytg ggaycayccn aaratgacng arytntayca rwsnytngcn 2580

gayytnaaya aygtnmgntt ywsngcntay mgnacngcna tgaarytnmg nmgnytncar 2640

aargcnytnt gyytngayyt nytnwsnytn wsngcngcnt gygaygcnyt ngaycarcay 2700

aayytnaarc araaygayca rccnatggay athytncara thathaaytg yytnacnacn 2760

athtaygaym gnytngarca rgarcayaay aayytngtna aygtnccnyt ntgygtngay 2820

atgtgyytna aytggytnyt naaygtntay gayacnggnm gnacnggnmg nathmgngtn 2880

ytnwsnttya aracnggnat hathwsnytn tgyaargcnc ayytngarga yaartaymgn 2940

tayytnttya arcargtngc nwsnwsnacn ggnttytgyg aycarmgnmg nytnggnytn 3000

[0057] ytnytncayg aywsnathca rathccnmgn carytnggng argtngcnws nttyggnggn 3060

wsnaayathg arccnwsngt nmgnwsntgy ttycarttyg cnaayaayaa rccngarath 3120

gargcngcny tnttyytnga ytggatgmgn ytngarccnc arwsnatggt ntggytnccn 3180

gtnytncaym gngtngcngc ngcngaracn gcnaarcayc argcnaartg yaayathtgy 3240

aargartgyc cnathathgg nttymgntay mgnwsnytna arcayttyaa ytaygayath 3300

tgycarwsnt gyttyttyws nggnmgngtn gcnaarggnc ayaaratgca ytayccnatg 3360

gtngartayt gyacnccnac nacnwsnggn gar 3393

<210> 3
<211> 1098
<212> PRT
<213> 人工序列

<220>
<223> 毫微-肌养蛋白相关蛋白 1 的氨基酸序列

<220>
<221> MISC_FEATURE
<222> (1)..(433)
<223> 与全长人肌养蛋白相关蛋白的 N-末端区域相同的氨基酸序列

<220>

<220>
<221> MISC_FEATURE
<222> (433)..(433)
<223> 位于所述超家族中蛋白的隐蔽马尔科夫模型（HMM）和所有晶体结构的核心处的“A”螺旋中的色氨酸残基

<220>
<221> MISC_FEATURE
<222> (434)..(477)
<223> 与全长人肌养蛋白相关蛋白的C-末端区域相同的氨基酸序列

<220>
<221> MISC_FEATURE
<222> (477)..(478)
<223> 侧接如图2F中所描绘的假设横切面平面的“B”螺旋的超家族HMM内的位置

<220>
<221> MISC_FEATURE
<222> (478)..(514)
<223> 与全长人肌养蛋白相关蛋白的N-末端区域相同的氨基酸序列

<220>
<221> MISC_FEATURE
<222> (514)..(514)
<223> 位于所述超家族中蛋白的隐蔽马尔科夫模型（HMM）和所有晶体结构的核心处的“C”螺旋中的色氨酸残基

<220>
<221> MISC_FEATURE
<222> (515)..(1098)
<223> 与全长人肌养蛋白相关蛋白的C-末端区域相同的氨基酸序列

[0058]

<400> 3

```
Met Ala Lys Tyr Gly Glu His Glu Ala Ser Pro Asp Asn Gly Gln Asn
1               5                  10                  15

Glu Phe Ser Asp Ile Ile Lys Ser Arg Ser Asp Glu His Asn Asp Val
            20                  25                  30

Gln Lys Lys Thr Phe Thr Lys Trp Ile Asn Ala Arg Phe Ser Lys Ser
        35                  40                  45

Gly Lys Pro Pro Ile Asn Asp Met Phe Thr Asp Leu Lys Asp Gly Arg
    50                  55                  60

Lys Leu Leu Asp Leu Leu Glu Gly Leu Thr Gly Thr Ser Leu Pro Lys
65                  70                  75                  80

Glu Arg Gly Ser Thr Arg Val His Ala Leu Asn Asn Val Asn Arg Val
                85                  90                  95

Leu Gln Val Leu His Gln Asn Asn Val Glu Leu Val Asn Ile Gly Gly
```

100 105 110

Thr Asp Ile Val Asp Gly Asn His Lys Leu Thr Leu Gly Leu Leu Trp
115 120 125

Ser Ile Ile Leu His Trp Gln Val Lys Asp Val Met Lys Asp Val Met
130 135 140

Ser Asp Leu Gln Gln Thr Asn Ser Glu Lys Ile Leu Leu Ser Trp Val
145 150 155 160

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

Thr Ser Trp Thr Asp Gly Leu Ala Phe Asn Ala Val Leu His Arg His
180 185 190

Lys Pro Asp Leu Phe Ser Trp Asp Lys Val Val Lys Met Ser Pro Ile
195 200 205

Glu Arg Leu Glu His Ala Phe Ser Lys Ala Gln Thr Tyr Leu Gly Ile
[0059] 210 215 220

Glu Lys Leu Leu Asp Pro Glu Asp Val Ala Val Gln Leu Pro Asp Lys
225 230 235 240

Lys Ser Ile Ile Met Tyr Leu Thr Ser Leu Phe Glu Val Leu Pro Gln
245 250 255

Gln Val Thr Ile Asp Ala Ile Arg Glu Val Glu Thr Leu Pro Arg Lys
260 265 270

Tyr Lys Lys Glu Cys Glu Glu Glu Ala Ile Asn Ile Gln Ser Thr Ala
275 280 285

Pro Glu Glu Glu His Glu Ser Pro Arg Ala Glu Thr Pro Ser Thr Val
290 295 300

Thr Glu Val Asp Met Asp Leu Asp Ser Tyr Gln Ile Ala Leu Glu Glu
305 310 315 320

Val Leu Thr Trp Leu Leu Ser Ala Glu Asp Thr Phe Gln Glu Gln Asp
325 330 335

Asp Ile Ser Asp Asp Val Glu Glu Val Lys Asp Gln Phe Ala Thr His
340 345 350

Glu Ala Phe Met Met Glu Leu Thr Ala His Gln Ser Ser Val Gly Ser
355 360 365

Val Leu Gln Ala Gly Asn Gln Leu Ile Thr Gln Gly Thr Leu Ser Asp
370 375 380

Glu Glu Glu Phe Glu Ile Gln Glu Gln Met Thr Leu Leu Asn Ala Arg
385 390 395 400

Trp Glu Ala Leu Arg Val Glu Ser Met Asp Arg Gln Ser Arg Leu His
405 410 415

Asp Val Leu Met Glu Leu Gln Lys Lys Gln Leu Gln Gln Leu Ser Ala
420 425 430

Trp Ile Gln Glu Ala Glu Thr Thr Val Asn Val Leu Val Asp Ala Ser
435 440 445

[0060] His Arg Glu Asn Ala Leu Gln Asp Ser Ile Leu Ala Arg Glu Leu Lys
450 455 460

Gln Gln Met Gln Asp Ile Gln Ala Glu Ile Asp Ala His Gln Val Lys
465 470 475 480

Val Asn Ser Leu Thr His Met Val Val Ile Val Asp Glu Asn Ser Gly
485 490 495

Glu Ser Ala Thr Ala Ile Leu Glu Asp Gln Leu Gln Lys Leu Gly Glu
500 505 510

Arg Trp Asn Asp Leu Lys Ala Lys Ser Ala Ser Ile Arg Ala His Leu
515 520 525

Glu Ala Ser Ala Glu Lys Trp Asn Arg Leu Leu Met Ser Leu Glu Glu
530 535 540

Leu Ile Lys Trp Leu Asn Met Lys Asp Glu Glu Leu Lys Lys Gln Met
545 550 555 560

Pro Ile Gly Gly Asp Val Pro Ala Leu Gln Leu Gln Tyr Asp His Cys

565 570 575

Lys Ala Leu Arg Arg Glu Leu Lys Glu Lys Glu Tyr Ser Val Leu Asn
580 585 590

Ala Val Asp Gln Ala Arg Val Phe Leu Ala Asp Gln Pro Ile Glu Ala
595 600 605

Pro Glu Glu Pro Arg Arg Asn Leu Gln Ser Lys Thr Glu Leu Thr Pro
610 615 620

Glu Glu Arg Ala Gln Lys Ile Ala Lys Ala Met Arg Lys Gln Ser Ser
625 630 635 640

Glu Val Lys Glu Lys Trp Glu Ser Leu Asn Ala Val Thr Ser Asn Trp
645 650 655

Gln Lys Gln Val Asp Lys Ala Leu Glu Lys Leu Arg Asp Leu Gln Gly
660 665 670

[0061] Ala Met Asp Asp Leu Asp Ala Asp Met Lys Glu Ala Glu Ser Val Arg
675 680 685

Asn Gly Trp Lys Pro Val Gly Asp Leu Leu Ile Asp Ser Leu Gln Asp
690 695 700

His Ile Glu Lys Ile Met Ala Phe Arg Glu Glu Ile Ala Pro Ile Asn
705 710 715 720

Phe Lys Val Lys Thr Val Asn Asp Leu Ser Ser Gln Leu Ser Pro Leu
725 730 735

Asp Leu His Pro Ser Leu Lys Met Ser Arg Gln Leu Asp Asp Leu Asn
740 745 750

Met Arg Trp Lys Leu Leu Gln Val Ser Val Asp Asp Arg Leu Lys Gln
755 760 765

Leu Gln Glu Ala His Arg Asp Phe Gly Pro Ser Ser Gln His Phe Leu
770 775 780

Ser Thr Ser Val Gln Leu Pro Trp Gln Arg Ser Ile Ser His Asn Lys
785 790 795 800

Val Pro Tyr Tyr Ile Asn His Gln Thr Gln Thr Thr Cys Trp Asp His
805 810 815

Pro Lys Met Thr Glu Leu Phe Gln Ser Leu Ala Asp Leu Asn Asn Val
820 825 830

Arg Phe Ser Ala Tyr Arg Thr Ala Ile Lys Ile Arg Arg Leu Gln Lys
835 840 845

Ala Leu Cys Leu Asp Leu Leu Glu Leu Ser Thr Thr Asn Glu Ile Phe
850 855 860

Lys Gln His Lys Leu Asn Gln Asn Asp Gln Leu Leu Ser Val Pro Asp
865 870 875 880

Val Ile Asn Cys Leu Thr Thr Thr Tyr Asp Gly Leu Glu Gln Met His
885 890 895

Lys Asp Leu Val Asn Val Pro Leu Cys Val Asp Met Cys Leu Asn Trp
900 905 910

[0062] Leu Leu Asn Val Tyr Asp Thr Gly Arg Thr Gly Lys Ile Arg Val Gln
915 920 925

Ser Leu Lys Ile Gly Leu Met Ser Leu Ser Lys Gly Leu Leu Glu Glu
930 935 940

Lys Tyr Arg Tyr Leu Phe Lys Glu Val Ala Gly Pro Thr Glu Met Cys
945 950 955 960

Asp Gln Arg Gln Leu Gly Leu Leu Leu His Asp Ala Ile Gln Ile Pro
965 970 975

Arg Gln Leu Gly Glu Val Ala Ala Phe Gly Gly Ser Asn Ile Glu Pro
980 985 990

Ser Val Arg Ser Cys Phe Gln Gln Asn Asn Asn Lys Pro Glu Ile Ser
995 1000 1005

Val Lys Glu Phe Ile Asp Trp Met His Leu Glu Pro Gln Ser Met
1010 1015 1020

Val Trp Leu Pro Val Leu His Arg Val Ala Ala Ala Glu Thr Ala

1025 1030 1035

Lys His Gln Ala Lys Cys Asn Ile Cys Lys Glu Cys Pro Ile Val
1040 1045 1050

Gly Phe Arg Tyr Arg Ser Leu Lys His Phe Asn Tyr Asp Val Cys
1055 1060 1065

Gln Ser Cys Phe Phe Ser Gly Arg Thr Ala Lys Gly His Lys Leu
1070 1075 1080

His Tyr Pro Met Val Glu Tyr Cys Ile Pro Thr Thr Ser Gly Glu
1085 1090 1095

<210> 4
<211> 3294
<212> DNA
<213> 人工序列

<220>
<223> 编码毫微-肌养蛋白相关蛋白-1蛋白的核酸序列

[0063]
<220>
<221> misc_feature
<222> (6)..(6)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (15)..(15)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (27)..(27)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (30)..(30)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (33)..(33)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (42)..(42)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (57)..(57)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (72)..(72)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (75)..(75)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (78)..(78)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (96)..(96)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (108)..(108)
<223> n是a、c、g或t

[0064]

<220>
<221> misc_feature
<222> (114)..(114)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (129)..(129)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (132)..(132)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (138)..(138)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (144)..(144)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (147)..(147)

<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (153)..(153)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (156)..(156)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (174)..(174)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (180)..(180)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (189)..(189)
<223> n是a、c、g或t

[0065]

<220>
<221> misc_feature
<222> (192)..(192)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (198)..(198)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (201)..(201)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (207)..(207)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (210)..(210)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (216)..(216)
<223> n是a、c、g或t

<220>

<220>
<221> misc_feature
<222> (219)..(219)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (222)..(222)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (225)..(225)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (228)..(228)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (231)..(231)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (234)..(234)
<223> n是a、c、g或t

[0066]

<220>
<221> misc_feature
<222> (237)..(237)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (246)..(246)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (249)..(249)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (252)..(252)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (255)..(255)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (258)..(258)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (261)..(261)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (267)..(267)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (270)..(270)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (279)..(279)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (285)..(285)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (288)..(288)
[0067] <223> n是a、c、g或t

<220>
<221> misc_feature
<222> (291)..(291)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (297)..(297)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (300)..(300)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (315)..(315)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (321)..(321)
<223> n是a、c、g或t

<220>
<221> misc_feature

<220>
<221> misc_feature
<222> (324)..(324)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (333)..(333)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (336)..(336)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (339)..(339)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (348)..(348)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (354)..(354)
<223> n是a、c、g或t

[0068]

<220>
<221> misc_feature
<222> (366)..(366)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (369)..(369)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (372)..(372)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (375)..(375)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (378)..(378)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (381)..(381)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (387)..(387)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (396)..(396)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (408)..(408)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (417)..(417)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (429)..(429)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (435)..(435)
<223> n是a、c、g或t

[0069]

<220>
<221> misc_feature
<222> (441)..(441)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (450)..(450)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (456)..(456)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (468)..(468)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (471)..(471)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (474)..(474)

<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (480)..(480)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (483)..(483)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (489)..(489)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (492)..(492)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (495)..(495)
<223> n是a、c、g或t

[0070]

<220>
<221> misc_feature
<222> (498)..(498)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (504)..(504)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (510)..(510)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (516)..(516)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (519)..(519)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (528)..(528)
<223> n是a、c、g或t

<220>

<220>
<221> misc_feature
<222> (531)..(531)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (534)..(534)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (540)..(540)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (546)..(546)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (549)..(549)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (552)..(552)
<223> n是a、c、g或t

[0071]

<220>
<221> misc_feature
<222> (561)..(561)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (564)..(564)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (567)..(567)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (573)..(573)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (582)..(582)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (588)..(588)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (594)..(594)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (606)..(606)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (609)..(609)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (618)..(618)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (621)..(621)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (630)..(630)
[0072] <223> n是a、c、g或t

<220>
<221> misc_feature
<222> (633)..(633)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (642)..(642)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (648)..(648)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (654)..(654)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (660)..(660)
<223> n是a、c、g或t

<220>
<221> misc_feature

<220>
<221> misc_feature
<222> (666)..(666)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (669)..(669)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (681)..(681)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (684)..(684)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (690)..(690)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (699)..(699)
<223> n是a、c、g或t

[0073]

<220>
<221> misc_feature
<222> (702)..(702)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (705)..(705)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (711)..(711)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (714)..(714)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (726)..(726)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (741)..(741)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (744)..(744)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (747)..(747)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (750)..(750)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (759)..(759)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (762)..(762)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (765)..(765)
<223> n是a、c、g或t

[0074]

<220>
<221> misc_feature
<222> (774)..(774)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (777)..(777)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (786)..(786)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (792)..(792)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (798)..(798)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (804)..(804)

<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (807)..(807)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (810)..(810)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (813)..(813)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (843)..(843)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (858)..(858)
<223> n是a、c、g或t

[0075]

<220>
<221> misc_feature
<222> (861)..(861)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (864)..(864)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (867)..(867)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (885)..(885)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (888)..(888)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (891)..(891)
<223> n是a、c、g或t

<220>

<220>
<221> misc_feature
<222> (894)..(894)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (900)..(900)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (903)..(903)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (906)..(906)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (909)..(909)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (912)..(912)
<223> n是a、c、g或t

[0076]

<220>
<221> misc_feature
<222> (915)..(915)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (921)..(921)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (933)..(933)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (939)..(939)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (951)..(951)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (954)..(954)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (963)..(963)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (966)..(966)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (969)..(969)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (975)..(975)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (978)..(978)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (981)..(981)
[0077] <223> n是a、c、g或t

<220>
<221> misc_feature
<222> (984)..(984)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (993)..(993)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1017)..(1017)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1026)..(1026)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1035)..(1035)
<223> n是a、c、g或t

<220>
<221> misc_feature

<220>
<222> (1050)..(1050)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1053)..(1053)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1062)..(1062)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1077)..(1077)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1080)..(1080)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1083)..(1083)
<223> n是a、c、g或t

[0078]

<220>
<221> misc_feature
<222> (1092)..(1092)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1095)..(1095)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1098)..(1098)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1101)..(1101)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1104)..(1104)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1107)..(1107)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1110)..(1110)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1116)..(1116)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1119)..(1119)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1128)..(1128)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1134)..(1134)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1140)..(1140)
<223> n是a、c、g或t

[0079]

<220>
<221> misc_feature
<222> (1143)..(1143)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1146)..(1146)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1149)..(1149)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1185)..(1185)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1188)..(1188)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1191)..(1191)

<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1197)..(1197)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1200)..(1200)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1209)..(1209)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1212)..(1212)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1215)..(1215)
<223> n是a、c、g或t

[0080]

<220>
<221> misc_feature
<222> (1218)..(1218)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1224)..(1224)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1233)..(1233)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1239)..(1239)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1242)..(1242)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1245)..(1245)
<223> n是a、c、g或t

<220>

<220>
<221> misc_feature
<222> (1254)..(1254)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1257)..(1257)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1266)..(1266)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1281)..(1281)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1290)..(1290)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1293)..(1293)
<223> n是a、c、g或t

[0081]

<220>
<221> misc_feature
<222> (1296)..(1296)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1311)..(1311)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1317)..(1317)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1320)..(1320)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1323)..(1323)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1329)..(1329)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1332)..(1332)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1335)..(1335)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1341)..(1341)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1344)..(1344)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1350)..(1350)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1359)..(1359)
[0082] <223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1362)..(1362)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1371)..(1371)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1377)..(1377)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1380)..(1380)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1383)..(1383)
<223> n是a、c、g或t

<220>
<221> misc_feature

<222> (1389)..(1389)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1416)..(1416)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1428)..(1428)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1437)..(1437)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1443)..(1443)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1449)..(1449)
<223> n是a、c、g或t

[0083] <220>
<221> misc_feature
<222> (1452)..(1452)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1455)..(1455)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1464)..(1464)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1467)..(1467)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1473)..(1473)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1485)..(1485)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1488)..(1488)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1494)..(1494)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1497)..(1497)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1500)..(1500)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1503)..(1503)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1509)..(1509)
<223> n是a、c、g或t

[0084]

<220>
<221> misc_feature
<222> (1521)..(1521)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1530)..(1530)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1533)..(1533)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1539)..(1539)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1551)..(1551)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1557)..(1557)

<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1563)..(1563)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1566)..(1566)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1569)..(1569)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1575)..(1575)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1578)..(1578)
<223> n是a、c、g或t

[0085]

<220>
<221> misc_feature
<222> (1584)..(1584)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1590)..(1590)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1593)..(1593)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1596)..(1596)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1611)..(1611)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1614)..(1614)
<223> n是a、c、g或t

<220>

<220>
<221> misc_feature
<222> (1617)..(1617)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1623)..(1623)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1626)..(1626)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1635)..(1635)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1647)..(1647)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1668)..(1668)
<223> n是a、c、g或t

[0086]

<220>
<221> misc_feature
<222> (1683)..(1683)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1689)..(1689)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1692)..(1692)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1698)..(1698)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1701)..(1701)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1704)..(1704)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1707)..(1707)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1713)..(1713)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1734)..(1734)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1737)..(1737)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1740)..(1740)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1743)..(1743)
[0087] <223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1749)..(1749)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1767)..(1767)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1770)..(1770)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1773)..(1773)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1779)..(1779)
<223> n是a、c、g或t

<220>
<221> misc_feature

<222> (1782)..(1782)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1791)..(1791)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1794)..(1794)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1797)..(1797)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1803)..(1803)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1806)..(1806)
<223> n是a、c、g或t

[0088] <220>
<221> misc_feature
<222> (1815)..(1815)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1824)..(1824)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1827)..(1827)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1836)..(1836)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1839)..(1839)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1842)..(1842)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1848)..(1848)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1854)..(1854)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1860)..(1860)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1866)..(1866)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1869)..(1869)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1872)..(1872)
<223> n是a、c、g或t

[0089]

<220>
<221> misc_feature
<222> (1881)..(1881)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1884)..(1884)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1896)..(1896)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1902)..(1902)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1908)..(1908)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1917)..(1917)

<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1920)..(1920)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1926)..(1926)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1944)..(1944)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1947)..(1947)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1953)..(1953)
<223> n是a、c、g或t

<220>
[0090] <221> misc_feature
<222> (1956)..(1956)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1959)..(1959)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1962)..(1962)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1980)..(1980)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1989)..(1989)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (1992)..(1992)
<223> n是a、c、g或t

<220>

<220>
<221> misc_feature
<222> (2001)..(2001)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2004)..(2004)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2010)..(2010)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2016)..(2016)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2019)..(2019)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2031)..(2031)
<223> n是a、c、g或t

[0091]

<220>
<221> misc_feature
<222> (2037)..(2037)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2052)..(2052)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2058)..(2058)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2061)..(2061)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2064)..(2064)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2070)..(2070)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2079)..(2079)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2082)..(2082)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2085)..(2085)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2091)..(2091)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2094)..(2094)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2103)..(2103)
[0092] <223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2106)..(2106)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2133)..(2133)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2139)..(2139)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2151)..(2151)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2154)..(2154)
<223> n是a、c、g或t

<220>
<221> misc_feature

<222> (2169)..(2169)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2175)..(2175)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2178)..(2178)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2187)..(2187)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2190)..(2190)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2193)..(2193)
<223> n是a、c、g或t

[0093] <220>
<221> misc_feature
<222> (2199)..(2199)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2202)..(2202)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2205)..(2205)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2208)..(2208)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2214)..(2214)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2220)..(2220)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2223)..(2223)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2226)..(2226)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2235)..(2235)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2238)..(2238)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2244)..(2244)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2253)..(2253)
<223> n是a、c、g或t

[0094]

<220>
<221> misc_feature
<222> (2262)..(2262)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2271)..(2271)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2274)..(2274)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2280)..(2280)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2283)..(2283)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2286)..(2286)

<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2295)..(2295)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2298)..(2298)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2307)..(2307)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2316)..(2316)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2322)..(2322)
<223> n是a、c、g或t

[0095]

<220>
<221> misc_feature
<222> (2331)..(2331)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2334)..(2334)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2337)..(2337)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2340)..(2340)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2352)..(2352)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2355)..(2355)
<223> n是a、c、g或t

<220>

<220>
<221> misc_feature
<222> (2358)..(2358)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2361)..(2361)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2364)..(2364)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2370)..(2370)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2373)..(2373)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2382)..(2382)
<223> n是a、c、g或t

[0096]

<220>
<221> misc_feature
<222> (2385)..(2385)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2391)..(2391)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2403)..(2403)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2406)..(2406)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2427)..(2427)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2433)..(2433)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2436)..(2436)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2451)..(2451)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2460)..(2460)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2466)..(2466)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2475)..(2475)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2478)..(2478)
[0097] <223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2481)..(2481)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2487)..(2487)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2496)..(2496)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2499)..(2499)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2505)..(2505)
<223> n是a、c、g或t

<220>
<221> misc_feature

<222> (2508)..(2508)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2514)..(2514)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2517)..(2517)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2520)..(2520)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2532)..(2532)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2535)..(2535)
<223> n是a、c、g或t

[0098] <220>
<221> misc_feature
<222> (2538)..(2538)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2547)..(2547)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2550)..(2550)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2556)..(2556)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2562)..(2562)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2565)..(2565)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2571)..(2571)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2574)..(2574)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2577)..(2577)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2580)..(2580)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2607)..(2607)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2625)..(2625)
<223> n是a、c、g或t

[0099]

<220>
<221> misc_feature
<222> (2628)..(2628)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2631)..(2631)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2634)..(2634)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2637)..(2637)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2643)..(2643)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2655)..(2655)

<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2658)..(2658)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2661)..(2661)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2664)..(2664)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2673)..(2673)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2676)..(2676)
<223> n是a、c、g或t

[0100]

<220>
<221> misc_feature
<222> (2697)..(2697)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2700)..(2700)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2706)..(2706)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2709)..(2709)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2712)..(2712)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2718)..(2718)
<223> n是a、c、g或t

<220>

<220>
<221> misc_feature
<222> (2730)..(2730)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2739)..(2739)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2742)..(2742)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2748)..(2748)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2757)..(2757)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2760)..(2760)
<223> n是a、c、g或t

[0101]

<220>
<221> misc_feature
<222> (2763)..(2763)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2766)..(2766)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2769)..(2769)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2778)..(2778)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2781)..(2781)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2787)..(2787)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2790)..(2790)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2799)..(2799)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2802)..(2802)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2808)..(2808)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2811)..(2811)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2814)..(2814)
[0102] <223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2820)..(2820)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2823)..(2823)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2826)..(2826)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2841)..(2841)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2847)..(2847)
<223> n是a、c、g或t

<220>
<221> misc_feature

<220>
<221> misc_feature
<222> (2859)..(2859)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2862)..(2862)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2865)..(2865)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2868)..(2868)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2871)..(2871)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2889)..(2889)
<223> n是a、c、g或t

[0103]
<220>
<221> misc_feature
<222> (2895)..(2895)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2898)..(2898)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2901)..(2901)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2904)..(2904)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2907)..(2907)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2916)..(2916)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2928)..(2928)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2931)..(2931)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2937)..(2937)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2940)..(2940)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2946)..(2946)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2949)..(2949)
<223> n是a、c、g或t

[0104]

<220>
<221> misc_feature
<222> (2952)..(2952)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2958)..(2958)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2961)..(2961)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2964)..(2964)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2976)..(2976)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2979)..(2979)

<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2982)..(2982)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2985)..(2985)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (2988)..(2988)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3015)..(3015)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3024)..(3024)
<223> n是a、c、g或t

[0105]

<220>
<221> misc_feature
<222> (3027)..(3027)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3054)..(3054)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3060)..(3060)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3066)..(3066)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3072)..(3072)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3078)..(3078)
<223> n是a、c、g或t

<220>

<220>
<221> misc_feature
<222> (3081)..(3081)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3084)..(3084)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3087)..(3087)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3093)..(3093)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3096)..(3096)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3099)..(3099)
<223> n是a、c、g或t

[0106]

<220>
<221> misc_feature
<222> (3102)..(3102)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3105)..(3105)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3111)..(3111)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3114)..(3114)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3126)..(3126)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3153)..(3153)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3159)..(3159)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3162)..(3162)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3168)..(3168)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3174)..(3174)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3177)..(3177)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3180)..(3180)
[0107] <223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3201)..(3201)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3210)..(3210)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3222)..(3222)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3225)..(3225)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3228)..(3228)
<223> n是a、c、g或t

<220>
<221> misc_feature

<220>
<221> misc_feature
<222> (3231)..(3231)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3234)..(3234)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3240)..(3240)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3249)..(3249)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3258)..(3258)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3264)..(3264)
<223> n是a、c、g或t

[0108]

<220>
<221> misc_feature
<222> (3279)..(3279)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3282)..(3282)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3285)..(3285)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3288)..(3288)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (3291)..(3291)
<223> n是a、c、g或t

<400> 4
atggcnaart ayggngarca ygargcnwsn ccngayaayg gncaraayga rttywsngay 60

athathaarw snmgnwsnga ygarcayaay gaygtncara araaracntt yacnaartgg 120

athaaygcnm gnttywsnaa rwsnggnaar ccnccnatha aygayatgtt yacngayytn 180

aargayggnm gnaarytnyt ngayytnytn garggnytna cnggnacnws nytnccnaar 240

garmgnggnw snacnmgngt ncaygcnytn aayaaygtna aymgngtnyt ncargtnytn 300

caycaraaya aygtngaryt ngtnaayath ggnggnacng ayathgtnga yggnaaycay 360

aarytnacny tnggnytnyt ntggwsnath athytncayt ggcargtnaa rgaygtnatg 420

aargaygtna tgwsngayyt ncarcaracn aaywsngara arathytnyt nwsntgggtn 480

mgncaracna cnmgnccnta ywsncargtn aaygtnytna ayttyacnac nwsntggacn 540

gayggnytng cnttyaaygc ngtnytncay mgncayaarc cngayytntt ywsntgggay 600

aargtngtna aratgwsncc nathgarmgn ytngarcayg cnttywsnaa rgcncaracn 660

tayytnggna thgaraaryt nytngayccn gargaygtng cngtncaryt nccngayaar 720

aarwsnatha thatgtayyt nacnwsnytn ttygargtny tnccncarca rgtnacnath 780

gaygcnathm gngargtnga racnytnccn mgnaartaya araargartg ygargargar 840

gcnathaaya thcarwsnac ngcnccngar gargarcayg arwsnccnmg ngcngaracn 900

ccnwsnacng tnacngargt ngayatggay ytngaywsnt aycarathgc nytngargar 960

[0109] gtnytnacnt ggytnytnws ngcngargay acnttycarg arcargayga yathwsngay 1020

gaygtngarg argtnaarga ycarttygcn acncaygarg cnttyatgat ggarytnacn 1080

gcncaycarw snwsngtngg nwsngtnytn cargcnggna aycarytnat hacncarggn 1140

acnytnwsng aygargarga rttygarath cargarcara tgacnytnyt naaygcnmgn 1200

tgggargcny tnmgngtnga rwsnatggay mgncarwsnm gnytncayga ygtnytnatg 1260

garytncara araarcaryt ncarcarytn wsngcntgga thcargargc ngaracnacn 1320

gtnaaygtny tngtngaygc nwsncaymgn garaaygcny tncargayws nathytngcn 1380

mgngarytna arcarcarat gcargayath cargcngara thgaygcnca ycargtnaar 1440

gtnaaywsny tnacncayat ggtngtnath gtngaygara aywsnggnga rwsngcnacn 1500

gcnathytng argaycaryt ncaraarytn ggngarmgnt ggaaygayyt naargcnaar 1560

wsngcnwsna thmgngcnca yytngargcn wsngcngara artggaaymg nytnytnatg 1620

wsnytngarg arytnathaa rtggytnaay atgaargayg argarytnaa raarcaratg 1680

ccnathggng gngaygtncc ngcnytncar ytncartayg aycaytgyaa rgcnytnmgn 1740

mgngarytna argaraarga rtaywsngtn ytnaaygcng tngaycargc nmgngtntty 1800

ytngcngayc arccnathga rgcnccngar garccnmgnm gnaayytnca rwsnaaracn 1860

garytnacnc cngargarmg ngcncaraar athgcnaarg cnatgmgnaa rcarwsnwsn 1920

gargtnaarg araartggga rwsnytnaay gcngtnacnw snaaytggca raarcargtn 1980

gayaargcny tngaraaryt nmgngayytn carggngcna tggaygayyt ngaygcngay 2040

atgaargarg cngarwsngt nmgnaayggn tggaarccng tnggngayyt nytnathgay 2100

wsnytncarg aycayathga raarathatg gcnttymgng argarathgc nccnathaay 2160

ttyaargtna aracngtnaa ygayytnwsn wsncarytnw snccnytnga yytncayccn 2220

wsnytnaara tgwsnmgnca rytngaygay ytnaayatgm gntggaaryt nytncargtn 2280

wsngtngayg aymgnytnaa rcarytncar gargcncaym gngayttygg nccnwsnwsn 2340

carcayttyy tnwsnacnws ngtncarytn ccntggcarm gnwsnathws ncayaayaar 2400

gtnccntayt ayathaayca ycaracncar acnacntgyt gggaycaycc naaratgacn 2460

garytnttyc arwsnytngc ngayytnaay aaygtnmgnt tywsngcnta ymgnacngcn 2520

athaarathm gnmgnytnca raargcnytn tgyytngayy tnytngaryt nwsnacnacn 2580

aaygaratht tyaarcarca yaarytnaay caraaygayc arytnytnws ngtnccngay 2640

gtnathaayt gyytnacnac nacntaygay ggnytngarc aratgcayaa rgayytngtn 2700

[0110] aaygtnccny tntgygtnga yatgtgyytn aaytggytny tnaaygtnta ygayacnggn 2760

mgnacnggna arathmgngt ncarwsnytn aarathggny tnatgwsnyt nwsnaarggn 2820

ytnytngarg araartaymg ntayytntty aargargtng cnggnccnac ngaratgtgy 2880

gaycarmgnc arytnggnyt nytnytncay gaygcnathc arathccnmg ncarytnggn 2940

gargtngcng cnttyggngg nwsnaayath garccnwsng tnmgnwsntg yttycarcar 3000

aayaayaaya arccngarat hwsngtnaar garttyathg aytggatgca yytngarccn 3060

carwsnatgg tntggytncc ngtnytncay mgngtngcng cngcngarac ngcnaarcay 3120

cargcnaart gyaayathtg yaargartgy ccnathgtng gnttymgnta ymgnwsnytn 3180

aarcayttya aytaygaygt ntgycarwsn tgyttyttyw snggnmgnac ngcnaarggn 3240

cayaarytnc aytayccnat ggtngartay tgyathccna cnacnwsngg ngar 3294

<210> 5
<211> 1108
<212> PRT
<213> 人工序列

<220>
<223> 毫微-肌养蛋白相关蛋白-2 的氨基酸序列

<400> 5

```
Met Ala Lys Tyr Gly Glu His Glu Ala Ser Pro Asp Asn Gly Gln Asn
1               5                  10                  15

Glu Phe Ser Asp Ile Ile Lys Ser Arg Ser Asp Glu His Asn Asp Val
            20                  25                  30

Gln Lys Lys Thr Phe Thr Lys Trp Ile Asn Ala Arg Phe Ser Lys Ser
        35                  40                  45

Gly Lys Pro Pro Ile Asn Asp Met Phe Thr Asp Leu Lys Asp Gly Arg
    50                  55                  60

Lys Leu Leu Asp Leu Leu Glu Gly Leu Thr Gly Thr Ser Leu Pro Lys
65                  70                  75                  80

Glu Arg Gly Ser Thr Arg Val His Ala Leu Asn Asn Val Asn Arg Val
                85                  90                  95

Leu Gln Val Leu His Gln Asn Asn Val Glu Leu Val Asn Ile Gly Gly
            100                 105                 110
```

[0111]

```
Thr Asp Ile Val Asp Gly Asn His Lys Leu Thr Leu Gly Leu Leu Trp
        115                 120                 125

Ser Ile Ile Leu His Trp Gln Val Lys Asp Val Met Lys Asp Val Met
    130                 135                 140

Ser Asp Leu Gln Gln Thr Asn Ser Glu Lys Ile Leu Leu Ser Trp Val
145                 150                 155                 160

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

Thr Ser Trp Thr Asp Gly Leu Ala Phe Asn Ala Val Leu His Arg His
            180                 185                 190

Lys Pro Asp Leu Phe Ser Trp Asp Lys Val Val Lys Met Ser Pro Ile
        195                 200                 205

Glu Arg Leu Glu His Ala Phe Ser Lys Ala Gln Thr Tyr Leu Gly Ile
    210                 215                 220

Glu Lys Leu Leu Asp Pro Glu Asp Val Ala Val Gln Leu Pro Asp Lys
```

225 230 235 240

Lys Ser Ile Ile Met Tyr Leu Thr Ser Leu Phe Glu Val Leu Pro Gln
245 250 255

Gln Val Thr Ile Asp Ala Ile Arg Glu Val Glu Thr Leu Pro Arg Lys
260 265 270

Tyr Lys Lys Glu Cys Glu Glu Glu Ala Ile Asn Ile Gln Ser Thr Ala
275 280 285

Pro Glu Glu Glu His Glu Ser Pro Arg Ala Glu Thr Pro Ser Thr Val
290 295 300

Thr Glu Val Asp Met Asp Leu Asp Ser Tyr Gln Ile Ala Leu Glu Glu
305 310 315 320

Val Leu Thr Trp Leu Leu Ser Ala Glu Asp Thr Phe Gln Glu Gln Asp
325 330 335

[0112] Asp Ile Ser Asp Asp Val Glu Glu Val Lys Asp Gln Phe Ala Thr His
340 345 350

Glu Ala Phe Met Met Glu Leu Thr Ala His Gln Ser Ser Val Gly Ser
355 360 365

Val Leu Gln Ala Gly Asn Gln Leu Ile Thr Gln Gly Thr Leu Ser Asp
370 375 380

Glu Glu Glu Phe Glu Ile Gln Glu Gln Met Thr Leu Leu Asn Ala Arg
385 390 395 400

Trp Glu Ala Leu Arg Val Glu Ser Met Asp Arg Gln Ser Arg Leu His
405 410 415

Asp Val Leu Met Glu Leu Gln Lys Lys Gln Leu Gln Gln Leu Ser Ala
420 425 430

Trp Ile Gln Glu Ala Glu Thr Thr Val Asn Val Leu Val Asp Ala Ser
435 440 445

His Arg Glu Asn Ala Leu Gln Asp Ser Ile Leu Ala Arg Glu Leu Lys
450 455 460

Gln Gln Met Gln Asp Ile Gln Ala Glu Ile Asp Ala His Gln Val Lys
465 470 475 480

Val Asn Ser Leu Thr His Met Val Val Ile Val Asp Glu Asn Ser Gly
485 490 495

Glu Ser Ala Thr Ala Ile Leu Glu Asp Gln Leu Gln Lys Leu Gly Glu
500 505 510

Arg Trp Asn Asp Leu Lys Ala Lys Ser Ala Ser Ile Arg Ala His Leu
515 520 525

Glu Ala Ser Ala Glu Lys Trp Asn Arg Leu Leu Met Ser Leu Glu Glu
530 535 540

Leu Ile Lys Trp Leu Asn Met Lys Asp Glu Glu Leu Lys Lys Gln Met
545 550 555 560

Pro Ile Gly Gly Asp Val Pro Ala Leu Gln Leu Gln Tyr Asp His Cys
565 570 575

[0113] Lys Ala Leu Arg Arg Glu Leu Lys Glu Lys Glu Tyr Ser Val Leu Asn
580 585 590

Ala Val Asp Gln Ala Arg Val Phe Leu Ala Asp Gln Pro Ile Glu Ala
595 600 605

Pro Glu Glu Pro Arg Arg Asn Leu Gln Ser Lys Thr Glu Leu Thr Pro
610 615 620

Glu Glu Arg Ala Gln Lys Ile Ala Lys Ala Met Arg Lys Gln Ser Ser
625 630 635 640

Glu Val Lys Glu Lys Trp Glu Ser Leu Asn Ala Val Thr Ser Asn Trp
645 650 655

Gln Lys Gln Val Asp Lys Ala Leu Glu Lys Leu Arg Asp Leu Gln Gly
660 665 670

Ala Met Asp Asp Leu Asp Ala Asp Met Lys Glu Ala Glu Ser Val Arg
675 680 685

Asn Gly Trp Lys Pro Val Gly Asp Leu Leu Ile Asp Ser Leu Gln Asp

690 695 700

His Ile Glu Lys Ile Met Ala Phe Arg Glu Glu Ile Ala Pro Ile Asn
705 710 715 720

Phe Lys Val Lys Thr Val Asn Asp Leu Ser Ser Gln Leu Ser Pro Leu
725 730 735

Asp Leu His Pro Ser Leu Lys Met Ser Arg Gln Leu Asp Asp Leu Asn
740 745 750

Met Arg Trp Lys Leu Leu Gln Val Ser Val Asp Asp Arg Leu Lys Gln
755 760 765

Leu Gln Glu Ala His Arg Asp Phe Gly Pro Ser Ser Gln His Phe Leu
770 775 780

Ser Thr Ser Val Gln Leu Pro Trp Gln Arg Ser Ile Ser His Asn Lys
785 790 795 800

[0114] Val Pro Tyr Tyr Ile Asn His Gln Thr Gln Thr Thr Cys Trp Asp His
805 810 815

Pro Lys Met Thr Glu Leu Phe Gln Ser Leu Ala Asp Leu Asn Asn Val
820 825 830

Arg Phe Ser Ala Tyr Arg Thr Ala Ile Lys Ile Arg Arg Leu Gln Lys
835 840 845

Ala Leu Cys Leu Asp Leu Leu Glu Leu Ser Thr Thr Asn Glu Ile Phe
850 855 860

Lys Gln His Lys Leu Asn Gln Asn Asp Gln Leu Leu Ser Val Pro Asp
865 870 875 880

Val Ile Asn Cys Leu Thr Thr Thr Tyr Asp Gly Leu Glu Gln Met His
885 890 895

Lys Asp Leu Val Asn Val Pro Leu Cys Val Asp Met Cys Leu Asn Trp
900 905 910

Leu Leu Asn Val Tyr Asp Thr Gly Arg Thr Gly Lys Ile Arg Val Gln
915 920 925

Ser Leu Lys Ile Gly Leu Met Ser Leu Ser Lys Gly Leu Leu Glu Glu
930 935 940

Lys Tyr Arg Tyr Leu Phe Lys Glu Val Ala Gly Pro Thr Glu Met Cys
945 950 955 960

Asp Gln Arg Gln Leu Gly Leu Leu Leu His Asp Ala Ile Gln Ile Pro
965 970 975

Arg Gln Leu Gly Glu Val Ala Ala Phe Gly Gly Ser Asn Ile Glu Pro
980 985 990

Ser Val Arg Ser Cys Phe Gln Gln Asn Asn Asn Lys Pro Glu Ile Ser
995 1000 1005

Val Lys Glu Phe Ile Asp Trp Met His Leu Glu Pro Gln Ser Met
1010 1015 1020

Val Trp Leu Pro Val Leu His Arg Val Ala Ala Ala Glu Thr Ala
1025 1030 1035

[0115] Lys His Gln Ala Lys Cys Asn Ile Cys Lys Glu Cys Pro Ile Val
1040 1045 1050

Gly Phe Arg Tyr Arg Ser Leu Lys His Phe Asn Tyr Asp Val Cys
1055 1060 1065

Gln Ser Cys Phe Phe Ser Gly Arg Thr Ala Lys Gly His Lys Leu
1070 1075 1080

His Tyr Pro Met Val Glu Tyr Cys Ile Pro Thr Thr Ser Gly Glu
1085 1090 1095

Asp Val Arg Asp Phe Thr Lys Val Leu Lys
1100 1105

<210> 6
<211> 3327
<212> DNA
<213> 人工序列

<220>
<223> 编码毫微-肌养蛋白相关蛋白-2 蛋白的核酸序列

<400> 6

atggctaagt atggcgagca tgaggcttct cccgacaacg gccagaacga gttctccgac 60

atcatcaagt cccgctccga cgagcacaac gacgtgcaga agaaaacctt caccaagtgg 120

atcaacgcca ggttctccaa gtccggcaag cctcctatca acgacatgtt caccgacctg 180

aaggacggcc gcaaactgct ggatctgctg gaaggactga ccggcacctc tctgcctaaa 240

gaaaggggct ctactagggt gcacgccctg aacaatgtga acagggtgct ccaggtgctg 300

caccagaaca acgtcgagct ggtcaatatc ggcggcaccg acatcgtgga tggcaaccat 360

aagctgaccc tgggcctgct gtggtccatc attctgcact ggcaggtcaa agacgtgatg 420

aaggacgtga tgtccgacct ccagcagacc aactccgaga agatcctgct gtcttgggtc 480

cgccagacca ctaggcccta ttctcaggtc aacgtgctga acttcaccac ctcttggacc 540

gacggactgg ccttcaatgc tgtgctgcac aggcacaagc ccgacctgtt ctcttgggac 600

aaagtggtca agatgagccc catcgagagg ctggaacacg ccttctctaa ggcccagacc 660

tacctgggaa tcgagaagct gctggaccct gaggatgtgg cagtgcagct gcccgacaag 720

aaatccatca tcatgtacct gaccagcctg ttcgaggtgc tgcctcagca ggtcacaatt 780

gatgccatcc gcgaggtgga aaccctgcct cgcaagtaca agaaagagtg cgaggaagag 840

[0116] gccatcaaca tccagtccac tgctcccgag gaagaacacg agtctccaag ggccgagact 900

ccttccaccg tgactgaagt ggacatggac ctggactcct accagatcgc tctggaagag 960

gtgctgactt ggctgctgtc cgccgaggat actttccagg aacaggacga catctccgac 1020

gacgtcgagg aagtgaagga ccagttcgcc actcacgagg ctttcatgat ggaactgact 1080

gcccaccagt cctccgtggg aagtgtcctc caggctggaa accagctgat cacccaggga 1140

accctgtccg atgaggaaga gttcgagatt caggaacaga tgaccctgct gaacgccaga 1200

tgggaagctc tgagggtcga gtccatggat aggcagtcta ggctgcacga cgtgctcatg 1260

gaactccaga agaagcagct ccagcagctg tccgcttgga tccaggaagc cgagactacc 1320

gtcaatgtgc tggtggacgc ctctcacaga gagaacgccc tccaggattc catcctggct 1380

agggaactga agcagcagat gcaggacatc caggccgaga tcgacgctca ccaggtcaaa 1440

gtgaactccc tgactcacat ggtggtcatc gtggacgaga attccggcga gtctgccacc 1500

gctatcctgg aagatcagct ccagaaactg ggcgagcgct ggaacgatct gaaggctaag 1560

tccgcctcta tcagggccca tctggaagcc tctgccgaga agtggaatag gctgctgatg 1620

tccctggaag aactgatcaa gtggctgaac atgaaggatg aggaactcaa gaaacagatg 1680

cccatcggag gcgacgtgcc agctctccag ctccagtatg atcattgcaa ggctctgagg 1740

cgcgagctga aagagaaaga gtactccgtc ctgaacgccg tggatcaggc tagggtgttc 1800

ctggccgatc agcctattga ggctcctgag gaacctaggc ggaacctcca gtccaagacc 1860

gagctgactc ctgaagagag ggcccagaaa atcgccaagg ctatgcgcaa gcagtccagc 1920

gaagtcaaag agaagtggga gtccctgaat gccgtgacct ccaactggca gaaacaggtg 1980

gacaaggccc tggaaaagct gagggatctc cagggcgcta tggacgacct ggacgccgat 2040

atgaaggaag ctgagtccgt cagaaacggc tggaagccag tgggagatct gctgattgat 2100

agcctccagg accacattga gaagatcatg gccttcaggg aagagatcgc ccctatcaac 2160

ttcaaagtca agaccgtgaa cgacctgtcc agccagctga gccctctgga tctccatcca 2220

agcctgaaga tgagcaggca gctggatgac ctgaacatgc gctggaaact cctccaggtg 2280

tccgtggacg ataggctgaa acagctccag gaagctcaca gagacttcgg ccctagctct 2340

cagcacttcc tgtctacctc tgtgcagctc ccttggcagc gctccatctc tcacaacaag 2400

gtgccctact acatcaacca ccagactcag accacctgtt gggatcaccc caagatgact 2460

gagctgttcc agagcctggc cgacctcaac aatgtgcgct tctccgccta taggaccgcc 2520

atcaaaatca ggcgcctcca gaaagccctg tgcctggacc tgctcgaact gtccaccacc 2580

[0117] aacgagatct tcaagcagca caagctgaac cagaacgacc agctgctgtc tgtgcccgac 2640

gtgatcaatt gcctgaccac cacctacgac gggctcgaac agatgcacaa ggatctcgtg 2700

aatgtgcctc tgtgcgtgga catgtgcctg aactggctgc tgaatgtgta cgataccggc 2760

aggaccggca agatcagggt gcagtctctg aagatcggcc tgatgagcct gagcaagggg 2820

ctgctcgaag agaagtaccg ctacctgttc aaagaggtgg caggccctac cgagatgtgc 2880

gatcagagac agctgggact gctgctgcac gatgccattc agattcctcg ccagctggga 2940

gaagtggccg ccttcggagg atctaatatc gagccttccg tgcgctcctg ctttcagcag 3000

aacaacaaca agcctgagat ctccgtgaaa gaattcatcg actggatgca cctggaacct 3060

cagagcatgg tctggctgcc agtcctgcat agagtggctg ctgctgagac tgccaagcac 3120

caggccaagt gcaacatctg caaagagtgc cctatcgtcg gcttccgcta tcgctccctg 3180

aagcacttca actacgacgt gtgccagagt tgcttctttt ccggacgcac cgccaagggc 3240

cacaagctgc attaccctat ggtcgagtac tgcatcccta ccacatccgg cgaggatgtg 3300

cgcgatttca ccaaggtgct gaagtga 3327

<210> 7
<211> 1142
<212> PRT

〈213〉 人工序列

〈220〉
〈223〉 毫微-肌养蛋白相关蛋白-3 的氨基酸序列

〈400〉 7

```
Met Ala Lys Tyr Gly Glu His Glu Ala Ser Pro Asp Asn Gly Gln Asn
1               5                  10                  15

Glu Phe Ser Asp Ile Ile Lys Ser Arg Ser Asp Glu His Asn Asp Val
            20                  25                  30

Gln Lys Lys Thr Phe Thr Lys Trp Ile Asn Ala Arg Phe Ser Lys Ser
        35                  40                  45

Gly Lys Pro Pro Ile Asn Asp Met Phe Thr Asp Leu Lys Asp Gly Arg
    50                  55                  60

Lys Leu Leu Asp Leu Leu Glu Gly Leu Thr Gly Thr Ser Leu Pro Lys
65                  70                  75                  80

Glu Arg Gly Ser Thr Arg Val His Ala Leu Asn Asn Val Asn Arg Val
                85                  90                  95

Leu Gln Val Leu His Gln Asn Asn Val Glu Leu Val Asn Ile Gly Gly
            100                 105                 110

Thr Asp Ile Val Asp Gly Asn His Lys Leu Thr Leu Gly Leu Leu Trp
        115                 120                 125

Ser Ile Ile Leu His Trp Gln Val Lys Asp Val Met Lys Asp Val Met
    130                 135                 140

Ser Asp Leu Gln Gln Thr Asn Ser Glu Lys Ile Leu Leu Ser Trp Val
145                 150                 155                 160

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

Thr Ser Trp Thr Asp Gly Leu Ala Phe Asn Ala Val Leu His Arg His
            180                 185                 190

Lys Pro Asp Leu Phe Ser Trp Asp Lys Val Val Lys Met Ser Pro Ile
        195                 200                 205
```

[0118]

Glu Arg Leu Glu His Ala Phe Ser Lys Ala Gln Thr Tyr Leu Gly Ile
210 215 220

Glu Lys Leu Leu Asp Pro Glu Asp Val Ala Val Gln Leu Pro Asp Lys
225 230 235 240

Lys Ser Ile Ile Met Tyr Leu Thr Ser Leu Phe Glu Val Leu Pro Gln
245 250 255

Gln Val Thr Ile Asp Ala Ile Arg Glu Val Glu Thr Leu Pro Arg Lys
260 265 270

Tyr Lys Lys Glu Cys Glu Glu Glu Ala Ile Asn Ile Gln Ser Thr Ala
275 280 285

Pro Glu Glu Glu His Glu Ser Pro Arg Ala Glu Thr Pro Ser Thr Val
290 295 300

Thr Glu Val Asp Met Asp Leu Asp Ser Tyr Gln Ile Ala Leu Glu Glu
305 310 315 320

[0119] Val Leu Thr Trp Leu Leu Ser Ala Glu Asp Thr Phe Gln Glu Gln Asp
325 330 335

Asp Ile Ser Asp Asp Val Glu Glu Val Lys Asp Gln Phe Ala Thr His
340 345 350

Glu Ala Phe Met Met Glu Leu Thr Ala His Gln Ser Ser Val Gly Ser
355 360 365

Val Leu Gln Ala Gly Asn Gln Leu Ile Thr Gln Gly Thr Leu Ser Asp
370 375 380

Glu Glu Glu Phe Glu Ile Gln Glu Gln Met Thr Leu Leu Asn Ala Arg
385 390 395 400

Trp Glu Ala Leu Arg Val Glu Ser Met Asp Arg Gln Ser Arg Leu His
405 410 415

Asp Val Leu Met Glu Leu Gln Lys Lys Gln Leu Gln Gln Leu Ser Ala
420 425 430

Trp Ile Gln Glu Ala Glu Thr Thr Val Asn Val Leu Val Asp Ala Ser

435 440 445

His Arg Glu Asn Ala Leu Gln Asp Ser Ile Leu Ala Arg Glu Leu Lys
450 455 460

Gln Gln Met Gln Asp Ile Gln Ala Glu Ile Asp Ala His Gln Val Lys
465 470 475 480

Val Asn Ser Leu Thr His Met Val Val Ile Val Asp Glu Asn Ser Gly
485 490 495

Glu Ser Ala Thr Ala Ile Leu Glu Asp Gln Leu Gln Lys Leu Gly Glu
500 505 510

Arg Trp Asn Asp Leu Lys Ala Lys Ser Ala Ser Ile Arg Ala His Leu
515 520 525

Glu Ala Ser Ala Glu Lys Trp Asn Arg Leu Leu Met Ser Leu Glu Glu
530 535 540

[0120] Leu Ile Lys Trp Leu Asn Met Lys Asp Glu Glu Leu Lys Lys Gln Met
545 550 555 560

Pro Ile Gly Gly Asp Val Pro Ala Leu Gln Leu Gln Tyr Asp His Cys
565 570 575

Lys Ala Leu Arg Arg Glu Leu Lys Glu Lys Glu Tyr Ser Val Leu Asn
580 585 590

Ala Val Asp Gln Ala Arg Val Phe Leu Ala Asp Gln Pro Ile Glu Ala
595 600 605

Pro Glu Glu Pro Arg Arg Asn Leu Gln Ser Lys Thr Glu Leu Thr Pro
610 615 620

Glu Glu Arg Ala Gln Lys Ile Ala Lys Ala Met Arg Lys Gln Ser Ser
625 630 635 640

Glu Val Lys Glu Lys Trp Glu Ser Leu Asn Ala Val Thr Ser Asn Trp
645 650 655

Gln Lys Gln Val Asp Lys Ala Leu Glu Lys Leu Arg Asp Leu Gln Gly
660 665 670

Ala Met Asp Asp Leu Asp Ala Asp Met Lys Glu Ala Glu Ser Val Arg
675 680 685

Asn Gly Trp Lys Pro Val Gly Asp Leu Leu Ile Asp Ser Leu Gln Asp
690 695 700

His Ile Glu Lys Ile Met Ala Phe Arg Glu Glu Ile Ala Pro Ile Asn
705 710 715 720

Phe Lys Val Lys Thr Val Asn Asp Leu Ser Ser Gln Leu Ser Pro Leu
725 730 735

Asp Leu His Pro Ser Leu Lys Met Ser Arg Gln Leu Asp Asp Leu Asn
740 745 750

Met Arg Trp Lys Leu Leu Gln Val Ser Val Asp Asp Arg Leu Lys Gln
755 760 765

Leu Gln Glu Ala His Arg Asp Phe Gly Pro Ser Ser Gln His Phe Leu
770 775 780

[0121] Ser Thr Ser Val Gln Leu Pro Trp Gln Arg Ser Ile Ser His Asn Lys
785 790 795 800

Val Pro Tyr Tyr Ile Asn His Gln Thr Gln Thr Thr Cys Trp Asp His
805 810 815

Pro Lys Met Thr Glu Leu Phe Gln Ser Leu Ala Asp Leu Asn Asn Val
820 825 830

Arg Phe Ser Ala Tyr Arg Thr Ala Ile Lys Ile Arg Arg Leu Gln Lys
835 840 845

Ala Leu Cys Leu Asp Leu Leu Glu Leu Ser Thr Thr Asn Glu Ile Phe
850 855 860

Lys Gln His Lys Leu Asn Gln Asn Asp Gln Leu Leu Ser Val Pro Asp
865 870 875 880

Val Ile Asn Cys Leu Thr Thr Thr Tyr Asp Gly Leu Glu Gln Met His
885 890 895

Lys Asp Leu Val Asn Val Pro Leu Cys Val Asp Met Cys Leu Asn Trp

900 905 910

Leu Leu Asn Val Tyr Asp Thr Gly Arg Thr Gly Lys Ile Arg Val Gln
915 920 925

Ser Leu Lys Ile Gly Leu Met Ser Leu Ser Lys Gly Leu Leu Glu Glu
930 935 940

Lys Tyr Arg Tyr Leu Phe Lys Glu Val Ala Gly Pro Thr Glu Met Cys
945 950 955 960

Asp Gln Arg Gln Leu Gly Leu Leu Leu His Asp Ala Ile Gln Ile Pro
965 970 975

Arg Gln Leu Gly Glu Val Ala Ala Phe Gly Gly Ser Asn Ile Glu Pro
980 985 990

Ser Val Arg Ser Cys Phe Gln Gln Asn Asn Asn Lys Pro Glu Ile Ser
995 1000 1005

[0122]
Val Lys Glu Phe Ile Asp Trp Met His Leu Glu Pro Gln Ser Met
1010 1015 1020

Val Trp Leu Pro Val Leu His Arg Val Ala Ala Ala Glu Thr Ala
1025 1030 1035

Lys His Gln Ala Lys Cys Asn Ile Cys Lys Glu Cys Pro Ile Val
1040 1045 1050

Gly Phe Arg Tyr Arg Ser Leu Lys His Phe Asn Tyr Asp Val Cys
1055 1060 1065

Gln Ser Cys Phe Phe Ser Gly Arg Thr Ala Lys Gly His Lys Leu
1070 1075 1080

His Tyr Pro Met Val Glu Tyr Cys Ile Pro Thr Thr Ser Gly Glu
1085 1090 1095

Asp Val Arg Asp Phe Thr Lys Val Leu Lys Asn Lys Phe Arg Ser
1100 1105 1110

Lys Lys Tyr Phe Ala Lys His Pro Arg Leu Gly Tyr Leu Pro Val
1115 1120 1125

Gln Thr Val Leu Glu Gly Asp Asn Leu Glu Thr Pro Ile Thr
1130 1135 1140

<210> 8
<211> 3429
<212> DNA
<213> 人工序列

<220>
<223> 编码毫微-肌养蛋白相关蛋白-3 蛋白的核酸序列

<400> 8

[0123]

atggctaagt atggcgagca tgaggcttct cccgacaacg gccagaacga gttctccgac 60

atcatcaagt cccgctccga cgagcacaac gacgtgcaga agaaaacctt caccaagtgg 120

atcaacgcca ggttctccaa gtccggcaag cctcctatca acgacatgtt caccgacctg 180

aaggacggcc gcaaactgct ggatctgctg gaaggactga ccggcacctc tctgcctaaa 240

gaaaggggct ctactagggt gcacgccctg aacaatgtga acagggtgct ccaggtgctg 300

caccagaaca acgtcgagct ggtcaatatc ggcggcaccg acatcgtgga tggcaaccat 360

aagctgaccc tgggcctgct gtggtccatc attctgcact ggcaggtcaa agacgtgatg 420

aaggacgtga tgtccgacct ccagcagacc aactccgaga agatcctgct gtcttgggtc 480

cgccagacca ctaggcccta ttctcaggtc aacgtgctga acttcaccac ctcttggacc 540

gacggactgg ccttcaatgc tgtgctgcac aggcacaagc ccgacctgtt ctcttgggac 600

aaagtggtca agatgagccc catcgagagg ctggaacacg ccttctctaa ggcccagacc 660

tacctgggaa tcgagaagct gctggaccct gaggatgtgg cagtgcagct gcccgacaag 720

aaatccatca tcatgtacct gaccagcctg ttcgaggtgc tgcctcagca ggtcacaatt 780

gatgccatcc gcgaggtgga aaccctgcct cgcaagtaca agaaagagtg cgaggaagag 840

gccatcaaca tccagtccac tgctcccgag gaagaacacg agtctccaag ggccgagact 900

ccttccaccg tgactgaagt ggacatggac ctggactcct accagatcgc tctggaagag 960

gtgctgactt ggctgctgtc cgccgaggat actttccagg aacaggacga catctccgac 1020

gacgtcgagg aagtgaagga ccagttcgcc actcacgagg ctttcatgat ggaactgact 1080

gcccaccagt cctccgtggg aagtgtcctc caggctggaa accagctgat cacccaggga 1140

accctgtccg atgaggaaga gttcgagatt caggaacaga tgaccctgct gaacgccaga 1200

tgggaagctc tgagggtcga gtccatggat aggcagtcta ggctgcacga cgtgctcatg 1260

gaactccaga agaagcagct ccagcagctg tccgcttgga tccaggaagc cgagactacc 1320

gtcaatgtgc tggtggacgc ctctcacaga gagaacgccc tccaggattc catcctggct 1380

agggaactga agcagcagat gcaggacatc caggccgaga tcgacgctca ccaggtcaaa 1440

gtgaactccc tgactcacat ggtggtcatc gtggacgaga attccggcga gtctgccacc 1500

gctatcctgg aagatcagct ccagaaactg ggcgagcgct ggaacgatct gaaggctaag 1560

tccgcctcta tcagggccca tctggaagcc tctgccgaga agtggaatag gctgctgatg 1620

tccctggaag aactgatcaa gtggctgaac atgaaggatg aggaactcaa gaaacagatg 1680

cccatcggag gcgacgtgcc agctctccag ctccagtatg atcattgcaa ggctctgagg 1740

cgcgagctga aagagaaaga gtactccgtc ctgaacgccg tggatcaggc tagggtgttc 1800

ctggccgatc agcctattga ggctcctgag gaacctaggc ggaacctcca gtccaagacc 1860

gagctgactc ctgaagagag ggcccagaaa atcgccaagg ctatgcgcaa gcagtccagc 1920

gaagtcaaag agaagtggga gtccctgaat gccgtgacct ccaactggca gaaacaggtg 1980

gacaaggccc tggaaaagct gagggatctc cagggcgcta tggacgacct ggacgccgat 2040

atgaaggaag ctgagtccgt cagaaacggc tggaagccag tgggagatct gctgattgat 2100

agcctccagg accacattga gaagatcatg gccttcaggg aagagatcgc ccctatcaac 2160

[0124] ttcaaagtca agaccgtgaa cgacctgtcc agccagctga gccctctgga tctccatcca 2220

agcctgaaga tgagcaggca gctggatgac ctgaacatgc gctggaaact cctccaggtg 2280

tccgtggacg ataggctgaa acagctccag gaagctcaca gagacttcgg ccctagctct 2340

cagcacttcc tgtctacctc tgtgcagctc ccttggcagc gctccatctc tcacaacaag 2400

gtgccctact acatcaacca ccagactcag accacctgtt gggatcaccc caagatgact 2460

gagctgttcc agagcctggc cgacctcaac aatgtgcgct tctccgccta taggaccgcc 2520

atcaaaatca ggcgcctcca gaaagccctg tgcctggacc tgctcgaact gtccaccacc 2580

aacgagatct tcaagcagca caagctgaac cagaacgacc agctgctgtc tgtgcccgac 2640

gtgatcaatt gcctgaccac cacctacgac gggctcgaac agatgcacaa ggatctcgtg 2700

aatgtgcctc tgtgcgtgga catgtgcctg aactggctgc tgaatgtgta cgataccggc 2760

aggaccggca agatcagggt gcagtctctg aagatcggcc tgatgagcct gagcaagggg 2820

ctgctcgaag agaagtaccg ctacctgttc aaagaggtgg caggccctac cgagatgtgc 2880

gatcagagac agctgggact gctgctgcac gatgccattc agattcctcg ccagctggga 2940

gaagtggccg ccttcggagg atctaatatc gagccttccg tgcgctcctg ctttcagcag 3000

aacaacaaca agcctgagat ctccgtgaaa gaattcatcg actggatgca cctggaacct 3060

cagagcatgg tctggctgcc agtcctgcat agagtggctg ctgctgagac tgccaagcac 3120

caggccaagt gcaacatctg caaagagtgc cctatcgtcg gcttccgcta tcgctccctg 3180

aagcacttca actacgacgt gtgccagagt tgcttctttt ccggacgcac cgccaagggc 3240

cacaagctgc attaccctat ggtcgagtac tgcatcccta ccacatccgg cgaggatgtg 3300

cgcgatttca ccaaggtgct gaagaacaag ttccgctcca agaagtattt cgctaagcac 3360

cccaggctgg gctatctgcc tgtccagact gtcctggaag gcgacaacct ggaaacccct 3420

atcacctga 3429

<210> 9
<211> 736
<212> PRT
<213> 人工序列

<220>
<223> 具有 GenBank 检索: AAS99264 的氨基酸序列 4

<400> 9

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1 5 10 15

[0125]

Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro
20 25 30

Lys Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg

130 135 140

Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly
145 150 155 160

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
180 185 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
195 200 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240

[0126] Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
260 265 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
275 280 285

Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
290 295 300

Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
305 310 315 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
325 330 335

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
340 345 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
355 360 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
370 375 380

Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385 390 395 400

Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu
405 410 415

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
420 425 430

Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser
435 440 445

Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
450 455 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465 470 475 480

[0127] Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
485 490 495

Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
500 505 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
530 535 540

Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545 550 555 560

Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser
565 570 575

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln
580 585 590

Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln

595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
610 615 620

Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met
625 630 635 640

Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
645 650 655

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr
660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700

Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val
[0128] 705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730 735

<210> 10
<211> 2211
<212> DNA
<213> 人工序列

<220>
<223> AAV9 衣壳 vp1 的编码序列

<400> 10
atggctgccg atggttatct tccagattgg ctcgaggaca accttagtga aggaattcgc 60

gagtggtggg ctttgaaacc tggagcccct caacccaagg caaatcaaca acatcaagac 120

aacgctcgag gtcttgtgct tccgggttac aaataccttg gacccggcaa cggactcgac 180

aaggggggagc cggtcaacgc agcagacgcg gcggccctcg agcacgacaa ggcctacgac 240

cagcagctca aggccggaga caacccgtac ctcaagtaca accacgccga cgccgagttc 300

caggagcggc tcaaagaaga tacgtctttt gggggcaacc tcgggcgagc agtcttccag 360

gccaaaaaga ggcttcttga acctcttggt ctggttgagg aagcggctaa gacggctcct 420

ggaaagaaga ggcctgtaga gcagtctcct caggaaccgg actcctccgc gggtattggc 480

aaatcgggtg cacagcccgc taaaaagaga ctcaatttcg gtcagactgg cgacacagag 540

tcagtcccag accctcaacc aatcggagaa cctcccgcag ccccctcagg tgtgggatct 600

cttacaatgg cttcaggtgg tggcgcacca gtggcagaca ataacgaagg tgccgatgga 660

gtgggtagtt cctcgggaaa ttggcattgc gattcccaat ggctggggga cagagtcatc 720

accaccagca cccgaacctg ggccctgccc acctacaaca atcacctcta caagcaaatc 780

tccaacagca catctggagg atcttcaaat gacaacgcct acttcggcta cagcaccccc 840

tgggggtatt ttgacttcaa cagattccac tgccacttct caccacgtga ctggcagcga 900

ctcatcaaca acaactgggg attccggcct aagcgactca acttcaagct cttcaacatt 960

caggtcaaag aggttacgga caacaatgga gtcaagacca tcgccaataa ccttaccagc 1020

acggtccagg tcttcacgga ctcagactat cagctcccgt acgtgctcgg gtcggctcac 1080

gagggctgcc tcccgccgtt cccagcggac gttttcatga ttcctcagta cgggtatctg 1140

acgcttaatg atggaagcca ggccgtgggt cgttcgtcct tttactgcct ggaatatttc 1200

ccgtcgcaaa tgctaagaac gggtaacaac ttccagttca gctacgagtt tgagaacgta 1260

[0129] cctttccata gcagctacgc tcacagccaa agcctggacc gactaatgaa tccactcatc 1320

gaccaatact tgtactatct ctcaaagact attaacggtt ctggacagaa tcaacaaacg 1380

ctaaaattca gtgtggccgg acccagcaac atggctgtcc agggaagaaa ctacatacct 1440

ggacccagct accgacaaca acgtgtctca accactgtga ctcaaaacaa caacagcgaa 1500

tttgcttggc ctggagcttc ttcttgggct ctcaatggac gtaatagctt gatgaatcct 1560

ggacctgcta tggccagcca caaagaagga gaggaccgtt tctttccttt gtctggatct 1620

ttaatttttg gcaaacaagg aactggaaga gacaacgtgg atgcggacaa agtcatgata 1680

accaacgaag aagaaattaa aactactaac ccggtagcaa cggagtccta tggacaagtg 1740

gccacaaacc accagagtgc ccaagcacag gcgcagaccg gctgggttca aaaccaagga 1800

atacttccgg gtatggtttg gcaggacaga gatgtgtacc tgcaaggacc catttgggcc 1860

aaaattcctc acacggacgg caactttcac ccttctccgc tgatgggagg gtttggaatg 1920

aagcacccgc ctcctcagat cctcatcaaa aacacacctg tacctgcgga tcctccaacg 1980

gccttcaaca aggacaagct gaactctttc atcacccagt attctactgg ccaagtcagc 2040

gtggagatcg agtgggagct gcagaaggaa aacagcaagc gctggaaccc ggagatccag 2100

tacacttcca actattacaa gtctaataat gttgaatttg ctgttaatac tgaaggtgta 2160

tatagtgaac cccgccccat tggcaccaga tacctgactc gtaatctgta a 2211

<210> 11
<211> 2211
<212> DNA
<213> 人工序列

<220>
<223> hu31 vp1 的编码序列

<400> 11
atggctgccg atggttatct tccagattgg ctcgaggaca accttagtga aggaattcgc 60

gagtggtggg ctttgaaacc tggagcccct caacccaagg caaatcaaca acatcaagac 120

aacgctcgag gtcttgtgct tccgggttac aaataccttg gacccggcaa cggactcgac 180

aaggggagc cggtcaacgc agcagacgcg gcggccctcg agcacgacaa ggcctacgac 240

cagcagctca aggccggaga caacccgtac ctcaagtaca accacgccga cgccgagttc 300

caggagcggc tcaaagaaga tacgtctttt gggggcaacc tcgggcgagc agtcttccag 360

gccaaaaaga ggcttcttga acctcttggt ctggttgagg aagcggctaa gacggctcct 420

ggaaagaaga ggcctgtaga gcagtctcct caggaaccgg actcctccgc gggtattggc 480

[0130] aaatcgggtg cacagcccgc taaaaagaga ctcaatttcg gtcagactgg cgacacagag 540

tcagtcccag accctcaacc aatcggagaa cctcccgcag ccccctcagg tgtgggatct 600

cttacaatgg cttcaggtgg tggcgcacca gtggcagaca ataacgaagg tgccgatgga 660

gtgggtagtt cctcgggaaa ttggcattgc gattcccaat ggctgggga cagagtcatc 720

accaccagca cccgaacctg ggccctgccc acctacaaca atcacctcta caagcaaatc 780

tccaacagca catctggagg atcttcaaat gacaacgcct acttcggcta cagcaccccc 840

tgggggtatt ttgacttcaa cagattccac tgccacttct caccacgtga ctggcagcga 900

ctcatcaaca acaactgggg attccggcct aagcgactca acttcaagct cttcaacatt 960

caggtcaaag aggttacgga caacaatgga gtcaagacca tcgccaataa ccttaccagc 1020

acggtccagg tcttcacgga ctcagactat cagctcccgt acgtgctcgg gtcggctcac 1080

gagggctgcc tcccgccgtt cccagcggac gttttcatga ttcctcagta cgggtatctg 1140

acgcttaatg atggaagcca ggccgtgggt cgttcgtcct tttactgcct ggaatatttc 1200

ccgtcgcaaa tgctaagaac gggtaacaac ttccagttca gctacgagtt tgagaacgta 1260

cctttccata gcagctacgc tcacagccaa agcctggacc gactaatgaa tccactcatc 1320

gaccaatact tgtactatct ctcaaagact attaacggtt ctggacagaa tcaacaaacg 1380

ctaaaattca gtgtggccgg acccagcaac atggctgtcc agggaagaaa ctacatacct 1440

ggacccagct accgacaaca acgtgtctca accactgtga ctcaaaacaa caacagcgaa 1500

tttgcttggc ctggagcttc ttcttgggct ctcaatggac gtaatagctt gatgaatcct 1560

ggacctgcta tggccagcca caaagaagga gaggaccgtt tctttccttt gtctggatct 1620

ttaatttttg gcaaacaagg aactggaaga gacaacgtgg atgcggacaa agtcatgata 1680

accaacgaag aagaaattaa aactactaac ccggtagcaa cggagtccta tggacaagtg 1740

gccacaaacc accagagtgc ccaagcacag gcgcagaccg gctgggttca aaaccaagga 1800

atacttccgg gtatggtttg gcaggacaga gatgtgtacc tgcaaggacc catttgggcc 1860

aaaattcctc acacggacgg caactttcac ccttctccgc tgatgggagg gtttggaatg 1920

aagcacccgc ctcctcagat cctcatcaaa aacacacctg tacctgcgga tcctccaacg 1980

gccttcaaca aggacaagct gaactctttc atcacccagt attctactgg ccaagtcagc 2040

gtggagatcg agtgggagct gcagaaggaa aacagcaagc gctggaaccc ggagatccag 2100

tacacttcca actattacaa gtctaataat gttgaatttg ctgttaatac tgaaggtgta 2160

tatagtgaac cccgccccat tggcaccaga tacctgactc gtaatctgta a 2211

[0131]

<210> 12
<211> 2211
<212> DNA
<213> 人工序列

<220>
<223> hu32 vp1 的编码序列

<400> 12
atggctgccg atggttatct tccagattgg ctcgaggaca ctctctctga aggaataaga 60

cagtggtgga agctcaaacc tggcccacca ccaccaaagc ccgcagagcg gcataaggac 120

gacagcaggg gtcttgtgct tcctgggtac aagtacctcg gacccggcaa cggactcgac 180

aaggggagc cggtcaacgc agcagacgcg gcggccctcg agcacgacaa ggcctacgac 240

cagcagctca aggccggaga caacccgtac ctcaagtaca accacgccga cgccgagttc 300

caggagcggc tcaaagaaga tacgtctttt gggggcaacc tcgggcgagc agtcttccag 360

gccaaaaaga ggcttcttga acctcttggt ctggttgagg aagcggctaa gacggctcct 420

ggaaagaaga ggcctgtaga gcagtctcct caggaaccgg actcctccgc gggtattggc 480

aaatcgggtt cacagcccgc taaaaagaaa ctcaatttcg gtcagactgg cgacacagag 540

tcagtccccg accctcaacc aatcggagaa cctcccgcag ccccctcagg tgtgggatct 600

cttacaatgg cttcaggtgg tggcgcacca gtggcagaca ataacgaagg tgccgatgga 660

gtgggtagtt cctcgggaaa ttggcattgc gattcccaat ggctggggga cagagtcatc 720

accaccagca cccgaacctg ggccctgccc acctacaaca atcacctcta caagcaaatc 780

tccaacagca catctggagg atcttcaaat gacaacgcct acttcggcta cagcaccccc 840

tgggggtatt ttgacttcaa cagattccac tgccacttct caccacgtga ctggcagcga 900

ctcatcaaca acaactgggg attccggcct aagcgactca acttcaagct cttcaacatt 960

caggtcaaag aggttacgga caacaatgga gtcaagacca tcgccaataa ccttaccagc 1020

acggtccagg tcttcacgga ctcagactat cagctcccgt acgtgctcgg gtcggctcac 1080

gagggctgcc tcccgccgtt cccagcggac gttttcatga ttcctcagta cgggtatctg 1140

acgcttaatg atgggagcca ggccgtgggt cgttcgtcct tttactgcct ggaatatttc 1200

ccgtcgcaaa tgctaagaac gggtaacaac ttccagttca gctacgagtt tgagaacgta 1260

cctttccata gcagctacgc tcacagccaa agcctggacc gactaatgaa tccactcatc 1320

gaccaatact tgtactatct ctcaaagact attaacggtt ctggacagaa tcaacaaacg 1380

ctaaaattca gcgtggccgg acccagcaac atggctgtcc agggaagaaa ctacatacct 1440

[0132] ggacccagct accgacaaca acgtgtctca accactgtga ctcaaaacaa caacagcgaa 1500

tttgcttggc ctggagcttc ttcttgggct ctcaatggac gtaatagctt gatgaatcct 1560

ggacctgcta tggccagcca caaagaagga gaggaccgtt tctttccttt gtctggatct 1620

ttaatttttg gcaaacaagg aactggaaga gacaacgtgg atgcggacaa agtcatgata 1680

accaacgaag aagaaattaa aactactaac ccggtagcaa cggagtccta tggacaagtg 1740

gccacaaacc accagagtgc ccaagcacag gcgcagaccg gctgggttca aaaccaagga 1800

atacttccgg gtatggtttg gcaggacaga gatgtgtacc tgcaaggacc catttgggcc 1860

aaaattcctc acacggacgg caactttcac ccttctccgc taatgggagg gtttggaatg 1920

aagcacccgc ctcctcagat cctcatcaaa aacacacctg tacctgcgga tcctccaacg 1980

gctttcaata aggacaagct gaactctttc atcacccagt attctactgg ccaagtcagc 2040

gtggagattg agtgggagct gcagaaggaa aacagcaagc gctggaaccc ggagatccag 2100

tacacttcca actattacaa gtctaataat gttgaatttg ctgttaatac tgaaggtgta 2160

tatagtgaac cccgccccat tggcaccaga tacctgactc gtaatctgta a 2211

<210> 13
<211> 1131
<212> PRT

〈213〉 人工序列

〈220〉
〈223〉 毫微-肌养蛋白的氨基酸序列

〈400〉 13

Met Leu Trp Trp Glu Glu Val Glu Asp Cys Tyr Glu Arg Glu Asp Val
1 5 10 15

Gln Lys Lys Thr Phe Thr Lys Trp Val Asn Ala Gln Phe Ser Lys Phe
20 25 30

Gly Lys Gln His Ile Glu Asn Leu Phe Ser Asp Leu Gln Asp Gly Arg
35 40 45

Arg Leu Leu Asp Leu Leu Glu Gly Leu Thr Gly Gln Lys Leu Pro Lys
50 55 60

Glu Lys Gly Ser Thr Arg Val His Ala Leu Asn Asn Val Asn Lys Ala
65 70 75 80

[0133] Leu Arg Val Leu Gln Asn Asn Asn Val Asp Leu Val Asn Ile Gly Ser
85 90 95

Thr Asp Ile Val Asp Gly Asn His Lys Leu Thr Leu Gly Leu Ile Trp
100 105 110

Asn Ile Ile Leu His Trp Gln Val Lys Asn Val Met Lys Asn Ile Met
115 120 125

Ala Gly Leu Gln Gln Thr Asn Ser Glu Lys Ile Leu Leu Ser Trp Val
130 135 140

Arg Gln Ser Thr Arg Asn Tyr Pro Gln Val Asn Val Ile Asn Phe Thr
145 150 155 160

Thr Ser Trp Ser Asp Gly Leu Ala Leu Asn Ala Leu Ile His Ser His
165 170 175

Arg Pro Asp Leu Phe Asp Trp Asn Ser Val Val Cys Gln Gln Ser Ala
180 185 190

Thr Gln Arg Leu Glu His Ala Phe Asn Ile Ala Arg Tyr Gln Leu Gly
195 200 205

Ile Glu Lys Leu Leu Asp Pro Glu Asp Val Asp Thr Thr Tyr Pro Asp
210 215 220

Lys Lys Ser Ile Leu Met Tyr Ile Thr Ser Leu Phe Gln Val Leu Pro
225 230 235 240

Gln Gln Val Ser Ile Glu Ala Ile Gln Glu Val Glu Met Leu Pro Arg
245 250 255

Pro Pro Lys Val Thr Lys Glu Glu His Phe Gln Leu His His Gln Met
260 265 270

His Tyr Ser Gln Gln Ile Thr Val Ser Leu Ala Gln Gly Tyr Glu Arg
275 280 285

Thr Ser Ser Pro Lys Pro Arg Phe Lys Ser Tyr Ala Tyr Thr Gln Ala
290 295 300

Ala Tyr Val Thr Thr Ser Asp Pro Thr Arg Ser Pro Phe Pro Ser Gln
305 310 315 320

[0134] His Leu Glu Ala Pro Glu Asp Lys Ser Phe Gly Ser Ser Leu Met Glu
325 330 335

Ser Glu Val Asn Leu Asp Arg Tyr Gln Thr Ala Leu Glu Glu Val Leu
340 345 350

Ser Trp Leu Leu Ser Ala Glu Asp Thr Leu Gln Ala Gln Gly Glu Ile
355 360 365

Ser Asn Asp Val Glu Val Val Lys Asp Gln Phe His Thr His Glu Gly
370 375 380

Tyr Met Met Asp Leu Thr Ala His Gln Gly Arg Val Gly Asn Ile Leu
385 390 395 400

Gln Leu Gly Ser Lys Leu Ile Gly Thr Gly Lys Leu Ser Glu Asp Glu
405 410 415

Glu Thr Glu Val Gln Glu Gln Met Asn Leu Leu Asn Ser Arg Trp Glu
420 425 430

Cys Leu Arg Val Ala Ser Met Glu Lys Gln Ser Asn Leu His Arg Val

435 440 445

Leu Met Asp Leu Gln Asn Gln Lys Leu Lys Glu Leu Asn Asp Trp Leu
450 455 460

Thr Glu Ala Glu Thr Thr Ala Asn Val Leu Gln Asp Ala Thr Arg Lys
465 470 475 480

Glu Arg Leu Leu Glu Asp Ser Lys Gly Val Lys Glu Leu Met Lys Gln
485 490 495

Trp Gln Asp Leu Gln Gly Asp Leu Glu Gln Glu Gln Val Arg Val Asn
500 505 510

Ser Leu Thr His Met Val Val Val Val Asp Glu Ser Ser Gly Asp His
515 520 525

Ala Thr Ala Ala Leu Glu Glu Gln Leu Lys Val Leu Gly Asp Arg Trp
530 535 540

[0135] Ala Asn Ile Cys Arg Trp Ser Glu Leu Arg Lys Lys Ser Leu Asn Ile
545 550 555 560

Arg Ser His Leu Glu Ala Ser Ser Asp Gln Trp Lys Arg Leu His Leu
565 570 575

Ser Leu Gln Glu Leu Leu Val Trp Leu Gln Leu Lys Asp Asp Glu Leu
580 585 590

Ser Arg Gln Ala Pro Ile Gly Gly Asp Phe Pro Ala Val Gln Lys Gln
595 600 605

Asn Asp Val His Arg Ala Phe Lys Arg Glu Leu Lys Thr Lys Glu Pro
610 615 620

Val Ile Met Ser Thr Leu Glu Thr Val Arg Ile Phe Leu Thr Glu Gln
625 630 635 640

Pro Leu Glu Gly Leu Glu Lys Leu Tyr Gln Glu Pro Arg Glu Leu Pro
645 650 655

Pro Glu Glu Arg Ala Gln Asn Val Thr Arg Leu Leu Arg Lys Gln Ala
660 665 670

Glu Glu Val Asn Thr Glu Trp Glu Lys Leu Asn Leu His Ser Ala Asp
675 680 685

Trp Gln Arg Lys Ile Asp Glu Thr Leu Glu Arg Leu Gln Glu Leu Gln
690 695 700

Glu Ala Thr Asp Glu Leu Asp Leu Lys Leu Arg Gln Ala Glu Val Ile
705 710 715 720

Lys Gly Ser Trp Gln Pro Val Gly Asp Leu Leu Ile Asp Ser Leu Gln
725 730 735

Asp His Leu Glu Lys Val Lys Ala Leu Arg Gly Glu Ile Ala Pro Leu
740 745 750

Lys Glu Asn Val Ser His Val Asn Asp Leu Ala Arg Gln Leu Thr Thr
755 760 765

Leu Gly Ile Gln Leu Ser Pro Tyr Asn Leu Ser Thr Leu Glu Asp Leu
770 775 780

[0136] Asn Thr Arg Trp Lys Leu Leu Gln Val Ala Val Glu Asp Arg Val Arg
785 790 795 800

Gln Leu His Glu Ala His Arg Asp Phe Gly Pro Ala Ser Gln His Phe
805 810 815

Leu Ser Thr Ser Val Gln Gly Pro Trp Glu Arg Ala Ile Ser Pro Asn
820 825 830

Lys Val Pro Tyr Tyr Ile Asn His Glu Thr Gln Thr Thr Cys Trp Asp
835 840 845

His Pro Lys Met Thr Glu Leu Tyr Gln Ser Leu Ala Asp Leu Asn Asn
850 855 860

Val Arg Phe Ser Ala Tyr Arg Thr Ala Met Lys Leu Arg Arg Leu Gln
865 870 875 880

Lys Ala Leu Cys Leu Asp Leu Leu Ser Leu Ser Ala Ala Cys Asp Ala
885 890 895

Leu Asp Gln His Asn Leu Lys Gln Asn Asp Gln Pro Met Asp Ile Leu

900 905 910

Gln Ile Ile Asn Cys Leu Thr Thr Ile Tyr Asp Arg Leu Glu Gln Glu
915 920 925

His Asn Asn Leu Val Asn Val Pro Leu Cys Val Asp Met Cys Leu Asn
930 935 940

Trp Leu Leu Asn Val Tyr Asp Thr Gly Arg Thr Gly Arg Ile Arg Val
945 950 955 960

Leu Ser Phe Lys Thr Gly Ile Ile Ser Leu Cys Lys Ala His Leu Glu
965 970 975

Asp Lys Tyr Arg Tyr Leu Phe Lys Gln Val Ala Ser Ser Thr Gly Phe
980 985 990

Cys Asp Gln Arg Arg Leu Gly Leu Leu Leu His Asp Ser Ile Gln Ile
995 1000 1005

[0137]

Pro Arg Gln Leu Gly Glu Val Ala Ser Phe Gly Gly Ser Asn Ile
1010 1015 1020

Glu Pro Ser Val Arg Ser Cys Phe Gln Phe Ala Asn Asn Lys Pro
1025 1030 1035

Glu Ile Glu Ala Ala Leu Phe Leu Asp Trp Met Arg Leu Glu Pro
1040 1045 1050

Gln Ser Met Val Trp Leu Pro Val Leu His Arg Val Ala Ala Ala
1055 1060 1065

Glu Thr Ala Lys His Gln Ala Lys Cys Asn Ile Cys Lys Glu Cys
1070 1075 1080

Pro Ile Ile Gly Phe Arg Tyr Arg Ser Leu Lys His Phe Asn Tyr
1085 1090 1095

Asp Ile Cys Gln Ser Cys Phe Phe Ser Gly Arg Val Ala Lys Gly
1100 1105 1110

His Lys Met His Tyr Pro Met Val Glu Tyr Cys Thr Pro Thr Thr
1115 1120 1125

Ser Gly Glu
1130

<210> 14
<211> 1131
<212> PRT
<213> 人工序列

<220>
<223> 毫微-肌养蛋白的氨基酸序列

<400> 14

Met Leu Trp Trp Glu Glu Val Glu Asp Cys Tyr Glu Arg Glu Asp Val
1 5 10 15

Gln Lys Lys Thr Phe Thr Lys Trp Val Asn Ala Gln Phe Ser Lys Phe
20 25 30

Gly Lys Gln His Ile Glu Asn Leu Phe Ser Asp Leu Gln Asp Gly Arg
35 40 45

Arg Leu Leu Asp Leu Leu Glu Gly Leu Thr Gly Gln Lys Leu Pro Lys
50 55 60

[0138]

Glu Lys Gly Ser Thr Arg Val His Ala Leu Asn Asn Val Asn Lys Ala
65 70 75 80

Leu Arg Val Leu Gln Asn Asn Asn Val Asp Leu Val Asn Ile Gly Ser
85 90 95

Thr Asp Ile Val Asp Gly Asn His Lys Leu Thr Leu Gly Leu Ile Trp
100 105 110

Asn Ile Ile Leu His Trp Gln Val Lys Asn Val Met Lys Asn Ile Met
115 120 125

Ala Gly Leu Gln Gln Thr Asn Ser Glu Lys Ile Leu Leu Ser Trp Val
130 135 140

Arg Gln Ser Thr Arg Asn Tyr Pro Gln Val Asn Val Ile Asn Phe Thr
145 150 155 160

Thr Ser Trp Ser Asp Gly Leu Ala Leu Asn Ala Leu Ile His Ser His
165 170 175

Arg Pro Asp Leu Phe Asp Trp Asn Ser Val Val Cys Gln Gln Ser Ala
180 185 190

Thr Gln Arg Leu Glu His Ala Phe Asn Ile Ala Arg Tyr Gln Leu Gly
195 200 205

Ile Glu Lys Leu Leu Asp Pro Glu Asp Val Asp Thr Thr Tyr Pro Asp
210 215 220

Lys Lys Ser Ile Leu Met Tyr Ile Thr Ser Leu Phe Gln Val Leu Pro
225 230 235 240

Gln Gln Val Ser Ile Glu Ala Ile Gln Glu Val Glu Met Leu Pro Arg
245 250 255

Pro Pro Lys Val Thr Lys Glu Glu His Phe Gln Leu His His Gln Met
260 265 270

His Tyr Ser Gln Gln Ile Thr Val Ser Leu Ala Gln Gly Tyr Glu Arg
275 280 285

[0139] Thr Ser Ser Pro Lys Pro Arg Phe Lys Ser Tyr Ala Tyr Thr Gln Ala
290 295 300

Ala Tyr Val Thr Thr Ser Asp Pro Thr Arg Ser Pro Phe Pro Ser Gln
305 310 315 320

His Leu Glu Ala Pro Glu Asp Lys Ser Phe Gly Ser Ser Leu Met Glu
325 330 335

Ser Glu Val Asn Leu Asp Arg Tyr Gln Thr Ala Leu Glu Glu Val Leu
340 345 350

Ser Trp Leu Leu Ser Ala Glu Asp Thr Leu Gln Ala Gln Gly Glu Ile
355 360 365

Ser Asn Asp Val Glu Val Val Lys Asp Gln Phe His Thr His Glu Gly
370 375 380

Tyr Met Met Asp Leu Thr Ala His Gln Gly Arg Val Gly Asn Ile Leu
385 390 395 400

Gln Leu Gly Ser Lys Leu Ile Gly Thr Gly Lys Leu Ser Glu Asp Glu

405 410 415

Glu Thr Glu Val Gln Glu Gln Met Asn Leu Leu Asn Ser Arg Trp Glu
420 425 430

Cys Leu Arg Val Ala Ser Met Glu Lys Gln Ser Asn Leu His Arg Val
435 440 445

Leu Met Asp Leu Gln Asn Gln Lys Leu Lys Glu Leu Asn Asp Trp Leu
450 455 460

Thr Glu Ala Glu Thr Thr Ala Asn Val Leu Gln Asp Ala Thr Arg Lys
465 470 475 480

Glu Arg Leu Leu Glu Asp Ser Lys Gly Val Lys Glu Leu Met Lys Gln
485 490 495

Trp Gln Asp Leu Gln Gly Glu Leu Glu Gln Glu Gln Val Arg Val Asn
500 505 510

Ser Leu Thr His Met Val Val Val Val Asp Glu Ser Ser Gly Asp His
[0140] 515 520 525

Ala Thr Ala Ala Leu Glu Glu Gln Leu Lys Val Leu Gly Asp Arg Trp
530 535 540

Ala Asn Ile Cys Arg Trp Ser Glu Leu Arg Lys Lys Ser Leu Asn Ile
545 550 555 560

Arg Ser His Leu Glu Ala Ser Ser Asp Gln Trp Lys Arg Leu His Leu
565 570 575

Ser Leu Gln Glu Leu Leu Val Trp Leu Gln Leu Lys Asp Asp Glu Leu
580 585 590

Ser Arg Gln Ala Pro Ile Gly Gly Asp Phe Pro Ala Val Gln Lys Gln
595 600 605

Asn Asp Val His Arg Ala Phe Lys Arg Glu Leu Lys Thr Lys Glu Pro
610 615 620

Val Ile Met Ser Thr Leu Glu Thr Val Arg Ile Phe Leu Thr Glu Gln
625 630 635 640

Pro Leu Glu Gly Leu Glu Lys Leu Tyr Gln Glu Pro Arg Glu Leu Pro
645 650 655

Pro Glu Glu Arg Ala Gln Asn Val Thr Arg Leu Leu Arg Lys Gln Ala
660 665 670

Glu Glu Val Asn Thr Glu Trp Glu Lys Leu Asn Leu His Ser Ala Asp
675 680 685

Trp Gln Arg Lys Ile Asp Glu Thr Leu Glu Arg Leu Gln Glu Leu Gln
690 695 700

Glu Ala Thr Asp Glu Leu Asp Leu Lys Leu Arg Gln Ala Glu Val Ile
705 710 715 720

Lys Gly Ser Trp Gln Pro Val Gly Asp Leu Leu Ile Asp Ser Leu Gln
725 730 735

Asp His Leu Glu Lys Val Lys Ala Leu Arg Gly Glu Ile Ala Pro Leu
740 745 750

[0141] Lys Glu Asn Val Ser His Val Asn Asp Leu Ala Arg Gln Leu Thr Thr
755 760 765

Leu Gly Ile Gln Leu Ser Pro Tyr Asn Leu Ser Thr Leu Glu Asp Leu
770 775 780

Asn Thr Arg Trp Lys Leu Leu Gln Val Ala Val Glu Asp Arg Val Arg
785 790 795 800

Gln Leu His Glu Ala His Arg Asp Phe Gly Pro Ala Ser Gln His Phe
805 810 815

Leu Ser Thr Ser Val Gln Gly Pro Trp Glu Arg Ala Ile Ser Pro Asn
820 825 830

Lys Val Pro Tyr Tyr Ile Asn His Glu Thr Gln Thr Thr Cys Trp Asp
835 840 845

His Pro Lys Met Thr Glu Leu Tyr Gln Ser Leu Ala Asp Leu Asn Asn
850 855 860

Val Arg Phe Ser Ala Tyr Arg Thr Ala Met Lys Leu Arg Arg Leu Gln

865 870 875 880

Lys Ala Leu Cys Leu Asp Leu Leu Ser Leu Ser Ala Ala Cys Asp Ala
885 890 895

Leu Asp Gln His Asn Leu Lys Gln Asn Asp Gln Pro Met Asp Ile Leu
900 905 910

Gln Ile Ile Asn Cys Leu Thr Thr Ile Tyr Asp Arg Leu Glu Gln Glu
915 920 925

His Asn Asn Leu Val Asn Val Pro Leu Cys Val Asp Met Cys Leu Asn
930 935 940

Trp Leu Leu Asn Val Tyr Asp Thr Gly Arg Thr Gly Arg Ile Arg Val
945 950 955 960

Leu Ser Phe Lys Thr Gly Ile Ile Ser Leu Cys Lys Ala His Leu Glu
965 970 975

Asp Lys Tyr Arg Tyr Leu Phe Lys Gln Val Ala Ser Ser Thr Gly Phe
[0142] 980 985 990

Cys Asp Gln Arg Arg Leu Gly Leu Leu Leu His Asp Ser Ile Gln Ile
995 1000 1005

Pro Arg Gln Leu Gly Glu Val Ala Ser Phe Gly Gly Ser Asn Ile
1010 1015 1020

Glu Pro Ser Val Arg Ser Cys Phe Gln Phe Ala Asn Asn Lys Pro
1025 1030 1035

Glu Ile Glu Ala Ala Leu Phe Leu Asp Trp Met Arg Leu Glu Pro
1040 1045 1050

Gln Ser Met Val Trp Leu Pro Val Leu His Arg Val Ala Ala Ala
1055 1060 1065

Glu Thr Ala Lys His Gln Ala Lys Cys Asn Ile Cys Lys Glu Cys
1070 1075 1080

Pro Ile Ile Gly Phe Arg Tyr Arg Ser Leu Lys His Phe Asn Tyr
1085 1090 1095

Asp Ile Cys Gln Ser Cys Phe Phe Ser Gly Arg Val Ala Lys Gly
1100 1105 1110

His Lys Met His Tyr Pro Met Val Glu Tyr Cys Thr Pro Thr Thr
1115 1120 1125

Ser Gly Glu
1130

<210> 15
<211> 1131
<212> PRT
<213> 人工序列

<220>
<223> 毫微-肌养蛋白的氨基酸序列

<400> 15

Met Leu Trp Trp Glu Glu Val Glu Asp Cys Tyr Glu Arg Glu Asp Val
1 5 10 15

Gln Lys Lys Thr Phe Thr Lys Trp Val Asn Ala Gln Phe Ser Lys Phe
20 25 30

[0143]

Gly Lys Gln His Ile Glu Asn Leu Phe Ser Asp Leu Gln Asp Gly Arg
35 40 45

Arg Leu Leu Asp Leu Leu Glu Gly Leu Thr Gly Gln Lys Leu Pro Lys
50 55 60

Glu Lys Gly Ser Thr Arg Val His Ala Leu Asn Asn Val Asn Lys Ala
65 70 75 80

Leu Arg Val Leu Gln Asn Asn Asn Val Asp Leu Val Asn Ile Gly Ser
85 90 95

Thr Asp Ile Val Asp Gly Asn His Lys Leu Thr Leu Gly Leu Ile Trp
100 105 110

Asn Ile Ile Leu His Trp Gln Val Lys Asn Val Met Lys Asn Ile Met
115 120 125

Ala Gly Leu Gln Gln Thr Asn Ser Glu Lys Ile Leu Leu Ser Trp Val
130 135 140

Arg Gln Ser Thr Arg Asn Tyr Pro Gln Val Asn Val Ile Asn Phe Thr
145 150 155 160

Thr Ser Trp Ser Asp Gly Leu Ala Leu Asn Ala Leu Ile His Ser His
165 170 175

Arg Pro Asp Leu Phe Asp Trp Asn Ser Val Val Cys Gln Gln Ser Ala
180 185 190

Thr Gln Arg Leu Glu His Ala Phe Asn Ile Ala Arg Tyr Gln Leu Gly
195 200 205

Ile Glu Lys Leu Leu Asp Pro Glu Asp Val Asp Thr Thr Tyr Pro Asp
210 215 220

Lys Lys Ser Ile Leu Met Tyr Ile Thr Ser Leu Phe Gln Val Leu Pro
225 230 235 240

Gln Gln Val Ser Ile Glu Ala Ile Gln Glu Val Glu Met Leu Pro Arg
245 250 255

[0144] Pro Pro Lys Val Thr Lys Glu Glu His Phe Gln Leu His His Gln Met
260 265 270

His Tyr Ser Gln Gln Ile Thr Val Ser Leu Ala Gln Gly Tyr Glu Arg
275 280 285

Thr Ser Ser Pro Lys Pro Arg Phe Lys Ser Tyr Ala Tyr Thr Gln Ala
290 295 300

Ala Tyr Val Thr Thr Ser Asp Pro Thr Arg Ser Pro Phe Pro Ser Gln
305 310 315 320

His Leu Glu Ala Pro Glu Asp Lys Ser Phe Gly Ser Ser Leu Met Glu
325 330 335

Ser Glu Val Asn Leu Asp Arg Tyr Gln Thr Ala Leu Glu Glu Val Leu
340 345 350

Ser Trp Leu Leu Ser Ala Glu Asp Thr Leu Gln Ala Gln Gly Glu Ile
355 360 365

Ser Asn Asp Val Glu Val Val Lys Asp Gln Phe His Thr His Glu Gly

370 375 380

Tyr Met Met Asp Leu Thr Ala His Gln Gly Arg Val Gly Asn Ile Leu
385 390 395 400

Gln Leu Gly Ser Lys Leu Ile Gly Thr Gly Lys Leu Ser Glu Asp Glu
405 410 415

Glu Thr Glu Val Gln Glu Gln Met Asn Leu Leu Asn Ser Arg Trp Glu
420 425 430

Cys Leu Arg Val Ala Ser Met Glu Lys Gln Ser Asn Leu His Arg Val
435 440 445

Leu Met Asp Leu Gln Asn Gln Lys Leu Lys Glu Leu Asn Asp Trp Leu
450 455 460

Thr Glu Ala Glu Thr Thr Ala Asn Val Leu Gln Asp Ala Thr Arg Lys
465 470 475 480

[0145] Glu Arg Leu Leu Glu Asp Ser Lys Gly Val Lys Glu Leu Met Lys Gln
485 490 495

Trp Gln Asp Leu Gln Gly Glu Ile Glu Gln Glu Gln Val Arg Val Asn
500 505 510

Ser Leu Thr His Met Val Val Val Val Asp Glu Ser Ser Gly Asp His
515 520 525

Ala Thr Ala Ala Leu Glu Glu Gln Leu Lys Val Leu Gly Asp Arg Trp
530 535 540

Ala Asn Ile Cys Arg Trp Ser Glu Leu Arg Lys Lys Ser Leu Asn Ile
545 550 555 560

Arg Ser His Leu Glu Ala Ser Ser Asp Gln Trp Lys Arg Leu His Leu
565 570 575

Ser Leu Gln Glu Leu Leu Val Trp Leu Gln Leu Lys Asp Asp Glu Leu
580 585 590

Ser Arg Gln Ala Pro Ile Gly Gly Asp Phe Pro Ala Val Gln Lys Gln
595 600 605

Asn Asp Val His Arg Ala Phe Lys Arg Glu Leu Lys Thr Lys Glu Pro
610 615 620

Val Ile Met Ser Thr Leu Glu Thr Val Arg Ile Phe Leu Thr Glu Gln
625 630 635 640

Pro Leu Glu Gly Leu Glu Lys Leu Tyr Gln Glu Pro Arg Glu Leu Pro
645 650 655

Pro Glu Glu Arg Ala Gln Asn Val Thr Arg Leu Leu Arg Lys Gln Ala
660 665 670

Glu Glu Val Asn Thr Glu Trp Glu Lys Leu Asn Leu His Ser Ala Asp
675 680 685

Trp Gln Arg Lys Ile Asp Glu Thr Leu Glu Arg Leu Gln Glu Leu Gln
690 695 700

Glu Ala Thr Asp Glu Leu Asp Leu Lys Leu Arg Gln Ala Glu Val Ile
705 710 715 720

[0146] Lys Gly Ser Trp Gln Pro Val Gly Asp Leu Leu Ile Asp Ser Leu Gln
725 730 735

Asp His Leu Glu Lys Val Lys Ala Leu Arg Gly Glu Ile Ala Pro Leu
740 745 750

Lys Glu Asn Val Ser His Val Asn Asp Leu Ala Arg Gln Leu Thr Thr
755 760 765

Leu Gly Ile Gln Leu Ser Pro Tyr Asn Leu Ser Thr Leu Glu Asp Leu
770 775 780

Asn Thr Arg Trp Lys Leu Leu Gln Val Ala Val Glu Asp Arg Val Arg
785 790 795 800

Gln Leu His Glu Ala His Arg Asp Phe Gly Pro Ala Ser Gln His Phe
805 810 815

Leu Ser Thr Ser Val Gln Gly Pro Trp Glu Arg Ala Ile Ser Pro Asn
820 825 830

Lys Val Pro Tyr Tyr Ile Asn His Glu Thr Gln Thr Thr Cys Trp Asp

835 840 845

His Pro Lys Met Thr Glu Leu Tyr Gln Ser Leu Ala Asp Leu Asn Asn
850 855 860

Val Arg Phe Ser Ala Tyr Arg Thr Ala Met Lys Leu Arg Arg Leu Gln
865 870 875 880

Lys Ala Leu Cys Leu Asp Leu Leu Ser Leu Ser Ala Ala Cys Asp Ala
885 890 895

Leu Asp Gln His Asn Leu Lys Gln Asn Asp Gln Pro Met Asp Ile Leu
900 905 910

Gln Ile Ile Asn Cys Leu Thr Thr Ile Tyr Asp Arg Leu Glu Gln Glu
915 920 925

His Asn Asn Leu Val Asn Val Pro Leu Cys Val Asp Met Cys Leu Asn
930 935 940

Trp Leu Leu Asn Val Tyr Asp Thr Gly Arg Thr Gly Arg Ile Arg Val
[0147] 945 950 955 960

Leu Ser Phe Lys Thr Gly Ile Ile Ser Leu Cys Lys Ala His Leu Glu
965 970 975

Asp Lys Tyr Arg Tyr Leu Phe Lys Gln Val Ala Ser Ser Thr Gly Phe
980 985 990

Cys Asp Gln Arg Arg Leu Gly Leu Leu Leu His Asp Ser Ile Gln Ile
995 1000 1005

Pro Arg Gln Leu Gly Glu Val Ala Ser Phe Gly Gly Ser Asn Ile
1010 1015 1020

Glu Pro Ser Val Arg Ser Cys Phe Gln Phe Ala Asn Asn Lys Pro
1025 1030 1035

Glu Ile Glu Ala Ala Leu Phe Leu Asp Trp Met Arg Leu Glu Pro
1040 1045 1050

Gln Ser Met Val Trp Leu Pro Val Leu His Arg Val Ala Ala Ala
1055 1060 1065

Glu Thr Ala Lys His Gln Ala Lys Cys Asn Ile Cys Lys Glu Cys
1070 1075 1080

Pro Ile Ile Gly Phe Arg Tyr Arg Ser Leu Lys His Phe Asn Tyr
1085 1090 1095

Asp Ile Cys Gln Ser Cys Phe Phe Ser Gly Arg Val Ala Lys Gly
1100 1105 1110

His Lys Met His Tyr Pro Met Val Glu Tyr Cys Thr Pro Thr Thr
1115 1120 1125

Ser Gly Glu
1130

<210> 16
<211> 1131
<212> PRT
<213> 人工序列

<220>
<223> 毫微-肌养蛋白的氨基酸序列

[0148]

<400> 16

Met Leu Trp Trp Glu Glu Val Glu Asp Cys Tyr Glu Arg Glu Asp Val
1 5 10 15

Gln Lys Lys Thr Phe Thr Lys Trp Val Asn Ala Gln Phe Ser Lys Phe
20 25 30

Gly Lys Gln His Ile Glu Asn Leu Phe Ser Asp Leu Gln Asp Gly Arg
35 40 45

Arg Leu Leu Asp Leu Leu Glu Gly Leu Thr Gly Gln Lys Leu Pro Lys
50 55 60

Glu Lys Gly Ser Thr Arg Val His Ala Leu Asn Asn Val Asn Lys Ala
65 70 75 80

Leu Arg Val Leu Gln Asn Asn Asn Val Asp Leu Val Asn Ile Gly Ser
85 90 95

Thr Asp Ile Val Asp Gly Asn His Lys Leu Thr Leu Gly Leu Ile Trp
100 105 110

Asn Ile Ile Leu His Trp Gln Val Lys Asn Val Met Lys Asn Ile Met
115 120 125

Ala Gly Leu Gln Gln Thr Asn Ser Glu Lys Ile Leu Leu Ser Trp Val
130 135 140

Arg Gln Ser Thr Arg Asn Tyr Pro Gln Val Asn Val Ile Asn Phe Thr
145 150 155 160

Thr Ser Trp Ser Asp Gly Leu Ala Leu Asn Ala Leu Ile His Ser His
165 170 175

Arg Pro Asp Leu Phe Asp Trp Asn Ser Val Val Cys Gln Gln Ser Ala
180 185 190

Thr Gln Arg Leu Glu His Ala Phe Asn Ile Ala Arg Tyr Gln Leu Gly
195 200 205

Ile Glu Lys Leu Leu Asp Pro Glu Asp Val Asp Thr Thr Tyr Pro Asp
210 215 220

[0149] Lys Lys Ser Ile Leu Met Tyr Ile Thr Ser Leu Phe Gln Val Leu Pro
225 230 235 240

Gln Gln Val Ser Ile Glu Ala Ile Gln Glu Val Glu Met Leu Pro Arg
245 250 255

Pro Pro Lys Val Thr Lys Glu Glu His Phe Gln Leu His His Gln Met
260 265 270

His Tyr Ser Gln Gln Ile Thr Val Ser Leu Ala Gln Gly Tyr Glu Arg
275 280 285

Thr Ser Ser Pro Lys Pro Arg Phe Lys Ser Tyr Ala Tyr Thr Gln Ala
290 295 300

Ala Tyr Val Thr Thr Ser Asp Pro Thr Arg Ser Pro Phe Pro Ser Gln
305 310 315 320

His Leu Glu Ala Pro Glu Asp Lys Ser Phe Gly Ser Ser Leu Met Glu
325 330 335

Ser Glu Val Asn Leu Asp Arg Tyr Gln Thr Ala Leu Glu Glu Val Leu

340 345 350

Ser Trp Leu Leu Ser Ala Glu Asp Thr Leu Gln Ala Gln Gly Glu Ile
355 360 365

Ser Asn Asp Val Glu Val Val Lys Asp Gln Phe His Thr His Glu Gly
370 375 380

Tyr Met Met Asp Leu Thr Ala His Gln Gly Arg Val Gly Asn Ile Leu
385 390 395 400

Gln Leu Gly Ser Lys Leu Ile Gly Thr Gly Lys Leu Ser Glu Asp Glu
405 410 415

Glu Thr Glu Val Gln Glu Gln Met Asn Leu Leu Asn Ser Arg Trp Glu
420 425 430

Cys Leu Arg Val Ala Ser Met Glu Lys Gln Ser Asn Leu His Arg Val
435 440 445

[0150] Leu Met Asp Leu Gln Asn Gln Lys Leu Lys Glu Leu Asn Asp Trp Leu
450 455 460

Thr Glu Ala Glu Thr Thr Ala Asn Val Leu Gln Asp Ala Thr Arg Lys
465 470 475 480

Glu Arg Leu Leu Glu Asp Ser Lys Gly Val Lys Glu Leu Met Lys Gln
485 490 495

Trp Gln Asp Leu Gln Gly Glu Ile Glu Ala Glu Gln Val Arg Val Asn
500 505 510

Ser Leu Thr His Met Val Val Val Val Asp Glu Ser Ser Gly Asp His
515 520 525

Ala Thr Ala Ala Leu Glu Glu Gln Leu Lys Val Leu Gly Asp Arg Trp
530 535 540

Ala Asn Ile Cys Arg Trp Ser Glu Leu Arg Lys Lys Ser Leu Asn Ile
545 550 555 560

Arg Ser His Leu Glu Ala Ser Ser Asp Gln Trp Lys Arg Leu His Leu
565 570 575

Ser Leu Gln Glu Leu Leu Val Trp Leu Gln Leu Lys Asp Asp Glu Leu
580 585 590

Ser Arg Gln Ala Pro Ile Gly Gly Asp Phe Pro Ala Val Gln Lys Gln
595 600 605

Asn Asp Val His Arg Ala Phe Lys Arg Glu Leu Lys Thr Lys Glu Pro
610 615 620

Val Ile Met Ser Thr Leu Glu Thr Val Arg Ile Phe Leu Thr Glu Gln
625 630 635 640

Pro Leu Glu Gly Leu Glu Lys Leu Tyr Gln Glu Pro Arg Glu Leu Pro
645 650 655

Pro Glu Glu Arg Ala Gln Asn Val Thr Arg Leu Leu Arg Lys Gln Ala
660 665 670

Glu Glu Val Asn Thr Glu Trp Glu Lys Leu Asn Leu His Ser Ala Asp
675 680 685

[0151] Trp Gln Arg Lys Ile Asp Glu Thr Leu Glu Arg Leu Gln Glu Leu Gln
690 695 700

Glu Ala Thr Asp Glu Leu Asp Leu Lys Leu Arg Gln Ala Glu Val Ile
705 710 715 720

Lys Gly Ser Trp Gln Pro Val Gly Asp Leu Leu Ile Asp Ser Leu Gln
725 730 735

Asp His Leu Glu Lys Val Lys Ala Leu Arg Gly Glu Ile Ala Pro Leu
740 745 750

Lys Glu Asn Val Ser His Val Asn Asp Leu Ala Arg Gln Leu Thr Thr
755 760 765

Leu Gly Ile Gln Leu Ser Pro Tyr Asn Leu Ser Thr Leu Glu Asp Leu
770 775 780

Asn Thr Arg Trp Lys Leu Leu Gln Val Ala Val Glu Asp Arg Val Arg
785 790 795 800

Gln Leu His Glu Ala His Arg Asp Phe Gly Pro Ala Ser Gln His Phe

805 810 815

Leu Ser Thr Ser Val Gln Gly Pro Trp Glu Arg Ala Ile Ser Pro Asn
820 825 830

Lys Val Pro Tyr Tyr Ile Asn His Glu Thr Gln Thr Thr Cys Trp Asp
835 840 845

His Pro Lys Met Thr Glu Leu Tyr Gln Ser Leu Ala Asp Leu Asn Asn
850 855 860

Val Arg Phe Ser Ala Tyr Arg Thr Ala Met Lys Leu Arg Arg Leu Gln
865 870 875 880

Lys Ala Leu Cys Leu Asp Leu Leu Ser Leu Ser Ala Ala Cys Asp Ala
885 890 895

Leu Asp Gln His Asn Leu Lys Gln Asn Asp Gln Pro Met Asp Ile Leu
900 905 910

Gln Ile Ile Asn Cys Leu Thr Thr Ile Tyr Asp Arg Leu Glu Gln Glu
[0152] 915 920 925

His Asn Asn Leu Val Asn Val Pro Leu Cys Val Asp Met Cys Leu Asn
930 935 940

Trp Leu Leu Asn Val Tyr Asp Thr Gly Arg Thr Gly Arg Ile Arg Val
945 950 955 960

Leu Ser Phe Lys Thr Gly Ile Ile Ser Leu Cys Lys Ala His Leu Glu
965 970 975

Asp Lys Tyr Arg Tyr Leu Phe Lys Gln Val Ala Ser Ser Thr Gly Phe
980 985 990

Cys Asp Gln Arg Arg Leu Gly Leu Leu Leu His Asp Ser Ile Gln Ile
995 1000 1005

Pro Arg Gln Leu Gly Glu Val Ala Ser Phe Gly Gly Ser Asn Ile
1010 1015 1020

Glu Pro Ser Val Arg Ser Cys Phe Gln Phe Ala Asn Asn Lys Pro
1025 1030 1035

Glu Ile Glu Ala Ala Leu Phe Leu Asp Trp Met Arg Leu Glu Pro
1040 1045 1050

Gln Ser Met Val Trp Leu Pro Val Leu His Arg Val Ala Ala Ala
1055 1060 1065

Glu Thr Ala Lys His Gln Ala Lys Cys Asn Ile Cys Lys Glu Cys
1070 1075 1080

Pro Ile Ile Gly Phe Arg Tyr Arg Ser Leu Lys His Phe Asn Tyr
1085 1090 1095

Asp Ile Cys Gln Ser Cys Phe Phe Ser Gly Arg Val Ala Lys Gly
1100 1105 1110

His Lys Met His Tyr Pro Met Val Glu Tyr Cys Thr Pro Thr Thr
1115 1120 1125

Ser Gly Glu
1130

[0153]

<210> 17
<211> 1131
<212> PRT
<213> 人工序列

<220>
<223> 毫微-肌养蛋白的氨基酸序列

<400> 17

Met Leu Trp Trp Glu Glu Val Glu Asp Cys Tyr Glu Arg Glu Asp Val
1 5 10 15

Gln Lys Lys Thr Phe Thr Lys Trp Val Asn Ala Gln Phe Ser Lys Phe
20 25 30

Gly Lys Gln His Ile Glu Asn Leu Phe Ser Asp Leu Gln Asp Gly Arg
35 40 45

Arg Leu Leu Asp Leu Leu Glu Gly Leu Thr Gly Gln Lys Leu Pro Lys
50 55 60

Glu Lys Gly Ser Thr Arg Val His Ala Leu Asn Asn Val Asn Lys Ala
65 70 75 80

Leu Arg Val Leu Gln Asn Asn Asn Val Asp Leu Val Asn Ile Gly Ser
85 90 95

Thr Asp Ile Val Asp Gly Asn His Lys Leu Thr Leu Gly Leu Ile Trp
100 105 110

Asn Ile Ile Leu His Trp Gln Val Lys Asn Val Met Lys Asn Ile Met
115 120 125

Ala Gly Leu Gln Gln Thr Asn Ser Glu Lys Ile Leu Leu Ser Trp Val
130 135 140

Arg Gln Ser Thr Arg Asn Tyr Pro Gln Val Asn Val Ile Asn Phe Thr
145 150 155 160

Thr Ser Trp Ser Asp Gly Leu Ala Leu Asn Ala Leu Ile His Ser His
165 170 175

Arg Pro Asp Leu Phe Asp Trp Asn Ser Val Val Cys Gln Gln Ser Ala
180 185 190

[0154] Thr Gln Arg Leu Glu His Ala Phe Asn Ile Ala Arg Tyr Gln Leu Gly
195 200 205

Ile Glu Lys Leu Leu Asp Pro Glu Asp Val Asp Thr Thr Tyr Pro Asp
210 215 220

Lys Lys Ser Ile Leu Met Tyr Ile Thr Ser Leu Phe Gln Val Leu Pro
225 230 235 240

Gln Gln Val Ser Ile Glu Ala Ile Gln Glu Val Glu Met Leu Pro Arg
245 250 255

Pro Pro Lys Val Thr Lys Glu Glu His Phe Gln Leu His His Gln Met
260 265 270

His Tyr Ser Gln Gln Ile Thr Val Ser Leu Ala Gln Gly Tyr Glu Arg
275 280 285

Thr Ser Ser Pro Lys Pro Arg Phe Lys Ser Tyr Ala Tyr Thr Gln Ala
290 295 300

Ala Tyr Val Thr Thr Ser Asp Pro Thr Arg Ser Pro Phe Pro Ser Gln

305 310 315 320

His Leu Glu Ala Pro Glu Asp Lys Ser Phe Gly Ser Ser Leu Met Glu
325 330 335

Ser Glu Val Asn Leu Asp Arg Tyr Gln Thr Ala Leu Glu Glu Val Leu
340 345 350

Ser Trp Leu Leu Ser Ala Glu Asp Thr Leu Gln Ala Gln Gly Glu Ile
355 360 365

Ser Asn Asp Val Glu Val Val Lys Asp Gln Phe His Thr His Glu Gly
370 375 380

Tyr Met Met Asp Leu Thr Ala His Gln Gly Arg Val Gly Asn Ile Leu
385 390 395 400

Gln Leu Gly Ser Lys Leu Ile Gly Thr Gly Lys Leu Ser Glu Asp Glu
405 410 415

Glu Thr Glu Val Gln Glu Gln Met Asn Leu Leu Asn Ser Arg Trp Glu
[0155] 420 425 430

Cys Leu Arg Val Ala Ser Met Glu Lys Gln Ser Asn Leu His Arg Val
435 440 445

Leu Met Asp Leu Gln Asn Gln Lys Leu Lys Glu Leu Asn Asp Trp Leu
450 455 460

Thr Glu Ala Glu Thr Thr Ala Asn Val Leu Gln Asp Ala Thr Arg Lys
465 470 475 480

Glu Arg Leu Leu Glu Asp Ser Lys Gly Val Lys Glu Leu Met Lys Gln
485 490 495

Trp Gln Asp Leu Gln Gly Glu Ile Glu Ala His Gln Val Arg Val Asn
500 505 510

Ser Leu Thr His Met Val Val Val Val Asp Glu Ser Ser Gly Asp His
515 520 525

Ala Thr Ala Ala Leu Glu Glu Gln Leu Lys Val Leu Gly Asp Arg Trp
530 535 540

Ala Asn Ile Cys Arg Trp Ser Glu Leu Arg Lys Lys Ser Leu Asn Ile
545 550 555 560

Arg Ser His Leu Glu Ala Ser Ser Asp Gln Trp Lys Arg Leu His Leu
565 570 575

Ser Leu Gln Glu Leu Leu Val Trp Leu Gln Leu Lys Asp Asp Glu Leu
580 585 590

Ser Arg Gln Ala Pro Ile Gly Gly Asp Phe Pro Ala Val Gln Lys Gln
595 600 605

Asn Asp Val His Arg Ala Phe Lys Arg Glu Leu Lys Thr Lys Glu Pro
610 615 620

Val Ile Met Ser Thr Leu Glu Thr Val Arg Ile Phe Leu Thr Glu Gln
625 630 635 640

Pro Leu Glu Gly Leu Glu Lys Leu Tyr Gln Glu Pro Arg Glu Leu Pro
645 650 655

[0156] Pro Glu Glu Arg Ala Gln Asn Val Thr Arg Leu Leu Arg Lys Gln Ala
660 665 670

Glu Glu Val Asn Thr Glu Trp Glu Lys Leu Asn Leu His Ser Ala Asp
675 680 685

Trp Gln Arg Lys Ile Asp Glu Thr Leu Glu Arg Leu Gln Glu Leu Gln
690 695 700

Glu Ala Thr Asp Glu Leu Asp Leu Lys Leu Arg Gln Ala Glu Val Ile
705 710 715 720

Lys Gly Ser Trp Gln Pro Val Gly Asp Leu Leu Ile Asp Ser Leu Gln
725 730 735

Asp His Leu Glu Lys Val Lys Ala Leu Arg Gly Glu Ile Ala Pro Leu
740 745 750

Lys Glu Asn Val Ser His Val Asn Asp Leu Ala Arg Gln Leu Thr Thr
755 760 765

Leu Gly Ile Gln Leu Ser Pro Tyr Asn Leu Ser Thr Leu Glu Asp Leu

770 775 780

Asn Thr Arg Trp Lys Leu Leu Gln Val Ala Val Glu Asp Arg Val Arg
785 790 795 800

Gln Leu His Glu Ala His Arg Asp Phe Gly Pro Ala Ser Gln His Phe
805 810 815

Leu Ser Thr Ser Val Gln Gly Pro Trp Glu Arg Ala Ile Ser Pro Asn
820 825 830

Lys Val Pro Tyr Tyr Ile Asn His Glu Thr Gln Thr Thr Cys Trp Asp
835 840 845

His Pro Lys Met Thr Glu Leu Tyr Gln Ser Leu Ala Asp Leu Asn Asn
850 855 860

Val Arg Phe Ser Ala Tyr Arg Thr Ala Met Lys Leu Arg Arg Leu Gln
865 870 875 880

[0157] Lys Ala Leu Cys Leu Asp Leu Leu Ser Leu Ser Ala Ala Cys Asp Ala
885 890 895

Leu Asp Gln His Asn Leu Lys Gln Asn Asp Gln Pro Met Asp Ile Leu
900 905 910

Gln Ile Ile Asn Cys Leu Thr Thr Ile Tyr Asp Arg Leu Glu Gln Glu
915 920 925

His Asn Asn Leu Val Asn Val Pro Leu Cys Val Asp Met Cys Leu Asn
930 935 940

Trp Leu Leu Asn Val Tyr Asp Thr Gly Arg Thr Gly Arg Ile Arg Val
945 950 955 960

Leu Ser Phe Lys Thr Gly Ile Ile Ser Leu Cys Lys Ala His Leu Glu
965 970 975

Asp Lys Tyr Arg Tyr Leu Phe Lys Gln Val Ala Ser Ser Thr Gly Phe
980 985 990

Cys Asp Gln Arg Arg Leu Gly Leu Leu Leu His Asp Ser Ile Gln Ile
995 1000 1005

Pro Arg Gln Leu Gly Glu Val Ala Ser Phe Gly Gly Ser Asn Ile
1010 1015 1020

Glu Pro Ser Val Arg Ser Cys Phe Gln Phe Ala Asn Asn Lys Pro
1025 1030 1035

Glu Ile Glu Ala Ala Leu Phe Leu Asp Trp Met Arg Leu Glu Pro
1040 1045 1050

Gln Ser Met Val Trp Leu Pro Val Leu His Arg Val Ala Ala Ala
1055 1060 1065

Glu Thr Ala Lys His Gln Ala Lys Cys Asn Ile Cys Lys Glu Cys
1070 1075 1080

Pro Ile Ile Gly Phe Arg Tyr Arg Ser Leu Lys His Phe Asn Tyr
1085 1090 1095

Asp Ile Cys Gln Ser Cys Phe Phe Ser Gly Arg Val Ala Lys Gly
1100 1105 1110

[0158]

His Lys Met His Tyr Pro Met Val Glu Tyr Cys Thr Pro Thr Thr
1115 1120 1125

Ser Gly Glu
1130

<210> 18
<211> 1131
<212> PRT
<213> 人工序列

<220>
<223> 毫微-肌养蛋白的氨基酸序列

<400> 18

Met Leu Trp Trp Glu Glu Val Glu Asp Cys Tyr Glu Arg Glu Asp Val
1 5 10 15

Gln Lys Lys Thr Phe Thr Lys Trp Val Asn Ala Gln Phe Ser Lys Phe
20 25 30

Gly Lys Gln His Ile Glu Asn Leu Phe Ser Asp Leu Gln Asp Gly Arg
35 40 45

Arg Leu Leu Asp Leu Leu Glu Gly Leu Thr Gly Gln Lys Leu Pro Lys
50 55 60

Glu Lys Gly Ser Thr Arg Val His Ala Leu Asn Asn Val Asn Lys Ala
65 70 75 80

Leu Arg Val Leu Gln Asn Asn Asn Val Asp Leu Val Asn Ile Gly Ser
85 90 95

Thr Asp Ile Val Asp Gly Asn His Lys Leu Thr Leu Gly Leu Ile Trp
100 105 110

Asn Ile Ile Leu His Trp Gln Val Lys Asn Val Met Lys Asn Ile Met
115 120 125

Ala Gly Leu Gln Gln Thr Asn Ser Glu Lys Ile Leu Leu Ser Trp Val
130 135 140

Arg Gln Ser Thr Arg Asn Tyr Pro Gln Val Asn Val Ile Asn Phe Thr
145 150 155 160

[0159] Thr Ser Trp Ser Asp Gly Leu Ala Leu Asn Ala Leu Ile His Ser His
165 170 175

Arg Pro Asp Leu Phe Asp Trp Asn Ser Val Val Cys Gln Gln Ser Ala
180 185 190

Thr Gln Arg Leu Glu His Ala Phe Asn Ile Ala Arg Tyr Gln Leu Gly
195 200 205

Ile Glu Lys Leu Leu Asp Pro Glu Asp Val Asp Thr Thr Tyr Pro Asp
210 215 220

Lys Lys Ser Ile Leu Met Tyr Ile Thr Ser Leu Phe Gln Val Leu Pro
225 230 235 240

Gln Gln Val Ser Ile Glu Ala Ile Gln Glu Val Glu Met Leu Pro Arg
245 250 255

Pro Pro Lys Val Thr Lys Glu Glu His Phe Gln Leu His His Gln Met
260 265 270

His Tyr Ser Gln Gln Ile Thr Val Ser Leu Ala Gln Gly Tyr Glu Arg

275 280 285

Thr Ser Ser Pro Lys Pro Arg Phe Lys Ser Tyr Ala Tyr Thr Gln Ala
290 295 300

Ala Tyr Val Thr Thr Ser Asp Pro Thr Arg Ser Pro Phe Pro Ser Gln
305 310 315 320

His Leu Glu Ala Pro Glu Asp Lys Ser Phe Gly Ser Ser Leu Met Glu
325 330 335

Ser Glu Val Asn Leu Asp Arg Tyr Gln Thr Ala Leu Glu Glu Val Leu
340 345 350

Ser Trp Leu Leu Ser Ala Glu Asp Thr Leu Gln Ala Gln Gly Glu Ile
355 360 365

Ser Asn Asp Val Glu Val Val Lys Asp Gln Phe His Thr His Glu Gly
370 375 380

Tyr Met Met Asp Leu Thr Ala His Gln Gly Arg Val Gly Asn Ile Leu
[0160] 385 390 395 400

Gln Leu Gly Ser Lys Leu Ile Gly Thr Gly Lys Leu Ser Glu Asp Glu
405 410 415

Glu Thr Glu Val Gln Glu Gln Met Asn Leu Leu Asn Ser Arg Trp Glu
420 425 430

Cys Leu Arg Val Ala Ser Met Glu Lys Gln Ser Asn Leu His Arg Val
435 440 445

Leu Met Asp Leu Gln Asn Gln Lys Leu Lys Glu Leu Asn Asp Trp Leu
450 455 460

Thr Glu Ala Glu Thr Thr Ala Asn Val Leu Gln Asp Ala Thr Arg Lys
465 470 475 480

Glu Arg Leu Leu Glu Asp Ser Lys Gly Val Lys Glu Leu Met Lys Gln
485 490 495

Trp Gln Asp Leu Gln Gly Glu Ile Glu Ala His Thr Val Arg Val Asn
500 505 510

Ser Leu Thr His Met Val Val Val Val Asp Glu Ser Ser Gly Asp His
515 520 525

Ala Thr Ala Ala Leu Glu Glu Gln Leu Lys Val Leu Gly Asp Arg Trp
530 535 540

Ala Asn Ile Cys Arg Trp Ser Glu Leu Arg Lys Lys Ser Leu Asn Ile
545 550 555 560

Arg Ser His Leu Glu Ala Ser Ser Asp Gln Trp Lys Arg Leu His Leu
565 570 575

Ser Leu Gln Glu Leu Leu Val Trp Leu Gln Leu Lys Asp Asp Glu Leu
580 585 590

Ser Arg Gln Ala Pro Ile Gly Gly Asp Phe Pro Ala Val Gln Lys Gln
595 600 605

Asn Asp Val His Arg Ala Phe Lys Arg Glu Leu Lys Thr Lys Glu Pro
610 615 620

[0161] Val Ile Met Ser Thr Leu Glu Thr Val Arg Ile Phe Leu Thr Glu Gln
625 630 635 640

Pro Leu Glu Gly Leu Glu Lys Leu Tyr Gln Glu Pro Arg Glu Leu Pro
645 650 655

Pro Glu Glu Arg Ala Gln Asn Val Thr Arg Leu Leu Arg Lys Gln Ala
660 665 670

Glu Glu Val Asn Thr Glu Trp Glu Lys Leu Asn Leu His Ser Ala Asp
675 680 685

Trp Gln Arg Lys Ile Asp Glu Thr Leu Glu Arg Leu Gln Glu Leu Gln
690 695 700

Glu Ala Thr Asp Glu Leu Asp Leu Lys Leu Arg Gln Ala Glu Val Ile
705 710 715 720

Lys Gly Ser Trp Gln Pro Val Gly Asp Leu Leu Ile Asp Ser Leu Gln
725 730 735

Asp His Leu Glu Lys Val Lys Ala Leu Arg Gly Glu Ile Ala Pro Leu

740 745 750

Lys Glu Asn Val Ser His Val Asn Asp Leu Ala Arg Gln Leu Thr Thr
755 760 765

Leu Gly Ile Gln Leu Ser Pro Tyr Asn Leu Ser Thr Leu Glu Asp Leu
770 775 780

Asn Thr Arg Trp Lys Leu Leu Gln Val Ala Val Glu Asp Arg Val Arg
785 790 795 800

Gln Leu His Glu Ala His Arg Asp Phe Gly Pro Ala Ser Gln His Phe
805 810 815

Leu Ser Thr Ser Val Gln Gly Pro Trp Glu Arg Ala Ile Ser Pro Asn
820 825 830

Lys Val Pro Tyr Tyr Ile Asn His Glu Thr Gln Thr Thr Cys Trp Asp
835 840 845

His Pro Lys Met Thr Glu Leu Tyr Gln Ser Leu Ala Asp Leu Asn Asn
[0162] 850 855 860

Val Arg Phe Ser Ala Tyr Arg Thr Ala Met Lys Leu Arg Arg Leu Gln
865 870 875 880

Lys Ala Leu Cys Leu Asp Leu Leu Ser Leu Ser Ala Ala Cys Asp Ala
885 890 895

Leu Asp Gln His Asn Leu Lys Gln Asn Asp Gln Pro Met Asp Ile Leu
900 905 910

Gln Ile Ile Asn Cys Leu Thr Thr Ile Tyr Asp Arg Leu Glu Gln Glu
915 920 925

His Asn Asn Leu Val Asn Val Pro Leu Cys Val Asp Met Cys Leu Asn
930 935 940

Trp Leu Leu Asn Val Tyr Asp Thr Gly Arg Thr Gly Arg Ile Arg Val
945 950 955 960

Leu Ser Phe Lys Thr Gly Ile Ile Ser Leu Cys Lys Ala His Leu Glu
965 970 975

Asp Lys Tyr Arg Tyr Leu Phe Lys Gln Val Ala Ser Ser Thr Gly Phe
980 985 990

Cys Asp Gln Arg Arg Leu Gly Leu Leu Leu His Asp Ser Ile Gln Ile
995 1000 1005

Pro Arg Gln Leu Gly Glu Val Ala Ser Phe Gly Gly Ser Asn Ile
1010 1015 1020

Glu Pro Ser Val Arg Ser Cys Phe Gln Phe Ala Asn Asn Lys Pro
1025 1030 1035

Glu Ile Glu Ala Ala Leu Phe Leu Asp Trp Met Arg Leu Glu Pro
1040 1045 1050

Gln Ser Met Val Trp Leu Pro Val Leu His Arg Val Ala Ala Ala
1055 1060 1065

Glu Thr Ala Lys His Gln Ala Lys Cys Asn Ile Cys Lys Glu Cys
1070 1075 1080

[0163]

Pro Ile Ile Gly Phe Arg Tyr Arg Ser Leu Lys His Phe Asn Tyr
1085 1090 1095

Asp Ile Cys Gln Ser Cys Phe Phe Ser Gly Arg Val Ala Lys Gly
1100 1105 1110

His Lys Met His Tyr Pro Met Val Glu Tyr Cys Thr Pro Thr Thr
1115 1120 1125

Ser Gly Glu
1130

<210> 19
<211> 3420
<212> DNA
<213> 人工序列

<220>
<223> 编码毫微-肌养蛋白相关蛋白的人工改造的序列

<220>
<221> CDS
<222> (1)..(3420)

<400> 19

```
atg gct aag tat ggc gag cat gag gct tct ccc gac aac ggc cag aac   48
Met Ala Lys Tyr Gly Glu His Glu Ala Ser Pro Asp Asn Gly Gln Asn
1               5                  10                  15

gag ttc tcc gac atc atc aag tcc cgc tcc gac gag cac aac gac gtg   96
Glu Phe Ser Asp Ile Ile Lys Ser Arg Ser Asp Glu His Asn Asp Val
            20                  25                  30

cag aag aaa acc ttc acc aag tgg atc aac gcc agg ttc tcc aag tcc  144
Gln Lys Lys Thr Phe Thr Lys Trp Ile Asn Ala Arg Phe Ser Lys Ser
        35                  40                  45

ggc aag cct cct atc aac gac atg ttc acc gac ctg aag gac ggc cgc  192
Gly Lys Pro Pro Ile Asn Asp Met Phe Thr Asp Leu Lys Asp Gly Arg
    50                  55                  60

aaa ctg ctg gat ctg ctg gaa gga ctg acc ggc acc tct ctg cct aaa  240
Lys Leu Leu Asp Leu Leu Glu Gly Leu Thr Gly Thr Ser Leu Pro Lys
65                  70                  75                  80

gaa agg ggc tct act agg gtg cac gcc ctg aac aat gtg aac agg gtg  288
Glu Arg Gly Ser Thr Arg Val His Ala Leu Asn Asn Val Asn Arg Val
                85                  90                  95

ctc cag gtg ctg cac cag aac aac gtc gag ctg gtc aat atc ggc ggc  336
Leu Gln Val Leu His Gln Asn Asn Val Glu Leu Val Asn Ile Gly Gly
            100                 105                 110
```

[0164]

```
acc gac atc gtg gat ggc aac cat aag ctg acc ctg ggc ctg ctg tgg  384
Thr Asp Ile Val Asp Gly Asn His Lys Leu Thr Leu Gly Leu Leu Trp
        115                 120                 125

tcc atc att ctg cac tgg cag gtc aaa gac gtg atg aag gac gtg atg  432
Ser Ile Ile Leu His Trp Gln Val Lys Asp Val Met Lys Asp Val Met
    130                 135                 140

tcc gac ctc cag cag acc aac tcc gag aag atc ctg ctg tct tgg gtc  480
Ser Asp Leu Gln Gln Thr Asn Ser Glu Lys Ile Leu Leu Ser Trp Val
145                 150                 155                 160

cgc cag acc act agg ccc tat tct cag gtc aac gtg ctg aac ttc acc  528
Arg Gln Thr Thr Arg Pro Tyr Ser Gln Val Asn Val Leu Asn Phe Thr
                165                 170                 175

acc tct tgg acc gac gga ctg gcc ttc aat gct gtg ctg cac agg cac  576
Thr Ser Trp Thr Asp Gly Leu Ala Phe Asn Ala Val Leu His Arg His
            180                 185                 190

aag ccc gac ctg ttc tct tgg gac aaa gtg gtc aag atg agc ccc atc  624
Lys Pro Asp Leu Phe Ser Trp Asp Lys Val Val Lys Met Ser Pro Ile
        195                 200                 205

gag agg ctg gaa cac gcc ttc tct aag gcc cag acc tac ctg gga atc  672
Glu Arg Leu Glu His Ala Phe Ser Lys Ala Gln Thr Tyr Leu Gly Ile
    210                 215                 220

gag aag ctg ctg gac cct gag gat gtg gca gtg cag ctg ccc gac aag  720
```

```
Glu Lys Leu Leu Asp Pro Glu Asp Val Ala Val Gln Leu Pro Asp Lys
225                 230                 235                 240

aaa tcc atc atc atg tac ctg acc agc ctg ttc gag gtg ctg cct cag      768
Lys Ser Ile Ile Met Tyr Leu Thr Ser Leu Phe Glu Val Leu Pro Gln
                245                 250                 255

cag gtc aca att gat gcc atc cgc gag gtg gaa acc ctg cct cgc aag      816
Gln Val Thr Ile Asp Ala Ile Arg Glu Val Glu Thr Leu Pro Arg Lys
            260                 265                 270

tac aag aaa gag tgc gag gaa gag gcc atc aac atc cag tcc act gct      864
Tyr Lys Lys Glu Cys Glu Glu Glu Ala Ile Asn Ile Gln Ser Thr Ala
        275                 280                 285

ccc gag gaa gaa cac gag tct cca agg gcc gag act cct tcc acc gtg      912
Pro Glu Glu Glu His Glu Ser Pro Arg Ala Glu Thr Pro Ser Thr Val
    290                 295                 300

act gaa gtg gac atg gac ctg gac tcc tac cag atc gct ctg gaa gag      960
Thr Glu Val Asp Met Asp Leu Asp Ser Tyr Gln Ile Ala Leu Glu Glu
305                 310                 315                 320

gtg ctg act tgg ctg ctg tcc gcc gag gat act ttc cag gaa cag gac     1008
Val Leu Thr Trp Leu Leu Ser Ala Glu Asp Thr Phe Gln Glu Gln Asp
                325                 330                 335

gac atc tcc gac gac gtc gag gaa gtg aag gac cag ttc gcc act cac     1056
Asp Ile Ser Asp Asp Val Glu Glu Val Lys Asp Gln Phe Ala Thr His
            340                 345                 350

gag gct ttc atg atg gaa ctg act gcc cac cag tcc tcc gtg gga agt     1104
Glu Ala Phe Met Met Glu Leu Thr Ala His Gln Ser Ser Val Gly Ser
        355                 360                 365

gtc ctc cag gct gga aac cag ctg atc acc cag gga acc ctg tcc gat     1152
Val Leu Gln Ala Gly Asn Gln Leu Ile Thr Gln Gly Thr Leu Ser Asp
    370                 375                 380

gag gaa gag ttc gag att cag gaa cag atg acc ctg ctg aac gcc aga     1200
Glu Glu Glu Phe Glu Ile Gln Glu Gln Met Thr Leu Leu Asn Ala Arg
385                 390                 395                 400

tgg gaa gct ctg agg gtc gag tcc atg gat agg cag tct agg ctg cac     1248
Trp Glu Ala Leu Arg Val Glu Ser Met Asp Arg Gln Ser Arg Leu His
                405                 410                 415

gac gtg ctc atg gaa ctc cag aag aag cag ctc cag cag ctg tcc gct     1296
Asp Val Leu Met Glu Leu Gln Lys Lys Gln Leu Gln Gln Leu Ser Ala
            420                 425                 430

tgg atc cag gaa gcc gag act acc gtc aat gtg ctg gtg gac gcc tct     1344
Trp Ile Gln Glu Ala Glu Thr Thr Val Asn Val Leu Val Asp Ala Ser
        435                 440                 445

cac aga gag aac gcc ctc cag gat tcc atc ctg gct agg gaa ctg aag     1392
His Arg Glu Asn Ala Leu Gln Asp Ser Ile Leu Ala Arg Glu Leu Lys
    450                 455                 460
```

[0165]

```
cag cag atg cag gac atc cag gcc gag atc gac gct cac cag gtc aaa      1440
Gln Gln Met Gln Asp Ile Gln Ala Glu Ile Asp Ala His Gln Val Lys
465                 470                 475                 480

gtg aac tcc ctg act cac atg gtg gtc atc gtg gac gag aat tcc ggc      1488
Val Asn Ser Leu Thr His Met Val Val Ile Val Asp Glu Asn Ser Gly
                485                 490                 495

gag tct gcc acc gct atc ctg gaa gat cag ctc cag aaa ctg ggc gag      1536
Glu Ser Ala Thr Ala Ile Leu Glu Asp Gln Leu Gln Lys Leu Gly Glu
            500                 505                 510

cgc tgg aac gat ctg aag gct aag tcc gcc tct atc agg gcc cat ctg      1584
Arg Trp Asn Asp Leu Lys Ala Lys Ser Ala Ser Ile Arg Ala His Leu
        515                 520                 525

gaa gcc tct gcc gag aag tgg aat agg ctg ctg atg tcc ctg gaa gaa      1632
Glu Ala Ser Ala Glu Lys Trp Asn Arg Leu Leu Met Ser Leu Glu Glu
    530                 535                 540

ctg atc aag tgg ctg aac atg aag gat gag gaa ctc aag aaa cag atg      1680
Leu Ile Lys Trp Leu Asn Met Lys Asp Glu Glu Leu Lys Lys Gln Met
545                 550                 555                 560

ccc atc gga ggc gac gtg cca gct ctc cag ctc cag tat gat cat tgc      1728
Pro Ile Gly Gly Asp Val Pro Ala Leu Gln Leu Gln Tyr Asp His Cys
                565                 570                 575
```

[0166]

```
aag gct ctg agg cgc gag ctg aaa gag aaa gag tac tcc gtc ctg aac      1776
Lys Ala Leu Arg Arg Glu Leu Lys Glu Lys Glu Tyr Ser Val Leu Asn
            580                 585                 590

gcc gtg gat cag gct agg gtg ttc ctg gcc gat cag cct att gag gct      1824
Ala Val Asp Gln Ala Arg Val Phe Leu Ala Asp Gln Pro Ile Glu Ala
        595                 600                 605

cct gag gaa cct agg cgg aac ctc cag tcc aag acc gag ctg act cct      1872
Pro Glu Glu Pro Arg Arg Asn Leu Gln Ser Lys Thr Glu Leu Thr Pro
    610                 615                 620

gaa gag agg gcc cag aaa atc gcc aag gct atg cgc aag cag tcc agc      1920
Glu Glu Arg Ala Gln Lys Ile Ala Lys Ala Met Arg Lys Gln Ser Ser
625                 630                 635                 640

gaa gtc aaa gag aag tgg gag tcc ctg aat gcc gtg acc tcc aac tgg      1968
Glu Val Lys Glu Lys Trp Glu Ser Leu Asn Ala Val Thr Ser Asn Trp
                645                 650                 655

cag aaa cag gtg gac aag gcc ctg gaa aag ctg agg gat ctc cag ggc      2016
Gln Lys Gln Val Asp Lys Ala Leu Glu Lys Leu Arg Asp Leu Gln Gly
            660                 665                 670

gct atg gac gac ctg gac gcc gat atg aag gaa gct gag tcc gtc aga      2064
Ala Met Asp Asp Leu Asp Ala Asp Met Lys Glu Ala Glu Ser Val Arg
        675                 680                 685

aac ggc tgg aag cca gtg gga gat ctg ctg att gat agc ctc cag gac      2112
```

```
Asn Gly Trp Lys Pro Val Gly Asp Leu Leu Ile Asp Ser Leu Gln Asp
    690                 695                 700

cac att gag aag atc atg gcc ttc agg gaa gag atc gcc cct atc aac      2160
His Ile Glu Lys Ile Met Ala Phe Arg Glu Glu Ile Ala Pro Ile Asn
705                 710                 715                 720

ttc aaa gtc aag acc gtg aac gac ctg tcc agc cag ctg agc cct ctg      2208
Phe Lys Val Lys Thr Val Asn Asp Leu Ser Ser Gln Leu Ser Pro Leu
                725                 730                 735

gat ctc cat cca agc ctg aag atg agc agg cag ctg gat gac ctg aac      2256
Asp Leu His Pro Ser Leu Lys Met Ser Arg Gln Leu Asp Asp Leu Asn
            740                 745                 750

atg cgc tgg aaa ctc ctc cag gtg tcc gtg gac gat agg ctg aaa cag      2304
Met Arg Trp Lys Leu Leu Gln Val Ser Val Asp Asp Arg Leu Lys Gln
        755                 760                 765

ctc cag gaa gct cac aga gac ttc ggc cct agc tct cag cac ttc ctg      2352
Leu Gln Glu Ala His Arg Asp Phe Gly Pro Ser Ser Gln His Phe Leu
    770                 775                 780

tct acc tct gtg cag ctc cct tgg cag cgc tcc atc tct cac aac aag      2400
Ser Thr Ser Val Gln Leu Pro Trp Gln Arg Ser Ile Ser His Asn Lys
785                 790                 795                 800

gtg ccc tac tac atc aac cac cag act cag acc acc tgt tgg gat cac      2448
Val Pro Tyr Tyr Ile Asn His Gln Thr Gln Thr Thr Cys Trp Asp His
                805                 810                 815

ccc aag atg act gag ctg ttc cag agc ctg gcc gac ctc aac aat gtg      2496
Pro Lys Met Thr Glu Leu Phe Gln Ser Leu Ala Asp Leu Asn Asn Val
            820                 825                 830

cgc ttc tcc gcc tat agg acc gcc atc aaa atc agg cgc ctc cag aaa      2544
Arg Phe Ser Ala Tyr Arg Thr Ala Ile Lys Ile Arg Arg Leu Gln Lys
        835                 840                 845

gcc ctg tgc ctg gac ctg ctc gaa ctg tcc acc acc aac gag atc ttc      2592
Ala Leu Cys Leu Asp Leu Leu Glu Leu Ser Thr Thr Asn Glu Ile Phe
    850                 855                 860

aag cag cac aag ctg aac cag aac gac cag ctg ctg tct gtg ccc gac      2640
Lys Gln His Lys Leu Asn Gln Asn Asp Gln Leu Leu Ser Val Pro Asp
865                 870                 875                 880

gtg atc aat tgc ctg acc acc acc tac gac ggg ctc gaa cag atg cac      2688
Val Ile Asn Cys Leu Thr Thr Thr Tyr Asp Gly Leu Glu Gln Met His
                885                 890                 895

aag gat ctc gtg aat gtg cct ctg tgc gtg gac atg tgc ctg aac tgg      2736
Lys Asp Leu Val Asn Val Pro Leu Cys Val Asp Met Cys Leu Asn Trp
            900                 905                 910

ctg ctg aat gtg tac gat acc ggc agg acc ggc aag atc agg gtg cag      2784
Leu Leu Asn Val Tyr Asp Thr Gly Arg Thr Gly Lys Ile Arg Val Gln
        915                 920                 925
```

[0167]

```
tct ctg aag atc ggc ctg atg agc ctg agc aag ggg ctg ctc gaa gag      2832
Ser Leu Lys Ile Gly Leu Met Ser Leu Ser Lys Gly Leu Leu Glu Glu
    930                 935                 940

aag tac cgc tac ctg ttc aaa gag gtg gca ggc cct acc gag atg tgc      2880
Lys Tyr Arg Tyr Leu Phe Lys Glu Val Ala Gly Pro Thr Glu Met Cys
945                 950                 955                 960

gat cag aga cag ctg gga ctg ctg ctg cac gat gcc att cag att cct      2928
Asp Gln Arg Gln Leu Gly Leu Leu Leu His Asp Ala Ile Gln Ile Pro
                965                 970                 975

cgc cag ctg gga gaa gtg gcc gcc ttc gga gga tct aat atc gag cct      2976
Arg Gln Leu Gly Glu Val Ala Ala Phe Gly Gly Ser Asn Ile Glu Pro
            980                 985                 990

tcc gtg cgc tcc tgc ttt cag cag  aac aac aac aag cct  gag atc tcc    3024
Ser Val Arg Ser Cys Phe Gln Gln  Asn Asn Asn Lys Pro  Glu Ile Ser
        995                 1000                1005

gtg aaa  gaa ttc atc gac tgg  atg cac ctg gaa cct  cag agc atg       3069
Val Lys  Glu Phe Ile Asp Trp  Met His Leu Glu Pro  Gln Ser Met
    1010                1015                1020

gtc tgg  ctg cca gtc ctg cat  aga gtg gct gct gct  gag act gcc       3114
Val Trp  Leu Pro Val Leu His  Arg Val Ala Ala Ala  Glu Thr Ala
    1025                1030                1035
```

[0168]

```
aag cac  cag gcc aag tgc aac  atc tgc aaa gag tgc  cct atc gtc       3159
Lys His  Gln Ala Lys Cys Asn  Ile Cys Lys Glu Cys  Pro Ile Val
    1040                1045                1050

ggc ttc  cgc tat cgc tcc ctg  aag cac ttc aac tac  gac gtg tgc       3204
Gly Phe  Arg Tyr Arg Ser Leu  Lys His Phe Asn Tyr  Asp Val Cys
    1055                1060                1065

cag agt  tgc ttc ttt tcc gga  cgc acc gcc aag ggc  cac aag ctg       3249
Gln Ser  Cys Phe Phe Ser Gly  Arg Thr Ala Lys Gly  His Lys Leu
    1070                1075                1080

cat tac  cct atg gtc gag tac  tgc atc cct acc aca  tcc ggc gag       3294
His Tyr  Pro Met Val Glu Tyr  Cys Ile Pro Thr Thr  Ser Gly Glu
    1085                1090                1095

gat gtg  cgc gat ttc acc aag  gtg ctg aag aac aag  ttc cgc tcc       3339
Asp Val  Arg Asp Phe Thr Lys  Val Leu Lys Asn Lys  Phe Arg Ser
    1100                1105                1110

aag aag  tat ttc gct aag cac  ccc agg ctg ggc tat  ctg cct gtc       3384
Lys Lys  Tyr Phe Ala Lys His  Pro Arg Leu Gly Tyr  Leu Pro Val
    1115                1120                1125

cag act  gtc ctg gaa ggc gac  aac ctg gaa acc tga                   3420
Gln Thr  Val Leu Glu Gly Asp  Asn Leu Glu Thr
    1130                1135
```

<210> 20
<211> 1139
<212> PRT
<213> 人工序列

<220>
<223> 合成的构建体

<400> 20

Met Ala Lys Tyr Gly Glu His Glu Ala Ser Pro Asp Asn Gly Gln Asn
1 5 10 15

Glu Phe Ser Asp Ile Ile Lys Ser Arg Ser Asp Glu His Asn Asp Val
20 25 30

Gln Lys Lys Thr Phe Thr Lys Trp Ile Asn Ala Arg Phe Ser Lys Ser
35 40 45

Gly Lys Pro Pro Ile Asn Asp Met Phe Thr Asp Leu Lys Asp Gly Arg
50 55 60

Lys Leu Leu Asp Leu Leu Glu Gly Leu Thr Gly Thr Ser Leu Pro Lys
65 70 75 80

[0169]

Glu Arg Gly Ser Thr Arg Val His Ala Leu Asn Asn Val Asn Arg Val
85 90 95

Leu Gln Val Leu His Gln Asn Asn Val Glu Leu Val Asn Ile Gly Gly
100 105 110

Thr Asp Ile Val Asp Gly Asn His Lys Leu Thr Leu Gly Leu Leu Trp
115 120 125

Ser Ile Ile Leu His Trp Gln Val Lys Asp Val Met Lys Asp Val Met
130 135 140

Ser Asp Leu Gln Gln Thr Asn Ser Glu Lys Ile Leu Leu Ser Trp Val
145 150 155 160

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

Thr Ser Trp Thr Asp Gly Leu Ala Phe Asn Ala Val Leu His Arg His
180 185 190

Lys Pro Asp Leu Phe Ser Trp Asp Lys Val Val Lys Met Ser Pro Ile
195 200 205

Glu Arg Leu Glu His Ala Phe Ser Lys Ala Gln Thr Tyr Leu Gly Ile
210 215 220

Glu Lys Leu Leu Asp Pro Glu Asp Val Ala Val Gln Leu Pro Asp Lys
225 230 235 240

Lys Ser Ile Ile Met Tyr Leu Thr Ser Leu Phe Glu Val Leu Pro Gln
245 250 255

Gln Val Thr Ile Asp Ala Ile Arg Glu Val Glu Thr Leu Pro Arg Lys
260 265 270

Tyr Lys Lys Glu Cys Glu Glu Glu Ala Ile Asn Ile Gln Ser Thr Ala
275 280 285

Pro Glu Glu Glu His Glu Ser Pro Arg Ala Glu Thr Pro Ser Thr Val
290 295 300

[0170] Thr Glu Val Asp Met Asp Leu Asp Ser Tyr Gln Ile Ala Leu Glu Glu
305 310 315 320

Val Leu Thr Trp Leu Leu Ser Ala Glu Asp Thr Phe Gln Glu Gln Asp
325 330 335

Asp Ile Ser Asp Asp Val Glu Glu Val Lys Asp Gln Phe Ala Thr His
340 345 350

Glu Ala Phe Met Met Glu Leu Thr Ala His Gln Ser Ser Val Gly Ser
355 360 365

Val Leu Gln Ala Gly Asn Gln Leu Ile Thr Gln Gly Thr Leu Ser Asp
370 375 380

Glu Glu Glu Phe Glu Ile Gln Glu Gln Met Thr Leu Leu Asn Ala Arg
385 390 395 400

Trp Glu Ala Leu Arg Val Glu Ser Met Asp Arg Gln Ser Arg Leu His
405 410 415

Asp Val Leu Met Glu Leu Gln Lys Lys Gln Leu Gln Gln Leu Ser Ala
420 425 430

```
Trp Ile Gln Glu Ala Glu Thr Thr Val Asn Val Leu Val Asp Ala Ser
        435                 440                 445

His Arg Glu Asn Ala Leu Gln Asp Ser Ile Leu Ala Arg Glu Leu Lys
    450                 455                 460

Gln Gln Met Gln Asp Ile Gln Ala Glu Ile Asp Ala His Gln Val Lys
465                 470                 475                 480

Val Asn Ser Leu Thr His Met Val Val Ile Val Asp Glu Asn Ser Gly
                485                 490                 495

Glu Ser Ala Thr Ala Ile Leu Glu Asp Gln Leu Gln Lys Leu Gly Glu
            500                 505                 510

Arg Trp Asn Asp Leu Lys Ala Lys Ser Ala Ser Ile Arg Ala His Leu
        515                 520                 525

Glu Ala Ser Ala Glu Lys Trp Asn Arg Leu Leu Met Ser Leu Glu Glu
    530                 535                 540
```

[0171]

```
Leu Ile Lys Trp Leu Asn Met Lys Asp Glu Glu Leu Lys Lys Gln Met
545                 550                 555                 560

Pro Ile Gly Gly Asp Val Pro Ala Leu Gln Leu Gln Tyr Asp His Cys
                565                 570                 575

Lys Ala Leu Arg Arg Glu Leu Lys Glu Lys Glu Tyr Ser Val Leu Asn
            580                 585                 590

Ala Val Asp Gln Ala Arg Val Phe Leu Ala Asp Gln Pro Ile Glu Ala
        595                 600                 605

Pro Glu Glu Pro Arg Arg Asn Leu Gln Ser Lys Thr Glu Leu Thr Pro
    610                 615                 620

Glu Glu Arg Ala Gln Lys Ile Ala Lys Ala Met Arg Lys Gln Ser Ser
625                 630                 635                 640

Glu Val Lys Glu Lys Trp Glu Ser Leu Asn Ala Val Thr Ser Asn Trp
                645                 650                 655
```

Gln Lys Gln Val Asp Lys Ala Leu Glu Lys Leu Arg Asp Leu Gln Gly
660 665 670

Ala Met Asp Asp Leu Asp Ala Asp Met Lys Glu Ala Glu Ser Val Arg
675 680 685

Asn Gly Trp Lys Pro Val Gly Asp Leu Leu Ile Asp Ser Leu Gln Asp
690 695 700

His Ile Glu Lys Ile Met Ala Phe Arg Glu Glu Ile Ala Pro Ile Asn
705 710 715 720

Phe Lys Val Lys Thr Val Asn Asp Leu Ser Ser Gln Leu Ser Pro Leu
725 730 735

Asp Leu His Pro Ser Leu Lys Met Ser Arg Gln Leu Asp Asp Leu Asn
740 745 750

Met Arg Trp Lys Leu Leu Gln Val Ser Val Asp Asp Arg Leu Lys Gln
755 760 765

[0172] Leu Gln Glu Ala His Arg Asp Phe Gly Pro Ser Ser Gln His Phe Leu
770 775 780

Ser Thr Ser Val Gln Leu Pro Trp Gln Arg Ser Ile Ser His Asn Lys
785 790 795 800

Val Pro Tyr Tyr Ile Asn His Gln Thr Gln Thr Thr Cys Trp Asp His
805 810 815

Pro Lys Met Thr Glu Leu Phe Gln Ser Leu Ala Asp Leu Asn Asn Val
820 825 830

Arg Phe Ser Ala Tyr Arg Thr Ala Ile Lys Ile Arg Arg Leu Gln Lys
835 840 845

Ala Leu Cys Leu Asp Leu Leu Glu Leu Ser Thr Thr Asn Glu Ile Phe
850 855 860

Lys Gln His Lys Leu Asn Gln Asn Asp Gln Leu Leu Ser Val Pro Asp
865 870 875 880

Val Ile Asn Cys Leu Thr Thr Thr Tyr Asp Gly Leu Glu Gln Met His
885 890 895

```
Lys Asp Leu Val Asn Val Pro Leu Cys Val Asp Met Cys Leu Asn Trp
            900                 905                 910

Leu Leu Asn Val Tyr Asp Thr Gly Arg Thr Gly Lys Ile Arg Val Gln
        915                 920                 925

Ser Leu Lys Ile Gly Leu Met Ser Leu Ser Lys Gly Leu Leu Glu Glu
    930                 935                 940

Lys Tyr Arg Tyr Leu Phe Lys Glu Val Ala Gly Pro Thr Glu Met Cys
945                 950                 955                 960

Asp Gln Arg Gln Leu Gly Leu Leu Leu His Asp Ala Ile Gln Ile Pro
                965                 970                 975

Arg Gln Leu Gly Glu Val Ala Ala Phe Gly Gly Ser Asn Ile Glu Pro
            980                 985                 990

Ser Val Arg Ser Cys Phe Gln Gln  Asn Asn Asn Lys Pro  Glu Ile Ser
        995                 1000                1005
```

[0173]

```
Val Lys  Glu Phe Ile Asp Trp  Met His Leu Glu Pro  Gln Ser Met
    1010                1015                1020

Val Trp  Leu Pro Val Leu His  Arg Val Ala Ala Ala  Glu Thr Ala
    1025                1030                1035

Lys His  Gln Ala Lys Cys Asn  Ile Cys Lys Glu Cys  Pro Ile Val
    1040                1045                1050

Gly Phe  Arg Tyr Arg Ser Leu  Lys His Phe Asn Tyr  Asp Val Cys
    1055                1060                1065

Gln Ser  Cys Phe Phe Ser Gly  Arg Thr Ala Lys Gly  His Lys Leu
    1070                1075                1080

His Tyr  Pro Met Val Glu Tyr  Cys Ile Pro Thr Thr  Ser Gly Glu
    1085                1090                1095

Asp Val  Arg Asp Phe Thr Lys  Val Leu Lys Asn Lys  Phe Arg Ser
    1100                1105                1110
```

Lys Lys Tyr Phe Ala Lys His Pro Arg Leu Gly Tyr Leu Pro Val
1115 1120 1125

Gln Thr Val Leu Glu Gly Asp Asn Leu Glu Thr
1130 1135

<210> 21
<211> 1208
<212> PRT
<213> 人工序列

<220>
<223> 人肌养蛋白相关蛋白突变体

<220>
<221> MISC_FEATURE
<222> (1)..(324)
<223> 与全长人肌养蛋白相关蛋白的N-末端区域相同的氨基酸序列

<220>
<221> MISC_FEATURE
<222> (311)..(417)
<223> 杂合三股螺旋

[0174]

<220>
<221> MISC_FEATURE
<222> (325)..(362)
<223> 与全长人肌养蛋白相关蛋白的C-末端区域相同的氨基酸序列

<220>
<221> MISC_FEATURE
<222> (363)..(401)
<223> 与全长人肌养蛋白相关蛋白的N-末端区域相同的氨基酸序列

<220>
<221> MISC_FEATURE
<222> (402)..(1208)
<223> 与全长人肌养蛋白相关蛋白的C-末端区域相同的氨基酸序列

<400> 21

Met Ala Lys Tyr Gly Glu His Glu Ala Ser Pro Asp Asn Gly Gln Asn
1 5 10 15

Glu Phe Ser Asp Ile Ile Lys Ser Arg Ser Asp Glu His Asn Asp Val
20 25 30

Gln Lys Lys Thr Phe Thr Lys Trp Ile Asn Ala Arg Phe Ser Lys Ser
35 40 45

Gly Lys Pro Pro Ile Asn Asp Met Phe Thr Asp Leu Lys Asp Gly Arg
50 55 60

```
Lys Leu Leu Asp Leu Leu Glu Gly Leu Thr Gly Thr Ser Leu Pro Lys
65                  70                  75                  80

Glu Arg Gly Ser Thr Arg Val His Ala Leu Asn Asn Val Asn Arg Val
                85                  90                  95

Leu Gln Val Leu His Gln Asn Asn Val Glu Leu Val Asn Ile Gly Gly
            100                 105                 110

Thr Asp Ile Val Asp Gly Asn His Lys Leu Thr Leu Gly Leu Leu Trp
        115                 120                 125

Ser Ile Ile Leu His Trp Gln Val Lys Asp Val Met Lys Asp Val Met
    130                 135                 140

Ser Asp Leu Gln Gln Thr Asn Ser Glu Lys Ile Leu Leu Ser Trp Val
145                 150                 155                 160

Arg Gln Thr Thr Arg Pro Tyr Ser Gln Val Asn Val Leu Asn Phe Thr
                165                 170                 175
```

[0175]

```
Thr Ser Trp Thr Asp Gly Leu Ala Phe Asn Ala Val Leu His Arg His
            180                 185                 190

Lys Pro Asp Leu Phe Ser Trp Asp Lys Val Val Lys Met Ser Pro Ile
        195                 200                 205

Glu Arg Leu Glu His Ala Phe Ser Lys Ala Gln Thr Tyr Leu Gly Ile
    210                 215                 220

Glu Lys Leu Leu Asp Pro Glu Asp Val Ala Val Gln Leu Pro Asp Lys
225                 230                 235                 240

Lys Ser Ile Ile Met Tyr Leu Thr Ser Leu Phe Glu Val Leu Pro Gln
                245                 250                 255

Gln Val Thr Ile Asp Ala Ile Arg Glu Val Glu Thr Leu Pro Arg Lys
            260                 265                 270

Tyr Lys Lys Glu Cys Glu Glu Glu Ala Ile Asn Ile Gln Ser Thr Ala
        275                 280                 285
```

Pro Glu Glu Glu His Glu Ser Pro Arg Ala Glu Thr Pro Ser Thr Val
290 295 300

Thr Glu Val Asp Met Asp Leu Asp Ser Tyr Gln Ile Ala Leu Glu Glu
305 310 315 320

Val Leu Thr Trp Leu Val Leu Ile Asp Gln Met Leu Lys Ser Asn Ile
325 330 335

Val Thr Val Gly Asp Val Glu Glu Ile Asn Lys Thr Val Ser Arg Met
340 345 350

Lys Ile Thr Lys Ala Asp Leu Glu Gln Arg Gln Ser Ser Val Gly Ser
355 360 365

Val Leu Gln Ala Gly Asn Gln Leu Ile Thr Gln Gly Thr Leu Ser Asp
370 375 380

Glu Glu Glu Phe Glu Ile Gln Glu Gln Met Thr Leu Leu Asn Ala Arg
385 390 395 400

[0176] Trp Asp Gly Thr Gln His Gly Val Glu Leu Arg Gln Gln Gln Leu Glu
405 410 415

Asp Met Ile Ile Asp Ser Leu Gln Trp Asp Asp His Arg Glu Glu Thr
420 425 430

Glu Glu Leu Met Arg Lys Tyr Glu Ala Arg Leu Tyr Ile Leu Gln Gln
435 440 445

Ala Arg Arg Asp Pro Leu Thr Lys Gln Ile Ser Asp Asn Gln Ile Leu
450 455 460

Leu Gln Glu Leu Gly Pro Gly Asp Gly Ile Val Met Ala Phe Asp Asn
465 470 475 480

Val Leu Gln Lys Leu Leu Glu Glu Tyr Gly Ser Asp Asp Thr Arg Asn
485 490 495

Val Lys Glu Thr Thr Glu Tyr Leu Lys Thr Ser Trp Ile Asn Leu Lys
500 505 510

Gln Ser Ile Ala Asp Arg Gln Asn Ala Leu Glu Ala Glu Trp Arg Thr
515 520 525

```
Val Gln Ala Ser Arg Arg Asp Leu Glu Asn Phe Leu Lys Trp Ile Gln
    530                 535                 540

Glu Ala Glu Thr Thr Val Asn Val Leu Val Asp Ala Ser His Arg Glu
545                 550                 555                 560

Asn Ala Leu Gln Asp Ser Ile Leu Ala Arg Glu Leu Lys Gln Gln Met
                565                 570                 575

Gln Asp Ile Gln Ala Glu Ile Asp Ala His Asn Asp Ile Phe Lys Ser
            580                 585                 590

Ile Asp Gly Asn Arg Gln Lys Met Val Lys Ala Leu Gly Asn Ser Glu
        595                 600                 605

Glu Ala Thr Met Leu Gln His Arg Leu Asp Asp Met Asn Gln Arg Trp
    610                 615                 620

Asn Asp Leu Lys Ala Lys Ser Ala Ser Ile Arg Ala His Leu Glu Ala
625                 630                 635                 640
```

[0177]

```
Ser Ala Glu Lys Trp Asn Arg Leu Leu Met Ser Leu Glu Glu Leu Ile
                645                 650                 655

Lys Trp Leu Asn Met Lys Asp Glu Glu Leu Lys Lys Gln Met Pro Ile
            660                 665                 670

Gly Gly Asp Val Pro Ala Leu Gln Leu Gln Tyr Asp His Cys Lys Ala
        675                 680                 685

Leu Arg Arg Glu Leu Lys Glu Lys Glu Tyr Ser Val Leu Asn Ala Val
    690                 695                 700

Asp Gln Ala Arg Val Phe Leu Ala Asp Gln Pro Ile Glu Ala Pro Glu
705                 710                 715                 720

Glu Pro Arg Arg Asn Leu Gln Ser Lys Thr Glu Leu Thr Pro Glu Glu
                725                 730                 735

Arg Ala Gln Lys Ile Ala Lys Ala Met Arg Lys Gln Ser Ser Glu Val
            740                 745                 750
```

```
Lys Glu Lys Trp Glu Ser Leu Asn Ala Val Thr Ser Asn Trp Gln Lys
        755                 760                 765

Gln Val Asp Lys Ala Leu Glu Lys Leu Arg Asp Leu Gln Gly Ala Met
    770                 775                 780

Asp Asp Leu Asp Ala Asp Met Lys Glu Ala Glu Ser Val Arg Asn Gly
785                 790                 795                 800

Trp Lys Pro Val Gly Asp Leu Leu Ile Asp Ser Leu Gln Asp His Ile
                805                 810                 815

Glu Lys Ile Met Ala Phe Arg Glu Glu Ile Ala Pro Ile Asn Phe Lys
            820                 825                 830

Val Lys Thr Val Asn Asp Leu Ser Ser Gln Leu Ser Pro Leu Asp Leu
        835                 840                 845

His Pro Ser Leu Lys Met Ser Arg Gln Leu Asp Asp Leu Asn Met Arg
    850                 855                 860
```

[0178]

```
Trp Lys Leu Leu Gln Val Ser Val Asp Asp Arg Leu Lys Gln Leu Gln
865                 870                 875                 880

Glu Ala His Arg Asp Phe Gly Pro Ser Ser Gln His Phe Leu Ser Thr
                885                 890                 895

Ser Val Gln Leu Pro Trp Gln Arg Ser Ile Ser His Asn Lys Val Pro
            900                 905                 910

Tyr Tyr Ile Asn His Gln Thr Gln Thr Thr Cys Trp Asp His Pro Lys
        915                 920                 925

Met Thr Glu Leu Phe Gln Ser Leu Ala Asp Leu Asn Asn Val Arg Phe
    930                 935                 940

Ser Ala Tyr Arg Thr Ala Ile Lys Ile Arg Arg Leu Gln Lys Ala Leu
945                 950                 955                 960

Cys Leu Asp Leu Leu Glu Leu Ser Thr Thr Asn Glu Ile Phe Lys Gln
                965                 970                 975

His Lys Leu Asn Gln Asn Asp Gln Leu Leu Ser Val Pro Asp Val Ile
            980                 985                 990
```

```
Asn Cys Leu Thr Thr Thr Tyr Asp  Gly Leu Glu Gln Met  His Lys Asp
        995                 1000                1005

Leu Val  Asn Val Pro Leu Cys  Val Asp Met Cys Leu  Asn Trp Leu
    1010                1015                1020

Leu Asn  Val Tyr Asp Thr Gly  Arg Thr Gly Lys Ile  Arg Val Gln
    1025                1030                1035

Ser Leu  Lys Ile Gly Leu Met  Ser Leu Ser Lys Gly  Leu Leu Glu
    1040                1045                1050

Glu Lys  Tyr Arg Tyr Leu Phe  Lys Glu Val Ala Gly  Pro Thr Glu
    1055                1060                1065

Met Cys  Asp Gln Arg Gln Leu  Gly Leu Leu Leu His  Asp Ala Ile
    1070                1075                1080

Gln Ile  Pro Arg Gln Leu Gly  Glu Val Ala Ala Phe  Gly Gly Ser
    1085                1090                1095
```

[0179]

```
Asn Ile  Glu Pro Ser Val Arg  Ser Cys Phe Gln Gln  Asn Asn Asn
    1100                1105                1110

Lys Pro  Glu Ile Ser Val Lys  Glu Phe Ile Asp Trp  Met His Leu
    1115                1120                1125

Glu Pro  Gln Ser Met Val Trp  Leu Pro Val Leu His  Arg Val Ala
    1130                1135                1140

Ala Ala  Glu Thr Ala Lys His  Gln Ala Lys Cys Asn  Ile Cys Lys
    1145                1150                1155

Glu Cys  Pro Ile Val Gly Phe  Arg Tyr Arg Ser Leu  Lys His Phe
    1160                1165                1170

Asn Tyr  Asp Val Cys Gln Ser  Cys Phe Phe Ser Gly  Arg Thr Ala
    1175                1180                1185

Lys Gly  His Lys Leu His Tyr  Pro Met Val Glu Tyr  Cys Ile Pro
    1190                1195                1200
```

Thr Thr Ser Gly Glu
1205

<210> 22
<211> 1237
<212> PRT
<213> 人工序列

<220>
<223> 突变人肌养蛋白蛋白

<220>
<221> MISC_FEATURE
<222> (1)..(354)
<223> 与全长人肌养蛋白的N-末端区域相同的氨基酸序列

<220>
<221> MISC_FEATURE
<222> (340)..(447)
<223> 杂合三股螺旋

<220>
<221> MISC_FEATURE
<222> (355)..(392)
<223> 与全长人肌养蛋白的C-末端区域相同的氨基酸序列

[0180]

<220>
<221> MISC_FEATURE
<222> (393)..(431)
<223> 与全长人肌养蛋白的N-末端区域相同的氨基酸序列

<220>
<221> MISC_FEATURE
<222> (432)..(1237)
<223> 与全长人肌养蛋白的C-末端区域相同的氨基酸序列

<400> 22

Met Leu Trp Trp Glu Glu Val Glu Asp Cys Tyr Glu Arg Glu Asp Val
1 5 10 15

Gln Lys Lys Thr Phe Thr Lys Trp Val Asn Ala Gln Phe Ser Lys Phe
20 25 30

Gly Lys Gln His Ile Glu Asn Leu Phe Ser Asp Leu Gln Asp Gly Arg
35 40 45

Arg Leu Leu Asp Leu Leu Glu Gly Leu Thr Gly Gln Lys Leu Pro Lys
50 55 60

Glu Lys Gly Ser Thr Arg Val His Ala Leu Asn Asn Val Asn Lys Ala
65 70 75 80

Leu Arg Val Leu Gln Asn Asn Asn Val Asp Leu Val Asn Ile Gly Ser
85 90 95

Thr Asp Ile Val Asp Gly Asn His Lys Leu Thr Leu Gly Leu Ile Trp
100 105 110

Asn Ile Ile Leu His Trp Gln Val Lys Asn Val Met Lys Asn Ile Met
115 120 125

Ala Gly Leu Gln Gln Thr Asn Ser Glu Lys Ile Leu Leu Ser Trp Val
130 135 140

Arg Gln Ser Thr Arg Asn Tyr Pro Gln Val Asn Val Ile Asn Phe Thr
145 150 155 160

Thr Ser Trp Ser Asp Gly Leu Ala Leu Asn Ala Leu Ile His Ser His
165 170 175

Arg Pro Asp Leu Phe Asp Trp Asn Ser Val Val Cys Gln Gln Ser Ala
180 185 190

[0181]

Thr Gln Arg Leu Glu His Ala Phe Asn Ile Ala Arg Tyr Gln Leu Gly
195 200 205

Ile Glu Lys Leu Leu Asp Pro Glu Asp Val Asp Thr Thr Tyr Pro Asp
210 215 220

Lys Lys Ser Ile Leu Met Tyr Ile Thr Ser Leu Phe Gln Val Leu Pro
225 230 235 240

Gln Gln Val Ser Ile Glu Ala Ile Gln Glu Val Glu Met Leu Pro Arg
245 250 255

Pro Pro Lys Val Thr Lys Glu Glu His Phe Gln Leu His His Gln Met
260 265 270

His Tyr Ser Gln Gln Ile Thr Val Ser Leu Ala Gln Gly Tyr Glu Arg
275 280 285

Thr Ser Ser Pro Lys Pro Arg Phe Lys Ser Tyr Ala Tyr Thr Gln Ala
290 295 300

Ala Tyr Val Thr Thr Ser Asp Pro Thr Arg Ser Pro Phe Pro Ser Gln
305 310 315 320

His Leu Glu Ala Pro Glu Asp Lys Ser Phe Gly Ser Ser Leu Met Glu
325 330 335

Ser Glu Val Asn Leu Asp Arg Tyr Gln Thr Ala Leu Glu Glu Val Leu
340 345 350

Ser Trp Leu Ser Leu Leu Asp Gln Val Ile Lys Ser Gln Arg Val Met
355 360 365

Val Gly Asp Leu Glu Asp Ile Asn Glu Met Ile Ile Lys Gln Lys Ala
370 375 380

Thr Met Gln Asp Leu Glu Gln Arg Gln Gly Arg Val Gly Asn Ile Leu
385 390 395 400

Gln Leu Gly Ser Lys Leu Ile Gly Thr Gly Lys Leu Ser Glu Asp Glu
405 410 415

[0182] Glu Thr Glu Val Gln Glu Gln Met Asn Leu Leu Asn Ser Arg Trp Asp
420 425 430

Glu Val Gln Glu His Leu Gln Asn Arg Arg Gln Gln Leu Asn Glu Met
435 440 445

Leu Lys Asp Ser Thr Gln Trp Leu Glu Ala Lys Glu Glu Ala Glu Gln
450 455 460

Val Leu Gly Gln Ala Arg Ala Lys Leu Glu Ser Trp Lys Glu Gly Pro
465 470 475 480

Tyr Thr Val Asp Ala Ile Gln Lys Lys Ile Thr Glu Thr Lys Gln Leu
485 490 495

Ala Lys Asp Leu Arg Gln Trp Gln Thr Asn Val Asp Val Ala Asn Asp
500 505 510

Leu Ala Leu Lys Leu Leu Arg Asp Tyr Ser Ala Asp Asp Thr Arg Lys
515 520 525

Val His Met Ile Thr Glu Asn Ile Asn Ala Ser Trp Arg Ser Ile His
530 535 540

```
Lys Arg Val Ser Glu Arg Glu Ala Ala Leu Glu Glu Thr His Arg Leu
545                 550                 555                 560

Leu Gln Gln Phe Pro Leu Asp Leu Glu Lys Phe Leu Ala Trp Leu Thr
                565                 570                 575

Glu Ala Glu Thr Thr Ala Asn Val Leu Gln Asp Ala Thr Arg Lys Glu
            580                 585                 590

Arg Leu Leu Glu Asp Ser Lys Gly Val Lys Glu Leu Met Lys Gln Trp
        595                 600                 605

Gln Asp Leu Gln Gly Glu Ile Glu Ala His Thr Asp Val Tyr His Asn
    610                 615                 620

Leu Asp Glu Asn Ser Gln Lys Ile Leu Arg Ser Leu Glu Gly Ser Asp
625                 630                 635                 640

Asp Ala Val Leu Leu Gln Arg Arg Leu Asp Asn Met Asn Phe Lys Trp
                645                 650                 655
```

[0183]

```
Ser Glu Leu Arg Lys Lys Ser Leu Asn Ile Arg Ser His Leu Glu Ala
            660                 665                 670

Ser Ser Asp Gln Trp Lys Arg Leu His Leu Ser Leu Gln Glu Leu Leu
        675                 680                 685

Val Trp Leu Gln Leu Lys Asp Asp Glu Leu Ser Arg Gln Ala Pro Ile
    690                 695                 700

Gly Gly Asp Phe Pro Ala Val Gln Lys Gln Asn Asp Val His Arg Ala
705                 710                 715                 720

Phe Lys Arg Glu Leu Lys Thr Lys Glu Pro Val Ile Met Ser Thr Leu
                725                 730                 735

Glu Thr Val Arg Ile Phe Leu Thr Glu Gln Pro Leu Glu Gly Leu Glu
            740                 745                 750

Lys Leu Tyr Gln Glu Pro Arg Glu Leu Pro Pro Glu Glu Arg Ala Gln
        755                 760                 765
```

```
Asn Val Thr Arg Leu Leu Arg Lys Gln Ala Glu Glu Val Asn Thr Glu
    770                 775                 780

Trp Glu Lys Leu Asn Leu His Ser Ala Asp Trp Gln Arg Lys Ile Asp
785                 790                 795                 800

Glu Thr Leu Glu Arg Leu Gln Glu Leu Gln Glu Ala Thr Asp Glu Leu
                805                 810                 815

Asp Leu Lys Leu Arg Gln Ala Glu Val Ile Lys Gly Ser Trp Gln Pro
            820                 825                 830

Val Gly Asp Leu Leu Ile Asp Ser Leu Gln Asp His Leu Glu Lys Val
        835                 840                 845

Lys Ala Leu Arg Gly Glu Ile Ala Pro Leu Lys Glu Asn Val Ser His
    850                 855                 860

Val Asn Asp Leu Ala Arg Gln Leu Thr Thr Leu Gly Ile Gln Leu Ser
865                 870                 875                 880
```

[0184]

```
Pro Tyr Asn Leu Ser Thr Leu Glu Asp Leu Asn Thr Arg Trp Lys Leu
                885                 890                 895

Leu Gln Val Ala Val Glu Asp Arg Val Arg Gln Leu His Glu Ala His
            900                 905                 910

Arg Asp Phe Gly Pro Ala Ser Gln His Phe Leu Ser Thr Ser Val Gln
        915                 920                 925

Gly Pro Trp Glu Arg Ala Ile Ser Pro Asn Lys Val Pro Tyr Tyr Ile
    930                 935                 940

Asn His Glu Thr Gln Thr Thr Cys Trp Asp His Pro Lys Met Thr Glu
945                 950                 955                 960

Leu Tyr Gln Ser Leu Ala Asp Leu Asn Asn Val Arg Phe Ser Ala Tyr
                965                 970                 975

Arg Thr Ala Met Lys Leu Arg Arg Leu Gln Lys Ala Leu Cys Leu Asp
            980                 985                 990

Leu Leu Ser Leu Ser Ala Ala Cys  Asp Ala Leu Asp Gln  His Asn Leu
        995                 1000                1005
```

```
Lys Gln  Asn Asp Gln Pro Met  Asp Ile Leu Gln Ile  Ile Asn Cys
    1010                 1015                 1020

Leu Thr  Thr Ile Tyr Asp Arg  Leu Glu Gln Glu His  Asn Asn Leu
    1025                 1030                 1035

Val Asn  Val Pro Leu Cys Val  Asp Met Cys Leu Asn  Trp Leu Leu
    1040                 1045                 1050

Asn Val  Tyr Asp Thr Gly Arg  Thr Gly Arg Ile Arg  Val Leu Ser
    1055                 1060                 1065

Phe Lys  Thr Gly Ile Ile Ser  Leu Cys Lys Ala His  Leu Glu Asp
    1070                 1075                 1080

Lys Tyr  Arg Tyr Leu Phe Lys  Gln Val Ala Ser Ser  Thr Gly Phe
    1085                 1090                 1095

Cys Asp  Gln Arg Arg Leu Gly  Leu Leu Leu His Asp  Ser Ile Gln
    1100                 1105                 1110
```

[0185]

```
Ile Pro  Arg Gln Leu Gly Glu  Val Ala Ser Phe Gly  Gly Ser Asn
    1115                 1120                 1125

Ile Glu  Pro Ser Val Arg Ser  Cys Phe Gln Phe Ala  Asn Asn Lys
    1130                 1135                 1140

Pro Glu  Ile Glu Ala Ala Leu  Phe Leu Asp Trp Met  Arg Leu Glu
    1145                 1150                 1155

Pro Gln  Ser Met Val Trp Leu  Pro Val Leu His Arg  Val Ala Ala
    1160                 1165                 1170

Ala Glu  Thr Ala Lys His Gln  Ala Lys Cys Asn Ile  Cys Lys Glu
    1175                 1180                 1185

Cys Pro  Ile Ile Gly Phe Arg  Tyr Arg Ser Leu Lys  His Phe Asn
    1190                 1195                 1200

Tyr Asp  Ile Cys Gln Ser Cys  Phe Phe Ser Gly Arg  Val Ala Lys
    1205                 1210                 1215
```

Gly His Lys Met His Tyr Pro Met Val Glu Tyr Cys Thr Pro Thr
1220 1225 1230

[0186]

Thr Ser Gly Glu
1235

螺旋结构域内的剪接位点

螺旋结构域之间的剪接位点

图 1A

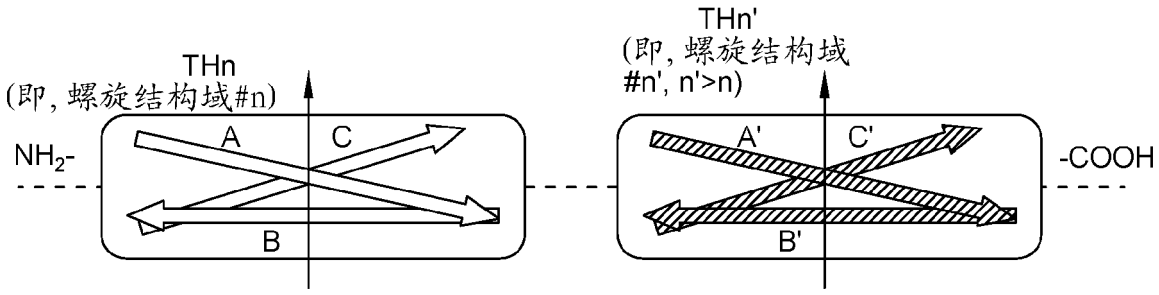

图 1B

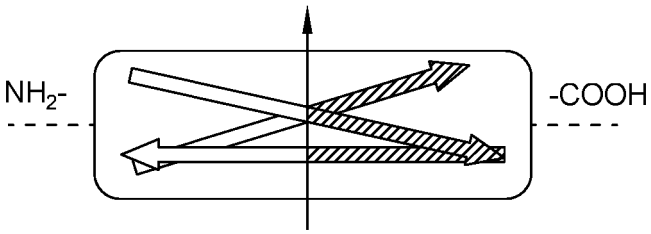

图 1C

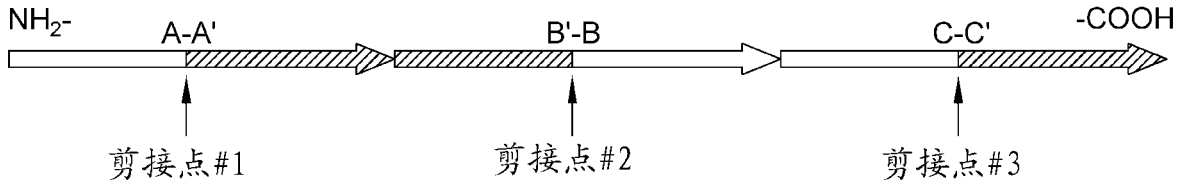

图 1D

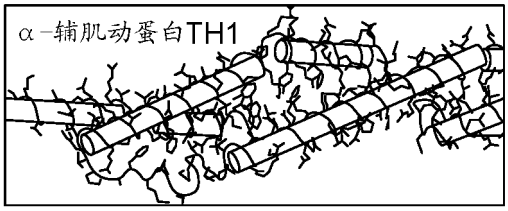

图 2A

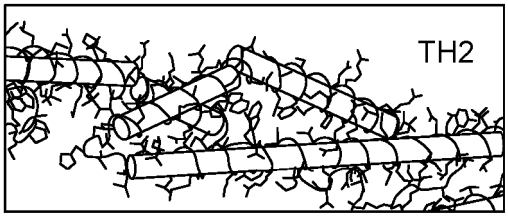

图 2B

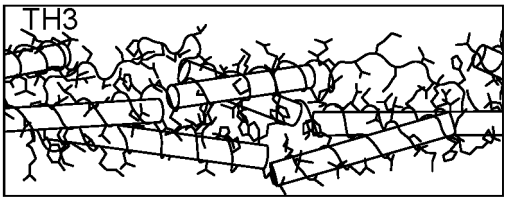

图 2C

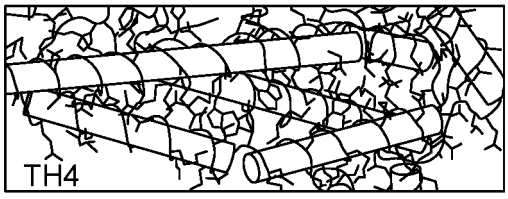

图 2D

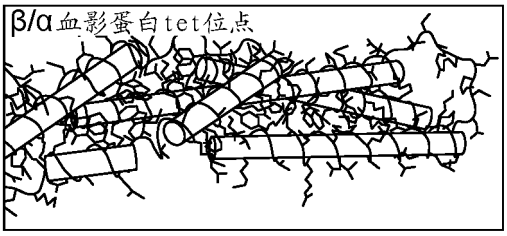

图 2E

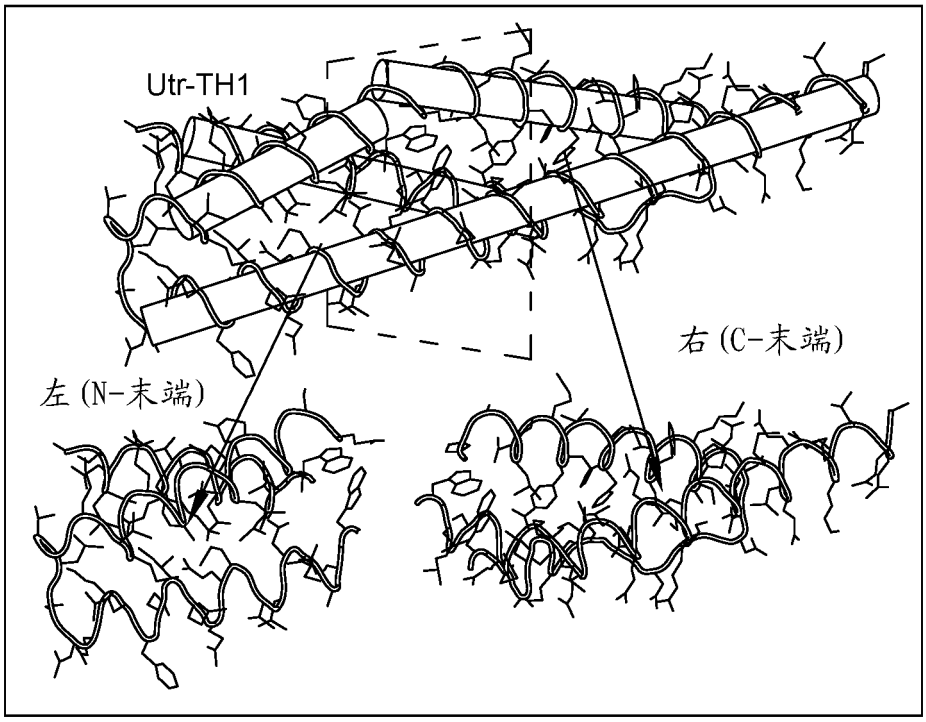

图 2F

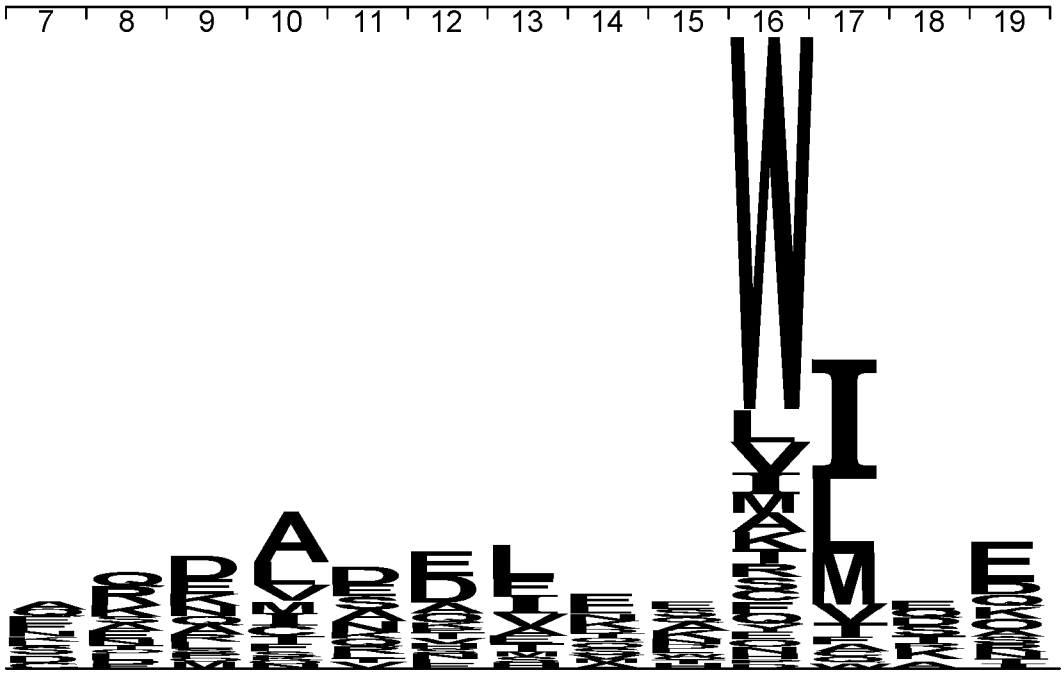

图 3

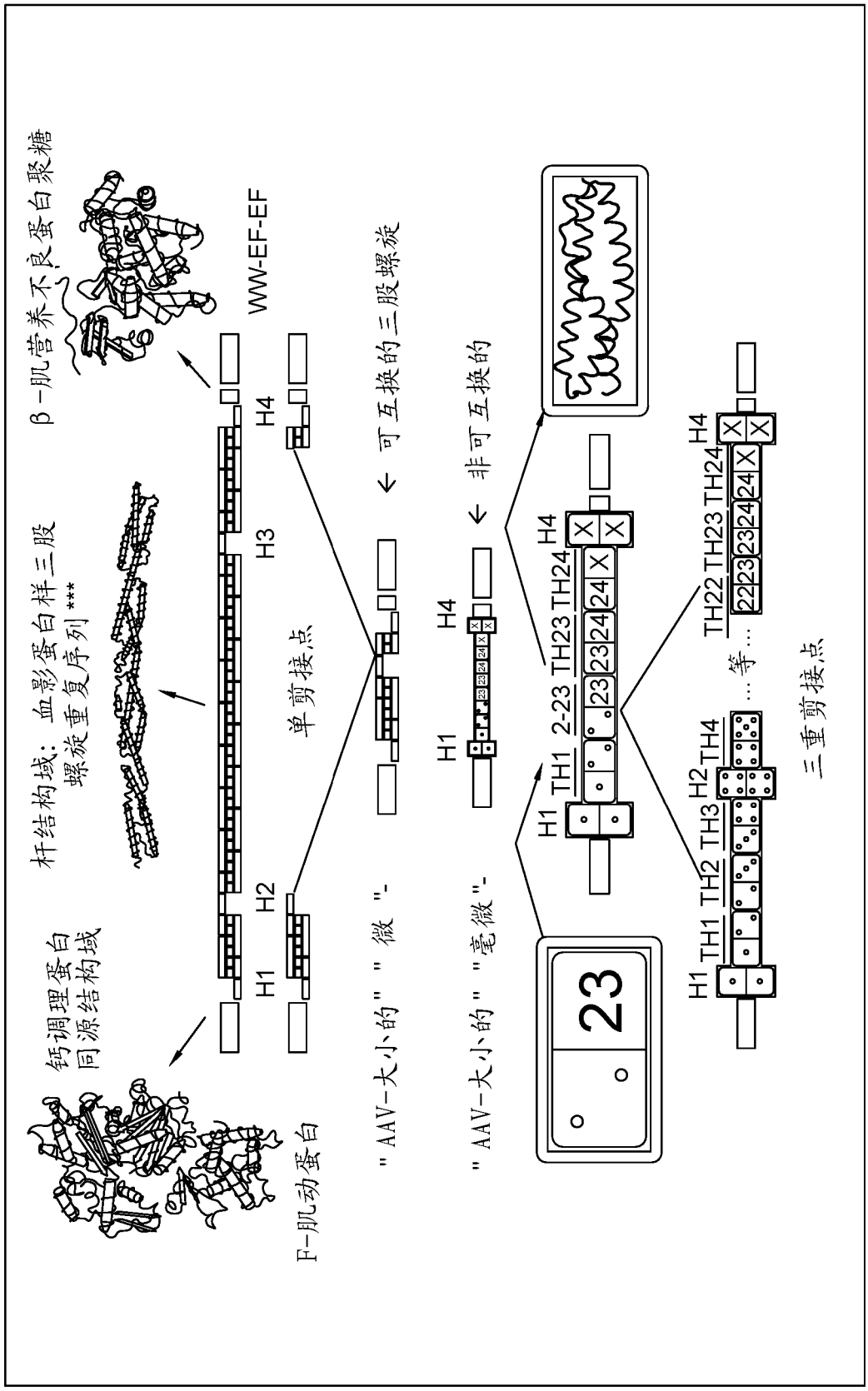
钙调理蛋白
同源结构域
杆结构域：血影蛋白样三股
螺旋重复序列***
β-肌营养不良蛋白聚糖
WW-EF-EF
H1
H2
H3
H4
F-肌动蛋白
单剪接点
"AAV-大小的" "微"-
↔ 可互换的三股螺旋
"AAV-大小的" "毫微"-
↔ 非可互换的
TH1
2-23
TH23
TH24
TH2
TH3
TH4
TH22
…等…
三重剪接点

图 4

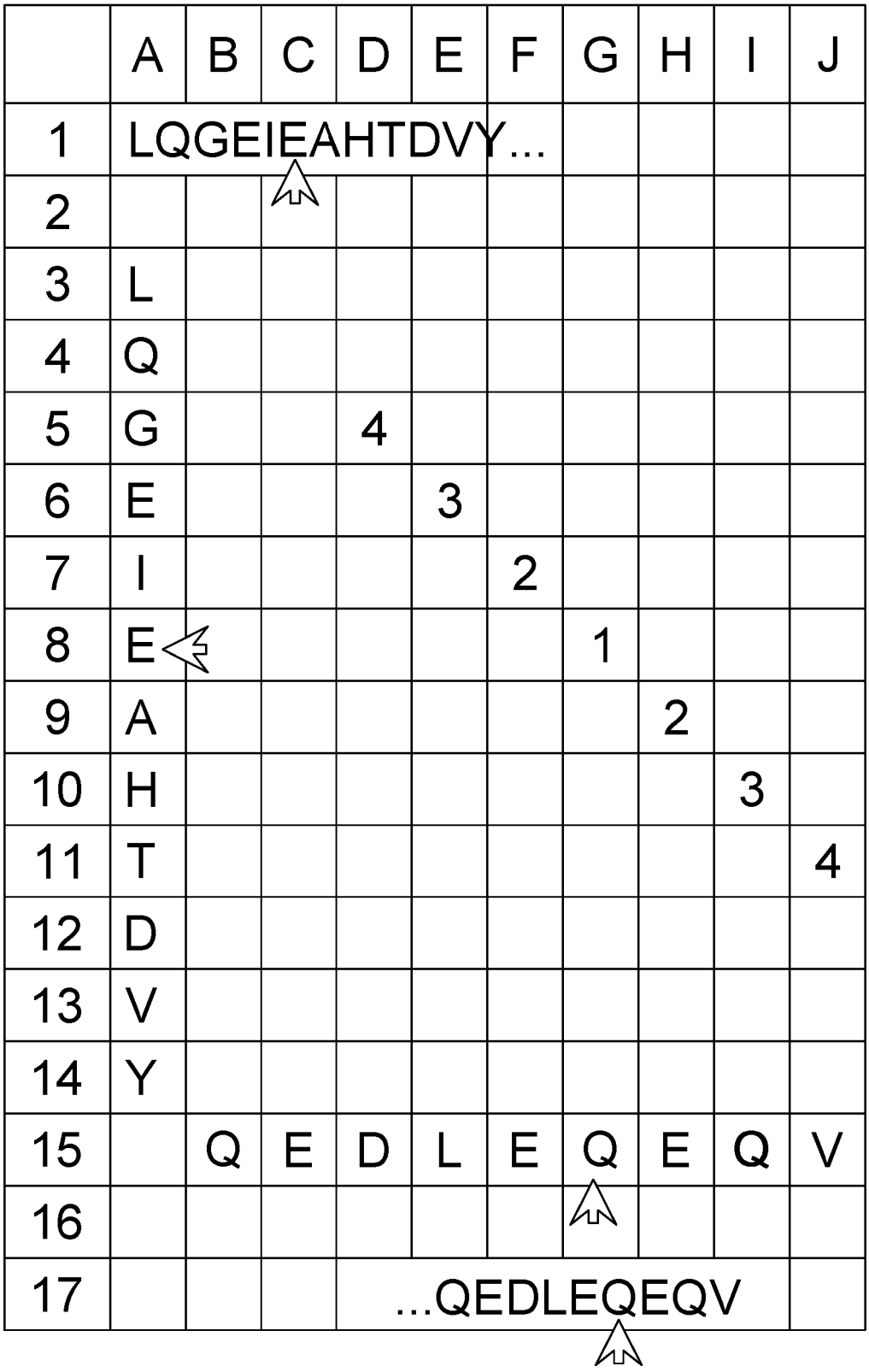
A B C D E F G H I J
1 LQGEIEAHTDVY...
2
3 L
4 Q
5 G 4
6 E 3
7 I 2
8 E 1
9 A 2
10 H 3
11 T 4
12 D
13 V
14 Y
15 Q E D L E Q E Q V
16
17 ...QEDLEQEQV

图 5

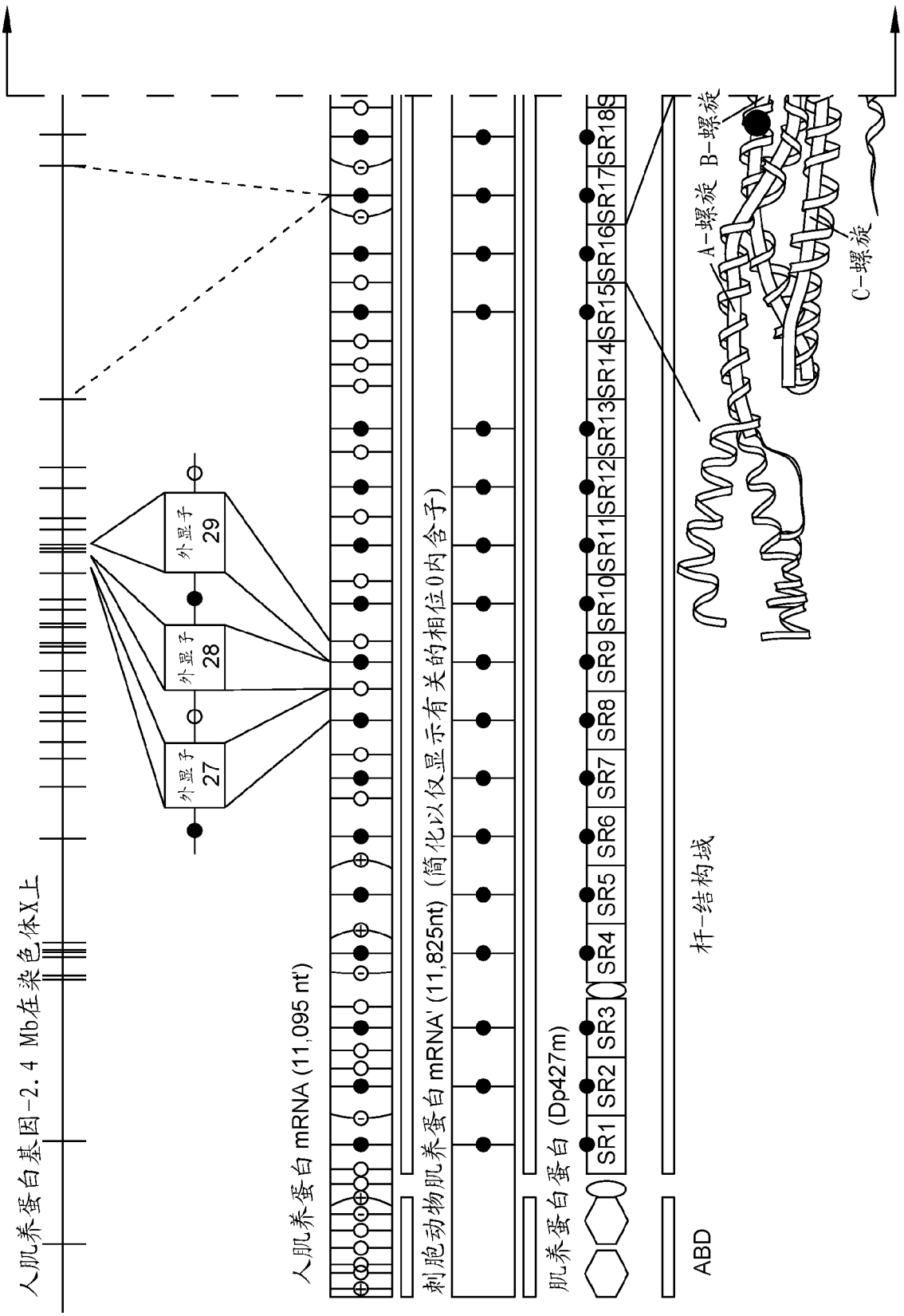
人肌养蛋白基因-2.4 Mb在染色体X上
外显子 27
外显子 28
外显子 29
人肌养蛋白 mRNA (11,095 nt)
刺胞动物肌养蛋白 mRNA (11,825nt) (简化以仅显示有关的相位0内含子)
肌养蛋白蛋白 (Dp427m)
SR1
SR2
SR3
SR4
SR5
SR6
SR7
SR8
SR9
SR10
SR11
SR12
SR13
SR14
SR15
SR16
SR17
SR18
ABD
杆-结构域
A-螺旋
B-螺旋
C-螺旋

图 6A(I)

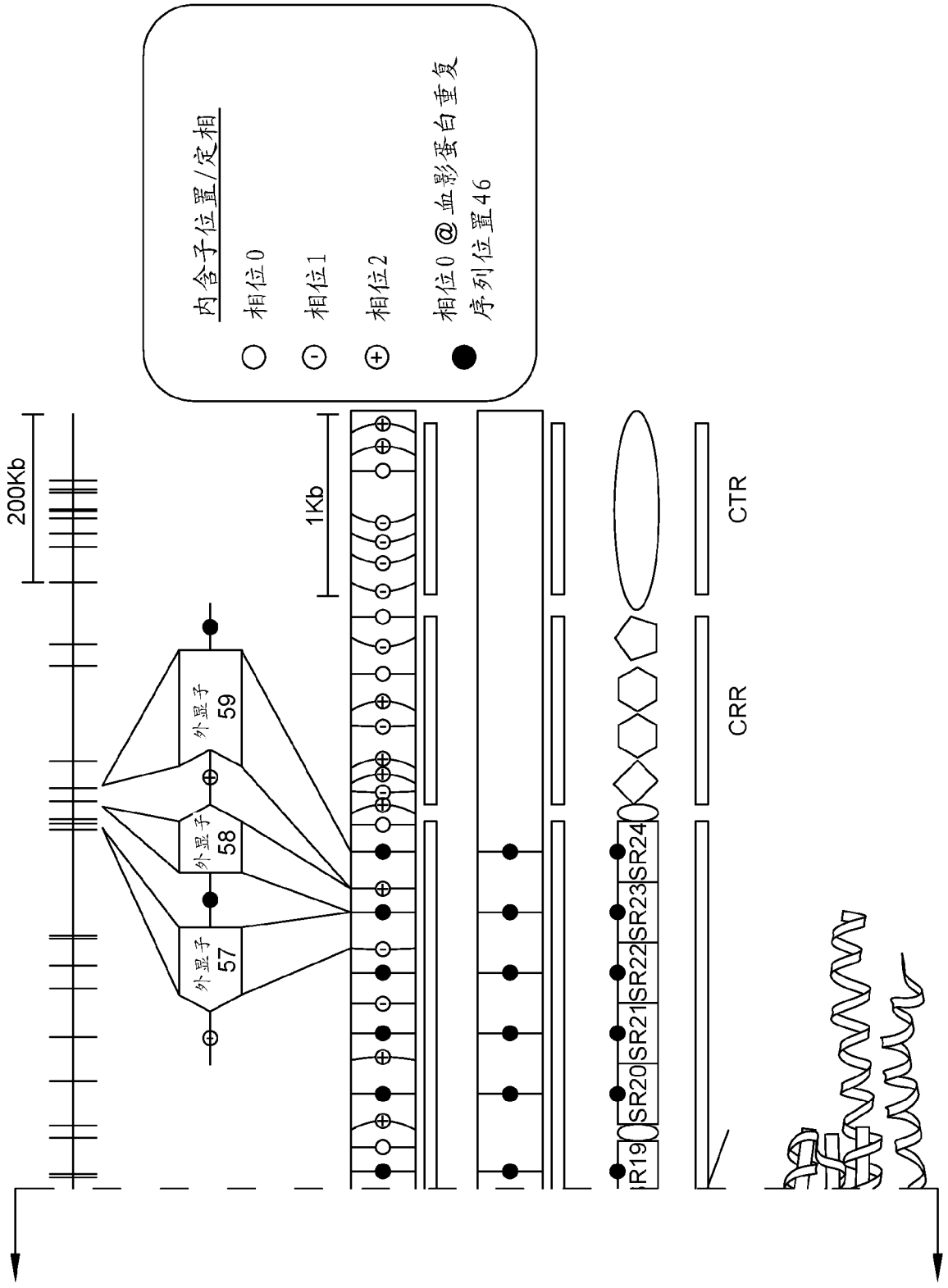
内含子位置/定相
相位0
相位1
相位2
相位0@血影蛋白重复
序列位置46
200Kb
1Kb
外显子
59
外显子
58
外显子
57
CTR
CRR
IR19
SR20
SR21
SR22
SR23
SR24

图 6A(II)

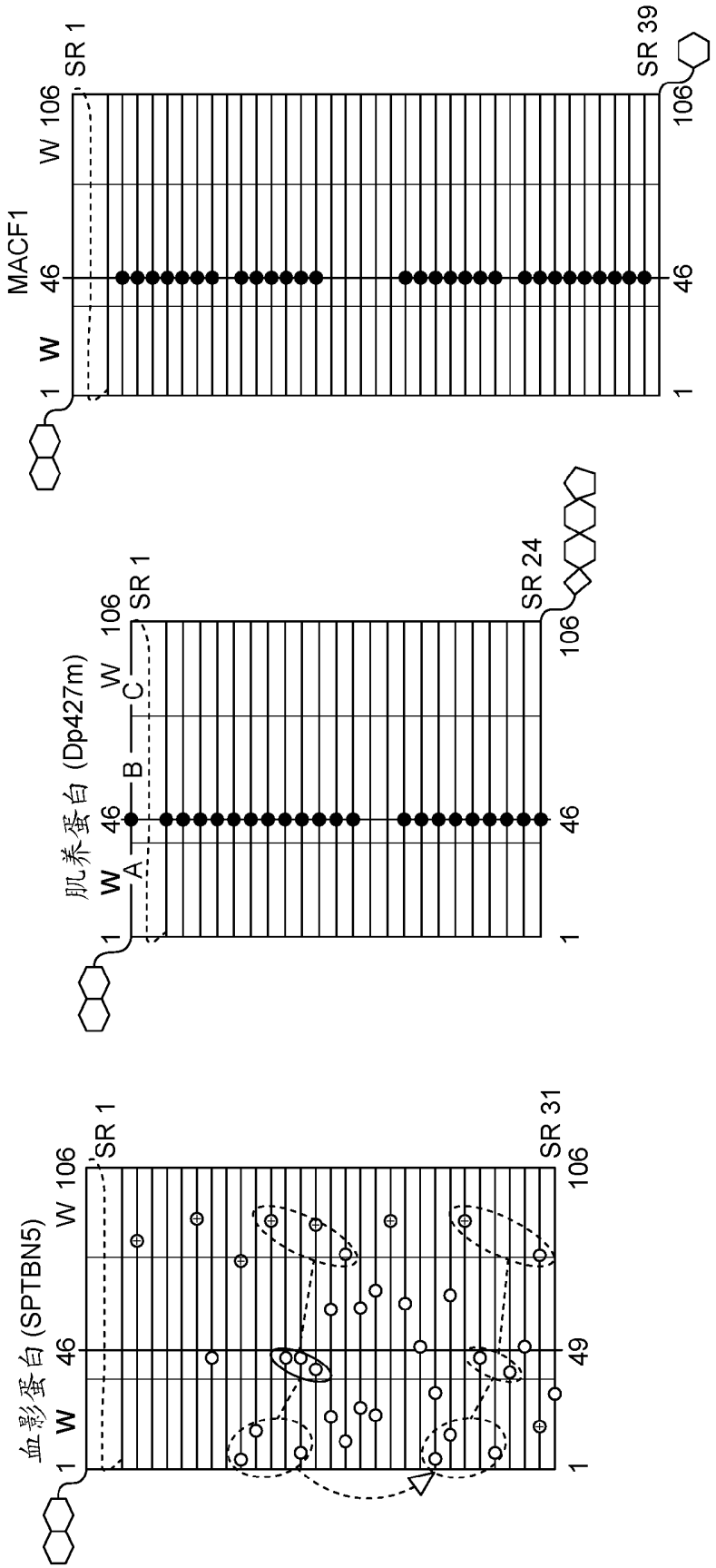
MACF1
W 106
SR 1
SR 39
106
46
1
W
肌养蛋白 (Dp427m)
A
B
C
SR 24
血影蛋白 (SPTBN5)
SR 31
49

图 6B

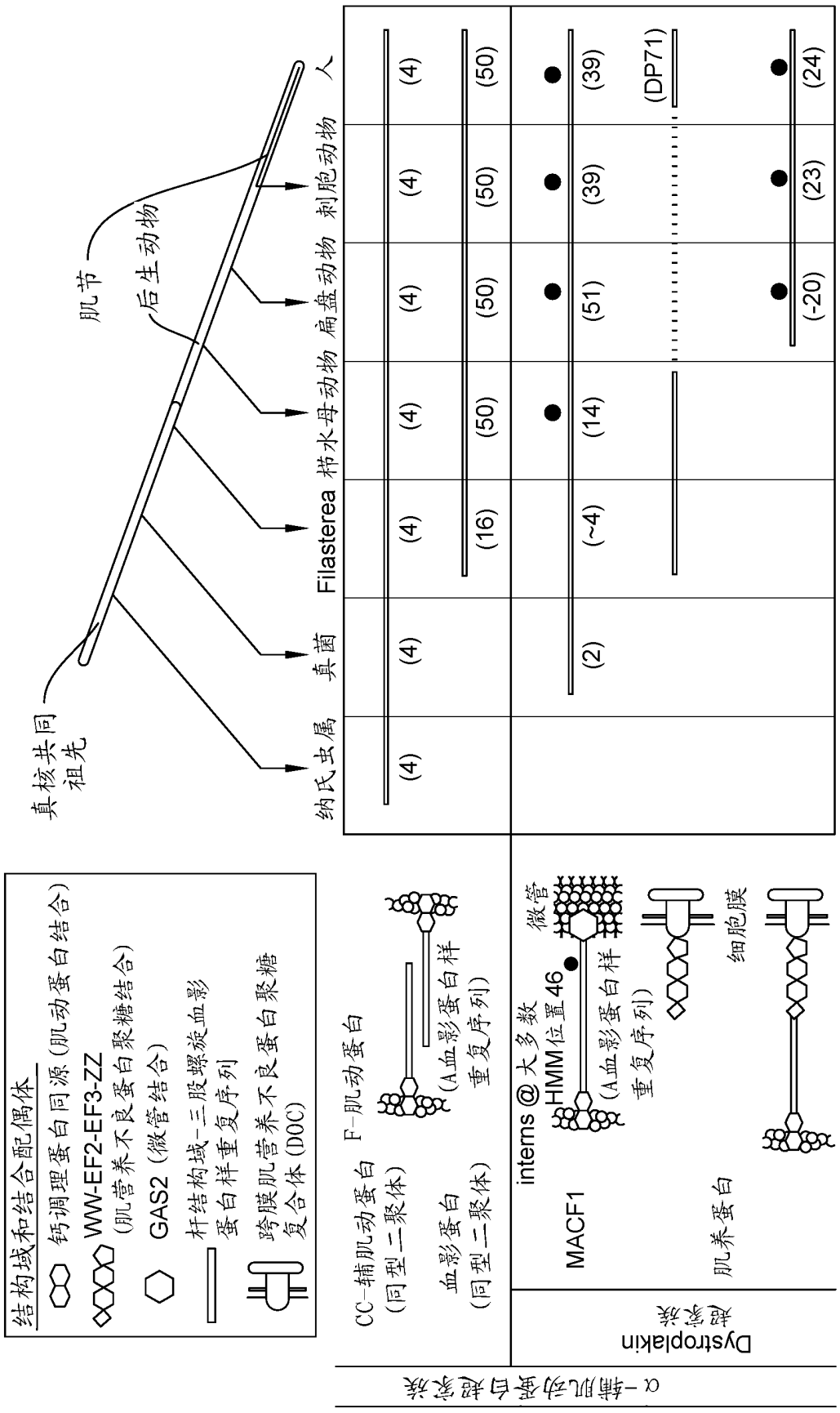
结构域和结合配偶体
钙调理蛋白同源(肌动蛋白结合)
WW-EF2-EF3-ZZ
(肌营养不良蛋白聚糖结合)
GAS2(微管结合)
杆结构域-三股螺旋血影蛋白样重复序列
跨膜肌营养不良蛋白聚糖复合体(DOC)
真核共同祖先
肌节
后生动物
纳氏虫属
真菌
Filasterea
栉水母动物
扁盘动物
刺胞动物
人
(4)
(4)
(4)
(4)
(4)
(4)
(16)
(50)
(50)
(50)
(50)
(2)
(~4)
(14)
(51)
(39)
(39)
(DP71)
(-20)
(23)
(24)
α-辅肌动蛋白超家族
Dystrophakin超家族
CC-辅肌动蛋白(同型二聚体)
F-肌动蛋白
血影蛋白(同型二聚体)
(A血影蛋白样重复序列)
MACF1
interns @大多数HMM位置46
微管
(A血影蛋白样重复序列)
细胞膜
肌养蛋白

图 6C

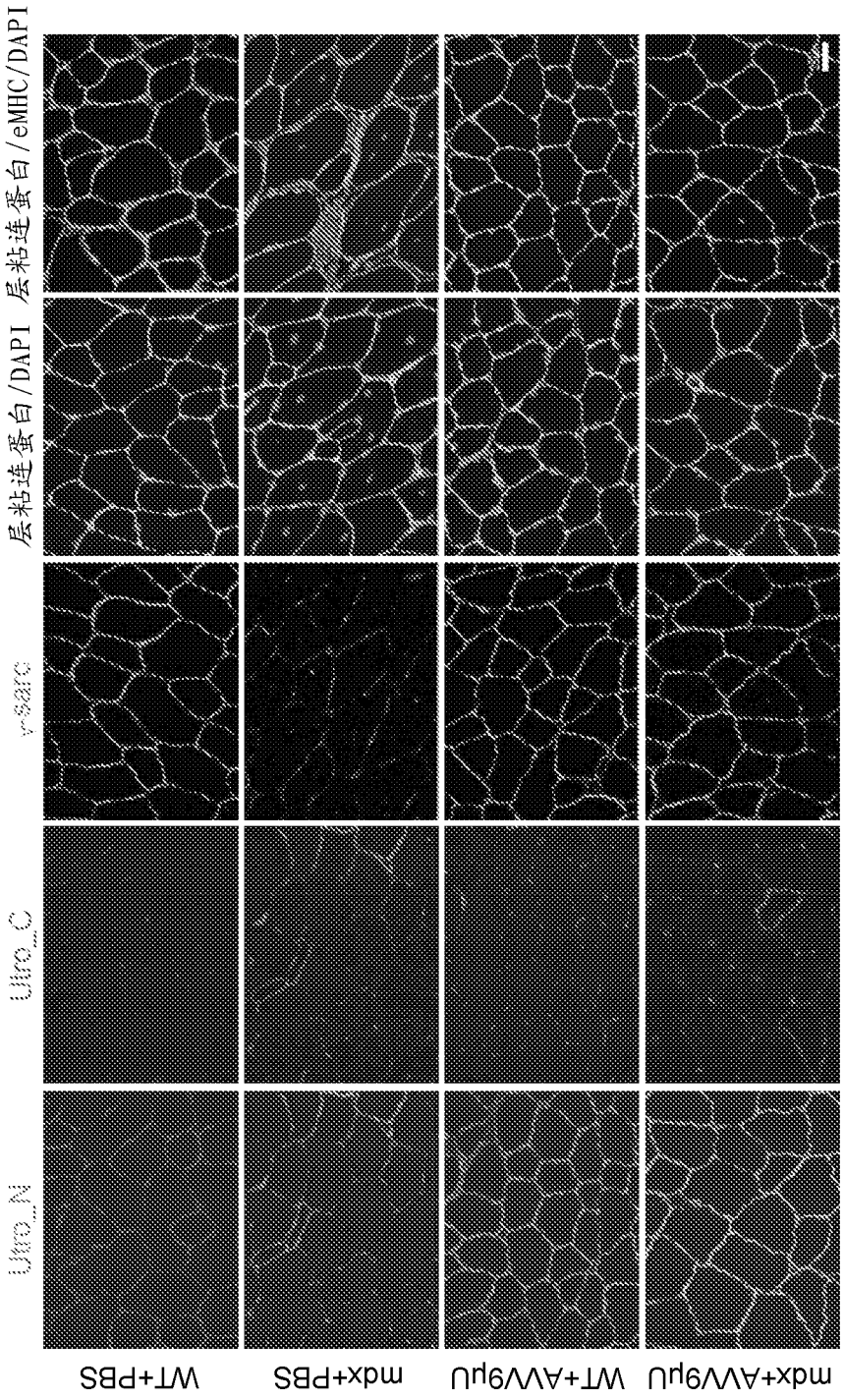
层粘连蛋白/eMHC/DAPI
层粘连蛋白/DAPI
γ-Sarc
Ubc_C
Ubc_N
WT+PBS
mdx+PBS
WT+AAV9μU
mdx+AAV9μU

图 7A

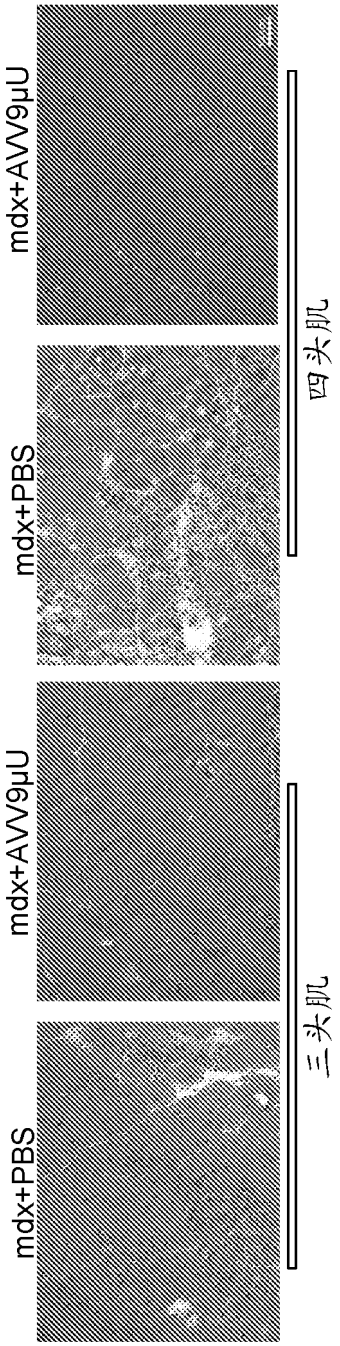

图 7B

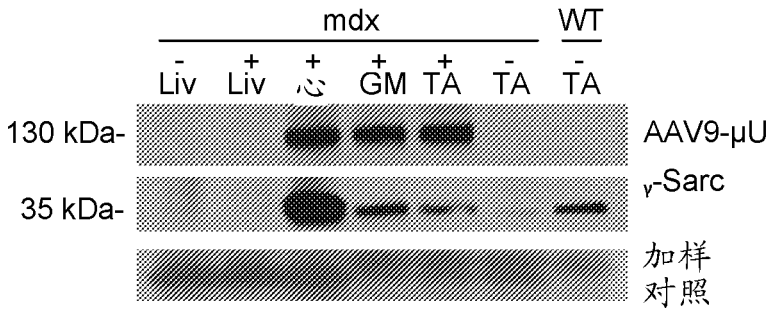
mdx
WT
- + + + + - -
Liv Liv 心 GM TA TA TA
130 kDa-
35 kDa-
AAV9-μU
γ-Sarc
加样
对照

图 7C

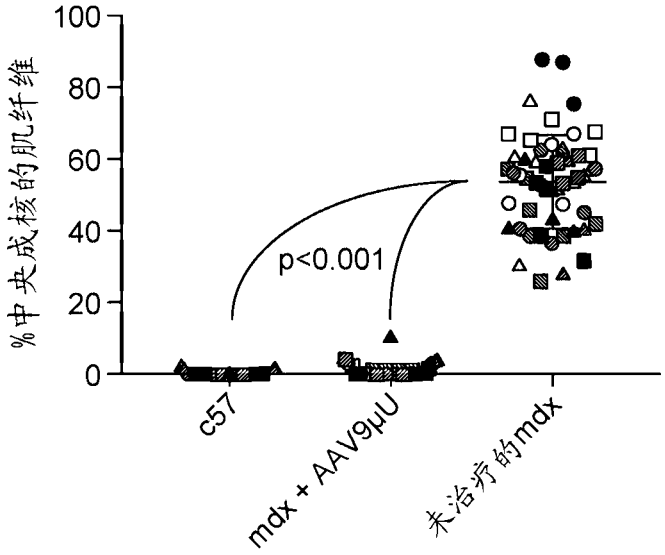
%中央成核的肌纤维
100
80
60
40
20
0
p<0.001
c57
mdx + AAV9μU
未治疗的mdx

图 7D

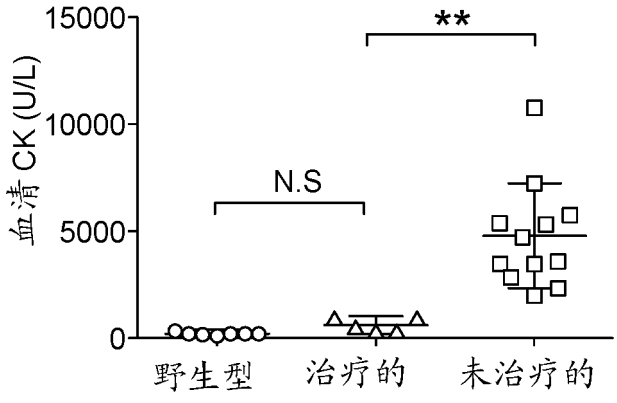
血清 CK (U/L)
15000
10000
5000
0
**
N.S
野生型
治疗的
未治疗的

图 7E

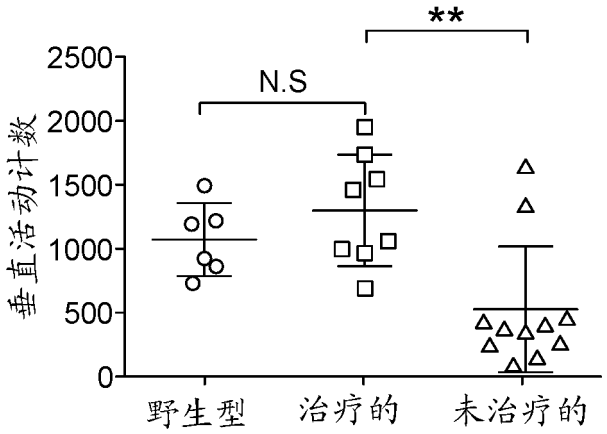
**
N.S
2500
2000
1500
1000
500
0
垂直活动计数
野生型
治疗的
未治疗的

图 7F

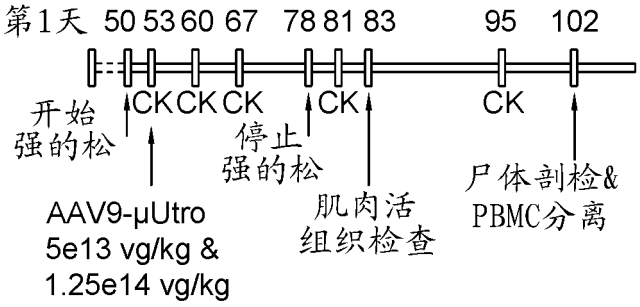
第1天 50 53 60 67 78 81 83 95 102
CK CK CK CK CK
开始强的松
停止强的松
AAV9-μUtro
5e13 vg/kg &
1.25e14 vg/kg
肌肉活组织检查
尸体剖检&
PBMC分离

图 8A

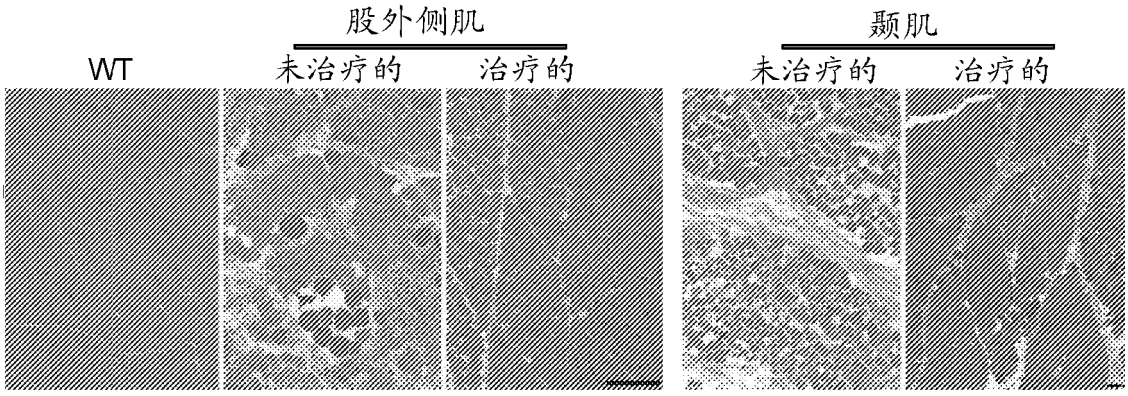
股外侧肌
颞肌
WT
未治疗的
治疗的
未治疗的
治疗的

图 8B

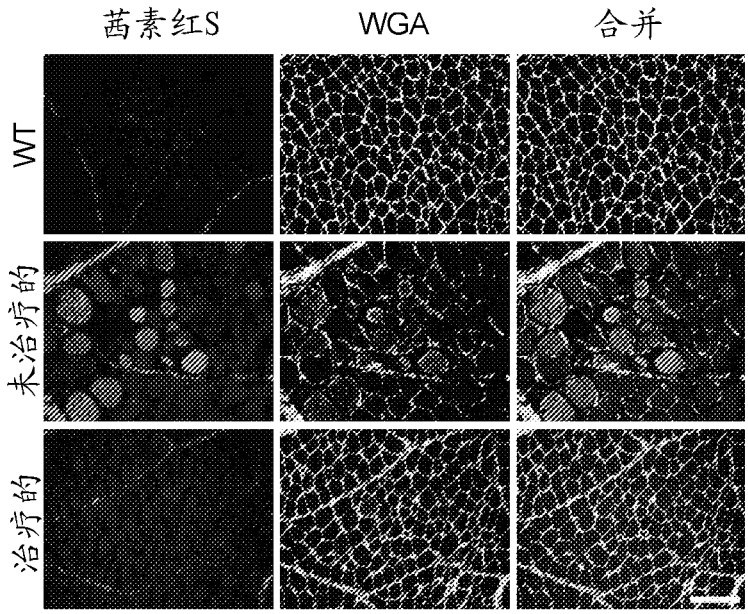

图 8C

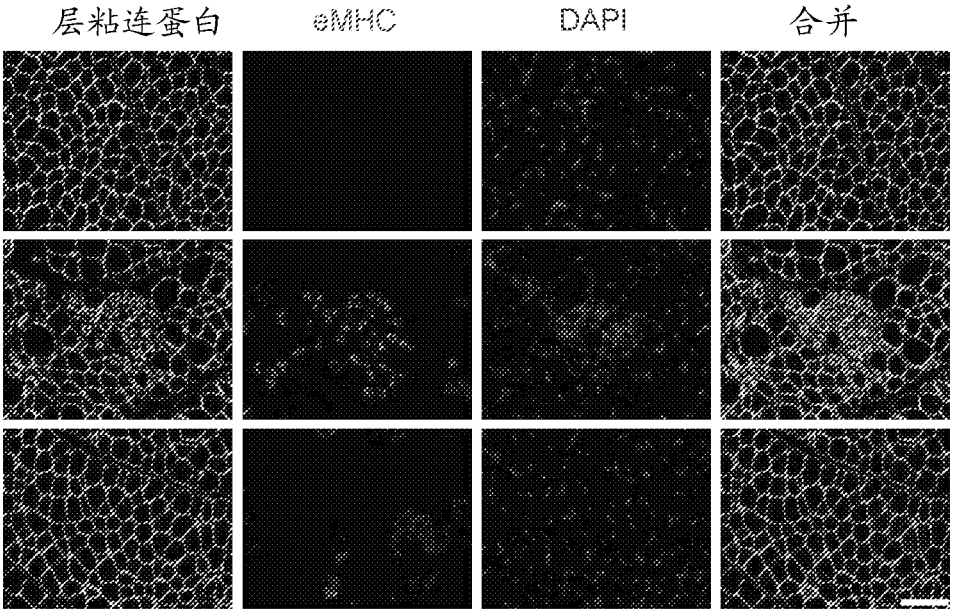

图 8D

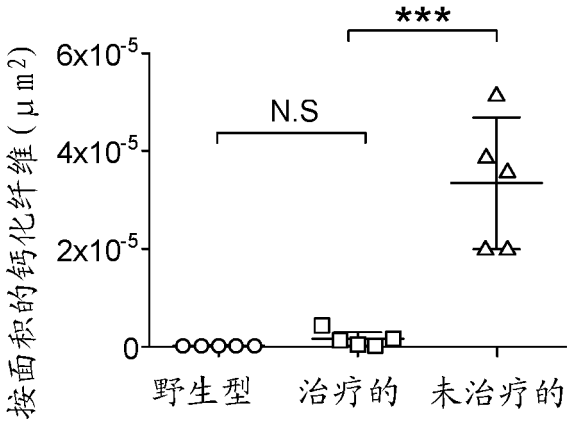
***
N.S
6x10-5
4x10-5
2x10-5
0
接面积的钙化纤维（μm2）
野生型
治疗的
未治疗的

图 8E

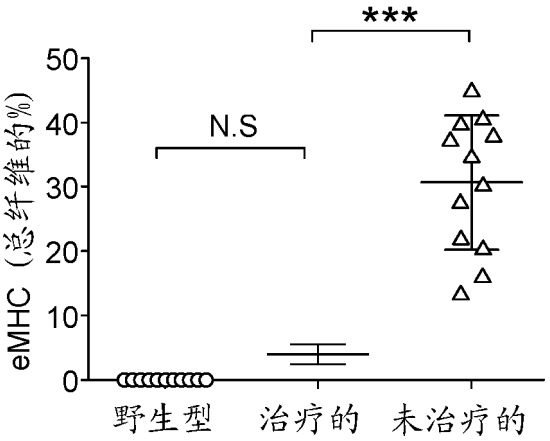
***
N.S
50
40
30
20
10
0
eMHC（总纤维的%）
野生型
治疗的
未治疗的

图 8F

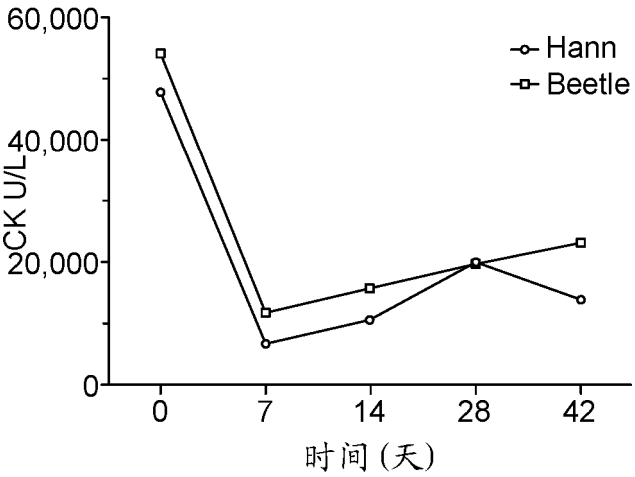
60,000
40,000
20,000
0
CK U/L
Hann
Beetle
0
7
14
28
42
时间(天)

图 8G

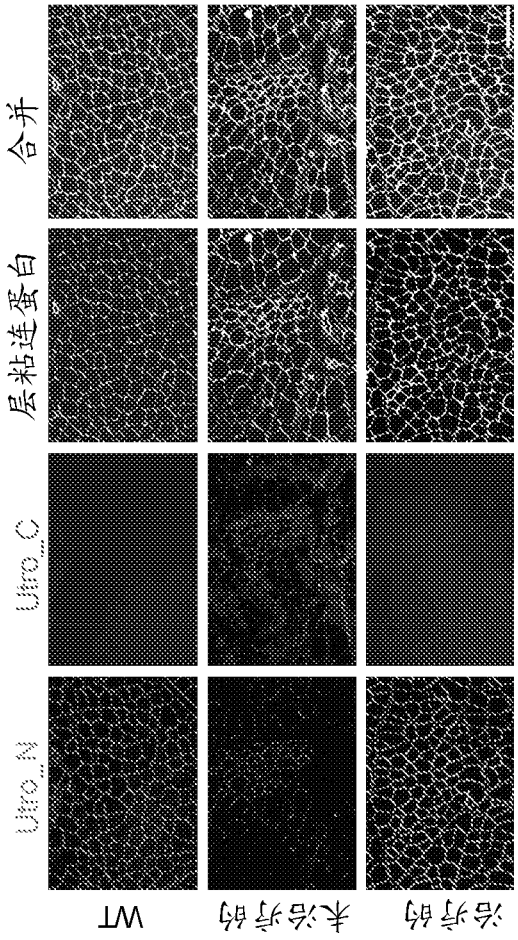

图 9A

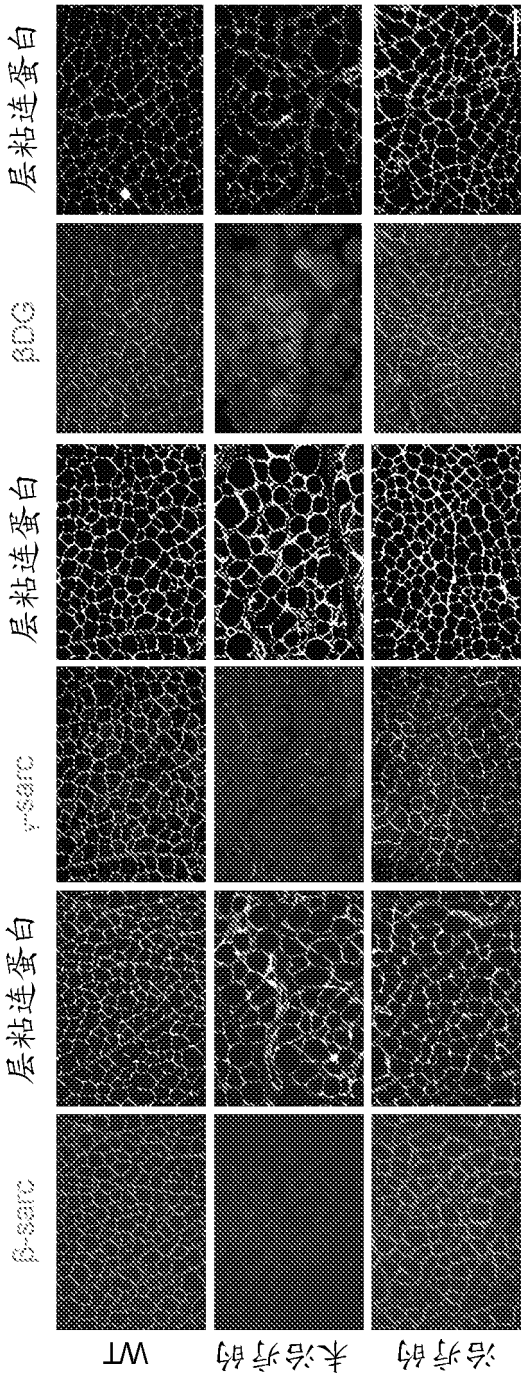

图 9B

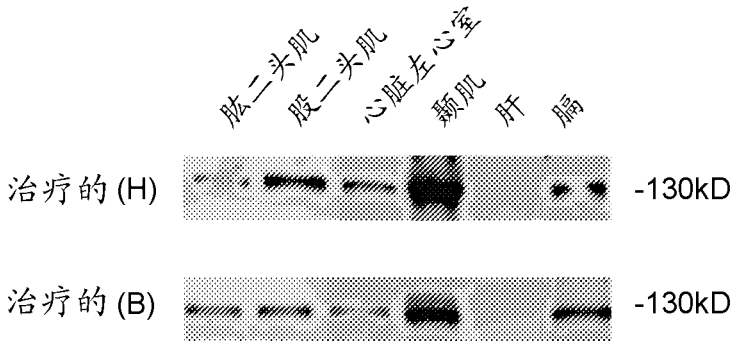
胫二头肌
股二头肌
心脏左心室
颞肌
肝
膈
治疗的 (H)
-130kD
治疗的 (B)
-130kD

图 9C

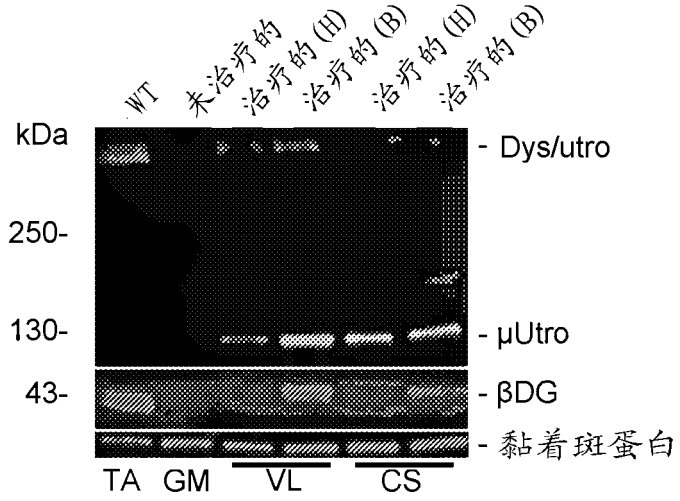
WT
未治疗的
治疗的 (H)
治疗的 (B)
治疗的 (H)
治疗的 (B)
kDa
250-
130-
43-
- Dys/utro
- μUtro
- βDG
- 黏着斑蛋白
TA GM VL CS

图 9D

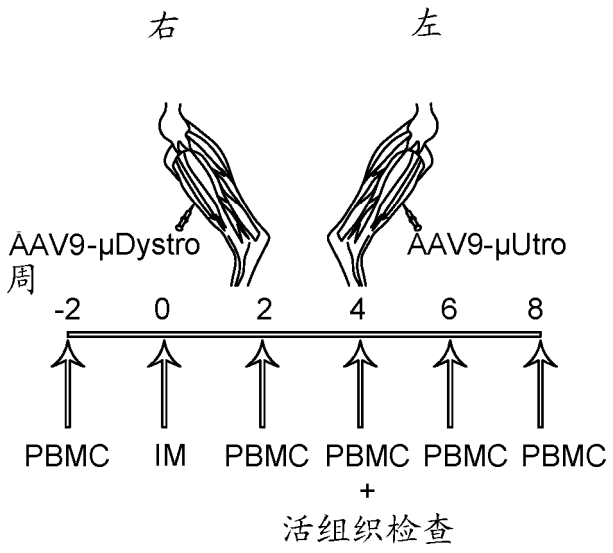
右
左
AAV9-μDystro
周
AAV9-μUtro
-2 0 2 4 6 8
PBMC IM PBMC PBMC PBMC PBMC
+
活组织检查

图 10A

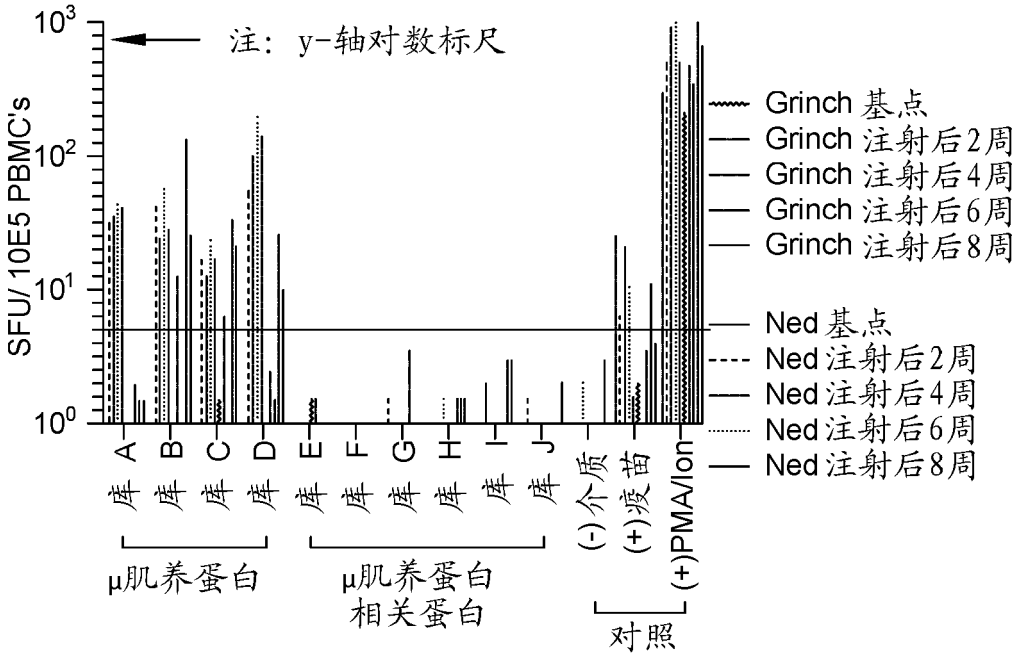
注：y-轴对数标尺
SFU/10E5 PBMC's
Grinch 基点
Grinch 注射后2周
Grinch 注射后4周
Grinch 注射后6周
Grinch 注射后8周
Ned 基点
Ned 注射后2周
Ned 注射后4周
Ned 注射后6周
Ned 注射后8周
A 库
B 库
C 库
D 库
E 库
F 库
G 库
H 库
I 库
J 库
(-)介质
(+)疫苗
(+)PMA/Ion
μ肌养蛋白
μ肌养蛋白相关蛋白
对照

图 10B

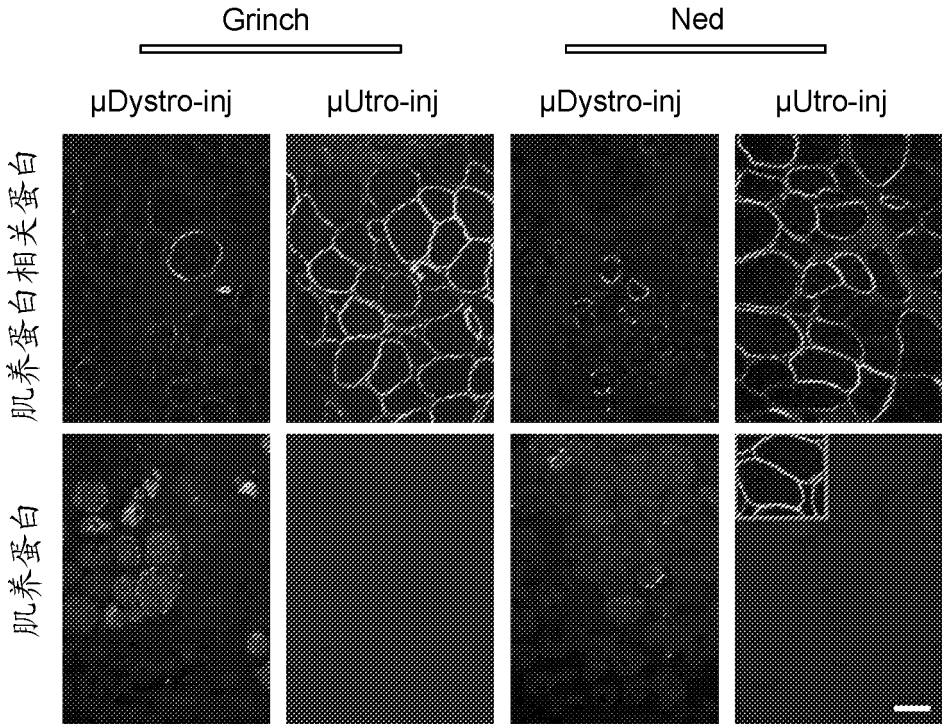
Grinch
Ned
μDystro-inj
μUtro-inj
μDystro-inj
μUtro-inj
肌养蛋白相关蛋白
肌养蛋白

图 10C

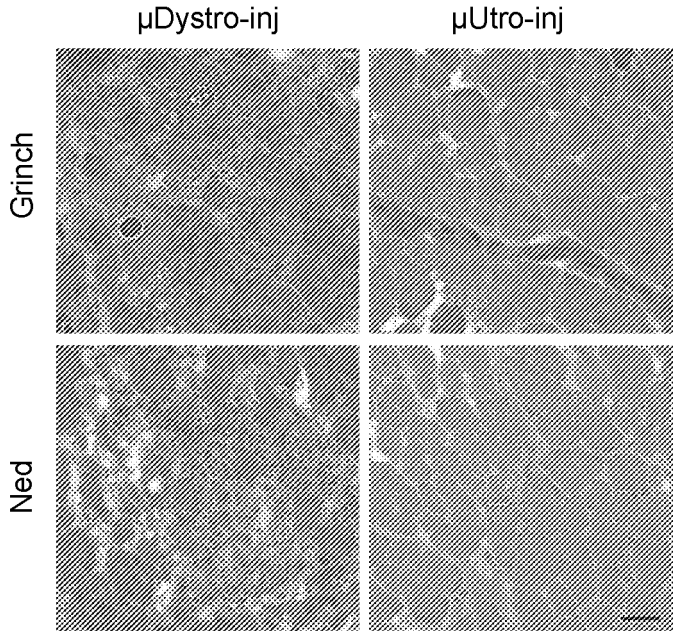

图 10D

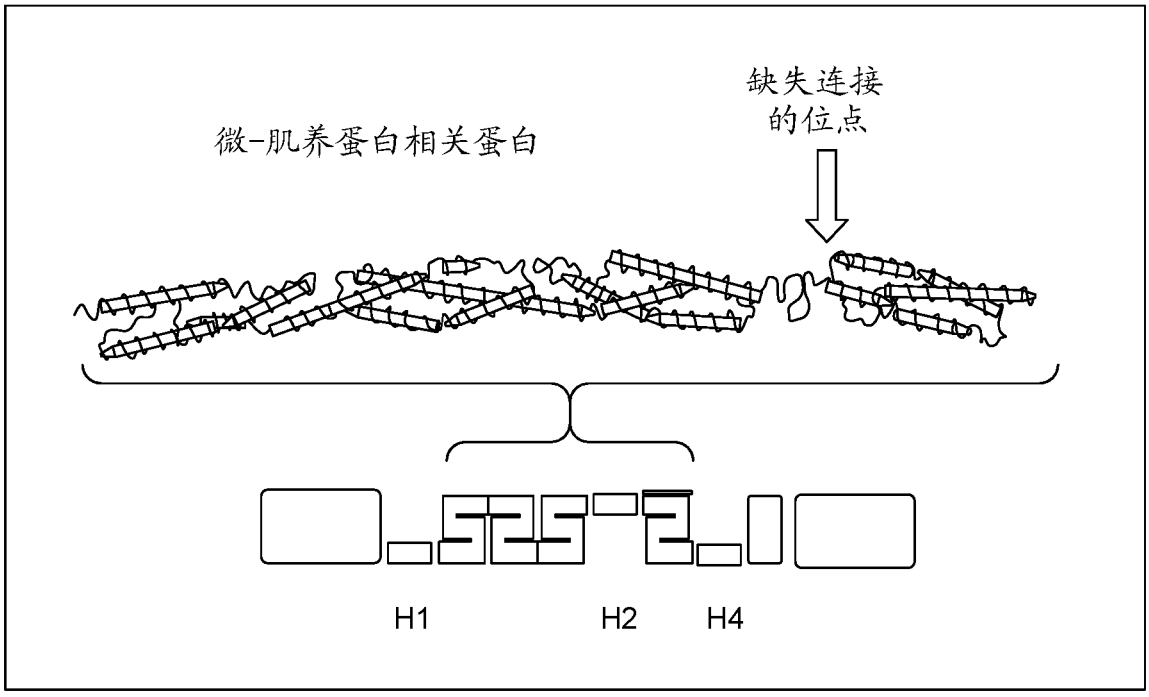

图 11

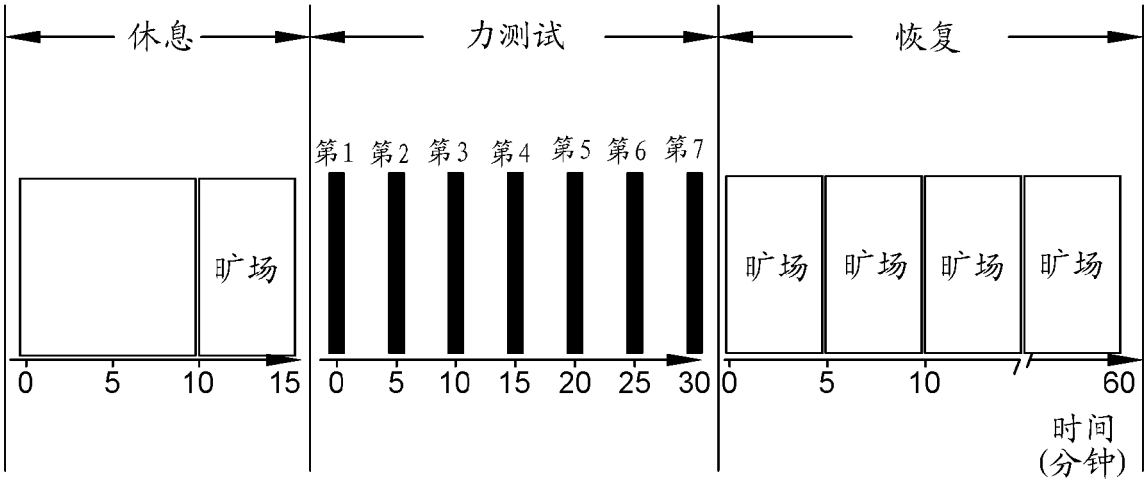

图 12A

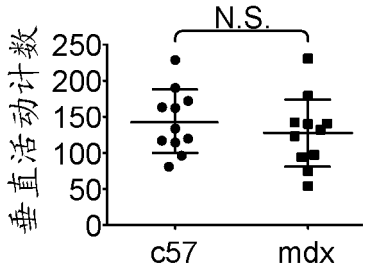

图 12B

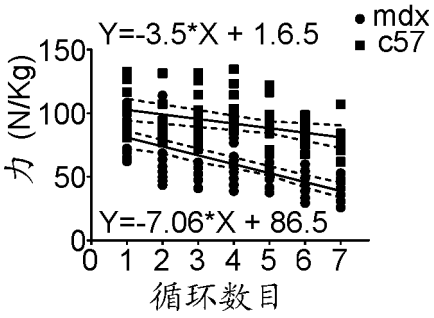

图 12C

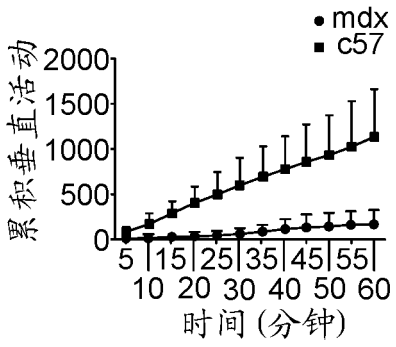

图 12D

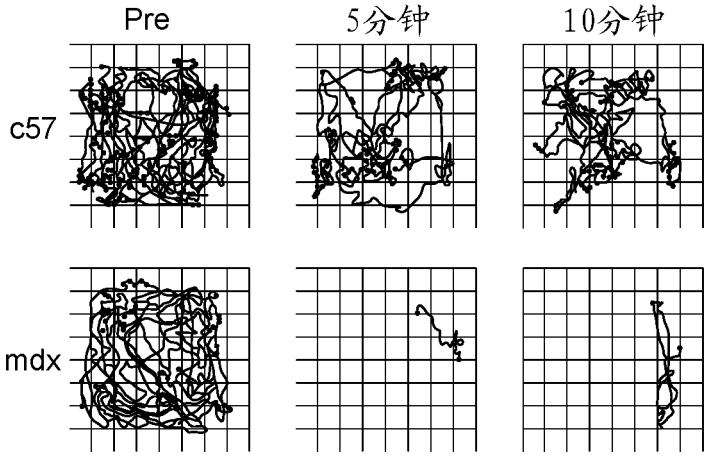

图 13

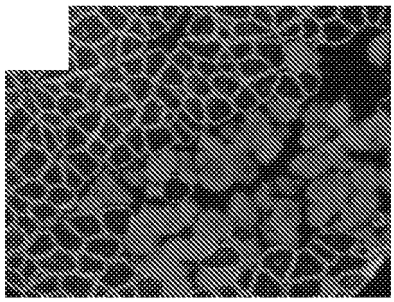

图 14A

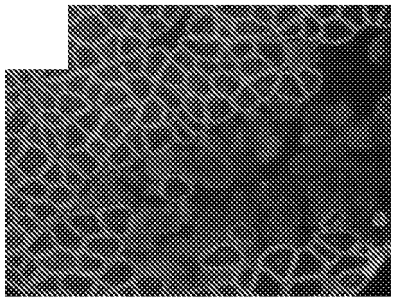

图 14B

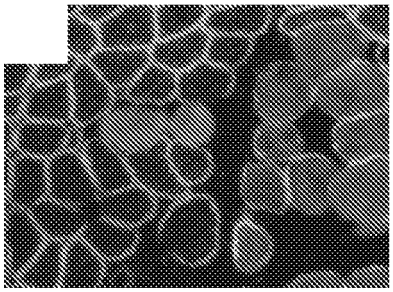

图 14C

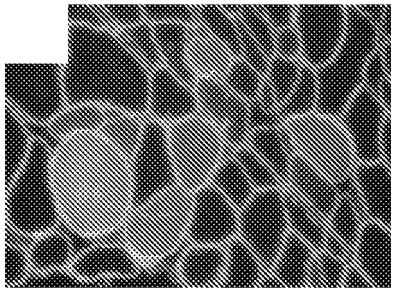

图 14D

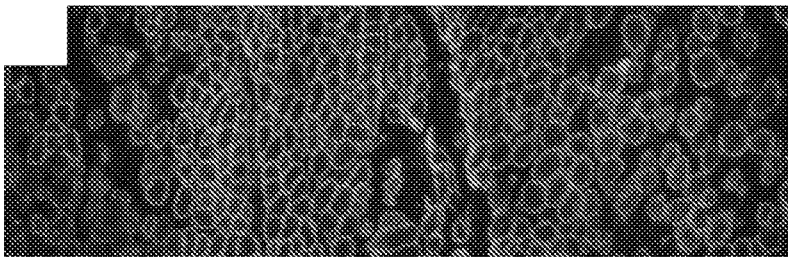

图 14E

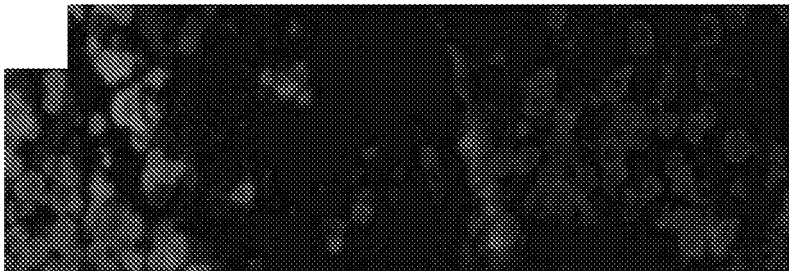

图 14F

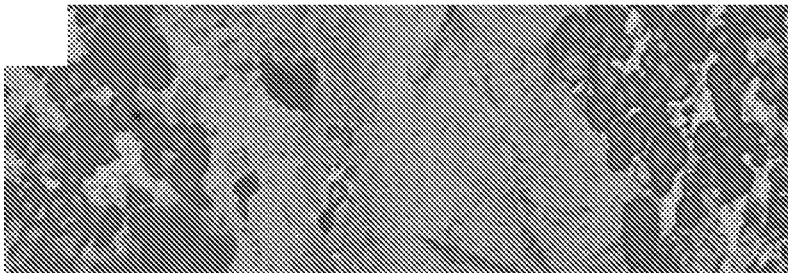

图 14G

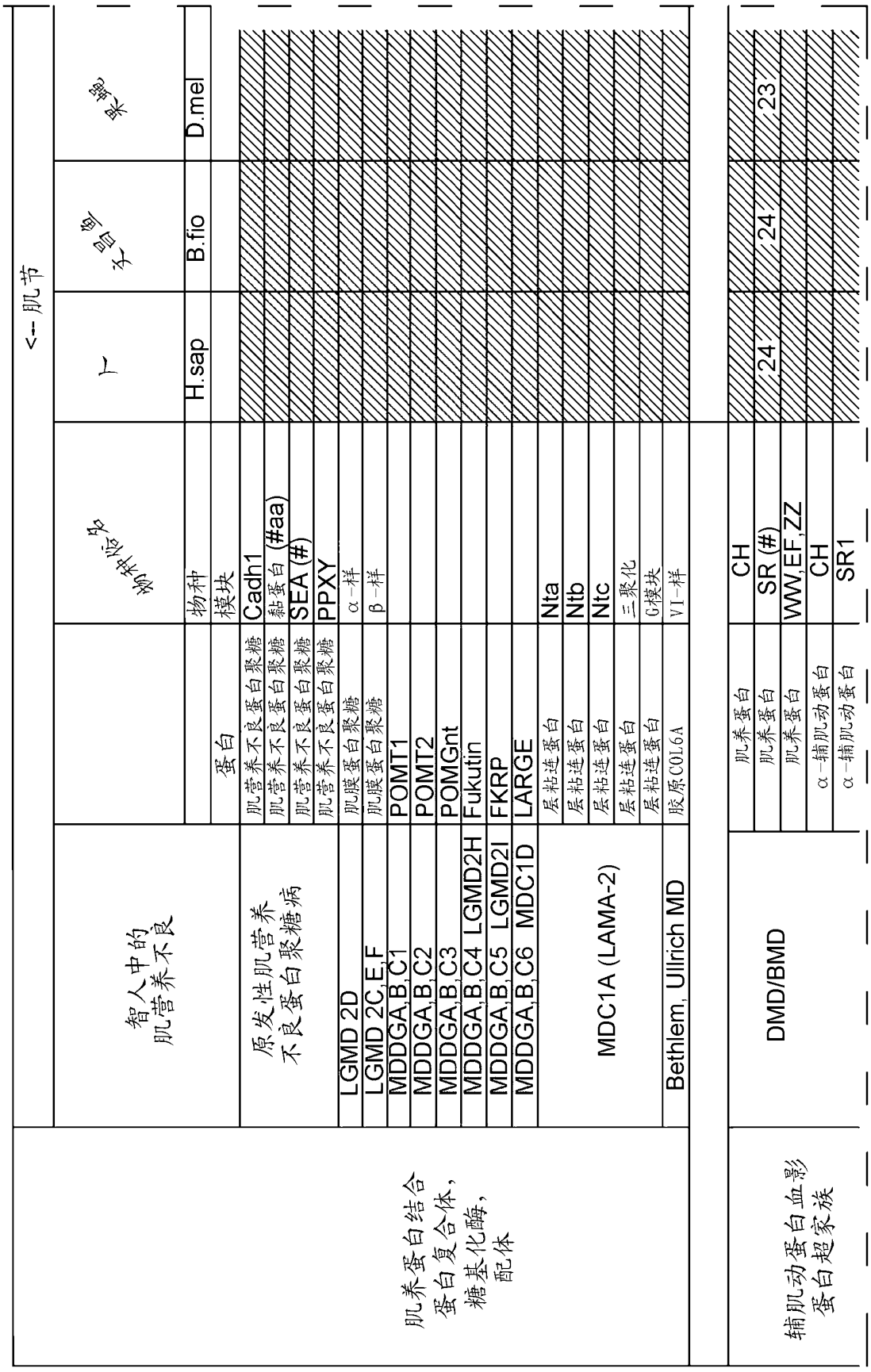
<-- 肌节
智人中的肌营养不良
物种
模块
H.sap
B.fio
D.mel
蛋白
原发性肌营养不良蛋白聚糖病
肌营养不良蛋白聚糖
Cadh1
黏蛋白 (#aa)
SEA (#)
PPXY
LGMD 2D
α-样
LGMD 2C,E,F
肌膜蛋白聚糖
β-样
MDDGA,B,C1
POMT1
MDDGA,B,C2
POMT2
MDDGA,B,C3
POMGnt
MDDGA,B,C4 LGMD2H
Fukutin
MDDGA,B,C5 LGMD2I
FKRP
MDDGA,B,C6 MDC1D
LARGE
MDC1A (LAMA-2)
层粘连蛋白
Nta
Ntb
Ntc
三聚化
G模块
Bethlem, Ullrich MD
胶原COL6A
VI-样
肌养蛋白聚糖病
肌养蛋白复合体，糖基化酶，配体
DMD/BMD
肌养蛋白
CH
肌养蛋白
SR (#)
WW,EF,ZZ
肌养蛋白
CH
α-辅肌动蛋白
α-辅肌动蛋白
SR1
辅肌动蛋白血影蛋白超家族
24
24
23

图 15(I)

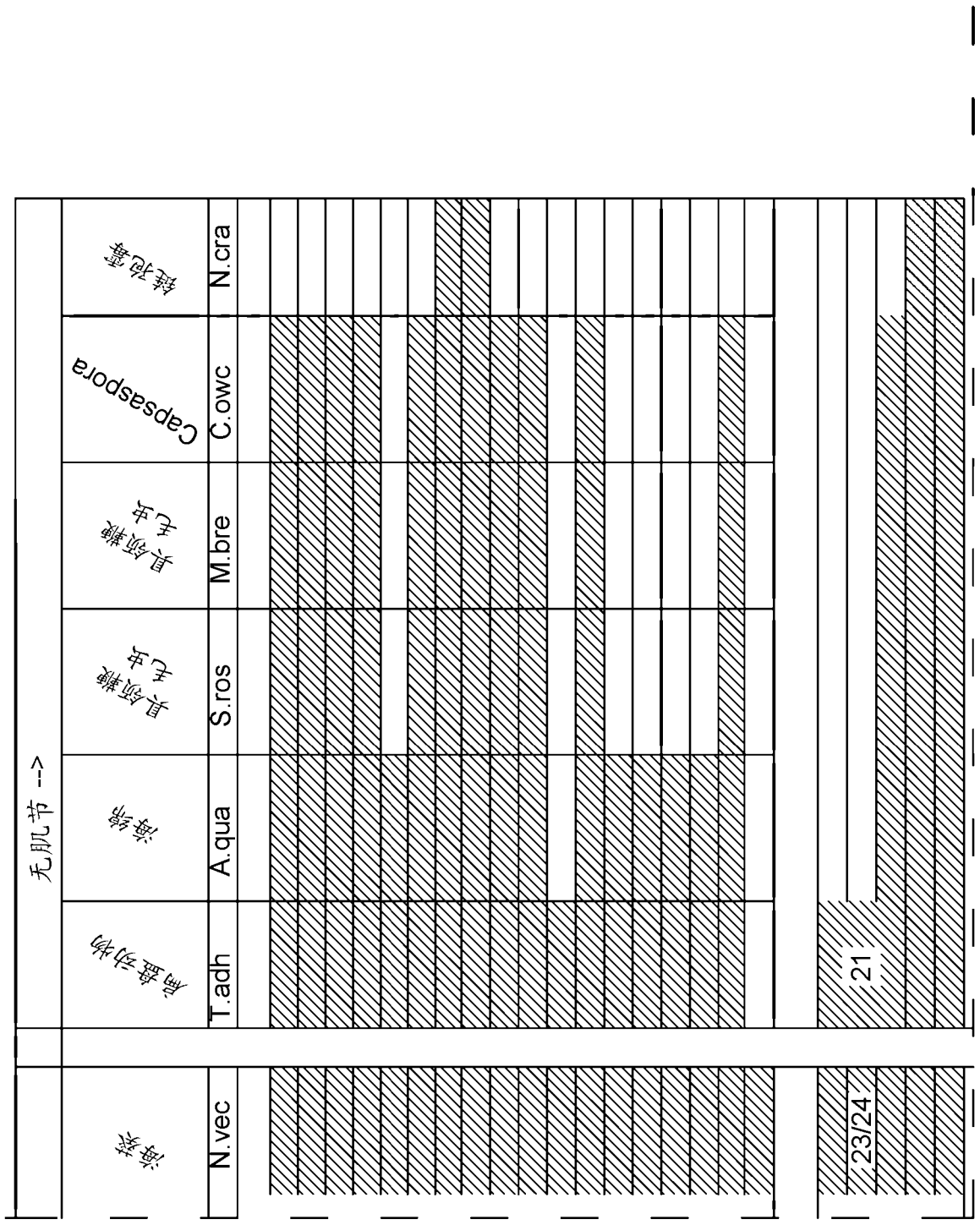
海绵
N.vec
角质动物
T.adh
海绵
A.qua
单领鞭毛虫
S.ros
单领鞭毛虫
M.bre
Capsaspora
C.owc
雄孢霉
N.cra
无肌节 -->
21
23/24

图 15(II)

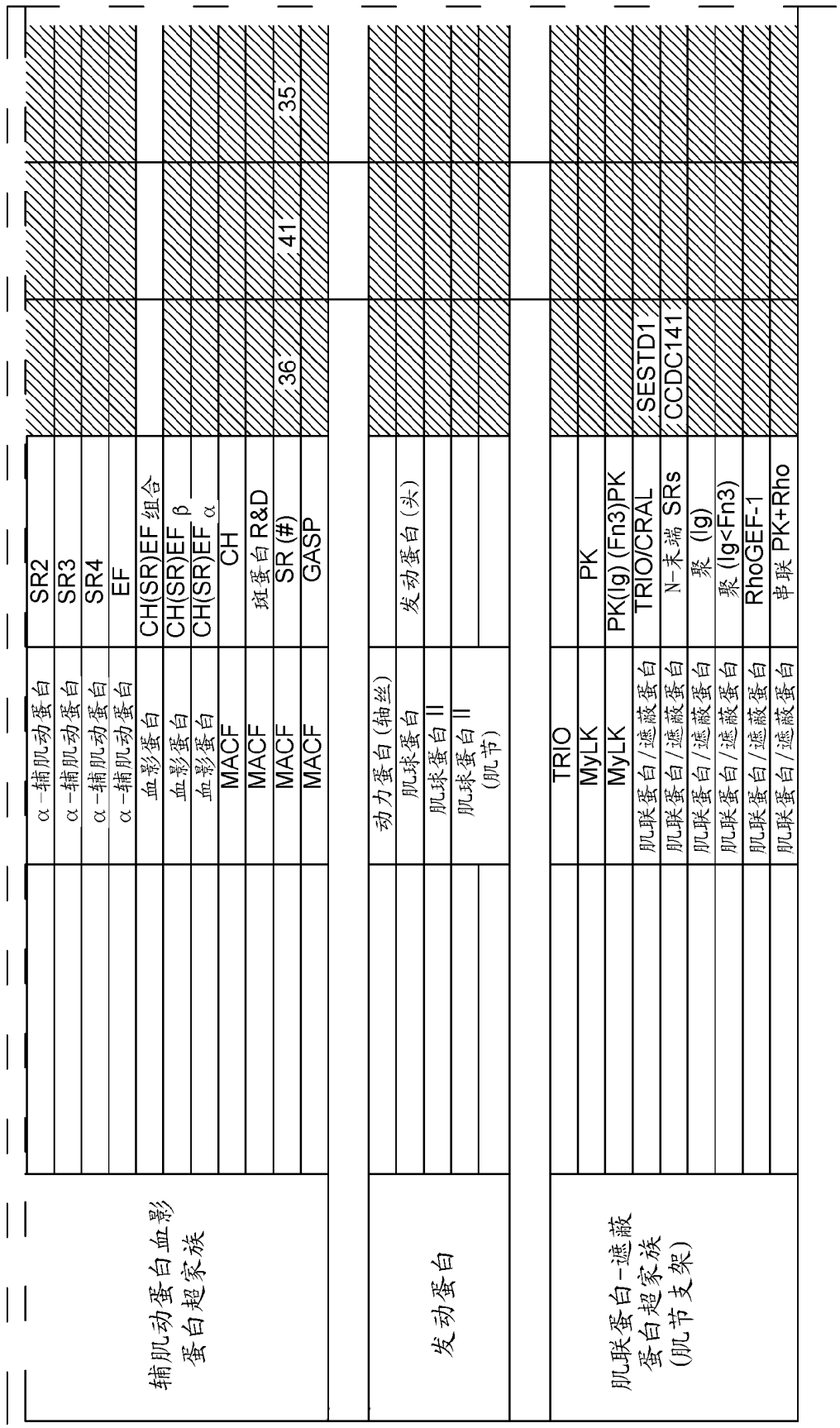
辅肌动蛋白血影蛋白超家族
α-辅肌动蛋白
α-辅肌动蛋白
α-辅肌动蛋白
α-辅肌动蛋白
血影蛋白
血影蛋白
MACF
MACF
MACF
MACF
SR2
SR3
SR4
EF
CH(SR)EF 组合
CH(SR)EF β
CH(SR)EF α
CH
斑蛋白 R&D
SR (#)
GASP
36
41
35
发动蛋白
动力蛋白(轴丝)
肌球蛋白
肌球蛋白 II
肌球蛋白 II (肌节)
发动蛋白(头)
肌联蛋白-遮蔽蛋白超家族(肌节支架)
TRIO
MyLK
MyLK
肌联蛋白/遮蔽蛋白
肌联蛋白/遮蔽蛋白
肌联蛋白/遮蔽蛋白
肌联蛋白/遮蔽蛋白
肌联蛋白/遮蔽蛋白
PK
PK(Ig)(Fn3)PK
TRIO/CRAL
N-末端 SRs
聚 (Ig)
聚 (Ig<Fn3)
RhoGEF-1
串联 PK+Rho
SESTD1
CCDC141

图 15(III)

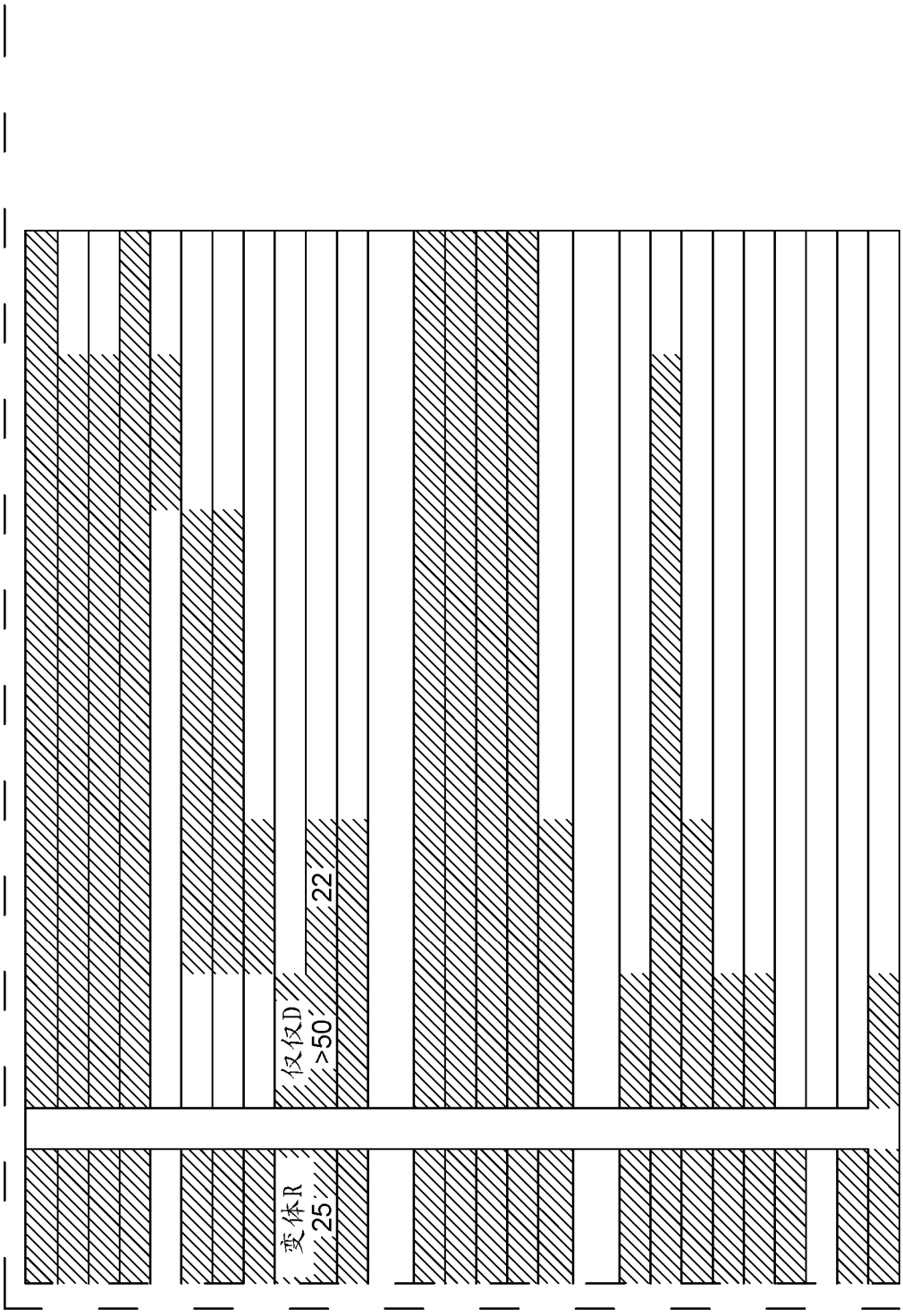

图 15(IV)

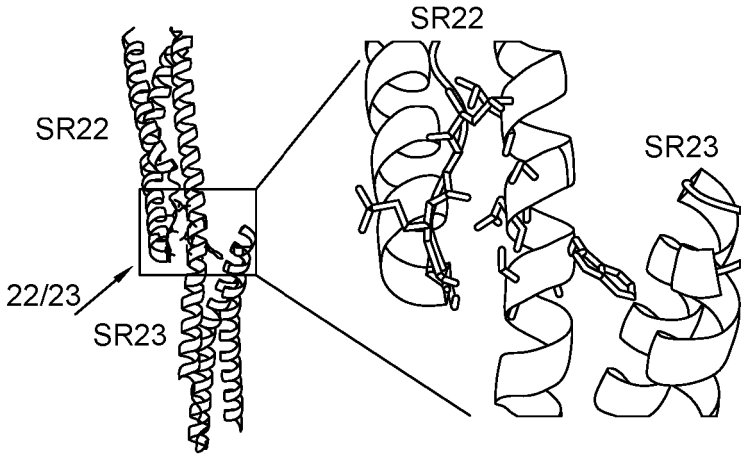

图 16A

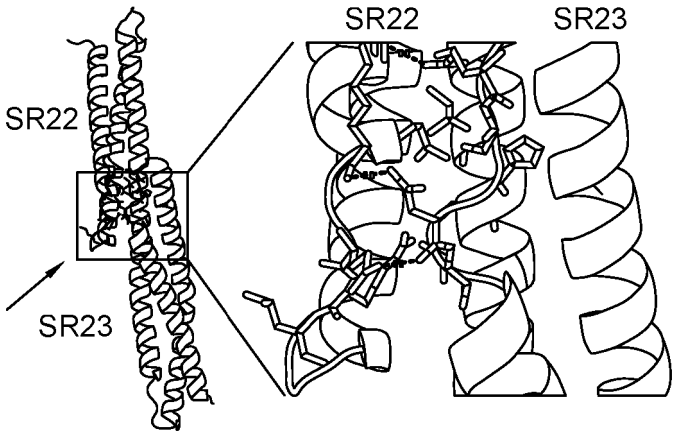

图 16B

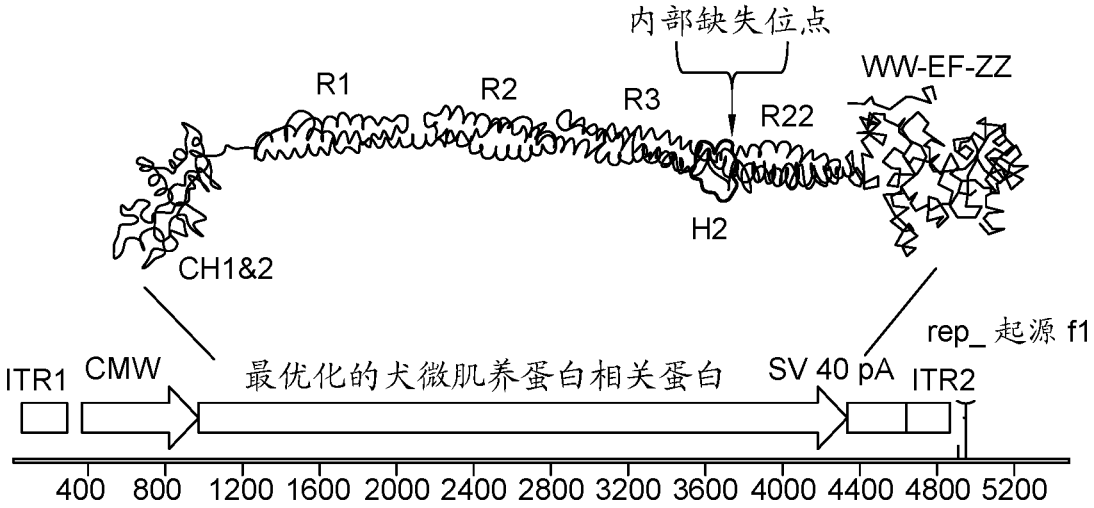

图 17

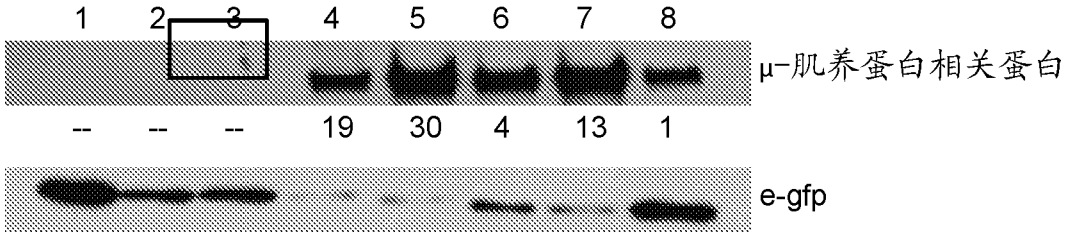
1
2
3
4
5
6
7
8
μ-肌养蛋白相关蛋白
--
--
--
19
30
4
13
1
e-gfp

图 18

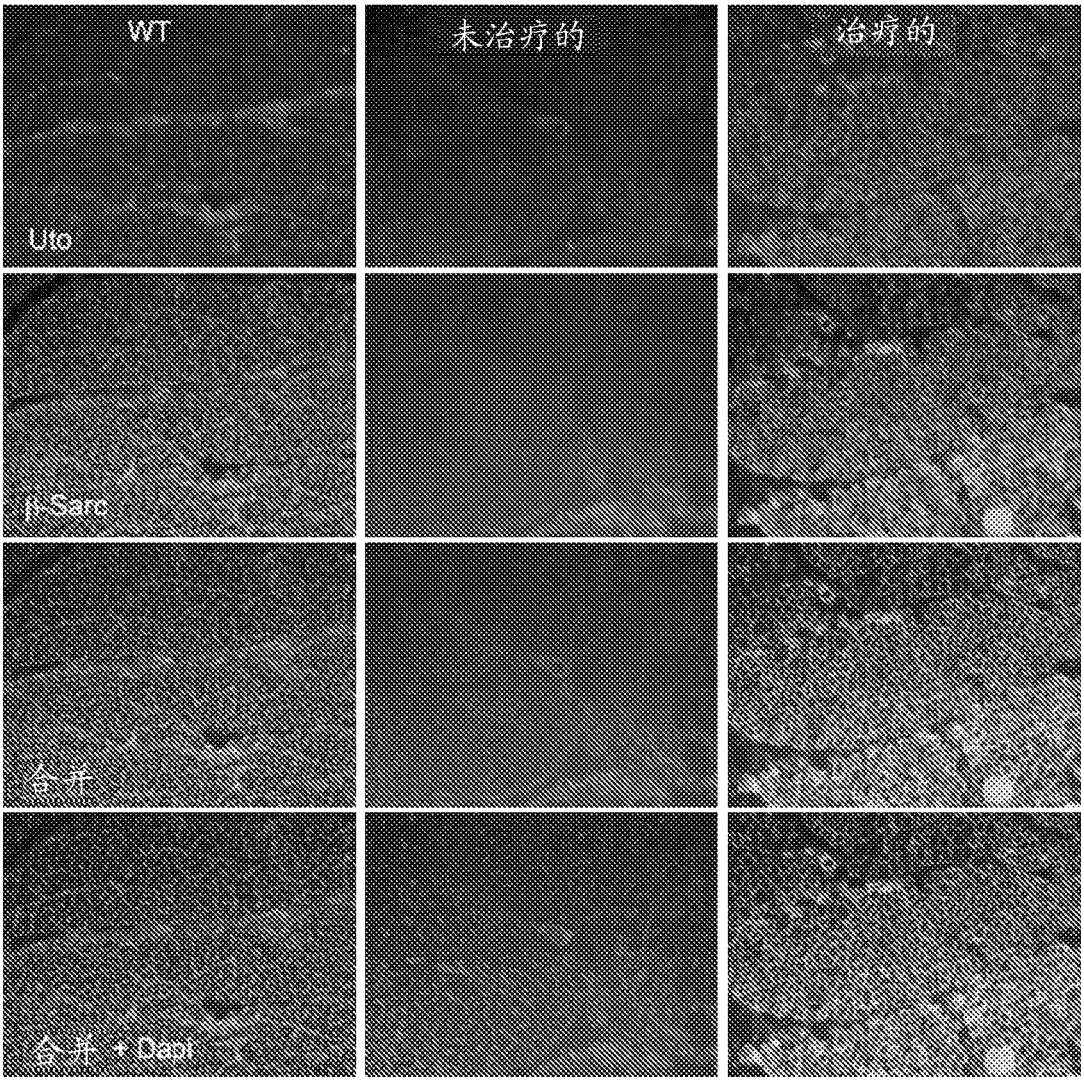
WT
未治疗的
治疗的
Uto
β-Sarc
合并
合并 + Dapi

图 19

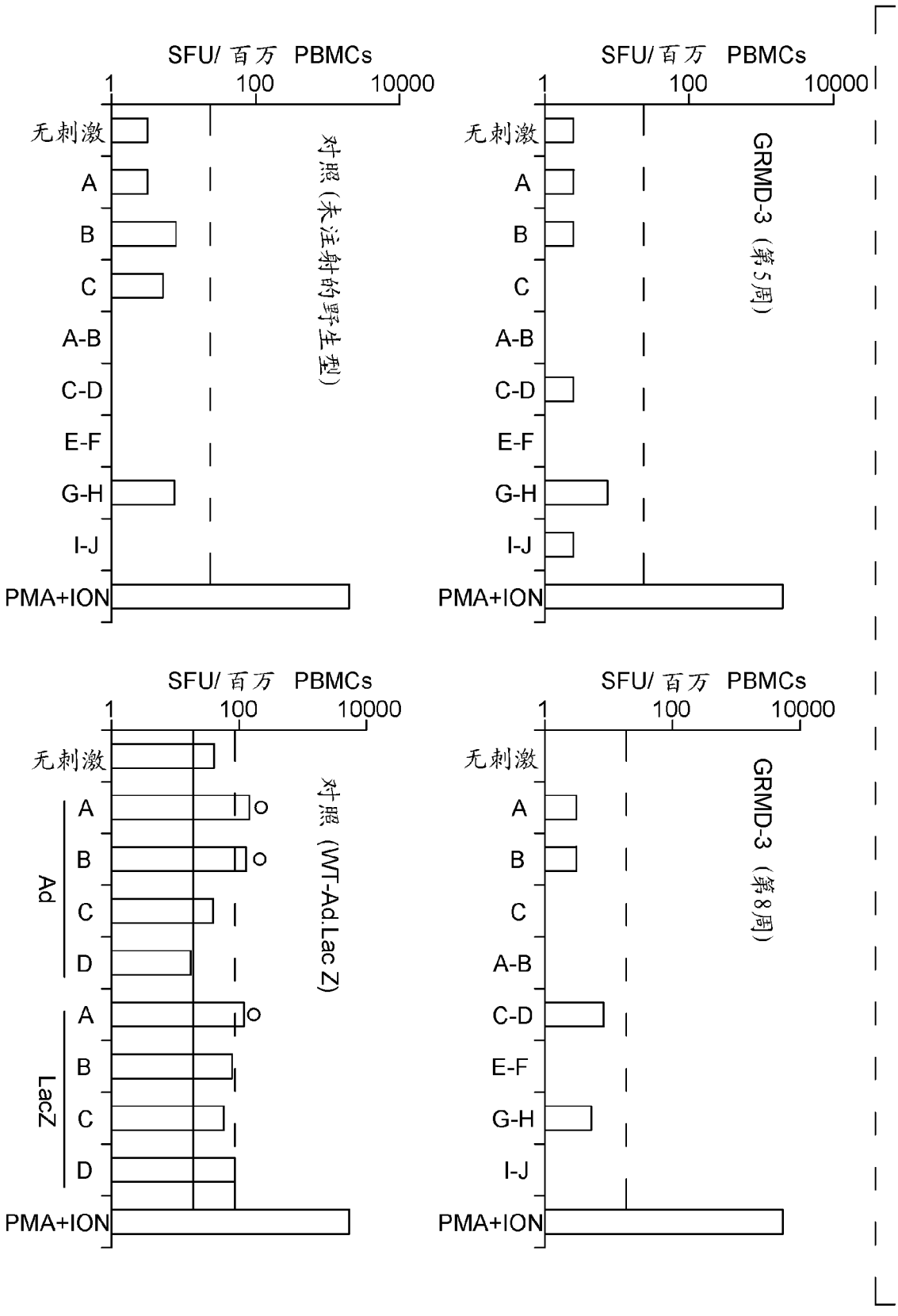
SFU/百万 PBMCs
1
100
10000
无刺激
A
B
C
A-B
C-D
E-F
G-H
I-J
PMA+ION
对照(未注射的野生型)
GRMD-3 (第5周)
Ad
LacZ
对照 (WT-Ad.Lac Z)
GRMD-3 (第8周)

图 20(I)

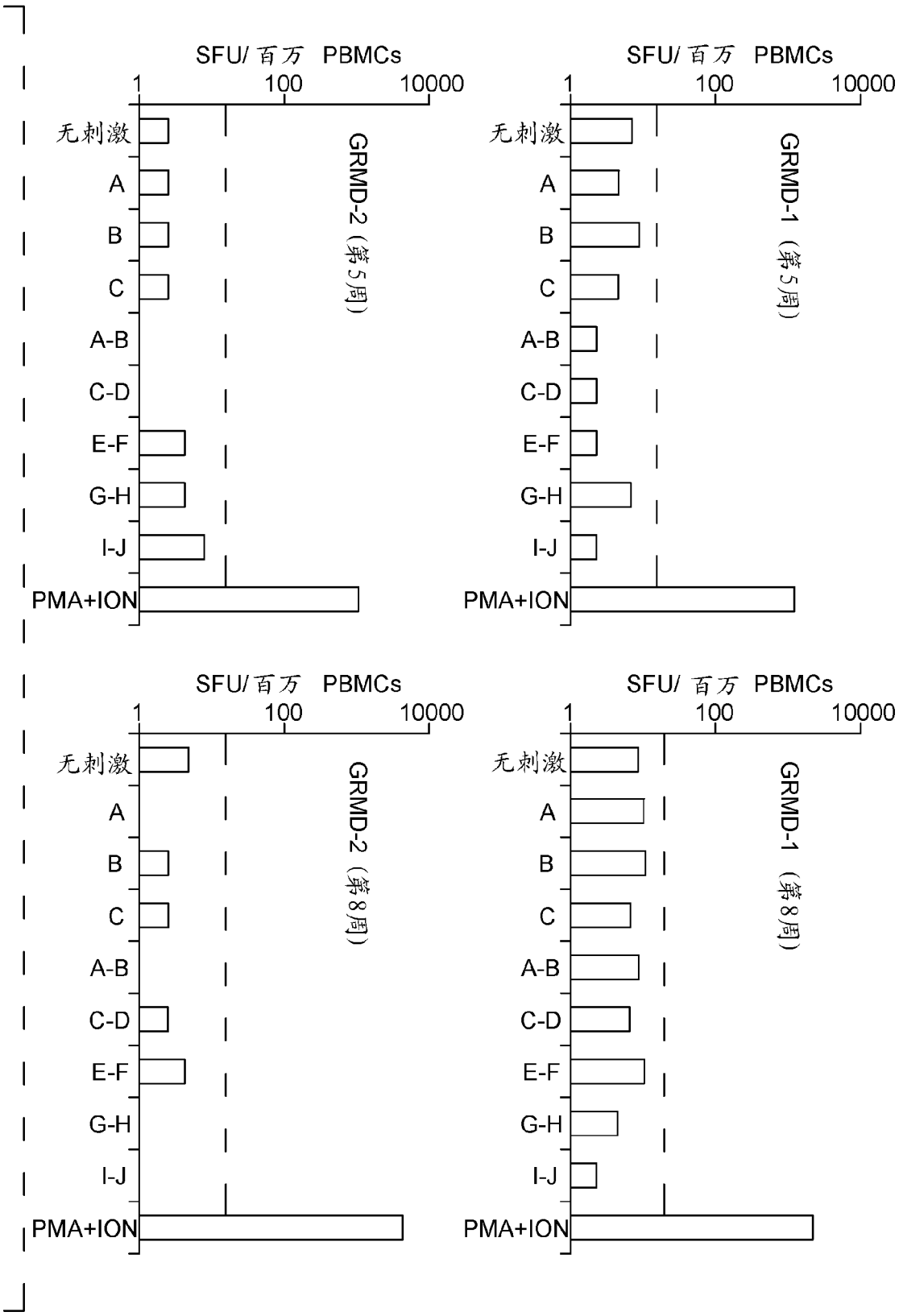
SFU/百万 PBMCs
1
100
10000
无刺激
A
B
C
A-B
C-D
E-F
G-H
I-J
PMA+ION
GRMD-2（第5周）
GRMD-1（第5周）
GRMD-2（第8周）
GRMD-1（第8周）

图 20(II)

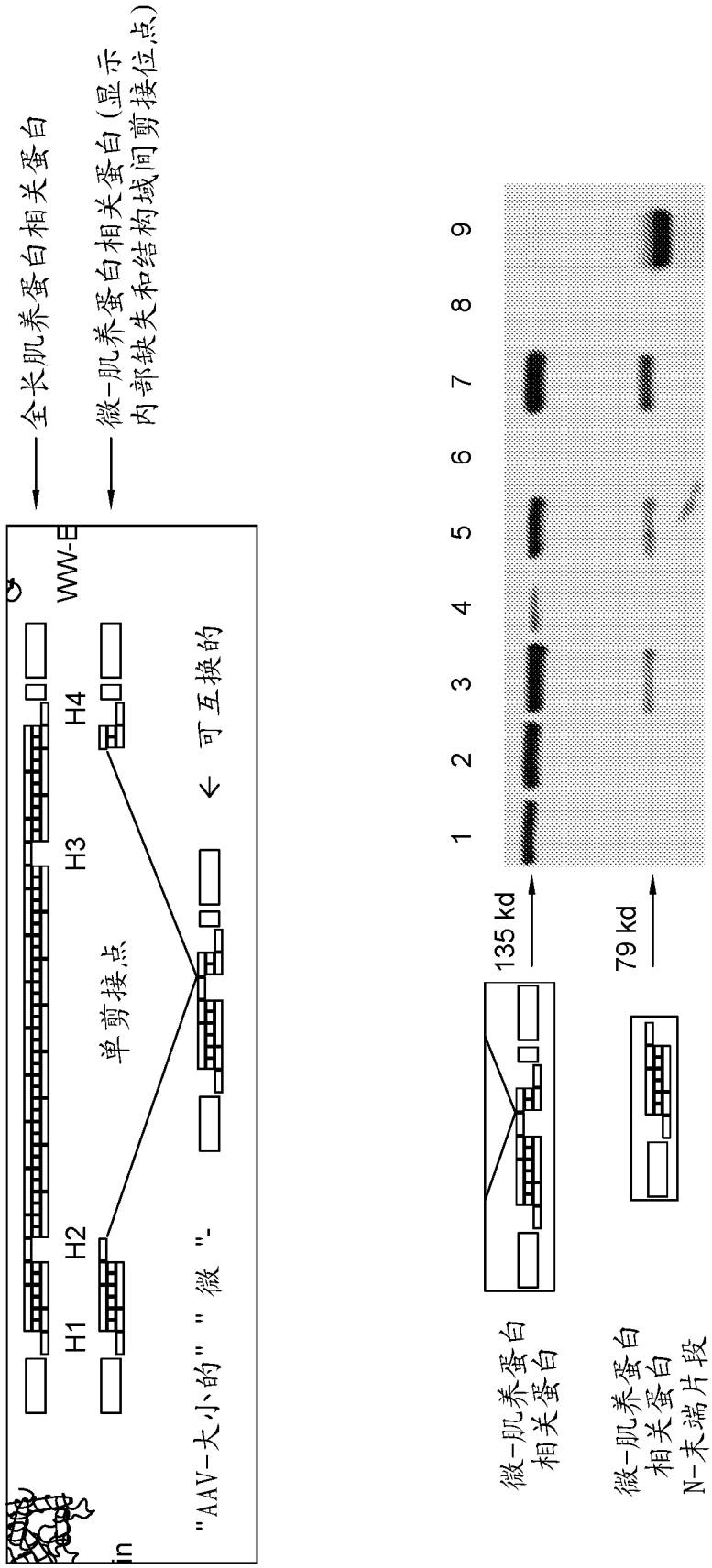
全长肌养蛋白相关蛋白
微-肌养蛋白相关蛋白(显示
内部缺失和结构域间剪接位点)
WW-E
H1
H2
H3
H4
单剪接点
"AAV-大小的" " 微 "-
↙ 可互换的
in
1 2 3 4 5 6 7 8 9
135 kd
79 kd
微-肌养蛋白
相关蛋白
微-肌养蛋白
相关蛋白
N-末端片段

图 21

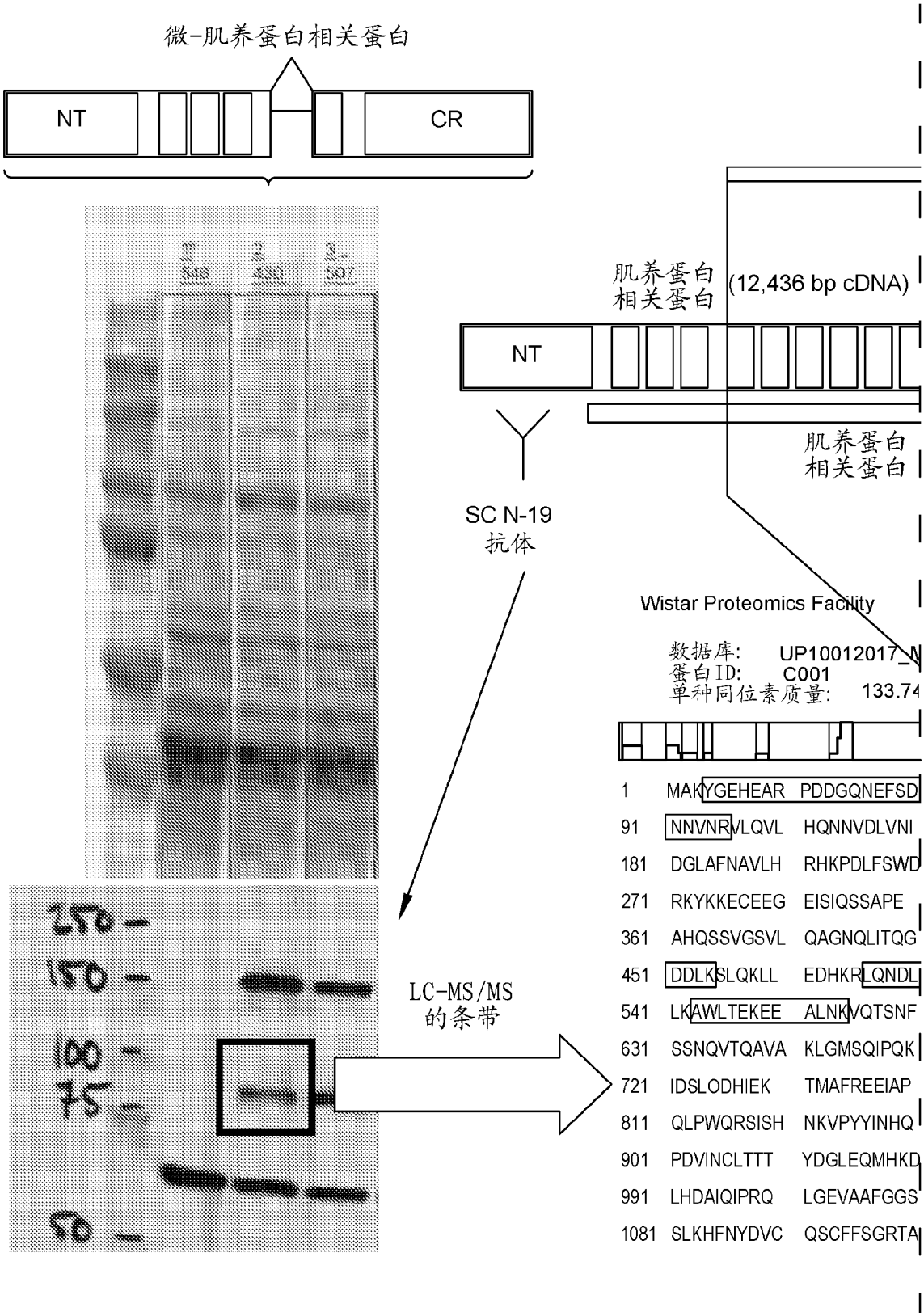
微-肌养蛋白相关蛋白
NT
CR
肌养蛋白
相关蛋白
(12,436 bp cDNA)
NT
肌养蛋白
相关蛋白
SC N-19
抗体
Wistar Proteomics Facility
数据库: UP10012017_N
蛋白ID: C001
单种同位素质量: 133.74
1 MAKYGEHEAR PDDGQNEFSD
91 NNVNRVLQVL HQNNVDLVNI
181 DGLAFNAVLH RHKPDLFSWD
271 RKYKKECEEG EISIQSSAPE
361 AHQSSVGSVL QAGNQLITQG
451 DDLKSLQKLL EDHKRLQNDL
541 LKAWLTEKEE ALNKVQTSNF
631 SSNQVTQAVA KLGMSQIPQK
721 IDSLODHIEK TMAFREEIAP
811 QLPWQRSISH NKVPYYINHQ
901 PDVINCLTTT YDGLEQMHKD
991 LHDAIQIPRQ LGEVAAFGGS
1081 SLKHFNYDVC QSCFFSGRTA
250
150
100
75
50
LC-MS/MS
的条带

图 22(I)

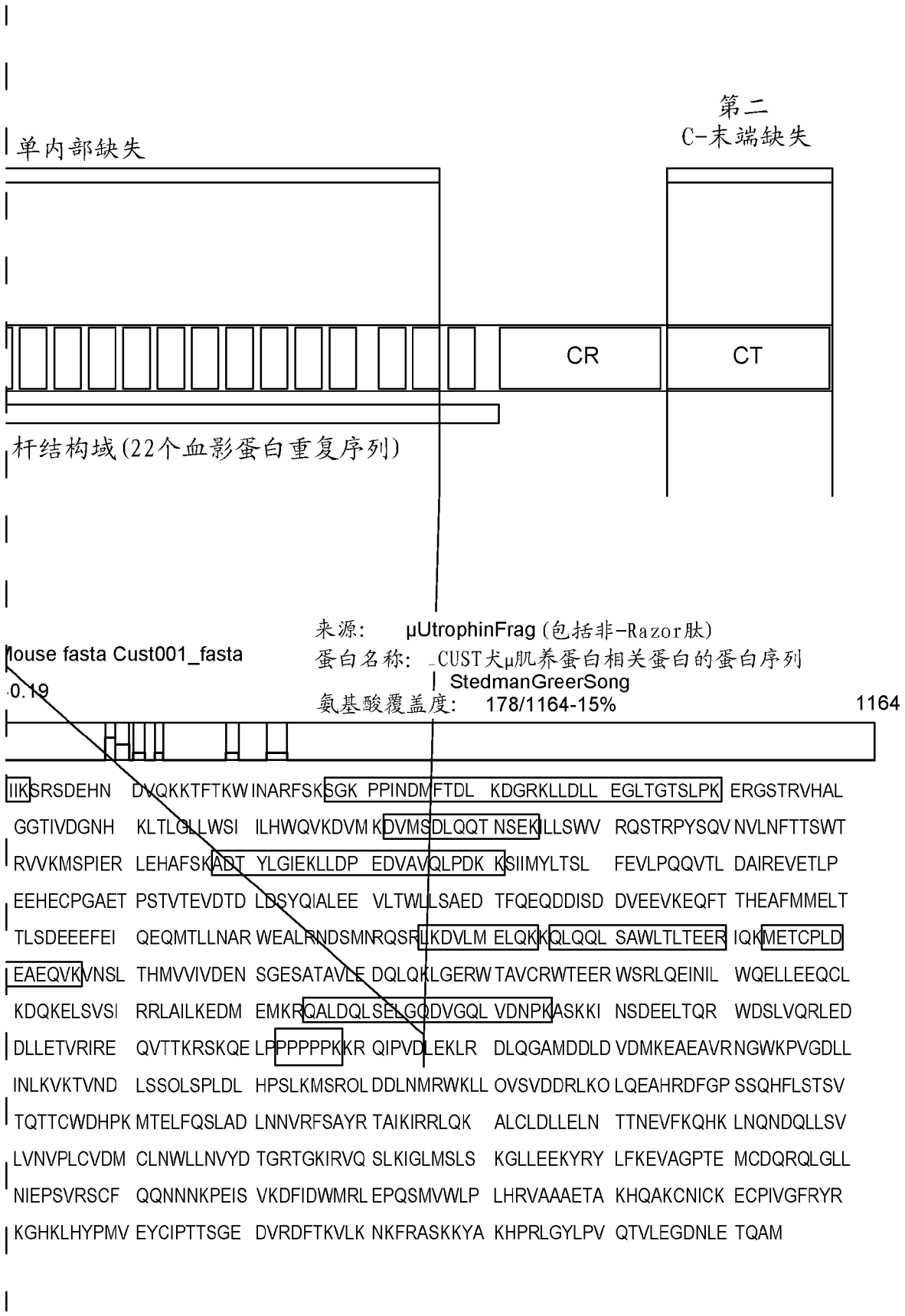
第二
C-末端缺失
单内部缺失
CR
CT
杆结构域(22个血影蛋白重复序列)
来源: μUtrophinFrag(包括非-Razor肽)
蛋白名称: _CUST犬μ肌养蛋白相关蛋白的蛋白序列
StedmanGreerSong
氨基酸覆盖度: 178/1164-15%
1164
Mouse fasta Cust001_fasta
0.19
IIKSRSDEHN DVQKKTFTKW INARFSKSGK PPINDMFTDL KDGRKLLDLL EGLTGTSLPK ERGSTRVHAL
GGTIVDGNH KLTLGLLWSI ILHWQVKDVM KDVMSDLQQT NSEKLLSWV RQSTRPYSQV NVLNFTTSWT
RVVKMSPIER LEHAFSKADT YLGIEKLLDP EDVAVQLPDK KSIIMYLTSL FEVLPQQVTL DAIREVETLP
EEHECPGAET PSTVTEVDTD LDSYQIALEE VLTWLLSAED TFQEQDDISD DVEEVKEQFT THEAFMMELT
TLSDEEEFEI QEQMTLLNAR WEALRNDSMN RQSRLKDVLM ELQKKQLQQL SAWLTLTEER IQKMETCPLD
EAEQVKVNSL THMVVIVDEN SGESATAVLE DQLQKLGERW TAVCRWTEER WSRLQEINIL WQELLEEQCL
KDQKELSVSI RRLAILKEDM EMKRQALDQL SELGQDVGQL VDNPKASKKI NSDEELTQR WDSLVQRLED
DLLETVRIRE QVTTKRSKQE LPPPPPPKKR QIPVDLEKLR DLQGAMDDLD VDMKEAEAVR NGWKPVGDLL
INLKVKTVND LSSQLSPLDL HPSLKMSROL DDLNMRWKLL QVSVDDRLKQ LQEAHRDFGP SSQHFLSTSV
TQTTCWDHPK MTELFQSLAD LNNVRFSAYR TAIKIRRLQK ALCLDLLELN TTNEVFKQHK LNQNDQLLSV
LVNVPLCVDM CLNWLLNVYD TGRTGKIRVQ SLKIGLMSLS KGLLEEKYRY LFKEVAGPTE MCDQRQLGLL
NIEPSVRSCF QQNNNKPEIS VKDFIDWMRL EPQSMVWLP LHRVAAAETA KHQAKCNICK ECPIVGFRYR
KGHKLHYPMV EYCIPTTSGE DVRDFTKVLK NKFRASKKYA KHPRLGYLPV QTVLEGDNLE TQAM

图 22(II)

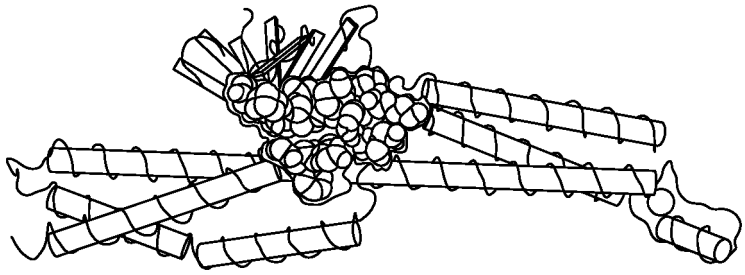

图 23

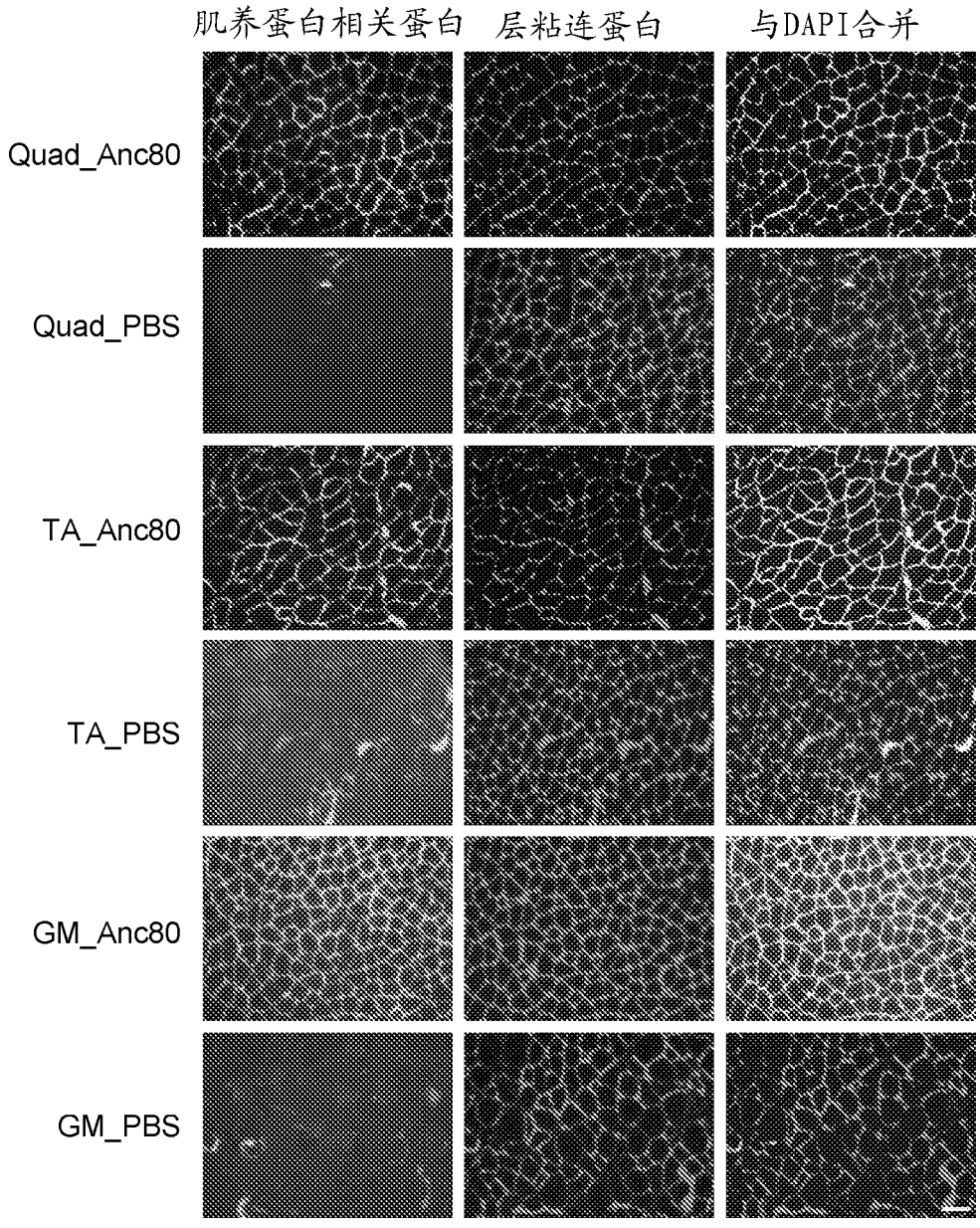
肌养蛋白相关蛋白
层粘连蛋白
与DAPI合并
Quad_Anc80
Quad_PBS
TA_Anc80
TA_PBS
GM_Anc80
GM_PBS

图 24A

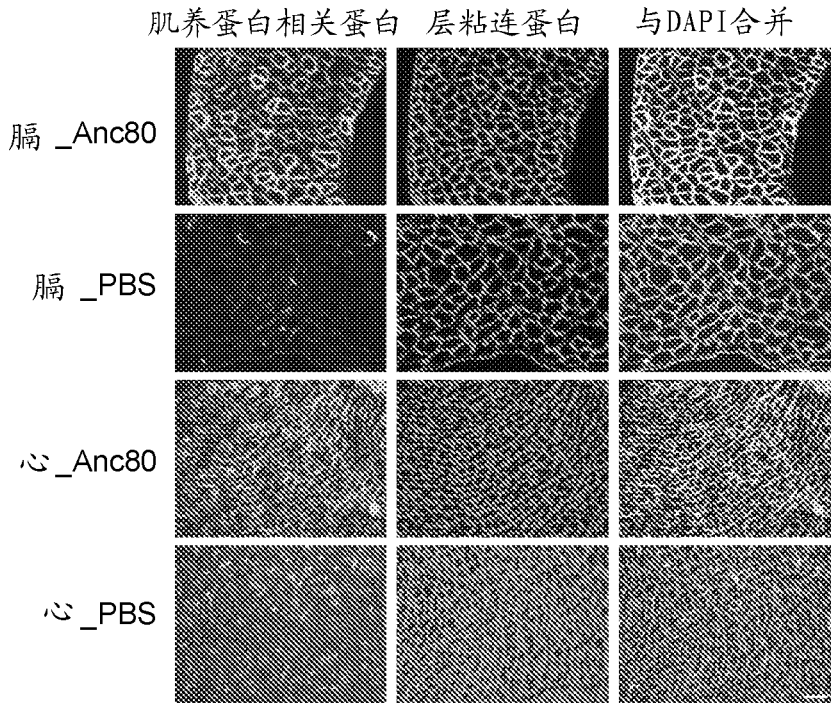
肌养蛋白相关蛋白
层粘连蛋白
与DAPI合并
膈 _Anc80
膈 _PBS
心_Anc80
心_PBS

图 24B

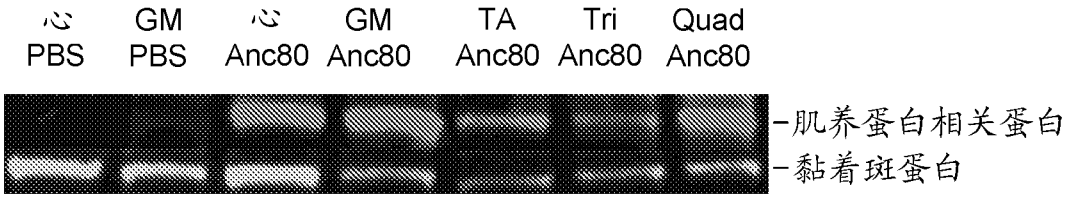
心 PBS
GM PBS
心 Anc80
GM Anc80
TA Anc80
Tri Anc80
Quad Anc80
-肌养蛋白相关蛋白
-黏着斑蛋白

图 24C

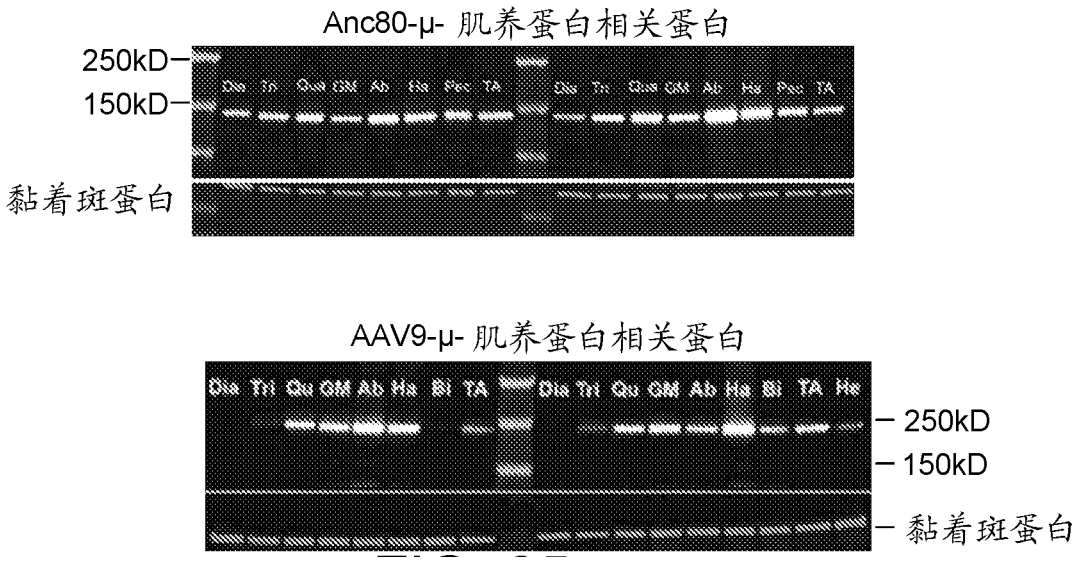
Anc80-μ- 肌养蛋白相关蛋白
250kD
150kD
黏着斑蛋白
AAV9-μ- 肌养蛋白相关蛋白
250kD
150kD
黏着斑蛋白

图 25

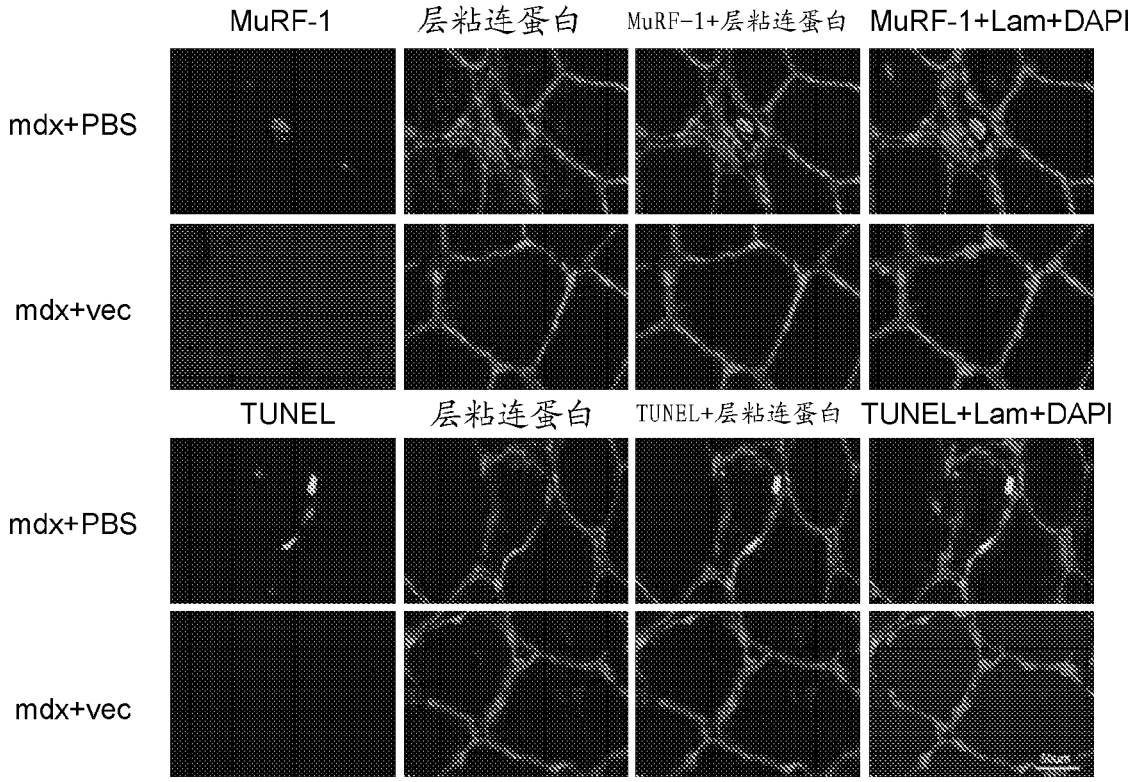
MuRF-1
层粘连蛋白
MuRF-1+层粘连蛋白
MuRF-1+Lam+DAPI
mdx+PBS
mdx+vec
TUNEL
层粘连蛋白
TUNEL+层粘连蛋白
TUNEL+Lam+DAPI
mdx+PBS
mdx+vec

图 26A

	中央成核的肌纤维(%)	MURF-1(+)肌纤维(%)	TUNEL (+)肌纤维(%)	胚胎MyHC (+)肌纤维 (%)
MDX+PBS	53.6+/-13.1	6.8+/-12.4	3.1+/-3.6	11.7+/-6.0
MDX+VEC	0.99+/- 1.9	0	0	0
MDX+VEC	0.15+/- .49	0	0	0

图 26B

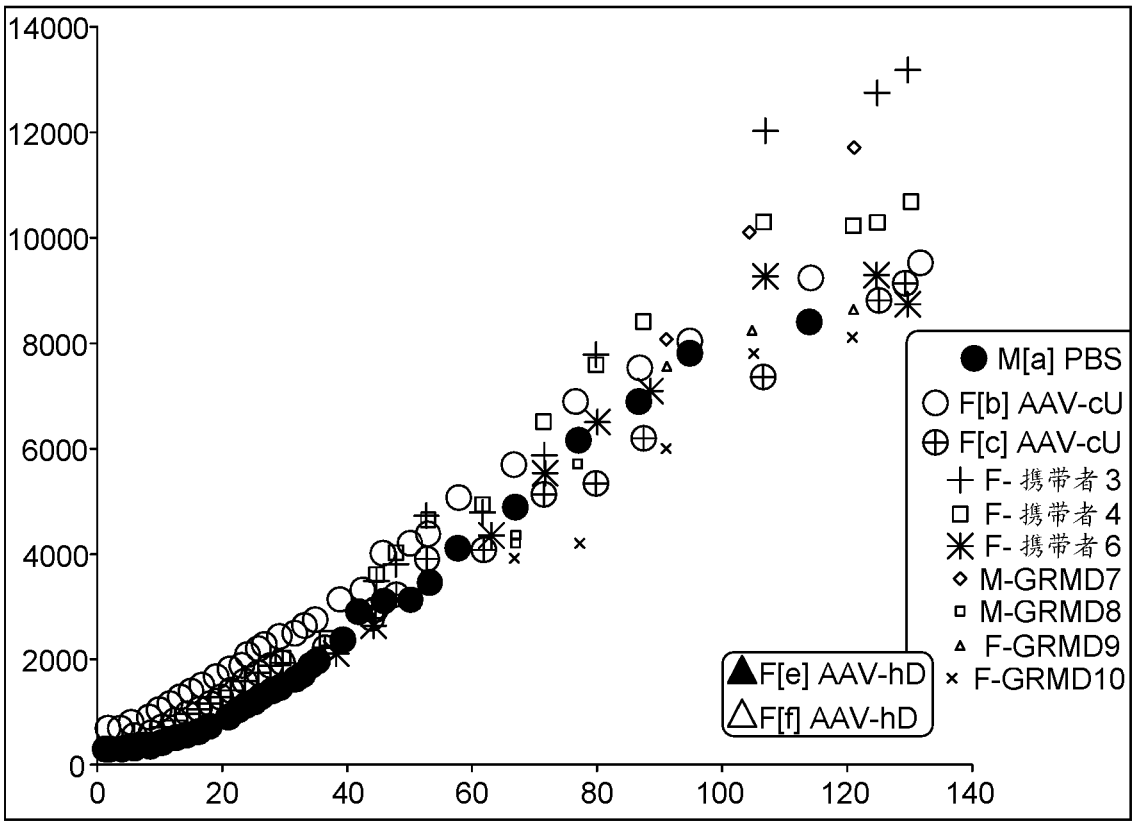

图 27

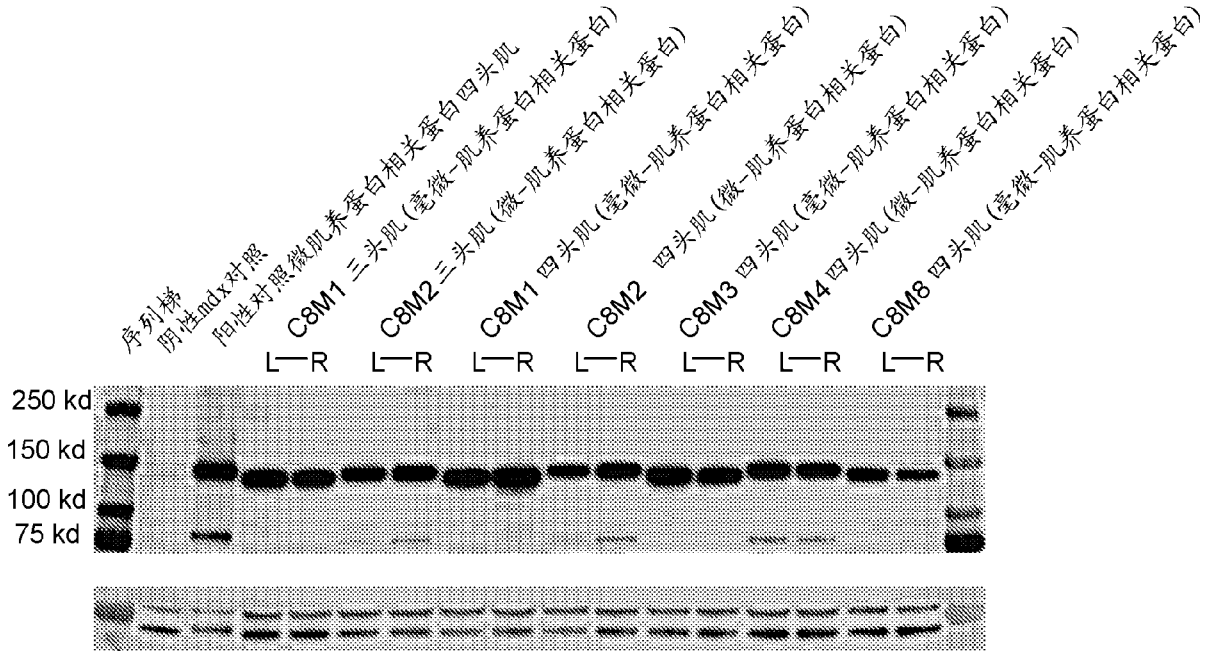

图 28

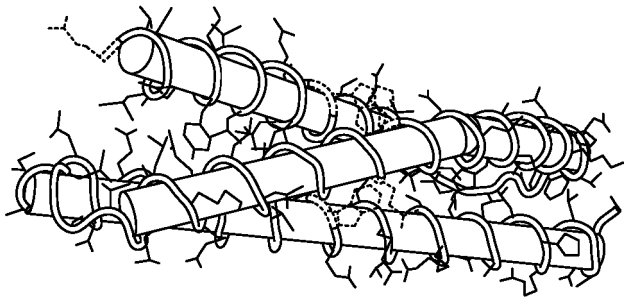

图 29A

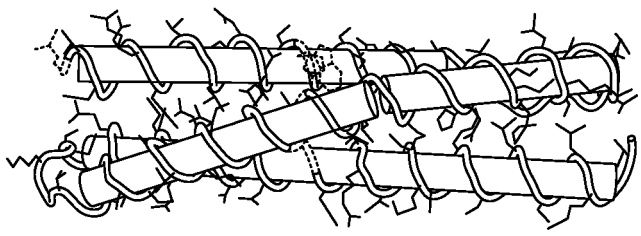

图 29B

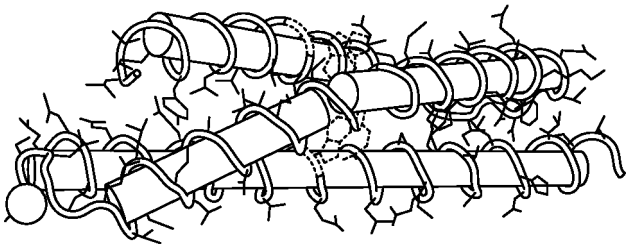

图 30A

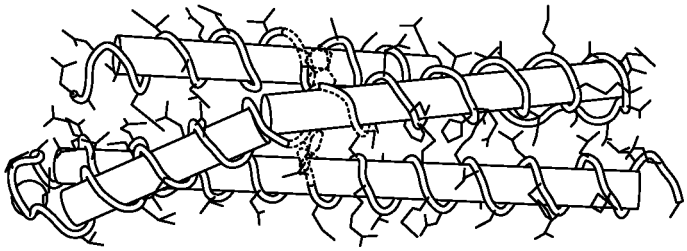

图 30B